Citi Bike Analysis (August 2019—January 2020)


**Your task in this assignment is to aggregate the data found in the Citi Bike Trip History Logs and find two unexpected phenomena.**
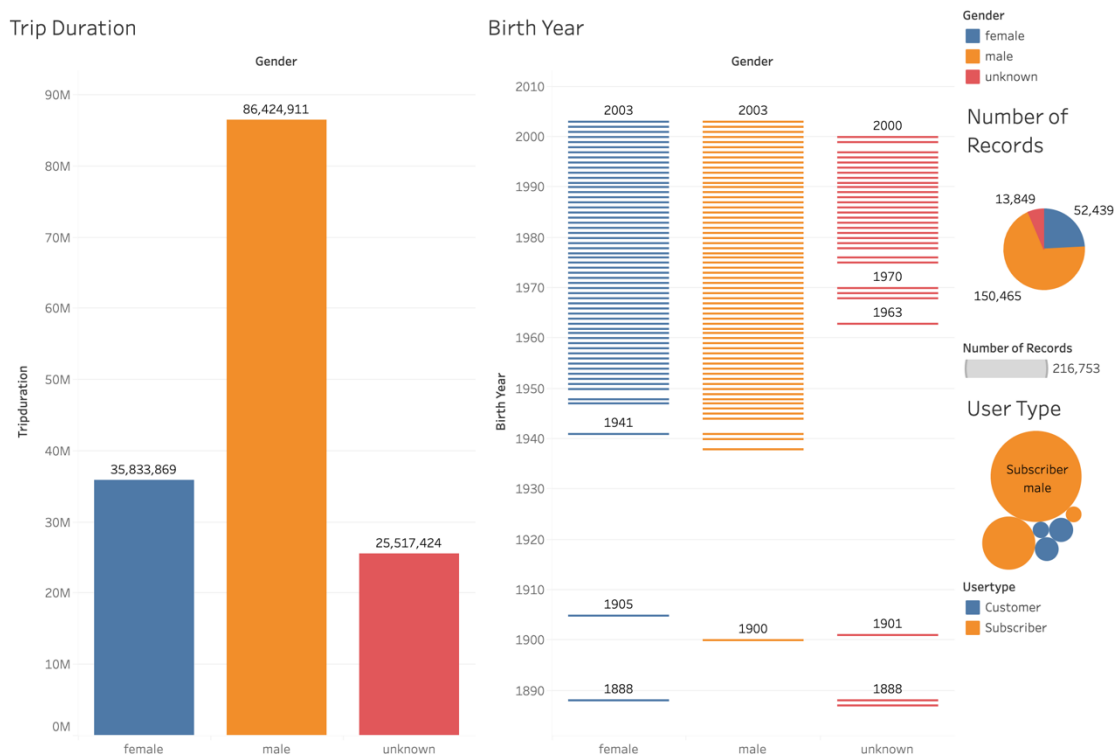
**Design 2-5 visualizations for each discovered phenomenon (4-10 total). You may work with a timespan of your choosing. Optionally, you may merge multiple datasets from different periods.**

**Tableau Link:**
https://public.tableau.com/profile/peirang.xu#!/vizhome/CitiBikeAnalysis_15826640950230/Story1?

**Github Link:** https://github.com/peirangxu/tableau-challenge.git
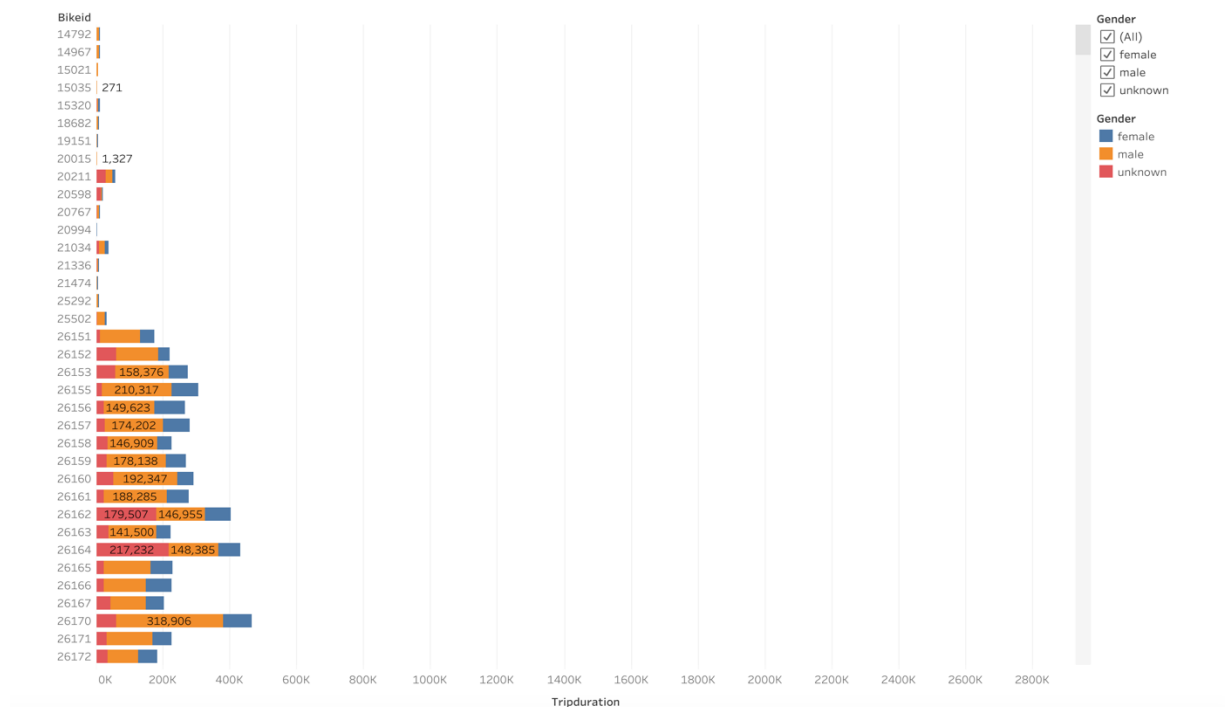

## Dashboard I: Gender Effects



Considering the gender effects on citi bike riders, first we can see that the trip duration for male riders is much longer than that for female riders (86,424,911 comparing to 35,833,869).
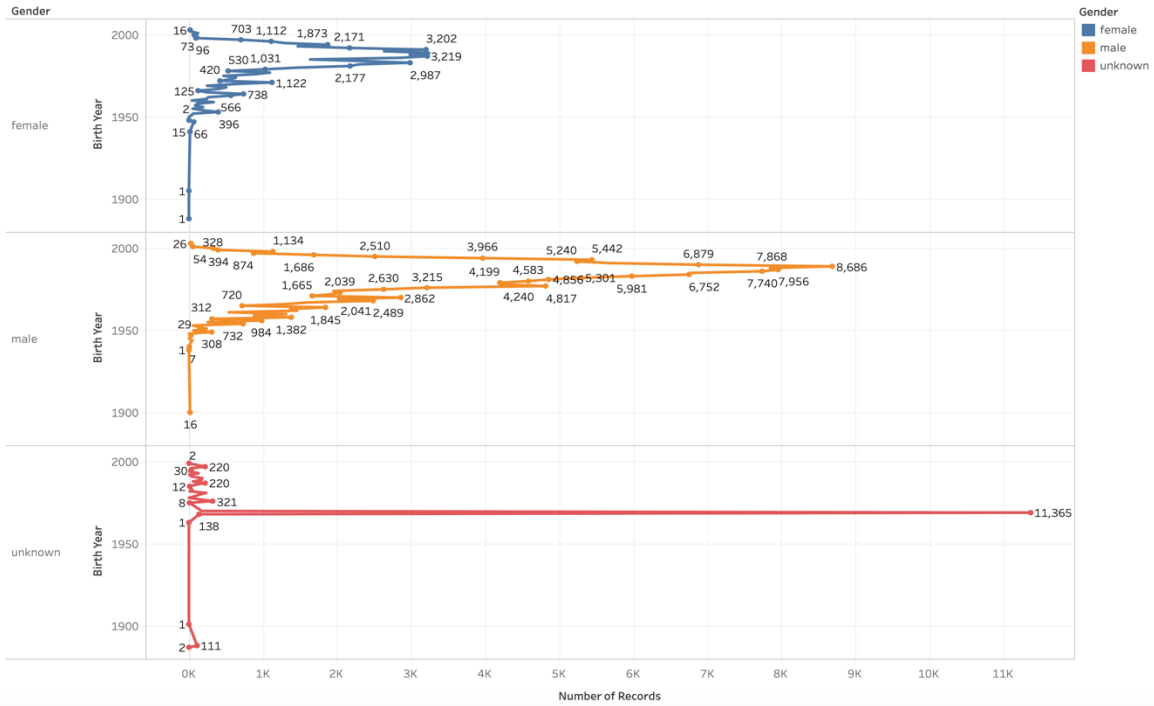
Looking at the number of records we get from the data, females rider records is only about ¼ of the data and male riders consist of almost ¾ of all the records.

Age-wise, the distribution is similar for male and female riders, as in most of them are born in years between 1950 and 2003, indicating a rider age range of 17-70 years old.
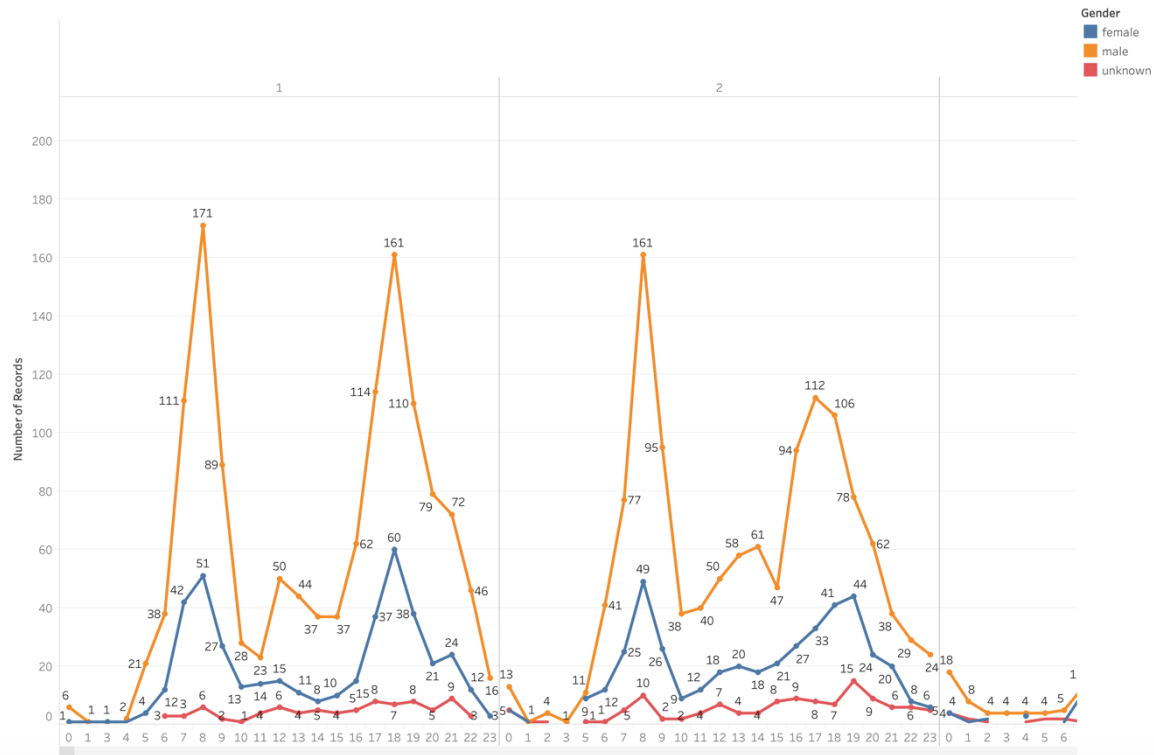
Also, we can see that most of the riders are subscribers (annual pass holders) and only a small portion of riders are customers.



Looking at the trip duration for each bike id, we can see that male riders are still the majority portion, and some of the bikes are used more often than the others.
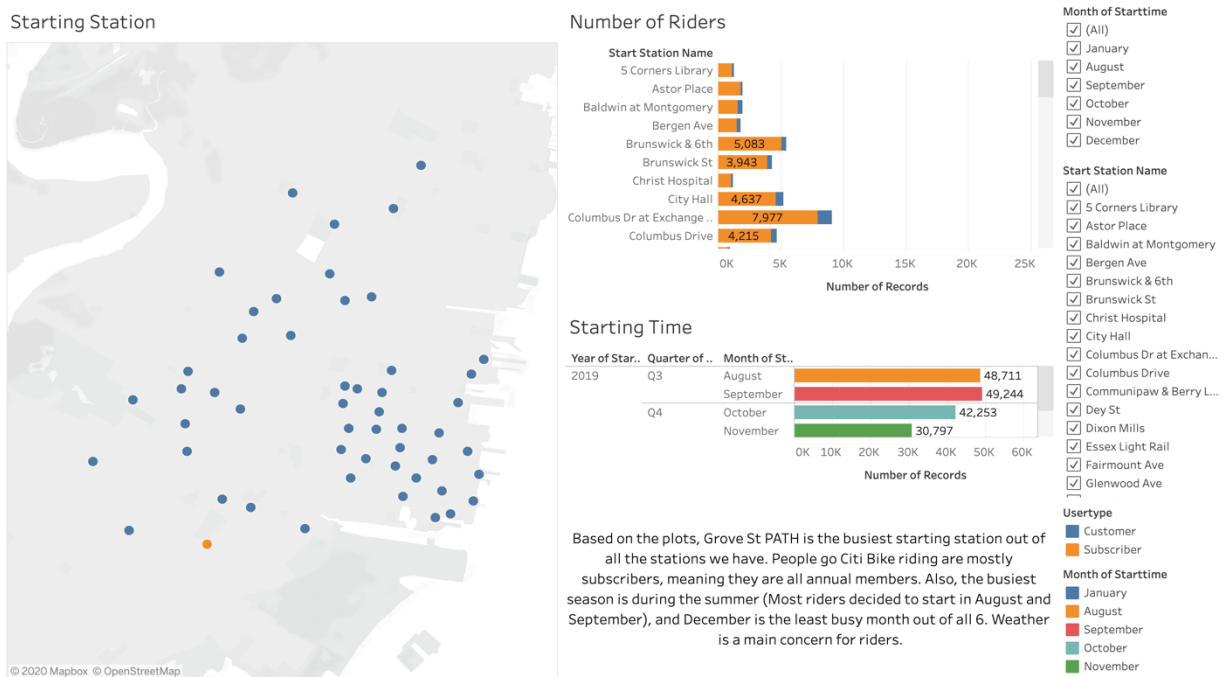
From this plot, we can see that male riders have a greater number of records than females. However, there is a peak for the gender reported as unknown corresponding to birth year of around 1960.
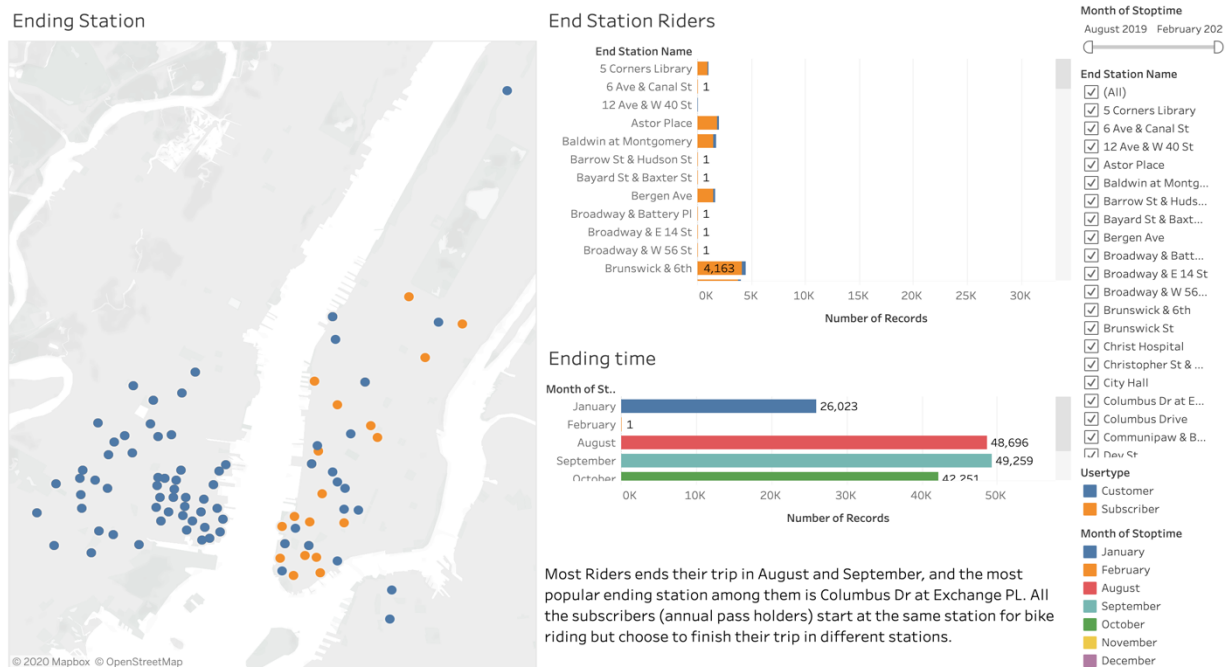
The number of records have peaks on the 8<sup>th</sup> and 18<sup>th</sup> of each month based on our data. Although males are still trending more records than females and the unknowns, their peaks follow a similar trend.

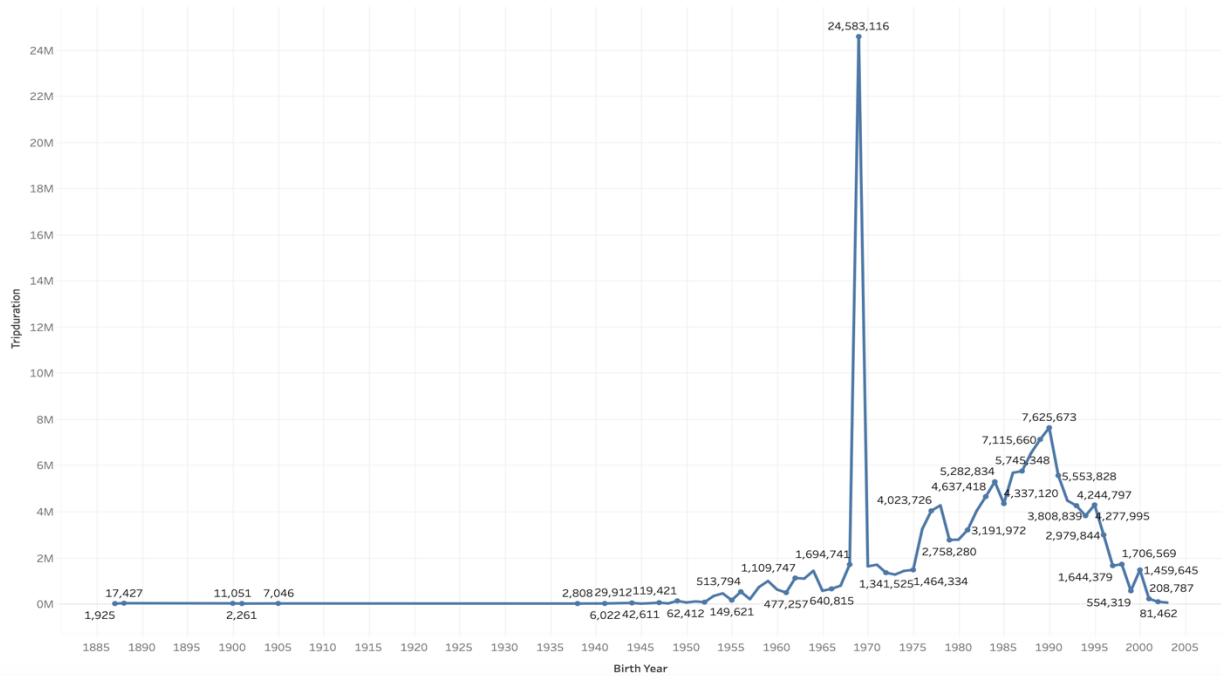**Dashboard II: Starting and Ending Station Analysis**



Based on the plots, Grove St PATH is the busiest starting station out of all the stations we have. People go bike riding are mostly subscribers, meaning they are all annual members. Also, the
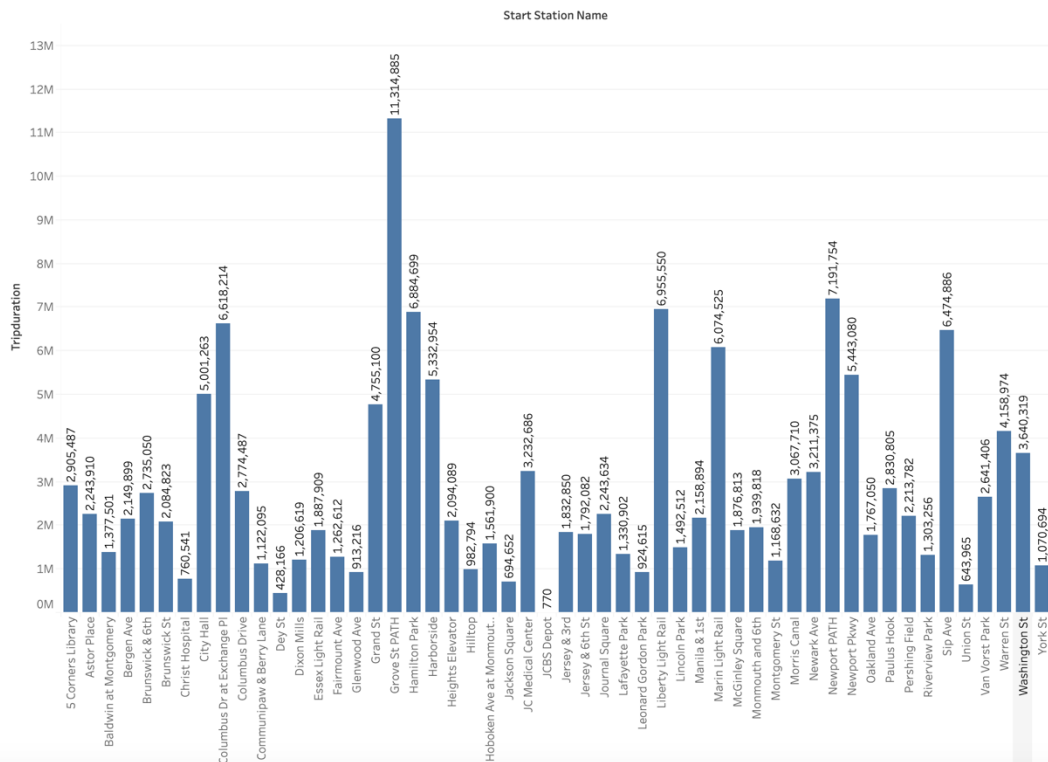
busiest season is during the summer (Most riders decided to start in August and September), and December is the least busy month out of all the 6 months we analyzed. And this could imply that weather could be a main concern for all the riders.
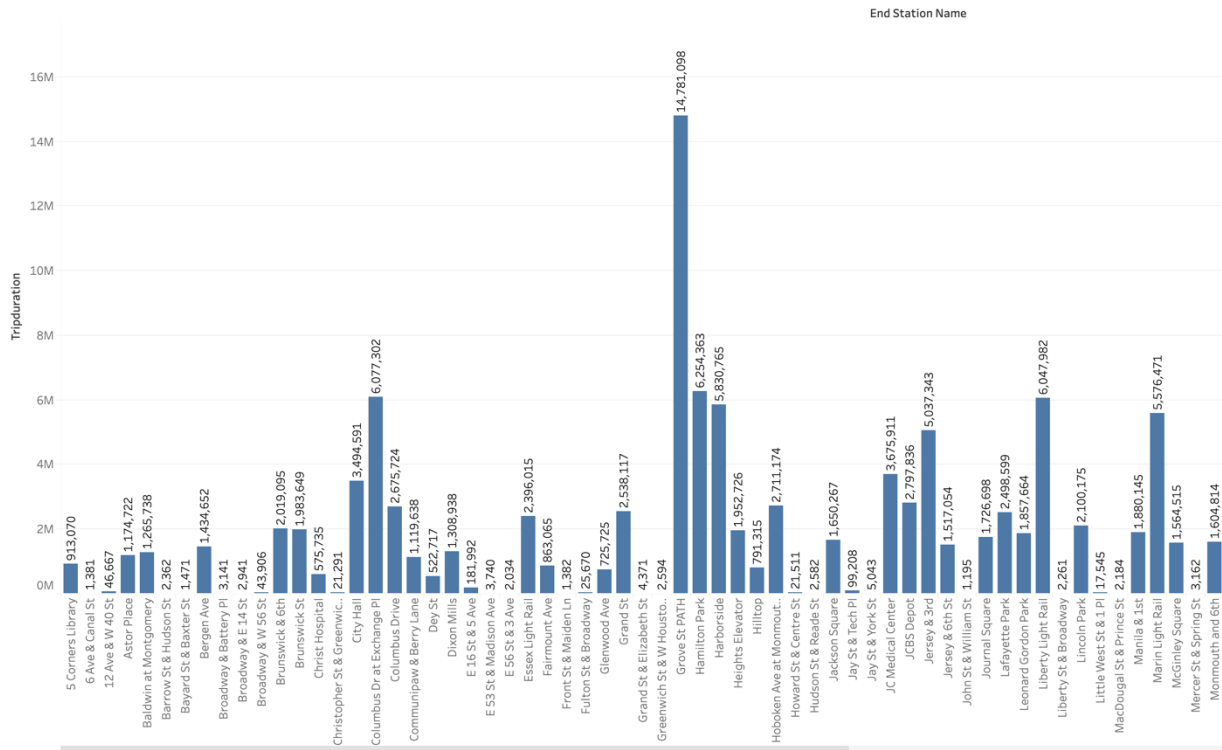


Most riders end their trip in August and September, and the most popular ending station among them is still Grove St. PATH, and Columbus Dr. at Exchange PL station is also a really busy one. All the subscribers (annual pass holders) start at the same station for bike riding but choose to finish trips in different station.

The riders born around 1970 (age around 50) have the longest trip duration based on this plot. The people born after 1970 generally have longer riding time comparing to the people born before, this could be due to age effects on riding time.



The most popular station to start a ride is Grove St. PATH, and there are some stations that are really unpopular, for example, JCBS Depot only has 770 riders to as a start station, comparing to 11,314,885 at station Grove St. PATH.

End Station Name

Looking at the plot for end stations, Grove St. PATH is still the most popular, and as the number of records change between start and end station, we can see that people do not generally start and end in the exact same station.