**Case Study: Analysis of Ride Booking Dataset**

**Understand Problem Statement**

The Problem statement contains the ride booking information including several attributes. the data has missing values outliers and inappropriate formats which may affect the analysis the problem is to clen and preprocess the data and get the required results and insights and obtained desired visuals.

**Data Collection / Sourcing**

* In this case the dataset is provided in the form of csv file, read the csv file import it to data frame for further operations.

**Methodology**

* **Data Cleaning / Understand your variables**

1. Start by importing all the required libraries-based n the operation needs to be performed: pandas, NumPy, matplotlib, and display.
2. Data Cleaning: Find all the missing values and fill them with appropriate data using the fillna () function.
3. Convert the date and Time column to the appropriate format.

**Results and Insights**

* **Descriptive Analysis:**

1. The most popular vehicle is an auto.
2. The least popular vehicle is Uber XL.
3. Average ride distance is 24.64.
4. Average booking value is 499.
5. Distribution of ratings across drivers and customers is visualized using a histogram.

* **Customer Behavior Insights:**

1. Drivers are the most frequent cancellers of rides.
2. Total number of rides canceled by drivers are 27,000.

* **Operational Matrices:**

1. Ashram is the location with high demand, with a total count of 936 rides.
2. 05:44:57 PM is the time slot with highest demand

Prajwal S. Chougule

Prajwalchougule41@gmail.com