Chicago Bike Share Analysis

Team "Small P-Values"
Members: Yongqi Liu, Wei-en Lu, Rui Qiang

# Heatmap of Rides Initiated

Heatmap of Ride Starts in Chicago

Data Cleaning

1. We deleted all rides with no return information.

2. We took out the faulty data:
   a. Negative durations.
   b. Extremely long single rentals (>20 hours)

3. We consolidated the station and trip information by:
   a. imputing any missing variable as long as one piece of identifying information (station name, station id, or long/lat) is available.
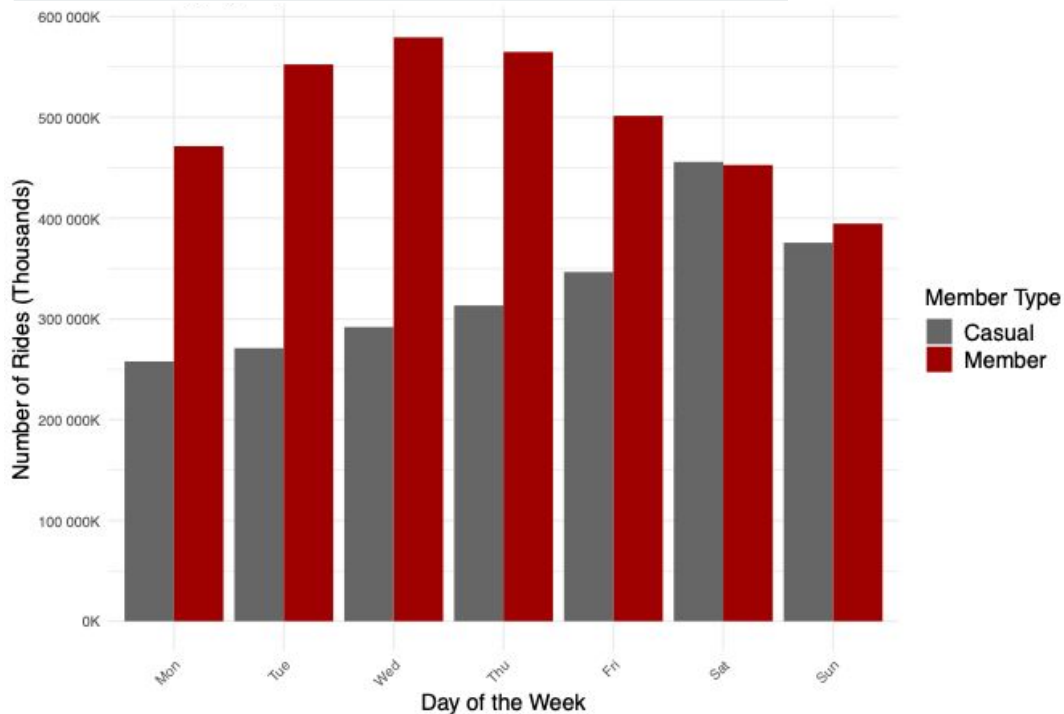
Bike Usage By Hour
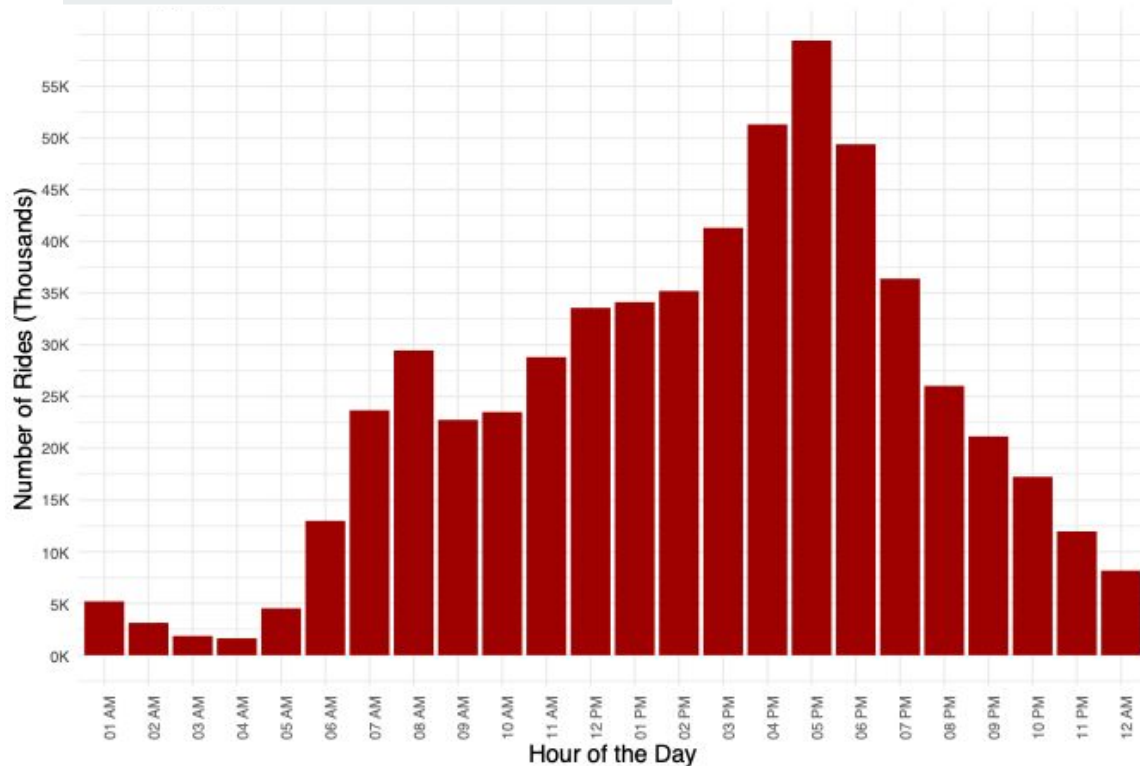
Here we plot the number of bike rentals by the hour of day for 2023.

1. Not accounting for weekday vs weekend effect, we notice that **most rentals start in the day time**, between 6AM and 9PM.
2. Bike rentals peak around early afternoon, at **5 PM**.

Bike Usage By Day of Week

Here we plot the number of bike rentals by the day of week for 2023, broken down by membership status.

1. Total bike rental numbers seem pretty consistent across the entire week.

2. Members tend to rent more during weekdays, and casual riders tend to rent more on weekends.

GLM Model Predictions

Day of the Week
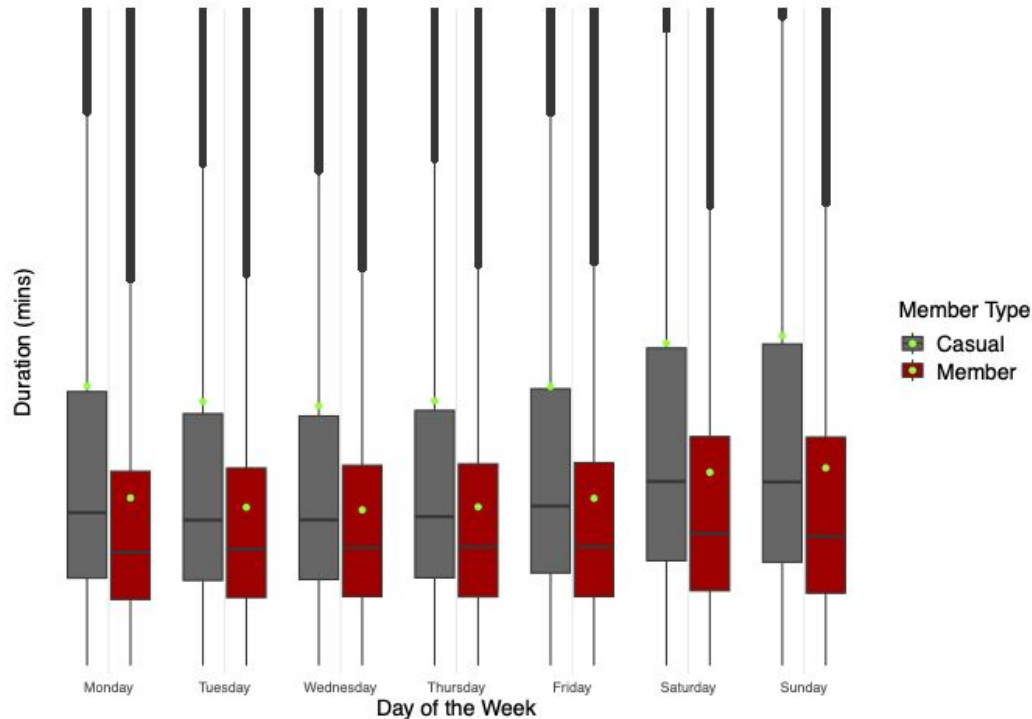
We fitted a Gaussian Generalized Linear Model (GLM) with a log link function

1. Response Variable:
   a. Duration in minutes.

2. Predictor Variables
   a. Day of week.
   b. Member vs Casual.

3. All coefficients are significant with p-values < 0.0001.

4. Bonferroni adjustment for multiple comparison still yields significant results.

Weekday Bike Usage by Hour and Rider Type
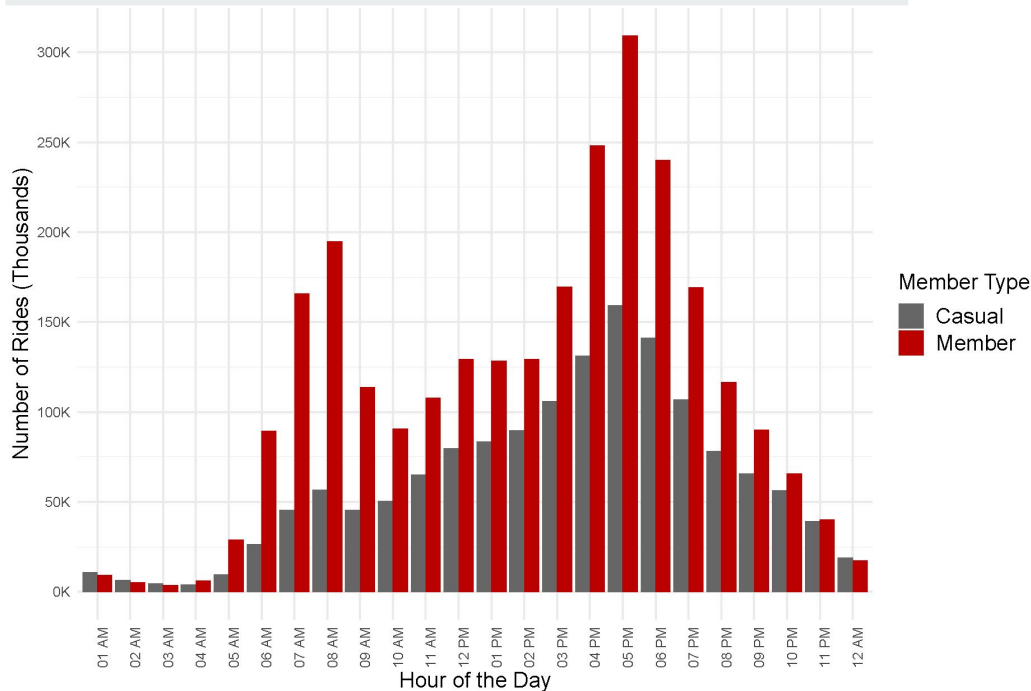
Here we plot the bike rental numbers by hour of day on **weekdays**, categorized by membership status.

1. In the commuting hours, we see a lot of more member riders vs casual riders. This is due to members consistently biking to commute to and from work.

2. For casual members, we expect them to ride the bikes for recreational purposes. Therefore, their usage numbers linearly increase from early morning to late afternoon as expected.
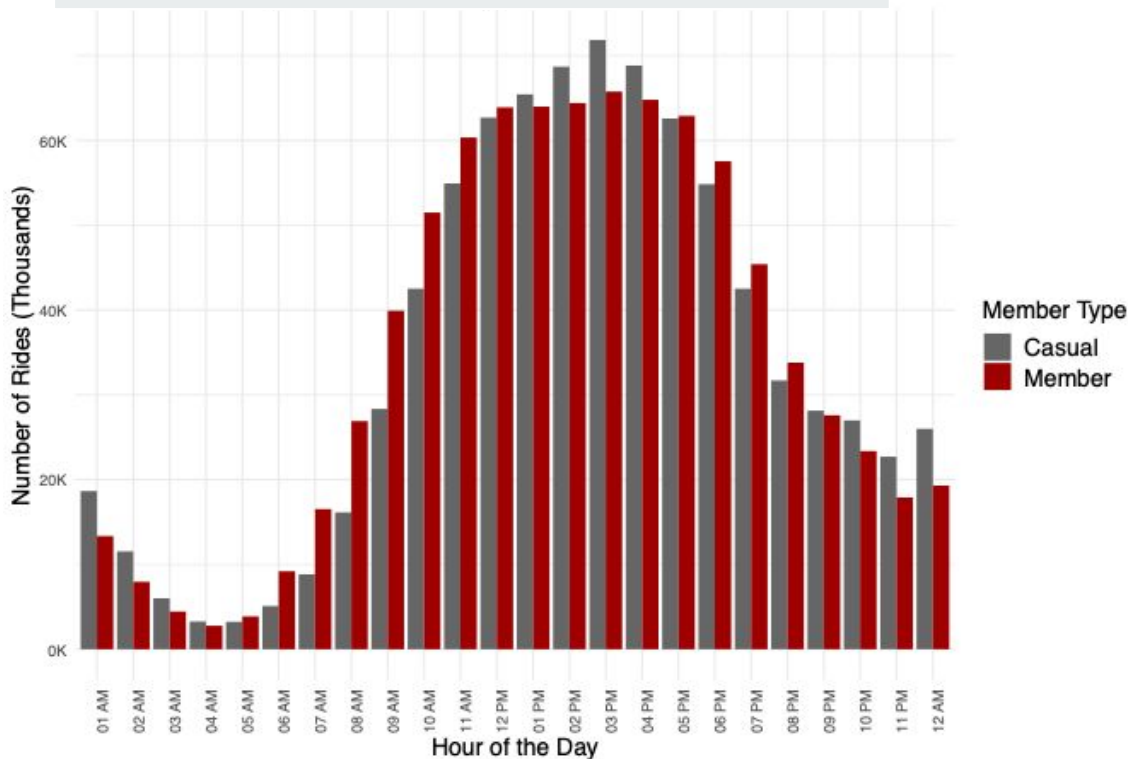
Weekend Bike Usage by Hour and Rider Type

Here we plot the bike rental numbers by hour of day on **weekends**, categorized by membership status.

1. Different from the previous plot, we see that member and non-member rides have approximately the same number across the entire day.

2. This further reaffirms the idea that most of members use bikes for daily commutes.

Business Insights

- We observe that **60%** of all rides are from member users, vs **40%** from casual riders..
  - Further marketing campaigns should aim at retaining current members, since it constitutes a significant proportion of our rides.

- On weekends, bike rental patterns between member and casual rentals are **similar**.
  - Provide new recreational plans at a discounted rate to attract more casual riders to become members.

- **45 minutes** is the **95%** percentile of all ride durations.
  - For riders who exclusively ride on weekdays for commute, develop a plan to give **two 45mins** rental periods on workdays.

Future Directions

- The dataset is has a limited number of covariates:
    - **USER_ID** can track individuals to provide each of them with a **selection of customizable subscription plans**.
    - **Bike_IDs** would help us understand maintenance schedule, total length of trips, etc.

- The **number of docks at each station** helps us understand any overflow problems (more bikes than docks), so that we can derive a plan to build more stations where demand is high.

Appendix