Case Study 1

# Introduction

This is my version of Google Data Analytics Capstone project – Case study 1. The full document to the case study can be founded in the [Google Data Analytics Capstone Project](https://in.coursera.org/learn/google-data-analytics-capstone) course.

It will follow the steps of data analysis process:

Ask, Prepare, Process, Analyze, Share and Act

Each step will proceed with its own set of tasks -

* Code required for the step
* Guiding questions with answers

**Ask**

For this step let’s get some context from the cyclist document:

# 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.

# Characters and teams

* **Cyclistic:** A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also oﬀering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can’t use a standard two-wheeled bike. The majority of riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.
* **Lily Moreno:** The director of marketing and your manager. Moreno is responsible for the development of campaigns

And initiatives to promote the bike-share program. These may include email, social media, and other channels.

* **Cyclistic marketing analytics team:** A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps guide Cyclistic marketing strategy. You joined this team six months ago and have been busy learning about Cyclistic’s mission and business goals — as well as how you, as a junior data analyst, can help Cyclistic achieve them.
* **Cyclistic executive team:** The notoriously detail-oriented executive team will decide whether to approve the recommended marketing program

|  |
| --- |
| **Guiding questions**   * What is the problem you are trying to solve?   The main objective is to build the best marketing strategy to turn casual bike riders into members and also record the patterns of the member bike riders in their business.   * How can your insights drive business decisions?   The insights that will be generated will assist in building marketing strategies. |
| **Key tasks**   1. Identify the business task 2. Consider key stakeholders |
| **Deliverable**  A clear statement of the business task  Generating the patterns of bike ride behavior of casual bike riders and member bike riders. |

# Prepare

This project will use the dataset provided by Google. [Download the previous 12 months of Cyclistic trip data](https://divvy-tripdata.s3.amazonaws.com/index.html) [here.](https://divvy-tripdata.s3.amazonaws.com/index.html)

|  |
| --- |
| **Guiding questions**   * Where is your data located?   The dataset is located in the Amazon web server and the link is provided by Google.   * How is the data organized?   The data is organized in csv format which stands for comma separated value.   * Are there issues with bias or credibility in this data? [Does your data ROCCC?](https://www.coursera.org/learn/data-preparation/lecture/lHirM/what-is-bad-data)   The data is reliable, original, comprehensive, current and cited.   * How are you addressing licensing, privacy, security, and accessibility?   The dataset is published by Google and has its own public license.   * How did you verify the data’s integrity?   Each column has correct data in it.   * How does it help you answer your question?   The data may assist in providing patterns of rider’s behavior.   * Are there any problems with the data?   There are no problems with the data as we are using this year’s data. |
| **Key tasks**   1. Download data and store it appropriately. 2. Identify how it’s organized. 3. Sort and ﬁlter the data. |

|  |
| --- |
| 4. Determine the credibility of the data. |
| **Deliverable**  A description of all data sources used |

The main data source has various formats of previous year’s data based on monthly data and also quarterly data.

Our focus is on current year data up-to month of September.

# Process

This step will prepare data for analysis.

|  |
| --- |
| **Guiding questions**   * What tools are you choosing and why?   Python is chosen for cleaning since it has pandas library which can help with cleaning and sorting.   * Have you ensured your data’s integrity? * What steps have you taken to ensure that your data is clean?   Dropping null values, sorting the data, changing the data type of certain columns.   * How can you verify that your data is clean and ready to analyze?   The data must not have any null values and columns must have their correct data type.   * Have you documented your cleaning process so you can review and share those results?   The code has been well documented but result screenshots are not attached due to some hardware problem. |
| **Key tasks**   1. Check the data for errors. 2. Choose your tools. 3. Transform the data so you can work with it electively. 4. Document the cleaning process. |
| **Deliverable**  Documentation of any cleaning or manipulation of data |

The python code is attached in the repository.

# Analyze

The analysis will be done by processing certain calculations to obtain numbers showcasing riders riding patterns.

|  |
| --- |
| **Guiding questions**   * How should you organize your data to perform analysis on it?   The organization should be done in ascending order.   * Has your data been properly formatted?   Yes it has been properly formatted.   * What surprises did you discover in the data?   Negative values were obtained when calculating distance and time taken for travelling which is practically impossible.   * What trends or relationships did you ﬁnd in the data?   Member count and casual rider count and their riding patterns such as time taken for travel and distance covered.   * How will these insights help answer your business questions?   This can help us to find mean time for riding bikes total days for when bikes are taken away and they don’t return. |
| **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**  A summary of your analysis |

# Share

Tableau is used to visualize data.

|  |
| --- |
| **Guiding questions**   * Were you able to answer the question of how annual members and casual riders use Cyclistic bikes diﬀerently?   Yes. The viz supports the answer to this question.   * What story does your data tell?   The riding patterns with respect to members and casual riders in space.   * How do your ﬁndings relate to your original question?   It shows the different trends based on hour of the day and also weekday in the usage of bikes by both members.   * Who is your audience? What is the best way to communicate with them?   The marketing analytics team and executive team is the audience here and the best way to communicate with them is by holding a meeting to share the visualization that is generated.   * Can data visualization help you share your ﬁndings?   Yes. It shows clear patterns between both the riders.   * Is your presentation accessible to your audience? |
| **Key tasks**   1. Determine the best way to share your ﬁndings. 2. Create elective data visualizations. 3. Present your ﬁndings. 4. Ensure your work is accessible. |
| **Deliverable**  Supporting visualizations and key ﬁndings |

[Bike-rider-summary](https://public.tableau.com/views/cyclist-bikeride/Story1?:language=en-US&:display_count=n&:origin=viz_share_link) is the link to the visualization on tableau.

# Act

|  |
| --- |
| **Guiding questions**   * What is your ﬁnal conclusion based on your analysis?  1. The casual riders have used/ taken away bikes for more no. of days than members. 2. Casual riders tend to ride more during the weekends whereas members have average bike riding. 3. Members cover more distance during mid-week and also more distance is covered per hour during afternoon time by them. 4. Casual rider cover more distance per hour during evening time. 5. Only casual riders prefer docked bike and member have greater preference for classic bike. There is an equal and rising share of both members preferring electric bike. 6. Average usage time of casual bike rider is more when compared to members.  * How could your team and business apply your insights?   Marketing strategy can be generated focusing on some of these insights.   * What next steps would you or your stakeholders take based on your ﬁndings?   Brainstorming the marketing strategy which should focus on both casuals as well as members. This can increase our customer base.   * Is there additional data you could use to expand on your ﬁndings?   The weather data can be used to see the affect of weather patterns to the bike riding patterns and their route patterns can be analyzed. |
| **Key tasks**   1. Create your portfolio. 2. Add your case study. 3. Practice presenting your case study to a friend or family member. |
| **Deliverable**  Your top three recommendations based on your analysis |

1. The marketing team should enable weekend passes to the customers who take up their membership. There can be additional food court coupons given to members which can be availed on weekends on their route.
2. Company can tie-up with premium music servicing company and its services can be given to members with focusing on some weekend playlist. Also more offers can be availed on docked bikes to turn casual riders into members.
3. Members only service can be provided for the members covering larger distances over the mid-week days. This can include free beverages to their offices.
4. With the growing interest of both members and riders in electric bikes offers can also be prevailed on electric bikes and more focus can be given on marketing strategy focusing on E-future.