

# Uso de información complementaria en las estadísticas de turismo

Fernando Cortina García, María Izquierdo Valverde, Jesús Prado Mascuñano y María Velasco Gimeno  
Subdirección de Estadísticas de Turismo, Ciencia y Tecnología. INE

## INTRODUCCIÓN

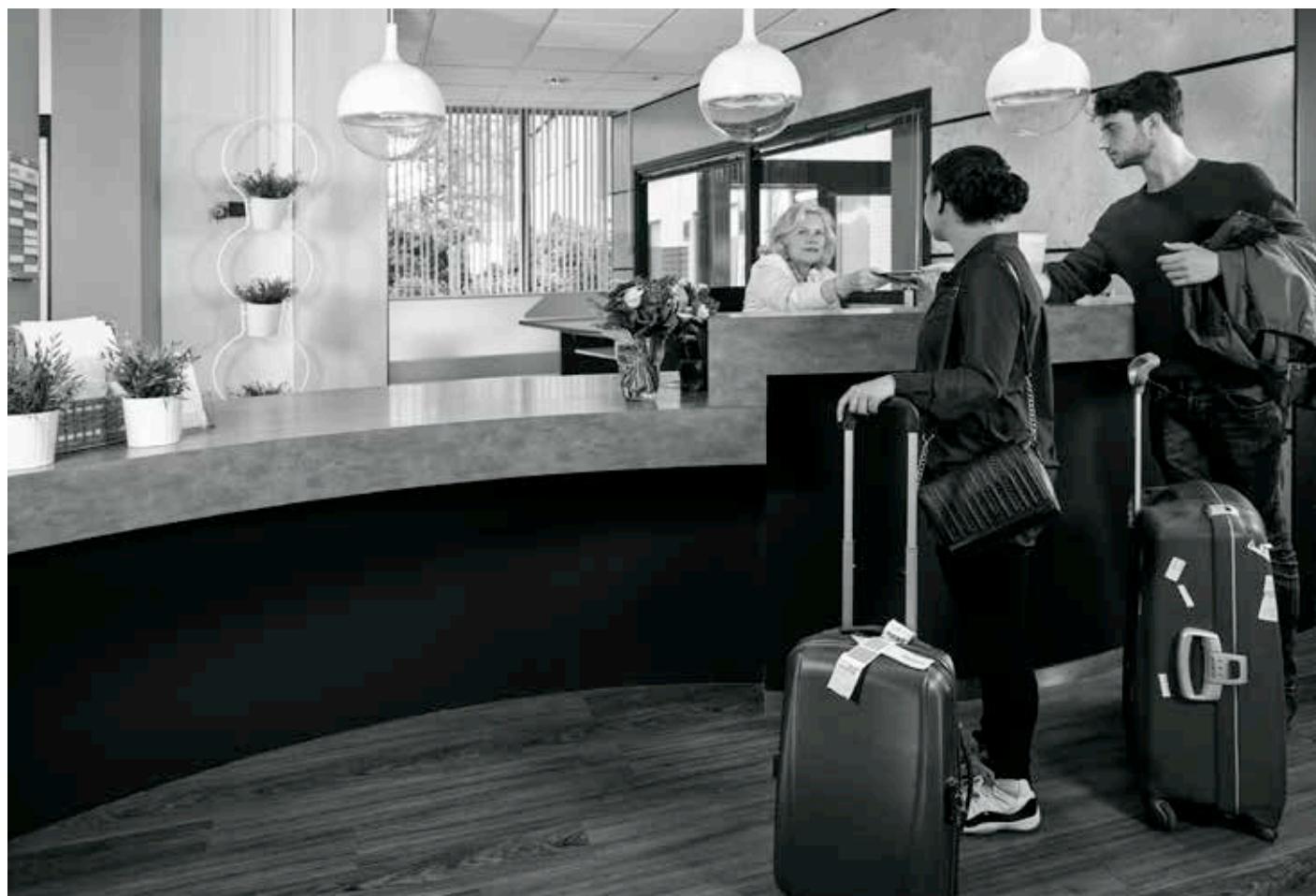
Según la Cuenta Satélite del Turismo de España, en 2014 el turismo representó un 10,9% del PIB y generó un 12,7% del total de puestos de trabajo. Hablamos, por tanto, de un sector cuyo seguimiento recibe una amplia atención por parte de los agentes económicos y sociales a nivel nacional e internacional, pero también en el ámbito regional y local.

España cuenta con un completo sistema de estadísticas de turismo conformado por las estadísticas de oferta (principalmente las Encuestas de Ocupación en Alojamientos Turísticos y sus índices asociados) y las estadísticas de demanda, que estiman los flujos de visitantes internacionales que recibimos cada mes (FRONTUR) y su gasto asociado (EGATUR), así como la actividad turística

de los residentes tanto dentro de España como en el extranjero (ETR/Familitur). La Cuenta Satélite del Turismo sintetiza además ambas perspectivas.

El componente interregional del turismo hace de la comparabilidad un rasgo esencial de la calidad de los datos. Nuestras estadísticas oficiales están plenamente adaptadas al marco de trabajo desarrollado por los manuales y recomendaciones de la Organización Mundial del Turismo<sup>1,2</sup> y de Eurostat<sup>3</sup>, por lo que cuentan con una fuerte base metodológica y un nivel de calidad ampliamente reconocido.

Sin embargo, en un contexto europeo de gran movilidad de la población, de ausencia de fronteras y de cambios en el comportamiento de los consumidores, las necesidades de información son muchas y el sistema debe prepararse para incorporar las nuevas fuentes de información disponibles.



Los datos generados por fuentes no puramente estadísticas ofrecen la posibilidad de mejorar la relevancia, oportunidad y puntualidad de los productos ofrecidos bajo el paraguas de la calidad de la estadística oficial. Ejemplos de estas nuevas fuentes de información son los dispositivos de control de tráfico, los datos de posicionamiento de dispositivos móviles o la actividad de las tarjetas de crédito durante un viaje.

Todas ellas han sido exploradas en mayor o menor medida en el marco del sistema español de estadísticas de turismo y en este artículo se resumen las distintas experiencias.

### EL USO DE LA INFORMACIÓN DE DISPOSITIVOS DE CONTROL DEL TRÁFICO

Una de las fuentes de *Big Data* cuyo potencial está siendo explorado en diferentes campos de la estadística oficial son los dispositivos de control de tráfico que posibilitan el cálculo automático del número de vehículos que pasan por un determinado punto.

En España, desde hace años se incorpora la información proporcionada por la Dirección General de Tráfico (DGT) en el cálculo del número de turistas internacionales que acceden a nuestro país por carretera. Esta fuente de información es imprescindible desde que, por el Acuerdo de Schengen, en vigor desde 1995, se eliminaron los controles en las fronteras internas de los países firmantes del mismo. Es bien conocido que Portugal, Francia y España permiten la libre circulación de personas entre sus fronteras.

Las espiras son unos dispositivos colocados bajo el pavimento que registran el número de vehículos que pasan sobre ellas y permiten clasificarlos según su longitud. Es la propia DGT la que procesa los registros generados en las espiras colocadas en la frontera y cada mes proporciona al INE los microagregados necesarios para el proceso de estimación.

Esta información se completa con la procedente de cámaras de control del tráfico equipadas con un lector de matrículas, también propiedad de la DGT. Las situadas en los puntos fronterizos proporcionan un registro completo, debidamente anonimizado, de los vehículos que visitan nuestro país clasificados por nacionalidad de la matrícula.

Combinando estas dos fuentes de información con otras operaciones de campo complementarias que permiten conocer la ocupación media de los vehículos y el tipo de visitante, se obtiene el núme-

ro de turistas y de excursionistas internacionales que entran por carretera cada mes en España.

Este es un ejemplo de éxito en el uso estadístico de información generada automáticamente por dispositivos diseñados y utilizados para otros fines. Una comunicación ágil y una voluntad decidida de cooperación a nivel institucional son esenciales en el funcionamiento de este proyecto, que depende del correcto funcionamiento de los dispositivos para asegurar la disponibilidad de la información.

Actualmente se están explorando las posibilidades que ofrecen las bases de datos de cámaras en la identificación del excursionismo, tanto en el caso de los visitantes internacionales como en las salidas al extranjero de los residentes en España. El seguimiento de una matrícula debidamente anonimizada a través de su huella en la base de datos podría servir para determinar el tipo de visitante. Disponer de información longitudinal y el seguimiento de la matrícula más allá de la frontera ayudaría a determinar el país de residencia habitual así como el entorno habitual del usuario del vehículo.

### ANÁLISIS PRELIMINAR DE LOS DATOS DE TELEFONÍA MÓVIL

Una de las fuentes de datos más populares cuando se habla de *Big Data* son las bases de datos de los operadores de telefonía móvil, donde se almacenan las interacciones de los dispositivos móviles con las antenas que componen las redes de telefonía. Cada dispositivo tiene asociado un número de identificación cuyo seguimiento, debidamente encriptado, puede aportar valiosa información sobre la movilidad del portador del aparato. Estos movimientos se podrían asociar además al perfil del usuario, tomando en todo momento las adecuadas medidas de protección de la confidencialidad.

No cabe duda de que para el análisis del turismo esta información ofrece un potencial enorme, ya que permitiría obtener datos con un grado de detalle geográfico que las encuestas no pueden alcanzar. En el escenario más favorable, se podrían reducir las muestras utilizadas en las operaciones actuales, que se limitarían a recoger información sobre el alojamiento utilizado, las actividades realizadas o el gasto efectuado.

Ante esta situación, Eurostat encargó un estudio, cuyos resultados se publicaron en 2014, sobre la viabilidad del uso de los datos de telefonía móvil en las estadísticas de turismo<sup>4</sup>. Lamentablemente, una de las principales conclusiones del estudio

fue la constatación de las dificultades para acceder a esta información, principalmente por restricciones legales. En un gran número de países de la Unión Europea los Institutos Nacionales de Estadística están trabajando conjuntamente con los operadores de telefonía móvil para tratar de solventar estas dificultades y encontrar una forma de colaboración que resulte ventajosa para todos los actores implicados. También se están haciendo esfuerzos a nivel europeo en este sentido.

En España, el INE ha realizado un primer análisis de los datos de telefonía móvil a partir de un conjunto de agregados solicitados a un operador nacional. El objetivo era estudiar si las cifras derivadas de este primer análisis resultaban comparables con las estimaciones oficiales. En concreto, los indicadores objeto de estudio fueron la distribución del total de turistas por origen y destino y la duración media de la estancia.

El INE obtuvo datos referidos a un mes determinado, clasificados por comunidad autónoma o país de origen y por destino.

La complejidad de los datos capturados por una antena requiere transformarlos en un conjunto de información válido para el tratamiento estadístico. Este proceso lo llevó a cabo el operador de telefonía móvil siguiendo sus propios algoritmos. A partir de ahí, el INE trató de adaptar las definiciones de turismo internacionalmente aceptadas a las posibilidades de la base de datos. Este es un trabajo que precisa una interacción continua con el operador, el análisis detallado de los resultados, la detección de diferencias sistemáticas y la elaboración de conclusiones que permitan el refinamiento de las condiciones inicialmente establecidas para empezar de nuevo el ciclo hasta alcanzar resultados satisfactorios. Este primer paso es fundamental para un futuro aprovechamiento de los datos.

Comparando los datos de los residentes obtenidos con los datos de posicionamiento móvil con los de la encuesta FAMILITUR, se observó una estructura similar en los principales destinos, pero en algún caso se encontraron diferencias importantes. Un análisis más detallado mostró que la proporción de viajes dentro de la comunidad autónoma es sistemáticamente superior en los datos de telefonía móvil que en la encuesta, pero en mayor medida en la comunidad donde se daban las mayores diferencias.

También se detectaron diferencias sistemáticas en la variable estancia media. En los viajes internos realizados por los residentes, los datos de telefonía móvil estiman una estancia media más de un 50% más larga que la cifra estimada por FAMILITUR.



En el caso de los no residentes que nos visitan, la distribución de turista por país de origen también resultó muy similar en ambas fuentes, con alguna discrepancia un poco más acentuada en algún mercado concreto. Sin embargo, la estancia media estimada fue sistemáticamente inferior en los datos de telefonía móvil que en las cifras oficiales.

Estos hechos parecen indicativos de la necesidad de mejorar la determinación de la residencia habitual, de identificar posibles ruidos en las zonas fronterizas y en las áreas metropolitanas, etc. Conviene ahondar también en las diferencias conceptuales que puede haber tras estas discrepancias y analizar críticamente cada una de las fases del proceso de extracción de información de los datos de posicionamiento móvil.

#### APROXIMACIÓN A LOS DATOS DE LOS SISTEMAS DE PAGO CON TARJETA DE CRÉDITO

Otra de las nuevas fuentes de información de gran interés son los registros de las transacciones efec-

*Las nuevas fuentes de información ofrecen un gran potencial en el ámbito del turismo. De ellas se podrán obtener datos con un nivel de detalle necesario para los usuarios pero que una operación por muestreo no puede cubrir*

*Los primeros análisis realizados ponen de manifiesto la necesidad de establecer un marco metodológico y conceptual que dote de validez a los nuevos indicadores y que establezca la convivencia de los mismos con los datos actuales*

tuadas con tarjetas de crédito, como son los pagos en los terminales de venta (TPV) de distintos establecimientos o extracciones de efectivo en cajeros automáticos.

El INE solicitó a un operador nacional una serie de agregados obtenidos de las transacciones efectuadas por sus clientes con las tarjetas de la entidad para un periodo de dos años. El objetivo del análisis era la comparación del gasto medio por viaje de los residentes.

Como en el caso anterior, el trabajo directo con la base de datos fue realizado por el operador, que proporcionó al INE resultados agregados. De nuevo, la traducción de las definiciones de turismo en un algoritmo fue, y sigue siendo, el principal reto a desarrollar.

Teniendo en cuenta que la cobertura del gasto recogido no es la misma en las dos fuentes, cabría esperar que de los datos de tarjetas se derivase un

gasto medio inferior al de la encuesta. En el análisis se observaron diferencias sistemáticas pero en dirección opuesta a la esperada.

## WEB SCRAPING

Otra técnica de obtención masiva de datos es el webscraping, que consiste en la recopilación de información de un sitio web de forma automática. El INE está desarrollando un proyecto piloto para la obtención de precios de los paquetes turísticos, por un lado, y del alojamiento y el transporte incluido en cada paquete si se adquiriesen de forma independiente. El objetivo es la descomposición del precio de los paquetes turísticos en sus principales componentes: alojamiento y transporte.

Este desglose es exigido a nivel europeo en las estadísticas de turismo de los residentes como consecuencia de las necesidades de la Cuenta Satélite del Turismo<sup>5</sup> y de la Balanza de Pagos. La incorporación de información auxiliar en los procedimientos actuales posibilitaría obtener información detallada por destinos e incluso distinguir la nacionalidad del proveedor del servicio, permitiendo estimar el porcentaje del importe del paquete turístico que repercute en la economía de origen o de destino.

## CONCLUSIONES

Las nuevas fuentes de información ofrecen un gran potencial en el ámbito del turismo. De ellas se podrán obtener datos con un nivel de detalle necesario para los usuarios pero que una operación por muestreo no puede cubrir.

Los primeros análisis realizados ponen de manifiesto la necesidad de establecer un marco metodológico y conceptual que dote de validez a los nuevos indicadores y que establezca la convivencia de los mismos con los datos actuales. Las estadísticas oficiales han alcanzado unos estándares de calidad muy altos que deben preservarse y, cómo no, actualizarse. Los esfuerzos dedicados a esta actualización y a la incorporación de estas nuevas fuentes de datos deben realizarse de forma coordinada en el Sistema Estadístico Europeo.

Este trabajo de conceptualización requiere, además, una gran colaboración entre las oficinas de estadística y los proveedores de datos. La información proporcionada por los mismos debe estar acompañada de una exhaustiva documentación auxiliar y la comunicación fluida entre todos los actores resulta esencial para salvaguardar la calidad de los resultados, especialmente en estas fases iniciales.

## Referencias

1. OMT, Recomendaciones internacionales para las estadísticas de turismo, (RIET, 2008).
2. OMT, International Recommendations for Tourism Statistics 2008 Compilation guide, (2014).
3. Regulation (EU) No 692/2011 of the European Parliament and of the Council of 6 July 2011 concerning European statistics on tourism.
4. EUROSTAT, Ahas R., Armoogum J., Esko S., Ilves M., Karus E., Madre JL., Nurmi O., Potier F., Schmücker D., Sonntag U., Tiru M. (2014), Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics, Consolidated Report, Publications Office of the European Union.
5. United Nations, World Tourism Organization, Eurostat, OECD (2010), Tourism Satellite Account: Recommended Methodological Framework 2008, United Nations publication.