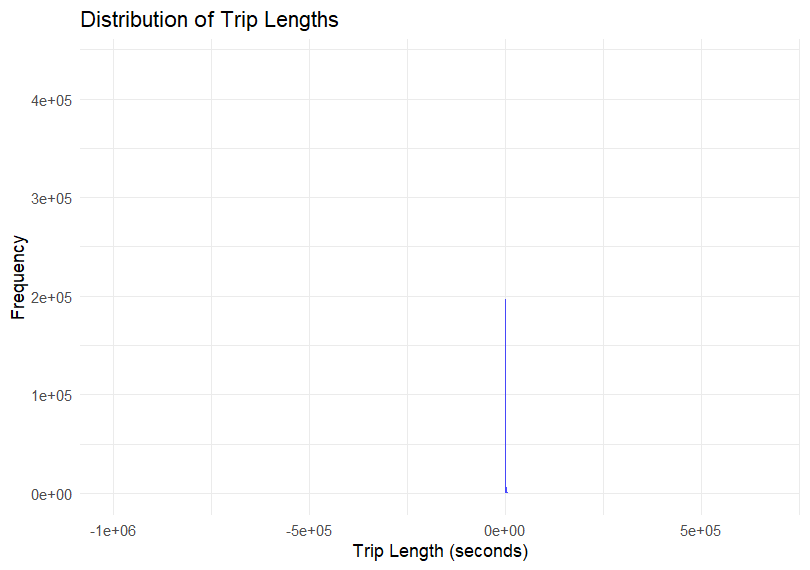
—------------------------------------------------------------------------------------------------------------------------------------------

ggplot(unique\_trips, aes(x = trip\_length)) +

geom\_histogram(binwidth = 60, fill = "blue", alpha = 0.7) +

labs(title = "Distribution of Trip Lengths", x = "Trip Length (seconds)", y = "Frequency") +

theme\_minimal()

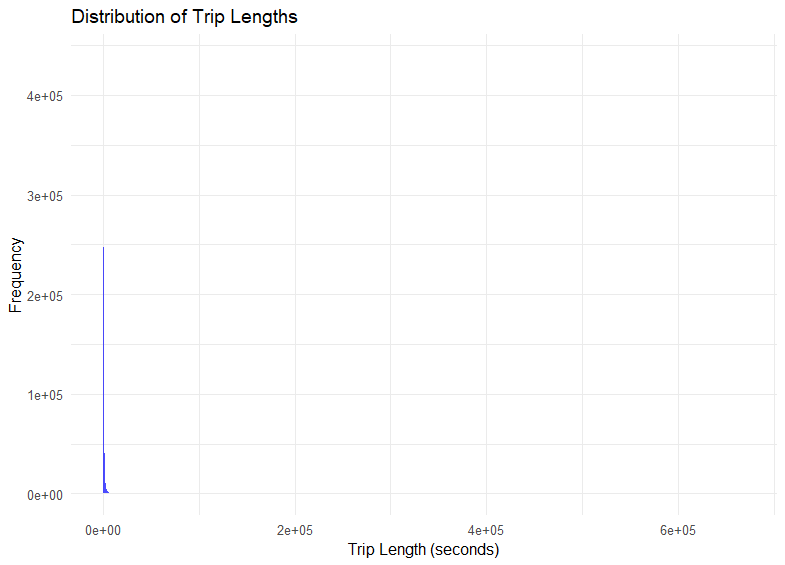


ggplot(filtered\_trips\_2, aes(x = trip\_length)) +

geom\_histogram(binwidth = 60, fill = "blue", alpha = 0.7) +

labs(title = "Distribution of Trip Lengths", x = "Trip Length (seconds)", y = "Frequency") +

theme\_minimal()



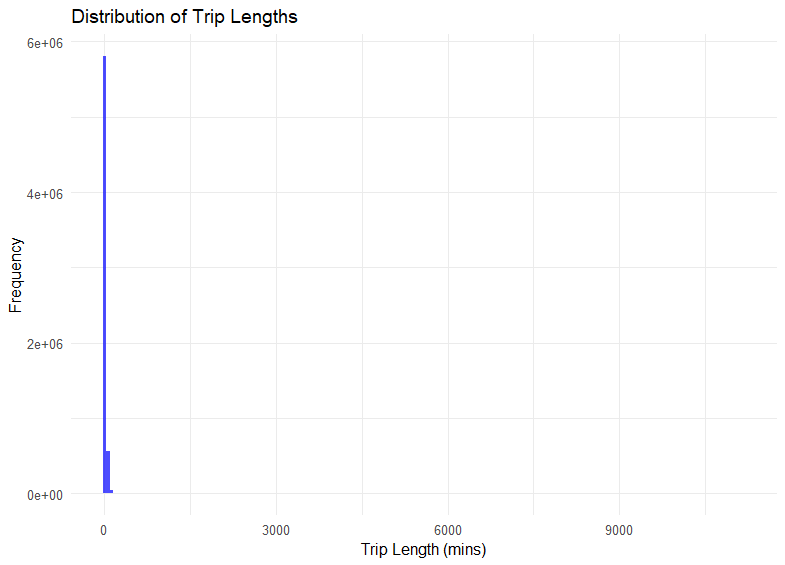
# Histogram of trip lengths in minutes

ggplot(filtered\_trips\_2, aes(x = trip\_length/60)) +

geom\_histogram(binwidth = 60, fill = "blue", alpha = 0.7) +

labs(title = "Distribution of Trip Lengths", x = "Trip Length (mins)", y = "Frequency") +

theme\_minimal()



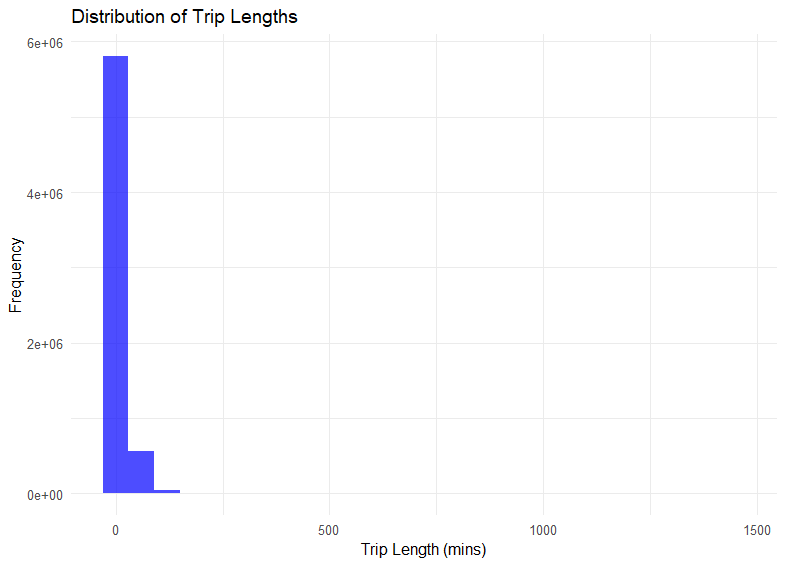
# Histogram of trip lengths less than 24 hours

ggplot(filtered\_trips\_2 %>% filter(trip\_length < 86400), aes(x = trip\_length/60)) +

geom\_histogram(binwidth = 60, fill = "blue", alpha = 0.7) +

labs(title = "Distribution of Trip Lengths", x = "Trip Length (mins)", y = "Frequency") +

theme\_minimal()



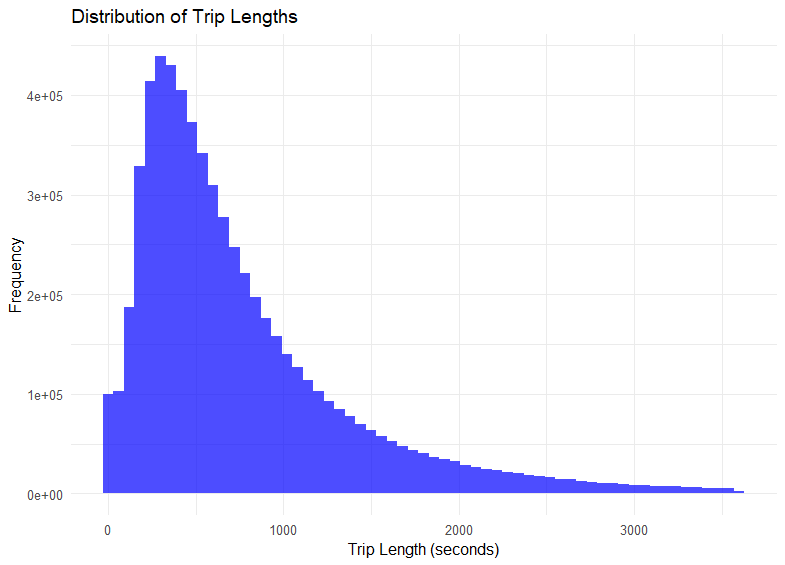
# Histogram of trip lengths less than 1 hour in seconds

ggplot(filtered\_trips\_2 %>% filter(trip\_length < 3600), aes(x = trip\_length)) +

geom\_histogram(binwidth = 60, fill = "blue", alpha = 0.7) +

labs(title = "Distribution of Trip Lengths", x = "Trip Length (seconds)", y = "Frequency") +

theme\_minimal()



—------------------------------------------------------------------------------------------------------------------------------------------

ggplot(all\_trips) +

geom\_density(aes(x = hour(started\_at), fill = "Start Time"), alpha = 0.5) +

geom\_density(aes(x = hour(ended\_at), fill = "End Time"), alpha = 0.5) +

labs(title = "Density of Trip Start and End Times",

x = "Hour of the Day",

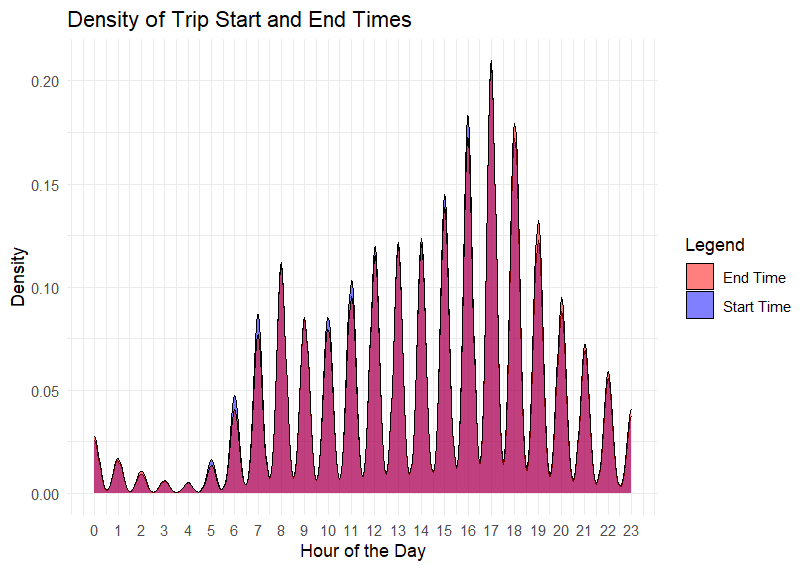
y = "Density",

fill = "Legend") +

scale\_x\_continuous(breaks = 0:23) +

scale\_fill\_manual(values = c("Start Time" = "blue", "End Time" = "red")) +

theme\_minimal()



# Extract hour and minute of the day from started\_at and ended\_at

filtered\_trips\_2 <- filtered\_trips\_2 %>%

mutate(

start\_time\_minute = hour(started\_at) \* 60 + minute(started\_at) + second(started\_at) / 60,

end\_time\_minute = hour(ended\_at) \* 60 + minute(ended\_at) + second(ended\_at) / 60

)

# Plotting density plots

ggplot(filtered\_trips\_2) +

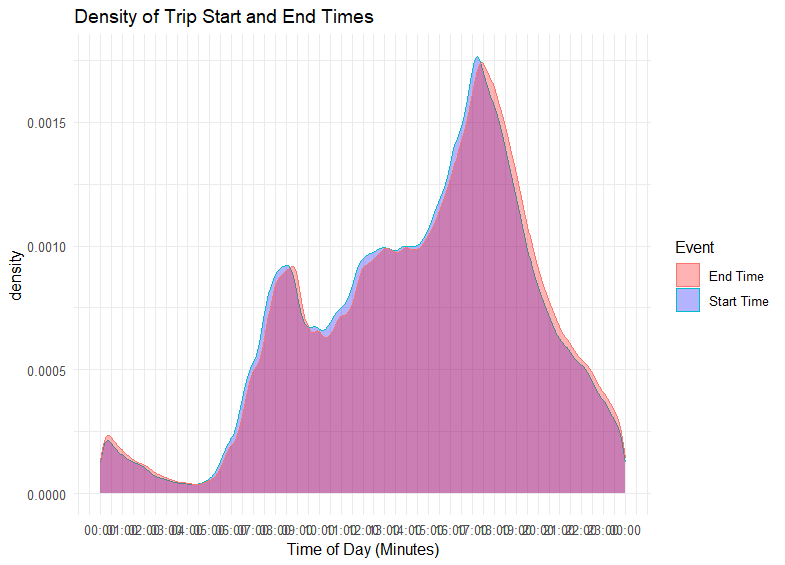
geom\_density(aes(x = start\_time\_minute, color = "Start Time"), fill = "blue", alpha = 0.3) +

geom\_density(aes(x = end\_time\_minute, color = "End Time"), fill = "red", alpha = 0.3) +

scale\_x\_continuous(name = "Time of Day (Minutes)", breaks = seq(0, 1440, by = 60), labels = function(x) format(as.POSIXct(x \* 60, origin = "1970-01-01", tz = "UTC"), "%H:%M")) +

labs(title = "Density of Trip Start and End Times", color = "Event") +

theme\_minimal()



—------------------------------------------------------------------------------------------------------------------------------------------

# Create pie charts for each rideable type

ggplot(membership\_by\_bike\_type, aes(x = "", y = percentage, fill = member\_casual)) +

geom\_bar(stat = "identity", width = 1) +

coord\_polar("y", start = 0) +

facet\_wrap(~ rideable\_type) +

labs(title = "Membership Distribution by Rideable Type",

fill = "User Type",

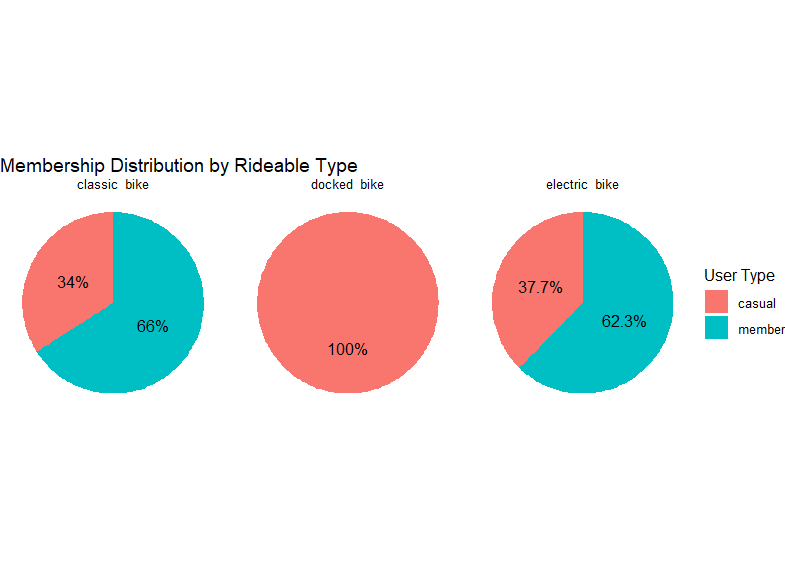
y = "Percentage") +

theme\_void() +

theme(legend.position = "right") +

geom\_text(aes(label = paste0(round(percentage, 1), "%")),

position = position\_stack(vjust = 0.5))



—-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

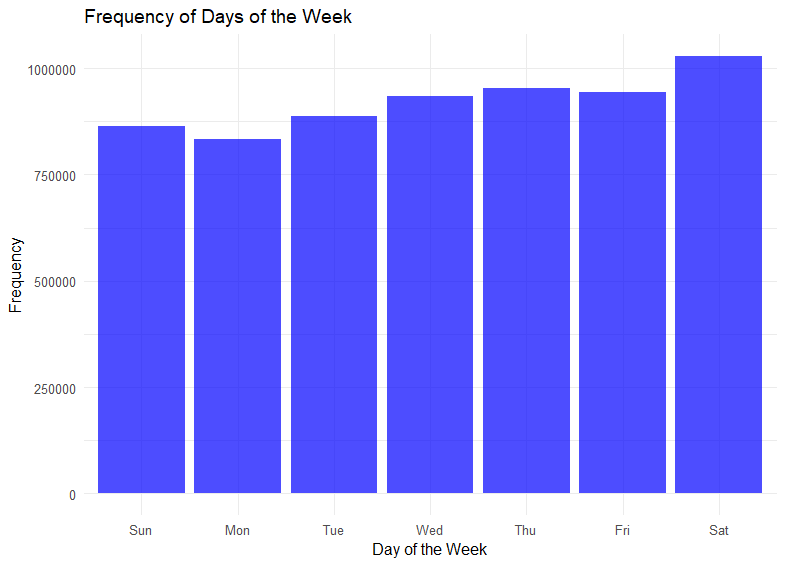
# histogram of days of week

ggplot(frequency\_day\_of\_week\_started, aes(x = day\_of\_week\_started, y = n)) +

geom\_bar(stat = "identity", fill = "blue", alpha = 0.7) +

labs(title = "Frequency of Days of the Week", x = "Day of the Week", y = "Frequency") +

theme\_minimal()



—-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

# map start\_lat and start\_lng to see if they cluster around specific points to indicate stations

ggplot(head(filtered\_trips\_2, 100), aes(x = start\_lat, y = start\_lng)) +

geom\_point() +

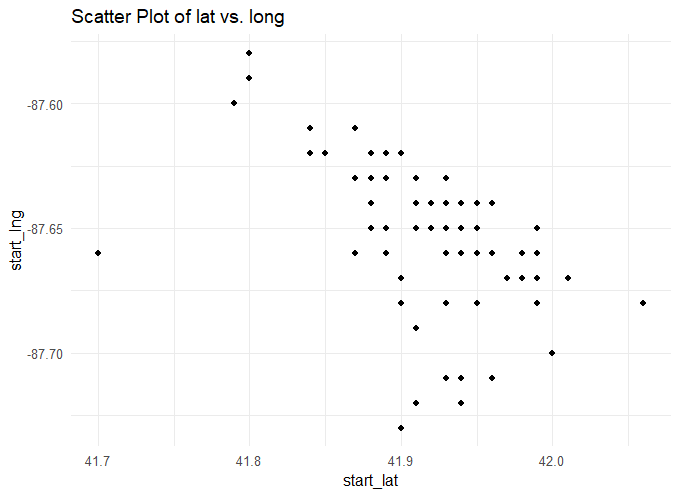
labs(title = "Scatter Plot of lat vs. long",

x = "start\_lat",

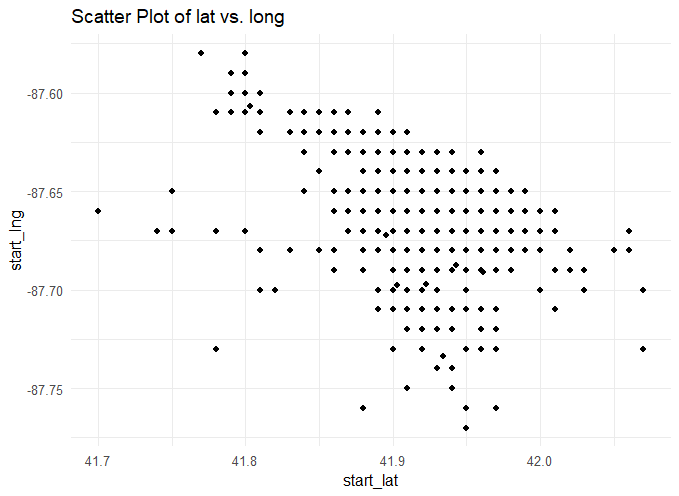
y = "start\_lng") +

theme\_minimal()

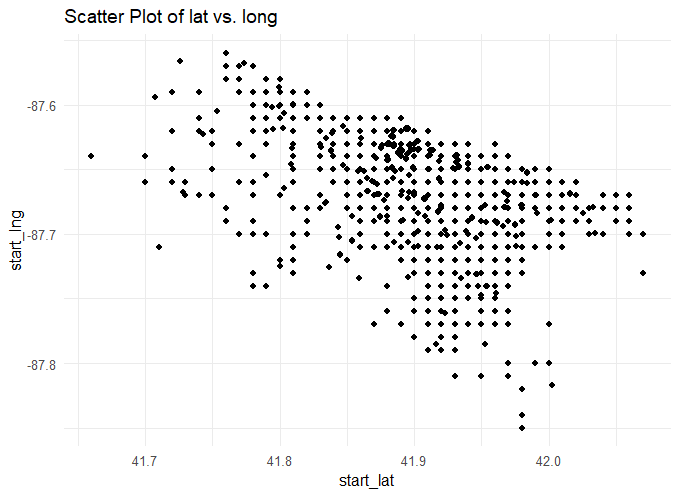
100 =



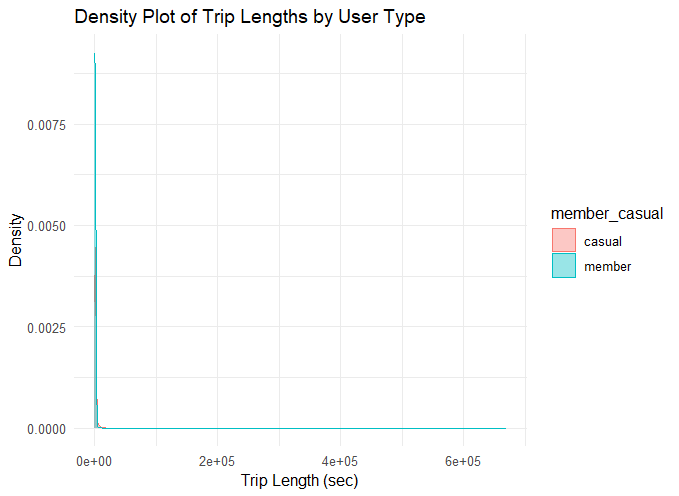
1000 =



10000 =



—-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



# group by user type density plot

ggplot(filtered\_trips\_4 %>% filter(trip\_length < 3600), aes(x = trip\_length, color = member\_casual, fill = member\_casual)) +

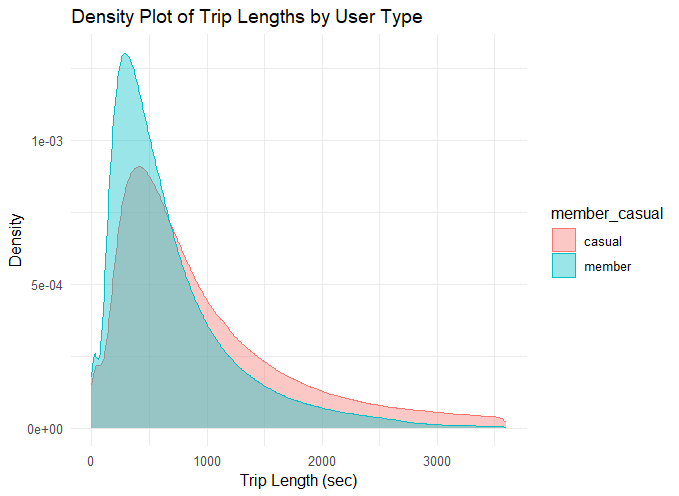
geom\_density(alpha = 0.4) +

labs(title = "Density Plot of Trip Lengths by User Type",

x = "Trip Length (sec)",

y = "Density") +

theme\_minimal()



—-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

# viz for the number of rides by rider type

filtered\_trips %>%

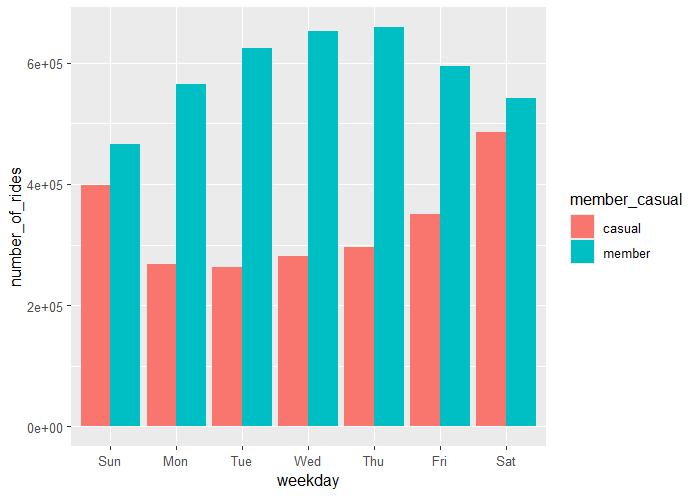
mutate(weekday = wday(started\_at, label = TRUE)) %>%

group\_by(member\_casual, weekday) %>%

summarise(number\_of\_rides = n(), average\_duration = mean(trip\_length)) %>%

arrange(member\_casual, weekday) %>%

ggplot(aes(x = weekday, y = number\_of\_rides, fill = member\_casual)) + geom\_col(position = "dodge")



# viz for average duration

filtered\_trips %>%

mutate(weekday = wday(started\_at, label = TRUE)) %>%

group\_by(member\_casual, weekday) %>%

summarise(number\_of\_rides = n(), average\_duration = mean(trip\_length)) %>%

arrange(member\_casual, weekday) %>%

ggplot(aes(x = weekday, y = average\_duration, fill = member\_casual)) + geom\_col(position = "dodge")

