

# UBER SUPPLY-DEMAND GAP CASE STUDY

## SUBMISSION

Submitted By : Barkha Garg

# Objective & Approach

The purposes of this Case Studies are the following:

- Analyse the high **Supply-Demand Gap** that is leading to discontent of the Ride Seekers.
- Analyse the trends and the influencing factors that lead to this gap and present them to visually assess the pain areas and **recommend some ways to resolve the supply-demand gap**

## Approach:

The analysis was carried out using the Uber request data of various trips to and from Airport. This dataset doesn't include other rides, but only the rides originating from City and going to the Airport and the vice-versa. (source of data: Uber). The language used to analyse and visualize is Python and the scripting tool used is Jupyter Notebook.

The entire process was divided into sequence of steps (individual tasks) with output of tasks helping us infer some useful insights about the problem and also look at them visually for a better understanding. A detailed EDA (Exploratory Data Analysis) was carried out on individual attributes and pairs of attributes to do the Univariate and Bivariate analysis.

At the end of these tasks/steps, some useful inferences were drawn to factor down the key problems that result in this huge Demand & Supply gap and also see what is the nature of these requests (What is the Pickup point, at what hour of the day, frequency of requests is higher and what was the status of each requests). This helped in thinking of some ways to overcome these problems.

# Data Cleaning

## Data Loading & Data Cleaning

Step 1: Loading the request data (request\_data.csv)

- a) The request data was loaded in Data Frames **request\_data**.
- b) The total count of records in request\_data data frame was 6745

Step 2: Cleaning the Data

- a) The naming of fields was standardised for ease of use in further code (converted column names to lowercase with underscore like request\_id)
- b) The drop timestamp and request timestamp fields were converted to appropriate datatype (datetime) to carry out calculations further
- c) Saw ~39% nulls in driver\_id code & ~58% in drop\_timestamp
- d) Rows with the above null values were not dropped from the data frame as these are genuine nulls that indicate that the requests were either Cancelled (Driver Id Present but drop\_timestamp null) or were not served due to cars not being available(Driver Id & Drop Timestamp both null)

# Deriving new Attributes

Deriving new attributes from Request Timestamp

a) Various new attributes were derived from the request timestamp viz. Day(Day of the month like 1, 5, 10 etc.), Month (like January, February etc.), Year (like 2019, 2016 etc.), Weekday(like Monday, Wednesday etc.), Hour of the day (like 6, 12, 18 etc.). Some of these attributes will help in deriving some metrics in later analysis.

b) Based on hour of the day, another attribute `time_slot` was derived that bucketed the hour of the day into 6 broad categories

- 1) between 4 AM and before 8 AM (Early Morning)
- 2) between 8 AM and before 12 PM (Morning)
- 3) between 12 PM and before 4 PM (Afternoon)
- 4) between 4 PM and before 8 PM (Evening)
- 5) between 8 PM and before 12 AM (Late Evening)
- 6) between 12 AM and before 4 AM (Late Night)

# Analysis Activities

Check the overall number of requests by categories:

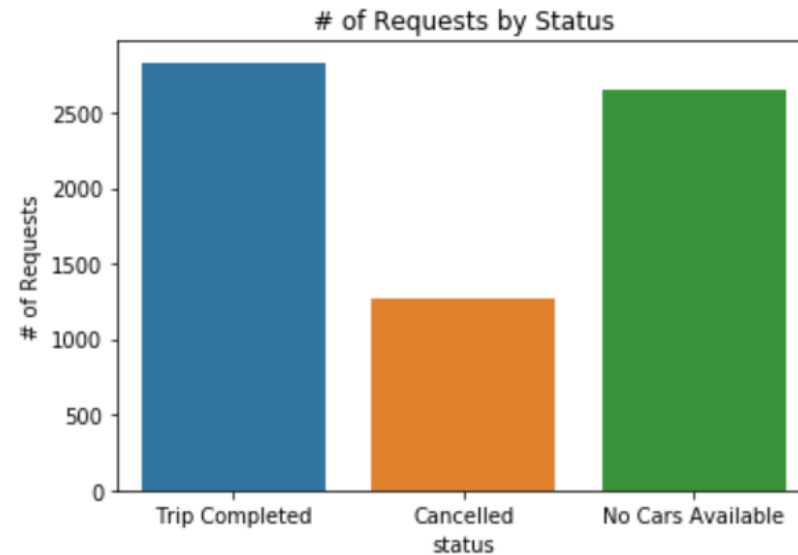
- This is just a summary of requests and will be analysed one by one to reveal useful insights in our analysis

	pickup_point	status	time_slot	#_of_requests
0	Airport	Cancelled	Early Morning	10
1	Airport	Cancelled	Morning	39
2	Airport	Cancelled	Afternoon & Evening	74
3	Airport	Cancelled	Late Evening	75
4	Airport	No Cars Available	Early Morning	125
5	Airport	No Cars Available	Morning	51
6	Airport	No Cars Available	Afternoon & Evening	574
7	Airport	No Cars Available	Late Evening	933
8	Airport	Trip Completed	Early Morning	246
9	Airport	Trip Completed	Morning	381
10	Airport	Trip Completed	Afternoon & Evening	317
11	Airport	Trip Completed	Late Evening	360
12	City	Cancelled	Early Morning	373
13	City	Cancelled	Morning	579
14	City	Cancelled	Afternoon & Evening	57
15	City	Cancelled	Late Evening	54
16	City	No Cars Available	Early Morning	288
17	City	No Cars Available	Morning	335
18	City	No Cars Available	Afternoon & Evening	164
19	City	No Cars Available	Late Evening	124
20	City	Trip Completed	Early Morning	280
21	City	Trip Completed	Morning	473
22	City	Trip Completed	Afternoon & Evening	368
23	City	Trip Completed	Late Evening	366

# Analysis Activities (Cntd..)

## 1) Analysing Requests with every Status (Completed, Cancelled, No Cars Available)

Grouped the request\_data based on the request Status and observed the following:



As we can see from above that more than 50% of the requests ( $1264 + 2650 = 3914$ ) were not fulfilled as the trip was either cancelled or the cars were unavailable. So there is definitely going to be some discontent among the ride seekers, so we need to analyze further and go to more details.

# Analysis Activities (Cntd..)

## 2) Analyze the average trip time of rides Completed

Found out the average time of completed trips using the difference of drop timestamp and request timestamps. This was done in order to see if the average time of a trip between City and Airport was too huge that lead to Drivers not accepting the requests.

	pickup_point	time_slot	avg_trip_time_in_seconds
0	Airport	Early Morning	00:52:21.617886
1	Airport	Morning	00:52:09.430446
2	Airport	Afternoon & Evening	00:51:49.791798
3	Airport	Late Evening	00:52:32.838889
4	City	Early Morning	00:53:46.742857
5	City	Morning	00:52:47.752643
6	City	Afternoon & Evening	00:51:36.769022
7	City	Late Evening	00:52:12.729508

The average trip time for any trip is ~50 minutes which is not so fluctuating. So this does not seem to be influencing the denial of requests

Lets dig deeper into analyzing the Status of these requests from different pickup points

# Analysis Activities (Cntd..)

## 3) Analyze requests by pickup point and request status

Considered all the requests and grouped them by pickup point and the trip status to see what type of requests are higher and what is its status



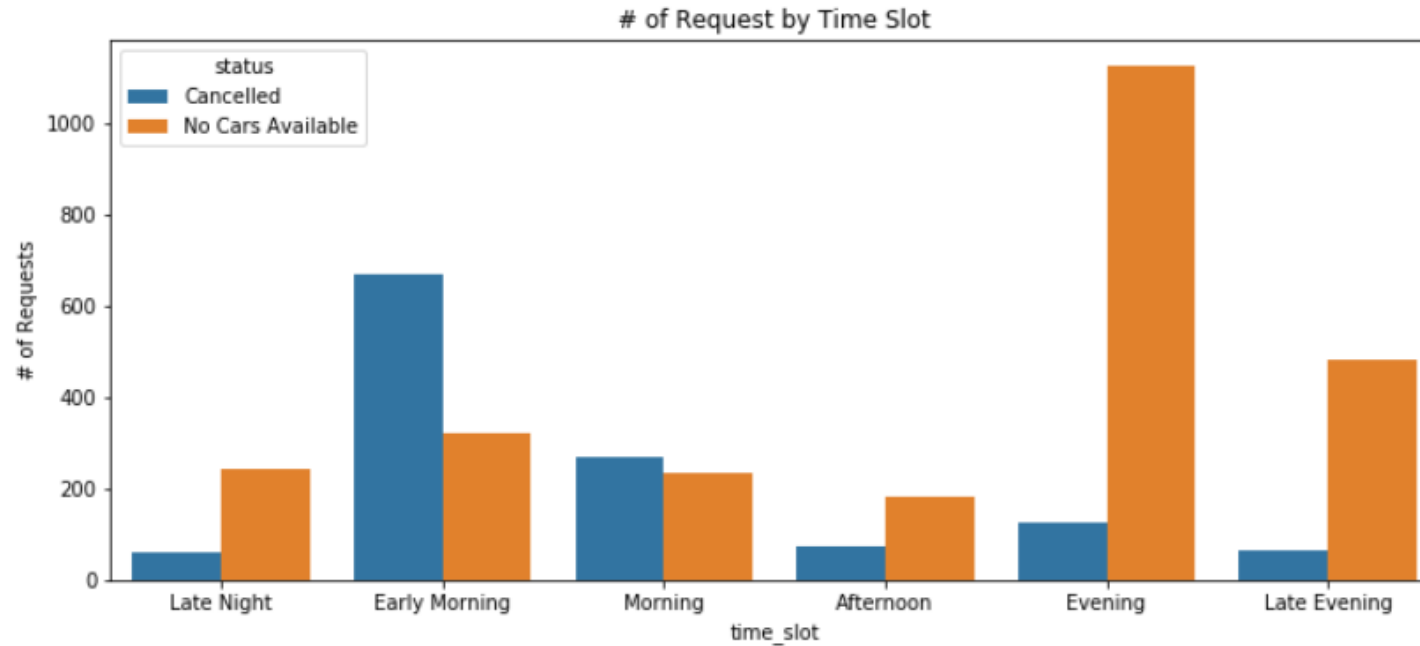
a) So as we see from above, most of the requests at the Airport that go unfulfilled are due to unavailability of the cars while a few get cancelled

b) Out of all unfulfilled requests in City area, most of them are cancelled that result in unavailability of Cars at the Airport. Quite a considerable count of requests also go unfulfilled due to No Cars Available



# Analysis Activities (Cntd..)

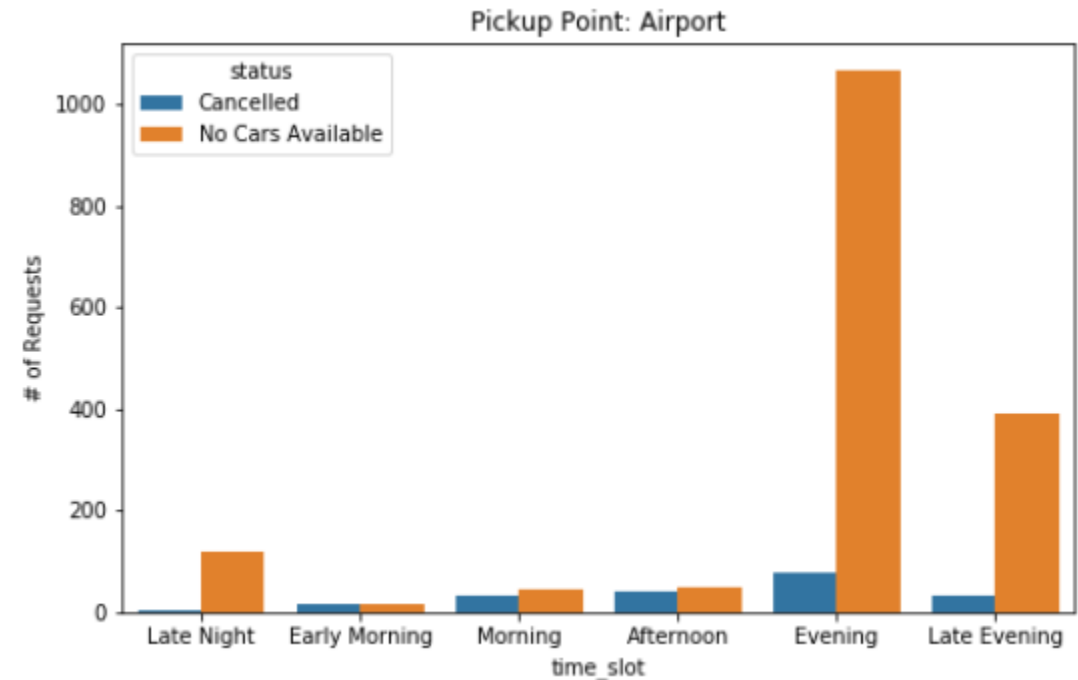
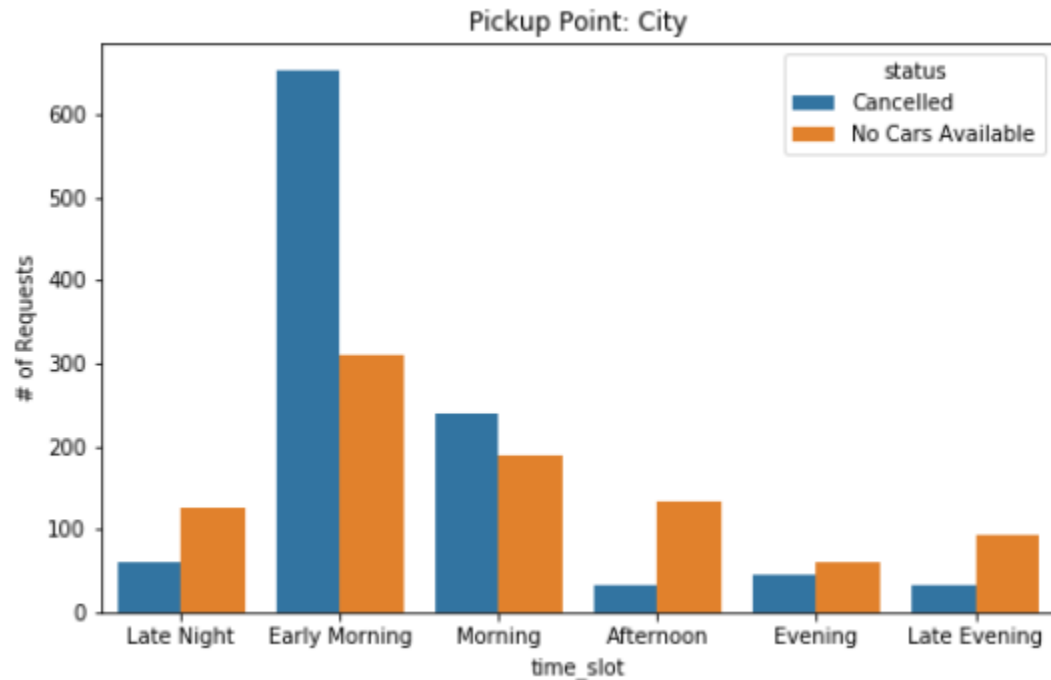
## 4) Analyze the usual hours of the day when the requests are cancelled , or show No Cars available



As we see in the above figure, the number of cars that are available reduces in evening and Late evening. This is quite a huge supply-demand gap. Lets check one more aspect of the nature of these requests: Availability of Cars/Status of trips during various time slots at different pick up points

# Analysis Activities (Cntd..)

## 5) Analyze the availability of Cars/Status of trips during various time slots at different pick up points



- As we see above, a large number of requests from City to Airport get cancelled during the Early Morning & Morning (4 AM to 12 PM).
- The above chart also tells us that the maximum number of trip requests at the Airport are seen in the Evenings & Late Evenings (4 PM to 12 AM) probably because a lot of flights land at that time. So the Demand is way higher than the Supply of cars at the Airport at this time.

# Analysis Activities (Cntd..)

## Supply-Demand Gap Analysis

- **Demand:** Demand is the Total number of requests made by ride seekers
- **Supply:** Supply is the Total number of Cars made available to fulfill the requests (Status = Trip Completed)
- **Gap** = Demand – Supply

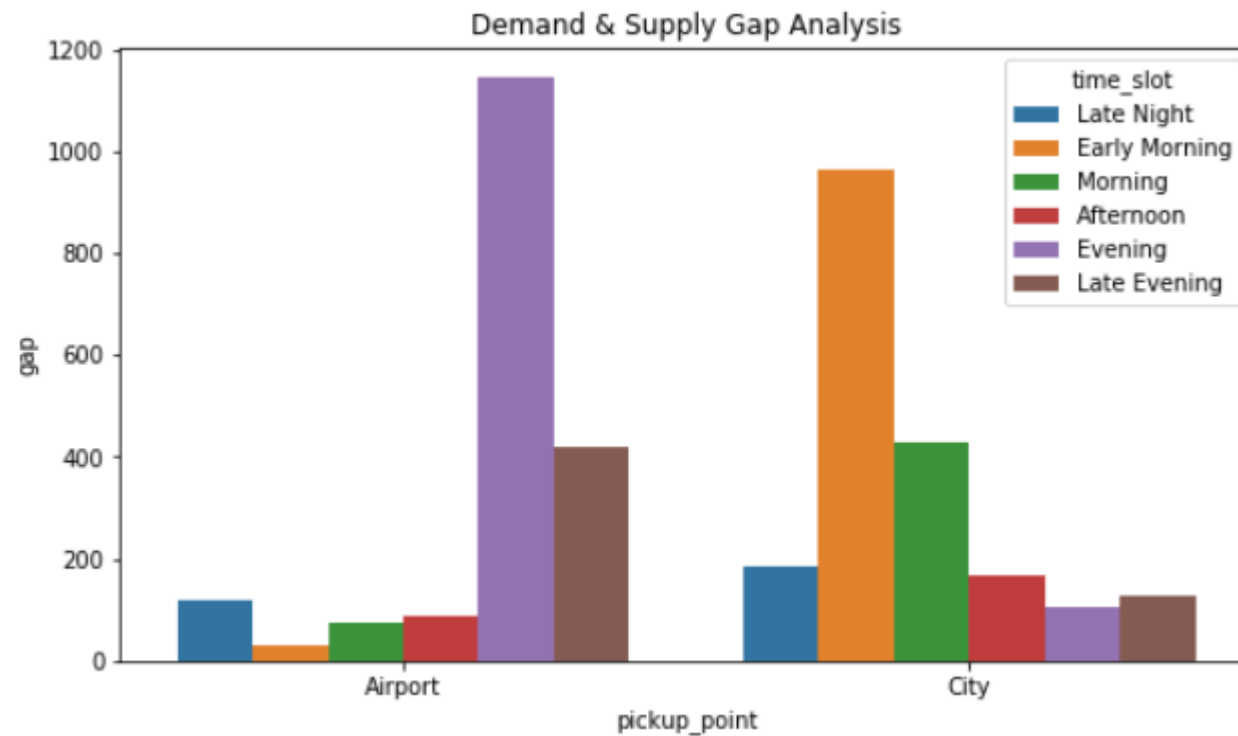
Lets see what was the Total Demand and what was the total Supply against the demand for the data available

	pickup_point	time_slot	demand	supply	gap
4	Airport	Evening	1457	312	1145
5	Airport	Late Evening	624	203	421
0	Airport	Late Night	200	80	120
3	Airport	Afternoon	252	162	90
2	Airport	Morning	315	239	76
1	Airport	Early Morning	337	308	29
7	City	Early Morning	1335	373	962
8	City	Morning	714	286	428
6	City	Late Night	279	94	185
9	City	Afternoon	374	208	166
11	City	Late Evening	323	196	127
10	City	Evening	436	330	106

# Analysis Activities (Cntd..)

## Supply-Demand Gap Analysis (Contd..)

Lets visualize the numbers we saw before graphically so that we can make more quick inferences



# Analysis Activities (Cntd..)

## Supply-Demand Analysis (Contd..)

- High Gap exists during Evening and Late Evening at the Airport
- High Gap exists during Early Morning & Morning in the City

The pattern is almost opposite at both pickup points.

- When the gap is high at the Airport it is minimal in the City and the vice-versa

# Reasons for the Supply-Demand Gap

- For trips starting in the Airport Area, the huge Supply Demand gap in Evenings and late night appears to be happening due to less number of cars being available. The probable reasons for these could be
  - Drivers going back home ending their day and hence not accepting more rides
  - Though many flights might be landing at this time resulting in high Demand, perhaps not many flights take off during these hours so the number of cars arriving at the Airport during this time is very less. This results in causing many requests going unserved as enough cars are not available to pick these requests
  
- For trips originating in City area, the huge Supply Demand gap in Early Morning and morning time could probably be due to following reasons:
  - The number of requests during these hours from Airport to City is relatively low. This results in many Drivers cancelling the trip from City to Airport during this time as they know that probability of them getting a return trip during this time is low and they would end up waiting for long to get a trip back. The reason is perhaps that many flights take off during these hours but not significant number of flight land during this time. So less passengers taking rides towards City.
  - Not many Drivers would have started their day during early morning
  - The Drivers could be probably getting more lucrative rides in terms of rides headed towards another location where the returning rides are also more and may be less travel time

# Ways to Improve Supply Demand Gap

- Give more incentives (or surges in trip amount) to Drivers accepting rides during the hours where huge Supply-Demand gap is observed in City & Airport
- More Incentives for Driver partners (if exists or may create partners) if they provide higher availability of cars to serve the requests during the time when high Supply-Demand gap exists