Cyclistic Bike Share - A Case Study

Aurin Bose

2023-01-06

I Introduction

The following Case Study has been done in order to determine how Annual member differ from casual riders on a fictional bike sharing service called Cyclistic which in 2016 launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

The analysis performed on their bike sharing data will help them determine how to convert casual riders into annual members which in turn will be the key to their future growth. The six key phases of data analysis have been used to complete this objective and will be highlighted in the rest of this report.

II Ask Phase

The Ask Phase is used by data analysts to determine the problem to be solved and how their insights could drive business decisions.

In this case the central query to be answered has been assigned to us namely: How do annual members and casual riders use Cyclistic bikes differently?

III Prepare Phase

The Prepare Phase is meant for collecting, organizing, sorting and filtering data. It is important in this phase to check where the data comes from, whether it meets the ROCCC principles and if the data can be used for the purpose of data analysis.

The data used here is based on the Divvy case study "'Sophisticated, Clear, and Polished': Divvy and Data Visualization" written by Kevin Hartman found here.

(*Note*: The datasets have a different name because Cyclistic is a fictional company. For the purposes of this case study, the datasets are appropriate and will enable you to answer the business questions. The data has been made available by Motivate International Inc. under this license)

Since the data has been obtained from a first hand source, it can classified as trustworthy and any licensing issues have been addressed as well. The data provided contains bike ride information from the past 12 months from when this Case Study is being performed, hence January to December 2022.

IV Process Phase

The Process Phase is used to check the data for errors, select the tools we wish to use for analysis, transform the data as per our requirements and clean the data.

For the purposes of this study since the dataset provided was of significant size, the best options were to use SQL or R programming and subsequently R has been deemed most appropriate.

The first step in R is to load the libraries necessary in the entire analysis process:

```
library(tidyverse)
library(lubridate)
library(ggplot2)
library(tidyr)
library(scales)
```

Then we check the current directory in R and set it to the folder where our dataset is located if it is not assigned already.

Next we can begin loading all our dataset into R which has been stored in the form of CSV files:

```
jan_22<- read.csv("202201-divvy-tripdata.csv")
feb_22<- read.csv("202202-divvy-tripdata.csv")
mar_22<- read.csv("202203-divvy-tripdata.csv")
apr_22<- read.csv("202204-divvy-tripdata.csv")
may_22<- read.csv("202205-divvy-tripdata.csv")
jun_22<- read.csv("202206-divvy-tripdata.csv")
jul_22<- read.csv("202207-divvy-tripdata.csv")
aug_22<- read.csv("202208-divvy-tripdata.csv")
sep_22<- read.csv("202209-divvy-tripdata.csv")
oct_22<- read.csv("202210-divvy-tripdata.csv")
nov_22<- read.csv("202211-divvy-tripdata.csv")
dec_22<- read.csv("202212-divvy-tripdata.csv")
```

Once loaded we can preview the columns in our data and check their basic structures as shown in the samples below:

```
colnames(jan_22)

##  [1] "ride_id"            "rideable_type"      "started_at"
##  [4] "ended_at"           "start_station_name" "start_station_id"
##  [7] "end_station_name"   "end_station_id"     "start_lat"
## [10] "start_lng"          "end_lat"            "end_lng"
## [13] "member_casual"

colnames(feb_22)

##  [1] "ride_id"            "rideable_type"      "started_at"
##  [4] "ended_at"           "start_station_name" "start_station_id"
##  [7] "end_station_name"   "end_station_id"     "start_lat"
```

```
## [10] "start_lng"          "end_lat"            "end_lng"
## [13] "member_casual"
```

```
str(jan_22)
```

```
## 'data.frame':    103770 obs. of  13 variables:
##  $ ride_id           : chr  "C2F7DD78E82EC875" "A6CF8980A652D272"
"BD0F91DFF741C66D" "CBB80ED419105406" ...
##  $ rideable_type     : chr  "electric_bike" "electric_bike" "classic_bike"
"classic_bike" ...
##  $ started_at        : chr  "2022-01-13 11:59:47" "2022-01-10 08:41:56"
"2022-01-25 04:53:40" "2022-01-04 00:18:04" ...
##  $ ended_at          : chr  "2022-01-13 12:02:44" "2022-01-10 08:46:17"
"2022-01-25 04:58:01" "2022-01-04 00:33:00" ...
##  $ start_station_name: chr  "Glenwood Ave & Touhy Ave" "Glenwood Ave &
Touhy Ave" "Sheffield Ave & Fullerton Ave" "Clark St & Bryn Mawr Ave" ...
##  $ start_station_id  : chr  "525" "525" "TA1306000016" "KA1504000151" ...
##  $ end_station_name  : chr  "Clark St & Touhy Ave" "Clark St & Touhy Ave"
"Greenview Ave & Fullerton Ave" "Paulina St & Montrose Ave" ...
##  $ end_station_id    : chr  "RP-007" "RP-007" "TA1307000001"
"TA1309000021" ...
##  $ start_lat         : num  42 42 41.9 42 41.9 ...
##  $ start_lng         : num  -87.7 -87.7 -87.7 -87.7 -87.6 ...
##  $ end_lat           : num  42 42 41.9 42 41.9 ...
##  $ end_lng           : num  -87.7 -87.7 -87.7 -87.7 -87.6 ...
##  $ member_casual     : chr  "casual" "casual" "member" "casual" ...
```

```
str(feb_22)
```

```
## 'data.frame':    115609 obs. of  13 variables:
##  $ ride_id           : chr  "E1E065E7ED285C02" "1602DCDC5B30FFE3"
"BE7DD2AF4B55C4AF" "A1789BDF844412BE" ...
##  $ rideable_type     : chr  "classic_bike" "classic_bike" "classic_bike"
"classic_bike" ...
##  $ started_at        : chr  "2022-02-19 18:08:41" "2022-02-20 17:41:30"
"2022-02-25 18:55:56" "2022-02-14 11:57:03" ...
##  $ ended_at          : chr  "2022-02-19 18:23:56" "2022-02-20 17:45:56"
"2022-02-25 19:09:34" "2022-02-14 12:04:00" ...
##  $ start_station_name: chr  "State St & Randolph St" "Halsted St &
Wrightwood Ave" "State St & Randolph St" "Southport Ave & Waveland Ave" ...
##  $ start_station_id  : chr  "TA1305000029" "TA1309000061" "TA1305000029"
"13235" ...
##  $ end_station_name  : chr  "Clark St & Lincoln Ave" "Southport Ave &
Wrightwood Ave" "Canal St & Adams St" "Broadway & Sheridan Rd" ...
##  $ end_station_id    : chr  "13179" "TA1307000113" "13011" "13323" ...
##  $ start_lat         : num  41.9 41.9 41.9 41.9 41.9 ...
##  $ start_lng         : num  -87.6 -87.6 -87.6 -87.7 -87.6 ...
##  $ end_lat           : num  41.9 41.9 41.9 42 41.9 ...
##  $ end_lng           : num  -87.6 -87.7 -87.6 -87.6 -87.6 ...
##  $ member_casual     : chr  "member" "member" "member" "member" ...
```

Now we can merge our individual datasets into a large dataset that will be used for the overall analysis process:

```
all_trip<- bind_rows(jan_22, feb_22, mar_22, apr_22, may_22, jun_22, jul_22,
                     aug_22, sep_22, oct_22, nov_22, dec_22)
```

Then we can preview the new merged data set by performing the below operations:

```
head(all_trip) #see the first 6 rows of data frame

##            ride_id rideable_type          started_at            ended_at
## 1 C2F7DD78E82EC875 electric_bike 2022-01-13 11:59:47 2022-01-13 12:02:44
## 2 A6CF8980A652D272 electric_bike 2022-01-10 08:41:56 2022-01-10 08:46:17
## 3 BD0F91DFF741C66D  classic_bike 2022-01-25 04:53:40 2022-01-25 04:58:01
## 4 CBB80ED419105406  classic_bike 2022-01-04 00:18:04 2022-01-04 00:33:00
## 5 DDC963BFDDA51EEA  classic_bike 2022-01-20 01:31:10 2022-01-20 01:37:12
## 6 A39C6F6CC0586C0B  classic_bike 2022-01-11 18:48:09 2022-01-11 18:51:31
##              start_station_name start_station_id
end_station_name
## 1      Glenwood Ave & Touhy Ave              525          Clark St & Touhy
Ave
## 2      Glenwood Ave & Touhy Ave              525          Clark St & Touhy
Ave
## 3 Sheffield Ave & Fullerton Ave     TA1306000016 Greenview Ave & Fullerton
Ave
## 4      Clark St & Bryn Mawr Ave     KA1504000151     Paulina St & Montrose
Ave
## 5   Michigan Ave & Jackson Blvd     TA1309000002        State St &
Randolph St
## 6         Wood St & Chicago Ave              637       Honore St &
Division St
##   end_station_id start_lat start_lng  end_lat   end_lng member_casual
## 1         RP-007  42.01280 -87.66591 42.01256 -87.67437        casual
## 2         RP-007  42.01276 -87.66597 42.01256 -87.67437        casual
## 3   TA1307000001  41.92560 -87.65371 41.92533 -87.66580        member
## 4   TA1309000021  41.98359 -87.66915 41.96151 -87.67139        casual
## 5   TA1305000029  41.87785 -87.62408 41.88462 -87.62783        member
## 6   TA1305000034  41.89563 -87.67207 41.90312 -87.67394        member

colnames(all_trip) #list of column names

##  [1] "ride_id"            "rideable_type"      "started_at"
##  [4] "ended_at"           "start_station_name" "start_station_id"
##  [7] "end_station_name"   "end_station_id"     "start_lat"
## [10] "start_lng"          "end_lat"            "end_lng"
## [13] "member_casual"

str(all_trip) #see list of columns and data types

## 'data.frame':    5667717 obs. of  13 variables:
##  $ ride_id           : chr  "C2F7DD78E82EC875" "A6CF8980A652D272"
```

```
"BD0F91DFF741C66D" "CBB80ED419105406" ...
## $ rideable_type     : chr  "electric_bike" "electric_bike" "classic_bike"
"classic_bike" ...
## $ started_at        : chr  "2022-01-13 11:59:47" "2022-01-10 08:41:56"
"2022-01-25 04:53:40" "2022-01-04 00:18:04" ...
## $ ended_at          : chr  "2022-01-13 12:02:44" "2022-01-10 08:46:17"
"2022-01-25 04:58:01" "2022-01-04 00:33:00" ...
## $ start_station_name: chr  "Glenwood Ave & Touhy Ave" "Glenwood Ave &
Touhy Ave" "Sheffield Ave & Fullerton Ave" "Clark St & Bryn Mawr Ave" ...
## $ start_station_id  : chr  "525" "525" "TA1306000016" "KA1504000151" ...
## $ end_station_name  : chr  "Clark St & Touhy Ave" "Clark St & Touhy Ave"
"Greenview Ave & Fullerton Ave" "Paulina St & Montrose Ave" ...
## $ end_station_id    : chr  "RP-007" "RP-007" "TA1307000001"
"TA1309000021" ...
## $ start_lat         : num  42 42 41.9 42 41.9 ...
## $ start_lng         : num  -87.7 -87.7 -87.7 -87.7 -87.6 ...
## $ end_lat           : num  42 42 41.9 42 41.9 ...
## $ end_lng           : num  -87.7 -87.7 -87.7 -87.7 -87.6 ...
## $ member_casual     : chr  "casual" "casual" "member" "casual" ...
```

```
summary(all_trip) #statistical summary of data
```

```
##    ride_id          rideable_type       started_at          ended_at
##  Length:5667717     Length:5667717     Length:5667717     Length:5667717
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  start_station_name start_station_id   end_station_name   end_station_id
##  Length:5667717     Length:5667717     Length:5667717     Length:5667717
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    start_lat       start_lng         end_lat         end_lng
##  Min.   :41.64   Min.   :-87.84   Min.   : 0.00   Min.   :-88.14
##  1st Qu.:41.88   1st Qu.:-87.66   1st Qu.:41.88   1st Qu.:-87.66
##  Median :41.90   Median :-87.64   Median :41.90   Median :-87.64
##  Mean   :41.90   Mean   :-87.65   Mean   :41.90   Mean   :-87.65
##  3rd Qu.:41.93   3rd Qu.:-87.63   3rd Qu.:41.93   3rd Qu.:-87.63
##  Max.   :45.64   Max.   :-73.80   Max.   :42.37   Max.   :  0.00
##                                   NA's   :5858    NA's   :5858
##  member_casual
##  Length:5667717
##  Class :character
##  Mode  :character
```

```
##
##
##
##

all_trip<- all_trip %>% rename (member_type=member_casual)
```

Before we begin our analysis there are a few of steps to be completed so that our data clean and suitable for making calculations:

- Firstly the formats of the started_at and ended_at columns were changed to Year-month-day and Hour-minute-seconds.

```
all_trip$started_at <- as.POSIXct(
  all_trip$started_at,
  format = "%Y-%m-%d %H:%M:%S"
)

all_trip$ended_at <- as.POSIXct(
  all_trip$ended_at,
  format = "%Y-%m-%d %H:%M:%S"
)

all_trip<- all_trip %>% arrange(started_at)
```

- Then a new column was created called ride length from calculating the difference between ride start time and end time.

```
all_trip$ride_length <- difftime(all_trip$ended_at,all_trip$started_at)
all_trip$ride_length <- as.numeric(as.character(all_trip$ride_length))
is.numeric(all_trip$ride_length)

## [1] TRUE
```

- Next separate columns were created for date, year, month, day, day of the week and even time of day.

```
all_trip$date <- as.Date(all_trip$started_at)
all_trip$month <- format(as.Date(all_trip$date), "%m")
all_trip$day <- format(as.Date(all_trip$date), "%d")
all_trip$year <- format(as.Date(all_trip$date), "%Y")
all_trip$day_of_week <- format(as.Date(all_trip$date), "%A")
all_trip$YMD <- format(all_trip$started_at, "%Y-%m-%d")
all_trip$ToD <- format(all_trip$started_at, "%H:%M:%S")
```

- The final steps involved removing any rows with negative ride lengths and empty rows in the start station name and end station name fields. To ensure that our original dataset is unchanged in case some chnages need to be reverted, we create a new dataset all_trip_v2.

```
all_trip_v2<- all_trip %>% filter (!(ride_length < 0))

all_trip_v2<- all_trip_v2 %>% filter(!(is.na(start_station_name) |
```

```
start_station_name == "")) %>%
  filter(!(is.na(end_station_name) | end_station_name == ""))
```

V Analysis Phase

The Analysis Phase is where we will aggregate our data, perform calculations on it and identify trends and relationships that could be useful towards accomplishing our business task.

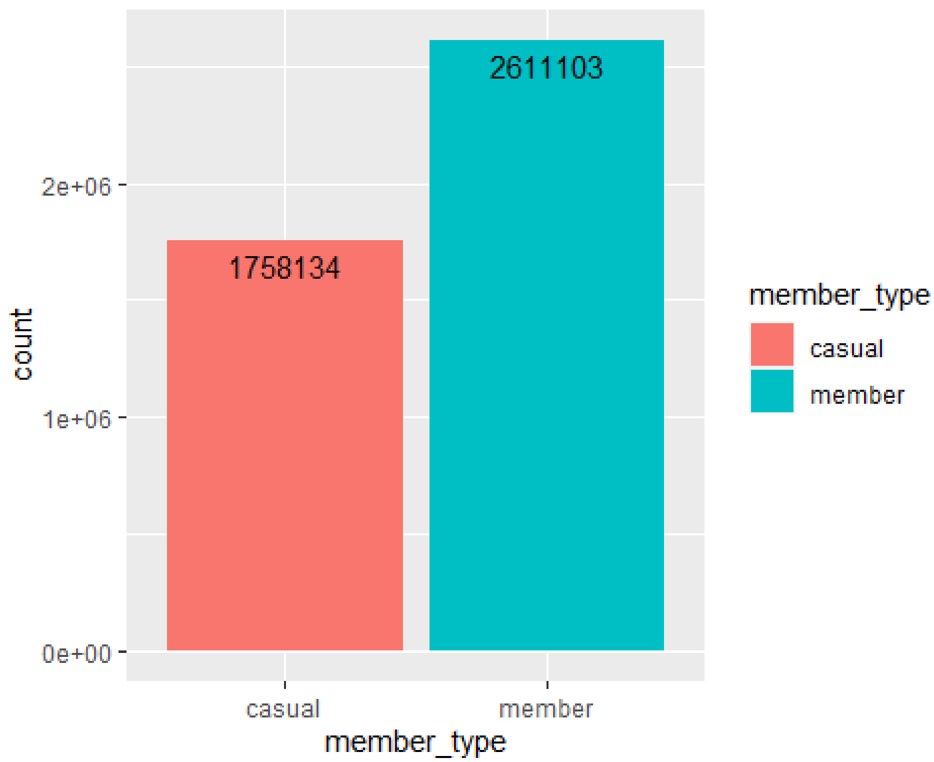
To start with we perform some basic descriptive analysis on our data:

```
mean(all_trip_v2$ride_length)

## [1] 1025.84

median(all_trip_v2$ride_length)

## [1] 636

max(all_trip_v2$ride_length)

## [1] 2057644

min(all_trip_v2$ride_length)

## [1] 0
```

We notice that the average ride length is 1025.84 secs, the midpoint number in the ascending array of ride lengths is 636 secs, maximum ride length is 2057644 secs and minimum is 0 secs.

Next we look at the total number of each type of rider (casual and annual member) and visualize it with the help of a bar chart:

```
ggplot(all_trip_v2, aes(member_type, fill = member_type))+
  geom_bar()+geom_text(aes(label=..count..),stat =
"count",vjust=1.5,color="black")
```



As illustrated above , there are 1758134 casual riders and 2611103 annual members present in the dataset.

Now we compare casual riders and annual members based on their ride lengths:

```
aggregate(all_trip_v2$ride_length ~ all_trip_v2$member_type, FUN = mean)
```

```
##   all_trip_v2$member_type all_trip_v2$ride_length
## 1                  casual               1439.7284
## 2                  member                747.1561
```

```
aggregate(all_trip_v2$ride_length ~ all_trip_v2$member_type, FUN = median)
```

```
##   all_trip_v2$member_type all_trip_v2$ride_length
## 1                  casual                     831
## 2                  member                     539
```

```
aggregate(all_trip_v2$ride_length ~ all_trip_v2$member_type, FUN = max)
```

```
##   all_trip_v2$member_type all_trip_v2$ride_length
## 1                  casual                 2057644
## 2                  member                   89594
```

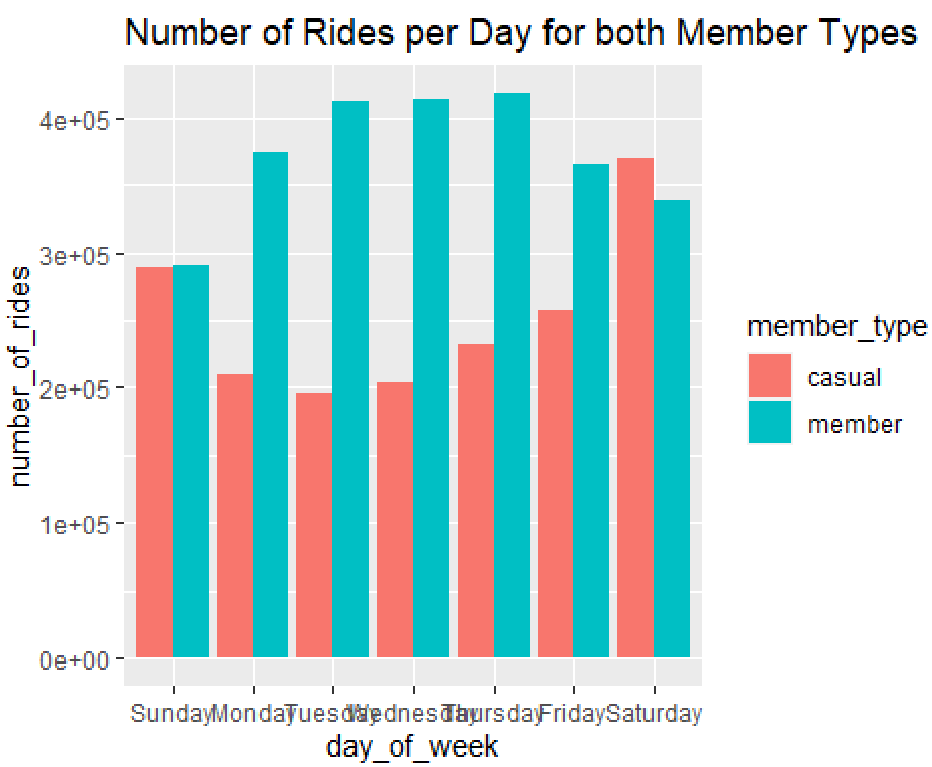
```
aggregate(all_trip_v2$ride_length ~ all_trip_v2$member_type, FUN = min)
```

```
##   all_trip_v2$member_type all_trip_v2$ride_length
## 1                  casual                       0
## 2                  member                       0
```

As observed, casual riders have a longer ride length on average at around **1440 secs** whereas annual members ride for an average of **747 secs**. Some other metrics noted were that casual riders have a median ride length of **831 secs** compared to the **539 secs** of annual members and the maximum ride length for them is **2057644 secs** compared to **89594 secs** for annual members.

Having calculated the above information we can begin visualizing this through a plot that illustrates the ride time per day of the week for both types of riders:

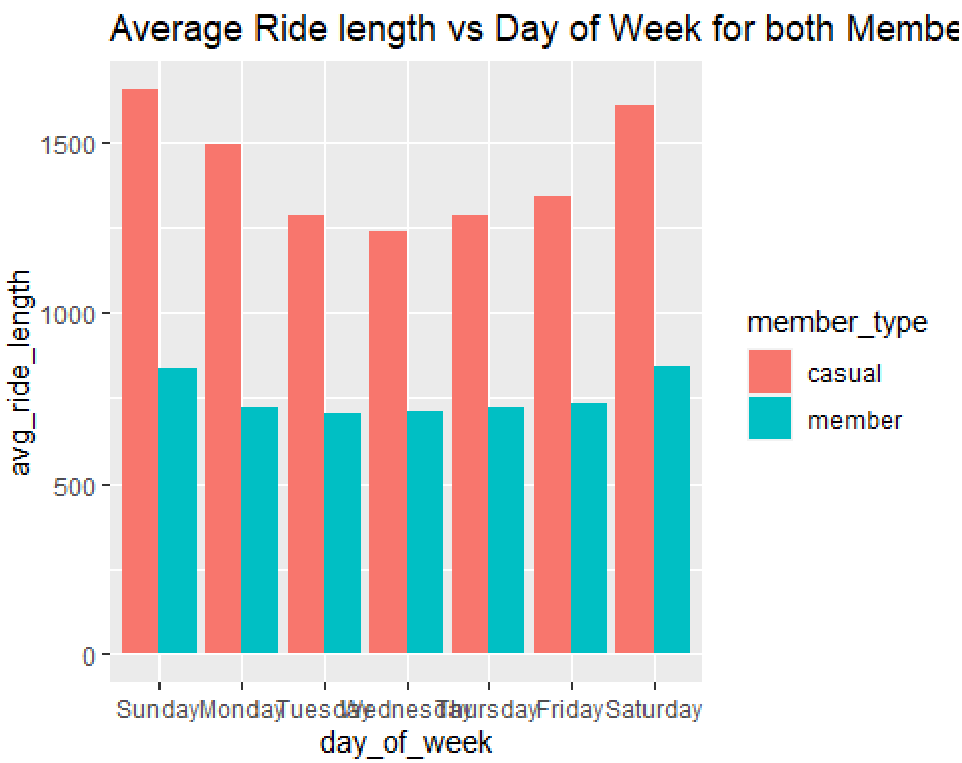
```
all_trip_v2$day_of_week <- ordered(all_trip_v2$day_of_week,
          levels=c("Sunday", "Monday","Tuesday", "Wednesday", "Thursday",
"Friday", "Saturday"))

all_trip_v2 %>%
  group_by(member_type, day_of_week) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_type, day_of_week) %>%
  ggplot(aes(x = day_of_week, y = number_of_rides, fill = member_type)) +
  geom_col(position = "dodge")+
  labs(title = "Number of Rides per Day for both Member Types")
```



From this column chart we notice that casual riders ride most frequently on the weekends and annual members are more likely to ride on the weekdays. In fact on Saturdays, casual riders complete a greater number of rides than annual members.

Having analyzed the rides per day for each member type we can now proceed to take a look at the average ride length for each day of the week:

```
all_trip_v2 %>% group_by(member_type, day_of_week) %>%
  summarise(avg_ride_length=mean(ride_length)) %>%
  arrange(member_type, day_of_week) %>%
  ggplot(aes(x=day_of_week, y=avg_ride_length, fill=member_type))+
  geom_col(position="dodge")+
  labs(title = "Average Ride length vs Day of Week for both Member Types")
```

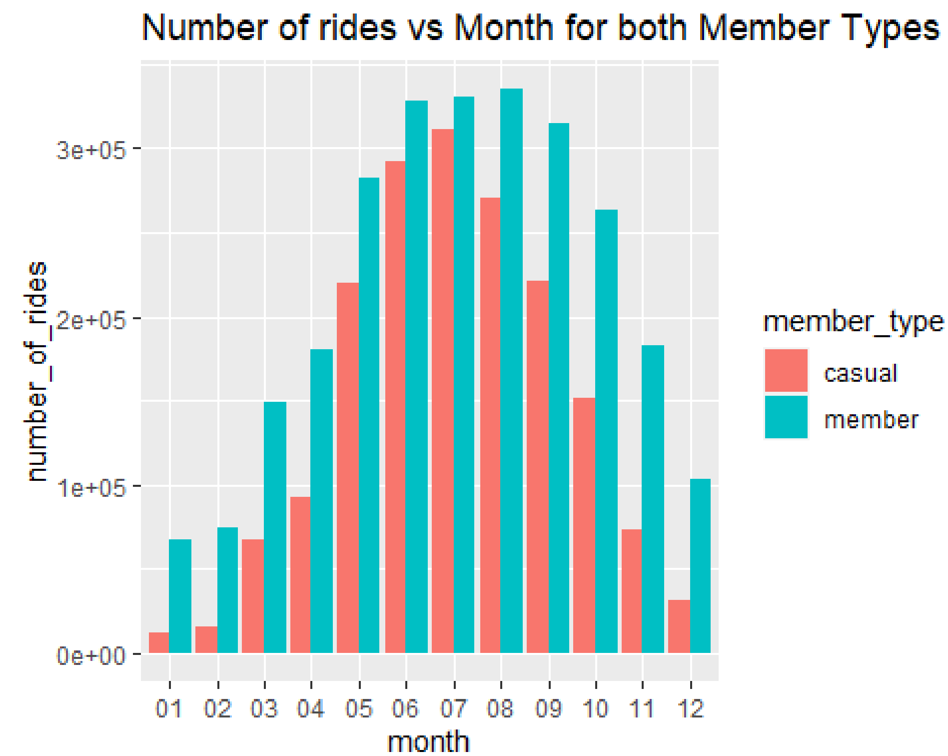


On an average casual riders seem to ride bikes longer than annual members.

To have a slightly deeper understanding of the riding habits of both member types we take a look at two column charts, the first for number of rides by member type per month and second for average ride length by each member type per month:

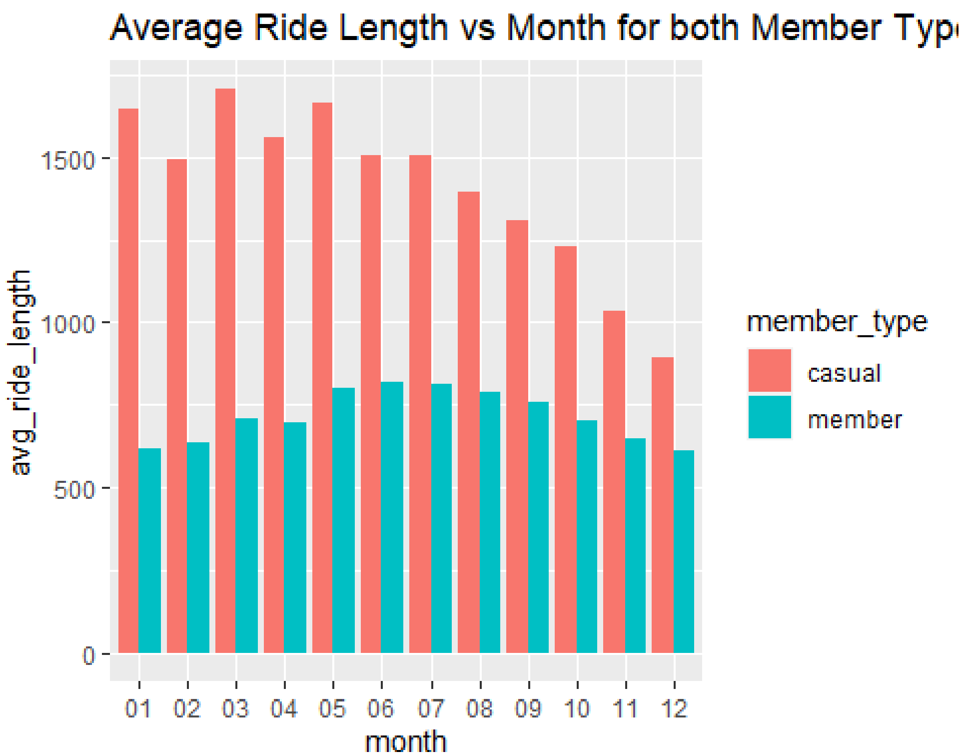
Number of Rides per Month:

```
all_trip_v2 %>% group_by(member_type, month) %>%
  summarise(number_of_rides = n()) %>%
  arrange(member_type, month) %>%
  ggplot(aes(x=month, y=number_of_rides, fill=member_type))+
  geom_col(position="dodge")+
  labs(title = "Number of rides vs Month for both Member Types")
```



Average Ride Length per Month:

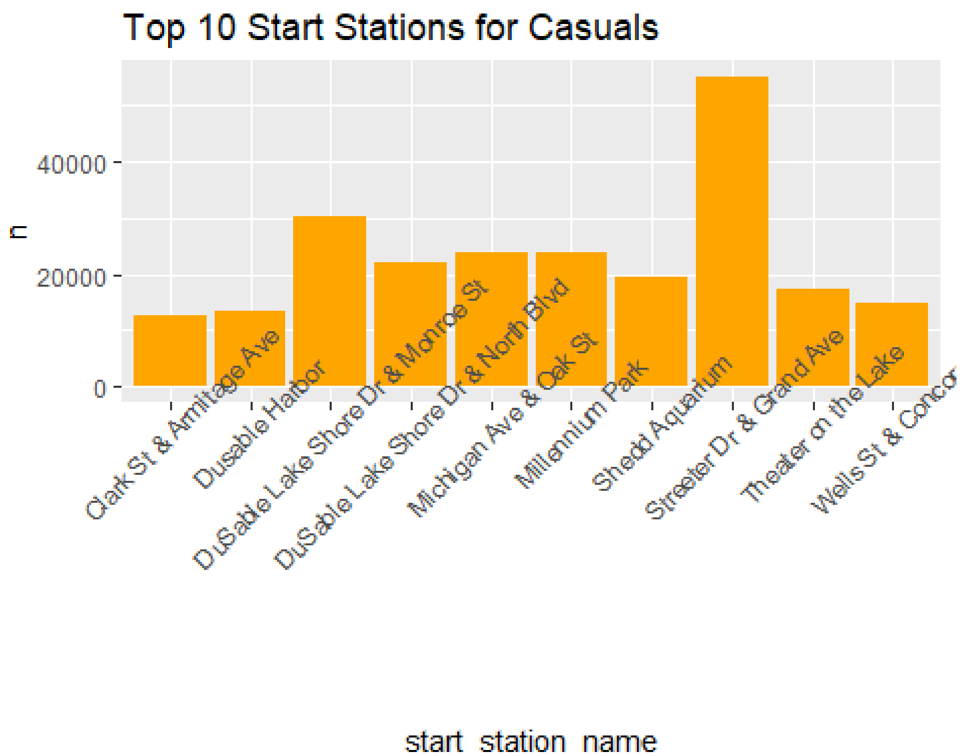
```
all_trip_v2 %>% group_by(member_type, month) %>%
  summarise(avg_ride_length=mean(ride_length)) %>%
  arrange(member_type, month) %>%
  ggplot(aes(x=month, y=avg_ride_length, fill=member_type))+
  geom_col(position="dodge")+
  labs(title = "Average Ride Length vs Month for both Member Types")
```



From our charts we can infer that peak ridership for casual riders is between June to August and for annual members this trend extends up until September. As we already know, the average ride length for casual riders is greater overall than annual members but we also observe that casual riders ride longer on average at the start of the year and then drop off towards the end of the year. Annual members remain quite consistent in terms of average ride length throughout the year.

Another interesting trend to take a look at would be the favorite stations for each member type from where they begin their journeys. This can be done as follows. First for Casual Riders:

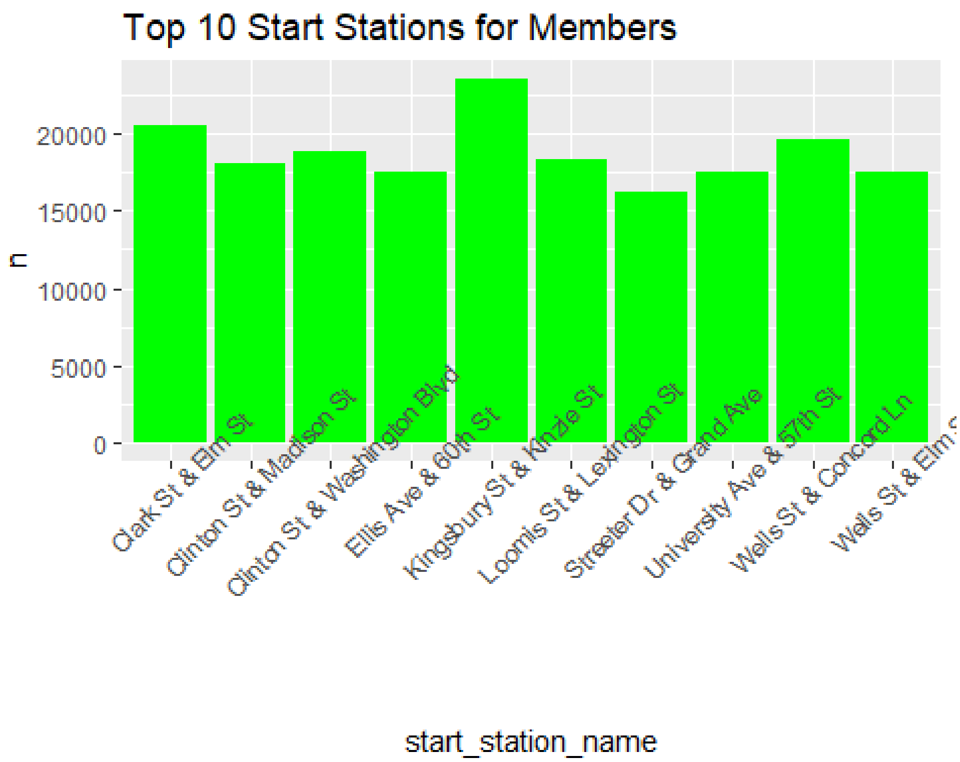
```
all_trip_v2 %>% filter(member_type == "casual") %>%
  group_by(start_station_name) %>%
  summarise(n=n()) %>% arrange(desc(n)) %>% slice_max(n,n=10) %>%
ggplot(aes(x = start_station_name, y = n))+
  geom_col(position = "dodge", fill="orange")+
  theme(axis.text.x = element_text(angle = 45))+
  labs(title="Top 10 Start Stations for Casuals")
```



From the above graph, casual riders start most of their journeys from **Streeter Dr & Grand Ave**

Then for Annual Members:

```
all_trip_v2 %>% filter(member_type == "member") %>%
  group_by(start_station_name) %>%
  summarise(n=n()) %>% arrange(desc(n)) %>% slice_max(n,n=10) %>%
  ggplot(aes(x = start_station_name, y = n))+
  geom_col(position = "dodge", fill="green")+
  theme(axis.text.x = element_text(angle = 45))+
  labs(title="Top 10 Start Stations for Members")
```



According to this analysis, annual members prefer to start from **Kingsbury St & Kinzie St**

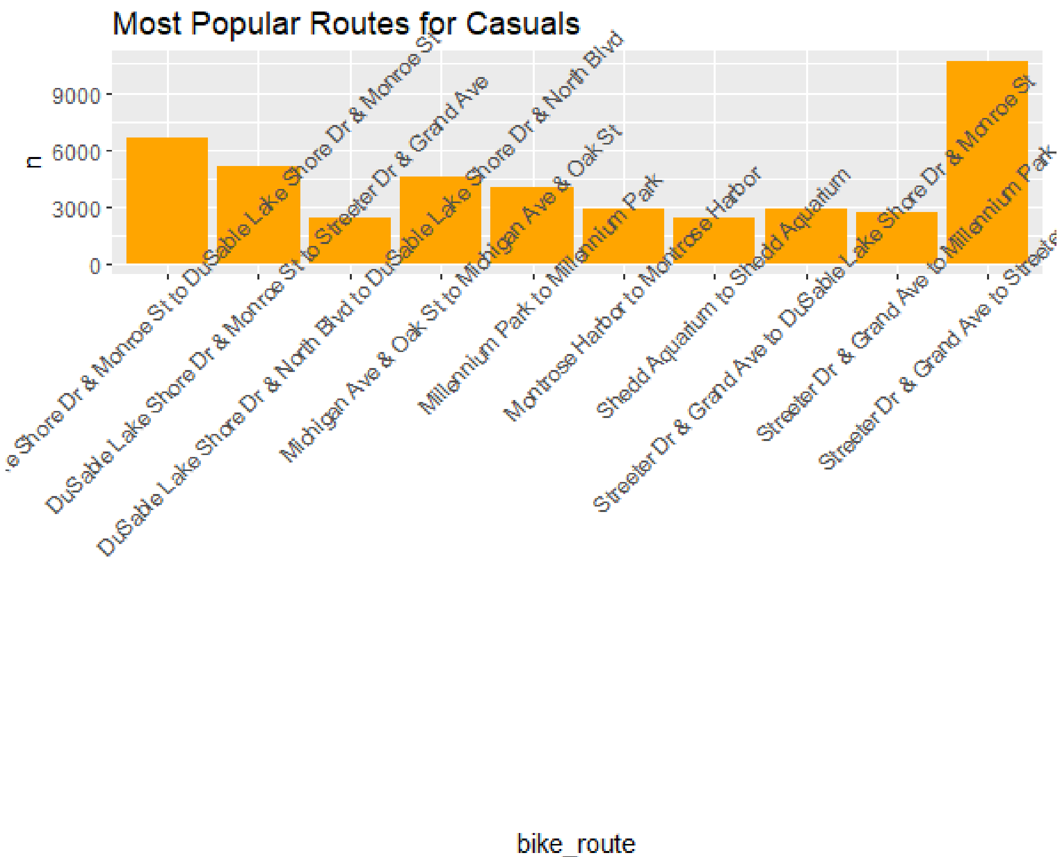
Having seen popular start stations among the different rider types, analyzing the routes taken would also help us get a clearer picture of how casual riders differ from annual members. For this first we create a new column route by merging start station and end station names:

```
all_trip_v2$bike_route = paste(all_trip_v2$start_station_name,
all_trip_v2$end_station_name,
                                            sep = " to ")
```

Once done, we can continue to analyze favourite routes for our two rider types:

Casual Riders:

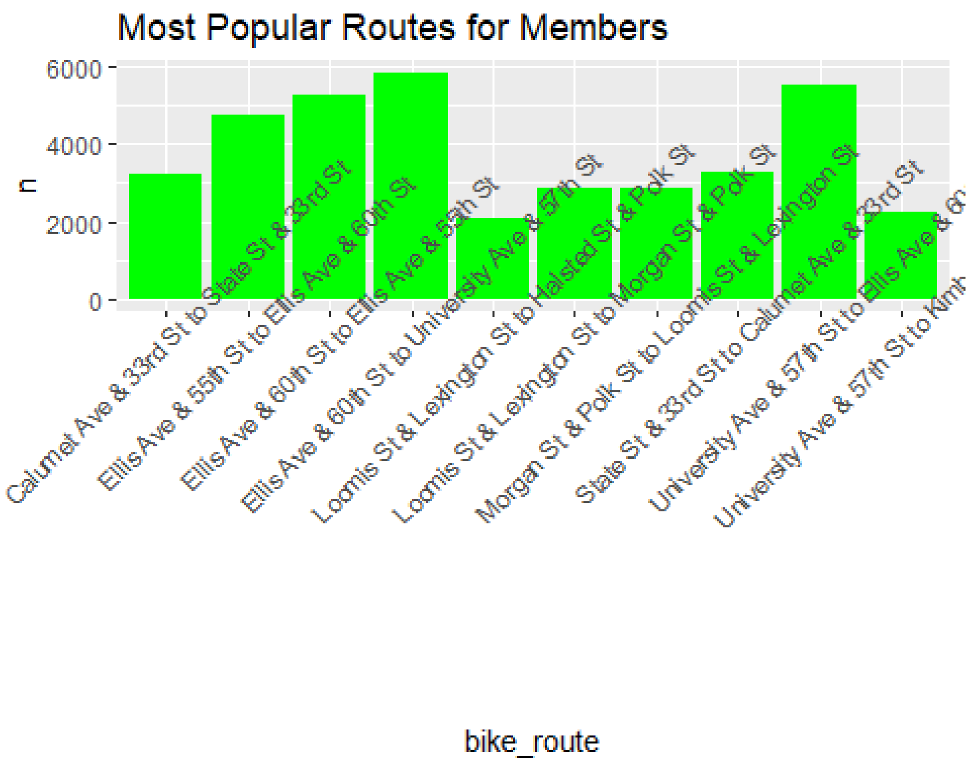
```
all_trip_v2 %>% filter(member_type == "casual") %>%
  group_by(bike_route) %>%
  summarise(n=n()) %>% arrange(desc(n)) %>% slice_max(n,n=10) %>%
  ggplot(aes(x = bike_route, y = n))+
  geom_col(position = "dodge", fill="orange")+
  theme(axis.text.x = element_text(angle = 45))+
  labs(title="Most Popular Routes for Casuals")
```



The most popular route as observed above for casual riders is from **Streeter Dr & Grand Ave** and back to the same station.

Annual Members:

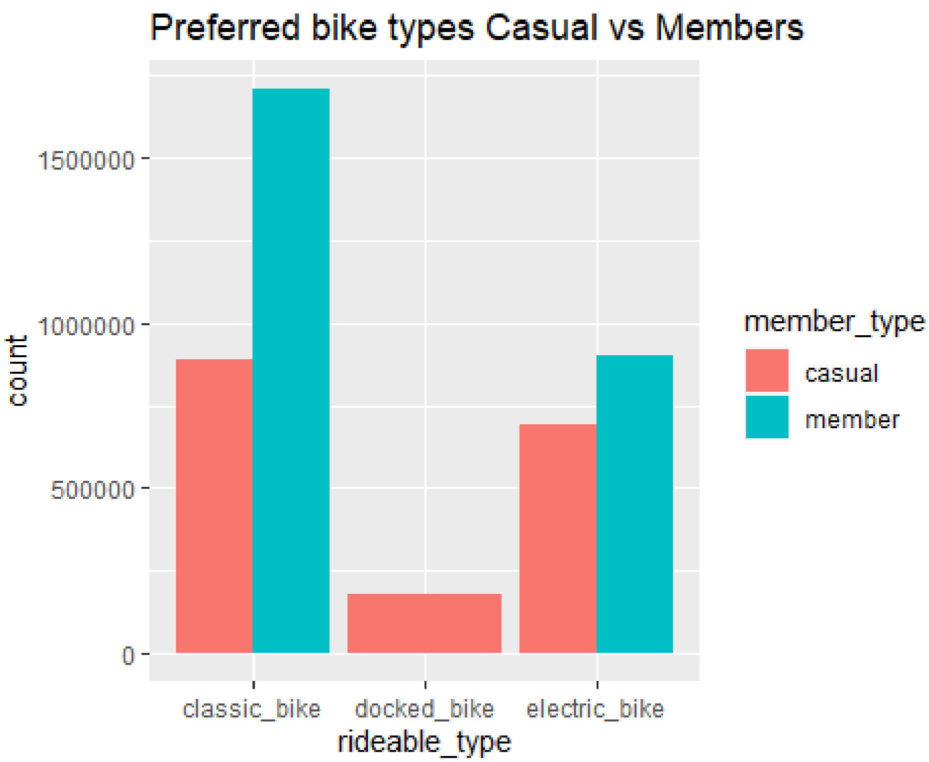
```
all_trip_v2 %>% filter(member_type == "member") %>%
  group_by(bike_route) %>%
  summarise(n=n()) %>% arrange(desc(n)) %>% slice_max(n,n=10) %>%
  ggplot(aes(x = bike_route, y = n))+
  geom_col(position = "dodge", fill="green")+
  theme(axis.text.x = element_text(angle = 45))+
  labs(title="Most Popular Routes for Members")
```



The most popular route as observed above annual members is from **Ellis Ave & 60th Street to University Ave & 57th Street**

For the final step of the analysis the bike type preferred by casual riders and annual members is identified and plotted as follows:

```
all_trip_v2 %>%
  group_by(member_type, rideable_type) %>%
  ggplot(aes(x = rideable_type, fill = member_type))+
  geom_bar(position = "dodge")+
  labs(title="Preferred bike types Casual vs Members")
```



This illustration clearly shows that annual members have a preference for classic bikes whereas electric bikes are equally preferred by both member types. There is no clear data on whether annual members use docked bikes or not hence a comparison is not possible at this time.

VI Conclusion

In this Case Study we began with a question that needed to be answered: How do annual members and casual riders use Cyclistic bikes differently? We then collected our data from a reputable data source and organized and sorted it into appropriate folders in the Prepare phase. Then in the Process phase we decided on R Studio as the tool of our choice and began load our data, check for errors or inconsistencies and clean the data. Finally in the Analysis phase we performed calculation on our data and reached the following major conclusions that define the differences between casual riders and Cyclistic annlual members:

- Casual riders have a longer ride length on average at around **1440 secs** whereas annual members ride for an average of **747 secs**. This could be because casual riders ride bikes for leisure or exercise whereas annual members might use the bikes for commuting purposes.
- Casual riders ride most frequently on the **weekends** and annual members are more likely to ride on the **weekdays**. This again could be because of the contrasting purposes for which the members might use bikes.

- Peak ridership for casual riders is between **June to August** and for annual members this trend extends up until **September**. The summer months are the most popular for bike riding as expected although annual members have a more even spread throughout the year due to the fact that they might be commuting by bikes.
- Casual riders ride longer on average at the **start of the year** and then drop off towards the end of the year. Annual members remain quite **consistent** in terms of average ride length throughout the year. The beginning of the year often is when people are motivated more than usual to exercise more frequently and lead a healthier lifestyle which could be why casual riders ride longer at that time. The weather could also influence this trend.
- Casual riders begin most of their rides from **Streeter Dr & Grand Ave** and their favorite route in turn is from Streeter Dr & Grand Ave back to the same station. Annual members start most frequently from **Kingsbury St & Kinzie St** but their most used route is from **Ellis Ave & 60th Street to University Ave & 57th Street**.
- Lastly, Annual members have a preference for **classic bikes** whereas **electric bikes** are almost equally preferred by both member types. Here an assumption could be made that either annual members belong to an older age group that prefer to not only commute by bikes but also prefer classic bikes over electric or that the stations they start from have more number of classic bikes available.