Case Study: How Does a Bike-Share Navigate Speedy Success?

This case study will be about a fictitious bike-share company based in Chicago.

I will be going through each process based on the terms Ask, Prepare, Process, Analyze, Share, and Act.

I will be asking some questions at each of these terms to show how my thought process was guided throughout the case study, and a core deliverable at the end of each section.

Scenario for case study:

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. Most 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.

Ask

Questions:

• **What is the problem I am trying to solve?**

-How do annual members and casual riders use Cyclistic bikes differently, and how would we incorporate an effective marketing strategy?

• **How can your insights drive business decisions?**

-My insights could help by showing the company, how annual members and casual riders use Cyclists bikes differently and therefor what could be done to be more appealing to a casual rider and make them into an annual member and thus increase revenue. Also identifying what features annual members like could also improve retention rate.

Deliverable

• **A clear statement of the business task**

-To find out what the core differences between the different member types, and what methods we could pursue to increase annual membership.

Prepare

The dataset here; https://divvy-tripdata.s3.amazonaws.com/index.html. The license can be viewed here; https://ride.divvybikes.com/data-license-agreement.

Questions

• **Where is your data located?**

The data is stored on Kaggle as a dataset named "divvy-tripdata".

• How is the data organized?

The data consists of 4 csv files. The file for each yearly quarter is compiled into its own separate CSV file, e.g., Quarter 1 consists of the months Jan-Apr.

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

The company Cyclist collected all the information firsthand; the credibility of the data is also high as it is the companies' own data. The integrity of the data is reliable, original, comprehensive, current, and cited, and can help me answer my previous question by giving the relevant information to see a trend and draw an informed conclusion and decision.

• How are you addressing licensing, privacy, security, and accessibility?

The data is free to the public as provided by Cyclistic and is covered by their license. Each of the rider's privacy is kept protected by having their personal information kept hidden.

• How did you verify the data's integrity?

The data was searched through and any changes in format such as the names of rows and columns that were updated later was changed so that the data remained consistent throughout the case study.

• How does it help you answer your question?

The data shows us the differences between how the various members of rides and gives us ideas on how we could better tailor the annual membership to their experience.

• Are there any problems with the data?

Having more information on the riders daily riding schedule could potentially increase the data's overall value.

Deliverable

• A description of all data sources used

The data source consists of 4 CSV files. Each quarter of the month is in an appropriately labeled CSV file. The four files together run through May 2019- Apr 2020.

Process

The individual CSV files will be combined into 1 to make it easier to manipulate and analyze. The combined file will be cleaned, and additional columns will be added.

Guiding questions

- What tools are you choosing and why?

I used R and Microsoft excel. I started working on the dataset in excel in order to do any preliminary data cleaning and sorting, I then proceeded to complete the project in R using R studio since the dataset was large and I may have missed something in excel I decided to double check everything in R studio.

- Have you ensured your data's integrity?

After double checking the dataset, I have verified the data's integrity and found that the data was consistent.

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

The NA's and duplicates were removed, the time and dates were formatted, and previous column titles were formatted to match 2020's column formats.

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

In the code I have added comments detailing the cleaning and formatting process throughout the code, to double check any data inconsistencies I may have overlooked when working in Microsoft excel.

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

The cleaning process has been documented throughout the code.

Deliverable

Documentation of any cleaning or manipulation of data which I put in the R code as comments.

Code

Available on my GitHub @
https://github.com/keeko411/PortfolioProjects/blob/main/Capstone%20Project%20code

Share

Guiding questions

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

I was able to answer the question of how annual and casual riders use Cyclistic bikes differently by being able to identify what the differences between the two types of members.

● What story does your data tell?

The data tells me that the casual riders end up having a much longer average ride time than an annual member. By adding ride time incentives for annual members, I believe this could boost the number of casual riders who are willing to get an annual membership.

● How do your findings relate to your original question?

The data answered all the original questions.

1. How do annual members and casual riders use Cyclistic bikes differently?
Casual riders use the rental bikes for longer periods of time.

2. Why would casual riders buy Cyclistic annual memberships?
In theory if they use the bike rental often and ride for longer duration, then they would want to be an annual member, however the data provided does not show this to be the case.

3. How can Cyclistic use digital media to influence casual riders to become members?
A digital media campaign can influence casual riders to become members by displaying what benefits an annual member has over a casual rider.

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

My audience is the key stakeholders identified earlier. They are Lily Moreno and the Cyclistic marketing analytics team.

● Can data visualization help you share your findings?

Data visualization provided a visual comparison of the differences between the two groups. This allows the stakeholders to quickly and easily understand the data provided.
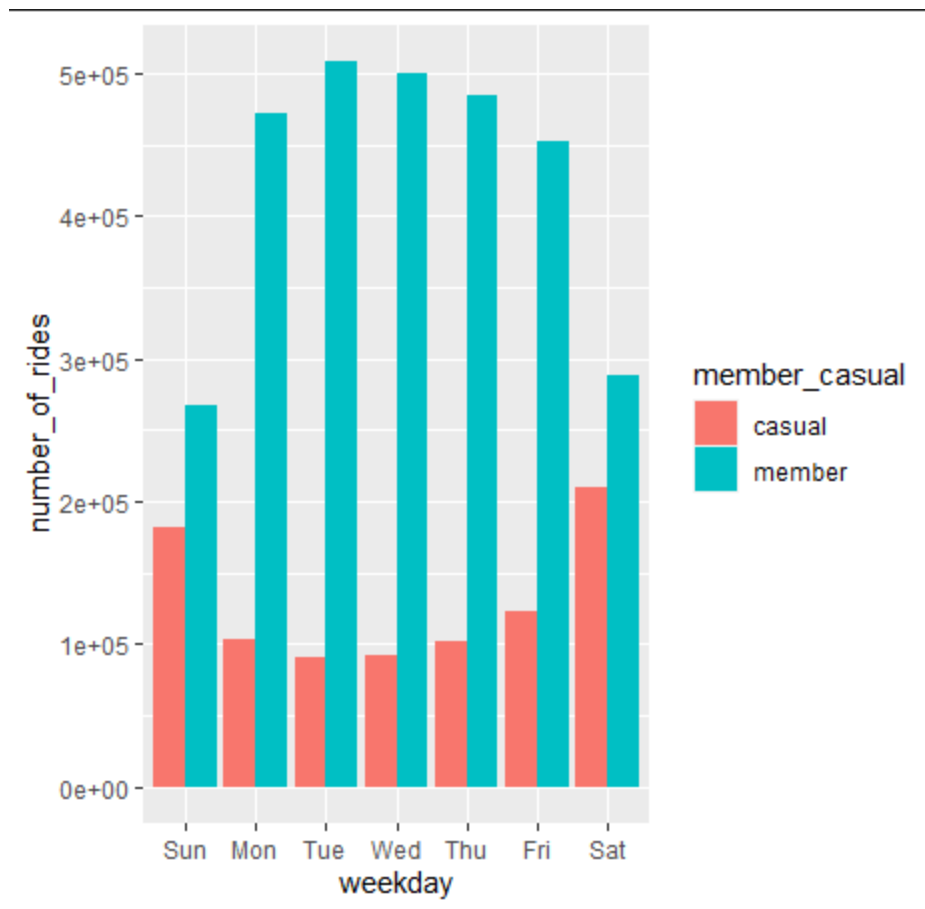
● Is your presentation accessible to your audience?

The data visualization is available to my audience (the stakeholders), in an easy-to-read manner so they can see the data for themselves.

Deliverable

Supporting visualizations and key findings
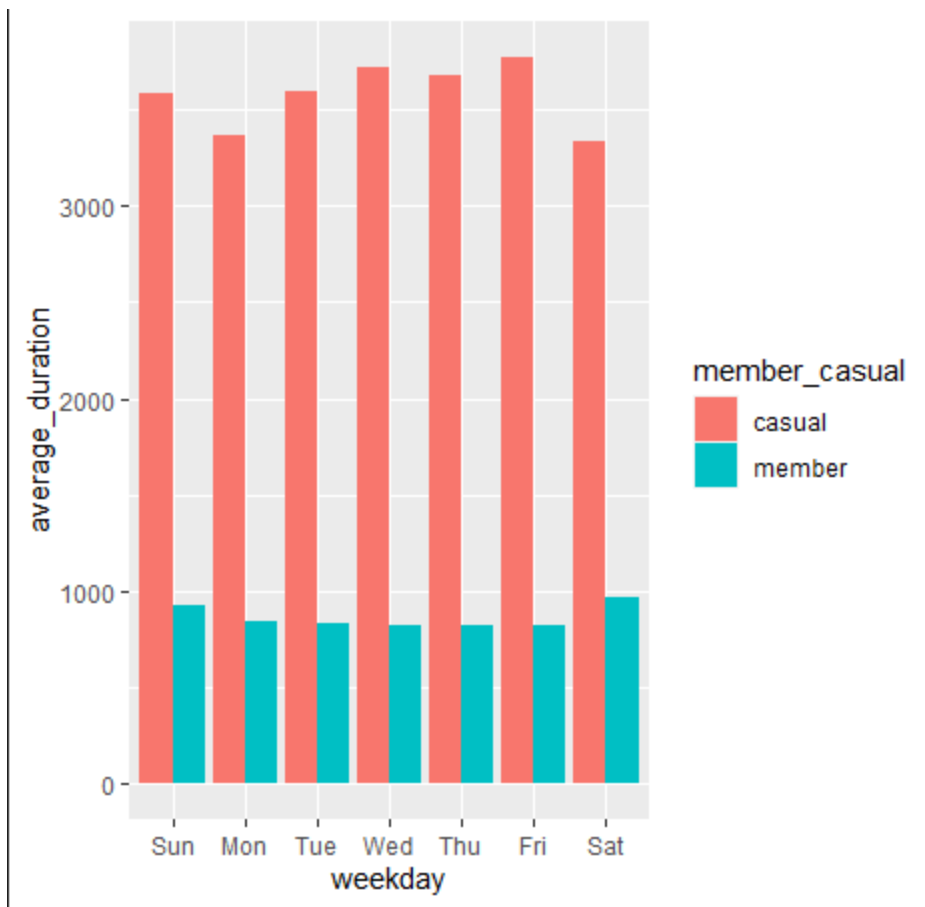
For the benefit of my audience, I have also added more interactive visualizations on tableau that go more in depth which can be located here.
https://public.tableau.com/app/profile/brandon.k.hunte/viz/Capstoneproject1_167218297711 40/Dashboard

Act

Questions

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

By adding incentives to annual members, I believe that we can convert casual members into annual members, as the data shows that many casual members use the rental bikes more than the annual members.

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

By adding additional incentives to annual membership based on how much they ride, I believe we could gain more annual members.

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

I would recommend launching a poll to see what possible incentives we could incorporate to catch potential annual membership applicants, and what could possibly retain interest with our current annual members.

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

Some more additional data we could use to expand upon our findings could be the climate as that could potentially influence how many people use the ride share, and how long they ride for.

Deliverable

My top three recommendations based on my analysis.

1. Setting up advertisements around the area to increase awareness of Cyclistic ride share.

2. Having a potential discount on less popular months to incentivize retaining a membership or getting a membership even on an off season.

3. Adding rewards and other incentives based on how long they ride for, and the distance ridden for annual members.