Cyclistic Bikeshare Member Usage

February 11, 2025

Table of Contents

Chapter organization places Suggestions and Summary Conclusions directly after the Introduction.
For readers wishing to follow the case study in analytical sequence, the chapter order is:

1. Introduction
2. Goals and Objectives
3. Data Testing, Transformation, and Validation
4. Validated dataset
5. Analyses, Data Summaries, and Results
6. Summary Conclusions
7. Suggestions

Michael H Haag
Senior Data Scientist
Tech Write LLC
www.techwrite2011.com
info@techwrite2011.com

1 Introduction

Cyclistic[1] operates a bike-share system in Chicago, IL offering more than 5,800 bicycles of various types and numerous docking or rental stations. The stations are located at selected locations throughout the Chicago urban area. Member customers consist of Annual members and Casual members.

The purpose of this study is to identify and quantify usage patterns between the two membership classes.

A number of parameters are explored and Membership class differences and usage patterns are identified.

1 Cyclistic is a fictional company; however, real-world data from Google Analytics is used in this study.

2 Suggestions

Suggestions are offered to highlight, and as possible ways to utilize, the conclusions and findings in the Summary Conclusions chapter. They are only suggestions, and may or may not be viable business decisions, investments, or actions. Those decisions, responsibilities, and risks therefrom are Cyclistic's alone. Cyclistic will need to pre-evaluate expected outcomes against expected costs of implementation.

1. Casual member ride-rentals are generally associated with tourist attractions and public parks. Consider co-operative promotions with these other entities.
2. Casual member's ride-rentals dominate the round-trip routes. While it is likely that their rides are for sight-seeing, pleasure riding, or exercise, those motivations cannot be determined from the ride data. Consider performing rider surveys before committing the promotion's focus on these activities.
3. Casual member's rides are notably longer than existing Annual members. Consider highlighting the bicycles' attributes such as easy operation, comfort, convenience, and availability in the promotion.
4. Both Casual and Annual members have about the same preference for the electric bicycles. But Casual members prefer the electric scooters at twice the rate of Annual members. Consider including, or test marketing, a rate discount on the electric bicycles and/or scooters for upgrading to Annual membership.
    1. Consider emphasizing the time-saving advantages of the electric bicycles and scooters to enhance electric rentals.
5. Casual member's round trips primarily occur at a select few stations. Consider focusing on those locations for advertising media.
6. Casual member's trips occur on Fridays and Saturdays. Consider timing advertising to the end of the workweek.
7. Casual member's ride-rentals peak in July. Consider timing promotions to start late Winter or early Spring.
8. Peak ride time-of-day coincides with the Noon hour. Consider timing advertising for late mornings.

3 Summary Conclusions

7.1 Ride-rentals by membership type: Annual members account for 64.1 % of bike-rentals while Casual members' share is 35.9%.

7.2.1 Rental time durations for all membership types: The majority of rides are short. The median time is only 10.2 minutes and 99% of all rides are completed within three hours.

7.2.2 Comparative rental time durations by membership type: Annual and Casual riders exhibit significantly different ride-rental time durations.

1. The median rental-ride durations for...
    1. Annual members is 8.9 minutes.
    2. Casual members is 13.4 minutes.
2. For longer ride-rentals the differences become more pronounced.
    1. Annual members completed 99% of ride-rental in 50.0 minutes or less.
    2. But, 99% of Casual members' ride-rentals were well over twice as long; 124.8 minutes.

7.3.1 Total stations and possible trips: There are 1,755 stations analyzed in the dataset and 3,080,025 possible trip routes including round-trips from, and returning to, the same station.

7.3.2.1 Most frequent trips by all members: The two most frequent trips by all members are round-trips; from, and returning to, the same station. Both stations are associated with tourist attractions and public parks:

1. Streeter Dr and Grand Av
2. Dusable Lake Shore Drive and E Monroe St

7.3.2.2 Most frequent round trips by all members: Round-trip stations are generally associated with tourist attractions or public parks.

7.3.2.3 Most frequent one-way trips by all members: The first two most frequent one-way trips by all members are a reverse trip pair: State St & 33rd St <==> Calumet Ave & 33rd St

7.3.2.4 Most frequent start stations used by all members: Almost 1.5% of all ride-rentals start at the Streeter Dr & Grand Ave station.

7.3.2.5 Most frequent end stations used by all members: The most frequent end station used by all members is Streeter Dr & Grand Ave at 1.5% of all ride-rentals.

7.3.3.1 Most frequent trips by combined membership types: The two most frequent trips are round-trips by Casual riders. Both station locations are in large public parks.

7.3.3.2 Most frequent round-trips by combined membership types: All top ten round-trips are by Casual members.

7.3.3.3 Most frequent one-way trips by combined membership types: Annual members represent seven out of the top ten positions.

7.3.3.4 Most frequent start stations by combined membership types: Six of the top ten most frequent start stations are attributed to Casual members.

7.3.3.5 Most frequent end stations by combined membership types: Five of the top ten end stations are attributed to casual members.

7.4.1 Bicycle type preference by membership type: Casual members prefer electric bicycles and, especially electric scooters, more than Annual riders:

1. Casual members rent electric bicycles for 34.34% of ride-rentals versus 33.18% for Annual members.
2. Casual members rent electric scooter for 1.68% of ride-rentals versus 0.81% for Annual members. Just over twice as often.

The two proportions are statistically significant to the 95% confidence level.

7.4.2.1 Average rental time by bicycle type (all member types): Rental time savings over classic bicycles are considerable: 25.9% for electric bicycles and 41.8% for electric scooters. A marketing advantage to be exploited.

7.4.2.2 Average rental time by bicycle type by membership type: Overall average rental times of Casual members, by bicycle type, are distinctly longer than those of Annual members: 95% longer for classic bicycles, 45% longer for electric bicycles, and 44% longer for electric scooters.

7.5.1.1 Rides by month - all members: The busiest months are July and August at 12.8% each. January is the least busy month at 2.6%. A ratio of about 4.9 to 1.

7.5.1.2 Rides by month - Annual members: September is the heaviest use month for Annual members at 7.6% of all rides. January reflects the lowest participation at 2.2 %. A ratio of about 3.4 to 1.

7.5.1.3 Rides by month - Casual members: July is the heaviest use month for Casual members at 5.5% of all rides. January reflects the lowest participation at 0.4 %. A ratio of about 13.8 to 1. A much sharper high-to-low ratio than that of Annual members.

7.5.1.4 Rides by month - combined member types: The highest ride-rental participation of both groups takes place during the warmer months. January is the most one-sided with Annual riders outnumbering Casual riders 5.5 to 1. For every month, Annual riders always outnumber Casual riders.

7.5.2.1 Rides by weekday (all members): Saturday is the busiest day for all members at 15.6% of total rides. Sunday is the least busiest day at 12.8%. Wednesday is the highest day at 15.2% for the five-day workweek.

7.5.2.2 Rides by weekday - Annual members: Wednesday is the busiest day for Annual members at 10.7%. Sunday is the lowest at only 6.9%. There is clearly more ride-rentals during the workweek than the weekend. This suggests Annual members are riding to work or school. Or, if they ride for exercise, do so mostly during the five-day workweek.

7.5.2.3 Rides by weekday - Casual members: Casual members ride activity is oriented towards weekends. The busiest days are Saturdays at 7.5%, followed by Sundays at 6.0%. For Casual members, the five-day workweek has generally lower ride participation. The one exception being Fridays at 5.5% which leads into the Saturday spike.

7.5.2.4 Rides by weekday - combined member types: Saturday is the busiest day for both member types combined at 15.6%3. This is due to the large spike of Casual riders making up for the lower weekend participation by Annual members. Although it's the second most active day for Casual riders, Sundays are the lowest overall at 12.8%.

7.5.3.1 Rides by hour - all members: The busiest hour is Noon at 9.9%. The least busiest hour is 10pm at 0.2%. For the most part, this plot is as expected: A steady increase during the morning, peaking late morning through the lunch hour, then followed by a steady decline. The notable exception is the early morning ramp-up. Rides begin increasing sharply starting at Midnight and continuing through 3am.

7.5.3.2 Rides by hour - Annual members: The busiest hour is Noon at 6.5%. The least busiest hours are 9 - 10pm.

The early morning ramp-up starting at Midnight from Figure 35 is also matched in this graph. Thus, these are mostly Annual members riding around Chicago at these times. Bars in Chicago close between 2am and 4am depending on their license. It seems logical to assume a relationship exists.

7.5.3.3 Rides by hour - Casual members: The busiest hour is Noon at 3.4%. The least busiest hours are 10 - 11pm at 0.1% each.

The steep early morning spike seen for Annual member is not present here, but there is a steady increase in rentals starting at Midnight which continues through the Noon peak.

7.5.3.4 Rides by hour - combined member types: The busiest hour for both member types combined is Noon at 9.9%. The least busiest hour is 10pm at 0.2%.

The early morning ramp-up for combined members peaks in the 3am hour at 5.5%.

4 Goals and Objectives

4.1 Purpose

Cyclistic's bicycle riders are classified as either Annual or Casual members. Cyclistic has commissioned this study to identify how Cyclistic members of the two membership classes differ in their use of the bicycles. Cyclistic plans to use these findings to support a future marketing program to convert Casual members to Annual members.

4.1.1 Definitions

Membership class, classification, or type refers to:

- **Annual members**: Cyclistic customers who purchase an annual membership.
- **Casual members**: Cyclistic customers who purchase either a...
    - Single-ride pass
    - Full-day pass
- **Rides:** Refers to the individual rides in the dataset. Rides may also be referred to as rental-rides, rental-trips, rentals, or trips. These are individual bicycle rentals, rides, or trips from a start (departure) station to an end (arrival) station.

4.2 Dataset description

The bicycle trip data is supplied as twelve monthly CSV files. Each ride is described as one file row in each file. There are thirteen data fields:

Table 1: Dataset column names and descriptions.

Column (variable)	**Description**
ride_id	A unique identifier for each bicycle ride (rental)
rideable_type	Describes the type of bicycle being rented
started_at	GPS rental start time UTC in date-time ISO format
ended_at	GPS rental end time UTC in date-time ISO format
start_station_name	The name of the departure bicycle rack location.
start_station_id	An identifier for the departure station
end_station_name	The name of the arrival bicycle rack location
end_station_id	An identifier for the arrival station
start_lat	GPS start station latitude
start_lng	GPS start station longitude
end_lat	GPS end station latitude
end_lng	GPS end station longitude
member_casual	The member type is either… • member: Pays an annual member subscription fee • casual: Pays a one-trip or one-day fee

4.3 Scope and analyses objectives of member type usage

The scope of this study the analysis of member usage: Specifically, evaluating how the two membership classes (Annual and Casual) use Cyclistic's bicycles. The analyses starts with broad-based measures of the entire dataset followed by progressively narrower measures of membership class usage. A number of measures and categories are evaluated. For this study, first-level analyses will be done using membership type as the categorical factor.

4.3.1 Analysis by membership types

Characterize differences in membership types with regard to:

- Rentals by each membership type
- Rental time statistics for each membership type
- Analysis of variance between membership types

4.3.2 Analyze favored stations by membership type

Determine individual stations favored by each membership type.

4.3.3 Analyze riders favored trips (station-to-station)

Determine the station-to-station trips by membership types

4.3.4 Analyze member choice of bicycle types

Cyclistic provides:

Classic bicycle

Electric bicycles

Electric scooters

Determine each membership type's preferred bicycle choice.

4.3.5 Analyze member rides by time periods

Determine each membership type use of the bike-share system by time periods:

1. By month
2. By weekday
3. By hour

5 Data Testing, Transformation, and Validation

5.1 Data import and consolidation

5.1.1 Import

The bicycle data is supplied as 12 monthly CSV files starting December 2023 and ending November 2024 imported as. Each file is imported to a separate temporary dataframe:

Dataframe	Import file	Rides
X202312_divvy_tripdata	DataFiles/202312-divvy-tripdata.csv	224,073
X202401_divvy_tripdata	DataFiles/202401-divvy-tripdata.csv	144,873
X202402_divvy_tripdata	DataFiles/202402-divvy-tripdata.csv	223,164
X202403_divvy_tripdata	DataFiles/202403-divvy-tripdata.csv	301,687
X202404_divvy_tripdata	DataFiles/202404-divvy-tripdata.csv	415,025
X202405_divvy_tripdata	DataFiles/202405-divvy-tripdata.csv	609,493
X202406_divvy_tripdata	DataFiles/202406-divvy-tripdata.csv	710,721
X202407_divvy_tripdata	DataFiles/202407-divvy-tripdata.csv	748,962
X202408_divvy_tripdata	DataFiles/202408-divvy-tripdata.csv	755,639
X202409_divvy_tripdata	DataFiles/202409-divvy-tripdata.csv	821,276
X202410_divvy_tripdata	DataFiles/202410-divvy-tripdata.csv	616,281
X202411_divvy_tripdata	DataFiles/202411-divvy-tripdata.csv	335,075
	Total:	**5,906,269**

5.1.2 Consolidate

For this analysis, the temporary monthly data frames are consolidated into a single dataframe (X12months) containing all 12 months of ride data.

5.1.2.1 Verify imported data rows

A row count of the consolidated dataframe verifies that all 5,906,269 rows were imported.

5.1.2.2 Remove the monthly dataframes

Once the data is consolidated, the temporary monthly dataframes are no longer needed and are removed. This step is taken to free up system resources.

5.2 Data Cleaning and Transformation

5.2.1 Remove duplicate rows

For the analyses to be valid, all duplicate rows are removed from the dataframe. The ride_id field can then be used as the Primary Key for linking with related datasets if they become available.

After de-duplication, 5,906,269 rows are in the dataset indicating there were no duplicate rows.

5.2.2 Convert UTC date-time to Chicago (CT) date-time

The time fields started_at and ended_at are produced by the GPS system and are in Universal Coordinated Time (UTC). The time-based portions of this study will need the date-times converted to Chicago (Central) date-times

To accomplish this, two new variables are added to the dataframe: ride_start and ride_end which will be used to store the calculated Chicago date-times. The calculations adjust for Standard and Daylight Savings time.

5.2.3 Remove early rides

After calculating the new time values, some data rows have ride_start values that are dated prior to the December 01, 2023 starting date of this study. These are early-morning rides now dated 2023-11-30 CT, that were originally date-time-stamped 2023-12-01 UTC. The rows are removed from the dataframe.

5.2.4 Ride times

The GPS system does not record ride tracks. The start and end times refer to the beginning and end of the rental periods, and will provide valuable insight into rider usage. To use this data, the time duration needs to be calculated for each ride-rental from the new ride_start and ride_end variables. A new variable, rental_mins is added to the dataframe to store the rental time duration.

5.2.4.1 Rental period (rental_mins) tests

5.2.4.1.1 Negative rental times

Summarizing the rental periods shows some of the ride times (rental_mins) are negative…

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-2748.32     5.55     9.70    17.32    17.23  1559.93
```

...which reveals invalid data. The rows are filtered from the dataframe.

5.2.4.1.2 Zero rental times

A subsequent data summary shows some ride rental times are zero minutes long.

```
  Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
 0.000    5.550    9.717   17.322   17.233 1559.933
```

This could indicate a glitch in the system. Another possible explanation might be, perhaps for mechanical reasons, such as low tire pressure or damage, some riders elected to immediately return the bicycle to the rack.

When the filtered rows are counted, 132,431 rows have ride times less than one minute long, and are removed from the dataframe.

```
  Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
 1.000    5.783    9.933   17.710   17.517 1559.933
```

5.2.5 Missing station data

In this sample, it is evident that many data rows are missing station names:

ride_id <chr>	start_station_name <chr>	end_station_name <chr>	start_lat <dbl>	start_lng <dbl>	end_lat <dbl>	end_lng <dbl>
C9BD54F578F57246	NA	NA	41.92	-87.66	41.92	-87.66
CDBD92F067FA620E	NA	NA	41.92	-87.66	41.89	-87.64
ABC0858E52CBFC84	NA	NA	41.89	-87.62	41.90	-87.64
F44B6F0E8F76DC90	NA	NA	41.95	-87.65	41.94	-87.65
3C876413281A90DF	NA	NA	41.92	-87.64	41.93	-87.64
28C0D6EFB81E1769	NA	NA	41.91	-87.63	41.88	-87.65

This is caused by poor GPS location data: The latitude and longitude values only have two digits after the decimal point. Five digits are needed to accurately identify the station location:

ride_id <chr>	start_station_name <chr>	end_station_name <chr>	start_lat <dbl>	start_lng <dbl>	end_lat <dbl>	end_lng <dbl>
84BFC1F137684EAB	DuSable Museum	Cottage Grove Ave & 51st St	41.79157	-87.60785	41.80304	-87.60662
EEC92D30A70471E5	California Ave & Division St	California Ave & Division St	41.90303	-87.69747	41.90303	-87.69747
1C33464DEEB1F23C	Chicago State University	Chicago State University	41.71895	-87.60831	41.71896	-87.60830
E0A61810C305E5EC	Cottage Grove Ave & 51st St	Cottage Grove Ave & 51st St	41.80304	-87.60662	41.80304	-87.60662
EB09035006DCCB2C	Chicago State University	Chicago State University	41.71896	-87.60834	41.71896	-87.60830
81EE8687F217E531	California Ave & Milwaukee Ave	California Ave & Milwaukee Ave	41.92269	-87.69715	41.92269	-87.69715

When the dataframe is filtered for the “NA” values in either the `start_station_name` or the `end_station_name` then counted, 1,570,791 rows have this unusable data and are removed from the dataframe.

Solar Activity And GPS
While this dataset does not contain information about Cyclistic’s systems, adverse Space Weather is known to cause reception problems for some GPS receivers. Calendar year 2024 (the year representing the bulk of data used in this analysis) was, or is close to, the peak of the current, eleven-year sunspot cycle (Cycle 25). More information is available at Space Weather and GPS Systems and Solar cycle 25.

5.2.6 Rename member_casual field values

In the member_casual data column, Casual members are denoted by the value "casual". But, annual members are called "members" leading to potential confusion. To clarify, "members" are renamed to "annual" in this study to designate the annual subscribers.

6 Validated dataset

6.1 Description and structure

The dataset, stored as a dataframe named X12months in this report, can be visualized as a large spreadsheet of rows and columns. The first few columns and first six rows look like:

```
ride_id     rideable_type started_at          ended_at            start_station_name start_station_id end_station_name end_station_id start_lat start_lng
<chr>       <chr>         <dttm>              <dttm>              <chr>              <chr>            <chr>            <chr>              <dbl>     <dbl>
84BFC1F137… classic_bike  2023-12-02 23:12:51 2023-12-02 23:21:01 DuSable Museum     KA1503000075     Cottage Grove A… TA1309000067        41.8     -87.6
EEC92D30A7… classic_bike  2023-12-14 13:43:14 2023-12-14 13:44:14 California Ave & … 13256            California Ave … 13256               41.9     -87.7
1C33464DEE… electric_bike 2023-12-04 11:57:04 2023-12-04 12:13:59 Chicago State Uni… 20106            Chicago State U… 20106               41.7     -87.6
E0A61810C3… classic_bike  2023-12-04 09:34:22 2023-12-04 09:35:56 Cottage Grove Ave… TA1309000067     Cottage Grove A… TA1309000067        41.8     -87.6
EB09035006… electric_bike 2023-12-02 06:06:32 2023-12-02 06:09:06 Chicago State Uni… 20106            Chicago State U… 20106               41.7     -87.6
81EE8687F2… classic_bike  2023-12-27 23:55:45 2023-12-28 01:43:13 California Ave & … 13084            California Ave … 13084               41.9     -87.7
i 4,202,501 more rows
i 6 more variables: end_lat <dbl>, end_lng <dbl>, member_casual <chr>, ride_start <dttm>, ride_end <dttm>, rental_mins <dbl>
```

Figure 1: X12months dataframe partial view

6.1.1.1 Rows – Observations

Each row is commonly referred to as an observation, record, or data row. Each row represents one ride-rental. Each ride is identified by the unique ride_id value.

6.1.1.2 Columns – Variables

The ride-rental is described by the values in the column variables.

Each column must have a unique name. Three columns (ride_start, ride_end, and rental_mins) have been added, for ease of calculation, from the original thirteen listed in Table 1.

Table 2: New dataframe columns

Column (variable)	Description
ride_start	Chicago date-time for started_at and adjusted for Daylight Savings Time
ride_end	Chicago date-time for ended_at and adjusted for Daylight Savings Time
rental_mins	Rental time in minutes calculated by: ride_end - ride_start

6.1.2 Data retrieval

Since each row and each column are uniquely identified, any individual data value can be retrieved by specifying the dataframe, row, and column: dataframe[row, column]

For example, the start_station_name value for ride_id = 978 is "University Ave & 57th St".

```
> X12months[978, "start_station_name"]
# A tibble: 1 × 1
  start_station_name
  <chr>
1 University Ave & 57th St
```

Figure 2: An individually selected data value

6.2 General statistical measures

Even though each variable can be summarized, an overall look at common statistics for the entire dataframe is useful:

```
   ride_id          rideable_type        started_at                          ended_at                         start_station_name
 Length:4202507     Length:4202507     Min.   :2023-12-01 06:01:22.00   Min.   :2023-12-01 06:09:43.00   Length:4202507
 Class :character   Class :character   1st Qu.:2024-05-06 17:54:04.50   1st Qu.:2024-05-06 18:10:56.00   Class :character
 Mode  :character   Mode  :character   Median :2024-07-13 10:34:12.68   Median :2024-07-13 10:56:03.93   Mode  :character
                                       Mean   :2024-07-02 14:49:18.01   Mean   :2024-07-02 15:06:05.99
                                       3rd Qu.:2024-09-10 17:00:02.42   3rd Qu.:2024-09-10 17:14:17.13
                                       Max.   :2024-11-30 23:50:53.45   Max.   :2024-11-30 23:57:43.00
 start_station_id   end_station_name   end_station_id       start_lat       start_lng        end_lat         end_lng
 Length:4202507     Length:4202507     Length:4202507     Min.   :41.65   Min.   :-87.86   Min.   :41.65   Min.   :-87.84
 Class :character   Class :character   Class :character   1st Qu.:41.88   1st Qu.:-87.66   1st Qu.:41.88   1st Qu.:-87.66
 Mode  :character   Mode  :character   Mode  :character   Median :41.89   Median :-87.64   Median :41.89   Median :-87.64
                                                          Mean   :41.90   Mean   :-87.64   Mean   :41.90   Mean   :-87.64
                                                          3rd Qu.:41.93   3rd Qu.:-87.63   3rd Qu.:41.93   3rd Qu.:-87.63
                                                          Max.   :42.06   Max.   :-87.53   Max.   :42.06   Max.   :-87.53
 member_casual        ride_start                         ride_end                          rental_mins
 Length:4202507     Min.   :2023-12-01 00:01:22.00   Min.   :2023-12-01 00:09:43.00   Min.   :   1.000
 Class :character   1st Qu.:2024-05-06 12:54:04.50   1st Qu.:2024-05-06 13:10:56.00   1st Qu.:   5.933
 Mode  :character   Median :2024-07-13 05:34:12.00   Median :2024-07-13 05:56:03.00   Median :  10.233
                    Mean   :2024-07-02 09:49:17.69   Mean   :2024-07-02 10:06:05.65   Mean   :  16.799
                    3rd Qu.:2024-09-10 12:00:02.00   3rd Qu.:2024-09-10 12:14:17.00   3rd Qu.:  18.317
                    Max.   :2024-11-30 17:50:53.00   Max.   :2024-11-30 17:57:43.00   Max.   :1509.367
```

Figure 3: Dataframe summary statistics

First, a summary like this is valuable as a "sanity check" that the data is what it is expected to be, and helps to identify anomalous values.

And, there are some oddities in this data to be explored. For example, the first ride-rental (ride_start) in the dataset started at 01:22 AM Chicago time on December 1.

The median (50% of all rentals) value of the rental_mins duration is only a bit longer than 10 minutes and 75% of all rentals only took 18.3 minutes; yet, the last 25% of all rental times extend to over 1509 minutes (over one full day).

The summary gives hints, but does not directly relate the ride data to the category variables. It is those relations that describe how riders are using Cyclistic's bike-share bicycles.

6.2.1 Total rides count

Several categorical variables describe rider usage of the system. The dataframe row count, after removing the unusable rows, is **4,202,507** ride-rentals[2]. That is an excellent sampling despite the large number of unidentified stations rows that were removed during the data-cleaning process.

2 Shown as "Length: 4202507" in Figure 3: Dataframe summary statistics

6.3 Categorical variables

Rider usage can be explored and characterized over several categorical variables. Categorical variables are used to store only a few specific and discrete values.

6.3.1 member_casual

This variable records the rider's membership type for each ride. The focus of this study is understanding the usage patterns between membership type. Cyclistic's objective is development of a marketing plan to promote their annual subscription plan. So, this is the key category variable that all other analyses will be based on. There are only two values:

1. annual (changed from "member" to "annual" in section 5.2.6)
2. casual

6.3.2 start_station_name and end_station_name

Stations can be either start stations or end stations. There are 1,755 distinct stations in the dataframe.

The various station ID, Latitude, and Longitude fields are used by the GPS system to identify the individual bike-share Stations (start and end). This report will use the Station names which are more descriptive and meaningful.

6.3.3 rideable_type

This category describes the bicycles. There are three types in the dataframe.

1. classic_bike
2. electric_bike
3. electric_scooter

6.4 Continuous variables

Continuous variables store character, numerical, or date-time information. They can take on any value and are not limited to a few discrete values like categorical variables are.

6.4.1 rental_mins

This is the ride-rental duration in minutes. rental_mins is a new added variable whose values are calculated by subtracting the ride_start time from the ride_end time.

6.4.2 ride_start and ride_end

These two date-time variables are calculated from the started_at and ended_at variables in the original data files which are expressed in Universal Coordinated Time (UTC). To be meaningful, the date-time information is needed in Chicago calendar dates and Chicago clock time, adjusted for Daylight Savings Time.

6.4.3 Unused variables

Other variables not listed above are redundant or have been converted to new variables. They are not needed for this study. To simplify and streamline calculations, the unused variables are removed from the dataframe.

6.4.4 Simplified dataframe summary

The full dataframe summary now looks like:

```
   ride_id          rideable_type      start_station_name end_station_name   member_casual
 Length:4202507     Length:4202507     Length:4202507     Length:4202507     Length:4202507
 Class :character   Class :character   Class :character   Class :character   Class :character
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character


   ride_start                         ride_end                          rental_mins
 Min.   :2023-12-01 00:01:22.00   Min.   :2023-12-01 00:09:43.00   Min.   :   1.000
 1st Qu.:2024-05-06 12:54:04.50   1st Qu.:2024-05-06 13:10:56.00   1st Qu.:   5.933
 Median :2024-07-13 05:34:12.00   Median :2024-07-13 05:56:03.00   Median :  10.233
 Mean   :2024-07-02 09:49:17.69   Mean   :2024-07-02 10:06:05.65   Mean   :  16.799
 3rd Qu.:2024-09-10 12:00:02.00   3rd Qu.:2024-09-10 12:14:17.00   3rd Qu.:  18.317
 Max.   :2024-11-30 17:50:53.00   Max.   :2024-11-30 17:57:43.00   Max.   :1509.367
```

Figure 4: Simplified data summary after removal of unused variables.

Comparing this summary with that of Figure 3 confirms that all 4,202,507 data rows are retained.

6.5 Limitations

6.5.1 Rental period

The GPS system records start and end times and, for most trips, the start and end stations. The GPS does not log actual travel paths or the travel routes of the bike rides. *The start and end times reflect the rental period not actual ride times.* Individual riders might peddle directly to their destination. Others, may go for a sight-seeing cruise stopping to take pictures, ride for exercise, or pause for coffee and doughnuts.

6.5.2 Ride distance

It is possible to calculate the station to station "distance" from the GPS coordinates. Such calculations are unsuitable for this study for three reasons:

1. The calculation would produce the direct, straight-line, as-the-crow-flies distance. But, Chicago is a large urban environment with a complex road system. For the majority of rides, it is not possible to travel in a straight line from one station to another. Therefore, GPS-based calculations would understate the actual travel distance.
2. These are bicycles. They are not constrained to a specific travel path. Riders are free to choose whatever route that suits them. Chosen routes could be direct, indirect, or simply taking laps around a city block.
3. In fact, as will be shown, some of the most popular trips are round trips; starting and ending at the same station. In these cases, a GPS-based calculation would give a travel distance of zero.

6.5.3 Ride vs. Rental

Despite the limitations mentioned above in section 6.5.2, the time data derived from the GPS fixes is very valuable information to understand how riders use the bicycles. This study uses the words ride, rental, rental-time, ride-time, ride-rental, bike-rental, trips, etc. somewhat interchangeably. The meaning is usually clear from the context, or an explanation provided.

7 Analyses, Data Summaries, and Results

This section begins analyses of the validated ride-rental records.

7.1 Ride-rentals by membership type

This section enumerates ride-rentals by the two membership types: Annual and Casual.

Membership types, levels, or classes are designated as…

- annual
    - Annual members are designated as "member" in the original dataset files. They have been renamed as "annual" to avoid potential confusion.
- casual

Annual members account for almost two-thirds of all ride-rentals as shown in Figure 5

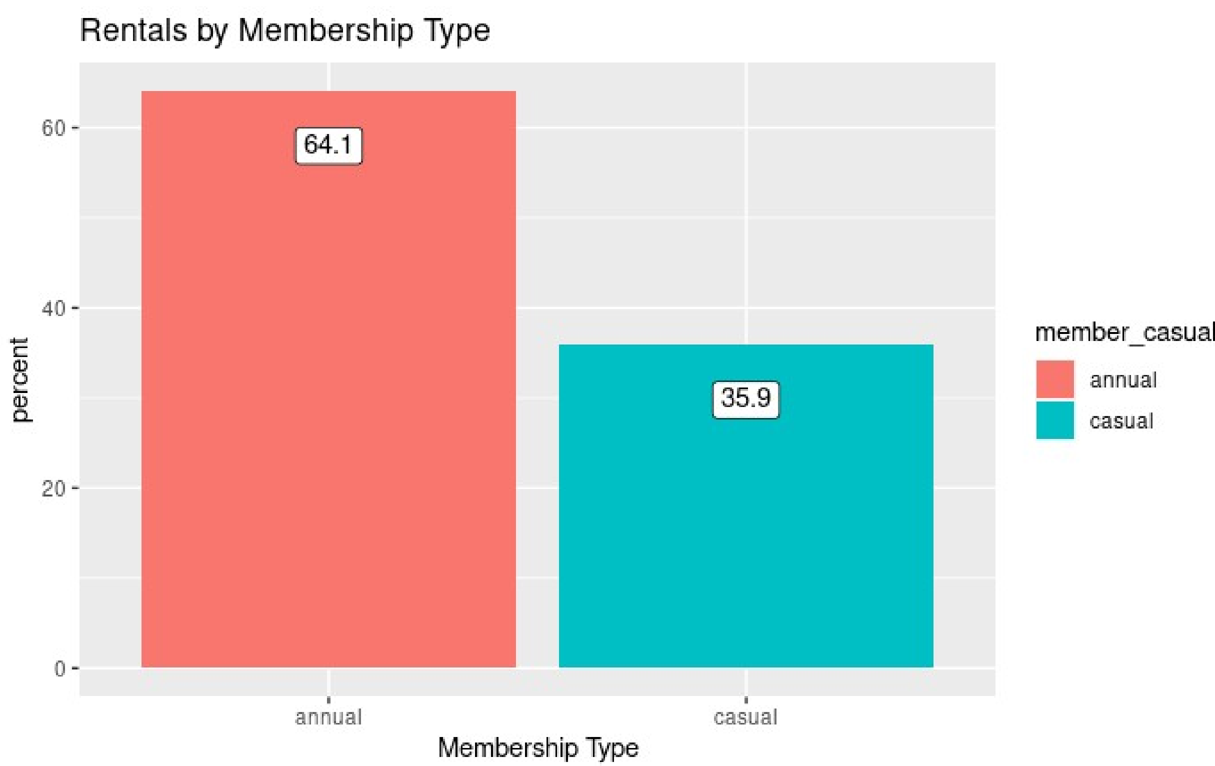


Figure 5: Rentals by membership type

7.1.1.1 Summary

- Annual members account for 64.1 % of bike-rentals.
- Casual members account for 35.9 % of bike-rentals.

7.2 Ride-rental time durations

This section analyzes patterns of member ride-rental time usage.

7.2.1 Rental time durations for all membership types

An overall rental times summary of rental_mins for all membership types provides the general rider usage pattern:

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    1.0     5.9    10.2    16.8    18.3  1509.4
```

Figure 6: Rental summary of all member ship types

Here, the **median** (middle) value of 10.2 minutes is the better measure of a typical rental. This is because the **mean** (average) is adversely affected by a few abnormally high values. This dataset certainly has high values: The maximum rental is 1509.4 minutes (25.2 hours).

The minimum rental is one minute. Ride-rentals of less than one minute were earlier filtered out to eliminate the large number of data rows where the rental time was zero.

The first **quartile** indicates that 25% of all rentals were 5.9 minutes long, or less.

The median value indicates that 50% of all rentals were completed in only 10.2 minutes or less.

The third quartile indicates that 75% of all rentals were completed in 18.3 minutes or less.

The rental times are heavily grouped towards the low end of the range as plotted in Figure 7, page 26.

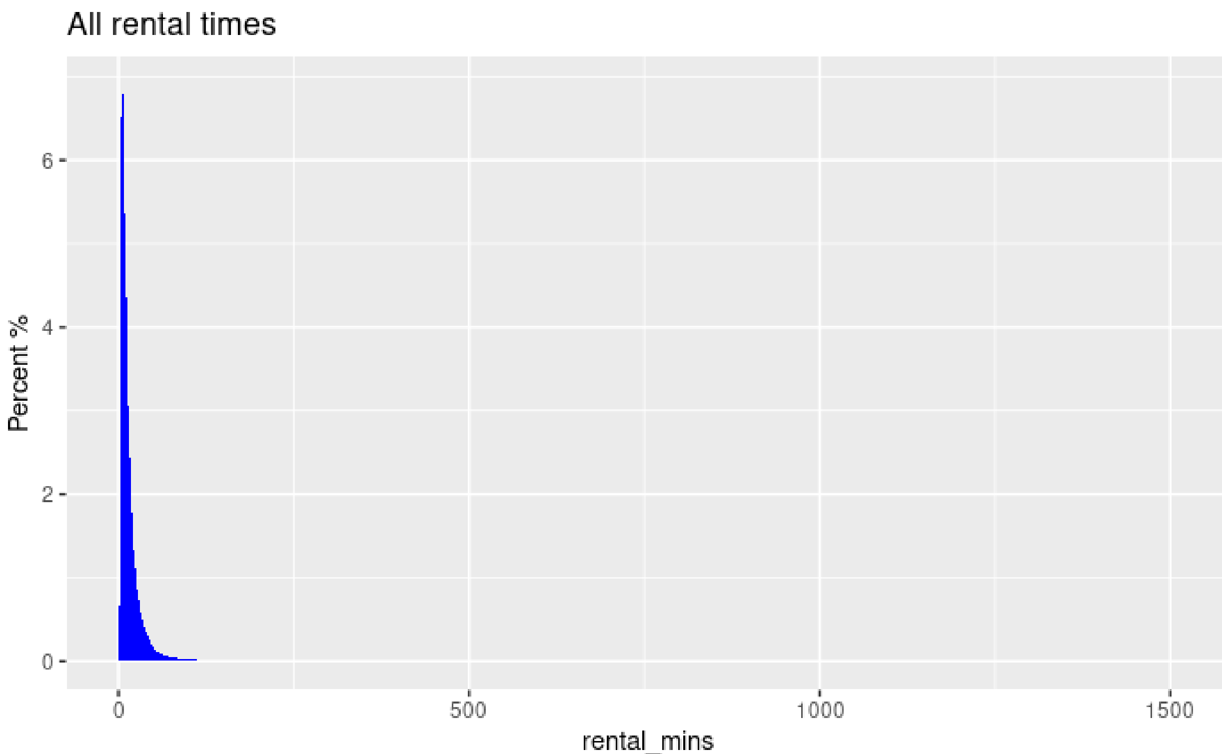


Figure 7: Rental time histogram for all rentals.

As seen in the summary (Figure 6) and the chart (Figure 7) rental times extend well beyond the median value of 10.2 minutes. But, the ride counts for these longer periods are so few that they do not even appear on the chart.

7.2.1.1 Extended rental times

The goal of this study is to understand how the majority of riders use the bicycles. Outlying values can easily distort the results. From inspection of the plot, most rides are less than 180 minutes long.

A conditional count is used to check the actual percentages:

rental_mins <= 180 <lgl>	n <int>	percent <chr>
TRUE	4190467	99.71 %
FALSE	12040	0.29 %

Which confirms that 99.71% of all rentals are 180 minutes or less. The dataframe is then filtered to remove the longer rentals and re summarized, Figure 8, page 27.

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.0     5.9    10.2    15.5    18.2   180.0
```

Figure 8: Summary after removal of rentals greater than 180 minutes

Which shows that the maximum rental is now limited to 180 minutes.

Comparing Figure 8 with Figure 7, notice how the median remained the same at 10.2 minutes, but the mean dropped from 16.8 to 15.5 minutes. This shows how easily a few high values can affect the mean, or average, calculation, and why the median is frequently the preferred measurement.

Additionally, the first and second quartile values are similar to their original values, except an expected slight change, 18.3 to 18.2, in the third quartile.

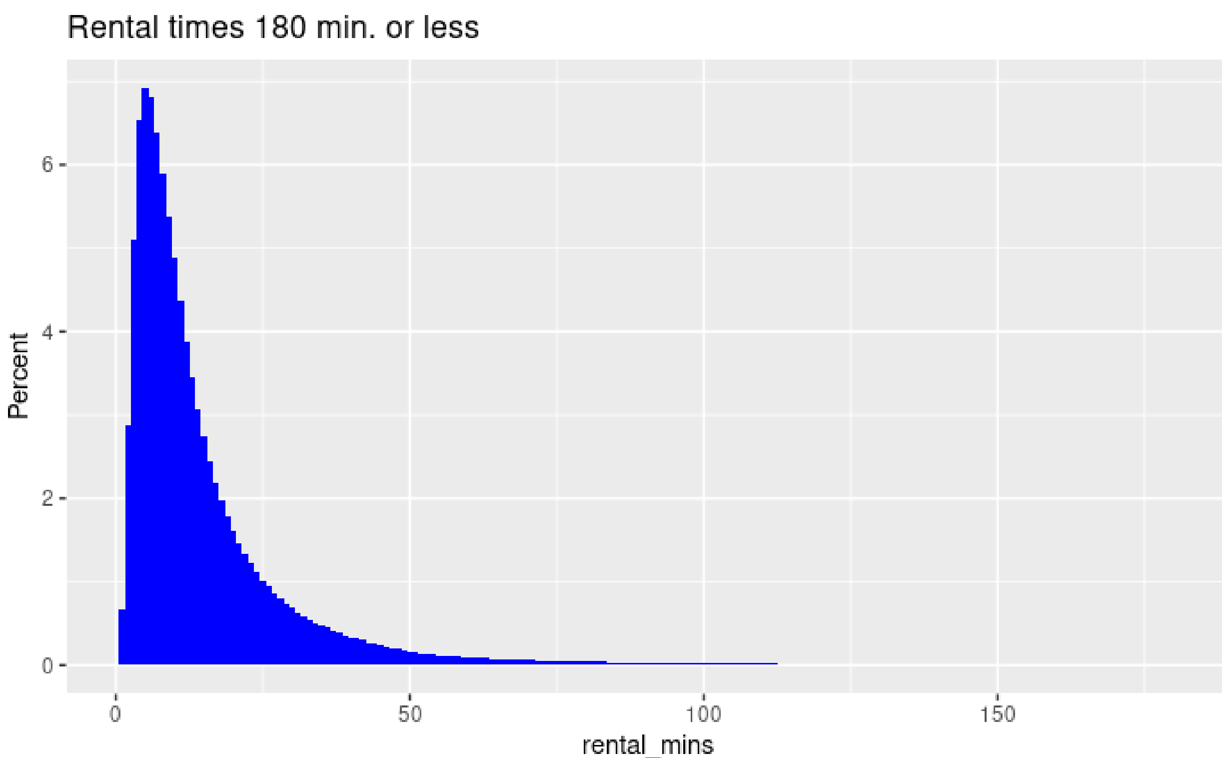


Figure 9: Rental time histogram for rentals <= 180 minutes

The histogram of Figure 7 is replotted as Figure 9 which now shows more detail.

7.2.1.2 Summary

This section has revealed the majority of ride-rentals are relatively short. The median rental time is only 10.2 minutes.

1. 25% of all rides are 5.9 minutes or less
2. 50% of all rides are 10.2 minutes or less

3. 75% of all rides are 18.2 minutes or less
4. 99% of all rides are 180 minutes or less

7.2.2 Comparative rental time durations by membership type

This section compares and analyzes the ride-rental durations by membership type. The analysis reveals significant time differences between the two membership classes.

Similar to the last section, the `rental_mins` time duration distribution graphs of both groups are plotted side-by-side in Figure 10:

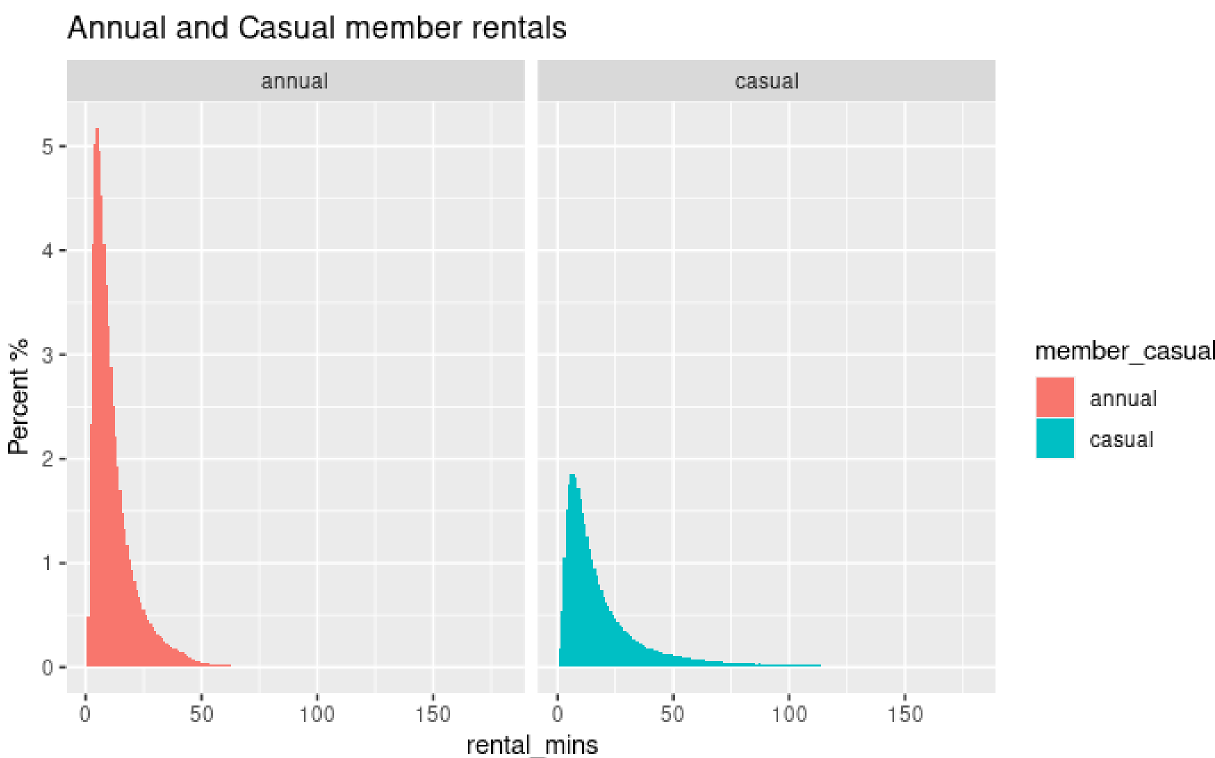


Figure 10: Ride-rental duration distribution plots of Annual and Casual members

As can be seen in these side-by-side charts, Casual riders have generally longer rental times than Annual riders. The following summary table quantifies the differences:

member_casual <chr>	min <dbl>	q25 <dbl>	q50 <dbl>	q75 <dbl>	q99 <dbl>	max <dbl>
annual	1	5.3	8.9	15.1	50.0	180
casual	1	7.6	13.4	25.4	124.8	180

Figure 11: Comparison or Annual and Casual member ride durations by percentile

At every percentile level, ride durations by Casual members are noticeably longer than those of the Annual members.

It's important to examine these differences statistically. An answer is needed to the important question: "Are these real differences between Annual and Casual rider usage, or are the differences caused simply by random variation or chance?"

7.2.2.1 Statistical measures by membership type

A closer look at ride-time differences is in the following table. The median, mean, and standard deviation (SD) are fundamental statistics for comparing groups of data.

member_casual	median	mean	sd
<chr>	<dbl>	<dbl>	<dbl>
annual	8.9	12.1	10.8
casual	13.4	21.5	23.7

Figure 12: Statistical measures of Annual and Casual member rental durations

The Standard Deviation (SD) is a measure of how "spread out" the data is. While the differences look obvious in Figure 11, the SD values in Figure 12 are rather large compared to their respective median values. That indicates a lot of variation within both groups.

That's to be expected in data like this. After all, these are individual riders, with individual destinations, with individual riding abilities, and with individual reasons for renting a bicycle in the first place.

The important question is the difference between the two membership groups just caused by random variation? Or, is the group difference *significant* meaning it's *not* due to just random chance?

7.2.2.2 Analysis of variance (ANOVA) of the means: rental_mins vs membership type

Those questions can be answered by performing an Analysis of Variance (ANOVA) test on the dataset. This calculates a statistical measure called the *p-value*. By generally accepted convention, a p-value of less than 0.05 indicates, to 95% confidence, that the two means *are* statistically different.

The results of that test are shown in Figure 13:

```
        One-way analysis of means (not assuming equal variances)

data:  X12months$rental_mins and X12months$member_casual
F = 215875, num df = 1, denom df = 1862925, p-value < 2.2e-16
```

Figure 13: ANOVA test or rental durations between membership types

The p-value of 2.2e-16 is almost zero, and much lower than the 0.05 threshold. That shows that the two membership groups *are* statistically different. This high confidence is due to the abundance of data, 4,190,467 rides in the dataframe. Very few analysis projects have the luxury of that many observations.

7.2.2.3 Summary

This section identified significant differences between Annual and Casual member ride-rental durations. The rental durations in minutes for each group are:

Table 3: Ride durations by membership type

Member type	Percentile:	25th%	Median (50th%)	75th%	99th%
Annual		5.3	8.9	15.1	50.0
Casual		7.6	13.4	25.4	124.8

Additionally, the ANOVA analysis showed that the duration differences between Annual and Casual members are statistically valid to better than a 95% confidence level.

7.3 Trips and stations usage

Each row of the dataframe equates to one ride-rental. This includes two variables, start_station_name and end_station_name, which can be used to further characterize member usage of Cyclistic's bike-share system.

7.3.1 Total stations and possible trips

The number of distinct stations listed in the dataset is counted then the number of different possible trips (start-station to end-station combinations) is calculated.

stations <int>	possible_trips <dbl>
1755	3080025

7.3.1.1.1 Summary

There are 1,755 stations analyzed in the dataset. The number of possible trips is the number of stations squared. This means there are 3,080,025 possible trip routes including round-trips from and returning to the same station. Reverse trips (where the start and end stations are reversed) are considered as two distinct trips.

7.3.2 Most trips and stations by all members

The most frequent trips as a percentage of all trips in the dataframe. Then, the most frequently used stations for starting, ending and round trips. With the large number of stations and possible combinations, the lists are too long to be printed in this report. Instead, the top most frequent trips or stations are plotted.

7.3.2.1 Most frequent trips by all members

The start_station_name and end_station_name variables are combined into a single trip description. Trips are sorted in descending order by the start station then the end station and plotted as a percent of the total trips count, Figure 14:

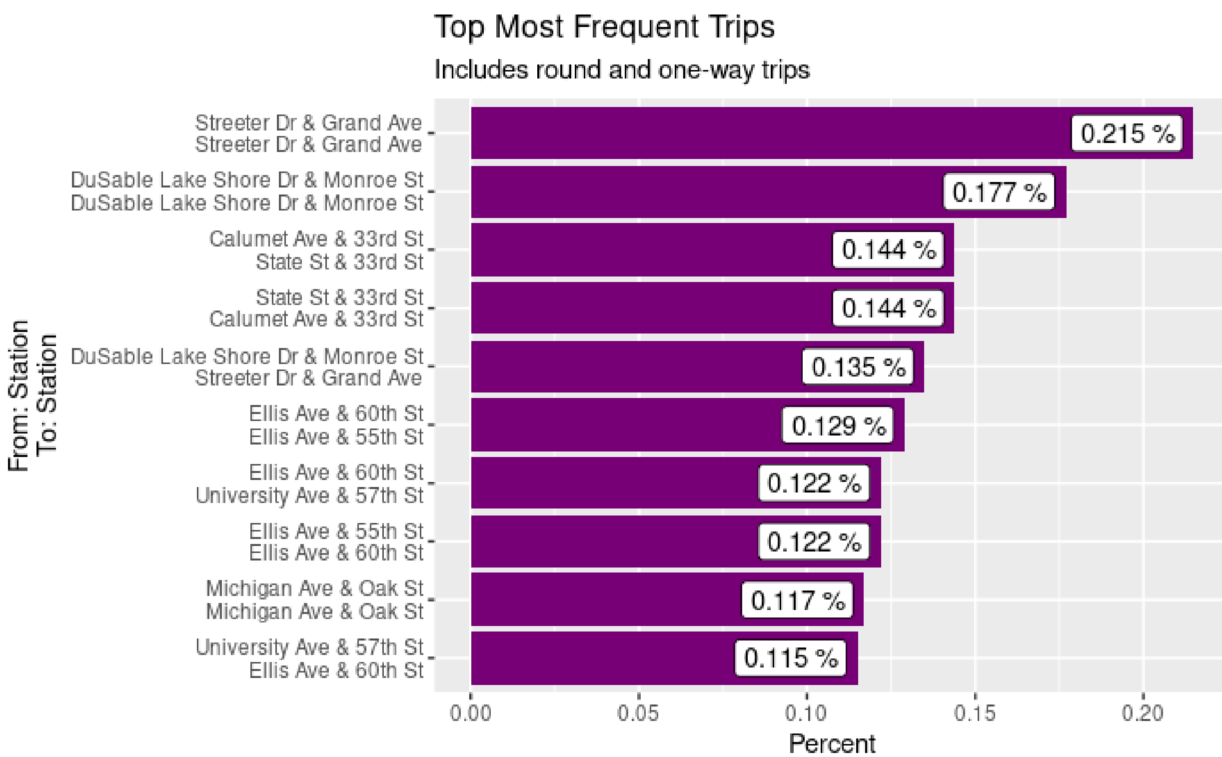


Figure 14: Top most frequent trips (percent of total)

Notice that three of the top ten, and the two most frequent trips are actually round-trips; where riders departed from and returned to the same station. Several remaining trips on this graph are reverse trips pairs. Reverse trips may or may not be return trips and could be by the same or different riders.

7.3.2.1.1 Summary

The two most frequent trips are round-trips; starting and ending at the same station:

- **Streeter Dr and Grand Av** is near several tourist attractions, including the Chicago Children's Museum, Navy Pier, several public parks, and the Ohio Street Beach.
- **DuSable Lake Shore Drive and E Monroe St** is also associated with several large public parks, The Art Institute of Chicago, and the Chicago Yacht Club.

In section 7.3.1 the number of possible routes was calculated as 3,080,025. That is all possible combinations. Not every possible route is ridden. In this compilation, there are a total of 178,927 unique, station-to-station route descriptions.

7.3.2.2 Most frequent round trips by all members

The most frequent round-trip stations are graphed in Figure 15, page 37.

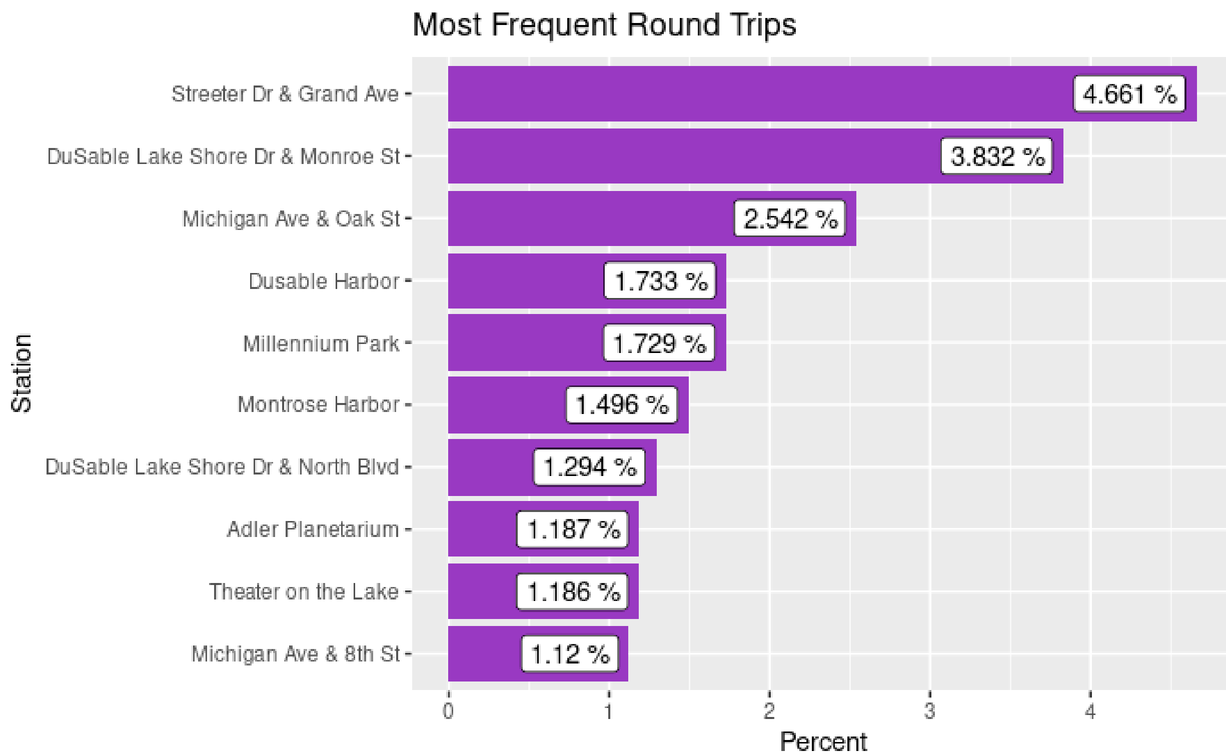


Figure 15: Most frequent round trips

7.3.2.2.1 Summary

DuSable Lake Shore Dr & Monroe St tops this list. Most, if not all, round-trip stations are associated with tourist attractions or public parks.

7.3.2.3 Most frequent one-way trips by all members

The most frequent one-way trips are graphed as Figure 16, page 38

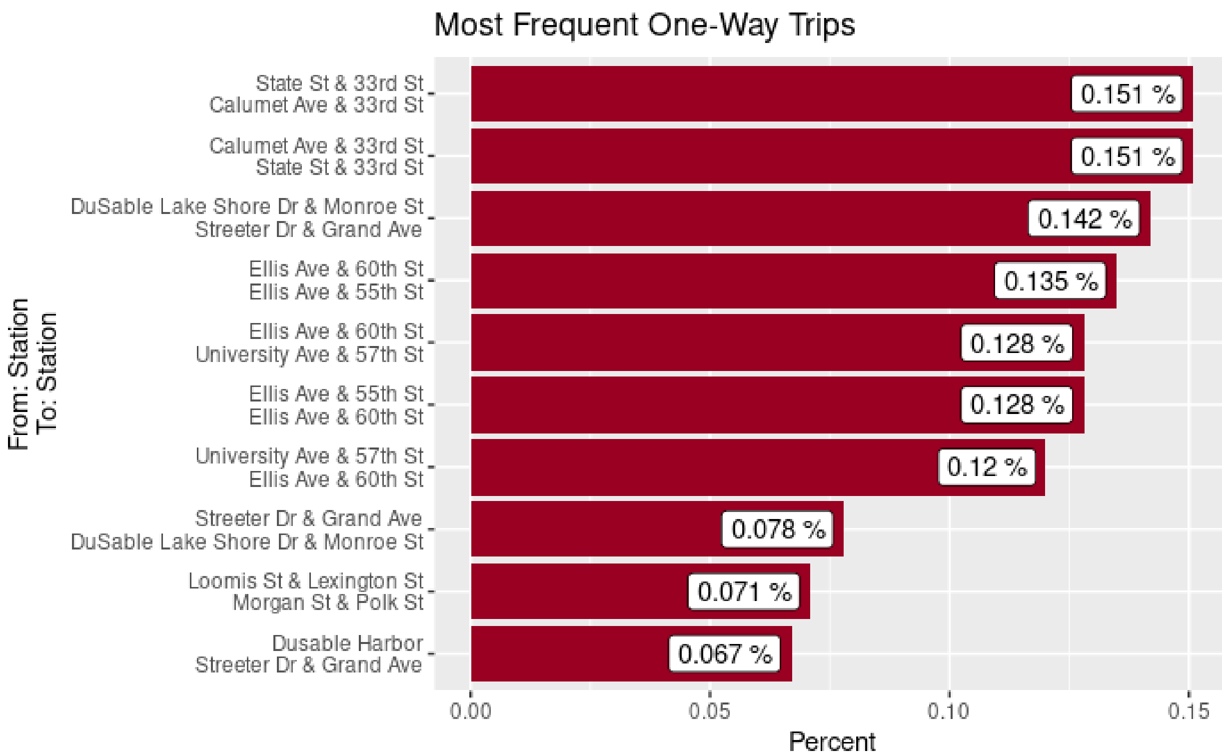


Figure 16: Most frequent one-way trips

7.3.2.3.1 Summary

The first two most frequent one-way trips by all members are a reverse trip pair:

1. State St & 33rd St TO Calumet Ave & 33rd St
2. Calumet Ave & 33rd St TO State St & 33 St

Several other reverse trip pairings appear in this top-ten graph.

7.3.2.4 Most frequent start stations used by all members

The top most frequent start stations are plotted in Figure 17:

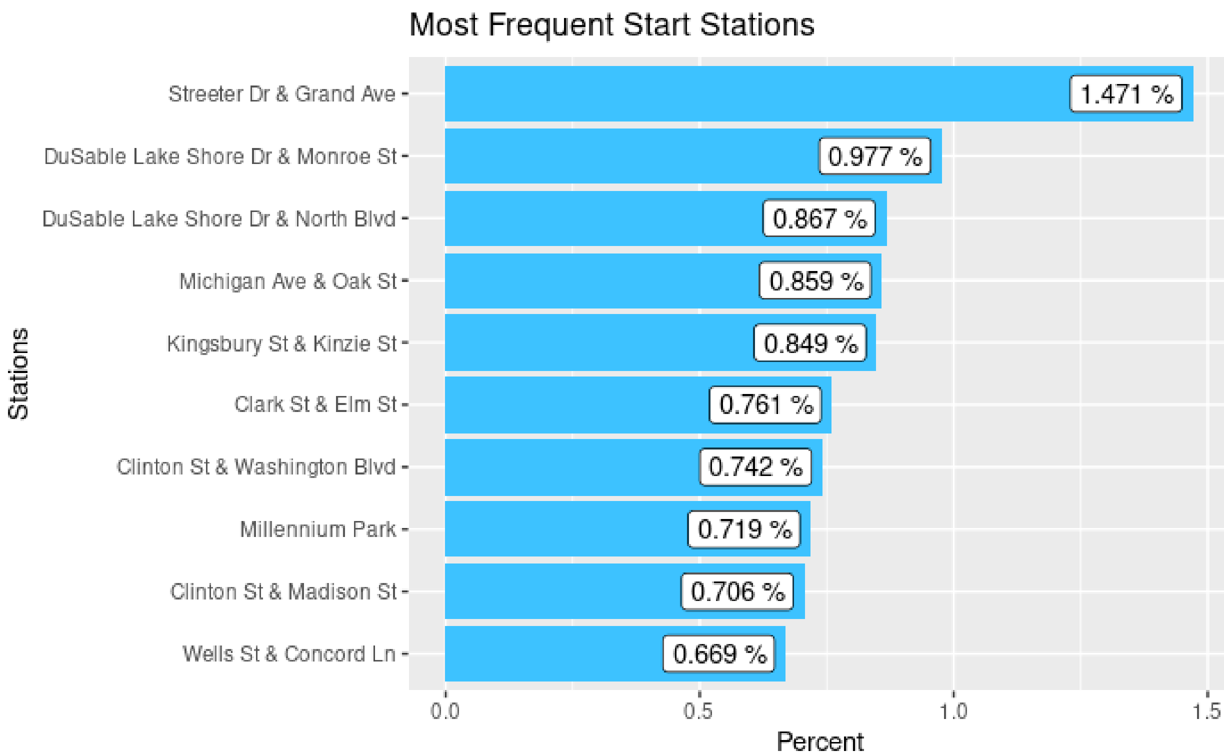


Figure 17: Most frequent start stations

7.3.2.4.1 Summary

Almost 1.5% of all rental-rides in Cyclistic's bike-share system start at the Streeter Dr & Grand Ave station. This is also the most frequent round-trip station (Section 7.3.2.1).

7.3.2.5 Most frequent end stations used by all members

The most frequent end stations are plotted as Figure 18, page 40:

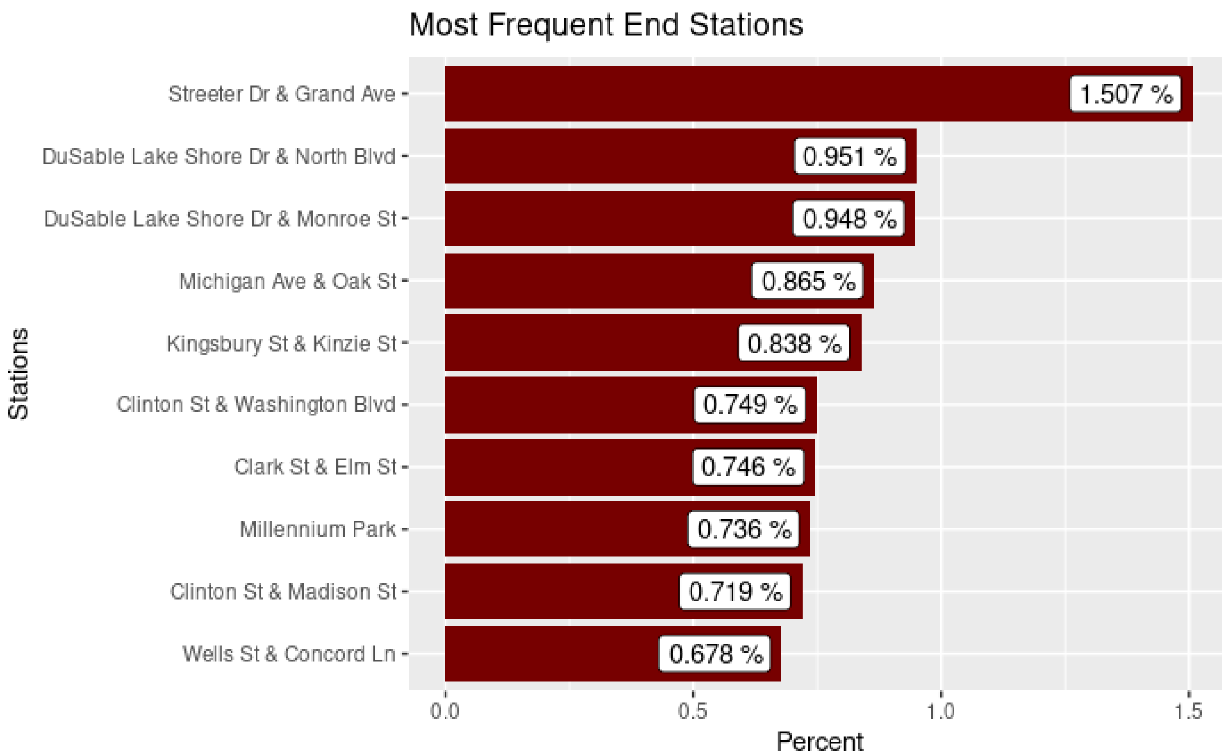


Figure 18: Most frequent end stations

7.3.2.5.1 Summary

The most frequent end station used by all members is Streeter Dr & Grand Ave at 1.507% of all ride-rentals.

7.3.3 Most trips and stations by combined membership types

7.3.3.1 Most frequent trips by combined membership types

The top ten most frequent trips, compiled by membership type, are listed below in decreasing order by percent of total trips starting at 0.199%.

start_station_name <chr>	end_station_name <chr>	member_casual <chr>	percent <dbl>
Streeter Dr & Grand Ave	Streeter Dr & Grand Ave	casual	0.199
DuSable Lake Shore Dr & Monroe St	DuSable Lake Shore Dr & Monroe St	casual	0.163
State St & 33rd St	Calumet Ave & 33rd St	annual	0.137
Calumet Ave & 33rd St	State St & 33rd St	annual	0.135
DuSable Lake Shore Dr & Monroe St	Streeter Dr & Grand Ave	casual	0.125
Michigan Ave & Oak St	Michigan Ave & Oak St	casual	0.105
Ellis Ave & 60th St	Ellis Ave & 55th St	annual	0.094
Ellis Ave & 60th St	University Ave & 57th St	annual	0.094
University Ave & 57th St	Ellis Ave & 60th St	annual	0.093
Ellis Ave & 55th St	Ellis Ave & 60th St	annual	0.091

1-10 of 274,176 rows Previous 1 2

Most frequent trips by membership type are plotted in Figure 19 on page 42. This graph includes all trips: round-trips, one-way trips, and reverse trips.

7.3.3.1.1 Summary

The two most frequent trips are round-trips by Casual riders. Both station locations are in large public parks. Several of these stations will continue to appear in the most frequent stations by combined membership types analyses in the following sections.

This suggests using these locations for local, at-station promotions to encourage Casual to Annual member upgrades.

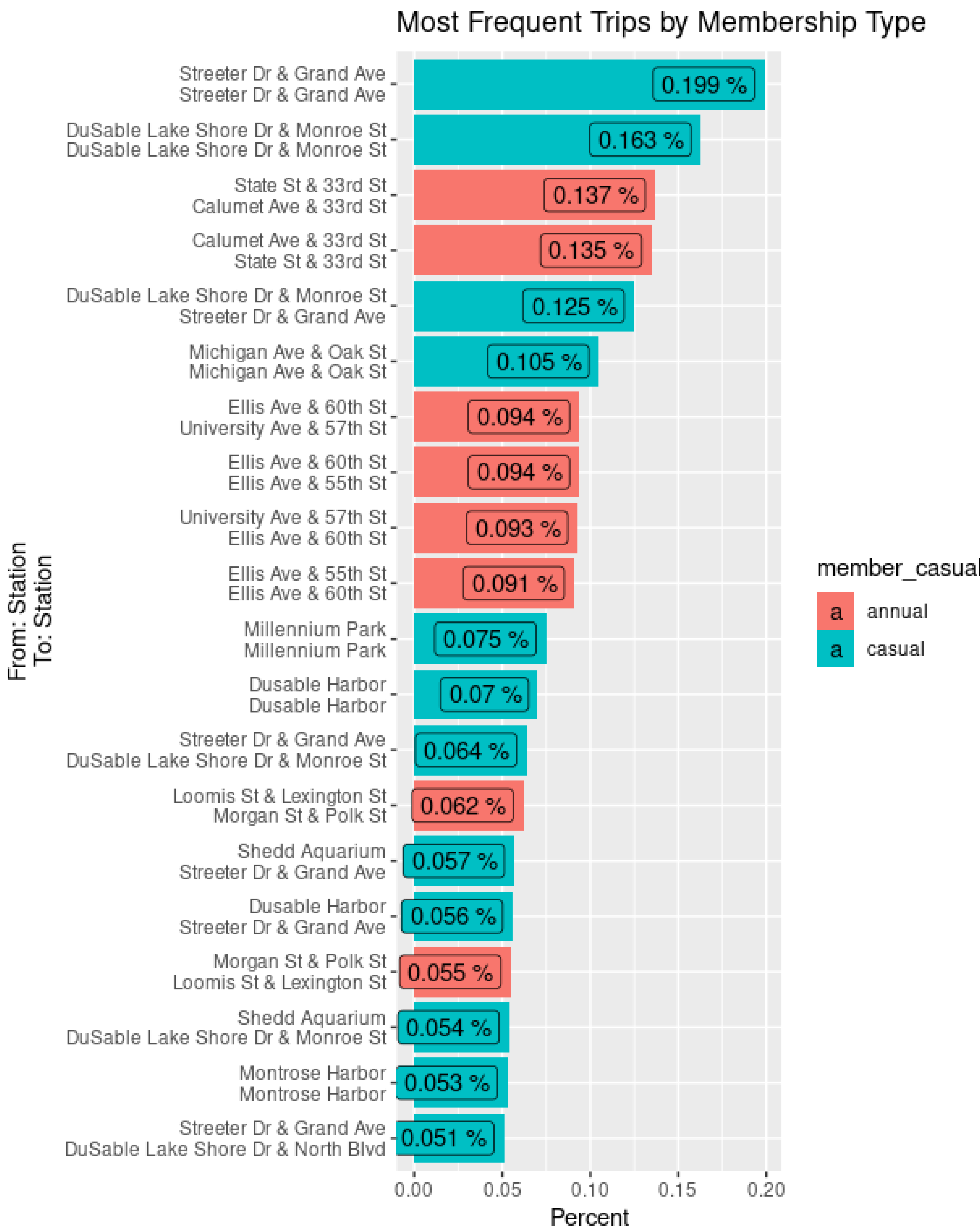


Figure 19: Most frequent trips by membership type

7.3.3.2 Most frequent round-trips by combined membership types

The top ten most frequent round-trips, compiled by membership type, are listed below in decreasing order by percent of total trips starting at 4.308%.

station <chr>	member_casual <chr>	percent <dbl>
Streeter Dr & Grand Ave	casual	4.308
DuSable Lake Shore Dr & Monroe St	casual	3.536
Michigan Ave & Oak St	casual	2.264
Millennium Park	casual	1.615
Dusable Harbor	casual	1.516
Montrose Harbor	casual	1.157
DuSable Lake Shore Dr & North Blvd	casual	1.062
Michigan Ave & 8th St	casual	1.036
Adler Planetarium	casual	1.025
Shedd Aquarium	casual	0.995

The most frequent trips by membership type are plotted in Figure 20 on page 44.

7.3.3.2.1 Summary

All top ten round-trips are by Casual members. That dominance is demonstrated in the graph where 19 of the 20 trips are attributed to Casual members.

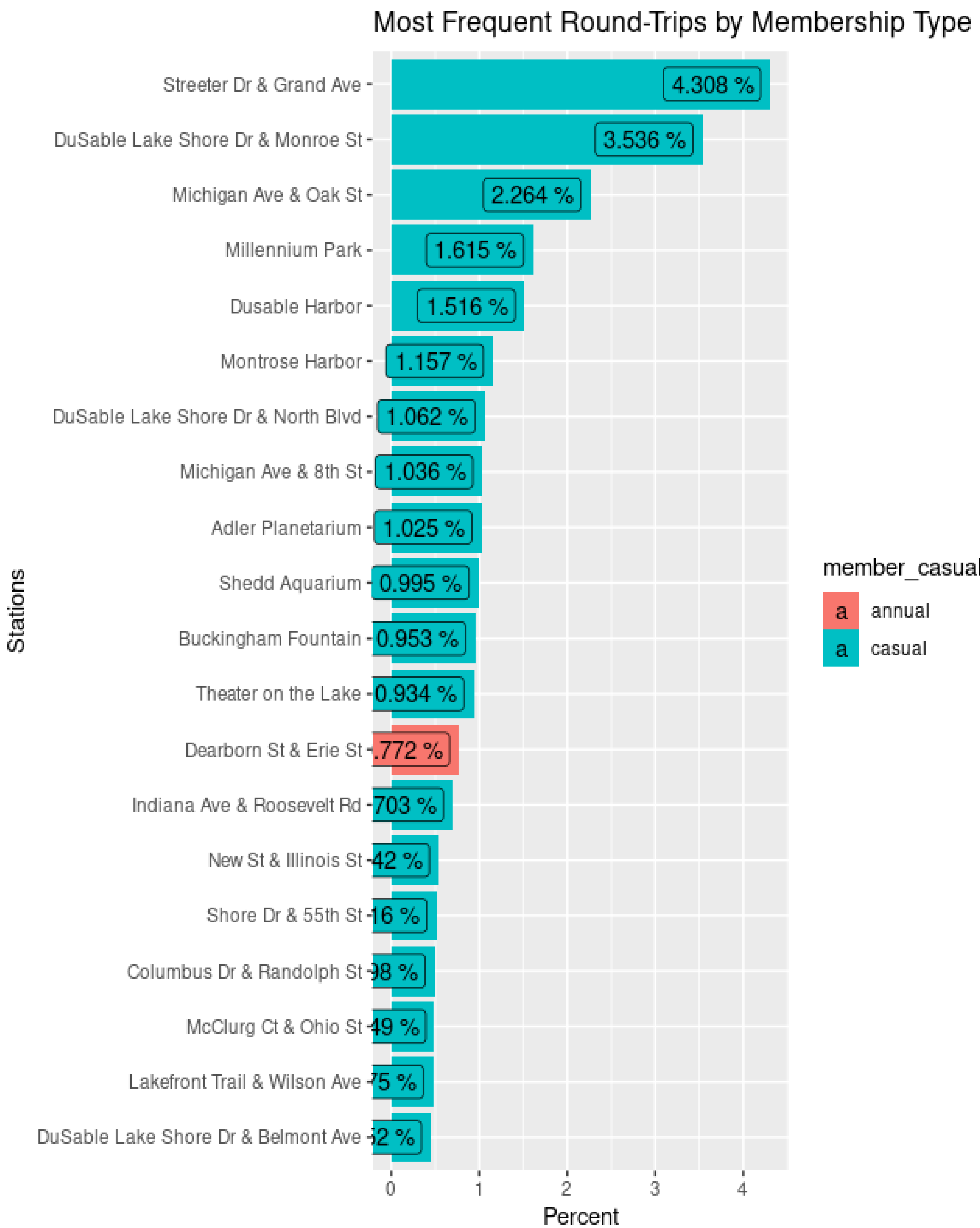


Figure 20: Most frequent round-trips by membership type

7.3.3.3 Most frequent one-way trips by combined membership types

The top ten most frequent one-way trips, compiled by membership type, are listed below in decreasing order by percent of total trips starting at 0.143%.

start_station_name <chr>	end_station_name <chr>	member_casual <chr>	percent <dbl>
State St & 33rd St	Calumet Ave & 33rd St	annual	0.143
Calumet Ave & 33rd St	State St & 33rd St	annual	0.142
DuSable Lake Shore Dr & Monroe St	Streeter Dr & Grand Ave	casual	0.131
Ellis Ave & 60th St	Ellis Ave & 55th St	annual	0.099
Ellis Ave & 60th St	University Ave & 57th St	annual	0.098
University Ave & 57th St	Ellis Ave & 60th St	annual	0.098
Ellis Ave & 55th St	Ellis Ave & 60th St	annual	0.095
Streeter Dr & Grand Ave	DuSable Lake Shore Dr & Monroe St	casual	0.067
Loomis St & Lexington St	Morgan St & Polk St	annual	0.065
Shedd Aquarium	Streeter Dr & Grand Ave	casual	0.060

The most frequent trips by membership type are plotted in Figure 21 on page 46.

7.3.3.3.1 Summary

For the most frequent one-way trips, Annual members represent seven out of the top ten positions. For Casual member trips, the station locations tend to be associated with tourist attractions.

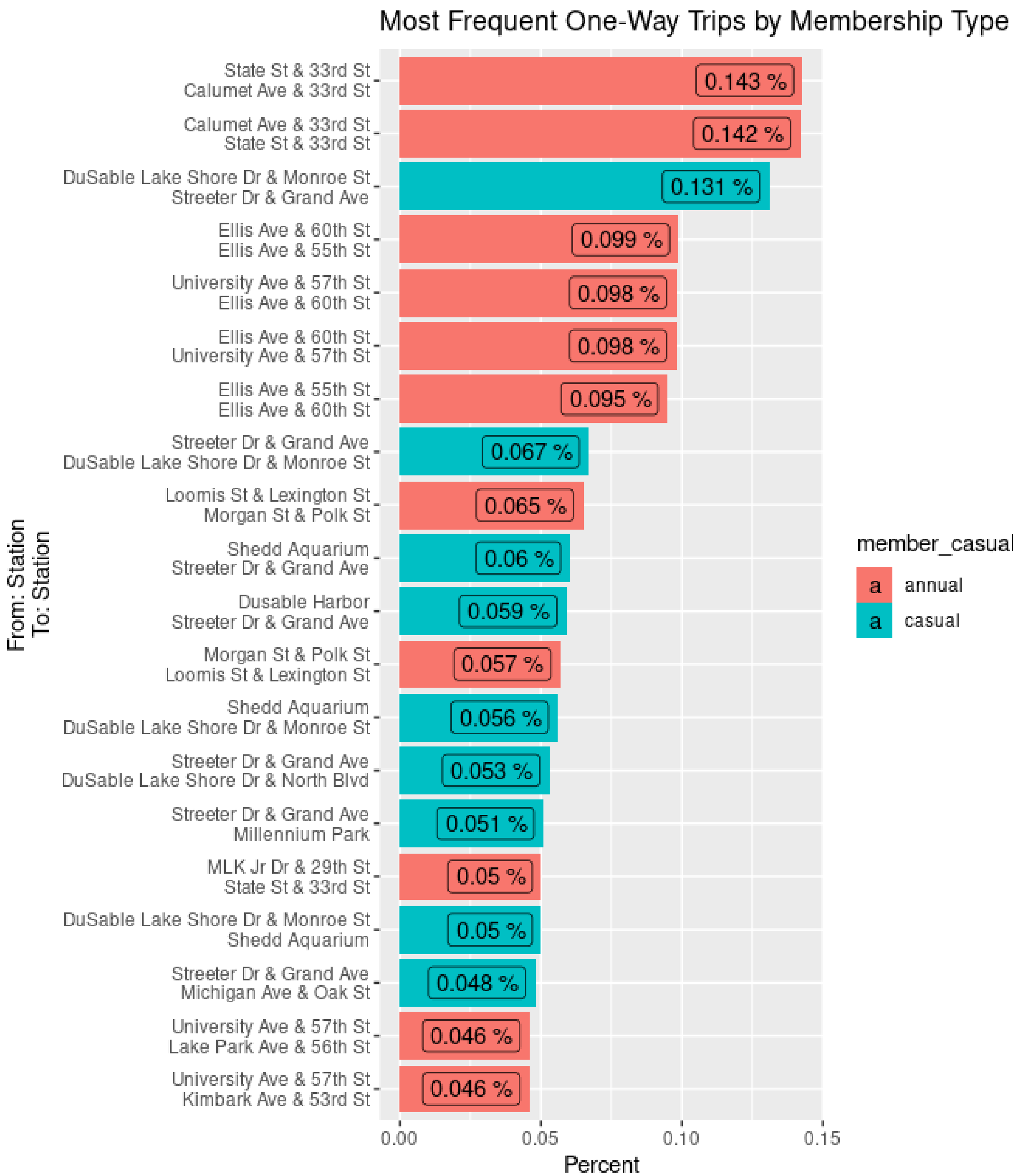


Figure 21: Most frequent one-way trips by membership type

7.3.3.4 Most frequent start stations by combined membership types

The top ten most frequent start stations, compiled by membership type, are listed below in decreasing order by percent of total trips starting at 1.137%.

start_station_name <chr>	member_casual <chr>	percent <dbl>
Streeter Dr & Grand Ave	casual	1.137
DuSable Lake Shore Dr & Monroe St	casual	0.756
Kingsbury St & Kinzie St	annual	0.633
Clinton St & Washington Blvd	annual	0.599
Michigan Ave & Oak St	casual	0.549
Clinton St & Madison St	annual	0.536
Clark St & Elm St	annual	0.533
DuSable Lake Shore Dr & North Blvd	casual	0.504
Millennium Park	casual	0.489
Shedd Aquarium	casual	0.475

1-10 of 3,211 rows

The top most frequent start stations by membership type are plotted in Figure 22 on page 48.

7.3.3.4.1 Summary

Six of the top ten most frequent start stations are attributed to Casual members. This suggests using these locations for local, at-station promotions to encourage Casual to Annual member upgrades.

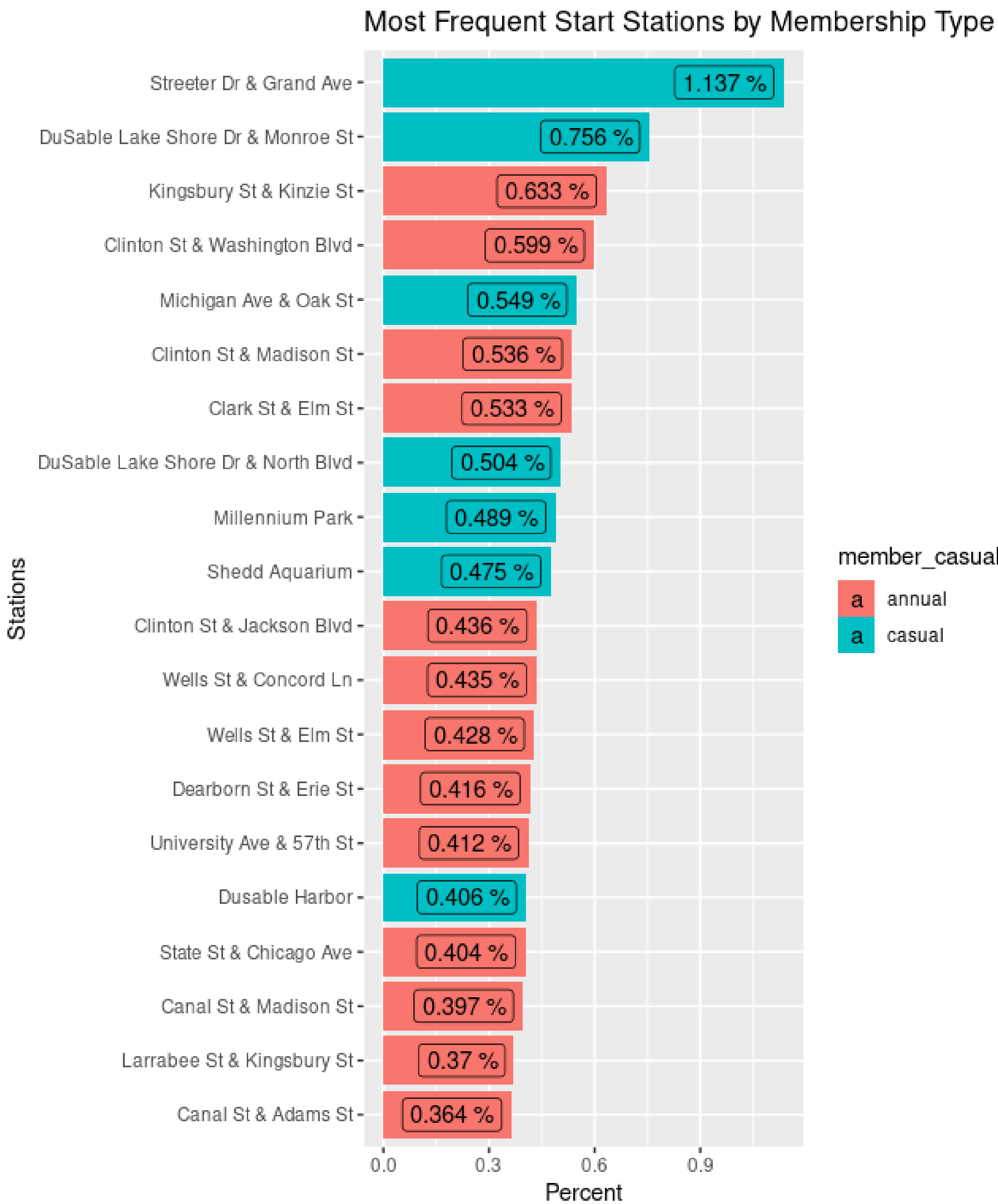


Figure 22: Most frequent start stations by membership type

7.3.3.5 Most frequent end stations by combined membership types

The top ten most frequent end stations, compiled by membership type, are listed below in decreasing order by percent of total trips starting at 1.235%.

end_station_name <chr>	member_casual <chr>	percent <dbl>
Streeter Dr & Grand Ave	casual	1.235
DuSable Lake Shore Dr & Monroe St	casual	0.707
Kingsbury St & Kinzie St	annual	0.636
Clinton St & Washington Blvd	annual	0.609
DuSable Lake Shore Dr & North Blvd	casual	0.593
Michigan Ave & Oak St	casual	0.570
Clinton St & Madison St	annual	0.556
Millennium Park	casual	0.539
Clark St & Elm St	annual	0.530
Wells St & Concord Ln	annual	0.438

1-10 of 3,260 rows

The top most frequent start stations by membership type are plotted in Figure 23 on page 50.

7.3.3.5.1 Summary

Five of the top ten end stations are attributed to casual members.

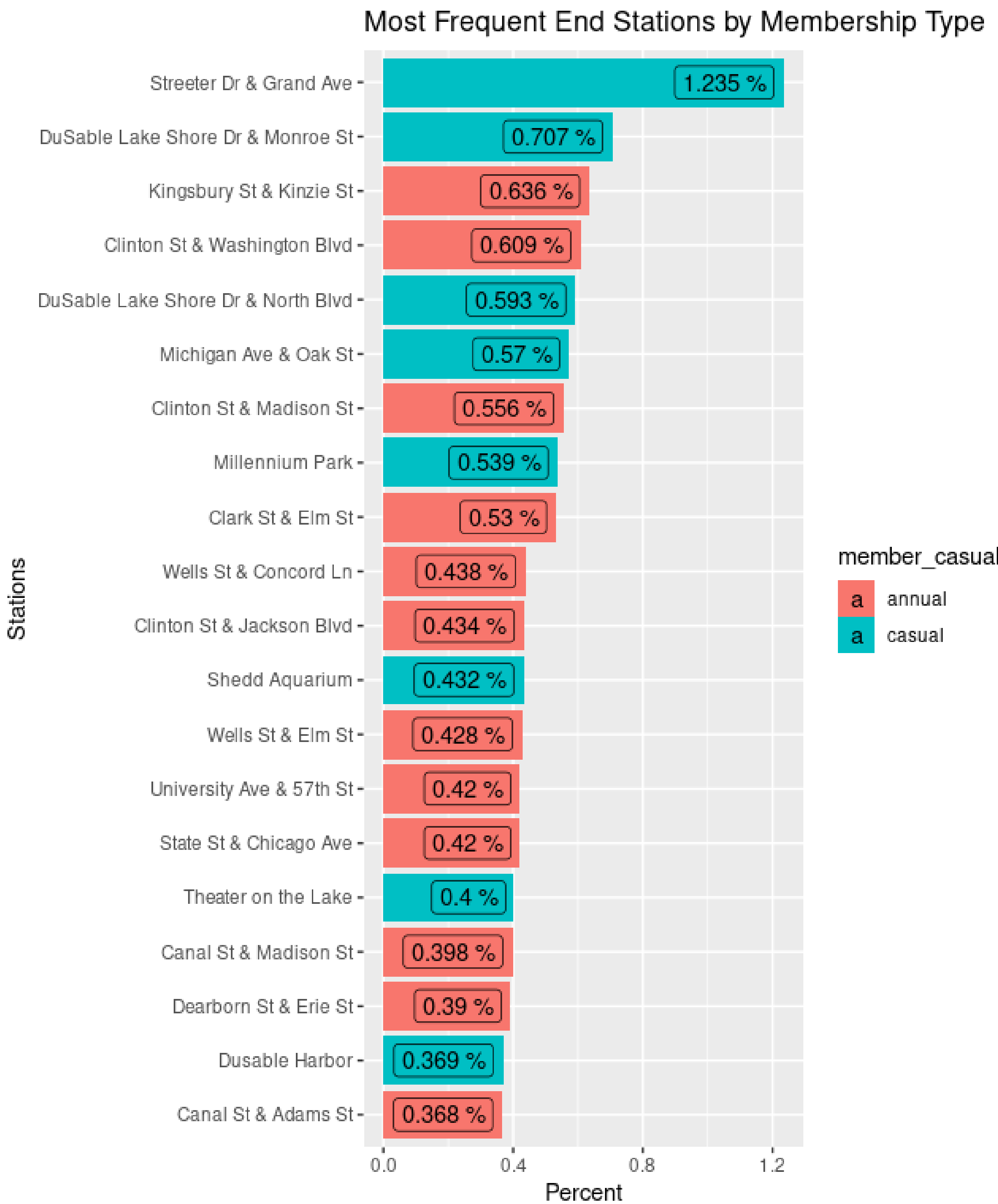


Figure 23: Most frequent end stations by membership type

7.4 Bicycle types

Cyclistic has several types of bicycles for riders to rent. This section will analyze the trip dataset to determine rider usage patterns and preferences.

7.4.1 Bicycle type preference by membership type

Figure 24 compares bicycle preference by membership types.

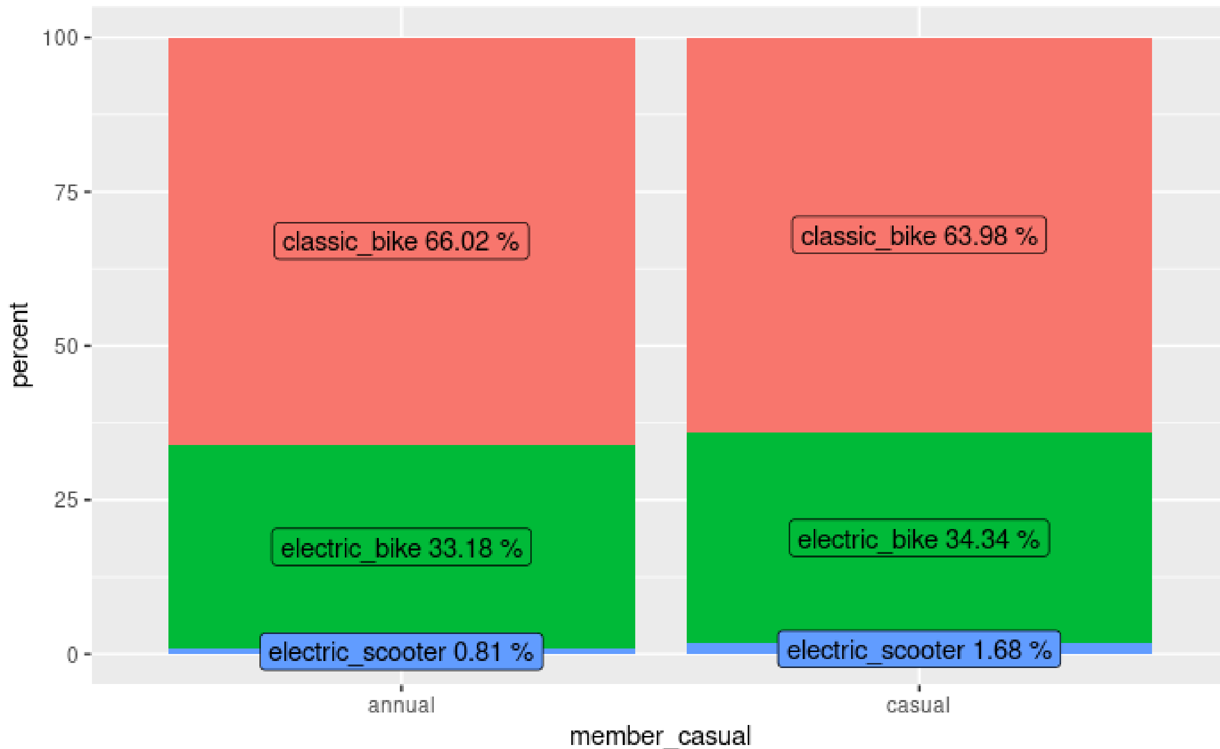


Figure 24: Bicycle type preferences by membership types

Other than the electric scooters, which Casual riders use twice as much, the proportions for electric bicycles (33.18% and 34.34%) are very similar. It appears the difference is easily just due to random chance.

It is not.

7.4.1.1 Testing for electric bike and scooters preference

This statistical test uses the **2-sample test for equality of proportions**. The test is run against both the electric bike and electric scooters proportions to ascertain if they are statistically significant; meaning not due to random variation.

The 2-sample test is run twice; once each for electric bicycles and electric scooters. The results appear in Figure 25, page 54. The test produces a statistical measure called the p-value. By convention, a value less than 0.05 is generally accepted as indicating statistical validity; meaning the differences are *not* due to random variation.

In both cases, the p-value obtained ($p < 2.2e-16$) is almost zero and much less than the 0.05 threshold indicating the proportions are valid to a 95% confidence level. This is due to very high number of observations, 4,190,467 rides, in the dataset.

7.4.1.1.1 Summary

The data shows Casual members have a preference for the electric ride options:

1. **Electric bicycles**: Casual members have a slightly higher, but statistically significant, preference for electric bicycles over Annual members.
    1. 34.34% of all Casual member ride-rentals use electric bicycles.
    2. 33.18% of all Annual member ride-rentals use electric bicycles.
2. **Electric scooters**: Casual members select electric scooters more than 2:1 over Annual members.
    1. 1.68% of all Casual member ride-rentals use electric scooters.
    2. 0.81% of all Annual member ride-rentals use electric scooters

The 2-sample test, Figure 25, shows both sets of rider preference proportions are statistically significant and are *not* due to random chance to better than 95% confidence.

This suggests some type of remuneration on one or both electric choices, such as a special first-ride discount for an electric bike/scooter rental, would be a positive inducement to sign up for Annual membership.

```
[1] "Electric bike"

        2-sample test for equality of proportions with continuity correction

data:  c(890922, 516825) out of c(n_annual, n_casual)
X-squared = 585.2, df = 1, p-value < 2.2e-16
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.01258009 -0.01068946
sample estimates:
   prop 1    prop 2 
0.3317616 0.3433964 

[1] "Electric scooter"

        2-sample test for equality of proportions with continuity correction

data:  c(21677, 25278) out of c(n_annual, n_casual)
X-squared = 6623.7, df = 1, p-value < 2.2e-16
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.008955535 -0.008491453
sample estimates:
     prop 1      prop 2 
0.008072084 0.016795578 
```

Figure 25: Electric bike and electric scooter proportions testing

7.4.2 Rental times by bicycle types

7.4.2.1 Average rental time by bicycle type (all member types)

As an overall baseline measurement, the average rental times, in minutes, by bicycle type is compiled from the ride dataset. The differences are shown, for both membership types, in Figure 26.

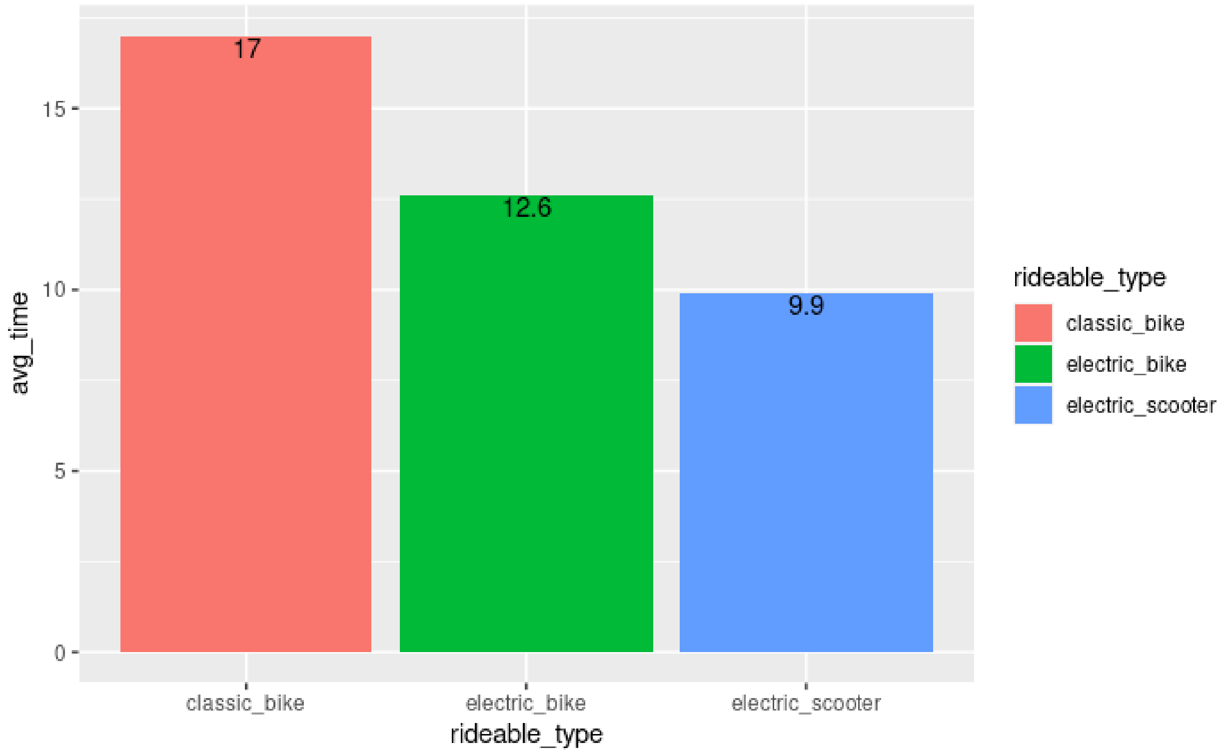


Figure 26: Average rental times by bicycle type for all member types.

Table 4: Time savings of electric bicycles and electric scooters over classic bicycles

Type	Avg. Time (min.)	Time Svgs.%
Classic	17	
Electric	12.6	25.9%
Scooter	9.9	41.8%

7.4.2.1.1 Summary

Overall ride-rental time durations are notably shorter for electric bicycles and electric scooters in the bike-share fleet. Rental time savings over classic bicycles are considerable: 25.9% for electric bicycles and 41.8% for electric scooters. A marketing advantage to be exploited.

7.4.2.2 Average rental time by bicycle type by membership type

The average rental times by bicycle types (Section 7.4.2.1) are extended to differentiate by the Annual and Casual membership types. The split chart shows clear differences:

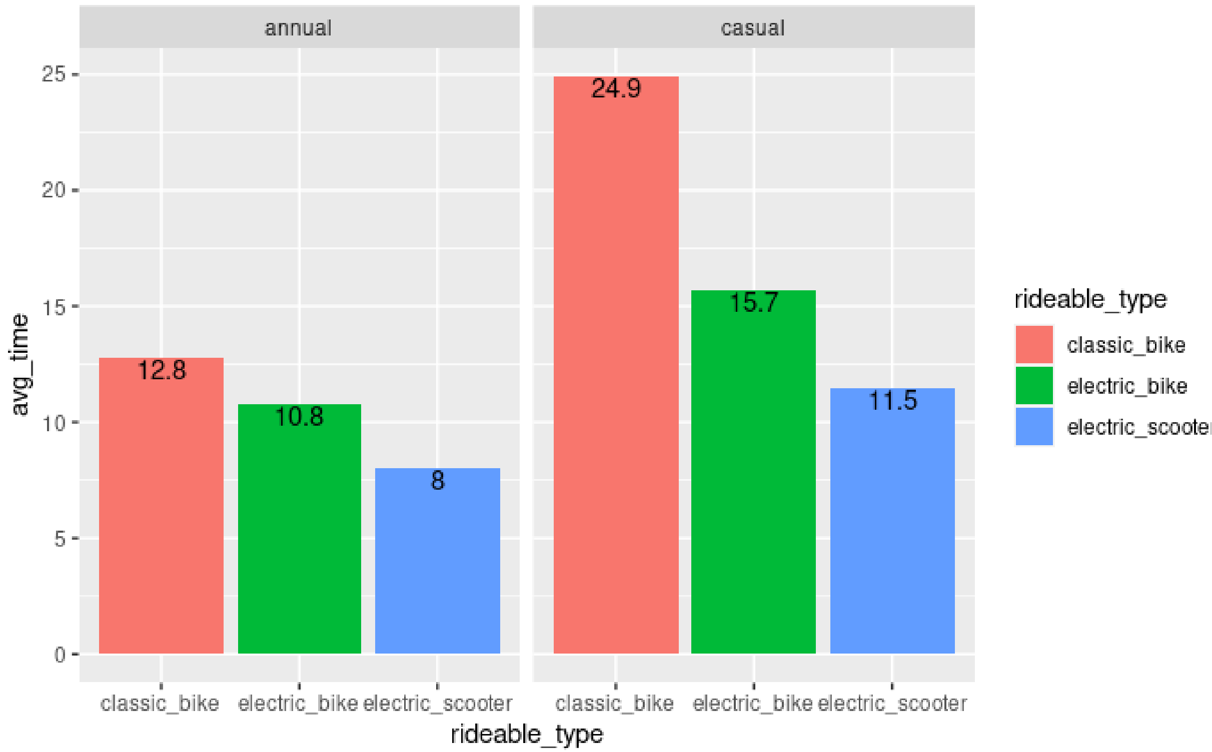


Figure 27: Average rental times by bicycle type by membership type

7.4.2.2.1 Summary

Overall average rental times of Casual members, by bicycle type, are distinctly longer than those of Annual members: 95% longer for classic bicycles, 45% longer for electric bicycles, and 44% longer for electric scooters.

7.5 Date and Time Usage

Future promotions for the member upgrade strategy will need to be scheduled when they are most effective.

This section presents several analyses of time-based rider patterns: Month, Weekday, and Hour.

7.5.1 Rides by Month

7.5.1.1 Rides by month - all members

All members ride-rentals by month are shown in Figure 28

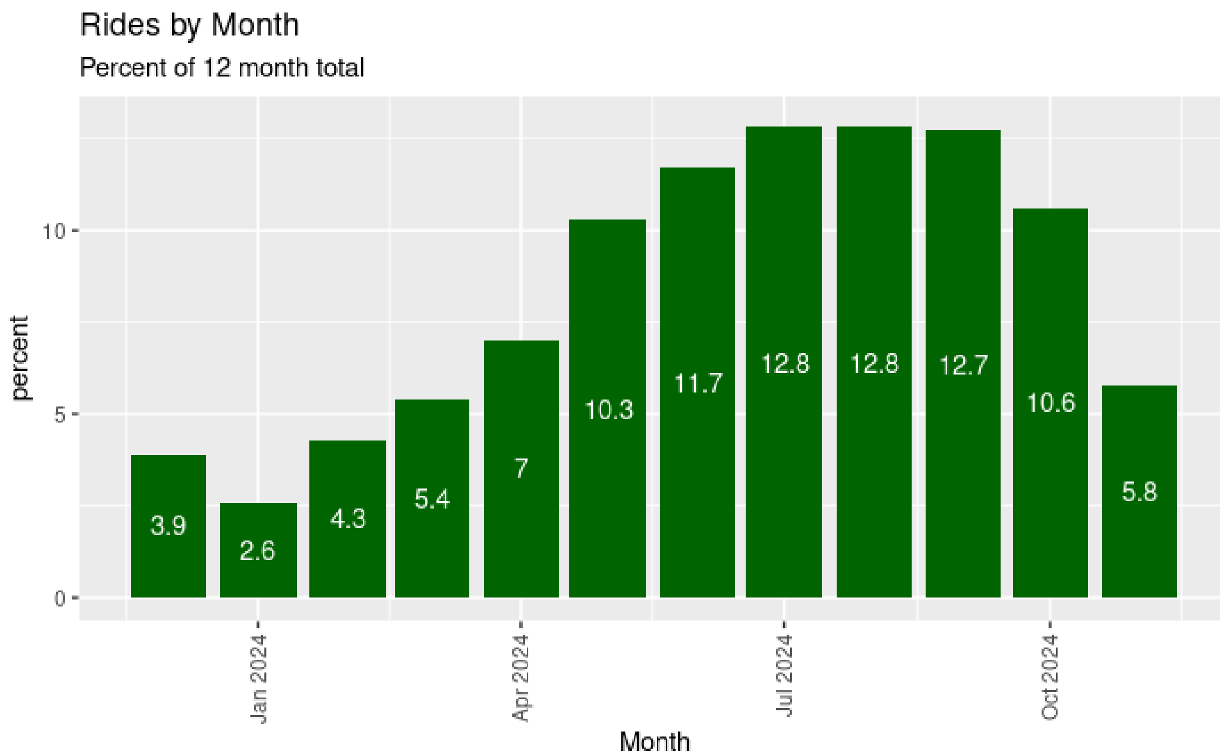


Figure 28: Rentals by month - all membership types

7.5.1.1.1 Summary

The busiest months are July and August at 12.8% each. January is the least busy month at 2.6%. A ratio of about 4.9 to 1.

7.5.1.2 Rides by month - Annual members

Rides by month for Annual members are shown in Figure 29.

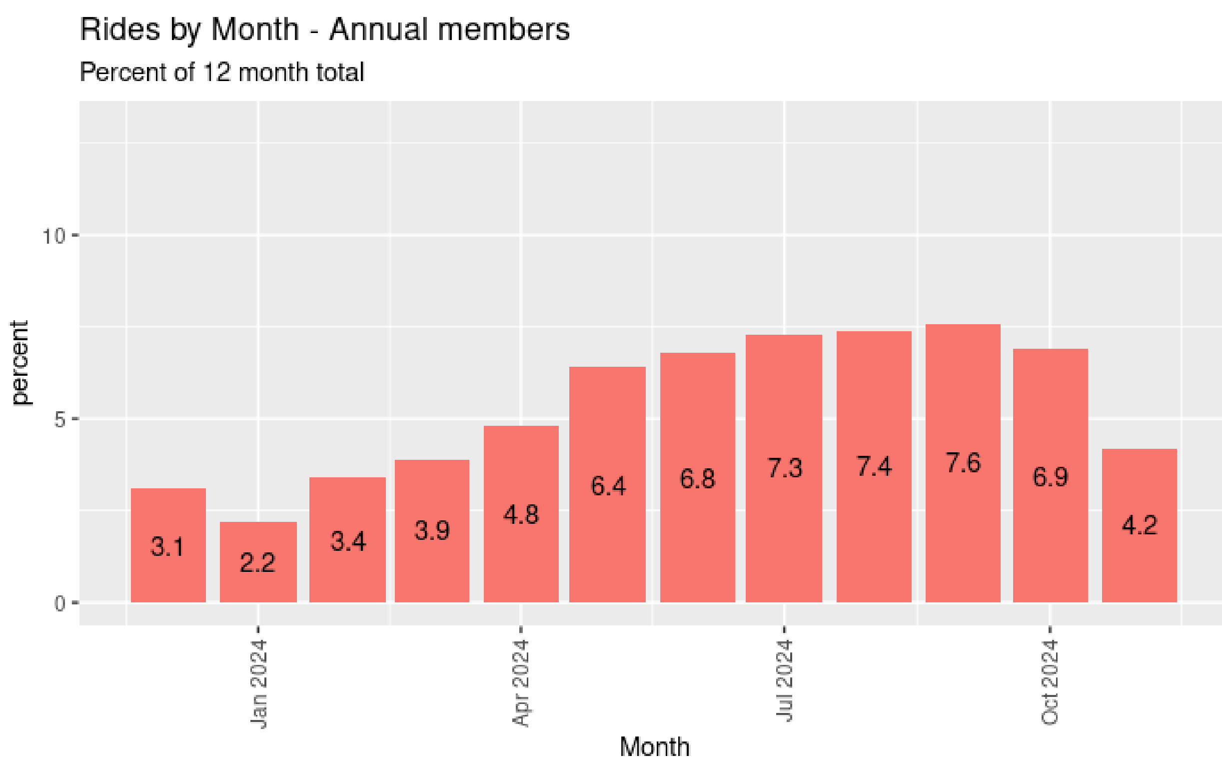


Figure 29: Rides by month - Annual members

7.5.1.2.1 Summary

September is the heaviest use month for Annual members at 7.6% of all rides. January reflects the lowest participation at 2.2 %. A ratio of about 3.4 to 1.

7.5.1.3 Rides by month - Casual members

Rides by month for Annual members are shown in Figure 30

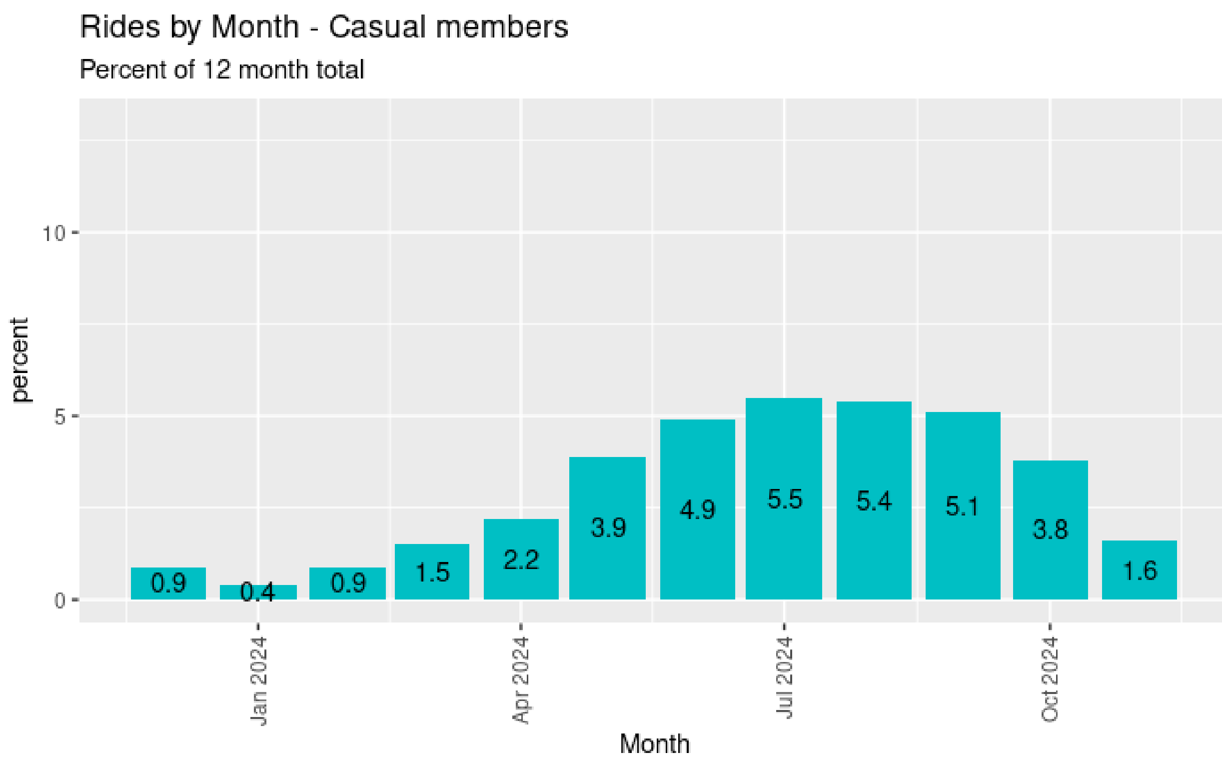


Figure 30: Rides by month - Casual members

7.5.1.3.1 Summary

July is the heaviest use month for Casual members at 5.5% of all rides. January reflects the lowest participation at 0.4 %. A ratio of about 13.8 to 1. A much sharper high-to-low ratio than that of Annual members.

7.5.1.4 Rides by month - combined member types

Rides by month for combined member types are shown in Figure 31.

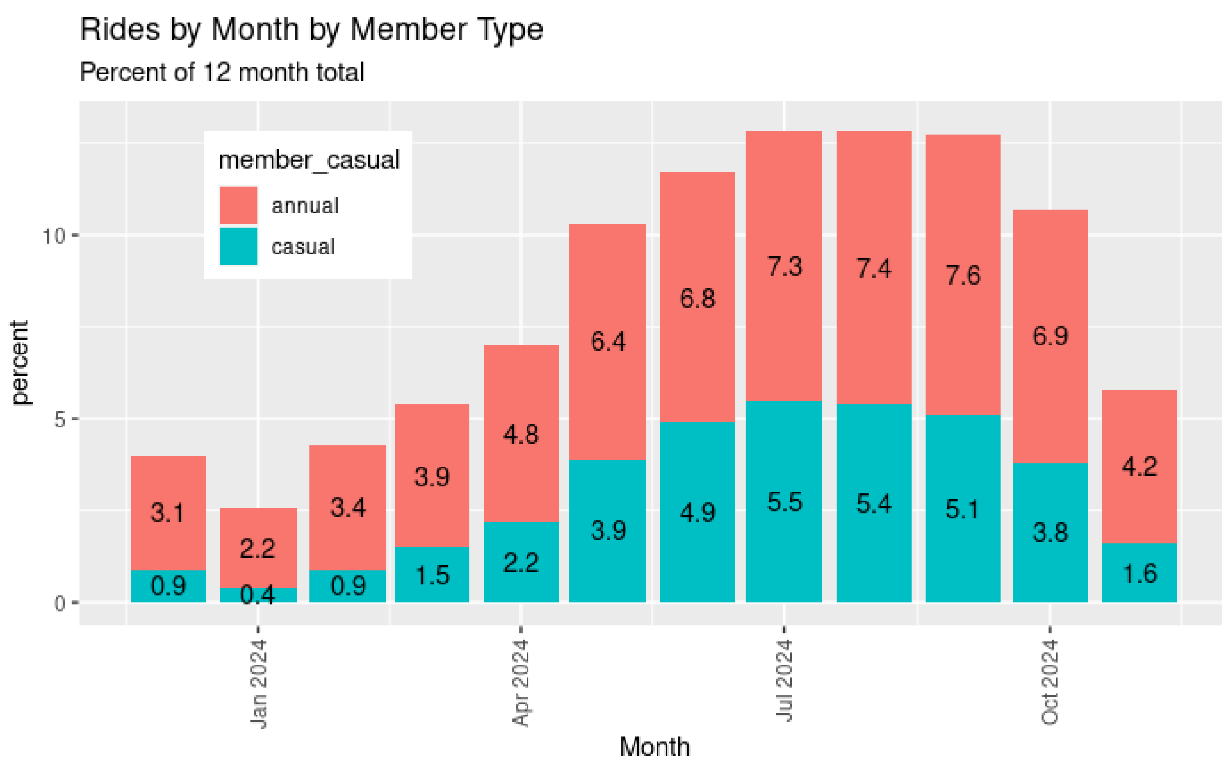


Figure 31: Rides by month - combined member types

This plot combines the plots of both member types (from sections 7.5.1.2 and 7.5.1.2) into the same graph to illustrate the relative proportions of ride-rentals by month for the entire twelve month time frame.

7.5.1.4.1 Summary

The highest ride-rental participation of both groups takes place during the warmer months. January is the most one-sided with Annual riders outnumbering Casual riders 5.5 to 1. For every month, Annual riders always outnumber Casual riders.

7.5.2 Rides by Weekday

7.5.2.1 Rides by weekday (all members)

All members ride-rentals by weekday are shown in Figure 32.

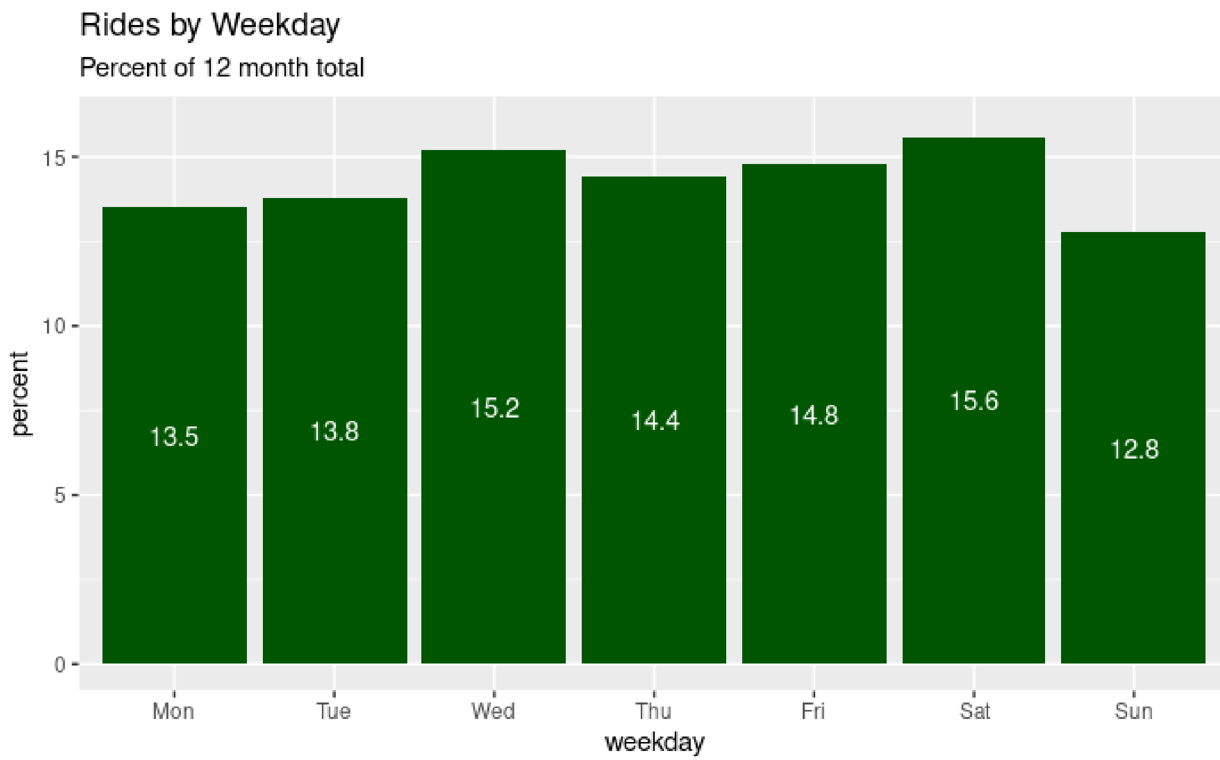


Figure 32: Rentals by weekday - all members

7.5.2.1.1 Summary

Saturday is the busiest day for all members at 15.6% of total rides. Sunday is the least busiest day at 12.8%. Wednesday is the highest day at 15.2% for the five-day workweek.

7.5.2.2 Rides by weekday - Annual members

Rides by weekday for Annual members are shown Figure 33.

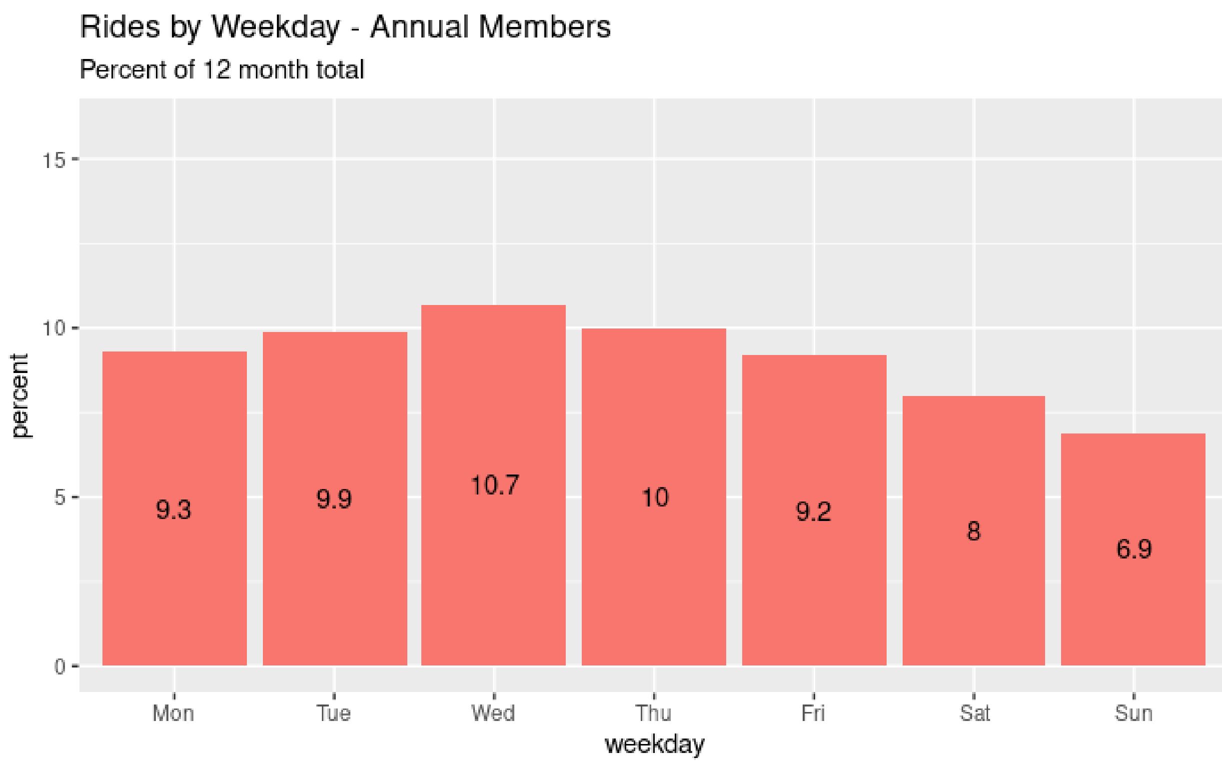


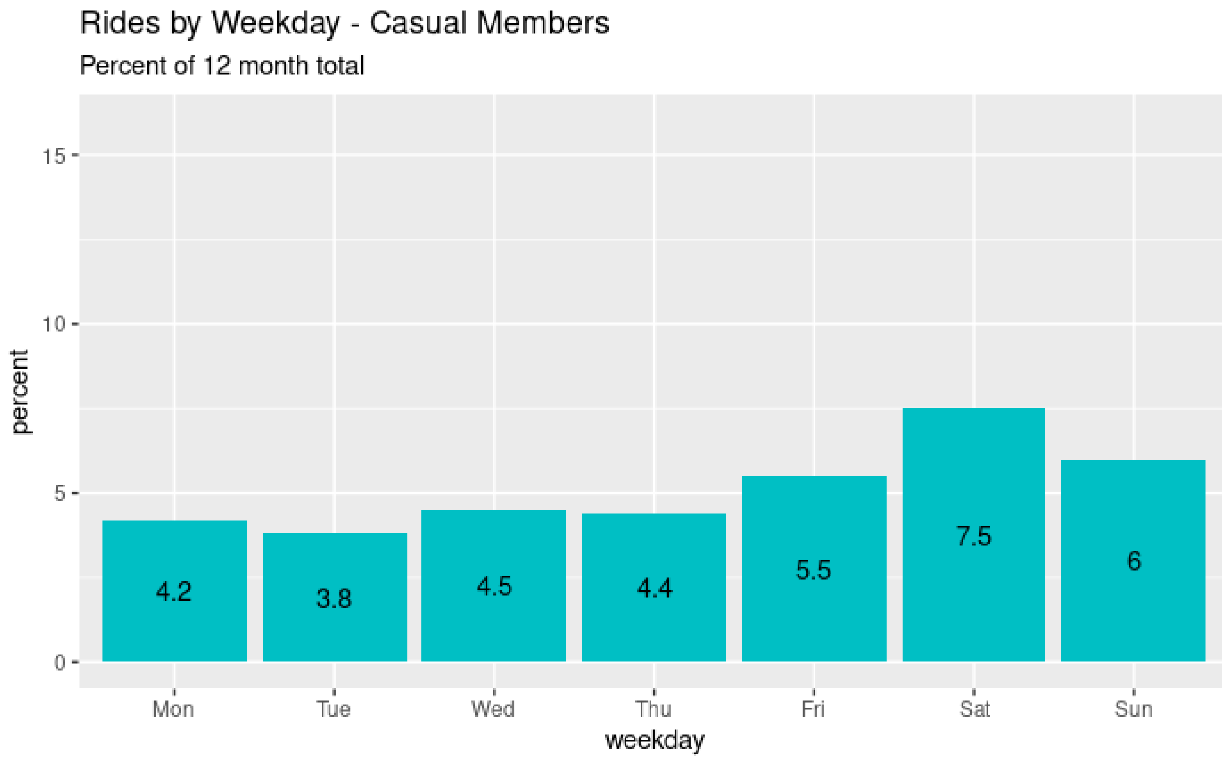
Figure 33: Rides by weekday - Annual members

7.5.2.2.1 Summary

Wednesday is the busiest day for Annual members at 10.7%. Sunday is the lowest at only 6.9%. There is clearly more ride-rentals during the workweek than the weekend. This suggests Annual members are riding to work or school. Or, if they ride for exercise, do so mostly during the five-day workweek.

7.5.2.3 Rides by weekday - Casual members

Rides by weekday for Casual members are shown Figure 34.



7.5.2.3.1 Summary

Casual members ride activity is oriented towards weekends. The busiest days are Saturdays at 7.5%, followed by Sundays at 6.0%. For Casual members, the five-day workweek has generally lower ride participation. The one exception being Fridays at 5.5% which leads into the Saturday spike.

7.5.2.4 Rides by weekday - combined member types

Rides by month for combined member types are shown in Figure 34.

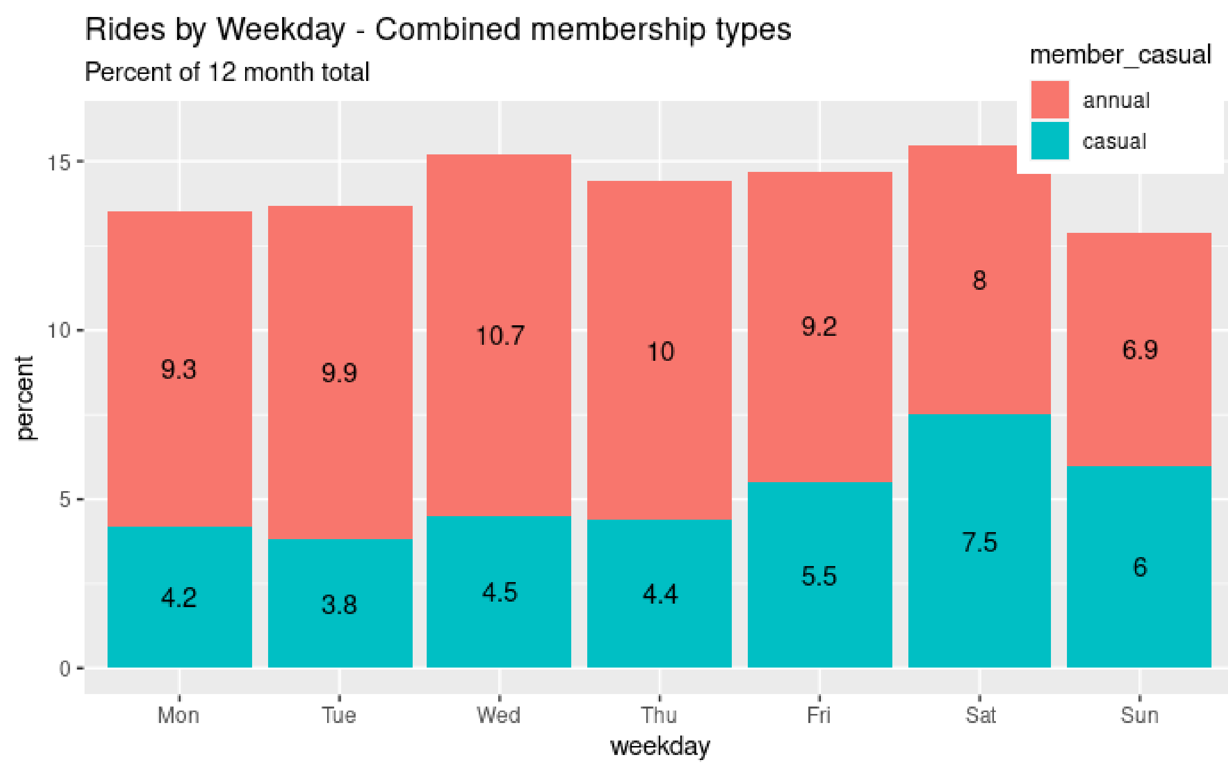


Figure 34: Rides by weekday - combined member types

7.5.2.4.1 Summary

Saturday is the busiest day for both member types combined at 15.6%[3]. This is due to the large spike of Casual riders making up for the lower weekend participation by Annual members. Although it's the second most active day for Casual riders, Sundays are the lowest overall at 12.8%.

3 There are the usual round-off discrepancies that show up in some of the value notations when combining categorical variables. Values used in the summary are from Figure 32.

7.5.3 Rides by Hour

7.5.3.1 Rides by hour - all members

All members ride-rentals by hour are shown in Figure 35

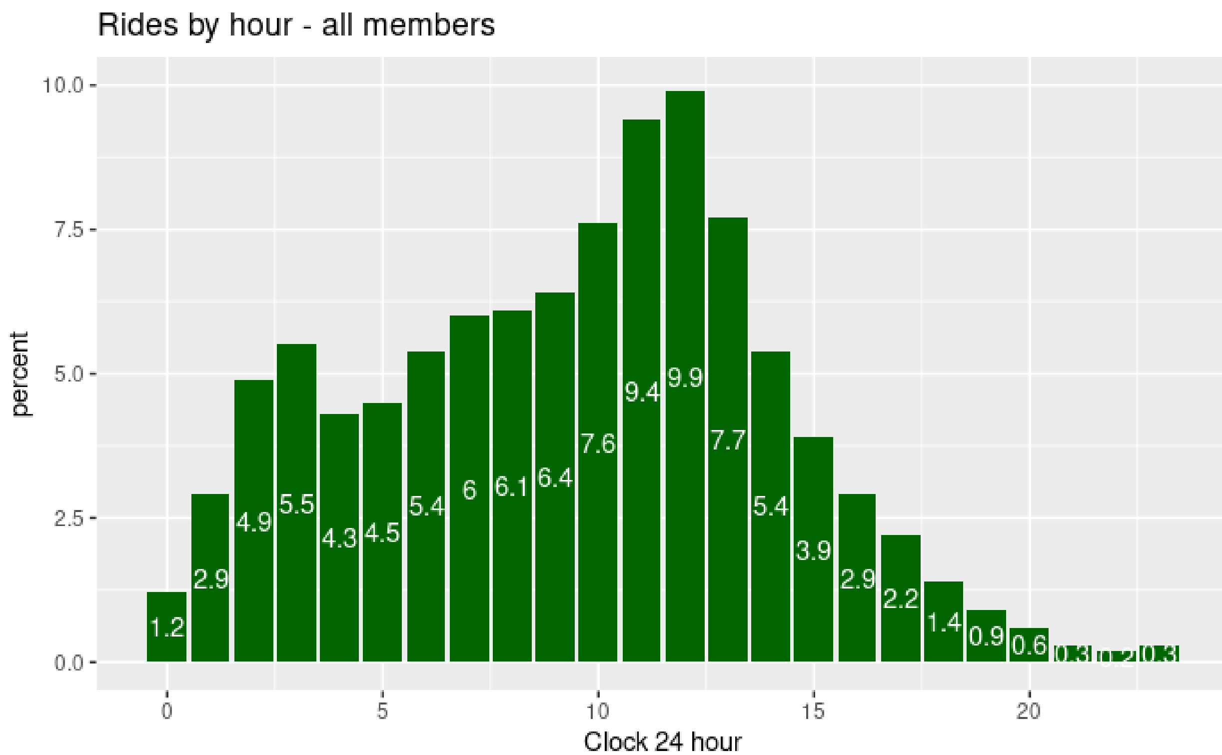


Figure 35: Rides by hour - all members

7.5.3.1.1 Summary

The busiest hour is Noon at 9.9%. The least busiest hour is 10pm at 0.2%. For the most part, this plot is as expected: A steady increase during the morning, peaking late morning through the lunch hour, then followed by a steady decline. The notable exception is the early morning ramp-up. Rides begin increasing sharply starting at Midnight and continuing through 3am.

7.5.3.2 Rides by hour - Annual members

Ride-rentals for Annual members are shown in Figure 36.

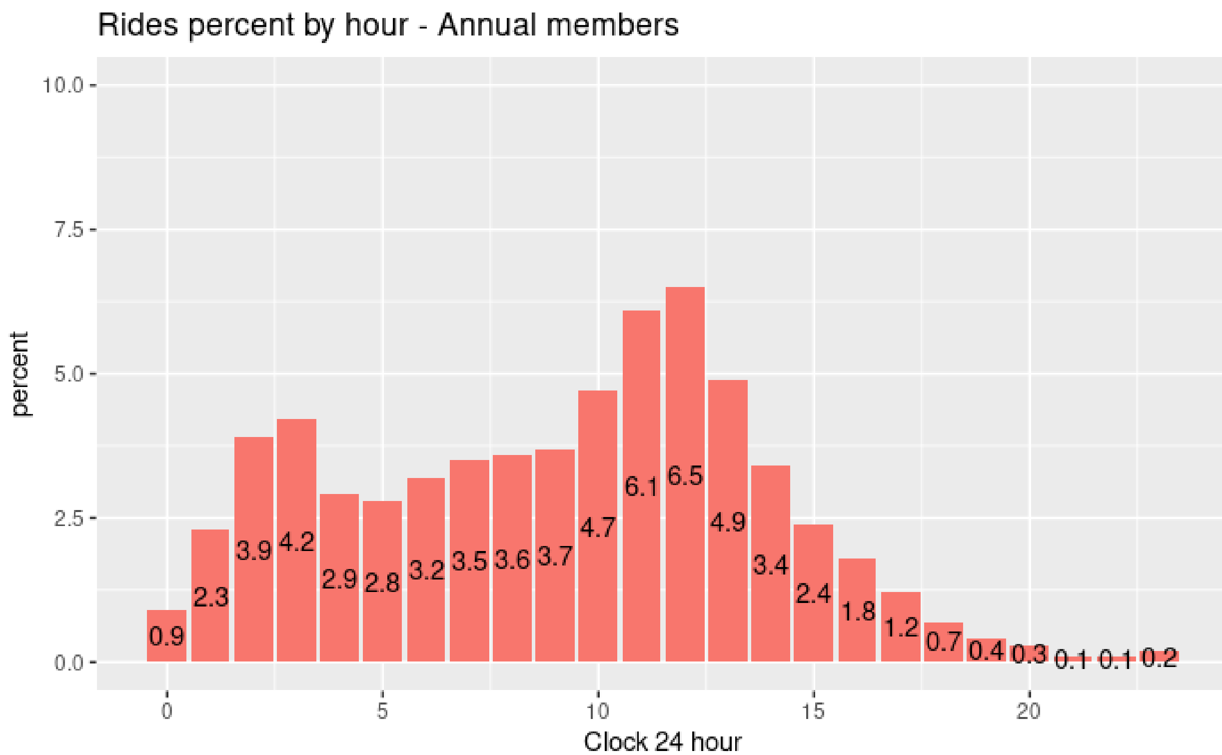


Figure 36: Rides by hour - Annual members

7.5.3.2.1 Summary

The busiest hour is Noon at 6.5%. The least busiest hours are 9 - 10pm.

The early morning ramp-up starting at Midnight from Figure 35 is also matched in this graph. Thus, these are mostly Annual members riding around Chicago at these times. Bars in Chicago close between 2am and 4am depending on their license. It seems logical to assume a relationship exists.

7.5.3.3 Rides by hour - Casual members

Ride-rentals for Casual members are shown in Figure 37.

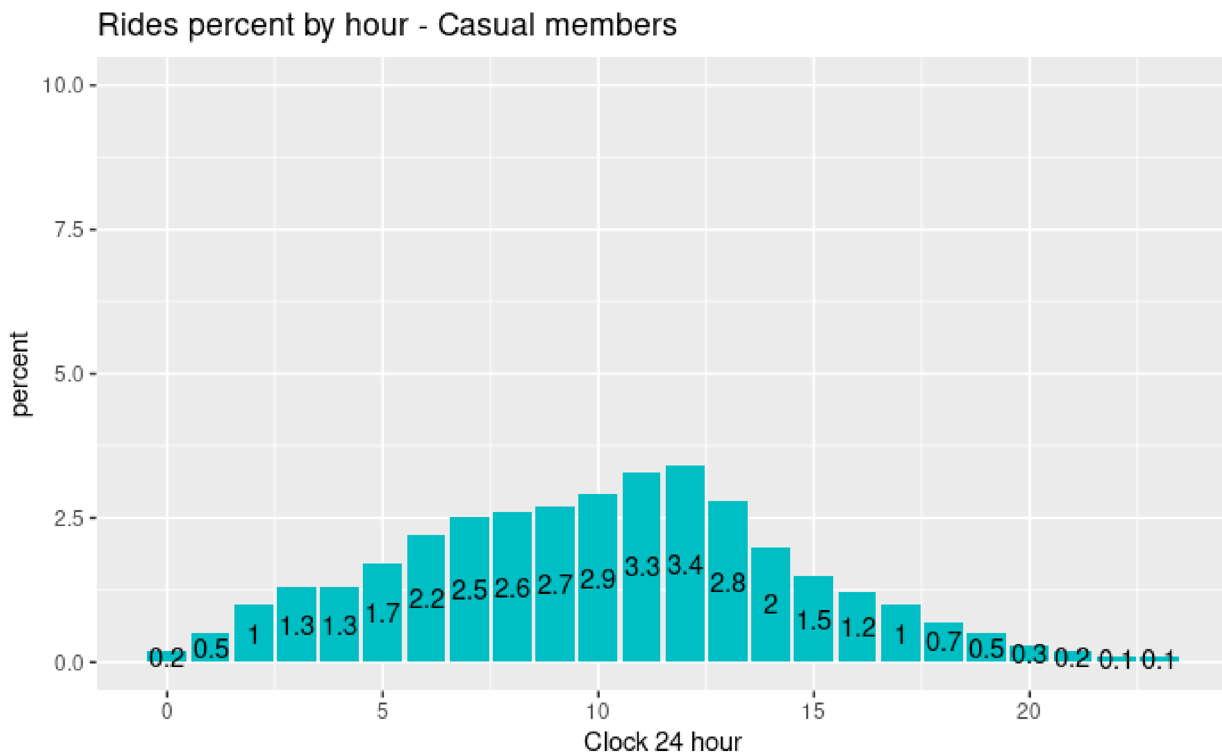


Figure 37: Rides by hour - Casual members

7.5.3.3.1 Summary

The busiest hour is Noon at 3.4%. The least busiest hours are 10 - 11pm at 0.1% each.

The steep early morning spike seen for Annual member is not present here, but there is a steady increase in rentals starting at Midnight which continues through the Noon peak.

7.5.3.4 Rides by hour - combined member types

Ride-rentals for combined member types are shown in Figure 38.

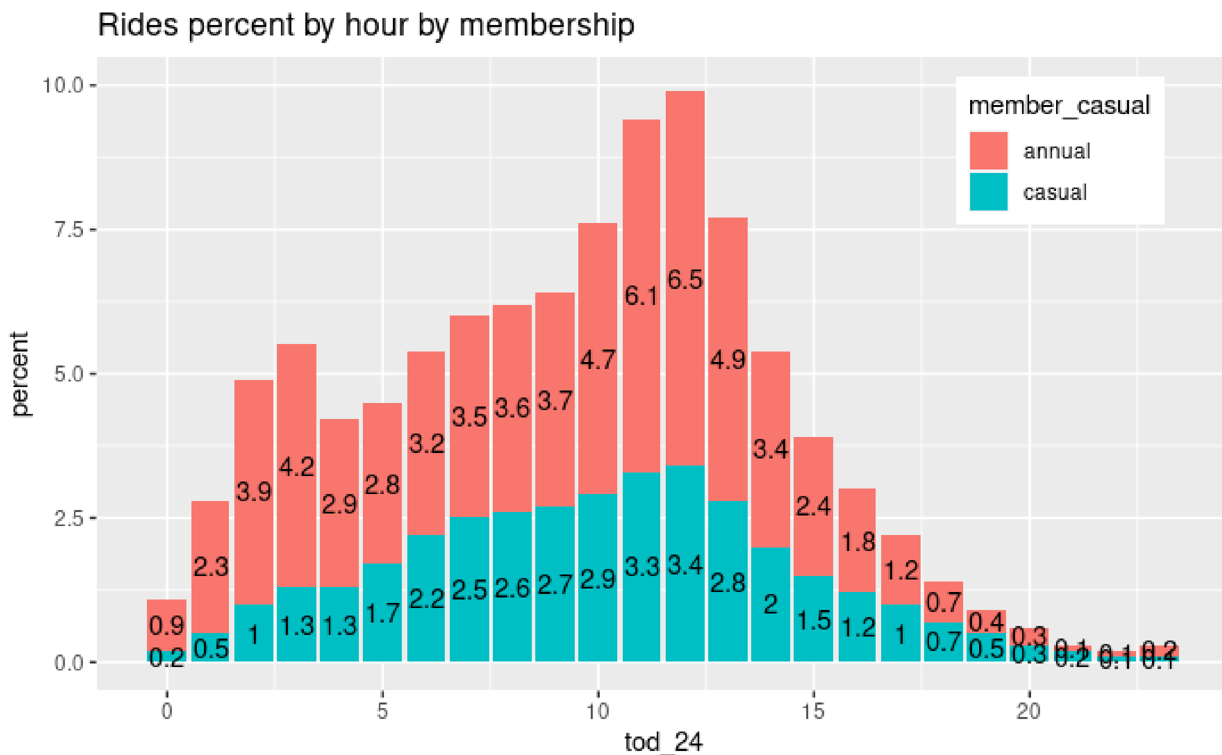


Figure 38: Rides by hour - combined member types

7.5.3.4.1 Summary

The busiest hour for both member types combined is Noon at 9.9%. The least busiest hour is 10pm at 0.2%.

The early morning ramp-up for combined members peaks in the 3am hour at 5.5%.