**Communicate-Data-Findings**

# by Apurva Verma

# Dataset- Ford GoBike System Dateset

Ford GoBike System Dateset is a dateset that contains trip data from lyft's bike service for public use. This project is divided into two major parts. In the first part, you will conduct an **exploratory** data analysis on a dataset of your choosing. This project is divided into two major parts. In the first part, you will conduct an **exploratory** data analysis on a dataset of your choosing. Variables including, trip duration, start time and end time with date, start station and end station names, start and end coordinates and customer types.

The Dataset can be found on the website  <https://www.fordgobike.com/system-data>.



# Main findings from the exploratory data analysis.

In the exploration, I found that there are two types of clients using the system: subscribers and customers. The subscribers are those who are mainly daily commuters, having short trips to and from work. The customers or occasional riders are those who use the system mainly on weekends to explore the Bay Area. The bike share system was used more often around summertime (May-October) with a clear drop from January to March, most probably due to the weather condition.

* The most number of bike trips toped around 8.00am and 17:00pm during the day, also there were more number of trips on the working days (Monday - Friday) than on the weekends(Saturday, Sunday). The riding trips tend to be shorter/quicker on the working days compared to weekends.
* There are more subscribers type usage than the customers type Subscribers are prone to use the bike sharing system for work purposes and especially during the rush hours ,while it seems that the customers are those who tend to use the service on any purpose rarely tend to ride for fun or recreation activities in the noon or in early evenings over the weekends preferably.

To conclude this analysis , I say that bike riding service was predominantly use by working professionals who got subscriptions for the rides on a daily basis to head to work and to return from work, while the non- subscribers used the service for preferably recreational activities in the evenings.

#### Key Insights for Presentation

#### In this project I've investigated the dataset of FordGoBike Data for the year 2019(January-December) and from the visualization drawn few conclusion listed below.

#### 

* From the two user types of FordGoBike data, the counts of Subscribers are around 3 times more than the Customers.
* 90% of the users tend to share their bike for all trips in which only Subscribers contribute. None of the Customers tend to share their bike for all trips.
* As a Start Day of the Trip, Tuesday and Thursday have the highest frequency of users. While the weekends have the lowest number of users.
* April and July are the months when the user count is the most while December has the lowest count.
* Most of the trip duration last within the range of 5-10 minutes.
* For different hours of the day, at 8am in morning and the evening 5pm are the hours where the bikes are most used by the users.
* Subscribers rides are much shorter compared to those of the customers on everyday of the week. The trend line seems to increase on weekends, especially for customer user type .
* The heatmap represents the Subscribers weekly distribution for Start Time of the hour. Most of the Subscribers prefer to start their ride between 7-9am in morning and 4-6pm at evening. Weekend activity of Subscribers are not very strong. The Customers weekly distribution for Start Time of the hour. Most of the Customers prefer to start their ride between 8-9am in morning and 4-7pm at evening. Also during weekends they ride between 11am-6pm.
* Most of the trip duration last within the range of 5-10 minutes. Also almost none of them are used for more than a hour. The box plots we have conclude that the ride duration(in minutes) for Customer typically have trip durations that are longer (8 to 18 minutes ) than for subscribers (5 to 12 ).

**References**

* Numpy Documentation
* Stack Overflow
* MatplotLib Documenation