

2024

BIG DATA 2 MIDTERM PROJECT

BICYCLE RENTAL ANALYSIS

OUR TEAM



HARDI MAJMUNDAR
N01498789



VRUSHALI PONKIA
N01530336



JANKI PATEL
N01533282



AYCAN LIZOR
N01534088

INTRODUCTION

- Analyzing a dataset using the provided parameters involves performing various operations and calculations on the columns Bike_id, Bike_type, Start_Time, End_Time, Start_Station, start_station_id, End_Station, end_station_id, Start_Lat, Start_Lng, End_Lat, End_Lng, and Rider_Type.



DATASETS OF CONTENT

- There are four data sets
- The total number of data is 2009713
- The total number of stations is 1475
- Bikes: classic, electric, and docked
- Members: member, casual
- The total number of ride is 2009713



WHAT WE DID



PROCESSING DATA

- Comparing data types of all 4 data frames.
- All of the data frames share the same structure, so integrate them.
- Convert blank value to NA



CLEANING DATA

- Removing the rows containing NA values.
- Remove columns start_station_id and end_station_id
- Rename column names for better understanding.
- Sort the data in ascending order by start time column.



ANALYSIS/VISUALIZATION

- Counting each bike type
- Percentage distribution of two group riders
- Stating the time of both riders on an hourly basis on a day
- Number of starts per hour of both the riders group
- Number of starts per hour every day

Story **TRIP DURATION**

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	6.00	10.00	16.19	18.00	12136.00



AFTER CLEANING DATA



Bike_ID	Bike_Type	Start_Time	End_Time
3F858645ACA7EB53	classic_bike	2023-03-01 0:00	2023-03-01 0:22
44563F7CDABE55B6	electric_bike	2023-03-01 0:04	2023-03-01 0:07
7491E3F211032D6A	classic_bike	2023-03-01 0:05	2023-03-01 0:10
92F4B2EE7B4CBA76	classic_bike	2023-03-01 0:06	2023-03-01 0:10
F73F4FE77F9B9A98	electric_bike	2023-03-01 0:06	2023-03-01 0:10
4E162F04422E6F57	electric_bike	2023-03-01 0:08	2023-03-01 0:12

Start_Station	End_Station	Start_Lat	Start_Lng	End_Lat
Racine Ave & Congress Pkwy	Wabash Ave & Grand Ave	41.87464 -	87.65703	41.89147
Public Rack -	Prairie Ave & 47th St Cottage	41.81000 -	87.62	4
Ellis Ave & 60th St	Harper Ave & 59th St	41.78510 -	87.60107	41.78794
Wells St & Elm St	Franklin St & Chicago Ave	41.90322 -	87.63432	41.89675
Shields Ave & 31st St	State St & 33rd St	41.83850 -	87.63532	41.83473
Daley Center Plaza	Orleans St & Hubbard St	41.88420 -	87.62972	41.89003

Bike_ID	Bike_Type	Start_Time	End_Time
3F858645ACA7EB53	classic_bike	2023-03-01 0:00	2023-03-01 0:22
44563F7CDABE55B6	electric_bike	2023-03-01 0:04	2023-03-01 0:07
7491E3F211032D6A	classic_bike	2023-03-01 0:05	2023-03-01 0:10
92F4B2EE7B4CBA76	classic_bike	2023-03-01 0:06	2023-03-01 0:10
F73F4FE77F9B9A98	electric_bike	2023-03-01 0:06	2023-03-01 0:10
4E162F04422E6F57	electric_bike	2023-03-01 0:08	2023-03-01 0:12

ANALYSIS



Counting of each rider type

member 968349

casual 553703



counting of the bike types used by the two groups of riders.

	Rider_Type	Bike_Type	n
1	member	classic_bike	582301
2	member	electric_bike	386048
3	casual	classic_bike	276317
4	casual	electric_bike	238575
5	casual	docked_bike	38811



Percentage distribution of two groups of riders.

	Rider_Type	n
1	member	63.62128
2	casual	36.37872

ANALYSIS

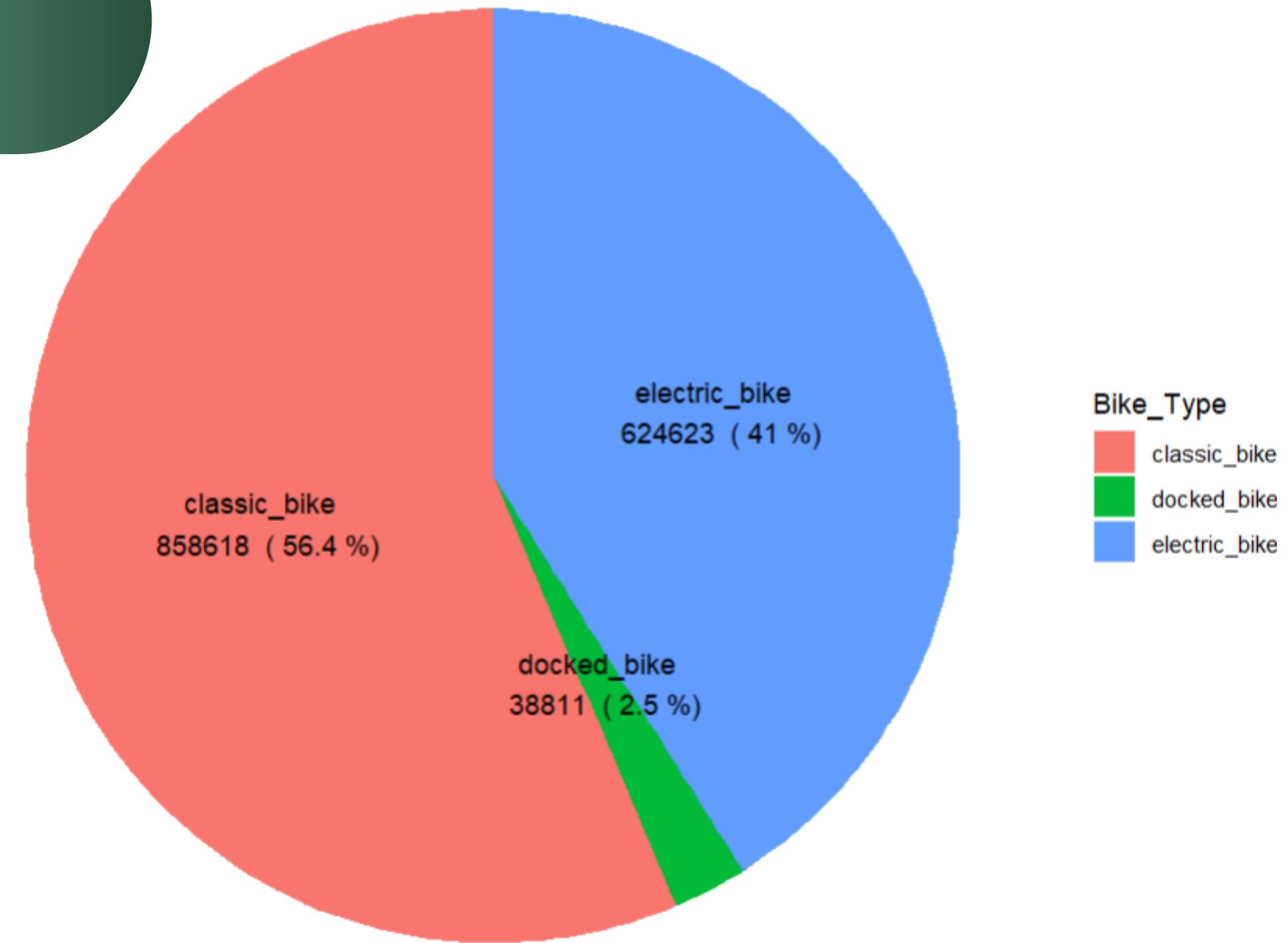


Percentage distribution of
bike type of two groups of
riders.

	Bike_Type	casual	member
1	classic_bike	49.9	60.1
2	docked_bike	7.01	0
3	electric_bike	43.1	39.9

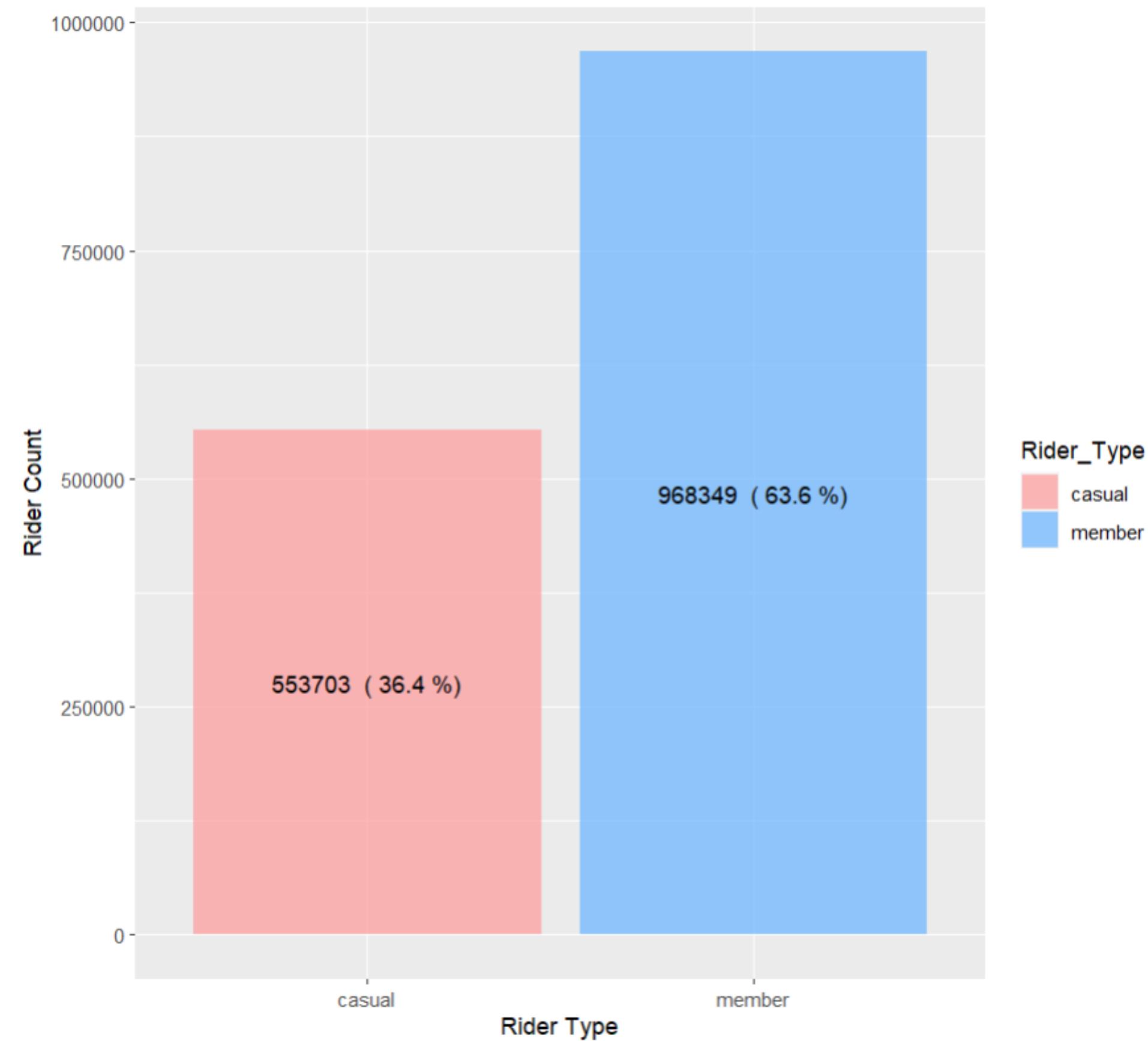
VISUALIZATION

DISTRIBUTION OF BIKE TYPES

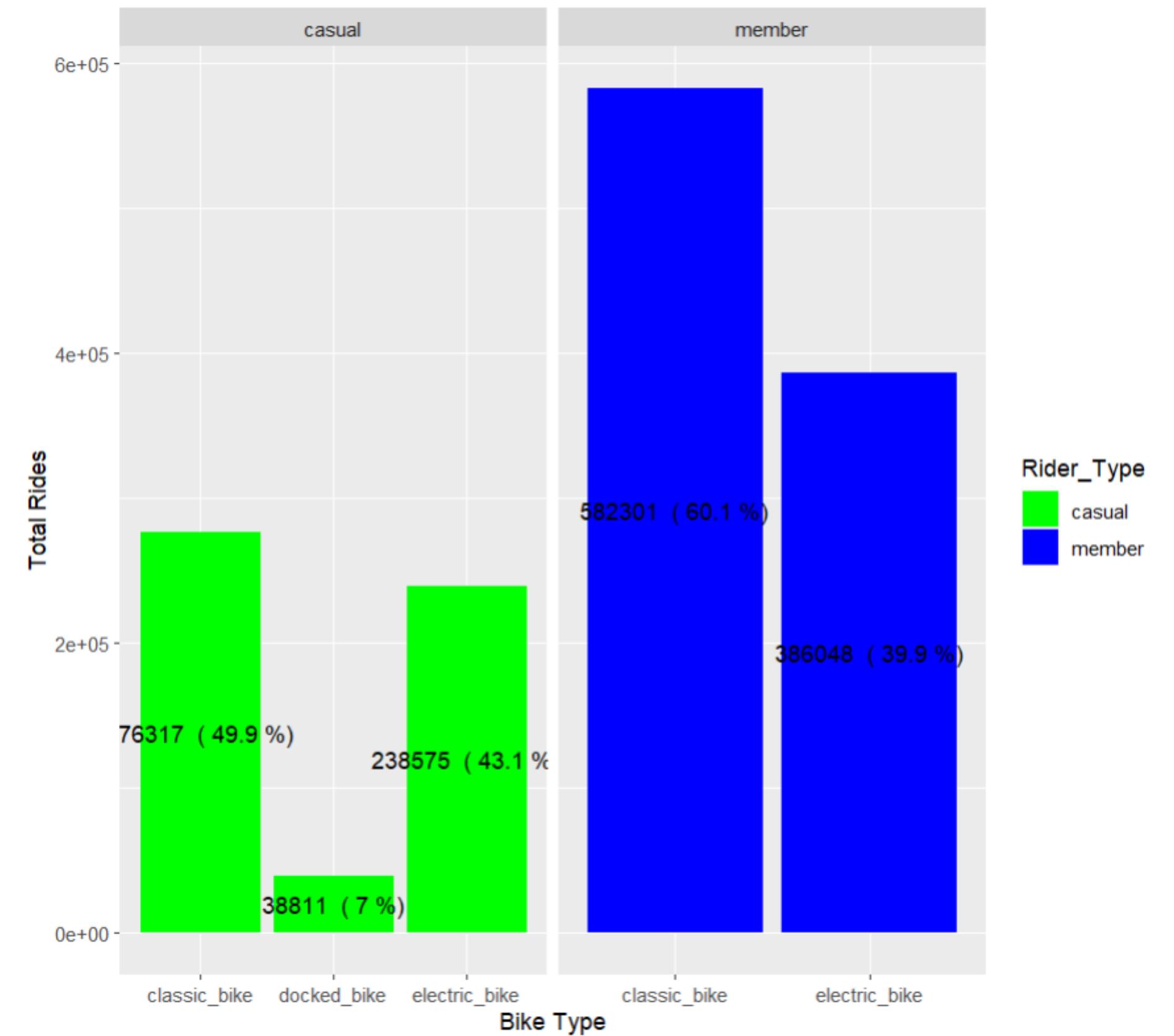


VISUALIZATION

DISTRIBUTION OF RIDER TYPES



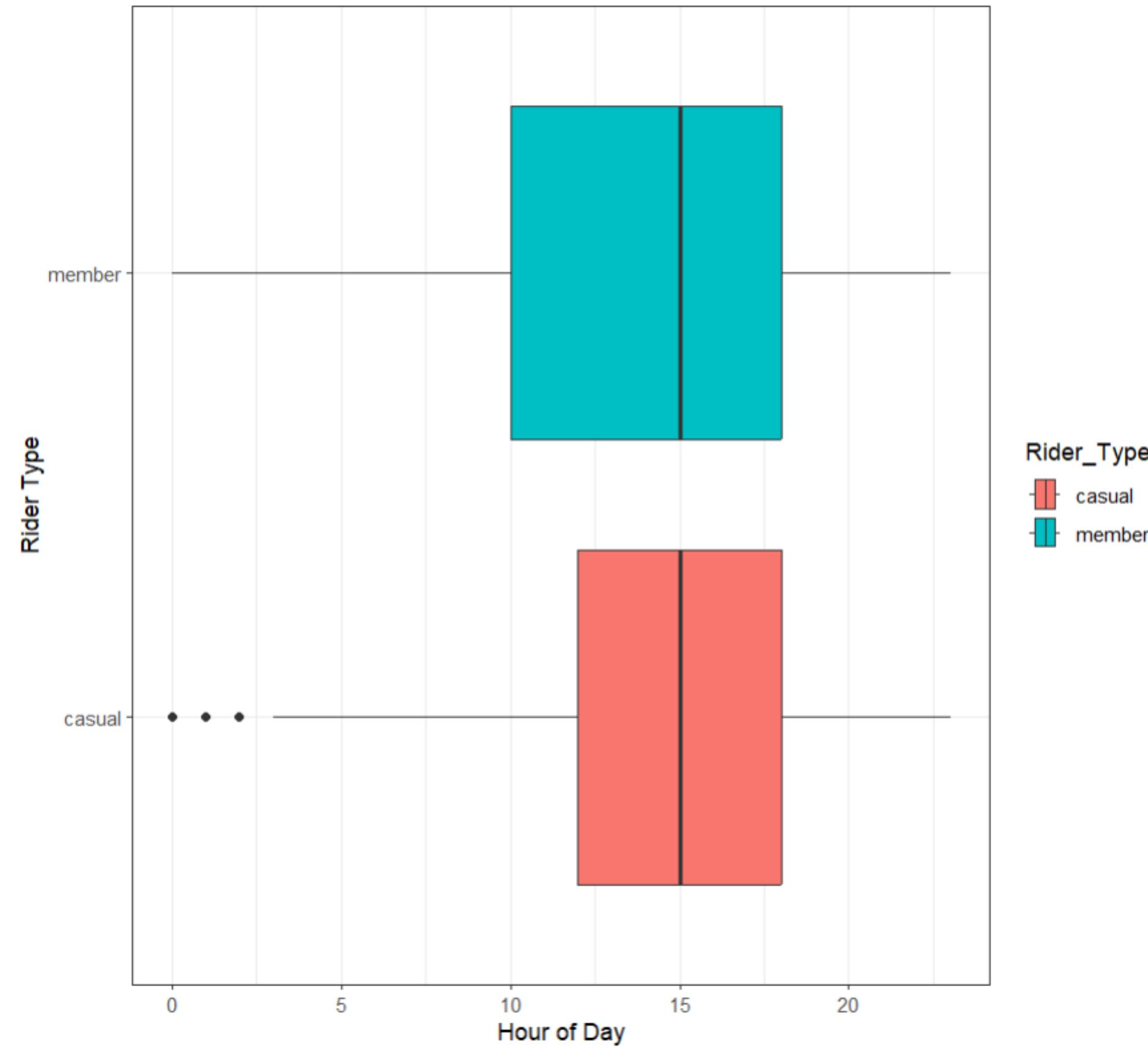
VISUALIZATION



BIKE TYPE PREFERENCES

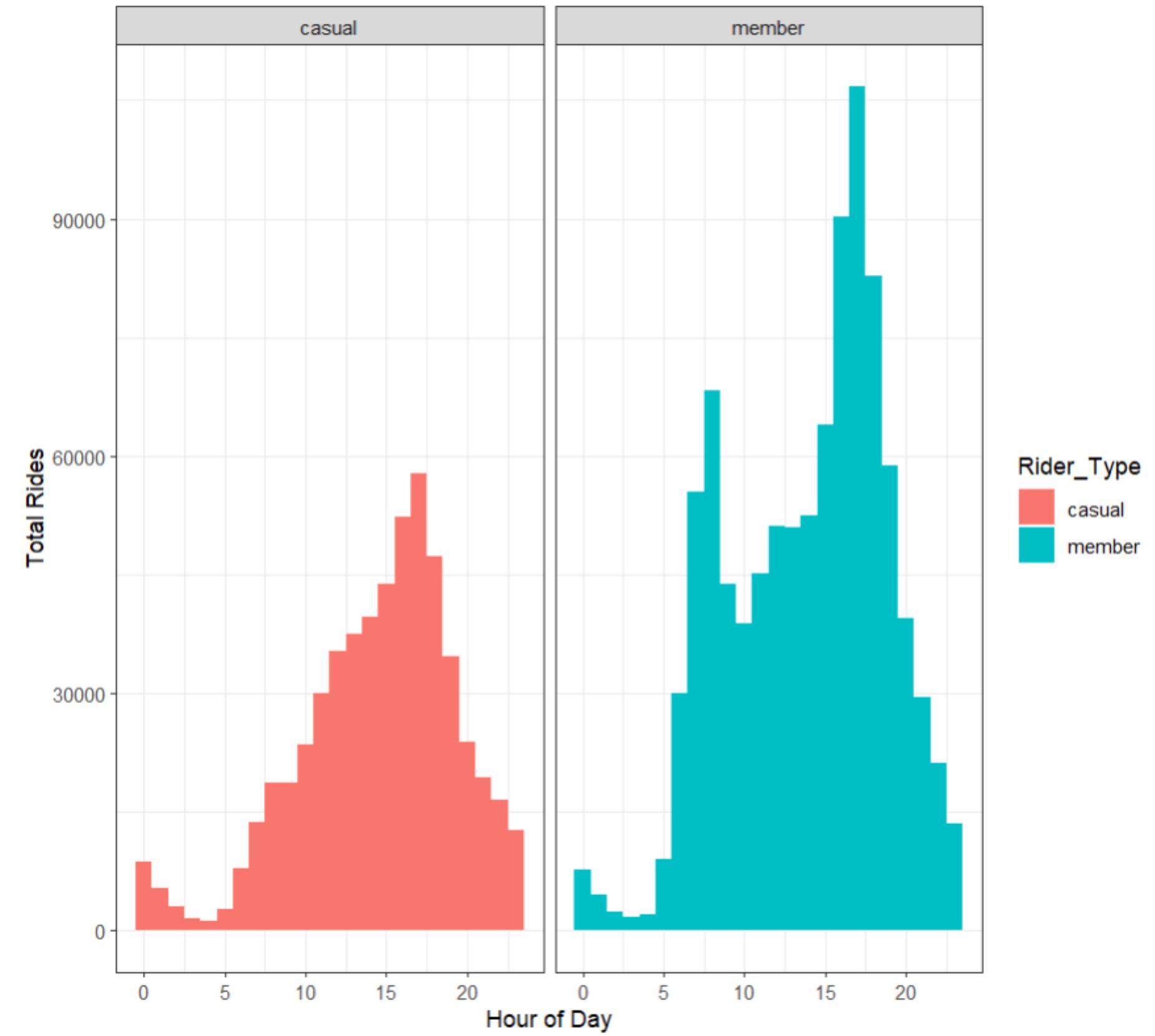
VISUALIZATION

START TIME TENDENCIES



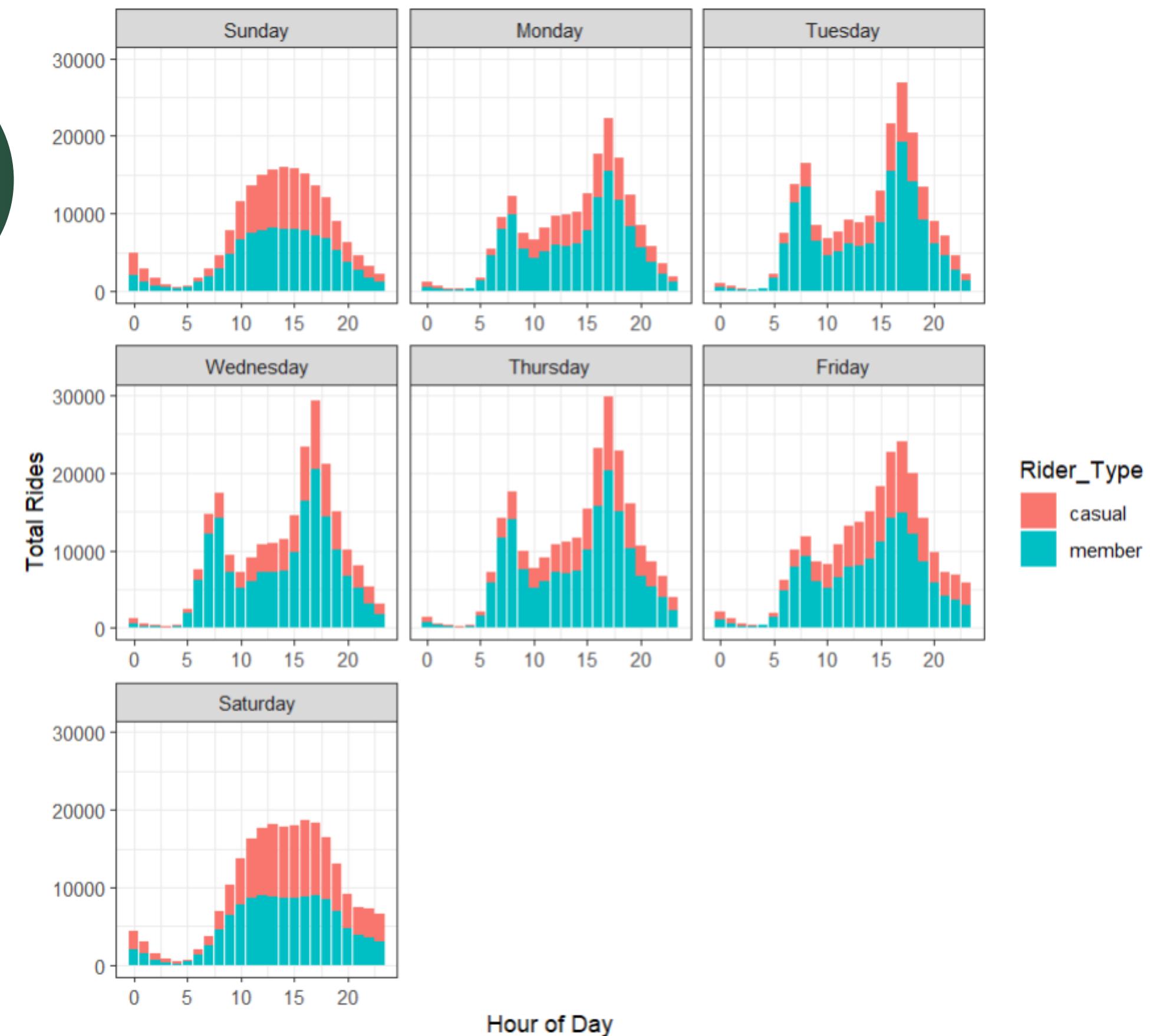
VISUALIZATION

NUMBER OF START PER HOUR FOR BOTH RIDERS GROUP



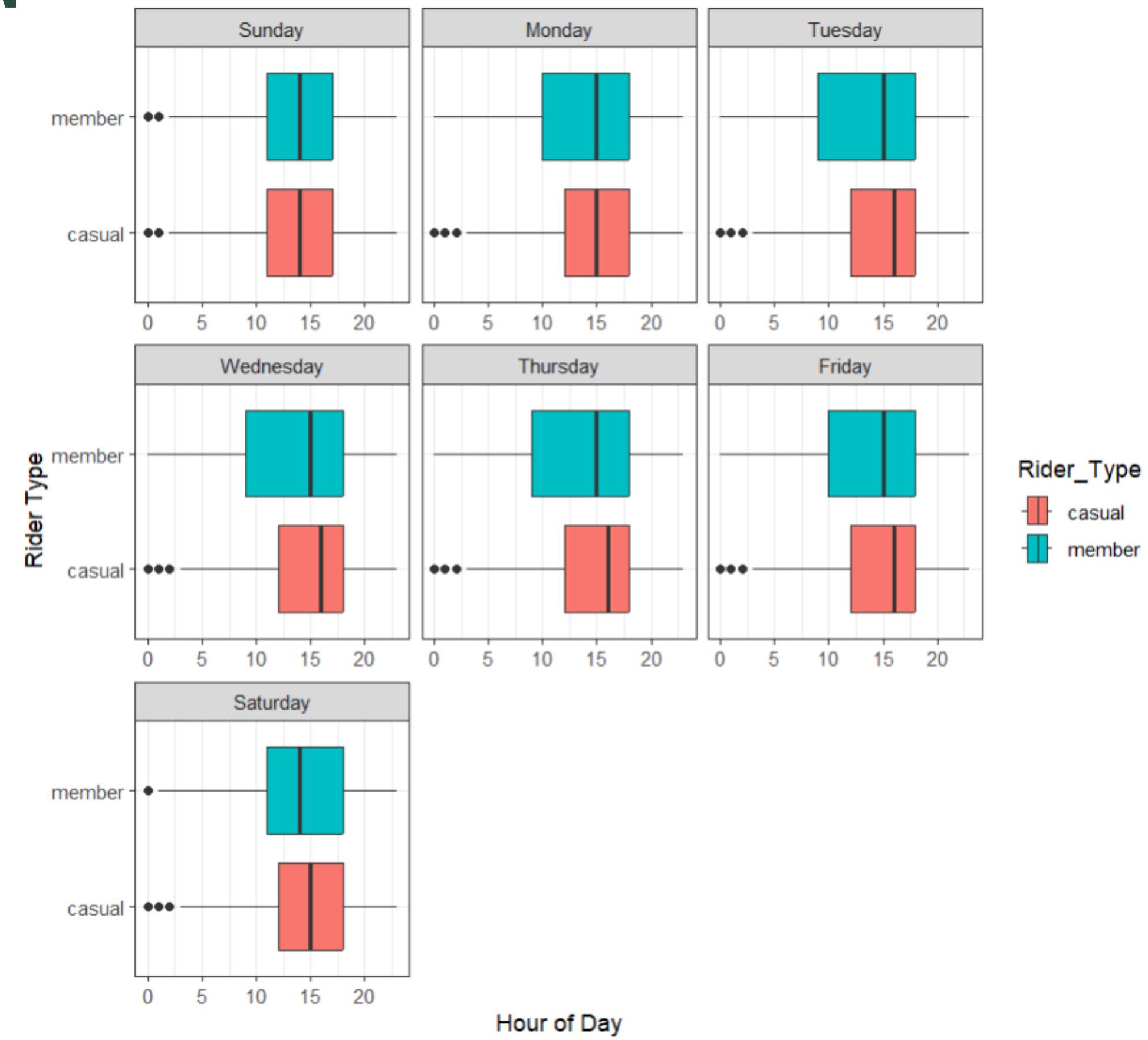
VISUALIZATION

NUMBER OF START PER
HOUR EVERYDAY

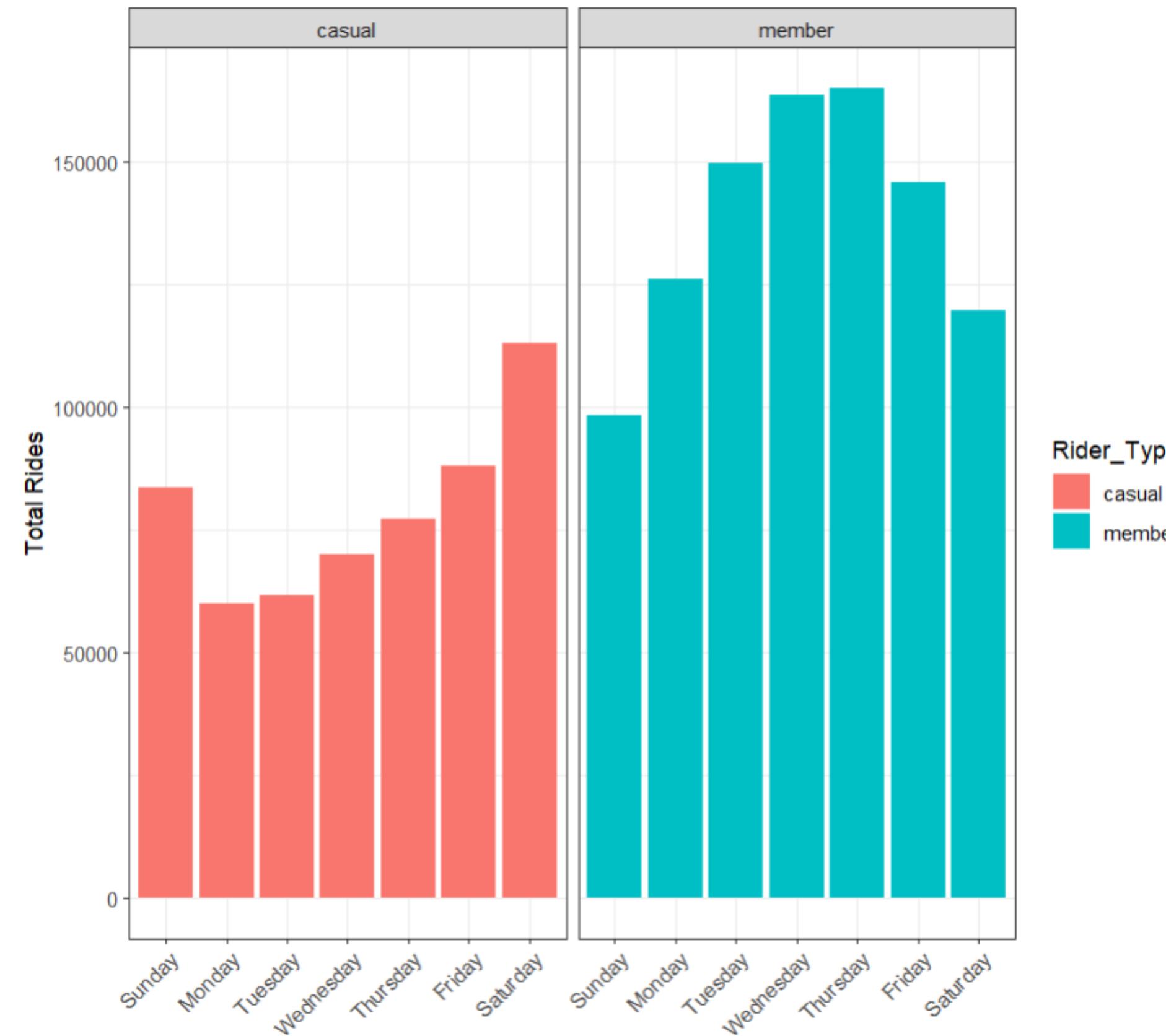


VISUALIZATION

STATING TIME OF BOTH RIDERS GROUP EVERYDAY

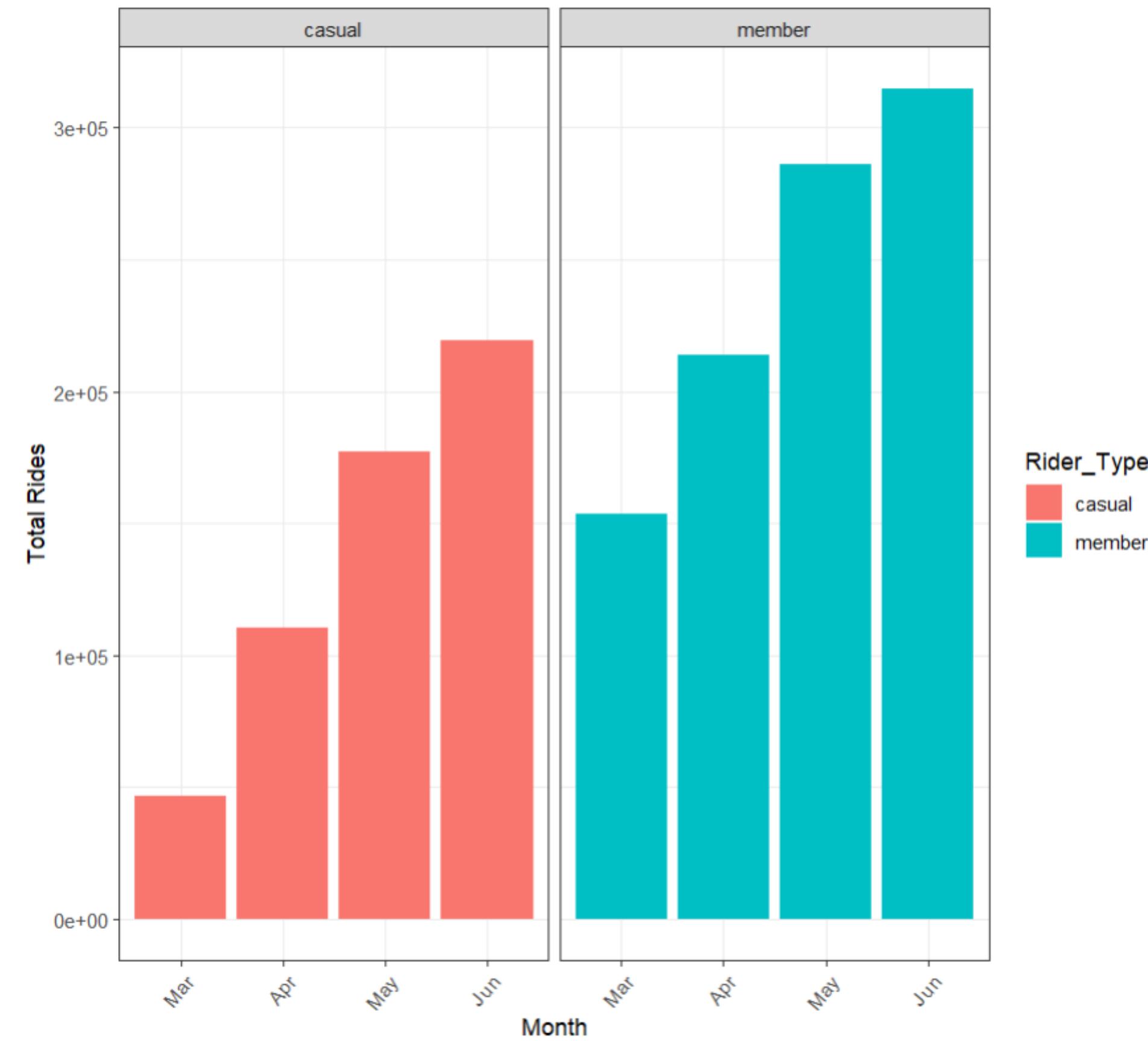


VISUALIZATION



**TOTAL RIDERS PER DAY
RELATIVE TO EACH RIDE TYPE**

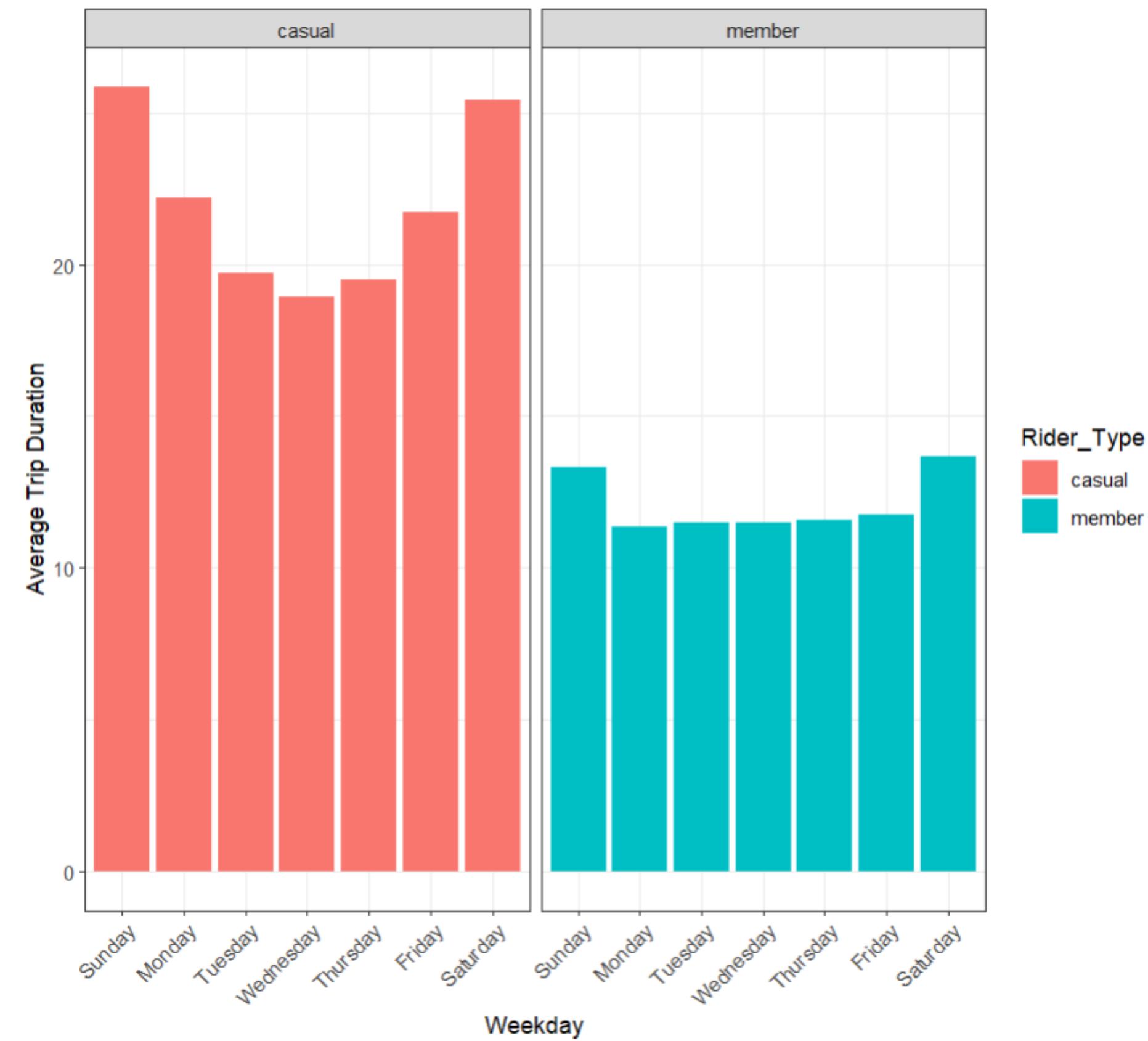
VISUALIZATION



TOTAL RIDERS PER MONTH
RELATIVE TO EACH RIDE TYPE

VISUALIZATION

AVERAGE TRIP DURATION
EVERYDAY



CONCLUSION

ANALYSIS BASED ON VSUALIZATION

- 1. The majority of annual members use the Cyclist Bike-Share services to commute to and from work as indicated by how consistent the start times are and the predictability of their respective average trip duration. Casual riders use these services for other casual purposes because of their longer average trip duration.
- 2. Also due to the short trip duration of annual riders, they may be living in city areas compared to casual riders(suburban areas).



THANK YOU