

Bus Transit Performance
Team D

Xavier Thomas, 2025, xthomas@bu.edu
Ketan Suhaas Saichandran, 2025, ketanss@bu.edu
James Xiao, 2025, jamxiao@bu.edu
Haoxiang Huo, 2024, stevehuo@bu.edu
Rishven K Pravin, 2025, rishven@bu.edu

Introduction

Problem Statement

Public transportation is a pivotal element in shaping the daily experiences and quality of life for residents across Massachusetts and the Greater Boston area. Yet, the assurance of equal and fair access to quality service across all reachable areas remains a question that demands investigation. Given its profound impact on the day-to-day lives of residents, it becomes imperative to quantify the equity and fairness embedded in Boston's public transportation system and to discern the varying perceptions of service quality among different neighborhoods.

Our analysis leverages a combination of data science methodologies, encompassing the extraction and examination of public transportation data, and demographic information. We explored key questions such as: How does the quality of public transportation services vary across different neighborhoods? Are there discernible disparities in service frequency, reliability, and accessibility? How do factors such as income levels, population density, and geographic location correlate with the perceived quality of public transportation?

Our analysis provides a comprehensive overview of the existing state of public transportation equity. The goal is to pinpoint areas where improvements can be made, ensuring that the benefits of an efficient and reliable transportation network are shared equitably among diverse communities.

Extension: Understand how Demographic Factors play a role in Bus Ridership/Accessibility.

Data Collection and Cleaning

Collected Data from the <https://mbta-massdot.opendata.arcgis.com/> website.

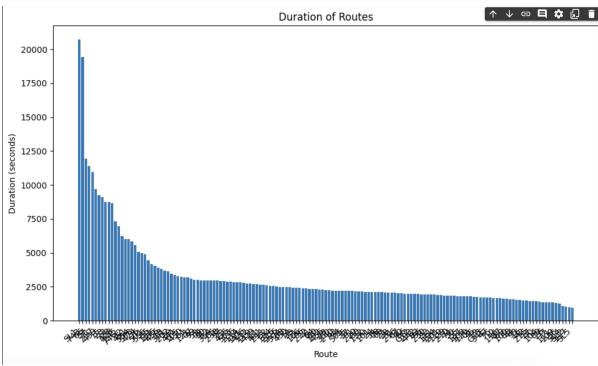
The following Datasets were used for analysis:

1. PATI Bus Stops
2. MBTA Bus Arrival Departure Times
3. Bus Reliability
4. Bus Ridership by Time Period, Season, Route Line, and Stop
5. Wheelchair/Accessibility
6. Boston Census Dataset

In addition, data from the <https://api-v3.mbta.com/> api was utilized to obtain bus level information such as "vehicle_number", "route_id" and the bus stops each particular bus would visit along with their coordinates.

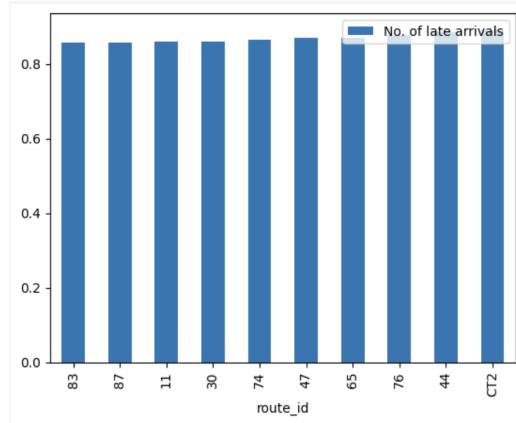
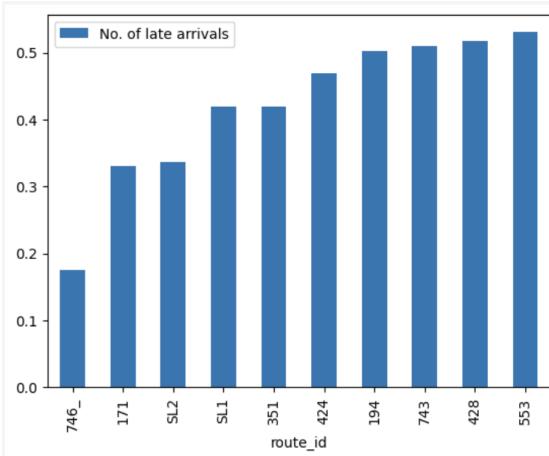
Base Analysis

1. End-to-end travel times for different bus routes

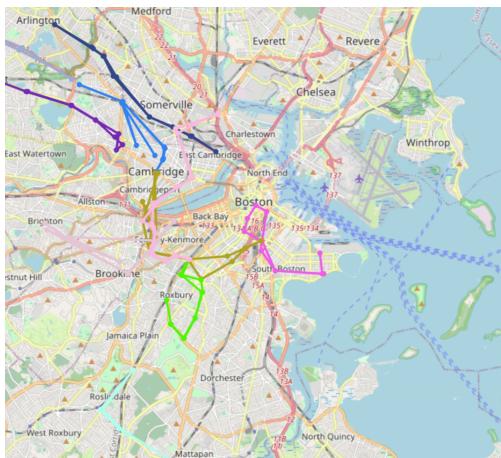


We notice that a majority of the lines have end-to-end travel times of < 5000 seconds

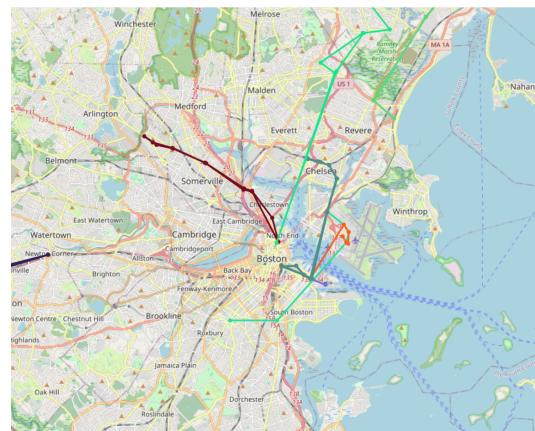
2. Disparities in the service levels of different routes



Top 10 least late routes



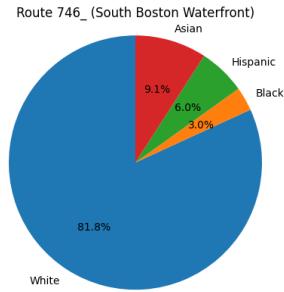
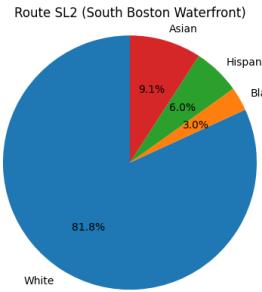
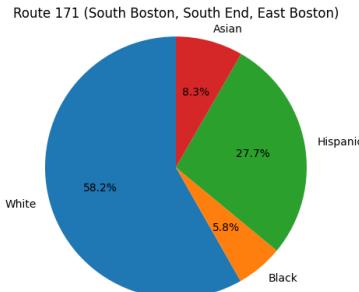
Top 10 most late routes.



The map shows the 10 most late routes. Somerville and Cambridge have higher percentages of non-minorities than Jamaica Plain, Roxbury, Dorchester, and Chinatown

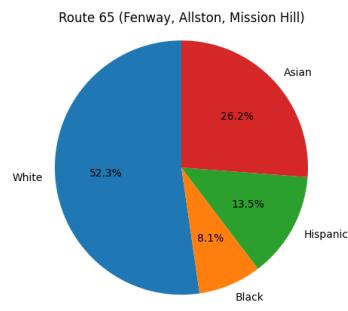
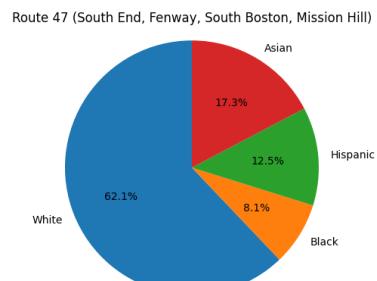
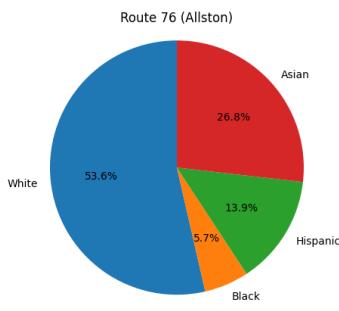
The map shows the 10 least late routes.

3. Disparity in quality of service by race



Top 3 Least Late Routes

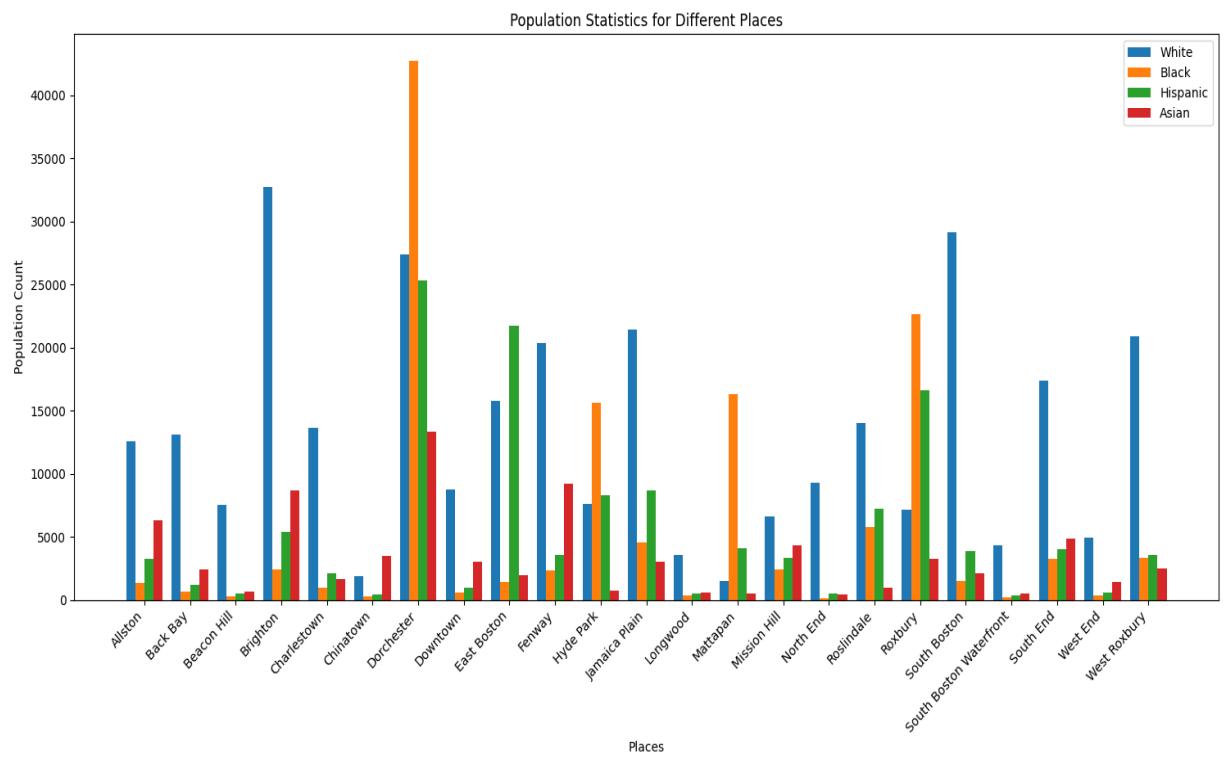
- We noticed that there was a higher percentage of white population in routes that were on time.



Top 3 Most Late Routes

- We notice that among the routes that are the most late, there is a higher proportion of the non-white population.
- We analyze more demographic factors contributing to disparity in quality of service in our extension analysis below

4. Wheelchair accessibility against Population Demographic



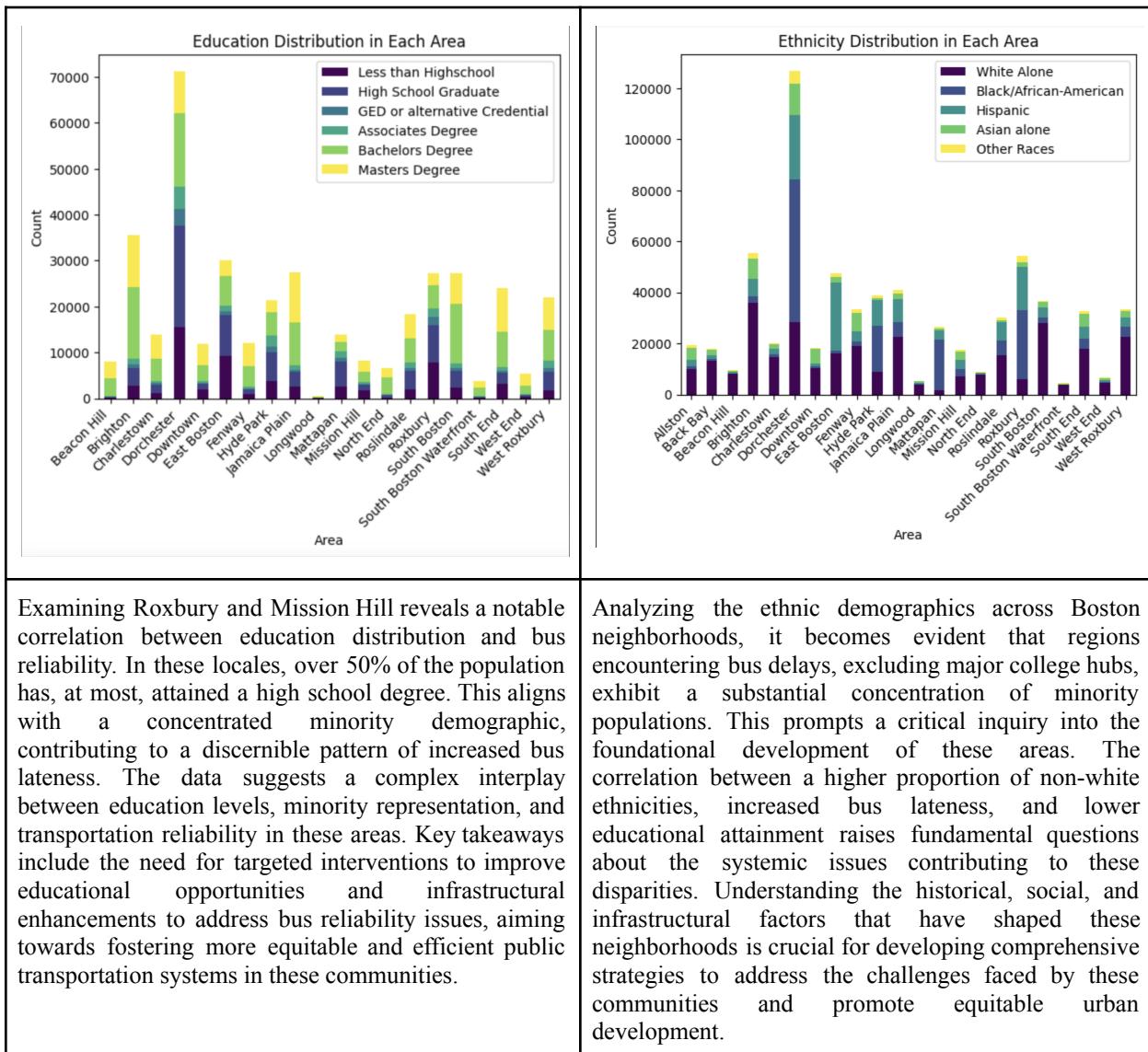
Neighborhoods sorted based on wheelchair accessibility:

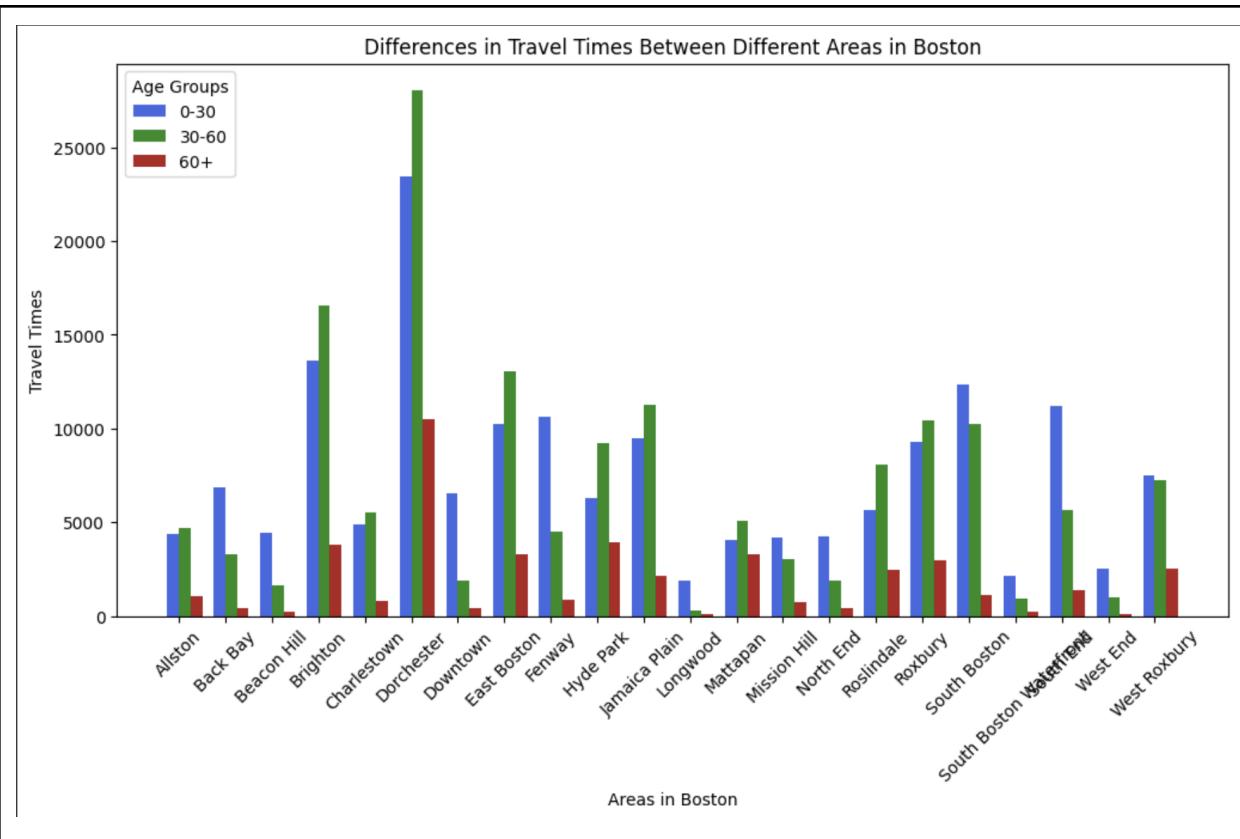
Charlestown, East Boston, Downtown, Brighton, Hyde Park, Allston, Dorchester, Mattapan, Back Bay, West End, Fenway, West Roxbury, South End, Jamaica Plain, Chinatown, South Boston, South Boston Waterfront, Mission Hill, North End, Longwood, Beacon Hill, Roxbury, Roslindale

Extension Analysis

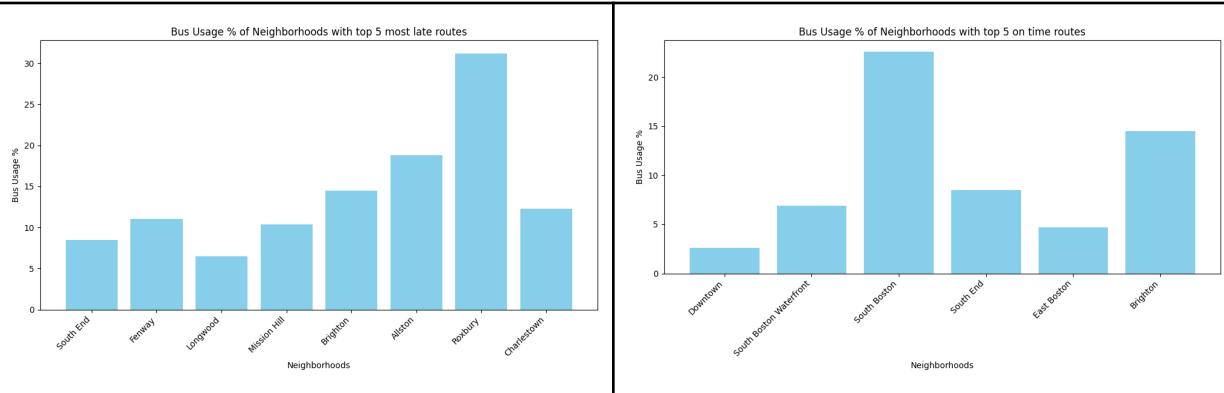
Demographic/Census Stat Plots:

We've identified Roxbury, Mission Hill, Fenway, Cambridge, Chinatown, and Allston as major "unreliable" bus performance areas. Let's take a look at some demographic statistics of Boston areas, and see if we can discover any correlations between demographic features and bus performance.

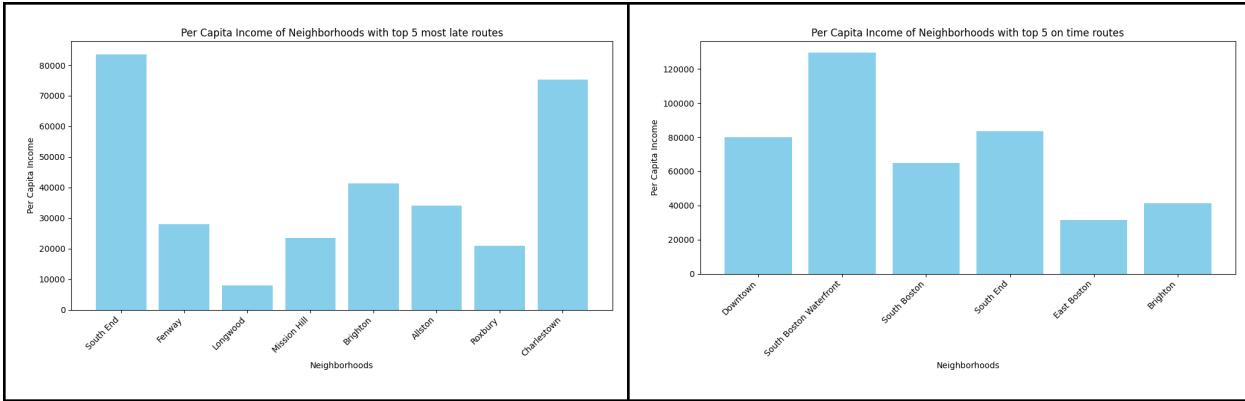




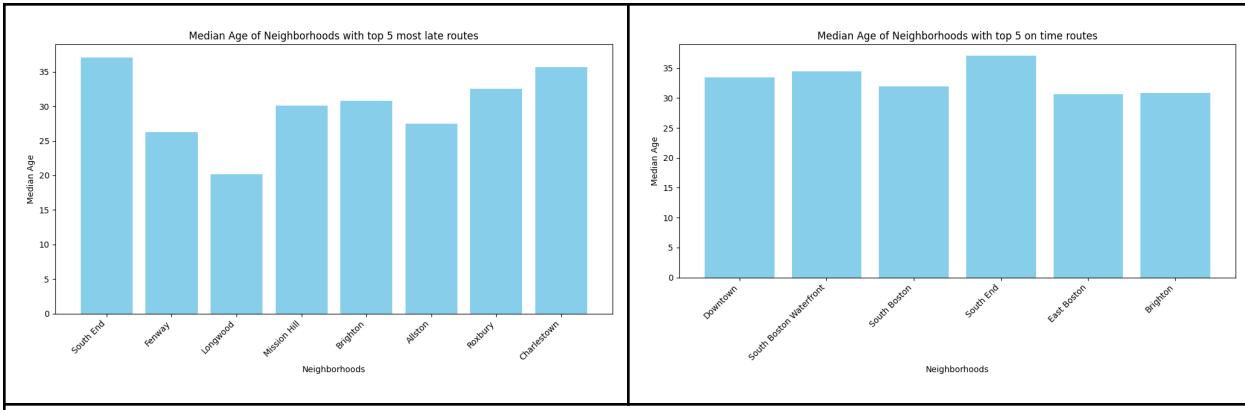
Let's examine the travel times for work (in minutes) of people in different areas. The observation of poorly-serviced neighborhoods having, on average, lower travel times as compared to other neighborhoods raises questions about the fairness of bus service as they experience more lateness than neighborhoods that have a higher travel time. See Dorchester, Allston, Mission Hill, and Roxbury.



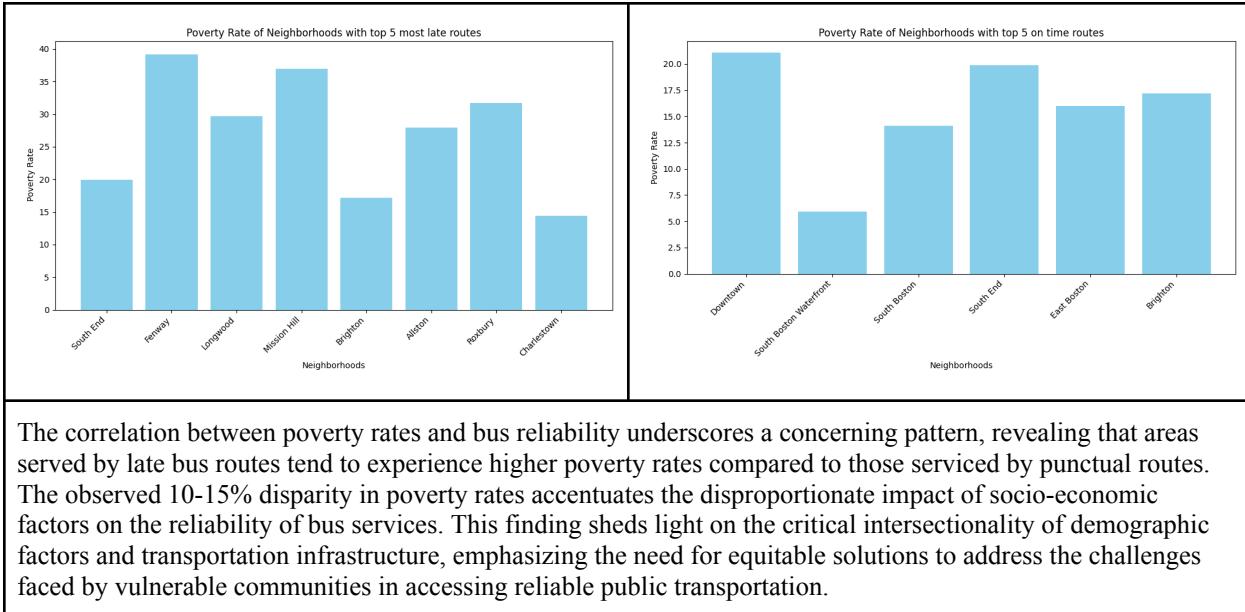
Roxbury emerges as a neighborhood with a significantly high bus usage despite experiencing several late bus routes. This observation points to a critical need for infrastructure enhancement within this area. On the other end of the spectrum, downtown stands out for having consistently punctual bus services, yet the percentage of the population utilizing buses remains notably low. There might be opportunities to reassess resources or redirect efforts toward other areas with more pressing transportation needs.



The results revealed a thought-provoking trend: neighborhoods characterized by higher wealth and affluence tend to experience faster bus services, while areas with lower income levels tend to have slower bus services. However, it should've been the other way as people with lower incomes tend to use the bus services mostly. This observation underscores a potential socio-economic aspect influencing the quality and efficiency of public transportation within our city.



We attempted to examine to see if age was a meaningful factor when it came to the relationship between demographic status and bus reliability. Evidently, the median age in late and timely neighborhoods was very similar. Thus, no strong correlations are concluded.



Contributions

Xavier Thomas - Data Collection, Data Preprocessing, Base Questions Visualizations, Extension Project Visualizations, Code for Base Questions and Extension Project.

Ketan Suhaas Saichandran - Data Collection, Data Preprocessing, Base Questions Visualizations, Extension Project Visualizations, Code for Base Questions and Extension Project.

James Xiao - Data Collection, Extension Project Visualizations, Extension Project Analysis

Haoxiang Huo - Data Collection, Data Preprocessing, Extension Project Visualizations, Extension Project Analysis

Rishven K Pravin - Data Collection, Data Preprocessing, Extension Project Analysis