

Exploring Member vs. Non-Member Bike Riding Trends

Background

To better understand the difference between riding trends at Divvy, a Chicago-based bike-share company, I studied the last 12 months of data to compare member and non-member (casual) riding behavior. Understanding the differences between the two rider types could help the company make better decisions regarding bike inventory, pricing, membership incentives and packages.

The original dataset was organized into monthly CSV files, and for this study I reviewed the information from 05/2022-04/2023. The data from the last 12 months contains thirteen variables and 5,859,061 observations. I used SQL, BigQuery, and Google Sheets to parse and understand the information. The data is made available under their [limited license agreement](#).

Notable Findings

1. While casual riders ride mostly on the weekends, members ride mostly during the week
2. Member rides peak at 8 AM and 5 PM, while casual rides steadily increase until 5 PM, and then decrease
3. Casual rides are more than twice as long as member rides
4. About 49.97% of member rides and 59.26% of casual rides are on an electric bike
5. Casual riders were riding almost the same number of rides as members in July, and about one quarter the amount in January
6. Members and casual riders are picking up their bikes at the docking station at almost the same rate.
7. Casual riders have a slightly higher rate of not returning their bikes to the docking station.

Sanitizing

First, I imported the data into 12 separate database tables and then combined those tables into a single table using the following query:

```
SELECT *
FROM `bike-share-385304.Bikeshare.04-2023`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.03-2023`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.02-2023`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.01-2023`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.12-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.11-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.10-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.09-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.08-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.07-2022`
UNION ALL
```

```
SELECT *
FROM `bike-share-385304.Bikeshare.06-2022`
UNION ALL
SELECT *
FROM `bike-share-385304.Bikeshare.05-2022`
```

Then I confirmed that there weren't any months left out:

```
SELECT EXTRACT (Month FROM started_at) AS month, EXTRACT(year FROM started_at) AS
year, COUNT (*) AS total
FROM `bike-share-385304.Bikeshare.052023-042022`
GROUP BY year, month
ORDER BY year, month
```

Next, I checked for any missing data:

```
SELECT DISTINCT member_casual
FROM `bike-share-385304.Bikeshare.052023-042022`

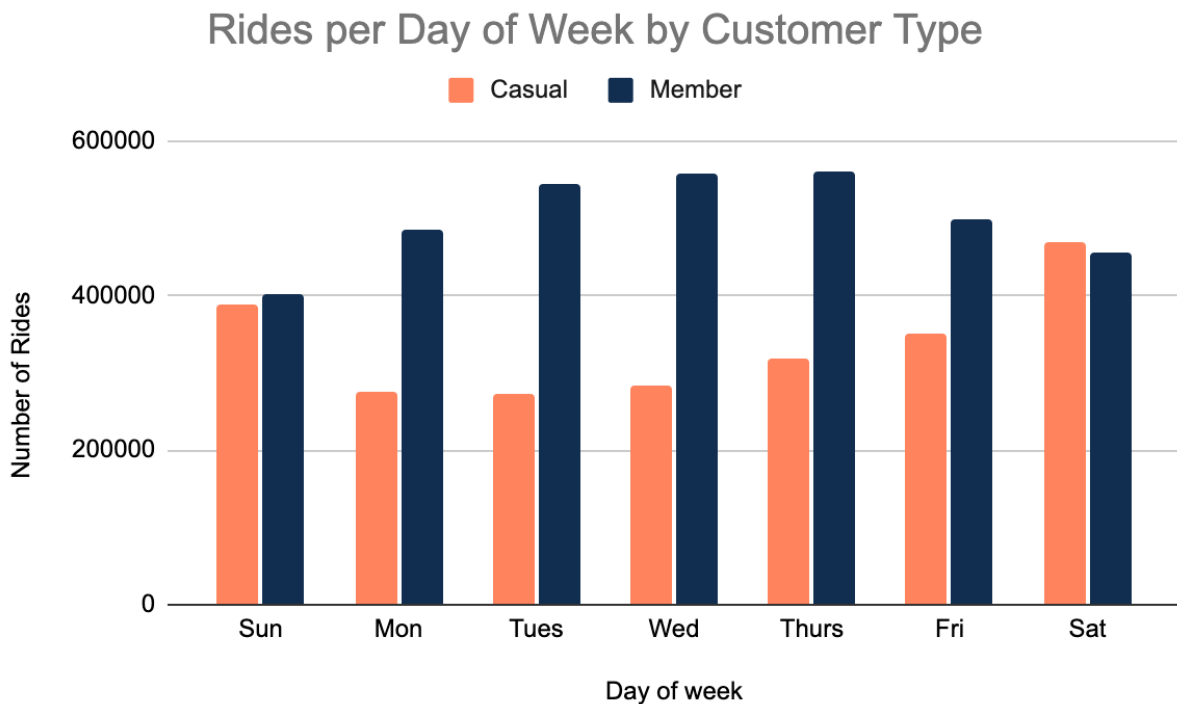
SELECT DISTINCT rideable_type
FROM `bike-share-385304.Bikeshare.052023-042022`
```

Analysis

Rides per Day of Week by Customer Type

After compiling and validating all the data, I started my analysis. First, I wanted to find riding trends per day of week based on customer type. Here is the query that I ran:

```
SELECT member_casual, EXTRACT (dayofweek FROM started_at) AS day_of_week, COUNT (*) AS  
number_of_rides  
FROM `bike-share-385304.Bikeshare.052023-042022`  
GROUP BY member_casual, day_of_week  
ORDER BY member_casual, day_of_week
```

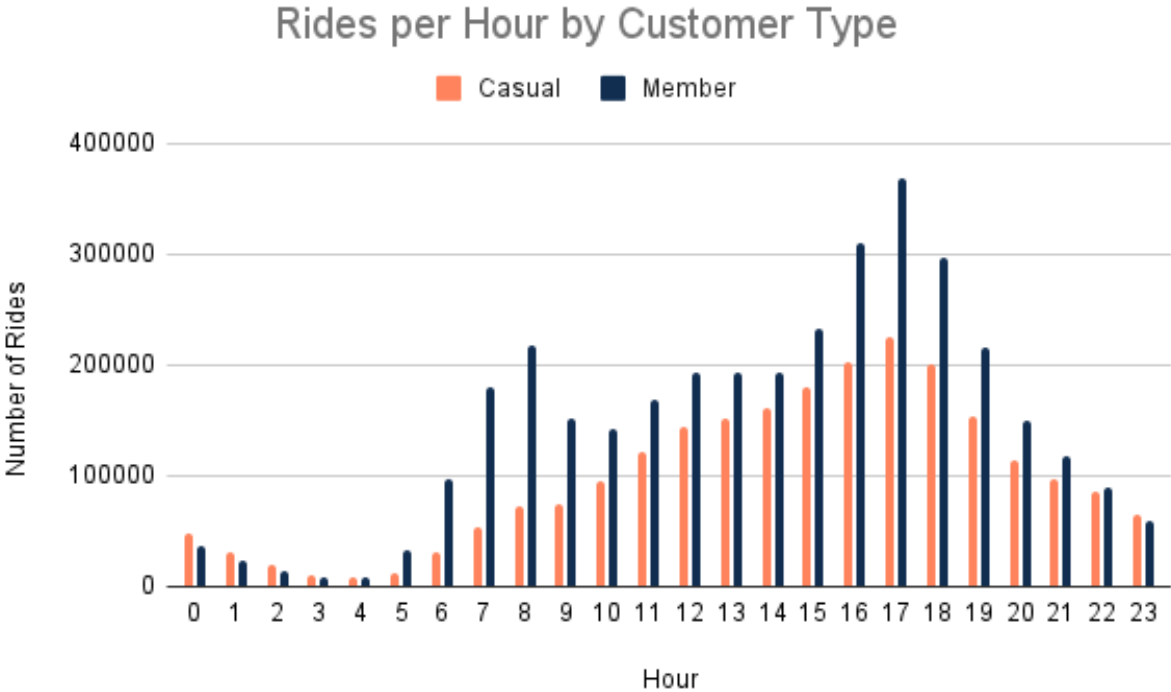


Member rides steadily increase throughout the week, while casual rides decrease throughout the week and are at their highest on the weekend.

Rides per Hour by Customer Type

To find the number of rides per hour based on customer type:

```
SELECT member_casual, EXTRACT (hour from started_at) AS hour, COUNT(*) AS
number_of_rides
FROM `bike-share-385304.Bikeshare.052023-042022`
GROUP BY hour, member_casual
ORDER BY member_casual, hour
```

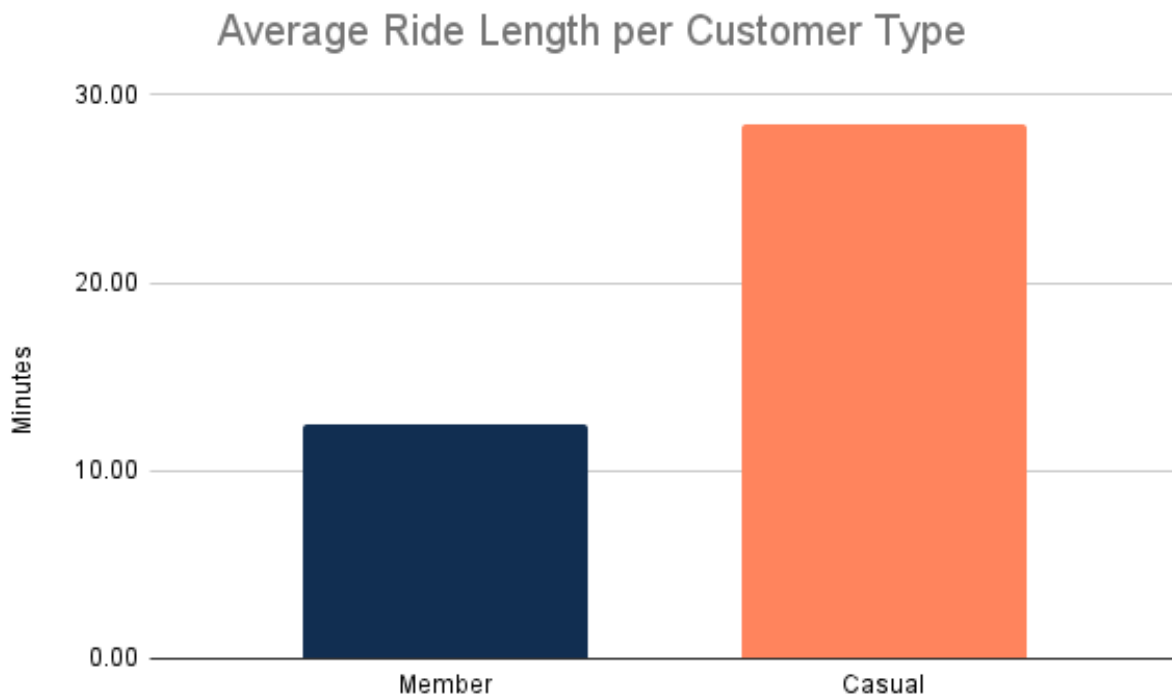


Member rides peak at the 8 AM and 5 PM hours and casual rides increase gradually throughout the day and then decrease at the 5 PM hour.

Average Ride Length per Customer Type

For average ride time based on customer type:

```
SELECT member_casual, AVG (ended_at - started_at)
FROM `bike-share-385304.Bikeshare.052023-042022`
GROUP BY member_casual
```

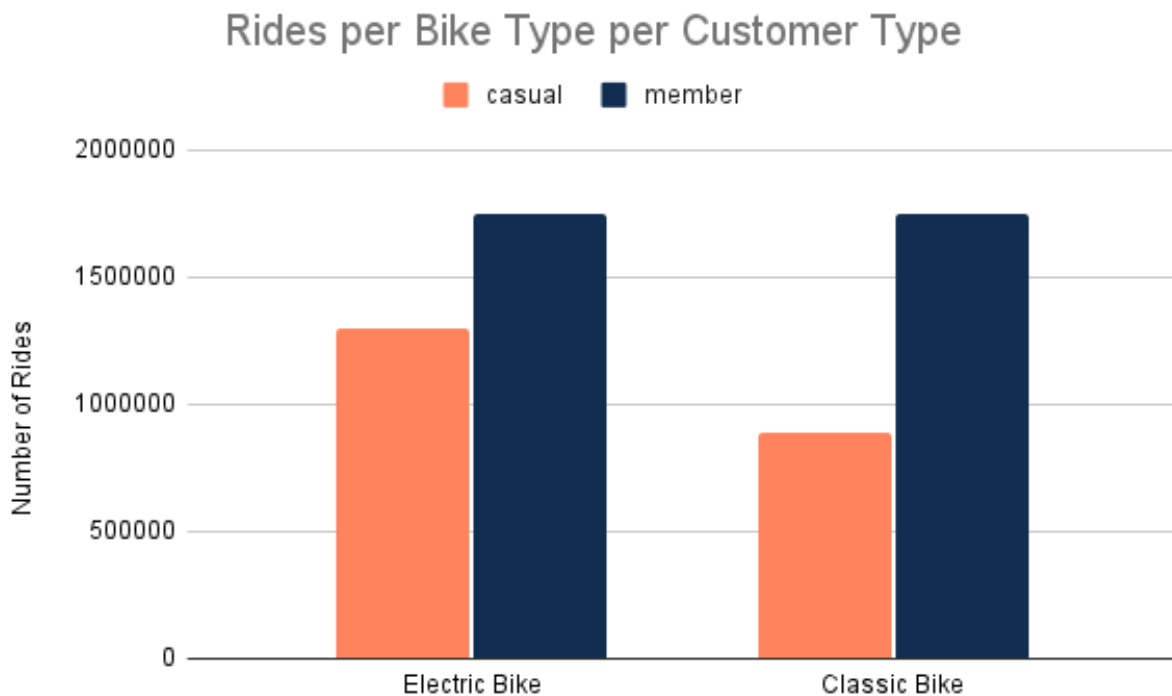


Members are riding an average of 12.48 minutes, while casual riders ride an average of 28.48 minutes.

Rides per Bike Type per Customer Type

To determine type of bike being ridden per customer type:

```
SELECT member_casual, rideable_type, COUNT (*)  
FROM `bike-share-385304.Bikeshare.052023-042022`  
GROUP BY member_casual, rideable_type
```



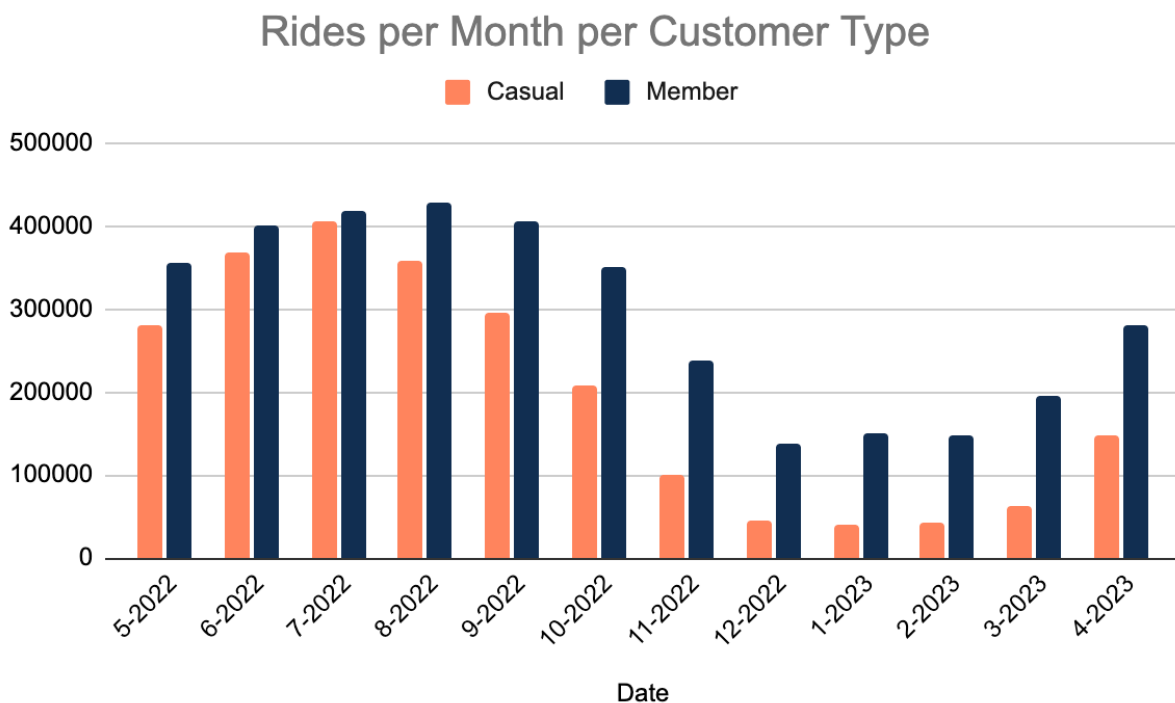
Rideable_type "docked" has not been included in this analysis due to an unclear definition. "Docked" type accounts for 2.9% of the data.

Casual riders are riding electric bikes 59.26% of the time, while members are only riding electric bikes 49.97% of the time.

Rides per Month per Customer Type

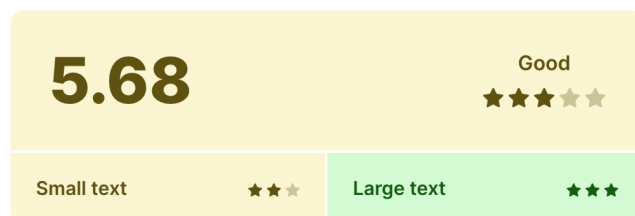
To determine number of rides per time of year:

```
SELECT EXTRACT (year FROM started_at) * 100 + EXTRACT( month FROM started_at) AS  
Year_Month, member_casual, COUNT (*) AS Rides_per_Month  
FROM `bike-share-385304.Bikeshare.052023-042022`  
GROUP BY member_casual, Year_Month  
ORDER BY member_casual, Year_Month
```



Rides are increasing during the summer and decreasing during the winter, however in the summer casual riders are riding 97.27% compared to members in July, but only 26.62% in January.

I checked the graph colors using a contrast checker for accessibility and they have a ratio of 5.68:1. Web Content Accessibility Guidelines require a ratio of at least 4.5:1.



Final thoughts

The most evident finding from this study is that members are riding predominantly at 8 AM and 5 PM on weekdays, and on average are only riding about 12 minutes. This suggests that members may be riding to and from work.

Casual riders are riding mostly on weekends. Their rides increase greatly in the summer months and on average last about 28 minutes. They are also riding electric bikes more than classic bikes. This data suggests they may be riding mostly for enjoyment.

Further insights

Due to privacy, this dataset did not contain Member ID, Customer ID, or credit card information that could be used to look at an individual's data. If that data were available, I would have liked to study member rider trends and repeat casual rider trends. This information could show us how much money the company is making or losing on memberships or how much they could save by converting casual riders to members.