Case Study

Report: Analysis of Cyclistic Bike-share Data

By: Breanyl Baker

07/29/2023

**Business Task**

Conduct a comprehensive analysis to identify the difference between casual riders and annual members' usage of Cyclistc bikes to present findings to the stakeholders.

**Data Sources**

The data used was sourced from Cyclistic’s historical trip data. The data sourced consists of the past 12 months of cyclists' trip data datasets. It stores information on the bike rides taken by both the annual members and the casual riders. Details include the start and end times and locations, ride ID, type of bike, station locations, membership type, latitude, and longitude.

**Cleaning and Manipulation**

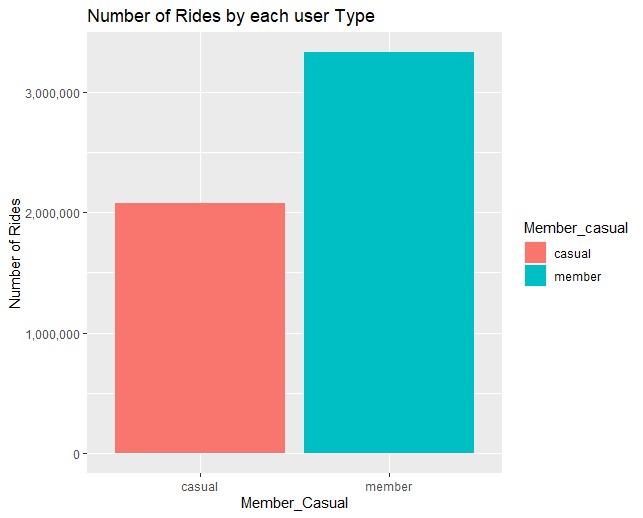
* In Excel
  + Created a column called ride\_length by subtracting the started\_at column from the ended\_at column. Then formatted the column as time 37:30:55.
  + Created a column called day\_of\_week using the weekday function. They are represented as numbers where 1- Sunday 7-Saturday.
  + Changed the format from general to text for the start and end Id.
* In R
  + I checked that the column names were consistent across all datasets.
  + I merged the 12 datasets in R because it was too large to do in Excel.
  + After which I removed the columns start\_station\_name, start\_station\_id, end\_station\_name, end\_station\_id, start\_lat, start\_lng, end\_lat, end\_lng.
  + Created a new column called ride\_length\_seconds because the mean could not be calculated in HH:MM: SS format.
  + Created a column called Season to group the month in their respective season extracting the month from the started\_at column.
  + Created a column called weekday\_vs\_weekend to group the day in their respective place using the day\_of\_week column.

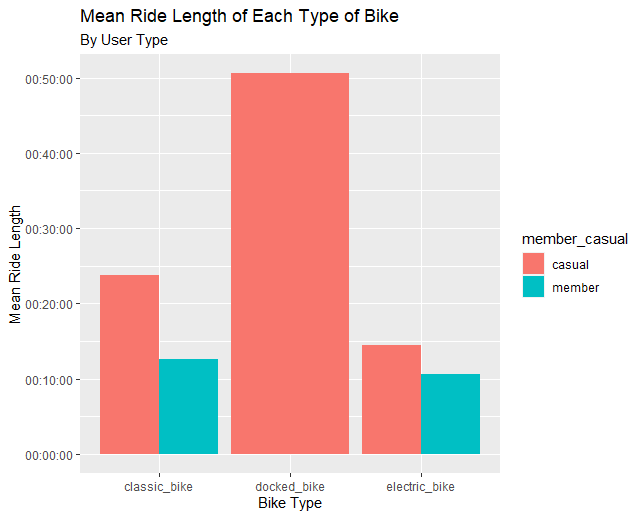
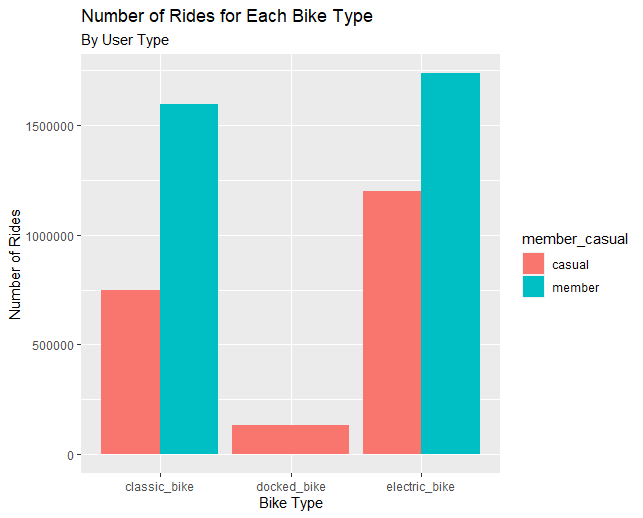
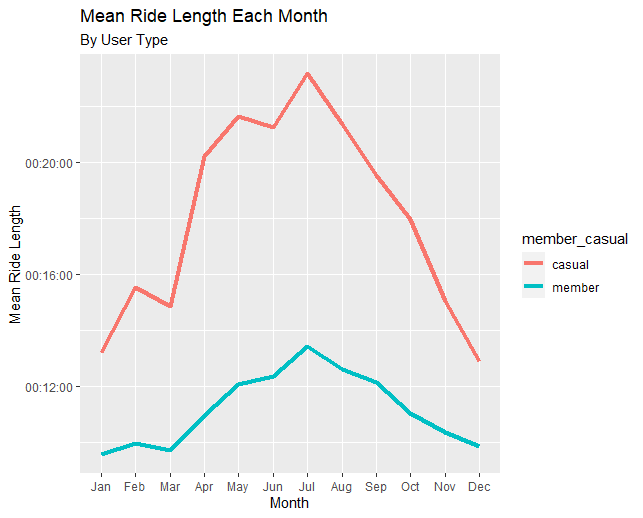
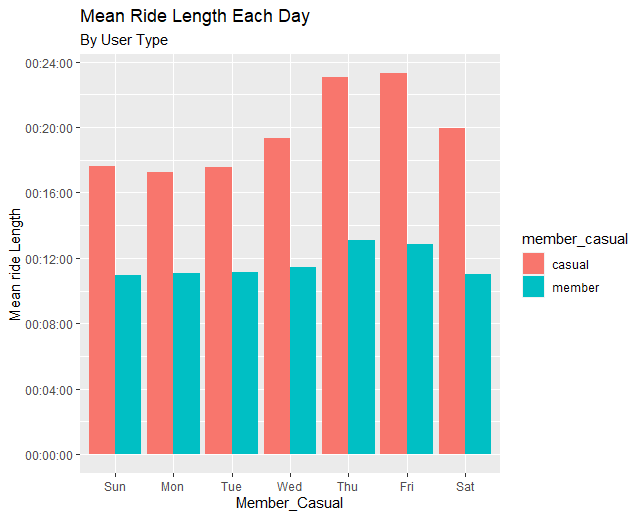
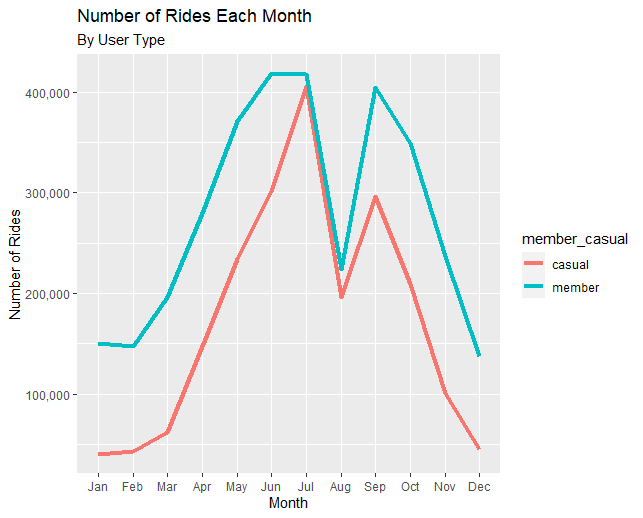
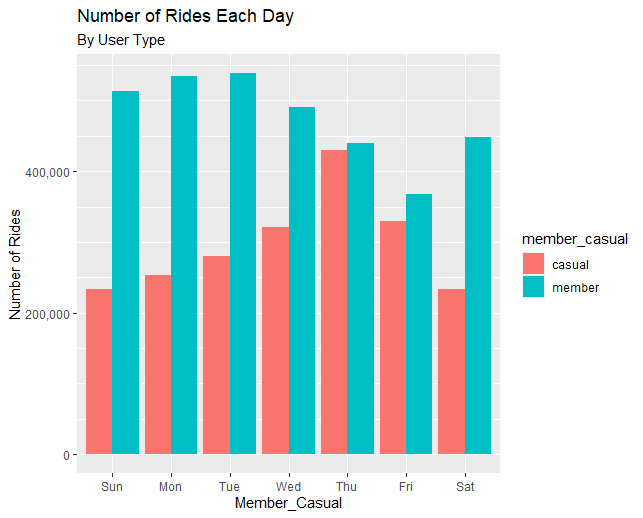
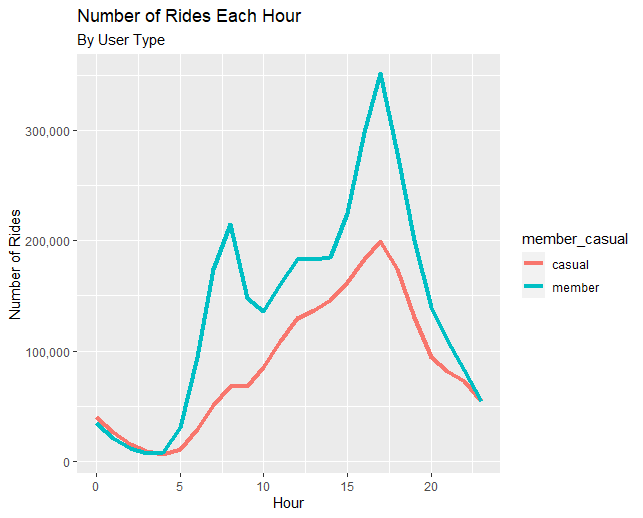
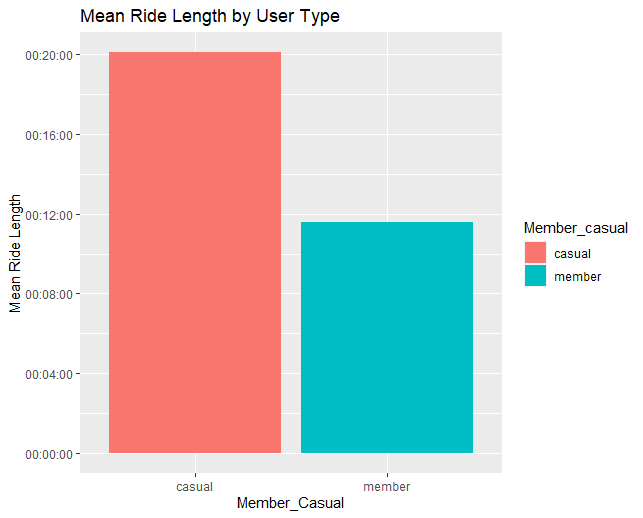
**Summary of Analyze**

* Members took more rides than casual users.
* Members took shorter rides while casual users tend to take longer rides.
* Members took more rides than casual users each hour of the day, day of the week, and month.
* Member took more rides than causal users each hour except from 12 am to 4 am.
* Casual users had longer mean ride lengths for each day of the week and month.
* Friday and Thursday had the longest mean ride lengths for casual users and members respectively.
* Monday and Sunday had the shortest mean length for casual users and members respectively.
* Casual and members both have the longest mean ride length in July but casual had the shortest mean ride in December while members had in January.
* Casual users used all bicycle types, with electric bikes being the most used and docked bikes the least used. Member users did not use docked bikes, with electric bikes being the most used and classic bikes the least used.
* Weekdays had more rides than weekends.

Overall, the analysis reveals that casual and member users exhibit distinct ride behaviors, with casual users taking longer and more diverse rides, while member users take more frequent but shorter rides, primarily using electric bikes.

**Supporting Visualization**





A graph of a person and person

Description automatically generated**Recommendations**

1. Conduct promotions membership during the weekdays for casual users when they have their longest mean ride length.
2. Can offer a time-limited discount on membership fee during the Summer which has the most rides of the year.
3. For casual users, since they take longer rides, promote the cost-effectiveness of a membership. Showing that cost reduces per minute for extended rides.