Deliverable 0

1. What is the project focus/overall goal?

The overall goal is to gather data, analyze it comprehensively, and provide insights that can help the City of Boston and the MBTA make informed decisions to enhance the quality of public transit services for residents.

2. Why is this project important?

This project is important because public transportation should not have disparities in terms of service pertaining to certain communities. The main goal of public transportation is to provide consistent and reliable means of transportation for residents. If certain demographics/areas experience an increase or decrease in quality of transportation CONSISTENTLY, then this is a fault in the system that should be addressed.

3. What type of data will you collect or be analyzing?

We should consider collecting data that allows us to draw inferences about Bus performance in relation to demographics. The google doc guide already provides us with a couple sources of information such as bus data (arrival/departure times, routes, schedules), demographic data, and performance data such as ridership, travel times and service data. We will also have to consider external factors for our extension project such as alternative transportation modes and weather conditions to determine the optimal route(s) for this project.

4. What are potential limitations of the project?

Among potential limitations for this project include the availability of data (not all bus routes/times may be available), the presence of confounding variables (such as the weather, and economic conditions) that may affect the data, temporal variations due to seasonal traffic pattern changes, the limited scope to just the City of Boston's public transportation, and complex nature of exploring and extension phase that is feasible.

5. What are your next steps? Divide tasks amongst the team
    1. Data collection: James Xiao
    2. Data cleaning and preprocessing: Haoxiang Huo
    3. Exploratory data analysis: Ketan Suhaas Saichandran
    4. Determining Key Performance Indicators: Rishven K Pravin
    5. Statistical Analysis: Xavier Thomas
    6. Application and evaluation of Machine Learning Models: Xavier Thomas