Uber Demand and Supply Case Study

An analytical approach to solving the demand and supply gap because of driver cancellation and non-availability of cars. Analysis of the root-causes and possible hypotheses of the problems and recommended ways to tackle them.

Prachi Agrawal

Analysis Objectives

Business Objectives:

1. Identify root cause for:
   - **Non-availability of Cars**
   - **Cancellation of Cars**
2. Possible hypotheses of the problem(s)
3. Recommend solutions for driving to tackle the issues.

Data Dictionary:

1. Trips are to and fro between city and airport
2. Trip Status- "Completed", "Cancelled" and "Not Available"
3. Data is of 5 weekdays spanning across all 24 hours
4. Information available is of 6700+ trips

Assumptions made for the analysis

1. Demand at any Pickup Point is equal to sum of Trips Completed, Trips Cancelled and No Cabs Available.

2. Supply at any Pickup Point is equal to sum of Trips Completed and Trips Cancelled.

3. Supply of cabs also includes Cancelled trips because the cabs are available but cancelled. Reasons for these cancellations would require more data for analysis.

4. Assuming the cancellation of the cars is by the drivers and not by the customers.

5. The timezone is IST.
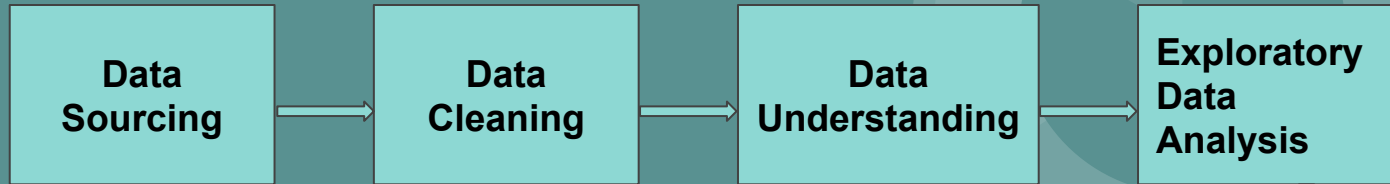
Data Sourcing and Cleaning

CSV based data:
- 6745 Rows, 6 Columns
- 2 Character Features and 4 Numeric Features

Analysis of null values and missing values in the data:
- Total null values= 6564
- Column Name- Driver_id= 2650
- Column Name- Drop_timestamp= 3914
- There are no missing values in the dataset.
- Not dropping or imputing it with values, as it will be used for further analysis.

Data Understanding

	Request id	Pickup point	Driver id	Status	Request timestamp	Drop timestamp
0	619	Airport	1.0	Trip Completed	11/7/2016 11:51	11/7/2016 13:00
1	867	Airport	1.0	Trip Completed	11/7/2016 17:57	11/7/2016 18:47
2	1807	City	1.0	Trip Completed	12/7/2016 9:17	12/7/2016 9:58
3	2532	Airport	1.0	Trip Completed	12/7/2016 21:08	12/7/2016 22:03
4	3112	City	1.0	Trip Completed	13-07-2016 08:33:16	13-07-2016 09:25:47
5	3879	Airport	1.0	Trip Completed	13-07-2016 21:57:28	13-07-2016 22:28:59
6	4270	Airport	1.0	Trip Completed	14-07-2016 06:15:32	14-07-2016 07:13:15
7	5510	Airport	1.0	Trip Completed	15-07-2016 05:11:52	15-07-2016 06:07:52
8	6248	City	1.0	Trip Completed	15-07-2016 17:57:27	15-07-2016 18:50:51
9	267	City	2.0	Trip Completed	11/7/2016 6:46	11/7/2016 7:25

Insights from initial look at the data:
1. The data is for trips in between the Airport and City Area
2. The Date columns i.e.., Request timestamp and Drop timestamp are not in consistent format.

Data Manipulation
Type/Business-Driven Metrics derived

1. Renaming Columns for efficient coding.
   New Column Names-- Request_id', 'Pickup_point', 'Driver_id', 'Status', 'Request_timestamp', 'Drop_timestamp'
2. Standardizing 'Request_timestamp' and 'Drop_timestamp' column to bring about consistency in the date values.
3. New Type Driven Metrics Derived-
   - Drop_date
   - Request_date
   - Drop_hour
   - Request_hour
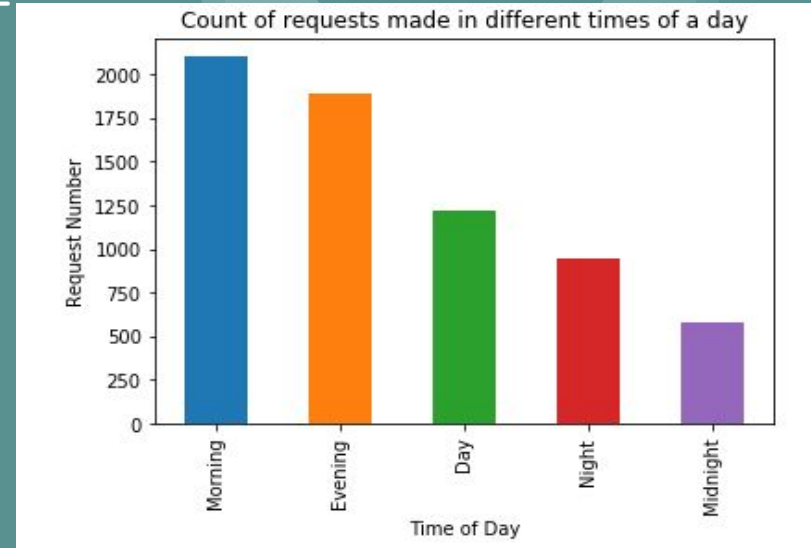4. Business Driven Metrics Derived-
   - Trip Time

Data Manipulation
Data-Driven Metrics derived

1. Slotting of day by time: Time_of_day
   Conditions applied for slotting- based on Request_hour
   Five slots have been created-
   - Midnight
   - Morning
   - Day
   - Evening
   - Night

Count of requests made in different times of a day

Insights

1. Count of cabs requested on each of six dates is approximately the same. The same can be inferred from the graph.

2. So, for further analysis, the demand and Supply gap will be analysed for a single day. And the same insights will be assumed to be applicable for all the dates.

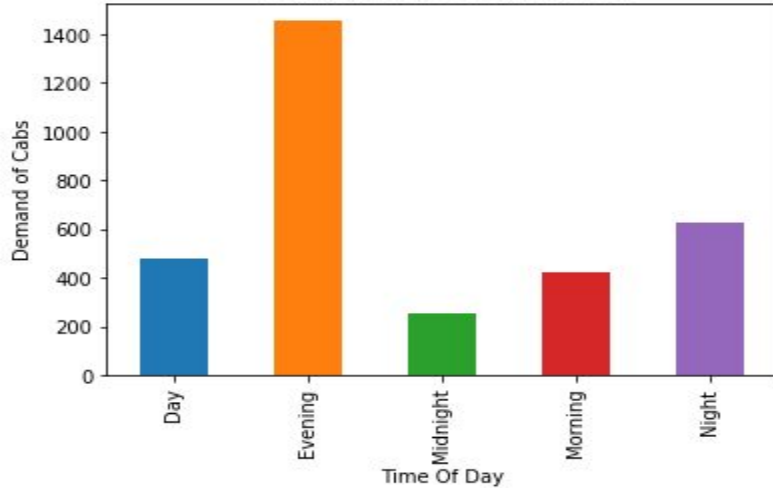3. Mean number of cabs requested each day =1349 trips

Count of cabs requested at different dates
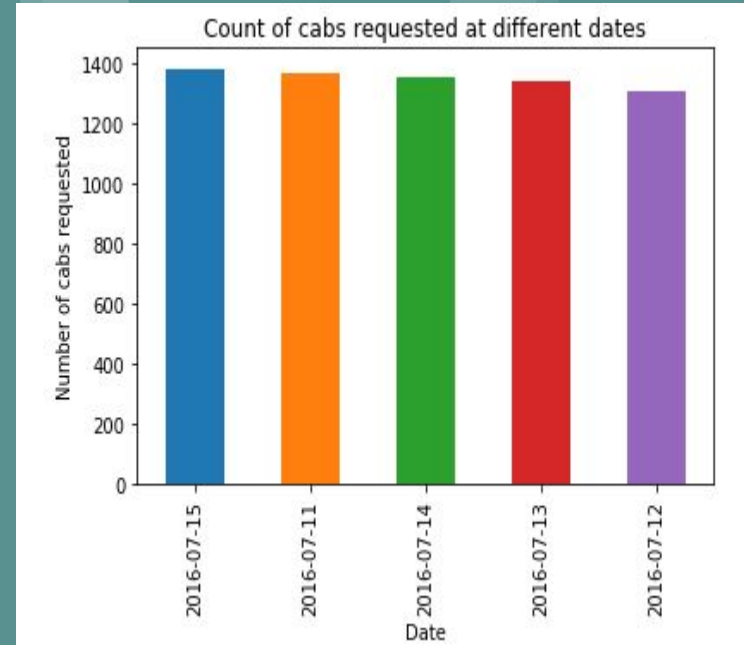
Segmented Bivariate Analysis

===

Further analysis of the given dataset will be divided into two segments
 1. Trips from Airport to City that is Pickup_point= 'Airport'
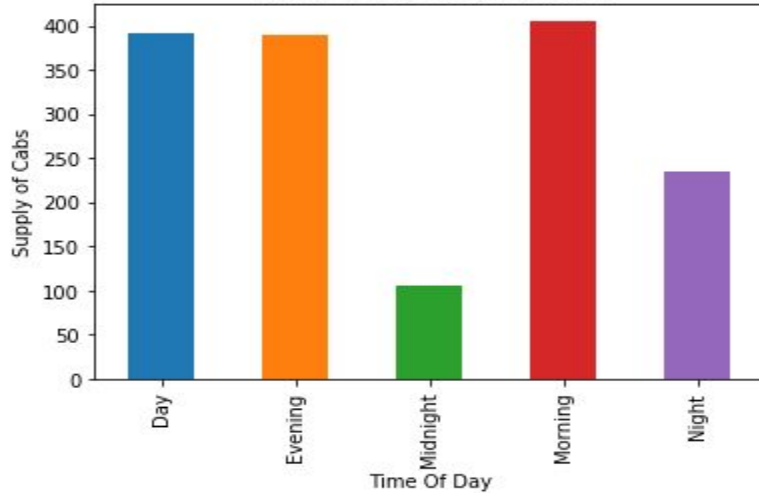 2. Trips from City to Airport that is Pickup_point='City'

===

Scenario 1: Trips from Airport to City
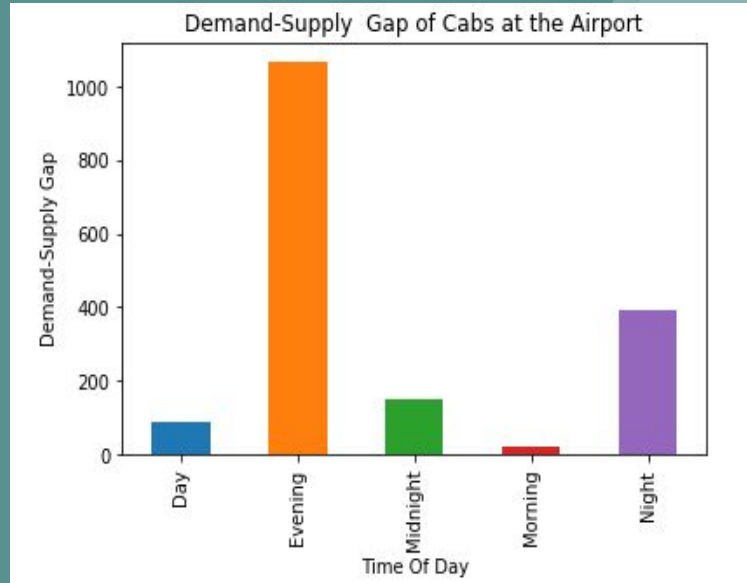## Calculating Demand and Supply at the Airport

Insights:
1. The demand for cabs at the airport is the highest at the evening time.
2. The supply of cabs at different times is somewhat comparable, but are much less than demand.

Scenario 1: Trips from Airport to City
Calculating Demand-Supply Gap
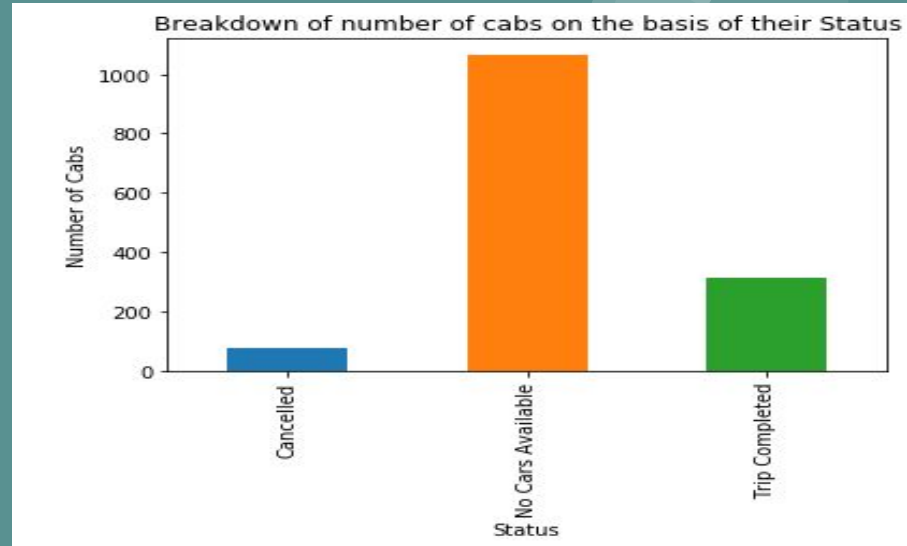

Demand-Supply Gap of Cabs at the Airport

Insights:
1. The gap in between demand and supply of cabs is highest in the evenings.
2. For other times of the day, the gap is not as significant.

Scenario 1: Trips from Airport to City
Root cause for high demand-supply gap in the evening

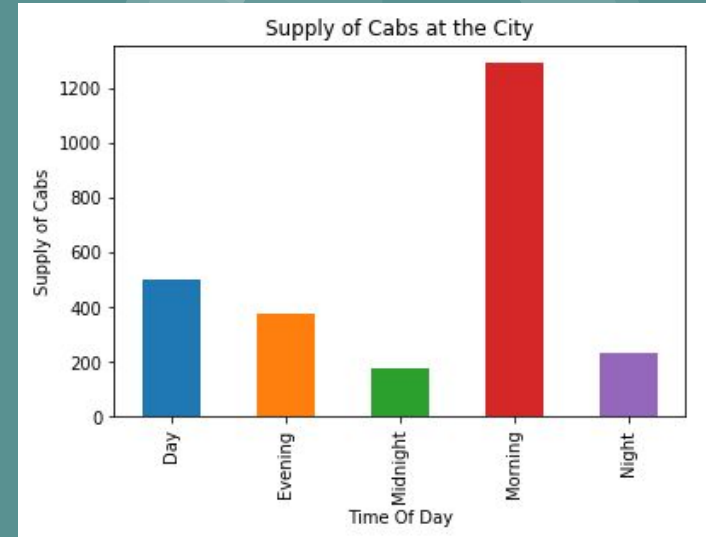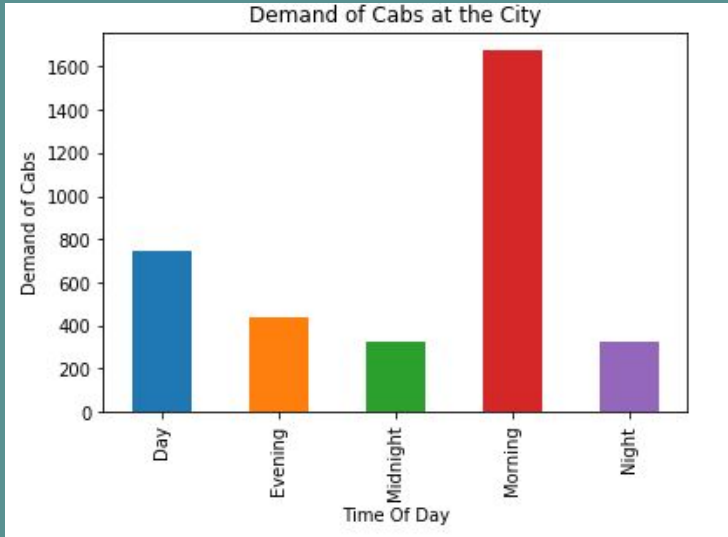Breakdown of number of cabs on the basis of their Status

Insights:
1. <u>Non-availability of cars</u> is the major reason for high demand-supply gap in the evenings at the airport
2. The number of trips cancelled is significantly less.

Scenario 2: Trips from City to Airport
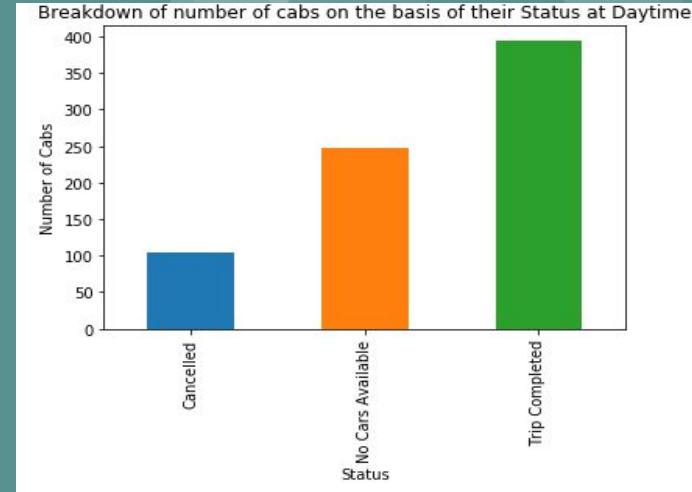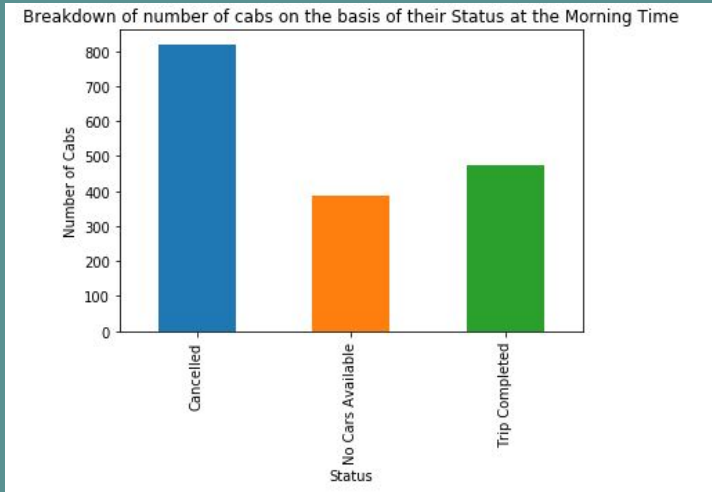## Calculating Demand and Supply at the City

Insights:
1. The demand and supply of cabs at the city area is high in the mornings.
2. Also both demand and supply for different times of the day are comparable.

Scenario 1: Trips from City to Airport
Root cause of high demand supply gap in the mornings

Insights:
1. Cancellation of trips by the driver is the root cause for high demand-supply gap in the city area in the morning time.
2. Further in the day-time the number of trips cancelled or non-availability of cabs is relatively insignificant.

Final Conclusions from the Analysis

Demand-Supply Gap--
- Significant Demand-Supply Gap for both pickup points
- The Demand-Supply Gap at the airport is greater than at the city area.

Reasons--
- At airport, major cause is non-availability of cabs
- At city, major cause is high number of car cancellations. May be because of high wait time at the airport for the next trip.

Recommendation--
- Reach out to more number of drivers and strengthen the Uber network to increase availability of cars both at the airport and city.
- Incentives and perks for drivers who wait for long times for trips or serve at odd times, so that number of trips cancelled by them is less.
- Add a tipping option like being done by a few food delivery apps. Will act as a motivation for the drivers.
- Rewarding drivers on the basis of their ratings.
- Partnerships with various tech/business parks; employees can pool together for cost-efficient and safe commute.