Bike sharing assignment

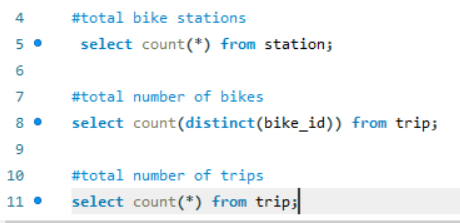
Task 1:

1. What are the total numbers of:
   * Bike stations?
   * Bikes?
   * Trips?

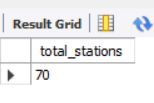
Answer:

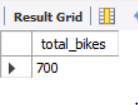
Total number of bike stations are 70, total number of bikes are 700 and the total number of trips are 669959.

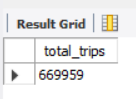
Input:



Output:

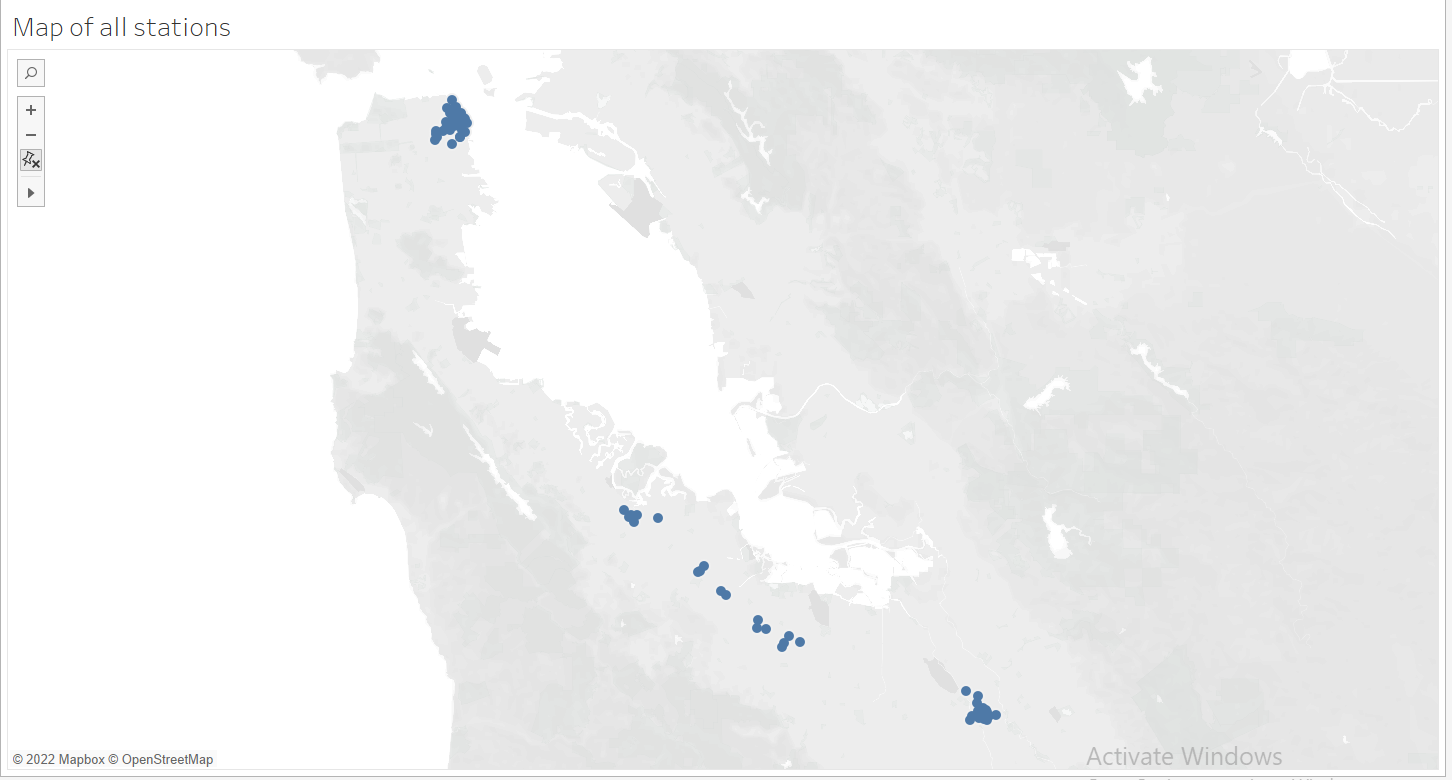






1. Construct a geographical plot to show the location of each bike station using the latitude and longitude provided under the Station table.

Answer:



Here is the map of all the stations

1. What is the relationship between the following columns (one to one, many to one, many to many)?
   1. bike\_id (Trip table) and start\_station\_id (Trip table)
   2. pincode (Weather table) and station location (latitude and longitude in Station table)
   3. 8/29/2013 (date column in Weather table) and mean wind speed (Weather table)

Answer:

many to one, one to one and one to one respectively.

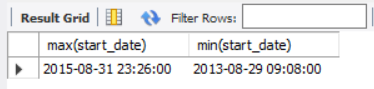
1. Find the first and the last trip in the data.

Answer: the first trip was made at 29th August 2013 at 09:08 a.m. and the last trip was made at 31st August 2015 at 11:26 p.m.

Input:



Output:

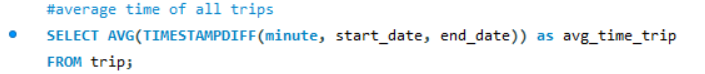


1. What is the average duration:
   1. Of all the trips?
   2. Of trips on which customers are ending their rides at the same station from where they started?

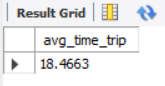
Answer:

1. The average time for all the trips is 18.4663 minutes
2. The time for the trips where the start stations and end stations were the same is 105.9443 minutes.

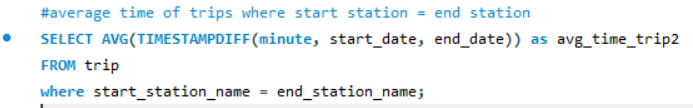
Input:

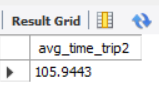


Output:



Input:





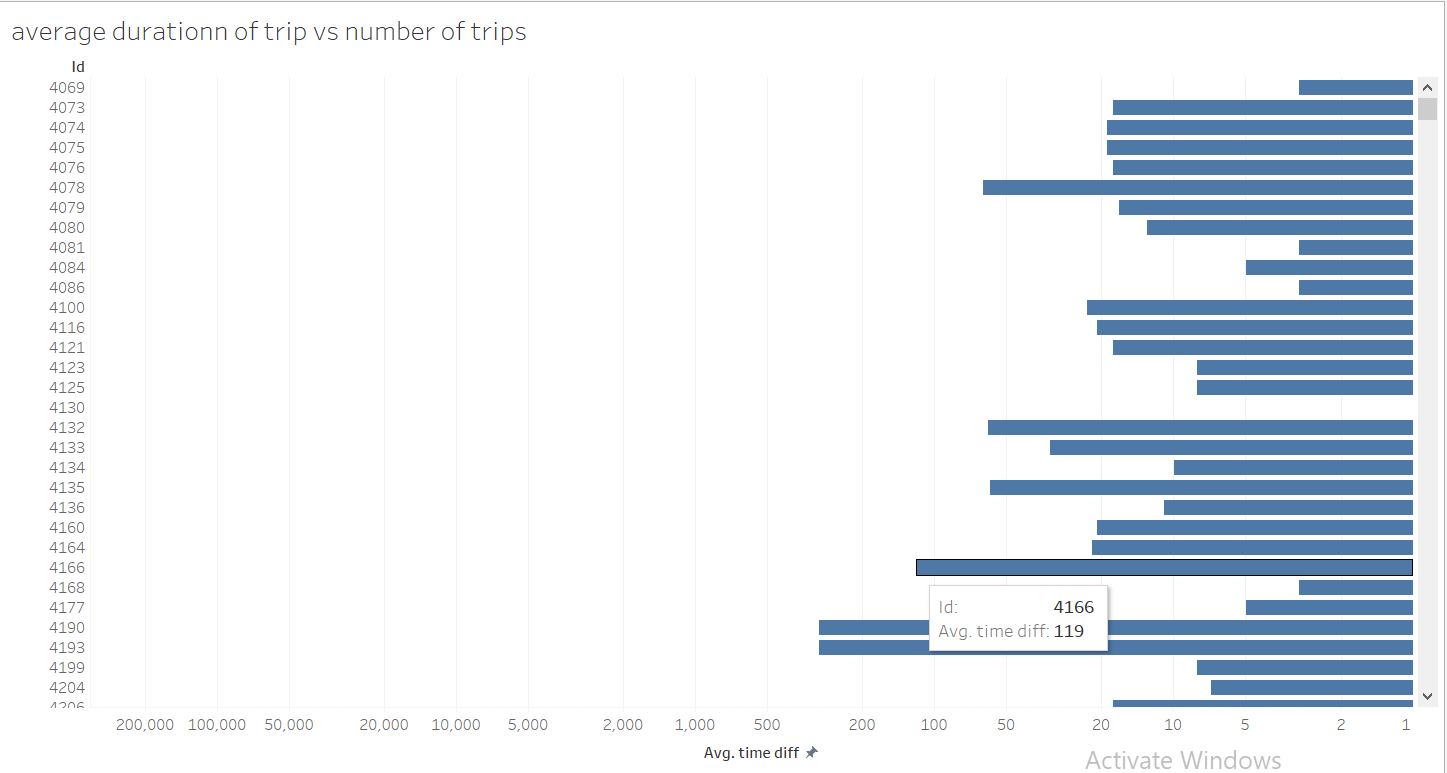
1. Which bike has been used the most in terms of duration?

Answer: Bike with bike\_id 535 is used the most which ended up to be 287899 minutes.

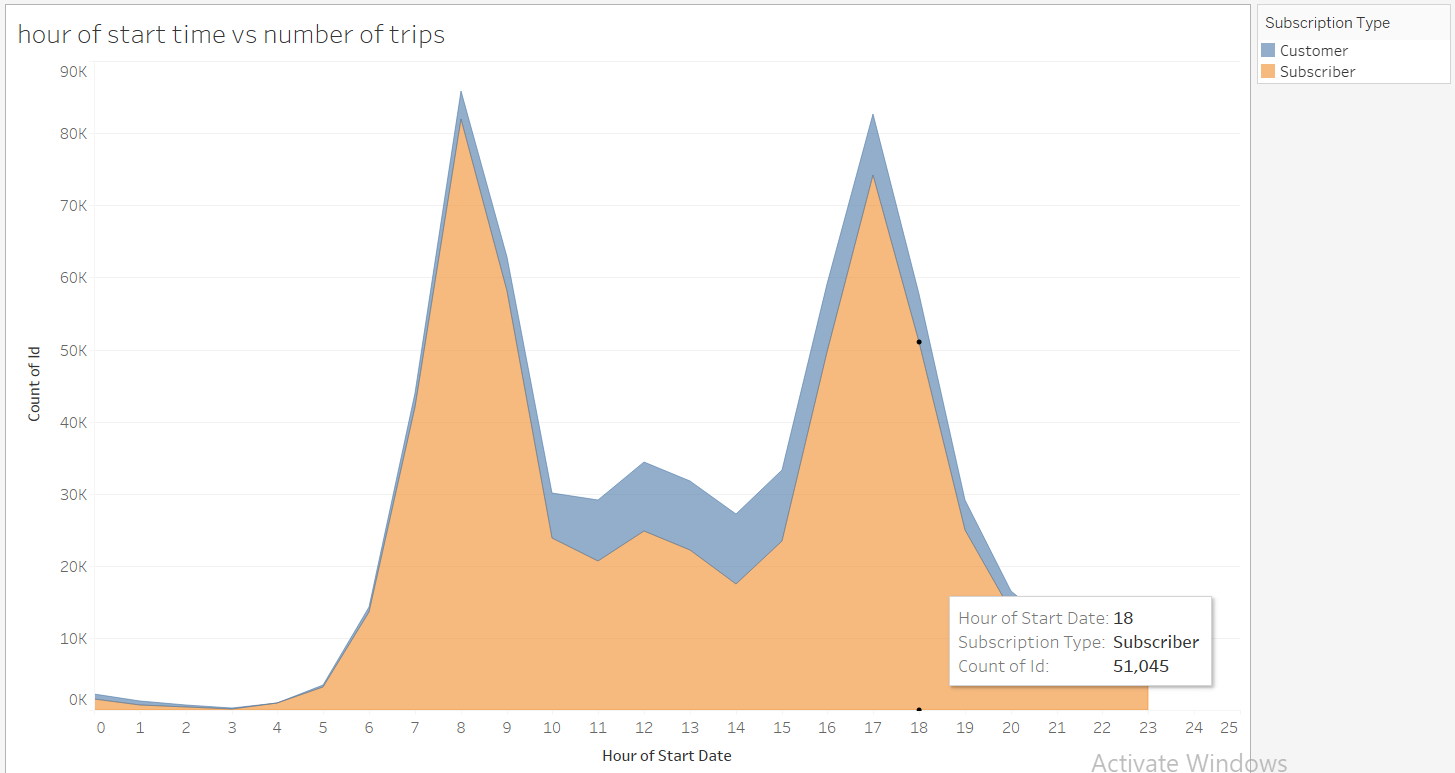
1. Plot the most suitable graph for the following:
   1. The average duration of a trip versus Number of trips
   2. Hour of start time versus No. of trips
   3. Day of the week versus No. of trips also denote subscribers and customers with different colors.

Answer:

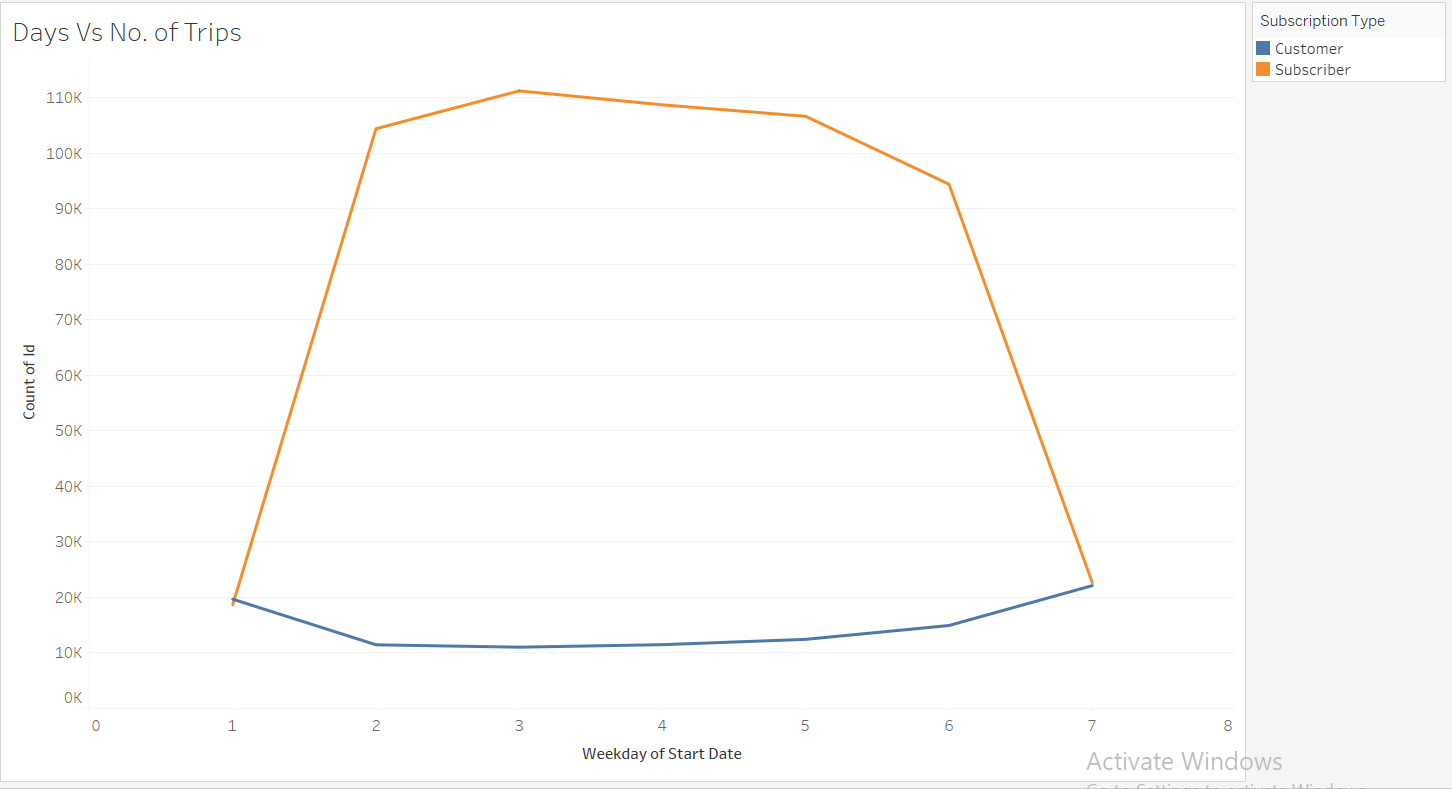
Here is the plot for the average duration of a trip versus Number of trips



Followed by the second plot which is the hour of start time versus No. of trips

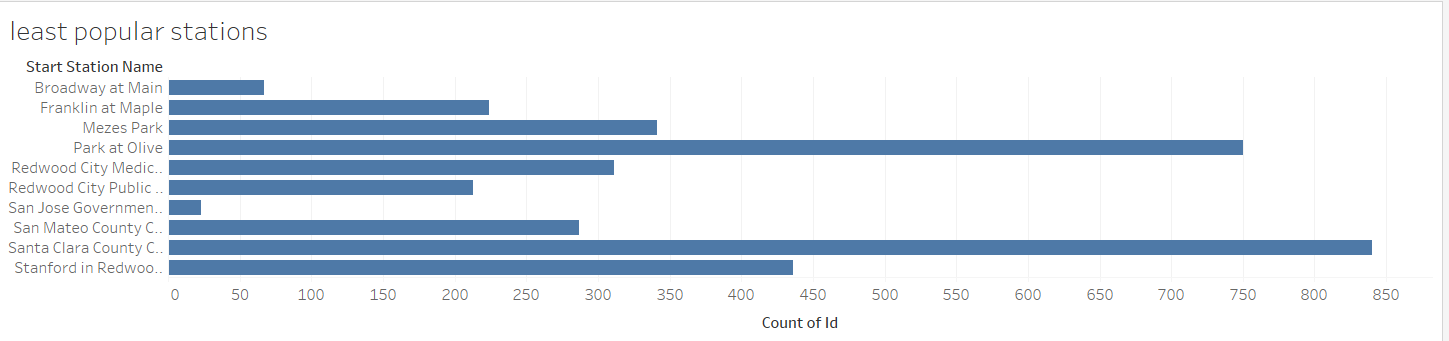


And finally, the last plot is day of the week versus No. of trips also denote subscribers and customers with different colors.



Task 2

1. What are the top 10 least popular stations? Hint: Find the least frequently appearing start stations from the Trip table.

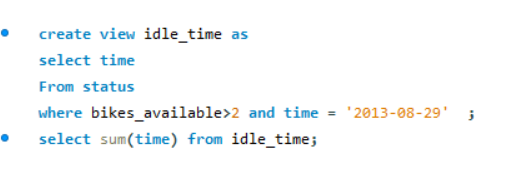


Here are the top 10 least popular stations.

1. Find the idle time for Station 2 on the date '2013/08/29'

Answer: The ideal time for station 2 on 29th august 2013 is 1 hour and 59 minutes

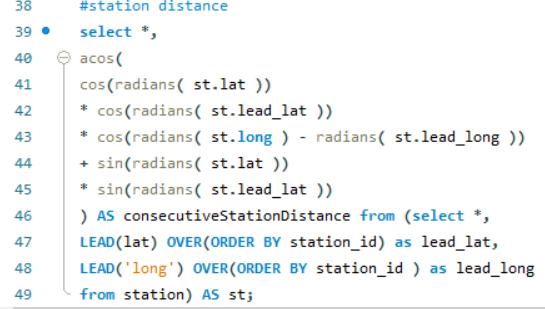
Input:



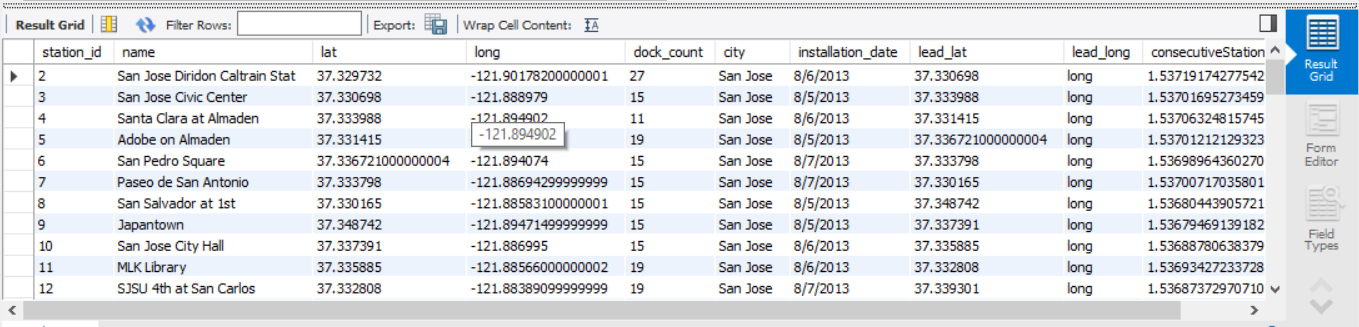
1. Find the distance between consecutive stations

Answer

Input:



Output:



1. Use the findings above to recommend three stations that can be shut.

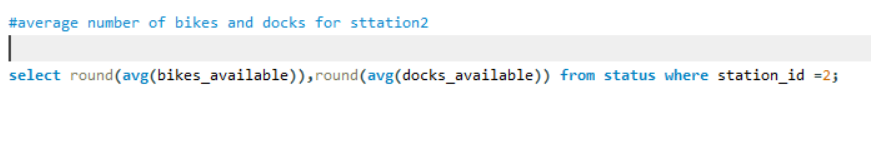
Answer: The three stations that we can shut down are the Broadway at main, San Jose government and Franklin at Maple as these are the least popular stations. Another reason to shut down these stations is because there are other stations which are close to these three stations and are preferred more by both consumers and subscribers.

Task 3

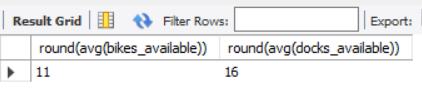
1. Calculate the average number of bikes and docks available for Station 2.

Answer: The average number of bikes and docks for station with id 2 are 11 and 16 respectively.

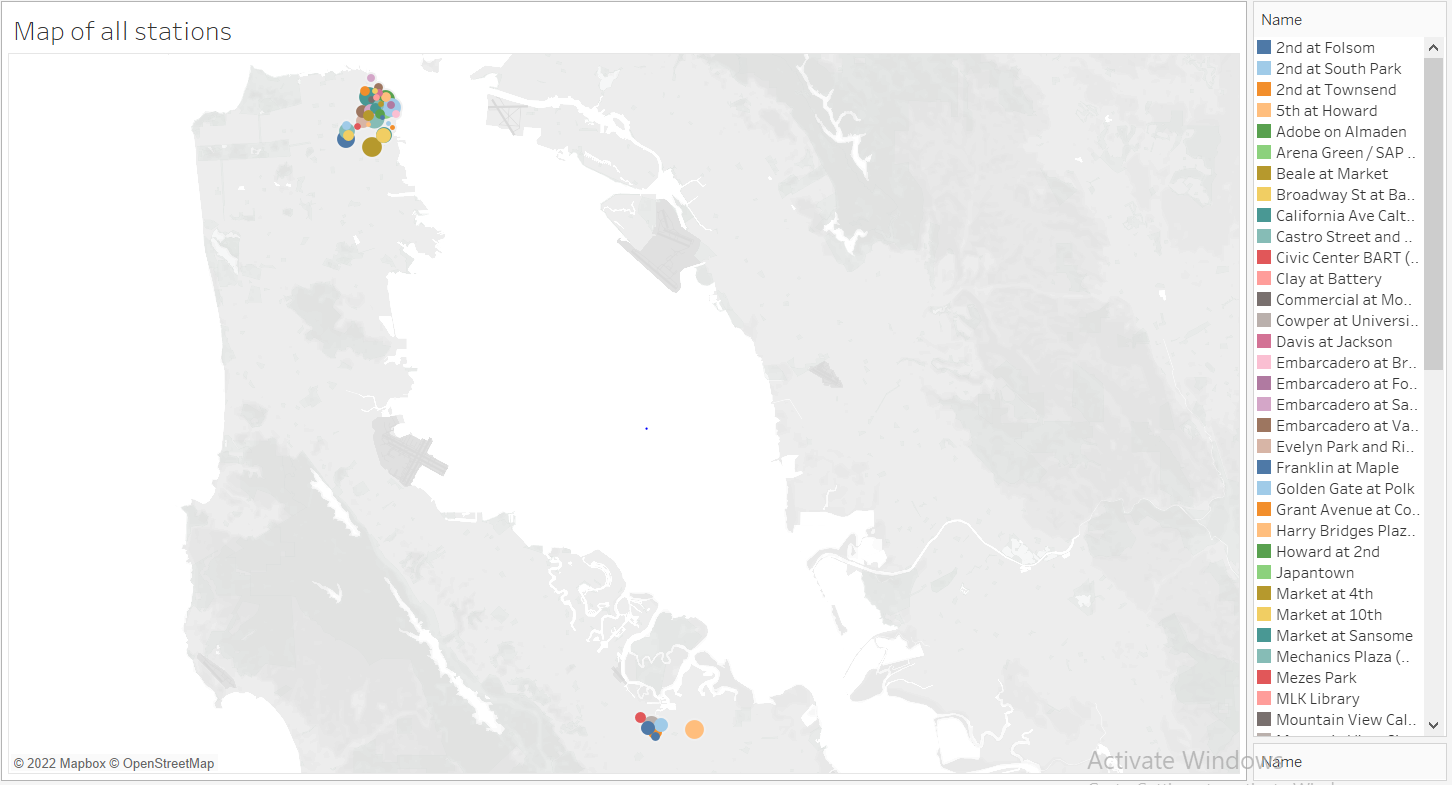
Input:



Output:



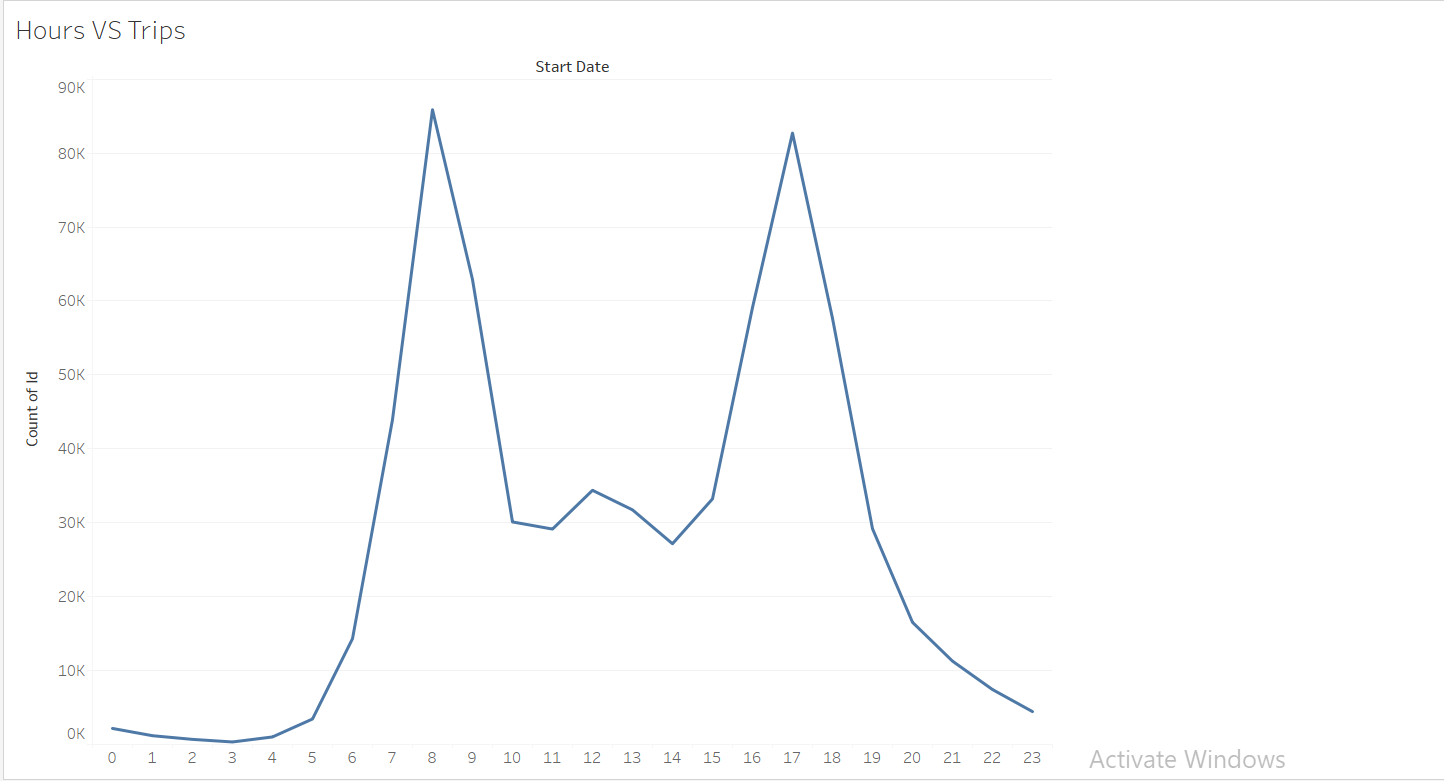
1. Plot the popularity of each station on a map for subscribers and customers.

Answer: 

Here is the colour coordinated map of all stations

1. Plot the number of trips per hour for all the data provided in the Trip table.

Answer:

3236 

Here is the plot for hours vs no of trips

1. Use the findings above to provide insights on how to optimize operations

Answer: To optimize operations, we can conclude from these three plots and data that the average number of docks and bikes available are 16 and 11 so we can easily shut down the stations which are not popular or are near more popular stations. We can also conclude that there are two clusters of stations in our map which means either the distance between two stations is a lot or is very less. We can also conclude that the mostly our bikes are not used in the time of night. Rather they are used the morning hours i.e., 5 to 10 AM and evening hours i.e.,3 to 11 PM which is the office commute time. We can also see that our subscribers prefer more to use our service on weekdays while consumers prefer on weekends.

Task 4

1. Zulip has decided to start a new product line called Couple Bikes. This will enable two persons to travel from one station to another at the same time. What are some of the factors that you will have to consider while validating the idea of couple bikes?

* Answer: The idea of couple bikes would both help the company and at the same time could bring a lot of challenges. The main factor to consider is that our customers are mainly travelling on weekdays and on office times that means our customers are mainly office workers and generally office workers don’t prefer taking couple bikes. Rather they prefer their commute alone. If the company wants to introduce couple bikes, then it is highly possible that the existing customers would be divided. We can also not keep the charges more than double the price of individual bike. That would totally fail the motive of introducing a couple bike. These are the factors that we have to consider while introducing couple bikes in our company.