Bay Wheels - Find Patterns in Bike Ridership

by Shilpa Madini

Investigation Overview

> In this investigation, I wanted to look at the Bay Wheels trip data and analyse how ridership pattern is affected by the subscription type and day of the week. Data set I used for this anlysic can be found at Bay Wheels System Data (https://www.lyft.com/bikes/bay-wheels/system-data)

Dataset Overview

Structure of the dataset?

> I have used 2017 year data for this analysis.There are 519700 observations with 13 features.

Each trip is anonymized and includes:

1. Trip Duration (seconds)
2. Start Time and Date
3. End Time and Date
4. Start Station ID
5. Start Station Name
6. Start Station Latitude
7. Start Station Longitude
8. End Station ID
9. End Station Name
10. End Station Latitude
11. End Station Longitude
12. Bike ID
13. User Type (Subscriber or Customer – “Subscriber” = Member or “Customer” = Casual)

Following features are engineered from the original dataset

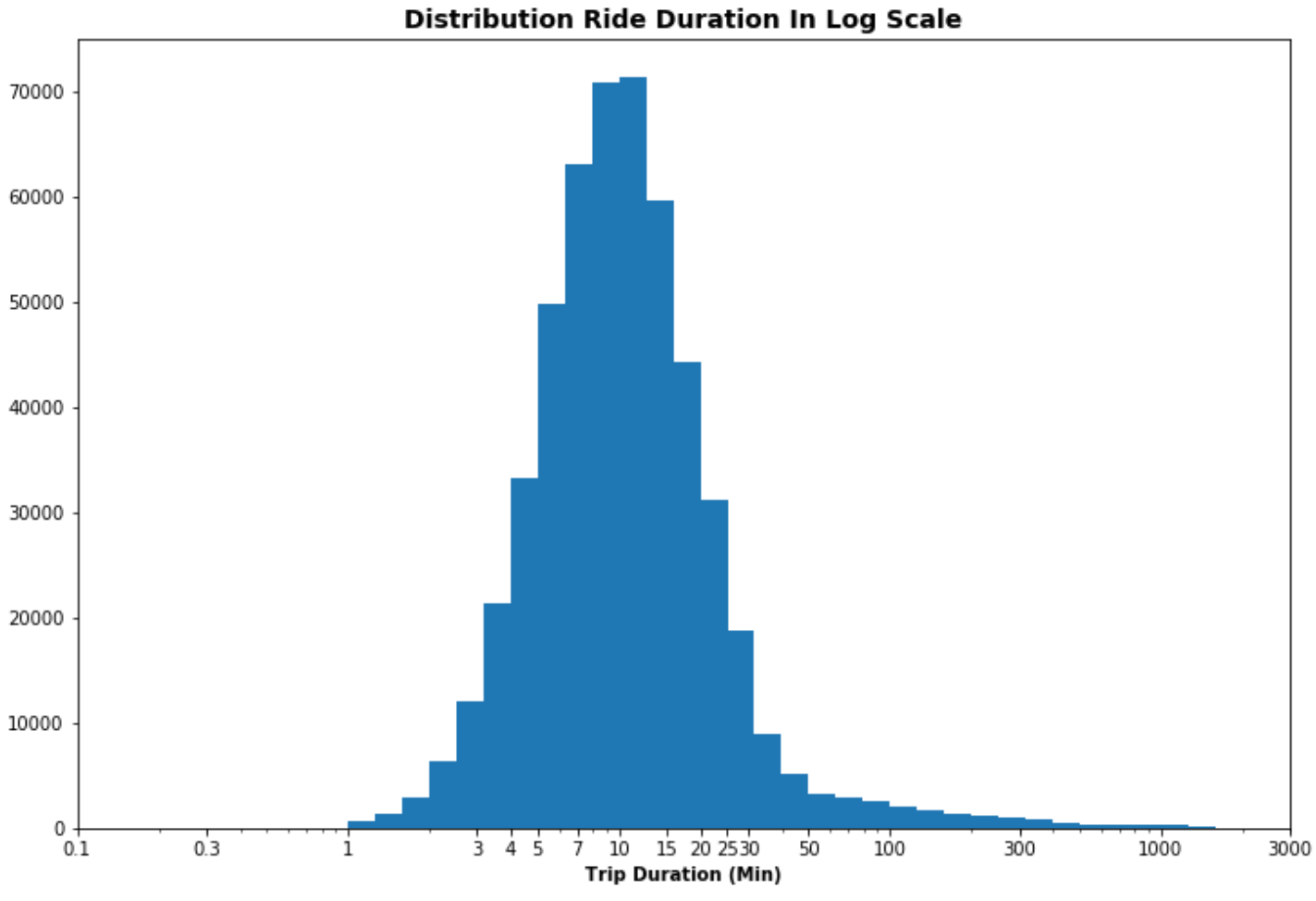
1. month (month of the year,extracted from start time and Date)
2. weekday (day of the week, extracted from start time and Date)
3. duration-min (calculated using duration-sec, represents the duration in minutes)

Main feature(s) of interest in the dataset?

1. When are most trips taken interms of day of the week, month of the year?
2. Top 10 station names where most trips are originated?
3. How long is the average trip taken?
4. Significance of user type on the number of rides and average duration?

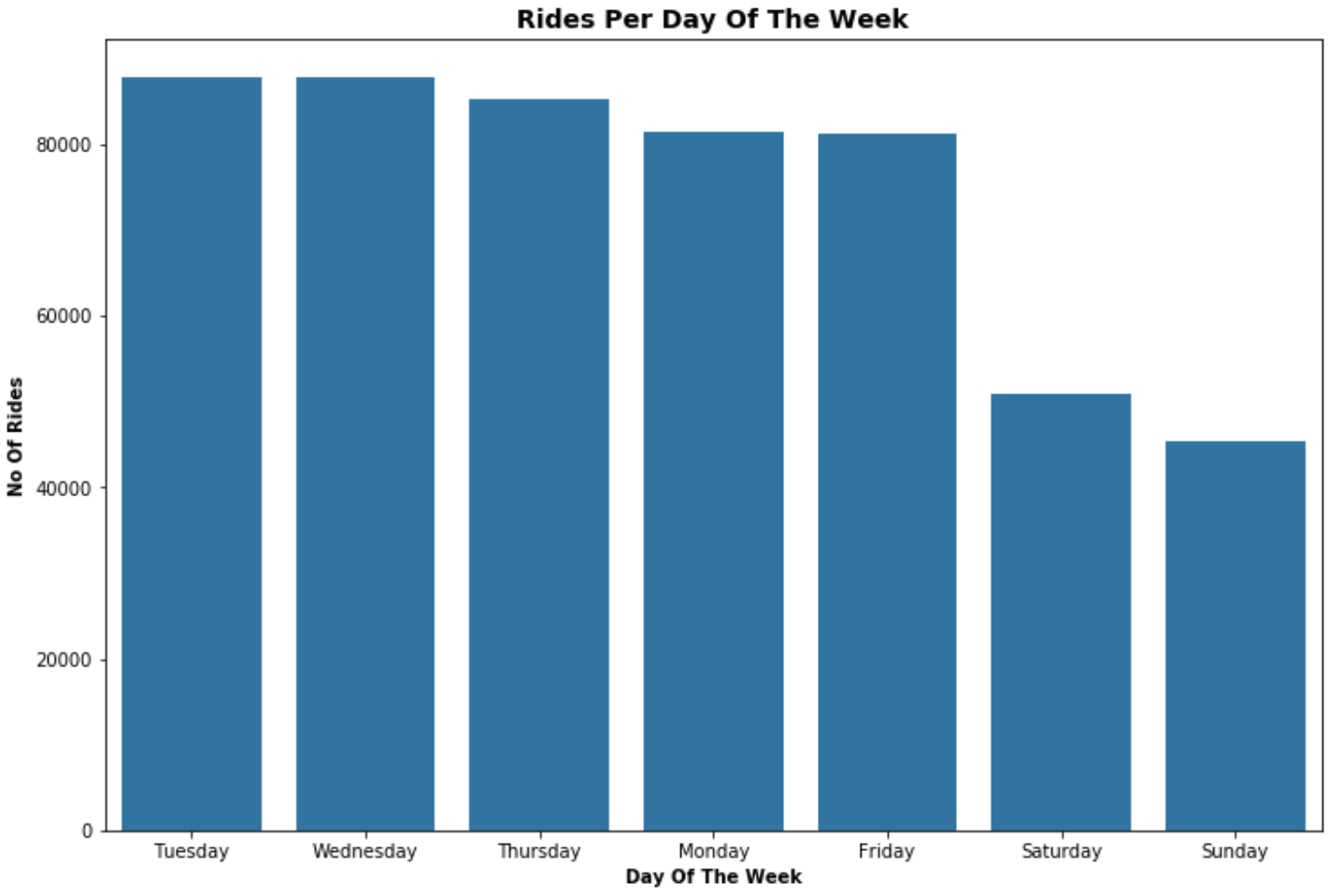
Distribution of Ride Duration

ride duration in the dataset takes on a very large range of values, from 1 min 1439 mins. But most of the data is less than 200 mins. After plotted on a logarithmic scale, the distribution of ride duration takes on a normal distribution with slight skew to the left.



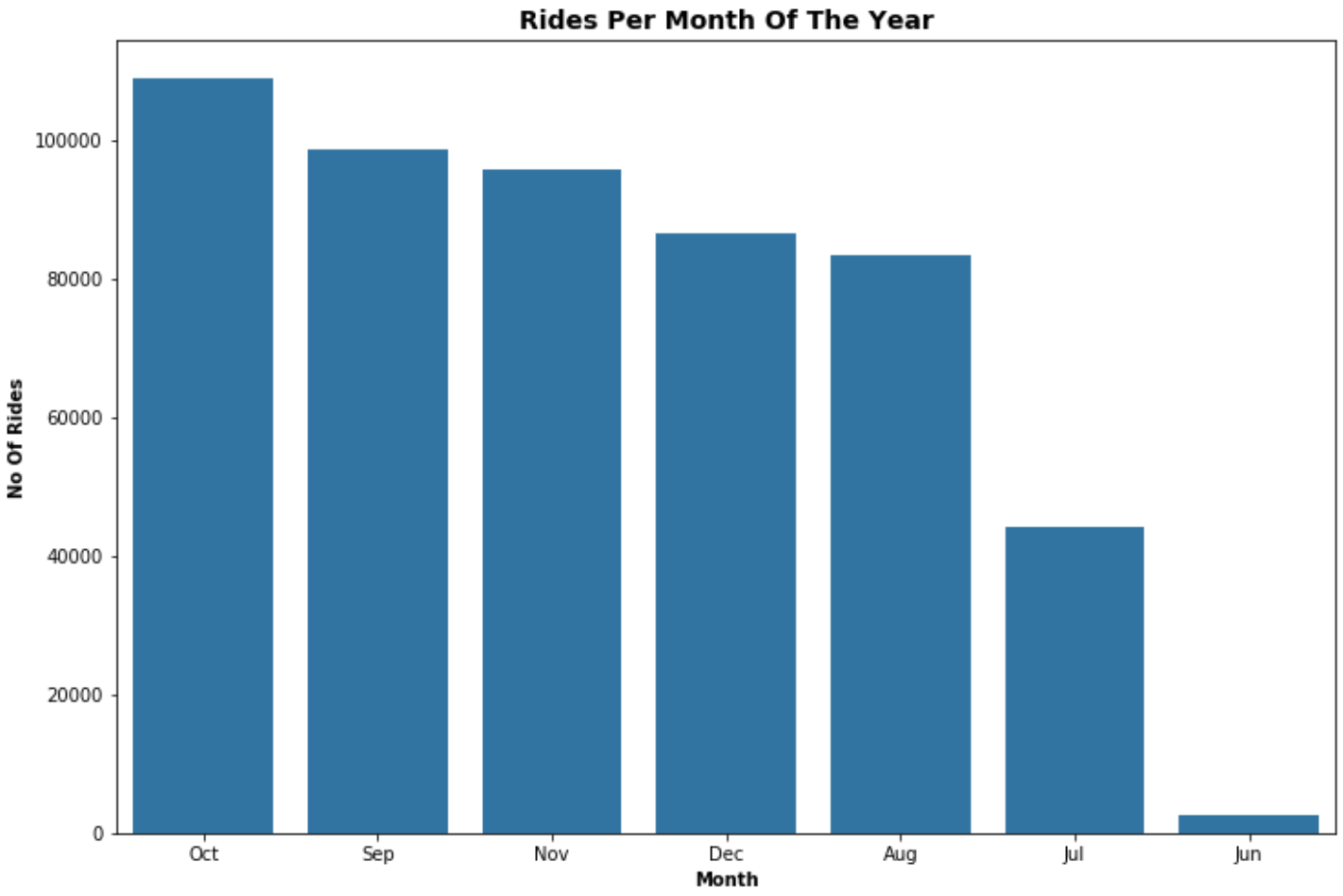
When are most rides taken in terms of day of the week?

This plot shows that most trips are taken on a tuesday. Weekdays seem to have more number of trips than weekends.



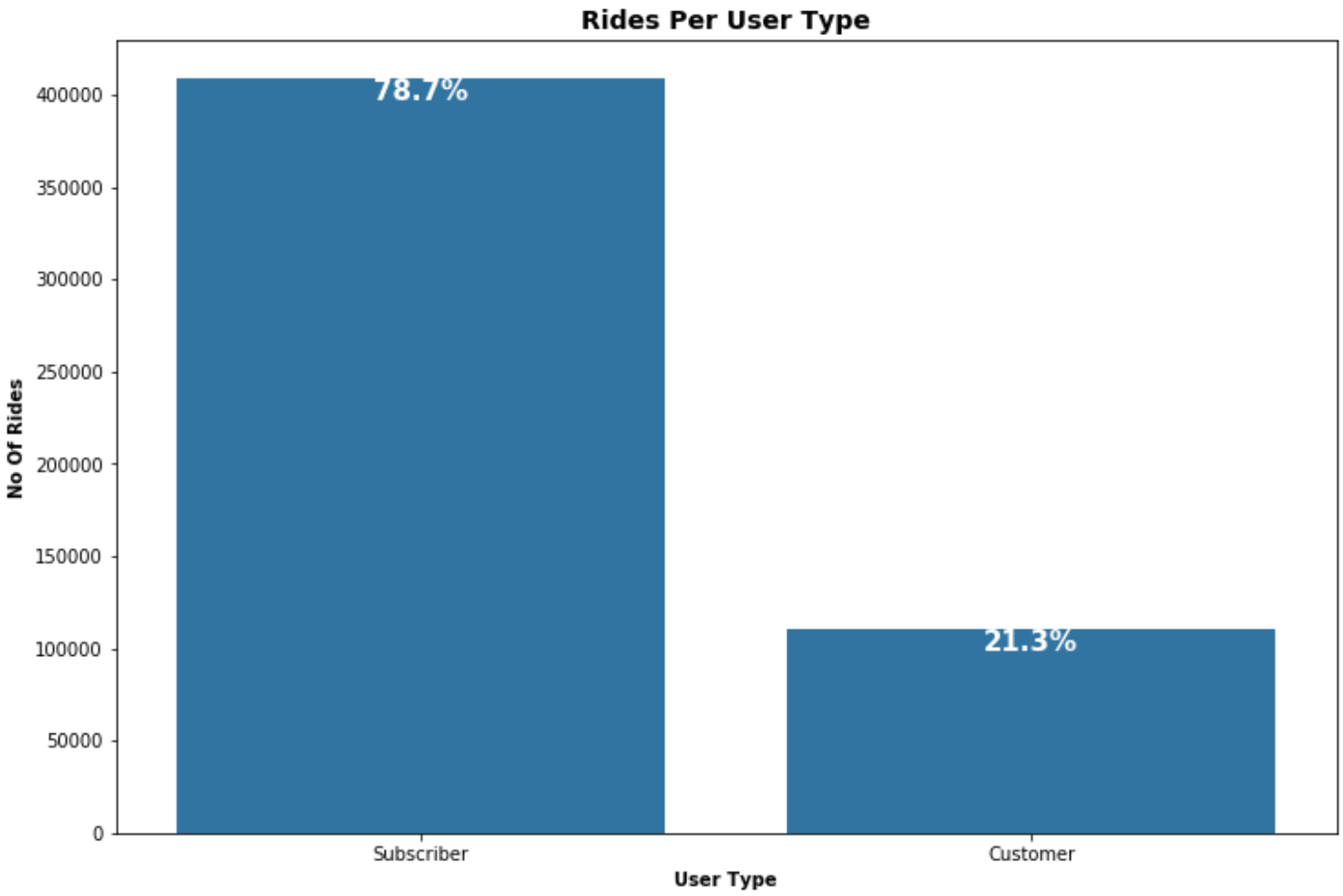
When are most rides taken interms of month of the year?

October is the month where most number of trips are recorded.



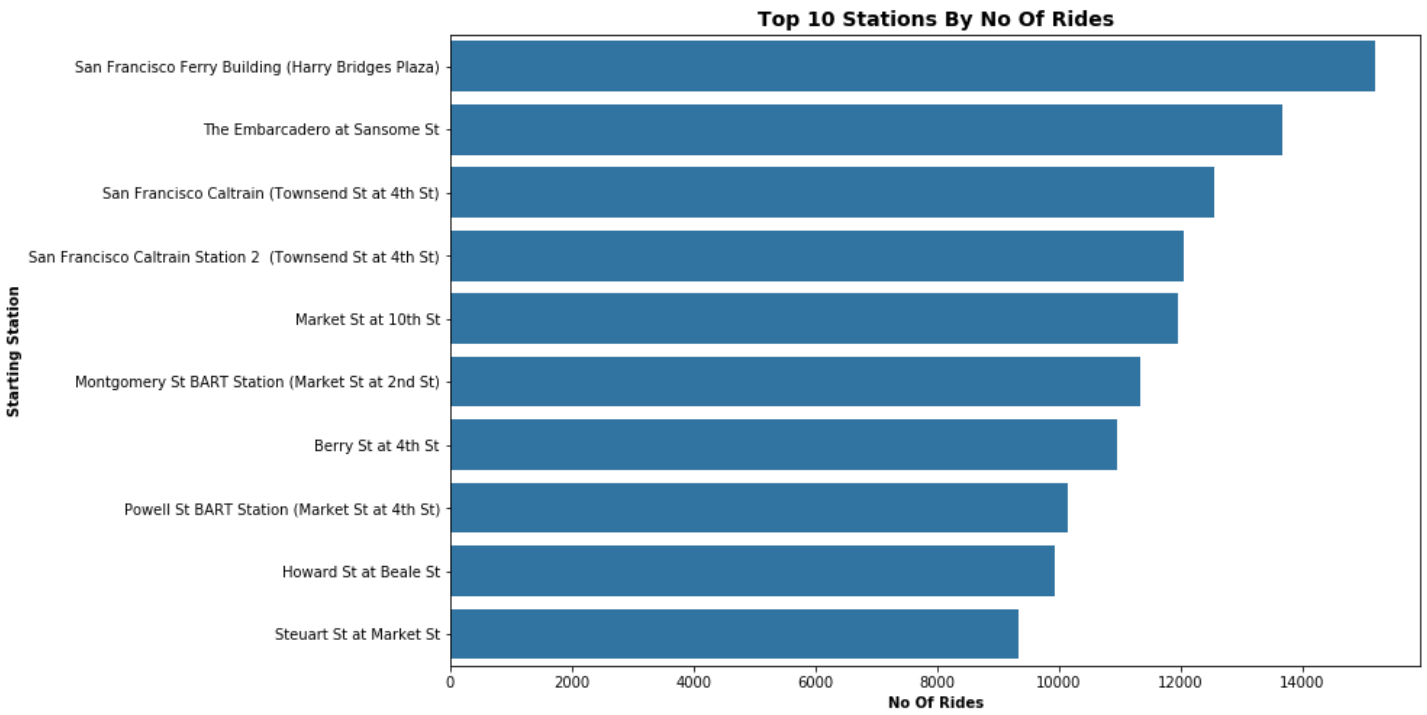
Which User Type Has Most Number of Rides?

Subscribers have the most number of rides than customers.



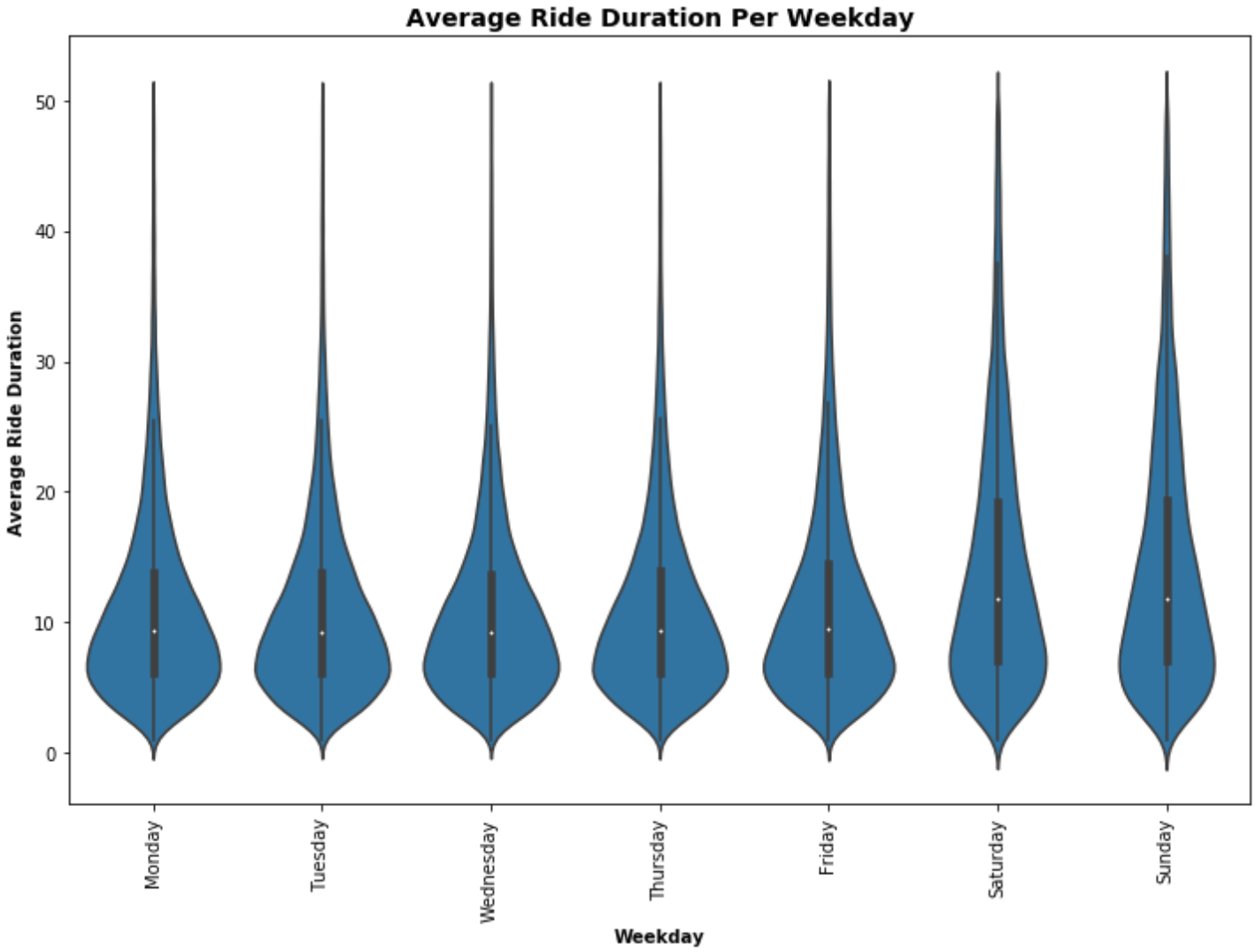
Top 10 station names where most trips are originated?

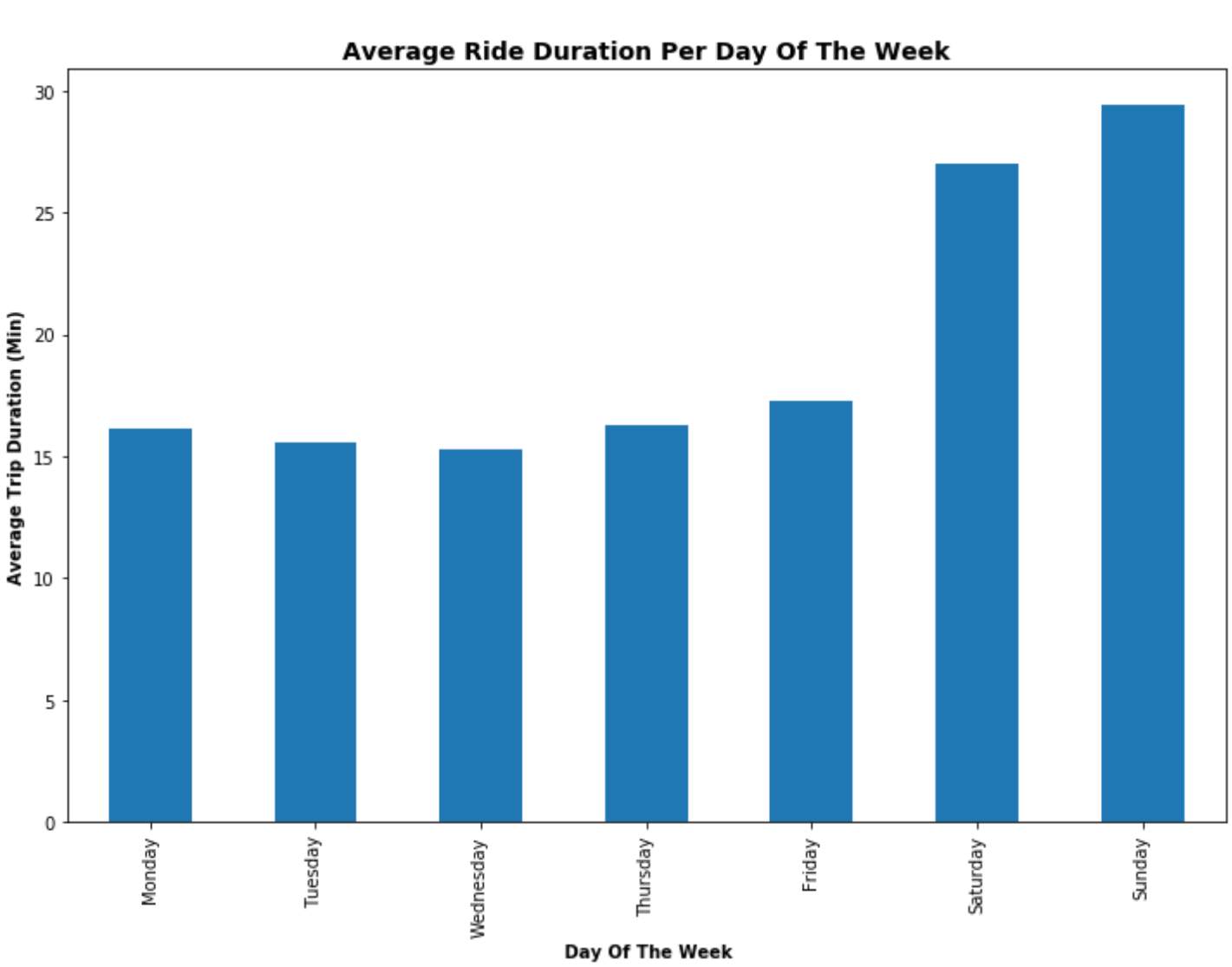
Plot below displays the top 10 stations where most of the trips are originated. This is an interesting obervation along with most rides being on weekdays. Bikeshare company can dispatch more bikes rentingstations at these staions to market more at these stations to attract more customers.



Average Ride Duration by Day Of The Week

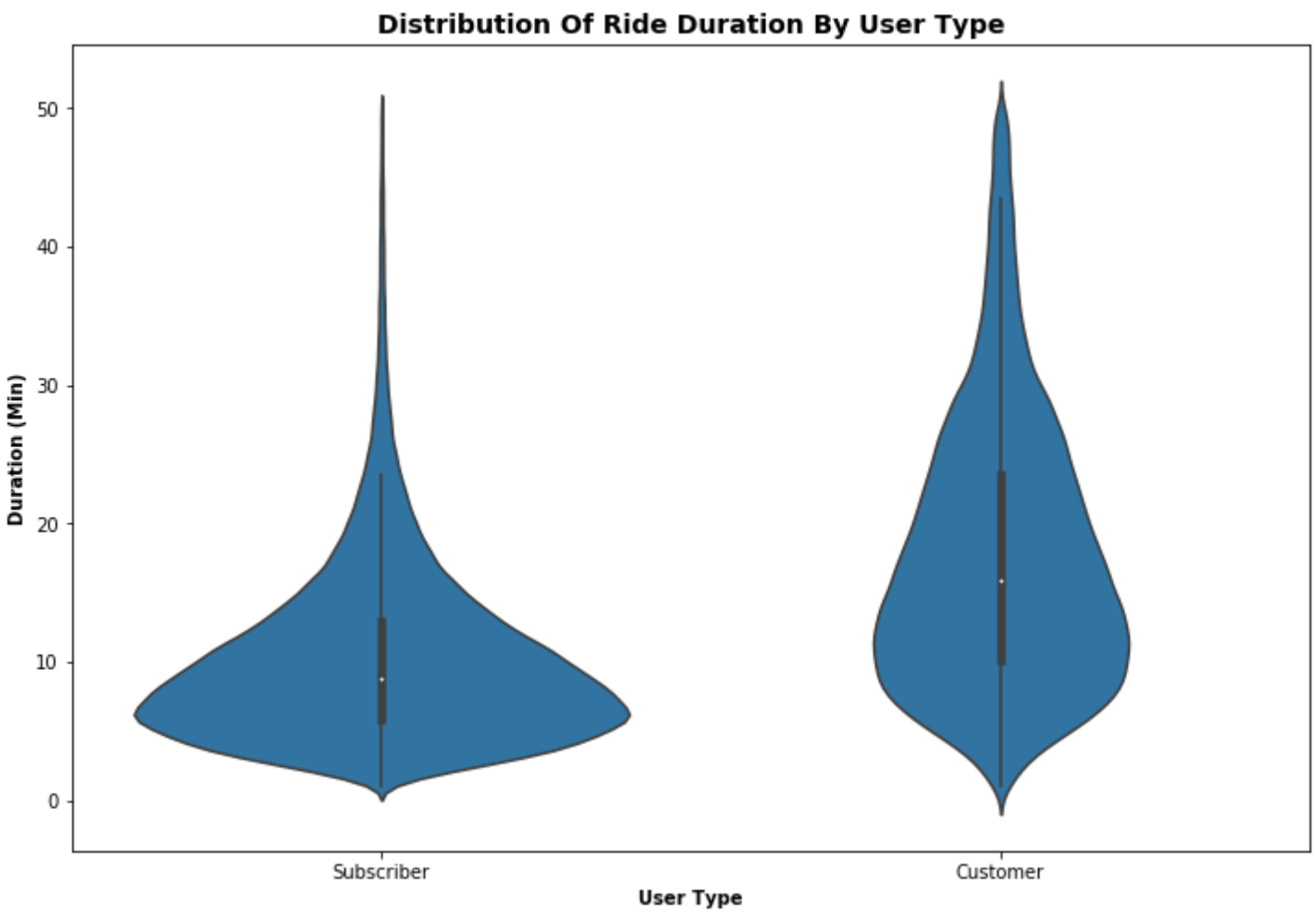
Average trip duration on weekdays (Mon to Fri) is pretty much consistant and value ranges between 15mins to 17 mins Average trip duration on weekends (Sat and Sun) ranges between 20 to 30 mins


Average Ride Duration Per Weekday
Average Ride Duration
50
40
30
20
10
0
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Weekday


Average Ride Duration Per Day Of The Week
Average Trip Duration (Min)
30
25
20
15
10
5
0
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Day Of The Week

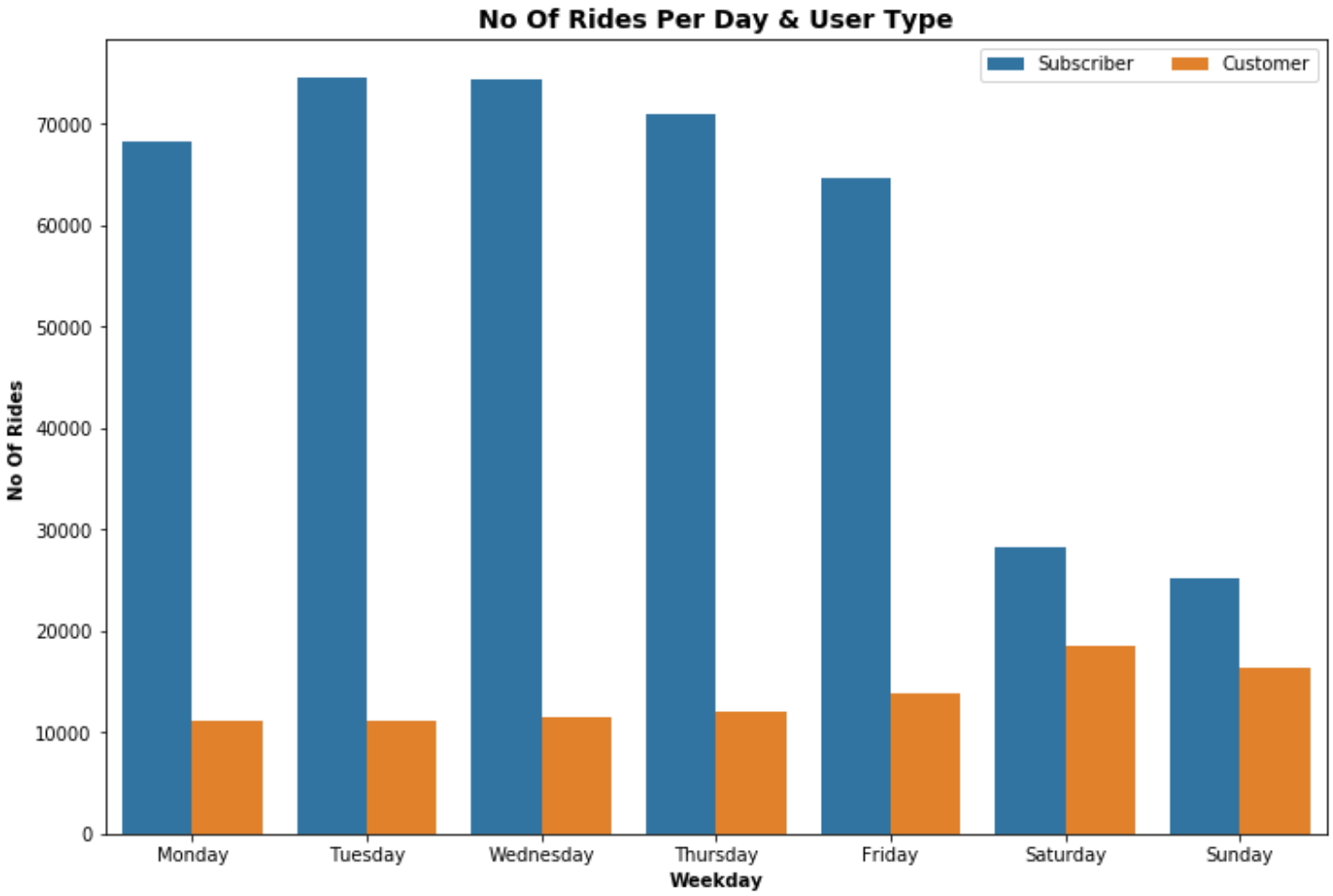
Average Ride Duration By User Type

Customers have longer ride duration than subscribers



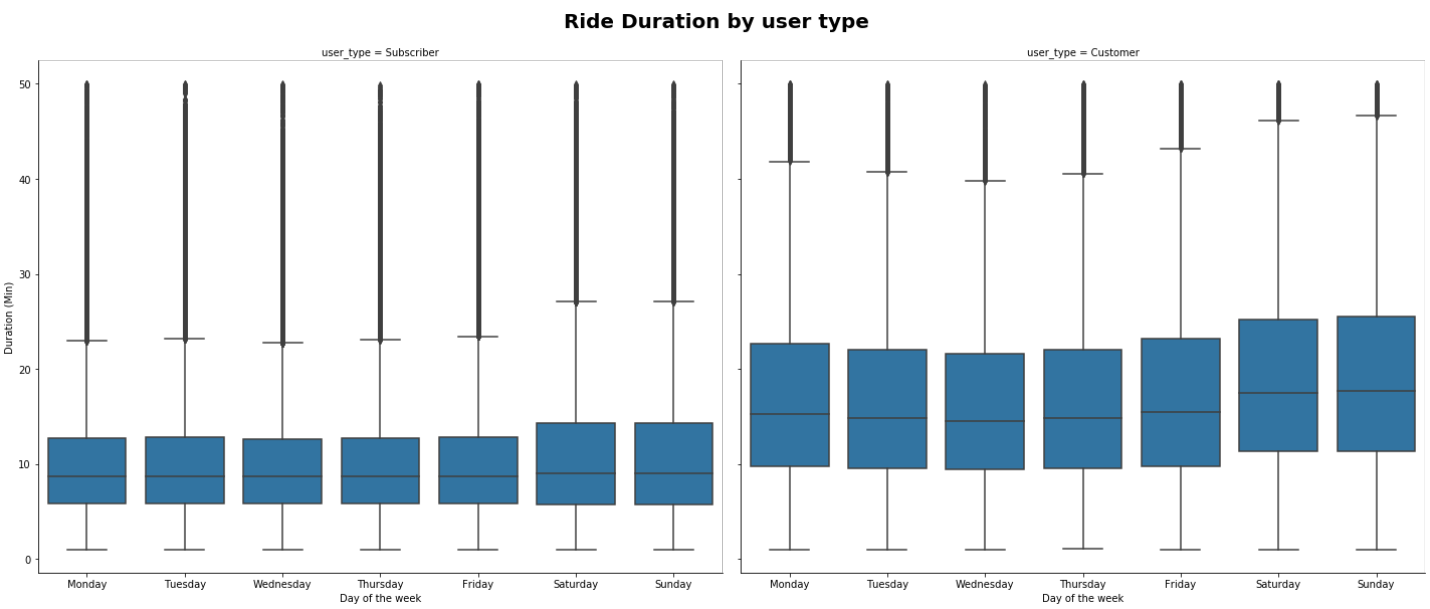
Number of rides per user type and per weekday

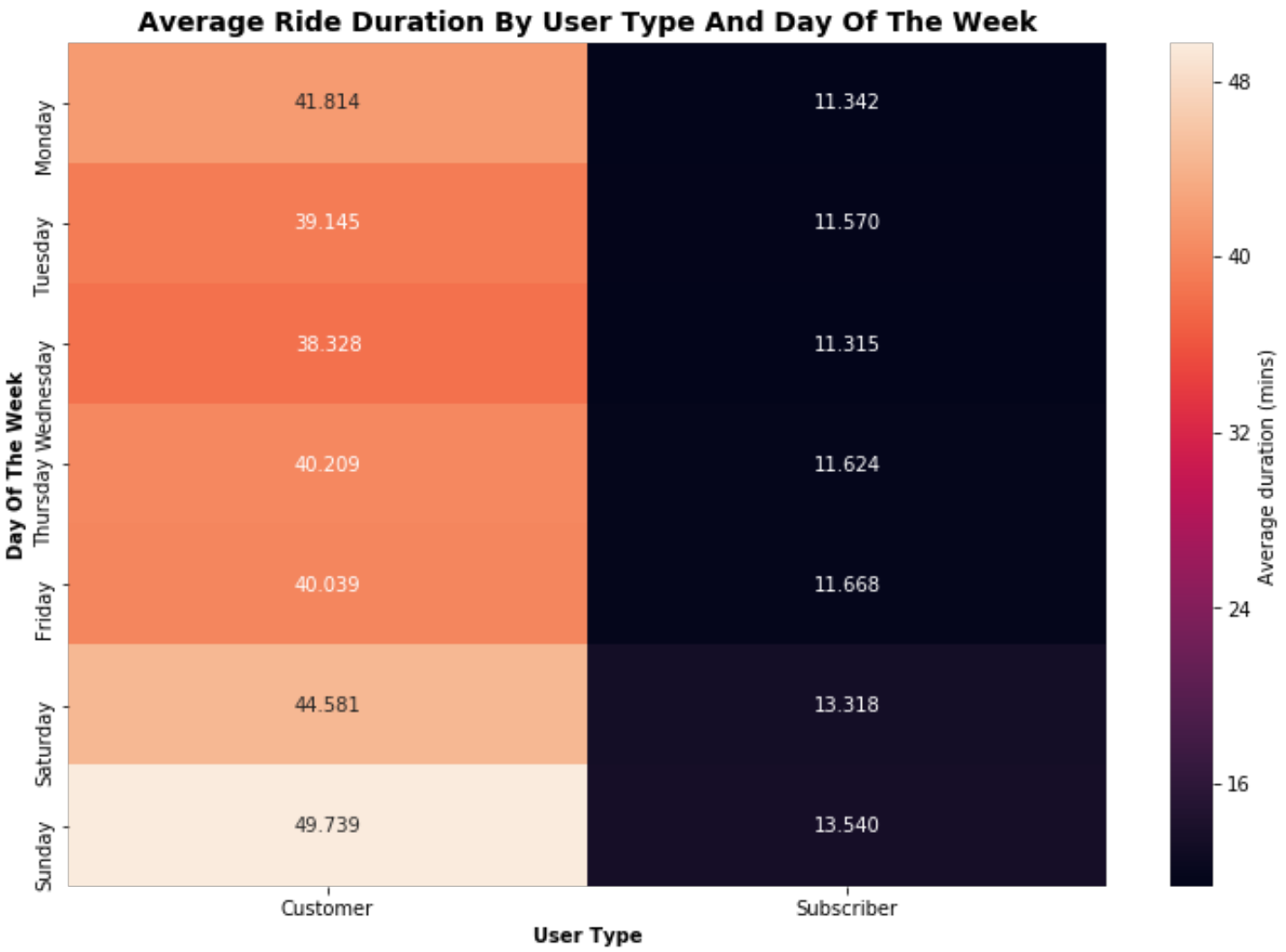
subscribers ride mostly on weekdays and customers mostly on weekends. Subscribers on weekdays record more number of rides



Average ride duration by weekday for each user type

This plot shows that Average trip duration depends on the user type. Customers although have less number of rides than subscribers in the rider ship they tend to have longer duration of ride.





Based on the above plots we can establish that average trip duration and number of rides vary by cutomer type and their usage. Subscriber have more trips and tend to have shorter duration of the trip. They also use bikes mostly on weekdays. Weekdays are typically working days and subscribers may be mostly using the bikes for commuting. Hence these trips tend to be shorted in duration. Customers have less trips than subscribers but they ride mostly on weekends. Average trip duration for these users is also high compared to subscribers. This is different pattern of consumption of bike ridership than subscribers. Customers may be using the bikes to explore the city and these activities late longer time.