**Citi Bike Program Analysis**

* Data Range January to March 2018 (Link source <https://s3.amazonaws.com/tripdata/index.html>)

These zipped files were used to download the CSV files

* <https://s3.amazonaws.com/tripdata/201701-citibike-tripdata.csv.zip>, <https://s3.amazonaws.com/tripdata/201702-citibike-tripdata.csv.zip>

<https://s3.amazonaws.com/tripdata/201703-citibike-tripdata.csv.zip>

## Goal

Analyze CitiBike data between January and March 2017 for trends and visualize them using Tableau Public. View the Tableau Public workbook

* After creating and analyzing our visualizations, there were many phenomena uncovered.
* The first being that the trips for the most used bikes are often started on or around the same stations, this may indicate that we need to increase our supply of bicycles in these areas. Adding more bikes to these stations will allow more users to have a bike available as well as preventing the same bikes from enduring heavy usage which will increase our repair fees, or worse, the risk of user injury. Pershing Square North is the Station with highest number of records for starting a journey (26,810) as well as ending a journey (26,599).
* The Static Map for the Start Station and Locations to end a journey have records ranging from 10,000 to 26,810 and 10000 to 26,599 respectively. The maps give idea about the many stations with higher number of records. The size of the circles for a particular station increases proportionately to the increasing number of records. We can get a very clear idea of the busier stations for the start and end journeys.
* Another phenomenon that was uncovered from our visualizations, is that most of our riders are between the ages of 20 and 40 However, there are some red flags raised as it appears that we do not confirm age in any way and that the user can say they were born in any year. This was noticed when we had multiple users over 100 years old riding bikes. This is extremely unlikely It may be wise to start having users validate their age, so we do not have invalid data or even have underage children lying about their age to ride our bikes. Also, to emphasize this phenomenon we saw that that the Birth year of 1915 shows an Average trip duration of 2678. This shows that the riders age is about 102 years. The data seems to be highly erroneous.
* The final phenomenon that I want to recognize is that the most popular start times are at 8 AM and 6 PM. This makes sense as it is similar to when rush hour occurs. People take a bike as they are going to work in the morning and then again when they leave work in the evening. The bump around 6 PM could also be from students leaving class for the day. Also, subscribers(male) contribute to very high total trip numbers.