UBER CASE STUDY

SUBMISSION

Akarshi Rastogi

Abstract

An uber case study is to understand the problem uber facing which may results into the loss of revenue.

We may have some experience of travelling to and from the airport and used Uber or any other cab service for this travel.

Did you at any time face the problem of cancellation by the driver or non-availability of cars?

Well, these are the problems faced by customers, these issues may impact the business of Uber. If drivers cancel the request of riders or if cars are unavailable, Uber loses out on its revenue

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 the analysis, we 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.

Problem solving methodology
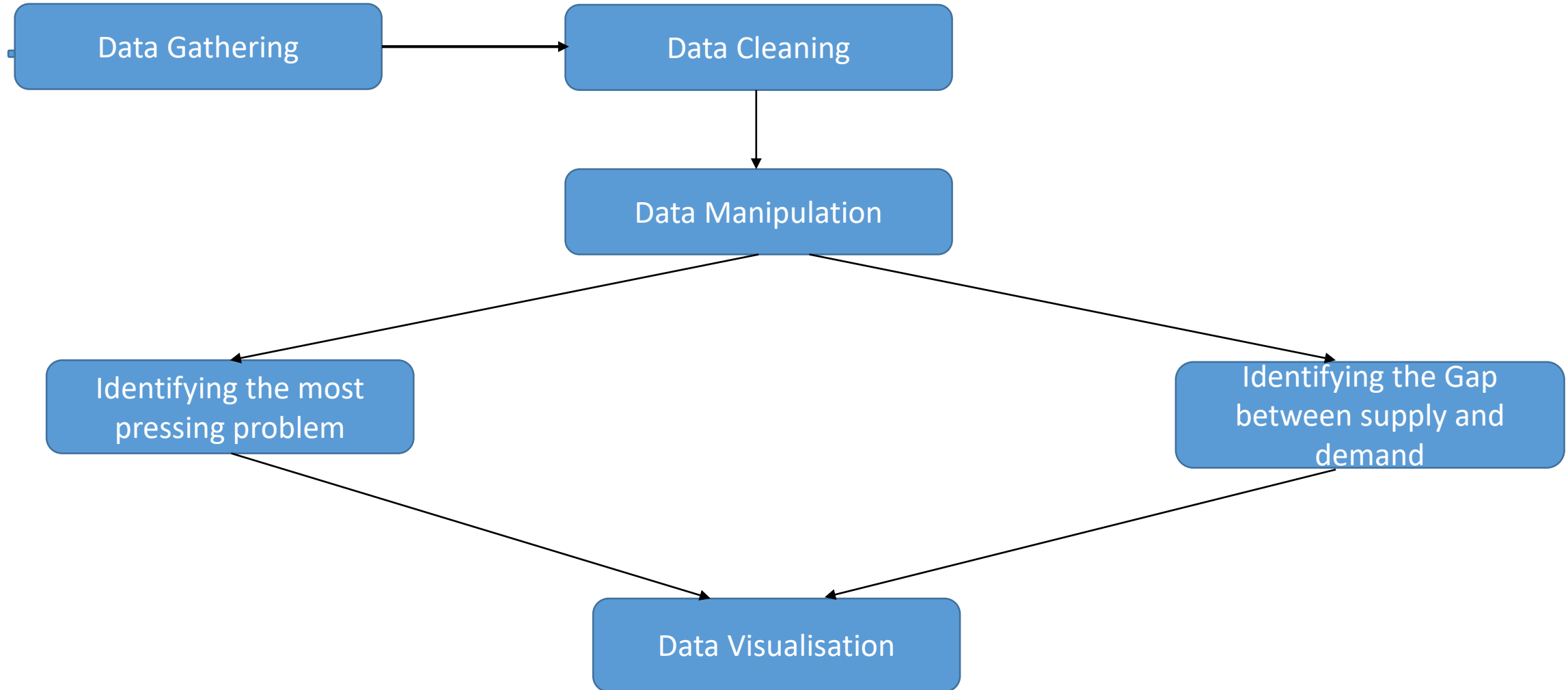
Analysis

To provide the best and profitable outcome, analysis is divided in further steps:

- **Data Gathering:** Data is categorise in file: Uber Request Data

- **Data Cleaning:**

    - the presence of different date and time format in the uber_request_data file in columns(Request timestamp and Drop timestamp). These are changed by using datetime function.

    - Null and Duplicate values have been removed from all the three data frames to reduce redundancy in the data.

    - Columns which contains the max NULL values are replaced by "None".

    - After data cleaning found that *6745* unique uber request data are present

- **Data Manipulation:**

    - Identifying the most pressing problem

    As we have seen as challenge while booking a uber cab to and fro from city and Airport. Many times the driver either cancelled the booking or show no car available

    In this case, we try to analyse the frequency or count of cabs status i.e. cancelled, no car available or trip completed

Analysis

From the plot between pickup point and status we analyse the below:

As we can see through above graph that count of `No Cars Available` from airport to city is near around 1700 which is far greater than that off city to Airport i.e 900. Also the count of `Trip Completed` from Airport is 1300 which is less than that off city i.e. 1500.

And for `cancelled`, the count is 200 for Airport and 1100 for city.

Cancelled + No car available = 1700 + 200 = 1900(Airport)

Cancelled + No car available = 1000 + 1100 = 2000(City)

 - As the `Cancelled + No car available` count is higher for city than airport hence city to airport is the problematic request

Then we divide the request time into different time slots: Early Morning, Morning, Afternoon, Evening, Late Evening and Night and plot the graph between timeslots and status.

From it, we can see that as per the time slots defined the frequency of the status i.e. can cancelled, trip completed and No cars available. As in the morning the count of cancelled and trip completed is max than rest of the day. Hence we can say that in this duration of day, the demand and supply is high.

Identifying the Gap between supply and demand:

Supply is defined as the total number of request or trip completed whereas Demand hold all the values whether cancelled, no car available or completed. We create new dataframe from the original with status, timeslots and the count.

Create a table of frequency of status along time slots :

Analysis

The timeslot when the highest GAP exists is the "MORNING" i.e. between 6AM to 11AM as we can see that the bar of Gap on morning is max then the rest. Also the demand is high than supply which results into the max Gap

The type of request for which the Gap is most severe as per the timeslots is "City to Airport". As the value of Gap is high in morning for city rather than Airport analysed from the plot between pickup point , Gap and Time slots

Status Timeslots	Cancelled	No Cars Available	Trip Completed	Total Supply	Total Demand	Gap
Aftnoon	63	174	414	414	651	237
Early Morning	241	383	400	400	1024	624
Evening	102	649	510	510	1261	751
Late Evening	65	575	323	323	963	640
Morning	729	388	785	785	1902	1117
Night	64	481	399	399	944	545

Results
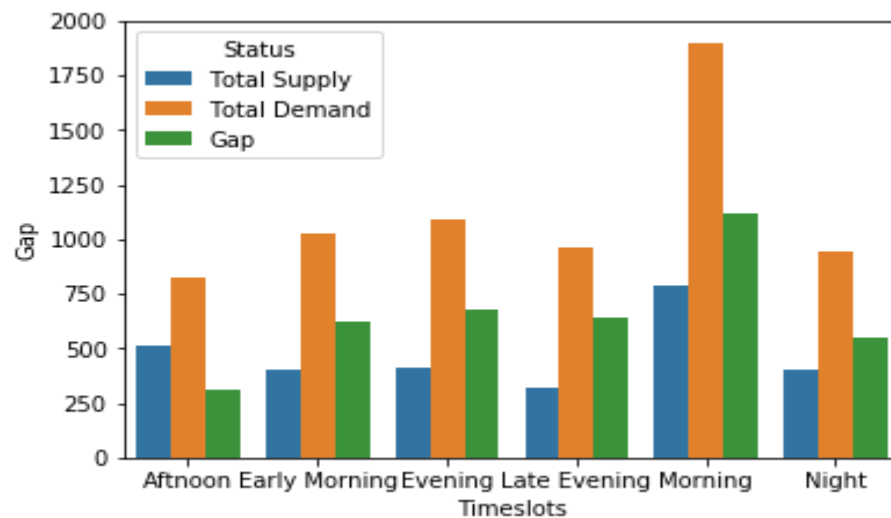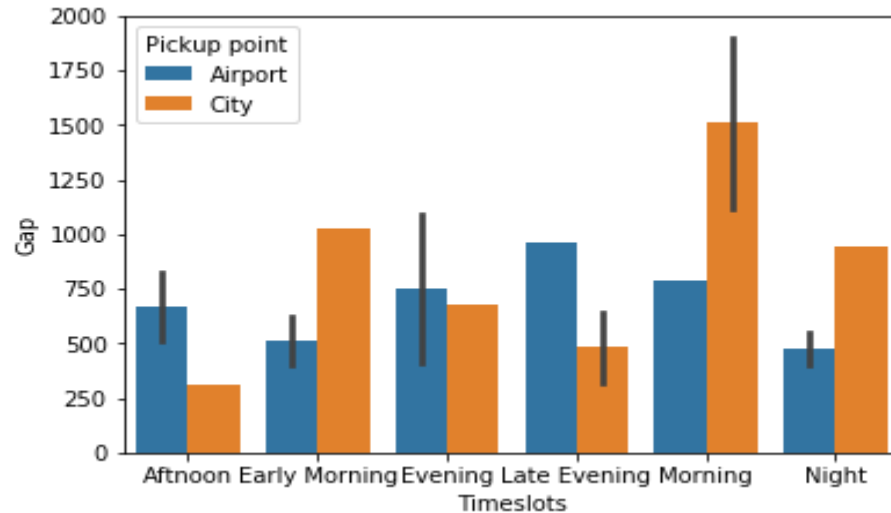
The reason for issue of supply-demand Gap is to analyse the time slot when the problem faced is maximum.
That means at which part of the day cabs are cancelled or no cars available at max.
Like through the graph present, we can analyse easily that during morning the Gap is maximum as the demand is more than the supply.
Also through it, we came to know that we face this problem maximum while searching for the cab in the city to airport.

Results

Some Ways to solve the supply demand gap:

As the problem faced is maximum during the morning so there can be few reasons for it:

- The traffic during morning (6AM to 11AM) is maximum as its peak office hour due to which cabs have to wait a long in the traffic.

- The cab drivers uses OLA and UBER together and accept the request which comes first. Hence it results into the unavailability in other.

- The hired driver must not have enough knowledge of all the roads leading to the same destination, so that if one is busy or max traffic then other way can be used.

- The driver preferred to accept the request of nearby rather than far, as through it they can complete many rides instead of few.

- The cabs may be just completed its one ride of long distance and don't want to go on another.

Results

Solutions can be:

- Drivers must be allowed to use only one application at a time that is either Uber or OLA

- They must be fixed minimum number of request from City to Airport or vice versa that driver have to accept I, before it he is not allowed to cancel or the few credit amount will get debited.

Once the min number of request completed then only he is authorised to cancel rest coming. And if he completes fixed minimum number of request then he must be rewarded.

For ex:

Min number of request = 3

Uber driver takes only 2 request from city to airport then on 3rd coming request if he cancel it then a amount say 50 will be deducted from his wallet. And this amount will debited till he accept third request.

And after third request, he is allowed to cancel 4th without paying any amount.

- There must be fix number of night shift and day shift that driver have to under go.

Conclusion

So as per the analysis we performed, we came to know that on Morning timeslot the Gap is maximum for city to airport as the demand frequency is higher than the supply frequency.

The timeslot when the highest GAP exists is the "MORNING" i.e. between 6AM to 11AM as we can see that the bar of Gap on morning is max then the rest. Also the demand is high than supply which results into the max Gap

The type of request for which the Gap is most severe as per the timeslots is "City to Airport". As the value of Gap is high in morning for city rather than Airport analysed from the plot between pickup point , Gap and Time slots