Uber supply-demand gap case study

SUBMISSION

Business understanding and objective

We are considering the request made to Uber to and from Airport to City for five days. Our main goal is identify the supply-demand gap. We are analyzing the provided data and try to find out the cause of the supply-demand gap. This case study is focused towards the analysis of supply and demand, we should be able to identify the root case and hypothesis recommendations to resolve the gap.

The objective is classified in the following sub-goals:

➤ Data cleaning , formatting and driving the metrics from existing columns
➤ Assumptions and Reasoning
➤ Data Analysis and reporting

Problem Solving Methodology

The approach for this project has been to divide the entire case study into various checkpoints to meet each of the sub-goals and give more meaning to the analysis.

The checkpoints are represented in a sequential flow as below:
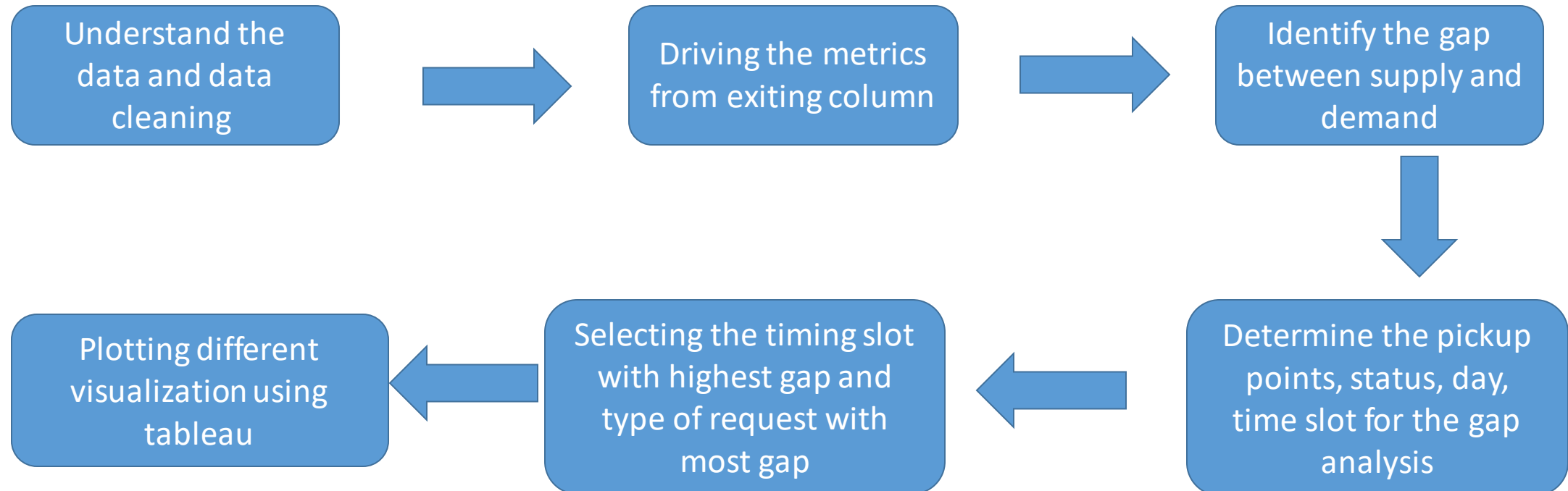
Data Cleaning and Derive metrics

The data provided had various consistency issues and discrepancies.

Some of these are enumerated below:

- The date and time were inconsistent, and also aren't of standard date-time format.

- There are NA values which were treated as a part of our assumptions.

- There were no duplicates found in the dataset.

Derived metrics:

- Served: This variable is a simpler factorized version of the Status column. This tells whether a trip was served or not
- Trip time: the trip duration for each trip
- Request day: This tells us which day of the week it is so that we could test our hypothesis that the days of the week impact the supply and demand
- Request date: This tells us which date of the year it is, so that we could test our hypothesis that the date
- time slot: This variable tells us whether the ride was made during a Peak/Rush hour or not. The peak and rush hours were calculated on the basis of a time plot further in the presentation.
- Request hour: the trip hour when the request was made

Identify the pressing problem

With the performance index trips completed relative to the total request received. Trip completed relative to the total request. 59% of the total request are denied for service either no cars available or request is being cancelled by drivers.
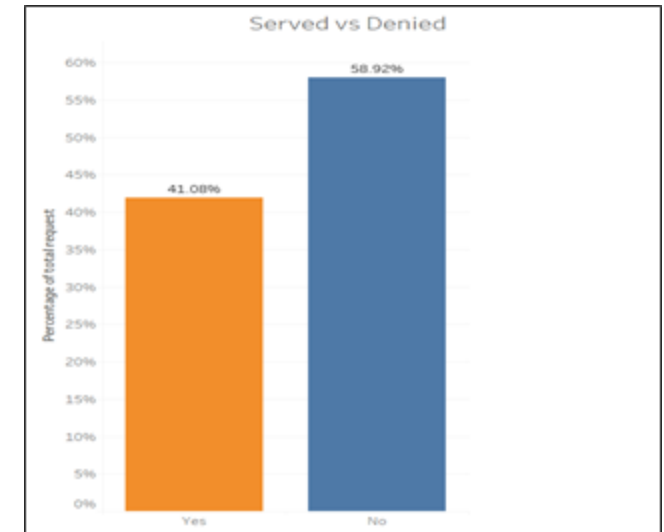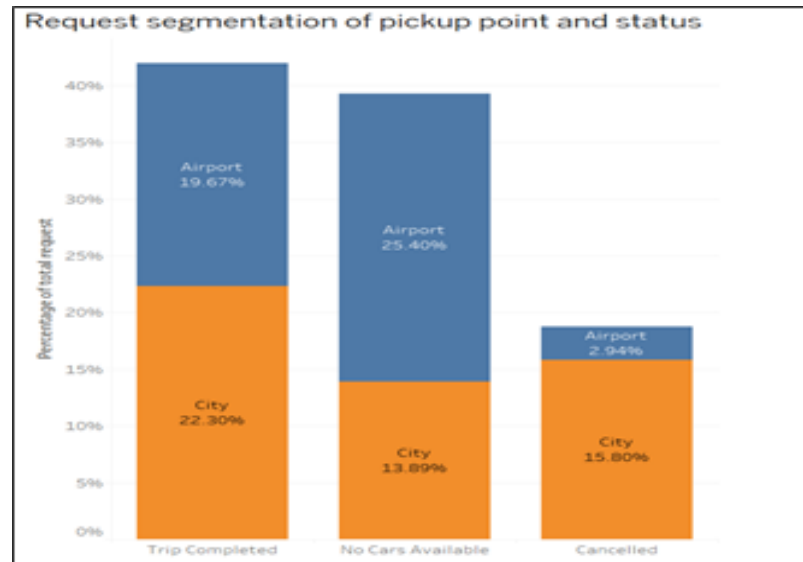
Shows the breakdown of the 59% of customer requests that were denied. Of which, 39% corresponds to no cars available at the time of customer request and 18.7%corresponds to requests that were cancelled by the driver.

Conclusion: The above two plots shows the service and the request of each factor top overall customer request denial.
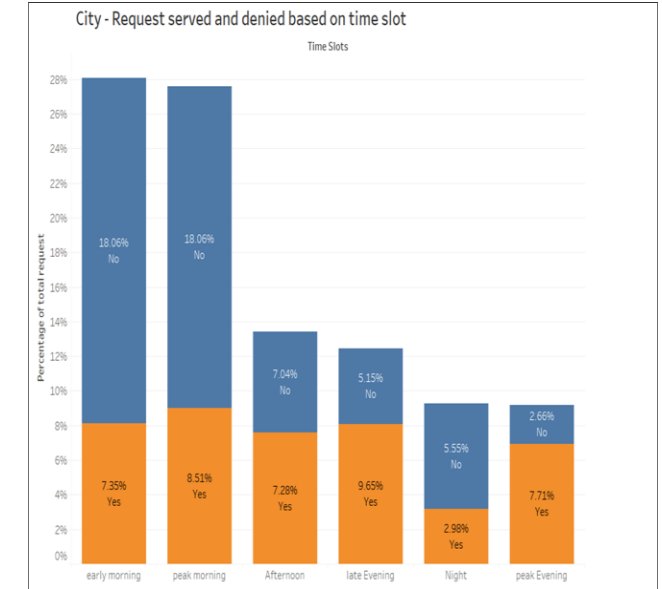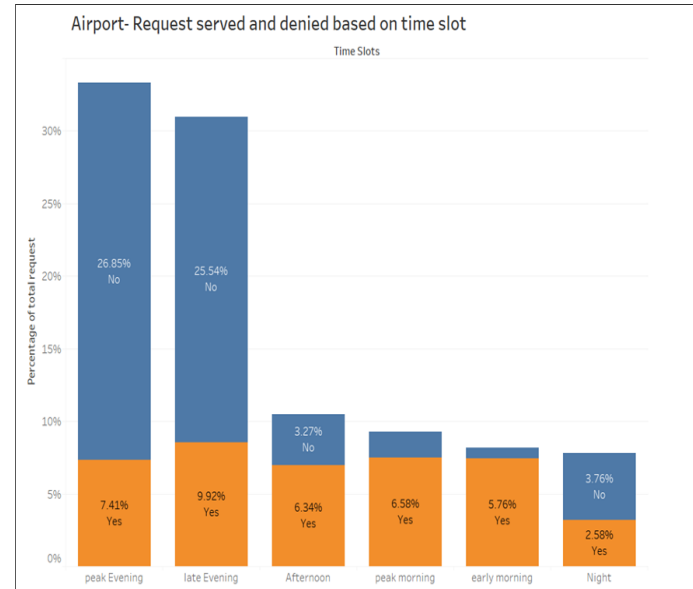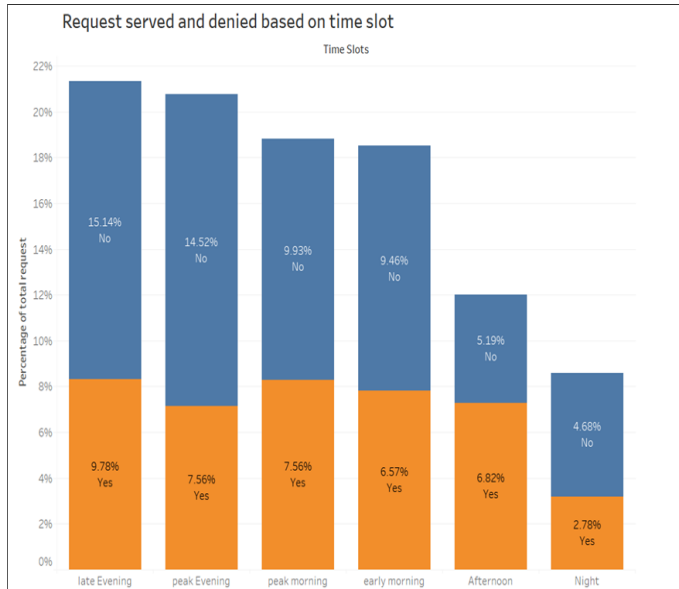
Identify the Most problematic type request
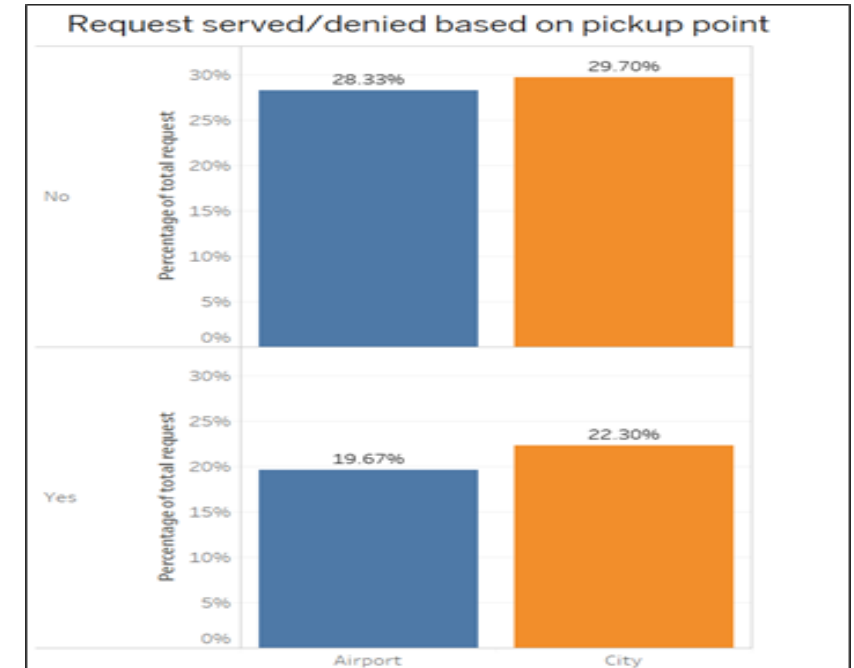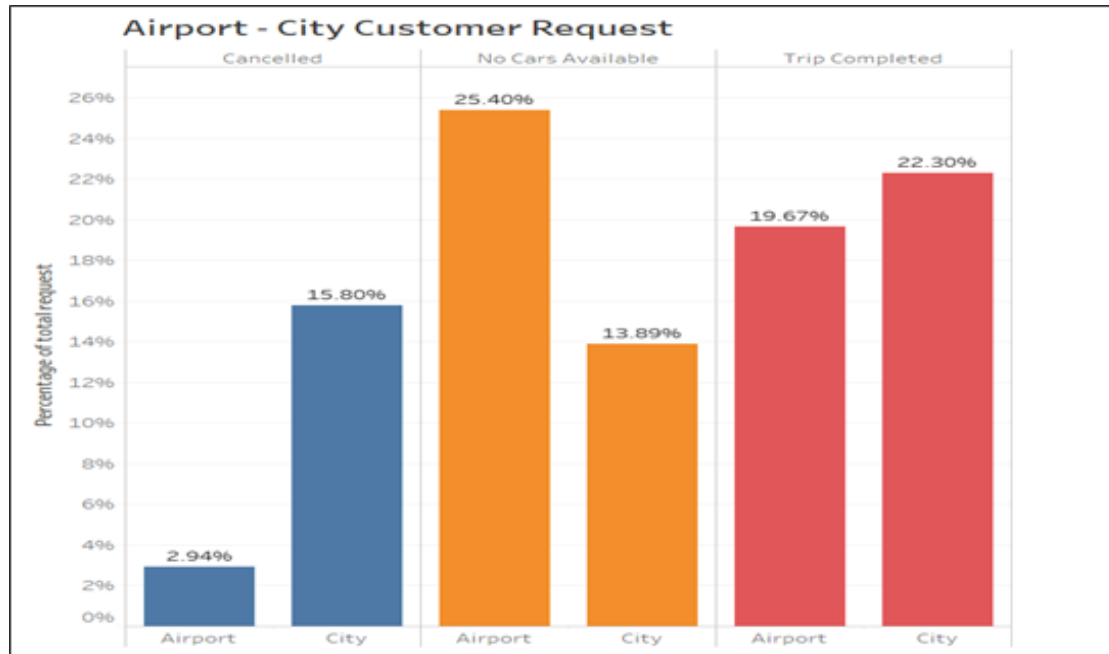
The above shows us that a higher number of requests are denied with pickup pint as city than that of airport. However, when we check the percentage of total requests served with respect to pick up point, we find that service requests made from the city is 1.9% higher than the service requests made from the airport.
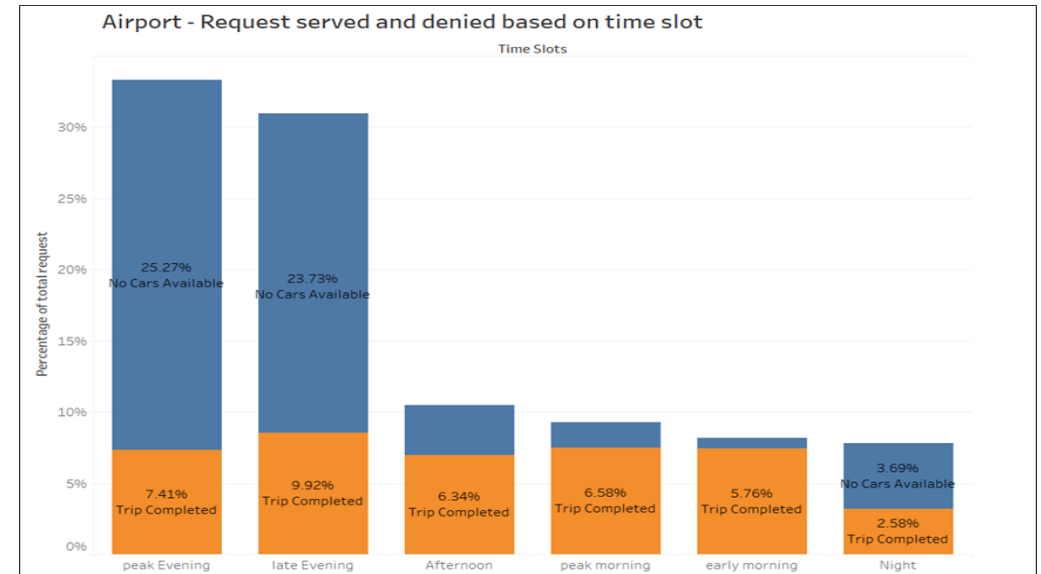
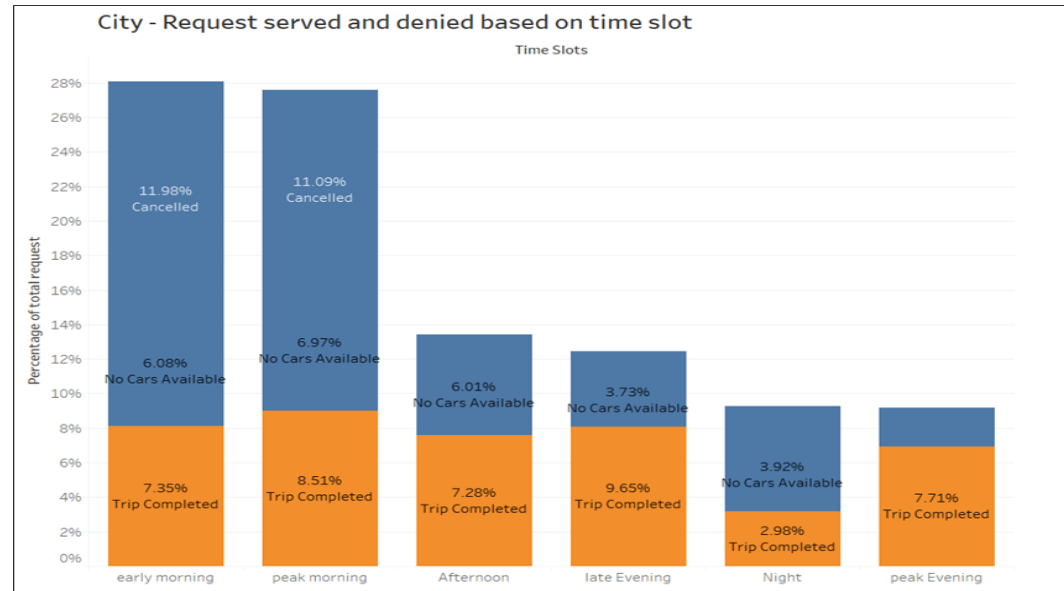Therefore, a higher percentage of request from the airport are denied for service. Hence, requests with pick-up location specified as airport are more problematic and are more likely to be denied for service

Identifying Most Problematic Time-Slot

Clearly shows that the "peak/early morning" and "peak/late evening" time slots are the most problematic segments contributing to the highest gap in overall service and demand.

Identifying Most Problematic Time-Slot

City - Request served and denied based on time slot

Airport - Request served and denied based on time slot

Problem Prognosis

1. The peak/late evening Time slot shows excessively high number of requests denial due to lack of organic flow, i.e. No cars available. This mainly impacts Uber from airport towards the city. This forms our prognosis for relevant business Problem

2. The peak/early morning Time slot shows excessively high number of request for service, cancelled by the driver. This mainly impacts Uber from the city to the airport. This forms our relevant business problem

Identifying Hourly Gap between served
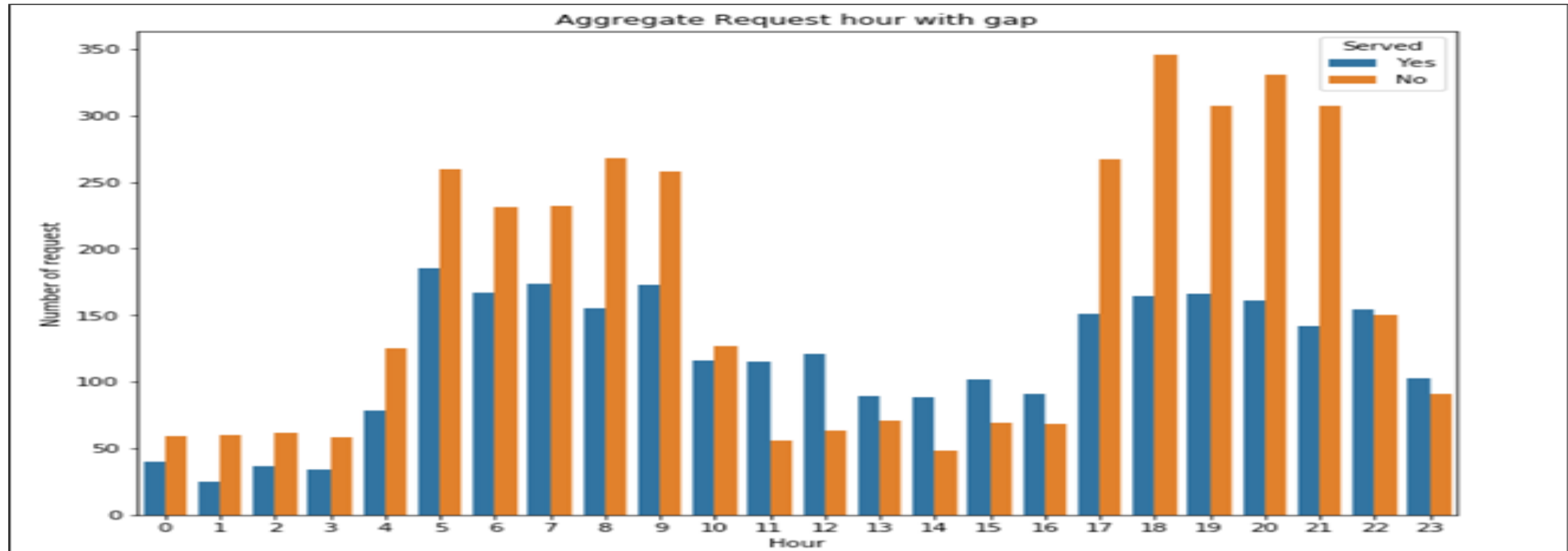
Shows the hourly trend of demand versus supply and the gap variation throughout the day. It can be observed that the demand remains fairly constant between 100 to 200 requests per hour with a sudden surge from 5am to 10am, followed by a return to normal range and surges again between 5pm to 10pm. With the above plot we have formulated problem statements

Conclusions for Analysis

From this analysis it is clear that there is supply-demand gap of Uber service:

1. Peak/early morning hour is most severe for requests from city to airport being cancelled by the driver.
Reason: The driver who makes a trip from the city to the airport has to inadvertently wait for a longer duration to receive a return trip request. Consequently this will lead to increase in driver idle time, which could have been utilized for other trips if he chose not to accept a city to airport request during morning hours. This could be the reason of high number of city-airport request cancelled by the driver.

2. Peak/late evening from airport to city due to no cars available.

Reason: This could be due to the high number of flights landing during the late evening to night time. Therefore, there is significant reduction in natural supply of taxi's to the airport during the late evening rush hour. This imbalance leads to a surge in the demand for cars from customers desiring to leave the airport.

Recommendations

1. Increasing the profit margin for drivers making trips from the city to the airport during morning hours

2. Surge pricing, by increasing the rate charged to the customer we can increase the revenue. The benefit of surge charge can be passed to driver

3. Reduce the share ride fare to encourage ride-sharing or Taxi pooling for customers leaving the airport during the evening hours

4. Surge Charge the customer for evening trips from the airport to encourage more drivers to initiate a trip toward the airport during the late evening
 rush hour.
5. Maintain transparency with respect to the surge charging of customers during evening/morning hour trips