Project: (Ford GoBike Data Exploration)

by Amal Almutairi

Dataset

I have used Ford GoBike Dataset, which are a dataset contains data about bike-racing attributes for approximately 519700 bike.

Summary of Findings

We discovered In **Univariate Exploration** the duration trips has an inverse relation with age, the age distrburtion is fair. the trips are less in colder months and higher in warmer months , The weekends has alot less trips than the weekday , According to this graph it is concluded that maximum bike rides are taken from may to october in year 2017 , According to this graph it is clearly visible that there are nearly 7 times more Subscriber than Customer , amd Most of the trips are initiated from The Embarcadero at Sansome St , The end of most trips san francisco ferry building .
And **we Bivariate Exploration** I will look at the relationship between (duration_sec) and (user_type) using violing plot , we see the (Subscribers) often finish the race faster than the (Customers) , We noticed that the subscribers has the most number of trips on October and on the other side where the customers has the most number of trips in September , and we can see that july has the most time spent in min .
And we **Multivariate Exploration** I will Customers have more distance than the subscriber for all the month .
The last we discovered Trip Duration in different months taking into account the user type, now we can see the Customer has much higher trip duration in average.

Key Insights for Presentation

In this presentation. Many visualizations on (Ford GoBike System Data characterics will be created to study their relationship on duration_sec The main characterics are start_time, end_time, start_station_name, end_station_name, start_station_name and, user_type , year-month ,start_station_latitude , start_station_longitude , end_station_latitude , end_station_longitude , is the most helpful features to help my go through the investigation part in our dataset