**Tableau Analysis**

Dashboards are created to provide insights into "Start Station" data, including location, the Top 10 station names, ridership type, and membership breakdown on an annual and quarterly basis.

**Change – Dashboard:**

As part of data preparation, a "Ride\_Month" column is added based on each file's title to identify the corresponding month. Tittle of the files are manipulated to show the right formatting to add as column. All files are then combined to create quarterly and annual datasets for this analysis. VBA scripting is used to complete the source cleanup.

Percentage Change Dashboard is created to indicate the two breakdowns of Membership and Rideable Type percentage change over months until September. As it indicates on both bar charts, the total bike rides are increasing from Jan to September, except a slight decline in August. Majority of bikers are members. Casual bikers are increasing after month of February. In march we see 52% increase, while member bikers are increasing at 11% in March. Overall, summer season and September, the biking is at its highest compare to winter season.

Similar analysis is also done in the Ridable Type bar chart. Overall, the electric bikes are utilized more than the classic bikes. However, there is 33% increase in classic bike ride in March while electric bikes ride only increased 7%. Biking trend continues to increase until June. As indicates in the dashboard, the biking is increasing during spring compare to winter season, and stays steady during summer and beginning of fall.

**Q1, Q2, Q3 Map and Start Station Analysis:**

For this analysis, “Started At Time” column separated into two columns as “Start\_Date” and “Start\_Time”. Start\_Time column is aggregated into below groups:

* Early Morning: 00:00 - 07:00
* Morning: 07:00 - 12:00
* Afternoon: 12:00 - 16:00
* Evening: 16:00 - 18:00
* Prime: 18:00 - 21:00
* Late Night: 21:00 – 24:00

Two interactive charts are created to indicate the location of the Top 10 Start Station and the relationship between the other variables (Start ride time grouping, Member and Ridable type) against the monthly data for each quarter. **The map** indicates the total count of the rides based on monthly breakdown for each start station.

**Box plot** is created to show how the count of the start ride time grouping is distributed amongst the Top 10 Start Stations. This allows to identify any outliers and the most used time (peak) for given station name.

**Observations:**

* Based on the observation from Map and Box plots, Grove St. PATH and Hoboken Terminal – Hudson st, are the two top stations where they can be considered as outliers compare the others start stations.
  + Prime, Evening and Afternoon are the peak times when these two stations are used the most as start station.
* Hamilton Park, Hoboken Terminal – River St. & Hudson PI stations are the most used start stations for the Morning time riders.
* Hoboken Terminal – River St. & Hudson PI and Grove St PATH are the highest two start stations for late night rides.
* Similar analysis can be expanded to include Membership and Ridable Types across Q1, Q2, Q3.