Bixi Project - Part 2 - Visual Analytics in Tableau

Daniel Mortensen

25-July-2021

Question 1

1. Build a visualization to contrast the total number of monthly trips for the calendar years of 2016 and 2017 by month. What differences do you notice about the usage of the Bixi service between the two years?

Answer: Figure 1.1 shows the total number of trips for each month in the years of 2016 and 2017. A significant increase is seen in the months of July through October. A milder increase is also observed in May and June.

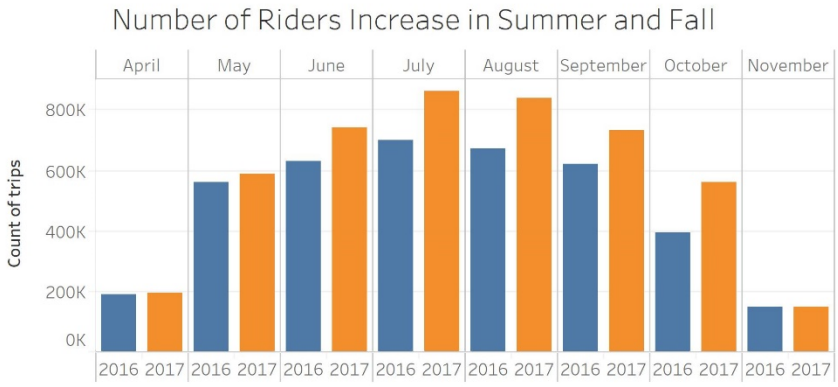

Figure 1.1 Total trips per month for 2016 (blue) and 2017 (orange).

2. Use a quick calculation to contrast the percentage of trips that occurred in each month per year, between 2016 and 2017; e.g. if 1000 trips occurred in 2017 in total, and 120 of them occurred in July, then July has 12% of trips in 2017. How does the proportional monthly usage differ between 2016 and 2017?

Answer: Only a modest shift is observed for most months when the percentage of trips taken over the whole year is analyzed on a month-by-month basis (Figure 1.2). There is a general trend of decreased percentages in the earlier months and increased percentages in the later months, starting with July. The only exception to this observation is in November, wherein a slight decrease is observed. Significant changes were only observed for May and October, wherein a substantial decrease and increase were observed, respectively. These shifts do not correlate with a decrease in popularity for the earlier months, but rather an increase in popularity for later months.

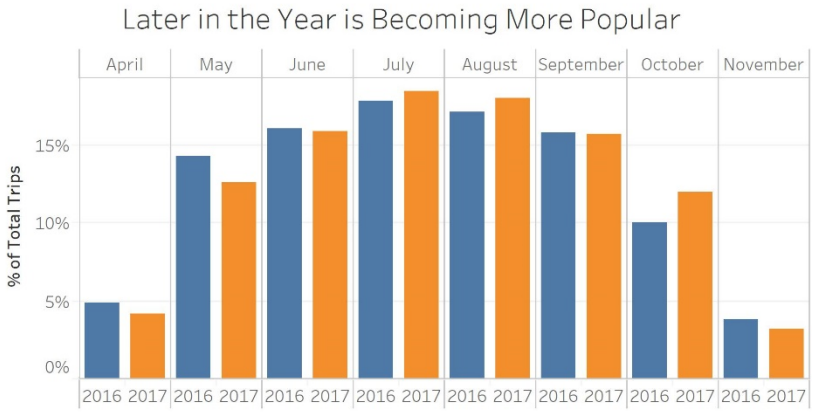

Figure 1.2 Percent of total trips taken for the years 2016 (blue) and 2017 (orange) broken down by month.

3. Make a calculated field to calculate the percentage of trips that were done by members, and using this, visualize what percentage of trips per month were member trips for the year 2017.

Answer: The membership status of every rider on every bike trip is recorded, thus allowing for the percentage of trips taken by members to be calculated (see the "Percentage by Members" calculation in the attached Tableau file). The percentage of trips taken by members has been broken down by month in Figure 1.3. Members make up the vast majority of riders, but the percentage of trips that are taken by members dips significantly in July and August, the most popular months for riding in general. Thus, it can be concluded that these months are most popular for nonmember riders.

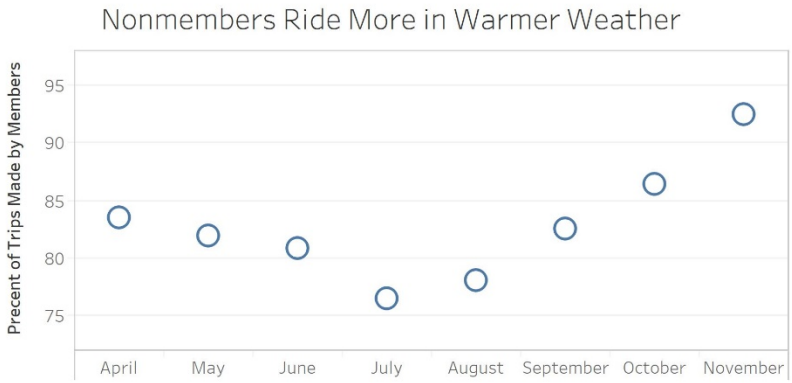

Figure 1.3 Precent of rides made by members for each month in 2017.

4. Create a calculated field for identifying round trips (hint: you will need to use a calculated field with conditionals). Create a visualization showing the top 10 stations by percentage of round trips.

Answer: The 10 stations with the highest percentage of trips resulting in a roundtrip are shown in Figure 1.4 in descending order (see the "Roundtrip" calculation in the attached Tableau file). Métro Jean-Drapeau has the highest percentage of roundtrips and is located on St. Helen's Island, which is part of Jean-Deapeau Park. Riders likely take rides in this location for leisure and for sightseeing and are less likely to need to get to a particular end location. Thus, these riders are more prone to return their bikes to the stations from which they were checked out.

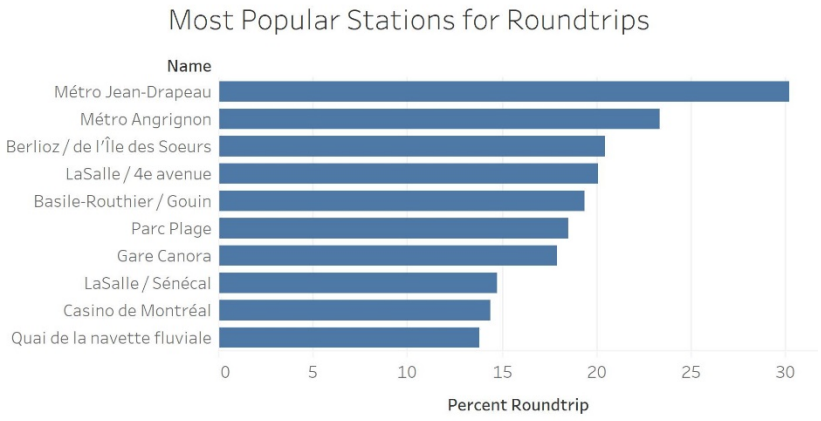

Figure 1.4 Percent of trips that are roundtrip for the top 10 stations with the highest percentage of roundtrips.

Question 2

1. Build a visualization for marketing showing the relationship between percentage of round trips and percentage of member trips by station. Comment upon/interpret any interesting patterns you see.

Answer: Figure 2.1 shows the percentage of roundtrips as a function of the percentage of rides made by members for each station. The black line in this figure corresponds to a best fit line. Although for most stations there does not appear to be a strong relationship between membership status and whether a trip is roundtrip or not, the percentage of trips made by members is significantly lower for stations where roundtrips are more common. For example, at the Métro Jean-Drapeau station, which has the highest percentage of roundtrips of any station, only about 24% of trips were made by members. These high-roundtrip, low-member stations are likely located near tourist attractions, where nonlocals are more likely to be sightseeing or engaging in other recreational activities, rather than seeing to either personal or professional needs.

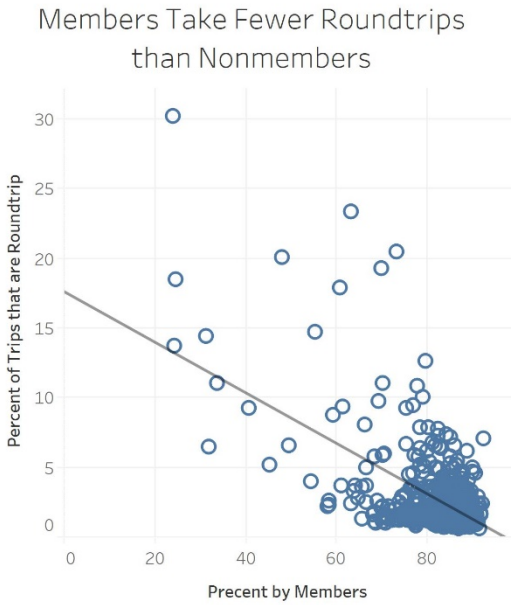

Figure 2.1 The percent of roundtrip for each station as a function of the percent of trips made by members at those stations. The black line is a best fit line to the data.

2. Make a histogram (or histograms) to visualize the distribution of all trips by duration in minutes, and contrast this between member and non-member trips. What can be said about the behavior of members vs. non-members in terms of trip length?

Answer: The top panel of Figure 2.2 shows a histogram of all trips broken down by trip duration in minutes. The average trip lasts about 13.7 minutes. The bottom panel of this same figure shows histograms of trips taken by members (orange) and nonmembers (red). Trip durations lasting longer than 50 minutes comprise only about 1% of all trips taken and are, thus, not included in this plot for simplicity of visualization. These longer trips are, however, included in the calculated average trip times shown on the figure. Although nonmembers take significantly fewer trips than members, the average trip length for a nonmember

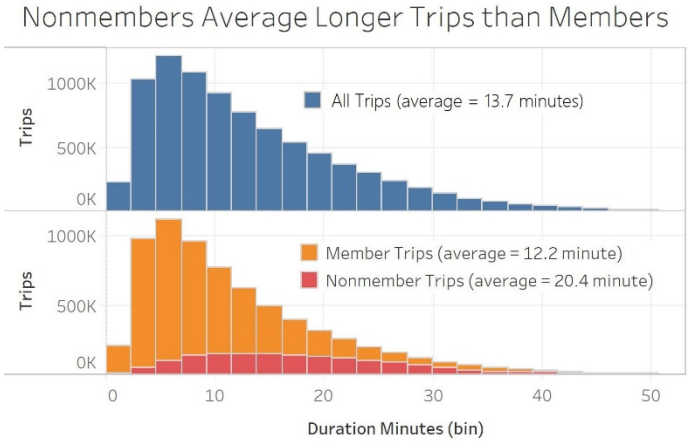

Figure 2.2 Histograms of trip duration in minutes for (top panel) all trips taken and (bottom panel) trips made by members (orange) and nonmembers (nonmembers). For ease of display, trip durations above 50 minutes were excluded from these data. Trips above 50 minutes account for ~1% of all trips. The maximum trip time is approxiamatly 120 minutes.

(20.4 minutes) is about 67% longer than the average trip taken by members (12.2 minutes). This difference likely occurs because nonmembers are more prone to use Bixi bikes for longer sightseeing rides, whereas members are more likely to simply use the bikes to get where they need to go in a rapid manner.

3. Create a map to visualize the average trip duration per station across the city. Are there any interesting geographic patterns you notice? Why do you think this might this be?

Answer: The requested map is shown in Figure 2.3, and an interactive version of this map has been included in the requested dashboard. A trend observed in this map is that shorter trips tend to occur in business areas, where people tend to use the bikes to get from one place to another quickly, whereas longer trips tend to occur more in residential or recreational areas.

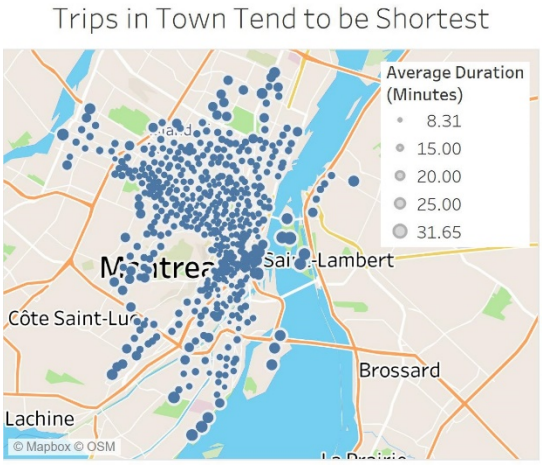

Figure 2.3 Map of all stations. Marker size corresponds to average trip length.

Question 3

1. Create a calculated field (or fields) to calculate the revenue generated by the pricing model.

Answer: Trips were segmented by trip length in the "Pricing Model Lengths" calculation in the attached Tableau file, and the corresponding fees were associated with these trip lengths using the "Pricing Model Fees" calculation in the same file.

2. What are the total dollar amounts and relative percentage of revenue from single trips up to an hour in length for each of the three different pricing buckets above?

Answer: Total revenue is broken down by category of trip length in Figure 3.2. Short trips, those lasting no more than 30 minutes, earn the lions share of the revenue, comprising 88.4% of all earnings. Medium length trips, those between 30 and 45 minutes, bring in only 7.9% of total revenue, and trips lasting longer than 45 minutes do not even earn break the million-dollar mark and bring in only about 3.6% of revenue. Clearly, the short trips category is the Bixi revenue powerhouse.

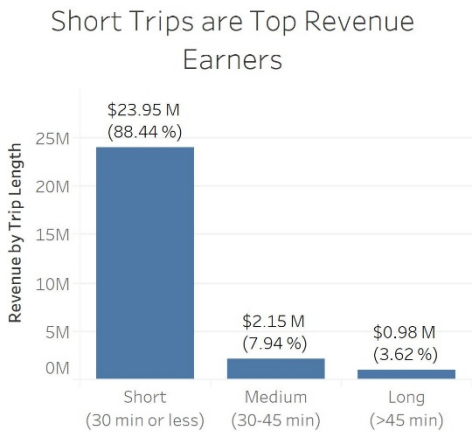

Figure 3.2 Total profit (and percent of total profit) broken down by trip duration.

3. The Director of Finance is not satisfied with the above insights, and wants very detailed information on exactly when they are seeing the most revenue from single trips 30 minutes or less. Create a visualization to show the total amount of flat rate revenue in the data for each hour and each day of week (Monday, Tuesday, Wednesday, etc.). At which days/times is Bixi generating the most revenue from their flat rate charge? You may use a new calculated field for this question

Answer: A breakdown of total revenue by hour of the day for each day of the week is shown in Figure 3.3. The pattern is significantly different for weekdays and weekends. For weekdays, the most revenue is earned around the hours of 8 and 17, which roughly correspond to the times many people begin and end work, respectively. There is also a small spike on weekdays around the lunch hour. For weekends, the greatest activity is seen around the hours of 15 and 16 and likely correspond more to leisure activities than do the behaviors observed midweek.

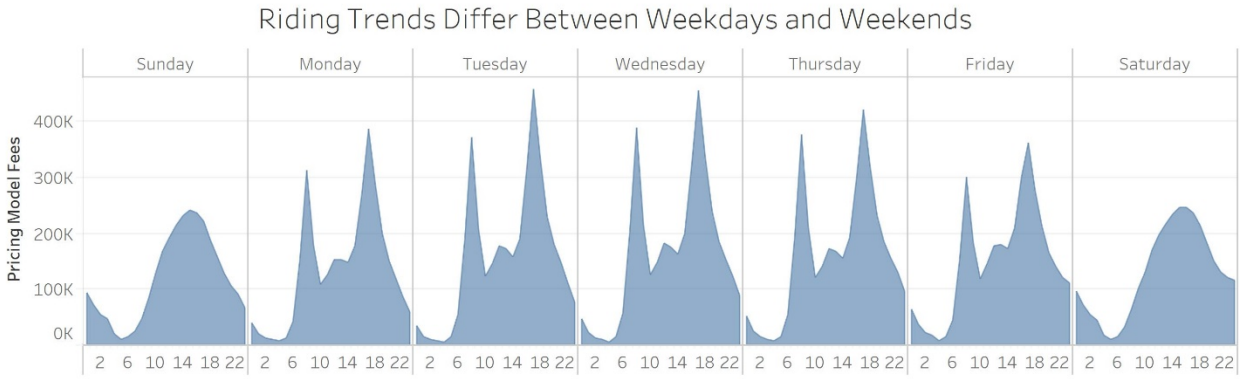

Figure 3.3 Total revenue by hour of the day for each day of the week for trips lasting no longer than 30 minutes.

Question 4

1. Create a dashboard containing at least 3 visualizations, using two you've already created thus far as well as one additional new one of your choosing.

Answer: A preview of the requested dashboard is shown below. Included in the dashboard are the breakdown of revenue by hour of the day for each day of the week, as shown in Figure 3.3, the map with markers scaled to represent the average trip duration for each station, as shown in Figure 2.3, and an additional plot of the average trip duration and percentage of trips taken by members broken down by hour of the day.