Introduction

We are analyzing data related to the New York Citi Bike program, specifically in the month of December 2018.
More information about the program can be found on the following page:
https://en.wikipedia.org/wiki/Citi_Bike
The link to the dataset that we are analyzing is:
https://s3.amazonaws.com/tripdata/index.html
(The specific file used is the "201812-citibike-tripdata.csv.zip" file).
Finally, the link to the visualization produced is:
https://public.tableau.com/profile/rohit.rao#!/vizhome/CitiBikeAnalytics_9/GenderBreakdown

Questions

1. What is the age distribution of our riders?
2. Is there a correlation between the age of our riders and how long their trips are?
3. What is the gender breakdown of our riders?
4. What is the gender breakdown of our customers/subscribers?
5. Where are most of our riders coming from?
6. Where are most of our riders going?
7. Which day(s) saw the most activity? Which day(s) saw the least activity?

Analyses

1. (Reference "Age Distribution (By Gender)" sheet) The majority of our male bikers are 30 years of age, while the majority of our female bikers are 28 years of age. This makes sense, as people who are in their late twenties/early thirties have three things: Disposable income, time, and energy, all of which are crucial to success in our program.
2. (Reference "Age vs. Trip Duration" sheet) There is neither a clear positive nor negative correlation between age and trip duration in our data, which at this time refutes the claim that those who are younger choose to bike longer than those who are older. A potential reason for this could be that our service, particularly on weekdays, is mainly used as a means for our riders to commute to work and back, as opposed to one used for recreational services; as such, it makes sense for our riders to have a destination in mind when they start biking, as opposed to biking long distance for fun or for exercise.
3. (Reference "Gender Breakdown" sheet) Within our data, we have many more male bikers than female bikers. This could either be a sampling bias (we have collected more data on males than females), or it could have something to do with how we market this program to female riders.

4. (Reference "Gender vs. Subscribers" sheet) We have many more subscribers than we do customers within our dataset. A possible reason for this could be because it was easier for us to collect data from on our subscribers than our customers, resulting in a sampling bias. Another reason for this disparity could involve the change in weather; Since this past December was extremely cold, people are less likely to want to rent a bicycle, unless they are already paying for a subscription service and want to make the most of it.

5. (Reference "Most Popular Start Stations" dashboard) The most popular start station is "Pershing Square North," with "W 21 Street & 6 Ave" at a close second.

6. (Reference "Most Popular End Stations" dashboard) The most popular end station is also "Pershing Square North," with "Broadway & E 22 Street," coming in at second.

7. (Reference "Daily User Activity" sheet) The 3rd of December recorded the highest level of user activity, while the 25th recorded the lowest level of activity. One can see a general downward trend in user activity after Friday the 21st. These observations make sense, as many people took time off from work to travel and/or spend time with family during the final week of the year. Since the 25th was a national holiday, people would not be commuting to work, which also explains the lack of user activity on that particular day.

Conclusion

One of the key takeaways from our dataset (for December 2018) is that we have significantly more male riders than we do female riders. This requires us to ensure we are doing all we can to market our services to female riders as well. The majority of our riders are in their late twenties/early thirties, which makes that our primary demographic. The fact that there isn't a correlation between age and trip duration leads us to believe that our service is seen by riders as a commuting service, as opposed to one used for recreation/exercise; This explains the decline in activity after the 21st of December, both from one-off customers and subscribers, as many people choose to take off from work during the last week of the year in order to travel and/or spend time with their families. Lastly, "Pershing Square North" is our most popular station, both as a starting and an ending point. It would be beneficial for us to analyze what we are doing to draw in customers in that particular location and focus our efforts either on replicating our marketing process with our less popular stations or on growing our user-base in that particular location and in neighboring areas.