

# **UBER SUPPLY-DEMAND GAP CASE STUDY SUBMISSION**

**Name : Ranip Hore**

**Cohort : 30<sup>th</sup> Sep , 2018**

# INTRODUCTION

## **Business Understanding :**

Uber Technologies Inc. face a loss of potential revenue due to supply shortages either caused by driver induced request cancellations or non-availability of cars.

## **Business Objectives**

- The aim of analysis is to identify the root cause of the problem (i.e. cancellation and non-availability of cars) and recommend ways to improve the situation. As a result of your analysis, you should be able to present to the client the root cause(s) and possible hypotheses of the problem(s) and recommend ways to improve them.

## **Data Understanding:**

- Request id: A unique identifier of the request
- Time of request: The date and time at which the customer made the trip request
- Drop-off time: The drop-off date and time, in case the trip was completed
- Pick-up point: The point from which the request was made
- Driver id: The unique identification number of the driver
- Status of the request: The final status of the trip, that can be either completed, cancelled by the driver or no cars available

We imported the dataset which had some data quality issues. The columns Request timestamp and Drop timestamp were having date and timestamp information in non-standard form. So we converted it into standard date-time form and then proceeded for analysis.

## Snapshot of dataset before cleansing :

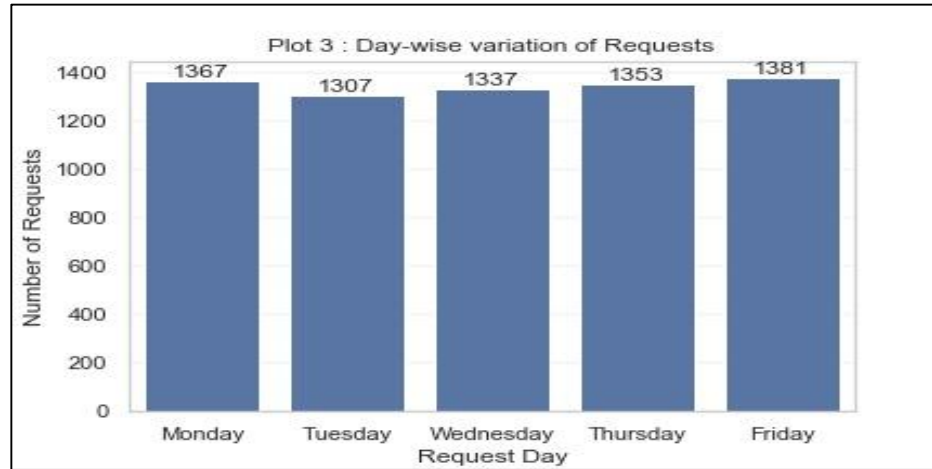
	Request id	Pickup point	Driver id	Status	Request timestamp	Drop timestamp
0	619	Airport	1.0	Trip Completed	11/7/2016 11:51	11/7/2016 13:00
1	867	Airport	1.0	Trip Completed	11/7/2016 17:57	11/7/2016 18:47
2	1807	City	1.0	Trip Completed	12/7/2016 9:17	12/7/2016 9:58
3	2532	Airport	1.0	Trip Completed	12/7/2016 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

## Snapshot of dataset after cleansing :

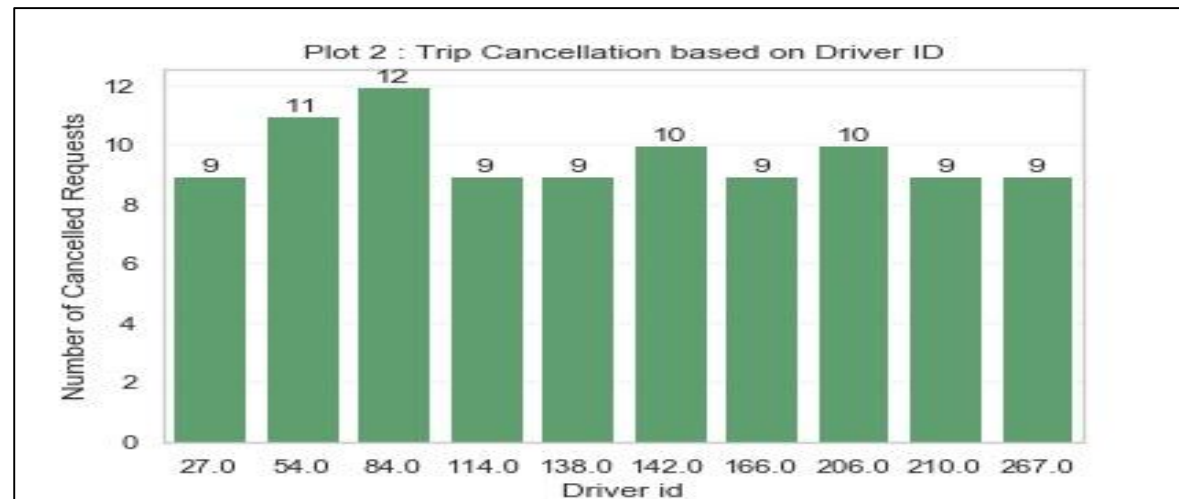
	Request id	Pickup point	Driver id	Status	Request timestamp	Drop timestamp
0	619	Airport	1.0	Trip Completed	2016-07-11 11:51:00	2016-07-11 13:00:00
1	867	Airport	1.0	Trip Completed	2016-07-11 17:57:00	2016-07-11 18:47:00
2	1807	City	1.0	Trip Completed	2016-07-12 09:17:00	2016-07-12 09:58:00
3	2532	Airport	1.0	Trip Completed	2016-07-12 21:08:00	2016-07-12 22:03:00
4	3112	City	1.0	Trip Completed	2016-07-13 08:33:16	2016-07-13 09:25:47

# ASSUMPTIONS AND REASONING

**Day-Wise Variation** : There is minimal variation in the number of requests generated with respect to the days of the week. Refer below plot.



**Driver Cancellation Trends** : From the dataset the driver id's having high number of cancellation trends for airport-city and city-airport trips has been shown in the below figure :



**Driver Idle Time Calculation** : As this dataset contains information only with respect to trips made to and from the airport, and the discrepancy with respect to registering information for certain trips [For example for numerous Driver ID's the city to airport is registered on the 11th and the subsequent trip is again registered as city to airport on 12th ] , so considering the scope of our analysis, driver idle time at the airport is not calculated as calculating idle time with missing information will generate misleading insights.

**Time-Slots** : We have grouped the whole day into different time-slots based on the below logic to study the supply-demand gap across various times of the day.

5 AM – 10 AM : Morning Peak Hour , 10 AM – 2 PM : Noon , 2 PM – 5 PM : Afternoon , 5 PM – 9 PM : Evening Rush , 9 PM – 12 AM : Night,  
12 AM – 5 AM : Midnight

**Request Hour** : For the scope of our analysis we have extracted the hour from the request timestamp to check when the request was initiated by the customer. For instance the any requests generated between 10:00:00 and 10:59:59 will be assigned the request hour attribute as 10. This will facilitate grouping and analysis.

**Demand Definition** : Demand is defined as the total number of requests received within a given period.

**Supply Definition** : Supply is defined as the total number of requests for which trip status is completed.

**Gap Definition** : Gap is defined as the number of requests for which there was no cars available or the driver has cancelled the requests. In terms of expression it can be written as : **GAP = DEMAND - SUPPLY**

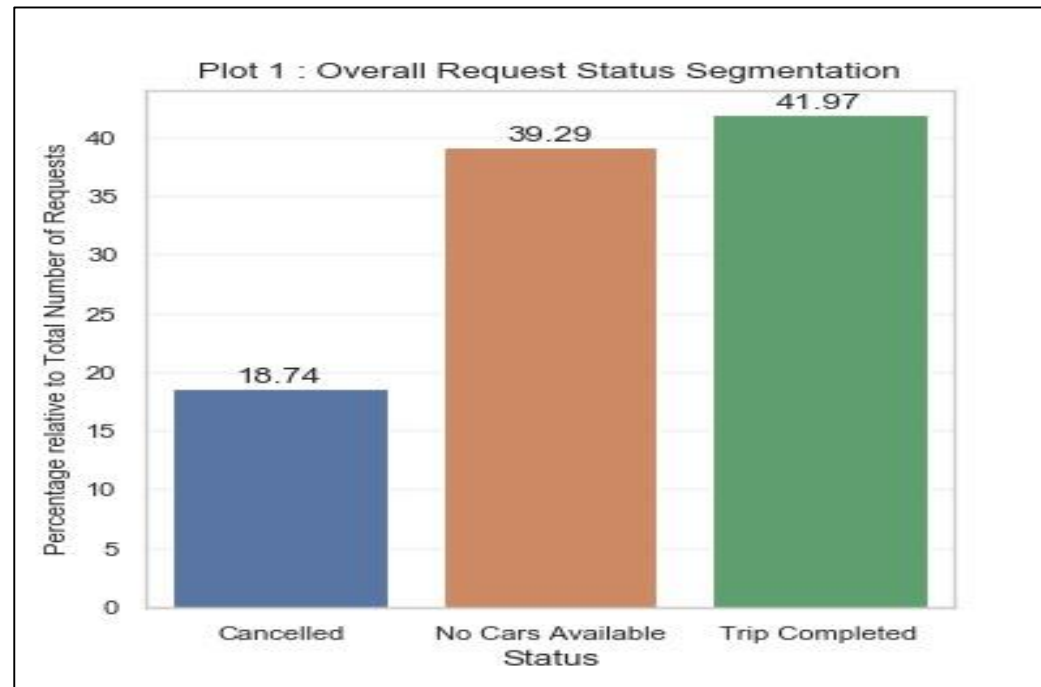
## 1. Identify the pressing problems :

### Analysis of overall request status :

Plot 1. shows the overall frequency of requests status provided in the dataset. It can be clearly seen that 41.97% of the requests has been served and rest has been denied due to cancellation or unavailability of cars.

**Customer Service Rate : 41.97%**

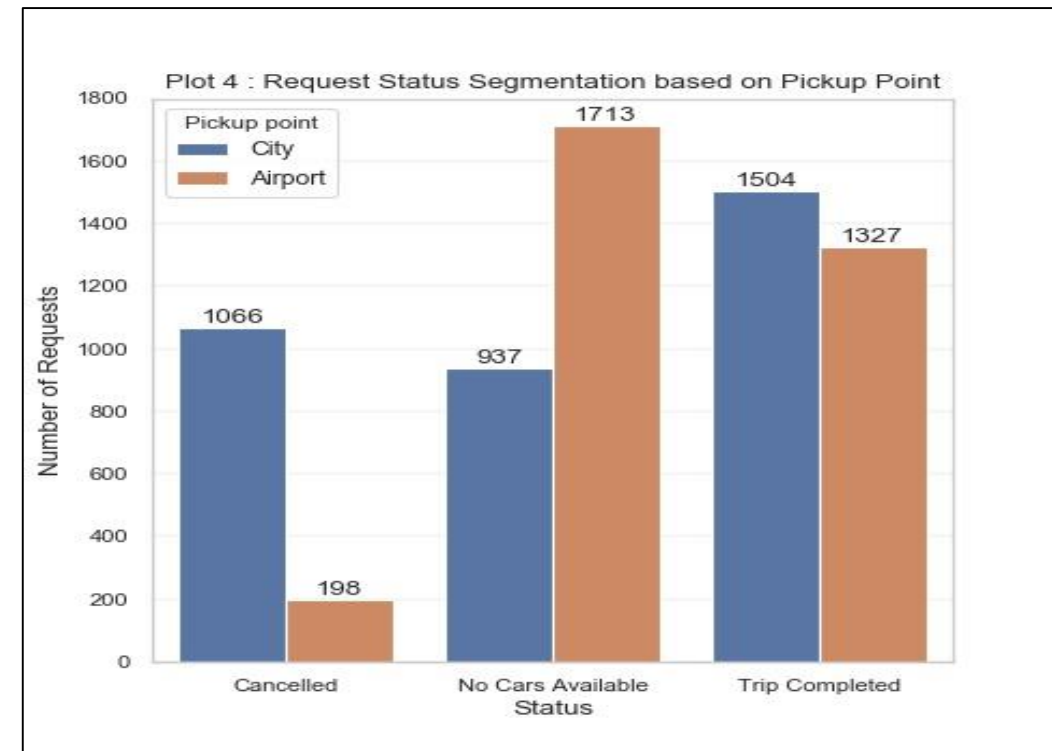
**Customer Denial Rate : 58.03%**



# DATA ANALYSIS [Contd.]

## a. Identifying Most Problematic Type of Requests [City to Airport / Airport to City] :

Plot 5. shows that a higher number of requests are denied with customer pickup location as city than that of airport. Therefore a higher percentage of customers booking Uber from airport are denied service. On further drill down we found that the high denial from airport to city trips can be attributed to unavailability of cars at airport.[Plot 4.]

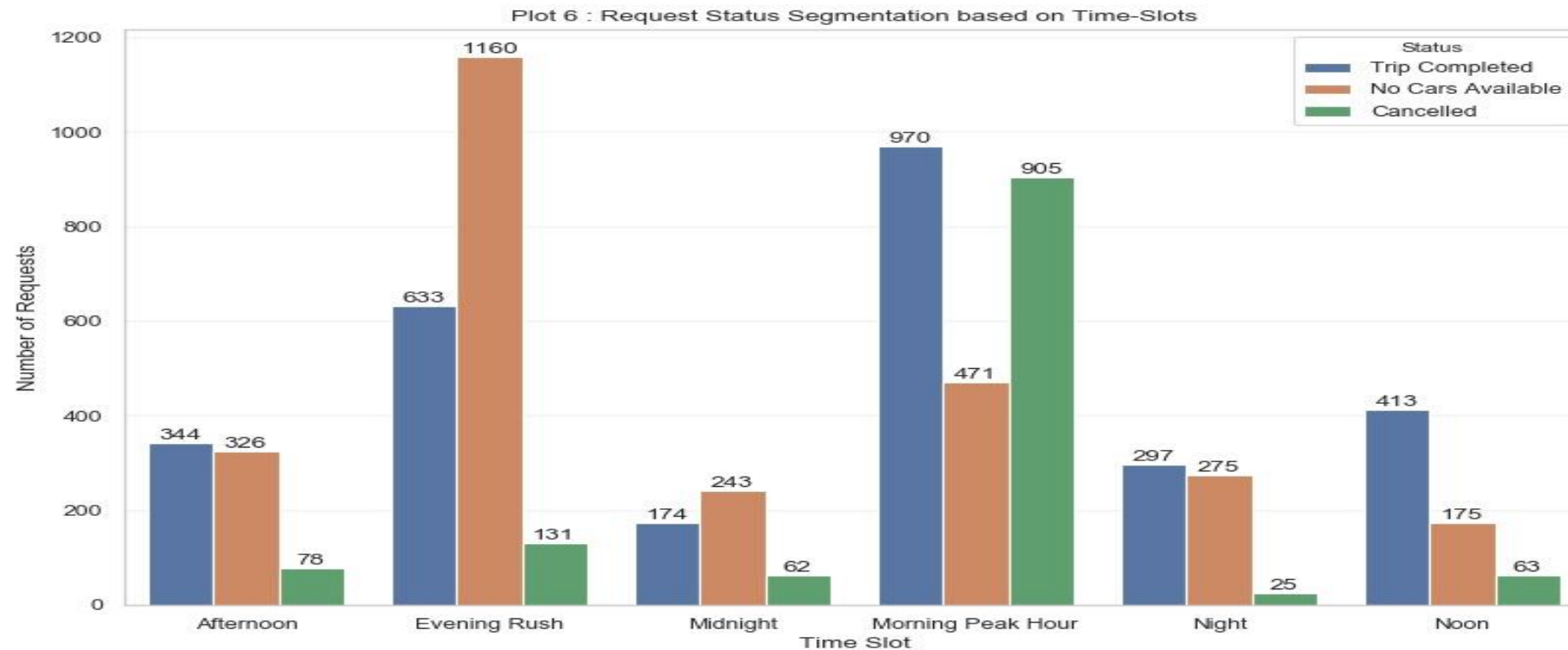


Hence we can conclude that customer requests with pickup location as Airport are more problematic and more likely to be denied service.

# DATA ANALYSIS [Contd.]

## a. Identifying Most Problematic Time Slot :

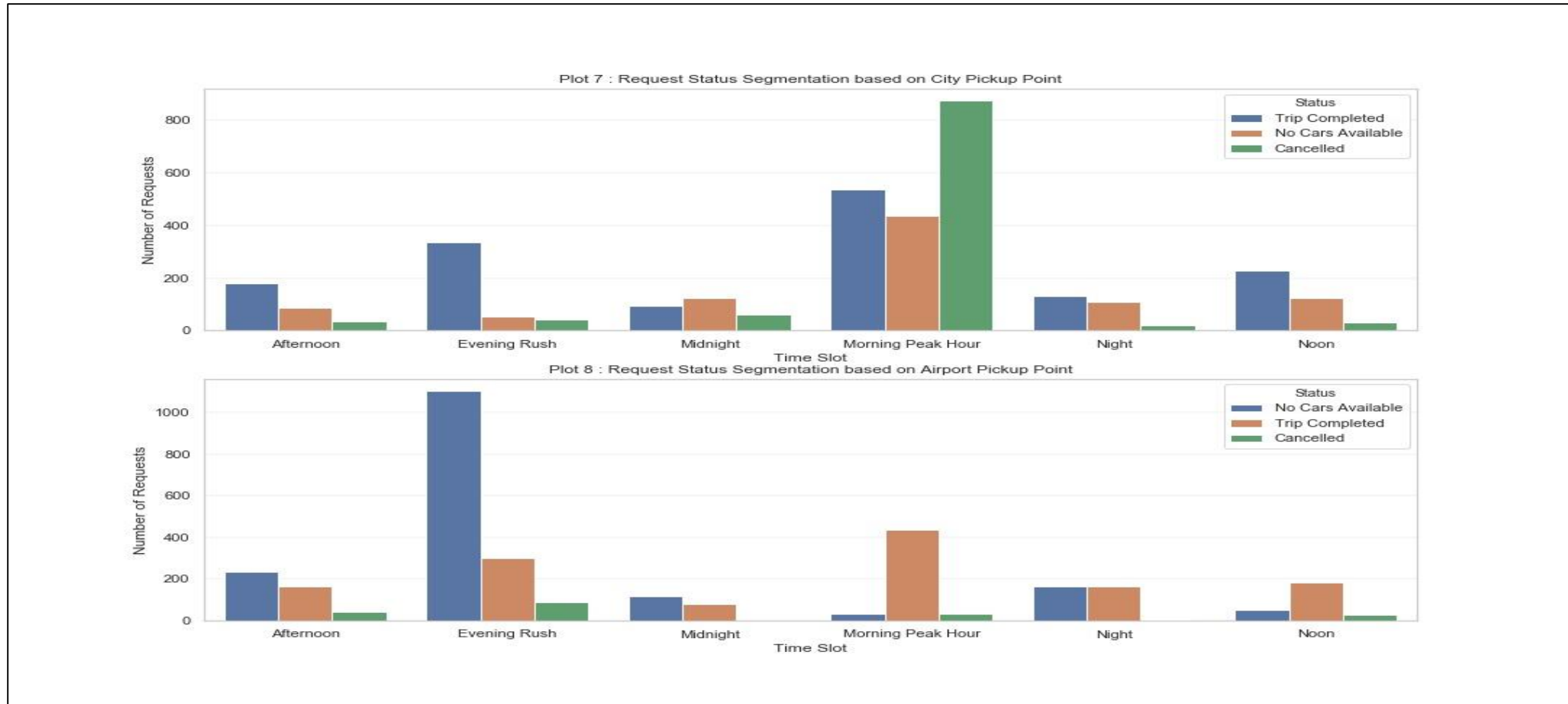
Plot 6. clearly shows that “Morning Peak Hour” and “Evening Rush” are the two most problematic segments contributing to the highest gap in supply and demand.





# DATA ANALYSIS [Contd.]

On further inspection of the problematic time slots based on pickup point we arrived at the below plots for city pickup and airport pickup:

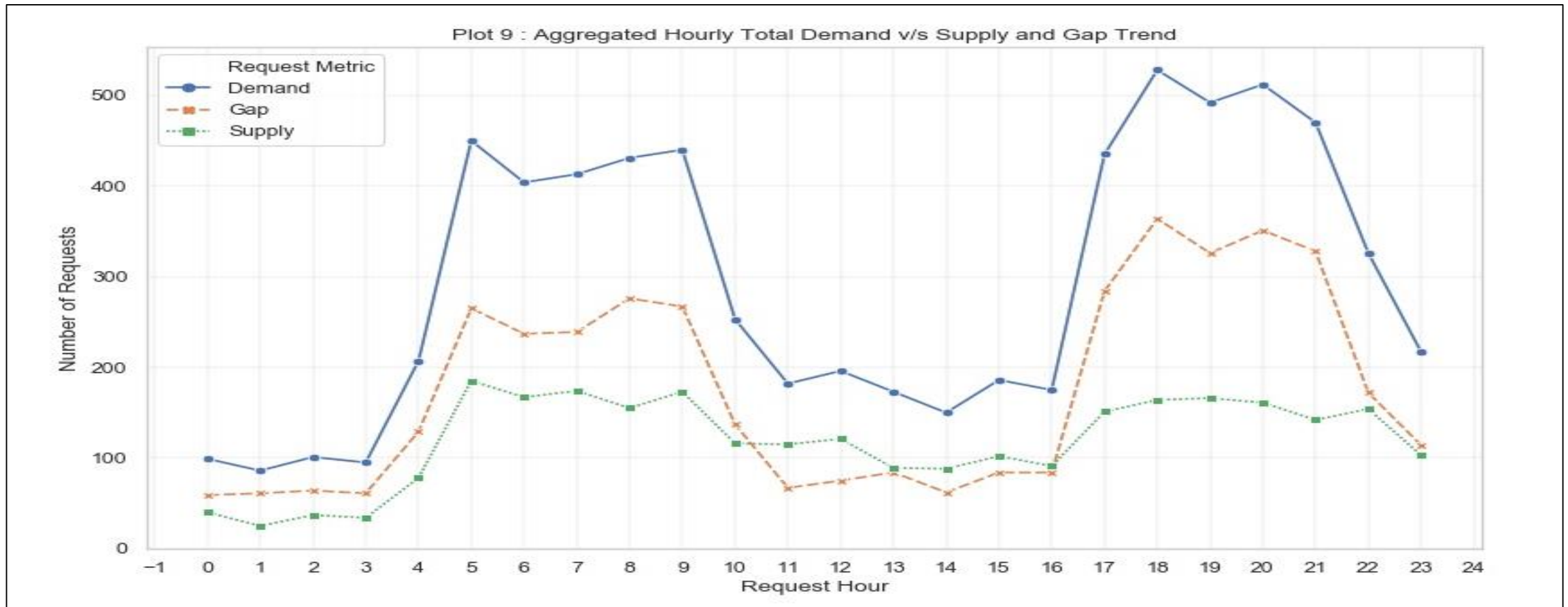


- From the above plots it can be seen that during **morning peak hour** , a high number of requests are being cancelled by driver. This mainly impacts a customer booking cab from City to Airport.
- During **evening rush** a large number of customers are denied service due to lack of organic flow i.e. unavailability of cars. This mainly impacts a customer booking cab from Airport to City.

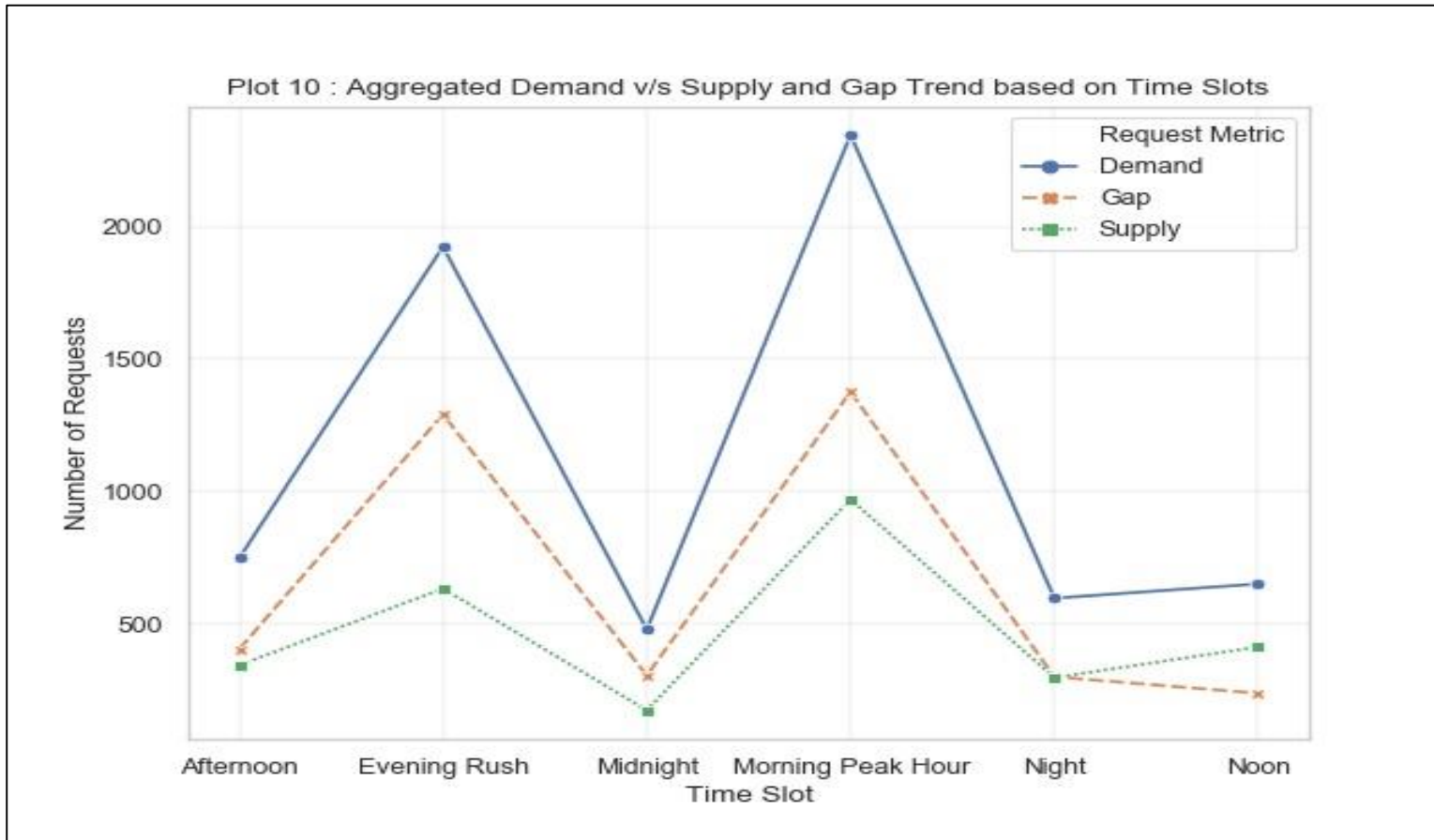
## 2. Supply – Demand Gap :

### a. Time Slot When Highest Gap Exists :

Plot 9. depicts the hourly trend of **demand vs. supply** and the gap variation throughout the day. It is clearly evident from the plot that the highest gap exists from 5 AM to 9 AM(morning peak hour) and between 5 pm and 9 pm in the evening(evening rush).



Plot 10. depicts the **demand vs. supply** and the gap variation throughout various timeslots.



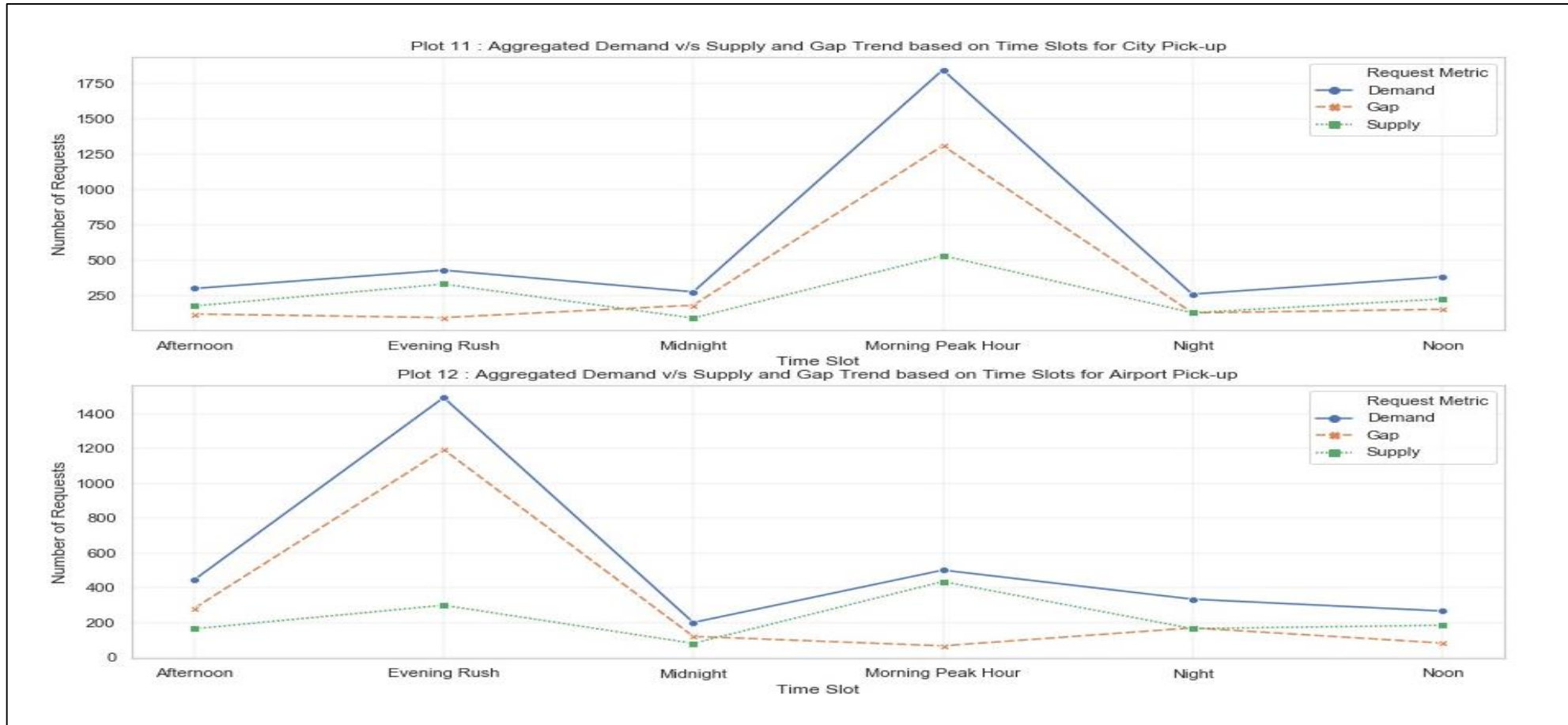
Time Slot	Demand	Supply	Gap
Afternoon	748	344	404
Evening Rush	1924	633	1291
Midnight	479	174	305
Morning Peak Hour	2346	970	1376
Night	597	297	300
Noon	651	413	238

It can be observed that the timeslot with the highest gap are **Morning Peak Hour** and **Evening Rush**.

# DATA ANALYSIS [Contd.]

## b. Type of Requests(City-Airport/Airport-City) for which Highest Gap Exists :

From Plot 12 and 13 it is evident that the Demand vs. Supply gap for Morning Peak Hour is most severe for requests from City to Airport. The gap is due to large number of requests being cancelled as depicted in Plot 6 and 7.



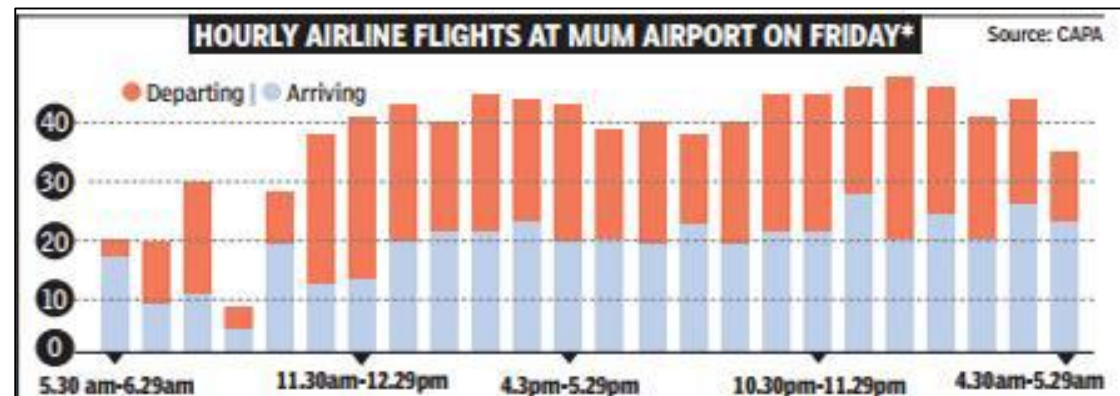
**Requests where Highest gap exists in Identified time-slots : City to Airport**

# REASON AND RECOMMENDATION

## Reason for Supply-Demand Gap :

Additional information has been sourced from a leading daily and attached below which shows the hourly number of flights arriving and departing in Mumbai Airport on a Friday.

- The image reveals an interesting trend that there are more outgoing flights than incoming flights in the early morning timeslot. For a driver who has to make a trip to airport from city has to inadvertently wait at airport for a longer period before getting the next ride back to city. Consequently increasing the driver idle time, which could have been utilised for other trips if he chose not to accept a city to airport ride in the morning hours. In some cases it might happen that the driver has to return to city without getting any booking and thus incurring a loss and waste of fuel. This might be the reason for high number of cancellation requests for city-airport trip in morning.
- Similarly it is also clear from the image that there are more number of incoming flights in the evening, which might be due to the high number of international flights landing during evening to night time. Since there is an increased demand and due to unavailability of cars to cater that demand, there is a huge surge in supply-demand gap during the evening rush hour.



## **Recommendation for Supply-Demand Gap :**

**Problem** : City-Airport Cancellation in Morning Hours

### **Recommendation** :

- Introduce surge pricing in the morning hour so that the drivers are encouraged to take up more city-airport trips.
- Proper training should be provided to drivers to make them aware of the no-refusal policy.
- Take appropriate actions against driver with high number of cancellations.
- Maintain transparency with respect to the surcharging of customers.

**Problem** : Airport-City Unavailability of Cars in Evening Hours

### **Recommendation** :

- Introduce ride-sharing option and also encourage customers to take shared rides. This will improve the number of customers served through less number of cabs.
- Introduce surge pricing for trips to airport so that more number of drivers are encouraged to go to airport during evening hours.
- Plan the distribution of cabs by analysing the data , so that areas with high demand in a specific timeslot can have more number of cabs.