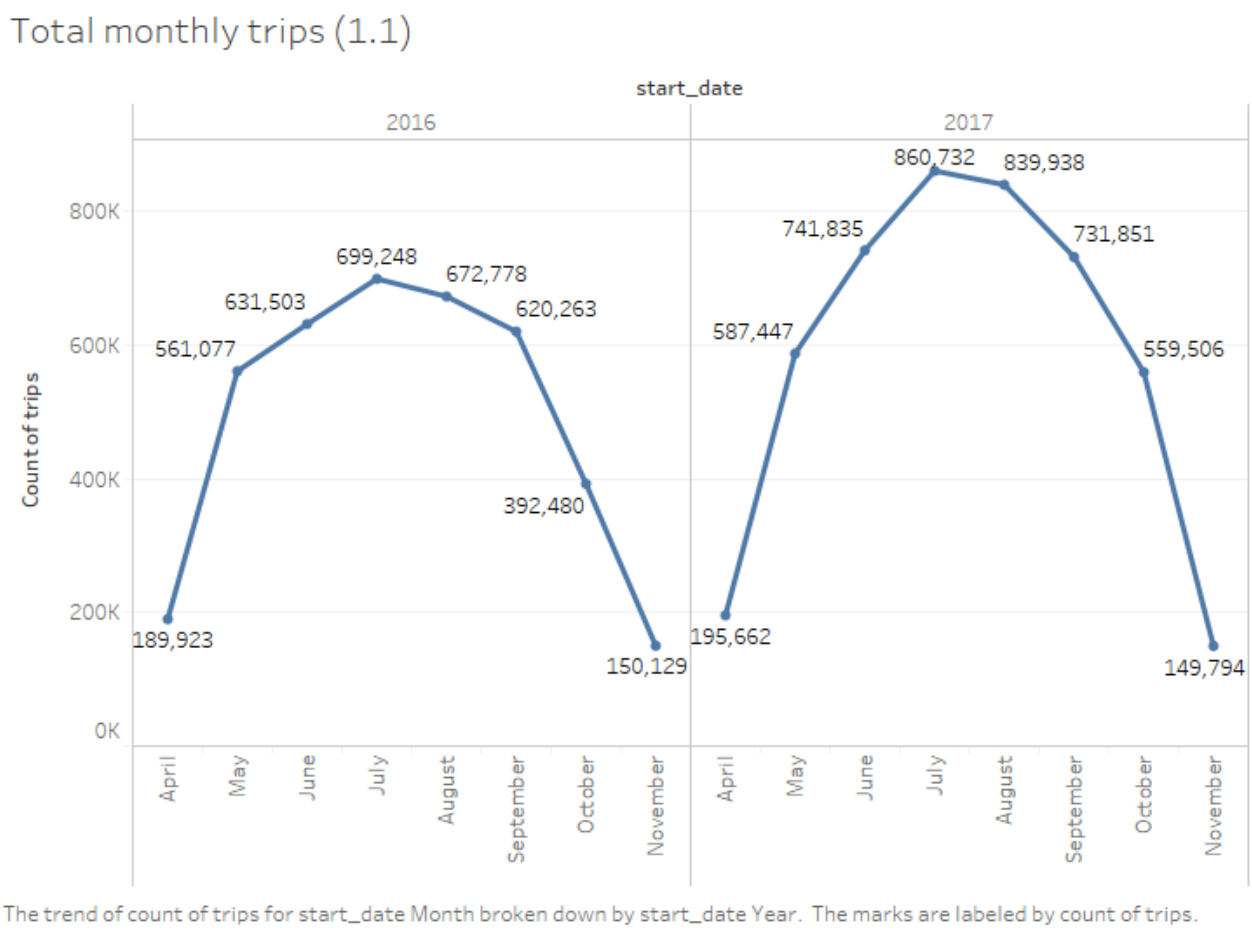
BIXI DELIVERABLE 2

Sohail Shaikh
12-017-2020



1.1 What differences do you notice about the usage of the Bixi service between the two years?

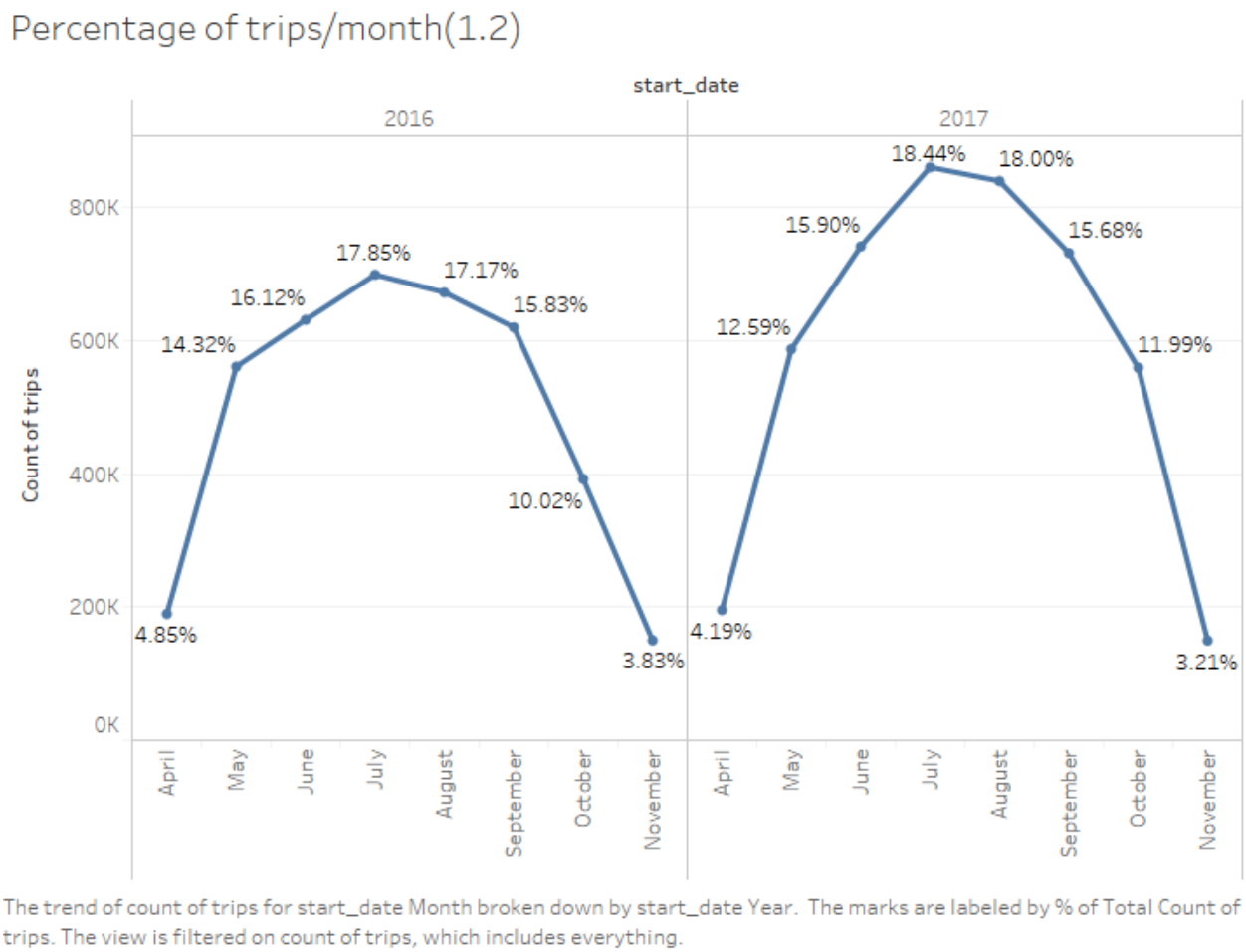
Answer: The following chart shows the total number of trips for each month for the years 2016 and 2017. The chart is showcased this way to allow viewers to easily draw conclusions based on the comparison of the data.

As summer approaches, for both 2016 and 2017, the number of trips reaches its peak in July. For both years, after July, the numbers begin to drop. I am assuming the number of total trips decreases after July due the end of summer approaching and the weather getting colder throughout the rest of the year.

Overall, the number of trips in 2017 per month were significantly higher than the same months in the previous year. The summer months of July to August showed increased numbers of around 200k trips in 2017 compared to 2016.

For both 2016 and 2017, there are no trips taken in the months of January, February March and December. I assume Bixi is not in service during this time due to winter weather.

In 2017, each month has a higher number of total trips than the same time the previous year, except for the month of November, where November 2017 had a lower amount of total trips vs November 2016.



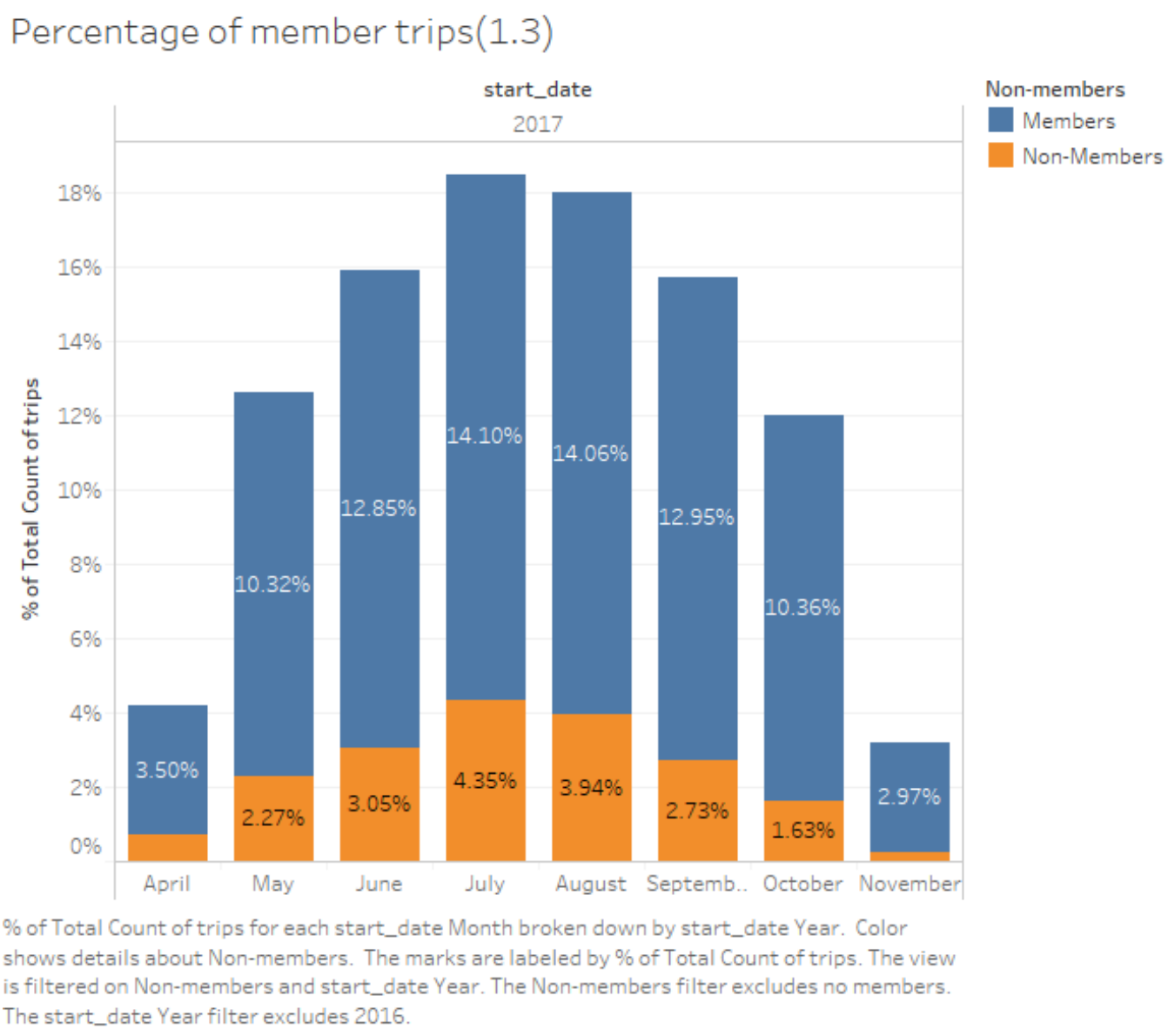
This chart depicts the percentage of trips per month for the years 2016 and 2017.

1.2 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: For 2016, the percentage increases from 4.85% in April to 17.85% in July whereas for the year 2017, the percentage increases from 4.19% and reaches to 18.44% for July.

For the month of November the percentage is the lowest for both years accounting 3.83% for 2016 and 3.21% for 2017. The same applies to April with April 2016 accounting for 4.85% and April 2017 accounting for 4.19%

For the remaining months, the percentage of trips in 2017 is significantly higher than the year 2016 except the months of January, February, March and December which accounts to 0% of trips.

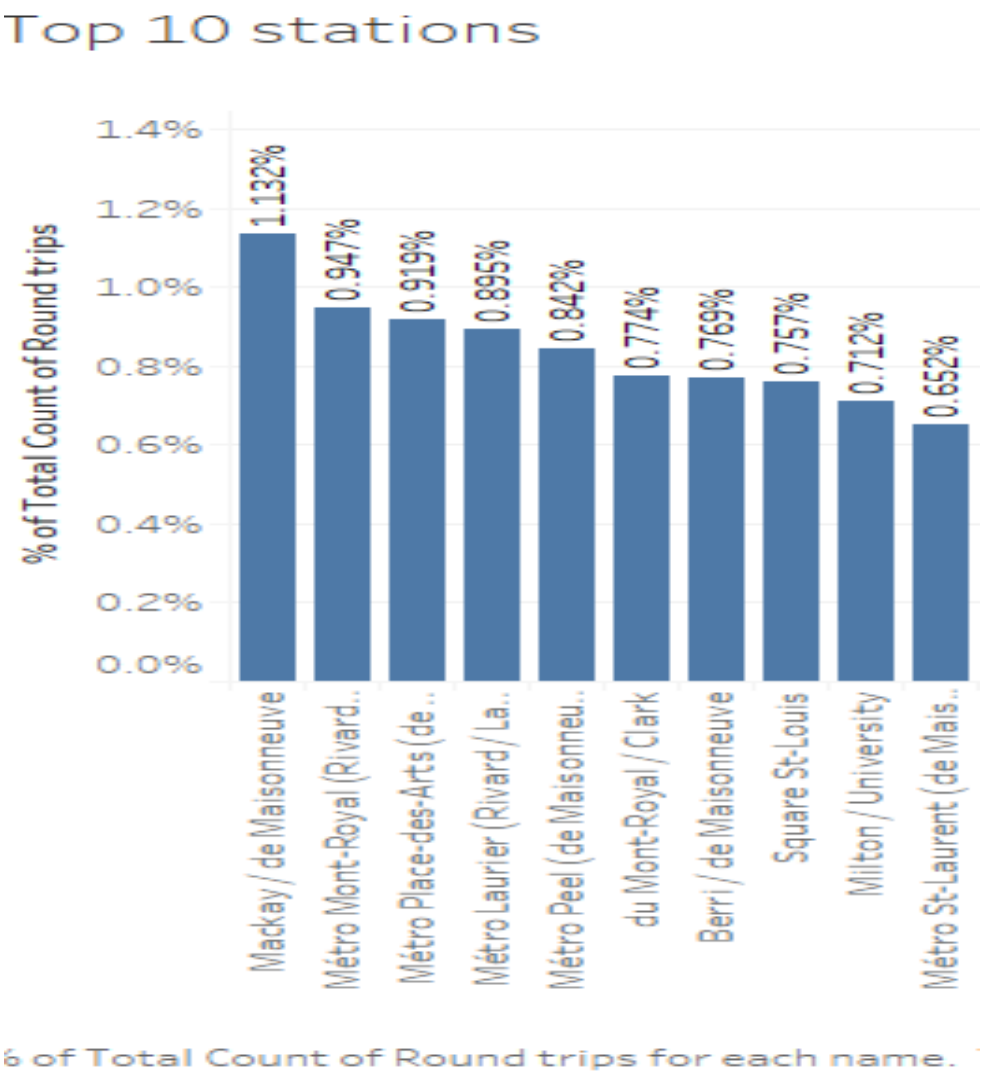


The chart shows the percentage of trips taken by members and non-members

1.3 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 field for members can be easily made by putting in [is_member] = 1. This will give a true value to Members and False value to non-members. The chart answers stating the percentages of the trips taken by members which is shown by blue parts of the bar graph. The non-members are shown by the orange part of the bars.

It can be seen that the member trips had higher percentage than non-members for every month.



The above chart shows the highest percentage of round trips by the top 10 stations

1.4 Create a calculated field for identifying round trips. Create a visualization showing the top 10 stations by percentage of round trips.

Answer: Round trips can be calculated by creating a calculated field

IF [start_station_code] = [end_station_code] THEN "Round trips" ELSE "NULL" END

As shown, the data shows the percentage from the overall percentage of all the stations for the top 10 stations.

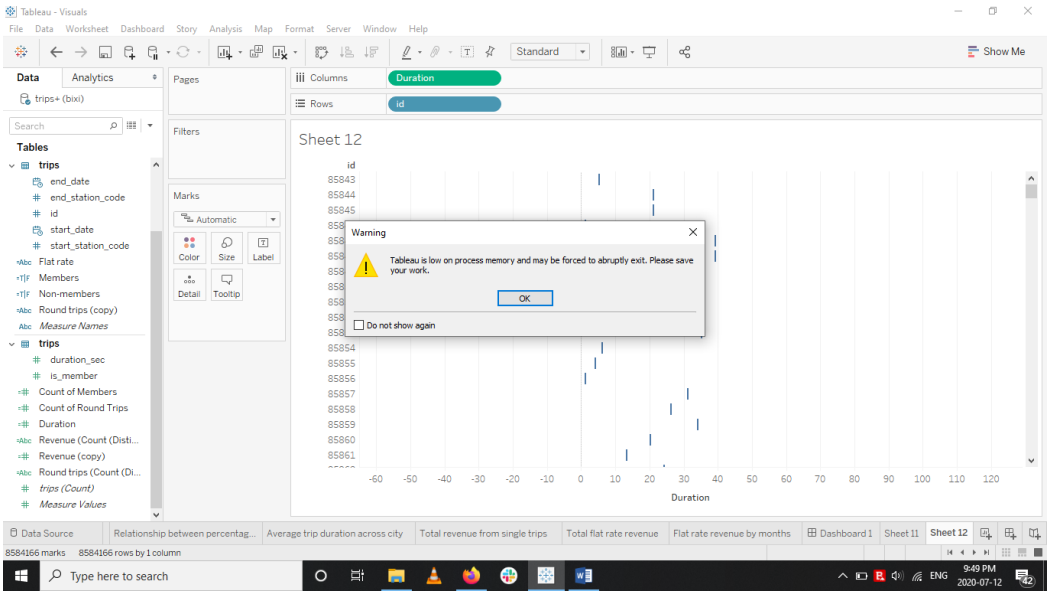
The most round trips were taken by the Mackay/ de Maisonneuve which was 1.132% from the total stations and so on.

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: The relationship can be obtained by taking into account the percentage of round trips and the number of trips taken by the members including all stations.

The visual as seen in the workbook is vast as it encompasses all the stations. The relationship between the round trips and member trips including all the stations is linear with values ranging from 0 to 0.2%. A majority of the values are lying from 0 to 1 range with some stations having value around 0.2%

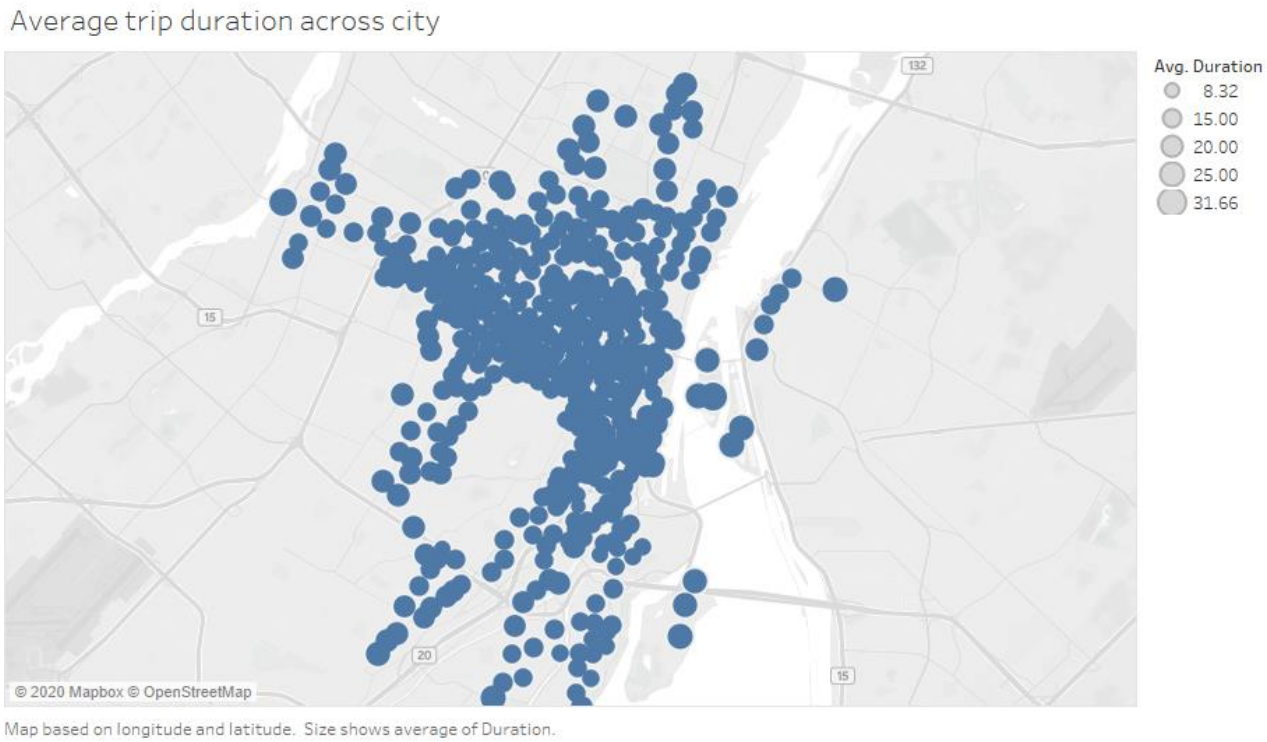
2.2 Make a histogram (or histograms) to visualize the distribution of all trips by duration in minutes, and contrast this between member and non-member trips. What can be said about the behaviour of members vs. non-members in terms of trip length?



The process of getting to the histogram would be to plot id for all the trips with the calculated duration field and adding members field to give data for members as well as non members.

Making an assumption the highest trips with more duration would be members and it would vary for each trip.

(Note: I couldn't complete the query because tableau couldn't respond with all the trips).



2.3 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 map can be obtained by taking into account the trip duration and using the map function to display the stations spread along the city.

It can be seen that there is a higher number of trips taken in the centre of the city. It can be assumed that the trips are mostly saturated in the downtown area.

As we go farther from the centre the symbols goes on the decreasing at the periphery of the city but it can be also noticed that the trips are of longer duration on the outskirts compared to the inner city short trips.

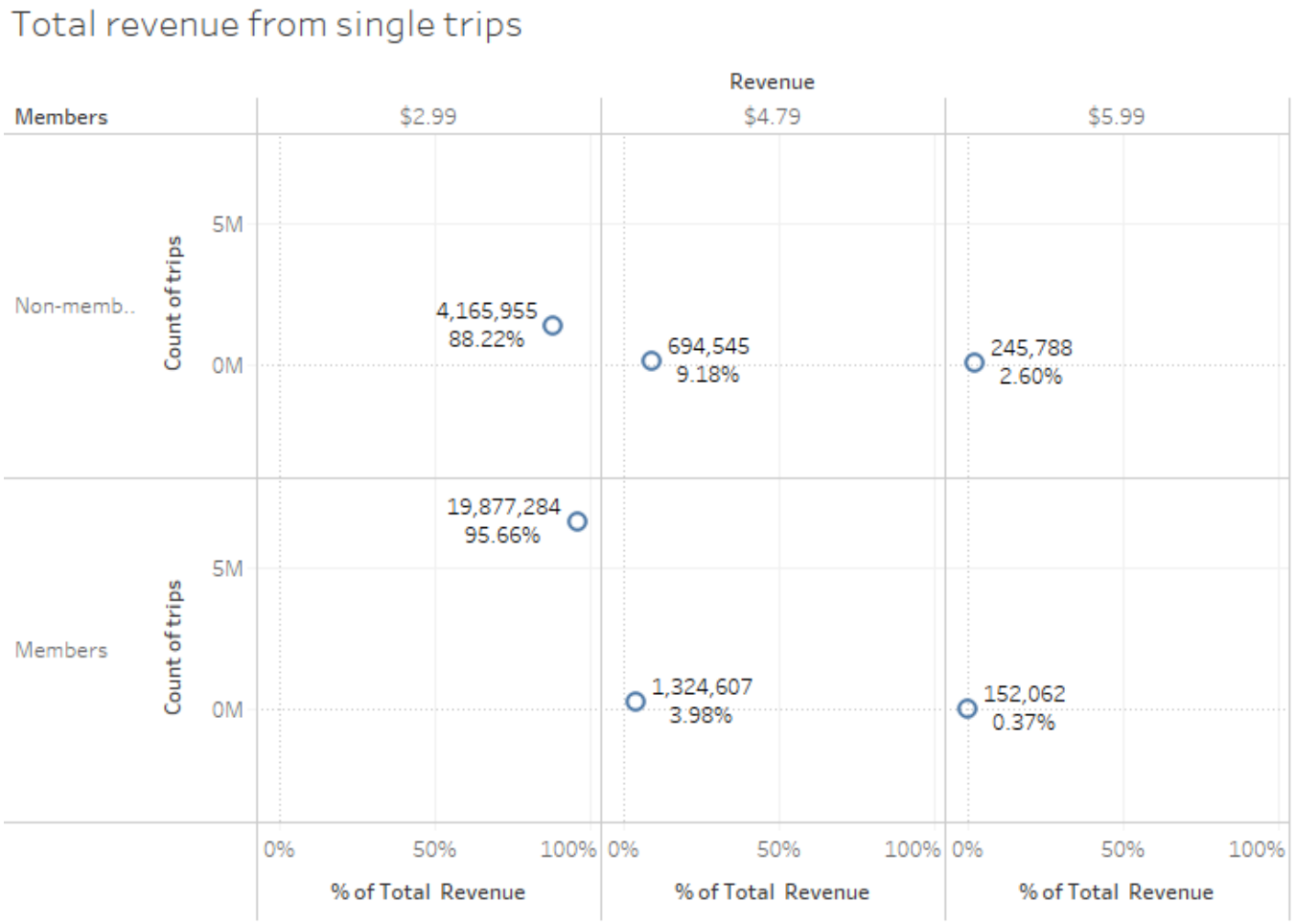
3.1 Create a calculated field for the revenue generated by flat rate trips (30 min or less)
3.2 Create a second calculated field for the revenue generated by trips 30-45 minutes in length.
3.3 Create a third calculated field for the revenue generated by trips 45-60 minutes in length.

Answer: All the calculated fields can be merged into one as

```
IF [Duration] <= 30 then '$2.99'
ELSEIF [Duration]<=45 then '$4.79'
ELSEIF [Duration]<=60 THEN '$5.99'
ELSEIF [Duration]<= 75 THEN '$8.99'
ELSEIF [Duration]<= 90 THEN '$11.99'
ELSEIF [Duration]<= 115 THEN '$14.99'
```

```
ELSEIF [Duration]<= 130 THEN '$17.99'
ELSEIF [Duration]<= 145 THEN '$20.99'
ELSE '$23.99'
END
```

This takes into account all the trips to be safe.



% of Total Count of Revenue vs. count of trips broken down by Revenue vs. Members. The marks are labeled by sum of Revenue (copy) and % of Total Count of Revenue. The view is filtered on Members and Revenue. The Members filter keeps Non-members and Members. The Revenue filter has multiple members selected.

3.4 What are the total dollar amounts and relative percentage of revenue from single trips less than an hour in length for each of the three different pricing buckets above?

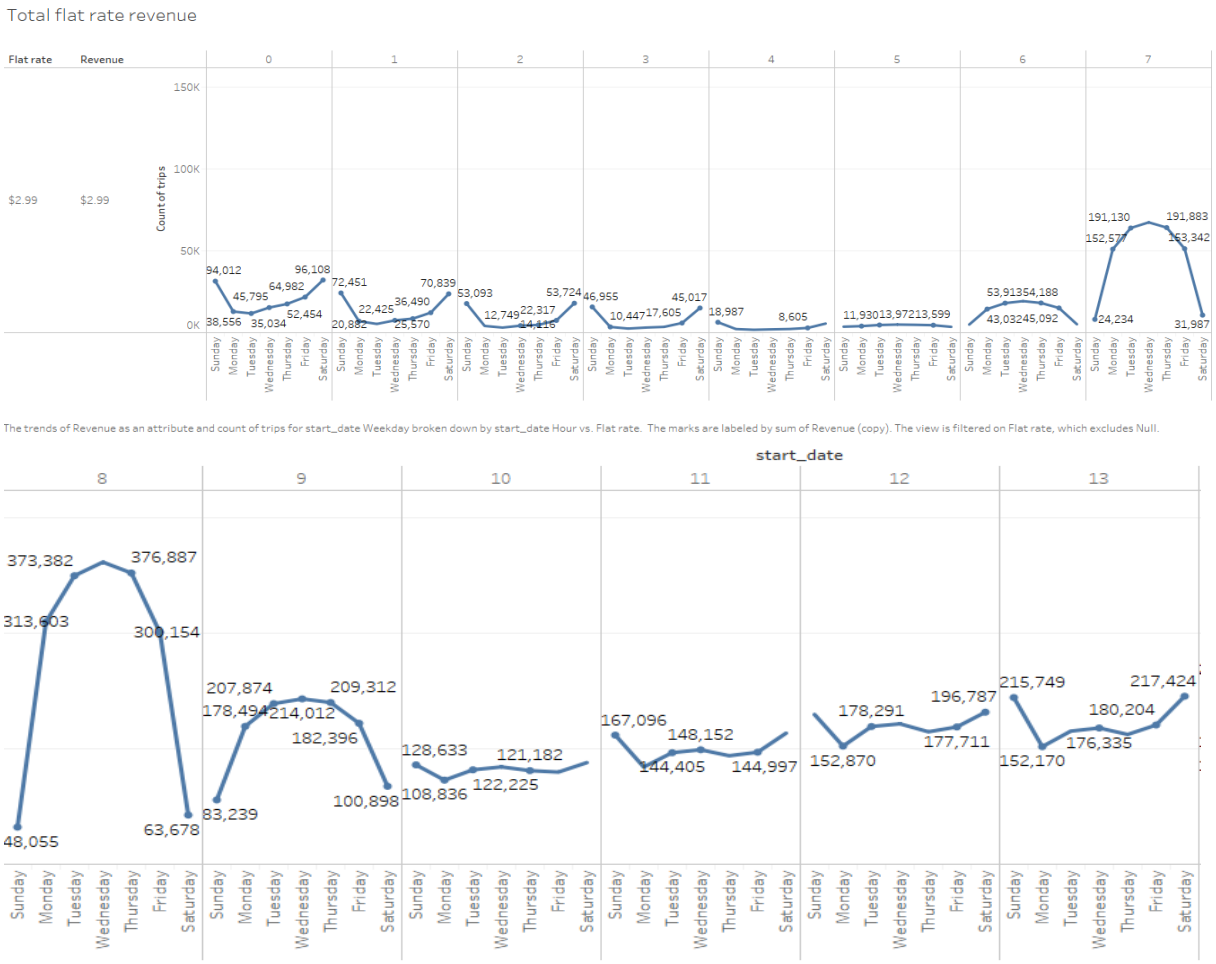
Answer: The chart above showcases the amount of revenue and relative percentage of revenue from single trips less than an hour in length for three different pricing buckets: $2.99, $4.79 and $5.99.

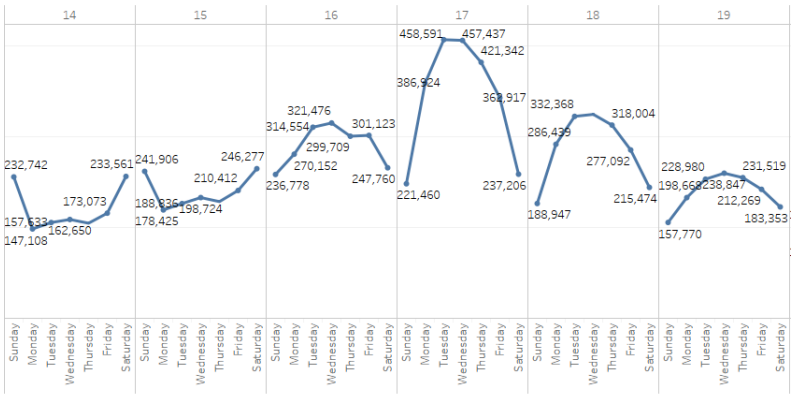
From the data in the chart, it can be seen that trips under 45 minutes generate more revenue as well as a larger percentage of the overall revenue. It can be assumed that this category generates more revenue because the price is lower.

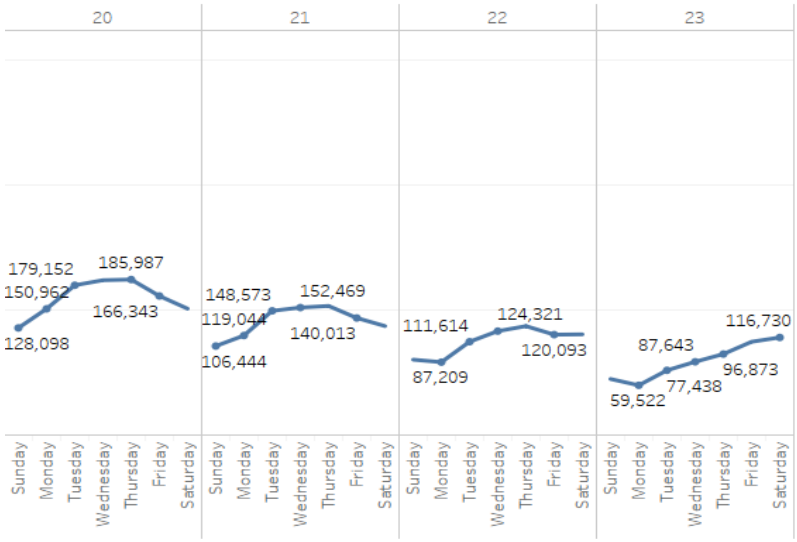
Trips between 45 minutes and 60 minutes, which cost $4.79, generate the second highest revenue while trips 60 minutes and above at $5.99 generate the least amount of revenue. It can be assumed from these numbers that customers prefer to take shorter trips at a lower cost than to pay more for a longer trip.

3.5 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?

Answer:



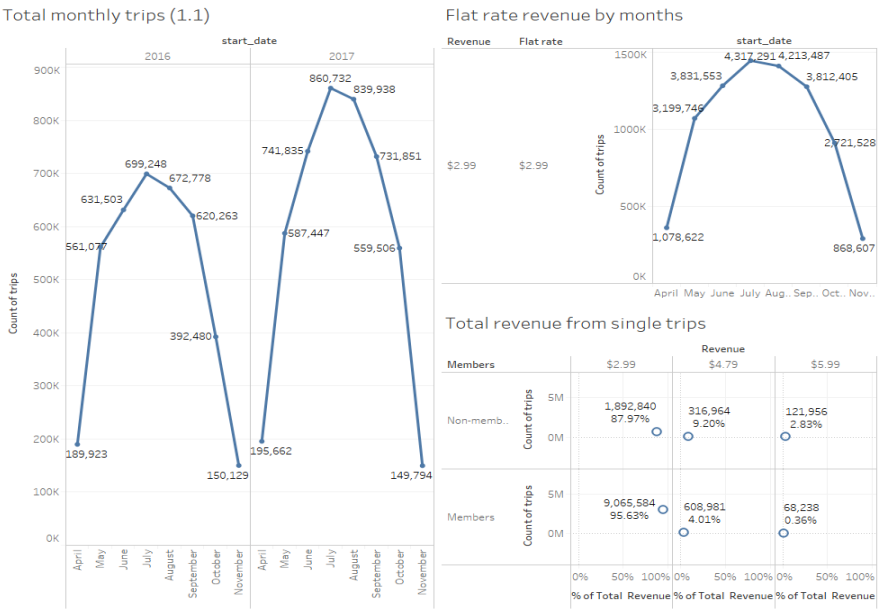




The following days/times are generating the most amount for the flat rate charge:

Wednesdays at 7 a.m., 8 a.m., 9 a.m., 5 p.m., 6 p.m. and 7 p.m.

From these numbers, it can be assumed that a majority of BIXI users are using the service to commute to and from work, as the most popular times are during the morning and evening rush hours when a majority of business professionals are working.



4. Finally, the operations team wants an interactive reporting dashboard they can use to drill into the data and get out insights as required, as well as fulfill lots of ad hoc data requests for data that they receive. 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: Based on this dashboard, it can be seen that the total number of trips per month peaks in the summer months, particularly in July. July is also where the most revenue is most generated.

Trips above $4.79 and $5.99, which are 45 – 60 minutes in length, are gaining more revenue in comparison to trips that are longer than an hour.

Trips that are 30 minutes or less are priced at $2.99 and generate the most revenue based on the categories of trip lengths.

Based on this information, it is recommended that BIXI give discounts to non-members in the summer months for longer trips to generate more revenue and gain more customers as BIXI's existing customer base gravitates towards shorter trips that are lower in cost during the summer months.