S&P Global Assessment

Daniel Anzures Fernández

January 22, 2026

1 Introduction

Urban mobility in crowded cities like New York has been a complex and ongoing challenge, particularly when viewed through the lens of transportation equity. This complexity arises from the asymmetric power held by different stakeholders. Riders, especially those in the outer boroughs, seek affordable, safe, and reliable transportation, while ride-hailing companies such as Uber and Lyft aim to maximize profit and market share. At the same time, city governments must ensure traffic flow and reduce congestion. The tension between these objectives complicates the design of transportation systems that provide fair and equitable access to all users.

The New York City Taxi and Limousine Commission (TLC) publishes detailed trip records for all for-hire vehicles and taxi services operating in the city. The High Volume For-Hire Vehicle (HVFHV) dataset includes trip-level data from ride-hailing companies such as Uber and Lyft, which together represented approximately 85% of all for-hire trips in 2024.

Most existing analyses of this data focus on demand prediction or congestion patterns, while relatively few examine issues of service equity. Simple trip counts by zone can be misleading, as they may reflect either low demand or systematic underservice. Atkinson-Palombo et al. (2019) found that rideshare usage in low-income NYC neighborhoods increased significantly from 2014 to 2017. Taking into account that before 2020 the primary ride-hailing apps Uber and Lyft were heavily subsidized by venture capital, it is worth revisiting service equity among New York's distinct boroughs to determine whether subsidies were masking service inequities that these zones may now experience. Additionally, Paithankar et al. (2025) found that fare levels and wait times are key drivers of demand, further motivating the analysis of service inequity.

This analysis uses machine learning to estimate expected rideshare demand for each zone in New York City and then identifies systematic service gaps and their relationship with income and demographic characteristics using data from the U.S. Census Bureau with the purpose of uncovering service inequity.

2 Exploratory Data Analysis

Thorough data cleaning was performed in the notebook *01_data_cleaning.ipynb*. This document omits those steps; the notebook itself contains the full explanations and thought process, it also includes the spatial join logic used to incorporate U.S. Census data.

For context, I used only data from Yellow/Green Taxis and HVFHV for the months of January, April, July, and October, all for the year 2024. I chose these months to capture as much seasonality as possible whilst minimizing the amount of data, given each month has over 20 million records.

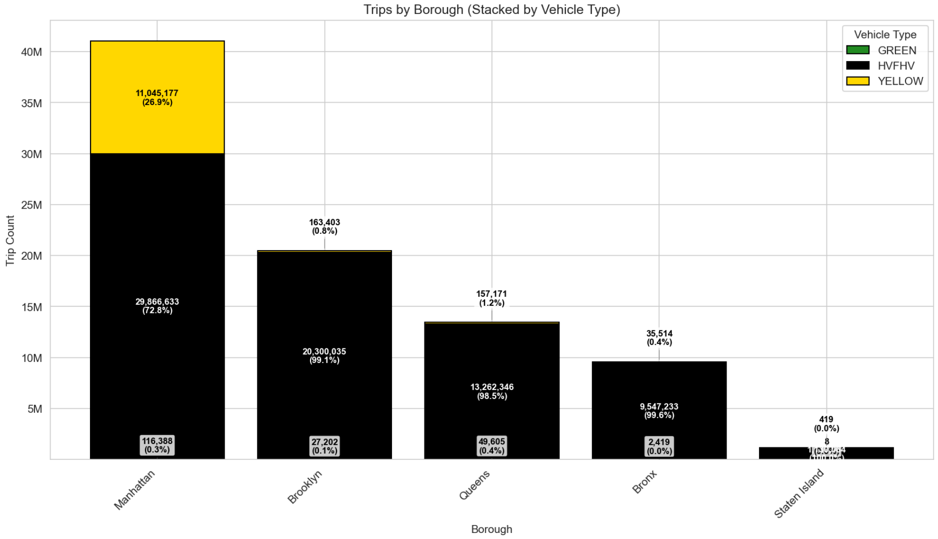

Figure 1: Number of Trips Separated by Borough

Diving into the TLC Trip Record Data, I immediately noticed the overwhelming market share that HVFHV has over New York (86.6% of the total rides made during the sampled months). Green Taxis' trip volume is practically negligible compared to that of both yellow taxis and HVFHV. I considered this sufficient evidence to exclude green taxi data and debated whether to include yellow taxi data. Looking at figure 2, you can clearly identify that yellow taxis operate primarily in inner Manhattan, which is the area with the most traffic flow. Given the scope of the analysis is to find any service inequities in the outer boroughs, yellow taxis clearly do not play an important role in them, in contrast with HVFHV. For that reason, I decided to also exclude yellow taxi trip records from the analysis, so from now on, every visualization and conclusion will be regarding only HVFHV data.

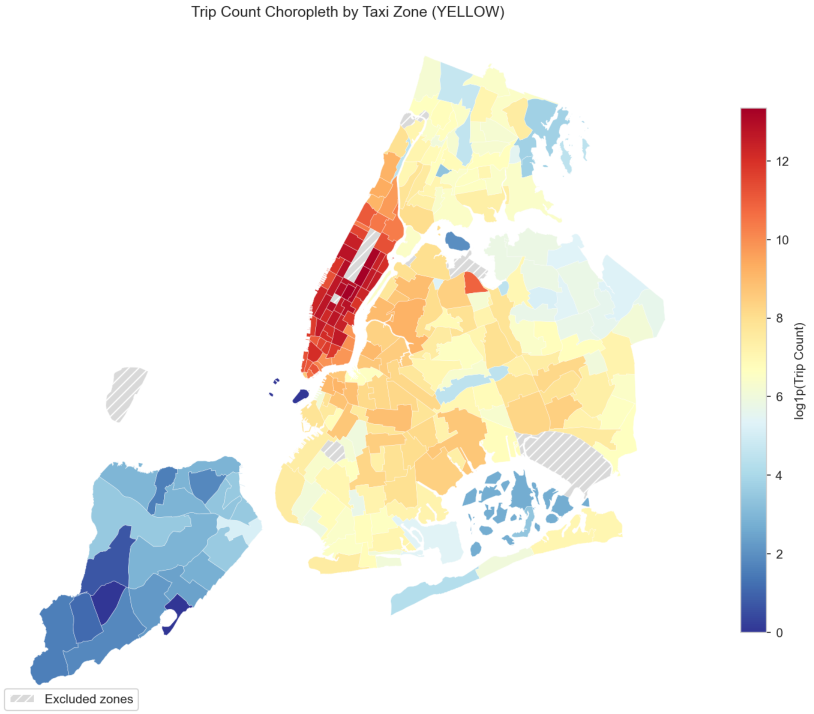

Figure 2: Choropleth For Yellow Taxi Number of Trips

Further exploring the data, I looked for clear evidence that would strengthen my hypothesis that lower-income or minority-populated zones received worse service on average when compared to the average zone. First, I calculated how many dollars per mile each zone paid on average, visualized in figure 3, and found opposing evidence for my hypothesis: lower-income zones actually pay fewer dollars per mile compared to other zones.

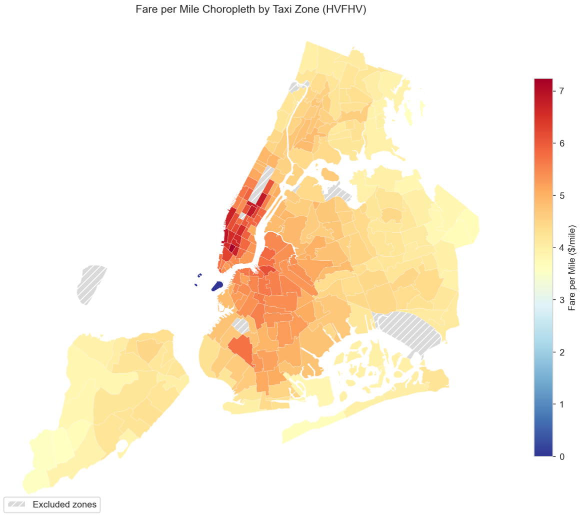

Figure 3: Dollar per Mile Choropleth

The notebook *02_data_exploration.ipynb* contains a more in-depth analysis of the findings I made during exploratory data analysis. At the end, I plotted a pair plot with a regression line and calculated the R^2 coefficient for all relationships.

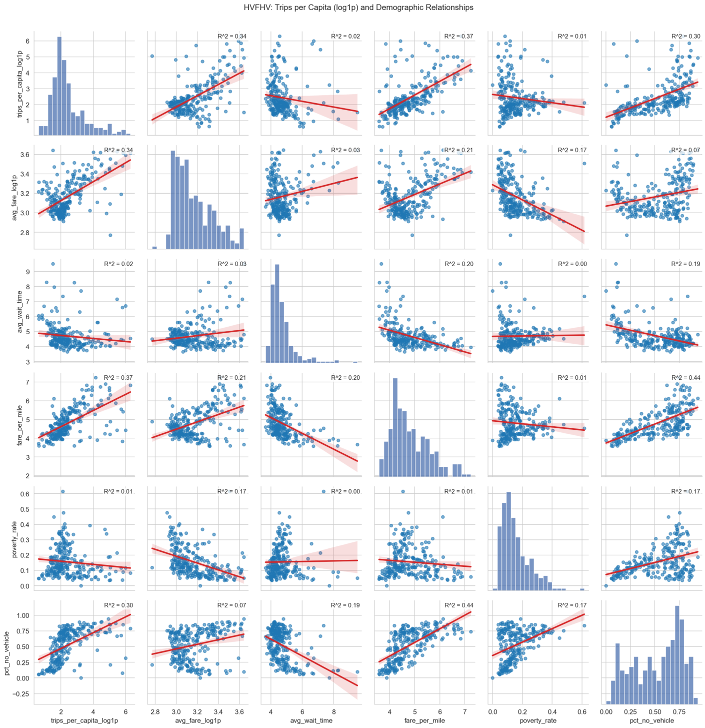

Figure 4: Pair Plot For Variables of Interest

Hoping to find some moderate to significant correlations between variables such as average wait time, trips per capita, and dollars per mile against poverty rate, the only moderate relationships I found ($R^2 \geq 0.15$) were poverty rate with dollars per mile, which had a negative slope consistent with my previous visual finding, and poverty rate with percentage of households with no vehicle, which makes sense but is in no way related to a service equity gap.

3 Model Pipelines

After finishing exploratory data analysis and finding no evidence supporting my hypothesis, I decided to proceed with the main test: fitting a machine learning model on the data I had gathered from the TLC Trip Record Data and the U.S. Census Bureau and computing the residuals to find any systematic patterns that could hint at service inequity. I decided to use two models: linear regression as a baseline, and given I did not find any convincing linear relationships, I also included a random forest algorithm to capture any non-linear relationships.

In *03_ml_underservice_detection.ipynb*, you can find a more thorough explanation of my thought process on feature engineering. I decided to use the following features:

- Percentage of flexible commuters (those likely to use a rideshare app)
- Labor force participation rate
- Percentage of households with no vehicle access
- Percentage of workers who commute via public transit
- Log1p transformation of total population

The variables I chose to explore to find equity gaps were the following:

- Poverty rate
- Percentage of low-income households
- Percentage of minorities

I decided to use trips per capita as my target variable. My logic was that if residuals correlate negatively with equity variables—meaning residuals decrease as poverty or minority percentage increases—this would imply that lower-income areas are underserved.

Afterwards, I performed a cross-validated grid search with 5 folds over a conservative parameter grid for both linear regression and random forest. For linear regression, I ultimately decided to exclude Lasso Regression from my parameter grid to avoid coefficients shrinking to zero and to improve interpretability of the results.

Finally, I calculated and plotted the scatter plots for the residuals for both Ridge Regression and Random Forest, which are shown in figure 5. Again, reiterating my findings from the exploratory analysis, I found no significant correlation between residuals and equity variables.

4 Conclusion

I found no evidence suggesting systematic underservice in lower-income areas. Contrary to concerns raised by ridesharing apps no longer being subsidized, we find no evidence of systematic rideshare underservice in lower-income NYC neighborhoods as of 2024. Service levels are well-predicted by vehicle ownership and transit access, with no residual correlation with poverty or race. While this solution does not prove inequity, it provides a service equity framework that demonstrates the system is functioning equitably as of 2024 for riders.

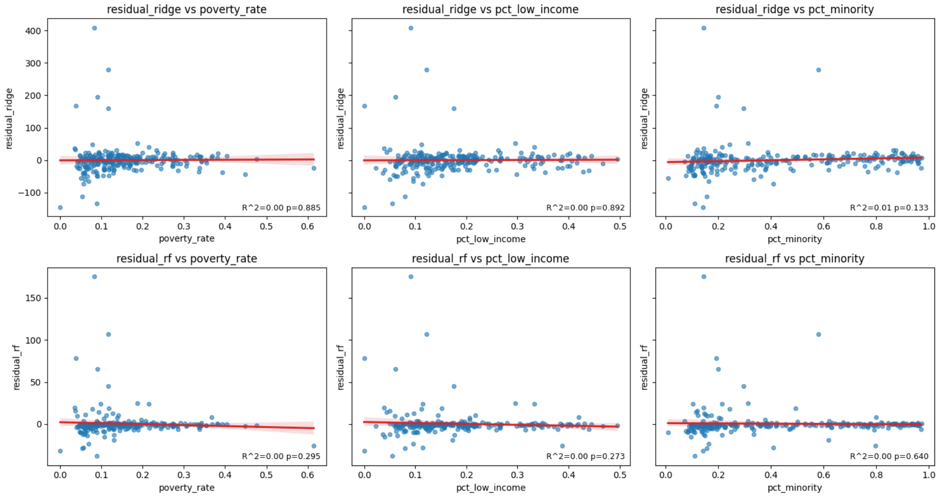

Figure 5: Residual Analysis

5 How to Leverage Generative AI to Develop an Even Better Solution

The framework we built works for a single point-in-time analysis (in this case, four months of 2024), but equity patterns can shift as fares change, new regulations take effect, or service areas expand. Manually re-running this analysis every month is not realistic for most organizations.

A generative AI system could operationalize this framework in several ways:

Automated reporting. When TLC releases new trip data each month, an LLM could ingest it, run the demand model, compute residuals, and generate a plain-language summary highlighting any zones where service levels dropped unexpectedly, hinting at service inequity. Instead of a data scientist spending a week on this, stakeholders get a monthly equity report in their inbox.

Early warning system. The model could flag zones where residuals trend negative over consecutive months, catching emerging service gaps before they become entrenched. A simple alert like ”Hunts Point has shown below-expected demand for three straight months despite stable population” gives regulators a head start on intervention.

Accessible querying. Not everyone at TLC or community advocacy organizations knows Python. An LLM interface could let non-technical users ask questions in plain English: ”Which Bronx neighborhoods have the longest wait times?” or ”How does service in East New York compare to Park Slope?” The model translates these into queries against the underlying data and returns interpretable answers.

None of this replaces the need for careful methodology—the AI would be running the same statistical framework we developed, just faster and more accessibly. The goal is to lower the barrier so equity monitoring actually happens on an ongoing basis, rather than as a one-off study.

References

Atkinson-Palombo, C., Varone, L., and Garrick, N. W. (2019). Understanding the surprising and oversized use of ridesourcing services in poor neighborhoods in new york city. *Transportation Research Record*, 2673(11):185–194.

Paithankar, P., Kockelman, K. M., and Gurumurthy, K. M. (2025). Ride-hailing fares and demand interactions: Insights from market analysis over space and time. Presented at Bridging Transportation Research Conference 2025, accepted for presentation at Transportation Research Board Annual Meeting (January 2026), under review for publication in Research in Transportation Economics.