Cyclistic usage patterns across Member and Casual riders

For Q1 of 2019

Business Task

To analyze Cyclistic's historical bike trip data to identify and understand the distinct usage patterns between annual members and casual riders for the first quarter of 2019. This analysis will be used for Cyclistic's marketing strategy, as the company aims to convert casual riders into annual members to drive future growth and profitability. The findings will inform the marketing team's approach and will be supported by visualizations for executive team approval.

Data Source and Details

Data Sources: We will be using Cyclistic's historical trip data for the first quarter of 2020.

License: Data made available by Motivate International Inc. under their license agreement

Format: .csv

Location: Securely hosted on my personal computer.

Key Data Points

Available Columns:

- Trip start and end times
- Day of week information
- Rider type (member vs casual)
- Start and end station information
- Gender
- Date of birth

Data Privacy Considerations

- No personally identifiable information (PII) is included
- Cannot link past purchases to credit card numbers
- Cannot determine if casual riders live in service area
- Cannot track multiple single-pass purchases by same user

Data Limitations

- Limited to just one quarter of a year.
- Privacy restrictions limit ability to track individual user behavior.
- Cannot directly connect purchase patterns to individual users.
- Incomplete gender and DOB information making columns unusable.
- There is no trip duration column.
- Cannot track multiple single-pass purchases by same user.

Does the data ROCCC?

- **R**eliable – The data comes from a trustworthy source and is consistent.
- **O**riginal – The data is from the original source, not second-hand or altered.
- **C**omprehensive – The data covers the full scope of what's needed, not just a partial view.
- **C**urrent – The data is up-to-date and relevant to the present context.
- **C**ited – The data includes proper references or citations to its source

Reliable

✓ YES

- Data comes from a credible source (Motivate International Inc.)
- Data is from the company's own tracking system
- Comprehensive system-wide data collection
- Automated data collection reduces human error

Original

✓ YES

- First-party data collected directly from Cyclistic's bike-share system
- We have access to the original source, not second-hand data
- Data comes straight from the bike tracking and station systems

Comprehensive

✓ MOSTLY

- Covers all rides in the system
- Includes key metrics (times, stations, membership types)
- Contains essential trip data

Current

✓ NO

- Covers data only from quarter one of 2019
- Regularly updated tracking system
- Reflects historic user behaviour patterns
- Real-time data collection system

Cited

✓ YES

- Clear data source (Motivate International Inc.)
- Proper licensing in place
- Official company data
- Traceable to its origin

Overall ROCCC Assessment: ★★★☆☆ (3/5)

Outside of this exercise, the data would provide questionable insights!

Data Cleaning and Preparation

- I imported the dataset into Google Sheets.
- I created a trip duration column and calculated the trip duration for each trip ID.
- Checked for duplicates and blanks in the critical trip ID column and eliminated them.
- Eliminated the latitude and longitude data as it is irrelevant for our analysis.
- Renamed the columns so that they are easier to understand.
- Eliminated the gender and birth year columns for privacy reasons and because it was blank ridden, therefore inappropriate for analysis.
- I eliminated the specific bike ID column because we are not trying to answer questions about specific bikes.

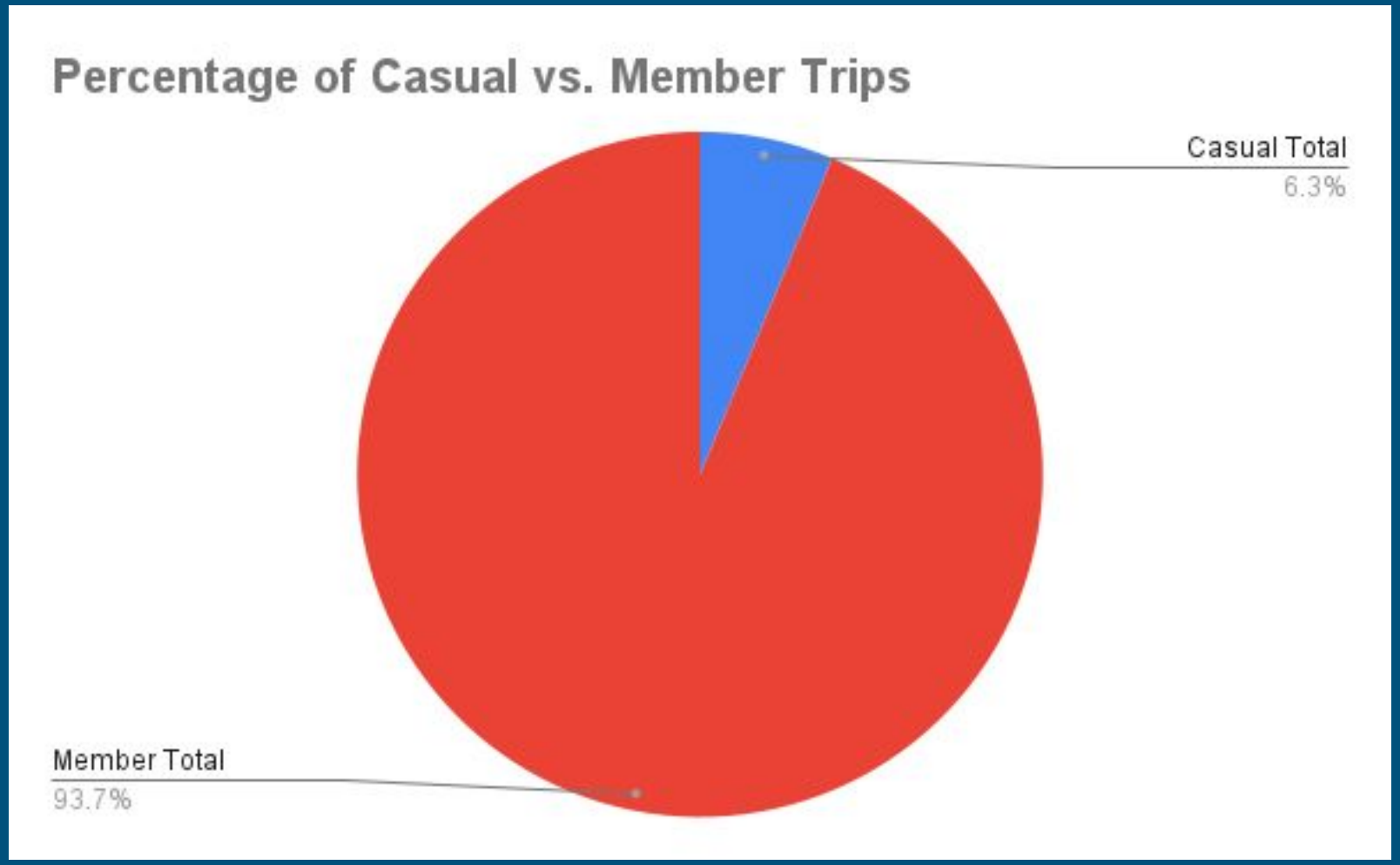
Data Cleaning and Preparation

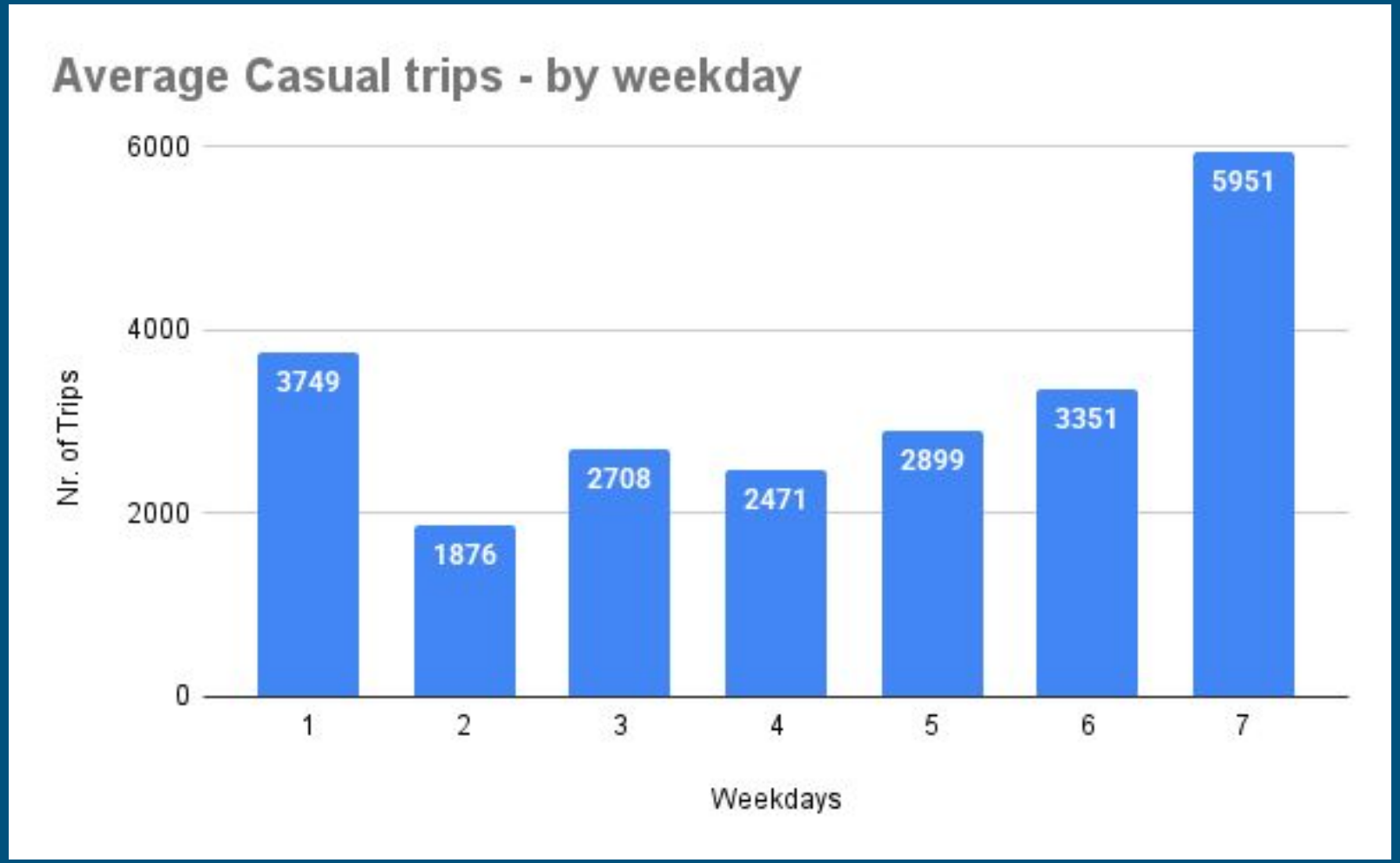
- I rearranged the column order to be easier to follow.
- I created a trip duration column and calculated the trip duration so that the dataset has them available.
- I started performing descriptive analysis to check for outliers.
- I checked for the average trip duration of members and casual riders.
- This led to me discovering outliers when the average trip of a casual rider came out at about 1.6 hours compared to 11 minutes for members.
- I checked for outlier values by sorting the length of trips and found trip lengths as long as 38 hours. This was probably the case due to improperly docked bikes and/or technical errors. I eliminated all trips longer than 6 hours to account for these errors without damaging legitimate trips.

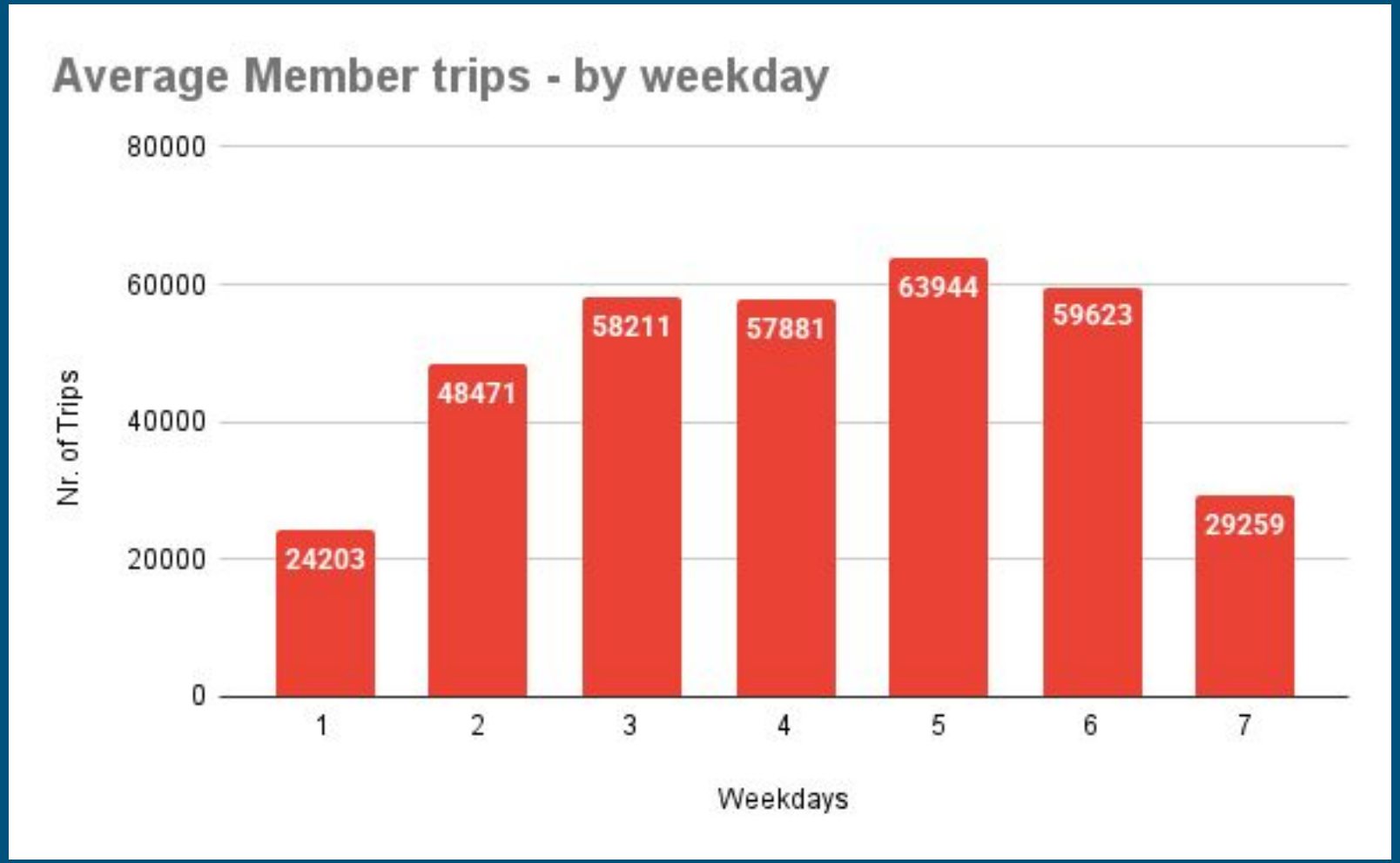
Analysis

- I separated the visualizations for members and casual riders as their volumes are drastically different and this ruins our ability to visualise their trends alongside each other.

The days of the week, Sun-Mon will be represented as 1-7

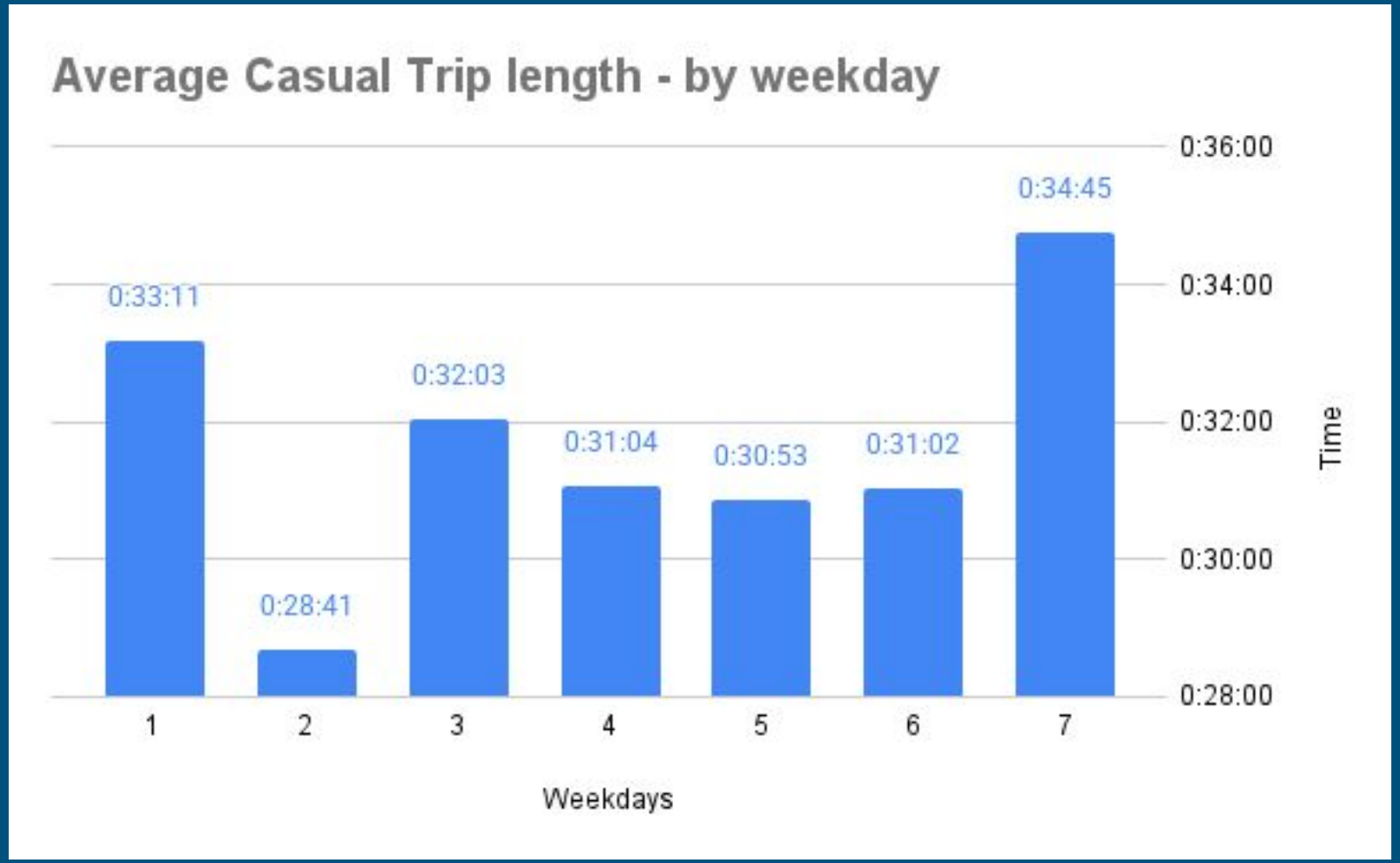

Percentage of Casual vs. Member Trips
Casual Total
6.3%
Member Total
93.7%

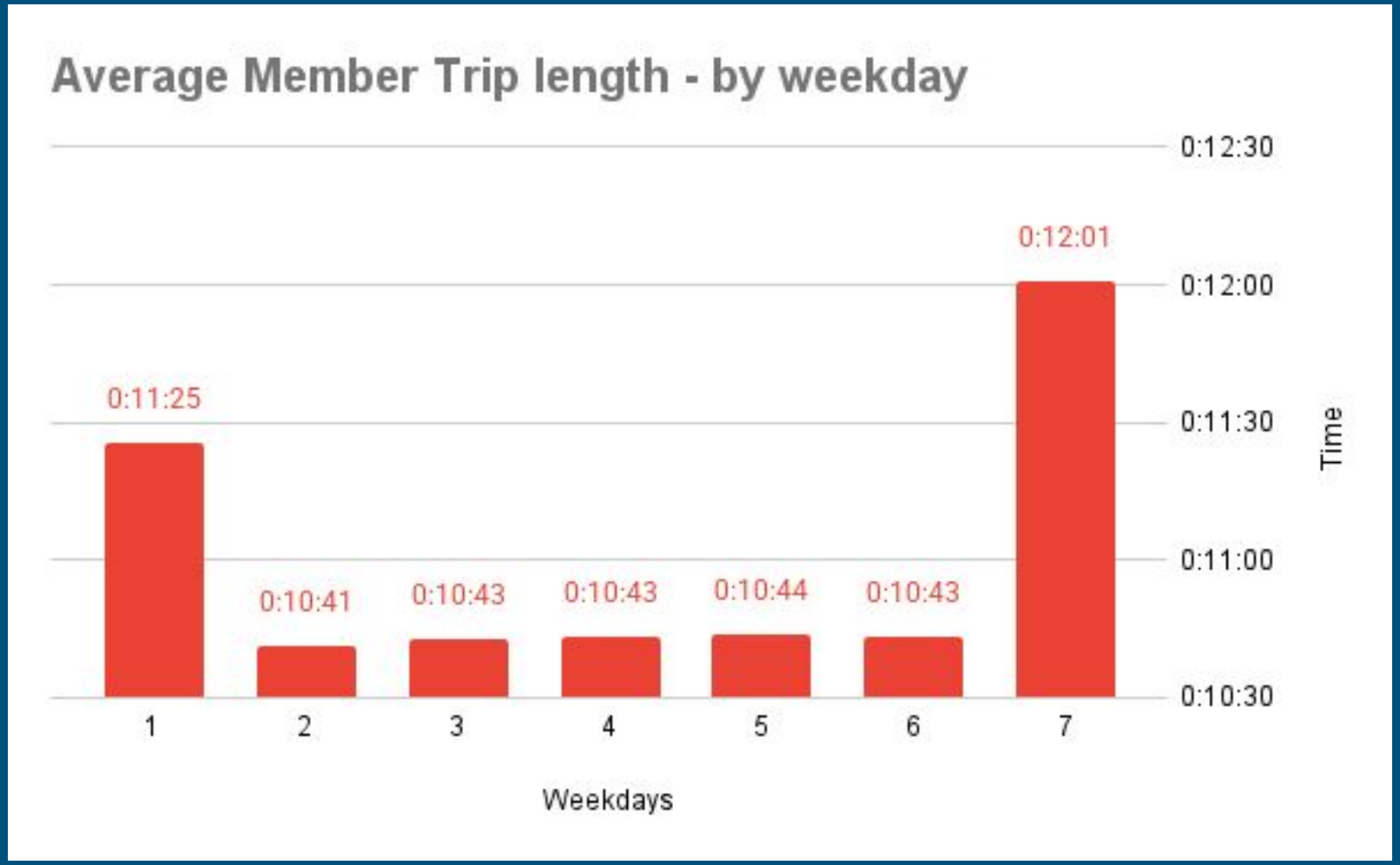

Average Casual trips - by weekday
6000
4000
2000
0
Nr. of Trips
3749
1876
2708
2471
2899
3351
5951
1
2
3
4
5
6
7
Weekdays


Average Member trips - by weekday
80000
60000
40000
20000
0
Nr. of Trips
24203
48471
58211
57881
63944
59623
29259
1
2
3
4
5
6
7
Weekdays

Trip Volume - Observations

- Sat, Sun, Friday, most trip intensive days for the casual riders. Mon, Wed, Tue see the least traffic for this group.
- Thu, Fri and Tue are the most trip intensive days for members with reliable steady high numbers throughout the working week but massive dips in usership during the weekend, with Sun, Sat being drastically lower in numbers followed by Mon.
- Casual trips pick up on the Friday leading into the weekend and peak on the weekend whereas the member trips pick up on the Monday and peak through the working week.

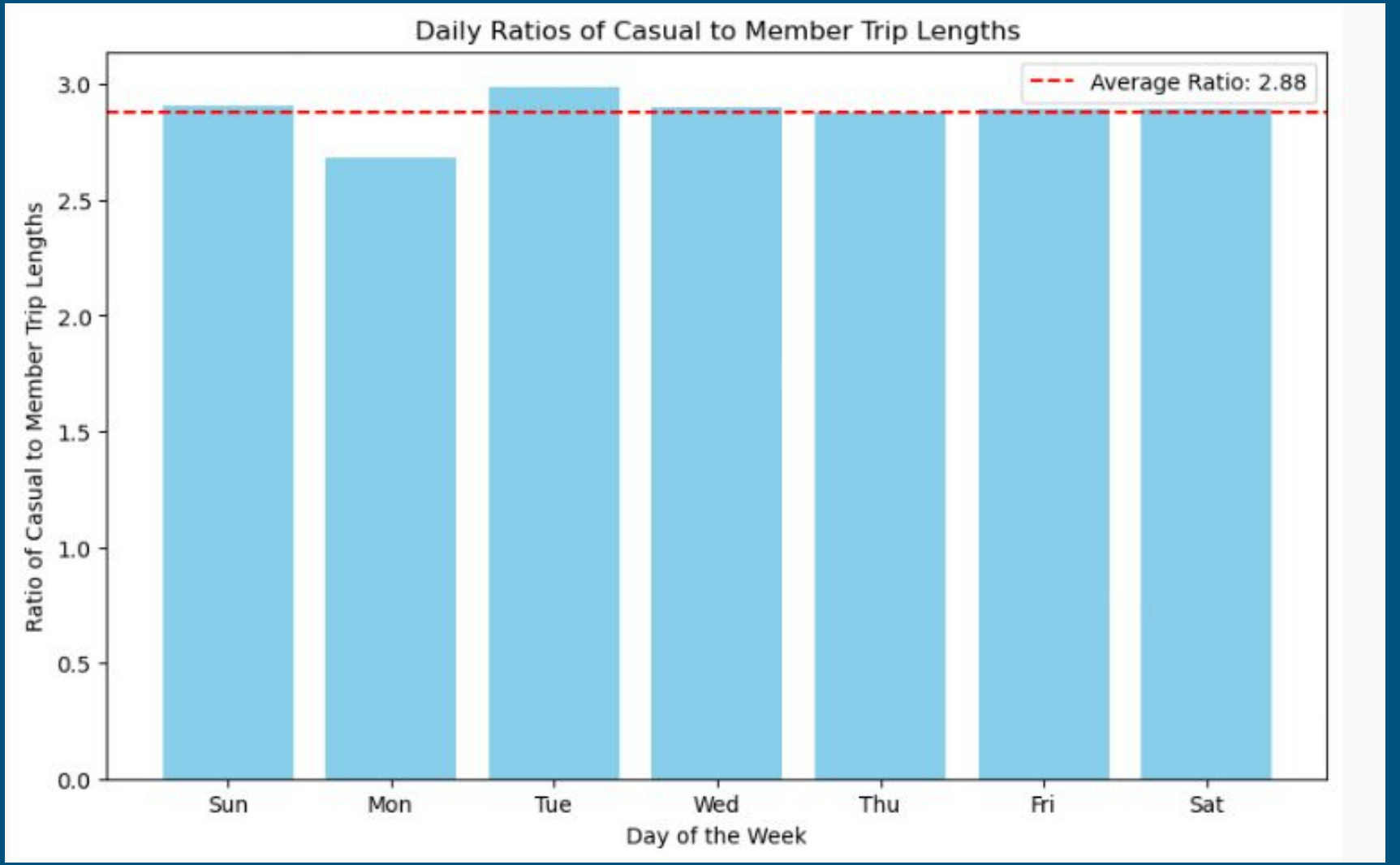

Average Casual Trip length - by weekday
0:36:00
0:34:00
0:32:00
0:30:00
0:28:00
Time
0:33:11
0:28:41
0:32:03
0:31:04
0:30:53
0:31:02
0:34:45
1
2
3
4
5
6
7
Weekdays

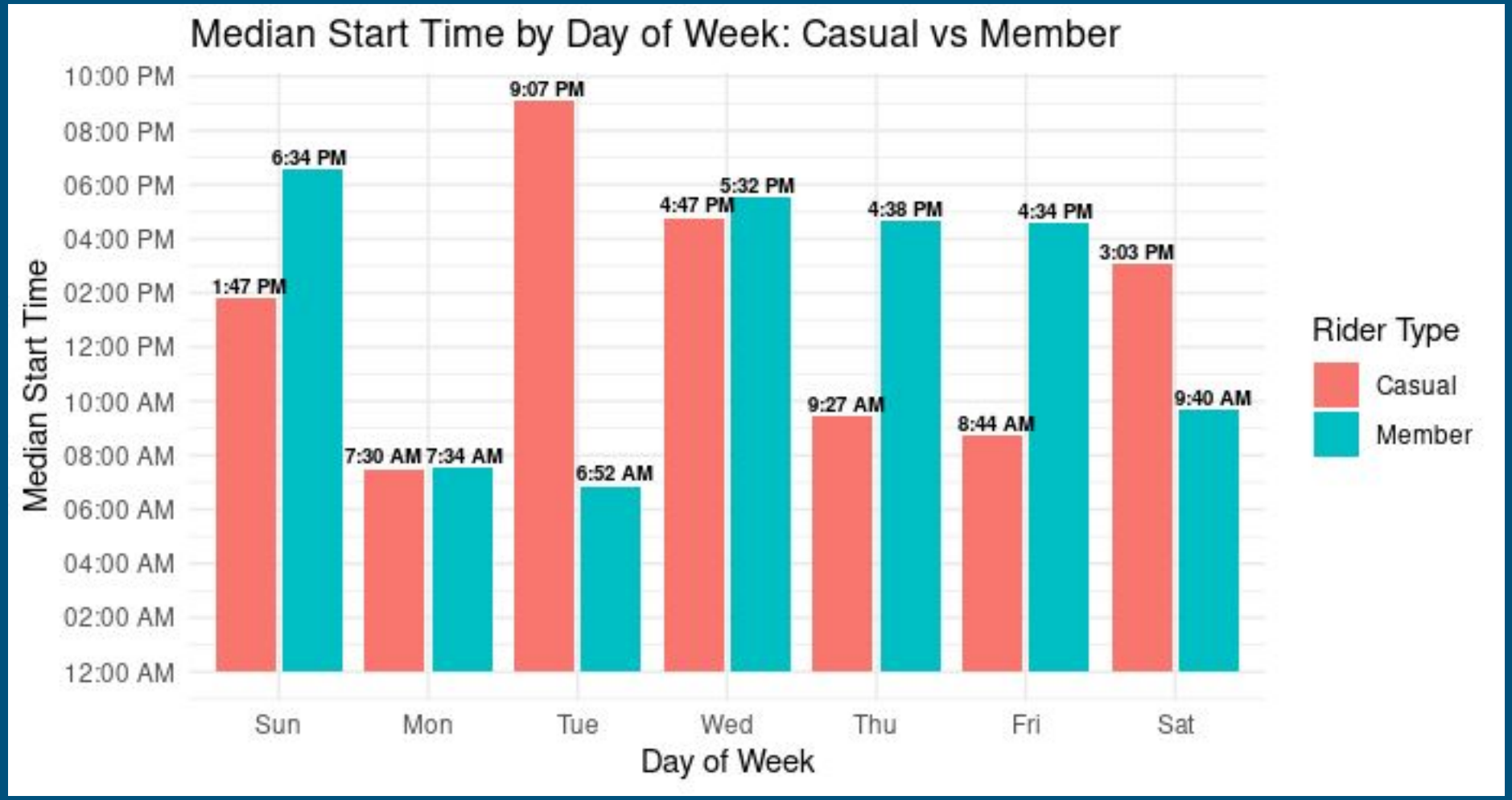

Average Member Trip length - by weekday
0:12:30
0:12:00
0:11:30
0:11:00
0:10:30
Time
0:11:25
0:10:41
0:10:43
0:10:43
0:10:44
0:10:43
0:12:01
1
2
3
4
5
6
7
Weekdays

Casual riders take their time!

- Overall Average Ratio for Casual to Member Trip lengths is: **2.88**

- This means that **the average trip of Casual riders is nearly 3 times longer than that of Members**


Daily Ratios of Casual to Member Trip Lengths
Average Ratio: 2.88
Ratio of Casual to Member Trip Lengths
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Sun
Mon
Tue
Wed
Thu
Fri
Sat
Day of the Week


Median Start Time by Day of Week: Casual vs Member
Median Start Time
10:00 PM
08:00 PM
06:00 PM
04:00 PM
02:00 PM
12:00 PM
10:00 AM
08:00 AM
06:00 AM
04:00 AM
02:00 AM
12:00 AM
1:47 PM
6:34 PM
7:30 AM
7:34 AM
9:07 PM
6:52 AM
4:47 PM
5:32 PM
9:27 AM
4:38 PM
8:44 AM
4:34 PM
3:03 PM
9:40 AM
Sun
Mon
Tue
Wed
Thu
Fri
Sat
Day of Week
Rider Type
Casual
Member

Start Time Analysis

- On average, member start times are slightly later in the day, most days, 5/7, and 4/5 if referring to the working week, the exception being Tuesday. In both cases, although more prevalent in Member Riders, the start times match what we can assume is a commute schedule around their working hours. Both have the same number of days where the median start times are in the later half of the day, 4/7.

Initial Analysis Outcomes

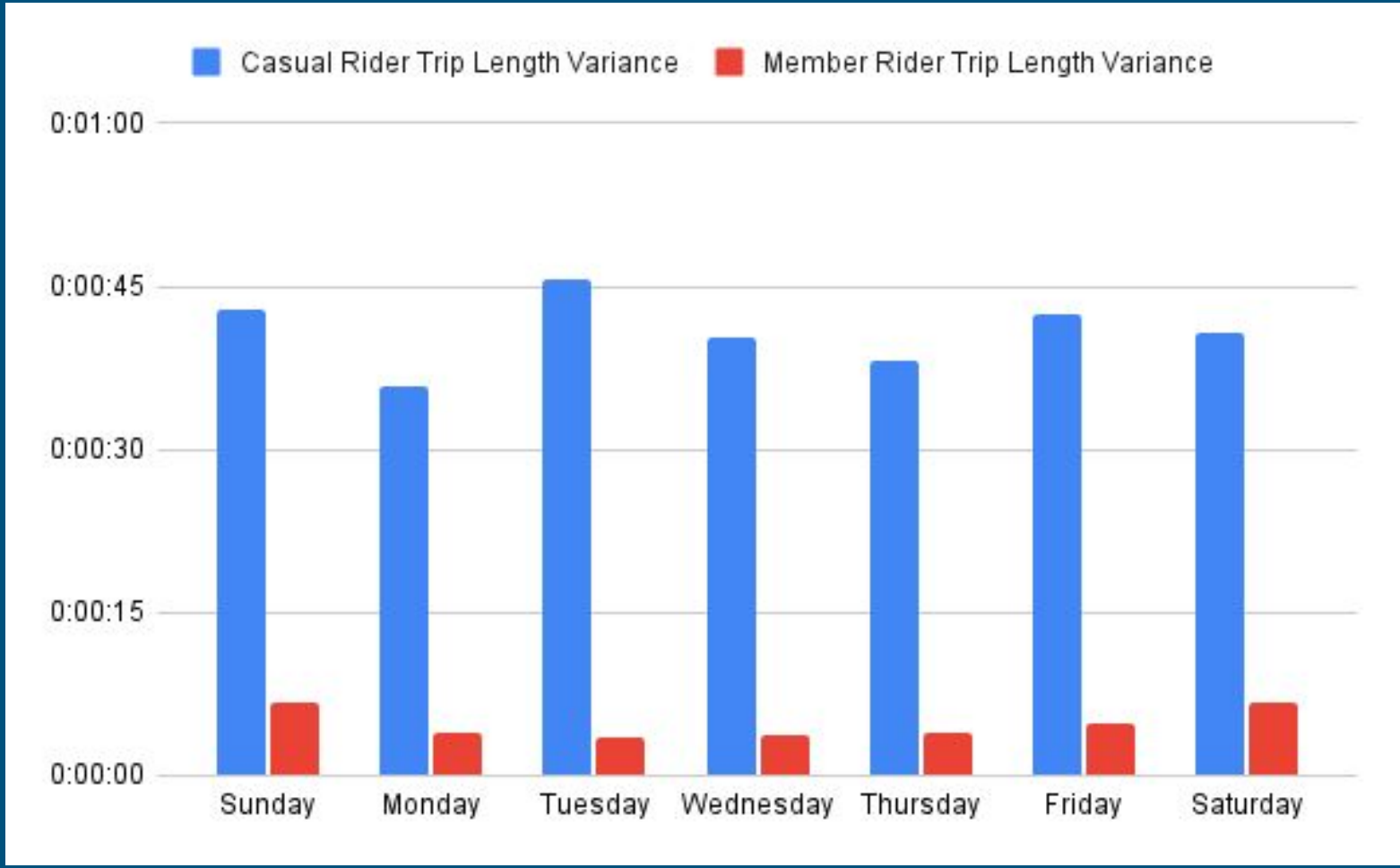
Based on the analysis so far we can deduce that Casual riders take far longer trips than Members and use the service particularly heavily during the weekend, implying that their trips may have other purposes such as pleasure, rather than business.

However, despite a drop in travel times during the work week, Casual rider trips remain far longer than Member trips and maintain similar ‘office hour’ oriented start times overall. This suggests something more interesting. It suggests that Casual riders also use the service for commuting to work but also, on average, live much further away from their workplace. This could explain their reluctance towards making this service their main, consistent mode of travel to work.

Final Steps

To test our preliminary analysis outcomes further with available data, I have calculated the variance of travel times between Member and Casual Riders and have determined their preferred start and end docking stations for each day of the week.

Rider Trip Length Variance



With the overall cross-week variance for Casual Riders being 41 seconds and 4 seconds for Members, you can tell that Members are 10 times more consistent with their ride length, once again supporting the idea that they use it mostly for commuting to work. However, it is notable that a 41 seconds variance for a trip is still narrow in real-life terms.

Popular Start/End Stations by Rider Type

Member Riders (Starts & Ends)

- **Monday to Friday:**
  The most popular start and end station is consistently **Clinton St & Washington Blvd**.
  This pattern is very strong and suggests that this station is a major hub for weekday commuting among members—likely located in a business or transit district.

Popular Start/End Stations by Rider Type

- **Saturday and Sunday:**
  The most popular start and end station shifts to **Kingsbury St & Kinzie St**.
  This indicates a significant change in member rider behavior on weekends, possibly reflecting leisure or residential area usage.

Popular Start/End Stations by Rider Type

Casual Riders (Starts & Ends)

- **Monday to Thursday, Saturday:**
  The most popular start and end station is **Streeter Dr & Grand Ave**.
  This station is likely close to tourist attractions, the waterfront, or entertainment areas, which casual riders frequent.

Popular Start/End Stations by Rider Type

- **Sunday and Friday:**
  The most popular start and end station is **Lake Shore Dr & Monroe St**.
  This suggests that on weekends and Fridays, casual riders are especially drawn to the lakefront, possibly for recreational activities.

Key insights - Starts & Ends

- **Distinct Usage Patterns:**
  Member and casual riders have different preferred stations, reflecting different purposes:
  - **Members:** Consistent weekday commuting, with a shift to leisure/residential areas on weekends.
  - **Casuals:** Strong preference for scenic or tourist-friendly locations, especially on weekends and Fridays.

Key insights - Starts & Ends

- **Station Roles:**
  - **Clinton St & Washington Blvd:** Major commuter hub for members (weekdays).
  - **Kingsbury St & Kinzie St:** Weekend hotspot for members.
  - **Streeter Dr & Grand Ave:** Main choice for casual riders most days.
  - **Lake Shore Dr & Monroe St:** Casual rider favorite on Sundays and Fridays.

Key insights - Starts & Ends

- **Behavioral Trends:**
  - Members are likely using the service for work commutes during the week and for leisure or errands on weekends.
  - Casual riders' station choices reflect recreational or tourist activity, peaking on weekends and Fridays.

Analysis Summary

Members seem to be primarily using the service for work commutes during the week and for leisure or errands on weekends. With the overall cross-week variance for Casual Riders being 41 seconds and 4 seconds for Members, you can tell that Members are 10 times more consistent with their ride length, once again supporting the idea that they use it mostly for commuting to work. This is also supported by their preferred start and end stations. However, it is notable that a 41 seconds variance for a trip is still narrow in real-life terms.

Analysis Summary

Casual riders, overall, use the service to travel to tourist attractions and the waterfront peaking on weekends and on Fridays. Casual riders take longer trips. Casual Rider start times during the work week may suggest some also use the service for commuting to work on occasion but are reluctant to use it as their consistent commuting option.

Suggestions

Based on our prior analysis, it seems that, in order to attract more Casual riders towards becoming Members, we need to tackle 2 things:

- Find ways of reducing travel times such as by employing electric bicycles and/or placing docking stations closer to where the Casual riders live and work.
- Make the subscription more attractive than other modes of transportation such as public transport by offering discounted subscriptions for those who travel over longer distances.

Suggestions

- Include weekend and route-specific deals with the Membership, particularly targeting the waterfront, entertainment areas, areas with tourist attractions and these two stations in particular: **Lake Shore Dr & Monroe St, Streeter Dr & Grand Ave.**

Further Research

- It would be valuable to analyse available data over all quarters of a more recent year to check for any seasonal or long-term fluctuation in our results.
- Considering holidays and special events of interest (Concerts, Sports Games) as part of the year-round analysis mentioned above will help reveal further trends and potential outliers.

Further Research

- Demographics data would also be valuable to test for any correlation or preferences that would help our team market more effectively to our desired audience.
- A survey is needed to gauge the opinion of Casual riders regarding the Membership. We need to identify their reasoning behind avoiding it, and whether pricing, infrequent use, weekend tourism, station availability or commute distance are determining factors.