

# Uber Supply-Demand Gap

A case study with visualization

By Puneet Agarwal

# Business Objective

- To identify the root cause of the problem of cancellation and non-availability of Uber cabs and suggest possible solutions to alleviate the problem.

# Data for Analysis

- There are six attributes in the data:
  - Request id: unique id of each request
  - Time of request : date and time at which request was raised
  - Drop-off time : date and time at which trip was completed
  - Pickup point : location of making request i.e. Airport and City
  - Driver id : unique id of Driver.
  - Status of Request: The final status of the ride i.e. Trip Completed, Cancelled and No Cars Available.

The data is for the month of July in 2016. The dates range from 11<sup>th</sup> to 15<sup>th</sup> of July. The hours are in 24 Hours format and there are 6745 entries.

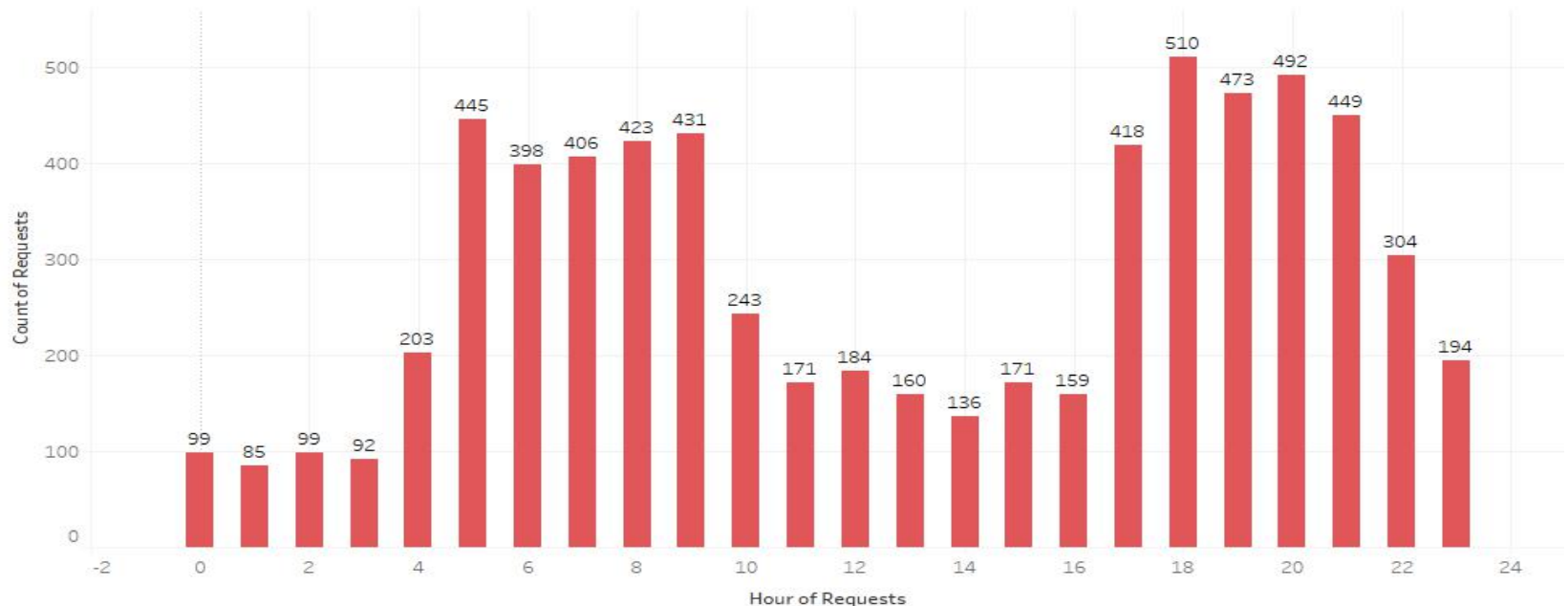
# Problem Analysis

- Getting the data
- Cleaning and properly formatting the data
- Dividing the hours into proper timeslot
- Visually identifying the most pressing problem for Uber on the basis of pickup location and status of requests.
- Finding gap between supply and demand using visualisation based on the timeslots, pickup location and status of trips.

# Result Analysis

- The most pressing problem for Uber :
- The plot between frequency of requests and hour of request shows two particular time-zone with high demand : 5-10 am and 17-22 pm.

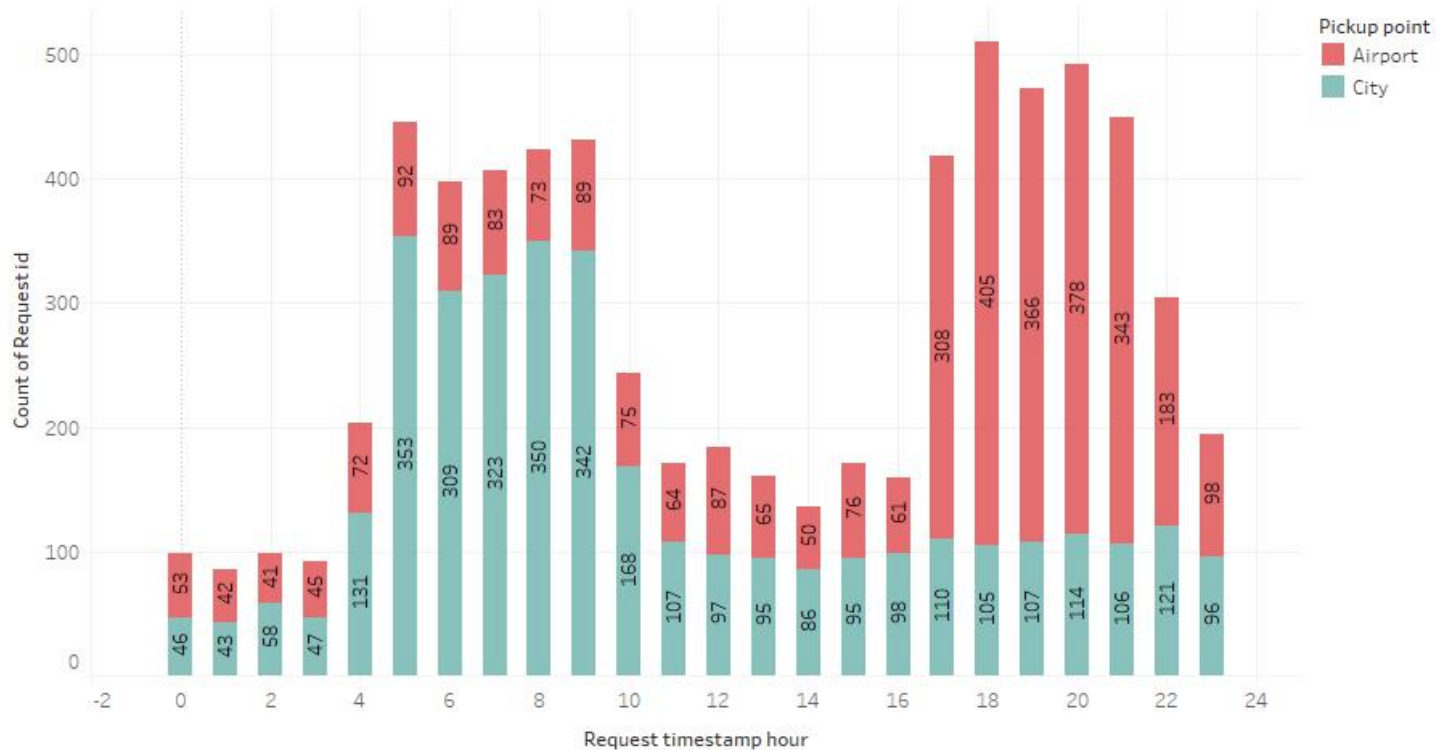
Requests made per hour



The plot of count of Request id for Request timestamp hour.

The plot shows the distribution of request for both Airport and City in each hour.

Frequency of requests from Airport and City made in each hour.

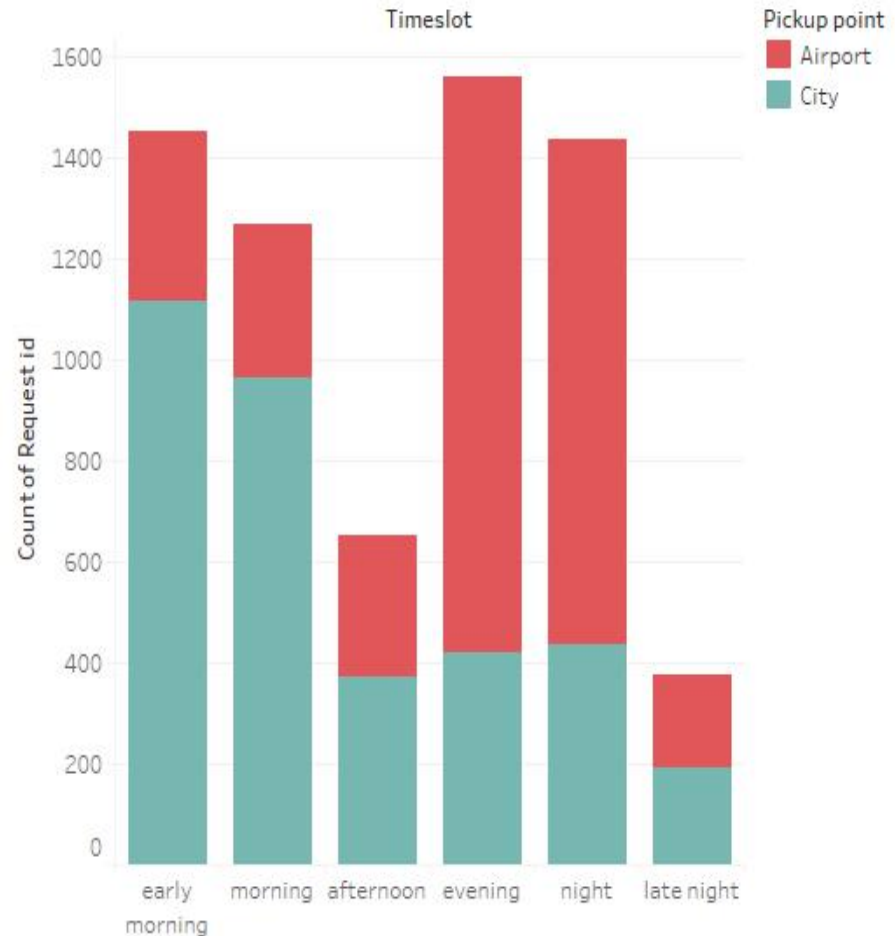


The plot of count of Request id for Request timestamp hour. Color shows details about Pickup point.

Most of requests made from city for airport are consolidated around 5-10 am and most of requests from airport to city are in the time range 17-22 pm

Most of requests made from city for airport are made around early morning and morning timeslot and most of requests from airport to city are in the timeslots evening and night

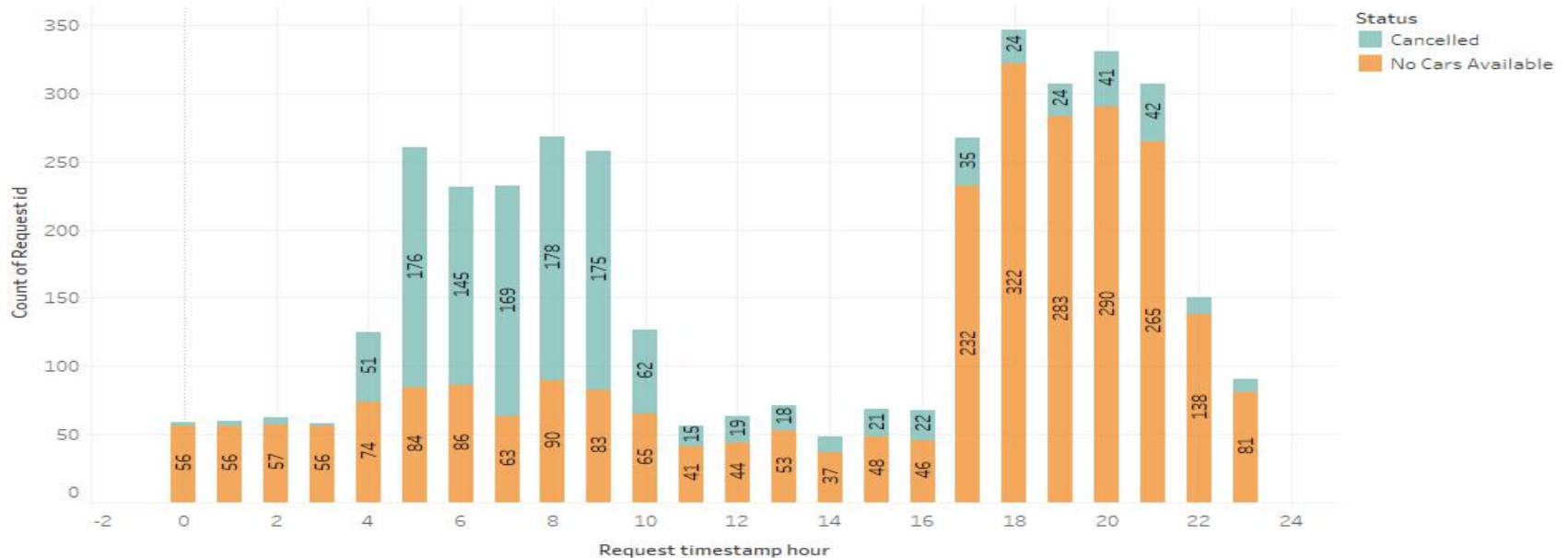
Frequency of requests from Airport and City made in each timeslot.



Count of Request id for each Timeslot. Color shows details about Pickup point.

The plot shows the distribution of request with both Cancelled and No cars available status in each hour.

Frequency of requests with Cancelled and No cars available status made in each hour.



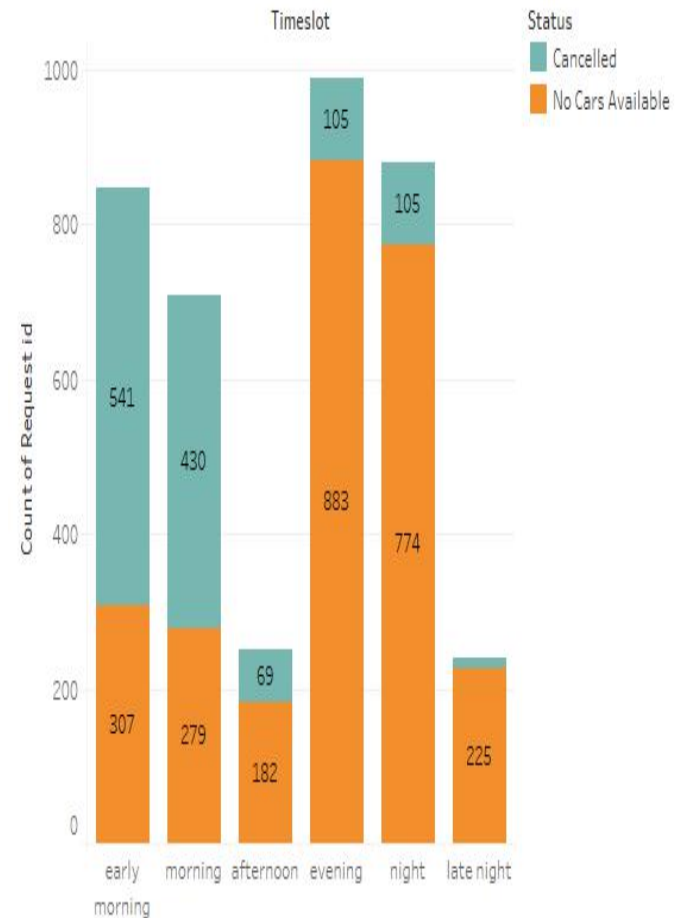
The plot of count of Request id for Request timestamp hour. Color shows details about Status. The view is filtered on Status, which excludes

Most of requests which are cancelled are consolidated around 5-10 am and most of requests showing no cars available are in the time range 17-22 pm



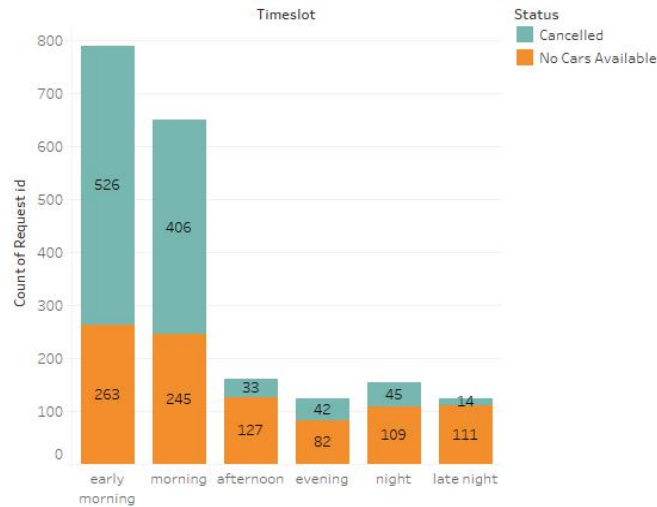
Most of the cancellation are seen in the time slots early morning and morning and the non availability of cars is mostly observed in the timeslot evening and night.

Frequency of Requests with status  
Cancelled and No cars available in different  
time slots



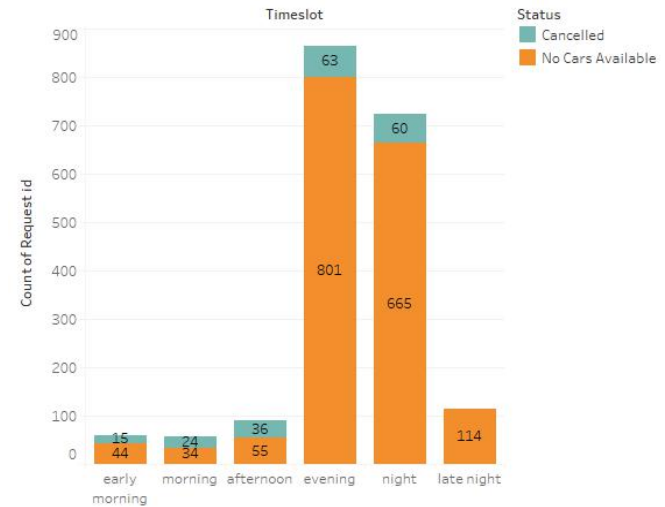
Count of Request id for each Timeslot. Color shows details about Status.  
The view is filtered on Status, which excludes Trip Completed.

Frequency of requests from City with status Cancelled/No cars available made in each timeslot.



Count of Request id for each Timeslot. Color shows details about Status. The data is filtered on Pickup point, which keeps City. The view is filtered on Status, which excludes Trip Completed.

Frequency of requests from Airport with status Cancelled/No cars available made in each timeslot.



Count of Request id for each Timeslot. Color shows details about Status. The data is filtered on Pickup point, which keeps Airport. The view is filtered on Status, which excludes Trip Completed.

From the plots above, it can be concluded that:

- A very high number of requests made from City for Airport are getting cancelled in the early morning and morning timeslots and
- A very high number of requests made from Airport for City are showing no cars available in the evening and night timeslots

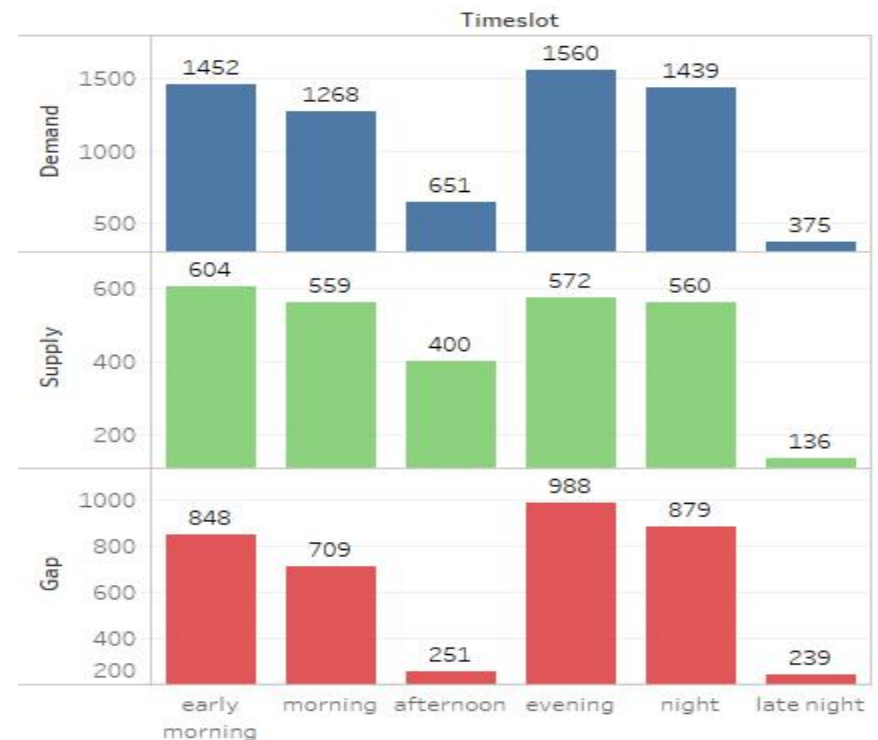
**And this is the most pressing problems for Uber as it can significantly affect the revenue of Uber.**

# • Supply-Demand Gap

- The total Demand is calculated by the requests having the status Trip Completed, Cancelled and No Cars available status
- The Supply is calculated by the requests having Trip Completed status.
- The Gap is the difference of the two.

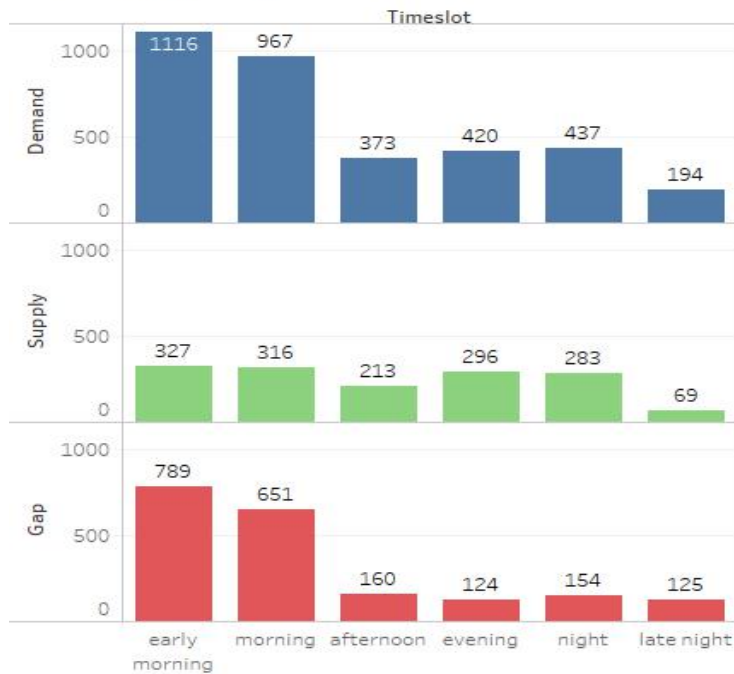
The highest gap exist in early morning, morning, evening and night timeslots. This is mainly because the number of requests are also high during these timeslots i.e Demand is very high and the cancellation and non availability of cars is even higher than supply.

Plot of total Demand, Supply and Gap for each timeslot



Demand (UberData.csv), Supply (UberData.csv) and Gap (UberData.csv) for each Timeslot.

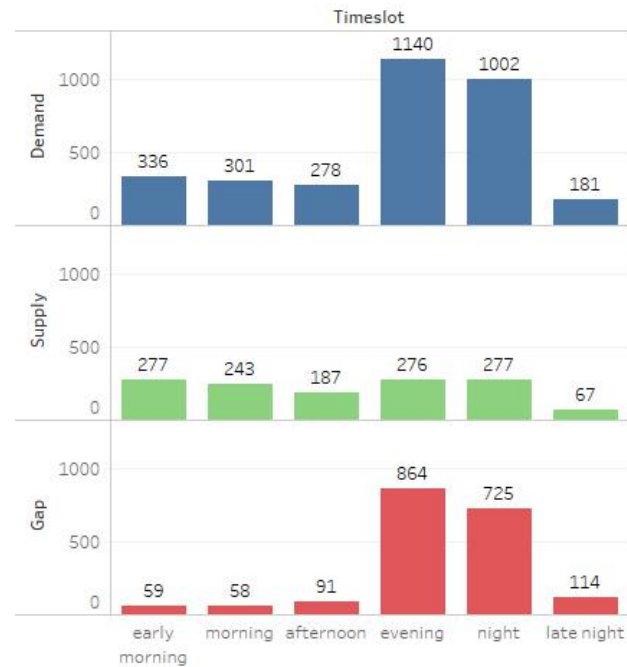
Plot of total Demand, Supply and Gap in City for each timeslot



Demand, Supply and Gap for each Timeslot. The view is filtered on Exclusions (Demand, Timeslot), which keeps 6 members.

For the slots early morning and morning the highest gap exists for the request type city-airport.

Plot of total Demand, Supply and Gap in Airport for each timeslot



Demand, Supply and Gap for each Timeslot. The view is filtered on Exclusions (Demand, Timeslot), which keeps 6 members.

For the slots evening and night the highest gap exists for the requests type airport-city

- There are four problematic timeslots early morning, morning, evening and night.
- For early morning and morning the requests from city are getting cancelled creating a huge gap between demand and supply of cabs from city to airport
- For evening and night the requests from airport are not getting any cabs assigned creating a huge gap between demand and supply of cabs from airport to city

# Possible Reasons behind Demand-Supply Gap

- Airports are mostly situated at the outskirts of cities. Going from city to airport may take considerable amount of time during which a driver can complete more rides and since Uber incentivize the drivers on the basis of number of trips it is beneficial from the perspective of driver to remain in the city.
- Most of the daily activities start in the morning hours and a driver is more likely to get more rides in the city during the same duration and from the plots it is evident that most of the requests from airport to city are in the evening, hence driver may have to wait for longer hours to get rides back to city. This may cause a loss to driver.
- Hence driver might prefer cancelling the ride.
- This is indirectly affecting the availability of cars at airport as drivers are not preferring to go from city to airport causing lack of supply at airport

# Possible solution for demand-supply gap

For the rides from city to airport :

- Uber can incentivize drivers taking ride from city to airport making the drivers financially satisfied.
- There should be a limit for the number of requests that can be cancelled during peak hours.
- There should be a penalty for cancelling requests more than the limit.
- So there should be a risk and reward framework to reduce the gap in Demand and Supply from city to airport

For the rides from airport to city :

- The reduction in the gap between supply and demand from city to airport will indirectly reduce the non availability of cabs from airport to city.
- There should be a threshold for the driver to complete rides from airport to city in a week during peak hours and the incentive should be higher for such rides. This will force the driver to be available at the airport for the required number of times thus reducing the non-availability of cars.