Do shared bike ridership patterns reveal income disparities in the District of Columbia?

Ilgi Toprak

Abstract

In this project, I aim to assess commuter bike activity density and income relationship in D.C. based on Capital Bikeshare data. I first investigate through exploratory analysis if shared bike ridership is more prevalent in census tracts associated with higher-income residents. In addition to Capital Bikeshare data, I use median household income and commuters' ridership data to assess income and preferences and other contributing factors, such as bike infrastructure throughout the city. I test the hypothesis that high and low-income census tracts have similar ridership activity. I use linear and multivariate regression methods to determine if income is a significant feature of bike-share ridership density. The study results show a weak positive relationship between activity density and income, while low and high-income areas have different patterns of activity density. There is a clear separation in how commuters from different income groups move.

1- Introduction

Bike sharing is becoming a popular micro-mobility tool in significant cities in the U.S. . Bikeshare's motivation is that it is convenient (Fishman, 2016), accessible and reliable transportation method, and a preferable feeder to the subway system, especially for middle-income residents. Although the initial investment for using the bike-share system remains low in many cities, many recent studies question the equity of bike-share systems in terms of income and demographics. This study will look into Washington DC's Capital Bikeshare to assess if residents' income is a significant feature of shared bike ridership.

In this regard, the study's research question is, **'Do shared-bike ridership patterns reveal income disparities in the District of Columbia? '.** To assess these patterns, I intend to make an in-depth exploratory analysis explaining the temporal and spatial use of share bikes throughout the city and consequently test the hypothesis that **shared bike ridership is more prevalent in higher-income than lower-income neighborhoods.** Using census tracts as spatial relationships, I will then make a regression analysis to assess shared-bike ridership features such as activity density, speed, distance, and duration vs. income in D.C.

The research mainly questions whether a medium in the city transportation system is fair and equitable among citizens of different income levels. This question is significant for research because the District of Columbia already has high income and racial disparities noticeable at the level of census tracts. This fact translates into urban spaces (see our previous research Toprak&Ekdi, 2022) and significant differences in infrastructure, including the lack of assigned bike lanes in lower-income neighborhoods. Secondly, shared bike ridership is a relatively new concept in the United States, and higher-income residents have better opportunities to try and implement novelties. Elaborating the research around this question will contribute to understanding inequities around micro-mobility patterns. The research problem is significant because it brings an equity perspective to an urban transport component. Findings might contribute to implementing policies for a more equitable bike-share system in the city.

2- Literature Review

The driving force behind using bike-share, is convenience; membership in the bike-share program is significantly influenced by how far away one lives from a docking station (Fishman, 2016). Various studies are concentrating on bike-share users' route choice behavior, bike-share station location optimization, the effectiveness of bike-share as a metro feeder system, and bike-share use patterns before and after the pandemic, among others. However, my approach to the literature review will be mainly based on how bike-share systems relate to inequality patterns.

Recent studies mention how the gender, income, and demographics of bike-share customers influence overall bike-share use patterns. In these studies, there is a consensus that bike-share users typically have higher average incomes, education status, and employment (Fishman, 2016; Fishman et al., 2015; Lewis, 2011; Shaheen et al., 2013; Woodcock et al., 2014). Evidence suggests that lower-income and people of color (POC) Americans utilize bike share less frequently than higher-income and white Americans (Dill et al.). These previous findings indicate that people from different income levels and demographics may have unequal access to bike share. Generally, bike-share dock allocation concentrates on stations in densely populated city centers (Ursaki and Aultman-Hall, 2015). This fact has had unintended consequences, such as limiting bike-share access for historically underserved groups (Ursaki and Aultman-Hall, 2015). Transit connectivity and extensive coverage areas walkable distances between stations are promoted qualities of bike-share systems, but very few established quotas for the proportion of stations in low-income and minority areas (Howland, 2017) that help close the exposure gap for lower-income individuals and POC (McNeil et al.,2018). The Covid-19 pandemic had significant implications for bike-share use and behaviors and changed how equitable bike-share systems are (Wang & Noland, 2021). Jobe and Griffin (2021) argue that bike-sharing operators should increase communication about community health policies and initiatives, look into the best ways to serve neighborhoods with unemployment and low incomes, and get ready for an equitable increase in ridership after the pandemic.

Previous studies for Washington, DC, measure the effectiveness and income equity of Capital Bikeshare. Buck et al. (2013) indicate that Capital Bikeshare had higher income levels compared to the median household income of the U.S. Census Bureau; however, they had lower income levels compared to regular cyclists. From a demographics perspective, 50% of the population of Washington, D.C., are African-American, and only 3% of Capital Bikeshare members were African-American, compared to 8% of all bicycle users in 2013 (Buck et al., 2013). Su et al. (2022) argue that the equity program for Capital Bikeshare in Washington, DC, now successfully encourages low-income use. However, Ebrahimi et al. (2022) indicate that docking stations are distanced apart, especially in some exclusively residential areas with limited access to bikeshare. A study investigating bike-share equity in seven cities across the US suggests that the most egalitarian cities were Washington DC and Arlington, sharing Capital Bikeshare(Ursaki and Aultman-Hall, 2015). This study will assess if income is a significant feature in bike-share ridership activity and assess bike-share equity in DC.

3- Data

Washington, DC, is the study context for analyzing the potential income inequities in the DC Capital Bikeshare system. Capital Bikeshare data includes every ride, duration, start and end times and locations, and member and casual user information and is free to access on the Capital bikeshare website (Capital Bikeshare data, 2022). A limitation of the dataset has been the change in the data types in March 2020. More data has been added after April 2020, such as latitude and longitude values. All data has been extracted for exploratory time-series analysis, covering a span between 2018-2022. All spatial analyses were made using data after March 2020 containing latitude and longitude data for spatial joins. As there is a focus on commuter data, 2020-2022 bikeshare data has been filtered by typical commute duration, OD trips (filtering out loop trips), trip distance (found from O-D longitude and latitude using the function haversine), and lastly by the typical speed of a commuter up to 18 mph.

Washington, DC, has eight wards and 179 census tracts. Data on a more granular level is available on Open Data DC (Open Data DC, 2022), mainly based on census tracts, including demographic and economic aspects. I investigate the geographic distribution of income, commuters' ridership preference, and working and resident population using this data accessible through the DC Economic Characteristics of Census Tracts dataset. I use a bike infrastructure dataset including separated and unseparated bike lanes, bike-share docking locations, and bike-share walkshed. I include a census tract shapefile for DC for spatial filtering of bike usage in only longitudes and latitudes into DC census tracts.

4- Methods

The methodology is based on three sections:

(1) Through exploratory analysis, I make a temporal assessment through a time-series analysis, month-of-the-year and day-of-the-month analyses, weekend-weekday differences, hour-of-the-day (rush hour and regular hour) differences, member and casual user differences. I make a spatial analysis of the bike-share ridership, all rides, member vs. casual, and commuter pattern activity densities by census tracts, as well as all contributing socio-economic factors such as income, commuter ridership preference, and bike infrastructure.

(2) I test the null hypothesis that low-income and high-income areas have the same underlying distributions in bike-share activity density, speed, duration, distance, and population. The top 25% high-income vs. the bottom 25% low-income census tracts will be tested as two samples using the t-test and the Kolmogorov-Smirnov test.

(3) I use linear and multivariate regressions to measure the correlation between bikeshare commuter activity density and income by census tracts. Activity density is the dependent variable, and all other contributing factors, including income, will be used as features. I use multivariate regression methods to assess how significant income is as a feature in different situations where other contributing factors are

present or not, and interpret using regression analysis writing guides (Allison, 1999; Miller, 2013) and research using similar methodologies (see Caspi & Noland, 2019).

5 -Results

5.1 - Exploratory analysis results

5.1.1 - Temporal analyses of bikeshare ridership

The time-series comparison of member vs. casual bike-share daily rides (see Figure 1) shows a distinction in riders' behavior before and after March 2020. With the start of the pandemic and the first lockdown, members have significantly decreased bike ridership activity, which can be explained by working from home. The commuter activity decreases and slowly rise again in 2021, with no significant recovery compared with the pre-pandemic era. Ridership among non-members increased during the pandemic as leisure rides became popular and almost equal to the daily member rides count.

The number of member rides exceeds the number of casual riders daily, weekly, and monthly. On an hourly basis, morning and evening rush hours are mostly constituted by member rides. Members are the most active from 6 AM to 10 AM and from 4 PM to 8 PM (Figure 2). The duration analysis (Figure 3) has a right-skewed distribution with an SD of 14.93; when the 5% and 95% quantiles are filtered out, the meaningful duration ranges between 3-45 min. These findings define commuters' morning rush hours filters used in later parts of this study.

5.1.2 - Spatial analyses of Income, commuters' data, and bike-share infrastructure

Income analysis (Figure 4) shows that the NW census tracts constitute the wealthiest neighborhoods, vs. SE neighborhoods have the lowest median income compared to others. This West/East divide in income level differences will be a significant input for later analyses. The commuters' ridership modes (Figure 5) show that their preference for driving could result from higher income levels and the distance to the center. Walking mode is concentrated in the center, and public transportation is concentrated along the metro and bus axes. Other means seemingly include bikeshare and personal bike data concentrated in places not covered by public transport. Commuters' ridership modes seem to show that distance and availability of transport services are major determinants of commute ride preference.

Bikeshare infrastructure analysis (Figure 6) indicates that bikeshare services are widely available in DC, although most docks concentrate in the city center. Low-income and high-income areas have equal access to bikeshare, and the bikeshare walkshed is available everywhere except in the north, northeast, and south. Most are bike lanes concentrated in the center of the city. However, signed bike lanes are more available in the peripheries linking the center to the major residential areas. The availability of infrastructure will be a feature in the multivariate regression model.

5.1.3- Spatial analyses of bikeshare ridership

Bikeshare activity density analysis (Figure 7) shows that most bikeshare activity concentrates in the city center. The distance from the center is the significant component in determining the density of bike use in the center vs. the boundary. The NW and N areas where ridership activity is of average density (yellow hues) are the axes where we encounter the most bikeshare infrastructure. Although there are no significant similarities with the median income heatmap, the SE – a low-income area – has less activity density than NE and NW areas. Bikeshare activity density analyses of member and casual users comparison (Figure 8) shows that central areas in the city, especially the mall, are used more by casual users. No significant difference explains the difference in member and casual uses for low or high-income residential areas.

Bikeshare commuters' morning activity density analysis (Figure 9) indicates commuters move mostly to the city's center as the destination heatmap has higher range activity density values in the center. The mean distance is the highest in commuters cycling from the farthest away from the center. Mean speed is the highest in green areas, parks, and waterfront biking areas. Mean speed and mean distance show an inverse trend on the heatmap (higher ranges at the boundaries of the city) because it is more likely that the distance and the speed of a commuter are higher if starting from farther away. No specific patterns indicate high or low-income areas differentiating from other areas.

In the last exploratory analysis, I explore where the morning commuters from the five highest and lowest income census tracts go (Figure 10). In terms of their destinations, bikeshare commuters either stay in similar income areas or go to the city center. Almost no users go from a high-income area to a low-income neighborhood and vice versa. This analysis indicates a restricted mix in the income profiles regarding O-D pairs. Distance and increased job availability in the city center might be the reasons behind this trend.

5.2 - Hypothesis testing

We can reject the hypothesis that the activity densities and the ride counts of the top 25% high income and bottom 25% of low-income census tracts belong to the same distribution at a 99.999…% confidence level. We keep the null hypothesis that the speed (k-s p-value:0.298, tt p-value:0.754), distance (k-s p-value: 0.795, tt p-value: 0.335), and duration (p-value: 0.256, tt p-value: 0.237) of the top 25% high income and bottom 25% low-income census tracts belong to the same distribution (Table 1).

An example histogram plot of distance distributions (Figure 11) in low vs. high-income neighborhoods shows that the distributions are similar. However, the volumes of the total distance for the two groups on the histogram are very different. This means high-income and low-income neighborhoods have different amounts of activities. High-income census tracts produce a larger activity volume than low-income census tracts.

5.3 - Correlations, linear and multivariate regression analyses

On the correlation table (Table 2), the strongest correlation is between duration and distance (0.91); the cyclists from the farthest away spend more time on rides. Speed and distance are also correlated (0.61), as the ones traveling farther are more likely to speed. Log activity density and income are also correlated (0.45), but they are weaker.

The linear regression table (Table 3) reports the coefficient of the independent variable and the r-squared values for each of the regressors. While all of the pairs have low r-squared values, meaning that any of the regressors, including income, do not explain much of the variation of activity density, I also observe that some coefficients are weak positives (income and infrastructure) and others are weak negatives (duration, distance, and speed). This means that a rise in income or infrastructure is likely to increase activity density in an area.

All regressors and the independent variable are standardized before multivariate regression analysis (Table 4). Speed, distance, and duration have a negative coefficient; the value of the activity density is anticipated to drop as the value of the activity density and income rises when speed, distance, and duration are the other predictors. Income has a weak positive connection with activity density in all multivariate cases. This indicates that activity density is expected to rise slightly when the median income rises. Infrastructure is the only other regressor that has a positive effect on activity density. Distance also has a positive coefficient when all the regressors are added to the model.

6 -Discussion and conclusion

Why is there only a weak positive relationship between bikeshare ridership activity density and income?

- First, there are other potential regressors that I have not considered. Some of these regressors with potential negative impacts on activity density could be traffic congestion or bike/vehicle collisions. The density of jobs and leisure activities could have been positive contributors.
- Having extreme activity density in the city center vs. residential areas prevents me from differentiating the impacts of low vs. high-income neighborhoods and seeing the patterns of how they contribute to the overall model.
- Even if income has a weak positive relationship with activity density, this does not mean that low and high-income areas have similar activity density patterns. I already found in hypothesis testing that activity densities of highest vs. lowest income census tracts do not have the same distribution and show different levels of activity density where high-income density is much higher.

A question that might interest is investigating the segregation effect in O-D pairs for high-income and low-income profiles. I shortly investigated this question in section 5.1.3 and Figure 10, rides starting high-income, and low-income census tracts only mix in the city center. There is a clear separation in the ways different income groups move, but this is more likely to be a problem of residential segregation and not the lack of access to bikeshare.

The main finding is that commuter activity density and income have a weak positive relationship. Because the activity density is very high in the city center, the impact of activity density in high vs. low-income areas does not seem to differentiate the overall multivariate model. However, movement patterns and histogram plots of commuter activity in low vs. high-income areas show that these groups show differences in bikeshare use. Disparities appear more in the movement patterns and less in the access opportunities.

Tables and figures (in order of appearance in the report)

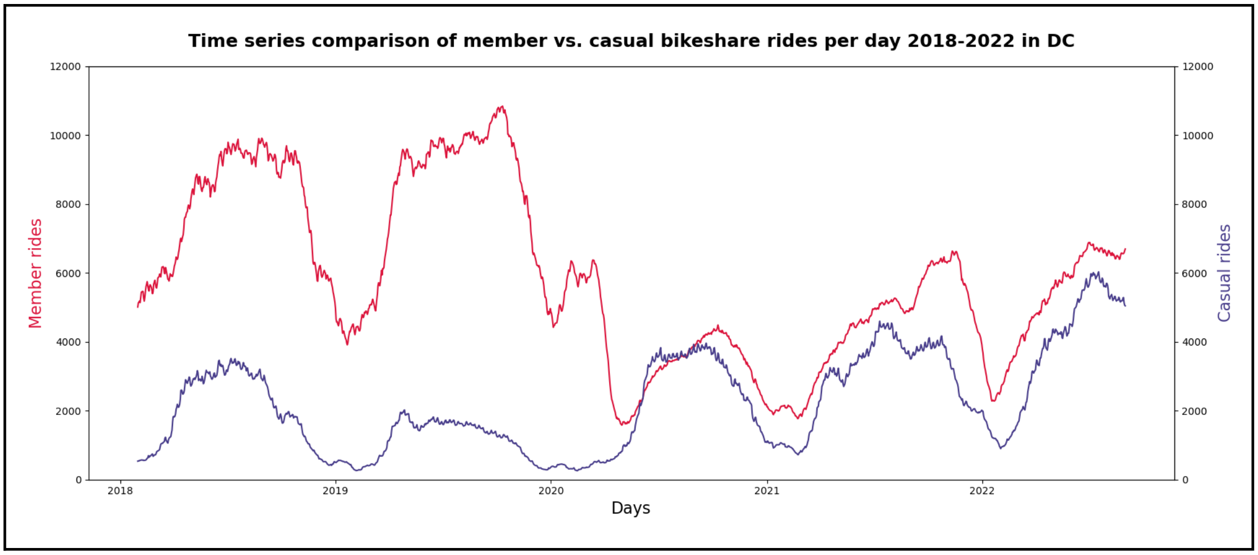


Figure 1 - The time-series comparison of member vs. casual bike-share daily rides (2018-2022)

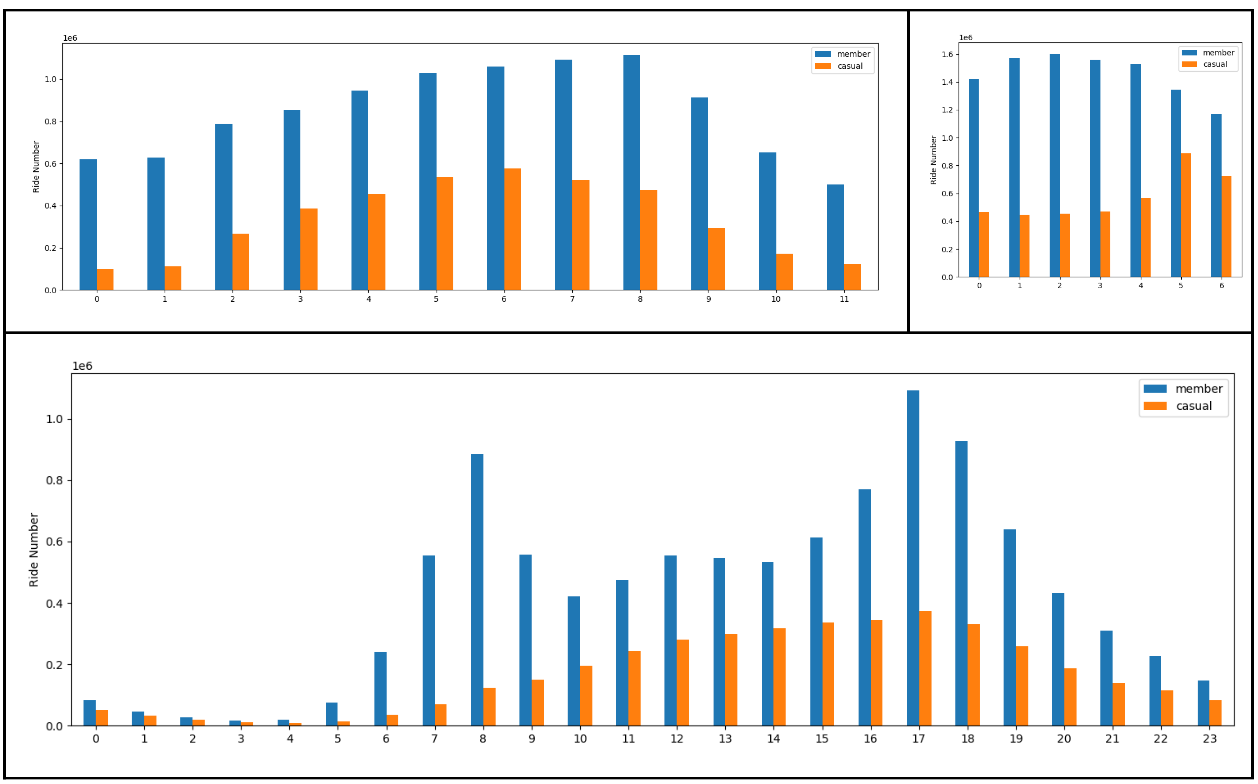
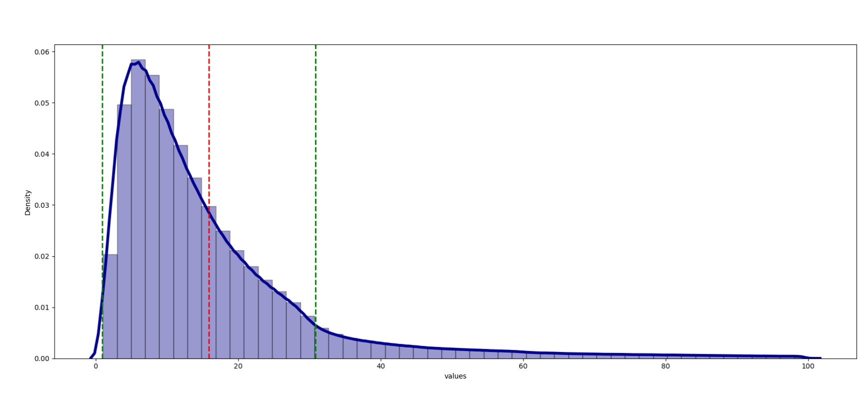


Figure 2 - (1)Month of the year, (2)day of the week and (3)hour of the day bar plots of member vs. casual bike-share daily rides (2018-2022)



Mean = 15.88
Standard deviation = 14.93
1 percentile = 1.00
5 percentile = 3.00
25 percentile = 6.00
50 percentile = 11.00
75 percentile = 20.00
95 percentile = 46.00
99 percentile = 80.00

Figure 3- Bike trip duration distribution analysis and statistical properties of the distribution.

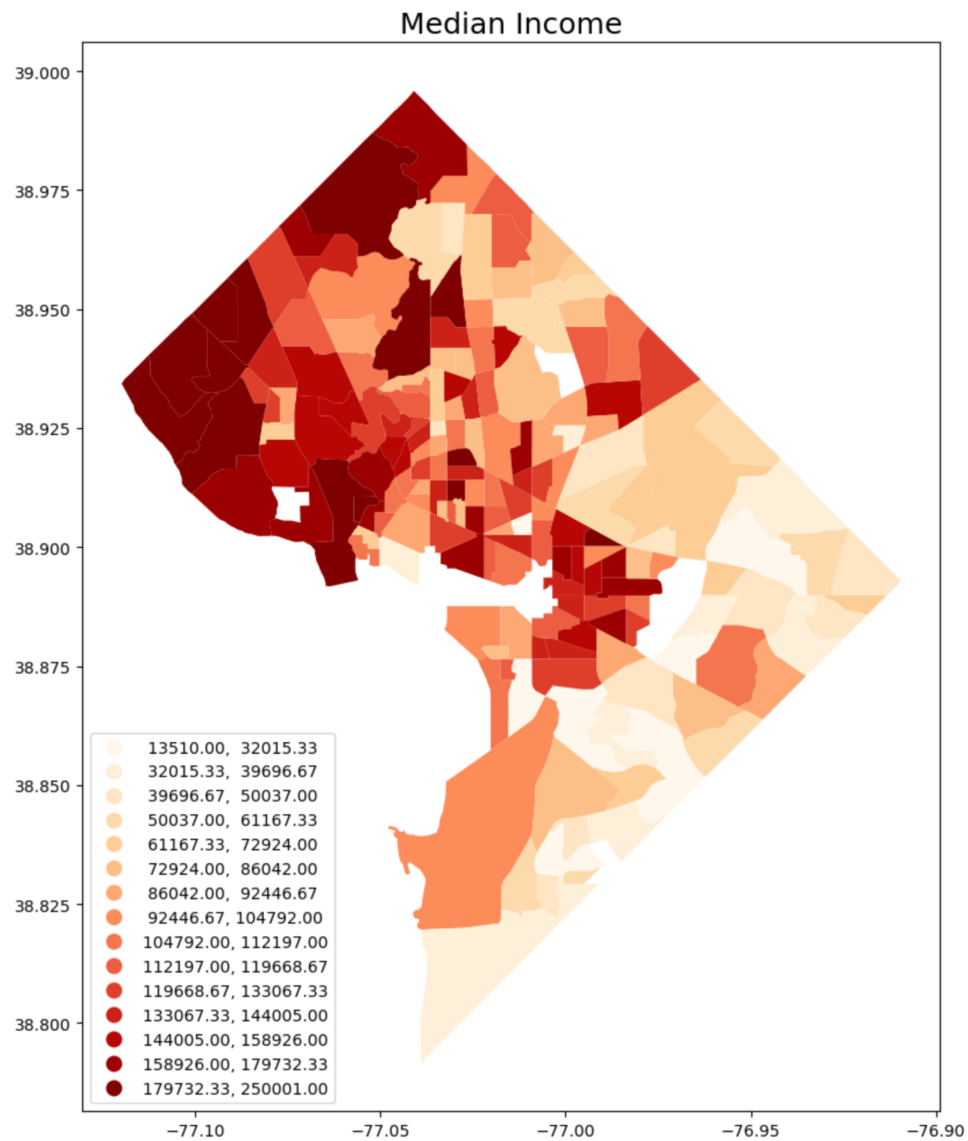


Figure 4- Median Household income per census tract

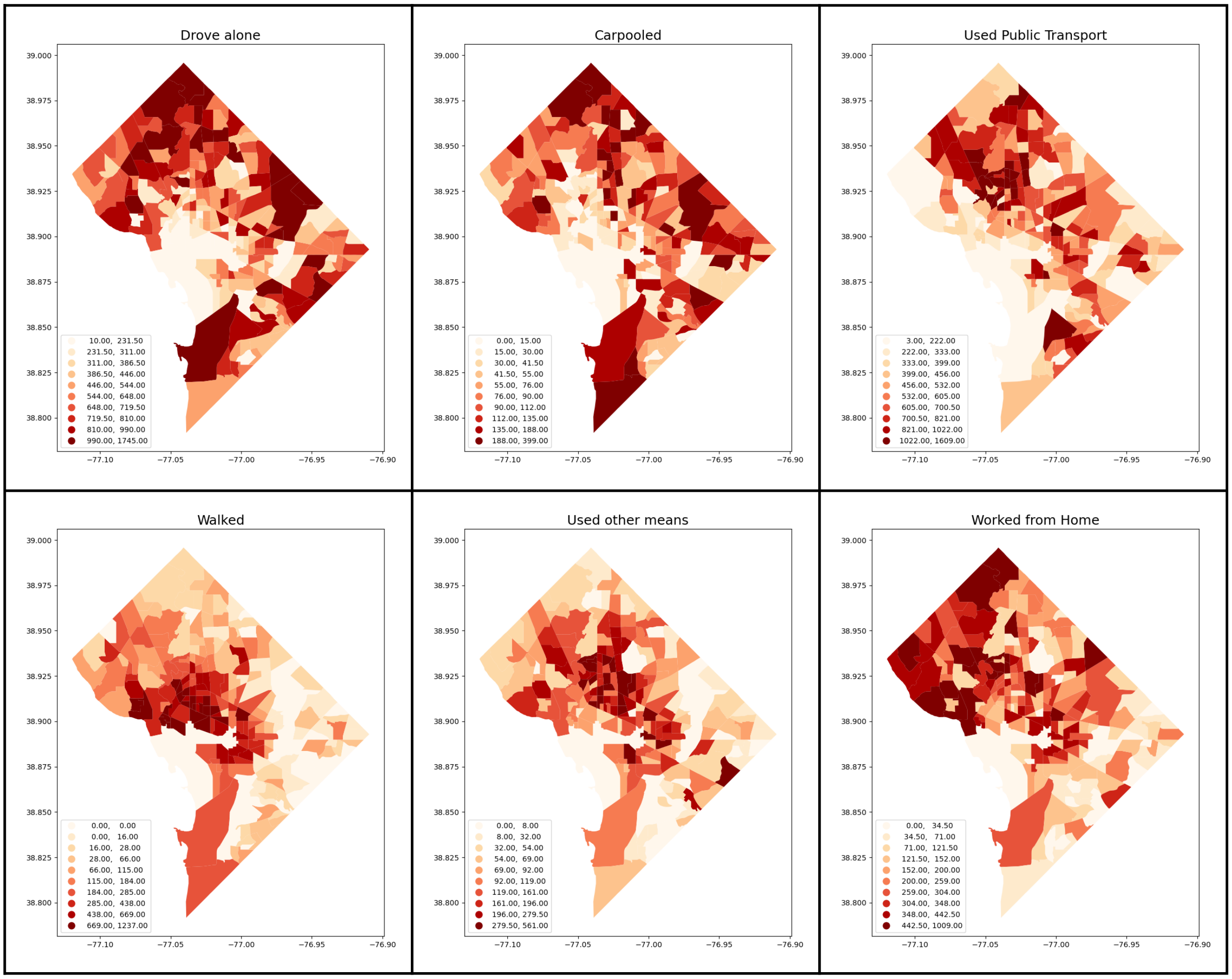


Figure 5- Commuters' ridership data per census tract

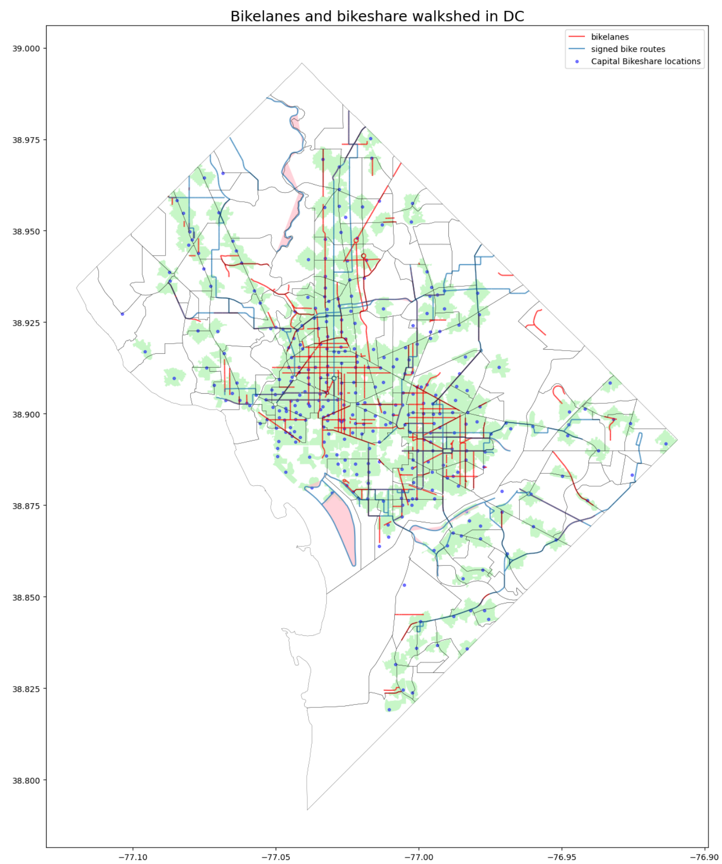


Figure 6- Bike infrastructure: bikelanes, signed bike routes and bikeshare locations

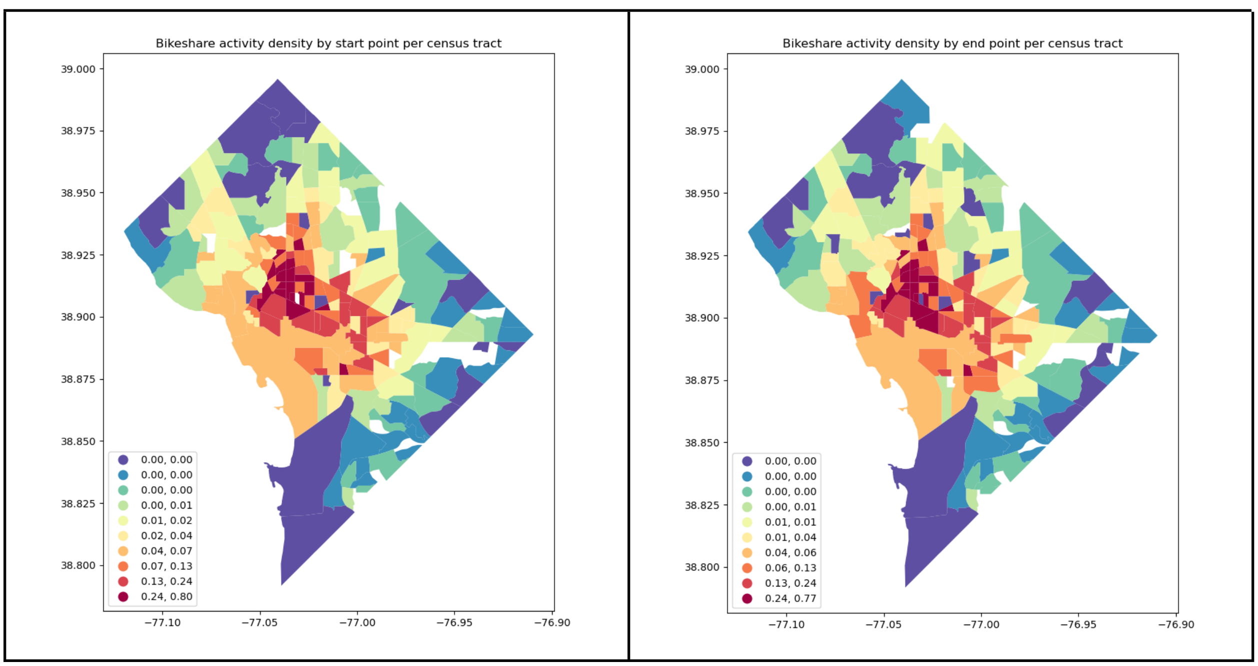


Figure 7- Comparison of the heatmaps of bikeshare activity density (1) by start points, (2) end points per census tract for all rides 2020-2022.

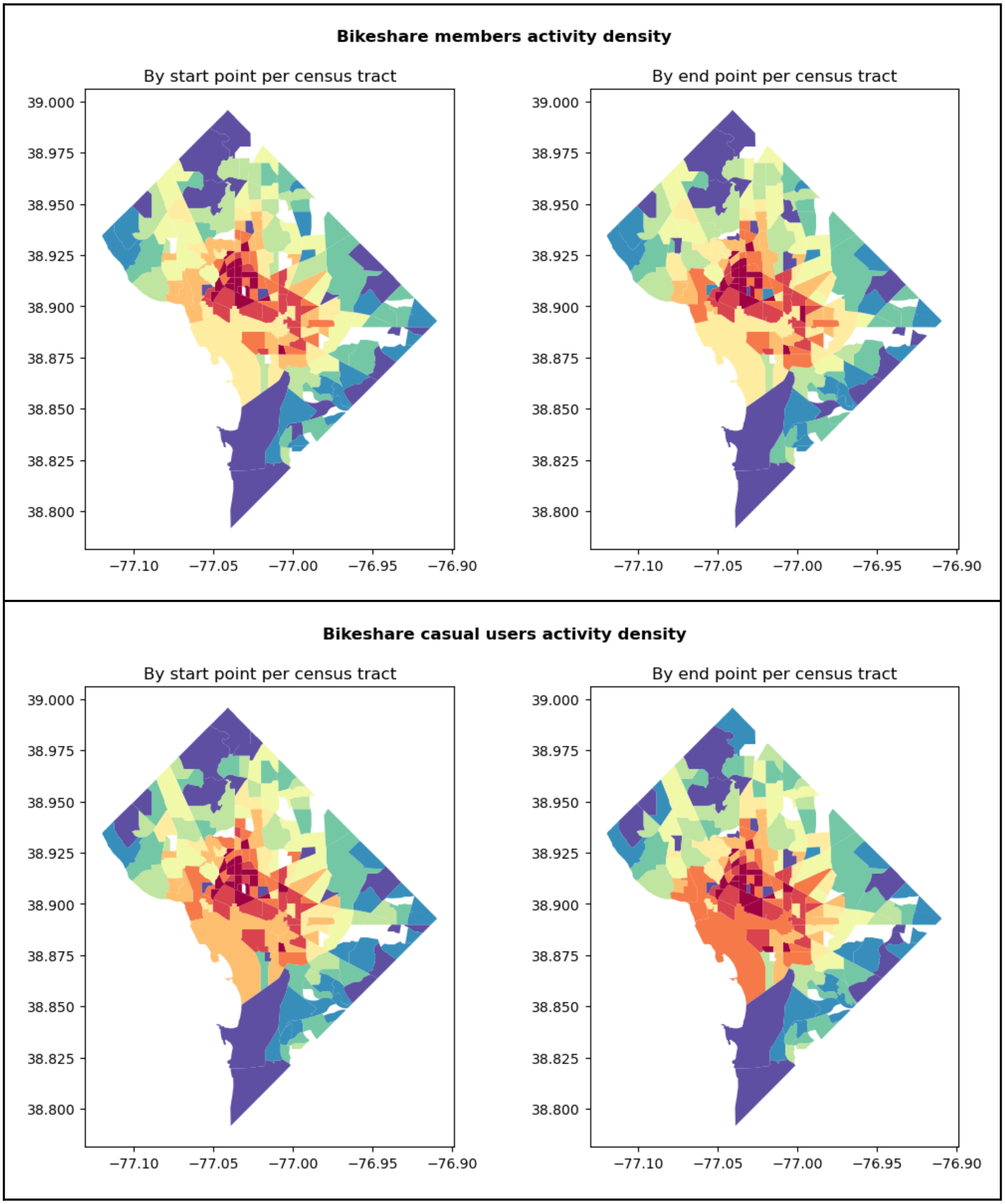


Figure 8- Comparison of the heatmaps of bikeshare activity density by census tracts, first two maps member (start/end) and last two maps casual users (start/end) 2020-2022.

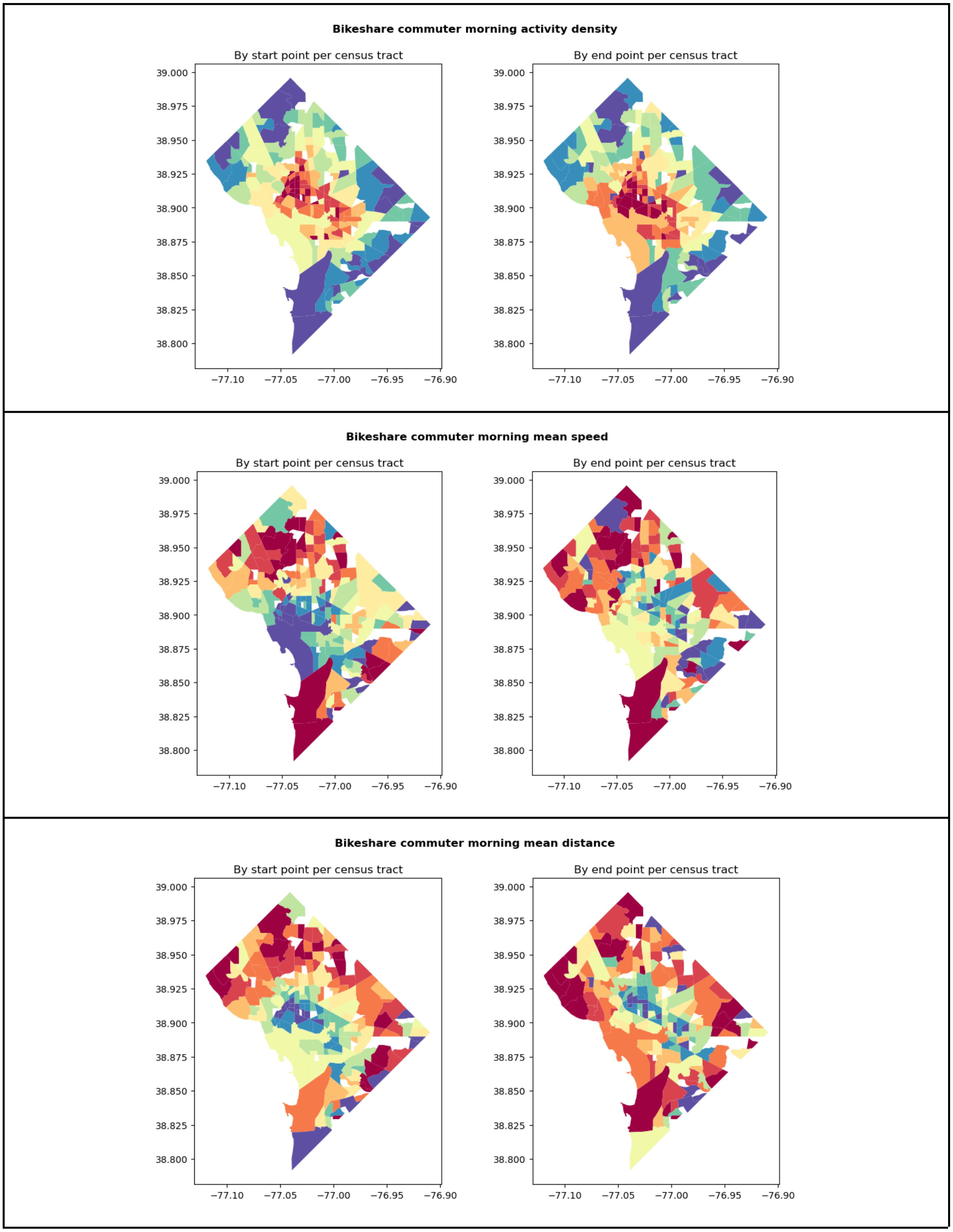


Figure 9- Comparison of the heatmaps of commuter bikeshare activity density, mean speed and mean distance by census tracts (by start/end points) 2020-2022.

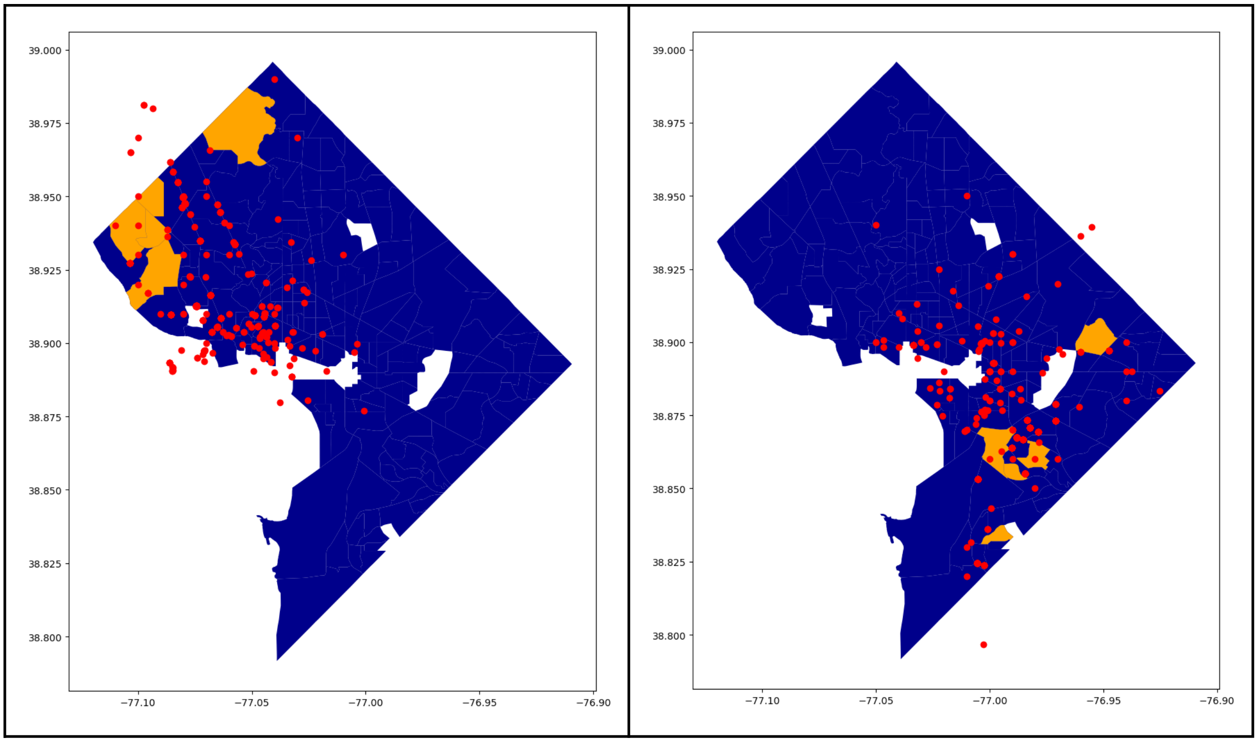


Figure 10 - Where do morning commuters from the five highest and lowest income census tracts go? (left) top 5 highest income commuters' end points (right) bottom 5 lowest income commuters' end points.

Table 1- Hypothesis testing p-value reporting table

H0: The activity density, ride count, speed, distance, duration, of the top 25% high income and bottom 25% of low-income census tracts belong to the same distribution.

	Keep or reject	p-value(t-test)	p-value(k-s test)
activity density	reject	0.0001	1.38 e-06
ride count	reject	2.31 e-05	1.01 e-07
speed	keep	0.75	0.29
distance	keep	0.23	0.79
duration	keep	0.33	0.25

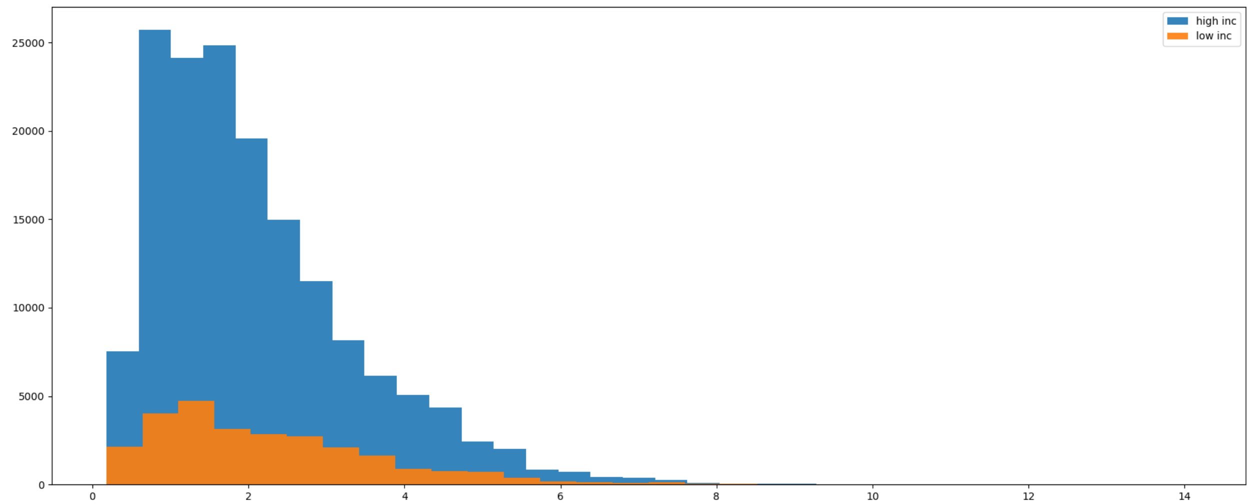


Figure 11 - Histogram plot for total distance produced by the top 25% high income and bottom 25% of low-income census tracts

Table 2- Correlation table for commuter activity density, income, distance, speed, duration, infrastructure and their log values.

	income	infrastructure	speed	distance	activitydensity	duration	log_Income	log_speed	log_distance	log_duration	log_Activity_Density	log_infrastructure
income	1	0.04	0.084	0.034	0.11	-0.027	0.95	0.095	0.035	-0.0085	0.33	0.082
infrastructure	0.04	1	-0.27	-0.27	0.14	-0.22	0.068	-0.27	-0.27	-0.21	0.3	0.92
speed	0.084	-0.27	1	0.61	-0.23	0.27	0.11	1	0.59	0.26	-0.19	-0.24
distance	0.034	-0.27	0.61	1	-0.47	0.91	0.02	0.6	0.98	0.9	-0.43	-0.27
activitydensity	0.11	0.14	-0.23	-0.47	1	-0.47	0.19	-0.22	-0.48	-0.47	0.69	0.24
duration	-0.027	-0.22	0.27	0.91	-0.47	1	-0.069	0.27	0.91	0.98	-0.47	-0.24
log_Income	0.95	0.068	0.11	0.02	0.19	-0.069	1	0.12	0.026	-0.04	0.45	0.13
log_speed	0.095	-0.27	1	0.6	-0.22	0.27	0.12	1	0.58	0.26	-0.17	-0.24
log_distance	0.035	-0.27	0.59	0.98	-0.48	0.91	0.026	0.58	1	0.93	-0.39	-0.27
log_duration	-0.0085	-0.21	0.26	0.9	-0.47	0.98	-0.04	0.26	0.93	1	-0.4	-0.22
log_Activity_Density	0.33	0.3	-0.19	-0.43	0.69	-0.47	0.45	-0.17	-0.39	-0.4	1	0.37
log_infrastructure	0.082	0.92	-0.24	-0.27	0.24	-0.24	0.13	-0.24	-0.27	-0.22	0.37	1

Table 3 - Coefficient and R2 table for linear regression, dependent variable commuter activity density, regressors are income, speed, duration, distance and infrastructure. All values were standardized before analysis.

DV:act. density	coefficient	R2
income	0.0828	0.007
speed	-0.2598	0.067
duration	-0.4049	0.164
distance	-0.3881	0.151
infrastruct.	0.1435	0.021

Table 4 - Coefficient table for multivariate regression of dependent variable commuter activity density, regressors are income, speed, duration, distance and infrastructure. All values were standardized before analysis.

	Activity density	Activity density	Activity density	Activity density	Activity density	Activity density
income	0.0828	0.0832	0.0721	0.0717	0.1010	0.1037
speed		-0.2599				-0.1306
duration			-0.4030			-0.4601
distance				-0.3860		0.0332
infrastruct.					0.1395	0.0095

References

Allison, P.D. (1999). Multiple Regression: A Primer, Pine Forge Press, Thousand Oaks, California.

Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., & Borecki, N. (2013). Are bikeshare users different from regular cyclists? Transportation Research Record: Journal of the Transportation Research Board, 2387(1), 112–119.

Capital Bikeshare Data (2022). [System Data | Capital Bikeshare]

Caspi, O, and Robert B. N, (2019). Bikesharing in Philadelphia: Do Lower-income areas generate trips?, Travel Behaviour and Society, 16 (2019), 143-152. https://doi.org/ 10.7282/t3-sbb4-6y09

Dehdari Ebrahimi Z. , Momenitabar M., Nasri, A.A. , Mattson, J. (2022). Using a GIS-based spatial approach to determine the optimal locations of bikeshare stations: The case of Washington D.C, Transport Policy, Volume 127, Pages 48-60, ISSN 0967-070X, https://doi.org/10.1016/j.tranpol.2022.08.008.

Dill, J, Ma, J McNeil, N, Broach, J MacArthur, J, 2022. Factors influencing bike share among underserved populations: Evidence from three U.S. cities, Transportation Research Part D: Transport and Environment, Volume 112,2022,103471,ISSN 1361-9209, https://doi.org/10.1016/j.trd.2022.103471.

Fishman E. (2016) Bikeshare: A Review of Recent Literature, Transport Reviews, 36:1, 92-113, DOI: 10.1080/01441647.2015.103303

Fishman, E., Washington, S., Haworth, N., & Watson, A. (2015). Factors influencing bike share membership: An analysis of Melbourne and Brisbane. Transportation Research Part A, 71, 17–30

Howland, S, (2017)."Current Efforts to Make Bike Share More Equitable: A Survey of System Owners and Operators" TREC Friday Seminar Series. 107.

Jobe, J., Griffin, G. P. (2021) Bike share responses to COVID-19, Transportation Research Interdisciplinary Perspectives, 10 (2021) 100353, https://doi.org/10.1016/j.trip.2021.100353

Lewis, T. (2011). Has London's cycle hire scheme been a capital idea? The Guardian. Retrieved from http://www.guardian.co.uk/uk/bike-blog/2011/jul/10/boris-bikes-hire-scheme-london?commentpage=all — start-of-comments

McNeil, N., Broach, N.J., Dill, J., (2018). Breaking Barriers to Bike Share: Lessons on Bike Share Equity, ITE Journal,February 18, p. 31-35.

Miller, J. E. (2013). The Chicago Guide to Writing about Multivariate Analysis, Second Edition.The University of Chicago Press, Chicago.

Open Data DC (2022). [Open Data DC]

Shaheen, S., Cohen, A. P., & Martin, E. W. (2013). Public bikesharing in North America: Early operator understanding and emerging trends. Transportation Research Record: Journal of the Transportation Research Board, 2387, 83–92. doi:10.3141/2387-10

Su, L., Yan, X., Zhao, X. (2022). Spatial Equity of Micromobility Systems: A Comparison of Shared E-scooters and Station-based Bikeshare in Washington DC

Toprak, I., Ekdi, F.P. (2022). Social and spatial characteristics of systemic inequality in three US cities: Washington DC, Baltimore and Philadelphia, Conference: 13th International Space Syntax Symposium, Bergen, Norway.

Ursaki and Aultman-Hall, (2015). QUANTIFYING THE EQUITY OF BIKESHARE ACCESS IN US CITIES,University of Vermont Transportation Research Center UVM TRC, Report 15-011 U.

Wang, H., Noland, R.B, (2021) Bikeshare and subway ridership changes during the COVID-19 pandemic in New York City,Transport Policy,Volume 106,2021,Pages 262-270,ISSN 0967-070X, https://doi.org /10.1016/j.tranpol.2021.04.004.

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., & Goodman, A. (2014). Health effects of the London bicycle sharing system: Health impact modelling study. BMJ, 348. doi:10.1136/bmj.g425