Deliverable 1, Transit Equity Team 1

Data Aggregation

The initial Step 1 was to create a spreadsheet of all the different bus stops in Massachusetts including MBTA, Regional Transit Authorities, and City/Town buses, but we narrowed down the scope by focusing on MBTA data. We explored what datasets are available to us and may be of interest. The sources we found are compiled here.

Data Manipulation

Step 2 was to assign an income level to each stop based on the census tract data. We had a source with the median household income for tracts in Massachusetts, and we needed a way to map those incomes to bus stops. Using MassGIS Data: **MBTA Bus Routes and Stops** and MassGIS Data: **Datalayers from the 2010 U.S. Census**, we used GeoPandas to find which tract each bus stop is located in. With this information, we assigned each bus stop the median household income of its respective tract.

To complete Step 2, we assigned income levels to each bus stop using the income group standard according to Pew Research:

LEVEL	INCOME GROUP	INCOME/\$
0	Lowest income	31,000 or less
1	Lower-middle income	31,000 - 42,000
2	Middle-income	42,000 - 126,000
3	Upper-middle income	126,000 - 188,000
4	Higher-income	188,000 or more

The resulting data can be found here, and a full walkthrough of the procedure can be found here.

This preliminary analysis can be used to answer two of the strategic questions:

1. What bus routes and stops, if made free, would most benefit low income riders in Massachusetts?
    - Filtering the data here will allow us to see the names and IDs of stops located in tracts that serve low-income riders.
2. Which towns (and districts) would most benefit by a policy change to the fare change to these routes?
    - See the figures below for a town-level breakdown showing the number of buses that serve low-income riders. Boston, Lynn, and Malden would be most impacted.

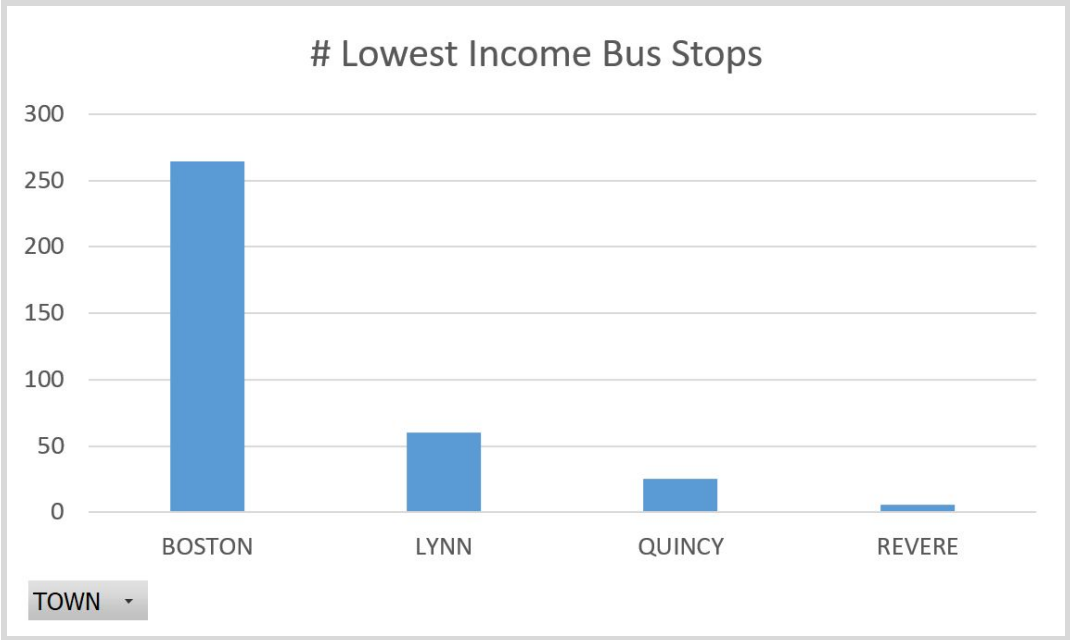

Figure 1. The number of bus stops located in tracts with an average household income of $31,000 or less, broken down by town.

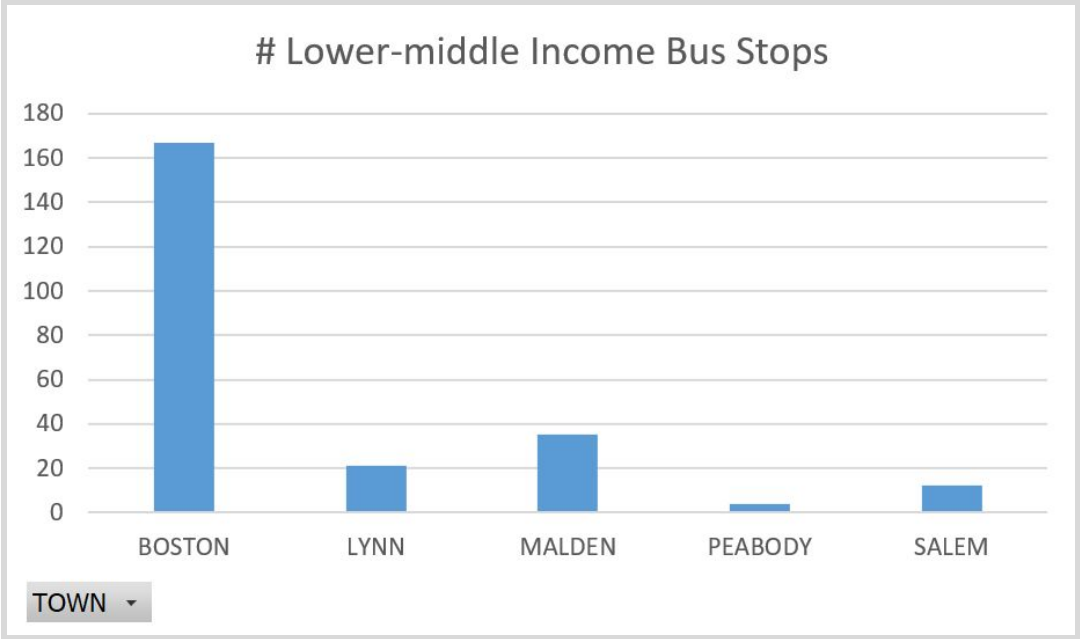

Figure 2. The number of bus stops located in tracts with an average household income between $31,000-$42,000, broken down by town.

Next Steps

The data has some missing values, and we are currently working on a solution to replace those with accurate data from other sources. We are also working on a more robust calculation of average household income, widening the scope to looking at the income of tracts within a 0.5 mile radius of each stop. Some initial code can be found here.

We have also begun work on Step 3 (determine average fare for each transit stop) and Step 4 (calculate bus ridership for each transit authority). However, we have run into an issue with calculating the average fare for each stop because of the volume of people who use monthly passes - this makes it difficult to use the initial calculation we had in mind:

$$revenue_i = \sum_{stop\_id=i} (average\_ons_i) \times fare_i$$

We have also found a promising dataset of self-reported demographic data for MBTA bus riders, which we will use in our analysis of ridership. Lastly, we hope to get more granular data about the cost of operation for a single bus/stop/route, which will help us more closely mirror the analysis conducted by Worcester, found here.