**Project 2**

**Aim:**

The aim is to collect and analyze the data in detail of the Uber Data. Through the analysis of data, we will find some important insights.

**Introduction:**

Uber provides the 2016 data. By using this data find out the demand and supply Analysis gap of the cab.

Provide Uber dataset have the history of pickup request genrated by customers.

**Problem Statement:**

We will find out how many request id is placed from the pickup point of the Airport and city, How many Drivers are present, How many trips are completed, and how many trips are cancelled also find out how many trips the car is not available for the trips.

**Methodology:**

In the given Uber Dataset the first column is Request id in this column the request-id is present the placed by the customer.

In the second column, the pickup point is present where only two Airport and City pickup point is present.

The next column is Driver ID in this column so many drivers are present each driver assigns a separate driver ID.

Another column is Status in this Status column three statuses are present completed, cancelled and No car available.

The last two columns are the Request Timestamp and Drop Timestamp in this column the request time and drop time are present.

The first step is cleaning the data. In data cleaning, we cleaned the data with the transform data option present in Power BI because the dataset was not cleaned and included some missing values, so with the help of the duplicates function in Power BI, we cleaned the data and replaced the null values with 0 and then started working on it. After Data cleaning perform the Data interpretation step.

Given dataset presents 6746 rows and 4 columns.

We have replaced the null value from zero to remove errors.

**Analysis:**

 We have prepared three sheets for visualization:

* Trip Status --> In this sheet, we made a stacked column chart to compare different trip status like completed, cancelled or no cars available with the numbers of trip and their pickup locations.
* Demand and Supply --> We made a comparison plot between how much the demand and how much supply we are providing. Also make a gauge to showcase how much the gap difference.
* Demands with time --> By using conditional formatting we have divided request time in 5 different sections which are early morning(00:00 to 5:00), morning(after 5:00 to till 11:00), Afternoon(after 11:00 to till 15:00), evening(after 15:00 to till 17:00) and night(after 17:00 to till 00:00).

**Insights:**

* Sheet 1--> Sufficient number of cabs are not present at the airport for pickups. People from the city area are likely to cancel more cabs than the airport area.
* Sheet 2 --> Total number of demand is 6745 but we are providing supply only 2831 which means we are lagging by 41.97%. So, to make more profit we need to reduce the lagging percentage.
* Sheet 3 --> Most of the trips are cancelled in night because no proper supply of cars are available at that time. During the afternoon, there is less chance for trips to get cancelled.

**Recommendation:**

 We need to increase the number of cars at airport and in night for both pickup locations (city and airport) as less cars available in night.

Like this dataset, we can Analyze the Data according to the Different Cities-Area it takes more clarity to how to fill the supply and demand gap.

**Conclusion:**

According to analysis, almost 1713 trips are cancelled for Airport trips. In airports, trips need to increase the drivers that help to increase the Business. The second point is that almost 1066 trips are cancelled by the customer in the city area and need to improve some features. A total of 1327 Drivers successfully completed the Airport trips and 1504 Drivers successfully completed the City trips.