**INTRODUCTION**

This case study is my capstone project for the Google Data Analytics Certificate. It involves analysis of historical data of a fictional company called Cyclistic, a bike sharing company located at Chicago, to make recommendation for an upcoming marketing campaign. Although, the company and scenario are fictitious the data used for the project are real data collected from a bike share program in Chicago. In this project I am assuming the role of the junior analyst.

**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 throughout the 8 courses in the Google Data Analytics program to answer key business questions. **Ask**, **prepare**, **process**, **analyze**, **share**, and **act**.

**ACT**

In this step, I asked myself questions such as ‘How do annual members and casual riders use Cyclistic bikes differently and why would casual riders buy Cyclistic annual memberships?’ to guide my analysis. 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 12 months the data used for this analysis were obtained from the Cyclistic bike-share program, a company employed by the City of Chicago to collect data on bike share usage. The [data](https://divvy-tripdata.s3.amazonaws.com/index.html) is collected directly by Motivate Inc., the company that runs the Cyclistic Bike Share program for the City of Chicago. The data has been made available by Motivate International Inc. under this [license.](https://ride.divvybikes.com/data-license-agreement) The data is comprehensive and consistent. The data is current as it had been released monthly. The City of Chicago makes the data available to the public. The data is organized in monthly csv files. The files from June 2020 – May 2021 were used for this project. The files consist of 13 columns containing information related to ride id, rideable type, ride time, start station name etc.

**PROCESS**

For this project I choose to use **MICROSOFT EXCEL, MySQL, POWER BI**.

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) and also added two additional columns. I imported the data into MySQL to clean the data faster. I then proceeded to combine all 12 files into one table where I deleted rows with NULLS in columns like station\_id, station\_name

**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 Microsoft Excel. MySql was used to perform descriptive analysis where I calculated and summarized the data using functions such as mean, max, min, and sum.

**SHARE**

I created data visualizations in Power BI. I then combined the visualizations on a dashboard that can be viewed [here](https://app.powerbi.com/view?r=eyJrIjoiN2M1YjE3YjUtM2RkNy00Zjk2LWJhNjYtMDE1MTA1N2NkOTI5IiwidCI6IjgyM2Q2NDg0LWJkOGMtNDZhNC1hMzQzLTgxZWNjNjZkNzg3YiJ9)

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**

1. From my analysis I found out that casual riders mostly take rides during weekends while members mostly take rides during weekdays this could be because the mostly use the bikes to commute to work as against casual riders who use the bikes for leisure.
2. During the winter season (December, January, February) there was a dip in ride rate in both casual riders and members because its snowy and cold during these period
3. During the summer season (July, August September) there was a spike in ride rate for all riders because of the weather and also due to the fact its holiday season
4. Noticed that both casual riders and members prefer using the docked bikes than other bike type.

**RECOMMENDATION**

1. There should be provision for part time membership for casual riders that may not want the yearly membership, I believe this will entice them
2. Adverts should be targeted at the top 5 start and end stations that casual riders frequent.
3. There should be some form of incentive like giving discount on the membership for a period of time to encourage casual riders.
4. Users that take long rides could be offered a discount on renting bikes

I’m confident these recommendations will encourage Casual riders to upgrade to annual membership thus increasing revenue and profits at Cyclistic bike-share