



Munich Personal RePEc Archive

Urban mobility and high frequency data

Gutiérrez, Antonio

Universidad de Zaragoza

3 October 2022

Online at <https://mpra.ub.uni-muenchen.de/114854/>
MPRA Paper No. 114854, posted 03 Oct 2022 14:10 UTC

Movilidad urbana y datos de alta frecuencia

Antonio Gutiérrez

Universidad de Zaragoza

RESUMEN:

Los patrones de movilidad urbana están cambiando en respuesta a nuevos comportamientos en las ciudades. El mayor número de trayectos, el aumento en la demanda de vehículos motorizados o el incremento en la distancia de los trayectos hacen necesario el estudio de la movilidad urbana para orientarla hacia un horizonte más sostenible.. El Big Data y la huella digital de personas y vehículos han configurado una nueva base de información de gran utilidad para el estudio de la movilidad urbana. Por este motivo, en este artículo se presentan las distintas herramientas que ofrecen datos de alta frecuencia y resolución espacio-temporal, junto con una revisión de la literatura que emplea estos datos en investigación sobre movilidad urbana.

ABSTRACT:

Urban mobility patterns are changing as a response to new behaviours in cities. With more journeys, increased demand for motorised vehicles and longer distances to travel the need to study urban mobility is necessary to guide society towards a more sustainable horizon. Big Data and the digital footprint of people and vehicles have created a new source of appropriate information for urban mobility studies. Therefore, this article presents the different tools that offer high-frequency and spatial-temporal resolution data along with a review of the literature that uses these datasets in urban mobility research.

Palabras clave: movilidad urbana, redes sociales, Big Data.

Clasificación JEL: R40, C80

1. Introducción

En los últimos 15 años ha surgido nueva literatura basada en el uso de datos de alta frecuencia en investigación. Esta línea, fundamentada en el aprovechamiento del “Big data”, promete un desarrollo considerable para la investigación en economía y política económica (Einav & Levin, 2013). Esta nueva fuente de información constituye una alternativa atractiva frente a los bancos de datos convencionales, los cuales se construyen a partir de métodos de procesamiento y recolección costosos. Por este motivo, series de datos como las que ofrece Google resultan interesantes debido a que son publicadas con mayor frecuencia, puntualidad y concreción.

Uno de los campos de aplicación para este tipo de datos es el de la movilidad urbana, el cual constituye una parte esencial de la vida en las ciudades. Ante el incremento de los desplazamientos dentro de la ciudad, el desarrollo del transporte motorizado y el aumento del tamaño de las ciudades, la movilidad urbana se enfrenta al reto de la sostenibilidad (Osorio-Arjona & García-Palomares, 2017). Tradicionalmente, el estudio de la movilidad ha sufrido de escasez en cuanto a fuentes de datos que proporcionen información detallada para su estudio. Estas fuentes de datos convencionales resultan de gran relevancia debido a su contenido, pues permiten el análisis de las características sociodemográficas que afectan a la movilidad general (Giménez et al., 2014, 2016, 2018, 2019a, 2019c, 2020, 2021, 2022a, 2022b) y a la movilidad sostenible (Belloc et al., 2022a, 2022b; Echeverría et al., 2019, 2022a, 2022b; 2022c; Giménez et al., 2019b; 2022c; Molina et al., 2021). No obstante, para conducir la movilidad urbana hacia un escenario de sostenibilidad se hace imprescindible el uso de nuevas bases de datos que nos ayuden a comprender mejor los desplazamientos en la ciudad.

El propósito fundamental de este artículo es realizar una revisión de la literatura sobre los datos de alta frecuencia y su aplicación al estudio de la movilidad urbana. De este modo, el trabajo se estructura de la siguiente manera. En la Sección 2 se presenta una revisión de la herramienta pionera en investigación con datos de alta frecuencia, Google Trends. En la Sección 3, se revisa la literatura sobre movilidad urbana que incorpora datos de la red social Twitter. En la Sección 4, se presentan las aplicaciones de los datos de teléfonos móvil en el ámbito de la movilidad urbana. En la Sección 5, se muestra la relevancia en la literatura de los datos obtenidos de tarjetas bancarias. En la Sección 6, se realiza una revisión de otras fuentes de datos de alta frecuencia en el contexto de la movilidad urbana. Finalmente, en la Sección 7 se plantean las conclusiones y consideraciones finales.

2. Google Trends y otras aplicaciones de Google

Una de las herramientas pioneras implicadas en este nuevo enfoque es Google Trends (GT), una aplicación gratuita ofrecida por el buscador de Google que indica el grado de popularidad que tiene una determinada consulta. Para ello utiliza el Search Volume Index (SVI), un índice presentado en serie temporal que refleja la evolución de la popularidad tomando valores del 0-100, dónde el valor 100 representa la mayor popularidad posible. El SVI se basa en la cuota de consultas, es decir, el volumen total de consultas para el término de búsqueda en una región geográfica concreta dividido por el número total de consultas en la región durante el periodo de tiempo examinado.

El uso de los datos de GT comenzó con investigaciones en epidemiología, uno de los estudios precursores fue Ginsberg et al. (2009), dónde estiman semanalmente la evolución de la gripe para EEUU. Más adelante, Choi & Varian (2012)¹ observaron que las búsquedas de los usuarios de Google permiten obtener información a tiempo real sobre el comportamiento de los consumidores en distintos mercados. De este modo, realizaron predicciones a corto plazo para el mercado inmobiliario, la venta de automóviles, las ventas minoristas y viajes. Para ello, Choi & Varian (2012) añaden los índices de GT a los mejores modelos de predicción que incorporan la variable endógena retardada, a continuación, evalúan a través del error absoluto medio las predicciones fuera de la muestra. Observan una mejora en las predicciones que varía sustancialmente dependiendo del caso de estudio. A partir de entonces, se incorporaron los datos de GT para la predicción de la tasa de paro y el estudio del mercado de trabajo (Askitas & Zimmerman, 2009; Fondeur & Karamé, 2013; D'Amuri & Marcucci, 2017; Baker & Fradkin, 2017), o la construcción de indicadores de confianza (Della Penna & Huang, 2010; Bontempi et al., 2016; Castelnuevo & Tran, 2017).

De acuerdo con el análisis realizado por Cebrián & Domenech (2022) y siguiendo el marco establecido por Karr, Sanil, & Banks (2006), los datos que ofrece GT son válidos para la investigación en ciencias sociales y economía, aunque pueden tener problemas de precisión que afecten a los resultados de las investigaciones. Estos autores, observan que los valores de SVI obtenidos varían en función de la fecha de obtención de los datos (Date of Retrieval). Esto indica la presencia de errores de medición, y, por consiguiente, errores

¹ Versión simplificada de sus trabajos anteriores, Choi & Varian (2009a,b)

de precisión de las series temporales. Este error – reconocido por Google - se produce debido al método utilizado por Google para seleccionar la muestra (Choi & Varian, 2012). Por otro lado, también pueden presentar problemas relacionados con la cobertura que ofrecen debido a que solo recogen las búsquedas realizadas en Google, y no en otros buscadores (Cebrián & Domenech, 2022), representando aquella parte de la población con acceso frecuente a la red (Steinmetz et al., 2014). Por otro lado, el distinto uso del lenguaje en cada país puede suponer errores en nuestras estimaciones puesto que las consultas relacionadas no se realizan de la misma manera en todos los países (Artola & Galan, 2012).

Sin embargo, los datos de GT presentan una serie de ventajas sobre las fuentes de datos convencionales. En primer lugar, es importante mencionar que se trata de datos de alta frecuencia, lo que elimina el retraso que se produce en las investigaciones que emplean fuentes de datos convencionales debido a la publicación oficial de los datos. En segundo lugar, datos como los de GT reflejan comportamientos privados. De esta manera y como refleja Owen & Wei (2021), podemos conocer aspectos culturales que los individuos no estén dispuestos a reconocer, como el sexismo.

Es importante matizar que algunos autores emplean una perspectiva distinta al enfoque longitudinal comúnmente utilizado, esto es, el enfoque de sección cruzada. Mientras que la perspectiva longitudinal analiza la evolución de un índice simple a lo largo del tiempo para unas consultas y región determinadas, el enfoque de sección cruzada permite la comparación de múltiples índices compuestos para distintas regiones y consultas (Siliverstovs & Wochner, 2018). Esta nueva perspectiva surge debido a algunos errores de medición que se pueden experimentar en el empleo del enfoque longitudinal. Debemos recordar que el SVI es un índice normalizado a la unidad, por tanto, elimina cualquier efecto sobre el volumen de búsquedas causado por el crecimiento de usuarios de Google (Carrière-Swallow & Labbé, 2013). No obstante, cabe la posibilidad de que surja un comportamiento tendencial no relacionado con la consulta realizada (Suhoy, 2009)². De este modo, aparece el enfoque de sección cruzada. En esta línea encontramos algunos trabajos que analizan el impacto sobre el bienestar de la Crisis Financiera y Económica de 2008 (Askitas & Zimmermann, 2015), las preferencias de los votantes durante el referéndum griego de 2015 (Askitas, 2015a) e Irlanda en el referéndum sobre el

² Puede suceder que la serie temporal del SVI disminuya, aunque el número bruto de consultas de búsqueda este aumentando debido a que otras consultas dispongan de mayor popularidad (Google, 2016).

matrimonio homosexual (Askitas, 2015b), las tendencias en el mercado de la vivienda (Askitas, 2016) o la demanda turística (Siliverstovs & Wochner, 2018).

Sin embargo, hasta dónde sabemos, los datos de GT apenas han sido utilizados en el estudio de la movilidad urbana. Encontramos un ejemplo en Lo et al. (2021), dónde analizan el impacto de la pandemia del covid-19 sobre atención pública y la movilidad utilizando datos de Google Trends y datos ofrecidos por Google LLC (2020) COVID-19 Community Mobility Report³. Mendolia et al. (2021) evalúa las diferencias entre la respuesta autoimpuesta e impuesta por el gobierno a la movilidad en la pandemia del COVID-19. Para ello utilizan datos obtenidos a través de Community Mobility Reports. Concluyen que la restricción voluntaria o autoimpuesta de la movilidad representa un 15% de la reducción total en la movilidad debido a la pandemia del coronavirus. Podemos encontrar otras herramientas alternativas de Google útiles para la investigación en la movilidad urbana. Un ejemplo es el caso de Google Places Application Programming Interface (API), aplicación que permite la búsqueda de información de lugares dentro de un área especificada. Otterbach et al. (2021) utilizan esta tecnología para elaborar un mapa de proximidad a supermercados y restaurantes de comida rápida con el fin de explorar el impacto de estos sobre la obesidad en Sudáfrica.

3. Twitter

Los datos de Twitter han conformado una base de datos muy relevante para los estudios de movilidad. Esta red social, es una de las más utilizadas y permite el acceso a sus datos de forma gratuita vía streaming. Generalmente, los datos de RRSS se pueden obtener a través de interfaces de programación, conocidas como API (Application Programming Interface). De esta manera, se accede a los datos de forma continua (streaming) y utilizando un código de programación en Python dónde se definen los parámetros básicos de descarga. No obstante, las posibilidades ofrecidas por la API de Twitter están restringidas por el tipo de acceso y perfil de desarrollador que tiene cada usuario de Twitter. Para obtener datos geolocalizados y más tipos de datos es necesario disponer de un perfil de desarrollador para la investigación académica. Para conseguir este perfil,

³ Disponibles en: <https://www.google.com/covid19/mobility>

debemos obtener el permiso de Twitter⁴ que se concede de acuerdo a una serie de preguntas sobre el contenido de la investigación y personal investigador.

Los datos de Twitter incorporan información muy valiosa para el investigador. Las RRSS permiten conocer aspectos relacionados con el entorno sociodemográfico de los individuos, como el lenguaje utilizado. Por otro lado, constituye una herramienta muy interesante para conocer las percepciones y opiniones de los usuarios a través del contenido de los textos publicados (Osorio-Arjona & García-Palomares, 2019). Por este motivo, encontramos distintos trabajos que validan los datos de Twitter como una herramienta eficiente y válida en el estudio la movilidad urbana (Gao et al., 2014; Lenormand et al., 2014; Lee et al., 2015; Osorio-Arjona & García-Palomares, 2019).

Observamos distintas aplicaciones de este tipo de datos en la literatura sobre movilidad urbana. Salas-Olmedo & Rojas Quezada (2017) analizan los patrones de movilidad asociados a los distintos espacios públicos con el propósito de prevenir la exclusión social en Chile. García-Palomares et al. (2018) utilizan datos de usuarios en Twitter para estudiar las dinámicas urbanas a lo largo del día y la actividad en distintas zonas de la ciudad. Observan que la actividad se reduce a lo largo del día en zonas de trabajo, educación o transporte mientras se mantiene constante en parques y se incrementa en zonas residenciales y comerciales. Osorio-Arjona & García-Palomares (2019a) realizan matrices Origen-Destino a través de una muestra de datos de Twitter geolocalizados⁵ de un periodo de 2 años, demostrando la precisión de este tipo de datos mediante comparación con la encuesta de movilidad realizada por el Consorcio Regional de Transporte de Madrid en 2014. Osorio-Arjona & García-Palomares (2019b) analizan la movilidad de los estudiantes universitarios del área Metropolitana de Madrid utilizando los datos geo-localizados de Twitter. Al mismo tiempo, combinan los datos con otras fuentes (nivel de renta (Instituto de Estadística de CAM) y tiempos de trayecto (GPS)) con el objetivo de analizar la duración del trayecto según modo de transporte y tipo de universidad.

Podemos encontrar literatura sobre otras temáticas como impactos de eventos en la movilidad urbana, la percepción sobre el servicio de transporte o estudios de accesibilidad. Condeço-Melhorado et al. (2020) utilizan los datos geolocalizados de

⁴ <https://developer.twitter.com/en/products/twitter-api/academic-research/application-info>

⁵ En Osorio-Arjona & García-Palomares (2019) describen el proceso preparación de los datos para el trabajo con ellos

Twitter para analizar las variaciones en la ciudad de Rio de Janeiro durante y después de las olimpiadas (llegada de visitantes, lugares de procedencia y distintos uso de la ciudad de residentes y visitantes). Osorio-Arjona et al. (2021) estudian la percepción de los usuarios de Twitter sobre el trayecto en el Metro de Madrid. De esta manera, evalúan la bondad de los datos de Twitter para conocer la percepción de los usuarios del transporte público. En este caso, no utilizan tuits geolocalizados, sino que utilizan un método para geolocalizar los tuits a través de palabras clave empleadas en el tuit.

4. Call Detail Records

Otra fuente de datos similar es la que se recopila a través de los Call Detail Records (CDR), que son los informes que crean las operadoras cuando se establece una llamada de teléfono. De esta manera, los CDR contienen información relacionada con la identificación temporal de los usuarios, la hora de inicio de las llamadas/mensajes y la ubicación geográfica de la estación base a la que se ha conectado el usuario durante el inicio de esa actividad de telecomunicaciones. Mediante la conexión de los usuarios a las estaciones base se puede observar cómo se mueven los interlocutores por la ciudad. Con esta fuente de datos las muestras que se obtienen son de gran tamaño. De acuerdo con Gutiérrez et al. (2019) aquellas operadoras que dispongan de una cuota de mercado del 30% tienen acceso a una muestra próxima a un tercio de la población. Sin embargo, esta fuente de datos presenta algunas limitaciones importantes. En primer lugar, las compañías telefónicas no suelen compartir los datos por motivos de confidencialidad. Por este motivo, la información relacionada con las características sociodemográficas de los usuarios y de los viajes es escasa (Gutiérrez et al., 2019). En segundo lugar, la información obtenida está limitada por el rango de cobertura de las estaciones base, lo que implica trabajar con diagramas de Voronoi. Por lo que el uso de los datos CDR puede proporcionar una resolución espacial baja que no permita analizar la distribución espacial de los usos de la tierra o zonas de transporte (Järv et al., 2017).

Los datos de teléfonos móviles se han utilizado previamente para analizar patrones de movilidad y la elaboración de matrices origen-destino (Caceres et al., 2007; Calabrese et al., 2013; Toole et al., 2015; Alexander et al., 2015). Por otro lado, estos datos permiten estudiar otros conceptos relacionados con la movilidad. En García-Albertos et al. (2019) estudian las dos dimensiones principales de la accesibilidad (tiempo de trayecto y

atractivo de los destinos) a través de datos obtenidos de Call Detail Records⁶ (CDR) y Google Maps API. Mediante este análisis, pueden conocer la variación temporal de accesibilidad al vehículo privado en la ciudad de Madrid. Por último, cabe mencionar el enfoque que plantean Dujardin et al. (2021) en una revisión del potencial de estos datos para la planificación urbana en adaptación al cambio climático. Estos autores exponen la utilidad de estos datos en las ciudades en vías de desarrollo, de tal manera que permiten el estudio de la exposición, susceptibilidad, resiliencia y capacidad adaptativa de estas ciudades a las consecuencias del cambio climático.

5. Tarjetas bancarias

El uso de tarjetas de débito o crédito también puede constituir una fuente de datos relevante. Estos datos provienen de la actividad de los consumidores al pagar en terminales punto de venta. Gracias a la geolocalización de estos terminales, estos datos constituyen una alternativa interesante para el estudio de la distribución espacial y temporal del gasto (Carpio-Pinedo et al., 2022). De manera similar a los datos CDR, los datos de transacciones son capaces de recopilar muestras de gran tamaño debido a la expansión del uso de tarjetas de débito/crédito. De acuerdo con los datos ofrecidos por el Banco Central Europeo en 2019, en la zona euro el 24% de las transacciones se realizan a través de tarjetas bancarias (European Central Bank, 2020). No obstante, la investigación con este tipo de datos es relativamente escasa. Esto es debido a la naturaleza de la información de las tarjetas bancarias, de contenido sensible, privada, confidencial y protegida por la legislación. A pesar de esta limitación, encontramos algunos estudios con estos datos que estudian las características socioeconómicas de las regiones (Hashemian et al., 2017), o la resiliencia económica ante desastres naturales (Martinez et al., 2016).

En relación a la movilidad, también encontramos distintos artículos. Lenormand et al. (2015) analizan los datos de tarjetas bancarias en Barcelona y Madrid a través de la geolocalización de las transacciones. Comprueban que los patrones de movilidad varían en función de la edad, género y ocupación. Carvalho et al. (2021) siguen el impacto de la pandemia del coronavirus en España a través de los datos que ofrecen las transacciones comerciales de tarjetas de crédito y débito del banco BBVA. De esta manera, construyen un indicador alternativo para medir el impacto de la pandemia sobre el consumo y

⁶ En (Gutiérrez et al., 2019) encontramos una guía para trabajar con estos datos.

demuestran la validez de estos datos como indicador alternativo para la movilidad geográfica. Campos-Vazquez & Esquivel (2021) estudian la relación entre el consumo en terminales punto de venta (TPV, point of sale POS) y la movilidad geográfica durante la pandemia en México. Para estudiar movilidad geográfica, utilizan datos de Apple Corporation (2020) y Google LLC (2020)⁷. Concluyen que en los países en desarrollo el gasto podría verse más afectado por las restricciones de movilidad, de modo que un indicador de movilidad puede funcionar como un indicador a tiempo real del consumo. Carpio-Pinedo et al. (2022) elaboran un análisis de la geografía del gasto en Madrid a partir de los datos obtenidos de las transacciones con tarjeta bancaria. El objetivo principal es entender la dispersión espacial de cada sector. Aparicio et al. (2022) analizan los patrones de gasto de los turistas utilizando los datos de transacciones bancarias (Banco BBVA). Concluyen que el gasto de los turistas se concentra en el centro de la ciudad, observando una disminución del gasto conforme aumenta la distancia al centro.

6. Otras fuentes de datos y estudios

A parte de las fuentes de datos mencionadas previamente, distintas aplicaciones móviles y tecnológicas permiten el acceso a información para el estudio de la movilidad urbana. Un ejemplo es la información que proporciona la empresa *TomTom* a través de *TomTom Speed Profiles*. Esta herramienta recoge la información sobre la velocidad en los distintos tramos de la red vial cada 5 minutos. Estos datos son muy relevantes para estudiar la accesibilidad desde una perspectiva dinámica (Gutiérrez et al., 2020). En este sentido, encontramos la investigación de Moya-Gómez et al. (2018) dónde analizan la accesibilidad⁸ con datos de *TomTom* y Twitter. Observan que las condiciones de accesibilidad varían significativamente a lo largo del día.

Con este tipo de datos, encontramos literatura sobre movilidad ciclista. Romanillos et al. (2015) realizan una revisión sobre lo investigado en movilidad ciclista con los nuevos datos que dividen en tres categorías: datos de *global positioning system* (GPS), *live point data* y *journey data*. Romanillos & Austwick (2015) presentan los objetivos, desarrollo y resultado de la iniciativa *Huella Ciclista de Madrid*⁹. Así, repasan los trabajos publicados

⁷ Los datos de Apple Corporation (2020) ya no están disponibles, por otro lado, los datos de Google LLC (2020) presentarán su último informe de movilidad en octubre de 2022.

⁸ Medida, en este caso, como la efectividad de las redes de transporte y el atractivo de los distintos destinos.

⁹ Una iniciativa impulsada con el objetivo de conocer las rutas ciclistas de voluntarios en Madrid.

de acuerdo a las distintas fuentes: datos de teléfonos móviles, redes sociales, comunidades de fotos geolocalizadas, datos de transacciones bancarias, tarjetas de transporte inteligentes o navegadores de automóviles. Por otro lado, consiguen elaborar un mapa para representar la distribución del flujo real de ciclistas en Madrid. Romanillos et al. (2018) estudian los flujos de ciclistas a partir de 250.000 rutas GPS de los servicios de bicicletas compartidas en Madrid con el objetivo de comprender la distribución espacial de los trayectos que siguen los ciclistas. Romanillos & Gutiérrez (2020) estudian la velocidad de los ciclistas en función de distintos parámetros relacionados con aspectos locales (infraestructuras, tráfico, número de semáforos, intersecciones...), género, edad, propósito del trayecto y las condiciones meteorológicas.

Asimismo, observamos literatura sobre un tema de especial relevancia en el contexto de la movilidad urbana sostenible, la movilidad compartida en la ciudad. Arias-Molinares et al. (2021a) ponen el foco en los trayectos (dimensión espacio-temporal) de los vehículos de movilidad compartida, específicamente los ciclomotores, en 7 ciudades de España¹⁰. Para este estudio, utilizan datos GPS instalados en los ciclomotores de la app *Moving*. En definitiva, se trata de un estudio sobre los patrones que siguen los usuarios de servicios de micromovilidad compartida (usuarios de ciclomotores) desde una perspectiva espacio-temporal. Arias-Molinares et al. (2021b) analizan los servicios de micromovilidad urbana¹¹ poniendo énfasis en los servicios de micromovilidad sin estación base. El objetivo principal es captar las diferencias entre los usuarios de servicios de micromovilidad con estación base y aquellos que funcionan sin estación. De esta manera, analizan las diferencias sociodemográficas entre los distintos individuos, entre ambos tipos de servicios y si la relación es de competitividad o complementariedad. En definitiva, realizan un análisis de la interacción entre los clientes de los dos tipos de servicios. Es importante mencionar el gran impacto de las aplicaciones móviles sobre la movilidad urbana y el ritmo de crecimiento de las mismas. Por este motivo, surge una necesidad de agrupar todas estas aplicaciones en una única plataforma, las llamadas apps de *Mobility as a Service* (MaaS) que sea capaz de gestionar y pagar cada uno de los múltiples servicios de movilidad. Arias-Molinares & García-Palomares (2020) contribuyen a este tema analizando cómo está contribuyendo la oferta de movilidad compartida en Madrid al desarrollo de MaaS.

¹⁰ Madrid, Valencia, Sevilla, Zaragoza, Málaga, Córdoba y Cádiz

¹¹ Uso compartido de bicicletas, patinetes, etc.

De gran interés son las tarjetas inteligentes o *Smart cards* del transporte urbano. Estas tarjetas son capaces de aglutinar grandes muestras de datos debido a la expansión en su uso, desplazando otros medios de pago del transporte público. De este modo, esa información ayuda a conocer los viajes y trayectos que realizan los ciudadanos en transporte público. Sin embargo, su uso es complejo puesto que cabe la posibilidad de cometer un error en la diferenciación de viaje y etapa. Es decir, debemos ser capaces de diferenciar distintas etapas que pueden componer el trayecto en una ciudad como es el caso de los transbordos (Gutiérrez et al., 2020). En Gschwender et al. (2016) muestran la utilidad de estos datos a partir del cálculo de matrices origen destino, velocidad de los autobuses y el análisis de calidad del servicio de transporte público en Santiago de Chile.

7. Conclusiones

La evolución de las ciudades y el avance tecnológico han provocado cambios en el comportamiento de la movilidad urbana. Estos cambios se resumen en mayor número de viajes, trayectos más largos y de mayor dispersión. Por este motivo, es importante estudiar de manera eficiente la movilidad urbana con el objetivo de conducirla hacia una movilidad sostenible. Para ello, las fuentes de datos convencionales resultan infructuosas debido a su construcción a partir de métodos de procesamiento y recolección costosos, así como la baja resolución espacio-temporal que presentan. En este contexto, observamos distintas herramientas que, a través del *Big Data*, proporcionan datos de alta frecuencia que permiten tener acceso a información para estudiar la movilidad urbana desde distintas perspectivas.

En este sentido, las primeras investigaciones, surgieron de la mano de Google Trends. Una herramienta gratuita ofrecida por el buscador de Google que permite conocer el grado de popularidad de determinadas consultas. La literatura ha demostrado que estos datos son útiles para el estudio de ciencias sociales y economía. Sin embargo, en el campo de la movilidad urbana no encontramos apenas contribuciones y aportaciones. Al contrario, los datos ofrecidos por Twitter han servido de base para distintas investigaciones en el ámbito de la movilidad. De esta manera, encontramos contribuciones sobre la percepción de los usuarios del transporte público, el impacto de los eventos en la movilidad o la elaboración de matrices origen-destino.

Por otro lado, en el estudio de la movilidad urbana aparecen otras fuentes de datos de alta frecuencia como la ofrecida por las tarjetas bancarias o la obtenida a través de los servicios de telecomunicaciones a través de los Call Detail Records. Sin embargo, estas alternativas presentan limitaciones importantes debido a aspectos relacionados con la confidencialidad y privacidad de los usuarios de los que se recoge pasivamente este tipo de información. También están disponibles otras fuentes de datos como las aplicaciones móviles que utilizan tecnología GPS, las cuales resultan muy relevantes en la movilidad ciclista o la movilidad compartida. En definitiva, el aprovechamiento del Big Data y las nuevas fuentes de datos de alta frecuencia han comenzado su incursión en la investigación sobre movilidad urbana dibujando un futuro prometedor en este campo.

8. Referencias

- Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation research part c: emerging technologies*, 58, 240-250.
- Aparicio, D., Hernández Martín-Caro, M. S., García-Palomares, J. C., & Gutiérrez, J. (2022). Exploring the spatial patterns of visitor expenditure in cities using bank card transactions data. *Current Issues in Tourism*, 25(17), 2770-2788. <https://doi.org/10.1080/13683500.2021.1991898>
- Apple Inc. (2020). “Apple Mobility Trends reports”. <https://www.apple.com/covid19/mobility/>
- Artola, C., & Martínez-Galán, E. (2012). Tracking the future on the web: construction of leading indicators using internet searches. Banco de España Occasional Paper, (1203).
- Arias-Molinares, D., & Carlos García-Palomares, J. (2020). Shared mobility development as key for prompting mobility as a service (MaaS) in urban areas: The case of Madrid. *Case Studies on Transport Policy*, 8(3), 846-859. <https://doi.org/10.1016/j.cstp.2020.05.017>
- Arias-Molinares, D., Julio, R., García-Palomares, J. C., & Gutiérrez, J. (2021). Exploring micromobility services: Characteristics of station-based bike-sharing users and their relationship with dockless services. *Journal of Urban Mobility*, 1(October 2020), 100010. <https://doi.org/10.1016/j.urbmob.2021.100010>

- Arias-Molinares, D., Romanillos, G., García-Palomares, J. C., & Gutiérrez, J. (2021). Exploring the spatio-temporal dynamics of moped-style scooter sharing services in urban areas. *Journal of Transport Geography*, 96(February). <https://doi.org/10.1016/j.jtrangeo.2021.103193>
- Askatas, N. (2015a). Calling the Greek Referendum on the nose with Google Trends. Available at SSRN 2633443.
- Askatas, N. (2015b). Predicting the Irish 'Gay Marriage' Referendum. Available at SSRN 2674243.
- Askatas, N. (2016). Trend-spotting in the housing market. *Cityscape*, 18(2), 165-178.
- Askatas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting.
- Askatas, N., & Zimmermann, K. F. (2015). Health and well-being in the great recession. *International Journal of Manpower*.
- Baker, S. R., & Fradkin, A. (2017). The impact of unemployment insurance on job search: Evidence from Google search data. *Review of Economics and Statistics*, 99(5), 756-768.
- Belloc, I., Gimenez-Nadal, J. I., & Molina, J. A. (2022a). The effect of diesel tax rates on the daily commuting of US workers: An effective instrument to promote sustainable mobility? IZA DP No. 15416.
- Belloc, I., Gimenez-Nadal, J. I., & Molina, J. A. (2022b). The gasoline price and the commuting behavior: Towards sustainable modes of transport. GLO Discussion Paper Series 1130.
- Bontempi, M. E., Golinelli, R., & Squadrani, M. (2016). A new index of uncertainty based on internet searches: A friend or foe of other indicators?
- Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin–destination data from a mobile phone network. *IET Intelligent Transport Systems*, 1(1), 15-26.
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira Jr, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation research part C: emerging technologies*, 26, 301-313.
- Campos-Vazquez, R. M., & Esquivel, G. (2021). Consumption and geographic mobility

in pandemic times. Evidence from Mexico. *Review of Economics of the Household*, 19(2), 353-371.

- Carpio-Pinedo, J., Romanillos, G., Aparicio, D., Martín-Caro, M. S. H., García-Palomares, J. C., & Gutiérrez, J. (2022). Towards a new urban geography of expenditure: Using bank card transactions data to analyze multi-sector spatiotemporal distributions. *Cities*, 131(November 2021). <https://doi.org/10.1016/j.cities.2022.103894>
- Carrière-Swallow, Y., & Labbé, F. (2013). Nowcasting with Google Trends in an emerging market. *Journal of Forecasting*, 32(4), 289-298.
- Carvalho, V. M., Garcia, J. R., Hansen, S., Ortiz, Á., Rodrigo, T., Rodríguez Mora, J. V., & Ruiz, P. (2021). Tracking the COVID-19 crisis with high-resolution transaction data. *Royal Society Open Science*, 8(8), 210218.
- Castelnuovo, E., & Tran, T. D. (2017). Google it up! a google trends-based uncertainty index for the united states and australia. *Economics Letters*, 161, 149-153.
- Cebrián, E., & Domenech, J. (2022). Is Google Trends a quality data source?. *Applied Economics Letters*, 1-5.
- Choi, H. and Varian, H. (2009a), 'Predicting the Present with Google Trends', Technical report, Google.
- Choi, H., & Varian, H. (2009b). Predicting initial claims for unemployment benefits. Google Inc, 1, 1-5.
- Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic record*, 88, 2-9.
- Condeço-Melhorado, A., Mohino, I., Moya-Gómez, B., & García-Palomares, J. C. (2020). The Rio olympic games: A look into city dynamics through the lens of Twitter data. *Sustainability (Switzerland)*, 12(17), 1-16. <https://doi.org/10.3390/su12177003>
- D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, 33(4), 801-816.
- Della Penna, N., & Huang, H. (2010). Constructing consumer sentiment index for US using Google searches (No. 2009-26).

- Dujardin, S., Jacques, D., Steele, J., & Linard, C. (2020). Mobile phone data for urban climate change adaptation: Reviewing applications, opportunities and key challenges. *Sustainability*, 12(4), 1501.
- Echeverría, L., Gimenez-Nadal, J. I., & Molina, J. A. (2021). Carpooling: User profiles and well-being. IZA DP No. 14736.
- Echeverría, L., Gimenez-Nadal, J. I., & Molina, J. A. (2022a). Who uses green mobility? Exploring profiles in developed countries. *Transportation research part A: policy and practice*, 163, 247-265.
- Echeverría, L., Gimenez-Nadal, J. I., & Molina, J. A. (2022b). Green mobility and well-being. *Ecological Economics*, 195, 107368.
- Echeverría, L., Gimenez-Nadal, J. I., & Molina, J. A. (2022c). Active commuting and the health of workers. IZA DP No. 15572.
- Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210), 1243089.
- Fondeur, Y., & Karamé, F. (2013). Can Google data help predict French youth unemployment?. *Economic Modelling*, 30, 117-125.
- Gao, S., Yang, J. A., Yan, B., Hu, Y., Janowicz, K., & McKenzie, G. (2014, September). Detecting origin-destination mobility flows from geotagged tweets in greater Los Angeles area. In *Proceedings of the eighth international conference on geographic information science* (pp. 1-4).
- García-Albertos, P., Picornell, M., Salas-Olmedo, M. H., & Gutiérrez, J. (2019). Exploring the potential of mobile phone records and online route planners for dynamic accessibility analysis. *Transportation Research Part A: Policy and Practice*, 125(February 2018), 294-307. <https://doi.org/10.1016/j.tra.2018.02.008>
- García-Palomares, J. C., Salas-Olmedo, M. H., Moya-Gómez, B., Condeço-Melhorado, A., & Gutiérrez, J. (2018). City dynamics through Twitter: Relationships between land use and spatiotemporal demographics. *Cities*, 72(September), 310-319. <https://doi.org/10.1016/j.cities.2017.09.007>
- Giménez, J.I. and Molina, J.A. (2014). Commuting time and labour supply in the Netherlands: a time use study. *Journal of Transport Economics and Policy*, 48 (3), 409-426.

- Giménez, J.I. and Molina, J.A. (2016). Commuting time and household responsibilities: evidence using propensity score matching. *Journal of Regional Science*, 56, 332-359. <https://doi.org/10.1111/jors.12243>.
- Giménez-Nadal, J.I. and Molina, J.A. (2019a). Daily feelings of US workers and commuting time. *Journal of Transport & Health*, 12, 21-33. <https://doi.org/10.1016/j.jth.2018.11.001>.
- Giménez-Nadal, J.I. and Molina, J.A. (2019b). Green commuting and gasoline taxes in the United States. *Energy Policy*, 132, 324-331. <https://doi.org/10.1016/j.enpol.2019.05.048>.
- Giménez, J.I. , Molina, J.A. and Velilla, J. (2018). The commuting behavior of workers in the United States: differences between the employed and the self-employed. *Journal of Transport Geography*, 66, 19-29. <https://doi.org/10.1016/j.jtrangeo.2017.10.011>.
- Giménez-Nadal, J.I., Molina, J.A. and Velilla, J. (2019c). Modeling commuting time in the US: Bootstrapping techniques to avoid overfitting. *Papers in Regional Science*, 98(4), 1667-1684. <https://doi.org/10.1111/pirs.12424>.
- Giménez-Nadal, J.I., Molina, J.A. and Velilla, J. (2020). Commuting and self-employment in Western Europe. *Journal of Transport Geography*, 88, 102856. <https://doi.org/10.1016/j.trangeo.2020.102856>.
- Giménez-Nadal, J.I., Molina, J.A. and Velilla, J. (2021). Two-way commuting: Asymmetries from time use surveys. *Journal of Transport Geography*, 95, 103146. <https://doi.org/10.1016/j.trangeo.2021.103146>.
- Giménez-Nadal, J.I., Molina, J.A. and Velilla, J. (2022a). Trends in commuting time of European workers: A cross-country analysis. *Transport Policy*, 116, 327-342. <https://doi.org/10.1016/j.tran-pol.2021.12.016>.
- Giménez-Nadal, J.I., Molina, J.A. and Velilla, J. (2022b). Commuting time and sickness absence of US workers. *Empirica*, forthcoming. <https://doi.org/10.1007/s10663-022-09534-z>.
- Giménez-Nadal, J.I., Gracia-Lázaro, C. and Molina, J.A. (2022c). Increasing the use of public bicycles: efficiency and demand. *Economic Analysis and Policy*, forthcoming. <https://doi.org/10.1016/j.eap.2022.09.015>.

- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232), 1012-1014.
- Google, 2016. Google Trends help – Trends graphs and forecasts, Retrieved from <https://support.google.com/trends/answer/4355164?hl=en&reftopic=4365530> (20.05.16, 14:51).
- Google LLC. (2020). Google COVID-19 Community Mobility Reports. <https://www.google.com/covid19/mobility/>
- Gschwender, A., Munizaga, M., & Simonetti, C. (2016). Using smart card and GPS data for policy and planning: The case of Transantiago. *Research in Transportation Economics*, 59, 242-249. <https://doi.org/10.1016/j.retrec.2016.05.004>
- Gutiérrez, J., Benítez, C., Leaño, J., García-Palomares, J., Condeço-Melhorado, A., Mojica, C., Scholl, L., Adler, V., Vera, F., Moya-Gómez, B., & Romanillos, G. (2019). *Cómo aplicar big data en la planificación del transporte urbano: El uso de datos de telefonía móvil en el análisis de la movilidad*. <https://doi.org/10.18235/0002009>
- Gutiérrez Puebla, J., Benítez, C., García Palomares, J. C., Romanillos Arroyo, G., Rubinstein, E., Leaño, J. M., Ribeiro, K. M., Scholl, L., Moya Gómez, B., & Condeço Melhorado, A. (2020). *Cómo aplicar Big Data en la planificación del transporte: El uso de datos de GPS en el análisis de la movilidad urbana*. 150. <https://doi.org/10.18235/0002487>
- Hashemian, B., Massaro, E., Bojic, I., Murillo Arias, J., Sobolevsky, S., & Ratti, C. (2017). Socioeconomic characterization of regions through the lens of individual financial transactions. *PloS one*, 12(11), e0187031.
- Järv, O., Tenkanen, H., & Toivonen, T. (2017). Enhancing spatial accuracy of mobile phone data using multi-temporal dasymetric interpolation. *International Journal of Geographical Information Science*, 31(8), 1630-1651.
- Karr, A. F., Sanil, A. P., & Banks, D. L. (2006). Data quality: A statistical perspective. *Statistical Methodology*, 3(2), 137-173.
- Lee, J. H., Gao, S., & Goulias, K. G. (2015). Can Twitter data be used to validate travel demand models. In 14th international conference on travel behaviour research.

- Lenormand, M., Picornell, M., Cantú-Ros, O. G., Tugores, A., Louail, T., Herranz, R., ... & Ramasco, J. J. (2014). Cross-checking different sources of mobility information. *PloS one*, 9(8), e105184.
- Lenormand, M., Louail, T., Cantú-Ros, O. G., Picornell, M., Herranz, R., Arias, J. M., ... & Ramasco, J. J. (2015). Influence of sociodemographic characteristics on human mobility. *Scientific reports*, 5(1), 1-15.
- Lo, K. L., Liu, H., Yang, M., & Mi, J. J. (2021). The relationship between public attention and COVID-19: Evidence from the big data analysis of Google trends. *Applied Economics Letters*, 1-8.
- Martínez, E. A., Rubio, M. H., Martínez, R. M., Arias, J. M., Patane, D., Zerbe, A., ... & Luengo-Oroz, M. (2016). Measuring economic resilience to natural disasters with big economic transaction data. *arXiv preprint arXiv:1609.09340*.
- Mendolia, S., Stavrino, O., & Yerokhin, O. (2021). Determinants of the community mobility during the COVID-19 epidemic: The role of government regulations and information. *Journal of Economic Behavior & Organization*, 184, 199-231.
- Molina, J.A., Giménez-Nadal, J.I. and Velilla, J. (2020). Sustainable commuting: Results from a social approach and international evidence on carpooling. *Sustainability*, 12(22), 9587. <https://doi.org/10.3390/su12229587>.
- Moya-Gómez, B., Salas-Olmedo, M. H., García-Palomares, J. C., & Gutiérrez, J. (2018). Dynamic Accessibility using Big Data: The Role of the Changing Conditions of Network Congestion and Destination Attractiveness. *Networks and Spatial Economics*, 18(2), 273-290. <https://doi.org/10.1007/s11067-017-9348-z>
- Osorio-Arjona, J., & García-Palomares, J. C. (2017). Nuevas fuentes y retos para el estudio de la movilidad urbana. *Cuadernos Geográficos*, 56(3), 247-267. <https://doi.org/10.30827/CUADGEO.V56I3.5352>
- Osorio-Arjona, J., & García-Palomares, J. C. (2019a). Social media and urban mobility: Using twitter to calculate home-work travel matrices. *Cities*, 89(January), 268-280. <https://doi.org/10.1016/j.cities.2019.03.006>
- Osorio-Arjona, J., & García-Palomares, J. C. (2019b). Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter. *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, (24),

- Osorio-Arjona, J., Horak, J., Svoboda, R., & García-Ruíz, Y. (2021). Social media semantic perceptions on Madrid Metro system: Using Twitter data to link complaints to space. *Sustainable Cities and Society*, 64(March 2020), 102530. <https://doi.org/10.1016/j.scs.2020.102530>
- Otterbach, S., Oskorouchi, H. R., Rogan, M., & Qaim, M. (2021). Using Google data to measure the role of Big Food and fast food in South Africa's obesity epidemic. *World Development*, 140, 105368.
- Owen, A. L., & Wei, A. (2021). Sexism, household decisions, and the gender wage gap. *Labour Economics*, 72, 102062.
- Romanillos, G., & Austwick, M. Z. (2015). *Madrid cycle track : visualizing the cyclable city*. 5647(October). <https://doi.org/10.1080/17445647.2015.1088901>
- Romanillos, G., Austwick, M. Z., Ettema, D., Kruijf, J. De, Romanillos, G., Austwick, M. Z., Ettema, D., & Kruijf, J. De. (2015). *Big Data and Cycling*. 1647(October). <https://doi.org/10.1080/01441647.2015.1084067>
- Romanillos, G., & Gutiérrez, J. (2020). Cyclists do better. Analyzing urban cycling operating speeds and accessibility. *International Journal of Sustainable Transportation*, 14(6), 448-464. <https://doi.org/10.1080/15568318.2019.1575493>
- Romanillos, G., Moya-Gómez, B., Zaltz-Austwick, M., & Lamíquiz-Daudén, P. J. (2018). The pulse of the cycling city: visualising Madrid bike share system GPS routes and cycling flow. *Journal of Maps*, 14(1), 34-43. <https://doi.org/10.1080/17445647.2018.1438932>
- Salas-Olmedo, M. H., & Rojas Quezada, C. (2017). The use of public spaces in a medium-sized city: From Twitter data to mobility patterns. *Journal of Maps*, 13(1), 40-45. <https://doi.org/10.1080/17445647.2017.1305302>
- Suhoy, T. (2009). Query indices and a 2008 downturn: Israeli data (No. 2009.06). Bank of Israel.
- Steinmetz, S., Bianchi, A., Tijdens, K., & Biffignandi, S. (2014). Improving web survey quality. Online panel research: A data quality perspective, 273-298.
- Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C.

(2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58, 162-177.