Bixi Project—Part 1—Data Analysis in SQL

Daniel Mortensen

18-July-2021

INTRODUCTION

Bike sharing services are a popular and growing trend in many large cities. Bixi is one such service, located in Montréal, Canada. This report is aimed at answering certain key business questions to help Bixi grow its userbase and better serve its customers. The specific questions being answered are as follows:

Question Summary
1. What is the total volume of trips made each year, the total trips broken down by month for each year, and the average number of trips per day broken down by month and year?
2. Do members and nonmembers behave differently in terms of rental usage? What percentage of trips are taken by members vs nonmembers over the course of the year and by month?
3. At what time of year is demand for bikes the highest? When should a promotion be run to get nonmembers to sign up for the service as members?
4. What are the 5 most popular stations? Does using a subquery help answer this question more quickly?
5. At what parts of the day are people starting and ending their trips and why?
6. Which stations with at least 500 starting trips have at least 10% of their trips as roundtrips? Where are these stations likely to be located?

Results Summary

The popularity of Bixi bikes grew by 19% in 2017 compared to 2016. Peak months occur for both members and nonmembers during summer months, so any special offers aimed at gaining new members should be released during these months, especially July, wherein the most trips are taken. The most popular stations service thousands of trips each year, so special attention should be given to these stations to ensure optimum performance and availability. Evenings are the most popular time for riders to begin and end their rides, perhaps because of the warmer weather or because of work schedules. Relatively few stations have at least 10% of their trips being roundtrips. These trips and their corresponding stations are likely to be found in areas where people live, work, or visit.

METHODS

Data were taken from the provided Bixi SQL database, which contains usage data for the years 2016-2017. These data were analyzed using MySQL version 8.0.21. A copy of the SQL file used for analysis is provided in the accompanying folder. Graphics for this report were generated using Microsoft Excel 2016.

RESULTS & DISCUSSION

Trip Volume by Year and Month

In the years 2016 and 2017 there were 3.92 million and 4.67 million trips taken on Bixi bikes, respectively. That is an increase of 19% in trips taken in 2017 compared to 2016. When the number of trips is broken down by month, as seen in the Figure 1, it is observed that bike rentals are most popular in the warmest months. For example, in 2017, 861 thousand and 840 thousand trips were taken in the warm months of July and August, respectively, whereas the number of trips dropped more than 4-fold in the colder months of November and April, wherein only 150 thousand and 196 thousand trips were taken, respectively. It is also observed in Figure 1 that the greatest increase in number of trips for 2017 compared to 2016 occurred between July and October.
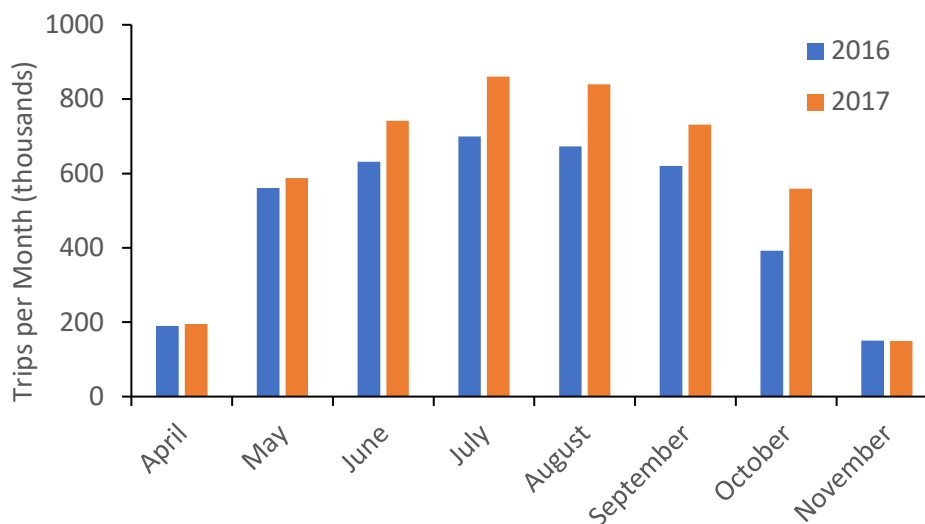

Figure 1. Bar graph showing the total trips per month in thousands of trips for the years 2016 (blue) and 2017 (orange).

A near identical trend is observed in the average number of trips per day for the individual months (Table 1). The highest number of average trips per day are seen in the warm summer months, the lowest average numbers of trips per day are in the cooler months, and the greatest growth by number of trips per day is seen the summer months.

Table 1. Average trips per day for each month and year. The color scale ranges from low values in yellow to high values in green.

	April	May	June	July	August	September	October	November
2016	6,331	18,099	21,050	22,556	21,703	20,675	12,661	5,004
2017	6,522	18,950	24,728	27,766	27,095	24,395	18,049	4,993

Member and Nonmember Trends

Member and nonmember trends were analyzed using the data from 2017. In this year, approximately 81% of all trips were made by members, meaning that only 882 thousand of the 4.67 million trips were made by nonmembers. In Figure 2, the percentage of trips made by members vs nonmembers is broken down by month. The month with the lowest percentage of trips taken by members, and conversely the month with the highest percentage of trip by nonmembers, is July. In this month, only about 76% of trips were made by members, whereas nonmembers took about 24% of the trips. The percentage of nonmember trips is lowest towards the opening and closing of the biking season, in the cooler months. In fact, in November fewer than 10% of trips were taken by nonmembers. This is likely because a member who relies on Bixi for regular transportation is more likely to keep using the service when it becomes uncomfortable to ride, compared to a nonmember who is not committed to the service in any tangible way and is thus more likely to choose a transportation method based on comfort.
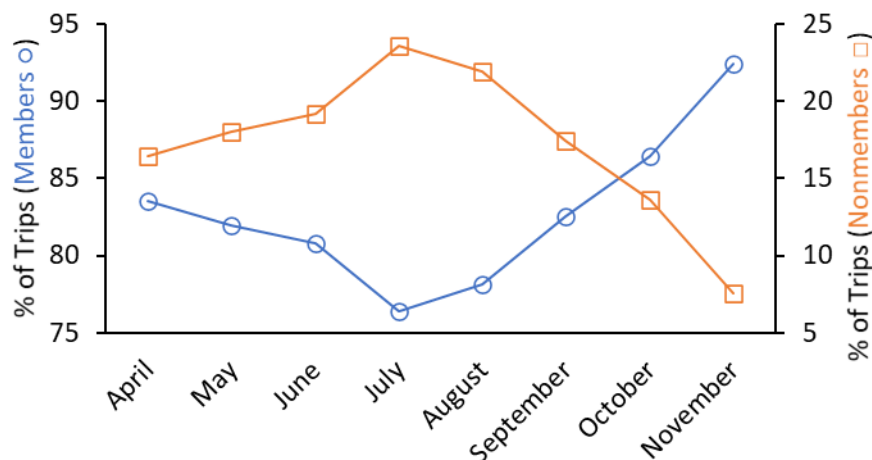

Figure 2. Number of trips made by members (blue circles, left axis) and nonmembers (orange squares, right axis) as a percentage of total trips per month for the year of 2017.

Promotion for Nonmembers

As has been noted previously, the summer months are the busiest for Bixi bikes, especially the months of July and August. Both members and nonmembers are more likely to use Bixi during these months as opposed to the cooler months. Because of the high number of nonmember riders using the bike service in July, I would recommend any special offer aimed at converting nonmembers to members be offered during this month. Although running a special offer during July would certainly reduce revenue for this month, this loss would hopefully be made up for by increased membership dues during less busy months, like November.

Most Popular Stations

The most popular stations were determined in order to optimize the supply of bikes and bike racks at these locations. The popularity of different stations was determined by counting the number of trips started from each station, including trips started in both 2016 and 2017. The top five stations determined in this way are listed in Table 2 in order of rank and total number of trips, with Mackay / de Maisonneuve being the most popular station at 97,150 trips.

Table 2. Most popular stations by total number of trips started between 2016 and 2017.

Rank	Station Name	Total Trips
1	Mackay / de Maisonneuve	97,150
2	Métro Mont-Royal (Rivard / du Mont-Royal)	81,279
3	Métro Place-des-Arts (de Maisonneuve / de Bleury)	78,848
4	Métro Laurier (Rivard / Laurier)	76,813
5	Métro Peel (de Maisonneuve / Stanley)	72,298

Station popularity was determined using SQL queries both excluding and including subqueries. These two methods resulted in identical results but dramatically different query run times (26.9 and 2.5 seconds, respectively), indicating that data is more quickly obtained when subqueries are used.

Behaviors by Time of Day

Trips starting and ending at the Mackay / de Maisonneuve station, the most popular Bixi station, were analyzed to determine when riders are using bike sharing services the most. 97,150 trips were both started and completed at this station between 2016 and 2017. For this analysis, the day was broken into four distinct time periods: "Morning" between the hours of 7 and 11, "Afternoon" between the hours of 12 and 16, "Evening" between the hours of 17 and 21, and "Night" between the hours of 22 and 6.

The percentage of trips started and ended during these different time periods is shown as a bar graph in Figure 3. There is a high similarity in the patterns for when riders start or stop

riding, indicating that most trips do not last for extended periods of time. The most popular times for trips are in the evening, when 38% of trips are started, and in the afternoon, when 32% of trips are started. These times are popular possibly because of the warmer temperatures experienced at these times, resulting in a more comfortable riding experience, or possibly because riders prefer to ride after getting off work. Further analysis is required to determine the exact reason for the observed trend. Night is the least popular time for riding, likely due to the lower visibility and cooler temperatures experienced during this time interval.
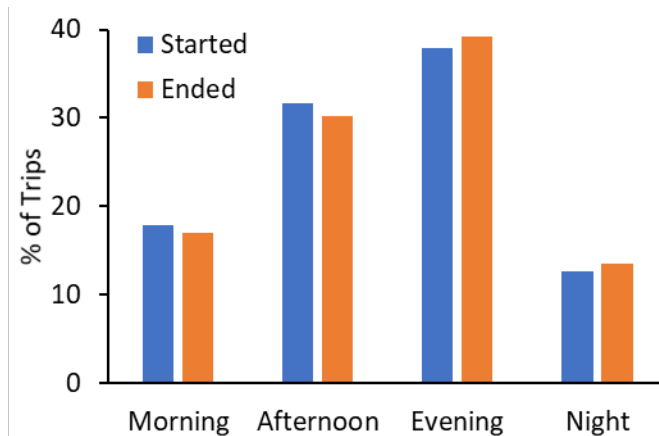

Figure 3. Percent of trips started (blue) and ended (orange) by time of day at Mackay / de Maisonneuve station. The different periods of time correspond to the following hours: "Morning" between 7 and 11, "Afternoon" between 12 and 16, "Evening" between 17 and 21, and "Night" between 22 and 6.

Popular Roundtrip Stations

      Rider trends were analyzed to determine which stations had the highest occurrence of roundtrips. For this analysis, stations were excluded if fewer than 500 trips were started from them. Trips were considered "roundtrip" if the starting and ending location of the individual trip were at the same station. Fourteen (14) station were found where at least 10% of trips were roundtrip (Table 3). The station with the highest percentage of roundtrips was the Métro Jean-Drapeau station, where 30.2% of trips were roundtrip. These stations are not likely located near other public transit stations, such as bus stops or train stations, where riders already have convenient transportation, but are more likely found near where people would need to start a trip, such as where people work, live, shop, or visit.

Table 3. Stations with at least 500 trips and at least 10% of trips being roundtrip, in order of highest percentage of roundtrips.

Station Name	Total Trips	% Roundtrip
Métro Jean-Drapeau	28,672	30.2
Métro Angrignon	2,398	23.3
Berlioz / de l'Île des Soeurs	5,246	20.4
LaSalle / 4e avenue	2,991	20.1
Basile-Routhier / Gouin	1,708	19.3
Parc Plage	6,201	18.5
Gare Canora	2,439	17.9
LaSalle / Sénécal	3,151	14.7
Casino de Montréal	6,138	14.4
Quai de la navette fluviale	6,417	13.8
de la Commune / Place Jacques-Cartier	50,822	11.1
Jacques-Le Ber / de la Pointe Nord	2,719	11.0
Place du Commerce	8,569	10.8
Collège Édouard-Montpetit	1,439	10.0

CONCLUSION

Bixi bike sharing services are extremely popular in Montréal, growing by nearly 20% in a single year. The majority of this growth was seen in summer months, which are the most popular months for riding in general. The popularity for these months is related to the local temperature, which dips too low during winter months to continue the service. Summer months are also when most nonmembers use Bixi services, suggesting that a promotion aimed at getting more users to sign up for memberships would be most successful during these months, especially during July and August, the most popular months.

The most popular stations tend to be in metro areas, where users tend to spend more time on foot. Special care should be given to optimize these locations so that rider happiness is assured and profits are maximized. The majority of riders prefer riding either in the evening or afternoon hours, probably because of the warmer temperatures experienced during these later hours of the day. The least popular time to ride is at night, when poor lighting and colder temperatures make bike riding inconvenient. The majority of trips are one-way; however, more than a dozen stations have at least 10% of their trips being roundtrip. These stations are likely located near where people live, work, or often visit.