

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

REDES SOCIALES GEOLOCALIZADAS Y COVID-19: ANÁLISIS DE LA ACTIVIDAD ESPACIOTEMPORAL DE LOS USUARIOS DE TWITTER DE ESPAÑA DURANTE LA PANDEMIA

Joaquín Osorio Arjona  

Departamento de Población. Centro de Ciencias Humanas y Sociales CSIC
C/ Albasanz 26, 28037, Madrid, España
joaquin.osorio@cchs.csic.es

RESUMEN

Este trabajo busca mostrar Twitter como fuente de datos alternativa para el estudio de la pandemia causada en España por el virus COVID-19. Para ello, se plantea un análisis de la distribución espacial y temporal de una muestra de usuarios obtenida en tres periodos diferentes del año 2020, y se comparan los resultados obtenidos con los mismos periodos del año anterior a la pandemia. También se elabora un análisis espaciotemporal del uso de términos asociados con la enfermedad, y se realizan mapas de calor para observar el impacto causado en dos ciudades de relevante peso turístico. Los resultados obtenidos indican una fuerte disminución del número de usuarios que publican *tweets* geolocalizados en todo el país a lo largo del año 2020, especialmente en la segunda mitad del año y en las provincias del interior peninsular. También se observa de forma menos pronunciada una disminución del número de usuarios en áreas costeras y provincias orientadas al sector turístico.

Palabras clave: Twitter; COVID-19; movilidad; vigilancia; geolocalización; distribución espaciotemporal

GEOLOCATED SOCIAL NETWORKS AND COVID-19: ANALYSIS OF THE TEMPORAL AND SPATIAL ACTIVITY OF TWITTER USERS IN SPAIN DURING THE PANDEMIC

ABSTRACT

This work seeks to show Twitter as an alternative data source for the study of the pandemic caused by the COVID-19 virus in Spain. For this work, an analysis of the spatial and temporal distribution of the sample of users obtained in three different periods of the year 2020 is proposed, and then the obtained results are compared with the same periods of the year prior to the pandemic. A space-time analysis of the use of terms associated with the disease is also elaborated, and heat maps are made to observe the impact

Recepción: 16/09/2022

Editor al cargo: Dr. Joan Cristian Padró

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

Aceptación definitiva: 03/11/2022

www.geofocus.org

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

caused in the activity of two cities of relevant tourist weight. The obtained results indicate a sharp decrease in the number of users who publish geolocated *tweets* in the country throughout 2020, especially in the second half of the year and in the interior provinces of the peninsula. A less pronounced decrease in the number of users is also observed in coastal areas and provinces oriented to the tourism sector.

Keywords: Twitter; COVID-19; mobility; surveillance; geolocation; spatial-temporal distribution

1. Introducción

En el año 2020, el planeta ha sufrido la mayor crisis sanitaria vivida en los últimos 100 años. A día 1 de noviembre de 2022, se han reportado 630.452.398 casos de la enfermedad COVID-19 (SARS-CoV 2) y 6.590.206 muertes en todo el mundo (Google, 2020). Esta enfermedad se originó en Wuhan, China en diciembre de 2019, y no fue hasta principios de 2021 cuando se empezaron a suministrar vacunas para combatir el virus. La pandemia ha cambiado radicalmente la vida diaria de las personas y sus patrones de movilidad. En el año 2020 se ha promovido constantemente la distancia social, el uso de mascarillas, el teletrabajo y salir del hogar lo estrictamente necesario (World Health Organization, 2020). La movilidad urbana en concreto suele jugar un rol clave en la transmisión de enfermedades. En esta situación, muchos ciudadanos han seguido desplazándose de forma necesaria al trabajo, ocasionándose varias veces la saturación de los sistemas de transporte público, lo cual ha puesto en riesgo la salud de las personas. Para solucionar estos problemas, se antoja de gran valor estudiar con el mayor detalle posible la movilidad ciudadana para poder facilitar medidas de movilidad sostenible mientras se busca mantener la distancia social y evitar aglomeraciones.

Monitorizar la difusión de una enfermedad epidémica es una tarea muy importante. Sin embargo, trabajar con muestras grandes de datos de individuos infectados y no infectados suele conllevar un alto coste económico. Estos datos pueden tardar meses en ser procesados, y los resultados obtenidos difícilmente llegan a los ciudadanos, quienes normalmente son incapaces de seguir el trabajo de la comunidad sanitaria (Sousa *et al.*, 2018). Hace falta una infraestructura avanzada y el desarrollo de nuevos modelos de datos espaciales que puedan incorporar la movilidad de individuos en riesgo e informar a tiempo real, y que permita mostrar a la población las acciones que toman los agentes de salud (Lamos, De Bie, & Cristianini, 2010; Souza, Neill, Assuncao, & Meira, Jr., 2019). Por tanto, es necesario disponer de datos que se puedan actualizar constantemente, rápidamente y a tiempo casi real.

A las pocas semanas de comenzar la pandemia causada por el COVID-19, las discusiones sobre la necesidad de utilizar grandes cantidades de datos para elaborar visualizaciones de los flujos de movilidad se convirtieron en un tema recurrente en los medios de noticias, y las agencias gubernamentales y las compañías empezaron a usar datos geolocalizados basados en Big Data para monitorear el avance de la enfermedad (Frith & Saker, 2020; Poom, Järv, Zook, & Toivonen, 2020). La Organización Mundial de la Salud (OMS) se ha apoyado en estas nuevas fuentes de datos y en los Sistemas de Información Geográfica (SIG) para monitorear la evolución de la enfermedad en plataformas webs interactivas (Kamel Boulos & Geraghty, 2020).

Cada vez hay un mayor número de investigaciones que usan Big Data para identificar y caracterizar enfermedades infecciosas en el contexto de vigilancia (Eysenbach, 2011). La salud pública se ha beneficiado

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

de esta vigilancia, que ha sido crucial para detectar enfermedades en la sociedad de la información moderna (Hartley *et al.*, 2010). Twitter en particular es una fuente de datos valiosa para obtener información sobre el comportamiento de los ciudadanos en temas sanitarios. Esta red social permite obtener resultados comparables a encuestas oficiales y además el valor semántico de los textos de los *tweets* puede dar información adicional sobre el impacto de una enfermedad en la sociedad (Zhang *et al.*, 2020). Twitter además permite la posibilidad de descargar de forma gratuita grandes muestras de datos con alto detalle espacial y temporal (Huang, Li, Jiang, Li, & Porter, 2020). Los *tweets* tienen geometría en forma de punto, lo que permite el diseño de bases de datos espaciales que pueden ser fácilmente actualizables, y el fácil tratamiento, análisis y visualización de los datos en SIG (Blanford, Huang, Savelyev, & MacEachren, 2015; Gutiérrez-Puebla *et al.*, 2016; Lansley & Longley, 2016; Osorio Arjona & García Palomares, 2020).

El objetivo de este trabajo es mostrar el valor de Twitter como fuente de datos para la monitorización de problemas sanitarios y su influencia en la movilidad. En este caso, los datos de Twitter son aplicados para estudiar la distribución espacial y temporal de sus usuarios en España durante el año 2020. Los resultados obtenidos son comparados con la distribución de usuarios dada en el año anterior. Además, se analiza la distribución temporal y espaciotemporal de los usuarios de una muestra según palabras claves asociadas al COVID-19, con el objetivo de ver la evolución del uso de su terminología y del impacto de la enfermedad en la red social a lo largo del año. También se elaboran mapas de calor para visualizar el impacto causado por el COVID-19 durante el verano en las áreas centrales de dos ciudades importantes en el ámbito del turismo como son Madrid y Málaga.

Esta investigación está dividida en cinco apartados. Después de esta introducción, el apartado 2 expondrá un estado del arte respecto al uso de datos de Twitter en el estudio y monitorización espacial de enfermedades. El apartado 3 presentará el área de estudio, datos y metodología empleados para este trabajo, mientras el apartado 4 recogerá y explicará los resultados obtenidos. Finalmente, el apartado 5 establecerá una serie de conclusiones y futuras líneas de investigación.

2. Estado del arte

El análisis espacial de datos ha tenido una gran importancia histórica en el análisis y monitorización de enfermedades con el objetivo de entender las causas y establecer acciones preventivas o medidas para combatir pandemias. Uno de los grandes hitos de la historia de la geografía y la cartografía fue la elaboración del mapa del cólera por el médico John Snow en 1854. En ese año Londres estaba azotada por un brote de la enfermedad del cólera. Snow cartografió el sistema de suministro de agua de Londres e identificó la relación espacial entre la distribución de la enfermedad y la ubicación de una bomba de uso público ubicada en Broad Street (Cerdeira L & Valdivia C, 2007; Kamel Boulos & Geraghty, 2020).

En la última década se ha aprovechado el gran detalle espacial y temporal de los datos de Twitter para estudiar las consecuencias que provoca la distribución espacial de determinadas enfermedades. Además, la facilidad de monitorizar los comentarios y preocupaciones de los usuarios a tiempo real, sugieren su utilidad para evaluar el estado de salud de la población (Barros *et al.*, 2018). La mayoría de las investigaciones que han usado datos de Twitter en temas de salud se han ceñido al contenido de los textos de los *tweets*, pero cada vez hay un mayor número de trabajos que también tiene en cuenta las coordenadas espaciales con las que cuentan los *tweets* geolocalizados. De forma general, Barros *et al.* (2018) han investigado la distribución

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

espacial de *tweets* durante un periodo de seis meses en todo el mundo para identificar en que lugares se hace un mayor número de menciones de enfermedades. Como resultado encontraron que al contrario de lo que se esperaba, los hospitales y aeropuertos no son lugares adecuados para recopilar *tweets* cuyo tema son las enfermedades. Además, también observaron que los *tweets* con menciones a enfermedades son principalmente comentarios de noticias, y que tiende a haber una representación de enfermedades no comunicables en los textos en contraste a enfermedades infecciosas. Khan *et al.* (2018) buscaron identificar posibles problemas de salud en un evento de masas celebrado en Toronto (Canadá). Encontraron una correlación entre los *tweets* de una muestra y el único problema de salud detectado durante el evento, los golpes de calor.

En la investigación de enfermedades más específicas, Cai *et al.* (2020) utilizaron *tweets* geolocalizados en Indiana (Estados Unidos) para estudiar el brote del virus de la inmunodeficiencia humana (HIV-1) en 2015. Lampos *et al.* (2010) descargaron *tweets* en 49 áreas urbanas de Reino Unido e identificaron mensajes relacionados con la gripe a partir de una serie de palabras clave. Young, Mercer, Weiss, Torrone, & Aral (2018) filtraron 8500 *tweets* geolocalizados que contenían palabras claves relacionadas con la enfermedad de la sífilis, y encontraron una correlación positiva entre los *tweets* de la muestra usada y los casos de sífilis reportados en Estados Unidos en 2012 y 2013. Zhang *et al.* (2020) utilizaron tres bases de datos de *tweets* publicados desde 2014 a 2018 en Estados Unidos para identificar 122 temas relacionados con el virus del papiloma humano (HPV), descubriendo que la mayoría de estos temas entre la información promocional y las discusiones de los consumidores están correlacionados. Souza *et al.* (2019) recopilaron *tweets* publicados en Brasil en 2015 para identificar áreas con un alto riesgo de ser afectadas por el virus del dengue, demostrando que Twitter puede ser una fuente de datos alternativa y eficaz para detectar zonas de alto riesgo, especialmente en zonas urbanas, pero también en lugares con un menor nivel de urbanización. Sousa *et al.* (2018) diseñaron una plataforma web que monitorea Twitter para encontrar *tweets* en cuyo texto se reportan picaduras de mosquito o enfermedades relacionadas con el virus del zika. Twitter también ha sido utilizada para estudios espaciales relacionados con otros temas de salud que no son enfermedades, como es el caso de la obesidad (Cesare, Dwivedi, Nguyen, & Nsoesie, 2019; Ghosh & Guha, 2013; Nguyen *et al.*, 2017).

En los dos últimos años ha habido una proliferación de investigaciones acerca del uso de datos de Twitter para analizar la distribución de la pandemia causada por el virus COVID-19. A nivel global, destaca el trabajo de Bisanzio *et al.* (2020), quienes han utilizado Twitter para predecir la dispersión del virus a escala global mediante la estimación de patrones de movilidad de usuarios que han publicado mensajes en la ciudad de Wuhan en un periodo de 2013 a 2015. Huang *et al.* (2020) utilizaron 580 millones de *tweets* publicados en todo el mundo para analizar en diferentes escalas espaciales como los esfuerzos colaborativos de reducir la movilidad son reflejados en los comportamientos espaciales de los usuarios de Twitter. Daughton *et al.* (2021) elaboraron un esquema de 6 categorías y 11 subcategorías asociadas con términos clave relacionados con la pandemia sobre una muestra de 228 millones de *tweets* de todo el mundo. Como resultado, identificaron comportamientos humanos relevantes, y concluyeron que los patrones de movilidad son consistentes en el tiempo con eventos específicos asociados a la enfermedad.

El uso de datos de Twitter para el estudio del impacto causado por el COVID-19 en la población y movilidad ha sido especialmente abundante en Estados Unidos. Zeng *et al.* (2021) estudiaron los *tweets* publicados desde marzo hasta noviembre de 2020 para realizar un estudio longitudinal en Carolina del Sur, obteniendo como resultados una asociación significativa y positiva entre la movilidad y los casos nuevos diarios causados por el COVID-19. Morshed, Khan, Tanvir, & Nur (2021) realizaron un análisis de sentimientos para entender el impacto del COVID-19 sobre los servicios de transporte en Florida (EEUU),

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

comparando una muestra de *tweets* recopilados entre febrero y abril de 2019 y otra muestra de mensajes publicados en el mismo periodo temporal en 2020. Observaron un aumento del número de *tweets* con sentimiento negativo durante las primeras fases de la pandemia. Li *et al.* (2020) elaboraron un cubo tridimensional (lugar de origen, lugar de destino, y fecha) de movilidad para extraer flujos de población en tiempo casi real a partir de datos de Twitter. Este cubo de movilidad fue combinado con datos diarios de infecciones causadas por el COVID-19, información socioeconómica y demográfica, y datos de movilidad humana de otras fuentes como los registros de telefonía móvil para crear una red neural que captura patrones espaciales y temporales, y permite elaborar mapas de riesgos y predicciones de casos diarios. Jiang, Huang, & Li (2021) emplearon los *tweets* publicados en Nueva York entre febrero y mayo y calcularon patrones de movilidad a partir de los cambios observados en el porcentaje de usuarios de las parcelas y usos del suelo de la ciudad. Al comparar los resultados con reportes de movilidad de Google, vieron que estos reportes indican un aumento de los usuarios en las parcelas residenciales del barrio de Manhattan, mientras que Twitter detectó una disminución del número de usuarios, asociándose la razón a que los principales usuarios de la red social en la ciudad son visitantes.

La elaboración de índices de movilidad o de distancia social ha sido una herramienta muy utilizada en este campo. Xu, Dredze, & Broniatowski (2020) midieron el índice de movilidad social en Estados Unidos a partir de *tweets* publicados en los últimos 16 meses. Los resultados obtenidos mostraron una gran reducción de la movilidad a media y larga distancia en el país después de la implementación de medidas de prevención contra el COVID-19. Esta reducción de la movilidad ha sido mayor en los estados que adoptaron pronto medidas de distancia social. Porcher & Renault (2021) elaboraron un índice diario de distancia social a nivel estatal con el objetivo de capturar comportamientos como el confinamiento o la reducción de la movilidad a partir de palabras claves en los *tweets*. Mediante un análisis de diversos modelos, observaron una correlación entre las creencias de comportamiento de distancia social en Twitter y la movilidad en Estados Unidos a nivel estatal.

A pesar del amplio abanico de trabajos que han empleado Twitter para el análisis espacial de comportamientos de población y movilidad asociados a los cambios causados por el COVID-19, la mayoría de estos trabajos han usado muestras de datos publicados hasta junio de 2020, habiendo todavía pocos estudios que hayan avanzado más en el tiempo. Además, el área de estudio de la mayoría de estos trabajos ha sido Estados Unidos. Aunque hay trabajos que han empleado registros de telefonía móvil para monitorizar los efectos causados por la enfermedad en España, apenas hay todavía investigaciones que hayan utilizado datos de Twitter.

3. Área de estudio, datos, y metodología

3.1. Área de estudio

Se ha seleccionado España como área de estudio al ser uno de los países más afectados por el virus COVID-19, con más de 2 millones de casos detectados a día 1 de enero de 2021 según el mapa web *COVID-19 Dashboard* diseñado por la *Johns Hopkins Coronavirus Resource Center* (2020). Se analizaron los años 2019 y 2020 para comparar la distribución y actividad espacial de los usuarios de una muestra en un periodo de movilidad normal inmediatamente anterior a la pandemia. En particular se han seleccionado los siguientes tres periodos temporales:

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

1. Meses de marzo, abril y mayo: periodo de gran interés ya que coincide con el momento en que empezó la crisis sanitaria causada por el COVID-19 en España y por el establecimiento del primer estado de alarma y el confinamiento en todo el país (Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática, 2020a).

2. Meses de julio y agosto: el interés radica en ser el periodo del año con mayor movilidad en el territorio, debido a las vacaciones laborales y al turismo extranjero que recibe el país. En 2020, las medidas de confinamiento y restricción de movilidad terminaron el día 21 de junio, por lo que muchas instituciones promovieron este periodo como de “nueva normalidad”.

3. Meses de septiembre, octubre y noviembre: esta franja temporal ha sido seleccionada debido al auge de la segunda ola de COVID-19 a lo largo del periodo. Al contrario que durante el primer estado de alarma, las medidas de restricción de movilidad fueron más flexibles, siendo la principal prohibición la limitación de la entrada o salida en las comunidades autónomas (Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática, 2020b).

Se han seleccionado en mayor detalle los centros de las ciudades de Madrid y Málaga para la elaboración de mapas de calor y comparación de la distribución espacial de los usuarios de Twitter localizados en el centro de las ciudades durante la pandemia y durante un mes de actividad normal. Madrid es la capital y ciudad más poblada de España, y por tanto la ciudad con mayor número de usuarios de Twitter (Fernández, R., 2022). Además, Madrid ha sido una de las regiones de Europa más afectadas por la pandemia de COVID-19. Por otra parte, Málaga es la sexta ciudad más poblada de España y es uno de los principales destinos turísticos y de playa del país durante el periodo estival.

3.2. Datos

La base de datos utilizada para este trabajo contiene un total de 2.754.308 *tweets* publicados en España por 199.645 usuarios durante el año 2019 y 1.653.104 *tweets* publicados por 107.025 usuarios durante el año 2020. Cada tweet tiene información relacionada con el número de identificación de usuario (ID), idioma de usuario, coordenadas espaciales de latitud y longitud, y registro temporal de fecha y hora. Los *tweets* se han descargado a partir de la API de Twitter que permite recolectar *tweets* en tiempo real. Para ello, se utilizó un script realizado en código *Python*. Los *tweets* fueron almacenados en una base de datos *MongoDB* e incorporados al SIG de escritorio *ArcGIS Pro*.

3.3. Metodología

Tras incorporar los *tweets* a un SIG, el primer paso fue eliminar cuentas con más de 5000 *tweets* (se asume que cuentas con un número mayor de mensajes son bots o cuentas publicitarias) y usuarios que han publicado todos sus mensajes en un mismo lugar (usuarios que recogen las mismas coordenadas de longitud y latitud en todos sus mensajes). Tras esta limpieza, se crearon dos campos nuevos para agregar los *tweets* tanto por mes como por año, y se realizó un resumen de datos para obtener el número de usuarios que han publicado mensajes en cada mes mediante el campo del identificador de usuario como clave. Además, mediante una consulta SQL sobre los datos de palabras clave relacionadas con el COVID-19 (COVID, coronavirus, pandemia, contagio, confinamiento, mascarilla, vacuna), se seleccionaron los *tweets* que contienen alguno de esos términos y se realizó otro resumen de datos para visualizar la evolución del uso de las palabras analizadas a lo largo del año.

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Cada *tweet* ha sido enriquecido con información a nivel provincial mediante una unión espacial de los datos con una capa de división administrativa del territorio español. Nuevamente, se realizó una metodología de resumen de datos para identificar el número de usuarios registrados por provincia a lo largo del año. Para la cartografía de resultados, se seleccionaron los *tweets* publicados en cada periodo temporal de estudio, y se exportaron en capas nuevas. Como resultado, se han obtenido las capas recogidas en la Tabla 1.

Tabla 1. Número de *tweets* y usuarios descargados en cada periodo temporal de estudio

Periodo temporal	Número de <i>tweets</i>	Número de usuarios
Marzo, abril y mayo 2019	551285	64346
Marzo, abril y mayo 2020	318147	33929
% pérdida marzo, abril y mayo	57,71 %	52,72 %
Julio y agosto 2019	298940	60494
Julio y agosto 2020	139887	29597
% pérdida julio y agosto	46,79 %	48,92 %
Septiembre, octubre y noviembre 2019	470417	55846
Septiembre, octubre y noviembre 2020	250491	27392
% pérdida septiembre, octubre y noviembre	53,24 %	49,04 %

Fte. Elaboración propia a partir de datos de Twitter

Por último, se han seleccionado los *tweets* publicados en los cascos históricos de Madrid y Málaga durante los meses de julio y agosto para realizar mapas de calor. Los valores de densidad se han calculado mediante una estimación de densidad de kernel, obteniendo valores entre 0 y 1 que representan una densidad relativa basada en el recuento de puntos. Se ha empleado un radio de búsqueda de puntos de 10 metros.

4. Resultados

4.1. Distribución temporal de los usuarios de Twitter

Los primeros resultados muestran una comparación del número de usuarios de Twitter en España durante los años 2019 y 2020. Se puede apreciar cómo ha habido un mayor número de usuarios detectados a lo largo de 2019 en comparación con 2020. Durante 2019 el número de usuarios crece en febrero, decrece bruscamente durante marzo y abril, pero crece de nuevo en mayo. Desde mayo hasta noviembre se da un decrecimiento suave de usuarios, y de nuevo en diciembre aumenta el número de usuarios. En cambio, en 2020 se da un decrecimiento brusco de marzo a mayo (los meses en los que se establece el confinamiento domiciliario en España). En junio el número de usuarios detectados sube muy ligeramente y se mantiene estable hasta noviembre, donde hay un nuevo descenso (Figura 1).

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

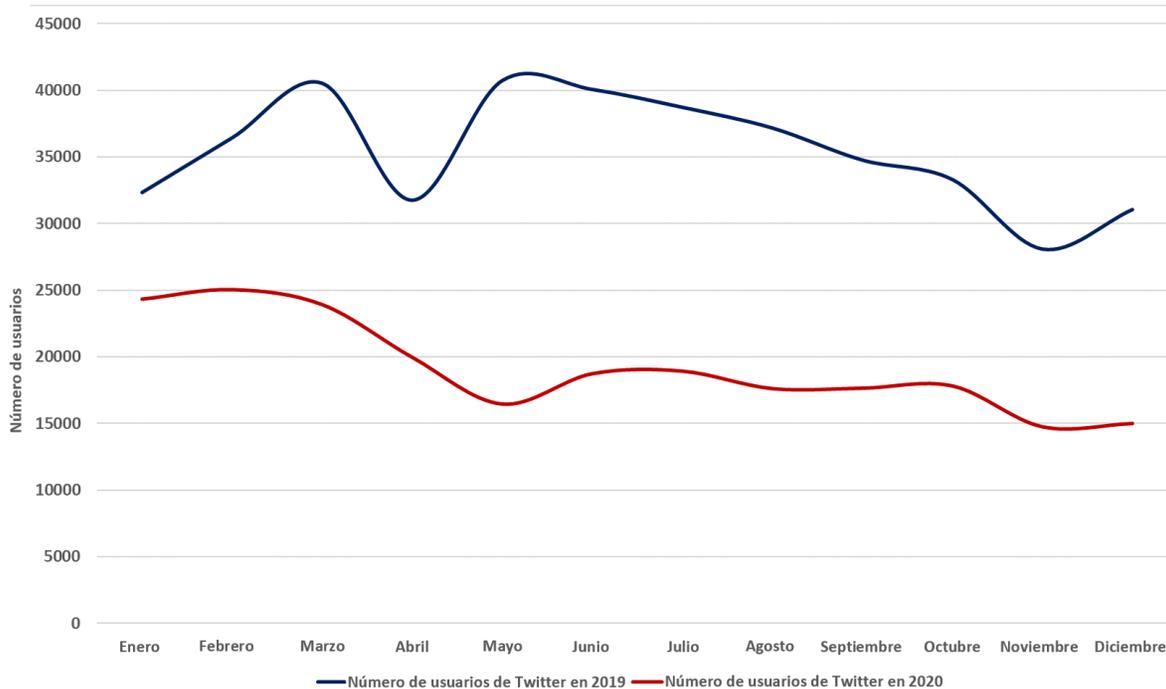


Figura 1. Número de usuarios de Twitter detectados en España a lo largo de los años 2019 y 2020
Fte. Elaboración propia a partir de datos de Twitter

La Tabla 2 muestra el número de *tweets* que contienen palabras claves relacionadas con el COVID-19 y el número de usuarios que han publicado dichas palabras. Se puede observar cómo al principio coronavirus fue el término más utilizado, pero rápidamente dejó paso a la palabra COVID. En casi todos los términos, el pico de volumen de usuarios y *tweets* ocurre en marzo, aunque se puede apreciar como el pico relacionado con la palabra confinamiento ocurre en abril, segundo mes de restricción de movilidad a nivel nacional, y que conllevó la anulación de las festividades de Semana Santa en todo el territorio. Además, el pico del término mascarilla sucede en julio, momento en el que hubo una mayor libertad de movimientos por lo que se las instituciones realizaron campañas de concienciación ciudadana para promover el uso de la mascarilla en la calle. A excepción de este último término, el resto de palabras claves sufre un descenso continuo y brusco de usuarios a lo largo de primavera y verano. Este número se estabiliza en los meses de otoño, donde se mantiene continuo. Finalmente, la palabra vacuna surge en diciembre, correspondiéndose con el momento en el que se empezaron a aprobar vacunas de diferentes laboratorios (Figura 2).

Tabla 2. Número de *tweets* geolocalizados que contienen una palabra clave relacionada con el COVID-19 y número de usuarios que han publicado *tweets* con dichas palabras

Palabra	Número de <i>tweets</i>	Núm. de usuarios
Coronavirus	6888	2457
COVID	8610	3603
Confinamiento	6283	3040

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Pandemia	1535	996
Contagio	266	211
Vacuna	195	146
Mascarilla	2873	1600

Fte. Elaboración propia a partir de datos de Twitter

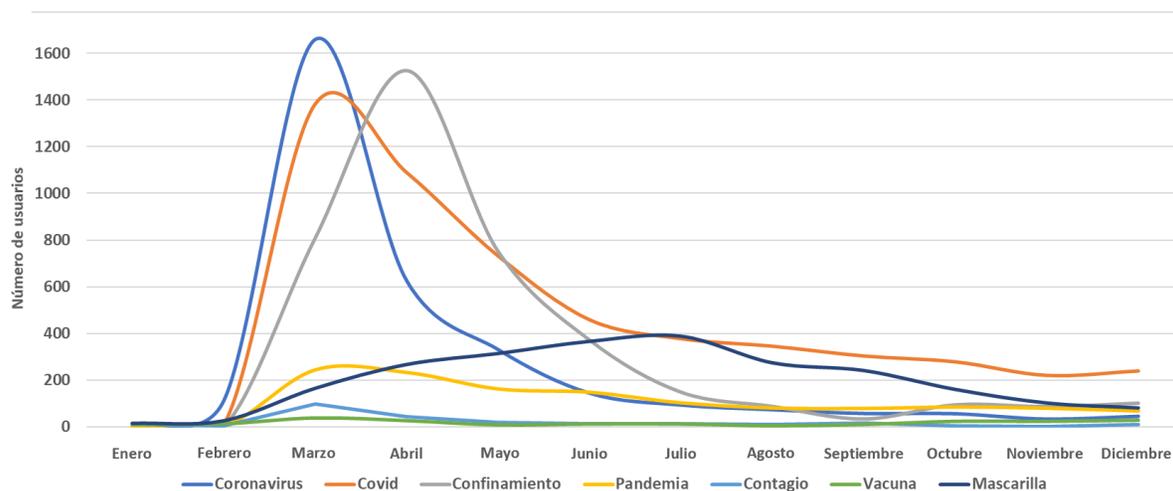


Figura 2. Número de usuarios que han publicado palabras claves relacionadas con el COVID-19 a lo largo del año 2020

Fte. Elaboración propia a partir de datos de Twitter

Se ha obtenido una aproximación al número de usuarios extranjeros no hispanohablantes que han visitado España a partir del campo de idioma de los usuarios de Twitter detectados en la muestra (si el idioma español es el valor más presente en el campo de idioma por cada identificador de usuario, se dice que ese usuario ha publicado principalmente en español). Mientras que en 2019 se ha obtenido de media un 35 % de usuarios que han publicado *tweets* principalmente en otro idioma, el porcentaje cae a aproximadamente un 28 % en 2020, lo cual concuerda con el descenso de la actividad turística en el país por la crisis sanitaria (Gutiérrez, H., 2020) (Tabla 3).

Tabla 3. Número y porcentaje de usuarios de Twitter que han publicado *tweets* principalmente en español y en otro idioma

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Periodo	Número usuarios	Número usuarios que publicaron principalmente en español	Número usuarios que publicaron principalmente en otro idioma	% usuarios que publicaron principalmente en español	% usuarios que publicaron principalmente en otro idioma
Marzo, abril y mayo 2019	64346	40426	23920	62,83 %	37,17 %
Marzo, abril y mayo 2020	33929	23782	10147	70,09 %	29,91 %
Julio y agosto 2019	60494	34354	26140	56,79 %	43,21 %
Julio y agosto 2020	19831	13290	6541	67,02 %	32,98 %
Septiembre, octubre y noviembre 2019	55846	35690	20156	63,91 %	36,09 %
Septiembre, octubre y noviembre 2020	27392	20288	7104	74,07 %	25,93 %

Fte. Elaboración propia a partir de datos de Twitter

4.2. Distribución espacial y espaciotemporal de los usuarios de Twitter

Al visualizar de forma general el número de usuarios de Twitter detectados en cada uno de los tres periodos temporales establecidos en la metodología, se observan en todos los casos un mayor número de usuarios en las provincias de Madrid y Barcelona, y en menor medida, en las provincias de Valencia, Málaga y Sevilla. Del mismo modo, se observa claramente un descenso del número de usuarios en todo el territorio durante el año 2020 respecto al año anterior.

Al comparar el periodo compuesto por los meses de marzo, abril y mayo de ambos años, se puede apreciar el descenso del número usuarios durante el confinamiento principalmente en las comunidades autónomas del interior peninsular (destacando Extremadura y Castilla y León), además de las comunidades de Navarra y País Vasco. Este descenso de usuarios es menos pronunciado en Andalucía. También se visualiza como la provincia de Cuenca mantiene un número destacado de usuarios detectados en ese periodo, posiblemente por la situación geográfica de la provincia entre Madrid y Valencia (Figura 3).

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

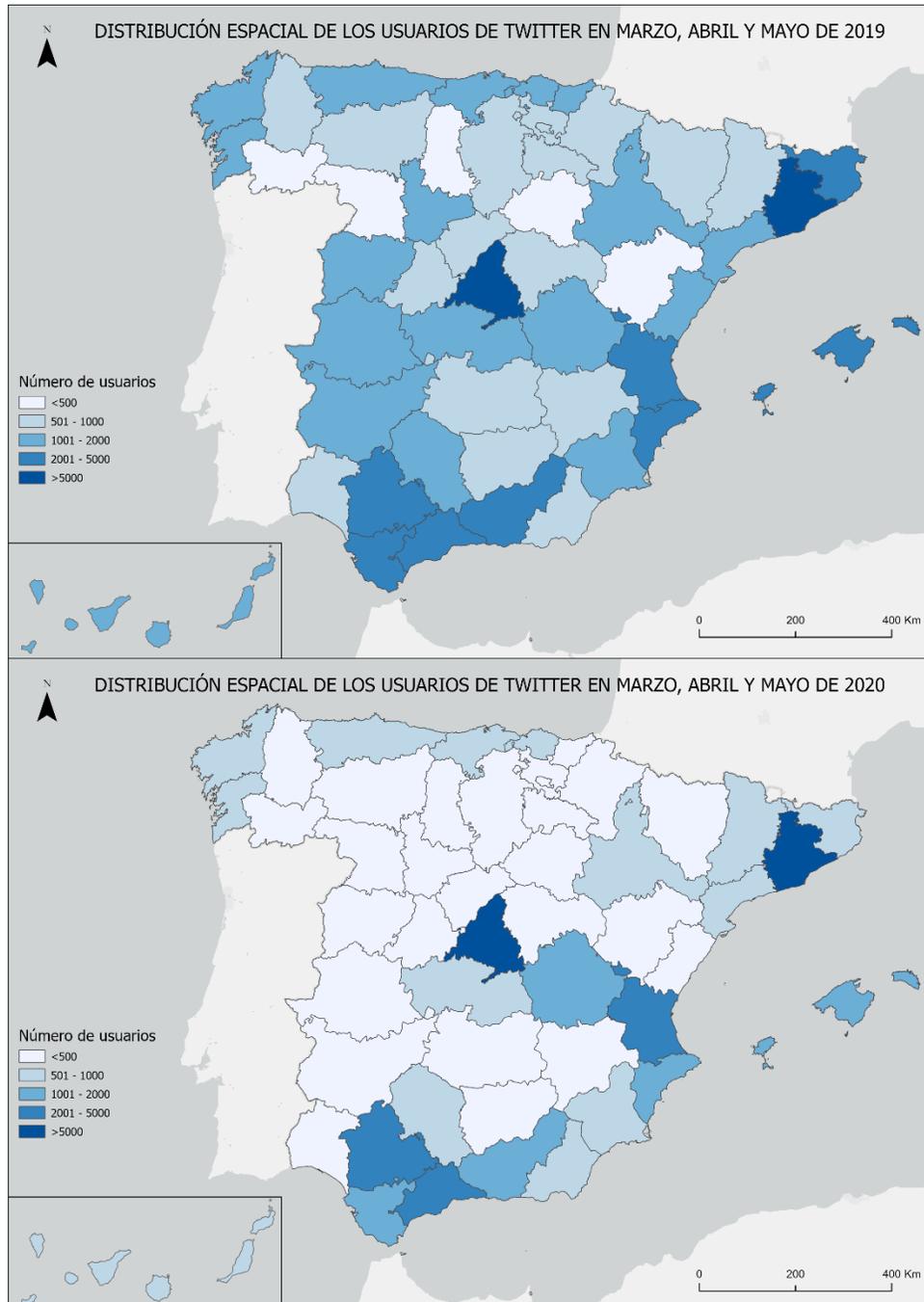


Figura 3. Número de usuarios de Twitter por provincia durante marzo, abril y mayo de 2019 (arriba) y durante el mismo periodo de 2020 (abajo)

Fte. Elaboración propia a partir de datos de Twitter

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Durante los meses de julio y agosto, se ha sufrido un descenso importante del número de usuarios detectados en casi todo el país. En este periodo de tiempo, la muestra de usuarios de Twitter suele ser menor debido a la disminución de flujos de viajes relacionados con trabajo o estudios. Este descenso de la actividad se vuelve muy pronunciado en el año 2020, ya que la pandemia también generó una importante reducción de flujos de viaje relacionados con la movilidad turística, reduciéndose el turismo en valores aproximados al 15 % en la mayor parte del país (Tabla 4). En cuanto a la distribución espacial de usuarios de Twitter, destacan el descenso de usuarios en las Islas Baleares y la provincia de Granada, importantes focos turísticos durante el periodo estival. También se observa un descenso de usuarios en las Islas Canarias, aunque este descenso es menor (Figura 4).

Tabla 4. Número de viajes realizados por comunidad autónoma de destino

Comunidad autónoma	Verano de 2019	Verano de 2020	Porcentaje de reducción de actividad turística
Andalucía	10981904	9268410	15,60 %
Aragón	1920899	1494747	22,19 %
Asturias	1759956	1435352	18,44 %
Islas Baleares	1283834	1091246	15,00 %
Islas Canarias	2261917	1784749	21,10 %
Cantabria	1539861	1533210	0,43 %
Castilla y León	5273611	4695376	10,96 %
Castilla la Mancha	3073841	2921949	4,94 %
Cataluña	7140158	6107141	14,47 %
Comunidad Valenciana	7514243	5747469	23,51 %
Extremadura	1223060	1101132	9,97 %
Galicia	3424200	2707298	20,94 %
Comunidad de Madrid	2601575	1509266	41,99 %
Comunidad de Murcia	1282774	1108350	13,60 %
Navarra	810887	699161	13,78 %
País Vasco	1209772	993347	17,89 %
La Rioja	478257	418828	12,43 %

Fte. Instituto Nacional de Estadística (Encuesta de turismo de residentes)

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

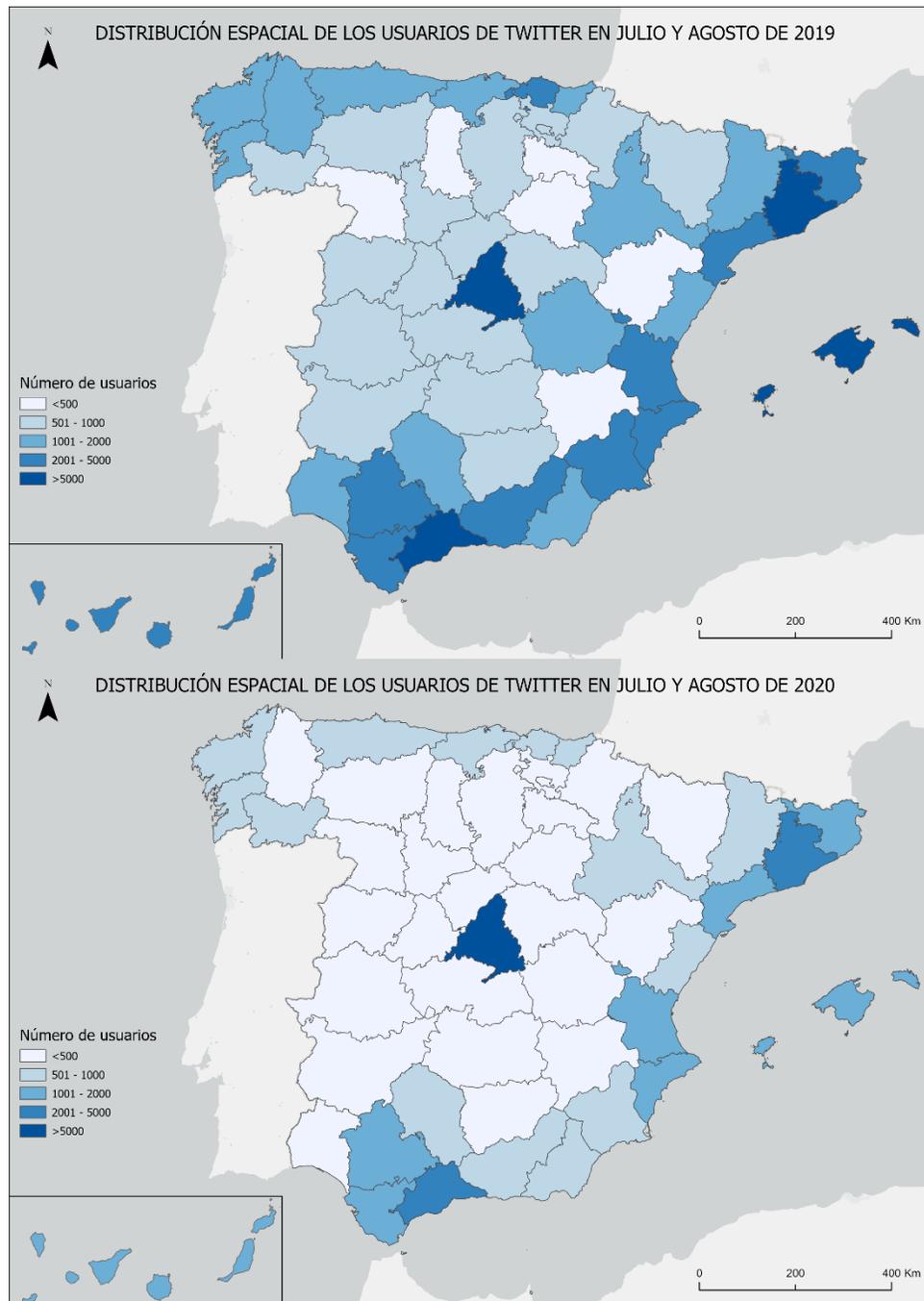


Figura 4. Número de usuarios de Twitter por provincia durante julio y agosto de 2019 (arriba) y durante el mismo periodo de 2020 (abajo)

Fte. Elaboración propia a partir de datos de Twitter

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Durante septiembre, octubre y noviembre, hay un ligero aumento de número de usuarios respecto al periodo estival, pero se observa también un descenso en todo el país respecto a la primavera, posiblemente debido a las medidas de restricción de movilidad que han impedido los desplazamientos entre comunidades autónomas o ciudades. Mientras que se observa un fuerte descenso del número de usuarios en el interior peninsular o el este de Andalucía, también se visualizan algunas zonas cuyo descenso de usuarios en comparación a los datos de 2019 es más suavizado (principalmente Asturias, y en menor medida Cantabria, las provincias costeras de País Vasco, o la Comunidad Valenciana) (Figura 5).

Al seleccionar y analizar las cinco provincias con mayor número de usuarios detectados en 2020 y extraer el número de usuarios que han publicado términos relacionados con el COVID-19, se aprecia cómo mientras Madrid ha sido la provincia con un mayor número de usuarios, Barcelona es la provincia donde se han publicado un mayor número de mensajes relacionados con la enfermedad, especialmente en el mes de julio. Sin embargo, esta tendencia cambia a partir de septiembre, mes donde Madrid pasa a convertirse en la provincia con mayor número de usuarios publicando *tweets* relacionados con la crisis sanitaria. El motivo puede deberse a un mayor impacto del virus ese mes, causando el inicio del segundo estado de alarma. Tanto en estas dos provincias, como en menor escala en las otras provincias destacadas, hay un gran pico de usuarios publicando mensajes específicos sobre la enfermedad en el mes de marzo, y a partir de entonces la pendiente baja drásticamente a lo largo de la primavera y verano hasta aplanarse en otoño (Figura 6).

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

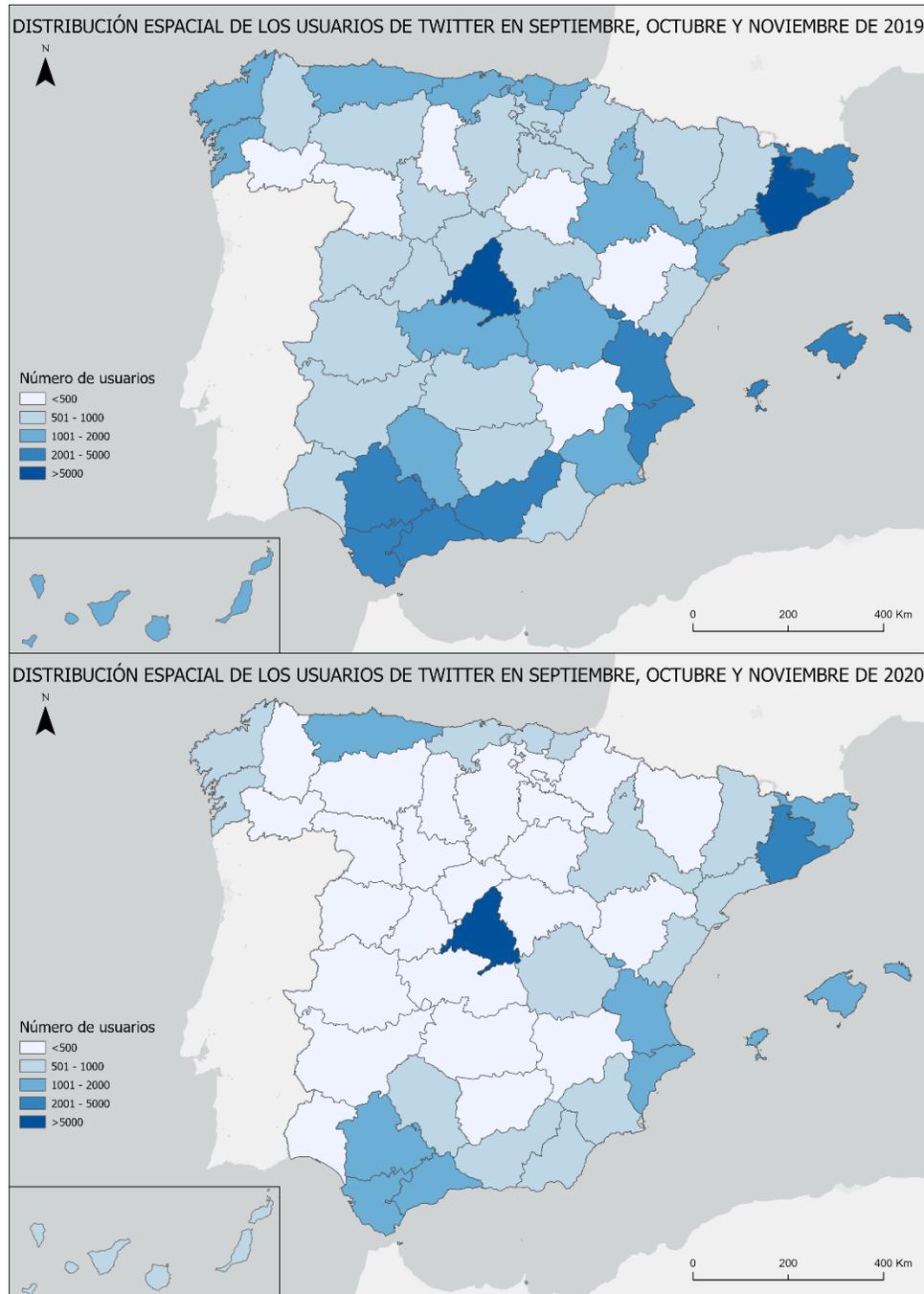


Figura 5. Número de usuarios de Twitter por provincia durante septiembre, octubre y noviembre de 2019 (arriba) y durante el mismo periodo de 2020 (abajo)

Fte. Elaboración propia a partir de datos de Twitter

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

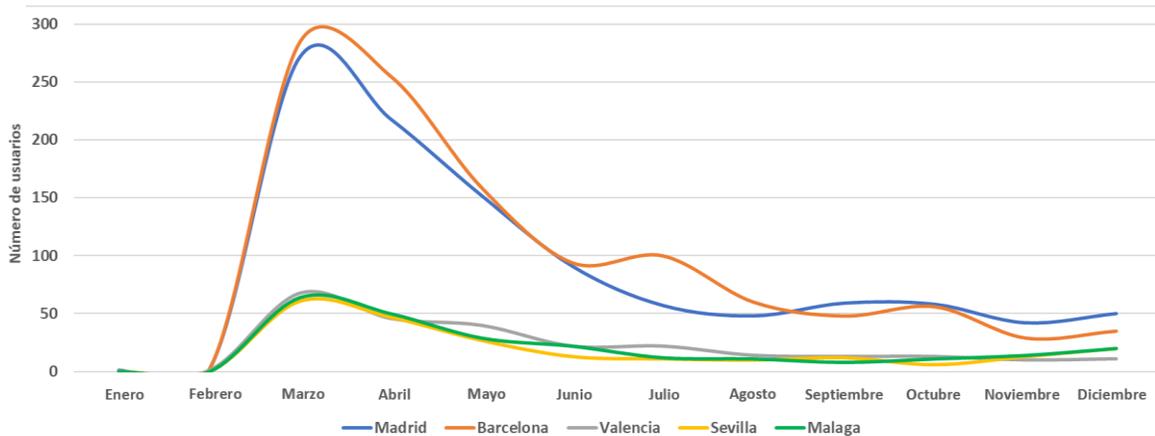


Figura 6. Número de usuarios que han publicado palabras claves relacionadas con el COVID-19 a lo largo del año 2020 por provincia

Fte. Elaboración propia a partir de datos de Twitter

4.3. Distribución espaciotemporal de los usuarios de Twitter en las áreas centrales de Madrid y Málaga

Para observar en mejor detalle la caída del número de usuarios detectados y del turismo en el año 2020 a raíz de la crisis sanitaria, se han elaborado dos mapas de calor durante los meses de julio y agosto de 2019 y 2020 de los cascos históricos de Madrid y de Málaga. En el caso de Madrid, se observa una distribución similar de usuarios, lo que indica unos comportamientos turísticos parecidos a los que se dan en un año de actividad normal. Los principales focos de calor se ubican en Puerta del Sol y el Parque del Retiro. Aun así, se observa un menor número y densidad de usuarios en el año 2020, además de la desaparición de puntos de calor asociados a actividad en importantes hitos turísticos como los ubicados en el Palacio Real o la Puerta de Alcalá (Figura 7).

Málaga, al contar con un número menor de usuarios detectados, presenta una distribución espacial más fácil de diferenciar, con un claro descenso de la densidad de usuarios durante el año 2020, sobre todo en el casco histórico, además de la desaparición de puntos asociados al área donde se encuentran el Teatro Romano, y la entrada a la Alcazaba. El único foco de usuarios importante que se da en la ciudad en julio y agosto del año 2020 se halla sobre la estación de trenes y autobuses de la ciudad (Figura 8).

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

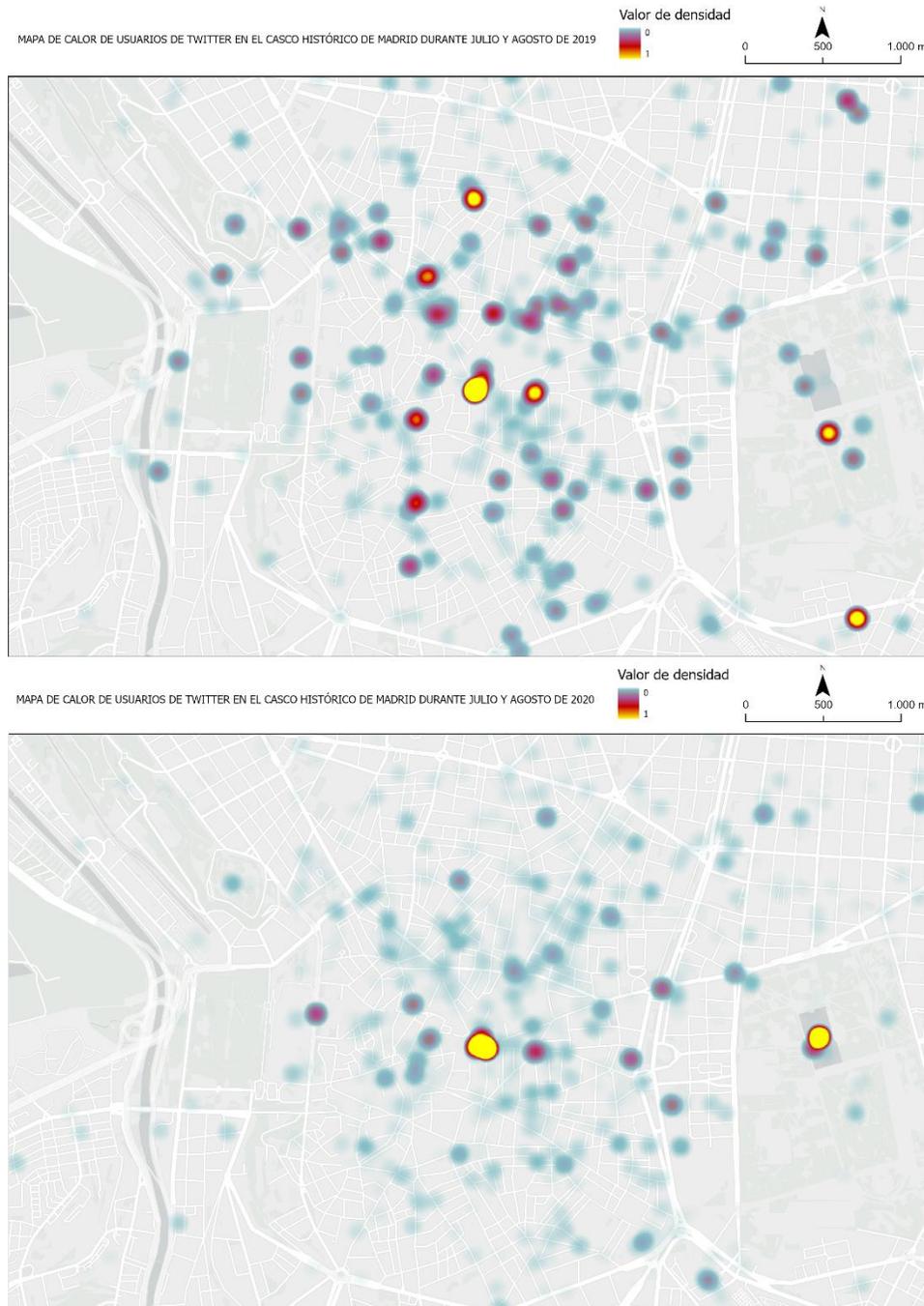


Figura 7. Mapa de calor de tweets por usuario publicados en el Casco Histórico de Madrid durante julio y agosto de 2019 (arriba) y julio y agosto de 2020 (abajo)

Fte. Elaboración propia a partir de datos de Twitter

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

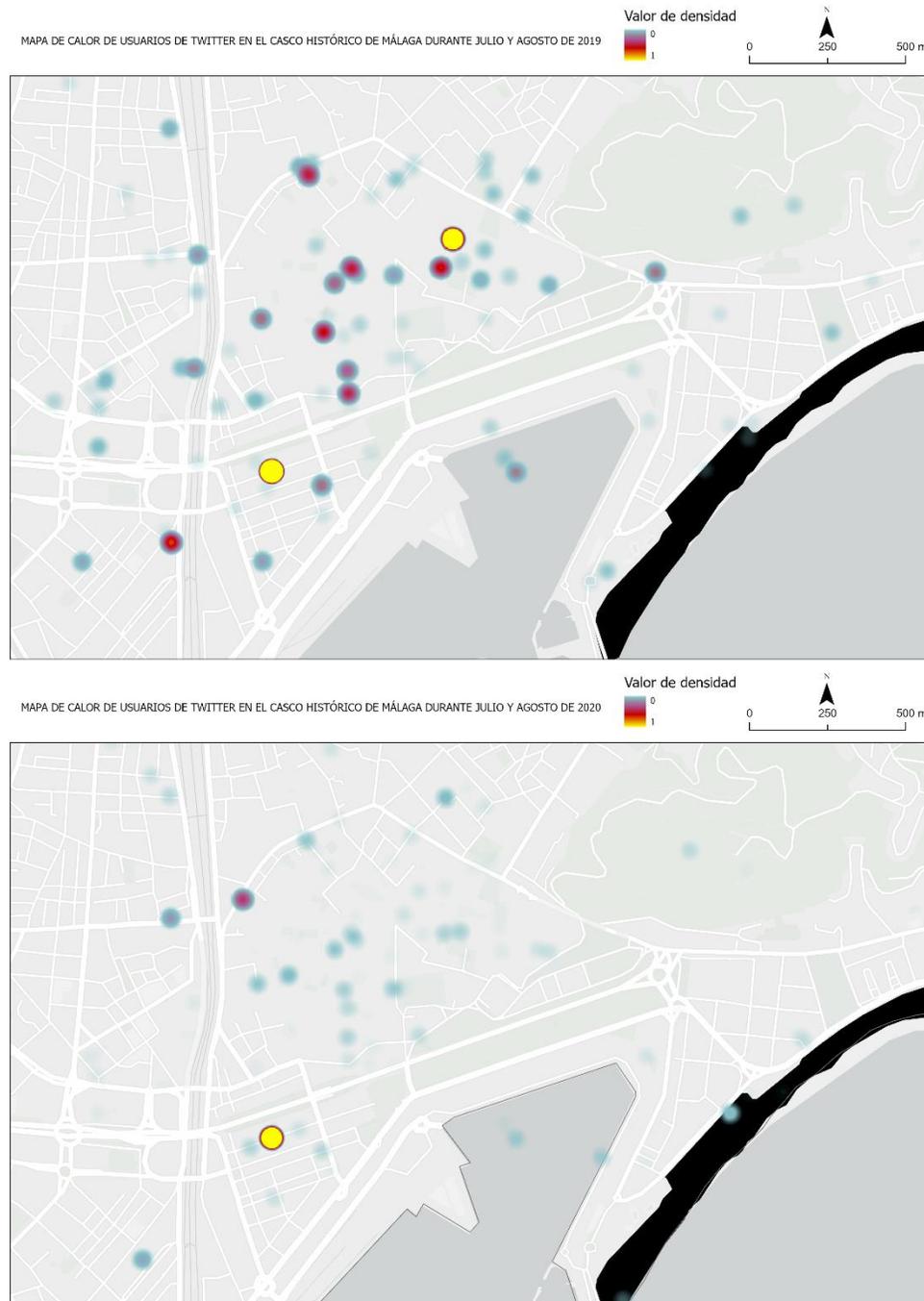


Figura 8. Mapa de calor de tweets por usuario publicados en el Casco Histórico de Málaga durante julio y agosto de 2019 (arriba) y julio y agosto de 2020 (abajo)

Fte. Elaboración propia a partir de datos de Twitter

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

5. Conclusiones

En este trabajo se han utilizado datos de Twitter para observar la distribución espacial y temporal de usuarios que han publicado mensajes geolocalizados en España durante 2020, año en el que la actividad y movilidad humana se han visto fuertemente afectadas por la irrupción de la pandemia ocasionada por la enfermedad del COVID-19. Para entender mejor como ha afectado esta crisis sanitaria a la movilidad, se ha comparado la distribución temporal y espacial de usuarios por periodos concretos de tiempo respecto al año anterior. Los resultados obtenidos muestran un pico de reacción semántica a la pandemia en el mes de marzo, mes donde la enfermedad del COVID-19 brotó en España provocando como consecuencia el confinamiento a nivel nacional. Se puede apreciar que, a lo largo del año, aunque sigue presente, ha bajado la reacción a la pandemia en los textos de los *tweets* de la muestra, señalando una situación en la que la población se ha empezado a acostumbrar al nuevo modo de vida causado por la enfermedad.

Normalmente, los usuarios de Twitter publican mensajes geolocalizados desde teléfonos móviles fuera del lugar de residencia. Al disminuir el número de viajes y desplazamientos a causa del COVID-19, la muestra obtenida de usuarios ha sufrido una fuerte bajada en el año 2020 respecto al año anterior. Además, el número de usuarios de la muestra decrece constantemente a lo largo del año pese a no haber restricciones de confinamiento y movilidad en comparación con el periodo primaveral. Este comportamiento puede indicar un aumento y consolidación del teletrabajo durante el segundo semestre del año.

Al analizar el área de estudio, se aprecia mejor el decrecimiento de la actividad causada por la pandemia en las provincias con un menor perfil demográfico. Se observa fácilmente como la distribución espacial de la muestra ha disminuido en el año 2020 en Andalucía o la Comunidad Valenciana, comunidades autónomas que cuentan con una población muy dinámica. En áreas de como las provincias interiores de la meseta peninsular, el número de usuarios obtenidos en el año 2020 ha sido bastante bajo. También se ha visualizado como la pandemia ha afectado al turismo durante el periodo estival, produciéndose un fuerte descenso de usuarios en provincias dedicadas al turismo como las Islas Baleares. Al analizar esta situación en dos ciudades con una fuerte actividad turística como es el caso de Madrid y Málaga, se observan huellas de calor asociadas a una menor densidad de usuarios respecto a un año de actividad turística normal.

Pese a que los datos de Twitter han mostrado un reflejo de actividad de usuarios durante la pandemia cercano a la realidad, hay que tener en cuenta una serie de limitaciones como la presencia de sesgos hacia una población con acceso a teléfonos móviles. Esta población suele tener mayores recursos económicos para desplazarse en viajes a media y larga distancia, un factor crítico de cara a la propagación del COVID-19. El peso demográfico de la zona a analizar es otro sesgo a tener en cuenta, ya que provincias con bajo peso demográfico (provincias de interior peninsular como Soria, Teruel o Zamora) suelen contar con población de mayor rango de edad y un menor acceso a las redes sociales. En este sentido, otro sesgo del uso de datos de Twitter radica en su uso principal por población joven en una franja de edad entre 20 y 39 años. Por último, existe un desafío a la hora de estudiar y controlar problemas sanitarios a partir de redes sociales como Twitter debido a la falta de recursos o habilidades técnicas requeridas para analizar redes sociales en este campo.

Este trabajo presenta futuras líneas de investigación a seguir. Por un lado, en esta investigación se ha hecho una introducción al uso de datos semánticos de los textos de los *tweets* para analizar la evolución temporal del uso de palabras claves asociadas al COVID-19. Siguiendo esta línea, una línea de investigación potencial consiste en el análisis y cartografía de sentimientos de los *tweets* relacionados con la pandemia.

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Otra línea de investigación es la comparación de los resultados obtenidos en España en este artículo con otros países de Europa o América Latina para observar distintas casuísticas y poder visualizar si las diferentes políticas de actuación respecto a la enfermedad han conllevado diferentes comportamientos espaciales. Finalmente, otro futuro trabajo radica en la aplicación de la metodología realizada en este artículo para comparar los datos descargados de Twitter en 2020 con datos de los años 2021 y 2022, con el objetivo de visualizar si hay una restauración de la actividad y distribución espacial de los usuarios de Twitter en el periodo posterior a la pandemia.

Referencias bibliográficas

Barros, J. M., Duggan, J., & Rebholz-Schuhmann, D. (2018). Disease mentions in airport and hospital geolocations expose dominance of news events for disease concerns. *Journal of Biomedical Semantics*, 9(1), 1–11. <https://doi.org/10.1186/s13326-018-0186-9>

Bisanzio, D., Kraemer, M. U. G., Bogoch, I. I., Brewer, T., Brownstein, J. S., & Reithinger, R. (2020). Use of Twitter social media activity as a proxy for human mobility to predict the spatiotemporal spread of COVID-19 at global scale. *Geospatial Health*, 15(1), 19–24. <https://doi.org/10.4081/gh.2020.882>

Blanford, J. I., Huang, Z., Savelyev, A., & MacEachren, A. M. (2015). Geo-Located Tweets. Enhancing Mobility Maps and Capturing Cross-Border Movement. *PLOS ONE*, 10(6), e0129202. <https://doi.org/10.1371/journal.pone.0129202>

Cai, M., Shah, N., Li, J., Chen, W. H., Cuomo, R. E., Obradovich, N., & Mackey, T. K. (2020). Identification and characterization of tweets related to the 2015 Indiana HIV outbreak: A retrospective infoveillance study. *PLoS ONE*, 15(8 August 2020), 1–13. <https://doi.org/10.1371/journal.pone.0235150>

Cerda L, J., & Valdivia C, G. (2007). John Snow, la epidemia de cólera y el nacimiento de la epidemiología moderna. *Revista Chilena de Infectología*, 24(4), 331–334. <https://doi.org/10.4067/S0716-10182007000400014>

Cesare, N., Dwivedi, P., Nguyen, Q. C., & Nsoesie, E. O. (2019). Use of social media, search queries, and demographic data to assess obesity prevalence in the United States. *Palgrave Communications*, 5(1), 1–9. <https://doi.org/10.1057/s41599-019-0314-x>

Daughton, A. R., Shelley, C. D., Barnard, M., Gerts, D., Ross, C. W., Crooker, I., ... Fairchild, G. (2021, May 1). Mining and validating social media data for COVID-19-related human behaviors between January and July 2020: Infodemiology study. *Journal of Medical Internet Research*. JMIR Publications Inc. <https://doi.org/10.2196/27059>

Eysenbach, G. (2011). Infodemiology and Infoveillance: Tracking Online Health Information and Cyberbehavior for Public Health. *American Journal of Preventive Medicine*, 40(5), S154–S158. <https://doi.org/10.1016/j.amepre.2011.02.006>

Fernández, R. (2022). *Ranking de las 10 ciudades españolas con mayor número de usuarios activos de Twitter en 2021*. <https://es.statista.com/estadisticas/520085/ciudades-con-mayor-numero-de-usuarios-activos-de-twitter/> (Última fecha de acceso: 01/11/2022).

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Frith, J., & Saker, M. (2020). It Is All About Location: Smartphones and Tracking the Spread of COVID-19. *Social Media and Society*, 6(3). <https://doi.org/10.1177/2056305120948257>

Ghosh, D., & Guha, R. (2013). What are we ‘tweeting’ about obesity? Mapping tweets with topic modeling and Geographic Information System. *Cartography and Geographic Information Science*, 40(2). <https://doi.org/10.1080/15230406.2013.776210>

Google (2020). *Coronavirus (COVID-19)*. <https://news.google.com/covid19/map?hl=es> (Última fecha de acceso: 01/11/2022).

Gutiérrez-Puebla, J., García-Palomares, J. C., & Salas-Olmedo, M. H. (2016). Big (Geo) Data en Ciencias Sociales: Retos y Oportunidades. *Revista de Estudios Andaluces*, 33(331), 1–23. <https://doi.org/10.12795/rea.2016.i33.0>

Gutiérrez, H. (2020). *El turismo sufre su peor verano con un 75% menos de viajeros extranjeros en julio*. <https://elpais.com/economia/2020-09-01/el-turismo-comienza-su-peor-verano-con-un-75-menos-de-viajeros-extranjeros-en-julio.html> Última fecha de acceso: 01/11/2022).

Hartley, D., Nelson, N., Walters, R., Arthur, R., Yangarber, R., Madoff, L., ... Lightfoot, N. (2010). Landscape of international event-based biosurveillance. *Emerging Health Threats Journal*, 3:e3. <https://doi.org/10.3134/ehj.10.003>

Huang, X., Li, Z., Jiang, Y., Li, X., & Porter, D. (2020). Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLoS ONE*, 15(11 November). <https://doi.org/10.1371/journal.pone.0241957>

Jiang, Y., Huang, X., & Li, Z. (2021). Spatiotemporal patterns of human mobility and its association with land use types during covid-19 in New York city. *ISPRS International Journal of Geo-Information*, 10(5). <https://doi.org/10.3390/ijgi10050344>

Kamel Boulos, M. N., & Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: How 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *International Journal of Health Geographics*, 19(1). <https://doi.org/10.1186/s12942-020-00202-8>

Khan, Y., Leung, G. J., Belanger, P., Gournis, E., Buckeridge, D. L., Liu, L., ... Johnson, I. L. (2018). Comparing Twitter data to routine data sources in public health surveillance for the 2015 Pan/Parapan American Games: an ecological study. *Canadian Journal of Public Health*, 109(3), 419–426. <https://doi.org/10.17269/s41997-018-0059-0>

Johs Hopkins Coronavirus Resource Center (2020). *COVID-19 Dashboard*. <https://coronavirus.jhu.edu/map.html> (Última fecha de acceso: 01/11/2022).

Lamos, V., De Bie, T., & Cristianini, N. (2010). Flu detector-tracking epidemics on twitter. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6323 LNAI(PART 3), 599–602. https://doi.org/10.1007/978-3-642-15939-8_42

Lansley, G., & Longley, P. A. (2016). The geography of Twitter topics in London. *Computers, Environment and Urban Systems*, 58, 85–96. <https://doi.org/10.1016/j.compenvurbsys.2016.04.002>

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Li, Z., Li, X., Porter, D., Zhang, J., Jiang, Y., Olatosi, B., & Weissman, S. (2020). Monitoring the spatial spread of covid-19 and effectiveness of control measures through human movement data: Proposal for a predictive model using big data analytics. *JMIR Research Protocols*, 9(12). <https://doi.org/10.2196/24432>

Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática (2020a). *Real Decreto 463/2020, de 14 de marzo, por el que se declara el estado de alarma para la gestión de la situación de crisis sanitaria ocasionada por el COVID-19*. <https://boe.es/buscar/act.php?id=BOE-A-2020-3692> (Última fecha de acceso: 01/11/2022).

Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática (2020b). *Real Decreto 926/2020, de 25 de octubre, por el que se declara el estado de alarma para contener la propagación de infecciones causadas por el SARS-CoV-2*. https://boe.es/diario_boe/txt.php?id=BOE-A-2020-12898 (Última fecha de acceso: 01/11/2022).

Morshed, S. A., Khan, S. S., Tanvir, R. B., & Nur, S. (2021). Impact of COVID-19 pandemic on ride-hailing services based on large-scale Twitter data analysis. *Journal of Urban Management*, 10(2), 155–165. <https://doi.org/10.1016/j.jum.2021.03.002>

Nguyen, Q. C., Brunisholz, K. D., Yu, W., McCullough, M., Hanson, H. A., Litchman, M. L., ... Smith, K. R. (2017). Twitter-derived neighborhood characteristics associated with obesity and diabetes. *Scientific Reports*, 7(1), 1–10. <https://doi.org/10.1038/s41598-017-16573-1>

Osorio Arjona, J., & García Palomares, J. C. (2020). Spatio-temporal mobility and Twitter: 3D visualisation of mobility flows. *Journal of Maps*, 16(1), 153–160. <https://doi.org/10.1080/17445647.2020.1778549>

Poom, A., Järv, O., Zook, M., & Toivonen, T. (2020). COVID-19 is spatial: Ensuring that mobile Big Data is used for social good. *Big Data and Society*. <https://doi.org/10.1177/2053951720952088>

Porcher, S., & Renault, T. (2021). Social distancing beliefs and human mobility: Evidence from Twitter. *PLoS ONE*, 16(3 March). <https://doi.org/10.1371/journal.pone.0246949>

Sousa, L., de Mello, R., Cedrim, D., Garcia, A., Missier, P., Uchôa, A., ... Romanovsky, A. (2018). VazaDengue: An information system for preventing and combating mosquito-borne diseases with social networks. *Information Systems*, 75, 26–42. <https://doi.org/10.1016/j.is.2018.02.003>

Souza, R., Neill, D. B., Assuncao, R. M., & Meira, Jr., W. (2019). Identifying High-Risk Areas for Dengue Infection Using Mobility Patterns on Twitter. *Online Journal of Public Health Informatics*, 11(1), 3–5. <https://doi.org/10.5210/ojphi.v11i1.9754>

Xu, P., Dredze, M., & Broniatowski, D. A. (2020). The Twitter Social Mobility Index: Measuring Social Distancing Practices With Geolocated Tweets. *Journal of Medical Internet Research*, 22(12), e21499. <https://doi.org/10.2196/21499>

World Health Organization (2020). *Advice for the public: Coronavirus disease (COVID-19)*. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public> (Última fecha de acceso: 01/11/2022).

Young, S. D., Mercer, N., Weiss, R. E., Torrone, E. A., & Aral, S. O. (2018). Using social media as a tool to predict syphilis. *Preventive Medicine*, 109(August 2017), 58–61. <https://doi.org/10.1016/j.ypmed.2017.12.016>

Osorio Arjona, J. (2022). Redes sociales geolocalizadas y COVID-19: análisis de la actividad espaciotemporal de los usuarios de Twitter de España durante la pandemia, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica (Artículos)*, 30, 25–47. <https://dx.doi.org/10.21138/GF.789>

Zeng, C., Zhang, J., Li, Z., Sun, X., Olatosi, B., Weissman, S., & Li, X. (2021). Spatial-temporal relationship between population mobility and COVID-19 outbreaks in south carolina: Time series forecasting analysis. *Journal of Medical Internet Research*, 23(4). <https://doi.org/10.2196/27045>

Zhang, H., Wheldon, C., Dunn, A. G., Tao, C., Huo, J., Zhang, R., ... Bian, J. (2020). Mining Twitter to assess the determinants of health behavior toward human papillomavirus vaccination in the United States. *Journal of the American Medical Informatics Association*, 27(2), 225–235. <https://doi.org/10.1093/jamia/ocz191>.

