Deliverable 1 City of Boston: Transit & Performance B
Suin Lee, Kevin Smith, Vishvakishore Venkatesan, Yu Han, Yufeng Song

Problem Statement:

Despite the significance of the MBTA in Boston's public transport history and its impact on over 1 million daily commuters, there is a need to delve deeper into its service performance. With the City of Boston partnering with BU Spark!, the goal is to analyze MBTA bus data to understand the performance trends based on geography. The primary aim is to identify in this deliverable:

1. The end-to-end travel times across different bus routes.
2. Potential disparities in service levels among routes.

        Leveraging the comprehensive MBTA Bus Arrival Departure Times 2022 dataset, we developed analytical functions to figure out the operational aspects of bus routes across the network. Our analysis aimed to check the end-to-end travel times for different bus routes, providing a clear picture of the transit system's efficiency. To answer the end-to-end travel times questions we filtered out the data so we were strictly looking at service date, route id, the point type, and the actual times. By inspecting service dates, and scheduled departure times, we were also able to determine the variances in travel times and identify patterns that suggest disparities in service levels.
        However, the dataset did have some missing information and mistakes so we have been careful to take that into consideration when looking at the results. One issue that we had to deal with was the incorrect dates in the dataset so we did have to clean those out within our functions.

Our Analytical Functions do the following things:
●   Creates a visual graph of end to end service times for month of January 2022
●   Calculates the end to end times in hours:minutes:seconds format and cleans all null values and stores them in corresponding .csv files for each month
●   Calculates the lateness as a percentage and outputs to .csv file

Yu and Vishvakishore worked together for the preliminary data cleaning and analysis. They sifted through the data, standardized formats, and addressed any missing or outlier values.

Yu, Vishvakishore, and Kevin also collaborated on computing, writing code, and visualizing the end-to-end travel time and lateness of the bus routes in the python notebook.

Suin and Yufeng focused on the write-up analysis of Deliverable 1, including the method, results, and initial conclusions.

1. What are the end-to-end travel times for different bus routes
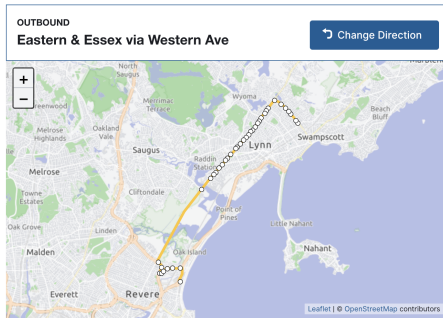   a. We will show .csv files containing this data during our presentation shown below

Average End-to-End Travel Times per Route
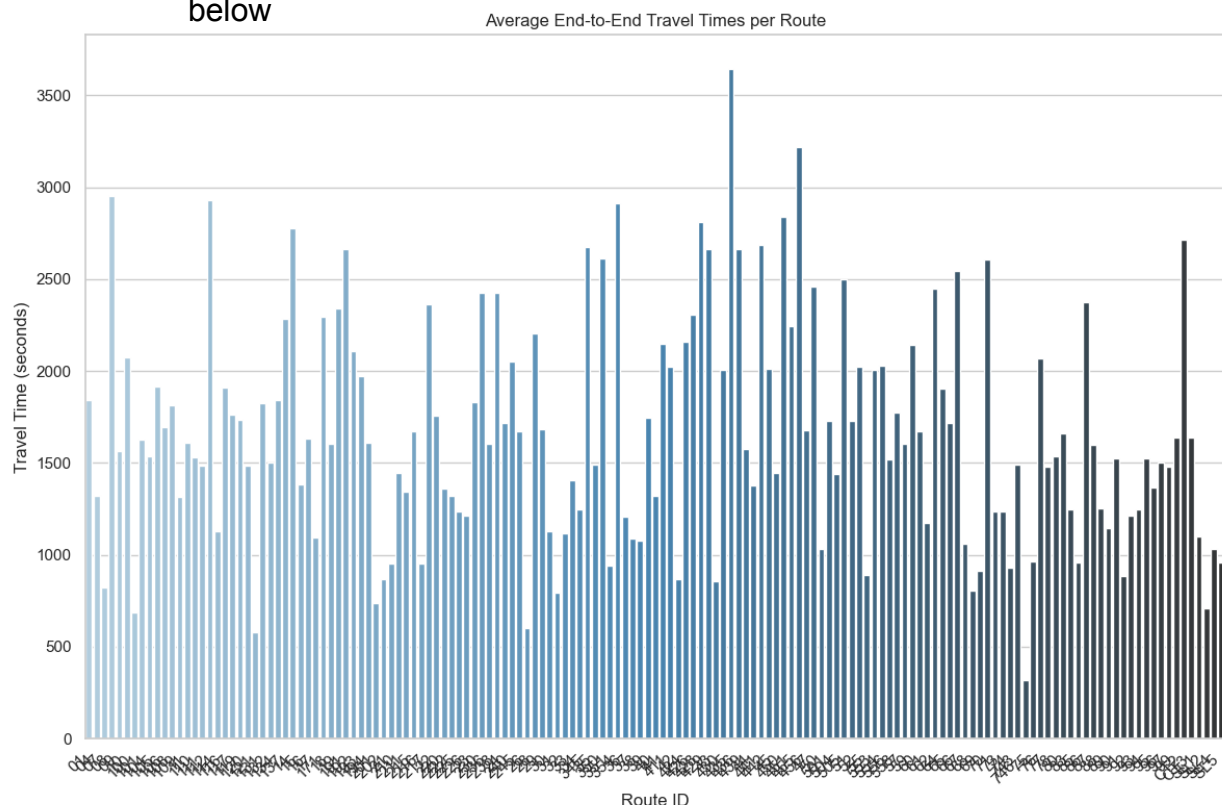
2. Are there disparities in the service levels of different routes? (which lines are late more often than others)

a. We will show a visual graph of which routes are generally later than others during presentation as shown on next page

b. We can see below that route 424 is on average EARLY in the year 2022 as opposed to route 501 which is on average ~550 seconds late in the year 2022

Average Lateness per Route in 2022