MBTA Data Analysis

Runci Hu

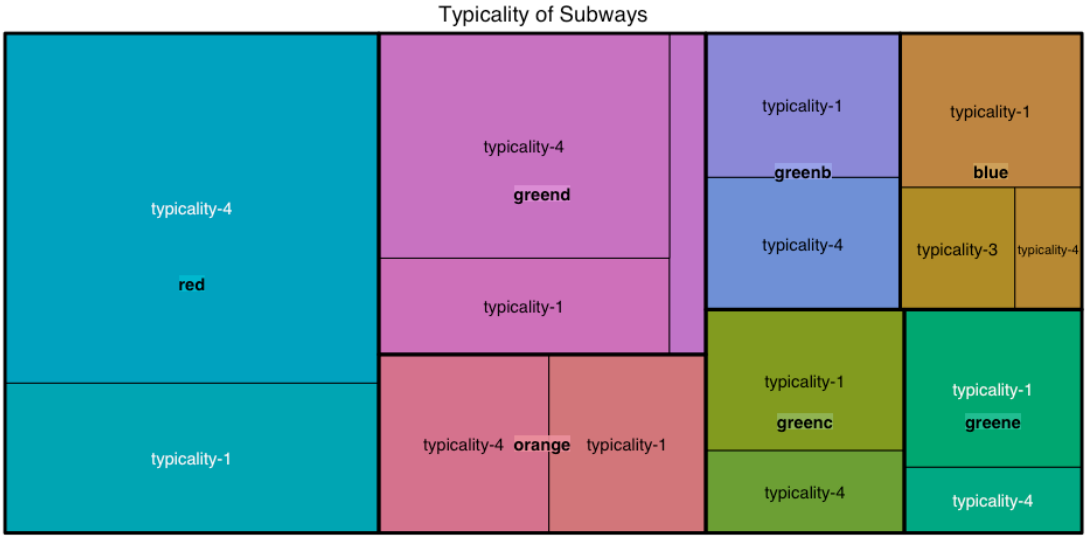
2022-12-13

Abstract

This report is studying MBTA data in Boston. The period of the research is 12 months from 2021 to 2022. The final result has two parts. One is this text file, based on the gtfs Documents, showing my analysis of topicalities of different types of MBTA transportation. Another result is a Shiny App. It is made from MBTA open data. The main idea of the Shiny App is to represent stations and routes on the map. Besides, when you click select bars and markers, it will show corresponding results with the conditions.

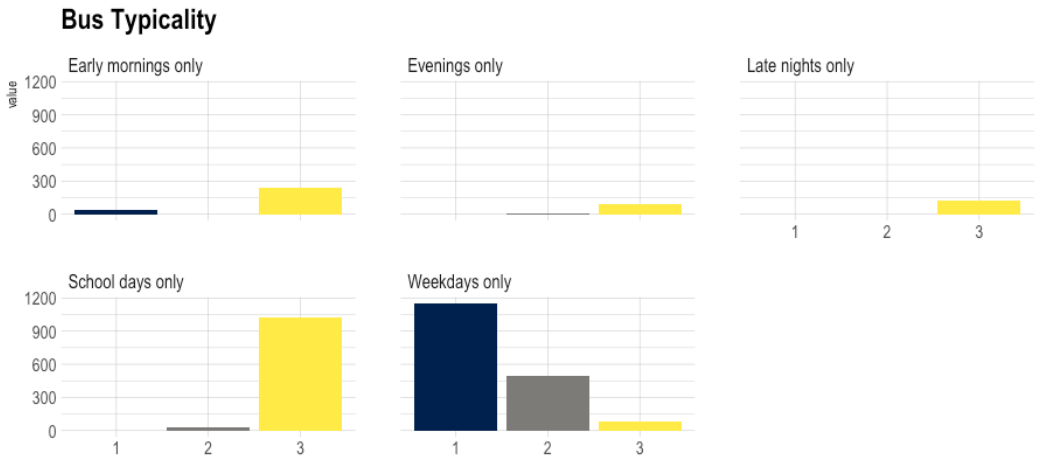
EDA

There is some information to help you understand plots. Typicality explains how common the route pattern is. There are four degrees, 0-4. 0 means not defined. 1 is the most typical pattern of a route, in which indicates that this route only has one pattern per direction. On the other hand, 2 shows a deviation from the regular route. For instance, Red Line has 2 routes in direction of southbound, Ashmont and Braintree. Pattern 3 represents a highly atypical pattern. It only runs a few times a day. At least, pattern 4 is the route that diverts from the regular way, such as shuttle buses and snow routes.



The plot describes the values of subway types with their typicality. We can see that pattern 1 and pattern 4 appear on each subway line. It is easy to understand because subway lines get their unique pattern in each direction. Pattern 1 is the

most common. The reason why pattern 4 shows a lot is subway repair. For example. The orange line was closed for renewing the pathway in the first half year of 2022. Hence, shuttle buses took place of transportation. Moreover, we can see that Green Lines are the most stable subway lines in general.



Afterwards, I explore the typicality of buses. Due to the large size of the dataset, I analysis data under service day condition. I combine data of “Weekdays only”, “School days only”, “early morning only”, “Late nigh only”, “Evenings only”. Watching the graph, we can get some useful information. Weekdays buses shows a normal pattern type. In most of the time, they keep their route pattern. Some buses have more than one route pattern. And they do not divert from their regular way. In contrast, school days buses are more often to divert from their regular way. I did not find evidence explaining why this happens. Early morning buses and late buses all have high possibility of changing routes. Above all, my suggestion for taking the MBTA routs is taking subway. In the reason that subway typicality has a smaller number of pattern 3 plus pattern 4. Smaller number represents that the route is more settled.

Shiny

The data I used for the Shiny app is from the open data of MBTA. I chose days between the 15th to the 21st day in each month from October 2021 to September 2022. Because my conclusion above is taking the subway rather than the bus and Green Lines are the steadiest routes. I filtered the original data to data of Green Lines only, and I chose one pair in each line as an example. In the Shiny App, people are able to choose which line they are taking. With this condition, they can see the location of the stations on the map, and the average travel time between the two stations. There is also a table below, which displays details information under the selected condition.