Cyclistic Analysis Case Study

Drew Mistelske

Updated: 2023-08-24

Introduction

The data analysis you are about to see was completed using a six phase data analysis process: Ask, Prepare, Process, Analyze, Visualize, and Act and is a case study from Google Data Analytics Capstone Course found on coursera.org.

Ask

Company Background

In 2016, Cyclistic launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

Involved Parties

Lily Moreno: The director of marketing.

Cyclistic Executive Team: Will decide whether to approve the recommended marketing program.

Drew Mistelske: Data Analyst.

Business Goal, Task, and Statement

Goal - Convert casual riders into annual members.

Task - How do annual members and casual riders use Cyclistic bikes differently?

Statement - I will find out differences in the way casual riders (single-day riders) use bikes with Cyclistic versus Cyclistic members (annual membership riders), to help the stakeholders decide the best route to take in converting the casual riders to a Cyclistic member.

Prepare

About the Data

The data is twelve internal, first party, and structured csv files with hundreds of thousands of rows per file and thirteen columns. The files are per month starting in July 2022 and ending in June 2023. The data contains ordinal and nominal qualitative data, continuous quantitative data. The data types are string, floating point, date and time, and list. The files can be found here. The data is original, reliable, current, cited, and comprehensive.

Data License

This is public data made available from Motivate International Inc. that can be used to explore how different customer types are using Cyclistic bikes. License.

Useful Data

This data contains the correct information needed to answer the business task at hand.

There are a couple extra pieces of data I would like to have to create a more complete picture such as, user id for statistics on repeated use and price points for spending trends or patterns.

Process

Tools Used

Microsoft Excel, SQLite, and RStudio.

Cleaning

Microsoft Excel I used Excel to clean the duplicates, check for spelling errors, make sure blank/empty cells were valid, and verify the integrity of the data.

SQLite Combined all csv files into one database table, turned the ride_id column into a primary key, not null, and unique column. Ran queries to make sure the correct number of observations were present, i.e. count(distinct ride_id) and the ride_id column had the correct 16-character length in all observations. Finally exported the table into one csv file.

RStudio Verified the data transferred correctly by comparing the total number of rows equaled each previous change to the files. Added six new columns for better analysis.

Analyze

Trends and Relationships

Casual

- Overall casual riders have an up and down trend with everything being up in summer and then down in winter.
- Weekly the casual riders have high numbers on the weekends and lower numbers on weekdays.
- The average duration of classic bike rides is triple that of the members' average duration.
- Classic bikes are preferred over electric when it comes to longer rides but electric is chosen more often.
- Casual riders account for 69.1 percent of the total duration of rides.

Members

- Overall member riders have an up (in summer) and down (in winter) trend, same as casual riders though members have a less extreme variation between the highs and lows.
- The weekdays have a higher number of rides than the weekends for the members.
- The member riders have a small spike on the weekends for the duration of their rides.
- Like casual classic bikes are used for the longest rides but electric is used more often.
- Members riders account for 61.2 percent of the number of rides.

Summary

There are a lot of similarities between the casual and member riders with patterns, trends, and relationships. There are three significant differences between the two groups. One is the duration of the rides. Causal riders have for the entire year on average slightly more than double the duration than the member riders. Two, is the number of rides during the weekdays is almost double for the member riders than the casual riders for the year. Three, year-round on the weekdays, members have a spike in usage around 7 A.M. - 8 A.M. (07:00 - 08:00) and from 4 P.M. - 6 P.M (16:00 - 18:00) suggesting they use the bikes for transportation to and from work. Casual riders have an uptrend till 5 P.M. (17:00) and then a downtrend till 4 A.M. (04:00), it should be noted that there is a small spike in casual usage on Monday - Thursday at 8 A.M. (08:00) and could be a small potential to pull those casual riders to memberships.

Visualize

Key Visualizations from the Analysis

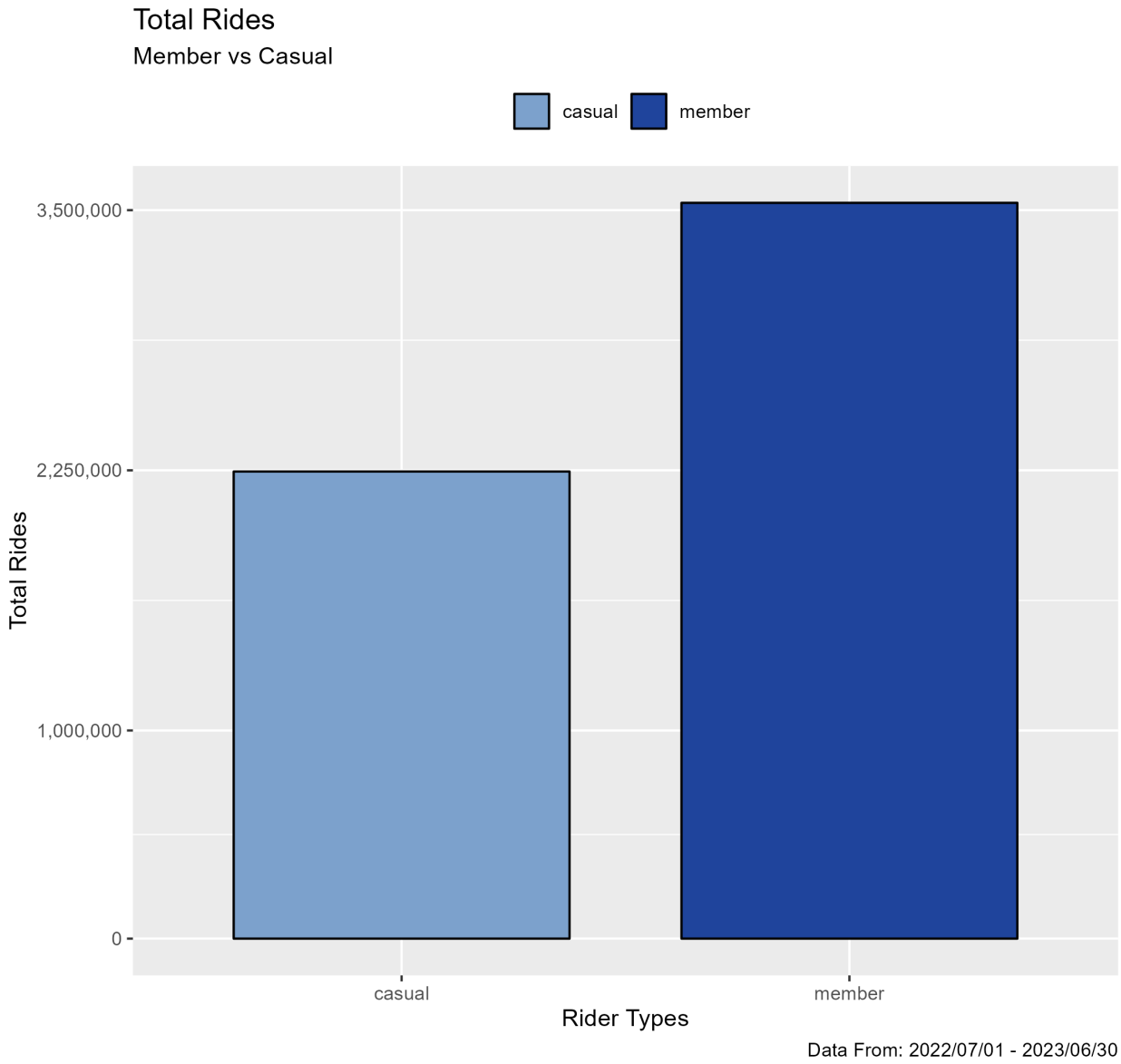

Figure 1: *Graph representing the difference between casual and member riders total rides.*

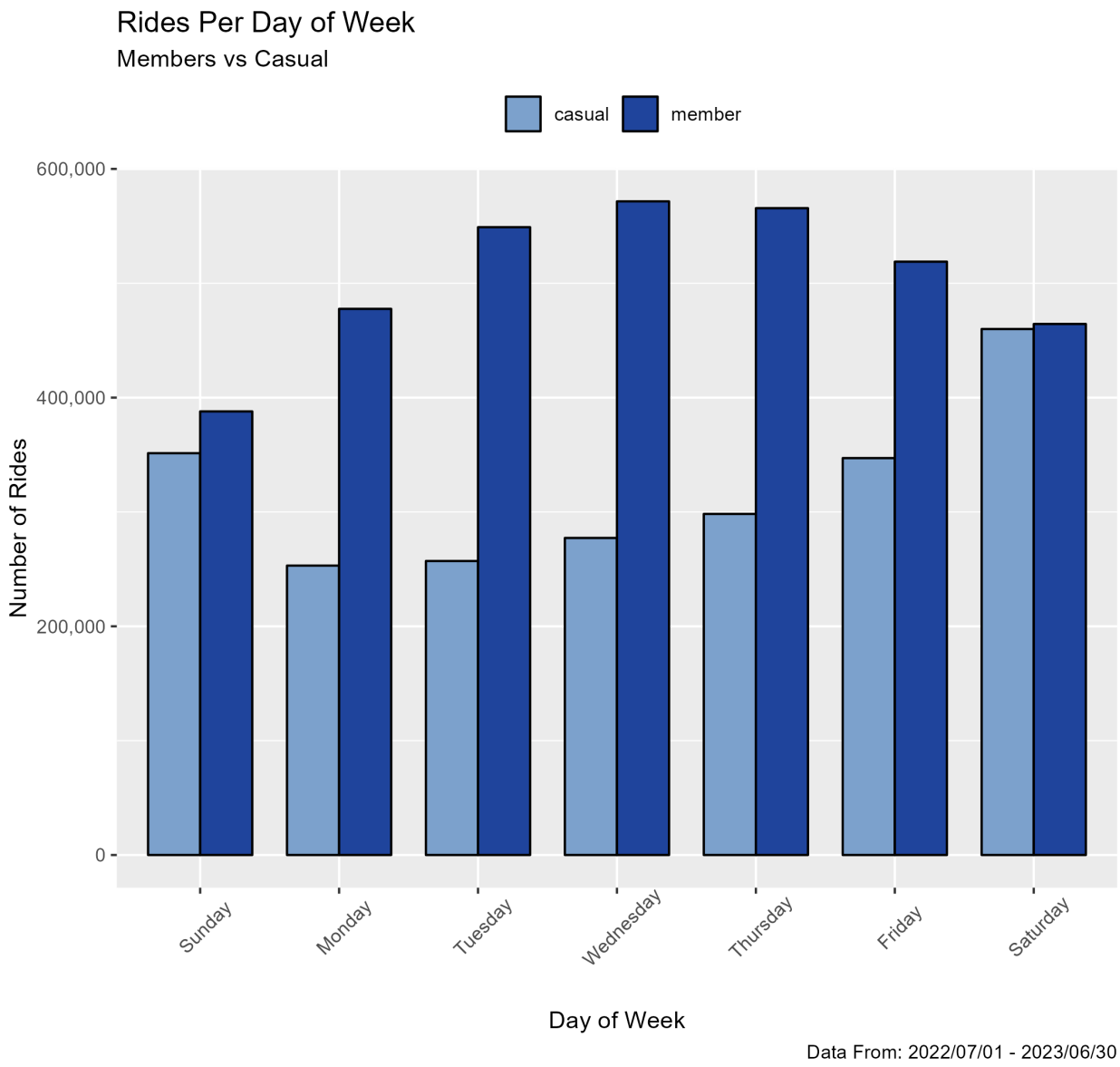

Figure 2: *Graph representing the difference between casual and member riders ride counts per day of week.*

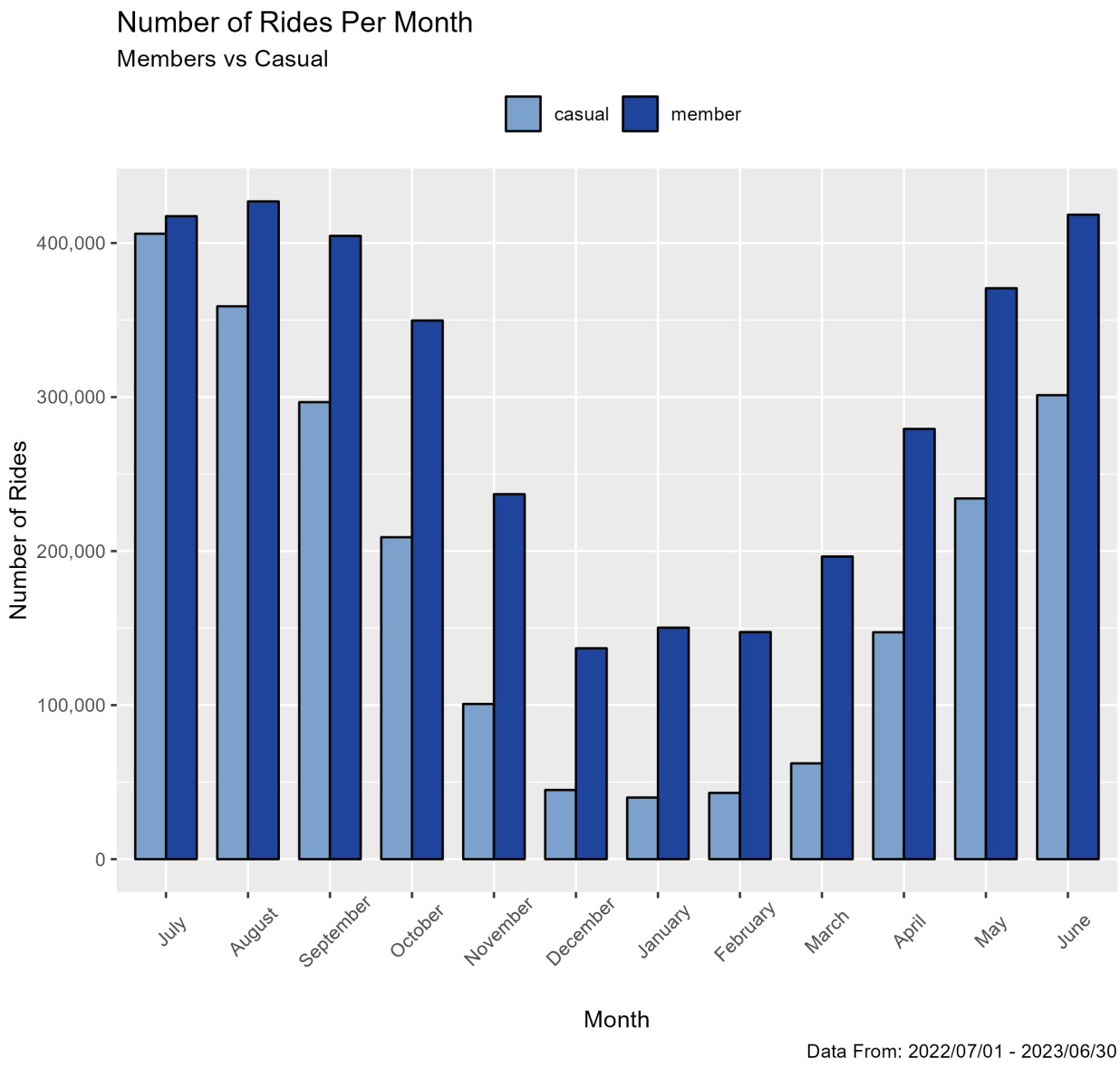

Figure 3: *Graph representing the difference between casual and member riders ride counts per month.*

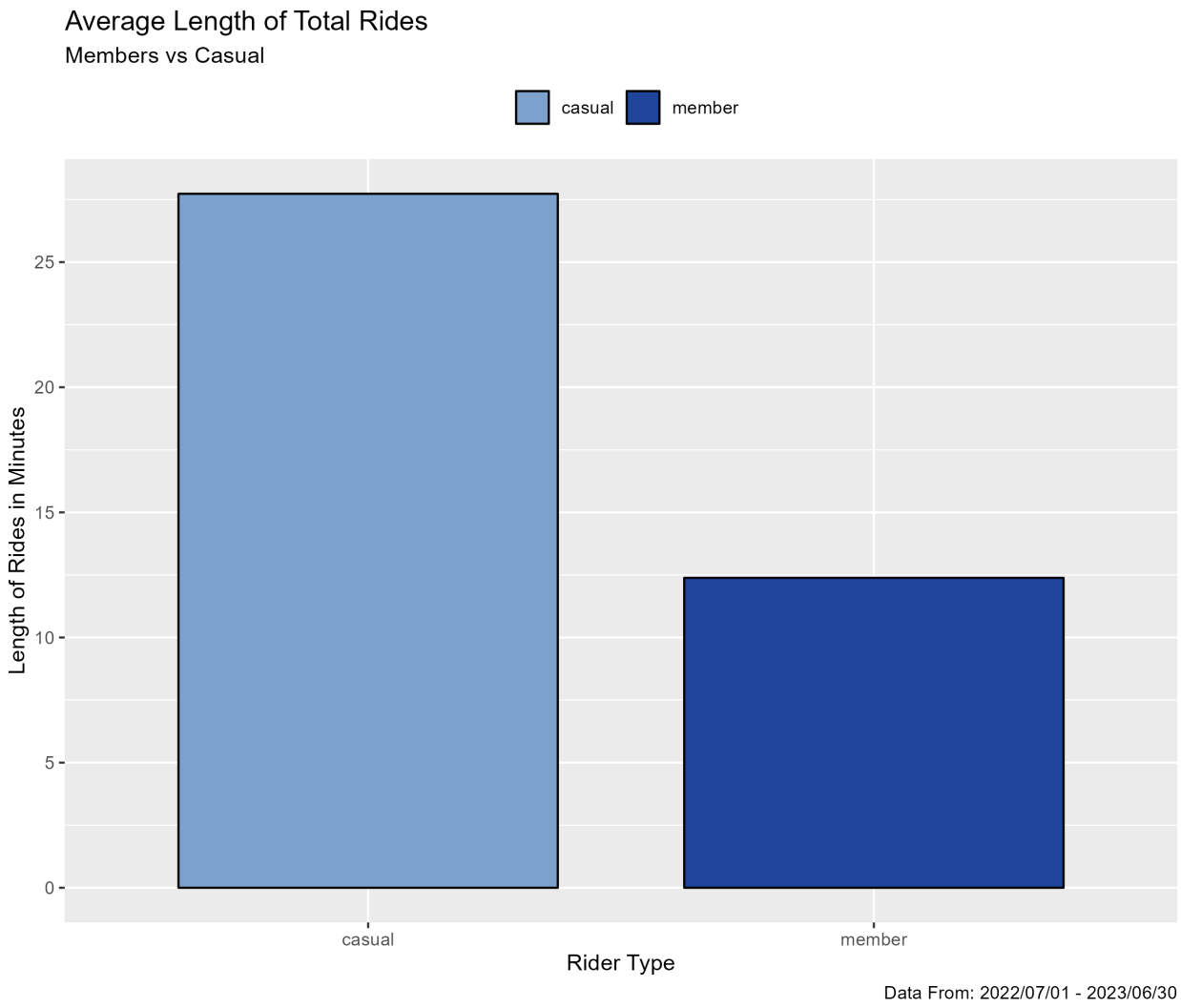

Figure 4: *Graph representing the difference between casual and member riders average length from the total number of rides.*

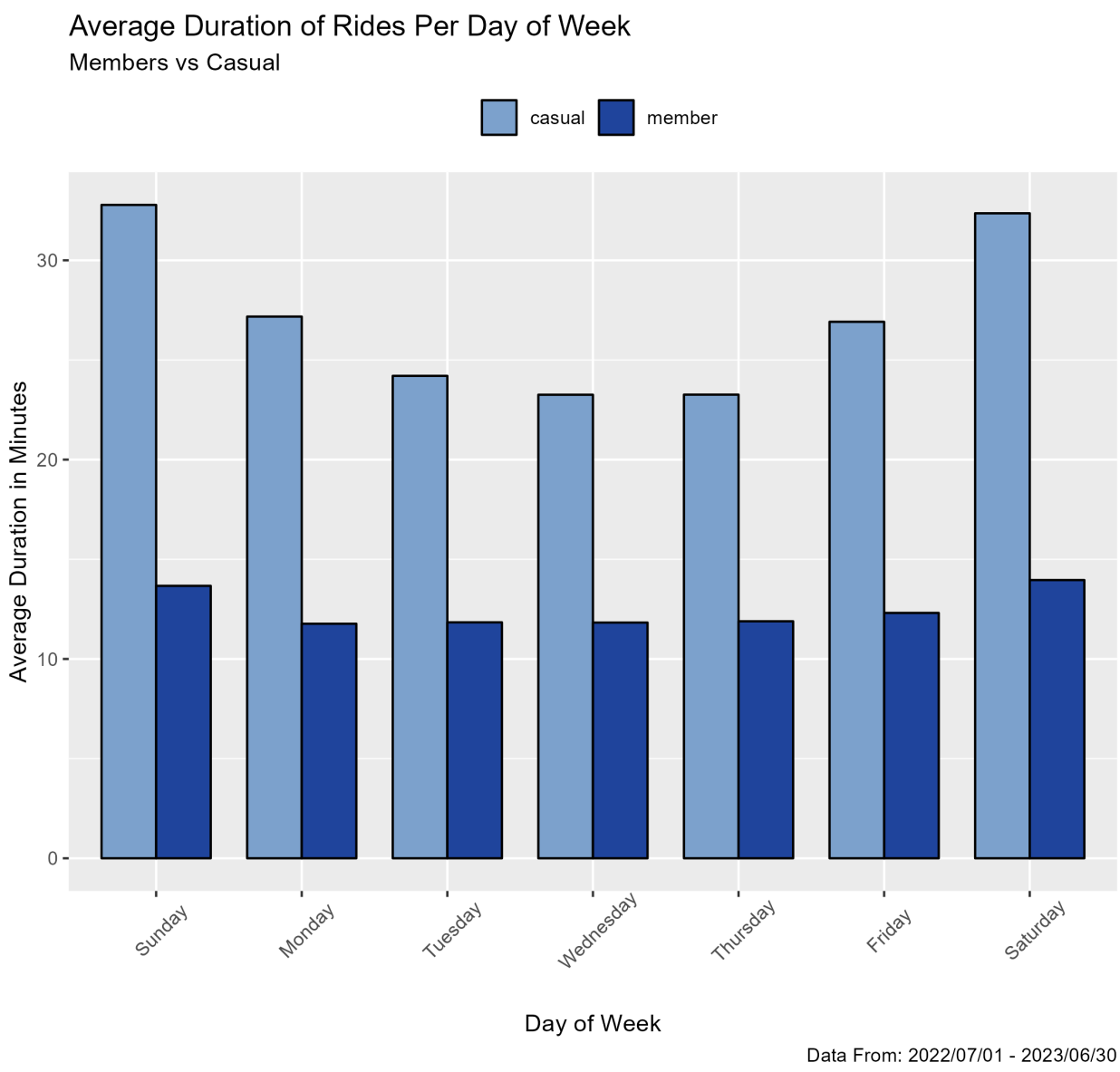

Figure 5: *Graph representing the difference between casual and member riders average length per day.*

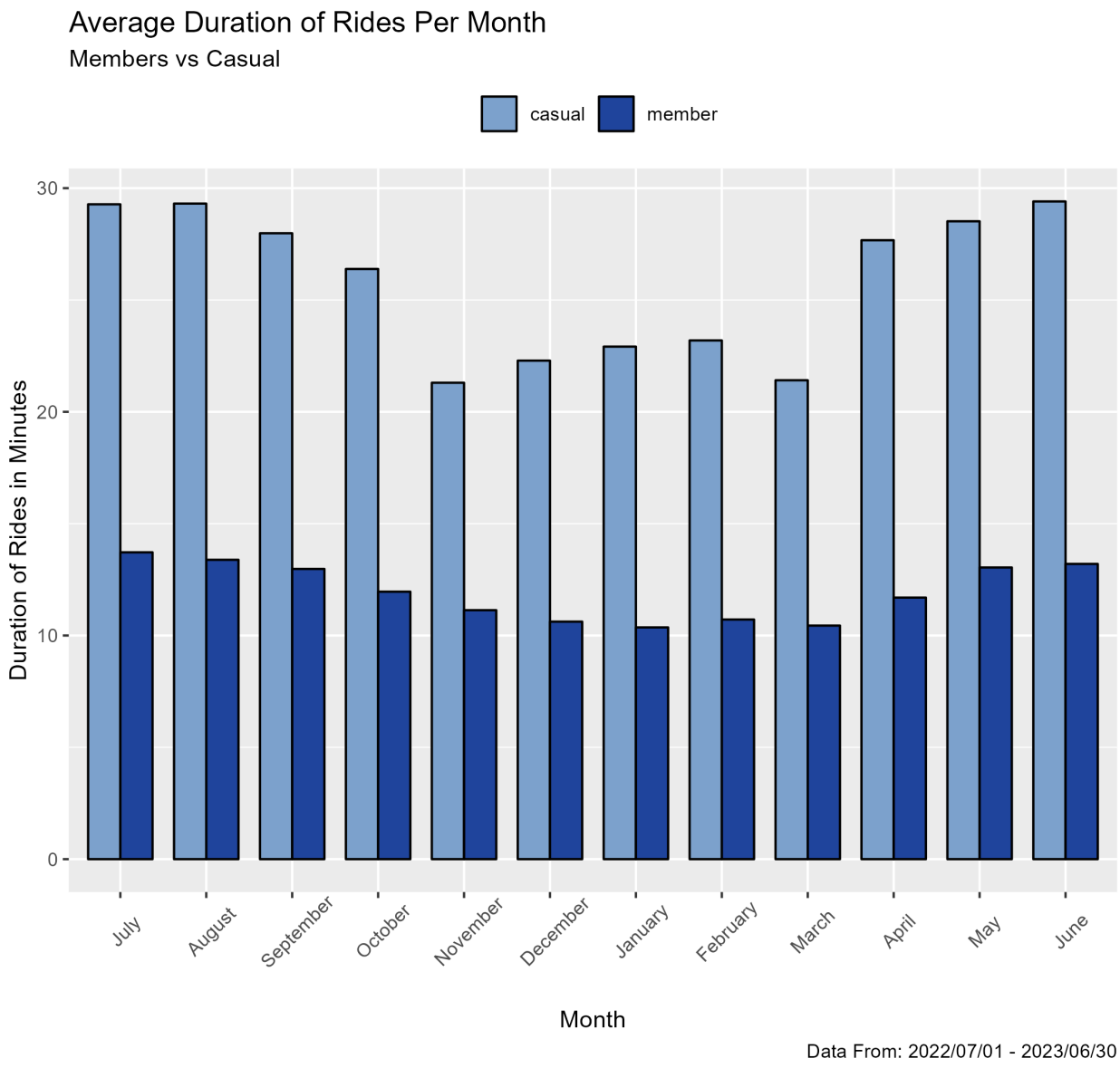

Figure 6: *Graph representing the difference between casual and member riders average length per month.*

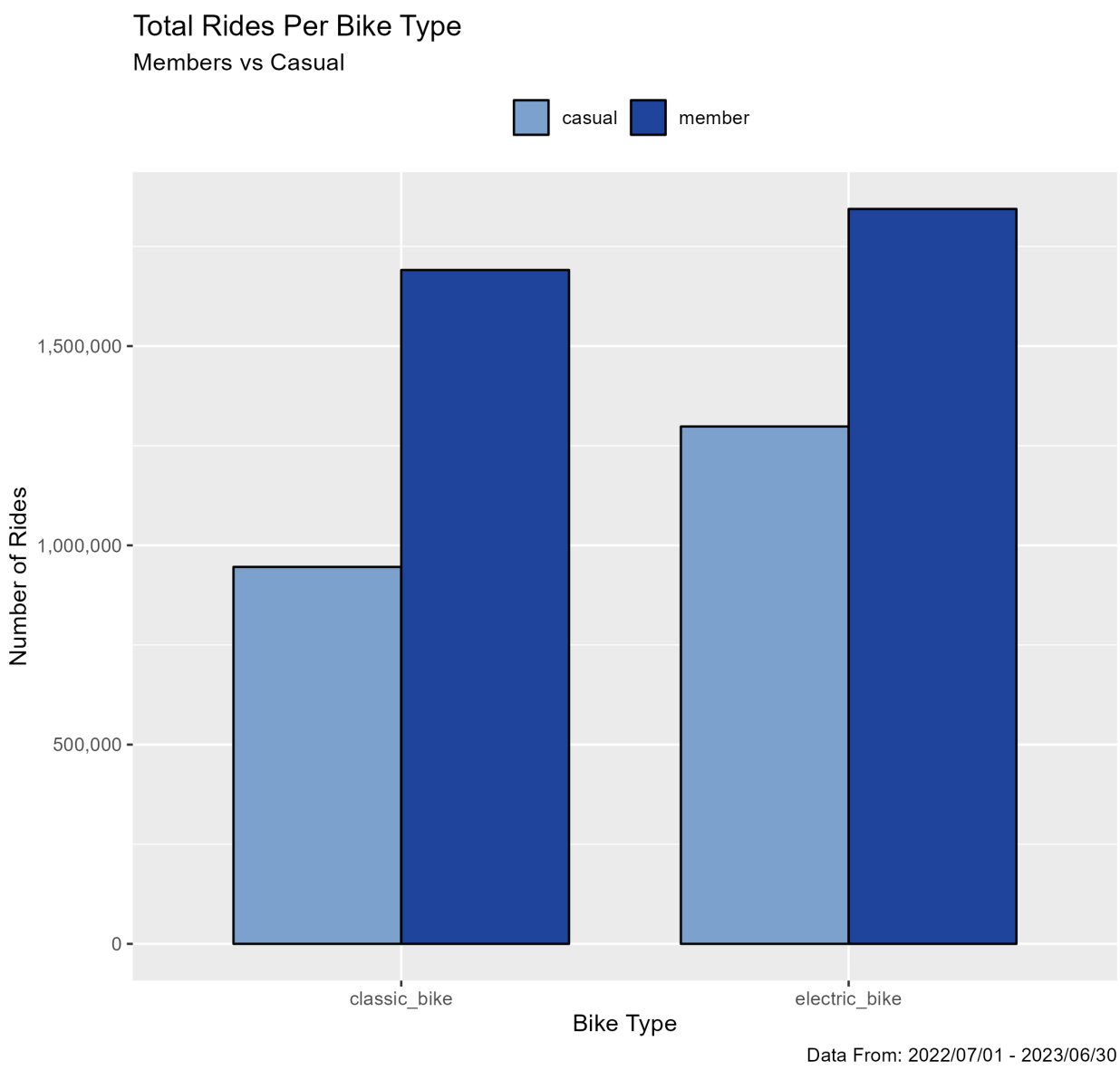

Figure 7: *Graph representing the difference between casual and member riders total rides per bike type*

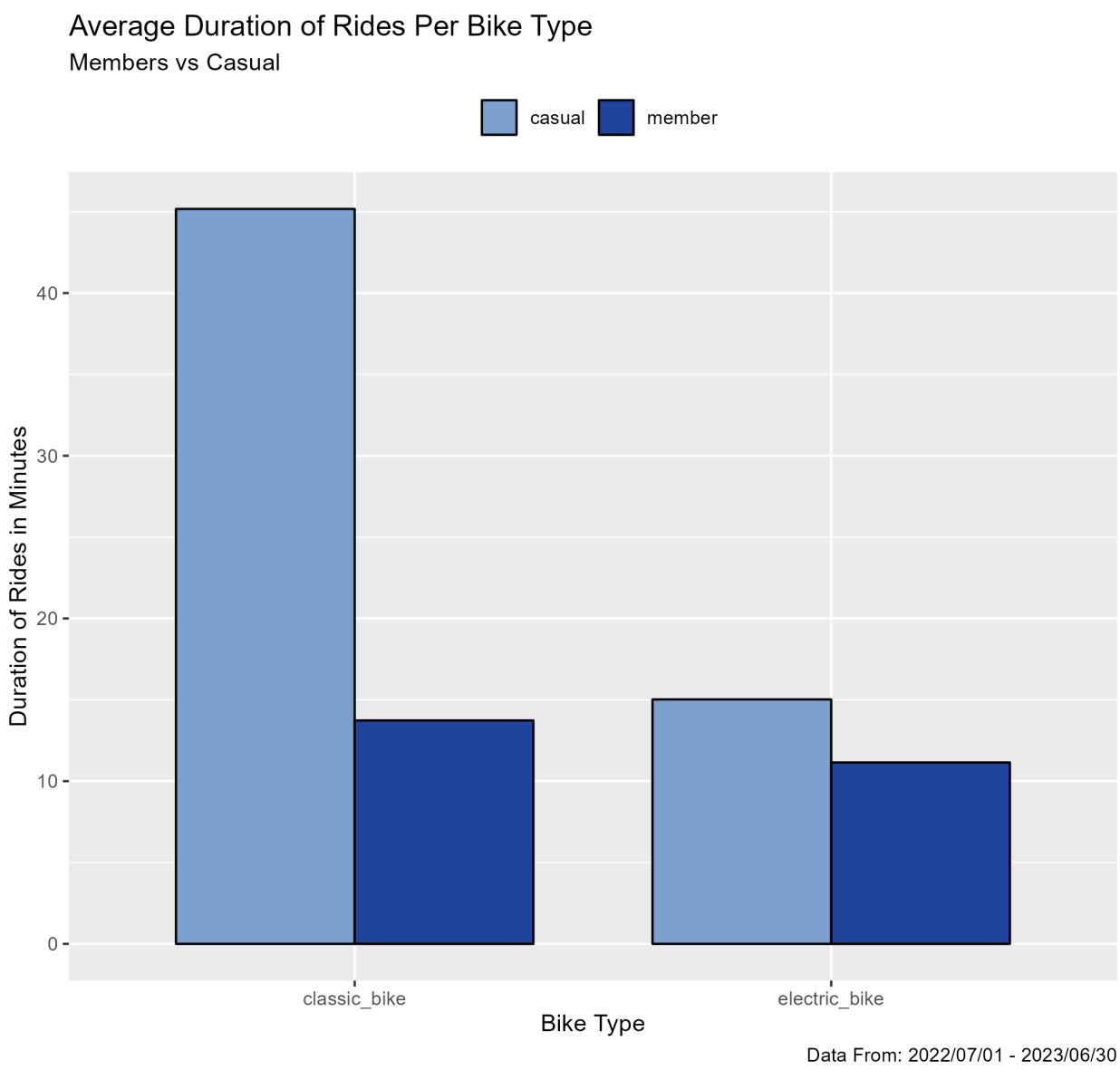

Figure 8: *Graph representing the difference between casual and member riders average ride length per bike type.*

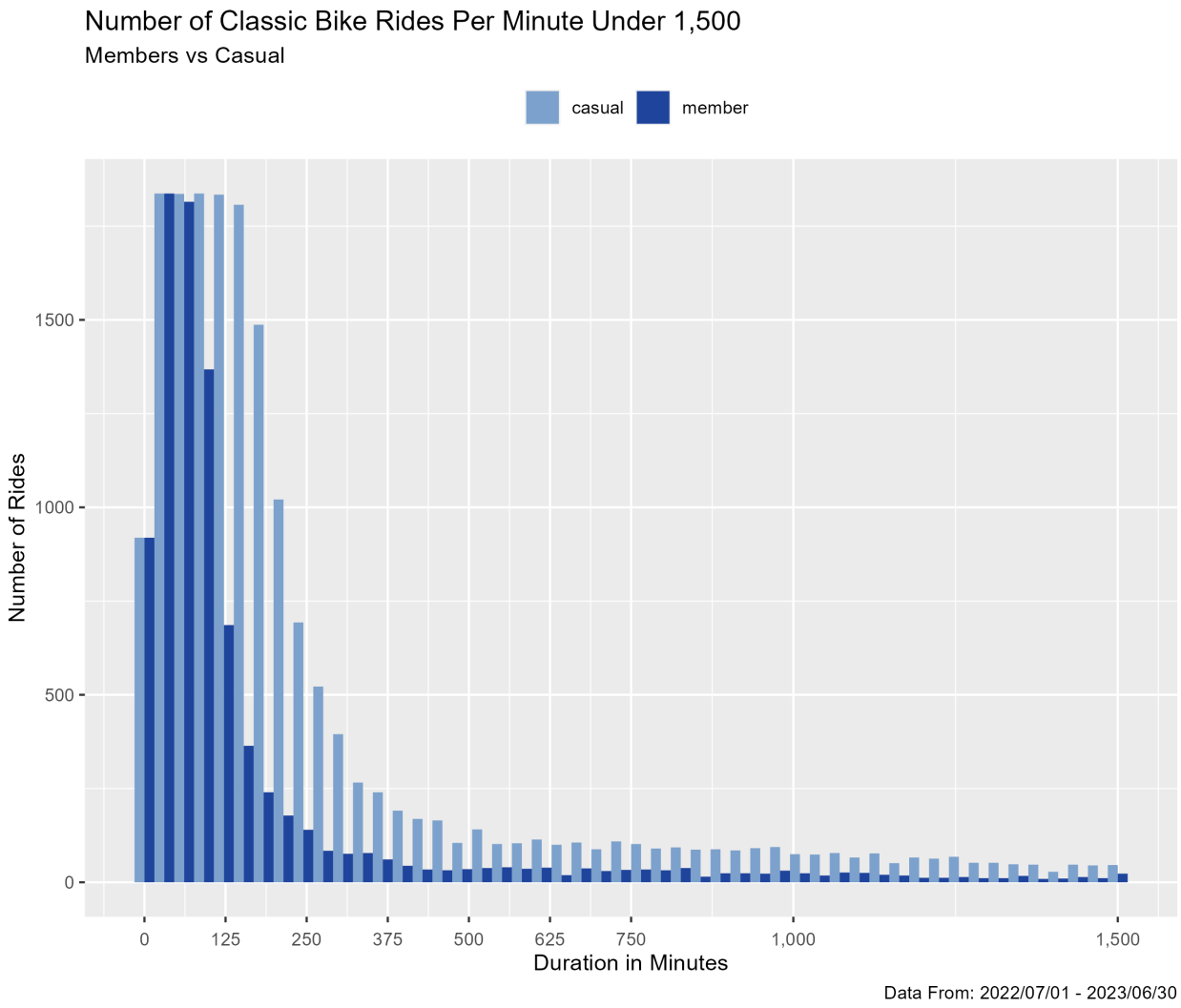

Figure 9: *Graph representing the difference between casual and member riders total number of rides per minute up till 1,500 minutes for the classic bike type.*

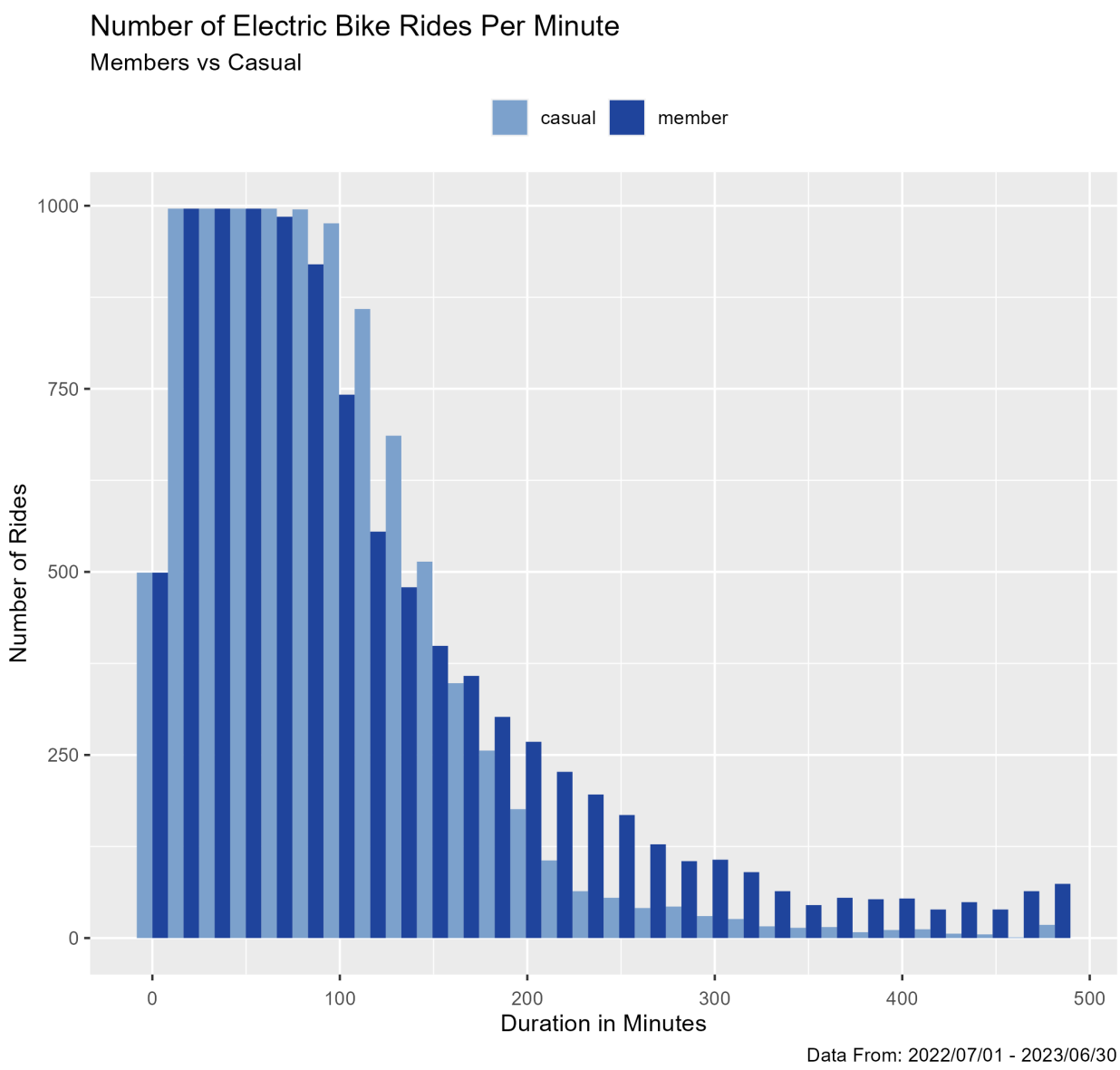

Figure 10: *Graph representing the difference between casual and member riders total number of rides per minute for the electric bike type.*

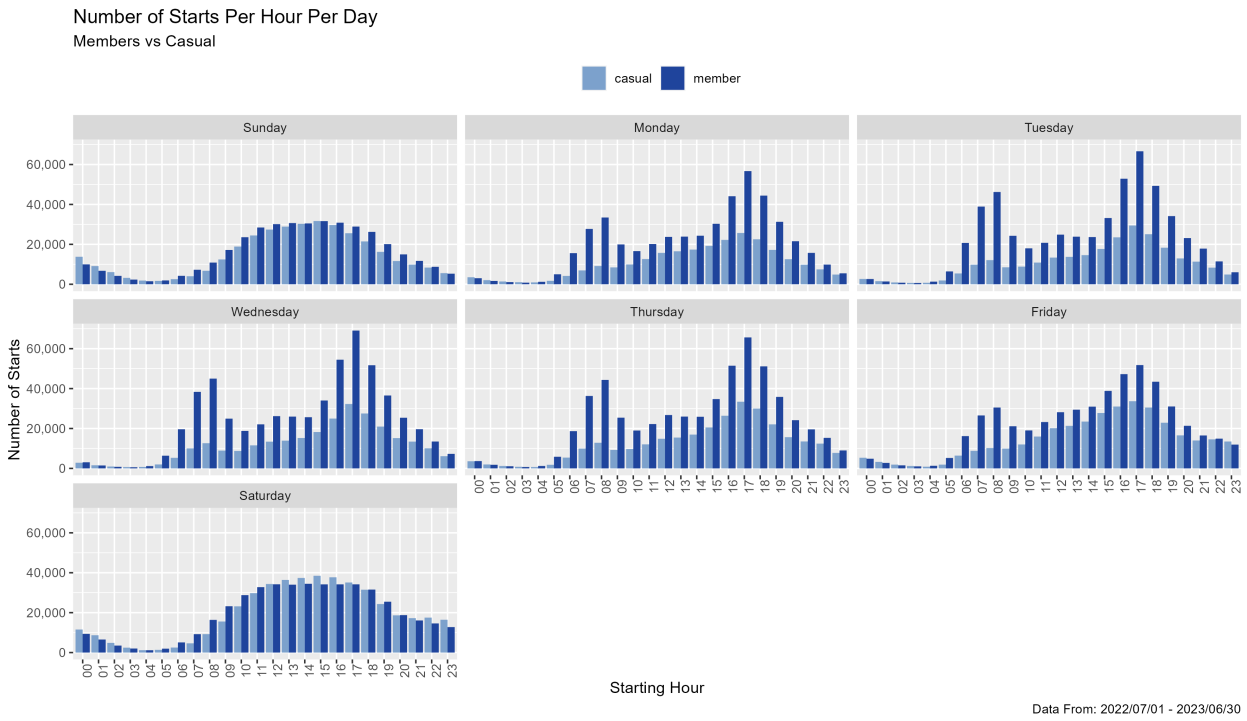

Figure 11: *Graph representing the difference between casual and member riders total number of rides per hour for each day of the week.*

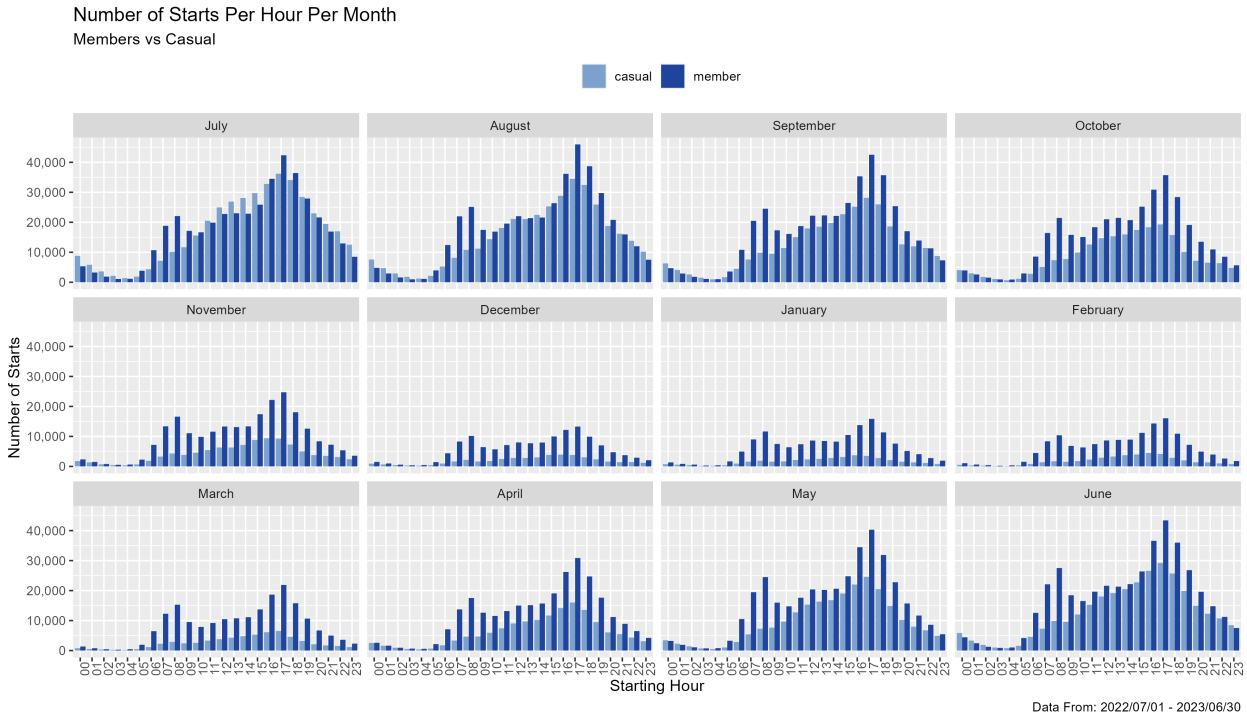

Figure 12: *Graph representing the difference between casual and member riders total number of rides per hour for each month.*

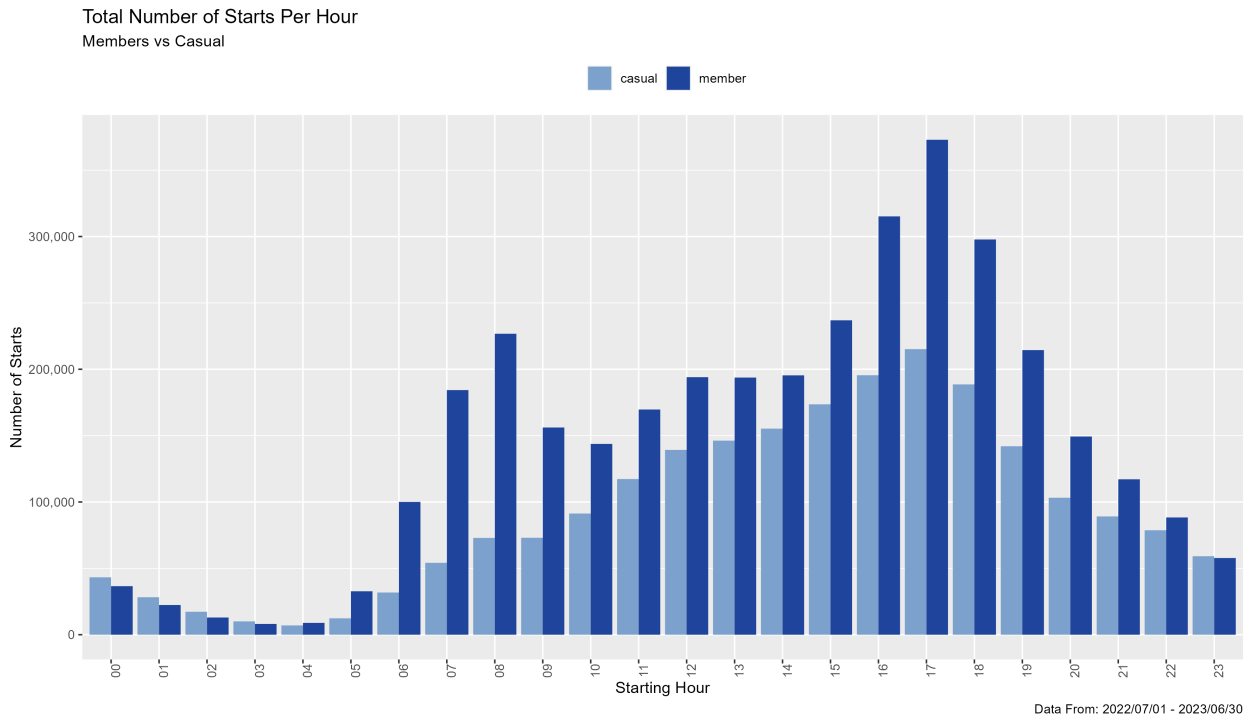

Figure 13: *Graph representing the difference between casual and member riders total number of rides per hour.*

Act

From the data that has been shown these are my suggestions for Lily Moreno to convert casual riders into member riders.

1. My first suggestion will target the higher duration rides of the casual riders, specifically those rides with the classic bike. Adding to the membership that after the two-hour mark (when the trend of members duration of their ride on classic bikes starts to trend down) the price per min drops as a member or the price for a member to rent a bike for the entire day is a set rate.

2. This suggestion is based on the times casual riders use the service. Add a weekday deal for those who only ride during business hours and/or a weekend deal starting Friday night and ending Sunday evening for those casual riders who use the bikes for long nights on the weekends. Also, we could add a summer seasonal deal.

3. For my final suggestion, I want to take suggestion number one and two and put them together and create an entirely different membership and/or add-on membership designed with the casual riders in mind. If made as an add-on this could also benefit member riders, bringing more revenue to the membership package.

Extra Documentation

Extra documentation for all the processes along with the code files please click this link