SUPPLY-DEMAND GAP

SUBMITTED BY-

GOVIND SONI

Uber

❖ Problem statement:-

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 INSPECTION PART

Uber

- There are 6745 rows and 6 columns in this dataset.
- These columns are as follows:

- 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.

Uber

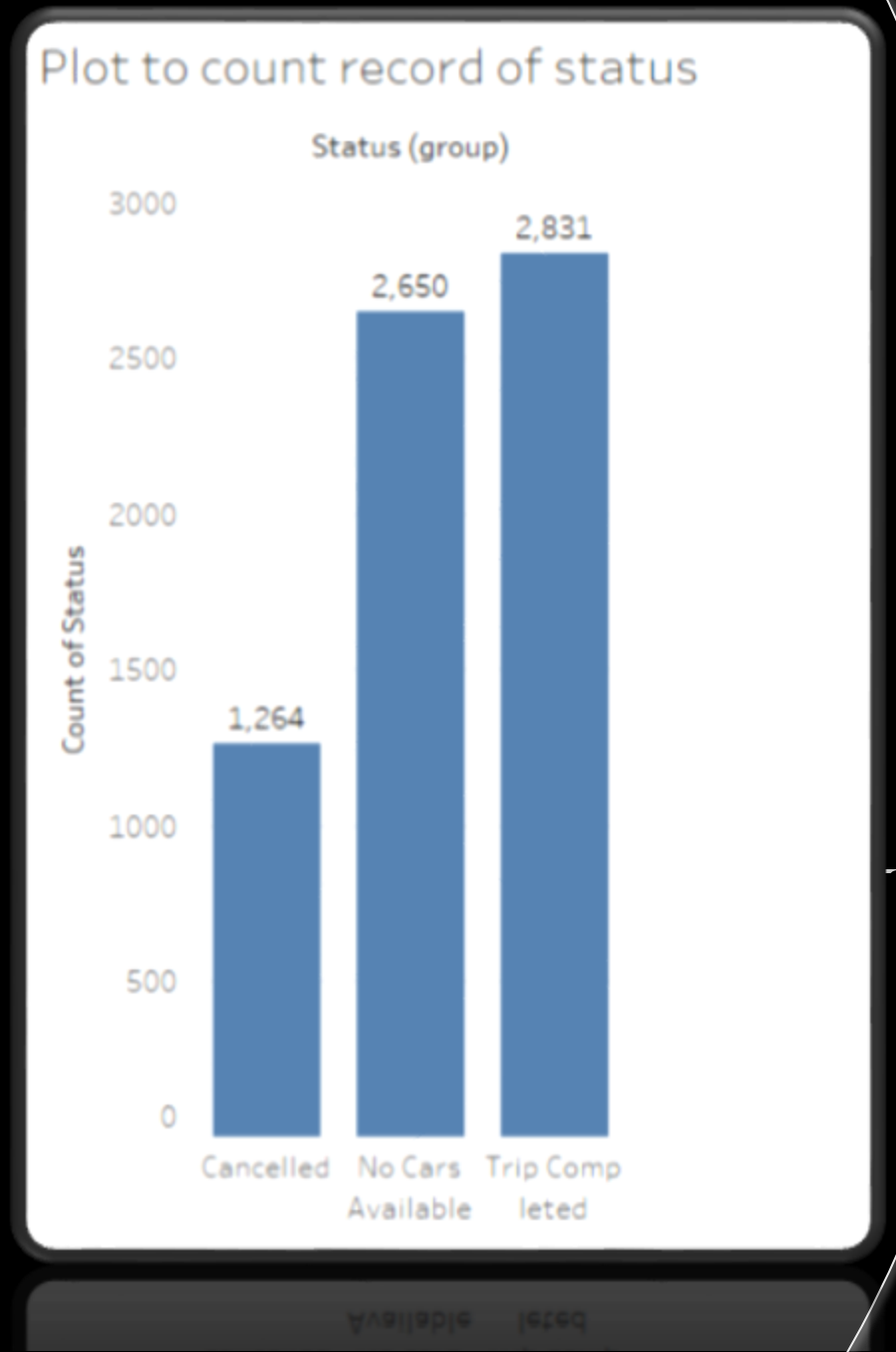
DATA MANIPULATION PART

- Column 'Request timestamp' and 'Drop timestamp' were changed from object data type to datetime64[ns] data type.
- Firstly, check for null values row-wise and it is found that few rows contains 0.03% of null values. So, these rows can be ignored.
- Secondly, check for null values column-wise and it is found that column 'Driver id' contain 39% null values and column 'Drop timestamp' contain 58% null values.
- Column 'Driver id' is dropped.
- Column 'Drop timestamp' contain null values due to trip cancelled by driver or no cars available at that time. This is important part of analysis that's why it is not dropped.

Uber

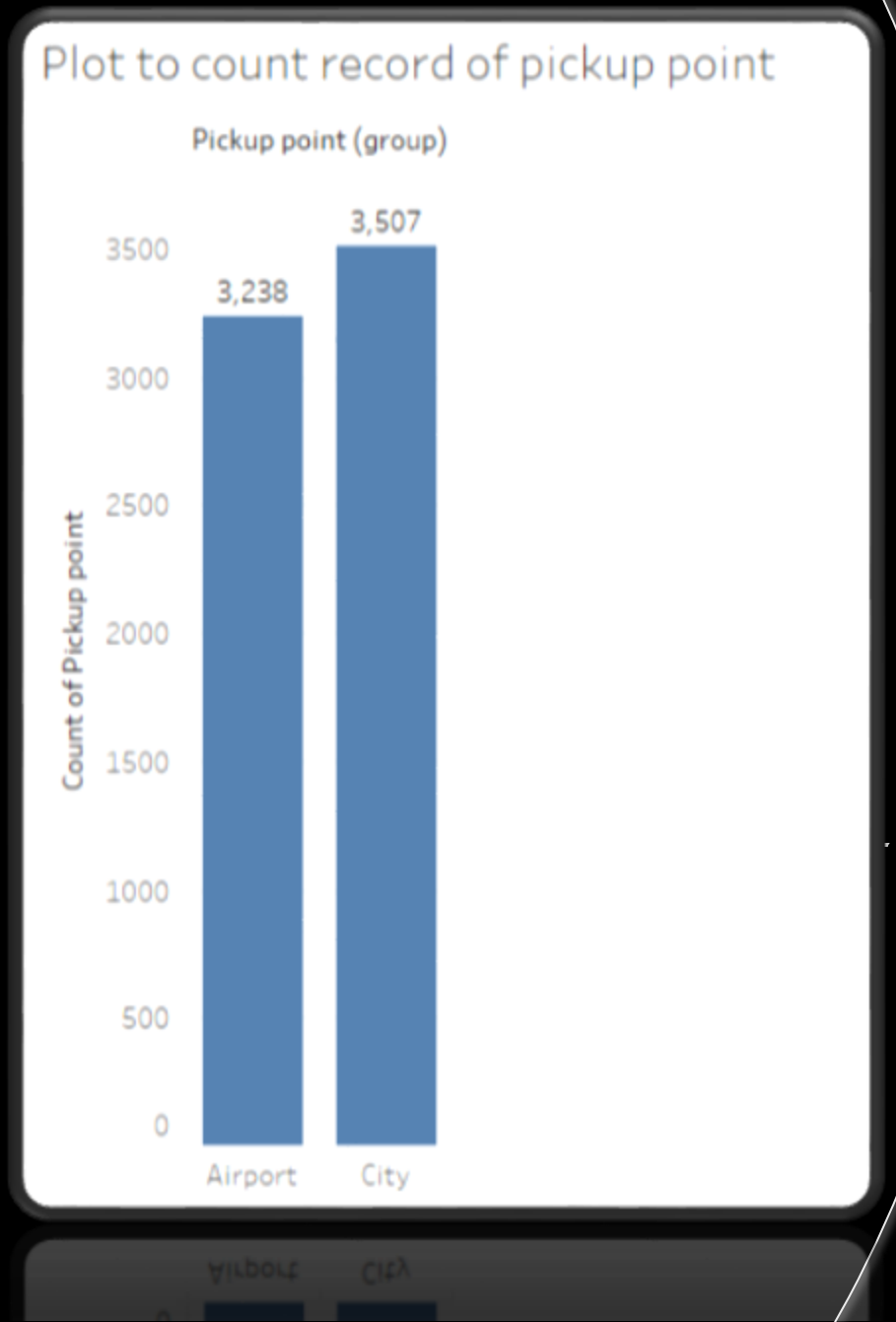
DATA ANALYSIS PART

- Analysis done using the exploratory data analysis technique.
- In this technique, univariate, segmented univariate and bivariate plots are formed to analyse the data.
- In this data set, univariate and bivariate analysis is used and some plots are formed from derived matrices technique.
- In derived matrices, request hours are extracted from request time and these hours are converted to their respective time slots.
- These time slots are plotted against 'Status' and 'Pickup Point' variables to get desired results.
- Few of the plots are attached here.



Uber

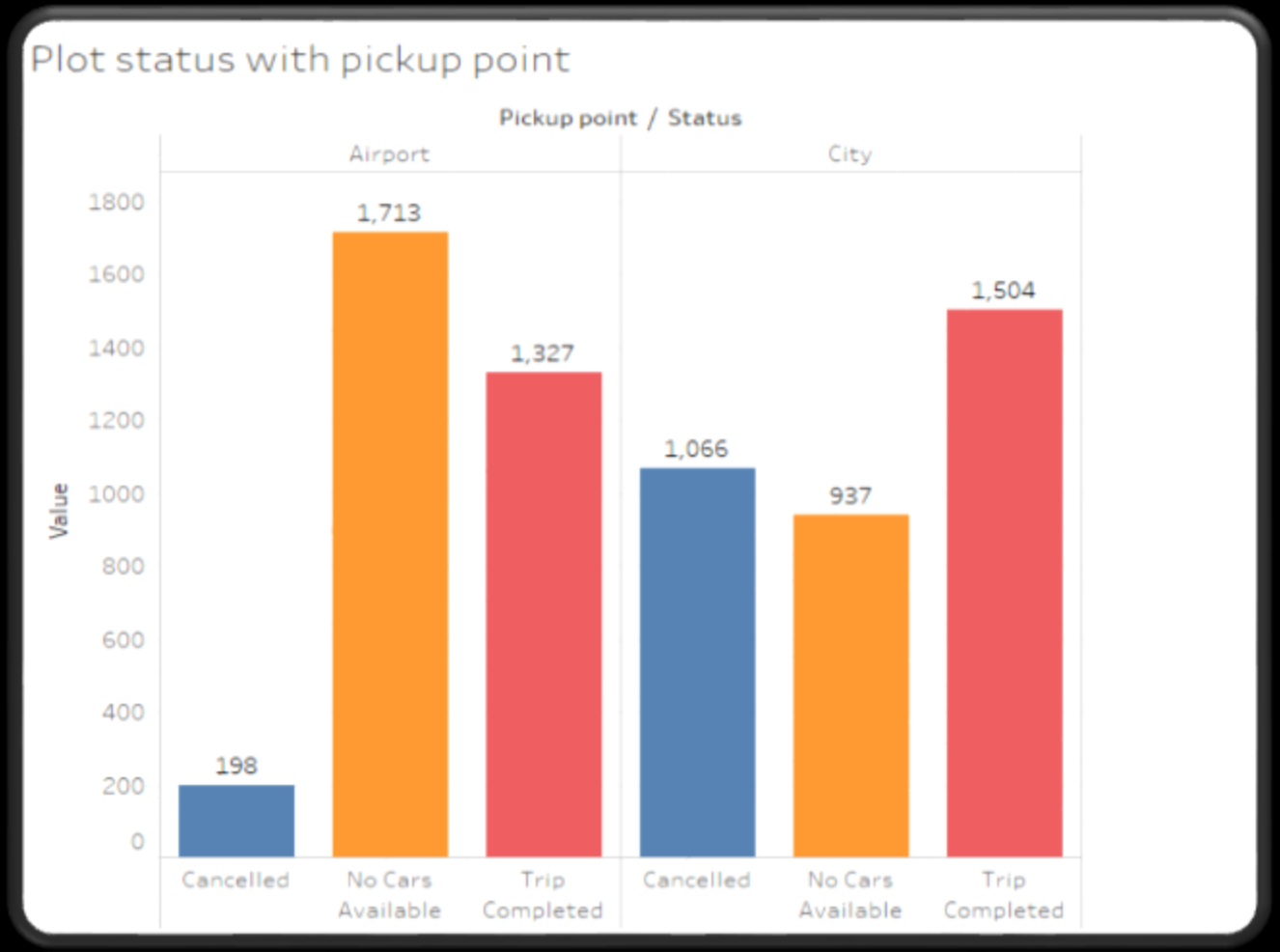
- **PLOT TO COUNT RECORD OF STATUS:-**
  - It is observed that no cars available is more than cancelled trips.
  - Status contain three categories:
    - Cancelled – 1264 records.
    - No cars available - 2650 records.
    - Trip completed - 2831records.



Uber

- **PLOT TO COUNT RECORD OF PICKUP POINT:-**
  - Count of pickup point in city is slightly higher than airport.
  - There are two pickup points in this data:
    - Airport – 3238 records.
    - City - 3507 records.

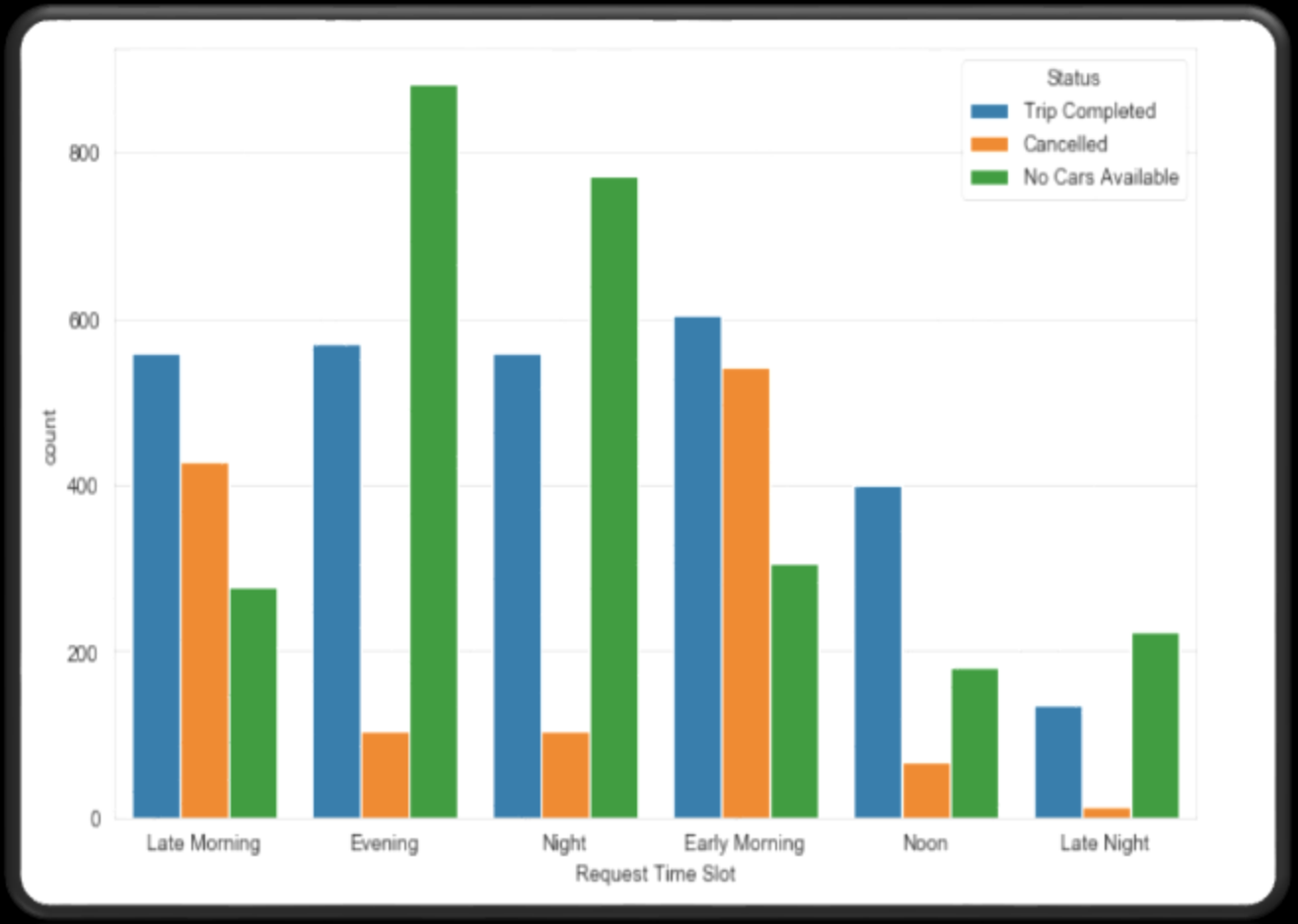
PLOT STATUS WITH PICKUP POINT:-



- No cars available is higher from airport to city.
- Most of the cancellation took place from city to airport.

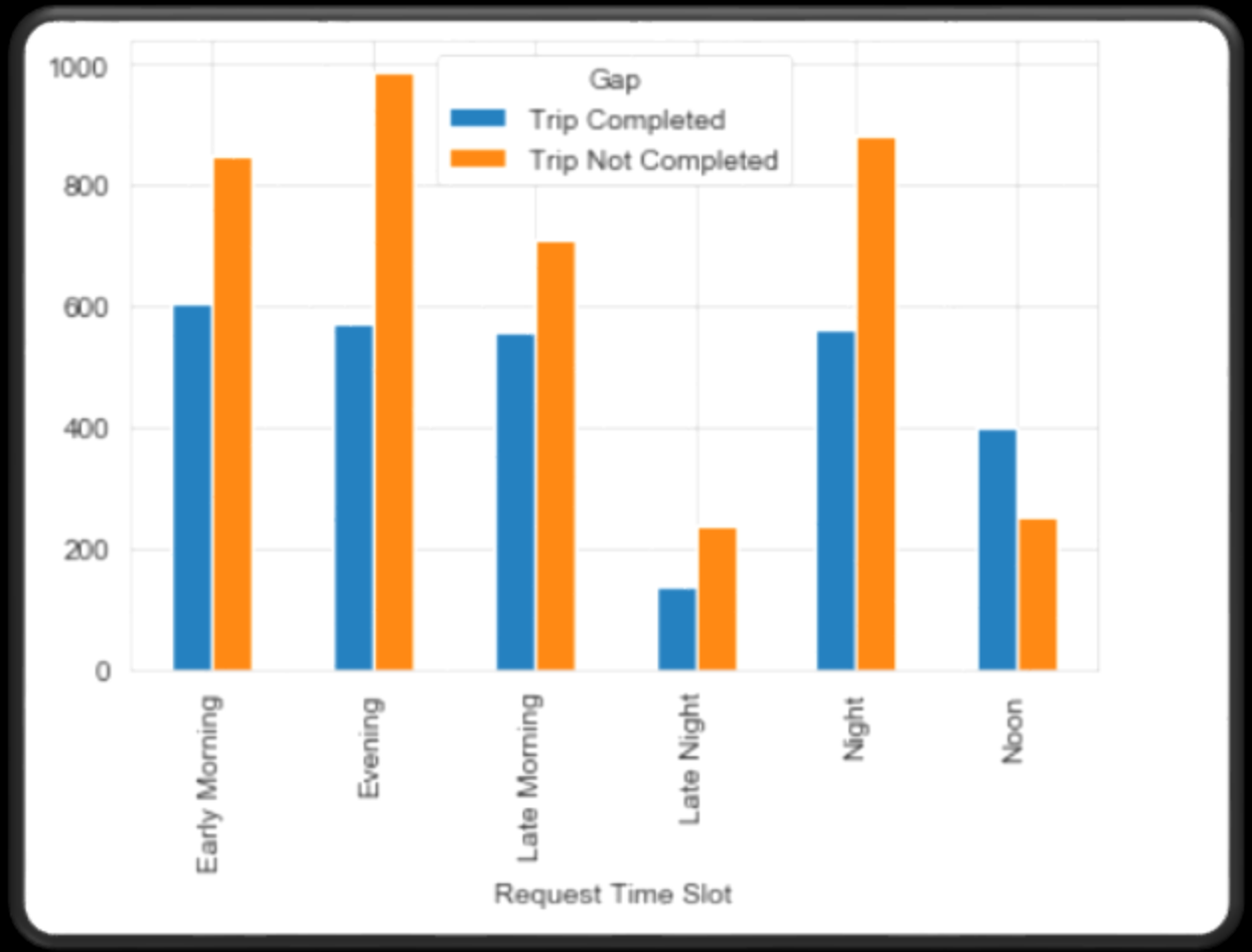
Uber

▪ PLOT STATUS WITH REQUEST TIME SLOT:-



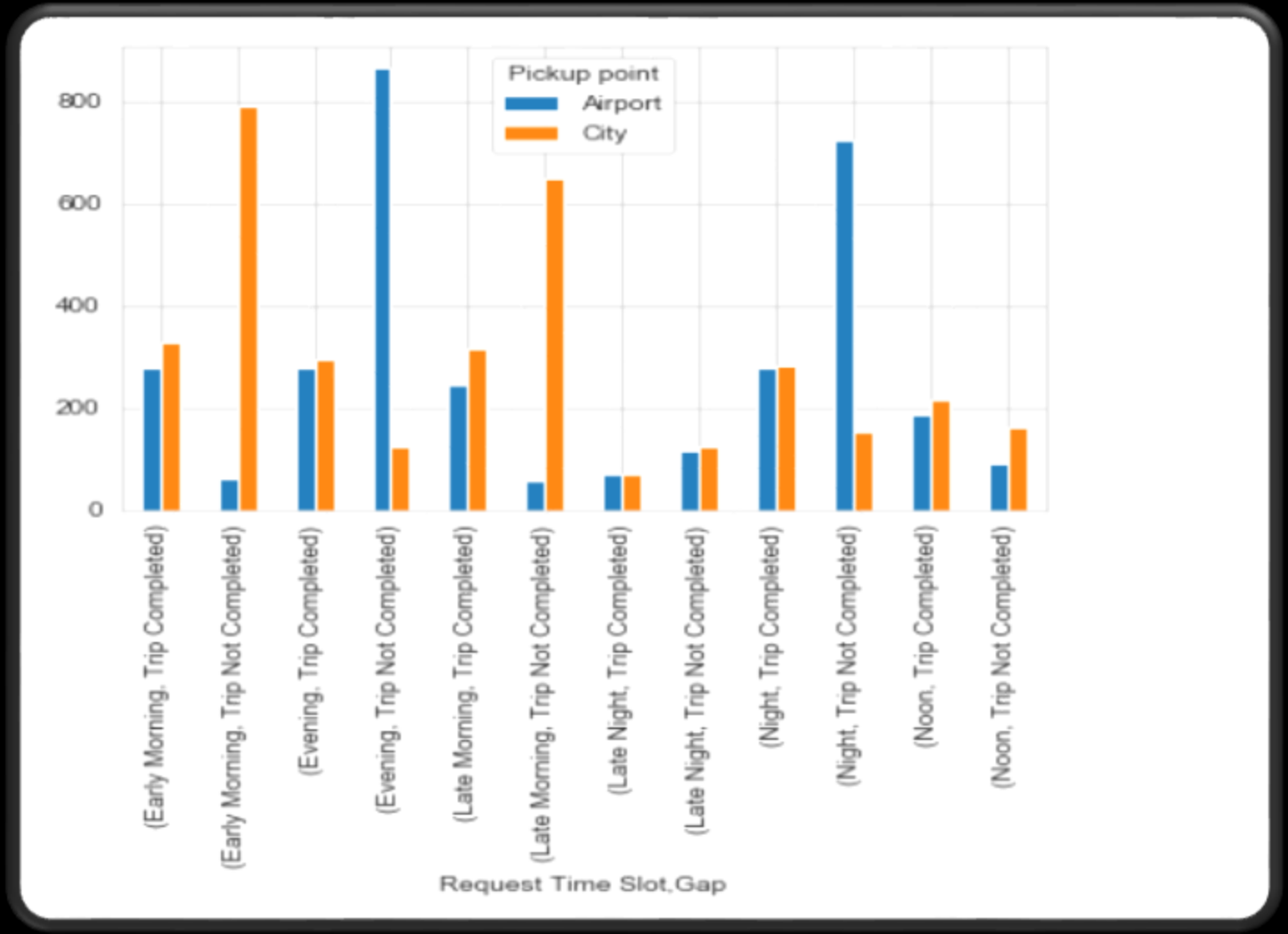
- Most of the cars were not available in evening and night time.
- Most of the cancellation took place in early morning and late morning.
- There is less cancellation in late night.

▪ PLOT GAP WITH REQUEST TIME SLOT:-



- Most of the trips were not completed in evening time.
- Less trips were completed during late night.
- In every case number of trips not completed is higher than trips completed except noon time.

▪ PLOT PICKUP POINT WITH RESPECT TO REQUEST TIME SLOT AND GAP:-



- Most of the trips were not completed from airport to city in evening time.
- In early morning, more trips were not completed from city to airport.

HYPOTHESIS

Uber

- **Pickup Point - City :-**

  As per the analysis, the morning time slot is most problematic where the requests are being cancelled. Most probably the requests are being cancelled by the drivers due to the morning rush as it being the office hours and seeing the destination as airport which would be too far, the driver would think to earn more for the shorter trips within the city.

- **Pickup Point - Airport :-**

  Upon analysis, the evening time slot seems to be most problematic for pickup points as airport where the requests being No Cars Available. The reason seems to be that not enough cars are available to service the requests as cars might not be available at the airport due to the cars serving inside the city.

CONCLUSION

Based on the data analysis performed, following recommendation can be used by Uber to bridge the gap between supply and demand:-

- For bridging the demand supply gap from airport to city, making a permanent stand in the airport itself where the cabs will be available at all times and the incomplete requests can come down significantly.
- Uber can provide some incentives to the driver who complete the trip from city to airport in the morning part. This might result the driver to not cancel the request from city to airport trips.
- Last but sure solution to bring down the gap is to increase the numbers of cab in its fleet.

Uber

THANKS