

# Proyecto de evaluación de la Movilidad Interprovincial de Viajeros aplicando la Tecnología Big Data

OTLE - Jornada anual 6ª edición



**Juan Carlos Huertas de Andrés**  
**Ministerio de Fomento**

**marzo de 2019**

# Índice

- **Situación de partida**
- **Objeto y alcance del proyecto**
- **Metodología: identificación de viajes, validación y ajuste**
- **Resultados**
- **Conclusiones y futuras líneas de trabajo**



Situación de partida

# Situación de partida

**Necesidad:** Disponer de datos actualizados y de calidad sobre la movilidad de la población, indispensables para planificar. Actualizar Movilia 2007.

**Oportunidad:** Estudiar nuevas fuentes de datos para caracterizar la movilidad frente a la técnica tradicional de encuestas. BIG DATA



# Situación de partida

## Nueva fuente de datos: huella digital

Dos grandes grupos de huella digital (cada vez más imbricados):

- Huella digital en el espacio “on line”
- **Huella digital en el espacio físico:** Las nuevas tecnologías permiten registrar datos geolocalizados de la población, tomados de la realidad, con un nivel de resolución espacio/temporal que permiten obtener resultados solventes sobre movilidad o sus pautas de comportamiento:
  - Tarjetas de transporte público: datos de acceso y salida de la red
  - navegadores: datos de su uso geolocalizados en carretera.
  - Apps: datos de recorridos geolocalizados...
  - Redes sociales: opción de geolocalizar mensajes
  - Telefonía móvil: datos de eventos registrados por las antenas (call detail records “CDR”) tanto activos y como pasivos.
  - .....

# Situación de partida

## Opción para el trabajo: posicionamiento de terminales de telefonía móvil

- Recogida pasiva de datos: sin molestias al usuario.
- Es oportunista: Los datos se han recogido para otro propósito. No requiere infraestructura adicional.
- Resolución espacio/temporal razonable para el reto planteado.
- Muestra muy grandes: 20-30% de la población. Móviles servidos por cada uno de los tres grandes operadores de telefonía.
- Los datos son longitudinales: continuos en el tiempo (segundo/minuto/.../año/...). No son “foto fija” como encuestas.
- Evita subjetividad: se analiza lo que el ciudadano hace, no lo que dice que hace.
- Permiten una importante reducción coste y tiempo en la elaboración del trabajo.



# Situación de partida

## Algunos datos del trabajo:

- Analizados más de 14 millones de móviles, 27,3% cuota mercado. Adjudicatario del trabajo Orange/Kineo.
- Volumen de datos: 1.000 M de registros/día medio Orange.
- Más de 3 teras/mes analizados para el trabajo

## Comparación con metodología clásica:

- Mayor dificultad en determinar el motivo. Sí movilidad obligada casa/trabajo.
- Dificultad en determinar modo autobús
- Imposibilidad de encuestas de preferencias declaradas. Valor del tiempo.



# Situación de partida

## Imperativo tecnología BIG DATA:

- La privacidad de las personas debe de ser absolutamente respetada.
  - Estricto cumplimiento normativa de carácter personal actualmente vigente, recogida principalmente en:
    - Reglamento (UE) 2016/679 General de Protección de Datos
    - Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales.
  - Seguimiento de las pautas establecidas en las guías elaboradas por la Agencia Española de Protección de Datos





# Objeto y Alcance del proyecto

# Objetivo del proyecto

Estudiar la movilidad de viajeros interprovincial, a nivel nacional, en cada uno de los cuatro modos de transporte (carretera, ferrocarril, marítimo y aéreo), aplicando la tecnología Big Data



- **Periodo de estudio:**
  - 2 periodos / 29 grupos

Julio  
Agosto

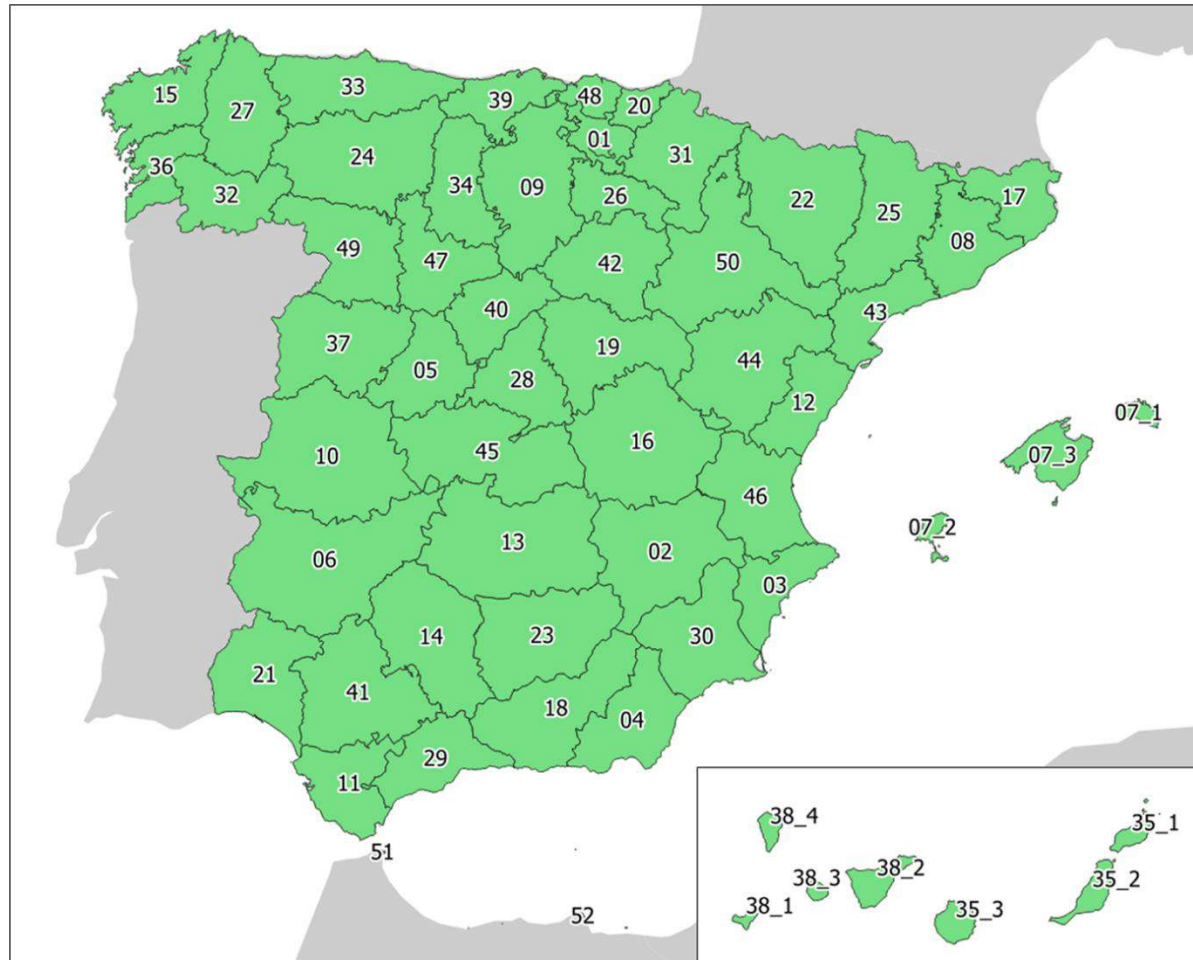
|       | L              | M               | X                  | J               | V                | S               | D                |
|-------|----------------|-----------------|--------------------|-----------------|------------------|-----------------|------------------|
| Mes 1 | 10             | 11              | 12                 | 13              | 14               | 15              | 16               |
|       | 17             | 18              | 19                 | 20              | 21               | 22              | 23               |
|       | 24             | 25              | 26                 | 27              | 28               | 29              | 30               |
|       | 31             | 1               | 2                  | 3               | 4                | 5               | 6                |
|       | 7              | 8               | 9                  |                 |                  |                 |                  |
|       | Lunes Promedio | Martes Promedio | Miércoles Promedio | Jueves Promedio | Viernes Promedio | Sábado Promedio | Domingo Promedio |

Octubre

|       | L              | M               | X                  | J               | V                | S               | D                |
|-------|----------------|-----------------|--------------------|-----------------|------------------|-----------------|------------------|
| Mes 2 |                |                 |                    |                 |                  |                 | 1                |
|       | 2              |                 | 4                  | 5               | 6                | 7               | 8                |
|       | 9              | 10              | 11                 | 12              | 13               | 14              | 15               |
|       | 16             | 17              | 18                 | 19              | 20               | 21              | 22               |
|       | 23             | 24              | 25                 | 26              | 27               | 28              | 29               |
|       | 30             | 31              |                    |                 |                  |                 |                  |
|       | Lunes Promedio | Martes Promedio | Miércoles Promedio | Jueves Promedio | Viernes Promedio | Sábado Promedio | Domingo Promedio |

Días a analizar de manera individual debido a sus previsibles características especiales

- Zonificación



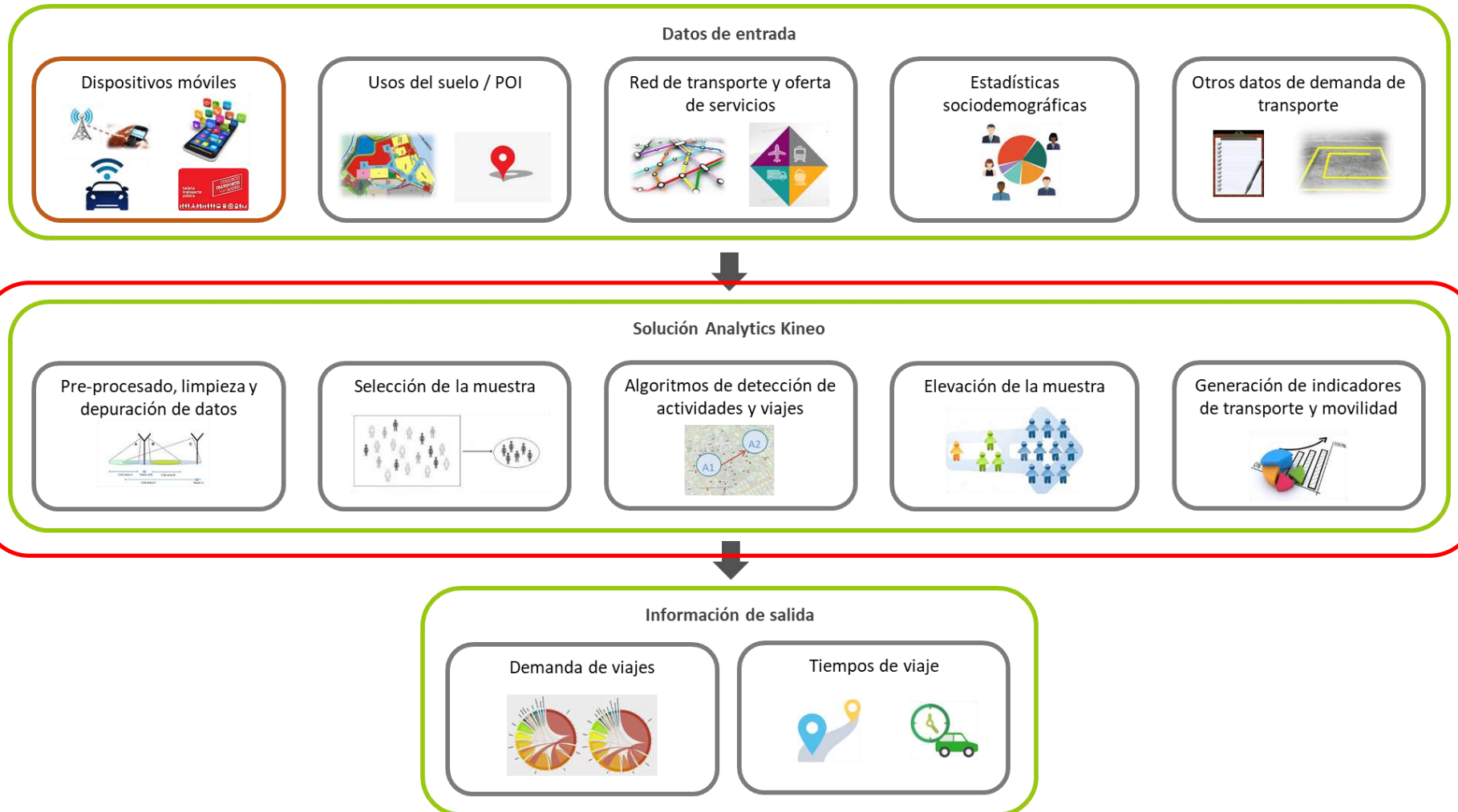
- **Viajes y etapas a analizar:**
  - De manera general: viajes cuyos itinerarios superan los 50 km de longitud
  - Madrid, Barcelona, Alicante y Vizcaya → viajes de entre 10 y 50 km
  - Etapas interprovinciales para los viajes de estudio
- **Segmentación por modos de transporte:**
  - De manera general: carretera, ferrocarril, marítimo y aéreo
  - Carretera segmentada en vehículo privado y autobús para ciertos pares OD
  - Identificación de vehículos pesados → eliminados de la muestra
- **Franjas horarias: según hora de inicio del viaje:**
  - P1: 00:00 – 06:00
  - P2: 06:00 – 10:00
  - P3: 10:00 – 17:00
  - P4: 17:00 – 00:00



- **Marco muestral:**
  - ***Residentes en España***: población residente en España mayor de 10 años reportada por el Padrón de Habitantes de 2017
  - ***Residentes en el extranjero***: encuesta FRONTUR del INE de los periodos de julio, agosto y octubre de 2017
- **Caracterización de actividades:**
  - ***Casa***: lugar de residencia del viajero
  - ***Trabajo/Estudio***: lugar de trabajo o estudio a jornada completa (6-8 horas)
  - ***Frecuente***: lugar donde se realiza una actividad de manera recurrente distinta de casa y trabajo
  - ***Esporádica***: lugar donde se realiza una actividad no recurrente
- **Recurrencia casa-trabajo/estudio (mensual)**
  - R0: usuarios que no realizan el viaje casa-trabajo ningún día de estudio
  - R1: usuarios que realizan el viaje casa-trabajo entre 1 y 5 días de estudio
  - R2: usuarios que realizan el viaje casa-trabajo entre 6 y 10 días de estudio
  - R3: usuarios que realizan el viaje casa-trabajo entre 11 y 15 días de estudio
  - R4: usuarios que realizan el viaje casa-trabajo entre 16 y 20 días de estudio
  - R5: usuarios que realizan el viaje casa-trabajo más de 20 días de estudio

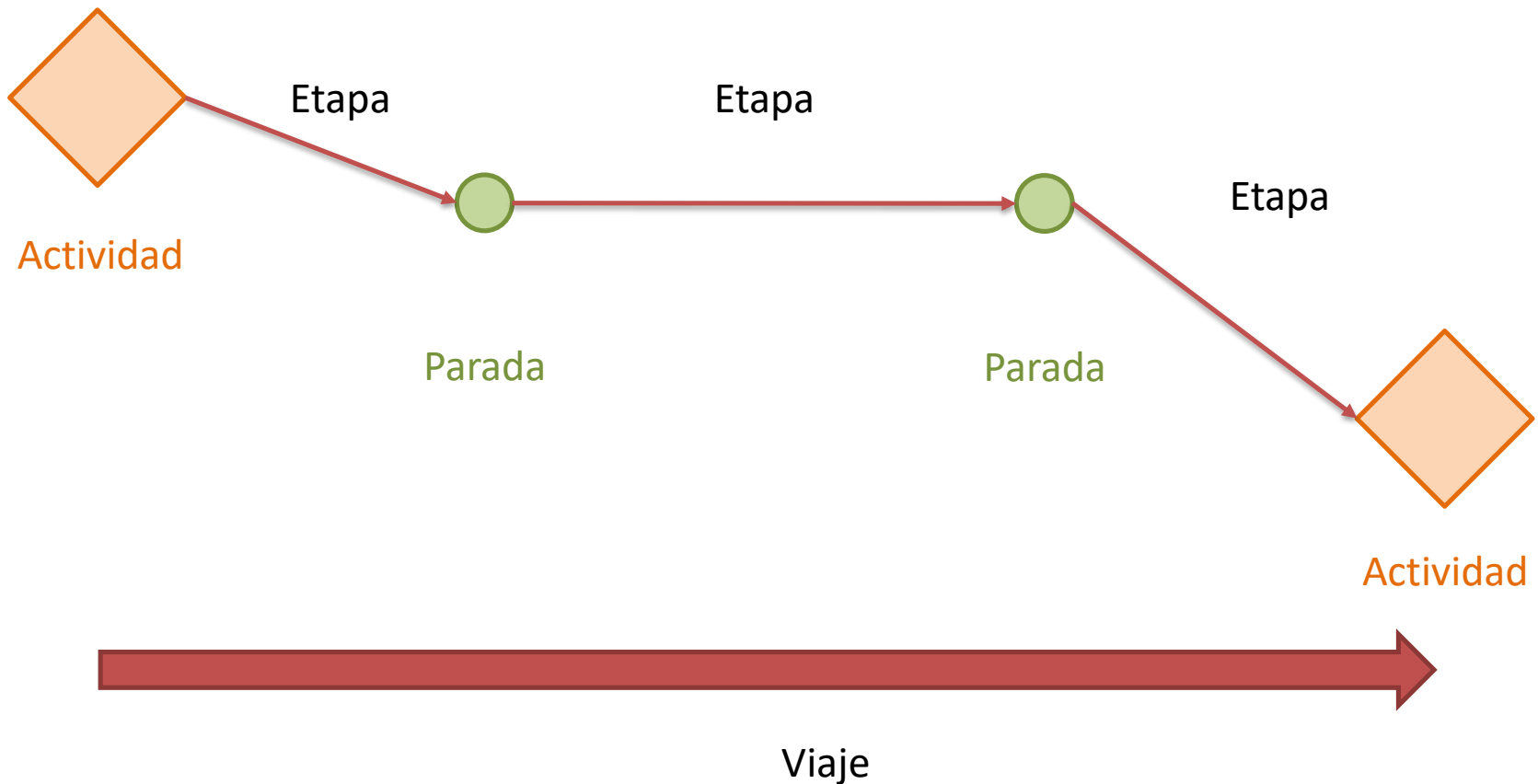
# Metodología

# Esquema general

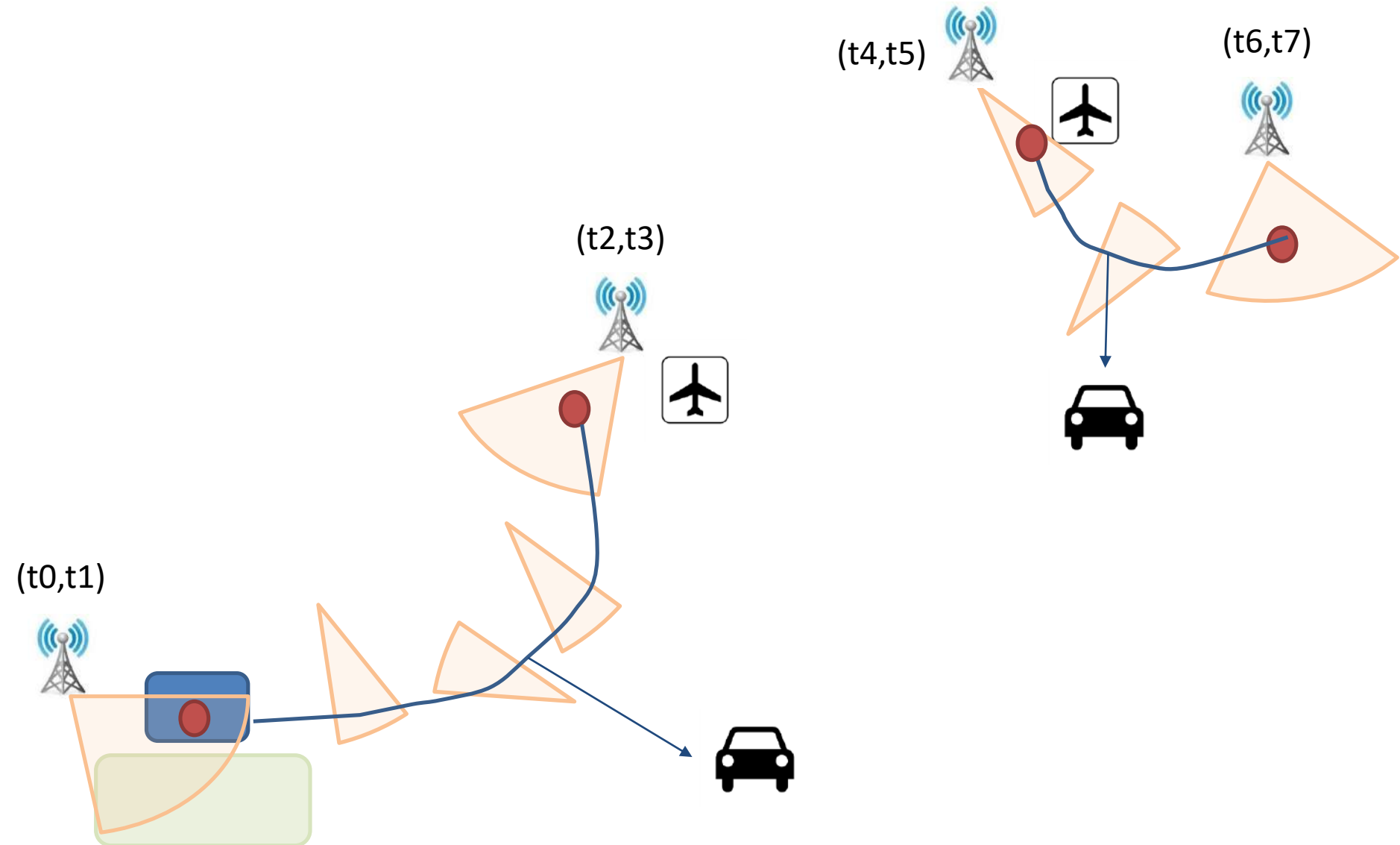




# Diario de actividades y viajes



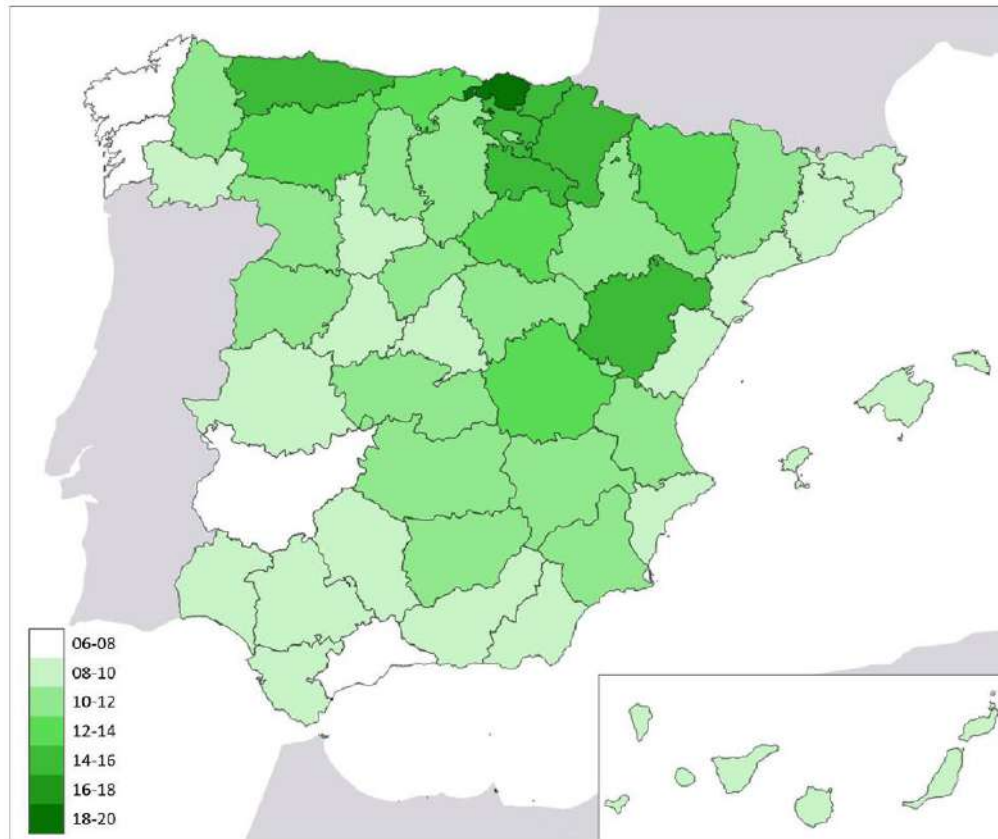
# Detección de actividades y viajes



# Elevación muestral

## Factor de expansión de la población residente

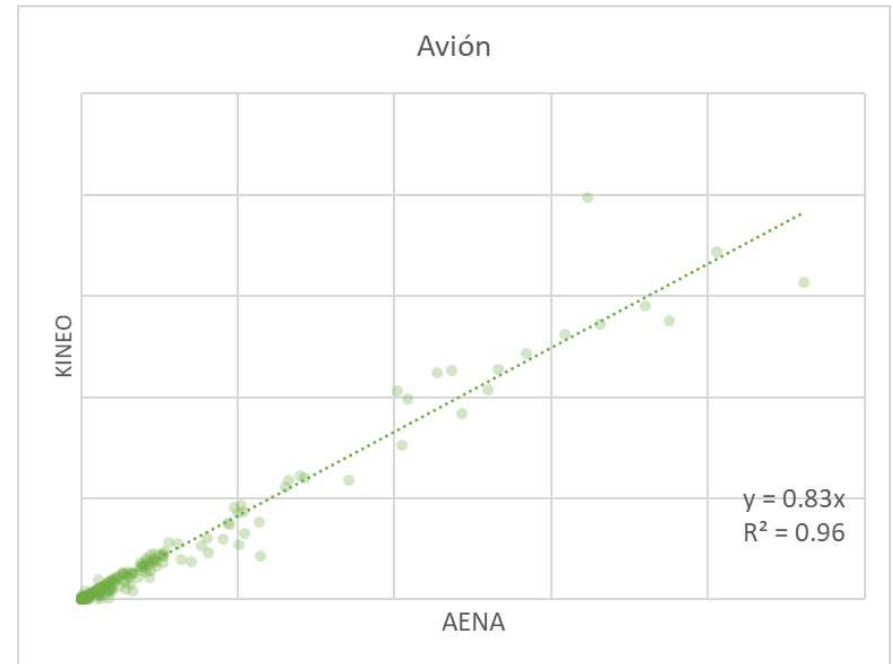
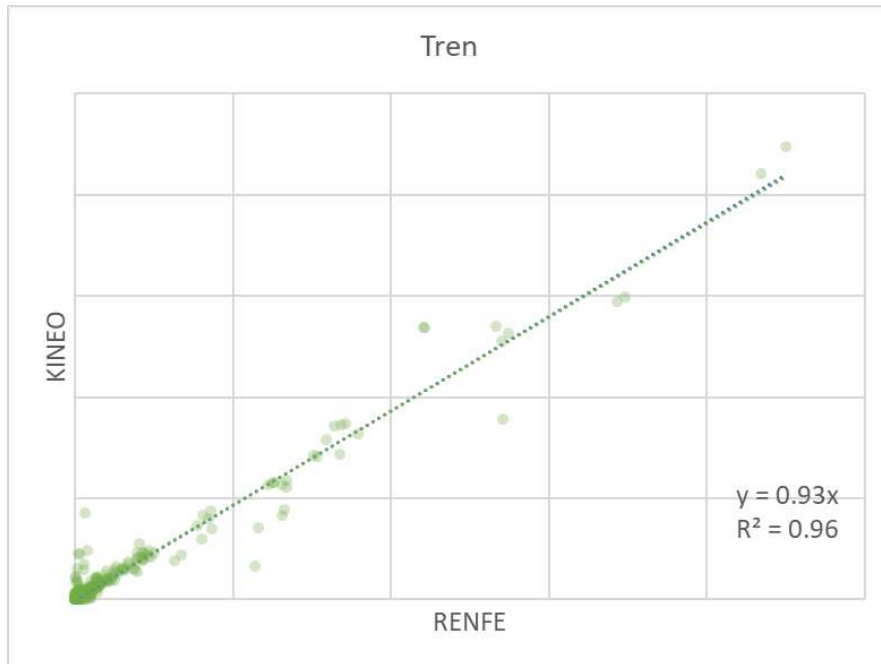
- Factor de expansión medio: 8,29



\* Resultados del miércoles promedio de octubre

# Validación y Ajuste

## Etapas billeteaje vs etapas telefonía



# Resultados

# Listado de entregables

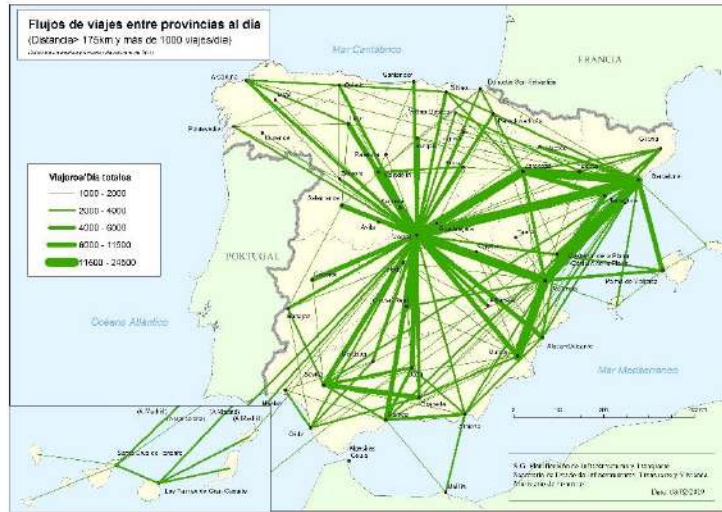
- Matriz de viajes
- Matriz de etapas
- Matriz de tours
- Recurrencia casa-trabajo
- Matriz de rutas

# Matriz de viajes: Especificación

| Matrices de viajes |                   |  |  |
|--------------------|-------------------|--|--|
| Columna            | Nombre            | Valores  | Descripción  |
| 1                  | Origen            | según los valores definidos en el fichero zonificación.shp<br>60: Extranjero                 | Identificador de la zona de origen del viaje   |
| 2                  | Destino           | según los valores definidos en el fichero zonificación.shp<br>60: Extranjero                 | Identificador de la zona de destino del viaje  |
| 3                  | Año               | 2017   | Año de estudio   |
| 4                  | Mes               | 07-08-10   | Meses de estudio   |
| 5                  | Día               | 01-31: días específicos<br>Lunes-domingo: días promedios                                     | Nombre del día de la semana para los días promedio y número de día para los días sueltos |
| 6                  | Periodo           | P1: 00:00 - 06:00<br>P2: 06:00 - 10:00<br>P3: 10:00 - 17:00<br>P4: 17:00 - 00:00             | Franja horaria correspondiente al inicio del viaje                                       |
| 7                  | Modo              | autobús/avión/barco/privado/tren   | Modo de transporte principal del viaje   |
| 8                  | Distancia         | D0: viajes de menos de 10 km.<br>D1: viajes entre 10 y 50 km.<br>D2: viajes de más de 50 km. | Rango de distancia al que pertenecen los viajes  |
| 9                  | Residencia        | Nacionalidad según los valores de la <i>Tabla 1</i>  | País o grupo de países de residencia de los viajeros                                     |
| 10                 | Actividad_Origen  | Casa/trabajo-estudio/frecuente/esporádica  | Tipo de actividad realizada en el lugar de origen del viaje                              |
| 11                 | Actividad_Destino | Casa/trabajo-estudio/frecuente/esporádica  | Tipo de actividad realizada en el lugar de destino del viaje                             |
| 12                 | Viajeros          | Número decimal (3 dígitos)   | Número de viajeros   |
| 13                 | Viajeros-km       | Número decimal (3 dígitos)   | Número de viajeros multiplicado por los kilómetros del viaje                             |

# Matriz de viajes: principales flujos

flujos



Nº de viajes por O/D

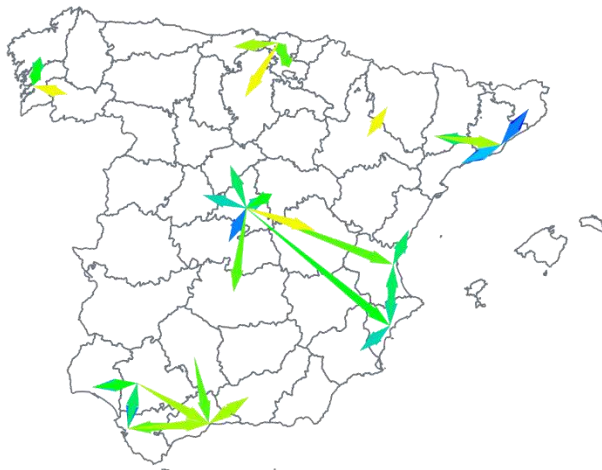




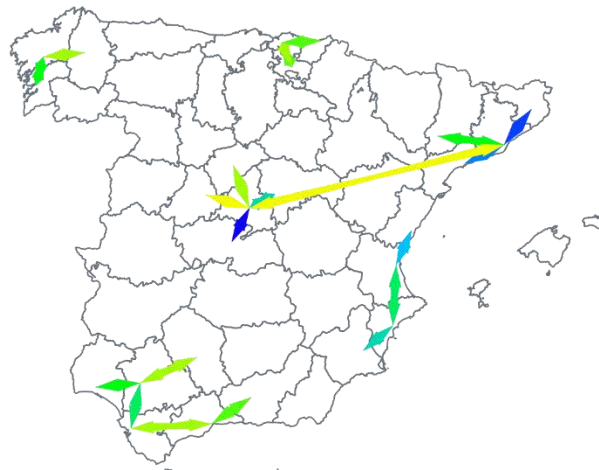
# Matriz de viajes: Principales flujos

- 40 flujos principales de viajes
- Viajes considerados: Viajes de más de 50 km

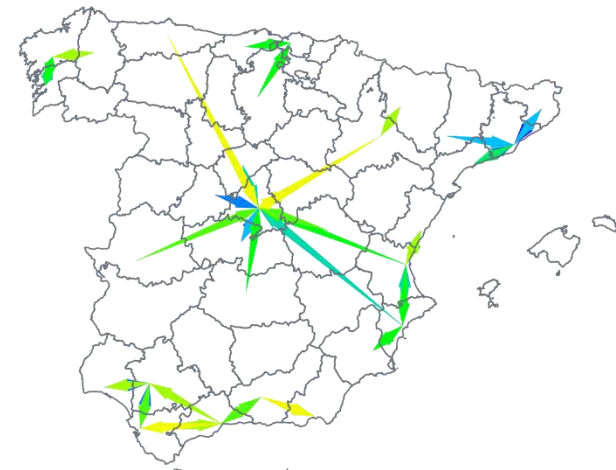
Viernes 14 de julio  
Operación salida



Jueves promedio de  
octubre



Domingo 15 de octubre  
Vuelta del puente



+ nº viajes



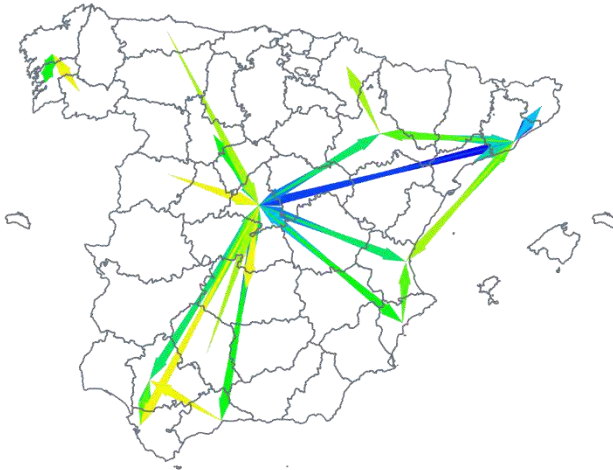
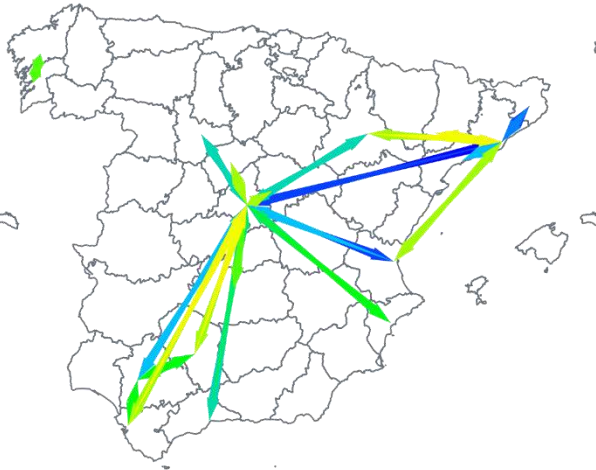
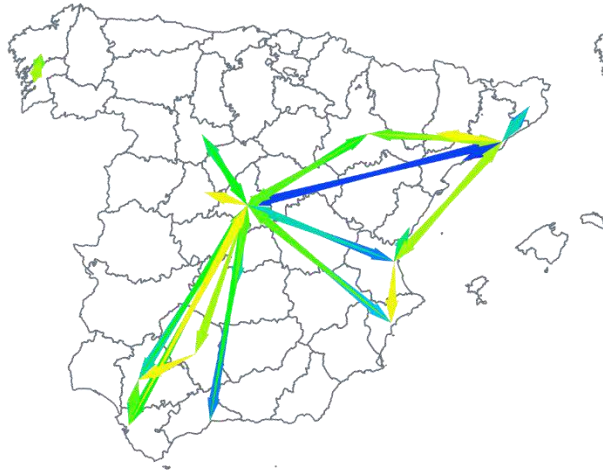
# Matriz de viajes: Principales flujos (modo)

Viernes 14 de julio  
Operación salida

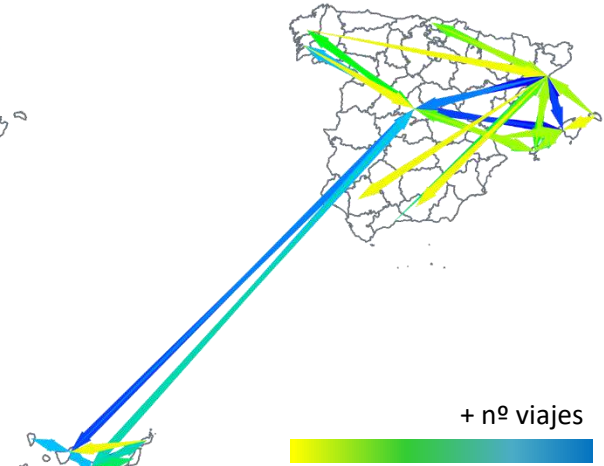
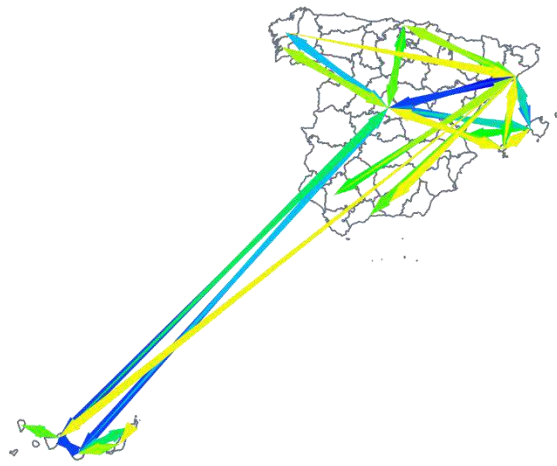
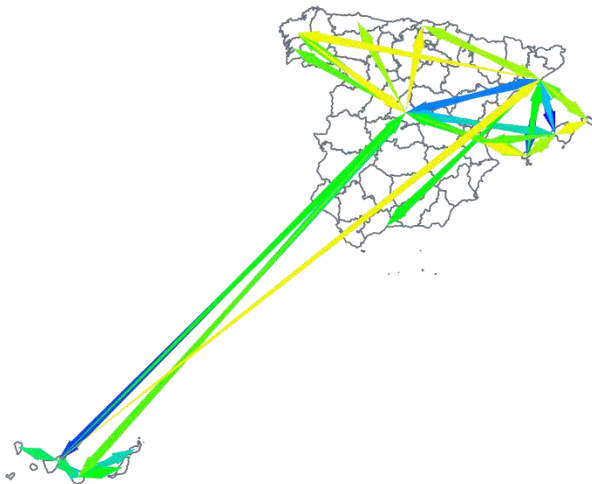
Jueves promedio de  
octubre

Domingo 15 de octubre  
Vuelta del puente

tren



avión

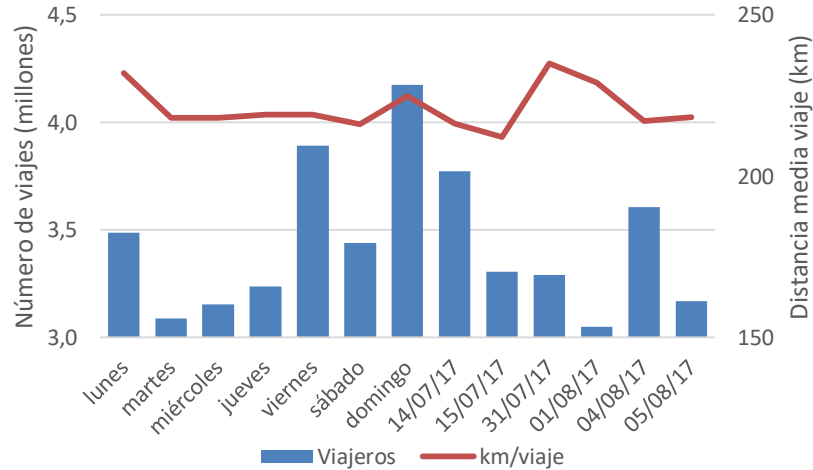


+ nº viajes

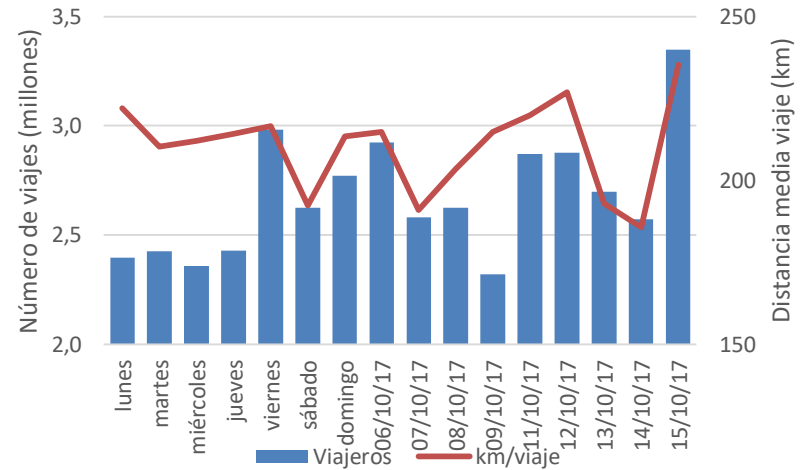


# Matriz de viajes: Volúmenes y distancias

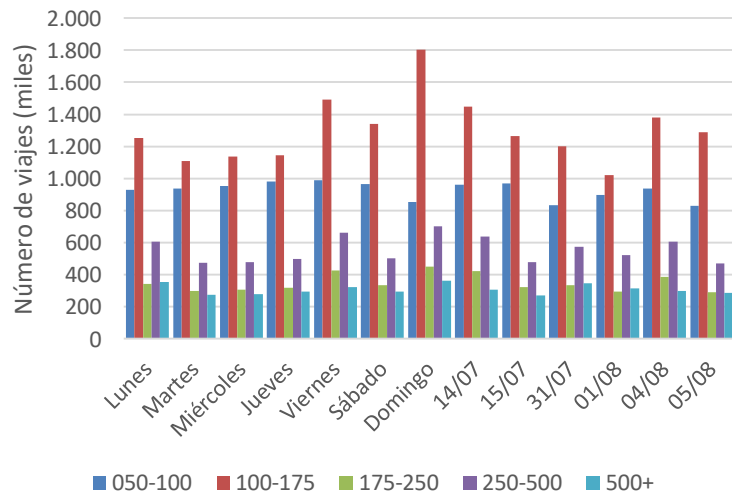
Julio/agosto



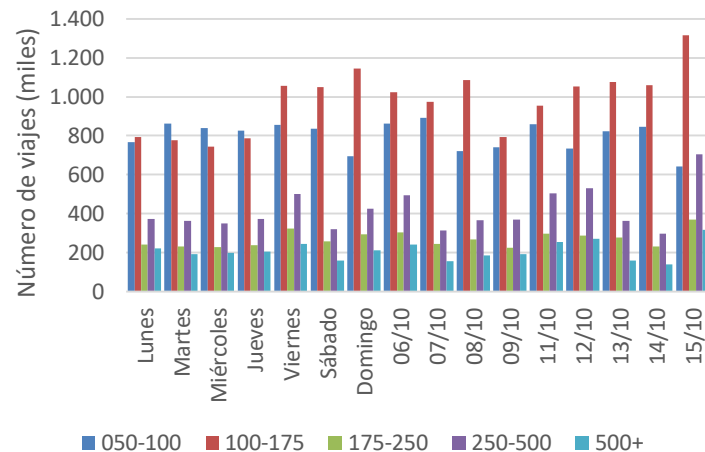
Octubre



Julio/agosto



Octubre

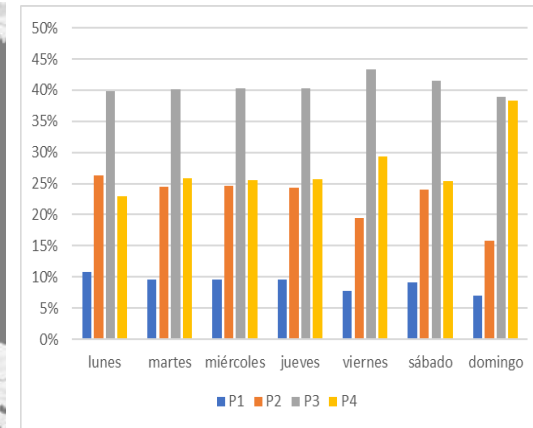
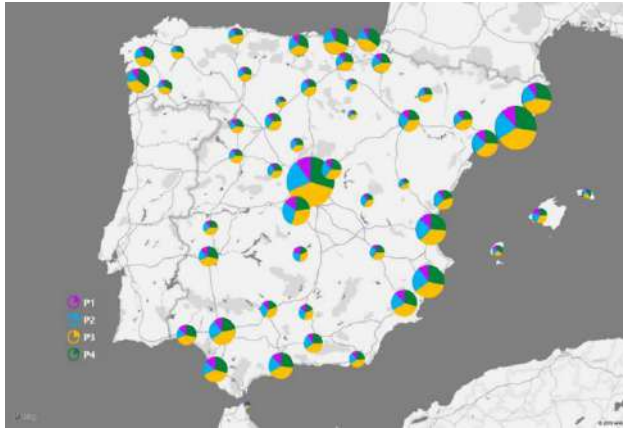


# Matriz de viajes: Reparto horario

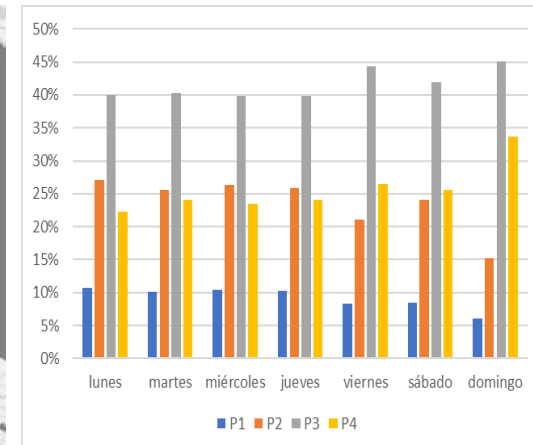
Laborable

Domingo

Julio/agosto



Octubre

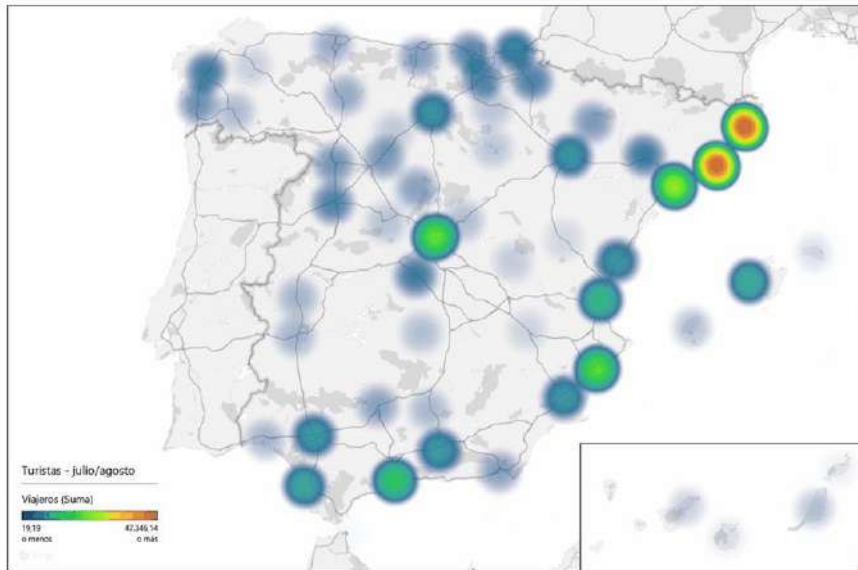


# Matriz de viajes: Viajes de extranjeros

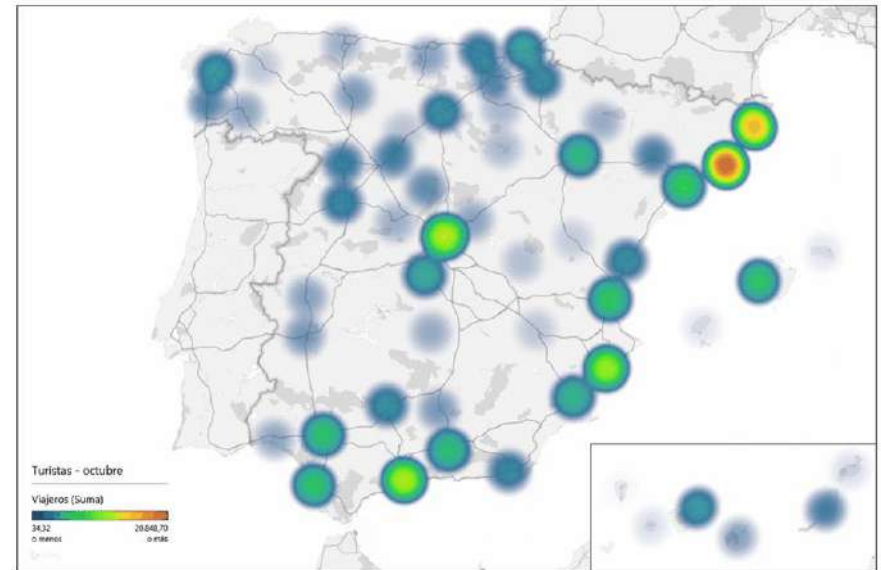
- Aproximación para conocer dónde se concentran los turistas por nacionalidad

Total turistas

julio/agosto



octubre



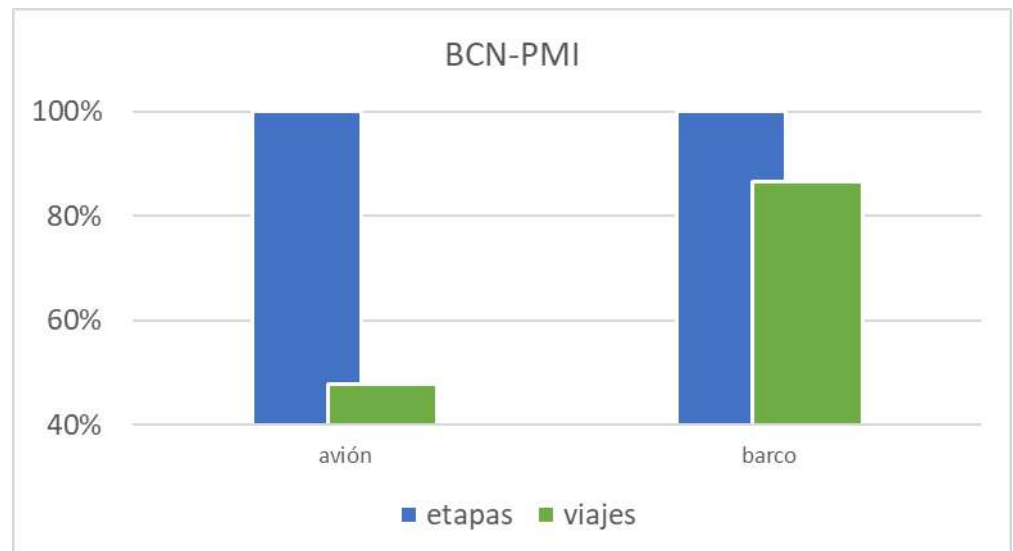
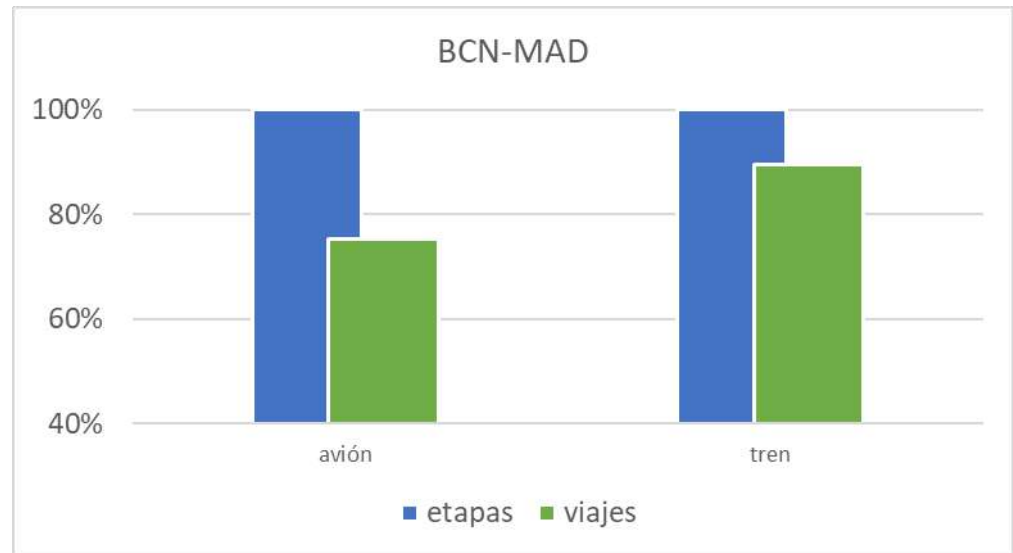


# Matriz de etapas: Especificación

| Matrices de etapas |                   |  |  |
|--------------------|-------------------|--|--|
| Columna            | Nombre            | Valores  | Descripción  |
| 1                  | Origen            | según los valores definidos en el fichero zonificación.shp / 60: Extranjero      | Identificador de la zona de origen del viaje   |
| 2                  | Destino           | según los valores definidos en el fichero zonificación.shp / 60: Extranjero      | Identificador de la zona de destino del viaje  |
| 3                  | Año               | 2017   | Año de estudio   |
| 4                  | Mes               | 07-08-10   | Meses de estudio   |
| 5                  | Día               | 01-31: días específicos<br>Lunes-domingo: días promedios                         | Nombre del día de la semana para los días promedio y número de día para los días sueltos |
| 6                  | Periodo           | P1: 00:00 - 06:00<br>P2: 06:00 - 10:00<br>P3: 10:00 - 17:00<br>P4: 17:00 - 00:00 | Franja horaria correspondiente al inicio del viaje                                       |
| 7                  | Modo              | autobús/avión/barco/carretera/privado/tren                                       | Modo de transporte principal de la etapa   |
| 8                  | Residencia        | Nacionalidad según los valores de la <i>Tabla 1</i>                              | País o grupo de países de residencia de los viajeros                                     |
| 9                  | Actividad_Origen  | Casa/trabajo-estudio/ frecuente/ esporádica                                      | Tipo de actividad realizada en el lugar de origen del viaje                              |
| 10                 | Actividad_Destino | Casa/trabajo-estudio/ frecuente/ esporádica                                      | Tipo de actividad realizada en el lugar de destino del viaje                             |
| 11                 | Viajeros          | Número decimal (3 dígitos)   | Número de viajeros   |
| 12                 | Viajeros-km       | Número decimal (3 dígitos)   | Número de viajeros multiplicado por los kilómetros del viaje                             |

# Relación matriz viajes - etapas

Comparando las matrices de viajes y matrices de etapas, se pueden observar aquellas provincias que actúan como *hub* para las provincias colindantes, como es el caso de Barcelona y Madrid



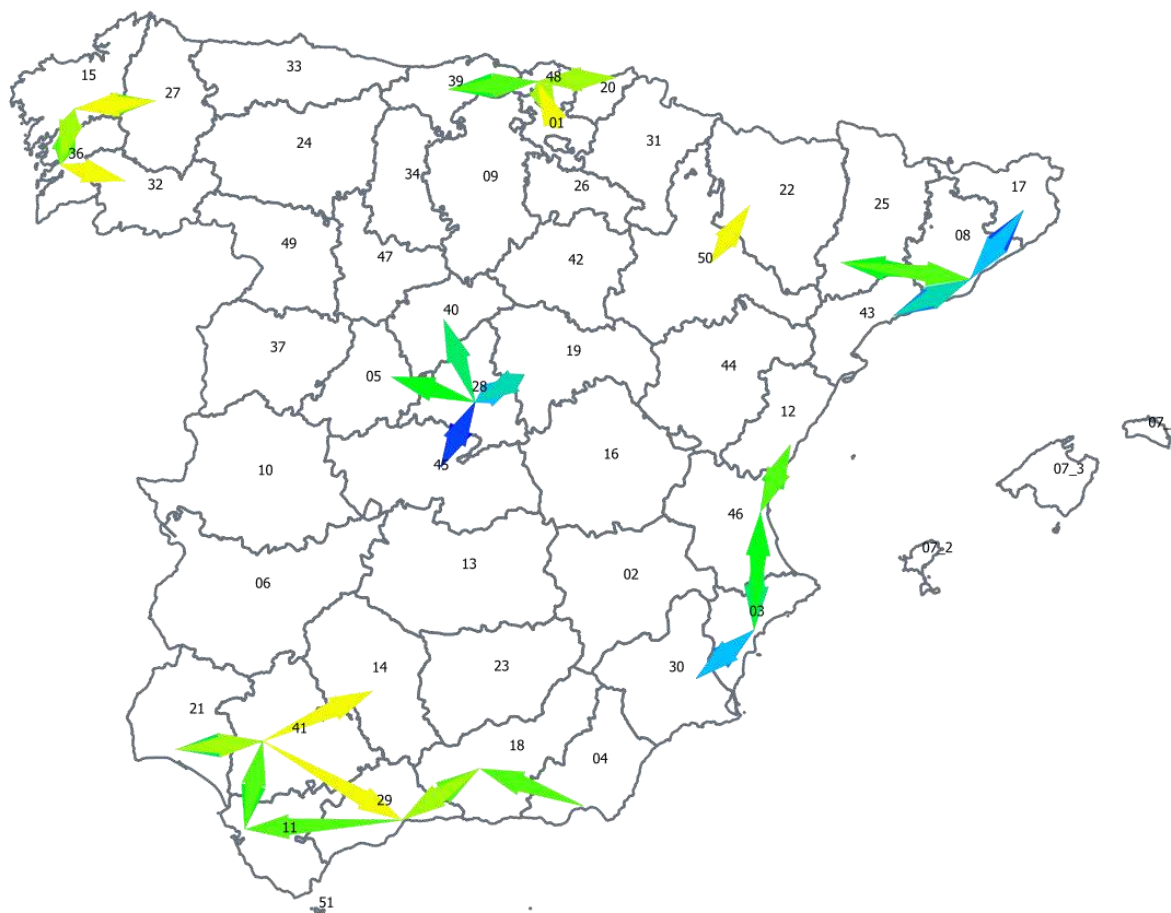
# Matriz de tours: Especificación

| Matrices de tours |             |  |  |
|-------------------|-------------|--|--|
| Columna           | Nombre      | Valores  | Descripción  |
| 1                 | Origen      | según los valores definidos en el fichero zonificación.shp<br>60: Extranjero     | Identificador de la zona de origen del viaje   |
| 2                 | Destino     | según los valores definidos en el fichero zonificación.shp<br>60: Extranjero     | Identificador de la zona de destino del viaje  |
| 3                 | Año         | 2017   | Año de estudio   |
| 4                 | Mes         | 07-08-10   | Meses de estudio   |
| 5                 | Día         | 01-31: días específicos<br>Lunes-domingo: días promedios                         | Nombre del día de la semana para los días promedio y número de día para los días sueltos   |
| 6                 | Periodo     | P1: 00:00 - 06:00<br>P2: 06:00 - 10:00<br>P3: 10:00 - 17:00<br>P4: 17:00 - 00:00 | Franja horaria correspondiente al inicio del viaje   |
| 7                 | Modo        | autobús/avión/barco/privado/tren<br>modos separados por ','<br>NA: No aplica     | Concatenación de los modos de transporte empleados en los distintos viajes del tour. En el caso de que el viaje no sea relevante, se proporciona el valor NA |
| 8                 | Residencia  | Nacionalidad según los valores de la <i>Tabla 1</i>                              | País o grupo de países de residencia de los viajeros   |
| 9                 | Viajeros    | Número decimal (3 dígitos)   | Número de viajeros   |
| 10                | Viajeros-km | Número decimal (3 dígitos)   | Número de viajeros multiplicado por los kilómetros del viaje   |



# Matriz de tours: principales flujos

- Las 40 relaciones principales se dan entre provincias limítrofes



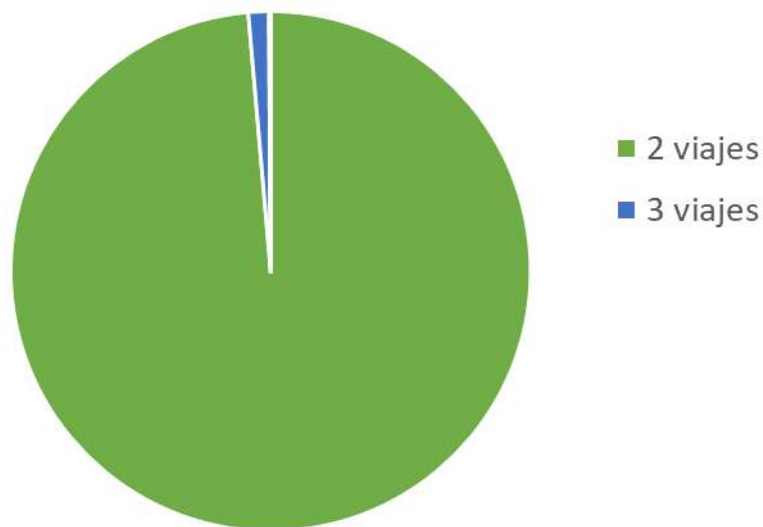
\* Resultados del martes promedio de octubre

+ nº tours



# Matriz de tours: principales flujos

- El 98% de los tours son de ida y vuelta ( 2 viajes)
- El 90% de los tours se realizan en vehículo privado



*\* Resultados del martes promedio de octubre*

# Recurrencia casa-trabajo: Especificación

| Recurrencia casa-trabajo |                 |  |   |
|--------------------------|-----------------|--|---|
| Columna                  | Nombre          | Valores  | Descripción   |
| 1                        | Año             | 2017   | Año de estudio  |
| 2                        | Mes             | 07-08-10   | Meses de estudio  |
| 3                        | Casa            | según los valores del fichero zonificación.shp   | Lugar de residencia a nivel de zona de estudio                              |
| 4                        | Trabajo/Estudio | según los valores del fichero zonificación.shp   | Lugar de trabajo/estudio a nivel de zona de estudio                         |
| 5                        | Distancia       | D0: viajes de menos de 10 km.<br>D1: viajes entre 10 y 50 km.<br>D2: viajes de más de 50 km.                                   | Rango de distancia en el que se encuentra la localización de casa y trabajo |
| 6                        | Recurrencia     | R0: 0 viajes<br>R1: 1-5 viajes/mes<br>R2: 6-10 viajes/mes<br>R3: 11-15 viaje/mes<br>R4: 16-20 viajes/mes<br>R5: >20 viajes/mes | Nivel de recurrencia  |
| 7                        | Personas        | Número decimal (3 dígitos)   | Número de personas  |

# Recurrencia casa-trabajo

40 primeras relaciones distancias >50Km. Octubre

Principales flujos



+ nº personas



Viajeros con recurrencia mayor de 16 viajes/mes

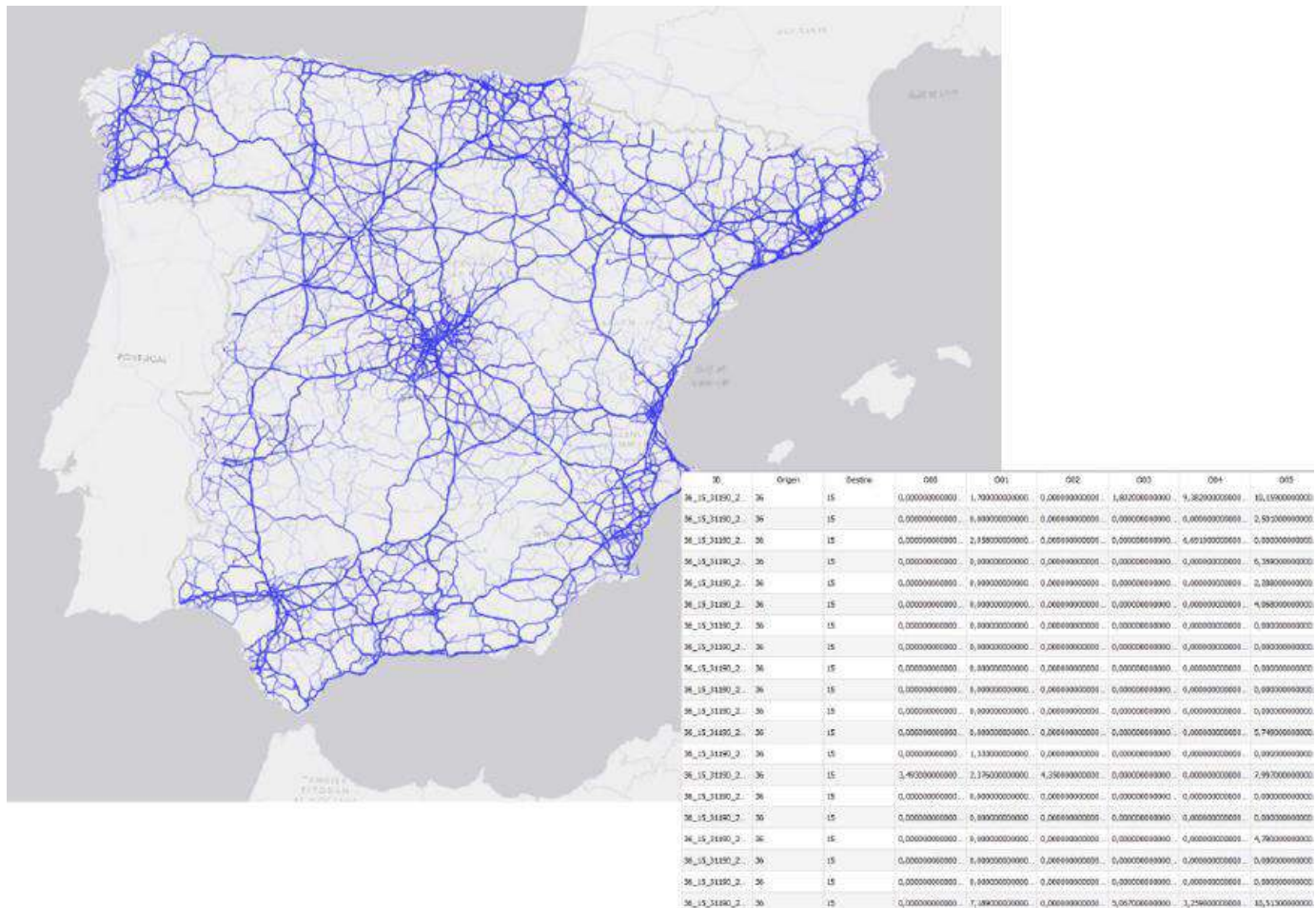


# Matriz de rutas (carretera): Especificación

| Matrices de rutas |          |   |  |
|-------------------|----------|---|--|
| Columna           | Nombre   | Valores   | Descripción  |
| 1                 | Origen   | según los valores definidos en el<br>fichero zonificación.shp<br>60: Extranjero | Identificador de la zona de origen<br>del viaje  |
| 2                 | Destino  | según los valores definidos en el<br>fichero zonificación.shp<br>60: Extranjero | Identificador de la zona de destino<br>del viaje   |
| 3                 | Año      | 2017  | Año de estudio   |
| 4                 | Mes      | 07-08-10  | Meses de estudio   |
| 5                 | Día      | 01-31: días específicos<br>Lunes-domingo: días promedios                        | Nombre del día de la semana para<br>los días promedio y número de día<br>para los días sueltos |
| 6                 | Ruta     | Código numérico asociado al<br>fichero rutas.shp                                | Identificador único de ruta  |
| 7                 | Viajeros | Número decimal (3 dígitos)  | Número de viajeros   |

# Matriz de rutas

- Análisis de 17M de rutas en territorio nacional
- Se proporciona la geometría de las rutas junto al número de viajes que hacen uso de cada ruta por mes analizado





# Matriz de rutas

Rutas posibles

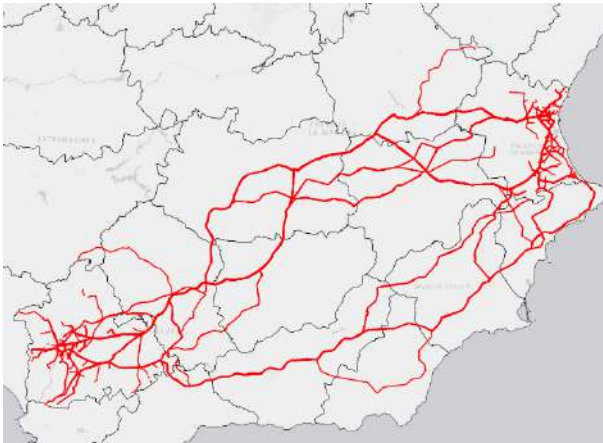
Madrid - Barcelona



Rutas según importancia



Valencia - Sevilla



# Conclusiones y Futuras líneas de trabajo



# Conclusiones

- La **información de demanda de transporte** generada a partir de los datos de telefonía móvil es de **calidad**. La metodología empleada permite identificar de manera satisfactoria el **modo de transporte** en la mayoría de los viajes
- **Valor añadido:**
  - Viajes **completos**, con distinción **de etapas**
  - volumen y estructura de los **viajes por carretera**.
  - **Tamaño de muestra** muy elevado (mejor estructura + mejor caracterización).
  - Posibilidad de analizar un **gran número de días** (demanda anual más precisa).
  - Reducción de **costes** y **plazos**.
- **Limitaciones**
  - Cualificación de la demanda
  - Resolución espacio-temporal de los datos

# Líneas de trabajo con los resultados

- Proceso de análisis de coherencia del trabajo con otras fuentes. Finalizada: publicación escalonada de resultados.
- Desarrollo de herramientas de exploración y visualización/analítica visual
- Integración con modelos de transporte a nivel nacional



# Proyecto de evaluación de la Movilidad Interprovincial de Viajeros aplicando la Tecnología Big Data

**MUCHAS GRACIAS**

