Investigating Relationship Between Monthly Bluebike and T Ridership

Evie deVos | DS340H Spring 2025 Capstone

Introduction

Background

Boston is making an effort to make its public transit more accessible. More specifically, subsidies for Bluebike for Boston residents were introduced in 2023. Did these subsidies have an effect on how desirable Bluebikes are as a method of communication?

Research Questions

Q1) Is there any correlation between the number of Bluebike rides and their average duration and the number of MBTA rides in a sample of Bluebike and MBTA rides from July 2018-January 2025?

Q2) Does the introduction of the Bluebike subsidy program in October of 2023 affect the amount of Bluebike rides taken?

Data & Methods

Data sources

Bluebike trip data from July 2018-January 2025

- ❏ Included start time, end time, station information,

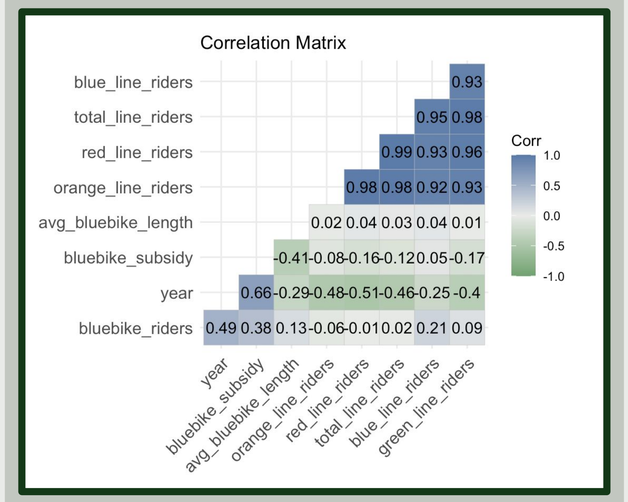
MBTA trip data from July 2018-January 2025

- ❏ Included total number of trips per month for each form of MBTA transportation

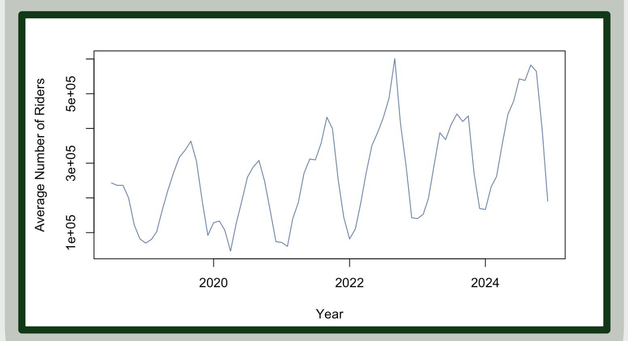
Data Manipulation

- Chose most popular "lines" based in Boston area
  - ❏ Ultimately selected Blue, Orange, Green, and Red lines
- Bluebike data was initially sorted into a spreadsheet for each month, code that iterated through each file and counted the number of rides within that given month
- Computed ride duration for Bluebikes by subtracting start times and end times

Visualizations

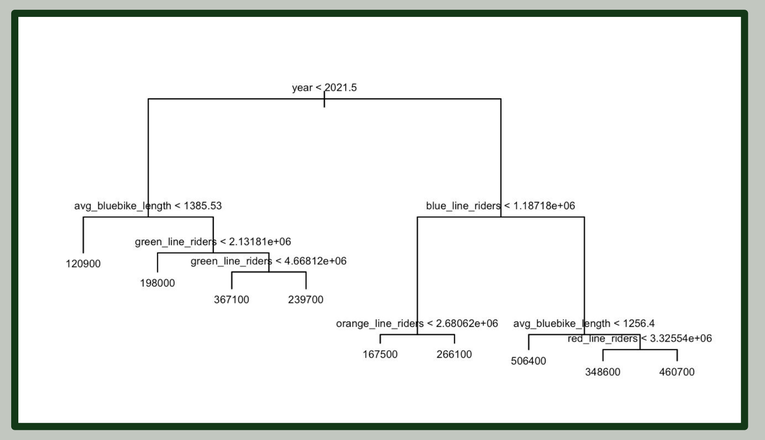


Correlation matrix

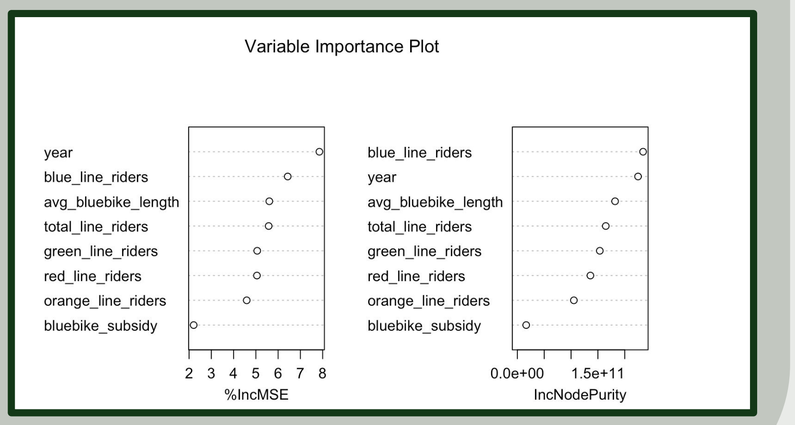


Change in Bluebike ridership over time

Model Selection



Full classification tree



Variable importance plot

Discussion & Conclusion

Results

Both the year and number of Blue line rides taken can help classify the number of Bluebike rides taken, as both have positive correlations and are ranked highly in the variable importance plot.

Limitations

Only a total of 78 data points could be used given that the MBTA data only provided month totals, and this does not account for variations that could exist on a day to day basis.