**CASE STUDY:** How Does a Bike-Share Navigate Speedy Success?

Author: Amandeep Singh Gill

Date: 16/10/2022

**Icon

Description automatically generated**

**Table of content:**

* Introduction
* Ask
* Prepare
* Process
* Analyze
* Share
* Conclusions

**Introduction:**

A fictional bike sharing program named Cyclistic offers over 5,800 bikes and 600 docking stations. It offers reclining bikes, hand tricycles, and cargo bikes for people with disabilities and riders who can't use standard two-wheeled bikes. In 2016, the company founded a fleet of bicycles that are geo-tracked and locked to a network of 692 stations across Chicago. Bikes can be unlocked at any station and returned to any other station in the system at any time.

Cyclistic’s marketing strategy previously attempted to build general awareness and appeal to a broad audience. Single-ride passes, full-day passes, and annual memberships are available. A casual rider is a person who purchases a one-day or full-day pass, while a Cyclistic member is someone who subscribes to a membership every year.

Goal: Develop marketing strategies to convert casual riders into annual members.

Business question: "How do annual members and casual riders use Cyclistic bikes differently?"

**Ask:**

* **Identifying the business task**

1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?

* **Stakeholders**

1. Lily Moreno - Director of marketing (Cyclistic)
2. Cyclistic executive team, Cyclistic market analytics team

* **Stakeholder perspective:**

Monero believes company’s future success depends on maximizing the number of annual memberships. She believes rather than creating a marketing campaign targeting all new customers, there is a very good chance to convert casual riders into members

**Prepare:**

1. **Data Source:** Past 8 months of original bike share dataset from 01/01/22 to 31/08/22 were extracted as 8zipped.csv files. A license is provided by Motivate International Inc. for the use of this data.
2. **Data organization and description:**

**File Naming convention:** Cyclistic\_TripData\_YYYYMM

**File content:** Each excel file contains 13 columns containing ride ids, ridership types, ride times, and locations. From different excel files, the number of rows varies between 49k and 531k.

1. **Data Security:**

* Through tokenization, personal information about riders is hidden.
* Separate folders are used to back up original files.

**Process:**

**Overview:**

**Tools used for data verification and cleaning:** Microsoft Excel, R, Tableau**.**

The data was first analyzed separately (month by month) in Excel, and then analyzed as a whole (one year) in R. The last step was to create a dashboard in Tableau to visualize our findings.

**Reasons:**

* Scanning through data in Excel worksheets allows me to gain a better understanding of the dataset by seeing the general outline and basic information. It is also possible for me to check the formatting, missing information, sorting and filtering from the spreadsheet.
* After aggregating the monthly data into a single csv we get approximately 3,000,000 rows of data and 18 columns. Processing of such huge data on excel would not be possible so I decided to switch to R for the statistical analysis.
* R-studio will also be useful for creating new tables with desired information which will be used for data visualization using tableau.

**Data Cleaning and Statistical Analysis using R:**

1. Load all the libraries I used: tidyverse, lubridate, hms, data. table
2. Uploaded all the original data from the data source divytrip into R using read\_csv function to upload all individual csv files and save them in separate data frames.
3. Merged the 8 months of data together using rbind to create a one-year view
4. Created new columns for:

* Ride Length - did this by subtracting end\_at time from start\_at time
* Day of the Week
* Month
* Day
* Year
* Time - convert the time to HH:MM: SS format
* Hour
* Season - Spring, Summer, Winter or Fall
* Time of Day - Night, Morning, Afternoon or Evening

1. Cleaned the data by:

* Removing duplicate rows
* Remove rows with NA values (blank rows)
* Remove where ride\_length is 0 or negative (ride\_length should be a positive number)
* Remove unnecessary columns: ride\_id, start\_station\_id, end\_station\_id, start\_lat, start\_long, end\_lat, end\_lng.

**Analyze and Share Insights:**

Tools used for data exploration & visualizations: R, Tableau

* I created a separate R code that made some changes for specifically the Tableau portion:
* For ride length I rounded the digits by 1. E.g.: 29.8, 12.5.
* Revised how I created my "month" column. I used mutate() to create a column that had the month in \_\_\_ format and not number format. So instead of 01 it would say "January"
* Cleaned the data: removed rows with NA values, removed duplicate rows, removed where ride\_length was 0 or negative and removed unnecessary columns like: ride\_id, start\_station\_id, end\_station\_id, start\_lat, start\_long, end\_lat, end\_lng
* Created a new data frame with this information so I could test the difference between the original data frame (cyclistic\_date) that I used for my analysis and the data frame I would use for Tableau (cyclistic\_tableau).
* In this new data frame I removed more columns to make calculations quicker in Tableau. I removed: start\_station\_name, end\_station\_name, time, started\_at, ended\_at
* Downloaded this data frame into a .csv file which I uploaded to Tableau

**Chart

Description automatically generated**

**Total Bike rides by user type:**

The classic bike is by far the most popular ride with 1,857,000 riders, it leads the rider count in both members and casuals.

**Total rides by user type:**

Members had more rides 57%(1,710,000) as compared to 43%(1,290,000) of casuals.

**Average ride length(member vs casual):**

Average ride length for casuals(25mins) is greater than members(13 mins).

**Chart, bar chart

Description automatically generated**

**Average length of ride(members vs casuals):**

As we had already concluded that the average length of ride for casuals is more than members, but it is interesting to note that the avergae ride length increases on weekends for both user types.

**No. of rides based on user type across a week:**

Highest number of rides were recorded on Saturday (480,000). Througout the week the count of rides is almost uniformly distributed.

**Chart, bar chart

Description automatically generated**

**No. of rides according to weather:**

Due to limitation of data availability till august we can conclude that July (641,000) is the most popular season followed by June (610,000) and August (600,000). Summertime has the busiest months followed by spring and winter. It can be predicted that winters will have less rides but without relevant data it is not possible to come to this conclusion.

**No of rides according to time of day:**

The number of rides start showing an upward trend from 8AM onwards and rise until 5PM following which it starts to decline. Afternoon is the busiest for both types of riders and constitutes 40% of total rides, night has the least no. of rides, 4AM having the least rides.

**Act:**

**Analysis:**

* Members had the most rides at 57% of total rides. However, the average ride length of casuals was greater than that of members by 14 mins.
* The most popular bike type for members and casuals was classic bike. The casuals used the docked bikes which were not used by members.
* The busiest time with maximum rides was afternoon(12PM-5PM) for both members and casuals, nights had least rides with 4AM being least busy.
* Saturday was the most popular weekday for both members & casuals. Members used the bikes equally throughout the week with slight increase on Wednesday, while casuals tend to have more rides on the weekends as compared to weekdays.
* Summer (June-August) was the busiest season with the maximum rides in July. The number of rides were minimum in winter months (January-February).

**Business Suggestions:**

I have the following suggestions for the marketing team to convert casual riders into annual members:

1. Personalized discounts and benefits for members based on their riding habits and preferences.
2. Users that take long rides could be offered a discount on renting bikes or better yet annual Members that travel more than a certain distance say for instance 10 miles should be offered a discount.
3. Make sure to highlight the benefits of memberships, including discounts during busy times like the summer or weekends.
4. By inviting existing members to share how Cyclistic's system has changed their lives to create a sense of community, and offer a discount if they do so, new riders will be more likely to join.
5. Cyclistic bike-share should introduce a weekend only membership that costs less than the current 7-day membership.

I’m confident these recommendations will encourage Casual riders to upgrade to annual membership thus increasing revenue and profits at Cyclistic bike-share.