



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Informática

Optimización multifactorial de rutas peatonales entre dos  
puntos de la ciudad de Valencia

Trabajo Fin de Grado

Grado en Ciencia de Datos

AUTOR/A: Borregón Sacristán, David

Tutor/a: Sebastiá Tarín, Laura

CURSO ACADÉMICO: 2023/2024



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica  
Superior d'Enginyeria  
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica  
Universitat Politècnica de València

# Optimización multifactorial de rutas peatonales entre dos puntos de la ciudad de Valencia

TRABAJO FIN DE GRADO

Grado en Ciencia de Datos

*Autor:* David Borregón Sacristán

*Tutor:* Laura Sebastiá Tarín

Curso 2023-2024



# Resumen

El cálculo de rutas entre dos puntos de una ciudad puede atender a múltiples factores. La minimización de la distancia recorrida suele ser el factor determinante, pero otros elementos, como la proximidad a parques, edificios emblemáticos y zonas de ocio, así como el nivel de ruido, también pueden ser de interés. Este proyecto aprovecha datos geoespaciales abiertos de diversas fuentes para realizar una optimización multifactorial de rutas turísticas peatonales entre dos puntos de la ciudad de Valencia. Para ello, se modela la ciudad como un grafo, tomando esquinas como nodos, y tramos de calles como aristas, a las que se les asignará un peso resultado del cómputo de todos los criterios considerados. Posteriormente, se aplica un algoritmo de determinación del camino más adecuado, buscando proporcionar una experiencia turística más agradable y personalizada.

**Palabras clave:** punto de interés, ruta peatonal, grafo

---

# Resum

El càlcul de rutes entre dos punts d'una ciutat pot atendre múltiples factors. La minimització de la distància recorreguda sol ser el factor determinant, però altres elements, com la proximitat a parcs, edificis emblemàtics i zones d'oci, així com el nivell de soroll, també poden ser d'interés. Aquest projecte aprofita dades geoespaciales obertes de diverses fonts per a realitzar una optimització multifactorial de rutes turístiques per als vianants entre dos punts de la ciutat de València. Per a això, es modela la ciutat com un graf, prenent cantonades com a nodes, i trams de carrers com a arestes, a les quals se'ls assignarà un pes resultat del còmput de tots els criteris considerats. Posteriorment, s'aplica un algorisme de determinació del camí més adequat, buscant proporcionar una experiència turística més agradable i personalitzada.

**Paraules clau:** punt d'interés, ruta per als vianants, graf

---

# Abstract

The calculation of routes between two points in a city can take into account a multitude of factors. The minimization of the distance traveled is typically the primary determining factor, but other elements, such as proximity to parks, emblematic buildings, and leisure areas, as well as noise level, can also be of interest. This project leverages open geospatial data from various sources to perform a multifactor optimization of pedestrian tourist routes between two points in the city of Valencia. To achieve this, the city is represented as a graph, with corners representing nodes and street sections representing edges. These edges are assigned a weight, resulting from the computation of all the criteria considered. Subsequently, an algorithm is applied to determine the most appropriate path, with the objective of providing a more pleasant and personalized tourist experience.

**Key words:** point of interest, pedestrian route, graph

---





# Índice general

---

Índice general	V
Índice de figuras	VII
Índice de tablas	VII
<hr/>	
<b>1 Introducción</b>	<b>1</b>
1.1 Contexto inicial . . . . .	1
1.2 Motivación . . . . .	1
1.3 Objetivos . . . . .	2
1.4 Estructura de la memoria . . . . .	3
<b>2 Estado del arte</b>	<b>5</b>
<b>3 Fundamentos teóricos</b>	<b>7</b>
3.1 Datos abiertos y software libre . . . . .	7
3.1.1 Datos abiertos . . . . .	7
3.1.2 Software libre . . . . .	11
3.2 Sistemas de Información Geográfica . . . . .	11
3.2.1 Datos geoespaciales . . . . .	11
3.2.2 Sistemas de coordenadas . . . . .	12
3.2.3 Herramientas para la manipulación de datos geoespaciales . . . . .	12
3.2.4 Formatos de datos geoespaciales . . . . .	13
3.2.5 Disponibilidad de datos geoespaciales . . . . .	13
3.2.6 Índices espaciales . . . . .	13
3.3 Teoría de grafos . . . . .	14
<b>4 Metodología</b>	<b>17</b>
4.1 Extracción de datos . . . . .	17
4.1.1 OpenStreetMap . . . . .	17
4.2 Transformación de datos . . . . .	20
4.2.1 Reducción de atributos . . . . .	20
4.2.2 Eliminación de datos duplicados . . . . .	20
4.2.3 Normalización de la estructura de datos . . . . .	21
4.2.4 Corrección de valores anómalos y valores faltantes . . . . .	22
4.2.5 Enriquecimiento de datos geoespaciales . . . . .	23
4.2.6 Asignación de pesos a las aristas . . . . .	24
4.2.7 Carga de datos . . . . .	27
4.3 Cálculo de la ruta . . . . .	28
4.3.1 Imputación de direcciones de origen y destino . . . . .	28
4.3.2 Algoritmo de búsqueda del camino . . . . .	30
4.4 Representación de la ruta . . . . .	30
4.4.1 Elección del mapa base . . . . .	30
4.4.2 Extracción de geometrías . . . . .	31
<b>5 Desarrollo</b>	<b>33</b>
5.1 Extracción de datos de la ciudad de Valencia . . . . .	33
5.1.1 Conjuntos de datos extraídos . . . . .	33

5.2	Transformación de los datos extraídos . . . . .	41
5.2.1	Eliminación de datos duplicados . . . . .	41
5.2.2	Imputación de datos faltantes . . . . .	43
5.2.3	Combinación de conjuntos de datos . . . . .	43
5.2.4	Normalización de las variables . . . . .	46
5.2.5	Definición de los perfiles de ruta . . . . .	46
5.2.6	Definición de coeficientes en función del perfil de ruta . . . . .	47
5.2.7	Asignación de pesos . . . . .	47
5.3	Carga de datos . . . . .	47
<b>6</b>	<b>Resultados</b>	<b>49</b>
6.1	Validación de las rutas verdes . . . . .	50
6.2	Validación de las rutas silenciosas . . . . .	51
6.3	Validación de las rutas turísticas . . . . .	53
6.4	Validación de las rutas sociales . . . . .	53
<b>7</b>	<b>Conclusiones</b>	<b>57</b>

---

Apéndice		
<b>A</b>	<b>Relación con el Grado en Ciencia de Datos</b>	<b>61</b>
A.1	Análisis del marco legal y ético . . . . .	61
A.2	Relación con los estudios cursados . . . . .	61
A.3	Relación con los Objetivos de Desarrollo Sostenible (ODS) . . . . .	62
A.3.1	Objetivo 3: Garantizar una vida sana y promover el bienestar para todos en todas las edades . . . . .	62
A.3.2	Objetivo 10: Reducir la desigualdad en y entre los países . . . . .	62
A.3.3	Objetivo 11: Lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles . . . . .	62
A.3.4	Objetivo 12: Garantizar modalidades de consumo y producción sostenibles . . . . .	62

## Índice de figuras

---

3.1	Página de inicio <a href="http://www.openstreetmap.org">www.openstreetmap.org</a> . . . . .	8
3.2	Iconos de los elementos definidos en OSM . . . . .	9
3.3	Página para un conjunto de datos del portal del Ayuntamiento de Valencia . . . . .	10
3.4	Principales entidades utilizadas para la representación geográfica . . . . .	12
4.1	Vista parcial de los jardines y reservas naturales de Valencia recuperados de OSM . . . . .	19
4.2	Cuadros delimitadores de la red peatonal y puntos de interés turístico de Valencia . . . . .	23
4.3	Buffer de 100 metros alrededor de una calle de Valencia. . . . .	24
4.4	Ejemplo de grafo no dirigido y ponderado . . . . .	29
4.5	Fragmento de la ciudad de Valencia estilizado con el mapa base “CartoDB Positron” . . . . .	31
4.6	Ejemplo de ruta peatonal calculada entre dos puntos de la ciudad . . . . .	31
5.1	Visualización sobre mapa del conjunto de la red peatonal de Valencia . . . . .	34
5.2	Visualización sobre mapa del conjunto final sobre monumentos turísticos . . . . .	37
5.3	Visualización sobre mapa del conjunto final sobre zonas verdes . . . . .	38
5.4	Visualización sobre mapa del conjunto final sobre locales sociales . . . . .	39
5.5	Visualización sobre mapa del conjunto final sobre niveles de ruido . . . . .	41
5.6	Ejemplos de niveles de ruido bajo y alto en la ciudad de Valencia . . . . .	41
5.7	Diagrama el proceso de asignación de un nivel de ruido a cada calle . . . . .	45
5.8	Visualización sobre mapa de las calles coloreadas según el nivel de ruido . . . . .	45
5.9	Subgrafo representante de los 2500 primeros nodos de la red y las aristas que los unen . . . . .	48
6.1	Representación de los distintos caminos generados en función del perfil de ruta . . . . .	50
6.2	Ejemplos de caminos generados siguiendo el perfil de ruta verde . . . . .	51
6.3	Ejemplos de caminos generados siguiendo el perfil de ruta silenciosa . . . . .	52
6.4	Ejemplos de caminos generados siguiendo el perfil de ruta turística . . . . .	54
6.5	Ejemplos de caminos generados siguiendo el perfil de ruta social . . . . .	55

## Índice de tablas

---

4.1	Ejemplo de datos descargados de OpenStreetMap a través de OSMnx . . . . .	18
4.2	Nodos de la red peatonal de Valencia, obtenidos a través de OSMnx . . . . .	21
4.3	Aristas de la red peatonal de Valencia, obtenidas a través de OSMnx . . . . .	21

4.4	Ejemplo de datos en formato JSON convertidos a GeoDataFrame . . . . .	22
4.5	Ejemplos de conjuntos de coeficientes . . . . .	25
5.1	Campos que componen el conjunto de datos original sobre monumentos turísticos del Ayuntamiento de Valencia . . . . .	34
5.2	Campos del conjunto de datos sobre monumentos turísticos del Ayunta- miento de Valencia utilizados en el análisis . . . . .	35
5.3	Campos que componen el conjunto de datos original sobre bienes de rele- vancia local del Ayuntamiento de Valencia . . . . .	35
5.4	Campos del conjunto de datos sobre bienes de relevancia local del Ayun- tamiento de Valencia utilizados en el análisis . . . . .	35
5.5	Campos que componen el conjunto de datos original sobre monumentos turísticos de OpenStreetMap . . . . .	36
5.6	Campos del conjunto de datos sobre monumentos turísticos de OpenS- treetMap utilizados en el análisis . . . . .	36
5.7	Extracto del conjunto final sobre monumentos turísticos, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia . . . . .	36
5.8	Campos que componen el conjunto de datos original sobre zonas verdes del Ayuntamiento de Valencia . . . . .	37
5.9	Campos del conjunto de datos sobre zonas verdes del Ayuntamiento de Valencia utilizados en el análisis . . . . .	37
5.10	Campos que componen el conjunto de datos original sobre zonas verdes de OpenStreetMap . . . . .	38
5.11	Campos del conjunto de datos sobre zonas verdes de OpenStreetMap uti- lizados en el análisis . . . . .	38
5.12	Extracto del conjunto final sobre zonas verdes, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia . . . . .	38
5.13	Extracto del conjunto final sobre locales sociales, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia . . . . .	39
5.14	Campos que componen el conjunto de datos original sobre niveles de rui- do del Ayuntamiento de Valencia . . . . .	40
5.15	Campos del conjunto de niveles de ruido del Ayuntamiento de Valencia utilizados en el análisis . . . . .	40
5.16	Cantidad de registros con nombre ausente por conjunto de datos . . . . .	42
5.17	Cantidad de registros con nombre ausente por conjunto de datos tras apli- car transformaciones para la reducción de duplicados . . . . .	42
5.18	Distribución de los valores de anchura de las calles . . . . .	43
5.19	Fragmento del conjunto base donde se combinan los datos extraídos . . . . .	44
5.20	Fragmento del conjunto base tras añadir los datos de ruido . . . . .	44
5.21	Fragmento del conjunto base tras añadir los datos de zonas verdes . . . . .	46
5.22	Fragmento del conjunto base tras añadir los datos de puntos de interés turístico . . . . .	46
5.23	Fragmento del conjunto base tras añadir los datos de puntos de interés social	46
5.24	Fragmento del conjunto base tras normalizar los factores implicados . . . . .	46
5.25	Esquema de coeficientes en función del perfil de ruta . . . . .	47
5.26	Fragmento del conjunto base tras calcular el peso asignado a cada calle . . . . .	47
5.27	Fragmento del conjunto base tras ajustar el peso asignado a cada calle . . . . .	47
A.1	Grado de relación de cada ODS con el proyecto . . . . .	63

---

---

# CAPÍTULO 1

## Introducción

---

### 1.1 Contexto inicial

---

Desde principios del siglo XXI, los programas de navegación urbana han experimentado un auge en su uso. Sus enormes beneficios y el fácil acceso a los mismos (siendo el único requisito estar en posesión de un dispositivo móvil capaz de conectarse a internet) han provocado que aplicaciones como Google Maps [7] encuentren usuarios en cualquier parte de la sociedad. Los individuos gozan de indicaciones para llegar de un punto A a un punto B en cuestión de segundos. Instituciones públicas, como las empresas municipales de transporte, integran estos servicios en sus aplicaciones para recomendar itinerarios de viaje a los ciudadanos. Las empresas privadas de distribución optimizan las rutas de sus transportistas para entregar el mayor número de paquetes en el menor tiempo posible.

Sea cual sea el ámbito de uso, parece que la tendencia es a la reducción de tiempos y distancias en el recorrido. El desarrollo de soluciones de navegación urbana se ha volcado con la optimización de la eficiencia, dejando de lado otros factores que pueden contribuir significativamente a la satisfacción con la ruta propuesta.

### 1.2 Motivación

---

A pesar de este énfasis en la eficiencia, las ciudades modernas están siendo testigo de un movimiento creciente que valora formas alternativas de recorrer los núcleos urbanos y busca la sostenibilidad. Valencia, como una ciudad rica en patrimonio cultural y natural, y recientemente nombrada Capital Verde Europea [2], es el escenario ideal para explorar cómo se pueden optimizar los recorridos peatonales teniendo en cuenta diversos factores que trascienden la mera minimización de la distancia. En palabras de ONU Turismo, el organismo de las Naciones Unidas especializado en turismo, “los destinos urbanos ofrecen un espectro amplio y heterogéneo de experiencias y productos culturales, arquitectónicos, tecnológicos, sociales y naturales para el tiempo libre y los negocios” [29], y por ello, la principal motivación de este trabajo es desarrollar una metodología para generar rutas peatonales que se beneficien de todo el valor que la ciudad puede ofrecer.

La misma entidad lista como deber del turismo sostenible “respetar la autenticidad sociocultural de las comunidades anfitrionas, conservar sus activos culturales y arquitectónicos y sus valores tradicionales, y contribuir al entendimiento y la tolerancia intercultural” [27]. Al priorizar el paso por zonas verdes, edificios emblemáticos y áreas de bajo ruido, se contribuye a este cometido, en línea con los principios del turista y ciudadano

contemporáneo. Además, esta iniciativa aceleraría la adhesión de la ciudad de Valencia al objetivo 11 de los ODS (Objetivos de Desarrollo Sostenible) de las Naciones Unidas, que expresa la necesidad de *“lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles”* [14].

En definitiva, este trabajo pretende cuestionar las nociones actuales en torno a la generación de rutas, y argumentar que una ruta óptima no es necesariamente la más corta o rápida, sino aquella que mejor equilibra los aspectos que el usuario valora, proporcionando una experiencia más enriquecedora y personalizada. Puesto de otra forma: se persigue generar rutas que integren las preferencias del peatón utilizando para ello datos que reflejan el atractivo de la ciudad.

En adición a la navegación y el turismo sostenible, el uso de datos abiertos se sitúa como tercer pilar fundamental de este proyecto, con el fin de demostrar el potencial de estos datos como fuente de información geográfica detallada y punto de partida para la innovación ciudadana.

Por último, el enfoque multifactorial propuesto en este proyecto pretende contribuir a la investigación en el ámbito de la movilidad urbana y el turismo, ofreciendo una metodología que puede ser replicada en otras ciudades.

### 1.3 Objetivos

---

El objetivo general de este proyecto es desarrollar una metodología para la generación de rutas peatonales en la ciudad de Valencia, que utilice datos abiertos y software libre para integrar diversos criterios relevantes para la experiencia del usuario, como la proximidad a puntos de interés y la minimización del ruido ambiental. Para lograrlo, se proponen diversos objetivos específicos:

- **Desarrollar una metodología estandarizada para el proceso de extracción, transformación y carga de los datos (ETL):** una definición adecuada, a la vez que flexible, del modelo de datos utilizado permite adaptar el proyecto al caso práctico de otras ciudades, atendiendo a detalles particulares de la ciudad estudiada en fases posteriores del proyecto.
- **Incorporar la existencia de puntos y otras variables de interés en el cálculo de rutas peatonales:** la proximidad de un punto de interés a un tramo de calle o factores como el nivel de ruido le otorgan un peso a la arista que la representa gráficamente. El uso de un grafo ponderado en el cálculo de una ruta recompensa a los tramos de calle considerados relevantes, logrando así una experiencia ajustada a los parámetros imputados.
- **Utilizar algoritmos de teoría de grafos para el cálculo de rutas entre dos puntos:** si el proceso ETL seguido proporciona un conjunto de datos de calidad, la fase del cálculo de la ruta se puede simplificar recurriendo a algoritmos preexistentes de cálculo del camino más corto en un grafo.

Se propone la utilización de datos geospaciales abiertos y software libre como medio para el cumplimiento de estos objetivos, pues eliminan las dependencias de terceros y aseguran la accesibilidad y reproducibilidad de la metodología. Asimismo, se garantiza que la información utilizada sea transparente y disminuye las barreras de adopción del sistema, facilitando la adaptación a otros contextos urbanos y promoviendo la colaboración y la innovación en el campo de estudio.

---

## 1.4 Estructura de la memoria

---

De este documento se pueden desprender tres grandes bloques de contenido. El primero de ellos explica el contexto en el que se desarrolla el proyecto: su razón de ser y qué carencias se pretenden suplir. Este apartado incluye un breve resumen del estado del arte, mencionando estudios que tienen relación con este, directa o tangencialmente. Se destacará su relación el caso específico de la ciudad de Valencia, pero también iniciativas globales vinculadas a la generación de rutas agradables. Por último, se presentan los conceptos teóricos necesarios para una lectura comprensiva del siguiente bloque, relacionados con los Sistemas de Información Geográfica (SIG) y la teoría de grafos.

El segundo bloque de esta memoria aborda la metodología desarrollada para la creación de un sistema de generación de rutas peatonales entre dos puntos de la ciudad de Valencia basado en la optimización de varios factores. Si bien se enfoca en el caso de la urbe valenciana, se proporcionan los pasos genéricos que debería seguir el proceso de extracción, transformación y carga de los datos (ETL) para obtener conjuntos de datos similares de otras ciudades. Es en este bloque en el que se obtiene un modelo gráfico de la ciudad de Valencia, sobre el que posteriormente se calcularán rutas peatonales entre dos puntos, que se ilustrarán sobre un mapa para ejemplificar los resultados.

En el último bloque del estudio se integran los apartados de resultados y conclusiones. Se evaluarán los resultados obtenidos, tanto las rutas calculadas como la adecuación de la metodología desarrollada y su viabilidad para aplicarse a otras ciudades.





---

---

## CAPÍTULO 2

# Estado del arte

---

El interés reciente por los sistemas de navegación ha supuesto una expansión de su campo de estudio. Si bien son los productos comerciales los que reciben la mayor parte de la atención de los usuarios, un número creciente de proyectos académicos y de investigación de este ámbito se enfoca en la optimización de rutas con objetivos más afines a la salud y el desarrollo social.

El presente proyecto se incluye en un conjunto de trabajos que buscan explotar el potencial de los datos abiertos, como los proporcionados por OpenStreetMap (OSM), para tener un impacto positivo en la sociedad y cubrir carencias detectadas en los sistemas de cálculo de rutas actuales.

En el estudio “Towards a Landmark-Based Pedestrian Navigation Service Using OSM Data” [24], los autores emplean puntos de referencia (*landmarks*) extraídos de OpenStreetMap para enriquecer y aclarar las indicaciones que guían a los usuarios. Este enfoque se basa en la premisa de que los peatones pueden orientarse mejor mediante referencias visuales significativas en su entorno, en lugar de depender únicamente de indicaciones direccionales y de distancia. Utilizando los datos de OSM, el sistema identifica y prioriza puntos de referencia relevantes, algo que es particularmente útil en áreas urbanas densamente construidas, donde los puntos de referencia pueden ser más reconocibles que las direcciones convencionales. El mismo trabajo aborda la completitud de los datos de OSM para evaluar su aptitud como fuente de datos única en este tipo de proyectos.

Por su parte, los desarrolladores del programa web descrito en “Journey Planners Can Promote Active, Healthy and Sustainable Urban Travel” [21] abogan por la generación de rutas peatonales y ciclistas con menos contaminación atmosférica y acústica en la región de Helsinki capital, argumentando que la incorporación de estos factores en los planificadores de rutas puede tener un impacto significativo en la calidad de vida urbana y en la salud pública. Este esfuerzo a nivel local es una prueba de concepto que, con la adecuada recepción y financiación podría ampliar su alcance de aplicación. Cabe destacar que esta herramienta también se basa en el uso de datos abiertos, aprovechando tanto la información de OpenStreetMap (OSM) para la red de calles, como fuentes de datos abiertas nacionales para obtener información sobre calidad del aire y niveles de ruido.

Otro proyecto titulado “A System for Generating Customized Pleasant Pedestrian Routes Based on OpenStreetMap Data” [15], desarrollado también a partir de datos extraídos exclusivamente de OSM, recoge la satisfacción de los usuarios al probar un sistema que personaliza rutas según las preferencias indicadas (en qué medida les gustaría que la ruta atravesara zonas verdes, como parques, o lugares sociales, como restaurantes), construido sobre datos extraídos de OSM. Los resultados obtenidos confirman que, a pesar de que las rutas son ligeramente más largas, presentan mayor valor social y natu-

ral. Además, se concluye que este sistema otorga a los peatones la sensación de control, valorada positivamente.

En “Solving the tourist trip design problem with time windows and variable profit using incremental local search” [13], los autores abordan el problema de diseño de viajes turísticos con ventanas de tiempo y beneficio variable mediante búsqueda local incremental. Si bien el presente proyecto no aborda explícitamente estos aspectos, la metodología diseñada para el cálculo de rutas entre dos puntos podría integrarse en su planteamiento, de manera que el camino entre los dos puntos de interés para el viajero se adapte a sus preferencias.

Aunque la lista continúa, es inevitable observar puntos comunes entre todos ellos. Es importante destacar que, aunque estas iniciativas han mostrado resultados prometedores en sus respectivos ámbitos académicos y locales, ninguna de ellas ha alcanzado aún una implementación comercial global. Por otro lado, generalmente, estos proyectos utilizan datos abiertos, como los proporcionados por OpenStreetMap, lo que subraya su compromiso con los principios de accesibilidad y transparencia en la información geográfica.

El proyecto descrito a lo largo de esta memoria comparte la filosofía *open source* en cuanto que se sirve de datos abiertos como los de OpenStreetMap y el Ayuntamiento de Valencia para la generación de rutas peatonales. No obstante, se distingue por su orientación específica hacia la mejora de la experiencia turística, considerando el atractivo turístico de los trayectos. Además, la creación de una metodología robusta y replicable facilita su aplicación en diversas ciudades del mundo, aumentando su potencial para convertirse en una solución ampliamente adoptada en el ámbito global.

---

---

## CAPÍTULO 3

# Fundamentos teóricos

---

En este capítulo se presenta el marco teórico que contextualiza el presente proyecto y proporciona los conceptos fundamentales necesarios para una comprensión adecuada del mismo.

Por su relación directa con el proyecto, se incide en tres grandes conceptos clave: los datos abiertos y el software libre, los Sistemas de Información Geográfica (SIG) y la teoría de grafos. Estos conceptos son esenciales para entender cómo se modela y analiza la red urbana de Valencia utilizando datos geoespaciales y técnicas de análisis de redes.

### 3.1 Datos abiertos y software libre

---

#### 3.1.1. Datos abiertos

Los datos abiertos son conjuntos de información que están disponibles públicamente para que cualquier persona pueda acceder, usar y compartir. La elección de fuentes de datos abiertos para esta proyecto está fundamentada en las múltiples ventajas que ofrecen, algunas de las cuales se listan a continuación:

- **Acceso libre y gratuito:** algo crucial para proyectos que disponen de un presupuesto limitado o inexistente.
- **Actualización y mantenimiento:** la relevancia de un proyecto está ligada a su antigüedad y precisión. Los datos de OSM son mantenidos y actualizados por una comunidad global de colaboradores. De igual manera, los portales de datos abiertos municipales suelen actualizar sus conjuntos de datos periódicamente (diariamente, en muchos casos), reflejando cambios recientes en la infraestructura y servicios de la ciudad.
- **Diversidad y riqueza de datos:** útil cuando se busca “cruzar” datos de distinta naturaleza. Muchas fuentes de datos propietarias ofrecen conjuntos de datos aislados, inconexos o que siguen distintos formatos. El repertorio de datos de OSM incluye no solo la red vial, sino también puntos de interés, información sobre el uso del suelo, y detalles sobre la infraestructura urbana. Por su parte, los portales de datos abiertos de municipios (como el del Ayuntamiento de Valencia), proporcionan una vista multidimensional de una ciudad específica, pudiendo obtener el nivel de ruido, la densidad del tráfico y la existencia de líneas de transporte público para un mismo punto determinado.

- **Facilidad de descarga y uso:** por diseño, las plataformas de datos abiertos facilitan la descarga y uso de sus datos. No es extraño que se desarrollen librerías que permiten la extracción, manipulación y análisis eficiente de datos geospaciales. Los portales municipales de datos abiertos municipales suelen ofrecer datos en formatos estándar como CSV, JSON, o GeoJSON, y ponen a disposición varios métodos de descarga, incluyendo APIs.
- **Licencias abiertas y libertad de uso:** los datos abiertos están generalmente sujetos a licencias abiertas, como la Open Database License (ODbL) [16] de OSM, que permiten el uso, modificación y distribución de los datos con pocas restricciones.
- **Reproducibilidad:** la accesibilidad y disponibilidad de los datos abiertos fomentan la reutilización y el intercambio de información. Además, al no estar mantenidos por un solo organismo, su probabilidad de perdurar en el tiempo es mayor.

Es importante destacar que la disponibilidad de datos está asociada, normalmente, a la relevancia de la ubicación de estudio. Valencia, afortunadamente, es una ciudad conocida internacionalmente, cuyas características geográficas y puntos de interés han sido documentados exhaustivamente, posibilitando este proyecto. Antes de embarcarse en un estudio similar, es indispensable consultar la existencia de los datos y recursos necesarios.

Del mismo modo, no todos los municipios ponen a disposición de sus ciudadanos un portal de datos abiertos. Incluso algunos que sí lo hacen, no ofrecen muchos conjuntos de datos y a veces estos no son de gran relevancia o calidad.

## OpenStreetMap

OpenStreetMap (OSM) es un proyecto colaborativo para crear mapas editables y libres. Sin embargo, en lugar de los mapas en sí, los datos generados son considerados su principal valor. Para un proyecto como este, los datos de interés son las calles y intersecciones de la red peatonal (su ubicación y atributos como el nombre o la anchura), así como los puntos definidos de interés (la ubicación y atributos de lugares de ocio, edificios de interés turístico, zonas verdes...).

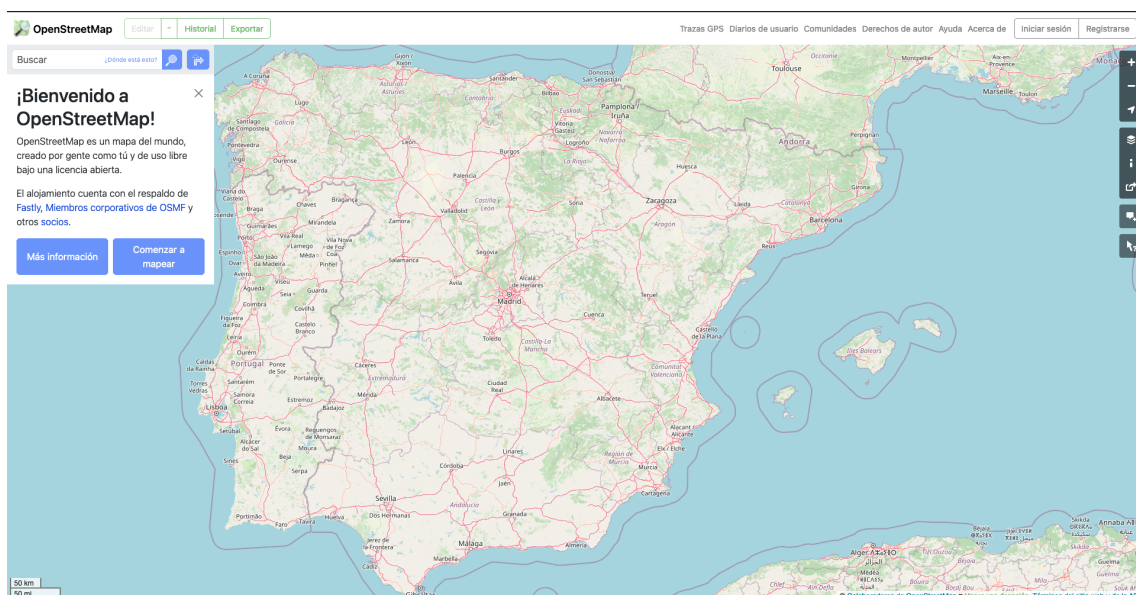


Figura 3.1: Página de inicio [www.openstreetmap.org](http://www.openstreetmap.org)

OpenStreetMap [18] es un excelente ejemplo de Información Geográfica Voluntaria (VGI), un término acuñado por Goodchild [6] que se refiere al caso geográfico del fenómeno web de contenido generado por el usuario, por el que cualquiera puede aportar información sobre el mundo geográfico. El crecimiento en la generación y distribución de datos de OSM en los últimos años se debe, precisamente, al aumento de la participación comunitaria global y al reconocimiento de su valor por parte de empresas, gobiernos y organizaciones, que contribuyen activamente con datos y utilizan la plataforma para diversos fines, desde navegación hasta planificación urbana y respuesta a desastres [12]. Además, su naturaleza abierta y colaborativa ha facilitado la expansión y actualización constante de su base de datos geospaciales a nivel mundial.

OpenStreetMap ofrece un amplia variedad de opciones para descargar sus datos. A fin de minimizar el número de herramientas a utilizar, se propone extraer datos de OSM través de la librería de Python OSMnx [1]. Esta librería permite obtener los nodos y aristas que conforman la red de una ciudad, pudiendo especificarse el tipo de red que sea desea, que en este caso es la peatonal (caminable), dada la naturaleza del proyecto. Además de esta red, que conforma la base de la cartografía de OSM, hay una amplia variedad de datos disponibles que describen diferentes aspectos geográficos y características del entorno, como edificios y estructuras, áreas naturales y parques, infraestructura de transporte, comercios y servicios, y límites administrativos, entre otros. Sea lo que sea que representen, todos los objetos en OSM toman una de las siguientes geometrías: nodo, vía o relación.

En OpenStreetMap, los objetos geográficos se representan mediante un sistema de etiquetas que permite describir de manera detallada las características y propiedades de cada elemento. Cada uno de estos objetos puede tener múltiples etiquetas, las cuales se componen de pares clave-valor. Las claves describen la característica general del objeto, como "highway" para carreteras o "building" para edificaciones, mientras que los valores especifican el tipo concreto, como "residential" para carreteras residenciales o "yes" para indicar que un objeto es un edificio. La *wiki* de OpenStreetMap ofrece una breve descripción, y algunos ejemplos, para las principales etiquetas en uso [11]. Este sistema flexible y extensible permite a los colaboradores de OSM añadir una gran variedad de detalles a los mapas, proporcionando una representación rica y precisa del mundo real. Aunque la vasta cantidad de etiquetas disponibles puede resultar abrumadora, afortunadamente existen herramientas como TagFinder, que facilitan la tarea proporcionando una lista de etiquetas relacionadas con un término buscado [19].

Los elementos principales en OSM son nodos, vías y relaciones. Los nodos son puntos definidos por su latitud, longitud e identificador, y se utilizan para representar elementos simples como árboles, semáforos o monumentos. Las vías son secuencias ordenadas de nodos que definen elementos lineales como carreteras, ríos y caminos. Las relaciones, por su parte, agrupan nodos, vías y otras relaciones para representar estructuras más complejas como rutas de transporte público o límites administrativos. La figura 3.4 muestra los iconos asociados a los distintos tipos de elementos de OSM.

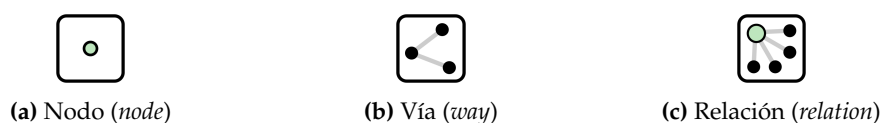


Figura 3.2: Iconos de los elementos definidos en OSM

## Portales municipales de datos abiertos

Los portales municipales de datos abiertos son plataformas digitales gestionadas por ayuntamientos (aunque también son ofrecidos por algunos gobiernos nacionales) que ponen a disposición del público una gama de datos generados por las administraciones locales. Estos portales permiten a los ciudadanos, investigadores, desarrolladores y empresas acceder a información diversa, que puede incluir desde estadísticas demográficas y económicas hasta datos sobre infraestructura, medio ambiente, transporte y servicios públicos. Al facilitar el acceso a estos datos, los portales municipales promueven la transparencia, la participación ciudadana y la innovación, permitiendo que los usuarios utilicen la información para crear aplicaciones, realizar análisis o contribuir al desarrollo de políticas públicas más informadas y efectivas.

La estructura de estos sitios web se suele componer de un catálogo, donde se listan los conjuntos de datos (a menudo clasificados por tipología o palabras clave, por ejemplo, “urbanismo”, “transporte” o “medioambiente”), y de una página dedicada a cada conjunto, que incluye atributos del mismo (identificador, una descripción y la frecuencia de actualización, entre otros) y e instrucciones sobre las opciones de descarga. Adicionalmente, algunas páginas generan visualizaciones simples, como tablas, gráficos de columnas y mapas (si se trata de datos geoespaciales) para un análisis exploratorio básico, como se muestra en la figura 3.3, donde se puede ver una página del portal de datos abiertos del Ayuntamiento de Valencia [22].

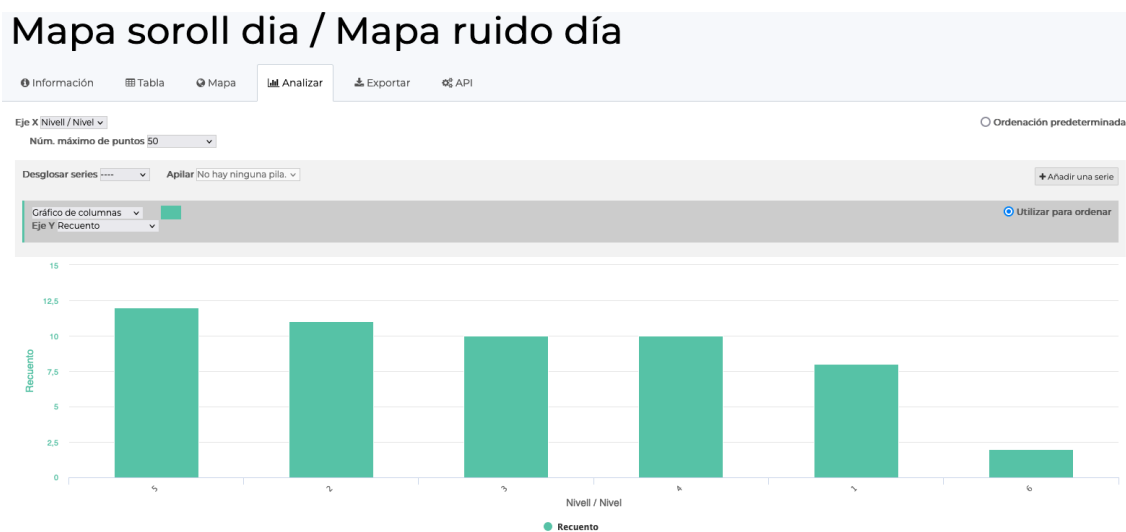


Figura 3.3: Página para un conjunto de datos del portal del Ayuntamiento de Valencia

Normalmente, los portales ofrecen la posibilidad de descargar los datos a través de la interfaz gráfica. No obstante, para quien desea hacerlo de forma programática, existe una sección en estas páginas web que instruye sobre cómo utilizar su API (*Application Programming Interface*) para comunicarnos con el servicio a través del protocolo HTTP y descargar los datos deseados. A pesar de que el uso de APIs REST es un método de descarga ampliamente preferido para la descarga de datos no estáticos (esto es, datos que se actualizan con algún tipo de frecuencia), a continuación se discuten algunos puntos a favor y en contra del uso de esta interfaz de comunicación en este contexto:

### ■ Ventajas

- **Actualización en tiempo real:** al usar una API, podemos acceder a la última versión publicada de los datos. No obstante, esta ventaja puede no suscitar

demasiado interés en proyectos como este, ya que el ritmo de modificación de datos urbanísticos no es tan alto como para compensar la configuración del sistema.

- **Automatización:** la naturaleza programática de las APIs permiten la automatización del proceso de extracción de datos, al delegar el trabajo en la ejecución de scripts que realizan peticiones.
- **Filtrado y selección de datos:** la mayoría de APIs ofrecen parámetros para filtrar y seleccionar únicamente los datos relevantes. Esto permite optimizar el volumen de datos descargados, haciendo las consultas más eficientes y manejables. Cuando los datos de interés son de gran volumen, esta capacidad resulta indispensable en los pasos iniciales de un proyecto, donde se busca experimentar con los datos y no se requiere el conjunto entero.

#### ■ Inconvenientes

- **Complejidad en la implementación:** el uso de APIs requiere de unos conocimientos programación básicos, pudiendo esto suponer una barrera de entrada para muchas personas.
- **Límites de tasa:** muchas APIs imponen límites en la cantidad de solicitudes que se pueden realizar en un determinado período de tiempo. Esto puede ser un inconveniente si se necesita acceder a grandes volúmenes de datos rápidamente.
- **Cambios en la API:** las APIs pueden sufrir cambios en su estructura o en sus puntos de acceso (*endpoints*), lo que puede requerir actualizaciones en el código del proyecto para mantener la funcionalidad y, consecuentemente, interrupciones en el acceso a los datos.

### 3.1.2. Software libre

El software libre se refiere a programas que se distribuyen con licencias que permiten a los usuarios ejecutarlos, estudiarlos, modificarlos y redistribuirlos libremente. En el ámbito de la programación en Python, muchas de las herramientas utilizadas, como librerías para el procesamiento y análisis de datos, son de software libre. Estas librerías, como Pandas [26] y GeoPandas [10], son colecciones de funciones y métodos que simplifican tareas complejas, como la manipulación de estructuras de datos. Pandas, por ejemplo, permite trabajar con DataFrames [20], que son estructuras bidimensionales que almacenan datos en formato de tabla, similar a una hoja de cálculo, con filas y columnas etiquetadas. Un GeoDataFrame [5] es una extensión de un DataFrame que añade soporte para datos geoespaciales, permitiendo almacenar y operar con geometrías como puntos, líneas y polígonos, lo que es esencial para análisis espaciales y mapas.

## 3.2 Sistemas de Información Geográfica

---

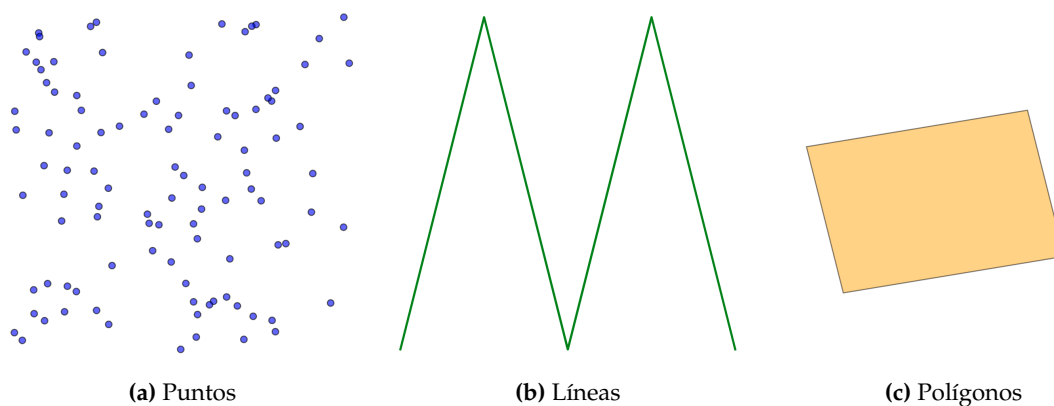
Los Sistemas de Información Geográfica (SIG) son herramientas fundamentales para la recopilación, gestión, análisis y visualización de datos geoespaciales.

### 3.2.1. Datos geoespaciales

Los datos geoespaciales son aquellos con un componente geográfico asociado, es decir, que puede ser localizados en un sistema de coordenadas espaciales. Pueden incluir



información sobre características físicas de la Tierra, como la ubicación de calles, edificios, ríos y montañas, así como datos socioeconómicos, ambientales y de infraestructura. Los datos geospaciales se representan comúnmente mediante puntos, líneas y polígonos que modelan entidades geográficas y sus relaciones espaciales.



**Figura 3.4:** Principales entidades utilizadas para la representación geográfica

En el contexto de los sistemas de información geográfica, un concepto clave es el de *buffer*. Un *buffer* es una zona de proximidad que se crea alrededor de una entidad geográfica, como un punto, una línea o un polígono, definida a una distancia especificada desde la entidad original. En relación a este proyecto, un *buffer* se puede interpretar como un área de influencia alrededor de una calle.

### 3.2.2. Sistemas de coordenadas

Los sistemas de coordenadas son marcos de referencia utilizados para identificar la ubicación de puntos en el espacio, ya sea en la superficie terrestre o en un mapa. Permiten describir precisamente las posiciones geográficas utilizando coordenadas como latitud, longitud y, en algunos casos, altura. Existen diferentes tipos de sistemas de coordenadas, entre los que se incluyen los sistemas de coordenadas geográficas (que utilizan grados de latitud y longitud) y los sistemas de coordenadas proyectadas (que proyectan la superficie curva de la Tierra en un plano).

Un estándar ampliamente utilizado para definir sistemas de coordenadas es el EPSG (European Petroleum Survey Group), que asigna códigos únicos a cada sistema de coordenadas y proyección. Estos códigos EPSG permiten a los usuarios especificar el sistema de coordenadas utilizado en sus datos geospaciales, garantizando la coherencia y la correcta interpretación de las ubicaciones.

Es importante destacar que no todos los sistemas de coordenadas miden las distancias en metros. Algunos, como el EPSG:4326 (WGS 84), que es un sistema de coordenadas geográficas, utilizan grados como unidades, lo que no representa distancias lineales directamente. Otros sistemas, como el EPSG:3857 (Web Mercator), que es un sistema de coordenadas proyectadas, sí miden las distancias en metros, facilitando cálculos más intuitivos y precisos en análisis espaciales. En última instancia, la elección del sistema de coordenadas depende de su finalidad.

### 3.2.3. Herramientas para la manipulación de datos geospaciales

Múltiples son las herramientas para manipular y analizar datos geospaciales. Los programas de escritorio, como ArcGIS y QGIS, se encuentran entre los más utilizados,

debido a su gran capacidad de procesamiento y análisis de datos geográficos. En un plano más humilde, aunque a la vez más accesible, se encuentran librerías como GeoPandas, Shapely y Pyproj, que permiten realizar operaciones geoespaciales en Python.

Para este proyecto, se emplea la herramienta OSMnx, un paquete de Python orientado a la descarga y manipulación de datos de OpenStreetMap. OSMnx permite convertir estos datos en estructuras de datos geoespaciales manejables mediante GeoDataFrames, con aspecto tabular, facilitando su análisis y visualización.

#### 3.2.4. Formatos de datos geoespaciales

El paso del tiempo ha dado lugar a diversos estándares para el almacenamiento de los datos geoespaciales, por lo que actualmente se puede escoger entre una variedad de formatos, cada uno con sus propias ventajas y aplicaciones. A continuación se listan algunos de los más comunes:

- **Shapefile:** un formato popular y ampliamente utilizado que puede almacenar geometrías y atributos. Es compatible con la mayoría de los SIG de escritorio.
- **GeoJSON:** un formato basado en JSON que es ideal para la web debido a su legibilidad y facilidad de uso.
- **KML (Keyhole Markup Language):** un formato utilizado para representar datos geoespaciales en aplicaciones como Google Earth.
- **GPKG (GeoPackage):** un formato abierto basado en SQLite que permite almacenar datos geoespaciales y atributos en un solo archivo.
- **TIFF (Tagged Image File Format):** comúnmente usado para imágenes ráster geoespaciales, como fotografías aéreas y mapas de elevación.

#### 3.2.5. Disponibilidad de datos geoespaciales

La proliferación de fuentes de datos abiertas y accesibles ha aumentado significativamente la disponibilidad de datos geoespaciales, y ha mejorado su accesibilidad. En este sentido, es difícil no reconocer la presencia de OpenStreetMap como una de las fuentes más importantes de datos geoespaciales libres, ofreciendo una base de datos global y colaborativa (se mantiene gracias a su comunidad de voluntarios) de información geográfica. Por otro lado, muchas instituciones gubernamentales y organizaciones internacionales proporcionan libre acceso a datos geoespaciales a través de portales y APIs.

En este proyecto, se utiliza la API de OSM a través de OSMnx para obtener datos de la red de calles de Valencia. Se integran datos complementarios sobre puntos de interés que el Ayuntamiento de Valencia pone a disposición a través su API.

#### 3.2.6. Índices espaciales

Los índices espaciales son estructuras de datos diseñadas para optimizar las consultas y operaciones sobre datos geográficos en los sistemas de información geográfica (SIG). En el mundo SIG, donde se manejan grandes volúmenes de información espacial, como puntos, líneas y polígonos, estos índices son cruciales para realizar operaciones de búsqueda, intersección y proximidad de manera eficiente. Sin un índice espacial adecuado, las consultas podrían volverse prohibitivamente lentas, especialmente cuando se trabaja con bases de datos geoespaciales de gran tamaño.

Entre los diferentes tipos de índices espaciales, uno de los más utilizados es el R-tree (del inglés “Rectangle Tree”). Un R-tree es una estructura de datos espacial utilizada para indexar información multidimensional, como datos geográficos. Es especialmente efectiva para consultar objetos espaciales, como rectángulos o polígonos, permitiendo una recuperación eficiente de objetos que intersectan con una región de consulta dada. La eficiencia de un R-tree radica en su capacidad para reducir el espacio de búsqueda mediante el uso de cajas delimitadoras. Cada nodo en un R-tree contiene una caja delimitadora que envuelve todos los objetos o nodos hijos en ese nodo, lo que permite descartar rápidamente grandes porciones de los datos que no intersectan con la región de consulta.

STRTree es una implementación específica de un R-tree optimizada para la indexación y consulta espacial. Basada en el algoritmo Sort-Tile-Recursive (STR), STRTree se utiliza ampliamente en bibliotecas como Shapely y GeoPandas en Python para operaciones espaciales. STRTree suele ser más eficiente que los R-trees genéricos, particularmente para conjuntos de datos grandes y cuando se realizan consultas que se benefician de la clasificación.

### 3.3 Teoría de grafos

---

La teoría de grafos es una rama de las matemáticas que estudia las propiedades y aplicaciones de los grafos. Un grafo es una estructura compuesta por nodos (también llamados vértices) y aristas que conectan pares de nodos. Esta teoría proporciona un marco sólido para modelar relaciones y conexiones en diversos contextos, incluyendo redes de transporte, redes sociales y redes de comunicación.

Este proyecto emplea la teoría de grafos como marco para modelar la red peatonal de la ciudad de Valencia. Los nodos representan intersecciones, mientras que las aristas representan segmentos de calles que conectan estos nodos. A estas aristas se les puede asignar un peso, convirtiendo la representación en un grafo ponderado. Esta modelización será de utilidad, por ejemplo, para el cálculo de rutas con mayor valor mediante la aplicación de algoritmos preexistentes.

A fin de evitar confusión en los subsecuentes capítulos, se definen los siguientes términos:

- **Peso:** el peso es un valor numérico asignado a una arista en un grafo, que representa el “costo”, “distancia” o “valor” asociado con transitar entre los nodos que conecta. Los pesos pueden ser positivos o negativos y se utilizan en algoritmos para determinar rutas óptimas.
- **Valor de la ruta:** el valor de la ruta es la suma de los pesos de todas las aristas que componen una ruta específica en un grafo. Dependiendo del contexto, se puede buscar maximizar o minimizar este valor para encontrar la ruta más favorable según el criterio definido (como menor costo o mayor beneficio).
- **Ciclo:** un ciclo en un grafo es una secuencia de aristas que comienza y termina en el mismo nodo, sin repetir ninguna arista. Es una ruta cerrada dentro del grafo.
- **Ciclo negativo:** un ciclo negativo es un ciclo en un grafo donde la suma de los pesos de las aristas que lo componen es negativa. Esto puede causar problemas en algoritmos que buscan el camino más corto, ya que puede llevar a bucles indefinidos en la búsqueda de una solución.

OSMnx utiliza NetworkX —una biblioteca de Python con funcionalidades de grafos que permite la creación, manipulación y análisis de estos, sean simples o complejos—

---

internamente para convertir los datos de OSM en grafos y aplicar técnicas avanzadas de análisis de redes. Entre las funcionalidades de grafos que ofrece se incluyen el cálculo de rutas y métricas de la red representada.



---

---

## CAPÍTULO 4

# Metodología

---

Valencia es la urbe tomada como referencia en este proyecto, pero los procesos explicados a lo largo de esta memoria posibilitan la aplicación del sistema desarrollado a otros casos prácticos. Cabe recordar que uno de los objetivos listados en este proyecto es la creación de una metodología que sirva como documentación para impulsar iniciativas similares. De acuerdo con este propósito, este capítulo detalla los procesos básicos a los que se someten los datos geoespaciales para acabar conformando un generador de rutas entre dos puntos de una ciudad, que atiende a múltiples factores.

Como la mayoría de proyectos en el campo de la ciencia de datos —independiente de su temática—, este comienza con un proceso conocido ETL, por sus siglas en inglés *Extract, Transform, Load* (Extracción, Transformación y Carga), por el cual se integran datos provenientes de una o varias fuentes con el fin de almacenarlos en un repositorio único. Este proceso se compone de tres etapas, que varían en su implementación en función de cuál sea el objetivo final. A continuación se explican las tres etapas y cómo aplican a este proyecto.

### 4.1 Extracción de datos

---

La fase de extracción implica la recolección de datos de diferentes fuentes. Conviene señalar que, a lo largo de este proyecto, se trabaja con dos categorías de datos: los referentes a la red urbana (las calles e intersecciones que componen la red peatonal de la ciudad) y los referentes a factores y puntos de interés que afecta a dicha red (zonas de ruido, locales de ocio, zonas verdes, monumentos turísticos, zonas).

Para el caso de Valencia, la estructura de calles e intersecciones se obtiene a través de OpenStreetMap. Los datos sobre ruido y puntos de interés se toman tanto de OpenStreetMap como del portal de datos abiertos del Ayuntamiento de Valencia, por lo que se hacen necesarias operaciones para combinar ambos conjuntos.

#### 4.1.1. OpenStreetMap

Al extraer datos de OpenStreetMap (OSM), es crucial considerar el nombre de la región especificado en la consulta, ya que puede tener implicaciones significativas en los resultados obtenidos. Por ejemplo, “Valencia, Spain” no devuelve necesariamente los mismos datos que “Valencia, Comunitat Valenciana, Spain”, pues la primera puede referirse a la provincia de Valencia, no a la ciudad de Valencia. Antes de proceder con la descarga, es fundamental verificar que el parámetro introducido coincida exactamente con la ubicación de interés para la cual se desean obtener los datos y que este sea geocodificable. Se

observa que el término ‘Valencia’ puede generar ambigüedad; por lo tanto, se opta por utilizar la designación completa ‘Valencia, Comunitat Valenciana, Spain’ para asegurar una mayor precisión en la selección de la región deseada.

Para obtener datos de OSM a través de Python, se explotan principalmente dos funciones del paquete NetworkX:

- `graph_from_place`: para descargar y crear un grafo del lugar o lugares ingresados como parámetro. Este grafo se convierte posteriormente a un formato tabular por medio de la función `graph_to_gdfs`, que devuelve un `GeoDataFrame` para las aristas (calles) y otro para los nodos (intersecciones) que representan la red urbana.
- `features_from_place`: para descargar en formato tabular (`GeoDataFrame`) los datos coincidentes con las propiedades y lugar (o lugares) especificados como parámetro.

A continuación se muestra un ejemplo de cómo utilizar el sistema de etiquetas de OSM y OSMnx para acceder solamente a los datos que se necesitan o se quieren incluir en el proyecto. En este caso, se quiere obtener información sobre jardines y reservas naturales en la ciudad de Valencia:

```
1 import osmnx as ox
2
3 place = 'Valencia, Comunitat Valenciana, Spain'
4 tags = {'leisure': ['garden', 'nature_reserve']}
5 features = ox.features_from_place(place, tags)
6 features.head()
```

El anterior código devuelve un `GeoDataFrame`. Por conveniencia, la tabla 4.1 muestra un extracto del resultado, limitando las columnas mostradas a las más relevantes.

	osmid	name	geometry	leisure
0	6843975152	El Parterre	POINT (-0.37040 39.47211)	garden
1	10259087	Jardín de la Generalitat	POLYGON ((-0.37573 39.4...	garden
2	209610562	Parc Natural de l'Albufera	POLYGON ((-0.23338 39.1...	nature_reserve
3	2169637	Marjal de Rafalell i Vistabella	POLYGON ((-0.29405 39.5...	nature_reserve
4	15974831	Mediterrànies	MULTIPOLYGON (((-0.38682 39.4...	garden

**Tabla 4.1:** Ejemplo de datos descargados de OpenStreetMap a través de OSMnx

Como se puede observar, los elementos coincidentes con las etiquetas especificadas pueden ser nodos, vías o relaciones, que tomarán la geometría de punto (*POINT*), polígono (*POLYGON*) o multipolígono (*MULTIPOLYGON*), respectivamente. Esto también se puede verificar gráficamente, observando la figura 4.1.

### Portal de datos abiertos del Ayuntamiento de Valencia

Como ya se ha comentado, la descarga de datos a través de este portal puede ser manual o programática. Tras evaluar las ventajas y desventajas del uso de APIs, se decide optar por este enfoque para generalizar el proceso de descarga de datos. Esta elección evita la dependencia de una versión específica del portal del Ayuntamiento Valencia, el cual podría ser modificado en el futuro. Además, permite ilustrar el uso de esta tecnología, que es ampliamente adoptada por la mayoría de los portales de datos abiertos a nivel mundial. Por otra parte, un código bien estructurado en módulos y que emplea funciones contribuye a una mayor versatilidad y facilita una rápida adaptación mediante el uso de parámetros. Para este caso, se define una función que lanza una petición al endpoint de la API especificado y guarda la respuesta en un archivo de texto del mismo nombre.





```
'type': 'Point'},  
'properties': {}},  
'geo_point_2d': {'lon': -0.37899494446702714, 'lat': 39.477910849714426}}}
```

## 4.2 Transformación de datos

---

Una vez extraídos los datos necesarios, el siguiente paso lógico es comprobar su estado y realizar las transformaciones oportunas. Los datos suelen estar en formatos diferentes y pueden contener inconsistencias, duplicados y errores. La fase de transformación se encarga de depurar los datos para mejorar su calidad y establecer una coherencia, haciéndolos útiles para el futuro análisis. A lo largo de esta subsección se comentan algunas transformaciones comunes en proyectos con datos geospaciales.

### 4.2.1. Reducción de atributos

A menudo se utilizan conjuntos de datos de distintas fuentes, como en este caso, lo que obliga a adoptar un estándar de formato y estructura para poder integrarlos. Partiendo del principio de que estos datos se representarán sobre un mapa, se puede establecer que un campo indispensable en cualquiera de los conjuntos debe ser aquel con la ubicación geográfica de los puntos o áreas de interés. Algunos atributos, como el nombre o el identificador del elemento, pueden ser también de utilidad. Se recomienda desechar el resto de campos, si no sirven un propósito particular.

### 4.2.2. Eliminación de datos duplicados

Al integrar datos de varias fuentes, especialmente si estas representan elementos de la misma naturaleza, es frecuente encontrar duplicidades. La redundancia puede deteriorar la calidad de un conjunto de datos, inflar innecesariamente su tamaño e introducir sesgos en los análisis y resultados, llevando a conclusiones erróneas. Por ejemplo, si en la generación de una ruta, la inclusión de una arista u otra depende de un índice calculado teniendo en cuenta el mismo dato dos o más veces, el peso asignado no refleja la realidad fielmente.

En el contexto de este proyecto, conviene comprobar la existencia de puntos o áreas de interés repetidas. Para ello, se puede comparar el nombre o la ubicación geográfica de las mismas. La comparación espacial es, sin duda, más costosa computacionalmente (cuánto más puntos conformen un polígono, más operaciones se requerirán, por ejemplo). Además, su desarrollo programático puede resultar más complejo, al requerir la creación de buffers (ya que, con casi total seguridad, las entidades repetidas se encontrarán próximas pero no exactamente en la misma ubicación, debido a fallos de precisión en la imputación). Por otro lado, habría que tener en cuenta las diferentes geometrías, obteniendo un punto representativo de los polígonos con el que comparar.

Por conveniencia y simplicidad, se recomienda recurrir a la comparación del nombre de la entidad, si este atributo está disponible. Si no dispone de él, se sugiere valorar su eliminación del conjunto. Al fin y al cabo, esta carencia se puede entender como una falta de relevancia (normalmente, los lugares reconocidos están extensamente documentados), lo que lleva a descartar la entidad como punto de interés. No obstante, conviene realizar un análisis exploratorio básico para determinar el porcentaje de registros faltos de nombre, y si este es alto, optar por eliminar algunos, en lugar de todos, para evitar empobrecer el conjunto de datos.

Si el nombre sí está disponible, bastaría con reducirlo a minúsculas para homogeneizar el formato y comparar los textos. Como un mismo punto puede recibir distintos nombres según la fuente de la que provenga, la comparación se puede basar en un umbral de similitud. Así, se determinaría que nombres como “Parc de l’Albufera” y “Parque Natural de la Albufera” hacen referencia a la misma entidad, a pesar de tener denominaciones distintas.

### 4.2.3. Normalización de la estructura de datos

Puesto que para el desarrollo de este proyecto se hace un uso intensivo de OSMnx, una buena práctica es moldear los datos de forma que tengan una estructura compatible con las funciones de este paquete. Un GeoDataFrame parece la elección ideal, al tratarse de una estructura tabular con con la capacidad de albergar, analizar y visualizar geometrías como puntos, líneas y polígonos.

OSMnx facilita la conversión a GeoDataFrame. Para los datos redes, como la peatonal, se puede utilizar la función “graph\_to\_gdfs”, que devuelve el grafo asociado en forma de dos GeoDataFrames: uno para los nodos y otro para las aristas.

```

1 import osmnx as ox
2
3 place = 'Valencia, Comunitat Valenciana, Spain'
4 graph = ox.graph_from_place(place_name)
5 nodes, edges = ox.graph_to_gdfs(graph)
6 nodes.head()

```

Para mayor claridad, la tabla 4.2 muestra el resultado del código anterior, un extracto del conjunto de nodos, en forma de tabla simple.

osmid	y	x	highway	street_count	ref	geometry
25767290	39.470348	-0.375915	crossing	4	NaN	POINT (-0.37592 39.47035))
25767290	39.469146	-0.375500	crossing	4	NaN	POINT (-0.37550 39.46915)
25767290	39.468287	-0.375265	NaN	4	NaN	POINT (-0.37527 39.46829)
25767290	39.467130	-0.375038	NaN	3	NaN	POINT (-0.37504 39.46713)
25767290	39.467780	-0.377841	NaN	3	NaN	POINT (-0.37784 39.46778)

**Tabla 4.2:** Nodos de la red peatonal de Valencia, obtenidos a través de OSMnx

Como se ha comentado anteriormente, la variable imprescindible es la geometría en sí. Otras, como el número de calles que intersectan en ese punto (“street\_count”), pueden ser también útiles para posteriores análisis. La información redundante, como las coordenadas “x” e “y”, puede ser descartada.

El conjunto de aristas incluye muchas más variables, pero en la tabla 4.3 se observan las consideradas más relevantes.

u	v	osmid	highway	oneway	length	lanes	name	geometry
25767290	11420770018	803049226	footway	False	8.795	NaN	NaN	LINestring (-0.37...
25767290	11420770019	803049226	footway	False	4.275	NaN	NaN	LINestring (-0.37...
25767290	11420770020	803466250	residential	True	3.947	3	Plaça de l’Ajuntament	LINestring (-0.37...
25767291	2213108120	583027120	footway	False	5.906	NaN	NaN	LINestring (-0.37...
25767291	9853663631	[5098786, 61720119]	residential	True	47.283	1	[Carrer de Roger de Llòria...	LINestring (-0.37...

**Tabla 4.3:** Aristas de la red peatonal de Valencia, obtenidas a través de OSMnx

Los datos de otras fuentes (como portales de datos abiertos), si se han obtenido mediante API, generalmente están formateados como JSON (JavaScript Object Notation) [9], un formato de texto ligero ampliamente adoptado en la transferencia de datos porque permite representar complejas estructuras anidadas de manera clara y comprensible.

No obstante, convertirlos es sencillo gracias a la función “GeoDataFrame” de la librería GeoPandas. Se propone una solución que reduce el proceso a tres pasos básicos: leer el archivo de texto, guardar los campos deseados y pasarlos como parámetro a la función mencionada. A continuación se muestra un ejemplo utilizando el fichero de monumentos turísticos.

```

1 import ast
2 import geopandas as gpd
3 from shapely.geometry import shape
4
5 with open('monumentos-turisticos.txt', 'r') as file:
6     data = file.read()
7
8 tourist_monuments = ast.literal_eval(data)
9
10 features = []
11
12 for result in tourist_monuments:
13     geometry = result["geo_shape"]["geometry"]
14     properties = {
15         "name": result["nombre"]
16     }
17     features.append({
18         "geometry": shape(geometry),
19         **properties
20     })
21
22 tourist_monuments_gdf = gpd.GeoDataFrame(features)
23 tourist_monuments_gdf.head()

```

El código utilizado en el ejemplo produce un GeoDataFrame con los campos de geometría y nombre, como se muestra en la tabla 4.4.

	<b>geometry</b>	<b>name</b>
0	POINT (-0.37899 39.47791)	PALAU DELS MARTÍNEZ DE RAGA
1	POINT (-0.37934 39.47647)	PALAU DELS QUEIXAL O DELS TRÉNOR
2	POINT (-0.36266 39.47345)	PALAU MUNICIPAL DE L'EXPOSICIÓ
3	POINT (-0.37072 39.47327)	CASA NATALÍCIA DE SANT VICENT FERRER - EL POUET
4	POINT (-0.37589 39.45430)	PARRÒQUIA DE SANTA CECÍLIA

**Tabla 4.4:** Ejemplo de datos en formato JSON convertidos a GeoDataFrame

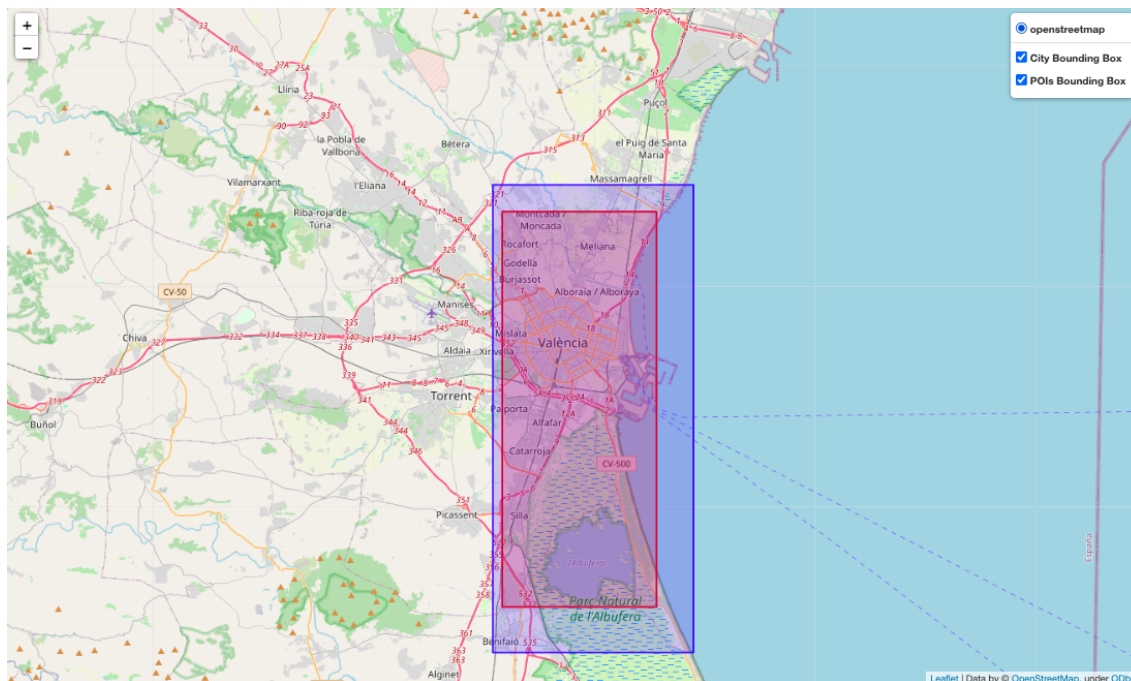
Tener todos los datos formateados con estructuras semejantes facilita la comprensión y agiliza la posterior labor de análisis.

#### 4.2.4. Corrección de valores anómalos y valores faltantes

Al tratarse de datos geospaciales, la variable que más concierne analizar en busca de valores anómalos es la geometría. Se recomienda, pues, realizar una inspección visual de los datos —tanto los de la red peatonal (calles e intersecciones) como de los puntos de interés (ubicación de zonas verdes, locales de ocio, etc.)— representándolos sobre un mapa.

El método “explore” de un GeoDataFrame facilita su visualización sobre un mapa. No obstante, para graficar varios GeoDataFrames simultáneamente, se recurre al paquete Folium [4], que subyace tras esta función. Un mapa de Folium permite representar múltiples capas y ofrece control sobre estas.

Disponer tantas geometrías sobre un mismo mapa puede ser computacionalmente costoso, y corre el riesgo de ensuciar la visualización. Por tanto, graficar los cuadros delimitadores de cada capa y limitarse a una comprobación aproximada (verificar que, más o menos, todas las geometrías caen dentro de la ciudad), resulta más razonable. La figura 4.2 ejemplifica el resultado plasmando la red peatonal (“City Bounding Box”) y los puntos de interés turístico (“POIs Bounding Box”), confirmando que, efectivamente, no hay valores anómalos (al ser solo el área de la ciudad de Valencia la cubierta por los cuadros delimitadores).



**Figura 4.2:** Cuadros delimitadores de la red peatonal y puntos de interés turístico de Valencia

Una variable que podría presentar valores faltantes en el conjunto de datos de la red de calles es la anchura de las mismas, un campo que se incluye por defecto en los datos de OpenStreetMap (OSM). En este caso, una solución sencilla para abordar la ausencia de estos datos consiste en sustituir los valores faltantes por un valor de 1, asumiendo de este modo que la calle tiene la anchura mínima correspondiente a un carril.

#### 4.2.5. Enriquecimiento de datos geoespaciales

Los datos existentes pueden ser manipulados y cruzados para integrar información adicional que proporciona contexto y relevancia. En este caso, se proponen dos técnicas para enriquecer los datos geoespaciales: la creación de buffers alrededor de las calles y la asignación de niveles de ruido a cada calle.

La creación de buffers permite identificar el área circundante a cada calle y las características cercanas que la afectan. Al crear un buffer alrededor de las calles, se puede, por ejemplo, asignar puntos de interés a estas áreas, facilitando la identificación de servicios y atractivos próximos a cada tramo de la red peatonal. Se recomienda realizar una unión espacial de dos conjuntos, el que contiene los buffers y el que contiene los puntos de interés, basándose en el solapamiento de sus geometrías. La figura 4.3 ejemplifica el área bajo análisis al crear un buffer de 100 metros alrededor de una calle cualquiera. La determinación del alcance del buffer es subjetiva y se basa en la estimación del radio dentro del cual se considera que una calle puede verse afectada.



la ruta óptima. Este cálculo se basa en una fórmula básica que combina de manera lineal los productos de coeficientes específicos por el valor de cada variable considerada.

### Fórmula básica

La fórmula básica para calcular el peso  $w_i$  de una arista  $i$  se expresa como una combinación lineal de los productos de coeficientes  $\alpha_j$  por los valores normalizados  $v_{ij}$  de las variables involucradas:

$$w_i = \sum_{j=1}^n \alpha_j \cdot v_{ij}$$

Donde:

- $\alpha_j$  es el coeficiente que indica la importancia relativa de la variable  $j$  en el cálculo del peso.
- $v_{ij}$  es el valor normalizado de la variable  $j$  para la arista  $i$ .

Este enfoque permite integrar múltiples variables, como la distancia, el tiempo, la seguridad, la agradabilidad del entorno, entre otras, en un único valor de peso que refleja las características globales de la arista en el contexto de la ruta peatonal.

### Ajuste de coeficientes según el perfil de ruta

Los coeficientes  $\alpha_j$  no son constantes, sino que se ajustan en función del perfil de ruta deseado. Por ejemplo, para una ruta turística, los coeficientes correspondientes a variables como la cercanía a puntos de interés turístico podrían tener un mayor peso en comparación con otros perfiles, como una ruta que maximice la exposición a espacios verdes. Este ajuste permite personalizar las rutas según las preferencias y necesidades del usuario, logrando un equilibrio adecuado entre los diferentes factores que afectan la experiencia peatonal.

Se plantea definir distintos perfiles de ruta y un conjunto de coeficientes asociados a cada uno de ellos. En la tabla 4.5a se propone un esquema de coeficientes por defecto, y en la tabla 4.5b su correspondiente modificación si se quiere obtener una ruta “verde”.

Factor	Valor
Longitud	-1
Anchura	0,5
Ruido	-0,5
Espacios verdes	0,5
Interés turístico	0,5
Interés social	0,5

(a) Propuesta de coeficientes por defecto

Factor	Valor
Longitud	-1
Anchura	0,5
Ruido	-0,5
Espacios verdes	2
Interés turístico	0,5
Interés social	0,5

(b) Propuesta de coeficientes ajustados para otorgar mayor relevancia a la cercanía a espacios verdes

**Tabla 4.5:** Ejemplos de conjuntos de coeficientes

El esquema de coeficientes por defecto otorga la misma relevancia a prácticamente todos los factores. El signo de cada coeficiente es positivo o negativo en función de su aporte. Por ejemplo: es positivo caminar por calles anchas (por lo que este factor se recompensará asignándole un valor positivo), pero negativo que el nivel de ruido sea alto

(por lo que este factor se castigará asignándole un valor negativo). El multiplicador de la distancia, con un valor de -1, es el único distinto a 0,5 o -0,5, y con él se penaliza a las calles más largas, en aras de optar por alternativas más cortas. Este ajusta no acota explícitamente la longitud de una ruta, pero sí evita que las rutas tomen una extensión excensiva dando rodeos innecesarios.

El segundo conjunto de coeficientes propuesto, destinado a generar rutas que pasen cerca de espacios verdes, cambia el valor del multiplicador en cuestión (el de espacios verdes, en este caso) para darle mayor importancia a este factor. Si los demás factores permanecen iguales y solo se modifica este, tomará más peso en el cálculo. El valor del nuevo coeficiente (2 en lugar de un 0,5) se ha asignado arbitrariamente, buscando un mayor protagonismo a la vez que limitando su distancia a otros coeficientes.

### Normalización de las variables

Para garantizar que las diferentes variables sean comparables y que ninguna domine el cálculo de los pesos debido a su escala, es necesario normalizar los valores de cada variable. Se propone una normalización sencilla que divide el valor de cada variable  $v_{ij}$  entre el valor máximo de esa variable en toda la red:

$$v_{ij} = \frac{x_{ij}}{x_j^{max}}$$

Donde:

- $x_{ij}$  es el valor original de la variable  $j$  para la arista  $i$ .
- $x_j^{max}$  es el valor máximo de la variable  $j$  entre todas las aristas.

Este proceso de normalización asegura que todas las variables estén en un rango común (entre 0 y 1), lo que permite una combinación justa y equilibrada en la fórmula de cálculo de pesos. Además, facilita la interpretación y el ajuste de los coeficientes  $\alpha_j$ , dado que estos se aplican sobre valores normalizados.

### Conversión a pesos positivos

En un grafo como el que representa la red peatonal en cuestión, la presencia de pesos negativos puede llevar a la aparición de ciclos negativos. Un ciclo negativo se produce cuando la suma de los pesos de las aristas que forman un ciclo es negativa. Esto puede ser problemático al calcular el camino más corto utilizando algoritmos como Bellman-Ford, ya que el algoritmo puede quedar atrapado en un bucle indefinido, intentando minimizar continuamente el costo al recorrer el ciclo negativo repetidamente.

Para evitar la aparición de ciclos negativos y asegurar la correcta ejecución del algoritmo de búsqueda del camino más corto, es posible transformar todos los pesos del grafo a valores positivos. Esto se logra sumando a cada peso el valor absoluto del peso mínimo en el grafo. De esta manera, todos los pesos se desplazan hacia valores no negativos, eliminando la posibilidad de ciclos negativos sin alterar la estructura relativa de los pesos. Matemáticamente, si  $w_{min}$  es el peso mínimo en el grafo, se transforma cada peso  $w_i$  en  $w'_i$  de la siguiente forma:

$$w'_i = w_i + |w_{min}|$$

Es importante destacar que, si  $w_{min}$  es negativo, sumarlo con su valor absoluto transformará  $w_{min}$  en 0. Esto puede generar un problema en la siguiente fase, ya que dividir por 0 durante la inversión de pesos causaría un error de división. Para prevenirlo, se establece que, en aquellos casos donde  $w'_i$  sea igual a 0, el valor resultante de la inversión se reemplaza por infinito. Tener pesos infinitos no representa un problema, ya que los algoritmos de búsqueda de caminos cortos están diseñados para encontrar los caminos con los menores pesos, y por lo tanto evitarán transitar por aristas con pesos infinitos.

### Inversión de pesos

Para el cálculo de la ruta normalmente se utilizan algoritmos diseñados para encontrar el camino más corto en un grafo. No obstante, como el objetivo es maximizar el valor de una ruta en lugar de minimizarlo (ya que los pesos de las aristas representan valor positivo que se desea maximizar), se deben invertir los pesos antes de aplicar el algoritmo.

$$w''_i = \frac{1}{w'_i}$$

En los casos donde  $w'_i = 0$ , como se mencionó anteriormente, el valor de  $w''_i$  se reemplaza por infinito para evitar errores de división. Al invertir los pesos, se transforma un problema de maximización en un problema de minimización, lo cual es adecuado para algoritmos como Dijkstra o Bellman-Ford, que están diseñados para encontrar el camino con el menor costo (aunque ahora corresponderá al camino que maximiza la suma de los pesos originales).

**Por qué buscar el camino más corto realmente maximiza el valor de la ruta** Después de invertir los pesos, el algoritmo de búsqueda del camino más corto buscará minimizar la suma de los pesos  $w''_i$ . Dado que estos pesos son inversamente proporcionales a los pesos originales  $w'_i$  (y por extensión, a los  $w_i$  iniciales), minimizar la suma de los  $w''_i$  equivale a maximizar la suma de los  $w_i$  originales.

En otras palabras, al transformar los pesos de manera que los valores altos se convierten en valores bajos y viceversa, el algoritmo de búsqueda del camino más corto naturalmente seleccionará el camino que maximiza los pesos originales, pues lo que minimiza en el grafo transformado es el equivalente a maximizar en el grafo original.

#### 4.2.7. Carga de datos

La fase final del proceso ETL es la carga de los datos transformados en el sistema de destino. Dado que la ruta se calcula a partir de un grafo, el proceso de preparación de datos debe concluir con la creación y carga de esta estructura.

#### Creación del grafo representante de la red

El grafo representante de la red se construye a partir de los datos tabulares obtenidos en las fases anteriores. Cada nodo del grafo corresponde a una intersección de la red peatonal, cada arista a un segmento de la misma, y su peso refleja las características multifactoriales previamente calculadas. Además de ser ponderado, el grafo es no dirigido, pues en el contexto de rutas peatonales, se asume que la capacidad de transitar entre dos puntos es bidireccional, es decir, que un peatón puede moverse en ambas direcciones entre dos nodos conectados sin restricciones direccionales.



NetworkX facilita la creación del grafo y sus componentes al iterar sobre el conjunto de calles de la red y especificar su nodo de origen, su nodo de destino y su peso como parámetros:

```

1 import networkx as nx
2
3 G = nx.Graph()
4 for idx, row in edges_with_data.iterrows():
5     G.add_edge(row['u'], row['v'], weight=row['weight'])

```

Una vez obtenido el grafo, mostrarlo por pantalla es sencillo. Aunque existen múltiples formas de hacerlo, a continuación se propone utilizando la librería Matplotlib [8] para generar un grafo aleatorio  $G$  (véase la figura 4.4) como ejemplo:

```

1 import networkx as nx
2 import matplotlib.pyplot as plt
3
4 pos = nx.spring_layout(G, k=20, seed=42) # Adjust 'k' for more spacing
5
6 nx.draw_networkx_nodes(G, pos, node_size=500, node_color='lightblue')
7 nx.draw_networkx_edges(G, pos, width=2, alpha=0.6, edge_color='gray')
8
9 plt.show()

```

## 4.3 Cálculo de la ruta

Una vez se ha obtenido el peso final de cada arista, se procede a la creación del grafo. Este proceso implica pasar del contexto tabular de los GeoDataFrames a una modelización basada en la teoría de grafos, que representa la red y sobre la cual se aplicará el algoritmo de búsqueda del camino.

### 4.3.1. Imputación de direcciones de origen y destino

Una vez creado y poblado el grafo, es necesario definir las posiciones de inicio y fin para la ruta a calcular. Este proceso se realiza en dos pasos principales: la geocodificación de las direcciones proporcionadas por el usuario y la búsqueda de los nodos más cercanos dentro del grafo.

#### Geocodificación

Solicitar una dirección en lugar de unas coordenadas geográficas mejora significativamente la experiencia del usuario porque las direcciones son más intuitivas y fáciles de recordar, lo que elimina la necesidad de que el usuario conozca o maneje información técnica como latitudes y longitudes. El proceso de traducir una dirección a coordenadas geográficas recibe el nombre de geocodificación, y es fundamental para traducir las entradas del usuario en ubicaciones precisas que puedan ser manejadas dentro del modelo de grafo. Afortunadamente, es un servicio que ofrecen múltiples librerías de Python, como Geopy [28]:

```

1 from geopy.geocoders import Nominatim
2
3 def geocode_address(address):
4     geolocator = Nominatim(user_agent="tutorial")
5     location = geolocator.geocode(address).raw
6     return (location['lon'], location['lat'])

```

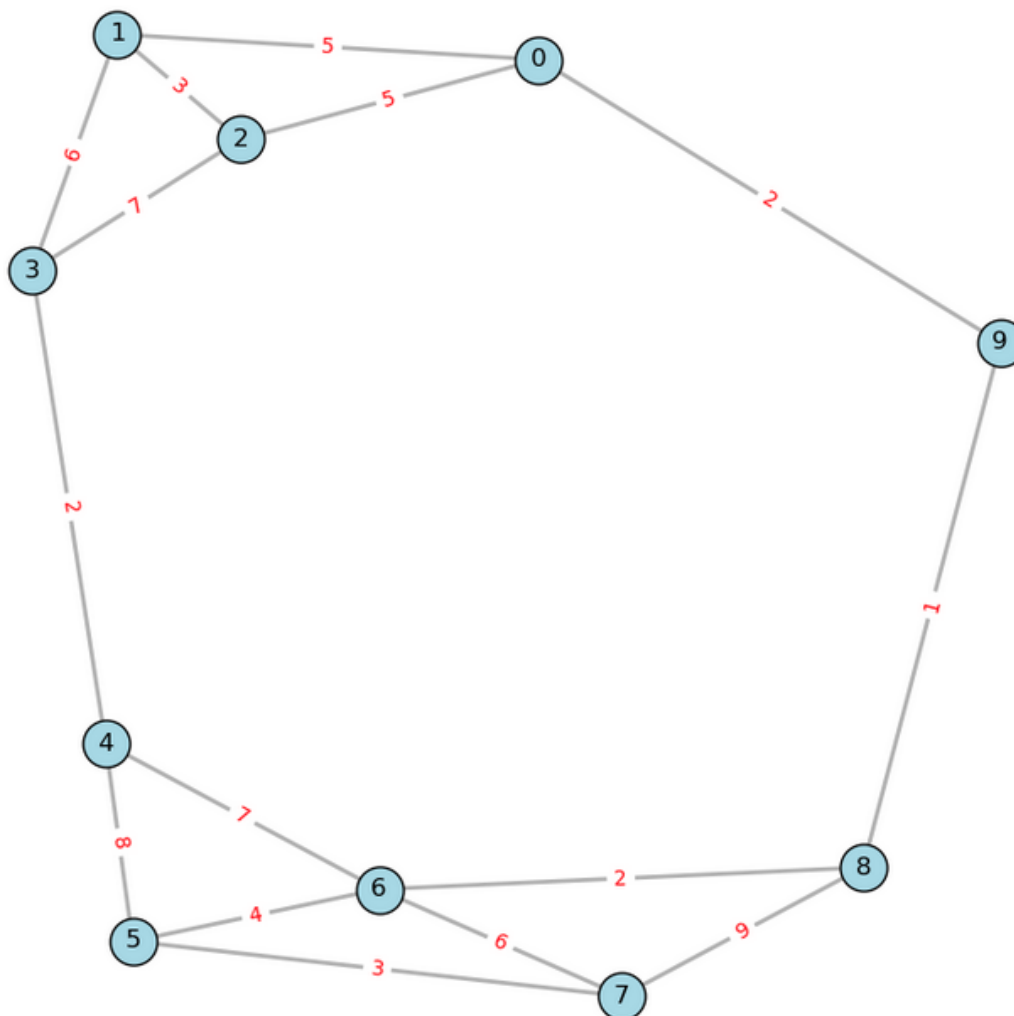


Figura 4.4: Ejemplo de grafo no dirigido y ponderado

```

7
8 origin_coords = geocode_address(
9     "Universidad Politecnica de Valencia, Comarca de Valencia, Valencia, Espana
10    ")
11 destination_coords = geocode_address(
12     "Plaza del Ayuntamiento, Comarca de Valencia, Valencia, Espana")
13 print(origin_coords, destination_coords)
14
15 '''
16 Resultado:
17 ('-0.3420311427062354', '39.4808517')
18 ('-0.3768264115438601', '39.470647299999996')
'''

```

### Búsqueda de los nodos más cercanos

Tras geocodificar las direcciones de origen y destino, el siguiente paso consiste en encontrar los nodos del grafo más cercanos a estas coordenadas. Esto asegura que los puntos de inicio y fin de la ruta se alineen con la estructura del grafo, permitiendo una correcta aplicación del algoritmo de búsqueda de caminos. Cabe destacar que, al tomar

como puntos de partida y destino los nodos más cercanos, se asume un umbral de imprecisión limitado por la distancia a la direcciones proporcionadas.

Un aspecto importante a considerar al buscar el nodo más cercano a unas coordenadas específicas es la necesidad de utilizar metros como unidad de medida para facilitar la interpretación de las distancias. Para ello, es fundamental asegurarse de que las geometrías estén proyectadas en un sistema de coordenadas que utilice el metro como unidad de distancia.

```
1 import geopandas as gpd
2 from shapely.geometry import Point
3
4 def find_nearest_node(coords, nodes_gdf):
5     point = Point(coords)
6     nodes_gdf.crs = 25830 # uses meter as distance unit
7     nodes_gdf['distance'] = nodes_gdf.geometry.distance(point)
8     nearest_node = nodes_gdf.loc[nodes_gdf['distance'].idxmin()]
9     nodes_gdf.crs = 4326 # back to original CRS
10    return nearest_node
11
12 origin_node = find_nearest_node(origin_coords, nodes)
13 destination_node = find_nearest_node(destination_coords, nodes)
```

### 4.3.2. Algoritmo de búsqueda del camino

Como se ha mencionado anteriormente, para el cálculo de la ruta se recurre a algoritmos de búsqueda del camino más corto en un grafo. Al invertir previamente los pesos de las aristas, estos algoritmos devuelven, de forma implícita, el camino con el mayor peso.

El algoritmo de Dijkstra [3], uno de los más estudiados y eficientes de este tipo, se puede aplicar en esta situación, ya que se cumple la condición de que todos los pesos de las aristas sean no negativos.

## 4.4 Representación de la ruta

---

### 4.4.1. Elección del mapa base

En el contexto de este proyecto, la elección de un mapa base es un paso fundamental para visualizar eficazmente las geometrías almacenadas, como los nodos y aristas de la red peatonal, los puntos de interés y los datos sobre ruido. Un mapa base adecuado proporciona un contexto geográfico claro y detallado sobre el que pueden superponerse estos elementos de forma comprensible y visualmente atractiva.

Las funciones de Folium, librería propuesta para la creación de mapas en este tipo de proyectos, disponen de un argumento 'tiles' que permite especificar diferentes tipos de mapa base, tales como "Stamen Terrain", "stamentoner" o su valor por defecto, "OpenStreetMap", entre otros. Algunos ofrecen opciones que destacan las características principales de la ciudad, como las calles, los edificios y los espacios públicos, mientras minimizan el desorden visual.

Para asegurar una visualización clara y concisa, se escoge "CartoDB Positron", ya que equilibra la claridad de la ciudad con la eliminación de detalles innecesarios que puedan distraer. La figura 4.5 captura un trozo de la ciudad bajo la lente del mapa base escogido.

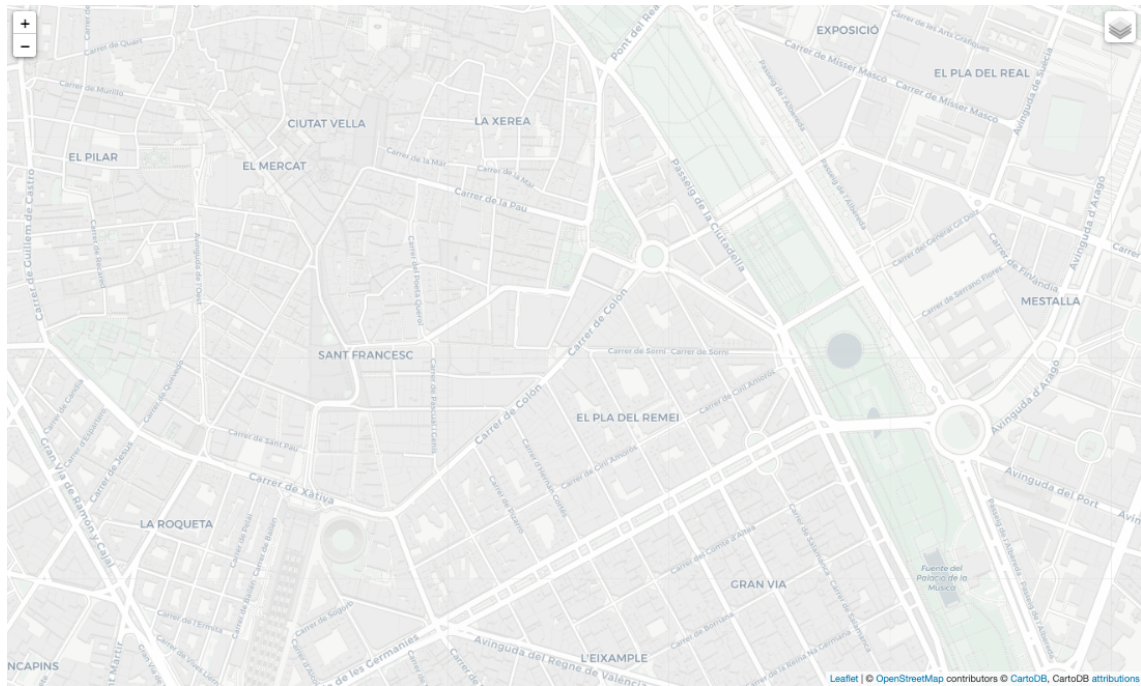


Figura 4.5: Fragmento de la ciudad de Valencia estilizado con el mapa base “CartoDB Positron”

#### 4.4.2. Extracción de geometrías

Una vez calculada la ruta óptima, la representación sobre el mapa se realiza seleccionando las aristas cuyo nodo de origen ( $u$ ) y nodo de destino ( $v$ ) se encuentran en la lista de nodos que componen el camino calculado. Las geometrías de dichas aristas se puede extraer del conjunto de aristas previamente descrito. La figura 4.6 muestra un ejemplo de ruta calculada entre la Universitat Politècnica de València y la Plaza del Ayuntamiento de la ciudad.

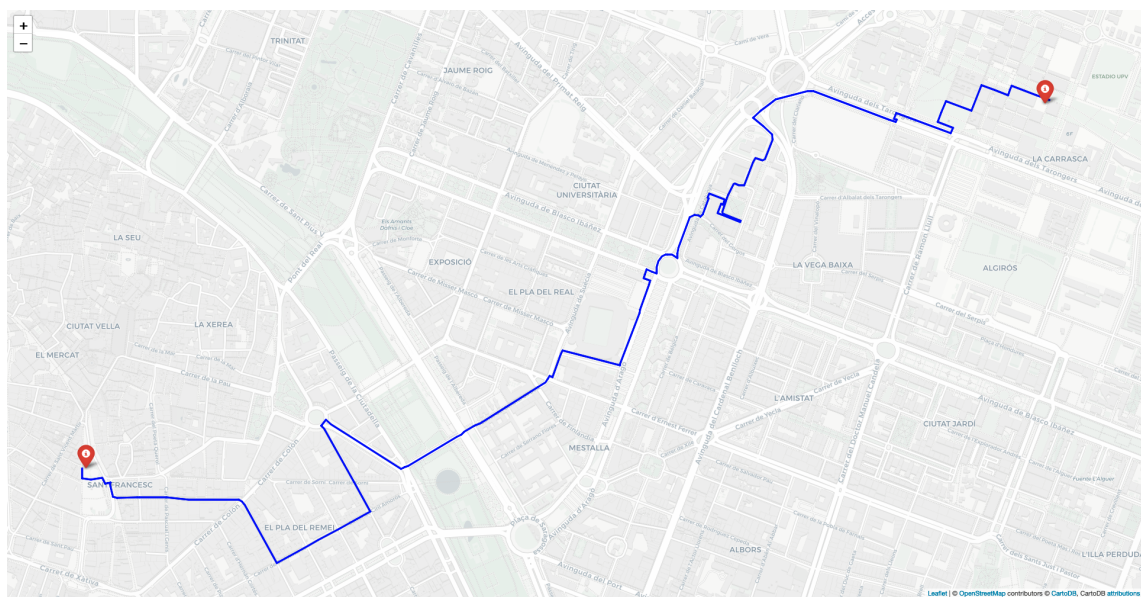


Figura 4.6: Ejemplo de ruta peatonal calculada entre dos puntos de la ciudad

Esta representación permite visualizar la ruta óptima sobre la red geográfica, facilitando su interpretación y uso práctico. Se recomienda utilizar algún tipo de marcador para indicar los puntos de origen y destino.



---

---

# CAPÍTULO 5

## Desarrollo

---

En el capítulo anterior se expuso la metodología propuesta, detallando sus componentes y ofreciendo sugerencias para su implementación. En las secciones que siguen, se describe la aplicación de esta metodología al caso concreto de la ciudad de Valencia, sirviendo como ejemplo ilustrativo.

### 5.1 Extracción de datos de la ciudad de Valencia

---

En esta sección se detalla el proceso de extracción de datos geoespaciales de la ciudad de Valencia, que constituye el primer paso en la implementación práctica de la metodología. Para un proyecto de esta índole, Valencia ofrece una amplia disponibilidad de datos geoespaciales, tanto a través de OpenStreetMap como del portal del Ayuntamiento.

#### 5.1.1. Conjuntos de datos extraídos

Previo a la presentación de los conjuntos de datos empleados, es relevante subrayar que la frecuencia de actualización de los mismos varía significativamente en función de la fuente de origen. Los datos obtenidos del portal de datos abiertos del Ayuntamiento de Valencia se actualizan de manera diaria, lo que garantiza su vigencia y precisión en tiempo real para el análisis. Por otro lado, los datos de OpenStreetMap (OSM) dependen de la participación voluntaria de la comunidad global de usuarios, lo que implica que su actualización no sigue una frecuencia estructurada ni bien definida. Esta naturaleza colaborativa de OSM puede resultar en variaciones en la actualidad de la información, dependiendo de la actividad y el enfoque de los contribuidores en cada región específica.

#### Datos de la red peatonal

A la hora de obtener la red de circulación de la ciudad, en adición a la región de interés, es indispensable puntualizar el tipo de red que se quiere descargar (no hacerlo supone la descarga por defecto de vías aptas para vehículos). Como el tipo que concierne en este caso es el peatonal, se especifica mediante el parámetro correspondiente 'network\_type' como se muestra a continuación:

```
1 import osmnx as ox
2
3 place = 'Valencia, Comunitat Valenciana, Spain'
4 graph = ox.graph_from_place(place_name, network_type='walk')
5 nodes, edges = ox.graph_to_gdfs(graph)
```



En las tablas 4.2 y 4.3 se mostraron fragmentos de las estructuras devueltas para nodos (intersecciones) y aristas (vías), respectivamente. Los GeoDataFrames completos para la red peatonal de la ciudad de Valencia se componen de 38936 nodos y 121898 aristas. Dada la gran magnitud de esta red, se vuelve imperativa la optimización de las operaciones que involucren a los nodos o a las aristas, a fin de reducir el coste computacional y el tiempo de ejecución de las mismas. El subconjunto de la red mostrado en la figura 5.1 refleja la densidad de calles y cruces en la ciudad.



**Figura 5.1:** Visualización sobre mapa del conjunto de la red peatonal de Valencia

El conjunto de aristas incluye información sobre la longitud (campo “length”) y la anchura (campo “lanes”) de las calles —factores de relevancia para el cálculo de la ruta—, por lo que no es necesaria una extracción independiente de estos datos.

### Datos de puntos de interés turístico

El catálogo de datos abiertos del Ayuntamiento de Valencia permite filtrar por categorías. Dentro de la dedicada al turismo, se selecciona el único conjunto de datos que ofrece la localización de los monumentos turísticos considerados por esta institución como los más importantes de la ciudad, cuyo identificador es *monuments-turisticos-monumentos-turisticos*. Se compone de 225 registros y 10 campos, mostrados en la tabla 5.1.

Nombre	Descripción	Ejemplo
gid	Identificador en el conjunto de datos	4
nombre	Nombre del monumento	PALAU DE MONTORNÉS
numpol	Número de portal en la calle	1, 3, 5
telefono	Teléfono de contacto	0
ruta	Desconocido	7
audio	Desconocido	0
escena	Desconocido	
nombreaudio	Desconocido	
geo_shape	Geometría compleja de su ubicación	{“coordinates”:[-0.3719090538302984, 39.473989548561875]}
geo_point_2d	Punto simple representativo de su ubicación	39.473989548561875, -0.37190905383

**Tabla 5.1:** Campos que componen el conjunto de datos original sobre monumentos turísticos del Ayuntamiento de Valencia

Como se puede ver, el conjunto presenta algunos datos faltantes. Para el registro expuesto, por ejemplo, los campos “escena” y “nombreaudio” están vacíos. También lo

están los campos “telefono” y “audio”, aunque en estos casos la ausencia de datos se representa mediante un ‘0’. Sin embargo, los campos más relevantes, como el del nombre y el referente a la ubicación geográfica del monumento, están disponibles en todos los casos; por ello, estos son los únicos atributos que no se desechan. La tabla 5.2 muestra de nuevo el mismo extracto del conjunto tras reducir el número de campos a los indispensables. Nótese que el campo “geo\_point\_2d” también se desecha, pues ofrece información redundante que ya se encuentra en “geo\_shape”, donde, además, la estructura de los datos es más compatible con las herramientas de análisis debido a su formato estándar en forma de diccionario. Por coherencia con el resto de conjuntos de datos, los campos restantes pasan a denominarse “name” y “geometry”, siendo este último además formateado según el estándar para geometrías Shapely [25].

Nombre	Ejemplo
name	PALAU DE MONTORNÉS
geometry	POINT (-0.37191 39.47399)

**Tabla 5.2:** Campos del conjunto de datos sobre monumentos turísticos del Ayuntamiento de Valencia utilizados en el análisis

Asimismo, es conveniente señalar que al obtenerse mediante API, la cual devuelve JSON, todos los datos descargados son de tipo texto. Esto no supone un problema, pues se les asigna el tipo de datos más adecuado en la posterior fase de transformación.

Del portal del Ayuntamiento se escoge un segundo conjunto de datos, que se identifica como *be-bien-brl* y se encuentra en la categoría de urbanismo e infraestructuras. Contiene la localización de 325 bienes de relevancia local en la ciudad de Valencia, y complementa al conjunto de monumentos. Se describen los campos que lo componen en la tabla 5.3.

Nombre	Descripción	Ejemplo
NOMBRE	Nombre del bien	IGLESIA PARROQUIAL DE SAN JUAN Y SAN VICENTE
FICHA	Nombre del documento oficial en el que consta	02-15 Iglesia Parroquial de San Juan y San Vicente_firmado.pdf
OBJECTID	Identificador en el conjunto de datos	1
geo_shape	Geometría compleja de su ubicación	{“coordinates”:[[-0.36979851944997333,39.46882824652696],...
geo_point_2d	Punto simple representativo de su ubicación	39.4688955625821, -0.3700387920490497

**Tabla 5.3:** Campos que componen el conjunto de datos original sobre bienes de relevancia local del Ayuntamiento de Valencia

De nuevo, los campos útiles se reducen al nombre y la geometría que indica la ubicación, que por coherencia con el resto de conjuntos han sido renombrados a “name” y “geometry” (formateado también como geometría Shapely), como se observa en la tabla 5.4.

Nombre	Ejemplo
name	IGLESIA PARROQUIAL DE SAN JUAN Y SAN VICENTE
geometry	POLYGON ((-0.36980 39.46883, -0.36993 39.46877...

**Tabla 5.4:** Campos del conjunto de datos sobre bienes de relevancia local del Ayuntamiento de Valencia utilizados en el análisis

Por lo que respecta a los datos turísticos de OpenStreetMap, el esquema de etiquetas escogido para la petición vía OSMnx es el siguiente:

```
{‘tourism’: [‘artwork’, ‘attraction’, ‘gallery’, ‘information’, ‘museum’,
‘viewpoint’]}
```



De entre todas las etiquetas listadas en la página web de OSM bajo la categoría “Tourism”, esas seis se escogen arbitrariamente al considerarse las más coherentes con su objetivo: hallar entidades de interés turístico. Otras, como ‘camp\_pitch’ (zona de acampada) no presentan necesariamente valor turístico, a pesar de estar agrupadas en la misma categoría.

La petición filtrando por las etiquetas arriba mencionadas devuelve un GeoDataFrame compuesto por 422 registros y 4 campos, mostrados en la tabla 5.5.

Nombre	Descripción	Ejemplo
osmid	Identificador en OSM	254021929
name	Nombre de la entidad	Monument a Joaquim Sorolla
geometry	Geometría compleja de su ubicación	POINT (-0.32981 39.46347)
tourism	Etiqueta que ha causado la coincidencia	artwork

**Tabla 5.5:** Campos que componen el conjunto de datos original sobre monumentos turísticos de OpenStreetMap

Al igual que con los datos provenientes del Ayuntamiento, la tabla 5.6 recoge los atributos considerados indispensables que se utilizan posteriormente en el análisis. El campo “geometry” es el homólogo del campo “geo\_shape” del conjunto anterior.

Nombre	Ejemplo
name	Monument a Joaquim Sorolla
geometry	POINT (-0.32981 39.46347)

**Tabla 5.6:** Campos del conjunto de datos sobre monumentos turísticos de OpenStreetMap utilizados en el análisis

Antes de realizar transformaciones sobre los datos, es preciso constituir un conjunto de datos de interés turístico único, que estará conformado por los provenientes de OSM y del Ayuntamiento. Como los tres conjuntos cuentan con los mismos campos, una simple concatenación bastará para integrarlos todos. La tabla 5.7 ofrece una vista previa del resultado de esta unión. También se puede visualizar el conjunto sobre un mapa, como se muestra en la figura 5.2. En esta visualización se aprecia que el conjunto está formado por varios tipos de geometría, incluyendo puntos y polígonos.

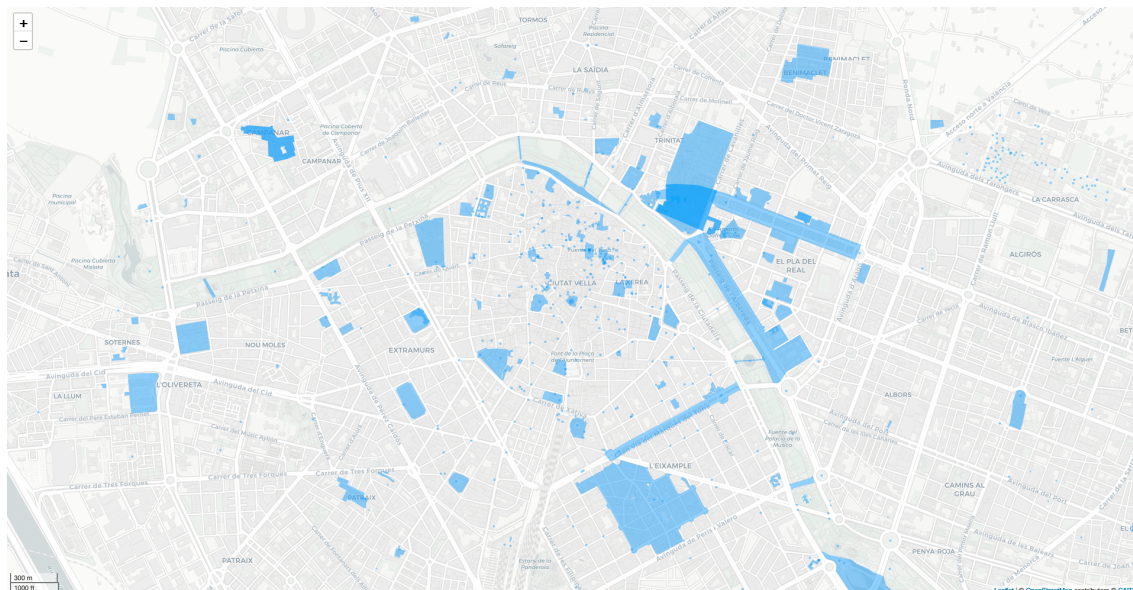
	name	geometry
0	PALAU DE MONTORNÉS	POINT (-0.37191 39.47399)
1	EDIFICIO CHAPA, G.V.MARQUES DEL TURIA 69	MULTIPOLYGON (((-0.36662 39.46840, -0.36664 39...
2	EDIFICIO CHAPA, GRABADOR ESTEVE 36	POLYGON ((-0.36643 39.46899, -0.36643 39.46899...
...	...	...
969	Gulliver	MULTIPOLYGON (((-0.35947 39.46270, -0.35947 39...
970	Museu del Corpus i Casa de les Roques	POLYGON ((-0.37692 39.47898, -0.37698 39.47900...
971	Monument a Joaquim Sorolla	POINT (-0.32981 39.46347)

**Tabla 5.7:** Extracto del conjunto final sobre monumentos turísticos, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia

Cuando se tiene el conjunto único, es conveniente estandarizar su estructura convirtiéndolo a un formato tabular manejable, preferentemente un GeoDataFrame (como se explicó en la subsección 4.2.3).

### Datos de zonas verdes

El procedimiento a seguir para la extracción de datos de zonas verdes y los subconjuntos es análogo al mostrado para los datos de puntos de interés turístico.



**Figura 5.2:** Visualización sobre mapa del conjunto final sobre monumentos turísticos

En este caso, se vuelven a tomar datos tanto del portal del Ayuntamiento como de OSM. Convenientemente, el Ayuntamiento pone a disposición un conjunto identificado como *espais-verds-espacios-verdes* que contiene la localización de espacios verdes en la ciudad de Valencia. Un segundo conjunto, dedicado a la localización individual de cada árbol registrado de la ciudad se descarta por su granularidad excesivamente fina. El conjunto escogido se compone de 772 y 14 campos, descritos en la tabla 5.8. La versión reducida posteriormente utilizada para el análisis se recoge en la tabla 5.9.

Nombre	Descripción	Ejemplo
OBJECTID	Identificador en el conjunto de datos	2.043
Id. Jardín	Identificador del jardín	99101506
Nombre	Nombre del jardín	Músico Espi Pl
Barrio	Barrio que alberga el jardín	TORREFIEL
Tipología	Tipo de espacio verde	Jardines Barrio Plaza
Área	Área cubierta por el jardín	3.749,720
Número Elementos Fitness	Desconocido	
Superficie Huerto Urbano	Desconocido	
Zona	Zona de la ciudad	NORTE
DM	Distrito	RASCANYA
Superficie Total	Superficie del recinto que alberga el jardín	3.888,6
Ud. Gestion	Unidad de gestión del jardín	SJ JARDINES NORTE
geo_shape	Geometría compleja de su ubicación	{“coordinates”:[[[[-0.3744873995245515,39.49645163879353]...]]
geo_point_2d	Punto simple representativo de su ubicación	39.482936824391665, -0.35839979049

**Tabla 5.8:** Campos que componen el conjunto de datos original sobre zonas verdes del Ayuntamiento de Valencia

Nombre	Ejemplo
name	Músico Espi Pl
geometry	MULTIPOLYGON (((-0.37449 39.49645...

**Tabla 5.9:** Campos del conjunto de datos sobre zonas verdes del Ayuntamiento de Valencia utilizados en el análisis

Para la obtención vía OSM se escoge el siguiente esquema de etiquetas, halladas bajo la categoría “leisure” (en español *ocio*):

```
{‘leisure’: [‘garden’, ‘nature_reserve’, ‘park’, ‘playground’]}
```

El GeoDataFrame devuelto por la petición se compone de 2037 registros y 4 campos, mostrados en la tabla 5.10. Su versión reducida se ve en la tabla 5.11.

Nombre	Descripción	Ejemplo
osmid	Identificador en OSM	1189715504
name	Nombre de la entidad	NaN
geometry	Geometría compleja de su ubicación	POINT (-0.39384 39.45502)
leisure	Etiqueta que ha causado la coincidencia	playground

**Tabla 5.10:** Campos que componen el conjunto de datos original sobre zonas verdes de OpenStreetMap

Nombre	Ejemplo
name	NaN
geometry	POINT (-0.39384 39.45502)

**Tabla 5.11:** Campos del conjunto de datos sobre zonas verdes de OpenStreetMap utilizados en el análisis

La unión de ambos conjuntos resulta en la estructura mostrada en la tabla 5.12 o su alternativa visual, la figura 5.3.

	name	geometry
0	PARQUES DE BENIFERRI	MULTIPOLYGON (((-0.40563 39.49405, -0.40558 39...
1	DELS LLANTERNERS C (MEDIANERA)	MULTIPOLYGON (((-0.40862 39.46199, -0.40862 39...
2	HORTICULTOR CORSET PL	POLYGON ((-0.39296 39.47331, -0.39250 39.47266...
...	...	...
2034	Ornamentals	MULTIPOLYGON (((-0.38726 39.47783, -0.38713 39...
2035	Desèrtiques	MULTIPOLYGON (((-0.38650 39.47796, -0.38648 39...
2036	NaN	MULTIPOLYGON (((-0.36688 39.49261, -0.36685 39...

**Tabla 5.12:** Extracto del conjunto final sobre zonas verdes, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia



**Figura 5.3:** Visualización sobre mapa del conjunto final sobre zonas verdes



## Datos de puntos de interés social

Para el perfil social de las rutas se buscan puntos que representen locales de ocio y otras amenidades, como cafeterías y restaurantes. En este momento, el portal del Ayuntamiento de Valencia no lista ningún conjunto de datos valioso al respecto, por lo que OpenStreetMap es la única fuente de datos utilizada. En consecuencia, no se precisa de la combinación de varios conjuntos.

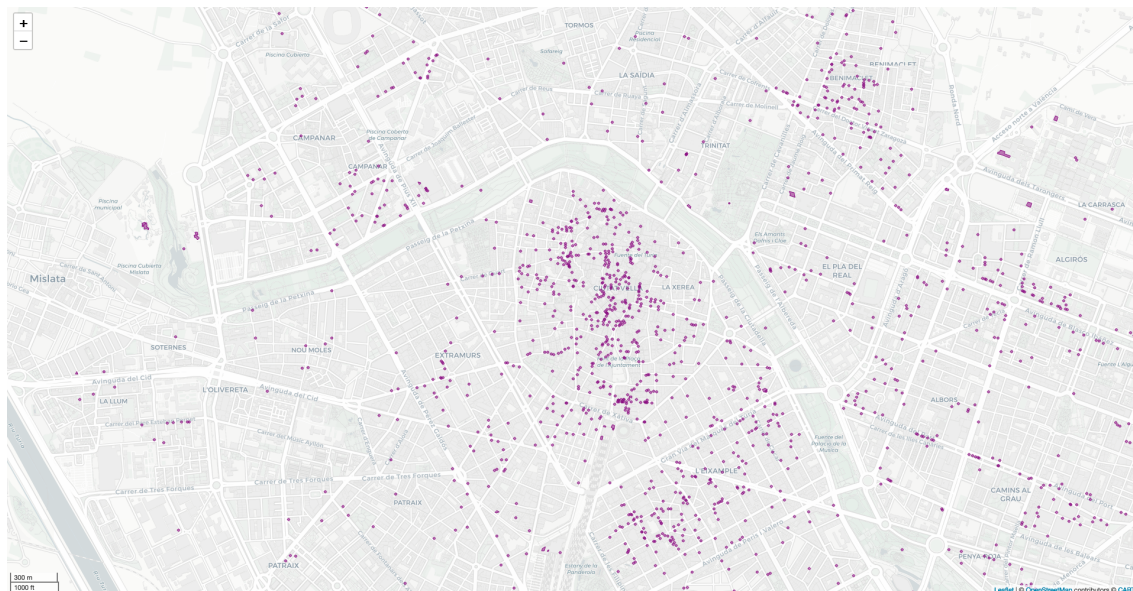
Se opta por la utilización de las siguientes etiquetas en la petición a OSM:

```
{'amenity': ['bar', 'cafe', 'pub', 'restaurant']}
```

La tabla 5.13 muestra un fragmento del conjunto final obtenido, formado por 1699 registros. Por su parte, la figura 5.4 permite observar que, en este conjunto concreto, la mayoría de elementos toman forma de punto, siendo pocos los polígonos visibles.

	name	geometry
0	Ni pa ti ni pa mi	POINT (-0.38186 39.47101)
1	Cervecería Velluters	POINT (-0.38040 39.47118)
2	Café Santos	POINT (-0.38088 39.47097)
...	...	...
1696	Anar i tornar	POLYGON ((-0.37759 39.46685, -0.37746 39.46682...
1697	Cafestore	POLYGON ((-0.37784 39.46615, -0.37778 39.46633...
1698	NaN	POLYGON ((-0.37577 39.44298, -0.37577 39.44295...

**Tabla 5.13:** Extracto del conjunto final sobre locales sociales, resultado de la unión de datos de OpenStreetMap y del Ayuntamiento de Valencia



**Figura 5.4:** Visualización sobre mapa del conjunto final sobre locales sociales

## Datos de ruido

Los datos de ruido obtenidos requieren de un tratamiento distinto a los anteriores conjuntos. En este caso, solo el Ayuntamiento dispone de esta información, por lo que su portal será la fuente exclusiva. El conjunto identificado como *mapa-soroll-dia-mapa-ruido-dia* ofrece información sobre 53 polígonos de ruido que cubren prácticamente la

totalidad de la superficie de la urbe valenciana. Los campos que componen este conjunto se describen en la tabla 5.14.

Cabe destacar que el conjunto escogido ofrece información de los niveles de ruido correspondientes al día de consulta. Por tanto, es posible que los datos sean distintos si han sido descargados en diferentes días. No obstante, no se esperan grandes variaciones de un para otro, por lo que se trabaja sobre el supuesto de que el día de descarga de los datos no tiene un impacto notable en el resultado de la generación de la ruta.

Asimismo, conviene hacer saber que en el portal del Ayuntamiento se lista otro conjunto similar, pero con una granularidad diaria. Dado que estudiar la evolución del ruido a lo largo del día no es uno de los objetivos de este proyecto, se descarta el uso de este conjunto alternativo.

Nombre	Descripción	Ejemplo
objectid	Identificador en el conjunto de datos	20
Nivel	Nivel de ruido (1 al 6)	3
geo_shape	Geometría del polígono	{"coordinates":[[[-0.338652954211283,39.422727580117716],...]
geo_point_2d	Punto representativo del polígono	-0.3265648600516565, 39.37797557989289

**Tabla 5.14:** Campos que componen el conjunto de datos original sobre niveles de ruido del Ayuntamiento de Valencia

Los campos de mayor relevancia y que se utilizan posteriormente en el análisis se recogen en la tabla 5.15. En interés de cumplir con la convención de nombres seguida hasta el momento, se renombran los campos a “level” (nivel de ruido) y “geometry” (geometría del polígono de ruido).

Nombre	Ejemplo
level	3
geometry	MULTIPOLYGON (((-0.33865,39.42273,...

**Tabla 5.15:** Campos del conjunto de niveles de ruido del Ayuntamiento de Valencia utilizados en el análisis

La figura 5.5 representa gráficamente los datos de ruido antes de aplicar transformaciones sobre ellos. Se ha utilizado un código de colores intuitivo para facilitar la interpretación visual de los datos. Las áreas con menor nivel de ruido están coloreadas en tonos verdes, simbolizando un entorno más tranquilo y menos ruidoso. A medida que aumenta el nivel de ruido, los colores transicionan hacia tonos marrones y naranjas, representando niveles moderados de ruido. Finalmente, las zonas más ruidosas están marcadas en tonos rojizos, indicando un mayor nivel de ruido ambiental. Este gradiente de color, que va del verde al rojo, proporciona una forma clara y efectiva de visualizar cómo varían los niveles de ruido en la ciudad, y permite apreciar que, generalmente, las áreas que registran mayores niveles de ruido coinciden con carreteras y vías para vehículos.

Si bien la visualización de los polígonos de ruido ofrece un vistazo rápido de cuáles son las áreas más afectadas por la contaminación acústica, el conjunto debe ser sometido a una serie de transformaciones antes de utilizar el nivel de ruido como un factor en la generación de rutas. Concretamente, se necesita asociar un nivel de ruido a cada de la red. La siguiente fase del proceso ETL cumple este requisito.

Las figuras 5.6a y 5.6b se utilizan para ejemplificar los distintos niveles de ruido posibles. En la primera de ellas toma el centro de la imagen una de las arterias de la ciudad: la Gran Vía. El concurrido tráfico en esta zona le otorga el nivel de ruido más alto, de ahí su color rojo tan intenso. Nótese cómo el antiguo cauce del río Turia, al final de la avenida (parte derecha de la imagen), registra tonos marrones a pesar de ser un importante



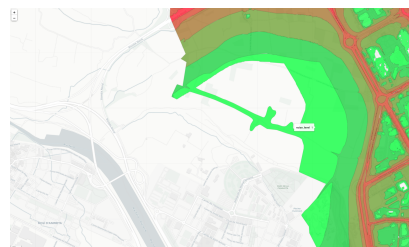
**Figura 5.5:** Visualización sobre mapa del conjunto final sobre niveles de ruido

espacio verde en la ciudad. Esto se debe a que está rodeado a ambos lados por calles que también atraen un gran flujo de vehículos, con su consecuente impacto acústico.

En cambio, en la segunda imagen podemos encontrar el extremo opuesto. En la parte central se observa una forma curva y verde. Se trata del Parque de Cabecera, un amplio espacio natural. A pesar de tener cerca carreteras muy ruidosas, al no verse influenciado por ningún polígono de ruido a su izquierda (pues ese área ya pertenece a otro término municipal, ajeno a la jurisdicción del Ayuntamiento de Valencia), el balance es positivo, recibiendo el nivel de ruido más bajo posible.



(a) Ejemplo 1: nivel de ruido alto en la Gran Vía



(b) Ejemplo 2: nivel de ruido bajo en el Parque de Cabecera

**Figura 5.6:** Ejemplos de niveles de ruido bajo y alto en la ciudad de Valencia

## 5.2 Transformación de los datos extraídos

### 5.2.1. Eliminación de datos duplicados

Siguiendo las indicaciones sobre el tratamiento de duplicados dadas en la subsección 4.2, se examina la cantidad de registros donde el campo referente al nombre del elemento está ausente, resumidos en la tabla 5.16. Quedan fuera de este estudio los conjuntos de la red peatonal, los niveles de ruido y los puntos de interés social, por las siguientes razones:

1. Son conjuntos extraídos de una única fuente, por tanto, no se esperan repeticiones al no haberse combinado con datos de otros repositorios que recogen información similar.
2. Carece de sentido eliminar calles del conjunto de datos, pues se estaría alterando la representación de la red peatonal de la ciudad, dando pie a rutas potencialmente erróneas y no óptimas.
3. El conjunto de niveles de ruido no presenta un campo dedicado al nombre. Además, si se elimina alguno de sus registros, es posible que determinadas porciones de la ciudad queden descubiertas sin ningún polígono al que asociarlas, no pudiendo entonces asignarles un nivel de ruido.

Conjunto de datos	Registros totales	Registros con nombre ausente
Puntos de interés turístico	972	91 (9.36 %)
Zonas verdes	2807	1913 (68.15 %)

**Tabla 5.16:** Cantidad de registros con nombre ausente por conjunto de datos

En vista de los resultados, se prescinde de todos los registros con nombre ausente del conjunto de puntos de interés turístico, ya que la eliminación de menos de un 10 % de los datos no supone una diferencia notable en un conjunto con esta magnitud. Esta decisión se ve apoyada por el descubrimiento adicional de que todos los ausentes provienen de OSM, lo cual no supone una sorpresa, pues es resulta razonable esperar que la entidad propietaria de los datos (el Ayuntamiento de Valencia) proporcione datos más fiables y completos que fuentes particulares pueden no disponer de toda la información (contribuyentes voluntarios de OSM). En cierto modo, la calidad global del conjunto aumenta con esta eliminación.

Por lo que respecta al conjunto de zonas verdes, se descarta la eliminación de registros con nombre ausente, dado el gran porcentaje del total que suponen (más de la mitad). A cambio, se opta por buscar duplicados comparando entre sí los nombres de los registros del conjunto. Es improbable que el nombre de dos registros, a pesar de representar el mismo elemento en la realidad, coincida literalmente si estos provienen de fuentes distintas (la escritura de los nombres puede presentar variaciones menores, como diferencias en el uso de mayúsculas, errores tipográficos, o la presencia de caracteres especiales y abreviaturas). Por ello, se plantea un enfoque que considere un umbral de similitud del 90 % para permitir identificar registros que, aunque no coincidan exactamente, representan el mismo elemento. De este modo, se garantiza una detección más precisa y exhaustiva de duplicados.

La tabla 5.17 muestra el estado actualizado de estos dos conjuntos tras el tratamiento aplicado con el fin de eliminar duplicados.

Conjunto de datos	Registros totales	Registros con nombre ausente
Puntos de interés turístico	881	0 (0.00 %)
Zonas verdes	2761	1913 (69.29 %)

**Tabla 5.17:** Cantidad de registros con nombre ausente por conjunto de datos tras aplicar transformaciones para la reducción de duplicados

Los resultados de esta tabla actualizada revelan dos sucesos relevantes:

- El conjunto de monumentos turístico ya no presenta ningún registro con el campo del nombre vacío.

- El conjunto de zonas verdes presenta el mismo número de registros con nombre ausente a pesar de que el número de registros totales se ha visto reducido, lo que lleva a la conclusión de que sí existían registros duplicados. Nótese que el porcentaje de ausentes ha incrementado ligeramente, pues la misma cantidad representa una porción mayor en el conjunto ahora reducido.

Aunque útil, el método de eliminación de duplicados por similitud no es infalible, pues pueden permanecer algunos duplicados que no disponen de nombre. Para deshacerse de ellos, habría que recurrir a una comparación espacial, pero se considera que el coste computacional y temporal de esa operación supera el impacto que unos pocos duplicados pueden tener en la calidad del análisis posterior.

### 5.2.2. Imputación de datos faltantes

Como se comentó en la subsección 5.1.1, el conjunto de datos de la red peatonal, concretamente el subconjunto de las calles, contiene información sobre la longitud y la anchura de las mismas. Tras un breve análisis exploratorio se observa que, aunque la longitud es un atributo disponible para todas las calles, tan solo el 12 % de ellas cuentan con un valor para la anchura.

La anchura de las calles en este conjunto se cuantifica numéricamente, de modo que a menor valor corresponde una anchura reducida, mientras que a mayor valor corresponde una anchura mayor. Como método de imputación para datos faltantes, se decide asignar un valor de '1' a todas aquellas calles con el campo vacío. Esto se realiza bajo la asunción de que las vías más importantes (avenidas, bulevares, paseos, etc.) tienen documentada su anchura, mientras que aquellas de menos relevancia (como callejones) no siempre presentan este valor. La tabla 5.18 muestra la distribución de esta variable en el conjunto.

Valor de anchura	Número de registros
1	111570 (91.53 %)
2	5295 (4.34 %)
3	2606 (2.14 %)
4	1608 (1.32 %)
5	508 (0.42 %)
6	140 (0.12 %)
7	70 (0.05 %)
8	104 (0.08 %)

Tabla 5.18: Distribución de los valores de anchura de las calles

### 5.2.3. Combinación de conjuntos de datos

Después de la extracción y preparación individual de los conjuntos de datos, se procede a su integración para proporcionar una visión global. El conjunto de datos principal, que servirá como base para los demás conjuntos, corresponde a las calles de la ciudad. Por consiguiente, se inicia con un GeoDataFrame que inicialmente contiene varias columnas vacías, las cuales se completarán de manera progresiva (véase la tabla 5.19).

Para mayor claridad, se aporta una descripción de cada campo:

- **u**: nodo de origen (un extremo de la calle). Corresponde con uno de los registros en el subconjunto de nodos (obtenido como parte del conjunto de la red peatonal).



	u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	3.947	3	None	None	None	None	None
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	13.755	3	None	None	None	None	None
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	56.653	1	None	None	None	None	None
...	...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	31.588	1	None	None	None	None	None

**Tabla 5.19:** Fragmento del conjunto base donde se combinan los datos extraídos

- **v:** nodo de destino (otro extremo de la calle). Corresponde con uno de los registros en el subconjunto de nodos (obtenido como parte del conjunto de la red peatonal).
- **osmid:** identificador de la vía en OSM.
- **name:** nombre de la vía.
- **length:** longitud de la vía.
- **width:** anchura de la vía (versión renombrada de “lanes”).
- **noise\_level:** nivel de ruido de la vía.
- **num\_green:** número de zonas verdes próximas a la vía.
- **num\_tourism:** número de puntos de interés turístico próximos a la vía.
- **num\_social:** número de puntos de interés social próximos a la vía.
- **weight:** peso asignado a la vía (arista), calculado a partir de varios factores.

### Asignación de un nivel de ruido a cada calle de la red

Uno de los mayores retos computacionales de este proyecto es la asociación de calles con niveles de ruido. En esencia, consiste en comprobar cuál es el valor de ruido más adecuado para cada vía de la red en base a su relación espacial con los polígonos de ruido. Como se puede intuir, este proceso requiere muchas operaciones de comparación entre geometrías, por lo que se introduce un índice espacial para reducir el número de intersecciones a calcular. Concretamente, se recurre a un STRTree. Tras probar dos implementaciones, con GeoPandas y Shapely, esta segunda resulta en una ejecución mucho más rápida.

La figura 5.7 recoge, en forma de diagrama, los pasos seguidos para la asignación utilizando indexación espacial.

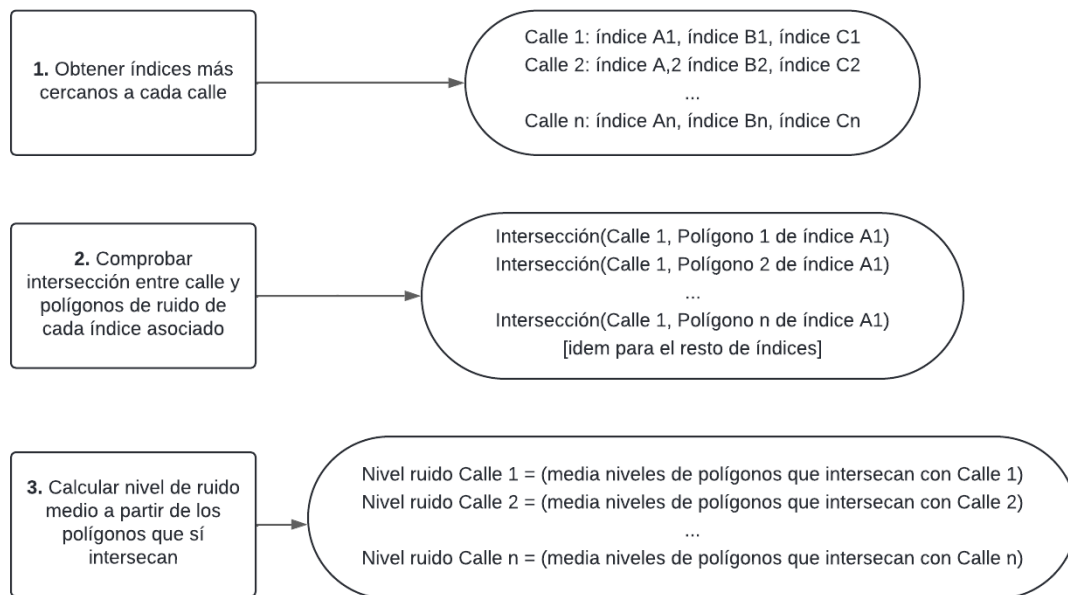
De esta forma, cada calle en nuestro conjunto queda asociada a un valor de ruido, como muestra la tabla 5.20. El mapa de la figura 5.8 permite confirmar que, en efecto, cada calle recibe un valor, el cual se refleja en su correspondiente coloración. Los polígonos que antes cubrían la ciudad se han convertido en valores dentro de cada calle.

	u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	3.947	3	5.0	None	None	None	None
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	13.755	3	5.0	None	None	None	None
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	56.653	1	1.5	None	None	None	None
...	...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	31.588	1	3.0	None	None	None	None

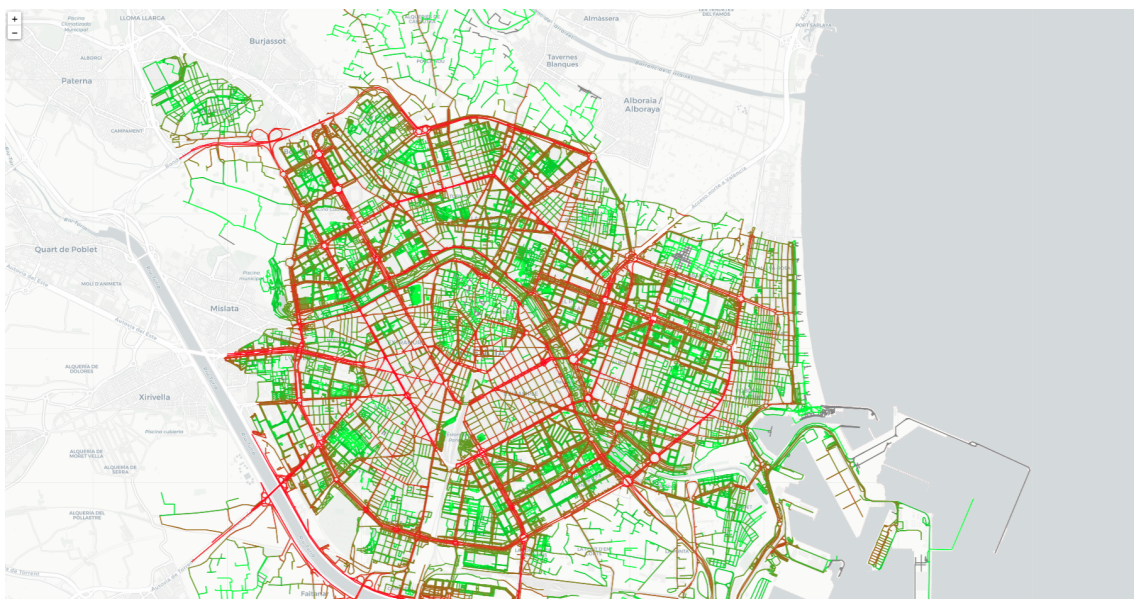
**Tabla 5.20:** Fragmento del conjunto base tras añadir los datos de ruido

### Asignación de número de zonas verdes próximas a cada calle de la red

Para contar el número de zonas verdes presentes en las inmediaciones de cada calle, se utiliza un buffer de 100 metros alrededor de cada una de ellas. Los buffers se pueden



**Figura 5.7:** Diagrama el proceso de asignación de un nivel de ruido a cada calle



**Figura 5.8:** Visualización sobre mapa de las calles coloreadas según el nivel de ruido

superponer, y por tanto, una zona verde puede ser contada para más de una calle. Otro matiz a destacar es el sistema de coordenadas utilizado para proyectar las geometrías de las calles en la creación de los buffers. Si se desea crear un perímetro en metros en torno a las calles, antes hay que cerciorarse de que el sistema en uso admite esa unidad de medida de la distancia.

Una vez asignadas las zonas a cada buffer, se puede añadir esta información a su correspondiente columna del conjunto base (véase figura 5.21).

u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	3.947	3	5.0	1	None	None
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	13.755	3	5.0	1	None	None
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	56.653	1	1.5	1	None	None
...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	31.588	1	3.0	1	None	None

Tabla 5.21: Fragmento del conjunto base tras añadir los datos de zonas verdes

### Asignación de un número de puntos de interés turístico a cada calle de la red

Para asignar puntos de interés turístico a cada buffer, se procede de la misma forma que con las zonas verdes. El resultado se recoge en la tabla 5.22.

u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	3.947	3	5.0	1	3	None
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	13.755	3	5.0	1	4	None
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	56.653	1	1.5	1	1	None
...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	31.588	1	3.0	1	6	None

Tabla 5.22: Fragmento del conjunto base tras añadir los datos de puntos de interés turístico

### Asignación de un número de puntos de interés social a cada calle de la red

Como última adición en este proceso iterativo, se asigna a cada calle el número de puntos de interés social próximos, completando así el último factor en la tabla 5.23 (el peso se calcula en un paso posterior).

u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	3.947	3	5.0	1	3	7
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	13.755	3	5.0	1	4	7
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	56.653	1	1.5	1	1	7
...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	31.588	1	3.0	1	6	32

Tabla 5.23: Fragmento del conjunto base tras añadir los datos de puntos de interés social

#### 5.2.4. Normalización de las variables

Como se explicó en la subsección 4.2.6, se normalizan las variables para hacerlas equiparables. La tabla 5.24 demuestra el aspecto del conjunto base tras aplicar esta transformación.

u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.001111	0.272727	0.833333	0.035714	0.103448	0.162791
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.003871	0.272727	0.833333	0.035714	0.137931	0.186047
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.037527...	0.015944	0.090909	0.250000	0.035714	0.034483	0.744186
...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	0.008890	0.090909	0.500000	0.035714	0.206897	0.255814

Tabla 5.24: Fragmento del conjunto base tras normalizar los factores implicados

#### 5.2.5. Definición de los perfiles de ruta

Uno de los objetivos de este proyecto es utilizar las variables recogidas para generar una ruta acorde a las preferencias del usuario. La forma escogida para ello es mediante la definición de cuatro perfiles de ruta que obedezcan a distintas prioridades:

- **Rutas verdes:** priorizan el contacto con áreas naturales, parques y espacios arbolados de la ciudad. Estas rutas están diseñadas para pasar por el mayor número posible de zonas verdes sin una desviación drástica del destino final.
- **Rutas silenciosas:** se enfocan en minimizar la exposición del usuario a entornos ruidosos, recompensando calles que registran niveles de ruido bajo y penalizando grandes avenidas con tráfico intenso y otros focos de ruido.

- **Rutas turísticas:** están diseñadas para maximizar el acceso a puntos de interés cultural e histórico de la ciudad. Estas rutas guían al usuario por los principales monumentos, museos, plazas y otros sitios relevantes, ofreciendo una experiencia rica en conocimiento y exploración del patrimonio local.
- **Rutas sociales:** se centran en el paso por áreas de alta interacción social, como restaurantes, cafeterías y bares.

### 5.2.6. Definición de coeficientes en función del perfil de ruta

Para cumplir con los objetivos de cada perfil de ruta, se definen los siguientes esquemas de coeficientes en la tabla 5.25.

Perfil	length	width	noise_level	num_green	num_tourism	num_social
Rutas verdes	-1	0.5	-0.5	10	0.5	0.5
Rutas silenciosas	-1	0.5	-10	0.5	0.5	0.5
Rutas turísticas	-1	0.5	-0.5	0.5	10	0.5
Rutas sociales	-1	0.5	-0.5	0.5	0.5	10

Tabla 5.25: Esquema de coeficientes en función del perfil de ruta

### 5.2.7. Asignación de pesos

La selección de un perfil de ruta u otro determina el esquema de coeficientes a utilizar en el cómputo del peso de cada calle. A modo de ejemplo, se presentan los resultados obtenidos con la selección de perfil de ruta verde en la tabla 5.26.

u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight	
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.001111	0.272727	0.833333	0.035714	0.103448	0.162791	0.208848
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.003871	0.272727	0.833333	0.035714	0.137931	0.186047	0.234957
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.0.37527...	0.015944	0.090909	0.250000	0.035714	0.034483	0.744186	0.650988
...	...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	0.008890	0.090909	0.500000	0.035714	0.206897	0.255814	0.097977

Tabla 5.26: Fragmento del conjunto base tras calcular el peso asignado a cada calle

Ahora que ya se tienen los pesos, se deben ajustar —como se instruye en las subsecciones 4.2.6 y 4.2.6— para cumplir con los requisitos que impone el algoritmo de cálculo de la ruta. La tabla 5.27 muestra el conjunto base con los pesos actualizados.

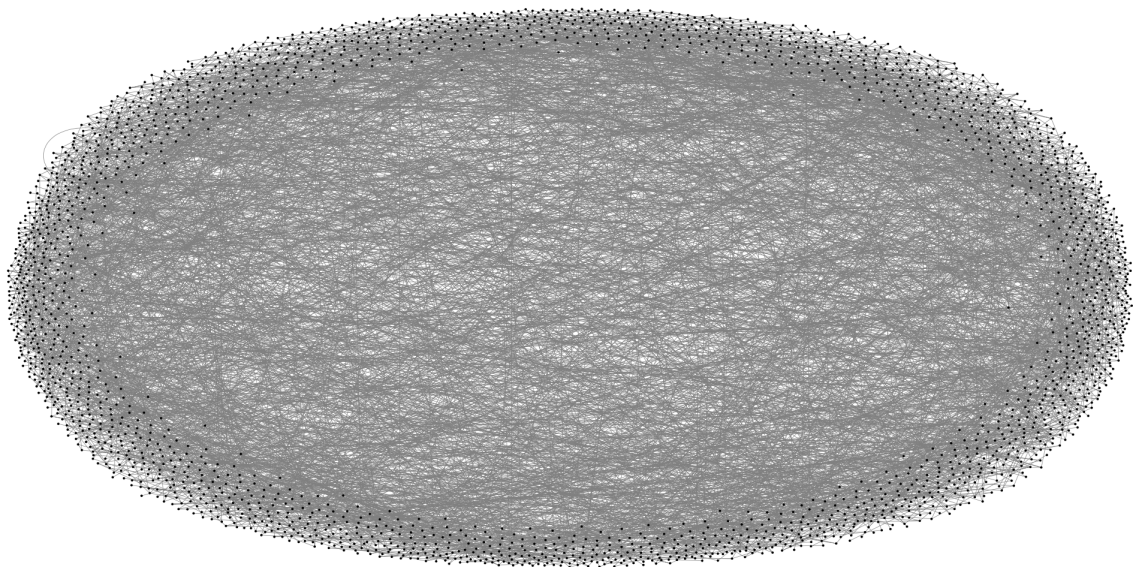
u	v	osmid	name	geometry	length	width	noise_level	num_green	num_tourism	num_social	weight	
0	25767290	11420770020	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.001111	0.272727	0.833333	0.035714	0.103448	0.162791	1.286369
1	25767290	524790982	803466250	Plaça de l'Ajuntament	LINestring (-0.37592...	0.003871	0.272727	0.833333	0.035714	0.137931	0.186047	1.244569
2	25767292	6766690060	5098789	Carrer de Martínez Cubells	LINestring (-0.0.37527...	0.015944	0.090909	0.250000	0.035714	0.034483	0.744186	0.819994
...	...	...	...	...	...	...	...	...	...	...	...	...
121897	25767296	25767297	5011431	Carrer del Periodista Azatti	LINestring (-0.37803...	0.008890	0.090909	0.500000	0.035714	0.206897	0.255814	1.059776

Tabla 5.27: Fragmento del conjunto base tras ajustar el peso asignado a cada calle

## 5.3 Carga de datos

Una vez determinado el peso definitivo para cada calle, se procede a la carga de los datos en una estructura de grafo no dirigido y ponderado, utilizando los pesos calculados. Para ilustrar la magnitud del grafo creado, se dispone un subgrafo formado por los 2500 primeros nodos (del total de 38936) y sus respectivas aristas en la figura 5.9.

El grafo se utiliza como parámetro en la función encargada de calcular el camino más corto (camino que, implícitamente, maximiza la suma de pesos de las aristas), utilizando el algoritmo de Dijkstra. Adicionalmente, se proporcionan como parámetros los nodos de origen y destino, derivados de las direcciones introducidas por el usuario.



**Figura 5.9:** Subgrafo representante de los 2500 primeros nodos de la red y las aristas que los unen

---

---

## CAPÍTULO 6

# Resultados

---

Finalizado el desarrollo de la herramienta generadora de rutas, se procede a la fase de prueba, en la cual se evaluarán los resultados obtenidos para algunos ejemplos de uso. Concretamente, se proponen las siguientes direcciones de origen y destino para la demostración:

1. **Ruta 1:** de Carrer de Xàtiva a Carrer del Governador Vell.
2. **Ruta 2:** de Plaça d'Artur Piera a Carrer del Beat Gaspar Bono.

Una de las acciones implícitas de la herramientas es la geocodificación. Al traducir las direcciones proporcionadas a coordenadas geográficas, se obtienen los siguientes pares de latitud y longitud:

1. **Ruta 1:** de (39.4674648, -0.3769474) a (39.4753963, -0.3709078).
2. **Ruta 2:** de (39.4703165, -0.3948204) a (39.4759246, -0.3872668).

Yendo un paso más allá, estas coordenadas se convierten a identificadores de nodos (*osmid*) de la red en cuestión, estructuras que ya pueden ser procesadas por el algoritmo que halla el camino:

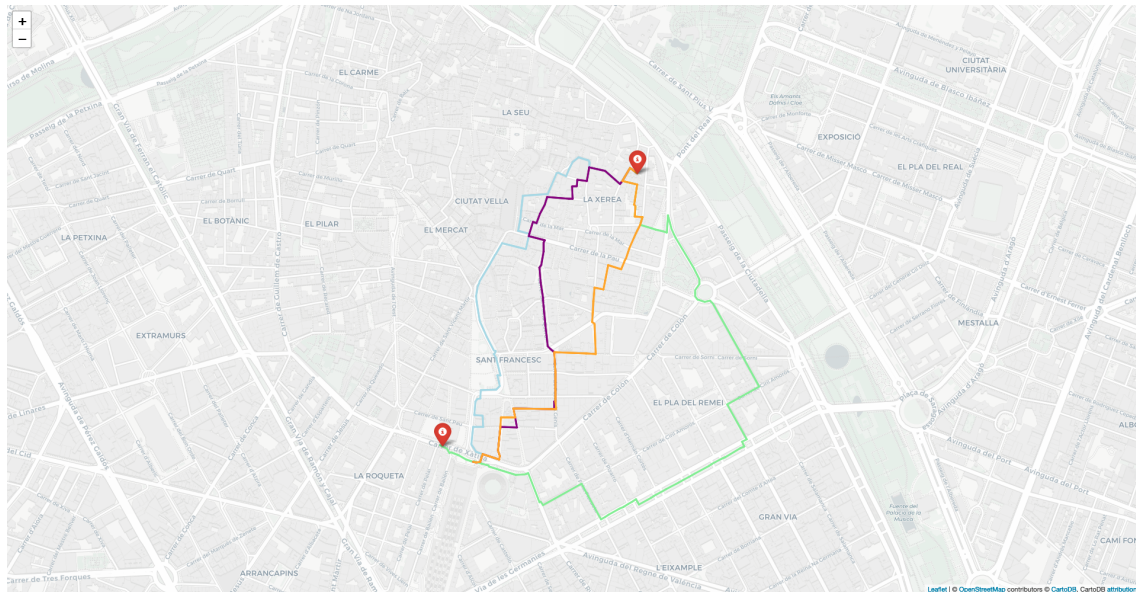
1. **Ruta 1:** de 33358134 a 60707971.
2. **Ruta 2:** de 10130674933 a 10963821043.

A pesar de tener el mismo origen y destino, el perfil de ruta determina el recorrido seguido para ir de un punto a otro. Dependiendo de la disponibilidad de datos (puntos de interés turístico, puntos de interés social, zonas verdes y niveles de ruido) en las zonas circundantes al origen y destino, las rutas de distintos perfiles pueden diferir totalmente, coincidir parcialmente o solaparse por completo.

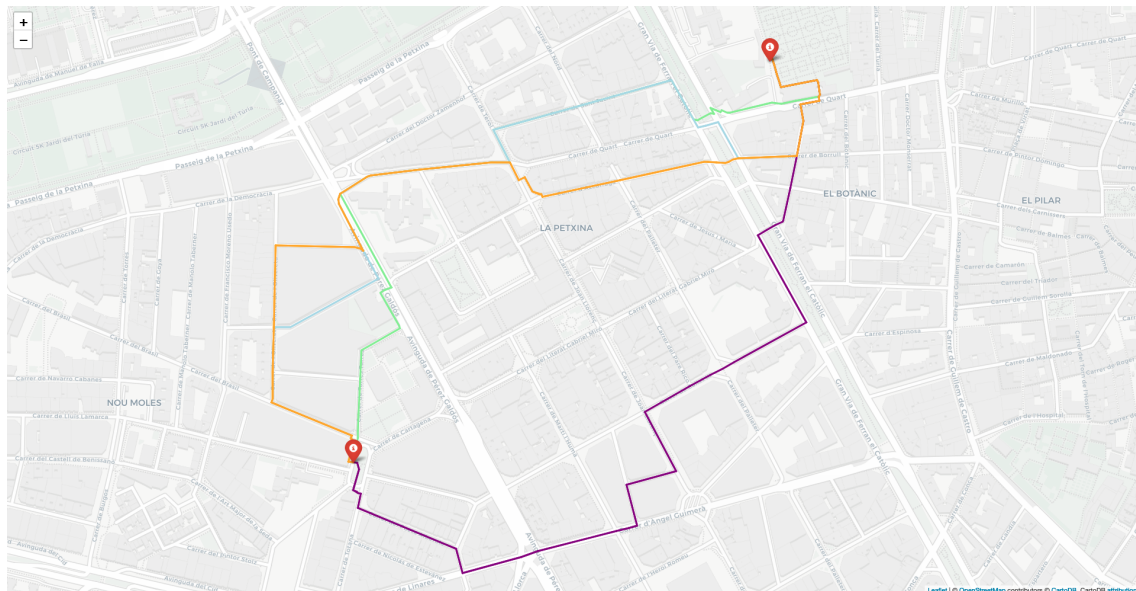
Las figuras 6.1a y 6.1b ilustran los caminos generados para la ruta 1 y la ruta 2, respectivamente. Se pueden observar las diferencias según el factor prioritario. Para facilitar la interpretación se aplica un código de colores: verde para las rutas verdes, azul para las rutas turísticas, morado para las rutas sociales y naranja para las rutas silenciosas. Los puntos de origen y destino se representan con marcadores rojos.

Si bien la existencia de recorridos distintos en ambas rutas confirma el efecto de los coeficientes definidos, es preciso validar también que estos recorridos responden a sus objetivos. Para ello, se grafica cada camino por separado, superponiendo los puntos de interés correspondientes en el mismo mapa.





(a) Caminos obtenidos para la ruta 1



(b) Caminos obtenidos para la ruta 2

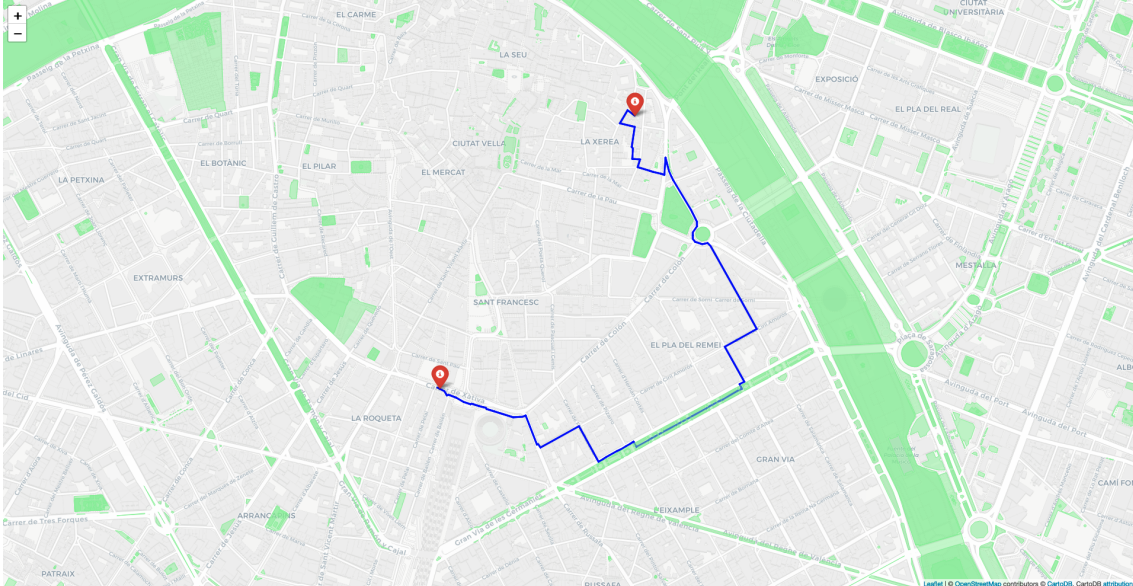
**Figura 6.1:** Representación de los distintos caminos generados en función del perfil de ruta

## 6.1 Validación de las rutas verdes

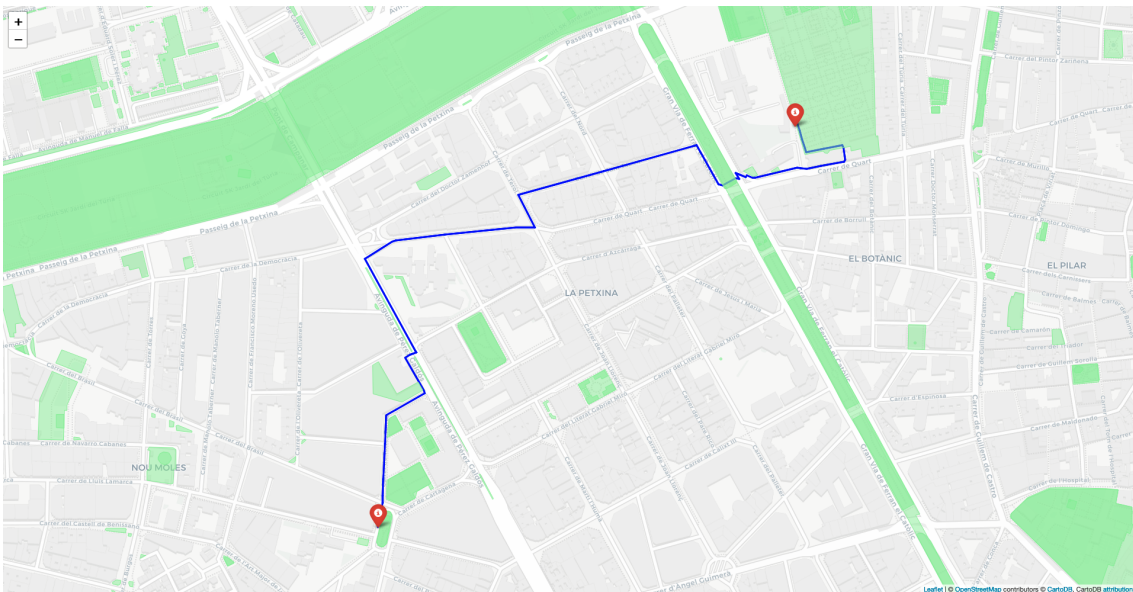
Cabe recordar que el perfil de ruta verde prioriza el contacto con jardines, parques y reservas naturales de la ciudad. Por tanto, se comprueba si, efectivamente los caminos generados pasan por el mayor número posible de zonas verdes (sin desviarse drásticamente de su destino).

Observando el ejemplo de la figura 6.2a se verifica que cumple su finalidad. Un camino recto entre el origen y el destino perdería supondría atravesar el casco antiguo de la ciudad de Valencia, mientras que la alternativa propuesta brinda la oportunidad de caminar bajo la zona arbolada de la Gran Vía. Al separarse de esta calle, también se puede apreciar su paso cercano al cauce del río y alrededor de los Jardines del Parterre.

Por lo que respecta al segundo ejemplo, representado en la figura 6.2b, parece que la zona abarcada atraviesa varios parques y jardines, especialmente al principio. De nuevo, frente a la posibilidad de un camino recto, se opta por bordear una gran avenida (Avenida de Pérez Galdós) debido a su alameda. Antes de llegar a su destino, también toma un pequeño desvío para pasar cerca de un par de parques alrededor del Jardín Botánico.



(a) Ejemplo 1: Carrer de Xàtiva a Carrer del Governador Vell



(b) Ejemplo 2: Plaça d'Artur Píera a Carrer del Beat Gaspar Bono

**Figura 6.2:** Ejemplos de caminos generados siguiendo el perfil de ruta verde

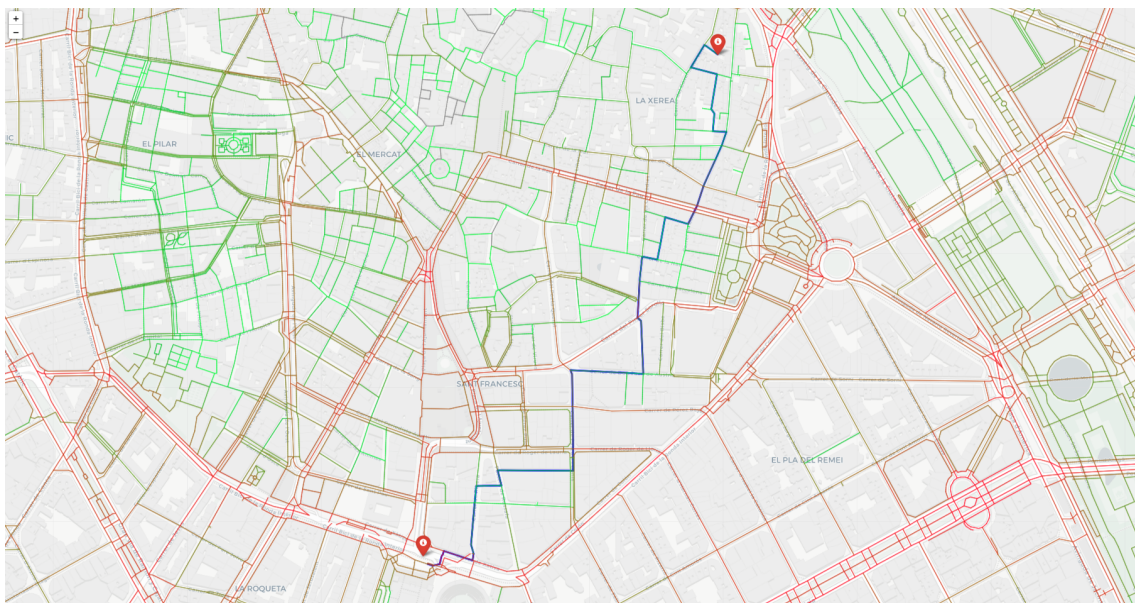
## 6.2 Validación de las rutas silenciosas

La adecuación de las rutas en base al nivel de ruido se puede juzgar fácilmente si se comprende el código de colores interno. Los tonos rojizos denotan niveles de ruido más alto que los verdosos. De forma opuesta a las rutas verdes previamente comentadas, las rutas silenciosas evitan las grandes vías.

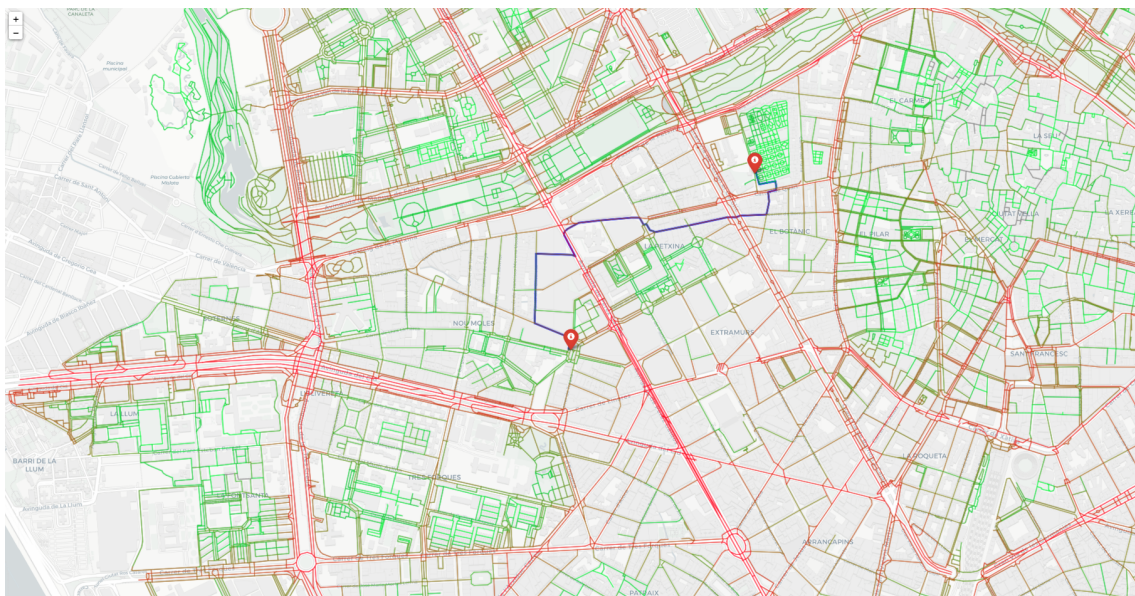


La figura 6.3a evidencia este rechazo a las avenidas: en lugar de recorrer la Gran Vía, como se hizo en la versión verde de la ruta, se observa un callejeo por el centro histórico de Valencia. Esto se debe a que esta es una zona de preferencia peatonal donde solos algunos vehículos autorizados tienen acceso, restricción que limita el nivel de ruido en comparación con las calles convencionales.

La zona abarcada por la ruta de la figura 6.3b se sitúa en el extrarradio de la ciudad antigua, lo cual dificulta el callejeo al no estar las calles tan próximas entre sí. No obstante, sí se observa que, para llegar de un punto al otro, es imperativo cruzar dos avenidas que registran altos niveles de ruido. Al no poder evitar esto, el algoritmo propone seguir otras calles y atravesarlas solo cuando sea estrictamente necesario, limitando el tiempo de exposición al ruido de estas grandes vías.



(a) Ejemplo 1: Carrer de Xàtiva a Carrer del Governador Vell



(b) Ejemplo 2: Plaça d'Artur Piera a Carrer del Beat Gaspar Bono

**Figura 6.3:** Ejemplos de caminos generados siguiendo el perfil de ruta silenciosa

---

## 6.3 Validación de las rutas turísticas

---

Como se ve a continuación, la interpretación de las rutas turísticas generadas tiene la dificultad añadida de integrar dos tipos de geometrías: puntos y polígonos. Por su tamaño, el paso próximo a los puntos no es tan evidente sobre el mapa, pero tiene el mismo valor que la proximidad a un polígono, pues ambos cuentan como un monumento turístico.

La primera ruta de ejemplo tiene lugar en el centro de la urbe valenciana, por lo que una ruta de este perfil es muy relevante. En la figura 6.4a se observa una mayor densidad de este tipo de monumentos, por lo que el camino al destino es más directo (al no requerir de desviaciones para pasar forzosamente junto a puntos de interés). Sin embargo, sí se pueden observar, en torno a la mitad del recorrido, algunos cambios de dirección repentinos en aras de aproximarse a espacios de interés turístico.

Desplazando la mirada a la figura 6.4b, se contempla una elección de ruta menos obvia, rodeando, en la primera mitad del recorrido, varios bloques de edificios grandes a fin de acercarse a algún punto de interés turístico. Durante la segunda mitad, a falta de polígonos, se pueden apreciar algunos puntos de interés en torno a la línea azul que representa la ruta.

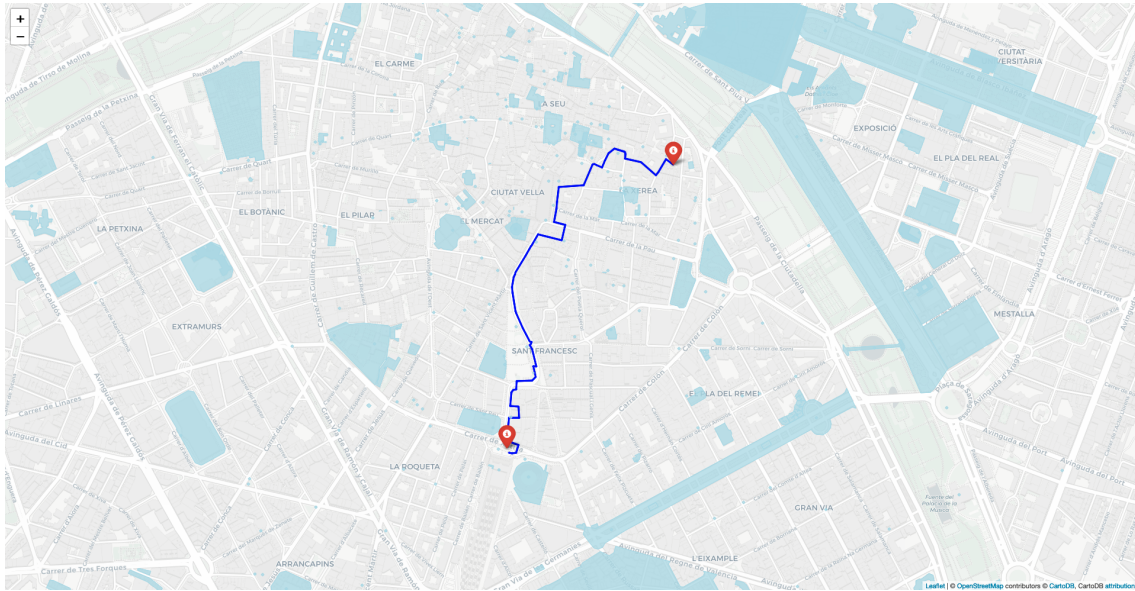
---

## 6.4 Validación de las rutas sociales

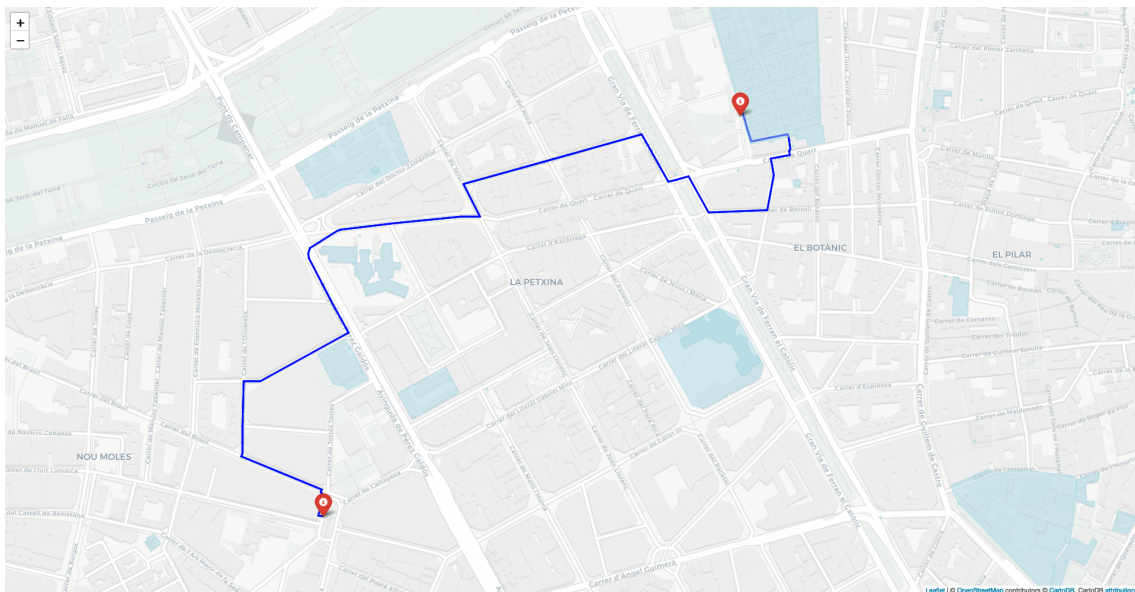
---

Si bien la presencia de puntos sobre el mapa dificultaba el análisis de las rutas turísticas, en el caso de las rutas sociales tiene el efecto contrario, pues al tomar casi todos los puntos de interés social esta forma, se conforman pequeñas nubes de puntos, candidatas a ser puntos de paso de las rutas generadas.

Y, en efecto, esto ocurre para la ruta mostrada en la figura 6.5a, como consecuencia de su entorno. Un efecto colateral del turismo masivo que experimenta el centro de Valencia es la aparición de bares y otros locales de ocio. Algunas calles albergan varios de ellos, incluso, como se ve en este mapa. Una vez más, la alta concentración de puntos disponibles produce una ruta más directa que la que se observa en la figura 6.5b. Este segundo camino, aunque probablemente registra un valor social total menor, tiene la ventaja de ofrecer una visualización más despejada: la falta de aglomeraciones de puntos permite distinguir más claramente los puntos que envuelven la línea azul dibujada. Sin embargo, sigue siendo perceptible la elección de calles sobre aquellas libres de ellos.



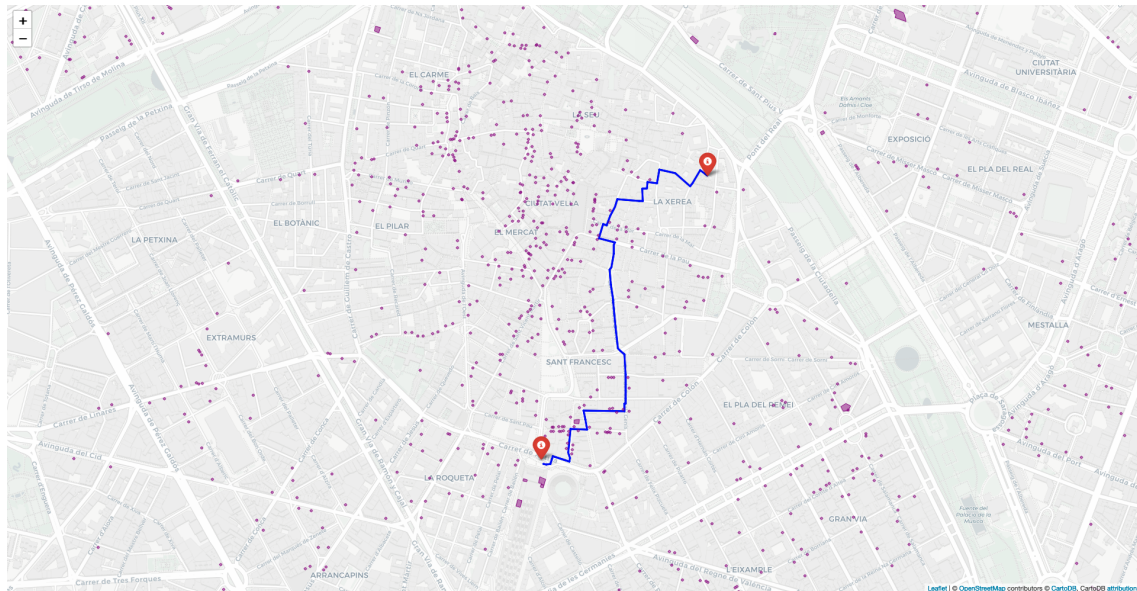
(a) Ejemplo 1: Carrer de Xàtiva a Carrer del Governador Vell



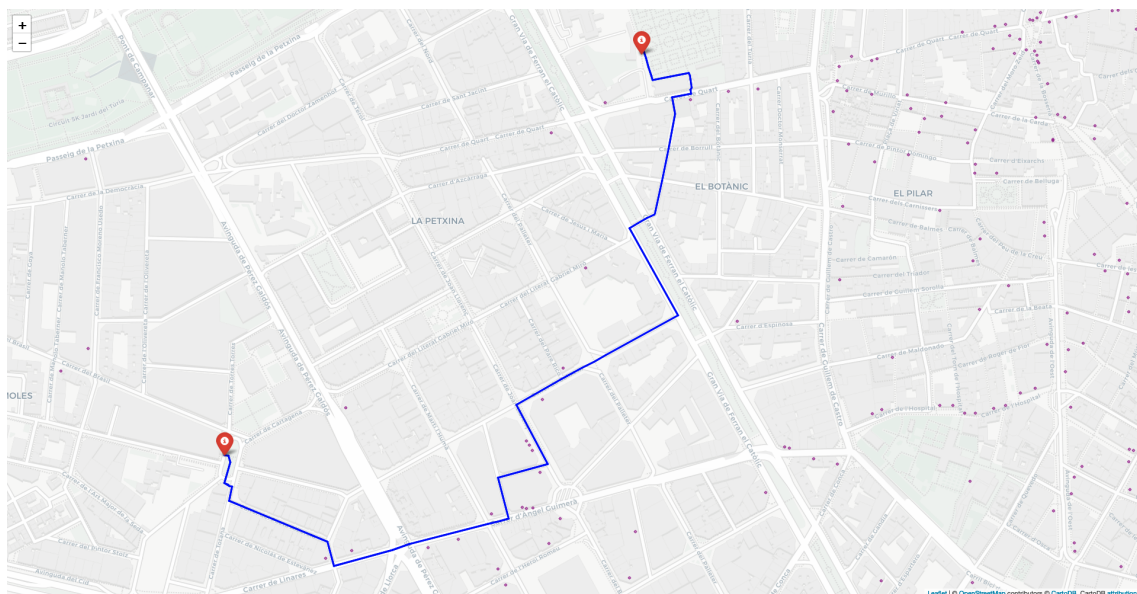
(b) Ejemplo 2: Plaça d'Artur Piera a Carrer del Beat Gaspar Bono

**Figura 6.4:** Ejemplos de caminos generados siguiendo el perfil de ruta turística





(a) Ejemplo 1: Carrer de Xàtiva a Carrer del Governador Vell



(b) Ejemplo 2: Plaça d'Artur Piera a Carrer del Beat Gaspar Bono

**Figura 6.5:** Ejemplos de caminos generados siguiendo el perfil de ruta social



---

---

## CAPÍTULO 7

# Conclusiones

---

El desarrollo de la herramienta generadora de rutas siguiendo la metodología propuesta ha dado como resultado una solución de alta calidad, lo que demuestra que esta metodología es una guía sólida y efectiva para proyectos similares. A lo largo del proyecto, se ha logrado cumplir con todos los objetivos establecidos, lo que demuestra su viabilidad.

En primer lugar, se ha conseguido desarrollar una herramienta funcional para el cálculo de rutas peatonales que utiliza exclusivamente datos abiertos y software libre, respetando así los principios de accesibilidad y reproducibilidad. La metodología estandarizada para el proceso de extracción, transformación y carga (ETL) de datos se ha mostrado robusta, facilitando la integración de diversas fuentes de información en un flujo de trabajo coherente y replicable.

Asimismo, se han integrado exitosamente puntos de interés y otras variables relevantes en el cálculo de rutas, lo que permite ofrecer resultados personalizados y ajustados a las necesidades del usuario. El uso de algoritmos de teoría de grafos para el cálculo de rutas ha sido clave para garantizar la eficiencia y precisión en la obtención de los caminos más óptimos entre dos puntos.

A pesar de los resultados positivos, se identifican algunas limitaciones inherentes al proyecto, principalmente relacionadas con los recursos disponibles, en términos de presupuesto, tiempo y, sobre todo, acceso a los datos. El catálogo de datos abiertos del Ayuntamiento de Valencia es algo limitado a día de hoy; la adición de nuevos conjuntos de datos podría impulsar la incorporación de nuevas funcionalidades en la herramienta. Además, se ha observado que la calidad de los datos depende en gran medida de la adherencia a estándares por parte de quienes los publican. En este sentido, los datos del Ayuntamiento presentaron inconsistencias que complicaron ligeramente su limpieza y procesamiento, mientras que los datos provenientes de OpenStreetMap, al ser más homogéneos y estandarizados, resultaron más fáciles de integrar y utilizar. Este aspecto subraya la importancia de promover y seguir estándares abiertos y uniformes para mejorar la calidad y utilidad de los datos en futuros proyectos.

Estas limitaciones han afectado, en cierta medida, al alcance del proyecto, y sugieren áreas de mejora y expansión en el futuro. Entre las propuestas de trabajo futuro, se contempla el desarrollo de una interfaz web que permita a los usuarios interactuar de manera más intuitiva con la herramienta, facilitando su acceso y uso. También se considera aumentar la transparencia en el proceso de cálculo de rutas, mostrando la generación de la ruta calle por calle. Esto permitiría a los usuarios comprender mejor las decisiones tomadas por el algoritmo, incrementando así la confianza en los resultados obtenidos.



# Bibliografía

---

- [1] Geoff Boeing. “OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks”. En: *Computers, Environment and Urban Systems* 65 (2017), págs. 126-139. ISSN: 0198-9715. DOI: <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>. URL: <https://www.sciencedirect.com/science/article/pii/S0198971516303970>.
- [2] *Designación València Capital Verde Europea #EUGreenCapital*. es-ES. URL: <https://www.valencia.es/-/capital-verda-europea> (visitado 02-08-2024).
- [3] E. W. Dijkstra. “A note on two problems in connexion with graphs”. en. En: *Numerische Mathematik* 1.1 (dic. de 1959), págs. 269-271. ISSN: 0945-3245. DOI: 10.1007/BF01386390. URL: <https://doi.org/10.1007/BF01386390> (visitado 29-08-2024).
- [4] *Folium — Folium 0.16.1.dev76+g2921126e documentation*. URL: <https://python-visualization.github.io/folium/latest/> (visitado 28-08-2024).
- [5] *geopandas.GeoDataFrame — GeoPandas 1.0.1+0.g747d66e.dirty documentation*. URL: <https://geopandas.org/en/stable/docs/reference/api/geopandas.GeoDataFrame.html> (visitado 18-08-2024).
- [6] Michael F. Goodchild. “Citizens as sensors: the world of volunteered geography”. en. En: *GeoJournal* 69.4 (ago. de 2007), págs. 211-221. ISSN: 1572-9893. DOI: 10.1007/s10708-007-9111-y. URL: <https://doi.org/10.1007/s10708-007-9111-y> (visitado 04-08-2024).
- [7] *Google Maps Usage Statistics*. URL: <https://trends.builtwith.com/mapping/Google-Maps> (visitado 02-08-2024).
- [8] John D. Hunter. “Matplotlib: A 2D Graphics Environment”. En: *Computing in Science & Engineering* 9.3 (mayo de 2007). Conference Name: Computing in Science & Engineering, págs. 90-95. ISSN: 1558-366X. DOI: 10.1109/MCSE.2007.55. URL: <https://ieeexplore.ieee.org/document/4160265> (visitado 28-08-2024).
- [9] *JSON*. URL: <https://www.json.org/json-en.html> (visitado 28-08-2024).
- [10] Kelsey Jordahl et al. *geopandas/geopandas: v0.8.1*. Jul. de 2020. URL: <https://doi.org/10.5281/zenodo.3946761> (visitado 02-08-2024).
- [11] *Map features - OpenStreetMap Wiki*. URL: [https://wiki.openstreetmap.org/wiki/Map\\_features](https://wiki.openstreetmap.org/wiki/Map_features) (visitado 02-08-2024).
- [12] *Map for UN Maps | UN Mappers*. URL: <https://maps.un.org/map-un-maps> (visitado 28-08-2024).
- [13] Eliseo Marzal y Laura Sebastia. “Solving the tourist trip design problem with time windows and variable profit using incremental local search”. En: *Applied Soft Computing* 155 (abr. de 2024), pág. 111399. ISSN: 1568-4946. DOI: 10.1016/j.asoc.2024.111399. URL: <https://www.sciencedirect.com/science/article/pii/S156849462400173X> (visitado 29-08-2024).



- [14] Mirtha Moran. *Objetivo 11: Lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles*. es. URL: <https://www.un.org/sustainabledevelopment/es/cities/> (visitado 02-08-2024).
- [15] Tessio Novack, Zhiyong Wang y Alexander Zipf. "A System for Generating Customized Pleasant Pedestrian Routes Based on OpenStreetMap Data". en. En: *Sensors* 18.11 (nov. de 2018). Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, pág. 3794. ISSN: 1424-8220. DOI: 10.3390/s18113794. URL: <https://www.mdpi.com/1424-8220/18/11/3794> (visitado 28-08-2024).
- [16] *Open Data Commons Open Database License (ODbL) — Open Data Commons: legal tools for open data*. URL: <https://opendatacommons.org/licenses/odbl/> (visitado 18-08-2024).
- [17] *Opendatasoft's Explore API Reference Documentation*. URL: <https://help.opendatasoft.com/apis/ods-explore-v2/> (visitado 28-08-2024).
- [18] *OpenStreetMap*. es. URL: <https://www.openstreetmap.org/> (visitado 01-08-2024).
- [19] *OSM TagFinder*. URL: <https://tagfinder.osm.ch/about> (visitado 20-08-2024).
- [20] *pandas.DataFrame — pandas 2.2.2 documentation*. URL: <https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html> (visitado 18-08-2024).
- [21] Age Poom, Joose Helle y Tuuli Toivonen. "Journey planners can promote active, healthy and sustainable urban travel". En: *Urbaria Summaries Series 2020/6*. University of Helsinki, sep. de 2020.
- [22] *Portal de Datos Abiertos del Ayuntamiento de València — Portal de l'Ajuntament de la ciutat de València*. URL: <https://valencia.opendatasoft.com/pages/home/> (visitado 08-08-2024).
- [23] *requests: Python HTTP for Humans*. URL: <https://requests.readthedocs.io> (visitado 08-08-2024).
- [24] Adam Rousell y Alexander Zipf. "Towards a Landmark-Based Pedestrian Navigation Service Using OSM Data". en. En: *ISPRS International Journal of Geo-Information* 6.3 (mar. de 2017). Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, pág. 64. ISSN: 2220-9964. DOI: 10.3390/ijgi6030064. URL: <https://www.mdpi.com/2220-9964/6/3/64> (visitado 28-08-2024).
- [25] Sean Gillies. *The Shapely User Manual*. Ago. de 2024. (Visitado 28-08-2024).
- [26] The pandas development team. *pandas-dev/pandas: Pandas*. Feb. de 2020. URL: <https://doi.org/10.5281/zenodo.3509134> (visitado 18-08-2024).
- [27] ONU Turismo. *Desarrollo Sostenible del Turismo | ONU Turismo*. URL: <https://www.unwto.org/es/desarrollo-sostenible> (visitado 04-08-2024).
- [28] *Welcome to GeoPy's documentation! — GeoPy 2.4.1 documentation*. URL: <https://geopy.readthedocs.io/en/stable/index.html> (visitado 28-08-2024).
- [29] World Tourism Organization (UNWTO), ed. *Recomendaciones de la OMT sobre turismo urbano*. World Tourism Organization (UNWTO), jul. de 2020. ISBN: 978-92-844-2203-6. DOI: 10.18111/9789284422036. URL: <https://www.e-unwto.org/doi/book/10.18111/9789284422036> (visitado 28-08-2024).

---

---

## APÉNDICE A

# Relación con el Grado en Ciencia de Datos

---

Esta porción de la memoria recoge consideraciones importantes para el desarrollo del trabajo presentado, que se enmarca en el contexto del Grado en Ciencia de Datos de la Universitat Politècnica de València.

### A.1 Análisis del marco legal y ético

---

El presente proyecto utiliza exclusivamente datos abiertos para su desarrollo, provenientes de fuentes públicas. El uso de datos abiertos no solo ofrece ventajas prácticas, como la facilidad de acceso y la reducción de costes, sino que también tiene un valor ético significativo. Al emplear datos que son accesibles a cualquier persona, sin restricciones basadas en factores sociales, económicos u otros, se promueve la igualdad de oportunidades en la investigación y el desarrollo tecnológico. Además, el uso de datos abiertos garantiza la reproducibilidad de los resultados, ya que cualquier interesado puede acceder a la misma información para verificar, replicar o ampliar los estudios realizados. Este enfoque contribuye a la transparencia, la inclusión y la democratización del conocimiento, valores fundamentales en la ciencia y la tecnología. De manera complementaria, las herramientas empleadas para el procesamiento de estos datos pertenecen a la categoría de software libre.

Además, es importante destacar que el proyecto se ha llevado a cabo sin la utilización de datos personales. La exclusión de información personal garantiza la protección de la privacidad y el cumplimiento de las normativas legales y éticas relacionadas con la gestión de datos sensibles.

### A.2 Relación con los estudios cursados

---

Como se espera de un trabajo de fin de grado, los conocimientos y habilidades adquiridas durante el Grado en Ciencia de Datos tienen un protagonismo patente en este proyecto. Los bloques de programación, gestión y visualización de datos destacan como los más influyentes. Más específicamente, las asignaturas con mayor relación son:

- Programación
- Matemática discreta
- Estructuras de datos

- Visualización
- Geospatial Data Processing

### **A.3 Relación con los Objetivos de Desarrollo Sostenible (ODS)**

El TFG “Optimización multifactorial de rutas peatonales entre dos puntos de la ciudad de Valencia” está alineado con varios Objetivos de Desarrollo Sostenible (ODS) establecidos por la Organización de las Naciones Unidas, contribuyendo a la agenda global para un desarrollo más equitativo y sostenible. A continuación, se detallan los ODS con los que el proyecto guarda una estrecha relación y cómo se vinculan con sus objetivos, mientras que la tabla A.1 resume el grado de relación de todos los ODS con el proyecto.

#### **A.3.1. Objetivo 3: Garantizar una vida sana y promover el bienestar para todos en todas las edades**

El sistema generador de rutas promueve la creación de rutas peatonales que priorizan la accesibilidad cultural, la reducción del ruido y la integración de espacios verdes. Estas características no solo facilitan un entorno urbano más saludable y agradable, sino que también fomentan un estilo de vida activo y el bienestar general de los ciudadanos al ofrecer opciones de movilidad que minimizan el impacto ambiental y promueven el disfrute del entorno urbano.

#### **A.3.2. Objetivo 10: Reducir la desigualdad en y entre los países**

El uso de datos abiertos en el proyecto está alineado con el ODS 10, que busca reducir las desigualdades. Al utilizar datos accesibles públicamente, el proyecto garantiza que la información necesaria para la optimización de rutas peatonales esté disponible para todos, independientemente de su situación económica o social. Esta apertura promueve la equidad en el acceso a la información y permite que diversas comunidades se beneficien de la mejora en la movilidad urbana al sistematizar los procedimientos mediante una metodología.

#### **A.3.3. Objetivo 11: Lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles**

Al optimizar las rutas peatonales en función de las preferencias de los ciudadanos, se fomenta la inclusividad y la adaptación a sus necesidades, contribuyendo a la creación de ciudades más inclusivas y sostenibles.

#### **A.3.4. Objetivo 12: Garantizar modalidades de consumo y producción sostenibles**

La promoción del uso de datos abiertos y software libre en detrimento de alternativas privadas asegura una mayor transparencia y accesibilidad en la gestión de la información, fomentando, asimismo, prácticas sostenibles en la producción y el consumo de los recursos tecnológicos.

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No procede
ODS 1. Fin de la pobreza				X
ODS 2. Hambre cero				X
ODS 3. Salud y bienestar	X			
ODS 4. Educación de calidad				X
ODS 5. Igualdad de género				X
ODS 6. Agua limpia y saneamiento				X
ODS 7. Energía asequible y no contaminante				X
ODS 8. Trabajo decente y crecimiento económico				X
ODS 9. Industria, innovación e infraestructuras				X
ODS 10. Reducción de las desigualdades		X		
ODS 11. Ciudades y comunidades sostenibles	X			
ODS 12. Producción y consumo responsables	X			
ODS 13. Acción por el clima				X
ODS 14. Vida submarina				X
ODS 15. Vida de ecosistemas terrestres				X
ODS 16. Paz, justicia e instituciones sólidas				X
ODS 17. Alianzas para lograr objetivos				X

**Tabla A.1:** Grado de relación de cada ODS con el proyecto