Cyclistic Bike-Share Case Study

Abstract

A data-driven analysis to understand rider behavior and support membership growth strategies

AMMAR JAFFAR

ammarjaffar100@gmail.com

Section 1: Business Task

Cyclistic is a Chicago-based bike-share company with 5,800+ bikes and 700 stations. It offers three pricing options: single-ride, day pass, and annual membership.

Riders fall into two categories:

- **Casual riders** – single-ride or day-pass users
- **Members** – annual subscribers

An internal report found that **members provide more long-term value** than casual riders. Now, the company wants to **convert more casual riders into annual members**.

Business Objective

Analyze historical ride data to identify key behavioral differences between casual riders and members.
This insight will help Cyclistic target the right audience with the right message—boosting membership, revenue, and customer loyalty.

Section 2: Prepare

I used Cyclistic's historical trip data from the past **6 months** (instead of 12, to avoid system limitations) available via Divvy's public dataset, provided by **Motivate International Inc.** under a public license.

The data was downloaded, unzipped, and organized into folders. Using **Power Query in Excel**, I merged all CSV files into one.

Then, I prepared the dataset for analysis by:

- Adding a ride_length column (ended_at - started_at)
- Adding a day_of_week column (WEEKDAY(started_at))

No personally identifiable information (PII) was used, in compliance with privacy guidelines.

Section 3: Process

To get the Cyclistic dataset ready for analysis, I used **MySQL** for data cleaning and transformation.

- Removed 99 rows with corrupted ride_length or invalid ride_id.
- Set ride_id as the **Primary Key** for uniqueness and data integrity.
- Converted data types for accurate analysis:
  - started_at, ended_at → DATETIME
  - ride_length → TIME
  - Coordinates → DOUBLE
- Created a new column day_of_week using started_at.
- Verified nulls, checked for format issues, and ensured consistency.
- Documented the entire cleaning process for transparency and reproducibility.

Section 4: Analyze

In this phase, I used **SQL** and **Excel** to identify key trends and patterns in the ride data. The goal was to better understand user behavior and ride distribution across different segments.

Analysis highlights:

- Calculated **average** and **maximum ride lengths** to understand ride duration trends.
- Identified the **most common day of the week** for rides using frequency counts.
- Analyzed ride distribution by **user type (member vs. casual)** across days of the week.
- Measured the **percentage share of total rides** taken by each user type.
- Used Excel to further calculate:
  - Average ride duration by user type
  - Most frequent ride day (mode) using pivot tables

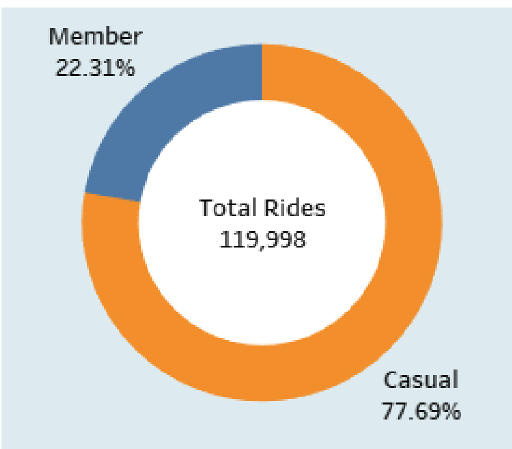
These analyses form the basis for identifying behavior differences between user groups and will guide the business recommendations.

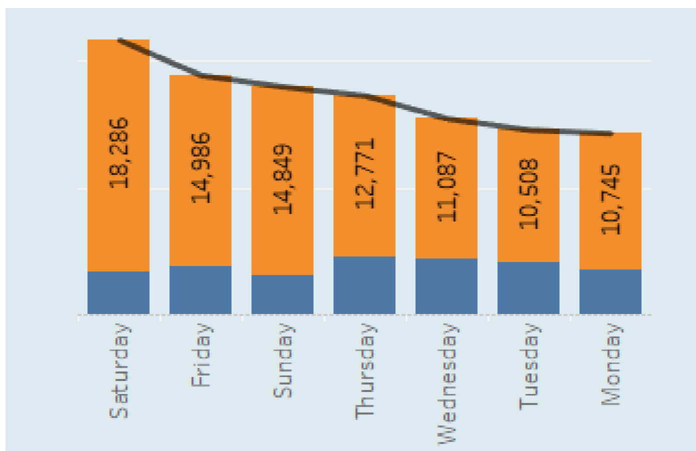
Section 5: Share

To effectively communicate the findings to the Marketing team, I created a donut chart that highlights the division of total rides between **Casual Riders** and **Members**. The visualization clearly shows that:

- **77.69%** of the total rides were taken by **Casual Riders**
- Only **22.31%** of the rides were contributed by **Members**, which is roughly one-fourth of the total share

This visual insight emphasizes the significant dominance of casual users in ride usage, offering a clear direction for tailoring marketing strategies, such as converting casual users into members or better targeting their preferences.

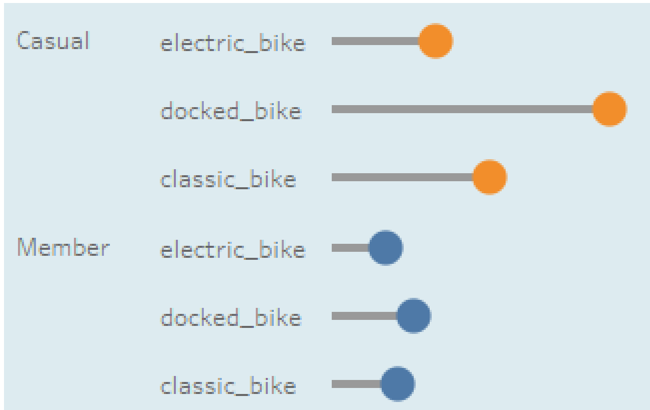




The second figure illustrates the **weekly ride patterns** of both **Casual Riders** and **Members**. It reveals a clear distinction in behavior:

- **Casual Riders** tend to ride **mostly on weekends**, showing higher activity on **Saturdays and Sundays**.
- In contrast, **Members** show increased usage during **weekdays**, aligning more with **commuting patterns**.

This comparison helps the Marketing team understand user behavior and optimize campaigns based on when each group is most active.



The third figure displays the **average ride duration (in minutes)** across different **user types** and **bike types**, revealing important trends:

- **Casual Riders** consistently have **longer ride durations** than **Members** across all bike types.
- While there are slight differences, the **ride duration pattern by bike type** generally mirrors the overall duration pattern:
    - **Docked bikes** have the **longest average ride time**
    - **Electric bikes** come next
    - **Classic bikes** have the **shortest duration**

These patterns suggest that both **user behavior** and **bike type** play roles in determining ride length, with casual users spending more time per ride—often for leisure—while members use bikes more efficiently for shorter, goal-oriented trips.

Additional analyses were conducted (not included in this report), focusing on location-based trends and change over time, which further supported the findings shared above.

Act: Turning Insights into Action

Based on the analysis, the Marketing team can take the following actions:

- **Convert Casual Riders to Members**: With 77.69% of rides from casual users, offer weekend membership deals or referral rewards to encourage conversion.

- **Weekend-Focused Campaigns**: Casual riders are most active on weekend launched promotions during these days to boost engagement.
- **Optimize Bike Availability**: Docked bikes have the longest ride durations. Ensure more docked and electric bikes are available on weekends.
- **Target Key Locations**: Watson St, Cler St, and Elm St show the highest casual rider activity. Focus localized campaigns and bike redistribution efforts here.
- **Next Steps**: Further analyze location trends and track membership conversion rates to refine future strategies.