

Exploratory Data Analysis: Using NYC MTA Turnstile Data to Reduce Carbon Footprint

Prathap Rajaraman



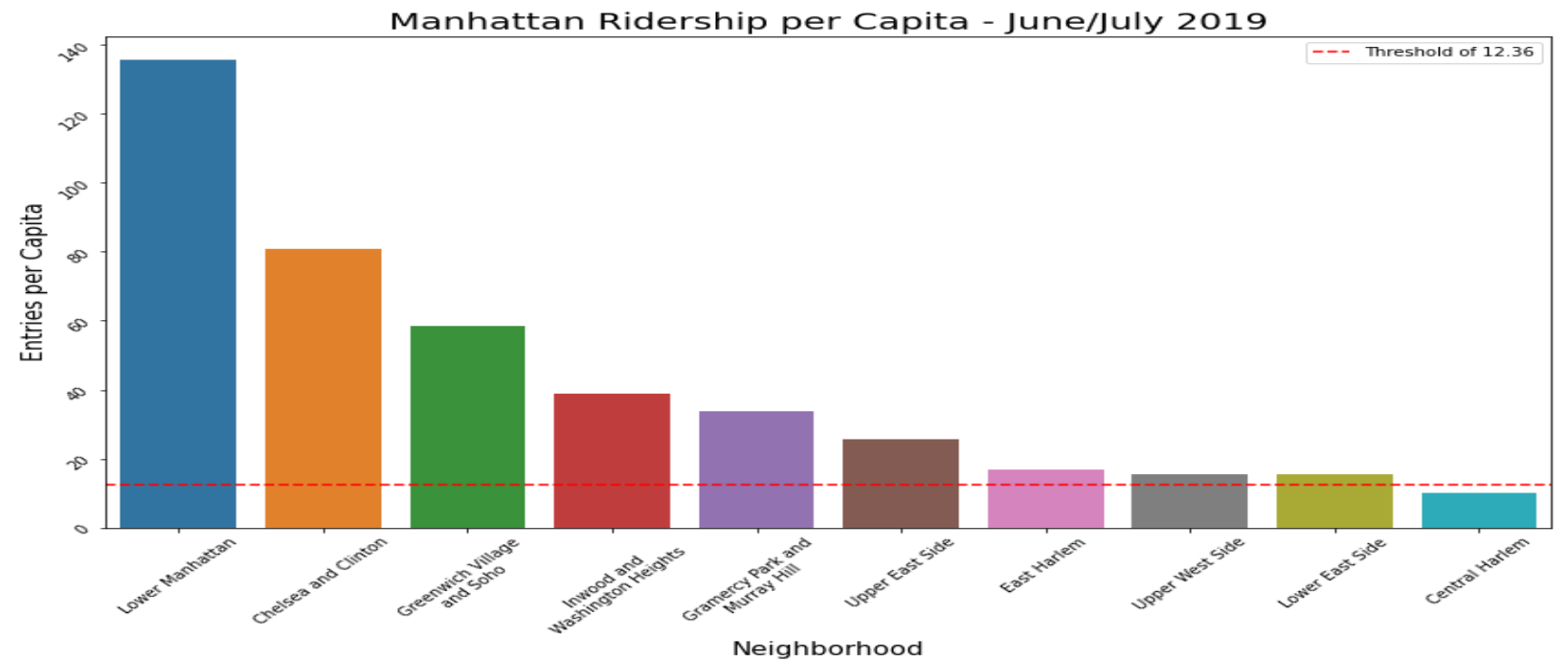
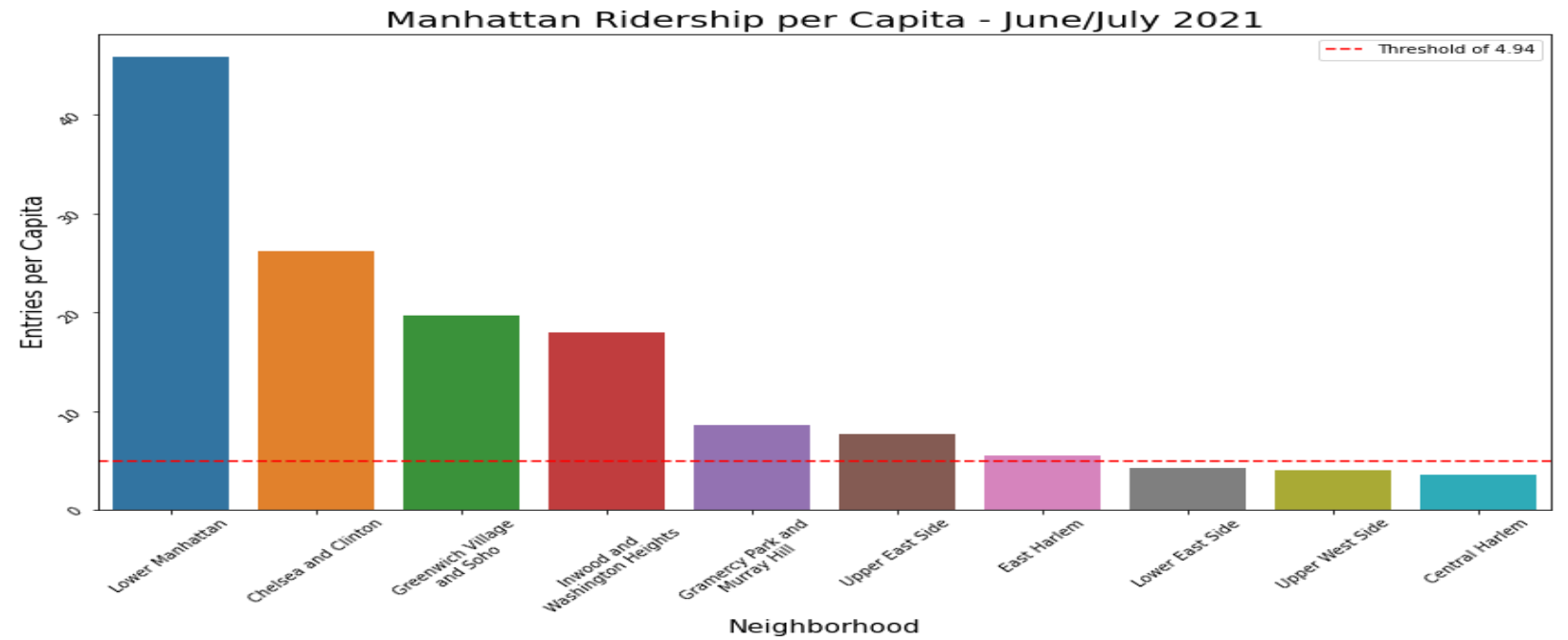
Background

- **Motivation**
 - NY Department of Transportation and Department of Environmental Protection wish to reduce congestion and carbon emissions
 - Only a few neighborhoods have a robust subway system
 - 45% of NYC residents own a car
- **Objective**
 - Explore MTA ridership levels and identify neighborhoods that have a low ridership per capita
 - Use data to determine which regions need further investment in subway lines, tracks, and stations

Methodology

- **Metrics**
 - Total ridership per capita, broken down by neighborhood
 - *All neighborhoods below threshold (4.94 entries per capita in 2021, 12.36 in 2019) are recommended for subsequent investment**
- **Data Sources**
 - MTA turnstile data
 - NYC Population broken down by zip code and neighborhood
- **Data Filtering**
 - June/July 2021 MTA data
 - June/July 2019 MTA data
 - Weekdays before noon
- **Tools used**
 - Pandas for data analysis and manipulation
 - SQL Alchemy for importing MTA data
 - Google API, Geopandas, and Geopy for location mapping
 - Matplotlib, Seaborn, and Plotly for visualization

Results



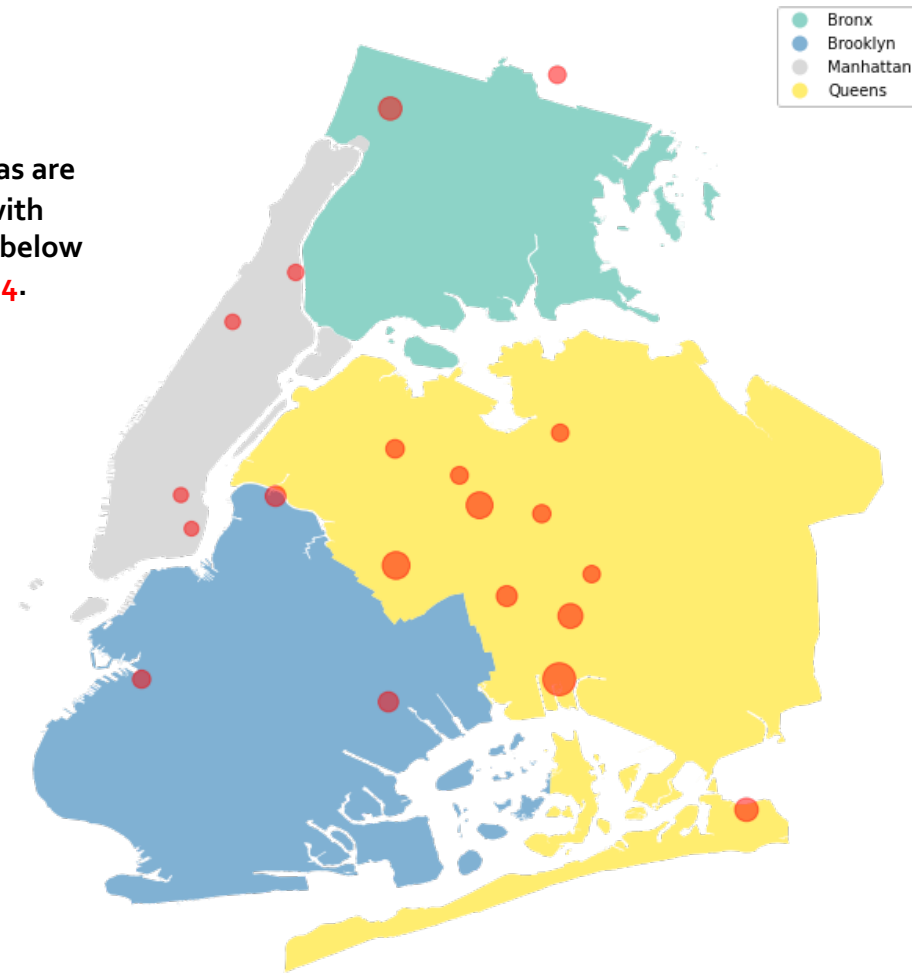
Results



Results

The pinpointed areas are neighborhoods with ridership per capita below threshold of 4.94.

Neighborhoods with Low Ridership per Capita



Conclusion

- 20 neighborhoods in NYC warrant a closer look in terms of infrastructure investment
- Invest in improving subway stations, lines, and tracks in Queens in order to reduce carbon emissions from vehicle usage
- COVID did not impact the trends between neighborhoods

Appendix

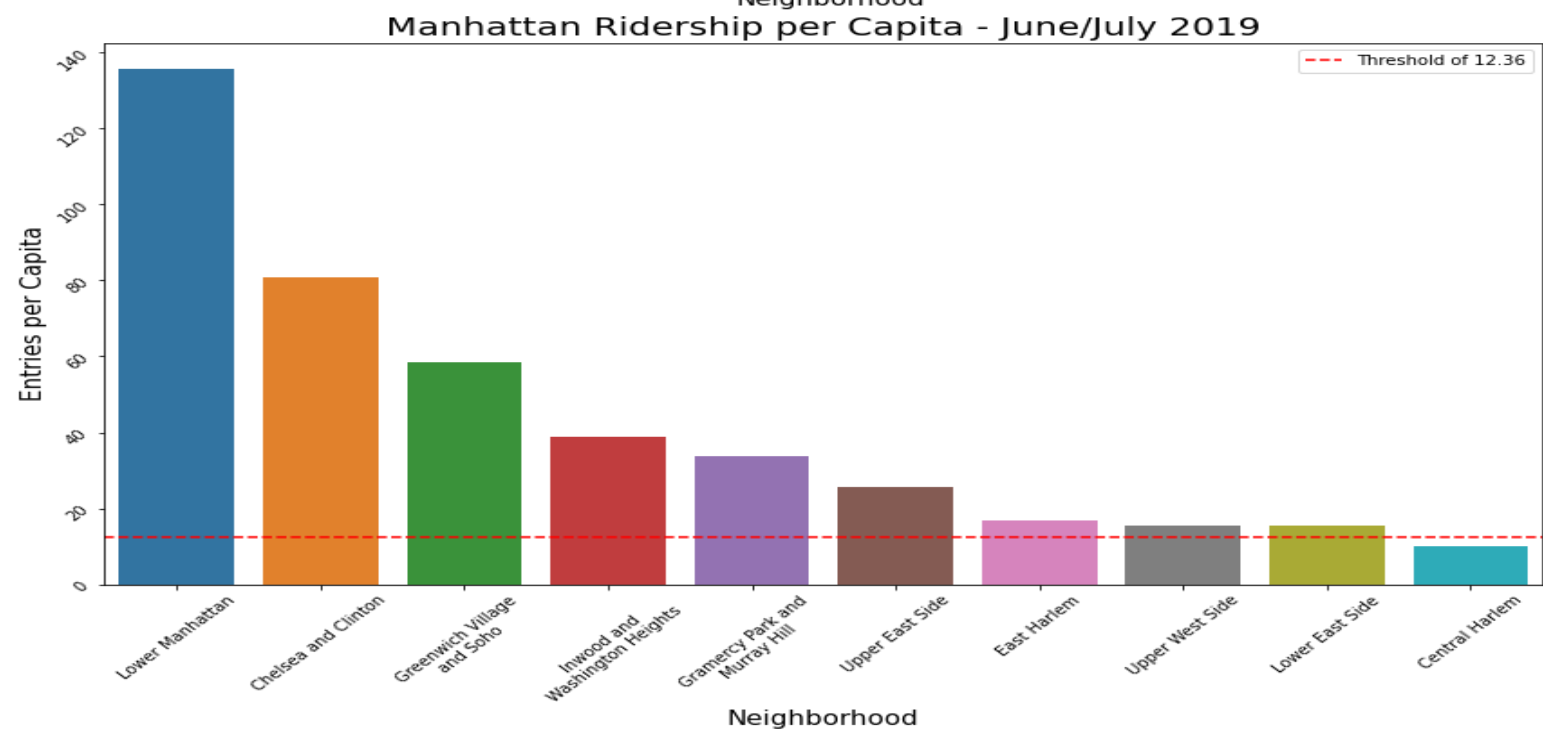
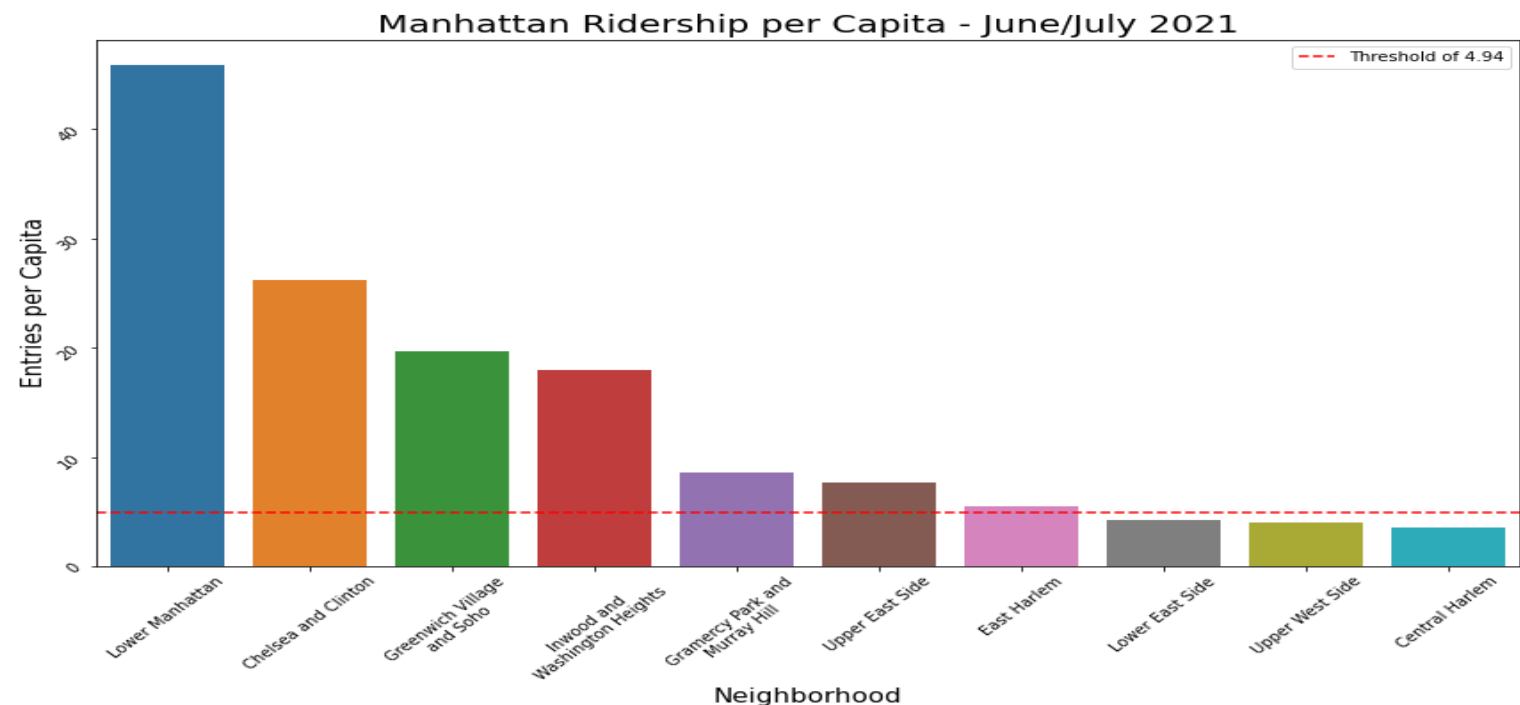
Calculation of Threshold

- The threshold of 4.94 entries per capita (12.36 for 2019) is the product of the following factors:
 - We ideally want NYC residents to use the subway instead of a car to commute to work daily, so we will start with the ideal threshold of 40, which is one subway entry per day for two months worth of business days.
 - 77% of New Yorkers are 18+, or working population that would likely use the subway on business days prior to noon
 - 55% of New Yorkers do not own a car. However, this varies by neighborhood so ideally, we would like to get to the 55% threshold for neighborhoods that fail to meet it.
 - Of these car owners, 73% use their car for daily commuting
 - For 2021 only, a COVID adjustment of 40%. I am assuming that in a post-pandemic world, a hybrid or fully remote work arrangement is embraced.

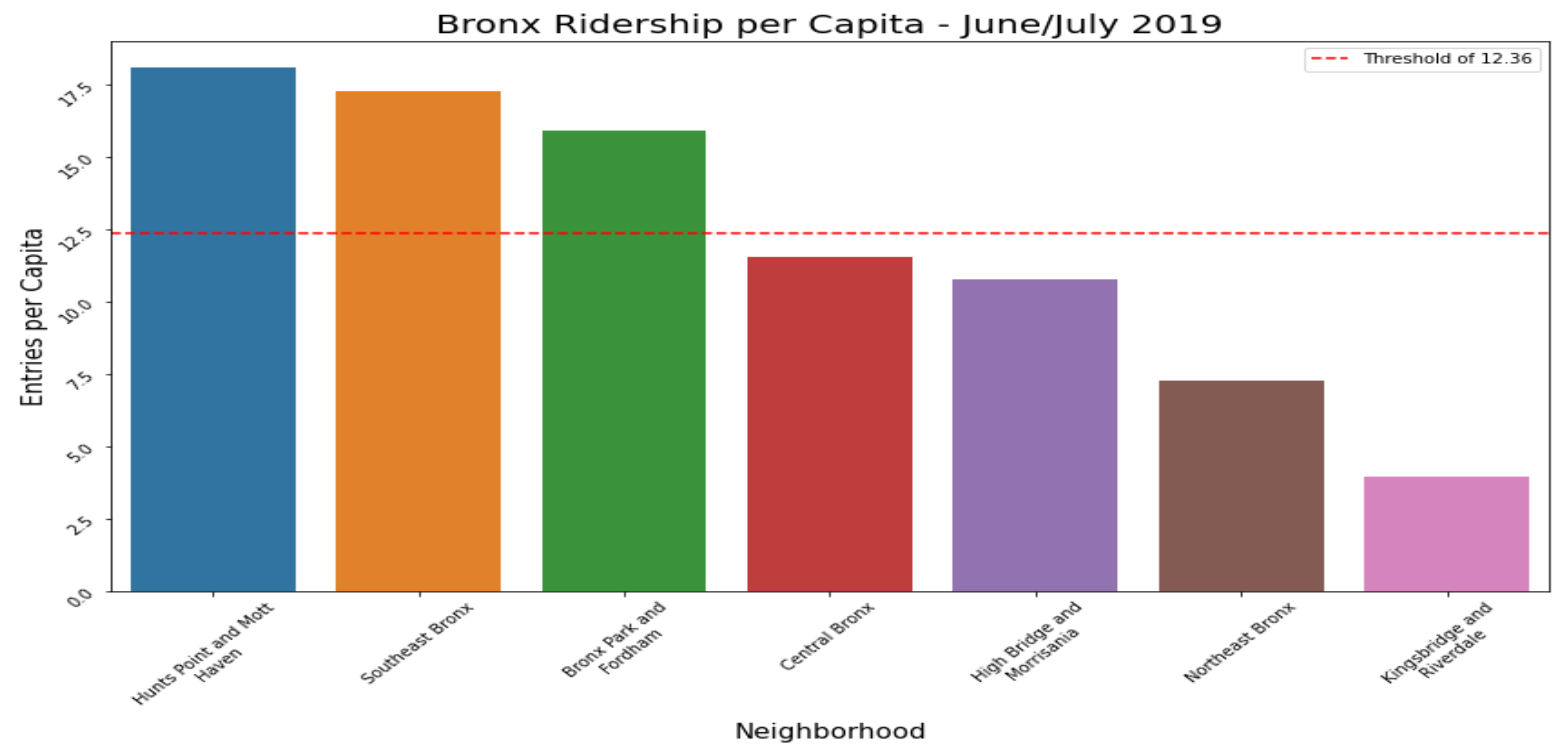
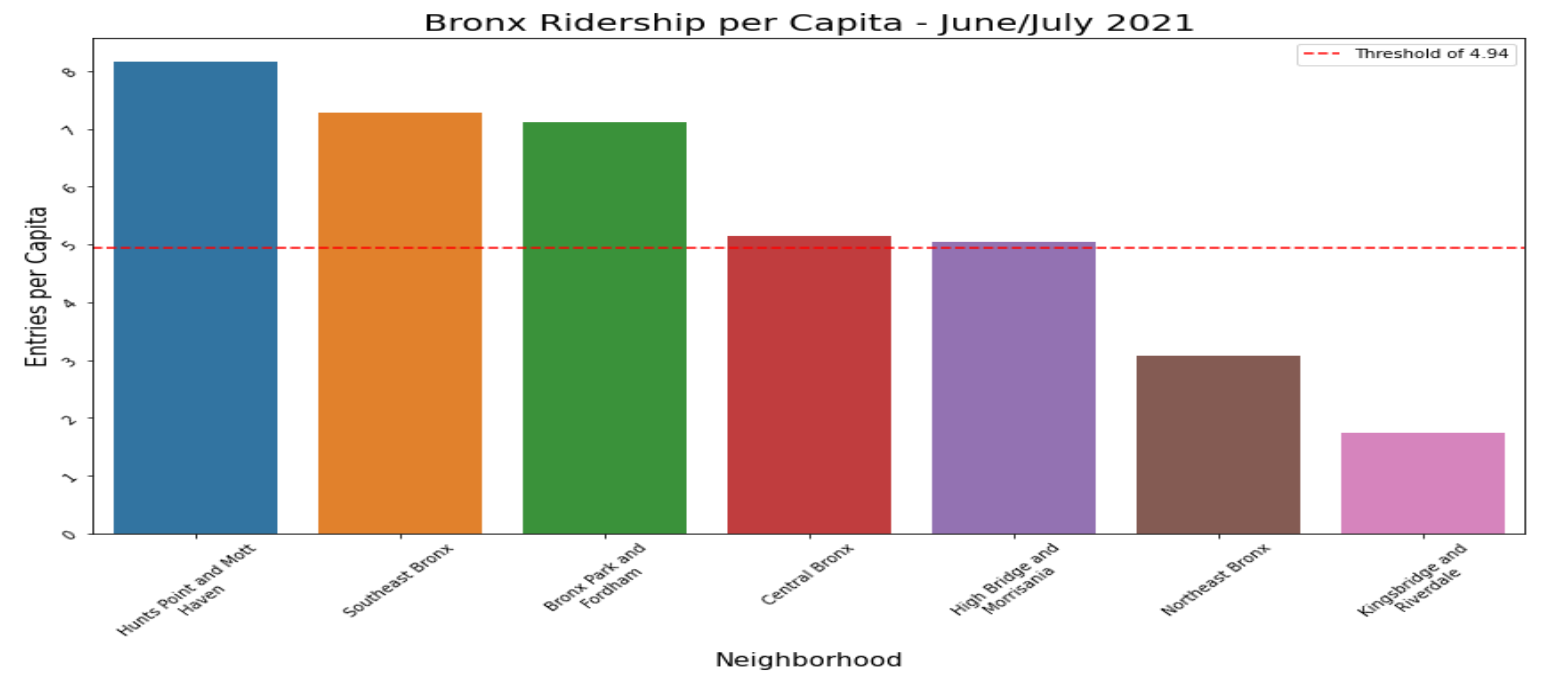
List of
Neighborhoods
that miss
threshold

Borough	Neighborhood
Bronx	Kingsbridge and Riverdale
Bronx	Northeast Bronx
Brooklyn	Canarsie and Flatlands
Brooklyn	Greenpoint
Brooklyn	Southwest Brooklyn
Brooklyn	Sunset Park
Manhattan	Central Harlem
Manhattan	Lower East Side
Manhattan	Upper West Side
Queens	Elmhurst
Queens	Flushing
Queens	Forest Hills
Queens	Howard Beach
Queens	Rego Park
Queens	Richmond Hill
Queens	Ridgewood
Queens	Rockaways
Queens	South Richmond Hill
Queens	Woodhaven
Queens	Woodside

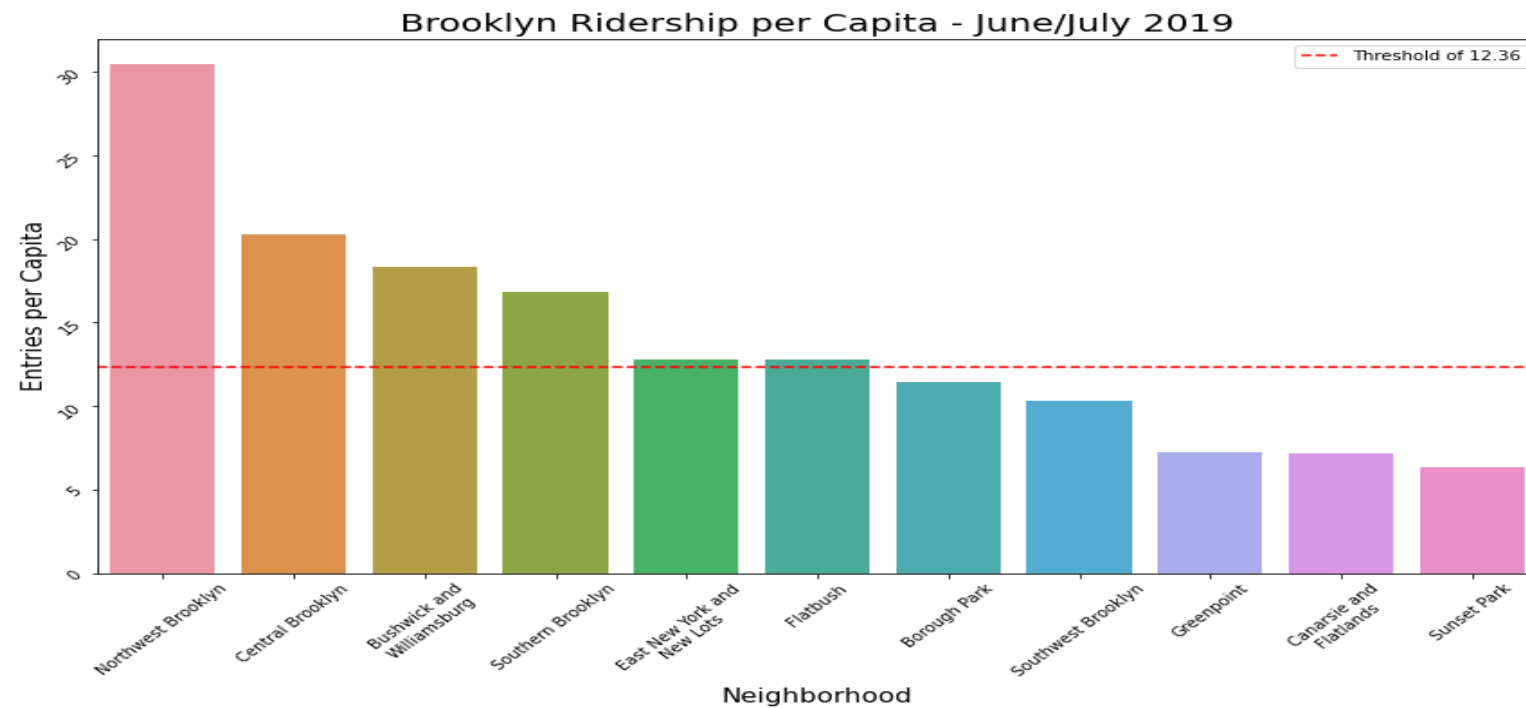
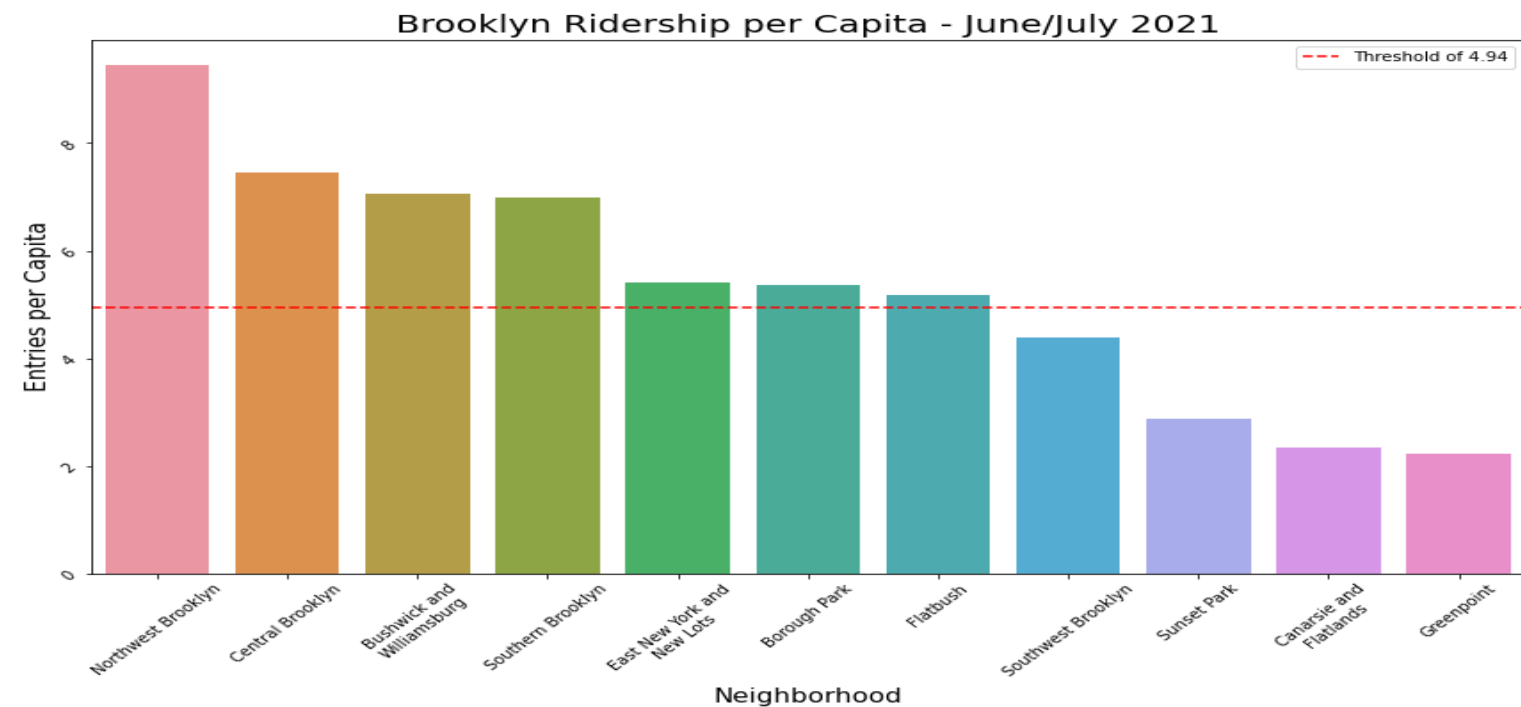
Manhattan



Bronx



Brooklyn



Queens

