

Article

Forecasting Moped Scooter-Sharing Travel Demand Using a Machine Learning Approach

Tulio Silveira-Santos ^{1,*} , Thais Rangel ^{1,2} , Juan Gomez ¹  and Jose Manuel Vassallo ¹ 

¹ Transport Research Center (TRANSyT), Universidad Politécnica de Madrid, 28040 Madrid, Spain; thais.rangel@upm.es (T.R.); juan.gomez.sanchez@upm.es (J.G.); josemanuel.vassallo@upm.es (J.M.V.)

² Department of Organizational Engineering, Business Administration and Statistics, Universidad Politécnica de Madrid, 28012 Madrid, Spain

* Correspondence: tulio.silveira@upm.es

Abstract: The increasing popularity of moped scooter-sharing as a direct and eco-friendly transportation option highlights the need to understand travel demand for effective urban planning and transportation management. This study explores the use of machine learning techniques to forecast travel demand for moped scooter-sharing services in Madrid, Spain, based on origin–destination trip data. A comprehensive dataset was utilized, encompassing sociodemographic characteristics, travel attraction centers, transportation network attributes, policy-related variables, and distance impedance. Two supervised machine learning models, linear regression and random forest, were employed to predict travel demand patterns. The results revealed the effectiveness of ensemble learning methods, particularly the random forest model, in accurately predicting travel demand and capturing complex feature relationships. The feature scores emphasize the importance of neighborhood characteristics such as tourist accommodations, public administration centers, regulated parking, and commercial centers, along with the critical role of trip distance. Users’ preference for short-distance trips within the city highlights the appeal of these services for urban mobility. The findings have implications for urban planning and transportation decision-making to better accommodate travel patterns, improve the overall transportation system, and inform policy recommendations to enhance intermodal connectivity and sustainable urban mobility.

Keywords: moped scooter-sharing; shared mobility; urban mobility; travel demand; machine learning; transportation management



Citation: Silveira-Santos, T.; Rangel, T.; Gomez, J.; Vassallo, J.M.

Forecasting Moped Scooter-Sharing Travel Demand Using a Machine Learning Approach. *Sustainability* **2024**, *16*, 5305. <https://doi.org/10.3390/su16135305>

Academic Editor: Zihan Kan

Received: 28 May 2024

Revised: 14 June 2024

Accepted: 19 June 2024

Published: 21 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rise of shared mobility services has become a popular subject in urban transportation studies in recent years. The emergence of new modes such as carsharing, bike-sharing, and moped scooter-sharing has transformed how people move around cities. Instead of requiring vehicle ownership, the notion of “shared mobility” has been defined as providing users with short-term access to shared automobiles based on their needs and convenience [1]. The term “micro-mobility” first appeared to refer to shared, low-speed vehicles like mopeds and kick-style scooters, which are becoming more and more popular [2]. This shift has led to new challenges and opportunities for city planners, policymakers, and researchers to address issues such as congestion, emissions, and equitable access to transportation.

Among these shared mobility services, moped scooter-sharing has gained popularity in many cities worldwide. These services allow users to rent and share motorcycles for short trips and have the potential to offer an alternative mode of transportation that is faster, more flexible, and cheaper than traditional modes such as public transit or taxi services [3]. However, understanding the travel demand patterns of moped scooter-sharing services is crucial for ensuring their sustainable growth and efficient use of resources and assessing their potential role within the urban transport system. This study aims

to bridge this gap by leveraging machine learning techniques to forecast travel demand, providing a nuanced understanding of travel patterns that can inform urban planning and transportation policies.

Accurately predicting travel demand for moped scooter-sharing can have significant policy implications. For urban planners and policymakers, such predictions can inform infrastructure planning, traffic management, and environmental policies. By understanding when and where the demand for these services is highest, cities can better allocate resources, optimize parking regulations, and enhance the integration of moped scooter-sharing with other public transportation options. This may result in less pollution, fewer traffic jams, and more fair access to transit for locals. Moreover, from an economic perspective, efficient deployment of moped scooter-sharing services can support local businesses by increasing accessibility and foot traffic.

Applications of big data and machine learning techniques are being employed extensively globally in a range of fields [4], offering more comprehensive data on transportation networks and assisting businesses and government agencies in better organizing and arranging their operations [5–7]. The use of machine learning techniques has been gaining traction in the transportation industry due to their ability to analyze and predict travel demand patterns.

Machine learning techniques have been explored over the last few years to estimate travel demand for a variety of transport modes, such as metro [8], taxi [9,10], ride-hailing [11], and bike-sharing [12,13]. However, the use of moped scooter-sharing data has received limited attention, despite its potential to provide valuable insights into travel demand patterns in urban areas. Moreover, the existing studies have focused primarily on predicting travel demand for specific periods or events, such as rush hour or major sports events. In contrast, this study seeks to explore the feasibility of predicting global travel demand for an origin–destination pair at the neighborhood level on an annual basis, which could inform long-term planning and decision-making.

This study aims to explore the use of machine learning techniques to forecast travel demand patterns in Madrid, Spain, based on historical data from the moped scooter-sharing service provided by a private mobility operator. Data have been collected by the company Acciona, with approximately 1.5 million moped scooter-sharing trips registered in the city in 2022.

To achieve this objective, this paper explores machine learning techniques to develop predictive models that can estimate the travel demand of moped scooter-sharing among origin and destination pairs in each neighborhood of the city of Madrid. These models have been based on a combination of travel demand data and additional explanatory variables, including (i) sociodemographic characteristics of the neighborhood, (ii) travel attraction centers, (iii) transportation network attributes, (iv) policy-related variables (such as the application of low-emission zones—LEZs—and regulated on-street parking areas), and (v) the distance impedance variable.

Using the scikit-learn open-source machine learning toolkit, two supervised machine learning models (such as random forest and linear regression) were employed to forecast travel demand trends. The former was used to visualize feature coefficients, and the latter was used to visualize feature importance scores. The findings could inform policy decisions related to transportation planning and infrastructure investment, as well as provide insights into the behavior of travelers in urban areas.

The paper is divided into seven sections in addition to this introduction. Background and literature review are provided in Section 2. Section 3 provides an overview of the study context, and Section 4 provides a thorough description of the data. Section 5 provides an overview of the methodology used to obtain the research results. Section 6 presents the results, and Section 7 offers a discussion that follows. Conclusions and policy recommendations are presented in Section 8 at the end.

2. Background and Literature Review

The scientific literature on travel demand prediction using machine learning has seen significant growth in recent years, driven by advancements in machine learning techniques [14] and the availability of large-scale datasets [15]. Traditional time series models, such as the auto-regressive integrated moving average (ARIMA), have long been used for travel demand forecasting [16–18]. However, these models have limitations in capturing complex nonlinear relationships and managing large datasets. In contrast, machine learning models, given their ability to capture complex interrelationships and nonlinearities, have shown superior predictive accuracy and model fit in travel demand estimates for various modes of transportation [15], including metro [8], bus [19], taxi [9,10], ride-hailing [11,20], bike-sharing [12,13], and scooter sharing [21].

For instance, Ma et al. [8] proposed a deep neural network architecture combining convolutional neural networks (CNNs) and bi-directional long short-term memory (BLSTM) to predict travel demand in the Beijing metro system. Their model outperformed traditional and deep-learning-based models in terms of accuracy. For buses, Zhao et al. [19] enhanced short-term bus trip demand prediction using an experiment in Shenzhen, China, by proposing a graph-deep-learning-based technique coupling with the spatiotemporal influence of the built environment (GDLBE). In the case of taxis, Liu et al. [9] demonstrated the superior predictive ability of a context-aware convolutional recurrent neural network (CACRNN) in forecasting taxi demand in New York City, incorporating spatio-temporal variables and outperforming traditional methods like historical average (HA) and ARIMA. Yao et al. [10] also developed a prediction model for taxi demand using a large-scale dataset from Guangzhou, China, and found that their dynamic multivariate spatio-temporal neural network (DMVST-Net) outperformed traditional time series and deep learning models.

Predicting travel demand for ride-hailing services has also demonstrated potential using machine learning approaches. Chen et al. [11] compared the performance of an ARIMA model with three different deep-learning models, finding that a hybrid CNN-LSTM model achieved better accuracy in capturing spatiotemporal patterns of ride-hailing demand. To perform intelligent prediction effectively and economically, Qiao et al. [20] presented ERPM, a three-in-one multi-agent reinforcement-learning-based online algorithm for ride-hailing demand prediction.

In the realm of bike-sharing, Lin et al. [12] proposed a graph convolutional neural network with a data-driven graph filter (GCNN-DDGF) for demand prediction at the station level in New York City, outperforming other models. For the city level, Giot and Cherrier [13] exploited data from individual trips in Washington, D.C., and found that Ridge Regression and Adaboost Regression were the most effective models for city-level demand prediction. In the case of scooter-sharing, Xu et al. [21] proposed a novel deep learning architecture named spatio-temporal multi-graph transformer (STMGT) to forecast the real-time spatiotemporal dockless scooter-sharing demand. The proposed approach was assessed for two real-world case studies, one in Austin, Texas, and the other in Washington, D.C. The findings indicate that STMGT significantly exceeds all of the chosen benchmark models for both case studies.

Despite the growing interest in machine learning for travel demand estimation across various travel modes, there is a notable research gap in predicting travel demand for moped scooter-sharing using machine learning techniques. Existing studies on shared mopeds have focused on accident analysis [22,23], user characteristics and preferences [3,24,25], and the spatiotemporal dynamics of mopeds [26], but not on predicting travel demand using machine learning.

Furthermore, there is ample evidence about the impact of the built environment and land use features on travel demand. For instance, Jun et al. [27] examined the land use features of the pedestrian catchment areas of Seoul subway stations and how they affected ridership at the station level. They discovered that intermodal connectivity, a diverse mix of land uses, and high employment and population densities all had a positive effect on subway ridership. Similarly, Wang et al. [28] explored the spatially varying relationships between the built environment and subway trips in Chengdu, highlighting the significant

impact of factors such as residential, commercial, and transport infrastructure on trip origins and destinations.

The purpose of this study is to fill this gap in understanding the patterns of travel demand patterns concerning moped scooter-sharing services.

3. Study Context

The study was carried out in Madrid, Spain, one of the most populated cities in the European Union with a variety of public and private transportation options. According to the last Madrid Mobility Survey [29], the average number of trips made in the city on a working day is 7.9 million. They are distributed among several means, with walking and bicycling accounting for the majority (34.6%), followed by public transportation (33.4%) and private vehicles/motorcycles (28.6%). Less than 3.4% of all shared modes are made up of taxis, moped scooters, and other vehicles.

Spanish city centers are typically distinguished by old, historic urban structures with narrow streets and ongoing traffic issues, which may favorably affect the use of shared mopeds. Regardless of the cause, the country has one of the largest fleets of shared mopeds in Europe, which results in about 9000 motorcycles [3,30].

In Spain, companies rent out basic electric mopeds with swappable batteries that do not require moving the vehicle. These moped scooter-sharing services can accommodate one or two people and run on a free-floating system. Through a mobile application, their users can book the service and find out where the nearest accessible mopeds are located. In Spain, anyone with a motorbike or moped driver's license or one who has owned a car license for longer than three years is permitted to drive this type of vehicle [25].

The city of Madrid has several moped scooter-sharing operators (such as Acciona, eCooltra, Movo, and Cabify). Acciona plays an important role as a shared mobility operator in Europe [26], with an important presence in Spain. Moreover, there are no restrictions on these services in Madrid because local agencies and operators work closely together to regulate them [31]. Acciona contributed to the dataset used for this study, which contains all travel to Madrid in 2022, as described in the data description (see Section 4).

Acciona launched its moped scooter-sharing business in 2018 and intends to accelerate the adoption and use of light electric vehicles to decarbonize urban centers, making it a relevant player in urban mobility, especially when the average distance traveled is less than 10 km per trip [32]. In Spain, Acciona has the largest market share, reaching 36.4% in 2022, surpassing its main competitor, eCooltra, which has a market share of over 30% [33]. The Acciona moped scooter-sharing service allows speeds of up to 50 km/h for city driving and offers flexible pricing options in Madrid, with EUR 0.33/min being the standard rate [34]. This competitive pricing makes it an appealing choice for users seeking efficient and convenient transportation in Madrid, making moped scooter-sharing a cost-effective alternative for short-distance trips, especially within urban areas.

Electric vehicles, including shared modes, are allowed to park anywhere in a city center in Spain and are typically granted free on-street parking. This makes them a desirable alternative for commuting within inner city districts, such as Madrid's M-30 highway, which is part of the first ring road [3]. It should be noted that in Madrid, a policy to promote sustainable mobility was implemented in 2003, called the Regulated Parking Service (SER by its Spanish acronym), which is explained in more detail in the data description (see Section 4). To rationalize and balance the use of public space and the parking of private automobiles, this SER policy is in charge of administering, regulating, and controlling vehicle parking on public roadways in the city's major areas [35]. Since Acciona (the dataset provider) only operates electric mopeds, its vehicles are immune from these regulations and enjoy free on-street parking throughout the city.

The introduction of low-emission zones (LEZs) is another significant legislative move in Madrid intended to promote sustainable mobility. These LEZs, which are further described in the data description (see Section 4), encourage the use of electric or low-emission vehicles by preventing high-pollution vehicles from entering areas of the city center.

4. Data Description

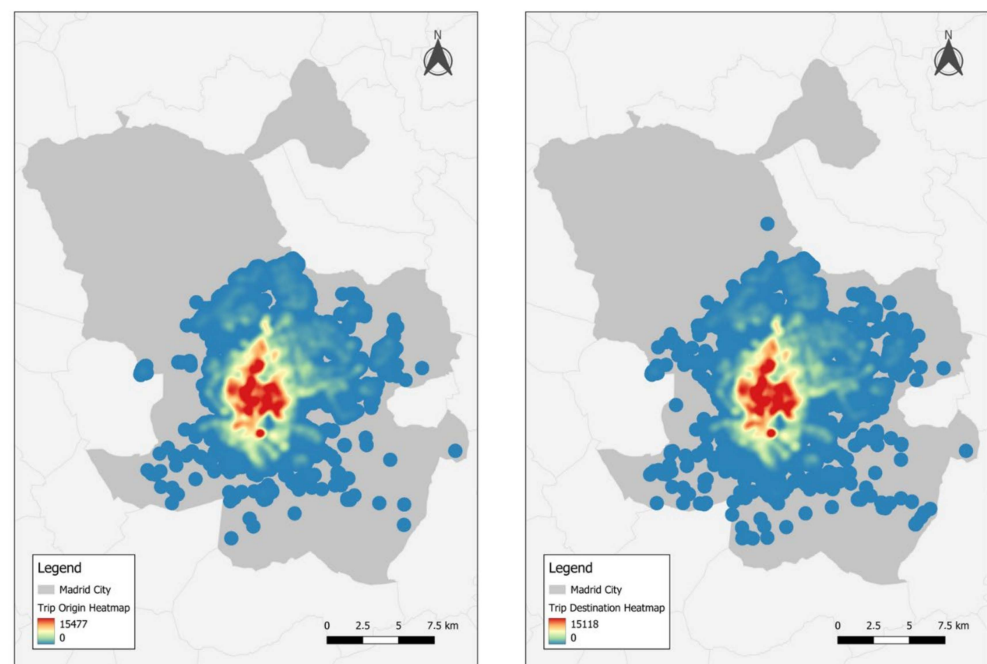
The dataset used in this study includes information gathered by GPS units mounted on Acciona moped scooters. Acciona's data collection processes adhere to industry standards, ensuring high reliability and accuracy. Every trip made in Madrid in 2022 was included in the CVS format of data, which also included the following details:

- Id_trip: an identification number for every trip;
- Id_user: an identification number for every user;
- Id_vehicle: a unique number assigned to every motorcycle;
- Start_time: the timestamp of the trip's beginning (yy-mm-dd hh:mm:ss);
- End_time: the timestamp of the trip's conclusion (yy-mm-dd hh:mm:ss);
- Start_latitude and longitude: the location in xy coordinates where the motorcycle was picked up;
- End_latitude and longitude: the location in xy coordinates where the motorcycle was dropped.

The dataset provided by Acciona included all trips in the entire metropolitan area of Madrid in 2022, with a total of 1,532,088 entries. Based on the information about the geographic coordinates, it was noted that 1,453,433 trips have their origin and destination within the city of Madrid (94.87% of the original dataset). This last dataset was used in this paper to estimate travel demand among origin and destination pairs in each neighborhood of the city, which is the dependent variable of this study.

Using feature engineering approaches, further exogenous factors were included, such as journey time and distance, origin and destination district, and origin and destination neighborhood. The city of Madrid has a total of 21 districts and 131 neighborhoods [36]. Data cleaning was performed in cases in which the travel time variable had a value less than or equal to zero. Following the data cleaning procedure, 1,453,395 entries (with 99.99% representativeness) were found in the final dataset.

The final dataset includes moped scooter-sharing trip data in Madrid, Spain, classified as inter-neighborhood (95%) and intra-neighborhood (5%) trips. On average, the distance covered is 2.89 km, with an average travel time of 13.52 min. The statistics underscore the dominance of short trips. Figure 1 illustrates the origin and destination trip heatmaps on an annual basis, with the intensity of travel represented by the heatmap color gradient.



(a). Trip origin heatmap

(b). Trip destination heatmap

Figure 1. Origin and destination trip heatmaps on an annual basis.

The heatmaps of the origin and destination trips are almost identical, verifying that the highest concentration of origin and destination points (highlighted in red) of the trips occurs in the central area of Madrid, inside the M-30 highway (see Figure 2a). There are high-demand areas in this zone, especially at the Atocha and Chamartín multimodal stations. These stations serve as major hubs for various modes of transportation, including buses, metros, and trains, making them significant nodes in the city's transport network.

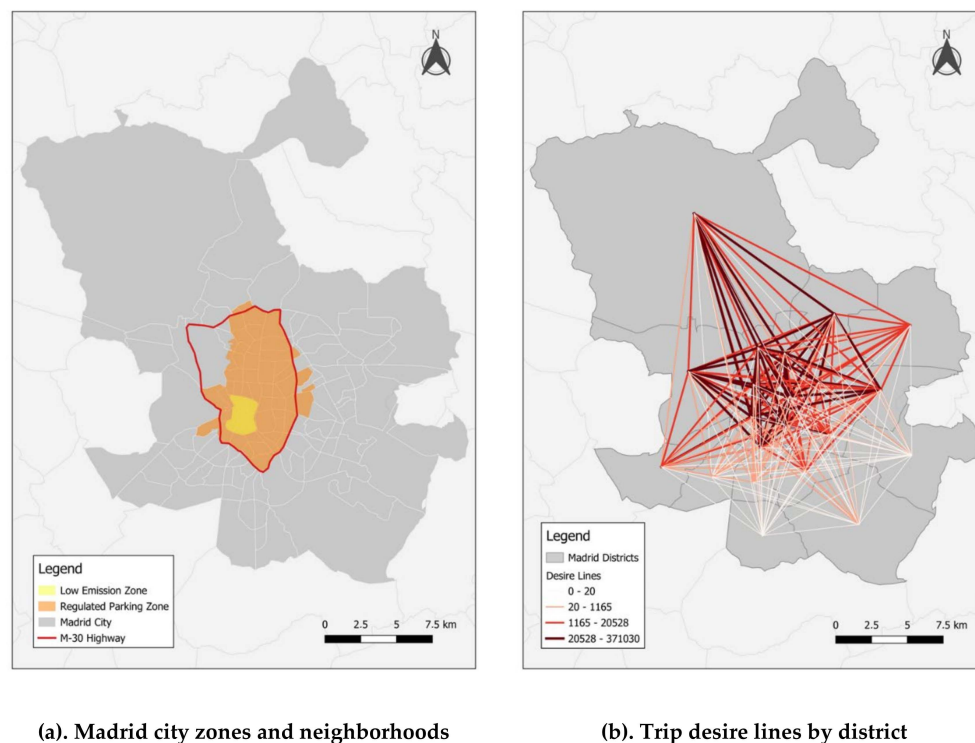


Figure 2. Madrid city zones and desire lines.

In the next step, trip data by district and neighborhood of Madrid were aggregated to calculate the total number of trips to and from each location, as illustrated by the trip desire lines by district (see Figure 2b). The areas of Madrid with the greatest travel demand are found in the downtown region (along the M-30 highway) and the regulated parking zone, particularly in the higher-income neighborhoods (see Appendix A).

The districts with the highest travel demand are in the central area of Madrid (within the M-30 highway) and within the regulated parking zone, especially in the higher-income areas (see Appendix A).

Moving forward, the analysis focuses on aggregating travel data per neighborhood across the 131 neighborhoods in Madrid. The analysis focuses on the yearly number of trips for every pair of origin and destination, and in 2022, the city recorded almost 1.5 million moped scooter-sharing trips. Figure 3 presents a histogram and a kernel density estimate (KDE) of travel demand between origin–destination neighborhoods throughout the analysis period. KDE is a non-parametric method utilized to estimate the probability density function of a random variable, providing a smoothed representation of the distribution of travel demand between neighborhoods. As can be observed, the histogram and KDE of the travel demand between origin–destination neighborhoods in the period have skewed distributions and contain many zero values.

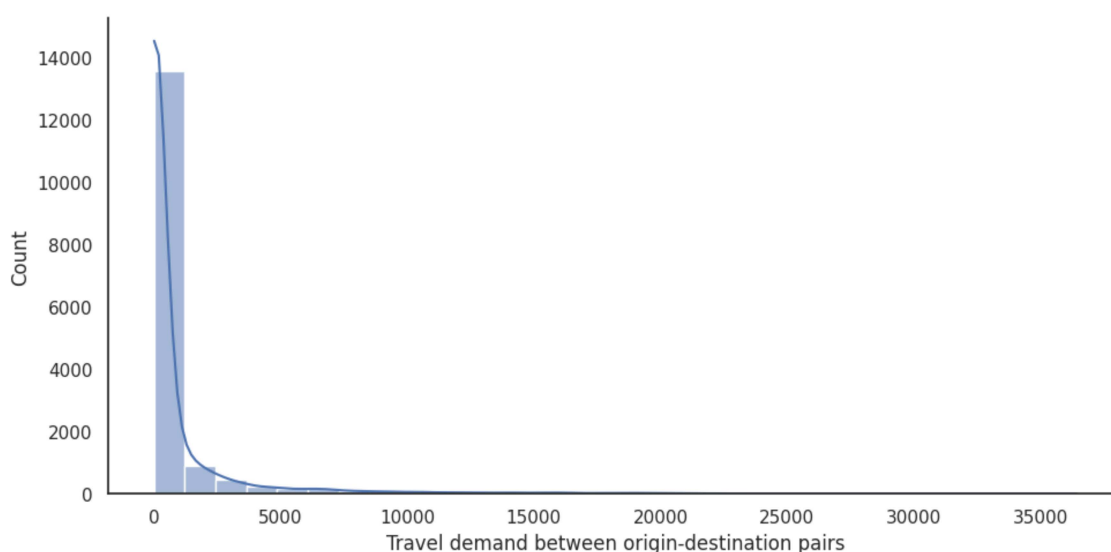


Figure 3. Travel demand between origin–destination neighborhoods in the period. The blue line represents KDE.

Additional explanatory variables were combined with travel demand data, including (i) sociodemographic characteristics of the neighborhood, (ii) travel attraction centers, (iii) transportation network attributes, and (iv) policy-related variables, such as the application of low-emission zones and regulated on-street parking areas. The values of all these attributes, which will be explained in greater detail below, as well as the travel demand data, are referred to in the year 2022.

The dataset with the sociodemographic characteristics of the neighborhood was collected on the website of the Open Data Portal of the City of Madrid through the panel of indicators of districts and neighborhoods in Madrid [36]. Data were provided in XLSX format, from the year 2022, and the following neighborhood indicators were used:

- Pop_Density: population density (inhabitant/Ha.);
- Pop_Avg_Age: population average age;
- Income: average annual net income of households;
- Unemployment: number of unemployed people (registered in February 2022).

Data on travel attraction centers were collected from the website of the Spanish Institute of Statistics [37]. The data collected only include the specific locations of these travel attraction centers, rather than their sizes or areas. The data were provided in SHP format, and a GIS tool was used to extract the magnitude of each variable by neighborhood in Madrid, namely:

- Public_Admin: points of interest of public administration, including (i) administration of justice; (ii) tax agency; (iii) city councils and ministries; (iv) embassies and consulates; (v) employment workshops; and (vi) social security;
- Commercial_Centers: commercial points of interest, including (i) shopping centers; (ii) food galleries; (iii) large, specialized areas; (iv) hypermarkets and markets; (v) supply markets; and (vi) banks;
- Education_Centers: education points of interest, including (i) university campuses; (ii) non-university educational centers—private and public centers; and (iii) non-university educational centers—educational services;
- Health_Centers: health points of interest, including (i) comprehensive care centers for drug addicts; (ii) specialty centers; (iii) health centers; (iv) health consultation; (v) district mental health service; and (vi) pharmacies;
- Cultural_Services: points of interest for cultural, recreational, and personal services, including (i) theaters; (ii) cinemas; (iii) museums; (iv) theme parks; (v) parishes;

(vi) bullrings; (vii) historic bridges; (viii) cemeteries; (ix) libraries; and (x) parks and gardens;

- Tourist_Accom: tourist points of interest, including (i) hotels (ii) hostels; (iii) rural accommodation; (iv) tourist apartments; (v) camping; and (vi) other tourist offices.

Data on transportation network attributes were collected from the Open Data website of the Madrid Regional Transport Authority [38]. The data were provided in SHP format, and a GIS tool was used to extract the magnitude of each variable at the neighborhood level in Madrid, namely:

- Bus_Stop: Madrid Urban Bus Network (EMT by its Spanish acronym) bus stops;
- Metro_Station: Madrid Metro Network metro stations.

Data on policy-related variables were collected from the Madrid City Hall Geoportal website [39]. Data were provided in SHP format (see Figure 2a), namely:

- Low_Emission_Zone: as part of its Madrid 360 Environmental Sustainability Strategy, the local government established a low-emission zone (LEZ) in the “Central District” to improve air quality and lessen the harmful effects of motorized traffic, thereby protecting human health and the urban environment. For a detailed description of this low-emission zone, the reader is referred to reports such as Geoportal del Ayuntamiento de Madrid [39] and Gonzalez et al. [35];
- Regulated_Parking: paid map service encompassing designated on-street parking areas, subdivided into neighborhoods, and differentiated zones.

In addition, a crucial impedance variable was introduced to the dataset. The distance between the centroids of the origin and destination pairs of neighborhoods in Madrid is represented by this variable. Since it sheds light on the spatial interactions across neighborhoods, the inclusion of this impedance component is crucial. The distance matrix, containing the distances between the centroids of all neighborhoods in Madrid, was obtained using a GIS tool. Employing the centroid-to-centroid distance metric provides a pragmatic approximation for capturing travel distances within the urban context, ensuring consideration of spatial accessibility between neighborhoods while minimizing computational complexity.

Table 1 shows the descriptive statistics of the additional explanatory variables considered in this research, concerning the 131 neighborhoods of Madrid. The appendix illustrates the geographical distribution of the sociodemographic characteristics, travel attraction centers, and transportation network attributes across neighborhoods.

Table 1. Summary statistics of explanatory variables, concerning the 131 neighborhoods of Madrid.

Attribute	Variable (Units)	Typology	Summary Statistics	
Sociodemographic characteristics of the neighborhood	Pop_Density (inhabitant/Ha.)	Continuous	Mean	187.26
			Max.	462.90
			Min.	0.19
			SD	123.65
	Pop_Avg_Age (years)	Continuous	Mean	44.05
			Max.	49.77
			Min.	32.57
			SD	3.46
	Income (EUR)	Continuous	Mean	44,275.59
			Max.	91,933.00
			Min.	21,070.67
			SD	17,192.94
	Unemployment (No. of unemployed people)	Continuous	Mean	1195.79
			Max.	3579.00
			Min.	35.00
			SD	852.52

Table 1. Cont.

Attribute	Variable (Units)	Typology	Summary Statistics	
Travel attraction centers	Public_Admin (No. of public administrations or offices)	Continuous	Mean	3.11
			Max.	36.00
			Min.	0.00
			SD	6.56
	Commercial_Centers (No. of commercial centers or establishments)	Continuous	Mean	10.55
			Max.	51.00
			Min.	0.00
			SD	8.53
	Education_Centers (No. of education centers or schools)	Continuous	Mean	15.50
			Max.	168.00
			Min.	0.00
			SD	16.62
	Health_Centers (No. of healthcare facilities)	Continuous	Mean	15.47
			Max.	41.00
			Min.	1.00
			SD	8.66
	Cultural_Services (No. of cultural facilities or services)	Continuous	Mean	9.41
			Max.	55.00
			Min.	0.00
			SD	9.43
	Tourist_Accom (No. of tourist accommodations)	Continuous	Mean	95.02
			Max.	1752.00
			Min.	0.00
			SD	268.84
Transportation network attributes	Bus_Stop (No. of bus stops)	Continuous	Mean	35.99
			Max.	134.00
			Min.	2.00
			SD	21.96
	Metro_Station (No. of metro stations)	Continuous	Mean	1.82
			Max.	8.00
			Min.	0.00
			SD	1.79
Policy-related variables	Low_Emission_Zone (Binary)	Categorical	1 (Yes)	6
			0 (No)	125
	Regulated_Parking (Binary)	Categorical	1 (Yes)	55
			0 (No)	76
Impedance variable	Distance * (km)	Continuous	Mean	7.38
			Max.	26.24
			Min.	0.44
			SD	3.99

* Distance between the centroids of the origin and destination pairs of Madrid neighborhoods.

5. Methodology

This section presents the methodology employed in this study, outlined in four main steps. The flowchart in Figure 4 provides an overview of the methodology.

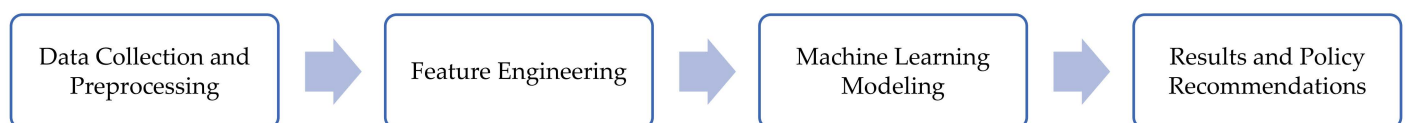


Figure 4. Methodological framework.

The first step of data collection and preprocessing involved gathering a comprehensive dataset of moped scooter-sharing trips in Madrid, Spain, and preparing the data for analysis. The second step of feature engineering involved extracting and constructing relevant features from the dataset. These first two steps have already been presented in the data description (see Section 4).

The third step of machine learning modeling shows the methods used to investigate travel demand patterns for moped scooter-sharing services. Machine learning techniques were employed to develop predictive models to estimate travel demand among origin and destination pairs in each neighborhood of Madrid.

To predict origin–destination, two supervised machine learning models were used: random forest and linear regression. Both models have strengths and are widely applicable for regression purposes, that is, to predict a continuous-value attribute associated with an object.

Ordinary Least Squares Linear Regression, or simply linear regression, fits a linear model with coefficients $w = (w_1, \dots, w_p)$ to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation. It is a widely used, straightforward statistical model. Additionally, it is among the simplest machine learning algorithms [40,41].

A Random Forest regressor, or simply random forest, is a meta-estimator that fits many classifying decision trees on various sub-samples of the dataset and uses averaging to improve predictive accuracy and control over-fitting [42]. Its key benefits include robustness, minimum hyper-parameter adjustment, short training time, and the ability to reduce overfitting without increasing bias-related error. It is one of the most accurate general-purpose machine learning techniques [43].

The models were based on a combination of travel demand data and additional explanatory variables (see Table 1). These various input features were standardized and used to train a model with travel demand. All models were compared using random data split into training and testing sets (adopted Train/Test equal to 80/20).

In terms of performance metrics, the models' predictive power can be validated. As can be seen in Figure 3, the study's target variable has a skewed distribution and contains zero values, making the Root Mean Squared Error (RMSE) and Mean Squared Error (MSE) two metrics that are frequently used to assess the accuracy of prediction models [44].

Firstly, the Root Mean Square Error (RMSE) quantifies the discrepancy between actual and expected values. The standard deviation of the prediction errors is used to compute the difference, or residuals (see Equation (1)). The dependent unit used in this metric is the same. In this study, the number of annual trips between origin and destination pairs in each Madrid neighborhood serves as the dependent unit.

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(Actual_i - Predicted_i)^2}{N}} \quad (1)$$

Secondly, the average of the squares of the errors is measured by the Mean Squared Error (MSE). Equation (2) shows that the MSE is expressed in squared units of the dependent unit.

$$MSE = \frac{1}{N} \sum_{i=1}^N (Actual_i - Predicted_i)^2 \quad (2)$$

Notice that the formulas are nearly identical. The RMSE is just the square root of the MSE. When assessing how well a model fits a dataset, the RMSE is more often used because it is measured in the same units as the response variable.

It is important to note that random forest is a tree-based ensemble machine learning method that does not have a direct notion of coefficient signs (positive or negative) of the relationship between the features and the dependent variable, unlike the linear regression model, in which the coefficients indicate the direction and strength of the relationship.

However, random forest models calculate feature importance based on how much a feature reduces the impurity of the nodes it splits in the decision trees. The feature

importance value captures the relative importance of features by measuring how much the performance of the model decreases when that feature is not available in the data.

Results and policy recommendations derived from machine learning models are discussed in the following sections (see Sections 6–8). These results include model performance, feature importance scores, and implications for urban planning and transportation policy.

6. Results

This section shows the use of machine learning approaches in the development of predictive models able to estimate travel demand patterns among origin and destination pairs in each area of Madrid for moped scooter-sharing services.

Travel demand patterns were predicted using two supervised machine learning models: random forest and linear regression. The models were based on a combination of travel demand data and additional explanatory variables (see Section 4). These variables were assigned to each origin and destination pair using the suffixes ‘_Origin’ and ‘_Destination’. Additionally, the dataset includes a distance impedance variable, representing the distance between the centroids of the origin and destination pairs of Madrid neighborhoods. Unlike the other variables, the distance impedance is not assigned using suffixes; rather, it captures the spatial relationship between neighborhoods and provides essential geographic information for the predictive capabilities of the models.

These various input features have been standardized to ensure they are on a similar scale. This can prevent certain features from dominating the model and improve its performance. The two machine learning models’ default hyperparameters were applied in this investigation. Subsequently, they were employed to train a model where the target variable was travel demand. Every additional variable shown in Table 1 was employed as a feature (i.e., the sociodemographic characteristics of the neighborhood, travel attraction centers, transportation network attributes, policy-related variables, and the distance impedance variable).

Notably, for comparison, all models employed the same random data split of the training and testing sets (i.e., Train/Test equal to 80/20). The number of trips made annually between origin and destination pairs in each Madrid neighborhood is taken into consideration when evaluating and comparing the predictive accuracy of each model (see Table 2).

Table 2. Comparison of the models’ indicators of performance.

Model	Performance Metrics	
	RMSE	MSE
Linear Regression	0.060	0.004
Random Forest	0.026	0.001

While the linear regression model provides valuable insights between independent variables and the dependent variable, its predictive accuracy falls short compared to the random forest model. The higher RMSE and MSE values associated with the linear regression model indicate greater disparity between predicted and actual values. This discrepancy underscores the limitations of the model in capturing the complex and non-linear relationships inherent in the dataset.

On the other hand, the random forest model outperformed the linear regression in terms of forecasting trip demand. The random forest model captured the variability of the target variable with a higher degree of precision due to its lower RMSE and MSE values. By considering interactions and non-linear relationships among the features, the random forest model outperformed the linear regression model.

As mentioned before, the linear regression model was used to visualize feature coefficients, and the random forest model was used to visualize feature importance scores. Figure 5 illustrates the feature coefficients of the linear regression model. The feature coefficients represent the relative importance of each feature in predicting travel demand. Positive coefficients (in blue color) indicate that an increase in the corresponding feature

value is associated with an increase in travel demand, while negative coefficients (in red color) suggest the opposite effect.

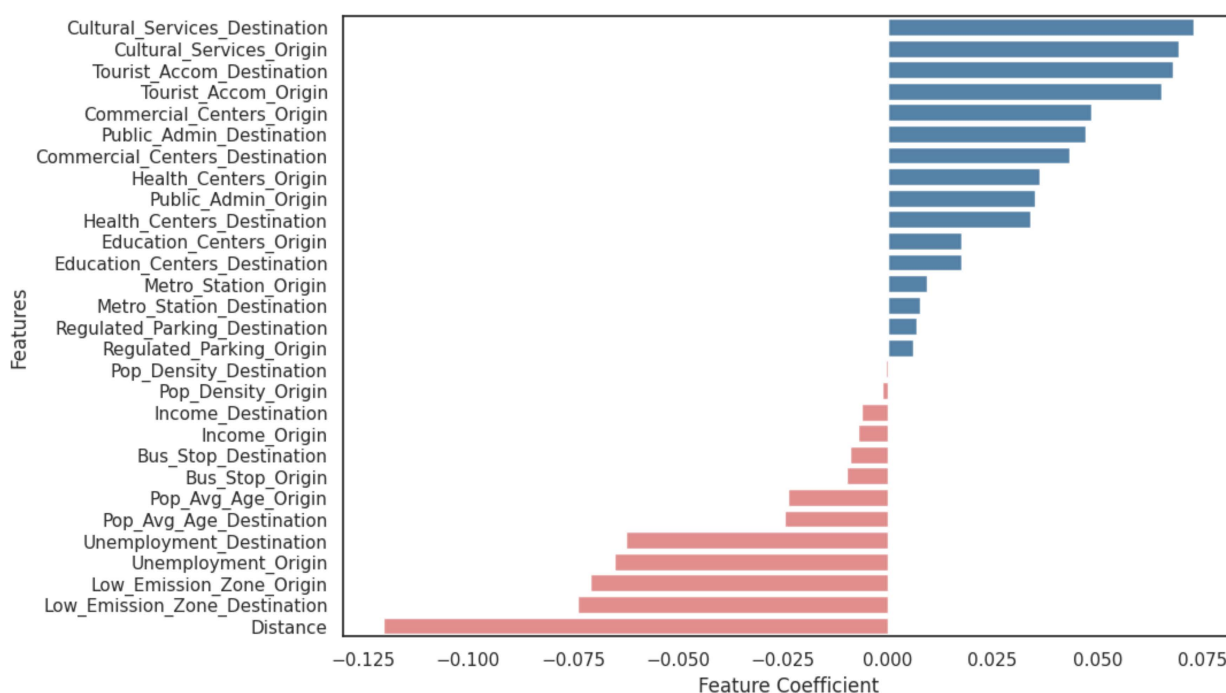


Figure 5. Feature coefficients of the linear regression model. Blue has a positive effect on motorcycle travel demand, and red has a negative effect.

Relevant features in the linear regression model include the presence of cultural services, tourist accommodations, commercial centers, and public administration centers in both the origin and destination neighborhoods. These features play crucial roles in attracting travel demand because they represent key points of interest and activities in the neighborhoods. Cultural services, such as museums and theaters, attract both residents and tourists, leading to increased travel demand. Tourist accommodations and commercial centers contribute to higher mobility as people move to and from these locations for leisure, shopping, and other activities. The importance of commercial areas in explaining the demand for moped scooter sharing is consistent with the analysis of spatiotemporal travel patterns of micromobility services (including shared mopeds) conducted by Arias-Molinares [45] in Madrid.

Similarly, public administration centers serve as hubs for administrative activities and services, attracting a significant number of people to the area. Likewise, the presence of health centers also contributes to increasing travel demand. This strong relationship between the demand for moped scooter sharing and the location of health centers has also been concluded by Arias-Molinares et al. [46].

As for transport-related variables, the number of metro stations also increases the volume of moped-sharing demand. Proximity to metro stations enhances accessibility and connectivity, encouraging people to travel more frequently. Moped scooter-sharing users may take advantage of that service to connect with metro stations for long-distance trips.

On the other hand, the results for the variables related to the presence of low-emission zones, unemployment, and the average age of the population indicate that these factors may have a limiting effect on the volume of travel demand. The presence of low-emission zones can have a negative effect on motorcycle travel demand, which can be attributed to the limited availability of spaces for parking in the area and the high supply of public transportation services. Higher unemployment rates can lead to reduced disposable income, which can prevent individuals from using motorcycles, which are much more expensive than public transportation. Likewise, areas with older populations are expected

to have a lower demand for motorcycles, as older individuals would rather choose other transportation options due to comfort or mobility restrictions.

The distance between the centroids of the origin and destination pairs of Madrid neighborhoods emerged as a significant negative factor, indicating that reasonably longer travel distances tend to decrease travel demand for moped scooter-sharing. This suggests that users prefer these services for short-distance trips within the city or to connect with other public transportation means, making them more appealing for urban mobility.

Figure 6 illustrates the feature importance scores of the random forest model. These scores indicate the relative importance of each feature in predicting travel demand. Higher scores show a stronger influence of the corresponding feature on the forecast of travel demand.

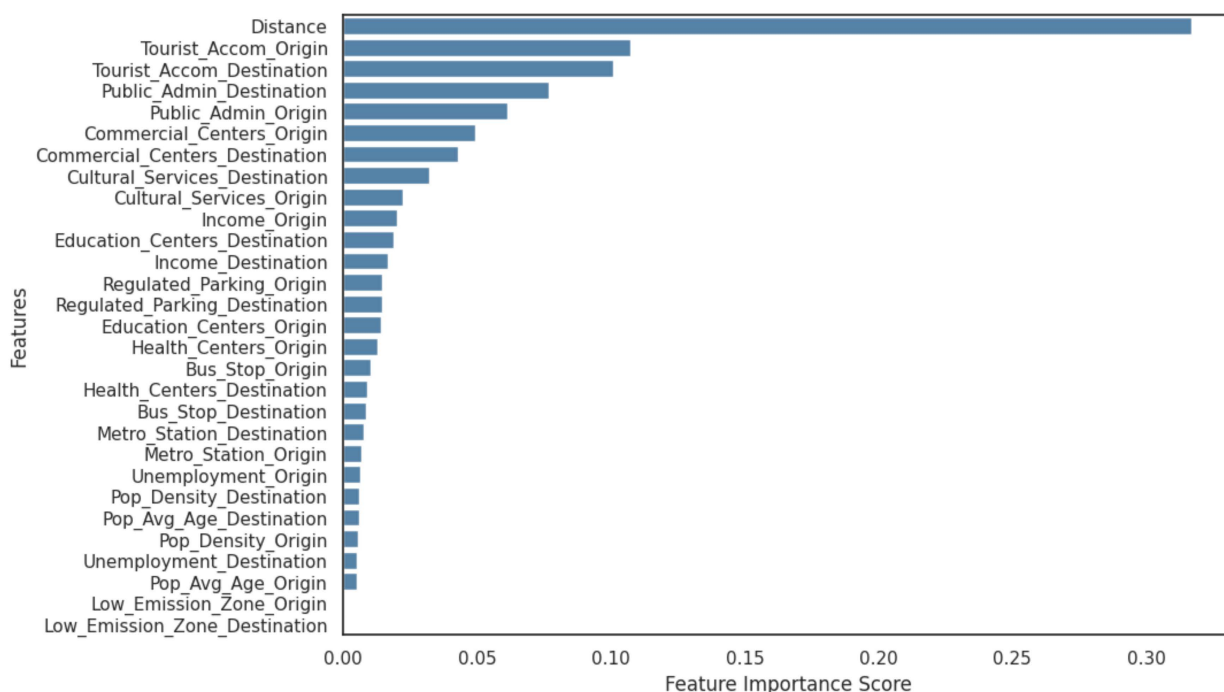


Figure 6. Feature importance scores of the random forest model.

The most influential feature in predicting travel demand is the distance between the centroids of the origin and destination pairs of Madrid neighborhoods, which accounts for approximately 31.7% of the model's prediction accuracy. This highlights the critical role of travel distance in shaping travel demand, also for moped scooter-sharing services. Users seem to prefer these services for shorter trips within the city or to connect with other public transportation means. Additionally, features related to the presence of tourist accommodations, public administration centers, commercial centers, and regulated parking also have notable importance in influencing travel demand.

On the other hand, features related to low-emission zones have negligible importance in the forecast of travel demand. This suggests that moped scooter-sharing services are not primarily being utilized as an alternative transport mode to access low-emission zones, potentially due to existing restrictions for car use or other preferences in those areas.

7. Discussion

The analysis of travel demand prediction for moped scooter-sharing services using both linear regression and random forest models provided valuable insights into the factors influencing travel demand in Madrid's neighborhoods.

The influence of sociodemographic characteristics on travel demand is a complex interplay of various factors. Variables such as the average age of the population and the number of unemployed individuals show limited impact on travel demand, primarily

due to their concentration in neighborhoods far away from the downtown area of the city (see Appendix A). The results also suggest that income has varying degrees of influence on travel patterns. Although this variable is less significant in the linear regression model, it gains more importance in the random forest model. This result highlights the nuanced relationship between sociodemographic factors and travel behavior, emphasizing the need to consider spatial variations and urban dynamics when formulating effective transport policies and interventions.

Travel attraction centers play a crucial role in shaping travel demand, influencing both the origin and destination neighborhoods in the analyzed models (see Figures 5 and 6). Notably, tourist accommodations, cultural services, public administration centers, and commercial centers emerge as key attributes driving travel patterns. These features are concentrated in the central areas of the city (see Appendix A), underscoring their potential impact on travel behavior and the significance of considering their spatial distribution in urban planning and transportation policies.

The analysis of transportation network attributes reveals distinct patterns for bus stops and metro stations. Bus stops are notably concentrated around the city center (see Appendix A), and while their influence on travel demand is evident, it appears relatively limiting, as suggested by the linear regression model. On the contrary, metro stations also exhibit a concentration in the city center (see Appendix A) and emerge as influential factors in predicting travel demand. These results highlight the multifaceted nature of transportation choices in Madrid. While bus stops impact travel demand, metro stations play a more pronounced role, likely due to the increased accessibility they provide for moped scooter-sharing users. This demonstrates that the proximity and ease of connecting to metro stations can significantly influence travel behavior, especially for longer-distance journeys.

The examination of policy-related variables uncovers distinctive trends concerning regulated parking zones and low-emission zones in Madrid (see Figure 2a). The implementation of on-street parking regulations has played a pivotal role in influencing travel patterns. This policy has successfully contributed to reducing car ownership rates among households situated within the designated areas and promoting the use of greener cars [35,47]. Notably, the absence of parking fees for mopeds within regulated parking zones highlights the importance of effective parking regulations when estimating travel demand in the central neighborhoods of Madrid. However, it is worth noting that the low-emission zone has little impact on travel demand prediction, which can be attributed to the limited availability of spaces for parking in the area and the high supply of public transportation services. These findings shed light on the multifaceted interplay between policy measures and travel behavior, underscoring the need for a comprehensive and integrated approach to urban mobility planning.

All these findings suggest that neighborhood characteristics and amenities are crucial in explaining travel demand. Understanding these relationships can inform urban planners and decision-makers to better accommodate travel patterns and improve the overall transportation system.

8. Conclusions and Policy Recommendations

This study aimed to investigate travel demand for moped scooter-sharing between neighborhoods in Madrid, Spain, using machine learning techniques. Through the analysis of various features and the application of machine learning models, valuable insights were obtained regarding the factors explaining travel demand in the city.

From the point of view of the methodology employed, the random forest model surpassed the linear regression model in predicting travel demand for moped scooter-sharing services, showcasing its effectiveness in capturing the complex relationships among the features. The superior performance of the random forest model emphasizes the importance of considering non-linear relationships and interactions in travel demand analysis.

The insights gained from this research can inform broader discussions on how to promote sustainable and efficient mobility options in cities. Notably, the distance between

the centroids of the origin and destination pairs of Madrid neighborhoods emerged as a significant factor, with longer travel distances negatively impacting moped scooter-sharing demand. This highlights the preference of users for these services on short-distance trips within the city, making them an attractive option for urban mobility. The presence of metro stations remains a crucial explanatory factor, emphasizing the importance of fostering intermodal connectivity between moped scooter-sharing services and public transportation. By encouraging seamless integration, cities can promote multimodal travel behavior, reducing reliance on private vehicles and contributing to a more sustainable urban transport landscape. Furthermore, the identified neighborhood characteristics that increase the demand for shared mopeds, such as the availability of tourist accommodations and commercial centers, can guide urban planners in creating vibrant and accessible neighborhoods that cater to diverse travel demands.

The findings from this study have important implications for urban planning, transportation management, and various stakeholders involved in the development of sustainable and efficient transportation systems. Based on the results, several policy recommendations can be made to different stakeholders (e.g., urban planners and policymakers, transportation authorities, and moped scooter-sharing service providers).

Urban planners and policymakers can (i) consider the identified important features such as tourist accommodations, public administration centers, and commercial centers when designing land use and zoning plans; (ii) encourage mixed land use development to create vibrant and diverse neighborhoods that can cater to different travel demands; and (iii) focus on enhancing accessibility to public services like health and education centers, which can positively impact travel behavior.

Transportation authorities can (i) use the insights from the Random Forest model to prioritize investment in public transportation infrastructure, particularly in areas with high importance scores for features such as metro stations and bus stops, and (ii) improve connectivity and integration between different modes of transportation to provide seamless and convenient travel options for residents.

Moped scooter-sharing service providers can (i) utilize the predictions of travel demand to optimize fleet allocation and ensure adequate availability of vehicles in high-demand neighborhoods and (ii) collaborate with local authorities to identify potential areas for expanding the service based on the identified features and travel patterns.

Overall, this study illustrated the potential of machine learning techniques to predict travel demand and provide valuable insights for urban planning and transportation decision-making. Accurate predictions of travel demand at the neighborhood level can inform infrastructure planning, optimize public transportation services, and support sustainable urban development. The insights gained from the random forest model can guide policymakers and stakeholders in making informed decisions to address transportation challenges and enhance the overall quality of life in Madrid. Future research can expand these findings by exploring additional variables, incorporating new models, considering the varying attraction capacities of different travel points, conducting comparative studies across different locations, and utilizing extensive travel databases where some user characteristics are known to increase the accuracy and applicability of the findings.

Author Contributions: Conceptualization, T.S.-S., J.G. and J.M.V.; methodology, T.S.-S. and T.R.; software, T.S.-S.; validation, T.S.-S.; formal analysis, T.S.-S.; investigation, T.S.-S. and T.R.; resources, J.G. and J.M.V.; data curation, T.S.-S.; writing—original draft, T.S.-S. and T.R.; writing—review and editing, T.S.-S., T.R., J.G. and J.M.V.; supervision, J.M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Science and Innovation of Spain, the State Research Agency, and the European Union, which financed Project PDC2022-133045-I00 (MICROMOV). Tulio Silveira-Santos is also grateful for his research grant (PRE2019-088587) funded by the Spanish Ministry of Science and Innovation and co-financed by the European Social Fund and the State Research Agency.

Institutional Review Board Statement: Not applicable.

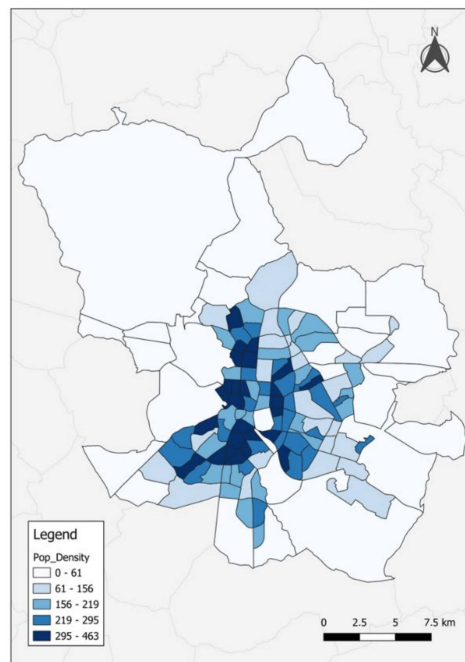
Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

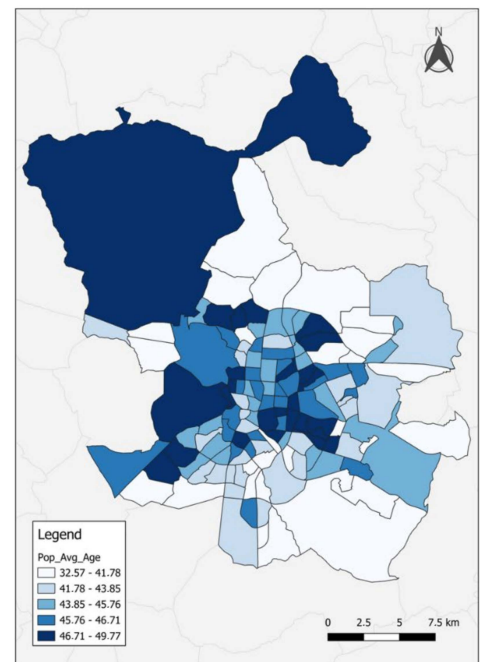
Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

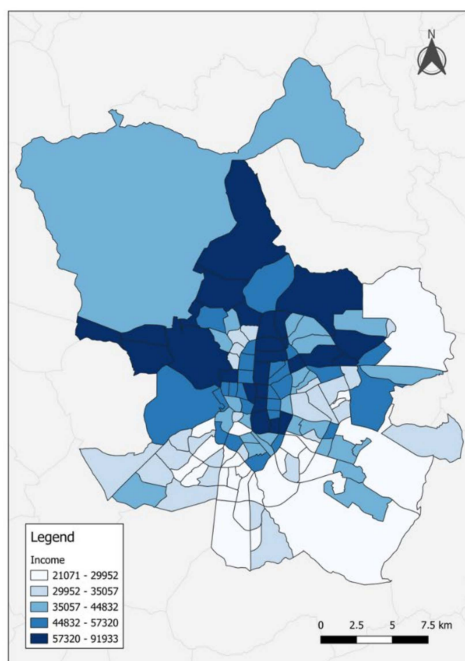
Sociodemographic characteristics of the neighborhood.



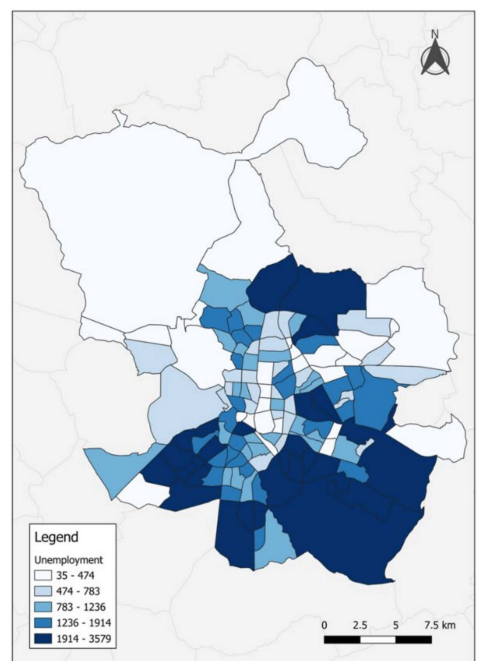
Population density



Population average age

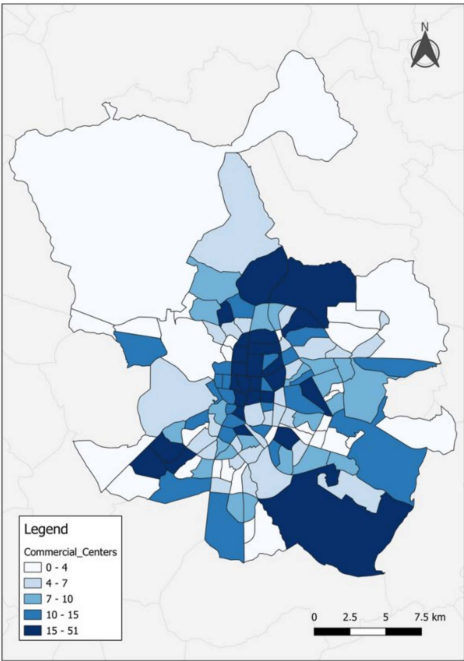
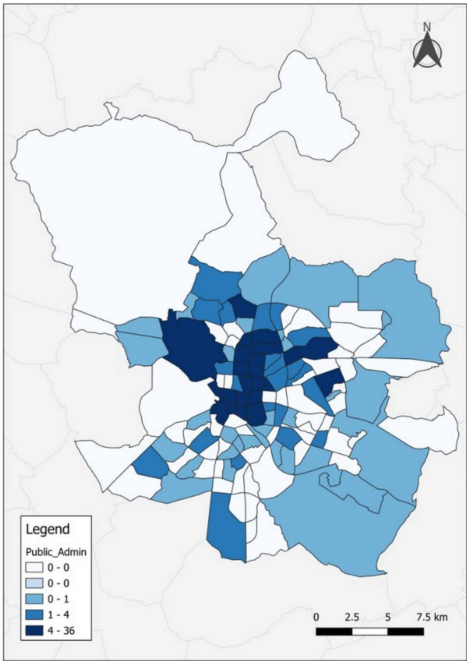


Income



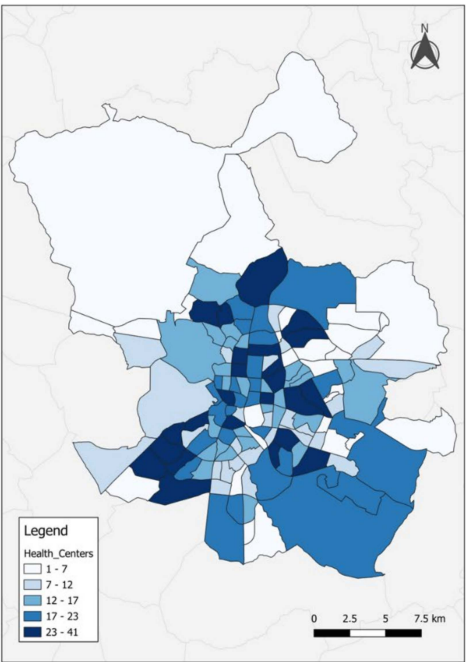
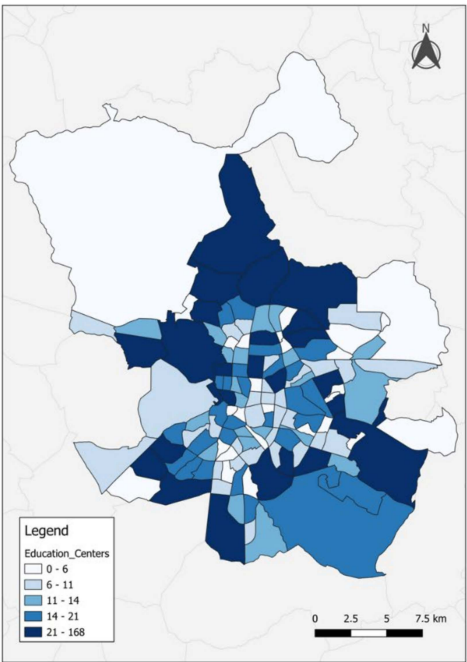
Unemployment

Travel attraction centers



Public administration sectors

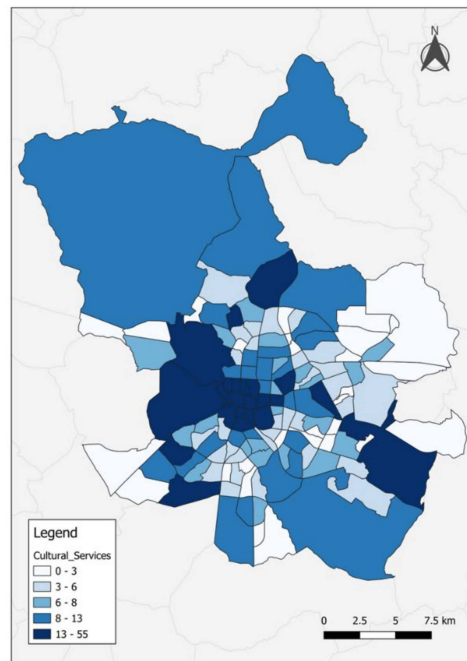
Commercial centers



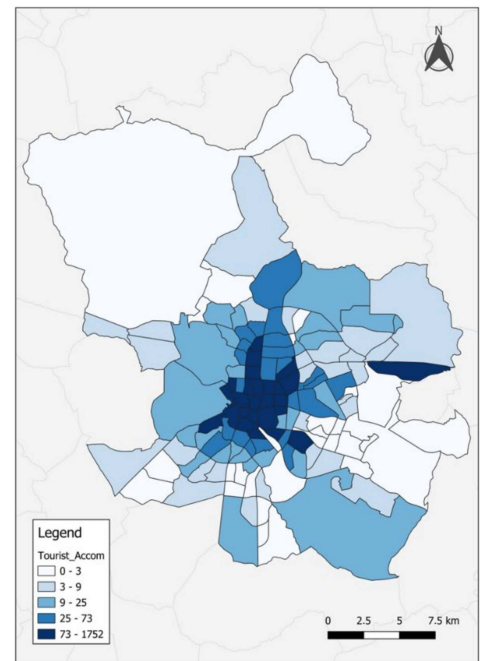
Education centers

Health centers

Travel attraction centers (Cont.).

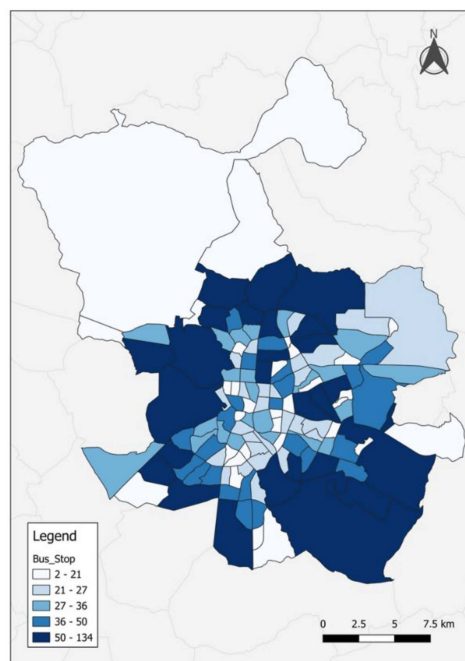


Cultural services

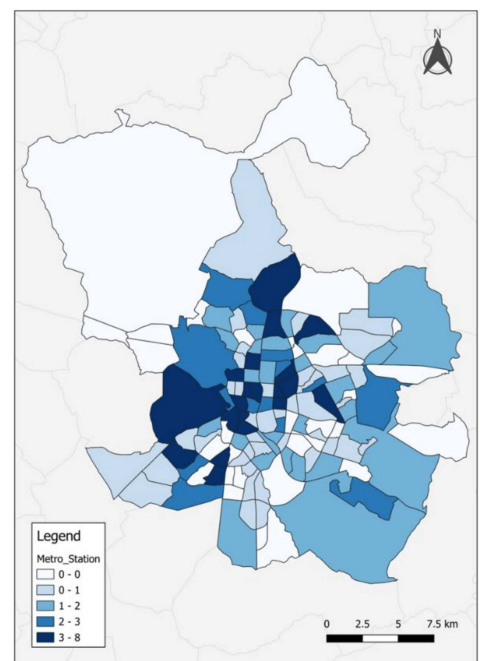


Tourist accommodations

Transportation network attributes.



Bus stops



Metro stations

References

1. Shaheen, S.; Cohen, A.; Zohdy, I. *Shared Mobility: Current Practices and Guiding Principles*; Report Number Fhwa-Hop-16-022; U.S. Department of Transportation. Federal Highway Administration: Washington, DC, USA, 2016; 120p.
2. Shaheen, S.; Cohen, A. Shared Micromobility Policy Toolkit: Docked and Dockless Bike and Scooter Sharing. *UC Berkeley Transp. Sustain. Res. Cent.* **2019**, 1–34.
3. Aguilera-García, Á.; Gomez, J.; Sobrino, N. Exploring the adoption of moped scooter-sharing systems in Spanish urban areas. *Cities* **2020**, *96*, 102424. [[CrossRef](#)]

4. Basu, R.; Ferreira, J. Understanding household vehicle ownership in Singapore through a comparison of econometric and machine learning models. *Transp. Res. Procedia* **2020**, *48*, 1674–1693. [CrossRef]
5. Gong, L.; Kanamori, R.; Yamamoto, T. Data selection in machine learning for identifying trip purposes and travel modes from longitudinal GPS data collection lasting for seasons. *Travel Behav. Soc.* **2018**, *11*, 131–140. [CrossRef]
6. Victoriano, R.; Paez, A.; Carrasco, J.A. Time, space, money, and social interaction: Using machine learning to classify people's mobility strategies through four key dimensions. *Travel Behav. Soc.* **2020**, *20*, 1–11. [CrossRef]
7. Zhao, X.; Yan, X.; Yu, A.; Van Hentenryck, P. Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behav. Soc.* **2020**, *20*, 22–35. [CrossRef]
8. Ma, X.; Zhang, J.; Du, B.; Ding, C.; Sun, L. Parallel Architecture of Convolutional Bi-Directional LSTM Neural Networks for Network-Wide Metro Ridership Prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 2278–2288. [CrossRef]
9. Liu, T.; Wu, W.; Zhu, Y.; Tong, W. Predicting taxi demands via an attention-based convolutional recurrent neural network. *Knowl. Based Syst.* **2020**, *206*, 106294. [CrossRef]
10. Yao, H.; Wu, F.; Ke, J.; Tang, X.; Jia, Y.; Lu, S.; Gong, P.; Ye, J.; Li, Z. Deep multi-view spatial-temporal network for taxi demand prediction. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; pp. 2588–2595.
11. Chen, Z.; Liu, K.; Feng, T. Examine the Prediction Error of Ride-Hailing Travel Demands with Various Ignored Sparse Demand Effects. *J. Adv. Transp.* **2022**, *2022*, 7690309. [CrossRef]
12. Lin, L.; He, Z.; Peeta, S. Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 258–276. [CrossRef]
13. Giot, R.; Cherrier, R. Predicting bikeshare system usage up to one day ahead. In Proceedings of the 2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS), Orlando, FL, USA, 9–12 December 2014; pp. 22–29.
14. Wang, F.; Ross, C.L. Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model. *Transp. Res. Rec.* **2018**, *2672*, 35–45. [CrossRef]
15. Bzdok, D.; Nichols, T.E.; Smith, S.M. Towards algorithmic analytics for large-scale datasets. *Nat. Mach. Intell.* **2019**, *1*, 296–306. [CrossRef]
16. Li, X.; Pan, G.; Wu, Z.; Qi, G.; Li, S.; Zhang, D.; Zhang, W.; Wang, Z. Prediction of urban human mobility using large-scale taxi traces and its applications. *Front. Comput. Sci.* **2012**, *6*, 111–121. [CrossRef]
17. Moreira-Matias, L.; Gama, J.; Ferreira, M.; Mendes-Moreira, J.; Damas, L. Predicting taxi-passenger demand using streaming data. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1393–1402. [CrossRef]
18. Varshavskiy, I.; Stavinova, E.; Chunaev, P. Forecasting railway ticket demand with search query open data. *Procedia Comput. Sci.* **2022**, *212*, 132–141. [CrossRef]
19. Zhao, T.; Huang, Z.; Tu, W.; He, B.; Cao, R.; Cao, J.; Li, M. Coupling graph deep learning and spatial-temporal influence of built environment for short-term bus travel demand prediction. *Comput. Environ. Urban Syst.* **2022**, *94*, 101776. [CrossRef]
20. Qiao, S.; Han, N.; Huang, J.; Peng, Y.; Cai, H.; Qin, X.; Lei, Z. An three-in-one on-demand ride-hailing prediction model based on multi-agent reinforcement learning. *Appl. Soft. Comput.* **2023**, *149*, 110965. [CrossRef]
21. Xu, Y.; Zhao, X.; Zhang, X.; Paliwal, M. Real-Time Forecasting of Dockless Scooter-Sharing Demand: A Spatio-Temporal Multi-Graph Transformer Approach. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 8507–8518. [CrossRef]
22. Blackman, R.A.; Haworth, N.L. Tourist use of mopeds in Queensland. *Tour. Manag.* **2013**, *36*, 580–589. [CrossRef]
23. Haworth, N. Powered two wheelers in a changing world—Challenges and opportunities. *Accid. Anal. Prev.* **2012**, *44*, 12–18. [CrossRef]
24. Degele, J.; Gorr, A.; Haas, K.; Kormann, D.; Krauss, S.; Lipinski, P.; Tenbih, M.; Koppenhoefer, C.; Fauser, J.; Hertweck, D. Identifying E-Scooter Sharing Customer Segments Using Clustering. In Proceedings of the 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Stuttgart, Germany, 17–20 June 2018.
25. Vega-Gonzalo, M.; Aguilera-García, Á.; Gomez, J.; Vassallo, J.M. Analysing individuals' use of moped-sharing and their perception about future private car dependency. *Cities* **2024**, *146*, 104741. [CrossRef]
26. Arias-Molinares, D.; Romanillos, G.; García-Palomares, J.C.; Gutiérrez, J. Exploring the spatio-temporal dynamics of moped-style scooter sharing services in urban areas. *J. Transp. Geogr.* **2021**, *96*, 103193. [CrossRef]
27. Jun, M.J.; Choi, K.; Jeong, J.E.; Kwon, K.H.; Kim, H.J. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *J. Transp. Geogr.* **2015**, *48*, 30–40. [CrossRef]
28. Wang, C.; Wang, X.; Pan, R.; Yan, Y. Influence of Built Environment on Subway Trip Origin and Destination: Insights Based on Mobile Positioning Data. *Transp. Res. Rec.* **2022**, *2676*, 693–710. [CrossRef]
29. Consorcio Regional de Transportes de Madrid. *Encuesta Domiciliaria de Movilidad de la Comunidad de Madrid 2018*; Consorcio Regional de Transportes de Madrid: Madrid, Spain, 2019.
30. Howe, E. *Global Scootersharing Market Report 2018*; InnoZ—Innovation Center for Mobility Societal Change GmbH: Berlin, Germany, 2018.
31. Bach, X.; Miralles-Guasch, C.; Marquet, O. Spatial Inequalities in Access to Micromobility Services: An Analysis of Moped-Style Scooter Sharing Systems in Barcelona. *Sustainability* **2023**, *15*, 2096. [CrossRef]
32. ACCIONA. ACCIONA Company Overview 2022. Available online: <https://mediacd.accion.com/media/qevnqurg/accion-overview-march-2022.pdf> (accessed on 25 July 2023).

33. Movilidad Eléctrica. Carsharing, Motosharing, Taxis, VTC y Transporte Público, Estas son las Aplicaciones de Movilidad que más Utilizan los Españoles. 2022. Available online: <https://movilidadelectrica.com/observatorio-movilidad-urbana-78458-2/> (accessed on 25 July 2023).
34. ACCIONA Movilidad. Moverse por Madrid Nunca fue tan Cómodo. 2023. Available online: https://movilidad.acciona.com/es_ES/madrid/ (accessed on 25 July 2023).
35. Gonzalez, J.N.; Perez-Doval, J.; Gomez, J.; Vassallo, J.M. What impact do private vehicle restrictions in urban areas have on car ownership? Empirical evidence from the city of Madrid. *Cities* **2021**, *116*, 103301. [CrossRef]
36. Portal de Datos Abiertos del Ayuntamiento de Madrid. Panel de Indicadores de Distritos y Barrios de Madrid. Estudio Sociodemográfico. 2023. Available online: <https://datos.madrid.es/sites/v/index.jsp?vgnextoid=71359583a773a510VgnVCM2000001f4a900aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD> (accessed on 11 July 2023).
37. Instituto de Estadística. Puntos de Interés. Nomecalles. 2023. Available online: <https://gestiona.comunidad.madrid/nomecalles/DescargaBDTCorte.icm> (accessed on 11 July 2023).
38. Datos Abiertos del Consorcio Regional de Transportes de Madrid. Portal de Datos Abiertos del Consorcio Regional de Transportes de Madrid. 2023. Available online: <https://data-crtm.opendata.arcgis.com/> (accessed on 11 July 2023).
39. Geoportal del Ayuntamiento de Madrid. Geoportal. 2023. Available online: https://geoportal.madrid.es/IDEAM_WBGEOPORTAL/index.iam (accessed on 11 July 2023).
40. Maulud, D.; Abdulazeez, A.M. A Review on Linear Regression Comprehensive in Machine Learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147. [CrossRef]
41. Wu, J.; Liu, C.; Cui, W.; Zhang, Y. Personalized Collaborative Filtering Recommendation Algorithm based on Linear Regression. In Proceedings of the 2019 IEEE International Conference on Power Data Science (ICPDS), Taizhou, China, 22–24 November 2019; pp. 139–142.
42. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
43. Yan, X.; Liu, X.; Zhao, X. Using machine learning for direct demand modeling of ridesourcing services in Chicago. *J. Transp. Geogr.* **2020**, *83*, 102661. [CrossRef]
44. Washington, S.P.; Karlaftis, M.G.; Mannering, F.L. *Statistical and Econometric Methods for Transportation Data Analysis*, 2nd ed.; Taylor & Francis Group: Abingdon, UK, 2011; 530p.
45. Arias-Molinares, D.; García-Palomares, J.C.; Gutiérrez, J. Micromobility services before and after a global pandemic: Impact on spatio-temporal travel patterns. *Int. J. Sustain. Transp.* **2022**, *17*, 1058–1073. [CrossRef]
46. Arias-Molinares, D.; Xu, Y.; Büttner, B.; Duran-Rodas, D. Exploring key spatial determinants for mobility hub placement based on micromobility ridership. *J. Transp. Geogr.* **2023**, *110*, 103621. [CrossRef]
47. Gonzalez, J.N.; Gomez, J.; Vassallo, J.M. Are low emission zones and on-street parking management effective in reducing parking demand for most polluting vehicles and promoting greener ones? *Transp. Res. Part A Policy Pract.* **2023**, *176*, 103813. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.