

# MBTA Travel Time Exploratory Data Analysis

Wang Jingyao

2022-12-13

## Introduction

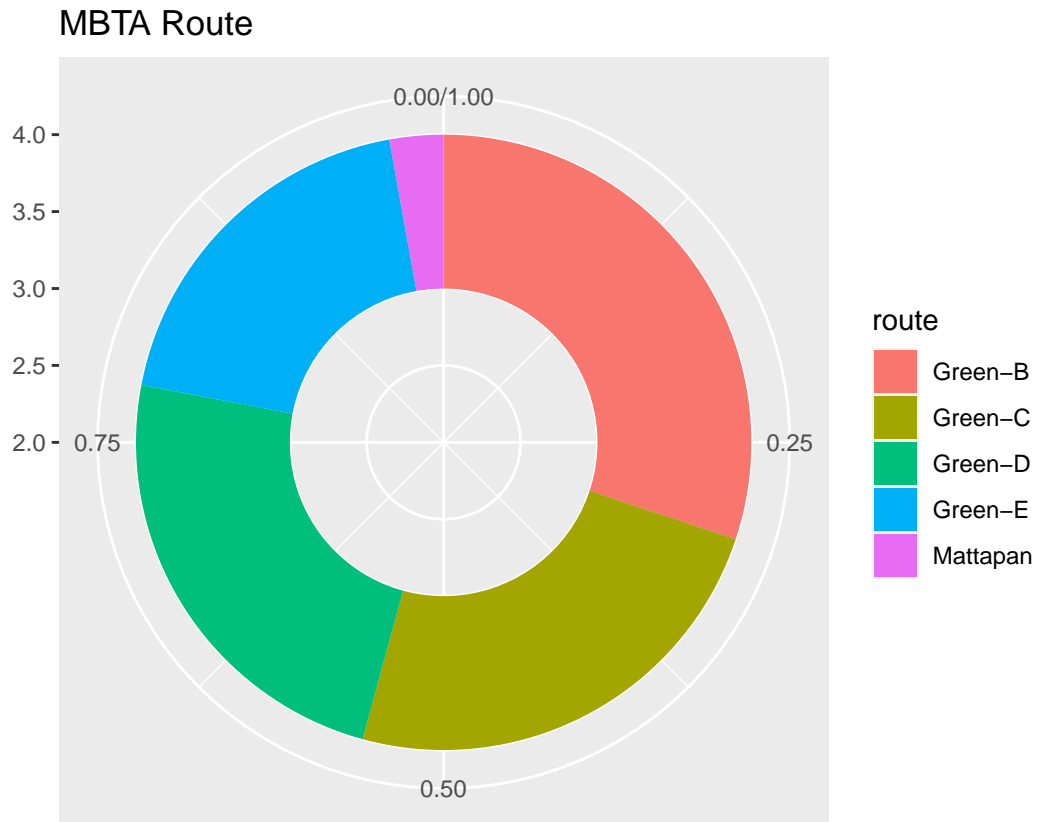
Public transportation has now become one of the main ways for people to travel. The development of public transportation in a city can reflect the development of the city. In this report, we will perform Exploratory Data Analysis on MTBA dataset.

## Data

We use the open data on MBTA website for our EDA. The dataset that we chose are **MBTA Travel Times 2021** and **MBTA Travel Times 2022**. These two datasets contains the data of MBTA from 2021 January to 2022 September. We select the time period from 2021 October to 2022 September, choose the first 7 days from each month, then bind these data to a new dataset named **LR\_New**. We are going to analyzing the travel time of Green line of MBTA.

## Exploratory Data Analysis

The green line of MTBA contains Green line B, C, D, E. Our data also contains Mattapan, which is an MBTA light rail station in Boston, Massachusetts. Our first exploratory data analysis is to count which trips appear the most in the Green Line. As shown in the chart, among the Green Line, the B Line has the most trips and the E Line has the least. In the whole data, Mattapan has the least number of trips.



Then we want to analyze the travel time of each route. We sorted the data by quarter, October to December is named 2021 Q4, January to March is named 2022 Q1, April to June is named 2022 Q2, July to September is named 2022 Q3. The travel time is the time that takes between the origin stop and the destination stop. We calculate the average travel time of each line route in different quarter.

According to the bar plot, we can notice that each line has a similar travel time in different quarter. The Green line D takes the longest time, which is around 1000 seconds and the Mattapan takes the shortest time, around 200 seconds. The route that took the second most time is Green line B, which is about 860 seconds.

