A
Case study

on

“HOW DOES A BIKE-SHARE NAVIGATE SPEEDY SUCCESS?”

for the company

“CYCLISTIC”

Submitted by

Mr. Viraj V Joshi
Junior Data Analyst
Marketing Team, Cyclistic

Under the guidance of

Mrs. Lily Moreno
Director
Marketing, Cyclistic

ACKNOWLEDGEMENT

I take the opportunity to show my immense gratitude to the "**Cyclistic**" pvt ltd Chicago, for granting me to complete my case study on the topic of **"HOW DOES A BIKE-SHARE NAVIGATE SPEEDY SUCCESS?"**

.

I extend my sincere thanks to respected madam **"Lily Moreno",** Director, Marketing, Cyclistic for encouraging, suggesting and guiding me in my complete case study process.

I extend my sincere thanks to all the respected members of the marketing analytics team for supporting and guiding me through the analysis process for the case study.

Mr. Viraj V Joshi
Junior Data Analyst
Marketing Team, Cyclistic

PREFACE

A broad description of “Cyclistic”, a bike share program, offering more than 5800 bicycles and 600 docking stations by the company “Cyclistic” has been made in this case study. What is the offering under this program, what are its future goals and how single-ride, monthly-ride passengers can convert themselves into annual membership is under study in this case. An effort has been made to exploit the merits and demerits of annual membership program. An attempt has been made to promote annual membership program as the best option for the customers of cyclistic, so as to satisfy all the requirements of customers. More emphasis has been created on annual membership so as to bring a win-win situation for the company and customers. A broad comparison of casual members and annual members has been carried out. It has been proposed that, casual members need to be converted to annual members so that Cyclistic earns profit and expand its business in other states. The three marketing strategy questions have been answered by the process of data analysis for the company so as to increase its profit and expand the business.

TABLE OF CONTENTS

CHAPTER 1

COMPANY PROFILE AND SHORT SURVEY

CYCLISTIC

In 2016, Cyclistic launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geotracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

Until now, Cyclistic's marketing strategy relied on building general awareness and appealing to broad consumer segments. One approach that helped make these things possible was the flexibility of its pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members.

SHORT SURVEY

Cyclistic's finance analysts have concluded that annual members are much more profitable than casual riders. Although the pricing flexibility helps Cyclistic attract more customers, Moreno believes that maximizing the number of annual members will be key to future growth. Rather than creating a marketing campaign that targets all-new customers, Moreno believes there is a very good chance to convert casual riders into members. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

Moreno has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. In order to do that, however, the marketing analyst team needs to better understand 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.

CHAPTER 2

BUSINESS TASK

The business task of the case study is to answer how does a bike-share navigate speedy success?

To find answer for this task, three prominent questions of the marketing analysts team needs to be answered. The questions clearly define the objectives of the case study. The questions defining the business task are as follows:

Three questions that 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?

In order to answer these three questions, the marketing analysts team has performed the data analysis for the "Cyclistic" company for the period of January 2021 to September 2021. The data analysis has been performed in six phases namely, ask, prepare, process, analyse, share and act. Each phase identifies and gathers data for answering the above stated questions. Currently, the business task and its questions fall in the ask phase of data analysis, where the marketing team has defined the questions that need to be solved for guiding the future marketing team. Further, we will proceed with the rest phases of the data analysis to solve the questions. The key stakeholders for the analysis are the members of the executive team whose approval for the findings in the study is required.

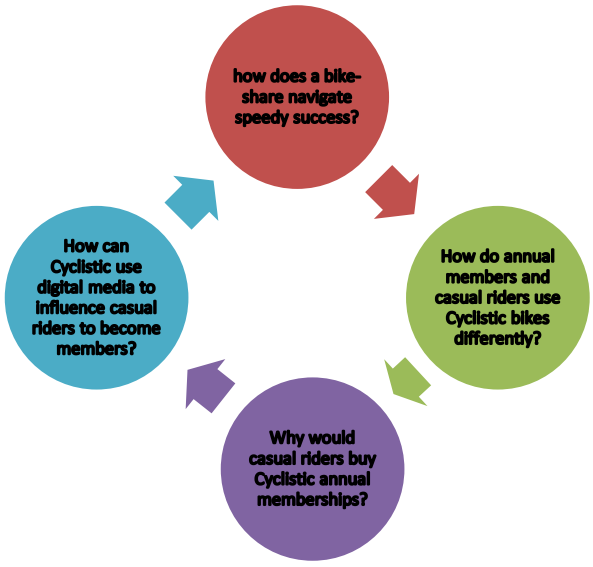

Figure 1 Business Task Diagram

CHAPTER 3

PREPARE PHASE

The prepare phase of the study takes into consideration the sources of data. From where the data has been made available, how is it organized, are there any bias or issues with the data, data safety/security, data integration verification. For our case study, the data has been made available by Motivate International Inc and the license for the data can be verified at https://www.divvybikes.com/data-license-agreement. The data is available at https://divvy-tripdata.s3.amazonaws.com/index.html. The data has been organized into 13 columns with approximate 99000 rows. This is applicable for each monthly data file. The data has been organized under the classifications of ID of the rider, type of ride, start and end date/time, start and end station name with ID and the type of member. The data made available is of the type Open data. It is accessible to everyone for free of cost. The security of the data lies in the hands of the people using the data as per the license policy. The data needs sorting of few columns for proper organization. For example, the date and time are made available in the same column. The alignment of the data for individual cells is not proper. The data available is as follows:

Table 1 Data for PREPARE PHASE

File Name	Data Description & link for the year 2021
202101-divvy-tripdata	Data for the month of January https://divvy-tripdata.s3.amazonaws.com/index.html
202102-divvy-tripdata	Data for the month of February https://divvy-tripdata.s3.amazonaws.com/index.html
202103-divvy-tripdata	Data for the month of March https://divvy-tripdata.s3.amazonaws.com/index.html
202104-divvy-tripdata	Data for the month of April https://divvy-tripdata.s3.amazonaws.com/index.html
202105-divvy-tripdata	Data for the month of May https://divvy-tripdata.s3.amazonaws.com/index.html
202106-divvy-tripdata	Data for the month of June https://divvy-tripdata.s3.amazonaws.com/index.html
202107-divvy-tripdata	Data for the month of July https://divvy-tripdata.s3.amazonaws.com/index.html
202108-divvy-tripdata	Data for the month of August https://divvy-tripdata.s3.amazonaws.com/index.html
202109-divvy-tripdata	Data for the month of September https://divvy-tripdata.s3.amazonaws.com/index.html

The credibility of data lies in the fact that the data has been collected, provided by Motivate International Inc. It is one of the oldest organizations in the field of data science performing collection & organization of the data. In this phase, the data has been sorted as per ridetype, start time, typer of member. Sorting is performed as ridetype: A-Z, start time: Oldest to newest, Type of member: A-Z. This is done for better clarification of the data.

CHAPTER 4

PROCESS PHASE

The process phase of the study takes into consideration the data integrity and cleaning. The data has already been sorted as per the requirement and organized accordingly. The next step is to clean data. We filter data for the BLANKS identified in the spreadsheet in individual columns after visual inspection. Filtering allows us to remove blanks from the field, so that the data becomes error free. The data was filtered by applying filter function for all the columns of all the spreadsheets and then deleting the BLANKS in individual columns. This created blank rows in the data sheet. To remove them, find & select option was selected followed by selecting blanks and then delete cells. This removed blank rows from data sheet. The data was also checked for duplicates. If any duplicate was found, it was removed. In order to find answer to the business task and its three questions, two new columns were created. The column names are RIDE_LENGTH and DAY_OF_WEEK. The first column data is obtained by subtracting start time from end time of the journey i.e. D2-C2 and the weekday was identified by applying WEEKDAY function to column C2 with format-cells-data type as number with zero decimals and 1=SUNDAY, 7=SATURDAY. Thus, we have cleaned the data and we proceed ahead to the next steps of the process. The cleaned data is available as follows:

Table 2 Data From Process Phase

File Name	Data Description & link for the year 2021
Cleaned 202101-divvy-tripdata	Data for the month of January https://drive.google.com/file/d/1UqOMFzM-vT--H7xYnml0qIlZoiqjosxV/view?usp=sharing
Cleaned 202102-divvy-tripdata	Data for the month of February https://drive.google.com/file/d/1EkBbZ3pBLSb7_JxNJENthMF1XRD9HYPT/view?usp=sharing
Cleaned 202103-divvy-tripdata	Data for the month of March https://drive.google.com/file/d/1rWqF6QR1YeRVfVViGnM6wByLbqAeo-3l/view?usp=sharing
Cleaned 202104-divvy-tripdata	Data for the month of April https://drive.google.com/file/d/1xQRpBZjChQbbEEE1N7ZqIV-F84eRQCIG/view?usp=sharing
Cleaned 202105-divvy-tripdata	Data for the month of May https://drive.google.com/file/d/1ykKQO4vrWpL2APcCq0Kngkp2oqXGqz6I/view?usp=sharing
Cleaned 202106-	Data for the month of June https://drive.google.com/file/d/1MHu-vypYzqscM4YS3I0-4Ot-Ih1EKTc6/view?usp=sharing

divvy-tripdata	
Cleaned 202107-divvy-tripdata	Data for the month of July https://drive.google.com/file/d/1QwjW8v4ADC2YJRl7X7Q8EP5RtGz1XasZ/view?usp=sharing
Cleaned 202108-divvy-tripdata	Data for the month of August https://drive.google.com/file/d/1UfdUGaaTFblQWs8CKj4N4fRSECXbBzUz/view?usp=sharing
Cleaned 202109-divvy-tripdata	Data for the month of September https://drive.google.com/file/d/1-k0lUo9r1kKt406gGr1grml1NVorpddF/view?usp=sharing

The process phase thus allowed us to clean, organize, format data. This data can now be utilized for analysis purposes. Since, the data didn't have any calculus, regression analysis was not performed on the data. Otherwise, we would have processed data using regression analysis for differential or integral methods.

CHAPTER 5

ANALYZE PHASE

The analyze phase of the study takes into consideration the data integrity, organization, trend and correlation analysis. It answers how well the data has been organised, is there any trend or pattern being followed, any surprises discovered in the data. The prepare and process phase allowed us to create an organized data with proper cleaning measures being adopted. The data is free of any null values, duplicates, empty columns, empty rows and is organized properly. The surprising factor from the data was that, it had many missing information with respect to starting station, ending station. The fields were blank. But they have been removed while cleaning the data. Now, in order to analyse the data for trend or patterns, we perform three calculations, namely a) average ride_length for members and casual riders, b) average ride_length for users by day_of_week, c) number of rides for users by day_of_week by adding Count of trip_id. These calculations are performed for all the nine months. After these calculations were performed, following data was available for trend and pattern analysis. The analysis calculations can be verified at the following link:

Table 3 Data From Analyze Phase

Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
January	00:22:19	Casual	00:14:27	Sunday	368604	Sunday
	00:12:47	Member	00:14:29	Monday	355839	Monday
			00:14:28	Tuesday	356188	Tuesday
			00:14:27	Wednesday	357345	Wednesday
			00:14:24	Thursday	358711	Thursday
			00:14:27	Friday	367761	Friday
			00:14:29	Saturday	370266	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
February	00:16:45	Casual	00:17:03	Sunday	199764	Sunday
	00:16:45	Member	00:17:29	Monday	195636	Monday
			00:17:24	Tuesday	196340	Tuesday
			00:17:20	Wednesday	195567	Wednesday
			00:17:16	Thursday	192960	Thursday
			00:17:11	Friday	194002	Friday
			00:08:54	Saturday	476061	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**

March	00:17:46	Casual	00:18:54	Sunday	182752	Sunday
	00:17:46	Member	00:17:47	Monday	395446	Monday
			00:17:48	Tuesday	288810	Tuesday
			00:17:46	Wednesday	290341	Wednesday
			00:17:41	Thursday	286212	Thursday
			00:17:46	Friday	297481	Friday
			00:18:26	Saturday	329118	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
April	00:21:15	Casual	00:20:55	Sunday	198254	Sunday
	00:21:15	Member	00:19:26	Monday	195898	Monday
			00:22:41	Tuesday	34011	Tuesday
			00:20:12	Wednesday	41680	Wednesday
			00:22:07	Thursday	343902	Thursday
			00:22:07	Friday	287551	Friday
			00:22:13	Saturday	277686	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
May	00:20:05	Casual	00:19:36	Sunday	218863	Sunday
	00:20:05	Member	00:16:44	Monday	214021	Monday
			00:16:34	Tuesday	214972	Tuesday
			00:15:21	Wednesday	30924	Wednesday
			00:15:53	Thursday	28000	Thursday
			00:16:18	Friday	46560	Friday
			00:20:51	Saturday	351844	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
June	00:21:22	Casual	00:24:34	Sunday	16384	Sunday
	00:21:22	Member	0	Monday	0	Monday
			00:18:53	Tuesday	34002	Tuesday
			00:17:27	Wednesday	46568	Wednesday
			00:19:37	Thursday	67615	Thursday
			00:19:59	Friday	87102	Friday

			00:24:32	Saturday	145390	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
July	00:22:33	Casual	00:26:17	Sunday	19956	Sunday
	00:22:33	Member	00:23:46	Monday	33764	Monday
			0	Tuesday	0	Tuesday
			0	Wednesday	0	Wednesday
			00:17:30	Thursday	77425	Thursday
			00:19:02	Friday	95082	Friday
			00:24:29	Saturday	140000	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
August	00:18:44	Casual	00:22:44	Sunday	18934	Sunday
	00:18:44	Member	00:18:21	Monday	28262	Monday
			00:17:35	Tuesday	43482	Tuesday
			00:17:03	Wednesday	60160	Wednesday
			00:17:14	Thursday	80315	Thursday
			00:18:18	Friday	57048	Friday
			00:00:00	Saturday	0	Saturday
Month	**Ride_Length (hh:mm:ss)**	**Member Type**	**Ride_Length (hh:mm:ss)**	**Weekday**	**Total Riders**	**Weekday**
September	00:19:50	Casual	00:25:28	Sunday	20745	Sunday
	00:19:50	Member	00:23:11	Monday	9152	Monday
			0	Tuesday	0	Tuesday
			00:15:55	Wednesday	64728	Wednesday
			00:15:53	Thursday	78460	Thursday
			00:17:13	Friday	83904	Friday
			00:21:34	Saturday	118937	Saturday

Next we move to the share phase of the project.

CHAPTER 6

SHARE PHASE

In this phase, we insert the trend and relationship diagrams to tell our story. The patterns we have identified and the answers to the business task that we created in the ask phase. Following are the trend and relationship diagrams:

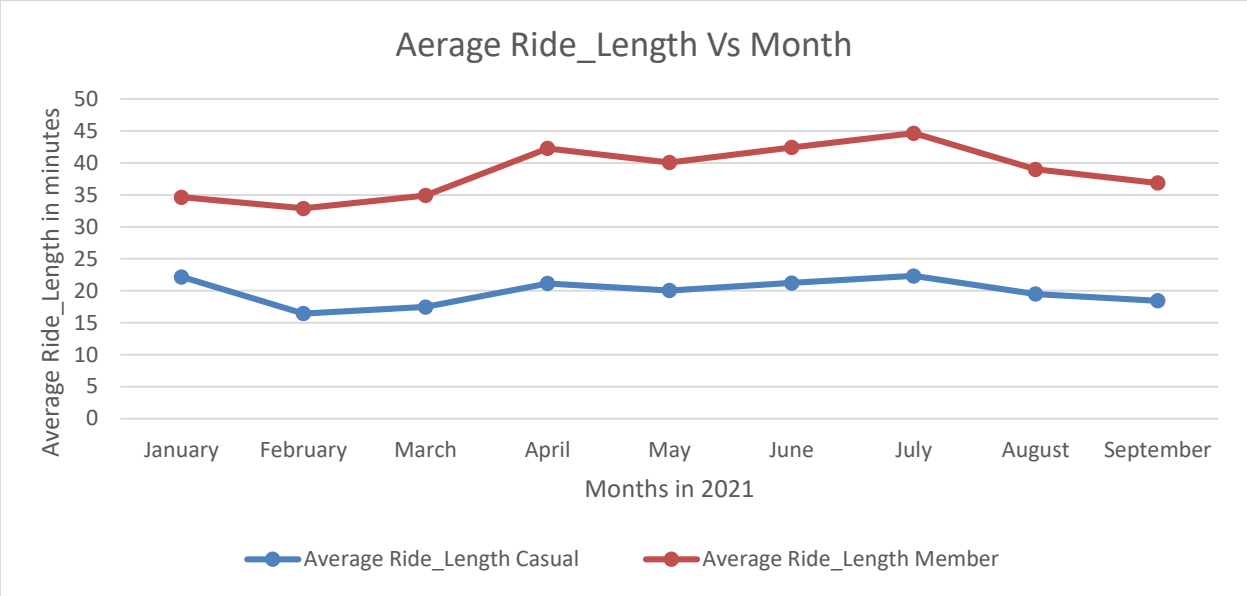

Figure 2 Average Ride_Length Versus Month

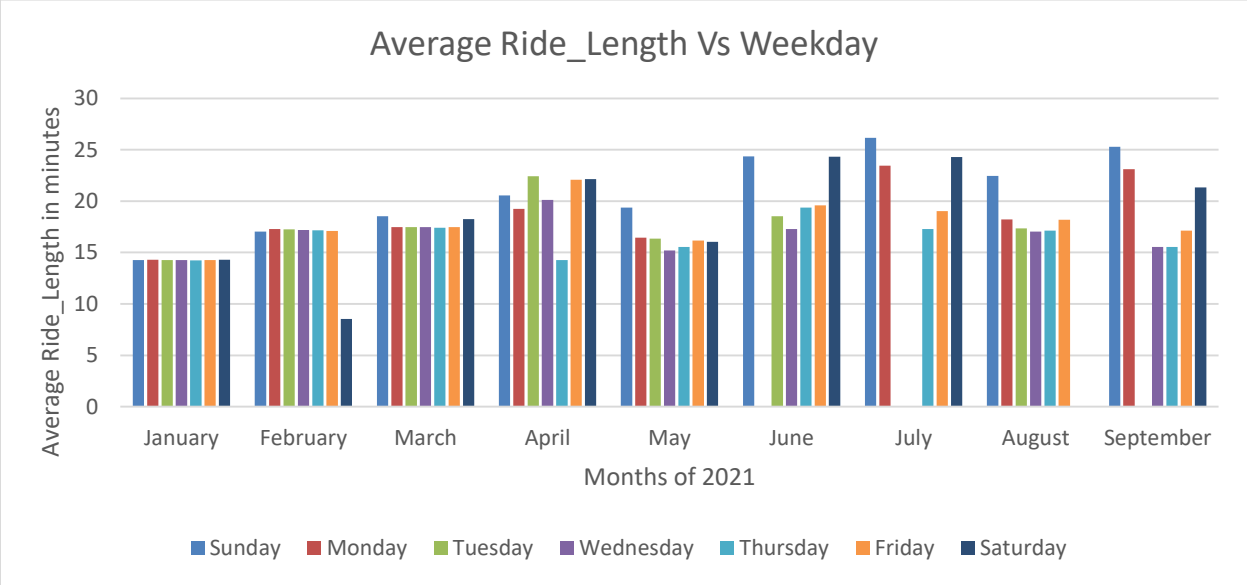

Figure 3 Average Ride_Length Vs Weekday

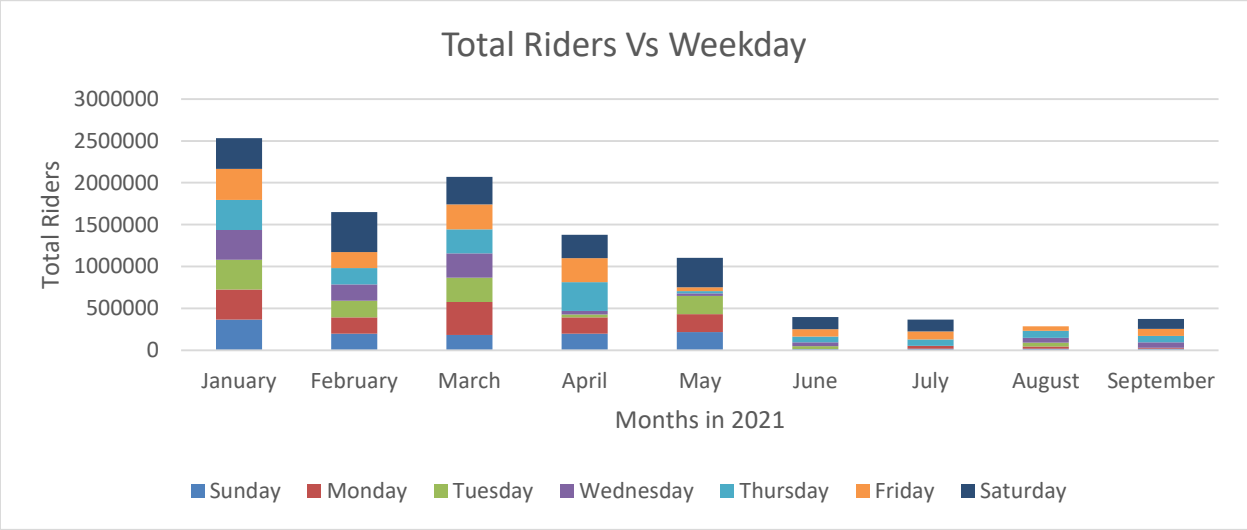

Figure 4 Total Riders Vs Weekday

Thus from figure number 2, we can now answer the first question of the business task. The question being how does annual members and casual members use cyclistic bikes differently. The answer is as follows: Annual members ride the bikes of cyclistic at a greater average ride length as compared to casual members. This is true for each month out of all the 9 months of the year 2021. The highest average ride length being 42.50 minutes for annual member in the month of april 2021 and for the casual member, it is 22.42 minutes. This also validates the claim of the director, marketing analyst team, “Moreno” that, if we convert our casual members into annual members, the company would achieve greater success of heights. This is confirmed as greater the ride length, greater will be the profits.

Another key finding from the data tell that, the average ride length is constant for the months of January, February, March and down the year it fluctuates a lot. Weekdays are more affected over weekends. Cause for his factor has to be found out. Refer figure 3.

Similar trend is seen for the total number of riders. The first quarter of the year, achieves maximum number of riders for the company, while down the year, it fluctuates. Analysts believe that, the specific reason behind fluctuation would be season weather trends affecting the choice of bike transportation. This is intuitive, scientifically other methods need to be adopted for evaluation.

Due to limitations of the software available, the completed data analysis was performed using MS-Excel 2021 version. The same data can be generated using SQL, R, Python programming software if needed.

The data for the share phase can be accessed at the following link:

CHAPTER 7

ACT PHASE

This is the final phase of the data analysis for our case study. Here, we conclude and offer suggestions to the executive team of cyclistic for achieving the marketing strategies for the year 2021. The conclusion drawn from the analysis is that, the marketing team needs to convert maximum members of the cyclistic to its annual membership program. In order to achieve this, Cyclistic can offer discounts to the members of second, third, fourth quarter so as to attract more and more customers. It can extend the membership of the members of first quarter for free and offer free service support. Another way to achieve this milestone is to announce bumper prizes for 10 lucky customers annually. And hand over them free tours to tourism destinations. Another, important work that the Cyclistic team needs to do, is to, make use of social media platforms for promoting schemes of annual membership. The company can target the audience of all states through this. For attracting local state customers, monthly advertisements in local magazines and on the festival boards can be done. Thus, this answers the last two questions of the business task. Furthermore, the marketing analytics team would like to make a cautionary advice for the stakeholders, the above based conclusion and suggestions are purely based on first three quarters of the year 2021. The data for the fourth quarter is under preparation stage. The trend and relationships may change after the data for the fourth quarter has been introduced in the analysis. The marketing analytics team would like to thank all stakeholders for analysing and verifying the data analysis for the case study "**HOW DOES A BIKE-SHARE NAVIGATE SPEEDY SUCCESS?".**

Thank You

Viraj Joshi

Junior Data Analyst

Marketing Analytics Team

Cyclistic, Chicago.

REFERENCES

1. **GOOGLE DATA ANALYTICS PROFESSIONAL certificate program : https://www.coursera.org/professional-certificates/google-data-analytics?utm_source=google&utm_medium=institutions&utm_campaign=GwG-IN-DR-Q32021-LI-workprof-course-readiness-jobready-courseralogo-analytics**