Increasing Bike Usage and Subscriptions for Bluebikes

Amitabh Agrawal

Dorina Alimadhi

Ereza Gjikolli

Dilip Jagannathan Seshadri

Elif Kaya

Divya Minocha

[*All team members contributed equally towards the group project work*]

IPM 652: Managing with Analytics

Dr. Martin Wiener

December 9, 2019

Executive Summary

The purpose of our analysis is to identify areas for potential growth and improvement for bike usage and subscriptions in Bluebikes and as well as to provide recommendations to the management of Lyft.

The bike share service is primarily used by the age groups of 20 and 40 years old, with an evident popularity among male users (65%). It is a seasonal business, peaking during the summer months (July – August) and becoming more dormant from December to February. The bikes are used mostly in metropolitan areas rather than suburbs (Exhibit 1).

Among subscribers (80.42%) and one-time customers (19.58%), subscribers mostly use the service during weekdays (8 AM and 5 PM), while one-time customers during weekends. The most popular stations for subscribers are close to campuses (MIT station being the most popular), while nonsubscribers are more drawn to the Downtown Boston Area (South Station).

43% of the trips taken are between 0 - 10 minutes. Based on this information, we can conclude that most of the bikes are used for less than 10-minute rides. However, this data also includes issue the users encounter when trying to dock their bikes. This problem occurs when the bikes are not docked in the stations properly and where the customers get charged by the minute. Additionally, the data retrieval method from customers is not efficient, as some of the fields are not mandatory on the registration form for one-time users.

To address some of the mentioned problems and leverage the uncovered insights, we recommend introducing RFIDs for bicycles, making them dock-less, adding safety measures to reach more female users and adopting targeted-promotion to increase usage among various age groups, subscribers and non-subscribers and suburban areas.

Introduction

Bluebikes is a public bicycle sharing system in the Boston, Massachusetts metropolitan area. It has a fleet of over 1,800 bikes with 200 stations stretching over Boston, Cambridge, Somerville, Brookline and most recently Everett. Owned by these respective municipalities and operated by Motivate, Blue bikes (initially Hubway) has been operating in Boston since 2011, becoming a crucial component of public transportation in the area.

Bluebikes' fleet is docking stations throughout the city, which can be unlocked from one station and locked into another. Bluebikes offers daily passes, annual or monthly memberships and corporate memberships.

As of July 2019, Motivate was purchased by Lyft, the ride-hailing company based in San Francisco (CA). This acquisition is an extension of Lyft's vision to improve transportation sustainability, access and affordability. The company has been focusing on reaching transportation equality, environment sustainability, transit integration and street safety through scooter and bike sharing services.

As these types of services are continually increasing in the US, with 35 million trips taken in 2017 alone, through this analysis, the goal is to provide Lyft with valuable insights and recommendations in increasing bike usage and subscriptions in Bluebikes.

Data

Bluebikes dataset is available to the public in the company's website and it is in agreement with the Bluebikes Data License Agreement. The data is published monthly, and it includes:

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name, ID & latitude/longitude
- End Station Name, ID & latitude/longitude
- Bike ID
- User Type (Subscriber/Non-subscriber)
- Birth Year
- Gender, self-reported by user

In accordance with the data standard recommended by the North American Bike Share Association (NABSA), Bluebikes publishes real-time system data in open General Bikeshare Feed Specification (GBFS) format. This information is used for deriving how many docks are available in each station to perform capacity estimate.

Methodology

The raw data contained multiple tables, extracted directly from Bluebikes' website, and were stored in separate excel sheets. As a first step, the data was merged into one single file using Python script. The dataset included general information about the trips and user demographic information. Later, an additional file with information about station bike capacity and utilization was added on to the main data source. The data was then analyzed using Tableau.

Upon first analysis, we noticed that Gender column was filled in with values of "0", "1" and "2". The data provided by the company did not have a legend with explanations, therefore we reached out to Bluebikes directly. The company responded and confirmed our inferences that "0" is for Other, "1" for Male and "2" for female.

Another inconsistency with the data is the recorded age for one-time customers within the dataset was the year 1969 as the highest. (Figure 2). This is a computer auto generated response. It is a missing time stamp and conversion issue from Unix epoch time. When there is a missing value, it gets saved in the system as 00:00:00 which later when gets converted, translates into the year 1969.

Findings

Figure 1. below represents the percentage of number of trips based on the duration of the trip (in buckets of 10 min). For example, 42.76% of the trips last 0-10 min which represents the major number of trips.

Figure 1: Percentage of number of trips based on trip duration

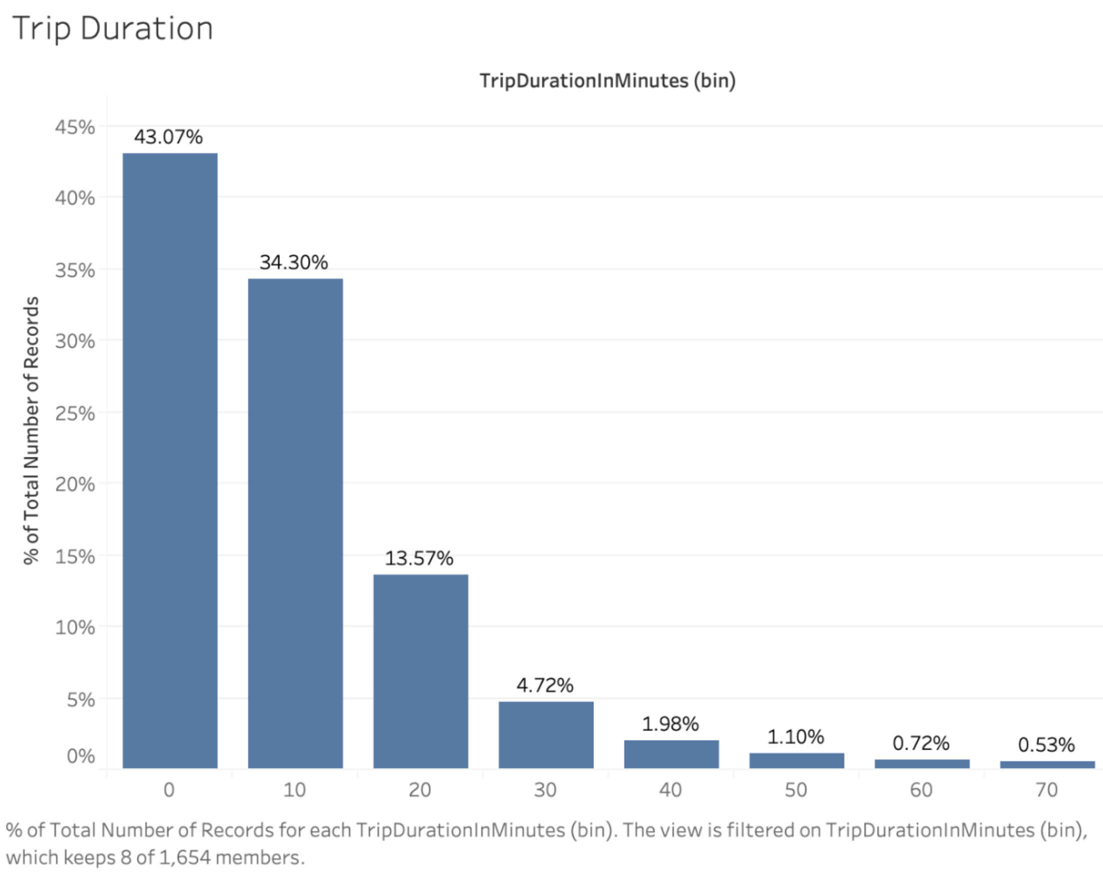


Figure 2. below provides the distribution of number of trips based on birth year and customer type. When we look at our below graph for age distribution among user types, we see a significant spike for the people who are not subscribers at the year of 1969. This is a computer auto generated response. If we need to explain it more technically, it is a missing time stamp and conversion issue from Unix epoch time. In this issue, when there is a missing value, it gets saved in the system as 00:00:00 which later when gets converted, translates into the year 1969. Therefore,

we would like to make our audience aware of this general technical assumption on the dataset to better understand the analysis.

Figure 2: Distribution of number of trips based on birth year and user type of the customer

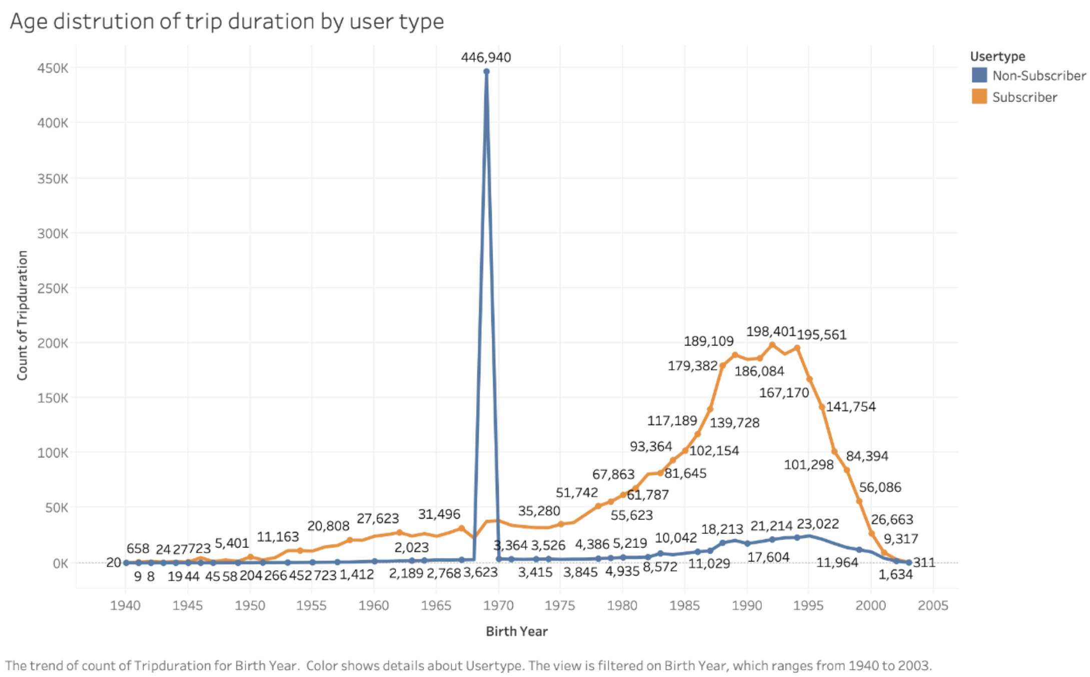


Figure 3: % of number of trips based on age buckets (size of 10) and customer type

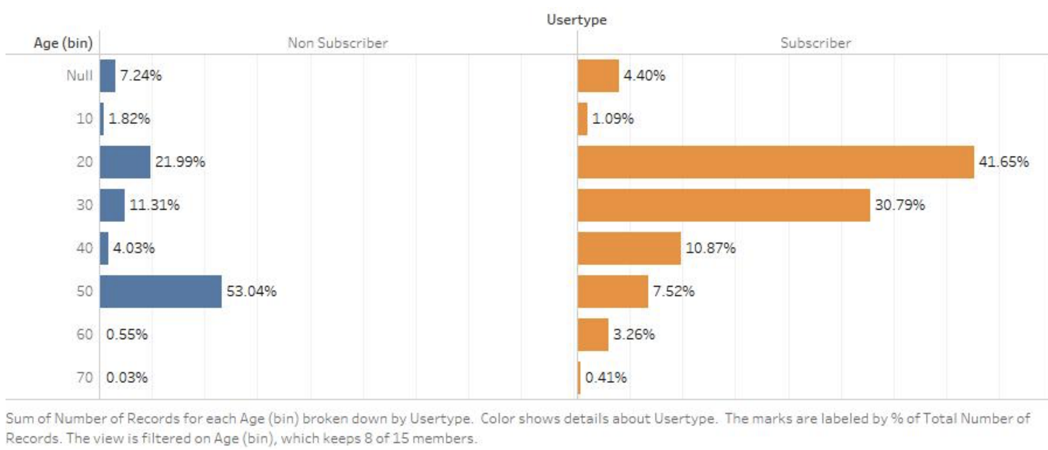


Figure 3 above displays the number of trips based on age groups and type of the customer. The above information shows that age groups between 20 and 40 are the typical users of this service. The non-subscribers showing 53.04% for age group between 50-60 which is due to the issue highlighted in figure 2.

Figure 4: Popular Metrics for Non-subscribers

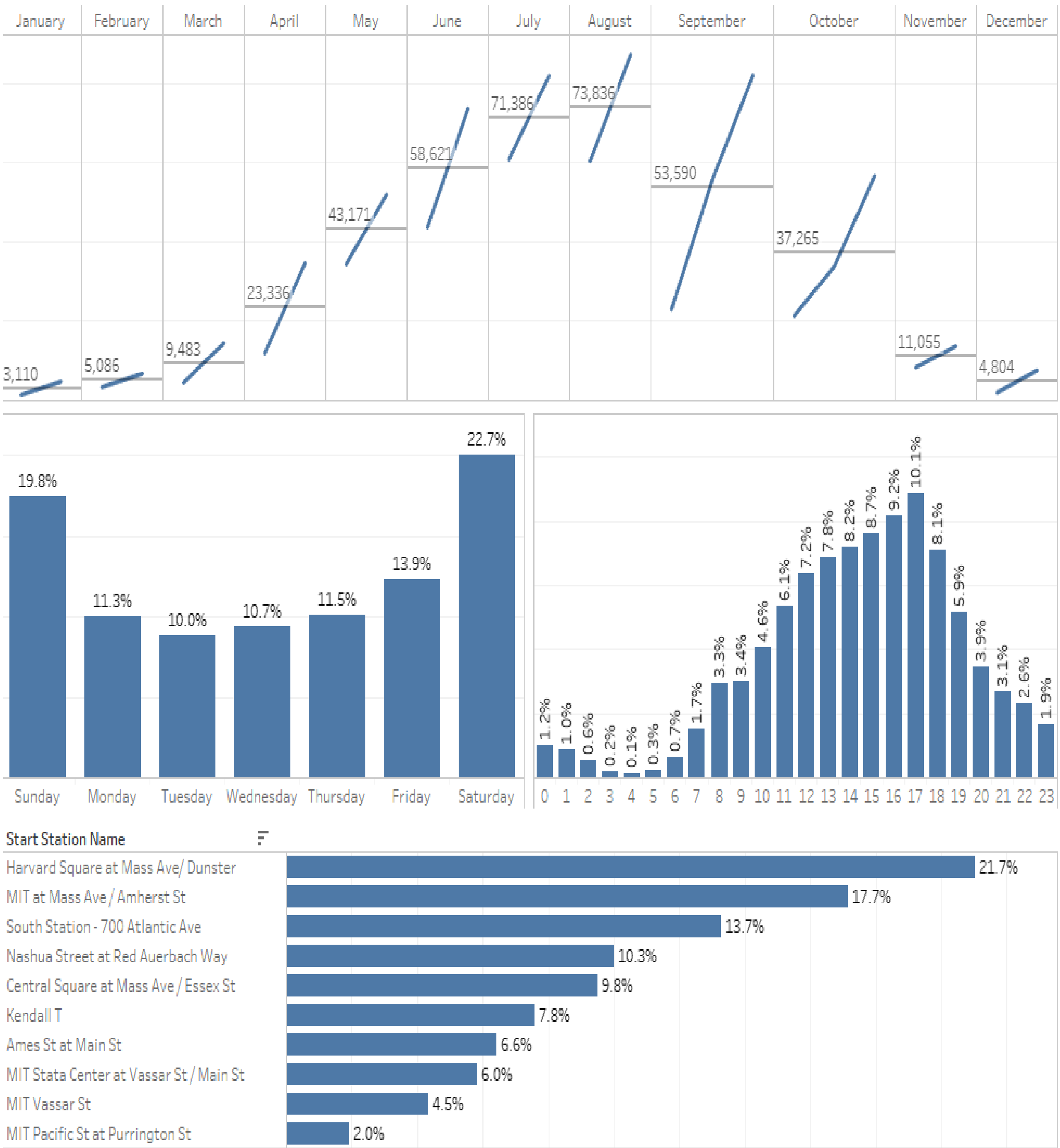


Figure 4 above provides the popular metrics for non-subscribers, seasonal view of number of trips based on months, percentage of trips based on days of the week, percentage of trips based on times in the day and top 10 start stations based on number of trips.

Figure 5: Popular Metrics for Subscribers

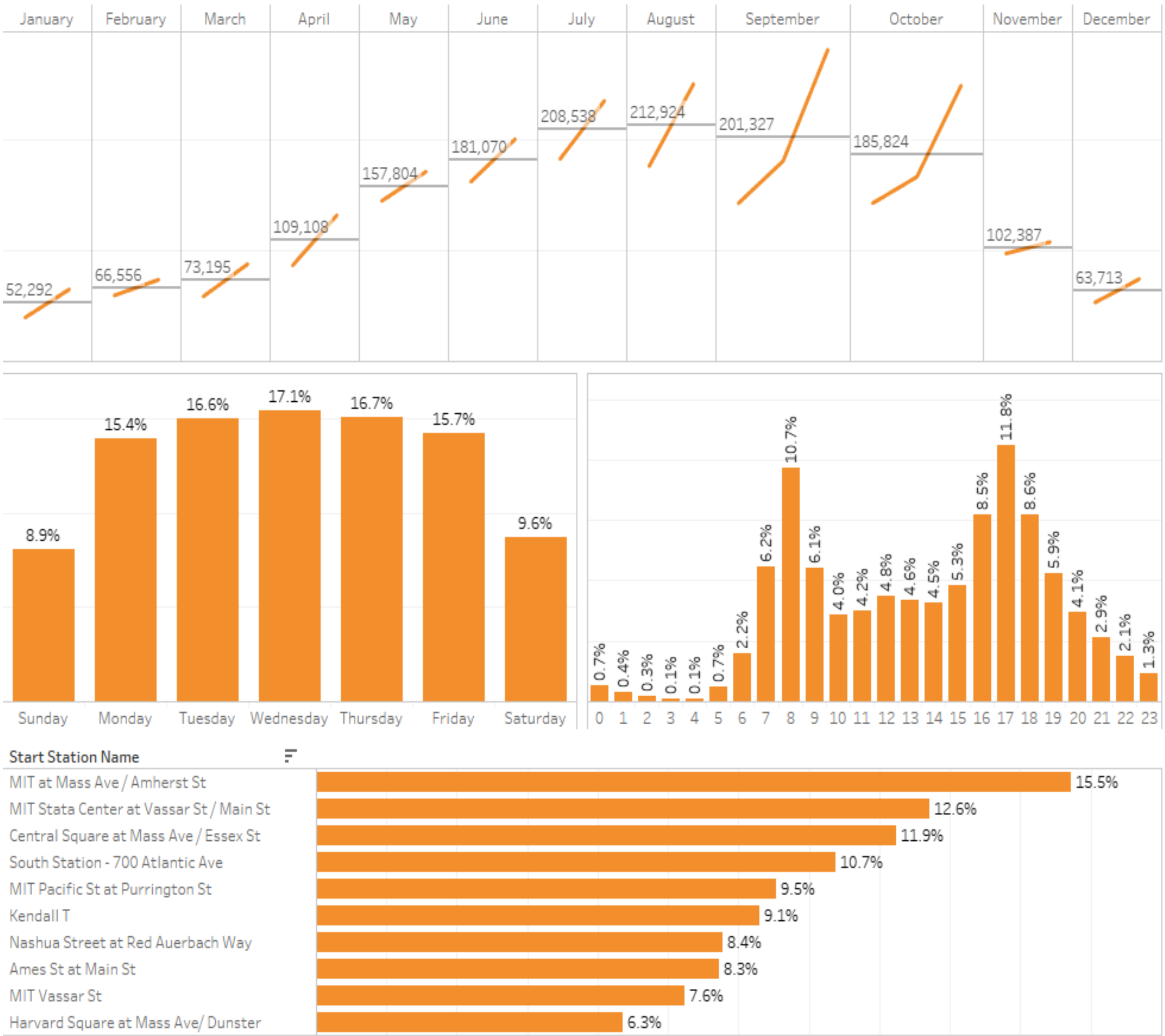


Figure 5 above provides the popular metrics for subscribers, seasonal view of number of trips based on months, percentage of trips based on days of the week, percentage of trips based on times in the day and top 10 start stations based on number of trips.

Recommendations

Following are our recommendations based on the analysis conducted above;

- Dock-less bikes: On our analysis, we can see that bike usage have increased over the years in the Boston area. We have touched upon one of the negative results of bike rental process as the final docking process. This is an issue where one-time users can get discouraged if the docking experience does not go well and they get charged by the minute until they are able to get back to the station to make sure the bike is docked properly. Making bikes dock-less will potentially increase the usability of the bike service. This will also improve the customer satisfaction of subscribers thereby creating lock-in opportunities in value-sources framework (Figure 1)
- Add more safety measures for bikes and improve lane conditions: There are more male riders in our data set than female riders. (Exhibit 2) However, in general assumption, females tend to ride bikes in areas where there is a better system of bike lanes. This was stated in an article written in 2017. Taking this into account, the new management can try and identify which routes are more common among female users, and with this data they may be able to generate a solution to the gender gap of riders. Also, female riders face number of obstacles as discussed in the article referenced and safety being one of the important ones. Increasing more safety measures in bikes such as adding a button to alert nearby police force etc. can help increase bike usage by female riders. We believe that, as the bike lanes increase, people may be more comfortable riding bikes. This can be an opportunity for the new management to work with city officials to provide more bike lanes in each city or neighborhood and work with health groups or environmental groups to promote its bike usage. We should not forget, not only do people not burn any gas when they bike but they also burn calories.
- Improve promotion strategies based on key insights: Improve promotion strategies based on key insights derived from findings:
  - When we look at the bike usage in a year, we can see that summer times are the busiest times for bike usage. (Figure 4 and Figure 5)
  - Non-subscribers typically use this service over the weekends, near popular landmarks and sight-seeing places in the evenings. (Figure 4)

- Subscribers typically use this service over weekdays, near colleges/offices during typical college / office hours (peaking at 8 AM and 5 PM). (Figure 5)
- Based on the data, people aged between 20 and 40 are the majority users of this service. Promote this service more in commercials targeting these age groups (Figure 3)
- The service is primarily used near Boston downtown and the usage decreases significantly in suburbs. Promote this service more in busy areas with short commutes and higher traffic (Exhibit 1)

- <u>Make key fields in forms mandatory for one-time users:</u> Eliminate errors in data for getting more meaningful insights of one-time users by making the following fields mandatory (age, gender etc.) (Figure 2 and Exhibit 2). This can help key insights in deriving better promotion strategies for one-time users

Conclusions

Through this research, the main goal was to derive meaningful and helpful insights in the usage and subscription patterns of Bluebikes. It is recommended that company management address issues such as usability, accessibility, safety, bike docking and data collection while maintaining the already well-established profile of the service. By improving safety measures, both in lanes and bikes, research suggests that the service will become more appealing to female riders. Additionally, reducing tedious issues such as improper bike docking will also impact the user usage and satisfaction thereby increasing lock-in opportunities in values-sources framework. Lastly, the company has enough information to implement targeted promotion by using age, subscription, gender, urban/suburban tiers. It can further enhance this type of promotion by improving the user registration forms and collecting necessary data.

<u>Directions for future research:</u>

Future research should examine capacity constraint in increasing bike usage and subscriptions based on real time data and examining the capacity used in each station vs capacity available in each station. This would help in identifying bottlenecks in popular stations based on the usage. Further research can be done on unlocking patterns between types of subscription and one-time passes and drawing inferences on gender, age groups, location, residence and more. Another key analysis would be on learning how the weather and temperature changes can affect bike usage during the peak and dormant periods.

References

Anzilotti, E. (2017, November 27). What Will It Take To Close The Gender Gap In Urban Cycling? Retrieved December 2, 2019 from https://www.fastcompany.com/40488970/what-will-it-take-to-close-the-gender-gap-in-urban-cycling

Bluebikes System Data. (n.d.). Retrieved October 2019, from https://gbfs.bluebikes.com/gbfs/gbfs.json.

Index of bucket "hubway-data". (n.d.). Retrieved October 2019, from https://s3.amazonaws.com/hubway-data/index.html

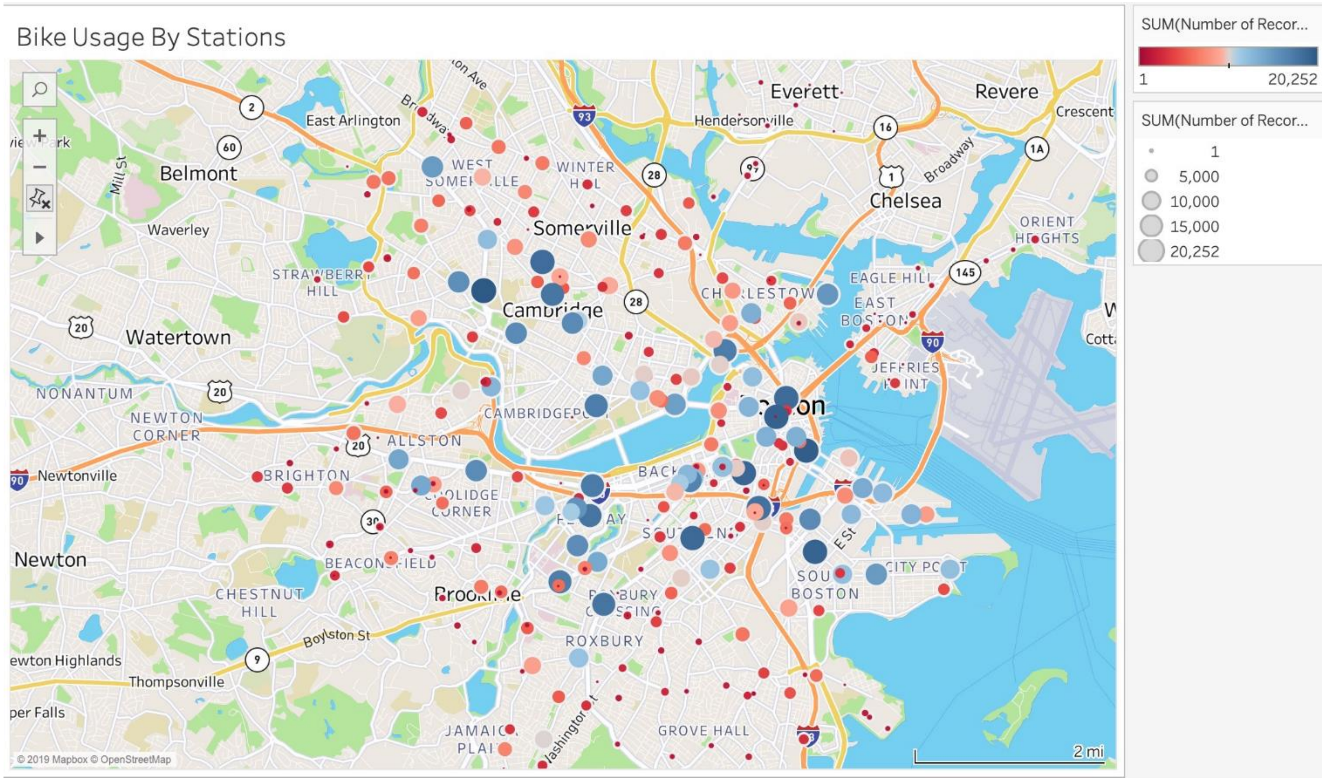
Marcotte., A. (2014, September 17). Today in Gender Gaps: Biking Retrieved December 2, 2019 from https://slate.com/human-interest/2014/09/gender-gap-alert-men-ride-bikes-way-more-than-women-do.html

Parker, K., Rice J., Gustat, J., Ruley, J., Spriggs, A., Johnson, C. (2012). Effect of Bike Lane Infrastructure Improvements on Ridership in One New Orleans Neighborhood https://link.springer.com/content/pdf/10.1007%2Fs12160-012-9440-z.pdf

O'Kane, S. (July 2, 2018). Lyft buys the biggest bike-sharing company in the US. Retrieved November, 2019 from https://www.theverge.com/2018/7/2/17526892/lyft-buys-motivate-bike-sharing-expansion

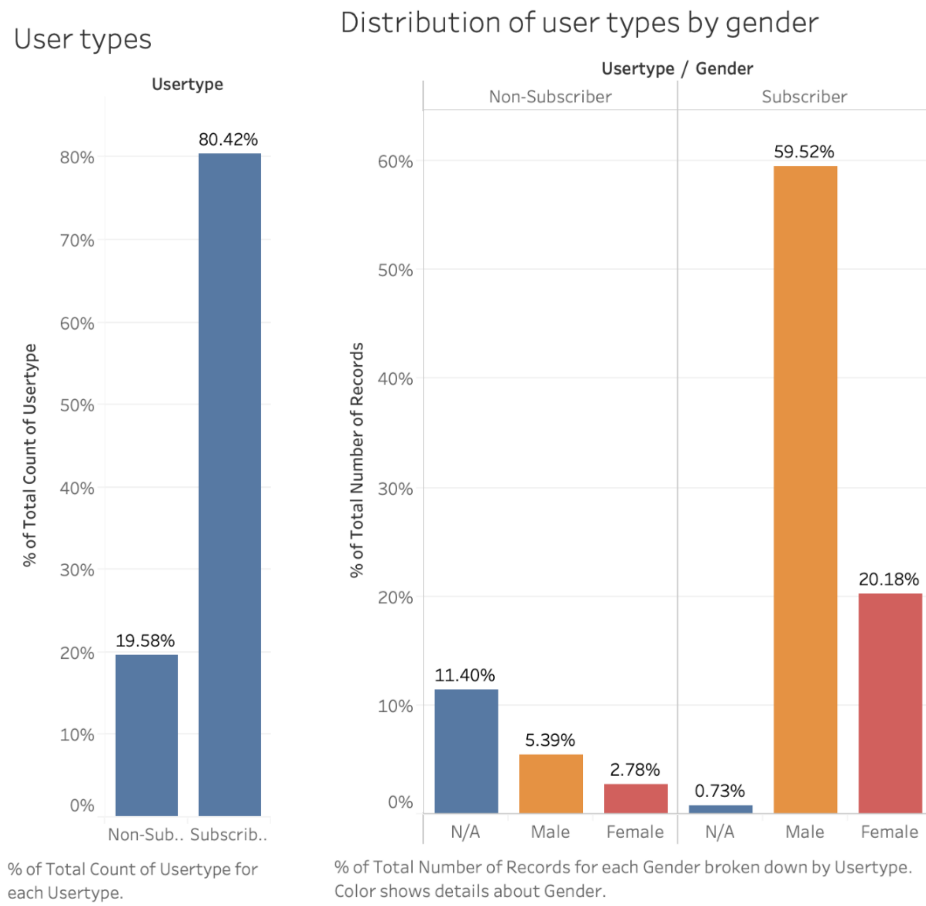
Exhibits

Exhibit 1: Number of trips per station



When we look at the bike usage by station, it clear that the most popular stations are close to the Boston downtown area and around university campuses (MIT and Harvard for Cambridge, Boston University (Kenmore), Suffolk (Downtown Boston)).

Exhibit 2: Percentage of number of trips based on gender and user type



When we look at the usage based on gender, male subscribers seem to be using this service the most. Similar analysis cannot be done for one-time users due to the issue of optional entries for indicating the gender in their forms.

Exhibit 3: Trip durations by user types

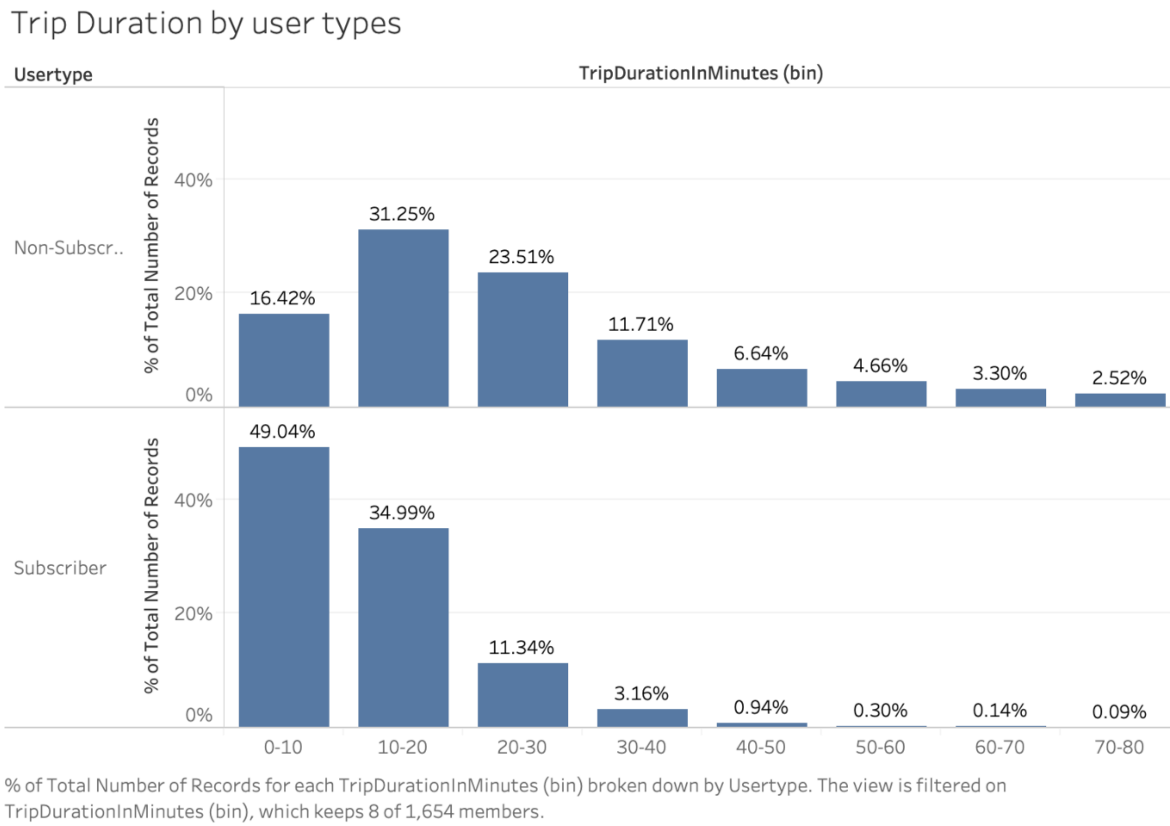


Exhibit 3 shows us that non-subscribers tend to use the service for longer than subscribers.