GDA\_CaseStudy\_Analyze

Fred

4/22/2021

#### Case Study Roadmap - Analyze

Ask Three questions will guide the future marketing program:

1. How do annual members and casual riders use Cyclistic bikes dierently?  
2. Why would casual riders buy Cyclistic annual memberships?   
3. How can Cyclistic use digital media to influence casual riders to become members?

You will produce a report with the following deliverables:

1. A clear statement of the business task - Completed  
2. A description of all data sources used - Completed  
3. Documentation of any cleaning or manipulation of data - Completed  
4. A summary of your analysis - In progress  
5. Supporting visualizations and key findings  
6. Your top three recommendations based on your analysis

Guiding questions

- How should you organize your data to perform analysis on it?  
- Has your data been properly formatted?  
- What surprises did you discover in the data?  
- What trends or relationships did you find in the data?  
- How will these insights help answer your business questions?

Key tasks

1. Aggregate your data so it’s useful and accessible.  
2. Organize and format your data.  
3. Perform calculations  
4. Identify trends and relationships.

Deliverable

- Here is a summary of the analysis of the Bike Share data from Jan, 2021 to March, 2021  
- The data cleaning has been performed in RStudio.  
- This analysis is for case study 1 from the Google Data Analytics Certificate (Cyclistic).

Setup environment in R:

- Install required packages and library for data cleaning, transformation, and visualization.

Start of data cleaning and transformation:

- Fist we need to load the csv files to a dataframe  
- Check data for consistency in reparation for cleaning and transoformation.  
- Inspect the new dataframe that has been created  
- Transform Necessary Columns for calculation  
- Create a ride length column  
- Remove unnecessary data.  
- Perform initial analysis of data

Here is the statistical summary of the data:

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0 382.8 674.0 1265.9 1242.0 1897299.0

##### Compare members and casual users using mean

## all\_trips\_v2$member\_casual all\_trips\_v2$ride\_length  
## 1 casual 2227.4068  
## 2 member 854.8531

##### Compare members and casual users using median

## all\_trips\_v2$member\_casual all\_trips\_v2$ride\_length  
## 1 casual 1031  
## 2 member 576

##### Compare members and casual users using max

## all\_trips\_v2$member\_casual all\_trips\_v2$ride\_length  
## 1 casual 1897299  
## 2 member 89997

##### Compare members and casual users using min

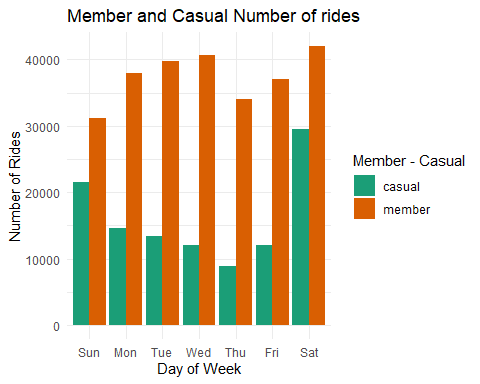
## all\_trips\_v2$member\_casual all\_trips\_v2$ride\_length  
## 1 casual 0  
## 2 member 0

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

## all\_trips\_v2$member\_casual all\_trips\_v2$day\_of\_week all\_trips\_v2$ride\_length  
## 1 casual Sunday 2390.3306  
## 2 member Sunday 959.8920  
## 3 casual Monday 2428.8481  
## 4 member Monday 867.0024  
## 5 casual Tuesday 2104.9577  
## 6 member Tuesday 819.6833  
## 7 casual Wednesday 1721.7870  
## 8 member Wednesday 828.2023  
## 9 casual Thursday 1616.5500  
## 10 member Thursday 765.1199  
## 11 casual Friday 1915.4691  
## 12 member Friday 803.3357  
## 13 casual Saturday 2581.7795  
## 14 member Saturday 943.1932

#### Here is the data analysis visualization using a bar chart.

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



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

