Project 5 --> Ford GoBike System analysis and visualization.

Project overview

This work on this project is divided in two main parts. For the first part, I will apply an exploratory data analysis on Ford GoBike System Data Flights dataset. I will use Python and data visualization libraries to investigate and extract all the dataset's variables and information, . The analysis in this section will be conducted as following:

1.The first section will be all about univariate exploration . 2.The second one will be all about investigating multivariate's relationships.

In the second part, I will pick my major findings from my exploration and carry my output to others through an explanatory analysis. At the end , I will create a slide deck that will be polished then i will be able to use my explanatory visualizations to communicate with my results. This second part of the project will be very useful in my conclusion

STEP 1: CHOOSE THE DATASET

I choose Ford GoBike bike ride data from 2019/02 for my project. The data was downloaded from this website https://www.fordgobike.com/system-data. Ford GoBike is the Bay Area's new bike share system, with thousands of public bikes for use across San Francisco, East Bay and San Jose.

STEP 2: EXPLORE THE DATA

I'll explore FOrd GoBike data and put my findings in a report. The report will have introduction of the dataset,and all the steps i will perform in order to reach my output.

STEP 3: DOCUMENT MY FINDING

At the end of the exploration, I'll have the results of my work that I’ve detected,and I will regulate my findings. Furthermore, Iwill write my main results that I have found and reflect on the steps I took in my data exploration in this readme document.

SUMMARY OF MY WORK

In the exploration, I discover that there are two types of clients using the system: subscribers The majority of users have 20-30 years old compared to other user groups. People hire a bike on weekdays more than weekends especially on Thursday. Percentage of subscribers is almost %90.4 while percentage of customers is almost 9.6%. Subscribers are the majority of users and their trip duration time is around 10 minutes while customers do not use bikes more often and they took along time to finish their trip , we can conclude here most of subscribers use this service for daily activity or going the same place everday like work or school.Finally, it appears that 20 to 30 years old age group use the service the most.

KEY INSIGHTS FOR PRESENTATION

For the presentation, I focus on the influence age group, weekdays, time effects on renting a bike.I started by plotting visualization of time distrubition, daily bike hiring trend, user type,gender followed by age group distribution, then plot the weekday, timeframe with age data. I used different color palettes for each variables to make sure that it is obvious ,interesting.