ASK

Guiding questions

**Q.What is the problem you are trying to solve?**

Answer the question How do annual members and casual riders use Cyclistic bikes differently?

And thus help create an effective marketing program to

**Q.How can your insights drive business decisions?**

Digging deep and finding the difference between the two types of riders will help increase the conversion rate from casual riders to annual members and thus help increase the profit margins

**Key tasks**

1. Identify the business task

2. Consider key stakeholders

**Deliverable**

A clear statement of the business task

Find the key differences between casual and member drivers to deliver an appropriate digital marketing scheme

PREPARE

Guiding Questions

**Q.Where is your Data located?**

The data is located in <https://divvy-tripdata.s3.amazonaws.com/index.html>

**Q.How is the Data organised?**

The data is organised in the form of a csv file

**Q Are there issues with bias or credibility in this data? Does your data ROCCC?**

The dataset is credible considering the fact that the company has collected the data from its own geolocation techniques and from its own users. And finally, it's ROCCC because it's reliable, original, comprehensive, current and cited.

**Q How are you addressing licensing, privacy, security, and accessibility?**

The company has their own licence over the dataset. Besides that, the dataset doesn't have any personal information about the riders.

**Q How did you verify the data’s integrity?**

All the files have consistent columns and each column has the correct type of data.A preliminary check was done using filters in spreadsheet tools

**Q How does it help you answer your question?**

The data may hold some valuable insights on rider information which can be found out in our analysis

**Q Are there any problems with the data?**

The data could have been updated to the current time scenario and more information regarding the riders will be useful to gain a deeper insight into the type of cyclists and member riders and thus give a better digital media marketing approach

## **Key tasks**

* [x] Download data and store it appropriately.
* [x] Identify how it’s organized.
* [x] Sort and filter the data.
* [x] Determine the credibility of the data.

## **Deliverable**[**¶**](https://www.kaggle.com/code/jhelisonuchoa/google-data-analytics-capstone-case-study-1/notebook#Deliverable)

* [x] A description of all data sources used

The main data source is 12 months (Between april 2020 and march 2021) of riding data provided by the Cyclistic company.

PROCESS

## **Guiding questions**

* **What tools are you choosing and why?**

I'm using python for the data cleaning process as the pandas library is my stronghold and R loads the dataset very slowly and changes are very slow . On the contrary I used R for data analysis as R is very versatile to plot charts and draw inferences

**Q Have you ensured your data’s integrity?**

Yes, the data is consistent throughout the columns.

**Q What steps have you taken to ensure that your data is clean?**

First the duplicated values where removed, then the columns where formatted to their correct format.

**Q How can you verify that your data is clean and ready to analyze?**

It can be verified by the python notebook I have uploaded

**Q Have you documented your cleaning process so you can review and share those results?**

Yes, it's all documented in a jupyter notebook

## **Key tasks**

* [x] Check the data for errors.
* [x] Choose your tools.
* [x] Transform the data so you can work with it eectively
* [x] Document the cleaning process.

## **Deliverable**

* [x] Documentation of any cleaning or manipulation of data

ANALYZE

## **Guiding questions**

**Q How should you organize your data to perform analysis on it?**

The data has been organized into a single CSV concatenating all the files from the dataset and then cleaning it

**Q Has your data been properly formatted?**

Yes, all the columns have their correct data type and if not have been typecasted

**Q What surprises did you discover in the data?**

One of the main surprises is how members differ from casuals when analysed from weekdays. Also that members have less riding time than casual and also that docked bikes were of very high frequency

**Q What trends or relationships did you find in the data?**

* + There are more members than casuals in the dataset.
  + There are huge differences between mid year and end or start year statistics regarding the bike rent
  + Members have less riding time.
  + Members tend to prefer docked bikes.

**Q How will these insights help answer your business questions?**

This insights helps to build a profile for members.

## **Key tasks**

* [x] Aggregate your data so it’s useful and accessible.
* [x] Organize and format your data.
* [x] Perform calculations.
* [x] Identify trends and relationships.

linkcode

## **Deliverable**

* [x] A summary of your analysis

SHARE

## **Guiding questions**

**Q Were you able to answer the question of how annual members and casual riders use Cyclistic bikes differently?**

Yes. The data points to several differences between casuals and members.

**Q What story does your data tell?**

The main story the data tells is that members have set schedules, and the data is highly representative of members. The months average give us an idea about the fact that the summer months in Chicago have a higher amount of members using their bikes . We also infer from the chart that Streeter Dr is the most popular station and docked bikes are the most used rideable type

**Q How do your findings relate to your original question?**

The findings build a profile for members, relating to "Find the keys differences between casuals and annual riders", also knowing whey they use the bikes helps to find "How digital media could influence them".

**Q Who is your audience? What is the best way to communicate with them?**

The main target audience is my cyclistic marketing analytics team and Lily Moreno. The best way to communicate is through a slide presentation of the findings.

**Q Can data visualization help you share your findings?**

Yes, the main core of the finds is through data visualization.

**Q Is your presentation accessible to your audience?**

Yes, the plots were made using vibrant colors, and corresponding labels were added to give perspective

## **Key tasks**

* [x] Determine the best way to share your findings.
* [x] Create effective data visualizations.
* [x] Present your findings.
* [x] Ensure your work is accessible.

linkcode

## **Deliverable**

* [x] Supporting visualizations and key findings

ACT

## **Guiding questions**

**Q What is your final conclusion based on your analysis?**

Members and casual have different habits when using the bikes and thus the same marketing strategy cannot be used for both

**Q How could your team and business apply your insights?**

The insights could be implemented when preparing a marketing campaign for turning casual into members. The marketing can focus in a green initiative . they can work on promotional offers on the famous station names analysed so that members will be willing to continue their subscriptions

**Q What next steps would you or your stakeholders take based on your findings?**

Further analysis could be done to improve the findings, besides that, the marketing team can take the main information to build a marketing campaign. More data can be derived for increasing the depth of knowledge

**Q Is there additional data you could use to expand on your findings?**

* + Improved climate data.
  + More information on other places

## **Key tasks**

* [x] Create your portfolio.
* [x] Add your case study.
* [x] Practice presenting your case study to a friend or family member.

linkcode

## **Deliverable**

* Your top three recommendations based on your analysis

1. Build a marketing campaign focusing on show how bikes help people to get to work, while maintaining the planet green and avoid traffic. The ads could be show on professional social networks.
2. Increase benefits for riding during cold months. Coupons and discounts could be handed out.
3. As the bikes are also used for recreations on the weekends, ads campaigns could also be made showing people using the bikes for exercise during the weeks. The ads could focus on how practical and **consistent** the bikes can be.