Collin McClendon

Cyclistic Bike-Share Case Study

Ask

-This is a case study of Cyclistic, a bike-sharing company in Chicago.

-There are casual riders and those who apply for annual memberships; the main goal is to convert casual riders to member riders.

-The problem I am solving is “how do annual members and casual riders use Cyclistic bikes differently”?

Prepare

-The data that is used is located at: https://divvy-tripdata.s3.amazonaws.com/index.html

-Each dataset covers a month and are organized by start date.

-A potential Bias for the data could be that while classic and electric bikes are used by both casual riders and members, docked bikes were only used by casual riders.

-Some Station Names/ Ids and Latitude and Longitude values were empty.

-A few Latitude and Longitude values were out of the Chicago range.

-The Case Study said the data was appropriate to use and would answer business questions.

-Each row has data assigned to either a casual rider or a member.

-A problem with the latitude and longitude values is that even if you calculate the distance between two points, the bike paths are usually not straight and may not be enough for an accurate travel distance.

Process

-I used a combination of Python and SQL (in this case Python 3 and sqlite 3)

-Python3 and sqlite3 can display and alter the data at good.

-The steps I have taken to ensure that the data is clean is checking for duplicates, end times that are earlier than start times, null values, latitude and longitude values that are not in the Chicago range, and trailing spaces in text values.

-Printing the number of rows before and after cleaning, as well as printing the number of rows deleted when taking a step to clean will ensure that the data is clean and analyze ready.

-The cleaning process is shown in the Python code.

Analyze

-The data should be organized in the order of start dates.

-The data is formatted by taking the 12 months of data (csv), put them together as one data in Python, then do the data cleaning and removed columns that are not need for analysis before converting the data as a big csv file.

-I also made calculations of the time difference of start time and end time in hours, day of the week and the year-month based on the date.

-The new csv file is taken to Tableau Public to visualize the data in different ways involving the two different rider types (casual and member).

-Line Graphs, Bar Charts, and Pie Charts were used.

Share

-Based on the visualizations for the Case Study:

\*Members have more bike rides than Casuals all year.

\*Casuals spend more time on bikes outside of winter.

\*Members ride more on a weekday while casuals ride more in the weekend.

\*Bike Rides are more common in the Summer and less common in the winter.

\*The most common Bike Stations for casuals are Streeter Dr & Grand Ave, DuSable Lake Shore Dr & Monroe St, and Millennium Park.

\*The most common Bike Stations for members are Kingsbury St & Kinzie, Clark St & Elm St, and Wells St & Concord Ln.

\*Not counting docked bikes (only used by causal riders), casual riders spend the most time on classic bikes. Casuals also have more bike trips on classic bikes than the other bike types.

-The findings above help with answering the question of “how do annual members and casual riders use Cyclistic bikes differently”?

-I did my findings based on using the data column for rider types (casual and member).

-The audience is the executive team for Cyclistic.

-I made use of line, bar, and pie chart data to help communicate the data to the audience.

-The visualizations were made in Tableau Public:

https://public.tableau.com/app/profile/collin.mcclendon/viz/CyclisticBikeData\_16690709427350/Sheet9#1

Act

-Conclusions on the data and visualizations:

-Recommendations:

\*Advertise the program during the weekends.

\*Offer discounts and/or trial for classic bikes in the Summer.

\*Focus on the common stations for casuals, particularly Streeter Dr & Grand Ave, DuSable Lake Shore Dr & Monroe St, and Millennium Park.

-The Stakeholders will make a business decision based on the findings for the analysis.

-An additional data piece that may help would be the precise number of miles traveled by bike.