



Islas Canarias
Del 15 al 19 de noviembre de 2021



Junta de Andalucía
Consejería de Transformación Económica,
Industria, Conocimiento y Universidades
Instituto de Estadística y Cartografía de Andalucía

MOVILIDAD URBANA A PARTIR DE DATOS MÓVILES

Cristina Fernández Álvaro

Instituto de Estadística y Cartografía de Andalucía

cristina.fernandez@juntadeandalucia.es

Introducción

La pandemia de COVID-19 y la declaración del estado de alarma en marzo de 2020 implicó desarrollar estrategias de análisis que tenían que ir más allá de las estadísticas tradicionales.

Aunque el estudio de la huella digital de los teléfonos móviles era algo ya analizado por empresas privadas y la estadística pública había comenzado a hacer incipientes trabajos, la necesidad de monitorizar la movilidad de la población, por ser un elemento clave en la transmisión de la enfermedad, provocó la liberación de estos datos y la apertura de nuevas oportunidades de análisis.

El uso de estos datos de movilidad ha supuesto para el Instituto de Estadística y Cartografía un doble reto: por un lado reforzar su papel en la sociedad como organismo generador de datos, y por otro lado ser capaz de dar una respuesta diaria a la ciudadanía, procesando datos que llegaban con una periodicidad mucho más frecuente que la habitual.

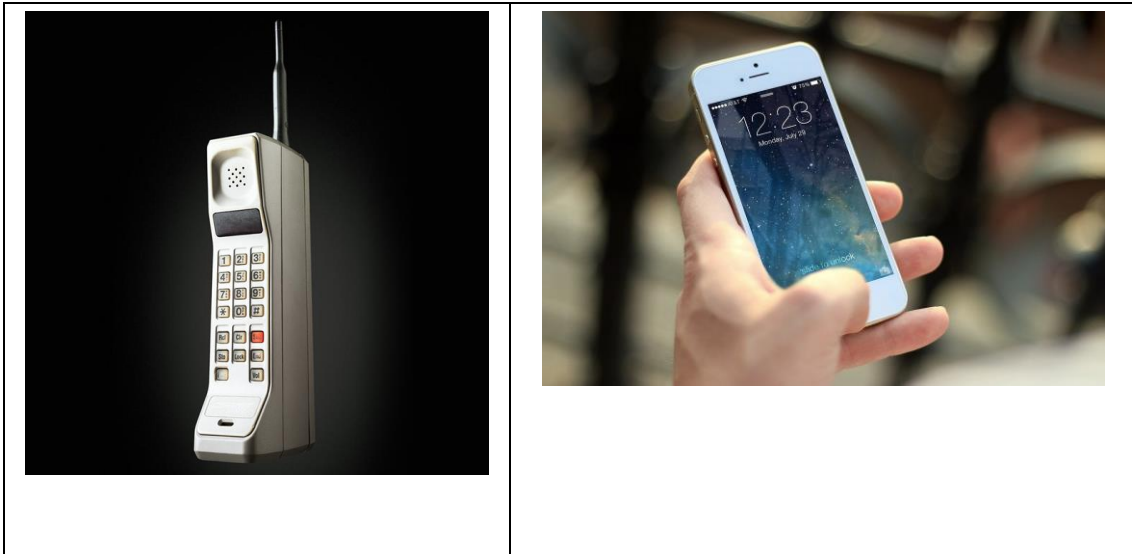
Además los análisis de estos datos han ayudado a comprender mejor cómo es y ha sido la movilidad en Andalucía actualmente y en periodos anteriores a la pandemia.

Objetivos

El objetivo de este trabajo es mostrar los distintos tipos de datos sobre movilidad procedentes de teléfonos móviles que pueden analizarse así como los análisis y los productos que el Instituto de Estadística y Cartografía han realizado en base a ellos.

Metodología

Origen de los datos



El primer teléfono considerado como móvil fue el Motorola Dyna TAC (izquierda). Pesaba 800gr y costaba 4.000€. 40 años después nuestros smartphones distan mucho de este primer teléfono, siendo la única función que realizaba el primer teléfono (llamar) de las menos utilizada en nuestros dispositivos.

El número de funcionalidades de los dispositivos se fue incrementando progresivamente (SMS, relojes y alarmas, pequeños video juegos, calculadora...) y con la llegada de internet y los smartphones se dispararon (gestor de correo, linterna, biblioteca musical, parte meteorológico o una forma de pago)

El uso de estas funcionalidades, en la medida en que hace uso de la red de telecomunicaciones, genera huellas digitales que se convierten en datos, cuyo aprovechamiento no suelen estar al alcance de los propietarios de esos teléfonos y por supuesto menos para las oficinas de estadística.

Clasificaremos la información obtenida de los teléfonos móviles en función del comportamiento que puede tener el titular del mismo frente a ella.

- Información procedente de aplicaciones, es decir procedente de apps del dispositivo. El individuo tiene un comportamiento activo frente a los datos pues decide qué aplicaciones se instala, a quien le da sus datos sobre edad, sus gustos y sus comentarios....

Esta es la información con más detalles de la persona registrada, información sobre sus características, sus acciones y sus gustos. Las propias app instaladas y utilizadas ya definen los intereses del individuo, además de los datos que éstas generan sobre temas relacionados con la app. De este tipo de información muchas empresas están haciendo su modelo de negocio, están convirtiendo información en valor económico.

Esta es la información más difícil y cara de conseguir, y a la que no acceden de momento los organismos estadísticos.

Quizás en un futuro, una nueva ley estadística nos podría dar acceso a esta información. Algo así se planteó recientemente en la Jornada del INE sobre Estadística y Sociedad. Sería una forma de devolverle a la sociedad, de forma gratuita, parte de la información que la sociedad genera (de forma gratuita) y de la que las empresas sacan rédito.

- Información procedente de Google.

Google se ha consolidado como un gestor de servicios para los teléfonos móviles que lo deja muy por encima del resto de aplicaciones. Es la base de gran parte de los teléfonos actuales y controla muchas de nuestras principales actividades (correos, agendas, contactos, fotos...).

En esta información generada el individuo tiene también un comportamiento activo, pero debemos admitir que el consentimiento es más inconsciente que en otras aplicaciones como Twitter o en Spotify.

- Información procedente de la necesidad de conexión.

Esta información no depende de las aplicaciones instaladas en el teléfono, su procedencia es la necesidad de estar conectado a una antena para que la tecnología del teléfono funcione. El individuo tiene por tanto un comportamiento pasivo.

En estos casos no hay información de carácter personal (salvo quizás la que se pudiera deducir por la tipología de terminal que se conecta, que puede decir mucho de la persona). Se trata de una información de carácter más tecnológico: información sobre las antenas a las que se conecta el teléfono, el tiempo de conexión, distancia de una conexión a otra...

Esta es la información con la que podemos empezar a trabajar los organismos de estadística.

Fuentes de datos de movilidad

La necesidad de conocer la movilidad diaria de la población y los motivos por los que la realiza es una materia que siempre ha preocupado y ocupado a muchos investigadores y gestores públicos.

La movilidad ha tenido siempre unos patrones bastante estables y, salvo diferencias en periodos concretos como los vacacionales, siempre se ha estudiado con estadísticas estructurales no anuales.

La fuente estadística principal ha sido tradicionalmente el Censo de Población y Viviendas, hasta que dejó de ser exhaustivo en 2001. La movilidad fue una de las variables que más sufrió la falta de información sobre toda la población.

Existieron otras encuestas como la Encuesta Movilia del Ministerio de Fomento en su día, hoy Ministerio de Transportes, Movilidad y Agenda urbana. La encuesta se llamaba Encuesta de Movilidad de las Personas Residentes. La segunda encuesta de movilidad que llevó a cabo el ministerio fue en el año 2006/2007, con el objetivo de poner a disposición de los usuarios, públicos y privados, un conjunto de información amplia, integrada y coherente sobre las pautas de movilidad de la población residente en España en viviendas familiares, sus características y sus determinantes.

El IECA en 2011 también llevó a cabo una encuesta sobre movilidad para Andalucía, concretamente la Encuesta Social 2011. Movilidad en las Regiones Urbanas de Andalucía. Esta encuesta presenta una panorámica general sobre las dinámicas y las pautas de movilidad física de la población, centrándose en los desplazamientos que se desarrollan en las zonas urbanas de movilidad intensa, denominadas en esta encuesta regiones urbanas de Andalucía, en las que residía el 71,3% de la población andaluza.

En el futuro los estudios de movilidad de las personas van a cobrar una mayor relevancia ya que otra de las consecuencias de la pandemia ha sido el cambio en los comportamientos de la población, ya sea por razones impositivas (restricciones a la movilidad intermitentes, aumento del desempleo) o cambios en las formas de empleo (teletrabajo).

Están comenzando a llegar a los institutos de estadística datos sobre movilidad de las personas obtenidos al procesar información procedente de los teléfonos móviles. Estos datos deberán ser analizados y comenzar a utilizarse con la finalidad de estudiar la movilidad de la población, a falta de otras fuentes que lo hagan.

La comparativa de estos datos con los obtenidos con las estadísticas tradicionales diseñadas para estudiar la movilidad plantea un debate similar al dilema que presentaba los registros administrativos frente a las encuestas.

Las nuevas fuentes de información

De todos los tipos de datos que pueden tener las compañías nos centraremos en los que actualmente podemos utilizar los institutos de estadística, que son los derivados del propio servicio que presentan las empresas de comunicación.

Estos datos existían, como puede intuirse, desde mucho antes de la pandemia, desde que los móviles existen, pero no ha sido hasta la pandemia cuando las compañías lo han ofrecido a los organismos públicos y a la ciudadanía para ayudar a la gestión de la crisis sanitaria.

Una vez que los organismos estadísticos hemos “descubierto” el potencial de estos datos no queremos y no debemos perderlos como fuente de información.

Actualmente las fuentes de información sobre movilidad que la ciudadanía tiene a su disposición son tres, cada una de ellas con distinto nivel de detalle y desagregación e incluso podríamos decir que cada uno mide aspectos distintos, todos ellos relacionados con la movilidad pero relativamente comparables.

Datos de Google:

Google consideró que los datos que tenía sobre las personas y que alimentaban productos como Google Maps podían ser útiles a las autoridades sanitarias de todo el mundo.

Por ello ofrece, por tiempo desconocido, información agregada y anonimizada sobre los cambios que se han producido en la movilidad de las personas como consecuencia de las políticas que se han establecido para combatir el COVID-19.

Los datos pueden encontrarse en <https://www.google.com/covid19/mobility/>

Los datos csv ofrecidos muestran las tendencias de movimiento a lo largo del tiempo ordenadas por zonas geográficas y clasificadas en diversas categorías de lugares, como tiendas y espacios de ocio, supermercados y farmacias, parques, estaciones de transporte, lugares de trabajo y zonas residenciales.

Con estos datos puede conocerse cómo cambia la cantidad de visitantes en los lugares categorizados (o el tiempo que pasan en ellos) en comparación con los días de referencia. Es información curiosa distinta de la que dan otras fuentes.

Los datos de google están muy tratados y no aportan información sobre movilidad entre áreas, pero si podrían ayudar a otros análisis al mostrarnos información sobre otros aspectos como los lugares comunes donde va la población, que no tenemos en otro tipo de información.

Estudio del INE:

Durante los años 2019, 2020 y 2021 el INE acometió un proyecto de medición de movilidad a partir de la telefonía móvil. Debido a la pandemia, la idea inicial del proyecto cambió, si bien la metodología ha seguido siendo la misma.

Inicialmente el estudio se diseñó para medir movilidad cotidiana tomando como referencia una semana concreta de noviembre de 2019 y cuatro días, también de 2019, para medir cómo se localiza la población fuera de su lugar habitual de residencia en fechas señaladas (dos días de verano, un fin de semana normal, el día de navidad).

El enorme impacto que supuso para la movilidad de la población el brote de coronavirus en marzo de 2020 y la consiguiente declaración del estado de alarma hizo que se planteara de forma urgente un segundo estudio de movilidad que proporcionó información diaria sobre movimientos entre las 3.214 áreas que se diseñaron para el proyecto.

Desde el 16 de marzo al 20 de junio (estado de alarma) los operadores de telefonía entregaron información diaria y gratuita que llegó al IECA a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial del Ministerio de Asuntos Económicos y Transformación Digital.

Terminado este periodo se plantea la necesidad de continuar con la serie de datos de movilidad para poder monitorizar lo que por entonces se denomina “vuelta a la normalidad”, lo que dio lugar al estudio EM-3. Vuelve a haber un cuarto estudio EM-4 para el año 2021 por un tiempo que es difícil de estimar pero que podría extenderse a todo el año 2021. En resumen los estudios de 2020 y 2021 han dado continuidad a la serie iniciada en 2019.

Actualmente el INE está publicando datos de movilidad cada semana, relativos a dos días concretos de la semana anterior (miércoles y domingo). El estudio se mantendrá previsiblemente durante todo el año 2021.

Los datos proceden de las tres principales compañías de móviles (Movistar 30%, Orange 25,5% y Vodafone 23%). Cada una hace una entrega al INE con la información ya tratada según la metodología adoptada. Para la obtención de los datos difundidos por el INE se agregarán la información de los tres operadores.

La información obtenida por el posicionamiento de los teléfonos móviles es la siguiente:

- Movilidad cotidiana: datos de todos los miércoles y domingos desde marzo de 2020. Se actualiza cada semana, los jueves no festivos.

- Movilidad estacional: datos de cuatro días concretos de 2020 (18 de julio y 15 de agosto, 22 de noviembre y 25 de diciembre), que se pueden comparar con los días similares de 2019, antes de la aparición del Covid-19 lo que permite una primera aproximación a conocer movimientos turísticos o estacionales. Estos datos se completarán con los correspondientes a fechas similares de 2021.

Para determinar el lugar de origen y el lugar de destino de cada terminal se plantea la siguiente metodología.

El área de residencia de cada teléfono móvil se determina en función de dónde “duerma” el terminal, es decir se considerará aquella en donde el teléfono móvil se encuentra durante más tiempo entre las 22:00 horas del día anterior y las 06:00 horas durante el día observado.

Asignada para cada teléfono móvil el área de residencia, se intenta determinar el área de destino cotidiano. El método es parecido al de asignación del área de residencia pero no dará siempre un valor: se deberán tomar todas las áreas (incluida la de residencia) en las que se encuentra el terminal durante el horario de 10:00 a 16:00 del día observado. Si hay más de una área, se tomará aquella más frecuente (en la que está más tiempo), que no sea la de residencia.

Si solo el área de residencia cumple esta condición, entonces el área de residencia será también el área de destino.

La información disponible se ofrece en ficheros por periodos. Los primeros hacían referencia a casi todos los días y luego se han pasado a analizar miércoles y domingos

- Flujos Origen Destino de ese día. Se trata de flujos de más de 15 personas
- Población por lugar de destino de ese día
- Población por lugar de origen de ese día

El Instituto de Estadística y Cartografía de Andalucía comenzó utilizando estos datos para los informes diarios de movilidad, pero luego se pasó a utilizar los del ministerio que vemos a continuación

Toda la información puede encontrarse en:

https://www.ine.es/experimental/movilidad/experimental_em.htm

Datos del ministerio:

El Ministerio de Transportes, Movilidad y Agenda Urbana también comenzó a generar información sobre movilidad durante la crisis del COVID-19, con la finalidad de obtener información que sirviese tanto para evaluar el efecto de las medidas de restricción de la movilidad impuestas a la ciudadanía como para realizar otros análisis y estudios que ayudasen a la gestión y a la posterior salida de esta crisis.

La solución que se consideró mejor para ello fue emplear soluciones basadas en el análisis de datos masivos, aprovechando la experiencia adquirida en el proyecto ‘Estudio de la Movilidad Interprovincial de Viajeros aplicando la Tecnología Big Data’ realizado en 2018.

La información procedente de las redes de telefonía móvil se ha fusionado con otras fuentes de datos como usos de suelo, puntos de interés, redes de transporte o información sobre demanda de transporte.

A partir de esa información se generan matrices origen-destino y otros indicadores de movilidad y presencia de población anónimos y agregados, garantizando el estricto cumplimiento con los requisitos de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPD-GDD).

Los datos pueden descargarse en: <https://www.mitma.es/ministerio/covid-19/evolucion-movilidad-big-data>

Se dispone de datos desde el 14 de febrero de 2020, siendo esa primera semana la de referencia. Desde esa fecha se tienen los ficheros de cada día hasta el día 9 de mayo, fecha en la que acabó el estado de alarma. El suministro de datos actualmente ha terminado, si bien se está gestionando un nuevo contrato que permitirá dar continuidad a los datos.

Las características más singulares de estos datos del ministerio son:

- Que proceden de fuentes privadas, concretamente de orange
- Los datos tienen un pretratamiento: existen variables establecida en función de variables auxiliares como por ejemplo la asignación de los motivos de trabajo, que se llevan a cabo en base a criterios espaciales de los lugares de destino
- Estudia fundamentalmente desplazamientos si bien hay también datos de personas
- El desplazamiento considera la variable hora en el que se inicia el mismo.

Los datos diarios nos muestran dos ficheros, uno de movilidad de personas y otros de desplazamientos.

- El de personas se llama maestra 2 y contiene el número de personas residentes en cada ámbito geográfico, que se ha desplazado (o no) y cuantos viajes ha realizado.
- El fichero de desplazamientos (maestra 1) es bastante grande, tiene de media de entre 6 y 4 millones de registros (para Andalucía 900.000- 700.000).

La información que contiene básicamente es la matriz origen destino, si bien con más características que detallan el desplazamiento:

- periodo de inicio del desplazamientos
- motivo del mismo (casa, trabajo, otros)
- edad (está en diseño de registro pero viene vacío)
- residencia
- distancia recorrida

La información geográfica cuenta con 2834 áreas o distritos. Para Andalucía se reduce a 504.

Limitaciones de este tipo de datos.

Sin olvidar que estas fuentes de datos no proceden de actividades estadísticas trataremos sacar de ellas la máxima información aunque nos encontramos con limitaciones como estas:

- Pueden existir teléfonos apagados o lugares con problemas de cobertura.
- Las personas mayores y los niños de menos de 12 años no suelen tener móviles (nuestra población objetivo se ve reducida)
- Dependemos de compañías privadas en la entrega de datos, en este sentido la jornada del INE Estadística y Sociedad ya dejó caer ese cambio en la mentalidad de que fueran los organismos estadísticos los que recogieran esos datos de estas empresas y de alguna forma le devolvieran a la sociedad la información que tan codiciada está en otros mercados.
- En referencia a los datos, no sabemos quien realiza exactamente cada el recorrido ni el motivo (el que existe se lo hemos asignado a posteriori según fuentes auxiliares) ni como se completa en un día todos los desplazamientos de una misma persona.
- Se trata de información pretratada, los datos no nos llegan en estado bruto.

Resultados

Diferencias entre las fuentes

En referencia a la **fente generador** de la información los datos del INE proceden de tres compañías, lo que ofrece una información más completa, pues las tres aportan más del 60% de los teléfonos móviles. Por su parte MITMA sólo considera un operador.

Ninguna de las instituciones realiza el **tratamiento de la información**, sino que éste viene dado por parte de las compañías de telefonía. En el caso de estudio del INE, cada compañía aporta las matrices y luego éstas son agregadas. La censura se realiza a nivel de cada compañía, por lo que hay sobre censura a nivel global. Los procesos que dan lugar a los datos de MITMA se llevan a cabo dentro de la infraestructura del operador móvil, de manera que la información generada y entregada al Ministerio es ya información agregada y anonimizada.

Desde el punto de vista del **contenido de la información** los datos del INE hacen referencia a número de personas, tanto desde un punto de vista de stock como de flujos entre origen (residencia) y destino. La información no añade más información sobre características de esas personas, cuando se han desplazado ni el motivo.

La información difundida por MITMA aporta información sobre personas en cada área y si estas se han desplazado o no, pero la información más relevante la ofrece en términos de número de desplazamientos, donde aporta además otras variables como lugar de residencia, distancia recorrida o actividad u el horario de inicio del desplazamiento

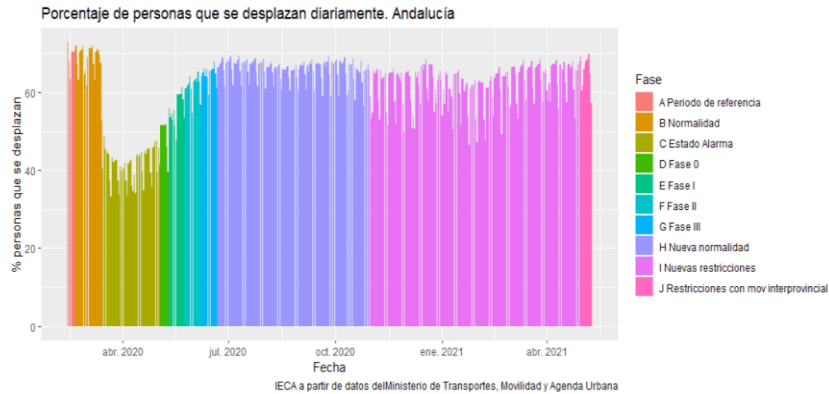
En referencia al tratamiento dado a los **extranjeros**, los datos procedentes del estudio experimental del INE no considera los teléfonos de numeración extranjera que operan en España en roaming, los cuales estarían en manos de turistas. El ámbito poblacional por tanto es la población residente en España. Los datos de los últimos proyectos de MITMA no ofrecen información sobre datos de no residentes en España ni de origen o destino fuera del país, aunque no fue así en el proyecto piloto.

Productos obtenidos con datos de movilidad

Aproximadamente al mes de declararse el estado de alarma, el IECA tuvo acceso a los datos de movilidad a partir de teléfonos móviles, tanto los del Ministerio como los del estudio del INE.

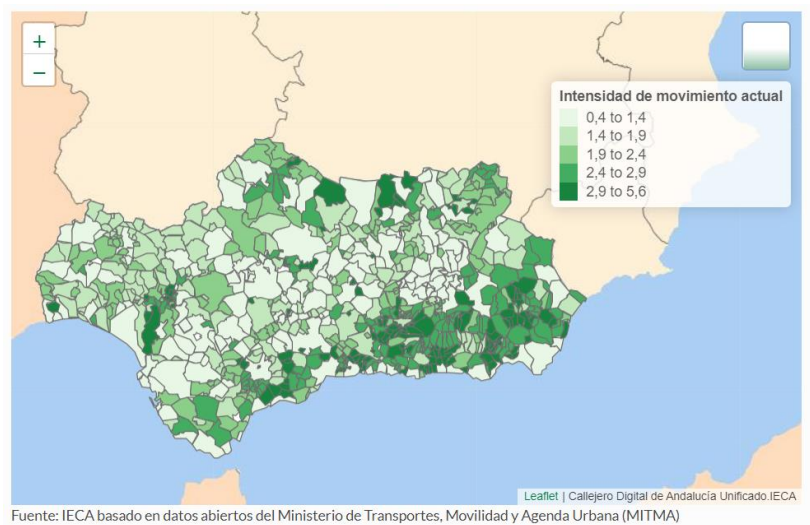
Inicialmente se planteó la elaboración de un informe diario que ayudara a la desescalada, para ello se utilizó inicialmente los datos del INE para luego centrarnos en los datos de MITMA. El primer informe se presentó el día 14 de mayo de 2020, y se optó por el formato HTML a partir de RmarkDown.

Con estos datos de movilidad se pretendía **monitorizar la evolución de la reactivación de la actividad económica**, para ello se consideró el indicador *% de población que se desplazaba*, como indicador de actividad económica.



Partiendo de una situación de “normalidad” (que los datos sitúan entre el 14 y el 20 de febrero) el seguimiento de los datos mostraría en qué medida éstos van regresando a unos valores similares a los del mes de febrero.

Para el informe también se construyó un **indicador de movilidad**, que a modo de catalizador identificaba los lugares donde había más movimientos, tanto de entrada como de salida, ya que se consideraba que tanto la llegada de personas a un lugar como la salida del mismo (que luego volverían a su residencia) podían ser focos de contagio.

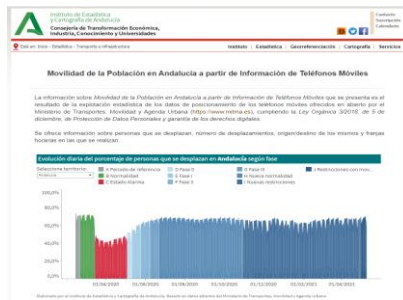


Este informe se alimentaba y actualizaba automáticamente todos los días con los últimos datos disponibles

Publicación de Movilidad

También se optó por darle más visibilidad a los datos de movilidad del informe y en el mes de marzo de 2021 se publicó un nuevo producto en el que se mostraba una visualización de los datos de movilidad del Ministerio de Transportes, Movilidad y Agenda Urbana. En estos momentos los datos del MITMA ya estaban en abierto es decir que todo el mundo tenía acceso a ellos, pero consideramos que era interesante hacer una

publicación fundamentalmente divulgativa que los diera a conocer más, por ello lo que primó fueron los gráficos de Tableau.



Se trata de una explotación básica, que pretende fundamentalmente mostrar los cambios diarios, no la estructura de los movimientos.

El IECA consideró que la gestión de los datos al nivel de detalle que ofrecía el ministerio era de difícil asimilación para usuarios no preparados (fundamentalmente por su volumen no por su complejidad). Por lo que para que los usuarios pudieran iniciarse en el mundo de los microdatos publicamos también los microdatos para Andalucía y agrupados por provincias y con menos variables que los datos originales, dejando siempre claro que los datos completos los tenían en abierto el ministerio a un nivel de detalle mayor.

El ministerio dejó de ofrecer datos el pasado 9 de mayo, con el fin del Estado de Alarma. Según consta en su página web se está gestionando un nuevo contrato que permitirá continuar con la serie. De momento esté producto está parado de actualizaciones.

Análisis de movilidad en la ciudad de Sevilla

Los datos de movilidad del ministerio permiten análisis territoriales más complejos que los descriptivos expuestos.

Como ejemplo de un estudio de movilidad urbana dejo el enlace a un blog en el que, utilizando la información de MITMA, se ha realizado un buen trabajo de análisis en base a los datos de posicionamiento de teléfonos móviles.

<https://jaimesierra.wordpress.com/2020/12/29/movilidad-y-big-data-asi-se-mueve-sevilla/>

Conclusiones

Los teléfonos móviles se han convertido en una fuente inagotable de datos, y el hecho de que siempre queramos llevarlos encima los convierten en un excelente sensor de movimiento y por tanto de análisis de la movilidad.

La pandemia ha favorecido romper con las reticencias existentes a la utilización de los datos obtenidos de ellos. Esa “necesidad” de monitorización que ha obligado la crisis sanitarias ha mostrado los datos como unos buenos aliados en la gestión de la misma.

Consideramos que los datos de movilidad obtenidos a partir de los teléfonos móviles han llegado para quedarse entre nosotros y formar parte de nuestras estadísticas sobre desplazamientos.

En general incluso podríamos afirmar que los datos procedentes de la huellas de nuestro comportamiento digital son una fuente de información muy interesante que debe convertirse en una nueva fuente de información.

Principales referencias bibliográficas

- Metodología datos telefonía del INE:
https://www.ine.es/experimental/movilidad/exp_em1_proyecto.pdf
- Información MITMA: <https://www.mitma.es/ministerio/covid-19/evolucion-movilidad-big-data>
- Ejemplo de análisis de movilidad:
<https://jaimesierra.wordpress.com/2020/12/29/movilidad-y-big-data-asi-se-mueve-sevilla/>