Uber Supply-Demand Gap Case study

By:- Amrendra Singh Rathore

Data Cleaning and Manipulation:-

Data set contains 6 columns:

1. Request.id
2. Pickup.point
3. Driver.id
4. Status
5. Request.timestamp
6. Drop.timestamp

Possible data inconsistencies:

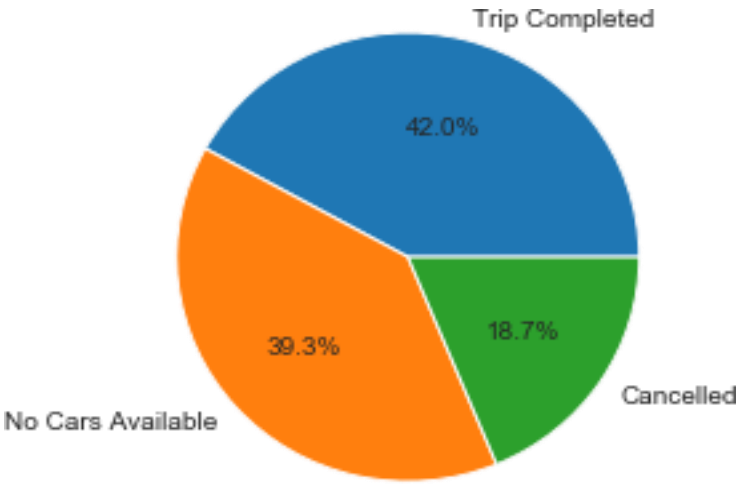
1) Duplicate values of Request ID
2) NA values in the columns of interest

Other Issues:

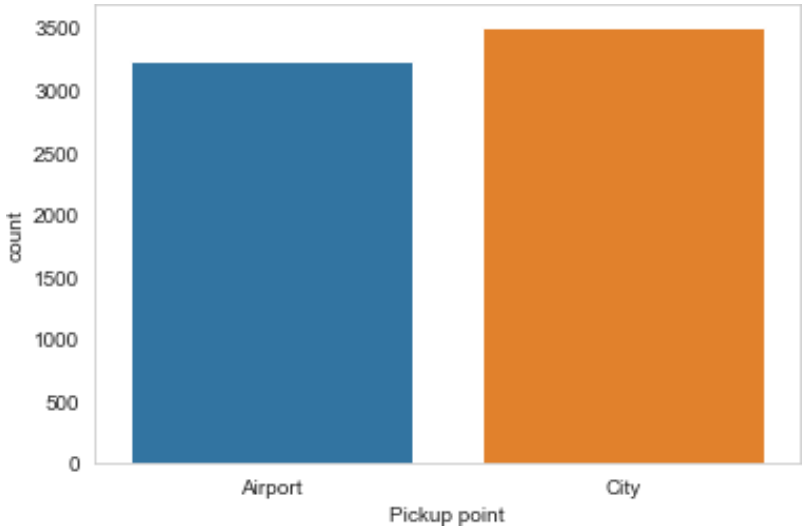
1. Request time stamp is object here. Convert it to datetime format.
2. Dates are separated by "/" and "-". Make this consistent for ease of data analysis.

Data Analysis:-

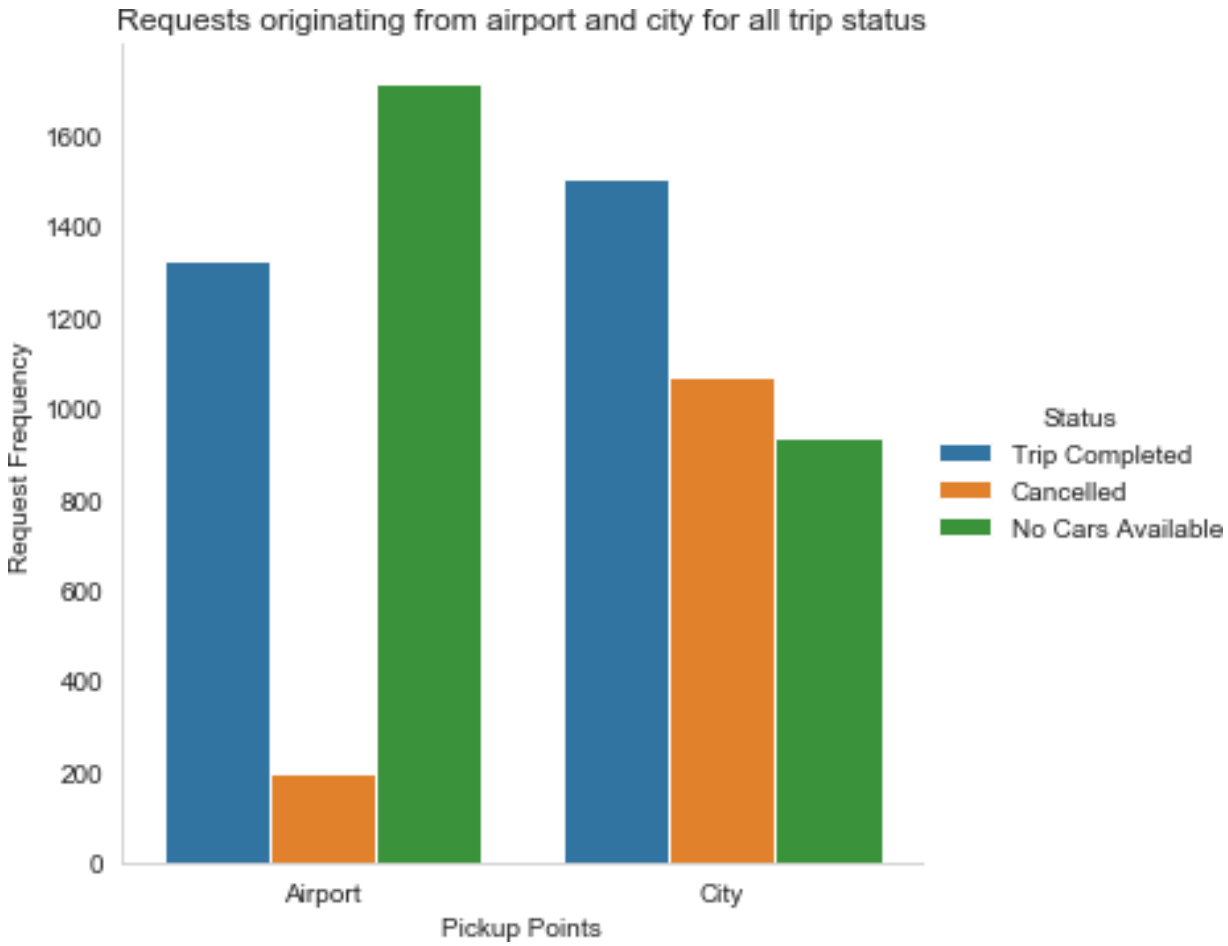
At this point the data is ready for analysis with all the required columns in place.

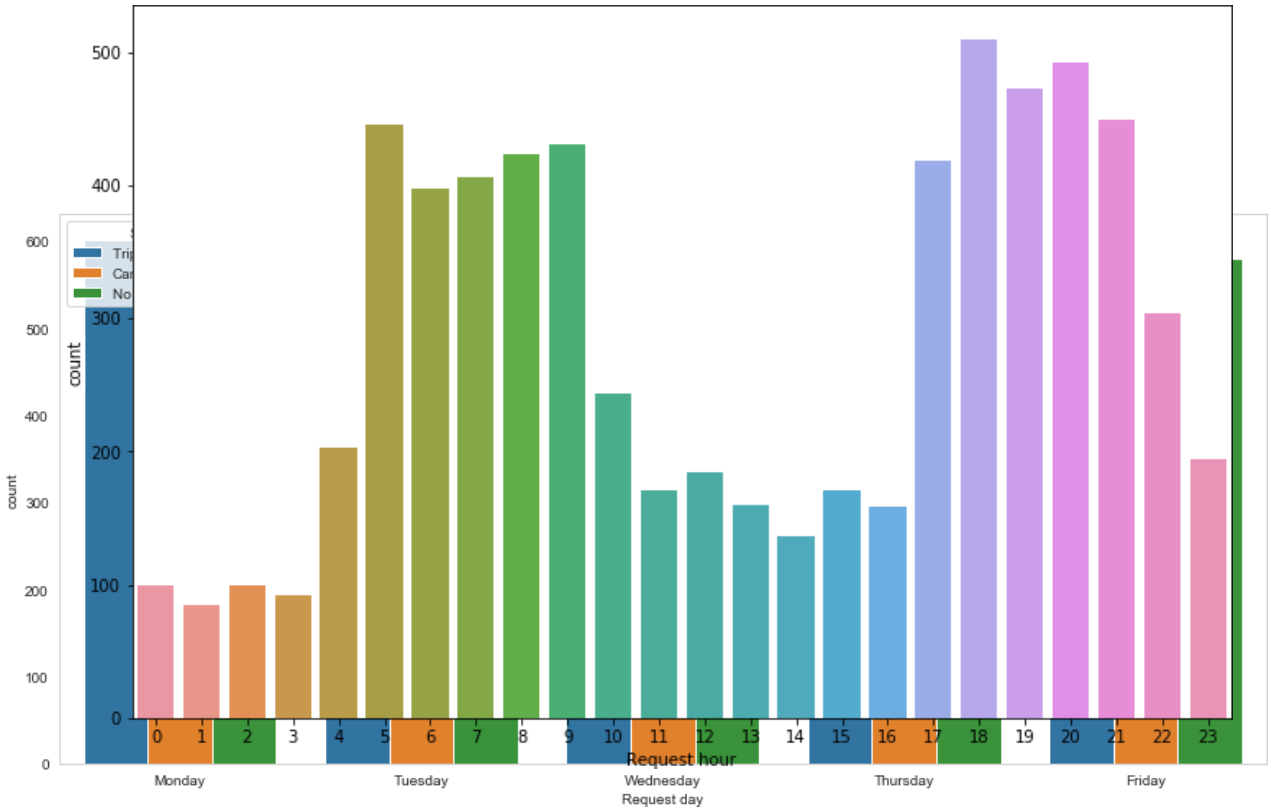


Viewing the frequency of request with their status shows us the above the figure. It is clear that most of the failed request are due to no cars being available at that moment. Still further analysis is need to give a concrete conclusion and the cause for this.



From above plot it can be seen that there are two pickup points Airport and City. They have almost equal request however city have slightly more requests than airport.



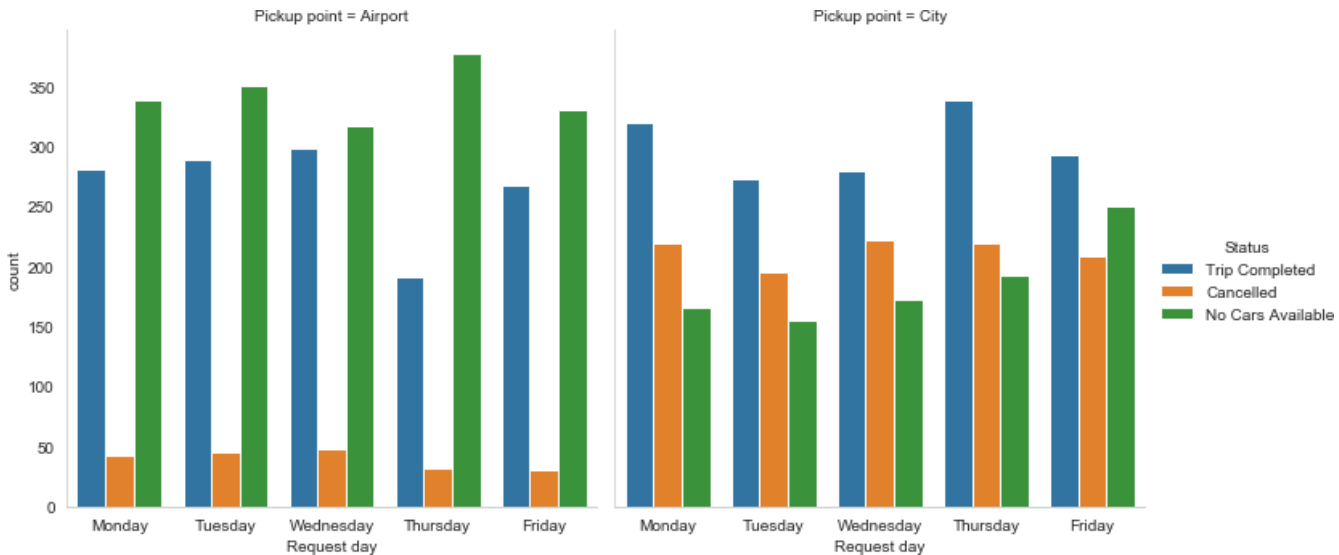


This figure shows us that in cites the greatest number of failed requests are due to cancelled trips. This might be due to personal reason of the driver. The car unavailable will require it plotting against time of the day. We will do this later.

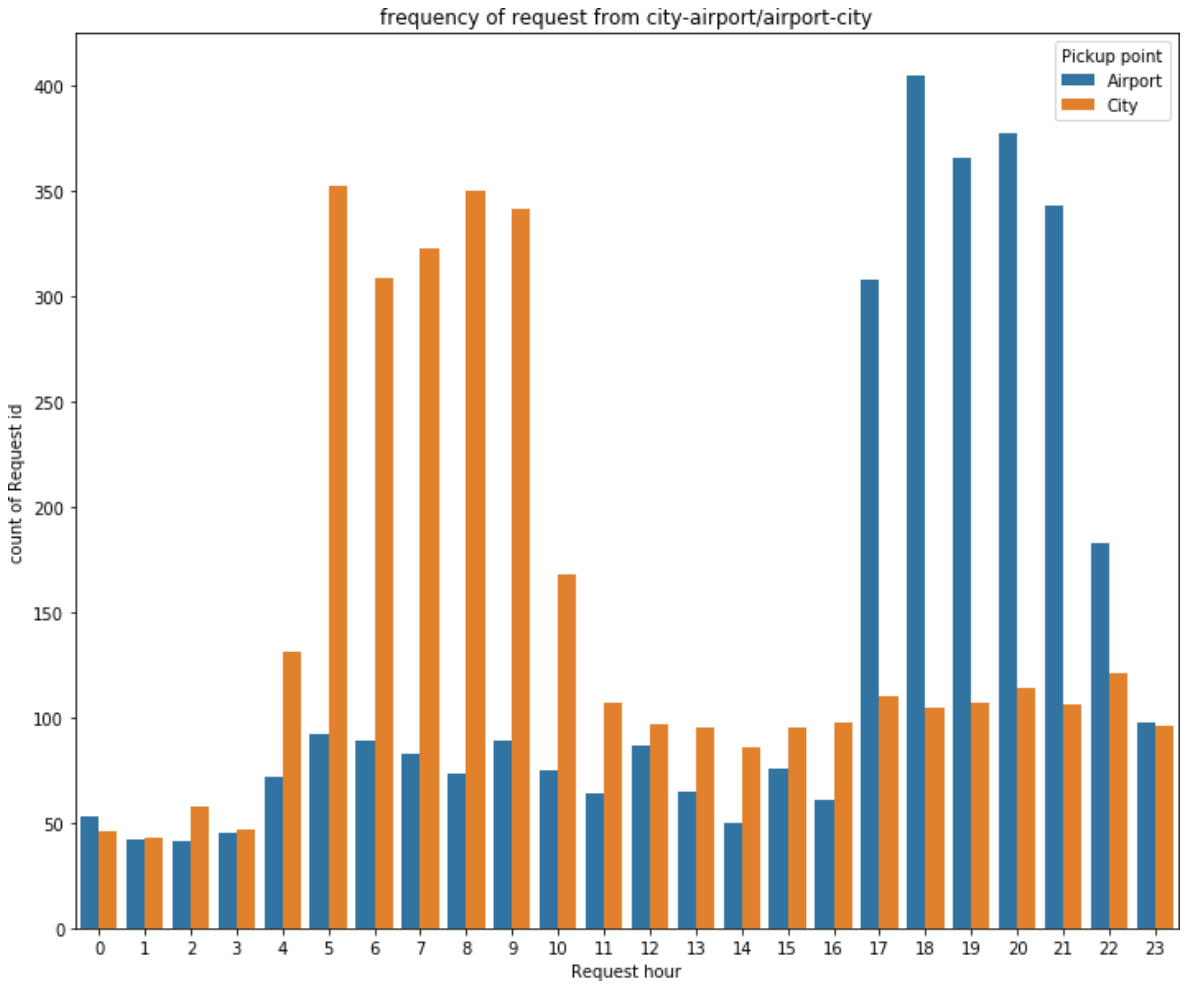
From above plot it can be see that the cancelled trip frequency is almost constant.

From above plot it can be seen that -

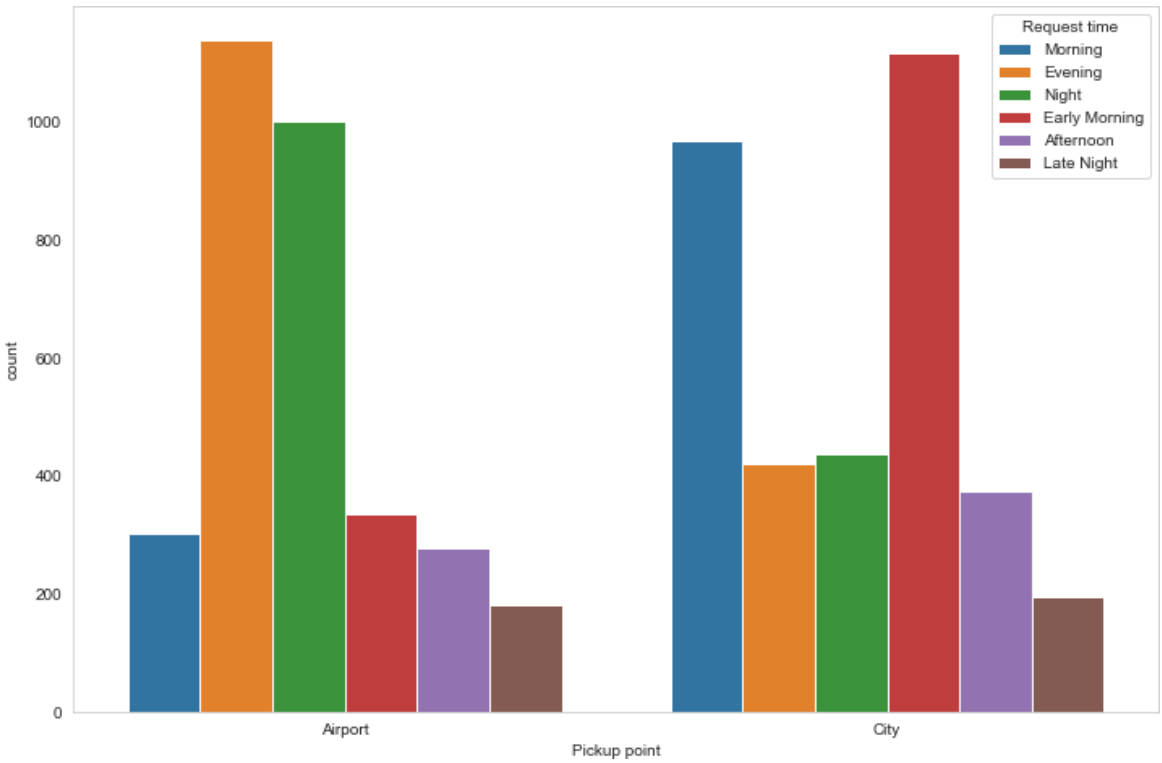
- For trips from city - airport, except for Friday, cancelled trips are more for all other days of week
- For trips from airport - city, for all days there are a greater number of No Cars Available



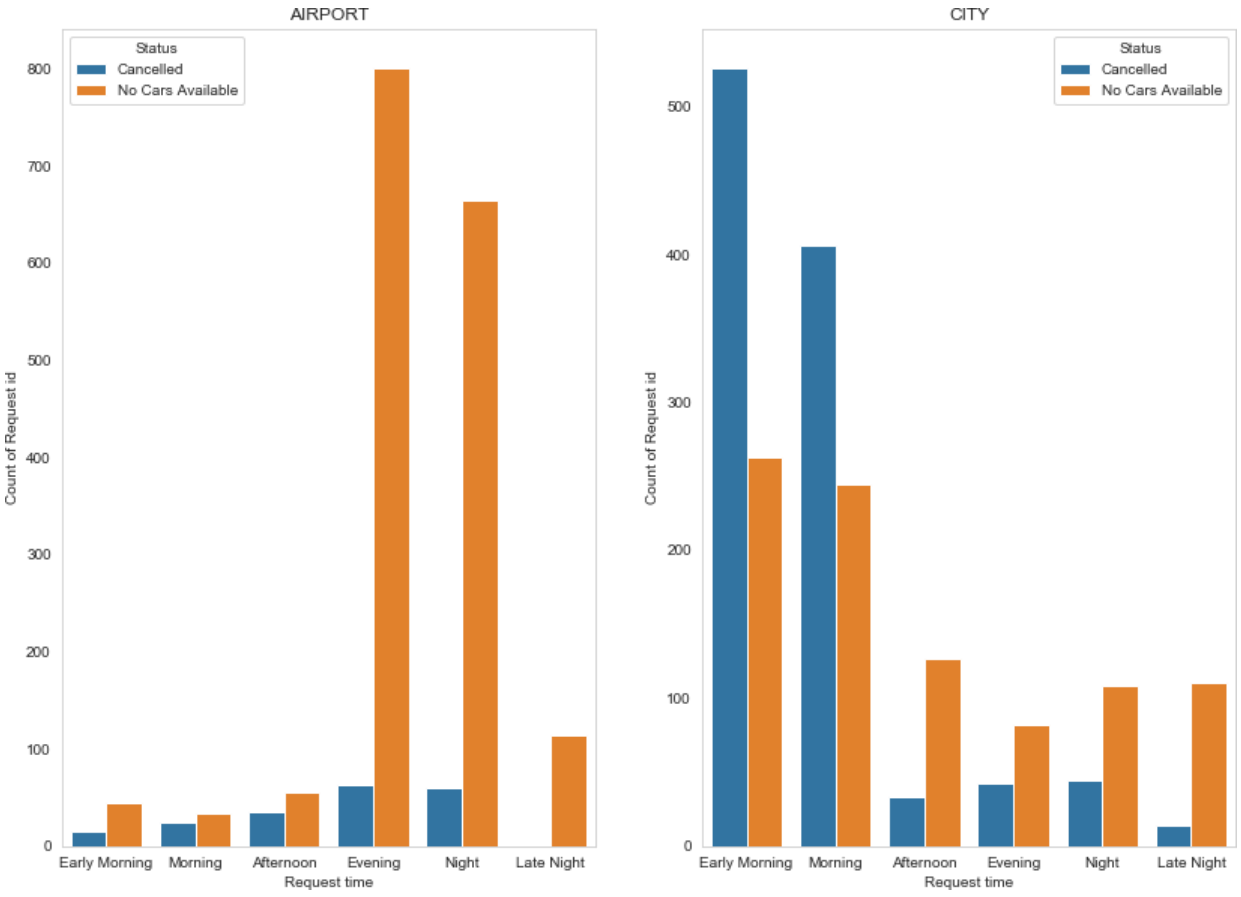
From above plot it can be seen that the requests in early and late hour is more.



From above plot it can be seen why that there are higher request during **Morning** hours at city and **Evening**
hours at airport which is expected as it is rush hour.



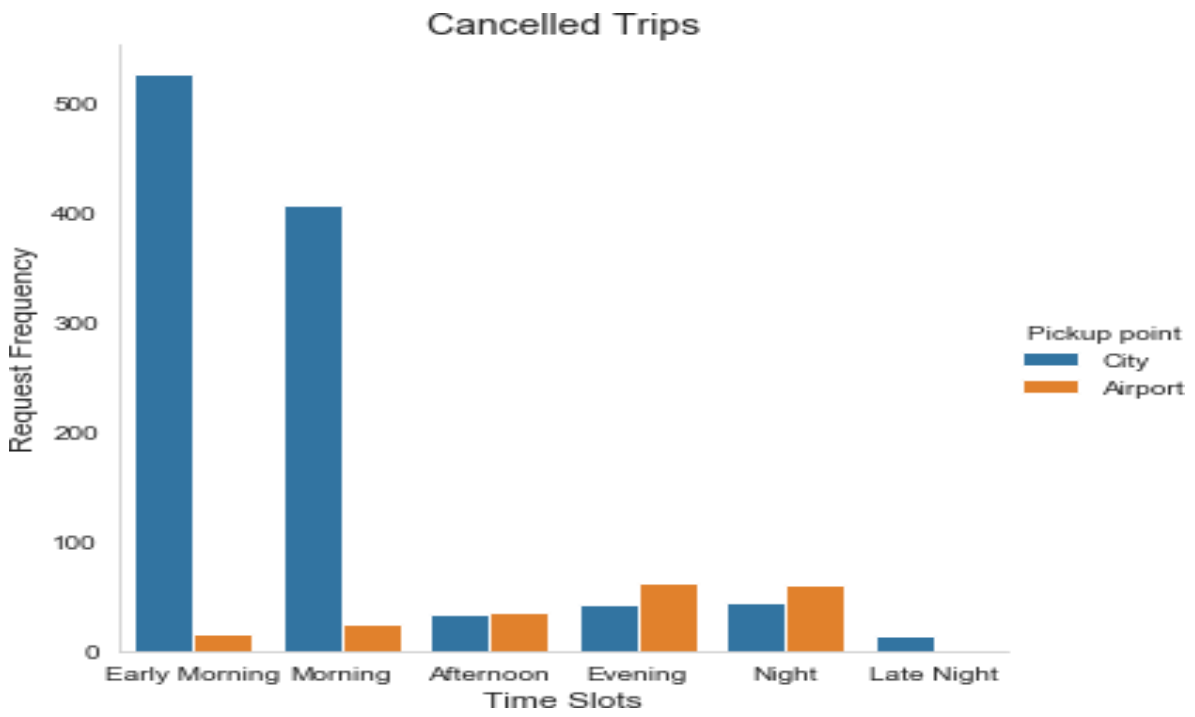
From above plot it can be see that **Airport** has higher request in **Evening** and **City** has higher request in the **Morning**.



The airport data is highly skewed in favor of no car being available this might me due to the fact that the airports are generally far away from the city and drivers generally do a one-way trip to airport.

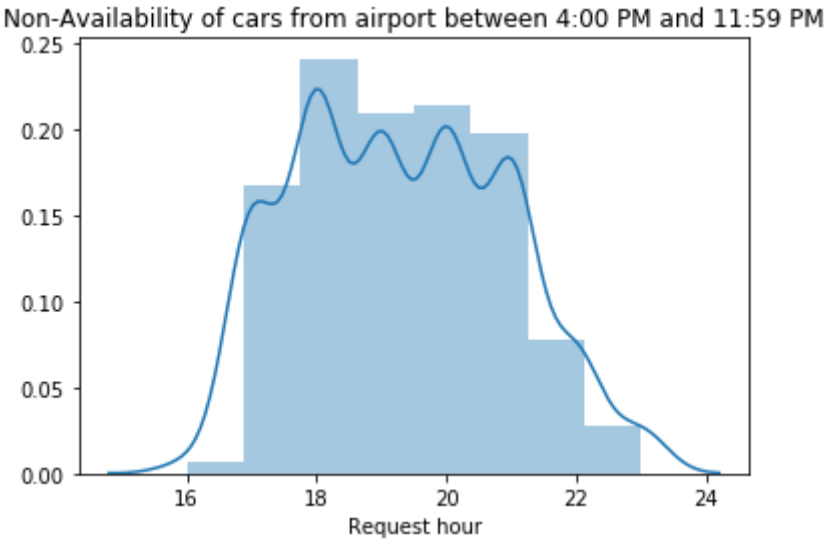
This figure plots the status (cancellation) against time. We can see that most of the trips in the city are cancelled in the early hours. At airport the same is for in the late hours this may be due to the fact the drivers aren't able to get a return booking.

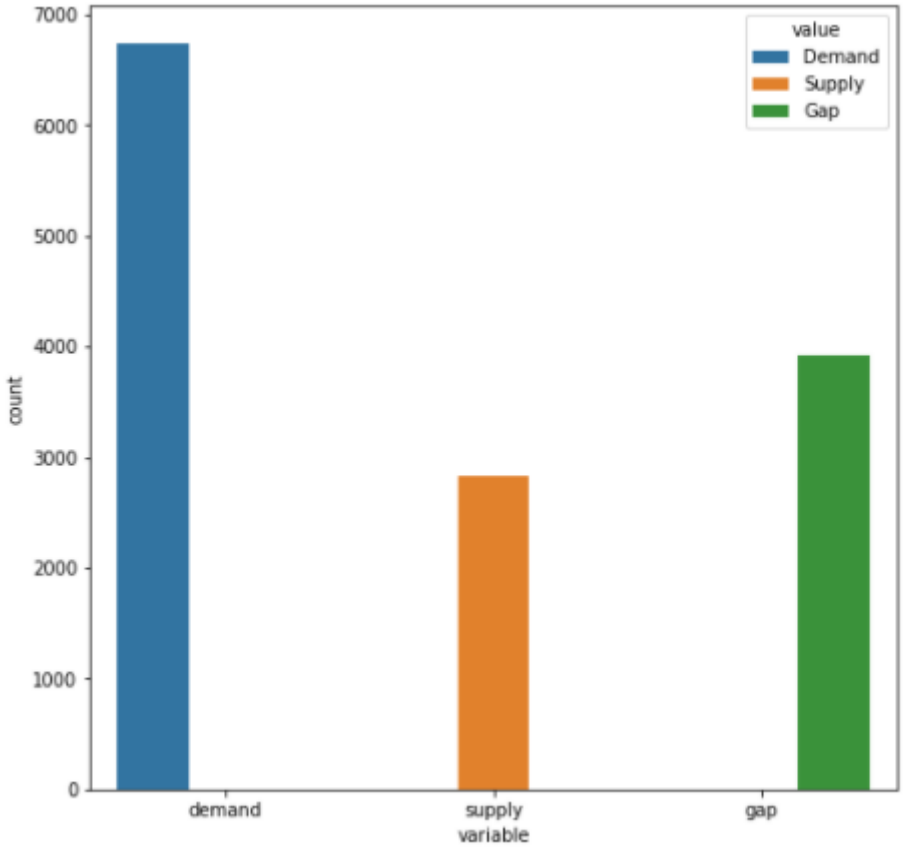
This graph clearly confirms our hunch that the drivers aren't available near airport during late night. This
results in failed trips.



The plot shows that more number of cars are unavailable around 6:00 PM followed by 8:00 PM, 7:00 PM, 9:00 PM and 5:00 PM.

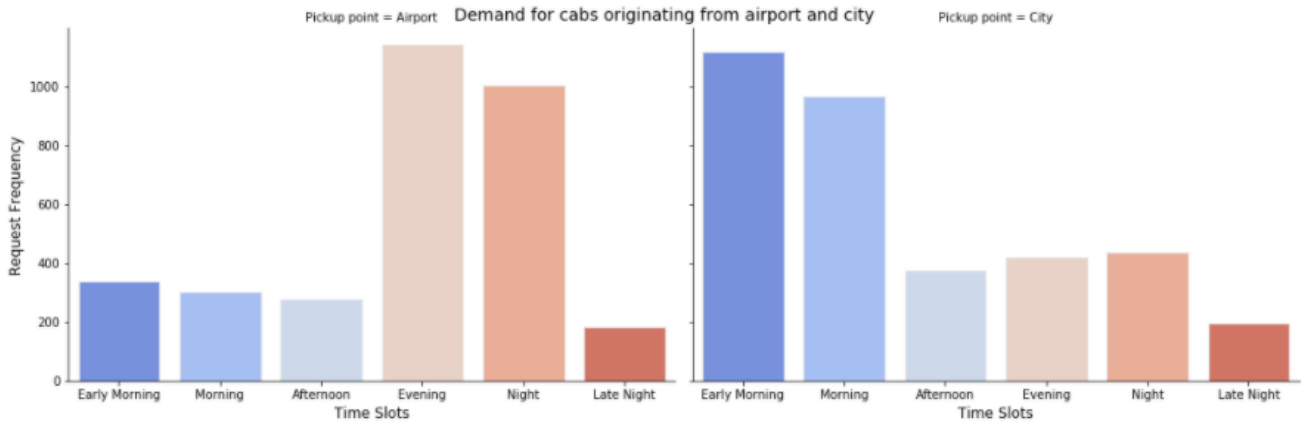
Hence, the most pressing problem is as follows :
- More cars are cancelled from city - airport, for timeslots between 4:00 AM and 11:59 AM
- More cars are unavailable from airport - city, for timeslots between 4:00 PM and 11:59 PM



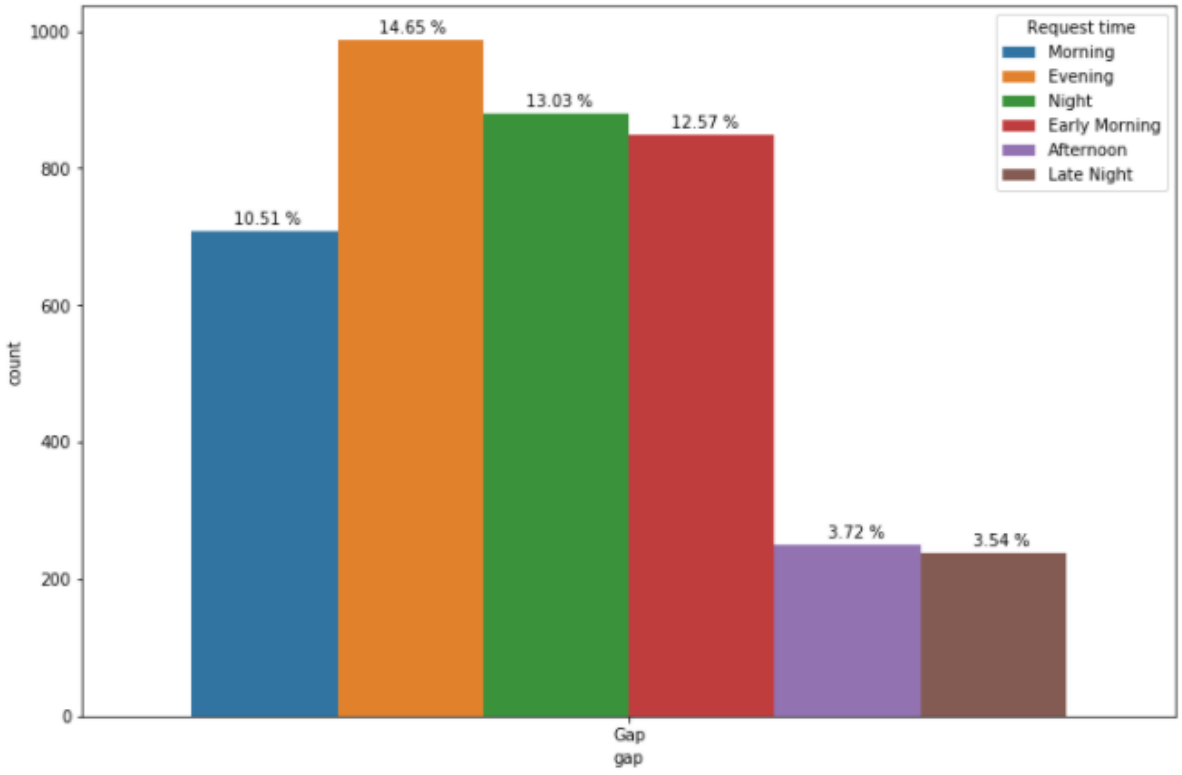


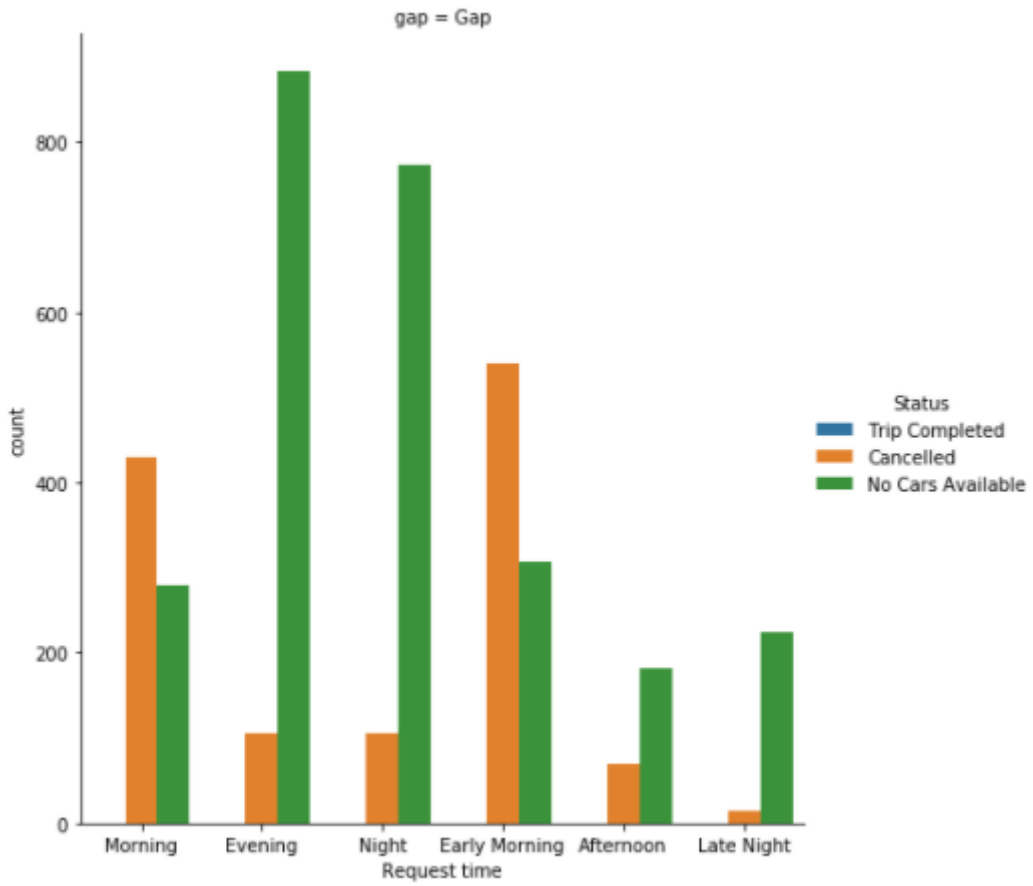
Irrespective of the trip status, following can be deduced from the above plots :-

- For airport pickups, demand is more during the evening and night hours.
- For city pickups, demand is more during the early morning and morning.

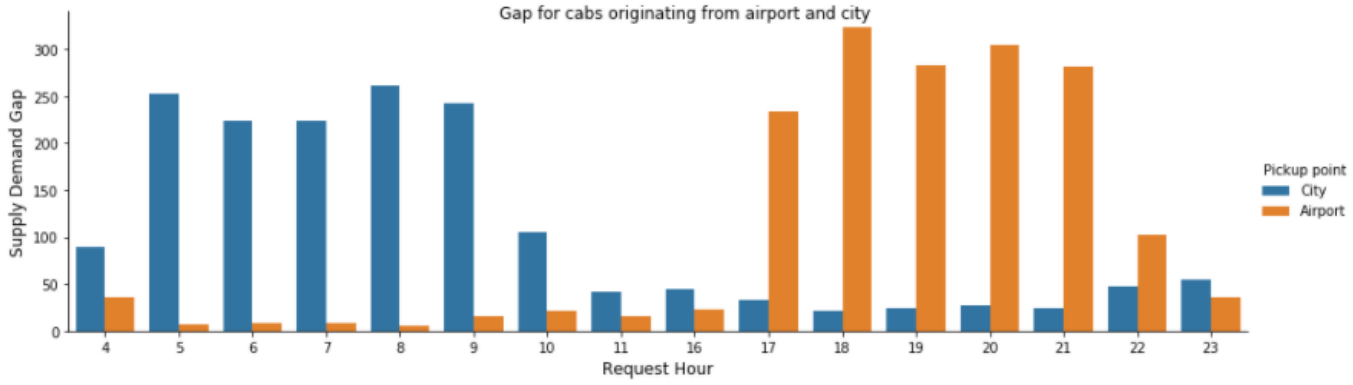


In the below graph - the percentage is taken only across the total gap percentage which is 58.02%.From the total gap percentage(58.02%) -around 15% gap exist during morning and evenings.





The below plot shows that the most severe supply demand gap is for the trip request from Airport - City, between 6:00 PM to 9:00 PM.



The above plot shows that the most severe supply demand gap is for the trip request from Airport - City, between 6:00 PM to 9:00 PM

Hypothesis :

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.

Conclusions :

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

i. 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.

ii. 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.

iii. Last but sure solution to bring down the gap is to increase the numbers of cab in its fleet.