



Facultad de Ciencias Económicas y Empresariales  
ICADE

# **ANÁLISIS DE LA DEMANDA DEL SISTEMA DE BiciMAD EN FUNCIÓN DE VARIABLES METEOROLÓGICAS EN LA CIUDAD DE MADRID**

**TRABAJO FIN DE GRADO**

Autor: Juan Carlos Jimena Ordoñez  
Director: Dr. Carlos Miguel Vallez Fernández  
Clave: 202014388

## **AGRADECIMIENTOS**

*Quisiera expresar mi más sincero agradecimiento a mi tutor, el Dr. Carlos Miguel Vallez Fernández, por su apoyo, orientación y dedicación durante el desarrollo de este Trabajo de Fin de Grado. Su acompañamiento ha sido fundamental para la realización de este proyecto.*

*Asimismo, quiero agradecer de manera especial a mi familia y seres queridos, cuyo apoyo incondicional, ánimo constante y confianza en mí han sido pilares esenciales a lo largo de todo este proceso académico y personal.*

*A todos ellos, gracias de corazón.*

## **RESUMEN:**

Este trabajo realiza un análisis detallado del sistema de bicicletas compartidas de Madrid (BiciMAD), explorando cómo las condiciones meteorológicas influyen en los patrones de uso del servicio. A partir del registro de millones de viajes realizados entre 2021 y 2023, y de los datos meteorológicos proporcionados por la red de estaciones del Ayuntamiento de Madrid, se ha desarrollado un modelo de integración que asigna a cada trayecto las condiciones climáticas más cercanas en tiempo y espacio. La metodología incluye procesos de extracción, depuración y transformación de datos (ETL), así como análisis descriptivos y exploratorios para identificar tendencias estacionales, temporales y geográficas. Entre los hallazgos más relevantes se encuentran la fuerte estacionalidad de la demanda, el impacto de variables como la temperatura, la lluvia o el viento en la duración y frecuencia de los trayectos, y la relevancia del uso del sistema en días laborables frente a fines de semana. El estudio contribuye a una mejor comprensión del comportamiento de los usuarios de BiciMAD y ofrece información útil para la planificación operativa y la mejora de la sostenibilidad del sistema.

**Palabras clave:** bikesharing, micromovilidad, movilidad sostenible, meteorología urbana, datos abiertos, análisis espacial, BiciMAD, Madrid, patrones de uso y transporte.

## **ABSTRACT:**

This study presents a detailed analysis of the public bike-sharing system in Madrid (BiciMAD), focusing on how weather conditions influence user behavior and service demand. Using real trip data from 2021 to 2023 and meteorological records from Madrid's municipal weather stations, a spatial-temporal integration methodology was developed to assign the closest weather observations to each bike trip based on its starting location. The project follows an ETL process for data extraction, cleaning, and transformation, followed by exploratory and descriptive analyses. The findings reveal strong seasonal demand patterns, significant influence of variables such as temperature, rainfall, and wind speed on the frequency and duration of trips, and a higher intensity of use on weekdays compared to weekends. These insights contribute to a better understanding of user behavior in shared mobility systems and provide practical recommendations for improving operational planning and system efficiency in the face of environmental variability.

**Key words:** bikesharing, micromobility, sustainable mobility, urban meteorology, open data, spatial analysis, BiciMAD, Madrid, usage patterns, transportation.

## Índice de contenidos

1.	Introducción .....	8
1.1.	Justificación del proyecto .....	8
1.2.	Objetivos específicos .....	9
1.3.	Metodología .....	9
1.4.	Estado de la cuestión .....	10
2.	Análisis de la industria de bike-sharing.....	11
2.1.	Contexto histórico de los sistemas de bike-sharing .....	11
2.2.	Evolución de la industria en España .....	12
2.3.	El caso de BiciMad .....	14
4.	Metodología de extracción, exploración, limpieza y creación de variables .....	16
4.1.	Extracción de los datos .....	16
4.1.1.	Base de datos de BiciMAD .....	17
4.1.2.	Datos de estaciones meteorológicas de Madrid .....	17
4.2.	Exploración de los datos descargados.....	18
4.2.1.	Descripción de las bases de datos .....	18
4.3.	Análisis exploratorio de los conjuntos de datos.....	23
4.3.1.	Primeras líneas de las bases de datos.....	24
4.3.2.	Medidas estadísticas de las variables más relevantes .....	26
4.4.	Transformación de los conjuntos de datos .....	28
4.4.1.	Transformación de valores nulos.....	29
4.4.2.	Transformación de valores atípicos .....	30
4.4.3.	Transformaciones y creación de variables .....	32
5.	Análisis detallado de las bases de datos .....	33
5.1.	Análisis detallado de la base de datos de BiciMad.....	33
5.1.1.	Demanda del servicio de BiciMad por ID .....	34
5.1.2.	Demanda del servicio de BiciMad por mes y día de la semana .....	35
5.1.3.	Análisis del servicio de BiciMad por duración de viaje .....	36
5.1.4.	Análisis del servicio de BiciMad por tipo de día de la semana.....	38
5.1.5.	Distribución geográfica de la demanda de BiciMad.....	40
5.2.	Análisis detallado de la base de datos de estaciones meteorológicas.....	43
5.2.1.	Análisis geográfico de las estaciones meteorológicas .....	44
5.2.2.	Análisis de la disposición del dato de las magnitudes geográficas .....	45
5.3.	Análisis detallado de la base de datos de la información meteorológica diaria .....	46
5.3.1.	Evolución temporal de las variables meteorológicas relevantes.....	47
5.3.2.	Análisis de correlación entre variables relevantes.....	49
6.	Integración de datos .....	50
7.	Análisis de la base de datos unificada.....	52

7.1.	Análisis exploratorio del Dataset unificado .....	53
7.2.	Análisis detallado del Dataset unificado.....	55
7.2.1.	Evolución mensual de la demanda y variables meteorológicas .....	56
7.2.2.	Evolución mensual de la demanda y variables meteorológicas .....	58
7.2.3.	Especificación cuadrática del efecto de la temperatura.....	60
7.3.	Conclusiones del estudio .....	61
8.	Limitaciones del estudio .....	62
9.	Implicaciones prácticas y futuras líneas de investigación.....	63
9.1.	Implicaciones prácticas.....	64
9.2.	Futuras líneas de investigación.....	64
10.	Anexos.....	66
11.	Declaración de uso de herramientas de inteligencia artificial generativa en trabajos fin de grado .....	67
12.	Referencias.....	68

## Índice de tablas

Tabla 1: Resumen de los datasets.....	18
Tabla 2: Resumen de las variables de las bases de datos de BiciMad .....	19
Tabla 3: Resumen de las variables de la base de datos de estaciones meteorológicas fijas.....	21
Tabla 4: Resumen de las variables de la base de datos de meteorológicos diarios .....	22
Tabla 5: Rangos válidos establecidos para las magnitudes meteorológicas.....	31
Tabla 6: Rangos válidos establecidos para las magnitudes meteorológicas.....	33
Tabla 7: Clasificación de franjas horarias en días laborables.....	39

## Índice de ilustraciones

Ilustración 1: Evolución temporal del porcentaje de valores nulos de variables address_unlock y address_lock .....	20
Ilustración 2: Evolución temporal del porcentaje de valores nulos de idTrip .....	21
Ilustración 3: Primeras líneas de la base de datos de BiciMad .....	24
Ilustración 4: Primeras líneas de la base de datos de estaciones meteorológicas .....	25
Ilustración 5: Primeras líneas de la base de datos de datos meteorológicos .....	25
Ilustración 6: Estadísticas principales de los datasets de BiciMad .....	27
Ilustración 7: Estadísticas principales de los datasets de datos meteorológicos diarios.....	27
Ilustración 8: Distribución de trayectos con duración superior a 300 minutos.....	30
Ilustración 9: Resultados de las 5 bicis más usadas por año.....	34
Ilustración 10: Resultados de las 5 bicis menos usadas por año .....	34
Ilustración 11: Distribución de la demanda según el mes y el día de la semana.....	35
Ilustración 12: Diagrama de violín sobre la duración del viaje por año (en segundos) .....	36
Ilustración 13: Media anual de duración de los viajes (en minutos) .....	37
Ilustración 14: Porcentaje de viajes por tipo de día .....	38
Ilustración 15: Distribución porcentual del uso de BiciMAD por estación y tipo de día .....	39
Ilustración 16: Distribución de viajes en días laborables por franja horaria .....	40
Ilustración 17: Mapa de calor de densidad de origen viajes BiciMAD.....	41
Ilustración 18: Media diaria de desbloques por estación de BiciMAD en días laborables .....	42
Ilustración 19: Media diaria de desbloques por estación de BiciMAD en fines de semana .....	43
Ilustración 20: Distribución general de las estaciones meteorológicas fijas en el municipio de Madrid y su núcleo urbano.....	44
Ilustración 21: Magnitudes meteorológicas registradas por estación.....	45
Ilustración 22: Evolución diaria de la temperatura media (°C).....	47
Ilustración 23: Evolución diaria de la humedad relativa media (%).....	47
Ilustración 24: Evolución diaria de la velocidad media del viento (m/s) .....	48
Ilustración 25: Evolución diaria de la precipitación acumulada (l/m <sup>2</sup> ) .....	48
Ilustración 26: Matriz de correlación entre variables meteorológicas diarias .....	49
Ilustración 27: Modelo entidad-relación entre estaciones meteorológicas y datos horarios.....	51
Ilustración 28: Modelo entidad-relación entre datos meteorológicos y viajes de BiciMad .....	52
Ilustración 29: Exploración inicial del dataset final.....	54
Ilustración 30: Verificación de la estación meteorológica más cercana al punto de desbloqueo del viaje .....	55
Ilustración 31: Demanda mensual y temperatura media .....	56
Ilustración 32: Demanda mensual y precipitación.....	56
Ilustración 33: Demanda mensual y humedad relativa .....	57

Ilustración 34: Demanda mensual y velocidad del viento .....	57
Ilustración 35: Matriz de correlación entre la demanda diaria y variables meteorológicas .....	58
Ilustración 36: Resultados del modelo de regresión lineal múltiple con num_viajes como variable dependiente .....	59
Ilustración 37: Modelo de regresión polinómica sobre la temperatura.....	61

## ***Acrónimos***

BSS: Bike Sharing System – Sistema de bicicletas compartidas.

BSP: Bike Sharing Program – Programa de bicicletas compartidas.

COVID-19: Coronavirus Disease 2019 – Enfermedad causada por el virus SARS-CoV-2.

CSV: Comma-Separated Values – Formato de archivo para datos separados por comas.

DBK Informa: Base de Datos de Información Empresarial de Informa D&B – Fuente de información empresarial y de mercado.

EMT Madrid: Empresa Municipal de Transportes de Madrid – Entidad pública que gestiona el transporte urbano de Madrid.

ETL: Extraction, Transformation and Loading – Metodología para la extracción, transformación y carga de datos.

PBSC: Public Bicycle System Company – Empresa canadiense especializada en tecnología para sistemas de bicicletas compartidas.

R Studio: RStudio – Entorno de desarrollo integrado para el lenguaje de programación R.

WRI: World Resources Institute – Organización global de investigación para la sostenibilidad.

# 1. Introducción

## 1.1. Justificación del proyecto

La movilidad urbana ha experimentado un cambio notable en la última década, impulsado por la necesidad de transformar las ciudades en entornos más sostenibles, eficientes y menos dependientes del vehículo privado. En este proceso, los sistemas de micromovilidad han adquirido especial protagonismo, consolidándose como alternativas flexibles, económicas y adaptables para cubrir trayectos urbanos de corta distancia (Bozzi & Aguilera, 2021).

La micromovilidad engloba un amplio abanico de vehículos ligeros, de tracción humana o eléctrica, que permiten desplazamientos a velocidad limitada, con el objetivo de ofrecer soluciones de transporte eficaces en el ámbito urbano. Entre ellos, destacan las bicicletas compartidas, cuyo uso se ha incrementado notablemente en los últimos años en muchas capitales europeas, incluida Madrid (Oeschger et al., 2020).

En particular, los sistemas de bikesharing han emergido como una de las opciones preferidas por su simplicidad de uso, bajo coste y gran versatilidad. Al permitir utilizar una bicicleta sin necesidad de poseerla, reducen las barreras de entrada y favorecen la integración de este modo de transporte en la rutina diaria (DeMaio, 2009; Faghih-Imani & Eluru, 2015). Este modelo se alinea perfectamente con los principios de la economía compartida, que redefine la manera en que los ciudadanos acceden a bienes y servicios a través de plataformas digitales (Zhu et al., 2018).

El crecimiento económico del sector también es notable: se estima que el mercado global de la micromovilidad alcanzó un valor de 40.000 millones de dólares en 2021, y podría superar los 186.000 millones en 2030, impulsado por políticas públicas y privadas orientadas a la sostenibilidad (Acumen, 2023).

Más allá de su comodidad, los sistemas de bikesharing aportan beneficios significativos a nivel ambiental y urbano. Al integrarse en la movilidad cotidiana, ayudan a reducir los desplazamientos en vehículo privado, lo que conlleva una menor emisión de gases contaminantes y una descongestión del tráfico. Ejemplos como el de Shanghái han demostrado que la implantación masiva de bicicletas compartidas puede reducir de forma significativa tanto el consumo de petróleo como las emisiones de CO<sub>2</sub> (Zhang & Mi, 2018). Asimismo, investigaciones como la de Hamilton (2018) señalan que este tipo de sistemas puede disminuir hasta en un 4% el volumen de tráfico urbano.

A ello se suman los efectos positivos sobre la salud pública: el uso habitual de bicicletas compartidas favorece la actividad física, con beneficios en la salud física y mental, y se asocia con reducciones en la mortalidad prematura (Clockston & Rojas Rueda, 2021; Ricci, 2015).

En el caso de Madrid, el sistema BiciMAD, implementado en 2014, ha supuesto un avance significativo en la estrategia de movilidad sostenible de la ciudad. Su apuesta por una flota 100% eléctrica ha permitido ampliar la accesibilidad a zonas con orografía compleja, promoviendo un cambio de hábitos en muchos usuarios que han sustituido el uso del coche por la bicicleta en trayectos urbanos cotidianos (García-Palomares et al., 2022).

En este contexto, resulta especialmente relevante profundizar en el conocimiento de los factores que influyen en el uso de BiciMAD, no solo desde una perspectiva temporal o geográfica, sino también considerando aspectos ambientales. La meteorología, en particular, juega un papel clave en el comportamiento de los usuarios, condicionando tanto la frecuencia de uso como la duración de los trayectos. Sin embargo, son escasos los estudios que integran con precisión espacial y temporal esta dimensión en el análisis de servicios de bikesharing.

El presente trabajo aborda precisamente ese vacío mediante una metodología de integración de datos meteorológicos y de trayectos de BiciMAD. Para ello, se ha desarrollado un sistema de asignación meteorológica que vincula cada viaje

con las condiciones climáticas registradas en la estación más cercana al punto de desbloqueo de la bicicleta. Esto permite analizar con mayor profundidad cómo influyen factores como la temperatura, la lluvia, la humedad o el viento en la demanda y el patrón de uso del servicio.

De este modo, el estudio no solo contribuye al conocimiento académico sobre movilidad urbana, sino que ofrece información valiosa para la toma de decisiones en materia de planificación operativa, distribución de bicicletas, diseño de campañas de promoción y adaptación del sistema ante fenómenos climáticos adversos.

## **1.2. Objetivos específicos**

El propósito principal de este trabajo consiste en analizar los patrones de uso del sistema de bicicletas compartidas BiciMAD en la ciudad de Madrid, integrando para ello información detallada sobre las condiciones meteorológicas registradas en el momento y lugar de inicio de cada viaje. Para alcanzar este objetivo general, el estudio se plantea los siguientes objetivos específicos:

1. Realizar un proceso riguroso de extracción, depuración y transformación de datos procedentes del sistema BiciMAD y de la red municipal de estaciones meteorológicas de Madrid, garantizando la integridad, consistencia y trazabilidad de la información analizada.
2. Realizar un análisis exploratorio y descriptivo de los datos del sistema, para identificar tendencias estacionales, diferencias entre días laborables y fines de semana, variaciones en la duración de los viajes y patrones espaciales asociados al uso del servicio.
3. Integrar las condiciones meteorológicas más cercanas a cada trayecto mediante asignación espaciotemporal, para estudiar el impacto de variables como la temperatura, la lluvia, la humedad relativa o la velocidad del viento sobre la demanda del sistema y la duración de los trayectos.
4. Detectar anomalías o irregularidades en los registros de uso del sistema (duraciones atípicas, errores de geolocalización o registros inconsistentes), y analizar los efectos de posibles cambios estructurales en la plataforma durante el periodo 2021–2023.
5. Desarrollar una visualización gráfica clara, eficaz y accesible que facilite la comprensión de los resultados obtenidos, apoyando la extracción de conclusiones útiles para la mejora del servicio y la toma de decisiones por parte de las autoridades de movilidad.

## **1.3. Metodología**

Con el objetivo de alcanzar los propósitos planteados en este trabajo, la investigación se estructura bajo el enfoque ETL (Extraction, Transformation and Loading), una metodología ampliamente utilizada en el procesamiento y análisis de datos, que permite garantizar su correcta depuración, transformación y organización para su posterior interpretación.

En primer lugar, se procederá a la extracción de datos relativos a los desplazamientos realizados en el sistema de bicicletas compartidas BiciMAD, en la ciudad de Madrid, a través del portal oficial de datos abiertos del Ayuntamiento. Junto a estos datos de movilidad, se recopilarán también los registros meteorológicos históricos correspondientes a las distintas estaciones de medición climática de la ciudad. Estos últimos serán obtenidos a través de fuentes oficiales como la Agencia Estatal de Meteorología (AEMET) o portales públicos de información meteorológica.

Los datos extraídos se someterán a un proceso de limpieza, transformación y validación para asegurar su calidad, consistencia y coherencia. Se eliminarán registros incompletos o duplicados, se corregirán errores de codificación y se ajustarán valores atípicos si es necesario. Además, se transformarán las variables necesarias para unificar criterios y

permitir el cruce de información entre los desplazamientos en bicicleta y las condiciones meteorológicas. Este proceso se realizará con herramientas de análisis de datos y programación, utilizando el lenguaje Python.

Una de las tareas clave en esta fase será la asignación de la estación meteorológica más cercana a cada punto de recogida de bicicleta, de forma que a cada trayecto se le asocien las condiciones climáticas más representativas del entorno en el momento en que se inició el desplazamiento.

Una vez finalizado el tratamiento de los datos, se procederá a realizar un análisis exploratorio de datos, orientado a identificar patrones de comportamiento, variaciones en la demanda del servicio según las condiciones meteorológicas (temperatura, lluvia, viento, humedad, etc.) y posibles tendencias estacionales. Este análisis permitirá comprender de qué manera influyen las variables climáticas en el uso de BiciMAD en los distintos momentos del año y a lo largo del día.

Por último, se organizarán y visualizarán los resultados de manera clara y estructurada, utilizando gráficos, tablas y mapas que faciliten la interpretación de los hallazgos, permitiendo así extraer conclusiones fundamentadas y coherentes con los objetivos del trabajo.

#### **1.4. Estado de la cuestión**

En los últimos años, los sistemas de bicicletas compartidas (BSS, por sus siglas en inglés) se han consolidado como una de las alternativas más relevantes dentro de las estrategias de movilidad sostenible en España. Su implantación responde tanto a las crecientes demandas sociales de transporte eficiente y respetuoso con el medioambiente como a las políticas públicas orientadas a reducir la dependencia del vehículo privado en entornos urbanos. Según el Observatorio de la Bicicleta Pública en España, en el año 2018 se contabilizaban 53 sistemas públicos de bicicletas compartidas activos, una cifra que ha ido variando con el tiempo, reflejo de la naturaleza aún dinámica de este modelo de movilidad (Anaya & Douch, 2019).

Entre los diferentes sistemas en funcionamiento, el caso de Madrid resulta especialmente relevante por su dimensión, su progresiva electrificación y su integración en la estrategia de movilidad de la capital. BiciMAD, el servicio municipal gestionado por EMT ha experimentado una notable evolución desde su implantación en 2014, incorporando mejoras tecnológicas, ampliaciones territoriales y un crecimiento sostenido en el número de desplazamientos registrados. La incorporación de bicicletas eléctricas ha sido especialmente determinante, ya que permite superar barreras derivadas de la orografía urbana y ampliar el perfil de usuarios potenciales (Grau-Escolano et al., 2024).

En paralelo, diferentes estudios han puesto de manifiesto la influencia de múltiples factores contextuales sobre el uso de estos servicios, incluyendo las características urbanas, el perfil socioeconómico de los usuarios, y, de manera significativa, las condiciones meteorológicas. La climatología es un factor clave para entender las fluctuaciones en la demanda, afectando directamente a la frecuencia de uso y a la duración de los trayectos. Temperaturas extremas, precipitaciones, o la presencia de viento, pueden disuadir a los usuarios, modificar sus patrones de comportamiento, o alterar las decisiones sobre el medio de transporte elegido.

El impacto de la pandemia de COVID-19 también ha modificado las dinámicas de uso del bikesharing en Madrid. Si bien se registró una caída brusca durante los confinamientos, los datos muestran una recuperación progresiva conforme se relajaron las restricciones, lo que sugiere una notable resiliencia del sistema ante contextos de crisis y refuerza su potencial como solución urbana sostenible (Hu et al., 2022).

En este contexto, y con el objetivo de profundizar en el conocimiento de los determinantes reales del uso del sistema BiciMAD, el presente trabajo se centra en el análisis de los desplazamientos registrados en Madrid durante el periodo

2021–2023, combinando dicha información con datos meteorológicos obtenidos de las estaciones de medición más próximas a los puntos de recogida. A través de esta aproximación, se pretende identificar patrones de comportamiento vinculados a las condiciones climáticas, analizar su influencia sobre la demanda del servicio y generar evidencia útil para la mejora de la planificación, operación y promoción de la bicicleta compartida como alternativa de transporte urbano.

## **2. Análisis de la industria de bike-sharing**

En las últimas décadas, la movilidad urbana ha experimentado transformaciones profundas impulsadas por factores demográficos, tecnológicos y medioambientales. La creciente urbanización, la congestión del tráfico y la necesidad imperiosa de reducir las emisiones contaminantes han generado un escenario propicio para el desarrollo de alternativas de transporte sostenibles y eficientes. En este contexto, la micromovilidad, entendida como el conjunto de modos de transporte de corta distancia, ligeros y de bajo impacto ambiental, ha ganado protagonismo, siendo el bikesharing uno de sus exponentes más relevantes (Galatoulas et al., 2020).

El bikesharing, o sistema de bicicletas compartidas, se ha consolidado como una solución innovadora que combina la economía colaborativa con la movilidad sostenible. Este modelo permite a los usuarios acceder a bicicletas de forma temporal y flexible, facilitando desplazamientos urbanos rápidos, saludables y respetuosos con el medio ambiente. Además, contribuye a la descongestión del tráfico y complementa el transporte público, especialmente en la denominada “última milla”, es decir, el tramo final del viaje que conecta al usuario con su destino (WRI, 2018).

La industria del bikesharing ha evolucionado significativamente desde sus inicios, incorporando avances tecnológicos que han mejorado la experiencia del usuario y la eficiencia operativa. La integración de bicicletas eléctricas, sistemas de geolocalización, aplicaciones móviles y modelos de negocio diversificados ha permitido su expansión en múltiples ciudades a nivel global (Chen, Shaheen & Martin, 2018). En España, esta evolución ha sido particularmente notable, con un crecimiento sostenido en la implantación de sistemas públicos y privados, así como en la producción y comercialización de bicicletas, especialmente eléctricas.

Este capítulo tiene como objetivo analizar en profundidad la industria del bikesharing y su evolución en España, abordando su contexto histórico, la evolución de la industria en el país y las principales empresas implicadas en el sector. Para ello, se estructura en tres secciones principales: en primer lugar, se presenta una revisión histórica que permite comprender los orígenes y las etapas de desarrollo del bikesharing a nivel global; en segundo lugar, se examina la evolución de la industria del bikesharing en España, destacando sus cifras, tendencias y políticas públicas; y finalmente, se realiza un análisis de las principales empresas y actores que operan en el mercado español, sus modelos de negocio y estrategias competitivas.

### **2.1. Contexto histórico de los sistemas de bike-sharing**

El concepto de bikesharing surgió en el siglo XX, concretamente en la década de 1960, coincidiendo con un aumento del interés por el uso de la bicicleta. Este desarrollo propició la rápida expansión de los programas de bicicletas compartidas (BSP, por sus siglas en inglés) en diversas ciudades europeas (DeMaio, 2009; Shaheen, Guzman, & Zhang, 2010). La primera generación de estos sistemas apareció bajo la denominación de “Bicicletas Blancas” (White Bikes), que se implementaron en Ámsterdam en 1965. Sin embargo, este sistema tuvo una vida corta debido a que muchas bicicletas fueron arrojadas a los canales de la ciudad, lo que evidenció problemas relacionados con la falta de control y seguimiento de los vehículos (DeMaio, 2009; Shaheen et al., 2010; eldiario.es, 2022).

Posteriormente, en 1993, en Dinamarca, nació la segunda generación de programas de bikesharing con el sistema “Nakskov”, que contaba con 26 bicicletas distribuidas en cuatro estaciones. Este sistema también terminó cerrando debido a los robos de bicicletas, lo que puso de manifiesto la necesidad de incorporar mecanismos de rastreo y control más efectivos (DeMaio, 2009; Shaheen et al., 2010). Este problema fue el punto de partida para el desarrollo de la tercera generación de sistemas, que integraron tecnologías de telecomunicaciones, acceso mediante smartphones, tarjetas inteligentes, quioscos y sistemas informatizados para la gestión y control del servicio (Shaheen, Guzman, & Zhang, 2012).

Durante la década de 2000, los sistemas de tercera generación, como “Call a Bike” en Múnich y “Vélo’v” en Francia, experimentaron un crecimiento sostenido y se extendieron a países fuera de Europa, incluyendo Brasil, China y Estados Unidos (DeMaio, 2009; Shaheen et al., 2012). Finalmente, la cuarta generación de programas de bikesharing se caracteriza por la incorporación de sistemas inteligentes que buscan maximizar la sostenibilidad, eficiencia y calidad del servicio, integrando tecnologías avanzadas como sensores, geolocalización y aplicaciones móviles para facilitar el acceso y uso de las bicicletas (DeMaio, 2009; Shaheen et al., 2010; Tang, Pan, & Shen, 2011).

En la actualidad, los sistemas de bicicletas compartidas se han difundido rápidamente en todo el mundo durante más de 20 años, consolidándose como una solución clave para la movilidad urbana sostenible (Tang et al., 2011). Cada dos años, el índice Copenhagenize clasifica las 20 ciudades más amigables con la bicicleta y sostenibles, evaluando criterios que incluyen la presencia y calidad de los programas de bikesharing (Design, 2014).

Respecto a los modos de operación, los sistemas de bikesharing se dividen principalmente en dos categorías: con estaciones fijas (station-based) y sin estaciones fijas (dock-less). Por ejemplo, en Estados Unidos, los sistemas dock-less como Mobike, Ofo y Spin comenzaron a operar en 2017. Aunque el 44 % de las bicicletas compartidas en ese país son dock-less, el 96 % de los viajes se realizan con bicicletas de sistemas basados en estaciones (Guide, 2011). Los sistemas dock-less presentan problemas asociados, como robos, vandalismo y falta de seguimiento de las bicicletas, por lo que su estudio se considera un tema de investigación separado (Sun, 2018).

Por otro lado, los sistemas con estaciones fijas disponen de numerosas estaciones y bicicletas listas para su uso, lo que facilita a los usuarios recoger y devolver las bicicletas en diferentes puntos de la ciudad. La “conveniencia” es el principal factor que motiva el uso de la bicicleta compartida, incluyendo aspectos como la accesibilidad a las estaciones, la simplicidad en los procedimientos de afiliación y pago, así como la disponibilidad de cascos (Leister, Vairo, Sims, & Bopp, 2018; Zanotto, 2014). En los últimos años, la literatura ha incorporado estudios orientados a incrementar el uso del bikesharing y a identificar los factores que facilitan la adopción de la bicicleta como medio de transporte urbano (Fishman, Washington, & Haworth, 2012; Hampshire & Marla, 2012).

## **2.2. Evolución de la industria en España**

La industria del bikesharing en España ha experimentado un crecimiento notable en la última década, consolidándose como un componente clave dentro de las estrategias de movilidad urbana sostenible que adoptan muchas ciudades para reducir la dependencia del vehículo privado y las emisiones asociadas al transporte motorizado. El auge de estos sistemas en el contexto español ha estado condicionado tanto por factores tecnológicos, como la electrificación de las flotas y la digitalización de las plataformas de alquiler, como por el impulso de políticas públicas que promueven un modelo de ciudad más accesible y menos contaminante.

España inició la implantación de sistemas de bicicleta pública a mediados de la década de 2000, con experiencias piloto en ciudades como Córdoba y Gijón, orientadas inicialmente al ocio y con flotas reducidas. En 2007, se produjo un despegue significativo con el lanzamiento de Bicing en Barcelona y Sevici en Sevilla, que lograron integrar la bicicleta

en el sistema de transporte urbano gracias a la construcción simultánea de redes de carriles bici. Posteriormente, otras ciudades como Zaragoza, Valencia y Santander sumaron sus propios sistemas, consolidando la presencia del bikesharing en ciudades medianas para 2010 (Ministerio de Transportes y Movilidad Sostenible, s.f.; Sanmiguel-Rodríguez, 2022).

Entre 2011 y 2014, la expansión continuó a otras ciudades, aunque algunas localidades pequeñas enfrentaron dificultades financieras y técnicas que llevaron al cierre o reinversión de sus sistemas. La entrada de Madrid en 2014 con BiciMAD, el primer sistema de bicicletas eléctricas asistidas a gran escala en España marcó un hito importante, consolidándose posteriormente con gestión municipal directa y ampliación progresiva (EMT Madrid, 2021). La pandemia de COVID-19 impactó temporalmente el uso, pero también impulsó políticas a favor del uso de la bicicleta y la electrificación de flotas, con un aumento significativo de bicicletas eléctricas en ciudades como Barcelona, Sevilla y Madrid (Ministerio para la Transición Ecológica, 2023).

El desarrollo de la bicicleta como medio de transporte urbano en España ha evolucionado en paralelo a la creciente concienciación social sobre sostenibilidad y salud pública, así como a la necesidad de buscar alternativas al tráfico rodado en entornos cada vez más saturados. Según el último Barómetro de la Bicicleta en España, impulsado por la Red de Ciudades por la Bicicleta en colaboración con la Dirección General de Tráfico, el uso habitual de la bicicleta ha crecido hasta alcanzar al 22,8% de la población española, mientras que un 48,2% ha utilizado la bicicleta en alguna ocasión (RedBici, 2024). Este incremento, sostenido desde 2008, pone de manifiesto la consolidación de la bicicleta como medio de transporte habitual en las ciudades, en buena parte favorecido por la expansión de los sistemas de bikesharing.

En términos de oferta, los datos recabados por DBK Informa señalan que en abril de 2024 la flota total de bicicletas compartidas en España alcanzaba las 42.110 unidades, representando aproximadamente el 85% de todos los vehículos de movilidad urbana compartida disponibles en el país, superando en volumen a opciones como los patinetes eléctricos (DBK Informa, 2024). Cataluña concentra más de un tercio de esta flota, impulsada principalmente por el sistema Bicing en Barcelona, mientras que Madrid y Valencia completan el podio nacional tanto en volumen de flota como en número de trayectos anuales registrados.

En cuanto al análisis académico de estos sistemas, algunas investigaciones han puesto de relieve que la electrificación ha sido uno de los factores decisivos para garantizar la adopción de las bicicletas compartidas en ciudades españolas, en especial aquellas con topografías que dificultan el uso de bicicletas convencionales. Grau-Escolano et al. (2024) destacan la relevancia de las nuevas tecnologías aplicadas al mantenimiento predictivo y a la gestión eficiente de las bicicletas compartidas en Barcelona, donde la automatización de estas tareas ha permitido optimizar la disponibilidad de vehículos y la eficiencia del sistema en su conjunto.

A pesar de su expansión, la industria del bikesharing en España enfrenta ciertos retos que condicionan su crecimiento futuro. Uno de los desafíos más relevantes ha sido la aparición de alternativas de micromovilidad como los patinetes eléctricos, que han generado una competencia directa en entornos urbanos, especialmente en trayectos cortos. Este fenómeno, junto con el auge del teletrabajo y la gratuidad del transporte público en algunas ciudades, ha provocado un ligero retroceso en el uso de la bicicleta en los últimos años, según informaba recientemente El País (2024).

Además, la infraestructura ciclista sigue siendo un factor limitante en muchas ciudades, donde la falta de carriles bici seguros o interconectados desincentiva a nuevos usuarios.

Desde una perspectiva económica, las previsiones apuntan a que la industria del bikesharing en España seguirá creciendo, gracias tanto a la electrificación de las flotas como a la expansión de sistemas híbridos que combinan

estaciones fijas con modelos dockless, adaptándose así a las necesidades cambiantes de las ciudades y de los usuarios. Según DBK Informa (2024), el mercado español de movilidad urbana compartida generó una facturación conjunta de aproximadamente 155 millones de euros en 2023, con perspectivas de crecimiento sostenido en los próximos años.

En este contexto, el sistema BiciMAD en Madrid se posiciona como una de las iniciativas más consolidadas y de mayor volumen operativo del país. Su carácter 100% eléctrico, su expansión territorial y su integración con la red de transporte público lo convierten en un referente dentro del ecosistema de micromovilidad urbana. Estas características, junto con la creciente disponibilidad de datos abiertos, hacen de Madrid un caso de especial interés para el estudio del impacto de factores externos (como la meteorología) sobre el comportamiento de los usuarios.

En conclusión, la industria del bikesharing en España se encuentra en un estado de madurez creciente, con una adopción social consolidada en las principales ciudades, pero también enfrenta retos significativos que requieren la adaptación de sus modelos de negocio y de las políticas públicas para garantizar su viabilidad y su integración definitiva como parte del sistema de transporte urbano.

### **2.3. El caso de BiciMad**

BiciMAD es el sistema público de bicicletas compartidas de la ciudad de Madrid, inaugurado el 23 de junio de 2014. En sus inicios contaba con 1.560 bicicletas eléctricas y 123 estaciones distribuidas en seis distritos centrales (Centro, Salamanca, Arganzuela, Retiro, Moncloa-Aravaca y Chamberí). Inicialmente la operación del servicio fue concedida a la empresa Bonopark, pero en 2016 su gestión pasó a manos del Ayuntamiento de Madrid a través de la Empresa Municipal de Transportes (EMT). Esta municipalización buscó mejorar la estabilidad financiera y operativa del sistema tras algunos problemas iniciales de mantenimiento y disponibilidad de bicicletas. Desde entonces, BiciMAD se ha consolidado como una iniciativa clave de movilidad sostenible en la capital, experimentando una expansión gradual de cobertura y capacidad a lo largo de la última década.

Entre 2015 y 2020, el número de estaciones y bicicletas de BiciMAD creció significativamente para atender la demanda en aumento. El servicio pasó de las 123 estaciones originales a 165 estaciones en 2015 y 207 en 2017, extendiéndose a distritos adicionales más allá del centro. Para 2020, BiciMAD ya contaba con 257 estaciones activas (+50 respecto a 2017) distribuidas en unos 15 distritos de la ciudad, con una flota cercana a las 4.000 bicicletas operativas (Empresa Municipal de Transportes de Madrid, 2022). Esta ampliación progresiva permitió acercar el servicio a nuevos barrios y usuarios, aunque la cobertura seguía concentrada principalmente en la zona central de Madrid.

En términos de uso, BiciMAD superó los 3,49 millones de viajes anuales en 2020, pero enfrentó retos en 2021 debido a circunstancias extraordinarias. En enero de 2021, la intensa nevada causada por la borrasca Filomena colapsó Madrid, dejando el sistema BiciMAD inoperativo durante varias semanas y dañando cientos de bicicletas (baterías, cuadros, sistemas eléctricos) así como numerosos anclajes y terminales en las estaciones. La EMT tardó más de un mes en restablecer por completo el servicio, evidenciando la vulnerabilidad de la infraestructura original ante fenómenos climáticos extremos. Consecuentemente, el año 2021 cerró con 3.445.232 viajes, una ligera disminución (-1,15% respecto a 2020) atribuida tanto al impacto de Filomena como a cambios de patrón modal durante la pandemia. A pesar de ello, la tendencia general de uso se mantuvo estable en torno a 3,5 millones de viajes anuales durante 2021 y 2022, señal de una demanda constante pero todavía limitada por la capacidad del sistema previo a su gran expansión. Cabe destacar que en este periodo la EMT implementó mejoras de seguridad, como la instalación de candados en bicicletas y cámaras de vigilancia en estaciones, ante un aumento de robos reportado en 2020-2021.

En 2022, el Ayuntamiento de Madrid anunció un ambicioso plan de renovación integral y ampliación de BiciMAD, financiado con 48,8 millones de euros (incluyendo 30 millones de fondos europeos Next Generation EU). El proyecto,

implementado en 2023, actualizó la tecnología del sistema: nuevas bicicletas, estaciones renovadas y una aplicación móvil mejorada. El objetivo central era escalar el servicio a toda la ciudad con equipamiento de última generación, corrigiendo las deficiencias del antiguo sistema (Movilidad Eléctrica, 2023).

La transformación de 2023 implicó también la mayor expansión territorial en la historia de BiciMAD. Entre marzo y junio de 2023 se llevó a cabo la instalación progresiva de cientos de nuevos puntos de préstamo: el sistema pasó de las 264 estaciones y ~3.000 bicis disponibles al inicio de 2023, a 611 estaciones operativas y 7.500 bicicletas hacia mediados de año. Esta expansión permitió que BiciMAD, antes circunscrito a áreas centrales y algunos barrios, alcanzara por primera vez la totalidad de los 21 distritos de Madrid, incluyendo zonas periféricas que hasta entonces carecían del servicio.

Durante la transición, que implicó la convivencia temporal de los dos sistemas, surgieron problemas operativos, diversas fuentes reportaron estaciones vacías o fuera de servicio, bicicletas abandonadas o vandalizadas y dificultades técnicas que incluso bloquearon temporalmente las cuentas de algunos usuarios. Este desorden transitorio evidenció los retos logísticos de reemplazar el sistema en vivo a gran escala. En respuesta, el Ayuntamiento impulsó medidas para incentivar la adopción del nuevo BiciMAD pese a las molestias iniciales. La más notable fue la gratuidad temporal del servicio: desde el lanzamiento del nuevo sistema (7 de marzo de 2023) todos los viajes de hasta 30 minutos fueron gratuitos durante varios meses. Al inicio la gratuidad se fijó hasta el 31 de julio de 2023, pero después la Junta de Gobierno decidió prorrogarla hasta fin de año 2023 dado el éxito de la medida.

Por primera vez BiciMAD ofrecía viajes subvencionados de forma generalizada durante casi diez meses consecutivos, lo que permitió familiarizar a la ciudadanía con el renovado sistema y compensar las molestias de la transición. En paralelo, se introdujo una nueva estructura tarifaria para la fase post-gratuidad, eliminando la distinción previa entre abonos anuales y usuarios ocasionales. Desde 2024 todos los usuarios pasan a tener un “contrato básico” sin cuota fija anual, pagando solo por uso (0,50 € por cada tramo de 30 min, similar al esquema anterior). Adicionalmente, se ofrece una tarifa plana mensual voluntaria (10 € al mes) que permite realizar viajes ilimitados de hasta 30 minutos sin coste por viaje. Este modelo mixto (pago por uso vs tarifa plana) busca flexibilidad y adaptación a distintos perfiles de usuario, y asegura una fuente de ingresos sostenible tras concluir el periodo promocional gratuito (Empresa Municipal de Transportes de Madrid, 2022).

Los impactos de la transformación de 2023 sobre el desempeño del sistema han sido notables. A finales de ese año, con el nuevo BiciMAD ya plenamente operativo, se registró un crecimiento exponencial en el uso: 7,61 millones de viajes en 2023, lo que supone un incremento del +120,6% respecto al año anterior. Este aumento extraordinario se explica tanto por la expansión de la cobertura geográfica y la ampliación de flota, como por la política de gratuidad que disparó la demanda. De hecho, en noviembre de 2023 BiciMAD alcanzó un récord histórico de 57.116 viajes en un solo día, cifra nunca antes vista.

No obstante, esta profunda transformación tuvo una consecuencia directa sobre el presente trabajo: la discontinuidad en los datos abiertos del sistema. Hasta febrero de 2023, el Ayuntamiento publicaba de forma mensual microdatos anonimizados de todos los viajes realizados. Sin embargo, la migración tecnológica interrumpió esta serie histórica. A partir de marzo, los datos dejaron de estar disponibles, lo que impide la continuidad del análisis en los mismos términos (Empresa Municipal de Transportes de Madrid, 2023). Como consecuencia, el análisis desarrollado en este trabajo se ha visto obligado a limitar su alcance entre enero de 2021 y febrero de 2023.

### **3. Hipótesis de investigación**

Numerosos estudios han puesto de manifiesto que las condiciones meteorológicas constituyen uno de los factores contextuales más relevantes a la hora de explicar las variaciones en la demanda de sistemas de bicicletas compartidas. A diferencia de otros modos de transporte más protegidos frente al entorno físico, la bicicleta está directamente expuesta a las condiciones ambientales, lo que convierte al clima en un elemento determinante del comportamiento de los usuarios.

En este sentido, diversos trabajos han demostrado que variables como la temperatura, la precipitación, la humedad relativa o la velocidad del viento influyen significativamente en la decisión de utilizar una bicicleta compartida. Por ejemplo, El-Assi et al. (2017) constataron que tanto las temperaturas extremas como los episodios de lluvia afectan de forma negativa al uso del sistema en Toronto, observando también un mayor uso en condiciones moderadas. De forma similar, Gebhart y Noland (2014) analizaron el sistema de bikesharing de Washington D.C. y encontraron que la probabilidad de uso disminuye notablemente en días lluviosos, mientras que temperaturas agradables la incrementan. Rixey (2013), por su parte, en un estudio comparado de varias ciudades de EE.UU., concluye que el clima es uno de los predictores más sólidos de la demanda a nivel de estación, junto con factores sociodemográficos y urbanísticos.

La evidencia empírica disponible sugiere, por tanto, que el comportamiento de los usuarios no solo responde a factores estructurales como la localización de las estaciones o la oferta del sistema, sino también a condiciones dinámicas y variables como el clima, que pueden modificar sustancialmente el patrón de uso en función del momento y del contexto.

A la luz de estos antecedentes, el presente trabajo plantea como hipótesis central que la demanda del sistema BiciMAD está significativamente influida por las condiciones meteorológicas. En particular, se espera que variables como la temperatura, la precipitación, la humedad relativa y la velocidad del viento presenten una relación estadísticamente significativa con el volumen diario de viajes registrados, y que dicha relación sea consistente con los patrones observados en la literatura internacional.

Para contrastar esta hipótesis, se ha llevado a cabo un proceso de integración espaciotemporal entre los datos meteorológicos registrados por estación y los trayectos iniciados en sus proximidades. A partir de esta base enriquecida, se han aplicado diferentes técnicas de análisis descriptivo, exploración gráfica y estimación econométrica, incluyendo el cálculo de coeficientes de correlación y la especificación de modelos de regresión lineal múltiple. Este enfoque permite no solo observar patrones agregados de comportamiento, sino también evaluar con precisión la magnitud e intensidad del efecto que cada variable atmosférica ejerce sobre la demanda, una vez controlados otros factores.

Con ello, se pretende aportar una visión rigurosa y contextualizada del papel que desempeña el clima en la movilidad ciclista urbana en Madrid, así como generar evidencia útil para la planificación operativa del sistema, la previsión de la demanda y el diseño de políticas públicas de movilidad sostenible más adaptativas frente a la variabilidad climática.

## **4. Metodología de extracción, exploración, limpieza y creación de variables**

Después de contextualizar la temática, en esta sección se detalla el procedimiento técnico seguido para obtener y preparar los datos que sustentan el análisis. Se abordan las fuentes de datos utilizadas (extracción), las tareas de análisis preliminar y limpieza realizadas, y la definición de nuevas variables derivadas que permiten responder a las preguntas de investigación.

### **4.1. Extracción de los datos**

En este trabajo se han utilizado dos tipos principales de fuentes de datos: por un lado, los registros de uso del sistema de bicicletas públicas BiciMAD, y por otro, diversas bases de datos meteorológicas del municipio de Madrid. Ambas proceden del portal de datos abiertos del Ayuntamiento de Madrid y han sido combinadas para poder analizar la relación entre el uso del servicio y las condiciones meteorológicas. A continuación, se describen en detalle ambas fuentes, así como el proceso seguido para su obtención y posterior integración.

#### **4.1.1. Base de datos de BiciMAD**

Los datos utilizados para el análisis de los desplazamientos en bicicleta eléctrica en la ciudad de Madrid proceden del sistema público BiciMAD, cuya información está disponible de forma abierta en el portal de datos del Ayuntamiento de Madrid. Estos archivos, que recogen los viajes realizados por los usuarios del servicio, se publican en formato CSV y están organizados de manera mensual.

Para este estudio, se han recopilado los ficheros disponibles correspondientes al periodo comprendido entre el año 2019 y febrero de 2023, siendo esta última fecha el punto más reciente accesible en el momento de la descarga, realizada durante el primer semestre de 2025.

Dado que la estructura y codificación de los archivos varía ligeramente de unos meses a otros, fue necesario desarrollar un proceso de integración previa de los datos. Esta tarea se llevó a cabo utilizando Python, mediante un script que automatiza la lectura y procesamiento de todos los ficheros seleccionados. El procedimiento incluyó la normalización de nombres de columnas, el tratamiento de codificaciones de texto, y la consolidación de todos los registros mensuales en un único conjunto de datos global.

Durante la integración, también se prestó especial atención al periodo de transición del sistema BiciMAD que tuvo lugar a lo largo de 2023. Esta renovación supuso cambios significativos como la introducción de nuevas estaciones y bicicletas, así como un reinicio de los identificadores utilizados en los registros. Por este motivo, fue necesario realizar algunos ajustes adicionales para asegurar la consistencia temporal y espacial del conjunto consolidado, especialmente en lo relativo a la identificación de estaciones y la continuidad de los datos históricos.

#### **4.1.2. Datos de estaciones meteorológicas de Madrid**

Para incorporar el contexto climático al análisis de la movilidad en bicicleta, se utilizaron varias fuentes meteorológicas disponibles en el portal de datos abiertos del Ayuntamiento de Madrid. A diferencia de la base de datos de BiciMAD, en este caso fue necesario combinar dos conjuntos de datos diferentes: uno de carácter estático, con la localización de las estaciones meteorológicas, y otro dinámico, que contiene los registros diarios de observaciones.

En primer lugar, se descargó un fichero con la información georreferenciada de las estaciones meteorológicas situadas en la ciudad. Esta base recoge datos de identificación y ubicación que permiten ubicar con precisión cada estación en el entramado urbano. Su incorporación fue esencial para poder emparejar, en etapas posteriores, los registros de movilidad con la estación meteorológica más próxima al lugar donde se inició cada viaje.

A continuación, se recopilaron los archivos mensuales que contienen las observaciones meteorológicas diarias en cada una de estas estaciones. Estos datos están organizados por día y por estación, y recogen distintas magnitudes registradas a lo largo de las 24 horas del día.

Dado que el volumen y estructura de esta información resultaba muy extenso para el análisis directo, se procedió a realizar un trabajo de transformación y simplificación. Este incluyó la creación de una única columna de fecha a partir

de las variables de año, mes y día, la identificación de cada estación mediante códigos comunes con la base de localización, y el cálculo de valores promedio diarios por magnitud y estación. Además, se consolidaron todos los archivos mensuales en un único conjunto de datos integrado, y se depuraron aquellos registros con formatos incorrectos o datos ausentes.

## 4.2. Exploración de los datos descargados

Tras detallar el proceso seguido para la recopilación e integración de las distintas bases de datos, en este apartado nos centraremos en examinar el contenido de la información recopilada. Para ello, dividiremos el análisis en dos bloques diferenciados: por un lado, se describirán las principales variables presentes en cada conjunto de datos; por otro, se llevará a cabo una primera exploración de los patrones y características generales que revelan los datos obtenidos.

### 4.2.1. Descripción de las bases de datos

Tras recoger la información necesaria para desarrollar el análisis, el siguiente paso consistió en visualizar y explorar los distintos conjuntos de datos. El objetivo de esta revisión inicial era comprobar si las estructuras de los archivos eran consistentes entre sí, identificar posibles diferencias en su formato, y comenzar a familiarizarnos con las variables incluidas en cada uno.

Para este proceso se empleó la herramienta Python, que permitió automatizar la carga y examen de los distintos ficheros. En primer lugar, se generó un resumen básico de cada conjunto de datos, detallando el número de registros y variables que contiene cada uno. Esta revisión resultó especialmente útil para detectar algunas inconsistencias estructurales entre archivos de diferentes años, como ligeras variaciones en la codificación o la presencia de columnas adicionales en determinados meses, aspectos que se abordarán más adelante.

En el caso de los datos meteorológicos, se comprobó también la distribución geográfica y temporal de las observaciones disponibles, lo que permitió verificar su cobertura dentro del municipio de Madrid y su compatibilidad con los registros de trayectos del sistema BiciMAD. Este paso fue clave para garantizar que ambos conjuntos de datos pudieran integrarse posteriormente de forma robusta en el análisis.

La siguiente tabla resume los principales conjuntos de datos utilizados, indicando para cada uno el número total de observaciones y variables:

**Tabla 1: Resumen de los datasets**

Nombre	N.º de observaciones	N.º de variables
Dataset BiciMad 2021	2.137.194	23
Dataset BiciMad 2022	4.142.143	23
Dataset BiciMad 2023	464.341	23

Nombre	N.º de observaciones	N.º de variables
Estaciones meteorológicas fijas	26	23

Nombre	N.º de observaciones	N.º de variables
Dataset Meteorológico 2021	17.774	60

Dataset Meteorológico 2022	26.510	60
Dataset Meteorológico 2023	4.405	60

**Fuente: elaboración propia**

Tal y como refleja la tabla, el conjunto de datos de BiciMAD se componía inicialmente de archivos mensuales, los cuales fueron unificados y agrupados por año para facilitar su análisis. Esta consolidación permitió observar diferencias notables en el número de observaciones disponibles: en 2021 se contabilizan algo más de dos millones de trayectos, en 2022 más de cuatro millones, y en 2023 apenas medio millón. Estas diferencias se explican por la cobertura temporal de cada bloque: en 2021 solo se dispone de datos desde junio, en 2022 el registro es completo durante todo el año, y en 2023 solo se pueden recuperar datos de enero y febrero. Además, durante el proceso de depuración se identificó que algunos archivos presentaban errores graves de estructura, lo que los hacía inservibles para el análisis. Es el caso del archivo correspondiente a octubre de 2021, que no pudo ser incorporado al dataset<sup>1</sup> final por sus problemas de estructura.

Por su parte, el dataset de las estaciones meteorológicas fijas contiene información descriptiva y geográfica de los 26 puntos de medida distribuidos por la ciudad, incluyendo variables como la dirección, altitud, coordenadas o tipo de estación. Este conjunto es esencial para contextualizar espacialmente los datos meteorológicos diarios.

En cuanto a los datos meteorológicos diarios, estos también se componían originalmente de archivos mensuales, unidos y agrupados posteriormente por año para facilitar su análisis conjunto. El resultado muestra un elevado número de variables (60 en total), debido a que las condiciones atmosféricas se registran 24 veces por día (una por cada hora), lo que genera múltiples columnas por cada magnitud. Cada registro diario incluye, además de la fecha, variables como temperatura, humedad relativa, velocidad del viento, presión barométrica y precipitación, entre otras. Esta granularidad horaria permitirá posteriormente analizar la influencia de factores meteorológicos puntuales en la demanda de bicicletas, más allá de promedios diarios o tendencias generales.

**Tabla 2: Resumen de las variables de las bases de datos de BiciMad**

Variable	Tipo de Variable	Definición
fecha	Fecha	Fecha en formato texto que indica el día en que se realizó el trayecto.
idBike	Numérica	Identificador numérico de la bicicleta utilizada.
fleet	Numérica	Código de la flota a la que pertenece la bicicleta. Puede hacer referencia al modelo o generación.
trip_minutes	Numérica	Duración del trayecto en minutos.
geolocation_unlock	Texto (diccionario)	Coordenadas en formato JSON del punto en el que se desbloqueó la bicicleta. Incluye latitud y longitud.
address_unlock	Texto	Dirección postal o aproximada donde se desbloqueó la bicicleta.
unlock_date	Fecha y hora	Fecha y hora exacta en la que el usuario desbloqueó la bicicleta para iniciar el viaje.
locktype	Texto	Tipo de cierre usado al finalizar el trayecto (ej. automático, manual).

<sup>1</sup> Un Dataset es un conjunto estructurado de datos, generalmente en formato tabular, utilizado para su análisis o procesamiento computacional.

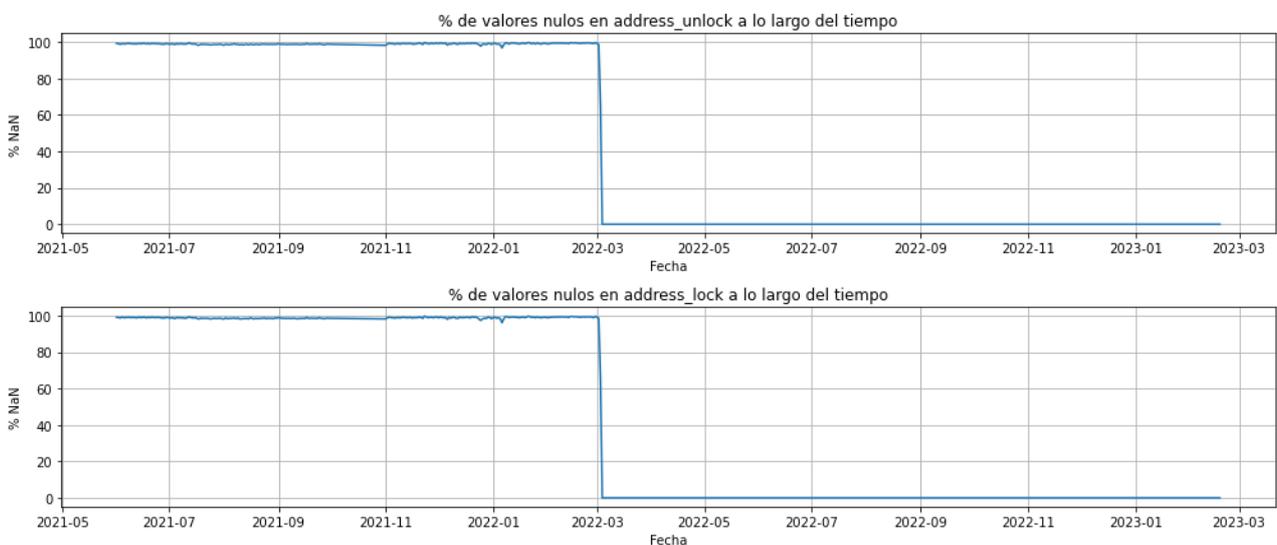
unlocktype	Texto	Tipo de desbloqueo utilizado (ej. app móvil, tarjeta, QR).
geolocation_lock	Texto	Coordenadas en formato texto tipo JSON del punto donde se bloqueó la bicicleta. Incluye latitud y longitud.
address_lock	Texto	Dirección aproximada donde se finalizó el trayecto.
lock_date	Fecha y hora	Fecha y hora en la que se bloqueó la bicicleta. Requiere transformación a formato datetime.
station_unlock	Texto	Código alfanumérico de la estación donde se inició el trayecto. Puede incluir valores nulos o inconsistentes.
dock_unlock	Numérica	Número de anclaje (dock) en el que estaba la bici al iniciarse el trayecto.
unlock_station_name	Texto	Nombre de la estación donde se inició el trayecto.
station_lock	Numérica	Código numérico de la estación donde se finalizó el trayecto.
dock_lock	Numérica	Número de anclaje (dock) donde se dejó la bicicleta al finalizar el trayecto.
lock_station_name	Texto	Nombre de la estación donde se finalizó el trayecto.
idTrip	Texto	Código alfanumérico único del trayecto realizado. En algunos casos se repite.

Fuente: elaboración propia

Como puede apreciarse en la tabla anterior, el conjunto de datos de BiciMAD incluye información detallada sobre cada trayecto realizado, desde identificadores de bicicleta y estaciones hasta coordenadas geográficas y duraciones. En el análisis exploratorio se identificaron aspectos relevantes de la calidad y disponibilidad de los datos.

En primer lugar, algunas variables no están de forma uniforme en todo el periodo analizado. Es el caso, por ejemplo, de `address_unlock` y `address_lock`, que recogen la dirección textual aproximada donde se inició y finalizó el trayecto. Estas variables presentan un porcentaje casi total de valores ausentes hasta febrero de 2022, momento desde el que comienzan a estar sistemáticamente registradas. Todo indica que estas columnas fueron incorporadas al sistema de recogida de datos a partir de marzo de 2022, lo que explica su comportamiento discontinuo en los registros más antiguos.

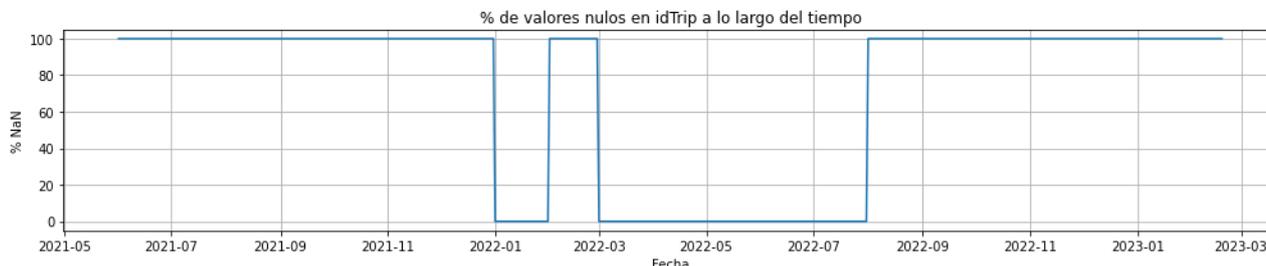
Ilustración 1: Evolución temporal del porcentaje de valores nulos de variables `address_unlock` y `address_lock`



Fuente: elaboración propia

Otro caso destacable es el de la variable idTrip, pensada para identificar cada trayecto de forma única. Su presencia es irregular a lo largo del tiempo: aparece en algunos meses y desaparece en otros sin un patrón aparente. Esta intermitencia impide emplearla como clave robusta en el análisis, aunque se ha mantenido en el dataset para conservar su trazabilidad cuando está disponible.

**Ilustración 2: Evolución temporal del porcentaje de valores nulos de idTrip**



**Fuente: elaboración propia**

Asimismo, las variables geolocation\_unlock y geolocation\_lock contienen las coordenadas de inicio y fin del trayecto, respectivamente, pero en un formato complejo (tipo JSON en texto) que impide su análisis directo. Fue necesario procesar estos campos para extraer de ellos la latitud y longitud de cada punto, generando así dos nuevas variables numéricas denominadas lon\_bici y lat\_bici, que sí pueden ser utilizadas en análisis geoespaciales o de correspondencia con estaciones meteorológicas.

Por último, antes de integrar el dataset en el análisis, se realizó una limpieza final eliminando aquellas filas que presentaban valores nulos en variables consideradas esenciales: duración del viaje (trip\_minutes), fecha (fecha) y coordenadas (lat\_bici y lon\_bici). Además, se descartaron los registros cuya duración del trayecto era negativa, por tratarse de datos erróneos o mal registrados.

Gracias a estas acciones, se obtuvo una base de datos coherente y estructurada, preparada para su integración con el resto de las fuentes de información. En la siguiente sección se describe el segundo conjunto de datos empleado en el análisis: las estaciones meteorológicas fijas del municipio de Madrid, que servirán como referencia para ubicar espacialmente las mediciones atmosféricas diarias.

**Tabla 3: Resumen de las variables de la base de datos de estaciones meteorológicas fijas**

Variable	Tipo de Variable	Definición
codigo_estacion	Numérica	Identificador numérico único asignado a cada estación meteorológica.
CÓDIGO_CORTO	Texto	Abreviatura o código simplificado de la estación. Suele usarse para visualizaciones o tablas resumidas.
ESTACION	Texto	Nombre completo de la estación meteorológica.
DIRECCION	Texto	Dirección física donde se ubica la estación.
LONGITUD_ETRS89	Numérica	Longitud geográfica de la estación en el sistema de referencia ETRS89.
LATITUD_ETRS89	Numérica	Latitud geográfica de la estación en el sistema ETRS89.
ALTITUD	Numérica	Altura sobre el nivel del mar en metros del punto donde está instalada la estación.
VV (81)	Numérica	Velocidad del viento (código de magnitud 81) medida en m/s o km/h, según el sistema.

DV (82)	Numérica	Dirección del viento (código 82), expresada en grados respecto al norte.
T (83)	Numérica	Temperatura del aire (código 83), normalmente en grados Celsius.
HR (86)	Numérica	Humedad relativa (código 86), expresada en porcentaje.
PB (87)	Numérica	Presión barométrica (código 87), medida en hPa.
RS (88)	Numérica	Radiación solar (código 88), expresada en W/m <sup>2</sup> .
P (89)	Numérica	Precipitación (código 89), acumulada en mm.
COD_VIA	Texto	Código del tipo de vía en la que se ubica la estación (calle, avenida, etc.).
VIA_CLASE	Texto	Clase de vía urbana según nomenclatura municipal.
VIA_PAR	Texto	Información sobre si la estación se encuentra en lado par o impar de la vía.
VIA_NOMBRE	Texto	Nombre de la vía (calle o avenida) donde está situada la estación.
NUM_VIA	Texto o numérica	Número de portal o ubicación en la vía donde se encuentra la estación.
COORDENADA_X_E TRS89	Numérica	Coordenada X (este-oeste) en el sistema de referencia ETRS89.
Longitud	Numérica	Longitud geográfica en formato estándar WGS84.
Latitud	Numérica	Latitud geográfica en formato estándar WGS84.

**Fuente: elaboración propia**

El segundo conjunto de datos empleado en este análisis corresponde a las estaciones meteorológicas fijas del municipio de Madrid. Este dataset no contiene registros temporales, sino información estática y descriptiva sobre la ubicación, características y capacidades de medición de cada estación. Entre los campos incluidos se encuentran identificadores, coordenadas geográficas (en diferentes sistemas de referencia), altitud, dirección física y tipo de vía. Asimismo, incorpora una serie de columnas numéricas que indican qué variables meteorológicas puede medir cada estación.

Este último punto es particularmente relevante, ya que no todas las estaciones registran todas las magnitudes atmosféricas. Para cada una de ellas, se incluyen campos como VV (81) para velocidad del viento, DV (82) para dirección del viento, T (83) para temperatura, HR (86) para humedad relativa, PB (87) para presión barométrica, RS (88) para radiación solar y P (89) para precipitación. La presencia o ausencia de valores en estas columnas permite identificar si la estación realiza o no esa medición concreta. Esta información será clave en el análisis posterior para vincular estaciones con registros meteorológicos específicos.

Dado que existe una gran diversidad de variables posibles y con el objetivo de enfocar el análisis en aquellas más directamente relacionadas con la movilidad urbana y la experiencia del usuario, se ha decidido seleccionar cuatro magnitudes principales para el análisis: velocidad del viento (81), temperatura (83), humedad relativa (86) y precipitación (89). Estas variables han sido elegidas tanto por su disponibilidad en la mayoría de las estaciones como por su potencial influencia en la decisión de utilizar o no el sistema BiciMAD.

A continuación, se describe el tercer conjunto de datos utilizado en este trabajo: el registro diario de datos meteorológicos observados por las estaciones activas, que permite vincular información atmosférica concreta con la fecha y ubicación de cada trayecto.

**Tabla 4: Resumen de las variables de la base de datos de meteorológicos diarios**

Variable	Tipo de Variable	Definición
----------	------------------	------------

PROVINCIA	Texto	Provincia en la que se ubica la estación. En este análisis, siempre corresponde a Madrid.
MUNICIPIO	Texto	Municipio de ubicación de la estación (Madrid).
ESTACION	Texto	Nombre completo de la estación meteorológica que realiza la observación.
MAGNITUD	Numérica	Código de la variable atmosférica registrada (por ejemplo, 81 para velocidad del viento, 83 para temperatura, etc.).
PUNTO_MUESTREO	Texto	Código que combina estación y magnitud medida, utilizado como identificador único del punto de muestreo.
ANO	Numérica	Año en el que se ha realizado la observación.
MES	Numérica	Mes de la observación.
DIA	Numérica	Día del mes correspondiente a la observación.
H01 – H24	Numérica	Valores horarios de la magnitud registrada para cada una de las 24 horas del día.
V01 – V24	V/N	Indicadores de validación de los valores horarios. Su función es indicar la calidad, consistencia y fiabilidad del dato medido. Los valores que comienzan por “V” indican datos válidos; los que comienzan por “N”, no válidos.
codigo_estacion	Numérica	Identificador numérico de la estación meteorológica.

**Fuente: elaboración propia**

El tercer conjunto de datos utilizado en este trabajo corresponde a los registros diarios de variables meteorológicas recogidas por las estaciones fijas del municipio de Madrid. Este dataset presenta una estructura muy particular: cada fila representa una combinación única de estación, magnitud atmosférica y fecha, mientras que los valores horarios correspondientes a esa magnitud se distribuyen en 24 columnas (una por cada hora del día). Adicionalmente, se incluyen otras 24 columnas que actúan como indicadores de calidad o validación de cada valor horario.

Para poder integrar este conjunto de datos con el resto de las fuentes utilizadas en el análisis, fue necesario realizar varias transformaciones. En primer lugar, se procedió a unificar el formato de fechas. Mientras que en otros datasets ya existía una variable temporal clara (unlock\_date, por ejemplo), en este caso los campos temporales se encontraban divididos en tres columnas: ANO, MES y DIA. Por este motivo, se creó una nueva variable llamada FECHA que combina estos tres elementos y que permite homogeneizar el formato con el resto de las tablas. Esta estandarización resultó fundamental para facilitar los procesos de cruce de información por fecha.

Por otro lado, teniendo en cuenta que cada observación incluía 24 valores correspondientes a una misma magnitud medida en diferentes horas del día, se optó por calcular una media diaria que resumiera esta información. Esta transformación se aplicó únicamente a las magnitudes seleccionadas para el análisis —velocidad del viento, temperatura, humedad relativa y precipitación—, tal como se justificó previamente. Para ello, se creó una nueva columna denominada avg, que recoge el valor medio de los registros horarios de cada magnitud en cada día y para cada estación. Este enfoque permite simplificar el tratamiento de los datos sin renunciar a la riqueza que aporta la granularidad horaria en futuras exploraciones más específicas.

Estas transformaciones constituyen una fase intermedia clave para garantizar la calidad, consistencia y compatibilidad de los datos antes de su integración definitiva en el análisis conjunto con los registros de trayectos de BiciMAD.

### **4.3. Análisis exploratorio de los conjuntos de datos**

Una vez descritas las variables clave que componen los tres conjuntos de datos utilizados en este trabajo, los trayectos del sistema BiciMAD, el inventario de estaciones meteorológicas fijas y los registros diarios de condiciones atmosféricas, es momento de realizar una primera aproximación a su contenido. Para ello, este apartado se divide en tres secciones principales.

En primer lugar, se mostrarán algunas capturas de las primeras líneas de cada uno de los datasets. Este paso tiene como objetivo ofrecer al lector una visión rápida de la estructura de las tablas, del tipo de variables que contiene cada conjunto y del formato en que se presentan los datos. Esta visualización inicial facilita la comprensión del tipo de información con el que se trabajará a lo largo del análisis.

A continuación, se abordará el análisis de las principales estadísticas descriptivas de las variables más relevantes, con el fin de identificar tendencias centrales, niveles de dispersión y rangos de valores. Este examen estadístico preliminar permite familiarizarse con la escala y el comportamiento general de los datos, así como detectar posibles inconsistencias o anomalías.

Finalmente, una vez concluido el análisis descriptivo, se explicará el proceso de unión de los distintos conjuntos de datos, describiendo las claves de cruce utilizadas y los criterios seguidos para asegurar la integridad de la información combinada. Este apartado incluirá también el diagrama entidad-relación (ER) que resume gráficamente la estructura del sistema de datos consolidado, las relaciones entre las tablas y la lógica de integración aplicada para construir el dataset final que se utilizará en el análisis posterior.

### 4.3.1. Primeras líneas de las bases de datos

Como se ha anticipado, esta sección muestra las primeras filas de cada uno de los conjuntos de datos utilizados. El propósito de esta visualización es doble: por un lado, permite al lector observar de forma directa cómo están estructuradas las variables; por otro, facilita la comprensión del tipo de información disponible en cada base y su nivel de detalle. Además, nos servirá para explicar las distintas transformaciones que ha sido necesario aplicar sobre las bases de datos originales, como la limpieza de valores nulos, la creación de columnas auxiliares como las coordenadas o la fecha unificada, o el cálculo de promedios diarios en el caso de las variables meteorológicas, y que permiten comprender cómo se ha llegado a construir la tabla final sobre la que se basa el análisis.

**Ilustración 3: Primeras líneas de la base de datos de BiciMad**

fecha	idBike	fleet	trip_minutes	geolocation_unlock	address_unlock	unlock_date	locktype	unlocktype	geolocation_lock	address_lock	lock_date	station_unlock	d
2021-06-01	5282.0	1.0	55.15	{'type': 'Point', 'coordinates': [-3.6977715, 40.4251906]}	NaN	2021-06-01 00:00:01	STATION	STATION	{'type': 'Point', 'coordinates': [-3.6991147, 40.4122047]}	NaN	2021-06-01T01:55:10	8.0	
2021-06-01	278.0	1.0	8.98	{'type': 'Point', 'coordinates': [-3.6853312, 40.445411]}	NaN	2021-06-01 00:00:17	STATION	STATION	{'type': 'Point', 'coordinates': [-3.6954983, 40.4218616]}	NaN	2021-06-01T00:09:16	152.0	
2021-06-01	6496.0	1.0	20.83	{'type': 'Point', 'coordinates': [-3.71019, 40.4364]}	NaN	2021-06-01 00:00:22	STATION	STATION	{'type': 'Point', 'coordinates': [-3.6945025, 40.3972616]}	NaN	2021-06-01T00:21:12	253.0	
2021-06-01	5766.0	1.0	10.03	{'type': 'Point', 'coordinates': [-3.680008, 40.421501]}	NaN	2021-06-01 00:00:38	STATION	STATION	{'type': 'Point', 'coordinates': [-3.698447, 40.424148]}	NaN	2021-06-01T00:10:40	90.0	

dock_unlock	unlock_station_name	station_lock	dock_lock	lock_station_name	archivo_origen	idTrip
15.0	Hortaleza	45.0	14.0	Ant�n Mart�n	trips_21_06_June.csv	NaN
8.0	Plaza de la Rep�blica Argentina	20.0	13.0	Prim	trips_21_06_June.csv	NaN
13.0	Galileo	136.0	15.0	Paseo de las Delicias	trips_21_06_June.csv	NaN
26.0	Puerta de Madrid	7.0	2.0	Colegio Arquitectos	trips_21_06_June.csv	NaN

Fuente: elaboraci n propia

Ilustraci n 4: Primeras l neas de la base de datos de estaciones meteorol gicas

codigo_estacion	C�DIGO_CORTO	ESTACION	DIRECCION	LONGITUD_ETRS89	LATITUD_ETRS89	ALTITUD	VV (81)	DV (82)	T (83)	...	P (89)
28079004	4	Plaza Espa�a	Plaza de Espa�a	3�42'43.91"O	40�25'25.98"N	637	NaN	NaN	X	...	NaN
28079008	8	Escuelas Aguirre	Entre C/ Alcal� y C/ O' Donell	3�40'56.22"O	40�25'17.63"N	672	NaN	NaN	X	...	NaN
28079016	16	Arturo Soria	C/ Arturo Soria esq. C/ Vizconde de los Asilos	3�38'21.17"O	40�26'24.20"N	695	NaN	NaN	X	...	NaN
28079018	18	Farolillo	Calle Farolillo - C/Ervigio	3�43'54.61"O	40�23'41.22"N	625	NaN	NaN	X	...	NaN

COD_VIA	VIA_CLASE	VIA_PAR	VIA_NOMBRE	NUM_VIA	COORDENADA_X_ETRS89	COORDENADA_Y_ETRS89
273600.0	PLAZA	DE	ESPA�A	NaN	439579.329073	4.475049e+06
18900.0	CALLE	DE	ALCALA	NaN	442117.236578	4.474771e+06
798700.0	CALLE	DEL	VIZCONDE DE LOS ASILOS	NaN	445786.172865	4.476796e+06
1903.0	CALLE	DEL	FAROLILLO	NaN	437891.696120	4.471833e+06

Fuente: elaboraci n propia

Ilustraci n 5: Primeras l neas de la base de datos de datos meteorol gicos

PROVINCIA	MUNICIPIO	ESTACION	MAGNITUD	PUNTO_MUESTREO	ANO	MES	DIA	H01	V01	...	H22	V22	H23	V23	H24	V24	FECHA	codigo_estacion
28	79	102	81	28079102_81_98	2022	4	1	2.17	V	...	2.12	V	0.0	N	0.0	N	2022-04-01	28079102
28	79	102	81	28079102_81_98	2022	4	2	2.7	V	...	2.78	V	3.3	V	0.0	N	2022-04-02	28079102
28	79	102	81	28079102_81_98	2022	4	3	2.22	V	...	0.0	N	0.0	N	0.0	N	2022-04-03	28079102
28	79	102	81	28079102_81_98	2022	4	4	5.72	V	...	0.48	V	1.28	V	3.15	V	2022-04-04	28079102
28	79	102	81	28079102_81_98	2022	4	5	2.6	V	...	2.8	V	2.68	V	2.05	V	2022-04-05	28079102

**Fuente: elaboración propia**

La visualización de las primeras filas de cada uno de los tres conjuntos de datos empleados, trayectos de BiciMAD, estaciones meteorológicas fijas y registros meteorológicos diarios, nos ha permitido obtener una primera impresión general sobre su estructura, tipo de variables y nivel de detalle. Este paso preliminar resulta clave para familiarizarse con la información disponible y anticipar posibles retos en su tratamiento.

Gracias a esta revisión inicial, hemos podido confirmar varios de los aspectos comentados previamente, como la ausencia parcial de algunas variables en determinadas fechas, la necesidad de unificar formatos de fecha o extraer coordenadas de campos complejos. Estas observaciones refuerzan la idea de que será necesario aplicar distintas transformaciones sobre los datos originales con el fin de obtener un conjunto homogéneo y robusto que permita realizar un análisis riguroso y obtener resultados concluyentes.

A continuación, en la siguiente sección, se procederá al análisis de las estadísticas descriptivas básicas de las variables numéricas más relevantes de cada Dataset. Este análisis nos permitirá identificar patrones generales, rangos de valores y posibles anomalías, sentando las bases para la posterior integración y modelización de los datos.

#### **4.3.2. Medidas estadísticas de las variables más relevantes**

Una vez visualizada la estructura general de cada uno de los datasets utilizados en este trabajo, el siguiente paso consiste en analizar las principales estadísticas descriptivas de algunas variables numéricas relevantes para el estudio. No se han incluido todas las variables, ya que muchos de los campos presentes en los datasets corresponden a valores categóricos, direcciones textuales o identificadores, cuya estadística básica no aporta información útil desde el punto de vista cuantitativo. Del mismo modo, aunque el conjunto de datos incluye variables geográficas como longitud y latitud, se ha optado por limitar su análisis descriptivo a una visión general, dado que medidas como la media o el valor máximo y mínimo en estos casos no reflejan información especialmente significativa en términos analíticos.

Antes de calcular las estadísticas, fue necesario aplicar una serie de transformaciones sobre el Dataset con el objetivo de garantizar la coherencia y la calidad de los datos analizados. En primer lugar, se eliminaron todos aquellos registros que contenían valores nulos en variables clave para el análisis, concretamente: `trip_minutes`, `fecha`, `lat_bici` y `lon_bici`. Estas variables son fundamentales tanto para el estudio de la duración de los trayectos como para la georreferenciación de su inicio.

Posteriormente, se procedió a depurar valores atípicos o incoherentes. Se eliminaron aquellos registros en los que la duración del trayecto (`trip_minutes`) era negativa o excesivamente alta. En concreto, se estableció un umbral de 300 minutos, considerando que duraciones superiores a este valor probablemente responden a errores de registro o a situaciones operativas no representativas del uso habitual del sistema (por ejemplo, traslados de bicicletas para mantenimiento o fallos en el cierre del trayecto).

Una vez aplicado este proceso de limpieza, se calcularon las estadísticas descriptivas de las variables seleccionadas. En la siguiente tabla se presentan las medidas básicas de tendencia central (media, mediana), dispersión (desviación estándar, cuartiles) y valores extremos (mínimo y máximo) de las variables `trip_minutes`<sup>2</sup>, `lon_bici` y `lat_bici`.

**Ilustración 6: Estadísticas principales de los datasets de BiciMad**

	<code>trip_minutes</code>	<code>lon_bici</code>	<code>lat_bici</code>
<b>count</b>	6717462.00	6717462.00	6717462.00
<b>mean</b>	11.94	-3.69	40.42
<b>std</b>	9.83	0.02	0.02
<b>min</b>	0.00	-3.83	40.34
<b>25%</b>	5.45	-3.71	40.41
<b>50%</b>	10.48	-3.70	40.42
<b>75%</b>	16.33	-3.68	40.43
<b>max</b>	59.98	-3.50	40.52

**Fuente: elaboración propia**

Los resultados descriptivos de las variables `trip_minutes`, `lon_bici` y `lat_bici` permiten obtener una primera aproximación al comportamiento espacial y temporal de los trayectos registrados en el sistema BiciMAD. La duración media de los viajes se sitúa en torno a los 11,94 minutos, con una mediana ligeramente inferior (10,48), lo que indica una distribución relativamente simétrica. La desviación estándar de casi 10 minutos refleja una cierta heterogeneidad en los desplazamientos, aunque los valores centrales muestran que la mayoría de los trayectos se concentran en un rango razonable de entre 5 y 16 minutos.

En cuanto a las coordenadas, tanto la latitud como la longitud presentan una variabilidad muy reducida (desviación estándar de apenas 0.02), lo cual es coherente con un sistema de movilidad localizado dentro del núcleo urbano de Madrid. Las coordenadas medias (40.42 de latitud y -3.69 de longitud) apuntan a una zona próxima al centro de la ciudad, lo que confirma la consistencia geográfica de los datos.

Por otro lado, el segundo Dataset, correspondiente a las estaciones meteorológicas fijas, no ha sido objeto de un análisis estadístico detallado, ya que contiene fundamentalmente información estática (códigos, direcciones, coordenadas y capacidades de medición), por lo que no aportaría valor adicional en términos descriptivos.

Pasamos, por tanto, al análisis del tercer conjunto de datos, que recoge información meteorológica diaria segmentada por horas. Este Dataset incluye variables numéricas de gran interés como la temperatura, la velocidad del viento, la humedad relativa o la precipitación, cuya evolución será clave para comprender su posible influencia sobre el uso del sistema de bicicletas compartidas.

**Ilustración 7: Estadísticas principales de los datasets de datos meteorológicos diarios**

<sup>2</sup> Variable que indica la duración del viaje

	Velocidad del viento (m/s)	Dirección del viento	Temperatura (°C)	Humedad relativa (%)	Presión barométrica (mb)	Radiación solar (W/m <sup>2</sup> )	Precipitación (l/m <sup>2</sup> )
count	126407.00	126432.00	299988.00	285457.00	97340.00	99558.00	130200.00
mean	1.21	117.01	16.22	52.88	946.27	199.12	0.04
std	0.92	96.40	9.57	23.40	13.68	287.14	0.40
min	0.00	0.00	-19.70	0.00	800.00	0.00	0.00
25%	0.57	29.00	8.80	35.00	941.00	1.00	0.00
50%	1.00	84.00	15.20	52.00	945.00	6.00	0.00
75%	1.63	206.00	23.50	71.00	950.00	360.00	0.00
max	39.45	359.00	47.90	100.00	1038.00	1403.00	26.70

**Fuente: elaboración propia**

La tabla anterior muestra las estadísticas descriptivas de las principales magnitudes meteorológicas registradas en la ciudad de Madrid, una vez aplicados los procesos de transformación, limpieza y agregación de datos horarios (H01–H24) para cada día y cada estación.

Previo al análisis, se llevó a cabo una depuración de los datos para eliminar valores atípicos o físicamente imposibles, que pudieran distorsionar la interpretación estadística. Se definieron rangos válidos específicos por cada magnitud meteorológica, eliminando por ejemplo temperaturas fuera del intervalo -20 °C a 50 °C, velocidades de viento superiores a 50 m/s, humedades relativas por encima del 100% o presiones barométricas fuera del rango 800–1100 Mb. Este filtrado permitió detectar y excluir valores claramente anómalos que habían sido identificados previamente (como máximos irreales de 270 m/s en viento o temperaturas extremas de hasta 98 °C). Asimismo, se transformaron en ceros aquellas celdas originalmente vacías que representaban la ausencia de precipitación, ya que eran interpretadas como valores nulos (NaN) en el Dataset original.

Una vez depurado el conjunto de datos, los valores estadísticos obtenidos reflejan rangos mucho más coherentes. La temperatura media diaria es de 16,22 °C, con un rango intercuartílico que va de 8,80 °C a 23,50 °C. La velocidad media del viento se sitúa en 1,21 m/s, con una desviación estándar moderada (0,92 m/s) y un máximo razonable de 39,45 m/s. La humedad relativa presenta una media del 52,88%, dentro de un rango lógico del 0% al 100%, y la precipitación muestra una media diaria muy baja (0,04 l/m<sup>2</sup>), reflejo de la alta frecuencia de días secos en Madrid, aunque con registros máximos que pueden alcanzar hasta 26,70 l/m<sup>2</sup> en episodios puntuales de lluvia intensa.

A partir de este análisis, y considerando tanto su relevancia teórica como su disponibilidad y calidad en el conjunto de datos, se han seleccionado como variables meteorológicas clave para el estudio las siguientes: velocidad del viento (m/s), temperatura (°C), humedad relativa (%) y precipitación (l/m<sup>2</sup>). Estas serán las variables utilizadas posteriormente para analizar su influencia sobre la demanda diaria del sistema de bicicletas públicas BiciMAD.

Concluido este análisis general y depurado del conjunto de datos, en la siguiente sección se realizará una exploración detallada de cada uno de los tres datasets principales utilizados en este trabajo: el registro de viajes de BiciMAD, información estática de las estaciones meteorológicas fijas y los datos meteorológicos diarios por estación. En cada uno de ellos profundizaremos en la estructura de las variables, su comportamiento a lo largo del tiempo y su relevancia para los objetivos del estudio. Este análisis más desagregado nos permitirá identificar patrones específicos, validar la coherencia interna de los datos y anticipar posibles retos metodológicos en el proceso de integración final de los datasets.

#### **4.4. Transformación de los conjuntos de datos**

Una vez presentadas las fuentes de datos utilizadas y descritas las principales variables de cada conjunto, es imprescindible proceder a un proceso de transformación y depuración que permita garantizar la calidad del análisis posterior. Esta fase es fundamental para eliminar inconsistencias, detectar valores atípicos, homogeneizar formatos y preparar las variables en un formato que facilite la explotación estadística y visual.

En esta sección nos centraremos exclusivamente en los datos de viajes registrados por el sistema BiciMAD, ya que constituyen la base principal para comprender los patrones de movilidad en la ciudad de Madrid. Las bases de datos meteorológicas y de estaciones fijas, por su parte, han requerido procesos mínimos de limpieza, dado que su origen institucional garantiza un alto grado de estandarización y fiabilidad.

El objetivo de este proceso no es únicamente eliminar errores, sino también crear nuevas variables derivadas (como el tipo de día, la franja horaria o la clave de unión con la meteorología) que resultan esenciales para enriquecer el análisis y permitir cruces significativos entre diferentes dimensiones. A continuación, se detallan los pasos más relevantes del tratamiento de datos, incluyendo la conversión de fechas, la extracción de horas, la clasificación temporal y la generación de identificadores únicos para integrar diferentes fuentes de información.

#### 4.4.1. Transformación de valores nulos

Uno de los primeros pasos críticos en el tratamiento de los datos fue la identificación y gestión de valores nulos, así como la detección de registros estructuralmente defectuosos o mal formateados. Este proceso resultó especialmente relevante en las bases de datos de viajes de BiciMAD, que presentaban ciertas irregularidades debido a la evolución del sistema de registro durante el periodo analizado (2021–2023).

En primer lugar, se decidió excluir por completo el conjunto de datos correspondiente al mes de octubre, debido a que presentaba errores graves de formato que impedían su correcta lectura e integración. Este fichero mostraba estructuras de columnas inestables, filas corruptas y una distribución de delimitadores inconsistente, lo cual imposibilitaba una limpieza automatizada con garantías.

Posteriormente, durante la carga y fusión de los archivos restantes, se aplicaron diferentes mecanismos para depurar los registros:

- **Eliminación de filas completamente vacías:** En los archivos originales de BiciMAD, entre registro y registro se incluía sistemáticamente una fila en blanco, lo que generaba ruido al intentar concatenar múltiples datasets. Este tipo de filas fueron eliminadas mediante una regla de limpieza basada en la ausencia total de contenido.
- **Corrección de cabeceras duplicadas:** En ciertos ficheros se detectó la repetición de la fila de encabezado dentro del propio cuerpo del archivo (probablemente por errores en la exportación). Además, algunas columnas presentaban nombres agrupados en una sola celda por errores de delimitación. Para solucionarlo, se identificaron estas cabeceras erróneas, se corrigieron los nombres de columna y se eliminaron las filas que replicaban la fila de encabezado dentro del contenido.

Durante la exploración inicial de los datos combinados, se identificaron también problemas más específicos:

- **Campos con formato incorrecto o incompleto:** Por ejemplo, la variable `geolocation_unlock` (pensada para contener las coordenadas del punto de desbloqueo) se encontraba en formato de tipo diccionario o texto mal estructurado, lo cual impedía extraer fácilmente la latitud y longitud necesarias para el análisis geográfico. Este problema fue documentado y tratado en etapas posteriores del preprocesado.

- **Variables en desuso o discontinuas:** Algunas columnas como `address_unlock`, `address_lock` o `id_trip` aparecían únicamente en ciertos periodos temporales, reflejando posiblemente cambios en la estructura interna de registro de datos por parte del sistema. Estas variables se descartaron del análisis principal al no aportar información sistemática y consistente.

Una vez fusionados y homogeneizados los registros válidos, se procedió a aplicar una limpieza más estricta en las variables clave para el análisis:

Se identificaron las columnas fundamentales para los análisis posteriores (concretamente `trip_minutes`, `fecha`, `lat_bici` y `lon_bici`), y se eliminaron únicamente aquellas filas que presentaban valores nulos en estos campos. Esta decisión permitió conservar un volumen amplio de registros válidos, evitando una pérdida masiva de datos por eliminar filas con nulos en campos irrelevantes.

Este enfoque selectivo y gradual en el tratamiento de los valores nulos permitió preservar la integridad de los datos realmente útiles, minimizando al mismo tiempo el impacto de registros defectuosos. El siguiente paso en la depuración será el análisis de outliers<sup>3</sup> y valores atípicos, que abordaremos en una sección específica dedicada a la detección de sesgos y valores atípicos.

#### **4.4.2. Transformación de valores atípicos**

Una vez identificados y eliminados los valores nulos más relevantes, fue necesario aplicar un tratamiento específico para gestionar los valores atípicos o incoherentes, tanto en las variables relacionadas con los viajes como en los registros meteorológicos. Este paso es esencial para garantizar la fiabilidad y la consistencia del análisis.

Uno de los aspectos clave en el proceso de depuración del conjunto de datos ha sido la identificación y eliminación de valores atípicos en la variable `trip_minutes`, correspondiente a la duración total del trayecto. Este tratamiento resulta esencial para preservar la validez estadística del análisis y asegurar que las conclusiones obtenidas representen de forma fiel el comportamiento típico de los usuarios del sistema BiciMAD.

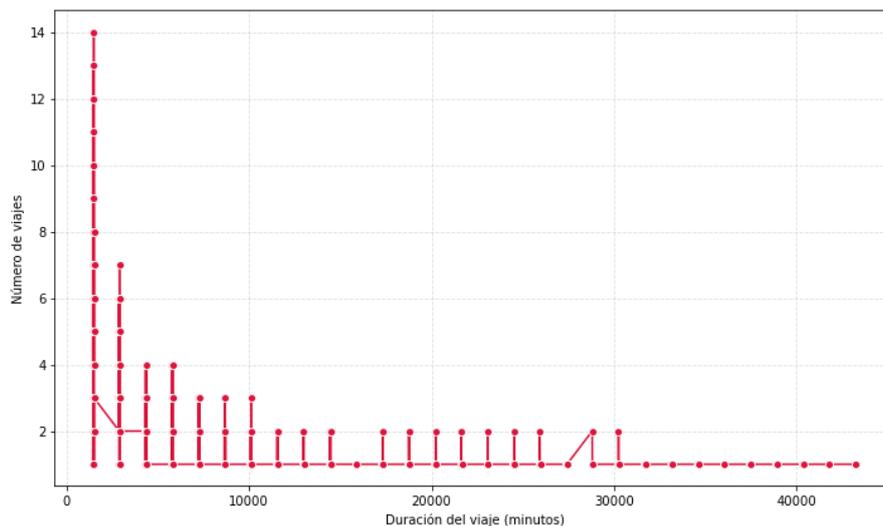
Para ello, se ha establecido un umbral máximo de 300 minutos (equivalente a cinco horas) como criterio de corte superior. Este límite se justifica tanto por su razonabilidad desde el punto de vista del uso ordinario del servicio, como por la necesidad de excluir registros extremos cuya validez operativa o representatividad es dudosa. En condiciones normales, un usuario medio no emplea una bicicleta compartida durante un tiempo tan prolongado, ya que el sistema está concebido para trayectos cortos, frecuentes y de carácter urbano.

Adicionalmente, se ha observado que algunos registros con duraciones superiores a dicho umbral pueden deberse a fallos técnicos, como errores en el cierre del candado, o a situaciones operativas especiales no vinculadas al uso personal (por ejemplo, traslados de bicicletas por parte del personal de mantenimiento). También se ha considerado la posibilidad de que determinados perfiles de usuario, especialmente durante campañas de gratuidad del servicio, hagan un uso intensivo y continuado de las bicicletas, como ocurre en el caso de repartidores de plataformas como Glovo o Uber Eats. Aunque estas situaciones generan datos técnicamente válidos, su inclusión en el análisis general podría desvirtuar los patrones de comportamiento que se pretende estudiar.

#### **Ilustración 8: Distribución de trayectos con duración superior a 300 minutos**

---

<sup>3</sup> Valor atípico



Fuente: elaboración propia

Para reforzar esta decisión metodológica, se ha representado gráficamente la distribución de los trayectos cuya duración supera los 300 minutos. En la ilustración 8 se aprecia claramente que el número de registros disminuye de forma muy acusada a medida que aumenta la duración, con apenas unos pocos viajes más allá de los 500 minutos y algunos casos extremos que alcanzan cifras completamente inverosímiles, como más de 40.000 minutos (equivalente a varias semanas). Este comportamiento confirma que los valores por encima del umbral establecido constituyen una fracción muy reducida del total, y que su distribución no responde a patrones propios del uso ciudadano habitual, sino más bien a situaciones anómalas o fuera de contexto.

La exclusión de estos valores atípicos permite, por tanto, mejorar la robustez del análisis posterior y garantizar que los resultados obtenidos se basen en datos representativos del funcionamiento real del sistema y del comportamiento medio de sus usuarios.

Por otro lado, se llevó a cabo un segundo proceso de limpieza de valores atípicos en el conjunto de datos meteorológicos. Durante la revisión de las magnitudes registradas por las estaciones de control, se detectaron valores claramente erróneos, como temperaturas de 99 °C o valores negativos extremos de -40 °C, así como velocidades de viento inverosímiles o porcentajes de humedad superiores al 100%. Para garantizar la coherencia de estas mediciones, se definieron rangos válidos teóricos para cada tipo de magnitud, como se resume a continuación:

**Tabla 5: Rangos válidos establecidos para las magnitudes meteorológicas**

Código de magnitud	Magnitud	Unidad	Rango válido
81	Velocidad del viento	m/s	0-50
82	Dirección del viento	grados (°)	0-360
83	Temperatura	grados Celsius	(-20)-50
86	Humedad relativa	%	0 – 100
87	Presión barométrica	Mb	800-1100
88	Radiación solar	W/m <sup>2</sup>	0-1500
89	Precipitación	l/m <sup>2</sup>	0-100

### Fuente: elaboración propia

A través de este procedimiento, se filtraron los registros meteorológicos con valores fuera de estos rangos y se marcaron como nulos para su posterior tratamiento. Esta limpieza fue crucial para asegurar la validez de los análisis correlacionales entre las condiciones climáticas y la demanda de bicicletas.

Gracias a estas acciones de depuración, se estableció una base sólida sobre la que desarrollar un análisis más profundo y fiable. En la siguiente sección se describirá el proceso de transformación y creación de nuevas variables a partir de los datos originales, orientado a enriquecer el análisis posterior.

#### 4.4.3. Transformaciones y creación de variables

Una vez completadas las tareas de exploración inicial, tratamiento de valores nulos y eliminación de outliers, resulta imprescindible llevar a cabo una serie de transformaciones sobre determinadas variables para que puedan ser utilizadas de forma eficaz en los análisis posteriores. Esta fase de transformación tiene como objetivo estandarizar, clarificar y derivar nuevas variables a partir de otras ya existentes, facilitando así la interpretación y explotación de la información contenida en los datos. A continuación, se detallan las principales operaciones realizadas.

Uno de los primeros desafíos encontrados en los datos de BiciMAD fue el tratamiento de las variables `geolocation_unlock` y `geolocation_lock`, que contenían las coordenadas de localización en un formato tipo JSON, poco adecuado para el análisis directo: `{'type': 'Point', 'coordinates': [-3.6977715, 40.4251906]}`. Este formato dificultaba la manipulación y visualización de los datos, por lo que se procedió a extraer las coordenadas explícitamente, creando dos nuevas columnas denominadas `lat_bici` y `lon_bici`. Mediante expresiones regulares se capturaron los valores numéricos y se convirtieron al tipo `float`<sup>4</sup>, asegurando su correcto tratamiento.

En el caso de las bases de datos meteorológicas diarias, se identificó una estructura inusual en la representación de la fecha, dividida en tres columnas independientes: `ANO`, `MES` y `DIA`. Para homogeneizar esta información con el resto de los datasets, se unificaron estas tres variables en una nueva columna denominada `FECHA`, con el formato `YYYY-MM-DD`. Este paso fue clave para posibilitar operaciones de fusión y comparación temporal entre las distintas fuentes de datos.

Otra transformación relevante consistió en la extracción del código de estación meteorológica a partir de la variable `PUNTO_MUESTREO`, que presentaba un formato compuesto como `28079102_81_98`. Mediante una separación por subrayado, se aisló el bloque numérico inicial de ocho cifras, correspondiente al identificador de la estación. Este valor se convirtió en numérico, permitiendo su uso como clave para unirlo con otras bases de datos.

Con el objetivo de sintetizar la información meteorológica registrada por hora, se calculó la media diaria de cada magnitud para cada estación. Esta media se utilizó como referencia en los análisis, si bien debe reconocerse que este método tiene limitaciones: por ejemplo, un evento puntual como una tormenta breve a las 9:00 de la mañana puede quedar diluido al considerar una media diaria. Aun así, esta transformación resulta útil para obtener una visión general de las condiciones climáticas asociadas a los registros de movilidad.

Finalmente, se abordó la asignación de nombres descriptivos a las magnitudes meteorológicas. Originalmente, estas venían codificadas numéricamente (por ejemplo, 81 para velocidad del viento o 83 para temperatura), lo que dificultaba la comprensión directa de los datos. Se implementó un diccionario de equivalencias para traducir estos

---

<sup>4</sup> El tipo `float` en Python representa números reales con decimales, lo que permite operar con valores continuos como coordenadas o duraciones.

códigos en etiquetas más comprensibles y facilitar su uso en gráficos y tablas. Las correspondencias definidas fueron las siguientes:

**Tabla 6: Rangos válidos establecidos para las magnitudes meteorológicas**

Código de magnitud	Magnitud
81	Velocidad del viento (m/s)
82	Dirección del viento
83	Temperatura (°C)
86	Humedad relativa (%)
87	Presión barométrica (Mb)
88	Radiación solar (W/m <sup>2</sup> )
89	Precipitación (l/m <sup>2</sup> )

**Fuente: elaboración propia**

Habiendo entendido todas las transformaciones realizadas con el objetivo de mejorar la comprensión y la calidad de los datos, se puede afirmar que el dataset está listo para el análisis detallado. Estas tareas han sido fundamentales para garantizar la coherencia estructural y semántica de la información y para permitir una interpretación rigurosa de los patrones de movilidad y su relación con las condiciones meteorológicas.

## **5. Análisis detallado de las bases de datos**

Una vez completadas las tareas de limpieza, transformación y estructuración de los datos, se procede al análisis individualizado de cada una de las bases que conforman el estudio. Este análisis tiene como objetivo comprender en profundidad la naturaleza y características específicas de cada fuente de información, evaluando la estructura interna de los datos, su comportamiento temporal y geográfico, y la calidad de las variables clave para el desarrollo posterior del trabajo. A lo largo de las siguientes secciones, se explorarán los patrones de uso del sistema de bicicletas compartidas, las mediciones meteorológicas recogidas por las estaciones de control atmosférico y su evolución diaria, así como la posible relación entre ambas dimensiones. Este enfoque analítico permitirá no solo detectar inconsistencias o valores atípicos, sino también sentar las bases para una integración rigurosa y significativa de las distintas fuentes en los modelos y visualizaciones posteriores. La comprensión detallada de cada conjunto de datos es, por tanto, esencial para garantizar la solidez metodológica y la validez de las conclusiones alcanzadas en las etapas finales del estudio.

### **5.1. Análisis detallado de la base de datos de BiciMad**

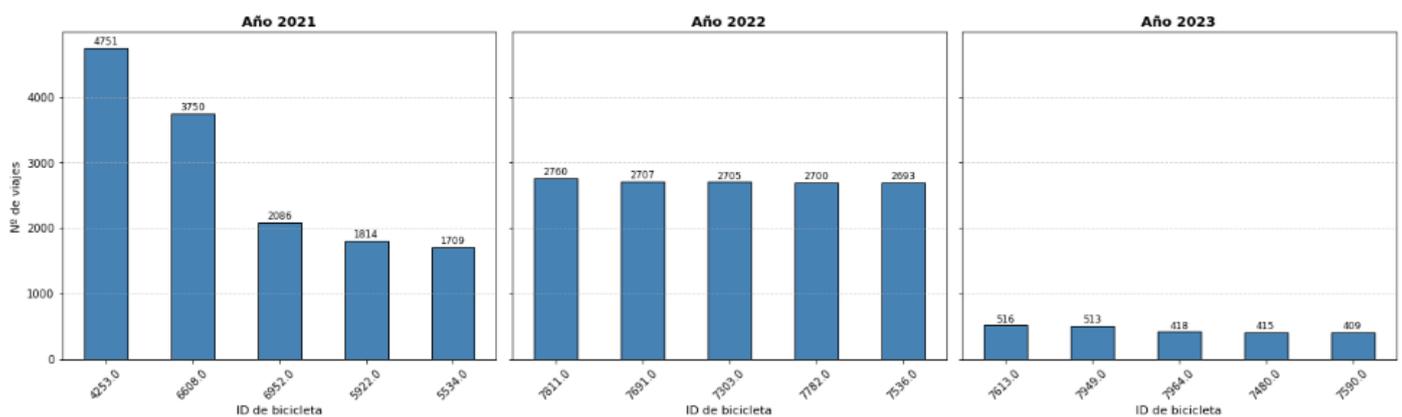
Una vez realizadas las primeras exploraciones generales y descriptivas de los conjuntos de datos, es momento de profundizar en el análisis específico del registro de trayectos de BiciMAD, el cual constituye el núcleo de este estudio. Esta base de datos recoge información detallada de cada viaje realizado a lo largo de los años 2021, 2022 y 2023, incluyendo variables como el identificador único de la bicicleta (idBike), la fecha y hora del inicio del trayecto, la duración del viaje, o las estaciones de origen y destino, entre otras.

### 5.1.1. Demanda del servicio de BiciMad por ID

El objetivo de este apartado es comprender en profundidad los patrones de uso del sistema, tanto a nivel agregado como desagregado. Para ello, se estudiará el comportamiento de las bicicletas de forma individual, identificando aquellas que presentan un mayor o menor nivel de utilización, así como la evolución temporal del número de viajes a lo largo del año. Estos análisis ayudan a detectar desequilibrios operativos, necesidades de redistribución o mantenimiento, e incluso errores de registro. Todo ello resulta útil para mejorar la gestión del servicio.

A continuación, se analiza qué bicicletas han tenido un mayor y menor uso durante los años 2021, 2022 y 2023, a partir de la variable idBike, que actúa como identificador único de cada unidad en el sistema BiciMAD. Este tipo de análisis permite observar no solo patrones de uso individual, sino también detectar posibles desequilibrios operativos en la distribución y rotación de las bicicletas.

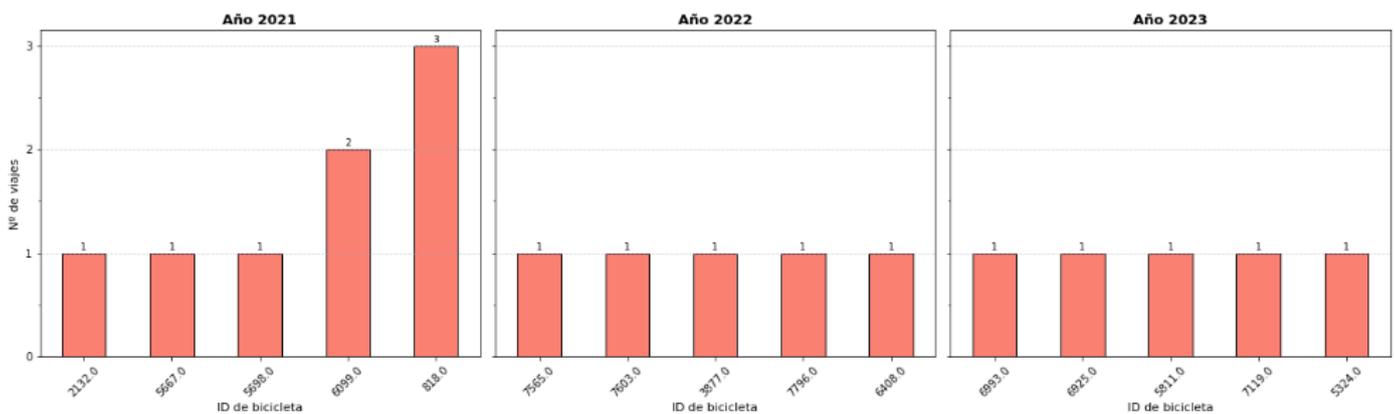
**Ilustración 9: Resultados de las 5 bicis más usadas por año**



Fuente: elaboración propia

En contraste, las gráficas inferiores muestran el top 5 de bicicletas menos utilizadas. En todos los años analizados, hay bicicletas que únicamente registraron uno o muy pocos trayectos. Este hecho puede deberse a múltiples causas: ubicación geográfica con baja demanda, fallos técnicos, mantenimiento prolongado, o incluso errores en el registro de datos.

**Ilustración 10: Resultados de las 5 bicis menos usadas por año**



Fuente: elaboración propia

Estos resultados pueden ser de gran utilidad para BiciMAD en varios aspectos clave:

- Planificación de mantenimiento y renovación de flota, priorizando aquellas bicicletas con mayor desgaste.
- Optimización de la rotación y redistribución, en base al uso real.
- Detección de bicicletas o estaciones inactivas, permitiendo reacciones operativas rápidas.

Ahora bien, al observar la homogeneidad de los valores en los años 2022 y 2023 —tanto en bicicletas más como menos utilizadas— surgen dudas razonables sobre la fiabilidad del sistema de identificación (idBike) en esos períodos. Esto puede estar relacionado con los cambios estructurales implementados por BiciMAD a partir de 2022, como la incorporación del nuevo modelo de bicicleta E-FIT, la sustitución de estaciones, y la reconfiguración tecnológica general del sistema. Durante este proceso, es probable que:

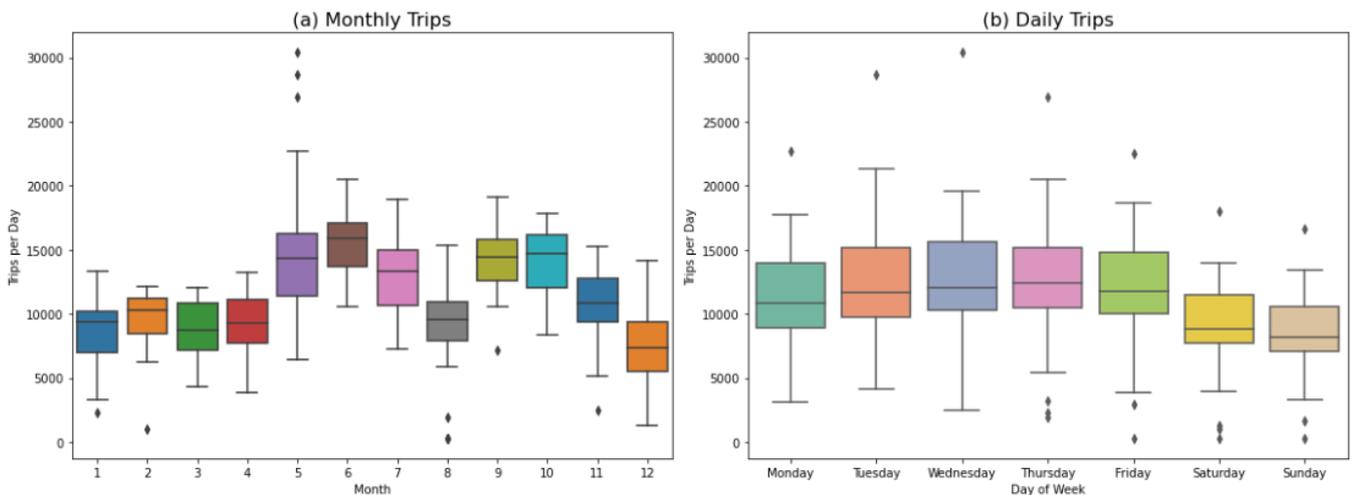
- Se produjera una reasignación de identificadores, lo que dificultaría el seguimiento individual de bicicletas a lo largo del tiempo.
- Se registrarán errores o datos incompletos durante la transición.
- El uso gratuito del servicio (vigente en 2023) contribuyera a alterar los patrones de uso habituales.

Por tanto, si bien estos gráficos permiten extraer conclusiones útiles para la gestión operativa, es importante tomar con cautela los datos de los años 2022 y 2023, especialmente en lo relativo al seguimiento individual de bicicletas, debido al contexto de renovación tecnológica y posibles limitaciones en la calidad del registro.

### 5.1.2. Demanda del servicio de BiciMad por mes y día de la semana

Para continuar esta sección dedicada al análisis detallado del comportamiento del sistema, presentamos una visualización que resume la distribución mensual de la demanda de viajes registrados en BiciMAD durante el periodo de estudio.

**Ilustración 11: Distribución de la demanda según el mes y el día de la semana**



**Fuente: elaboración propia**

Esta primera visualización representa, mediante diagramas de caja, la variabilidad del número de viajes por día a lo largo de los diferentes meses del año. En ella se pueden observar tendencias estacionales muy marcadas, que ofrecen una visión clara del impacto del calendario y el clima sobre la utilización del servicio.

Los meses con mayor mediana de viajes diarios son mayo, junio y julio, lo cual es consistente con una mayor propensión al uso de la bicicleta durante la primavera y el inicio del verano. Estos meses presentan, además, un rango intercuartílico más amplio, indicando una mayor dispersión en la demanda diaria, probablemente influida por condiciones meteorológicas, eventos urbanos o días festivos.

Por el contrario, los meses con menor número de viajes son enero, febrero y diciembre, lo que sugiere una reducción de la demanda en invierno, coherente con el efecto disuasorio del frío y otras condiciones climáticas adversas. También se observan valores atípicos (outliers), especialmente en los meses centrales del año, donde se registraron días con un número de viajes notablemente superior al promedio, posiblemente vinculados a campañas promocionales o eventos de ciudad

Además de esta estacionalidad mensual, el gráfico comparativo por día de la semana permite observar un patrón muy definido: la demanda es significativamente mayor de lunes a viernes, con una mediana que ronda los 12.000 a 14.000 viajes diarios, mientras que los sábados y domingos registran valores considerablemente más bajos. Esta diferencia refleja con claridad el carácter utilitario y vinculado a la movilidad cotidiana del servicio, utilizado mayoritariamente en días laborables, posiblemente por personas que lo emplean para desplazamientos al trabajo o estudios. Los fines de semana, en cambio, el uso parece orientarse más hacia el ocio, con un volumen de viajes menor y más variable.

En conjunto, ambas visualizaciones evidencian patrones estacionales y semanales relevantes que deben ser tenidos en cuenta para la planificación operativa de BiciMAD. La distribución temporal de la demanda permite anticipar mejor la carga del sistema, gestionar los recursos de forma eficiente y ajustar estrategias de redistribución o mantenimiento.

Cabe señalar que la posible influencia de factores meteorológicos sobre estos patrones de demanda será objeto de análisis específico más adelante, donde se evaluará si variables como la temperatura, la humedad o la precipitación tienen un impacto significativo sobre el uso del servicio.

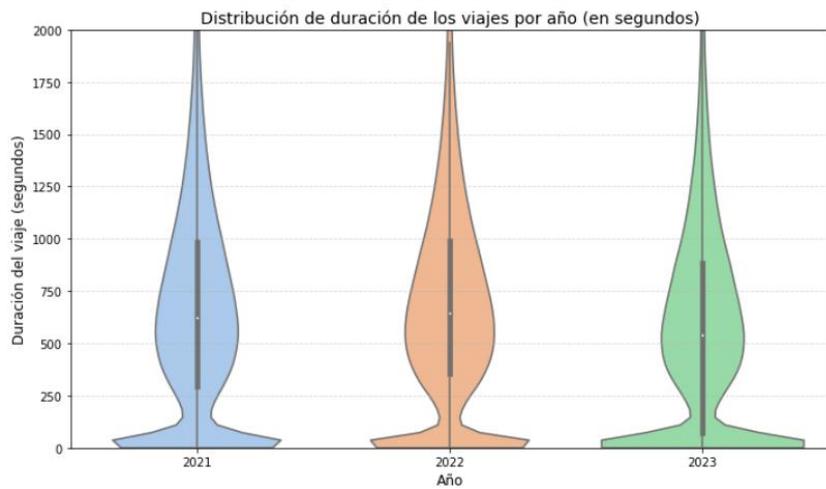
En la siguiente sección, se abordará el estudio de la duración de los trayectos, analizando su comportamiento global, su evolución temporal a lo largo de los años y su distribución entre diferentes tipos de usuarios o momentos del calendario.

### **5.1.3. Análisis del servicio de BiciMad por duración de viaje**

La siguiente visualización muestra la distribución de la duración de los viajes realizados en el sistema BiciMAD durante los años 2021, 2022 y 2023. La duración se ha representado en segundos para una mayor precisión y se han empleado diagramas de violín que permiten visualizar tanto la densidad de los valores como los rangos más frecuentes y extremos dentro de cada año.

Antes de generar esta gráfica, se llevó a cabo un proceso de limpieza y tratamiento de datos atípicos, eliminando del análisis aquellos registros cuya duración excedía los 300 minutos (18.000 segundos), por considerarse valores poco realistas o asociados a trayectos interrumpidos o errores del sistema. Además, se excluyeron los viajes con duración negativa o igual a cero, y se creó una nueva variable `trip_seconds` a partir de la conversión de los minutos registrados en `trip_minutes`.

#### **Ilustración 12: Diagrama de violín sobre la duración del viaje por año (en segundos)**



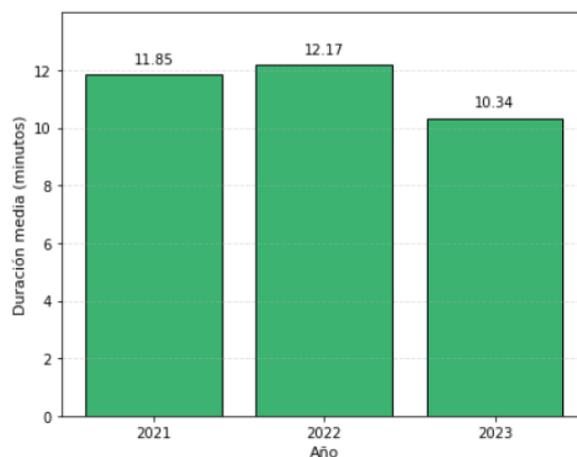
Fuente: elaboración propia

La primera visualización, mediante un diagrama de violín, refleja un patrón bastante estable en la distribución de los viajes entre los tres años analizados, con una mediana que se mantiene aproximadamente entre los 500 y 600 segundos (8 a 10 minutos). La forma de los diagramas revela también una alta concentración de viajes de corta duración, especialmente en la franja entre los 200 y 900 segundos, lo cual es coherente con un uso principalmente urbano y de trayectos breves.

En los tres años se observan colas alargadas hacia la derecha, que indican la presencia de un número reducido de viajes significativamente más largos. No obstante, gracias al filtrado aplicado, estos valores extremos ya no distorsionan la visualización general, permitiendo apreciar con claridad el comportamiento predominante de los usuarios.

Para complementar esta visualización, se ha generado el siguiente gráfico de barras que representa la media anual de duración de los viajes, expresada en minutos:

Ilustración 13: Media anual de duración de los viajes (en minutos)



Fuente: elaboración propia

Como se puede observar, las medias se mantienen en torno a los 10–12 minutos por trayecto. En particular, el valor medio más bajo se registra en 2023 con 10,34 minutos, seguido de 11,85 en 2021 y 12,17 en 2022. Sin embargo, es

importante matizar que los datos correspondientes a 2023 solo incluyen los meses de enero y febrero, que como ya se ha señalado previamente son meses con menor demanda estacional, lo que podría afectar a la interpretación de la media. Por tanto, cualquier comparación con años completos debe hacerse con cautela.

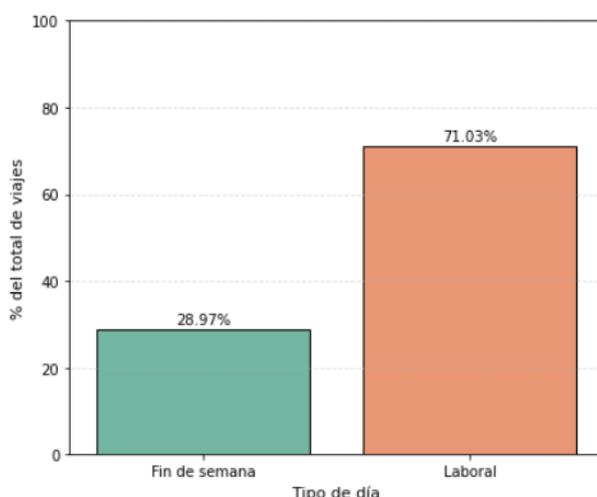
Este análisis sugiere que, independientemente del año, la duración de los trayectos tiende a mantenerse relativamente estable, lo que podría reflejar hábitos de desplazamiento consistentes entre los usuarios del sistema. En futuras secciones se explorará si esta duración varía en función de otras variables como el tipo de día (laborable vs. fin de semana), la hora del viaje o incluso las condiciones meteorológicas, cuya influencia será analizada en profundidad más adelante.

#### 5.1.4. Análisis del servicio de BiciMad por tipo de día de la semana

Con el objetivo de comprender mejor los patrones de uso del sistema BiciMAD, se ha realizado un análisis detallado de la demanda segmentada por tipo de día y franjas horarias. Para ello, se ha partido de una definición operativa precisa del tipo de día, que se mantiene coherente en todos los análisis realizados a lo largo del estudio. En concreto, se consideran días laborables aquellos comprendidos desde el lunes a las 00:00 hasta el viernes a las 15:59 horas, mientras que se clasifica como fin de semana el periodo que va desde el viernes a las 16:00 hasta el lunes a las 00:00. Esta segmentación responde a la necesidad de diferenciar de forma más precisa los desplazamientos asociados a la rutina laboral de aquellos que tienen un carácter más recreativo o vinculado al tiempo libre, especialmente en un contexto urbano como el de Madrid, donde las pautas de movilidad cambian notablemente a partir del viernes por la tarde.

En primer lugar, se ha analizado la distribución general de viajes entre ambos tipos de día. Los resultados indican que aproximadamente un 71 % de los trayectos se concentran en los días laborables, frente a un 29 % registrados durante el fin de semana. Esta diferencia cuantitativa refuerza la hipótesis de que el sistema BiciMAD cumple fundamentalmente una función de movilidad funcional a lo largo de la semana, siendo utilizado por los ciudadanos para desplazamientos asociados a estudios, trabajo u otras obligaciones cotidianas.

**Ilustración 14: Porcentaje de viajes por tipo de día**



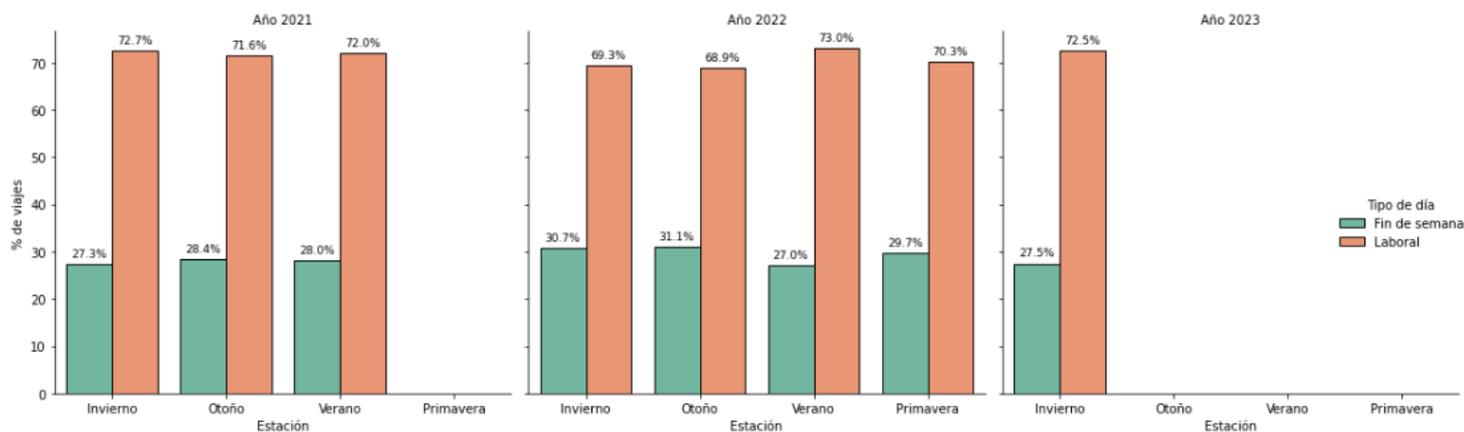
**Fuente: elaboración propia**

En segundo lugar, se ha desagregado el análisis por año (2021-2023) y estación del año, lo que ha permitido identificar ligeras variaciones estacionales. Durante los meses fríos (invierno y otoño), se observa un incremento de la proporción

de viajes laborables (por encima del 72%), mientras que en primavera y verano crece la participación del fin de semana. Este patrón estacional podría explicarse por una mayor propensión al uso recreativo de la bicicleta en los meses cálidos, aprovechando las mejores condiciones meteorológicas.

La distribución anual también se mantiene relativamente estable: en 2021 un 72,1% de los viajes se produjeron en días laborables, en 2022 un 70,5%, y en 2023 un 72,5%. Estas cifras consolidan la idea de una demanda fundamentalmente funcional, aunque con una base significativa de uso durante los fines de semana.

**Ilustración 15: Distribución porcentual del uso de BiciMAD por estación y tipo de día**



Fuente: elaboración propia

Para profundizar en la comprensión del comportamiento en días laborables, se ha incorporado un análisis adicional que clasifica los viajes en función de la hora de inicio del trayecto, agrupándolos en franjas horarias representativas de los diferentes momentos de la jornada. La siguiente tabla muestra la clasificación utilizada:

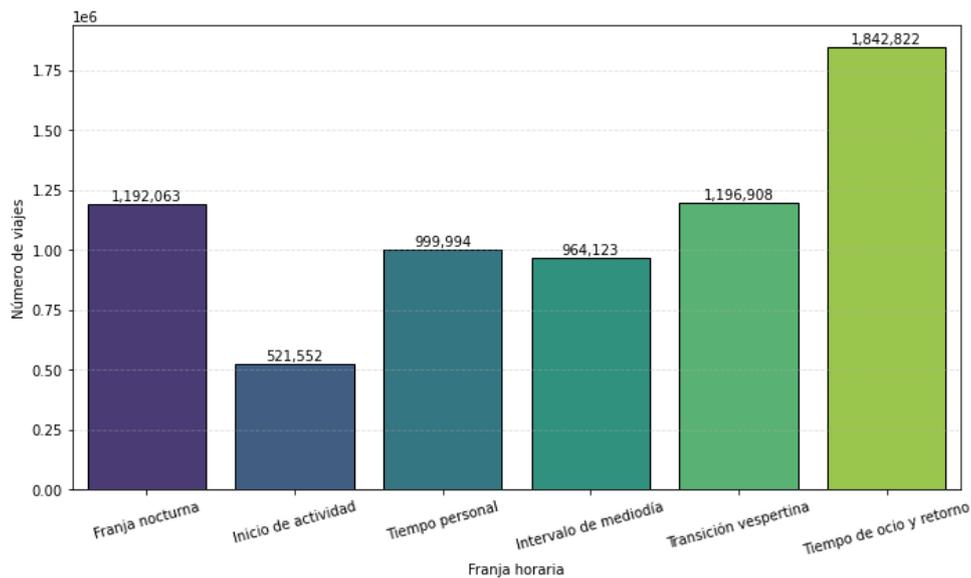
**Tabla 7: Clasificación de franjas horarias en días laborables**

Franja horaria	Descripción
06:00 – 09:59	Inicio de actividad diaria
10:00 – 12:59	Tiempo personal
13:00 – 15:59	Intervalo de mediodía
16:00 – 18:59	Transición vespertina
19:00 – 22:59	Tiempo de ocio y retorno
23:00 – 5:59	Franja nocturna

Fuente: elaboración propia

Tras aplicar esta segmentación, los resultados muestran una distribución de la demanda claramente orientada a los tramos de tarde, especialmente al "tiempo de ocio y retorno", que concentra el mayor número de trayectos. Le siguen en volumen los desplazamientos de "transición vespertina" y "intervalo de mediodía". En contraste, los viajes realizados durante el inicio de actividad diaria (6:00–9:59), si bien son significativos, representan una proporción inferior, lo que podría reflejar una combinación de factores como la competencia con otros modos de transporte o el impacto persistente del teletrabajo.

**Ilustración 16: Distribución de viajes en días laborables por franja horaria**



**Fuente: elaboración propia**

El análisis por franjas horarias revela una distribución claramente desigual en la demanda del sistema BiciMAD a lo largo del día. Destaca de forma muy significativa el uso intensivo durante el “tiempo de ocio y retorno” (de 19:00 a 22:00 h), que concentra el mayor número de viajes registrados, seguido por la “transición vespertina” y la “franja nocturna”. En cambio, el “inicio de actividad”, correspondiente a las primeras horas de la mañana, registra un volumen de uso notablemente inferior, lo que sugiere una menor adopción del sistema como medio de transporte habitual para desplazamientos a primera hora.

Estas pautas permiten afinar la planificación operativa del servicio, especialmente en lo relativo a la redistribución de bicicletas, el mantenimiento predictivo y la disponibilidad de personal técnico en horarios críticos. A diferencia de otros sistemas de bikesharing con un patrón más centrado en las horas punta laborales, BiciMAD muestra un fuerte componente de uso vinculado al ocio y al retorno al hogar en horarios vespertinos y nocturnos.

En consecuencia, se recomienda priorizar el refuerzo del sistema en los tramos de mayor intensidad (especialmente entre las 18:00 y las 22:00 h), asegurar la disponibilidad de anclajes libres y bicicletas operativas, e introducir medidas adaptadas a un uso más recreativo, como campañas estacionales o ajustes en las tarifas. Esta comprensión granular de la demanda horaria resulta esencial para garantizar una experiencia de usuario satisfactoria y una gestión eficiente del sistema.

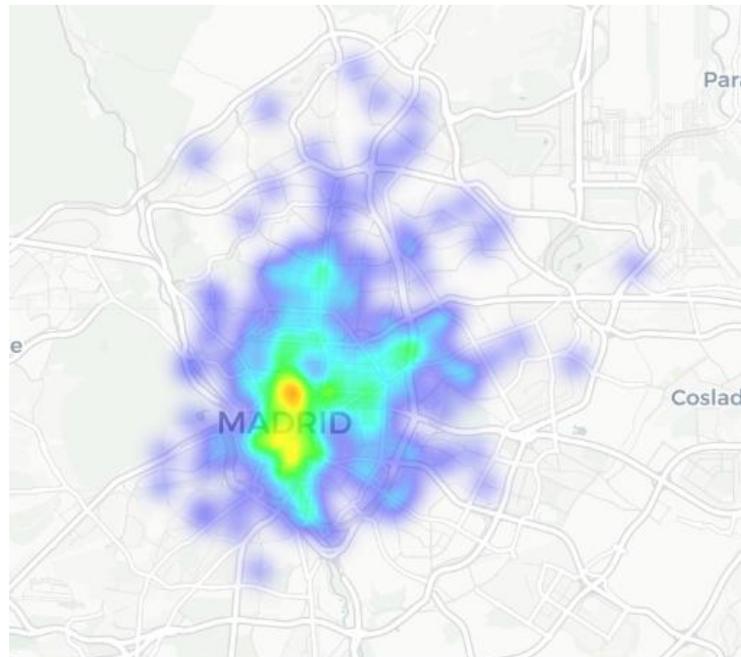
En la siguiente sección se abordará el análisis de la distribución geográfica de los viajes, con el fin de identificar las zonas con mayor concentración de actividad y evaluar posibles desigualdades territoriales en el acceso y uso del sistema BiciMAD.

### **5.1.5. Distribución geográfica de la demanda de BiciMad**

Con el objetivo de profundizar en los patrones espaciales de uso del sistema BiciMAD, se ha llevado a cabo un análisis de la distribución geográfica de los viajes realizados. Conocer qué zonas concentran un mayor número de desbloques resulta fundamental no solo para entender las dinámicas de movilidad urbana, sino también para optimizar la oferta del servicio, detectar posibles desequilibrios y planificar futuras ampliaciones de estaciones.

En una primera aproximación, se ha generado un mapa de calor de densidad del origen de los viajes (ver ilustración 17), sin aplicar segmentación por tipo de día. Este mapa permite identificar de forma visual las zonas con mayor intensidad de uso. Los resultados revelan una concentración muy marcada de la actividad en el área central de Madrid, especialmente en los barrios de Centro, Chamberí y Salamanca. Este patrón refuerza la hipótesis de que BiciMAD es una herramienta clave para la movilidad de corta y media distancia en entornos densamente poblados, con múltiples puntos de atracción como centros de trabajo, comercios, instituciones y nodos de transporte.

**Ilustración 17: Mapa de calor de densidad de origen viajes BiciMAD**



**Fuente: elaboración propia**

Para profundizar en las diferencias de uso según el tipo de día, se ha realizado un análisis comparativo mediante mapas de puntos proporcionales al número medio diario de desbloques por estación. En este caso, se han generado dos visualizaciones diferenciadas: una correspondiente a los días laborables (definidos como de lunes a viernes hasta las 16:00 horas) y otra para los fines de semana, entendidos como el periodo comprendido entre el viernes a partir de las 16:00 y el domingo por la noche, tal y como se ha aplicado en análisis anteriores.

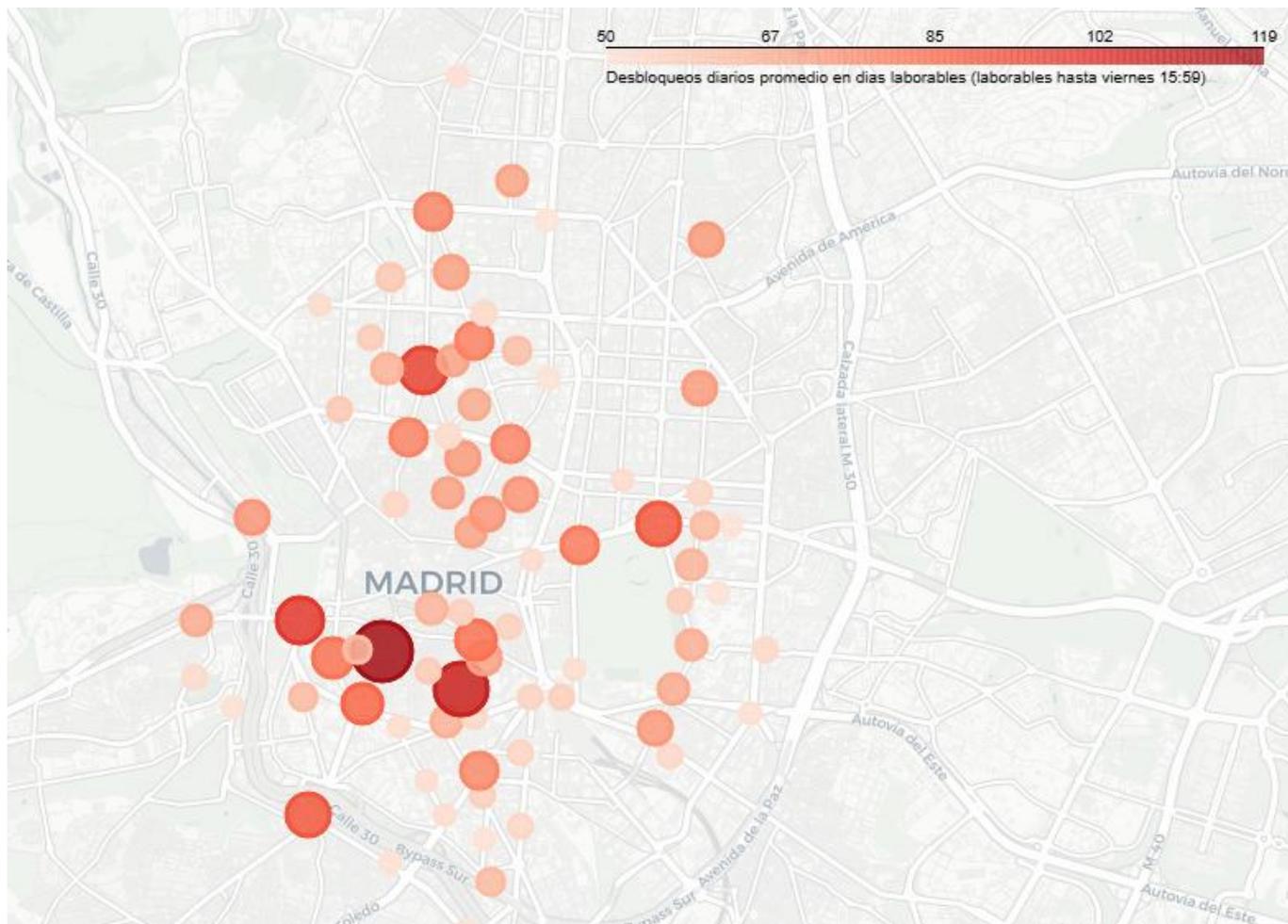
En la Ilustración 18, se muestra el resultado correspondiente a los días laborables. Para asegurar la comparabilidad y relevancia de las estaciones representadas, se ha aplicado un filtro previo que excluye aquellas ubicaciones con un número de anclajes inferior a 50, garantizando así que el análisis se centre en estaciones con capacidad operativa suficiente.

Cada punto en el mapa representa una estación del sistema BiciMAD, cuya superficie está escalada en función de la media diaria de desbloques registrada en días laborables. La intensidad del color, visible en la leyenda situada en la esquina superior derecha, permite interpretar el rango de valores observados: desde tonos más claros para estaciones con una actividad media reducida (por debajo de 50 desbloques diarios), hasta tonos rojos más intensos para aquellas que superan los 120 desbloques diarios, valor que representa el extremo superior de la escala observada.

Los resultados muestran una utilización más extendida del sistema en estos días, con una clara concentración a lo largo del eje norte-sur del centro urbano de Madrid. Destacan especialmente aquellas estaciones situadas en áreas

con fuerte componente funcional, como entornos de oficinas, intercambiadores de transporte, zonas comerciales o campus universitarios. Este patrón sugiere que el uso del sistema en días laborables responde principalmente a fines de movilidad obligada, ya sea como solución de primer o último kilómetro o como alternativa intermodal al transporte público.

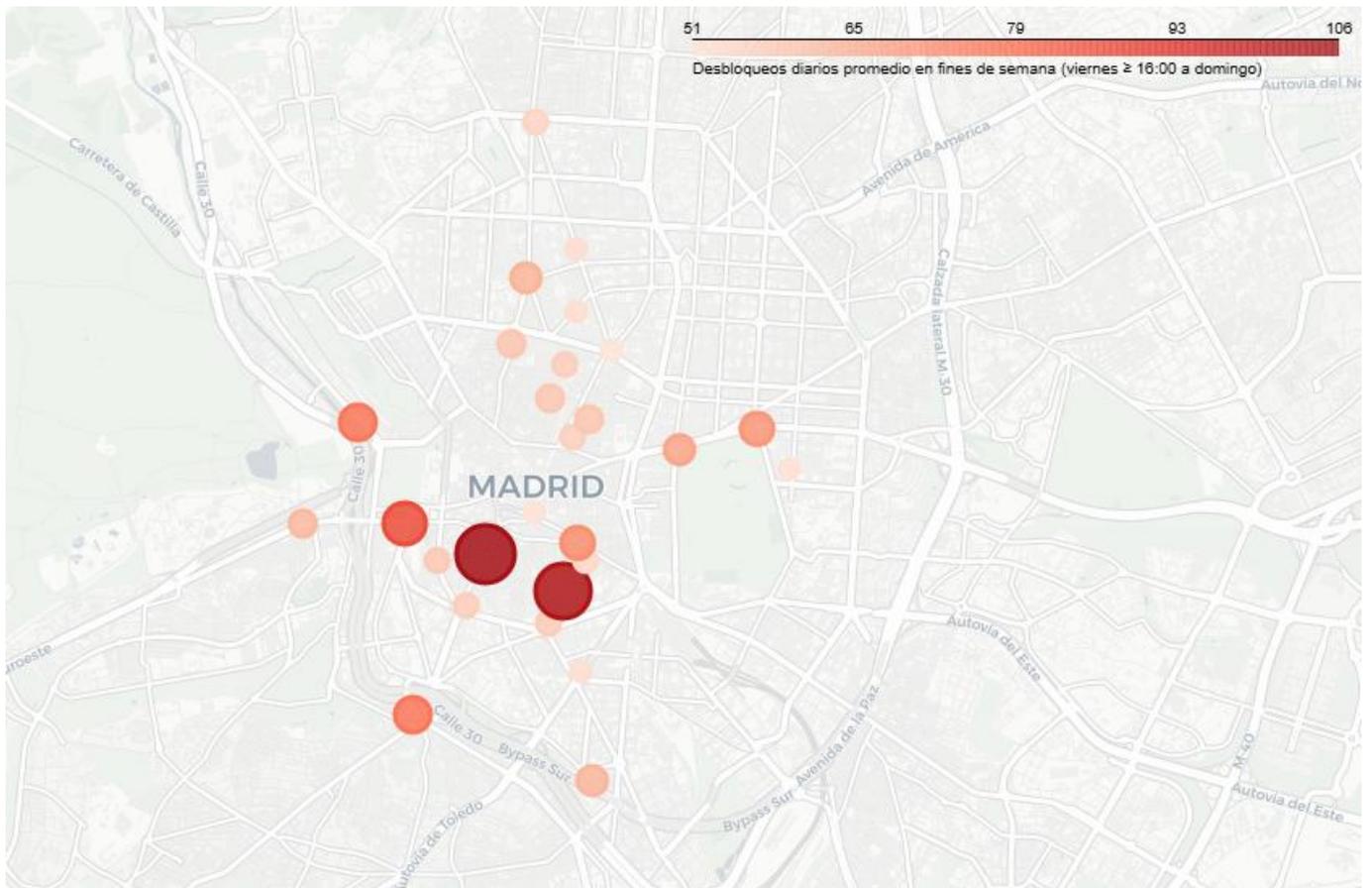
**Ilustración 18: Media diaria de desbloques por estación de BiciMAD en días laborables**



**Fuente: elaboración propia**

En cambio, durante los fines de semana en la ilustración 18, se aprecia una concentración aún mayor en el núcleo central de la ciudad, especialmente en áreas de alto atractivo turístico y de ocio. Zonas como Gran Vía, Sol o el entorno del parque de El Retiro registran una densidad elevada de viajes, lo que apunta a un uso más recreativo del sistema durante estos días. La demanda en fin de semana tiende a centrarse en desplazamientos más espontáneos y vinculados al disfrute del espacio urbano.

### Ilustración 19: Media diaria de desbloques por estación de BiciMAD en fines de semana



Fuente: elaboración propia

Este análisis territorial permite extraer varias conclusiones relevantes. En primer lugar, la localización geográfica de los desbloques se alinea con una lógica de densidad de actividad urbana, siendo el centro de la ciudad el principal núcleo de movilidad. En segundo lugar, se identifican ciertos puntos con baja actividad que podrían corresponder a estaciones infrutilizadas o directamente ausentes, lo que abre la posibilidad de reubicar recursos o priorizar nuevas instalaciones en zonas desatendidas. Finalmente, la segmentación por tipo de día sugiere que la planificación operativa debería adaptarse a estos patrones, reforzando la redistribución en horas punta entre semana y mejorando la cobertura en puntos de ocio durante el fin de semana.

En la siguiente sección se abordará el análisis de la red de estaciones meteorológicas fijas, con el objetivo de integrar las condiciones climáticas en el estudio de la demanda del sistema.

## 5.2. Análisis detallado de la base de datos de estaciones meteorológicas

Después de haber analizado en profundidad el conjunto de datos correspondiente a los registros de viajes del sistema BiciMAD, procedemos ahora a explorar la base de datos que recoge la información relativa a las estaciones meteorológicas fijas del municipio de Madrid. Esta base de datos nos permitirá entender el sistema de medición climática disponible en la ciudad y servirá como punto de partida para vincular, en etapas posteriores, los desplazamientos en bicicleta con las condiciones meteorológicas en las que se produjeron.

La información contenida en este Dataset es de carácter estático y describe las características fundamentales de cada una de las estaciones meteorológicas gestionadas por el Ayuntamiento de Madrid. En concreto, incluye variables como

la identificación de la estación, su localización geográfica (coordenadas en distintos sistemas de referencia), la dirección postal, la altitud sobre el nivel del mar y el conjunto de magnitudes meteorológicas que cada estación es capaz de registrar.

Esta primera aproximación al conjunto de estaciones resulta fundamental para determinar qué puntos de medición estarán disponibles a lo largo del análisis y con qué nivel de detalle se podrá asociar cada viaje de BiciMAD a las condiciones climáticas del entorno en el que se inicia. Asimismo, permitirá identificar posibles limitaciones en cuanto a cobertura territorial o disponibilidad de variables meteorológicas específicas, aspectos que condicionarán la calidad y alcance de los análisis posteriores.

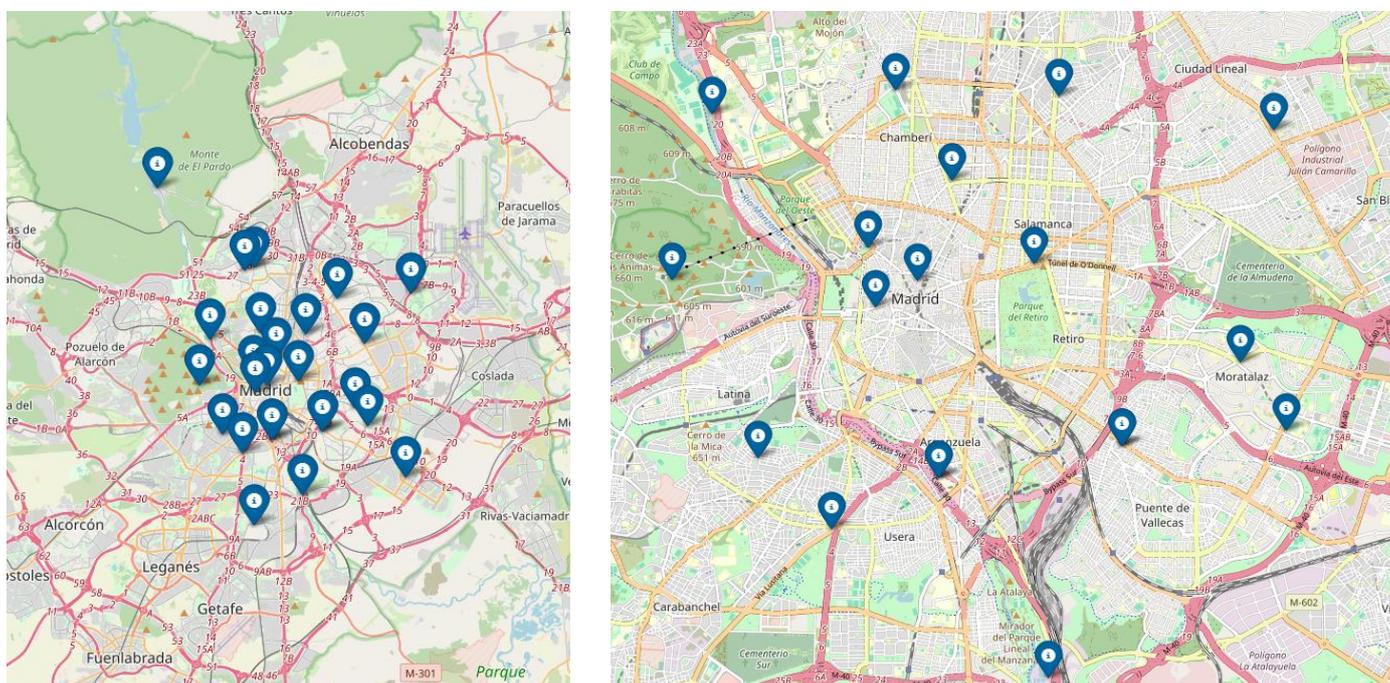
### 5.2.1. Análisis geográfico de las estaciones meteorológicas

Una vez descrita la estructura general del conjunto de datos correspondiente a las estaciones meteorológicas fijas del municipio de Madrid, se procede a examinar su distribución geográfica. Este análisis preliminar resulta de especial relevancia en el contexto del presente trabajo, dado que la ubicación exacta de cada estación será utilizada posteriormente para asignar, a cada trayecto registrado en BiciMAD, las condiciones meteorológicas más representativas del entorno en que se inició.

Dado que cada viaje de BiciMAD contiene información georreferenciada de su punto de partida, la estrategia metodológica que se adoptará en fases posteriores se basa en la asignación de variables climáticas por proximidad espacial. Para ello, resulta imprescindible disponer de una red de estaciones suficientemente distribuida por el término municipal, que permita cubrir de forma equilibrada los distintos distritos de la ciudad y minimizar las distancias entre los puntos de origen de los trayectos y los puntos de medición climática.

Las figuras que se presentan a continuación permiten visualizar, en distintos niveles de zoom, la localización espacial de las estaciones meteorológicas disponibles en el municipio.

**Ilustración 20: Distribución general de las estaciones meteorológicas fijas en el municipio de Madrid y su núcleo urbano.**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023).**

En la primera de ellas se ofrece una visión general de toda la ciudad de Madrid y su entorno metropolitano. Se aprecia una cobertura razonablemente densa en el interior del anillo de la M-30, con presencia de estaciones también en zonas periféricas como el distrito de Villaverde al sur, el entorno del Monte de El Pardo al noroeste, o áreas próximas a Coslada y Vicálvaro al este. Esta distribución permite anticipar una cobertura adecuada para la mayoría de los trayectos, aunque pueden detectarse ciertas áreas urbanas más alejadas con menor densidad de medición climática.

La segunda imagen, centrada en el núcleo urbano de Madrid, permite identificar con mayor claridad la presencia de estaciones en distritos tradicionalmente asociados a una elevada concentración de trayectos en BiciMAD. Zonas como Chamberí, Centro, Arganzuela, Retiro o Salamanca presentan estaciones relativamente próximas entre sí, lo que refuerza la representatividad de las variables meteorológicas en estos entornos. Este hecho resulta especialmente relevante si se tiene en cuenta que, históricamente, la demanda del sistema BiciMAD ha estado más concentrada en el centro de la ciudad.

Del análisis visual de la distribución geográfica pueden extraerse varias conclusiones preliminares. En primer lugar, la cobertura de estaciones en el centro urbano garantiza una adecuada capacidad de asignación meteorológica para un elevado número de trayectos. En segundo lugar, la existencia de zonas con menor densidad de estaciones, especialmente en distritos exteriores, deberá tenerse en cuenta a la hora de evaluar posibles sesgos o limitaciones del análisis. Por último, la presencia de estaciones relativamente próximas en algunas áreas permitirá implementar modelos de asociación basados en criterios de distancia, ya sea a través de cálculos euclídeos o mediante sistemas de redes viales.

En definitiva, el análisis de la localización geográfica de las estaciones constituye un paso esencial para garantizar la coherencia espacial entre los trayectos analizados y las condiciones meteorológicas que se les asignarán. A partir de esta base, será posible implementar un esquema riguroso de emparejamiento entre viajes y clima, con el objetivo de estudiar cómo varía la demanda del sistema BiciMAD en función de las condiciones atmosféricas de cada jornada.

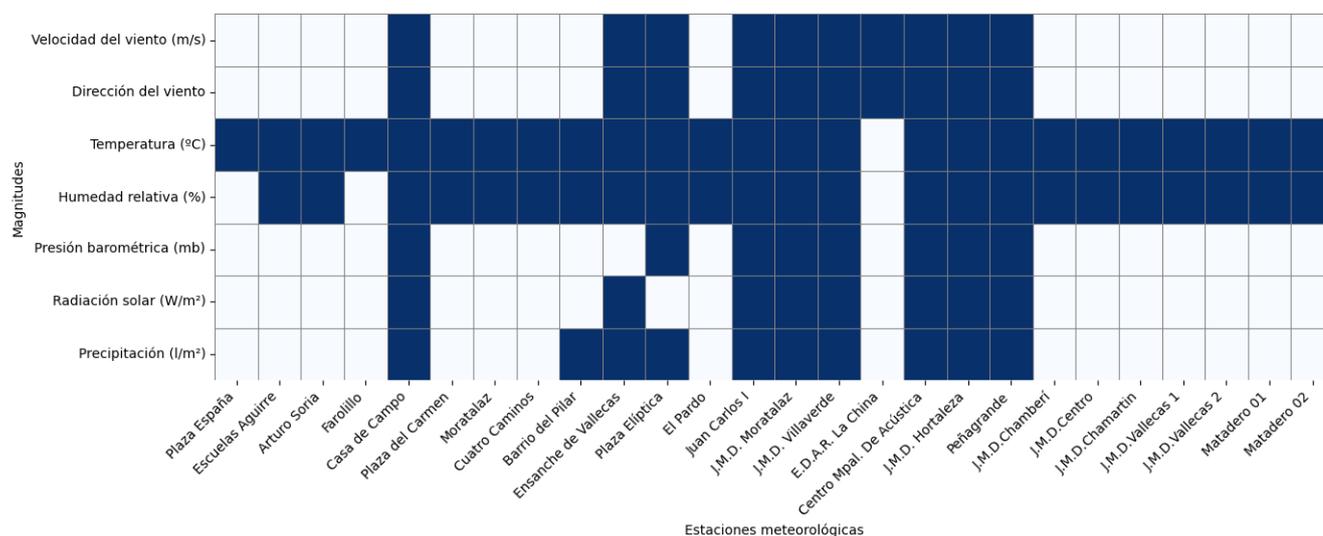
### **5.2.2. Análisis de la disposición del dato de las magnitudes geográficas**

Una vez analizada la localización espacial de las estaciones meteorológicas fijas del municipio de Madrid, se procede a examinar qué magnitudes climáticas son registradas por cada una de ellas. Este análisis resulta clave para valorar la calidad y consistencia del cruce de datos entre los trayectos del sistema BiciMAD y las condiciones meteorológicas presentes en el entorno urbano en el momento de realización de cada viaje.

La base de datos consultada proporciona información relativa a siete variables meteorológicas: velocidad del viento, dirección del viento, temperatura, humedad relativa, presión barométrica, radiación solar y precipitación. Sin embargo, la capacidad de medición no es uniforme en el conjunto de estaciones, por lo que resulta necesario evaluar de forma visual y sistemática la cobertura efectiva de cada magnitud a lo largo de la red de sensores instalados.

Con este fin, se ha elaborado una matriz de cobertura que permite representar, mediante un diagrama de calor, qué variables son medidas por cada estación meteorológica. Las celdas sombreadas indican la presencia de medición, mientras que las celdas en blanco corresponden a variables no registradas en dicha estación.

#### **Ilustración 21: Magnitudes meteorológicas registradas por estación**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

Tal y como se observa en la Figura 20, aunque el número total de estaciones disponibles es considerable, la mayoría de ellas no miden la totalidad de las variables contempladas. Las magnitudes más ampliamente registradas son la temperatura y la humedad relativa, seguidas a cierta distancia por la precipitación y la velocidad del viento. En cambio, variables como la presión barométrica, la radiación solar o la dirección del viento presentan una cobertura parcial o localizada, lo que limita su aplicabilidad en análisis comparativos o de carácter longitudinal.

A la luz de estos resultados, se ha considerado metodológicamente apropiado restringir el análisis a aquellas variables que presentan una mayor disponibilidad, una distribución geográfica más equilibrada y una relevancia contrastada en estudios previos sobre movilidad ciclista. De este modo, el conjunto de variables seleccionadas para el cruce con los registros de viajes de BiciMAD queda constituido por:

- Temperatura (°C)
- Humedad relativa (%)
- Precipitación (l/m²)
- Velocidad del viento (m/s)

La elección de estas variables responde tanto a criterios técnicos (presencia en la mayoría de las estaciones meteorológicas del municipio) como analíticos (evidencia empírica sobre su influencia en la movilidad ciclista urbana). Esta decisión permite preservar la solidez metodológica del estudio y asegurar la representatividad de los resultados obtenidos en las secciones posteriores.

### 5.3. Análisis detallado de la base de datos de la información meteorológica diaria

El tercer conjunto de datos empleado en el presente trabajo corresponde al registro diario de condiciones meteorológicas en el municipio de Madrid, con cobertura temporal desde junio de 2021 hasta febrero de 2023. Esta base de datos, procedente del Portal de Datos Abiertos del Ayuntamiento, contiene información detallada sobre diversas magnitudes climáticas recogidas por las estaciones fijas distribuidas por el término municipal.

A diferencia del fichero de estaciones meteorológicas (de carácter estático), este dataset presenta una estructura dinámica, en la que cada fila representa una combinación única de estación, fecha y variable meteorológica. Se incluyen mediciones horarias para cada día del periodo analizado, lo que permite construir indicadores diarios (por

ejemplo, medias, máximos o acumulados) que podrán asociarse a los trayectos registrados en BiciMAD en fechas coincidentes.

En términos estructurales, el Dataset se compone de columnas que identifican la estación de medición, la variable climática registrada (magnitud), el punto de muestreo, y los valores horarios correspondientes a las 24 horas del día. Esta organización posibilita una exploración detallada de la evolución diaria de variables como la temperatura, la humedad relativa, la velocidad del viento o la precipitación, todas ellas de interés en el marco del presente estudio.

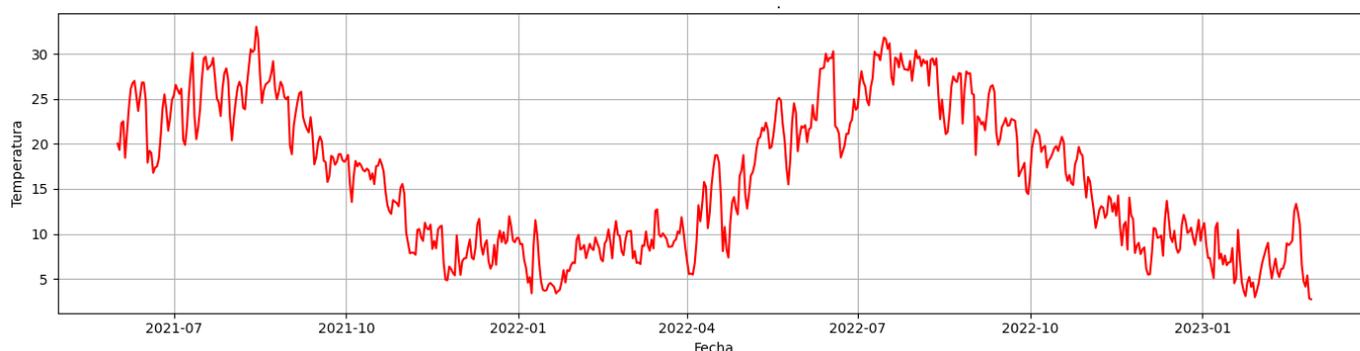
Este conjunto de datos constituye la base sobre la que se construirá, en secciones posteriores, el análisis cuantitativo de la influencia de las condiciones meteorológicas sobre el uso del sistema de bicicletas compartidas en la ciudad de Madrid.

### 5.3.1. Evolución temporal de las variables meteorológicas relevantes

Con el objetivo de comprender los patrones climáticos que podrían haber condicionado el uso del sistema BiciMAD, se ha realizado un análisis exploratorio de la evolución diaria de las principales magnitudes meteorológicas a lo largo del periodo comprendido entre junio de 2021 y febrero de 2023. Para ello, se ha representado gráficamente la serie temporal de cada variable mediante un gráfico de líneas, facilitando así la identificación de tendencias, estacionalidades y posibles eventos extremos.

Las variables consideradas han sido: temperatura media diaria (°C), humedad relativa media (%), velocidad media diaria del viento (m/s) y precipitación acumulada diaria (l/m<sup>2</sup>). A continuación, se describen los resultados observados para cada una de ellas.

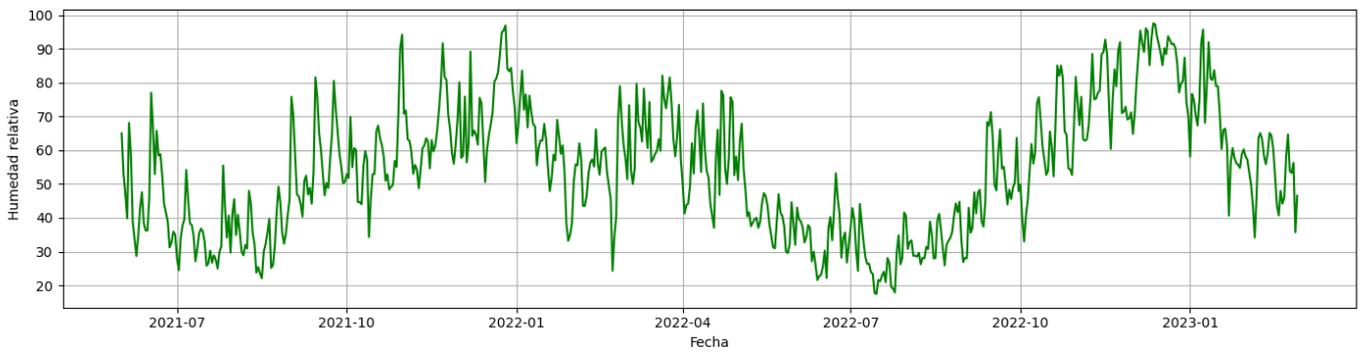
**Ilustración 22: Evolución diaria de la temperatura media (°C)**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

La evolución de la temperatura muestra un claro patrón estacional, con máximos durante los meses de verano y mínimos en invierno. Se observa una marcada oscilación anual, con valores que alcanzan o superan los 30 °C en los meses de julio y agosto, y descensos hasta los 3-5 °C durante los inviernos de 2021-2022 y 2022-2023. Este comportamiento es coherente con el clima continental propio de la ciudad de Madrid, y puede tener una relación directa con la intensidad del uso del sistema BiciMAD en los meses intermedios, cuando las condiciones térmicas son más favorables para el desplazamiento ciclista.

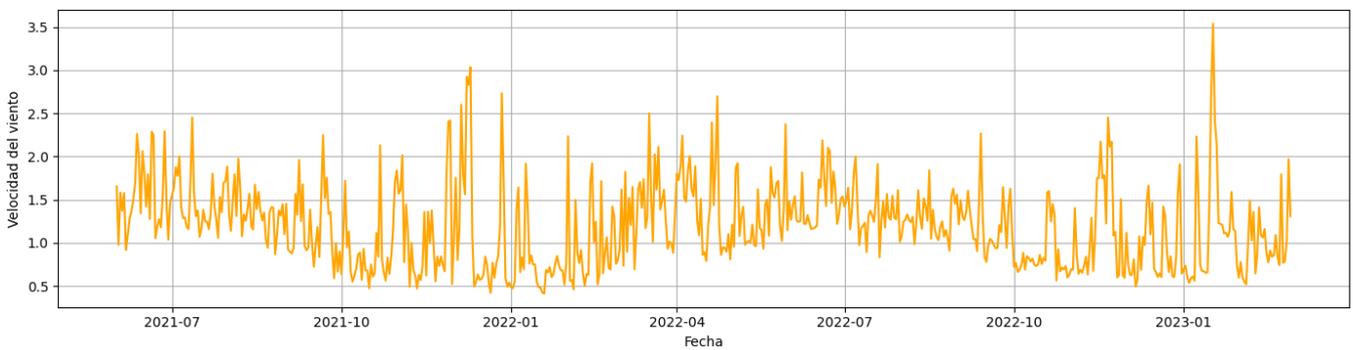
**Ilustración 23: Evolución diaria de la humedad relativa media (%)**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

La serie temporal de la humedad relativa presenta una variabilidad considerable a lo largo del tiempo, aunque también se aprecia cierta estacionalidad inversa a la temperatura: los valores más elevados se concentran en los meses fríos, superando frecuentemente el 80–90 %, mientras que los más bajos coinciden con el periodo estival, con registros por debajo del 30 % en varios días de los meses de julio, agosto y septiembre. Estos valores extremos de sequedad ambiental pueden influir en la percepción de confort térmico de los usuarios, especialmente en combinación con altas temperaturas.

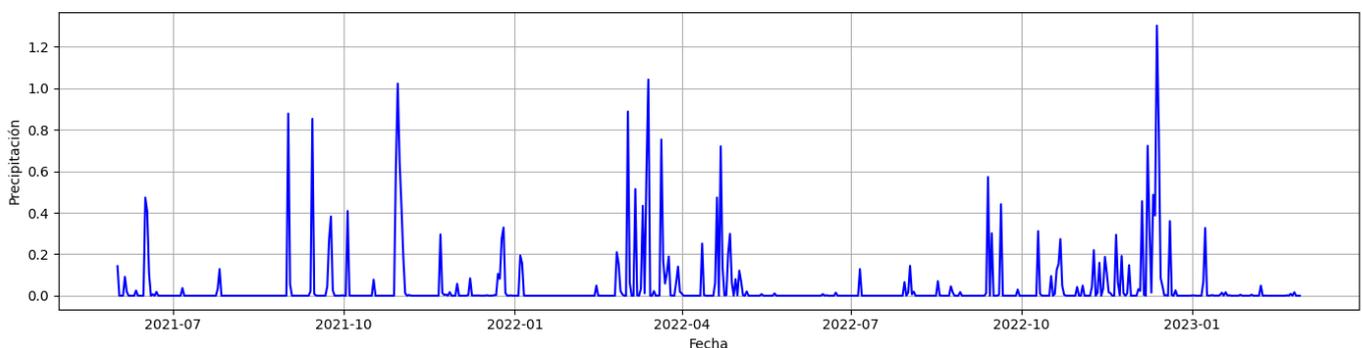
**Ilustración 24: Evolución diaria de la velocidad media del viento (m/s)**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

La velocidad del viento muestra una distribución más dispersa, con picos puntuales que en raras ocasiones superan los 3 m/s. No se observan patrones estacionales tan definidos como en las variables anteriores, aunque sí se detectan episodios de viento moderado en los meses invernales, posiblemente asociados a frentes atlánticos o borrascas. Dada su baja magnitud media, la velocidad del viento no parece representar un factor limitante significativo para la movilidad ciclista en la ciudad durante el periodo analizado.

**Ilustración 25: Evolución diaria de la precipitación acumulada (l/m²)**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

La variable de precipitación presenta una distribución marcadamente irregular, caracterizada por la alternancia de largos periodos secos y episodios aislados de lluvia. La mayoría de los días presentan valores nulos, y los registros de precipitación se concentran en días concretos, con intensidades que rara vez superan los 1 l/m<sup>2</sup>. Se aprecian algunos eventos más intensos entre los meses de otoño e invierno (octubre, noviembre y enero), aunque sin una estacionalidad definida. Este comportamiento refuerza la necesidad de considerar esta variable como un factor puntual más que continuo, y de explorar su impacto sobre la demanda diaria de BiciMAD a través de análisis específicos por umbrales.

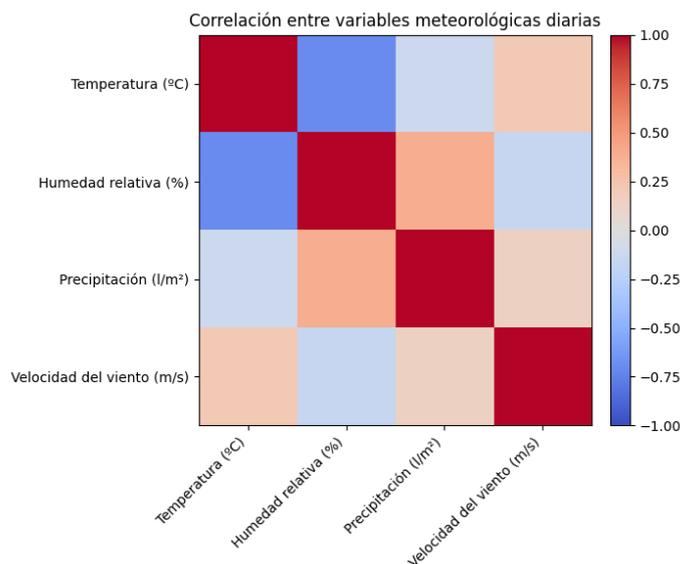
En este sentido, y con el objetivo de comprender mejor las interrelaciones existentes entre las distintas condiciones climáticas, el siguiente apartado aborda el estudio de las correlaciones entre variables meteorológicas, proporcionando una visión cuantitativa de su comportamiento conjunto en el contexto urbano de Madrid.

### 5.3.2. Análisis de correlación entre variables relevantes

Tras haber analizado individualmente la evolución temporal de las principales magnitudes meteorológicas resulta pertinente estudiar la relación existente entre estas variables. Este análisis de correlación permite identificar patrones conjuntos, redundancias o interacciones relevantes que pueden tener implicaciones tanto en la interpretación climática como en el estudio del impacto que estas condiciones puedan tener sobre el uso del sistema BiciMAD.

La correlación entre variables meteorológicas aporta una visión complementaria al análisis descriptivo, permitiendo detectar hasta qué punto ciertas magnitudes evolucionan de forma paralela o inversa. Además, este tipo de análisis es especialmente útil para reducir la multicolinealidad en modelos posteriores, seleccionar variables explicativas relevantes, y comprender mejor el comportamiento sistémico de la atmósfera urbana. Para ello, se ha calculado la matriz de correlación de Pearson entre las siguientes variables: temperatura media diaria, humedad relativa media, precipitación acumulada diaria y velocidad media del viento. Los resultados se presentan a continuación en la Figura 25.

**Ilustración 26: Matriz de correlación entre variables meteorológicas diarias**



**Fuente: Elaboración propia a partir de datos del Ayuntamiento de Madrid (2023)**

La matriz de correlación muestra varios patrones relevantes:

- Se observa una correlación negativa moderada entre la temperatura y la humedad relativa ( $r \approx -0.55$ ), lo cual es consistente con el comportamiento climático típico de Madrid: a mayor temperatura, menor contenido de humedad en el aire. Esta relación inversa puede tener implicaciones directas sobre la percepción térmica de los usuarios del sistema BiciMAD, al intensificarse la sensación de calor en días calurosos y secos.
- La precipitación presenta una correlación positiva débil con la humedad relativa ( $r \approx +0.40$ ), lo cual resulta esperable, ya que las condiciones húmedas suelen preceder o acompañar eventos de lluvia. Sin embargo, esta correlación no es extremadamente fuerte, posiblemente debido a la baja frecuencia de días lluviosos en el conjunto de datos.
- La velocidad del viento muestra una correlación débil o muy baja con el resto de las variables, lo que sugiere que se comporta de manera más autónoma respecto a la temperatura, la humedad o la lluvia. Esta independencia relativa refuerza su valor como variable complementaria en futuros modelos explicativos, aunque su variabilidad también la convierte en un predictor menos estable.
- Por último, no se observan valores de correlación fuertes entre ninguna pareja de variables ( $r > 0.75$  o  $r < -0.75$ ), lo cual indica que las cuatro magnitudes analizadas pueden considerarse estadísticamente no redundantes y, por tanto, potencialmente útiles para ser incluidas de forma conjunta en modelos explicativos del comportamiento de la demanda ciclista.

En conjunto, este análisis sugiere que existe una coherencia meteorológica esperada entre ciertas variables, como la temperatura y la humedad relativa, pero que también se preserva una diversidad de comportamientos entre las distintas magnitudes, lo que enriquece las posibilidades de análisis multivariable. Este diagnóstico será de utilidad en los próximos apartados del trabajo, en los que se estudiará cómo estas condiciones meteorológicas se relacionan con el volumen de trayectos registrados por el sistema BiciMAD.

Para ello, resulta imprescindible unificar en una única estructura analítica las tres fuentes de datos empleadas hasta ahora: el registro de viajes individuales, el conjunto de estaciones meteorológicas fijas, y el histórico diario de condiciones climáticas. En la siguiente sección se describe detalladamente el proceso de integración, emparejamiento y limpieza de estos datasets, así como los criterios metodológicos adoptados para asociar a cada viaje la información meteorológica correspondiente más precisa en términos espaciales y temporales.

## 6. Integración de datos

Con el objetivo de analizar el impacto de las condiciones atmosféricas sobre la demanda del sistema BiciMAD, fue necesario integrar diferentes fuentes de datos meteorológicos y de movilidad que originalmente se encontraban separadas. Este proceso de integración se articuló en dos fases fundamentales, permitiendo construir una única tabla final que contiene, para cada viaje, las condiciones climáticas existentes en el momento y lugar de su inicio.

La primera fase consistió en combinar dos conjuntos de datos meteorológicos complementarios. Por un lado, los datos de mediciones horarias incluían registros detallados de magnitudes como temperatura, humedad relativa, velocidad del viento o precipitación, captadas hora a hora y organizadas por estación y fecha. Este Dataset presentaba variables como ANO, MES, DIA, MAGNITUD, PUNTO\_MUESTREO y múltiples columnas de horas (H01, V01, ..., H24, V24), además de una columna agregada (avg) que promediaba los valores diarios por magnitud.

Por otro lado, los datos de estaciones meteorológicas fijas ofrecían información descriptiva y georreferenciada de cada estación de medición. Entre sus columnas destacaban `codigo_estacion`, `ESTACION`, `DIRECCION`, `LATITUD`, `LONGITUD` y otras variables técnicas vinculadas a la infraestructura de sensores.

El principal reto en esta fase era encontrar un identificador común que permitiera cruzar ambas fuentes. Aunque el campo `codigo_estacion` estaba disponible en la tabla de estaciones, no lo estaba explícitamente en el dataset horario. Sin embargo, a partir de la variable `PUNTO_MUESTREO` —una cadena del tipo `28079102_81_98`— fue posible extraer el código de ocho dígitos correspondiente a la estación. Esta transformación permitió normalizar el campo `codigo_estacion` en ambos conjuntos y realizar una fusión (`left join`) que enriqueció los registros horarios con la ubicación física de cada estación.

El resultado de esta primera fase fue un dataset meteorológico unificado, que contenía tanto los valores atmosféricos diarios por magnitud como la ubicación precisa de la estación que los midió.

**Ilustración 27: Modelo entidad-relación entre estaciones meteorológicas y datos horarios**



**Fuente: Elaboración propia**

Para facilitar la comprensión de esta estructura de unión, se presenta a continuación un modelo entidad-relación que representa gráficamente la relación entre ambas fuentes de datos. En este modelo, la entidad "Estaciones met." actúa como tabla maestra, identificada por la clave primaria `codigo_estacion`, mientras que "Datos met. diarios" recoge múltiples observaciones de magnitudes atmosféricas asociadas a cada estación a lo largo del tiempo. La relación entre ambas es de tipo uno-a-muchos, lo que permite asociar un número indefinido de mediciones meteorológicas diarias a una única estación meteorológica fija.

Una vez consolidada la base meteorológica, la segunda fase consistió en emparejar estos datos con los registros individuales de trayectos del sistema BiciMAD. El objetivo era asignar a cada viaje las condiciones meteorológicas más representativas del momento y lugar en que se inició.

Este emparejamiento se realizó en dos etapas. Primero se aseguró la coincidencia temporal exacta entre ambas fuentes: tanto la fecha de inicio del viaje como la de la medición meteorológica fueron transformadas al mismo formato (`date`) para poder filtrarlas por igualdad. Solo se consideraron aquellas observaciones meteorológicas que coincidieran exactamente con la fecha del trayecto.

La segunda etapa fue más compleja, ya que implicaba localizar la estación meteorológica más cercana al punto de desbloqueo de cada viaje. Para ello se utilizó la función `geodesic()`<sup>5</sup> de la librería `geopy`, que calcula la distancia

<sup>5</sup> La función `geodesic()` de la librería `geopy` calcula distancias entre coordenadas geográficas considerando la curvatura de la Tierra.

geográfica teniendo en cuenta la curvatura terrestre. Este cálculo se aplicó a un subconjunto reducido de coordenadas únicas (latitud, longitud y fecha) para optimizar el rendimiento del proceso. Para cada punto, el algoritmo identificaba la estación más próxima con registros válidos en ese día y extraía sus valores correspondientes de temperatura, humedad relativa, precipitación y velocidad del viento.

Estos datos fueron almacenados en un diccionario, convertido más tarde en un DataFrame estructurado. Para poder realizar la fusión con los trayectos, se generó una clave compuesta a partir de la latitud, longitud y fecha. Esta clave se usó tanto en la tabla de trayectos como en la de datos meteorológicos, permitiendo así una unión eficiente mediante merge<sup>6</sup>.

Finalmente, se añadió una última unión con el Dataset meteorológico original, a través del identificador de estación (ESTACION\_y) y la fecha, con el fin de conservar también la información espacial y administrativa de cada estación.

**Ilustración 28: Modelo entidad-relación entre datos meteorológicos y viajes de BiciMad**



**Fuente: Elaboración propia**

Para reflejar esta integración compleja, a continuación, se muestra un segundo modelo entidad-relación que ilustra gráficamente cómo se relacionan los datos meteorológicos con los viajes de BiciMAD. En este caso, la unión se produce por dos criterios simultáneos: coincidencia exacta de la fecha (FECHA ↔ fecha) y proximidad geográfica entre las coordenadas del punto de inicio del trayecto (lat\_bici, lon\_bici) y las coordenadas de la estación (LATITUD, LONGITUD). Esta lógica de emparejamiento se ha representado mediante una línea comentada, ya que no se trata de una relación física directa, sino de una asignación dinámica basada en distancia geográfica.

El resultado final fue una única tabla (df\_resultado) que contiene, para cada viaje de BiciMAD, su duración, localización, hora de inicio y las condiciones meteorológicas que se registraban en ese lugar y momento. Este conjunto de datos, limpio y enriquecido, se convierte así en la base operativa de todos los análisis posteriores del presente estudio, tanto desde un enfoque descriptivo como inferencial.

Este enfoque de integración asegura que el análisis sea temporal y espacialmente coherente, clave para obtener conclusiones robustas sobre la relación entre movilidad urbana y clima.

## 7. Análisis de la base de datos unificada

<sup>6</sup> Proceso de unión de datasets

Una vez completado el proceso de integración entre los registros de trayectos del sistema BiciMAD y los datos meteorológicos diarios proporcionados por las estaciones fijas de la red municipal, se ha generado una base de datos unificada que constituye el eje central de los análisis posteriores. Esta nueva estructura de datos incorpora, para cada viaje, no solo sus atributos básicos (como la fecha, duración, hora de inicio o coordenadas geográficas), sino también las condiciones climáticas registradas en la estación más próxima al punto de desbloqueo y en la misma fecha del trayecto.

El presente bloque analítico tiene como objetivo examinar en profundidad este conjunto de datos enriquecido. En primer lugar, se llevará a cabo una exploración descriptiva general del Dataset, orientada a validar su estructura, evaluar la distribución de sus variables clave y detectar posibles regularidades o anomalías relevantes desde el punto de vista analítico.

Posteriormente, se desarrollarán distintos análisis específicos destinados a investigar la relación entre la demanda de movilidad ciclista y diversas dimensiones contextuales, con especial énfasis en las condiciones atmosféricas. Entre otros aspectos, se analizará cómo varía el número y tipo de trayectos según la clasificación del día (laborable o fin de semana), las franjas horarias en las que se concentran los desplazamientos, y los valores de temperatura, precipitaciones, humedad relativa y velocidad del viento registrados en la ciudad de Madrid.

A través de este enfoque, se pretende no solo caracterizar el comportamiento de los usuarios del sistema BiciMAD, sino también evaluar en qué medida las variables meteorológicas inciden en la intensidad y patrón de uso del servicio, proporcionando evidencia empírica que pueda ser de utilidad tanto para la gestión operativa como para la formulación de políticas públicas orientadas a la promoción de una movilidad urbana sostenible.

### **7.1. Análisis exploratorio del Dataset unificado**

Una vez completado el proceso de integración entre los datos de trayectos del sistema BiciMAD y las condiciones meteorológicas diarias registradas en las estaciones fijas de Madrid, se ha generado una tabla unificada que recoge, para cada viaje, tanto sus atributos propios como las variables climáticas más representativas en el lugar y momento de su inicio. Este nuevo Dataset —base sobre la que se estructurará el análisis empírico posterior— requiere una revisión exploratoria inicial para verificar su consistencia y completitud.

En la Figura 1 se presentan dos vistas complementarias de este conjunto de datos final. En la parte superior se muestran los registros correspondientes a viajes individuales, donde se observan variables clave como la fecha (`fecha_x`), duración del trayecto (`trip_minutes`), estación de origen y destino, coordenadas geográficas del punto de desbloqueo, hora de inicio, y las variables meteorológicas integradas (temperatura, humedad relativa, precipitación, velocidad del viento), así como la estación meteorológica de referencia (`ESTACION_y`). Esta tabla confirma que la asignación de condiciones climáticas a cada trayecto se ha realizado correctamente, respetando tanto la proximidad espacial como la coincidencia temporal.

En la parte inferior de la figura se visualizan los registros meteorológicos diarios por estación para un mismo día. Aquí puede observarse cómo, para una determinada fecha (2021-06-01), algunas estaciones no disponen de valores para todas las magnitudes, especialmente en el caso de la "Precipitación" o la "Velocidad del viento". Este fenómeno responde a ausencias puntuales en los registros originales y será objeto de tratamiento posterior si se identifican patrones sistemáticos de datos faltantes.

### Ilustración 29: Exploración inicial del dataset final

fecha_x	idBike	trip_minutes	unlock_station_name	lock_station_name	lon_bici_x	lat_bici_x	hora_inicio	clave_union	lat_bici_y	...	fecha_y	Temperatura
2021-06-01	5282.0	55.15	Hortaleza	Antaño Martián	-3.697771	40.425191		0	(40.4251906, -3.6977715, 2021-06-01)	...	2021-06-01	20.779
2021-06-01	278.0	8.98	Plaza de la República Argentina	Prim	-3.685331	40.445411		0	(40.445411, -3.6853312, 2021-06-01)	...	2021-06-01	20.438
2021-06-01	6496.0	20.83	Galileo	Paseo de las Delicias	-3.710190	40.436400		0	(40.4364, -3.71019, 2021-06-01)	...	2021-06-01	20.058
2021-06-01	5766.0	10.03	Puerta de Madrid	Colegio Arquitectos	-3.680008	40.421501		0	(40.421501, -3.680008, 2021-06-01)	...	2021-06-01	20.479
2021-06-01	5407.0	11.82	Velázquez-Diego de León	Menéndez Pelayo	-3.683583	40.434333		0	(40.4343333, -3.6835833, 2021-06-01)	...	2021-06-01	21.167

Precipitación	Humedad relativa	Velocidad del viento	ESTACION_y	FECHA	codigo_estacion	LONGITUD	LATITUD
NaN	62.250	NaN	Plaza del Carmen	2021-06-01	28079035	-3.703166	40.419209
NaN	63.583	NaN	J.M.D.Chamartin	2021-06-01	28079111	-3.677813	40.444782
NaN	64.375	NaN	Cuatro Caminos	2021-06-01	28079038	-3.707130	40.445544
NaN	65.042	NaN	Escuelas Aguirre	2021-06-01	28079008	-3.682316	40.421553
NaN	59.833	NaN	J.M.D.Chamberí	2021-06-01	28079109	-3.697151	40.433041

Fuente: Elaboración propia

Estos resultados iniciales permiten confirmar que el procedimiento de fusión entre los trayectos y las mediciones meteorológicas se ha ejecutado de forma correcta a nivel técnico.

Para garantizar la fiabilidad del proceso de integración entre los trayectos de BiciMAD y los registros meteorológicos, se ha llevado a cabo una validación específica de la asignación espacial de las estaciones climáticas. El objetivo de este análisis es confirmar que, para cada viaje, se ha vinculado correctamente la estación meteorológica más próxima al punto de inicio del trayecto.

En la ilustración 29 se muestra el resultado de esta comprobación para un viaje concreto. Sobre un mapa de la ciudad de Madrid se ha representado, mediante un marcador rojo, la ubicación exacta donde se produjo el desbloqueo de la bicicleta. A su vez, las estaciones meteorológicas activas ese día se han representado con marcadores circulares de color azul, salvo la estación seleccionada por el algoritmo de integración, que aparece destacada en color verde.

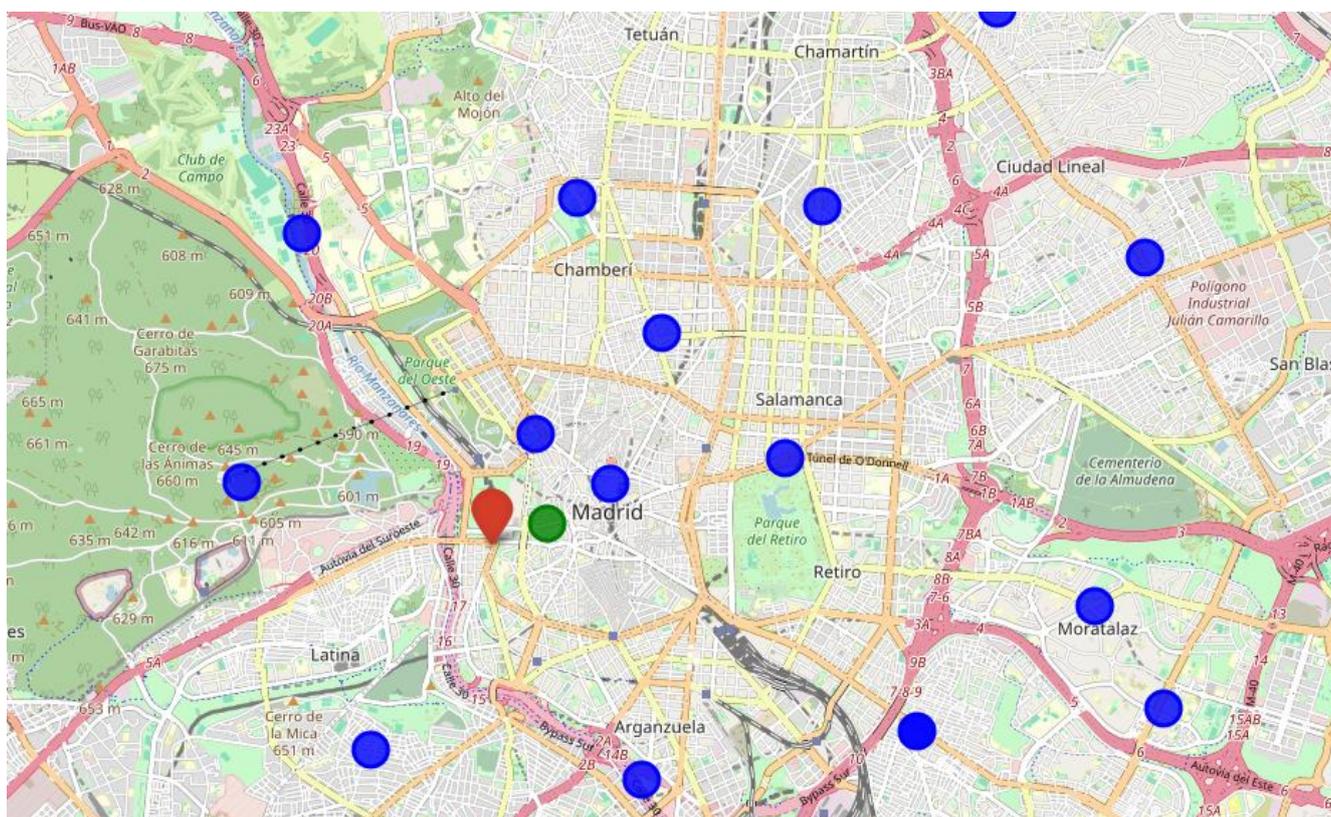
La asignación de esta estación se ha realizado utilizando la función `geodesic()` de la librería `Geopy`, que permite calcular la distancia más corta entre dos coordenadas geográficas teniendo en cuenta la curvatura terrestre. Este enfoque garantiza una mayor precisión que un cálculo de distancia euclidiana en un plano, lo cual es especialmente importante en entornos urbanos con una alta densidad de estaciones.

La visualización confirma que la estación seleccionada corresponde efectivamente a la más cercana al punto de inicio del trayecto, validando así la lógica de emparejamiento espacial implementada en el algoritmo. Esta revisión proporciona evidencia adicional de que el Dataset final está correctamente estructurado y que los datos

meteorológicos asociados a cada viaje reflejan con fidelidad las condiciones atmosféricas experimentadas por los usuarios del sistema en el momento del uso.

En la imagen se representa gráficamente este proceso de asignación. El marcador rojo indica el punto exacto en el que se desbloqueó la bicicleta, mientras que el marcador verde señala la estación meteorológica asignada a dicho trayecto, en base a su proximidad geográfica. Los puntos azules restantes corresponden al conjunto de estaciones meteorológicas disponibles en la red. La distancia mínima entre el punto de inicio del viaje y la estación meteorológica seleccionada refuerza la validez del criterio de emparejamiento empleado, asegurando que las condiciones climáticas asociadas a cada registro reflejen con precisión la realidad del entorno más inmediato.

**Ilustración 30: Verificación de la estación meteorológica más cercana al punto de desbloqueo del viaje**



**Fuente: Elaboración propia**

Esta comprobación se ha replicado de forma sistemática sobre un subconjunto representativo de trayectos, obteniéndose resultados coherentes en todos los casos analizados. A partir de esta verificación, se puede proceder con confianza al análisis detallado de la relación entre la demanda de viajes y las condiciones meteorológicas.

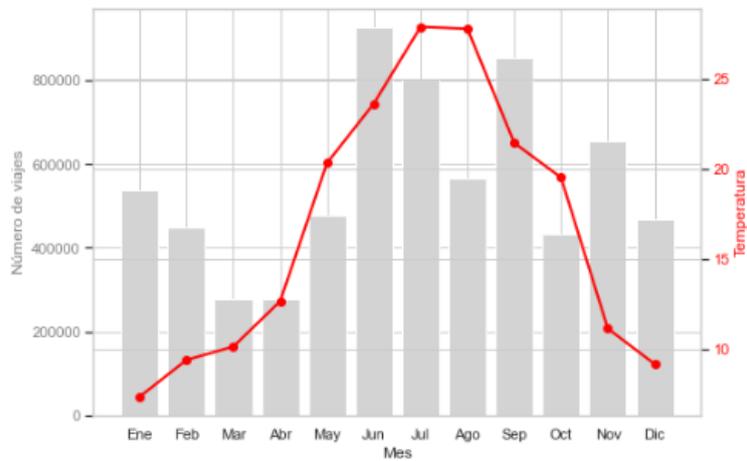
## **7.2. Análisis detallado del Dataset unificado**

Una vez completado el proceso de integración de los datos de movilidad y meteorología, se inicia el análisis detallado del Dataset unificado. El primer paso consiste en una exploración descriptiva inicial con el objetivo de observar la evolución mensual de la demanda de BiciMAD y su posible relación con las principales variables meteorológicas.

### 7.2.1. Evolución mensual de la demanda y variables meteorológicas

Para ello, se han elaborado cuatro gráficos combinados, en los que se representa el número total de viajes mensuales mediante un gráfico de barras y, sobre el mismo eje temporal, se superpone una línea correspondiente a la media mensual de cada una de las siguientes variables meteorológicas: temperatura, precipitación, humedad relativa y velocidad del viento.

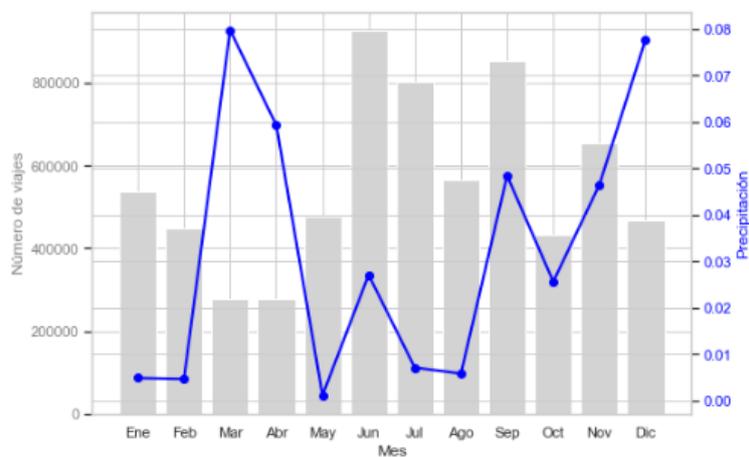
**Ilustración 31: Demanda mensual y temperatura media**



Fuente: Elaboración propia

El número de viajes muestra una clara estacionalidad, con picos entre los meses de mayo y septiembre, que coinciden con los valores más elevados de temperatura media. La relación positiva entre temperatura y uso del sistema es evidente en este gráfico, lo que sugiere que el clima cálido favorece significativamente la demanda de BiciMAD

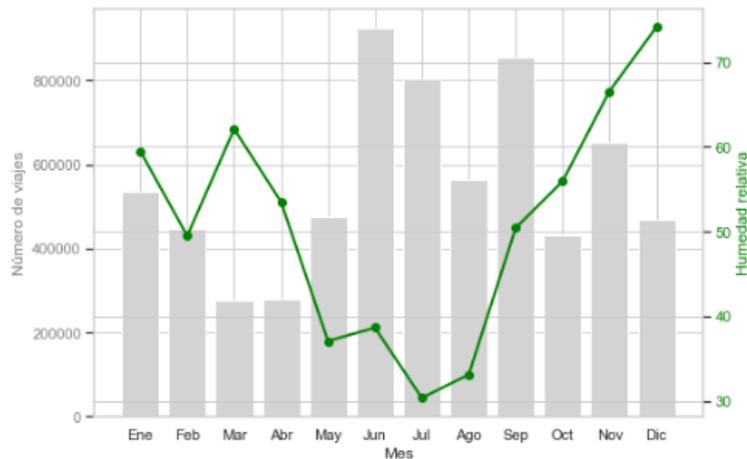
**Ilustración 32: Demanda mensual y precipitación**



Fuente: Elaboración propia

En este caso se observa una aparente relación inversa: los meses con menor precipitación (particularmente julio y agosto) tienden a registrar una mayor demanda. Por el contrario, los meses más lluviosos (como marzo, noviembre o diciembre) coinciden con un menor uso del sistema, lo que sugiere que la lluvia podría tener un efecto disuasorio relevante.

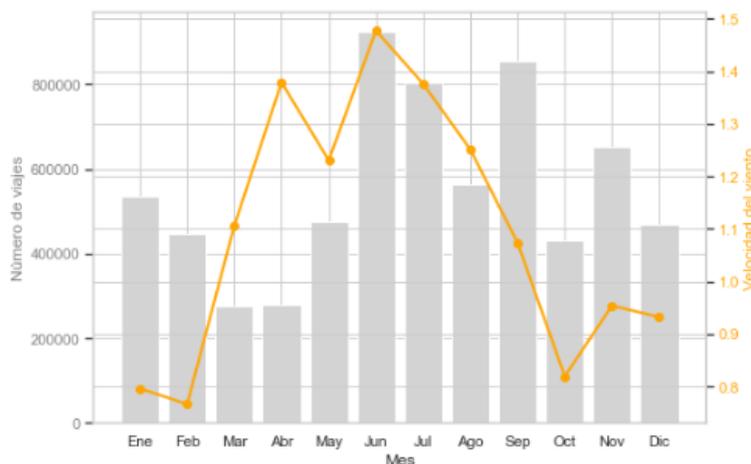
**Ilustración 33: Demanda mensual y humedad relativa**



**Fuente: Elaboración propia**

El comportamiento de la humedad relativa a lo largo del año es más irregular, pero se aprecia una cierta correspondencia negativa entre los valores más altos de humedad (invierno y finales de otoño) y los niveles bajos de uso del sistema. Aunque la relación no es tan evidente como en el caso de la temperatura o la precipitación, podría haber cierta influencia que tener en cuenta.

**Ilustración 34: Demanda mensual y velocidad del viento**



**Fuente: Elaboración propia**

La velocidad media del viento presenta una evolución relativamente estable a lo largo del año, con una ligera tendencia ascendente en los meses centrales. Sin embargo, no se aprecia una relación clara con la demanda. En este gráfico, el número de viajes parece responder más a otros factores (como temperatura y lluvia) que a la velocidad del viento.

Los gráficos exploratorios presentados permiten intuir que existe una posible relación entre determinadas condiciones meteorológicas (especialmente temperatura y precipitación) y la demanda de bicicletas compartidas. No obstante, esta visualización no permite establecer relaciones causales ni cuantificar con precisión la fuerza o significación de dichas asociaciones.

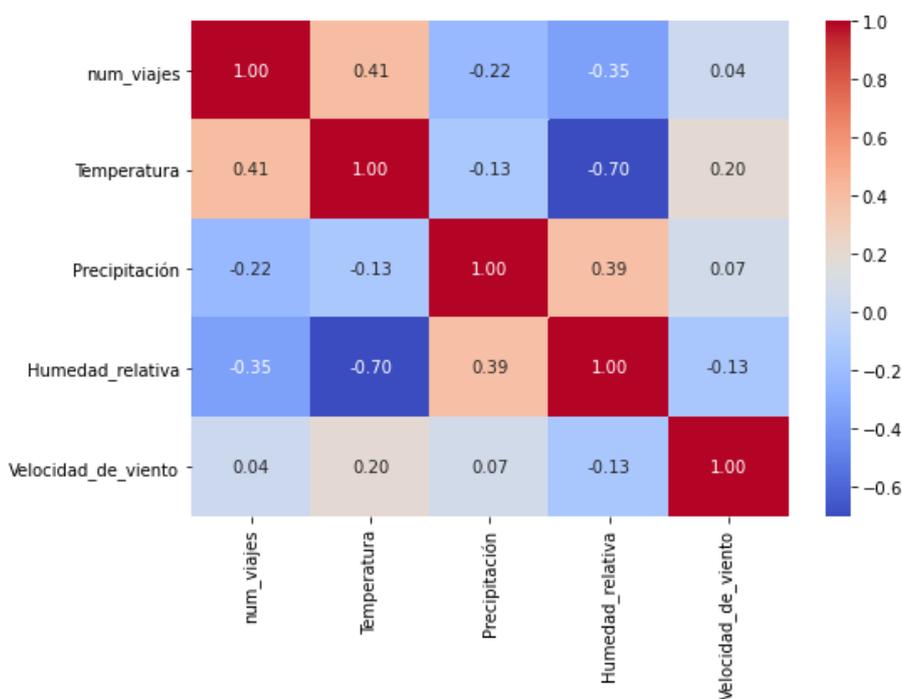
Por este motivo, en la siguiente sección se procederá a realizar un análisis de correlación y un modelo de regresión que permita verificar de manera más rigurosa si las relaciones observadas visualmente son estadísticamente significativas y si pueden considerarse factores explicativos consistentes del comportamiento de la demanda.

### 7.2.2. Evolución mensual de la demanda y variables meteorológicas

A partir de las evidencias visuales obtenidas en la fase exploratoria, resulta necesario avanzar hacia un análisis cuantitativo que permita medir con mayor precisión la intensidad y dirección de las relaciones entre las condiciones meteorológicas y la demanda del sistema BiciMAD. Para ello, se recurre en primer lugar al coeficiente de correlación de Pearson, que evalúa la fuerza de la asociación lineal entre dos variables, y posteriormente a un modelo de regresión lineal múltiple, que permite analizar el efecto conjunto de varias variables independientes sobre una variable dependiente, controlando posibles solapamientos o interacciones entre ellas.

Como primer paso, se ha calculado la matriz de correlaciones entre la variable num\_viajes (número de trayectos diarios) y las principales variables meteorológicas disponibles: temperatura, precipitación, humedad relativa y velocidad del viento.

**Ilustración 35: Matriz de correlación entre la demanda diaria y variables meteorológicas**



**Fuente: Elaboración propia**

Esta matriz permite una visualización clara de las relaciones existentes entre las variables analizadas. Los principales resultados obtenidos son los siguientes:

- La temperatura presenta una correlación positiva moderada con la demanda ( $r = 0.41$ ), lo cual respalda la hipótesis de que los días más cálidos están asociados con un mayor número de trayectos.
- La precipitación muestra una correlación negativa ( $r = -0.22$ ), indicando que los días lluviosos tienden a coincidir con una menor utilización del sistema.

- De forma similar, la humedad relativa también presenta una correlación negativa ( $r = -0.35$ ), lo que sugiere un posible efecto disuasorio sobre el uso de la bicicleta en días con elevados niveles de humedad ambiental.
- En cambio, la velocidad del viento apenas muestra relación con la demanda ( $r = 0.04$ ), lo que indica que su influencia directa podría ser marginal o inexistente en el contexto del sistema BiciMAD.

Si bien estas correlaciones permiten identificar tendencias y relaciones directas entre pares de variables, no permiten aislar el efecto individual de cada una cuando actúan simultáneamente. Por esta razón, es necesario avanzar hacia un modelo de regresión múltiple que permita evaluar el impacto neto de cada variable meteorológica sobre la demanda.

Se ha estimado un modelo de regresión lineal múltiple tomando como variable dependiente `num_viajes` (demanda diaria), y como variables independientes: temperatura, precipitación, humedad relativa y velocidad del viento.

**Ilustración 36: Resultados del modelo de regresión lineal múltiple con `num_viajes` como variable dependiente**

```

=====
                        OLS Regression Results
=====
Dep. Variable:          num_viajes      R-squared:                0.202
Model:                 OLS             Adj. R-squared:           0.196
Method:                Least Squares   F-statistic:              37.36
Date:                  Sun, 11 May 2025 Prob (F-statistic):       7.14e-28
Time:                  12:42:19        Log-Likelihood:          -5708.8
No. Observations:     596             AIC:                     1.143e+04
Df Residuals:         591             BIC:                     1.145e+04
Df Model:              4
Covariance Type:      nonrobust
=====
                        coef      std err      t      P>|t|      [0.025      0.975]
-----
const                9158.4723   1006.767    9.097    0.000    7181.196    1.11e+04
Temperatura          176.7674    25.557     6.916    0.000    126.573    226.962
Precipitación       -4869.4735  1283.259   -3.795    0.000   -7389.776   -2349.171
Humedad_relativa    -8.9093     11.956    -0.745    0.456    -32.392    14.573
Viento              -228.2361   280.978    -0.812    0.417    -780.074    323.601
=====
Omnibus:              43.339     Durbin-Watson:           0.482
Prob(Omnibus):        0.000     Jarque-Bera (JB):        143.434
Skew:                 0.255     Prob(JB):                7.14e-32
Kurtosis:             5.349     Cond. No.                538.
=====

```

Notes:  
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

**Fuente: Elaboración propia**

Del modelo obtenemos las siguientes conclusiones:

- La temperatura se confirma como una de las variables con mayor capacidad explicativa de la demanda. Su coeficiente positivo y altamente significativo ( $p < 0.001$ ) refuerza la hipótesis de que el clima cálido favorece el uso del sistema. El valor estimado (176.76) implica que, manteniéndose constantes el resto de las variables, un incremento de un grado centígrado en la temperatura media diaria se asocia, en promedio, con 177 trayectos adicionales por día. Este resultado es coherente con la intuición y con los patrones observados en el análisis exploratorio, y pone de manifiesto la relevancia del confort térmico como factor facilitador de la movilidad ciclista en entornos urbanos.
- En el caso de la precipitación, el modelo arroja un efecto negativo considerable y estadísticamente significativo (coef. = -4869.47,  $p < 0.001$ ). Este resultado refleja que la presencia de lluvia, incluso en cantidades moderadas, puede reducir de forma sustancial la demanda de bicicletas compartidas. En la práctica, este coeficiente indica que un aumento de un milímetro en la media diaria de precipitación puede traducirse en una disminución de

casi 5.000 viajes, lo cual subraya el fuerte carácter disuasorio de este factor climático. Este efecto resulta especialmente importante desde la perspectiva de la planificación operativa del sistema.

- Por el contrario, ni la humedad relativa ni la velocidad del viento resultan ser variables explicativas significativas en este modelo ( $p > 0.05$ ). Aunque ambas habían mostrado ciertas tendencias en la matriz de correlación y en los gráficos exploratorios, su efecto desaparece al considerar simultáneamente otras variables meteorológicas. Esto podría deberse a varios factores: por un lado, es posible que su impacto sea más indirecto o actúe en combinación con otras variables (por ejemplo, viento fuerte + lluvia); por otro, puede que su efecto solo sea relevante en determinados rangos extremos o contextos específicos, lo que escapa al alcance de un modelo lineal simple.
- El coeficiente de determinación ajustado ( $R^2$  ajustado = 0.196) indica que el modelo explica aproximadamente el 20 % de la variabilidad diaria de la demanda a partir exclusivamente de las condiciones meteorológicas. Aunque esta cifra pueda parecer limitada, debe interpretarse con cautela. El comportamiento de los usuarios de un sistema de transporte compartido depende también de factores estructurales (ubicación de estaciones, disponibilidad de bicicletas), sociales (eventos especiales, hábitos laborales o escolares) y contextuales (festivos, huelgas, pandemias, etc.), que no han sido incluidos en este modelo. En este sentido, el resultado obtenido es sólido y consistente, dado que se trata de un análisis centrado exclusivamente en variables climáticas.

En conjunto, los resultados del modelo permiten confirmar de forma estadísticamente significativa la relevancia de dos variables climáticas clave —temperatura y precipitación— en la explicación del uso diario del sistema BiciMAD. Este hallazgo refuerza la utilidad del análisis meteorológico para la predicción de la demanda, la planificación operativa y el diseño de políticas públicas de movilidad sostenible, especialmente en contextos urbanos sensibles a las variaciones estacionales.

### **7.2.3. Especificación cuadrática del efecto de la temperatura**

La variable temperatura constituye un caso singular dentro del conjunto de condiciones meteorológicas consideradas en este estudio, ya que su relación con la demanda de bicicletas compartidas no sigue necesariamente una lógica lineal. A diferencia de otras variables atmosféricas como la precipitación, cuyo efecto disuasorio es más uniforme, en el caso de la temperatura se puede observar un doble comportamiento: mientras que valores templados y agradables pueden incentivar el uso del sistema, temperaturas extremas —particularmente en rangos muy elevados— tienden a reducir la propensión a desplazarse en bicicleta.

Esta hipótesis de relación no lineal se alinea con lo observado en otros estudios empíricos (El-Assi et al., 2017; Gebhart & Noland, 2014), donde se constata que la demanda de sistemas de bikesharing tiende a aumentar con la temperatura hasta un umbral determinado, tras el cual comienza a descender. Este comportamiento se justifica por el mayor confort térmico en rangos moderados y el efecto disuasorio del calor intenso sobre la actividad física al aire libre.

Para capturar este fenómeno de forma más precisa, se ha estimado un modelo de regresión cuadrática, en el que se incorpora tanto el valor de la temperatura diaria como su término cuadrático. Esta especificación permite reflejar la posible existencia de una curva de tipo parabólico invertido (U invertida), donde el coeficiente lineal positivo se ve contrarrestado por un coeficiente cuadrático negativo, representando así un punto óptimo a partir del cual la relación se invierte.

El modelo se ha construido sobre la base de datos diarios, donde la variable dependiente es el número de viajes (`num_viajes`) registrados por día, y las variables independientes son la temperatura media diaria (`Temperatura`) y su cuadrado (`Temperatura_cuadrado`). El resultado del modelo es el siguiente:

### Ilustración 37: Modelo de regresión polinómica sobre la temperatura

OLS Regression Results						
Dep. Variable:	num_viajes	R-squared:	0.246			
Model:	OLS	Adj. R-squared:	0.243			
Method:	Least Squares	F-statistic:	96.79			
Date:	Fri, 23 May 2025	Prob (F-statistic):	4.13e-37			
Time:	09:26:58	Log-Likelihood:	-5704.9			
No. Observations:	597	AIC:	1.142e+04			
Df Residuals:	594	BIC:	1.143e+04			
Df Model:	2					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3075.2579	688.074	4.469	0.000	1723.904	4426.612
Temperatura	889.2374	88.589	10.038	0.000	715.252	1063.223
Temperatura_cuadrado	-19.5343	2.447	-7.982	0.000	-24.341	-14.728
Omnibus:	36.445	Durbin-Watson:	0.611			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	129.132			
Skew:	0.085	Prob(JB):	9.11e-29			
Kurtosis:	5.272	Cond. No.	2.30e+03			

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.  
 [2] The condition number is large, 2.3e+03. This might indicate that there are strong multicollinearity or other numerical problems.  
 Temperatura óptima estimada: 22.76 °C

Fuente: Elaboración propia

Los resultados confirman con claridad la existencia de un efecto parabólico de la temperatura sobre la demanda: a medida que la temperatura se incrementa hasta alrededor de 22.8 °C, la demanda aumenta; sin embargo, a partir de ese punto, nuevas subidas térmicas reducen el uso del sistema. Esta dinámica resulta especialmente relevante en contextos como el de Madrid, donde los meses de verano presentan temperaturas que frecuentemente superan los 35 °C, generando un entorno poco favorable para el uso de la bicicleta.

Este hallazgo refuerza la necesidad de tener en cuenta las relaciones no lineales en el análisis de movilidad urbana y pone de relieve la importancia de diseñar modelos que se ajusten al comportamiento real de los usuarios. Asimismo, aporta evidencia directa que matiza la hipótesis central del estudio, demostrando que la influencia de las variables meteorológicas no es siempre directa ni uniforme, sino que puede responder a curvas de comportamiento más complejas que deben modelarse adecuadamente para obtener conclusiones robustas.

### 7.3. Conclusiones del estudio

El presente trabajo ha permitido explorar con profundidad los patrones de uso del sistema BiciMAD en la ciudad de Madrid, integrando datos operativos del servicio con información meteorológica y temporal. A lo largo del análisis se ha buscado contrastar la hipótesis de que la demanda del sistema de bicicletas compartidas está significativamente influida por las condiciones meteorológicas, considerando como variables explicativas la temperatura, la precipitación, la humedad relativa y la velocidad del viento.

Los resultados obtenidos respaldan de forma consistente esta hipótesis. El análisis descriptivo y gráfico muestra una asociación visible entre las condiciones climáticas y la evolución mensual del número de viajes, con un incremento de la demanda en los meses con temperaturas más agradables y un descenso durante períodos con precipitaciones o niveles elevados de humedad. Los análisis de correlación y el modelo de regresión lineal múltiple inicial confirmaron que la temperatura media diaria tiene un efecto positivo y estadísticamente significativo sobre la demanda, mientras

que la precipitación actúa como un factor claramente disuasorio. En cambio, la humedad relativa y la velocidad del viento no mostraron efectos significativos una vez controladas las demás variables, lo que sugiere que su influencia es más marginal o dependiente de otros factores contextuales.

Con el objetivo de refinar el análisis, se incorporó un modelo de regresión con componente cuadrático para la variable temperatura, lo que permitió captar una relación no lineal más ajustada a la realidad del comportamiento observado. Esta especificación mostró que la demanda aumenta con la temperatura hasta alcanzar un umbral, a partir del cual comienza a disminuir, probablemente como consecuencia del efecto disuasorio del calor extremo sobre la actividad física. Esta mejora metodológica refuerza la validez del modelo y aporta evidencia empírica en línea con otros estudios internacionales.

Más allá de la dimensión meteorológica, el estudio ha revelado otros patrones estructurales de interés. El análisis por franjas horarias confirma que el sistema se utiliza de forma más intensiva en los tramos comprendidos entre las 16:00 y las 22:00 horas, siendo el denominado “tiempo de ocio y retorno” el periodo con mayor volumen de desplazamientos. Asimismo, la segmentación por tipo de día indica que los días laborables concentran más del 70 % de los viajes, reflejando un uso funcional y cotidiano del sistema, vinculado a desplazamientos obligados, mientras que los fines de semana presentan una demanda más dispersa y asociada al ocio.

A nivel espacial, los mapas generados evidencian una concentración del uso en el eje central de la ciudad, con especial intensidad en áreas próximas a intercambiadores de transporte, centros educativos, zonas comerciales y barrios con alta densidad de población activa. Estos resultados refuerzan la idea de que BiciMAD funciona como un sistema complementario al transporte público y al tejido urbano consolidado, adaptándose a las necesidades de movilidad del día a día en entornos urbanos densos.

No obstante, el estudio también ha puesto de relieve importantes limitaciones estructurales en los datos disponibles, especialmente en lo referente a su continuidad, formato y calidad. La interrupción en la publicación de datos tras la implementación del nuevo sistema BiciMAD en marzo de 2023, la variabilidad en la estructura de los datasets históricos, y las dificultades para asignar con total precisión las condiciones meteorológicas a cada momento del trayecto, limitan en parte el alcance del análisis. Aun así, la metodología desarrollada ha demostrado ser útil para identificar relaciones significativas y para generar una base sólida sobre la que construir futuros trabajos más amplios y detallados.

En conjunto, puede concluirse que el uso de sistemas de bicicletas compartidas como BiciMAD responde de forma clara tanto a patrones climáticos como temporales y espaciales, y que estos sistemas deben considerarse una pieza clave en el diseño de estrategias de movilidad urbana sostenible. En contextos donde la intermodalidad, la flexibilidad y la adaptación al entorno cobran protagonismo, comprender el impacto de factores como el clima resulta fundamental para optimizar su planificación, operación y promoción futura.

## **8. Limitaciones del estudio**

Como en toda investigación basada en datos reales y procesos de integración de fuentes heterogéneas, el presente estudio presenta una serie de limitaciones que deben tenerse en cuenta a la hora de interpretar sus resultados.

Una de las principales limitaciones ha sido la disponibilidad y continuidad de los datos del sistema BiciMAD. A partir de marzo de 2023, el sistema experimentó una transformación profunda en su infraestructura, incluyendo un cambio de modelo tecnológico y de estaciones, lo que afectó a la disponibilidad de los datos públicamente. Como consecuencia de esta transición, no ha sido posible incorporar al análisis trayectos posteriores a febrero de 2023, lo

que limita la representatividad temporal del estudio y la posibilidad de capturar los efectos de los nuevos patrones de uso tras la modernización del sistema.

En relación con lo anterior, otra dificultad relevante ha sido la heterogeneidad en la estructura de los datasets históricos, que presentan cambios frecuentes en las variables registradas, tanto en nombre como en contenido. Algunas columnas aparecen o desaparecen sin notificación clara, y en muchos casos contienen un porcentaje muy elevado de valores nulos, lo que ha obligado a realizar tareas complejas de limpieza, imputación o directamente exclusión de variables. Es especialmente destacable el caso de la variable `idBike`, cuya ausencia o presencia masiva de valores nulos en determinados meses ha dificultado el análisis de trayectos individuales o por tipo de bicicleta.

Asimismo, se han identificado archivos con datos incompletos o inservibles, como ocurre en el caso del mes de octubre de 2021, cuyo contenido fue descartado por carecer de registros válidos o presentar errores graves de formato. Esta discontinuidad ha afectado al equilibrio temporal del conjunto de datos y ha exigido decisiones de exclusión que deben ser tenidas en cuenta a la hora de valorar los resultados agregados.

Otro aspecto que considerar es la forma en que se ha realizado la asignación de variables meteorológicas a los trayectos. En este estudio, a cada viaje se le han asignado las condiciones meteorológicas registradas en la estación meteorológica más cercana al punto de desbloqueo de la bicicleta, considerando la fecha y hora de inicio. Si bien esta metodología permite una aproximación razonablemente precisa, presenta dos limitaciones fundamentales: en primer lugar, no se tiene en cuenta la variación de las condiciones climáticas durante la duración completa del trayecto, ya que se asigna una única medida por viaje. En segundo lugar, se asume que la estación más cercana al punto de partida representa de forma fidedigna las condiciones reales del trayecto, lo cual no siempre es exacto dada la escasa densidad de estaciones meteorológicas en el área urbana de Madrid y la posible variabilidad climática entre distritos o zonas con microclimas diferenciados.

Además, no se han considerado otras variables exógenas que podrían influir en la demanda del servicio, como el calendario laboral, eventos extraordinarios (huelgas, festividades, actos públicos), calidad del aire, presencia de alternativas modales o incluso aspectos del diseño urbano (pendiente, segregación de carriles bici, etc.). Aunque la finalidad del estudio era focalizarse en las condiciones meteorológicas como factor explicativo principal, es importante reconocer que la demanda real del sistema está determinada por múltiples factores que interactúan entre sí, y cuya omisión limita el poder explicativo de los modelos desarrollados.

Por último, cabe mencionar que los modelos estadísticos empleados, en particular la regresión lineal múltiple, ofrecen una visión agregada de la relación entre variables, pero no capturan posibles efectos no lineales, estacionales o dependientes del contexto, como puede ser la interacción entre meteorología y hora del día, o entre clima y tipo de día (laborable o festivo). Aunque se han planteado futuras extensiones metodológicas en este sentido, el modelo actual debe entenderse como una primera aproximación exploratoria.

En conjunto, estas limitaciones no invalidan los resultados obtenidos, pero sí invitan a interpretarlos con cautela y abren interesantes líneas de trabajo para futuras investigaciones con bases de datos más estables y modelos analíticos más avanzados.

## **9. Implicaciones prácticas y futuras líneas de investigación**

El análisis realizado en el presente trabajo proporciona no solo una mejor comprensión del comportamiento de la demanda del sistema de bicicletas compartidas BiciMAD en función de las condiciones meteorológicas, sino también

una serie de implicaciones relevantes desde el punto de vista operativo, estratégico y de planificación de políticas públicas.

### **9.1. Implicaciones prácticas**

Los resultados obtenidos en este estudio confirman que la demanda de BiciMAD está significativamente influida por variables meteorológicas, siendo la temperatura y la precipitación los factores que muestran mayor relevancia estadística. Este hallazgo tiene importantes repercusiones para la gestión eficiente del sistema y para el diseño de estrategias que permitan mejorar su rendimiento y calidad del servicio.

En primer lugar, el conocimiento anticipado de las condiciones meteorológicas permite adoptar medidas proactivas de optimización logística, particularmente en relación con el reparto y reequilibrio de bicicletas entre estaciones. En situaciones en las que se prevean condiciones meteorológicas favorables (como temperaturas templadas y ausencia de lluvia) es previsible un aumento significativo en la demanda, lo que justifica una planificación específica de la redistribución de bicicletas hacia las estaciones con mayor presión de uso, como las ubicadas en entornos laborales, universitarios o nodos de transporte intermodal.

Asimismo, estos resultados permiten orientar la gestión operativa del sistema en términos de mantenimiento, recarga de baterías y disponibilidad de anclajes, aspectos que cobran especial importancia en jornadas de alta demanda inducida por condiciones meteorológicas óptimas. La anticipación de estos picos facilita una mejor asignación de recursos y una reducción de posibles incidencias para el usuario final.

Por otro lado, la variabilidad climática identificada puede tener utilidad en el desarrollo de estrategias tarifarias diferenciadas o campañas promocionales adaptadas a la meteorología, contribuyendo a suavizar la estacionalidad del uso del sistema y a incentivar la movilidad ciclista en franjas temporales o condiciones menos favorables.

Desde una perspectiva más general, estos hallazgos también pueden servir de base para la integración de información meteorológica en las aplicaciones móviles del sistema, con el fin de proporcionar recomendaciones personalizadas o advertencias preventivas a los usuarios, mejorando así su experiencia y seguridad.

### **9.2. Futuras líneas de investigación**

El trabajo realizado abre diversas líneas de investigación que podrían profundizar en la relación entre condiciones climáticas y demanda, así como mejorar la robustez y aplicabilidad de los modelos desarrollados.

Una primera dirección consiste en replicar este estudio en otras ciudades, tanto nacionales como internacionales, con especial interés en aquellas que presentan condiciones climáticas distintas a las de Madrid (por ejemplo, climas más fríos, húmedos o ventosos), a fin de analizar el grado de generalización de los efectos observados.

En segundo lugar, resulta pertinente incorporar variables exógenas de naturaleza contextual, como la celebración de eventos especiales, festivos, huelgas, restricciones de movilidad o situaciones excepcionales como la pandemia de COVID-19, que pueden modificar sustancialmente los patrones de uso del sistema y cuya omisión puede limitar la capacidad explicativa de los modelos actuales.

Asimismo, una posible evolución metodológica consiste en desarrollar modelos de predicción más granulares, que analicen la demanda en franjas horarias concretas o en estaciones individuales, lo cual permitiría ajustar con mayor precisión la toma de decisiones operativas.

También resulta de interés analizar con mayor profundidad el comportamiento diferencial de usuarios según su perfil de uso (frecuente, ocasional, turístico, etc.), así como estudiar la posible interacción entre las variables meteorológicas y otras dimensiones temporales, como el tipo de día o la estación del año.

Finalmente, una línea de trabajo relevante sería actualizar y comparar los resultados del presente estudio con datos posteriores a marzo de 2023, momento en el cual BiciMAD experimentó una profunda transformación tecnológica y organizativa. Evaluar el impacto de estos cambios (incluyendo la gratuidad del servicio, la ampliación de estaciones y la renovación de la flota) en la sensibilidad del sistema frente a factores externos como la meteorología, permitiría generar una visión longitudinal del fenómeno.

En conjunto, estas propuestas no solo enriquecerían el conocimiento actual sobre movilidad ciclista urbana, sino que también contribuirían al diseño de sistemas más adaptativos, resilientes y orientados al usuario en un contexto de creciente necesidad de movilidad sostenible.

## 10. Anexos

El código fuente del análisis de datos y visualización se encuentra disponible en el siguiente repositorio público de GitHub: <https://github.com/juancar02/TFG-BiciMad-Juan-Carlos-Jimena.git>

## 11. Declaración de uso de herramientas de inteligencia artificial generativa en trabajos fin de grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo de Fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Juan Carlos Jimena Ordoñez, estudiante de Administración de Empresas y Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “ANÁLISIS DE LA DEMANDA DEL SISTEMA DE BiciMAD EN FUNCIÓN DE VARIABLES METEOROLÓGICAS EN LA CIUDAD DE MADRID”, declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a:

1. **Generación y estructuración de ideas de investigación:** Utilizado como apoyo para identificar temas relevantes y organizar posibles enfoques de análisis.
2. **Contraste de argumentos:** Empleado para analizar diferentes perspectivas sobre la hipótesis principal y enriquecer la discusión crítica del trabajo.
3. **Asistencia en la interpretación de resultados de código:** Usado para comprender mejor salidas de análisis preliminares de datos y orientar su interpretación adecuada.
4. **Mejora del estilo académico y claridad expositiva:** Utilizado para optimizar la redacción, asegurando un lenguaje preciso, coherente y adaptado al registro académico.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (se incluyen las referencias adecuadas en el TFG y he explicitado para qué se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

## 12. Referencias

- Acumen. (2023). Micromobility Market Size, Share & Trends Analysis Report. Recuperado de <https://www.acumen.com>
- Agencia Estatal de Meteorología (AEMET). (2023). Banco Nacional de Datos Climatológicos (OpenData). <https://opendata.aemet.es>
- Anaya, E., & Douch, D. (2019). Observatorio de la Bicicleta Pública en España: Análisis y tendencias del sistema de bicicletas compartidas. IDAE – Instituto para la Diversificación y Ahorro de la Energía.
- Ayuntamiento de Madrid. (2023). Datos abiertos del sistema BiciMAD. Plataforma de datos abiertos del Ayuntamiento de Madrid. <https://datos.madrid.es>
- Ayuntamiento de Madrid. (2023). Plan de renovación integral de BiciMAD.
- Bozzi, A., & Aguilera, A. (2021). Micromobility and urban sustainability: A review of recent developments. *Transportation Research Procedia*, 50, 127–134.
- Castro, A., et al. (2021). Determinantes de la viabilidad de los sistemas de bicicletas compartidas en España. *Transport Reviews*, 41(4), 533–552.
- Chen, Y., Shaheen, S., & Martin, E. (2018). Assessing the Safety Impacts of E-Scooter Sharing Programs. *Transportation Research Board*.
- Chen, Y., Shaheen, S. A., & Martin, E. (2018). The evolution of bikesharing: 10 years of experience and lessons learned. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(42), 57–68. <https://doi.org/10.1177/0361198118757381>
- Clockston, C., & Rojas Rueda, D. (2021). Health benefits of bikesharing systems: A systematic review. *Public Health Reviews*, 42(7), 1–12. <https://doi.org/10.1186/s40985-021-00156-4>
- DBK Informa. (2024). Informe sobre movilidad urbana compartida en España. Informa D&B.
- DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation*, 12(4), 41–56.
- Design, C. (2014). Copenhagenize Index 2015: Bicycle-friendly cities. Copenhagenize Design Co. <https://copenhagenizeindex.eu>
- El-Assi, W., Mahmoud, M. S., & Habib, K. N. (2017). Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation*, 44(3), 589–613. <https://doi.org/10.1007/s11116-015-9669-z>
- eldiario.es. (2022). La historia de las Bicicletas Blancas de Ámsterdam. Recuperado de <https://www.eldiario.es>
- El País. (2024, marzo). El auge del patinete y el teletrabajo frenan el uso de la bici en las ciudades. El País. <https://elpais.com>
- El País. (2024). El reto de la bicicleta pública frente a la movilidad gratuita. Recuperado de <https://elpais.com>
- EMT Madrid. (2023). Revolución BiciMAD: renovación integral del sistema de bicicletas públicas en Madrid.

EMT Madrid. (2021). Informe anual de sostenibilidad 2021.

Empresa Municipal de Transportes de Madrid. (2022). Resumen ejecutivo 2021.

<https://www.emtmadrid.es/Ficheros/Portal-Transparencia-2022/Resumen-Ejecutivo-2021.aspx>

EMT Madrid. (2021). Informe de gestión del servicio BiciMAD. Empresa Municipal de Transportes de Madrid.

Faghih-Imani, A., & Eluru, N. (2015). Analyzing bicycle-sharing system user destination choice preferences: Effects of user and station attributes. *Journal of Transport Geography*, 44, 218–225.

Fishman, E., Washington, S., & Haworth, N. (2012). Barriers and facilitators to public bicycle scheme use: A qualitative approach. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(6), 686–698.

<https://doi.org/10.1016/j.trf.2012.08.002>

Galatoulas, N., Papadopoulos, E., & Basbas, S. (2020). Micromobility and sustainable urban mobility: A review and agenda for future research. *Sustainable Cities and Society*, 61, 102343. <https://doi.org/10.1016/j.scs.2020.102343>

García-Palomares, J. C., Gutiérrez, J., & Latorre, M. (2022). Urban mobility transitions: The case of BiciMAD and electric bikesharing in Madrid. *Sustainability*, 14(9), 5027. <https://doi.org/10.3390/su14095027>

Gebhart, K., & Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation*, 41(6), 1205–1225. <https://doi.org/10.1007/s11116-014-9530-7>

Grau-Escolano, M., Rodríguez-Valverde, I., & Sánchez-García, L. (2024). Evaluación del impacto de la electrificación en sistemas públicos de bicicletas compartidas: El caso de BiciMAD. *Observatorio de Movilidad Metropolitana*.

Grau-Escolano, J., Bassolas, A., & Vicens, J. (2024). Cycling into the workshop: Predictive maintenance for Barcelona's bike-sharing system. arXiv:2404.17217.

Guide, R. (2011). *Bike-sharing in the United States: State of the Practice and Guide to Implementation*. Federal Highway Administration.

Hamilton, T. (2018). The impact of bike-sharing systems on urban traffic congestion. *Transportation Research Part A*, 118, 352–366.

Hampshire, R., & Marla, L. (2012). An analysis of bikeshare systems using big data. *Transportation Research Board Annual Meeting*. <https://doi.org/10.48550/arXiv.1210.0010>

Hu, Y., Chen, X., & Xie, K. (2022). Impacts of the COVID-19 pandemic on bike-sharing usage in urban environments: Evidence from Madrid. *Cities*, 120, 103456. <https://doi.org/10.1016/j.cities.2021.103456>

JCDecaux. (2023). Informe Valenbisi: Evolución y retos del sistema. JCDecaux España.

Leister, E., Vairo, R. M., Sims, S., & Bopp, M. (2018). Convenience, cost, and culture: Factors influencing bike share use among students and staff at a university. *Transportation Research Part A: Policy and Practice*, 118, 434–442. <https://doi.org/10.1016/j.tra.2018.09.005>

Ministerio de Transportes y Movilidad Sostenible. (s.f.). *Historia e implantación de los sistemas públicos de bicicleta en España*. Gobierno de España. <https://www.mitma.gob.es>

Ministerio para la Transición Ecológica y el Reto Demográfico. (2023). *Informe sobre políticas de electrificación de la micromovilidad urbana*. Gobierno de España.

Movilidad Eléctrica. (2023). El nuevo BiciMAD: inversión, expansión y desafíos. *Revista especializada en movilidad sostenible*. <https://movilidadelectronica.com>

- Oeschger, G., et al. (2020). Micromobility and urban transport: A review. *Transportation Research Part D: Transport and Environment*, 86, 102418.
- Open Data Madrid. (2025). Portal de Datos Abiertos del Ayuntamiento de Madrid. <https://datos.madrid.es>
- RedBici. (2024). Barómetro de la Bicicleta en España 2024. Red de Ciudades por la Bicicleta.
- Rixey, R. A. (2013). Station-level forecasting of bikesharing ridership. *Transportation Research Record: Journal of the Transportation Research Board*, 2387(1), 46–55. <https://doi.org/10.3141/2387-06>
- Ricci, M. (2015). Bike sharing: A review of evidence on impacts and processes of implementation and operation. *Research in Transportation Business & Management*, 15, 28–38.
- Rixey, R. A. (2013). Station-level forecasting of bikesharing ridership. *Transportation Research Record*, 2387(1), 46–55.
- Sanmiguel-Rodríguez, A. (2022). Evolución del uso de la bicicleta pública en España: factores determinantes y políticas públicas. *Revista de Estudios Urbanos*, 18(2), 45–67.
- Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future. *Transportation Research Record*, 2143(1), 159–167.
- Shaheen, S. A., Guzman, S., & Zhang, H. (2012). Public Bikesharing in North America: Early Operator and User Understanding. *Transportation Research Record*, 2387(1), 83–92.
- Sun, F. (2018). Problems and prospects of dockless bike-sharing in China: Evidence from policy documents. *Sustainability*, 10(6), 1939. <https://doi.org/10.3390/su10061939>
- Tang, Z., Pan, H., & Shen, Q. (2011). Bike-sharing systems in China and their implications for urban transportation planning. *Transport Reviews*, 31(6), 733–754. <https://doi.org/10.1080/01441647.2011.618203>
- Verenzuela Gómez, R. (2023). Impacto de la infraestructura ciclista en la adopción del bikesharing: el caso de Barcelona. *Revista de Movilidad Urbana*, 11(1), 62–79.
- WRI – World Resources Institute. (2018). The last mile by bike: Integrating bikesharing into urban transport. WRI Ross Center for Sustainable Cities. <https://wrirosscities.org/news/last-mile-bike>
- Zanotto, M. (2014). Public bicycle systems and their impact on user behaviour: A case study in Barcelona. *Transportation Research Procedia*, 4, 198–211. <https://doi.org/10.1016/j.trpro.2014.11.016>
- Zhang, Y., & Mi, Z. (2018). Environmental benefits of bike sharing: A big data-based analysis. *Applied Energy*, 220, 296–301.
- Zhu, Y., Yan, X., Liu, H., & Wang, L. (2018). Understanding bike-sharing travel behavior: A perspective of the extended theory of planned behavior. *Journal of Advanced Transportation*, 2018, Article ID 7407638.