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

1. A clear statement of the business task.

Cyclistic, a ride share company that starts operating in 2016. Since then it has grown to a fleet of 5824 bicycles with 692 Docker stations. Cyclistic have both the traditional bicycles, tricycles, and cargo bikes. People who can’t use regular bi-cycle use these tri-cycles. 8% of the total user use these tri-cycles. Normally people use these cycles for leisure rides but 30% of the riders use the cycles for regular commuting to work too.

The riders can use variety of packages like single-day passes, full-day passes, and annual membership. Riders who use single-day passes and full-day passes are referred to as casual riders.

Cyclistics finance analysts have concluded that annual members are more profitable than casual riders. That’s why marketing executives and managers wants to convert those casual riders to annual riders by marketing campaigns. But before doing that, the management and executives need to know how annual members and casual riders do use cyclistic bikes differently by using historical data.

1. A description of all the data source used

For answering the business task, historical data’s are needed. The dataset used for answering the questions are the last 12 months Cyclistics trips. All the datasets are organized as a csv file. The datasets are Reliable, Original, Comprehensive, Current and Cited. There is no bias or credibility issues with the dataset. The dataset has been made available by Motivate International Inc. under public license for using in any case study related to bike trips. The dataset do not contain any personal information of the users, in that case the data has no integrity or privacy issues. Also, for answering the business question the data are good enough. I need to do some sorting, filtering, cleaning and processing before the dataset can be used to answer the questions.

1. Documentation of any cleaning or manipulation of data

After downloading all the data and unzipping it I stored the data in a folder with relevant names. The data are in csv format. I then create another folder and convert those csv formatted data to Excel Workbook formatted and stored in that folder. Thus the data stored in both csv and Excel Workbook format. After doing so, I checked for data integrity and then find out that the data are accurate, complete, consistent, and trusted. I then checked for source of errors, misspelled words, extra spaces in cells, duplicates, mismatch of data types and found out everything is alright. I used Microsoft Excel for doing all these cleaning and cleaning verifications as it is convenient to use for data cleanliness verification. I then create two labels named ‘ride\_length’ and ‘day\_of\_week’. The ‘ride\_length’ column stored the total ride time which is end time – start time of any ride. The data formatted as HH:MM:SS format. The ‘day\_of\_week’ columns determines which day it is (ride start day), considering Sunday as the day 1 and so on. The data is clean and ready to analyze.

1. A summary of this analysis.

For performing analysis on the dataset I used bigquery sandobox for writing sql queries on the data. By writing sql queries on the dataset I discover various trends on the dataset which may help find the business questions that I am trying to find. The data was properly formatted to use as csv for creating bigquery table for performing queries. While writing queries on the data I found out minimum ride length is 0 and max ride length surpasses a whole day in some months. Also, found out casual members used the cycles most on Saturday and Sunday while annual members used it more during weekdays.

There are several key findings:

* Average ride length is always twice for the casual members than the annual members throughout the year.
* Annual member do not use docked bikes. Only casual member use them.
* Day by day casual members are converting into annual members.
* From November 21 to February 22 total rides for both the casual and annual member is significantly low.
* Casual members rides for on average 30 hours throughout the year while annual members rides for only 12 hours throughout the year.
* Casual members stopped using docked bike day by day.

1. Supporting visualizations and key findings

I used tableau Desktop for creating the visualization from the dataset. In Tableau I used bar chart, table, line graph and created a dashboard for every months to find out the relationship between the dataset. These visualization includes number of trips by member types in any month, average hour of ride length for both casual and annual members, start and end station for both members type and finally a dashboard which tells the story. The dashboard differentiate and answer the business question clearly for the audience how the annual and casual members are using the cyclistic bikes differently.

1. My top three recommendation based on my analysis:

There are several findings and I want to recommend several things to answer the business question about how casual and annual members are using cyclistic bikes differently.

1. There are several start and end station where riders are mostly casual. To convert those riders to annual member proper campaign should need at those stations.
2. Annual members use these bike during workdays, so, its better to target people who are doing jobs and their office is at cycling distance so they use the cycles regularly for commuting.
3. From November to February people use cycles less, because of heavy winter. So, proper measure should taken.