

**TRABAJO FIN DE MÁSTER UNIVERSITARIO EN INGENIERÍA  
GEOMÁTICA Y GEOINFORMACIÓN**

# **ANÁLISIS DE LA MOVILIDAD DE LOS TURISTAS URBANOS BASADO EN LA ACTIVIDAD DE LOS USUARIOS EN REDES SOCIALES**

**AUTOR: CLAVICILLAS CABALTERA, GLENE**

**TUTORA: SEBASTIÁ TARÍN, LAURA**

**Curso Académico: 2020-2021**





## RESUMEN

El sector turístico contribuye en gran medida al desarrollo de la economía de los países y por tanto es objeto de diversos estudios con el fin de mejorar la calidad de dicho sector. Los estudios de movilidad se realizan con la finalidad de analizar el comportamiento de las personas en un lugar y detectar las necesidades que estos experimentan, para proponer medidas que repercutan en la mejora de la economía de la zona. Dichos estudios se generan, actualmente, a partir del contenido que suben las personas en internet a través de las redes sociales. En el caso del sector turístico, dichos estudios ayudan a mejorar la calidad de los servicios, transportes e infraestructuras de la zona, entre otros.

El objetivo de este trabajo consiste fundamentalmente en obtener un modelo de comportamiento o de movilidad de los turistas de la ciudad de Valencia a través del contenido multimedia que suben en sus redes sociales, en concreto, las fotografías geolocalizadas que suben a la plataforma *Flickr*. El primer paso consiste en, a partir de los datos iniciales, diferenciar entre turistas y residentes utilizando, en combinación, las consultas de datos en postgres y la programación en python. A continuación, se lleva a cabo una representación gráfica de las rutas seguidas por los turistas a través del contenido geolocalizado subido a la plataforma mediante el uso de grafos para extraer modelos de movilidad. En el siguiente paso, se crean mapas de calor anuales y se comparan entre ellos para ver la evolución temporal del comportamiento de los turistas en la ciudad. Además, se crean comunidades entre los diferentes lugares visitados con el objetivo de determinar características similares. Finalmente, una vez determinados y analizados estos caminos, se extrae información acerca de los lugares visitados como, por ejemplo, el orden en el que se visitan, el tipo de ruta escogida, etc.

La metodología empleada ha permitido ver que durante el periodo comprendido entre los años 2007 y 2013 los lugares más visitados fueron La Ciudad de las Artes y las Ciencias, las zonas del centro de la ciudad como el Ayuntamiento, la Plaza de la Virgen y, la zona de la playa (Las Arenas y Malvarrosa). En cuanto a los lugares con menor frecuencia de turistas, estos fueron zonas perimetrales de Valencia, así como algunos pueblos vecinos y la zona de l'Albufera.

Los resultados obtenidos ponen de manifiesto la robustez de la metodología empleada y su potencialidad para el análisis del comportamiento de la movilidad de las personas. Esta metodología puede ser empleada en datos geolocalizados actuales para obtener patrones de comportamiento actualizados y detectar nuevas necesidades.

**Palabras clave:** Turismo urbano, SIG, Análisis espacial, Flickr, redes sociales, detección de comunidad, grafo, estudios de movilidad.



## ABSTRACT

The tourism sector contributes greatly to the development of the economy of countries and therefore is subject of several studies to improve the quality of this sector.

Mobility studies are carried out to analyze the behavior of people in a place and detect the needs they experience, to propose measures that have an impact on the improvement of the economy of the area. These studies are currently generated from the content that people upload on the internet through social networks. In the case of the tourism sector, these studies help to improve the quality of services, transport network and infrastructure in the area, among others.

The objective of this work is mainly to obtain a model of behaviour or mobility of tourists in the city of Valencia through the multimedia content that they upload on their social networks, specifically, the geolocated photographs that are uploaded to the Flickr platform.

The first step consists of, from the initial data, differentiating between tourists and residents using, in combination, the data queries in postgres and the programming in python. Next, a graphical representation of the routes followed by tourists is carried out through the geolocated content uploaded to the platform using graphs to extract mobility models. In the next step, heat maps are created and compared to each other to see the temporal evolution of the behavior of tourists in the city. In addition, communities are created between the different places visited to determine similar characteristics. Finally, once these paths have been determined and analyzed, information is extracted about the places visited, such as the order in which they are visited, the type of route chosen, etc.

The methodology used has allowed us to see that during the period between 2007 and 2013 the most visited places were the City of Arts and Sciences, the areas of the city center such as the Ayuntamiento, Plaza de la Virgen and, the beach area (Las Arenas and Malvarrosa). As for the places with the least frequency of tourists, these were the perimeter areas of Valencia, as well as some neighboring towns and the L'Albufera area.

The results show the robustness of the methodology used and its potential for analyzing the behavior of people's mobility. This methodology can be used in current geolocated data to obtain updated behavior patterns and detect new needs.

**Keywords:** urban tourism, GIS, spatial análisis, Flickr, social media, community detection, graph



## RESUM

El sector turístic contribueix en gran mesura al desenvolupament de l'economia dels països i per tant és objecte de diversos estudis per tal de millorar la qualitat d'aquest sector. Els estudis de mobilitat es realitzen amb la finalitat d'analitzar el comportament de les persones en un lloc i detectar les necessitats que aquests experimenten, per proposar mesures que repercutixin en la millora de l'economia de la zona. Aquests estudis es generen, actualment, a partir del contingut que pugen les persones a internet a través de les xarxes socials. En el cas del sector turístic, aquests estudis ajuden a millorar la qualitat dels serveis, transports i infraestructures de la zona, entre altres.

L'objectiu d'aquest treball consisteix fonamentalment en obtenir un model de comportament o de mobilitat dels turistes de la ciutat de València a través del contingut multimèdia que pugen a les seves xarxes socials, en concret, les fotografies geolocalitzades que pugen a la plataforma Flickr. El primer pas consisteix en, a partir de les dades inicials, diferenciar entre turistes i residents utilitzant, en combinació, les consultes de dades en postgres i la programació en python. A continuació, es duu a terme una representació gràfica de les rutes seguides pels turistes a través del contingut geolocalitzat pujat a la plataforma mitjançant l'ús de grafs per extreure models de mobilitat. En el següent pas, es creen mapes de calor anuals i es comparen entre ells per veure l'evolució temporal del comportament dels turistes a la ciutat. A més, es creen comunitats entre els diferents llocs visitats amb l'objectiu de determinar característiques similars. Finalment, un cop determinats i analitzats aquests camins, s'extrau informació sobre els llocs visitats, com ara, l'ordre en què es visiten, el tipus de ruta escollida, etc.

La metodologia emprada ha permès veure que durant el període comprès entre els anys 2007 i 2013 els llocs més visitats van ser La Ciutat de les Arts i les Ciències, les zones de centre de la ciutat com l'Ajuntament, la Plaça de la Mare de Déu i, la zona de la platja (Les Arenes i la Malva-rosa). Pel que fa als llocs amb menor freqüència de turistes, aquests van ser zones perimetrals de València, així com alguns pobles veïns i la zona de l'Albufera.

Els resultats obtinguts posen de manifest la robustesa de la metodologia emprada i la seva potencialitat per a l'anàlisi del comportament de la mobilitat de les persones. Aquesta metodologia pot ser emprada en dades geolocalitzats actuals per obtenir patrons de comportament actualitzats i detectar noves necessitats.

**Paraules claus:** Turisme urbà, SIG, Anàlisi espacial, Flickr, xarxes socials, detecció de comunitat, graf



## INDICE DE FIGURAS DE LA MEMORIA

Figura 1. Metodología empleada para la realización de un estudio de movilidad turística mediante tweets geolocalizados. Fei Hu et al. (2018) .....	4
Figura 2. Grafo no dirigido. ....	7
Figura 3. Grafo dirigido.....	7
Figura 4. Representación de ruta mediante NetworkX, librería de Python.....	9
Figura 5. Representación de ruta con nodos en común mediante NetworkX, librería de Python. ....	9
Figura 6. Representación de ruta aislada mediante NetworkX, librería de Python.....	10
Figura 7. Lugar de búsqueda de los diferentes data sets disponibles. Webscope Yahoo! .....	13
Figura 8. Esquema de la base de datos Template_postgis. ....	14
Figura 9. Tabla Grid_Valencia.....	15
Figura 10. Tabla Multimedia_city.....	15
Figura 11. Tabla photo_grid_valencia .....	15
Figura 12. Metodología del estudio de movilidad propuesto.....	16
Figura 13. Fichero de trabajo. Datos de partida para el estudio de movilidad.....	17
Figura 14. Resultado de consulta. Número de fotografías por año. ....	18
Figura 15. Resultado de consulta. Número de usuarios por año y mes.....	19
Figura 16 Mapa de calor. 2007.....	22
Figura 17. Mapa de calor. 2008.....	22
Figura 18. Mapa de calor. 2009.....	23
Figura 19. Mapa de calor. 2010.....	23
Figura 2011.....	23
Figura 21. Mapa de calor. 2012.....	23
Figura 22. Mapa de calor. 2013.....	24
Figura 23. Grafo obtenido para el mes de julio de 2013.....	25
Figura 24. Detección de comunidades. Mes de enero 2013.....	26
Figura 25. Mapa de Valencia. Desplazamientos del mes de enero 2013.....	28
Figura 26. Grafo mes de febrero 2013. ....	29
Figura 27. Mapa de distribución de comunidades durante el mes de febrero 2013.....	30
Figura 28. Resultado de detección de comunidades del mes de marzo de 2013.....	30
Figura 29. Mapa de distribución de comunidades durante el mes de marzo 2013.....	31
Figura 30. Mapa de distribución de comunidades durante el mes de abril 2013.....	33



Figura 31. Grafo mes de mayo. ....	33
Figura 32. Mapa de distribución de comunidades durante el mes de mayo 2013.....	34
Figura 33. Grafo mes de mayo. ....	35
Figura 34. Mapa de distribución de comunidades durante el mes de junio 2013.....	35
Figura 35. Mapa de distribución de comunidades durante el mes de Julio 2013 .....	36
Figura 36. Mapa de distribución de comunidades durante el mes de agosto 2013.....	38
Figura 37. Mapa de distribución de comunidades durante el mes de septiembre 2013 .....	39
Figura 38. Mapa de distribución de comunidades durante el mes de septiembre, ampliado, 2013 .....	39
Figura 39. Mapa de distribución de comunidades durante el mes de octubre 2013.....	40
Figura 40. Mapa de distribución de comunidades durante el mes de diciembre 2013 .....	42



## INDICE DE TABLAS DE LA MEMORIA

Tabla 1. Estadístico de centralidad para el mes de enero de 2013.....	26
Tabla 2. Estadístico de centralidad para el mes de febrero de 2013.....	29
Tabla 3. Estadístico de centralidad para el mes de marzo de 2013.....	31
Tabla 4. Estadístico de centralidad para el mes de abril 2013.....	32
Tabla 5. Estadístico de centralidad para el mes de mayo 2013.....	33
Tabla 6. Estadístico de centralidad para el mes de Julio 2013.....	35
Tabla 7. Estadístico de centralidad para el mes de agosto 2013.....	37
Tabla 8. Estadístico de centralidad para el mes de septiembre de 2013.....	38
Tabla 9. Estadístico de centralidad para el mes de octubre de 2013.....	40
Tabla 10. Estadístico de centralidad para el mes de diciembre de 2013.....	41
Tabla 11. Duración en horas de cada tarea realizada.....	43
Tabla 12. Coste del personal.....	43



## INDICE DE LA MEMORIA

1. Introducción .....	1
1.1. Objetivos .....	2
1.2. Motivación .....	3
1.3. Estado del arte. ....	3
2. Conceptos .....	6
2.1. Teoría de grafos.....	6
2.1.1. Python y teoría de grafos.....	8
2.2. Detección de comunidades.....	11
2.2.1. Python y detección de comunidades .....	12
3. Datos .....	13
4. Metodología .....	16
4.1. Obtención de datos de trabajo. ....	17
4.2. Filtrado. ....	18
4.3. Enfoque. ....	19
4.4. Representación de resultados. ....	20
5. Resultados .....	22
5.1. Análisis anual periodo 2007-2013. Mapas de calor .....	22
5.2. Análisis mensual año 2013.....	24
6. Presupuesto. ....	43
6.1. Costes de personal.....	43
7. Conclusiones .....	44
8. Bibliografía .....	45
9. Anexos.....	46
Grafos año 2013 .....	46
10. Cartografía.....	53





## 1. Introducción

El turismo es, a nivel mundial, un sector en continua expansión que forma parte de los principales sectores económicos. En el caso de la economía española, en el año 2019, antes de la pandemia por la COVID-19, el turismo representaba el 12.4% del PIB estando por encima de otros sectores como la construcción (14%), el comercio (12%) o la sanidad (12%).

Por esta razón, este sector es objeto de diversos estudios de movilidad, los cuales tienen como finalidad observar el comportamiento de los turistas en una ciudad, generar patrones de movimientos y detectar así, las necesidades de estos durante su estadía para, entre otras cosas, reforzar las conexiones de transporte existentes entre los diversos puntos de interés de una ciudad, llevar a cabo el desarrollo de productos turísticos, ofrecer nuevos servicios, diseñar nuevas infraestructuras, mejorar los servicios de hostelería, etc. Todas estas acciones influyen positivamente en la economía local de la zona.

Las redes sociales son plataformas que permiten crear, almacenar y compartir múltiples contenidos en línea a través de Internet como fotografías, vídeos, etc. e intercambiar información con otros usuarios lo que genera datos masivos diarios lo que se conoce como *Big Data*. Hoy en día, millones de usuarios generan gran cantidad de datos geolocalizados a partir del contenido que suben en Internet.

Una forma de llevar a cabo los estudios de los patrones de movilidad de los turistas es mediante el análisis de la ruta realizada por estos, la cual es obtenida a partir del contenido geolocalizado que suben a sus redes sociales. A partir de estos datos masivos, se puede extraer información acerca de los lugares visitados de una ciudad y el orden en el que se lleva a cabo, lo que puede ayudar en la toma de decisiones para la mejora de los servicios, transporte, etc. Este método sustituye al método tradicional de investigación turística, técnica laboriosa en la cual se extrae información a partir de encuestas o entrevistas de movilidad y que debido a la limitación en el conjunto de datos disponibles acerca del comportamiento de los turistas, puede llegar a ser complicado en la perspectiva de planificación o gestión de rutas o destino turísticos, resultando no ser una solución eficiente.

Dentro de las redes sociales, destaca *Flickr*, una plataforma que permite almacenar y compartir fotografías o vídeos en línea a través de Internet. Su primer lanzamiento fue el



10 de febrero de 2004 y actualmente cuenta con millones de cuentas. Los usuarios comparten sus fotografías y la plataforma permite a éstos administrarlas, etiquetarlas y comentar fotografías de otros usuarios.

La gran cantidad de datos que la plataforma puede almacenar es el objetivo de diversos estudios, en este caso, para realizar análisis de comportamiento de movilidad de los usuarios en una ciudad a partir de las fotografías y videos que suben.

Esta información se puede tratar utilizando lenguajes de programación enfocados al tratamiento de datos masivos. En este caso, *Python* es un lenguaje de programación de código abierto, orientado a objetos, muy fácil de dominar, pensado para trabajar con un volumen de datos enorme. Cuenta con muchas librerías y funciones que permiten realizar análisis de manera eficaz. En el apartado de conceptos teóricos se hará una breve explicación sobre los conceptos en los que se basan las librerías de Python utilizadas en este trabajo que han permitido obtener información sobre los lugares mas visitados y las rutas seguidas a partir de los datos proporcionado por Flickr.

## 1.1. Objetivos

El objetivo principal del trabajo consiste en realizar un estudio de movilidad de los usuarios a través del contenido geolocalizado que suben a sus redes sociales y a partir de los resultados de dicho estudio, determinar qué lugares reciben más visitas, qué rutas realizan los usuarios para llegar a dichos lugares, etc. Adicionalmente se analizarán los patrones de movilidad obtenidos con la finalidad de detectar las necesidades que experimentan los turistas en los lugares que visitan.

Para la realización de este trabajo, se hará uso del contenido fotográfico y multimedia de de la plataforma *Flickr* sobre la ciudad de Valencia disponible en un host del Departamento de Sistemas de Información y Computación de la Universitat Politècnica de València desde los años 2007 a 2013.

Para la consecución de los objetivos principales se han propuesto los siguientes objetivos específicos:

- Estudiar y analizar el conjunto de datos geolocalizados de *Flickr*, el cual se encuentra disponible en una base de datos del DSIC.



- Filtrar para eliminar datos no representativos mediante el uso de un gestor de bases de datos.
- Decidir el enfoque temporal a seguir (estudio mensual, estudio anual) y distinguir entre usuarios turistas y usuarios residentes mediante el uso de un lenguaje de programación, en concreto Python.
- Determinar las rutas que siguen los usuarios turistas a lo largo del día, mes y año mediante la aplicación de la teoría de grafos y detección de comunidades.
- Visualizar dichos resultados a través de grafos y mapas de calor.
- Analizar y detectar las posibles causas de la existencia de los lugares de mayor y menor frecuencia de visita detectados.

## 1.2. Motivación

La principal motivación para llevar a cabo el trabajo final de máster se debe al interés acerca de del uso de la programación enfocada al tratamiento y análisis de datos. Aprender a programar de manera correcta y fluida requiere práctica y constancia. En el comienzo de la carrera se empieza a programar partiendo de un nivel básico y, en la etapa final de la carrera, se pueden adquirir sólidos conocimientos tanto teóricos como prácticos que permiten llevar a cabo cualquier trabajo aplicado al campo de la geomática. Con este trabajo se pretende demostrar el potencial que tiene la programación para la realización de análisis de datos masivos. Para mí, resulta gratificante aprender nuevos aspectos que se desarrollarán en este trabajo y serán de gran utilidad en mi futuro profesional.

## 1.3. Estado del arte.

Los métodos actuales para la realización de estudios de movilidad están enfocados en el uso del Big Data como fuente de información. Existen muchas formas de aprovechar la huella digital que dejan los usuarios en las redes sociales. Por ejemplo, Fei Hu et al. (2018) llevan a cabo un estudio de movilidad en la ciudad de Nueva York (NYC) con el objetivo de analizar zonas donde existe mayor presencia de turistas, determinar los patrones de comportamiento de éstos y trazar nuevas rutas. Dicho estudio comprende el periodo entre el 1 de julio de 2016 y el 30 de abril de 2017. Los autores extraen sus datos de partida desde la plataforma de Twitter gracias a los tweets geolocalizados que los usuarios publican.

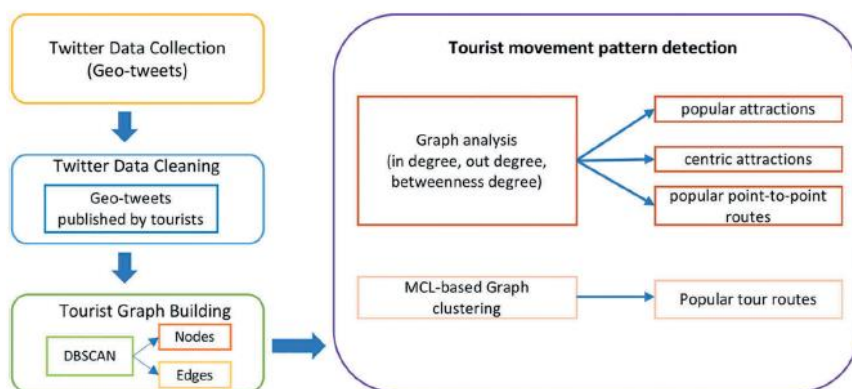


Figura 1. Metodología empleada para la realización de un estudio de movilidad turística mediante tweets geolocalizados. Fei Hu et al. (2018)

La metodología empleada, la cual se puede observar en la *figura 1*, consiste en:

- Filtrado de datos. Establecer condiciones que permitan diferenciar entre un residente y un turista.
- Realización de un análisis espacio temporal basado en el uso de un algoritmo de agrupamiento espacial (DBSCAN) para identificar atracciones turísticas y representar dichos resultados mediante grafos.
- Utilización de un método analítico de red para aplicarlo al grafo obtenido y detectar patrones de movimiento, las rutas más concurridas y las principales atracciones turísticas.
- Construcción de un gráfico de probabilidad para la detección de rutas más populares utilizando el algoritmo *Markov Clustering* y otros estadísticos como *betweenness centrality* para saber la importancia de las atracciones de un lugar, las cuales se consideran nodos.

Los resultados obtenidos se clasifican en atracciones más populares, atracciones céntricas, es decir, los lugares que los turistas prefieren visitar cuando se mueven en una atracción a otra y, rutas punto a punto más populares. Dichos resultados se comparan con las rutas realizadas por los buses turísticos de la ciudad para descubrir y proponer nuevas rutas.

Las redes sociales son un mundo por explorar y se ha vuelto el punto de mira para muchos investigadores. Gracias al contenido geolocalizado subido a dichas plataformas se pueden llevar a cabo estudios de movilidad con diferentes criterios. Otro procedimiento similar



es el llevado a cabo por Oriol Romero et. al (2018) donde extraen la información a partir de Yelp, una plataforma donde los usuarios pueden dejar comentarios, valoraciones, reseñas etc. sobre los lugares que visitan y, adoptan otro enfoque para extraer patrones de comportamiento. El estudio de movilidad realizado pretende identificar el comportamiento de turistas y residentes de diferentes núcleos urbanos de EEUU, Canadá y UK y, la posible presencia de una relación entre el comportamiento de ambos grupos. Como elemento diferenciador hacen uso de un sistema de información geográfica, GIS, para comprobar que las reseñas indicadas sobre un lugar o negocio se realizan desde el mismo.

Se puede observar a partir de los dos artículos mencionados anteriormente que existen diferentes metodologías o enfoques para la realización de estudios movilidad y que los resultados obtenidos dependen en gran medida de la plataforma elegida y de los criterios de cada investigador para el tratamiento de sus datos de partida.



## 2. Conceptos

En este apartado se explicarán los conceptos teóricos que han permitido alcanzar los objetivos planteados, en concreto para determinar patrones de comportamiento a partir de los datos geolocalizados.

### 2.1. Teoría de grafos

La teoría de grafos es una rama de las matemáticas discretas y las ciencias de computación que estudia las relaciones entre objetos o entidades. Tiene múltiples aplicaciones en diferentes disciplinas como la química, investigación, sociología, informática, etc.

En el caso del presente trabajo, los grafos se utilizan para estudiar el comportamiento de los turistas en función de las actividades que desarrollan en las redes sociales. El grafo ofrece un mecanismo para determinar flujos o rutas a partir de los datos los cuales se encuentran organizados por usuario y fecha. Su implementación en este trabajo es importante ya que ayuda a resolver problemas de manera eficiente y proporciona una alternativa para la determinación de patrones de comportamiento.

Un grafo es un conjunto de objetos llamados nodos o vértices y una secuencia de vértices llamados aristas. Dicho grafo permite estudiar las relaciones que pueden existir entre los elementos que interactúan en una red.

De manera formal, se puede decir que un grafo  $G$  es un par de conjuntos  $G = (V, E)$  donde  $V$  es el conjunto de vértices y  $E$  es el conjunto de aristas que conectan los vértices. Existen diversos tipos de grafos, pero este trabajo se centra exclusivamente en grafos dirigidos y no dirigidos.

#### Grafo no dirigido

Un grafo no dirigido, *figura 2*, está formado por un conjunto de vértices conectados a las aristas sin especificar su dirección. Las aristas se representan con una línea simple y se pueden recorrer desde cualquier dirección y cualquier vértice.

### Grafo no dirigido

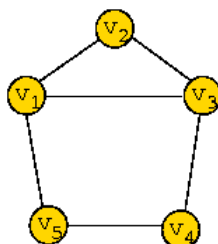


Figura 2. Grafo no dirigido.

Formalmente, se dice que un grafo no dirigido,  $G$ , es un par de conjuntos  $G = (V, E)$ , donde  $V$  es el conjunto de vértices y  $E$  es el conjunto de aristas (*edges*) las cuales se forman por los pares de elementos desordenados de  $V$ . Un ejemplo de este tipo de grafos sería un grafo de Facebook donde un usuario envía una petición de amistad a otro y estos se vuelven amigos sin que el otro usuario envíe otra petición.

### Grafo dirigido

Un grafo dirigido conocido como *DiGraph* está formado por un conjunto de vértices y aristas donde cada una se encuentra representada con una flecha que indica la dirección del flujo, *figura 3*.

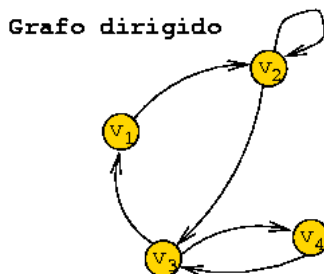


Figura 3. Grafo dirigido

Normalmente se utiliza el término *arcos* para indicar la dirección de una arista. Un dígrafo también es un par de conjuntos  $D = (V, A)$  donde  $V$  es el conjunto de vértices y  $A$  es el conjunto de aristas que conectan un conjunto de nodos. A diferencia del grafo no dirigido, este grafo hace distinción entre los nodos de partida y salida. Así pues, no es lo mismo decir  $(u, v)$  que  $(v, u)$ . Normalmente, en el grafo dirigido las aristas se suelen denominar arcos para indicar la orientación de ellas. Un ejemplo de este tipo sería un grafo de *Twitter* o *Instagram* donde un usuario sigue a otro y este no necesariamente tiene que seguirlo.



### 2.1.1. Python y teoría de grafos.

#### NetworkX. Representación de grafos.

Esta librería de Python sirve para crear grafos, así como para manipularlos y proporciona algoritmos para realizar análisis complejos. Dependiendo del tipo de grafo se utilizan las siguientes sentencias para definirlos:

```
G=networkx.Graph(),
```

para el grafo no dirigido y

```
G=networkx.DiGraph()
```

para el grafo dirigido el cual tiene además la opción *self-loops* que permite que una arista pueda conectar al mismo nodo.

Entre las clases de *Networkx* se encuentran:

- *add\_node* y *add\_nodes\_from*. El primero añade un nodo singular al grafo y el segundo se utiliza para insertar múltiples nodos.
- *add\_edge* y *add\_edges\_from*. El primero se utiliza para añadir una arista simple especificando los nodos U y V a los que conectar, mientras que el segundo se puede utilizar para añadir varias a la vez. En este caso se puede utilizar una lista que contiene un par de nodos para conectar.

Suponiendo que se usan los datos de *Flickr*, se debe tener en cuenta que están ordenados por usuario y fecha. De esta manera se puede crear la ruta que sigue cada usuario. Los nodos corresponden a cada uno de los lugares visitados, o *grids*, por los usuarios durante un determinado mes. Por lo general, los nodos no se duplican una vez se han creado. Para cada usuario, se crea un enlace entre nodos para determinar el desplazamiento que realiza. Si dos diferentes usuarios coinciden en algún nodo, entonces se crea una conexión entre ellos. A continuación, se explica un ejemplo de cómo se crean los nodos y enlaces y cómo se conectan entre sí.

- El usuario 1 sigue el camino  $1 \rightarrow 2 \rightarrow 3$ . Al tratarse de un grafo dirigido se crea mediante la opción correspondiente. Para añadir los enlaces entre *grids* se utiliza la función *add\_edges()*. El resultado de estas sentencias se puede ver en la *figura 4*.



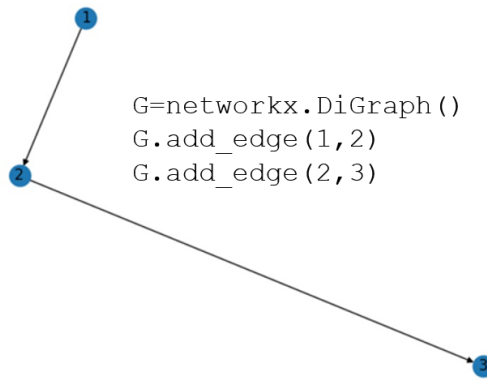


Figura 4. Representación de ruta mediante NetworkX, librería de Python.

- El usuario 2 sigue la ruta  $1 \rightarrow 2 \rightarrow 5$ . Por tanto, se añade una nueva arista que conecta en el grafo los vértices o grids 1 - 2 y 2 - 5. El nodo 2 ya se ha creado puesto que el usuario 1 ya ha pasado por ese nodo y, por tanto, no se vuelve a crear, figura 5.

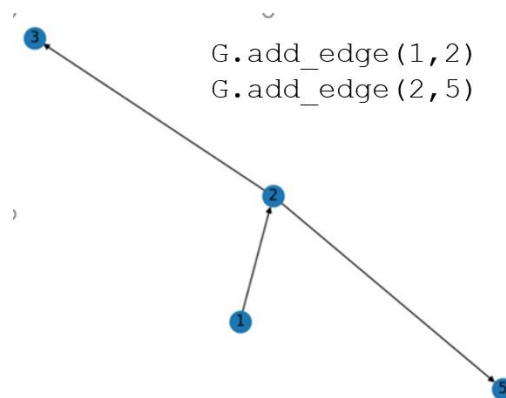


Figura 5. Representación de ruta con nodos en común mediante NetworkX, librería de Python.

- El usuario 3 sigue otra ruta diferente a los anteriores  $4 \rightarrow 6$ . Dado que son nodos distintos a los ya creados y no conectan con ninguno de ellos se crea una nueva ruta sin conectarse a la ya existente, figura 6.

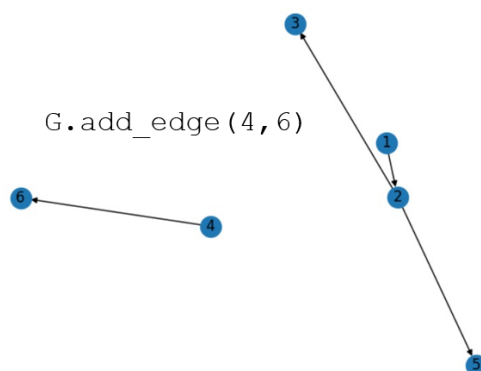


Figura 6. Representación de ruta aislada mediante NetworkX, librería de Python.

### NetworkX. Estadísticos.

Existen muchos conceptos interesantes que se pueden utilizar para realizar algunos análisis sobre los grafos. Srivatsa, (2018) menciona algunos de los algoritmos, los cuales se incluyen en la librería NetworkX y que pueden ser de utilidad para extraer estadísticas a partir de grafos.

#### - Average shortest path length

El promedio de la longitud de la ruta más corta es la suma de las longitudes de la ruta  $D(u, v)$  entre todos los pares de nodos normalizada por  $n*(n-1)$  donde  $n$  es el número total de nodos en el grafo  $G$ . Representa el camino más corto entre dos nodos y es una medida que ayuda a entender la eficiencia del flujo de la red.

#### - Breadth first search (BFS) y Depth first search (DFS)

El primer algoritmo determina la accesibilidad a un nodo desde uno dado de la red. A partir de un nodo dado el algoritmo comienza a buscar conexiones entre otros nodos siguiendo el criterio de que tiene que encontrar los nodos más cercanos posibles del nodo raíz.

En el segundo algoritmo se indica también el nodo raíz donde comienza la búsqueda, pero a diferencia del primero este algoritmo busca encontrar los nodos que están más alejados del nodo raíz hasta llegar al último nodo sin conexión con otros ya visitados.



## - **Centrality**

La centralidad tiene como objetivo determinar la importancia de los nodos de la red. Saber esta información da una idea de las atracciones turísticas más visitados por turistas y por tanto las más importante. Existen varios tipos de algoritmo de centralidad y los más comunes son:

- *Degree centrality*

Asigna un peso a cada nodo en función de los enlaces que tiene entre otros nodos. Este peso se obtiene en función de las aristas conectadas a un nodo específico.

- *Closeness centrality*

Asigna un peso a cada nodo en función de su proximidad a todos los demás nodos de la red.

- *Betweenness centrality*

Asigna un peso a un nodo según la cantidad de veces que ese nodo se encuentra en el camino más corto entre otros nodos.

- *Edge\_Betweenness centrality*

Al igual que los otros algoritmos, asigna un peso a un nodo en función de los enlaces que tiene con otros nodos. Además de asignar un valor, tiene en cuenta el número de enlaces entre conexiones de la red.

## 2.2. Detección de comunidades

La detección de comunidades es una herramienta útil para el análisis de redes porque permite agrupar nodos de la red que compartan alguna característica o propiedad similar, una información que a simple vista es difícil de desvelar. Con la ayuda de los algoritmos se pueden determinar propiedades comunes de los nodos de la red y agruparlos en comunidades de modo que los nodos de una comunidad están densamente conectados y débilmente conectados con otras comunidades (Thamindu Dilshan Jayawickrama [s. f]).

La detección de comunidad se puede dividir en dos grupos:



## Método aglomerativo

En este método las aristas se añaden una por una al grafo que solo contiene nodos. Se trata de un proceso iterativo donde las aristas se añaden desde la más fuerte hasta la más débil existente siguiendo el criterio de la similitud, es decir, los nodos se unen a otros formando una comunidad si comparten algún aspecto similar. Finalmente, se obtienen comunidades donde los nodos de una comunidad tienen una fuerte relación con los otros nodos de la misma comunidad y una conexión débil con los nodos de otras comunidades.

## Método divisivo

Este método realiza un proceso opuesto al método aglomerativo, las aristas se encuentran conectadas en el grafo al principio y se van eliminando una por una. También, se trata de un proceso iterativo en el cual se dividen las aristas con algún enlace o conexión con otra y que estas comparten poca similitud.

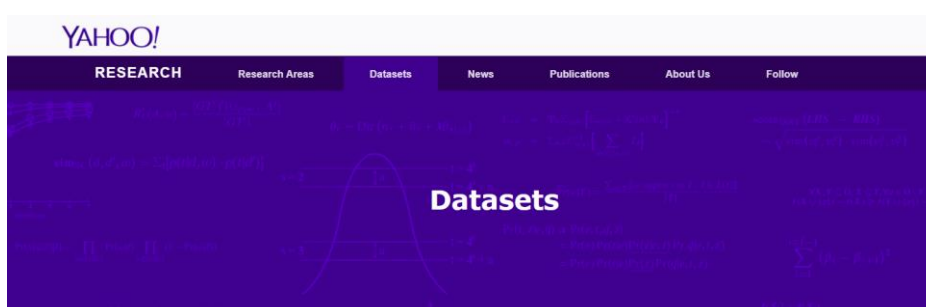
### 2.2.1. Python y detección de comunidades

Para la detección de comunidades, Python hace uso de la librería Community, la cual contiene la función Louvain, un algoritmo de detección de comunidades aglomerativo basado en el método de Louvain, creado por Blondel et al., el cual permite extraer comunidades de redes grandes mediante la optimización de la modularidad, un número entre -1 y 1 que compara la densidad de aristas dentro y fuera de una comunidad. Teóricamente, optimizando este valor iteración a iteración se obtiene la mejor posible agrupación de los nodos de una red. Cuando se obtiene una comunidad se está indicando que los nodos que la conforman comparten una característica. Esto aplicado a este trabajo quiere decir que una comunidad incluye el conjunto de nodos o *grids* por los cuales varios usuarios circulan.

Como se ha indicado anteriormente el uso de los grafos y de los algoritmos de detección de comunidades es interesante dado que se puede determinar la ruta que sigue un usuario en la red y plasmar dicha información en un Sistema de Información Geográfica, SIG, para tener una mejor interpretación de los resultados.

### 3. Datos

Para la realización de este análisis de movilidad urbano se ha hecho uso de los datos disponibles en la plataforma *Flickr*. Esta red social, lanzada en 2004, permite almacenar, etiquetar, buscar, geolocalizar, vender y compartir fotografías o videos en línea, a través de internet entre la comunidad de usuarios registrada. En 2014, la empresa liberaliza un *dataset* de aproximadamente 12 GB el cual contiene información como *photo\_id*, *url* de las fotos y videos almacenados, títulos, tipo de cámara, etiquetas, etc. Además de este *data set* también se encuentran millones de fotografías geolocalizadas, datos de utilidad en los estudios de movilidad. Toda esta información tiene una licencia *Creative Commons* por lo que es de dominio público y se puede copiar, modificar o distribuir sin la necesidad de atribuir al propietario. Se pueden obtener de manera gratuita utilizando la API de *Flickr* o bien accediendo a la nube de Amazon siempre y cuando se inicie sesión previamente en *yahoo!*, *figura 7*.



*Figura 7. Lugar de búsqueda de los diferentes data sets disponibles. Webscope Yahoo!*

Toda esta información ha sido obtenida por el Departamento de Sistemas Informáticos y Computación y aloja en un servidor de la universidad Politècnica de València, en concreto, en el host llamado *gtirouter.dsic.upv.es*.

Este *host* tiene instalado un sistema contenedor y gestor de bases de datos relacional y orientado a objetos, *PostgreSQL*. Dicho gestor posee una licencia de código abierto, multiplataforma y es mantenido por una comunidad de desarrolladores, colaboradores y organizaciones comerciales llamada PDGD (PostgreSQL Global Development Group). Es un potente sistema de base de datos relacional con más de 30 años de desarrollo activo el cual ha ganado reputación por su solidez y rendimiento. Adicionalmente, se encuentra instalado *PostGis*, una extensión de *PostgreSQL* que añade soporte para gestionar los datos geográficos de manera que convierte el sistema

PostgreSQL en una base de datos espacial para su uso mediante un sistema de información geográfica añadiendo nuevas características como por ejemplo, nuevos tipos de datos geométricos (puntos, líneas, polígonos, multipuntos, multilíneas, multipolígonos, y *geometrycollection*), soporte de datos tanto vectoriales como tipo ráster, índices espaciales, funciones espaciales, etc.

Dentro de dicho gestor existen varias bases de datos, las cuales están organizadas en diferentes esquemas (*schema*) para permitir mejor su administración y evitar conflictos con otras bases existentes. La base de datos en la que se encuentra la información necesaria para llevar a cabo este estudio se llama *template-postgis* (figura 8). Dentro de esta base de datos se encuentra el *schema multimedia\_tourism\_dbgis* y, dentro de este *schema* se encuentran las tablas que contienen información de interés.

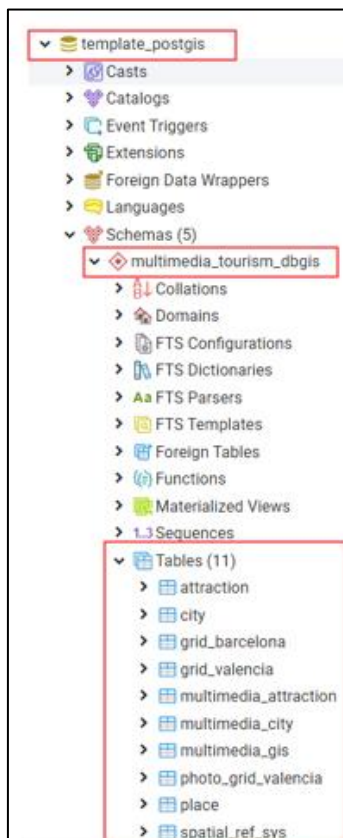


Figura 8. Esquema de la base de datos *Template\_postgis*.

A continuación, se explicará brevemente el contenido de las tablas que han sido utilizadas en este proceso.

## - Grid\_Valencia

Almacena información de una malla de 112m2 que cubre la totalidad de Valencia. La malla se creó desde QGIS utilizando la herramienta “Crear cuadrícula” que se encuentra en las herramientas de investigación del programa. Luego, ha sido importado a postgres y se ha asignado WGS84 como sistema de referencia (SRID 4326). Esta tabla se crea con la finalidad de agrupar las fotografías en mallas de forma que facilita la interpretación de estas. En ella se encuentran los campos (figura 9):

grid_valencia
Columns (7)
gid
id
x_min
x_max
y_min
y_max
geom

Figura 9. Tabla Grid\_Valencia

- *gid*: un campo autoincremento y único
- *id*: dato de tipo entero. Indica el nombre de cada malla en forma de números.
- *xmin, xmax, ymin, ymax*: indican las coordenadas de las esquinas de cada malla
- *geom*: almacena la geometría de la malla.

## - Multimedia\_city

multimedia_identifier
user_nsid
user_nickname
date_taken
date_uploaded
capture_device
title
description
user_tags
machine_tags
longitude
latitude

Figura 10. Tabla  
Multimedia\_city

Contiene información relacionada con las fotografías relativa al identificador de usuario, identificador de la fotografía, la fecha de toma, entre otros. De esta tabla se utilizan los campos (figura10):

- *multimedia\_identifier*: identificador de cada fotografía.
- *user\_nsid*: identificador de cada usuario.
- *date\_taken*: fecha en la que se tomó la fotografía.

A partir de estas dos tablas se crea, mediante intersección espacial, la tabla **photo\_grid\_valencia** la cual contiene la información necesaria para comenzar el análisis. De esta tabla contiene los campos *user\_nsid, grid\_id* y *date\_taken* (figura 11).

photo_grid_valencia
Columns (4)
grid_id
multimedia_identifier
user_nsid
date_taken

Figura 11. Tabla photo\_grid\_valencia

## 4. Metodología

En este apartado se explicará el proceso realizado para la consecución del objetivo principal de este trabajo. Para una mejor comprensión de éste, se explicará en cada paso los programas utilizados. A continuación, en la *figura 12* se presenta un esquema de la metodología llevada a cabo.

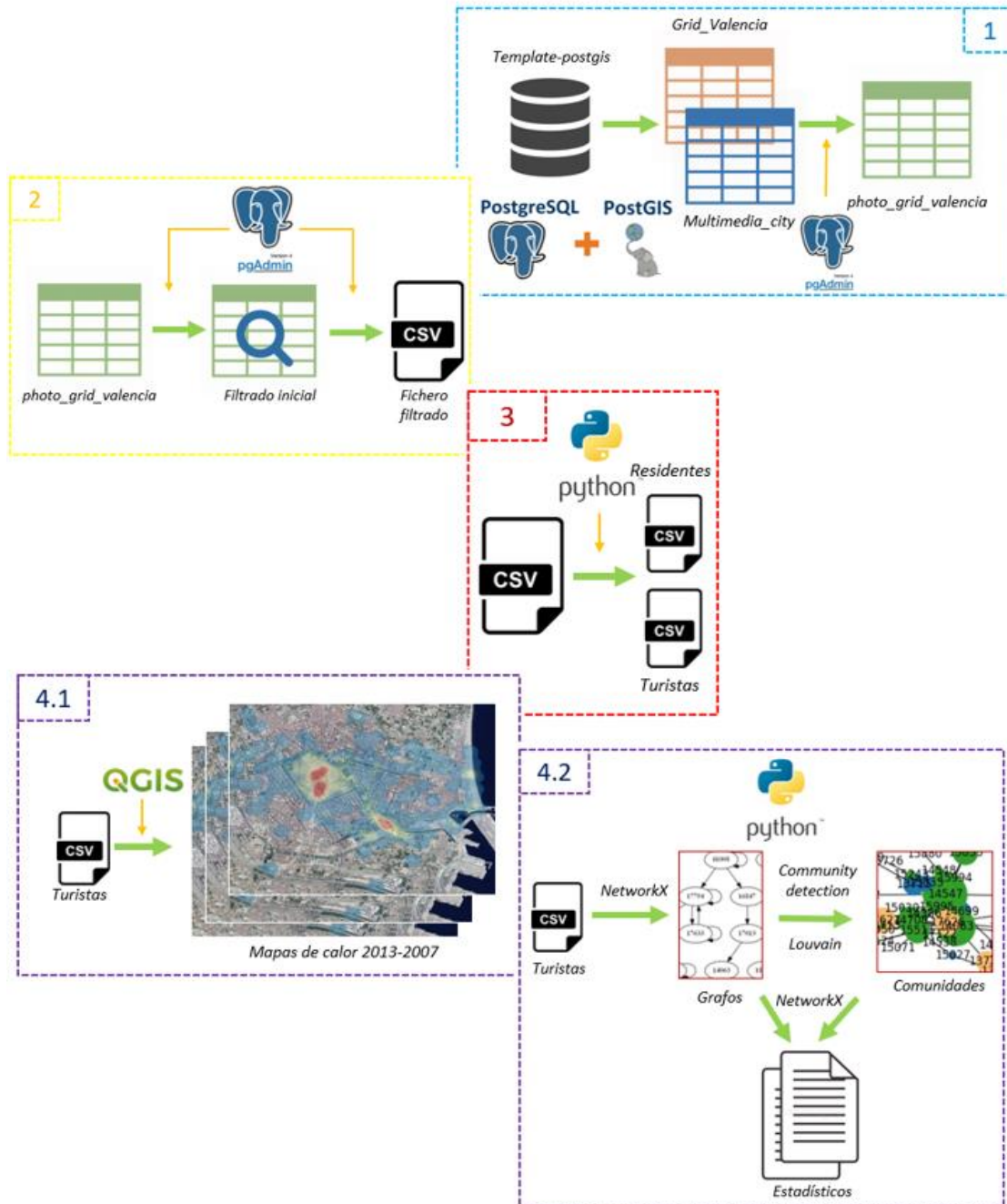


Figura 12. Metodología del estudio de movilidad propuesto.



## 4.1. Obtención de datos de trabajo.

El primer paso de este trabajo ha consistido en obtener el fichero de trabajo con las variables de interés necesarias para la realización del estudio desde la base de datos *template-postgis* en la que se encuentra disponible la información ofrecida por Flickr. Como se ha indicado anteriormente, las variables de interés escogidas a partir de *Photo\_grid\_Valencia*, intersección entre las tablas *grid\_Valencia* y *multimedia\_city*, han sido las siguientes:

- *Grid\_id*: Campo de tipo entero que indica el número de la malla.
- *User\_nsid*: Campo de tipo *string* que almacena el usuario de *Flickr*
- *Date\_taken*: Campo de tipo *timestamp* que almacena la fecha del día en el que la fotografía fue realizada, así como la hora de ese día.

Estas variables han sido ordenadas por *user\_nsid* y *date\_taken* con el propósito de poder seguir la ruta que hace un usuario. En la *figura 13* se puede observar el fichero de trabajo con el que se pretende realizar el estudio de movilidad. Gracias a que las entradas están ordenadas por usuario y tiempo de realización de la fotografía, se facilita la comprensión de los datos. Por ejemplo, las tres primeras entradas corresponden a un mismo usuario y se puede observar cómo ha realizado tres fotografías en diferentes días. Dichas entradas tienen un valor de *grid* distinto, el cual está asociado a un lugar y área de la ciudad, con lo cual, se puede deducir que el usuario se ha movido.

	user_nsid character varying (51)	grid_id numeric (10)	date_taken timestamp without time zone	
1	10006319@N00		14866	2009-04-11 11:27:01
2	10006319@N00		15189	2009-05-06 12:39:59
3	10006319@N00		14866	2009-05-07 12:53:25
4	100097189@N05		13610	2013-05-25 10:46:14
5	10028574@N02		15506	2012-11-29 17:09:02
6	10028574@N02		15506	2012-11-29 17:12:09
7	10028574@N02		14863	2012-11-29 17:18:17
8	10028574@N02		14863	2012-11-29 17:18:49
9	10028574@N02		14863	2012-11-29 17:22:30
10	10028574@N02		14863	2012-11-29 17:25:01
11	10028574@N02		14863	2012-11-29 17:27:57
12	10028574@N02		15506	2012-11-30 12:59:09
13	10028574@N02		15511	2012-11-30 13:13:52
14	10028574@N02		15511	2012-11-30 13:15:07
15	10028574@N02		15511	2012-11-30 13:18:59

Figura 13. Fichero de trabajo. Datos de partida para el estudio de movilidad.

Para el desarrollo de este apartado ha sido necesario el uso de PGADMIN 4, una aplicación de código abierto y multiplataforma que permite la gestión y administración de base de datos gestionadas por PostgreSQL.

## 4.2. Filtrado.

Una vez obtenido el fichero con los datos de partida, es necesario realizar un primer filtrado en PostgreSQL mediante PGADMIN 4. Los objetivos son saber la cantidad de fotos existentes por año y el número de usuarios por mes y año. Esto nos permitirá descartar del estudio aquellos años que no contengan información relevante.

El primer filtro, se consigue mediante la siguiente sentencia, la cual devuelve una tabla con el número de fotografías por año cuyo resultado se puede ver en la figura 14.

```
Select extract(year from date_taken) as año, count(date_taken) as fotos from  
multimedia_tourism_dbgis.photo_grid_valencia group by año order by fotos desc.
```

Año	Nº fotografías
2018	2
2014	1103
2013	4086
2012	6218
2011	8475
2010	6204
2009	7311
2008	6149
2007	4728
2006	2224
2005	1159
2004	690
2003	138
2002	73
2001	61
2000	29
1999	7
1998	1
1997	5
1989	8
1980	4
1966	1
1937	1

Figura 14. Resultado de consulta. Número de fotografías por año.

Los resultados obtenidos permiten descartar, por ejemplo, el año 2018, el cual solo contiene dos fotografías además del periodo comprendido entre 1937 y 2006, el cual no aporta un numero relevante de fotografías.

A continuación, el siguiente filtro consiste en saber para cada año del periodo elegido el número usuarios por mes. Esto permite descartar los años que no aporten datos

significativos ya que se pretende entre otras cosas ver la evolución mensual dentro de un año. La sentencia que se utiliza es la siguiente:

```
select user_nsid, extract(year from date_taken) as año, count(multimedia_identifier) as  
fotos from multimedia_tourism_dbgis.photo_grid_valencia where extract(year from  
date_taken) = 2013 group by user_nsid,año order by fotos desc;
```

Los resultados obtenidos (*figura 15*) permiten descartar el año 2014 puesto que a partir de mayo no se devuelve ningún número de usuarios por mes.

	2014	2013	2012	2011	2010	2009	2008	2007
Año	60	195	262	344	346	309	292	282
Enero	22	25	38	46	40	30	30	26
Febrero	15	23	40	58	34	36	32	24
Marzo	28	45	55	79	42	52	52	53
Abril	15	28	36	44	35	42	29	41
Mayo	0	29	52	55	47	44	41	35
Junio	0	17	33	58	62	34	26	32
Julio	0	25	34	47	78	39	42	38
Agosto	0	19	34	34	34	50	53	29
Septiembre	0	28	37	33	39	48	36	27
Octubre	0	23	36	50	41	41	43	37
Noviembre	0	21	32	30	44	46	40	29
Diciembre	0	19	45	38	45	45	36	41

Figura 15. Resultado de consulta. Número de usuarios por año y mes.

Una vez hechas estas comprobaciones, los datos del periodo elegido se exportan en formato .CSV ya que es un formato sencillo de utilizar por programas de tratamiento de datos masivos como Python. Se obtiene un fichero .CSV para cada año (2007-2013).

### 4.3. Enfoque.

Una vez obtenidos y filtrados los datos necesarios para realizar el estudio de movilidad, el siguiente paso consiste en determinar el tipo de enfoque que se pretende llevar a cabo. En este trabajo se decide realizar dos enfoques:

- El primer análisis está enfocado en el estudio anual del periodo de tiempo elegido.
- El segundo análisis se centra en un año, concretamente el año 2013 ya que es el año más reciente de los que se han elegido. La finalidad es obtener información acerca de la evolución mensual a lo largo de dicho periodo de tiempo.

En ambas opciones es necesario distinguir entre usuarios turistas y usuarios residentes locales. Para ello se hace uso de Python y de sus diferentes librerías. Python es un lenguaje de programación multiparadigma, ya que soporta parcialmente la orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje



interpretado, dinámico, multiplataforma y con licencia de código abierto. Las librerías son módulos con funciones específicas de un área que aportan más funcionalidad al programa. En este trabajo se ha hecho uso de la versión 3.7.

Dado que se ha utilizado para la mayoría de los pasos seguidos en el desarrollo de este proyecto, se explicará en cada apartado si se ha hecho uso de Python y en caso de hacerlo, se indicarán las librerías utilizadas. Por lo que respecta a este apartado no ha sido necesario el uso de librerías específicas ya que, para diferenciar entre turistas y residentes, dependiendo del tipo de análisis escogido (anual o mensual), se implementan unas condiciones u otras en las cuales permiten la clasificación de los datos.

Por ejemplo, en el enfoque anual la condición implementada permite clasificar los datos entre turistas y residentes en función del tiempo que tarda en volver un usuario al mismo sitio. El umbral escogido ha sido de seis meses. Es decir, si un usuario realiza una foto en un lugar dado y en una fecha concreta y se encuentra otra foto del mismo usuario, aunque sea en otro lugar, pero dentro de un intervalo de tiempo menor a seis meses, se considera residente ya que se parte de la premisa de que un turista no permanece en un lugar por mucho tiempo. Cada fichero .CSV es clasificado y los resultados, uno por cada año, son exportados a QGIS para la visualización de estos mediante mapas de calor con la finalidad de identificar los lugares con mayor concentración de fotografías por año.

En cuanto al segundo enfoque la condición utilizada para diferenciar entre residentes y turistas se basa en el número mínimo de fotografías publicadas por un usuario, cuyo umbral es 2 y, en el tiempo que permanece en el mismo lugar, cuyo umbral es de dos semanas. Por ejemplo, si un usuario, durante un mes dado solo realiza una foto se descarta o si realiza una foto y después de dos semanas vuelve a realizar otra, se considera residente y por tanto también se descarta. En este caso, se utiliza el fichero anual del 2013 y se obtienen doce ficheros a los cuales se les aplica técnica de representación mediante grafos y obtener rutas mediante nodos y aristas.

#### 4.4. Representación de resultados.

Como se ha explicado en el apartado anterior dependiendo del enfoque utilizado representaremos los resultados de una u otra forma.

Por lo que respecta al enfoque anual, se hace uso de QGIS para representar los resultados mediante mapas de calor. QGIS es un sistema de información geográfica multiplataforma



y de código abierto. Es capaz de visualizar, gestionar y analizar datos y visualizarlos de múltiples maneras como, por ejemplo, en modo de mapas a escala. La versión que se utiliza en este trabajo es la 3.18.2-Zürich. Presenta novedades y mejoras en interfaz de usuario, simbología, renderizado, creación de vistas 3D, integración de nubes de puntos, entre otras.

En cuanto al segundo enfoque, se decide realizar un tratamiento más profundo. A los datos clasificados se les aplica la teoría de grafos y detección de comunidades mediante las librerías *NetworkX* y *Community* de Python explicadas en el apartado de conceptos teóricos. Los resultados obtenidos de la aplicación de estas librerías se representan gráficamente mediante la librería *Matplotlib*. Para el estudio, se utilizan los datos del año 2013 puesto que es el año más cercano a la actualidad que contiene datos relevantes para utilizar en este proyecto. Se realiza un análisis mensual para determinar qué rutas escogen en diferentes meses y se visualizan los resultados en QGIS para una mejor interpretación.

Adicional a esto se han obtenido para ambos enfoques los estadísticos *betweennes centrality* y *Edge betweennes centrality* con la finalidad de obtener las zonas céntricas y las aristas más concurridas. Esta información ayuda a identificar las principales atracciones turísticas y corroborar lo que se obtiene mediante grafos y comunidades.

## 5. Resultados

En este apartado se presentan los resultados obtenidos para ambos estudios.

### 5.1. Análisis anual periodo 2007-2013. Mapas de calor

A continuación, se presentan los resultados obtenidos para el análisis realizado durante el periodo comprendido entre los años 2007 y 2013. Como se comentó con anterioridad a partir de los datos obtenidos, se distinguió entre residentes y turistas. Los años no relevantes fueron descartados y los años que contenían información se representaron en mapas de calor mediante QGIS. En las siguientes figuras se pueden observar los mapas de calor de cada año analizado. En cada mapa se puede observar el número de fotografías por área, siendo el color rojo representativo de los lugares con mayor densidad y azul de las áreas menos fotografiadas.

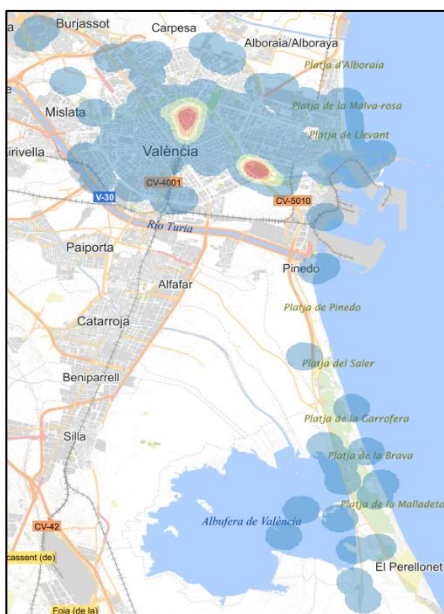


Figura 16 Mapa de calor. 2007

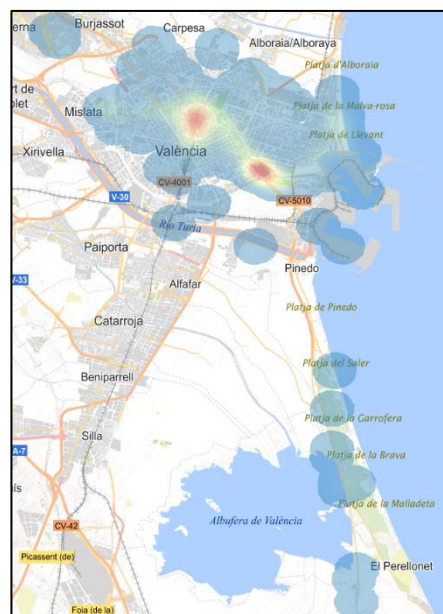


Figura 17. Mapa de calor. 2008

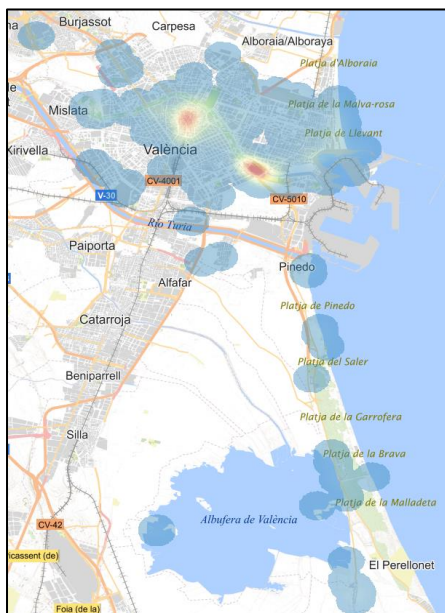


Figura 18. Mapa de calor. 2009

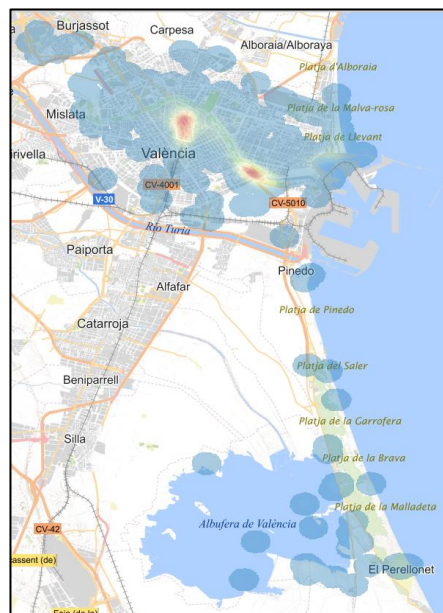


Figura 19. Mapa de calor. 2010

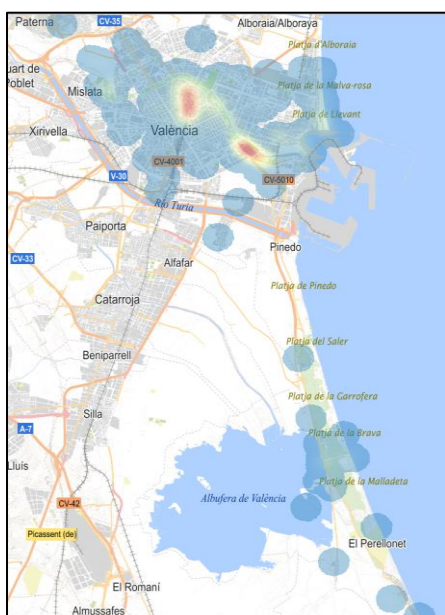


Figura 2011

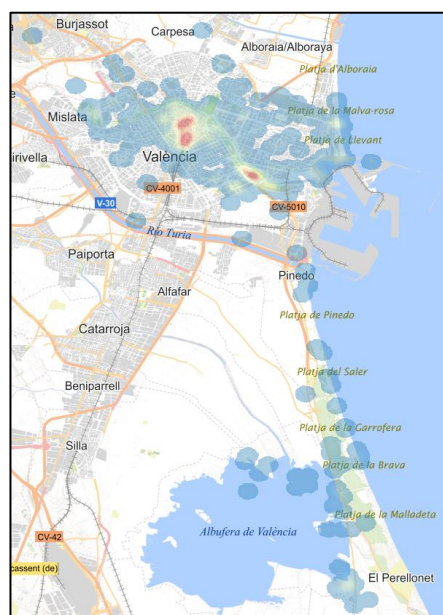


Figura 21. Mapa de calor. 2012

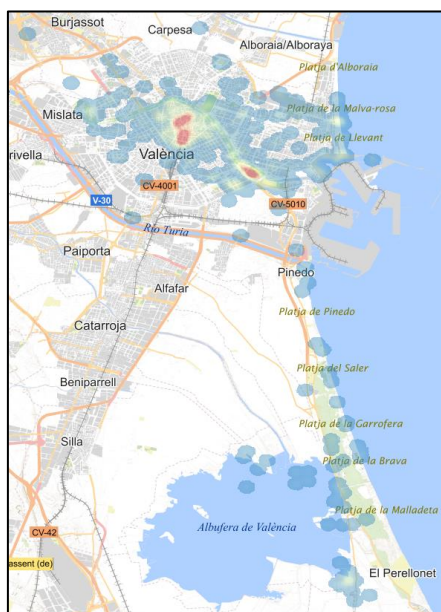


Figura 22. Mapa de calor. 2013

Gracias a estos mapas de calor se puede ver que las principales zonas de interés para los turistas a lo largo del periodo de tiempo elegido son la zona de la Ciudad de las Artes y las Ciencias y la zona del centro de la ciudad, Ayuntamiento, Plaza de la Virgen, Casco histórico, etc. Por otro lado, también hay otras zonas a las que los turistas llegan y realizan fotografías, aunque con menor frecuencia. Dichas zonas se corresponden con la zona de la playa, la zona del parque de Cabecera, L'Albufera y Blasco Ibáñez entre otras. A lo largo del año se puede ver una ligera dispersión de las fotografías por lo que puede significar que los turistas visitan más lugares turísticos.

## 5.2. Análisis mensual año 2013.

En este apartado se presentan los resultados que se obtienen tras aplicar a los datos del año 2013 los algoritmos *de NetworkX* para la representación de grafos y el algoritmo *Community Louvain* para la detección de comunidades.

Por lo que respecta a los grafos se puede observar en la *figura 23* el resultado obtenido para el mes de Julio. Se puede observar cómo hay rutas aisladas, remarcada en color rojo, que pueden dar lugar a pensar que solo las realiza un usuario y, rutas más concurridas, remarcada en verde, donde seguramente habrá más de un usuario realizando dicha ruta o serán rutas que tienen nodos en común. Para consultar el resto de los grafos ver ANEXO.



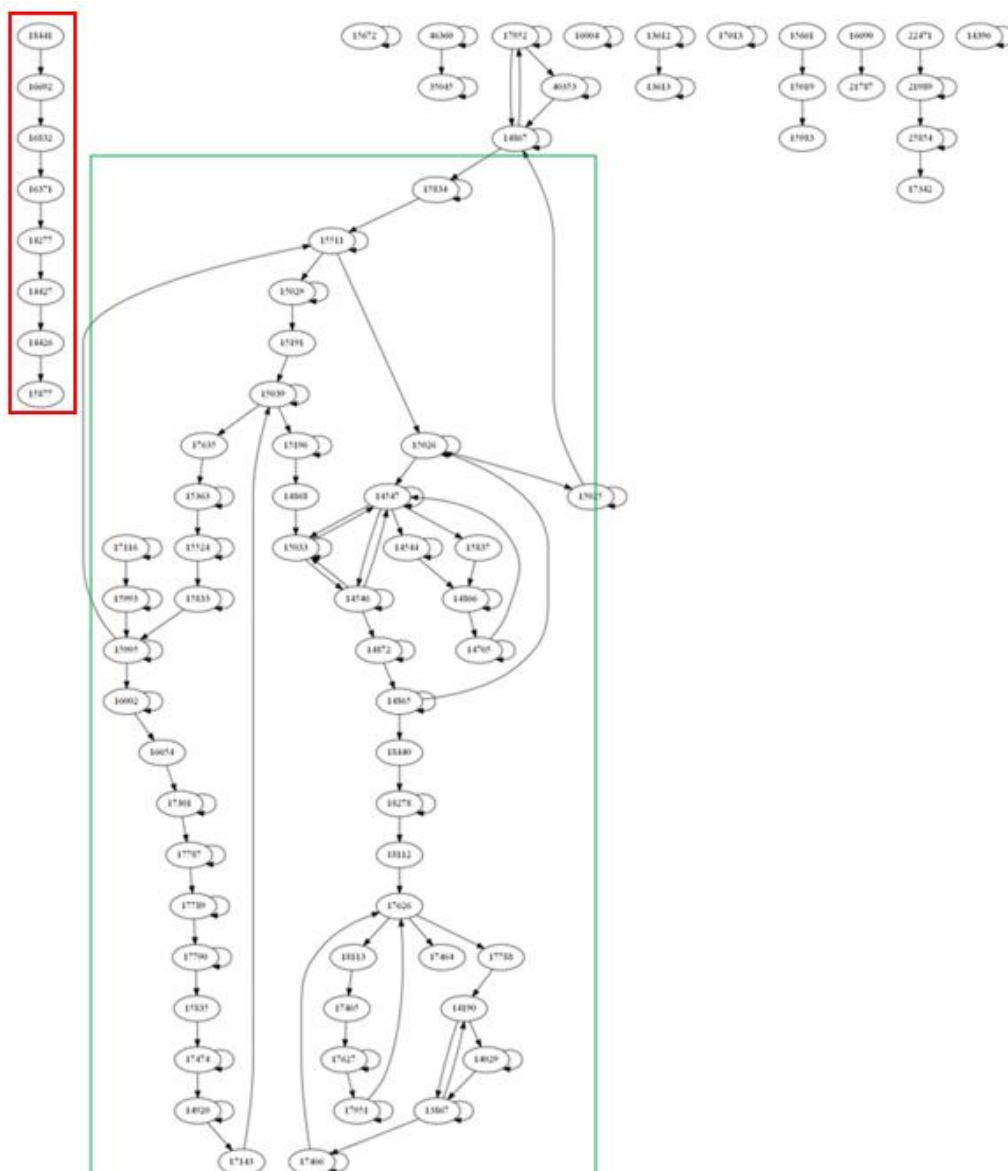


Figura 23. Grafo obtenido para el mes de julio de 2013

A partir de los grafos dirigidos se aplica la detección de comunidades de Louvain para encontrar alguna similitud entre las aristas y agrupar todos los vértices que estén fuertemente relacionados. Este algoritmo tiene una cierta componente de aleatoriedad y en cada ejecución se obtiene un resultado diferente que el anterior. Sin embargo, si los nodos mantienen una estrecha relación es muy probable que el algoritmo los agrupe a la misma comunidad. El tamaño de los nodos viene determinado en función de su centralidad, es decir, un nodo o vértice es más grande si varios usuarios pasan por dicho nodo. La visualización de las comunidades ayuda a detectar posibles rutas que siguen los usuarios durante un determinado periodo de tiempo.

La *figura 24* representa el resultado obtenido para el mes de enero. En una primera aproximación se puede decir que el algoritmo ha encontrado 3 comunidades (rosa, verde oscuro, verde claro). El *grid* 41008 (El Perellonet) parece ser la atracción más popular dado su tamaño, seguido de los *grids* 17794 (Ciudad de las Artes y las Ciencias) y 39575 (Zona Perellonet/ L'Albufera).

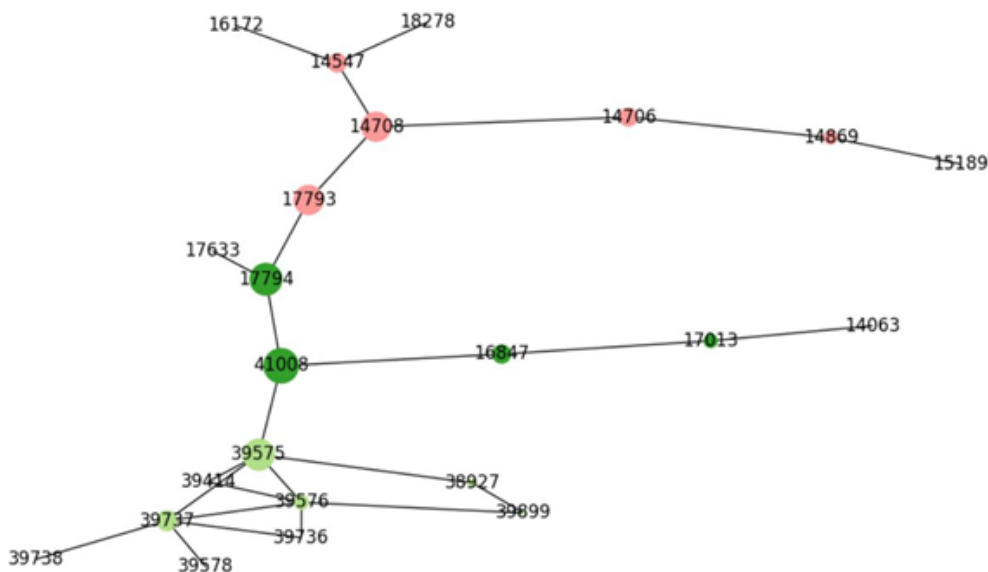


Figura 24. Detección de comunidades. Mes de enero 2013.

Para corroborar este resultado se presenta en la Tabla 1 el resultado del estadístico *betwenness centrality* de la librería Networkx, el cual determina terminar qué *grids* son los más concurridos por los turistas. Se puede observar en dicha tabla las diez aristas más populares y su nivel de centralidad.

Tabla 1. Estadístico de centralidad para el mes de enero de 2013.

RUTAS ENERO				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
17793	14708	0.04036036	Ciudad de las Artes y Ciencias	El Miguelete
17794	17793	0.043243243	Ciudad de las Artes y Ciencias	Ciudad de las Artes y Ciencias
39575	41008	0.045405405	El Palmar	El Perellonet
14708	14547	0.021621622	Iglesia del Sagrado Corazón de Jesús, La Lonja	El Miguelete
14708	14706	0.021621622	Iglesia del Sagrado Corazón de Jesús, La Lonja	El Miguelete
39737	39575	0.020660661	El Perellonet	El Perellonet
41008	16847	0.021621622	El Palmar	Puerto de Valencia, Zona Las Arenas



14706	14869	0.015135135	La Lonja	Plaza de la Virgen
16847	17013	0.015135135	Puerto de Valencia	Puerto Copa América
39576	39575	0.011111111	El Perellonet	El Perellonet

A la vista de los resultados indicados en la *figura 24* y la *Tabla 1*, se puede inferir que es probable que un turista que se encuentra en la zona de L'Albufera o El Perellonet se desplace hacia las zonas del puerto de Valencia para realizar visitas en la zona. Por ello, una posible ruta para el mes de enero podría ser la que comprende L'Albufera, Puerto de Valencia y sus alrededores como Puerta Copa América, etc. Por otro lado, se observa la existencia de otras comunidades que no se mueven por la costa valenciana y de las que se puede extraer otra posible ruta. Esta ruta la formarían la Ciudad de las Artes y las Ciencias y las zonas del centro y casco histórico de Valencia (Zona de la plaza del mercado, la Lonja, etc.).

Exportando dichos resultados a QGIS y visualizándolos en un mapa de la zona, *figura 25*, se pueden observar conexiones entre L'Albufera, zonas de playa y centro de Valencia. Esto podría ser debido a la existencia de una línea de bus que recorre la costa valenciana y está directamente conectada con el centro de la ciudad permitiendo a los turistas desplazarse con mayor fluidez por estas zonas.



Figura 25. Mapa de Valencia. Desplazamientos del mes de enero 2013.

A continuación, se presentan los resultados obtenidos tras la aplicación del algoritmo de detección de comunidades y estadísticos para el resto de meses junto con una interpretación de los mismos.

## Febrero

A partir del grafo obtenido para el mes de febrero, *figura 26*, se observa muy poca conexión entre las rutas lo cual podría ser indicativo de que los turistas toman rutas diferentes para recorrer la ciudad de Valencia

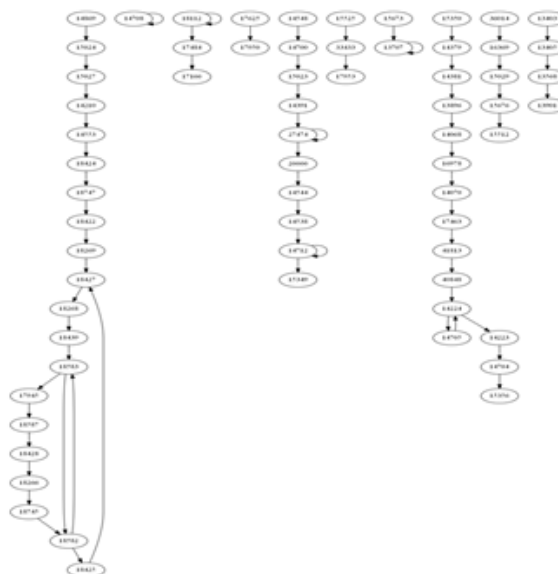


Figura 26. Grafo mes de febrero 2013.

Tabla 2. Estadístico de centralidad para el mes de febrero de 2013.

RUTA POPULAR FEBRERO				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
18422	18269	0.01978973	Polideportivo, La Fonteta	Centro Comercial, El Saler
18747	18422	0.01875902	Polideportivo, La Fonteta	Restauración, C/ Antonio Ferrandis
18424	18747	0.01731602	Polideportivo, Cuatro Carreres	Restauración, C/ Antonio Ferrandis
14553	18424	0.01546073	Puente de Real	Polideportivo, Cuatro Carreres
18427	18425	0.01401773	Zona Centro Comercial, El Saler	Polideportivo, Cuatro Carreres
14210	14553	0.01319316	Puente de Ademuz	Puente de Real
18582	18425	0.01319316	Polideportivo, La Fonteta	Polideportivo, Cuatro Carreres
14070	17463	0.01154401	Jardines del Real o Viveros	Hemisfèric
17463	41813	0.01154401	Hemisfèric	El Palmar
41813	40848	0.01113173	El Palmar	El Palmar

Observando los valores de centralidad obtenidos para este mes, Tabla 2, no existe una centralidad evidente puesto que, como bien se ha comentado, los nodos están muy dispersos y sin conectarse con los otros. Los valores de centralidad para todos los nodos arrojan un valor entre  $[0.02, 0]$  lo cual indicaría que los turistas no siguen las mismas rutas o llegan a visitar atracciones turísticas diferentes. De igual manera, si se visualizan los resultados en un *mapa*, figura 27, se podría decir que los usuarios permanecen en el mismo lugar para hacer turismo y no se desplazan con demasiada frecuencia. Esto podría



Según la Tabla 3, los visitantes de la ciudad prefieren moverse por la zona del centro y sus alrededores. Para este mes, la zona de las torres de Serranos es la más frecuentada por los turistas el motivo podría ser el hecho de que en esta zona se celebran numerosos actos para dar comienzo a las Fallas y muchas personas se aglomeran para inmortalizar el momento. Además, con la inauguración de las Fallas se suman los fuegos pirotécnicos los cuales llaman la atención, no solo a los falleros sino a los turistas que visitan la ciudad.

Tabla 3. Estadístico de centralidad para el mes de marzo de 2013.

RUTA POPULAR MARZO				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
14063	14708	0.07547231	Las Torres de Serranos	El Miguelete
14547	14707	0.07292961	Plaza de la Virgen	El Miguelete
14708	17627	0.07272904	Plaza de la Virgen	Ciudad de las Artes y Ciencias
14547	15833	0.07188147	Plaza de la Virgen	Zona Estación del Norte
14868	14707	0.06873706	Plaza de la Reina	Plaza de la Reina
12269	11946	0.06770186	Estadio Nou de Mestalla	Zona Estadio Nou de Mestalla
17627	17626	0.06061724	Ciudad Artes y Ciencias	Ciudad de las Artes y Ciencias
17626	11946	0.06004788	Ciudad Artes y Ciencias	Zona Estadio Nou de Mestalla
14063	15038	0.05435041	Torres de Serranos	Puente de la Exposición
15833	14857	0.05163043	Zona Estación del Norte	Zona Angel Guimerà, Falla Borull Socors

En general, *figura 29*, se observa que existe una agrupación de rutas en la zona centro de la ciudad y en la Ciudad de las Artes y las Ciencias. Las zonas de interés son las Torres de Serrano, la plaza de la Virgen, la plaza de la Reina, zonas alrededor del ayuntamiento y estación del norte. Durante este mes, en dichos lugares y a raíz de la festividad se pueden encontrar espectáculos callejeros y una gran de puestos de comida, souvenirs, etc.

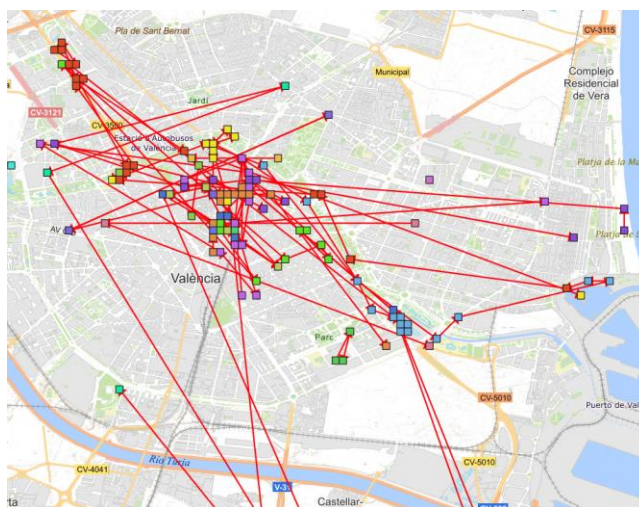


Figura 29. Mapa de distribución de comunidades durante el mes de marzo 2013.



## Abril

En el mes de abril el lugar más visitado es la Ciudad de las Artes y las Ciencias seguido por la zona del Mercado Central. Hay dos rutas bastante marcadas en abril. La primera no indica una ruta a seguir, sino que los turistas prefieren quedarse en la ciudad de las Artes y las Ciencias. La segunda es hacer visitas en la zona del Mercado Central y La plaza de la Reina.

Tabla 4. Estadístico de centralidad para el mes de abril 2013.

RUTA POPULAR ABRIL				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
17627	17790	0.07086835	El Hemisfèric	Ciudad Artes y las Ciencias
17790	17950	0.06722689	Ciudad Artes y Ciencias	Ciudad Artes y las Ciencias
17950	17953	0.06302521	Ciudad Artes y Ciencias	Puente L'Assut de l'Or
17627	18277	0.0627451	El Hemisfèric	L'Oceanogràfic
17953	17952	0.05826331	Puente L'Assut de l'Or	Ciudad Artes y las Ciencias
17952	18115	0.05294118	Ciudad Artes y Ciencias	L'Oceanogràfic
18115	18441	0.04705882	Ágora	L'Oceanogràfic
18441	17788	0.04061625	Ágora	El Hemisfèric
14707	15026	0.03361345	Mercado Central	El Miguelete
14707	14867	0.03361345	Mercado Central	Plaza Redonda/ Iglesia de Santa Caterina

Como se puede observar en la *figura 30* y, como se ha indicado antes, los turistas prefieren quedarse en una zona sin tener que desplazarse mucho. Un ejemplo de esto ocurre con los turistas que se quedan en el centro. En el caso de visitar la Ciudad de las Artes y las Ciencias existe una ruta desde el lugar hasta la estación del norte pasando por la Avenida Reino de Valencia. Esta ruta podría ser una de las populares por su comunicación con el centro de la ciudad debido a que algunas líneas de buses llegan de un extremo a otro conectando ambos lugares.



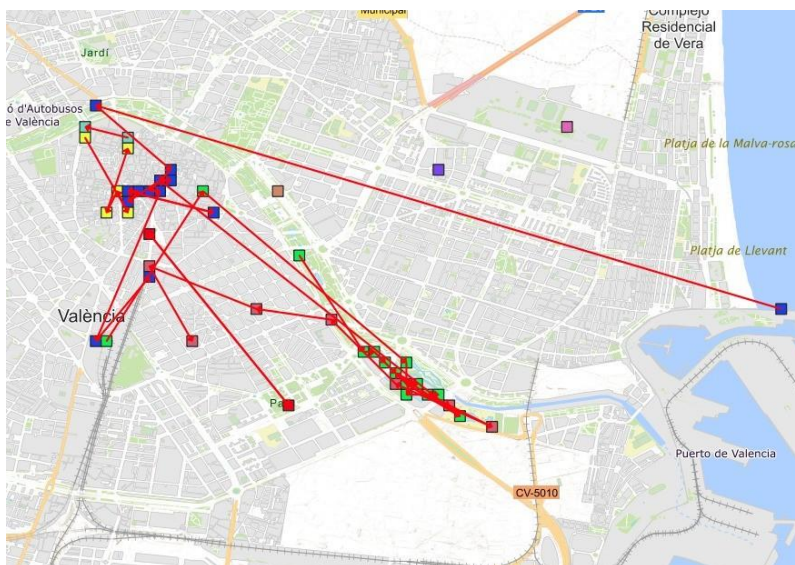


Figura 30. Mapa de distribución de comunidades durante el mes de abril 2013

## Mayo

Para este mes, gracias al grafo obtenido, *figura 31*, se puede observar que los turistas siguen rutas diferentes y hay pocas conexiones de nodos con los otros turistas.

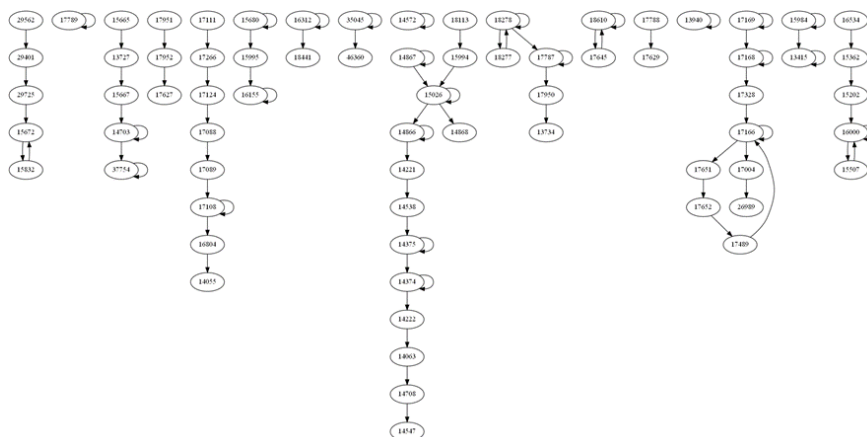


Figura 31. Grafo mes de mayo.

La Tabla 5 muestra valores muy pequeños de centralidad y por tanto no se tiene certeza de si las rutas que siguen son populares. No obstante, se puede decir que entre los lugares más visitados de este mes se encuentran el Jardín Botánico, Las Torres de Serranos y la plaza de Mercado y sus alrededores, *figura 32*.

Tabla 5. Estadístico de centralidad para el mes de mayo 2013.

RUTA POPULAR MAYO				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino

17328	17166	0.00336072	Edificio del reloj (Puerto)	Puerto de Valencia
15026	15994	0.00448096	Mercado Central	Estación del Norte
14063	14708	0.00448096	Torres de Serranos	El Miguelete
14222	14063	0.00616131	Barrio El Carmen	Torres de Serranos
14374	14222	0.00746826	Jardín botánico	Barrio El Carmen
15026	14866	0.00840179	Mercado Central	Mercado Central, La Lonja
14375	14374	0.00840179	Jardín botánico	Jardín botánico
14866	14221	0.00896191	Plaza del mercado, Mercado Central, La Lonja	Plaza Tavernes de Valldigna
14538	14375	0.00896191	Torres de Quart	Jardín botánico
14221	14538	0.00914862	Plaza del mercado, Mercado Central, La Lonja	Torres de Quart

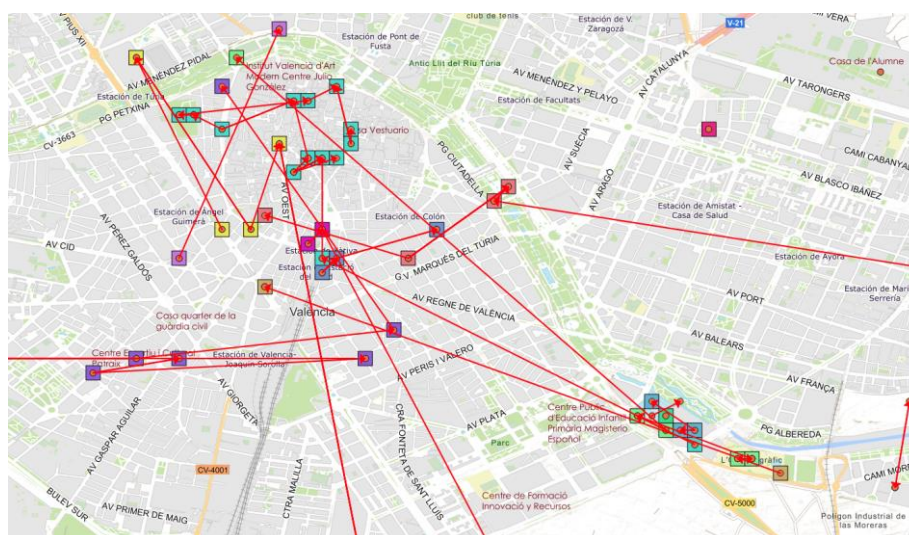


Figura 32. Mapa de distribución de comunidades durante el mes de mayo 2013

## Junio

Para el mes de junio, no se puede extraer rutas populares debido a la falta de información. Como se muestra en el grafo de la *figura 33* y en el mapa de distribución de comunidades de la *figura 34*, hay una escasez de datos y no se observan conexiones entre grids. Por tanto, se descarta el análisis para este mes.

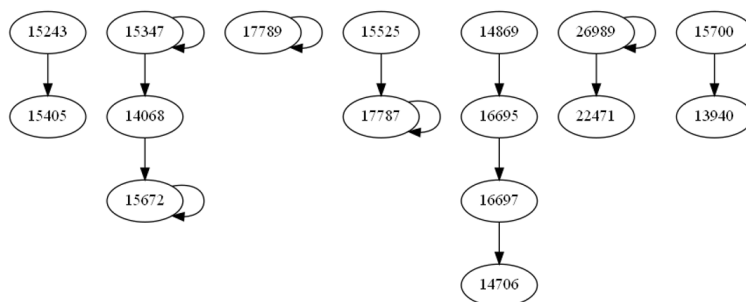


Figura 33. Grafo mes de mayo.

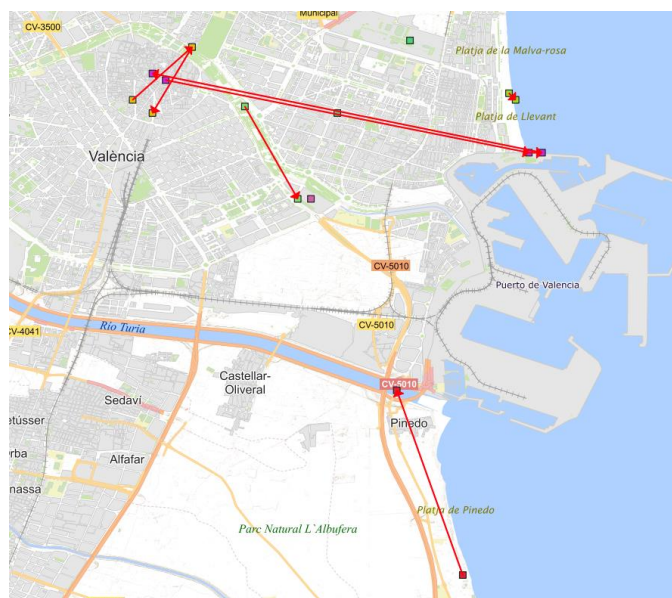


Figura 34. Mapa de distribución de comunidades durante el mes de junio 2013

## Julio

Viendo el grafo de la *figura 23* puede observar un grupo de nodos conectados que necesitan especial enfoque para extraer posibles rutas turísticas.

Los lugares más céntricos visitados por turistas en el mes de julio, Tabla 6, son las zonas del centro que comprenden la Plaza del Mercado Central, Mercado Central, La Lonja, la Plaza del Ayuntamiento, la Plaza de toros, entre otros.

Tabla 6. Estadístico de centralidad para el mes de Julio 2013.

RUTA POPULAR JULIO

Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
14865	15026	0.08558559	Plaza del mercado, Mercado central, La Lonja	Plaza del mercado, Mercado central, La Lonja
18440	18278	0.08304668	L'Oceanogràfic	L'Oceanogràfic
15026	15511	0.08149058	Plaza del mercado, Mercado central, La Lonja	Plaza del Ayuntamiento
18278	18112	0.07862408	L'Oceanogràfic	Ciudad de las Artes y las Ciencias
18112	17626	0.07387387	Ciudad de las Artes y las Ciencias	L'Hemisfèric
15995	15511	0.06519247	Plaza de toros	Plaza del Ayuntamiento
14547	15026	0.04103194	Plaza de la Virgen	Mercado central
15995	16002	0.03882064	Plaza de toros	Gran Vía Marqués de Túria
15511	15029	0.03865684	Plaza del Ayuntamiento	Plaza Redonda
15029	15191	0.03390663	Plaza Redonda	Iglesia de San Juan de la Cruz

Una vez más, la zona de la Ciudad de las Artes y las Ciencias es la preferida por los turistas para visitar.

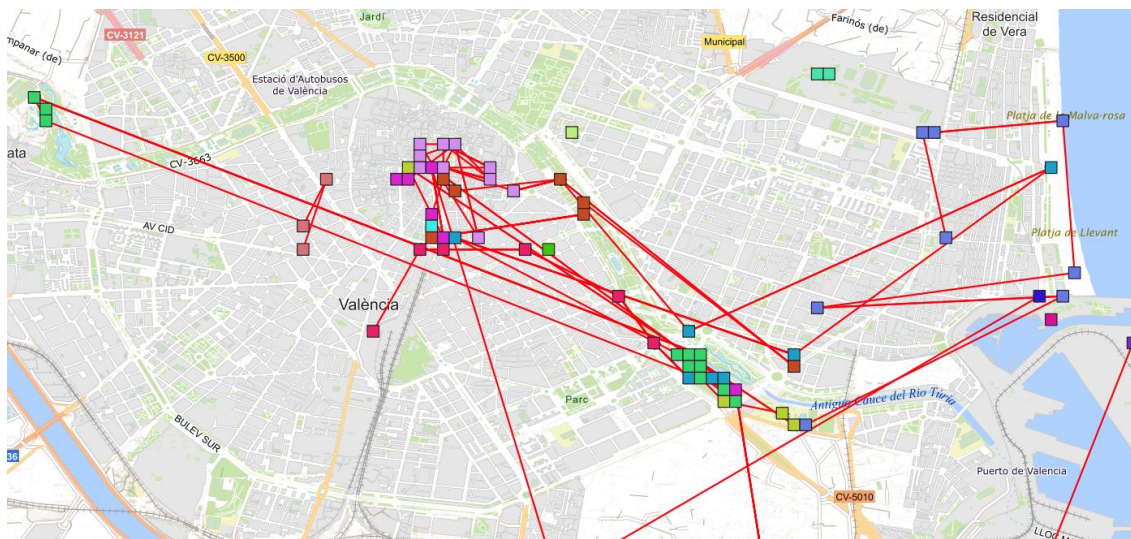


Figura 35. Mapa de distribución de comunidades durante el mes de Julio 2013

De este análisis se pueden sacar posibles rutas populares, figura 35. Una posible ruta consiste en realizar visitas en el centro de la ciudad pasando por la Plaza del Mercado, Plaza del Ayuntamiento, Plaza de Toros, etc. y la segunda posible ruta consiste en visitar las atracciones que se encuentran en el antiguo cauce del Río Turia como la Ciudad de las Artes y las Ciencias y el Parque de Cabecera.

## Agosto

Para el mes de agosto los turistas recorren las atracciones turísticas vía diferentes caminos y no se observan rutas similares. La Tabla 7 muestra unos valores de centralidad muy pequeños. Los lugares más visitados para este mes son la Ciudad de las Artes y las Ciencias, Plaza del Mercado, L'Albufera, etc.

Tabla 7. Estadístico de centralidad para el mes de agosto 2013.

RUTA POPULAR AGOSTO				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
16655	14865	0.00925583	Puente del Ángel Custodio, Parque Gulliver	Mercado central
16655	17627	0.0088856	Puente del Ángel Custodio, Parque Gulliver	Ciudad de las Artes y las Ciencias
17627	17626	0.0077749	Ciudad de las Artes y las Ciencias	Ciudad de las Artes y las Ciencias
14865	14868	0.00592373	Plaza del mercado, Mercado central	Plaza de la Reina
17117	15512	0.0044428	Estación Joaquin Sorolla	Plaza del Ayuntamiento
15483	17117	0.00370233	Av. Del Cid, Nou D'Octubre	Estación Joaquin Sorolla
14865	38107	0.0033321	Plaza del mercado, Mercado central	L'Albufera
14865	15357	0.0033321	Plaza del mercado, Mercado central	Plaza de Alfonso el Magnánimo
14544	14868	0.0033321	Iglesia Parroquial de San Nicolás, Barrio el Carmen	Plaza de la Reina
17626	17949	0.00296187	L'Hemisfèric	L'Umbracle

De las comunidades obtenidas se pueden detectar varias zonas de interés empezado por el centro de Valencia, *figura 36*. Algunos turistas no se desplazan más allá del centro y prefieren descubrir atracciones turísticas en la zona y otros prefieren desplazarse del centro a la Ciudad de las Artes y las Ciencias. Los turistas en la zona de playa tienen el mismo comportamiento y no se mueven del entorno donde se encuentran.

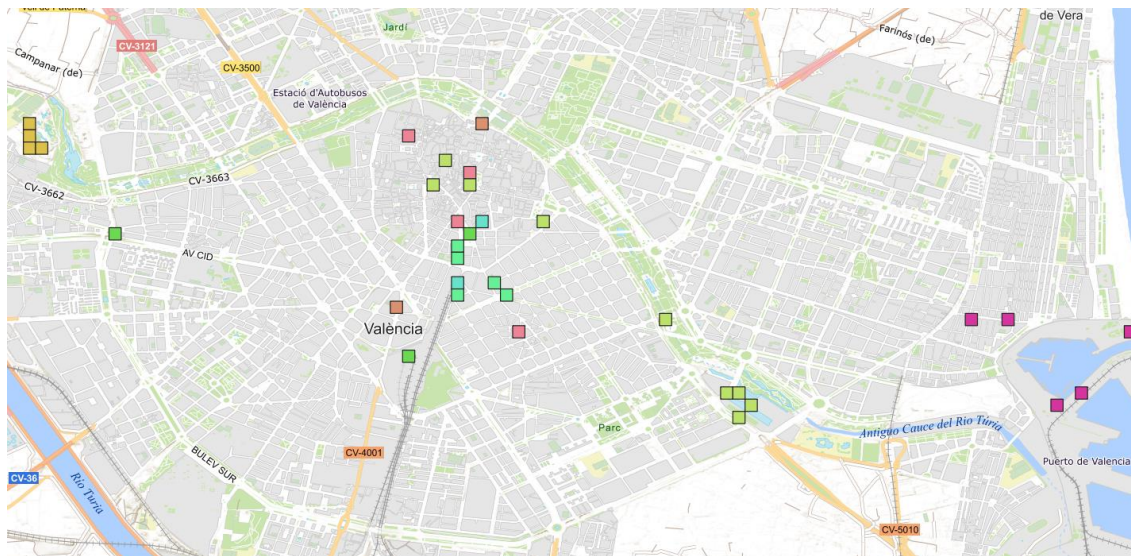


Figura 36. Mapa de distribución de comunidades durante el mes de agosto 2013

## Septiembre

Según la Tabla 8 la principal atracción o zona turística es la plaza de la Virgen y sus alrededores y la *figura 37* muestra, de forma general, las conexiones existentes entre nodos. Se puede observar que las zonas fuera de la ciudad como L'Albufera y el Perellonet están conectados de alguna manera con la zona centro de Valencia y los usuarios son capaces de desplazarse siguiendo esta ruta hacia ese sitio. Además, se puede ver grupos de turistas que se desplazan entre el centro del Valencia, la ciudad de las Artes y las Ciencias y la zona de la playa. Existen turistas que se quedan en el mismo sitio y concluyen su visita sin desplazarse mucho, pero también existen rutas más dispersadas, *figura 38*.

Tabla 8. Estadístico de centralidad para el mes de septiembre de 2013.

RUTA POPULAR SEPTIEMBRE				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
15672	16315	0.04890545	Pl. Ayuntamiento	Estación del Norte
15030	15189	0.04816799	Plaza de la Reina	Plaza de España
14865	14547	0.04527876	Mercado central	Plaza de la Virgen
40845	16315	0.04285049	L'Albufera	Estación del Norte
15189	14548	0.04013352	Plaza de España	Plaza de la Virgen
15189	14869	0.03990064	Plaza de España	Plaza de la Virgen
40684	40845	0.03648502	L'Albufera	L'Albufera
15403	17787	0.03648502	Zona playa la Malvarrosa	Ciudad de las Artes y las Ciencias
14547	14869	0.03640554	Plaza de la Virgen	Plaza de la Reina
14547	14548	0.03433732	Plaza de la Virgen	Plaza de la Virgen



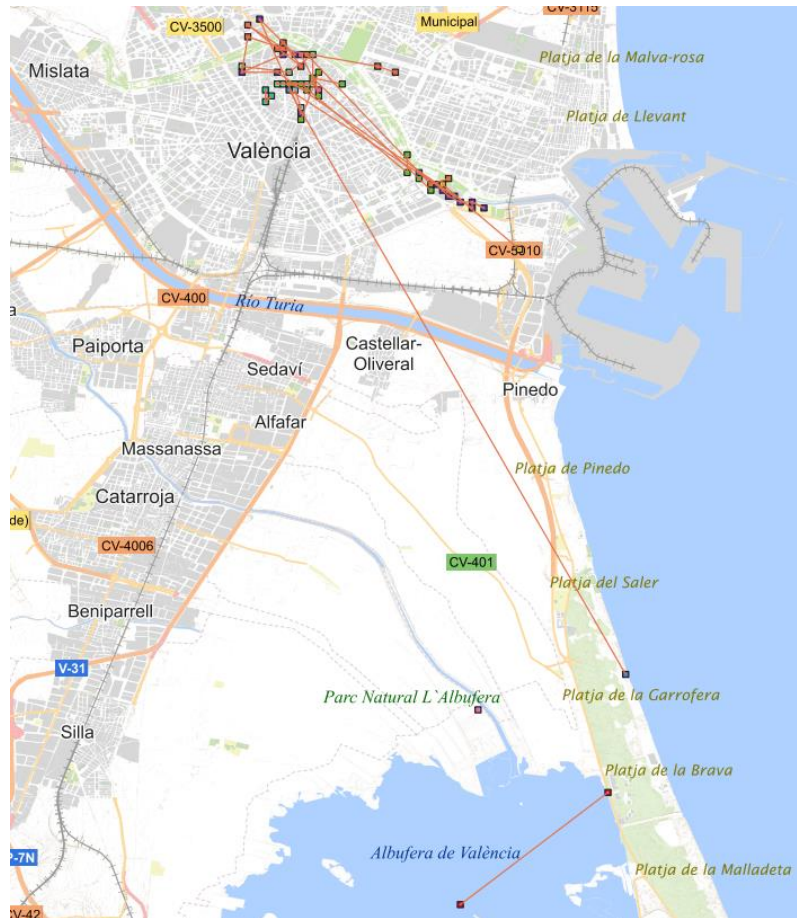


Figura 39. Mapa de distribución de comunidades durante el mes de octubre 2013

En este mes, los turistas prefieren visitar las zonas del centro y sus alrededores, Tabla 9, pero es cierto que existe más desplazamientos en el entorno en comparación con los meses ya comentados. Existe una ruta que empieza desde la ciudad de las Artes y las Ciencias hacia las zonas de interés como el jardín botánico, las Torres de Serranos, Plaza de la Virgen y Mercado Central y alrededores.

Tabla 9. Estadístico de centralidad para el mes de octubre de 2013.

RUTA POPULAR OCTUBRE				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
13570	14535	0.0366049	Iglesia Parroquial Juan Bosco	Jardín botánico
13898	13570	0.03363075	Plaza del Portal Nou	Iglesia Parroquial Juan Bosco
14535	14374	0.03363075	Jardín botánico	Jardín botánico
14541	18277	0.03363075	Institut Valencià de Cultura	L'Oceanogràfic
13248	13898	0.03019904	Bombas Gens Centre d'Art	Plaza del Portal Nou
14543	14374	0.03019904	Iglesia Parroquial de San Nicolás	Jardín botánico
18277	18116	0.03019904	L'Oceanogràfic	L'Oceanogràfic
14384	13248	0.02630977	Torre de San Bartolomé	Bombas Gens Centre d'Art





14064	14543	0.02630977	Torres de Serranos	Iglesia Parroquial de San Nicolás
18116	18114	0.02196294	L'Hemisfèric	El Àgora

## Noviembre

En el anexo. Grafo del mes de noviembre, se observan pocas rutas y, por tanto, no se puede realizar un análisis coherente ya que no existen conexiones entre grids..

## Diciembre

Para el mes de diciembre, Tabla 10, los turistas suelen moverse por la ciudad de las Artes y las Ciencias, zona que obtiene el mayor valor de centralidad en comparación con el resto. Durante ese mes se observa, *figura 40*, una concentración de rutas en el centro de Valencia, casco histórico y zona de la estación del Norte.

Tabla 10. Estadístico de centralidad para el mes de diciembre de 2013.

RUTA POPULAR DICIEMBRE				
Grid origen	Grid destino	Centralidad	Zona turística origen	Zona turística destino
18274	17788	0.07301587	El Àgora	L'Hemisfèric
18438	18274	0.07032967	L'Hemisfèric	El Àgora
13466	18438	0.06715507	Zona playa. La Patacona	L'Hemisfèric
17627	17951	0.06605617	L'Hemisfèric	Ciudad Artes y las Ciencias
14707	13466	0.06349206	Plaza de la Reina	Zona playa. La Patacona
17790	15506	0.06304436	Ciudad Artes y Ciencias	MUVIM
15026	15506	0.06060236	Mercado central	MUVIM
14869	14707	0.05934066	Plaza de la Reina	El Miguelete
17627	17790	0.05636956	L'Hemisfèric	Ciudad Artes y las Ciencias
15353	14869	0.05470085	Teatro principal	Plaza de la Reina

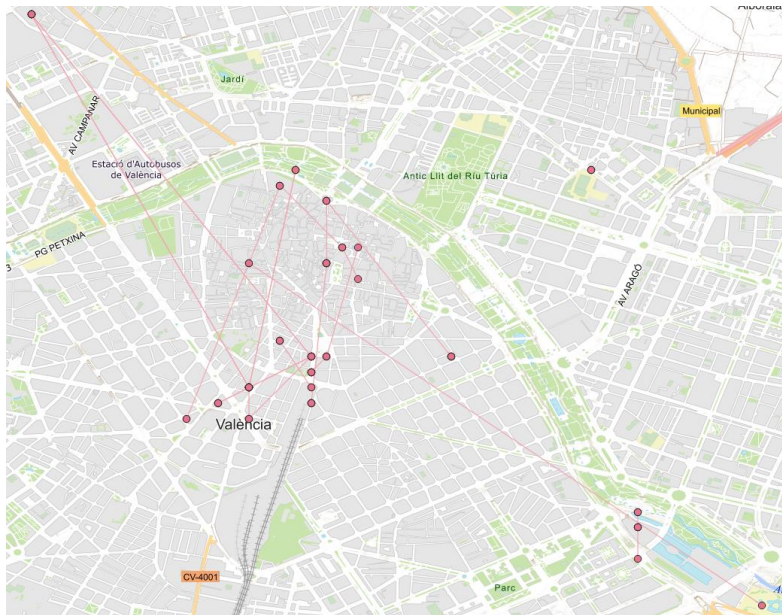


Figura 40. Mapa de distribución de comunidades durante el mes de diciembre 2013

Finalmente realizando una breve discusión de todos los resultados obtenidos, se puede destacar que, los mapas de calor desde el año 2007 hasta 2013 indican una cierta tendencia acerca de los lugares visitados. En líneas generales, los lugares más visitados de la ciudad han resultado ser la zona del centro de Valencia, la Ciudad de las Artes y las Ciencias y la zona de playa. Estas 3 zonas principales muestran una clara conexión entre ellas dejando ver que los turistas cuentan con varias opciones de desplazamiento lo cual permite que se muevan de forma sencilla. Se determinó también que la zona de L'Albufera, El Perellonet está, de alguna manera, conectada con la zona puerto de Valencia y la zona centro y que dicha conexión puede ser debido a la existencia de una línea de bus que conectan estos lugares. Durante algunos meses del año 2013 como febrero o julio no se ha comprobado de manera exitosa las rutas más populares debido a la falta de datos. No obstante, las rutas obtenidas del resto del mes son coherentes y los mapas de calor corroboran estos resultados.

## 6. Presupuesto.

### 6.1. Costes de personal.

En la siguiente tabla se muestran las tareas realizadas en orden cronológico junto con las horas invertidas en cada una para llevarla a cabo.

Tarea	Duración (h)
<b>Búsqueda de información</b>	<b>100</b>
• Búsqueda de información y aprendizaje de algoritmos Python	100
<b>Procesamiento de los datos y análisis</b>	<b>150</b>
• Gestión, análisis y tratamiento de los datos	75
• Obtención y comparación de los resultados	75
<b>Realización de cartografía</b>	<b>20</b>
• Representación de los resultados mediante un mapa	20
<b>Redacción de la memoria</b>	<b>80</b>
• Redacción	80
<b>Supervisión y resolución de dudas</b>	<b>40</b>

Tabla 11. Duración en horas de cada tarea realizada.

**Ingeniero en geomática:** se encargará de realizar todo el trabajo. Llevará a cabo la toma de datos en campo, su procesado, análisis e interpretación de los resultados. Además, se encargará también de realizarla cartografía final.

Perfil	Número de horas	Coste unitario	Coste imputable)
Ingeniero en Geomática	300	40 €/hora	16.500 €
		<b>Total</b>	<b>16.500 €</b>

Tabla 12. Coste del personal.

Consideraciones:

Dado que la totalidad de los softwares utilizados en este proyecto son de licencia gratuita no ha hecho falta considerarlo en el presupuesto.



## 7. Conclusiones

Los objetivos propuestos se han llevado a cabo de manera correcta y la metodología seguida da una idea de una alternativa de realizar análisis de movilidad mediante fuentes de información que no son las tradicionales. El uso de la información de las redes sociales es una fuente de datos que muchas personas quieren aprovechar puesto que existen muchas posibilidades y se puede enfocar a muchas aplicaciones.

Big Data, en combinación con los grafos, es un enfoque potente para realizar estudios de movilidad. Permite ver gráficamente las rutas realizadas por los usuarios, así como extraer valores estadísticos de estas rutas. Además, ha sido de gran ayuda la utilización de los mapas de calor ya que da una aproximación de las rutas más concurridas y los lugares más visitados.

La realización de este trabajo ha resultado gratificante y bastante retador. Para empezar, dominar la programación requiere esfuerzo, tiempo y práctica y con este trabajo se han demostrado estos valores plasmando los resultados obtenidos. La mayor inversión del tiempo para elaborar el proyecto se ha empleado en el aprendizaje más profundo de los conceptos de programación y en la búsqueda de información sobre las funcionalidades de las nuevas herramientas que se han utilizado. En un principio, los resultados obtenidos fueron difíciles de interpretar, pero con la ayuda de un sistema de información geográfica la interpretación de los resultados ha sido más sencillo.

Durante el desarrollo del trabajo han surgido ideas para dar otro enfoque al trabajo. La detección de comunidades permite ver la relación existente entre las rutas de los turistas y una de las ideas sería la utilización de otros algoritmos. En este trabajo se ha utilizado el algoritmo de detección de comunidades Louvain y los resultados conseguidos son adecuados. Sin embargo, existen más algoritmos que se podrían utilizar para contrastar los resultados con los obtenidos en este trabajo. Otra de las ideas es utilizar datos recientes y realizar un estudio que se adapte a la actualidad. En este trabajo se ha limitado a utilizar los *datasets* que liberalizó Flickr hace unos años y los datos más actuales son del 2013. Existen muchas plataformas como Facebook o Instagram que contienen millones de usuarios y, por tanto, millones de datos geolocalizados que podrían ser de utilidad para realizar un estudio de movilidad más actualizado.



## 8. Bibliografía

Hu, Fei & Li, Zhenlong & Yang, Chaowei & Jiang, Yongyao. (2018). A graph-based approach to detecting the tourist movement patterns using social media data. *Cartography and Geographic Information Science*. 46. 10.1080/15230406.2018.1496036.

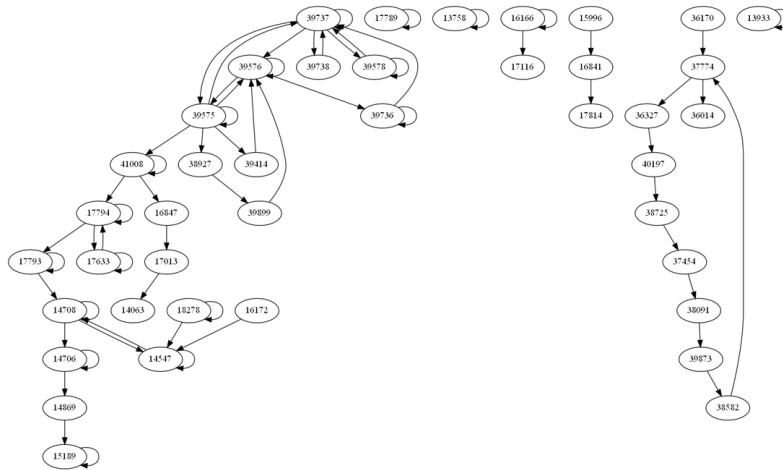
Romero, Oriol & Cerdan, Aurélie. (2018). Location-Based Social Network Data for Exploring Spatial and Functional Urban Tourists and Residents Consumption Patterns. 8. 35-52.



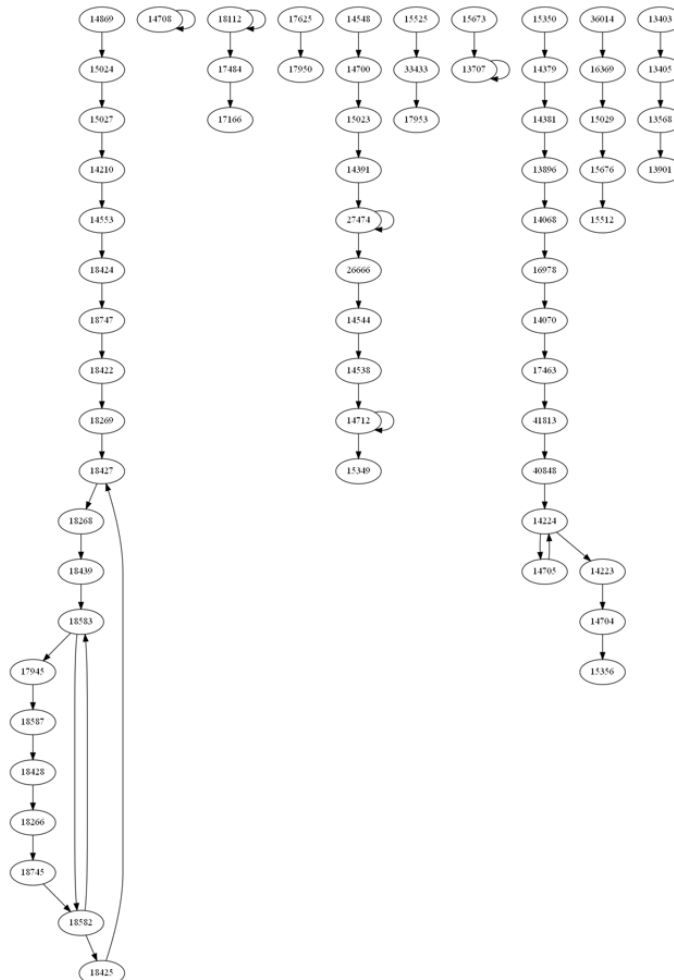
## 9. Anexos

### Grafos año 2013

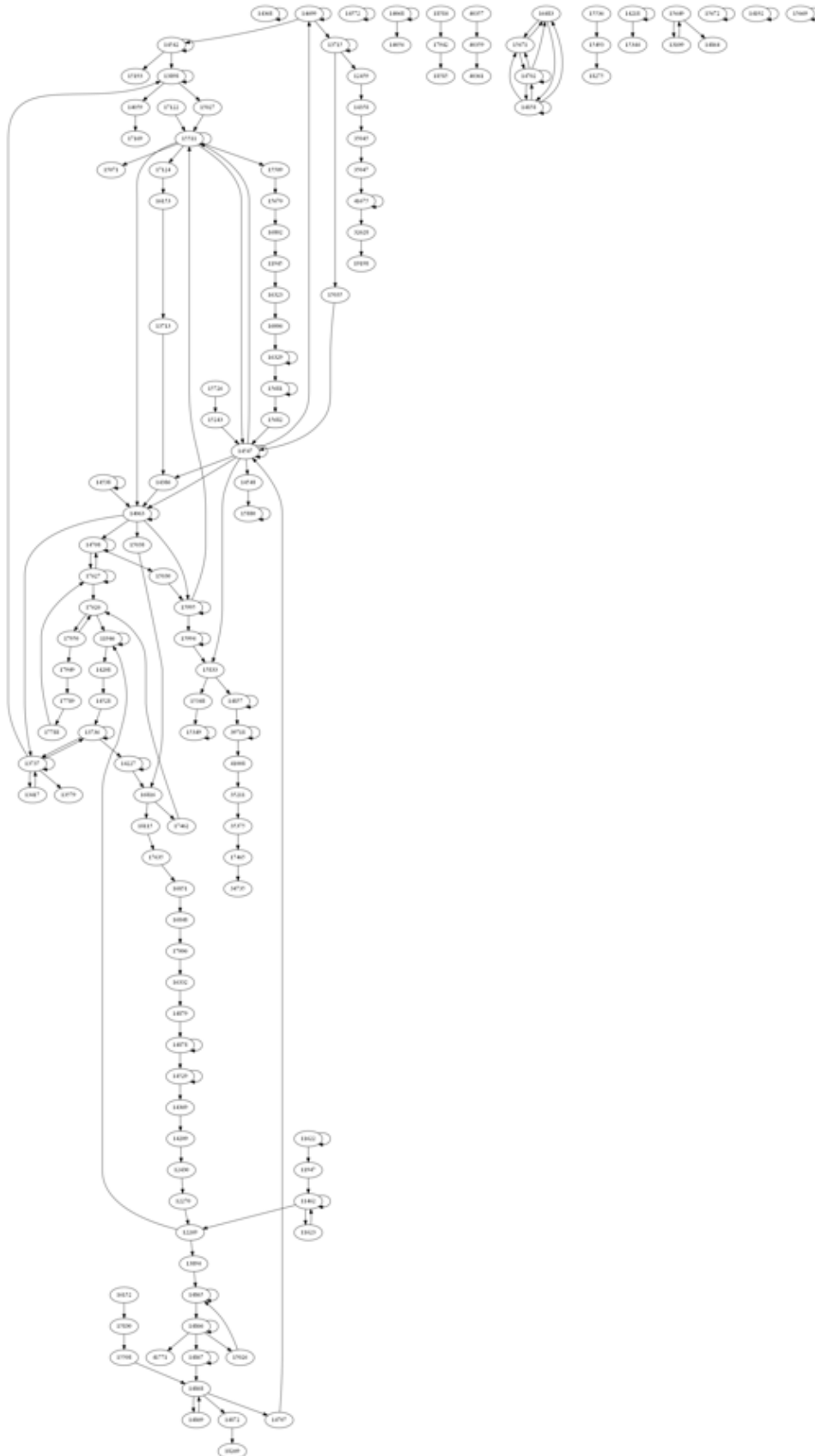
#### Enero



#### Febrero



## Marzo

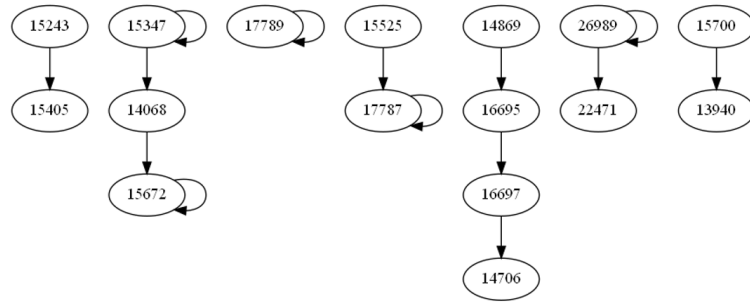




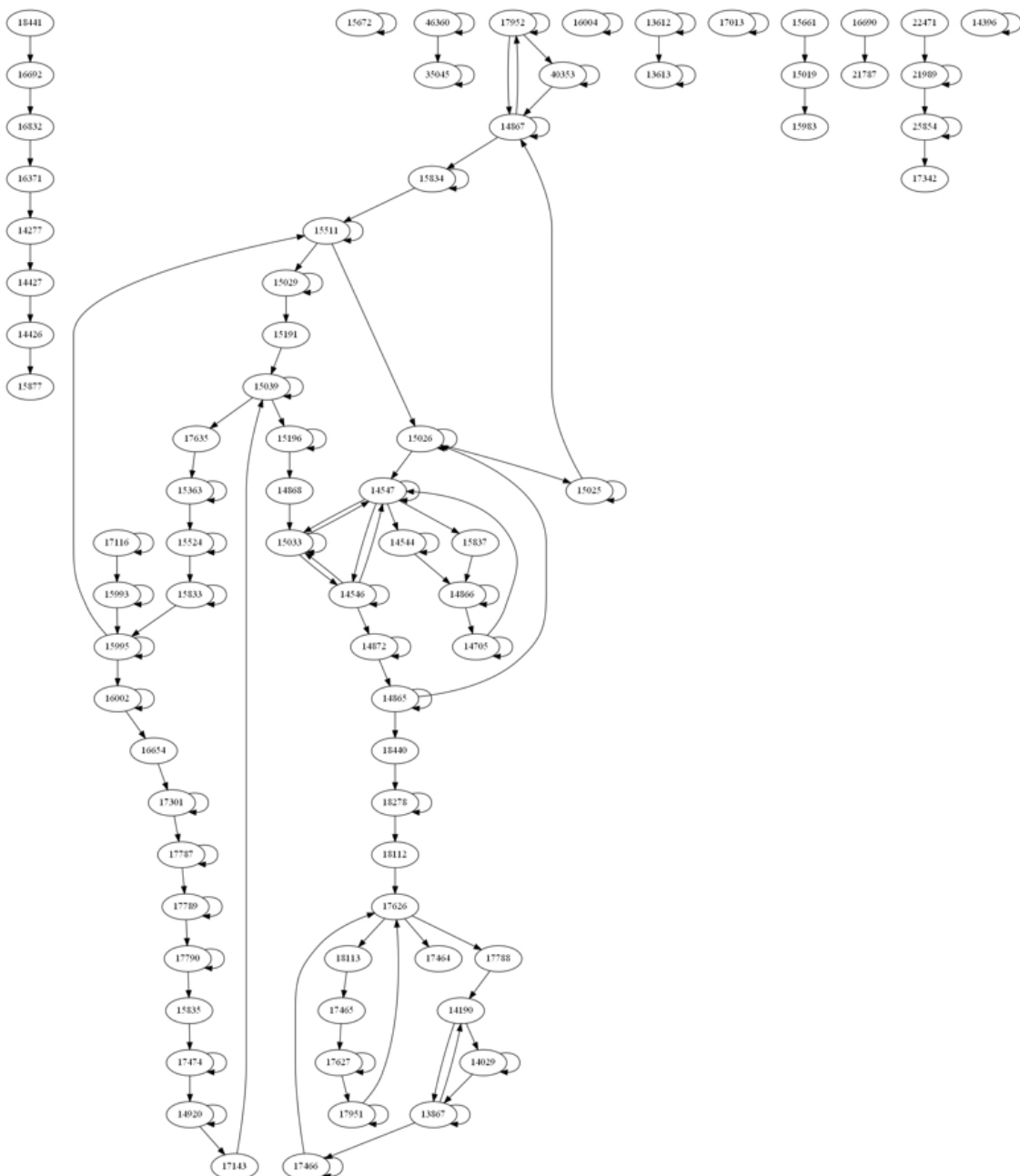




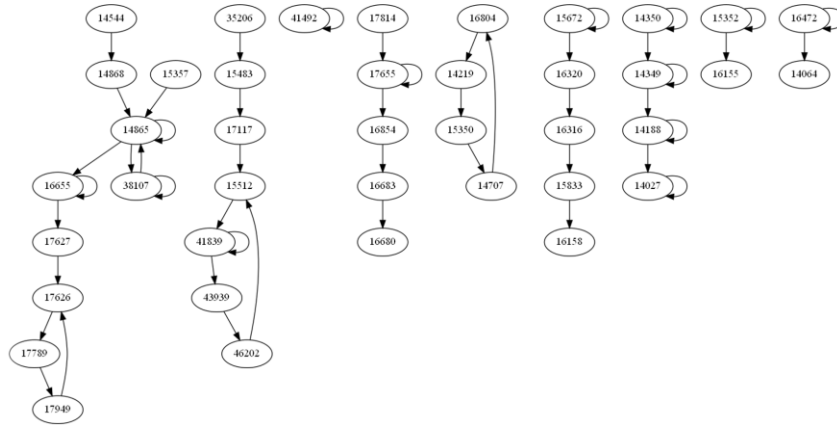
### Junio



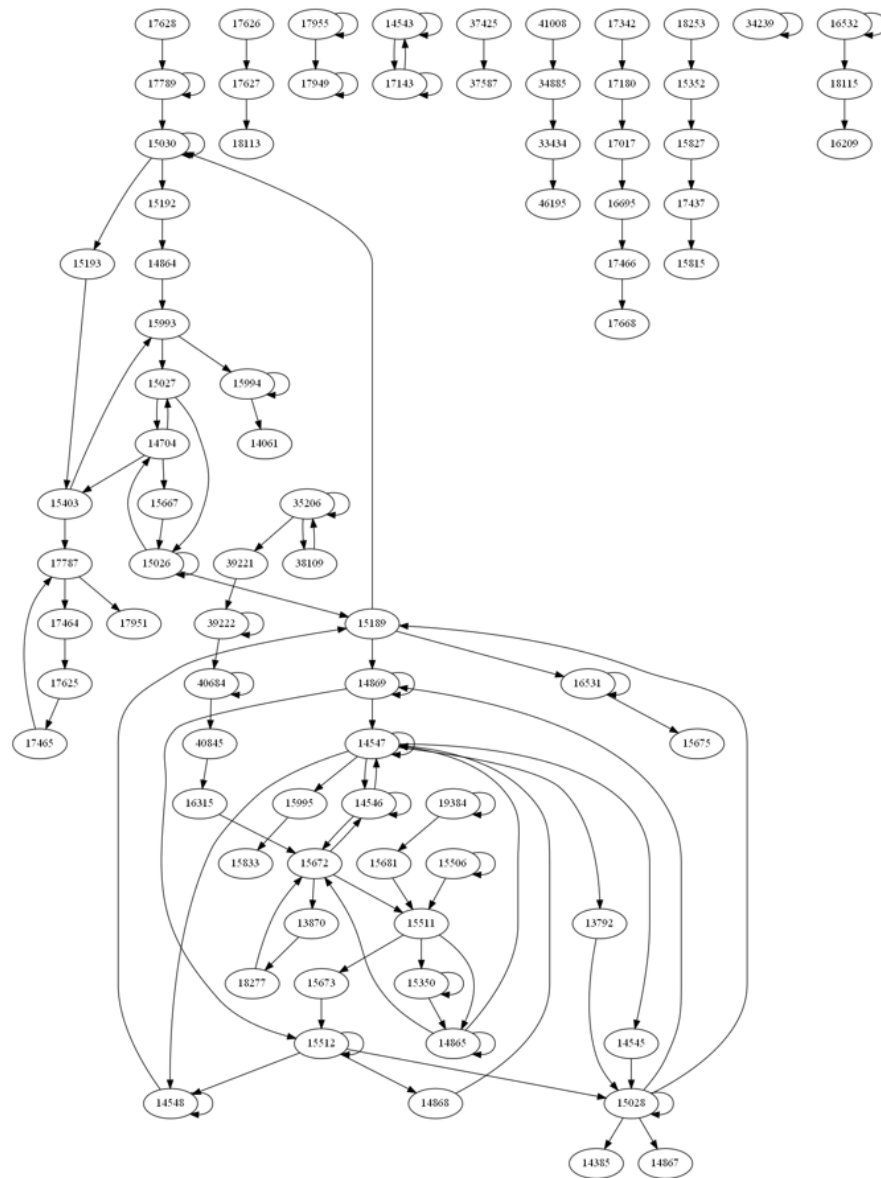
### Julio



### Agosto



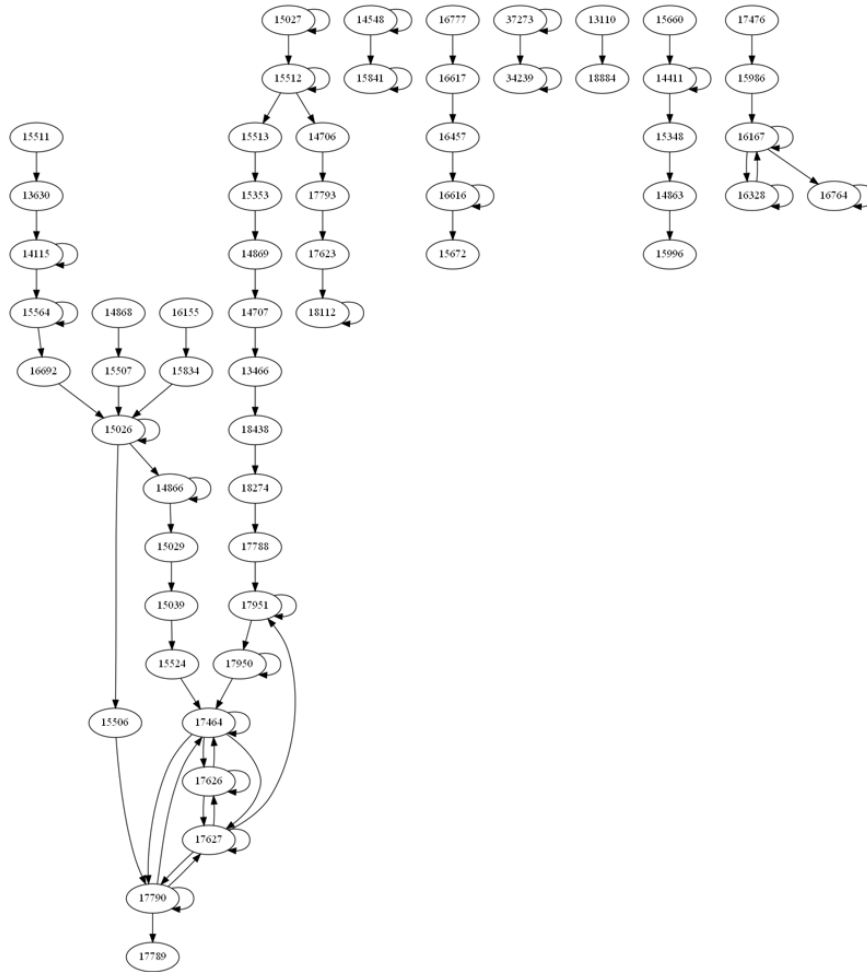
### Septiembre





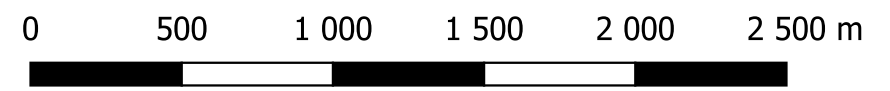
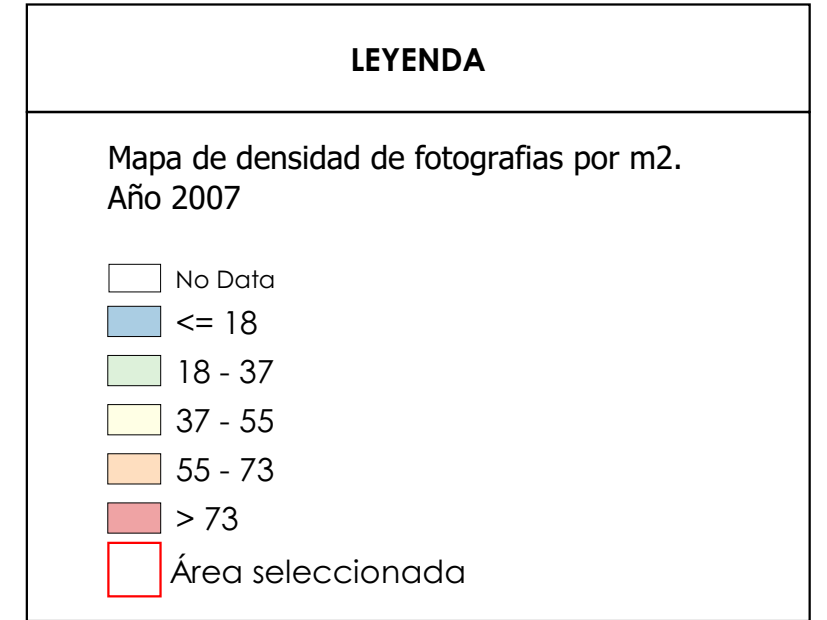
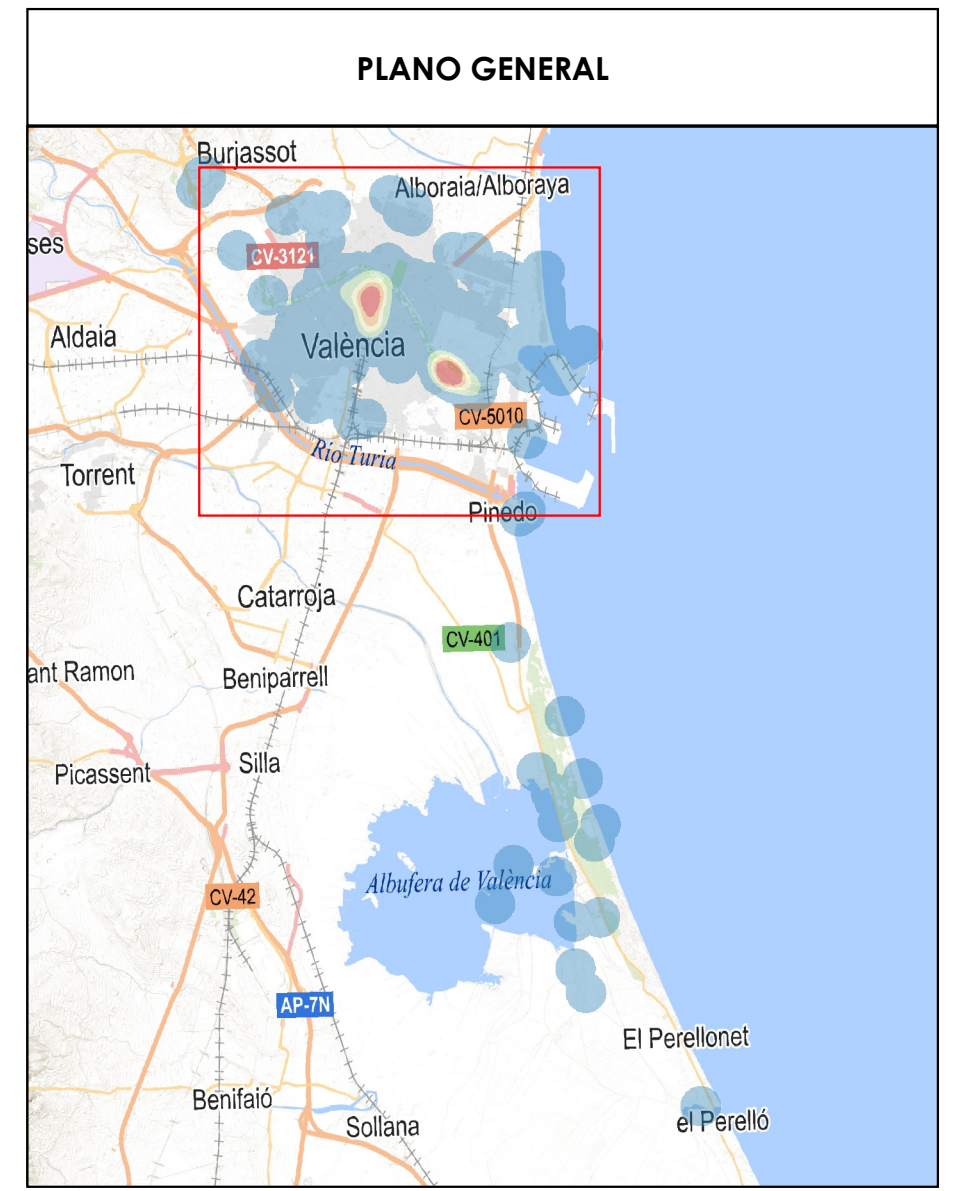


## Diciembre



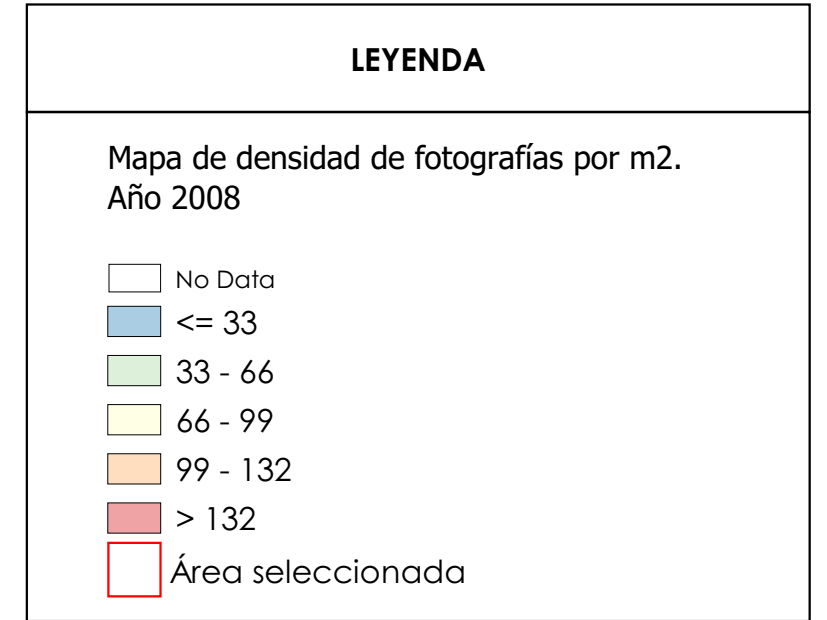
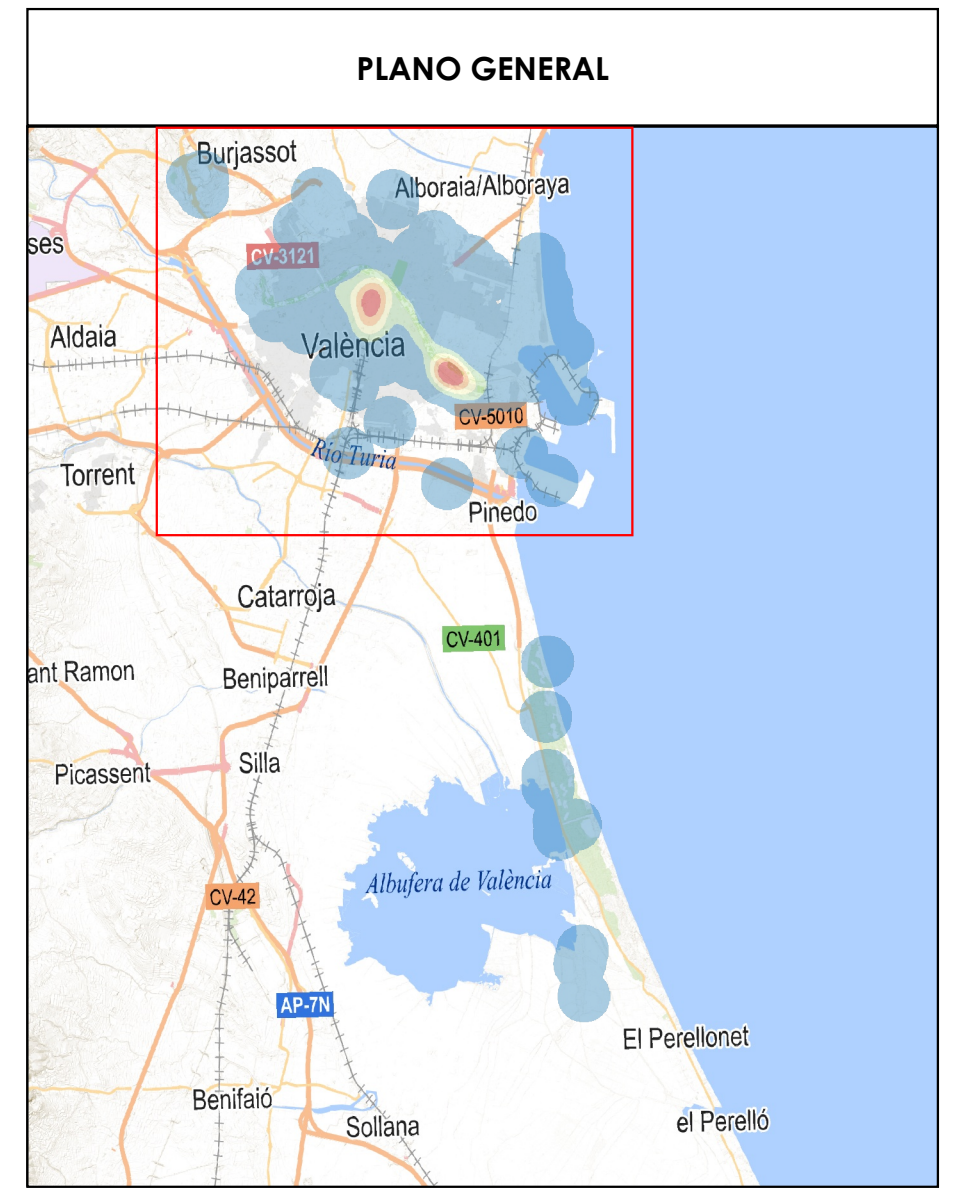
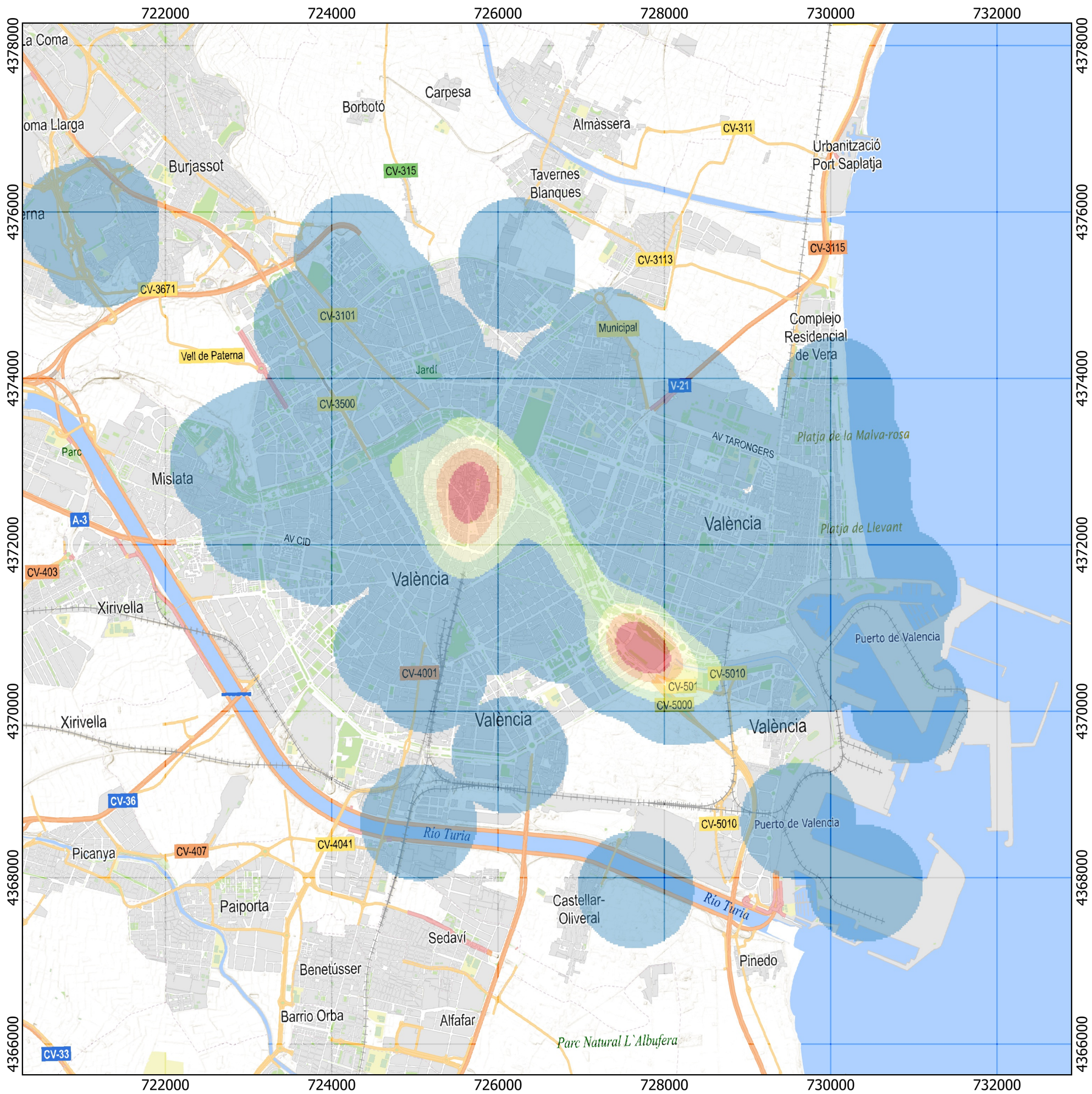


## 10. Cartografía



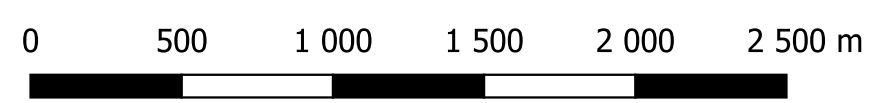
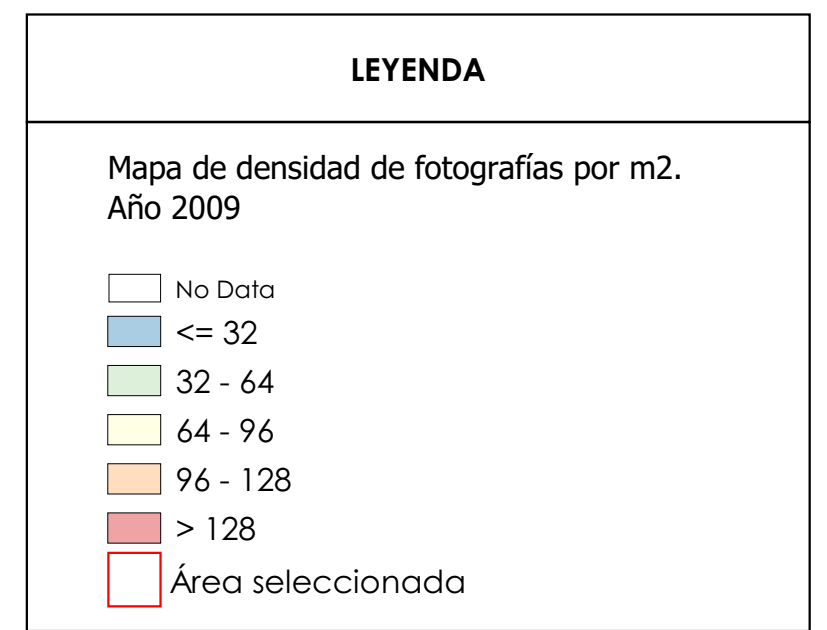
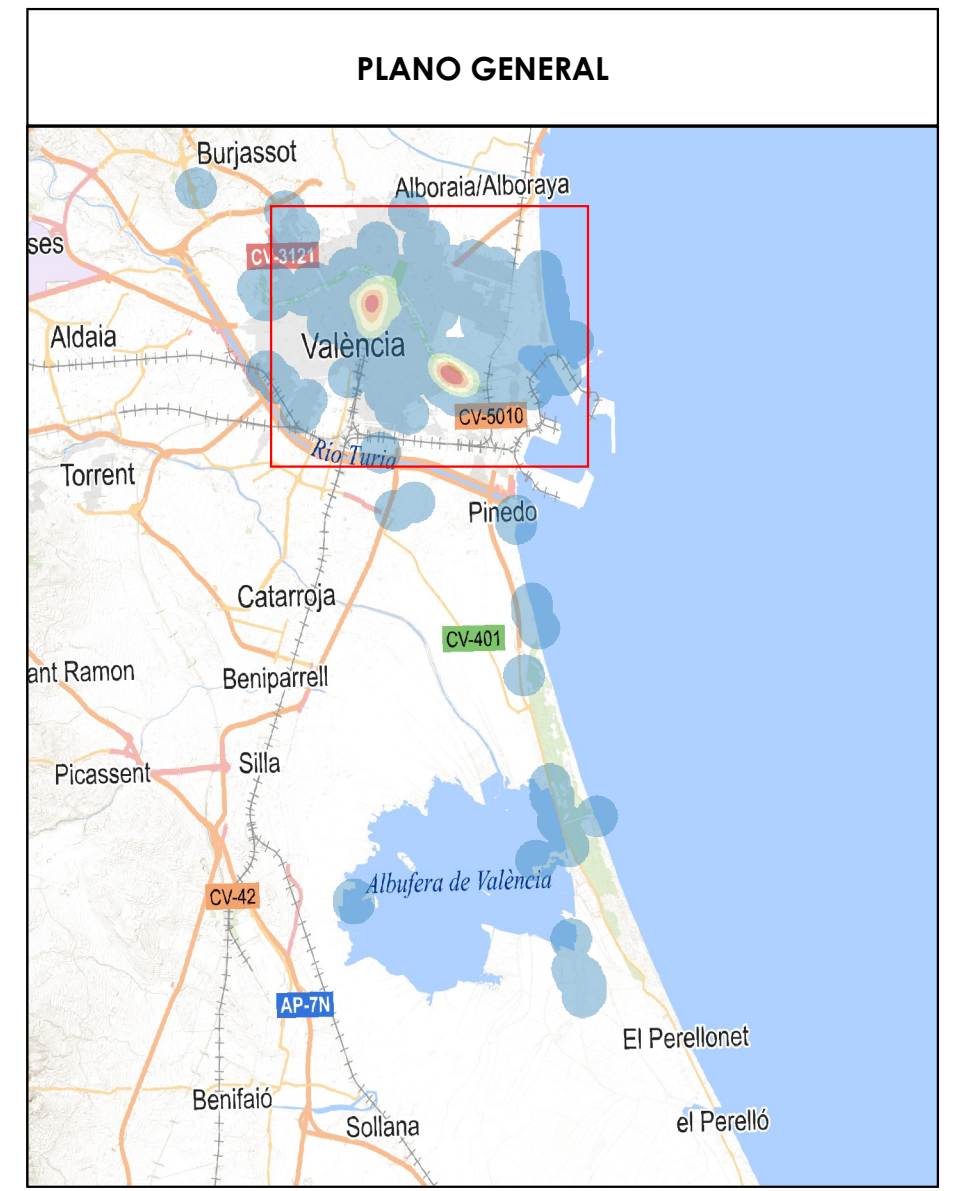
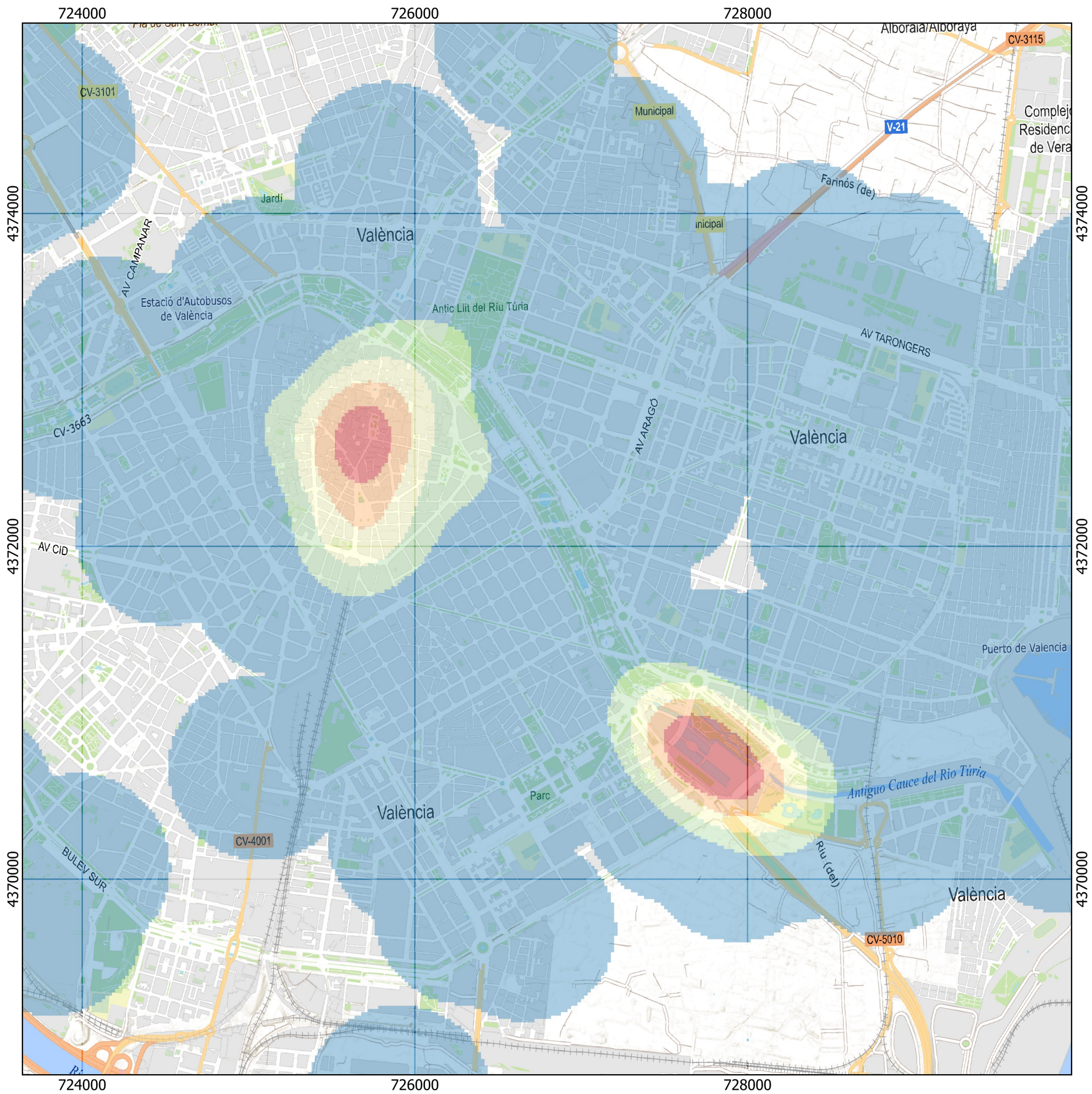
Fuente mapa base: Instituto geográfico nacional (IGN)

	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General	1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA		Detalle	1:25.000		
				Sistema de referencia ETRS89 UTM Zona 30		AUTOR: GLENE CLAVICILLAS CABALTERA	



Fuente mapa base: Instituto geográfico nacional (IGN)

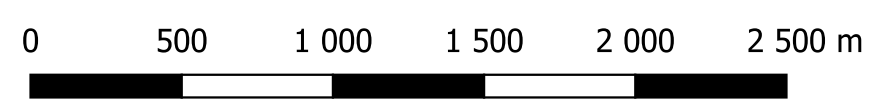
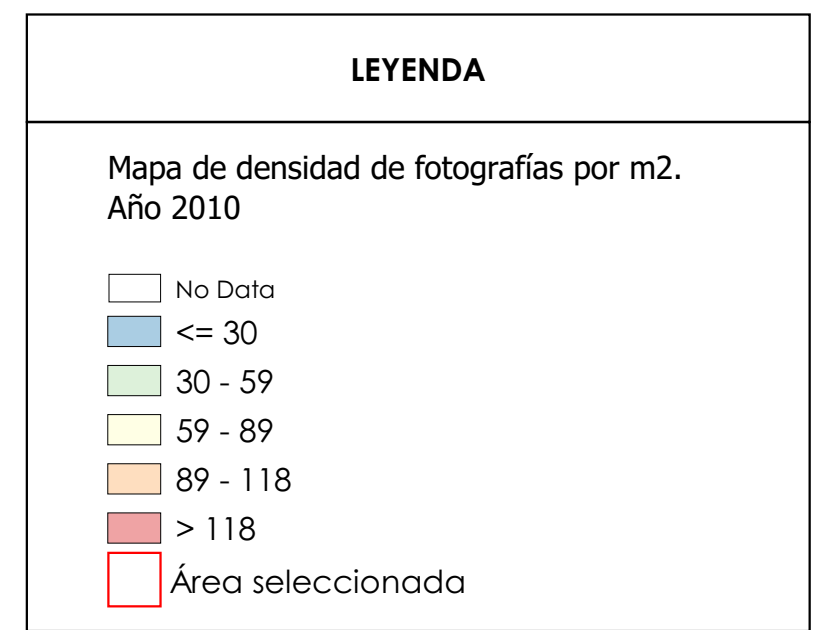
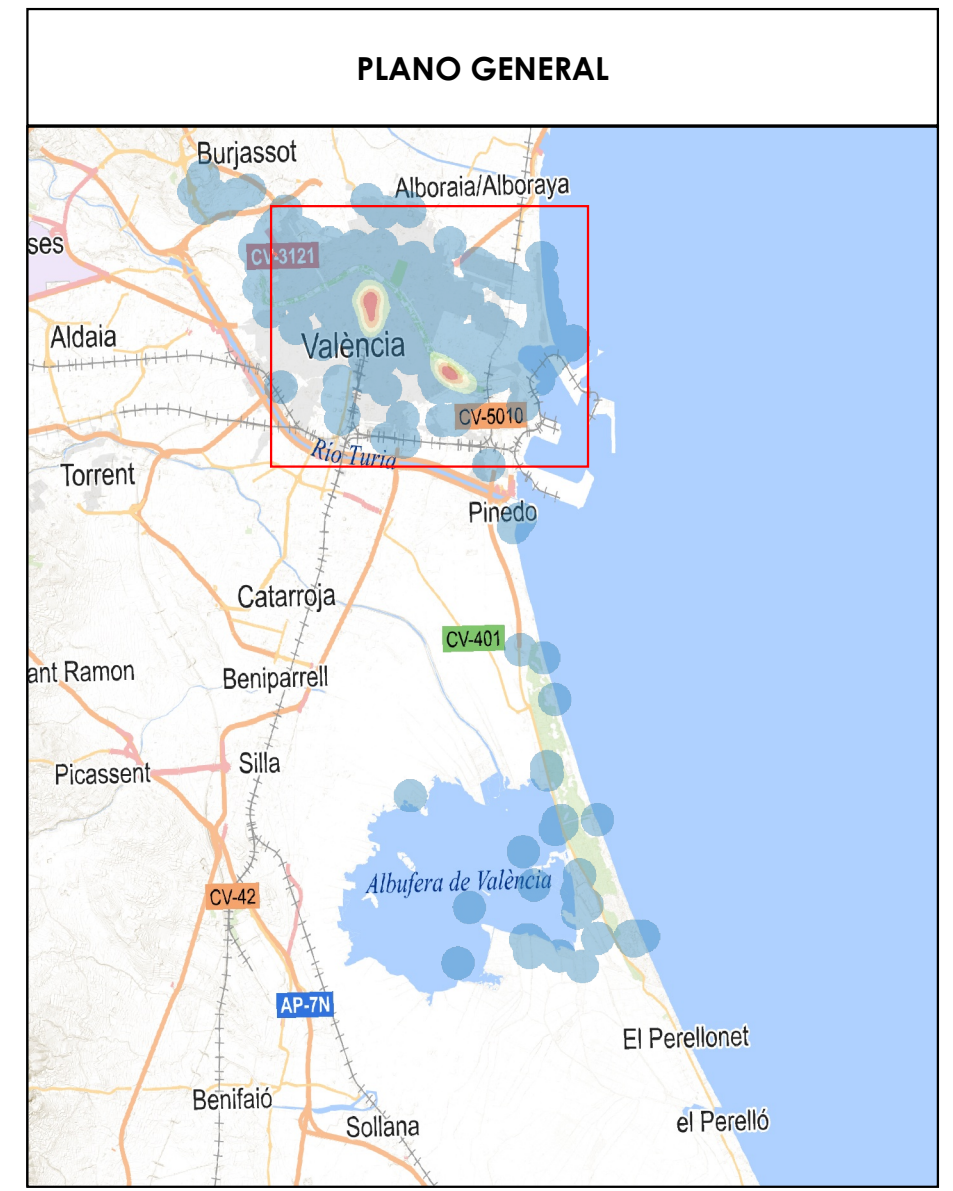
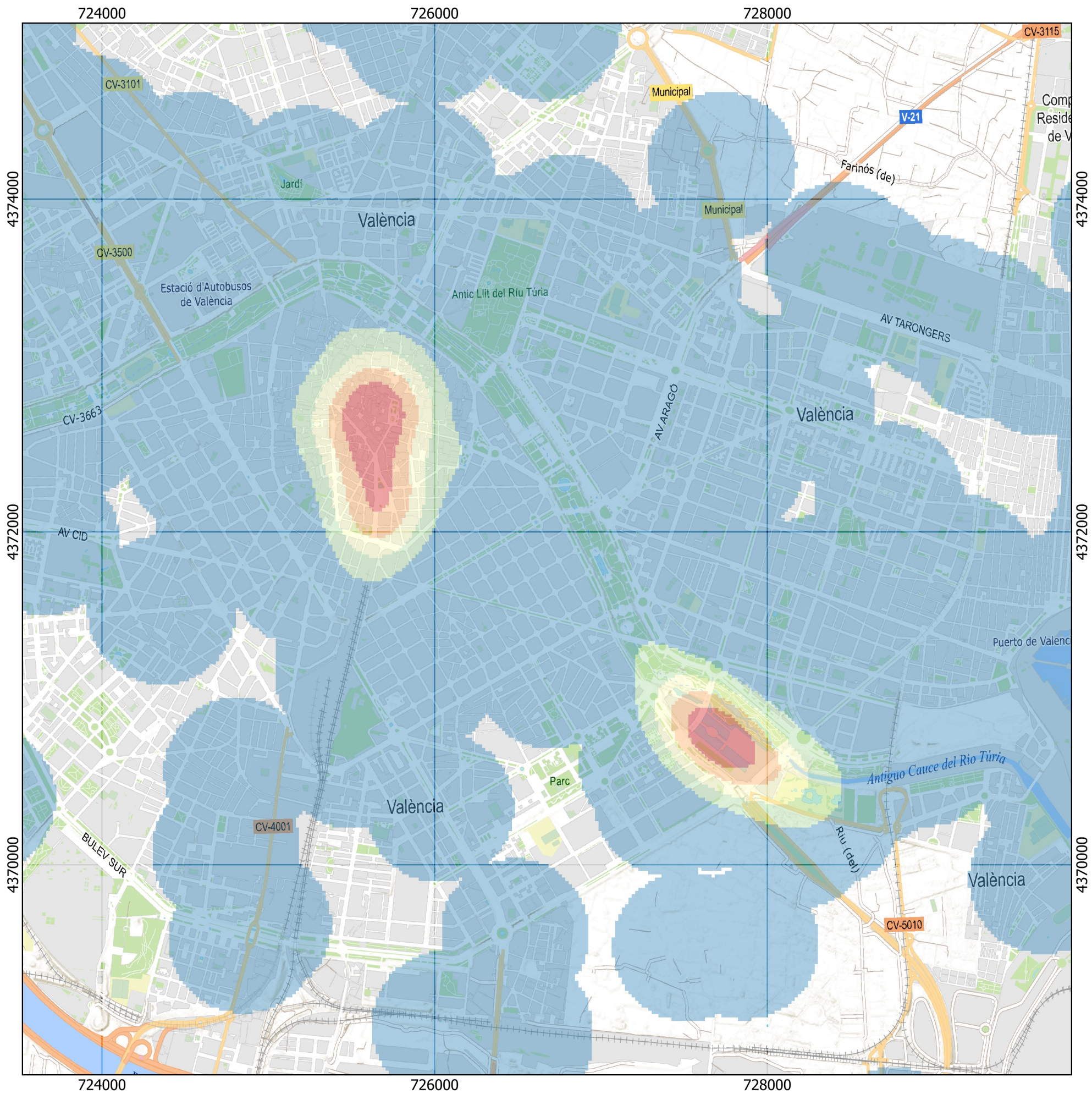
	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General 1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA	Detalle 1:50.000	Sistema de referencia ETRS89 UTM Zona 30		
				<b>AUTOR:</b> GLENE CLAVICILLAS CABALTERA		



Fuente mapa base: Instituto geográfico nacional (IGN)

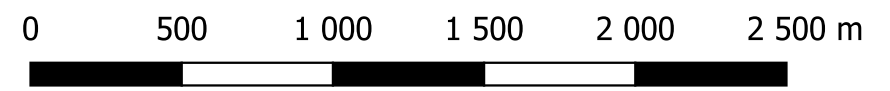
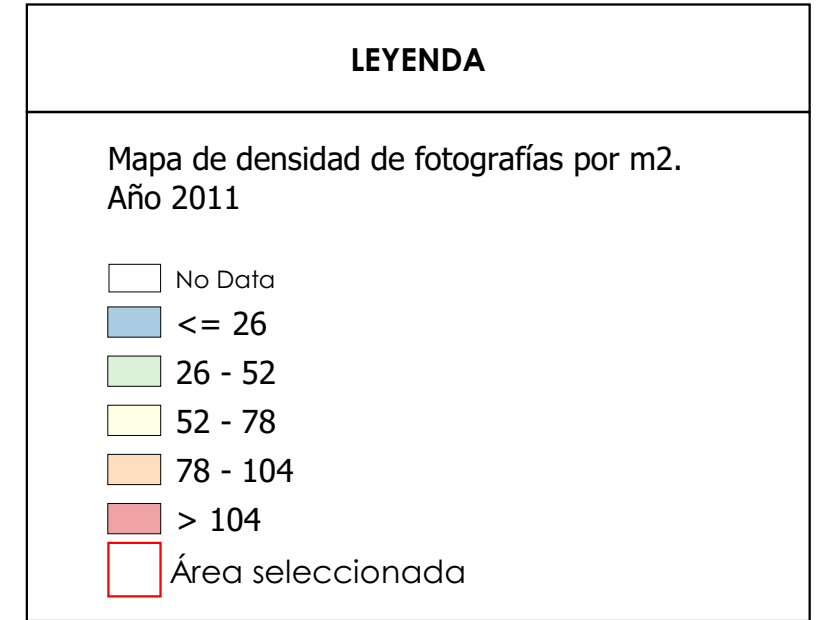
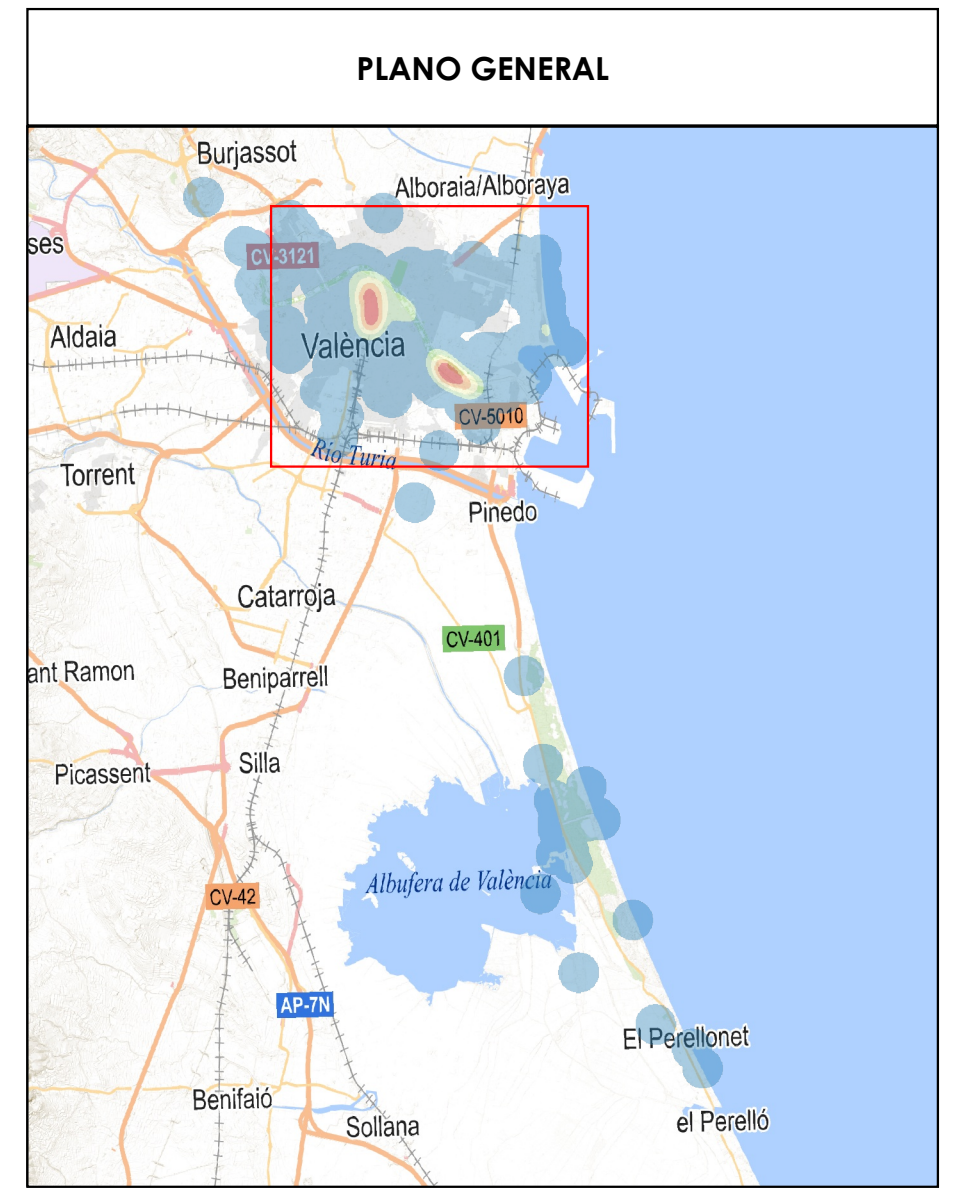
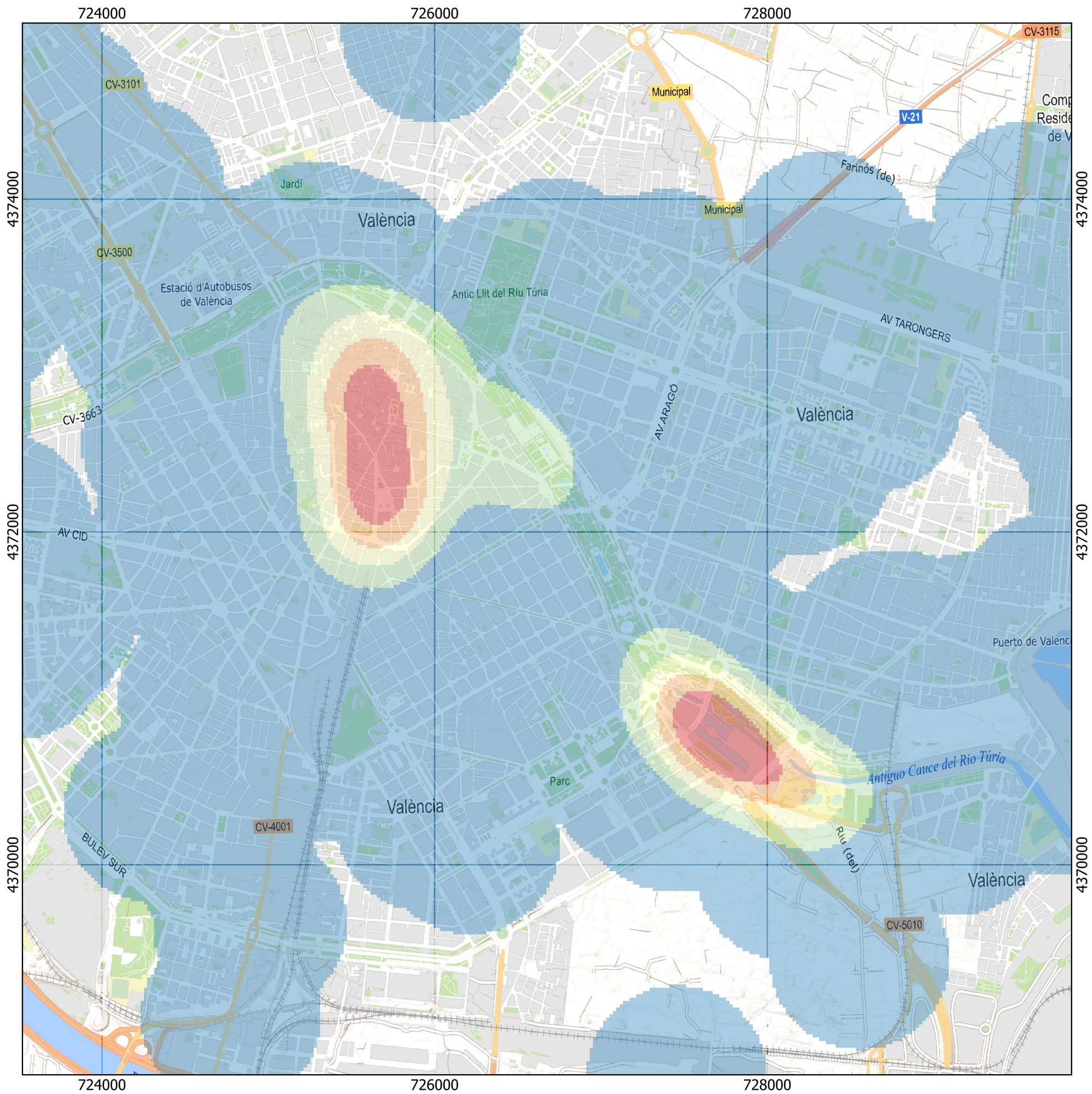
	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General	1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA		Detalle	1:25.000		
			Sistema de referencia		AUTOR: GLENE CLAVICILLAS CABALTERA		
			ETRS89 UTM Zona 30				





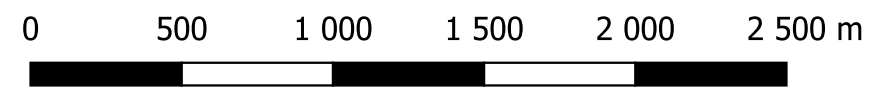
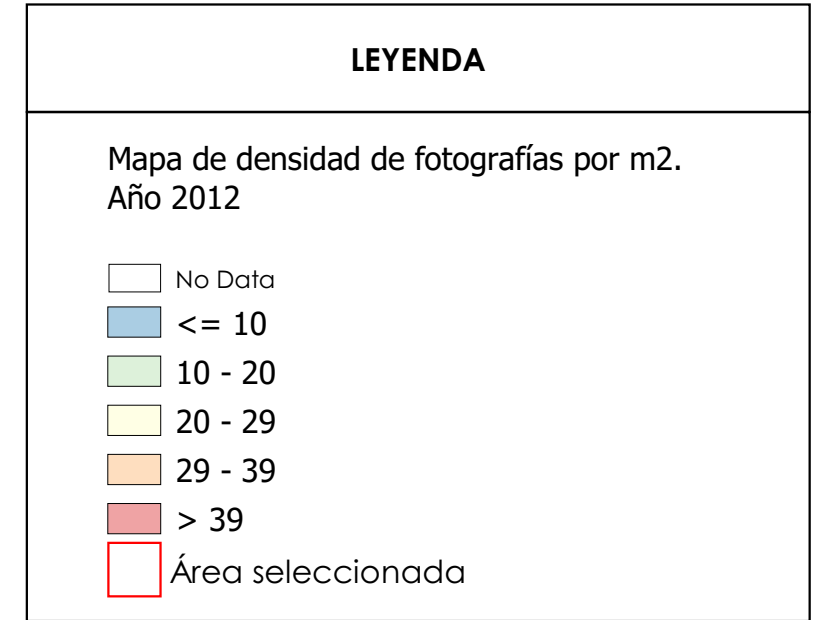
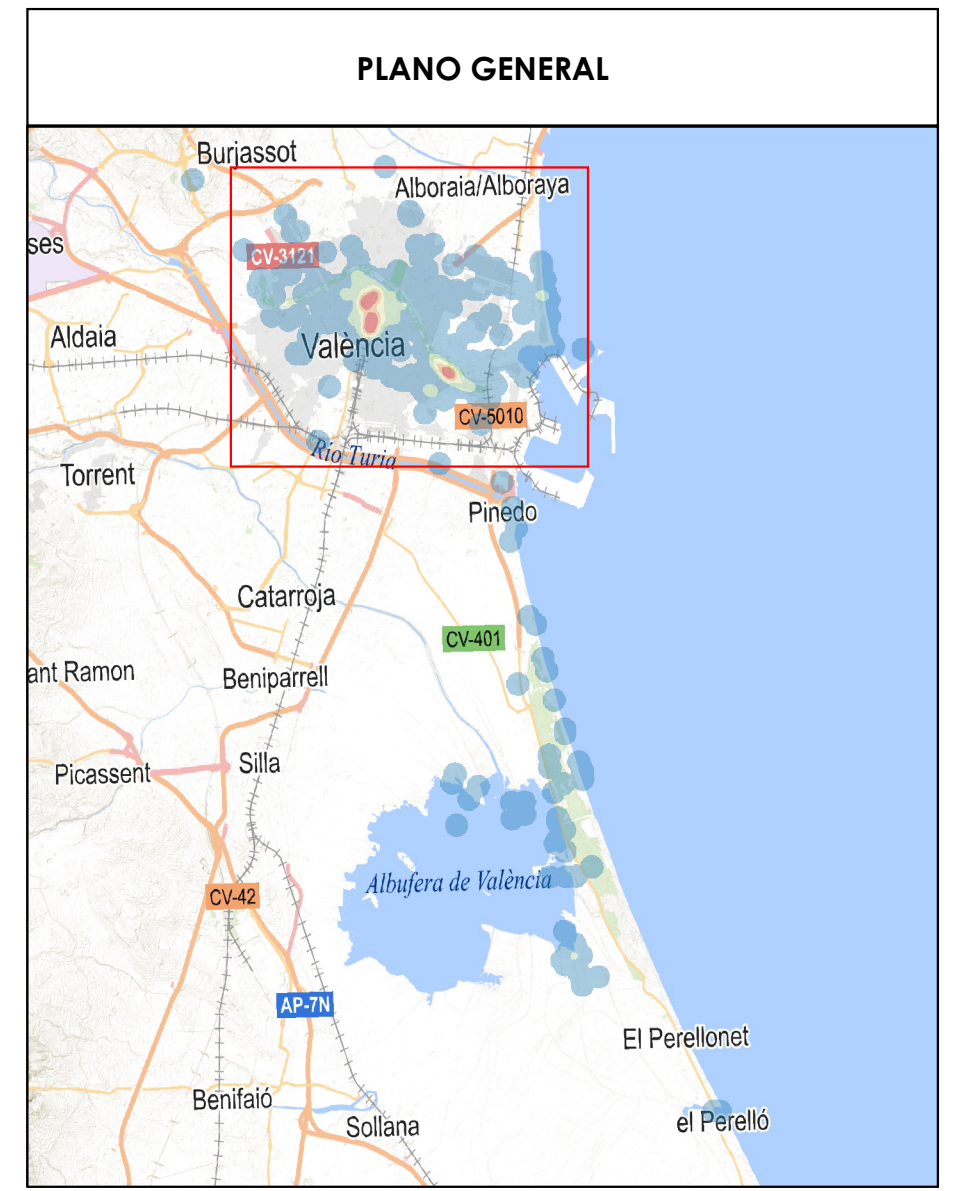
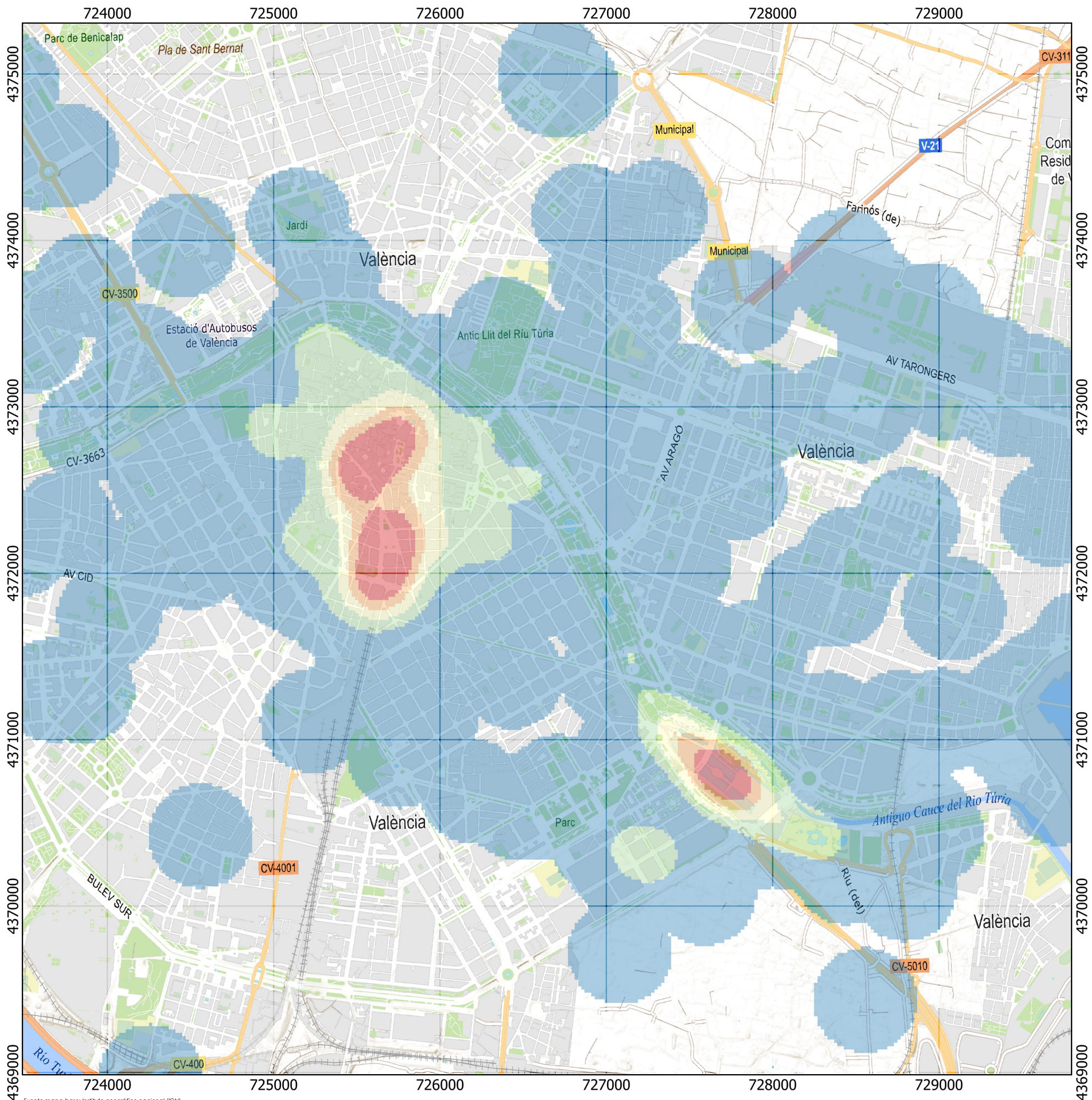
Fuente mapa base: Instituto geográfico nacional (IGN)

	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General	1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA		Detalle	1:25.000		
			Sistema de referencia		AUTOR: GLENE CLAVICILLAS CABALTERA		
			ETRS89 UTM Zona 30				



Fuente mapa base: Instituto geográfico nacional (IGN)

	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General	1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA		Detalle	1:25.000		
			Sistema de referencia		AUTOR: GLENE CLAVICILLAS CABALTERA		
			ETRS89 UTM Zona 30				



Fuente mapa base: Instituto geográfico nacional (IGN)

	<b>PROYECTO:</b>	SEPTIEMBRE 2021	<b>ESCALA</b>	General	1:200.000	<b>TÍTULO DEL PLANO:</b>	<b>Nº PLANO:</b>
	Análisis de la movilidad de los usuarios urbanos basado en la actividad de los usuarios en las redes sociales	VALENCIA, VALENCIA		Detalle	1:25.000		
				Sistema de referencia		AUTOR: GLENE CLAVICILLAS CABALTERA	1/1
				ETRS89 UTM Zona 30			