MBTA Fare Model Reconstruction
Mali Akmanalp, Nathaniel Burbank, Xinyuan Wang
Harvard AC297r Capstone

Introduction

The MBTA’s current fare model uses a single average fare by mode of transit for all trips. However, both costs and rider behaviors vary significantly by route, time of transit, and other temporal and geographic conditions. Our goal was to build a more precise model of revenues, and therefore to provide a more accurate view of the MBTA’s allocation of resources.

Datasets

Swipe Dataset:
Our primary dataset contained 275 million ‘boardings’ for MBTA subway and bus trips taken during the 2016 calendar year, including a (hashed) card id, payment amount, fare type, etc.

GTFs (General Transit Feed Specification) Dataset: The MBTA provided an estimate of variable costs associated with operating its fleet of buses. For each bus route, the dataset included the total trip count, total distance, and total duration for all routes on a day of a given day type (e.g. Summer-Weekdays) in 2016.

MBTA Cost Dataset:
The MBTA provided an estimate of variable costs associated with operating its fleet of buses. For each bus route, the dataset included the total trip count, total distance, and total duration for all routes on a day of a given day type (e.g. Summer-Weekdays) in 2016.

Data Preprocessing

1. We removed cancelled transactions.
2. For pass users, we imputed the cost of each transaction by distributing the revenue of daily, weekly and monthly passes across each user’s transactions.
3. We grouped transfers together and re-split revenue across all portions of a given trip, reassigned these values back.
4. We corrected the route names in the dataset with the MBTA official route list.

Model Construction

We used our model to estimate the revenue generated by the 64% of MBTA customers who use a monthly, weekly or daily pass. Of the 5 million boardings that occur each week, 1.7 million are bus trips, while the remaining 3.2 million represent subway fares. While more than 1/3 of monthly trips are taken on buses, bus revenue only accounts for 25% of the $34.9 million generated by passenger fares each month.

Analysis

Following the system-wide fare increases in July of 2016, according to our model the average bus customer generates $1.28 per trip in revenue. This is significantly higher than the MBTA’s previous estimate that bus passengers generate $0.90 in revenue.

Commuter trends dominate the hourly revenue patterns. However, the total estimated subsidies required to operate the bus system over the course of a week remain relatively consistent.

While revenues correlate consistently with higher mean passenger loads, estimated profits are affected by the length of the route and the type of vehicle used.

Regression Results

We found that the per-trip revenue generated by the mean monthly pass customer is equal to, or slightly higher than the average revenue generated by a single fare user. However, students, seniors, and disabled passengers can purchase passes at a significant discount. Bus routes that attract higher than average numbers of these passengers are associated with lower revenues per half trip.

Interactive Interface

We built an interactive interface that can be used by both the MBTA and the public. By simply toggling, hovering, zooming, and brushing, users can explore both aggregate trends and route-by-route comparisons.

Future work

- If provided with more granular transaction data, we could investigate similar questions at a per bus-stop level, which could help inform decisions around new routes locations.
- We could analyze passenger profiles in more detail, creating customer archetypes for commuters, tourists, and infrequent transit users.