Cyclistic Bike-Share Analysis (Capstone Project)

1. Business Task

Cyclistic, a bike-share company in Chicago, wants to increase revenue by converting **casual riders** into **annual members**.

The business question is:
How do annual members and casual riders use Cyclistic bikes differently?

2. Data & Preparation

•Data source: Cyclistic trip data (202408-divvy-tripdata) ~750,000 rows.

•Tools: Excel (data cleaning), Google Sheets (analysis & charts).

•Cleaning steps:

•Removed rides shorter than 1 minute and longer than 24 hours.

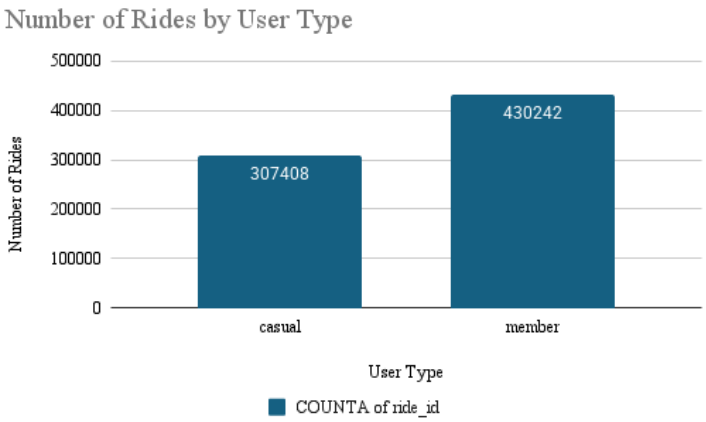
•Removed blank values and duplicates.

•Added calculated fields: ride_length, day_of_week.

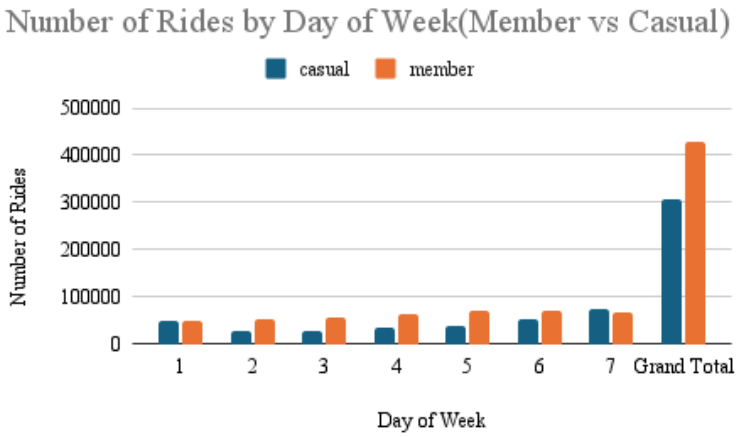
•Trimmed dataset to 7 relevant columns: ride_id, rideable_type, started_at, ended_at, member_casual, ride_length, day_of_week.

•Note: This analysis uses August only (representative sample). Other months are archived in data/archive/ and the full 12-month dataset can be processed using SQL/R for scalability.

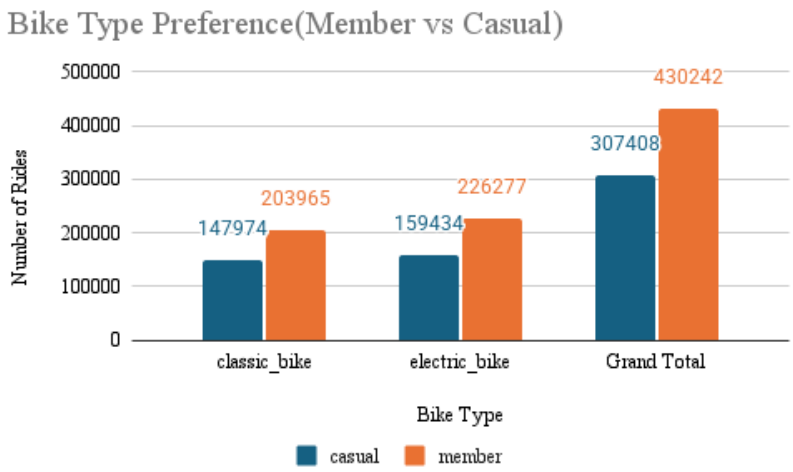
3. Analysis & Findings



Members completed significantly more rides overall (~430k) compared to casual riders (~307k) in August. This indicates that members are the more consistent and frequent users of the service.



Members ride most often during **weekdays (Mon–Fri)**, which reflects commuting behavior. Casual riders are most active on **weekends**, pointing to leisure and tourism-driven usage.



Casual riders show a stronger preference for **electric bikes**, while members mostly use **classic bikes**. This suggests casual riders prioritize convenience or leisure riding, whereas members rely more on traditional bikes for regular commuting.

4. Recommendations

- Offer **weekend discounts or trial memberships** to convert casual riders.
- Promote memberships as a **commuting solution** for frequent weekday riders.
- Adjust bike availability by **bike type preferences** (electric for casuals, classic for members).

5. Limitations & Next Steps

- This analysis covers only **August 2023** due to Excel/Sheets row limits.
- Seasonal patterns (e.g., summer vs winter) are not captured.
- The same cleaning + analysis process can be scaled to 12 months (~5M rows) using SQL or R.