# Analysis of dataset UBER

**Uber Technologies, Inc. (Uber) is an American mobility as a service provider. It is based in San Francisco with operations in approximately 72 countries and 10,500 cities. Its services include ride-hailing, food delivery (Uber Eats and Postmates), package delivery, couriers, freight transportation, electric bicycle and motorized scooter rental via a partnership with Lime, and ferry transport in partnership with local operators. Founded in 2009.**

## Dataset

The dataset used for this project was found on Kaggle. The basic idea of analyzing the IPL dataset is to get a fair idea about to understand the purpose of uber Drive. To anlyze the miles frequent used route.

## Tools & Libraries

• Python • Jupyter Notebook • Pandas • Numpy • Seaborn • Matplotlib • Plotly & Cufflinks

## Data Description

- Start_Date
- End_Date
- Category
- Start
- Stop
- Miles
- Purpose

## Data Cleaning

1) To find of NaN values
2) To Drop NaN Values

# EDA

I looked at the different data and below is a few highlights of theanalysis.

- Import Libraries
- To read file using
- Size of columns and rows
- Top 5-Observation
- Bottom 5-Observation
- Information of Data
- Name of columns
- To find out missing values
- Heatmap for NaN values
- Drop NaN Values
- Heatmap after drop  NaN Values
- Some statistical Inference
- Unique start destinations
- Unique  stop destinations
- How long people travel with uber?
- Most Popular Starting point Uber Driver(Top5)
- Most Popular Stop Point for Uber Driver(Top5)
- Most Frequent Route for UBER
- Purpose for UBER Drive
- Bar Plot for Purpose vs Distance
- Heatmap Using Correlation
- Bar Plot for category.