**TABLEAU CHALLENGE DESCRIPTION & ANALYSIS**

The data I used from the NYC Citibike website [Citi Bike System Data | Citi Bike NYC | Citi Bike NYC](https://citibikenyc.com/system-data) was the monthly csv data from December in the years 2016, 2017, 2018. I combined these 3 CSV files using a union in Tableau.

I created 2 versions of the same map showing the location of all bike stations from 2016-18. The definition of most popular stations I chose to use was top 30. I was unable to find a way to produce a street view map with zip codes so I chose to produce 2 maps, note one shows most popular departures and one shows destinations. Interestingly they were mostly the same stations, all located on Manhattan inner city, fairly close together.

The first “phenomena” I chose to look at was trip duration in particular was there significant variance in trip times for different destination stations. Initially I was looking to show average(mean) trip time but there a few major outliers skewing the data as shown in the bubble chart. One outlier was a trip of around 40 days apparently. I choose to display median trip times for a minimum of 10 trips and show both the greatest median trip and lowest median trip. Data was originally in seconds but I decided to create a calculate field so I could display in in minutes. Lowest median trips were between 2.15 and 6 minutes , while longer median trips were between 13 and 21 minutes. So 1 station has a median trip 10 times that of thee station with the shortest trip. It does seem surprising that trips and generally quite short in duration but I guess December being winter in NYC its probably not that surprising. Maybe it would have been good to compare winter and summer if I had more time.

The second “phenomena” I choose to investigate was looking at time of the day variance. This dashboard shows the variance in trips for time of the day for both 2017 and 2018. The seem to show a similar pattern and there doesn’t appear to be any major change in the way people used the bikes in December of these 2 years. Due to the difference of some field names from 2016 vs the other 2 years and time constraints I was unable to find a way to include 2016 as third option. Another interesting observation was that use seems to peak around start and end of the workday which seems to imply that local residents are the predominant users of the bikes as opposed to tourists and visitors to NYC. I then produced 2 more visualisations then utilized “time bins” calculate value I created. I had morning midnight to 11:00, lunch 11-2, afternoon 2-5 and then evening 5-midnight. There wasn’t much difference between the 2 years but lunch 11-2 was the time of day the bikes were used the least. I also looked at the impact of gender (1= male, 2= female). There was no obvious difference in the way the 2 genders used the bikes by time of day. Interestingly though about 75% of trips were made by males.