

BIG DATA Y UNIVERSIDADES: ANÁLISIS DE MOVILIDAD DE LOS ESTUDIANTES UNIVERSITARIOS A PARTIR DE DATOS DE TWITTER

JOAQUÍN OSORIO-ARJONA, JUAN CARLOS GARCÍA-PALOMARES
Departamento de Geografía, Universidad Complutense de Madrid
Plaza de Menéndez Pelayo s/n., 28040, Madrid, España.
joaquoso@ucm.es, jcgarcia@ucm.es

RESUMEN

Este trabajo investiga la movilidad a los campos universitarios en el Área Metropolitana de Madrid a partir de datos geolocalizados de Twitter, aprovechando su alto uso por la población joven. A partir de la identificación de usuarios de Twitter, de sus campus y lugares de residencia, se estiman áreas de influencia de las distintas universidades, y se combinan los datos obtenidos con otras fuentes como ficheros de tiempos de viaje o datos de nivel de renta para analizar tiempos según modo de transporte y tipo de universidad. Los resultados muestran que los estudiantes tienden a residir cerca del campus al que asisten y la tendencia de los estudiantes de universidades privadas a residir en las zonas con mayor nivel de renta.

Palabras clave: Twitter; movilidad; universidades; transportes; modelo gravitacional.

BIG DATA AND UNIVERSITIES: ANALYSIS OF UNIVERSITY STUDENTS MOBILITY BY TWITTER DATA

ABSTRACT

This work investigates university mobility in the Metropolitan Area of Madrid using geolocated Twitter data, taking advantage of its high use by the young population. From the identification of users, their campuses and residences, influence areas of the different universities are estimated, and the obtained data is combined with other sources like time travel files or income level data to analyse the influence of the transport mode, the type of university, or the place of residence in university mobility. The results show that students tend to reside near the campus they attend and the tendency of students from private universities to reside in areas with higher income levels.

Keywords: Twitter; mobility, universities, transports, gravitational model.

1. Introducción

Las universidades son centros de atracción de viajes que requieren de una infraestructura necesaria para soportar grandes volúmenes de viajeros (Whalen *et al.*, 2013). El transporte a los campus está afectado por múltiples causas, que se pueden resumir en factores de localización espacial, factores socioeconómicos de la población universitaria y factores de comportamiento social (Soria-Lara *et al.*, 2017). El análisis de la movilidad universitaria es importante para los gestores de transporte y para las propias universidades. Comprender la movilidad universitaria puede ayudar a crear un modelo sostenible medioambiental y que ayude a disminuir la exclusión social urbana.

Tradicionalmente, los estudios relacionados sobre la movilidad universitaria se han realizado a partir de encuestas como fuente de datos. Aunque estas fuentes de datos tradicionales permiten obtener una información muy rica, son generalmente costosas, lo que repercute en una baja posibilidad de actualización de datos, en que se realicen a partir de muestras pequeñas, y que requieran un tiempo importante de preparación. Sin embargo, la irrupción del Big Data facilita el acceso a datos que permiten complementar la información obtenida en encuestas tradicionales de forma rápida, barata, con muestras grandes, y con un alto detalle espacio-temporal.

Entre las nuevas fuentes de datos, las redes sociales son una de las herramientas con las que se puede conseguir datos masivos y de alta resolución espacio-temporal a tiempo casi real. Entre ellas destaca Twitter por su gran difusión, la facilidad de descarga de datos y la posibilidad de contar con datos georreferenciados de forma gratuita (Osorio Arjona y García-Palomares, 2017). Un aspecto a tener en cuenta respecto a las redes sociales es el sesgo producido por el uso diferencial de las mismas según franjas de edad. Casi el 40 % de usuarios de Twitter tienen entre 18 y 30 años¹. Sin embargo, este sesgo es una ventaja a la hora de estudiar la movilidad universitaria, ya que justo el grupo de edad más representado es en el que se encuentran los estudiantes universitarios.

En este artículo se propone detectar población universitaria a partir de datos obtenidos por Twitter, y valorar hasta qué punto Twitter puede ser una herramienta de utilidad en estas y otras cuestiones relacionadas con la localización y movimientos de sus usuarios. Para ello, se ha trabajado en el área metropolitana de Madrid y el conjunto de su red universitaria, tratando de establecer las áreas de influencia de las distintas universidades madrileñas y estudiar las diferencias según el tipo de universidad. A continuación, se busca estudiar factores de la movilidad a los campus analizando los datos de tiempos de viaje según el modo de transporte, para a continuación realizar un análisis socioeconómico que muestre datos de tiempos de viaje según el tipo de universidad y el nivel adquisitivo de la zona donde residan los estudiantes. Finalmente, se pretende crear un modelo gravitacional con el que comparar un modelo teórico de asignación de estudiantes a las universidades y compararlo con los resultados de la asignación identificada a partir de Twitter. En este caso, partimos de la premisa de que los datos de Twitter se aproximan más a la realidad, que ese modelo teórico, que únicamente considera la proximidad y el tamaño de los campus, lo que permite mostrar que existen otros factores más allá de estos dos en la elección de los centros universitarios.

La utilidad de esta investigación es obtener datos del lugar de residencia de los universitarios. Los datos del lugar de residencia de los estudiantes existen, las universidades lo tienen cuando estos realizan la matrícula, pero a los que con frecuencia no es fácil acceder. A

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

partir de estos datos se puede obtener información complementaria a las encuestas y estudios de movilidad, pero con un coste mucho menor y con la posibilidad de rápida y continua actualización.

El siguiente trabajo está dividido en seis secciones. A continuación de esta introducción, en la sección 2 se realizará un breve estado del arte, en la sección 3 se presentará el área de estudio y los datos utilizados, mientras que la sección 4 revisará la metodología empleada en la investigación. Finalmente, en la sección 5 se analizarán los resultados obtenidos, para establecer una serie de conclusiones en la sección 6.

2. Estado del arte

Prácticamente todos los estudios de movilidad en el ámbito universitario han utilizado encuestas. Hasta hace pocos años, estas encuestas se han realizado in-situ. Tenemos en España ejemplos de estudios realizados a partir de encuestas in-situ con el propósito de describir los patrones de movilidad de los estudiantes, trabajadores y residentes en un campus integrado en un entorno urbano (Lucas-García *et al.*, 2016 en la Universidad de Sevilla; de las Rivas *et al.*, 2010 en la Universidad de Valladolid), o de medir la accesibilidad para poder realizar propuestas de sostenibilidad, como los trabajos realizados en la (Saladié y Jurado, 2015 en la Universitat Rovira i Virgili; Miralles-Guasch y Domene, 2010 en la Universitat Autònoma de Barcelona; Albertos *et al.*, 2008 en la Universitat de València).

Con el desarrollo de internet, estas encuestas presenciales han pasado a realizarse utilizando páginas webs o correos electrónicos con el objetivo de aumentar las muestras y abaratar costes. Los trabajos basados en encuestas por internet se pueden clasificar en tres categorías. Un primer grupo tiene como objetivo elaborar una descripción sobre las características de transporte de los estudiantes universitarios (Moravec *et al.*, 2018 en la Universidad Tecnológica Nacional de Avellaneda (Argentina); Gutiérrez *et al.*, 2016 en la Universidad de Extremadura; Volosin *et al.*, 2013 en la Arizona State University (Estados Unidos); Whalen *et al.*, 2013 en la McMaster University (Hamilton, Canadá; Wang *et al.*, 2012 en la Dominion University (Virginia, Estados Unidos)). Una segunda categoría analiza las características socioeconómicas y de comportamiento social que influyen en la elección del modo de transporte (Soria-Lara *et al.*, 2017 y Miralles-Guasch *et al.*, 2014 en la Universitat Autònoma de Barcelona; Zhou, 2014 en la University of California, Los Angeles (Estados Unidos); Seguí-Pons *et al.*, 2013 en la Universidad de las Islas Baleares; Delmelle & Delmelle, 2012 en la University of Idaho (Estados Unidos). Una tercera sección busca examinar las emisiones de gases generadas por los viajes diarios a una universidad de estudio (Soria-Lara *et al.*, 2017 en la Universitat Autònoma de Barcelona; Davison *et al.*, 2015 en varias universidades de Reino Unido e Irlanda). En general, las encuestas de todos estos trabajos tienen una estructura similar, con un bloque de preguntas sobre los estudiantes y sus características socioeconómicas y otro sobre los hábitos de movilidad diaria y modos de viaje empleados. Habitualmente, el campo de estudio es una única universidad. Todas estas encuestas presentan un periodo temporal no superior a un mes. De forma general, del 10 al 30 % de las encuestas enviadas fueron contestadas. En algunos casos se han combinado las encuestas con otros métodos como entrevistas (Miralles-Guasch *et al.*, 2014). Las muestras son mayores respecto a las encuestas realizadas in-situ, pero siguen presentando un bajo detalle temporal y dificultad de actualización de los datos.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

El uso de fuentes basadas en Big Data para el estudio de la movilidad universitaria ha empezado a entrar muy recientemente. Con el objetivo de entender las características de los desplazamientos a la Universitat Autònoma de Barcelona, Delclòs-Alió y Miralles-Guasch (2017), después de realizar una encuesta por internet, tomaron una muestra de 71 personas y les suministraron una aplicación de tracking basada en el GPS, dotando al estudio de datos con una alta resolución espacial. Finalmente, seleccionaron a 25 personas para realizar entrevistas personales, realizando un estudio muy completo, donde se integran las ventajas de métodos tradicionales y nuevas fuentes. Sin embargo, los estudios que emplean nuevas fuentes de datos para el estudio de la movilidad universitaria siguen siendo bastante minoritarios.

La originalidad del presente trabajo respecto a estudios anteriores radica en el uso de las redes sociales geolocalizadas como fuentes de datos para identificar estudiantes universitarios, los campus a los que asisten y sus lugares de residencia. Este método permite obtener de forma rápida y barata una muestra mayor que las obtenidas por encuestas tradicionales. Hasta donde los autores han encontrado no consta que se hayan usado datos de Twitter previamente para estudiar la movilidad universitaria.

3. Área de estudio y datos

3.1. Área de estudio

El Área Metropolitana de Madrid está compuesta por 52 municipios y tiene una población estimada en 6 millones de habitantes en el año 2018². Es la principal área en cuanto a población, actividad y servicios en España, y una de las principales áreas metropolitanas europeas. Cuenta con seis universidades públicas y siete universidades privadas. En el curso 2017/18 las universidades madrileñas contaron con más de 315.000 estudiantes inscritos, de los cuales aproximadamente más de 210.000 se encontraban matriculados en universidades públicas o centros adscritos, unos 65.000 en universidades privadas, y casi unos 40.000 restantes en universidades especializadas en estudios a distancia³.

En este artículo se ha trabajado con todas las universidades madrileñas, ya sean públicas o privadas, tanto a nivel de universidad como de campus universitario, y se ha identificado la residencia de los estudiantes encontrados a nivel de municipio, excepto en el caso del municipio de Madrid que se ha trabajado a nivel de distritos. El periodo temporal empleado consiste en datos de los cursos escolares 2016/17 y 2017/18. La Figura 1 muestra la ubicación de los campus universitarios en el Área Metropolitana de Madrid.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

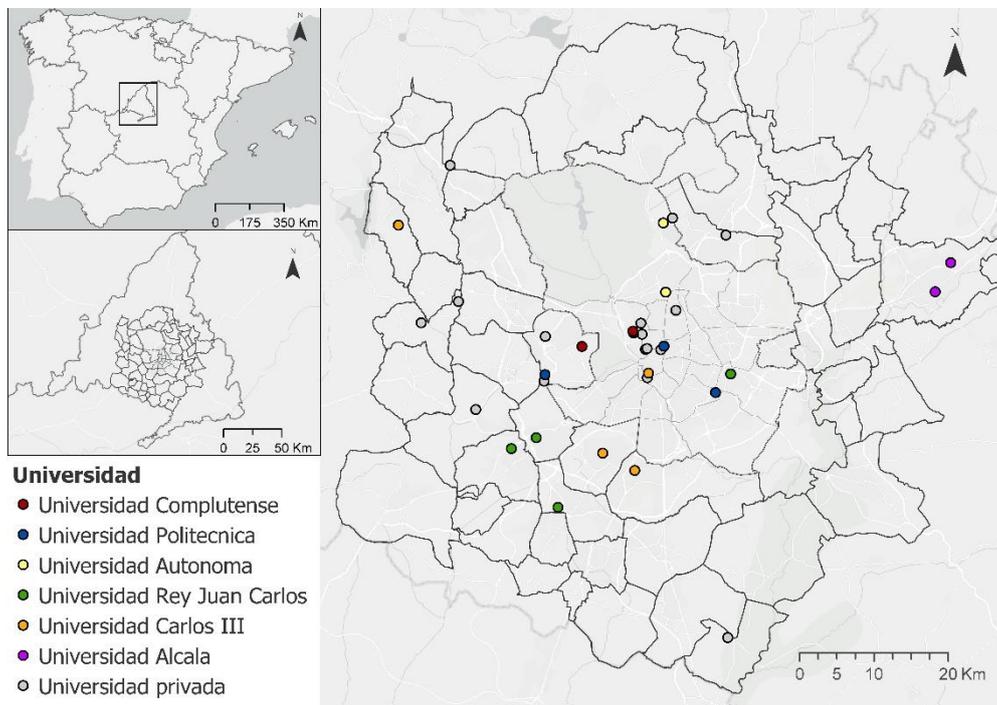


Figura 1: Ubicación de los campus universitarios en el área de estudio

Fuente: Elaboración propia.

3.2. Datos de Twitter

La fuente de datos principal utilizada para este artículo consiste en mensajes de Twitter, georreferenciados y localizados en la Comunidad de Madrid en el periodo comprendido entre el 1 de septiembre de 2016 y el 30 de junio de 2018. La base de datos inicial constaba de 2.229.753 tweets, realizados por un total de 171.631 usuarios. Cada tweet cuenta con información relativa al identificador de usuario, el nombre de usuario, la latitud y longitud, fecha y hora, idioma y los *hashtags* que incluye el tweet.

3.3. Otras fuentes de datos

Las localizaciones de los tweets fueron cruzadas con el uso del suelo. Para ello se ha trabajado con la información del Catastro de 2017. A partir de las parcelas de catastro se ha diseñado una capa de parcelas residenciales y residencias universitarias para localizar las viviendas de los estudiantes, y otra capa de facultades universitarias.

Para analizar el tiempo de viaje desde la residencia hasta la universidad en transporte privado se han obtenido los tiempos de viaje a partir de la red de carreteras de la compañía *TomTom*. La red es muy detallada, cubre toda el área de estudio y cuenta con una conectividad total. En el caso del tiempo en transporte público se han usado ficheros *GTFIS* suministrados por el Consorcio de Transportes de la Comunidad de Madrid. Estos ficheros cuentan con datos de tiempos viaje del metro y metro ligero de Madrid, el servicio de trenes de Cercanías, los autobuses municipales, urbanos, e interurbanos.

Los datos utilizados para estudiar los tiempos de viaje según el nivel económico del lugar de residencia han sido los datos de 2016 del Instituto de Estadística de la Comunidad de Madrid

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

(datos de nivel de renta bruta de cada municipio de la Comunidad de Madrid) y datos de 2015 del Urban Audit del Ayuntamiento de Madrid (datos de nivel de renta neta de cada distrito del municipio de Madrid, que fueron a continuación transformados en datos de renta bruta).

Por último, la página web del Ministerio de Educación, Cultura y Deporte⁴ cuenta con un apartado de datos estadísticos de los que se han obtenido datos pertinentes al curso 2017/18, tanto del número de alumnos matriculados por grado, master y doctorado a nivel de universidad y de campus, como datos de número de plazas de oferta, demanda y matrícula que ofertan cada universidad y campus. Estos datos han sido empleados para analizar la precisión de los datos de Twitter y para el cálculo del modelo de Huff.

4. Metodología

4.1. Recogida, limpieza y filtrado de datos e identificación de estudiantes universitarios

Los tweets utilizados fueron descargados a través de la API de Twitter mediante un código *Python*, y después fueron recopilados e incorporados en un Sistema de Información Geográfica (GIS). Se utilizaron dos softwares GIS: *QGIS* y *ArcGIS Pro*. Para este trabajo se emplearon los tweets publicados en días laborables, eliminando los mensajes publicados en fines de semana o vacaciones. Para limpiar la base de datos de cuentas bots, se calculó el número de mensajes publicados por cada usuario y se eliminaron los usuarios con más de 1.000 tweets con contenido semántico idéntico o similar. Además, se filtraron los usuarios con baja actividad en la red social (menos de 5 tweets en días laborables), y se eliminaron los usuarios con poca movilidad espacial (todos sus tweets están localizados en un radio de menos de 50 metros entre ellos), o con poca movilidad temporal (todos sus tweets fueron publicados en un periodo menor a dos semanas consecutivas).

Una vez realizados los procesos de limpieza y filtrado de usuarios, el siguiente paso fue detectar usuarios de los campus universitarios, en su mayoría estudiantes. Para ello, se clasificaron los tweets en función de si fueron publicados en el horario de día (8 de la mañana a 8 de la tarde) o de noche (el resto). A continuación, se construyó una capa de polígonos de las facultades universitarias de la Comunidad de Madrid a partir del Nomenclátor oficial y Callejero de la Comunidad de Madrid, y se corrigieron posibles errores tomando como referencia información del catastro y del callejero *OpenStreetMap*. Después, se realizó un buffer para seleccionar tweets publicados durante el día a unas determinadas distancias de los edificios universitarios. Para facultades fuertemente integradas en el tejido urbano de la ciudad, se empleó un buffer de 20 metros, para facultades dentro del tejido urbano pero con cierta separación respecto a otros edificios y usos, la distancia realizada fue de 50 metros, y para facultades integradas en áreas específicas de campus universitarios, se buscaron los tweets a una distancia máxima de 100 metros. Como resultado, se obtuvieron 15.380 mensajes realizados por 5.296 personas (el 3 % de usuarios respecto a la base inicial). A cada usuario se le asignó la universidad y campus desde la cual publicó un mayor número de mensajes.

4.2. Identificación del lugar de residencia los usuarios

A continuación, se expandió la base de datos a partir de la búsqueda de los últimos 3.200 tweets de cada uno de los 5.296 posibles estudiantes universitarios encontrados previamente, con el objetivo de aumentar la precisión espacial y temporal de los movimientos individuales de

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

cada uno de estos usuarios (Huang y Wong, 2016). Como resultado, la base de datos alcanzó un total de 18.951 mensajes.

Tras dicha expansión, se procedió a buscar el lugar de residencia de cada una de estos usuarios. Para ello, los tweets publicados por estos usuarios en el periodo de noche fueron cruzados con una capa de usos del suelo de catastro, y se seleccionaron los mensajes ubicados a una distancia máxima de 10 metros respecto a parcelas cuyo uso principal fuese residencial. A continuación, los tweets elegidos fueron agregados a nivel de parcela catastral. Por cada ID de usuario, se seleccionó la parcela residencial con mayor número de tweets (Ahas et al, 2010). En caso de que un usuario tuviese dos o más parcelas con el mismo máximo de tweets, se calculó el centro de mediana de todos sus tweets y se escogió la parcela más cercana a dicha mediana. Con esta metodología, se encontró la zona de residencia de 2.912 usuarios. Los lugares de residencia a nivel de parcela fueron agregados a nivel de municipio o distrito en caso de la ciudad de Madrid para posteriores análisis, mitigando así los problemas de privacidad que pudiese ocasionar la detección de parcelas de residencia.

4.3. Áreas de influencia de las universidades y matrices Origen-Destino

Para cartografiar las áreas de influencia de cada una de las universidades de la Comunidad de Madrid, se ha usado el lugar de residencia de los 2.912 estudiantes encontrados según municipios y distritos. Cada uno de estos usuarios tiene un distrito o municipio asignado como residencia, además de una universidad y campus de destino. Con esta información se ha obtenido una matriz residencia-campus. La matriz cuenta con 70 orígenes (21 distritos de la ciudad de Madrid y 49 municipios del área metropolitana), y 34 destinos (campus universitarios). Sin embargo, del total de 2.380 relaciones, 680 tenían datos de usuarios. Esta matriz ha sido expandida a partir de los datos de estudiantes según campus usando la siguiente fórmula:

$$E_{ij}^e = E_{ij} \cdot \frac{p_j}{\tilde{p}_j}, \forall j \in N,$$

donde $\frac{p_j}{\tilde{p}_j}$ son los pesos calculados para cada campus N , basándose en el número total de estudiantes por campus según las fuentes oficiales p_j y la muestra de usuarios de Twitter por campus \tilde{p}_j . Estos pesos son multiplicados por el valor de los viajes identificados con Twitter en cada flujo E_{ij} , siendo el resultado flujos expandidos E_{ij}^e .

4.4. Cálculo de distancias ponderadas

El cálculo de distancias medias recorridas se elaboró a partir de la matriz obtenida en el paso anterior. A partir de los datos de Twitter no es posible identificar el modo de transporte utilizado desde el lugar de residencia al campus. Sin embargo, se puede hacer una estimación del tiempo de viaje desde cada municipio al campus tanto en transporte público como privado. A partir de esos datos se calcularon distancias ponderadas por el número de estudiantes residentes en el municipio usando como fuente de datos los tiempos de viaje obtenidos a partir de ficheros *GTFS* para el cálculo de distancias ponderadas a partir de tiempos de transporte público, y datos de transporte privado *TomTom* para la obtención de distancias ponderadas mediante tiempos de transporte privado. El tiempo ponderado para cada universidad se calculó a partir de la siguiente fórmula:

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

$$T_u = \sum_{i \in N} \left(\sum_{c} T_{ic} \frac{\tilde{p}_c}{\sum \tilde{p}_u} \right) \frac{\tilde{p}_i}{\sum \tilde{p}_i}, \forall u \in U,$$

donde T_u es la media del tiempo ponderado de cada universidad u , N es el conjunto de todos los municipios y distritos del área de estudio, T_{ic} es el tiempo registrado de transporte público o privado una ruta entre un municipio i y un campus c , \tilde{p}_i es la muestra de usuarios de Twitter por municipio, y $\frac{\tilde{p}_c}{\sum \tilde{p}_u}$ es el valor de ponderación de un campus a partir de \tilde{p}_c (el número de usuarios de Twitter obtenidos en un campus c) y \tilde{p}_u (el número de usuarios de Twitter de una universidad u).

4.5. Análisis de tiempos de viaje a los campus según niveles de renta

Para el análisis de tiempos de acceso según el nivel de renta del lugar de origen se han agrupado los distritos y municipios por cuartiles, clasificándolos según su cuartil en municipios/distritos de rentas alta, media-alta, media-baja, y baja (Figura 2). Para cada uno de estos grupos se ha calculado el porcentaje de alumnos cuya residencia ha sido estimada, y la media de tiempos de viaje a partir de los datos obtenidos previamente a nivel de municipio. El porcentaje de usuarios cuya residencia ha sido estimada mediante datos de Twitter corresponde al 0,62 % respecto al total de la población joven entre 18 y 25 años según el censo (Tabla 1).

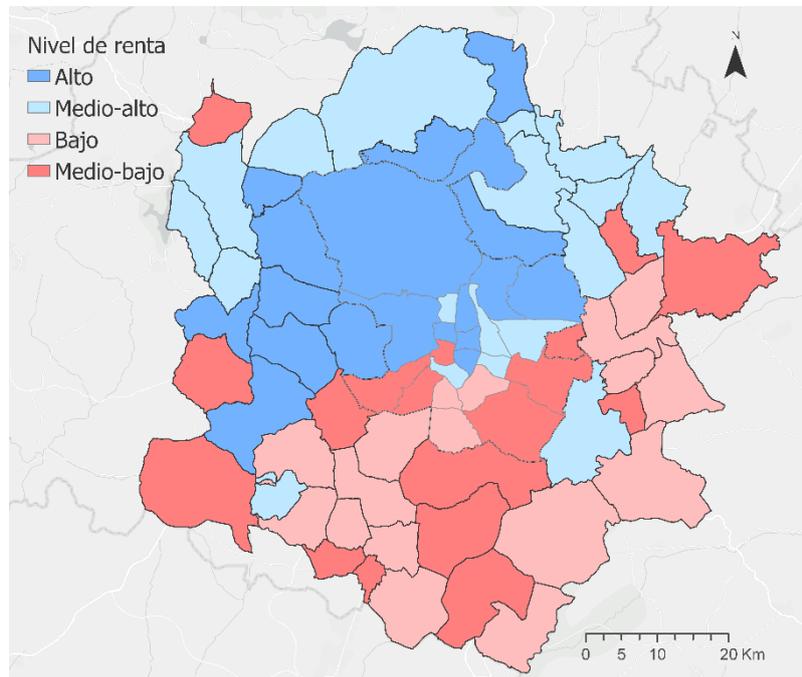


Figura 2: Municipios y distritos del Área Metropolitana de Madrid según nivel de renta por cuartiles.

Fuente: Elaboración propia a partir de los datos de 2016 del Instituto de Estadística de la Comunidad de Madrid y datos de 2015 del Urban Audit del Ayuntamiento de Madrid.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

Tabla 1: Porcentaje de universitarios encontrados en Twitter respecto a datos censales por cuartil de renta.

Nivel de renta	% Universitarios Twitter sobre población total 18-25 años
Alto	0,78
Medio-alto	0,42
Medio-bajo	0,95
Bajo	0,27
Total	0,62

Fuente: Elaboración propia a partir de Twitter.

4.6. Modelo de asignación de estudiantes (modelo de Huff)

Con el objetivo de comparar los datos obtenidos de Twitter con un modelo teórico de asignación de los estudiantes a los campus, se elaboró un modelo gravitacional de *Huff* que permite obtener una asignación de estudiantes universitarios a cada campus a partir de plazas ofertadas y tiempos de transporte de acceso a los mismos, a partir de la siguiente fórmula:

$$N_{iu} = \frac{O_u \cdot \frac{1}{T_u}}{\sum_{O \in u} O_u \cdot \frac{1}{T_u}} \cdot D_i, \forall u \in U,$$

donde el número de universitarios N_{iu} del municipio i asignado al campus u se calcula a partir de la oferta de plazas de cada universidad O_u , la inversa del tiempo desde cada municipio o distrito a cada universidad T_u , y el producto del número de habitantes (entre 18 y 25 años) por municipio i obtenidos mediante datos censales.

Una vez obtenido el número estimado de estudiantes asignados desde cada municipio o distrito a cada universidad por el modelo de Huff, se establece una relación con el número de usuarios residentes obtenidos en Twitter, pudiendo así conocer en que municipios y distritos se han obtenido en Twitter un mayor número de usuarios por universidad respecto al modelo gravitacional.

5. Resultados

Tras los procesos de limpieza y filtrado de datos, la base final de tweets alberga 5.296 usuarios, cada uno con una universidad y campus asignados. El coeficiente de correlación (r^2) entre el número de estudiantes detectado con los datos de Twitter y los datos oficiales de estudiantes por universidades muestra un ajuste elevado, de 0,95 a nivel de universidad y 0,92 a nivel de campus. La Figura 3 muestra la distribución de los estudiantes según los datos oficiales y según la muestra de estudiantes detectados en Twitter, corroborando la alta correlación entre la muestra obtenida y los datos oficiales del número de estudiantes. De estos 5.300 usuarios se encontró el lugar de residencia de 2.912 usuarios (el 55% de los posibles usuarios de campus universitarios encontrados). De nuevo, el coeficiente de correlación entre los usuarios de Twitter con residencia estimada y los datos oficiales de estudiantes por universidades muestra un ajuste elevado (0,96 a nivel de universidad y 0,91 a nivel de campus).

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

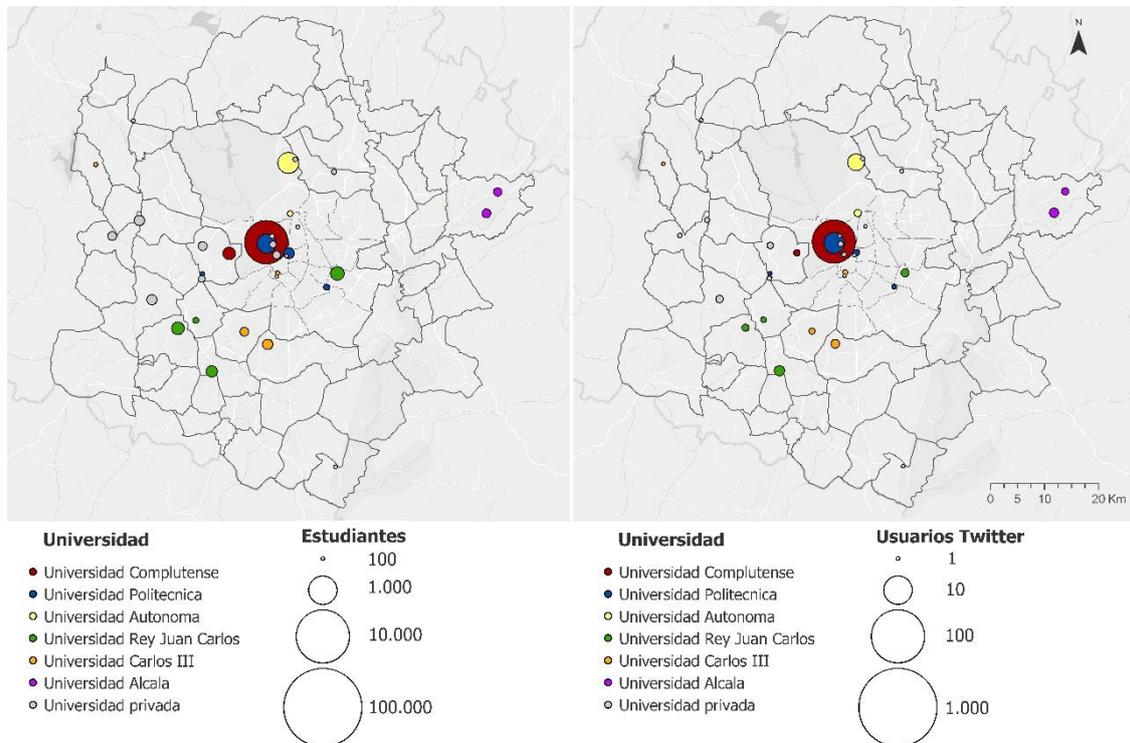


Figura 3: Distribución por campus de los alumnos universitarios en el Área Metropolitana de Madrid a partir de datos oficiales del Ministerio de Educación (izquierda) y número de usuarios detectados en Twitter (derecha).

Fuente: Elaboración propia a partir de datos del Ministerio de Educación, Cultura y Deporte, y Twitter.

5.1. Lugares de residencia de los estudiantes y áreas de influencia de las universidades

Los principales lugares de procedencia de los usuarios identificados con Twitter son los distritos centrales de la ciudad de Madrid, en especial los distritos de Centro y Chamberí (distritos con barrios y servicios orientados a la población joven), y Ciudad Universitaria (distrito que cuenta con un importante número de residencias universitarias). También hay un número importante de residentes en distritos periféricos del norte y sur de la ciudad (zonas dormitorio de la ciudad que albergan campus universitarios como Fuencarral-El Pardo y Vallecas, o se encuentran cerca como Carabanchel). Finalmente, destacan otros municipios con elevado número de habitantes y con importantes campus universitarios, como Alcalá de Henares, o los principales municipios del sur del área metropolitana (Getafe, Leganés, Alcorcón, Móstoles, y Fuenlabrada) (Figura 4).

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

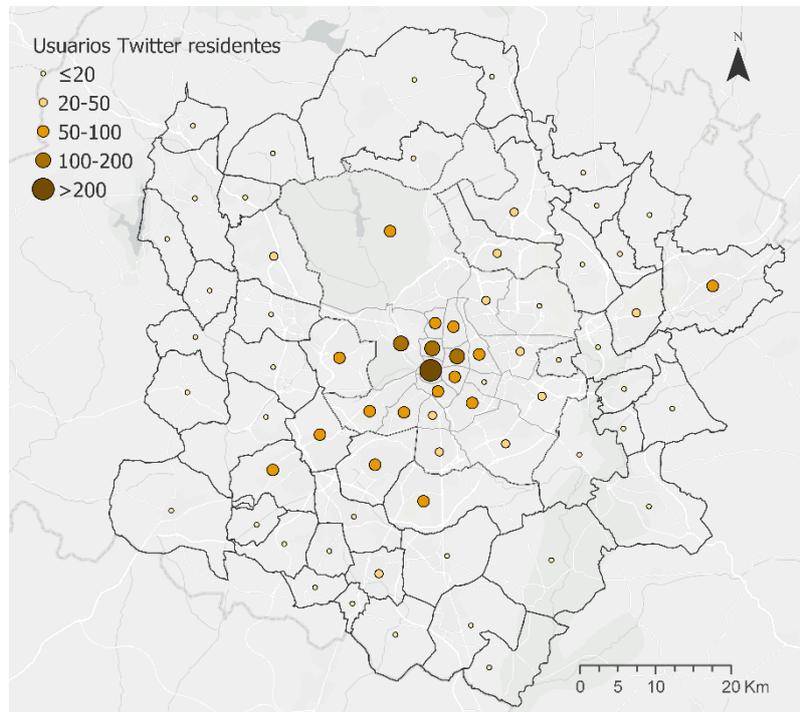


Figura 4: Número de usuarios residentes detectados en el Área Metropolitana de Madrid a partir de Twitter.

Fuente: Elaboración propia a partir de Twitter.

A partir de los datos de estudiantes residentes en cada distrito o municipio se pueden establecer áreas de influencia de las distintas universidades. La Universidad Complutense es la que cuenta con una mayor área de influencia que incluye casi toda el área metropolitana. El área de influencia de la Universidad Politécnica es similar, pero con menores porcentajes de alumnos. La Universidad Autónoma cuenta con una influencia importante en el norte del área metropolitana, mientras que las universidades Rey Juan Carlos y Carlos III destacan en el sur metropolitano y la Universidad de Alcalá al este en el corredor del Henares. Las universidades privadas tienen áreas de influencia menores, generalmente en el oeste metropolitano, y menor porcentaje de alumnos (Figura 5). Por tanto, destaca la importancia de la proximidad de la residencia respecto a la universidad, coincidiendo a grandes rasgos la ubicación del campus con el área de influencia de la universidad: el campus de Cantoblanco de la Universidad Autónoma está situado en el norte metropolitano, los principales campus de las universidades Rey Juan Carlos y Carlos III se hallan en los municipios más poblados del sur, la Universidad de Alcalá se sitúa en el este, y los principales campus periféricos de las universidades privadas están en el oeste.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

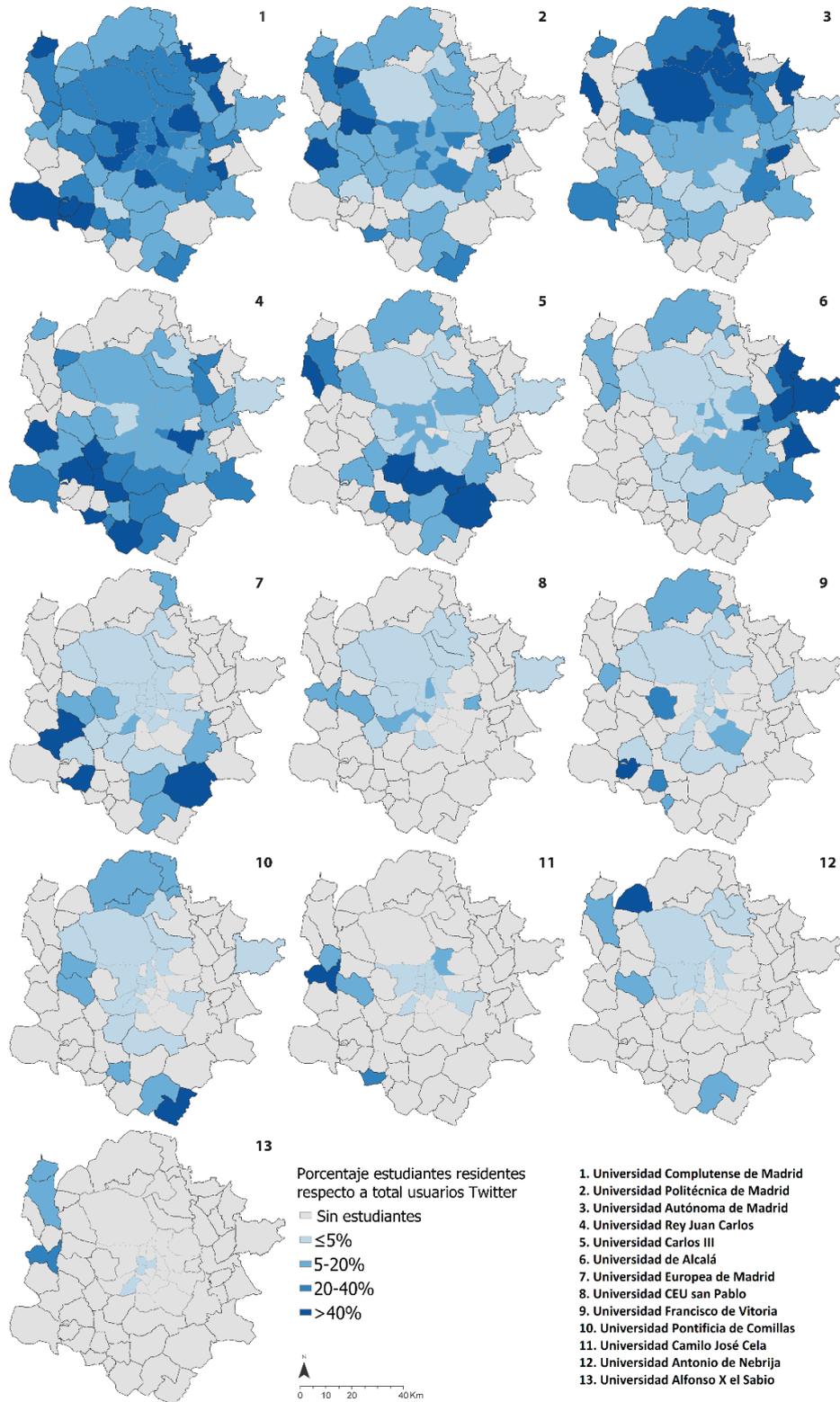


Figura 5: Porcentaje de estudiantes universitarios detectados en Twitter por universidad en cada municipio y distrito.

Fuente: Elaboración propia a partir de Twitter.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

5.2. Tiempos de viajes a partir de transporte privado y público

Los tiempos ponderados obtenidos a partir de datos de *TomTom* para el uso del coche y de ficheros *GTFIS* para el sistema de transporte público muestran que hay una gran diferencia en los tiempos de viaje en transporte privado respecto al transporte público: 23,12 minutos de media en coche, mientras el tiempo mediante transporte público aumenta a 56,34 minutos. Las universidades con campus centrales, principalmente las universidades Complutense y Politécnica, están mejor conectadas tanto en transporte público como por carretera respecto a universidades periféricas, por lo que cuentan con tiempos menores de acceso. Mientras, las universidades cuyos campus se sitúan en zonas periféricas del área de estudio (Alcalá, Alfonso X el Sabio), muestran mayores tiempos de viaje (Tabla 2).

Tabla 2: Tiempo medio ponderado de viaje a universidades por tipo de transporte.

Universidad	Tiempo medio ponderado transporte público (minutos)	Tiempo medio ponderado transporte privado (minutos)
Universidad Complutense de Madrid	38,24	16,61
Universidad Politécnica de Madrid	37,47	16,94
Universidad Autónoma de Madrid	53,06	26,26
Universidad Rey Juan Carlos	58,81	22,65
Universidad Carlos III	48,93	19,62
Universidad Alcalá	80,58	33,76
Universidad Europea de Madrid	70,77	23,92
Universidad CEU San Pablo	45,97	18,25
Universidad Francisco de Vitoria	68,43	21,48
Universidad Pontificia de Comillas	49,04	21,85
Universidad Camilo José Cela	57,31	27,86
Universidad Antonio de Nebrija	45,05	19,18
Universidad Alfonso X el Sabio	78,82	32,20
Total	56,34	23,12

Fuente: Elaboración propia.

Se puede apreciar mayor dispersión de tiempos de viaje a partir de transporte público, mientras que los tiempos de viaje por transporte privado están muy concentrados. Esto se debe a que mientras el acceso a la red de carreteras del Área Metropolitana de Madrid es homogéneo, hay una diversa variedad de situaciones a la hora de viajar en transporte público a diferentes universidades (variedad de tipos de transporte público, diversidad de horarios y flotas de vehículos, proximidad al tejido urbano) (Figura 6).

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

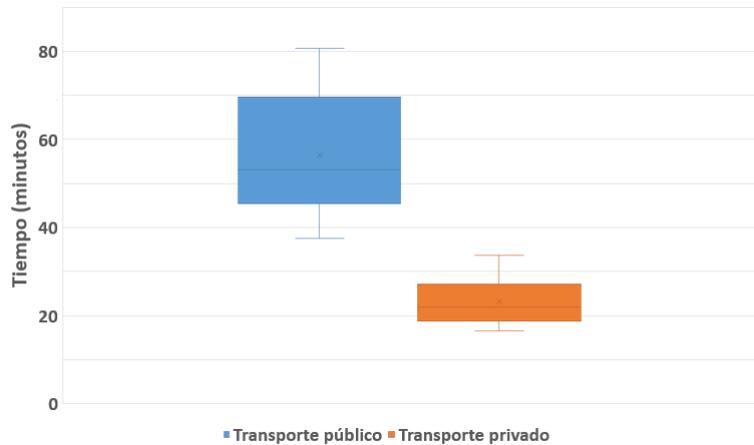


Figura 6: Diagrama de caja de tiempos ponderados de viaje a universidad por tipo de transporte.

Fuente: Elaboración propia.

5.3. Comparación de resultados entre zonas de distintos niveles adquisitivos

El porcentaje de usuarios de Twitter asignados a una universidad pública es del 87,5% mientras que solo el 12,2% de los usuarios encontrados asisten a una universidad privada. Analizando los distritos y municipios del área de estudio clasificados por cuartiles según su nivel de renta, se puede observar que el grupo formado por los municipios de mayor nivel de renta cuenta con el mayor porcentaje de alumnos en universidades privadas (un 23,3%). En el resto de cuartiles se puede observar que los municipios con un nivel de media medio-alto también cuenta con un porcentaje de alumnos asistentes a universidades públicas ligeramente por debajo de la media, mientras que los municipios con nivel de renta medio-bajo o bajo cuentan con un porcentaje de alumnos de universidades públicas superior a la media (Figura 7).

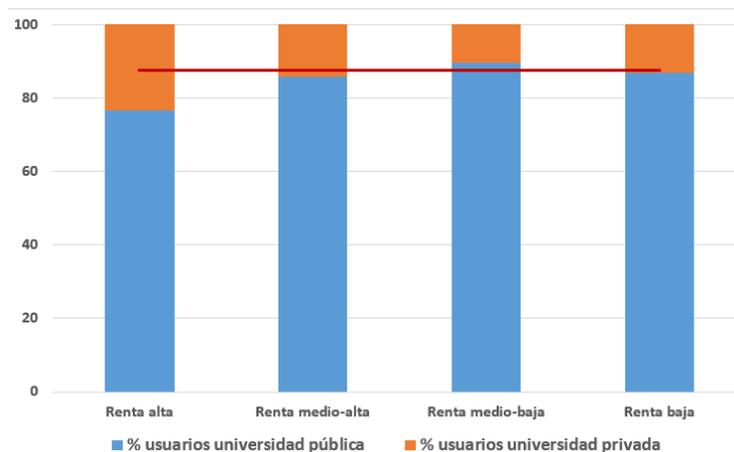


Figura 7: Porcentaje de alumnos residentes por tipo de universidad según grupos de municipios y distritos por nivel de renta.

Fuente: Elaboración propia.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

Analizando los tiempos medios de viaje a partir de los municipios y distritos clasificados según nivel de renta se puede apreciar que aquellos con mayor nivel de renta presentan los tiempos de desplazamiento más bajos. Tanto los tiempos en transporte privado como público están por debajo del nivel medio. En el resto de grupos se observa que los tiempos de viaje a partir de transporte público son superiores a la media, mientras que los tiempos de viaje usando transporte privado están ligeramente sobre la media (Figura 8).

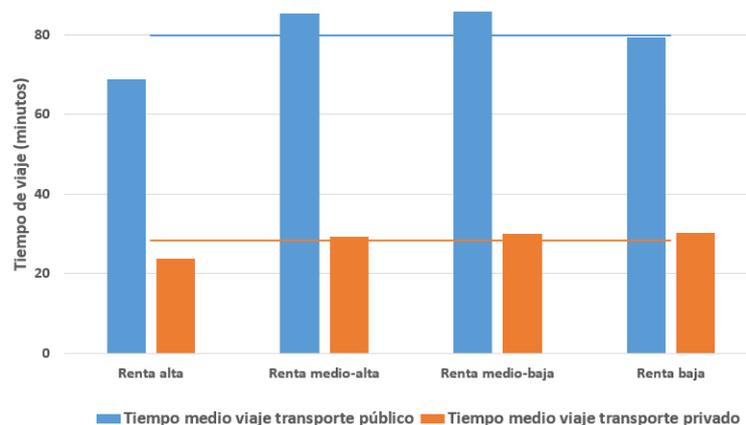


Figura 8: Tiempos de transporte privado y público según grupos de municipios y distritos por nivel de renta.

Fuente: Elaboración propia.

5.4. Modelo de Huff y relación con la asignación de Twitter

El modelo de Huff asume que los usuarios tienden a escoger la localización de un servicio, en este caso una universidad, basándose en la distancia a ese destino y su capacidad de atracción (aquí el número de plazas de una universidad). Este modelo estima una probabilidad de usar cada uno de los destinos ofertados (Rodríguez et al, 2016). Obviamente, en la elección de la universidad en la que se estudia participan otros muchos motivos. En este trabajo, comparando los modelos resultados de la asignación de usuarios a los campus según Twitter y los asignados por el modelo de Huff, se quiere ver qué universidades están atrayendo más estudiantes y cuáles menos de los estimados por Huff, es decir, la influencia de esos otros motivos.

Para ello se ha calculado el número de estudiantes por municipio y distrito y universidad a partir de este modelo gravitacional. Se han obtenido modelos basados en los dos tipos de transporte tratados con unos porcentajes de asignación muy similares, aunque con algunas diferencias. Se puede observar un número de universidades cuyo porcentaje de personas asignadas en transporte público es mayor respecto al transporte privado (Politécnica, Rey Juan Carlos, Alcalá, Europea de Madrid, Francisco de Vitoria), y universidades con mayores porcentajes asignados en transporte privado respecto al transporte público (Complutense, Autónoma, Carlos III, la mayoría de las universidades privadas) (Tabla 3). El número de plazas ofertadas, la ubicación de la universidad y la proximidad a redes de carreteras o de servicios de transporte público influyen en estos porcentajes.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

Tabla 3: Porcentaje de estudiantes asignados por universidad a partir del modelo de Huff.

Universidad	% transporte público	% transporte privado
Universidad Complutense de Madrid	27,08	27,92
Universidad Politécnica de Madrid	15,79	15,21
Universidad Rey Juan Carlos	16,18	14,69
Universidad Autónoma de Madrid	9,36	10,66
Universidad Carlos III	6,58	7,02
Universidad de Alcalá	4,45	4,06
Universidad Europea de Madrid	4,46	3,85
Universidad Camilo José Cela	3,17	3,83
Universidad CEU San Pablo	3,60	3,57
Universidad Pontificia de Comillas	2,68	2,93
Universidad Alfonso X el Sabio	2,43	2,53
Universidad Francisco de Vitoria	2,85	2,26
Universidad Antonio de Nebrija	1,38	1,48

Fuente: Elaboración propia.

Observando la relación entre usuarios de Twitter y usuarios asignados al modelo de Huff por universidad desde cada municipio y distrito, se observa que en todos los casos en los distritos centrales de Madrid hay un número de usuarios de Twitter superior al número asignado por el modelo de Huff. Igualmente, en cada universidad se observa un número mayor de usuarios de Twitter respecto al modelo gravitacional en los municipios y distritos con un campus perteneciente a la universidad de estudio, y en los municipios y distritos adyacentes. En cambio, se puede apreciar un número mayor de usuarios asignados al modelo de Huff respecto a Twitter en los municipios periféricos del área de estudio. Finalmente, se observa una mayor aproximación entre usuarios de Twitter y usuarios asignados al modelo de Huff en las universidades públicas, mientras que las universidades privadas presentan en su mayoría un número bajo de usuarios de Twitter respecto al modelo de Huff (Figura 9).

Osorio-Arjona, J. y García-Palomares, J. C. (2019): "Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter", *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

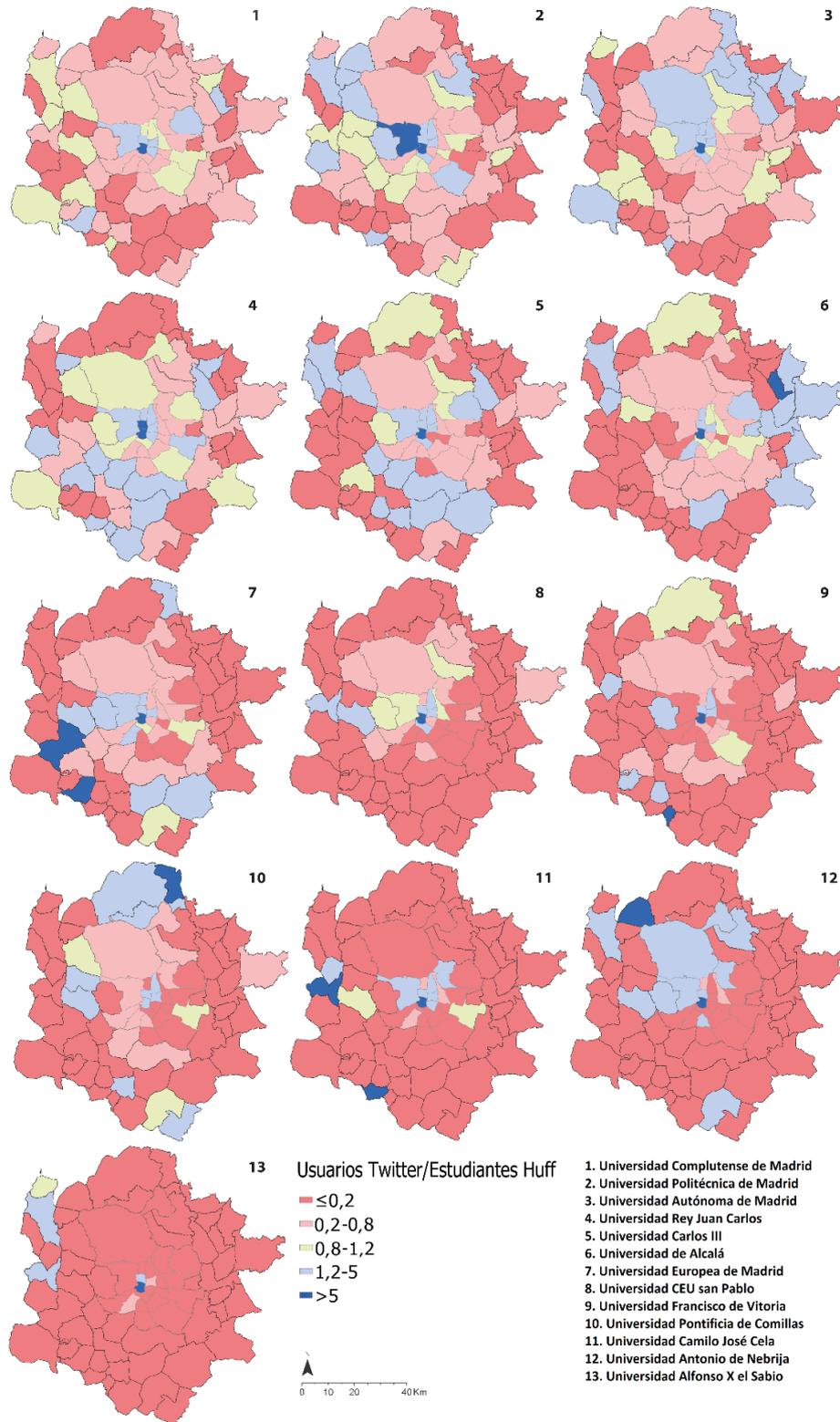


Figura 9: Relación entre número de usuarios Twitter y estudiantes estimados a partir del modelo de Huff.

Fuente: Elaboración propia a partir de Twitter.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

6. Conclusiones

El estudio de la movilidad de los estudiantes universitarios es un tema de creciente interés vinculado a ámbitos como la promoción de la movilidad sostenible o la lucha contra la desigualdad social. Para este tipo de estudios se ha empleado tradicionalmente encuestas de movilidad. Sin embargo, estas encuestas son costosas y requieren mucho tiempo. Ante este paradigma, el Big Data y las redes sociales se antojan como fuentes de datos novedosas con las que poder estudiar la movilidad universitaria.

Este trabajo ha pretendido explorar el uso de la red social Twitter para el estudio de la movilidad de los estudiantes universitarios, aprovechando que la población joven es la principal consumidora de este *microblogging*. Partimos de la premisa de que los datos de Twitter son una representación aproximada a la realidad. A partir de la muestra obtenida se ha podido observar una buena correlación entre los usuarios cuyo campus y residencia han sido estimados, y los datos oficiales obtenidos a partir del Ministerio de Educación. La estimación de la residencia a nivel de parcela puede conllevar implicaciones sobre la privacidad y confidencialidad de los usuarios de Twitter, por lo que estos lugares de residencia son agregados a un nivel mayor, de municipio o distrito, mitigando así los problemas de privacidad y confidencial que se pudieran causar.

Los resultados obtenidos permiten observar la universidad y campus de los alumnos encontrados en Twitter. Destaca la importancia de las universidades públicas, principalmente la Universidad Complutense de Madrid y el Campus de Ciudad Universitaria. En cuanto al lugar de residencia, este trabajo da una aproximación acerca de los municipios y distritos donde residen los estudiantes universitarios, aportación importante ya que no hay fuentes oficiales accesibles al respecto. Los resultados muestran que los usuarios registrados suelen habitar en los distritos centrales de Madrid (zonas con servicios orientados a la población universitaria), o en distritos o municipios cercanos al campus al que viajan, siendo la proximidad un factor importante tanto en la elección del campus en el que se estudia como también en la elección del lugar de residencia cuando la entrada a la universidad supone un cambio en la misma. También se pueden obtener de forma general las áreas de influencia de cada universidad. El número de alumnos matriculados en un campus, la cercanía del lugar de residencia, el tamaño y caracterización socioeconómica del municipio, el prestigio de una universidad, y la facilidad de acceso a redes de transporte son los principales factores que definen todas esas áreas. Las universidades generan áreas de influencia en torno a los distritos o municipios cercanos a sus principales campus universitarios.

Analizando las distancias obtenidas mediante ficheros de transporte público *GTFS*, y datos de transporte privado *TomTom*, y ponderando estos tiempos por el reparto de los universitarios a nivel de municipio o distrito, los resultados mostraron como los tiempos son mucho más bajos en transporte privado. Además, se observa que los tiempos de viaje de transporte público por universidad son dispares mientras que los tiempos a partir de transporte privado son más homogéneos. En general, cuanto más cerca del centro está el campus, mejores tiempos de transporte tiene al conllevar los desplazamientos un menor tiempo de viaje y al tener acceso a un mayor servicio y frecuencia de transportes. Clasificando los distritos y municipios del área metropolitana por cuartiles en base al nivel de renta bruta, se ha observado que los estudiantes asignados a universidades privadas destacan principalmente en las zonas con un nivel alto de renta como el oeste metropolitano. Por último, los tiempos de transporte registran

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

mejores resultados desde el cuartil que agrupa a los municipios con mayor nivel de renta, sobre todo en transporte público.

Finalmente, pese a la demostración de Twitter como fuente veraz para aplicar en este tipo de estudios, hay que resaltar algunas limitaciones de cara a futuras líneas de investigación. Entre estas limitaciones están la baja cantidad de mensajes con la que se ha trabajado, el sesgo de la red social (mientras se ha aprovechado para obtener potenciales alumnos, el sesgo sigue dificultando la obtención de otros grupos como profesores o trabajadores) o la imposibilidad de tener datos acerca del reparto modal de los viajeros a partir del tipo de transporte realizado.

Referencias bibliográficas

Ahas, R., Silm, S., Järv, O., Saluveer, E., & Tiru, M. (2010): “Using mobile positioning data to model locations meaningful to users of mobile phones”, *Journal of Urban Technology*, 17(1), pp. 3–27. <http://doi.org/10.1080/10630731003597306>

Albertos, J.M., Noguera, J., Pitarch, M.D., & Salom, J. (2008). “Los hábitos de movilidad en la Universitat de València (2005-2006): Problemas de acceso a los campus y sostenibilidad”, Valencia, Universitat de València.

Davison, L., Ahern, A., & Hine, J. (2015): “Travel, transport and energy implications of university-related student travel: A case study approach”, *Transportation Research Part D: Transport and Environment*, 38, pp. 27-40. <http://doi.org/10.1016/j.trd.2015.04.028>

De las Rivas, J.L., Iglesias, F., & Lalana, J.L. (2011): “Campus universitario de Valladolid. Integración urbana y movilidad”, *Bitácora*, 18(1), PP. 139-156.

Delclòs-Alió, X., & Miralles-Guasch, C. (2017): “Suburban travellers pressed for time: Exploring the temporal implications of metropolitan commuting in Barcelona”, *Journal of Transport Geography*, 65, pp. 165-174. <http://doi.org/10.1016/j.jtrangeo.2017.10.016>

Delmelle, E. M., & Delmelle, E. C. (2012): “Exploring spatio-temporal commuting patterns in a university environment”, *Transport Policy*, 21, pp. 1–9. <http://doi.org/10.1016/j.tranpol.2011.12.007>

Gutiérrez, J., Ruiz, E., & de Rodrigo, R. (2016). “Diagnóstico de la movilidad en los campus de la universidad de Extremadura”, *XVII Congreso Nacional de Tecnologías de la Información Geográfica*, pp. 140-154.

Huang, Q., & Wong, D. W. S. (2016): “Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us?”, *International Journal of Geographical Information Science*, 30(9), pp. 1873–1898. <https://doi.org/10.1080/13658816.2016.1145225>

Lucas-García F., RAcero-moreno, J., Torrecillas C., & García-Sánchez J.M. (2016). “Análisis de movilidad en campus universitarios integrados en zonas urbanas”, *Dyna*, 91(3), pp. 1-10. <http://dx.doi.org/10.6036/7660>

Ministerio de Fomento (2018): Las Grandes Áreas Urbanas y sus municipios (ordenadas por población)2018. https://www.fomento.gob.es/recursos_mfom/comodin/recursos/listado_areas_urbanas_por_municipios_2018.pdf (consultado 10-07-2019).

Miralles-Guasch, C., & Domene, E. (2010): “Sustainable transport challenges in a suburban

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

university: The case of the Autonomous University of Barcelona”, *Transport Policy*, 17(6), pp. 454-463. <http://doi.org/10.1016/j.tranpol.2010.04.012>

Miralles-Guasch, C., Martínez Melo, M., & Marquet Sarda, O. (2014): “On user perception of private transport in Barcelona Metropolitan area: An experience in an academic suburban space”, *Journal of Transport Geography*, 36, pp. 24-31. <http://doi.org/10.1016/j.jtrangeo.2014.02.009>

Moravec Giormenti, B., López Dentone, F., Bossio, D., Filgueira, E. M., Gurrera, W., Piccirillo, J. M. Rolón, H., Caprelli, C. (2018): “La movilidad de los estudiantes de la Facultad Regional Avellaneda de la Universidad Tecnológica Nacional”, Avellaneda, Universidad Tecnológica Nacional. Disponible en www.c3t.fra.utn.com.ar

Osorio Arjona, J., & García-Palomares, J. C. (2017): “Nuevas fuentes y retos para el estudio de la movilidad urbana”, *Cuadernos Geográficos*, 56(3), pp. 247–267. Disponible en <http://revistaseug.ugr.es/index.php/cuadgeo/article/view/5352/5858>

Rodrigue, J.P., Comtois, C., & Slack, B. (2016): “*The geography of transport systems: City Logistics*”, Nueva York, Routledge.

Saladié, O., & Jurado, J. (2015): “La movilidad en el Campus Vila-seca de la URV: propuestas para una movilidad más sostenible”, *Investigaciones Geográficas*, 64, pp. 163-182. <http://doi.org/10.14198/INGEO2015.64.10>

Seguí-Pons, J.M., Ruiz, M., & Luna, M. (2013). Movilidad y transportes en el acceso al Campus de la Universitat de les Illes Balears: una perspectiva de género, *XIII Congreso de Geógrafos Españoles*, pp. 685-695.

Shannon, T., Giles-Corti, B., Pikora, T., Bulsara, M., Shilton, T., & Bull, F. (2006): “Active commuting in a university setting: Assessing commuting habits and potential for modal change”, *Transport Policy*, 13(3), pp. 240–253. <http://doi.org/10.1016/j.tranpol.2005.11.002>

Soria-Lara, J. A., Marquet, O., & Miralles-Guasch, C. (2017): “The influence of location, socioeconomics, and behaviour on travel-demand by car in metropolitan university campuses”, *Transportation Research Part D: Transport and Environment*, 53, pp. 149-160. <http://doi.org/10.1016/j.trd.2017.04.008>

Soria-Lara, J. A., Miralles-Guasch, C., & Marquet, O. (2017): “The influence of lifestyle and built environment factors on transport CO2 emissions: The case study of Autonomous University of Barcelona”, *ACE: Architecture, City and Environment*, 12(34), pp. 11-28. <http://doi.org/10.5821/ace.12.34.4756>

Vij, A., & Shankari, K. (2015): “When is big data big enough? Implications of using GPS-based surveys for travel demand analysis”, *Transportation Research Part C: Emerging Technologies*, 56, pp. 446-462. <http://doi.org/10.1016/j.trc.2015.04.025>

Volosin, S. E., Paul, S., Pendyala, R. M., Livshits, V., & Maneva, P. (2013): “Activity-Travel characteristics of a large university population”, *Urban Transportation Data and Information Systems*, 7, pp. 1–19.

Wang, X., Khattak, A., Son, S., & Agnello, P. (2012): “What can be learned from analysing university student travel demand?”, *Transportation Research Record: Journal of the Transportation Research Board*, 2322, pp. 129–137. <http://doi.org/10.3141/2322-14>

Whalen, K. E., Páez, A., & Carrasco, J. A. (2013): “Mode choice of university students commuting to school and the role of active travel”, *Journal of Transport Geography*, 31, pp.

Osorio-Arjona, J. y García-Palomares, J. C. (2019): “Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter”, *GeoFocus (Artículos)*, n° 24, p.37-57. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.648>

132–142. <http://doi.org/10.1016/j.jtrangeo.2013.06.008>

Zhou, J. (2014): “From better understandings to proactive actions: Housing location and commuting mode choices among university students”, *Transport Policy*, 33, pp. 166–175. <http://doi.org/10.1016/j.tranpol.2014.03.004>

¹ <https://www.omnicoreagency.com/twitter-statistics/>

² http://www.madrid.org/iestadis/fijas/estructu/demograficas/padron/estructupopc_prov.htm

³ <http://www.comunidad.madrid/servicios/educacion/sistema-universitario-madrileno>

⁴ <https://www.mecd.gob.es/servicios-al-ciudadano-mecd/estadisticas/educacion/universitaria/estadisticas/alumnado/2016-2017.html>