Cyclistic Case

1. Business Task

Identify the similarities and differences between the behaviors of the two consumer groups of the company, also, as trends and spending behaviors of the two groups in order to provide data-supported recommendations about the company's marketing campaign to be presented to the executive team, which is designing marketing strategies aimed at converting casual riders into annual members.

Key Stakeholders

- Marketing Director (Lily Moreno),
- Executive Team,
- Marketing Analytics Team,
- Cyclistic Users

2. Data Sources Used

The main data source used is the publicly available "12 months of Cyclistic trip data" for the time period 2013-2021. Data is stored on the local computer and organized by year. These data can be used to complete the business task since it includes data for both groups (causal and annual).

Integrity

Data integrity is good, however, there are some missing data points that can be corrected or removed during data cleaning. Some files have different headings which should be fixed during the cleaning. There seems no bias in the initial preview and data is structured in the same way for different tables.

Limitations

In order to protect the privacy of the users, personally identifiable information is not available to use which prevents connecting pass purchases to credit card numbers to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

3. Documentation of any cleaning or manipulation of data

1. Headings are separated by delimiter "_" and standardized for all of the files.
2. All data sorted according to "start_time" and "trip_duration_s" columns.
3. Dates are formatted as "d.mm.yyyy hh:mm".
4. New columns added as "ride_length" (hh:mm:ss) and "trip_day" (general).
5. Trips with inconsistent dates are deleted.
6. Tables merged into a single file.
7. Trips longer than 120 seconds have been analyzed.

Further information is available in the SQL log.

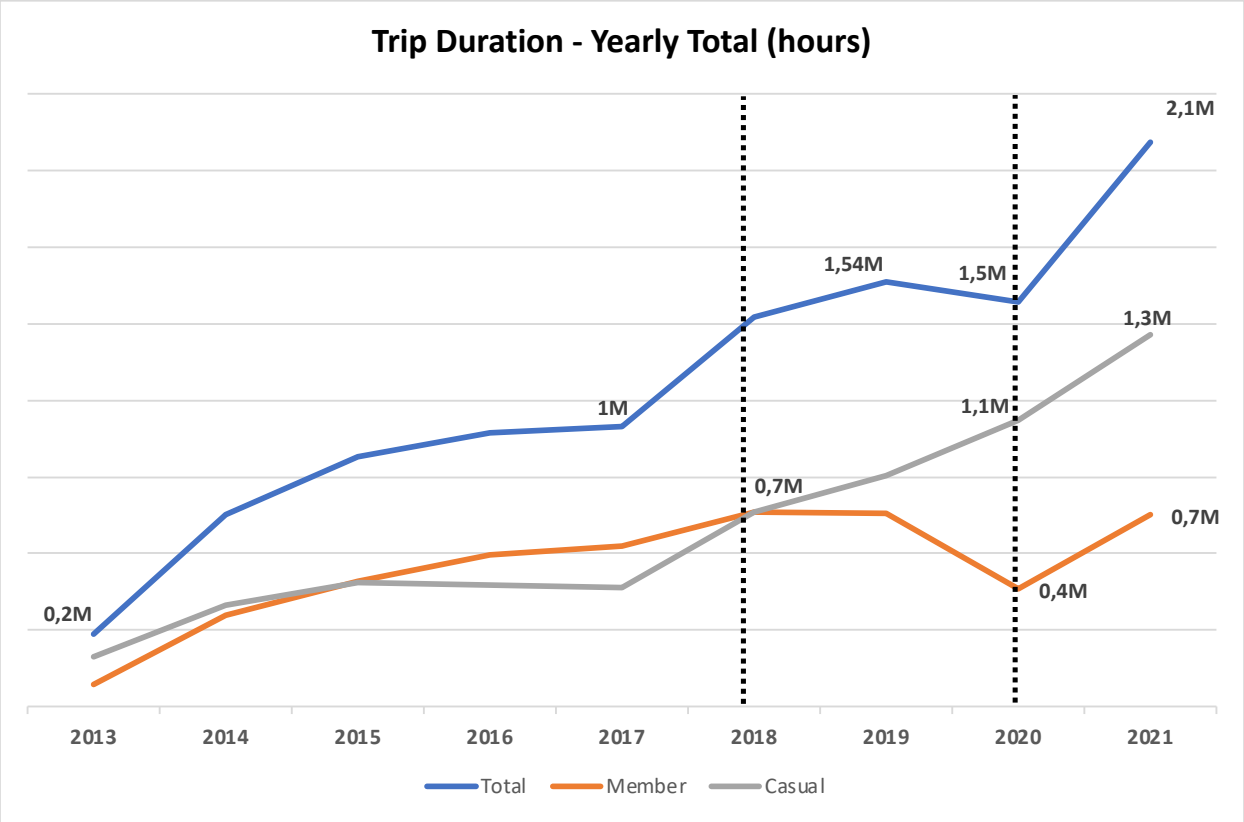
4. Summary of Analysis

After investigating the data from 2013-2021 which contains more than 30M observations, behavioral differences between the two customer groups have been realized. Also, it is obvious that the sales of the company are increasing over time. However, after 2019, Covid-19 period, there has been a slight decrease in the total number of trips and total trip duration for a year. It is

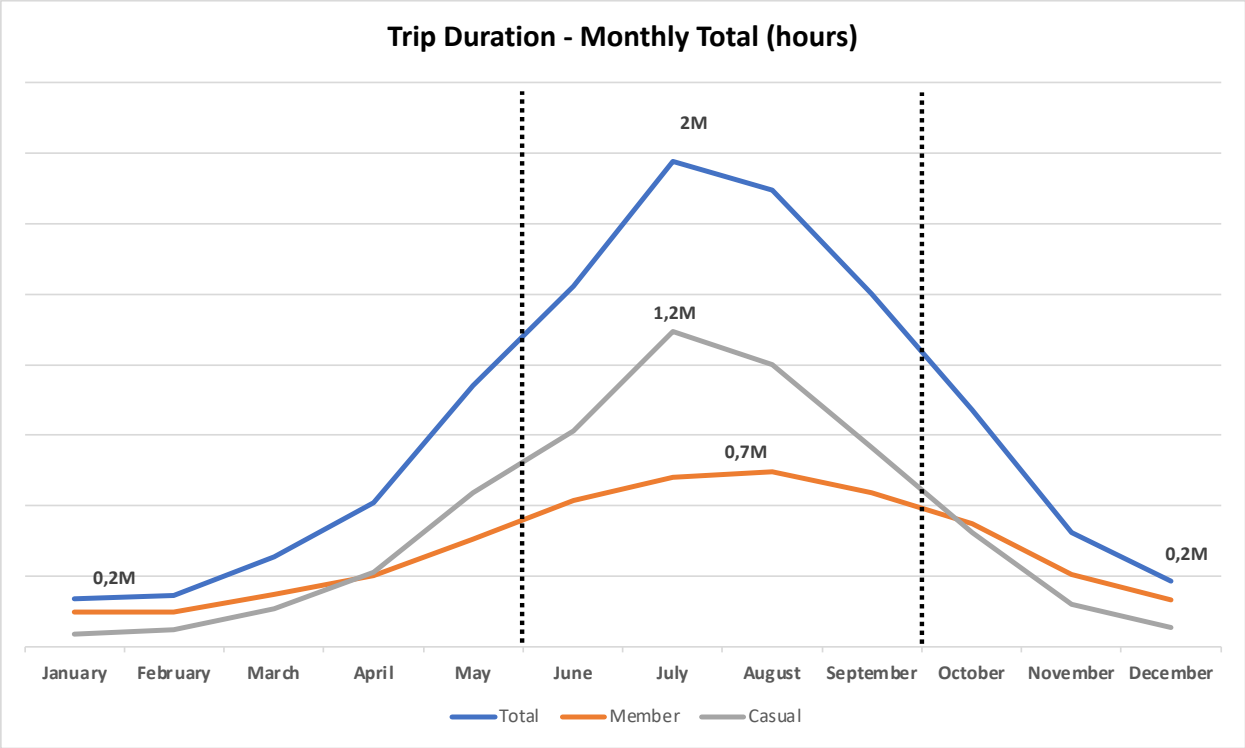
also interesting that the number of members realized a sudden decrease during Covid-19 while the number of casual riders increased.

Monthly analyzes show us that the summer period is the peak period of any year in terms of the number of trips and total trip duration for both customer groups. On the other hand, customer groups differ in daily usage. Casual riders seem to use the bikes mostly on weekends while members are using them slightly more on weekdays than on weekends.

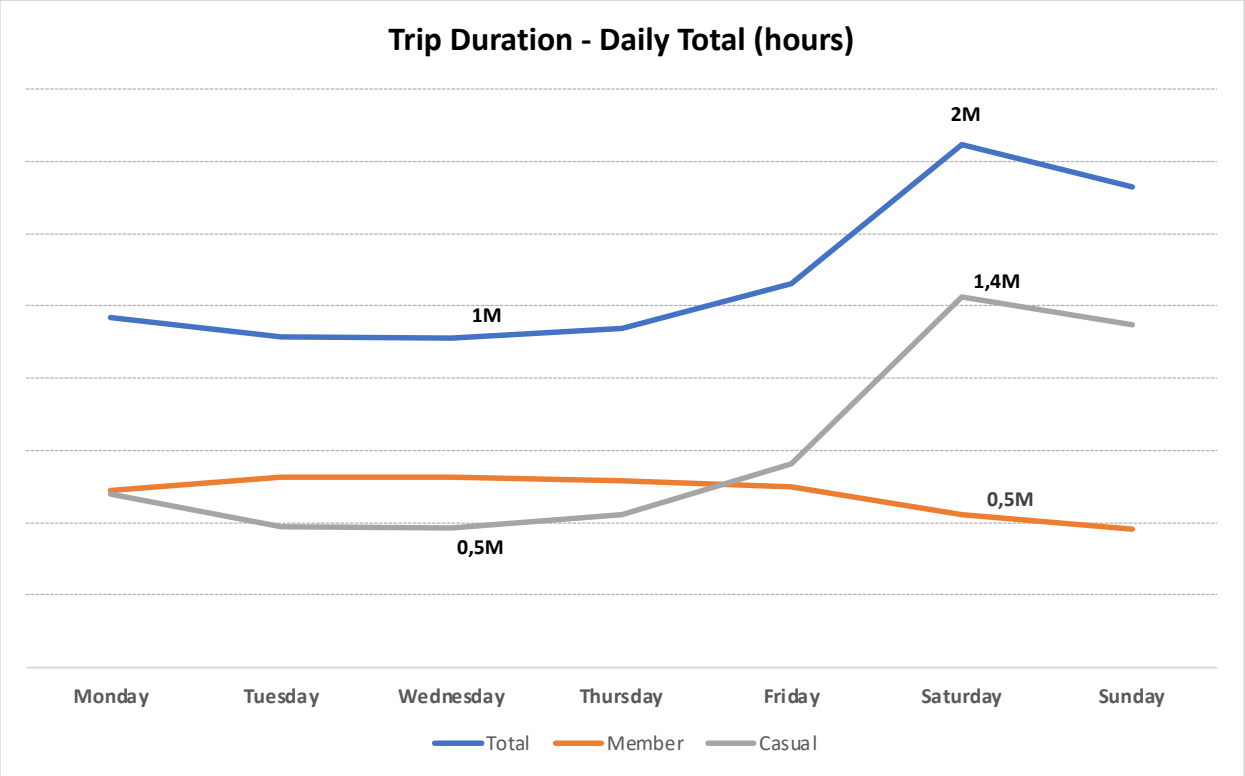
5. Supporting Visualizations and Key Findings



The visual above demonstrates the yearly total trip durations for different customer groups. Until Covid-19, both customer groups were contributing similarly but after 2019, trips taken by members decreased. Hence, total trip duration stayed stable because the increase in casual riders during the period compensated for the decrease in members.

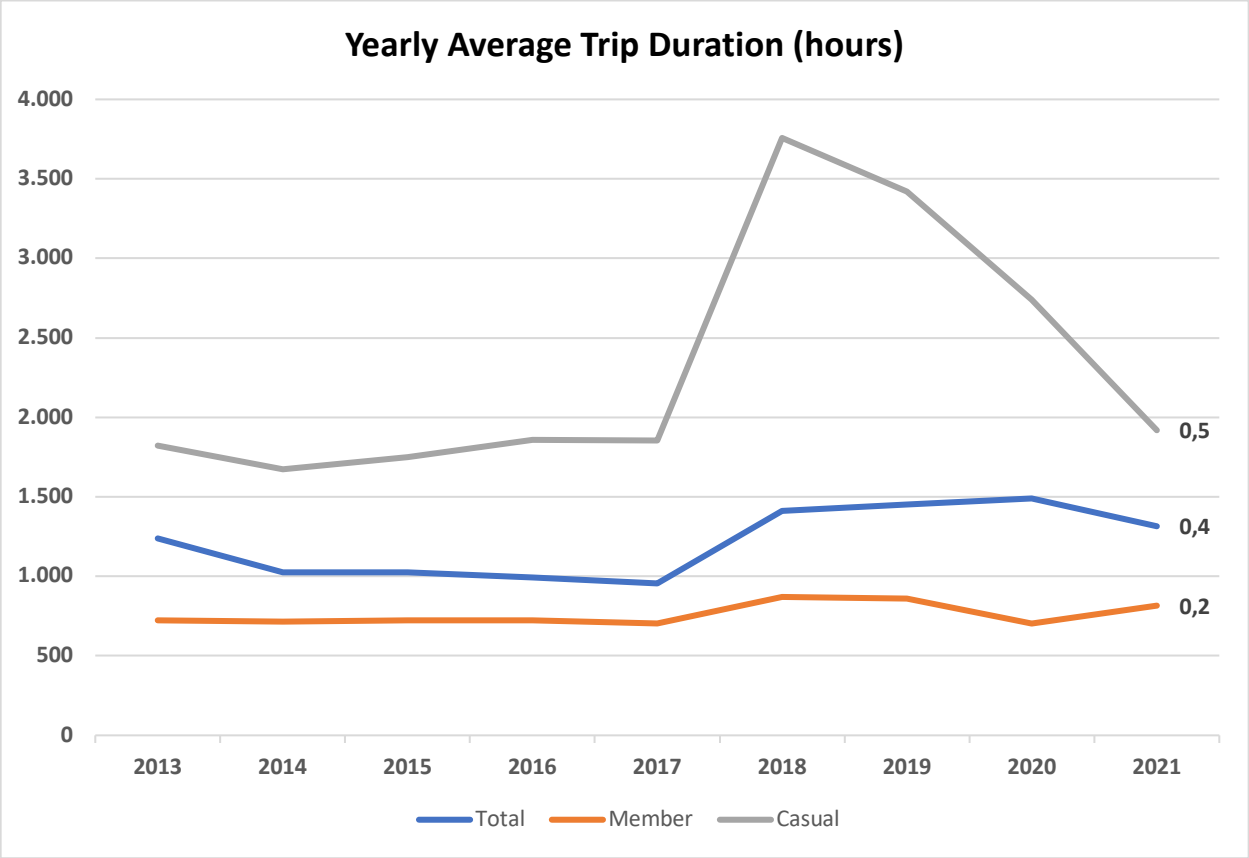


Summer has the highest demand for bikes for both customer groups. However, casual riders have a special interest during the summer.

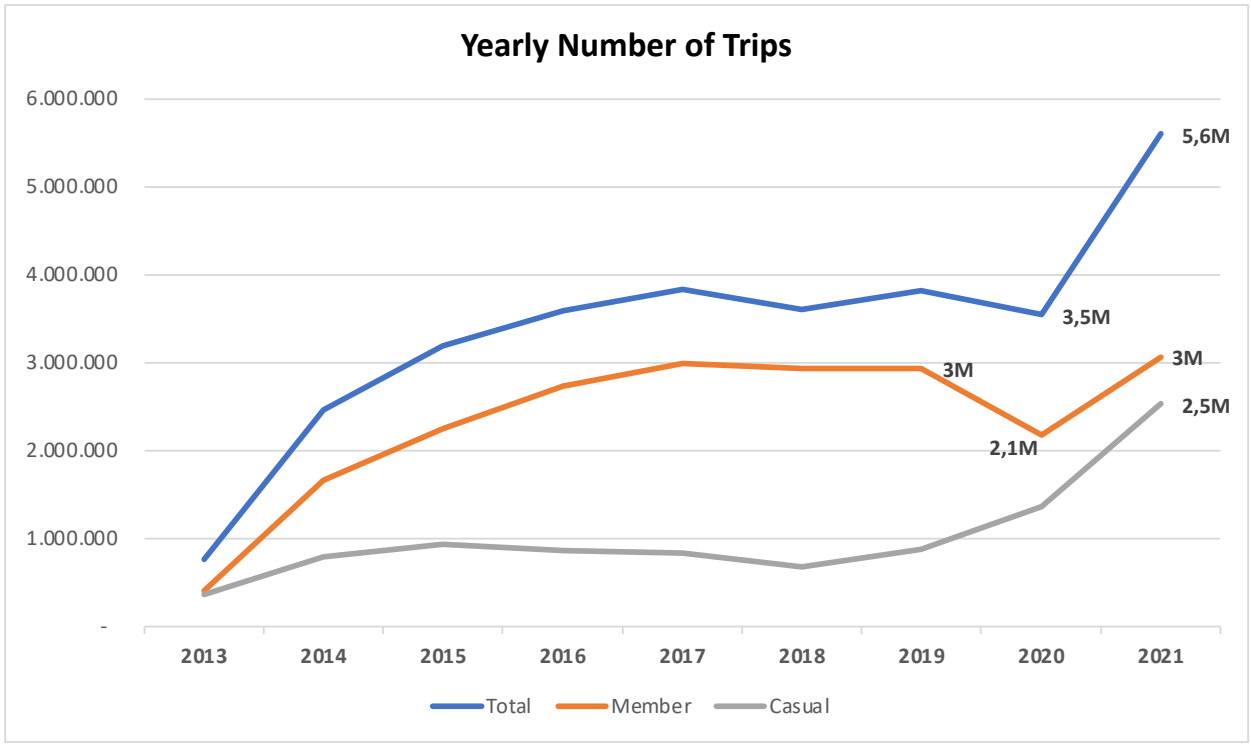


Members are using the bikes with a similar frequency during the week, however, casual riders are preferring the weekends. In total, weekends are the most used time of the week.

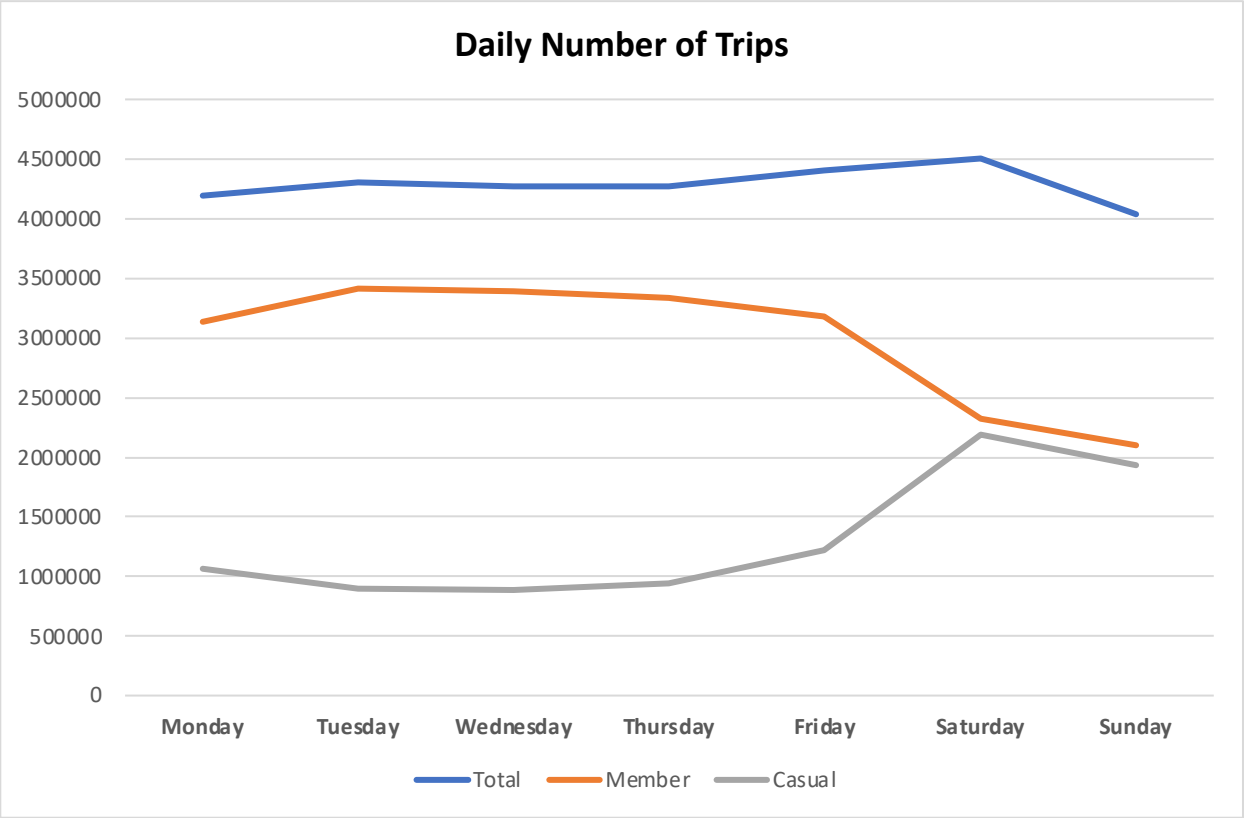
From the visuals, it seems that casual riders are mostly using bikes for leisure while members are utilizing bikes as a means of transport.



Casual riders spent 2.3 times more time with the bikes on average in 2021 which means 2.3 times more revenue per ride for the company.



This visual is showing that more than half of the customers of Cyclistic are members. The only exception is the Covid-19 period but the number of members is still higher than casual riders.



Similarly, the daily number of different customer groups shows that the number of members is always higher than the casual riders. Interestingly, the number of members during weekends is lower than on weekdays but it is vice versa for the casual riders.

To sum up, two customer groups were contributing the revenue equally until 2019. Since then, the revenue contribution of the casual riders are higher than members even though the number of rides for casual riders is lower than for the members. This is coming from the high average trip duration of casual riders.

6. Recommendations

Since the total trip duration is what matters in terms of revenue, for the last 3 years casual riders are accounting for ~60% of the total revenue. Increasing the number of members will have little or no effect on the revenue since the total trip duration will not be affected by the type of customers. However, members are showing stable use which means steady cash flows.

1. Summer has the highest demand, prepare campaigns for summer focusing on both customer groups in order to increase the number of total users instead of only focusing on the number of members. Because this attempt may cause a decrease in the number of casual riders which has a higher marginal effect on the revenue.
2. Prepare campaigns for weekends since the demand is at the top on weekends.
3. Casual riders have a higher average trip duration, therefore, increasing the number of casual riders will have a more significant marginal effect than the members. Preparing special campaigns targeting casual riders will be more profitable for the company.