Project Title: Uber supply demand gap analysis.

Problem definition: App-based taxis/cabs like Uber are used daily, for traveling towards or from an airport. As customers we often face the issue of ride cancellation (by driver) or non-availability of the cars. These issues also impact Uber's business resulting in significant revenue loss. Following are some of the key issues:

- Cancellation of rides going towards the airport were higher than regular trips.
- Trips to and from an airport resulted in high consumption of fuel and time. Hence, a trip back to the city without a rider is not economically beneficial for the driver.
- Due to high variance in flight arrivals (higher during evening, late night hours) the driver idle time is higher in morning. As a result, no cars are available during peak night hours because working hours ends for majority drivers at night.

Solution Approach: Exploratory data analysis on a masked data set using standard python libraries (e.g., NumPy, Pandas, Matplotlib, etc.) to identify the root cause of the problem and recommend ways to improve the situation.

Data Cleaning and Preparation:

- Convert request and drop time stamps to standard datetime format.
- Remove null values in the data.
- Derive new variables/metrics (e.g., 'Request Hours' and 'Drop Hours') using timestamp.
- Divide all requests into time slots (e.g., 'Early Morning', 'Peak Morning Hours', etc.).

Plots:

- Count plot for all time slot bins identified above to determine frequency of request for each status (Cancelled/Completed/No cars available).
- Graph showing demand and supply gap for all time slot bins for each type of ride (i.e., City to Airport and Airport to City).

Tools: Jupyter notebook, PyCharm, Python libraries – NumPy, Pandas, Matplotlib, Seaborn.

Deliverables: Jupyter notebook, Data set, Bibliography, Project video, Elevator speech.

Dataset: Uber requests data.

- Region – India (Only the trips to and from the airport are being considered.)
- Number of requests - 6475
- Key Features - Time of request, Drop-off time, Pick-up point, Driver id, Status of the request (Completed/Cancelled by the driver/No cars available).

Expected Results: Visually identify the most pressing problems for Uber such as:

- Identify the most problematic request (e.g., city to airport / airport to city, etc.) and their time slots (e.g., early mornings, late evenings etc.).
- Find the gap between supply and demand for different time slots. Also determine the time slot where maximum gap exists.
- Find the type of requests (e.g., city to airport or airport to city) for which the gap is the most severe in the identified time slots.
- Identify the reason behind the supply-demand gap and recommend possible resolutions.