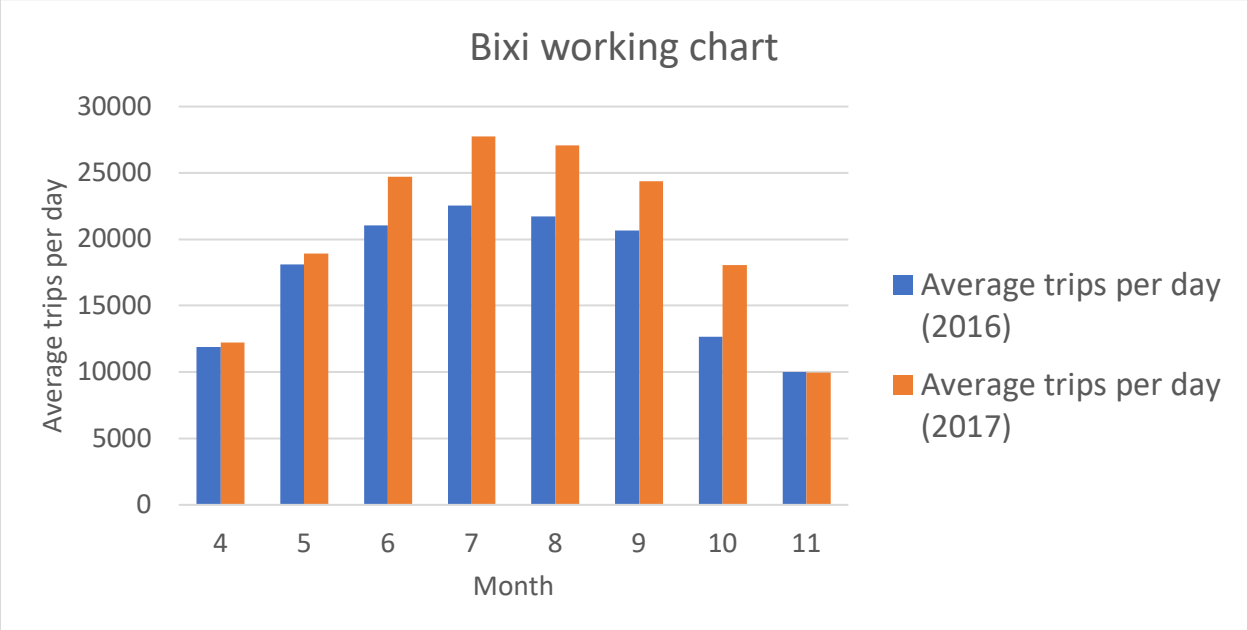
Bixi Project Part 1
Data Analysis in SQL

This report explores the data that was made public by Bixi Montreal, Quebec. The goal of the report is to find answers of crucial questions, find insights and general analysis of the operations for the years of 2016-2017 as part of an SQL assignment for the Data Science program at Brainstation.

Question:3.1 Which time of the year the demand for Bixi bikes is at its peak?

To answer the question, we will take a look the below graph.



We can establish that demand for the Bixi bikes follows a similar pattern for both years throughout the season, with the year 2017 peaking higher than 2016 clearly stating a noticeable growth in demand overall.

We can also notice that the months of July and August has the highest number of average trips per day in the whole season of operation.

Question:3.2 If you were to offer non-members a special promotion in an attempt to convert them to members, when would you do it?

As we can see in the graph the overall trips done by Non-members are significantly low, with slightly growing in the high demand months of July and August.
If the company wants to attract more members since it is their main source of revenue, they need to have their special offer's campaigns running during the months of April and May for the following reasons.

(a) These months are less popular with non-members having less 20% of share of the total trips.
(b) It is better to run promotions in the beginning of the season so that the new members can have enough time to use the service. And make the best out of their membership plans.

Question:4.1 What are the names of 5 most popular stations?

Rank	Name	Total Trips
1	Mackay / de Maisonneuve	97150
2	Metro Mont-Royal (Rivard / du Mont-Royal)	81279
3	Metro Place-des-Arts (de Maisonneuve / de Bleury)	78848
4	Metro Laurier (Rivard / Laurier)	76813
5	Metro Peel (de Maisonneuve / Stanley)	72298

Question:5.1 How is the number of starts and ends distributed for the station *Mackay / de Maisonneuve* throughout the day?

Time of Day	Total Start trips	Total end trips
evening	36781	31983
afternoon	30718	30429
night	12267	10326
morning	17384	26390

From the numbers we know that in the evening the traffic is the highest, followed by the afternoon time. The trips seem to slow down in the night probably being in the non-residential area. In the morning, number of start and end trips vary which confirms the point that the station is not in the residential zone. Most of the trips in the morning are ending at the station, staring from other stations.

5.2 explain the differences you see and discuss why the numbers are the way they are.

The high number of end trips in the morning represents that the station is in the centre parts of the city where there can be many workplaces and universities, where there is a lot of incoming traffic than outgoing. Evening numbers matches with the hypothesis too since people who came in the morning need to go back so start trips are ranking the highest in the evening.

The afternoon trips numbers can prove presence of a large-scale education institution and shopping and leisure centres in the vicinity.

Question:6.5 Where would you expect to find stations with a high fraction of round trips?

name	fract_percent
Metro Jean-Drapeau	30.1967
Metro Angrignon	23.3111
Berlioz / de ille des Soeurs	20.4346
LaSalle / 4e avenue	20.0602
Basile-Routhier / Gouin	19.3208
Parc Plage	18.4648
Gare Canora	17.9172
LaSalle / Senecal	14.7255
Casino de Montreal	14.3695
Quai de la navette fluviale	13.7603
de la Commune / Place Jacques-Cartier	11.0621
Jacques-Le Ber / de la Pointe Nord	11.0335
Place du Commerce	10.8181
CollVege edouard-Montpetit	10.0069

Station with high fraction of roundtrips can be expected in the area of tourist attraction, picnic spots and leisure centres. Because of a lot of trips are done by tourists and holidaymakers who are exploring and roaming the in the same vicinity.

Ankit Solanki
Data Science Program
Brainstation