Cyclistic Case Study

Sumaiya Mohammed Hanif

2025-07-15

# 1. Business Task

The business task is to analyze how annual members and casual riders use Cyclistic bikes differently. The goal is to identify actionable insights that can help convert casual riders into annual members and increase long-term revenue.

# 2. Data Sources

Data was obtained from the Divvy Bike Share system’s public trip data: - **Divvy 2019 Q1** CSV file - **Divvy 2020 Q1** CSV file

These datasets include ride IDs, timestamps, stations, and user types (member or casual). The data is publicly available and anonymized to protect user privacy.

# 3. Data Cleaning and Preparation

Below I show the steps taken to clean and prepare the data for analysis:

library(tidyverse)  
library(lubridate)

q1\_2019 <- read\_csv("Divvy\_Trips\_2019\_Q1 - Divvy\_Trips\_2019\_Q1.csv")  
q1\_2020 <- read\_csv("Divvy\_Trips\_2020\_Q1 - Divvy\_Trips\_2020\_Q1.csv")

q1\_2019 <- q1\_2019 %>%  
 rename(  
 ride\_id = trip\_id,  
 started\_at = start\_time,  
 ended\_at = end\_time,  
 start\_station\_id = from\_station\_id,  
 start\_station\_name = from\_station\_name,  
 end\_station\_id = to\_station\_id,  
 end\_station\_name = to\_station\_name,  
 member\_casual = usertype  
 ) %>%  
 mutate(  
 rideable\_type = "classic\_bike",  
 start\_lat = NA,  
 start\_lng = NA,  
 end\_lat = NA,  
 end\_lng = NA,  
 started\_at = ymd\_hms(started\_at),  
 ended\_at = ymd\_hms(ended\_at)  
 ) %>%  
 select(  
 ride\_id, rideable\_type, started\_at, ended\_at,  
 start\_station\_name, start\_station\_id,  
 end\_station\_name, end\_station\_id,  
 start\_lat, start\_lng, end\_lat, end\_lng,  
 member\_casual  
 )

q1\_2019 <- mutate(q1\_2019, ride\_id = as.character(ride\_id)  
 ,rideable\_type = as.character(rideable\_type))

q1\_2020 <- q1\_2020 %>%  
 mutate(  
 started\_at = ymd\_hms(started\_at),  
 ended\_at = ymd\_hms(ended\_at))

all\_trips <- bind\_rows(q1\_2019, q1\_2020)

all\_trips <- all\_trips %>%  
 drop\_na() %>%  
 mutate(  
 ride\_length = as.numeric(difftime(ended\_at, started\_at, units = "mins")),  
 day\_of\_week = wday(started\_at, label = TRUE, abbr = FALSE)  
 ) %>%  
 filter(ride\_length >= 1, ride\_length < 1440)

# 4. Analysis Summary

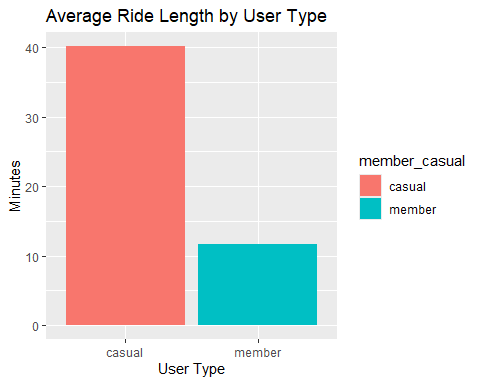
summary\_stats <- all\_trips %>%  
 group\_by(member\_casual) %>%  
 summarise(  
 mean\_ride\_length = mean(ride\_length),  
 median\_ride\_length = median(ride\_length),  
 max\_ride\_length = max(ride\_length),  
 min\_ride\_length = min(ride\_length)  
 )

rides\_by\_day <- all\_trips %>%  
 group\_by(member\_casual, day\_of\_week) %>%  
 summarise(num\_rides = n(), avg\_duration = mean(ride\_length)) %>%  
 arrange(member\_casual, day\_of\_week)

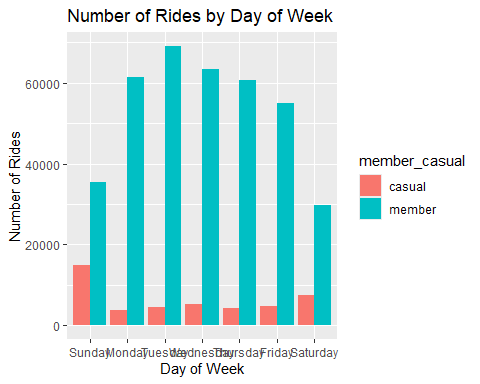
## `summarise()` has grouped output by 'member\_casual'. You can override using the  
## `.groups` argument.

# 5. Visualizations and Key Findings

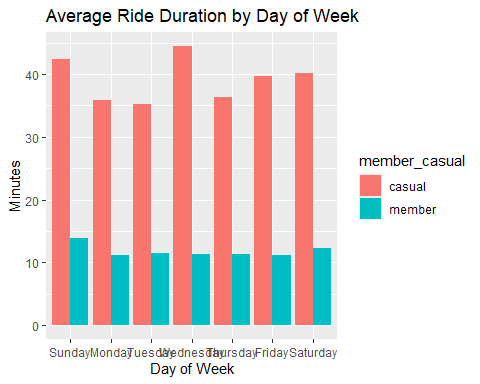
all\_trips %>%  
 group\_by(member\_casual) %>%  
 summarise(avg\_ride\_length = mean(ride\_length)) %>%  
 ggplot(aes(x=member\_casual, y=avg\_ride\_length, fill=member\_casual)) +  
 geom\_col() +  
 labs(title="Average Ride Length by User Type", x="User Type", y="Minutes")



rides\_by\_day %>%  
 ggplot(aes(x=day\_of\_week, y=num\_rides, fill=member\_casual)) +  
 geom\_col(position="dodge") +  
 labs(title="Number of Rides by Day of Week", x="Day of Week", y="Number of Rides")



rides\_by\_day %>%  
 ggplot(aes(x=day\_of\_week, y=avg\_duration, fill=member\_casual)) +  
 geom\_col(position="dodge") +  
 labs(title="Average Ride Duration by Day of Week", x="Day of Week", y="Minutes")



# 6. Recommendations

Based on the analysis, I propose the following recommendations:

Weekend Promotions: Offer discounted membership promotions targeted at casual users who primarily ride on weekends.

Commuter-Focused Advertising: Highlight cost savings and convenience of annual memberships for weekday commuters.

Personalized Offers: Target high-frequency casual riders with tailored email campaigns offering membership discounts or perks.

# 7. Exporting Data

Below is code to export cleaned data and analysis summaries for sharing or future use.

write\_csv(all\_trips, "cleaned\_all\_trips.csv")  
write\_csv(summary\_stats, "summary\_stats.csv")  
write\_csv(rides\_by\_day, "rides\_by\_day.csv")