Part II - (Ford GoBike system Data Exploratory Analysis and Visualization)

by (Amulya Kandukuri)

Investigation Overview

I am investigating the details of Bike trips in Ford GoBike dataset. I am exploring the details about the trips that happen during particular hour, time of day and weekday. I intend to compare trips taken by subscribers and customers in terms of duration in seconds and minutes. I intend to visualize the findings about longest trips, busiest trip taken hours, age of members whether that is customer or subscriber.
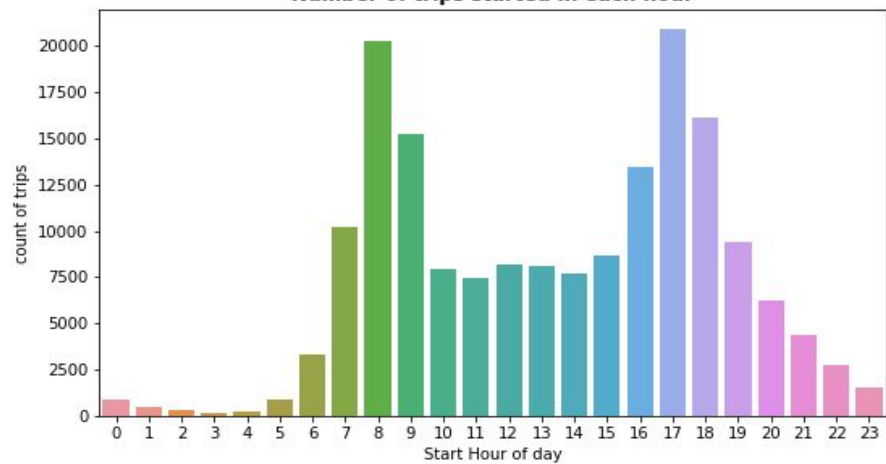
Dataset Overview

The dataset includes information about bike trips covering San Francisco area. This dataset includes information about various trips like duration_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, bike_share_for_all_trip. This dataset mainly focus on rides in February 2019.
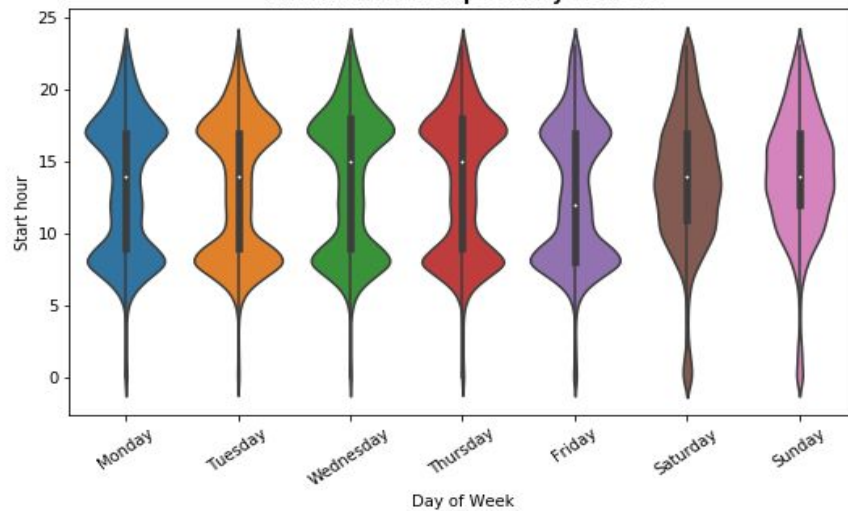
Distribution of Bike Trips in day hours

- There is a peak in number of trips taken at 17 th hour and 8th hour of day with more than 20000 times.
- The next most trips are at 9th hour and 18th hour of day with more than 15000 times.
- The least number of trips were taken during 3rd and 4th hours of day with less than 1000 trips
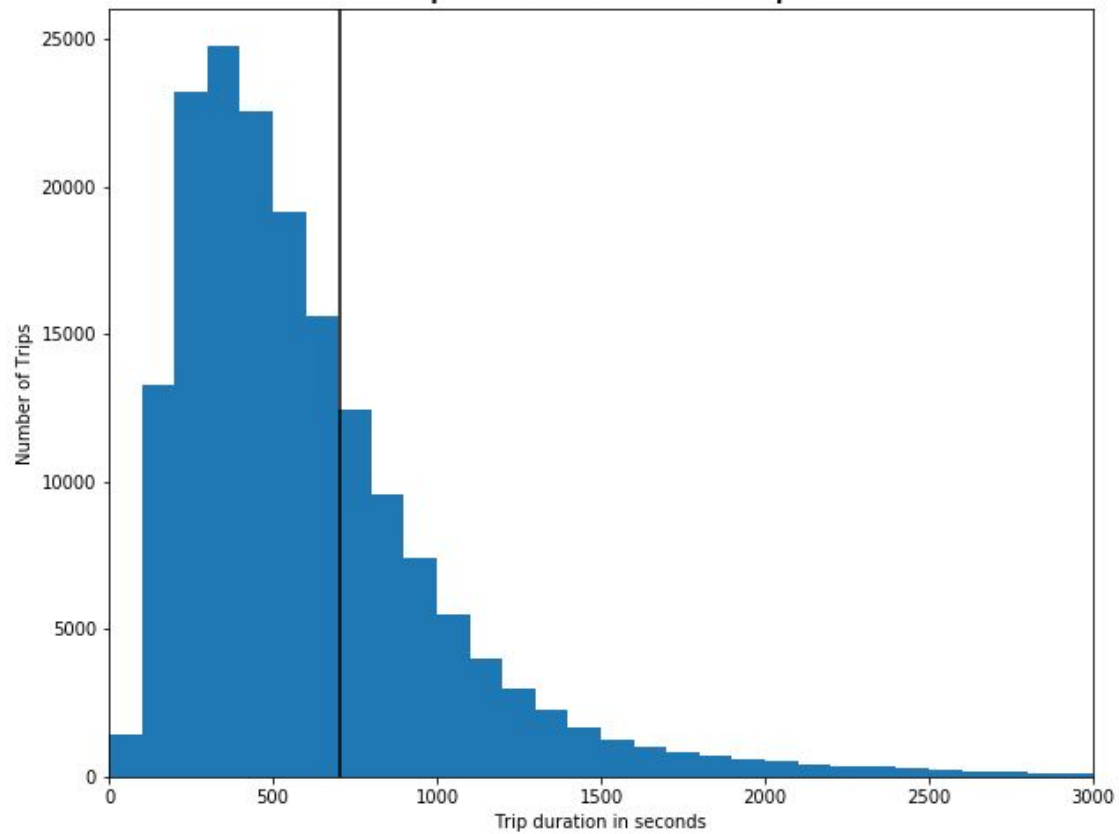
Number of trips started in each hour

Distribution of Trips in Day of week
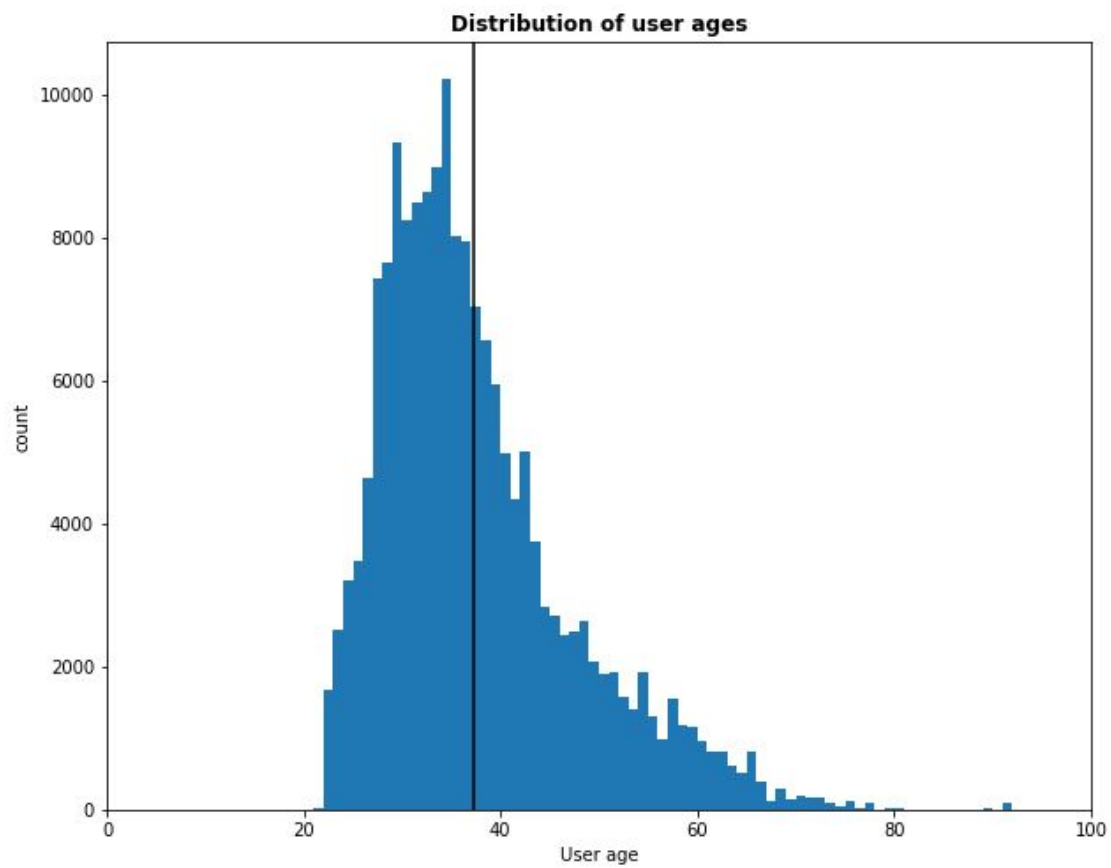
Trip duration and Number of Trips

- The histogram is right-skewed with most number of trips below 726 seconds.
- The most number of trips has duration less than 700 seconds.
- There are approximately 25000 trips with duration around 400 seconds.

Trip duration and Number of trips

Number of Trips (y-axis)
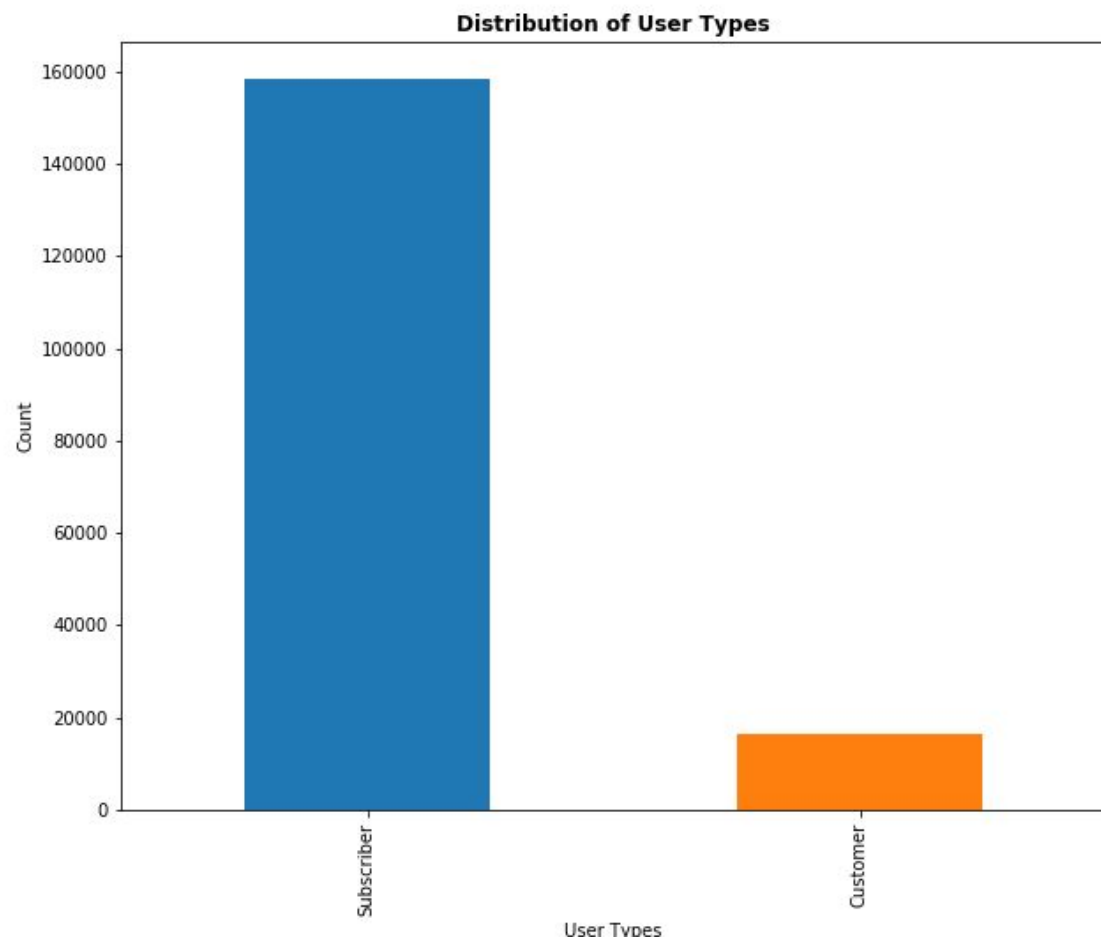
Trip duration in seconds (x-axis)

Distribution of age among members

- The most number of members are in the age 34 years with more than 10000 members.
- The average age of members is 37 years. The minimum age of the members is 21 years.
- The least number of members are in the age above 65 years with less than 1000 members in each year.
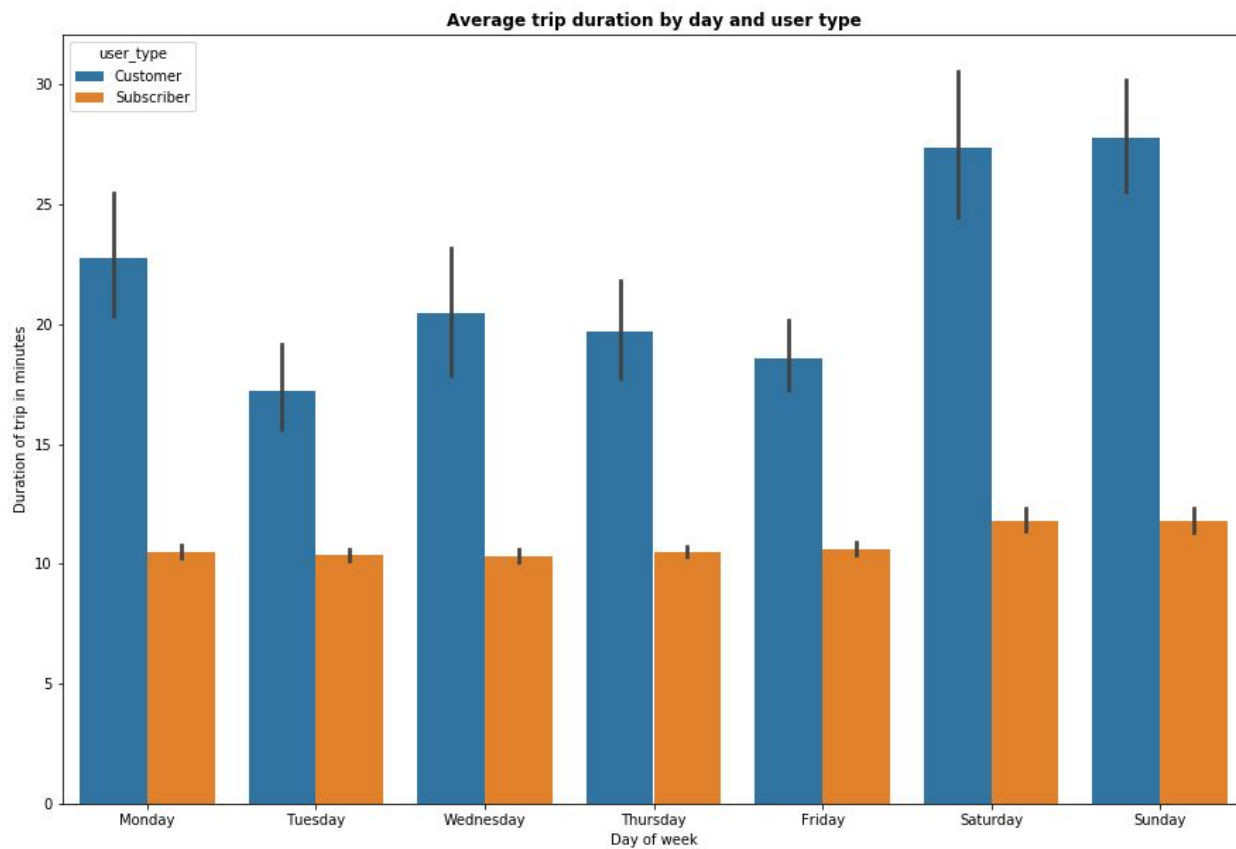
Distribution of user ages

Distribution of User Types

- There are most number of trips occurred due to subscribers which are approximately 160000.
- The number of customers are below 20000.

Distribution of User Types

Average trip duration by day and user type

- Most trip duration occurred during Saturday and Sunday at 25-30 minutes with a standard deviation little more than 30 minutes by customers.
- Most trip duration occurred during Saturday and Sunday at 10-11 minutes with a standard deviation little more than 10 minutes by subscribers.
- The least trip duration is on Tuesday by customers with 15-17 minutes trip duration.
- The least trip duration is on Tuesday by subscribers with 10 minutes trip duration.

Average trip duration by day and user type

Conclusions

- Busiest trip hours are dependent on day of week.
- Most of the customers does trips during Saturday and Sunday.
- Largest trip duration occurs with members of age 41 years.
- Thursdays have highest number of trips.
- Subscribers number is very high when compared to customers.