



TAXI

A Case study on
Uber's Demand vs Supply



Business Problem

- **High cancellation or non availability** of cars for customers travelling to and from the airport.

Impacts

- If driver cancels the request of customers or if cars are unavailable, **Uber loses out on its revenue.**
- **Customer dissatisfaction**
- **Brand name losing value.**

Assumptions

- The drop point for a request made in city is assumed to be airport and vice versa.
- The given data is only for 5 days. Assumption is the sample is a good representation of the entire population of Uber requests.

The Data in hand

| | |
|-------------------------------|--|
| Description | Masked data set containing Uber cab request details |
| Source | UpGrad case study assignment |
| Format | .csv (comma separated values) |
| Number of observations | 6745 |
| Number of variables | 6 |
| Each row is | A cab request made by a customer |
| Sampling method | All cab requests made for 5 days (Days 11 through 15 in the month of June for the year 2016) |

| Column Name | Type | Description | Missing | Uniques | Top |
|-------------------|----------------|--|---------|---------|---|
| Request id | Number | A unique identifier of the request | 0 | 6745 | 619, 867, 1807, 2532, 3112, 3879, ... |
| Pickup point | Text | The point from which the request was made | 0 | 6745 | "Airport", "Airport", "City", "Airport", "City", "Airport", ... |
| Driver id | Number | The unique identifier of the driver | 3914 | 300 | 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, ... |
| Status | Text | The final status of the trip, that can be either completed, cancelled by the driver or no cars available | 0 | 3 | "Trip Completed", "Trip Completed", "Trip Completed"... |
| Request timestamp | Date timestamp | The date and time at which the customer made the trip request | 0 | 5 | "11/7/2016 11:51", "11/7/2016 17:57", "12/7/2016 9:17", ... |
| Drop timestamp | Date timestamp | The drop-off date and time in case the trip was completed | 2650 | 5 | "11/7/2016 13:00", "11/7/2016 18:47", "12/7/2016 9:58"... |

The Demand vs Supply



What is the **average demand** from airport to city and city to airport?

| | Airport | City |
|--------------|---------|------|
| Total demand | 3238 | 3507 |
| Percentage | 48% | 52% |



What is the **total # of Cabs/Drivers** available?

| Number of Cars/Drivers |
|------------------------|
| 300 |



What is the **expected average # of trips** a driver should complete per week in order to meet the demand we have

| Airport | City |
|---------|------|
| 11 | 12 |



Are there any hours during which there is **more demand than the available Cabs/Drivers**

| 5 AM | 6 AM | 7 AM | 8 AM | 9 AM |
|------|------|------|------|------|
| 445 | 398 | 406 | 423 | 431 |



What is the **current average # of trips** a driver completes per week?

| Airport | City |
|---------|------|
| 4 | 5 |

| 5 PM | 6 PM | 7 PM | 8 PM | 9 PM | 10 PM |
|------|------|------|------|------|-------|
| 445 | 398 | 406 | 423 | 431 | 304 |



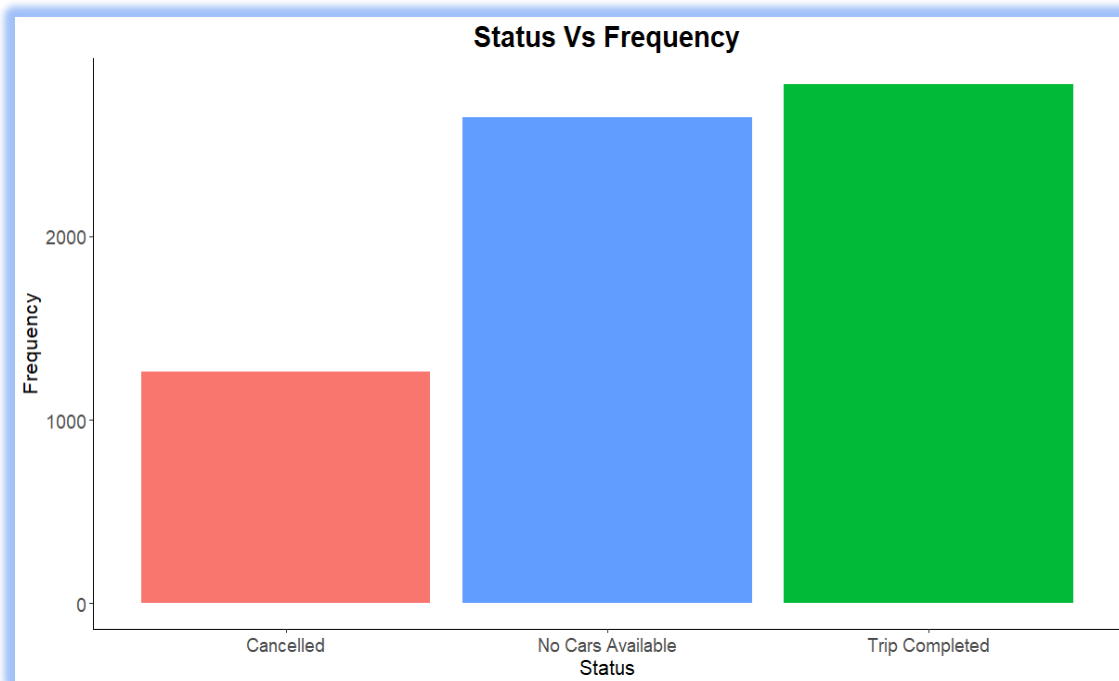
To meet the available demand, a driver should complete **at least 23 trips(11 Airport, 12 City) in a week (Mon - Fri)** between airport and city whereas the **current average is only 9 trips (4 Airport, 5 City)**.



With the available 300 drivers it wouldn't be possible to meet the peak demand. We would **need additional Cabs/Drivers** to meet the peak demand. If additional cabs/Drivers added, then the expected average seen on the left might become less.



How **many(or what % of) trips** got completed /cancelled or trips for which no cars where available?



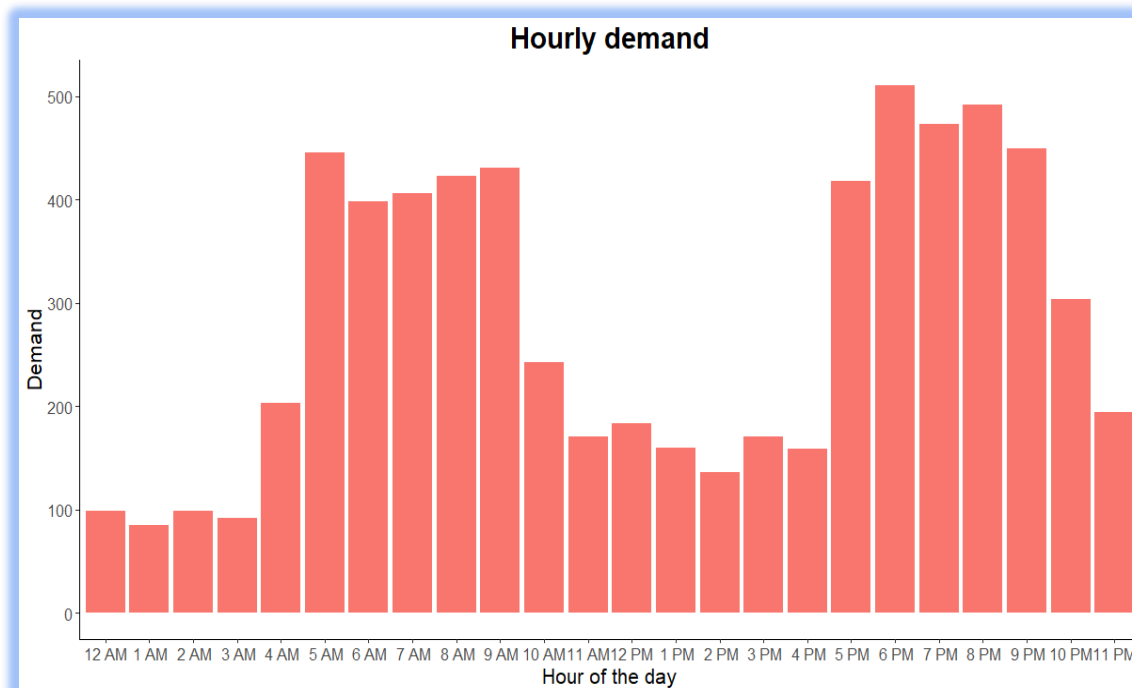
| Cancelled | No Cars Available | Trip Completed |
|-----------|-------------------|----------------|
| 19% | 39% | 42% |



- Less than 50%(~42%) of the trips got completed.
- For **~40%** of the trips, demand is not being met as no cars are available.
- **~20%** of the trips are being cancelled by the drivers



How does the **demand** look like for **each hour**?

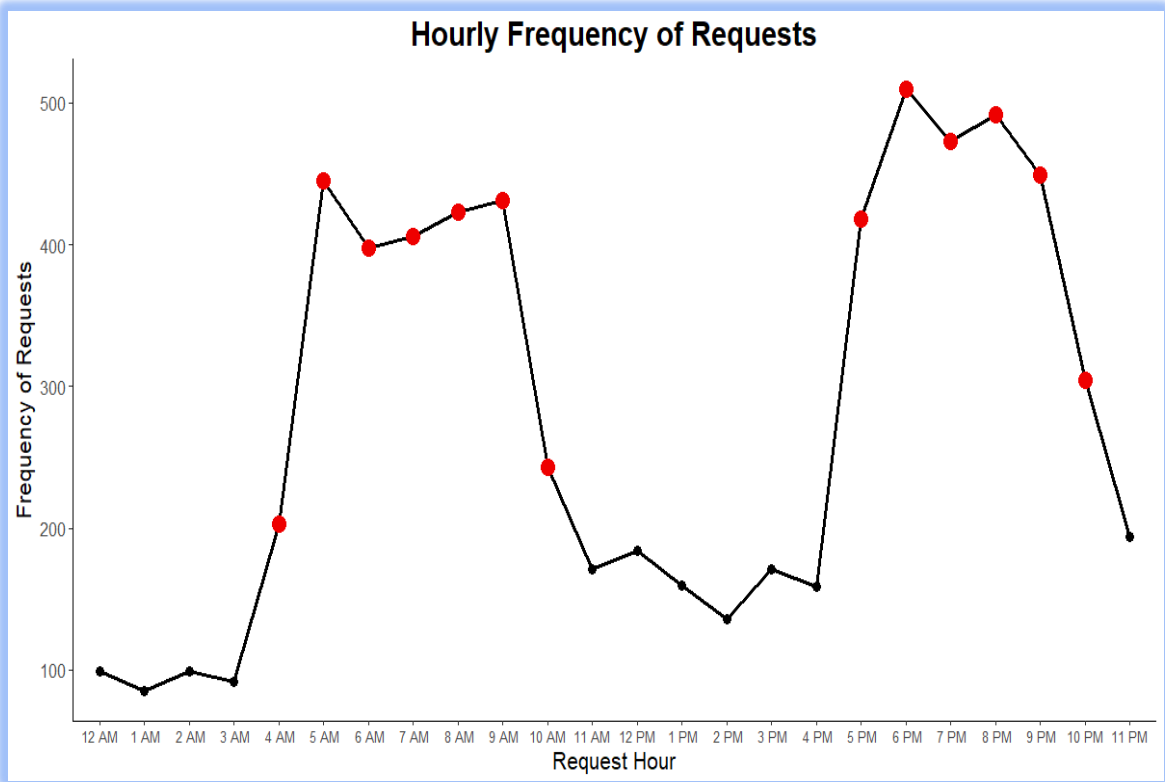


There are **2 spikes** in demand at certain hours of the day.

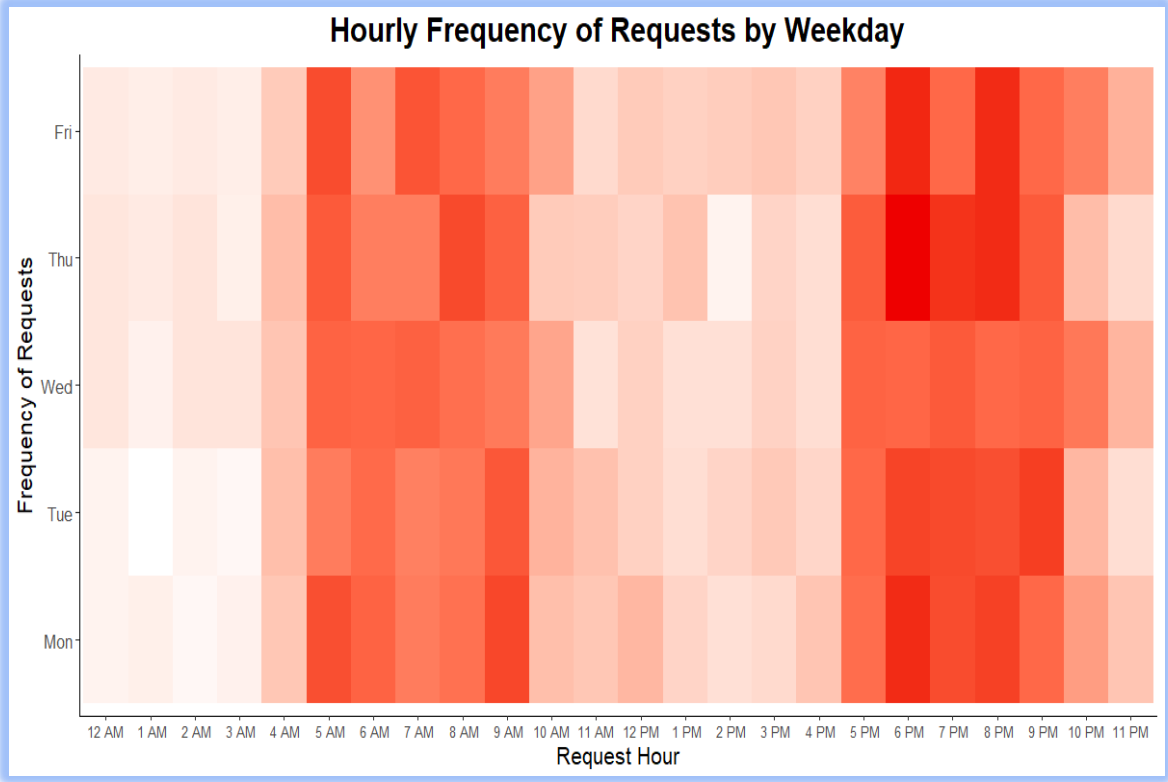
Identifying the Peak hours



What are the **peak hours** of the requests?



How does the **peak hours** look across weekdays?



Peak times : 4 AM to 10 AM & 5 PM to 10 PM across all weekdays



How does the proportion of various statuses look like during the peak hours?

| Request.hour | Cancelled | No Cars Available | Trip Completed | Total Demand |
|--------------|-----------|-------------------|----------------|--------------|
| 12 AM | 3% | 57% | 40% | 99 |
| 1 AM | 5% | 66% | 29% | 85 |
| 2 AM | 5% | 58% | 37% | 99 |
| 3 AM | 2% | 61% | 37% | 92 |
| 4 AM | 25% | 36% | 38% | 203 |
| 5 AM | 40% | 19% | 42% | 445 |
| 6 AM | 36% | 22% | 42% | 398 |
| 7 AM | 42% | 16% | 43% | 406 |
| 8 AM | 42% | 21% | 37% | 423 |
| 9 AM | 41% | 19% | 40% | 431 |
| 10 AM | 26% | 27% | 48% | 243 |
| 11 AM | 9% | 24% | 67% | 171 |
| 12 PM | 10% | 24% | 66% | 184 |
| 1 PM | 11% | 33% | 56% | 160 |
| 2 PM | 8% | 27% | 65% | 136 |
| 3 PM | 12% | 28% | 60% | 171 |
| 4 PM | 14% | 29% | 57% | 159 |
| 5 PM | 8% | 56% | 36% | 418 |
| 6 PM | 5% | 63% | 32% | 510 |
| 7 PM | 5% | 60% | 35% | 473 |
| 8 PM | 8% | 59% | 33% | 492 |
| 9 PM | 9% | 59% | 32% | 449 |
| 10 PM | 4% | 45% | 51% | 304 |
| 11 PM | 5% | 42% | 53% | 194 |



Peak times > 200 Requests

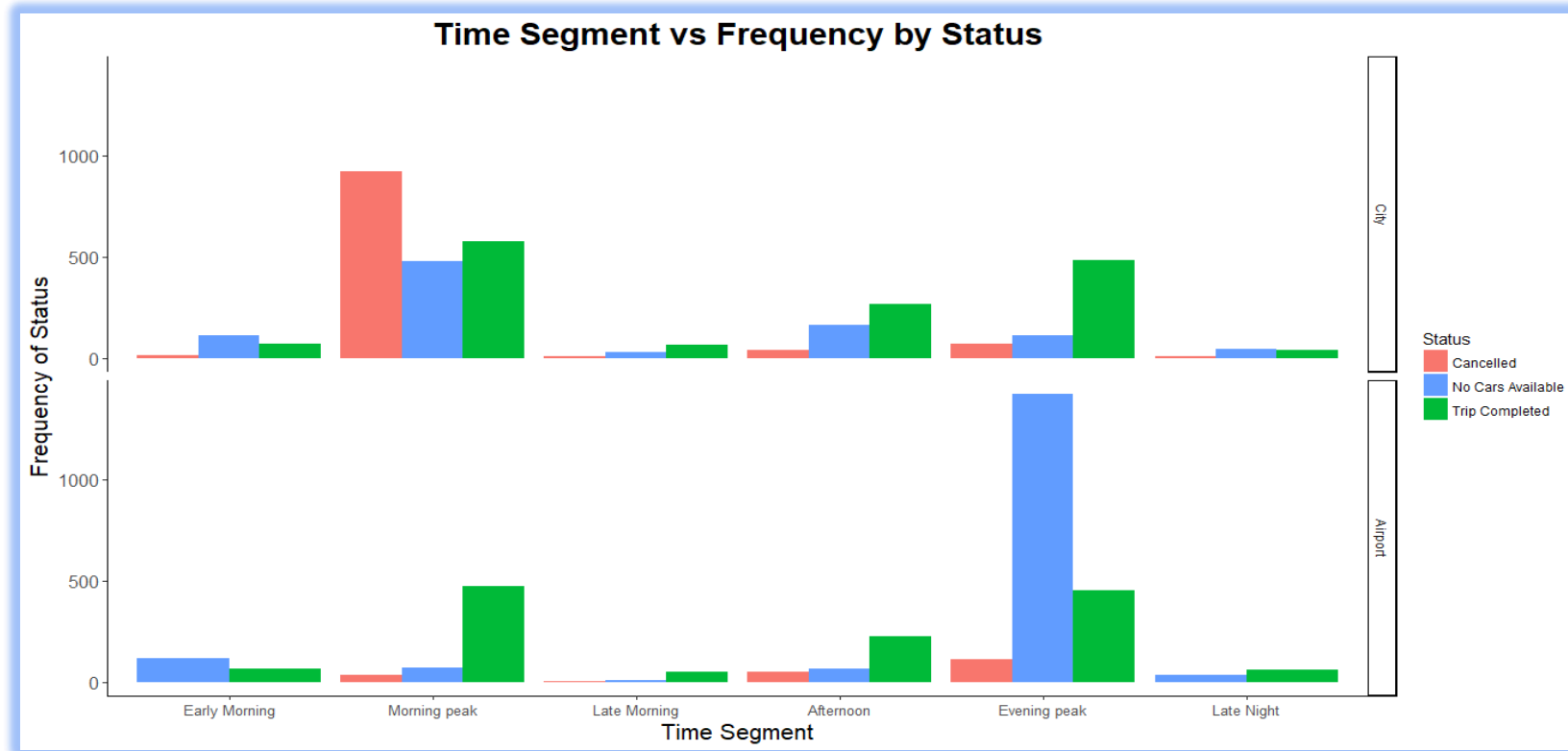
- > 25% Cancellations @ 4 AM to 10 AM (Morning peak time)
- > 45% unavailability of cars @ 5 PM to 10 PM (Evening peak time)

(Based on the total demand, we will focus our analysis on peak hours considering hours where total demand is >200)

How does status vary for different time segments especially during peak hours?



How does the demand vary for different time segments based on pickup points especially during peak hours?

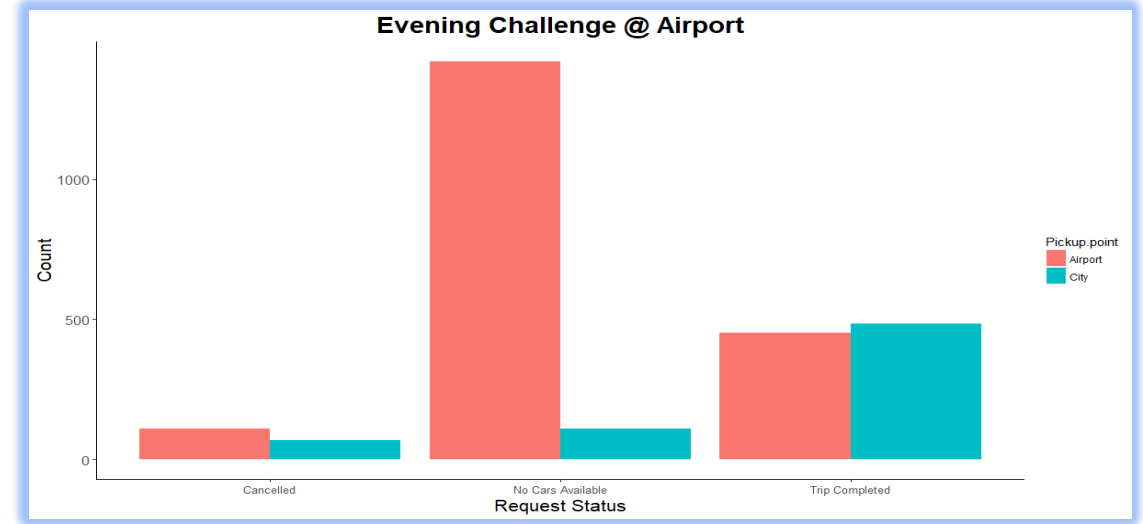
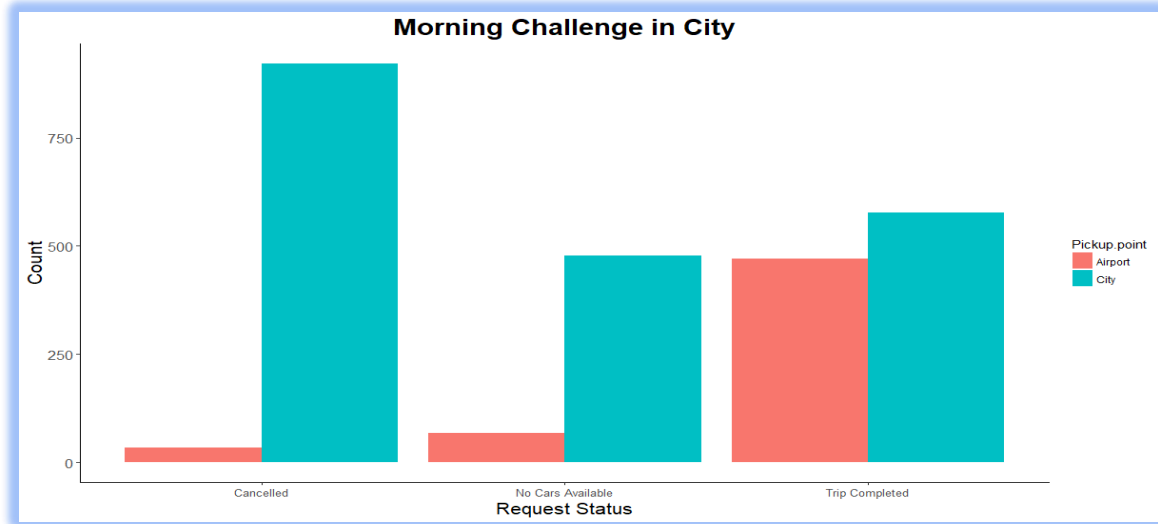


Challenge 1 : Morning peak hour window(4 AM to 10 AM) - Demand is higher in **City. Cab Cancellations** are higher.

Challenge 2 : Evening peak hour window(5 PM to 10 PM) - Demand is higher in **Airport. Cabs unavailability** is higher.

Early Morning : 12 AM - 5 AM | Morning peak : 4 AM - 10 AM | Late Morning : 10 AM - 11:59 AM
 Afternoon : 12 PM - 5 PM | Evening peak : 5 PM - 10 PM | Late Night : 09 PM - 11:59 PM

What does more (cancellations/unavailability of cars) mean? Can we quantify? How much are the numbers?



| Cancelled | No Cars Available | Trip Completed |
|-----------|-------------------|----------------|
| 47% | 24% | 29% |



No. of trips completed is only 29% and the cancellations are around 47%.

| Cancelled | No Cars Available | Trip Completed |
|-----------|-------------------|----------------|
| 6% | 72% | 22% |



No. of trips completed is only 22% whereas the cab unavailability is around 72%.

What is the expected average & actual average # of trips a driver takes in a week during the Morning/Evening peak hours



Expected average trips during morning and evening peak hours is **7** (or less if we add additional cabs/drivers) **per week** (5 days Mon – Fri) whereas the current average trips is only **2 per week** (5 days – Mon - Fri)

Challenges in City:

During Morning peak hour timeslot in the city:

- Average trips completed is only 2 per week whereas expected is around 7 per week
- Demand more than the available Cabs/Drivers.
- **47%** cancellations by the drivers.

Possible causes :

- More outbound flights from the airport during Morning peak hours
- Demand to reach airport from the city is more.
- Very little inbound flights at the airport.
 - **A driver who is taking a pickup to the airport at this time may not have a customer to pickup from the airport back to the city or may have to wait a considerable amount of time to get a return pickup.**
 - During this idle time he might make money by doing other trips in the city. Hence the driver prefers not to go to the airport and cancels the trips

Proposed Solution :

- **Peak hour charges (1.25x, 1.5x, 1.75x or 2x times the base price based on the hour demand)** for the customers.
- Use the additional income from the peak hour charges
 - **to provide Weekly Incentives** for drivers who take trips from city to airport during the morning peak hours.
 - **To compensate** for the idle time at the airport.
- Levy **cancellation charges** to drivers if they cancel more than 2 times in a week during the peak time slots.
- Add additional Cabs/Drivers to meet the peak hour demand.
- No. of trips based incentives if the driver completes more than the expected average # of 7 trips per week (5 trips if additional cabs added) from city to Airport during the peak hours.

Challenges @ Airport:

During Evening peak hour timeslot @ the Airport:

- Average trips completed is only 2 per week whereas expected is around 7 per week.
- Demand more than the available Cabs/Drivers.
- **72%** unavailability of cabs.

Possible causes :

- More incoming flights at the airport during Evening peak hours.
- Demand @ the airport is higher.
- Very little outbound flights from the airport.
 - **The number of cabs reaching the airport during this timeslot is very less resulting in the shortage of the cabs at the Airport.**
 - The drivers prefer not to go to the airport empty (without any passengers)

Proposed Solution :

- Starting 4 PM every day, identify cabs that are in and around a set (say 5 KM) radius from airport. Flash a message on the Mobile App, indicating that they are eligible for a weekly incentive if they reach Airport and take the pickup during the peak hours at the Airport.
- Add additional Cabs/Drivers to meet the peak hour demand.
- No. of trips based incentives if the driver completes more than the expected average # of 7 trips per week (5 trips if additional cabs added) from city to Airport during the peak hours.