Citi Bike 2018 Data

I utilized 12 CSVs for each month of 2018, I then put them in pandas to concatenate all the files after verifying the same column headings. By having one meta csv file it was easier to load into Tableau.

Citi Bike Analysis:

1. Total Trips – There were over 350K total trips over the span of last year
2. Percent of Total Ridership – Customers and Subscribers have a positive correlation but there are significantly more subscribers than customers. Subscribers shoot up towards summer months
3. Average Trip Duration- Trip durations peak at May and hits its low in February. Warmer months do better than colder ones.
4. Peak Hours – This visualization tells us a lot of info on the daily trip numbers. Peaks are around start and end times of the work day 8am and 6pm. 3am is the lowest hour of the day.
5. Age vs Avg. Trip Duration – Top 5 birth years (largest to smallest) that had the longest average trip durations were 2001, 1939, 2000, 1999, and 2002. This brings an interesting analytical point that 4 of the top 5 were people born at the turn of the new century but one of the birth years is 1939 which is very abnormal.
6. Total Trips by Age – Top 5 birth years (largest to smallest) that had the highest amounts of trips were 1969, 1988, 1986, 1989, and 1987. This shows an interesting observation about 4 of the top 5 as they are from the mid to late 1980s. The abnormal result is the number one birth year of 1969.