Google Data Analytics Capstone

Cyclistics Case Study

Ankush Koul

2022-03-09

#### Tools Used

* Excel
* MySQL
* Tableau
* RStudio

## Scenario

You are a junior data analyst working in the marketing analyst team at Cyclistic, a bike-share company in Chicago. The director of marketing believes the company’s future success depends on maximizing the number of annual memberships. Therefore, your team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, your team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.



I have used the six steps of the data analysis process learned through out the 8 courses in the Google Data Analytics program to answer key business questions. Ask, prepare, process, analyze, share, and act.

## Ask

In this step, I had to find out ‘How do annual members and casual riders use Cyclistic bikes differently?’ to guide my analysis. Lily Moreno, the director of marketing believes annual Members are much more profitable than Casual riders. As a junior data analyst working at Cyclistic, my task is to help the marketing team analyze the differences between Casual riders and Members with the goal of providing recommendations that will help convert Casual riders into annual Members to maximize profit.

## Prepare

I downloaded the dataset that is used for this case study and stored it. I was required to download data for the last 6 months. I downloaded trip data between January 2021-June 2021. You can access the datasets [here](https://divvy-tripdata.s3.amazonaws.com/index.html). It is a public dataset provided by Motivate International Inc.

## Process

I combined the separate files into one excel workbook. I then cleaned the data by removing duplicates, deleting rows that had ride start time later than ride end time (ie started\_at > ended\_at). I imported the data into Microsoft SQL server to clean the data faster. I deleted rows with NULLS in columns like station\_id, station\_name. I exported the data back into Excel using Power Query to merge all tables into a single table.

## Analyze

I created the ride\_length column to calculate the trip duration, Day\_of\_week to populate what day of the week the ride was taken in Excel. Pivot tables were used to help me summarize data. I then moved to RStudio to create data frames and visualize data with ggplot2. I calculated and summarized the data using functions such as mean, median, max, and min in R

## Share

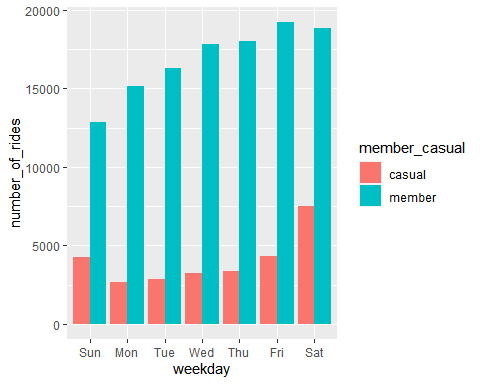
I created more data visualizations in Tableau. I then combined the visualizations on a dashboard that can be viewed [here](https://public.tableau.com/app/profile/ankush.koul/viz/CyclistsCasestudy1/Cyclisticbike_share). It was interesting to see similar patterns in all tools; the spreadsheet, Tableau, and in R. To view my R code in GitHub, click [here](https://github.com/Ankush231994/Case-study).

## Visulization

Here we will go through the series of visualizations

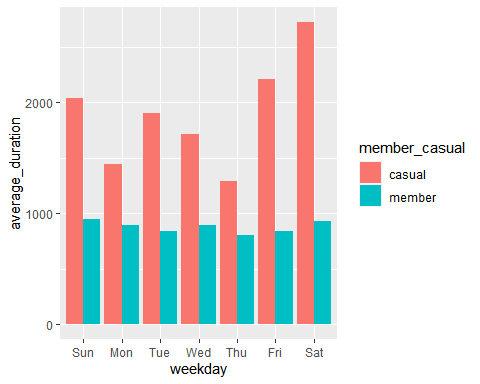
## Number of rides by rider type

Here is the plot for number of rides taken by causal and subscribed members



## Average duration of ride by the rider type

Here is the plot for average duration of ride by causal and subscribed members



## Act

In this final step, my task was to state and act on key findings by providing recommendations that will help the marketing director (my manager) maximize profits.

## Key Findings

* Casual riders were found to take longer trips or rides on average than annual Members. This could mean Casual riders used Cyclistic bikes for leisure.
* Casual riders often ride on weekends whereas annual Members use the program more over the week than on weekends. This could indicate that annual Members are using the bikes to commute to work.
* Casual riders preferred using electric bikes to other ride types like classic and docked bikes. Annual Members used classic bikes more than other types. The reason Casual riders mostly rode electric bikes could be because they rode longer so, electric bikes eased propulsion.

## Deliverables (Recommendations)

* Users that take long rides could be offered a discount on renting bikes or better yet annual Members that travel more than a certain distance say for instance 10 miles should be offered a discount.
* Cyclistic bike-share should introduce a weekend only membership that costs less than the current 7-day membership.
* A rebate program in form of refunds to annual Members that use a certain ride type more others should be introduced. It could be a small percentage of the cost.