**Report on Citi Bike Trip:**

My time window for analyzing the bike riding share data was from January 2018-January 2019. The total number of records recorded during this time window is a total of 373, 568. When we plot this data over months we see that the usage of bikes peaks during the summer months, starting around May and dropping right after October. The highest used months is in August with a total of 44, 432 rides taken, compared to January, with only 12, 667. The most used days are during the weekdays from Monday through Friday. The peak hours of usage are not surprisingly at 8am and then again at 6pm both during summer and winter months. The main difference between summer and winter months being that overall ridership is significantly reduced. When we look at the percent of difference in number of records throughout the months of the year we see that it peaks during August, with a 250% difference from January. In addition, there is a 55% increase between January 2018 and January 2019, showing that ridership increase from 2018 to 2019.

The percent of annual riders and short term customers is quite different between time. You have many short term customers specifically during the summer months. Yearly subscribers increase to about 200% over the summer months while short-term customers increases by 2000%. Today, the top 10 stations to start a Journey are: 1) Grove Street Path, 2) Sip Ave, 3) Exchange place, 4) Hamilton Park, 5) Newport PATH, 6) Dixon Mills, 7) Brunswick and 6th, 8) Newark Ave, 9) Monmouth and 6th, and 10) Harbor Side. Also, the top 10 stations to end a ride are: 1) 1) Grove Street Path, 2) Sip Ave, 3) Hamilton Park, 3) Exchange place, 5) Brunswick and 6th, 6) Harbor Side, 7) Newport PATH, 8) City Hall, 9) Jersey and 6th and 10) Newark Ave. By contrast the bottom 5 stations to end a Journey are: 5 Corner Library, Communipaw and BerryLane, Leonard Gordon Park, and Washington St. and the bottom 5 stations to start a Journey are: 5 Corner Library, Christ Hospital, Essex Light Rail, Fairmount Ave, and Lincoln Park. Today the gender breakdown is 299 bike records for male vs 66 records for females, showing that there is a much bigger usage of ridership for male drivers than female drivers. When we compare of usage overtime based on gender, we see that there is a 70% increase of usage among females from January 2018 to January 2019, compared to 34.6% among male riders, suggesting that female riders is increasing at a higher percentage that of males, suggesting that outreach for female riders might be making a difference.

Interestingly, when we look at average trip duration by age, we see that the biggest riders are born between the 1998-2002, showing that the average trip duration, and most likely correlating with distance, is among young people. When we make a bubble plot for bikes by id vs the sum of total trip duration, we can clearly visualize the bikes that are most used. Using this data, we should be able to track them and provide them with preventive maintenance. In addition, we can see that the utilization of bike IDs is quite variable, most bike IDs clustering between the digits of 26000-26500, 29000-30000, and 33500-34000. Finally, we are able to take advantage of mapping bikestations by lat and long in NYC maps and overlay it with zipcode. This way we can visualize the most used bikes to start a trip and end a trip in NYC.