# Data Exploration

- Data set contains 6 columns:
  1. Request.id
  2. Pickup.point
  3. Driver.id
  4. Status
  5. Request.timestamp
  6. Drop.timestamp

- Problem Statement: Find out the supply demand gap and suggest way to improve the situation.

- Based on the problem statement, we should focus on pickup point, status and driver id.

UpGrad

Data Cleaning and Manipulation

- Possible data inconsistencies:
  1. Duplicate values of Request ID.
  2. NA values in the columns of interest.

- Other Issues:
  1. Request time stamp is in factor format. Convert it to date time format.
  2. Dates are separated by "/" and "-". Make this consistent for ease of data analysis.

Requests Analysis– Cancellations in morning, unavailable in the evening

- The pattern of requests is common for all the days for the status of requests.

Request Data - Hourly Number of Requests Day Wise
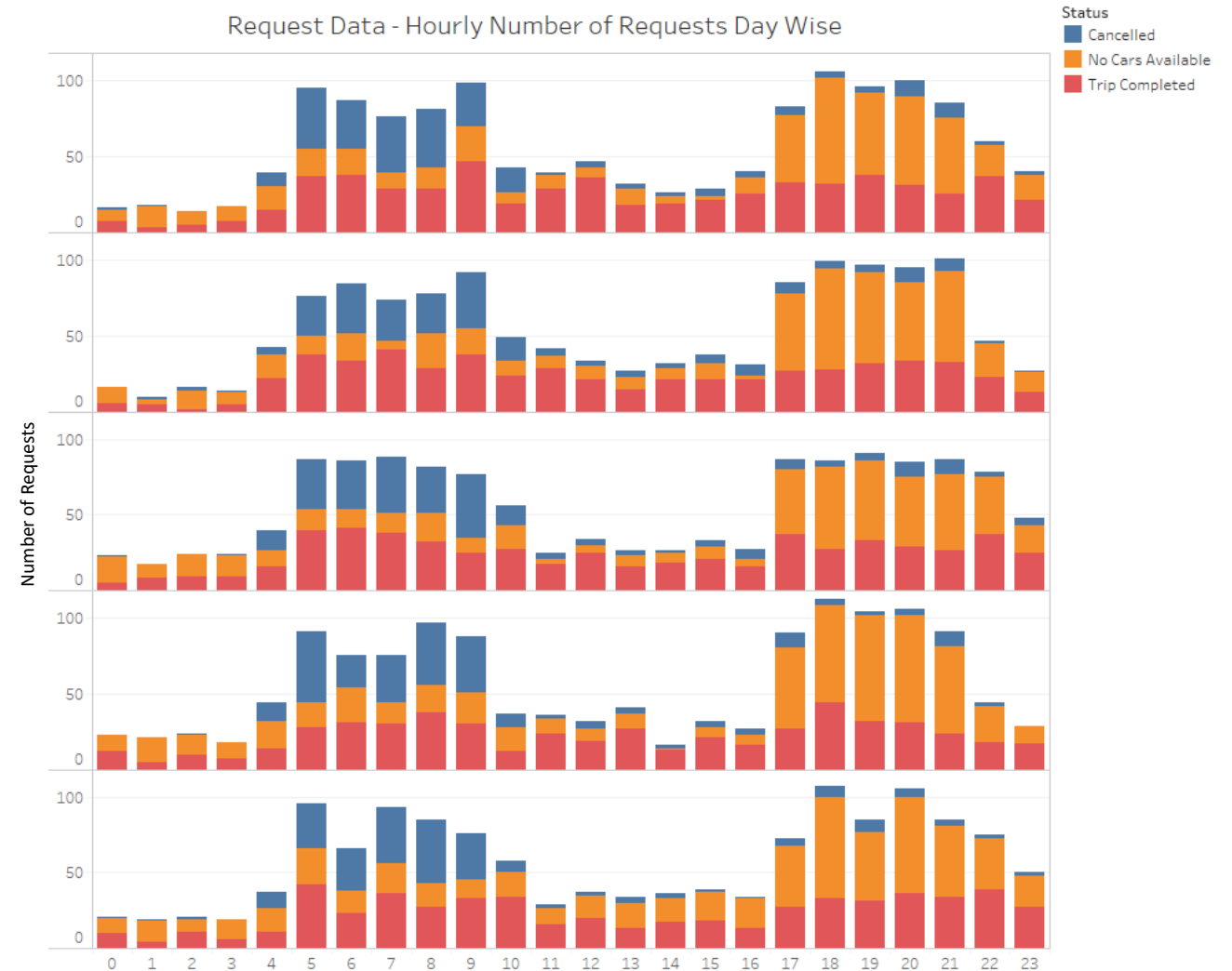
Requests Analysis – City and Airport requests in morning and evening respectively

- The pattern of requests is common even for all the days for the pick up point where the requests have been generated.

Request Data - Hourly Number of Requests Day Wise
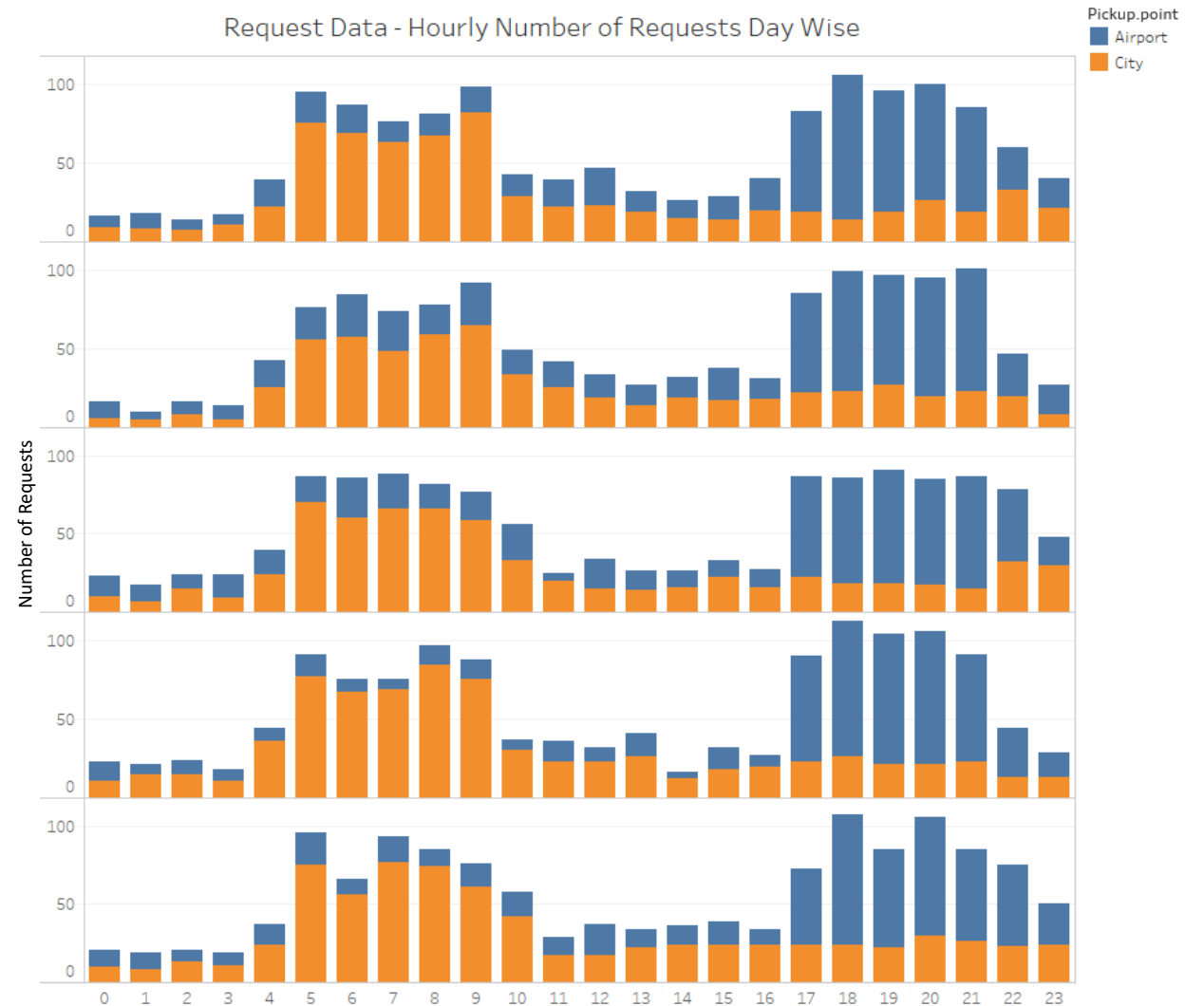
Combining Data For All Days

- Previous graphs show that all the days show common trends.

- Hence the number of requests can be clubbed together for further analysis.

- Conclusion:
  1. Number of trips in morning are high from the city.
 2. Number of trips from the evening are high from the airport.

Request Data - Hourly Number of Requests

UpGrad
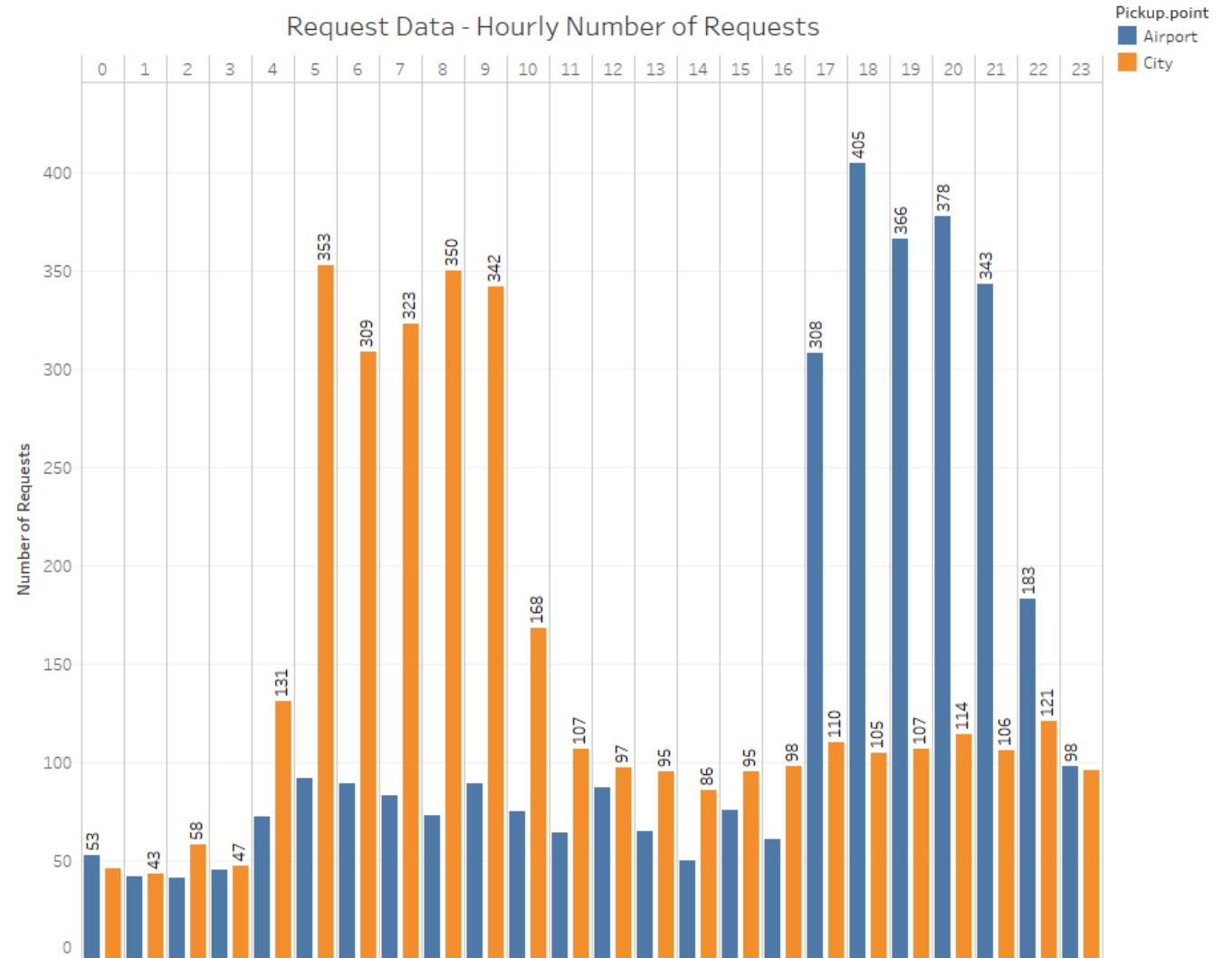
Binning Time Into 5 Categories

- Requests generated are divided into 5 homogenous categories based on the time of request

| Time	Category
12 AM – 5 AM	Pre_Morning
5 AM – 10 AM	Morning_Rush
10 AM – 5 PM	Day_Time
5 PM – 10 PM	Evening_Rush
10 PM – 12 AM	Late_Night

UpGrad

Problem Identification – Morning and Evening

- Graph clearly shows that the major problems are:
  1. Cancelled trips during the morning rush
  2. Unavailability of cars during evening rush

Number of Trips Vs Time Slot

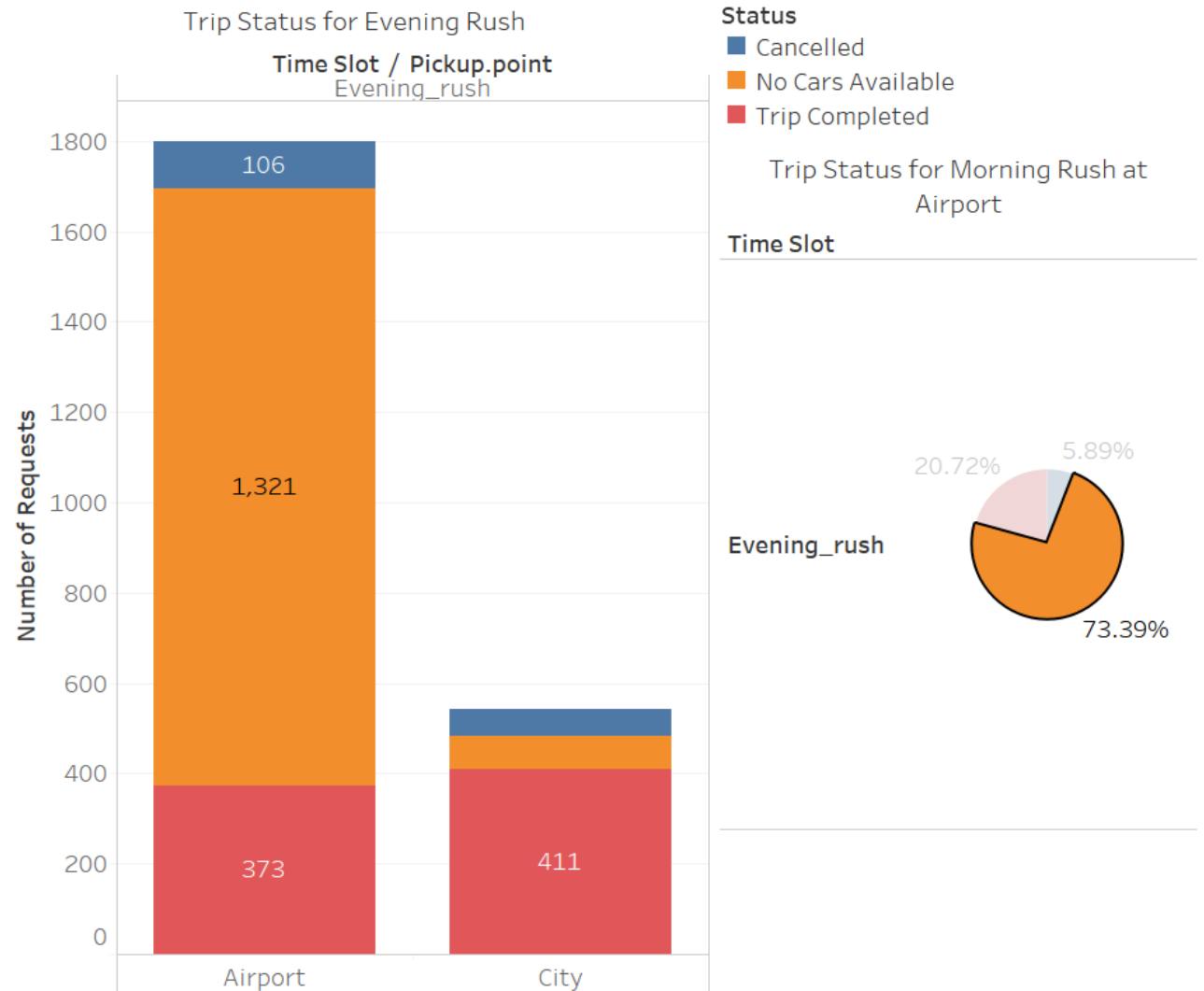**Status**
- Cancelled
- No Cars Availa..
- Trip Completed
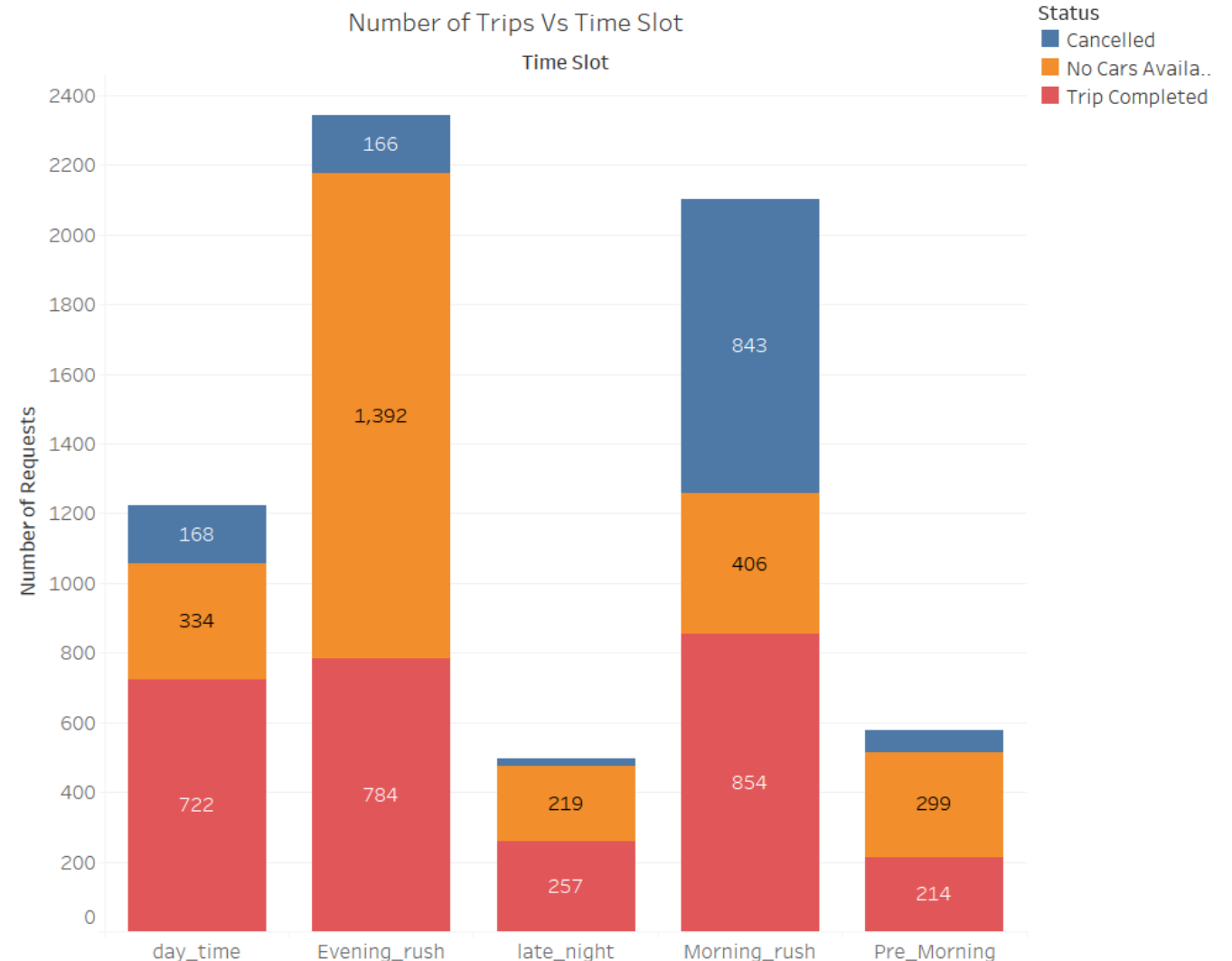
Cancelled trips - 50 % of total trips from city

- The supply from the city is 472, while the demand is 472 + 285 + 820 = 1677.

- The difference between the demand and the supply is 1205.

UpGrad
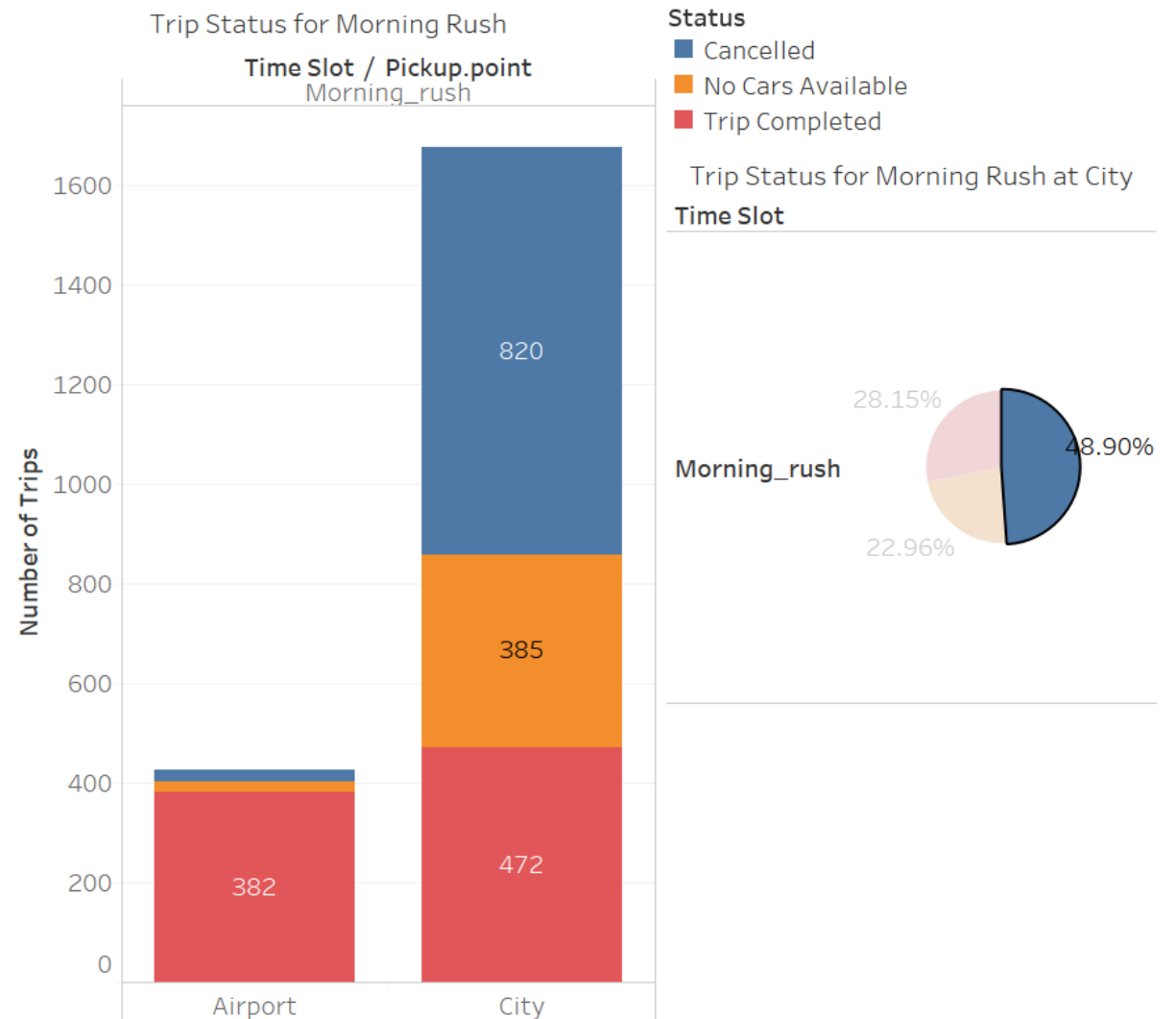
No Cars - 70 % of total trips from airport

- The supply from the airport is 373, while the demand is 373 + 1321 + 106 = 1800.

- The difference between the demand and the supply is 1427.

Trip Status for Evening Rush

Status
- Cancelled
- No Cars Available
- Trip Completed

Trip Status for Morning Rush at Airport

UpGrad

Recommendations

- For the trips in the morning, drivers can be incentivised to make those trips.
  1. They could be given a bonus for each trip they complete from the city to the airport in the morning rush. This will ensure that less number of trips are cancelled.
  2. Uber can pay for the gas mileage of drivers to come back to the city without a ride.
  3. Uber can increase the demand at the airport to reduce idle time – by increased marketing and price cuts for the passengers.

- For the evening, since the number of drivers is less, some of the ways are:
  1. Drivers can again be given a bonus to complete a trip from the airport in the evening. This will ensure that the supply increases at the airport.
  2. Uber can also pay drivers to come without a passenger to the airport.
  3. Another innovative way can be to pool the rides of passengers so that lesser number of cars can serve more passengers.

UpGrad