

# Udacity Data Visualization Project

## Ford GoBike System Data



Photo source: Sunset Magazine

**This presentation summarizes the results of data of the Ford GoBike System exploration.**

### **Dataset Overview:**

The dataset is showing the data for the trips made using Ford GoBike during February of 2019. It includes the data of 183412 Ford Gobike trips in the dataset.

There are 16 factors within trips:

[duration in sec, start time, end time, start station id, start station name, start station latitude, start station longitude, end station id, end station name, end station latitude, end station longitude, bike id, user type, member birth year, member gender and bike share for all trip].

### **Investigation Overview:**

We try here to see the relationship between trips durations in seconds and other factors.

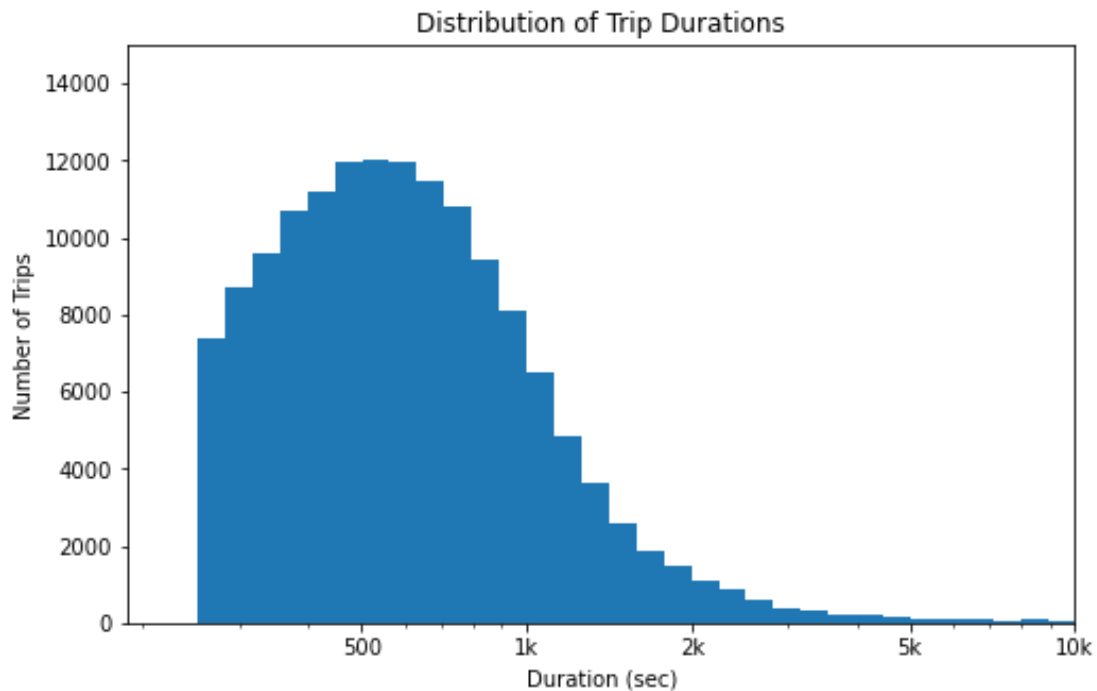
### **Key Insights for Presentation:**

I have chosen the appropriate plots to clarify the distribution of coefficients and show the relationships between them and how they affect each other, and I have used the logarithm scale to clarify and fix the spread to be easily understood.

## Summary of Findings and plots:

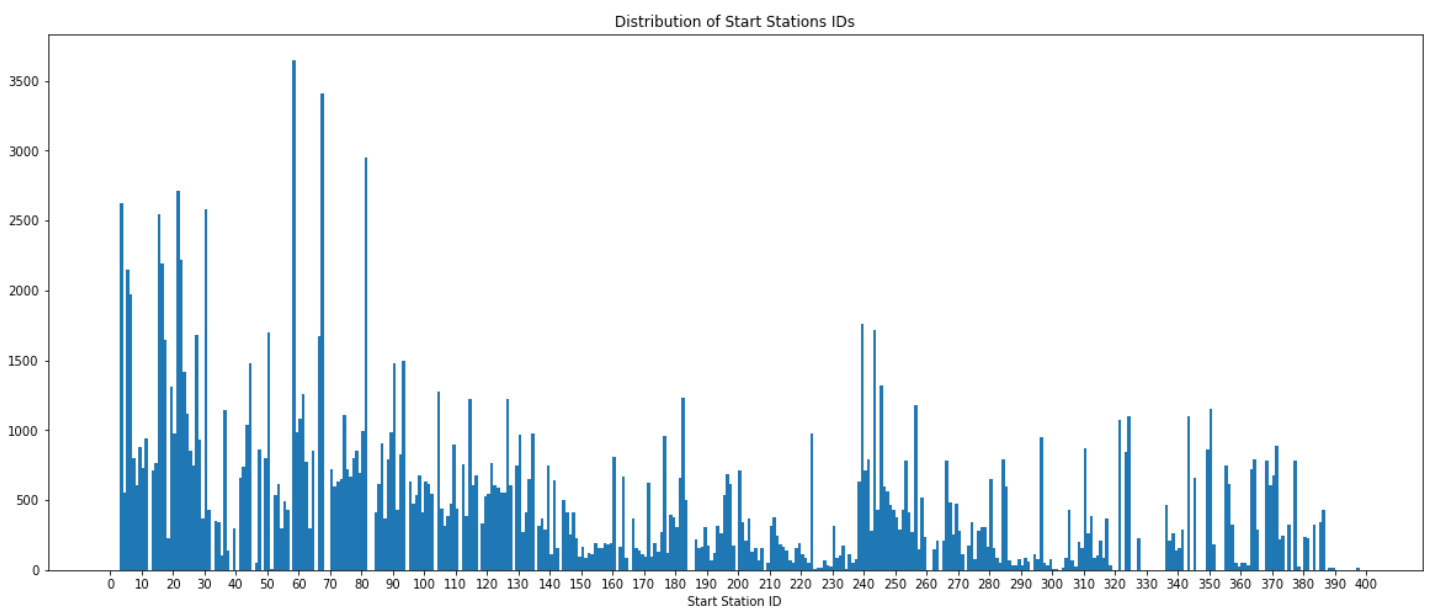
### 1- Number of trips Vs trips durations:

Most of trips' durations are less than 2000 seconds with peak around 550 seconds. around 7200 trips have a duration of about 200 sec, and around 12000 trips have the peak of durations which is around 550 sec. less than 2000 trips have a duration of less than 2000 sec.

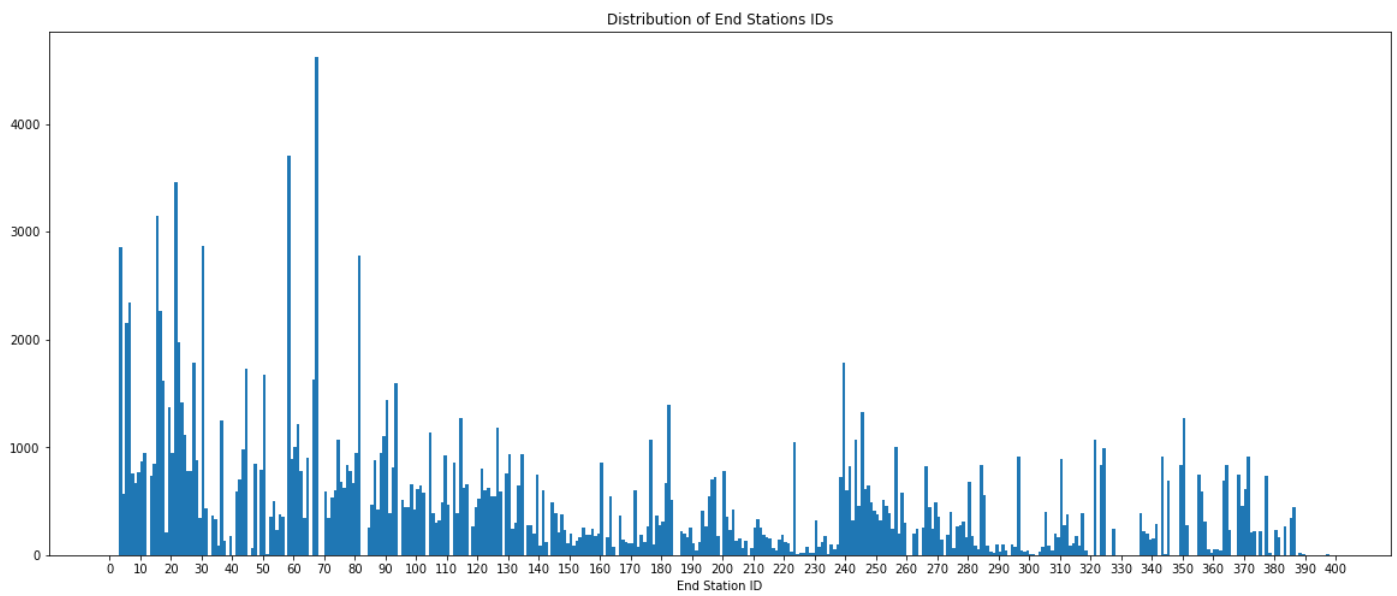


### 2- Distribution of Start and End Station Id vs Trips duration

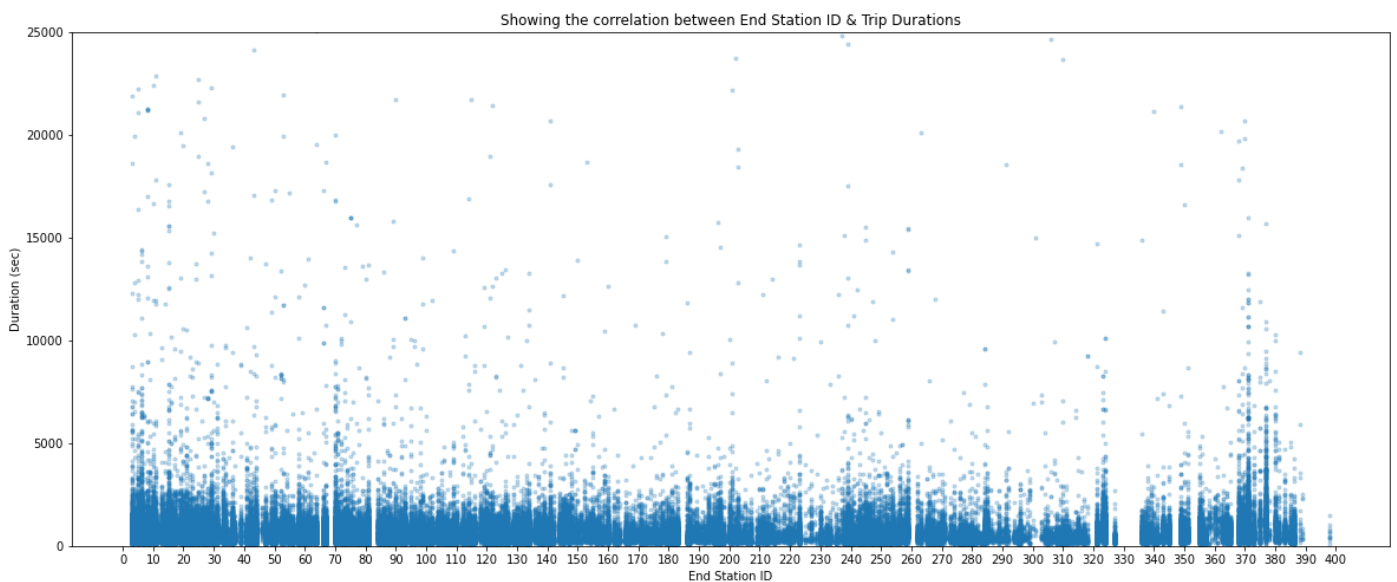
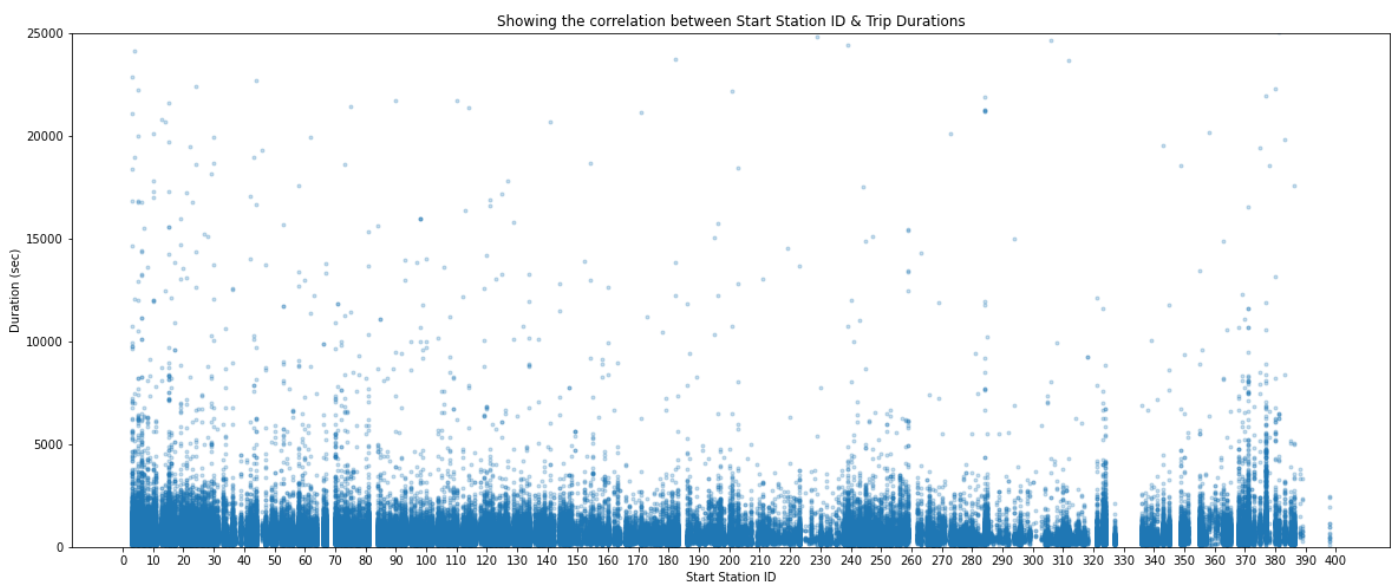
\* The station from which has the largest number of trips start from it is the stations with the ID of 8.



\* The station from which has the largest number of trips end in is the stations with the ID of 7.

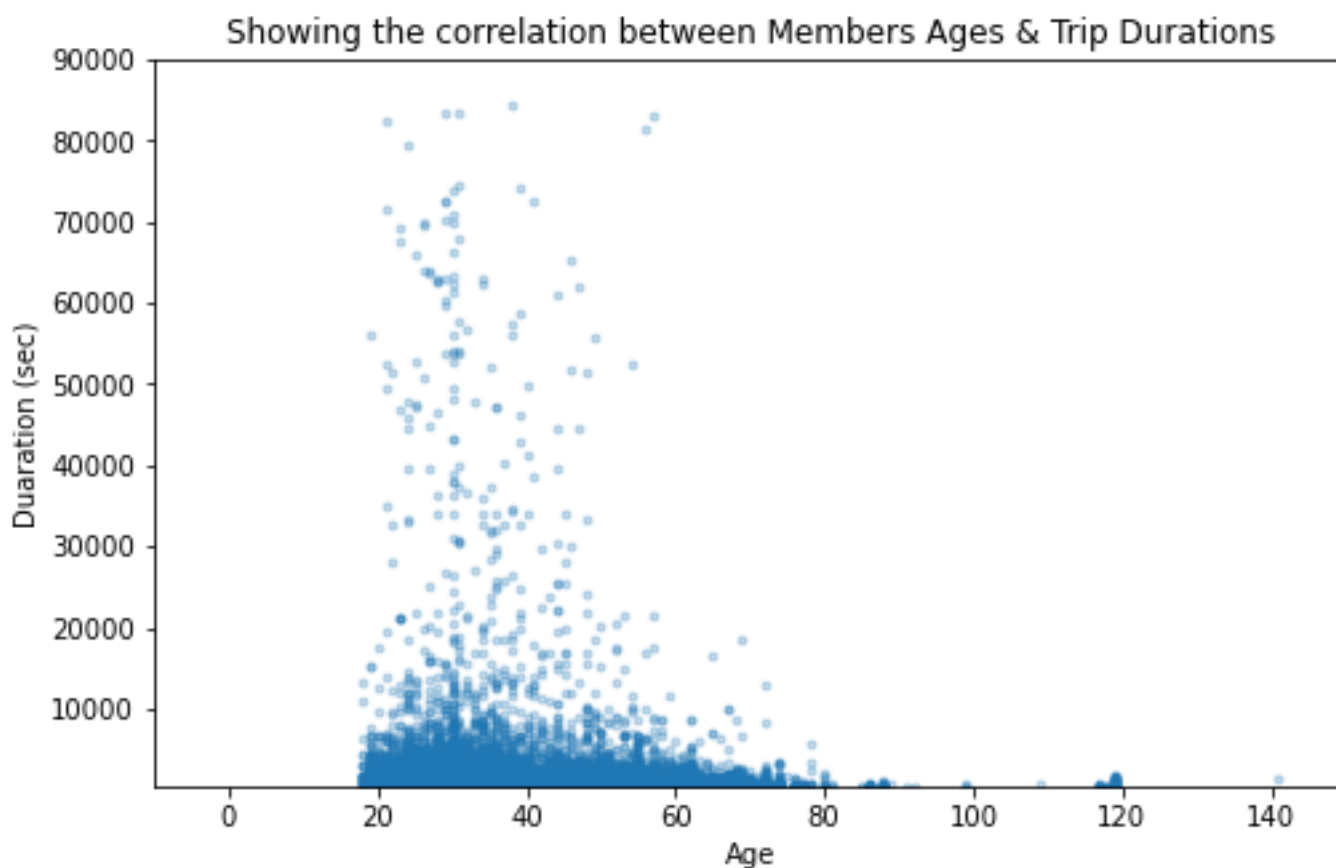
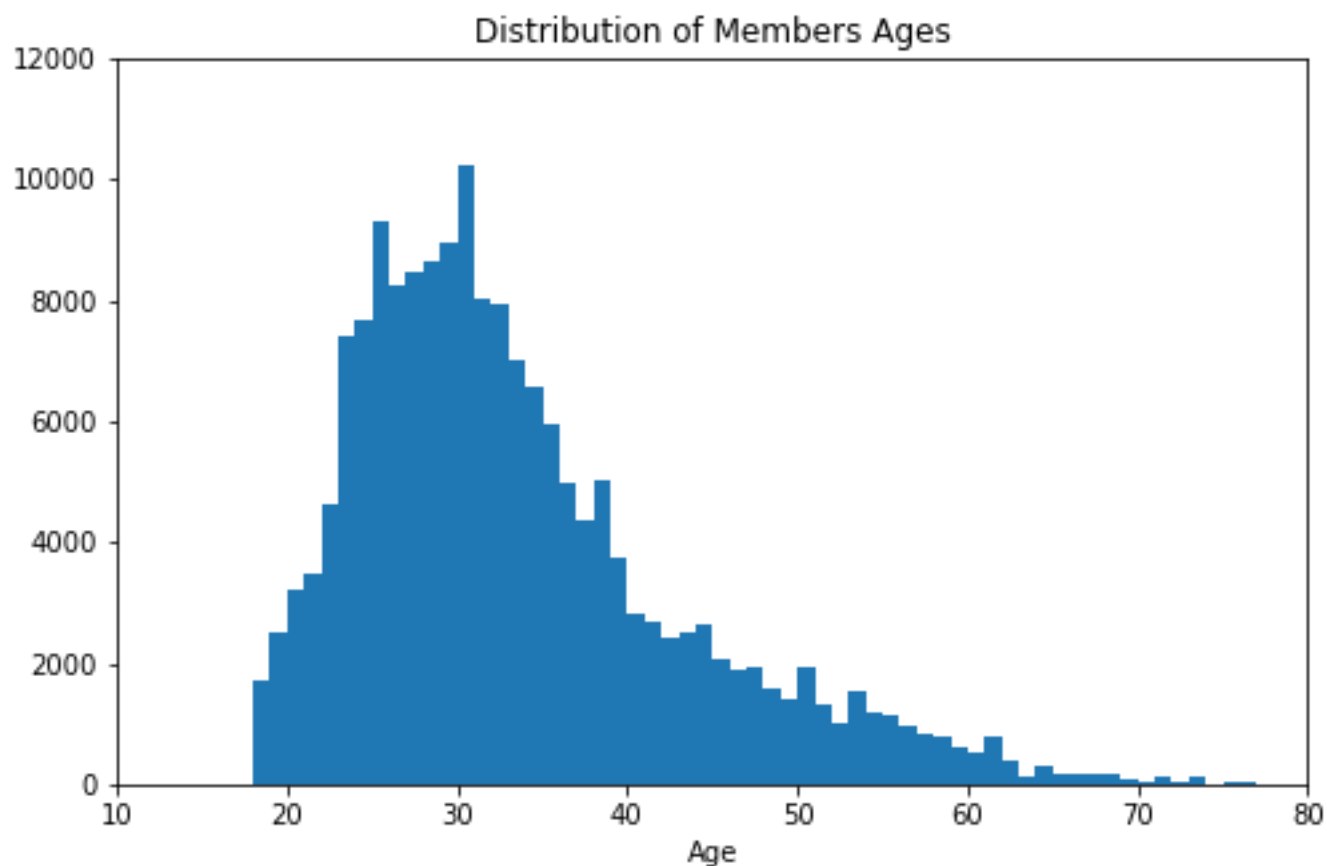


\* The most frequent users of bikes make a trip duration about 5200 sec for start and end stations.



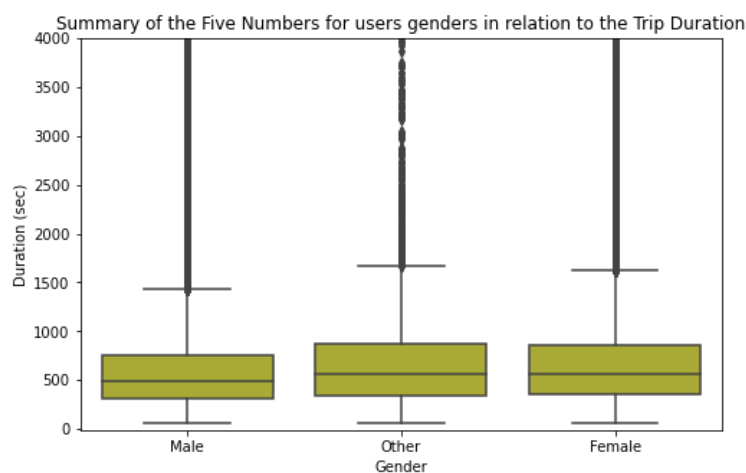
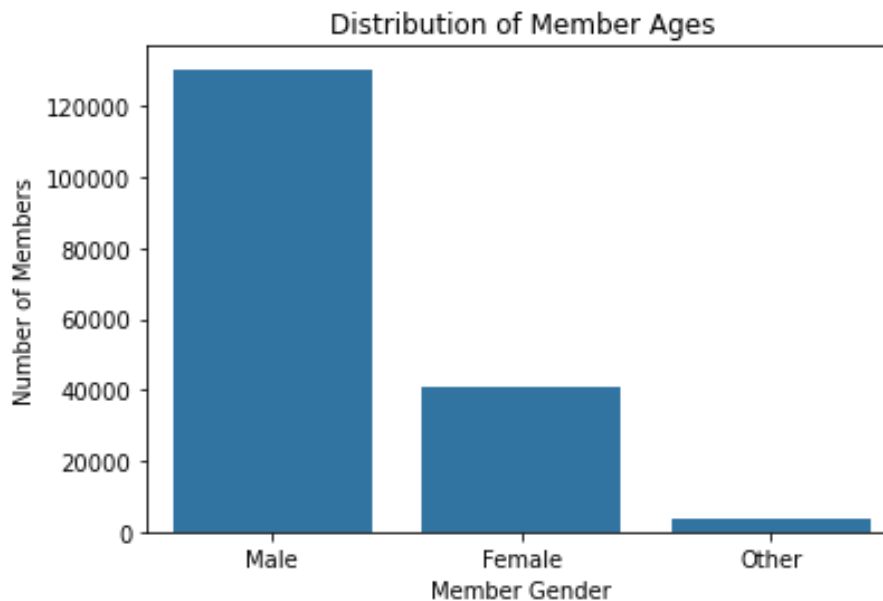
### 3- Finding the distribution of members ages vs Trips duration:

\* Most members are aged between 20 and 45, and most of the durations are below 6000 and age is below 80. Higher duration is clocked by younger members.



#### 4- Finding the distribution of members genders vs Trips duration:

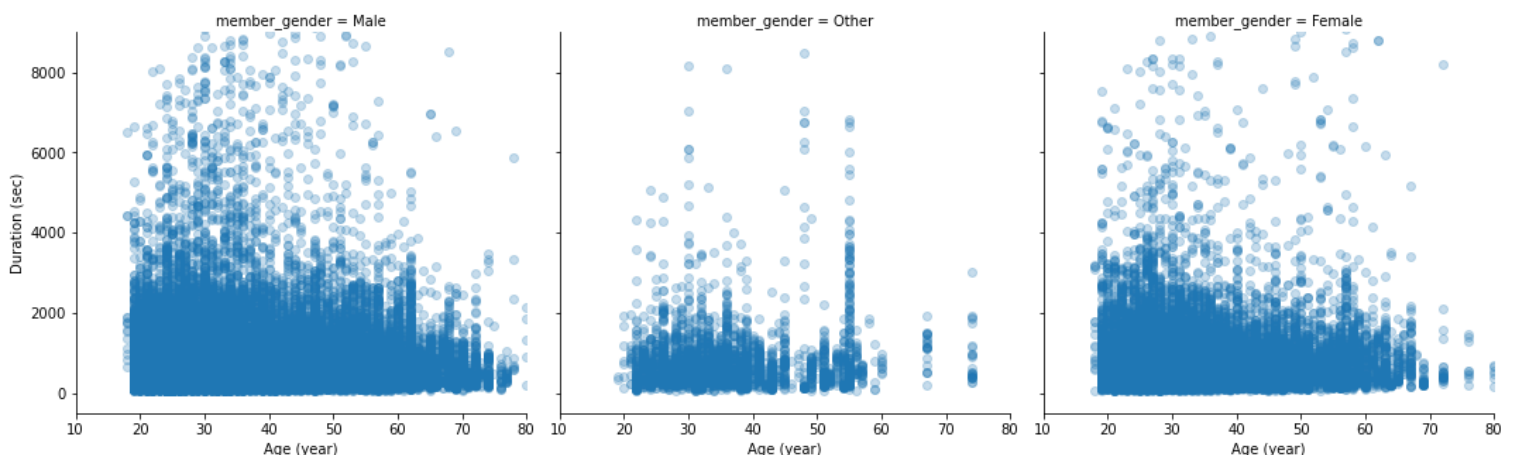
Most of Members are males and females are more than (other) genders. Though quantity of males riders are very high then other and females but we can see that higher percentage of female and other rides a longer mean of trips durations than males



\* We can see the correlation between Members Ages & Trip Durations in relation to users' genders:

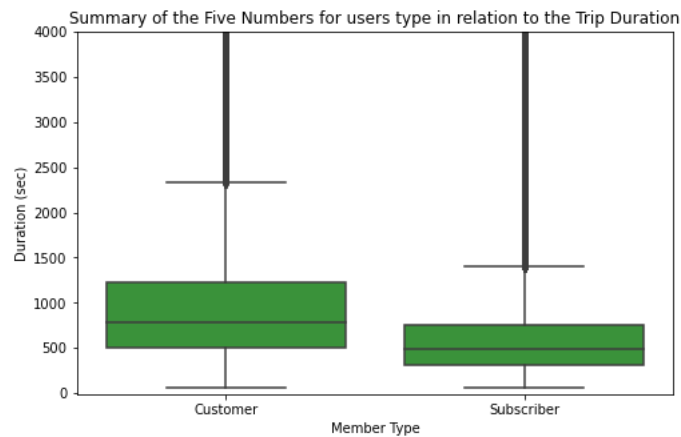
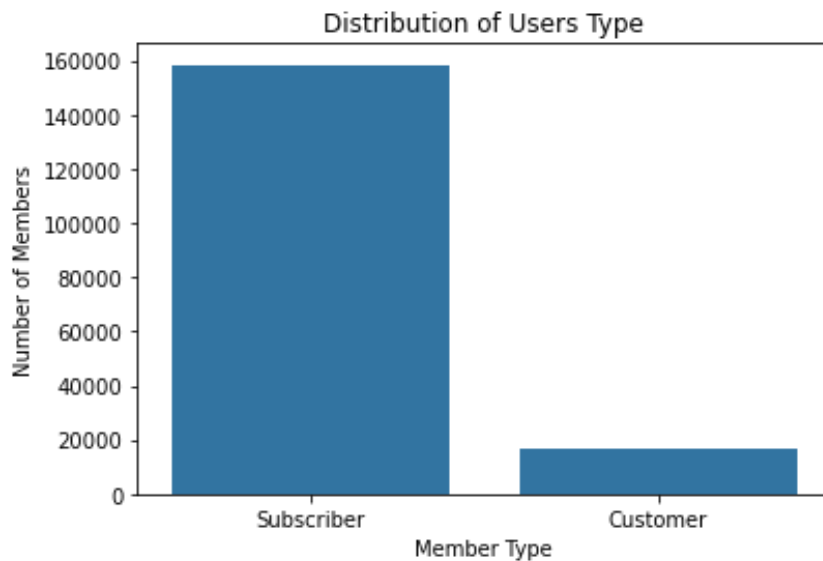
We can see a jump in duration for others at an older age which around 60 years.

The correlation between members ages and trip durations in seconds in relation to members genders



## 5- Finding the distribution of members Type vs Trips duration:

Most of Members are subscriber type. We also can see that higher percentage of customers are taking longer trips then compared to subscribers.



\* We can see the correlation between Members Ages & Trip Durations in relation to users' type:

Both Customer and Subscriber are showing similar trends for age and trip duration, but there is a little bit higher age for subscribers.

The correlation between members ages and trip durations in seconds in relation to members types

