**CASE STUDY DOCUMENTATION**

Moreno (Director of Marketing/Manager) has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. In order to do that, however, the team needs to understand better how annual members and casual riders differ, why casual riders would buy a membership, and how digital media could affect their marketing tactics. Moreno and her team are interested in analyzing the Cyclistic historical bike trip data to identify trends.

**Report Description for this Case Study**

1. A clear statement of the business task2. A description of all data sources used3. Documentation of any cleaning or manipulation of data4. A summary of your analysis5. Supporting visualizations and key findings6. Your top three recommendations based on your analysis

**ASK**

Three questions will guide the future marketing program:1. How do annual members and casual riders use Cyclistic bikes differently?2. Why would casual riders buy Cyclistic annual memberships?3. How can Cyclistic use digital media to influence casual riders to become members?

**Note:** Moreno asked firstly “What is the difference between the usage of Cyclistic bikes by annual members and causal riders”.

**1. ANALYTICAL REPORT FOR THE FIRST QUESTION**

Step1: Ask

**Answers to Guiding Questions** -The problem I am trying to solve is to find key insights or trends to what makes annual members different from casual riders. This insight yet to be gotten can help me determine how to improve the benefits of annual members which will invariably convert casual riders to full members.

**Business Task** – (i) Identify various sections that distinguish the usage of Cyclistic bikes of annual members and casual riders

**Key Stakeholders** – (i) Lily Moreno (Marketing Director)

(ii) Executive Team

**Objective Statement:**

(i) Thoroughly read through the brief description of the company given in this Case Study. It was deduced from historical trip data that annual members have proven to secure more revenue for the company. This brings us to the aim of the first business task to get a better overview of the operations of the two segments and make comparisons.

Step 2: Prepare

**Answers to Guiding Questions** – The data was obtained from the Index of bucket "divvy-trip data" an HTML file. The data is organized into datasets designated for each month in a zip file. There are no issues with biasing or credibility because this data source is reliable, original, comprehensive, current and cited. Also, the licensing, privacy, security and accessibility of the data source were checked and will be adhered to strictly. I verified the data’s integrity by ensuring the reliability of the source (Motivate International Inc.) This provides the datasets on where I will be able to run my analysis. After careful consideration, there are no problems with this data.

**Key Tasks to be Done** - Download the data and store it appropriately. Identify how it’s organized. Sort and filter the data. Determine the credibility of the data.

**Deliverable:** All data collected are presented in CSV format.

Step 3: Process

**Answers to Guiding Questions** –

* The analytical tool I will use in this stage is SQL (This is because of the large number of records in each dataset. There are thousands of records in each dataset which is too much to be done using a spreadsheet). However, it is much easier to use a spreadsheet to clean our data before the analysis.
* I have ensured data integrity earlier in step 2. It is from a reliable source.
* I have ensured all rows of empty cells found in each table are cleared by filtering out empty cells in Excel and deleting all. I also ensured there weren’t any form of duplicates in each table of data.
* To ensure my data is clean and ready for analysis I double-check all, my cleaning processes one step at a time by doing exact same cleaning steps.
* This is my documentation of my cleaning process done right here.
* I used SQL to find the length of rides by subtracting the started\_at column from the ended\_at column.
* Using Excel, I was able to get the exact day of the week using the (=WEEKDAY) function for each data in the started\_at column.

Step 4: Analyze

**Answers to Guiding Questions** –

-To organize my data before analysis, I ensured the values in each column were clean and consistent before importing the datasets in my BigQuery on my Google Workspace.

-I have ensured my data has been properly formatted using my Excel worksheet.

**Method of Analysis**

1. **Applying SQL and Excel for (Year 2024\_01)**

(i) Count all records of rides

SELECT COUNT(\*) AS *Total\_Records*FROM *the­\_table\_name*;.

(ii) No of member records

(iii) No of casual records

SELECT COUNT(\*) AS *Total­\_Records*FROM *the­\_table\_name*;.WHERE *member\_casual = member*;

SELECT COUNT(\*) AS *Total­\_Records*FROM *the­\_table\_name*;.WHERE *member\_casual = casual*;

(iv) No of Weekday rides

SELECT COUNT(\*) AS *Total\_Records*FROM *the\_table\_name*WHERE *day\_of\_week IN (2, 3, 4, 5, 6);*

(v) No of Weekends rides

SELECT COUNT(\*) AS *Total\_Records*FROM *the\_table\_name*WHERE *day\_of\_week IN (1,7);*

(vi) No of classic bike

SELECT COUNT(\*) AS *Total­\_Records*FROM *the­\_table\_name*;.WHERE *rideable\_type = classic*;

(vii) No of electric bike

SELECT COUNT(\*) AS *Total­\_Records*FROM *the­\_table\_name*;.WHERE *rideable\_type = electric*;

(viii) Average Journey of bikes, in general

SELECT AVG*(TIMESTAMPDIFF(MINUTE, started\_at, ended\_at))* AS *Average\_Journey*

FROM *the\_table\_name;*

(ix) Average Journey of bikes, for member riders

SELECT AVG*(TIMESTAMPDIFF(MINUTE, started\_at, ended\_at))* AS *Average\_Journey*

FROM *the\_table\_name;*

WHERE *member\_casual = member*;

(x) Average Journey of bikes, in causal

SELECT AVG*(TIMESTAMPDIFF(MINUTE, started\_at, ended\_at))* AS *Average\_Journey*

FROM *the\_table\_name;*

WHERE *member\_casual = casual*;

(xi) Longest time spent on the road (member)

SELECT MAX*(TIMESTAMPDIFF(MINUTE, started\_at, ended\_at))* AS *Average\_Journey*

FROM *the\_table\_name;*

WHERE *member\_casual = member*;

(xii) Longest time spent on the road (casual)

SELECT AVG*(TIMESTAMPDIFF(MINUTE, started\_at, ended\_at))* AS *Average\_Journey*

FROM *the\_table\_name;*

WHERE *member\_casual = casual*;

(xi) Latest time of arrival (Members)

SELECT MAX*(ended\_at) AS Maximum\_Arrival\_Date*

FROM *the\_table\_name*;

WHERE *member\_casual = member*;

(xii) Latest time of arrival (Casual)

SELECT MAX*(ended\_at) AS Maximum\_Arrival\_Date*

FROM *the\_table\_name*;

WHERE *member\_casual = casual*;

(xi) Start station ID for member

SELECT COUNT(\*) FROM *the­\_table\_name*;.WHERE start\_station\_id *= member*;

(xii) Start station ID for casual

SELECT COUNT(\*) FROM *the­\_table\_name*;.WHERE start\_station\_id *= casual*;

Step 5: Share

**Answers to guiding questions**

* I was able to answer the question postulated by the business task.
* My audience includes my Marketing Director and the board of Executive team.
* The best way to communicate with them is using two forms issuing out a report which is what I am compiling right now and creating a presentation slide to summarize and explain those key insights for better understanding.
* My data viz will help me share my findings most appropriately.

**Deductions**

1. A steady increase from February and reaches its peak at mid-year and then declines again at a steady rate.
2. The visual shows that member riders were higher each month than casual riders. Therefore, it indicates members had more rides than casuals in the calendar year.

**Note:** Generally, it is obvious there are more weekday rides (84%) than weekend rides (16%)

1. The average number of member rides on weekdays is much greater than that of casual rides. The visual shows there was a spike increment in the average rides in August for both ride types.
2. Both from the beginning and end of the year the average member rides were slightly greater than casual rides. However, there was a distortion at mid-year where casual rides lead member rides with a slight difference. This difference occurred between the May and September
3. Just like the number of rides taken there are more trips taken during the middle of the early than at the start or end of the year.
4. Considering the average journeys, there was a sudden change to the progressive order of trends shown in the previous slide. At this junction, the average trips taken by casual riders were significantly greater than member riders.
5. From the chat, it is shown that casual riders spend more time on the road than members. Also, from April to July, there was a drastic increase in the time spent on the road by casual riders. A slight drop was noticed for the next month before attaining a steady state for the rest of the year.

1. From the chart, casual riders arrive late at their destination compared to member riders. Perhaps the cost of membership is more than casual rides depending on how long you stay on the road. However, it was reversed for November and January. January was exceedingly high compared to November which may be due to probably a reduction in the cost of membership.

Step 6 Act

**Answers to guiding questions**

My conclusion and recommendation include:

* In this case study, I used cyclistic bike company data to evaluate the variation that exists between annual member riders and casual riders.
* There is a large margin between the average number of members and casuals during weekdays indicating how effective being a member will be for swift transit to work and good service rendered which is a priority over weekends.
* Average time spent on the trips is more common to casual riders because of the low charge to using casual rides than membership.
* However, for the start of 2024 there was a drastic increase in the time spent on trips for members than casuals which means a fair and affordable cost has been reached.

*Further analysis is required to verify the last conclusion which is a hypothesis yet to be confirmed.*

1. The Marketing Director can apply these insights by bringing a captivating narrative into the populace using media and other advertising agencies.

2. It shows that reducing the cost of membership to a reasonable form may eventually increase the time spent on trips to a significant period. It will be more effective to generate large revenue in the long run.

If true, January 2024 shows briefly what is like to happen in coming months.