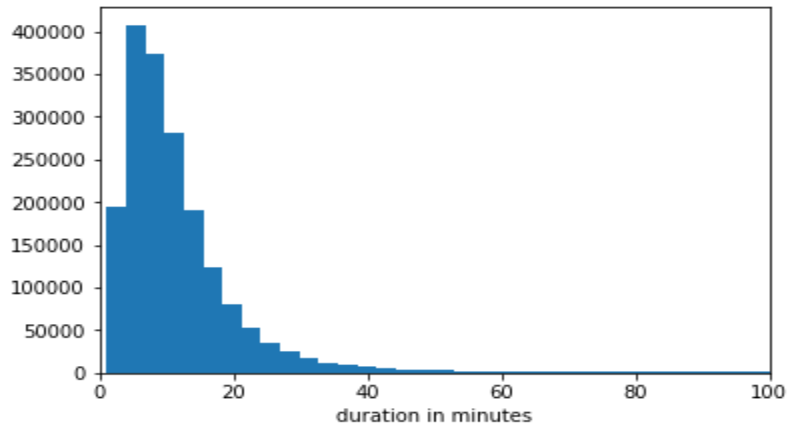
Lyft Systems Data

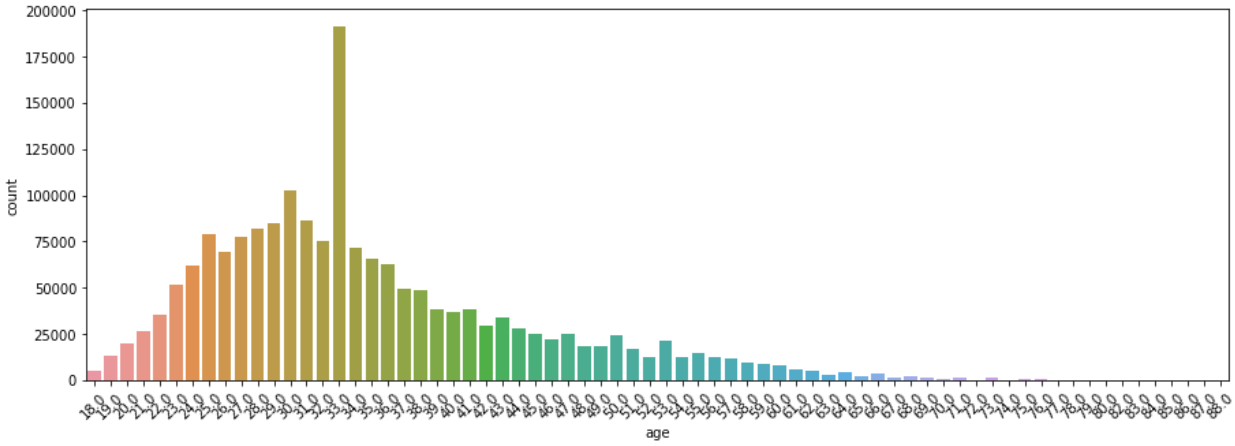
Below is the exploration of the data that was retrieved from Lyft's bikeshare data was cleaned .

This graph shows the **Distribution of the ride's duration in minutes.**



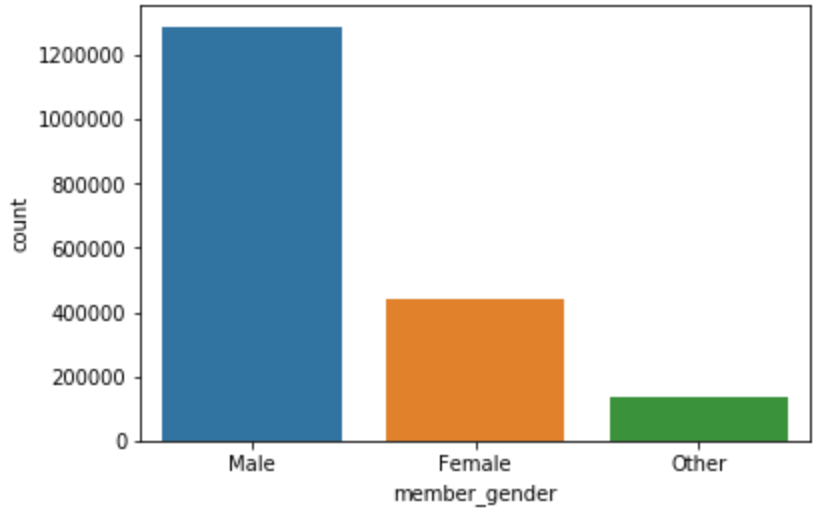
From the graph we can see that the duration in minutes is right skewed. Most of the riders rode the bike from between 0-20 minutes.

This graph shows the **Distribution of the rider's Age.**



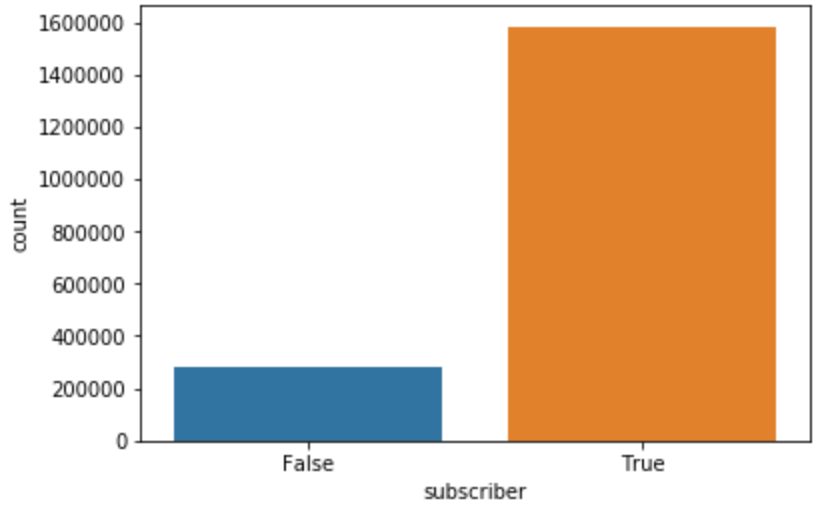
From the graph we can see that the distribution of age is right skewed. However there are many more 33 year olds than any other ages. Could this possibly mean that that 1985 was a default year that was used when users had to apply for the bikes?

This graph shows the **Distribution of genders** in the bikeshare data.



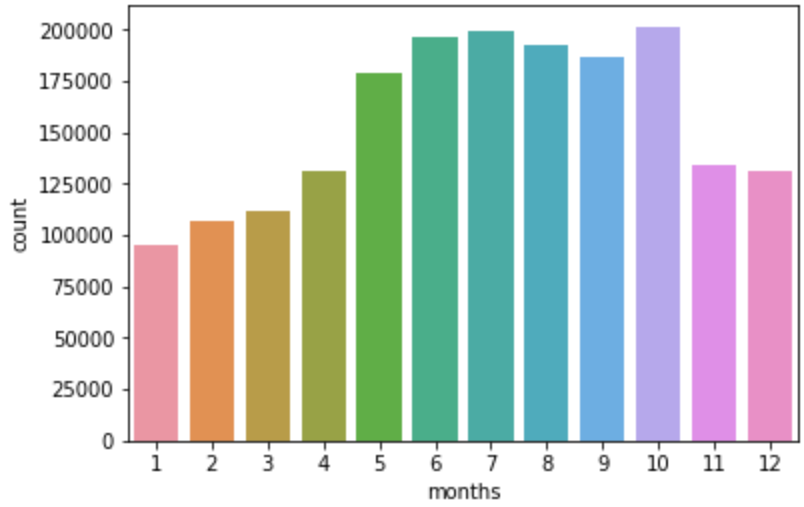
From the data we can see that there were many more males than females. There were some users who preferred not to put their gender, which is indicated by ‘Other’.

This graph shows the **distribution of the subscribers**



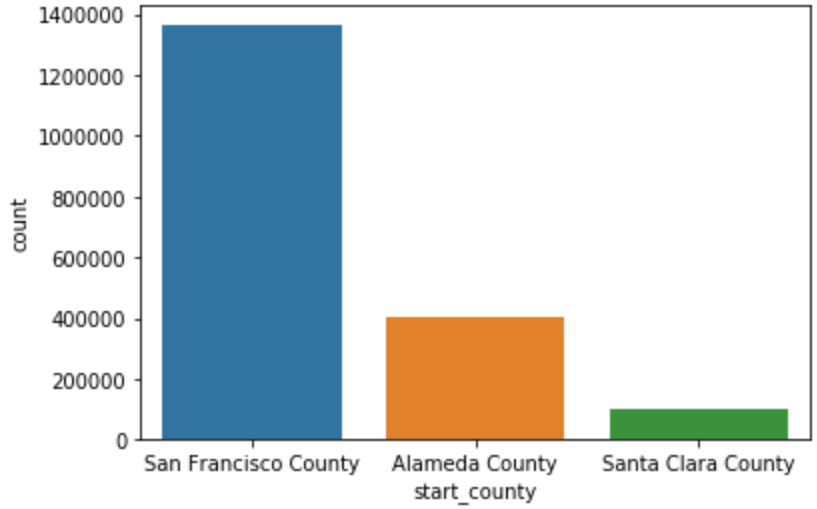
False means not a subscriber, True means is a subscriber. We can see that most of our users are subscribers. There are almost 5x more subscribers than non subscribers.

This graph shows the distribution of **riders that ride in each month**



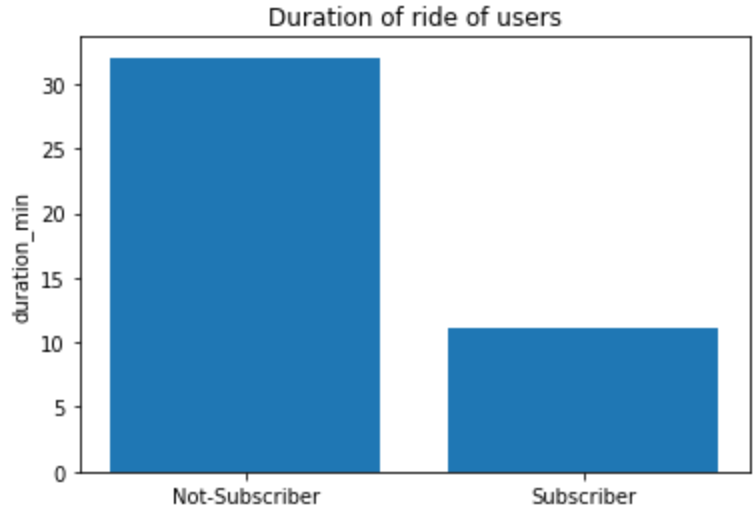
From the graph we can see that most of our riders ride in may, june, july, august, september, october. We can see a drastic decrease in riders in the other months, where the temperature is usually much colder.

This graph shows the distribution of the **amount of users in each county.**



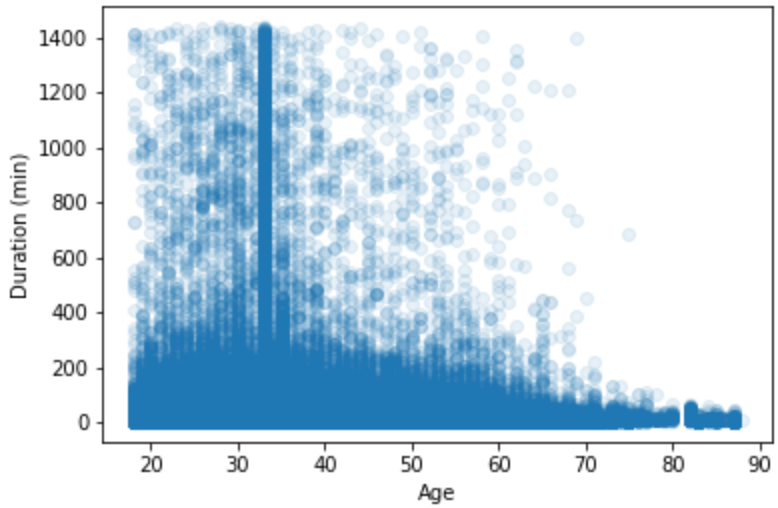
This graph shows the county of where the user started their bike ride. In the graph we can see that most of the users start from San Francisco County. Note that some county's were left out as users were less than 1000.

This graph shows the **User-Types & Duration of their rides.**



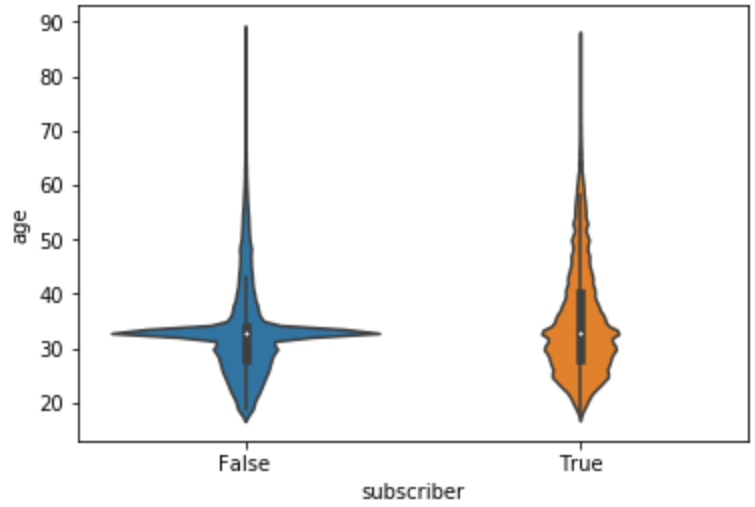
We can see that non-subscribers tend to ride for an average length of 30 minutes and more while subscribers ride for around 11 minutes.

This graph shows **Duration vs Age**



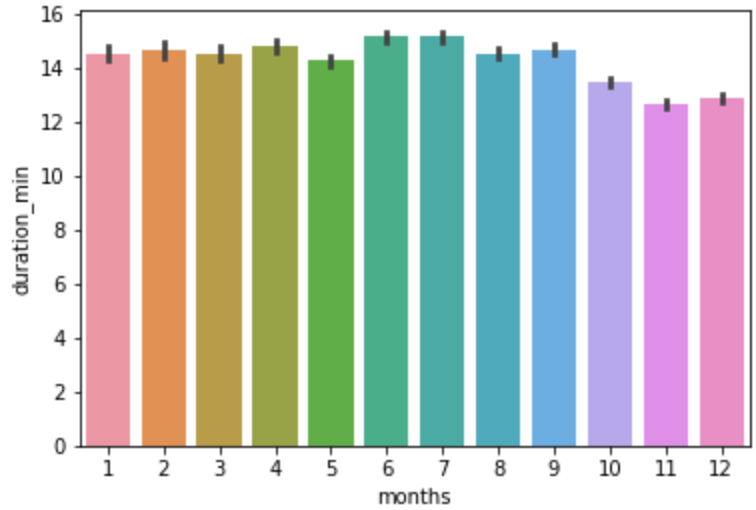
This graph shows that most of the riders rode under 200 minutes. Those that were 70 and above did not really ride for any longer than 200 minutes. There is a long dark line between 30 and 40 where these riders rode at similar distribution of minutes. This long line is probably the riders that were 33 years old.

This graph shows the **distribution of the age on the user-types**



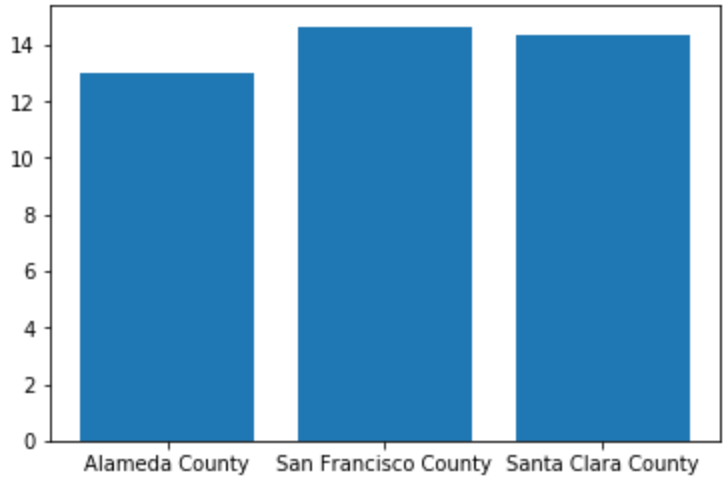
Most of those that were not subscribers were around 33 years old. Subscribers were more evenly distributed in terms of age.

This graph shows **average duration of rides and months**



For each month the average duration is similar. However, the duration of rides is much less in November and December.

This graph shows the **duration of ride in minutes and the three start county's**



We can see that throughout the three county's the duration of the rides is similar. With the averages being between 13.5 -14.5 minutes.

Summary

From the results we can see that there were many more subscribers than non subscribers. However, non subscribers seemed to ride for an average duration that was longer than that of subscribers. Many of the riders were from San Francisco County, however when finding the average minutes throughout the three counties, the average minutes were similar throughout the three. When looking at the distribution of gender in the riders, most of the users were males. When looking at the age, there were many more 33 year olds than any other ages.