Case study: How does a bike-share navigate speedy success?

—----------

- Problem statement: To design a new marketing strategy to convert Casual rides into annual members.
- Problem solution: To generate recommendations from the problem statement, the generated recommendations must be backed up with compelling data insights and professional data visualizations in order to be approved by the cyclist executives team.
- Stakeholder: Lily Moreno, Manager of analytics team and the director of marketing in developing campaigns and initiatives to promote the bike-share program.

Data Teams involved:

- Cyclistic marketing analytics team: team of data analysts responsible for collecting, analyzing, and reporting data that guide cyclists marketing strategy.
- Cyclistic executive team: Detail-oriented executive team to decide whether to approve the recommended marketing program.

—----------

Implemented by following Six step data analysis phases

—----------

1. ASK: Problem definition, Smart questions, structural thinking, Stakeholder Moreno's considerations from the provided document.

ASK PHASE
Questions: ● What problem does this analysis solve? ● How can insights for the following problem drive business decisions? ● How do annual members and casual riders use Cyclistic bikes differently? ● Why would casual riders buy Cyclistic annual memberships? ● How can Cyclists use digital media to influence casual riders to become members?
Key task: ● Connecting with stakeholder, manager, subject matter experts, fellow team data analysts ● Checking data fairness.. ● To design a new marketing strategy to convert Casual rides into annual members.
Out comes: ● Found understanding for problem questions to provide recommendations from cyclists_trip

data through easy understandable visualizations by which solutions from insights must be effective for the marketing team to make business decisions in order to convert casual riders into annual members.

2. PREPARE: Data gathering, Data governance.

PREPARE PHASE

Questions:

- What type of data?
- What is the size of the data?
- Is the data consistent and accurate for analysis?
- What is the license of the dataset?
- Does data cleaning have to be performed or not?
- What type of analysis wants to be performed?
- What are the data types of columns in the dataset?
- What relation is observed between columns?
- What are necessary columns to perform analysis?
- Does any data have to be entered manually or not?
- Is data sufficient for analysis?

Key tasks:

- Interacting with stakeholders, fellow data analysts, members, conducting meetings, sending emails and discussing business problems.
- Data Collection, Extraction.
- How data is organized.
- Checking bias, credibility and ethics of data.

Out comes:
Downloaded dataset from the provided link and loaded into spreadsheets, bigquery and colab by checking bias, credibility and ethics of data under the licence provided in the document.

3. PROCESS: Data integrity, Data Cleaning, Data modifications.

PROCESS PHASE

Questions:

- Which tool is used to perform analysis?
- Does the dataset contain any outliers?
- How to treat the outliers?
- How many numerical and categorical columns were present in the dataset?
- How to handle numerical columns?
- How to handle categorical columns?
- What calculations to be performed on the dataset?

Key Task: Transforming data to work effectively

Out comes:
Found inconsistent data with file size exceeds as an error, while loading into spreadsheet. So, loaded the dataset into the database and performed a few sql queries. Then loaded the dataset into a colab environment to perform python queries and found null values, unique values and outliers in numerical data, handheld them with methods like fill na, Interquartile range, boxplots and heatmaps. There by finding inconsistency in the few categorical columns, handled them by one hot encoding, label encoding, by functions like split, trim, cast, concat. Finally added columns into the dataset as trip_length by subtracting Ended_at and Started_at, Days_of_Week by extracting only the number of days from start_at and normalized data for further analysis.

4. ANALYSE: Data aggregation, sorting, filtering, Data organization, Data validation.

ANALYSE PHASE

Questions:

- How does the data have to be aggregated?
- How the data has to be organized?
- How does the data have to be analysed?
- Is the data properly formatted?
- What trends and relationships were found in the data?
- What insights are generated?
- Do insights drive business decisions?

Key tasks: To perform data aggregation, organization, calculations, trend and relationships.

Out comes:
Aggregated data into cleaned spreadsheets for performing calculations like mean ride length, mode of day of week, max ride length by functions such as SUM,AVG, MAX, MIN,COUNT. Filtered and Sorted data into ascending order. Created pivot tables for data calculations.

5. SHARE: Design thinking, Data visualizations, matrices, Dashboards, presentation.

SHARE - PHASE

Questions:

- What are key performance indicators?
- What is the total count of cyclists?
- How many types of cyclists?
- What is the total count of bikes?
- What are different types of bikes?
- What is the count of bikes for each member type?
- Were all the questions answered?
- What story does the data tell?

- How are insights related to problem questions?
- Who is your audience?
- What is the best way to communicate with them?
- Do data visualizations share findings?
- Is the presentation accessible to the audience?

Key tasks:
To Determine the best way to share findings and Create effective visualizations for presenting findings by ensuring work is accessible.

Out comes:
Created a data visualization dashboard with metrics, charts, graphs, plots from calculated pivot tables data to share insights of problem solution to audience and stakeholder.

6. ACT: Feedback and Actions on results

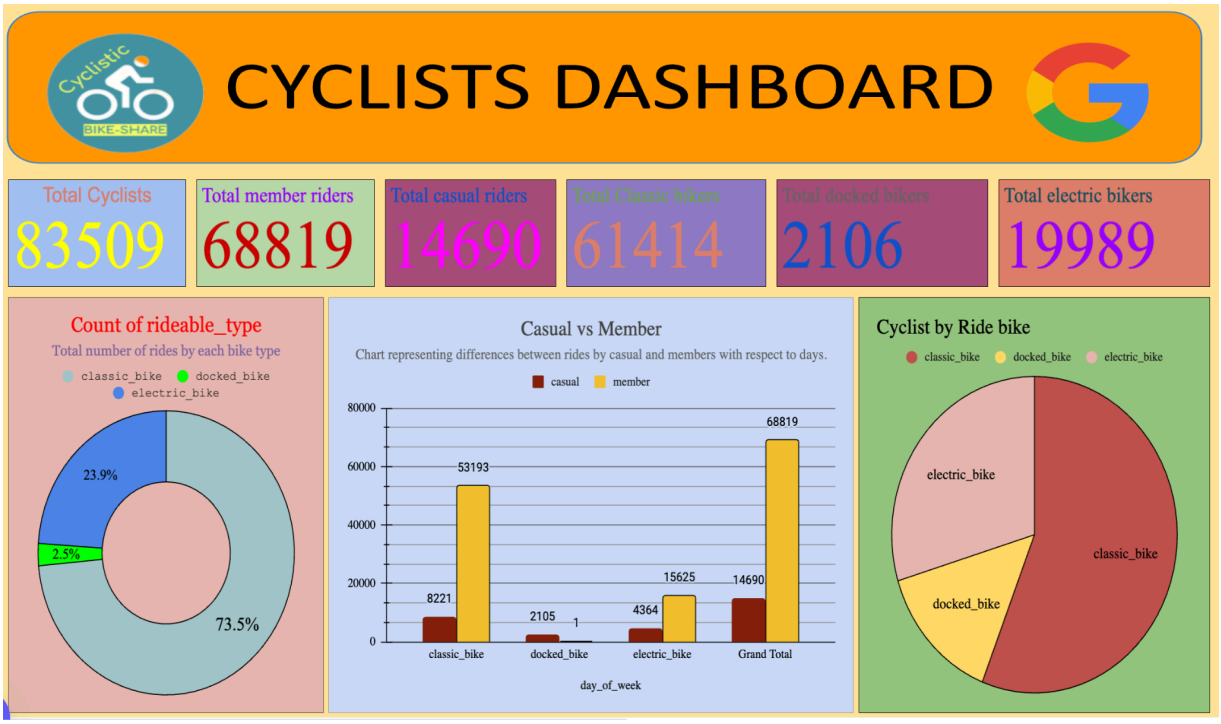
ACT - PHASE

Questions:
- What does the final conclusion represent?
- How can the insights be applied by team and business?

Key tasks: To create a portfolio to case study and present to a friend.

Out comes: Created Top three recommendations based on analysis and told a story from insights to a friend.

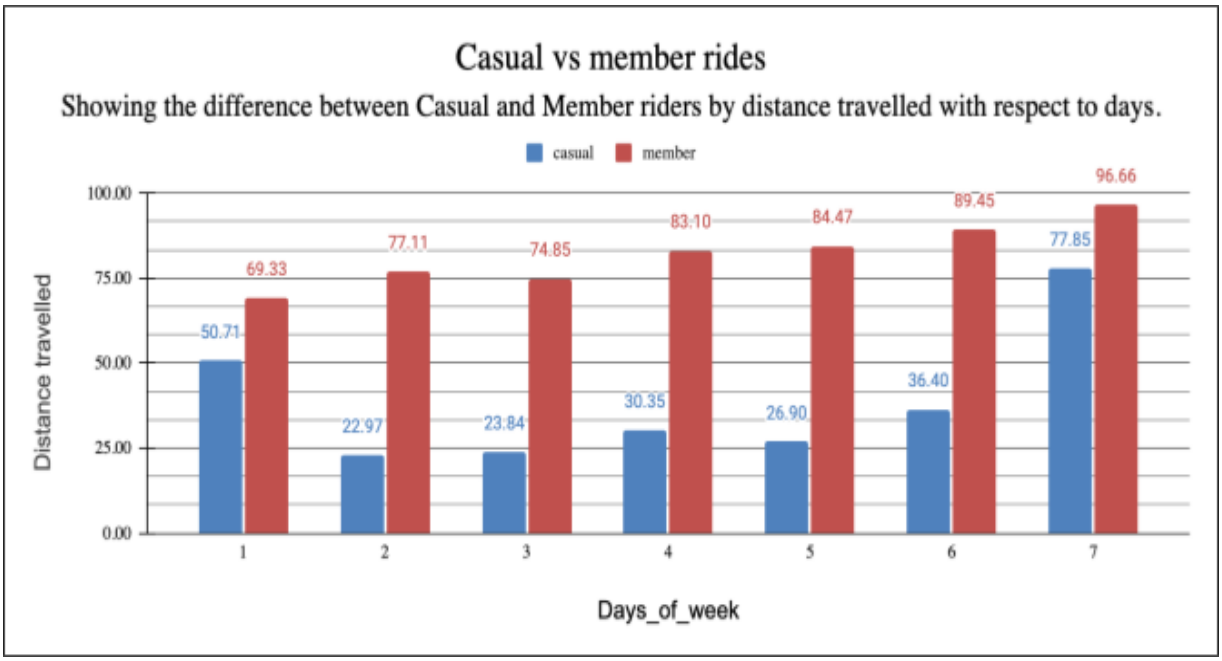
Dashboard:



CONCLUSIONS:

Answers to the questions concentrated to solve problem by stakeholder:

1. How do annual members and casual riders use Cyclistic bikes differently?

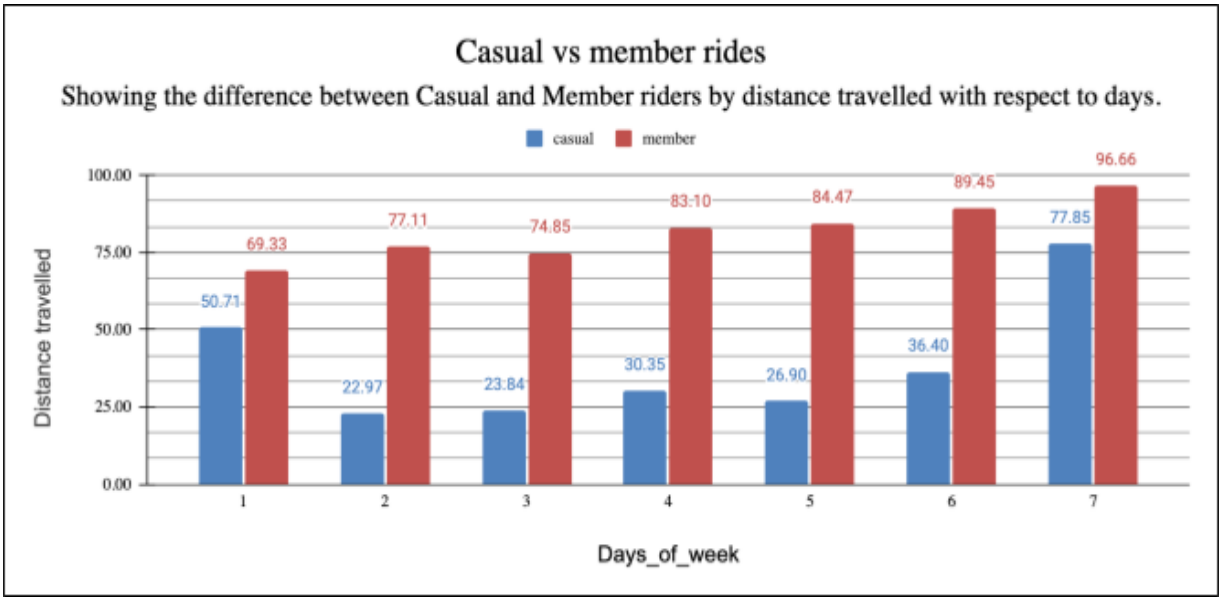


The above visualized column chart represents the comparison of Casual cyclist rides by Member Cyclist rides with respect to Days_of_week and ride length as distance travelled.

- Based on the insight generated on days_of_the week, the member cyclists travelled more distance at 69.33% through 96.66% whereas the Casual cyclists travelled less distance at 50.71% through 77.85%.
- By looking at all the days_of_week the Member Cyclists were at high with 96.66% as distance travelled but the Casual cyclists were at low with 50.71% as distance travelled.

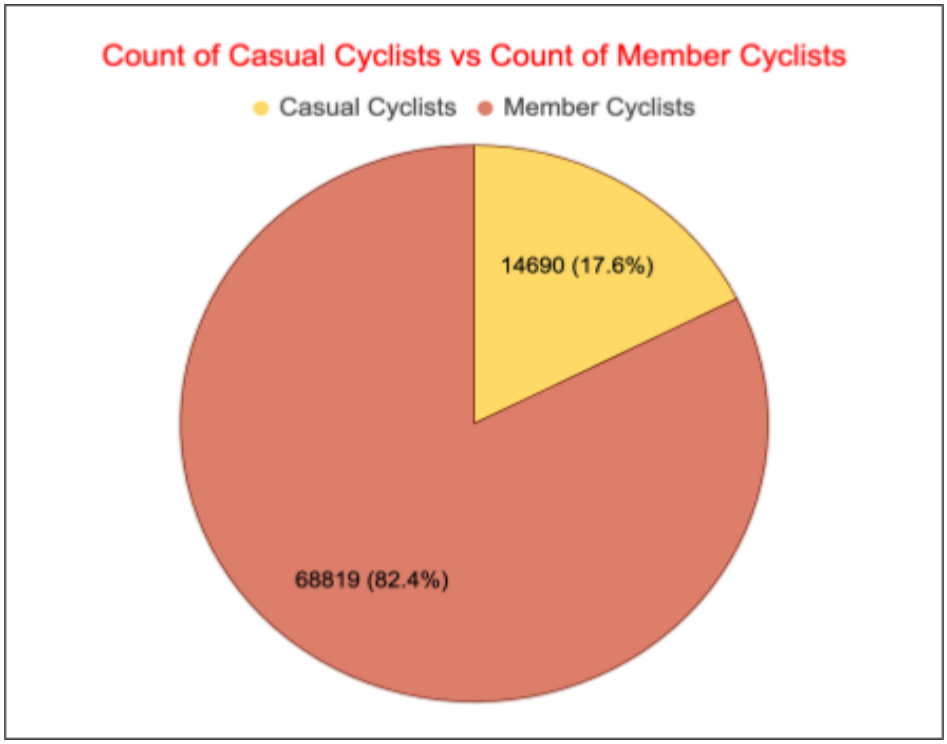
By looking at the factors from the chart with respect to the days of week it is concluded that Member cyclists on annual memberships use bikes more often than the casual riders on single and one day passes.

2. Why would casual riders buy Cyclistic annual memberships?



From the above chart it is clearly visualized as Member cyclists in red colored columns cover more travelling distance than casual members in blue colored columns. With the generated information marketing campaign can be performed by stating and representing visually as member cyclists were travelling more distances with annual memberships. So, if you are a casual cyclist who rides bikes on single or one day passes, you can benefit more from choosing an annual membership pass by which you can enjoy traveling more miles per day.

3. How can Cyclists use digital media to influence casual riders to become members?



- By updating and uploading an effective visualization to a digital media app by clearly representing benefits to an annual membership pass, benefits as referral bonus for both member cyclist and casual cyclist results in converting casual riders into member riders.

- From the above graph it is represented as at present there were 17.6% of the cyclists on single, one day passes with daily usage of cycle's at 30% and by performing marketing campaigns and promoting through digital with effective visualizations will lead to an increase of 50% in future. So, which results in the solution of converting casual cyclists to annual members.