

Uber Case Study

(11 July 2016 - 15 July 2016)

Prepared by,

Hariharan Subramaniam

Problem Statement

Uber customers experience two problems, a lot of cancellation of cars by drivers and unavailability of cars. This is a problem customers experience when they wanted to travel from city to airport and airport to city.

Business Objective

- a) To identify the root cause of Cancellations and Unavailability of cars from city to airport and vice-versa.
- b) To recommend ways to improve the situation.

Questions to be answered

- a) Visually identify the most pressing problems for Uber.
- b) Find out the gap between supply and demand and show the same using plots.
- c) What do you think is the reason for this issue for the supply-demand gap?
- d) Recommend some ways to resolve the supply-demand gap.

Problem solving methodology

Following steps will be taken ,

- 1.Create a metadata page ,detailing the basic information about the dataset.
- 2.Perform univariate analysis on appropriate columns and identify patterns.
- 3.Plot graphs to demonstrate findings.
- 4.Summarise the finds along with recommendations to improve the situation.

Meta Data

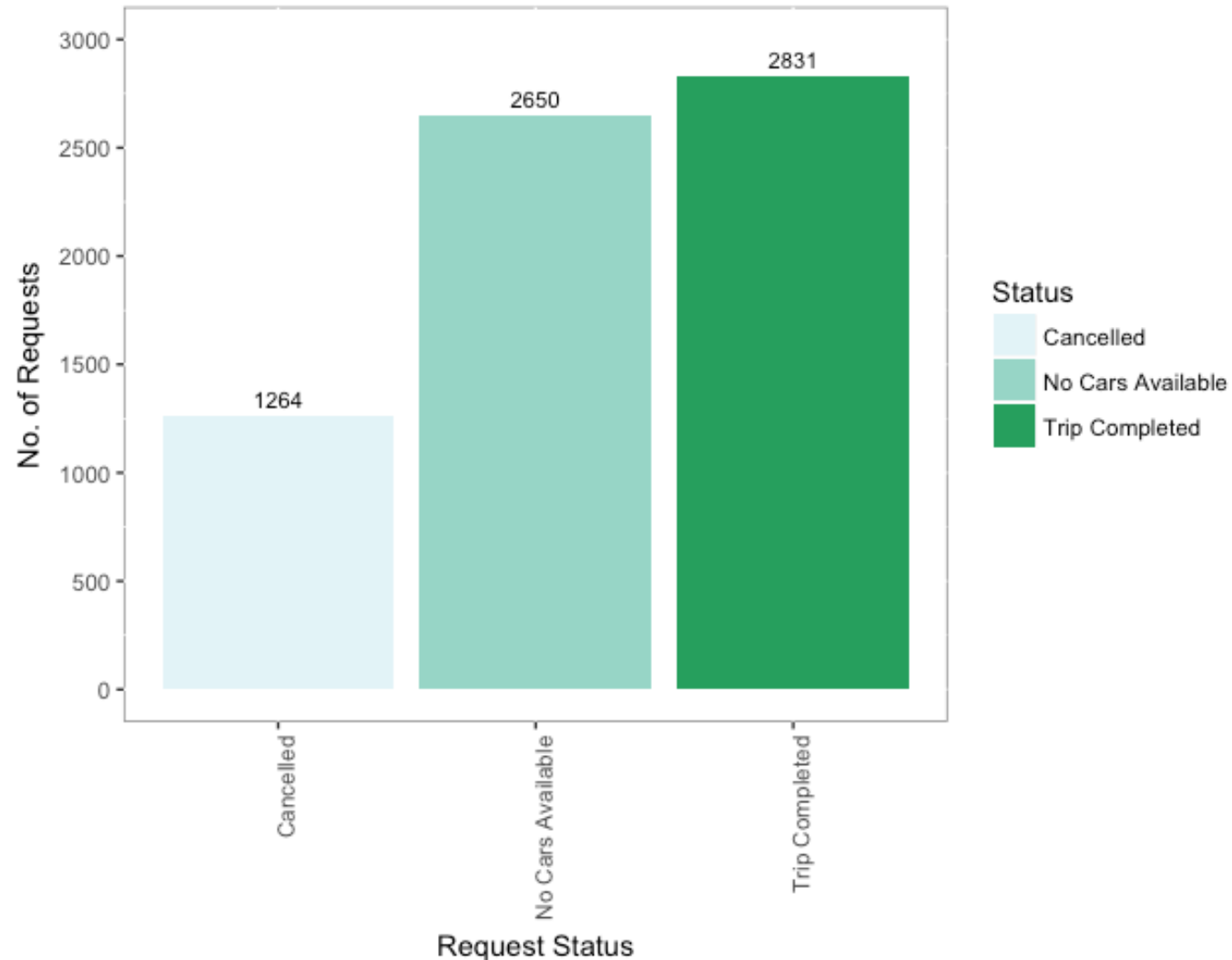
Meta Data	Description
Description	Uber data between 11-July-2016 to 15-July-2016.A subset of data containing details of requests from City to Airport and Vice-versa including its statuses such as fulfilled, cancelled and unavailability of cars.
Source	Upgrad
Format	csv
Number of rows	6745
Each row is	A request for a cab to travel from City to Airport or Vice Verse
Subset	The dataset is a subset of cab requests between airport and city between the dates 11July2016 - 15July2016

Column Name	Type	Description	Uniqu e Value	Number of NAs	Number of Blanks
Request Id	Integer	Unique request id of every request	6745	Nil	Nil
Pickup.point	Char	Pickup point(City,Airport)	2	Nil	Nil
Driver id	Integer	Unique id of drivers	301	2650	Nil
Status	Char	Statuses such as 'Trip completed', 'Cancelled', 'No Cars Available'	3	Nil	Nil
Request timestamp	Date Time	Time at which the car was requested at pickup points City ,Airport	5618	Nil	Nil
Drop timestamp	Date Time	Time at which the request was completed	2599	3914	Nil

(58% of the requests are not fulfilled which is a potential market to catch up for Uber)

Overall City/Airport Status Distribution

It appears only 42 % of the requests are fulfilled and the rest are either cancelled or are not fulfilled due to unavailability of cars.



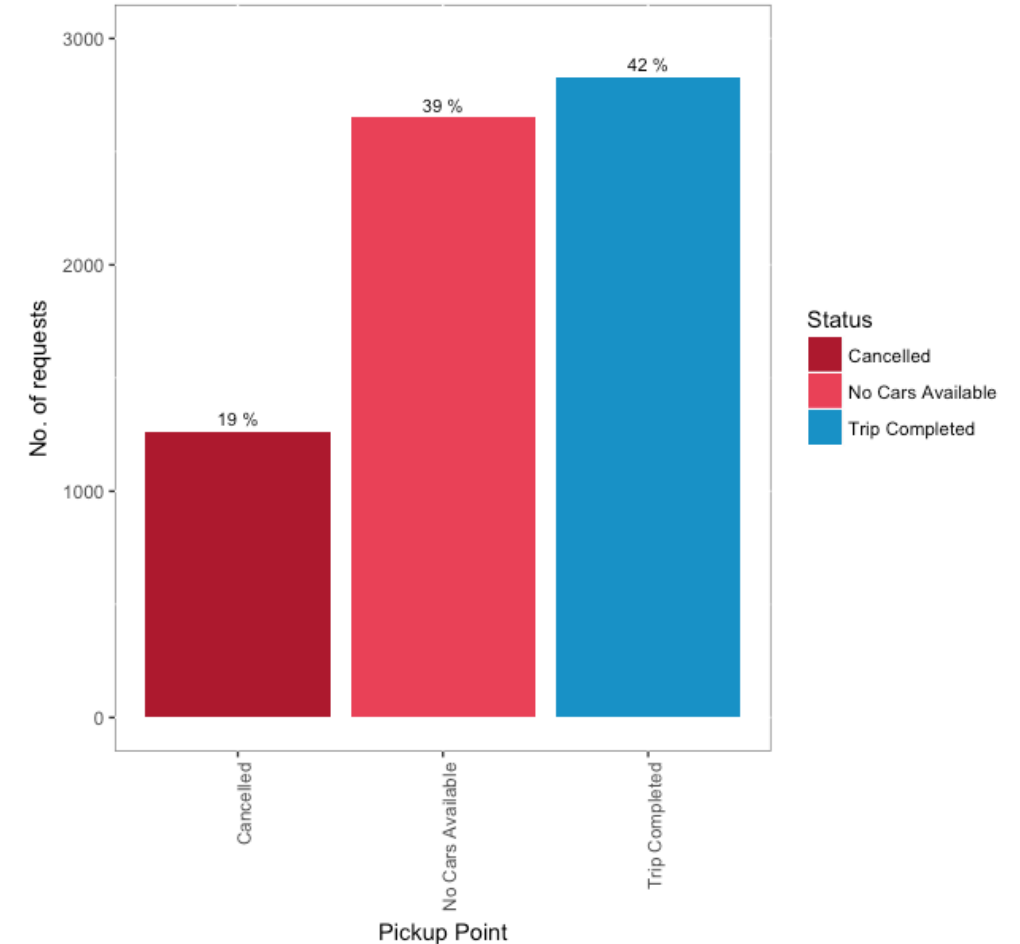
Uber case study for requests between 11July2016 - 15July2016

Overall request status - Airport/City pickup points

a) Only 42% of overall request is fulfilled.

b) 39% of requests are not fulfilled due to unavailability of cars.

c) 19% of requests are not fulfilled due to cancellation of request by car drivers.

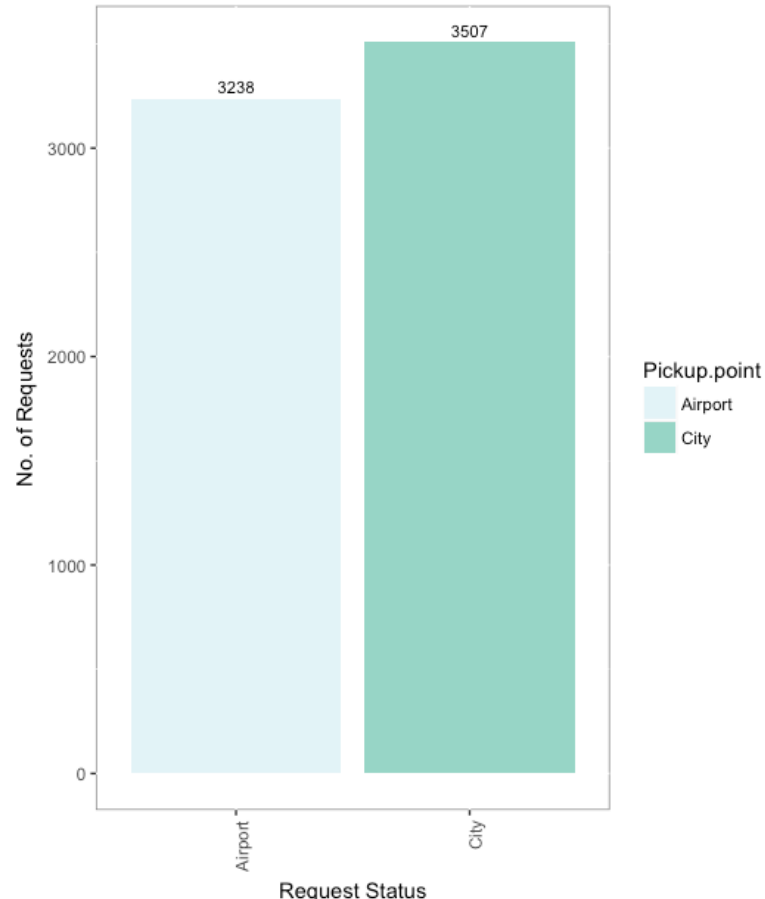


Uber case study for requests between 11July2016 - 15July2016

Univariate Analysis - Pickup Points

Overall City/Airport Status Distribution

It appears only 42 % of the requests are fulfilled and the rest are either cancelled or are not fulfilled due to unavailability of cars.



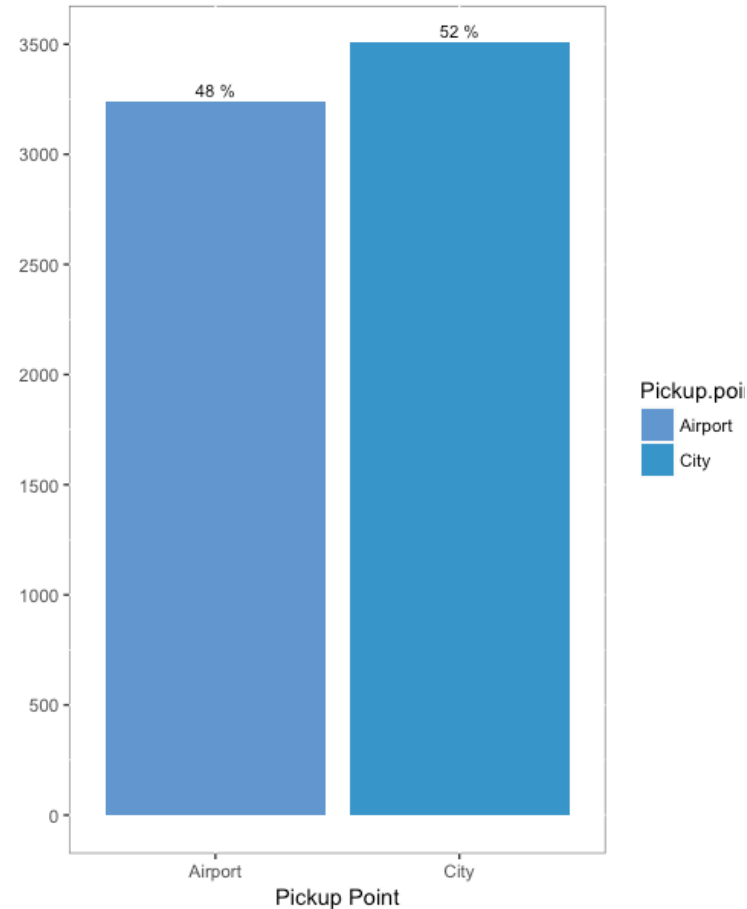
Uber case study for requests between 11July2016 - 15July2016

Overall City/Airport Pickup - Request Distribution

Nearly, there is equal amount of requests flowing for City/Airport pickups

a) Airport pickup points received 48% of total requests

b) City pickup points received 52% of total requests



Uber case study for requests between 11July2016 - 15July2016

Key Points

When performing univariate analysis on the column 'Pickup Points', the following was observed,

1. From the total 6745 requests, 3507 requests come from the City which is around 52%.

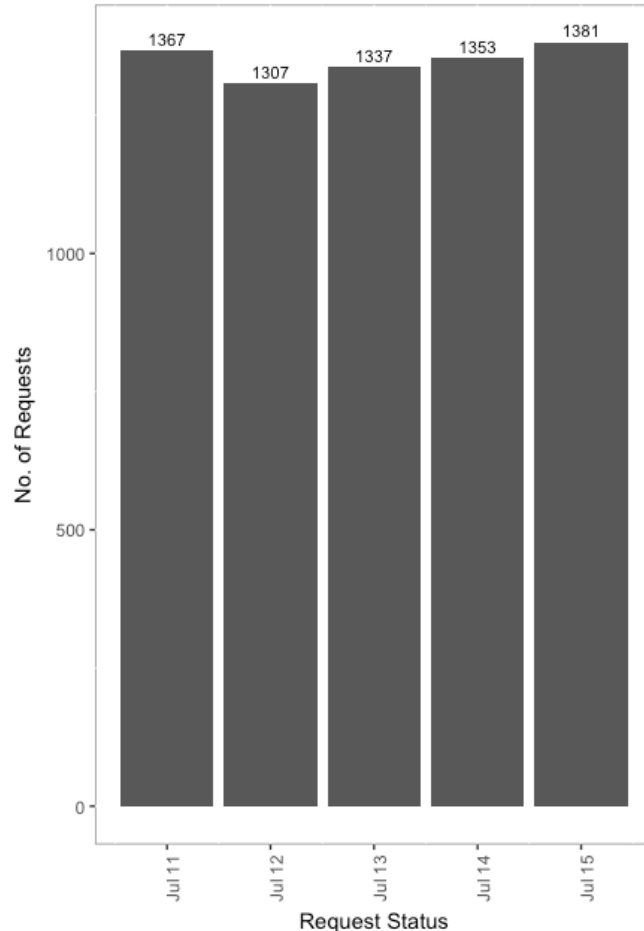
2. From the total 6745 requests, 3238 requests come from the Airport which is around 48%.

There seems to be almost equal demand from City and Airport

Univariate Analysis - Request Dates

Overall City/Airport Status Distribution

It appears all days have almost equal demands for cabs.

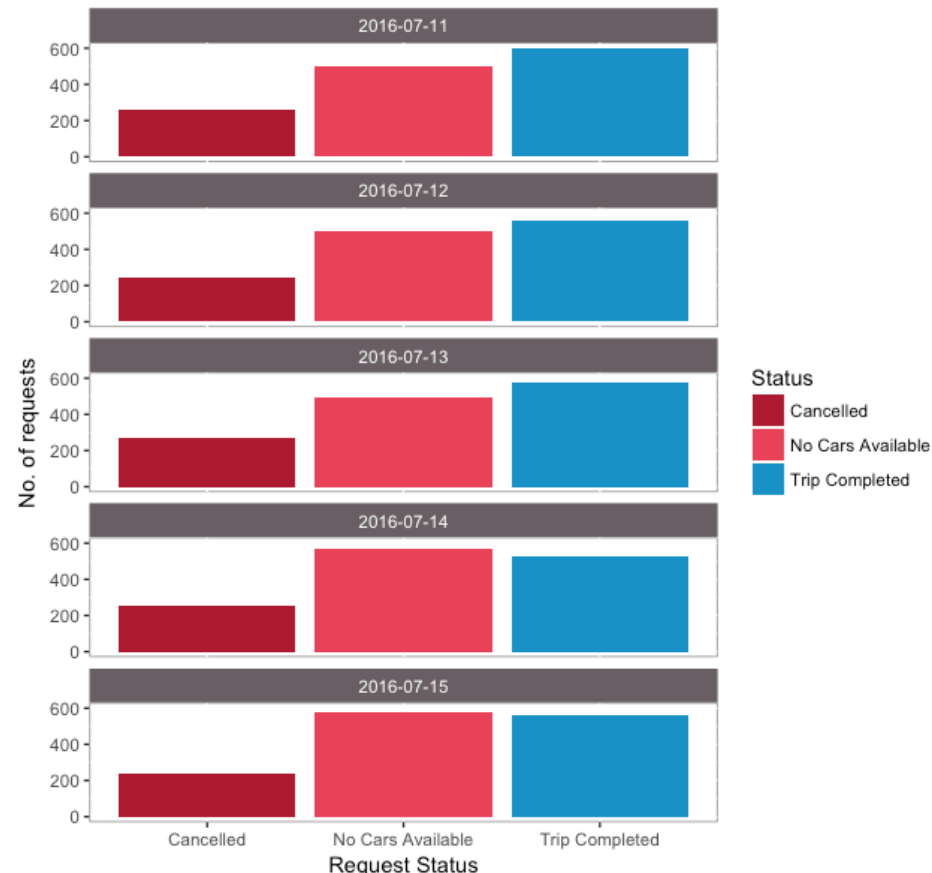


Uber case study for requests between 11July2016 - 15July2016

Overall request status - From Pickup points City/Airport

There seems to equal distribution of requests against various statuses on all days

- a) As shown only ~40% of requests are fulfilled
- b) Unavailability of Cars is the major concern on all days
- c) Cancellation is another major concern affecting the revenue for Uber



Uber case study for requests between 11July2016 - 15July2016

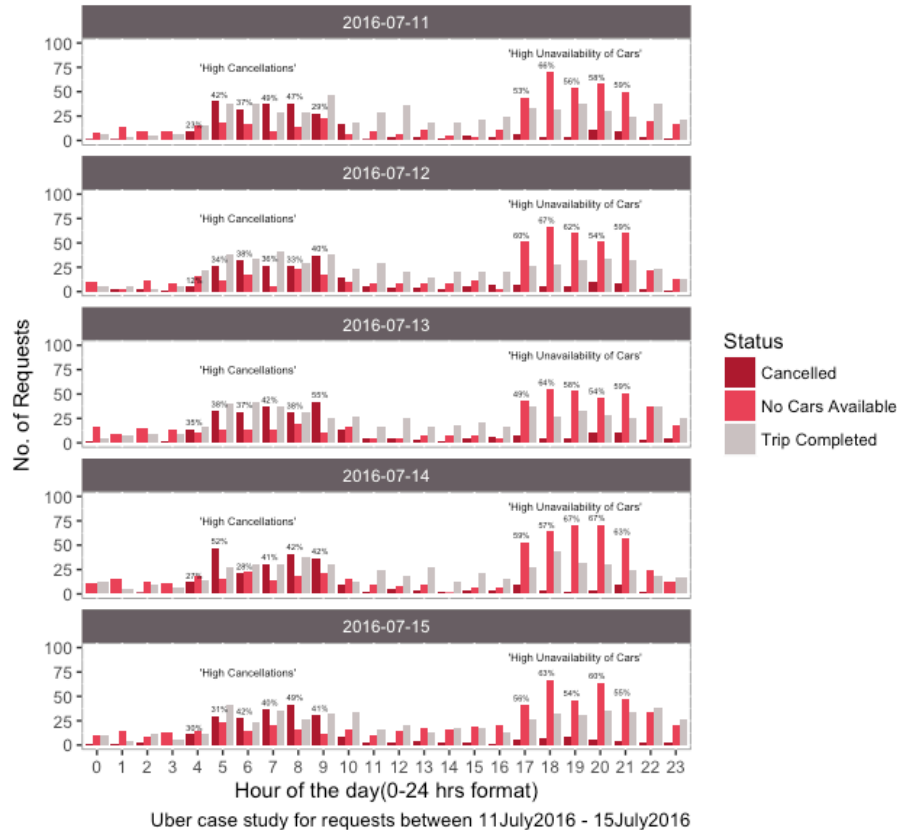
Key Points

When performing univariate analysis on the column 'Request date' which was derived from the column 'Request timestamp' it is observed that all days from 11th July to 15th July have almost equal demand for cabs.

It is also observed that the statuses also follow a similar pattern on all days

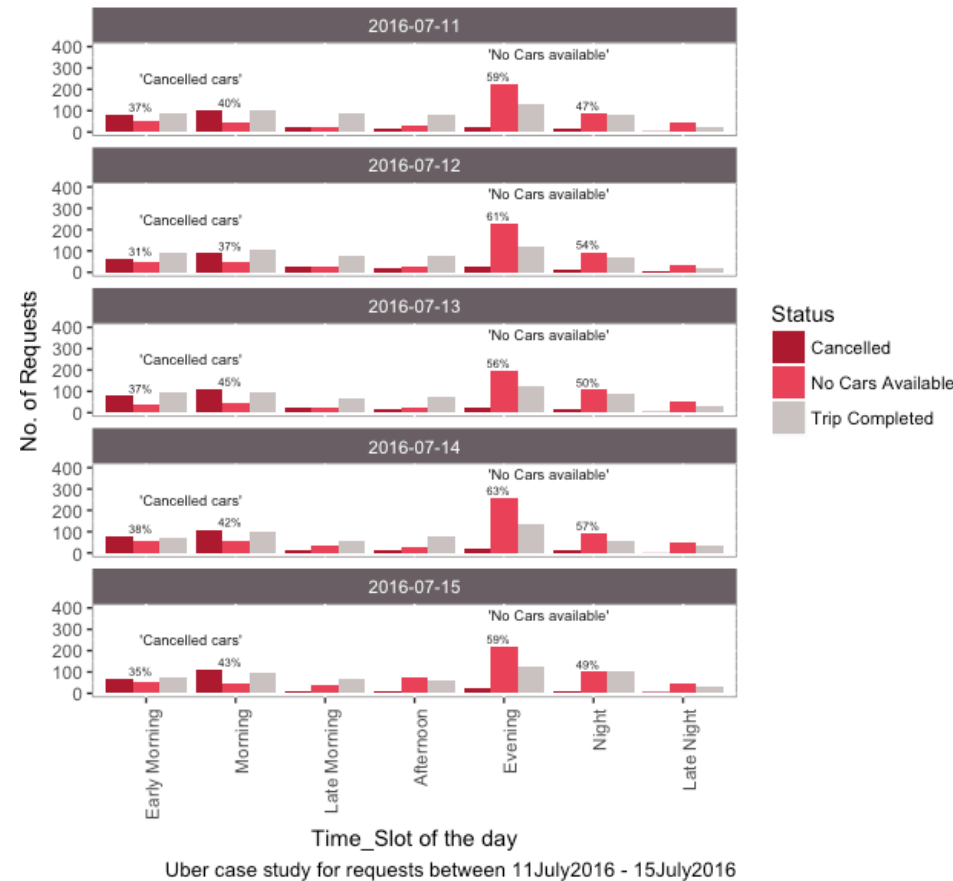
Overall Analysis of requests - hourly basis across all dates

- a) Between 5AM-9AM there are a lot of cancellations.
Peaks between 4AM-9AM - Annotated with Percentage of Cancellations at a particular hour as "High Cancellations"
- b) Between 5PM-11PM there are a lot of unavailability of cars.
Peaks between 5PM-9PM - Annotated with percentage of unavailability of cars at a particular hour as "High Unavailability of cars"



Overall Analysis of requests - Time_Slot basis across all dates

- a) Early morning and Morning relatively shows a lot of cancellations. Annotated and percentage of cancellations highlighted in the graph as "Cancelled cars"
- b) Evening and Night shows relatively a lot of unavailability of cars. Annotated and percentage of unavailability of cars highlighted in the graph as "No Cars available"



Key points to note

From the two plots it is observed that,

1. More cancellations of car requests are observed during 'Early Morning' and 'Morning' hours i.e 4AM - 9AM. This pattern is observed on all days.

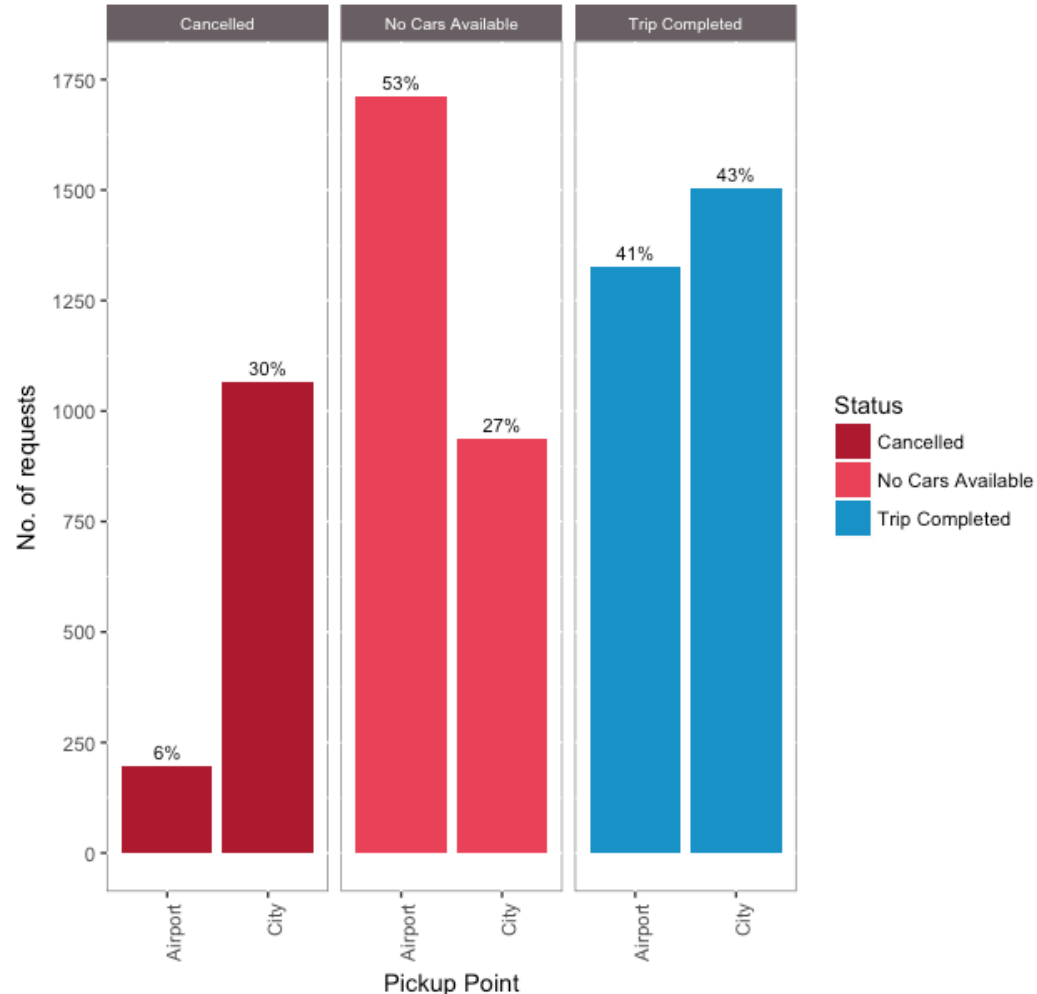
2. A lot of unavailability of cars are observed during 'Evening' and 'Night' hours i.e 5pm-11pm.

Please see the annotations added in the plots highlighting the problem area 'cancellation' and 'unavailability of cars' with percentage of the respective statuses.

Note : Percentage only pertaining to problems such as 'cancellations' and 'unavailability of cars' are shown in the plots.

City/Airport Pickup Point - Percentage Distribution

- a) Airports suffers from 53% of 'Unavailability of Cars'.
- b) Cities suffer from 30% of 'Cancellation' of requests.
- c) Roughly, only ~40% of Airport pick-up/drop requests are fulfilled.



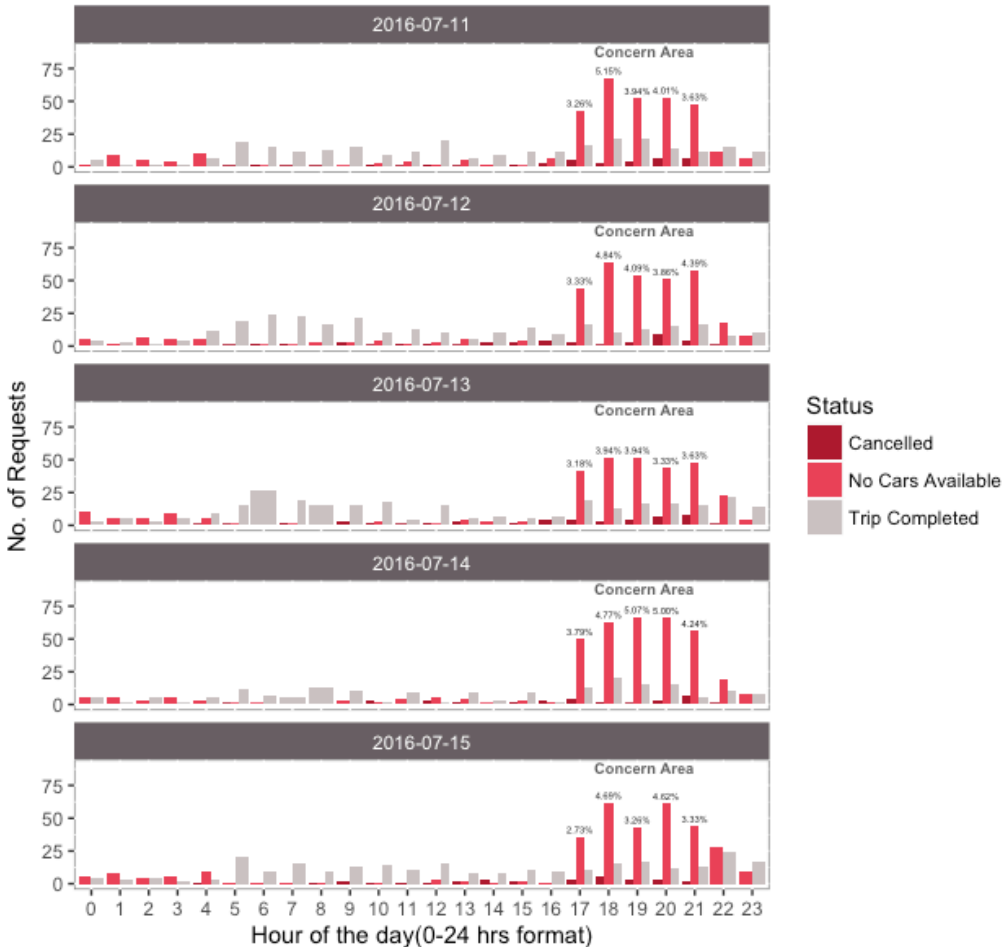
Uber case study for requests between 11 July 2016 - 15 July 2016

When the status of the request is analysed against each pickup point ie City and Airport, it is understood that the pickup points have different problems from each other,

1. **Airport** seems to suffer from more **unavailability of cars** compared to the City.
2. **City** seems to suffer from more **cancellations** of cars compared to Airport.
3. There is less than 50% of requests that are fulfilled in both City and Airport.
4. There is enough scope for Uber to generate revenue by addressing the issues in Airport and City.

Hourly Analysis of requests - Airport Pickup point

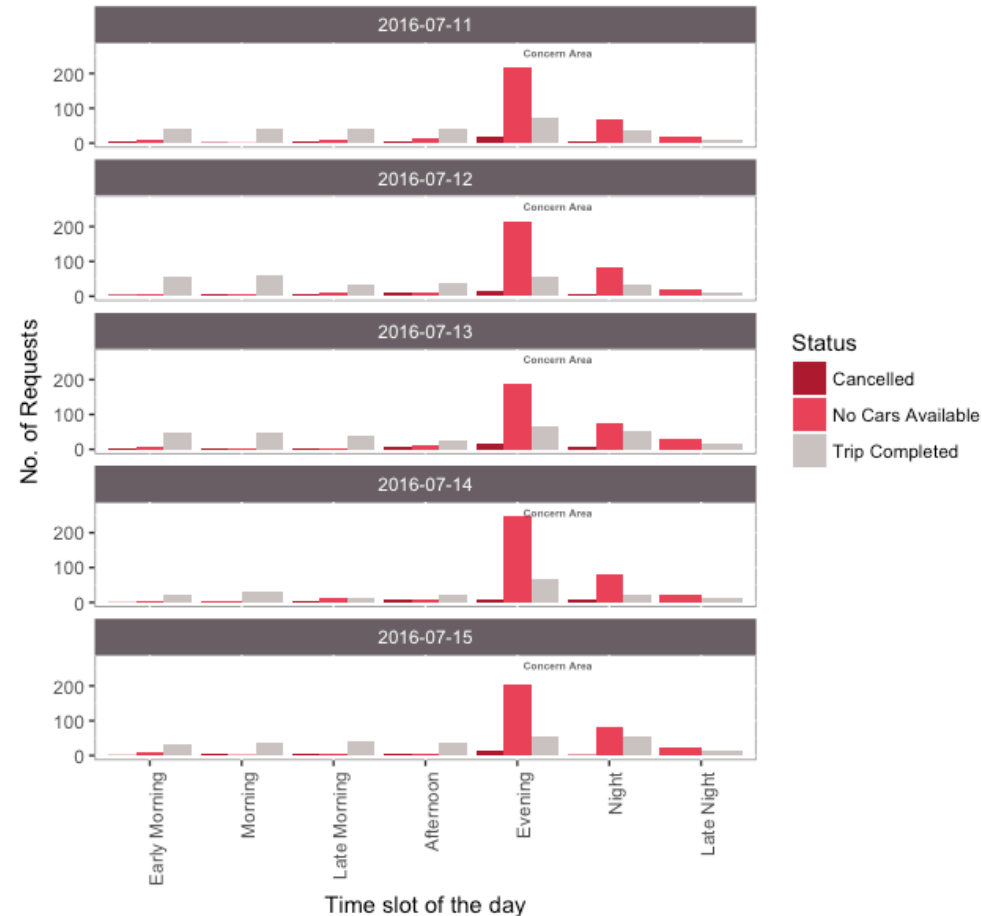
- a) Between 5PM-11PM there are a lot of unavailability of cars (Peaks between 5PM-9PM, Annotated as "Concern Area")
- b) There are a very few cancellations.
- c) A very few requests are fulfilled for the customers



Uber case study for requests between 11 July 2016 - 15 July 2016

Time_Slot wise Analysis of requests - Airport Pickup point

- a) Evening and Night show a lot of unavailability of cars (Peaks during Evening, Annotated as "Concern Area")
- b) There are a very few cancellations.
- c) A very few requests are fulfilled for the customers



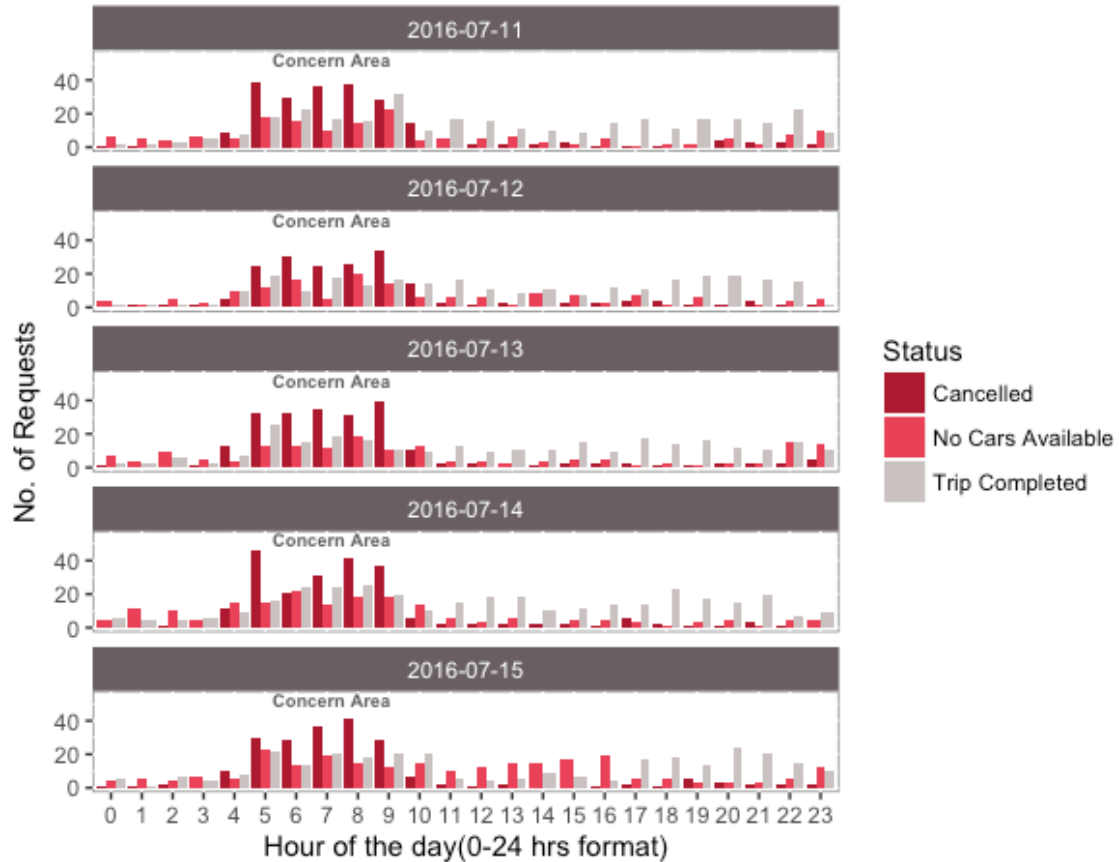
Uber case study for requests between 11 July 2016 - 15 July 2016

The plots give a view of cab requests made from Airport. The high bars seen in the plot across hours and time slot clearly indicates there is not enough cabs available in the Airport during Evening and Night I.e between 5PM-11PM to meet the demand. (refer to the annotation in the graph).

By addressing this issue Uber can generate more income by 100% as unavailability of cars in airport during evening and night hours account to nearly 50% of loss on Uber's revenue.

Hourly Analysis of requests - City Pickup point

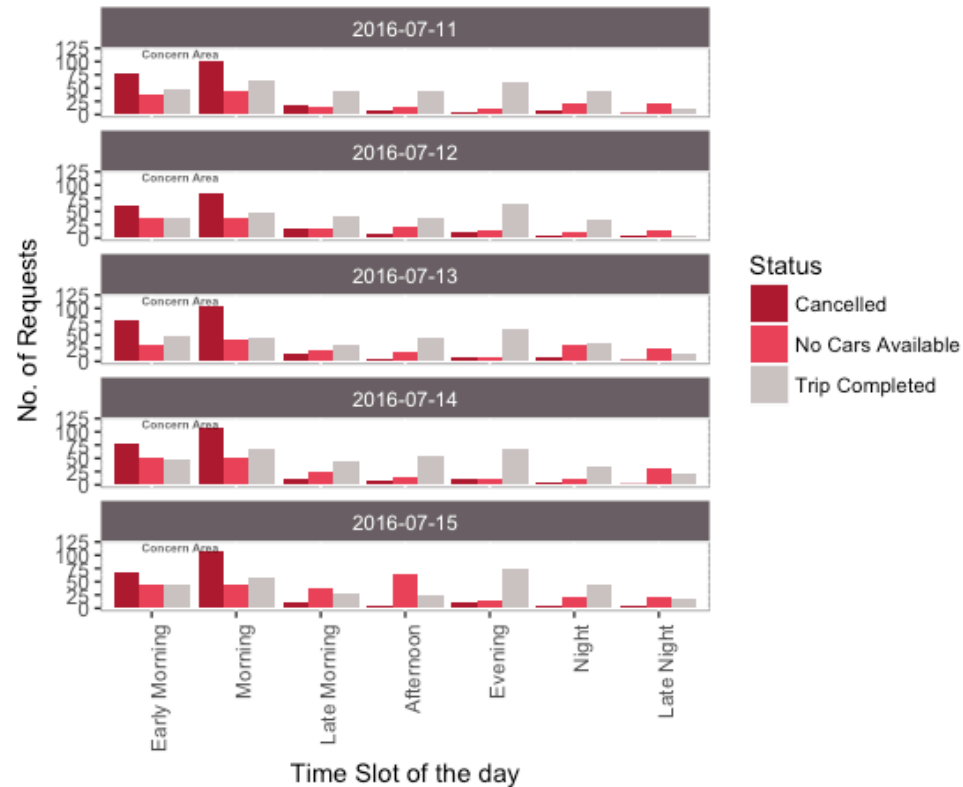
- a) Between 5AM-9AM there are a lot of cancellations. (Annotated as "Concern Area")
- b) Between this time there are a unavailability of cars as well which is relatively lesser than cancellations.
- c) Some requests are fulfilled for the customer.



Uber case study for requests between 11July2016 - 15July2016

Time_Slot wise Analysis of requests - City Pickup point

- a) Early morning and Morning show a lot of cancellations. (Annotated as "Concern Area")
- b) Between this time slot there are a unavailability of cars as well which is relatively lesser than cancellations.
- c) Some requests are fulfilled for the customer.



Uber case study for requests between 11July2016 - 15July2016

Key Points

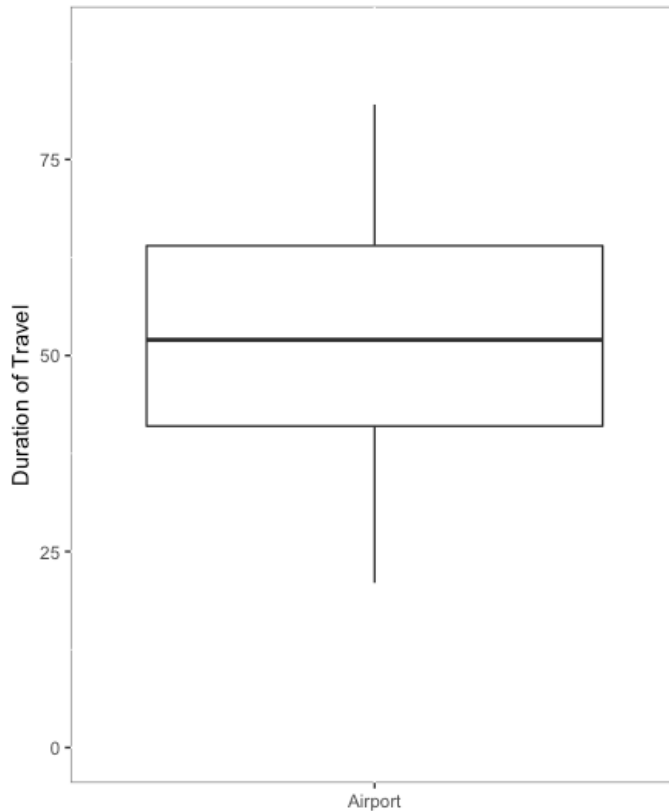
The plots give a view of cab requests made from City. The high bars seen in the plot across hours and time slot clearly indicates that the cab drivers cancel cab requests heavily during the 'early morning' and 'morning' hours i.e. between 4AM - 9AM (refer to the annotation in the graph).

By addressing this issue Uber can generate more income by 30% as Cancellation of cars in City during Early morning and morning hours account to nearly 30% of loss on Uber's revenue.

Duration of travel from Airport to City

It appears the median journey from City to Airport is 52 Mins.

Min : 21
1st Quartile : 41
Median : 52
Mean : 52.24
3rd Quartile : 64
Max : 82

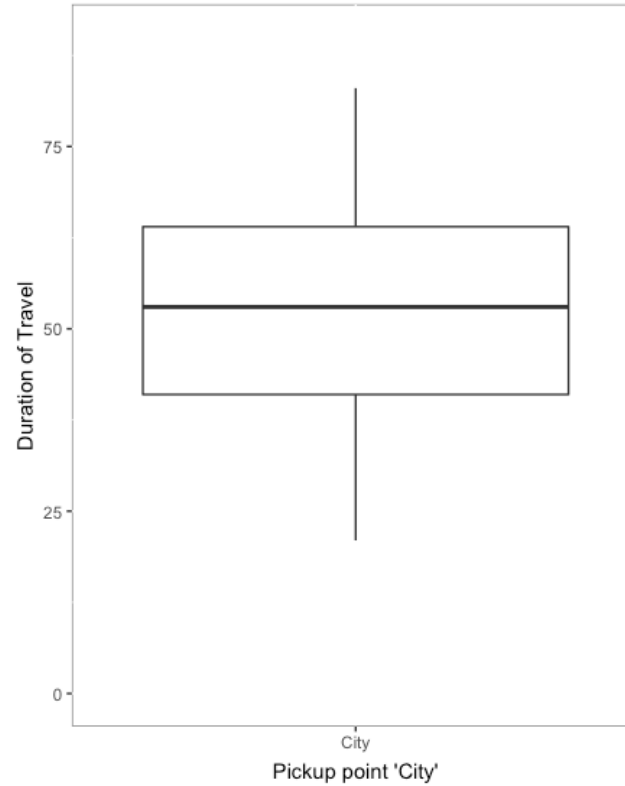


Uber case study for requests between 11July2016 - 15July2016

Duration of travel from City to Airport

It appears the median journey from City to Airport is 53 Mins.

Min : 21
1st Quartile : 41
Median : 53
Mean : 52.57
3rd Quartile : 64
Max : 83



Uber case study for requests between 11July2016 - 15July2016

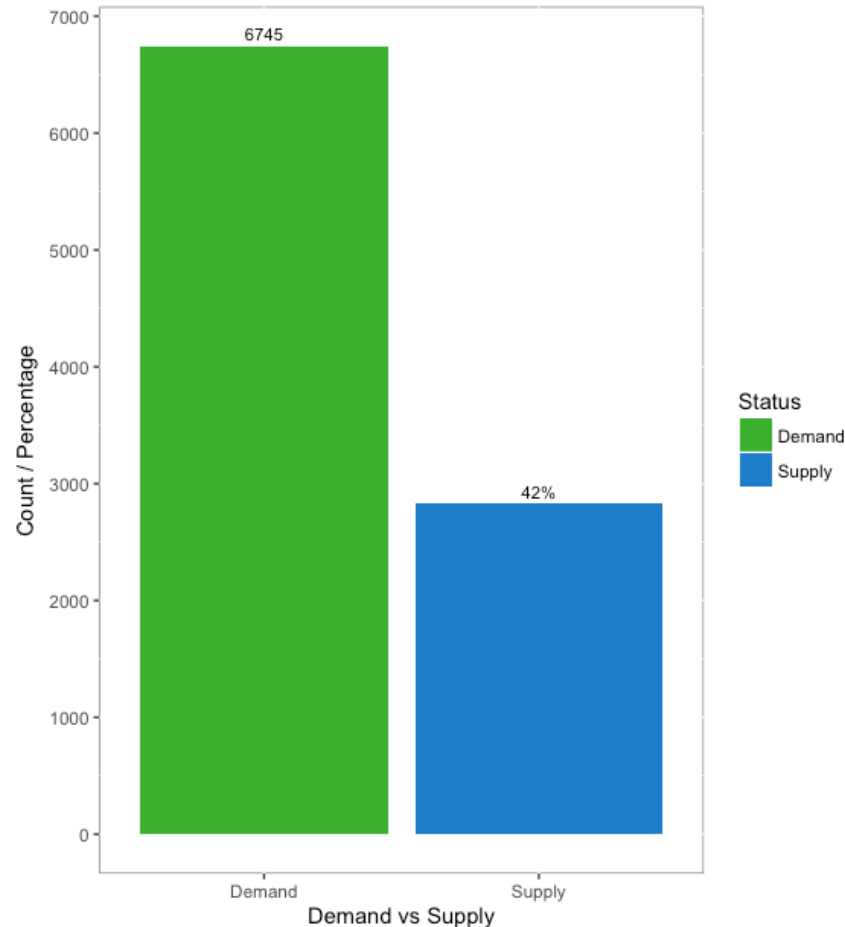
Key Points

It appears the duration of journey between City and Airport is around ~52 mins.

The duration is key to note because this might be one of the reason why the drivers cancel the requests and cabs not available in the Airport.

Overall Demand Vs Supply

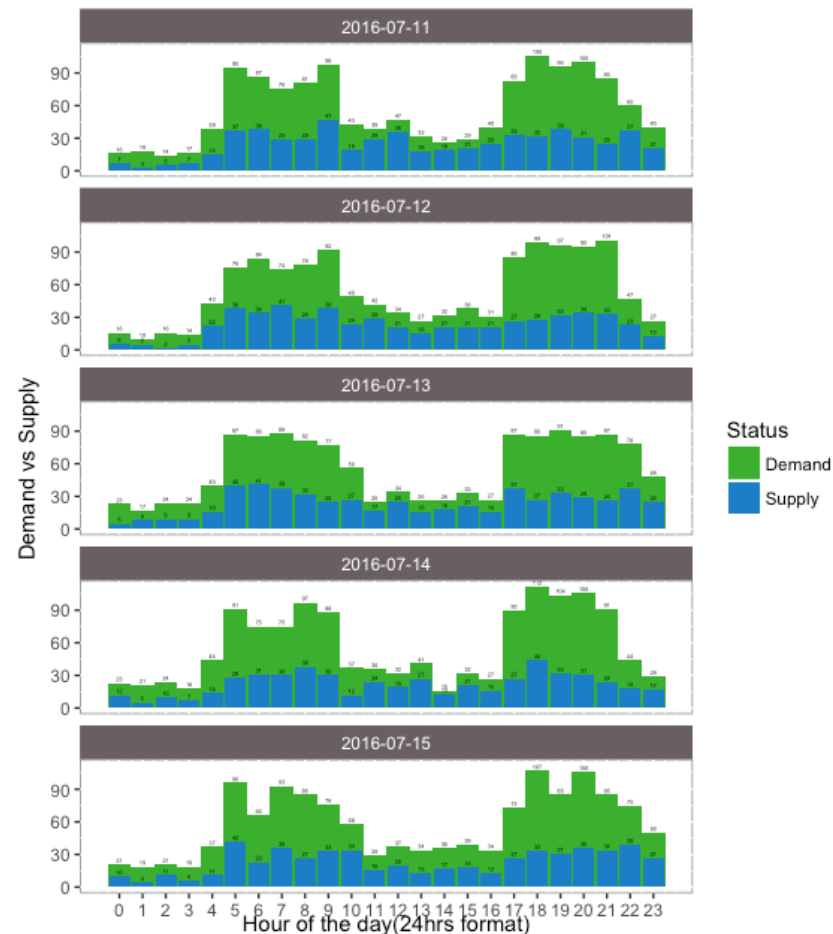
- a) Demand is twice than Supply
b) It appears only 42% of the demand is supplied



Uber case study for requests between 11July2016 - 15July2016

Demand Vs Supply - Hourly basis across all Days

Demand is twice than Supply



Uber case study for requests between 11July2016 - 15July2016

Key Points

Demand vs Supply plot clearly indicates that Uber has met only 42% of the total demand. Which means it needs to catch up with 58% of the demand. This is an opportunity for Uber to increase its revenue by 100%

The demand vs Supply plot across all days shows a similar pattern where more than 50% of the demand is not met

The time slots when the highest gap exist is during Early morning, Morning i.e nothing but (4AM-9AM), Evening and Night hours i.e nothing but (5PM-11PM)

From various analysis performed ,it is understood that,

- a)City to Airport and Airport to City has two different issues.Where City to Airport suffers high cancellations of cars and Airport to city suffers high unavailability of cars.
- b)The average duration of journey between City and Airport is around 50 mins which requires considerable amount of fuel.
- c)It also appears that the demand for cars from city is very high during early morning and morning hours (4AM-9AM) from City and Evening and Night(5PM-11PM) from Airport. However rest of the day do not show this high demand.
- d)It clearly appears there are more outgoing flights during early morning and morning hours, and more incoming flights flights during Evening and Night.
- e)Which means cabs going from City to Airport during Early morning and morning hours will not get a return trip back as there are barely few incoming flights.So, most cab drivers do not opt to fulfil requests during these hours as coming back to city empty will cost them causing less profit.
- f)As less cars go to the Airport, their availability is less during the evening and night hours.Due to which almost 50% of the request are not fulfilled.

Uber needs to address two problems,

1.High cancellations of cabs from City,

As drivers do not get a return trip without a long wait time, drivers do not prefer going to Airport during morning hours.However this is a considerable loss for Uber .Uber has a potential to increase its revenue by 100%.In order to achieve this,

- a)The trip cost during morning hours (peaks hours between 4AM-9AM) from City to Airport should be doubled so the driver does not experience any losses when he has to come with an empty trip back from Airport to City.
- b)Encourage drivers to accept cab requests from City during peak hours by providing them bonus points.

2.Unavailability of cars in Airport,

As there are barely few cabs available in the airport during the evening, and night hours(5PM-11PM).Uber has a potential to increase its revenue by 100% by addressing the Demand Vs Supply gap in Airport during these hours.

- a)Increase the availability of cars during the peak hours .
- b)Charge higher for the trips from Airport to City during peak hours to compensate the cost of driving from City to Airport empty.
- c)Encourage drivers to be available in airport during peak hours by providing them bonus points.

END