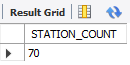
**Task 1: Get to Know Your Company**

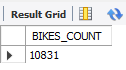
1. What are the total numbers of:
2. Bike Stations ?





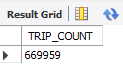
1. Bikes ?





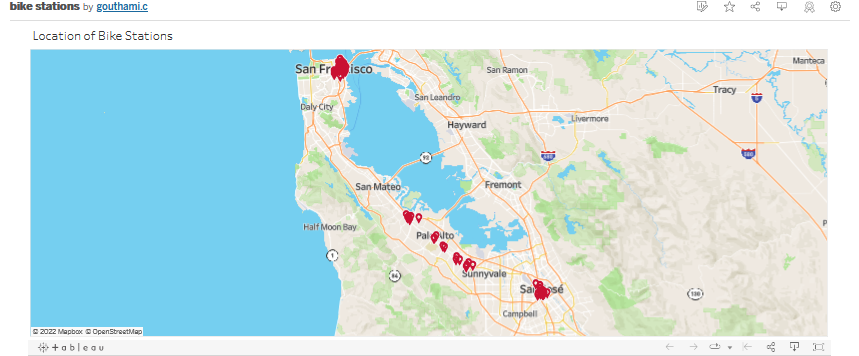
1. Trips ?





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

<https://public.tableau.com/app/profile/gouthami.c/viz/bikestations_16472381933850/bikestations>



1. What is the relationship between the following columns (one to one, many to one, many to many)?
2. Bike\_id (Trip table) and start\_station\_id (Trip table)

Each bike can have one start station at a time 🡪 one to one relationship

Each bike station can have multiple bikes at a time -> one to many relationships

1. Pincode (weather table) and station location (latitude and longitude in station table)

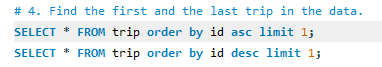
Each pincode can have multiple station location -> one to many relationships

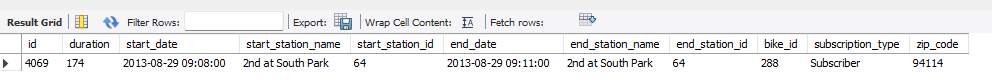
Each station can have one pincode -> one to one relationship

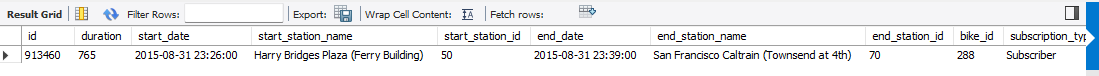
1. 8/29/2013 (date column in weather table) and mean wind speed (Weather table)

Date column in weather table and mean wind speed -> one to one relationship

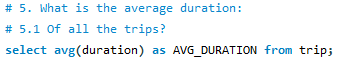
1. Find first and last trip from the trip table

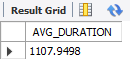




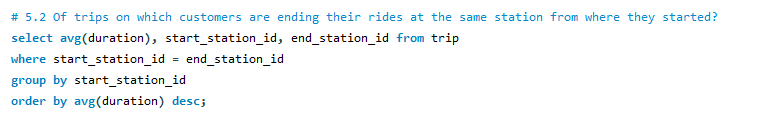


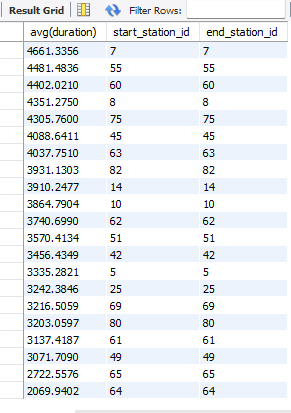
1. What is the average duration:
2. Of all the trips?





1. Of trips on which customers are ending their rides at the same station from where they started?

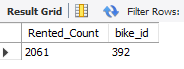




1. Which bike has been used the most in terms of duration? (Answer with the Bike ID)

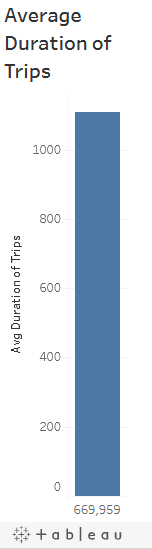
BIKE ID : 392





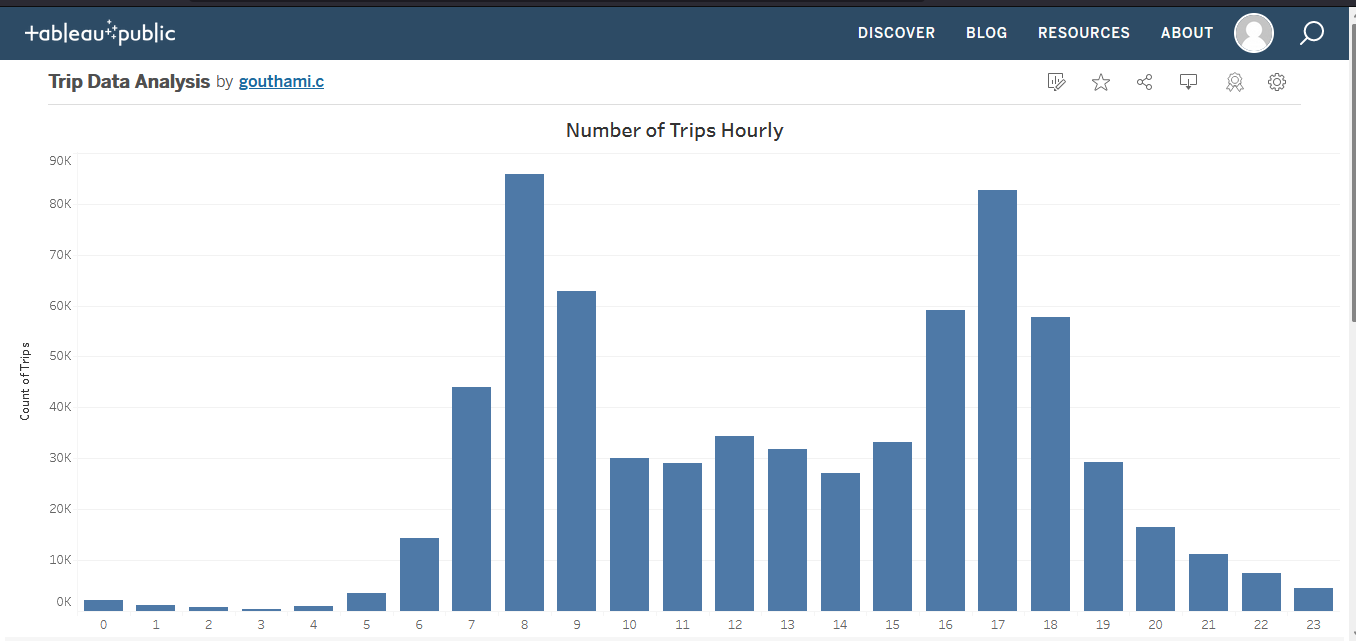
1. Plot the suitable graph for the following:
2. The average duration of trips vs Number of trips

<https://public.tableau.com/app/profile/gouthami.c/viz/TripDataAnalysis_16472514158640/AvgDurationofTrips>



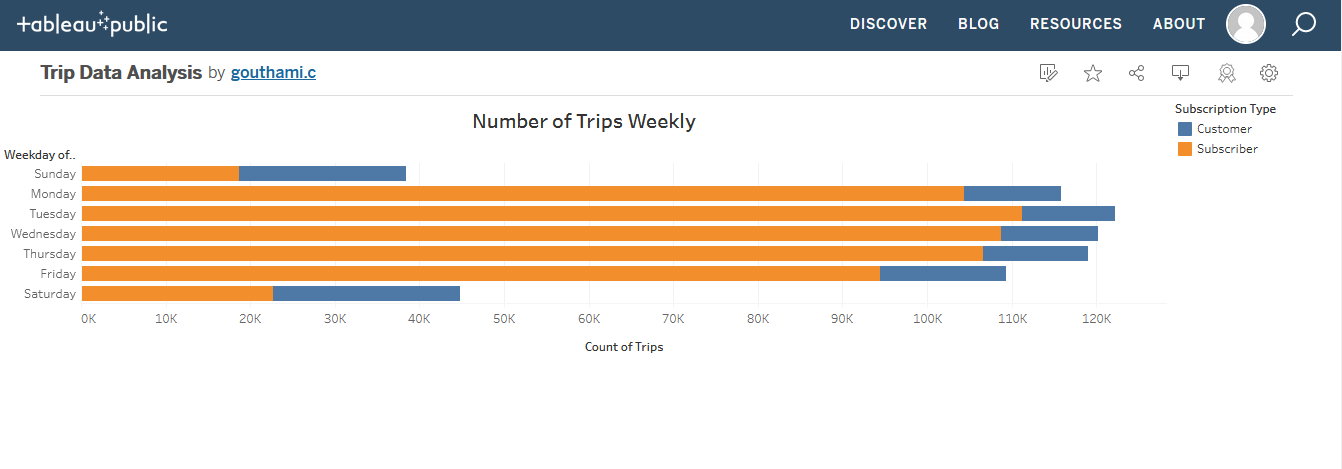
1. Hour of start time versus number of trips

<https://public.tableau.com/app/profile/gouthami.c/viz/TripDataAnalysis_16472514158640/NumberofTripsHourly>



1. Days of the week versus No of trips also denote subscribers and customers with different colors

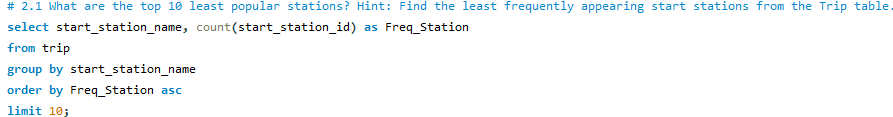
<https://public.tableau.com/app/profile/gouthami.c/viz/TripDataAnalysis_16472514158640/NumberofTripsWeekly>

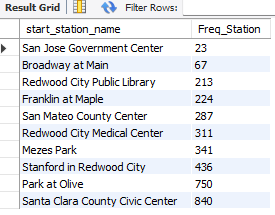


**Task 2: Demand Prediction**

Zulip is running under a loss and has decided to shut operations for three of its stations. You have to use the data provided to help Zulip decide which three stations should be shut.

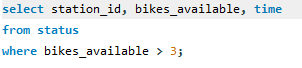
1. What are the top 10 least popular stations? Hint: Find the least frequently appearing start stations from the Trip table

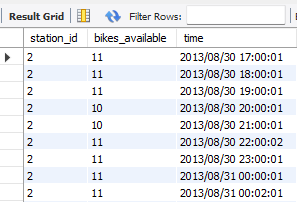




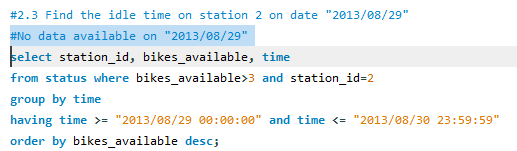
1. Idle time is the duration for which a station remains inactive.

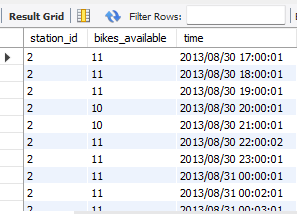
You can consider this as the time for which a station has more than 3 bikes available





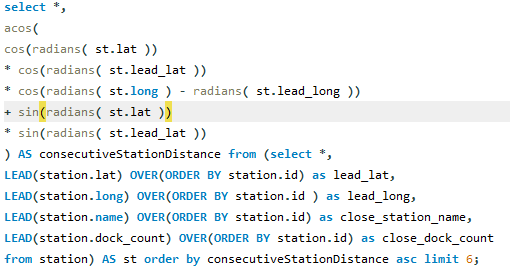
* 1. Find the idle time on station 2 on the date “2013/08/29”

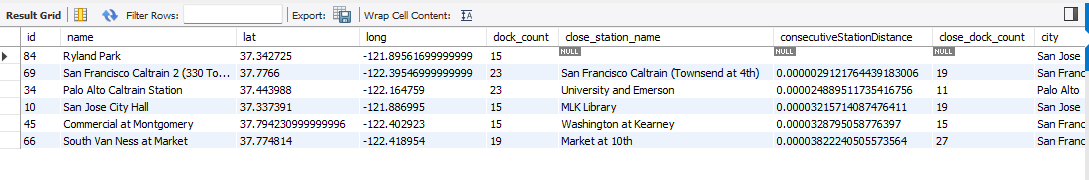




1. In case two stations are nearby, it might be possible to shut one down. Find the distance between consecutive stations (between Stations 1 and 2, Stations 2 and 3, and so on). The Haversine formula 2arcsin(√sin2(ϕ2−ϕ12)+(1−sin2(ϕ2−ϕ12)−sin2(ϕ2+ϕ12).sin2(λ2−λ12)) is used to find the distance between two points on a sphere given their longitude and latitude. (ϕ1,λ1) is the latitude–longitude pair for the first station, and (ϕ2,λ2)is the latitude–longitude pair for the second station. You can find the SQL code for this formula given below

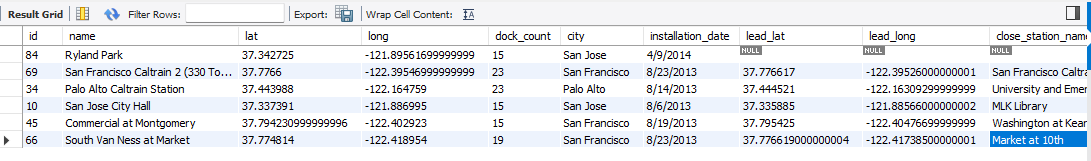
Added the station name to which distance is calculated and it’s respective dock count for comparison.





1. Use the findings above to recommend three stations that can be shut. (open ended) For example, if the Japantown and Ryland stations are nearby, and the Japantown is not as popular as the Ryland station, then it can be recommended to shut.

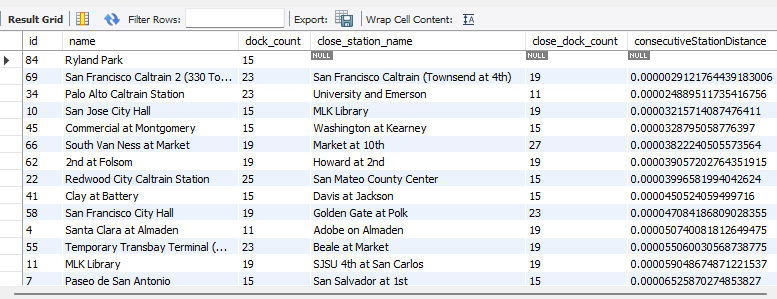
From the above query we can find out the stations which are closer by



From the trip table find out top stations by finding number of trips starting from start\_station\_name.



Compare the results to find out which stations can be shut down based on dock count, frequency of station used, popularity of closer station;



From the above table:

1. **MLK Library** with dock count 19 and frequency of trips 2034 closer to

**SJSU 4th at San Carlos** with dock count 19 and frequency of trips 1170

Based on above statistics **SJSU 4th at San Carlos** can be shut down.

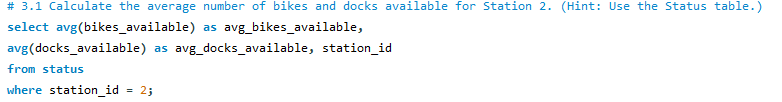
1. **Redwood City Caltrain Station** with dock count25and frequency of trips 1554 closer to **San Mateo County Center** with dock count 15 and frequency of trips 287 can be shut down **.**
2. **Commercial at Montgomery** with dock count 15 and frequency of trips 11888 is closer to **Washington at Kearney** with dock count 15 and frequency of trips 1472 can be shut down.

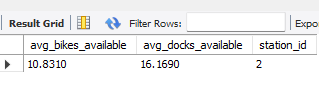
From above details **SJSU 4th at San Carlos, San Mateo County Center and Washington at Kearney** can be shut down**.**

**Task 3: Optimizing Operations**

Throughout the day, bikes keep moving around the city due to the trips. Zulip has to find out how to effectively move bikes around to ensure the demand is met with adequate supply. This is to ensure that at any time, there are sufficient bikes available at a given station. Here are some points that you will have to consider while deciding on the transportation of bikes from one place to another:

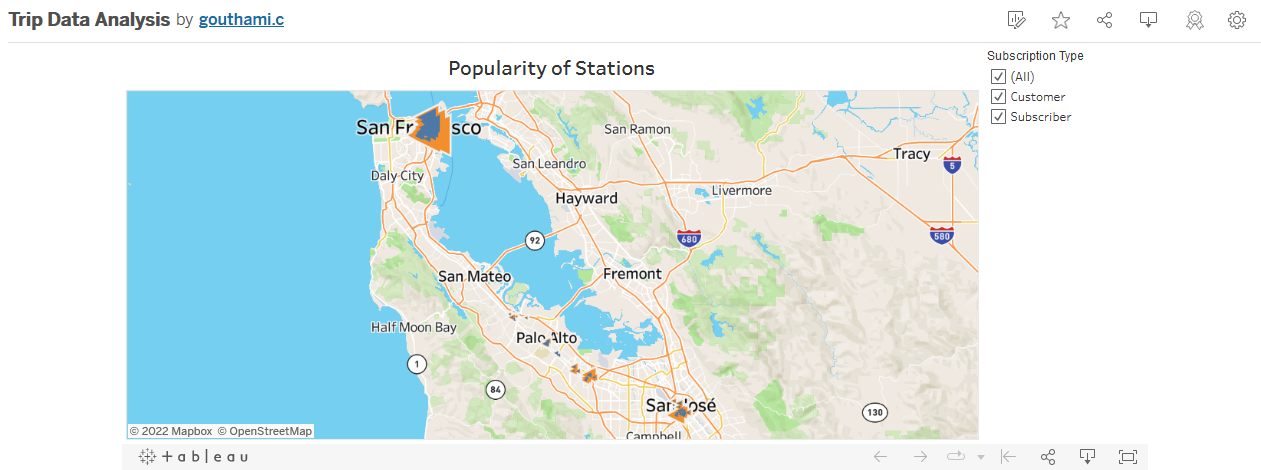
1. Calculate the average number of bikes and docks available for Station 2. (Hint: Use the Status table.)





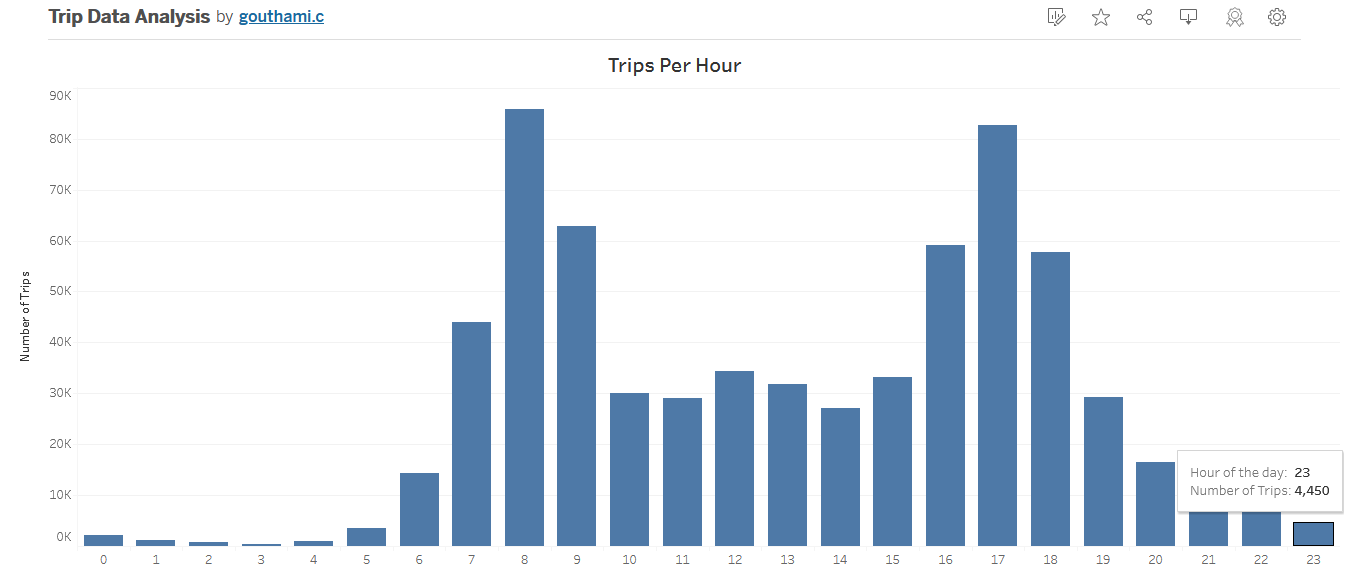
1. Plot the popularity of each station on a map for subscribers and customers. (Hint: Popular stations appear most frequently under the column start\_station\_name in the Trip table)

<https://public.tableau.com/app/profile/gouthami.c/viz/TripDataAnalysis_16472514158640/PopularityofStations>



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

<https://public.tableau.com/app/profile/gouthami.c/viz/TripDataAnalysis_16472514158640/TripsPerHour>



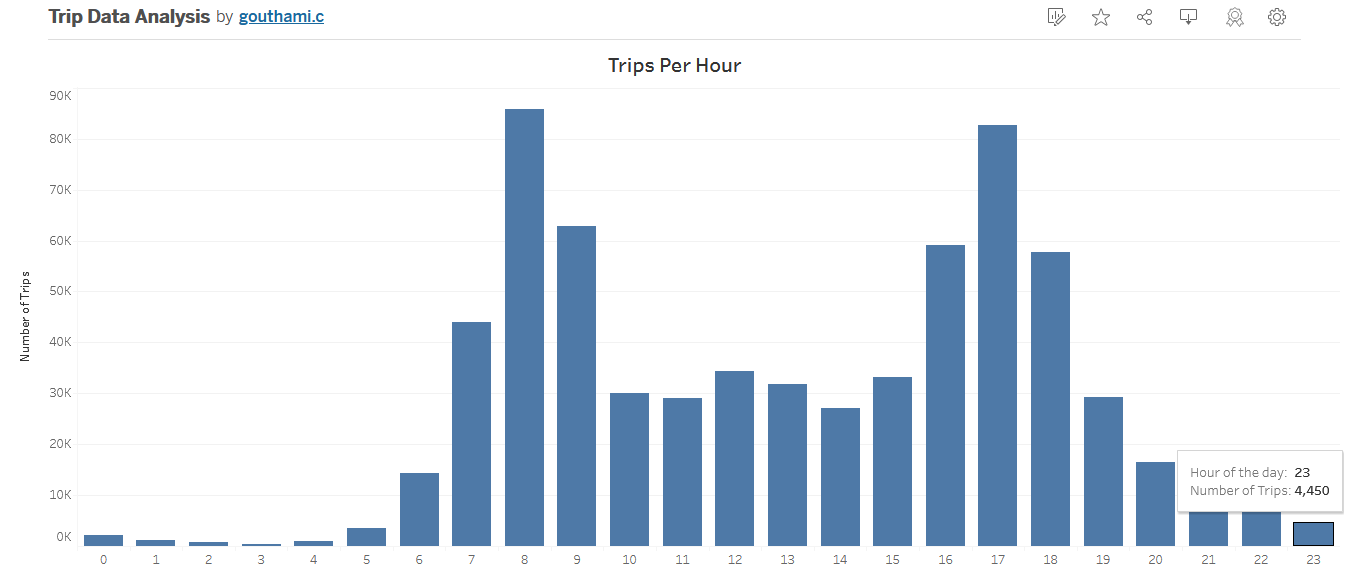
1. Use the findings above to provide insights on how to optimize operations. (open ended)

Based on the above data we can see first five hours of the day has minimum number of trips

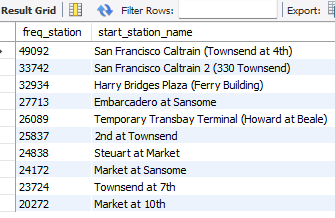
So we need to make use of this 5 hours to move the bikes to required stations. Required stations can be found checking the bikes\_available from status table and Popular stations can be found using trips table with statrt\_station\_name.

**Task 4: Couple Bikes? (Bonus)**

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?







From the above details:

Peak hour of bike rentals is from morning 8 AM to evening 6 PM

So we can start the couple bikes in the above 10 popular stations as a trial during peak hour and collect the data and analyze available data to make further decisions.