



CENTRO DE ESTUDIOS DEMOGRÁFICOS URBANOS Y AMBIENTALES

**GENTRIFICACIÓN DIGITALMENTE AUMENTADA. TWITTER COMO  
HERRAMIENTA DE DIFERENCIACIÓN DEL ESPACIO URBANO EN  
LA CIUDAD DE MÉXICO**

Tesis presentada por  
ALEJANDRO SÁNCHEZ ZÁRATE

Para optar por el grado de  
DOCTOR en  
ESTUDIOS URBANOS Y AMBIENTALES

Directora de tesis  
DRA. LANDY SÁNCHEZ PEÑA

CIUDAD DE MÉXICO 19 DE JUNIO DE 2020





CENTRO DE ESTUDIOS DEMOGRÁFICOS URBANOS Y AMBIENTALES

**Doctorado en Estudios Urbanos y Ambientales**

Constancia de aprobación de tesis

Ciudad de México, 19 de junio de 2020

Directora de tesis : Dra. Landy Sánchez Peña

Aprobada por el jurado examinador

Presidente

Dr. Alfonso Medina Urrea

---

Primer vocal

Dr. Alejandro Mercado Celis

---

Vocal secretario

Dr. Djamel Eddine Toudert

---

Sinodal Suplente

Dra. Alejandra Trejo Nieto

---



A mi mamá,  
mi hermana,  
y mi tío. Mi amada familia



## Agradecimientos

El trabajo de investigación que aquí desarrollo es resultado de un gran conjunto de personas e instituciones que me han apoyado y han sido copartícipes en este camino. Tengan siempre presente que los llevo en mi mente y mi corazón. Quisiera agradecerle al CONACyT por haberme otorgado el apoyo financiero para poder realizar mis estudios de posgrado, tanto en México como una estancia de investigación en el extranjero. Gracias al CEDUA y El Colegio de México por haberme abierto las puertas, formarme bajo los más altos estándares de excelencia académica, a fortalecer el pensamiento crítico y sobre todo, por darme la oportunidad de tener la experiencia más gratificante de mi vida. Gracias a Fundación Kaluz por haberme otorgado un apoyo para mi movilidad estudiantil.

Mi más grande y sincero agradecimiento a mi asesora de tesis, Landy Sánchez. Gracias por haber creído en este proyecto y en mi, por ofrecerme todas tus enseñanzas, por la infinita paciencia y consejos. Sin duda alguna, gracias a ti he crecido como profesional y persona. Landy, ten por seguro que tus palabras siempre resonaran en mi cabeza. También quiero reconocer y agradecer a Djamel Toudert, mi lector. Sin sus consejos en torno a la dimensión digital de la geografía esta investigación hubiera tomado otro rumbo, debido a él, puedo decir que, este trabajo es un granito de arena para entender el fascinante mundo de las geografías digitales. Mi gratitud a mis lectores externos, Alejandro Mercado y Alfonso Medina. Su revisión crítica, cuidadosa, puntual y propositiva hizo que repensara y fortaleciera el trabajo, sin duda alguna, gracias a ustedes este trabajo creció mucho.

Tengo una deuda enorme con mi madre, mi hermana y mi tío. Mi familia. Este trabajo es reflejo de toda mi historia de vida, es reflejo de todo lo que me han enseñado e influenciado; la preocupación por la desigualdad social, el apoyo al otro, el amor y pasión por las matemáticas, la computación, relatos de robots y un impulso por saber. Gracias también por todo el apoyo emocional que me han dado, desde los momentos de desesperación, frustración, enojo, tristeza y también

los momentos de felicidad, alegría y flojera. Les amo infinitamente.

Gracias Kary, la persona que llegó a mi vida en este proceso y me ha hecho desarrollar mi vida personal, más allá de los libros o computadoras. Debido a ti he tenido la suficiente fortaleza para arriesgarme y crecer. Te amo tanto como los números reales en un intervalo de 0 a 1. Gracias a mis amigas/os del doctorado, Daniel, Erick, Laura, Lucy, Luz, Nacho, Victoria, Vampo, ustedes le dieron un toque especial a estos años. No me olvido de todas la personas especiales con las he compartido mi estancia en el colegio, Carmen, Mariana, Rosy, Bere, Beca, Dardón, Jakob, Ramón, Gustavo, una disculpa si no los menciono a todos, pero los agradecimientos no pueden ser igual de extensos que la tesis 😊.

Mi especial reconocimiento y agradecimiento a Claudia Escobar. Gracias Clau por tu apoyo y las buenas pláticas, por enseñarme todo sobre el manejo eficiente de la biblioteca y por resolver mis dudas. También quiero agradecer Carlos Sarmiento. Charly sin tu ayuda la recolección de tweets y por ende, esta tesis no se hubiera logrado.

Finalmente quisiera agradecer a todas las personas que me apoyaron y me comentaron el trabajo en MUSSI-Maynooth, Rob Kitchin, Martin Charlton y Chris Brunson, Alistair, Bob y Caspar. Ustedes me hicieron conocer un mundo nuevo y lleno de oportunidades.

## Resumen

La presente investigación tiene como objetivo explorar los mecanismos que intervienen en la conformación de una gentrificación aumentada digitalmente en la Ciudad de México, por medio de Twitter. Para lograr dicho propósito la presente investigación es propositiva en dos campos. Por un lado se propone un marco teórico que permite ensamblar elementos del enfoque de la Realidad Geográfica Aumentada Digitalmente (RGAD) con elementos del desarrollo desigual urbano de la gentrificación. Por el otro lado se desarrolla el uso de grandes datos, proveniente de la plataforma de Twitter, desde un enfoque crítico de los datos.

En el presente trabajo se trabaja bajo la idea del proceso de gentrificación como un fenómeno que estructura de manera desigual el espacio urbano. Por ello, la investigación retoma la propuesta estructura de la gentrificación como desarrollo desigual urbano desarrollada por Neil Smith. En ella se reconocen tres fuerzas de igualación/diferenciación, valorización/desvalorización del espacio y desarrollo desigual de las ciudades. Particularmente se enfocan las dos primeras fuerza por medio de la aproximación culturalista sobre la gentrificación la cual ha explorado, los medios simbólicos y culturales del fenómeno de gentrificación. Estos elementos se ensamblan a la propuesta de la RGAD, la cual explora la producción de realidades digitales aumentadas por medio de tres poderes que generan diferentes espacialidades digitales. Dichos poderes son, el poder de la distribución desigual en el acceso a lo digital, el poder desigual de comunicación en la esfera digital y finalmente el poder del código. En consecuencia, el ensamblaje teórico apunta a la interacción de cada una de las fuerzas de la gentrificación como desarrollo desigual con los tres poderes de la RGAD.

El segundo tema transversal en la investigación es la aplicación de los grandes datos a los estudios urbanos desde una perspectiva crítica. En este punto se desarrolla la idea de que los grandes datos no son reflejo inerte de la digitalización de la sociedad, sino que son actantes que intervienen en la conformación de nuevas realidades mediante las diferentes dimensiones de lo digital. Para

el caso concreto del trabajo, se exploran los mecanismos por los cuales Twitter no solo refleja el proceso de gentrificación dentro de la propia dimensión digital. En este sentido, se enfatizan elementos como el algoritmo que subyace a las plataformas digitales, el cual está lejos de ser apolítico y neutral. También se enfatizan que los grandes datos enfrentan grandes sesgos debido a que no son representativos para toda la población y, porque intervienen elementos algorítmicos.

El trabajo de tesis se divide en 5 capítulos. En el primer capítulo se desarrollan los principales conceptos sobre gentrificación, los grandes datos, espacialidades digitales y el ensamblaje teórico sobre una gentrificación aumentada digitalmente, dicha propuesta funciona para guiar la investigación a lo largo de toda la tesis. El segundo capítulo se enfoca en investigar los determinantes sociodemográficos que condicionan, en primera instancia, el acceso a Internet en México y la Zona Metropolitana de la Ciudad de México, para después indagar sobre los factores que inciden en el uso de Twitter en ambas escalas. El tercer capítulo se centra en explorar la distribución espacio temporal de la emisión de tweet en la Ciudad de México y su correlación espacial con espacios identificados como gentrificados bajo un enfoque comercial. El cuarto capítulo se interesa en estudiar los principales tópicos léxicos latentes en la Ciudad de México emanados del corpus de tweets, con el objetivo de encontrar cómo se manifiestan prácticas de consumo gentrificadoras en las publicaciones de los usuarios de Twitter que hace uso del espacio en colonias gentrificadas. En el último capítulo se exponen los principales hallazgos, elementos poco desarrollados en el trabajo, retos a futuro en el campo de las geografías digitales y los grandes datos en estudios urbanos.

Los principales hallazgos de esta investigación se encaminan a explorar empíricamente la gentrificación digitalmente aumentada en la ciudad. El primero de ellos la exploración empírica del poder de la distribución desigual de lo digital en México y en la Zona Metropolitana de la Ciudad de México. Al respecto se ha encontrado que los perfiles de usuario de Twitter responden a estratos socioeconómicos medios y altos, los cuales tienen las mayor probabilidades de acceder a Internet y redes sociodigitales. Así, lo observado en la red de Twitter se encuentra sesgado por los intereses de aquellas personas con acceso, en este caso personas privilegiadas. La distribución digital de Twitter se manifiesta espacial y temporalmente bajo un patrón de concentración en colonias gentrificadas en la Ciudad de México. En contra parte, en espacios no gentrificados se manifiestan concentraciones de emisión que sugieren alguna relación con patrones de consumo y diversión en la ciudad de México.

Basado en un análisis sociolingüístico, se explora la dimensión de la comunicación digital para aumentar el espacio gentrificado. En este rubro se plantea que las principales palabras en Twitter se relacionan con estrategias de elitización del espacio central de la ciudad, esto por medio del uso de anglicismo y términos identificados con el consumo y vida *hipster*. En consecuencia, se plantea que lingüísticamente hablando, la gentrificación es un síntoma que explica el uso de signos, o palabras, ligadas con la gentrificación en la Ciudad de México.

Por último, se arrojan resultados en el terreno metodológico y empírico, al hacer uso de los datos proveniente de Twitter se expone a lo largo de la tesis los principales retos y ventajas que se enfrentan en el uso de estas bases de datos para estudios urbanos. De manera especial, la forma de utilizar las dimensiones espaciales y temporales de estos datos, así como el análisis de texto. Así la presente tesis propone un avance en el entendimiento de la dimensión digital de la sociedad, los fenómenos urbanos y el uso de grandes datos. Lo anterior, bajo una perspectiva transformadora en la forma de conocer y estudiar la ciudad en años venideros.



# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Gentrificación digitalmente aumentada. Hacia una propuesta teórica para entender la dimensión digital de la gentrificación</b>	<b>23</b>
1.1. Introducción . . . . .	23
1.2. Sobre el concepto de gentrificación . . . . .	24
1.2.1. El proceso de gentrificación en el siglo XXI . . . . .	26
1.2.2. Gentrificación comercial y cultural . . . . .	27
1.2.3. Gentrificación guiada por el Estado . . . . .	28
1.2.4. Gentrificación como desarrollo desigual del espacio urbano . . . . .	31
1.3. Grandes Datos y las Ciencias Sociales. Hacia una ciencia social computacional . .	34
1.3.1. Aproximaciones críticas a los Grandes Datos . . . . .	35
1.4. Los grandes datos en estudios urbanos . . . . .	39
1.4.1. El reto de los grandes datos en geografía . . . . .	46
1.5. Abordajes teóricos sobre las geografías digitales . . . . .	49
1.6. Realidad geográfica aumentada digitalmente (RGAD) . . . . .	53
1.7. Gentrificación digitalmente aumentada . . . . .	58
1.8. Consideraciones finales. Hacia un estudio del proceso de gentrificación digitalmente aumentado . . . . .	65
<b>2. Desigualdad Digital y perfiles de uso de Redes Sociales y Twitter. El poder distribuido de la dimensión digital</b>	<b>69</b>
2.1. Introducción . . . . .	69
2.2. Desigualdad digital en México . . . . .	70

2.3.	Perfiles de usuarios de Twitter . . . . .	76
2.4.	Definición de variables . . . . .	80
2.4.1.	Desigualdad Digital y género . . . . .	81
2.4.2.	DD y edad . . . . .	81
2.4.3.	Desigualdad digital e ingreso . . . . .	83
2.4.4.	Descripción metodológica de la ENDUTIH . . . . .	84
2.5.	Resultados . . . . .	100
2.5.1.	Caracterización sociodemográfica de usuarios de Internet . . . . .	103
2.5.2.	Caracterización sociodemográfica de usuarios de Redes Sociodigitales . . .	109
2.5.3.	Modelos explicativos . . . . .	115
2.6.	Consideraciones finales . . . . .	119
<b>3.</b>	<b>Gentrificación aumentada en la Ciudad de México. Análisis de la distribución y comunicación digital mediante los patrones de emisión espacio-temporal de Twitter</b>	<b>123</b>
3.1.	Introducción . . . . .	123
3.2.	Twitter en Estudios Urbanos. Fase experimental en el uso los grandes datos para entender la ciudad . . . . .	125
3.2.1.	Acercamientos sociales teóricos a Twitter . . . . .	126
3.2.2.	Twitter en los Estudios Urbanos en el siglo XXI . . . . .	130
3.3.	Apartado metodológico . . . . .	132
3.3.1.	Descripción de la base de Twitter para la ZMCM . . . . .	132
3.3.2.	Identificación de colonias gentrificadas . . . . .	138
3.3.3.	Construcción de indicadores espacio-temporales de la emisión de tweets en la ZMCM . . . . .	144
3.4.	Concentración espacio-temporal de Twitter y su co-incidencia en espacios gentrificados . . . . .	147
3.5.	Consideraciones finales . . . . .	162
<b>4.</b>	<b>El poder de la comunicación en la gentrificación aumentada digitalmente en la Ciudad de México</b>	<b>167</b>
4.1.	Introducción . . . . .	167

4.2.	Gentrificación aumentada digitalmente por el poder de la comunicación . . . . .	169
4.3.	Elementos para y entorno al análisis de texto de Twitter . . . . .	171
4.3.1.	Análisis de texto en Twitter en los Estudios Urbanos . . . . .	175
4.4.	Metodología y modelización de temas . . . . .	180
4.4.1.	Depuración de texto en los tweets de la ZMCM . . . . .	180
4.4.2.	Métodos para el análisis de texto. Frecuencias y Modelo de Asignación Latente de Dirichlet . . . . .	184
4.5.	Resultados empíricos. Una aproximación al poder de la comunicación ante una gentrificación digitalmente aumentada. . . . .	190
4.5.1.	Las principales palabras de Twitter en la ZMCM . . . . .	191
4.5.2.	Los temas de Twitter en la ZMCM . . . . .	201
4.6.	Consideraciones finales . . . . .	213
<b>5.</b>	<b>Reflexiones finales</b>	<b>217</b>
	<b>Consideraciones finales, preguntas sin responder y temas a futuro</b>	<b>217</b>
5.1.	Principales hallazgos y propuestas teóricas sobre la gentrificación aumentada digitalmente . . . . .	218
5.2.	Sobre la desigualdad digital y el perfil sociodemográfico en el uso de redes socio-digitales . . . . .	221
5.3.	Temporalidad y espacialidad de Twitter en la Ciudad de México . . . . .	223
5.4.	Sobre una propuesta sociolingüística para entender el contenido del texto en Twitter	225
5.5.	Los grandes datos en los estudios urbanos . . . . .	227
5.6.	Tópicos sin desarrollar y retos a futuro en la investigación de la geografía digital, los grandes datos y las ciudades . . . . .	229
	<b>Referencias</b>	<b>233</b>
<b>A.</b>	<b>Anexo</b>	<b>251</b>

## Índice de figuras

1.1. Relación de las 3 Vs del paradigma Big Data . . . . .	37
1.2. Aplicaciones de los grandes datos en estudios urbanos . . . . .	41
1.3. Geografía de Internet . . . . .	48
1.4. Búsquedas en diferentes motores sobre Tiananmen, China . . . . .	56
1.5. Búsqueda de restaurantes veganos en la Ciudad de México en inglés y español . . .	57
1.6. Modelo teórico del proceso de gentrificación como realidad geográfica aumentada digitalmente . . . . .	60
2.1. Estudios previos sobre Twitter . . . . .	79
2.2. Esquema de tamaño muestral y poblacional de la ENDUTIH, 2018 . . . . .	98
2.3. Población usuaria de Internet en México, 2001-2018 (Porcentajes) . . . . .	101
2.4. Población usuaria de Internet en México y Zona Metropolitana de la Ciudad de México, 2018 (Porcentajes) . . . . .	102
2.5. Porcentaje de acceso según dispositivo, 2018 . . . . .	107
2.6. Número de dispositivos y estrato socioeconómico en la ZMCM, 2018 . . . . .	107
2.7. Gasto mensual en telefonía móvil en México y ZMCM, 2018 . . . . .	108
2.8. Gasto mensual en telefonía móvil de usuarios de Twitter en y ZMCM, 2018 . . . .	109
2.9. Proporción de uso de Twitter según estrato económico (base estrato muy bajo) . . .	113
3.1. Número de publicaciones según usuario (escala logarítmica) . . . . .	136
3.2. Ubicación geográfica de cuentas automatizadas . . . . .	138
3.3. Colonias identificadas como gentrificadas . . . . .	142
3.4. Porcentaje de emisión en la ZMCM según día y hora . . . . .	148
3.5. Porcentaje de emisión de tweet georreferenciados en la ZMCM según día y hora . .	149

3.6. Núcleos de emisión de tweets en la ZMCM . . . . .	150
3.7. Núcleos de emisión de tweets en horarios de 7:00 - 12:59 hrs en la ZMCM . . . . .	152
3.8. Núcleos de emisión de tweets en horarios de 13:00 - 17:59 hrs en la ZMCM . . . . .	155
3.9. Núcleos de emisión de tweets en horarios de 18:00 - 21:59 hrs en la ZMCM . . . . .	158
3.10. Núcleos de emisión de tweets en horarios de 22:00 - 00:00 hrs en la ZMCM . . . . .	160
4.1. Modelo gráfico de la ALD para la creación de temas latentes . . . . .	187
4.2. Modelo gráfico de la ALD para la asignación de temas latentes . . . . .	188
4.3. Perplejidad de los ajustes ALD dado diferente número de clústers . . . . .	189
4.4. Principales palabras de los tweets de la ZMCM, según colonia de emisión . . . . .	193
4.5. Principales palabras de los tweets de la ZMCM, según colonia de emisión (días laborales, 7:00 hrs - 12:59 hrs) . . . . .	195
4.6. Principales palabras de los tweets de la ZMCM, según colonia de emisión (días laborales, 18:00 hrs - 21:59 hrs ) . . . . .	196
4.7. Principales palabras de los tweets de la ZMCM, según colonias de emisión (fines de semana, 18:00 hrs - 21:59 hrs ) . . . . .	198
4.8. Principales palabras de los tweets de la ZMCM, según colonias de emisión (fines de semana, 22:00 hrs - 23:59 hrs ) . . . . .	200
4.9. Núcleos de densidad de emisiones del tema latente de “Cultura hipster” . . . . .	210
4.10. Núcleos de densidad de emisiones del tema latente de “Espacio público de consumo”	211
4.11. Núcleos de densidad de emisiones del tema latente de “Comida y bebida” . . . . .	212
A.1. Principales palabras de los tweets en colonias gentrificadas(50 palabras) . . . . .	254
A.2. Principales palabras de los tweets en colonias no gentrificadas (50 palabras) . . . . .	254
A.3. Principales palabras de los tweets en colonias gentrificadas (días laborales, 7:00 hrs - 12:59 hrs) . . . . .	255
A.4. Principales palabras de los tweets en colonias no gentrificadas (días laborales, 7:00 hrs - 12:59 hrs) . . . . .	255
A.5. Principales palabras de los tweets en colonias gentrificadas (días laborales, 18:00 hrs - 21:59 hrs) . . . . .	256

A.6. Principales palabras de los tweets en colonias no gentrificadas (días laborales, 18:00 hrs - 21:59 hrs) . . . . .	256
A.7. Principales palabras de los tweets en colonias gentrificadas (fines de semana, 18:00 hrs - 21:59 hrs) . . . . .	257
A.8. Principales palabras de los tweets en colonias no gentrificadas (fines de semana, 18:00 hrs - 21:59 hrs) . . . . .	257
A.9. Principales palabras de los tweets en colonias gentrificadas (fines de semana, 22:00 hrs - 23:59 hrs) . . . . .	258
A.10. Principales palabras de los tweets en colonias no gentrificadas (fines de semana, 22:00 hrs - 23:59 hrs) . . . . .	258

## Índice de cuadros

2.1. Variables para el estudio de Twitter . . . . .	87
2.2. Variables relacionadas con el entorno de la vivienda . . . . .	90
2.3. Años acumulados al inicio de cada nivel de escolaridad . . . . .	91
2.4. Distribución de las variables de construcción según estrato socioeconómico (ZMCM)	96
2.5. Características sociodemográficas y cociente de localización de usuarios/no usuarios de Internet, 2018 (ZMCM y nacional) . . . . .	104
2.6. Características sociodemográficas y cociente de localización de usuarios/no usuarios de redes sociodigitales, 2018 (ZMCM y nacional) . . . . .	110
2.7. Características sociodemográficas de usuarios de Redes Sociodigitales en la ZMCM, 2018 . . . . .	111
2.8. Resultado de los modelo logísticos de uso de Internet, Redesociodigitales y Twitter en la ZMCM, 2018 . . . . .	116
3.1. Usuarios con más de 3000 publicaciones . . . . .	137
4.1. Texto hipotético de un tweet . . . . .	181
4.2. Temas latentes en el texto de tweets provenientes de las colonias gentrificadas de la ZMCM . . . . .	203
A.1. Descripción de clases económicas asociadas a gentrificación . . . . .	252
A.2. Ejemplo de entidades nombradas en los tweets de la ZMCM . . . . .	253



## Introducción

En la actualidad la distinción entre el mundo real y el digital se ha desvanecido, para dar paso a una constitución de la vida indisoluble entre ambas esferas. Así, de manera creciente las actividades cotidianas son mediadas por un dispositivo digital con conexión a Internet. La concentración, casi absoluta, de las tareas cotidianas en teléfonos inteligentes es clara muestra de ello. Ahora estos dispositivos disponen de mayor capacidad de procesamiento de datos que las computadoras primigenias; además, brindan desde herramientas simples como agenda, reloj o alarma, hasta la posibilidad de realizar transacciones bancarias en tiempo real vía Internet. Este cambio tecnológico y social se ha basado, por un lado, en el desarrollo y abaratamiento de los dispositivos digitales y, por otro lado, en el ascenso de la conectividad de las personas por medio de Internet. Ante la transformación inminente, se ha comenzado a entender la sociedad contemporánea como una sociedad de la información, de los datos, digital o de la cuarta revolución industrial.

El interés de distinguir las transformaciones sociales basadas en el avance digital no es solamente en términos de nombramiento. Frecuentemente se observa cómo la manera de realizar tareas fundamentales es mediada y transformada por las plataformas digitales. Por ejemplo: el desarrollo de las criptomonedas o el pago por medio de dinero electrónico; la manera de relacionarse cotidianamente con las personas conocidas mediante las plataformas de redes sociodigitales como Facebook, Twitter, Instagram, Whatsapp; la búsqueda de noticias en estas mismas redes o en los *blogs* digitales de los diarios que, antaño, eran impresos diariamente.

Sin embargo, la transición hacia la vida mediada por la dimensión digital también conlleva consecuencias negativas. En este rubro se habla de un nuevo capitalismo de hipervigilancia. Esto se debe a que todas las acciones realizadas por medio de una computadora e Internet pueden ser rastreadas, analizadas y utilizadas a favor de unos pocos que ostentan el poder tecnológico. Este punto es uno de los puntos más debatidos por los filósofos contemporáneos de lo digital y los datos

(Floridi, 2014; Habermas, 2018; Kitchin y Perng, 2016; Lewis, 2015; Pasquale, 2016).

Los cambios tecnológicos transforman el territorio y no se manifiestan ubicuamente en él. De esta manera, las ciudades son un espacio particular donde se expresan los cambios tecnológicos y la vida mediada por lo digital. Basta pensar en la tendencia creciente de pago electrónico de cualquier medio de transporte; así como, las crisis políticas o económicas que giran en torno al servicio de taxis alrededor del mundo, provocadas por la aparición y masificación de plataformas como UBER, Cabify o Blablacar. La pérdida del monopolio de la búsqueda de vivienda de agencias de bienes raíces ante la aparición de páginas de internet de anuncios clasificados, o bien, por medio de plataformas como Airbnb (Yrigoy, 2016). El cambio en la manera de habitar y vivir la ciudad gracias a los desarrollos de plataformas de tránsito y embotellamientos en tiempo real, tales como Waze o Google Maps. Desde un punto de vista político, también se ha visto el auge de movimientos sociales y la ocupación de espacios públicos urbanos gracias a plataformas como Twitter (AlSayyad y Guvenc, 2015; Castells, 2012; M. Graham, De Sabbata, y Zook, 2015).

Al mismo tiempo, estos fenómenos también implican una transformación en el modo de vida urbano, que no necesariamente se traducen en el mejoramiento de la ciudad. Por ejemplo, la hiper-vigilancia hacia los ciudadanos en ciertos países que se ha comenzado a observar recientemente; el incremento de la precariedad laboral gracias al empleo informal que generan aplicaciones como Uber eats o Rappi; la omisión de determinados lugares que no aparecen en la base de datos de Google Maps al realizar una búsqueda (Friedman, 2014). En pocas palabras, en el entorno urbano se experimentan las ventajas y desventajas de la transformación digital de la sociedad.

De manera paralela al avance digital, las ciudades en el siglo XXI se erigen como espacios de extrema desigualdad social. Por un lado, se tienen los núcleos de concentración de poder económico, político y social en los centros financieros y sociales de las grandes ciudades. En este sentido, las ciudades son nodos en un complejo entramado de poder económico (Sassen, 2001). Así, se ha establecido la idea de ciudades globales, urbes que controlan las decisiones sobre la economía y la política mundial y, de otras ciudades de menor jerarquía dentro del sistema.

Por otro lado, al interior de las urbes también se pueden encontrar espacios de extrema diferenciación social, basta con ver los puntos de concentración económica y social en cada ciudad, lugares arquitectónicamente bien desarrollados, con todas las facilidades y ventajas de la urbanización, lugares ícono del desarrollo urbano. También se tienen lugares periféricos, espacios donde

habitan personas en condición de pobreza y espacialmente segregadas, lugares que han sido desarrollados por medio de la autoconstrucción de sus viviendas y, que no tienen acceso a servicios públicos básicos.

En este contexto, el proceso de gentrificación ha sido foco de atención de los estudios urbanos en las últimas décadas. Ya sea desde el ámbito político, académico o de justicia social, la gentrificación es concebida como un proceso de transformación urbana que estructura de manera desigual el espacio urbano (Lees, Shin, y López-Morales, 2016). Así, la gentrificación es vista como una estrategia del sistema capitalista contemporáneo en los espacios urbanos para la generación e incremento de las ganancias económicas (N. Smith, 2002). En este sentido, el resultado del proceso de gentrificación a nivel planetario implica, en la mayoría de los casos, la reconfiguración social de espacios revitalizados, en detrimento de clases trabajadoras. La gentrificación, como fenómeno y proceso urbano, es uno de los temas más discutidos en la literatura especializada, ha sido abordado desde diferentes perspectivas y continua en constante discusión por su cualidad de adaptabilidad ante los cambios sociales (Butler, 2005; Hammel, 2009; Janoschka, Sequera, y Salinas, 2014; Krijnen, 2018; Sabatini, Rasse, Cáceres, Robles, y Trebilcock, 2017; Slater, 2012; Wyly, 2018).

La idea general que este trabajo de investigación plantea es la dimensión digital naciente de la gentrificación en la Ciudad de México. Esta idea se sustenta en tres premisas básicas. En primer lugar, la relevancia de estudiar el proceso de gentrificación en la Ciudad de México, ya que se trata de un proceso en desarrollo, con características propias de ciudades latinoamericanas, pero también con rasgos que lo asemejan a la gentrificación en ciudades como Nueva York, Londres o Hong Kong. En segundo lugar, el vínculo que la revolución digital establece con un proceso urbano como la gentrificación. Esta idea se basa en que la dimensión digital es poco explorada desde una perspectiva crítica de la geografía y las diferentes espacialidades que puede generar. En tercer lugar, la relación entre ambos fenómenos surge del interés de comprender cómo un proceso urbano previo a la explosión de la era digital se manifiesta, interactúa y se transforma ante la omnipresencia de la dimensión digital.

La gentrificación, como una tendencia planetaria, es generalmente asociada al cambio social y económico de determinados espacios urbanos. Desde sus inicios hasta la fecha, la literatura internacional ha enfatizado los mecanismos y causas de la expulsión de personas de clase traba-

jadora en favor de personas de estrato medio-alto, generalmente en los centros deprimidos de las ciudades. Ya sea que el foco sean los mecanismos políticos, coercitivos violentos o de mercado inmobiliario que provocan el desalojo de las personas de estrato bajo, hasta los mecanismos culturales o comerciales, la mayoría de trabajos ha enfatizado la desigualdad social que se encuentra intrínseca dentro del fenómeno. En este sentido, se ha documentado, de manera extensiva, la intervención del Estado en el desalojo violento de clases populares en favor de intereses inmobiliarios (Janoschka, 2016; Grange y Pretorius, 2016). Otro frente de análisis ha sido estudiar el papel de las inversiones del Estado en la renovación o ‘sanitización’ del espacio público. O bien, entender la gentrificación como un fenómeno que refleja las leyes internas del sistema capitalista en cuanto a estrategia de acumulación de capital en espacios urbanos (Crossa, 2009; di Virgilio y Guevara, 2015; Hernández-Cordero, 2015).

En contraposición, la gentrificación también se ha entendido como un síntoma de crecimiento y desarrollo urbano, ya que funge como atractor de clases creativas, las cuales pueden ser consideradas como médula la sociedad del conocimiento (Florida, 2005). Ante la multiplicidad de enfoques que han abordado la gentrificación, el cuerpo teórico internacional sobre el tema converge hacia una visión de estructuración urbana y social desigual y, que se manifiesta a nivel mundial (Lees y cols., 2016).

El proceso de gentrificación, en su carácter histórico, se ha visto transformado a lo largo de las últimas seis décadas. En este sentido, la gentrificación ha pasado por diferentes olas, desde 1960 hasta principios del siglo XXI (Hackworth y Smith, 2001). Cada una de ellas se ha caracterizado por la intervención y preponderancia de diferentes actores y acciones que han llevado a la fragmentación del espacio urbano, vía la gentrificación. En cada ola, el papel del Estado, la inversión privada y las fuerzas del mercado han actuado de manera particular. Para la década de 1990 se identificó que el proceso de gentrificación era un fenómeno de carácter global, con la intervención de capitales transnacionales. Los trabajos clásicos desarrollados en la vertiente estructural han analizado diferentes factores que impulsan e intervienen en la gentrificación (Krijnen, 2018; Slater, 2012; N. Smith, 1990).

En décadas más recientes, otro cuerpo de literatura se ha preguntado por factores culturales y comerciales asociados al fenómeno. Al respecto, este tipo de trabajos enfatizan el papel que tiene la transformación económica en el cambio poblacional de los centros urbanos. Dicho cambio ha

virado hacia una estructuración comercial dirigida al comercio elitizado, tales como restaurantes internacionales, comercio justo, establecimientos veganos, tiendas de antigüedades, etc. De igual manera, se ha estudiado que la dimensión cultural de la gentrificación se ha centrado en los patrones culturales, signos y símbolos que se relacionan con un modo de vida gentrificado. En este sentido, de nueva cuenta se asocian elementos de consumo elitizado, principalmente en el sector de comercio. O bien, se ha estudiado la manera de hablar, la distinción por medio del uso del lenguaje en el paisaje urbano gentrificado. Ejemplo de esto es la coexistencia de varios idiomas en un mismo espacio, generalmente de origen europeo como inglés, francés o italiano (Lyons y Rodríguez-Ordóñez, 2017). Las vertientes comercial y culturalista han abonado a las aproximaciones estructurales al enfatizar los elementos locales y subjetivos que también son sustento del proceso de gentrificación (Zukin, 1987, 1991; Zukin y cols., 2009) .

En el terreno empírico, la medición de la gentrificación es un debate abierto. De manera clásica, para identificar si una colonia ha experimentado un proceso con características gentrificadoras se estudia el cambio poblacional en el tiempo. Es decir, los estudios empíricos se han centrado en indagar si hay un desplazamiento de personas de estratos bajos y el consecuente aumento de población de estratos medios y altos. Otro acercamiento ha sido la evolución de los precios de la vivienda, ya que la gentrificación se ha asociado al alza de este indicador (Bournazou, 2016; Sobrino, 2016). Estos estudios han apuntado a la operacionalización de los enfoques clásicos del fenómeno.

En menor medida, se ha estudiado empíricamente la dimensión cultural, comercial y simbólica. En torno a la última se han realizado esfuerzos para identificar el uso de ciertas palabras de origen anglosajón o la combinación de dos o más idiomas en colonias gentrificadas (Lyons y Rodríguez-Ordóñez, 2017; Papen, 2012; Vandenbroucke, 2016). Las múltiples aproximaciones de este tipo han reportado un problema común, la falta información para dar cuenta del fenómeno. Esta limitación ha sido consistente debido a la escala territorial de la unidad de análisis, generalmente las colonias o los barrios, los cuales por lo general, no reúnen las características necesarias para la generación de información estadística oficial en muchos países. Ante esto se han hecho aproximaciones a escalas de municipalidad o incluso a escalas de zonas postales (Preis, Janakiraman, Bob, y Steil, 0)

Como se ha mencionado, existe una amplia literatura, tanto teórica como empírica, sobre

el fenómeno de gentrificación. Dentro de este cuerpo se puede identificar una evolución temporal en el concepto y en el fenómeno, en sí mismo. En este sentido, en años recientes, con el auge de Internet a comenzado a surgir un interés por entender el vínculo entre el proceso de gentrificación y las revolución digital de la sociedad. Por ejemplo, en la década de 2000, Belkind (2008) comenzó a vincular el rol de los blogs en Internet para buscar viviendas y comentarios sobre un barrio en proceso de gentrificación en Nueva York. La autora ya apuntaba a la naturaleza particular de Internet, especialmente de la Web 2.0, y la posibilidad de las personas de reflejar y crear imaginarios sobre la gentrificación en un sitio de Internet.

Por su parte, desde una perspectiva culturalista Zukin, Lindeman, y Hurson (2017) apuntan al análisis de Yelp, una plataforma digital destinada a las reseñas de restaurantes, como herramienta para estudiar las maneras en las que el proceso de gentrificación, en sus diversas etapas, se manifiesta en los textos de los comensales que emiten su opinión en dicha plataforma. Las autoras, apuntan a que las reseñas no solo dan cuenta de la elitización del consumo de comida en barrios en transición o plenamente gentrificados, sino que además, en dichas reseñas se manifiestan signos del proceso urbano que ocurre más allá de los muros de los restaurantes. En el caso particular para México, Díaz (2015) ha indagado sobre el papel de las redes sociodigitales como Facebook en la organización vecinal ante el proceso de gentrificación en un barrio en la Ciudad de México. Para Wyly (2018) la dimensión digital y la revolución tecnológica es un elemento ineludible si que quiere analizar los patrones emergentes en el proceso de gentrificación. Para este autor, es necesario revisar las teorías clásicas sobre el fenómeno, a la luz de la revolución digital.

Otro cuerpo de literatura se ha preguntado entre la relación de la revolución digital y la gentrificación desde una aproximación de los grandes datos y las técnicas estadísticas de frontera. Dentro de estos trabajos se pueden mencionar los esfuerzos de Gibbons, Nara, y Appleyard (2018); C.-C. Huang (2017); Jansson (2019); Mermet (2017); Rahimi, Andris, y Liu (2017) quienes han utilizado las fuentes de grandes datos como Airbnb, Twitter, Instagram o Yelp para estudiar por medio de ellas el fenómeno de gentrificación. Al igual que la literatura clásica sobre la gentrificación, estos trabajos se han preguntando cuestiones como la ruptura de la comunidad ante la irrupción de personas de estratos medios-altos, el papel de Airbnb en el alza de precios de las viviendas en áreas de transición hacia la gentrificación o de la búsqueda vivienda en plataformas digitales de bienes raíces. Estas investigaciones no solo han vinculado el uso de nuevas fuentes de información,

ajustes de aprendizaje de máquina y nuevas maneras de medición del fenómeno también reconocen explícitamente el papel activo de la dimensión digital en el fenómeno. Sin embargo, un vacío teórico que estos trabajos presentan es la falta de reflexión sobre la naturaleza misma de los grandes datos y su conexión con la revolución digital de la sociedad. En este sentido, los grandes datos solo son entendidos como reflejo inocuo de la transformación digital, aunque sean un pilar fundamental en dicho cambio.

Los grandes datos o *big data* han sido el elemento más explorado, tanto teórica como empíricamente, en el estudio de la digitalización de la sociedad desde las ciencias sociales. Los grandes datos son consecuencia del avance vertiginoso de la computación y de la revolución digital. Actualmente, la generación de información diaria implica alrededor de 2 Petabytes de almacenamiento y continuamente se encuentra en aumento. Ante el volumen inconmensurable de información, los grandes datos han sido entendidos en función de su volumen, velocidad de generación y variedad de tipos de datos (Holmes, 2017a; Kitchin, 2014a) . Esta definición ha sido la más extendida en varios campos del conocimiento.

Sin embargo, desde las ciencias sociales se ha reflexionado sobre el poco alcance analítico que dicha conceptualización representa. De esta manera, para Salganik (2018) los grandes datos adquieren relevancia por su cualidad de responder preguntas más que por su naturaleza de volumen. En esta misma discusión se encuentra el trabajo de Kitchin y McArdle (2016) quienes discuten desde un punto de vista socio-técnico la naturaleza de los grandes datos, basado en la multiplicidad de naturalezas de las bases de datos, potenciales usos e interrelación entre dichas bases. En este sentido, el presente trabajo se aleja de una visión reducida de los grandes datos y, al mismo tiempo, recupera una postura de exploración de este tipo de nuevas fuentes para conocer sus ventajas y desventajas y aportar en el avance del entendimiento de éstas.

En este tenor, una fuente particular, dentro del universo de los grandes datos, es la información proveniente de las redes sociodigitales <sup>1</sup>. Las redes sociodigitales se han convertido en el ícono de la revolución digital y en un hito dentro del estudio de los grandes datos. Ya sea por su carácter transformador, liberador y democrático, como fue el caso de la organización de movimientos sociales tales como la primavera árabe, *#OccupyWallStreet* y *#YoSoy132* (Castells, 2012). Otro ejemplo

---

<sup>1</sup>Como se explica en el trabajo de investigación se opta por el término de redes sociodigitales para distinguirlo del concepto de redes sociales en el sentido sociológico

de esto fue la gestación de acciones ciudadanas en respuesta a desastres de origen natural, tales como la propuesta de etiquetado #Verificado19S para coordinar ayuda humanitaria desde la sociedad civil después del sismo del 19 de septiembre de 2017 en la Ciudad de México (Salgado-Andrade, 2018).

Las redes sociodigitales también se han convertido en un foco rojo para atraer la atención ante la pérdida de libertades, privacidad, seguridad e incluso de información veraz. Ejemplo de este rubro fue el caso de *Cambridge Analytica*, estrategia que utilizó información de Facebook para generar tendencias de voto en las elecciones presidenciales de Estados Unidos en 2016; también los casos de generación y difusión de noticias falsas en Twitter.

En el ámbito de los estudios urbanos, los datos de redes sociodigitales han sido un nicho para explorar fenómenos ya existentes y nuevos fenómenos. Por ejemplo, por medio del uso de Twitter se han realizado estudios de movilidad cotidiana, regionalización del territorio nacional y urbano (Cheng, Gould, Han, y Jin, 2016; Hao, Zhu, y Zhong, 2015; Thakuriah, Tilahun, y Zellner, 2017; Wu, Wang, y Dai, 2016). Por el lado de nuevos fenómenos y dimensiones urbanas, los datos de redes sociodigitales han ayudado a explorar elementos simbólicos y los sentimientos de los ciudadanos (Liu, 2015; Mitchell, Frank, Harris, Dodds, y Danforth, 2013). Aunque el avance en la incorporación de los datos de redes sociodigitales y, en general, de los grandes datos a los estudios urbanos ha sido grande, aun queda mucho por explorar para identificar las bondades y las limitaciones de este tipo de información.

Dentro del conjunto de las redes sociodigitales una de las más estudiadas ha sido Twitter. Esto se debe, básicamente, a la facilidad que esta plataforma ofrece para cosechar la información de su sitio y a la naturaleza de esta red. Con respecto al primer punto, Twitter ofrece la posibilidad obtener gratuitamente una muestra de las emisiones diarias de los usuarios. Algunos estudios han calculado que el tamaño de la muestra gratuita es alrededor de 1 % del total de las emisiones (Neuhaus, 2015). La segunda razón ha sido la naturaleza de Twitter, debido a su particularidad como red social que la diferencia de otros sitios como blogs u otras redes como Facebook, el texto corto de Twitter permite emplear técnicas de procesamiento de texto relativamente sencillas. A pesar de las diversas ventajas que Twitter ofrece como fuente de información, se tienen brechas en el conocimiento acerca de esta plataforma, tales como los sesgos y limitaciones de la información obtenida (Sloan y cols., 2013; Sloan, Morgan, Burnap, y Williams, 2015). Particularmente, una de las preguntas

hechas recientemente versa sobre cuál es el tipo de personas usuarias de Twitter (Hargittai y Litt, 2012).

Ante estos cuestionamientos, al analizar los grandes datos de Twitter es ineludible un acercamiento desde la estructura social y las desigualdades socioeconómicas. Así, es preciso apuntar que el uso de Twitter responde, en primera instancia, a la posibilidad de acceder a Internet. En este sentido, tener en cuenta la desigualdad digital en el acceso a Internet puede arrojar luz sobre las limitaciones, naturaleza y politización de Twitter como fuente de información y reflexión social. Al respecto, trabajos sobre desigualdad digital apuntan a conocer los factores determinantes del uso de plataformas digitales. Por ejemplo, Hargittai y Litt (2011, 2012) ha identificado que el uso de Twitter responde a características sociodemográficas particulares, tales como el ingreso, la edad, el sexo y el uso de otras redes sociodigitales. Aunque estos estudios arrojan un primer haz de luz sobre el uso de la red social de interés, han sido de carácter exploratorio con fuentes de información limitadas tanto en términos de generalización poblacional como territorial. Lo cual da muestra de lo poco que se sabe de una red social ampliamente difundida.

Otro de los elementos poco explorado en cuanto al uso de redes sociodigitales y las desigualdades digitales, en general, ha sido la dimensión territorial de éstas. Desde el punto de vista de las desigualdades digitales, la dimensión territorial solo ha sido analizada como una variación en la disponibilidad y uso de tecnologías digitales a lo largo del espacio. En este sentido, la representación cartográfica del uso de Internet o Twitter se relaciona con el nivel de renta y de infraestructura en telecomunicaciones de los países (M. Graham, Hale, y Stephens, 2012; M. Graham, Stephens, y Hale, 2013; Nishida, Pick, y Sarkar, 2014; Pick y Sarkar, 2016) Sin embargo, este acercamiento deja de lado el papel activo de la revolución digital en la sociedad. Al respecto Cairncross (2001) identificó una posible muerte de la distancia ante los avances en las tecnologías de telecomunicaciones. Para el autor la fricción de la distancia y las barreras geográficas desaparecerían gracias a la posibilidad de la comunicación en tiempo real. En respuesta al planteamiento de la muerte de la distancia y, por ende, la muerte de la geografía y del espacio para el análisis social autores como Castells (1995); S. Graham y Marvin (2001); S. Graham (2002) apuntaron a la importancia de concebir los avances tecnológicos y su configuración espacial como un hecho social y vinculado a las desigualdades geográficas preexistentes. Así, los autores identificaron que la infraestructura informacional se encontraba inequitativamente distribuida en el territorio, favoreciendo a polos urbanos

social y políticamente bien posicionados y, que esta circunstancia conllevaría en una ventaja tecnológica de las ciudades. Recientemente Takhteyev, Gruzd, y Wellman (2012) han mostrado, para el caso de Twitter, que la geografía física sigue configurando las redes sociales en plataformas digitales. En su estudio, los autores muestran que las redes de los usuarios de Twitter alcanzan, como máximo, 100 kilómetros de radio y, que estas redes se establecen primordialmente entre habitantes de núcleos urbanos. De esta manera, lo digital continúa vinculado a lo espacial.

La irrupción de los avances digitales no solo se vio manifestada en la distribución desigual de la infraestructura tecnológica, también ha sido un parteaguas de la manera en cómo se conceptualiza y se construye cartografía en la actualidad. Para Elwood, Goodchild, y Sui (2012); Goodchild (2007); Goodchild, Aubrecht, y Bhaduri (2017) la revolución de los grandes datos ha permitido avanzar en la generación de cartografía digital y en las técnicas para su análisis.

De esta manera, los grandes datos basados en plataformas de Internet y la explosión de los Sistemas de Posicionamiento Global (GPS, por su siglas en inglés) han permitido generar un tipo de Información Geográfica Voluntaria, en la que los usuarios de Internet, por medio de su GPS, se han convertido en sensores humanos que generan información geolocalizada de sus actividades cotidianas. En este sentido, las redes sociodigitales han sido uno de los pilares de esta tendencia de nueva información; ya que, por medio de los datos generados por redes sociodigitales es posible conocer de manera espacial y temporal las actividades cotidianas de las personas en un espacio determinado (Elwood, 2010).

Finalmente, la geografía no solo ha modificado su método en torno a la generación de cartografía tras la revolución digital y de los grandes datos. En las últimas décadas se han generado aproximaciones teóricas desde la geografía crítica que han avanzado en el entendimiento del papel de la dimensión digital en la geografía. En esta vertiente se encuentran posturas como las de Shaw y Graham (2017) que plantean un derecho informacional a la ciudad. Los autores parten de una idea lefebvriana del derecho a la ciudad y a la producción social del espacio, con ello estudian la manera en la que empresas como Google, por medio del poder que les otorga el algoritmo, pueden generar espacios urbanos desiguales, al incrementar o disminuir la presencia de determinados espacios en plataformas digitales y cómo eso se vincula con poderes económicos o políticos.

La vertiente de las geografías digitales también ha teorizado sobre la dimensión digital y el espacio geográfico. Al respecto propuestas como las de Kitchin y Dodge (2011) o Leszczynski

(2019) apuntan a la construcción de diferentes espacialidades en función de la dimensión de lo digital que se analice. Para los primeros autores, el poder del código o el algoritmos y la omnipresencia de éste produce de manera diferenciada geografías digitales. Por su parte, para Leszczynski (2019) la generación de espacialidades en la actualidad se basa en el papel activo que tienen las redes sociodigitales como Twitter o Instagram para co-producir espacios e imaginarios en las personas. Además, dentro de esta corriente, se puede mencionar el trabajo de M. Graham, Zook, y Boulton (2013a) el cual apunta a una Realidad Aumentada Digitalmente.

En la propuesta de la Realidad Geográfica Aumentada Digitalmente (RGAD) la constitución de un lugar por medio de las redes sociodigitales no solo refleja el lugar físico, sino que lo produce en el terreno de lo digital. Así, dicho lugar en lo digital es sustancialmente diferente al espacio físico sin interacción de los medios digitales. Para ello, la propuesta se basa en tres poderes que explican la manera en la que reflejan y se producen espacialidades por medio de la intervención de herramientas digitales, como las plataformas de redes sociodigitales. En primera instancia, los autores apuntan a que existe un poder desigualmente distribuido en el acceso a Internet y a redes sociodigitales. Con respecto a esto, la RGAD retoma la distribución desigual del acceso a Internet entre la población. Desigualdad que se encuentra relacionada con desigualdades socioeconómicas preexistentes.

En segundo lugar, la RGAD propone entender una realidad aumentada digitalmente por medio del poder de la comunicación de la redes sociodigitales. En este punto, la propuesta establece una conexión entre la función principal de las redes sociodigitales, la cual es la comunicación entre personas, y los mecanismos que posibilitan el reflejo del espacio y las experiencias vividas; al mismo tiempo, se produce una realidad digital diferente a la física. Así, la propuesta apunta a que lo transmitido mediante plataformas digitales tiene un impacto diferenciado en función del emisor y sus palabras. De esta manera, si una celebridad o figura del mundo político comunica mediante Twitter un contenido referente a un restaurante en una colonia gentrificada causará mayor impacto que si lo hace un usuario de menos alcance público.

La tercera dimensión de la RGAD apunta al papel, poco explorado, que el código subyacente ejerce en la creación de realidades digitales. Al respecto se puede mencionar, por ejemplo, que el código de Google para mostrar los resultados de una búsqueda en Google Maps no se encuentra exento de interés políticos y económicos. Así, la propuesta de la RGAD recupera una perspectiva

política y social sobre el papel del código en la conformación de la revolución digital. Así, en esta aproximación teórica, la combinación de los tres poderes permite entender la manera en la que la dimensión digital interviene para que la realidad 'física' no solo se reproduzca, sino que se produzca una realidad ontológicamente diferente a la que existe sin la intervención de lo digital (M. Graham y Zook, 2013; M. Graham, Zook, y Boulton, 2013a, 2013b; M. Graham y cols., 2015; M. Graham, 2017; Shaw y Graham, 2017).

A manera de recapitulación, se han expuesto cuatro puntos fundamentales que conforman los antecedentes teóricos y epistemológicos del presente trabajo de investigación. En primer lugar se expuso el proceso de gentrificación como un fenómeno urbano contemporáneo que configura la ciudad bajo una lógica del sistema económico capitalista que, a su vez, es una expresión de la desigualdad espacial en los entornos urbanos a nivel mundial. En este hilo conductor, se ha problematizado la emergencia de una nueva dimensión en dicho fenómeno, la revolución digital. Así, el segundo punto central ha sido exponer las aproximaciones teóricas que, desde la geografía, han reflexionado sobre la interacción entre espacio y lo digital. Al respecto, desde la geografía se han desarrollado algunas aproximaciones teóricas que han reflexionado sobre la modificación, producción y reproducción del espacio en función de la dimensión digital de interés, ya sean los dispositivos móviles, las redes sociodigitales o el código. En tercer lugar, se ha mencionado el auge de los grandes datos como expresión concreta de la revolución digital en la sociedad. En este sentido, se ha apuntado a que éstos son una nueva fuente de información generada por la creciente hiperconectividad en Internet. Pero, al mismo tiempo, no solo son reflejo inocuo de las actividades en dicha esfera, sino que se encuentran económica y políticamente constituidos. Ejemplo de ello son las redes sociodigitales, las cuales no solo han sido herramienta de comunicación entre las personas, sino plataformas de control político, agenda pública e, incluso, de vigilancia. Finalmente, se ha expuesto la aproximación teórica de la RGAD, la cual considera la producción/reproducción desigual del espacio por medio de los medios digitales. En este sentido la RGAD, al igual que las posturas críticas de la gentrificación, considera que la conformación social del espacio es un fenómeno que habla sobre las desigualdades sociales en el espacio urbano. Esto por medio de la desigualdad en el acceso a la dimensión digital, expresada en la desigualdad en el uso de Internet. Aunado a eso, el poder que la comunicación en el terreno digital tiene para producir/reproducir realidades geográficas y, finalmente en el poder del código que subyace a la dimensión digital.

Dentro de este marco, se han encontrado diferentes brechas en el conocimiento acerca de la interacción entre el proceso de gentrificación y los mecanismos para que ésta se manifieste en el ámbito digital. Con esto en mente, la presente investigación tiene una pregunta general que guía y enmarca la investigación: **¿Cómo el proceso de gentrificación en la Ciudad de México se ve aumentado digitalmente mediante el uso de Twitter?**

Dada la complejidad de la pregunta general, que involucra todas las dimensiones y problemáticas expuestas anteriormente, ésta se apoya de las siguientes preguntas específicas:

1. ¿Cómo el proceso de gentrificación contemporáneo puede verse aumentado por las dimensiones de la Realidad Geográfica Aumentada Digitalmente?
2. ¿Cómo los *grandes datos* pueden construir un cambio en la nueva aproximación epistemológica para entender los fenómenos urbanos, particularmente la gentrificación?
3. ¿Cuáles son los determinantes sociodemográficos que explican la desigualdad digital – uso de Internet, redes sociodigitales y Twitter – en México y, específicamente en la Zona Metropolitana?
4. ¿Cómo esta desigualdad digital en el uso de redes sociodigitales se entrelaza con el proceso de gentrificación aumentada digitalmente en la Ciudad de México?
5. ¿Cuál es el papel de los grandes datos, particularmente los recolectados de Twitter, en la generación de una gentrificación digitalmente aumentada en la Ciudad de México?
6. ¿Cuáles son los patrones de emisión espacio-temporal de los tweets georreferenciados en la Ciudad de México?
7. ¿Cuáles son las colonias gentrificadas comercialmente en la Ciudad de México?
8. ¿Los patrones de emisión se relacionan espacial y temporalmente con colonias y horarios vinculados con prácticas gentrificadas en la Ciudad de México?
9. ¿Cómo la emisión espacio-temporal desigual de tweets en la Ciudad de México aumenta digitalmente el proceso de gentrificación ?
10. ¿ Las palabras emitidas en Twitter en la Ciudad de México pueden ser signos lingüísticos explicados bajo el del proceso de gentrificación?
11. ¿Dichas palabras se relacionan de manera espacio-temporal con lugares y horarios vinculados con el proceso de gentrificación en la ciudad?
12. ¿Cómo la distribución territorial de tópicos identificados del corpus de tweets se vincula con

prácticas de gentrificación?

Cómo se ha mencionado a lo largo de esta introducción, uno de los pilares de esta investigación se propone profundizar en el conocimiento entre los grandes datos y la dimensión digital de los procesos urbanos. En este sentido, para el presente proyecto se hará uso de datos provenientes de la plataforma de Twitter. En consecuencia, se tienen preguntas transversales en el terreno de la fuente de información, ya que ésta es de carácter novedoso y es otro aporte a los estudios urbanos:

13. ¿Cuáles son las ventajas y limitaciones de los datos cosechados de Twitter para los fenómenos urbanos como la gentrificación?
14. ¿Cuáles son los principales retos técnicos y epistemológicos para trabajar con este tipo de información dentro del campo de los estudios urbanos?
15. ¿Cómo se puede vincular el papel no neutral de los grandes datos con las transformaciones sociales relacionadas con la revolución digital?

Dada la complejidad de la pregunta de investigación, en el desarrollo del trabajo, se han seguido diferentes aproximaciones metodológicas. Éstas han sido guiadas por la necesidad de establecer, por un lado, la interacción entre la dimensión digital y los fenómenos urbanos, discusión que se encuentra en fases iniciales dentro de la literatura internacional. Y, por otro lado, la interacción entre las dimensiones netamente teóricas y la implementación empírica de los grandes datos en estudios urbanos.

De esta manera, en el capítulo 1 se presenta una propuesta metodológica de carácter teórico que busca construir un andamiaje conceptual que amalgame el fenómeno de gentrificación, los grandes datos y la Realidad Geográfica Aumentada Digitalmente. Por su parte, en los capítulos 2 y 3 se desarrollan elementos teóricos específicos de la propuesta teórica para entender la gentrificación aumentada digitalmente y, con base en dichos conceptos, construir indicadores empíricos para realizar las aproximaciones pertinentes. En otras palabras la construcción de indicadores empíricos se encuentra supeditada al andamiaje conceptual. Esto se ha sido posible ya que los dos apartados previos han sido contruidos con base en un conjunto amplio de literatura que se ha aproximado desde diversos enfoques y posturas a la gentrificación digital. Y, han propuesto metodologías empíricas que han sido reproducidas y han dado pie a evidenciar las fortalezas y limitaciones de éstas.

En el capítulo 4 el enfoque teórico-metodológico se construye con base en la exploración

empírica guiada por una interpretación teórica. Es decir, en este capítulo no se construyen indicadores específicos para explorar empíricamente el argumento teórico, sino que el argumento teórico ha sido herramienta fundamental para analizar lo que ha sido revelado por los datos. El cambio metodológico en dicho acápite se debe a dos razones fundamentales. La primera de ellas es la complejidad que ha implicado el manejo de los grandes datos vinculados con el Procesamiento de Lenguaje Natural y los pocos trabajos que han versado al respecto. La segunda razón es la elección de los algoritmos de procesamiento para los datos que se han mencionado. A pesar de que este capítulo puede parecer un guiño postpositivista de los grandes datos, la interpretación ha sido guiada por una construcción teórica en términos sociales y lingüísticos para fortalecer el argumento de la gentrificación digitalmente aumentada.

Dados los ejes teóricos, la pregunta general, las aproximaciones metodológicas y los retos que esta investigación implica, a continuación se expone el desarrollo de cada uno de los capítulos.

En el capítulo 1 titulado **“Gentrificación digitalmente aumentada. Hacia una propuesta teórica para entender la dimensión digital de la gentrificación”** se establecen el marco teórico y conceptual de la propuesta para entender una gentrificación digitalmente aumentada. En primera instancia, el capítulo desarrolla los elementos que han sido discutidos en torno a la problemática de la gentrificación. En este sentido, se exponen las principales corrientes que han versado sobre dicha temática, desde las primeras aproximaciones en la década de 1960 en Londres hasta la perspectiva de la gentrificación guiada por el Estado. De manera particular, se exponen dos enfoques que serán pilares para el ensamblaje teórico de la gentrificación digitalmente aumentada como caso particular de la Realidad Geográfica Aumentada Digitalmente (RGAD). Estos enfoques son la teoría culturalista desarrollada por Sharon Zukin y la gentrificación como expresión del desarrollo urbano desigual propuesta por Neil Smith. La revisión conceptual sobre la gentrificación no solo permite el ensamblaje teórico, sino que permite dialogar empíricamente con indicadores de gentrificación que serán desarrollados en capítulos posteriores.

Como segundo gran apartado de este capítulo se expone una visión crítica de los grandes datos y su aproximación con los estudios urbanos. En él se desarrollan algunas ideas sobre los grandes datos que han emergido desde las ciencias computacionales y de datos, así como de algunos intentos conceptuales por entender la aplicación de este tipo de datos en las ciudades. Dentro de este apartado, se expone una visión crítica sobre los grandes datos que será recuperada a lo largo de

todo el trabajo de investigación. Esta postura señala que los grandes datos se construyen de manera política y social y en ese tenor deben ser utilizados, con el fin de conocer sus potencialidades y sus sesgos de origen.

En tercer lugar se desarrolla la aproximación teórica de la RGAD. En este punto, el capítulo se enfoca en exponer las aproximaciones que han emergido dentro de la corriente de las geografías digitales y, la propuesta de la RGAD como una de ellas. Particularmente, se desarrollan los argumentos básicos de cómo dicha aproximación propone entender la interacción entre los lugares 'físicos' y la producción de una realidad digital que no solo reproduce la realidad geográfica sino que, gracias a la mediación de los elementos digitales, se puede hablar de una nueva realidad digital que interactúa y refuerza los fenómenos proyectados en medios digitales. Esta postura se ha elegido para una aproximación de la gentrificación debido a que dicha postura hace explícitos elementos de desigualdad social y algorítmicos que intervienen en la sinergia de lo digital y lo real. Este enfoque propone identificar una RGAD por medio de tres poderes que interactúan entre ellos y que se encuentran distribuidos de manera desigual en la sociedad. Estos son el poder de la distribución del acceso a la dimensión digital, el poder de la comunicación y el poder del código.

Finalmente, en el capítulo se expone una propuesta teórica original que apunta al ensamblaje entre las teorías culturalistas y de desarrollo desigual de la gentrificación y cómo éstas se ven aumentadas por la intervención de los diferentes poderes de la RGAD. En este orden de ideas, se desarrolla cómo la distribución desigual en el acceso a Internet y a las plataformas digitales pueden intervenir en la producción de concentración o igualación de tendencias gentrificadoras en las ciudades. O bien, cómo el poder de la comunicación que se encuentra concentrado en unos cuantos usuarios de redes sociodigitales puede favorecer la valorización o revalorización de determinados espacios gentrificados. Al mismo tiempo, cómo dicha valorización del espacio gentrificado por medio digital se vincula con la generación y consolidación de símbolos y signos que aluden a una cultura de la gentrificación.

Este desarrollo teórico será la base para evaluar empíricamente cómo el fenómeno de la gentrificación en la Ciudad de México se ve aumentado digitalmente por el uso de Twitter. Dicho argumento servirá para comprender por qué las personas acceden de manera diferenciada a las redes sociodigitales y si esto se vincula con desigualdades sociales preexistentes; además si el uso de estas redes se ve reflejado en los patrones espaciotemporales de emisión de tweets en la ciudad. Y, por

último, se podrá evaluar si existe la generación de símbolos y síntomas lingüísticos en Twitter que co-incidan en las colonias gentrificadas. Con lo anterior, se plantea que la gentrificación aumentada digitalmente es un caso particular de la RGAD, pero, al mismo tiempo, es una propuesta novedosa para entender las diferentes dimensiones del proceso de gentrificación actual y su posible evolución ante la inevitable irrupción de lo digital en todos los aspectos de la vida urbana contemporánea.

El segundo capítulo, denominado **“Desigualdad Digital y perfiles de uso de Redes Sociales y Twitter. El poder distribuido de la dimensión digital”**, tiene como objetivo identificar las principales características sociodemográficas de los usuarios de Twitter en la Zona Metropolitana de la Ciudad de México. Con esto en mente, el capítulo desarrolla a profundidad la discusión sobre las desigualdades digitales y los determinantes sociodemográficos que intervienen en el acceso inequitativo de las personas a Internet y, particularmente para el uso de Twitter. En este sentido, el acápite también profundiza en el entendimiento de la desigualdad digital territorial en México. Para ello en dicho capítulo se exponen las aproximaciones teóricas que versan sobre los determinantes sociales en el acceso a la dimensión digital.

En este tenor, se identifican elementos estructurales como la desigualdad de género, edad, escolaridad y estrato económico como los principales elementos que impiden el acceso a la dimensión digital. Al mismo tiempo, el enfoque del capítulo contextualiza el uso de Twitter como una herramienta particular dentro del vasto universo en el uso de Internet. Esto permite comprender que el uso de Twitter enfrenta una doble barrera en el acceso. Es decir, la primera barrera que se debe superar es el acceso a Internet y la segunda es el uso de redes sociodigitales.

En el terreno empírico, para la evaluación de los perfiles sociodemográficos de los usuarios de Internet, redes sociodigitales y Twitter se construyen indicadores basados en las aproximaciones teóricas con la ayuda de la Encuesta Nacional de Disponibilidad y Uso de Tecnologías de Información en los Hogares (ENDUTIH) 2018, generada por el INEGI. De esta manera se seleccionan variables relacionadas con los diferentes usos y los determinantes sociodemográficos que ha sido posible construir con la información disponible. Todo esto, para culminar con la aplicación de modelos logísticos binomiales cuyas variables a explicar son el uso de Internet- redes sociodigitales y Twitter en función de las variables sociodemográficas.

Por último, el capítulo apunta hacia la identificación del uso estratificado y desigual de la dimensión digital en México y en la Zona Metropolitana de la Ciudad de México. Es decir, se

muestra que la principal barrera para ser partícipe de la dimensión digital se relaciona con el estrato económico, el sexo y la escolaridad de las personas. Una vez que se ha superado dicha barrera, el uso de Twitter se ve acentuado por el estrato socioeconómico y la escolaridad. En este sentido, el estudio arroja luz de los perfiles de uso de dicha red sociodigital. Y, al mismo tiempo, permite vislumbrar que la dimensión digital no es homogénea en su uso en toda la población. Con ello, se apuntala el argumento de que una realidad geográfica como la gentrificación se aumenta digitalmente en función de las personas que hacen uso y emiten contenido en las plataformas digitales.

El tercer capítulo se denomina **“Gentrificación aumentada en la Ciudad de México. Análisis de la distribución y comunicación digital mediante los patrones de emisión espacio-temporal de Twitter”**. En este capítulo se expone el comportamiento de la emisión espacio-temporal de Twitter en la Ciudad de México en el contexto de colonias identificadas como gentrificadas bajo criterios comerciales y de intervención estatal.

De manera teórica, el capítulo apuntala a la distribución digital desigual en términos espaciales de las emisiones de Twitter en el espacio urbano de la Ciudad de México, con el objetivo de identificar patrones de centralización de emisión digital de Twitter asociados con espacios gentrificados. En primera instancia, en el capítulo se ofrece una discusión teórica de los abordaje teóricos desde la sociología y la geografía sobre los mecanismos de Twitter en la sociedad contemporánea. Particularmente, se desarrolla el papel de Twitter como un espacio social explicado bajo las interacciones de los usuarios. Al mismo tiempo, se expone el papel que Twitter ha significado como nueva fuente de información para la generación de datos georreferenciados. Y finalmente, se analiza como Twitter es una herramienta en ciernes en el campo de los estudios urbanos.

En el terreno empírico, en el capítulo se desarrollan dos retos metodológicos. Por un lado, la identificación de colonias gentrificadas en la Ciudad de México. En este punto, el trabajo ofrece una identificación que retoma elementos de la gentrificación comercial y la guiada por el Estado como centralidad de ciertos comercios vinculados con prácticas de consumo gentrificadas y la presencia de políticas de movilidad no motorizada impulsadas por el gobierno de la ciudad. Todos estos indicadores han sido construidos a nivel de colonia, lo cual es poco común en los estudios de identificación de espacio gentrificados en la Ciudad de México. Por otro lado, se mencionan los retos en el manejo de grandes bases de datos como Twitter y las características de éstos en cuanto al pre-procesamiento de dichas bases de datos. Igualmente, se detalla la generación de indicadores

espacio temporales basados en la emisión de tweets. Es importante señalar que la dimensión temporal cobrará relevancia debido a que la construcción de espacios digitales se basa en la dinámica temporal urbana.

Como principales resultados del capítulo 3 se muestran la concentración de emisión de tweets en las colonias gentrificadas de la Ciudad de México y de algunos puntos comerciales y de entretenimiento en los municipios de la Zona Metropolitana de la Ciudad de México. En términos temporales se muestra que la emisión de tweets en colonias gentrificadas y puntos seleccionados de la ciudad se asocia a actividades laborales y de consumo-entretenimiento en la ciudad. A través del análisis detallado en términos temporales se identifica que el principal centro de emisión de Twitter en la ciudad de México se co-localiza, principalmente en las colonias gentrificadas a lo largo del día. Aunque, este patrón se intensifica en los horarios y días de esparcimiento, es decir, durante las noches de fines de semana. Fuera de dichas colonias se identifican núcleos de concentración en centros comerciales y centros masivos de entretenimiento como estadios o recintos de conciertos. Estos hallazgos apuntan a pensar en el uso selectivo de Twitter en la ciudad con fines de manifestación de determinadas prácticas en la esfera digital. Así, se muestra que Twitter, no solo tiene un sesgo en términos del perfil sociodemográfico de sus usuarios, sino en su patrón espacial y temporal.

En el capítulo 4 titulado **“El poder de la comunicación en la gentrificación aumentada digitalmente en la Ciudad de México”** se muestra un análisis lingüístico de las emisiones de Twitter, con el objetivo de indagar sobre el poder de la comunicación digital en la generación de símbolos de gentrificación en la esfera de lo digital. Como se ha mencionado anteriormente, este capítulo es de naturaleza deductiva, en cuanto a la modelación de la información de Twitter y su interpretación teórica.

A pesar de esto en el capítulo se desarrolla, en primer lugar, una sección teórica donde se exponen los principales trabajos que han comenzado a establecer el vínculo entre lo digital y el fenómeno de gentrificación, en el terreno de la formación de signos lingüísticos relacionados con dicho proceso. En el mismo aparatado teórico, se retoma la propuesta teórica de la función lingüística de Karl Bühler con el fin de ofrecer un desarrollo e interpretación teórica a las palabras emanadas del corpus de tweets. Esta propuesta pone sobre la mesa de discusión la posibilidad de interpretar los temas identificados en tweets como signos lingüísticos de un síntoma social, que en este caso

es el proceso de gentrificación en la ciudad. De esta manera, se propone dar una base teórica desde un enfoque sociolingüístico a los resultados del análisis de procesamiento de lenguaje natural.

El capítulo también abona al conocimiento en cuanto al manejo de corpus de Twitter para los estudios urbanos. En este campo, se mencionan los principales obstáculos y limitaciones de la base de texto de tweets recolectados. Al mismo tiempo, se propone identificar entidades nombradas relacionadas con lugares en el texto de los tweets, con el fin de proporcionar mayor capacidad interpretativa en el modelaje dicho texto. Por último, se explica el funcionamiento del algoritmo de aprendizaje de máquina de Asignación Latente de Dirichlet. Dicho algoritmo es una herramienta de aprendizaje de máquina no supervisado que permite construir campos léxicos latentes emanados del corpus de tweets.

Los resultados de este capítulo apuntan a la identificación de los principales temas léxicos en los tweets emanados de las colonias gentrificadas de la Ciudad de México. En este sentido, se pueden señalar el dominio de temas vinculados con procesos de gentrificación, tales como el tema vinculado a la cultura *hipster*, el uso constante de anglicismos y las referencias al consumo en espacios cerrados, como centros comerciales. Al mismo tiempo, al seguir una herramienta no supervisada se identifican otros temas que no se vinculan directamente con el proceso de gentrificación, pero que emergen de dichos espacios. Ejemplo de esto son temas relacionados con actividades de oficina. Este hallazgo abre la puerta a la discusión sobre la formación de diversos signos lingüísticos que coexisten en lugares gentificados o en proceso de valorización. Finalmente, en el acápite se muestra la representación cartográfica de los principales temas lingüísticos emanados de los tweets en la ciudad. Este hallazgo lleva a pensar en la conformación de espacio lingüísticos que se basan en los fenómenos urbanos estructurales.

A lo largo del desarrollo de los cuatro capítulo se ensamblan dos elementos particulares de la gentrificación aumentada digitalmente. El primero de ellos es la exploración de la desigual distribución en el acceso a Internet. Esto se muestra por medio del hallazgo de perfiles sociodemográficos de usuarios de Twitter pertenecientes a los estratos medio y alto. Desde la perspectiva espacial se ha mostrado que la distribución en el territorio en la emisión de tweets en la Ciudad de México se encuentra concentrada en las colonias que han sido identificadas como gentrificadas. Así, se establece un vínculo entre los usuarios y el espacio usado por ellos. En otras palabras, lo visto y producido en Twitter en la ciudad de México se encuentra sesgado por sus productores digitales,

usuarios de alto estrato que emiten, principalmente, desde espacios centrales.

En segundo lugar, se ha apuntalado la dimensión del poder de la comunicación por medio de los tweets. Esta dimensión ha implicado hacer uso de técnicas de frontera dentro de la ciencia de datos y su interpretación teórica desde una postura sociolingüística. Al igual que la distribución desigual de emisión, los espacios lingüísticos generados mediante Twitter responden a prácticas espaciales vinculadas con el proceso de gentrificación en la ciudad. En este sentido, más allá de reproducir inocuamente lo vivido en dichos espacios, los usuario y el código de Twitter intervienen con elementos léxicos y de código para generar tendencias. Así, el proceso de gentrificación se aumenta digitalmente por medio de un paisaje lingüístico digital explicado por experiencias en los espacios identificados.

Al mismo tiempo, el trabajo de investigación es propositivo en el ámbito teórico y empírico. Por un lado, si bien el ensamblaje entre las teóricas previas sobre la gentrificación y la RGAD no se ha desarrollado cabalmente, sí se han desarrollado los puntos principales para entender la gentrificación digitalmente aumentada bajo un cuerpo teórico que recupera la complejidad de la nueva realidad digital. Por otro lado, se exponen los principales retos y bondades del uso de grandes bases de datos, como las que se pueden obtener de Twitter, para estudios urbanos. En este sentido se muestran diversas herramientas provenientes de la estadística probabilística clásica hasta los avances en el aprendizaje de máquina. Así, el presente trabajo también puede ser interpretado como un avance metodológico en el uso de nuevas herramientas de orden cuantitativos en los estudios urbanos. De manera que, la presente investigación aborda dos grandes preocupaciones en la discusión contemporánea sobre los estudios urbanos: primero, la necesidad de teorizar sobre los fenómenos urbanos clásicos y su cambio ante la revolución de la sociedad digital; y segundo, avanzar en el uso de nuevas herramientas cuantitativas ante la revolución de los grandes datos.



# Capítulo 1

## Gentrificación digitalmente aumentada. Hacia una propuesta teórica para entender la dimensión digital de la gentrificación

### 1.1 Introducción

El propósito de este capítulo es sentar las bases teórico-metodológicas para entender el proceso de gentrificación como una Realidad Geográfica Aumentada Digitalmente. Para lograr este objetivo a lo largo de este capítulo se ofrecen diversos elementos que permiten llegar a dicha postulación teórica y que posteriormente serán utilizados para comprobarla de manera empírica con un caso de estudio. En primero lugar, se apuntala el concepto de *gentrificación*. Dicho concepto es uno de los más polémicos dentro del campo de los Estudios Urbanos. En particular, el presente trabajo se aleja de la concepción de desplazamiento de población de clase trabajadora en respuesta a la llegada de clase media alta a los espacios centrales de la ciudad. Así, el propósito teórico de discutir la gentrificación tiene dos objetivos particulares. Por un lado, desarrollar los elementos teóricos conceptuales que permitan identificar espacios gentrificados de manera empírica, desde una perspectiva de gentrificación comercial y dirigida por el Estado. El segundo objetivo particular es exponer una perspectiva teórica de corte estructural que permita vincular dicho proceso urbano con la propuesta de Realidades Geográficas Aumentadas. En concreto, en este punto se retoma la propuesta de Neil Smith sobre la gentrificación como reflejo de un desarrollo desigual en el espacio urbano.

En segundo lugar, se desarrolla el concepto de la revolución de los Grandes Datos en geografía y en particular en Estudios Urbanos. Este apartado tiene, a su vez, dos objetivos particulares. El primero de ellos es dar un breve estado del arte de la situación de las Ciencias Sociales ante la revolución de los Grandes Datos. En este apartado se pone especial énfasis en los retos particulares de la geografía en esta área, por medio del concepto de la *Geoweb*. El segundo objetivo específico de esta sección es exponer las principales posturas teóricas que vinculan la geografía con la dimensión

digital. En este sentido se presentan las aproximaciones teóricas que versan sobre la constitución de espacialidades digitales. Dentro de las anteriores, se retoma, particularmente, la propuesta de la Realidad Geográfica Aumentada Digitalmente.

En tercer lugar, se desarrolla una propuesta teórica para entender la gentrificación desde la perspectiva de la Realidad Aumentada y la conformación de una espacialidad digital particular. Para ello, se recuperan las propuestas del desarrollo urbano desigual de Smith y se compagina con el aporte de la Realidad Aumentada, de tal suerte que ambas conforman un ensamblaje teórico-metodológico para conceptualizar el problema de interés. Así, se expone la manera cómo se entrelazan las dimensiones del poder de la distribución, comunicación y del código propuestas por la Realidad Aumentada con el proceso de desarrollo desigual en sus elementos constitutivos de creación de fuerzas centro/periférica y valorización/revalorización del espacio. En consecuencia, se propone una aproximación teórica novedosa para entender cómo el proceso de gentrificación se ve aumentado digitalmente y con este todas su repercusiones.

Por último, se presentan las consideraciones finales del capítulo. En éstas se desarrolla el vínculo entre el marco teórico y el desarrollo del caso particular de estudio de esta investigación. Es decir, la manera en la que las categorías desarrolladas a lo largo del acápite serán examinadas en el caso del proceso de gentrificación en la Zona Metropolitana de la Ciudad México mediante la plataforma de Red Sociodigital de Twitter.

## **1.2 Sobre el concepto de gentrificación**

El concepto de gentrificación ha sido uno de los más debatidos en los últimos 50 años en el campo de los Estudios Urbanos. Desde que Ruth Glass acuñó el término para dar cuenta del proceso que observó en las zonas centrales de Londres en los años 1960, la literatura internacional ha abordado, explicado y criticado el concepto. A lo largo de la última década, el concepto de gentrificación ha sido nutrido con diferentes dimensiones y posturas teóricas que han visto en éste, un elemento explicativo de la dinámica contemporánea de las ciudades a nivel mundial. De tal manera que, el concepto y el fenómeno de gentrificación en la actualidad puede ser analizado desde muchas vertientes. Por ejemplo, desplazamiento de población o reconfiguración social, teoría de la renta diferencial del suelo, la intervención estatal, aspectos culturales y de consumo, gentrificación planetaria, etc. Dada la variedad de elementos teóricos para entender la gentrificación en el siglo XXI,

solo se retoman ciertos elementos de algunas aproximaciones teóricas para entender cómo la gentrificación contemporánea se trastoca con el desarrollo y la omnipresencia de las tecnologías de la computación y, en particular de las redes sociodigitales.

En términos esquemáticos, el concepto de gentrificación puede ser estudiado desde dos perspectivas de análisis. La primera apunta al estudio de la evolución del fenómeno. Abordar el desarrollo histórico de la gentrificación permite clarificar la evolución de éste y la aparición de nuevas dimensiones que intervienen en su producción y reproducción (Hackworth y Smith, 2001). Entre estos elementos se identifica el papel del Estado, los elementos culturales o los nuevos desarrollos tecnológicos. En estos últimos se ubica el papel de Internet y las Redes Sociodigitales en la conformación de nuevas espacialidades gentrificadas digitalmente.

El segundo gran eje de análisis sobre la gentrificación son los métodos empíricos para estudiarla. Estas metodologías se encuentran limitadas a los abordajes teórico-epistemológicos de la gentrificación, así como a la disponibilidad de fuentes de información. Los abordajes epistemológicos de la gentrificación han guiado gran parte de las metodologías empíricas para su análisis, particularmente, se han centrado en aproximaciones de índole cualitativo con entrevistas a profundidad y acción participativa. Esto debido a que gran parte de la preocupación ha sido sobre el desplazamiento de población trabajadora en favor de clase media-alta. Por otro lado, los abordajes cuantitativos también han sido limitados debido a que la complejidad del fenómeno impide el uso de estadísticas oficiales de carácter estático. Sin embargo, la evolución del abordaje empírico de la gentrificación también refleja el cambio en el fenómeno.

En este sentido, el presente trabajo también pretende abonar al campo de discusión sobre la identificación de la gentrificación y las constantes transformaciones sobre éste. En primera instancia, debido al uso de un complejo ensamblaje de fuentes de información, tanto oficiales como provenientes de Internet. Con base en este ensamblaje de información el trabajo de investigación pretende aportar una metodología de aproximación cuantitativa en la identificación de colonias gentrificadas y la representación de dimensiones culturales de consumo en estos espacios. En segunda instancia, en este complejo entramado de información subyacen acercamientos epistemológicos sobre las transformaciones tecnológicas que posibilitan la generación de fuentes de información en Internet y que al mismo tiempo añaden nuevos elementos al proceso de gentrificación.

### *1.2.1 El proceso de gentrificación en el siglo XXI*

De manera general, la gentrificación se entiende como el proceso en el cual, áreas deprimidas de la ciudad comienzan a ser pobladas por estratos medios-altos. Esto trae consigo múltiples consecuencias en dichos espacios. El alza del precio de las rentas en la habitación ante una nueva demanda con mayor poder adquisitivo, en primera instancia. Y, como consecuencia, el desplazamiento de la población originaria ante la imposibilidad de pagar dichos precios. Es así que, “La gentrificación como concepto quedó asociada entonces a los efectos sociales y urbanos negativos de los procesos de renovación urbana en áreas centrales. Desde esta perspectiva, la clase trabajadora se ve desplazada por una naciente clase media vinculada a la economía de servicios” (di Virgilio y Guevara, 2015, p: 35). Así, la gentrificación es de interés debido a que el proceso de desplazamiento de población de estrato bajo en favor de los estratos altos, la cual lleva consigo una homogeneización social y cultural de los espacios transformados, conlleva la generación de espacios de emancipación y disputa en la construcción de urbanidad (Lees y cols., 2016).

La gentrificación ha sido analizada primordialmente desde dos vertientes, desde el cambio de los precios de vivienda basado en el diferencial de la renta del suelo y, segundo desde el desplazamiento de la población de estratos bajos. El trabajo de N. Smith (1979) apuntó a explicar la gentrificación, no como una causa de oferta y demanda de vivienda ante la recomposición poblacional de la postguerra en países del norte global. Sino que, propuso una explicación estructural de la configuración de la ciudad basada en el diferencial de la renta entre el centro y periferia del espacio urbano. En otras palabras, el diferencial de la renta es aquel que se muestra en el precio de suelo entre su valor capitalizado y su valor presente. El autor explica que, en primera instancia, los espacios centrales de la ciudad se encuentran deteriorados lo que hace que su valor presente sea relativamente bajo. Ante esto, el sector de inversión inmobiliaria aprovecha este valor presente para capitalizarlo por medio de la reinversión y el re-desarrollo del espacio y de esta manera incrementar el precio del suelo después del proceso de revalorización. A decir de Smith, una vez que el diferencial de la renta es suficiente, se puede iniciar el proceso de gentrificación por diferentes actores conductores tales como el mercado de suelo y de vivienda. Así, la gentrificación se explica por factores de precio de la vivienda a nivel de colonia y no simplemente por acciones individuales como la corriente económica neoclásica sugería hasta entonces (N. Smith, 1979).

La segunda vertiente más estudiada de la gentrificación ha sido el desplazamiento de la po-

blación de estrato bajo ante el incremento del diferencial de la renta, mencionado anteriormente. La gentrificación ha sido analizada como un proceso que incrementa la injusticia espacial de la ciudad. En ese sentido, la gentrificación provoca que espacios privilegiados de la ciudad sean construidos, habitados, vividos y consumidos solo por una élite, mientras estos espacios son ‘limpiados’ de los habitantes de estratos bajos. De esta manera, el estudio de la gentrificación adquiere un componente político y reflexivo sobre los efectos sociales del desarrollo capitalista en las ciudades (Balzarini y Shlay, 2016; Chaskin y Joseph, 2013; Hernández-Cordero, 2015; Janoschka y cols., 2014; Janoschka, 2016; Salinas, 2013a). Si bien, la renta del suelo y el desplazamiento de población son dos aspectos fundamentales en el estudio de la gentrificación, el fenómeno está en constante evolución y se caracteriza por diversas aristas. Por ende, se han estudiado otras aproximaciones teóricas que robustecen el análisis del fenómeno. En este ámbito se tienen aproximaciones de gentrificación comercial, cultural y guiada por el Estado. Todas estas propuestas permiten delimitar mejor en términos teóricos y empíricos la gentrificación.

### *1.2.2 Gentrificación comercial y cultural*

En el fenómeno de la gentrificación intervienen otros elementos que constituyen el fenómeno en su conjunto. Al respecto, la gentrificación ha sido analizada desde las personas con perfil gentrificador. Al mismo tiempo, estas personas materializan determinados estilos de vida y preferencias de consumo que coadyuvan a la gentrificación. En esta línea de ideas, la gentrificación ha sido explorada desde su dimensión cultural, comercial y de consumo.

La literatura especializada habla sobre un perfil sociodemográfico de personas gentrificadoras. Esta postura fue propuesta desde una aproximación neoclásica que explicaba el proceso de gentrificación con base en el cambio poblacional de los centros deprimidos de las ciudades en la etapa de la postguerra. Desde esta perspectiva, la gentrificación se explicaba por el crecimiento proporcional de la población joven que buscaba rentas más baratas en los barrios deprimidos del centro de las ciudades. En este sentido, las personas que se interesaban en habitar estos espacios eran jóvenes de clase media, profesionistas dedicados a las artes o actividades creativas (Florida, 2005; Hae, 2011; O’Sullivan, 2012). Desde este enfoque, la gentrificación se explica por la demanda de vivienda menos costosa con todas las facilidades de la ciudad central. Aunque el perfil de personas que habitan espacios gentrificados se mantiene constante en la actualidad, existen otros elementos que

han criticado esta visión de la gentrificación atomizada, de corto aliento y neoclásica.

Como una de las respuestas críticas a esta perspectiva se encuentra la realizada por Sharon Zukin, la cual señala que la gentrificación tiene elementos de consumo, creación de símbolos y homogeneización cultural del espacio urbano. Desde esta aproximación, la autora enmarca que el desarrollo económico de las ciudades a finales del siglo XX y principios del XXI se caracteriza por la preponderancia de los servicios y comercio. En particular, los espacios gentrificados se distinguen por la presencia de oficinas gubernamentales, oficinas centrales de empresas transnacionales y comercio en general. En este orden de ideas, para Zukin, el perfil sociodemográfico de las personas gentrificadoras explica solo una parte, ella propone que de manera general se puede hablar de la homogenización del espacio urbano por medio de la racionalidad económica del consumo (Zukin, 1987, 1998).

El proceso de gentrificación se asocia con la tendencia homogeneizadora de los espacios urbanos por medio de los estilos de vida y consumo. Zukin (1991) señala que el estilo de vida de la clase media se impone en los espacios gentrificados. En este sentido, el espacio urbano central se transforma no solo en términos de la vivienda sino comerciales. La gentrificación también se manifiesta en la transformación de espacios comerciales tradicionales en favor de espacios comerciales donde se satisfacen los intereses comerciales de la clase media. Este estilo de vida de la clase media se caracteriza por el consumo en restaurantes exóticos, bares o boutiques. Estos cambios comerciales causan, en el largo plazo, el desplazamiento de la población en espacios gentrificados dado que ni los habitantes, ni comerciantes originarios de los barrios centrales pueden sostener el alza en los precios de locales comerciales (Jayne, 2006). Por último, la creación de identidad simbólica de la gentrificación comercial se basa en la construcción cultural de la centralidad del entorno construido en la ciudad central (Zukin, 1991).

### *1.2.3 Gentrificación guiada por el Estado*

De manera histórica Hackworth y Smith (2001) identifican 3 olas en el estudio y desarrollo de la gentrificación. La tercera de ellas se desarrolla a partir de la década de 1990. Esta ola es caracterizada por la mayor participación del Estado en la generación y producción de políticas de desarrollo urbano en favor de procesos de gentrificación. Al respecto, Slater (2009) demuestra que la elección de habitación de la clase media-alta en espacios renovados no se explica solamente por una

elección individual. Por el contrario, dicha elección es condicionada por la intervención del Estado, al favorecer desarrollos inmobiliarios dirigidos a dichos estratos socioeconómicos. De tal manera que, en el proceso de gentrificación no solo intervienen actores privados tales como los productores de vivienda o los compradores de ésta, sino que el gobierno interviene al proporcionar facilidades para que esta dinámica se realice. Al respecto, Wyly (2018), acota que la “gentrificación realmente nunca ha sido un proceso geográfico que pueda ser entendido dentro de espacios específicos y acotados; siempre ha sido un proceso geopolítico, una expresión localizada de tensiones dialécticas multiescalares del urbanismo transnacional”(Wyly, 2018, p: 14-15).

A la participación activa del Estado en el proceso de gentrificación se añade la reconfiguración de éste ante el proceso de urbanización planetaria. Por un lado, el Estado se ha transformado de un rol benefactor guiado por políticas de corte proteccionista hacia modelos de intervención urbana guiadas por el emprendedurismo urbano (Harvey, 1989). En este proceso, el Estado ha pasado de la administración del espacio urbano hacia una comercialización de éste, acompañado de la intervención de entidades público-privadas. Mientras que el papel del Estado se ha transformado hacia esta vertiente de manera global, también ejecuta acciones locales en función de la valorización de espacios urbanos. En este sentido, se encuentran acciones guiadas por el Estado con el fin de valorizar determinados espacios urbanos por medio de la intervención pública, renovación del espacio público, la participación de megaproyectos, etc (Brenner, 2013; Lees, Shin, y López-Morales, 2015; Lees y cols., 2016).

El papel del Estado ha sido diferenciado en la constitución de la gentrificación en la última década. Por ejemplo, en países europeos la preocupación se centra en el papel del Estado como actor promotor de procesos de gentrificación, aunque, al mismo tiempo, su papel fundamental sea la provisión de calidad de vida para las personas. Este es el caso de Amsterdam, en el cual se ha abandonado la estrategia de Estado de bienestar encargado de organizar el desarrollo inmobiliario y se ha pasado a dejar en manos del libre mercado esta área (Uitermark y Bosker, 2014). En este escenario, el Estado se debate entre una paradoja de mantener su papel histórico de proveedor de bienestar y de liberar la producción de la ciudad al mercado, esto en el marco de las consecuencias sociales negativas asociadas a la gentrificación. Por su parte, Grange y Pretorius (2016) exponen las causas de la gentrificación guiada por el Estado en el contexto específico de Hong Kong. Los autores identifican que el proceso de gentrificación en dicha ciudad cumple con los rasgos gene-

rales de el fenómeno en las urbes occidentales, tales como el alza de los precios de la renta y el desplazamiento de población de bajos recursos. Para ellos, el papel del Estado puede ser entendido desde tres vertientes. La naturaleza del gobierno y sus acciones de re-desarrollo urbano, las instituciones subyacentes en la administración del suelo en la ciudad y la morfología misma de la ciudad (Grange y Pretorius, 2016). Finalmente, se tiene el papel del Estado en el fenómeno de la gentrificación en ciudades latinoamericanas. Al respecto, la principal preocupación del Estado y su intervención en procesos gentrificadores es la interacción entre éste y el emprendedurismo urbano.

La gentrificación guiada por el Estado en América Latina ha seguido la pauta de la promoción de la actividad económica bajo el modelo de libre mercado. En este sentido, el Estado se ha convertido en el principal promotor y generador de condiciones favorables para que la inversión privada promueva el crecimiento en la ciudad. En otras palabras, los gobiernos locales absorben los riesgos económicos para la renovación y reactivación económica, incrementando de esta manera, las ganancias para las empresas. Estas acciones, se pueden esquematizar en tres grandes dimensiones que se caracterizan por concentrarse solo en las áreas a valorizar. En primera instancia, el Estado incentiva particularmente el sector inmobiliario o el turístico. Segundo, proporcionar la mejor calidad de servicios públicos, tales como agua potable, drenaje, manejo de residuos sólidos o la generación y mantenimiento de áreas verdes. Tercero, la construcción de vialidades y accesibilidad vía transporte público masivo, como metro o *Bus Rapid Transit*. A diferencia de los enfoques europeos o asiáticos, el papel del Estado en América Latina ha sido analizado desde las acciones particulares que sustentan el incremento del precio del suelo, basada en las últimas dos estrategias (di Virgilio y Guevara, 2015; Vergara-Constela y Casellas, 2016).

Aunque se tienen diferencias del papel del Estado en la gentrificación en función del espacio específico, también existen similitudes entre todos los casos. La primera es que el Estado realiza acciones focalizadas y concentradas en ciertos espacios dentro de la ciudad con el fin de promover su valorización. Es decir, el Estado ha dejado de lado su papel de proveedor de bienestar para la mayoría de la población urbana en favor de la conformación de un pequeño segmento de la población. El segundo elemento común, es el discurso oficial de la gentrificación procedente del Estado. En la mayoría de los casos, no se habla directamente de gentrificación, sino que se utilizan términos como renovación, revitalización, generación o renacimiento. Esto, a decir de algunos autores, se ha realizado con el fin de ocultar la esencia de clase que conlleva el proceso de inversión en ciertas

zonas de la ciudad, el desplazamiento de población de estrato bajo y la posible movilización social antigentrificación (di Virgilio y Guevara, 2015; Grange y Pretorius, 2016; Lees y cols., 2015, 2016; Uitermark y Bosker, 2014; Vergara-Constela y Casellas, 2016).

Mención especial merece el papel de las políticas del Estado en la provisión de servicios públicos de transporte, y en particular, lo referente a la dotación de transportes como la bicicleta. En este aspecto, la literatura reciente sobre gentrificación ha comenzado a mirar la relación existente entre dicho proceso y la movilidad no motorizada en las ciudades. Al respecto, algunos autores señalan que se encuentran asociados los procesos de gentrificación y renovación de las ciudades con el discurso modernizador de la movilidad en éstas. De esta manera, el uso de bicicleta en las urbes ha interpretado como un símbolo de progresismo y ha sido adoptado por las clases medias y altas como tal. Sin embargo, este discurso y la materialización de éste ha sido ejecutada en espacios primordialmente centrales y propensos a la gentrificación o plenamente gentrificados (Cebollada, Sillero, y Anaya, 2018; Danyluk y Ley, 2007; Vith y Mössner, 2017) <sup>1</sup>

#### *1.2.4 Gentrificación como desarrollo desigual del espacio urbano*

La gentrificación ha sido expuesta desde sus vertientes comerciales y económicas y guiadas por el Estado, cada una de estas perspectivas abonan al entendimiento del proceso de gentrificación. Sin embargo, la propuesta estructural de N. Smith (1982) permite entender la gentrificación como un fenómeno estructural que configura el espacio urbano contemporáneo. En este trabajo se retoma la propuesta de Smith para estudiar el fenómeno de la gentrificación desde dos vertientes. La primera, desde una postura crítica de la propuesta, la cual suma las dimensiones de la gentrificación en el siglo XXI, tales como la gentrificación comercial y de consumo así como la guiada por el Estado. Y, al mismo tiempo, esta propuesta será la base para entender la gentrificación como realidad aumentada digitalmente.

Una de la grandes críticas al fenómeno de gentrificación desde su gestación, como concepto, ha sido la necesidad de explicarla teóricamente como fenómeno social urbano. En este marco, Neil Smith, desde una perspectiva marxista estructuralista, propone estudiar la gentrificación como manifestación del desarrollo desigual del capitalismo en el entorno urbano. Para esto, el autor

---

<sup>1</sup>La presencia de bicicletas públicas y el fenómeno de gentrificación se menciona dad que más adelante, este criterio se utilizará para identificar colonias gentrificadas en la Zona Metropolitana de la Ciudad de México

desagrega el proceso en tres grandes dimensiones, tendencia de igualación/diferenciación, valorización/desvalorización del capital y, el desarrollo desigual del capitalismo (N. Smith, 1982).

Primero, el tendencia hacia la igualación y diferenciación se manifiesta en el espacio urbano por medio de la conformación de dos espacios en la ciudad, la ciudad central y los suburbios. De manera sintética, Smith expone que la igualación de los espacios se da mediante fuerzas del desarrollo urbano como el transporte. Éstas, por ejemplo, ayudan a igualar el estilo de vida urbano a lo largo de toda la ciudad. Contrariamente, estas fuerzas ayudan a diferenciar los espacios en función de la renta del suelo que, a su vez se basan en la división funcional y social del espacio urbano. Es decir, la diferenciación en la ciudad se basa en la separación entre suelo habitacional, industrial, comercial, de entretenimiento y administrativo. El primero, en particular, genera su división con base en la segmentación social, por ingreso, raza o estrato social. La segunda tendencia es la valorización/desvalorización del espacio construido. Smith sustenta su análisis en un modelo centro periferia de la ciudad que se constituye históricamente. De esta manera, el autor expone que en una primera etapa la valorización del centro de la ciudad transfiere su valor a los inmuebles. Con el auge histórico de los suburbios, el entorno construido de la ciudad central comienza a deteriorarse y perder valor como elemento dentro del proceso de producción. Finalmente, en una tercera etapa del ciclo, los espacios de la ciudad central se revalorizan. Como tercer elemento, Smith propone que las crisis son un elemento intrínseco al sistema capitalista y que por esta razón las fuerzas de consumo y producción producen la recuperación de espacios de la ciudad central en busca de mayor ganancia en estos espacios (Redfern, 1997; N. Smith, 1982, 1979).

De manera estructural, la gentrificación como desarrollo desigual permite entender la evolución del espacio urbano en los últimos 50 años. Sin embargo, esta propuesta ha sido ampliamente criticada y sometida a perfeccionamiento. Una de las principales críticas ha sido su visión monocéntrica de la ciudad, asumiendo que las dinámicas urbanas se manifiestan exclusivamente entre el centro y los suburbios. Al respecto, han surgido otras propuesta teóricas como la gentrificación rural o bien la dispersión y contagio de la gentrificación más allá de las zonas centrales de la ciudad (Lees y cols., 2016). En este punto, vale la pena aclarar que la diferenciación del espacio basada en el estrato social no se da exclusivamente en la ciudad central. Otra crítica importante es el peso explicativo que Smith le otorga a la valorización y revalorización del suelo basado únicamente en el diferencial de la renta. En esta dirección, las críticas apuntan a que Smith confunde el valor de

uso del espacio con el valor económico. Así, el sustento de la valorización de los espacios no se encuentra únicamente en el diferencial de la renta, sino en la valoración simbólica de éstos. Esta vertiente es la que expone Sharon Zukin con su propuesta de gentrificación comercial . Por último, la desigualdad en el desarrollo del capitalismo como explicación de la gentrificación ha encontrado interlocución en las propuestas de la gentrificación planetaria y el neocolonialismo. En esta crítica se pueden encontrar las distintas vertientes que critican el proceso de gentrificación a nivel global como nuevo mecanismo de las potencias económicas para homogeneizar un patrón de desarrollo urbano a nivel internacional (Díaz-Parra, 2013; Lees, 2012; López-Morales, Shin, y Lees, 2016).

En este punto, se ha hablado de diferentes elementos de la gentrificación que ayudarán a entender la propuesta teórica de este trabajo de investigación, una gentrificación digitalmente aumentada. Al respecto, Wyly (2018) apunta a entender el fenómeno de la gentrificación como algo estructural que se desarrolla en un proceso de urbanización planetaria, transeccional y dentro de una etapa del capitalismo cognitivo. En este sentido, el autor señala que “la innovación informacional reproduce drásticamente la dinámica, y las formas cibernéticas del capitalismo cognitivo, la gentrificación llega a ser una manifestación del desarrollo cognitivo desigual, de la acumulación de capital mediada por un universo de código autorreplicado que altera la dinámica coevolutiva de lo que Harvey ha teorizado como la urbanización de la conciencia”(Wyly, 2018). En otras palabras, para el autor el proceso de gentrificación puede ser interpretado como la representación espacial en el espacio urbano de las diferencias creadas por el desarrollo desigual del conocimiento a nivel internacional. En su ejemplo, Wyly expone de qué manera se generan rentas exorbitantes en lugares cercanos a empresas de desarrollo tecnológico. Para el autor, la omnipresencia de las tecnologías son un elemento fundamental para entender la evolución de la gentrificación y de la urbanización en el siglo XXI. Aunque Wyly vislumbra la importancia de la tecnología y del código en el proceso urbano, solo se está enfocando en las disparidades que se generan entre aquellas personas que generan dicho código y las que no. Sin embargo, en la actualidad el código es omnipresente en todo lo que se realiza por medio de una computadora, ya sea desde una simple llamada telefónica o una publicación en sitios de la red sociodigital de preferencia. De esta manera, todo usuario de código es susceptible de ser parte de dicho proceso de urbanización guiada por la dimensión digital.

Así, el presente trabajo apunta a profundizar el estudio de la dimensión digital de la gentrificación. Como se ha desarrollado, la gentrificación es un fenómeno que configura el espacio

urbano de manera desigual basado en fuerzas estructurales y particulares. En este sentido, se retomaron los avances realizados hasta el momento, pero se considera la estructuración desigual del espacio urbano basado en la gentrificación y la co-producción de ésta con los elementos digitales contemporáneos, particularmente mediante los sitios de redes sociales en Internet. Así, en el ámbito metodológico, la gentrificación aumentada digitalmente no se centra en el desplazamiento o reconfiguración social de las zonas gentrificadas o bien en el análisis del precio de la vivienda o especulación inmobiliaria. Por el contrario, la dimensión digital de la gentrificación puede ser vista, de manera metodológica, por medio del análisis de los Grandes Datos, los cuales pueden permitir observar los elementos culturales, simbólicos y comerciales asociados a la gentrificación y, cómo estos se magnifican en el terreno de lo digital. Por esta razón, a continuación se presenta una discusión rápida de los elementos sustantivos de los **Grandes Datos** para ciencias sociales y, posteriormente se desarrollan con más profundidad los enfoques teóricos que desde el campo de la geografía digital pueden ayudar a entender el proceso de gentrificación digital.

### **1.3 Grandes Datos y las Ciencias Sociales. Hacia una ciencia social computacional**

La tendencia generalizada a la digitalización de la mayoría de las tareas de la vida cotidiana ha generado la disponibilidad de cantidades enormes de datos y por ende de información. A este cambio se le ha denominado la Revolución de los Grandes Datos (Kitchin, 2014b). Dicha revolución incluye todas las tareas automatizadas que se realizan gracias a la interconectividad computacional. Es decir, todas las actividades que van desde las publicaciones diarias de las personas en los sitios de redes sociodigitales, las transacciones bancarias internacionales hasta las imágenes satelitales, datos meteorológicos, edificios inteligentes hasta los datos generados por el CERN. La revolución de los grandes datos ha significado retos y desafíos para múltiples sectores de la sociedad. Por ejemplo, para los analistas de datos y computólogos que se han enfrentado al manejo, curaduría y preservación de los datos. Para los estadísticos y científicos sociales ha abierto un debate epistemológico entre las ventajas y desventajas de los grandes datos contra los ‘pequeños datos’. También para filósofos y sociólogos de la ciencia quienes se han visto ante críticas de un posible post-positivismo guiado por los datos, o bien, a cuestionar los posibles riesgos éticos ante una posible hipervigilancia digital.

En este entorno, las ciencias sociales han comenzado a mirar las potencialidades de los gran-

des. Por un lado para aproximar empíricamente fenómenos que tienen una larga tradición de estudio, por ejemplo, movilidad poblacional, mitigación de riesgos socioambientales, identificación de proliferación enfermedades, entre otras. Por otro lado, se han comenzado a formular nuevas interrogantes y la aproximación de nuevos fenómenos sociales ante el potencial transformador de esta revolución de la información, tal ha sido el caso de las criptomonedas, la realidad virtual o el Internet de las cosas.

Ante ese escenario, en esta sección se exponen algunas aproximaciones teóricas para entender el concepto de Big Data y sus posibles aplicaciones a las ciencias sociales, de manera específica en el campo de los estudios urbanos. Para el caso de este trabajo de tesis es importante conceptualizar el paradigma de los grandes datos desde un postura crítica, ya que este tipo de información no solo es reflejo de inocuo de los fenómenos sociales, sino que al estudiar la dimensión digital de los fenómenos sociales, la información también se convierte en vector de intervención para la conformación de la transformación que digital. Además, considerar una postura crítica sobre los Grandes Datos y la gentrificación permite evitar caer en análisis postpositivistas que no consideran la construcción social de la información.

### *1.3.1 Aproximaciones críticas a los Grandes Datos*

El *Big Data* o Grandes Datos se puede encontrar en la actualidad como sinónimo de avance tecnológico. En la vida cotidiana se puede encontrar las referencias a los Grandes Datos en televisión, periódicos, revistas de divulgación entre otros. En este campo, muchas de las referencias hablan de las bondades y potencialidades de este paradigma. Dentro del campo de las ciencias sociales esta euforia también ha permeado, tan solo en los últimos años los artículos publicados en revistas especializados se ha incrementado exponencialmente. Sin embargo, al igual que los Grandes Datos ofrecen bondades también presentan limitaciones y riesgos para el análisis. Así, un acercamiento crítico a este campo naciente de estudio se esquematiza en dos vertientes. La primera es la conceptualización técnica y informática de la naturaleza, construcción y aplicación de este tipo de información. La segunda aproximación, dotar a los Grandes Datos de un paradigma crítico de análisis, donde éstos no sean entes inocuos para estudios sociales, sino considerarlos actantes en la construcción de nuevas realidades sociales.

El origen del concepto de Grandes Datos es todavía discutido. Algunos consideran que John

Mashey, investigador en jefe de *Silicon Graphics* fue quien acuñó el término a mediados de la década de 1990 (Diebold, 2012). Mientras otros consideran difuso el origen del término (Holmes, 2017b). Sin embargo, mucha de la comunidad especialista en informática concuerda en que la primera característica a resaltar es el volumen de la información. A raíz de esto una de las aproximaciones más difundida sobre los Grandes Datos enfatiza las características, retos y ventajas de las grandes bases de datos.

Se consideran grandes bases de datos si éstas cumplen con, al menos, 4 características, resumidas en 3 Vs<sup>2</sup>. Estas Vs significan volumen, velocidad, variedad y valor. Como se ha mencionado, el volumen de las bases de datos es el primer rasgo que llamó la atención de los científicos de datos. Esto debido a que computadores promedio ya no podían almacenar ni procesar la enorme cantidad de datos que se generan día con día. Aunque cotidianamente se puede considerar a una base de datos como grandes datos aquella que no es manejable en computadoras comerciales, no existe un tamaño de base estandarizado para catalogar como Grandes Datos, dichas bases pueden ser del orden de Petabytes o Zettabytes (Holmes, 2017b). El segundo rasgo de los Grandes Datos es la velocidad tanto de generación como de procesamiento de la información. En este punto se relaciona con la naturaleza misma de los datos contemporáneos. Es decir, que las grandes bases de datos son consecuencia de la explosión y avance de Internet, desde la transformación de una Web 2.0 hasta el desarrollo contemporáneo del gran ancho de banda y de la interconectividad de dispositivos que desencadenará en el llamada Internet de las Cosas. El Internet permite generar información en tiempo real de manera prácticamente inmediata. Aunado a la rapidez de generación ésta conlleva a la velocidad en el procesamiento de la gran cantidad de información. En este punto vale la pena resaltar que ante el volumen y rapidez en la producción de datos crece en paralelo la necesidad de nuevos métodos para procesar y analizar de manera eficiente y eficaz con el objetivo de analizar la información. Ante este escenario, algoritmos de Aprendizaje de Máquina o análisis NoSQL han cobrado relevancia.

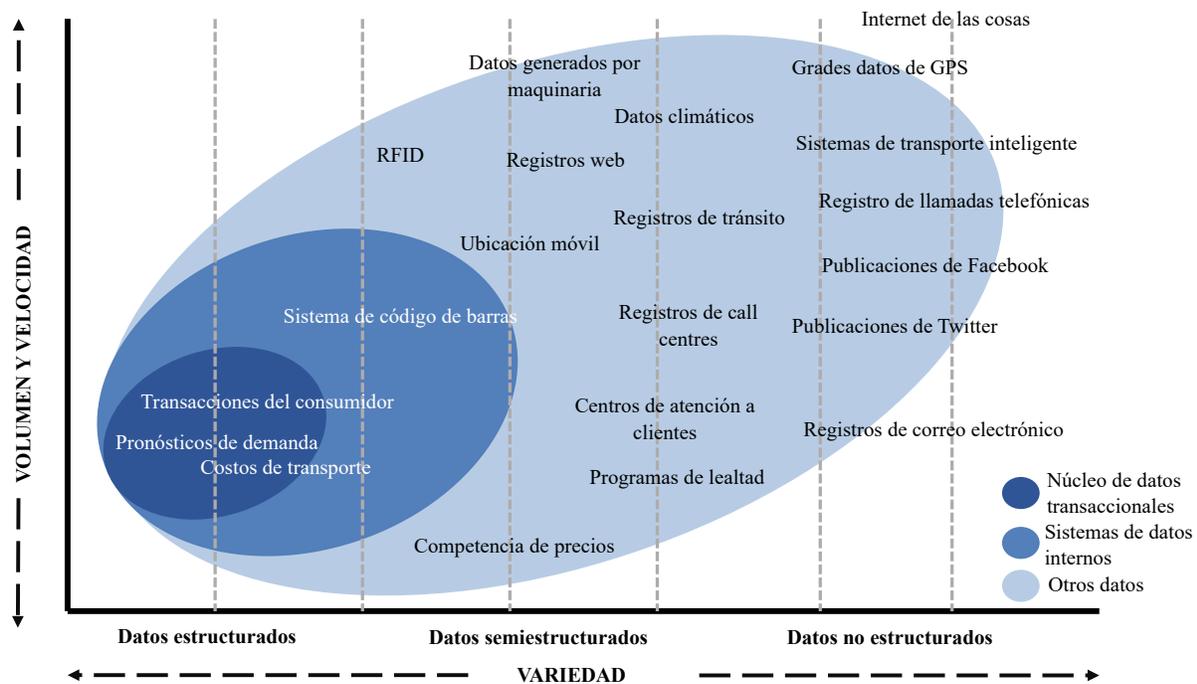
El tercer elemento es la variedad en el tipo de información. La característica de variedad enfatiza la diversidad de formatos y estructuras de las bases de datos. Previo a la explosión de la digitalización de las actividades y al crecimiento de Internet las bases de datos se construían de manera relacional. Es decir, se contaba con el registro de cada observación, este podía ser desde

---

<sup>2</sup>Aunque algunos proponen hasta 7 dimensiones del paradigma, estos se pueden sintetizar en 3.

una persona hasta un país y, a éste se le asociaba algunos de los rasgos de interés. Ante el auge de Internet los formatos de datos se han diversificado y ahora se cuentan con bases de no estructuradas, un ejemplo de ello son el audio y video. La variedad es una característica menor, ya que mucho del contenido cotidiano se genera por medio de formatos no alfanuméricos sino audiovisuales. La figura siguiente muestra la relación que guardan las distintas características de los Grandes Datos.

Figura 1.1: Relación de las 3 Vs del paradigma Big Data



Fuente: Jaggi y Kadam (2016, p:1016)

En la figura 1.1 se relaciona la variedad en los tipos de formato y de información en relación al volumen y a la velocidad con la que se generan cada tipo de datos. En la primera parte se encuentran los datos estructurados, esto quiere decir que este tipo de información se almacena en bases de datos relacionales o matriciales. Este tipo de datos tiene relacionado un volumen de información relativamente bajo y la velocidad con la que se genera es proporcionalmente lenta. La gráfica ofrece ejemplos de pronóstico de demanda, encuestas origen destino, modelos de base de datos de transacciones financieras. En este apartado se pueden incluir las encuestas realizadas por organismos de estadística nacional, tales como los censos poblacionales o encuestas probabilísticamente construidas. En el lado opuesto de la gráfica se encuentra el paradigma de Big Data. Por un lado, se tienen bases de datos no estructuradas, es decir información que no se encuentra en

tablas relacionales. Este tipo de información se caracteriza por generarse en grandes volúmenes y de manera casi instantánea, o en tiempo real. Ejemplos de estos datos son los generados en blogs, correos electrónicos, llamadas telefónicas, publicaciones de Facebook o Twitter, o datos de geolocalización. Además, se puede agregar los datos visuales que provienen de plataformas sustentadas en imágenes, tales como Instagram. La gráfica permite ver una relación prácticamente directa entre el volumen creciente de información y la velocidad con la que se genera y su cualidad de no ser estructurada. En otras palabras, la mayor cantidad de información será generada en bases de datos no estructuradas y en tiempo real.

La revolución de los Grandes Datos no solo implica un cambio en el términos computacionales, también implica retos en la manera de realizar investigación en ciencias sociales (Kitchin, 2014b; Lewis, 2015; Salganik, 2018). Por un lado, representa ventajas, desventajas y oportunidades que ofrecen los Grandes Datos como fuente de información y análisis con respecto las Bases de pequeños datos. Por otro lado, los grandes datos implican no solo nuevas fuentes de información y la aplicación de nuevas técnicas de análisis, éstos representan el reflejo del cambio hacia sociedad digital, donde el papel de los algoritmos está lejos de ser inocuo.

Los grandes datos implican retos para el análisis social por su carácter de fuentes de información novedosas y Además de las 3Vs mencionadas anteriormente, los grandes datos disponen de alta granularidad de observación (Kitchin y McArdle, 2016). Por ejemplo, la información financiera puede ser rastreada debido a que cada banco tiene identificado a cada usuario por medio de un número único de cuenta, cada publicación de cualquier sitio de red sociodigital está asignado a un nombre de usuario y cada dispositivo conectado a Internet se puede rastrear por medio su dirección MAC. En otras palabras, los registros que se tienen de internet pueden ser asignado a un individuo único. Esta particularidad ha hecho que los grandes datos sean fuentes alternativas para aproximar empíricamente muchos de los fenómenos que se analizaban por medio de encuestas o censos (Kitchin y McArdle, 2016; Salganik, 2018).

Los grandes datos pueden ser complementarios a las bases de pequeños datos, encuestas o censos tradicionales (Crampton y cols., 2013; Holmes, 2017b; Salganik, 2018). Si bien en cierto que los grandes datos son producidos en tiempo real, en comparación con las encuestas, los primeros son generados con motivos comerciales mas que de investigación de algún fenómeno en particular (Salganik, 2018). Esto, generalmente, provoca un problema de sesgo en los grandes da-

tos. Lo anterior implica la información de los grandes datos no es suficiente para realizar inferencias por sobre el total de la población, debido a las barreras de acceso a la dimensión digital <sup>3</sup>. Ante lo anterior, la propuesta metodológica es realizar análisis sociales que conjunten ambos tipos de datos.

El segundo punto a resaltar sobre los grandes datos es su carácter de fuente de información ante la digitalización del mundo y, al mismo tiempo un vector de digitalización de la sociedad. Con respecto al primer punto, se ha hablado que los grandes datos son novedosas bases de datos que son resultado de la revolución digital. En contraparte, se ha propuesto que los grandes datos no solo son reflejo inocuo de dicha transformación, sino que éstos se construyen con intereses específicos, políticos, económicos o sociales. En este sentido, los grandes datos también son actantes en la transformación digital de la sociedad (Crampton y cols., 2013).

Los algoritmos que subyacen a todas las plataformas digitales son políticamente contruídos (Lewis, 2015; Shaw y Graham, 2017). Dichos algoritmos son pasos consecutivos de acciones que una computadora puede interpretar y ejecutar. Este nivel de la revolución digital es poco estudiado ya que, en la generalidad, se considera como el algoritmo detrás de las plataformas digitales son elementos neutrales que son guiados por principios apolíticos de la la lógica y la programación computacional. Sin embargo, autores como Crampton y cols. (2013); Elwood y Leszczynski (2018); Lewis (2015); Shaw y Graham (2017) han argumentado que los algoritmos tienen un fuerte componente político. Un ejemplo de esto es el algoritmo de ordenamiento de páginas que utiliza Google para mostrar resultados en su página, ya que primeros resultados pueden ser pagados por empresas específicas. De igual manera, con los avances de Inteligencia Artificial, los algoritmos se modifican constantemente ante el comportamiento de los los usuarios de plataformas digitales. En este sentido, los grandes datos que se obtienen de Internet no solo son reflejo de la revolución digital sino también productores.

#### **1.4 Los grandes datos en estudios urbanos**

La ciudad contemporánea se constituye como un complejo ensamblaje de elementos de *hardware* y *software*. Esto va desde millones de kilómetros de fibra óptica, cámaras de circuito cerrado, antenas receptoras y emisoras hasta las computadoras o teléfonos inteligentes en los hogares. De manera paralela, se constituye de elementos de *software* tales como todos los programas para identificar

---

<sup>3</sup>este punto se desarrolla más adelante

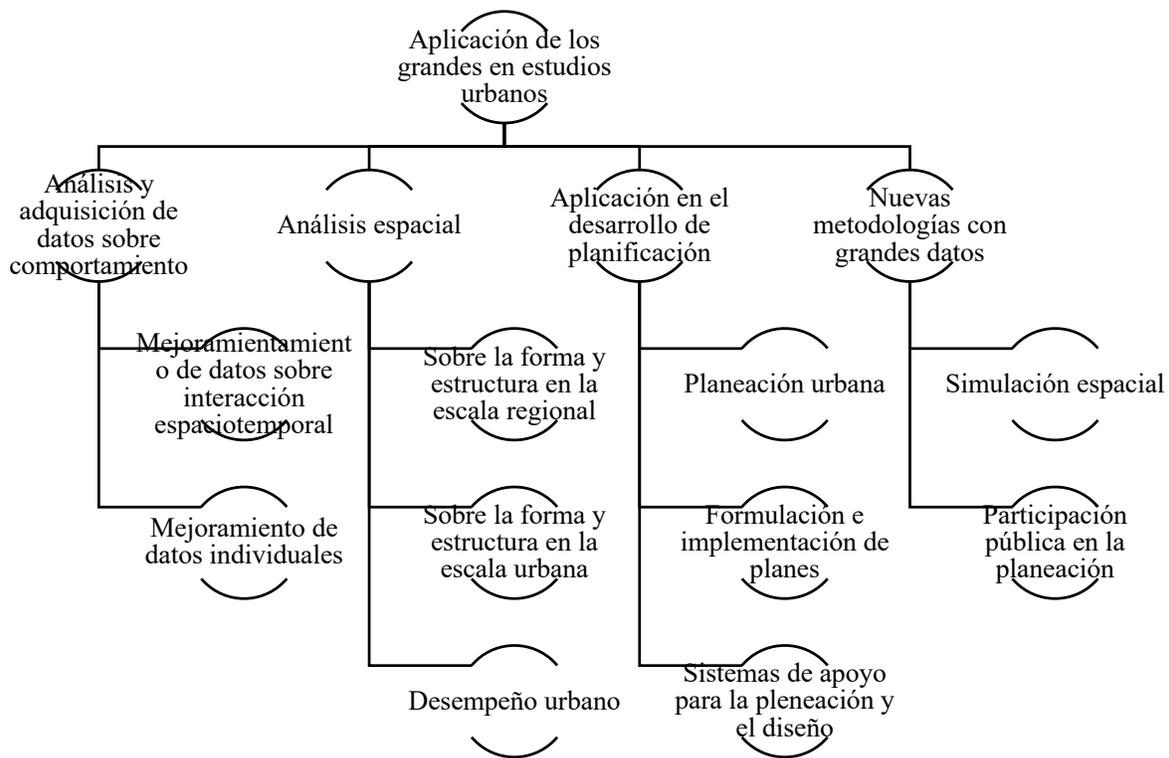
rostros o matrículas de automóviles en las cámaras, el programa que permite pagar electrónicamente el transporte público o la aplicación móvil que proporciona, en tiempo real, el clima en la ciudad. Otro elemento que se ha visto trastocado es la sociabilidad de las personas en entornos urbanos, para esto hay numerosos ejemplos. Uno de ellos es la transformación de las relaciones sociales mediadas por los sitios de redes sociodigitales, tales como Facebook, Twitter, Instagram, etc. Otro más es la búsqueda de hospedaje o vivienda por medio de Airbnb o el cambio en la movilidad cotidiana basada en aplicaciones como Uber, Cabify o Blablacar. La búsqueda de lugares para consumir o esparcirse en plataformas de reseñas como Google Maps o Yelp. En suma, el modo de vida urbano está cambiando y con ello, también la manera de estudiar estos fenómenos desde una perspectiva de lo digital (Kitchin y Dodge, 2011; Kitchin y Perng, 2016).

En este complejo entramado entre la constitución material y digital de la ciudad, así como de la constitución entre el espacio construido y social los grandes datos implican oportunidades y retos en la investigación de fenómenos urbanos. La emergencia de lo digital en lo urbano ha provocado dos maneras de estudiar la ciudad desde los grandes datos y la digitalización de la sociedad. Una de ellas ha sido la aplicación de nuevos y sofisticados modelos matemáticos que contemplan la complejidad de la ciudad y su representación por medio de los grandes datos. A decir de Cheng y cols. (2016) los grandes datos no solo permiten analizar las dimensiones temporales y espaciales de la ciudad, también posibilitan ver las interacciones sociales y emocionales. La otra manera de analizar la revolución de los grandes datos ha sido por medio de los efectos que plataformas digitales han tenido en los fenómenos urbanos preexistentes. En este punto, los casos paradigmáticos han sido los cambios la economía urbana debido a las plataformas de economía compartida, tales como Uber o Airbnb.

Gran parte de los grandes datos provenientes de dispositivos con acceso a Internet y de las aplicaciones digitales de la ciudad se encuentran dotados con GPS. Esta característica implica que la naturaleza de los grandes datos sea, en gran medida, de índole geográfico y propicia para el estudio de fenómenos urbanos. En este sentido, varios autores concuerdan en que los grandes datos en estudios urbanos surgen del desarrollo tecnológico del GPS, de la recolección de datos en tiempo real y, también apuntan a las ventajas asociadas a estas características. Así, los estudios que han reflexionado sobre la transformación de los estudios urbanos a raíz de los grandes datos se dirigen a tres grandes cambios en términos epistemológicos. El primero de ellos es que, con los grandes datos

se mejora la granularidad de la información, es decir se dispone de datos más precisos en términos espaciales y temporales. Por ejemplo, con imágenes satelitales se puede conocer la expansión física y transformación en el uso de suelo de las urbes en áreas de  $100m^2$ . Resultado de este cambio se apuntan a realizar mejores aproximaciones en el terreno espacial, ya que ahora es posible identificar cambio en la estructura urbana con alto detalle geográfico. El tercer cambio es la posibilidad de realizar estudios en tiempo real, ya que muchos de los datos generados a través de Internet tienen esta característica. Ejemplos de esto es el estudio de la movilidad cotidiana de las personas (Cheng y cols., 2016; Hao y cols., 2015; Long y Liu, 2016; Thakuria y cols., 2017).

Figura 1.2: Aplicaciones de los grandes datos en estudios urbanos



Fuente: Hao y cols. (2015, p:96)

La figura 1.2 propuesta por Hao y cols. (2015) es un punto de partida en el debate sobre el papel transformador de los grandes datos en estudios urbanos. Los autores identifican cuatro ejes de la aplicación de los grandes datos en estudios urbanos, análisis y adquisición de datos sobre comportamiento de las personas, análisis espacial, aplicaciones a la planificación urbana y nuevas metodologías con el uso de grandes datos. Cada una de estas dimensiones se conforma de un con-

junto de bases de datos específicas y de usos particulares para el campo de estudio. En el primer rubro los autores identifican los datos provenientes del sistema GPS, publicaciones de sitios de redes sociodigitales y de tarjetas de pago digital de transporte público. Dentro de esto, los autores identifican usos potenciales en delimitación metropolitana, viajes cotidianos, estructuración social del espacio construido, identificación de redes sociales, comportamiento de movilidad no laboral de las personas, entre otras. En segundo lugar, el análisis espacial, los autores identifican posibles usos de análisis de redes para estudios de movilidad, estructura y desempeño económico de las ciudades. En tercer lugar, los grandes pueden ser aplicados para la planificación urbana dado que por medio de grandes datos es posible identificar y pronosticar ritmo y dirección del crecimiento urbano. Finalmente, los grandes datos permiten la implementación de nuevas metodologías cuantitativas en los estudios urbanos, ejemplos de ello son la modelación de autómatas, modelos basados en agentes, modelos bayesianos y aprendizaje de máquina (Hao y cols., 2015).

Las posibilidades de mejorar y expandir las metodologías de estudio en el ámbito urbano son crecientes gracias a los grandes datos, como se ha argumentado, éstos surgen de una revolución digital que tiene como base la revolución tecnológica y el desarrollo de algoritmos. Al mismo tiempo, los elementos subyacentes a dicha revolución no son inocuos, ni apolíticos. Es así, que los grandes datos y a digitalización de la sociedad también le implican a los estudios urbanos el cambio constante del objeto de estudio, la ciudad.

Cada uno de las facetas de la digitalización de la vida urbana tiene características, motores y consecuencias particulares. En este sentido, los grandes datos son solo una vertiente de un marco general del proceso de digitalización de la sociedad urbana. Así, cada plataforma o herramienta digital tienen causas y consecuencias específicas en el cambio de las ciudades. A continuación se dan algunos ejemplos paradigmáticos en este cambio digital, particularmente la economía colaborativa y plataformas como Airbnb o Uber, además de los sitios de redes sociodigitales como Facebook o Twitter.

La economía ha sido uno de las dimensiones que más ha sido trastocado por el desarrollo digital, específicamente por la ‘economía colaborativa’ o *sharing economy*. “El término de economía colaborativa ha sido frecuentemente usado para describir a diferentes organizaciones que conectan a usuarios/arrendadores y propietarios/proveedores a través de plataformas consumidor a consumidor, (*consumer-to-consumer*) como es el caso de Airbnb o Uber, o de empresa a consumidor

(*business-to-consumer*), permitiendo que las rentas sean más flexibles, en términos de interacción social” (Parente, Geleilate, y Rong, 2018, p: 53). En particular, este tipo de plataformas en Internet se basan en el principio de que las personas poseen activos subutilizados, como el automóvil o casa, están dispuestas a ‘compartir’ y que los desarrolladores de aplicaciones utilizan este principio para realizar negocios (Belk, 2014; Parente y cols., 2018). Solo por mencionar un par de compañías que se han desarrollado en este rubro, se encuentran Airbnb y Uber, que actualmente son las dos empresas con la mayor rentabilidad de Estados Unidos dentro del rubro de compañías tecnológicas (Bloomberg, 2017). Airbnb ha crecido de manera notable desde su creación en 2008, esta empresa tiene presencia en 34 000 ciudades en 191 países (Lutz y Newlands, 2018). Dicha empresa reportó ganancias por 100 millones de dólares en 2017 (Times, 2017). Al mismo tiempo, se puede hablar de Uber. Desde 2009 cuando la empresa fue creada en San Francisco California y hasta la fecha Uber tiene presencia en 741 ciudades en todo el mundo (UBER, 2018). En 2016, las ventas de esta empresa ascendieron a 7.5 miles de millones de dólares. En pocas palabras, la economía colaborativa ha generado nuevas formas de generar ganancias y crear nuevos gigantes económicos.

Una de las dimensiones que notoriamente ha cambiado Internet, en específico las redes socio-digitales, es el político. Es sabido que por medio de este tipo de sitios se comparte contenido, ideas o material audiovisual. Al mismo tiempo, las plataformas digitales de redes sociales dan la ventaja de que la información que circula en ellas es prácticamente instantánea, aunque no necesariamente veraz. A pesar de estos inconvenientes, sitios como Facebook o Twitter se han convertido en elementos fundamentales en la sociedad actual en la arena política. A tal grado que, tanto académicos como políticos reconocen su importancia.

En general, Internet ha sido interpretado por los especialistas como una herramienta que puede ser un espacio extendido de la esfera pública (Trejo, 2015). De esta manera, por medio de la red, se pueden poner en la discusión pública temas de relevancia para la ciudadanía, más que la agenda preestablecida de los gobiernos. Al mismo tiempo, los sitios de redes sociodigitales y audiovisuales, como Facebook, Twitter o Youtube, son vistos por los políticos como medios para mantener un contacto directo con la población. Aunado a la perspectiva de ciudadanía extendida y de la utilización de Internet por personajes políticos, también se ha intentado promover la digitalización de la interacción entre el gobierno y la población por medio de plataformas que agilicen los trámites gubernamentales. En general, la era digital ha sido bien recibida en los ámbitos políticos, aunque esta

visión positiva es contrastada por prácticas antidemocráticas, de las cuales se hablará más adelante.

Se han documentado muy bien la importancia que tuvieron las redes sociodigitales como Twitter o Facebook para organizar manifestaciones en los primeros años de la década. Ejemplos de esto son la primavera árabe, el movimiento #Occupy Wall Street en Estados Unidos, #yosoy132 en México (AlSayyad y Guvenc, 2015; Castells, 2012; Gleason, 2013; Penney y Dadas, 2014). En América Latina los casos son menos documentados, sin embargo, existen movilizaciones que, en mayor o menor medida, han sido posibles por las redes sociodigitales. En Colombia, Twitter también fue arena de discusión sobre el acuerdo de paz entre el gobierno de ese país y las FARC (Marín-Cortés y Quintero-Jurado, 2018), en Argentina movimientos sociales en protesta por los feminicidios acontecidos en 2015, (#niunamenos) (Laudano, 2017), en Venezuela en 2014 (Said-Hung y Valencia-Cobos, 2017). En general, las redes sociodigitales son vistas como un espacio donde la esfera pública puede realizarse, se conceptualiza como una posibilidad de ciudadanía y participación política (González-Bailón, 2014).

En contrapartida, las diferentes herramientas de la revolución digital presentan consecuencias negativas, en el terreno social. Cada uno de los puntos expuestos anteriormente presentan serios cuestionamientos sobre la igualdad social, política o incluso ética en la era digital. Es importante mencionar que estos impactos negativos de las plataformas de Internet no son, en sí mismos, aleatorios ni se originan de manera espontánea. Las consecuencias negativas se vinculan a los usos políticos y económicos que Internet representa para conservar ganancias y cotos de poder.

En la problematización de las transformaciones digitales en la sociedad se tiene a la economía colaborativa. Si bien se ha mencionado que este tipo de economía aprovecha el principio de compartir activos excedentes también ha provocado que se generen segmentación de mercados y alza de los precios de la vivienda en diversas ciudades. Para Lutz y Newlands (2018), la economía colaborativa de Airbnb en la práctica reproduce elementos de discriminación, por edad, sexo y actividad económica. En su estudio, los autores encuentran que los arrendadores de Airbnb prefieren la población mayor para evitar tener a adolescentes y niños como huéspedes. Al mismo tiempo, los arrendatarios dan prioridad a personas con alta escolaridad y hacen explícitas sus preferencias de tener arrendadores mujeres u hombres. En consecuencia, estas limitaciones provocan la segmentación de mercado, que en primera instancia tiene que pasar por el acceso a Internet <sup>4</sup> y después en

---

<sup>4</sup> Antes de poder participar de la sociedad digital es necesario tener acceso a Internet, lo cual es una primera barrera

cumplir todos los requerimientos de los arrendadores. Al mismo tiempo, Airbnb ha provocado el aumento en el precio de la vivienda en diversas ciudades en el mundo. Esto ha llevado a que las habitaciones sean destinadas exclusivamente para renta y no para los habitantes cotidianos de la ciudad. Lee (2016) ha observado que el incremento de Airbnb en la ciudad de Los Angeles está asociado con procesos de gentrificación y reforzamiento de segregación racial en dicha ciudad. Esto debido a que los propietarios de departamentos, en busca de mayores ganancias, han dejado los departamentos solo para renta temporal en la plataforma de Airbnb. A decir del autor, Airbnb también ha provocado el reforzamiento de la segregación racial, debido a que los consumidores de la plataforma son, en su mayoría, blancos que prefieren hospedarse en vecindarios de la misma clase. Cabe mencionar que Airbnb por sí solo no implica un alza en los precios de la vivienda y mucho menos que se provoque gentrificación. Existen otros factores asociados con el fenómeno de economía colaborativa de Airbnb. Por ejemplo, la estrategia de propietarios de mantener departamentos en zonas con alta demanda de renta temporal, además de la propia demanda de departamentos propios para personas de altos ingresos y finalmente, estrategias que incentivan el turismo en zonas históricas y de reciente renovación (Füller y Michel, 2014; Ioannides, Röslmaier, y van der Zee, 2018).

Otra de las consecuencias del uso de Internet es la violación de la privacidad de las personas por medio de redes sociodigitales, ya sea por motivos de espionaje o políticos. En este sentido, el uso de la tecnología no es una herramienta nueva para el control de las población. En los últimos años del siglo pasado y los primeros años del presente ya se identificaba a Internet como una herramienta para el espionaje. Internet se visualizaba como un peligro constante para secretos de estado, acciones de inteligencia y tácticas militares. Tan solo en 2001 la Agencia Central de Inteligencia de Estados Unidos reportaba 350 000 intentos de hackeo anuales (Wettering, 2001). A decir de Wettering (2001), el objetivo principal del espionaje se enfocaba en los ciudadanos o empresas civiles. Sin embargo, a casi dos décadas de esa primera advertencia, el panorama no ha cambiado. Se ha documentado que Barack Obama utilizó la plataforma de Twitter como punto decisivo durante su campaña en 2008 (Mancera y Pano, 2013). En marzo de 2018 el periódico británico *The Guardian*, reveló que Facebook junto con Cambridge Analytica<sup>5</sup> utilizaron información de los perfiles

---

en países no desarrollados. Esta barrera se discutirá con más profundidad a lo largo de todo el trabajo de investigación debido a que es una de las dimensiones que explica la desigualdad digital

<sup>5</sup>Una empresa dedicada a la minería de datos que se vincula directamente en campañas políticas brindando asesoría.

de usuarios de Facebook sin autorización desde principios de 2014 para construir un sistema capaz de crear perfiles de votantes en Estados Unidos. Esto con el objetivo de hacer blanco a los usuarios de dicha red sociodigital de propaganda política personalizada, y con ello influir en la elección de Donald Trump en la elección presidencial de 2017 en Estados Unidos. Cambridge Analytica recolectó más de 50 millones de perfiles y ha posibilitado la negociación con varios procesos electorales en el mundo y la venta de dicha información a compañías con motivos de publicidad (Guardian, 2018). Este hecho ha desatado cuestionamientos fuertes acerca de vulnerabilidad de la población en Internet y sobre todo en sitios de redes sociodigitales.

A la luz de estos hechos, se cuestiona el papel democratizador de Internet de sitios como Facebook o Twitter, que lejos de permitir la discusión y libre pensamiento, el diálogo y la creación de esfera pública, la está destruyendo. Esto se debe a que las grandes empresas de tecnología no se preocupan, *per se* por los usuarios, sino por la información que estos dan voluntariamente y que es aprovechada para elevar las ganancias y ganar elecciones. Al respecto, Habermas (2018), menciona que todavía no se tiene el manejo suficiente sobre Internet y sus amenazas. Por ejemplo, una vez que se inventó la imprenta todas las personas se convirtieron en lectores potenciales. Ahora, Internet ha convertido a todos en usuarios potenciales de las redes, sin embargo, falta que todos lo sean de manera adecuada.

#### *1.4.1 El reto de los grandes datos en geografía*

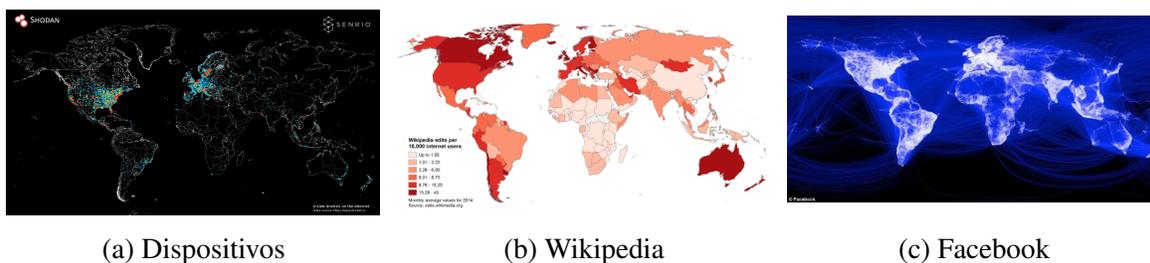
Los datos generados diariamente en Internet dejan rastros de lo que somos en la vida fuera de la red. Y, a pesar de que dicha información es generada para fines comerciales, lo cierto es que dan cuenta del comportamiento humano y además se puede disponer de ella. De tal manera que, ante la explosión de información y ante el paradigma de los grandes datos, las ciencias sociales han comenzado a mirar hacia la minería de datos. En este marco, la geografía, como ciencia social, desde principios de la década de 2010 se comenzó a cuestionar sobre los cambios epistemológicos y metodológicos que podría provocar el paradigma los grandes datos en esta ciencia. Ya que, se parte de una premisa fundamental “la información siempre tiene una geografía” (Poorthuis, Zook, Shelton, Graham, y Stephens, 2014). En este sentido, cualquier publicación realizada en Facebook o Twitter representa, no solo la realidad de una persona, sino su entorno social y geográfico, sus intereses y afinidades definidas por su realidad espacial que le rodea. Al mismo tiempo, la in-

intervención del algoritmo provoca que lo observado en el dominio de lo digital sea un reflejo del comportamiento de las personas y del cambio social provocado por la revolución digital. Gracias a la integración de Sistemas de Posicionamiento Global (GPS, por su siglas en inglés) en dispositivos móviles con acceso a Internet, muchas de las acciones realizadas a través de ellos pueden ser traducidas a coordenadas geográficas (Goodchild y cols., 2017; Sui, Goodchild, y Elwood, 2013). Este avance tecnológico también ha implicado un campo social y un cambio en términos epistémicos, a tal grado de plantear la generación de geografías digitales. Uno de los primeros acercamientos a esto campo ha sido por medio del concepto de Geoweb, el cual ha nacido de la intersección de de los grandes datos georreferenciados y la geografía crítica sobre datos. La Geoweb se basa tanto en herramientas como Google Maps o Google Earth como la información geolocalizada que voluntariamente se otorga (Elwood, 2010; Goodchild, 2007; Shelton, 2015).

Todo lo que se hace en Internet puede ser rastreado geográficamente. Es por ello que Elwood (2010) apunta que la Geoweb puede ser una herramienta política, debido a que, ahora más que nunca, es posible conocer con precisión la localización de cualquier persona que tenga un dispositivo conectado a Internet, ya sea computadora o teléfono inteligente. Ejemplos de visualización del comportamiento en Internet hay bastantes y todos reflejan un patrón de concentración geográfica en países desarrollados. Por ejemplo, en la figura 1.3 se muestran tres mapas mundiales que reflejan distintas dimensiones de Internet. En la figura 1.3a muestra el número de dispositivos conectados a Internet, esto incluye, no solo computadoras, también dispositivos móviles, así como televisores, consolas de videojuegos, refrigeradores o lavadoras inteligentes o incluso sistemas de seguridad. En la figura 1.3b se muestra el número promedio de ediciones en Wikipedia por cada 10 000 usuarios. Mientras que en la figura 1.3c se aprecia el número de usuarios de Facebook y sus interacciones. Estos mapas refuerzan la idea de la constitución de geografías de digitales. En todos los ejemplos anteriores se manifiesta una clara concentración geográfica del contenido digital en países del norte global, esto refuerza la idea de que la dimensión digital también es atravesada por elementos políticos, económicos y sociales.

A pesar de que la georreferenciación de las actividades en Internet proporcionan un primer acercamiento al fenómeno, la sola representación no aporta al entendimiento de los fenómenos representados. Ya que estos fenómenos son atravesados por causas sociales, económicas y políticas que dan sentido y sustento a lo que se visualiza. En este sentido, entender los grandes datos desde

Figura 1.3: Geografía de Internet



Fuente a: <https://dlink-report.shodan.io/>  
Fuente b: <http://www.markgraham.space/blog/the-geography-of-wikipedia-edits>  
Fuente c: <https://brilliantmaps.com/facebook-map/>

una perspectiva crítica y analítica permite abundar en los determinantes sociales y tecnológicos que implican la conformación de las diversas geografías digitales.

Con los elementos anteriores, se puede proponer que los grandes datos no solo son información exenta de valores políticos o económicos. Por el contrario, el contenido digital se convierte en una expresión de la realidad mediada por la dimensión digital y a su vez, es componente constitutivo de ésta. Al respecto, Crampton y cols. (2013) problematizan el uso de datos georreferenciados de Internet, en específico de Twitter. En su trabajo cuestionan el término de grandes datos para el análisis social. La crítica y propuesta de análisis se divide en dos grandes apartados, la primera refiere al abordaje teórico de los datos georreferenciados en Twitter, y el segundo concierne al propio término de grandes datos. En términos teóricos, los autores proponen abordar la geolocalización de Twitter desde cinco aristas para hacer análisis más allá de las coordenadas. En primera instancia, el estudio de las prácticas de la Geoweb debería de ir más allá de la simple visualización del contenido usando latitud y longitud. Segundo, la investigación de la Geoweb debe promover una perspectiva más allá del 'aquí y ahora', un enfoque que da cuenta de la relevancia de las relaciones espaciales y de cómo éstas evolucionan en el tiempo. Tercero, el estudio de la Geoweb no está restringido a las dimensiones geográficas, también abre la puerta para el análisis de redes. Cuarto, el contenido de la Geoweb no es producido solamente por usuarios humanos, sino es producto de un ensamblaje complejo y más que humano, que involucra una diversidad de actores, incluyendo los productores de contenido automático, como los robots spam de Twitter. Por último, se resalta la importancia de incluir datos no generados por los usuarios, tales como fuentes de datos gubernamentales o corporativas, como un apoyo a la investigación de la Geoweb (Crampton y cols., 2013, p: 131).

En conclusión, el paradigma de los grandes datos es una ventana de oportunidad para las

ciencias sociales y, particularmente para las ciencias como la geografía o los estudios urbanos. Sin embargo, para el análisis de los fenómenos digitales desde una perspectiva social se debe tener en cuenta que este paradigma no solo es una nueva fuente de información, sino que, es parte del objeto de estudio. En este sentido, a continuación se presentan algunas propuestas teóricas para entender la conformación de espacialidades digitales.

### **1.5 Abordajes teóricos sobre las geografías digitales**

Los geógrafos han reconocido el papel de la tecnología en la producción del espacio y en la construcción de relaciones socioespaciales (espacialidades). En una primera instancia, con la construcción de mapas o cartografía marítima, seguido de los avances en las comunicaciones, hasta el desarrollo de las computadoras y de los Sistemas de Información Geográfica. En la construcción y desarrollo de esta representación del mundo siempre ha quedado patente que la relación tecnología y espacio está mediada por condiciones políticas, económicas y sociales. En este sentido, la geografía ha comenzado a teorizar las diferentes maneras de construcción de espacialidades mediadas y formadas por la dimensión digital. Así, Leszczynski (2019) esquematiza las propuestas de la producción del espacio digital en cuatro vertientes: espacios híbridos, código/espacio, espacialidades mediadas y realidades aumentadas digitalmente. Cada una de ellas aborda diferentes aspectos de la dimensión digital y la constitución geográfica.

La primera aproximación teórica sobre las geografías digitales es la de *espacios híbridos*. Los espacios híbridos son teorizados como dos distintos reinos del espacio y de la espacialidad. Por un lado, los espacios digitales o virtuales son inmateriales y alejados de lo que ocurre en el mundo material. Mientras, los espacios reales (físicos o materiales) se constituyen como reales o tangibles. Esta separación epistemológica se originó por la escasa accesibilidad a la dimensión digital en la década de 2000. Sin embargo, esto ha cambiado gracias a la omnipresencia de dispositivos móviles que están conectados de manera continua a Internet. En este hilo de ideas, los espacios digital y real han configurado como un espacio híbrido donde coexisten ambas dimensiones (de Souza e Silva, 2006; Leszczynski, 2019). Aunque, para Leszczynski (2019) el enfoque de espacios híbridos no permite teorizar elementos las redes sociales ya que la separación ontológica de las espacialidades se restringe a entender el mundo digital en función de las condiciones materiales de accesibilidad. Que si bien son importantes, existen otras dimensiones en la formación de espacio digital.

La segunda aproximación teórica es la propuesta por Kitchin y Dodge (2011) y se refiere a espacios codificados y código/espacio. La primera diferencia con la propuesta de espacio híbridos radica en que la espacialidad no se constituye por medio de dispositivos tangibles, como computadoras o teléfonos celulares, sino que, la producción de espacio se basa en el rol del código computacional que subyace a dichos dispositivos. Al respecto, los autores señalan que el código es la célula primigenia de la revolución digital. Para los autores, el código o las instrucciones que hacen funcionar a los dispositivos digitales se encuentran inmersos en un halo de neutralidad, debido a su elevado nivel de abstracción matemática y técnica. Sin embargo, dicho código, en realidad, no es neutro y apolítico. En este sentido, Kitchin y Dodge (2011) señalan que el código también responde a intereses de clase y sociales. Una vez que los autores han señalado el papel político y social del código toman esto como base para explicar la conformación de distintos espacios en función del nivel de penetración del código. En primer lugar, los autores proponen que un espacio que funciona gracias al código, pero que no depende totalmente de éste es un espacio codificado. Es decir, en la actualidad muchos espacios cotidianos funcionan gracias a tareas automatizadas. Por ejemplo, los sistemas de cobro en centros comerciales o sistemas de transporte. Si el código llegase a fallar en estos espacios, la espacialidad no se perdería debido a las relaciones sociales formadas en dicha espacialidad podrían seguir existiendo. Por otro lado, cuando el código es omnipresente y de éste depende la producción misma del espacio, los autores le denominan *código/espacio*. Un ejemplo de este tipo de espacialidad son los aeropuertos y espacios aéreos. La primera característica del código/espacio del aeropuerto radica en que todas las actividades se encuentran computarizadas. La segunda característica se centra en la omnipresencia y omnipotencia del código para crear espacialidades. En el caso de los aeropuerto y espacio aéreo, todas las acciones que se llevan a cabo se dependen de la codificación y la digitalización, escáner, revisiones de seguridad, compras, abordaje, rutas aéreas, etc. Al mismo tiempo, la omnipotencia del código se muestra en la imposibilidad de sustituir la acción por medio análogos. Con respecto al aeropuerto, si el sistema de seguridad al abordaje fallase, toda la constitución del aeropuerto y vuelos no se llevaría a cabo ya que el registro depende enteramente de la dimensión digital. La propuesta de código/espacio apunta hacia una realidad donde todas las acciones humanas dependan absolutamente del código para poder llevarse a cabo. Ejemplos de esto, son el Internet de las cosas, los automóviles autónomos o la Inteligencia Artificial.

Dentro de la propuesta de código/espacio, los autores proponen entender la configuración de espacialidades basadas en elementos tecnológicos por medio de la ‘transducción del espacio’ (Dodge y Kitchin, 2005; Kitchin y Dodge, 2011). “La transducción es un proceso de ontogénesis, genera de manera diferente un dominio de individuaciones transformativas y reiterativas”(Dodge y Kitchin, 2005, p: 170). Primero, la individuación son las acciones cotidianas que todos hacemos en un entorno cotidiano determinado, transportarse, comunicarse, trabajar, etc. Estas acciones se ven transformadas de manera continua por medio de la intervención de tareas codificadas. En otras palabras, el código ayuda a realizar las tareas pero al mismo tiempo las modifica. Un caso de lo anterior son las compras realizadas por medios digitales, por ejemplo en Amazon. En una primera instancia, las personas realizan una tarea cotidiana de comprar para satisfacer una necesidad, al realizarlo por medio de Amazon se puede satisfacer dicha necesidad que antes se cumplía por medio análogos. En segundo lugar, dicha compra está mediada por códigos que establecen preferencias y necesidades de las personas. En tercer lugar, el código de Amazon puede pronosticar posibles necesidades con el objetivo de continuar con la tarea de compras. En este caso, el código de Amazon transduce el espacio de compras ya que se basa en una necesidad original, la realiza y al mismo tiempo la está modificando al generar nuevas maneras de generar necesidad de compras. Esta transducción se encuentra también en la generación de espacios en la vida cotidiana.

El tercer acercamiento teórico es el denominado *espacialidades mediadas*. Esta propuesta avanza en la teorización sobre la generación del espacio en la era de nativos digitales. “La *media-ción* se refiere a la condición contemporánea de la manera en la que experimentamos ser, tanto humanos como otros digitalmente no-humanos, en espacio y prácticas de la vida cotidiana, en la cual, nosotros vamos a entender los espacios, experiencias e interacciones como los efectos de los innumerables resultados conjuntos de la tecnología, las sociedad y las relaciones espaciales”(Leszczynski, 2019, p. 19-19). Es decir, las espacialidades mediadas son constituidas como efecto del ensamblaje conjunto de condiciones sociales, humanas y no humanas, tecnológicas y no tecnológicas. Un claro ejemplo de esto en la generación de espacios por medio de redes sociodigitales eminentemente audiovisuales. Instagram es un ejemplo de la generación de espacialidades *media-das*.

Cada una de las propuestas antes mencionadas recuperan elementos de la dimensión digital y cómo se construyen espacialidades basadas en éstas. En este sentido, cada una de ellas puede dar

cuenta de fenómenos geográficos o urbanos particulares basado en las características que resaltan. Así, la propuesta de espacios híbridos se ha utilizado para entender la conformación de nuevas interacciones sociales en los espacios públicos urbanos, particularmente, basado en el uso de telefonía móvil. También, los espacios híbridos han sido estudiados para dar cuenta de las transformaciones en la sociabilidad espacial basadas en la emergencia de videojuegos móviles (de Souza e Silva, Duarte, y Damasceno, 2017; de Souza e Silva, 5). En este sentido, esta propuesta enfatiza el papel de las comunicaciones digitales para la creación de espacialidades. Por su parte, la propuesta de espacio/código enfatiza el papel del código y software como actante en la creación de una espacialidad. Así, dicha propuesta deja de lado los dispositivos o aplicaciones específicas y centra su análisis en el papel del software. En consecuencia, esta perspectiva no centra su análisis en una espacialidad específica, sino que proporciona elementos para entender cualquier espacialidad digital particular, debido a que el código subyace a todo dispositivo digital. Por su parte, la espacialidad digitalmente *media-da* se centra en la generación de espacio por medio de las plataformas digitales dirigida a redes sociales.

Aunque todas las propuestas se interesan por la configuración geográfica de la dimensión digital y enfatizan los elementos políticos de éstas, ninguna de las anteriores permite entender fenómenos urbanos estructurales que se modifican y exacerban gracias a la dimensión digital. De esta manera, a continuación se desarrolla una cuarta perspectiva de creación de espacialidad digital, la cual puede ayudar a entender cómo fenómenos urbanos, se ven afectados y aumentados por la dimensión digital. De manera particular, cabe la pregunta sobre desarrollos teóricos desde las geografías digitales que permitan entender el proceso de digitalización de fenómenos de estructuración social urbana, como es el caso de este trabajo sobre gentrificación. En este sentido, buscar cómo los postulados de las diferentes geografías digitales pueden compaginar con los postulados teóricos de la gentrificación, tales como los del desarrollo urbano desigual o la gentrificación planetaria. Ante este, en la siguiente sección se desarrolla con mayor profundidad una de las vertiente de las espacialidades digitales que pueden ayudar a avanzar en el conocimiento de la transformación y representación de la gentrificación en la era de la digitalización.

## **1.6 Realidad geográfica aumentada digitalmente (RGAD)**

El concepto de Realidad Aumentada (RA) ha cobrado importancia y popularidad debido a la comercialización de aplicaciones móviles y no móviles. Algunos investigadores de ciencias computacionales definen a la RA como cualquier sistema que cumpla con tres características, que combine lo real y lo virtual, que sea interactivo en tiempo real y que su registro sea en 3 dimensiones (Azuma, 1997). Con el desarrollo de los dispositivos móviles y de la tecnología de GPS, la RA se popularizó por medio de juegos en tiempo real que conjuntaban por medio del dispositivo elementos digitales enmarcados en entornos físicos. El ejemplo más conocido de este tipo de aplicaciones es el juego de *Pokémon Go* (Rauschnabel, Rossmann, y Tom Dieck, 2017; Walker, McMahon, y Rosenblatt, 2017). Sin embargo, en términos académicos, la RA refiere a la constitución política y social de la realidad en su interacción con la dimensión digital, para tales propósitos de diferenciación, en este trabajo se referirá al concepto de RA por medio de Realidad Geográfica Aumentada Digitalmente (RGAD).

La RGAD intenta capturar los mecanismos con los cuales lugares y espacios son reflejados en la dimensión digital y, al mismo tiempo, cómo estos espacios adquieren nuevas características mediadas por la dimensión digital. Para M. Graham, Zook, y Boulton (2013a) el término de *Realidad Aumentada*<sup>6</sup> refiere a los “los nexos materiales/virtuales mediados a través de la tecnología, información y código y representados en configuraciones espacio-temporales específicas e individualizadas” (M. Graham, Zook, y Boulton, 2013a, p:465). Este concepto desarrolla la construcción política y desigual del espacio basado en las plataformas digitales, esto se basa en la idea de que la representación geográfica de cualquier lugar es siempre subjetiva. En este sentido, “las *realidades aumentadas* son más que la simple representación de un lugar, éstas son parte del lugar en sí mismo” (M. Graham y cols., 2015, p: 89). M. Graham y cols. (2015) definen la RGAD o RA con base en tres características primordiales, *poder de la distribución*, *poder de la comunicación* y el *poder del código*. En conjunto, estas tres dimensiones constituyen la conformación de la RGAD como una espacialidad digital

En primera instancia, el poder de la distribución hace referencia a la primera barrera social que las personas deben superar para participar del proceso de RGAD, el acceso a Internet. El acceso a Internet no se distribuye de manera homogénea, ni entre las personas, ni en el espacio. A nivel

---

<sup>6</sup>La propuesta original de esta espacialidad digital retoma el concepto de Realidad Aumentada, por esa razón se mantiene esta nomenclatura para exponer sus características, aunque a lo largo de este trabajo se refiera como Realidad Geográfica Aumentada Digitalmente

mundial solo la mitad de la población tiene acceso a Internet y esta cifra crece considerablemente en países desarrollados económicamente, Europa, Norteamérica y algunos países de Asia. Este es un primer obstáculo que impide a las personas generar y participar de las realidades aumentadas. En términos de poder, las personas que poseen las condiciones materiales para lograr participar de manera continua en el terreno digital son las que generan de manera más eficiente la transformación y configuración de espacialidades con base en su sesgo político y social (M. Graham y cols., 2012).

Desde otro punto de vista, el poder de la distribución de la RGAD se vincula con lo que otros autores han denominado como Desigualdad Digital (DD). La DD es definida como el diferencial en el acceso a Tecnologías de la Información y Comunicación, en otras palabras, las personas que tienen o no acceso a dichas tecnologías. Dicho diferencial en el acceso no es explicado solo por el porcentaje de penetración de la tecnologías en la sociedad, sino que se encuentra sustentando en disparidades socioestructurales previas. En este sentido, la DD ha sido estudiada desde sus dimensiones de género, estrato social, raza, lenguaje, escolaridad, capacidades emocionales y cognitivas o culturales (DiMaggio y Hargittai, 2001; M. Graham y cols., 2012; S. Graham, 2002; van Dijk, 2005, 2012). Tanto la dimensión del poder de distribución de la RGAD como la desigualdad digital apuntan a la desigualdad social en una sociedad donde la omnipresencia de los dispositivos digitales es vista como la panacea de los problemas contemporáneos.

El segundo elemento constituyente de la RGAD es el *poder de la comunicación*. El poder de la comunicación o representación apunta a dos elementos dentro de la construcción de espacialidades digitales. El primero es el poder que se basa en la habilidad de usar de manera efectiva y eficiente los canales de comunicación para promover objetivos particulares de representación. El segundo aspecto del poder de la comunicación es la habilidad, no solo de crear e interpretar contenido, sino recircularlo, codificarlo y de disputarlo. Esta dimensión del poder en la RGAD permite a determinadas personas y grupos filtrar y promover la representación y cierto significado establecido sobre algún lugar (M. Graham, Zook, y Boulton, 2013a, p: 469).

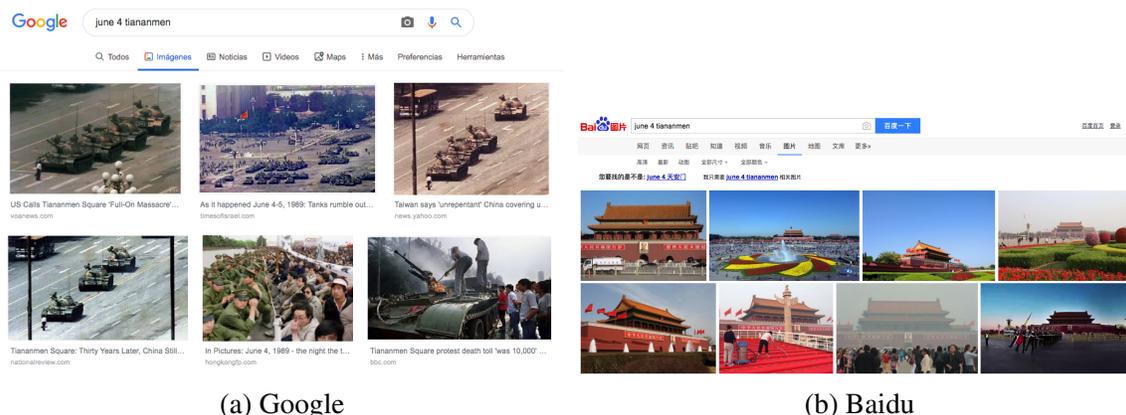
El poder de la comunicación de RGAD se basa en el auge de dispositivos móviles con conexión a Internet y su capacidad de ser geográficamente localizados por medio de geolocalización (GPS). En la última década el uso de teléfonos móviles se ha incrementado exponencialmente, prácticamente existe una suscripción a teléfono móvil por habitante a nivel mundial. Como consecuencia de esta tendencia y el desarrollo de tecnologías de transmisión de datos móviles, como

3G, 4G o 5G el acceso a Internet por medio de teléfono celular también se ha incrementado. Para 2018, a escala mundial, alrededor de 70% de la población accede de los usuarios de Internet por medio de alguna tecnología de acceso móvil (ITU, 2018). Aunado a este crecimiento, se vincula la posibilidad de poder generar información geolocalizada en tiempo real, Información Geográfica Voluntaria (IGV)

La IGV es un caso particular de la información generada en Internet (Goodchild, 2007). En general, la IGV hace referencia a los contenidos multimedia que se producen por medio de los usuarios de Internet que comparten su ubicación por medio de los datos asociados al GPS. Ejemplos de este tipo de información se pueden encontrar en plataformas de colaboración entre pares como Wikipedia, OpenStreetMaps, etc. En el marco de la IGV, Goodchild propone entender que las personas se han convertido en sensores gracias a la capacidad tecnológica de generar y transmitir información en tiempo real. De igual manera, existen motivos para que las personas inviertan gran cantidad de su tiempo para generar información geográfica voluntariamente. Una de las explicaciones que ofrece Goodchild es la auto-realización del individuo, elementos narcisistas y aprobación del individuo ante la sociedad. En contraste con esta idea, se tiene la experiencia de este tipo de información para dar cuenta de desastres naturales en tiempo real, lo cual no se explica por motivos egoístas sino por altruismo y cooperación (Elwood y cols., 2012). La IGV es una característica que subyace al poder de la comunicación dentro de la RGAD debido a que la posibilidad de crear contenido georreferenciado en tiempo real se encuentra sujeto a condiciones sociales y políticas específicas.

Al respecto, el poder político y económico es un actor crea y oculta espacios digitales. Este tipo de poder produce espacio por medio del lenguaje, ya que la mayor cantidad de contenidos se generan en idioma inglés. M. Graham (2017) muestra cómo la mayoría de artículos de Wikipedia son escritos en países Europeos, Estados Unidos y Japón, la mayoría de ellos en inglés. Al mismo tiempo, el autor muestra cómo por medio de realizar búsquedas en diferentes servidores se pueden obtener resultados controlados políticamente 1.4. Por ejemplo, cuando se realiza una búsqueda por medio de motores como Google sobre el 4 de junio en Tiananmen, China se obtienen resultados sobre las protestas sociales y masacre en dicha plaza de China (ver figura 1.4a). Mientras que, si se usan buscadores controlados oficialmente por el gobierno de China se obtienen resultados positivos sobre la plaza (ver figura 1.4b). Esto da muestra la posibilidad de construir diferentes

Figura 1.4: Búsquedas en diferentes motores sobre Tiananmen, China



Fuente: elaboración propia con base en M. Graham (2017)

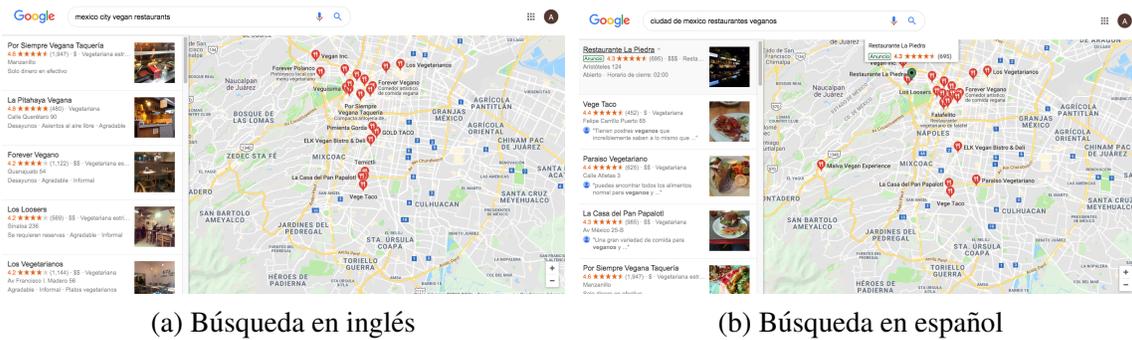
tipos de espacialidades mediante el contenido digital, ya que éste responde también a intereses políticos

En este mismo orden de ideas, la construcción de RGAD por medio de la comunicación se expresa a través del poder del idioma. En internet, la *lingua franca* es inglés. Por lo tanto, los resultados de búsquedas son diferentes si se realiza la consulta en inglés o en otro idioma. Por ejemplo, si se buscan restaurantes en una zona determinada. En la figura 1.5a y 1.5b se muestran los resultados de la búsqueda de restaurantes veganos en inglés y español en la Ciudad de México mediante Google Maps. La búsqueda difiere en dos grandes aspectos. El primero el orden diferenciado de los primeros lugares en la lista. El segundo, muestra que las búsquedas en idioma anglosajón se concentran en el centro de la ciudad. Por su parte, los resultados en español arrojan restaurantes en zonas periféricas de la urbe. En otras palabras, el poder de la comunicación basado en el idioma permite ocultar o enfocar solo determinados lugares.

Como se ha visto, la construcción de una RGAD se establece con base en el poder de acceso a Internet y a la dimensión digital, en general. Segundo, la construcción de RGAD responde a diferentes maneras de comunicar un mismo contenido. Por último, éstas dos dimensiones interactúan y se fortalecen por medio de una característica *invisible* para el ojo humano, pero sustancial para las computadoras, el código.

En conjunto, el poder de la comunicación para crear una RGAD se sustenta en dos grandes dimensiones. La primera de ellas, los lugares que se representan por medio de la IGV. Es decir, por medio de la información que las personas emiten constantemente en la red se expresan sus

Figura 1.5: Búsqueda de restaurantes veganos en la Ciudad de México en inglés y español



(a) Búsqueda en inglés

(b) Búsqueda en español

Fuente: elaboración propia

posiciones políticas y económicas, al evidenciar más unos lugares y ocultar otros. Otro elemento es la manera en la que se representan los lugares en Internet gracias al lenguaje. En este sentido, el idioma o la manera de búsqueda también representa u oculta lugares. En particular, las maneras diferenciadas de comunicación basadas en el lenguaje se encuentran íntimamente vinculadas con la última dimensión de la RGAD, el poder del código.

El tercer poder que caracteriza a la RGAD es el poder del código. Esta dimensión se basa en la propuesta realizada por Dodge y Kitchin (2005) sobre la no neutralidad del código y la transducción del espacio generado por éste. En particular, en el caso del código para generar RGAD la literatura especializada se ha centrado en demostrar la manera en la que las compañías rectoras de la vida digital condicionan la conformación de espacio. El caso más analizado ha sido el papel que juega el código de *Page-Ranking* de Google para arrojar resultados de una búsqueda en Internet. Por ejemplo, (M. Graham, 2017) muestra cómo cuando se realiza una búsqueda de pizzerías en inglés se obtienen resultados diferentes en comparación de la búsqueda realizada en otro idioma. En consecuencia, los autores demuestran cómo por medio de la búsqueda en Internet en Google se representan unos lugares y se ocultan otros, ya sea por motivos, sociales, económicos o políticos.

El código ha pasado a formar parte de la producción del espacio. Al respecto, Shaw y Graham (2017) argumentan que el código interactúa en la producción del espacio desde una perspectiva Lefebvriana. En otras palabras, la producción del espacio desde las dimensiones de concepción, realización y representación se magnifican y afectan por medio de los medios digitales. Para los autores, Google es un claro ejemplo de este tipo de producción y de poder. En este sentido mencionan que, el control que Google tiene sobre la nuevas formas omnipresentes del espacio digital posibilita que

la empresa reproduzca y controle el espacio en sí mismo. Google es capaz de controlar la importancia de las agendas de planeadores o desarrolladores en términos de su poder sobre la ciudad y la producción del espacio. Como base de este poder, mencionan los autores, está la fuente misma del poder de Google, la información de personas y de los datos de los fenómenos sociales urbanos en tiempo real, así como su cuantificación, procesamiento y análisis (Shaw y Graham, 2017). Un ejemplo del poder de la empresa en entornos urbanos es la capacidad de analizar en tiempo real el tránsito por medio de Google Maps, esto basado en la información recolectada en tiempo real de los teléfonos móviles de las personas.

A manera de conclusión, la Realidad Geográfica Aumentada Digitalmente, es una propuesta teórica que se enfoca en explicar cómo se constituyen ciertas espacialidades digitales por medio de la interacción de los diferentes poderes que la intervienen. En concreto, la RGAD se conforma por el poder de la distribución, comunicación y del código. En conjunto, los tres poderes intervienen para la generación de RGAD, siendo este tipo de espacialidades no solo reflejo de lugares físicos en el dominio de lo digital, sino una espacialidad construida en términos políticos y sociales por medio de lo digital. Con esto en mente, en este trabajo de investigación se retoma el concepto de la RGAD como herramienta teórica desde las geografías digitales para proponer que el proceso de gentrificación en la era digital se encuentra aumentado digitalmente.

### **1.7 Gentrificación digitalmente aumentada**

Hasta el momento se han expuesto los elementos teóricos y empíricos del proceso de gentrificación. Al mismo tiempo se ha hablado del enfoque de los grandes datos para entender la modificación de los procesos sociales ante el avance de la era digital. En este sentido, también se han mencionado las diferentes aproximaciones teóricas que desde la geografía se proponen para interpretar la realidad desde una perspectiva de espacialidades digitales. Dentro de estas aproximaciones se ha comentado que existen, al menos, cuatro que proponen diferentes elementos para entender la constitución de fenómenos geográficos en su vertiente digital. En particular, en este trabajo se ha escogido la propuesta de Realidades Geográficas Aumentadas Digitalmente para tratar de aproximar cómo se constituye la dimensión digital del proceso de gentrificación.

Esta elección se sustenta en el hecho de que ambos conceptos recuperan cuestiones.

La gentrificación es un proceso que da cuenta de la diferenciación social y política del espa-

cio urbano. Esto se da por medio de elementos económicos como la renta del suelo, por elementos simbólicos y culturales como los patrones de consumo y patrones comerciales o bien por la intervención del Estado, esto se ve en el proceso de renovación de determinadas áreas de las urbes. Sin embargo, dentro de este conjunto de enfoques que han abordado dicho fenómeno se agrega la presencia y uso de dispositivos digitales, Internet y los grandes datos. De esta manera, uno de los objetivos principales de esta tesis es indagar sobre una posible dimensión digital de la gentrificación. Es decir, se pretende explorar la posibilidad de que el proceso de gentrificación, particularmente en la ciudad de México, se ve aumentando digitalmente a través del uso de dispositivos móviles y de redes sociodigitales como Twitter. Con este propósito se retoma una de las propuestas de la corriente de geografías digitales, la Realidad Geográfica Aumentada Digitalmente, la cual proporciona algunos elementos para entender la interacción entre las características de la espacialidad digital y las características del proceso de gentrificación.

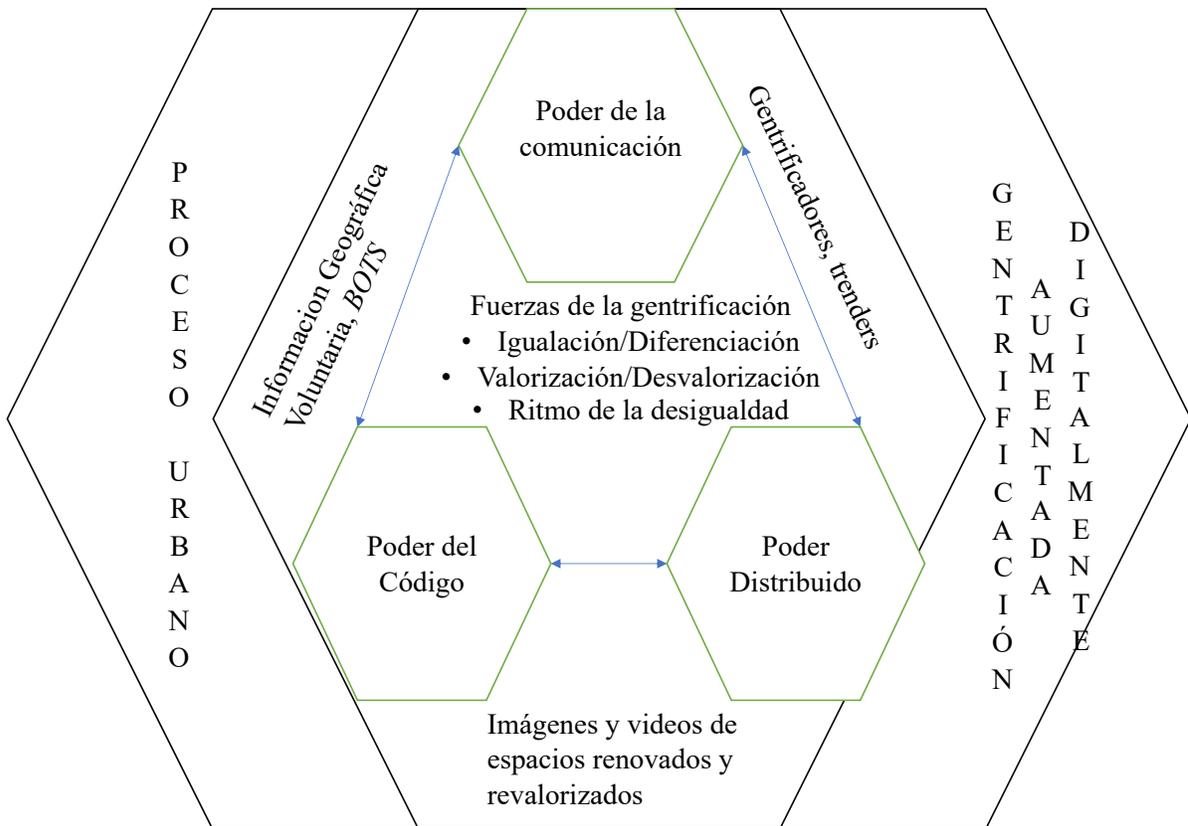
El presente trabajo propone analizar una dimensión poco estudiada de la gentrificación en la Ciudad de México, la dimensión digital. Es importante mencionar que, aunque se ha avanzado en la comprensión de fenómenos geográficos desde una mirada digital, la investigación sobre los mecanismos articuladores de las múltiples espacialidades digitales que se pueden configurar son aun incipientes. En este sentido, la aproximación teórica que en este trabajo se propone es conjuntar elementos de la gentrificación que han sido estudiados a lo largo de la evolución de dicho proceso y someterlos a los mecanismos de creación de una espacialidad digital particular, la RGAD. Este ensamblaje permitirá, a lo largo de los capítulos subsecuentes, explorar de manera empírica algunas de las dimensiones de la RGAD y del proceso de gentrificación para conocer en qué medida se puede hablar de una gentrificación digitalmente aumentada.

La propuesta teórica que aquí se construye amalgama dos apartados. En primera instancia, se recupera los elementos propuestos por N. Smith (1982) para entender el proceso de gentrificación como una manifestación del desarrollo desigual del sistema económico contemporáneo (Ver sección 1.2.4). De manera particular, se pone especial énfasis en una aproximación culturalista de la valorización y revalorización del espacio en la ciudad, esto basado en las propuestas de gentrificación cultural de Sharon Zukin. En segunda instancia, se articulan las tres dimensiones que componen la RGAD como marco de entendimiento en dicho proceso de valorización/revalorización del espacio con el fin de evaluar la capacidad teórica de dicha aproximación para estudiar la

gentrificación digital. Esto con el fin indagar sobre algunas manifestaciones de la gentrificación aumentada digitalmente.

El orden expositivo del modelo teórico propuesto se basa en las los tres elementos que componen la gentrificación como desarrollo desigual urbano propuesto por Smith <sup>7</sup> y cómo éstas se ensamblan en las tres características de la RGAD. La figura 1.6 muestra cómo las fuerzas de la gentrificación pueden ser transformadas por la dimensión digital. En el centro de la figura se encuentra la interacción entre las características de la RGAD y los elementos del desarrollo desigual que constituyen la diferenciación espacial. De manera integral, todas las dimensiones interactúan mutuamente, pero existen ciertos elementos de la gentrificación que se ven más afectados por las determinadas características de la RGAD.

Figura 1.6: Modelo teórico del proceso de gentrificación como realidad geográfica aumentada digitalmente



Fuente: elaboración propia con base en M. Graham, Zook, y Boulton (2013b); N. Smith (1982)

<sup>7</sup>En esta propuesta solo se desarrollan las dos primeras características de la gentrificación como desarrollo desigual, igualación/diferenciación y valorización/desvalorización del espacio. Aunque la tercera dimensión permite entender la gentrificación desde un enfoque totalmente sistémico, ésta se omitirá por razones de objetivos de la investigación

Para Smith, una primera esquematización teórica sobre la gentrificación es la construcción estructural del entorno urbano como lucha de fuerzas entre la ciudad central y la periferia <sup>8</sup>. De esta manera el proceso de gentrificación conlleva la diferenciación del espacio urbano por medio de la división social del trabajo, y al mismo tiempo se sigue la tendencia en la igualación de diferentes espacios a través de diferentes mecanismos económicos, como la igualación de la renta. Con base en esta idea, la propuesta de RGAD puede ayudar a interpretar de qué manera este proceso de igualación/diferenciación se lleva a cabo en espacios centrales/periféricos de la ciudad y, a su vez, refleja las tendencias gentrificadoras en las ciudades.

La RGAD puede ayudar a entender la manera en la que el proceso de diferenciación del espacio urbano se aumenta digitalmente gracias al conjunto de los tres elementos que componen a la realidad aumentada. En primer lugar la diferenciación del espacio urbano obedece a la división social del trabajo, ésta, a su vez, ha favorecido la concentración de las ramas productivas con mayor tasa de ganancia en determinados espacios, dichas ramas se asocian a empresas de tecnología o financieras (Sassen, 1991) Ciudades con núcleos de alto nivel tecnológico han comenzado a experimentar procesos de gentrificación. Un ejemplo de esto se ha mostrado en Barcelona o en San Francisco, donde a raíz de la concentración de empresas de alta tecnología se han incrementado las rentas para la habitación (Gdaniec, 2000; Wyly, 2018). En este sentido, las personas que laboran en actividades tecnológicas tienen un perfil determinado dentro del poder de la distribución digital. Es decir, asociado a las empresas de tecnologías se encuentran personas que han desarrollado habilidades digitales y tienen acceso a estas herramientas. Así, la concentración de empresas tecnológicas en puntos específicos de las ciudades se asocia a la concentración del poder en el acceso a la dimensión digital. Otro caso de concentración del poder digital distribuido en las ciudades es el patrón de emisión de redes sociodigitales, por ejemplo Twitter. En este punto, los trabajos empíricos que han dado cuenta del patrón de concentración de dichas emisiones apuntan a los centros de las ciudades o los distritos centrales de negocios (Croitoru, Wayant, Crooks, Radzikowski, y Stefanidis, 2015; García-Palomares, Salas-Olmedo, Moya-Gómez, Condeço-Melhorado, y Gutiérrez, 2018; Gibbons y cols., 2018; Lansley y Longley, 2016; Shelton, 2017, 2015).

La segunda dimensión de RGAD, el poder de la comunicación, puede ayudar a entender cómo

---

<sup>8</sup>Aunque este modelo centralista de la gentrificación ha sido discutido, argumentando que existen otros tipos de gentrificación en la periferia o en espacios rurales (Lees y cols., 2015), se retoma el modelo por simplicidad analítica

este proceso de diferenciación urbana se ve aumentado por la interacción digital. Recientemente, se ha comenzado a explorar la manera en la que los contenidos emitidos en redes sociodigitales y plataformas con lógica geográfica en Internet ayudan a representar o formar una concepción sobre determinados lugares. De manera particular, se ha indagado sobre la comunicación de orden visual en áreas gentrificadas. En este orden de ideas, algunos autores han mostrado cómo a través de fotografías de plataformas audiovisuales tales como Google Maps o Instagram se observa la reconversión del espacio áreas gentrificadas (C.-C. Huang, 2017; Ilic, Sawada, y Zarzelli, 2019). Otro elemento que se ha explorado es como por medio de la búsqueda de vivienda en Internet se reflejan el diferencial de la renta por medio de páginas web de búsqueda inmobiliaria (Boeing, Besbris, Schachter, y Kuk, 2020; Shamsuddin y Srinivasan, 2020). Finalmente, otra manera en la que la centralidad de espacios gentrificados se aumenta es por medio de la dimensión léxica en los sitios de redes sociodigitales. Un ejemplo de este punto, los trabajos de Lai, Cheng, y Lansley (2017); Lansley y Longley (2016) muestran la manera en la que por medio del texto emitido en Twitter, los espacios centrales o gentrificado de Londres se manifiestan en las emisiones cotidianas de la ciudad, así, puntos como el Soho, Oxford Street o El Museo Británico.

Este punto teórico es de suma importancia, debido a que será uno de los elementos que serán abordados en el desarrollo empírico en capítulos posteriores de este trabajo de investigación. Aunque el trabajo previo para identificar una geografía digital léxica de la gentrificación no ha sido abordada explícitamente, los trabajos de Lai y cols. (2017); Lansley y Longley (2016) apuntan a la posibilidad de identificar la centralidad de espacios gentrificados. Sin embargo, este cuerpo de literatura ha señalado la centralidad de la emisión de Twitter, sin enfocar, a su vez, la emisión en espacios más allá de los identificados como gentrificados. De esta manera, una de las hipótesis empíricas que guían la gentrificación aumentada digitalmente es que el poder de la comunicación permite que la centralidad de espacios gentrificados trascienda sus propios límites espaciales. En otras palabras, el poder de la comunicación aumenta la gentrificación mediante las redes sociodigitales en espacios no gentrificados, esto por medio de los mensajes léxicos de los usuarios.

El segundo elemento de la gentrificación que se ve aumentado por la mediación digital es la valorización/desvalorización del espacio. Como se ha mencionado en secciones precedentes, este elemento de la gentrificación señala que los espacios urbanos se insertan en un ciclo de valorización/desvalorización económica a lo largo del tiempo. Aunque Smith primordialmente refiere este

proceso a la valorización económica de los lugares también abre la puerta a la valorización simbólica de los espacios gentrificados. En este punto, desde la corriente culturalista se ha analizado los determinantes culturales y simbólicos que intervienen en la valorización de espacios gentrificados o en la desvalorización de espacios peligrosos. De esta manera, lo que se desarrolla en este punto es cómo los poderes de RGAD podrían ayudar a aumentar digitalmente la dimensión de valorización/desvalorización del espacio.

La actual valorización del espacio vinculada al poder de la distribución digital se ha dado por medio de dos vías. La primera es la localización de la infraestructura necesaria para acceder a la dimensión digital, es decir, la presencia de fibra óptica, antenas receptoras o repetidoras, conectividad de banda ancha, entre otras. En este sentido, S. Graham y Marvin (2001) argumentan que la dotación desigual de infraestructura tecnológica conllevaría al desarrollo de un urbanismo fragmentado. En el caso de la ZMCM Garnica (2014) encuentra que la dotación del servicio de Internet se concentra en las zonas centrales de la ciudad, que a su vez, son las de mayor precio en la renta muestran. La segunda se manifiesta por medio de la concentración geográfica de actividad en ciertas plataformas digitales. En este punto, varios estudios han mostrado la concentración de emisión a escala intraurbana de redes sociodigitales como Twitter o Instagram a lo largo del mundo. En este sentido, los caso más estudiados de concentración de emisión se han dado en las principales ciudades, Nueva York, Londres, Barcelona, Tokio, Paris (García-Palomares y cols., 2018; Lansley y Longley, 2016; Neuhaus, 2015).

Otro punto a considerar es la capacidad que el poder de la comunicación digital podría tener para valorizar determinados espacio en términos temporales. Al respecto, bajo los lentes de la RGAD, se podría vislumbrar cómo los lugares se valorizan temporalmente diferente a lo largo del día y de la semana. Por ejemplo, se ha encontrado que la concentración geográfica de temas particulares en Twitter no es homogénea a lo largo del día, el estudio de Lansley y Longley (2016) muestra que lugares vinculados a entretenimiento y esparcimiento se manifiestan en mayor medida durante las noches y fines de semana. En este sentido, el análisis de la dimensión digital de la gentrificación podría arrojar luz más detallada de los componentes para valorizar determinados espacios gentrificados o en proceso de gentrificación (Zukin y cols., 2017).

Recientemente se han comenzado a investigar otras maneras en las que la dimensión digital ayuda a valorizar espacios gentrificados, específicamente por medio de aplicaciones móviles y de

economía compartida. En este punto se pueden mencionar dos casos. Por un lado, el desarrollo de plataformas móviles dirigidas exclusivamente al comercio *hipster* (Jansson, 2019). Por otro lado, las plataformas de economía compartida enfocadas a la vivienda, particularmente, Airbnb, ya que a raíz del auge de esta plataforma para la búsqueda de alojamiento en varias ciudades alrededor del mundo se ha observado que el precio de la renta promedio de estos inmuebles se ha incrementado causando procesos de expulsión de los habitantes de estrato bajo y clase trabajadora (Santos y Sequera, 2018; Mermet, 2017; Wachsmuth y Weisler, 2017)

Hasta el momento se ha desarrollado cómo los elementos de la gentrificación, igualación/diferenciación y valoración/desvalorización del espacio, propuestos por la teórica del desarrollo desigual podrían ser aumentados por dos poderes de la RGAD (poder de la distribución digital y de la comunicación). Aunado a esto falta explicar la manera en la que se podría esperar que el poder del código intervenga en los elementos de la gentrificación. Cabe mencionar que el poder del código refiere a la no neutralidad del algoritmo subyacente a todas las plataformas y desarrollos digitales. Sin embargo, acceder al código subyacente muchas veces es imposible de realizar, ya sea porque está protegido por derechos de autor o es ilegible por una persona promedio. En este sentido, el código computacional es una especie de caja negra a la cual pocos tienen acceso (Pasquale, 2016).

El presente trabajo de investigación no se desarrollará a profundidad cómo aumenta el poder del código los procesos de gentrificación. Ya que eso implicaría conocer el código de Twitter y cómo éste se ha ido modificando para favorecer determinados intereses económicos y políticos en la ciudad. Sin embargo, se puede ofrecer algunos esbozos que permitirían establecer la hipótesis de que el código enfatiza ciertos lugares gentrificados. En este orden de ideas, se ha analizado la manera en la que el algoritmo de Twitter construye las tendencias generales dentro de su sitio. Así, algunos lugares se ven sobrerrepresentados en momentos determinados y debido a sucesos específicos. De esta manera, el código asigna una importancia relativa a las publicaciones que refieren a los sitios que son tendencia. Un ejemplo de esto es el analizado por Crampton y cols. (2013) quienes identifican como una revuelta en la Universidad de Kentucky dio pie a que la ciudad fuese tendencia en Estados Unidos y que apareciera en el mapa de la discusión pública. Aunque este estudio da muestra de la construcción de espacios por medio del código, el estudio de las tendencias se analiza en temporalidades muy cortas. En consecuencia, se podría lanzar una pregunta acerca

de la manera de reflejar gentrificación por medio de Twitter o cualquier plataforma en términos temporales más amplios.

En conclusión, en este trabajo de investigación se propone utilizar la propuesta teórica de la RGAD con sus tres componentes; poder distribuido digital, poder de la comunicación y poder del código, como una aproximación teórica y metodológica para explorar cómo el proceso de gentrificación en la Ciudad de México podría verse aumentado digitalmente. En términos específicos, se retoma el concepto marxista del proceso de gentrificación como reflejo del desarrollo urbano desigual el cual propone entender la gentrificación como una dinámica de interacción entre procesos de igualación/diferenciación y valorización/desvalorización del espacio urbano. Como una limitación de este trabajo de investigación, solo se explora la manera en la que las fuerzas de gentrificación son aumentadas digitalmente por la distribución desigual de lo digital y de la comunicación. Lo anterior, solamente mediante una red sociodigital, Twitter. Por su parte, explorar cómo interviene el poder del código es imposible en esta etapa de la investigación sobre las geografías digitales, ya que implicaría una investigación exclusivamente sobre tal interacción.

### **1.8 Consideraciones finales. Hacia un estudio del proceso de gentrificación digitalmente aumentado**

En el presente capítulo se han expuesto los conceptos teóricos que permiten entender el fenómeno de gentrificación en su dimensión de realidad digitalmente aumentada. Para ello, el desarrollo expositivo se ha dividido en dos grandes apartados que se han ensamblado para lograr el objetivo antes mencionado. Por un lado, se mencionaron las aproximaciones teórico-metodológicas que permiten entender la gentrificación con base en algunas de las aproximaciones teóricas que se han propuesto desde la literatura internacional. En segunda instancia, se ha expuesto la importancia de la dimensión digital en la sociedad contemporánea y cómo ésta se ha comenzado a estudiar desde un enfoque geográfico. En conjunción, ambos apartados son ensamblados en una propuesta teórica para entender un fenómeno urbano particular en su interacción con la dimensión digital.

En lo que concierne las bases teóricas sobre la gentrificación se han elegido solo algunos de sus determinantes, así como un enfoque teórico específico. En particular, se han recuperado las propuestas teórico-metodológicas del perfil sociodemográfico de personas gentrificadoras, así como el acercamiento de gentrificación comercial y guiada por el Estado. Al mismo tiempo, como

eje articulador el presente trabajo entiende la gentrificación desde la propuesta de N. Smith (1982, 2008). Este enfoque funciona como eje ensamblador entre el fenómeno urbano de la gentrificación y la dimensión digital. Estas acotaciones son importantes debido a que la discusión sobre gentrificación, en el debate contemporáneo, es inacabada y llena de particularidades. El presente trabajo no ha podido recuperar la complejidad del fenómeno y continua siendo un tema pendiente en futuras investigaciones.

Por su parte, la dimensión digital se expone como una realidad que es eminentemente geográfica. Al respecto se han expuesto algunos de los casos más representativos de la transformación de la realidad social por medio de plataformas digitales. Por ejemplo las Redes Sociodigitales, como Facebook, Twitter o bien los efectos de las plataformas ‘colaborativas’ como *Airbnb* o *UBER*. En este sentido, han retomado las propuestas de la geografía digital para entender la conformación de diferentes espacialidades digitales. Dentro de esta discusión, el presente trabajo retoma la propuesta de las Realidades Geográficas Digitalmente Aumentadas.

De manera concreta, la gentrificación digitalmente aumentada conjuga los elementos de la teoría del desarrollo urbano desigual con los elementos de la Realidad Geográfica Aumentada Digitalmente. Este trabajo de investigación propone entender cómo se aumenta digitalmente el proceso de gentrificación en la Ciudad de México. Para ello proponen estudiar la manera en la que el poder distribuido de lo digital y de la comunicación favorecen el reflejo y la creación de una espacialidad gentrificada basada en la igualación/diferenciación y en la valorización/desvalorización de espacios en la ciudad. Este marco analítico permitirá desarrollar a lo largo de los siguientes capítulos algunos apuntes sobre la hipótesis de que el proceso de gentrificación se aumenta digitalmente gracias a las diferentes dimensiones de la revolución digital.

Así, la pregunta rectora es ¿Cómo se manifiesta la gentrificación digitalmente aumentada mediante Twitter en la Zona Metropolitana de la Ciudad de México?. Para esto, el primer objetivo particular es conocer la manera en la que el poder de la distribución se manifiesta en México y particularmente en la Ciudad de México. En este mismo rubro, se propone indagar si existen diferencias socioestructurales que expliquen el poder de la distribución en Twitter. En otras palabras conocer, si el uso de Twitter se manifiesta con mayor recurrencia en ciertos estratos sociodemográficos, particularmente en los más privilegiados. El segundo objetivo para apuntalar la hipótesis es conocer la manifestación espacio-temporal del poder de la distribución y comunica-

ción de la Realidad Geográfica Aumentada Digitalmente para evidenciar procesos de centralidad y valorización de espacios gentrificados en la ciudad. Es decir, si la emisión de Twitter se asocia espacial y temporalmente con espacio gentrificados en la ciudad. Finalmente, para llevar a cabo empíricamente el desarrollo teórico de este capítulo se tiene como objetivo demostrar que el poder de la comunicación en Twitter en la ciudad refleja el proceso de gentrificación en las áreas centrales de la ciudad. En conjunción, el desarrollo posterior de este trabajo de investigación apuntalará cada una de las dimensiones de la propuesta teórica desarrollada durante el presente capítulo.

Finalmente, cabe mencionar que el presente capítulo también presente limitaciones en términos epistemológicos y de comprobación empírica. Por un lado, se ha dejado de lado la tercera fuerza del desarrollo desigual sobre la gentrificación. Es decir, no se ha podido construir un razonamiento de largo alcance temporal en el que la dimensión digital pueda ser factor determinante de la gentrificación. En este sentido, el estudio de la dimensión digital desde la geografía tiene un campo fértil de investigación, sobretodo a la luz del rápido avance de la revolución digital en las vidas cotidianas. El segundo elemento teórico que no se ha desarrollado ampliamente ha sido la interacción entre el poder del código con todas las fuerzas de la gentrificación. Esta limitación también responde al naciente campo de estudio en el que se inserta este trabajo de investigación. Aunado a lo anterior, realizar estudios a profundidad sobre el papel del código requeriría del desarrollo de habilidades etnográficas y de programación para conocer de primera mano la construcción de algoritmos específicos y cómo estos permiten el aumento digital de los fenómenos sociales.



## Capítulo 2

# Desigualdad Digital y perfiles de uso de Redes Sociales y Twitter. El poder distribuido de la dimensión digital

### 2.1 Introducción

El proceso de gentrificación es uno de los más relevantes y fragmentadores del espacio urbano que se experimentan en la actualidad. Al proceso subyacen tensiones de clase y la diferenciación de poderes que intervienen en los elementos económicos y urbanos. Así, la gentrificación, lejos de entenderse exclusivamente desde la estructuración urbana solo por medio de elementos económicos, se configura por medio de dimensiones sociales, de renovación urbana y políticos (López-Morales, 2016). Entre los elementos no económicos, el presente trabajo tiene como objetivo indagar sobre los elementos digitales de dicho fenómeno. Lo anterior bajo los argumentos expuestos en el capítulo anterior. Al igual que la gentrificación, los fenómenos digitales se constituyen por elementos de diferenciación social, económica y política. El presente capítulo tiene como objetivo principal indagar sobre el poder distribuido de la dimensión digital, con el fin de aproximarse al entendimiento de una gentrificación aumentada digitalmente en la Zona Metropolitana de la Ciudad de México. Dado que se ha seleccionado Twitter como una fuente de información para explorar la posibilidad de una gentrificación aumentada digitalmente, el capítulo se centra en los principales determinantes sociodemográficos que permiten explicar el proceso de selectividad que este sitio de red sociodigital presenta. Aunado a lo anterior y compaginado con el andamiaje de la RGAD, en cuanto a las 3 dimensiones de poder que la constituyen, se propone entender el poder distribuido de lo digital por medio del marco de las desigualdades digitales.

Con esto en mente, el propósito de este capítulo es indagar sobre las desigualdades digitales existentes en México y en la ZMCM como base que refleja las desigualdades sociales para participar de las diferentes maneras de interactuar en la creación de espacialidades digitales. En particular, la desigualdad para acceder a Internet o plataformas de redes sociodigitales es atravesada por ele-

mentos socioestructurales que condicionan la constitución del poder distribuido de lo digital en el proceso de gentrificación aumentado digitalmente. Para tal propósito, el presente capítulo se divide en cinco subapartados. En el primero se exponen las principales aproximaciones conceptuales acerca de la desigualdad digital en términos abstractos y cómo ha sido abordada para el caso mexicano. En el segundo apartado se presentan los elementos fundamentales para entender las desigualdades digitales en las plataformas de redes sociodigitales, tales como sexo, edad, estrato socioeconómico, entre otros. En tercer lugar se presenta la construcción de indicadores sociodemográficos que sirven para entender la desigualdad en el acceso a Internet y Redes Sociodigitales basado en la Encuesta Nacional de Disponibilidad y Uso de Tecnologías de la Información en Hogares, generada por el Instituto Nacional de Estadística y Geografía (INEGI) en 2018. Enseguida, se presentan los principales resultados estadísticos que apuntan a proponer que existe una desigualdad digital en México y, particularmente, en la ZMCM. De manera general se presenta que es poco conocido el perfil sociodemográfico del usuario promedio de Twitter, en este aspecto, el capítulo avanza en la exploración de este elemento. Aunado a lo anterior, se muestra que el perfil de usuario de Twitter corresponde a un usuario de estrato económico medio y alto, lo cual haría muy probable que se relacione buena parte de los habitantes de colonias gentrificadas de la ZMCM con el usuario de esta plataforma. Finalmente se desarrollan algunas consideraciones finales que apuntan a la desigualdad digital en el uso de Twitter en la ZMCM como primer elemento constitutivo de una gentrificación aumentada digitalmente en la ciudad.

## **2.2 Desigualdad digital en México**

La Desigualdad Digital (DD) es un concepto que da cuenta de las disparidades en el acceso y uso de las Tecnologías de la Información y Comunicación (TIC) y los elementos socioestructurales que la explican. De manera específica, la DD surge a raíz de la preocupación durante el gobierno de Bill Clinton para reducir la disparidad de acceso a computadoras e Internet entre la población estadounidense (Kuttan y Peters, 2003). Dicha preocupación se caracterizaba por su visión tecnodeterminista del uso de Internet como elemento para mejorar las condiciones de vida de las personas. A lo largo de 20 años, el estudio sobre las desigualdades digitales ha transitado hacia una concepción multidimensional. En la actualidad, la DD se enfoca en las dificultades para lograr el ideal de equidad en la apropiación de las TIC y los escenarios resultantes de dicha apropiación.

En este sentido, las desigualdades digitales reconocen las múltiples aristas del uso de las TIC (DiMaggio y Hargittai, 2001; Hargittai, 2010; Hargittai y Hsieh, 2013; van Deursen y van Dijk, 2010; van Dijk, 2005, 2012). Así, la desigualdad digital “aparece como un concepto nebuloso de difícil abordaje epistémico” (Toudert, 2014, p.559). Esta naturaleza compleja de la DD ha condicionado que se generen diversas formas de abordar el tema, éstas van desde la concepción primigenia de la DD, en la cual solo se consideraba el acceso/no acceso a Internet, hasta los desarrollos teóricos sobre la motivación, capacidades y adopción de estas tecnologías (Toudert, 2014). Aunque en términos generales la DD se ha enfocado a explicar el diferencial en el acceso y uso de TIC entendidas como acceso a dispositivos digitales tales como computadoras, teléfonos móviles o Internet, en años recientes la discusión ha mirado a entender el diferencial de uso de herramientas digitales tales como las Redes Sociodigitales, comercio o gobierno electrónico.

De manera general, en México, la DD se ha estudiado de manera incipiente. El interés sobre el uso y apropiación de Internet se ha hecho por pocos académicos, el gobierno federal y algunas fundaciones. Dichos estudios, a su vez, se pueden catalogar desde tres abordajes diferentes. El primero responde a la preocupación clásica entre el acceso/no acceso a Internet, esta perspectiva es recuperada tanto por académicos como el gobierno. Esto se debe a que, en un país como México la desigualdad digital se manifiesta en la imposibilidad de acceder a una computadora o Internet, ya que las desigualdades socioeconómicas que imperan en el país son factor determinante para explicar desigualdades subsecuentes, como la digital. Un segundo enfoque obedece a la preocupación de reducir las disparidades digitales por medios de políticas públicas y programas gubernamentales. En tercer lugar, la DD se analiza desde perspectivas más contemporáneas que incorporan elementos como el territorio, la apropiación, habilidades y usos digitales, es en esta vertiente desde donde se analiza este trabajo. A continuación se detallan cada uno de estos esfuerzos.

El estudio de la DD en México, como en otros países, comenzó por la perspectiva dualista y el enfoque tecnodeterminista del acceso a las TIC. Los primeros trabajos que se preocuparon sobre dicho fenómeno en el país se remontan a la primera década del siglo XXI. En ese periodo, el interés se centraba en conocer la situación del acceso a las TIC en la población, el lugar que ocupaba México en comparación con su contexto latinoamericano y qué relación tenían estas tecnologías en el crecimiento económico o su impacto en la pobreza. Bajo la perspectiva de acceso/no acceso se comenzó a pensar la DD en México.

Los primeros trabajos sobre la DD en México se enfocaron en el acceso a tecnologías en general, teléfono fijo o móvil, computadora e Internet. En el contexto latinoamericano México tenía mayor nivel de penetración de TIC en la población en comparación con países con PIB per cápita semejante en la región (Mariscal, 2005). Sin embargo, cuando se desglosan las TIC por tipo, ya sea teléfono fijo, computadora e Internet, la historia es diferente. Debido a que la mayor parte de penetración en México se le atribuía a la telefonía fija y a la par, la telefonía móvil creció de manera sustancial en el mismo periodo de análisis, ambas tasas de penetración alcanzaron 54% y 11% para telefonía fija y móvil respectivamente (Mariscal, 2009b). Por su parte la penetración de computadora e Internet era prácticamente nula. En 2006 el porcentaje de usuarios de Internet en México era 22.1% mientras que solo 5.48% de hogares era suscriptor de algún servicio de Internet, por lo tanto, solo 16% de la población tenía acceso a Internet en sus hogares (Mariscal, 2009a). En estos trabajos pioneros, la tasa de penetración en el uso de Internet se encontraba sustentada por las desigualdades sociales del país.

Una de las primeras relaciones que se exploraron fue la asociación entre la DD y los niveles de pobreza, marginación y la distribución territorial de éstas. Así, se comenzó a enfatizar la posición privilegiada de las ciudades con respecto a zonas rurales. Una de las disparidades más marcadas es la diferencia en el acceso de la Ciudad de México con respecto al resto del país, ya que esta ciudad por su tamaño e importancia ha figurado como la más sobresaliente en términos de adopción de tecnología en la población. Esta situación contrasta con el poco grado de penetración de TIC en estados del sur como Oaxaca, Guerrero y Chiapas. Ante este panorama, la conclusión se dirigió a la generación de políticas públicas que ayudaran a superar esta brecha tanto territorial como social, y que el acceso y uso de las TIC fueran detonantes para superar dichas desigualdades (Mariscal, 2005, 2009a, 2009b; Mariscal, Gil-García, y Aldama-Nalda, 2011; Mariscal y Martínez, 2013).

Al mismo tiempo que se consideró el nivel de acceso en la población, se comenzaron a hacer los primeros estudios en otras dimensiones de la DD, así como en la apropiación y uso de Internet. Ejemplos de esto son algunos esfuerzos por analizar la producción de contenidos en el país, en específico, la producción de sitios Web. La incipiente génesis del contenido web en México se caracterizó por la marcada diferencia en la creación de páginas web en comparación con países como Estados Unidos o Europa. Dichas páginas solo le correspondían a las grandes empresas, principalmente las dedicadas al entretenimiento (Toudert, 2000, 2001). De esta manera, la DD en

México se caracterizaba por el bajo nivel de penetración de Internet entre la población y la poca y concentrada creación de contenido en la red. Cabe aclarar que, las mejores condiciones de uso de Internet correspondían a la población de estratos altos y la poca creación y concentración de contenidos pertenecía a grandes empresas. Es decir, la DD era reflejo de las desigualdades sociales y económicas del país. Ante este panorama, se construyeron iniciativas de políticas públicas con el fin de cerrar la brecha.

Debido a la necesidad de incorporar a México en la era digital y en la economía de la información, las políticas públicas en torno al tema cobraron relevancia. En esta dimensión, vale la pena aclarar que la elaboración e implementación de políticas públicas enfocadas a reducir la desigualdad en el acceso y uso de TIC deben ser estudiadas bajo el matiz de las diferentes definiciones intrínsecas que se tiene sobre el concepto de la DD. De esta manera, el enfoque de políticas públicas parte de reconocer que los criterios para definir la DD se modifican entre comunidades. Así, es necesario hablar de diversas DD. En el mismo sentido, la desigualdad digital ‘se determina por los factores siguientes: la tecnología disponible; la apropiación social de las TIC, como resultado de su utilidad o valor social; la capacidad de las personas para beneficiarse de las TIC; las condiciones de desarrollo económico de las comunidades; la distribución geográfica de las mismas y por último las características demográficas de la población’ (la Selva, 2012, p. 56).

Desde la perspectiva de acción gubernamental, la DD puede abordarse desde dos enfoques. El primero de ellos “el llamado por Servaes *difusionista* (cuyo referente es el paradigma de la modernización) y el modelo *participativo-otro desarrollo* (que incorpora los concepto del esquema emergente de la multiplicidad/otro desarrollo)” (la Selva, 2012, p. 48). Aunque solo el primero ha marcado el proceder del gobierno mexicano para contrarrestar la DD, en específico, con programas como *e-Mexico*.

El proyecto e-Mexico fue el primer intento del gobierno federal por reducir la DD en el país. El programa tenía como objetivo ofrecer acceso gratuito a Internet a la población en lugares públicos, escuelas, oficinas de gobierno, bibliotecas. e-Mexico no solo pretendía fortalecer la masificación del acceso sino incrementar los usos de Internet. Para este objetivo, e-Mexico se enfocaba en cuatro dimensiones: educación, salud, economía (privilegiaba comercio y turismo) y gobierno. En el último rubro es donde se generó mayor cantidad de contenido, ya que un gran número de trámites gubernamentales se comenzaron a alojar en este sitio, uno de ellos fue la Clave Única

de Registro de Población (CURP). Aunque el programa se concibió como una solución integral al incorporar a los sectores clave de la sociedad, éste no tomó en cuenta la necesidad del acceso de alta velocidad o la falta de habilidades de la población para el manejo de Internet (Mariscal y cols., 2011). En otras palabras, no se contempló las condiciones sociales previas para la adopción y manejo de Internet.

Como se ha mencionado, el enfoque predominante de la DD ha sido sobre el acceso/no acceso a Internet. Sin embargo, también se han desarrollado posturas críticas acerca de esta visión sobre la temática. Dichas perspectivas han incorporado las desigualdades sociales, individuales o territoriales en la explicación del fenómeno. En concreto, sobre la dimensión territorial de la DD, Toudert (2016a) señala que “Independientemente si la brecha digital puede aludir a las barreras individuales, sociales y territoriales que limiten la accesibilidad a la red, la disponibilidad del servicios es quizá la condición inicial de toda apropiación socioterritorial de la Internet” (Toudert, 2016a).

A decir de Toudert (2015), la DD es un problema de escala geográfica. Esto debido a que el acceso básico a Internet responde de manera diferente en los diferentes niveles geográficos que se analice. Por ejemplo, el impacto que produce la falta de Internet en vecindarios es diferente si se habla del servicio de Internet en escalas metropolitanas. En este sentido, se recupera el concepto de urbanismo fragmentado de S. Graham y Marvin (2001), el cual da cuenta de las desigualdades territoriales que emergen en la era digital debido al aprovechamiento diferenciado de la tecnología a lo largo del territorio. Por ende, el acceso y uso de Internet configura el territorio y éste configura el acceso y ventajas que se obtienen de las tecnologías. En otras palabras, la tesis de la pérdida de la fricción de la distancia ante el avance tecnológico se cuestiona y debe superarse (Toudert, 2000).

Para el caso mexicano se ha avanzado en la perspectiva territorial del acceso, aprehensión y uso de Internet. Al respecto, se ha explorado la relación entre el territorio y la aprehensión y generación de contenido web en la frontera norte del país. Este estudio enfatiza la generación de contenido por medio de Internet y no solo a la disponibilidad del servicio (Toudert, 2000). Otro ejemplo de la importancia del territorio es el efecto que éste tiene en la inclusión de Internet en las escuelas. Debido a que la inaccesibilidad geográfica de las escuelas es un elemento importante para el éxito de Internet en el ámbito educativo (Toudert, 2016a). Aunado a lo anterior, en otro estudio se analiza la relación entre marginación socioterritorial y DD desde un enfoque de tamaño de ciudades (Toudert, 2015). Todos estos trabajos han recuperado y enfatizado la dimensión territorial

del fenómeno digital en el país. Sumado a lo anterior, se ha comenzado a probar la DD en México desde los usuarios, sus motivaciones y habilidades.

Otro enfoque para analizar la DD en México ha sido la Teoría Unificada de la Aceptación y Uso de la Tecnología (UTAUT por su siglas en inglés United Theory of Acceptance and Use of Technology). Desde esta perspectiva el nivel psicológico el nivel individual y psicológico es un elemento fundamental para explicar el uso de las tecnologías. Desde esta aproximación se le otorga más peso a los condicionantes individuales más que socialmente estructurales (Pick y Sarkar, 2016; Venkatesh, Morris, Davis, y Davis, 2003). Al respecto, a decir de Toudert (2016b), la limitante de explicar la desigualdad digital mediante la motivación de las personas se subsana, en cierta medida, por medio de entender las motivaciones y los niveles individuales como expresiones de la desigualdad social preexistente (van Dijk, 2012)

En la misma línea argumentativa van Dijk (2005, 2012) desarrolla un cuerpo teórico para explicar las desigualdades digitales. El centro de este enfoque radica en entender dichas desigualdades como un ensamblaje de la dimensión individual y estructural. Dicho ensamblaje se constituye circularmente acumulado. Es decir, el nivel estructural establece el estado de desigualdad originario, éste a su vez explica las disparidades a nivel individual en el acceso, para que, en última instancia la desigualdad individual sea un factor determinante de las disparidades estructurales (Pick y Sarkar, 2016). Así el modelo de Van Dijk se resume en etapas: la primera es la motivación individual, la segunda es el acceso En este apartado se encuentra la disponibilidad de infraestructura y dispositivos; las habilidades que finalmente condicionan el uso de las tecnologías.

Este modelo ha sido probado en México, de manera parcial. En 2012 Toudert (2016b) comprueba que la motivación de las personas de 6 años y más es factor determinante para el uso de TIC. Asimismo, se encuentra que las habilidades en el uso de tecnologías son acumuladas y explican el uso y no a la inversa. De esta manera, modelos teóricos de enfoque sociológico de apropiación y difusión de tecnologías son de suma importancia ante un escenario en el cual el acceso a Internet está creciendo rápidamente e invita a investigar las desigualdades digitales más profundas.

Hasta el momento se ha hablado de los aspectos generales de la desigualdad digital. Ésta se ha entendido, de manera amplia, desde los diferentes enfoques que han tratado de explicar la desigualdad en el acceso a TIC. Es importante remarcar que la DD implica ámbitos políticos y de diferenciación social que se relacionan con el poder de acceso a la dimensión digital y la generación

de contenidos en dicha esfera. Al mismo tiempo, gracias al avance en la penetración de Internet en la generalidad de la sociedad, la DD se constituyen al interior de la esfera digital. En este sentido, es de especial interés algunas de las herramientas digitales generalizadas, tal es el caso de los sitios o plataformas de redes sociodigitales. El uso de estas plataformas digitales se ha convertido en uno de los principales motores para entender cómo interactúan las vivencias cotidianas de las personas en el espacio urbano en la dimensión digital. De esta manera, al hablar de desigualdad digital en el acceso y uso de Redes Sociodigitales se podría indagar sobre los determinantes socioestructurales que condicionan el uso/no uso de redes digitales y cómo éste da muestra del poder de la distribución digital o acceso al momento de aumentar digitalmente procesos urbanos como la gentrificación. De esta manera, el poder distribuido de la dimensión digital se encuentra, en cierta medida, relacionado con la desigualdad digital existente en la realidad mexicana y de la ZMCM.

### **2.3 Perfiles de usuarios de Twitter.**

Como se ha mencionado, la desigualdad digital no solo radica en el acceso, sino en la aprehensión diferenciada de las tecnologías. En este marco, las plataformas de redes sociodigitales (RS)<sup>1</sup> son parte fundamental del uso cotidiano de los usuarios de Internet. En este sentido, las RS se han convertido en elemento articulador de movimientos sociales, termómetro político, plataformas de difusión y espacios de información y diversión. Por un lado, las RS han sido foco de atención por lo dicho anteriormente. Por otro lado, se ha reflexionado poco sobre la posible diferencia entre usuarios éstas, ya que se parte del supuesto del uso de Internet y de que éste es homogéneo. Sin embargo, poco se sabe del perfil sociodemográfico de usuarios de RS, así como sus elementos explicativos

Desde el trabajo académico, el uso de RS se ha abordado desde tres vertientes explicativas, por motivación personal-social, por complemento con otros usos de Internet y en menos medida desde factores sociodemográficos. El primero de ellos, los factores personales-sociales, son los más explorados. De manera general se ha encontrado que las RS son utilizadas por la necesidad de relacionarse con su entorno inmediato, es decir, extender la interacción social al entorno digital

---

<sup>1</sup>Se retoma el concepto de Redes Sociodigitales, cuando se refiere a los sitios de Internet como Facebook, Twitter, Instagram, LinkedIn, entre otros. Este concepto es propuesto por Winocur-Iparraguirre y Sánchez-Martínez (2015) con el objetivo de diferenciar las redes sociodigitales del concepto de Redes Sociales en los estudios sociológicos o antropológicos.

(Cheung, Chui, y Lee, 2011; Lin y Lu, 2011; Nadkarni y Hofmann, 2012) El segundo factor que se ha estudiado es la correlación existente entre la experiencia en otros usos de Internet y las RS. Se ha encontrado que tener experiencia en la producción de contenidos, como blogs o generación de páginas web, así como realizar compras por internet también favorece el uso de RS (Atil-Bulut y Dogan, 2017; Constantinides y Zinck-Stagno, 2011). Finalmente, el tercer enfoque ha sido caracterizar sociodemográficamente a los usuarios. En este enfoque, se ha encontrado que el uso de RS se asocia a las personas jóvenes de estrato medio-alto con grados altos de escolaridad.

Las RS comenzaron a cobrar importancia debido a que se convirtieron en factores articuladores de movimientos sociales como la primavera árabe, #Yosoy132 u *Occupy WallStreet*. Estos movimientos dejaron patente que un nuevo factor estaba emergiendo en los fenómenos sociales del siglo XXI, la redes sociodigitales. En este marco, muchos de los trabajos académicos comenzaron a analizar estas plataformas y sus consecuencias. Sin embargo, pocos han indagado los factores explicativos del uso de estas redes, ya que esto es un paso fundamental para entender la configuración de sus consecuencias sociales. De igual manera, éstas conllevan enormes beneficios, permiten estar en contacto con las personas del pasado y con personas que compartan intereses semejantes. Además, las RS están transformando nuestras vidas, nuestra manera de relacionarnos, comprar, trabajar y estar en contacto con la gente que amamos (Andrews, 2011).

En la actualidad las existen numerosos sitios y plataformas de RS. Cada uno de éstos se constituye de manera particular y se enfoca en un aspecto específico. Por ejemplo, Facebook es el sitio de RS más usado del planeta, y éste tiene como propósito brindar conexión entre amigos mediante contenido audiovisual. Por su parte YouTube es una red sociodigital enfocada al desarrollo de contenido exclusivamente audiovisual. Otro caso es WhatsApp, esta aplicación se enfoca en mensajería instantánea de persona a persona. En lo que concierne a Instagram, ésta se dirige a la participación de las personas por medio de fotografías. Cada una de las redes sociodigitales tiene un propósito particular, y por ende, su uso y consecuencias sociales y políticas son diferentes.

Un caso particular en el contexto de la RS es Twitter. Twitter es una plataforma de microblogging creada en 2006 por Jack Dorsey. En un principio, los mensajes emitidos contaban solo con 140 caracteres, ya en 2017 dicha cifra se duplicó a 280<sup>2</sup>. El sitio se caracteriza por emitir contenido que puede ser visto públicamente. Esto lo diferencia de otras RS como Facebook, las cuales requieren

---

<sup>2</sup>Dichos mensajes se les denominan *tweets* o tuits.

de reciprocidad entre sus participantes para poder ver el contenido completo. Dicho elemento lo ha caracterizado como termómetro social y político a nivel mundial (D'heer y Verdegem, 2014; Lassen y Brown, 2011; Murthy, 2012; Rodríguez, César, 2015; Sagolla, 2009).

Los estudios sobre el perfil de los usuarios de Twitter son numerosos, básicamente por dos razones. La primera razón se debe a la facilidad en la obtención de la información de esta RS. Twitter proporciona información gratuita de la emisión de tweets de los usuarios. Esto ha permitido identificar de manera cuantitativa y cualitativa el perfil de usuarios de esta RS. Los usuarios de twitter a nivel mundial se distinguen por ser jóvenes profesionistas, principalmente de habla inglesa (Mislove, Lehemann, Ahn, Onnela, y Rosenquist, 2011; Sloan y cols., 2015; Sloan y Morgan, 2015; Sloan y cols., 2013). Por otra parte, los métodos cualitativos, como la etnografía digital se han centrado en el perfil de médicos, artistas y políticos (Álvarez-Bornstein y Montesi, 2016; Rodríguez-Martín y Castillo, 2017; Rodríguez, César, 2015).

La segunda razón que explica el interés en Twitter es porque ésta ha cobrado relevancia política a nivel internacional. Por ejemplo, Twitter fue el principal motor para organizar movilizaciones sociales como la *primavera árabe* y *Occupy Wall Street*. Al mismo tiempo, el contenido emitido en esta RS ha sido catalogado como crucial en las campañas políticas de Barack Obama y Donald Trump en los comicios presidenciales de Estados Unidos en 2008 y 2016 respectivamente (AlSayyad y Guvenc, 2015; Mancera y Pano, 2013; Parnelee y Bichard, 2012).

A decir de Hargittai y Litt (2011, 2012), se sabe bastante sobre las personas que ya han adoptado la plataforma de Twitter como RS, pero poco se sabe de la etapa previa. Es decir, se desconocen los factores condicionantes para usar Twitter y en general de cualquier RS. En este marco, son pocos los estudios que analizan el perfil de social de usuarios de Twitter y, la mayoría de éstos se ha realizado en Estados Unidos. El Pew Research Center se ha interesado en identificar el perfil sociodemográfico de los usuarios de Internet y de las distintas RS (Greenwood, Perrin, y Duggan, 2016; A. Smith y Brenner, 2012; A. Smith, 2010), así como Hargittai y Litt (2011, 2012). Tanto el Pew Research Center como Hargittai y Litt llegan a conclusiones semejantes.

En la figura 2.1 se muestran los elementos básicos de los estudios del Pew Research Center y de Hargittai y Litt (2012). Cabe resaltar que estos son los únicos estudios que se han investigado sobre el perfil sociodemográfico de los usuarios de Twitter. El objetivo de comparar ambos estudios es dar cuenta de la construcción metodológica y de las fortalezas y limitaciones de los datos que

Figura 2.1: Estudios previos sobre Twitter

Estudio	Tamaño de muestra	Variabes	Método	Representatividad	Temporalidad	Población objetivo
Pew Research Center	Alrededor de 2000	<ul style="list-style-type: none"> <li>• Sexo</li> <li>• Edad</li> <li>• Ingreso</li> <li>• Grupo étnico</li> <li>• Nivel educativo</li> </ul>	Llamadas telefónicas	Nacional	2010, 2012, 2016	18 años y más
Hargittai y Litt	505	<ul style="list-style-type: none"> <li>• Sexo</li> <li>• Edad</li> <li>• Grupo étnico</li> <li>• Nivel de educación de los padres</li> <li>• Uso de Internet (años)</li> <li>• Número de lugares de acceso a Internet</li> <li>• Horas semanales</li> <li>• Rango de habilidades</li> <li>• Mensajes diarios</li> <li>• Acceso a Internet por medio de celular</li> </ul>	Encuestas en papel	Primer año de universidad	2009, 2010	18-24 años

Fuente: Elaboración propia

dichos trabajos presentan. Los dos trabajos toman como variables elementales sexo, edad, grupo racial, ingreso y nivel educativo. De igual manera, los estudios comparten temporalidad y población objetivo, personas mayores de 18 años. Se puede observar que estos trabajos tratan de explicar el uso de Twitter con base en las diferenciales sociales elementales.

Por otro lado, los estudios presentan algunas discrepancias que posibilitan reflexionar sobre los alcances y limitaciones de cada uno. Una de las diferencias más importantes es la representatividad estadística. Mientras que el estudio del Pew Research Institute arroja resultados representativos a nivel nacional, el estudio de Hargittai y Litt se limita a la población de primer año de universidad de la Universidad de Illinois en Chicago. La segunda diferencia metodológica radica en el método de obtención de la información. Hargittai y Litt (2011, 2012) argumentan que seleccionaron un método impreso para obtener información ya que así se elimina un posible sesgo en cuanto a las habilidades de los jóvenes para utilizar Internet. Por su parte, el estudio de Pew Research Institute obtiene la información vía telefónica. Al respecto, el propio instituto reconoce que esta manera de obtener información sobre temas digitales sesga la información ya que solo se está recopilando información sobre personas que ya tienen al menos un uso tecnológico en el teléfono.

Los trabajos anteriormente expuestos apuntan a perfiles determinados de los usuarios de Twitter. Ambos refuerzan la idea de que Twitter no es usada de manera homogénea entre la población.

El primer elemento a señalar es que alrededor de 20% de los usuarios de Internet son usuarios de Twitter. Enseguida, ambos estudios muestran que el sexo de los usuarios no es determinante para entender el uso de Twitter. El tercer punto esencial del perfil demuestra que Twitter es mayormente usado por personas jóvenes. Los usuarios de Twitter muestran una tendencia a tener mayor escolaridad en comparación con los que no son usuarios.

Como se ha mencionado, el objetivo de este capítulo es indagar sobre los elementos socioestructurales involucrados en la desigualdad digital en México y la ZMCM. Este objetivo se dirige a indagar en la desigualdad digital el primer elemento de la diferenciación en el poder digital distribuido inequitativamente. Para (M. Graham, Zook, y Boulton, 2013a) el poder de la distribución de elementos digitales es el primer elemento que ayuda a explicar cómo se constituye la realidad geográfica aumentada digitalmente. Esto bajo la premisa de que la construcción de dicha espacialidad digital se basa en elementos políticos de acceso diferenciado a dispositivos digitales como teléfonos móviles e Internet. En este sentido, basado en la idea de poder de la distribución para estudiar la gentrificación aumentada digitalmente, a continuación se estudian los determinantes socioestructurales que intervienen en la conformación desigual en el acceso a Internet, Redes Sociodigitales y Twitter. Lo anterior con la hipótesis de que este acceso se vincula a personas con características sociales de élite o de privilegio.

## **2.4 Definición de variables**

En esta sección se expone la delimitación de variables para conocer el estado del uso a Internet, Redes Sociodigitales y Twitter. Para lograr este objetivo se propone realizar tres modelos logísticos binomiales para dar cuenta del efecto que tienen las características sociodemográficas en las diferentes dimensiones de uso. Para este motivo se enmarca la construcción y elección de cada una de las variables en la discusión sobre DD desde una perspectiva territorial y estructural. Después de esto, se da paso a la presentación de los resultados de los modelos logísticos. Antes de definir técnicamente las variables, se expone brevemente el contexto conceptual de cada una de las variables involucradas.

#### *2.4.1 Desigualdad Digital y género*

La diferencia en el acceso y uso de TIC entre hombres y mujeres es una de las dimensiones sociales que dominan la explicación de la DD. Dicha dimensión tiene dos vertientes de análisis, una cuantitativa y la segunda desde una perspectiva crítica.

La primera dimensión de la DD basada en el género se enfoca, primordialmente, en observar la proporción de mujeres que disponen de TIC con respecto a hombres. Este indicador ha evolucionado favorablemente en años recientes. A nivel global, en la década de 1990 solo 1 de cada 3 personas que tenían acceso a TIC eran mujeres. Incluso en regiones como la europea la tasa descendía a 1 de cada 5 o en Estados Unidos, país más avanzado en estos indicadores, la proporción era 4 de cada 10 (Rodríguez, 2006). Aunque esta proporción ha mejorado sustancialmente, en 2015 el porcentaje de hombres y mujeres con acceso a Internet es prácticamente la misma (UIT, 2015). En apariencia, la DD se ha eliminado por motivos de género. Sin embargo, existen posturas críticas en torno a estos indicadores positivos.

En respuesta a lo anterior, la segunda perspectiva en torno al género y la DD es más crítica, debido a que los indicadores de igualdad en el acceso esconden elementos asociados al género. En este sentido, se ha encontrado que una explicación del aumento en el acceso a Internet por parte de las mujeres se debe al aumento de este servicio en los hogares, de esta manera las mujeres han incrementado la proporción de su acceso. Pero esto no implica que las mujeres accedan con mayor frecuencia. Es decir que, el uso es limitado y, en comparación con los hombres, las mujeres pasan menos horas en Internet. Otro ejemplo del género como factor de la DD se presenta en la predominante presencia masculina en áreas de estudio referentes a la computación, lo que refleja un sesgo por género en estas áreas (Gill-Juárez, Feliu, y Vitores, 2012; Sánchez, Ortega, y Valllovera, 2012). Más allá de las cifras de igualdad en el acceso las condicionantes sociales siguen explicando las diferencias detrás de los reportes, la doble jornada laboral que las mujeres deben realizar, las desventajas laborales o los roles culturales en torno a las tecnologías son determinantes para el uso diferencial entre hombres y mujeres (Castaño, Martín, y Vázquez, 2008).

#### *2.4.2 DD y edad*

La asociación entre rangos etáricos y la DD es otra de las asociaciones clásicas entre elementos demográficos y las disparidades en el acceso a las TIC. Esto debido que casi todos los trabajos que

versan sobre el tema identifican que la DD se asocia a la brecha generacional. Elementos como la motivación, las habilidades tecnológicas o laborales, se han explorado como explicaciones para esta dimensión.

Debido a la brecha generacional en la adopción de tecnologías, en términos generales, se observa una tendencia decreciente en el uso de tecnologías ante mayor edad de las personas. Los jóvenes entre 15 y 30 años es la población que más utiliza Internet, debido a que su porcentaje de penetración oscila entre 85% y 97%. Mientras que en adultos mayores a 60 años, el indicador se mueve entre 15% en países en desarrollo y 25% en países avanzados, como Estados Unidos, Alemania o Suiza (Friemel, 2016; Loges y Jung, 2001; Rodríguez, 2006).

La DD se explica por la edad de los usuarios en combinación con motivaciones personales. Una de las hipótesis en esta relación menciona que las personas mayores tienen menos incentivos que los jóvenes para utilizar computadoras o Internet, ya que los primeros generan resistencia al uso de nuevas tecnologías. A pesar de que los adultos mayores dispongan de los elementos físicos o de ingreso, éstos no están dispuestos a utilizar las TIC por lo dicho anteriormente. Por el otro lado, los jóvenes tienen mayor acceso a pesar de tener condiciones económicas adversas (Loges y Jung, 2001; Rodríguez, 2006) y el uso de las TIC se ha convertido en un medio para generar identidad entre jóvenes (Davis, 2013)

La pertenencia de la etapa laboral también aporta al entendimiento de la diferencia de uso entre rango etarios. Esta explicación no solo incluye la motivación o las habilidades relacionadas con el manejo de la tecnología, además se apunta a los procesos de tecnificación de las actividades económicas. La presencia de TIC en el uso cotidiano de la mayoría de empleos ha generado que las personas en edad de trabajar desarrollen habilidades relacionadas con las TIC, ya que de otra manera se enfrentan a salarios bajos y condiciones desfavorables de empleo. Esto, a su vez, condiciona el uso cotidiano de TIC entre las personas en edad de trabajar (Rodino-Colocino, 2006).

Al igual que la DD por género, la DD asociada a la edad también ha evolucionado desde los primeros años de las TIC. En este sentido, en términos relativos, la DD entre personas mayores y jóvenes se ha cerrado. A pesar de que en los jóvenes se tiene la mayor penetración y uso de TIC, el porcentaje de penetración de las TIC en personas mayores ha ido en aumento. Debido a que son personas que comenzaron a utilizar TIC en entornos laborales o en la vida cotidiana hace una década o motivados por las redes sociales o de comunicación constante se han convertido en

prioridad para las personas mayores (Friemel, 2016).

### 2.4.3 *Desigualdad digital e ingreso*

La relación entre la DD y el ingreso es la dimensión más explorada, debido a su intuitiva asociación y por las múltiples implicaciones que conlleva. Tanto las referencias teóricas como empíricas concuerdan en que el ingreso determina positivamente el acceso a las TIC, ya sea que la unidad de análisis sea nacional o individual. Asimismo, elementos como la oferta de servicios de telecomunicaciones, la mala distribución del ingreso o el abaratamiento de los dispositivos son elementos subyacentes a la relación ingreso y DD. Por ello, vale la pena mencionar rápidamente algunos de los elementos que intervienen en esta dimensión.

La asociación entre ingreso y acceso a TIC se ubica a diversas escalas de análisis. Ésta se ha estudiado a escala internacional, donde se ha encontrado que los países con mayor Producto Interno Bruto (PIB) tienen mayor nivel de penetración de las TIC. Bajo esta perspectiva han trabajado muchas de las instituciones internacionales como el Banco Mundial, la Organización para la Cooperación y el Desarrollo Económico (OCDE), Comisión Económica para América Latina (CEPAL) (CEPAL, 2013; Mundial, 2016; OCDE, 2001; Rodríguez, 2006). A nivel de los individuos la situación es semejante. Debido a que el acceso a las TIC atraviesa por una condición material, que es la posibilidad de pagar por una computadora o por el servicios de Internet, es coherente pensar que el ingreso sea uno de los factores con mayor capacidad explicativa de la DD. Dentro de este apartado se pueden hacer tres consideraciones que matizan la relación de la que se está hablando. La primera tiene que ver con la distribución inequitativa del ingreso, enseguida la oferta y los costos asociados a las TIC y finalmente el enfoque de políticas públicas que contrarrestan la DD por ingreso.

En primera instancia, aunque el poder adquisitivo de las personas o la cantidad de riqueza generada por un país son determinantes para la DD, también es necesario mencionar que la distribución de ese ingreso es un matiz que debe considerarse. Se ha propuesto que a mayor desigualdad en el ingreso en el país la DD será mayor (Mariscal, 2005; Zhang, 2013).

El segundo matiz en torno al ingreso se refiere a la oferta de TIC. Por un lado, el precio de los dispositivos vinculados a las telecomunicaciones, tanto teléfonos celulares como computadoras personales y *tablets* ha permitido que la proporción de población que dispone de estos medios se

incremente. Por el otro lado, el precio de servicios de proveedores de Internet. En este aspecto, se ha analizado el impacto de la competencia económica en el sector de telecomunicación, que conlleva a la reducción de precios en el servicios y, por ende, que la disponibilidad de Internet se incrementa (Hilbert, 2010; Mariscal, 2005; Mariscal, Bonina, y Luna, 2009).

Finalmente, los subsidios o políticas públicas enfocadas a la reducción de la DD por medio del acceso gratuito a Internet han sido propuestas y analizadas por diversos autores. A pesar de que los costos de Internet se han reducido, también es cierto que la población más pobre no tiene las posibilidades monetarias para acceder a servicios de TIC por sí mismos. De esta manera, se ha pretendido disminuir el efecto negativo de la falta de recursos para el acceso a Internet (Hilbert, 2010; Mariscal, 2005). Sin embargo, se ha demostrado que el principal factor de acceso a TIC es el ingreso (Hilbert, 2010).

Hasta el momento se han expuesto las dimensiones sociodemográficas estructurales que explican la DD, éstas son cuatro, sexo, edad, escolaridad e ingreso. Con esto en mente la siguiente sección se expone el desarrollo empírico para el caso mexicano de estas dimensiones con base en fuentes de información oficial.

#### *2.4.4 Descripción metodológica de la ENDUTIH*

La Encuesta Nacional de Disponibilidad y Uso de Tecnologías de Información en los Hogares (ENDUTIH) es una encuesta cuyo “objetivo central es obtener información sobre la disponibilidad y el uso de las tecnologías de información y comunicaciones en los hogares y su utilización por los individuos de seis años y más de edad en México, para generar información estadística en el tema y apoyar la toma de decisiones en cuestión de políticas públicas; asimismo, ofrecer elementos de análisis a estudios nacionales e internacionales y público en general interesado en la materia”(INEGI, 2016a).

La ENDUTIH encuentra su antecesor en el Módulo de Disponibilidad y Uso de Tecnologías de la Información en los Hogares (MODUTIH). El MODUTIH comenzó a publicarse en 2001 hasta 2014, a partir 2015 se cambió la metodología para constituirse como encuesta, la actual ENDUTIH. El objetivo en 2001 fue recopilar información sobre el uso y disponibilidad de computadoras debido la reciente importancia del manejo de información. Este primer esfuerzo se caracterizó por enfatizar la disponibilidad de la computadora en el hogar, así como en hardware y un uso básico

de esta tecnología (INEGI, 2003). A lo largo de 14 años el MODUTIH ha ampliado su abanico de preguntas en función del nivel de penetración creciente de las tecnologías en los hogares, así como las habilidades requeridas para su uso. De esta manera, el módulo en 2014 incluyó preguntas sobre los diferentes usos de una computadora, además enfatizó el uso de Internet. Se recuperaron frecuencia de uso, lugares de acceso y principales usos de Internet, tales como compras on-line o algunos trámites bancarios (INEGI, 2014b).

El MODUTIH tuvo como objetivo obtener información del uso de tecnologías de información en los hogares. Éste se dirigía a personas de 6 años y más que fuesen habitantes regulares en las viviendas seleccionadas. El tamaño de muestra en el levantamiento de 2014 recuperó la información de 43 822 viviendas (INEGI, 2014b).

El MODUTIH presentaba algunas limitaciones para el estudio a profundidad del uso de tecnologías en el país. La primera era la cobertura geográfica, ya que este módulo solo ofrecía resultados estadísticamente confiables a nivel nacional y entidad federativa. Se privilegió el análisis a nivel del usuario y de hogar y de esta manera, la dimensión geográfica del fenómeno no se podía recuperar con este módulo. La segunda limitante del MODUTIH era alcance en la caracterización sociodemográfica de los residentes y el uso específico de Internet de los usuarios. Estas limitaciones se han visto subsanadas, en cierta medida, a partir de 2015 cuando el MODUTIH se convirtió en encuesta.

La ENDUTIH, conservó varios rasgos metodológicos que se utilizaron en el MODUTIH. La primera es su temporalidad, esta encuesta se levanta anualmente. La segunda característica es la población objetivo, ya que ésta continua siendo aquellas personas de 6 años y más que habitualmente residen en las viviendas seleccionadas. Sin embargo, la ENDUTIH se distingue de su predecesor en dos puntos fundamentales. Por un lado, la ENDUTIH incrementó su representatividad estadística. Esto permite que a partir de 2015 se pueden obtener resultados representativos para el país, estados y zonas metropolitanas. En la actualidad, la ENDUTIH cuenta con una serie histórica desde 2015 hasta 2018. Para lograr esto, el tamaño de viviendas seleccionadas aumentó a 134 079 viviendas. Aunado a esto, la encuesta es representativa para 49 ciudades mexicanas. Para ello, se realizaron sobremuestras en las tres ciudades más grandes del país: Zona Metropolitana de la Ciudad de México, Guadalajara y Monterrey. El tamaño de esta sobremuestra fue de 5 000 viviendas para la primera y 3 000 viviendas para la segunda y tercera, mientras que para las demás 46 ciudades la muestra

fue de 2 000 viviendas. Así, la ENDUTIH recupera, en cierta medida, la dimensión geográfica y urbana del acceso y uso de las tecnologías.

En segunda instancia, “a partir de la ENDUTIH 2015, la información es proporcionada por un informante directo dentro del hogar, el cual es seleccionado aleatoriamente, y describe su propia experiencia en el uso de TIC a diferencia del MODUTIH, en donde un solo miembro proporcionaba información del resto de los integrantes”(INEGI, 2016a, 2016b). Esta diferencia es fundamental para dar cuenta de habilidades específicas de cada persona. Por último, la ENDUTIH permite caracterizar a los hogares y los usuarios de tecnologías y presenta información acerca de las características de la vivienda, hogares y usuarios, para de esta manera conocer el contexto de vivienda y hogar de cada persona usuaria de tecnologías.

En el presente trabajo se ha utilizado la ENDUTIH 2018. La encuesta recuperó información a nivel nacional sobre viviendas (130 928), hogares (133 745) y habitantes de hogares (473 606). Para el caso la ZMCM el tamaño de muestra fue de 4 459 viviendas, 4 504 hogares y 15 612 habitantes de hogares. Como se ha mencionado previamente, la encuesta elige aleatoriamente a un miembro del hogar para conocer su experiencia en el uso de Internet y tecnologías. En consecuencia, se tienen muestras sobre usos y experiencias específicas de las tecnologías igual al tamaño de muestra de hogares. Es decir, a nivel nacional la ENDUTIH permite conocer a profundidad la experiencia de usuario o no usuarios de 133 745 personas y 4 504 para la ZMCM. Con el fin de conocer el estatus de uso de Internet, Redes Sociodigitales y Twitter en la ZMCM, se recuperan ambas dimensiones geográficas y, se realiza un análisis a nivel de individuo.

De los habitantes seleccionados se tienen tres grandes secciones de información de tres principales tecnologías. Las tres principales tecnologías son computadora, Internet y teléfono móvil. La primera de las secciones se concentra en el acceso o no acceso de alguna de las tres tecnologías, es decir, si han usado dentro o fuera del hogar computadora o internet. La segunda sección se enfoca a los lugares y tipo de dispositivos para acceder a las tecnologías. Por ejemplo, si lo usuarios tienen disponibilidad de uso en el hogar, escuela, trabajo, lugar público entre otros. En cuanto al tipo dispositivo se tienen computadora de escritorio, laptop, teléfono móvil, consola de videojuegos o en el caso del teléfono móvil si éste es *smartphone* o analógico. En la tercera sección se recuperan los usos y experiencias en el uso de tecnologías. Para el caso de computadora, si ésta se utiliza para procesamiento de datos o texto, consulta de información, correo electrónico y si el uso es para fines

laborales, escolares entre otros. En cuanto a Internet, es la sección más extensa, ya que da cuenta de los principales usos y actividades en la web, redes sociodigitales, entretenimiento entre muchas otras. En cuanto al teléfono móvil, si éste es de prepago o plan tarifario y en dado caso de tratarse de *smartphone* si se accede a Internet por medio de datos móviles o por medio de red inalámbrica. En resumen, la ENDUTIH permite extraer información acerca de la desigualdad digital en México, más allá de la sola disponibilidad de tecnologías, también posibilita conocer las capacidades y los múltiples fines que se les pueden atribuir.

La ENDUTIH es una herramienta útil para conocer el estado de la DD en México. La DD atraviesa por diferentes niveles de selectividad. El primer paso que una persona debe cubrir para superar la DD es poder tener acceso a Internet. Una vez que se ha accedido a Internet, éste ofrece diversos usos, entre ellos, ser usuario de redes sociodigitales, como Facebook, Twitter, etc. En este orden, no todos los usuarios de redes sociodigitales son usuarios de la mismas. Este orden de acceso es el que se recupera para entender la DD, tanto en el país como en la ZMCM (ver cuadro 2.1).

Cuadro 2.1: Variables para el estudio de Twitter

<b>Variables</b>	<b>Tipo</b>	<b>Rango</b>
<b>Dependientes</b>		
Usuarios de Internet	Dicotómica	[0, 1]
Usuarios de redes sociales	Dicotómica	[0, 1]
Usuarios de Twitter	Dicotómica	[0, 1]
<b>Independientes</b>		
<b>Características del hogar</b>		
Estrato socioeconómico	Categórica	[1, 4]
Disponibilidad de internet en el hogar	Dicotómica	[0, 1]
Tasa de dependencia	Continua	[0, 1]
<b>Características individuales</b>		
Sexo	Categórica	[0, 1]
Edad	Categórica	[1, 9]
Escolaridad acumulada	Categórica	[1, 4]
Número de dispositivos (acceso)	Continua	[1, 6]
Gasto mensual en teléfono móvil	Continua	[10, 10000]

Fuente: elaboración propia

A continuación se detalla cada una de las variables propuestas en la cuadro anterior. En primera instancia, las variables independientes

Usuarios de Internet:

$$ui_i = \begin{cases} 1 & \text{si es usuario de Internet} \\ 0 & \text{si no es usuario de Internet} \end{cases} \quad (2.1)$$

Usuarios de Redes Sociodigitales:

$$urs_i = \begin{cases} 1 & \text{si } ui_i = 1 \text{ y es usuario de alguna red sociodigital} \\ 0 & \text{si } ui_i = 1 \text{ y no es usuario de alguna red sociodigital} \end{cases} \quad (2.2)$$

Usuarios de Twitter:

$$ut_i = \begin{cases} 1 & \text{si } ui_i = 1 \text{ y } urs_i = 1 \text{ y es usuario de Twitter} \\ 0 & \text{si } ui_i = 1 \text{ y } urs_i = 1 \text{ y no es usuario de Twitter} \end{cases} \quad (2.3)$$

Por medio de esta ruta se puede observar el cambio en el perfil de usuario según sus características sociodemográficas. Una vez definidas las variables dependientes, a continuación se presentan las definiciones y construcción de las variables explicativas o dimensiones sociales. Estas variables tratan de recuperar los niveles, que a decir de van Dijk (2005), influyen en la explicación de la DD, estos son, la dimensión individual y estructural, así como sus habilidades e intereses para el uso de las tecnologías digitales. De esta manera, a continuación se definen las variables que intervienen a nivel individual en la DD.

Sexo:

$$s_i = \begin{cases} 1 & \text{si mujer} \\ 0 & \text{si hombre} \end{cases} \quad (2.4)$$

Edad:

$$e_i = \begin{cases} 1 & \text{si 6 - 14 años} \\ 2 & \text{si 15 - 24 años} \\ 3 & \text{si 25 - 34 años} \\ 4 & \text{si 35 - 44 años} \\ 5 & \text{si 45 - 54 años} \\ 6 & \text{si 55 - 64 años} \\ 7 & \text{si 65 y años} \end{cases} \quad (2.5)$$

Nivel educativo:

$$edu_i = \begin{cases} 1 & \text{si primaria o menos} \\ 2 & \text{si secundaria} \\ 3 & \text{si medio superior} \\ 4 & \text{si superior y más} \end{cases} \quad (2.6)$$

Ahora se presenta la construcción de las variables independientes a nivel de hogar, las cuales dan cuenta del nivel estructural de la DD.

Tasa de dependencia del hogar:

$$td_{ih} = \frac{p_{0-14h} + p_{65}}{N_{nh}} \quad (2.7)$$

Donde:

$td_{ih}$ =Tasa de dependencia del hogar  $h$  del individuo  $i$

$p_{0-14h}$ =Total de personas menores de 14 años en el hogar  $h$

$p_{65h}$ =Total de personas de 65 y más años en el hogar  $h$

$N_{nh}$ =Total de personas en el hogar  $h$

La tasa de dependencia indica la proporción de habitantes que dependen de personas económicamente activas. Este indicador puede tomar valores de 0 si no existe algún habitante en los rangos etarios señalados y puede ser 1 si todos los habitantes son menores de 14 y mayores a 65.

Disponibilidad de Internet en el hogar:

$$I_h = \begin{cases} 1 & \text{si dispone de internet en el hogar} \\ 0 & \text{no dispone de internet en el hogar} \end{cases} \quad (2.8)$$

Finalmente, se propone construir un índice de estrato social. Como se ha mencionado anteriormente, aunque el estudio de las desigualdades digitales ha reconocido otros elementos para el uso de tecnologías, tales como la motivación, y uso de otras tecnologías, lo cierto es que las condiciones socioeconómicas de origen conservan un papel importante al momento de explicar las diferencias en el acceso. A la luz de esto, la construcción de una variable de estratificación socioeconómica de los usuarios es indispensable. Sin embargo, diferenciar a los usuarios enfrenta un reto en términos de la fuente de información, ya que la ENDUTIH no dispone de una pregunta

sobre ingreso del hogar o usuario. Por el otro lado, la construcción de estratificación responde a la necesidad de diferenciar a los usuarios de acuerdo a su entorno inmediato, el hogar. Ya que en éste es el lugar donde se cubren las necesidades básicas de protección, servicios y educación. Además, otro elemento a considerar es que la variable ingreso es más volátil temporalmente, mientras que las características del hogar son de mayor constancia a lo largo del tiempo (Echarri-Cánovas, 2008b).

De esta manera se construye un indicador de estrato socioeconómico basado en la propuesta de (Echarri-Cánovas, 2008b, 2008a). El indicador consta de tres dimensiones. La primera responde a la calidad de la vivienda, principalmente tipo de piso y disponibilidad de servicios públicos. La segunda dimensión refiere a la educación de los habitantes en el hogar. El tercer aspecto es la condición de actividad. Finalmente, una vez construidas las tres dimensiones del indicador se aplica el método de análisis factorial para sintetizar las tres dimensiones.

Cuadro 2.2: Variables relacionadas con el entorno de la vivienda

<b>Variables</b>	<b>Codificación ENDUTIH</b>	<b>Código</b>
Vivienda con piso de tierra	P1_1	$v_1$
Agua disponible dentro de la vivienda	P1_2	$v_2$
Drenaje conectado a la calle	P1_3	$v_3$
Energía eléctrica	P1_4	$v_4$
Disponibilidad de bienes	P1_5_1 + P1_5_2	$v_5$
Disponibilidad de refrigerador	P1_5_1	
Disponibilidad de lavadora	P1_5_2	

Fuente: elaboración propia

Dichas variables se codificaron de la siguiente manera

$$v_1 = \begin{cases} 1 & \text{si la vivienda tiene piso diferente de tierra} \\ 0 & \text{si la vivienda tiene piso de tierra} \end{cases} \quad (2.9)$$

$$v_2 = \begin{cases} 1 & \text{si dispone de agua potable dentro de la vivienda} \\ 0 & \text{de cualquier otro modo} \end{cases} \quad (2.10)$$

$$v_3 = \begin{cases} 1 & \text{si dispone de drenaje conectado a la calle} \\ 0 & \text{de cualquier otro modo} \end{cases} \quad (2.11)$$

$$v_4 = \begin{cases} 1 & \text{dispone de energía eléctrica} \\ 0 & \text{de cualquier otro modo} \end{cases} \quad (2.12)$$

$$v_5 = \begin{cases} 1 & \text{si dispone de refrigerador y lavadora} \\ 0,5 & \text{si dispone de alguno de los dos} \\ 0 & \text{si no dispone de bienes} \end{cases} \quad (2.13)$$

Basado en estas definiciones se construye una variable  $iv$ :

$$iv = \frac{\sum_{i=1}^5 iv_i}{5} \quad (2.14)$$

La segunda dimensión del índice refiere a la escolaridad de los individuos. Para la construcción de esta categoría se obtiene el número de años cursado según sexo, por medio de la combinación de dos variables de la ENDUTIH. La primera es el NIVEL de escolaridad, el cual refiere al último grado que se ha aprobado y el GRADO de escolaridad, y éste hace referencia al último grado escolar aprobado. Para obtener el número de años de escolaridad acumulados se construye una variable que da cuenta de los años de escolaridad al inicio de cada nivel escolar.

Cuadro 2.3: Años acumulados al inicio de cada nivel de escolaridad

Nivel de escolaridad	Años acumulados
Ninguno	0
Preescolar	0
Primaria	0
Secundaria	6
Estudios técnicos o comerciales con primaria terminada	6
Preparatoria o bachillerato	9
Estudios técnicos o comerciales con secundaria terminada	9
Normal básica	9
Estudios técnicos o comerciales con preparatoria	12
Normal de licenciatura	12
Licenciatura, profesional o especialidad	12
Maestría	17
Doctorado	19

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

Con base en la figura anterior, se define el total de años de escolaridad de cada individuo, según sexo:

$$e_{ij} = \text{años al empezar el nivel} + \text{último grado aprobado}$$

Una vez que se ha definido  $e_{ij}$  se procede a obtener el promedio de escolaridad según edad y sexo. Es decir, se obtiene el promedio de escolaridad para hombres o mujeres según su edad

$$ep_j = \frac{\sum_{i=1}^{N_j} e_{ij}}{N_j} \quad (2.15)$$

Donde:

$ep_j$ = Escolaridad promedio en la edad  $j$

$e_{ij}$ =Escolaridad en años de la persona  $i$  en la edad  $j$

$N_j$ = Total de personas en la edad  $j$

Basado en esta escolaridad promedio por edad, se estima la escolaridad estándar para cada edad y cada sexo, ésta se define como:

$$estd_j = \frac{ep_{j-1} + ep_j + ep_{j+1}}{3} \quad (2.16)$$

De esta manera se obtiene un nivel de escolaridad estándar para cada edad y sexo. Una vez que se obtuvo la  $estd_j$  se obtiene la escolaridad relativa para cada persona.

$$er_{ij} = \frac{e_{ij} - estd_j}{estd_j} \quad (2.17)$$

El valor mínimo de  $er_{ij}$  es de -1, cuando la persona  $i$  en el cohorte etario  $j$  no haya realizado algún año de escolaridad. Mientras que  $er_{ij}$  toma el valor de 0 cuando la persona  $i$  ha estudiado la misma cantidad de años que la media para el mismo rango de edad. Finalmente,  $er_{ij}$  será mayor a 1 si la persona  $i$  estudió más años que el promedio de su edad.

La tercera dimensión que Echarri-Cánovas (2008b) propone es la referente al aspecto laboral. En este caso, el autor jerarquiza la condición y posición laboral basado en el ingreso, para formar una sola variable con 11 categorías

Condición laboral  $cl_i$ :

$$cl_i = \left\{ \begin{array}{ll} 1 & \text{Buscó trabajo} \\ 2 & \text{Estudiante} \\ 3 & \text{Trabajador sin pago} \\ 4 & \text{Quehaceres del hogar} \\ 5 & \text{Incapacitado} \\ 6 & \text{No trabaja} \\ 7 & \text{Jornalero o peón} \\ 8 & \text{Empleado u obrero} \\ 9 & \text{Jubilado o pensionado} \\ 10 & \text{Trabajador por su cuenta} \\ 11 & \text{Patrón o empresario} \end{array} \right. \quad (2.18)$$

Una vez que se han construido las variables en las tres dimensiones a nivel de individuos, éstas son sometidas al método de Análisis Factorial (AF)

$$AF = f(iv, er_{ij}, cl_i) \quad (2.19)$$

Por medio del AF se obtiene un factor que sintetiza la varianza de las tres dimensiones:  $AF_i$ . una vez que se ha construido el indicador  $AF_i$  se estratifica por medio del método *Dalenius-Hodges* (Cortés y Vargas, 2011; Dalenius y Hodges, 1959; INEGI, 2010b). Este método de estratificación “permite la formación de grupos de acuerdo con la construcción de intervalos equidistantes para un número fijo de grupos. El procedimiento se resume en cuatro pasos: 1) la selección de la variable a estratificar, 2) la selección del número de estratos, 3) la determinación de la forma en que estratificar y, 4) la selección del número de casos por estrato” (Cortés y Vargas, 2011, p. 376).

El primer paso es estandarizar la variable  $AF_i$  mediante la siguiente función, la variable a estratificar:

$$X_i = \frac{AF_i - \min(AF_i)}{\max(AF_i) - \min(AF_i)} \quad (2.20)$$

El segundo paso es determinar el número de estratos  $L$ , en este caso  $L = 4$ . Como tercer etapa es la determinación de la forma de estratificar. Para este procedimiento se ordenan de manera

ascendente  $X_i$ . y se agrupan las observaciones en  $J$  clases donde:

$$J = \min(10L, n) \quad (2.21)$$

Una vez establecidas las  $J$  clases se agrupan las observaciones siguiendo la siguiente definición:

$$\liminf C_k = \min(X_i) + (k-1) \left( \frac{\max(X_i) - \min(X_i)}{J} \right) \quad (2.22)$$

$$\liminf C_k = \min(X_i) + k \left( \frac{\max(X_i) - \min(x_i)}{J} \right) \quad (2.23)$$

Con base en estos límites se obtiene la frecuencia de cada clase  $f_j (j = 1, \dots, J)$  El siguiente paso es obtener la suma de la raíz cuadrada de las frecuencias

$$C = \sum_{j=1}^J \sqrt{f_i} \quad (2.24)$$

Posteriormente se establecen los puntos de corte con base en  $L$  establecido previamente:

$$A = \frac{C}{L} \quad (2.25)$$

Finalmente se establecen los puntos de estratificación sobre la clase que tenga la suma de la raíz cuadrada de las frecuencias.

$$L_1 = A \approx \sum_{j=1}^{1j} \sqrt{f_i} \quad (2.26)$$

$$L_2 = 2A \approx \sum_{j=1}^{2j} \sqrt{f_i} \quad (2.27)$$

Así sucesivamente hasta  $C$ .

Para el presente estudio la construcción de la variable de estrato socioeconómico se recupera cuatro categorías: muy bajo, bajo, medio, alto. Por lo tanto  $L = 4$  y  $J = 40$ . Una vez construido

esta variable se categoriza de la siguiente manera:

$$es_i = \begin{cases} 1 & \text{Muy bajo} \\ 2 & \text{Bajo} \\ 3 & \text{Medio} \\ 4 & \text{Alto} \end{cases} \quad (2.28)$$

El indicador  $es_i$  categoriza a cada habitante en alguna categoría de estrato según sus características de vivienda, educativas y laborales. Sin embargo, como menciona (Echarri-Cánovas, 2008b), la unidad de análisis para estudiar la estratificación social es el hogar. Por ende y como último paso, se realiza la estratificación a nivel de hogar por medio de la siguiente expresión:

$$es_h = \text{moda}(es_{ih}) \quad (2.29)$$

Donde  $es_h$  es el valor más repetido el  $es_i$  en el hogar  $h$ . De esta manera, se propone un estrato socioeconómico a nivel de hogar. Para corroborar que dicha categorización refleja los perfiles socioeconómicos de los habitantes, se presentan los siguientes cuadros, donde se da cuenta de la distribución de las tres dimensiones en función del estrato socioeconómico.

En el cuadro 2.4 se muestra la distribución porcentual de la población según las características socioeconómicas utilizadas para la construcción del índice de estrato socioeconómico para la ZMCM. De manera general, se describe que la mayoría de la población se concentra en los estratos bajos y medios, por su parte, el estrato en el estrato muy bajo se ubican alrededor de 6% de las personas y aproximadamente 1 de cada 5 personas en el estrato alto. Esto refleja la distribución desigual de la sociedad en la ZMCM. La distribución según sexo, sigue el mismo comportamiento de la tendencia general, además de tener un distribución equitativa entre hombres y mujeres. En cuanto a las características de la vivienda, la mayor parte de viviendas con carencia de algún tipo de servicio público o con piso de tierra se encuentran en los estratos muy bajo y bajo. En lo que respecta a la provisión completa de servicios en la vivienda, las mejores condiciones se concentran en los estratos medios y altos. Esto habla directamente del entorno urbanizado de la ZMCM.

El cuadro también muestra la distribución porcentual de la ocupación laboral, según estrato. La tendencia en esta variable indica que mejor posición laboral se encuentra relacionada con

Cuadro 2.4: Distribución de las variables de construcción según estrato socioeconómico (ZMCM)

Variable	Estrato socioeconómico				Total
	Muy bajo	Bajo	Medio	Alto	
<b>Sexo</b>					
Mujer	7.56	43.17	30.21	19.07	100
Hombre	5.91	43.25	31.27	19.57	100
<b>Actividad económica</b>					
Menor de 12 años.	9.36	67.02	23.46	0.16	100
Buscó trabajo	9.96	52.69	30.35	7.00	100
Estudiante	2.89	52.27	44.24	0.60	100
Trabajador sin pago	0.00	37.80	47.81	14.40	100
Quehaceres del hogar	12.44	52.14	21.37	14.05	100
Incapacitado	28.75	40.88	12.32	18.06	100
No trabaja	18.21	42.82	21.75	17.23	100
Jornalero o Peón	37.43	49.81	10.94	1.82	100
Empleado u obrero	4.79	36.48	32.42	26.31	100
Jubilado o pensionado	2.00	19.46	22.31	56.23	100
Trabajador por su cuenta	6.18	34.34	32.86	26.62	100
Patrón o empresario	0.00	14.32	17.44	68.24	100
<b>Características de la vivienda</b>					
Vivienda con piso de tierra	62.41	28.09	9.51	0.00	100
Vivienda con piso de no tierra sin agua	27.91	69.36	2.24	0.49	100
Viviendas con todos los servicios	1.16	36.22	38.23	24.38	100
Años promedio de escolaridad	4.18	7.66	11.05	15.45	-

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

mejores estratos. Por ejemplo, las personas que no tienen un empleo remunerado (categorías como buscó trabajo, estudiantes, o trabajadores sin pago) se encuentran mayormente representados en los estratos muy bajo y bajo. En contraparte, las posiciones laborales mejor posicionadas se encuentran asociadas con los estratos altos y medios, así sucede con las categorías de jubilados, trabajadores por cuenta propia o empresarios. Por último, el nivel educativo promedio se incrementa conforme aumenta el estrato. En promedio las personas identificadas en el estrato muy bajo solo cursaron hasta el cuarto año de la educación primaria. El estrato muy bajo reporta una escolaridad promedio del primer año de secundaria y el estrato alto hasta el último año de educación media. Esto indica una relación directa entre escolaridad y estrato socioeconómico. Estos resultados son consistentes

con Echarri-Cánovas (2008b, 2008a).

Las variables anteriores dan cuenta de los factores que explican la DD. Sin embargo, dado que el interés del presente trabajo es indagar sobre las variables que explican el uso de redes sociales y Twitter se refieren a la capacidad de ser usuarios de estas dimensiones. Estos dispositivos pueden ser: computadora fija, laptop o netbook, tablet, smartphone, consola de video juegos o *smartTV*. En este sentido, se propone la variable de dispositivo de acceso. Ésta se define como:

$$ac_i = \text{número de dispositivos para acceder a Internet} \quad (2.30)$$

La naturaleza actual de la dimensión digital implica la interacción en ese terreno casi en tiempo real. De esta manera, como se verá más adelante, el principal medio de acceso a Internet en México y en la ZMCM en 2018 es el teléfono móvil o smartphone. Por ende, otro de los factores estructurales e individuales es el gasto que una persona realiza para navegar en Internet. En consecuencia, otra de las variables que se propone como explicativa en los niveles de uso de Internet es el gasto mensual en este rubro. Esta variable se define de la siguiente manera:

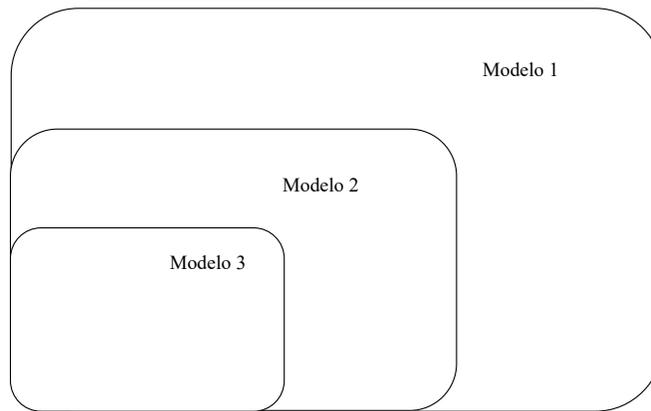
$$gmtf_i = \text{Gasto mensual en telefonía móvil (pesos mexicanos)} \quad (2.31)$$

Las variables antes definidas son utilizadas como dimensiones explicativas de la DD en el uso de Internet, redes sociodigitales y Twitter. Éstas se han construido con base en las consideraciones teóricas y metodológicas propuestas por la literatura sobre Desigualdad Digital. Ahora bien, en términos de información estadística la ENDUTIH permite analizar los subgrupos de población en función de la selectividad de uso. Es decir, la encuesta posibilita analizar en primera instancia los usuarios y no usuarios de Internet, ésta es la primera barrera a superar en la DD. Una vez que se han identificado los usuarios de Internet, se establecen aquellos que son usuarios de RS. Y finalmente aquellos usuarios de redes sociodigitales que utilizan particularmente Twitter. Esto se constituye así debido a la hipótesis de que existe una selectividad en el uso de herramientas digitales. Y, en consecuencia es necesario saber los determinantes sociodemográficos que explican los diferentes niveles de uso. En términos técnicos, a continuación se presentan las diferentes submuestras de cada nivel.

En la figura 2.2 se muestran los 3 subconjuntos para cada modelo. Al mismo tiempo, se pre-

Figura 2.2: Esquema de tamaño muestral y poblacional de la ENDUTIH, 2018

Nivel geográfico	Tamaño	Modelo 1	Modelo 2	Modelo 3
Nacional	Muestral	133 745	89 467	71 193
	Poblacional	113 003 500	74 325 379	57 849 065
ZMCM	Muestral	4 504	3 349	2 599
	Poblacional	18 669 413	14 354 211	11 183 168



Fuente: Elaboración propia con base ENDUTIH 2018, INEGI

senta la cantidad de habitantes que se entrevistaron así como su equivalencia poblacional. Además, se hace énfasis en las cantidades a nivel nacional y de zona metropolitana. Aunque solo se realizan las regresiones logísticas para la ZMCM, se agregan los totales nacionales por motivo de comparación con la zona metropolitana, para de esta manera, contextualizar los tres subconjuntos de población de interés.

Con base en todos los elementos anteriores, se plantean tres modelos logísticos. Cada uno de éstos da cuenta de tres variables dependiente que responden a cada uno de los niveles de selectividad de la DD, usuario/no usuario de Internet (modelo 1), usuarios/no usuarios de Redes Sociodigitales (modelo 2) y usuarios/no usuarios de Twitter (modelo 3). Para cada modelo, se establecen variables, tanto individuales como estructurales, que explican la selectividad de uso digital. Con base en estos elementos y la naturaleza de las variables dependientes de interés, la propuesta de esta investigación es modelar el peso de los factores sociales de la brecha digital por medio de modelos logísticos binomiales. En este sentido Agresti (2013) señala que ‘Para una respuesta binaria de la variable independiente  $Y$  y una variable explicativa  $X$ , sea  $\pi(x) = P(Y = 1|X = x) = 1 - P(Y = 0|X = x)$ . El modelo regresión logística es’ (Agresti, 2013, p.163):

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad (2.32)$$

De manera equivalente, el *logit* (momios) tiene la relación lineal:

$$\text{logit}[\pi(x)] = \log \frac{\pi(x)}{1 - \pi(x)} = \alpha + \beta x \quad (2.33)$$

De esta manera, los modelos queda definidos de la siguiente manera: Modelo 1:

$$\text{logit}[\pi(ui_i)] = \log \frac{\pi(ui_i)}{1 - \pi(ui_i)} = \alpha + \beta_1 s_i + \beta_2 e_i + \beta_3 edu_i + \beta_4 td_i + \beta_5 esh_i + I_h + e_i \quad (2.34)$$

Modelo 2:

$$\begin{aligned} \text{logit}[\pi(urs_i)] = \log \frac{\pi(urs_i)}{1 - \pi(urs_i)} = \alpha + \beta_1 s_i + \beta_2 e_i + \beta_3 edu_i + \beta_4 td_i + \beta_5 esh_i \\ + \beta_6 ac_i + \beta_7 ci_i + \beta_8 gmt f_i + e_i \end{aligned} \quad (2.35)$$

Modelo 3:

$$\begin{aligned} \text{logit}[\pi(ut_i)] = \log \frac{\pi(ut_i)}{1 - \pi(ut_i)} = \alpha + \beta_1 s_i + \beta_2 e_i + \beta_3 edu_i + \beta_4 td_i + \beta_5 esh_i \\ + \beta_6 ac_i + \beta_7 ci_i + \beta_8 gmt f_i + e_i \end{aligned} \quad (2.36)$$

Hasta este punto, se ha hablado de los factores constitutivos de la desigualdad digital y la manera de estudiarla empíricamente en el entorno urbano de la ZMCM. Por un lado, la DD de manera tanto teórica como analítica debido a que se ha reflexionado sistemáticamente sobre los determinantes sociales asociados a la diferencia en el acceso. En este sentido, la DD sirve de punto de partida para indagar sobre las diferencias sociales estructurales de la creación de diferentes poder en la dimensión digital. Por otro lado, el estudio realizado en el presente capítulo pretende abonar en la discusión sobre la desigualdad digital y la conformación de poderes digitales en dos puntos particulares. El primero de ellos es el aporte empírico en el estudio del perfil sociodemográfico de Twitter. En este punto, el presente análisis es algo que no se ha realizado a escala internacional, es decir,

investigar sobre el perfil de usuarios de Twitter en entornos urbanos, con el nivel de representatividad suficiente para ofrecer inferencias confiables. En segundo lugar, el estudio empírico aportará a entender en qué medida los perfiles sociodemográficos de uso se relacionan con la constitución del poder distribuido de lo digital. En otras palabras, a manera de hipótesis, se espera encontrar que la primera barrera para participar en la conformación de realidades geográficas aumentadas se vincule al acceso a Internet. Esta limitante ha sido bien documentada en otras ciudades (M. Graham y Zook, 2013; S. Graham, 2015). Pero este trabajo apunta que el poder distribuido de lo digital no solo se explica por la primera barrera, sino que existe un segundo nivel de conformación de poder de lo digital y que se explica por quienes utilizan determinadas redes sociodigitales.

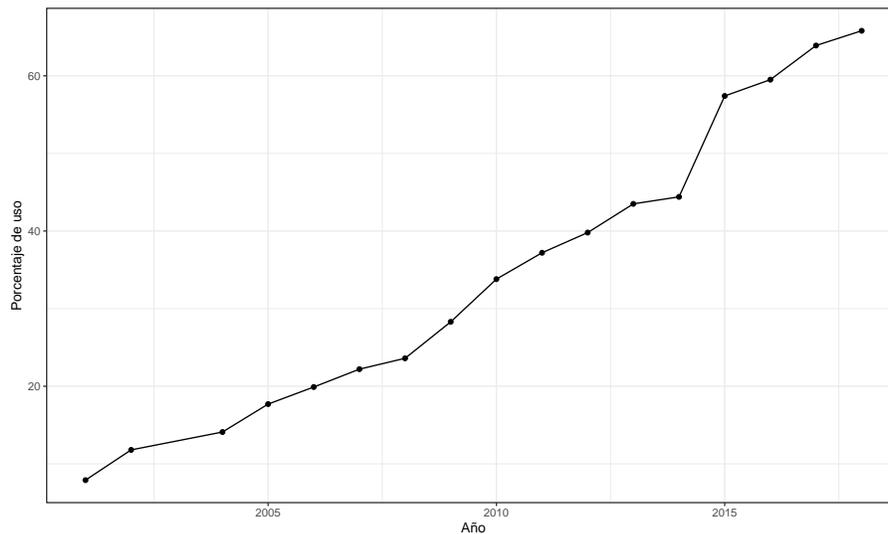
## **2.5 Resultados**

La desigualdad digital en México ha sido un tema poco estudiado. Como se ha mencionado, en el país se ha prestado interés en el avance de la penetración de TIC, principalmente, desde una perspectiva de la DD clásica. Sin embargo, es importante recordar que no solo el acceso a TIC garantiza el uso eficiente de las tecnologías. De esta manera, se debe entender la DD y la desigualdad digital como un proceso continuo, que parte de condiciones sociales y territoriales que interactúan con elementos de habilidades y motivación.

En esta sección se exponen los resultados del análisis empírico basado en la ENDUTIH con el objetivo de estudiar los determinantes sociodemográficos relacionados con la DD que ayuden a aproximar el poder digital distribuido en el uso de Twitter. Para ello, se propone avanzar en el análisis de las desigualdades digitales por medio de la caracterización de los usuarios y no usuarios de Internet, Redes Sociodigitales y Twitter. Esto con el propósito de entender que el diferencial en el acceso a distintas dimensiones de lo digital están condicionadas por las estructuras sociales. Al mismo tiempo, se pone particular énfasis en la dimensión territorial del fenómeno digital, por medio de la comparación de las tendencias nacional con respecto a lo que ocurre en la ZMCM. Esto bajo la premisa de que, el fenómeno digital también tiene dimensión territorial y, en particular, ésta se configura en el espacio urbano. Así, los resultados se dirigen a apuntalar la hipótesis de que el uso de ciertas tecnologías como las Redes Sociodigitales, y en específico Twitter, es selectivo por parte de la población basado en sus características individuales y estructurales. Dicha selectividad se acentúa en entornos urbanos, como la ZMCM.

El acceso a Internet es la dimensión digital primigenia y es la que más ha avanzado en las últimas décadas. En este sentido, la brecha en el uso de Internet ha disminuido sustancialmente en los últimos 15 años. En 2001 solo 7.9% de la población en México era usuaria de Internet mientras que en 2015 el nivel de penetración de Internet en el país alcanzó 57.4% (INEGI, 2004, 2015). Esto implica que, en promedio, los usuarios de Internet en México crecieron 15.22% anualmente. La figura 2.3 muestra la evolución del incremento de usuarios en el periodo señalado.

Figura 2.3: Población usuaria de Internet en México, 2001-2018 (Porcentajes)

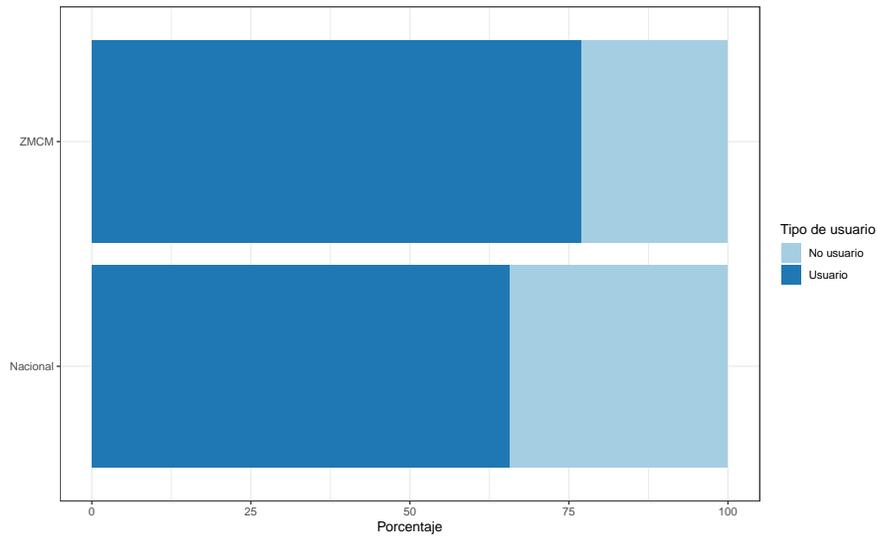


Fuente: Elaboración propia con base en (INEGI, 2004, 2006, 2007, 2008, 2009, 2010a, 2011, 2012, 2013, 2014a, 2015)

La figura 2.3, muestra la evolución a nivel nacional de los usuarios en México. Aunque la tendencia es positiva, el análisis al interior del país cambia sustancialmente, sobre todo si se distingue entre entorno rural, urbano e incluso entre ciudades, según jerarquía urbana. En este sentido es importante conocer la condición de usuarios de Internet en las zonas metropolitanas. Una limitante, para entender la evolución de este indicador a nivel metropolitano es la fuente de información, la cual se volvió representativa a escala metropolitana a partir de 2015. Sin embargo, es posible comenzar a analizar, de manera transversal, la situación del acceso a Internet en las ciudades del país.

La Zona Metropolitana de la Ciudad de México (ZMCM) tiene el mayor nivel de usuarios, proporcionalmente hablando, en comparación con el conjunto del país. La figura 2.4, muestra que en la ZMCM 76.9% de personas mayores a 6 años ha utilizado Internet, este porcentaje representa 14 354 211 personas en la zona metropolitana. Mientras que a nivel nacional son 74.3 millones los

Figura 2.4: Población usuaria de Internet en México y Zona Metropolitana de la Ciudad de México, 2018 (Porcentajes)



Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

usuarios de Internet, es decir, 65 %. Cabe resaltar que la ZMCM a pesar de ser la más importante en términos económicos y poblacionales, no es la ciudad con mayor proporción de habitantes usuarios de Internet. En particular, las ciudades que superan el umbral de 80 % de población usuaria de Internet se localizan en el norte del país. Entre éstas destacan Hermosillo, Chihuahua, Ciudad Obregón, Culiacán, Durango Ensenada, entre otras. A pesar de que la ZMCM no sea la que mayor nivel porcentual de usuarios posea, sí es la que mayor nivel de usuarios representa en el país. Tan solo una quinta parte de los usuarios de Internet en México habita en la ZMCM. En conclusión, la ZMCM es la ciudad más importante en términos de usuarios de Internet.

Las dos gráficas anteriores muestran la importancia creciente del uso de Internet en México. Por un lado, a nivel país, la tendencia creciente en el uso de esta tecnología. Por otro lado, dentro de esta tendencia, la ZMCM se posiciona como un lugar particular al momento de hablar de las DD y territorio. En este sentido, por medio de la caracterización sociodemográfica de los usuarios de Internet en la zona metropolitana permite subsanar un vacío en los trabajos sobre DD en México, desde una perspectiva espacial.

### 2.5.1 Caracterización sociodemográfica de usuarios de Internet

Como se ha expuesto, el uso de Internet se explica por las desigualdades de sexo, estrato socio-económico, educativas y de edad. Estas dimensiones han sido los elementos clásicos para explicar la DD. En primera instancia, se presentan las características sociodemográficas del uso de Internet. Al mismo tiempo, se pone énfasis en la dimensión territorial por medio del uso del Cociente de Localización (QL) para dar cuenta del peso que cada variable tiene en la ZMCM con respecto al nacional La figura 2.5 muestra este perfil para la ZMCM en 2018.

Cuadro 2.5: Características sociodemográficas y cociente de localización de usuarios/no usuarios de Internet, 2018 (ZMCM y nacional)

Variables	Delimitación espacial	Usuario	No usuario	Total	QL usuario
<b>Características del hogar</b>					
Tasa media de dependencia	ZMCM	0.23	-	-	-
	Nacional	0.23	-	-	-
Estrato socioeconómico					
Muy bajo	ZMCM	49.65	50.35	100	1.45
	Nacional	34.17	65.83	100	-
Bajo	ZMCM	71.75	28.25	100	1.08
	Nacional	66.06	33.94	100	-
Medio	ZMCM	83.88	16.12	100	0.99
	Nacional	83.94	16.06	100	-
Alto	ZMCM	86.02	13.98	100	1.00
	Nacional	85.56	14.44	100	-
<b>Características individuales</b>					
Sexo					
Hombre	ZMCM	80.28	19.72	100	1.19
	Nacional	67.37	32.63	100	-
Mujer	ZMCM	74.17	25.83	100	1.14
	Nacional	64.50	35.50	100	-
Edad					
6 - 14 años	ZMCM	82.73	17.27	100	1.24
	Nacional	66.49	33.51	100	-
15 - 24 años	ZMCM	95.96	4.04	100	1.06
	Nacional	90.02	9.98	100	-
25 - 34 años	ZMCM	93.94	6.06	100	1.14
	Nacional	82.08	17.92	100	-
35 - 44 años	ZMCM	84.42	15.58	100	1.20
	Nacional	69.93	30.07	100	-
45 - 54 años	ZMCM	74.21	25.79	100	1.27
	Nacional	58.15	41.85	100	-
55 - 64 años	ZMCM	53.23	46.77	100	1.33
	Nacional	39.83	60.17	100	-
65 y más años	ZMCM	28.55	71.45	100	1.77
	Nacional	16.68	83.32	100	-
Nivel de escolaridad					
Primaria	ZMCM	61.93	38.07	100	1.19
	Nacional	51.65	48.35	100	-
Secundaria	ZMCM	89.82	10.18	100	1.01
	Nacional	88.64	11.36	100	-
Medio Superior	ZMCM	95.29	4.71	100	1.00
	Nacional	94.66	5.34	100	-
Superior y más	ZMCM	98.55	1.45	100	1.00
	Nacional	97.83	2.17	100	-

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

En la ZMCM los factores sociodemográficos clásicos de la DD son relevantes. En general, se tiene mayor proporción de usuarios en la ZMCM según estrato socioeconómico en comparación con el nacional. En esta dimensión, sobresalen los estratos muy bajo y bajo, 50% y 70% de la población de la ciudad en estos estratos es usuaria de Internet, mientras que en las cifras nacionales solo 34% de las personas de muy bajo estrato y 60% de las de bajo declararon ser usuarias. Por su parte, los estratos medios y altos presentan la misma proporción de internautas tanto en el nivel nacional y metropolitano. Esto apuntaría a que en la ZMCM existen condiciones particulares que propician que estratos muy bajos y bajos de la población para el uso de Internet. En términos territoriales, los cocientes de localización reafirman la situación particular de usuarios de estrato muy bajo. El QL de este segmento social muestra que en la ZMCM se tiene 45% más usuarios de muy bajo estrato en comparación con el nacional. En lo que respecta a los estratos bajos, la ciudad tienen 8% usuarios de Internet, en comparación con la proporción de usuarios en el país. Finalmente, los estratos medios y altos muestran que no existen diferencias sustanciales en la proporción de usuarios en la ciudad en comparación el total poblacional.

Con respecto a las diferencias por sexo, se identifica una pequeña diferencia entre hombres y mujeres. Los hombres son más propensos a utilizar internet que las mujeres. En la ZMCM 80% de los hombres son usuarios mientras que 74% de las mujeres tienen esta condición. En la misma dirección se comporta la división por sexo en el país, 67% de los hombres son usuarios y 64% de las mujeres. El QL para sexo muestra que en la zona metropolitana se tienen, en términos proporcionales, más hombres y mujeres usuarios de Internet que en México. En la ZMCM existen 19% más hombres usuarios, mientras que las mujeres esta proporción es de 14%.

En el cuadro 2.5 se muestran las proporciones de uso según edad. En este rubro, se refuerza la hipótesis de la brecha generacional como factor para el uso de Internet. Esto se demuestra tanto para la ciudad de México como para el país, los rangos etarios que más usan internet son los jóvenes. Por grupo de edad, las personas jóvenes son las que reportan mayor proporción de uso, es decir, más del 90% de personas entre 15 y 34 años es usuaria de Internet. Esta proporción desciende conforme aumenta la edad, así, alrededor de 1 de cada cuatro personas de 65 y más años es usuaria en la ZMCM y, esta proporción es menor a nivel nacional, solo 16% de adultos mayores declararon ser usuarios. En la ciudad de México, se tiene 3% más usuarios de 55 a 64 años y 77% más usuarios de 65 y más años cuando se compara con los usuarios de la misma edad del total nacional. De nueva

cuenta, el entorno urbano condiciona que haya más usuarios de edades avanzadas y de manera absoluta exista un uso universal de los jóvenes en esta ciudad.

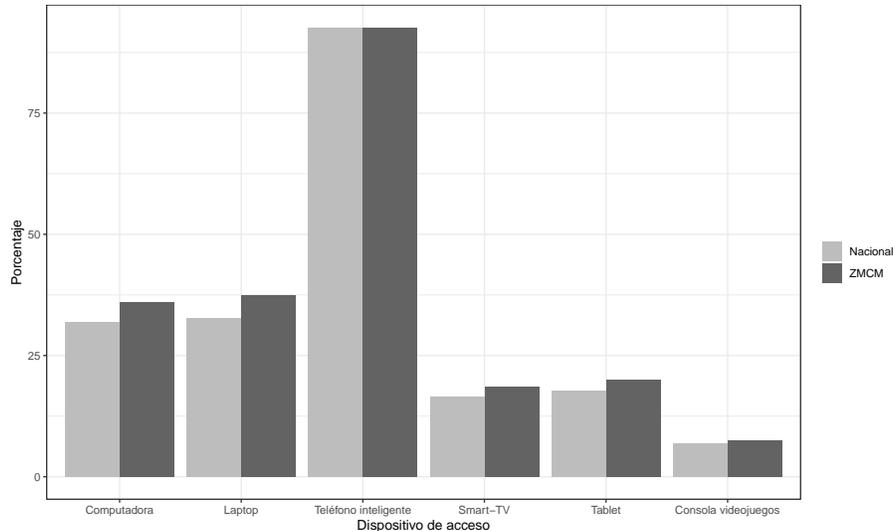
El nivel educativo es una variable que se relaciona positivamente con el porcentaje de uso de Internet. Es decir, tanto para la ZMCM como para el país, el porcentaje de usuarios incrementa conforme más educación se tiene. Las cifras muestran que para 2018 prácticamente todas las personas con algún grado de educación media o superior son internautas. En contra parte, 1 de cada 2 personas con educación básica es usuaria. En términos territoriales, la ZMCM sobresale solamente en los usuarios con educación básica ya que el cociente reporta que en la ciudad existen 19% más usuarios en comparación con el nacional. Por su parte, los niveles de secundaria, medio superior y superior se comportan proporcionalmente igual que el resto del país. Este es un hallazgo interesante puesto que la variable educacional no caracteriza territorialmente el fenómeno de la DD entre la ZMCM y México.

Con lo anterior, se puede ver que los usuarios de Internet en la ZMCM se caracterizan por dos tendencias. La primera indica que la ZMCM tiene mayor cantidad de usuarios según lo esperado por los estudios de la Brecha Digital. El uso de Internet se asocia positivamente con respecto al estrato social y al nivel educativo. Y de manera negativa en relación a la tasa de dependencia y edad. La segunda tendencia indica que la ZMCM es un espacio particular de uso de Internet. Ya que el QL para la ZMCM indica que en este espacio hay proporcionalmente más usuarios en las características menos favorables. De esta manera, se tienen más personas usuarias en estratos bajo, con mayor edad y con menos educación. Estos indicadores, llevan a plantear a la ciudad como espacio particular en la dimensión digital.

Los párrafos anteriores muestran las características que explican el ser o no usuario de Internet. Esta es la primera barrera que las personas tienen que superar para ser parte del dominio digital. En este aspecto, vale la pena mencionar que el acceso y uso de Internet se ha vuelto cada vez más sencillo en términos de dispositivos. Es decir, gracias al abaratamiento de las computadoras personales y al auge de dispositivos inteligentes. En este ámbito la figura 2.5 muestra la proporción de personas que acceden a Internet según tipo de dispositivo. La gráfica muestra que el principal medio para conectarse a Internet es el teléfono móvil o smartphone, alrededor de 80% de los usuarios de Internet lo hacen por este medio. Por su parte, las computadoras de escritorio y laptop son el segundo dispositivo con mayor cantidad de usuarios, prácticamente 1 de cada 3 usuarios utilizan

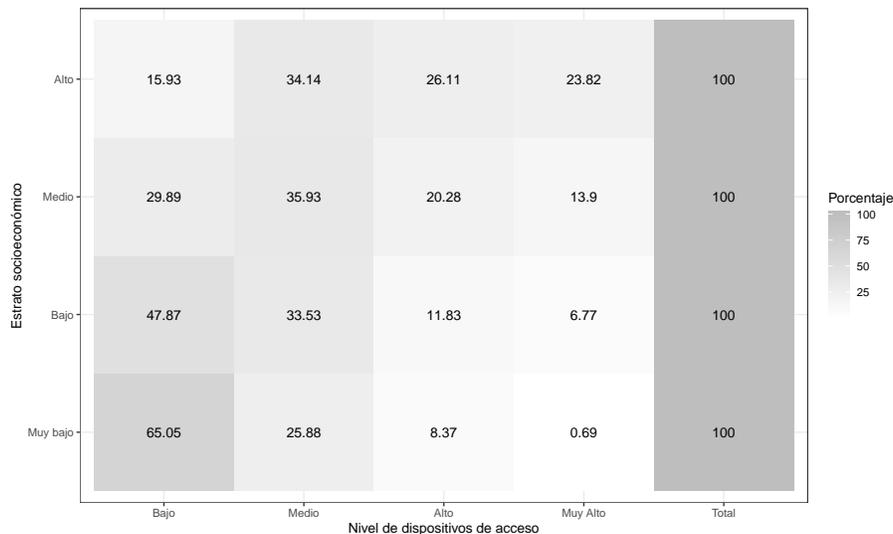
estos dispositivos para acceder a Internet. Finalmente, la menor proporción de usuarios lo hacen por medio de SmartTV, Tablet o consola de videojuegos.

Figura 2.5: Porcentaje de acceso según dispositivo, 2018



Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

Figura 2.6: Número de dispositivos y estrato socioeconómico en la ZMCM, 2018



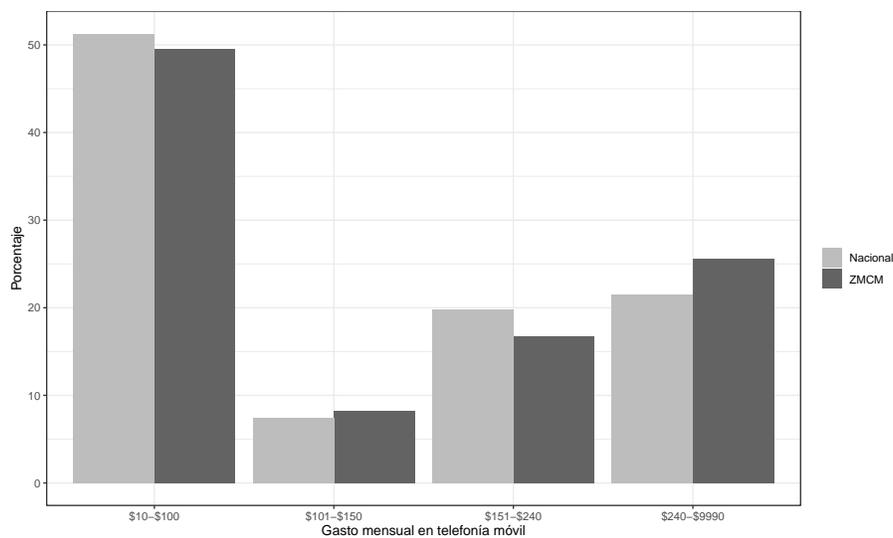
Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

Aunado a que la mayor proporción de personas accede por medio de smartphone, se debe considerar que no exclusivamente se puede acceder por medio de éste. Es decir, la personas pueden acceder por medio de smartphone y laptop. Pero la cantidad de dispositivos también se encuentra sujeta a condiciones sociodemográficas. En este sentido, la figura 2.6 muestra como se correlaciona

la cantidad de dispositivos usados y el estrato socioeconómico. Esta gráfica sugiere una relación positiva entre ambas dimensiones. Por una parte, los estratos bajos acceden a Internet, en su mayoría, solo mediante un dispositivo, primordialmente smartphone. Mientras que la mayor parte de personas en estrato alto accede con dos y más dispositivos. Esto podría sugerir que a mayor grado sociodemográfico se tiene mayor independencia para el acceso a Internet.

Otro elemento a considerar sobre el uso de Internet es el gasto que las personas realizan en este rubro. Se ha mostrado que la mayor parte de personas accede, al menos, mediante smartphone y, que éste es el único medio para estratos muy bajos y bajos. En consecuencia, la revisión del gasto en telefonía móvil es un buen indicador para caracterizar los patrones de uso de Internet. En la figura 2.7 se muestra el gasto promedio mensual de los usuarios de Internet en México y ZMCM.

Figura 2.7: Gasto mensual en telefonía móvil en México y ZMCM, 2018

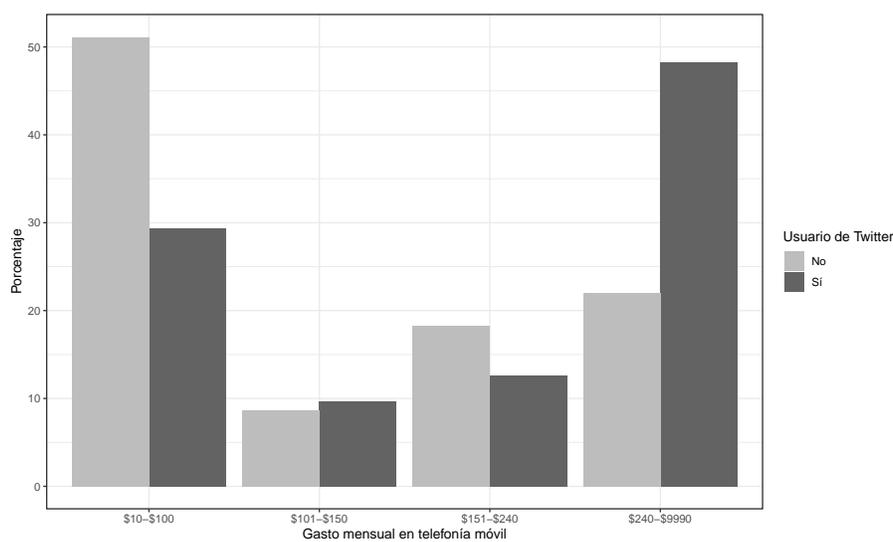


Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

La gráfica anterior muestra dos tendencias. La primera implica que tanto a nivel nacional como metropolitana la proporción de gasto en teléfono móvil es semejante. Segundo, por nivel de gasto, se observa que la mitad de la población solamente gasta entre \$10 y \$100 pesos al mes. Lo anterior sugiere relación con el estrato socioeconómico. Por su parte, alrededor de 25% de la población gasta más \$240 pesos al mes. Por otro lado, dada la naturaleza de las RS y de Twitter, particularmente, el gasto en telefonía móvil cambia con respecto al general.

Como se ha mencionado, metodológicamente, las personas tienen que superar diferentes barreras para participar de la dimensión digital. Así, dadas las distintas naturalezas de los diferentes

Figura 2.8: Gasto mensual en telefonía móvil de usuarios de Twitter en y ZMCM, 2018



Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

usos, éstas se basan en condiciones materiales previas, en este caso, la capacidad de pago por el acceso a tecnología. En este sentido, el uso de RS y Twitter se configuran como un uso particular de Internet que requiere estar siempre conectado. Por esta razón, la capacidad de pagar telefonía móvil condiciona el uso de Twitter. Esta idea se analiza en la figura 2.8, la cual muestra la proporción de usuarios de Twitter y su gasto mensual en este rubro. La gráfica permite observar que existe una relación positiva entre el gasto y ser o no usuario de Twitter. Por ejemplo, mientras que la mitad de los no usuarios de Twitter gastan entre \$10 y \$100, la misma proporción de usuarios gasta más de \$240 pesos al mes. De esta manera, se sugiere que existe, también asociación entre el gasto y el estrato socioeconómico.

### 2.5.2 Caracterización sociodemográfica de usuarios de Redes Sociodigitales

En esta subsección se analiza el perfil sociodemográfico de los usuarios de Redes Sociodigitales. Este uso es el siguiente nivel en la selectividad de uso en la dimensión digital. Para este propósito, según la figura 2.2, el universo de análisis se reduce, ya que las personas deben de superar primero la barrera de acceso a Internet para después ser usuarios de algunas RS. Para el propósito de estudiar el perfil sociodemográfico se presenta el siguiente cuadro.

Cuadro 2.6: Características sociodemográficas y cociente de localización de usuarios/no usuarios de redes sociodigitales, 2018 (ZMCM y nacional)

Variables	Delimitación espacial	Usuario	No usuario	Total	QL usuario
<b>Características del hogar</b>					
Tasa media de dependencia	ZMCM	0.21	-	-	-
	Nacional	0.23	-	-	-
Estrato socioeconómico					
Muy bajo	ZMCM	67.40	32.60	100	0.91
	Nacional	73.71	26.29	100	-
Bajo	ZMCM	76.67	23.33	100	1.00
	Nacional	76.02	23.98	100	-
Medio	ZMCM	80.70	19.30	100	0.99
	Nacional	81.51	18.49	100	-
Alto	ZMCM	78.07	21.93	100	0.98
	Nacional	79.25	20.75	100	-
<b>Características individuales</b>					
Sexo					
Hombre	ZMCM	76.74	23.26	100	1.00
	Nacional	76.53	23.47	100	-
Mujer	ZMCM	79.24	20.76	100	1.00
	Nacional	79.04	20.96	100	-
Edad					
6 - 14 años	ZMCM	44.44	55.56	100	0.98
	Nacional	45.22	54.78	100	-
15 - 24 años	ZMCM	96.16	3.84	100	1.01
	Nacional	94.74	5.26	100	-
25 - 34 años	ZMCM	92.16	7.84	100	1.02
	Nacional	89.85	10.15	100	-
35 - 44 años	ZMCM	84.05	15.95	100	1.02
	Nacional	81.83	18.17	100	-
45 - 54 años	ZMCM	73.56	26.44	100	0.99
	Nacional	74.19	25.81	100	-
55 - 64 años	ZMCM	62.33	37.67	100	0.94
	Nacional	65.83	34.17	100	-
65 y más años	ZMCM	44.79	55.21	100	0.85
	Nacional	52.64	47.36	100	-
Nivel de escolaridad					
Primaria	ZMCM	66.52	33.48	100	0.98
	Nacional	67.70	32.30	100	-
Secundaria	ZMCM	87.85	12.15	100	0.99
	Nacional	88.61	11.39	100	-
Medio Superior	ZMCM	84.49	15.51	100	0.96
	Nacional	87.82	12.18	100	-
Superior y más	ZMCM	82.25	17.75	100	0.95
	Nacional	86.39	13.61	100	-

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

El principal uso de los internautas, en México y la ZMCM, se da en los sitios de RS. En primera instancia, de la totalidad de usuarios de Internet, 98 % de ellos utiliza alguna red sociodigital. Entre los que captura la ENDUTIH se encuentran Facebook, Twitter, Instagram, LinkedIn, Snapchat y otras. De éstas, la que tienen la mayoría de usuarios, ya que casi la totalidad de las personas declaró hacer uso de está. La segunda RS más utilizada en México es Twitter. Los sitios de RS se han convertido en una de los usos prominentes en Internet, esto condiciona que las características tanto demográficas como territoriales no varíen a lo largo de las diferentes categorías de análisis, ni a través del espacio.

En el cuadro 2.6 se muestran las características sociodemográficas, así como el QL, para los usuarios no usuarios de RS en la ZMCM. El cuadro muestra, a nivel nacional y de manera general que no existe una diferenciación, en términos proporcionales, de los usuarios de RS según su estrato socioeconómico. Mientras que, para la ZMCM esta tendencia se observa solo a partir del estrato bajo en adelante, ya que entre los dos estratos más bajos existe una diferencia de casi 10% de población usuaria. Por su parte, la condiciones de sexo no condiciona el uso o no uso de RS, ya que la proporción de usuarios de RS según sexo varía entre 76% y 79%. En lo que respecta a los rango etarios, de nueva cuenta, el uso no se diferencia, con excepción de la tendencia de la brecha generacional. En el caso del nivel de escolaridad, solo el estrato muy bajo reporta menor proporción de uso. Pero, en el caso de los estratos de secundaria y más, el nivel de uso ronde entre 84% y 87%. En cuestión territorial, el QL sugiere que no existe alguna concentración de usuarios de RS independientemente de sus características sociodemográficas.

Cuadro 2.7: Características sociodemográficas de usuarios de Redes Sociodigitales en la ZMCM, 2018

VARIABLES	Facebook	Twitter	Instagram	LinkedIn
Hombre	97.73	27.27	21.23	4.95
Mujer	99.04	25.10	23.72	3.36
Edad promedio	33	32	28	38
Escolaridad promedio	11.95	13.38	12.88	15.64
Estrato sociodemográfico	Todos	Medio y Alto	Medio y Alto	Alto

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

En el cuadro anterior se presenta un resumen de las principales características de los usuarios de RS en la ZMCM. Esto con el propósito de dar cuenta de la diferenciación social entre cada una de ellas.

El caso de Twitter se distingue por ser una red sociodigital para estratos altos, principalmente adultos jóvenes y con alto grado de escolaridad que viven en hogares con pocos dependientes económicamente. Así lo demuestra la baja tasa promedio de dependencia, 0.19 dependientes en el hogar. Ahora bien, el estrato alto es donde proporcionalmente se utiliza más esta RS, 41 % de este estrato reporta ser usuario, además de que en la ciudad de México esta proporción es mayor con respecto al nacional, 34 %. Por su parte los estratos muy bajo y bajo, apenas tienen 13 % y 20 % de usuarios al interior del grupo, porcentajes muy semejantes a lo reportado nacionalmente. Finalmente, los usuarios son en promedio personas de 32 años y de grado de escolaridad alto.

En lo que se refiere a Facebook, esta red social es la que reporta el uso más ubicuo de todas. En general, casi 100 % de personas a lo largo de todos los estratos la utilizan. Así mismo, tanto hombres como mujeres son usuarios de esta RS. En cuanto a la edad promedio, ésta se ubica en 33 años. Sin embargo, algo a destacar del uso de Facebook es la escolaridad. Tanto para la ZMCM como el país, el usuario promedio de esta RS no supera la educación media superior. Y al mismo tiempo, los no usuarios de Facebook, son personas con al menos un año de escolaridad superior. Esto indica que, Facebook es la RS por excelencia tanto en México como en la ZMCM. Lo anterior contrasta con RS más específicas como Instagram o LinkedIn.

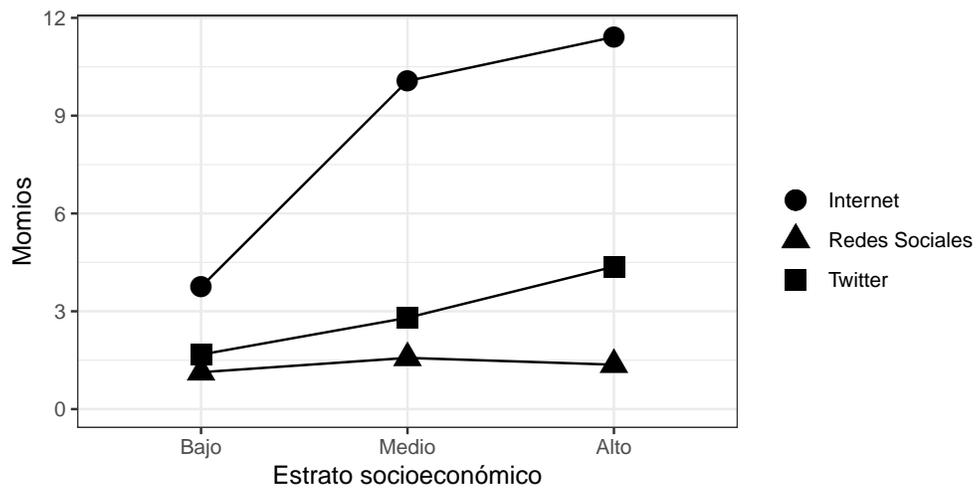
En el caso de Instagram, el perfil de usuario es parecido al perfil de Twitter. Se trata de una red que es usada por jóvenes con alta escolaridad de estratos medios y altos con bajas tasas de dependencia en el hogar. En términos de la estratificación, entre 20 % y 30 % de cada estrato desde el estrato bajo hasta el alto es usuaria de Instagram. A diferencia de las otras RS, las mujeres utilizan más esta red social. Otra peculiaridad de Instagram es la edad de sus usuarios, en promedio el usuario tiene 28 años en la ZMCM y 25 en el país. Finalmente, en promedio en Instagram se tienen usuarios con al menos un año en educación superior.

Por último, LinkedIn es una red que se enfoca a construir redes sociales entre personas con perfil profesional y de servicios especializados. Este sesgo se refleja en los años promedio de escolaridad, aproximadamente 15 años tanto para la ZMCM como el nacional. Al mismo tiempo, es la RS con usuarios de mayor edad de todas, 37 años en la ciudad de México y 32 en el país. El perfil especializado de la RS condiciona que sea la red menos utilizada. En LinkedIn se manifiesta una brecha de género, solo 3 % de las mujeres lo utiliza, mientras que en los hombres esta cifra asciende a cerca de 5 %. En cuanto al estrato socioeconómico, los principales usuarios son las personas de

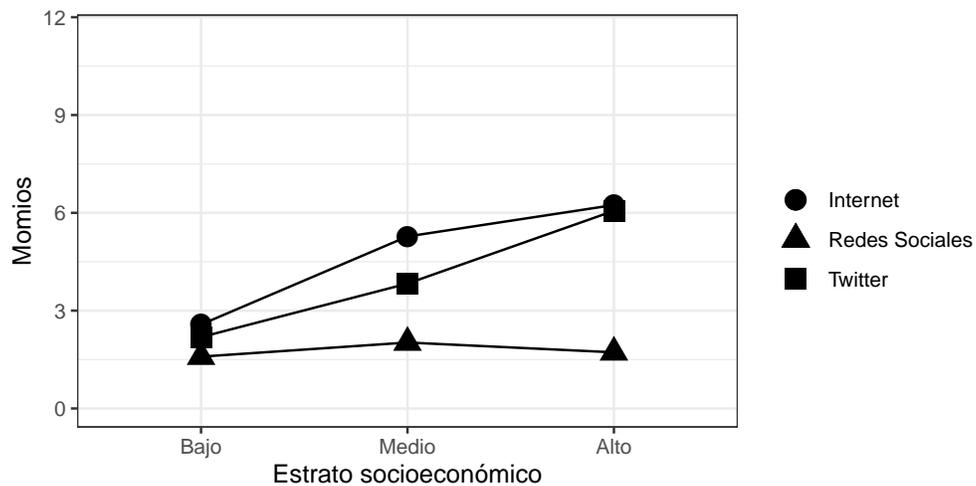
estratos altos, 14% de personas de estrato alto son usuarios, mientras que en estratos bajos su uso es mínimo. Es notable que, LinkedIn al ser la RS más específica, sea la red social en la cual se marquen con mayor claridad las diferencias sociales en el uso.

Para finalizar la caracterización de las personas según su nivel de selectividad a continuación se presentan la razón de momios de ser usuario o no usuarios de los tres niveles de acceso que se han analizado.

Figura 2.9: Proporción de uso de Twitter según estrato económico (base estrato muy bajo)



(a) Nacional



(b) ZMCM

Fuente: Elaboración propia con base en ENDUTIH, 2018; INEGI

La figura anterior muestra la razón de momios de ser usuario de internet, redes sociodigitales y Twitter en comparación con no ser usuario, esta proporción toma como punto de referencia el

estrato muy bajo <sup>3</sup>

En primera instancia, la figura 2.9a confirma, de nueva cuenta, que la barrera más discriminante en la dimensión digital es la de acceso a Internet. Esto se observa en el hecho de que los estratos bajos tienen 3 veces más oportunidades de acceder a Internet que los estratos muy bajos. Enseguida, los estratos medios tienen un salto cualitativo en las probabilidades de acceso ya que éstas son de 10 veces más con respecto al estrato más bajo. En lo que respecta a los estratos más altos, éstos muestran cerca de 12 veces más probabilidades de ser usuarios con respecto a los muy bajos. Una vez que se ha superado la primera barrera en el acceso, las probabilidades de acceso se comportan de una manera menos discriminatoria, aunque se continúa con la segmentación según estrato en el uso. En lo que respecta al nivel nacional, el uso de RS es el uso más homogéneo en Internet. Según la gráfica 2.9a muestra que a lo largo de todos los estratos se tienen las mismas probabilidades de ser usuario. Esto se explica por el hecho de que 98 % de los usuarios de Internet, usa alguna RS. Sin embargo, como se ha mencionado anteriormente, el uso mayoritario se concentra en Facebook. En lo que respecta a Twitter, en México, la probabilidad de ser usuario incrementa conforme aumenta el estrato, llegando a ser de 4 veces más en estratos altos con respecto al estrato muy bajo.

En segunda instancia, la gráfica 2.9b refuerza la idea de selectividad digital. De manera semejante a la escala nacional, en la ZMCM la probabilidad de ser usuarios de Internet incrementa con el estrato, para estratos bajos se tiene 3 veces más, para estratos medios es de 5.5 veces más y para estratos altos es de 6 veces más con respecto al estrato de referencia. De igual manera, las probabilidades de usar RS en la ZMCM son las mismas a lo largo de todos los estratos sociodemográficos. Finalmente, en particular, los momios de uso de Twitter en la ZMCM muestran un comportamiento semejante al reportado por el acceso a Internet. Es decir, el uso de Twitter es selectivo según estrato socioeconómico y, éste es más propenso en estratos más altos.

---

<sup>3</sup>La razón de momios de construye:

$$RM_i = \frac{\frac{u_i}{nu_i}}{\frac{u_{mb}}{nu_{mb}}}$$

Donde:  $u_i$  = usuarios en el estrato  $i$

$nu_i$  = no usuarios del estrato  $i$

$u_{mb}$  = usuarios en el estrato muy bajo

$nu_{mb}$  = no usuarios del estrato muy bajo

### 2.5.3 Modelos explicativos

Como se explicó más arriba, se proponen tres ajustes de tipo binomial para entender el comportamiento de todas las variables antes expuestas. El propósito central de estos modelos es obtener información sobre el peso de los factores estructurales, tales como el estrato socioeconómico en conjunción con los niveles individuales, tales como el sexo, edad o escolaridad para explicar el uso de las dimensiones digitales en los diferentes niveles de su selectividad.

El cuadro 2.8 presenta los resultados de los ajustes binomiales para los tres niveles de selectividad en el uso de la dimensión digital. Estos modelos corresponden a los resultados de las ecuación planteadas en las expresiones 2.34, 2.35, 2.36. El primer modelo explica la probabilidad de ser usuario de Internet. El segundo modelo explica la probabilidad de ser usuario de alguna RS, una vez que ya es usuario de Internet. Y, el tercer ajuste explica la probabilidad de ser usuario de Twitter, una vez que ya es usuario de alguna RS. De esta manera se puede observar las diferencias explicativas de las variables independientes ante cada nivel de selectividad.

El modelo 1 hace referencia a la ecuación 2.34 la cual establece como variable dependiente ser usuario o no de Internet. El primer punto a resaltar es la representatividad estadística de los coeficientes asociados a las variables explicativas. Este modelo muestra que todas las variables influyen estadísticamente en la probabilidad de ser usuario de Internet, excepto si se pertenece al estrato bajo, así como el nivel de la tasa de dependencia del hogar. Enseguida, las variables que afectan positivamente la probabilidad de ser usuario son el estrato sociodemográfico y el nivel de escolaridad. En otras palabras, al aumentar el estrato sociodemográfico o el nivel de escolaridad se espera mayor probabilidad de usar Internet. Por otro lado, el sexo y la edad se relacionan negativamente con el uso de Internet. Es decir, al aumentar la edad la probabilidad de uso disminuye, asimismo ante la condición de mujer.

Cuadro 2.8: Resultado de los modelo logísticos de uso de Internet, Redesociodigitales y Twitter en la ZMCM, 2018

Variables	Modelo 1			Modelo 2			Modelo 3		
	$\beta$	$exp(\beta)$		$\beta$	$exp(\beta)$		$\beta$	$exp(\beta)$	
Constante	1.19	3.30	***	2.25	9.56	***	-3.00	0.04	***
<b>Características del hogar</b>									
Tasa de dependencia	-0.07			-0.026	0.76		-0.0002	0.99	
Estrato socioeconómico									
Muy bajo		(Ref)			(Ref)			(Ref)	
Bajo	0.48	1.62	.	0.017	1.18		0.03	1.47	
Medio	1.03	2.80	***	0.023	1.26		0.05	1.80	
Alto	1.03	3.71	***	0.013	1.14		0.06	1.93	
Internet en hogar	1.86	6.43	***	-	-		-	-	
<b>Características individuales</b>									
Sexo									
Hombre		(Ref)			(Ref)			(Ref)	
Mujer	-0.37	0.68	**	0.026	1.29	*	-0.003	0.73	*
Edad									
6 - 14 años	-1.01	0.36	***	-3.025	0.04	***	-0.04	0.64	
15 - 24 años		(Ref)			(Ref)			(Ref)	
25 - 34 años	-0.63	0.52	*	-1.05	0.35	***	-0.003	0.97	
35 - 44 años	-1.61	0.19	***	-1.76	0.17	***	-0.04	0.62	*
45 - 54 años	-2.73	0.06	***	-2.42	0.08	***	-0.03	0.67	*
55 - 64 años	-3.71	0.02	***	-2.92	0.05	***	-1.22	0.29	***
65 y más años	-4.94	0.007	***	-3.37	0.03	***	-1.87	0.15	***
Nivel de escolaridad									
Primaria		(Ref)			(Ref)			(Ref)	
Secundaria	1.11	3.04	***	0.014	1.15		0.03	1.40	
Media superior	2.20	9.10	***	0.017	1.19		1.00	2.72	***
Superior y más	3.09	22.11	***	0.006	1.06		1.88	6.60	***
Números de dispositivos	-	-		4.7	1.60	***	0.04	1.51	***
Gasto en telefonía	-	-		-7.7	0.99	**	0.000006	1.00	***

Fuente: elaboración propia con base en ENDUTIH 2018, INEGI

Nota de niveles de significancia: \*\*\*  $p > 0,0001$ , \*\*  $p > 0,01$ , \*  $p > 0,05$ , ·  $p > 0,1$

Ahora bien, en términos de probabilidad, se presenta el antilogaritmo de los coeficientes para analizar el impacto en la probabilidad de cada variable. En lo que respecta al estrato sociodemográfico, el ajuste muestra que las personas de estrato medio tienen 2.8 veces más oportunidades de utilizar Internet, con respecto al estrato de referencia, estrato muy bajo. Al mismo tiempo, personas que pertenecen al estrato alto reportan 3.71 veces más oportunidad de ser usuarios frente al estrato de referencia. Por su parte, si la persona cuenta con Internet en el hogar, las probabilidades de uso se incrementan en 600.43%. En lo que respecta al nivel de escolaridad, personas con escolaridad profesional tienen 22.11 veces más oportunidades de usar internet frente a personas con escolaridad básica.

Por el contrario, las mujeres tienen 32% menos probabilidad de ser usuarias de Internet en comparación con los hombres. De igual manera, al incrementar la edad la probabilidad de ser usuario disminuye. El rango etario que presenta menor decremento es la cohorte de 25 a 34 años, ya que personas en esta edad tienen 48% menos probabilidades de ser usuario con respecto a jóvenes de 15 a 24 años, rango de referencia. En contraposición, las personas de 65 años y más tienen 99.3% menos probabilidades de ser usuarios de Internet.

A manera de resumen, los elementos sociodemográficos que explican la primera barrera para participar de la dimensión digital se reafirman en la ZMCM. La brecha de género, generacional y socioeconómica son dimensiones de relevancia estadística para explicar la primera diferencia en el acceso a Internet en la ciudad. Dentro de este análisis, se puede rescatar que la dimensión que afecta más para ser o no usuario de Internet es la escolaridad, ya que ante el mayor nivel de escolaridad, profesional, se tiene el mayor nivel de probabilidad para ser usuario.

El modelo 2 representa los resultados de la ecuación 2.35, la cual explica la probabilidad de ser usuario o no de alguna Red Sociodigital. En este modelo se utilizan las mismas variables sociodemográficas del modelo 1 y se añaden dos más referentes al uso y aprovechamiento de tecnologías digitales. En primera instancia, el modelo 2 muestra que las variables que son estadísticamente significativas son la edad y sexo, en lo que refiere a la dimensión sociodemográfica, mientras que los números de dispositivos y el gasto en telefonía móvil resultan relevantes para la explicación de uso. En este punto llama la atención la nula importancia estadística de las variables de estrato y escolaridad, ya que éstas fueron relevantes en el ajuste del modelo 1. Esto puede sugerir que superar la primera barrera digital está íntimamente vinculado a la estructura socioeconómica. Una vez que se

ha superado dicha barrera no importa estrato, ni nivel de escolaridad.

Dentro de las variables con coeficientes significativos se tiene que el sexo y el número de dispositivos para conectarse a Internet afectan positivamente la probabilidad de uso. En contra parte, la edad y el gasto en telefonía móvil reducen las probabilidades. En este punto, se resalta que las mujeres tienen 1.29 veces más probabilidades de ser usuarias de RS con respecto a los hombres. En este mismo hilo, por cada dispositivo móvil que se disponga para usar internet la probabilidades aumentan 1.60 veces. Es decir, una persona que acceda a Internet por medio de un teléfono inteligente, computadora portátil y tablet tienen 3 veces más probabilidades de ser usuario de RS en comparación con la persona que solo accede por medio de celular. La brecha generacional se mantiene en el uso de RS, debido a que ante el aumento de edad la probabilidad de uso de RS disminuye. Caso particular es el gasto en telefonía móvil.

El modelo 3 se basa en la ecuación 2.36 la cual considera todas las variables explicativas del modelo dos, con la salvedad de que en este ajuste se estima la probabilidad de ser usuario de Twitter. En primer lugar, las variables que explican estadísticamente la condición de ser usuario de Twitter son las relacionadas con sexo, edad nivel de escolaridad, cantidad de dispositivos y gasto en telefonía móvil. En cuanto a la dirección de la causalidad, las variables que afectan negativamente la probabilidad de uso son las referentes al sexo, y edad. Estos coeficientes apuntan a reafirmar un sesgo en el uso de Twitter en las personas más jóvenes y en hombres, lo que indica la presencia de desigualdad en el acceso a Twitter marcada por el género y la edad. Por otra parte, las variables que afectan positivamente al uso de Twitter son el nivel de escolaridad, número de dispositivos y gasto en telefonía móvil. En cuanto a la magnitud del impacto, los momios asociados al sexo indican que una mujer tiene 27% menos probabilidades de usar Twitter con respecto a un hombre. Asimismo, conforme aumenta la brecha generacional se encuentra marcada en el uso de Twitter. Para personas con edades entre 35 y 54 años se tiene entre 33% y 38% menos probabilidades de usuario de esta RS con respecto a los jóvenes de 15 a 24 años. En el mismo rubro, personas de 65 años y más tienen 85% menos probabilidades de ser usuarios de Twitter con respecto a los jóvenes.

En cuanto a las variables positivas, los niveles altos de escolaridad son significativos para explicar el uso de Twitter. En particular, las personas con escolaridad media superior tienen 2.72 veces más probabilidades de ser usuarias de Twitter con respecto a las personas con educación básica. Esta probabilidad aumenta considerablemente cuando la persona posee un nivel de escola-

ridad de superior y más, ya que la probabilidad de uso incrementa más de 6 veces con respecto al estrato más bajo de dicha variable. El modelo 3 muestra que ante el aumento de un dispositivo con conexión a Internet, la probabilidad de ser usuario de Twitter aumenta 50%. Finalmente, el gasto en telefonía móvil indica que ante un aumento de \$1.00 en el gasto en este rubro la probabilidad de ser usuario no cambia, con excepción de gastar grandes cantidades en esta dimensión.

En conclusión la comparación de los tres modelos implican consideraciones a la manera de constitución del acceso a la dimensión digital en la ZMCM. En primer lugar, el modelo 1 muestra que la principal barrera para acceder a Internet es la socioeconómica y que estratos altos tanto de escolaridad como sociodemográficos son los más favorecidos para acceder a Internet. El modelo 2 en comparación con el modelo 1 indica que las variables estructurales pierden peso en el uso de Redes Sociodigitales. Esto se debe a que 98% de los usuarios de Internet usan alguna red social, y dado los resultados descriptivos, se puede inferir que este efecto es explicado por el uso generalizado de Facebook. Finalmente, el modelo 3 comprueba que el uso de Twitter presenta selectividad al interior de los usuarios de Internet, en primera basado en la escolaridad, edad, sexo y en las variables de dispositivos y gasto. En esta última, es notable el cambio de dirección explicativa entre los modelos 2 y 3. Mientras que para el uso de RS el gasto implica decrecimiento en las probabilidades de uso, en el modelo 3 se relaciona positivamente, es decir al aumentar el gasto aumenta la probabilidad de uso. En términos concisos, ser usuario de Twitter implica doble selectividad, en primera instancia se debe ser privilegiado para poder acceder a Internet y en segunda instancia se debe ser privilegiado para ser usuario de Twitter.

## **2.6 Consideraciones finales**

El presente capítulo tuvo como objetivo estudiar a profundidad la dimensión del poder digital distribuido para crear una gentrificación digitalmente aumentada en la Ciudad de México. Dada la poca literatura sobre el desarrollo de la realidad geográfica aumentada digitalmente, en este trabajo se propuso una aproximación teórica y empírica desde el enfoque de las desigualdades digitales, dado que éste último tiene mayor desarrollo. Con base en esto, se ha planteado una estrategia metodológica para entender las desigualdades en el acceso y uso de tecnologías digitales desde dimensiones estructurales tales como el entorno del hogar, la estructura familiar o la territorialidad. Al mismo tiempo, se ha incorporado dimensiones de tipo personal, tales como sexo, educación,

edad y dispositivos de acceso a Internet. Todo lo anterior apunta a entender la desigualdad digital no solo como el diferencial entre los que usa o no usan herramientas digitales, sino como un fenómeno de tipo social que se relacionan con condiciones de poder y escenarios favorables para realizar un ejercicio de comunicación en la dimensión digital. Todo lo anterior apunta a elementos para entender la construcción de realidades geográficas aumentadas digitalmente desde su primera dimensión, el acceso a tecnologías digitales y quienes acceden a esta herramientas.

En este sentido, el capítulo abona en el estudio de las condiciones previas sociales previas para el uso de Internet y además, ofrece elementos para entender la construcción de las desigualdades digitales, ya que éstas son el primer elemento a superar para acceder a las realidades geográficas digitalmente aumentadas. En el primer orden de ideas, los hallazgos acerca de la doble selectividad en el uso de Twitter dialogan con la incipiente literatura que ha indagado sobre los determinantes en el uso de redes sociodigitales, más allá de los determinantes psicológicos. Con respecto al segundo argumento, se ha mostrado que el uso Internet y Twitter responde a perfiles específicos de personas, asociadas a perfiles privilegiados social y económicamente.

El diferencial en el acceso a Internet es una agenda pendiente en México. Como se ha demostrado, en país la barrera más complicada de superar para participar de la dimensión digital es el acceso a Internet. Este acceso está condicionado, principalmente por dos variables estructurales, el estrato socioeconómico y la educación. A lo largo de este capítulo se puso énfasis en la relación directa que tiene la posibilidad de acceder a Internet y el estrato sociodemográfico. Este es un primer elemento que habla del poder que tienen las personas que participan de la vida digital por medio de la red. Refuerza la idea de que los contenidos encontrados y generados responden a realidades de un sector privilegiado económica y socialmente.

El segundo nivel de selectividad en la dimensión digital es el uso de Redes Sociodigitales. Con respecto a esto, se ha encontrado el principal uso dentro de Internet es tener un perfil en alguna RS, siendo la principal Facebook. Por su parte, otras redes como Twitter, Instagram o LinkedIn sí tienen un perfil de usuario particular. En este sentido, se ha mostrado que los usuarios de Twitter son generalmente jóvenes con alta escolaridad. Adicionalmente, los usuarios de sitios como Instagram son principalmente adolescentes mujeres con escolaridad media. Esto da indicios de que una vez que se ha accedido a Internet existes usos diferenciados dentro de la red y, que éstos también tienen explicación en el terreno social. De esta manera, las RS son campos donde se crean realidades

digitales basadas en las experiencias subjetivas de sus usuarios.

En conclusión este capítulo permite establecer que el uso de Twitter en la ZMCM no es homogéneo en la población. Por el contrario, el perfil sociodemográfico del usuario de esta red sociodigital podría formar un subgrupo particular de las personas que han sido identificadas con un perfil gentrificador activo (Ver sección sobre perfil sociodemográfico del perfil gentrificador en el capítulo 1) En este sentido, para que una persona pueda llegar a usar Twitter debe primero tener la posibilidad de acceder a Internet, esta barrera es la más importante y difícil de superar. Una vez que una persona haya accedido a Internet la siguiente barrera a superar es la de uso de Twitter. Este trabajo ha demostrado que el uso de Twitter, probabilísticamente hablando, tiene sesgo de género, es decir, las mujeres tienen menor representatividad en esta RS. Aunado al género, las variables más representativas para explicar el uso de Twitter son: alta escolaridad, varios dispositivos para acceder a Internet y gasto en telefonía móvil. Por su parte, los altos niveles de escolaridad sugieren un perfil de persona profesionista, que se puede llegar a vincular con actividades de servicios o profesionales, tal como se expuso en los primeros aparados de este capítulo. En lo que respecta a la tasa de dependencia, ésta no es representativa, lo que implica que personas que habiten en hogares constituidos por jóvenes tienen la misma probabilidad de ser usuarios de Twitter que personas que habiten en hogares con personas mayores o niños. De manera general se puede concluir que Twitter es una herramienta digital con alto grado de selectividad y que esto implica un sesgo importante para construir realidades aumentadas digitalmente, en particular en procesos de gentrificación.

El presente capítulo arroja algunos indicios sobre el proceso de gentrificación aumentado digitalmente en la ZMCM. Como se ha mencionado en el primer capítulo y al inicio del presente el proceso de gentrificación es un fenómeno de segmentación espacial fuertemente influenciado por la dimensión de clase social y de reproducción del capital en el espacio urbano. En este sentido, en el capítulo teórico se ha propuesto que una aproximación teórica para entender cómo la gentrificación se aumenta digitalmente es la teoría de la Realidad Geográfica Aumentada Digitalmente. Dicha teoría compagina con los postulados analítico de la gentrificación en el hecho de que la RGAD también plantea que una espacialidad digital aumentada atraviesa elementos políticos y de diferenciación social. Así, se ha apunado que el acceso a TIC y, específicamente a RS en la realidad mexicana se explica por elementos de desigualdad social por lo tanto el estudio de la gentrificación digital encuentra en estos elementos una base para el posterior estudio de la conformación de es-

pacialidades por medio de Twitter. En este punto, cabe mencionar que Twitter es una plataforma digital elitizada que es reflejo del poder de distribución digital, al igual que espacios gentrificados se dirigen a personas de clase media alta.

## Capítulo 3

# Gentrificación aumentada en la Ciudad de México. Análisis de la distribución y comunicación digital mediante los patrones de emisión espacio-temporal de Twitter

### 3.1 Introducción

El objetivo central de este trabajo de investigación es indagar las maneras en las que el proceso de gentrificación en la Ciudad de México se ve aumentado digitalmente mediante Twitter. Hablar de la dimensión digital de los procesos urbanos es todavía un campo naciente en el campo de los estudios urbanos. Para subsanar esta brecha en el conocimiento de la gentrificación contemporánea se ha propuesto ensamblar la propuesta de la construcción de Realidades Geográficas Aumentadas Digitalmente (RGAD) para evaluar si las fuerzas de igualación/diferenciación y valorización/desvalorización del espacio urbano, propuestas por Neil Smith, se aumentan digitalmente por medio de Twitter en el entorno urbano.

La aproximación de la RGAD se construye por medio de tres poderes que interactúan para construir espacios digitales desiguales. Estos poderes son el poder de la distribución digital, el poder de la comunicación y el poder del código. El primero de éstos apunta a que el acceso a la dimensión digital no es homogénea entre la población, ni en el espacio. En este sentido, el poder distribuido de la dimensión digital se condiciona, en primera instancia, por quienes pueden acceder a él y los lugares con la mejor infraestructura para el desarrollo pleno del uso de Internet. Así, el poder de la distribución digital se compagina con las tendencias igualadoras/diferenciadoras del espacio urbano. Tanto la infraestructura como las personas con habilidades digitales se correlacionan espacialmente con lugares centrales, económicamente y socialmente importantes importantes. De la misma manera, la distribución digital establece diferencias entre los espacios. Esto, en primera instancia, podría parecer una contradicción a la luz de la hiperconectividad de Internet, sin embargo, en términos geográficos, existen diferencias en el acceso y producción digital a lo largo

del territorio.

La igualación/diferenciación del espacio por medio de lo digital se relaciona directamente con la valorización/desvalorización del espacio y la representación de ciertos lugares en el territorio. En contraposición a la hipótesis de Cairncross (2001) sobre la desaparición de la distancia geográfica gracias al avance en las telecomunicaciones y, especialmente de Internet, se ha mostrado que los desarrollos tecnológicos generan geografías distintas a lo largo del territorio (Dodge y Kitchin, 2001). Es por que ello que M. Graham, Zook, y Boulton (2013a) proponen analizar como se distribuye de manera desigual la emisión digital de contenidos en el espacio. Esto como parte de la conformación desigual de una realidad geográfica aumentada digitalmente desigual.

Para el caso particular de este trabajo de investigación esta conformación desigual de una RGAD sirve de puntal para explorar la manera en la que el proceso de gentrificación se aumenta digitalmente. Para ello, el presente capítulo busca llevar a cabo el objetivo de explorar distintas dimensiones que pueden ayudar a aproximar una gentrificación aumentada digitalmente en la ZMCM. Con este propósito, el capítulo se estructura en cuatro grandes apartados.

En la primera sección de este acápite se exponen algunos elementos teóricos generales para entender cultural y sociológicamente el uso de Twitter, así como su aplicación en estudios urbanos. Cabe mencionar que en este punto se puntualiza que el uso de los grandes datos, particularmente de Twitter se encuentra en una fase inicial en su aplicación en el área. En este sentido, se exponen un par de teóricas que pueden ayudar a comprender el funcionamiento particular de Twitter y el mecanismo social y cultural de construcción de geografías particulares. En esta misa sección se exponen los elementos teóricos y conceptuales de Twitter en estudios urbanos, enfocado en el uso de la geolocalización de tweets. Así se desarrollan elementos teóricos sobre las nuevas fuentes de información generadas por usuarios, la llamada Información Geográfica Voluntaria.

La segunda sección del capítulo se enfoca en la descripción técnica de las fuentes de información e identificación empírica de las unidades espaciales de estudio vinculadas con el proceso de gentrificación. En primera instancia se expone el proceso de recolección de tweets con criterios geográficos y los múltiples pasos que se han seguido para depurar la información recolectada de Twitter. En este punto, no solo se pone énfasis en las dimensiones clásicas sobre el manejo de bases de datos, tales como tamaño de muestra, sesgo estadístico y representatividad, también se menciona el papel de la dimensión no neutral del código computacional para generar este tipo de informa-

ción. En segunda instancia se desarrolla la identificación empírica de colonias gentrificadas. Como se ha comentado en el capítulo 1, este trabajo de tesis retoma una corriente comercial y culturalista de la gentrificación. Es así, que las variables que se discuten en esta sección apuntan a la selección de colonias por medio de indicadores relacionados con actividades comerciales e infraestructura pública de renovación de espacios. Finalmente en esta sección se describen los indicadores construidos para el estudio. En este punto se detalla la dimensión temporal de la emisión de tweets y la construcción de núcleos de densidad de emisión.

En el tercer apartado se presentan los principales resultados. En este apartado presenta el nivel de emisión de tweets georreferenciados y no, así como su representación cartográfica. En este punto se comienza a vislumbrar que existe una concentración significativa de tweets en las colonias centrales de la ciudad, identificadas como gentrificadas. Además, se tiene concentración en puntos específicos fuera de estas colonias, principalmente con zonas relacionadas con centros comerciales, actividades de esparcimiento o turismo. Este patrón general se explora en cada segmento temporal del día. Los principales hallazgos apuntan a que las concentraciones de emisión en días laborales se relacionan con actividades de oficina y servicios. Mientras que los fines de semana sobresalen los lugares gentrificadas en horarios de esparcimientos.

Para finalizar se desarrollan algunas consideraciones finales. Especialmente se mencionan algunas enseñanzas en el manejo de tweets para un caso urbano. De igual manera se describe cómo el desarrollo empírico de este capítulo apunta a la gentrificación aumentada digitalmente.

### **3.2 Twitter en Estudios Urbanos. Fase experimental en el uso los grandes datos para entender la ciudad**

Como se ha mencionado, el objetivo de este capítulo es explorar la relación entre la distribución geográfica de la emisión de Twitter con respecto a espacios que han sido renovados o gentrificados en la Ciudad de México. Por un lado, en el capítulo 1 se han mencionado elementos teóricos que apoyan la identificación empírica de la gentrificación. Entre éstos se puede mencionar indicadores como la actividad comercial ligada a comercios, restaurantes o bares, o bien, elementos de intervención estatal en favor de la renovación y valorización del espacio urbano, tales como la colocación de bicicletas públicas. Al mismo tiempo, en el capítulo mencionado se introdujeron las nociones generales sobre los grandes datos y los procesos urbanos. En este sentido se comentó

que la gran cantidad de datos no solo permite realizar mejores aproximaciones a los fenómenos urbanos estudiados tradicionalmente, sino que, dichos datos por su naturaleza digital forma parte de la constitución de la dimensión digital de procesos en la ciudad. Sin embargo, los grandes datos no son homogéneos tanto en su origen como en sus consecuencias. Bajo esta premisa, se analiza particularmente Twitter como fuente de información y como mediador de la dimensión digital del proceso de gentrificación en la Ciudad de México.

Con lo anterior en mente, en la presente sección se exponen elementos particulares de Twitter tanto desde enfoques de análisis sociológico para entender el papel de dicha red sociodigital en la constitución de nuevas realidades sociales, como de Twitter como nueva fuente de información. De esta manera, a continuación se desarrollan algunos enfoques teóricos emanados desde la sociología que permiten entender el rol de Twitter para la conformación de la gentrificación aumentada digitalmente. Particularmente se pone énfasis en la generación de espacios centrales y su co-incidencia con la emisión de publicaciones en Twitter. Enseguida se exponen algunos estudios empíricos que han utilizado Twitter tanto en términos analíticos como empíricos para procesos urbanos, en general.

### *3.2.1 Acercamientos sociales teóricos a Twitter*

Como se ha comentado en anteriores capítulos, la dimensión de lo digital es enorme y, en la actualidad, es imposible separarla de lo no digital, ya que se han convertido en una realidad híbrida. Aunado a lo anterior, se ha mostrado que la principal barrera para acceder a la dimensión digital es tener acceso a Internet. Dentro de Internet los usos se diversifican, por ejemplo, el comercio, banca o gobierno electrónicos, entretenimiento por medio de plataformas de *streaming*, etc. A su vez, el principal uso que las personas le otorgan a Internet son las redes sociodigitales. Para el caso de México y de la ZMCM la red más utilizada es Facebook. En este sentido, el uso de los grandes datos provenientes de redes sociales como la mencionada Facebook o Twitter, proporcionan un buen acercamiento a la naturaleza de la dimensión digital en la vida de las personas. Por ejemplo, a decir de (Cheng y cols., 2016) las redes sociales para el estudio de lo urbano ayuda a visibilizar dimensiones como el estado de ánimo de los habitantes con respecto a un tema particular o puede ayudar a aproximar la movilidad de las personas.

Dentro del conjunto de redes sociodigitales disponibles, la segunda más utilizada es Twitter.

Igualmente, los datos obtenidos son los más utilizados por estudios académicos, políticos y comerciales. Esto se debe a la manera en cómo configura las redes sociales en el mundo virtual, así como la estructura de *microblogging*. Estas características se explican más detalle en este capítulo. Aunado a esto, Twitter es la red social que ofrece sus datos de manera gratuita y de fácil acceso de manera pública. Al mismo tiempo, esta plataforma tiene algunas limitaciones para el estudio social. En primera instancia, el sesgo de sus usuarios. Como se ha mostrado en el capítulo 2 de este trabajo, dichos usuarios son primordialmente personas de estrato alto jóvenes con nivel de escolaridad alto. En segundo lugar, los datos gratuitos de Twitter son limitados.

A pesar de estas limitaciones y de la configuración particular de las dinámicas en Internet la plataforma ofrece una puerta de entrada para los estudios urbanos en varios niveles. Uno de estos niveles puede ser netamente empírico, ya que sea que se interpreten los datos de Twitter como fuente alternativa de información o como apoyo a las ya existentes. Otro nivel puede ser más teórico en tanto Twitter es una pequeña parte de la dimensión digital y, por ende, es motor de la transformación social digital.

Desde una aproximación sociológica Twitter ha sido teorizado desde dos vertientes. La primera utiliza una postura basada en los planteamientos de teoría interaccionista de Erving Goffman para teorizar el funcionamiento de esta Red Sociodigital. Por el otro lado, se encuentra la postura de Twitter como medio de *prosumo* de los usuarios. Ambas posturas aportan elementos para entender el comportamiento de Twitter desde una mirada geográfica, y particularmente para entender el proceso de gentrificación en su nivel digital.

El primer acercamiento para entender Twitter con base en teoría social ha sido desarrollado por Murthy (2012). El autor propone conceptualizar la conformación de esta red sociodigital desde el enfoque interaccionista y de formas del discurso planteado por Goffman (1981). Con base en los postulados de este autor, Murthy entiende Twitter teóricamente desde tres conceptos básicos: el ritual, el marco de participación y el anclaje, todos éstos, elementos del discurso. El primer punto, el ritual, refiere a los gestos o símbolos que se realizan al momento de hablar. Para el caso de Twitter, el autor señala que estos gestos trascienden la comunicación escrita y se manifiestan por medio de símbolos o elementos no textuales pero con significado, tales como los *emojis* o las fotografías vinculadas al tweet <sup>1</sup>. El segundo elemento corresponde al marco de referencia del mensaje. Este

---

<sup>1</sup>En este trabajo se refiere a las publicaciones en Twitter como *tweets*. Cabe mencionar que no existe una aproxima-

concepto de la teoría de Goffman adquiere relevancia en las interacciones mediadas por computadora debido a que da cuenta de la posición que los usuarios ocupan en la red social. Al respecto, el autor apunta a la construcción de discurso basado en el lugar simbólico que determinados usuarios poseen. Este valor simbólico se identifica por medio del número de seguidores a su cuenta, o bien, la cantidad de veces que otros usuarios comparten sus publicaciones. Este tipo de usuarios generan influencia en los temas que se desarrollan en la esfera de Twitter (Lahuerta-Otero y Cordero-Gutiérrez, 2016; Alp y Ögüdücü, 2018). Dicha posición, generalmente, la poseen personas de la esfera pública, tales como artistas o políticos. El tercer elemento es el anclaje. Este elemento permite entender la estructuración del discurso basado en el espacio situacional de las personas y el mensaje que éstas quieren transmitir en función de la audiencia espectadora. El anclaje, por ende, refiere al ensamblaje discursivo de los usuarios de Twitter. En otras palabras, los usuarios emiten mensajes para aquellas personas con las cuales se puede generar una posible interacción. Esto se ha comprobado para círculos especializados tales como el artístico, musical o médico (Borgmann y cols., 2016; Expósito, Rodríguez, y Barrau, 2017; Krause, North, y Heritage, 2018). Aunque la propuesta de Murthy recupera elementos de una teoría sociológica del discurso para entender el funcionamiento de Twitter, cabe mencionar que el sesgo que subyace a este enfoque es considerar a Twitter como un nuevo medio de comunicación con características de horizontalidad, en contraposición de medios de comunicación tradicionales como televisión, prensa o radio. En este sentido, se podría mencionar que el enfoque deja de lado otras consecuencias sociales de Twitter.

El segundo enfoque teórico para entender los sitios de redes sociales en Internet es el *prosumo*. Este concepto refleja la conjunción de ambas actividades dentro del ciclo de producción, la producción y el consumo al mismo tiempo. Por un lado, durante el desarrollo temprano del capitalismo, el motor de la economía se sustentaba en la producción de bienes y servicios y, ésta subordinaba al proceso de consumo como consecuente del primero. Con el desarrollo del sistema capitalista durante el siglo XX y el incremento de la producción de dichos bienes gracias a los avances técnicos y tecnológicos, el papel fundamental de la producción dejó de ser importante y cedió el paso al consumo de los bienes producidos masivamente. Este cambio de enfoque en la economía se basó estrictamente en acelerar el tiempo del ciclo de producción-consumo (Fuchs, 2014; Ritzer y Jurgenson, 2010). Ante la premisa de acelerar dicho proceso, Ritzer y Jurgenson (2010)

---

ción clara sobre la manera correcta de mencionar este concepto en idioma español.

identifican un elemento intermedio entre la producción y el consumo, al cual denominan *prosumo*. Este prosumo se observa en el papel que el consumidor asume al comprar un bien y al mismo tiempo interviene en su proceso de producción. Los autores ejemplifican esta propuesta por medio del mecanismo de autoservicio en las cadenas de comida rápida. En éste, los consumidores consumen un alimento y, al mismo tiempo fungen como meseros. A decir de los autores que sustentan este enfoque, esta es una estrategia de las empresas para reducir costos e incrementar sus ganancias.

Para el caso particular de Twitter, un usuario es tanto consumidor como productor de contenido digital. Para Fuchs (2014), el trabajo digital y la organización económica se puede esquematizar en tres esferas, la económica, política y cultural, las cuales, a su vez, forman iguales esferas de poder. En este sentido, los usuarios de redes sociales adquieren un papel de productores de contenido digital el cual sustenta el poder de llegar a posibles usuarios, por ejemplo, los usuarios que generan influencias en Twitter. Así, estos usuarios que se ubican en la cima de la estructura social de Twitter generan un poder comercial que es aprovechado para publicidad. Al mismo tiempo, cada usuario al realizar un *click* en Internet también está generando información para que sea analizada y se mejoren los procesos de publicidad (Nishimura, Sukegawa, Takano, y Iwanaga, 2018).

Tanto el enfoque interaccionista del discurso como el de prosumidor digital posibilitan analizar la construcción de la gentrificación digitalmente aumentada. Por un lado, el enfoque interaccionista proporciona elementos para entender cómo el poder de la comunicación puede encontrar sustento en la formación de un discurso específico, ya sea por medio de símbolos concretos, o bien, por medio de la conformación de redes en la dimensión digital. En este sentido, el poder de la comunicación puede explicar cómo se crean símbolos a partir de lugares gentrificados y su reflejo en Twitter. Al mismo tiempo, dichos símbolos no son elementos neutros en Twitter, éstos son emitidos para determinados interlocutores que entienden y reproducen el mensaje, en este caso de espacios y prácticas gentrificados. Por otro lado, la teoría del prosumo en Twitter permite entender que las publicaciones en dicha red están vinculadas con espacios económicos concretos. En este aspecto se puede postular que las publicaciones de Twitter co-inciden en establecimientos económicos localizados en colonias gentrificadas y, además dirigidos a prácticas relacionadas con una élite. En este sentido, la desigualdad de emisiones de tweets podría encontrar elementos que lo expliquen en el papel de las personas como prosumidores de espacios y prácticas de gentrificación.

### 3.2.2 *Twitter en los Estudios Urbanos en el siglo XXI*

Desde el área de Estudios Urbanos, Geografía o Sistemas de Información Geográfica, Twitter ha sido de interés por su carácter novedoso como fuente de información. Así lo apuntan la mayoría de trabajos sobre la temática, ya sea como Información Geográfica Voluntaria, información sobre desastres de origen natural o identificación de desigualdades territoriales. En este sentido, los grandes datos y, particularmente Twitter, han sido considerados en el área de información geográfica como una nueva dimensión constitutiva de los fenómenos geográficos. Este hueco en el desarrollo teórico de la dimensión digital y los grandes datos es, quizá, uno de los grandes retos para la disciplina en años recientes. En este marco, a continuación se presentan, de manera breve, algunos temas que engloban el uso de Twitter desde una perspectiva geográfica o urbana.

La primera consideración sobre los grandes datos y geografía fue propuesta por Goodchild (2007). El autor propuso el término de Información Geográfica Voluntaria (IGV) basado en el volumen creciente de información con carácter geográfico en Internet y, sobre todo generada no por organismos oficiales, sino por personas. Así, Goodchild observó en plataformas como Wikimedia, Open Street Maps, Flirck nuevas fuentes de información geográfica. A decir del autor, la IGV se sustenta en un conjunto de condiciones previas. La primera de ellas es la democratización de los Sistemas de Información Geográfica (SIG) por medio de la Web 2.0. El segundo punto que sustenta la IGV es el sistema de coordenadas, es decir la homologación de medición sobre el globo terrestre en latitud y longitud. El tercer punto que enfatiza el autor es la mejora en la generación de gráficos y visualización de la información geográfica debido al avance computacional. Como cuarto elemento, se tiene el incremento de la velocidad de transferencia de datos por Internet, es decir, el incremento de banda ancha. Como última característica, la IGV se sustenta en la capacidad de geoetiquetar el contenido de internet. La geoetiqueta <sup>2</sup> es un código estandarizado que puede ser insertado en la información de la publicación en Internet, con el propósito de dar cuenta de su ubicación.

La IGV, reconoce el autor, presenta varios problemas en cuanto a la calidad de la información. Por un lado, Goodchild menciona que ésta es generada con sesgo debido a la disparidad en el acceso de Internet. Aunque, en el caso particular de Twitter, menciona Goodchild y cols. (2017), tiende a superarse, debido a que en países del norte global la disparidad en el acceso se ha reducido y Twitter se ha convertido en una red sociodigital de uso generalizado. Sin embargo, en el capítulo

---

<sup>2</sup>En la terminología computacional se reconoce como *Geotag*

anterior se ha desarrollado la idea que, para el caso mexicano, la disparidad en el acceso a Internet sigue continua siendo factor importante para explicar el fenómeno digital en el país. Por otro lado, la IGV es susceptible a las motivaciones que las personas tienen para publicar su localización al momento de generar contenido digital. Al respecto, el autor menciona que esto puede ser explicado por exhibicionismo, necesidad de pertenecer a un grupo social, o bien, por la satisfacción personal de participar en la conformación de proyectos colectivos, como es el caso de *Open Street Maps*. En este punto, desde posturas críticas, se cuestiona que la geoetiqueta genera diferentes espacialidades en función de la temporalidad y motivos de emitir públicamente la ubicación del emisor. Así, la geoetiqueta refleja la constitución social de la publicación digital (Shelton, 2017).

Otra vertiente de análisis de Twitter ha sido la construcción de ciudadanía en espacios urbanos (Daume y Galaz, 2016). En este punto, los estudios sobre la generación en tiempo real por parte de los ciudadanos ante desastres naturales ha sido el tema más analizado. En el caso particular de la Ciudad de México ante el sismo del 19 de septiembre de 2017, Twitter fue una de las herramientas digitales utilizadas para conocer, de manera veraz, los puntos de siniestro en la ciudad. En este marco, la plataforma #Verificado19S surgió como una propuesta ciudadana por medio de tweets geolocalizados para verificar puntos de siniestro y evitar la desinformación colectiva horas y días después de ocurrido el sismo (Mekler, 2018; Salgado-Andrade, 2018).

Por último, Twitter ha sido utilizado para analizar cómo se configura el espacio urbano a lo largo del día. En este tema Neuhaus (2015) desarrolla el análisis temporal de la emisión de tweets en distintas ciudades alrededor el mundo. En primera instancia, el autor encuentra que la emisión durante el día no es homogénea. El autor explica que esto se puede deber al tipo de ciudad que se analiza. Ejemplo de esto es la configuración espacio-temporal de la emisión en Nueva York en comparación con Londres. Asimismo, el autor propone entender la emisión de Twitter basado en el concepto de distinción social de Pierre Bourdieu. De igual manera, estipula que la configuración de la ciudad a lo largo del día crea diferentes espacios sociales en el mismo espacio físico debido a la dinámica temporal de las personas. En este sentido, el autor recupera la idea teórica del análisis del ritmo urbano de Lefebvre (Lefebvre, 2004). Cabe mencionar que la dimensión temporal de la ciudad ha sido poco estudiada empíricamente debido a la dificultad de recabar información con esta dimensión analítica. Sin embargo, Twitter, y en general los grandes datos, permite que se realicen análisis temporales debido que la información es recabada en tiempo real. Esta dimensión es la que

se desarrolla a lo largo de este capítulo para dar cuenta de la conformación de espacios digitales de poder en la ZMCM a diferentes horas en el día.

Hasta este punto se han mencionado algunos marcos conceptuales para el estudio de Twitter como herramienta metodológica. Enseguida, el apartado metodológico apunta a la descripción de Twitter como una fuente de información para estudiar fenómenos urbanos. En particular, se tiene como objetivo analizar cómo se distribuye espacial y temporalmente la emisión de tweets en la ZMCM. Se espera que dicha distribución no se homogénea en el espacio, ni en el tiempo. En este sentido, se podría postular que las emisiones de Twitter se vinculan con espacios centrales gentrificados o que han pasado por un proceso de renovación en la ciudad. De igual manera, se espera que esta diferenciación de emisión también sea guiada por elementos temporales del ritmo de vida urbano, tales como los horarios laborales y esparcimiento, que en mayor o menor medida se relacionan con el fenómeno de gentrificación. En conjunto, la emisión diferenciada de tweets tanto temporal como espacialmente se podrían entender como un reflejo del poder de la distribución digital para aumentar digitalmente el proceso de gentrificación en la ciudad.

### **3.3 Apartado metodológico**

En esta sección se presenta la descripción de las fuentes de datos, tanto de Twitter como fuentes oficiales y la construcción de indicadores. Asimismo se presenta la unidad espacial de observación. Para tal propósito, en primera instancia se habla sobre la recolección de tweets, el tipo de información que se puede obtener de dichos tweets, la manera en la que se ha organizado la información. Enseguida, se desarrolla la identificación de colonias gentrificadas en de la ZMCM, esto con base en variables obtenidas del DENUE e información abierta gubernamental sobre bicicletas compartidas en la ciudad (ECOBICI). En tercer lugar, se expone la construcción de variables el análisis de distribución espacio temporal de emisiones de Twitter.

#### *3.3.1 Descripción de la base de Twitter para la ZMCM*

En el capítulo 2 se desarrolló la idea de los grandes datos en la actualidad y cómo éstos pueden afectar nuestra realidad social. Adicionalmente, se ha comentado que una de las características de la revolución de los grandes datos es su carácter semi-estructurado o no estructurado (Kitchin, 2014b).

Sin embargo, como lo menciona Salganik (2018), este tipo de información no es directamente generada para realizar estudios en ciencias sociales. Dicha información es propiedad de empresas privadas tales como Facebook, Twitter y la mayoría de sitios web. En consecuencia, a pesar de que se tiene gran cantidad de datos, éstos no son públicos, en su mayoría. Este es el caso de Facebook, esta empresa no permite de manera pública obtener los datos de su plataforma. El ejemplo contrario es Twitter, ya que ésta permite obtener una pequeña muestra de las publicaciones que sus usuarios realizan día a día. Esta es una de las razones por las cuales Twitter es la red sociodigital más estudiada en el campo de los grandes y su intersección con las ciencias sociales.

Trabajar con grandes datos es un proceso que es complicado y no estandarizado que se encuentra en construcción. El manejo de grandes bases de datos, de estructura no relacional y de fuentes de Internet ha sido ampliamente abordado por científicos de datos o computólogos. Sin embargo, el manejo en ciencias y sociales es escaso y mucho más en referencia a estudios urbanos. Por ello, a continuación se expone, los pasos que se siguieron en la recolección, depuración y análisis de la base de datos de Twitter. Los pasos son los siguientes, recolección, depuración y ensamblamiento con otras bases de datos. A continuación se describe la primera fase, la recolección de la información.

En general, se puede obtener información de sitios de redes sociales o de cualquier sitio web, que lo permita, por medio de una API. Una interfaz de programación de aplicaciones (API por sus siglas en inglés, *application programming interface*), consiste en un conjunto de comandos que pueden ser usados para recuperar datos almacenados en bases de datos de servidores públicos (Kitchin, 2014b). Otra definición de API es el “conjunto de herramientas de programación que posibilita a un programa la comunicación con otro programa o un sistema operativo, y eso ayuda a los desarrolladores de software crear sus propias aplicaciones (=piezas de software)” (Oxford, s.f.). En otras palabras, una API es un programa que le solicita continuamente información a un sitio web público para almacenarla. Para el presente estudio se ha utilizado la API de Twitter que permite recolectar en tiempo real una pequeña muestra de los tweets emitidos diariamente<sup>3</sup>. La API de Twitter permite recolectar tweets bajo criterios específicos, por ejemplo, usuario, localización geográfica de la emisión, idioma del tweet, etc. En el código A.1 del aparatado de anexos se presenta

---

<sup>3</sup>Existen diversos software disponibles para configurar la API de Twitter. Entre ellos se encuentran Python o R. En este caso se ha utilizado la solución LEK(Logstash-Elasticsearch-Kibana), el cual es un pipeline que referencia y visualiza consultas.

Los pasos para la descarga se pueden consultar en Coronado (2015)

la programación completa de la solicitud de información a Twitter, para el caso de la ZMCM. Lo más importante de esta pieza de código es el criterio geográfico, ya que con base en éste se realizó la recolección de tweets. En particular, la API se ha construido tomando como delimitación geográfica las las coordenadas que conforman el rectángulo que enmarca el polígono de la ZMCM

Para este estudio se ha recolectado toda la información que permite Twitter, entre las variables públicas se encuentran, día y hora de la emisión, identificación y nombre del usuario, texto publicado, sistema operativo, retweets, hashtags o bien coordenadas y lugares de la emisión. Aunque la información proporcionada por la API es vasta, para el presente trabajo solo se recuperan 5 variables: día y hora de la emisión, nombre de usuario, texto publicado y coordenadas geográficas. En este punto es importante mencionar que no todos los usuarios permiten conocer públicamente su geolocalización exacta. En otras palabras, las coordenadas no se encuentran disponibles para todos los registros obtenidos. Al respecto, estudios previos han estimado que la API gratuita de Twitter recupera solo 1 % del total de los tweets que se emiten con los criterios seleccionados (Pfeffer, Mayer, y Morstatter, 2018).

La API de Twitter que se ha seleccionado permite identificar cuántos tweets son georreferenciados con respecto al total. En este terreno, los estudios previos reportan una relación directa entre países desarrollados y la tasa de usuarios que permiten conocer su localización exacta en Twitter, por medio de su GPS. Para países europeos el porcentaje de tweets georreferenciados ronda 20 %. Por ejemplo, Londres, París o Hong Kong reportan que los tweets georreferenciados son entre 20 % y 30 %. Mientras que ciudades como Nueva York o Amsterdam esta cifra asciende a 50 % (Neuhaus, 2015). En contraste, países en vías de desarrollo, el porcentaje promedio apenas alcanza 10 % del total de tweets con esta característica.

La representatividad estadística de la base de datos georreferenciados proveniente de Twitter es un tópico a considerar dentro de los estudios sociales. Los grandes datos, a diferencia de los pequeños datos provenientes de encuestas estadísticamente diseñadas, enfrentan limitaciones para uso como fuentes de información sobre el comportamiento general del conjunto de la población. Particularmente, estudios previos han indagado sobre la capacidad de estimación población de los datos de Twitter, así como de su sesgo estadístico (Achrekar, Gandhe, Lazarus, Ssu-Hsin Yu, y Liu, 2011; Filho, Almeida, y Pappa, 2015; ?, ?). Este sesgo estadístico en la información ha sido mostrado para las bases de datos de tweets en general. En particular, el sesgo de los tweets geo-

rreferenciados ha sido estudiado en los trabajo de Sloan y Morgan (2015); Sloan y cols. (2015) quienes muestran que las personas jóvenes con profesiones cargadas a los servicios son quienes usualmente muestran su geolocalización al publicar en Twitter. A decir de los autores, este sesgo hacia este perfil de personas vuelve complicado el uso de las grandes bases de datos como herramientas de inferencia poblacional. En este mismo orden de ideas, se han realizado algunos estudios que consideran la limitante del sesgo muestral de Twitter para estudios territoriales, por ejemplo de movilidad de personas (Hawelka y cols., 2014).

La información recolectada de Twitter específicamente de la ZMCM corresponde al periodo del 1 de abril de 2017 al 31 de marzo de 2019<sup>4</sup>. Durante el periodo de recolección se logró recuperar 19.4 millones de tweets. En promedio, en el polígono seleccionado de la ZMCM, se recuperaron 29 538 tweet diarios. En particular, del total de tweets capturados por la API, 2 923 543 millones fueron georreferenciados, esta cifra representa 14.94%. Esta cifra se asemeja a lo reportado en países desarrollados, como Inglaterra o Alemania. Por su parte, en términos de usuarios, en la muestra se tienen 471 260 usuarios. Debido al interés de este trabajo radica en la dimensión digital de la gentrificación, la muestra de interés para este estudio solo considera los tweets que permiten conocer la georreferencia de la publicación. A continuación se describe el proceso de depuración.

Un elemento importante para el trabajo con grandes datos es la depuración de la información. En el caso la información de Twitter se ha identificado que existen determinados elementos computacionales que están presentes en la recolección de tweets. En particular, en los tweets existe la presencia de cuentas de usuarios automatizadas que bien pueden introducir ruido en los análisis o bien modificarlos por completo, por la misma naturaleza de dichas cuentas (Chavoshi, Hamooni, y Mueen, 2016; Mønsted, Sapieżyński, Ferrara, y Lehmann, 2017; Stella, Ferrara, y De Domenico, 2018).

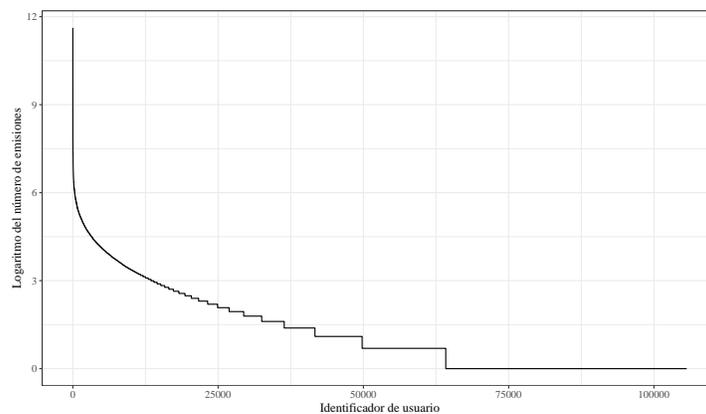
En este sentido, una primera tarea sobre la base de datos de los tweets ha sido la identificación de posibles *bots*. Como se ha mencionado anteriormente, los *bots* son cuentas automatizadas que son creadas con el fin de generar tendencias y resaltar determinados temas en el espacio de Twitter, y por lo tanto en la dimensión digital. En este tema, se ha explorado que el papel de estas cuentas automatizadas sesga el estudio de cualquier tipo de tendencia en la base de datos de Twitter. Al

---

<sup>4</sup>Cabe mencionar que no se tiene el total de días, debido a que en el transcurso del periodo de recolección se experimentaron problemas con la conexión al servidor de Twitter. Por lo tanto, se obtuvieron 658 días efectivos de recolección.

respecto, se han desarrollado varios algoritmos para identificarlos y tratar de eliminar el sesgo de éstos. Por ejemplo, Stella y cols. (2018) identifican las cuentas automatizadas por medio de la interacción con otros usuarios, en particular por su índice de tweets y retweets. Por su parte, Vosoughi, Roy, y Aral (2018) encuentran tweets con el objetivo de la difusión de noticias falsas por medio de Twitter. Los autores se enfocan en el contenido del tweet en lugar de la interacción de los usuarios. Estos trabajos dan cuenta de la dificultad técnica y analítica de la identificación y eliminación de *bots*, sin embargo, dado los alcances de este trabajo se ha optado por una criterio netamente de frecuencia de emisión. Es decir, se ha inferido que aquellos usuarios que emitan en exceso serán considerados como *bots*, para esto se ha establecido un umbral de más de 3 000 emisiones

Figura 3.1: Número de publicaciones según usuario (escala logarítmica)



Fuente: elaboración propia

En la figura 3.1 se observa el número total de emisión por usuario durante el periodo de recolección, en escala logarítmica. En términos descriptivos, se tiene que la mediana de emisión por usuarios es de 3.12 tweets diarios, en escala logarítmica esto implica 1.32. Con esto en mente, el primer elemento a destacar es que muy pocos usuarios emiten cantidades superiores a la reportada por la medida de tendencia central antes dada. Dado el sesgo que este tipo de usuarios puede causar en el análisis, las emisiones de usuarios que superen el umbral de 3 000 <sup>5</sup> han sido eliminadas. Dado que el comportamiento promedio de un humano oscila en 3 publicaciones, es considerable asumir que las cuentas que emitan más de 3 000 veces son usuarios automatizados. Para corroborar

<sup>5</sup>Lansley y Longley (2016) proponen un umbral máximo de 3 000 emitidas por un solo usuario como indicador de usuarios que pueden llegar a influenciar el análisis de tweets tanto en emisión como en análisis de contenido de texto.

este supuesto se presentan los nombres de las cuentas de usuarios eliminadas.

Cuadro 3.1: Usuarios con más de 3000 publicaciones

Usuario	Número de publicaciones
trendinaliaMX	111327
TrendsMexico	14885
PairsonnalitesE	7331
Pairsonnalites	7196
pairsonnalitesB	7043
trendinaliaMEX	5206
PairsonnalitesE	11128
pairsonnalitesB	11092
Pairsonnalites	7536

Fuente: Elaboración propia

El cuadro 3.1 muestra los nombres de usuario que han emitido más de 3 000 publicaciones en el periodo de recolección. Como se observa, los nombres pertenecen a usuarios no humanos o automatizados. En particular, las cuentas de ‘Trendinalia’ se dedican a publicar de manera repetida las tendencias que se generan a nivel mundial y nacional. Como se observa, estos usuarios no producen contenido analítico sobre la interacción entre los espacios urbanos y las prácticas de las personas. Sin embargo, sí pueden dar cuenta de la dimensión digital en términos del código que subyace a las publicaciones de Twitter para resaltar determinadas tendencias o temas. El último elemento que confirma la naturaleza automatizada de estas cuentas es su localización geográfica. En este ámbito la figura 3.2 muestra que todas las emisiones de los usuarios descritos en el cuadro 3.1 emiten desde los mismos puntos.

Otro elemento que puede ser obtenido de la gráfica 3.1 es la cantidad de usuarios que solo han emitido una vez. La cantidad de usuarios con esta característica representa alrededor de una cuarta parte de la muestra. La justificación de eliminar este tipo de usuarios radica en la posibilidad de estudiar aquellos perfiles que son recurrentes en sus publicaciones y no usuarios que han creado una cuenta pero no lo utilizan como medio de interacción en el terreno digital. Esta metodología de depuración es propuesta por Lansley y Longley (2016), quienes sugieren eliminar aquellos usuarios que solo emitieron una publicación, así como aquellos tweets que tengan menos de tres palabras, ya que éstos no aportan mucha información analítica. En particular, esta depuración se dirige al estudio de texto de los tweets. Dimensión que se desarrollará en el siguiente capítulo.

Figura 3.2: Ubicación geográfica de cuentas automatizadas



Fuente: elaboración propia

### 3.3.2 Identificación de colonias gentrificadas

La identificación cuantitativa de espacios gentrificados es un elemento muy debatido en la literatura especializada sobre el tema. En particular, el primer obstáculo que se enfrenta al llevar a cabo este ejercicio es la escala geográfica de los datos disponibles. Como se ha mencionado en el capítulo 1, la escala territorial para entender la gentrificación es la colonia, barrio o vecindario. El segundo problema que se ha identificado para la cuantificación ha sido la traducción del fenómeno de desplazamiento de población de los barrios gentrificados, en otras palabras, el cambio poblacional. En tercer lugar, la falta de estadísticas particulares sobre las dimensiones teorizadas sobre la gentrificación (Atkinson, 2000). Para el caso de la Ciudad de México, Bournazou (2016) ha registrado las Áreas Geoestadísticas Básicas y su nivel de gentrificación con base en características sociodemográficas como edad, disponibilidad de computadora, número de habitantes por recámara, entre otras. Sin embargo, el proceso de gentrificación no solo comprende las variables las transformaciones sociodemográficas o de precio de la vivienda, sino dimensiones culturales, de intervención estatal o renovación del espacio urbano.

Como se mencionó en el capítulo 1 el fenómeno de gentrificación ha sido analizado desde diversas aristas, cada una de ellas ha apuntado a distintos aspectos de la gentrificación como segmentador y diferenciador del espacio urbano. En este sentido, se ha mencionado previamente que el presente trabajo se distancia de la definición de gentrificación como solo expulsión de población de estratos bajos en favor de población de estratos altos en las colonias gentrificadas. Aunque no se

niega esta característica, el presente estudio se enfoca hacia la caracterización de la gentrificación desde perspectivas comerciales y de intervencionismo estatal como indicadores de gentrificación en las colonias. Así, en la determinación de colonias gentrificadas se recuperan variables empíricas que aproximen dimensiones como la renovación del espacio urbano construido o la vocación comercial de éstas

Por un lado, se propone una aproximación a la gentrificación económica por medio del índice de centralidad económica. Esta propuesta recupera la noción del proceso de centro/periferia propuesta por N. Smith (1982) como fuerza de la gentrificación en el espacio urbano y se conjunta con la literatura sobre centralidad económica como polos de atracción y concentración de determinadas actividades económicas en una ciudad (García-López y Muñiz, 2013; Garrocho-Rangel y Campos, 2007). En particular, el índice de centralidad económica busca identificar colonias en las cuales la concentración de servicios vinculados con el proceso de gentrificación sea considerable. Para construir el dicho índice, se ha tomado el personal ocupado de las actividades económicas seleccionadas, del Directorio Estadístico Nacional de Unidades Económicas (DENUE) de 2017. Estas actividades tratan de reflejar el consumo servicios de restaurantes, bares, servicios personales como salones de belleza, masajes, comercio al por menor de ropa y accesorios, ver figura A.1 del apartado de Anexos.

La selección de comercios para la construcción del índice de centralidad se basa en la literatura sobre gentrificación comercial y cultural. La literatura internacional y nacional apunta a que el comercio en los espacios gentrificados adquieren un papel dual en el imaginario de consumo. Por un lado, este tipo de espacios recuperan la noción de comercio local, en términos espaciales y económicos (Benediktsson, Lamberta, y Larsen, 2016; de Lourdes Cordero-Gómez y Salinas, 2017; Díaz-Parra y Salinas, 2016). Así, en numerosas áreas alrededor del mundo salones de belleza o masajes de élite se han convertido en ícono de este tipo de espacios. Por otro lado, las zonas gentrificadas manifiestan cierto nivel de internacionalización en sus comercios, de esta manera se recuperan bares o restaurantes internacionales. Aunque de manera individual los restaurantes o salones de belleza se localicen de manera generaliza a lo largo del espacio, la coexistencia del conjunto de actividades seleccionadas puede ser indicador de cierto nivel de gentrificación.

Para determinar si una colonia es central en comercio y servicios gentrificados se ha utilizado el índice de doble umbral. De acuerdo a este indicador, la colonia es considerada central si cumple

con dos criterios. El primero de ellos es que la densidad de empleo por kilómetro cuadrado en los servicios de interés sea mayor a la densidad promedio de dichos servicios en la totalidad del área de la ZMCM. El segundo criterio que debe cumplirse es que la magnitud de empleo en la colonia sea mayor que el promedio reportado en toda la ZMCM más una desviación estándar. En términos formales el índice de doble umbral para identificar centralidad económica se define como (McDonald, 1987; Méndez, 1996):

$$c_i = \begin{cases} 1 = \begin{cases} \text{si } D_{ij} > D_j \text{ y} \\ \text{si } M_{ij} > E(M_j) + DE(M_j) \end{cases} \\ 0 = \text{de cualquier otro modo} \end{cases} \quad (3.1)$$

Donde:

$c_i$  = centralidad económica de servicios gentrificados de la colonia  $i$

$D_{ij}$  = Densidad de empleo en la colonia  $i$  en los servicios  $j$  en  $km^2$

$M_{ij}$  = la magnitud del empleo en la colonia  $i$  en los servicios  $j$

$E(M_j)$  = valor esperado de la magnitud del empleo en los servicios  $j$  en toda la ZMCM

$DE(M_j)$  = desviación estándar de la magnitud del empleo en los servicios  $j$  en toda la ZMCM

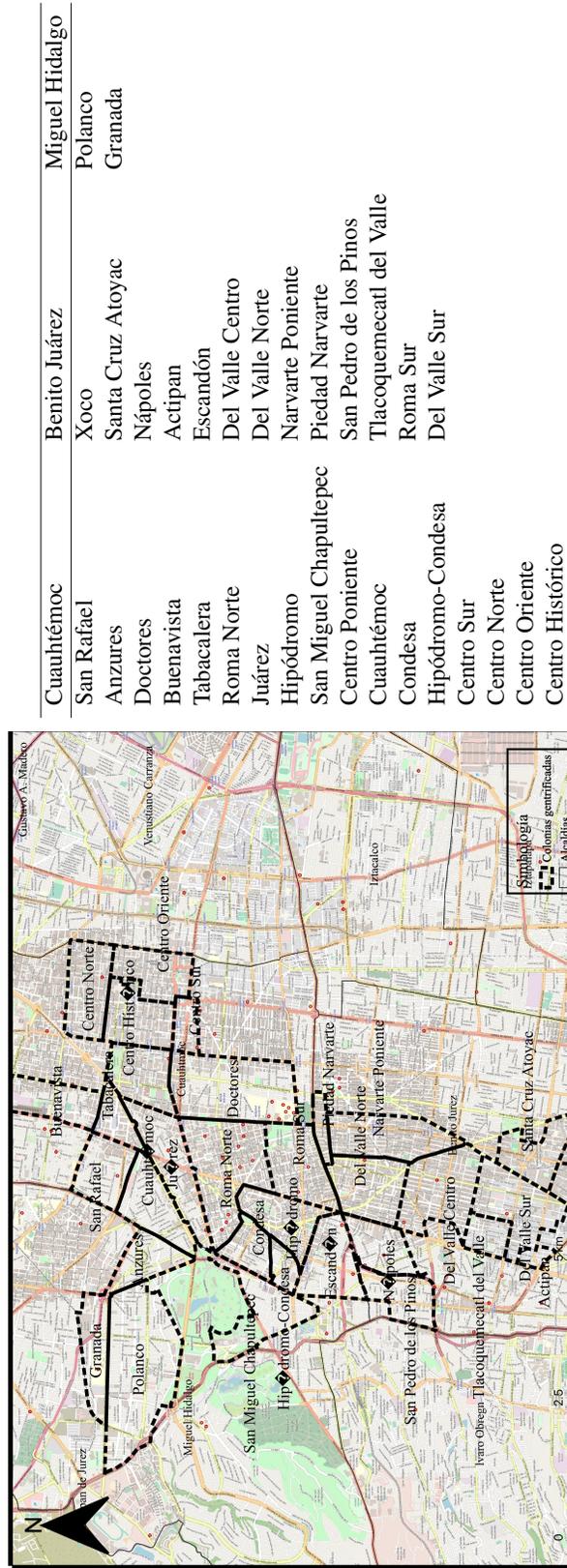
Por su parte, la gentrificación guiada por el Estado es mucho más difícil de medir empíricamente. Una propuesta ha sido medir el incremento la recaudación de impuestos en colonias que han sufrido procesos de renovación (Grange y Pretorius, 2016). Para el caso de este trabajo, el indicador de gentrificación guiada por el Estado se han tomado las estaciones del programa de bicicletas públicas compartidas, ECOBICI. Esta información ha sido recabada del sitio de datos abiertos de la Ciudad de México (ECOBICI, s.f.). Esto con la premisa de que el Estado ha colocado estas estaciones, no en función de los puntos más favorables para la movilidad intraurbana, sino para favorecer revalorización de espacios centrales en la ciudad.

Con ambas dimensiones se ha construido un criterio de identificación de colonias gentrificadas basado en la combinación del índice de centralidad económica de servicios gentrificados y la presencia de estaciones de bicicletas públicas gubernamentales. De esta manera, una colonia gentrificada debe de cumplir con el índice de doble umbral y, al menos, tener una estación de ECOBICI. Con estos criterios, se han identificado 32 colonias. colonias ubicadas en tres alcaldías centrales de la Ciudad de México, Cuauhtémoc, Benito Juárez y Miguel Hidalgo. La ubicación de estas colonias

se presenta en la figura 3.3.

Los estudios previos sobre la identificación de espacios gentrificados en la Ciudad de México se caracterizan por dos vertientes. Por un lado se tienen los estudios abocados a la delimitación cuantitativa los espacios de interés. Dichos estudios se enfocan en construir indicadores de cambio poblacional basado en estadísticas oficiales. En este sentido, se tiene el trabajo de Sobrino (2016) quien identifica el cambio de estructura poblacional en cuatro alcaldías de la Ciudad de México, dos de ellas centrales. Aunque la identificación del cambio poblacional se basa en la concepción clásica de gentrificación una limitación del estudio es la escala geográfica de análisis, ya que, se realiza a escala de municipalidad, una escala demasiado grande para identificar los cambios poblacionales de las colonias. En este mismo esfuerzo de identificación cuantitativa se tiene el trabajo de Bournazou (2016) quien identifica AGEBs en la Ciudad con procesos de cambio sociodemográfico. La autora identifica las AGEB con mayor estabilidad en las variables sociodemográficas y de viviendas seleccionadas y por el otro las AGEB con mayor cambio en estas dimensiones.

Figura 3.3: Colonias identificadas como gentrificadas



Fuente: Elaboración propia, con base en DENUE, 2017 y estaciones de ECOBICI

Por otro lado, se encuentran los estudios cualitativos que han estudiado a profundidad espacios y colonias determinadas de la ciudad. Con respecto a este vertiente de trabajos se ubican los trabajos de (Hernández-Cordero, 2015; Salinas, 2013a) quienes han estudiado el proceso de transformación, reconversión y valorización de espacios en las colonias centro, específicamente la Alameda central. Otro espacio de interés ha sido la colonias Condesa, Condesa-Hipódromo y Roma, espacios identificados como icónicos del proceso de gentrificación por la concentración de restaurantes *gourmet*, centro de diversión, y por la recuperación de su espacio construido (Díaz-Parra y Salinas, 2016; Duhau y Giglia, 2016; Salinas, 2013b). O el desarrollo de nuevos espacios gentrificados que han experimentado la dispersión y contagio de espacio de élite, tales como el Nuevo Polanco, espacio adyacente a la zona de Polanco o bien la colonia Doctores, adyacente a la colonia Roma (Aguayo-Ayala, 2016; Rodríguez, 2013). Una ventaja de este tipo de aproximaciones es la identificación del proceso *in situ* y la demarcación espacial del proceso de gentrificación a escala de barrio, colonia o, incluso, calle.

La identificación realizada en este trabajo comprende 33 colonias de tres alcaldías de la Ciudad de México, todas ellas contiguas y continuas espacialmente. Esta delimitación enfrenta una limitación en comparación con las realizadas por Bournazou (2016); Sobrino (2016) y es su carácter temporalmente estático, ya que solo da muestra de la centralidad comercial y distribución de bicicletas públicas en un punto en el tiempo, 2017, mientras que los estudios previos construyen indicadores diacrónicos del fenómeno. Sin embargo, este trabajo resalta la unidad espacial de colonia como primordial para la identificación del proceso de gentrificación, al mismo tiempo, se enfoca en variables de tipo comercial, mas que de tipo sociodemográfico y de vivienda. Esto posibilita vislumbrar colonias que en el estudio de Bournazou (2016) no son identificadas como gentrificadas, tal es el caso de Polanco o el centro o algunas del corredor Insurgentes Sur. En este sentido, la identificación hecha aquí concuerda con elementos resaltados en los estudios cualitativos. Así espacios como Polanco, Centro Histórico o Nápoles representan puntos de concentración de poder económico, político y comercial en la ciudad, lo que compagina empíricamente con los desarrollo teóricos de la gentrificación como desarrollo urbano desigual de Neil Smith.

Finalmente, la identificación aquí propuesta no es definitiva y está sujeta a múltiples críticas y mejoras. Sin embargo, ésta fortalece la noción de gentrificación comercial, económica y cultural como mecanismos paralelos al recambio población, previamente estudiado. Al mismo tiempo,

recupera la noción de colonia como unidad de análisis espacial fundamental para el estudio de la gentrificación.

### *3.3.3 Construcción de indicadores espacio-temporales de la emisión de tweets en la ZMCM*

Uno de los elementos que Twitter permite resaltar, en cuanto a la disponibilidad de su información, es la temporalidad y localización de las emisiones. Así, estudios como los de Crampton y cols. (2013); García-Palomares y cols. (2018); Shelton (2015); Shelton, Poorthuis, y Zook (2015); Neuhaus (2015) enfocan sus propuestas metodológicas hacia instrumentos para estudiar la dimensión temporal de la expresión espacial de Twitter. Lo que caracteriza a cada uno de estos trabajos es el objetivo de analizar el patrón de emisiones en función de fenómenos urbanos específicos. Así, para este trabajo se recupera la dimensión espacial de la emisión de tweet debido a que el objetivo es analizar la posible relación entre la emisión de tweets y el proceso de gentrificación en colonias centrales, previamente identificadas, de la ZMCM. En este sentido, la dimensión temporal de la emisión de tweets relacionado con el proceso de gentrificación permitiría estudiar la configuración espacio-temporal de los espacios gentrificados.

Para la construcción de indicadores espaciales y temporales de la emisión de tweets, este trabajo retoma la propuestas metodológicas de García-Palomares y cols. (2018) y Shelton y cols. (2015). Por su parte, García-Palomares y cols. (2018) analizan la dinámica diaria en Ciudad de Madrid y la relación entre ésta y el uso del suelo urbano. Para ello clasifican las emisiones de Twitter de la ciudad en distintos horarios con el objetivo de encontrar diferencias temporales en la emisión y, de esta manera explicar dicho patrón con los distintos usos de suelo urbano. De esta manera, los autores proponen que las zonas vinculadas a oficinas, actividades educativas o de salud muestran un patrón decreciente de emisión de tweets a lo largo del día. Mientras que, zonas con uso de suelo comercial, particularmente de comercio al menudeo, o bien, parques reportan emisión constante independientemente de la variable temporal. Aunque los autores utilizan la información de Twitter para dar cuenta de la dinámica urbana en función del espacio concreto, este estudio sugiere indirectamente la relación entre la dimensión digital, medida a través de Twitter, y el uso del espacio en la ciudad. Por su parte, Shelton y cols. (2015) exploran los imaginarios espaciales y su relación con los procesos de segregación y movilidad en la ciudad de Louisville, Kentucky. Los autores proponen realizar este estudio basado en un análisis crítico de la GeoWeb basado en la teoría

relacional socio-espacial. Para ello, analizan la emisión de tweets mediante unidades geográficas distintas a las unidades político-administrativas de la ciudad. En este sentido, el trabajo recupera la emisión como un *continuum* basado en unidades de  $1km^2$ . Los principales hallazgos del estudio apuntan a que los espacios diarios de actividad de usuarios de Twitter en la ciudad no reflejan segregación socioespacial en dicha ciudad mediante los tweets. En otras palabras, en Louisville se propone una mezcla social de estratos en la dimensión digital. En conclusión, los autores enfatizan la importancia de analizar el contexto social que se encuentra subyacente a la emisión de los tweets.

Con base en los dos estudios mencionados, el presente estudio propone la siguiente metodología para el análisis de Twitter en la ZMCM. En cuanto a la dimensión temporal de los tweets, se retoma la propuesta de García-Palomares y cols. (2018) para la categorización de horarios de la emisión de tweets, esto con el objetivo de identificar diferencias en los patrones de emisión a lo largo del día. La siguiente ecuación muestra que esta categorización responde a un criterio funcional, es decir, se considera las actividades realizadas durante un día. Así, la primera categoría responde al primer segmento laboral del día, de 7:00 a 12:59 hrs. La segunda categoría da cuenta de la segunda parte del día laboral, de 13:00 a 17:59 hrs. Por su parte, las categorías 3 y 4 tratan de dar cuenta de la vida después del trabajo en la ciudad, esto de 18:00 a 00:00 hrs. Por último, la categoría 5 captura la emisión en la madrugada <sup>6</sup>.

$$h = \begin{cases} 1 = & \text{mañana} & (\text{ si } 7:00 \text{ hrs a } 12:59\text{hrs}) \\ 2 = & \text{tarde} & (\text{ si } 13:00 \text{ hrs a } 17:59\text{hrs}) \\ 3 = & \text{tarde-noche} & (\text{ si } 18:00 \text{ hrs a } 21:59\text{hrs}) \\ 4 = & \text{noche} & (\text{ si } 22:00 \text{ hrs a } 23:59\text{hrs}) \\ 5 = & \text{madrugada} & (\text{ si } 00:00 \text{ hrs a } 06:59\text{hrs}) \end{cases} \quad (3.2)$$

Al mismo tiempo, se ha dividido temporalmente en función de los días de la semana:

$$d = \begin{cases} 1 = & \text{días laborales} & (\text{ si } \text{lunes a } \text{miércoles}) \\ 2 = & \text{fines de semana} & (\text{ si } \text{viernes a } \text{domingo}) \end{cases} \quad (3.3)$$

Esta clasificación se ha hecho con el objetivo de analizar cómo el fenómeno de gentrificación

---

<sup>6</sup>La API de Twitter proporciona la hora de emisión de la publicación en el horario del Meridiano cero de Greenwich. Es decir en horario 00:00:00 GMT. Por lo tanto, para poder realizar el análisis temporal se ha tenido que ajustar el horario de emisión a la zona horaria de México -06:00:00 GMT

se aumenta digitalmente en colonias identificadas como gentrificadas en términos temporales. En este sentido, esta clasificación se basa en los trabajos de García-Palomares y cols. (2018) Shelton y cols. (2015) quienes apuntan al análisis de la dimensión digital considerando el ensamblaje social y urbano que le subyace. En este caso, dicho ensamblaje se relaciona con el fenómeno de gentrificación y sus elementos temporales. Lo anterior se debe a que muchos de los estudios sobre gentrificación cultural, simbólica y comercial basan su análisis en determinadas actividades como el consumo en restaurantes, tiendas de ropa o bares (Minner y Shi, 2017; Zukin y cols., 2017). Sin embargo, dichas actividades se explican en función del tiempo disponible para el consumo, recreación o esparcimiento. De esta manera, al observar la conformación temporal de emisión en el mismo espacio gentrificados se podría llegar a observar, de mejor manera, la creación de espacio gentrificado mediante Twitter en ciertos horarios.

Por último, Shelton y cols. (2015) proponen obtener la densidad de los tweet sin constreñir estas a unidades espaciales político-administrativas, debido que este fenómeno no sigue estas unidades preestablecidas. Con esto en mente, se ha utilizado una función generadora de núcleos de densidad. “La función generadora de núcleos de densidad opera promediando una serie de pequeños bultos (funciones de probabilidad en dos dimensiones) centradas en cada punto observado”(Brunsdon y Comber, 2019, p: 196). La función estima la formación de núcleos de concentración para cada punto  $(x,y)$  De manera formal la función se expresa de la siguiente manera.

$$f(X) = f(x,y) = \frac{1}{nh_xh_y} \sum_i k \left( \frac{x-x_i}{h_x}, \frac{y-y_i}{h_{xy}} \right) \quad (3.4)$$

La expresión  $k \left( \frac{x-x_i}{h_x}, \frac{y-y_i}{h_{xy}} \right)$  representa la formación de núcleos basados en una distancia dada. La ecuación 3.4 representa el proceso de estimación de núcleos promedio en toda la zona. Por su parte, el parámetro  $h$  refiere al radio de construcción del núcleo. Para el caso de este trabajo se ha seleccionado una distancia euclidiana de 500 metros.

Con todos los elementos empíricos tanto de identificación de colonias gentrificadas en la ZMCM como de una dimensión de su manifestación digital por medio de Twitter se han hecho las respectivas estimaciones de densidad de emisión de tweets para examinar dos elementos. El primero de ellos es examinar si las colonias identificadas como gentrificadas tienen mayores intensidades de emisión de Twitter. En segundo lugar, si esta emisión diferenciada relacionada con espacios

gentrificados difiere en el uso del espacio en términos temporales, esto con el propósito de encontrar indicios de procesos de ocupación del espacio urbano en la ZMCM.

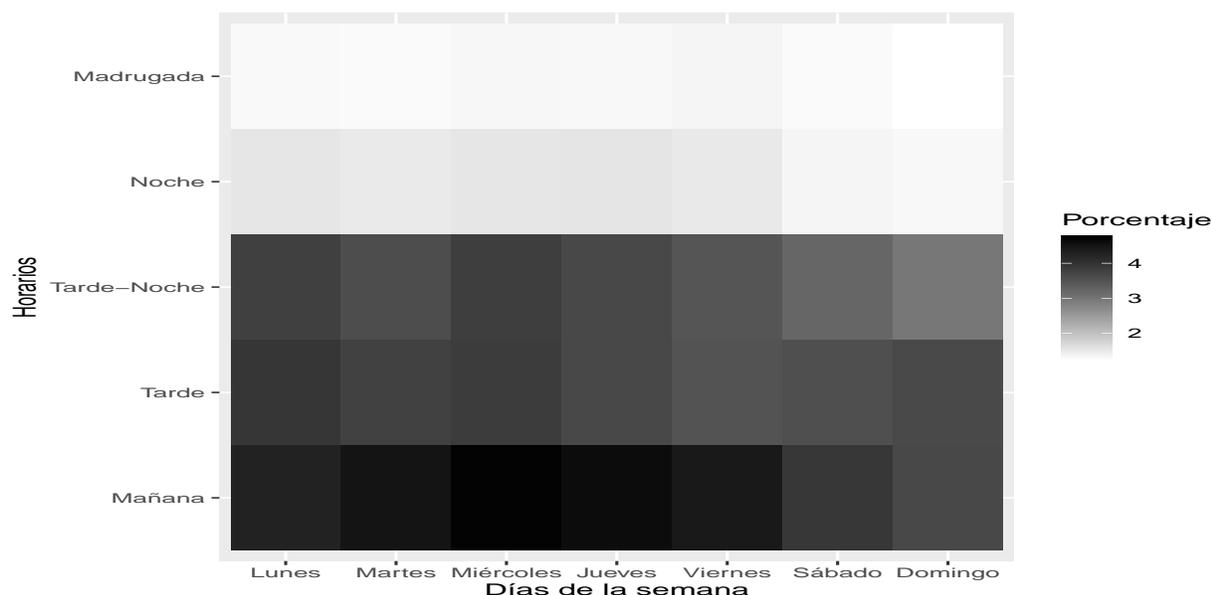
### **3.4 Concentración espacio-temporal de Twitter y su co-incidencia en espacios gentrificados**

La ciudad es un complejo entramado de interacciones que construyen diferentes espacios. Estas interacciones y entramados no solo se componen de elementos espaciales y sociales, sino también de dimensiones temporales. En este sentido, como se ha mencionado anteriormente, la ciudad se transforma a lo largo del día. Esta dimensión se expresa de manera palpable en las dinámicas de emisión de Twitter en la ZMCM. Así, el propósito de esta sección es dar cuenta de la conformación del poder de distribución espacial y de comunicación en la emisión de Twitter en la ciudad y cómo ésta aumenta digitalmente el fenómeno de gentrificación.

En este trabajo se analizan las combinaciones temporales de emisión en la ZMCM. Como se ha visto en el apartado metodológico de este capítulo, se han clasificado los tweets en función de la hora de su emisión, así como del día de la misma. De esta manera, se tienen 10 posibles combinaciones temporales. Estas combinaciones se explican con base en que el fenómeno de gentrificación tiene una fuerte carga temporal en cuanto a la construcción de dinámicas sociales de gentrificación en colonias identificadas con dicho proceso. Sin embargo, en primera instancia se analiza la temporalidad del total de emisiones en la ZMCM.

En la figura 3.4 se muestra el porcentaje de emisión de tweets en el total ZMCM, esto contempla tanto tweets georreferenciados como los que no cumplen con esta característica. El primer elemento a destacar es que la emisión sigue un comportamiento diurno, ya que la mayoría de publicaciones se generan durante horas con actividad social y económica. Dentro de los horarios diurnos resalta el horario que comprende de 7:00 a 13:00 hrs, ya que este captura el mayor porcentaje de emisiones a lo largo de toda la semana. Dentro de este horario también sobresalen días específicos, éstos son de miércoles a viernes, días que llegan a tener más de 4% de la emisión total. Por su parte, el horario que reporta de manera constante emisión de tweets durante todos los días de la semana es de 13:00 a 17:59 hrs (tarde), siendo este horario el que puede llegar a combinar actividades tanto laborales como recreativas. En la misma tendencia se tiene el horario de tarde-noche, el cual reporta una emisión de tweets de alrededor de 3% durante los días laborales, descendiendo durante sábado y domingo.

Figura 3.4: Porcentaje de emisión en la ZMCM según día y hora



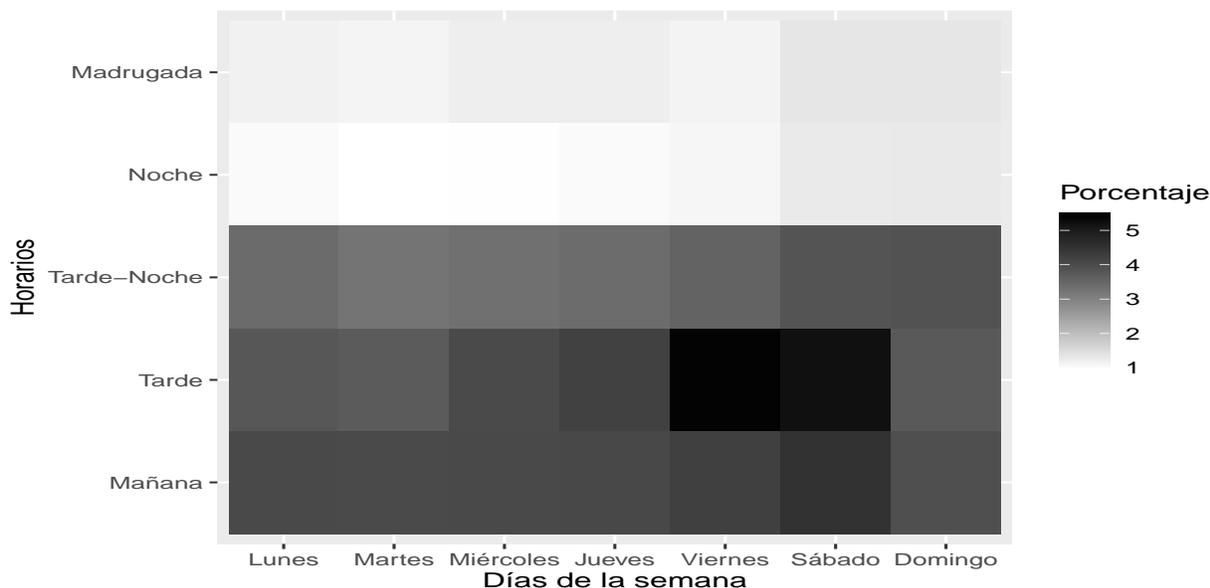
Fuente: Elaboración propia, con base API de Twitter

En este punto, se podría hablar de la configuración del poder digital distribuido temporalmente. Ya que, en primera instancia, se tiene que el comportamiento digital se asocia con el ciclo de vida diario de las personas. Por otro lado, dentro del comportamiento diurno la emisión también presenta heterogeneidad cuando se analiza a detalle. Esta simple representación sugiere que existen otros elementos que explican este diferencial a lo largo del día y la semana en horarios diurnos. De otra manera, la emisión digital debería de ser constante durante todo el día y la semana.

Dentro de este universo de emisión se encuentran las publicaciones que son explícitamente georreferenciadas. Como se ha mencionado anteriormente, este porcentaje representa solo 16% de la emisión total y la explicación de que las personas otorguen este tipo de información, aun siguen siendo poco exploradas. Pese a esto, la información que emana de los tweets georreferenciados permite observar patrones de comportamiento diferentes con respecto al total de tweets. En la figura 3.5 se muestra el porcentaje de emisión en la ZMCM solo de tweets georreferenciados.

El comportamiento de emisión entre los tweets georreferenciados con respecto del total difiere sustancialmente. En primer lugar, ya se ha hablado de las motivaciones de las personas para reportar voluntariamente información georreferenciada. Aunque la distribución de los tweets georreferenciados no se distribuye homogéneamente a lo largo del tiempo. La figura 3.5 muestra que la

Figura 3.5: Porcentaje de emisión de tweet georreferenciados en la ZMCM según día y hora



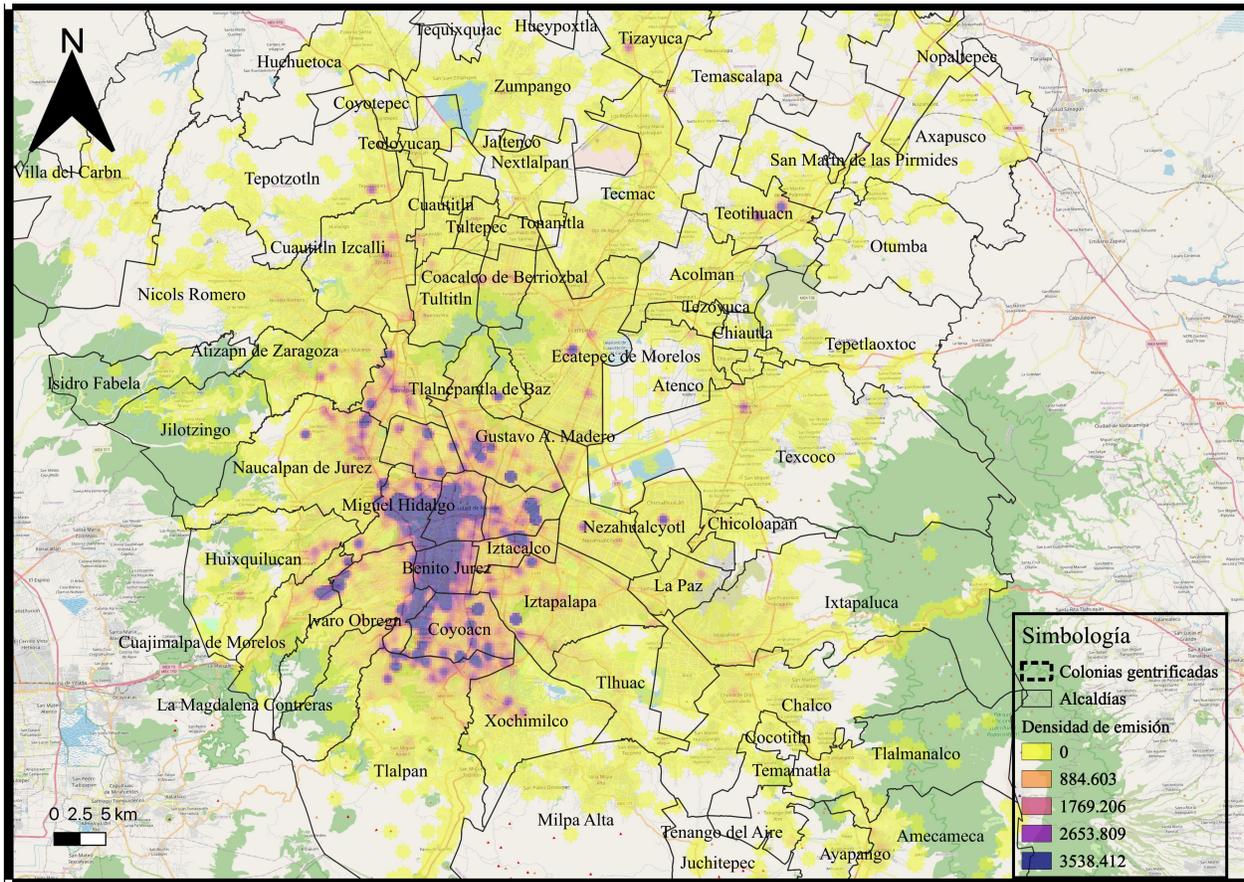
Fuente: Elaboración propia, con base API de Twitter

mayoría de tweets que contienen su ubicación exacta se emiten durante el día. Por ejemplo, durante los días viernes y sábados en el horario de 13:00 a 18:00 horas se llega a alcanzar alrededor de 10% de las publicaciones totales georreferenciadas. Otro dato a resaltar es el incremento de publicaciones durante la tarde-noche. Mientras que, de lunes a jueves el porcentaje de publicaciones oscila entre 2% y 3%, en este mismo horario los días de fines de semana se incrementa hasta 4%.

En resumen, la figura 3.5 permite observar asociación temporal entre la emisión de días y horas particulares de la semana. Esto posibilita sugerir que Twitter se relaciona con actividades cotidianas de sus usuarios en el espacio de ZMCM. Para sustentar esta idea, se analiza el comportamiento espacio-temporal de la emisión en la ZMCM. Con base en esto, a continuación, se presentan los diferentes niveles de concentración de emisión en la ciudad. Para realizar este análisis se exponen los 5 diferentes horarios y se compara la manera en que cambian las concentración en función del periodo de la semana, es decir, si las emisiones se realizan en días laborales o fines de semana.

La figura 3.6 muestra los principales núcleos de emisión de Twitter en la ZMCM durante los dos años de recolección. Con base en la delimitación espacial hecha en la identificación de colonias gentrificadas, el análisis de las concentraciones de emisión apunta en dos vertientes. Por un lado se enfatiza la gran concentración de la emisión en las zonas centrales de la zona metropolitana.

Figura 3.6: Núcleos de emisión de tweets en la ZMCM



Son tres las alcaldías que concentran la emisión de tweet en la ciudad, éstas son Miguel Hidalgo, Cuauhtémoc y Benito Juárez. Si se observa la identificación de colonias gentrificadas, éstas comparten con el principal núcleo de emisión de Twitter. Aunado a eso, se puede observar un fenómeno de contagio a las zonas contiguas a estas colonias, sobre todo hacia el sur de las alcaldías centrales, Coyoacán y Álvaro Obregón.

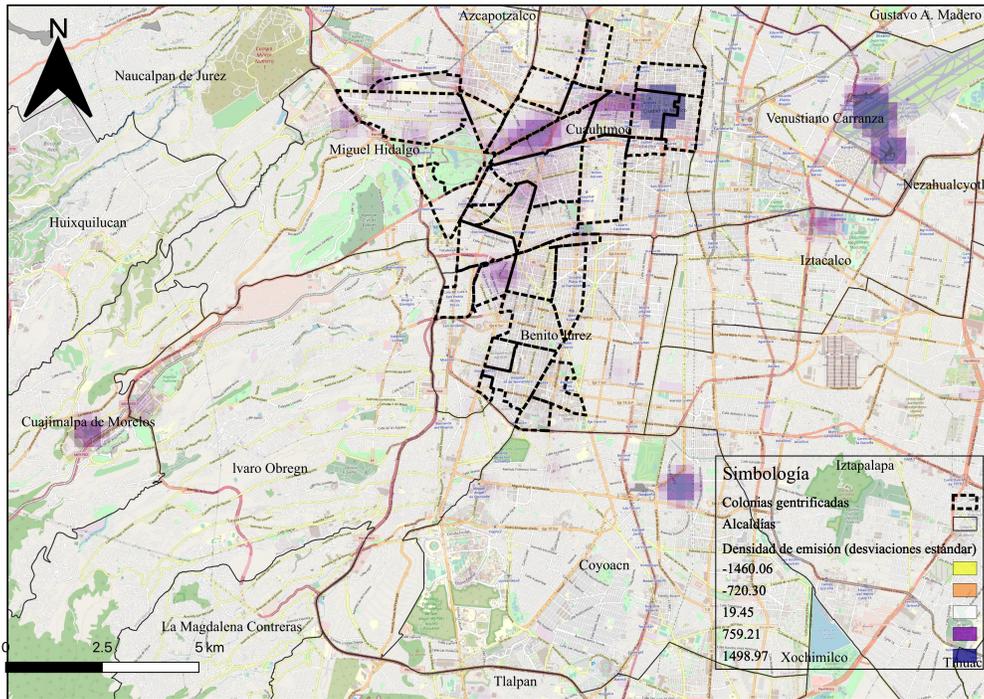
Por otro lado, se observan puntos de concentración fuera de las colonias centrales. Algunos puntos sobresalientes en este patrón de emisión son la zona de Santa Fe al occidente de la ZMCM, al sur sobre el corredor de Periférico y al oriente de la ciudad, particularmente en los puntos del Aeropuerto Internacional de la Ciudad de México, la zona del Palacio de los Deportes, Foro Sol y Autódromo Hermanos Rodríguez. Un patrón de concentración a resaltar es el observado en las delegaciones del norte de la Ciudad de México y municipios del Estado de México. En esto se observa una dispersión de puntos de concentración, particularmente en las áreas que son

cabeceras municipales como en Cuautitlán Izcalli, Ecatepec o Tlalnepantla. Este punto puede estar relacionado con dos aspectos del poder de la RGAD. Por un lado la distribución de recursos de infraestructura y de poder político en los centros de los municipios, particularmente en las oficinas de gobierno municipal que permite tener acceso a Internet. El segundo grupo de emisiones fuera de las colonias centrales se encuentran las emisiones que se asocian a actividades comerciales específicas. Dentro de este punto se encuentra el clúster de la zona de Naucalpan y Ciudad Satélite al nororiente de la ZMCM. En este punto se visualiza un corredor de emisiones que se localizan en zonas comerciales y educativas de la zona metropolitana. Otro ejemplo de emisión relacionado con actividades específicas son los pequeños núcleos que se muestran en el municipio de Teotihuacán, particularmente en la zona arqueológica de dicho municipio.

Con estos primeros hallazgos se comienza a vislumbrar, que la distribución digital de Twitter en el espacio urbano no es homogénea. En primera instancia, que la distribución de la emisión digital se concentra en determinadas zonas de la ciudad con vocaciones económicas y políticas específicas. En este mismo sentido, que Twitter puede ser un reflejo de estos núcleos de poder. En segundo término, se comienza a observar algunas de las motivaciones de la IGV para entender por qué las personas emiten públicamente su geolocalización en Twitter. En este punto se observa que está relacionado con actividades de consumo, ya sea en centros comerciales o turísticos. Sin embargo, la dimensión digital no solo reproduce las actividades en el espacio, también se producen imaginarios digitales de dichos espacios de concentración de emisión.

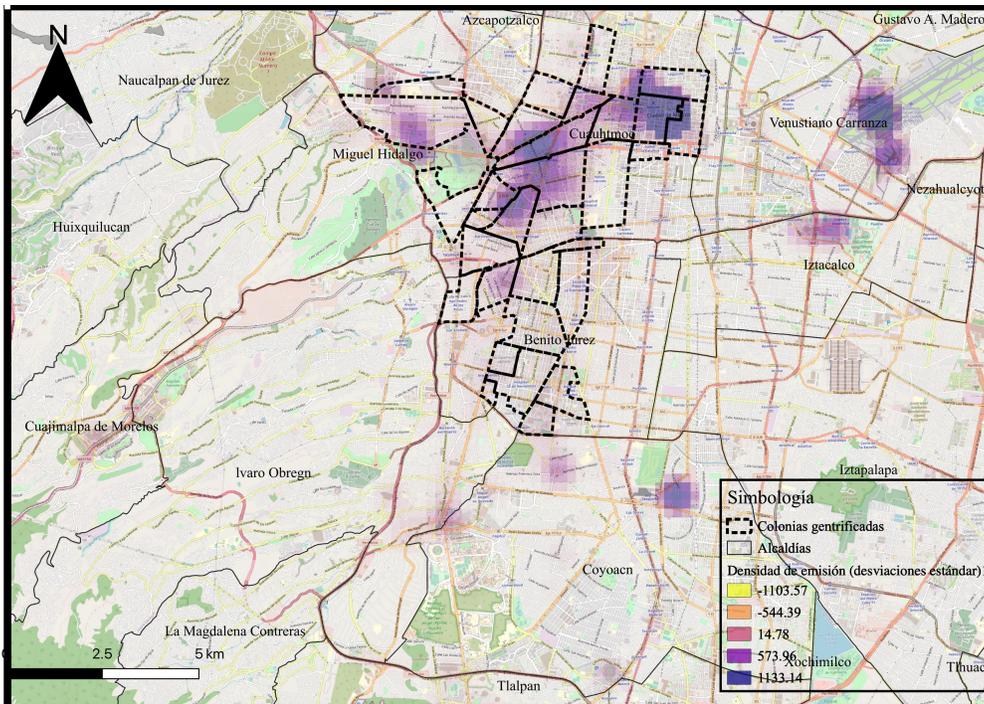
Finalmente, en la figura 3.6 no se distingue ninguna dimensión temporal expuesta en la sección metodológica de este capítulo. Este análisis se realiza a continuación, con el objetivo de controlar las variables temporales y observar como se configura la geografía de la emisión de twitter en la ciudad en diferentes horarios. Esto permitirá ver cómo la dimensión digital de los procesos urbanos en Twitter se compagina con las actividades cotidianas de la vida urbana. En este sentido, el análisis espacial controlado por el aspecto temporal posibilita observar la conformación de espacios sociales y digitales diferenciados en el mismo espacio físico y, la conformación de desigualdad espacial en la distribución espacial de la dimensión digital de manera más específica.

Figura 3.7: Núcleos de emisión de tweets en horarios de 7:00 - 12:59 hrs en la ZMCM



Fuente: Elaboración propia

(a) Días laborales



Fuente: Elaboración propia

(b) Fines de semana

Las figuras 3.7a y 3.7b muestran los niveles de concentración de tweets durante la mañana<sup>7</sup>. Por un lado, la concentración de emisiones durante las mañanas de días laborales se concentra en cuatro zonas de la ciudad (ver figura 3.7a). En la parte oriente se concentra en la alcaldía Venustiano Carranza. Esta emisión es explicada por el Aeropuerto Internacional de la Ciudad de México. El siguiente punto a resaltar es el ubicado en la parte sur-oriente de la ciudad en la alcaldía de Coyoacán, este punto corresponde al Centro de Transferencia Modal (CETRAM) de Taxqueña. En este CETRAM confluyen medios de transporte como Metro, Tren Ligero, autobuses urbanos y foráneos. La cuarta concentración a destacar es la ubicada en el occidente de la ciudad, en la alcaldía Cuajimalpa. Dicha concentración se explica por la ubicación del Centro Comercial Santa Fe, así como la presencia de diversos corporativos en la zona. La quinta concentración se ubica en el corredor de oriente-poniente que se ubica en la parte norte de las colonias identificadas como gentrificadas. Este corredor de alta concentración de tweets abarca colonias como el Centro Histórico, Polanco, Juárez o Cuauhtémoc.

Dentro de las colonias gentrificadas se resaltan dos colonias particulares en el horario laboral. Por una lado tres puntos de emisión en la colonia Polanco. Al occidente de la colonia, en el área del centro comercial Palmas Park, lugar caracterizado por localizar a gimnasios y cafeterías, actividades cotidianas. Otro punto de emisión significativa es al norponiente de la colonia en la colonia Polanco y Granada, este punto refleja la localización de Plaza Antara y Carso, lugares de concentración de oficinas centrales y servicios. Por último, al sur de la colonia Polanco se tiene el punto del Hotel Marriot polanco y sus zonas contiguas. Algunos usuarios realizan explícitamente el registro de actividad en Por otro lado, se muestra la concentración de gran emisión en el límite de las colonias Cuauhtémoc y Juárez, particularmente el Paseo de la Reforma.

Por otro lado, la figura 3.7b muestra los niveles de emisión durante horarios matutinos en fines de semana, el patrón espacial cambia con respecto a la figura 3.7a En primera instancia, se conservan *grosso modo* las áreas con mayor emisión en la ZMCM. Es decir, la zona del Aeropuerto Internacional, Taxqueña y algunas colonias identificadas como gentrificadas. En particular las colonias del Centro Histórico y Polanco. La zona de Santa Fe durante los fines de semana, reduce la concentración de emisión. De igual manera desaparecen los núcleos de concentración del sur de la

---

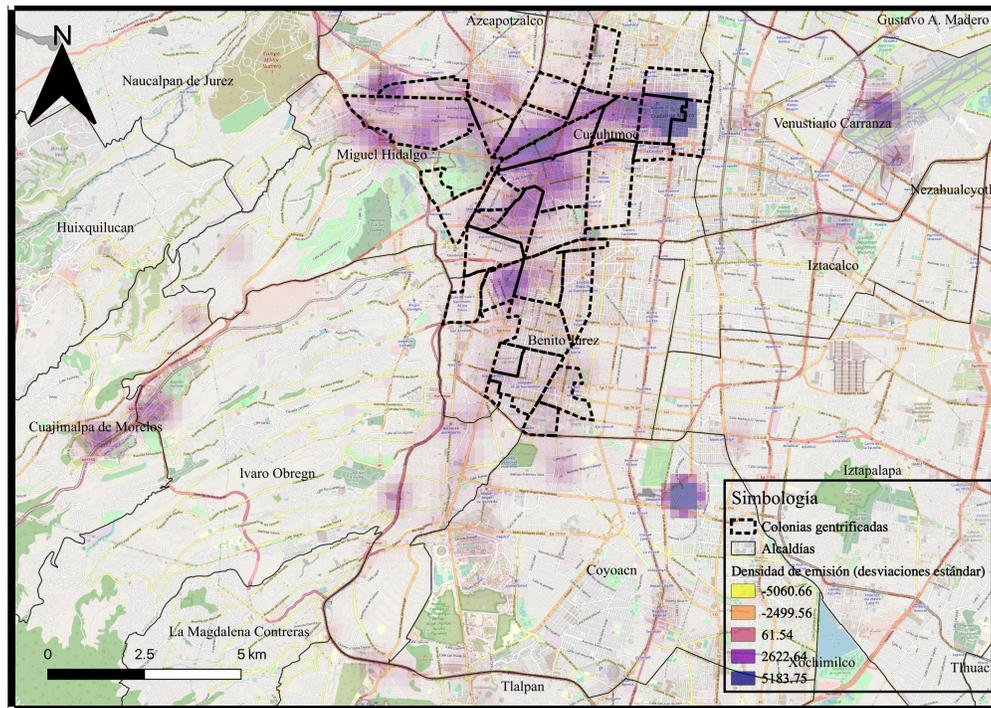
<sup>7</sup>La representación cartográfica de los núcleos de densidad serán mostradas en desviaciones estándar con respecto a la media de emisiones en cada horario y días de la semana

colonia Polanco y se concentra únicamente en la parte sur, cerca del hotel Marriot y del auditorio nacional. Por su parte, el núcleo formado desde la colonia Juárez se extiende desde el pase de la Reforma, pasando por la colonias Roma e Hipódromo-Condesa.

Fuera de las colonias centrales se forman núcleos de emisión importante. Al sur de la ZMCM se tienen tres densidades importantes. Las dos primeras dentro de la delegación Coyoacán, éstas corresponden al centro de la delegación y la segunda al estadio olímpico universitario. La tercera aglomeración se encuentra en los límites de alcaldías de Coyoacán y Tlalpan y corresponde al centro comercial Plaza Acoxta.

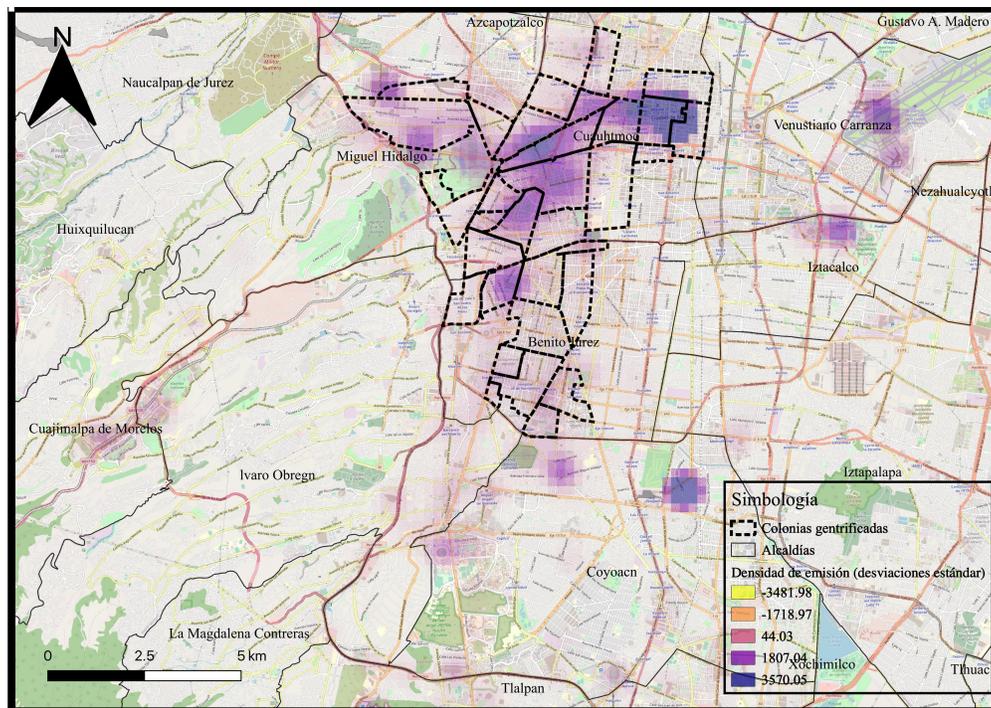
Los mapas de la figura 3.7 muestran que la emisión de tweets durante la mañana siguen dos dinámicas claramente marcadas. En primera instancia, durante las mañanas de días laborales la emisión se concentra en zonas vinculadas a dinámicas laborales, particularmente de servicios corporativos. Por otro lado, la emisión en el mismo horario, pero en fines de semana, se relaciona con actividades recreativas y esparcimiento. En este rubro, se encuentran lugares de consumo y recreación propios de horarios diurnos, tales como parques o centros comerciales, ya sea en las zonas gentrificadas o bien en puntos específicos fuera de éstas.

Figura 3.8: Núcleos de emisión de tweets en horarios de 13:00 - 17:59 hrs en la ZMCM



Fuente: Elaboración propia

(a) Días laborales



Fuente: Elaboración propia

(b) Fines de semana

La figura 3.8 muestra los núcleos de densidad de la emisión de tweets en el horario de 13:00 a 17:59 horas. Por una parte, el mapa 3.8a representa los principales núcleos de emisión en días laborales. En este figura se puede observar la formación de núcleos dentro y fuera de las colonias gentrificadas. Dentro de las colonias centrales se resaltan las emisiones en las colonias Polanco, Centro, Juárez, Tabacalera, Roma Norte Hipódromo-Condesa y la colonia Nápoles. En el caso de la colonia Polanco muestra emisión considerable en todo el polígono que la compone, esto da muestra que la actividad de digital de Twitter en esta colonia esta asociada directamente con actividad de servicios y oficinas. El mismo caso se tiene en el correo de Paseo de la Reforma. Este avenida muestra un corredor de emisión debido a la ubicación de las oficinas centrales de varias empresas internacionales como HSBC, BBVA, la Bolsa Mexicana de Valores y hoteles de lujo. El caso particular de las colonias centro y el corredor que conecta la plaza principal o zócalo se asocia a publicaciones de tipo turístico. La colonia Nápoles muestra otro punto importante de emisión, esto se debe a la ubicación del *World Trade Center* que muchos usuarios reportan haber visitado en sus cuentas de Twitter.

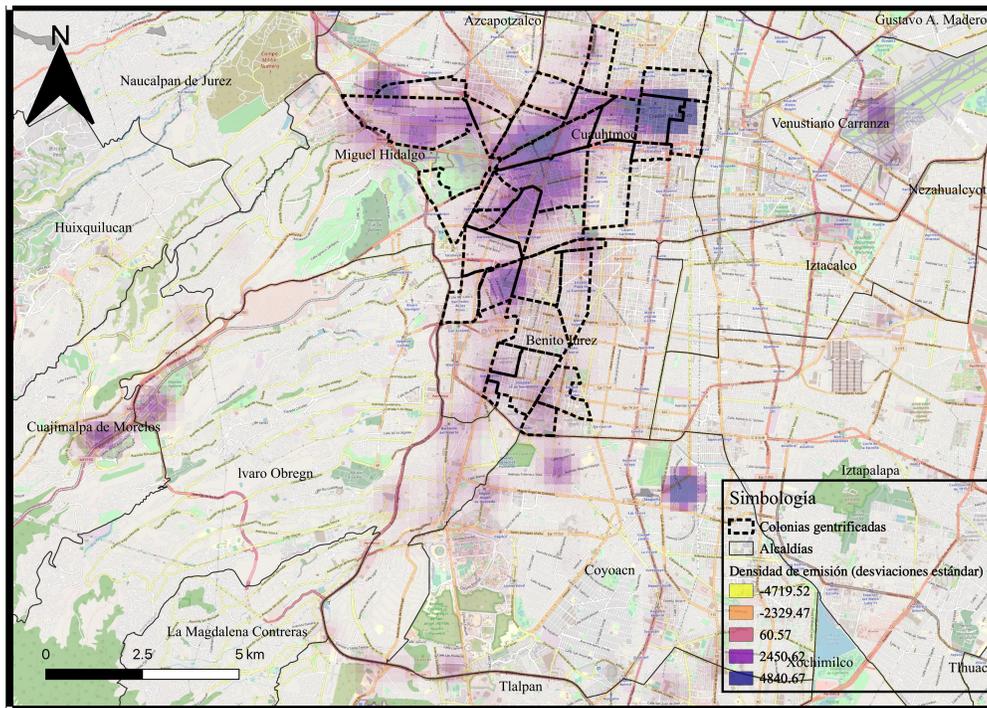
Por otro lado, el mapa 3.8a manifiesta la formación de núcleos de emisión fuera de las colonias gentrificadas. Al poniente de la ciudad se encuentra identificado la zona de Santa Fe, núcleo que se mantiene constante en horarios y días laborales. Por el lado oriente se conserva el Aeropuerto de la Ciudad de México, punto que tiene alta emisión de tweets debido a que los usuarios registran su actividad en ese lugar de manera automática. En la zona sur de la ciudad se tiene la mayor cantidad de puntos de emisión y la mayor dispersión geográfica. Al suroriente de la ciudad se mantiene el punto del CETRAM de Taxqueña. Al sur de la ciudad se encuentran ligeramente mostrados los puntos del centro de la delegación Coyoacán, la zona de Barranca del Muerto y las dos televisoras Televisa San Ángel y TV Azteca.

La emisión de Twitter en días no laborales durante la tarde muestra un patrón más concentrado con respecto a su par en días laborales. Dicha tendencia se puede observar la figura 3.8b. En primer lugar, la concentración en las colonias del Centro, Juárez, Roma Norte e Hipódromo Condesa se incrementa. Por su parte la colonia Polanco en este horario solo reporta dos puntos principales de emisión, Plaza Carzo y la zona del Auditorio Nacional. Fuera de las colonias gentrificadas se presentan los principales cambios. En primer lugar los puntos de emisión ligados con actividades laborales, principalmente servicios desaparecen. En su lugar, se refuerzan los puntos de

emisión vinculados con esparcimiento, tal es el caso del centro de Coyoacán y el estadio olímpico universitario. El mismo patrón se muestra al oriente de la ciudad, en esta zona aparece no solo el aeropuerto internacional sino la zona del Foro Sol y el Palacio de los deportes.

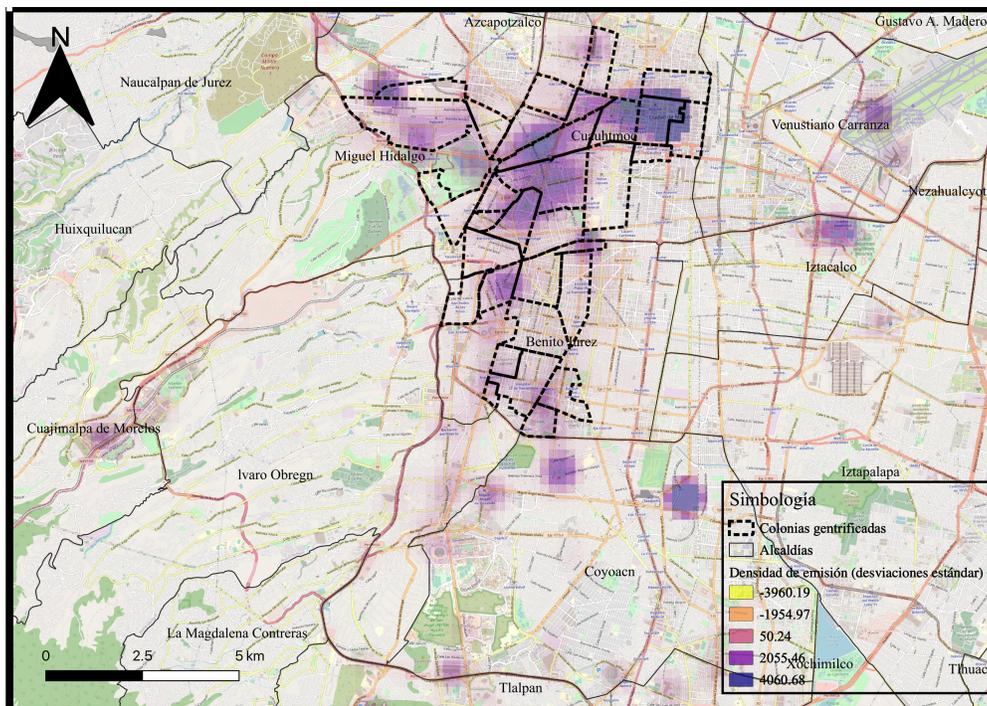
La gráfica 3.5 y el mapa 3.8b son complementarios en la conformación del poder de la distribución digital en términos temporales. Mientras que la gráfica señala que la mayor concentración de emisión de tweets georreferenciados se realizaba en horarios de tarde los días viernes y sábados, el mapa 3.8b muestra que esta emisión no se encuentra homogéneamente distribuida en el espacio y se vincula con lugares relacionados con esparcimiento. En este sentido, este horario puede ayudar a apuntalar la hipótesis que la distribución del poder digital de las dinámicas en la ciudad de México se encuentran concentradas en determinados espacios y enfocadas a ciertas temáticas.

Figura 3.9: Núcleos de emisión de tweets en horarios de 18:00 - 21:59 hrs en la ZMCM



Fuente: Elaboración propia

(a) Días laborales



Fuente: Elaboración propia

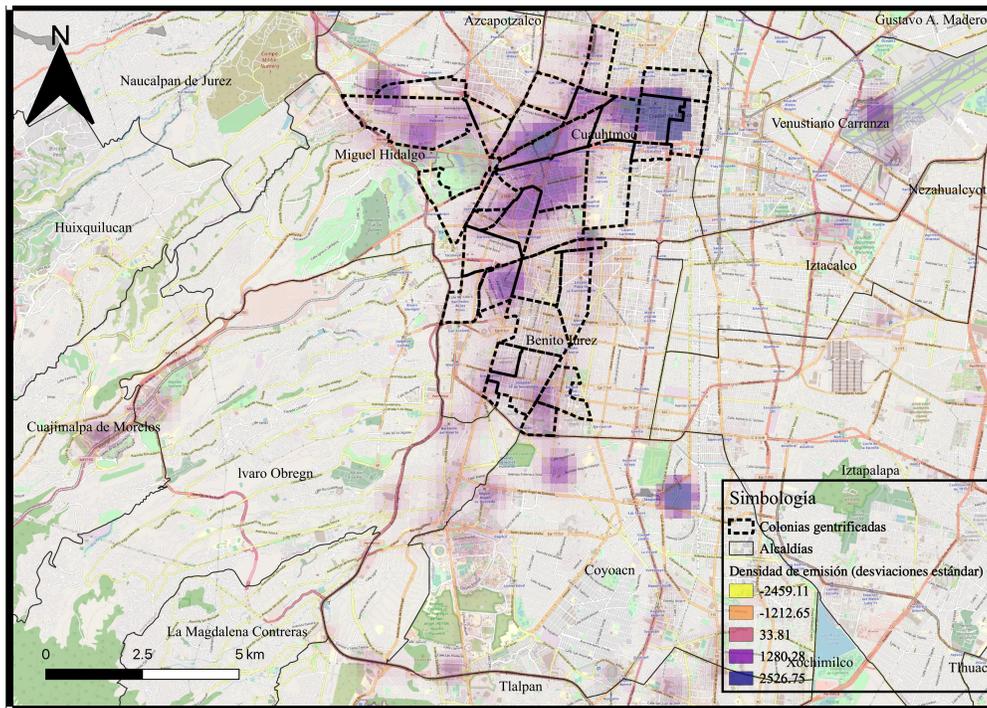
(b) Fines de semana

El siguiente horario a analizar comprende de 18:00 a 21:59 horas, los mapas 3.9a y 3.9b muestran el comportamiento espacial de la emisión en este horario. La concentración del mapa 3.9a en las colonias centrales y fuera de éstas reproduce el patrón encontrado durante los días laborales en el horario previo (ver mapa 3.8a). Es decir, la emisión durante la tarde-noche de lunes a jueves se concentra en las colonias gentrificadas, así como en las colonias adyacentes a éstas. De igual manera, se reproducen los puntos fuera de las colonias gentrificadas en el horario previo. De esta manera, aparecen como núcleos de emisión Santa Fe, Coyoacán, San Ángel, Taxqueña.

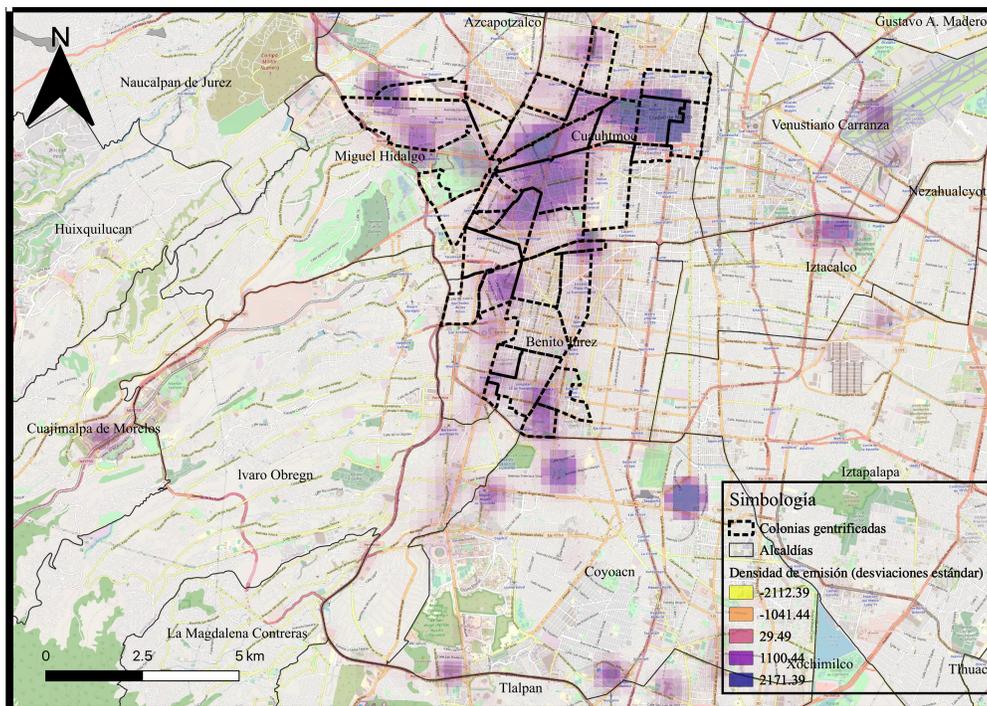
En lo que respecta al mapa 3.9b se muestra un resultado relevante al mostrar un cambio sustancial en los núcleos de emisión en días de fin de semana. Primero, los mapas 3.7b y 3.8b señalan que la emisión de tweets en sus respectivos horarios se concentra exclusivamente en algunas de las colonias identificadas como gentrificadas, este es el caso de las colonias del Centro Histórico, Juárez y Roma Norte. Sin embargo, en comparación con estos patrones, el mapa 3.9b muestra la propagación de emisión a lo largo de toda la zona central de la ZMCM. Esto podría sugerir la asociación entre la emisión de contenido en tweets y actividades específicas realizadas durante la tarde y las primeras horas de la noche. Dichas actividades se puede especular que son asociadas a consumo tales como restaurantes, bares o lugares de esparcimiento. Lo anterior debido a que este ni el horario ni los días se relacionan con actividades laborales cotidianas.

La idea del uso de Twitter vinculado a actividades de recreación y consumo, generalmente en espacios públicos de índole privado como centros comerciales, lo sugieren los núcleos de emisión fuera de las colonias gentrificadas. Por ejemplo al norte de la ciudad sobresalen tres núcleos en la alcaldía de Gustavo A. Madero. Estos núcleos corresponden a tres de los principales centros comerciales de la zona norte, Plaza Vía Vallejo, Parque Linda Vista y el Atrio de las Américas. Al oriente de la ciudad se puede encontrar una situación semejante. Resaltan, en primera instancia, los puntos del Aeropuerto, así como el punto correspondiente a la zona del Palacio de los Deportes, Foro Sol y Autódromo Hermanos Rodríguez. De igual manera en la alcaldía de Iztapalapa resalta el núcleo de emisión de Plaza Oriente. Al sur, el patrón de emisión asociado a consumo y esparcimiento se conserva. En esta zona la emisión se concentra en el corredor del centro de Coyoacán hacia la plaza comercial Oasis. Asimismo, se muestra concentración en el área de San Ángel y San Ángel Inn. Finalmente, la concentración de tweets al occidente, de nueva cuenta se relaciona con la zona de Santa Fe.

Figura 3.10: Núcleos de emisión de tweets en horarios de 22:00 - 00:00 hrs en la ZMCM



(a) Días laborales



(b) Fines de semana

Fuente: Elaboración propia

Las figuras anteriores muestran la formación de núcleos de emisión de 22:00 a 00:00 horas. El mapa 3.10a corresponde a la emisión durante días laborales, mientras que el mapa 3.10b da cuenta de la emisión en fines de semana. De manera general, ambos mapas reportan el mismo patrón espacial de conformación de núcleos. En este punto, Twitter muestra que la constitución de la ciudad aumentada digitalmente es constante a lo largo de la semana. Esto puede indicar que los usuarios de Twitter realizan actividades semejantes en horario nocturno independientemente del día. Y, además, se realizan con mayor frecuencia en las colonias centrales de la ZMCM.

En este punto, cabe resaltar que la identificación de colonias gentrificadas en este estudio se basa en la definición de centralidad urbana de actividades comerciales vinculadas a consumo de estratos elitizados o gentrificadores. Aunado a esto, el principal horario de funciones de estas actividades se realiza durante la noche (Mercado-Celis, 2018). Por lo tanto, se sugiere que la conformación constante de emisión de tweets en horario nocturno está relacionada con actividades comerciales vinculadas a patrones de gentrificación en la ciudad.

Por ejemplo, una usuaria de Twitter publicó el siguiente mensaje en días laborales en horario nocturno <sup>8</sup>. “*Cenita a gusto con la Principessa, Ana Campana, Bichito y el Maradonio #comportatemaradonio...*”. Por otro lado, otro usuario publicó en el mismo horario, pero en fin de semana: “*Tour friends en nuestra #CDMX ☺ #Chocolate #ChurrosElMoro #Zocalo...*”. Ambos testimonios de publicaciones dan cuenta del consumo en restaurantes en este horario en las colonias gentrificadas y además muestran una interacción social grupal en dicho espacios. En este punto, solo se puede conjeturar que la similitud en el patrón de emisión se explica por la semejanza en las actividades de los usuarios en colonias gentrificadas. El análisis de contenido de estos tweets se desarrolla más adelante.

La concentración de emisión en el área central de la ZMCM es constante a lo largo de todos los días y horas analizadas. En particular, la concentración de emisión se asocia espacialmente con los vecindarios identificados como gentrificados. Este patrón de concentración muestra la configuración del espacio digital y su vínculo con la construcción de espacialidades de centro periferia en la ZMCM. Según la propuesta teórica desarrollada en el capítulo 1, el fenómeno de gentrificación se ve aumentado digitalmente por medio de la interacción de fuerzas estructurales de la gentrificación, propuestas por N. Smith (1982) con las dimensiones de la realidad geográfica digitalmente

---

<sup>8</sup>Se han omitido dar usuarios de Twitter para mantener la privacidad de éstos

aumentada. En el caso del comportamiento espacio-temporal de la emisión de tweets, se puede postular la interacción entre centro/periferia y los poderes de distribución y comunicación.

La emisión de Twitter en la ZMCM es un vector que aumenta digitalmente la gentrificación en la ciudad por medio de la contradicción centro/periferia en la ciudad. En primer lugar, dado el potencial de ubicuidad de la emisión de contenido digital es contradictorio pensar que la emisión de tweets siga un patrón centro/periferia. Sin embargo, esta contradicción se subsana al mostrar que dicho patrón concuerda con los espacios comercialmente activos y que han experimentado intervención estatal en el espacio público. De esta manera, la emisión de tweets es un reflejo de la constitución del espacio basado en una de sus fuerzas primigenias, el centro contra la periferia. La emisión en el área central de la ciudad es reflejo, a su vez, del poder de distribución con el cual se constituye el uso de Twitter. En otras palabras, la centralidad de emisión corresponde a espacios de consumo comercial asociados a estratos altos. Ejemplo de esto son las colonias, Condesa, Roma, Polanco. En conclusión, el centro/periferia de la ZMCM se encuentra aumentado digitalmente debido a la alta concentración de emisiones de Twitter que acontecen en las colonias centrales y a la baja emisión en el resto de la zona metropolitana.

Los hallazgos permiten establecer que la conformación de núcleos de emisión en áreas no gentrificadas corresponden a lugares concretos de consumo y dispersión. En este sentido sobre salen los centros comerciales como puntos de emisión, Plaza Vallejo, Tecnoplaza, Centro de Coyoacán y la zona de eventos masivos del Foro Sol. Esto es indicador de que en estos espacios es importante aumentar la experiencia que se vive cotidianamente por medio de las herramientas digitales, como Twitter.

### **3.5 Consideraciones finales**

A lo largo de este capítulo se ha trabajado, implícitamente, con la idea de la gentrificación aumentada digitalmente en su dimensión del poder de la distribución. Para ello se ha propuesto una metodología teórico-empírica que sustenta la premisa del poder de la dimensión digital para generar espacios centrales, en una dicotomía centro/periferia y, de valorización de dichos espacios mediante Twitter. Así, la información de Twitter no se ha visto como una fuente adicional de información para estudiar el proceso urbano, sino como un elemento constitutivo de la generación de éste. En este sentido, se recuperan las ideas de Shelton (2017) y Crampton y cols. (2013) quienes argu-

mentan que la información georreferenciada de Twitter posee elementos sociales que trascienden la posición geográfica de la emisión.

El presente capítulo apunta a abonar en el campo de la discusión teórica sobre Twitter para estudios urbanos. En este sentido, la conceptualización de Twitter desde este ángulo permite discutir con los pocos desarrollos teóricos que se han propuesto desde las ciencias sociales sobre esta red. En particular, desde la postura de (Murthy, 2012) Twitter no solo se explica por medio del discurso emitido al interior de los mensajes, ni por el anclaje que los usuarios tienen en ésta. También se explica porque la emisión de Twitter presenta un anclaje geográfico, es decir, éstas no son aleatorias en el espacio. Al mismo tiempo, se refuerza la idea de que esta los tweets dan cuenta de las actividades cotidianas de los estratos altos de población, siguiendo el *ritmo* cotidiano de la ciudad. Esto se puede observar en el incremento de los tweets georreferenciados en fines de semana y en las diferentes concentraciones de emisión a lo largo del día, ya que, no se muestra la misma intensidad durante los días laborales que en las noches de fines de semana.

De igual manera, el estudio se vincula con los mencionados por (Goodchild, 2007; Elwood y cols., 2012) sobre las motivaciones de los individuos para la emisión georreferenciadas de sus actividades cotidianas. En este punto, se ha mostrado que las emisiones tienen un sesgo geográfico, ya que las emisiones no recuperan no se distribuyen homogéneamente en el espacio, por lo tanto se resaltan algunos fenómenos, mientras que otros siguen estando ocultos en la dimensión digital. De esta manera, la IGV puede ser una herramienta novedosa y que abra muchas posibilidades para el campo de los Sistemas de Información Geográfica, pero debe ser tomada con precaución.

En el apartado metodológico se han realizado dos aportes importantes en el campo de los estudios urbanos. El primero es dar cuenta de las vicisitudes del trabajo con los grandes datos y, particularmente con Twitter. Desde un punto de vista crítico de los grandes datos se ha tratado de recuperar todas las ventajas que este tipo de información representa. Ejemplos de esto ha sido su geolocalización de alta resolución, su alta granularidad temporal y la enorme cantidad de usuarios y publicaciones. Sin embargo, también se han considerado las limitaciones de esta información, tales como su carácter político que subyace a elementos como las cuentas automatizadas o la publicidad en Twitter, que a su vez provoca sesgo probabilístico. En este sentido, se puede mencionar que este no es una delimitación o uso definitorio de la información y es un campo naciente de investigación en los estudios urbanos.

El otro aporte empírico del capítulo ha sido una manera alternativa de identificación de colonias gentrificadas. En este punto se ha utilizado la unidad básica de análisis espacial para el estudio de la gentrificación, la colonia. De igual manera, se ha tomado distancia de las definiciones de cambio poblacional y se ha enfocado a la delimitación por medio de la centralidad de comercios vinculados con gentrificación y su coexistencia. Asimismo, se ha incorporado una variable aproximada para la dimensión de la gentrificación guiada por el Estado, asumiendo que la localización de las estaciones de bicicletas públicas responden más a una mentalidad de valorización del espacio urbano que a una solución de movilidad cotidiana. Con todo esto, se han confrontado los resultados obtenidos en esta identificación con estudios previos. La identificación presentada es consistente con dichos estudios y aporta una identificación cuantitativa a nivel de colonia. Sin embargo, esta identificación tiene una limitación importante, ya que la información recabada ha sido de carácter estático, lo cual no recupera el carácter de cambio constante que se vincula con la gentrificación, aunque esta limitación podría subsanarse con posteriores análisis.

En términos de la gentrificación, la emisión de tweets se asocia espacialmente con ésta. Como se ha mostrado en todos los mapas anteriores, independientemente de la hora o del día la principal concentración de emisiones se presenta en las colonias identificadas como gentrificadas. Sin embargo, al realizar el análisis temporal de la concentración de emisión se ha mostrado que ésta aumenta en días y horas ligados a actividades de esparcimiento, tales como consumo en restaurantes, bares.

Otro de los puntos que refuerzan que la georreferencia de un tweet se explica por los fenómenos urbanos es la concentración de éstos fuera de las colonias gentrificadas. En este punto, sobresalen los espacios vinculados al consumo en espacios públicos de índole privado, tales como los centros comerciales. Ejemplos de lo anterior son Industrial Vallejo, al norte de la ciudad, el Centro Comercial Santa Fe o Plaza Oriente. En este mismo sentido, se observan altas concentraciones en puntos importantes de la ciudad como el Aeropuerto Internacional o lugares de espectáculos masivos, el Foro Sol y el Palacio de los Deportes son muestra de esto. Estos hallazgos dialogan con las posturas sobre la información de Twitter como Información Geográfica Voluntaria, y la motivación de las personas para emitir. Para el caso de la Ciudad de México se puede inferir que la motivación de la georreferenciación de tweets se asocia a prácticas de consumo y entretenimiento, más que una construcción de ciudadanía.

La asociación de la dinámica urbana con la emisión de Twitter abona a la explicación de los

motivos para geoetiquetar publicaciones, la cual sustenta la postura de la Información Geográfica Voluntaria. Es decir, para el caso de Twitter, la información recopilada de esta red se explica no por elementos individuales, sino por las dinámicas estructurales cotidianas de las personas. En este sentido, la propuesta de las personas como prosumidores en la dimensión digital, permite extender el argumento a la producción y consumo del espacio al mismo tiempo mediante la dimensión digital. Aunque, este prosumo se manifiesta, únicamente, en espacios centrales con dinámicas de valorización económica y simbólica.

Finalmente, el poder de la distribución en el acceso digital no es la única dimensión para hablar de la realidad geográfica aumentada digitalmente. Para constituir la de manera completa, se tiene que analizar el poder de la comunicación. En palabras de Shelton (2017) y como se ha demostrado en el capítulo anterior, dado el sesgo que presentan los usuarios de Twitter, éstos tienen el poder de imponer determinados temas en la esfera pública digital, en detrimento de otros que atañen a las personas y espacios que no tienen el mismo poder de acceso y comunicación. De esta manera, en el siguiente capítulo se desarrolla el poder de la comunicación en el fenómeno analizado.



## Capítulo 4

# El poder de la comunicación en la gentrificación aumentada digitalmente en la Ciudad de México

### 4.1 Introducción

La manera de hablar y de comunicarnos nos distingue. Las palabras que usamos en determinados contextos pueden adquirir diferentes significados e interpretaciones. En la era de lo digital, la comunicación es continua e inmediata. El objetivo general de este trabajo de investigación es explorar cómo el proceso de gentrificación en la Ciudad de México se ve aumentado digitalmente mediante Twitter. Para ello, en el desarrollo de la investigación se ha propuesto la teoría de la Realidad Geográfica Aumentada Digitalmente (RGAD) como propuesta interpretativa para entender los mecanismos de la dimensión digital que ayudan la construcción de una espacialidad digital vinculada con la gentrificación. En este sentido, se recuperan las ideas básicas de Smith sobre la gentrificación como proceso del desarrollo urbano desigual en conjunción con la propuesta de la RGAD. Así, a lo largo de los capítulos de esta tesis se ha apuntado a los mecanismos de valorización y centralización económica que impulsan colonias gentrificadas en la Ciudad de México y cómo éstos se representan por medio de Twitter. En consecuencia, en el capítulo 2 se desarrolló la idea de que el uso de la red sociodigital de interés es utilizada primordialmente por un perfil de persona de estratos medios y altos, con grados de escolaridad elevados y primordialmente jóvenes. Es decir, Twitter no es utilizado de manera universal. En este orden de ideas, en el capítulo 3 se ha apuntado a que la distribución espacial de las emisiones de Twitter se concentra en espacios relacionados con el consumo, esparcimiento y actividades empresariales y de servicios, principalmente en colonias identificadas como gentrificadas. Por ende, la distribución digital de Twitter apunta a la centralización de actividad digital en espacios gentificados.

Con esto en mente el presente capítulo se enfoca a analizar el poder de la comunicación de la RGAD y su relación con el proceso de gentrificación en la Ciudad de México. Para ello se utiliza

el principal medio de comunicación de Twitter, el mensaje de texto de emitido por sus usuarios. Para ello, el presente acápite se divide en 6 apartados. El primero de ellos contextualiza el estado del arte de trabajos que se han preguntado sobre la relación entre el proceso de gentrificación y plataformas digitales, tales como Yelp, Instagram o Twitter. El segundo apartado expone algunos elementos del modelo lingüístico de Karl Bühler, el cual se propone como un acercamiento teórico para interpretar los resultados del poder de la comunicación de Twitter para construir una gentrificación digitalmente aumentada. Esta propuesta interpretativa es uno de los principales aportes de esta investigación ya que hasta el momento la mayoría de trabajos que han analizado el texto de Twitter para el caso de estudios urbanos lo han hecho desde la técnica de análisis de sentimientos y con ello no le han otorgado a las palabras un papel sociolingüístico en la conformación de espacialidades lingüísticas.

En el tercer apartado se exponen con detalle los pasos en el pre-procesamiento del *corpus* de Twitter. En esta sección se exponen tanto los pasos a seguir como las limitaciones técnicas y cognitivas que se enfrentaron para preparar los datos de tal manera de otorgar al análisis de este trabajo una visión territorial del lenguaje de Twitter. De manera específica se subraya la identificación de entidades nombradas de carácter espacial; aspecto que no ha sido reportado hasta el momento en la literatura internacional especializada. En el mismo apartado se expone el funcionamiento del algoritmo de aprendizaje de máquina no supervisado para la modelación de texto, denominado Asignación Latente de Dirichlet. En este sentido, cabe mencionar que se ha realizado un giro metodológico en comparación con los capítulos empíricos anteriores. Ya que, mientras la evaluación empírica había estado guiada por categorías teóricas que ayudaban a construir indicadores enfocados en explorar las dimensiones de la gentrificación digital, en el presente capítulo se ha volcado a un método de carácter deductivo. Esto debido a que la naturaleza de la técnica permite identificar temas latentes que no necesariamente se vinculen con el proceso de gentrificación.

El cuarto y quinto apartado se enfocan en la presentación de resultados, en ellos se pone atención a la variación de principales palabras en función de la temporalidad y de la colonia de emisión, ya sea gentrificada o no. De manera particular se presentan los resultados del modelo de Asignación Latente de Dirichlet, la cual arroja la identificación de 20 tópicos semánticos latentes en los tweets de la ciudad. En este sentido se identifican dos grandes grupos de temas. El primero de ellos representan los temas asociados al procesos de gentrificación y por el otro lado temas que no

se vinculan directamente con el proceso. En el último apartado se exponen algunas consideraciones principales de los hallazgos.

## **4.2 Gentrificación aumentada digitalmente por el poder de la comunicación**

Como se ha mencionado en el capítulo 1, el proceso de gentrificación no solo se puede entender por medio del desplazamiento de población o el incremento del precio de la renta del suelo. Este proceso implica dinámicas estructurales del sistema económico actual, tales como la renovación y revalorización de espacios centrales de la ciudad, en detrimento de las zonas periféricas. Al mismo tiempo, el fenómeno estructural de la gentrificación se aumenta por la presencia de la dimensión digital. Dicho aumento se basa en tres tipos de poderes, el poder de la distribución, el cual se basa en la disparidad para acceder a Internet y a los elementos de éste; el poder de la comunicación, el cual se sustenta en el diferencial del poder que las personas tienen de representar intereses en la dimensión digital por medio de discursos o símbolos; y finalmente, por el poder del código el cual establece que el código que sustenta a la dimensión digital no es neutro. De manera general, todas las aristas de la dimensión digital interactúan para reforzar los elementos estructurales de la gentrificación. Hasta el momento se ha desarrollado cómo el poder de la distribución de la realidad aumentada digitalmente interviene en el proceso de centralización y valorización del espacio urbano. En consecuencia, en este capítulo se ofrece el análisis del poder de la comunicación de la realidad aumentada digitalmente y cómo se entrelaza con los procesos de centralización y revalorización del espacio gentrificado.

En primera instancia, la gentrificación ha sido vista desde su dimensión comercial, simbólica y cultural. Los trabajos de Sharon Zukin han sido piedra angular de la construcción teórica y metodológica para entender la gentrificación desde esta perspectiva. Dicho enfoque discute que no solo el precio de la vivienda es el único determinante para el desplazamiento de personas de bajos recursos, sino que los elementos culturales y creación de identidades funcionan como mecanismos de diferenciación social que impiden el acceso a ciertas personas a espacios gentrificados. Así, la autora identifica que el consumo vinculado con cafés, galerías de arte, cines, teatros, restaurantes *gourmet* son factores que influyen en el proceso de fragmentación del espacio urbano por medio de la gentrificación (Zukin, 1987, 1998; Zukin y cols., 2009)

En esta corriente se pueden identificar los trabajos relacionados con la gentrificación económi-

ca y comercial. Al igual que la formación de identidades de la gentrificación fragmenta los espacios renovados, también lo hace la localización de servicios y comercio que atienden a una demanda elitizada. En este sentido, se ha mostrado cómo los espacios céntricos de algunas ciudades han transformado su vocación económica de locales de consumo inmediato y poco especializado hacia establecimientos de consumo dirigidos a las élites que ocupan y usan los espacios (Batres-Guadarrama, 2016; Díaz-Parra y Salinas, 2016; Duhau y Giglia, 2016). De esta manera, la perspectiva culturalista y comercial de la gentrificación establece que el espacio gentrificado se produce por todas las personas que comparten ciertos símbolos y prácticas, independientemente si residen o no en dichos espacios.

Esta formación de símbolos se ha trastocado con el auge de las redes sociodigitales e Internet. Los primeros indicios de una relación entre la dimensión digital, particularmente con la Web 2.0, y el proceso de gentrificación se registraron en la ciudad de Nueva York. En su estudio prolegómeno, Belkind (2008) analiza el vínculo entre el blog *curbed.com* y el cambio poblacional y en el espacio construido del Este de Nueva York. La autora apunta al papel del blog escrito en Internet como un medio para retroalimentar ideas culturales sobre el espacio gentrificado. En este sentido, el trabajo propone interpretar el papel de *curbed* como un mecanismo del sistema económico para valorizar el espacio. Aunque, el trabajo proporciona algunos elementos para entender el papel del poder de la comunicación, no desarrolla una idea de co-producción entre la dimensión digital y el fenómeno de gentrificación.

En este campo, algunos estudios han comenzado a explorar los mecanismos de gentrificación de consumo y simbólica por medio de Internet. Al respecto, Zukin y cols. (2017) analizan procesos de gentrificación en Manhattan que se ven reflejados por medio de las valoraciones de restaurantes internacionales en dos barrios de la isla. Las autoras encuentran que las reseñas escritas por comensales emitidas en el sitio Yelp reflejan los distintos niveles de gentrificación que dos barrios. Por un lado, describen como las reseñas que hacen referencia a los restaurantes ubicados en una colonia gentrificada consolidada y de cocina polaca refieren el entorno urbano como seguro y con aires europeos. Por otro lado, las reseñas que hablan sobre un barrio afro-caribeño que está en transición en el proceso de gentrificación hacen referencia al entorno como inseguro y poco amable. En este trabajo, las autoras enfatizan el papel de consumidores y productores del espacio urbano gentrificado por parte de los usuarios del sitio Yelp.

Por su parte, Gibbons y cols. (2018) estudian la manera en que el proceso de gentrificación en Washington refleja la transformación de identidades en los barrios afectados por medio de discurso y lugares nombrados en Twitter. Los autores cuestionan la manera en que la gentrificación rompe con la identidad original de los habitantes de los espacios renovados. Sin embargo, encuentran que por medio de la red sociodigital se construyen nuevas identidades. Esto lo hacen por medio del análisis léxico de los mensajes de tweets publicados en zonas que identifican como gentrificadas.

Por último, en México se tiene un primer esfuerzo de entender cómo la dimensión digital actúa en este fenómeno. Díaz (2015), desde un enfoque de etnografía digital, estudia cómo vecinos de la colonia Condesa, en la Ciudad de México, se organizan por medio de un grupo de Facebook para expresar sus inconformidades por la conversión que ha sufrido su colonia. El autor concluye que el proceso de gentrificación en la Ciudad de México, que ha sido emprendido por el gobierno y por corporativos privados ha encontrado en el crecimiento de las redes sociodigitales otro actor a considerar. El centro de la ciudad se ha poblado de personas jóvenes que tienen más facilidad para el uso de los instrumentos digitales y, al mismo tiempo, por este medio se ha comenzado a construir un espacio donde se vierten ideas sobre un espacio idílico sobre una de las colonias ícono de la gentrificación en la ciudad (Díaz, 2015).

En términos metodológicos, los trabajos de Zukin y cols. (2017) y Díaz (2015) analizan de manera cualitativa las conductas y los discursos sobre la gentrificación en diferentes plataformas digitales. En el caso de Gibbons y cols. (2018) utiliza la fortaleza de los grandes datos para realizar un análisis de texto computarizado. Todos estos trabajos utilizan la dimensión lingüística de las plataformas digitales para analizar el proceso gentrificador. Es por ello que, para realizar estudios sobre el poder de la comunicación de la gentrificación aumentada digitalmente se propone entenderla desde un enfoque lingüístico que permita entender cómo estas dos dimensiones se entrelazan. Con esto en mente, a continuación se expone, brevemente, algunos conceptos de teoría del lenguaje, con el objetivo de apuntalar el análisis del poder de la comunicación.

### **4.3 Elementos para y entorno al análisis de texto de Twitter**

Como se expuso en el apartado 3.2.1 un primer acercamiento social a Twitter es por medio de la teoría del discurso de Goffman. Esta aproximación teórica de Twitter enfatiza la construcción de la red sociodigital desde los diferentes niveles de interacciones de los usuarios, el uso de determinados

signos y símbolos. Dicha interacción se sustenta en un anclaje social que sustenta la interacción de los usuarios. Es decir, un mensaje en Twitter es emitido en un proceso bidireccional entre el emisor y sus seguidores. Sin embargo, a este enfoque teórico propuesto por (Murthy, 2012) se le pueden hacer dos críticas. La primera crítica apunta a un aspecto empírico, ya que para estudiar los símbolos, signos y anclaje social en Twitter son necesarias herramientas computacionales provenientes de los grandes datos, y no tanto una aproximación metodológica basada en análisis críticos del discurso (Locke, 2004). La segunda se sustenta en el ámbito teórico, ya que propone identificar un anclaje social que sustenta la interacción, pero nunca deja de manifiesto que dicho anclaje no solo es la interacción entre personas, sino la interacción entre ellas basada en el uso de espacios determinados.

En respuesta a la primera crítica, en este capítulo se propone estudiar el texto de Twitter mediante herramientas de la ciencia de datos, tales como el Aprendizaje de Máquina. En este punto se debaten algunas posturas y aproximaciones dentro del análisis léxico de Twitter. En respuesta a la segunda crítica, este trabajo también propone dar un sustento teórico desde el campo de la lingüística para entender la formación de signos y síntomas lingüísticos en relación al proceso urbano de gentrificación.

De manera generalizada los trabajos que han abordado empíricamente el mensaje textual de Twitter no han reflexionado sobre el carácter lingüístico de las palabras como signos. Por un lado el análisis de grandes volúmenes de texto en Twitter se han enfocado en la identificación de opiniones de cada uno de los tweets, por lo general, de carácter positivo, negativo o neutro. El objetivo central de esta aproximación es catalogar a cada tweet dentro de los sentimientos predefinidos, en este sentido se le han denominado estrategias de análisis de texto corto (Thelwall, Buckley, Paltoglou, Cai, y Kappas, 2010). La base epistemológica de este tipo de aproximaciones es la psicología de las emociones. Es por ello que el mejoramiento de las técnicas han sido destinadas a ampliar el abanico de sentimientos identificados por el algoritmo (Thelwall y cols., 2010; Thelwall, Buckley, y Paltoglou, 2011). En este sentido, el análisis sobre el texto parte de una postura de identificar *a priori*, y no desde una aproximación sociolingüística del texto. En consecuencia, el presente trabajo reflexiona desde una postura sociolingüística que conlleva a cambiar el punto de partida para el análisis léxico. En este sentido, primero se sustentan las bases desde una teoría del lenguaje para después dar paso al desarrollo técnico.

En este trabajo se propone una aproximación interpretativa basada en la teoría lingüística de las funciones del lenguaje (Bühler, 1961). Este enfoque permite interpretar las palabras que se encuentran en Twitter en consideración a su función lingüística. Para este motivo, primero hay que establecer que existe una diferencia entre el discurso y la palabra. En este sentido, el discurso da cuenta de algo o alguien, mientras que la segunda solamente nombra, es decir, atribuye nombre a algún objeto. Dentro de esta perspectiva, la palabra tiene tres funciones dentro de la teoría del lenguaje, la función referencial, sintomática y apelativa (Lara, 2015). A continuación se expone cada una de ellas.

La función referencial de la palabra se centra en el papel de las palabras para referir a un objeto o situación de la realidad, o bien nombra. Esta función se encuentra en el proceso de comunicación y se puede verificar el contenido del mensaje. En este nivel Lara (2015) ejemplifica, que la palabra *indio* hace referencia a una persona proveniente de la India o bien a una persona nativa del continente americano antes de la llegada de los españoles a éste. En este nivel, la palabra no contiene alguna carga política, positiva o negativa, es solo la manera de referir, por medio del habla o escritura ha un objeto. La siguiente función es la sintomática. Esta función hace énfasis el origen de quién emite la palabra. De nueva cuenta, Lara (2015), propone entender esta función de la palabra por medio del signo *banqueta*, ya que al decir o escribir dicha palabra cualquier hablante de lengua española sabrá que el emisor proviene de México, ya que dicho país es el único lugar donde se refiere a la acera de esta manera. La función sintomática de la palabra permite, entonces, contextualizar la palabra en sí misma por medio de su origen. La tercera función es la apelativa. Dicha función corresponde a la interpretación del receptor sobre la palabra, en esta función es cuando se presentan las diferentes interpretaciones que una palabra pueda tener, ya sea negativa o positiva.

En este trabajo se propone utilizar la función de las palabras encontradas en los tweets de colonias gentrificadas como contextualizarlas y asume una perspectiva de que las palabras son signos lingüísticos del síntoma de la gentrificación en la ciudad. En este punto se puede decir que la intersección entre el campo de la lingüística y la gentrificación es relativamente nueva. Sin embargo, se tienen algunos esfuerzos en ciudades del norte global que se han enfocado en identificar cómo el lenguaje es reflejo y motor de gentrificación.

Desde el campo de la lingüística se han tienen algunos enfoques que se han mirado al proceso de la gentrificación. Entre los enfoques que relacionan el fenómeno urbano con la manera

de comunicarse de manera oral y textual se pueden esquematizar tres aproximaciones. La primera de ellas ha indagado la relación entre la configuración de estrato o clase económica con determinado uso de lenguaje, de manera particular, con el uso de palabras en inglés o idiomas europeos (Vandenbroucke, 2016). Por ejemplo, se ha identificado que zonas gentrificadas de mayor ingreso se vinculan con educación bilingüe, en Estados Unidos o Europa. En este sentido, la segunda aproximación lingüística ha sido por medio del uso y diversidad de idiomas en determinados espacios. Dicha diversidad lingüística ha sido foco de atención debido a que se relaciona con el proceso de globalización y el desarrollo económico mundial. Este hecho ha llevado a plantear si la coexistencia y uso de varios idiomas en un mismo espacio como indicador de procesos sociales, económicos y culturales de escala mundial (Flores y Lewis, 2016; Lyons y Rodríguez-Ordóñez, 2017). Una tercera aproximación ha sido por medio de estudio en los cambios de entonación en la manera de hablar en zonas gentrificadas. En este caso, Médéric Gasquet-Cyrus (2015) vincula la reestructuración poblacional del proceso de gentrificación en Marsella, Francia con el cambio de entonación de los habitantes de dicha ciudad. Esto debido a que dicho cambio atrajo a personas jóvenes vinculadas a actividades artísticas de varias partes del país y Europa.

De manera particular la lingüística ha estudiado los mecanismos lingüísticos que permiten identificar espacios en transición o consolidados como gentrificados. En este rubro se identifica dos enfoques, el marco contextual y el paisaje lingüístico. El primero se ha concentrado en la premisa de que en espacios gentrificados coexisten signos que son reflejo de diferentes marcos culturales y sociales. Por ejemplo, Lyons y Rodríguez-Ordóñez (2017) identifican que un barrio en Nueva York en proceso de identificación coexisten diferentes marcos que se relacionan la multiplicidad de signos que se encuentran en sus calles. En este sentido, analizan la presencia de personas migrantes de reciente llegada al barrio por medio de anuncios en español y vinculados con trámites migratorios. Mientras que identifican a personas de reciente llegada al barrio jóvenes, generalmente blancos. Los autores identifican que los signos que se encuentran en las calles dirigidos a estas personas tienen que ver con signos vinculados a 'lo local', orgánico, cultural, además, se recuperan muchos signos relacionados con lo tradicional. El trabajo enfatiza que este tipo de personas son identificadas como '*hispters*' por las personas de más antigüedad en la zona.

El segundo enfoque es el denominado paisajes lingüísticos. Esta aproximación se enfoca en estudiar la conformación de paisajes un espacios determinados por medio de símbolos y signos. De

manera general, los paisajes lingüísticos se asocian al conjunto de palabras y mensajes semióticos utilizados por el Estado para construir opiniones y conductas en el espacio público. Para el caso del proceso de gentrificación se ha mirado cómo el paisaje lingüístico es guiado por estrategias comerciales. Así, estudios como los de Lyons y Rodríguez-Ordóñez (2017); Papen (2012) muestran que en colonias gentrificadas los establecimientos utilizan mensajes en inglés o idiomas europeos para dar promoción a sus tiendas. En este sentido, Papen (2012) menciona que los negocios utilizan nombres y anuncios ‘elegantes’. Los autores identifican que esto se puede deber a dos grandes tendencias. La primera de ellas la presencia de marcas internacionales vinculadas al consumo de élite, tales como Pull&Bear, The Body Shop. La segunda causa se puede deber a la disponibilidad de cadenas de comida rápida como McDonald’s. Para los autores esta estrategia se vincula con el proceso de gentrificación comercial.

Como se ha mostrado, las aproximaciones desde la sociolingüística apuntan a que los signos encontrados en el paisaje urbano se relacionan, refuerzan y co-producen el fenómeno de gentrificación. En este sentido, valdría la pena pensar que el espacio digital es una nueva espacialidad de conformación paisajes lingüísticos vinculados con la gentrificación. Así, tomando de nueva cuenta la propuesta teórica de Bühler, los signos o palabras encontradas en los tweets son reflejo de un síntoma socioterritorial, en este caso la gentrificación. Y, en última instancia, esta construcción lógica podría ayudar a entender la manera en la que se constituye un poder diferenciado en la comunicación para constituir una gentrificación digitalmente aumentada en la Ciudad de México.

#### *4.3.1 Análisis de texto en Twitter en los Estudios Urbanos*

Una de las principales fortalezas que ofrece Twitter como fuente de información es la posibilidad de obtener las opiniones de sus usuarios por medio de texto, el cual está restringido a 240 caracteres. Esta cualidad, aunada a la generación en tiempo real de los datos, ha convertido a Twitter en una fuente de información para áreas como la ciencia política, comunicación o periodismo. Las oportunidades de aplicación de Twitter se enfocan, principalmente en el análisis de tendencias políticas, preferencias en el voto, temas de agenda pública o quejas ciudadanas. Al mismo tiempo, ciencias de la computación y científicos de datos han explorado la posibilidad de utilizar estos datos para realizar análisis de minería de texto con el fin de encontrar patrones de comportamiento en los usuarios de la plataforma. En este contexto, herramientas como el *aprendizaje de máquina* (Machine

Learning) o el análisis de sentimientos han mostrado su capacidad y limitaciones para el estudio de las palabras recolectadas en Twitter. Para el caso del análisis de texto de Twitter y los estudios urbanos, los intentos por utilizar estas herramientas se han enfocado en tres grandes vertientes, la primera de ellas, el análisis del texto por medio de diccionarios geográficos (*gazetteers*), la segunda por medio del análisis de sentimientos basados en técnicas de aprendizaje de máquina y el estudio de prácticas sociales en el espacio.

Los diccionarios geográficos se definen como índices que conjuntan características y estadísticas que pueden ser vinculados a espacios particulares. En la antigüedad se utilizan para mantener registro del nombre de las capitales, lugares sagrados, dioses locales y otros datos sobre lugares (M. Graham y Sabbata, 2015). Estos diccionarios son utilizados en la actualidad por lingüistas e historiadores para el estudio y reconstrucción de lugares históricos y aspectos geográficos por medio del análisis del lenguaje natural. Recientemente los diccionarios geográficos han adquirido importancia debido a que se han convertido en una herramienta fundamental para motores de búsqueda de lugares en Internet (M. Graham y Sabbata, 2015). En específico, en el estudio de ciudades, los *gazetteers* en la era de los grandes datos se utilizan para diversos propósitos. A continuación se mencionan algunos de éstos.

Aunque el uso de diccionarios geográficos es común en estudios históricos, el empleo de éstos en los estudios urbanos es reciente y permite reflexionar sobre varias problemáticas urbanas contemporáneas por medio del lenguaje natural. Un ejemplo de esto es la creación del nombre de ciudad como marca bajo el contexto del emprendedurismo urbano. En este punto, Sevin (2014) interpreta la aparición del nombre de algunas ciudades estadounidenses como referencia comercial para determinados eventos deportivos o artistas. De esta manera, el autor señala la creación de la ciudad como una marca comercial y que ésta sea reconocida abiertamente en Twitter permite potenciar las actividades comerciales tales como el turismo, exportaciones e inversión extranjera. Desde una perspectiva crítica M. Graham y Sabbata (2015) exponen el papel de los *gazetteers* en las búsquedas de Internet como elementos de creación de desigualdad geográfica. Al respecto, los autores puntualizan que los diccionarios geográficos de las grandes ciudades y lugares reconocidos a nivel mundial se encuentran sobrerrepresentados. A decir de los autores, la disparidad en los lugares nombrados es un elemento para considerar que Internet y las búsquedas que en éste se realizan no son neutras, ya que se basan en la desigualdad geográfica existente.

De manera muy cercana al enfoque de los diccionarios geográficos para el estudio del texto en Redes Sociodigitales, se encuentra la construcción de regiones y territorialidades a través del análisis léxico. En estos estudios, se parte de la premisa de la manifestación del espacio heterogéneo que se vive cotidianamente por las personas mediante el texto que éstas emiten en los sitios de redes sociales en Internet (Kelley, 2014). Ejemplos de esto son los esfuerzos por regionalizar Estados Unidos por medio de las diferencias lingüísticas que los usuarios de Twitter emitieron durante un año (Y. Huang, Guo, Kasakoff, y Grieve, 2016). Otro esfuerzo lo aportan Zhou, Xu, y Kimmons (2015), quienes mediante el análisis de palabras identifican lugares turísticos en algunas ciudades y la ocurrencia de eventos importantes mediante Twitter. En contraparte, algunos trabajos argumentan que la creación y reproducción del espacio vivido en el entorno urbano no se duplica automáticamente en el entorno digital. Al respecto, Croitoru y cols. (2015) exploran la manera en que eventos inesperados o manifestaciones no se reproducen automáticamente en Twitter. Los autores estudian el movimiento de *Occupy Wall Street* en 2011 y de los bombardeos en el maratón de Boston en 2013 para identificar que los comentarios surgidos de manera inmediata en ambos eventos se fueron transformado durante horas próximas hasta formar entidades discursivas que hicieron alusión a los hechos ocurridos horas previas. En el mismo hilo conductor, Jiao, Holmes, y Griffin (2018) analizan el texto de tweets durante el gran tazón de 2012 y cómo se representaba la imagen urbana a través de dichas publicaciones. Los autores encuentran que los usuarios preponderaron el evento ante el lugar. Las referencias a la percepción de la ciudad se dieron días antes, lo que implicaría la presencia de turistas. Ambos trabajos, puntualizan la formación temporal de la imagen urbana en Twitter y su carácter asincrónico al tiempo real.

Otra de las perspectivas que se han estudiado para interpretar las dinámicas urbanas por medio del texto en Redes Sociodigitales por medio de identificar las actividades económicas y prácticas que se realizan en lugares determinados. En esta vertiente se encuentran los trabajos realizados por Lansley y Longley (2016), quienes analizan las menciones y patrones de publicaciones que refieren a compras, ropa, visita de museos y galerías, vida nocturna y entretenimiento en la ciudad de Londres. De la misma manera, se ha intentado determinar de mejor manera la dinámica de residencia y trabajo por medio de los tweets en la misma ciudad (Steiger, Westerholt, Resch, y Zipf, 2015). Por su parte, Gibbons y cols. (2018) se enfocan en identificar los discursos en Twitter que reflejan procesos de gentrificación en la ciudad de Washington DC. En conjunto, se ha explorado,

no solo la frecuencia de los lugares nombrados mediante el texto en Redes Sociodigitales, sino las prácticas que se asocian a ciertos lugares determinados por dinámicas urbanas.

Finalmente, una tercer vertiente de estudios se enfoca en la utilización de la herramienta del análisis de sentimiento <sup>1</sup> con el propósito de identificar las percepciones de los habitantes acerca de los lugares urbanos. Este tipo de análisis es empleado por Hollander y Renski (2015) para estudiar el ánimo y percepciones de algunas ciudades de Estados Unidos. Emociones como preocupación o molestia son recuperados por medio de dicho análisis de texto. Otro estudio es el realizado por Roberts, Sadler, y Chapman (2018) quienes se enfocan en estudiar los sentimientos de aprobación o desaprobación de personas que frecuentan áreas verdes y parques en la ciudad de Birmingham en Reino Unido. Dichos estudios permiten conocer en el potencial del análisis de texto para identificar dimensiones subjetivas de los habitantes de la ciudad.

En el caso de México, el INEGI ha desarrollado el análisis de sentimientos con base en el texto de los usuarios de Twitter en el trabajo “Estado de ánimo de los tuiteros en México” (INEGI, s.f.). El estudio realiza la recolección, depuración, geolocalización y clasificación de la emisión de tweets diarios en el territorio mexicano. Con base en esto, el INEGI recupera las publicaciones por medio de la API gratuita de Twitter. Dicha publicaciones, son georreferenciadas y clasificadas en función de su ubicación estatal y temporal. Por último, todas la publicaciones son etiquetadas por medio de un procedimiento de *aprendizaje de máquina* en dos categorías, tweets positivos o negativos. La herramienta de INEGI ha sido empleada como una aproximación del estado de ánimo de los mexicanos después de los sismos del 19 de septiembre de 2017 (CNN, s.f.; Salgado-Andrade, 2018). A pesar de que INEGI realiza la distinción espaciotemporal de las publicaciones de Twitter en el país, el ejercicio carece de un análisis más profundo sobre las causas geográficas y sociales de sus resultados.

Aunque el análisis de lenguaje natural de tweets aplicado a los fenómenos urbanos es relativamente reciente, existen algunas aproximaciones que pueden ayudar a indagar cómo el poder

---

<sup>1</sup>“El análisis de sentimiento es también llamado *minería de opinión* es el campo de estudio que analiza las opiniones, sentimientos, valoraciones, actitudes y emociones de las personas hacia entidades y sus atributos expresados en un texto. Las entidades pueden ser productos, servicios, organizaciones, individuos, eventos, problemas, o temas. El campo implica una amplia variedad de casos.” (Liu, 2015, p.1).

“Aunque el análisis de sentimientos se originó en las ciencias de la computación, en años recientes, ha sido ampliado a ciencias de la administración y ciencias sociales debido a su importancia para los negocios y la sociedad como un todo. Así, la investigación del análisis de sentimientos no solo abarca el procesamiento de lenguaje natural, sino también aquellos campos referentes a las opiniones públicas y de consumidores”(Liu, 2015, p.4)

diferenciado de la comunicación en Twitter puede aumentar en el terreno digital el proceso de gentrificación. Para profundizar esta premisa, se retoman algunos elementos expuestos anteriormente para estudiar este fenómeno. En primer lugar, se retoma la propuesta de la teoría de la función de lenguaje de Bühler. Esto se hace con el propósito de interpretar los resultados del procesamiento de lenguaje natural desde una perspectiva sociolingüística. Dado que las herramientas computacionales de grandes datos no permiten realizar un Análisis Crítico de Discurso se propone utilizar las categorías de símbolo y síntoma de las palabras para indagar sobre el proceso de gentrificación en la ZMCM mediante el texto de Twitter. En este sentido, se recupera la propuesta de las entidades nombradas para explorar si existen lugares explícitamente nombrados identificados como lugares referentes al proceso de gentrificación. Así, las entidades nombradas serán entendidas en su función de referencia y síntoma de la gentrificación. Esto se llevará a cabo por medio de la comparación de entidades nombradas en tweets emitidos en colonias identificadas como gentrificadas (ver capítulo 3) en comparación con las entidades nombradas en colonias que no lo son. Al respecto, se espera encontrar el texto de Twitter refleje los espacios reconocidos como íconos del proceso de gentrificación en la ciudad en colonias gentrificadas y que en colonias que no lo son se nombren lugares que, sintomáticamente, representen algunos procesos relacionados como gentrificados. Por ejemplo, se espera encontrar que algunas de las palabras en colonias gentrificadas se relacionen con cierto tipo de consumo de élite, ya sea de ropa, alimentos o esparcimiento. Por su parte, en colonias no gentrificadas se pudiesen llegar a encontrar palabras relacionadas con dicho proceso, pero con menor importancia relativa y en espacios específicos, como centros comerciales. Finalmente, se busca encontrar diferencias léxicas en las palabras entre los dos tipos de espacios. Es decir, que en las colonias gentrificadas se construye un determinado uso del lenguaje que responde a la generación de distinción y elitización de parte de los usuarios de Twitter que emiten desde esos espacios.

Para analizar cada uno de estos aspectos, a continuación se exponen los aspectos técnicos y metodológicos. Primero, para realizar el análisis de texto por medios computacionales se debe tener en cuenta que éste debe ser depurado y preparado para tal propósito. En este sentido, se presentan algunos elementos que se han realizado para poder estudiar el texto de Twitter, entre ellos la identificación de entidades nombradas. Segundo, con el objetivo de explorar la importancia que tienen las palabras en cada espacio de la ciudad se realiza un análisis de frecuencia. Es decir, se analiza la cantidad proporcional de repetición de las principales palabras con el objetivo de

encontrar significados y elementos sintomáticos de gentrificación en dichas palabras. En este punto, también se distingue entre las palabras más frecuentes en tweets emitidos en colonias gentrificadas y no. Finalmente, con base en uno de los algoritmos de *aprendizaje de máquina* se construyen temas semánticos latentes emanados de los tweets en colonias gentrificadas, de esta manera se puede comenzar a vislumbrar los diferentes tópicos que emergen en el espacio gentrificado.

#### **4.4 Metodología y modelización de temas**

En esta sección se expone la estrategia metodológica para el análisis de los principales temas en Twitter en la ZMCM. Además, se presenta el ajuste de ‘aprendizaje de máquina’ de Asignación Latente de Dirichlet (ALD) (LDA por sus siglas en inglés, *Latent Dirichlet Allocation*).

##### *4.4.1 Depuración de texto en los tweets de la ZMCM*

El análisis de texto de tweet implica una diversidad de retos. Ya sea retos lingüísticos, computacionales, o interpretativos. Al igual que los datos de tipo numérico, los datos provenientes de texto, y analizados bajo el paradigma del Procesamiento de Lenguaje Natural, requieren de una preparación previa para depurar la información bruta y realizar análisis confiables. En este sentido, “el pre-procesamiento de texto es una parte esencial dentro de cualquier sistema de Procesamiento de Lenguaje Natural, desde caracteres, palabras y oraciones identificadas en esta etapa son unidades fundamentales en las etapas subsecuentes del procesamiento” (Palmer, 2010, p. 9). El pre-procesamiento de texto implica convertir una secuencia de texto crudo en una cadena de caracteres bien definida lingüísticamente significativa. Así, los pasos de pre-procesamiento dependen del tipo de texto, objetivo de estudio y modelado posterior. Ejemplos de esto pueden ser la conversión del texto en minúsculas, eliminar signos de puntuación, la vectorización de texto o identificar conjuntos de palabras con sentido unitario, remover nombres propios o títulos (Bird, Klein, y Loper, 2009; Kwartler, 2017).

Con esto en mente, la depuración de texto de Twitter en la ZMCM ha tenido el objetivo de buscar patrones léxicos de emisión diferenciados, emanados de tweets en colonias gentrificadas en comparación con las no gentrificadas. De igual manera, se ha pre-procesado el texto para modelarlo por medio del algoritmo de Asignación Latente de Dirichlet. En consecuencia, la depuración de

tweets para análisis de texto tiene las siguientes características. Como se ha mencionado en la subsección 3.3.1, el número total de tweets recolectados durante el periodo de estudio han sido 3.16 millones. En el capítulo anterior, también se hizo notar la presencia de *bots* en la emisión metropolitana de tweets. Esto ha llevado a reducir la cantidad de tweets a un número alrededor de 1.5 millones de emisiones. Aunada a la depuración en función de usuarios hecha en la subsección ya citada, para el caso de obtener la mejor calidad de texto se ha realizado una segunda depuración. De manera puntual, la depuración final de tweets, se esquematiza en los siguiente puntos:

- Eliminación de *bots*, Aquellos usuarios que tengan más de 3 000 publicaciones.
- Eliminación de usuarios no asiduos. Aquellas usuarios que solo publicaron una vez
- Publicaciones con menos de tres palabras.

En particular, el último punto permite eliminar tweets que no contengan información sustancial acerca de las prácticas de los usuarios de Twitter en la ciudad. Este criterio ha considerado elementos de texto para la depuración de la información y solo ha transformado la base de datos en términos del número de observaciones. Una vez en este punto, el siguiente paso ha sido depurar el texto, debido a que existen palabras o elementos textuales que impiden realizar un buen análisis sobre los tweets. A continuación se enumeran los criterios de depuración, la justificación de éstas y algunos de sus inconvenientes. Para ejemplificar cada uno de los pasos realizados en la depuración se toma como base el siguiente tweet hipotético.

Cuadro 4.1: Texto hipotético de un tweet

Yaaaaaaa ven a la Ciudad de México xfa, para comer en elmichoacanisimo #noeslomismosinti 😞😞😞😞 <a href="https://www.youtube.com/watch?v=-FM6KmpvfMk">https://www.youtube.com/watch?v=-FM6KmpvfMk</a>
--

Fuente: elaboración propia

El primer paso en la depuración del texto ha sido eliminar palabras con más de 16 caracteres. Esto implica la eliminar palabras que probablemente no tengan algún significado para el análisis de contenido. En el tweet hipotético se puede observar que la palabra *#noeslomismosinti* contiene más de 16 caracteres, por lo tanto es eliminada del análisis final. En estricto sentido, ésta no es una palabra, ya que se encuentra precedida de un #, símbolo utilizado en los mensajes de Twitter para denotar etiquetas particulares sobre un tema. Estas etiquetas son usadas para nombrar un tópico en específico y pueden ser una sola palabra o frases sin espacios, como en este caso. Una posible desventaja de eliminar texto bajo este criterio es quitar etiquetas que, en conjunto, hagan referencia

a algún tema importante dentro de la ciudad. Otro elemento a considerar es el criterio arbitrario de 16 caracteres, ya que éste puede eliminar palabras de gran extensión con significado. En el caso de la palabra *elmichoacanisimo* hubiese sido eliminada de implementar un umbral de 15 caracteres. A pesar de los posibles problemas que implica la eliminación de palabras de una longitud de 16 caracteres, esta estrategia permite modelar los contenidos de manera eficiente.

El segundo elemento de depuración ha sido reducir las palabras con una misma letra repetida. En el caso del tweet de muestra se tiene la palabra *Yaaaaaaa*, la cual puede ser reducida simplemente a la palabra ‘Ya’. Esto se ha hecho porque, muchas personas utilizan la repetición de una vocal para enfatizar alguna palabra o acción. Sin embargo, no utilizan la misma cantidad de letras repetidas, por lo tanto puede provocar tener diferentes palabras que en realidad sean una sola. Una limitación de la reducción de las letras repetidas hacia la palabra original es la pérdida de énfasis en el sentido de la oración. Aunque, esta dimensión escapa al análisis de este estudio, la modelación de este tipo de expresiones, continua siendo un reto.

La tercera acción ha sido la eliminación de los *emojis*. “Un *emoji* en particular cubre un amplio y creciente rango de áreas y temas y, son usados por 92 % de la población internauta, con mujeres y jóvenes menores de 30 como sus usuarios más recurrente. Los *emojis* son empleados de manera creciente, especialmente entre los jóvenes, como una forma rápida de transmitir emociones y actitudes (por ejemplo, sarcasmo) en los servicios de mensajes instantáneos y en redes sociales. Desde un punto de vista lingüístico los *emoticones* y *emojis* aparecen en lugar de las palabras y sirven como una señal subrogada y no verbal y, contribuye a resaltar el significado del mensaje escrito” (Marengo, Giannotta, y Settanni, 2017, p: 74) <sup>2</sup>. Aunque los *emojis* son representaciones gráficas de expresiones o palabras, éstos son interpretados por las computadoras por medio de una cadena particular de números y letras, a estas combinaciones se le denomina texto *UNICODE*. Dado que la interpretación computacional de texto recupera la codificación *UNICODE* es complicado realizar análisis de texto sobre estos elementos. En cuando a una desventaja, se tiene la pérdida de significado en el conjunto de la oración. Tal como menciona Marengo y cols. (2017), la función de los *emojis* es reafirmar o transmitir un significado de manera más rápida y eficaz. De tal manera que, al eliminar estos elementos, se está perdiendo de esta riqueza de significados.

El cuarto paso para depurar texto ha sido la eliminación las direcciones URL. En el caso del

---

<sup>2</sup>Traducción libre el autor

tweet hipotético, el usuario ha compartido un enlace de YouTube. Al igual que en el caso de los *emojis*, los enlaces URL no aportan ningún significado interpretativo a la oración. Pero, esto, al igual que los elementos gráficos, refuerzan significados, mensajes y prácticas entre los usuarios de Twitter. Recientemente, las plataformas como Twitter permiten compartir contenidos audiovisuales, esta estrategia ha permitido que los usuarios se interrelacionen no solo por medio de texto, sino por elementos audiovisuales y esto ha complejizado el análisis de la dimensión digital.

El quinto elemento eliminado son las palabras vacías. Las palabras vacías son definidas como aquellas palabras con poco o nulo significado. Éstas son usualmente removidas del texto durante el procesamiento con el propósito de mantener el máximo de significado y contexto. En la generalidad estas palabras son las más frecuentes en cualquier idioma. Ejemplos de éstas son las preposiciones, artículos, números, pronombres, algunos verbos, etc (Sarkar, 2016). La lista de palabras vacías no se define claramente en algún idioma. Sin embargo, para este estudio se han utilizado las palabras vacías en inglés y español propuestas por Feinerer y Hornik (2018); Feinerer, Hornik, y Meyer (2008). Por ejemplo en el tweet del cuadro 4.1 se elimina las palabras: ya, a, la, para.

Previamente a eliminar las palabras vacías se identificaron las ‘entidades nombradas’. Las entidades nombradas son aquellos términos particulares que representan entidades que son más informativas y tienen un contexto único comparado con el resto del texto. Estas entidades refieren a términos que representan en el mundo real objetos como: personas, lugares, organizaciones, etc. (Sarkar, 2016). Para el caso particular de la ZMCM se hace énfasis en las entidades nombradas que refieren a lugares<sup>3</sup>. En el caso del tweet del cuadro 4.1 la entidad nombrada es ‘Ciudad de México’. Dado que esta entidad nombrada se encuentra unida por la preposición ‘de’, ésta se conserva dada esta entidad nombrada en específico.

En general, en el campo de la lingüística computacional, el reconocimiento de entidades nombradas es un campo amplio y de creciente interés dentro de la minería de datos. En este sentido, se han desarrollado herramientas computacionales sofisticadas para detectar de manera automatizada dichas entidades (Lample, Ballesteros, Subramanian, Kawakami, y Dyer, 2016). Sin embargo, para el desarrollo de este trabajo se ha optado por la identificación manual de estas entidades nombradas, esto se ha debido a dos razones. La primera de ellas por el objetivo de identificar lugares nombrados en estas entidades. Es decir, encontrar combinaciones de palabras que aludieran a sitios

---

<sup>3</sup>Para ver la lista completa de entidades nombradas en la Ciudad de México, consultar, en el cuadro A.2 del Anexo

físicos en la ciudad. La segunda razón ha sido de carácter técnico, ya que la identificación de estas entidades nombradas por medio de algoritmos requiere de mayor conocimiento en el terreno del procesamiento de lenguaje natural. En consecuencia la identificación de entidades nombradas se ha realizado por medio de encontrar los bigramas o trigramas de palabras con mayor frecuencia relativa y con alusión a lugares en la ciudad. Esta estrategia ha permitido acotar la identificación de entidades de sentido espacial, pero ha sido poco eficiente en términos del número de entidades recuperadas.

#### *4.4.2 Métodos para el análisis de texto. Frecuencias y Modelo de Asignación Latente de Dirichlet*

Como se ha mencionado en la subsección anterior, existen algunos elementos estandarizados acerca del pre-procesamiento del texto, entre ellos la conversión de todos los caracteres a minúsculas o la eliminación de palabras vacías. En contraste, también existen procesos específicos en función del objetivo de análisis y, una vez que se ha depurado el texto se procede al análisis y modelado de éste. En esta subsección se presentan las dos técnicas que se utilizan para indagar sobre el poder de la comunicación para generar una gentrificación aumentada digitalmente por medio del texto de Twitter. Para ello se tienen dos propósitos, proponer que las palabras más repetidas en los tweets emitidos en colonias gentrificadas difieren sustancialmente del texto emanado de colonias no gentrificadas. Y, además, que las palabras más repetidas en los primeros espacios son referenciales y sintomáticas del proceso de gentrificación. Aunado a lo anterior, dado que la frecuencia de palabras, por sí misma, no proporciona elementos suficientes para estudiar los temas en torno a los que gira un *corpus* determinado, se ofrece la modelación de tópicos latentes por medio de la Asignación Latente de Dirichlet (ALD), con el objetivo de identificar los principales temas de Twitter y conocer si éstos tienen un patrón geográfico determinado.

El texto es considerado como una fuente de información no estructurada (Kitchin, 2014b). Es decir, obtener el mensaje de una frase conlleva análisis de tipo cualitativo. Sin embargo, con el avance de la computación se analiza este tipo de información desde una vertiente denominada, procesamiento de lenguaje natural (PNL). Este tipo de análisis ha intentado compaginar los elementos estadísticos y computacionales de las fuentes de información numérica para extraer información valiosa de fuentes de información textuales o lingüísticas.

El primer acercamiento al análisis de texto es la frecuencia relativa de las palabras. Es decir,

la frecuencia de las palabras es un indicador de la importancia que tienen éstas en un entorno determinado. En ese sentido, dado un conjunto finito de palabras únicas en el conjunto de todos los tweets recolectados, la frecuencia relativa de una palabra será las veces que ha aparecido con respecto al total de las palabras captadas.

$$fr_i = \frac{f_i}{N} \quad (4.1)$$

Donde:  $fr_i$  = Frecuencia relativa de la palabra  $i$  en el conjunto de palabras.

$f_i$  = Frecuencia absoluta de la palabra  $i$

$N$  = Número total de palabras

La frecuencia de palabras es analizada en función de su localización y horario. Es decir, se ofrece un estudio comparativo de las principales palabras según tipo de colonia y horario con el propósito de identificar si dichas palabras muestran un patrón diferenciado en función del tiempo y espacio en que se emiten. Sin embargo, la frecuencia de una palabra no da cuenta del complejo universo de las temáticas que pueden emerger en Twitter. Ante esto, se han comenzado a desarrollar técnicas que permiten categorizar el conjunto de palabras en temas latentes. En este ámbito se encuentra la técnica de la ALD. La ALD permite la conformación de tópicos semánticos latentes con el objetivo de identificar temas referentes al fenómeno de gentrificación. A diferencia del análisis de frecuencia de palabras aisladas, el modelo de ALD permite construir conjuntos de palabras semánticamente cercanas.

La ALD se encuentra inserta en el paradigma del *aprendizaje de máquina* (AM). Dentro del paradigma del *aprendizaje de máquina* se encuentran cuatro tipos de enfoques. El primer enfoque es el AM supervisado. En este enfoque se tiene un conjunto de etiquetas que serán aplicadas a un conjunto de datos, dichas etiquetas son previamente conocidas. Por ejemplo, bajo el paradigma de AM se pueden realizar ajustes de tipo lineal o binomial. El segundo enfoque es el aprendizaje no supervisado. En este caso, el AM no es provisto de una respuesta previa para evaluar los datos, para ello el paradigma se basa en alguna medida de similaridad o distancia entre cada observación, de esta manera se obtiene la posible asociación de los datos. Ejemplo de este tipo de AM es el análisis de clústers. El tercer tipo de AM se le denomina semi-supervisado. Este tipo de algoritmos recuperan se construyen con base en algunas de las etiquetas de manera *a priori*, ya que el proceso de conoci-

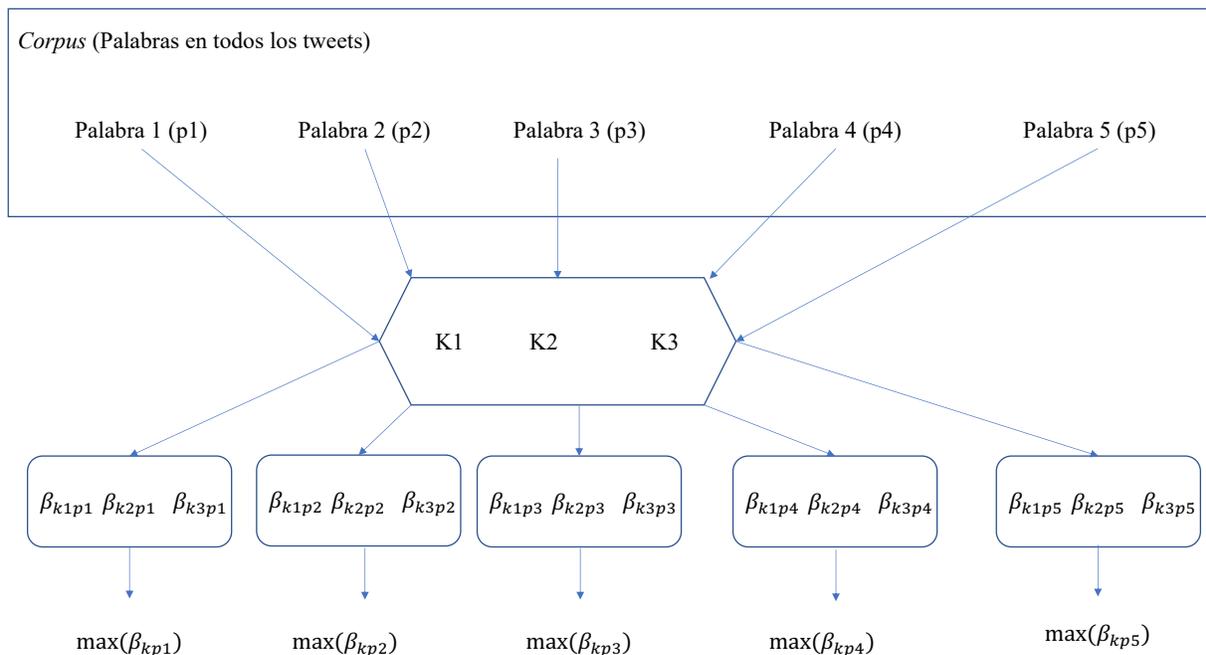
miento exhaustivo de etiquetas previamente puede ser costoso, y se utilizan técnicas de clúster para conocer la totalidad de las etiquetas. Por último, el AM de aprendizaje reforzado para enfoque en realizar continuamente los ajustes sobre el aprendizaje del algoritmo cuando los enfoques supervisado o no supervisado llegan a ser muy costosos o toman mucho tiempo. Este tipo de paradigma es el utilizado en la programación de AM en automóviles autónomos (Ramasubramanian y Singh, 2017). En particular, la Asignación Latente de Dirichlet se ubica en el AM no supervisado. Esto se debe a que no se le imponen etiquetas existentes al *corpus* de interés.

La ALD es una técnica probabilística que permite clasificar un documento particular dentro de un conjunto de temas latentes (Blei, Ng, y Jordan, 2003). El ALD se basa en la idea de creación de campos semánticos que se asocian con la construcción de modelos sobre temas. Para Jurafsky y Martin (2018) la ALD es un algoritmo de AM no supervisado que permite extraer un conjunto de palabras asociadas a un número determinado de temas provenientes de conjuntos grandes de texto. En otras palabras, es una herramienta de modelado generativo de texto; reduce a un pequeño grupo de temas un conjunto complejo de texto (Griffiths y Steyvers, 2004).

De manera formal, el modelo de ALD se constituye con los siguientes elementos. Primero, se considera como unidad central a las palabras ( $p_q$ ), así con el conjunto de todos los documentos se tendrá el total de las palabras posibles. Segundo, un tweet o documento ( $t_i$ ), es un conjunto determinado de palabras  $p_q$ . El tercer elemento, el conjunto de todos los tweets conforma el *corpus*. En cuarto lugar se determina un número fijo de clústers o temas latentes  $K$ . Una vez establecidos estos elementos a cada palabra  $p_q$  se le asigna un tema latente  $K$  por medio de un vector  $\beta$  que representa la probabilidad de que cierta palabra pertenezca a cada tema latente  $K$ . De esta manera, la palabra se asigna al  $\beta_k$  que tenga la mayor probabilidad. Este procedimiento permite la conformación de campos semánticos interpretativos. Finalmente, para cada tweet ( $t_i$ ) del *corpus*, y dado el conjunto de  $p_q$  que lo compone, se genera un vector  $\alpha$  que significa la probabilidad de que esa publicación pertenezca a cada tema latente. Así, el tweet se clasifica en función de la maximización de  $\alpha_k$  dado el conjunto de palabras en  $t_i$ .

Para ejemplificar el funcionamiento de la ALD, se supone un *corpus* que se compone de 2 documentos o tweets. En éstos, se pueden encontrar 5 palabras diferentes, mientras que el número de campos semánticos que se desea construir son  $K = 3$ . De esta manera, la creación de temas latentes con base en las palabras será la siguiente.

Figura 4.1: Modelo gráfico de la ALD para la creación de temas latentes

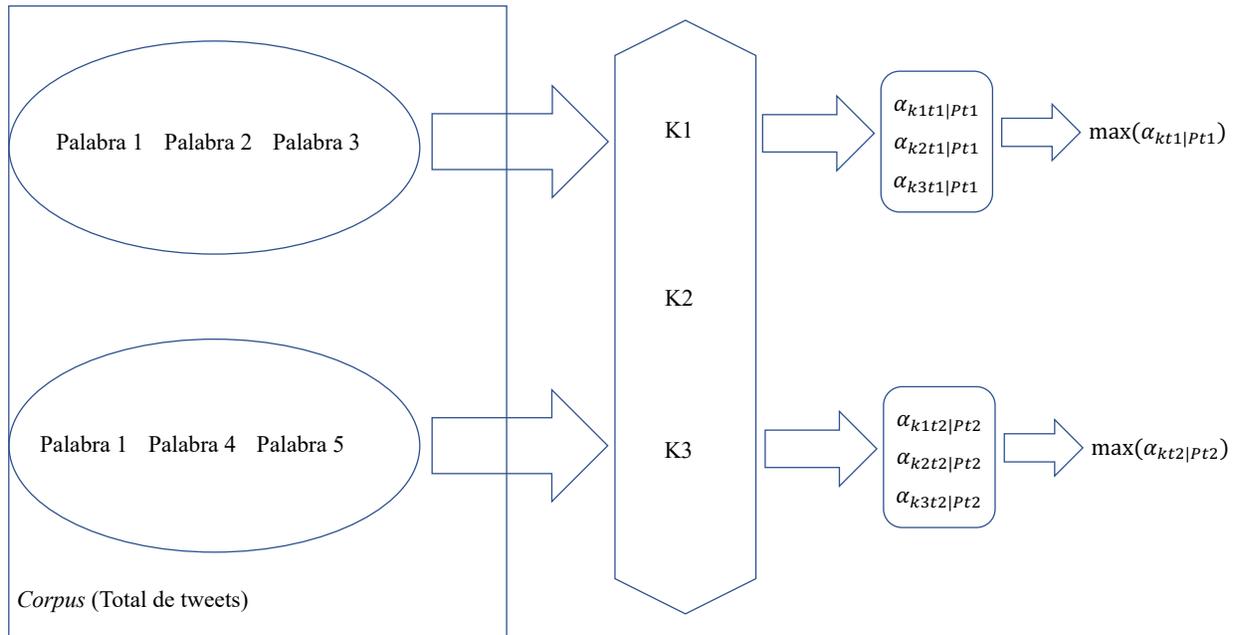


Fuente: Elaboración propia con base en Blei y cols. (2003)

La figura 4.1 representa la creación de 3 temas latentes dado un conjunto de 5 palabras. En la parte superior del diagrama se tiene el conjunto de las 5 diferentes palabras, éstas son sometidas a la pertenencia de alguno de los 3 tópicos. Esta etapa es iterativa, hasta el punto en que se garantice la máxima homogeneidad al interior de cada tópico y la máxima heterogeneidad entre ellos. En el siguiente paso, para cada palabra se le asigna un valor del vector  $\beta$  que representa la probabilidad de la palabra de pertenecer a cada tema. Finalmente, la asignación temática se determina por medio de la maximización del valor del vector  $\beta$  de cada palabra. Dado, que dichas probabilidades se consiguen por la cercanía de aparición de cada palabra con respecto a las demás, es posible tener una misma palabra en más de un tema, pero con diferente probabilidad de pertenecer a éste. La conformación de temas latentes de manera interpretativa se da por medio del conjunto de palabras que los componen. Aunado a esto, la otra ventaja de la estimación permite categorizar cada documento en cada uno de los campos semánticos encontrados. Esto se representa en el siguiente diagrama.

La figura 4.2 muestra el método para categorizar los tweets en cada campo semántico conformado. En el cuadro se encuentra el *corpus* completo conformado por dos tweets. Entre ellos,

Figura 4.2: Modelo gráfico de la ALD para la asignación de temas latentes



Fuente: Elaboración propia con base en Blei y cols. (2003)

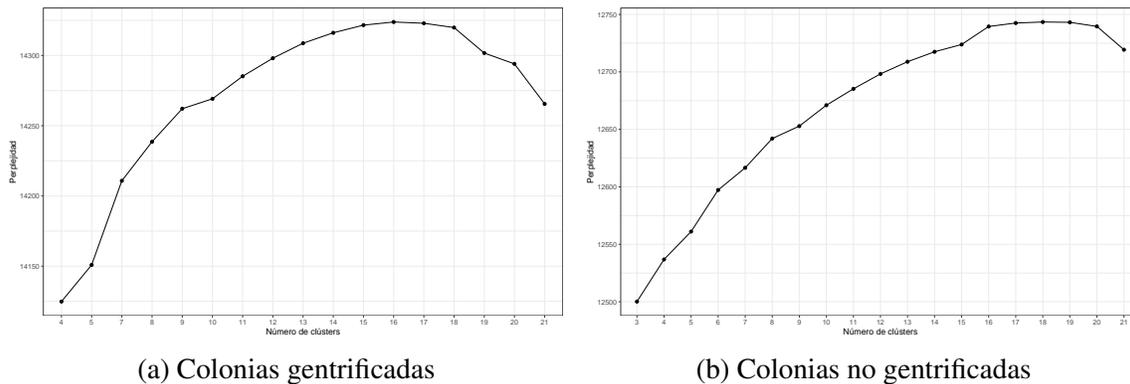
cada tweet se constituye por un conjunto dado de palabras. Dadas estas características, el tweet se somete a la clasificación de los 3 campos semánticos preestablecidos, esto a su vez, genera el vector  $\alpha$  que representa la probabilidad de dicho tweet de pertenecer a cada uno de los temas. Por último, la asignación del tweet a cada campo semántico se realiza mediante el valor máximo del vector  $\alpha$  del tweet.

En resumen, el modelo de ALD arroja dos resultados importantes. Por un lado, el procedimiento categoriza cada entrada o tweet dentro de alguno de los temas formados. Es decir, se le asigna a cada tweet un campo semántico dada la combinación de palabras que lo componen. Por otro lado, la ALD genera tópicos basado en el conjunto de las principales palabras las cuales son semánticamente diferentes entre cada grupo formado. Y, la formación de los clústers se basa en la máxima homogeneidad semántica al interior de cada uno de estos y la máxima heterogeneidad entre ellos (Stehle, 2017). Para el caso de la ZMCM se han tomado las muestras descritas en el capítulo anterior y se han obtenido los ajustes de ALD para las submuestras de tweets emitidos en

colonias gentrificadas y los tweets emitidos en colonias no gentrificadas <sup>4</sup>.

Como se ha mencionado, uno de los componentes de ALD es el número de clústers. Éstos se encuentran dados *a priori* para la realización de la formación de campos semánticos. De esta manera, el siguiente paso es la determinación del número de temas a formar. Al respecto, la perplejidad del modelo ayuda a medir el ajuste desde el paradigma del procesamiento de lenguaje natural. La perplejidad indica la incertidumbre de predecir una palabra dentro del modelo, mientras la perplejidad del modelo sea menor, mejor será la capacidad predictiva del ajuste temático. Sin embargo, hay que mencionar que mientras el número de temas incrementa la interpretación analítica del modelo se torna complicada. Por esta razón, en el presente caso se ha decidido combinar el análisis de la perplejidad de los modelos de ALD desde 3 clústers hasta 21 con la propuesta de Lansley y Longley (2016) que proponen analizar 20 temas para la ciudad de Londres. A continuación se presenta la perplejidad para cada submuestra de tweets.

Figura 4.3: Perplejidad de los ajustes ALD dado diferente número de clústers



Fuente: Elaboración propia

En la figura 4.3 se muestra la perplejidad para ambas submuestras. Las gráficas muestran que alrededor de 16 temas dados se alcanza la mayor perplejidad, lo que implica que este número sería el peor ajuste posible. El comportamiento de este indicador muestra una tendencia decreciente a partir de este punto. Por lo tanto, se ha elegido 20 tópicos para ambas submuestras. Es decir, para el caso de las colonias gentrificadas y no gentrificadas de las ZMCM se analizarán 20 temas latentes en los textos de Twitter.

<sup>4</sup>El ajuste de la Asignación Latente de Dirichlet se ha realizado por medio del paquete “*topicmodels*” dentro del lenguaje de programación R (Grün y Hornik, 2011)

La elección del modelo de ALD se ha debido a tres elementos. En primera instancia, como se ha explicado el algoritmo pertenece la familia de AM no supervisado, esto ha facilitado la construcción *a priori* de etiquetas para categorizar los tweets. Es decir, en contraposición al algoritmo más utilizado para el análisis de texto de grandes datos, el análisis de sentimientos, el ALD permite no imponer ninguna estructura previa a la información, en este sentido, el enfoque de ALD permite visualizar si entre los principales tópicos se encuentran los relacionados con elementos de gentrificación. El segundo elemento importante para elección de la ALD ha sido por motivos de recursos computacionales y temporales. Dado que métodos supervisados implican un conocimiento previo de los datos esto conlleva a la necesidad de tener un conjunto de tweets entrenados para identificar temas de gentrificación, este conjunto de palabras es inexistente hasta la fecha. En consecuencia, la elección de un método no supervisado ha ahorrado recursos en el sentido de identificación de categorías léxicas. Finalmente, se ha elegido el modelo por la capacidad de poder identificar otros temas en espacios gentrificados de los cuales no se tiene registro en la literatura. En este sentido, los métodos supervisados ofrecen una ventaja dado que le dan más robustez a la identificación y análisis, pero dejan de lado posibles temas relacionados que puedan estar en auge. Por otro lado, el ALD abre la posibilidad de encontrar temas emergentes en espacios gentrificados y por ende proponer la coexistencia de paisajes lingüísticos hasta la fecha no estudiados.

#### **4.5 Resultados empíricos. Una aproximación al poder de la comunicación ante una gentrificación digitalmente aumentada.**

Una de las dimensiones de la gentrificación aumentada digitalmente es el poder de la comunicación. Esta dimensión se encuentra vinculada al poder de distribución y del código. En el capítulo anterior se ha desarrollado la idea del poder de distribución desde una perspectiva espacio-temporal en la ZMCM. Los resultados apuntan a la conformación de espacios gentrificados en función de la hora y del día de emisión en Twitter. Sin embargo, esta dimensión se ve reforzada por los mensajes o discursos que las personas emiten en dicha plataforma. Con esto en mente, en este subapartado, se exponen los resultados del análisis de ALD bajo el marco interpretativo de la función del lenguaje de Bühler. En este sentido, las palabras que se encuentran en el texto de los usuarios de Twitter reflejan un síntoma, en este caso el proceso de gentrificación. Al mismo tiempo, las palabras hacen referencia a hechos concretos de la realidad (función referencial) y las palabras son interpretadas

por receptores que les asignan un determinado significado (función apelativa), aunque esta última no será retomada en este estudio. En consecuencia, los resultados son interpretados bajo la premisa de que las palabras y temas de Twitter en la ZMCM son referenciales a cierto conjunto de hechos que se relacionan con el proceso de gentrificación.

El análisis de fenómenos urbanos por medio de la información contenida en el texto de Twitter ha comenzado a crecer en los últimos años. En este marco, Lansley y Longley (2016), ofrecen un análisis de los temas de Twitter en la ciudad de Londres, los autores identifican temas que reflejan la composición social y económica de dicha ciudad. Por ejemplo, encuentran que los principales temas se asocian a moda y diseño, deportes, espectáculos, vida nocturna o museos. En este sentido, los autores ofrecen la ubicación geográfica de dichos temas, de esta manera, encuentran que estos temas se vinculan espacialmente con lugares icónicos a cada uno de ellos. Este es el caso del tema de la vida nocturna que se relaciona con el barrio del Soho. En otra dimensión del análisis, los autores señalan que el campo semántico relacionado con insultos e injurias se ubica en áreas de la ciudad deprimidas económicamente o donde habitan personas obreras. Este ejemplo da cuenta del síntoma de las palabras, particularmente, las ofensas en Twitter se relacionan con estratos bajos. Por su parte, Lai y cols. (2017) continúan la base de análisis semántico de Twitter en la capital inglesa, aunque exploran otra dimensión, la publicidad. En dicho trabajo, el principal hallazgo es la correlación espacial entre los temas emanados de Twitter que vinculan a publicidad con espacios de la ciudad que son concurridos por un gran número de personas, tal es el caso del metro de dicha urbe. Cabe resaltar que ambos trabajos utilizan la técnica de ALD para obtener tales resultados. Sin embargo, Zhong, Zeng, Tu, y Yoshida (2018) mencionan que los resultados del ALD se modifican según la escala geográfica de análisis. Así, los autores identifican que al reducir la escala espacial de estudios de los textos de Twitter se obtienen mayor diversidad de temas en comparación con una escala superior. Ellos mencionan que esto se debe a la naturaleza de la información, debido a que ésta se encuentra condicionada por el tipo de personas que emiten y lo que quieren expresar.

#### *4.5.1 Las principales palabras de Twitter en la ZMCM*

Como se expuso en el apartado metodológico de este capítulo, el primer elemento a considerar para el análisis de texto es la frecuencia relativa de las palabras. De esta manera, se presentan las principales palabras de los tweets emitidos en colonias gentrificadas y no gentrificadas. En términos

temporales se comparan el patrón de palabras según cuatro horarios basados en la metodología de clasificación temporal del capítulo anterior. En específico, los horarios de días laborales de 7:00 a 12:59 horas, días laborales de 18:00 a 21:59 horas, fines de semana de 13:00 a 17:59 horas y de 18:00 a 21:59 horas. Esta selección se ha hecho basado en el comportamiento de los núcleos de emisión en dichos periodos de tiempo. En primera instancia, se presentan las principales palabras en ambos tipos de colonia de manera general.

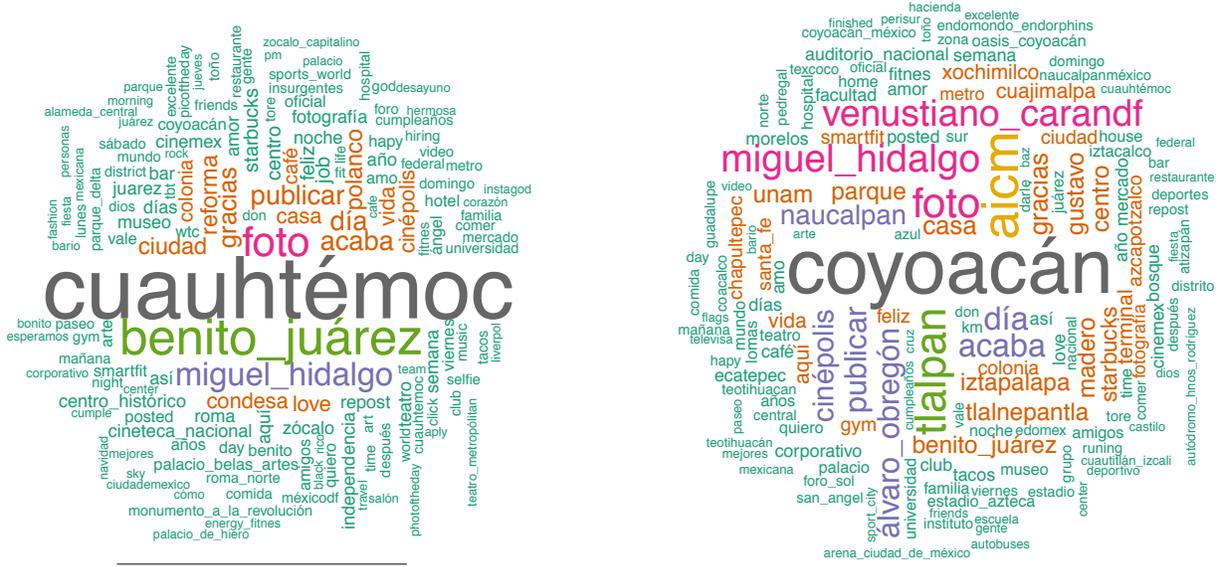
En la figura 4.4 se presentan las nubes de palabras principales de los tweets emitidos en colonias gentrificadas y no gentrificadas <sup>5</sup>. En primer lugar, se observa que las palabras más frecuentes cambian en función de la colonia de emisión. Este comportamiento diferenciado de las principales palabras indicaría que el poder de la comunicación de la Realidad Geográfica Aumentada Digitalmente interviene en la conformación de una gentrificación aumentada digitalmente. Esto bajo la premisa de que espacios gentrificados emiten discursos diferentes en Twitter en comparación con espacios no gentrificados. Sin embargo, independientemente de la colonia el principal tipo de palabra refiere a las alcaldías de la Ciudad de México o algunos municipios del Estado de México. Tal es el caso que, las principales palabras en colonias gentrificadas son Cuauhtémoc, Benito Juárez y Miguel Hidalgo (figura 4.4a), mientras que en las colonias no gentrificadas sobresalen las alcaldías de Tlalpan, Coyoacán, Miguel Hidalgo y Venustiano Carranza y, en menor medida Naucalpan, Tlalnepantla y Ecatepec (figura 4.4b). La frecuencia de estas palabras indica que el texto de los tweets se refiere más a las alcaldías centrales de la Ciudad de México y en el caso del Estado de México a los municipios contiguos a la Ciudad de México. La frecuencia de las palabras permite identificar la inexistencia de las menciones a los municipios periféricos de la ZMCM.

La figura 4.4 también muestra la mención a entidades nombradas dentro de las alcaldías. Por un lado, los lugares más nombrados en colonias gentrificadas hacen alusión a las colonias Roma, Condesa y Polanco, mientras que en el conjunto léxico de las no gentrificadas la mención a colonias es inexistente. Al mismo tiempo, ambos tipos de colonias presentan, entre sus palabras más frecuentes, lugares de consumo. Este es el caso de *Starbucks* o Cinépolis, ambos lugares de consumo vinculados al entretenimiento o esparcimiento. Esta primera aproximación permite plantear que el lugar o espacio de emisión es factor fundamental para el contenido digital encontrado en Twitter.

---

<sup>5</sup>Para generar este tipo de presentación gráfica se han utilizado los paquetes estadísticos desarrollados por Feinerer y cols. (2008) y Fellows (2018)

Figura 4.4: Principales palabras de los tweets de la ZMCM, según colonia de emisión



Palabras	Porcentaje
cuauhtémoc	1.79
benito_juárez	0.93
foto	0.71
miguel_hidalgo	0.5
acaba	0.39
día	0.38
publicar	0.38
gracias	0.33
reforma	0.3
ciudad	0.29
polanco	0.29
cinépolis	0.26
condesa	0.26
casa	0.25
vida	0.24
love	0.24
café	0.23
colonia	0.23
centro	0.21
starbucks	0.21
job	0.21
feliz	0.21
aquí	0.19
roma	0.19
cineteca_nacional	0.17

(a) Colonias gentrificadas

Palabras	Porcentaje
coyoacán	1.15
aicm	0.71
tlalpan	0.59
foto	0.56
miguel_hidalgo	0.5
venustiano_carandf	0.48
acaba	0.37
día	0.37
publicar	0.37
álvaro_obregón	0.35
cinépolis	0.33
naucalpan	0.31
casa	0.29
gracias	0.27
iztapalapa	0.26
tlalnepantla	0.26
madero	0.25
gustavo	0.25
parque	0.24
benito_juárez	0.24
unam	0.24
centro	0.24
starbucks	0.22
vida	0.2
aquí	0.2

(b) Colonias no gentrificadas

Fuente: Elaboración propia.  
 La gráficas se muestran la sección de anexos figuras A.1 y A.2

Aunque, como se desarrolla a continuación, este patrón léxico cambia en función del tiempo; esta dimensión se analiza a continuación.

De manera general, el patrón léxico diferenciado entre tipo de colonia muestra un primer indicio del poder diferenciado de la comunicación entre espacios gentrificados y no gentrificados.

En particular, las palabras más frecuentes en colonias gentrificadas muestran el nombre explícito de éstas como indicador de distinción espacial en los mensajes de Twitter. En este sentido, se podría interpretar que la gentrificación se aumenta digitalmente ante la mención en Twitter de colonias identificadas como elitizadas. Como se ha expuesto en el capítulo anterior, la emisión de tweets difiere a lo largo del día en función de la dinámica cotidiana urbana. Para el caso del poder de la comunicación reflejada en los mensajes de texto se puede encontrar un patrón semejante.

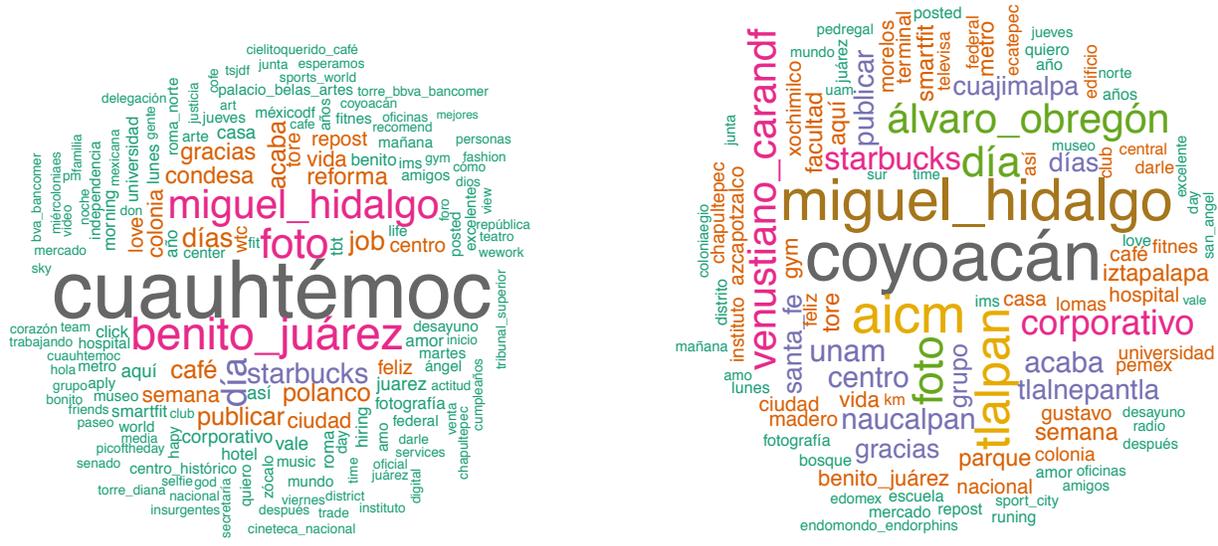
Como se muestra en la figura 4.5, las principales palabras según tipo de colonia muestran tendencia a mencionar actividades laborales. En las colonias no gentrificadas muestran mención a corporativos y, particularmente, a la zona de Santa Fe. Vinculado a esto, la Universidad Nacional Autónoma de México (UNAM) y la palabra facultad resaltan entre las principales. Estas menciones sugieren el uso por parte de personas con escolaridad profesional y actividades vinculadas con ello <sup>6</sup>. Las colonias no gentrificadas muestran actividades relacionadas con ejercicio físico, esto se observa en las palabras que hacen referencia a gimnasios en la Ciudad, éstas son Smartfit, Sport City, *fitness*. En menor medida, los usuarios en estos tipos de colonia emiten palabras relacionadas con el estado de ánimo, feliz o amo. Este tipo de palabras reportan frecuencia importante en colonias gentrificadas, éste puede ser un indicador del análisis de sentimiento entre el espacio urbano. Mientras en las colonias centrales los usuarios autorreportan ser “felices”, los usuarios de colonias no gentrificadas reportan en menor medida palabras asociadas a este sentimiento. Otro aspecto lingüístico a resaltar, es la presencia de palabras vinculadas con trabajo en idioma inglés, éstas son, *job*, *hiring*, *coffee*. Dicho elemento se explica por el hecho de los deslizamientos connotativos de palabras de otros idiomas. Es decir, el uso de palabras de origen externo al español señalan cierto nivel de estilismo y esnobismo al usarlas en español, aunque en su entorno original no denoten dicho prestigio (Thibault, 2004 citado por Franco, 2016).

El tema laboral en este horario es una constantes en ambos tipos de colonias en Twitter. Sin embargo, este punto debe ser matizado en torno al poder de la comunicación para generar realidades aumentadas digitalmente y por la concentración de tweets en este horario. En el primer punto, como se mencionó en el capítulo teórico, la construcción de realidades por medio de Google debido a los diferentes resultados que arroja según el idioma de búsqueda. En este sentido, para el caso de Twitter, se muestra que las colonias gentrificadas centrales apuntan a crear realidad

---

<sup>6</sup>Como se apuntó en la sección 2.5.2

Figura 4.5: Principales palabras de los tweets de la ZMCM, según colonia de emisión (días laborales, 7:00 hrs - 12:59 hrs)



Palabras	Porcentaje
cuauhtémoc	1.53
benito_juárez	0.75
foto	0.68
miguel_hidalgo	0.65
día	0.56
starbucks	0.39
días	0.36
job	0.35
café	0.34
acaba	0.34
publicar	0.33
polanco	0.33
reforma	0.3
gracias	0.3
condesa	0.27
semana	0.26
vida	0.25
colonia	0.24
ciudad	0.24
feliz	0.23
tore	0.21
centro	0.21
repost	0.21
love	0.2
wtc	0.19

(a) Colonias gentrificadas

Palabras	Porcentaje
coyoacán	0.92
miguel_hidalgo	0.73
aicm	0.68
tlalpan	0.6
foto	0.5
día	0.5
álvaro_obregón	0.47
venustiano_carandf	0.46
corporativo	0.39
starbucks	0.36
unam	0.32
naucalpan	0.31
centro	0.31
acaba	0.3
publicar	0.29
tlalnepantla	0.27
santa_fe	0.27
días	0.26
cuajimalpa	0.26
gracias	0.25
grupo	0.23
iztapalapa	0.23
tore	0.22
smartfit	0.22
parque	0.22

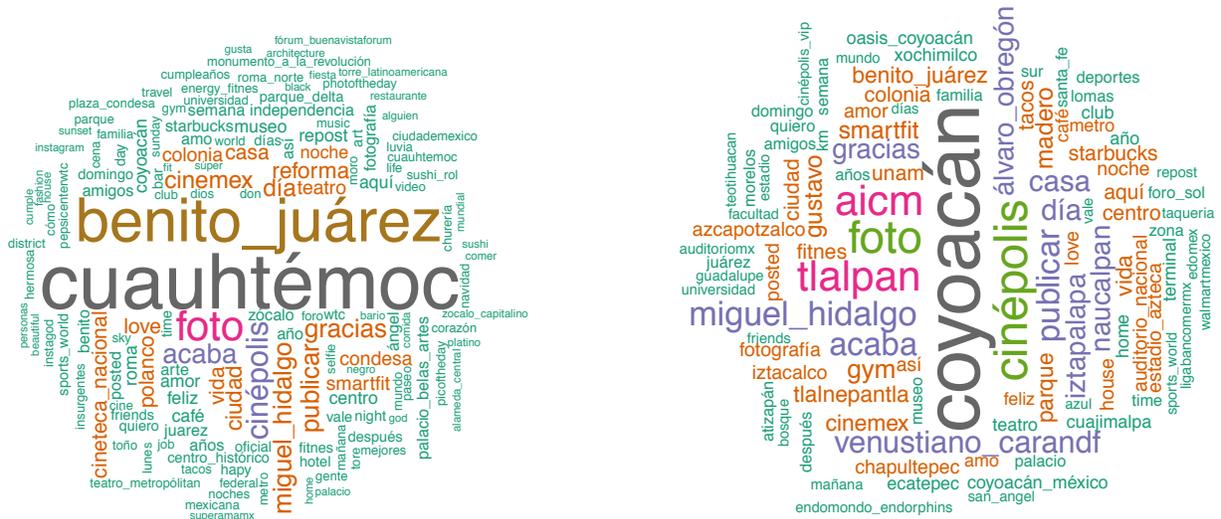
(b) Colonias no gentrificadas

Fuente: Elaboración propia.  
 La gráficas se muestran la sección de anexos figuras A.3 y A.4

por medio de menciones en idioma inglés en comparación con las palabras de la misma temática en colonias fuera de las centrales que son emitidas en español. Por otro lado, los signos que se emiten en cada lugar son diferenciados por la dimensión de volumen de publicaciones. Se debe

recordar que este horario genera núcleos de concentración en las colonias gentrificadas y en puntos estratégicos fuera de éstas, tal es el caso de la Zona de Santa Fe al poniente de la ciudad, esta entidad nombrada aparece en la figura 4.5b.

Figura 4.6: Principales palabras de los tweets de la ZMCM, según colonia de emisión (días laborales, 18:00 hrs - 21:59 hrs )



Palabras	Porcentaje
cuauhtémoc	1.58
benito_juárez	1.2
foto	0.77
cinépolis	0.51
acaba	0.4
miguel_hidalgo	0.39
publicar	0.39
gracias	0.38
día	0.33
cinemex	0.33
ciudad	0.31
reforma	0.3
love	0.26
cineteca_nacional	0.26
vida	0.25
casa	0.25
polanco	0.25
teatro	0.22
colonia	0.22
smartfit	0.22
condesa	0.21
noche	0.21
feliz	0.19
repost	0.19
amor	0.19

(a) Colonias gentrificadas

Palabras	Porcentaje
coyoacán	1.16
cinépolis	0.62
foto	0.62
aicm	0.54
tlalpan	0.54
miguel_hidalgo	0.41
acaba	0.4
publicar	0.39
venustiano_carandf	0.36
día	0.34
casa	0.32
iztapalapa	0.31
gracias	0.31
naucalpan	0.3
álvaro_obregón	0.29
gym	0.27
madero	0.26
gustavo	0.25
smartfit	0.24
parque	0.23
tlalnepantla	0.23
vida	0.23
cinemex	0.22
benito_juárez	0.22
unam	0.21

(b) Colonias no gentrificadas

Fuente: Elaboración propia.  
 La gráficas se muestran la sección de anexos figuras A.5 y A.6

La figura 4.6 muestra las principales palabras emitidas durante días labores en el horario de

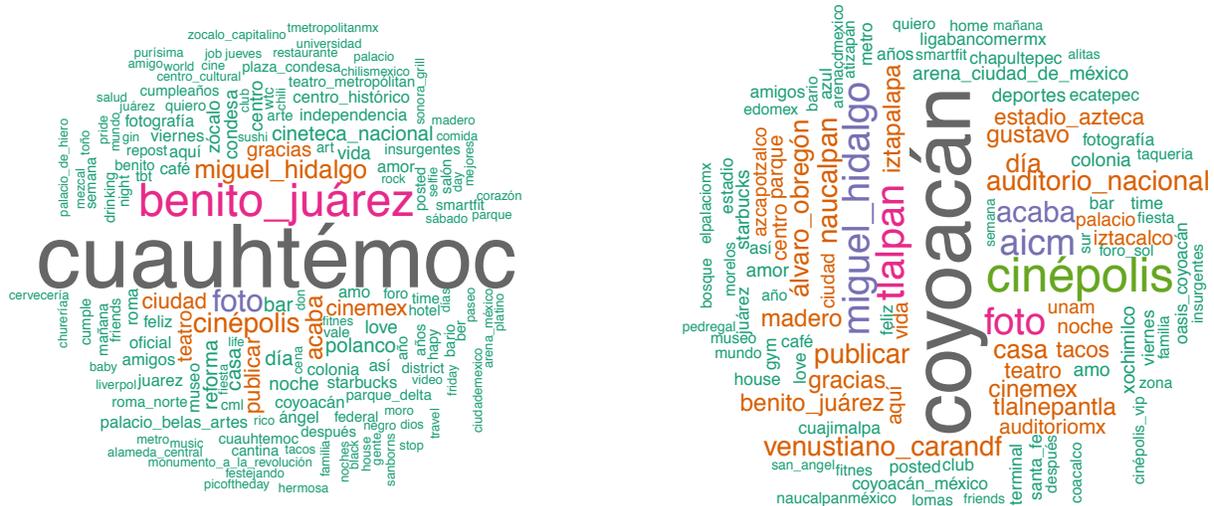
tarde-noche. Una primera impresión es que el patrón léxico de este horario con respecto a las figuras antecedente a éstas ha cambiado. Esto apunta a que los temas en Twitter varían no solo en el espacio, sino en el tiempo. En ésta se observa que las principales palabras continúan siendo los nombres de la delegación de emisión, aunque las palabras que le siguen demuestran ciertas prácticas cotidianas. Dichas prácticas se pueden agrupar en tres grandes direcciones, entretenimiento, ejercicio y comida. En lo que respecta a entretenimiento, sobresale las menciones a las multisalas de cine en la ciudad, en particular, Cinépolis para el caso de las emisiones de colonias no gentrificadas. Por su parte, este mismo rubro se reproduce en el otro tipo de colonias, se encuentra la palabra de Cinépolis y se agrega la entidad nombrada de Cineteca Nacional (figura 4.6a). Al respecto, este patrón se explica porque este tipo de salas cinematográficas se han convertido en la actividad de entretenimiento por excelencia en la zona metropolitana (Duhau y Giglia, 2016). Otro tipo de lugares de entretenimiento nombrados son espacios multitudinarios para espectáculos, tal es el caso del Auditorio Nacional, Foro Sol, Estadio Azteca o Buena Vista Fórum. El segundo tema que resalta en este horario es el referente al ejercicio. Las palabras relacionadas a éste son las marcas de gimnasios que se encuentran a lo largo de toda la ciudad. Ejemplo de esto son Smartfit, gym, Sports World. El tercer tema en común apunta al consumo de comida. Las palabras que se relacionan con este tema son tacos, café, Sushi Roll.

La diversidad léxica de las principales palabras en este horario permite postular dos puntos de análisis. El primero, las actividades reportadas en Twitter son contantes a lo largo de la ZMCM. Es decir, no importa si un usuario se encuentra en una colonia gentrificada o no, el uso reportado que se hace del espacio es similar. El segundo punto es la producción de publicidad de marcas por los propios usuarios. Esto se puede analizar bajo la interpretación del prosumo propuesto por Ritzer y Jurgenson (2010). Al respecto, los signos que refieren a marcas particulares son síntoma de un consumo determinado o actividades particulares, o bien, estas palabras, al referir a una marca particular adquieren un significado más allá del referencial y se transforma en un elemento económico para las empresas (Lara, 2015).

Por su parte, dicha homogeneidad léxica, vista desde la teoría de la gentrificación aumentada digitalmente puede indicar que los procesos de gentrificación que ocurren físicamente en colonias centrales de la ciudad encuentran eco más allá de los límites espaciales de dichos espacios. En otras palabras, por medio del poder de la comunicación el proceso de gentrificación digital reproduce

algunas de las actividades vinculadas al proceso de gentrificación fuera de espacios identificados como tales. De esta manera, el poder de la comunicación aumenta espacios por medio de la mención de símbolos gentrificadores en Twitter.

Figura 4.7: Principales palabras de los tweets de la ZMCM, según colonias de emisión (fines de semana, 18:00 hrs - 21:59 hrs )



Palabras	Porcentaje
cuauhtémoc	2.41
benito_juárez	1.07
foto	0.65
cinépolis	0.52
miguel_hidalgo	0.46
acaba	0.37
publicar	0.37
cinemex	0.36
ciudad	0.34
gracias	0.3
teatro	0.3
día	0.3
bar	0.29
reforma	0.29
polanco	0.28
cineteca_nacional	0.27
casa	0.26
condesa	0.26
noche	0.25
centro	0.25
love	0.23
vida	0.23
zócalo	0.2
colonia	0.2
aquí	0.2

(a) Colonias gentrificadas

Palabras	Porcentaje
coyoacán	1.43
cinépolis	0.76
tlalpan	0.59
foto	0.54
miguel_hidalgo	0.53
aicm	0.52
acaba	0.36
publicar	0.36
auditorio_nacional	0.34
naucalpan	0.34
venustiano_carandf	0.34
casa	0.32
álvaro_obregón	0.3
día	0.29
iztapalapa	0.28
madero	0.28
gustavo	0.28
gracias	0.27
cinemex	0.27
tlalnepantla	0.27
benito_juárez	0.26
teatro	0.24
estadio_azteca	0.24
tacos	0.23
centro	0.22

(b) Colonias no gentrificadas

Fuente: Elaboración propia.  
La gráficas se muestran la sección de anexos figuras A.7 y A.8

En las nubes de palabras de la figura 4.7 se muestra la diversidad léxica de las principales

palabras durante las tardes-noches de fines de semana. Por un lado, en las colonias no gentrificadas se muestran palabras referentes a sentimientos positivos, este es el caso de feliz, amo, *love*, querer. Otro tema sobresaliente son los deportes, así lo demuestra la propia palabra deportes, así como la marca de “Liga Bancomer MX”. Al mismo tiempo, existen algunas semejanzas con las palabras encontradas en el mismo horario durante días laborales. Este es el caso de los cines o la comida, la cual se interpreta a partir de la palabras “tacos”. Este horario muestra la configuración de espacialidades en la ciudad vinculadas al entretenimiento. Sin embargo, el entretenimiento que se puede encontrar en colonias gentrificadas y no gentrificadas es distinto. Esto se puede deber a dos factores, el primero a la distinción de las personas para divertirse en espacios elitizados y según por la infraestructura existente en cada tipo de colonia. Ejemplo de esto son las menciones a la liga de fútbol nacional, la cual se lleva a cabo generalmente en el Estadio Azteca. Este hecho compagina con los núcleos de emisión encontrados en el capítulo anterior fuera de colonias gentrificadas. Por otro lado, se aprecia la distinción del hecho de divertirse según el espacio de emisión. Mientras que los espacio de diversión fuera de colonias gentrificadas aluden a espacios hogareños y familiares, la diversión vinculada a colonias gentrificadas reflejan lugares de música, restaurantes o tiendas departamentales.

En este horario se podrían postular algunos apuntes teóricos de la RGAD basados en los hallazgos empíricos. Por un lado, la tendencia, en colonias gentrificadas, a mostrar signos que pueden ser interpretados como positivos, sentimiento que se podría vincular a espacios seguros y renovados. Por el otro lado, en las colonias no gentrificadas no resaltan las problemáticas de la periferia, sino actividades de esparcimientos y entretenimiento. En cierta medida, estas actividades reportadas en colonias no gentrificadas reproducen algunas de las actividades identificadas en colonias gentrificadas. En este sentido, la dimensión digital y, en particular, Twitter magnifican solo las actividades ‘positivas’ del espacio urbano, ocultando en la dimensión digital las problemáticas cotidianas de la ciudad.

Las nubes de palabras de la figura 4.8 dan clara muestra de la diferencia de los temas, que a su vez forman diferentes espacialidades. Este horario, es quizá, el más sintomático de generación de una realidad geográfica aumentada digitalmente vinculada con el proceso de gentrificación en la ciudad. Por un lado, las principales palabras de las colonias no gentrificadas apuntan a signos que se pueden interpretar como positivos, esto se sustenta en las palabras, gracias, *love*, quiero, vida, feliz,



muestra de la actividad económica central en dichos espacios, el consumo en espacios como bares o restaurantes. En este caso, aparecen de nueva cuenta palabras en inglés, tales como *drinking*, *party*, *friends* que encuentran su contraparte en colonias no gentrificadas, lo cual es síntoma del espacio elitizado que representan dichas colonias. En cuanto a las palabras que dan cuenta de los espacios de consumo elitizado se encuentran palabras como *club*, *pub* o pulquería. Esta última es significativa y sintomática del proceso de gentrificación cultural y simbólica de los espacios centrales. Ya que en últimos años el consumo elitizado ha recuperado las raíces de consumo alcohólico tradicional de México, en este caso las pulquerías y, las ha tomado como signo de distinción (Uribe, 2019).

Las últimas figuras (4.5,4.6, 4.7, 4.8) dan muestra de la conformación de espacialidades diferenciadas por medio de Twitter. Por ejemplo, sobresalen los las palabras con apelativo positivo en ambos tipos de colonias, aunque en diferentes horarios y días. Mientras la función apelativa vinculada a lo positivo se da en colonias gentrificadas durante las mañanas de los días laborales, esta misma aparece en los horarios de la tarde-noche los fines de semana en las colonias no gentrificadas. Otra diferencia de mismos temas aunque diferenciados por el tipo de colonia es la mención de espacio de cine. Mientras que en la generalidad las cadenas de cine en la ciudad aparecen de manera casi omnipresente, tanto en tiempo como en espacio, las referencia a la Cineteca Nacional es un síntoma claro del proceso de elitización del espacio aunque sea el mismo tipo de consumo. Ya que, la Cineteca tiene una única ubicación y se distingue por proyectar películas de corte internacional y del género, cine de arte. Esto es representativo debido a que aparece, prácticamente, con la misma frecuencia que las cadenas del ramo.

Aunque la frecuencia de palabras en el análisis espacio-temporal ha arrojado algunos indicios de la diferenciación, éstas no permiten identificar temáticas claras en los diferentes tipos de espacios en la ciudad. Por ello, a continuación se presentan los resultados del modelo de Asignación Latente de Dirichlet con el propósito de reforzar el argumento de la construcción de una gentrificación aumentada digitalmente por medio del poder de la diferenciación en la comunicación.

#### 4.5.2 *Los temas de Twitter en la ZMCM*

Las frecuencias relativas proporcionan un primer acercamiento a las temáticas de Twitter en la ZMCM. Aunque, este análisis solo se basa en la interpretación de las principales palabras fuera del contexto del mensaje. En el caso de las entidades nombradas que se han encontrado en el análisis

previo, no proporcionan información más allá del lugar. Por ello, para la identificación de espacialidades digitales y en particular de gentrificación aumentada digitalmente es necesario realizar un análisis más interpretativo de los mensajes escritos. Esta dimensión es complicada debido al inmenso volumen de texto con el cual se ha trabajado. Para subsanar dicho reto se ha utilizado el modelo de ALD expuesto en líneas anteriores. Con esto, el siguiente cuadro muestra solamente los temas latentes conformados del texto emitido en colonias gentrificadas de la ZMCM.

Cuadro 4.2: Temas latentes en el texto de tweets provenientes de las colonias gentrificadas de la ZMCM

Tema 1	Tema 2	Tema 3	Tema 4	Tema 5	Tema 6	Tema 7	Tema 8	Tema 9	Tema 10
Cultura hipster	Sentimientos/Celebración	Cultura gay	Espacio público de consumo	Clase creativa	Viajes/Turismo	Ejercicio	Comida y bebida	Tribunal	Trabajo de oficina
Benito Juárez	Cuauhtémoc	día	acaba	Miguel Hidalgo	Cuauhtémoc	Cuauhtémoc	Cuauhtémoc	Cuauhtémoc	Cuauhtémoc
Roma	district	picofteday	publicar	Cuauhtémoc	Benito Juárez	foto	Benito Juárez	job	Benito Juárez
Condesa	federal	colonia	foto	Cinépolis	foto	km	Miguel Hidalgo	foto	vida
Cineteca nacional	Benito Juárez	tacos	Cuauhtémoc	feliz	Cinépolis	centro	Condesa	click	foto
colonia	juarez	Cinépolis	Benito Juárez	Benito Juárez	Polanco	ángel	bar	hiring	día
Starbucks	benito	gayboy	Polanco	casa	love	endomondo_endorphins	feliz	méxicodf	ciudad
días	amor	Cuauhtémoc	parque_delta	año	casa	publicar	ciudad	god	Miguel Hidalgo
vale	Roma	instagay	ciudad	ángel	día	viernes	Starbucks	morning	gracias
Cuauhtémoc	Starbucks	Coyoacán	zócalo	sushi	Condesa	Reforma	Polanco	acaba	aquí
casa	cumpleaños	quiero	museo	domingo	gracias	finished	semana	justicia	centro
viernes	Cineteca nacional	café	Centro Histórico	Cineteca nacional	amo	runing	art	aply	noche
vida	colonia	vida	cinemex	Reforma	zócalo	independencia	cumpleaños	fit	amp
noche	gracias	Roma	chesscake	cinemex	dulce	acaba	Reforma	publicar	Centro Histórico
team	Reforma	Miguel Hidalgo	factory	centro	colonia	Miguel Hidalgo	arte	details	mytravelgram
parque_delta	Reforma	foto	semana	hrs	Centro Histórico	Benito Juárez	amigos	read	vale
vintage	feliz	Reforma	madero	independencia	teatro	palacio_belas_artes	zócalo	gracias	cuauhtemoc
happy	amo	arquitectura	Miguel Hidalgo	publicar	museo	amo	amo	zócalo	corporativo
parque	años	wtc	gracias	café	selfie	cycling	palacio	tribunal_superior	time
comer	time	gay	chesscakefmx	travel	año	ramen	burger	manager	ojo
teatro	aquí	aquí	video	happy	año	Starbucks	love	aquí	colonia
zócalo	center	Benito Juárez	Reforma	acaba	juárez	sports_world	posted	company	independencia
hair	noche	design	vida	teatro	Reforma	pet	centro	año	agua
follow me	follow me	ciudad	Cineteca nacional	gracias	friends	posted	gracias	familiar	tacos
mezcal	foto	instagram	café	vale	noche	smartfit	comida	fotografía	zócalo_capitalino
bar	mañana	street	venta	juárez	tbt	paseo	Cinépolis	tsjdf	hospital
mañana	Insurgentes	happy	monumento_a_la_revolución	fashion	bar	house	foto	sunrise	wtc
juarez	torre	años	palacio_belas_artes	colonia	hotel	vale	smartfit	posted	seguir
dogs	Polanco	love	arte	bar	Starbucks	museo	Roma	sales	ela
Reforma	Roma Norte	publicar	food	mextagram	años	Starbucks	cinemex	recomend	Cineteca nacional
fiesta	ciudad	instagood	quiero	cómo	travelgram	universidad	club	casa	secretaría
fotografía	café								

Fuente: Elaboración propia

Cuadro 4.2 (continuación): Temas latentes en el texto de tweets provenientes de las colonias gentrificadas de la ZMCM

Tema 11	Tema 12	Tema 13	Tema 14	Tema 15	Tema 16	Tema 17	Tema 18	Tema 19	Tema 20
Espectáculos visuales	Espectáculos musicales	Vida cotidiana	Diversión días laborales	Diversión diurna	Diversión de fines de semana	Lugares icónicos	Temas automatizados	Otros	Otros
Miguel Hidalgo	Cuauhtémoc	Benito Juárez	Miguel Hidalgo	Benito Juárez	Cuauhtémoc	Cuauhtémoc	Temas automatizados	Cuauhtémoc	Cuauhtémoc
mundial	love	día	semana	día	Miguel Hidalgo	Benito Juárez	repost	Benito Juárez	Benito Juárez
vida	Benito Juárez	amor	ciudad	rock	foto	foto	Cinépolis	bazar	vida
oficial	music	black	días	Cuauhtémoc	Polanco	café	gracias	centro	día
Benito Juárez	Roma	Miguel Hidalgo	Cuauhtémoc	cinemex	gracias	independencia	Miguel Hidalgo	casa	super
Cuauhtémoc	wtc	gracias	caja	gourmet	smartfit	aquí	feliz	posted	fotografía
aquí	festival	pride	Reforma	Reforma	love	día	teatro	Miguel Hidalgo	Polanco
Coyoacán	día	cinemex	teatro	Cinépolis	ciudad	cultura	vida	fotografía	museo_nacional
casa	Reforma	Reforma	amp	smartfit	sports.world	ángel	toprankrepost	foto	foto
Reforma	teatro	viernes	Benito Juárez	café	Benito Juárez	Polanco	ciudad	ciudad	posted
museo	gracias	días	amo	puro	bar	Miguel Hidalgo	Reforma	Benito Juárez	arte
gracias	amor	inktober	wtc	gracias	Condesa	Miguel Hidalgo	Reforma	mexicanitas	café
amigos	cover	rico	day	después	Coyoacán	Cineteca nacional	centro	días	Condesa
teatro	singer	foro	año	rol	amigos	amor	noche	art	Reforma
Condesa	cinemex	loveislove	hotel	casa	bnw	paseo	foto	Polanco	Starbucks
después	Cinépolis	foto	Centro Histórico	foro	año	time	bar	especiales	museo
consejo	teatro_metropolitano	juarez	Coyoacán	colonia	vale	Condesa	Polanco	Centro Histórico	feliz
Polanco	vale	mañana	lunes	paseo	café	Reforma	benito	regalos	quitero
cmi	life	noche	dios	Miguel Hidalgo	zócalo	fondo	juarez	quitero	inba
lucha libre	mercado	Starbucks	plaza_Condesa	happy	independencia	publicar	pedroprietotv	foodporn	acaba
fitness	song	casa	style	love	domingo	hotel	love	años	casa
wtc	colonia	ángel	ángel	ciudad	miyana	Ciudad de Mexico	palacio.belas.artes	travel	comer
foto	casa	fotografía	model	feliz	colonia	viernes	Cuauhtémoc	Insurgentes	independencia
Starbucks	semana	próximamente	gente	menstyle	oficial	centro	palacio.belas.artes	fitness	año
music	universidad	Cuauhtémoc	acaba	Condesa	home	universidad	cuauhtémoc	food	love
pinos	metropolitano	Rebel	selfie	rebel	amor	delegación	happy	bar	gracias
Cinépolis	model	Benito Juárez	día	palacio.belas.artes	tato	música	job	love	negro
domingo	Miguel Hidalgo	centro	Starbucks	fotografía	don	torre	delegación	arte	Centro Histórico
dr	familia	año	mexicana	Coyoacán	sol	cinemex	recomend	club	noche
arte	art	parque_delta	palacio	museo	metro	monumento	wallace	día	cervecería

Fuente: Elaboración propia

El cuadro 4.2 muestra las palabras con la mayor probabilidad de pertenecer a cada uno de los 20 clústers conformados por el modelo de ALD. Sin embargo, el modelo no asigna automáticamente el nombramiento de cada clúster. Al respecto, se han enfrentado varios aspectos para poder designar los nombres a los campos semánticos. En este sentido, dicho etiquetado ha seguido la intención de dar cuenta de actividades, lugares o palabras relacionadas al proceso de gentrificación en la ciudad. Pero, como era de esperarse, no todos los temas identificados refieren a dicho proceso. Esto debido a la naturaleza del algoritmo de AM, al cual pertenece la ALD. Ya que, al ser un algoritmo no supervisado, éste abre la posibilidad de obtener otras temáticas que no fueron previamente previstas. En este sentido, el análisis realizado en este trabajo se guía plenamente por la naturaleza de los datos, sin embargo, trata de dar una interpretación teórica-conceptual de los resultados basado en los marcos analíticos del poder de la comunicación de RGAD que puede constituir una gentrificación digitalmente aumentada y, del proceso de gentrificación como síntoma lingüístico que subyace a los signos encontrados en el modelo.

Así, el primer paso para nombrar los temas latentes ha sido identificar los campos semánticos que ofrecen un conjunto de palabras que refieren, sin ambigüedades, a prácticas determinadas. Este ha sido el caso de los temas del 1 al 9. Por ejemplo, el tema 1 se ha nombrado cultura hipster debido a que son palabras que son identificadas a prácticas de dicho grupo de población, tales como *vintage*, *mezcal* o *bar*. Por su parte, el conjunto de palabras del tema 9 ha sido un hallazgo puesto que hace referencia al Tribunal Superior de Justicia de la Ciudad de México. Espacio ubicado en las colonias que ha experimentado renovación urbana pero donde se encuentran oficinas de dicha entidad gubernamental.

El siguiente grupo de temas ha seguido una estrategia diferente para nombrarlo. Este grupo de temas comprende del tema 10 al 16. En este caso, los campos semánticos no se identifican de manera sencilla, debido a que a lo largo de estos grupos se tienen algunas palabras que se repiten entre ellos. Sin embargo, también se identifican palabras claves que diferencian cada tema latente. En consecuencia, algunos temas latentes de este segundo grupo han sido nombrados como referentes a la vida cotidiana o diversión, pero se han diferenciado por sus menciones a la temporalidad en la que se realizan. Ejemplo de lo anterior se encuentran en los temas de diversión en días laborales, diversión diurna o diversión en fin de semana.

Por último, los grupos semánticos finales no han mostrado un patrón claro para definirlos o

bien hacen referencia a lugares sin alguna actividad explícita. Así, el tema 17 ha sido nombrado como ‘lugares icónicos’ y los temas 19 y 20 se han nombrado como ‘otros’. En lo que respecta al tema 18 se ha nombrado como ‘temas automatizados’ debido a que su palabra más probable es *repost* o *toprankrepost* las cuales refieren a publicaciones con el objetivo de volverlas más visibles. Pero, en este estudio no se alcanzan a analizar a profundidad.

En el conjunto de los temas llama la atención la ausencia de palabras o conformación de temas latentes que refieran a aspectos políticos. Esto puede matizar los estudios que utilizan Twitter como nueva fuente de información para analizar temas políticos bajo la premisa de que en ésta se presentan discusiones sobre la vida pública y política de la sociedad. Dicha idea se generó con los movimientos sociales de la Primavera Árabe y de *Occupy Wall Street* (Castells, 2012). En México, particularmente se han tomado las redes sociodigitales como expresión de la esfera pública (Reguillo, 2019). Sin embargo, los resultados de este análisis apuntan a que en el conjunto de los temas de interés en Twitter no se ubican los tópicos políticos. Esto indicaría que, a nivel político, Twitter no es un espacio para la discusión política abierta, por el contrario, ésta, como casi todas las redes sociodigitales tienen un alto grado de homofilia (Bernie Hogan, 2018; Uzzi, 2010). Es decir, las redes sociodigitales se conforman por redes de individuos semejantes entre ellos, reproduciendo, así, su propias ideas y no abriendo el espectro a otras temáticas ajenas a las de su interés.

Los 20 temas latentes denotan agrupamiento entre ellos. El primer gran grupo refiere a campos semánticos referentes con el proceso de gentrificación en sus diferentes acepciones. El tema 1 se demarca por palabras referentes a colonias gentrificadas y a las actividades que se realizan en ellas, este es el caso de la colonia Condesa o Roma. Dentro de éstas se muestra el consumo en bares durante los fines de semana. En esta agrupación semántica se tienen palabras en inglés tales como *vintage*, y *dogs*. Al mismo tiempo, se tiene la recuperación de estéticas de décadas anteriores es una característica de lo jóvenes identificados como *hipster* (CNN, 2014). En esta misma tendencia se encuentra la palabra *mezcal* y que ha sido recuperada como signo de distinción (Martineau, 2016).

Es grupo daría indicios de una gentrificación aumentada digitalmente por medio del poder del lenguaje. Por un lado, desde las aproximaciones previas de la sociolingüística, se ha encontrado que en espacios plenamente identificados como gentrificados o en proceso de gentrificación se asocian palabras como *hispter*, la recuperación de palabras que aluden a tradiciones o consumos tradicio-

nales o la coexistencia de paisajes lingüísticos de dos o más idiomas (Lyons y Rodríguez-Ordóñez, 2017; Papien, 2012). En este sentido, los hallazgos de este grupo semántico son consistentes con la literatura previa sobre la sociolingüística de la gentrificación, ya que se han identificado palabras en idioma inglés, o bien, relacionadas con consumos de bebidas tradicionales, como el mezcal. Con esto en mente, este tipo de palabras se convierten en signos particulares del proceso espacial del cual emanan, el proceso de gentrificación. Así, desde la aproximación de la teoría del lenguaje se pueden interpretar estos signos en relación al síntoma social de la gentrificación. Aunado a esto, no se debe olvidar que dichos signos han sido encontrados en un medio de comunicación particular, Twitter, el cual tiene características particulares y un papel determinado en el proceder de las personas en la dimensión digital. En consecuencia, se podría pensar en que el proceso de gentrificación, se refleja mediante el lenguaje utilizado en Twitter. Pero no solo lo refleja, sino que en Twitter mismo se construye una espacialidad digital particular, la cual apunta a ciertos signos y símbolos vinculados con el imaginario de la gentrificación.

El segundo grupo semántico puede ser interpretado como sentimientos y celebración. En esta se encuentran palabras que se han mencionado anteriormente, tales como celebración, feliz, amor, cumpleaños, gracias, etc. Este tema, en particular, ha sido el más reportado por la literatura que ha utilizado técnicas de análisis de texto para realizar análisis de sentimiento mediante Twitter. De igual manera, este tema también ha sido identificado en la ciudad de Londres por Lansley y Longley (2016). El tema 3 refiere a signos de la cultura *gay*. La combinación de palabras en este tema permiten proponer actividades vinculadas con esta comunidad. Tal es la palabra *design*, y sus publicaciones en otra plataforma digital, *Instagram*, la cual se enfoca en contenido audiovisual. Otro tema muy cercano a éste es el tema 5, el cual es catalogado como “clase creativa”. Las palabras clave en este tópico son *mextagram*, página particular de *Instagram* que se enfoca a la presentación de ideas audiovisuales originales en México, *fashion*, palabra relacionada una actividad identificada con dicha clase creativa (Mercado-Celis, 2016) y por último la palabra *café*, ya se ha estudiado que las cafeterías son lugares de distinción para las personas con profesiones vinculadas a la creatividad y la innovación (Yodanis, 2006).

Por su parte, los temas 4 y 8 hacen referencia a consumo específico en espacios públicos. El tema 4 muestra el consumo de espacio públicos con vocación comercial. Ejemplos de esto son El Zócalo, Centro Histórico o lugares icónicos de la ciudad, tales como El Palacio de Bellas Artes

o el Monumento a la Revolución. Este tema se ha nombrado como espacio público de consumo debido a que muestran palabras referentes a centros comerciales cerrados, como el nombrado Parque Delta. Otro caso semejante es a la palabra Madero, que hace referencia la calle peatonal del centro histórico que tiene vocación netamente comercial de élite. En cuanto al tema 8, éste se demarca semánticamente a la comida y bares. Dichas actividades vinculadas al esparcimiento en la Ciudad o bien expresiones de procesos de renovación a decir de Zukin y cols. (2017). Al mismo tiempo, esto concuerda con los estudios sobre sociolingüística de la gentrificación analizados líneas más arriba en este capítulo, los cuales vinculan un conjunto de estrategias lingüísticas al proceso de gentrificación comercial. En este punto se puede enfatizar, no solo la producción del proceso de gentrificación digital por medio de las personas, sino también por medio de estrategias publicitarias de determinados comercios.

Existen otros conjuntos de campos semánticos que dan cuenta de la vida cotidiana de las personas, así se puede observar en los temas 9 y 10. Éstos aluden a lugares de trabajo y prácticas dentro de dichos lugares. En particular, el tema 9 hace referencia la llamada Ciudad Judicial, espacio donde se concentran tribunales y oficinas de impartición de justicia en la colonia Doctores. En el campo semántico 10 se pueden encontrar palabras como corporativo, Secretaría, *Word Trade Center* (WTC). Este tema en particular lo reafirma la palabra día, debido a que son actividades que se realizan en horarios diurnos. Los temas 13 al 16 también se catalogan en las actividades de la vida cotidiana. Sin embargo, éstos refieren claramente a actividades de diversión con la especificidad temporal, ya sea en días laborales, durante la noche o fines de semana. En este punto se puede postular que los usuarios enfatizan la hora del día de sus actividades cotidianas. En otras palabras, el tiempo es parte fundamental en la construcción de espacialidades dentro de Twitter en las colonias gentrificadas.

Mención especial merecen los temas 17 y 18, lugares icónicos de la ciudad y temas automatizados. En el caso del tema de lugares icónicos se refiere a la presencia explícita de los lugares nombrados más importantes de la ciudad. Mientras tanto, el tema 18 se ha denominado temas automatizados por diversas razones. La primera, se han encontrado dos cuentas que hacen referencia a la construcción del tendencias por medio del código de Twitter. Este es el caso de *toprankpost* el cual es una etiqueta para consolidar y denotar las principales tendencias temporales en la plataforma. El segundo ejemplo de la automatización dentro de este campo semántico es la presencia

de una cuenta de un personaje público, *pedroprietotv*, ya que es la cuenta oficial de un actor de televisión y potencial usuario con influencia en el mundo de Twitter.

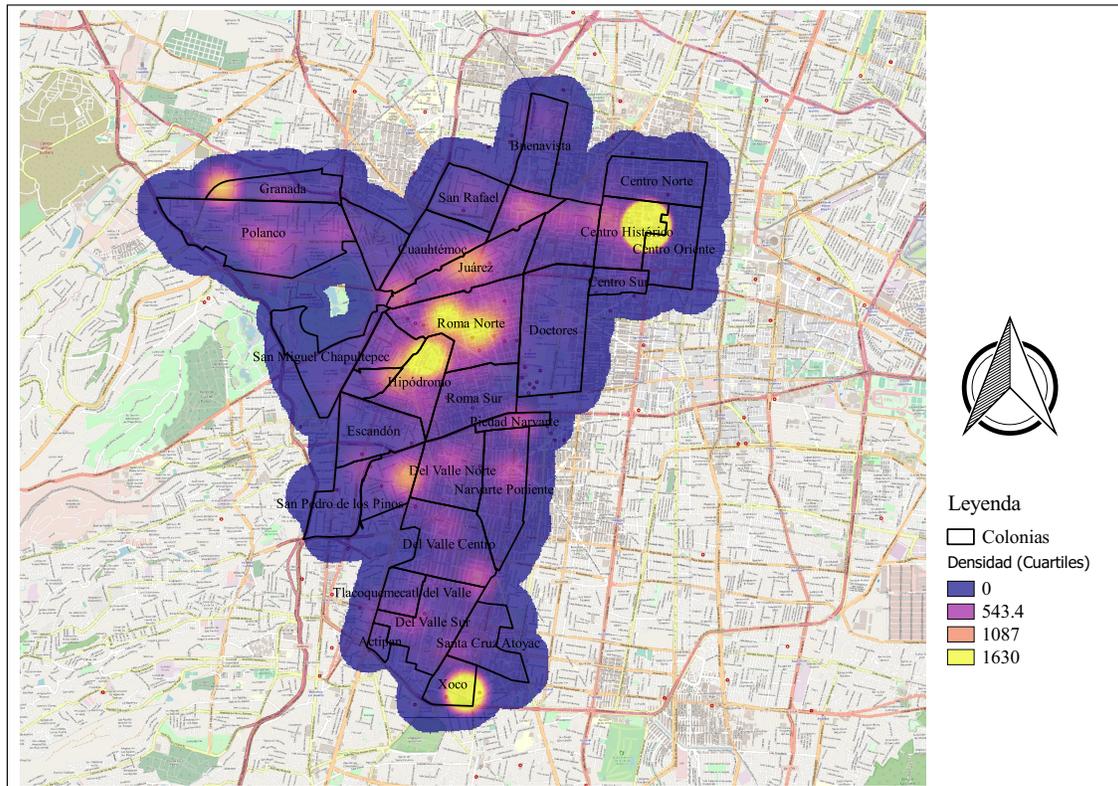
Los temas 11 y 12 refieren a temas sobre espectáculos, en particular sobre espectáculos visuales y musicales. En ambos temas, se observa la presencia de entidades nombradas donde se llevan a cabo dichas actividades. Ejemplo de esto es la palabra “teatro”, o “lucha libre” en el caso del tema 11. Por su parte, sobresale la entidad nombrada de “Teatro Metropolitano” en el tema 12 que se vincula con espectáculos musicales. Por último, se tienen los temas 19 y 20 que se han catalogado como otros debido a que no se ha encontrado un campo semántico que contenga a las palabras más probables de dichos agrupamientos.

Se ha mencionado en el apartado metodológico de este capítulo que la construcción de la ALD posibilita que una misma palabra pertenezca a más de un tema debido a la combinación de distintas palabras que, en conjunto, conforman temas homogéneos semánticamente hablando al interior y heterogéneos entre ellos. Esto permite identificar las palabras que aparecen en más de un campo semántico y observar como éstas varían su significado entre cada grupo. En específico se resaltan dos palabras Starbucks y Parque Delta. La palabra Starbucks aparece en 7 de los 20 temas formados. Su aparición se vincula a temas como la cultura hipster, diversión en días laborales, celebraciones, ejercicio o en general la vida cotidiana. Lo anterior se puede explicar por el hecho de que la marca de Starbucks se ha convertido en sinónimo de elitización del espacio y de distinción para los consumidores. Así lo demuestran Herman y Hadden (2019) quienes han encontrado que existe una asociación positiva entre la presencia de un establecimiento de Starbucks y mayores precios de la vivienda. En el caso de “Parque Delta” se refiere a un gran centro comercial en la colonia Narvarte. Se ha considerado importante para este análisis dicha entidad nombrada porque es ícono de espacios públicos jurídicamente privados. Este tipo de espacios a decir de Duhau y Giglia (2008) son espacios que generan segregación y segmentación social debido a su característica de usar dichos espacios de manera controlada y bajo las reglas de dichos establecimientos.

Cada uno de los campos semánticos latentes dan cuenta de diferentes marcos referenciales en función del conjunto de sus signos o palabras. Es decir, el conjunto de palabras de cada grupo es síntoma de lo que ocurre en las colonias gentrificadas de la ZMCM. Para propósitos de este trabajo resaltan algunos grupos más que otros, debido a que se vinculan más con el proceso de gentrificación y elitización de las colonias consideradas. En este sentido, los temas asociados a la

cultura hipster, espacio público de consumo, clase creativa, comida y bebida o, cultura gay son de particular interés para mostrar la manera en que ciertas prácticas del espacio se reflejan geográficamente para construir realidad geográfica aumentada. De esta manera, a continuación, se presentan los núcleos de densidad de los temas más importantes.

Figura 4.9: Núcleos de densidad de emisiones del tema latente de “Cultura hipster”

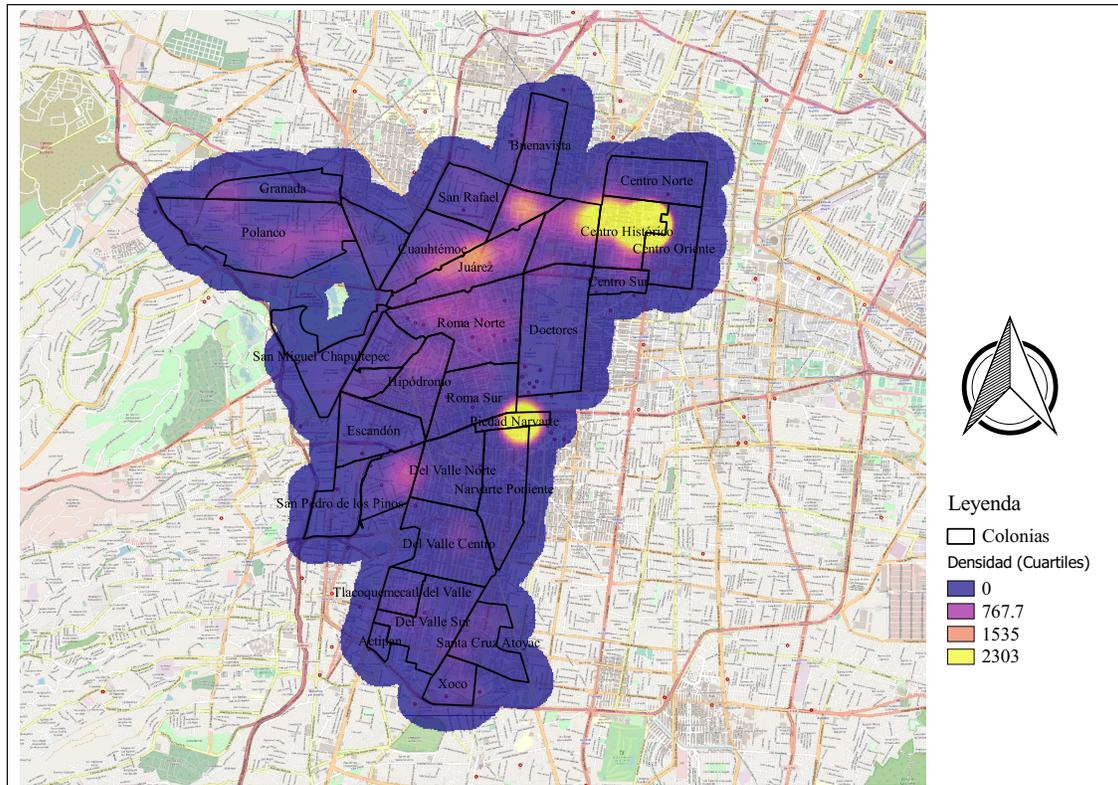


Fuente: Elaboración propia

En la figura 4.9 se aprecian los núcleos de emisión de publicaciones catalogadas como “cultura hipster”. Como se observa en el mapa, este tipo de publicaciones se encuentra bien identificada en 5 puntos de las colonias gentrificadas, colonia Centro Histórico, Juárez, Roma Norte, Hipódromo, Granada y Xoco. En particular, se ha mencionado que las colonias Juárez, Roma Norte e Hipódromo son las principales colonias identificadas con perfil gentrificador. Al sur de las colonias gentrificadas se encuentra la colonia Xoco, en dicha colonia se ubica la Cineteca Nacional. Esta entidad nombrada fue de importancia relativa al analizar la emisión de palabras por horario. Otro punto de alta densidad de emisión del campo semántico es la Plaza Antara, ubicada al poniente de las colonias gentrificadas y contigua a la colonia Polanco. Al mismo tiempo, se observa densidad media a lo largo de todas las colonias. Este tema, como se desarrolla a continuación es el más

extendido a lo largo de las colonias gentrificadas.

Figura 4.10: Núcleos de densidad de emisiones del tema latente de “Espacio público de consumo”

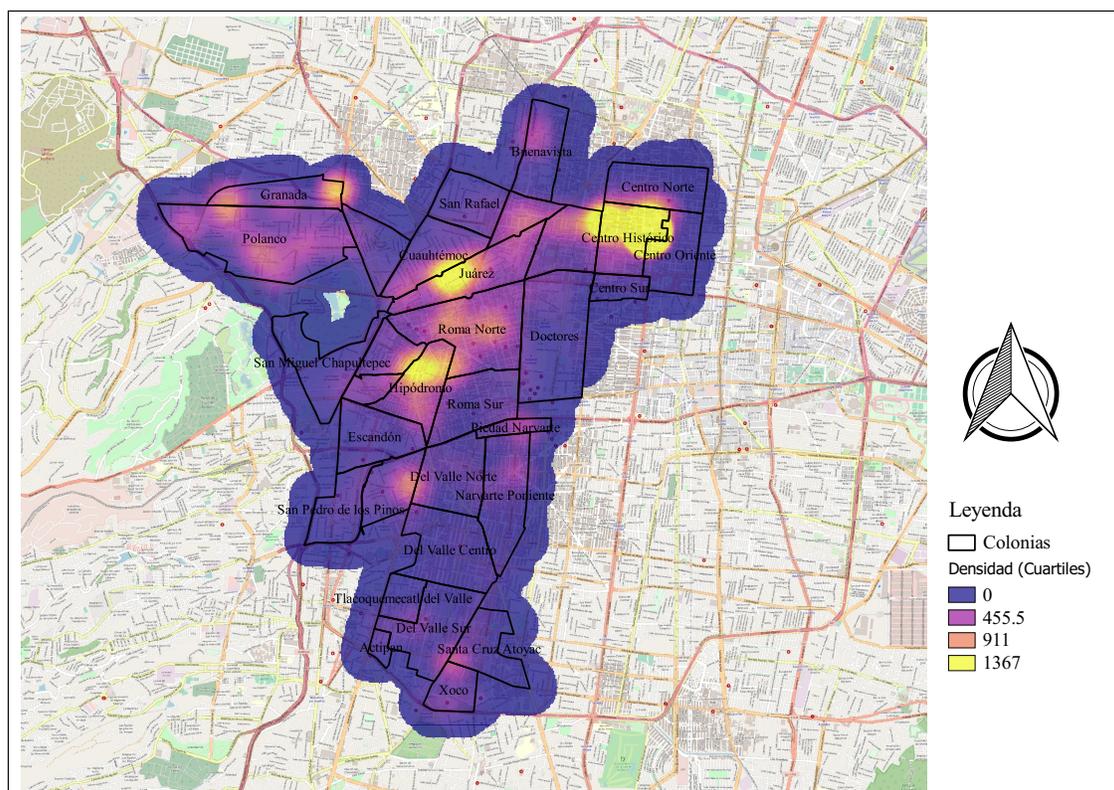


Fuente: Elaboración propia

En el mapa 4.10 se presentan los principales puntos de densidad bajo la temática de espacio público de consumo. En este mapa sobresalen cuatro puntos particulares. El primero de ellos ubicado en el Centro Histórico, dicho núcleo de densidad se extiende desde el zócalo de la ciudad hasta la Alameda Central pasando por la Calle de Madero. Dicho corredor refuerza la idea de un espacio público abierto dirigido al consumo. Al sur del Centro Histórico se observa el punto de Parque Delta en la colonia Piedad Narvarte. El tercer punto a mencionar es el ubicado sobre la Avenida Paseo de la Reforma entre las colonias Juárez y Cuauhtémoc. En este núcleo de emisión se puede ubicar centros comerciales cerrados como el Reforma 222. Enseguida, el cuarto punto de densidad es el ubicado en la colonia Tabacalera, en este punto se encuentra el Monumento a la Revolución.

En el mapa anterior se presentan los principales núcleos de emisión vinculados con el tema de “Comida y bebida”. En general, se observa que las emisiones que latentemente refieren a este tema se concentran en la colonia Centro, Juárez, Hipódromo y en menor medida en la colonia Del Valle Norte, Santa Cruz Atoyac, Polanco y de nueva cuenta en el punto de Parque Delta. Es-

Figura 4.11: Núcleos de densidad de emisiones del tema latente de “Comida y bebida”



Fuente: Elaboración propia

tos hallazgos refuerzan la idea de Duhau y Giglia (2016) sobre la acción de *salir a comer* como acción fundamental de la cotidianidad en la ZMCM. Los autores encuentran que dicha actividad se encuentra estratificada socialmente, ya que para estratos socioeconómicos altos el salir a comer implica una actividad de diversión y esparcimiento, sobre todo en fines de semana. En este caso, se puede observar que las publicaciones que refieren a esta acción se vinculan con espacios claramente comercializados. Es decir, el mapa permite proponer que la acción de comer en colonias con proceso de valorización se vincula con espacios, ya sea públicos o privados, dirigidos al consumo.

La combinación entre los núcleos de emisión de los temas latentes y la distribución desigual de las palabras se puede sugerir que el poder de la comunicación se dirige a la concentración de ciertas prácticas en espacios centrales, que a su vez valorizan espacios públicos dirigidos al consumo. En este sentido, el poder de la comunicación digital fortalece la valorización/revalorización del espacio urbano por medio de las publicaciones en redes sociodigitales como Twitter. En este nivel interactúan múltiples fuerzas, tanto económicas, culturales, políticas y, se incluyen las digitales. Así, por medio de la facilidad de emitir un tweet, el usuario no emite palabras aleatorias, éstas

responden a un contexto social, ideológico y espacial, tal es el caso de espacios gentrificados. De manera particular, se ha apuntado en los hallazgos de este apartado que puede ser interpretado el valor simbólico que las palabras tienen en función del contexto de la gentrificación, considerando a éste como un síntoma, en el sentido lingüístico.

Los resultados de esta investigación apuntan a que el proceso de gentrificación aumentado digitalmente en la ZMCM particularidades semánticas, dado el contexto concreto de la ciudad, aunque comparte similitudes con los hallazgos encontrados en ciudades europeas, específicamente Londres. En primera instancia, temas de consumo elitizados son claramente vectores de segmentación del espacio digitalmente aumentado que se encuentran en ambas ciudades. En este rubro se pueden contemplar los temas latentes denominados como espacios públicos de consumo, cultura hipster o bien clase creativa. Por otro lado, los temas particulares que refieren a este proceso en la ZMCM se pueden ubicar los temas referentes a asuntos laborales, en concreto a los corporativos y oficinas y a los tribunales de justicia. La identificación de temas latentes por medio de la técnica de ALD ha permitido profundizar en un marco interpretativo de las palabras de Twitter desde el punto de vista de su función referencial, dado que éstas nombran cosas, personas o entes. Al mismo tiempo se ha recuperado la noción de que estas palabras se encuentran enmarcadas bajo un síntoma, en este caso se ha propuesto que el proceso de gentrificación es un síntoma que puede dar explicación a ciertas palabras emitidas en la red sociodigital.

#### **4.6 Consideraciones finales**

El principal objetivo de este capítulo ha sido mostrar cómo el poder de la comunicación mediante Twitter aumenta digitalmente el proceso de gentrificación en la Ciudad de México mediante símbolos y signos relacionados con dicho proceso. Para ello, el capítulo ha retomado la teoría de la función del lenguaje, en la cual todas las palabras hacen referencia a un hecho de la realidad, sin alguna carga de valor. La segunda función retomada en este estudio refiere al síntoma del lenguaje, así las palabras enmarcadas en un contexto dan síntoma de un lugar, cultura o valor del emisor. Así, el poder de la comunicación apunta a que estos signos y símbolos en Twitter son reflejo del síntoma del proceso de gentrificación en la ciudad.

En términos generales, la aproximación a la gentrificación desde su vertiente de realidad aumentada aporta al análisis realizado previamente sobre el fenómeno y su relación con el mundo

digital. Al respecto, el uso de grandes datos permite abonar sobre el estudio de la gentrificación desde su dimensión simbólica y de creación de identidades. Esto se ha desarrollado por medio del análisis de las palabras y de temas latentes en el texto de los tweets. En este sentido, se ha visto que Twitter es un medio por el cual las personas reflejan su afinidad a determinados patrones de consumo y actividades en las colonias gentrificadas de la Ciudad de México.

Esto se refuerza al haber encontrado que los patrones léxicos de colonias no gentrificadas difieren del patrón identificado en colonias identificadas como gentrificadas. Dado este objetivo el principal hallazgo del apartado es mostrar que el lenguaje emitido en Twitter es reflejo del proceso de gentrificación, en particular por medio del uso de un lenguaje que implica elitización, se puede proponer que, para el caso del proceso de gentrificación en la ZMCM se tienen símbolos concretos como entidades nombradas y prácticas de consumo. En particular, esta elitización del lenguaje de Twitter se vincula con la dinámica laboral realizada en colonias gentrificadas de la ciudad, esto se da por medio del uso de expresiones en inglés, las cuales son usadas, en términos lingüísticos, para reflejar un mayor estatus.

La depuración del texto en términos técnicos y conceptuales es un reto metodológico. El primer reto se enmarca en la estandarización del lenguaje encontrado en Twitter para poder implementar técnicas del campo del procesamiento del lenguaje natural. Esto se debe a que el último, en su mayoría, se enfoca al análisis de texto estructurado, ya sea de libros, ensayos, artículos de periódicos, etc. En este punto, el análisis de Twitter se podría mejorar al establecer mejores técnicas de pre-procesamiento del texto. Otro punto a resaltar es la identificación de entidades nombradas. El presente trabajo se dio a la tarea de identificar las entidades nombradas de tipo espacial, lo que permitió encontrar patrones de actividades de manera más clara. En este sentido, el presente análisis aporta un elemento sustancial a la literatura reportada, debido a que en ninguno de los trabajos previos que han utilizado el análisis de texto para temas urbanos indican haber identificado dichas entidades nombradas para sus estudios. Este ejercicio ha posibilitado encontrar patrones más claros para estudiar el fenómeno de gentrificación, ya que muchas de las actividades identificadas con este fenómeno se encuentran ancladas a los lugares reconocidos comúnmente por sus nombres. De manera paralela, la depuración sistemática del texto en los tweets ha hecho que se pierda información relevante en un sentido cualitativo de las prácticas. Al respecto, haber eliminado las etiquetas o *hashtags* ha implicado que se pierdan los vínculos que las personas manifiestan hacia sus activi-

dades en el espacio. Del mismo modo, la depuración de los hipervínculos a otras plataformas ha implicado omitir cómo se relacionan las plataformas digitales entre ellas. Un ejemplo de esto se manifiesta en que muchos de los tweets se vinculan con plataformas como Yelp o Instagram. Por último, la eliminación de palabras vacías en español implica un reto constante para este tipo de análisis, debido a que la mayoría de avances sobre la temática se han hecho para el idioma inglés, lo que significa un sesgo al momento de interpretar los resultados.

Por su parte, el modelo de Asignación Latente de Dirichlet ha permitido una primera aproximación a los temas semánticamente latentes en Twitter en las colonias gentrificadas. Dicho modelo ha reafirmado la hipótesis de una gentrificación aumentada digitalmente por medio de la plataforma digital estudiada, esto se ha hecho en dos niveles diferentes. El primero responde a la reafirmación de las palabras identificadas como más frecuentes en términos aislados y su relación con otras palabras semánticamente vinculadas que no aparecían en el primer análisis. En segundo lugar, por medio del modelo semántico se han construido campos semánticos que no fue posible identificar bajo el análisis de frecuencias. Este es el caso del tema vinculado con el Tribunal Superior de Justicia o el tema relacionado con turismo y viajes en la ciudad. De esta manera, el modelo se presenta como una herramienta útil para hallar temas que bajo un análisis de frecuencias de palabras se ocultan ante la preeminencia de las palabras más usadas.

Otro hallazgo importante del análisis léxico de Twitter en su expresión geográfica es la representación cartográfica de los temas latentes. En este sentido, por medio de esta representación se puede observar que existe una construcción del espacio urbano por medio de elementos lingüísticos. En este terreno, el enfoque de grandes datos permite corroborar empíricamente proposiciones teóricas que se han realizado en el campo de la sociolingüística o desde la geografía digital antropológica y radical. En concreto, aunque se han identificado colonias gentrificadas en función de su centralidad de servicios e intervención estatal para renovar los espacios, se muestra, por medio del análisis lingüístico, que dentro de estos espacios existe heterogeneidad en las prácticas. De este manera, las colonias identificadas tradicionalmente como gentrificadas son las que presentan mayor cantidad la concentración de tweets que reflejan el proceso de gentrificación.

Tanto el análisis de frecuencias como los temas latentes arrojan la presencia de marcas comerciales y espacios de consumo públicos de índole privado. Específicamente, la aparición de marcas de restaurantes y cafeterías con presencia internacional es un patrón recurrente en las menciones

de Twitter. Dicho patrón abre la puerta a una discusión mucho más amplia de la conformación del espacio urbano; por un lado, cómo la dinámica de empresas como Starbucks aprovechan la dimensión digital para conformar espacios elitizados en la ciudad; por el otro, las motivaciones objetivas y subjetivas para reportar voluntariamente y de manera geográficamente precisa que hacen uso y consumen en dichos espacios. La segunda discusión tiene que ver con la mención de prácticas de consumo en centros comerciales. Este hecho apunta a indagar sobre la socialización en espacios públicos controlados que impiden la creación de ciudadanía en la ciudad. En otras palabras, las menciones en Twitter indican que se construye un espacio urbano despolitizado y aparentemente neutro. Lo anterior puede ser explicado por el hecho del uso mayoritario de personas de estratos altos de dicha red sociodigital.

Finalmente, los hallazgos de este capítulo apuntan a la exploración de cómo la comunicación por medio de Twitter se relaciona con espacios socialmente diferenciados. Para ello, se ha hecho un esfuerzo para entender desde diferentes aproximaciones cómo el poder de la comunicación puede aumentar digitalmente la gentrificación. Así, los resultados mostrados son consistentes con estudios sociolingüísticos sobre la relación entre gentrificación y lenguaje. Sin embargo, la diferencia es que este paisaje se construye de manera digital, sustentado en prácticas físicas en espacios gentrificados. Por otro lado, el capítulo abona a la propuesta de la realidad geográfica aumentada digitalmente propuesta por M. Graham, Zook, y Boulton (2013a) al dar un sustento lingüístico a la conformación del poder de la comunicación como motor constructor de realidades digitales en el espacio. En este sentido, el capítulo también abona a los trabajos que han utilizado el algoritmo de ALD para identificar tópicos en el espacio urbano, al otorgarle un marco interpretativo desde una perspectiva de la naturaleza epistemológica de la fuente de datos, los tweets.

## Capítulo 5

### Reflexiones finales

En la presente investigación se han explorado diversos mecanismos que intervienen en la generación de una gentrificación aumentada digitalmente por medio de Twitter en la Ciudad de México. Con esto se ha aproximado una respuesta a la pregunta general que ha guiado este trabajo. Para lograr esto, se han desarrollado cuatro capítulos que han documentado, ampliamente, como se ha abordado, tanto teórica como empíricamente, la construcción de geografías digitales y las potencialidades y limitaciones del uso de grandes datos en los estudios urbanos. En este sentido, se han expuesto un conjunto de propuestas conceptuales y empíricas que ayudan a cerrar la brecha en el conocimiento sobre el papel de la sociedad digital y los grandes datos en entornos urbanos, dimensiones que son emergentes actualmente.

Así, el objetivo de las consideraciones finales es ofrecer una visión amplia y crítica de las diferentes dimensiones que se han expuesto en la investigación y, que apoyan el argumento de una gentrificación aumentada digitalmente en la Ciudad de México. Con esto en mente, el presente capítulo se estructura en seis partes: principales hallazgos y propuestas teóricas expuestas a lo largo de la tesis en torno a la gentrificación aumentada digitalmente; principales hallazgos en el conocimiento sobre la desigualdad digital y el perfil sociodemográfico en el uso de redes sociodigitales; discusión del vínculo entre la emisión de Twitter con temporalidad y espacialidad de las actividades en la Ciudad de México; la propuesta interdisciplinaria del análisis sociolingüístico del corpus de los tweets emanados de colonias gentrificadas; los grandes datos en estudios urbanos; por último, las principales limitaciones y temas poco desarrollados, así como una agenda de investigación a futuro.

#### **5.1 Principales hallazgos y propuestas teóricas sobre la gentrificación aumentada digital-**

## mente

La preocupación de estudiar la dimensión digital de la gentrificación ha surgido de la curiosidad de explorar el ensamblaje de dos grandes fenómenos que se observan cotidianamente en las grandes ciudades a nivel mundial. Por un lado, el proceso de gentrificación, en sí mismo, con sus particularidades contemporáneas. Por el otro, el auge y la omnipresencia de elementos de la revolución digital en la sociedad. Elementos tan cotidianos son el acceso a Internet, el uso de dispositivos móviles e inteligentes, el uso de las redes sociodigitales o, para el caso de las ciudades, se puede hablar de un complejo conjunto de elementos, que van desde fibra óptica hasta videocámaras de vigilancia. En este sentido, la conjunción entre los elementos digitales y el fenómeno de gentrificación requiere de marcos conceptuales particulares que consideren la naturaleza del fenómeno urbano y la naturaleza de las geografías digitales y su interacción. Así, se puede avanzar hacia un mejor entendimiento de una realidad que no se encuentra dividida entre lo digital y lo no digital, sino que es un *continuum* que incesantemente se experimenta (Floridi, 2014).

El fenómeno de gentrificación ha sido abordado desde distintas perspectivas a lo largo de las últimas décadas. Esto ha sido debido a la relevancia que tal proceso implica en el desarrollo de las ciudades en el mundo contemporáneo y por la multiplicidad de dimensiones que se involucran en él. Estas perspectivas han apuntado a los factores políticos (Janoschka, 2016; Slater, 2009), económicos (Hackworth y Smith, 2001; N. Smith, 1979, 1990, 2002), sociales (Gómez-González, 2016; Zukin, 1991), gubernamentales (di Virgilio y Guevara, 2015), comerciales (Batres-Guadarrama, 2016), y sociolingüísticos (Lyons y Rodríguez-Ordóñez, 2017; Papen, 2012). La multiplicidad de perspectivas es un indicador de la complejidad que involucra el estudio de la gentrificación a nivel mundial. Sobre todo cuando estrategias gentrificadoras han proliferado tanto en el norte como en el sur global, en economías de oriente y occidente. Bajo este marco, el presente trabajo de investigación se inserta en una prolija discusión sobre la gentrificación, sus determinantes y consecuencias asociadas.

La gentrificación es un fenómeno que estructura las ciudades de manera desigual, ya sea por mecanismos políticos, violentos o económicos. Así, se han aproximado los factores de la dimensión digital que intervienen en la gentrificación como fenómeno que configura las urbes bajo lógicas de desigualdad. Recientemente la discusión internacional sobre la gentrificación ha comenzado a preocuparse por la relación de las tecnologías de la información, el auge de la sociedad digital y las

plataformas de Internet como nuevos vectores dentro del fenómeno.

En este sentido, Wyly (2018) se ha cuestionado cómo la gentrificación está cambiando ante el auge de la economía digital. A decir del autor, la revisión de N. Smith (1982) es indispensable para entender cómo se interrelacionan los elementos económicos y culturales, ampliamente estudiados sobre el fenómeno, con las nuevas características del auge de empresas digitales. En este mismo hilo, Zukin y cols. (2017); Zukin (2020) han explorado, desde una postura culturalista sobre la gentrificación, la manifestación en blogs y páginas de reseñas en la web de símbolos e ideologías vinculadas con la gentrificación y la renovación de espacio urbanos. Aunque ambos autores rescatan que los elementos digitales han comenzado a formar parte de la explicación de la gentrificación contemporánea, sus aproximaciones teóricas no recuperan la ontología propia de la dimensión digital, sino que, la incluyen en sus análisis como un factor que solo refuerza los mecanismo ya estudiados, sin proponer que se puedan generar sinergias entre ambas dimensiones.

Este trabajo de investigación toma distancia de estas posturas, ya que, propone una aproximación teórica desde la corriente crítica de la construcción de las espacialidades digitales. El desarrollo de las espacialidades digitales es una corriente de reciente desarrollo que propone interpretar la conformación de diferentes espacios en función de la diversidad de naturalezas que conforman la revolución digital en la sociedad. Por mencionar un ejemplo, se puede señalar que la espacialidad generada por las redes sociodigitales es sustancialmente diferente a la espacialidad producida gracias a la automatización del pago de un transporte público. Así, las espacialidades digitales dialogan con la producción constante del espacio en términos de la producción social del espacio propuesta por Lefebvre (1991).

Dentro del cuerpo teórico de las espacialidades digitales se ha recuperado la propuesta de la Realidad Geográfica Aumentada Digitalmente (RGAD). Dicha aproximación se ha elegido por su diálogo a nivel teórico con los desarrollos previos sobre el fenómeno de gentrificación, desde un punto de vista de la constitución desigual del espacio urbano. Al mismo tiempo, se ha optado por este enfoque porque las aproximaciones empíricas que se han llevado a cabo han sido realizadas por medio del análisis de información de sitios de redes sociodigitales tales como Facebook o Twitter.

Para la propuesta de la RGAD la constitución de una espacialidad digital se sustenta en un conjunto de poderes que generan de manera desigual una realidad aumentada por medio de mecanismos digitales. En primera instancia, el poder de la distribución digital apunta a la desigualdad en

el acceso a Internet. En la mayoría de los textos que han versado sobre la RGAD se ha analizado la desigualdad en el acceso entre el norte y sur global y se ha enfatizado el vínculo entre el acceso y el desarrollo económico y político de los países (M. Graham, Zook, y Boulton, 2013a; M. Graham, 2014, 2017). Sin embargo, poco se ha analizado dicha desigualdad desde factores sociodemográficos.

En segundo lugar, el poder de la comunicación en la dimensión digital. Al respecto, los estudios empíricos de la propuesta han analizado cómo se construyen espacialidades en el mundo digital gracias a la presencia de lugares nombrados y que estos lugares nombrados pueden ser visibles en búsquedas en Google gracias a factores políticos o económicos (Crampton y cols., 2013; M. Graham y Zook, 2013; M. Graham, Zook, y Boulton, 2013a; M. Graham y cols., 2015; M. Graham y Sabbata, 2015; Shelton y cols., 2015). Aunque estos estudios han puesto en la palestra la importancia diferenciada que tienen los usuarios al emitir mensajes en la dimensión digital, poco se ha avanzado en una interpretación sociolingüística de esta dimensión.

En tercer lugar, el poder del código produce realidades digitales diferenciadas. En este punto se ha avanzado en conocer el papel fundamental que el código subyacente a las plataformas digitales ejerce en la conformación de realidades. Al respecto se ha hablado de los intereses económicos y políticos que Google tiene para colocar en primer lugar de sus búsquedas determinados resultados o la capacidad de *bots* para construir tendencias en Twitter (Crampton y cols., 2013; Shaw y Graham, 2017).

En este trabajo de investigación se ha avanzado en someter estas dimensiones, basadas en una distribución desigual de poderes, a la interacción con los factores de desigualdad que constituye el proceso de gentrificación desde una perspectiva estructural. En este sentido, se han analizado los factores de la dimensión de acceso desigual en la generación de tendencias de igualación y valorización de espacios gentrificados, poniendo énfasis en las perspectivas comerciales y simbólicas de la gentrificación. Así se ha explorado cómo la desigual distribución en el acceso a Internet y Twitter se vincula a personas con perfiles sociodemográficos específicos.

De igual manera, se ha indagado en cómo los mensajes emitidos en Twitter, como reflejo del poder de la comunicación, construyen paisajes lingüísticos que apuntan a la valorización de espacios gentrificados por medio del uso de palabras anglosajonas o de la presencia de comercios vinculados con la gentrificación comercial. Por último y, con menos desarrollo, se han investigado

los mecanismos que el poder del código genera para la producción de ciertas tendencias vinculadas con la gentrificación. En este sentido, a lo largo del desarrollo de los capítulos se ha visto como los diferentes poderes digitales pueden ayudar a entender las nuevas formas de producción de espacialidades gentrificadas. En consecuencia, es posible proponer una gentrificación aumentada digitalmente, no solo como reflejo inerte de un fenómeno *offline*, sino como un proceso de co-producción e co-incidencia espacial y social.

## **5.2 Sobre la desigualdad digital y el perfil sociodemográfico en el uso de redes sociodigitales**

En la actualidad en medios de comunicación o conversaciones cotidianas se identifican referencias a las redes sociodigitales. En especial, estos comentarios refieren a estas plataformas como medios de información, ya sea de carácter político o en la cultura popular. Sin embargo, poco se sabe sobre los determinantes sociales para el uso de estas plataformas. Al respecto, el trabajo de investigación ha avanzado en el conocimiento sobre los factores sociodemográficos que explican el uso de Internet y redes sociodigitales como Twitter en México y, particularmente, en la Zona Metropolitana de la Ciudad de México. Esto, con el objetivo de aportar un marco de entendimiento más amplio sobre las desigualdades digitales y determinantes subyacentes en la distribución inequitativa del acceso a la dimensión digital.

Al respecto, el trabajo de tesis dialoga directamente con los estudios sobre las desigualdades digitales y los determinantes sociales que explican su uso entre la población. Así, se ha identificado un vasto cuerpo de literatura que ha analizado las desigualdades digitales determinadas por desigualdades sociales preexistentes, tales como las desigualdades de ingreso, escolaridad, edad, habilidades, motivaciones y sexo (DiMaggio y Hargittai, 2001; Hargittai y Hsieh, 2013; van Dijk, 2012).

En este orden de ideas se ha identificado que, a pesar del interés en la dimensión digital y de sus determinantes sociales, poco se ha avanzado hacia los determinantes sociodemográficos del uso de herramientas dentro de Internet. De esta manera, son pocos los estudios que se han preguntado sobre cuáles son los perfiles de usuarios de Twitter o Facebook (Greenwood y cols., 2016; Hargittai y Litt, 2011, 2012; A. Smith, 2010; A. Smith y Brenner, 2012). Aunado a esto, poco se ha indagado sobre la dimensión territorial de las desigualdades digitales (Pick y Sarkar, 2016; Toudert, 2015). En este contexto este trabajo de investigación ofrece un poco de luz sobre ambas dimensiones de

lo digital en México.

La desigualdad digital en México y en la Zona Metropolitana de la Ciudad de México se ha vinculado con el poder desigualmente distribuido del acceso a la dimensión digital. Por ende, el estudio ha propuesto una aproximación de selectividad de los usuarios. Es decir, se ha mostrado que para el caso mexicano y metropolitano la principal barrera para participar de la dimensión digital es tener acceso a Internet. Esta primera barrera se encuentra asociada al estrato socioeconómico de las personas. En esta instancia se han mostrado las desigualdades sociales para el acceso a Internet en ambos niveles geográficos, ya que, mientras en el conjunto del país alrededor de 60 % tiene acceso a Internet, en la zona metropolitana esta cifra asciende a 70 %.

La segunda barrera de selectividad que las personas deben enfrentar en la dimensión digital es el uso de alguna red sociodigital. Al respecto, se ha mostrado que el uso de redes sociodigitales es generalizado entre los internautas, siendo Facebook la plataforma por excelencia entre los usuarios de Internet. En este nivel, no se encontraron diferencias sustanciales entre el uso nacional y el metropolitano. Esto es un indicador de que las redes sociales, particularmente Facebook, se usa homogéneamente en el territorio, considerando la salvedad de la primera barrera en el acceso.

Finalmente, se ha realizado un análisis detallado del perfil sociodemográfico de los usuarios de Twitter. Al respecto, se ha visto que esta red sociodigital en el nivel nacional es usada por alrededor de 13 % de los internautas, mientras que su uso asciende en la zona metropolitana a cerca de 22 % de los usuarios de Internet. Particularmente, se ha visto que el uso de esta plataforma se encuentra vinculado con los estratos medios y altos de la población. De igual manera, se relaciona con perfiles con escolaridad alta y adultos jóvenes. Estos hallazgos apuntan a la presencia de un sesgo de Twitter como una posible fuente de información.

La identificación de los perfiles sociodemográficos en entornos urbanos, junto con el énfasis en las barreras al acceso es una perspectiva que poco se ha analizado dentro del campo de la desigualdad digital. Los hallazgos realizados en este rubro no solo apuntan a la construcción de poderes de distribución basados en desigualdades sociales, sino a dotar del análisis de Twitter de un contexto sociodemográfico poco visto en estudios previos. Al mismo tiempo, enfatizar el comportamiento diferenciado de la desigualdad digital en los distintos usos entre el contexto nacional y urbano ha proporcionado un andamiaje empírico para el análisis profundo de los grandes datos recolectados en la Ciudad de México.

Finalmente, con respecto a este tópico, es importante mencionar la necesidad de estudiar a profundidad el comportamiento general de las redes sociodigitales desde la aproximación de la desigualdad. Ya que éstas seguirán siendo foco de atención para los estudiosos y tomadores de decisiones en el terreno de lo digital a nivel mundial.

### **5.3 Temporalidad y espacialidad de Twitter en la Ciudad de México**

Twitter se ha convertido la red sociodigital más analizada en las diferentes esferas de la sociedad: iniciativa privada, organismos públicos, academia, medios de comunicación, etcétera. Esto, en gran medida, se debe a la configuración de dicha plataforma y a la posibilidad del manejo gratuito de una muestra de su información. Twitter se configura como una plataforma de interacción social donde se limita el texto a 240 caracteres y no es necesaria la reciprocidad de la conexión entre usuarios para acceder al perfil de una persona determinada. En lo que respecta a la posibilidad de acceder a su información, Twitter permite cosechar una muestra de las publicaciones diarias de manera gratuita, lo que se ha convertido en una ventana de oportunidad para el uso y exploración de la información generada en dicha plataforma.

A pesar de que Twitter ha sido ampliamente usado en estudios referentes a tendencias políticas (Baviera, Sampietro, y García-Ull, 2019; Filho y cols., 2015; Mancera y Pano, 2013; Marín-Cortés y Quintero-Jurado, 2018), el enfoque de este trabajo de investigación ha sido desde una perspectiva sociológica de la red y, particularmente, en su uso en ciernes en el campo de los estudios urbanos. En este sentido, se ha señalado que en Twitter subyacen elementos sociológicos, económicos y geográficos para entender su naturaleza.

De manera específica, Twitter es una red sociodigital que se consolida por medio del anclaje social de sus usuarios. Es decir, gran parte de lo que constituye este tipo de plataformas es la homofilia de sus usuarios. En otro terreno, se ha visto que Twitter no solo se constituye como una red de usuarios que muestran su cotidianidad, sino que por medio de Twitter se pueden acelerar procesos de compra-venta gracias a la publicidad involuntaria que las personas realizan en sus publicaciones, en otras palabras se convierten en prosumidores (Fuchs, 2014; Rosegger, 2013).

En términos geográficos Twitter tiene una dimensión geográfica (Takhteyev y cols., 2012). En este campo, Twitter es visto como uno de las herramientas que están haciendo avanzar a la geocomputación y la nueva revolución cuantitativa de geografía, basada en los Sistemas de Información

Geográfica. Esto gracias a la presencia de GPS en los teléfonos móviles se pueden obtener datos a nivel de coordenadas puntuales de los usuarios de redes sociodigitales. Ante esto, la geografía y los SIG tienen enormes retos por delante.

El uso de Twitter como fuente de información permite analizar la dimensión temporal de las emisiones y por ende de su contenido. Dicha ventaja es un aporte que los grandes datos posibilitan, ya que toda su información es generada y recolectada con dimensiones temporales precisas. Gracias a este avance técnico, ahora se pueden analizar desde un punto de vista cuantitativo la temporalidad de la vida cotidiana de las ciudades (Neuhaus, 2015). Dicho avance ha llevado a plantear la necesidad de realizar estudios en tiempo real con un fundamento teórico del tiempo y el ritmo de las ciudades. En consecuencia, muchos de los trabajos teóricos que han comenzado a apuntar a esta dimensión temporal han recuperado una propuesta de Henri Lefebvre poco explorada hasta el momento, ésta es el análisis del ritmo de la vida urbana y la vida cotidiana (Lefebvre, 2004).

Para la Ciudad de México, la temporalidad de tweets y su concentración espacial apunta a una desigualdad en la distribución de la emisión. En primera instancia, se ha encontrado que los usuarios de Twitter emiten con mayor frecuencia publicaciones georreferenciadas durante los fines de semana y en horarios nocturnos. En este sentido, se avanza en el conocimiento de las causas que Elwood y cols. (2012) para entender por qué las personas deciden otorgar voluntariamente su ubicación geográfica.

Así, las emisiones de la Ciudad de México pueden sugerir que no solo es un elemento de narcisismo, como lo señalan los autores, sino que se vincula con su cotidianidad y su anclaje social en el espacio que habitan. Por ejemplo, el aumento de tweets georreferenciados en horarios de esparcimiento apunta a la necesidad de reflejar en Twitter las actividades positivas realizadas en determinados espacios. Dichos espacios de concentración de emisión se asocian espacialmente con espacios identificados como gentrificados.

En este punto, la identificación de colonias gentrificadas bajo criterios comerciales y de renovación urbana ha permitido vincular estos espacios con el tipo de concentración de emisión. Por ejemplo, se ha mostrado que la mayor concentración de emisiones en la Ciudad de México, independientemente del momento del día, se ubica en las colonias gentrificadas. Este hecho no es azaroso, se propone entenderlo a la luz de la producción digital de las tendencias de centralización y valorización del espacio urbano que forman parte del proceso de producción del fenómeno de

gentrificación. En otras palabras, la emisión de tweets en las colonias centrales de la ciudad es una expresión de la producción de una realidad aumentada digital vinculada con la gentrificación.

#### **5.4 Sobre una propuesta sociolingüística para entender el contenido del texto en Twitter**

. Entender la dimensión digital de la gentrificación desde la mirada de la RGAD lleva a indagar sobre el proceso de comunicación que se establece en las plataformas digitales. En este sentido, la comunicación, en este caso escrita, es otro medio que interviene en el proceso de valorización e igualación del espacio gentrificado. Al respecto estudios anteriores han avanzado en diferentes niveles de análisis, ya sea bajo la perspectiva sociolingüística del paisaje lingüístico o el análisis cualitativo de blogs de reseñas de restaurantes (Díaz, 2015; Lyons y Rodríguez-Ordóñez, 2017; Zukin y cols., 2017). Otra vertiente de trabajos ha aproximado la comunicación escrita de Twitter y su vínculo con la gentrificación por medio de la pérdida de comunidad (Gibbons y cols., 2018). Sin embargo, los estudios mencionados solo utilizan las plataformas digitales como una nueva fuente de información, sin considerar la interacción entre el proceso urbano y la dimensión digital, y cómo esta última interviene.

Ante el gran volumen de datos no numéricos que implica el análisis de texto de Twitter, en entornos urbanos se han realizado aproximaciones utilizando algoritmos de Aprendizaje de Máquina. Dentro de estos trabajos se destacan dos clases de análisis. Los primeros se han enfocado a aproximaciones de Aprendizaje de Máquina supervisados, específicamente el análisis de sentimientos. El segundo cuerpo de literatura se ha volcado a algoritmos no supervisados, como la Asignación Latente de Dirichlet.

En este trabajo de investigación se ha decidido tomar distancia del primer enfoque metodológico. En primera instancia porque los algoritmos supervisados requieren de mayor tiempo y recursos para lograr obtener resultados satisfactorios. En segundo lugar, porque el andamiaje conceptual del análisis de sentimientos parte de una postura psicoanalítica que se enfoca en la interpretación del texto en Twitter bajo etiquetas de sensaciones (Liu, 2015; Thelwall y cols., 2011). De esta manera, la aproximación de análisis de sentimientos no permitiría evidenciar el papel activo que los mensajes de Twitter tienen en la generación de espacialidades digitales vinculadas con la gentrificación. En este sentido, se hace un primera crítica a este tipo de trabajos, en tanto no reflexionan sobre el rol del lenguaje en la producción del espacio, en este caso, por medio de Twitter.

En consecuencia, se ha propuesto un acercamiento interdisciplinario que conjunta aproximaciones de la geografía digital, los estudios urbanos y la sociolingüística. En este sentido, el trabajo de investigación ha considerado la propuesta teórica de Bühler (1961) que permite entender a las palabras como símbolos de cosas o hechos concretos. Así, a las palabras encontradas en los tweets de la Ciudad de México se les ha considerado signos que reflejan lo que ocurre cotidianamente en la urbe. En segunda instancia, se ha retomado la dimensión de síntoma lingüístico, la cual apunta a que las palabras son reflejo del lugar, tiempo o cultura donde son emitidas. En este tenor, se conjuntan las visiones geográfica y lingüística, en tanto que las palabras encontradas en colonias gentrificadas son reflejo del síntoma del proceso urbano.

De manera empírica, el estudio ha arrojado resultados relevantes en torno a la identificación de espacios lingüísticos vinculados a la gentrificación. Por ejemplo, se ha mostrado que en espacios gentificados se tiene una tendencia al uso de anglicismos en Twitter. Esto compagina con estudios sociolingüísticos que han apuntado al uso de palabras de origen extranjero en comercios que experimentan el fenómeno de gentrificación alrededor del mundo (Papen, 2012). Para el caso de la Ciudad de México resalta la misma tendencia al uso de anglicismos en los tweets emitidos en las colonias centrales. Este hallazgo apunta a la misma tendencia en el uso del lenguaje tanto en los espacios físicos como en los digitales.

Otro hallazgo relevante ha sido la identificación de patrones léxicos vinculados con la gentrificación en tweets emanados fuera de dichos espacios. En este rubro, al haber realizado un análisis comparativo entre los patrones de palabras emitidas en colonias gentrificadas contra las no gentrificadas, se han desvelado fuera de los espacios gentificados palabras que apuntan a la reproducción del consumo característico en ellos. Esto abre la puerta al vínculo simbólico y cultural de consumir productos de boutiques o cafeterías en espacios comerciales controlados. Por ejemplo, se compara el patrón léxico de la calle de Madero, símbolo de gentrificación comercial en la ciudad y el patrón léxico encontrado en centros comerciales como Parque Delta, ambos espacios controlados y dirigidos a consumo de alto ingreso.

La implementación de un método no supervisado ha permitido encontrar los principales temas latentes de Twitter en las colonias gentrificadas. En este punto se pueden resaltar dos aspectos. Primero, a la luz de una interpretación de las palabras como signos y reflejo del síntoma de gentrificación se han analizado los principales tópicos que semánticamente refieren al proceso urbano.

Así, se han identificado tópicos que también han sido identificados en estudios previos sobre el paisaje lingüístico en colonias gentrificadas y que hacen referencia a la cultura *hispter*, la cultura gay o comida-bebida. Esta misma herramienta ha permitido observar la geografía de los espacios léxicos. Así, se ha encontrado que los temas latentes íconos de la gentrificación en Twitter se relacionan geográficamente con los espacios físicos. En este sentido, se propone entender el texto y las palabras en Twitter como símbolos que se encuentran anclados a un síntoma cultural, económico y urbano, que en este caso son también muestra de las prácticas de gentrificación.

El segundo punto es la identificación de temas que no se relacionan directamente con el fenómeno de gentrificación, pero que coexisten en el mismo espacio físico. Ejemplo de esto es el tópico latente referente al Tribunal Judicial en la Ciudad de México. Este tema ha sido poco abordado en el trabajo de investigación, pero abre preguntas sobre la coexistencia de otros mecanismos para la valorización del espacio urbano.

Finalmente el estudio léxico de los temas latentes de Twitter en la ciudad ha arrojado luz en cuanto a las principales preocupaciones de sus usuarios. A diferencia del conjunto de trabajos académicos y no académicos que utilizan la red sociodigital para analizar temas políticos, se ha encontrado que éstos, ante la mirada espacial, pierden relevancia ante tópicos ligados al consumo. Estos es un descubrimiento interesante y permite lanzar cuestionamientos ante los estudios políticos basados en esta red sociodigital.

## **5.5 Los grandes datos en los estudios urbanos**

Uno de los principales retos que ha abordado este trabajo ha sido el manejo de los grandes datos. Al respecto se ha intentado avanzar en dos grandes vertientes. La primera es dar una interpretación teórica a los grandes datos, partiendo de la premisa de que éstos no solo son reflejo de la realidad, sino que son producto de ella y al mismo tiempo son actantes en el proceso de generación de realidades. La segunda vertiente ha girado en torno a los retos empíricos que los grandes datos implican para la investigación en estudios urbanos y la compaginación de estas nuevas fuentes de información con fuentes de datos oficiales para incrementar el potencial de los primeros.

En cuanto al primer tópico, la investigación ha tomado una postura crítica de los grandes datos. En este sentido, se ha planteado la necesidad de pensar los grandes datos no como reflejo estéril de las actividades cotidianas, sino como herramientas políticas y económicas de producción

de espacialidades. Así, el trabajo propone abandonar una definición descriptiva de los grandes datos basada en las características computacionales de éstos, la definición de las 3 Vs. En su lugar, se piensa que los grandes datos son socialmente y políticamente construidos y requieren de cuerpos teóricos de diversas disciplinas para su análisis.

La importancia de una postura crítica ante los grandes datos es la tendencia generalizada del poder absoluto de los datos. Ante esto, es necesario la reflexión profunda en el manejo de este tipo de bases, la consideración de las cuestiones éticas y las repercusiones sociales en dicho manejo. Al mismo tiempo, es necesario el diálogo multidisciplinario de los grandes datos y de los estudios digitales, ya que, en la actualidad se encaran fronteras en el conocimiento.

Los grandes datos tienen fuertes limitaciones. Como se ha mostrado a lo largo de la tesis, el principal obstáculo que estos datos enfrentan es su sesgo para representar a la totalidad de la sociedad, particularmente sociedades como la mexicana, en la cual el acceso a Internet es limitado. De igual manera, el análisis de los resultados provenientes de grandes datos pocas veces pasa por un filtro de quiénes son los emisores de estos datos. Por ejemplo, en el caso de Twitter se están representando intereses de personas de estratos medios y altos. Otra de las limitaciones es la propiedad de esta información.

El segundo gran eje sobre los grandes datos han sido los retos enfrentados para su manejo. Al respecto, el campo de los estudios urbanos tiene una gran cantidad de esfuerzos por realizar, esfuerzos que pueden encaminar nuevas vertientes epistemológicas en el área de estudio. Por ejemplo, la necesidad de tener marcos conceptuales y herramientas estadísticas que permitan analizar los datos espacio-temporales, sobre todo ante la posibilidad de manejar información que detalla, segundo a segundo, de las actividades de las personas.

Otro punto relevantes es la búsqueda de métodos híbridos para el análisis de grandes datos. Por ejemplo, buscar la combinación entre análisis léxico y cuantitativo. En este sentido, un importante reto es el diálogo entre el manejo de grandes datos y fuentes de datos tradicionales. Para este trabajo se ha logrado hacer una combinación que ha permitido complementar y profundizar el estudio, gracias a la utilización de fuentes de información generadas por INEGI y Twitter. Finalmente, el manejo de técnicas de frontera basadas en aprendizaje de máquina, aprendizaje profundo o redes neuronales. En este punto, cabe mencionar la relevancia de conocer las limitaciones y ventajas de cada una de estas técnicas y la complementariedad con técnicas del campo de la estadística

inferencial.

Por último, dentro de este apartado es necesario mencionar la necesidad de incluir en los planes docentes a nivel de pregrado y posgrado la enseñanza de las técnicas relacionadas con el manejo de grandes bases de datos en disciplinas como la economía, la sociología, el urbanismo, las ciencias políticas, la geografía y otras ciencias relacionadas con los estudios urbanos. Dado que la generación de información georreferenciada se incrementará en los próximos años con aplicaciones como Google Maps o OpenStreetMaps, serán necesarios cuadros de profesionales en estudios urbanos y otras disciplinas capaces de procesar y analizar este tipo de información de manera crítica. En este campo, se están generando nuevos paradigmas epistemológicos y técnicos en las ciencias que involucran la reflexión del espacio, como la geografía, los SIG o los estudios urbanos

### **5.6 Tópicos sin desarrollar y retos a futuro en la investigación de la geografía digital, los grandes datos y las ciudades**

La evolución de esta investigación ha sido un camino de descubrimientos hacia un campo en reciente desarrollo: las geografías digitales, los estudios urbanos y los grandes datos. La interacción de estos tres temas ha llevado a investigar el proceso de gentrificación digitalmente aumentada y los factores que se asocian a ello. Esto ha implicado un desafío metodológico que se ha expuesto anteriormente, la conjunción entre la propuesta de la RGAD y las fuerzas estructurales del proceso de gentrificación. En este sentido, se ha propuesto un marco analítico para entender la gentrificación aumentada. Es decir, cómo los tres poderes de la gentrificación intervienen en aumentar las tres fuerzas del proceso de gentrificación. Esto llevaría a entender que existen 9 mecanismos básicos de interacción para la gentrificación digital.

Sin embargo, no ha sido posible desarrollar a cabalidad una de las interacciones en el presente trabajo. Ésta cuestiona ¿cómo el poder del código interviene en el proceso de centralización, valorización o generación de tendencias desiguales? La limitación se debe a la imposibilidad de conocer el código subyacente a las publicaciones de Twitter, por un lado. Por el otro, esto requeriría abordar y conocer el papel activo que las cuentas automatizadas o *bots* ejecutan para la generación de espacialidades gentrificadas, ya sea por medio de la distribución espacial de las emisiones o bien por el contenido léxico de los mensajes. En este sentido, otro factor ha sido la naturaleza de la fuente de información, ya que Twitter es una empresa privada que no proporciona la programación

subyacente.

De manera general, es importante considerar en futuras investigaciones, que relacionen la dimensión digital y la ciudad, el papel activo del algoritmo. Ya que en la actualidad gobiernos locales alrededor del mundo están desarrollando plataformas digitales hacia teléfonos móviles para mejorar trámites, denuncias o vigilancia. De manera que, los resultados que dichas plataformas puedan tener serán consecuencia del código de dichos desarrollos. Por tanto, es importante empujar hacia una visión del código como una herramienta de generación de espacios urbanos.

Otro de los elementos poco desarrollados han sido las motivaciones y habilidades digitales de los usuarios de Twitter en México. En este sentido, se abre la puerta a futuras investigaciones sobre el entorno físico y social como factor para el uso de ciertas plataformas digitales. En el mismo hilo de ideas, la identificación de motivaciones para el uso de estos sitios puede aportar al entendimiento de la información geográfica basada en los ciudadanos y que se genera en plataformas digitales.

En términos del procesamiento de datos se tienen varios temas pendientes para futuras investigaciones. El primero de ellos es la identificación de entidades nombradas en Twitter. Este ha sido el primer estudio que ha enfocado las entidades nombradas en el corpus de Twitter desde una perspectiva urbana. Así, las pocas entidades nombradas identificadas pueden aumentar por medio de la implementación de algoritmos de identificación automatizada de las mismas, siempre bajo la perspectiva espacial. El segundo reto refiere al uso de métodos supervisados de Aprendizaje de Máquina, particularmente basados en la técnica de análisis de sentimientos. Se ha mencionado que esta herramienta se basa en teoría psicoanalítica para identificar emociones, pero ha sido una de las más desarrolladas en el campo del procesamiento del lenguaje natural para grandes datos. En este sentido, el reto es aprovechar las ventajas que estos avances implican y dotarla de una base teórica sustentada en una reflexión social de los fenómenos urbanos.

También ha quedado en el tintero el análisis de las geografías de tópicos no referentes al proceso de gentrificación. Al respecto, conocer la concentración de temas que coexisten en espacios predominantemente gentrificados permitiría conocer, con mayor profundidad, las sinergias espaciales que se realizan en los espacios gentrificados. En este mismo tenor, queda pendiente el análisis de los temas latentes de las colonias no gentrificadas y si éstos compaginan o no con los identificados en su contraparte.

Finalmente, los temas a futuro sobre la dimensión digital y la ciudad son innumerables. Uno

de ellos es el vínculo entre la desigualdad y la automatización de los procesos, por ejemplo, gubernamentales, financieros o comerciales. En este sentido, indagar más sobre la desigualdad digital es más importante que nunca. Con el avance del Internet de las Cosas y la Inteligencia Artificial se puede analizar ¿cómo estas dimensiones podrán mejorar o magnificar las tendencias de producción de espacio desigual en las décadas próximas?. Otro tópico es la gobernanza de los datos, por ejemplo, cómo las herramientas de análisis en este campo servirán para dictar comportamientos o procesos en las ciudades. Por último, se encuentra el tema de las consecuencias éticas por el uso masificado de plataformas digitales, al respecto, se han planteado posturas críticas para un nuevo capitalismo basado en la vigilancia digital. Ante el escenario de ruptura global generado por la pandemia de causada por el SARS-CoV2, se incrementan las preocupaciones sobre las consecuencias sociales que la omnipresencia del código y lo digital pueda causar para la vida venidera y el desarrollo de los espacio urbanos.



## Referencias

- Achrekar, H., Gandhe, A., Lazarus, R., Ssu-Hsin Yu, y Liu, B. (2011). Predicting flu trends using twitter data. En *2011 IEEE Conference on Computer Communications Workshops (InfoCom Wkshps)* (p. 702-707).
- Agresti, A. (2013). *Categorical Data Analysis*. Hoboken, Nueva Jersey: WILEY.
- Aguayo-Ayala, A. (2016). Nuevo polanco: renovación urbana, segregación y gentrificación en la ciudad de México. *Iztapalapa. Revista de ciencias sociales y humanidades*, 37, 101 - 123.
- Alp, Z. Z., y Ögüdücü, Ş. G. (2018). Identifying topical influencers on Twitter based on user behavior and network topology. *Knowledge-Based Systems*, 141, 211 -221. doi: <https://doi.org/10.1016/j.knosys.2017.11.021>
- AlSayyad, N., y Guvenc, M. (2015). Virtual uprisings: On the interaction of new social media, traditional media coverage and urban space during the 'arab spring'. *Urban Studies*, 52(11), 2018-2034.
- Álvarez-Bornstein, B., y Montesi, M. (2016). Researchers' communication on Twitter: A virtual ethnography in the area of information science. *Revista Española de Documentación Científica*, 39(4).
- Andrews, L. (2011). *I know who you are and i saw what you did. social networks and the death of privacy*. Free Press.
- Atil-Bulut, Z., y Dogan, O. (2017). The abc typology: profile and motivations of turkish social network sites users. *Computers in Human Behavior*, 67, 73-83.
- Atkinson, R. (2000). Measuring Gentrification and Displacement in Greater London. *Urban Studies*, 37(1), 149-165. doi: 10.1080/0042098002339
- Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355-385. doi: 10.1162/pres.1997.6.4.355
- Balzarini, J. E., y Shlay, A. B. (2016). Gentrification and the right to the city: Community conflict and casinos. *Journal of Urban Affairs*, 38(4), 503-517.
- Batres-Guadarrama, L. (2016). Gentrificación comercial en la ciudad de México y nuevas formas de lucha. En *Contested cities. congreso internacional madrid 2016* (p. 1-10).
- Baviera, T., Sampietro, A., y García-Ull, F. J. (2019). Political conversations on twitter in a disruptive scenario: The role of "party evangelists" during the 2015 spanish general elections. *The Communication Review*, 22(2), 117-138.
- Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Contents lists available at ScienceDirect*, 67, 1595-1600.
- Belkind, L. (2008). The internet and the city: blogging and gentrification on new york's lower east side. En D. E. Davis y C. Proenza-Coles (Eds.), *Political power and social theory* (Vol. 19, p. 207-233). Emerald Group Publishing Limited.

- Benediktsson, M. O., Lamberta, B., y Larsen, E. (2016). Taming a “chaotic concept”: gentrification and segmented consumption in brooklyn, 2002–2012. *Urban Geography*, 37(4), 590-610. Descargado de <https://doi.org/10.1080/02723638.2015.1096113> doi: 10.1080/02723638.2015.1096113
- Bernie Hogan. (2018). Digital Traces in Context— Social Media Giveth, Social Media Taketh Away: Facebook, Friendships, and APIs. *International Journal of Communication*, 12(0).
- Bird, S., Klein, E., y Loper, E. (2009). *Natural Language Processing with Python*. Estados Unidos de América: O’Reilly.
- Blei, D. M., Ng, A. Y., y Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bloomberg. (2017). *Airbnb enters the land of profitability*. Descargado 3 de mayo de 2018, de <https://www.bloombergquint.com/technology/2017/01/26/airbnb-enters-the-land-of-profitability>
- Boeing, G., Besbris, M., Schachter, A., y Kuk, J. (2020). Housing search in the age of big data: Smarter cities or the same old blind spots? *Housing Policy Debate*, 0(0), 1-15. Descargado de <https://doi.org/10.1080/10511482.2019.1684336>
- Borgmann, H., Loeb, S., Salem, J., Thomas, C., Haferkamp, A., Murphy, D. G., y Tsaour, I. (2016). Activity, content, contributors, and influencers of the twitter discussion on urologic oncology. *Urologic Oncology: Seminars and Original Investigations*, 34(9), 377-383. doi: <https://doi.org/10.1016/j.urolonc.2016.02.021>
- Bournazou, E. (2016). Cambios socioterritoriales e indicios de gentrificación. Un método para su medición. *Academia XXII · UNAM*, 6(12), 47-59.
- Brenner, N. (2013). Tesis sobre la urbanización planetaria. *Nueva Sociedad*, 243(1), 38-66.
- Brunsdon, C., y Comber, L. (2019). *An introduction to R for spatial analysis & mapping*. Londres: SAGE.
- Bühler, K. (1961). *Teoría del lenguaje*. Madrid: Revista de Occidente.
- Butler, T. (2005). Gentrification. En N. Buck, I. Gordon, A. Harding, y I. Turok (Eds.), *Changing cities: Rethinking urban competitiveness, cohesion and governance* (p. 172-187). Palgrave MacMillan with ESRC Cities Programme.
- Cairncross, F. (2001). *The death of the distance: how the communications revolution is changing our lives*. Boston Mass and Harvard Bussiness School.
- Castaño, C., Martín, J., y Vázquez, S. (2008). La brecha digital de género: acceso, uso y habilidades. En C. Castaño (Ed.), *La segunda brecha digital* (p. 55-132). Ediciones Cátedra.
- Castells, M. (1995). *La ciudad informacional : tecnologías de la información, reestructuración económica y el proceso urbano-regional*. Madrid: Madrid : Alianza, 1995.
- Castells, M. (2012). *Redes de indignación y esperanza*. Madrid: Alianza Editorial.
- Cebollada, À., Sillero, A., y Anaya, E. (2018). Indagando en Sant Antoni (Barcelona) en torno a la relación de la bicicleta con la gentrificación. En T. S. Marques, H. Santos, y M. P. Alonso (Eds.), *Viii jornadas de geografía económica: La geografía de las redes económicas y la geografía económica en redes* (p. 169-177). Faculdade de Letras da Universidade do Porto,.
- CEPAL. (2013). *Economía digital para el cambio estructural y la igualdad* (Inf. Téc.). Autor.
- Chaskin, R., y Joseph, M. (2013). ‘Positive’ Gentrification, Social Control and the ‘Right to the City’ in Mixed-Income Communities: Uses and Expectations of Space and Place. *International Journal of Urban and Regional Research*, 37(2), 480–502. doi: 10.1111/j.1468-2427.2012.01158.x

- Chavoshi, N., Hamooni, H., y Mueen, A. (2016). Identifying correlated bots in twitter. En E. Spiro y Y.-Y. Ahn (Eds.), *Social informatics* (p. 14-21). Cham: Springer International Publishing.
- Cheng, J., Gould, N., Han, L., y Jin, C. (2016, 07). Big data for urban studies: Opportunities and challenges: A comparative perspective. En (p. 1229-1234). doi: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0189
- Cheung, C., Chui, P.-Y., y Lee, M. (2011). Online social networks: why do students use facebook? *Computers in Human Behavior*, 27, 1337-1343.
- CNN. (s.f.). *México mide el estado de ánimo de sus tuiteros: ¿son más positivos o negativos?*
- CNN. (2014). *African hipsters take vintage fashion back to the future.*
- Constantinides, E., y Zinck-Stagno, M. (2011). Potential of the social media as instruments of higher education marketing: a segmentation study. *Journal of Marketing for Higher Education*, 21(1), 7-24.
- Coronado, A. (2015). *Big data en 30 min. (recolectando tuits).* Descargado 8 de mayo de 2018, de <https://abxda.wordpress.com/2015/12/13/big-data-en-30-min-recolectando-tuits/>
- Cortés, F., y Vargas, D. (2011). Marginación en México a través del tiempo: a propósito del índice de Conapo. *Estudios Sociológicos*, 29(86), 361-387.
- Crampton, J., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wilson, M., y Zook, M. (2013). Beyond the geotag: situating 'big data' and leveraging the potential of the web. *Cartography and Geographic Information Science*, 40(2), 130-139.
- Croitoru, A., Wayant, N., Crooks, A., Radzikowski, J., y Stefanidis, A. (2015). Linking cyber and physical spaces through community detection and clustering in social media feeds. *Computers, Environment and Urban Systems*, 53, 47-64.
- Crossa, V. (2009). Resisting the entrepreneurial city: street vendor struggle in mexico city's historic center. *International Journal of Urban and Regional Research*, 33(1), 43-63.
- Dalenius, T., y Hodges, J. (1959). Minimum variance stratification. *Journal of the American Statistical Association*, 54(285), 88-101.
- Danyluk, M., y Ley, D. (2007). Modalities of the New Middle Class: Ideology and Behaviour in the Journey to Work from Gentrified Neighbourhoods in Canada. *Urban Studies*, 44(11), 2195-2210. Descargado de <https://doi.org/10.1080/00420980701520277> doi: 10.1080/00420980701520277
- Daume, S., y Galaz, V. (2016). "anyone know what species this is? twitter conversations as embryonic citizen science communities. *PLoS ONE*, 11.
- Davis, K. (2013). Young people's digital lives: the impact of interpersonal relationships and digital media use on adolescents of identity. *Computers in Human Behavior*, 29, 2281-2293.
- de Lourdes Cordero-Gómez, L., y Salinas, L. (2017). Gentrificación comercial. Espacios escenificados y el modelo de los mercados gourmet. *Revista de Urbanismo*(37), 1-12.
- de Souza e Silva, A. (2006). From cyber to hybrid: Mobile technologies as interfaces of hybrid spaces. *Space and Culture*, 9(3), 261-278. doi: 10.1177/1206331206289022
- de Souza e Silva, A. (5). Pokémon Go and HRG: Mobility, sociability and surveillance in hybrid spaces. *Mobile Media & Communication*, 1(20-23).
- de Souza e Silva, A., Duarte, F., y Damasceno, C. (2017). Creative appropriation in hybrid spaces. Mobiles interfaces in art and games in Brazil. *International Journal of Communication*, 11, 1705-1728.

- D'heer, E., y Verdegem, P. (2014). Conversations about the elections on Twitter: Towards a structural understanding of Twitter's relation with the political and the media field. *European Journal of Communication*, 29(6), 720-734.
- Díaz, J. (2015). Gentrificación por la red: nuevos actores de clase en el centro histórico de la ciudad de México. En V. Delgadillo, Ibán, y L. Salinas (Eds.), *Perspectivas del estudio de la gentrificación en México y América Latina* (p. 303-322). México: UNAM and UNAM posgrado de Urbanismo and Contested Cities and Instituto de Geografía UNAM.
- Díaz-Parra, I. (2013). La gentrificación en la cambiante estructura socioespacial de la ciudad. *Revista Bibliográfica de Geografía y Ciencias Sociales*, 18(1030).
- Díaz-Parra, I., y Salinas, L. (2016). La producción del consumidor. valorización simbólica y gentrificación en el centro de la ciudad de México. *Andamios*, 13(32), 107-130.
- Diebold, F. X. (2012). *A Personal Perspective on the Origin(s) and Development of 'Big Data': The Phenomenon, the Term, and the Discipline, Second Version*.
- DiMaggio, P., y Hargittai, E. (2001). *From the digital divide to digital inequality: studying internet use as penetration increases* (Working Paper n.º 15). Center for Arts and Cultural Policy Studies, Princeton University.
- di Virgilio, M., y Guevara, A. (2015). Gentrificación liderada por el Estado y empresarismo urbano en la Ciudad Autónoma de Buenos Aires. En V. Delgadillo, I. Díaz-Parra, y L. Salinas (Eds.), *Perspectivas del estudio de la gentrificación en México y América Latina* (p. 11-30). Ciudad de México: UNAM and UNAM posgrado de Urbanismo and Contested Cities and Instituto de Geografía UNAM.
- Dodge, M., y Kitchin, R. (2001). *Mapping cyberspace*. Londres: Routledge.
- Dodge, M., y Kitchin, R. (2005). Code and the transduction of space. *Annals of the Association of American Geographers*, 95(1), 162-180. Descargado de <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8306.2005.00454.x> doi: 10.1111/j.1467-8306.2005.00454.x
- Duhau, E., y Giglia, Á. (2008). *Las reglas del desorden: habitar la metrópoli*. México: Editorial Siglo XXI and UAM Azcapotzalco.
- Duhau, E., y Giglia, A. (2016). *Metrópolis, espacio público y consumo*. México: Fondo de Cultura Económica.
- Echarri-Cánovas, C. (2008a). Construcción del estrato socioeconómico en la ENDIREH 2006. En R. Castro y I. Casique (Eds.), *Violencia de género en las parejas mexicanas. Análisis de resultados de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2006. Componente para mujeres unidas o casadas de 15 años y más* (p. 170-175). Inmujeres, CRIM-UNAM.
- Echarri-Cánovas, C. (2008b). Desigualdad socioeconómica y salud reproductiva: una propuesta de estratificación social aplicable a encuestas. En S. Lerner y I. Szasz (Eds.), *Salud reproductiva y condiciones de vida en México* (p. 59-113). México: El Colegio de México.
- ECOBICI. (s.f.). *Mapa de cicloestaciones*. Descargado 13 de agosto de 2019, de <https://www.ecobici.cdmx.gob.mx/es/mapa-de-cicloestaciones>
- Elwood, S. (2010). Geographic information science: Visualization, visual methods, and the geo-web. *Progress in Human Geography*, 35(3), 401-408.
- Elwood, S., Goodchild, M. F., y Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers*, 102(3), 571-590.

- Elwood, S., y Leszczynski, A. (2018). Feminist digital geographies. *Gender, Place & Culture*, 25(5), 629-644. Descargado de <https://doi.org/10.1080/0966369X.2018.1465396>  
doi: 10.1080/0966369X.2018.1465396
- Expósito, V. I., Rodríguez, P. A., y Barrau, A. N. (2017). Comunicación y divulgación de contenidos artísticos a través de las redes sociales: Facebook y twitter. *Estudios sobre el Mensaje Periodístico*, 23(2), 1161-1178.
- Feinerer, I., y Hornik, K. (2018). tm: Text Mining Package [Manual de software informático].
- Feinerer, I., Hornik, K., y Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software*, 25(5), 1-54.
- Fellows, I. (2018). wordcloud: Word clouds [Manual de software informático].
- Filho, R. M., Almeida, J. M., y Pappa, G. L. (2015, Aug). Twitter population sample bias and its impact on predictive outcomes: A case study on elections. En *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (p. 1254-1261).  
doi: 10.1145/2808797.2809328
- Flores, N., y Lewis, M. (2016). From truncated to sociopolitical emergence: A critique of superdiversity in sociolinguistics. *International Journal of the Sociology of Language*, 2016(241), 97-124.
- Florida, R. (2005). *Cities and the creative class*. Nueva York: Routledge.
- Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Inglaterra: Oxford University Press.
- Franco, E. D. (2016). *El anglicismo en el Español nacional de México* (Tesis Doctoral no publicada). El Colegio de México; Centro de Estudios Lingüísticos y Literarios, Ciudad de México.
- Friedman, G. (2014). Workers without employers: shadow corporations and the rise of the gig economy. *Review of Keynesian Economics*, 2(2), 171 - 188.
- Friemel, T. (2016). The digital divide has grown old: determinants of a digital divide among seniors. *New Media and Society*, 18(2), 313-331.
- Fuchs, C. (2014). Digital prosumption labour on social media in the context of the capitalist regime of time. *Time & Society*, 23(1), 97-123.
- Füller, H., y Michel, B. (2014). 'Stop Being a Tourist!' New Dynamics of Urban Tourism in Berlin-Kreuzberg. *International Journal of Urban and Regional Research*, 38(4), 1304-1318.
- García-López, M.-À., y Muñoz, I. (2013). Urban spatial structure, agglomeration economies, and economic growth in Barcelona: An intra-metropolitan perspective. *Papers in Regional Science*, 92(3), 515-534. doi: 10.1111/j.1435-5957.2011.00409.x
- García-Palomares, J. C., Salas-Olmedo, M. H., Moya-Gómez, B., Condeço-Melhorado, A., y Gutiérrez, J. (2018). City dynamics through twitter: Relationships between land use and spatio-temporal demographics. *Cities*, 72, 310-319.
- Garnica, R. (2014). La provisión de servicios de internet en México: ¿tiene un dominio nacional de este servicio la ZMCM? En B. Graizbord (Ed.), *Metrópolis. estructura urbana, medio ambiente y política pública* (p. 175-191). El Colegio de México.
- Garrocho-Rangel, C., y Campos, J. (2007). Dinámica de la estructura policéntrica del empleo terciario en el área metropolitana de Toluca, 1994-2004. *Papeles de Población*, 13(52), 109-135.
- Gdaniec, C. (2000). Cultural industries, information technology and the regeneration of post-industrial urban landscapes. Poblenou in Barcelona - a virtual city? *GeoJournal*, 50(4), 379-

- Gibbons, J., Nara, A., y Appleyard, B. (2018). Exploring the imprint of social media networks on neighborhood community through the lens of gentrification. *Environment and Planning B: Urban Analytics and City Science*, 45(3), 470–488.
- Gill-Juárez, A., Feliu, J., y Vitores, A. (2012). Género y tic: en torno a la brecha digital de género. *Athenea Digital*, 12(3), 3-9.
- Gleason, B. (2013). #Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter. *American Behavioral Scientist*, 57(7), 966-982.
- Goffman, E. (1981). *Forms of talk*. Philadelphia, Pa: Philadelphia, Pa. : University of Pennsylvania, 1981.
- Gómez-González, A. (2016). *El patrimonio edificado a debate: entre la renovación urbana y la pérdida de su identidad. el caso de los corredores peatonales gante-filomeno mata, regina y madero del centro histórico de la ciudad de méxico* (Tesis de Master no publicada). Maestría en Estudios de la Ciudad, Universidad Autónoma del Estado de México.
- González-Bailón, S. (2014). Online social networks and bottom-up politics. En M. Graham y W. Dutton (Eds.), *Society and the internet. how networks of information and communication are changing our lives* (p. 209-222). Oxford University Press.
- Goodchild, M. F. (2007, 01). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211–221. doi: 10.1007/s10708-007-9111-y
- Goodchild, M. F., Aubrecht, C., y Bhaduri, B. (2017). New questions and a changing focus in advanced vgi research. *Transactions in GIS*, 21(2), 189-190.
- Graham, M. (2014). Internet geographies: data shadows and digital divisions of labor. En M. Graham y W. Dutton (Eds.), *Society and the internet. how networks of information and communication are changing our lives* (p. 99-116). Reino Unido: Oxford University Press.
- Graham, M. (2017). Digitally augmented geographies. En R. Kitchin, T. P. Lauriault, y M. W. Wilson (Eds.), *Understanding spatial media* (p. 44-55). SAGE Publications.
- Graham, M., De Sabbata, S., y Zook, M. A. (2015). Towards a study of information geographies: (im)mutable augmentations and a mapping of the geographies of information. *Geo: Geography and Environment*, 2(1), 88-105.
- Graham, M., Hale, S., y Stephens, M. (2012). Featured graphic: Digital divide: The geography of internet access. *Environment and Planning A*, 44(5), 1009-1010.
- Graham, M., y Sabbata, S. D. (2015). Mapping information wealth and poverty: the geography of gazetteers. *Environment and Planning A*, 47, 254 –1264.
- Graham, M., Stephens, M., y Hale, S. (2013). Featured graphic. mapping the geoweb: A geography of twitter. *Environment and Planning A*, 44, 100-102.
- Graham, M., y Zook, M. (2013). Augmented realities and uneven geographies: Exploring the geolinguistic contours of the web. *Environment and Planning A: Economy and Space*, 45(1), 77-99. Descargado de <https://doi.org/10.1068/a44674> doi: 10.1068/a44674
- Graham, M., Zook, M., y Boulton, A. (2013a). Augmented reality in urban places: contested content and the duplicity of code. *Transactions of the Institute of British Geographers*, 38(3), 464–479. doi: 10.1111/j.1475-5661.2012.00539.x
- Graham, M., Zook, M., y Boulton, A. (2013b, 2019/04/17). Augmented reality in urban places: contested content and the duplicity of code. *Transactions of the Institute of British Geographers*, 38(3), 464–479. Descargado de

<https://doi.org/10.1111/j.1475-5661.2012.00539.x> doi: 10.1111/j.1475-5661.2012.00539.x

- Graham, S. (2002). Bridging the urban digital divides? urban polarisation and information and communication technologies (icts). *Urban Studies*, 39(1), 33-56.
- Graham, S. (2015). Information geographies and geographies of information. *New Geographies*, 7, 159-166.
- Graham, S., y Marvin, S. (2001). *Splintering urbanism. networked infrastructures, technological mobilities and the urban condition* (Vol. Estados Unidos). Routledge.
- Grange, A. L., y Pretorius, F. (2016). State-led gentrification in Hong Kong. *Urban Studies*, 53(3), 506-523.
- Greenwood, S., Perrin, A., y Duggan, M. (2016). *Social media update 2016. Facebook usage and engagement is on the rise, while adoption of other platforms holds steady* (Inf. Téc.). Pew Research Center.
- Griffiths, T. L., y Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5228-5235. doi: 10.1073/pnas.0307752101
- Grün, B., y Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13), 1-30.
- Guardian, T. (2018, marzo). *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*. Descargado 8 de mayo de 2018, de <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Habermas, J. (2018, mayo). *Jürgen habermas: "¡por dios, nada de gobernantes filósofos!"*. Descargado 8 de mayo de 2018, de [https://elpais.com/elpais/2018/04/25/eps/1524679056\\_056165.html?id\\_externo\\_rsoc=tw\\_CM](https://elpais.com/elpais/2018/04/25/eps/1524679056_056165.html?id_externo_rsoc=tw_CM)
- Hackworth, J., y Smith, N. (2001). The changing state of gentrification. *Tijdschrift voor economische en sociale geografie*, 92(4), 464-477.
- Hae, L. (2011). Dilemmas of the nightlife fix: Post-industrialisation and the gentrification of nightlife in new york city. *Urban Studies*, 48(16), 3449-3465.
- Hammel, D. (2009). Gentrification. En R. Kitchin y N. Thrift (Eds.), *International encyclopedia of human geography* (p. 360 - 367). Oxford: Elsevier. doi: <https://doi.org/10.1016/B978-008044910-4.01051-8>
- Hao, J., Zhu, J., y Zhong, R. (2015). The rise of big data on urban studies and planning practices in china: Review and open research issues. *Journal of Urban Management*, 4(2), 92 - 124.
- Hargittai, E. (2010). Digital na(t)ives? variation in internet skills and uses among members of the 'net generation'. *Sociological Inquiry*, 80(1), 92-113.
- Hargittai, E., y Hsieh, Y. P. (2013). Digital inequality. En W. Dutton (Ed.), *The oxford handbook of internet studies* (p. 130-151). Oxford University Press.
- Hargittai, E., y Litt, E. (2011). The tweet smell of celebrity success: explaining variation in Twitter adoption among a diverse group of young adults. *New Media and Society*, 13(5), 824-842.
- Hargittai, E., y Litt, E. (2012). Becoming a tweep. *Information, Communication & Society*, 15(5), 680-702.
- Harvey, D. (1989). From managerialism to entrepreneurialism: The transformation in urban governance in late capitalism. *Geografiska Annaler: Series B, Human Geography*, 71(1), 3-17. Descargado de <https://doi.org/10.1080/04353684.1989.11879583> doi: 10.1080/04353684.1989.11879583

- Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., y Ratti, C. (2014). Geo-located twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260-271. Descargado de <https://doi.org/10.1080/15230406.2014.890072> (PMID: 27019645) doi: 10.1080/15230406.2014.890072
- Herman, D., y Hadden, L. T. (2019). Does the Starbucks effect exist? Searching for a relationship between Starbucks and adjacent rents. *Property Management*, 37(4), 562-578.
- Hernández-Cordero, A. (2015). Gentrificación y desplazamiento: la zona de La Alameda, Ciudad de México. En V. Delgadillo, I. Díaz-Parra, y L. Salinas (Eds.), *Perspectivas del estudio de la gentrificación en México y América Latina* (p. 255-274). UNAM y UNAM posgrado de Urbanismo and Contested Cities and Instituto de Geografía UNAM.
- Hilbert, M. (2010). When cheap, cheap enough to bridge the digital divide? Modeling income related structural challenges of technology diffusion in Latin America. *World Development*, 38(5), 756-770.
- Hollander, J., y Renski, H. (2015). *Measuring urban attitudes Using Twitter: And Exploratory Study* (Working Paper n.º 1). Lincoln Institute of Land Policy.
- Holmes, D. E. (2017a). *Big Data: A very short introduction*. Oxford University Press.
- Holmes, D. E. (2017b). *Big data : a very short introduction*. Oxford: Oxford University Press.
- Huang, C.-C. (2017). *Developing a data mining framework to identify a sense of gentrification through social media data: A case study using instagram posts in salt lake city, utah* (Tesis de Master no publicada). San Diego University, Estados Unidos de América.
- Huang, Y., Guo, D., Kasakoff, A., y Grieve, J. (2016). Understanding U.S. regional linguistic variation with Twitter data analysis. *Computers, Environment and Urban Systems*, 59, 244-255.
- Ilic, L., Sawada, M., y Zarzelli, A. (2019). Deep mapping gentrification in a large Canadian city using deep learning and Google Street View. *PLoS ONE*, 14(3).
- INEGI. (s.f.). *Estado de ánimo de los tuiteros en México*.
- INEGI. (2003). *Módulo nacional de computación. principales características metodológicas. monaco-2001* (Inf. Téc.). Autor.
- INEGI. (2004). *Disponibilidad y uso de tecnologías de información en los hogares en México. presentación de resultados de las encuestas, 2001, 2002 y 2004* (Inf. Téc.). Autor.
- INEGI. (2006). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2005* (Inf. Téc.). Autor.
- INEGI. (2007). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2006* (Inf. Téc.). Autor.
- INEGI. (2008). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2007* (Inf. Téc.). Autor.
- INEGI. (2009). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2008* (Inf. Téc.). Autor.
- INEGI. (2010a). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2009* (Inf. Téc.). Autor.
- INEGI. (2010b). *Nota técnica. estratificación multivariada* (Inf. Téc.). Autor.
- INEGI. (2011). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2010* (Inf. Téc.). Autor.

- INEGI. (2012). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2011* (Inf. Téc.). Autor.
- INEGI. (2013). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2012* (Inf. Téc.). Autor.
- INEGI. (2014a). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2013* (Inf. Téc.). Autor.
- INEGI. (2014b). *Módulo sobre disponibilidad y uso de tecnologías de información en los hogares, 2014* (Inf. Téc.). Autor.
- INEGI. (2015). *Estadísticas sobre disponibilidad y uso de tecnología de información y comunicaciones en los hogares 2015* (Inf. Téc.). Autor.
- INEGI. (2016a). *Encuesta nacional sobre disponibilidad y uso de tecnologías de la información en hogares, 2016. diseño muestral* (Inf. Téc.). Autor.
- INEGI. (2016b). *Encuesta nacional sobre disponibilidad y uso de tecnologías de la información en hogares, 2016. nota técnica* (Inf. Téc.). Autor.
- Ioannides, D., Röslmaier, M., y van der Zee, E. (2018). Airbnb as an instigator of ‘tourism bubble’ expansion in Utrecht’s Lombok neighbourhood. *Tourism Geographies*, 0(0), 1-19.
- ITU. (2018). *Measuring the information society report* (Reporte n.º Volumen 1). Ginebra: International Telecommunication Union.
- Jaggi, H. S., y Kadam, S. (2016). Integration of spark framework in supply chain management. *Procedia Computer Science*, 79, 1013-1020.
- Janoschka, M. (2016). Gentrificación, desplazamiento, desposesión: procesos urbanos claves en América Latina. *Revista INVI*, 31(88), 27-71.
- Janoschka, M., Sequera, J., y Salinas, L. (2014). Gentrificación en España y América Latina. Un diálogo crítico. *Revista de Geografía Norte Grande*, 58, 7-40.
- Jansson, A. (2019). The mutual shaping of geomedial and gentrification: The case of alternative tourism apps. *Communication and the Public*, 4(2), 166-181.
- Jayne, M. (2006). *Cities and consumption*. London New York, N.Y: London : New York, N.Y. : Routledge, 2006.
- Jiao, J., Holmes, M., y Griffin, G. P. (2018). Revisiting image of the city in cyberspace: Analysis of spatial twitter messages during a special event. *Journal of Urban Technology*, 25(3), 65-82.
- Jurafsky, D., y Martin, J. H. (2018). *Speech and language processing an introduction to natural language processing, computational linguistics, and speech recognition* (3ª [Borrador] ed.). Stanford University.
- Kelley, M. J. (2014). The semantic production of space: pervasive computing and the urban landscape. *Environment and Planning A*, 46, 837 – 851.
- Kitchin, R. (2014a). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1).
- Kitchin, R. (2014b). *The Data Revolution*. Los Angeles, California: SAGE Publications.
- Kitchin, R., y Dodge, M. (2011). *Code/space: software and everyday life*. Cambridge, Mass: MIT Press.
- Kitchin, R., y McArdle, G. (2016). What makes big data, big data? exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1).
- Kitchin, R., y Perng, S.-Y. (2016). Code and the City. Introduction. En R. Kitchin y S.-Y. Perng (Eds.), *Code and the City* (p. 1-12). Routledge.
- Krause, A. E., North, A. C., y Heritage, B. (2018). Musician interaction via social networking sites: Celebrity attitudes, attachment, and their correlates. *Music & Science*, 1, 1-11.

- Krijnen, M. (2018). Gentrification and the creation and formation of rent gaps. *City*, 22(3), 437-446.
- Kuttan, A., y Peters, L. (2003). *From digital divide to digital opportunity*. Scarecrow Education.
- Kwartler, T. (2017). *Text Mining in Practice with R*. Reino Unido: WILEY.
- Lahuerta-Otero, E., y Cordero-Gutiérrez, R. (2016). Looking for the perfect tweet. The use of data mining techniques to find influencers on Twitter. *Computers in Human Behavior*, 64, 575 - 583. Descargado de <http://www.sciencedirect.com/science/article/pii/S0747563216305258> doi: <https://doi.org/10.1016/j.chb.2016.07.035>
- Lai, J., Cheng, T., y Lansley, G. (2017). Improved targeted outdoor advertising based on geotagged social media data. *Annals of GIS*, 23(4), 237-250.
- Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., y Dyer, C. (2016). Neural architectures for named entity recognition. En *Proceedings of the 2016 conference of the north American chapter of the association for computational linguistics: Human language technologies* (p. 260-270). San Diego, California: Association for Computational Linguistics.
- Lansley, G., y Longley, P. A. (2016). The geography of Twitter topics in London. *Computers, Environment and Urban Systems*, 58, 85-96.
- Lara, L. F. (2015). *Temas del español contemporáneo : cuatro conferencias en el colegio nacional*. México, D.F. : El Colegio de México, Centro de Estudios Lingüísticos y Literarios : El Colegio Nacional, 2015.
- la Selva, A. R. A.-D. (2012). *Brecha e inclusión digital en México: hacia una propuesta de políticas públicas*. Universidad Nacional Autónoma de México.
- Lassen, D. S., y Brown, A. R. (2011). Twitter. the electoral connection? *Social Science Computer Review*, 29(4), 419-436.
- Laudano, C. (2017). Movilizaciones #niunamenos y #vivasnosqueremos en Argentina. Entre el activismo digital y #elfeminismolohizo. En *Seminário internacional fazendo gênero 11 & 13th women's worlds congress*.
- Lee, D. (2016). How Airbnb Short-Term Rentals Exacerbate Los Angeles's Affordable Housing Crisis: Analysis and Policy Recommendations. *Harvard Law & Policy Review*, 10, 229-253.
- Lees, L. (2012). The geography of gentrification: Thinking through comparative urbanism. *Progress in Human Geography*, 36(2), 155-171. doi: 10.1177/0309132511412998
- Lees, L., Shin, H. B., y López-Morales, E. (2015). Conclusion: global gentrifications. En L. Lees, H. B. Shin, y E. López-Morales (Eds.), *Global gentrifications. uneven development and displacement* (p. 441-452). Gran Bretaña: Policy Press.
- Lees, L., Shin, H. B., y López-Morales, E. (2016). *Planetary gentrification*. Cambridge, Reino Unido y Malden, Estados Unidos: Polity Press.
- Lefebvre, H. (1991). *The production of space*. Oxford, UK : Cambridge, Massachusetts, USA : Blackwell, 1991, ©1991.
- Lefebvre, H. (2004). *Rhythmanalysis. space, time and everyday life*. London New York, N.Y: Continuum.
- Leszczynski, A. (2019). Spatialities. En J. Ash, R. Kitchin, y A. Leszczynski (Eds.), *Digital geographies* (p. 13-23). United Kingdom: Sage.
- Lewis, K. (2015). Three fallacies of digital footprints. *Big Data & Society*, 2(2). doi: 10.1177/2053951715602496

- Lin, K.-Y., y Lu, H.-P. (2011). Why people use social networking sites: an empirical study integrating network externalities and motivation theory. *Computers in Human Behavior*, 27, 1152-1161.
- Liu, B. (2015). *Sentiment Analysis. Mining opinions, sentiments, and emotions*. Reino Unido: Cambridge University Press.
- Locke, T. (2004). *Critical discourse analysis*. London New York, N.Y: London : New York, N.Y. : Continuum, c2004.
- Loges, W., y Jung, J.-Y. (2001). Exploring the digital divide. Internet connectedness and age. *Communication Research*, 28(4), 536-562.
- Long, Y., y Liu, L. (2016, octubre). Transformations of urban studies and planning in the big/open data era: a review. *International Journal of Image and Data Fusion*, 7(4), 295-308.
- López-Morales, E. (2016, 11). Acerca de una gentrificación “planetaria”, políticamente útil. *Revista INVI*, 31, 217 - 240.
- López-Morales, E., Shin, H. B., y Lees, L. (2016). Latin american gentrifications. *Urban Geography*, 37(8), 1091-1108. doi: 10.1080/02723638.2016.1200335
- Lutz, C., y Newlands, G. (2018). Consumer segmentation within the sharing economy: The case of Airbnb. *Journal of Business Research*, 88, 187-196.
- Lyons, K., y Rodríguez-Ordóñez, I. (2017). Quantifying the linguistic landscape: A study of spanish-english variation in pilsen, chicago. *Spanish in Context*, 14(3), 329-362.
- Mancera, A., y Pano, A. (2013). *El discurso político en Twitter. Análisis de mensajes que trinan*. Barcelona: Anthropos.
- Marengo, D., Giannotta, F., y Settanni, M. (2017). Assessing personality using emoji: An exploratory study. *Personality and Individual Differences*, 112, 74 - 78. doi: <https://doi.org/10.1016/j.paid.2017.02.037>
- Marín-Cortés, A. F., y Quintero-Jurado, J. M. (2018). Confianza en el proceso de paz en Colombia en Twitter. *Revista Mexicana de Sociología*, 80(1), 115-137.
- Mariscal, J. (2005). Digital divide in a developing country. *Telecommunications Policy*, 29, 409-428.
- Mariscal, J. (2009a). *Inclusión de jóvenes al mundo laboral vía capacitación en tic: el papel de ong en cuatro centros comunitarios* (Documento de trabajo n.º 234). CIDE.
- Mariscal, J. (2009b). *Oportunidades móviles: Pobreza y acceso telefónico en Latinoamérica y el Caribe. El caso de México* (Documento de trabajo n.º 222). CIDE.
- Mariscal, J., Bonina, C., y Luna, J. (2009). Los nuevos escenarios de mercado en América Latina. En H. Galperin y J. Mariscal (Eds.), *Pobreza digital. Perspectivas de América Latina y el Caribe* (p. 85-116). CIDE.
- Mariscal, J., Gil-García, R., y Aldama-Nalda, A. (2011). Policies on access to information technologies: The case of e-Mexico. *Information Technologies and International Development*, 7(2), 1-16.
- Mariscal, J., y Martínez, M. A. (2013). *La vida informacional de los marginados: un estudio sobre acceso digital en tres localidades mexicanas* (Documento de trabajo n.º 281). CIDE.
- Martineau, C. (2016). *Pricey mezcal, 'mother' of tequila, steals hipster crown*.
- McDonald, J. (1987). The identification of urban employment subcenters. *Journal of Urban Economics*, 21(2), 242-258.
- Médéric Gasquet-Cyrus. (2015). Changements urbains et conflits sociolinguistiques: l'impact de la gentrification sur le français de Marseille. *International Journal of the Sociology of*

- Language*, 2015(235), 77-101.
- Mekler, P. (2018). *Verificado19s: pipeline de datos*. Descargado April 4th, 2019, de <https://es.slideshare.net/RevistaSG/verificado19s-pipeline-de-datos-paco-mekler>
- Méndez, E. (1996). Cambios de los principales elementos de la estructura urbana de una ciudad fronteriza: Mexicali b.c. *Sociedad y Region*, 3(12), 3-27.
- Mercado-Celis, A. (2016). Distritos creativos en la Ciudad de México en la segunda década del siglo xxi. *Territorios*, 30, 183-213.
- Mercado-Celis, A. (2018). *Gobernanza de la economía nocturna en la ciudad de México*.
- Mermet, A.-C. (2017). Airbnb and tourism gentrification. critical insights from the exploratory analysis of the ‘airbnb syndrome’ in reykjavík. En M. Gravari-Barbas y S. Guinand (Eds.), *Tourism and gentrification in contemporary metropolises. international perspectives*. Routledge.
- Minner, J. S., y Shi, X. (2017). Churn and change along commercial strips: Spatial analysis of patterns in remodelling activity and landscapes of local business. *Urban Studies*, 54(16), 3655-3680. doi: 10.1177/0042098016684274
- Mislove, A., Lehemann, S., Ahn, Y.-Y., Onnela, J.-P., y Rosenquist, J. N. (2011). Understanding the demographics of twitter users. En *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (pp. 554–557). AAAI Press.
- Mitchell, L., Frank, M., Harris, K. D., Dodds, P. S., y Danforth, C. (2013). The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, 8(5).
- Mørsted, B., Sapieżyński, P., Ferrara, E., y Lehmann, S. (2017). Evidence of complex contagion of information in social media: An experiment using twitter bots. *PloS one*, 12(9).
- Mundial, B. (2016). *Dividendos digitales* (Inf. Téc.). Autor.
- Murthy, D. (2012). Towards a sociological understandig of social media: Theorizing Twitter. *Sociology*, 46(6), 1059-1073.
- Nadkarni, A., y Hofmann, S. (2012). Why do people use facebook? *Personality and Individual Differences*, 52, 243-249.
- Neuhaus, F. (2015). *Emergent Spatio-temporal Dimensions of the City. Habitus and Urban Rhythms* (University College London, Ed.). Reino Unido: Springer.
- Nishida, T., Pick, J., y Sarkar, A. (2014). Japan’s prefectural digital divide: A multivariate and spatial analysis. *Telecommunications Policy*, 38(11), 992-1010.
- Nishimura, N., Sukegawa, N., Takano, Y., y Iwanaga, J. (2018). A latent-class model for estimating product-choice probabilities from clickstream data. *Information Sciences*, 429(C), 406–420. doi: 10.1016/j.ins.2017.11.014
- OCDE. (2001). *Understanding the digital divide* (Inf. Téc.). Autor.
- O’Sullivan, A. (2012). *Urban economics*. 8ª edición: McGraw-Hill.
- Oxford, D. (s.f.). *Api*. Descargado 7 de mayo de 2018, de <https://www.oxfordlearnersdictionaries.com/us/definition/english/api?q=API>
- Palmer, D. D. (2010). Text preprocessing. En N. Indurkha y F. J. Damareau (Eds.), *Handbook of Natural Language processing* (2a ed., p. 9-30). Estados Unidos de América: CRC Press.
- Papen, U. (2012). Commercial discourses, gentrification and citizens’ protest: The linguistic landscape of prenzlauer berg, berlin1 uta papen. *Journal of Sociolinguistics*, 16(1), 56–80.

- Parente, R., Geleilate, J.-M., y Rong, K. (2018). The sharing economy globalization phenomenon: A research agenda. *Journal of International Management*, 24, 52-64.
- Parnelee, J., y Bichard, S. (2012). *Politics and the Twitter revolution: how tweets influence the relationship between political leaders and the public*. Lexington Books.
- Pasquale, F. (2016). *Black box society : the secret algorithms that control money and information*. Estados Unidos de América: Harvard University Press.
- Penney, J., y Dadas, C. (2014). (Re)Tweeting in the service of protest: Digital composition and circulation in the Occupy Wall Street movement. *New Media & Society*, 16(1), 74-90.
- Pfeffer, J., Mayer, K., y Morstatter, F. (2018). Tampering with Twitter's Sample API. *EPJ Data Science*, 7(50), 1-21. doi: 10.1140/epjds/s13688-018-0178-0
- Pick, J., y Sarkar, A. (2016). Theories of the digital divide: Critical comparison. En *IEEE Computer Society* (p. 3888-3897).
- Poorthuis, A., Zook, M., Shelton, T., Graham, M., y Stephens, M. (2014). Using geotagged digital social data in geographic research. En N. Clifford, S. French, M. Cope, y S. Gillespie (Eds.), *Key methods in geography*.
- Preis, B., Janakiraman, A., Bob, A., y Steil, J. (0). Mapping gentrification and displacement pressure: An exploration of four distinct methodologies. *Urban Studies*, 0(0).
- Rahimi, S., Andris, C., y Liu, X. (2017). Using yelp to find romance in the city: A case of restaurants in four cities. En *Proceedings of the 3rd acm sigspatial workshop on smart cities and urban analytics* (p. 8). New York, NY, USA: Association for Computing Machinery.
- Ramasubramanian, K., y Singh, A. (2017). *Machine learning using R*. New Delhi, Delhi, India : Apress, [2017], ©2017.
- Rauschnabel, P. A., Rossmann, A., y Tom Dieck, M. C. (2017). An adoption framework for mobile augmented reality games: The case of Pokémon Go. *Computers in Human Behavior*, 76(C), 276-286. doi: 10.1016/j.chb.2017.07.030
- Redfern, P. A. (1997). A new look at gentrification: 1. gentrification and domestic technologies. *Environment and Planning A: Economy and Space*, 29(7), 1275-1296. Descargado de <https://doi.org/10.1068/a291275> doi: 10.1068/a291275
- Reguillo, R. (2019). *Redes y esfera pública*. Descargado 18 de noviembre de 2019, de <https://signalab.mx/2019/11/14/redes-y-esfera-publica/>
- Ritzer, G., y Jurgenson, N. (2010). Production, Consumption, Prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, 10(1), 13-36.
- Roberts, H., Sadler, J., y Chapman, L. (2018). The value of Twitter data for determining the emotional responses of people to urban green spaces: A case study and critical evaluation. *Urban Studies*, February, 1-18.
- Rodino-Colocino, M. (2006). Laboring under the digital divide. *New Media and Society*, 8(3), 487-511.
- Rodríguez, A. (2006). *La brecha digital y sus determinantes*. Universidad Nacional Autónoma de México.
- Rodríguez, R. (2013). *Cambios socio-espaciales en la zona central de la ciudad de México 2000-2010. la colonia doctores: ¿un espacio urbano en proceso de gentrificación?* (Tesis de Master no publicada). Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México, Mexico.
- Rodríguez, César. (2015). Articulación y contrapoder. Los protagonistas del activismo digital en México (2009-2014). En R. W. y José Alberto Sánchez (Ed.), *Redes sociodigitales en México*

- (p. 81-114). Fondo de Cultura Económica y Consejo Nacional para la Cultura y las Artes.
- Rodríguez-Martín, B., y Castillo, C. A. (2017). Carbohydrates and sports practice: a Twitter virtual ethnography. *Nutricion Hospitalaria*, 34(1), 144-153.
- Rosegger, R. (2013). Urban prosumers: Network generated urban spaces. En H. T. Andersen y R. Atkinson (Eds.), *Production and use of urban knowledge. european experiences* (p. 55-80). Springer.
- Sabatini, F., Rasse, A., Cáceres, G., Robles, M. S., y Trebilcock, M. P. (2017). Promotores inmobiliarios, gentrificación y segregación residencial en Santiago de Chile. *Revista Mexicana de Sociología*, 79(2), 229-260.
- Sagolla, D. (2009). *140 caracteres. a style guide for the short form*. Nueva Jersey: WILEY.
- Said-Hung, E., y Valencia-Cobos, J. (2017). Twitter y movilización en Venezuela. *Revista Brasileña de Ciencias Sociales*, 32(94), 1-18.
- Salgado-Andrade, E. (2018). Del #SimulacroCDMX al #MéxicoNosNecesita: 19S en Twitter. *Revista Mexicana de Sociología*, 80(especial (septiembre2018)), 121-147.
- Salganik, M. (2018). *Bit by bit. social research in the digital age*. Estados Unidos de América: Princeton University Press.
- Salinas, L. (2013a). Gentrificación en la ciudad latinoamericana. El caso de Buenos Aires y la Ciudad de México. *GeoGraphos*, 4(44), 281-204.
- Salinas, L. (2013b). La gentrificación de la colonia condesa, ciudad de México. aporte para una discusión desde Latinoamérica. *Revista Geográfica de América Central*, 2(51), 145-167.
- Sánchez, N., Ortega, O., y Vall-llovera, M. (2012). Romper la brecha digital de género. Factores implicados en la opción por una carrera tecnológica. *Athenea Digital*, 12(3), 115-128.
- Santos, J. A., y Sequera, J. (2018). Expansión de la ciudad turística y nuevas resistencias el caso de Airbnb en Madrid. *Empiria. Revista de Metodología de Ciencias Sociales*, 41, 15-32.
- Sarkar, D. (2016). *Text analytics with Python : a practical real-world approach to gaining actionable insights from your data*. New York, NY: Apress.
- Sassen, S. (1991). *The global city: New York, London, Tokyo*. Princeton, New Jersey: Princeton University.
- Sassen, S. (2001). *The global city*. Princeton, NJ.: Princeton Univ. Press.
- Sevin, H. E. (2014). Understanding cities through city brands: City branding as a social and semantic network. *Cities*, 38, 47-56.
- Shamsuddin, S., y Srinivasan, S. (2020). Just smart or just and smart cities? assessing the literature on housing and information and communication technology. *Housing Policy Debate*, 0(0), 1-24. Descargado de <https://doi.org/10.1080/10511482.2020.1719181> doi: 10.1080/10511482.2020.1719181
- Shaw, J., y Graham, M. (2017). An informational right to the city? code, content, control, and the urbanization of information. *Antipode*, 0(0), 1-21.
- Shelton, T. (2015). *Toward a relational socio-spatial analysis of geotagged social media data* (Tesis Doctoral no publicada). Clark University.
- Shelton, T. (2017). The urban geographical imagination in the age of big data. *Big Data & Society*, 4(1), 1-14.
- Shelton, T., Poorthuis, A., y Zook, M. (2015). Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information. *Landscape and Urban Planning*, 142, 198-211.
- Slater, T. (2009). Missing Marcuse: On gentrification and displacement. *City*, 13(2-3), 292-311.

- Slater, T. (2012). Missing marcuse: on gentrification and displacement. En N. Brenner, P. Marcuse, y M. Mayer (Eds.), *Cities for people nor for profit. critical urban theory and the right to the city* (p. 171-196). Routledge.
- Sloan, L., y Morgan, J. (2015). Who tweets with their location? understanding the relationship between demographic characteristic and the use of geoservices an geotagging on Twitter. *PLoS ONE*, 6, 1-15.
- Sloan, L., Morgan, J., Burnap, P., y Williams, M. (2015). Who tweets? deriving demographic characteristics of age, occupation and social class from Twitter user Meta-Data. *PLoS ONE*, 10(3), 1-20.
- Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P., y Rana, O. (2013). Knowing the Tweeters: Deriving sociologically relevant demographics from Twitter. *Sociological Research Online*, 18(3), 1-11.
- Smith, A. (2010). *8% of online Americans use Twitter* (Inf. Téc.). Pew Research Center.
- Smith, A., y Brenner, J. (2012). *Twitter use 2012* (Inf. Téc.). Pew Research Center.
- Smith, N. (1979). Toward a theory of gentrification a back to the city movement by capital, not people. *Journal of the American Planning Association*, 45(4), 538-548. doi: 10.1080/01944367908977002
- Smith, N. (1982). Gentrification and uneven development. *Economic Geography*, 58(2), 139-155. doi: 10.2307/143793
- Smith, N. (1990). *Uneven development. nature, capital, and the production of space* (Tercera Edición ed.). Estados Unidos de América: The Universtiy of Geogia Press.
- Smith, N. (2002, Julio). New globalism, new urbanism: Gentrification as global urban strategy. *Antipode*, 34(3), 427-450.
- Smith, N. (2008). *Uneven development. nature, capital, and the production of space* (3ª ed.). Georgia: The University of Georgia Press.
- Sobrinho, J. (2016). Estudio de la gentrificación en la ciudad de méxico. *Coyuntura Demográfica*(9), 69-75.
- Stehle, S. K. (2017). *Mapping Semantic and Spatial Mediascapes in the Catalanian Independence Movement: Geopolitics, Sports, and Black Boxes* (Tesis Doctoral no publicada). The Pennsylvania State Univesity, The Graduate School College of Earth and Mineral Sciences, Pennsylvania, Estados Unidos.
- Steiger, E., Westerholt, R., Resch, B., y Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Computers, Environment and Urban Systems*, 54, 255-265.
- Stella, M., Ferrara, E., y De Domenico, M. (2018). Bots increase exposure to negative and inflammatory content in online social systems. *Proceedings of the National Academy of Sciences*, 115(49), 12435–12440. doi: 10.1073/pnas.1803470115
- Sui, D., Goodchild, M., y Elwood, S. (2013). Volunteered geographic information, the exaflood, and the growing digital divide. En D. Sui, S. Elwood, y M. Goodchild (Eds.), *Crowdsourcing geographic knowledge: Volunteered geographic information (vgi) in theory and practice* (p. 1-12). Dordrecht: Springer Netherlands.
- Takhteyev, Y., Gruzd, A., y Wellman, B. (2012). Geography of twitter networks. *Social Networks*, 34, 73-81.
- Thakuriah, P., Tilahun, N. Y., y Zellner, M. (2017). Big data and urban informatics: Innovations and challenges to urban planning and knowledge discovery. En P. Thakuriah, N. Y. Tilahun,

- y M. Zellner (Eds.), *Seeing cities through Big Data. Research, Methods and Applications in Urban Informatics* (p. 11-49). Glasgow y Chicago: Springer.
- Thelwall, M., Buckley, K., y Paltoglou, G. (2011). Sentiment in twitter events. *Journal of the American Society for Information Science and Technology*, 62(2), 406-418.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., y Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558. Descargado de <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21416> doi: 10.1002/asi.21416
- Times, F. (2017). *Airbnb marks first full year of profitability in 2017*. Descargado 3 de mayo de 2018, de <https://www.ft.com/content/96215e16-0201-11e8-9650-9c0ad2d7c5b5>
- Toudert, D. (2000). La www en la frontera norte mexicana: Hacia el descubrimiento de una dimension desconocida. *Revista Frontera Norte*, 13(24), 7-33.
- Toudert, D. (2001). Los portales de las televisoras en México: que continua el espectáculo en el ciberespacio. *Revista del Observatorio de la Cibersociedad*, 2.
- Toudert, D. (2014). Evolución de la polarización territorial de la producción web: ¿para cuándo la esperada descentralización? *Economía, Sociedad y Territorio*, 14(45), 549-580.
- Toudert, D. (2015). Brecha Digital y marginación socioterritorial: el caso de México. En C. Garrocho-Rangel y G. Buzai (Eds.), *Geografía aplicada en iberoamérica. avances, retos y perspectivas* (cap. 343-370). El Colegio Mexiquense.
- Toudert, D. (2016a). Aislamiento geográfico y la disponibilidad de la internet en las escuelas mexicanas. *Revista Electrónica Nova Scientia*, 17(8).
- Toudert, D. (2016b). Teoría del recurso y la apropiación: un acercamiento empírico a partir de las etapas del modelo de acceso digital en México. *Acta Universitaria*, 26(4), 79-90.
- Trejo, R. (2015). México enredado: Auge ligereza y limitaciones en el uso político de las redes sociodigitales. En R. W. y José Alberto Sánchez (Ed.), *Redes sociodigitales en México* (p. 17-39). Fondo de Cultura Económica y Consejo Nacional para la Cultura y las Artes.
- UBER. (2018). *Ciudades con presencia de uber*. Descargado 3 de mayo de 2018, de <https://www.uber.com/es-MX/CITIES/>
- UIT. (2015). *Statistics*. Descargado 1 de marzo 2017, de <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>
- Uitermark, J., y Bosker, T. (2014). Wither the 'Undivided City'? An Assessment of State-Sponsored Gentrification in Amsterdam. *Tijdschrift voor economische en sociale geografie*, 105(2), 221-230. Descargado de <https://onlinelibrary.wiley.com/doi/abs/10.1111/tesg.12072> doi: 10.1111/tesg.12072
- Uribe, J. P. (2019, julio 24). *Ciudad de México. La Catedral del pulque [audio podcast]*. Descargado de <https://open.spotify.com/show/02CtHFN3erzaohUAD2PCAg>
- Uzzi, M. T. R. S. B. S. B. (2010). Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology*, 36, 91-115.
- Vandenbroucke, M. (2016). Socio-economic stratification of english in globalized landscapes: A market-oriented perspective. *Journal of Sociolinguistics*, 20(1), 86-108.
- van Deursen, A., y van Dijk, J. (2010). Internet skills and the digital divide. *New Media and Society*, 13(6), 893-911.

- van Dijk, J. (2005). *The deeping digital divide. inequality in the information society*. SAGE Publications.
- van Dijk, J. (2012). The evolution of the digital divide. the digital divide turns to inequality of skills and usage. En I. Press (Ed.), *Digital enlightenment yearbook 2010* (p. 57-75). Washington D.C.: Jaques Bus and Malcolm Crompton and Mirelle Hildebrandt and George Metakides.
- Venkatesh, V., Morris, M. G., Davis, G. B., y Davis, F. D. (2003). User acceptance of information technology: towards a unified view. *MIS Quartely*, 27(3), 425-478.
- Vergara-Constela, C., y Casellas, A. (2016). Políticas estatales y transformación urbana: ¿hacia un proceso de gentrificación en Valparaíso, Chile? *EURE*, 42(126), 123-144.
- Vith, T. A., y Mössner, S. (2017). Contesting sustainable transportation: bicycle mobility in Boston and beyond. *Journal of Geographical Society of Berlin*, 148(4), 229-237.
- Vosoughi, S., Roy, D., y Aral, S. (2018). The spread of true and false news online. *Science*, 359, 1146-1151.
- Wachsmuth, D., y Weisler, A. (2017). *Airbnb and the rent gap: Gentrification through the sharing economy* [Documento de trabajo]. School of Urban Planning, McGill University.
- Walker, Z., McMahon, D., y Rosenblatt, K. (2017). Pokemon go, augmented reality, and universal design for learning. *Sage Open*.
- Wettering, F. (2001). The internet and the spy business. *International Journal of Intelligence and Counterintelligence*, 14, 342-365.
- Winocur-Iparraguirre, R., y Sánchez-Martínez, J. A. (2015). Introducción. En R. Winocur-Iparraguirre y J. A. Sánchez-Martínez (Eds.), *Redes sociodigitales en México* (p. 9-16). Fondo de Cultura Económica y Consejo Nacional para la Cultura y las Artes.
- Wu, W., Wang, J., y Dai, T. (2016). The geography of cultural ties and human mobility: Big data in urban contexts. *Annals of the American Association of Geographers*, 106(3), 612-630.
- Wyly, E. (2018). The evolving state of gentrification. *Tijdschrift voor economische en sociale geografie*, 110. doi: 10.1111/tesg.12333
- Yodanis, C. (2006). A place in town: doing class in a coffee shop.(author abstract). *Journal of Contemporary Ethnography*, 35(3), 341.
- Yrigoy, I. (2016). The impact of airbnb in the urban arena: towards a tourism-led gentrification? the case-study of palma old quarter (mallorca, spain). En *From contested cities to global urban justice*. Madrid.
- Zhang, X. (2013). Income disparity and digital divide: The internet consumption model and cross-country empirical research. *Telecommunications Policy*, 37, 515-529.
- Zhong, C., Zeng, S., Tu, W., y Yoshida, M. (2018). Profiling the Spatial Structure of London: From Individual Tweets to Aggregated Functional Zones. *ISPRS International Journal of Geo-Information*, 7(10), 386-400.
- Zhou, X., Xu, C., y Kimmons, B. (2015). Detecting tourism destinations using scalable geospatial analysis based on cloud computing platform. *Computers, Environment and Urban Systems*, 54, 144-153.
- Zukin, S. (1987). Gentrification: culture and capital in the urban core. *Annual Review of Sociology*, 13, 129-147.
- Zukin, S. (1991). *Landscapes of power : from Detroit to Disney World*. Berkeley, Calif: Berkeley, Calif. : University of California, 1991.
- Zukin, S. (1998). Urban lifestyles: diversity and standardisation in spaces of consumption. *Urban Studies*, 35(5-6), 825-839.

- Zukin, S. (2020). *The innovation complex. cities, tech, and the new economy*. Nueva York: Oxford University Press.
- Zukin, S., Lindeman, S., y Hurson, L. (2017). The omnivore's neighborhood? online restaurant reviews, race, and gentrification. *Journal of Consumer Culture*, 17(3), 459–479.
- Zukin, S., Trujillo, V., Frase, P., Jackson, D., Recuber, T., y Wlaker, A. (2009). New retail capital and neighborhood change: boutiques and gentrification in new york city. *City and Community*, 8(1), 47-64.

# Capítulo A

## Anexo

Listing A.1: Código de Logstash para descarga de tweets mediante API gratuita de Twitter

```
\label{}

input{
  twitter{
    consumer_key=>"Primera_clave"
    consumer_secret => "Segunda_clave"
    oauth_token => "Tercera_clave"
    oauth_token_secret => "cuarta_clave"
    locations => "-99.654,18.922,-98.609,20.067"
    full_tweet => true
  }}

filter{ }

output{
  stdout{
    codec => dots
  }
  file{
    codec => json_lines
    path => "/home/json/twitter-%{+YYYY.MM.dd}.json"}}}
```

Cuadro A.1: Descripción de clases económicas asociadas a gentrificación

Clave de clase	Descripción	Clave de clase	Descripción
462210	Comercio al por menor en tiendas departamentales	711510	Artistas, escritores y técnicos independientes
463211	Comercio al por menor de ropa, excepto de bebé y lencería	712111	Museos del sector privado
463212	Comercio al por menor de ropa de bebé	712112	Museos del sector público
463213	Comercio al por menor de lencería	713943	Centros de acondicionamiento físico del sector privado
463215	Comercio al por menor de bisutería y accesorios de vestir	713944	Centros de acondicionamiento físico del sector público
463216	Comercio al por menor de ropa de cuero y piel y de otros artículos de estos materiales	722411	Centros nocturnos, discotecas y similares
463310	Comercio al por menor de calzado	722412	Bares, cantinas y similares
465111	Comercio al por menor de artículos de perfumería y cosméticos	722511	Restaurantes con servicio de preparación de alimentos a la carta o de comida corrida
465112	Comercio al por menor de artículos de joyería y relojes	722512	Restaurantes con servicio de preparación de pescados y mariscos
465211	Comercio al por menor de discos y casetes	722513	Restaurantes con servicio de preparación de antojitos
465214	Comercio al por menor de equipo y material fotográfico	722514	Restaurantes con servicio de preparación de tacos y tortas
465215	Comercio al por menor de artículos y aparatos deportivos	722515	Cafeterías, fuentes de sodas, neverías, refresquerías y similares
465216	Comercio al por menor de instrumentos musicales	722516	Restaurantes de autoservicio
465312	Comercio al por menor de libros	722517	Restaurantes con servicio de preparación de pizzas, hamburguesas, hot dogs y pollos rostizados para llevar
711111	Compañías de teatro del sector privado	722518	Restaurantes que preparan otro tipo de alimentos para llevar
711112	Compañías de teatro del sector público	722519	Servicios de preparación de otros alimentos para consumo inmediato
711121	Compañías de danza del sector privado	812110	Salones y clínicas de belleza y peluquerías
711122	Compañías de danza del sector público	812120	Baños públicos

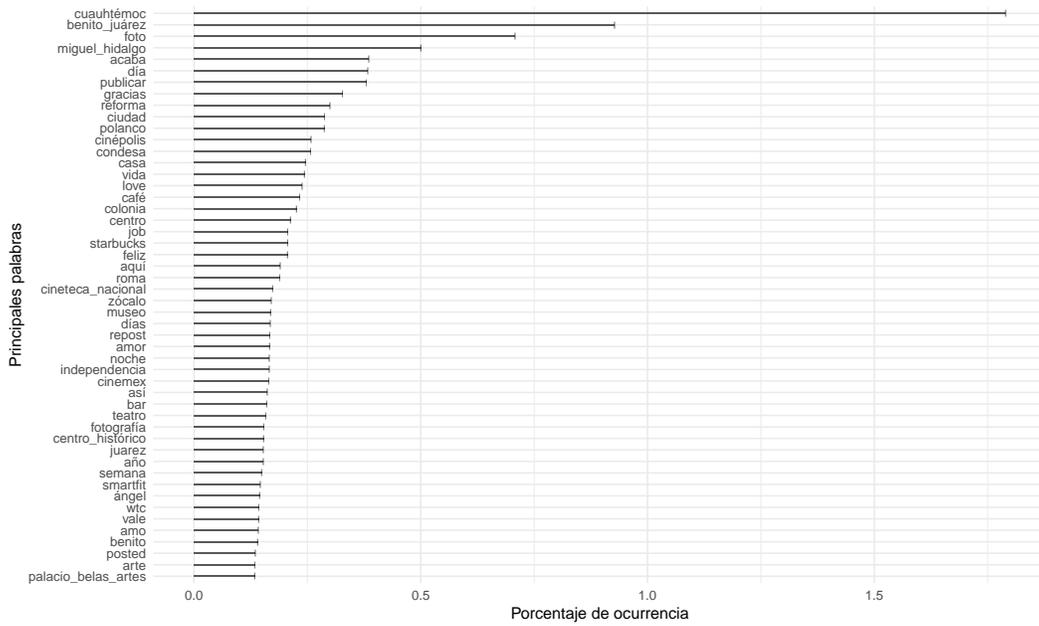
Cuadro A.2: Ejemplo de entidades nombradas en los tweets de la ZMCM

separadas	entidad_nombrada
ciudad de méxico	ciudad_de_méxico
estado de méxico	estado_de_méxico
distrito federal	distrito_federal
benito Juárez	benito_juárez
miguel hidalgo	miguel_hidalgo
mex	méxico
mx	méxico
colonia roma	colonia_roma
colonia condesa	colonia_condesa
colonia Juárez	colonia_juárez
colonia san rafael	colonia_san_rafael
colonia escandón	colonia_escandón
colonia xoco	colonia_xoco
colonia narvarte	colonia_narvarte
colonia polanco	colonia_polanco
plaza antarapolanco	plaza_antarapolanco
plaza san jacinto	plaza_san_jacinto
plaza de toros méxico	plaza_de_toros_méxico
plaza méxico	plaza_méxico
plaza universidad	plaza_universidad
plaza de zocalo	plaza_de_zocalo
plaza polanco	plaza_polanco
plaza condesa	plaza_condesa
p universidad	plaza_universidad

Fuente: Elaboración propia

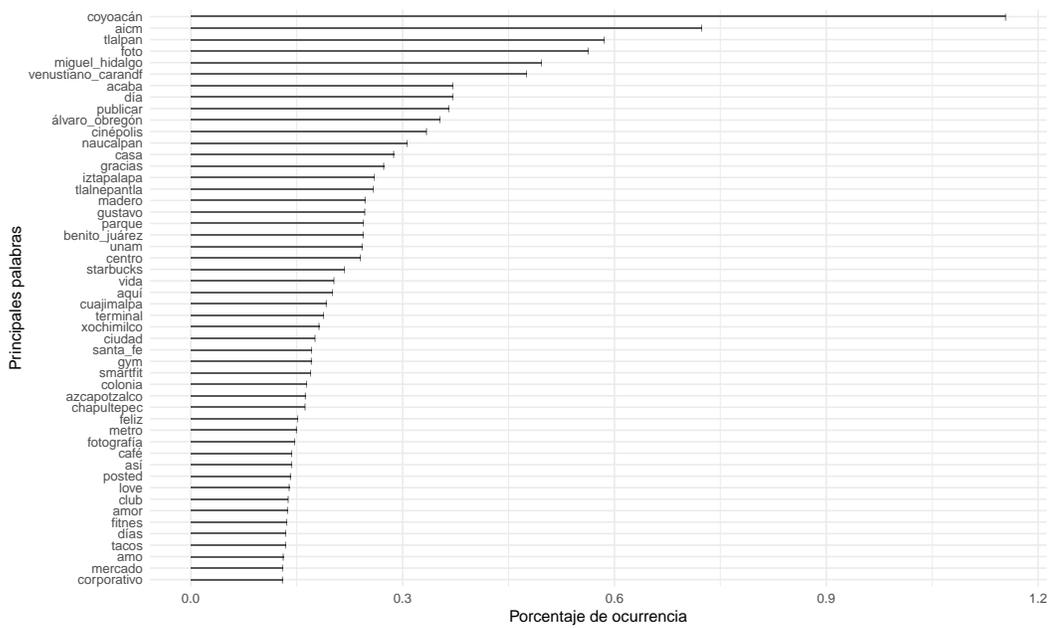
La lista completa se puede consultar en <https://bit.ly/2ZM5aDj>

Figura A.1: Principales palabras de los tweets en colonias gentrificadas(50 palabras)



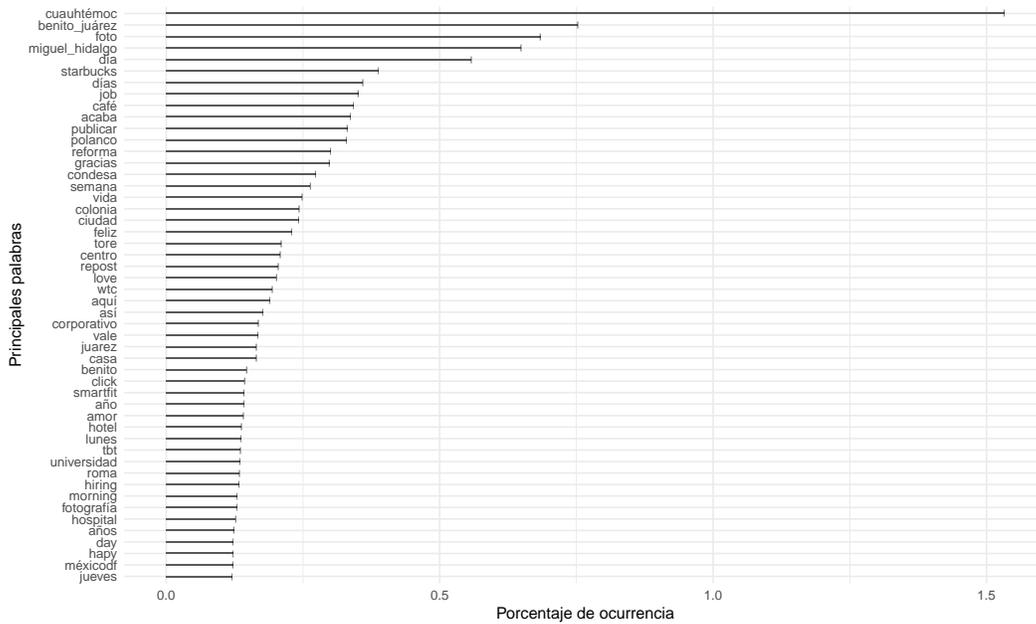
Fuente: Elaboración propia

Figura A.2: Principales palabras de los tweets en colonias no gentrificadas (50 palabras)



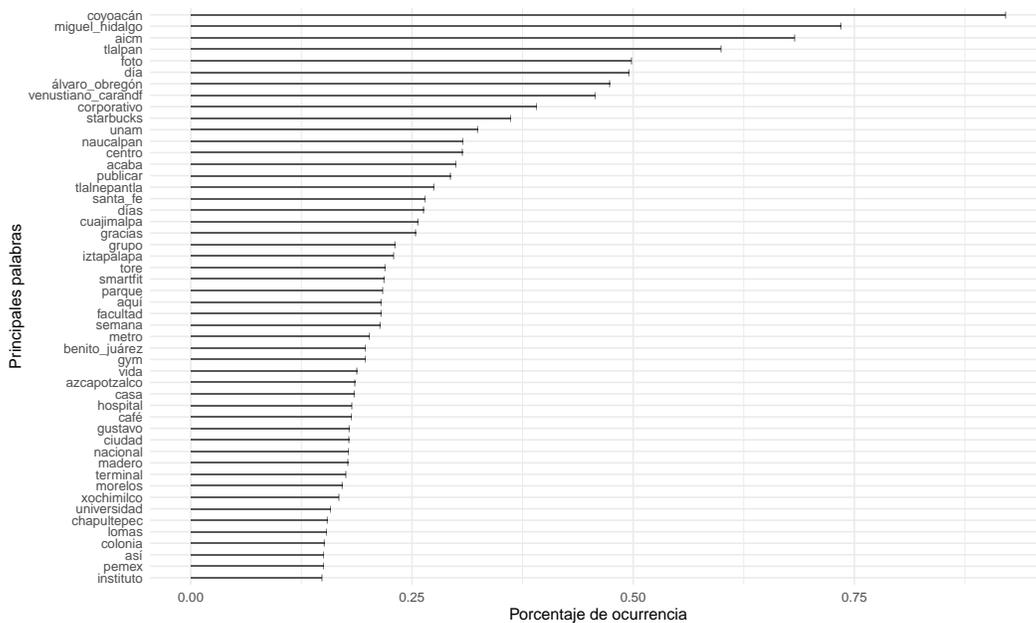
Fuente: Elaboración propia

Figura A.3: Principales palabras de los tweets en colonias gentrificadas (días laborales, 7:00 hrs - 12:59 hrs)



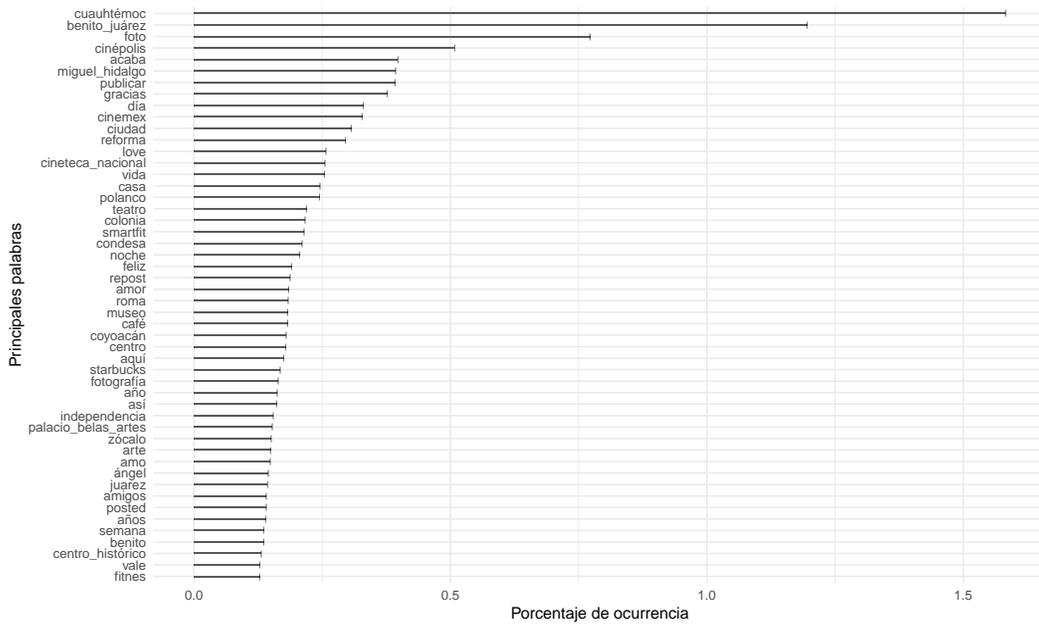
Fuente: Elaboración propia

Figura A.4: Principales palabras de los tweets en colonias no gentrificadas (días laborales, 7:00 hrs - 12:59 hrs)



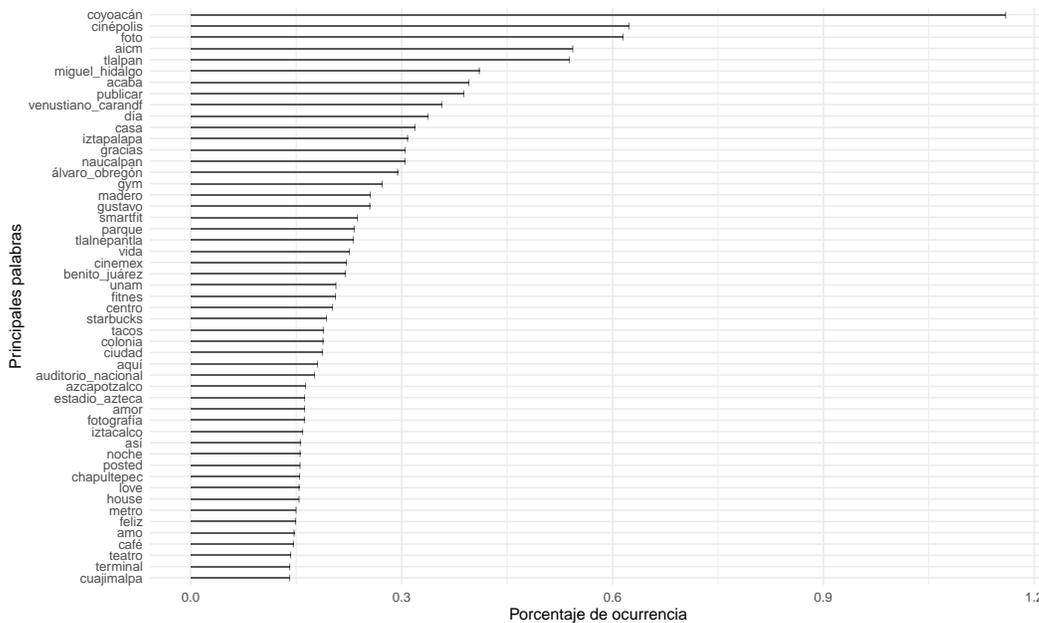
Fuente: Elaboración propia

Figura A.5: Principales palabras de los tweets en colonias gentrificadas (días laborales, 18:00 hrs - 21:59 hrs)



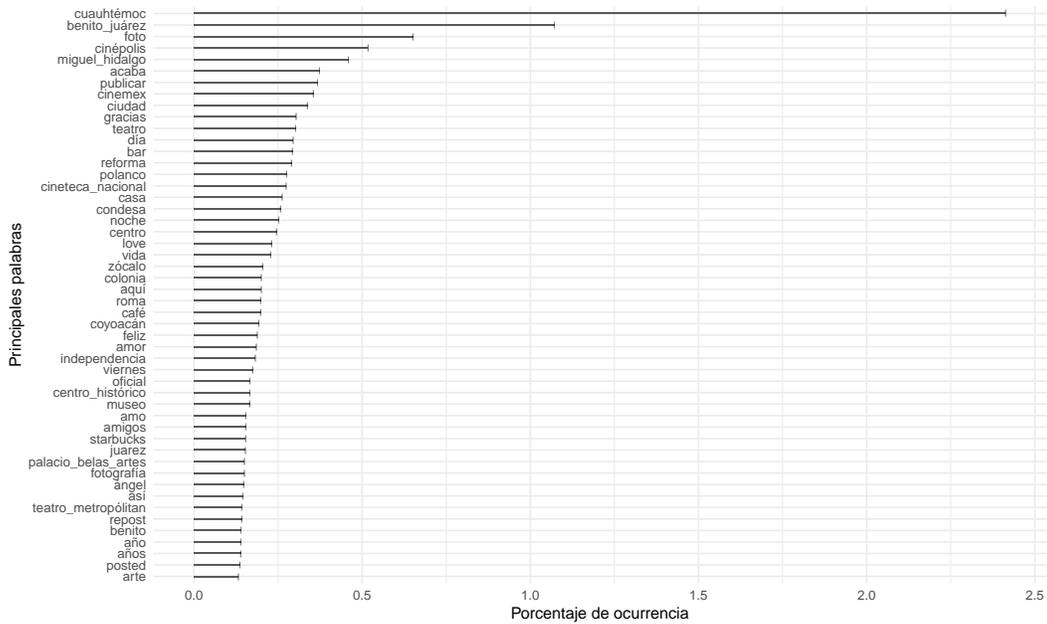
Fuente: Elaboración propia

Figura A.6: Principales palabras de los tweets en colonias no gentrificadas (días laborales, 18:00 hrs - 21:59 hrs)



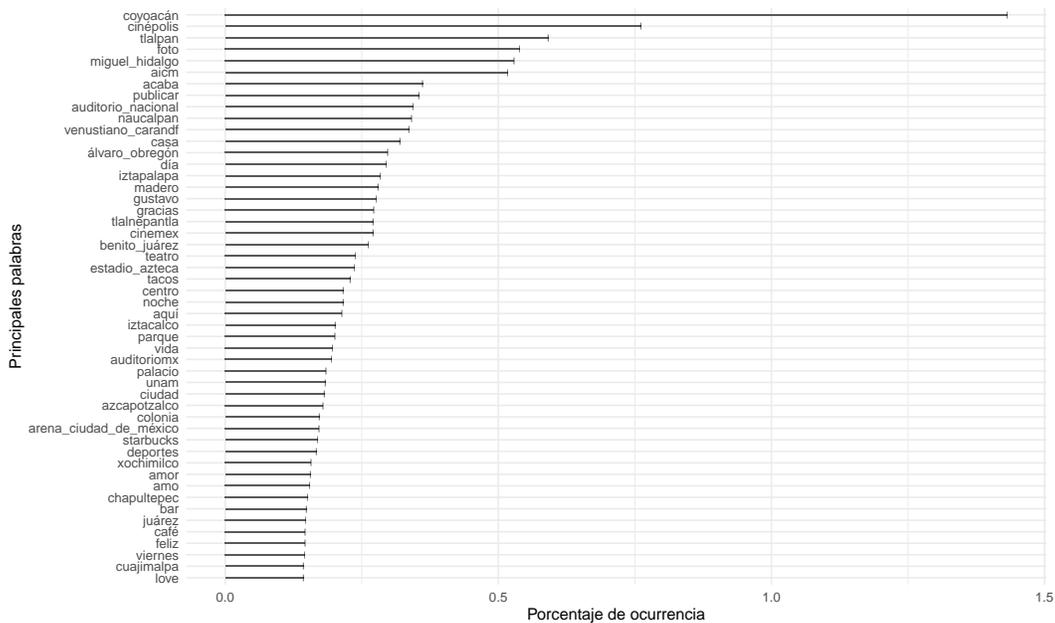
Fuente: Elaboración propia

Figura A.7: Principales palabras de los tweets en colonias gentrificadas (fines de semana, 18:00 hrs - 21:59 hrs)



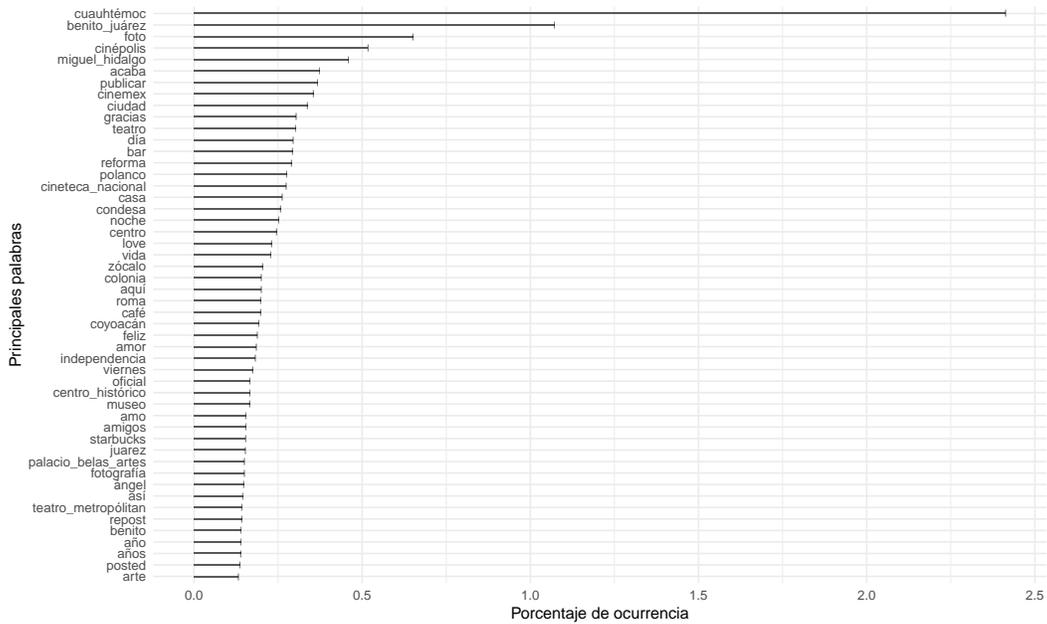
Fuente: Elaboración propia

Figura A.8: Principales palabras de los tweets en colonias no gentrificadas (fines de semana, 18:00 hrs - 21:59 hrs)



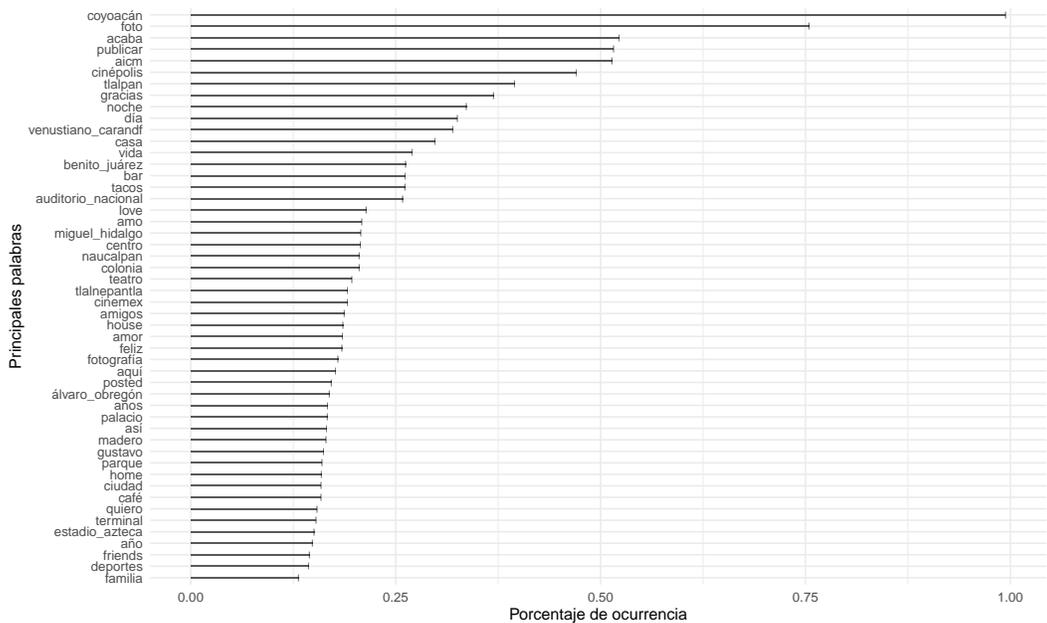
Fuente: Elaboración propia

Figura A.9: Principales palabras de los tweets en colonias gentrificadas (fines de semana, 22:00 hrs - 23:59 hrs)



Fuente: Elaboración propia

Figura A.10: Principales palabras de los tweets en colonias no gentrificadas (fines de semana, 22:00 hrs - 23:59 hrs)



Fuente: Elaboración propia