CS506 Bus Transit Performance Team D -

James Xiao, Xavier Thomas, Haoxiang Huo, Rishven Pravin, Ketan Suhaas Saichandran

Problem Motivation

Public transportation plays a crucial role in the quality of life for residents in Massachusetts and the Greater Boston area. However, how confident can we be that the quality of service is equal for all reachable areas? With such a substantial impact on resident day-to-day life, it is necessary to quantify how "fair" Boston's public transportation is, and which areas perceive better/worse quality of service.

Project Goal

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, racial demographic, 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.

Data Collection and Cleaning

Collected Data from the https://mbta-massdot.opendata.arcgis.com/ website. The following 5 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

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.

Division of Work

- Rishven and Haoxiang - Data preprocessing, base questions 1, 2
- Xavier and Ketan - Data collection, cleaning, preprocessing, base questions 1, 2, 3, 5
- James - Data preprocessing, extension project, presentation and slides

Base Questions:
1. What are the end-to-end travel times for different bus routes
2. Are there disparities in the service levels of different routes? (which lines are late more often than others)
3. What are the population sizes and characteristics of the communities serviced by different bus routes (e.g. race, ethnicity, age, people with disabilities/ vulnerabilities)?
4. If there are service level disparities, are there differences in the characteristics of the people most impacted?
5. Which routes are better/worse, differences in quality of service by class/race, contributing variables, ect.

Population Statistics for Different Places
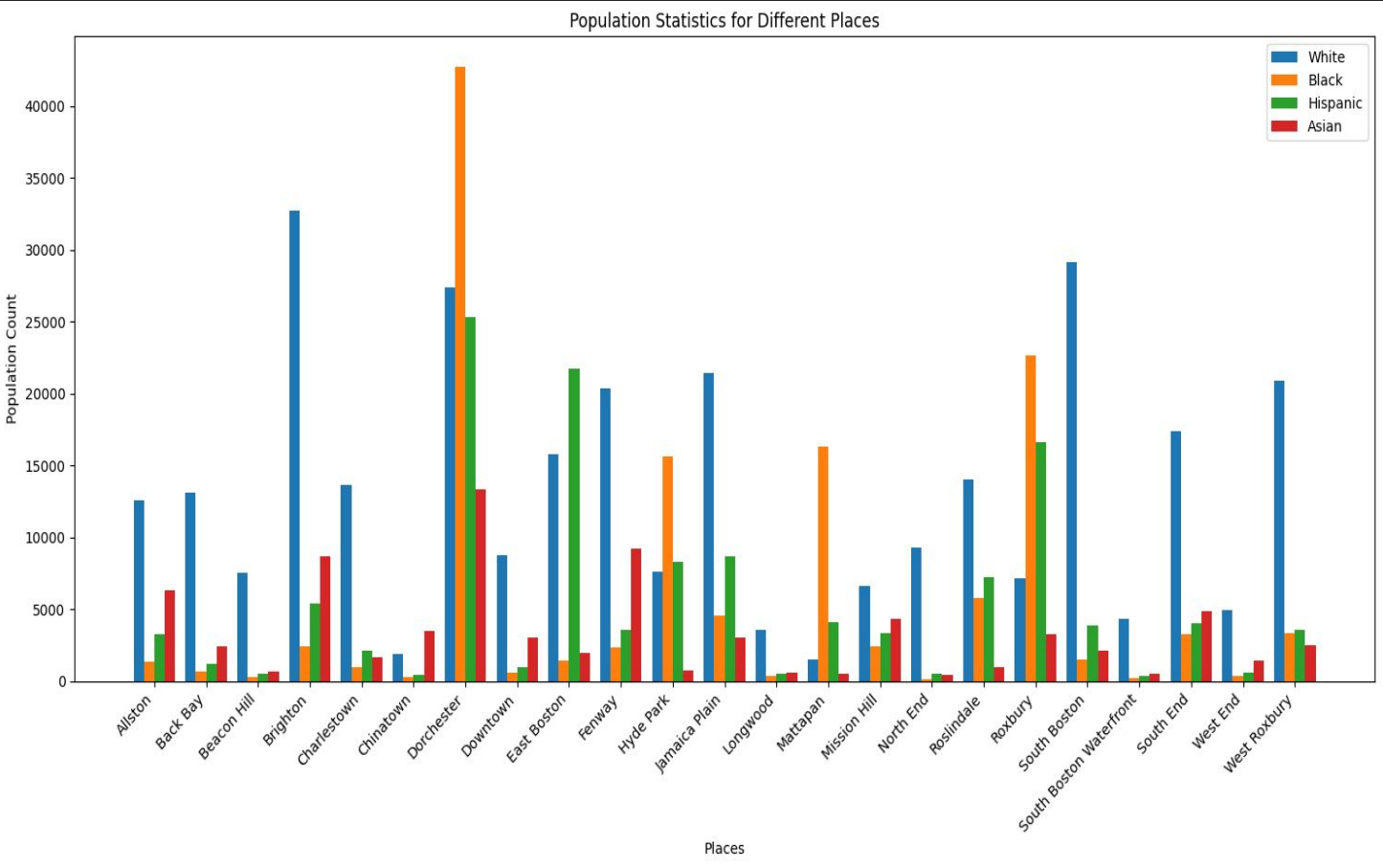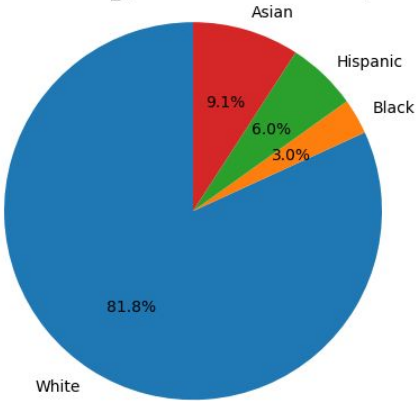
Neighborhoods sorted based on wheelchair accessibility:
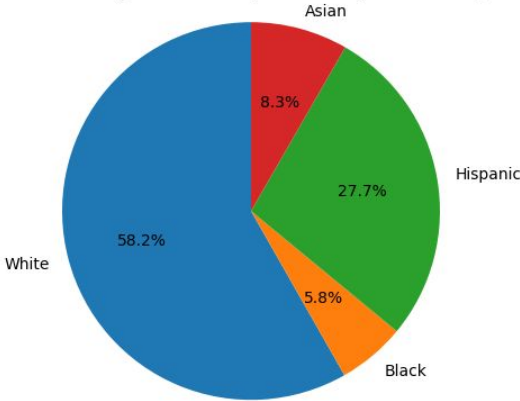
Allston
East Boston
Mattapan
West Roxbury
West End
Beacon Hill
Back Bay
South Boston
Waterfront
North End
Fenway
South Boston
South End
Jamaica Plain
Mission Hill
Roslindale

Top 3 Least Late Routes

Top 3 Most Late Routes
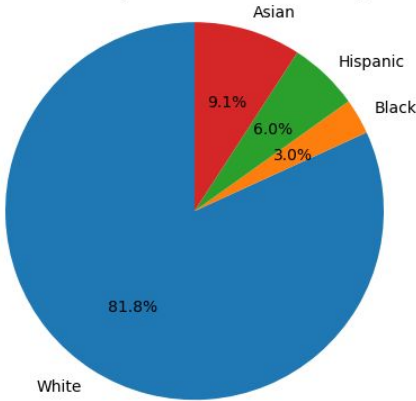
Number of bus stops against racial demographic
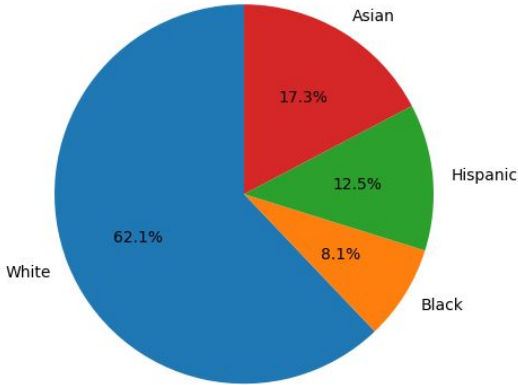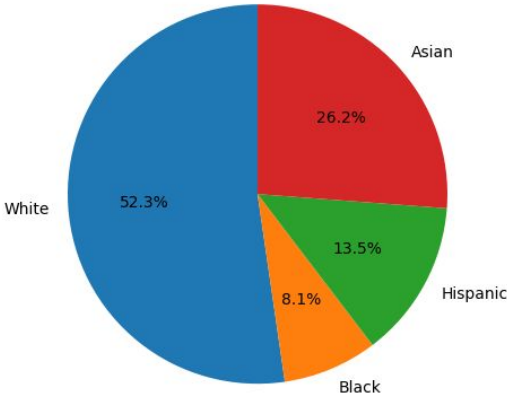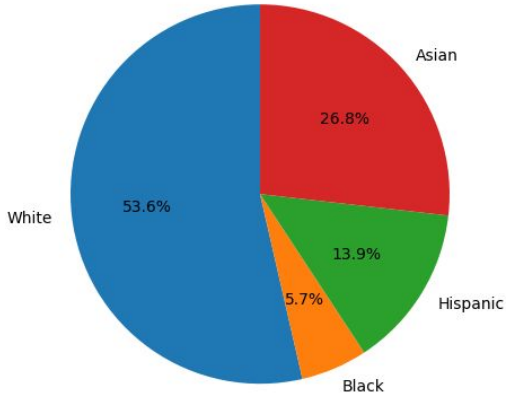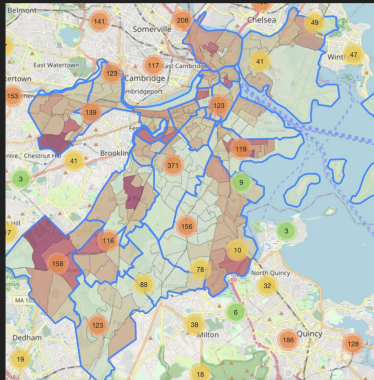
WHITE POPULATION: Roxbury, with 60,705 residents, has a diverse population with a median age of 34, primarily relying on buses for transportation, while Jamaica Plain, home to 41,112 residents, boasts a higher median income of $106,153, with a majority holding college degrees and utilizing various modes of transportation including cars, walking, biking, and buses.
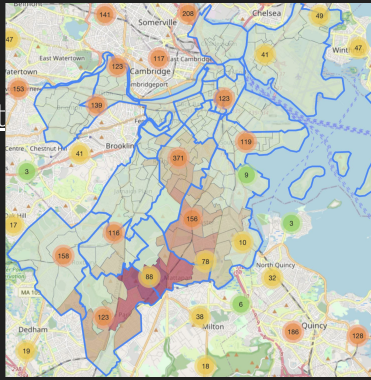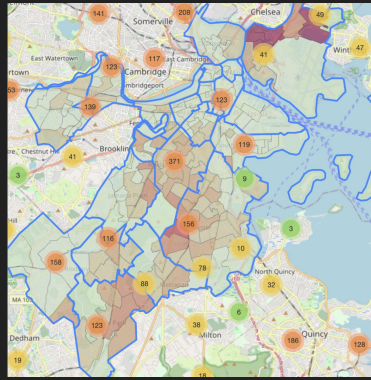
BLACK POPULATION: Hyde Park, home to 38,402 residents, features a diverse population with a median age of 37 and a preference for car transportation, while Mattapan, with 35,997 residents, has a median age of 36, and residents primarily rely on cars followed by buses, with an average household income of $81,033.
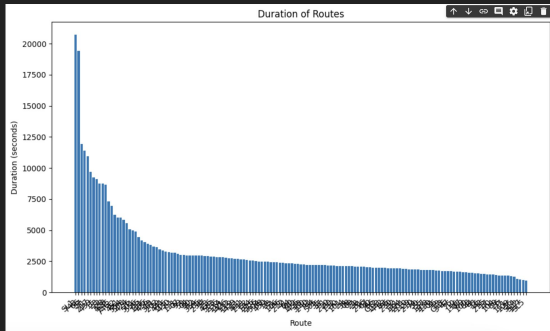
HISPANIC POPULATION: Dorchester, with 85,854 residents and a median age of 35.1, exhibits a diverse population with 54% holding college degrees, an average household income of $93,069, and a transportation preference for cars followed by buses. East Boston, home to 45,501 residents with a median age of 33.8, has a mixed demographic, and residents mainly rely on cars before using buses, with an average household income of $98,782.
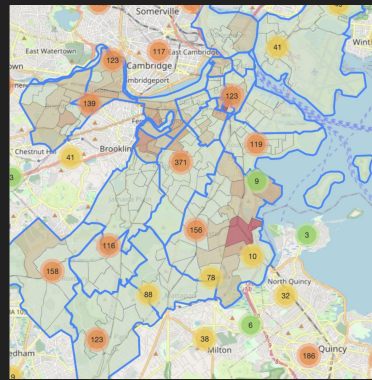
ASIAN POPULATION: Chinatown, with a population of 6,546, features a diverse demographic, with 50% holding some college degree and an average income of $50,652. Meanwhile, Dorchester, has 85,854 residents, a median age of 35.1, 54% college degree holders, and an average household income of $93,069, with a transportation preference for cars followed by buses.

We notice that a majority of the lines have end to end travel times of < 5000 seconds
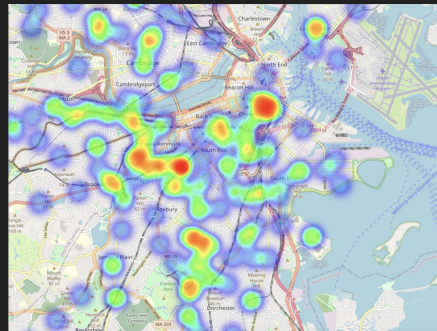
A heat map of late bus stops in Boston provides a visual representation of spatial and temporal patterns of delays, revealing clusters of high-delay areas, potential correlations with traffic congestion and infrastructure issues, and insights into route-specific challenges.
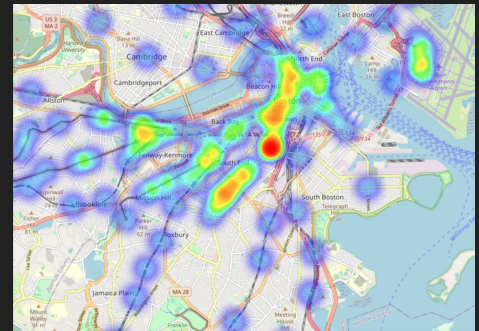
Heatmap of Disability-Accessible Bus Utilization

Reflection

Challenges:
- Lack of information in ridership csv.
- Inability to utilize car crash csv.

Limitations:
- Incompleteness and inconsistency in datasets.
- Assume the data is accurately reported and synchronized.

Next Steps:
- Further examine intense areas in the lateness heatmap.

Conclusions So Far

- There evidently is a larger population of Black, Asian, and Hispanic people In communities that experience more lateness.
- The lateness heat map shows a high concentration of delayed stops running across Commonwealth, Fenway, Roxbury, Mission Hill, and Chinatown.
- Commonwealth, Fenway, and Mission Hill are college hubs. .
- However, looking into Roxbury and Chinatown should yield more information.
- Not only does Roxbury have a higher minority population, it also has less bus stops as its neighboring communities that are predominantly white.
- This trend is apparent across Boston area when we examine the plot "Number of Bus Stops vs. Demographic".
- Chinatown and neighboring districts within the heart of the city experience constant congestion. As a result, bus delays are more frequent. Recently in 2022, the MBTA decided to not include Chinatown in the Orange Line Shuttle Route, which affects the lives of many people who depend on shuttles to travel to and from Chinatown & Tufts.

Looking at this data the people of South Boston, Roxbury, Dorchester, Mattapan, Chelsea, Revere, East Somerville and Lynn are feeling the brunt of unreliable service. We notice that a majority of people in these community do not own personal vehicles, and thus rely on public transportation (e.g. the bus) in the absence of rapid transit.

Extension Proposal

Extension Pitch

Exploring the Intersection of Disability Access and Bus Ridership

Rationale

This extension is crucial for uncovering the nuanced interplay between disability access features on buses, ridership patterns, and socioeconomic factors. It seeks to reveal how income levels, employment rates, and education within specific neighborhoods correlate with the utilization of disability-accessible public transportation. It is important for promoting equity in public transit and understanding the social dynamics influencing accessibility usage. The extension aligns with the original dataset on bus performance by deepening the understanding of the factors influencing ridership, particularly within the context of disability access.

Questions for Analysis

Socioeconomic Disparities in Ridership:
How do socioeconomic factors such as income and education levels impact the utilization of buses with disability access features?

Identification of Accessibility Deserts in Low-Income Areas:

Are there regions with a high prevalence of disabilities but lower accessibility usage due to socioeconomic factors?

Correlation Between Employment Hubs and Disability Access Usage:

Does the presence of employment hubs in certain areas correspond to higher usage of disability-accessible buses, possibly indicating a need for accessible commuting options?

Data Sets & Sources
- Bus Performance Data
- Census Data
- Disability Demographics

Visualization and Insights for Extension Proposal

- **Heatmap of Disability-Accessible Bus Utilization:**

A heatmap overlaying disability-accessible bus utilization on socioeconomic indicators, highlighting areas with high usage and potential disparities.

- **Scatter Plots of Ridership vs. Socioeconomic Factors:**

Scatter plots showing the correlation between ridership levels on disability-accessible buses and socioeconomic factors like income and education.

- **Bar Charts Comparing Accessibility Usage Across Employment Hubs:**

Bar charts comparing the percentage of disability-accessible bus ridership in areas with significant employment centers.

Understanding the socioeconomic dimensions of disability access and ridership is crucial for designing targeted interventions and ensuring that public transportation serves all community members equitably. This extension recognizes the importance of socioeconomic factors in shaping transit choices and aims to contribute insights that foster a more inclusive and accessible public transportation system.
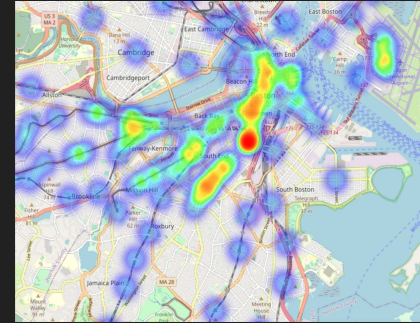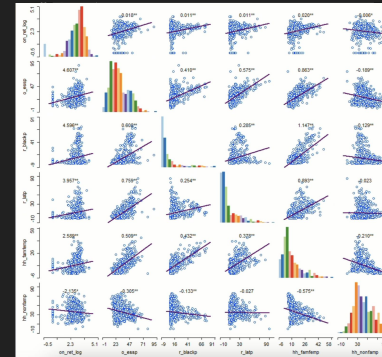

Figure.
Heatmap of Disability-Accessible BUs Utilization

Scatterplot Matrix: Boardings, Essential Workers, Race, Household Composition

The End