Google Analytics Capstone : Bike Sharing Case Study

Russhel De Leon

3/10/2022

How Does a Bike-Share Navigate Speedy Success?

Google Analytics Capstone

Scenario The director of marketing believes the company's future success depends on maximizing the number of annual memberships. Therefore, your team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights,your team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.

Ask

The business task is to better understand how annual members and casual riders differ. A descriptive analysis will help them to understand the difference between the two types of riders and to develop a new marketing strategy in converting casual riders into annual members.

Prepare

The data source can be found at https://divvy-tripdata.s3.amazonaws.com/index.html. Data is structured and organized in a wide format where each row is an observation and each column is a variable.

The datasets have a different name because Cyclistic is a fictional company. For the purposes of this case study, the datasets are appropriate and will enable us to answer the business questions. The data has been made available by Motivate International Inc. under this license: https://www.divvybikes.com/data-license-agreement.

Process

This step involves examining and cleaning the data. The cleaning involves standardize each column type, making sure that each column is consistent, removing duplicates and etc. The files are large and consist of million rows. RStudio is a flexible tool that will help us to clean and analyze these files.

Step 1: Setting up our environment.

Tidyverse helps us wrangle data.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.3     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.0     v forcats 0.5.1
```

```
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Lubridate helps us to format and clean date.

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

Janitor package helps us to clean column.

```
library(janitor)
```

```
##
## Attaching package: 'janitor'
```

```
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```