Unlocking Shared E-Scooter Potential: Insights from Austin's Micro-Mobility Landscape

Jiaqi Sun

April 2024

Abstract

The usage and profitability of shared E-scooters is one of the issues that scooter companies and investors are concerned about. We explore factors that influence E-scooter usage based on the trip data of Austin, Texas residents from March 2021 to February 2022. We analyzed the usage patterns of E-scooters at multiple time scales and used K-means algorithm to analyze the usage characteristics in different regions. We also fitted a negative binomial regression model to explore the relationship between E-scooter usage and environmental and social factors in different clustered regions. We found that people use E-scooters more for recreational activities after work and on weekends. The most intensive area for E-scooter use is downtown Austin and surrounding residential areas. Students and young people are the main users. Currently, the number of rides per month is not enough to recover the cost of the equipment. Based on the research results, we provided some suggestions that may increase the number of use, such as developing usage methods that do not rely on mobile apps and manually moving E-scooters to areas with dense trip starting points and schools. Our research may provide a basis for the location and number of E-scooters to be deployed in cities in the future, creating more profit and reducing the use of private transportation.

Keywords: E-scooter; Usage Pattern; K-Means; Negative Binomial Regression;

1 Introduction

Shareable electric scooters (E-scooters from now on) were introduced to the U.S in July 2017 in Santa Monica, California by a company called Bird. E-scooters were rapidly expanding around the world in 2018, such as some

European cities, Paris, Berlin and Madrid. E-scooter service claims to solve the problem of "the first/last mile" trips, replacing the use of private vehicles and reducing road congestion. The base price for a ride in the U.S is $1, and each minute costs 15 cents. E-scooters are extremely popular, with more than 150,000 rides per day in the U.S [6]. As of June 2023, there are 252 E-scooter systems in 51 cities across the U.S [9]. Some similar dockless sharing devices, such as shared bikes, are also popular. By the end of 2021, more than 10 million shared bikes have been deployed in 1,000 cities in China, with more than 30 million rides per day [12]. The popular use brought rich investment to the companies. Only one year after its establishment, Bird's valuation reached 2 billion U.S dollars.

There are many controversies about the shared E-scooter service. E-scooters are accused of being parked arbitrarily, occupying sidewalks, entrances and other places. Some users ride on sidewalks at high speeds, disturbing pedestrians and creating dangers. In terms of environmental protection, Hollingsworth et al. found that in 65% of simulations, E-scooters emitted more greenhouse gases than the replaced mode of transportation [5]. This situation was alleviated, when a two-year lifetime is assumed. Investors questioned the profitability of shared E-scooter companies, believing that shared equipment is difficult to recover costs. The CEO of Segway Ninebot (the main manufacturer of E-scooters) also questioned whether this is a sustainable business. Computation shows that an E-scooter needs to be put into operation for more than 5 months and used 5 times a day to recover its cost [1]. But the early micro-mobility data from Louisville, Kentucky, shows that the average lifetime of a scooter is 28.8 days, 3.5 uses a day [2]. Many devices are deliberately damaged and discarded in rivers or lakes.

The main method to accelerate the cost recovery and gain profitability is to increase the number of times an E-scooter is used during its lifespan. We tried to find ways to increase the usage of E-scooters. This article mainly analyzed the usage pattern of E-scooters in Austin, Texas. We first explored the usage characteristics of E-scooters per month, week, and day. K-means clustering method was used to explore the geographical usage characteristics. Based on the geographical pattern, we proposed strategies for re-deploying the E-scooters in some specific areas. E-scooters need to be recollected and charged regularly. After they are collected and charged by the operating company or "chargers" (Individuals that are paid to collect and charge the scooters at home, and redistribute them.), they can be re-deployed to some areas where the trips are dense to increase the usage rate. We also fitted a Negative Binomial generalized linear regression model to find out the relation between the number of times E-scooters are used and environmental (area) and social (income, education level, etc.) factors. These studies can provide a basis for the deployment of shared E-scooters in other cities in the future.

The analysis was based on the usage of micro-mobility devices by the residents in Austin, Texas, provided by the City of Austin Transportation Department [3]. We focused on the trips from March 2021 to February 2022. During this one year period, there are more than 3.7 million trips, 98% of which are traveled by E-scooters. The data set contains information such as the ID of each trip, ID of shared devices, the start and end time and location, etc. To protect privacy, the start and end times of the trip are accurate to 15 minutes. The originated and ended places of the trip are accurate to the census tracks in 2010 United States Census [10]. This geographic segmentation divides the city of Austin into approximately 170 areas, each with an area of approximately 1 square mile and approximately 5,000 people (median). We also introduced the American Community Survey 2022 (ACS 2022) data set [11], published by the Census Bureau, which contains data such as population, gender ratio, education level, and income level of each census track.

The article is organized as follows. In Section 2, we present the temporal and spatial analysis of the trip data. In Section 3 and 4, we describe the regression model and summarize the model fitting results. Finally, we conclude the study and analyze the limitations of the model and possible improvements.

2 Data Analysis

2.1 Temporal Data Analysis

We selected the shared E-scooter usage data of Austin residents from March 2021 to February 2022. Some abnormal trips, such as trips with too short or too long travel distance and duration, were removed, and 3,675,710 valid trips were left. The average trip duration is 12.7 minutes and the average trival distance is 1.14 miles.

We first analyzed how people use E-scooters in different months. The number of trips per month, total mileage, average mileage and duration, and the activity of E-scooters are summarized in Table 1. There are about 300,000 trips per month. May to October are the peak months for E-scooter use, with significantly higher number of trips and total mileage than the others. The month with the most usage is October, reaching over 470,000 trips, and the months with the least usage are January and February (160,000 trips). The longest average travel distance happens in April, reaching 1.32 miles per trip. The average travel distance in the remaining months is about 1.13 miles per trip. The total number of E-scooters available every months is more than 10,000, except March 2021. The most E-scooters are put into use in September (over 23,000). We define scooters that are used more than 20 times each month as active devices. The monthly activity of E-scooters is shown in Figure 1. Except for September (the number of available devices suddenly increased), December, January and February,

Table 1: Monthly Statistics of E-scooters Trips

Month	Trips	Total / Active Scooters	Total Mileage Traveled	Average Mileage Traveled	Average Duration Traveled
Mar 2021	260,633	9,331/6,177	342474	1.31	13.8
Apr 2021	282,560	10,889/6,358	373839	1.32	13.8
May 2021	371,285	10,745/8,157	437599	1.18	13.4
Jun 2021	352,150	11,726/7,652	400771	1.14	12.9
Jul 2021	397,041	11,838/8,460	436400	1.10	13.2
Aug 2021	350,823	11,637/7,993	379707	1.08	12.5
Sep 2021	397,064	23,401/8,665	410033	1.03	11.7
Oct 2021	476,851	12,465/9,818	529231	1.11	12.5
Nov 2021	248,684	11,255/5,925	265817	1.07	12.3
Dec 2021	213,087	10,881/5,225	243366	1.14	12.3
Jan 2022	161,759	11,808/2,968	181408	1.12	11.7
Feb 2022	163,773	11,841/3,173	186425	1.14	11.6
Average	306,309	12,318/6,714	348922	1.15	12.6

Note that, active scooter is defined as 20 times or more usage in a month.

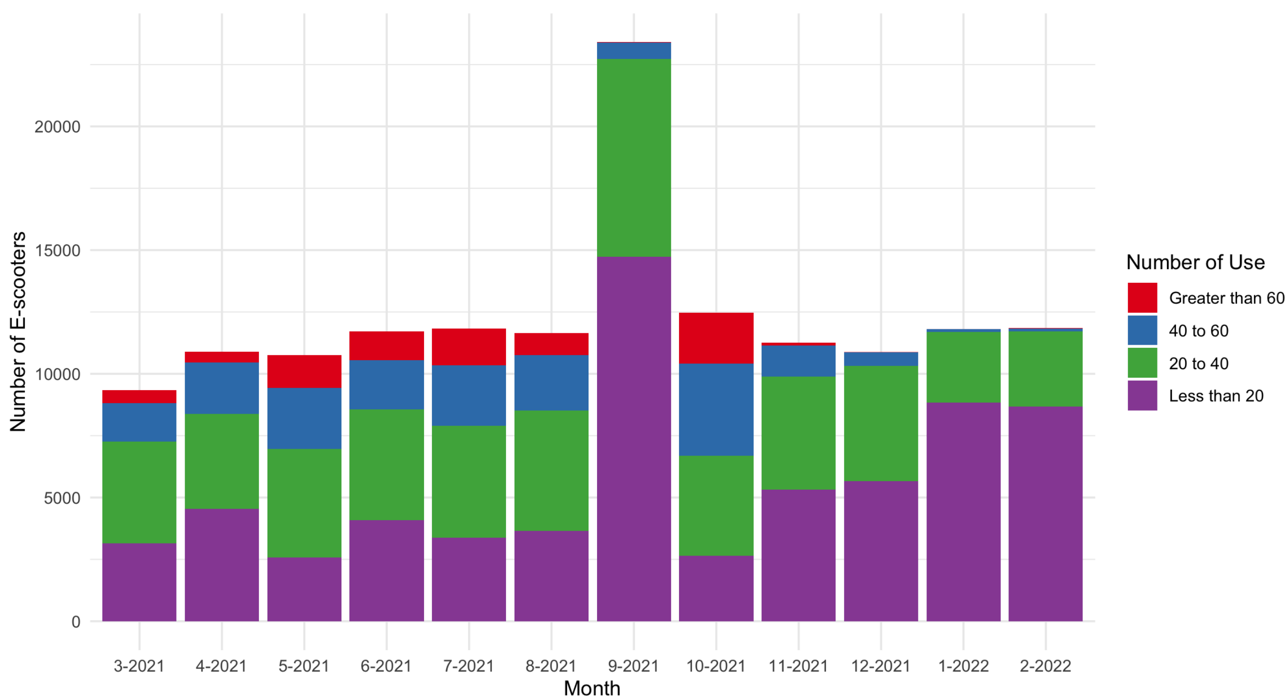

Figure 1: Number of (Active) E-scooters Available per Month

the proportions of active E-scooters in other months all exceed 50%. If the active standard of devices is raised to more than 60 times a month, only May, June, July and October have a proportion of more than 10%, which shows that the current daily usage of E-scooters is far from enough to recover the cost of the equipment.

We secondly explored the differences in usage within a day and a week and used a heat map to visualize it (Figure 2). On weekdays, people are more likely to use E-scooters from 9 a.m to 10 p.m. The peak usage on weekends lasts until 1 a.m. During a week, usage from Friday to weekends is higher than from Monday to Thursday. Friday afternoon and Saturday night are the busiest periods in a week. It is strongly recommended that the company should fully charge the vehicles before these two periods. From the different pattern of weekdays and weekends, we can conclude that people use E-scooters more for entertainment rather than commuting to work.

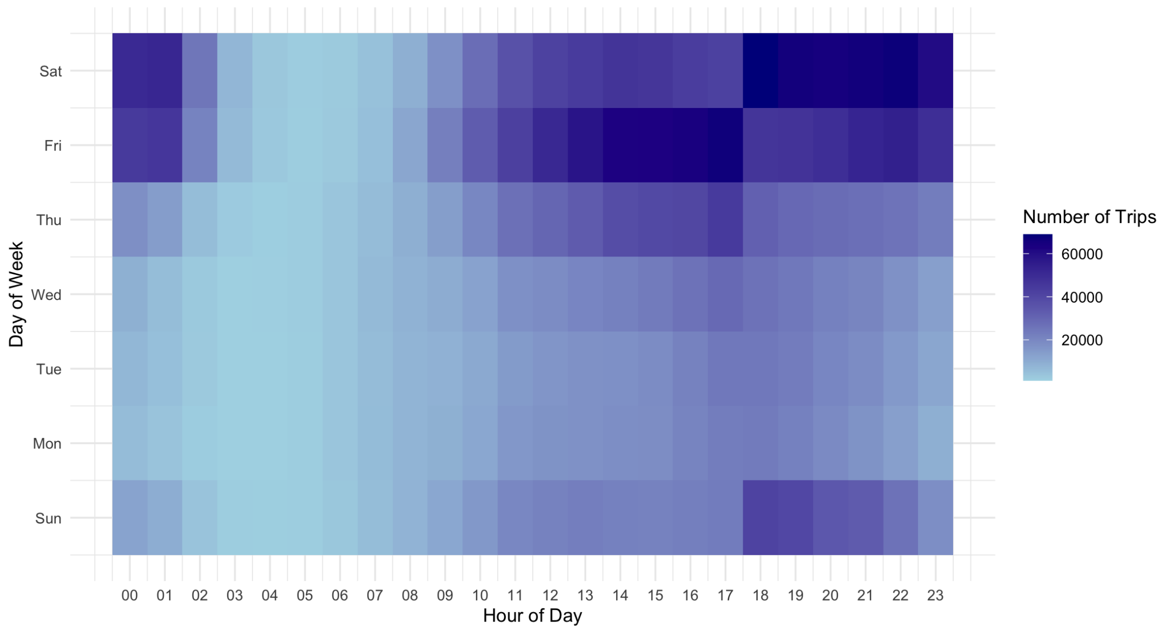

Figure 2: Heat map for Number of E-scooter Trips

2.2 Geographical Clustering Analysis

We used the K-means clustering algorithm to explore the spacial pattern of the usage. Three variables, the number of starting trips, the number of ending trips and the number of net trips (end - start) in each census track, are used as input. We also implemented Silhouette Method [8], which evaluates clustering quality based on data point's similarities within cluster and dissimilarities to the nearest neighboring cluster, to choose the number of clusters. Comparison of Silhouette Scores is in Appendix Figure 1. The 168 census tracks are divided into 4 clusters. The center of each cluster together with some socioeconomic characteristics (to be discussed in Section

3) are summarized in Table 2. Based on these statistics, we color the map of Austin (Figure 3) and conclude the four categories as

1. Extremely high traffic, high net outflow areas of E-scooters (indicated by red). This cluster only includes the downtown of Austin, where E-scooters are most intensively used and is the place where many trips start.
2. High traffic, relatively balanced inflow and outflow (indicated by yellow). The census tracks in this cluster surround the downtown and have high population density.
3. Medium traffic, net inflow areas (indicated by blue). These census tracks are distributed around the second cluster and are the destinations of many trips.
4. Low traffic, net inflow areas (indicated by white). These areas are distributed around the city, and the usage of E-scooters is relatively low. They are all net inflow areas.

Table 2: Center (Mean) of each Cluster

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Number of Start Trips	1,813,097	176,434	35,265	1,565
Number of End Trips	1,749,659	175,028	38,622	1,919
Number of Net Trips	-63,438	-1,406	3,356	354
Area	0.9	0.7	0.62	2.07
Population Density (1000 People/mile^2)	7.88	10.4	9.41	4.8
Gender Ratio	1.33	1.03	1.1	1.08
Young Proportion	0.54	0.65	0.63	0.45
Educated Proportion	0.77	0.76	0.69	0.53
log Median Income	11.9	10.9	11.1	11.3

Therefore, companies should encourage chargers to collect E-scooters in the census tracks in the third and fourth clusters and manually transfer them to the first and second clusters areas to increase the usage of E-scooters. The five census tracks with the highest net inflows and outflows are summarized in Appendix Table 1, 2. Based on this and the geographic location, we suggest to move E-scooters from census track 13.03 to 14.01, 10.00 to 9.02, 9.01 and 7.00 to 11.00, and 6.04 to 6.03 and 11.00.

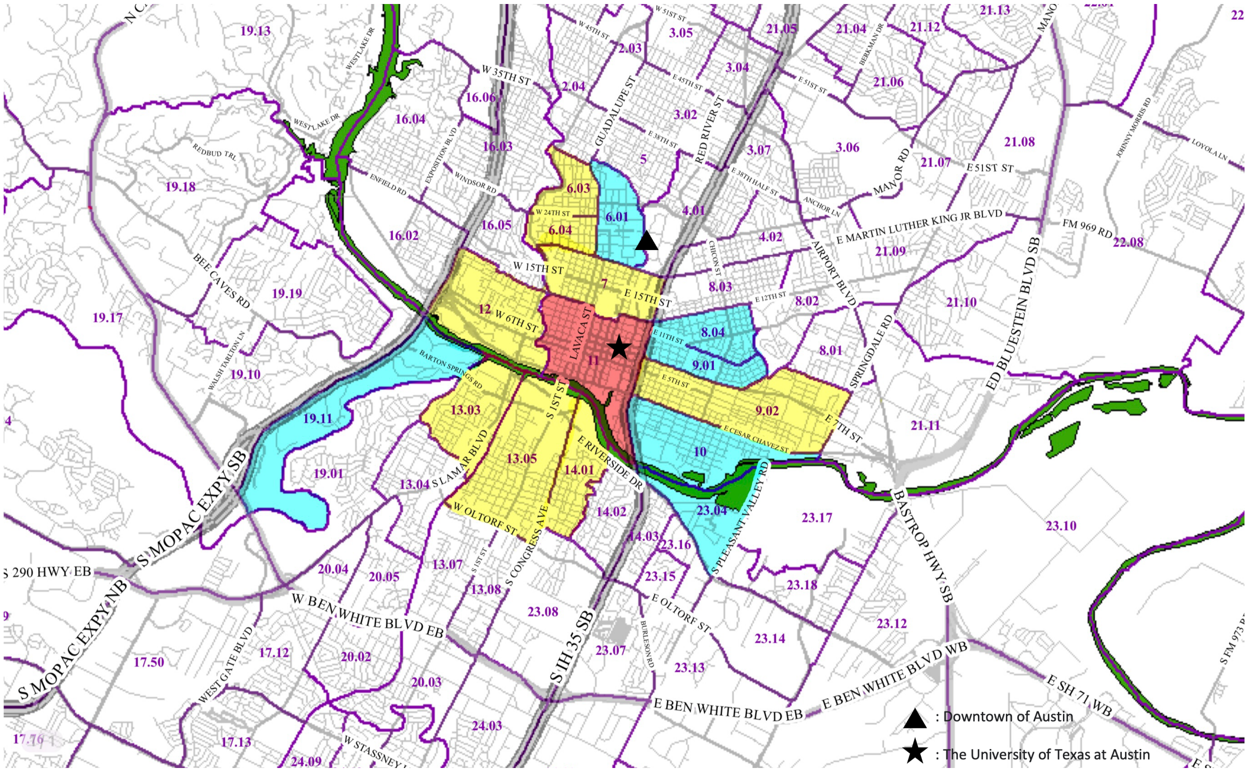

Figure 3: Clustering Results of Census Tracks in Austin TX

3 Model-Based Analysis

To better explore the relationship between socioeconomic factors and E-scooter usage, based on the clustering results of Section 2.2, we fitted a Negative Binomial generalized linear model. The number of trips (summation of start and end trips) in each month of each census track was used as the dependent variable. Compared with Poisson regression with log link, the negative binomial regression can better handle count data with large variance [7, 4] (In our situation, the mean and variance are 5077, 854283428, respectively). The mean and variance are modeled as

$$\begin{aligned} \mathrm{E}(Y) &= \mu = \exp(\beta^T X), \\ \mathrm{Var}(Y) &= \mu + r\mu^2 = \mu(1 + r\mu), \end{aligned} \tag{1}$$

where β is the regression parameter, and the multiplicative factor $1 + r\mu$ models the over-dispersion. Comparing with Quasi-Poisson regression, the multiplicative factor depends on μ [4].

The variables used in the regression model are summarized in Table 3, where the data on the environmental and social variables comes from ACS 2022. The divisions of the census tracks are different in the two data sets. We mapped the corresponding data according to the map. A total of 8 independent variables were selected for the model. Area (square mile) is the environmental variable. Societal variables include population density (1,000 people per square mile), proportion of young people, gender ratio, income and education level. We defined the young population as those aged 15 to 39. Gender ratio is the population of males over females. The median household income within the census track is used as the income data. The education level is the proportion of the population with a bachelor's degree or above. We also tried whether to add the clustering groups obtained by the K-means in Section 2.2 as independent variable. R (4.3.2) was used to build the model.

Table 3: Dependent and independent variables in the Generalized Linear model

Variable	Type	Median	Mean	SD	Min	Max
Total Trips per Month	Numeric	118	5,077	29,228	1	427,429
Month	Categorical	N/A	N/A	N/A	1	12
Area (mile^2)	Numeric	0.9	1.38	2.15	0.2	27.5
Population Density (1000 People/mile^2)	Numeric	5.09	5.75	3.93	0.30	25.26
Gender Ratio	Numeric	1.05	1.08	0.19	0.66	1.79
Young Proportion	Numeric	0.45	0.49	0.15	0.15	0.99
Educated Proportion	Numeric	0.56	0.55	0.21	0.06	0.93
log Median Income	Numeric	11.31	11.27	0.49	9.11	12.39
Cluster	Categorical	N/A	N/A	N/A	1	4

4 Regression Results

The estimated parameters, standard errors and z-values of the two models are summarized in Table 4. By adding the clustering variable, two variables, population density and gender ratio, became insignificant, and AIC reduced form 21199 to 20103. Therefore, we use the full model for analysis.

The parameters for month from May to October are significant (p-value < 0.01), and the estimated parameters are approximately between 0.5 and 0.9. This shows that people are more inclined to use E-scooters during these months, same as the analysis in Section 2.1. Area is significantly negatively correlated with E-scooter usage. Large census tracks are located on the edge of Austin city, and people's willingness to use E-scooters for long-distance travel is low, which matches with intuition. The proportion of young people and education level are significantly positively correlated with the number of scooter uses (with estimated parameters 4.42, 2.69, respectively), indicating that young and well educated people use E-scooters more. Since finding, unlocking

Table 4: Regression Results for Negative Binomial Models

Model	Model With Cluster			Model Without Cluster		
AIC	20103			21199		
Variable	Parameter	Std Error	Z-Value	Parameter	Std Error	Z-Value
Intercept	5.69	1.77	-0.114	-19.82	2.06	-9.643***
Month Apr	0.21	0.18	1.142	0.15	0.23	0.618
Month May	0.51	0.18	2.836**	0.45	0.23	1.913.
Month Jun	0.50	0.18	2.813**	0.37	0.23	1.636
Month Jul	0.51	0.18	2.856**	0.45	0.23	1.957.
Month Aug	0.56	0.18	3.155**	0.40	0.23	1.765.
Month Sep	0.66	0.18	3.701***	0.48	0.23	2.107*
Month Oct	0.88	0.18	4.881***	0.71	0.23	3.070**
Month Nov	0.20	0.19	1.088	0.11	0.24	0.481
Month Dec	0.06	0.18	0.319	-0.06	0.23	-0.269
Month Jan	-0.12	0.18	-0.673	-0.28	0.23	-1.211
Month Feb	-0.11	0.18	-0.603	-0.21	0.23	-0.920
Area	-0.15	0.02	-7.726***	-0.11	0.02	-4.427***
Population Density	-0.01	0.02	-0.696	0.17	0.02	9.271***
Gender Ratio	0.26	0.20	1.273	1.04	0.26	3.956***
Young Proportion	4.42	0.37	12.040***	6.10	0.46	13.176***
Educated Proportion	2.69	0.25	10.654***	5.08	0.31	16.128***
log Median Income	0.17	0.14	1.166	1.68	0.18	9.307***
Cluster 2	-2.17	0.43	-5.016***	N/A	N/A	N/A
Cluster 3	-3.48	0.44	-7.987***	N/A	N/A	N/A
Cluster 4	-5.87	0.41	-14.259***	N/A	N/A	N/A

'***' p-value < 0.001, '**' p-value < 0.01, '*' p-value < 0.05, '.' p-value < 0.1

E-scooters and all the rest steps require the use of the mobile app, this creates obstacles for the elderly and those without mobile phones to ride E-scooters. Companies need to consider providing more convenient ways for these people to expand their user base.

The population density is not significant (p-value = 0.48), which may be related to the fact that the downtown is more commercial than residential. Although there is a positive correlation between the number of uses and income, it is not significant (p-value = 0.24). The current price is not a deterrent to trips with E-scooters. The parameter of gender ratio is not significant, indicating that there is no difference in the gender of users. However, without any user information in the data set, more accurate conclusion about gender-equal usage cannot be drawn. Since the first cluster is used as the basis, the other three clusters are significantly negatively correlated with the response.

5 Conclusions

This article explored the use of shared E-scooters in Austin Taxes from March 2021 to February 2022. People use approximately 12,000 E-scooters to travel more than 300,000 trips every month. Each trip lasts approximately 13 minutes and travels 1.1 miles. About 50% of the E-scooters are used more than 20 times a month, and less than 10% are used more than 60 times a month. May to October is the peak period for E-scooter usage. We found that more trips occurred on Friday to Sunday during the week, and more trips within a day occurred in the afternoon and evening. From geographical point of view, downtown and its surroundings are areas where E-scooters are intensively used, and a large number of trips depart from here. The periphery of the city is a net inflow area for E-scooters. The regression model shows that the number of E-scooters used is positively correlated with summer season, the proportion of young people, and education level, and negatively correlated with area. Surprisingly, population density is not significantly related to the usage. Gender ratio and household income are also found no significant relationship.

Overall, shared E-scooter services do greatly facilitate people's first/last mile travels. But they are used more for recreational travel than daily commuting. The utilization rate is insufficient to recover the equipment cost and still needs to be improved. We give the following possible suggestions to increase the number of times an E-scooter is used. Companies can reduce the amount of equipment they put in during the winter to extend equipment's lifespan. Charging on Fridays and weekends should be prioritized. Some usage methods that do not rely on mobile apps should be considered to increase the number of potential users. More vehicles could be collected from the edges of the city and redistributed to downtown, near the University of Texas at Austin, and other schools.

6 Limitations

The study acknowledge certain limitations, such as the potential exclusion of latent variables, leading to unaccounted-for factors. For example, the land use pattern, average number of cars per household, and the convenience of public transportation. The lack of these variables in the regression model may affect the analysis. The start and end points of each trip in the data set are accurate to census tracks, each of which is about 1 square mile. This area is relatively large, which may lead to insufficient precision in the study. GIS software may be used to perform more detailed segmentation and geographical analysis.

Appendix

Table

Census Track ID	Net Trips
1100	-63438
1401	-20156
902	-6817
1200	-6184
603	-667

Table 1: The 5 Largest Net Outflow Census Tracks

Census Track ID	Net Trips
1303	10996
1000	8259
604	4973
901	4884
700	4491

Table 2: The 5 Largest Net Inflow Census Tracks

Figure

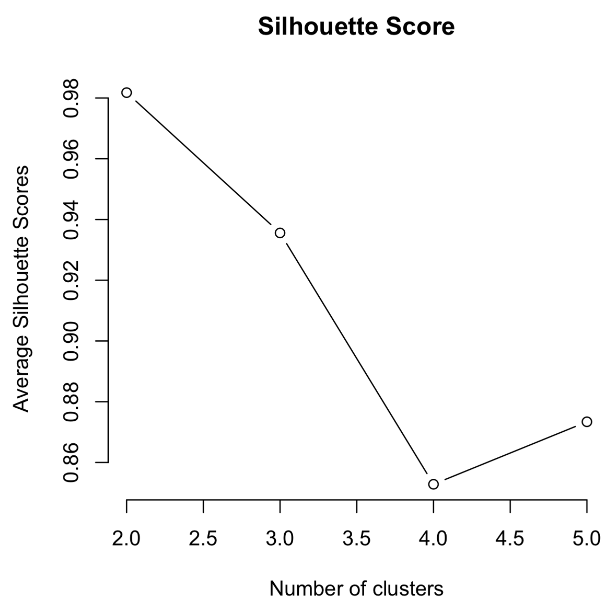

Figure 1: Silhouette Score for Different Number of Clusters

References

[1] Alison Griswold. Scooter Unit Economics Are Worse Than They Seemed

[2] Alison Griswold. Shared Scooters Don't Last Long

[3] City of Austin Transportation Department. Shared Micromobility Vehicle Trips

[4] Jay M. Ver Hoef, Peter L. Boveng. Quasi-Poisson V.S Negative Binomial Regression: How Should We Model Overdispersed Count Data?

[5] Joseph Hollingsworth et al 2019 Environ. Res. Lett. 14 084031. Are E-scooters Polluters? The Environmental Impacts of Shared Dockless Electric Scooters

[6] National Association of City Transportation Officials. Shared Micromobility in 2022

[7] NCCS Statistical Software. Negative Binomial Regression

[8] Peter J. Rousseeuw. Silhouettes: A Graphical Aid to The Interpretation and Validation of Cluster Analysis

[9] U.S. Department of Transportation Bureau of Transportation Statistics. Bikeshare and E-scooter Systems in the U.S.

[10] United States Census Bureau. 2010 Census Tract Reference Maps

[11] United States Census Bureau. Census Reporter for Census Tracks in Austin Taxes

[12] 中国城市规划设计研究院. 2022年度中国主要城市共享单车、电单车骑行报告