

Cyclist_analytics

Kondapudi Santhosh Kumar

15/02/2022

Analysis Summary(Part of Google Data Analytics Certification)

- From this analysis The people Who subscribed for **Annual membership** are more beneficial than casual riders.
- Major amount of riders booked by their Annual membership than normal
- Finally, **Average duration per ride** for a casual rider is almost **three times** of mean of ride duration for annual member every day.

Task Performed

Cyclistic marketing team needs to design marketing strategy to convert casual riders into annual members as it is believed that maximizing the number of annual members will be the key to future growth.

Source of Data

- The cyclistic data is an open source data, that can be accessed from the link <https://divvy-tripdata.s3.amazonaws.com/index.html> provided by the Coursera platform.
- From the Abundance Quantity of data I had chosen the **Past 12 months Data**
- Every month Data Contains **13 Variables** *Data is reliable, original, comprehensive, cited and current.

Cleaning and Transforming Data

- Merged all the 12 months data in to a Single Data frame.
- Removed the Nulls and Empty cells from the Data Frame.
- Also Removed Some attributes To create new reliable dataset.

Here goes the data analysis

Firstly loading the downloaded libraries of R

Importing the last one year data(12 months Data)

Combining all the 12 Months Data into a Single Dataframe named
“bike_rides”

Once the Primary dataset is ready,The Cleaning process will begins...

Removing the empty Rows and Columns from the Dataframe

Filtering the Date,time and Creating the Useful Data

Creating the Starting and Ending time

Converting the Hours time into Minutes

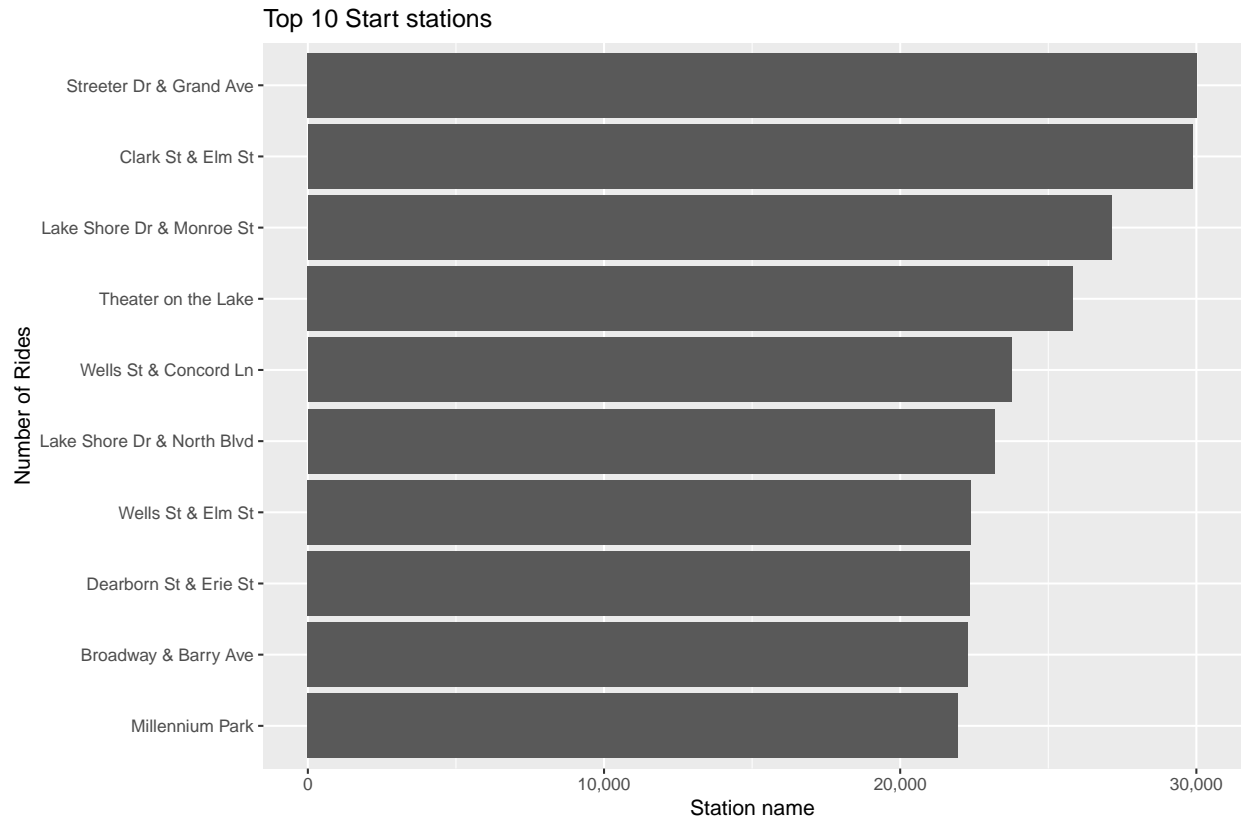
Creating the second Dataframe with Required attributes

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##           2      383    1422    2415    3320    15459

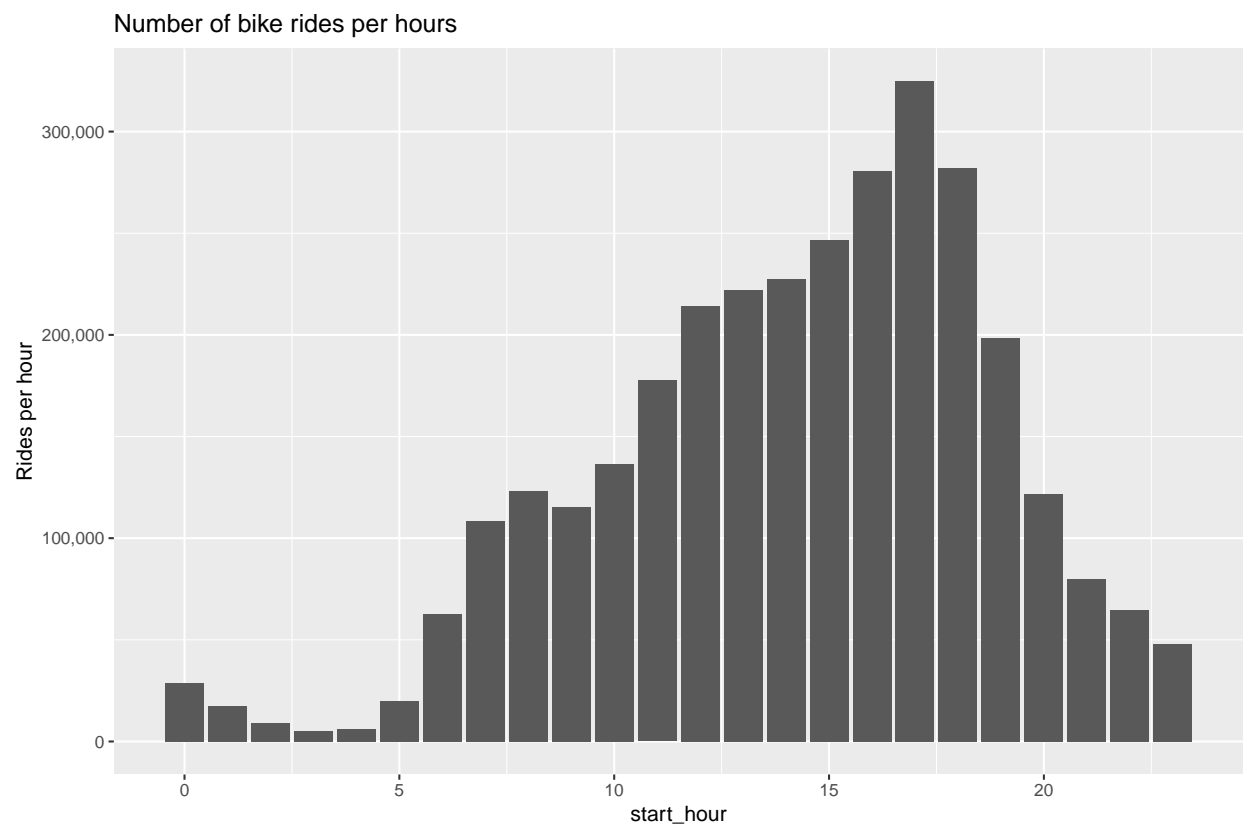
## bike_rides2$start_hour
##           0           1           2           3           4           5           6           7           8           9          10
##  28756  17188   8850   5100   5961  19935  62869 108084 122873 115188 136261
##          11          12          13          14          15          16          17          18          19          20          21
## 177798 214199 222006 227267 246678 280239 324640 282093 198139 121632  80046
##          22          23
##  64732  47773
```

Visualizing to Gain more insight

Plotting Bargraph for the Top 10 Stations based on Riders count

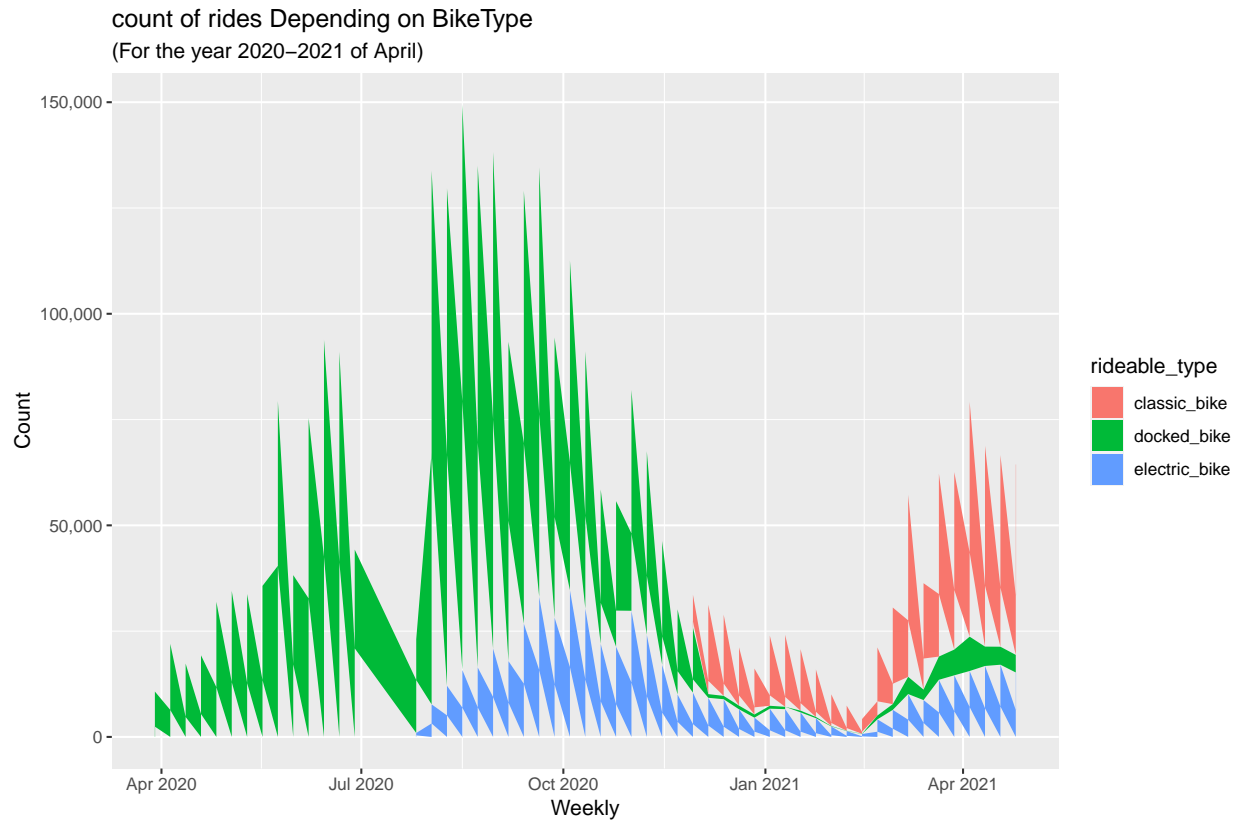


Plotting The Riders count per Hour

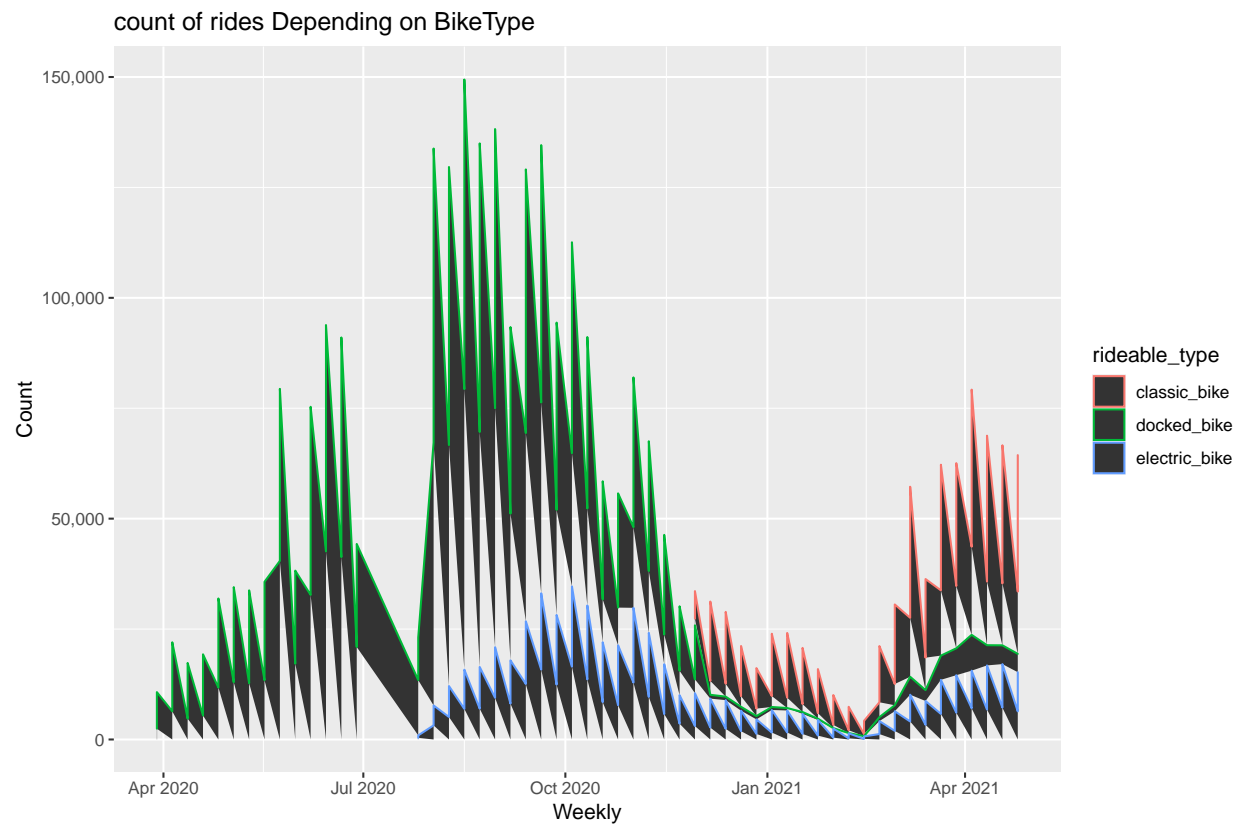


Creating the another Dataframe With required filtered attriutes

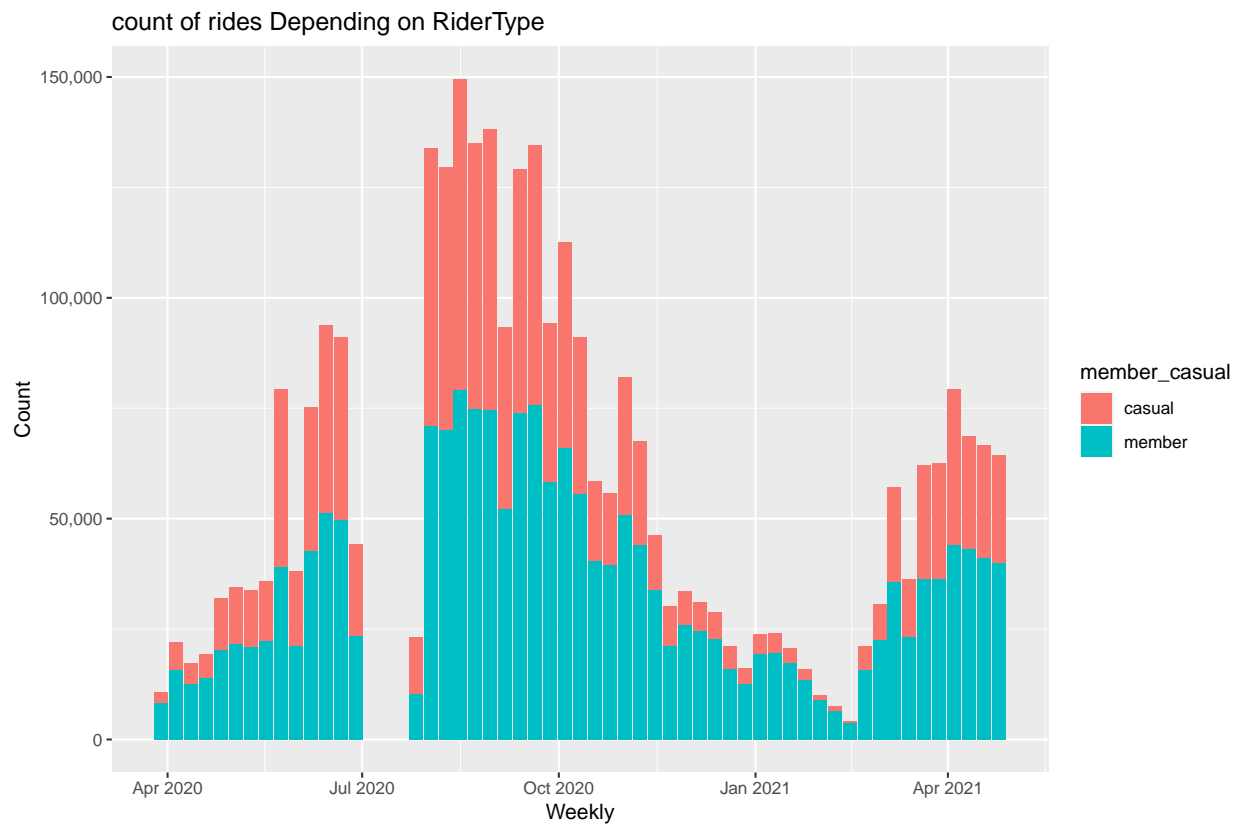
Plotting the Graph for Rides count depending on BikeType



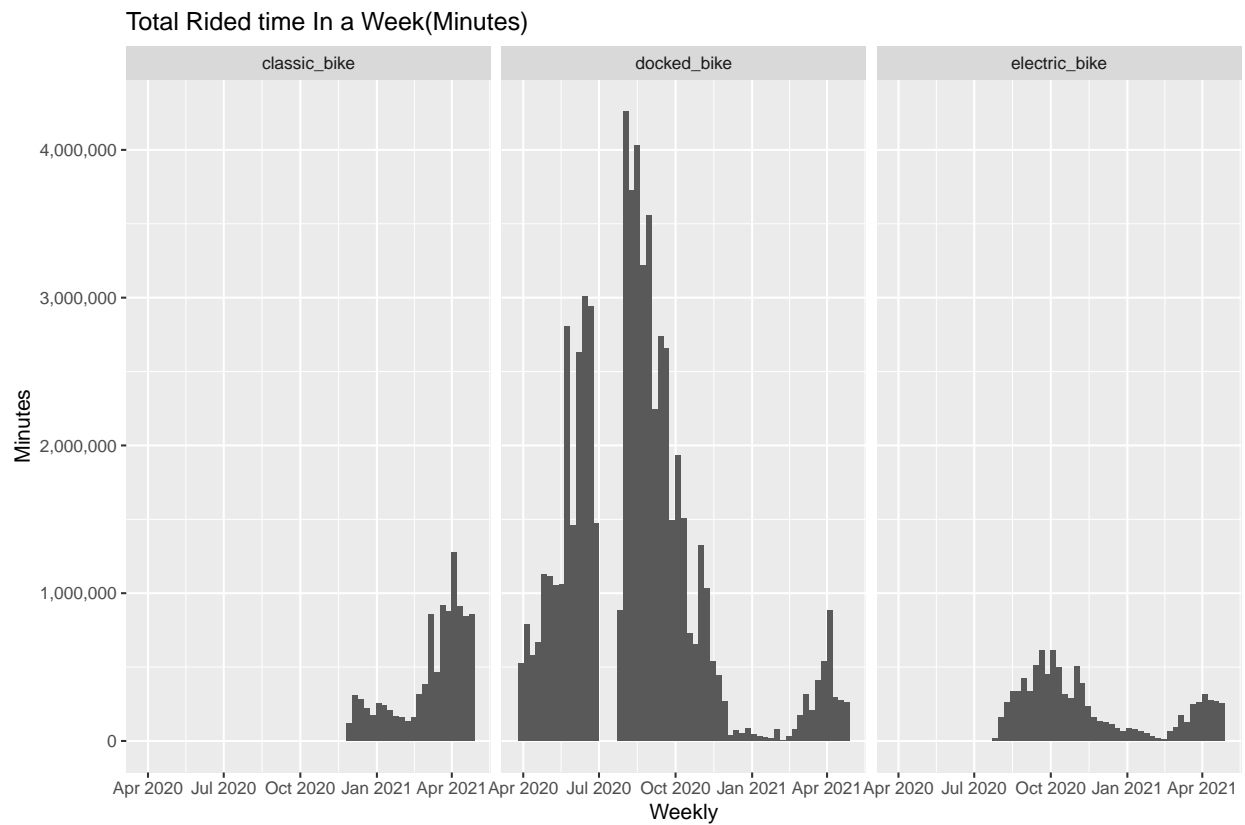
Plotting Rides count depending on BikeType



Plot to show the Rides Count depends on RiderType



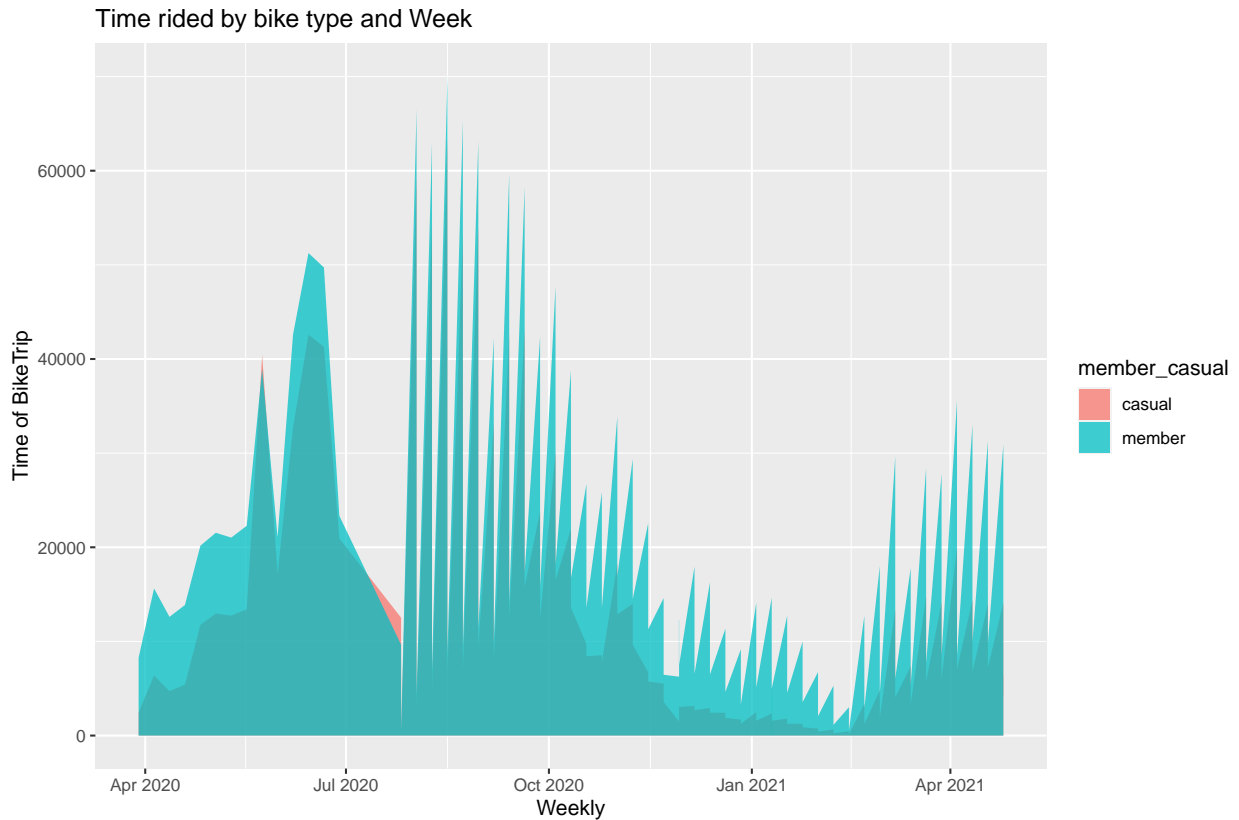
Visualizing Total Rides time in a Week



Plotting the time and biketype of a Ride for an One Week

```
## Warning: Ignoring unknown parameters: start
```

```
## Warning: Width not defined. Set with 'position_dodge(width = ?)'
```

Analysis and Observations :

- Casual users have a higher avg of ride length service members
- The density plot shows that most riders for both members and casual riders usually use the service for less than 10 mins on average
- The members seem to be utilizing the service consistently through out the week while the casual users utilize it more on the weekends

ACT and Recommendation Step

based on the observations from the data we can conclude the following:

- The marketing campaign should focus on the list of members with the 5 - 10 mins ride times and further deep dive into their habits
- perhaps some promotions for extending the their rides to have more time using the service
- Weekend centric promotions or service discounts for during week non members

further deeper analysis should be done to understand the location and route behavior the user take , that can lead to more thoughtful insights on the station placement.