

# Uber Supply-Demand Gap analysis

Prepared by:  
Poovarasan



# Abstract

## Objective:

To identify the root cause of the cancellations & non-availability of cars and recommend ways to improve the situation.

## Strategy:

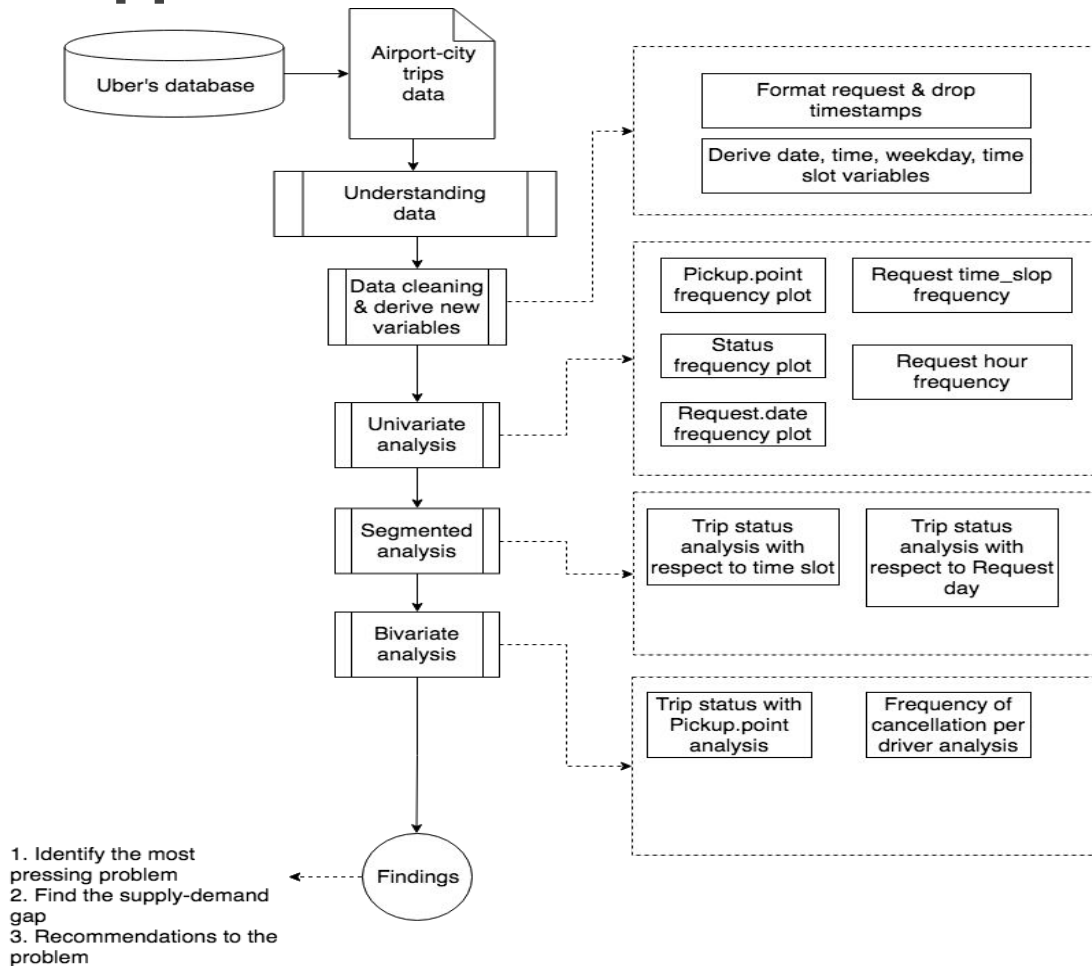
To reduce the supply - demand gap.

## Assumptions:

A 24 - hour day is split into 8 time slots as follows

00:00 - 2:59 - Night, 3:00 - 5:59 - Late night, 6:00 - 8:59 - Early morning, 9:00 - 11:59 - Morning, 12:00 - 14:59 - Afternoon, 15:00 - 17:59 - Evening, 18:00 - 20:59 - Late evening, 21:00 - 23:59 - Early night

# Solution approach





# Data sourcing, cleaning & derive new columns

Data:

The masked data set of Airport - City & City - Airport trip information of 3rd week of **July** 2016.

Data cleaning:

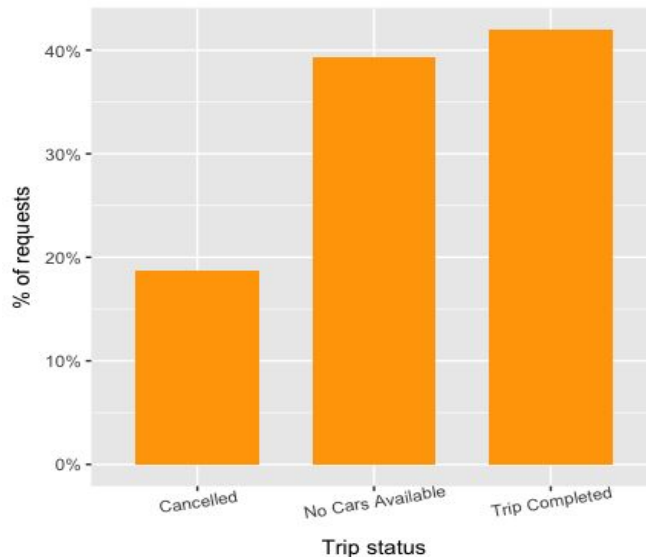
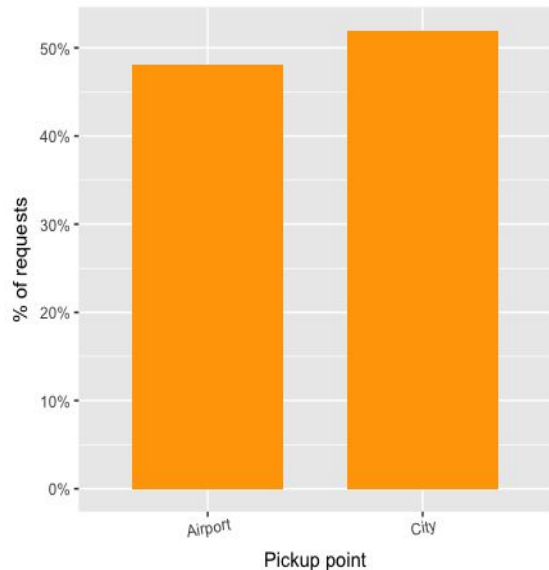
1. The data source supplied has everything(row, columns, values) cleaned except the Request.timestamp & Drop.timestamp
2. Format the Request.timestamp & Drop.timestamp columns

Derive new columns:

1. Request date, Request time, Request day(week day), Request time slot, approx travel time.

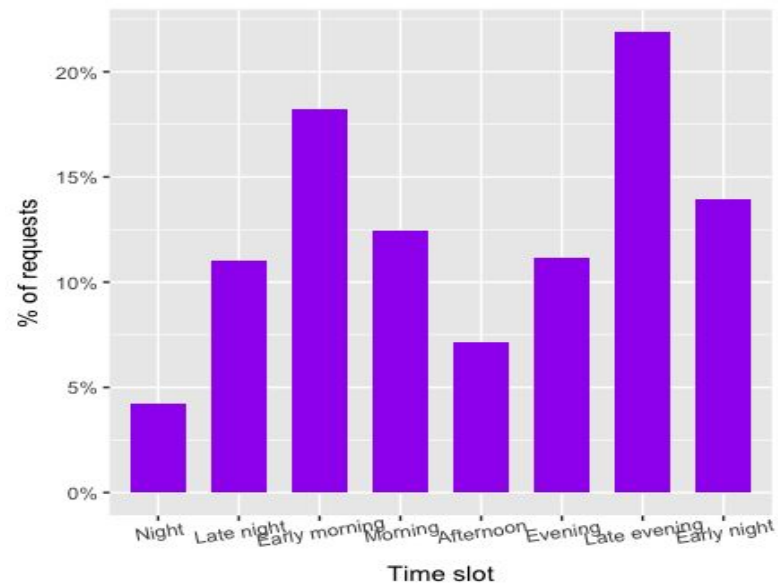
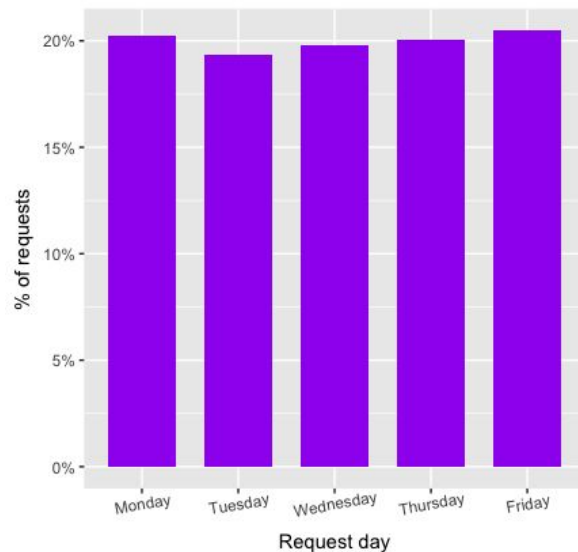
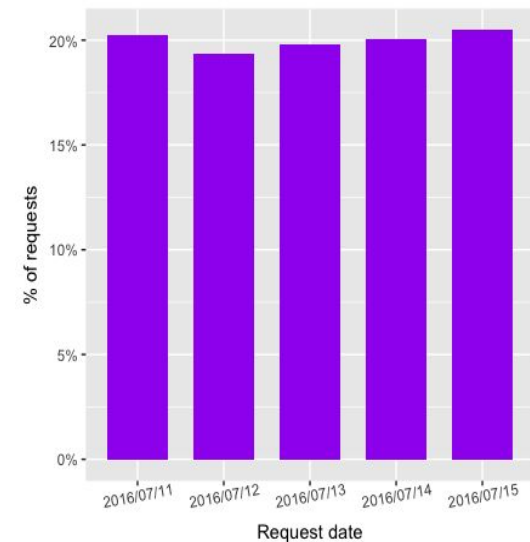
# Data analysis - Univariate analysis

- Requests received from Airport is 48%( 3238) & City is 52%( 3507)
- Trips completed are 42% (2831) & 58% of requests are in demand (39% no availability & 19% cancelled)



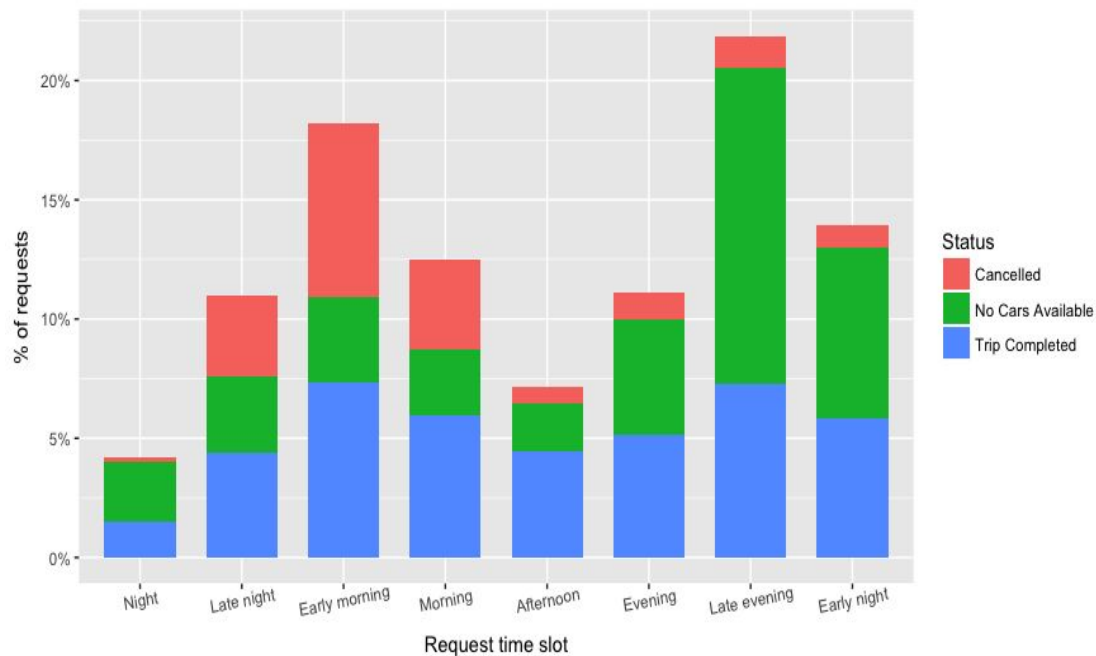
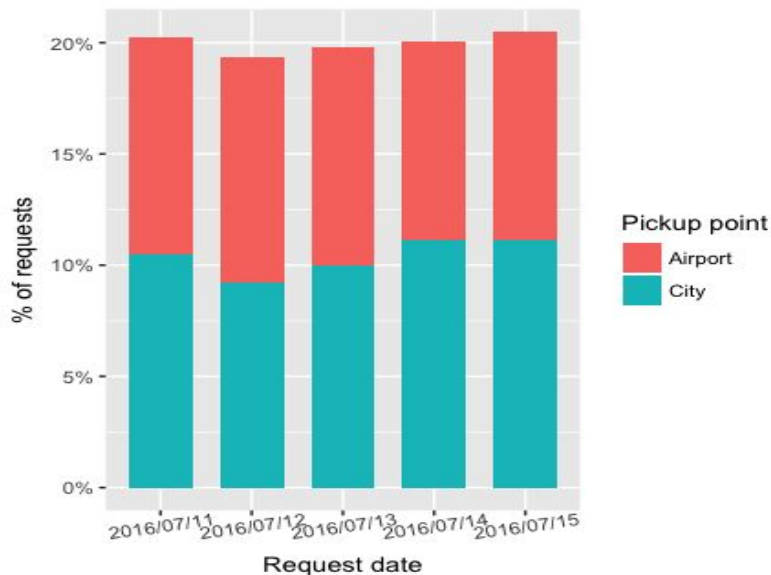
# Data analysis - Univariate analysis contd...

- Total no. of requests received per day are nearly same (between 19% - 20%) which is possible **hypothesis**
- Early morning & late evening received the highest no. of requests (6 AM - 7.59 AM, 6 PM - 8:59 PM)
- Received less no. of request at the night (Midnight - 2.59 AM)



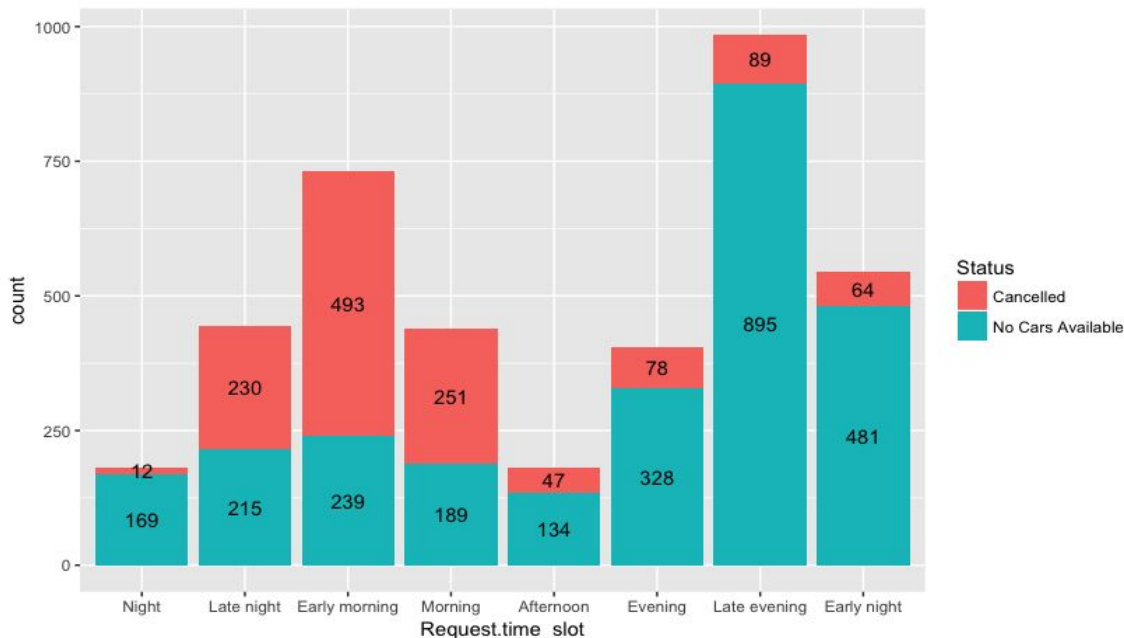
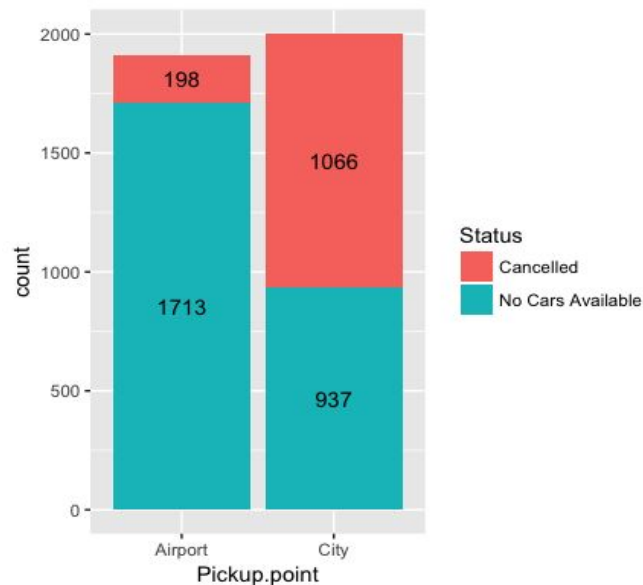
# Data analysis - Segmented analysis

- The no. of requests received from Airport & City are equal on 14 & 15th of July '16 and other days are nearly equal
- Demand** for cars are high at **late evenings** & more **cancellation** done at **early morning**



# Data analysis - Bivariate analysis

- **Airport** requests show **no cars available** for the maximum numbers( 1713)
- More **cancellation** done at the **City** (1066)
- More **demand** at the **evenings, late evenings & early night**( 328, 895 & 481)
- More **cancellation** done at **late night, early morning & morning**( 230, 495 & 252)

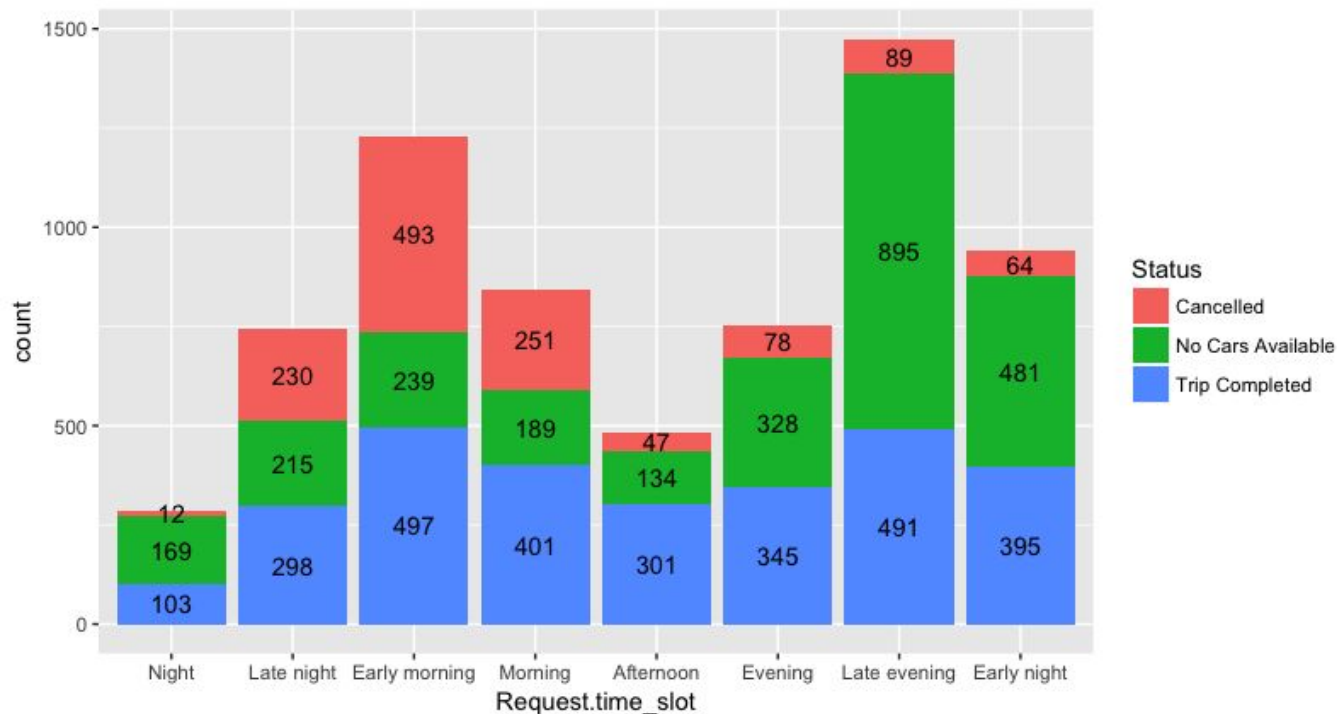




# Data analysis - Bivariate analysis contd...

Supply demand gap:

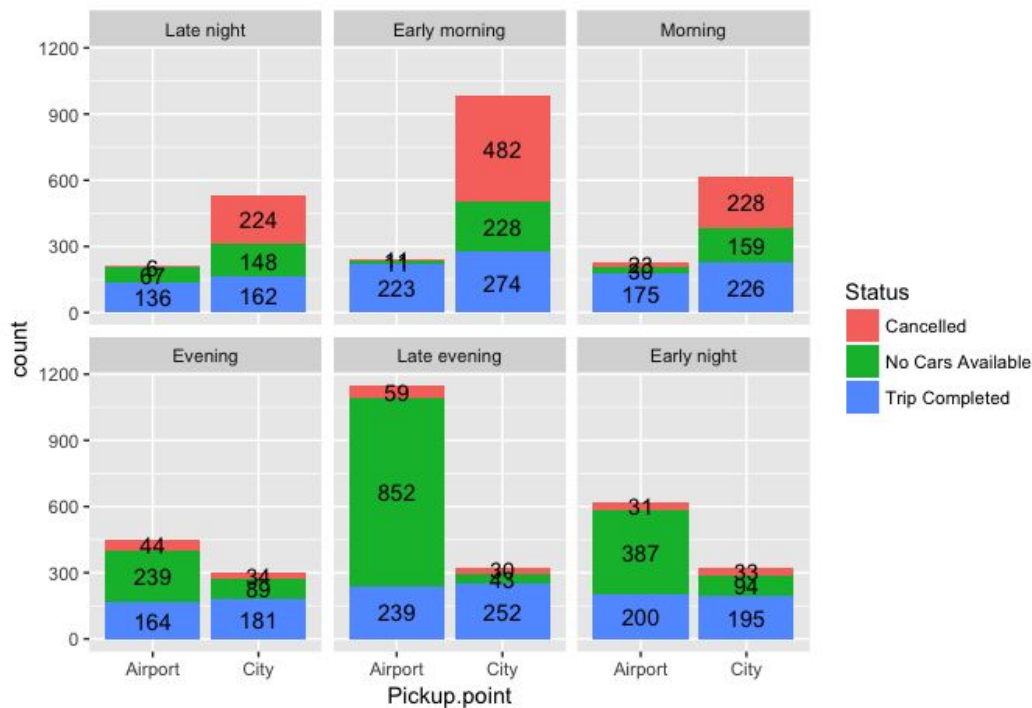
- The **highest demand** for the cars found at **Airport** whereas **highest cancellation** done at the **City**.



# Data analysis - Bivariate analysis contd...

Find the type of requests for which the gap is most severe in the identified time slots:

- Most cancellation done at **City** in the **Early morning**
- More demand at the **late evening** in the **Airport**



# Conclusion

## Reasons for the supply - demand gap:

The more demand at the Airport could be because of the more cars are available at City(1000 cancellation), and the drivers are cancelling the requests.

Most no. of demand in the late evening at the Airport could be because of the traffic at City in the evening. And the same could be the reason for more cancellation of City requests in the early morning.

There may also be the high/dynamic charge at the late evening would be the reason for high demand.

## Recommendations:

1. Introduce the incentives for the early morning & late evening trips. The ratio of cancellation & completion per driver can be considered for the incentives.
2. Introduce the offers for the late evening and early morning trips.

**Further analysis recommendations:** Considering the charges, dynamic fares & cancellation initiators( driver/customer) and also the traffic trends would provide more insights to identify & rectify the supply - demand gaps