Bike Rides Analysis

Pooja Rathee

2022-04-08

Load Libraries.

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union

##
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':
##
##     discard

## The following object is masked from 'package:readr':
##
##     col_factor
```

Read all data files separately.

```
dt1 <- read_csv("data/202104-divvy-tripdata.csv")
```

```
## Rows: 337230 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt2 <- read_csv("data/202105-divvy-tripdata.csv")
```

```
## Rows: 531633 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt3 <- read_csv("data/202106-divvy-tripdata.csv")
```

```
## Rows: 729595 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt4 <- read_csv("data/202107-divvy-tripdata.csv")
```

```
## Rows: 822410 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt5 <- read_csv("data/202108-divvy-tripdata.csv")
```

```
## Rows: 804352 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
```

```
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt6 <- read_csv("data/202109-divvy-tripdata.csv")
```

```
## Rows: 756147 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt7 <- read_csv("data/202110-divvy-tripdata.csv")
```

```
## Rows: 631226 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt8 <- read_csv("data/202111-divvy-tripdata.csv")
```

```
## Rows: 359978 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt9 <- read_csv("data/202112-divvy-tripdata.csv")
```

```
## Rows: 247540 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt10 <- read_csv("data/202201-divvy-tripdata.csv")
```

```
## Rows: 103770 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt11 <- read_csv("data/202202-divvy-tripdata.csv")
```

```
## Rows: 115609 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
dt12 <- read_csv("data/202203-divvy-tripdata.csv")
```

```
## Rows: 284042 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): ride_id, rideable_type, start_station_name, start_station_id, end_...
## dbl  (4): start_lat, start_lng, end_lat, end_lng
## dttm (2): started_at, ended_at
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

Combine All files into `all_trips`

```
all_trips <- rbind(dt1, dt2, dt3, dt4, dt5, dt6, dt7, dt8, dt9, dt10, dt11, dt12)
```

Convert ride_id and rideable_type to character so that they can stack correctly

```
all_trips <-  mutate(all_trips, ride_id = as.character(ride_id)
                   ,rideable_type = as.character(rideable_type))
```

Inspect the new table that has been created

```
colnames(all_trips)  #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"
```

```
nrow(all_trips)  #How many rows are in data frame?
```

```
## [1] 5723532
```

```
dim(all_trips)  #Dimensions of the data frame?
```

```
## [1] 5723532      13
```

```
head(all_trips)  #See the first 6 rows of data frame.  Also tail(all_trips)
```

```
## # A tibble: 6 x 13
##   ride_id rideable_type started_at          ended_at            start_station_n~
##   <chr>   <chr>         <dttm>              <dttm>              <chr>           
## 1 6C992B~ classic_bike  2021-04-12 18:25:36 2021-04-12 18:56:55 State St & Pear~
## 2 1E0145~ docked_bike   2021-04-27 17:27:11 2021-04-27 18:31:29 Dorchester Ave ~
## 3 E498E1~ docked_bike   2021-04-03 12:42:45 2021-04-07 11:40:24 Loomis Blvd & 8~
## 4 188726~ classic_bike  2021-04-17 09:17:42 2021-04-17 09:42:48 Honore St & Div~
## 5 C12354~ docked_bike   2021-04-03 12:42:25 2021-04-03 14:13:42 Loomis Blvd & 8~
## 6 097E76~ classic_bike  2021-04-25 18:43:18 2021-04-25 18:43:59 Clinton St & Po~
## # ... with 8 more variables: start_station_id <chr>, end_station_name <chr>,
## #   end_station_id <chr>, start_lat <dbl>, start_lng <dbl>, end_lat <dbl>,
## #   end_lng <dbl>, member_casual <chr>
```

```
str(all_trips)  #See list of columns and data types (numeric, character, etc)
```

```
## tibble [5,723,532 x 13] (S3: tbl_df/tbl/data.frame)
##  $ ride_id           : chr [1:5723532] "6C992BD37A98A63F" "1E0145613A209000" "E498E15508A80BAD" "188
##  $ rideable_type     : chr [1:5723532] "classic_bike" "docked_bike" "docked_bike" "classic_bike" ...
##  $ started_at        : POSIXct[1:5723532], format: "2021-04-12 18:25:36" "2021-04-27 17:27:11" ...
##  $ ended_at          : POSIXct[1:5723532], format: "2021-04-12 18:56:55" "2021-04-27 18:31:29" ...
##  $ start_station_name: chr [1:5723532] "State St & Pearson St" "Dorchester Ave & 49th St" "Loomis Bl
##  $ start_station_id  : chr [1:5723532] "TA1307000061" "KA1503000069" "20121" "TA1305000034" ...
##  $ end_station_name  : chr [1:5723532] "Southport Ave & Waveland Ave" "Dorchester Ave & 49th St" "Lo
##  $ end_station_id    : chr [1:5723532] "13235" "KA1503000069" "20121" "13235" ...
##  $ start_lat         : num [1:5723532] 41.9 41.8 41.7 41.9 41.7 ...
##  $ start_lng         : num [1:5723532] -87.6 -87.6 -87.7 -87.7 -87.7 ...
##  $ end_lat           : num [1:5723532] 41.9 41.8 41.7 41.9 41.7 ...
##  $ end_lng           : num [1:5723532] -87.7 -87.6 -87.7 -87.7 -87.7 ...
##  $ member_casual     : chr [1:5723532] "member" "casual" "casual" "member" ...
```

```
summary(all_trips)  #Statistical summary of data. Mainly for numerics
```

```
##    ride_id          rideable_type        started_at                 
##  Length:5723532     Length:5723532     Min.   :2021-04-01 00:03:18  
##  Class :character   Class :character   1st Qu.:2021-06-22 15:20:26  
##  Mode  :character   Mode  :character   Median :2021-08-17 18:25:49  
##                                        Mean   :2021-08-26 22:25:18  
##                                        3rd Qu.:2021-10-14 19:48:10  
##                                        Max.   :2022-03-31 23:59:47  
##                                                                     
##     ended_at                   start_station_name start_station_id  
##  Min.   :2021-04-01 00:14:29   Length:5723532     Length:5723532    
##  1st Qu.:2021-06-22 15:47:37   Class :character   Class :character  
##  Median :2021-08-17 18:44:32   Mode  :character   Mode  :character  
##  Mean   :2021-08-26 22:46:50                                        
##  3rd Qu.:2021-10-14 20:03:28                                        
##  Max.   :2022-04-01 22:10:12                                        
##                                                                     
##  end_station_name   end_station_id       start_lat       start_lng     
##  Length:5723532     Length:5723532     Min.   :41.64   Min.   :-87.84  
##  Class :character   Class :character   1st Qu.:41.88   1st Qu.:-87.66  
##  Mode  :character   Mode  :character   Median :41.90   Median :-87.64  
##                                        Mean   :41.90   Mean   :-87.65  
##                                        3rd Qu.:41.93   3rd Qu.:-87.63  
##                                        Max.   :45.64   Max.   :-73.80  
##                                                                        
##     end_lat         end_lng       member_casual     
##  Min.   :41.39   Min.   :-88.97   Length:5723532    
##  1st Qu.:41.88   1st Qu.:-87.66   Class :character  
##  Median :41.90   Median :-87.64   Mode  :character  
##  Mean   :41.90   Mean   :-87.65                     
##  3rd Qu.:41.93   3rd Qu.:-87.63                     
##  Max.   :42.17   Max.   :-87.49                     
##  NA's   :4716    NA's   :4716
```

Examine the table

```
table(all_trips$member_casual)
```

```
## 
##  casual  member 
## 2546542 3176990
```

Formatting the date and time

```
all_trips$date <- as.Date(all_trips$started_at) #The default format is yyyy-mm-dd
all_trips$month <- format(as.Date(all_trips$date), "%m")
```

```
all_trips$day <- format(as.Date(all_trips$date), "%d")
all_trips$year <- format(as.Date(all_trips$date), "%Y")
all_trips$day_of_week <- format(as.Date(all_trips$date), "%A")
all_trips$start_hour <- lubridate::hour(all_trips$started_at)
all_trips$end_hour <- lubridate::hour(all_trips$ended_at)
```

Add a “ride_length” calculation to all_trips (in seconds)

```
all_trips$ride_length <- difftime(all_trips$ended_at,all_trips$started_at)

# Inspect the structure of the columns
str(all_trips)
```

```
## tibble [5,723,532 x 21] (S3: tbl_df/tbl/data.frame)
##  $ ride_id           : chr [1:5723532] "6C992BD37A98A63F" "1E0145613A209000" "E498E15508A80BAD" "188
##  $ rideable_type     : chr [1:5723532] "classic_bike" "docked_bike" "docked_bike" "classic_bike" ...
##  $ started_at        : POSIXct[1:5723532], format: "2021-04-12 18:25:36" "2021-04-27 17:27:11" ...
##  $ ended_at          : POSIXct[1:5723532], format: "2021-04-12 18:56:55" "2021-04-27 18:31:29" ...
##  $ start_station_name: chr [1:5723532] "State St & Pearson St" "Dorchester Ave & 49th St" "Loomis Bl
##  $ start_station_id  : chr [1:5723532] "TA1307000061" "KA1503000069" "20121" "TA1305000034" ...
##  $ end_station_name  : chr [1:5723532] "Southport Ave & Waveland Ave" "Dorchester Ave & 49th St" "Lo
##  $ end_station_id    : chr [1:5723532] "13235" "KA1503000069" "20121" "13235" ...
##  $ start_lat         : num [1:5723532] 41.9 41.8 41.7 41.9 41.7 ...
##  $ start_lng         : num [1:5723532] -87.6 -87.6 -87.7 -87.7 -87.7 ...
##  $ end_lat           : num [1:5723532] 41.9 41.8 41.7 41.9 41.7 ...
##  $ end_lng           : num [1:5723532] -87.7 -87.6 -87.7 -87.7 -87.7 ...
##  $ member_casual     : chr [1:5723532] "member" "casual" "casual" "member" ...
##  $ date              : Date[1:5723532], format: "2021-04-12" "2021-04-27" ...
##  $ month             : chr [1:5723532] "04" "04" "04" "04" ...
##  $ day               : chr [1:5723532] "12" "27" "03" "17" ...
##  $ year              : chr [1:5723532] "2021" "2021" "2021" "2021" ...
##  $ day_of_week       : chr [1:5723532] "Monday" "Tuesday" "Saturday" "Saturday" ...
##  $ start_hour        : int [1:5723532] 18 17 12 9 12 18 16 16 15 15 ...
##  $ end_hour          : int [1:5723532] 18 18 11 9 14 18 16 17 16 15 ...
##  $ ride_length       : ’difftime’ num [1:5723532] 1879 3858 341859 1506 ...
##   ..- attr(*, "units")= chr "secs"
```

```
# Convert "ride_length" from Factor to numeric so we can run calculations on the data
is.factor(all_trips$ride_length)
```

```
## [1] FALSE
```

```
all_trips$ride_length <- as.numeric(as.character(all_trips$ride_length))
is.numeric(all_trips$ride_length)
```

```
## [1] TRUE
```

Reassign the data

```
all_trips_v2 <- all_trips[!( all_trips$ride_length<0),]
```

Find the mean, median, max and min

```
mean(all_trips_v2$ride_length) #straight average (total ride length / rides)
```

```
## [1] 1292.602
```

```
median(all_trips_v2$ride_length) #midpoint number in the ascending array of ride lengths
```

```
## [1] 703
```

```
max(all_trips_v2$ride_length) #longest ride
```

```
## [1] 3356649
```

```
min(all_trips_v2$ride_length) #shortest ride
```

```
## [1] 0
```

summarize the data

```
summary(all_trips_v2$ride_length)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##       0     394     703    1293    1280 3356649 
```

Compare members and casual users

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual, FUN = mean)
```

```
##   all_trips_v2$member_casual all_trips_v2$ride_length
## 1                     casual                 1904.427
## 2                     member                  802.187
```

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual, FUN = median)
```

```
##   all_trips_v2$member_casual all_trips_v2$ride_length
## 1                     casual                      946
## 2                     member                      562
```

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual, FUN = max)
```

```
##   all_trips_v2$member_casual all_trips_v2$ride_length
## 1                     casual                  3356649
## 2                     member                    93594
```

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual, FUN = min)
```

```
##   all_trips_v2$member_casual all_trips_v2$ride_length
## 1                     casual                        0
## 2                     member                        0
```

See the average ride time by each day for members vs casual users

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual + all_trips_v2$day_of_week, FUN = mean)
```

```
##    all_trips_v2$member_casual all_trips_v2$day_of_week all_trips_v2$ride_length
## 1                      casual                   Friday                1806.2163
## 2                      member                   Friday                 788.3755
## 3                      casual                   Monday                1888.9498
## 4                      member                   Monday                 778.1150
## 5                      casual                 Saturday                2056.8695
## 6                      member                 Saturday                 899.6031
## 7                      casual                   Sunday                2244.4133
## 8                      member                   Sunday                 920.6725
## 9                      casual                 Thursday                1672.8766
## 10                     member                 Thursday                 754.2578
## 11                     casual                  Tuesday                1646.1100
## 12                     member                  Tuesday                 751.3135
## 13                     casual                Wednesday                1665.9587
## 14                     member                Wednesday                 755.3658
```

Notice that the days of the week are out of order. Let's fix that.

```
all_trips_v2$day_of_week <- ordered(all_trips_v2$day_of_week, levels=c("Sunday", "Monday", "Tuesday", "
```

Now, let's run the average ride time by each day for members vs casual users

```
aggregate(all_trips_v2$ride_length ~ all_trips_v2$member_casual + all_trips_v2$day_of_week, FUN = mean)
```

```
##    all_trips_v2$member_casual all_trips_v2$day_of_week all_trips_v2$ride_length
## 1                      casual                   Sunday                 2244.4133
## 2                      member                   Sunday                  920.6725
## 3                      casual                   Monday                 1888.9498
## 4                      member                   Monday                  778.1150
## 5                      casual                  Tuesday                 1646.1100
## 6                      member                  Tuesday                  751.3135
## 7                      casual                Wednesday                 1665.9587
## 8                      member                Wednesday                  755.3658
## 9                      casual                 Thursday                 1672.8766
## 10                     member                 Thursday                  754.2578
## 11                     casual                   Friday                 1806.2163
## 12                     member                   Friday                  788.3755
## 13                     casual                 Saturday                 2056.8695
## 14                     member                 Saturday                  899.6031
```

analyze ridership data by type and weekday

```
all_trips_v2 %>%
  mutate(weekday = wday(started_at, label = TRUE)) %>%  #creates weekday field using wday()
  group_by(member_casual, weekday) %>% #groups by member_casual and weekday
  summarise(number_of_rides = n()    #calculates the number of rides and average duration
  ,average_duration = mean(ride_length)) %>% # calculates the average duration
  arrange(member_casual, weekday)    # sorts
```

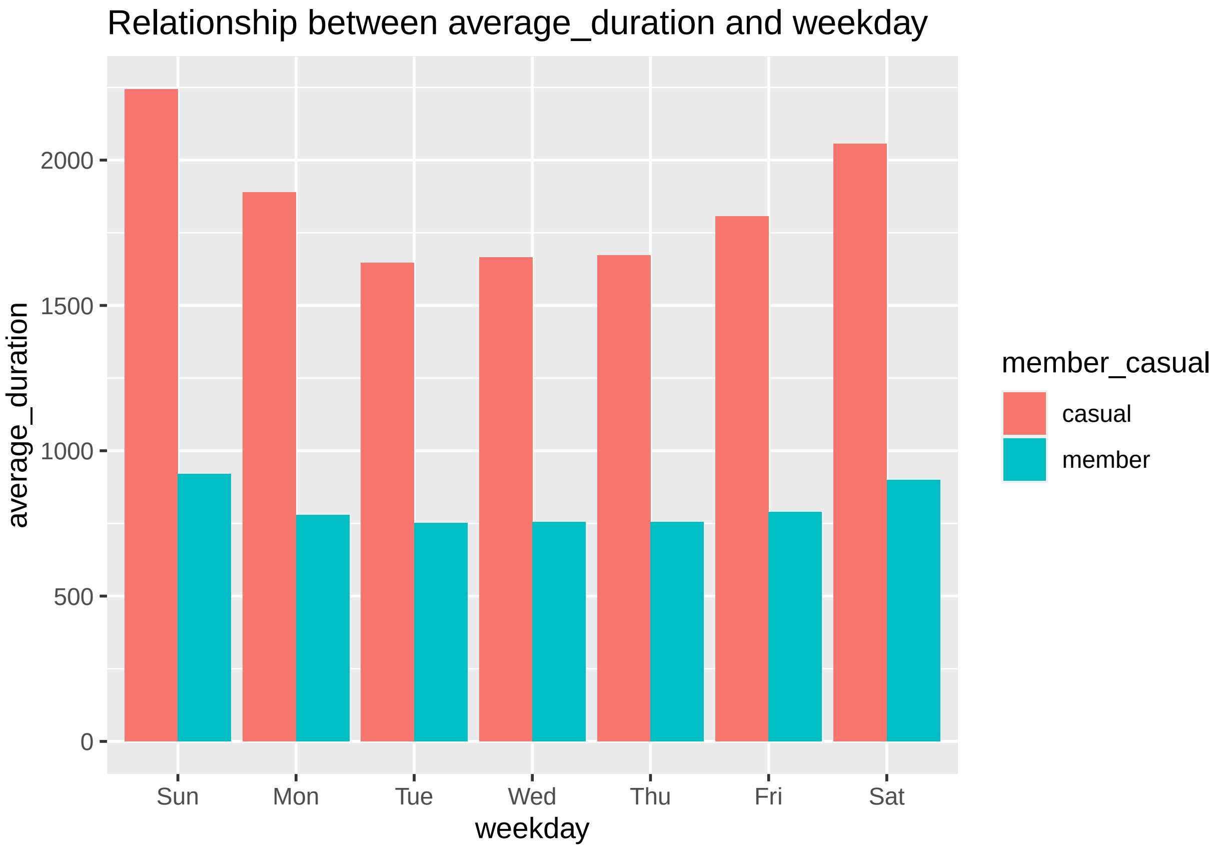
```
## `summarise()` has grouped output by 'member_casual'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 14 x 4
## # Groups:   member_casual [2]
##    member_casual weekday number_of_rides average_duration
##    <chr>         <ord>             <int>            <dbl>
##  1 casual        Sun              482801            2244.
##  2 casual        Mon              292993            1889.
##  3 casual        Tue              276371            1646.
##  4 casual        Wed              286400            1666.
##  5 casual        Thu              293632            1673.
##  6 casual        Fri              364277            1806.
##  7 casual        Sat              550008            2057.
##  8 member        Sun              387717             921.
##  9 member        Mon              439428             778.
## 10 member        Tue              490095             751.
## 11 member        Wed              499901             755.
## 12 member        Thu              475330             754.
## 13 member        Fri              453108             788.
## 14 member        Sat              431326             900.
```

Let’s create a visualization for average duration

```
all_trips_v2 %>%
  mutate(weekday = wday(started_at, label = TRUE)) %>%
  group_by(member_casual, weekday) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_casual, weekday)  %>%
  ggplot(aes(x = weekday, y = average_duration, fill = member_casual)) +
  geom_col(position = "dodge") +
labs(title = "Relationship between average_duration and weekday")
```

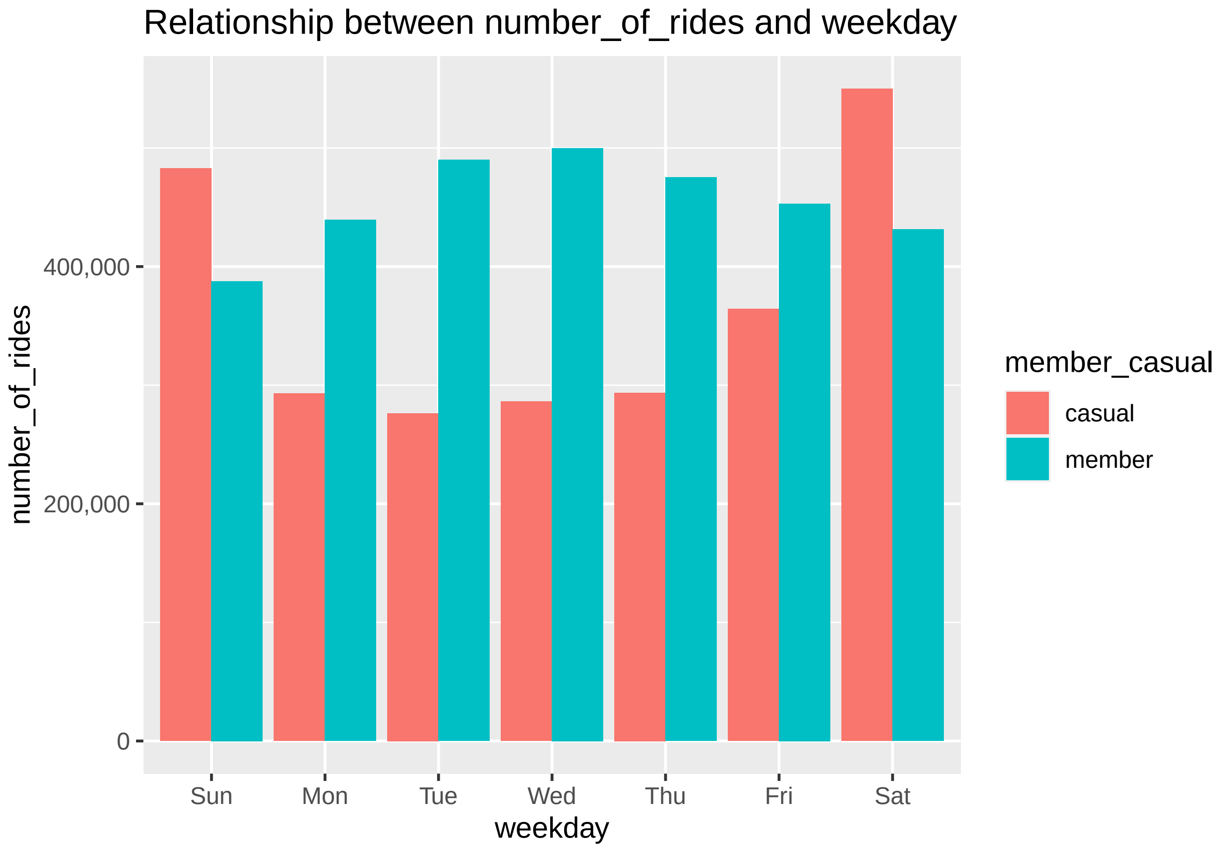
```
## ‘summarise()‘ has grouped output by ’member_casual’. You can override using the
## ‘.groups‘ argument.
```



Average duration of rides generally increases on Sunday and Saturday as compare to weekdays.

```
all_trips_v2 %>%
  mutate(weekday = wday(started_at, label = TRUE)) %>%
  group_by(member_casual, weekday) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_casual, weekday)  %>%
  ggplot(aes(x = weekday, y = number_of_rides, fill = member_casual)) + scale_y_continuous(labels =comma
  geom_col(position = "dodge") +
labs(title = "Relationship between number_of_rides and weekday")
```

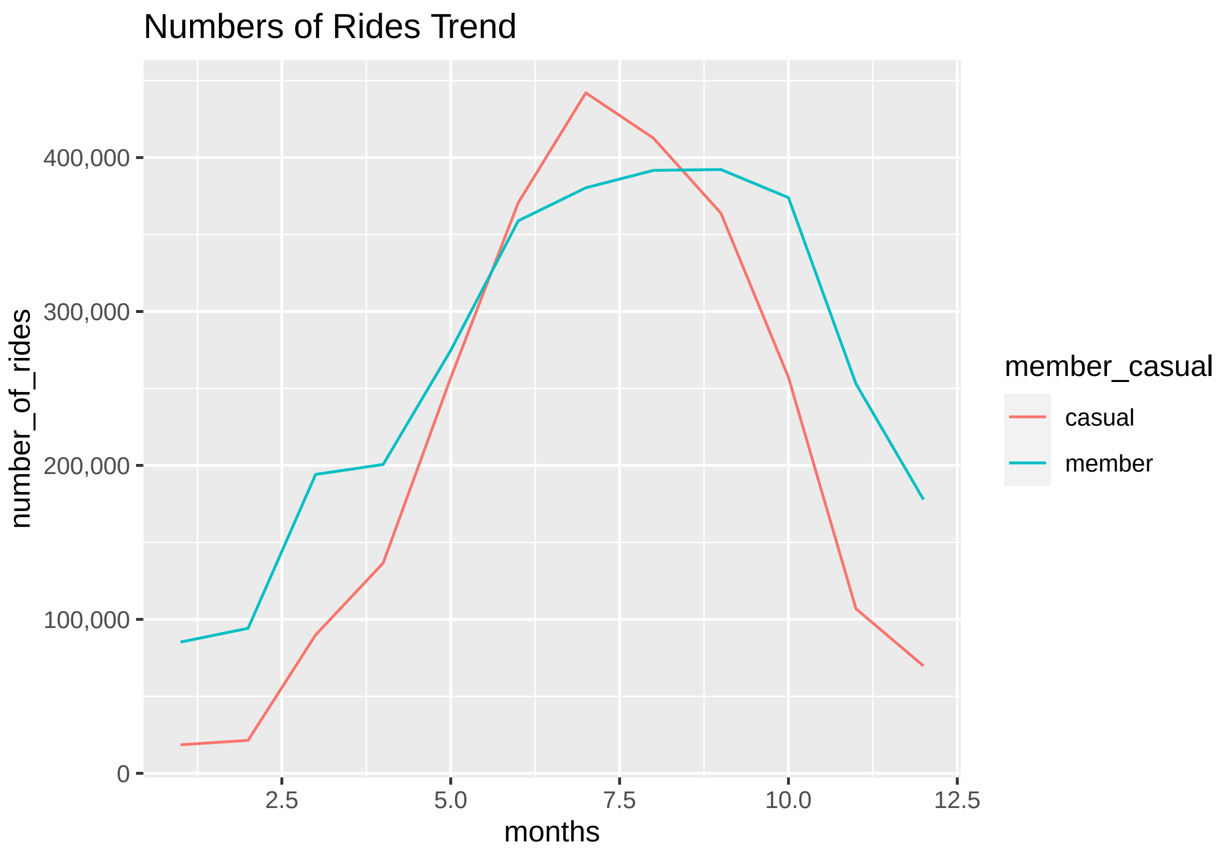
```
## ‘summarise()‘ has grouped output by ’member_casual’. You can override using the
## ‘.groups‘ argument.
```



# The number of rides are maximum on weekends for casual riders as compare to annual members.

```
all_trips_v2 %>%
  mutate(months = month(date, label = FALSE)) %>%
  group_by(member_casual, months) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_casual, months)  %>%
  ggplot(aes(x = months, y = number_of_rides, fill = member_casual, color=member_casual)) +
  scale_y_continuous(labels = comma) +
  geom_line() +
labs(title = "Numbers of Rides Trend")
```

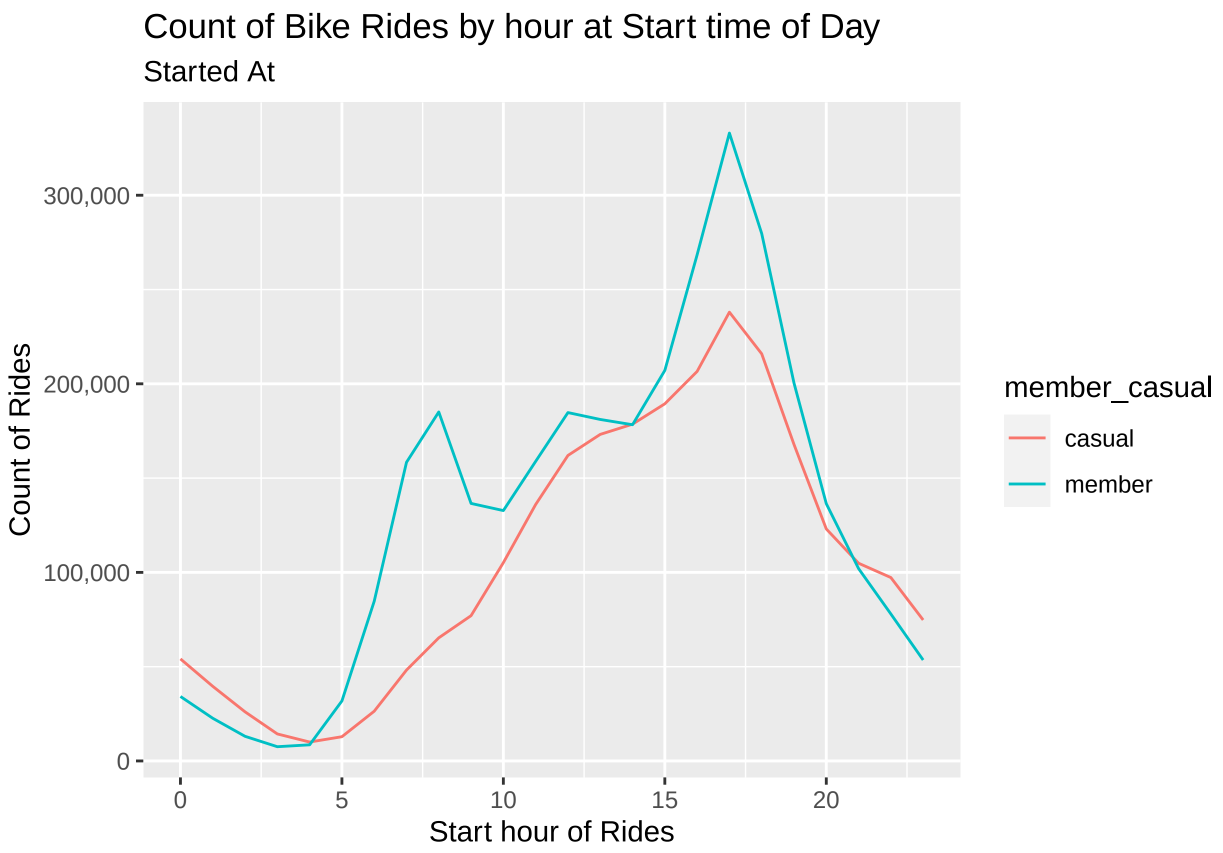
```
## ‘summarise()‘ has grouped output by ’member_casual’. You can override using the
## ‘.groups‘ argument.
```



# Number of rides of casuals increases between 6 to 8 months of the year.

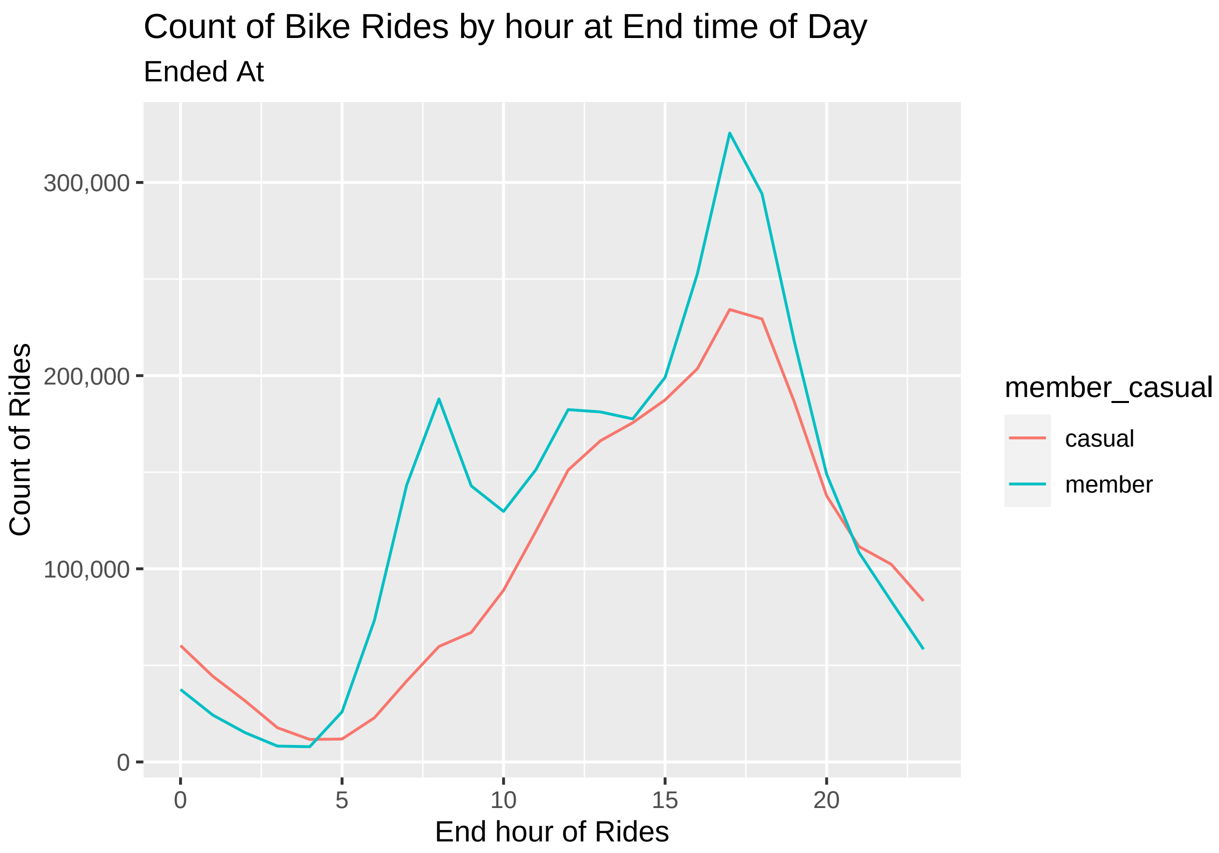
```
all_trips_v2 %>%
  mutate(start_hour = hour(all_trips_v2$started_at)) %>%
  group_by(member_casual,start_hour) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_casual, start_hour)  %>%
  ggplot(aes(x=start_hour,y=number_of_rides, fill = member_casual, color = member_casual)) +
  geom_line() +
  scale_y_continuous(labels = comma) +
  labs(title = "Count of Bike Rides by hour at Start time of Day", x = "Start hour of Rides", y="Count o
```

```
## `summarise()` has grouped output by 'member_casual'. You can override using the
## `.groups` argument.
```



```
all_trips_v2 %>%
  mutate(end_hour = hour(all_trips_v2$ended_at)) %>%
  group_by(member_casual,end_hour) %>%
  summarise(number_of_rides = n()
            ,average_duration = mean(ride_length)) %>%
  arrange(member_casual, end_hour)  %>%
  ggplot(aes(x=end_hour,y=number_of_rides, fill = member_casual, color = member_casual)) +
  geom_line() +
  scale_y_continuous(labels =comma) +
  labs(title = "Count of Bike Rides by hour at End time of Day", x = "End hour of Rides", y="Count of Ri
```

```
## ‘summarise()‘ has grouped output by ’member_casual’. You can override using the
## ‘.groups‘ argument.
```



Marketing memberships to casual_riders could begin on weekends from 4 PM to 6 PM.

Memberships to casual riders could be promoted with weekend-only memberships as number of rides surge on weekends.