# CS506 Bus Performance Team D

Deliverable 1

# Problem Statement

Public transportation is plays a crucial role in the quality of life for residents in Massachusetts and the Greater Boston area. However, how sure 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.

# Exploratory Data Analysis

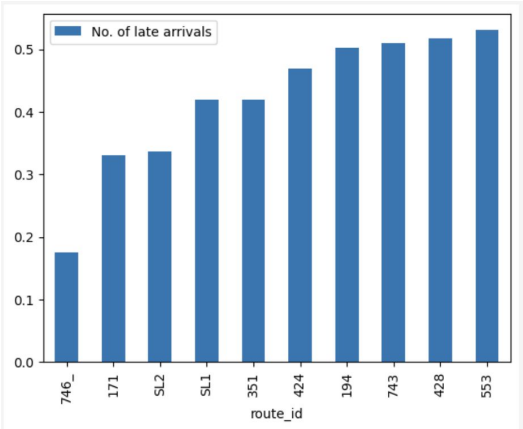
1) Pre-existing code on github which interfaces V3 API. Includes functions that get vehicles, predictions, stops for a given trip, and alerts.
2) CSV's From MBTA
    a) Bus Arrival Departure Times
    b) Bus Ridership By Time Period
    c) Bus Reliability
    d) Monthly Ridership
    e) Facilities (Wheelchairs)

# Question 1: What are the End to End Travel Times?

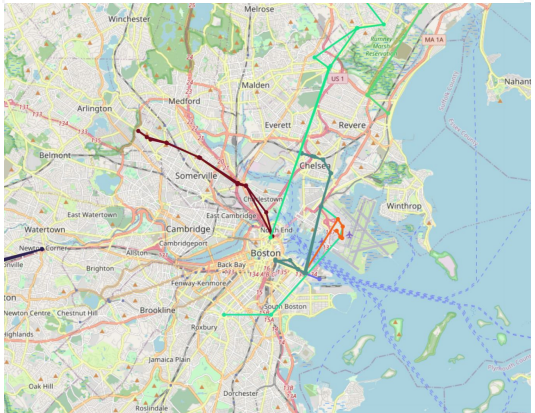
```
route_id
111   0 days 00:08:50.784762590
32    0 days 00:09:27.525725743
28    0 days 00:10:18.854261763
23    0 days 00:11:43.171046651
66    0 days 00:11:47.948230150
SL5   0 days 00:12:05.557564405
22    0 days 00:12:18.731593627
01    0 days 00:12:29.897617697
SL2   0 days 00:12:58.020416402
743   0 days 00:13:02.580268498
Name: actual, dtype: timedelta64[ns]
```

A sample of 10 end-to-end travel times from the MBTA provided CSVs sorted in ascending order.

# Lateness

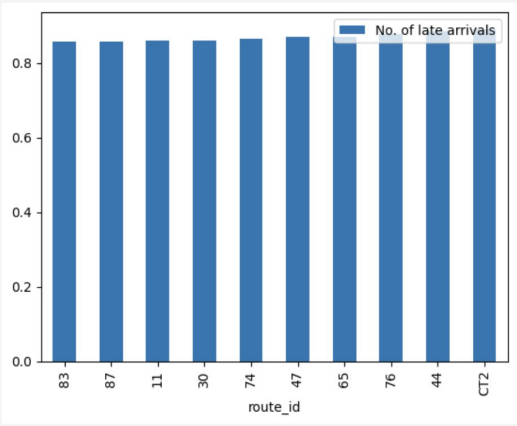



The bar graph on the left shows the 10 least late routes.

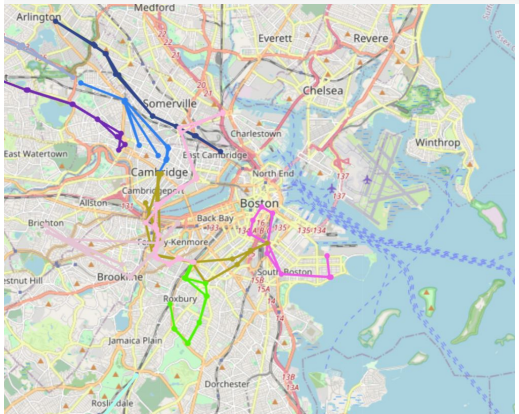



The map shows the 10 least late routes.

# Lateness (continued)




The bar graph shows the 10 most late routes.




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

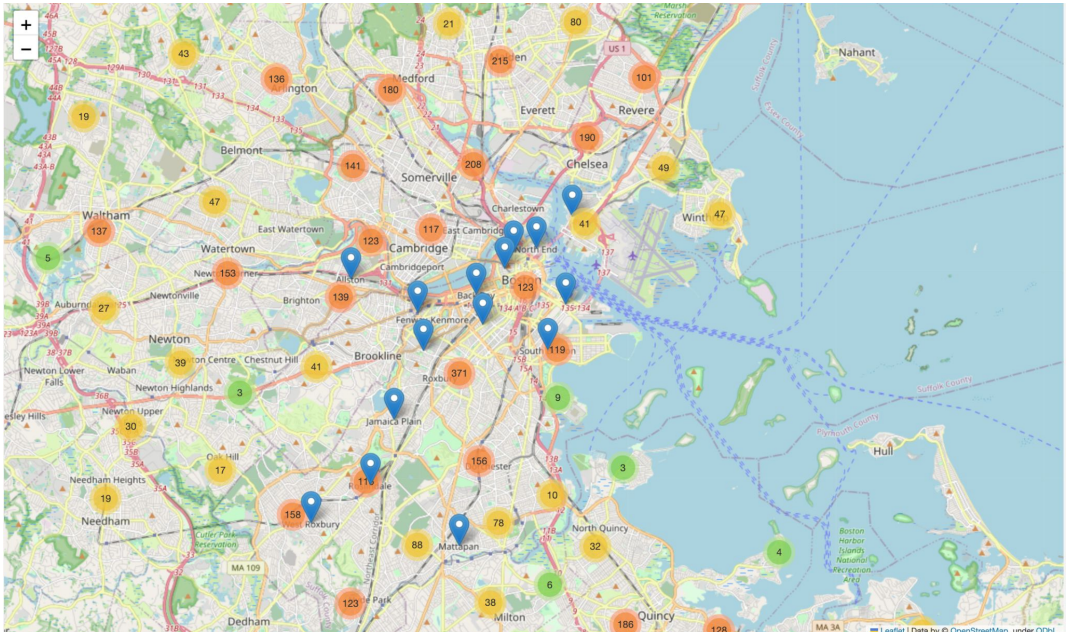
# Population Statistics

Get population count for each area

Display population against number of Stops

We would like to overlay the “most late” stops over the population map, but this csv doesn’t have metrics on the areas we are interested in (e.g.Cambridge)

Late routes are dense in Cambridge,Somerville, Roxbury, Dorchester, Jamaica Plain, and Chinatown. A majority of these areas have ~20% less White people than “on-time” routes.



# Answering the Questions

**Q: What are the End-To-End Times?**

A: The end-to-end times vary from 8 minutes to 2 hours and 45 minutes.

**Q: Are there disparities in the service levels of different routes? (which lines are late more often than others)**

A: There are lines that are consistently more late than others. The top 10 most late lines are late 80% of the time.

**Q: What are the population sizes and characteristics of the communities serviced by different bus routes (e.g. race, ethnicity, age, people with disabilities/ vulnerabilities)? (this can be an extension to #2)**

**Late routes are dense in Somerville, BU Bridge, Roxbury, Chinatown.**

Somerville: population 79k, 9.3% elderly, 74% white, 9% asian, 5% black, 12% latino

BU Bridge: College students

Roxbury: 650k, 11% eldery, 50% white, 9% asian, 23% asian, 19% latino

Chinatown: 5460, 10% elderly, 40% white, 50% asian, 4% black, 3% latino

**On time routes are dense in Charleston, North End, South End, and Saugus.**

Charleston: 16k, 10% elderly, 70% white, 9% black, 11% hispanic, 9% asian

North End: 8k, 11% eldery, 84% white, 8% black, 9% asian, 5% latino

South End: 6k, 10% eldery, 76% white, 9% black, 4% asian, 5% latino  Saugus: 28k, 20% eldery, 86.7% white, 3% black, 5% asian, 2% latino