UBER SUPPLY DEMAND GAP ASSIGNMENT SUBMISSION

GAURAV SACHDEVA

PGDDS COHORT 7

09/09/2018

Uber Supply-Demand Gap Issue

Problem Statement:

The issue being faced by Uber involves the cancellation by the driver or non-availability of cars. These problems are not only faced by customers but these also impact the business of Uber. If drivers cancel the request of riders or if cars are unavailable, Uber loses out on its revenue.

Considerations:

Only the trips to and from the airport are being considered.

Analysis Checkpoints:

1. Identify the most pressing problems for Uber
2. Find out the gap between supply and demand
3. Reason for the supply-demand gap

Business Objectives:

1. Identifying the root cause of the problem (i.e. cancellation and non-availability of cars)
2. Recommend ways to improve the situation.

Data Cleaning ,Preparation, Univariate and Bivariate analysis

Data quality is crucial for this step. Ride request information is in csv file.

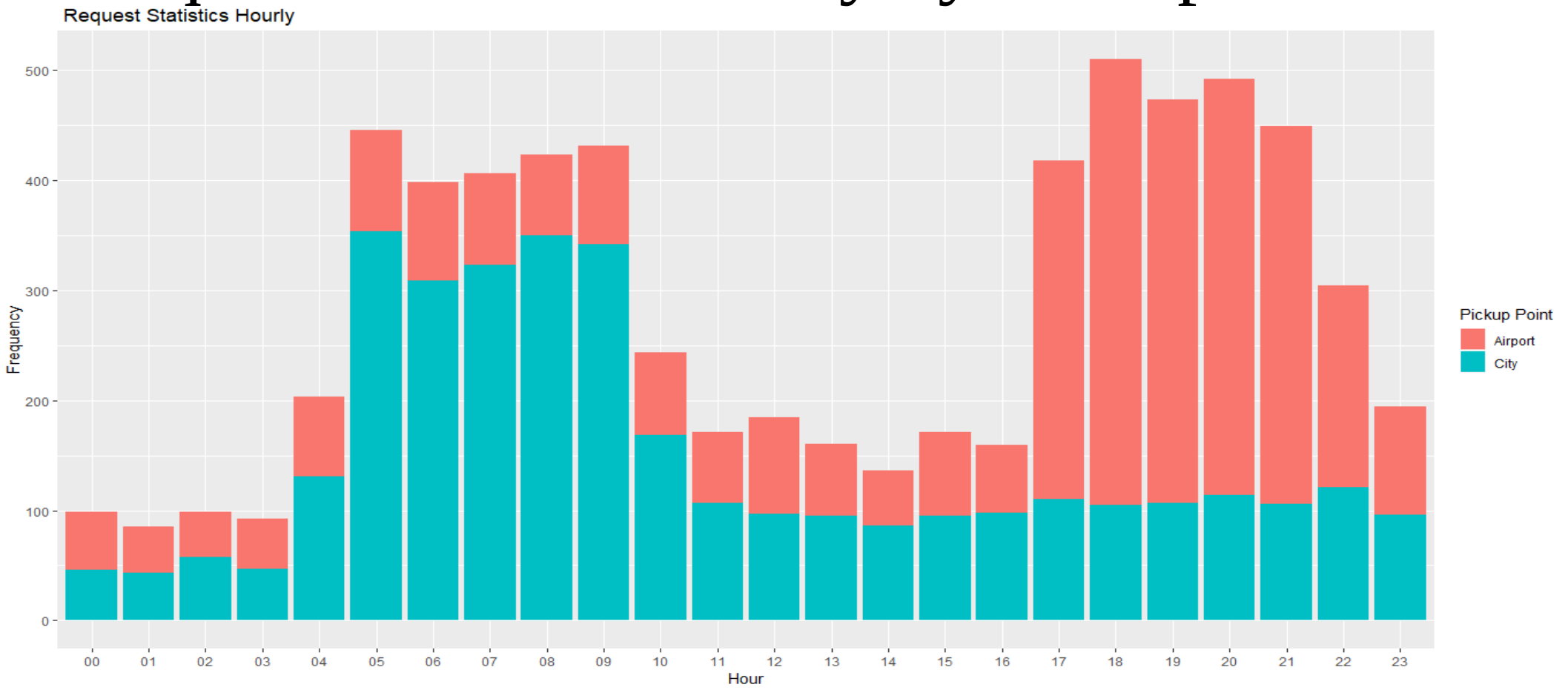
RStudio is extensively used and following stages are addressed in data cleaning, data preparation and analysis:

- Data Import – Import data from files.
- Data Cleansing – Removing or correcting dirty data. Handling of NA values. Removing Duplicate records
- Data Preparation(1)– Parsing of timestamp data. Creating columns for date, hour, minute and second for timestamp data.
- Data Preparation (2)– Creating columns for trip duration, timeslot and service status for analysis.

- Univariate and Bivariate analysis by using the following variables appropriately

1. Pickup Point
2. TimeSlot
3. Request Status

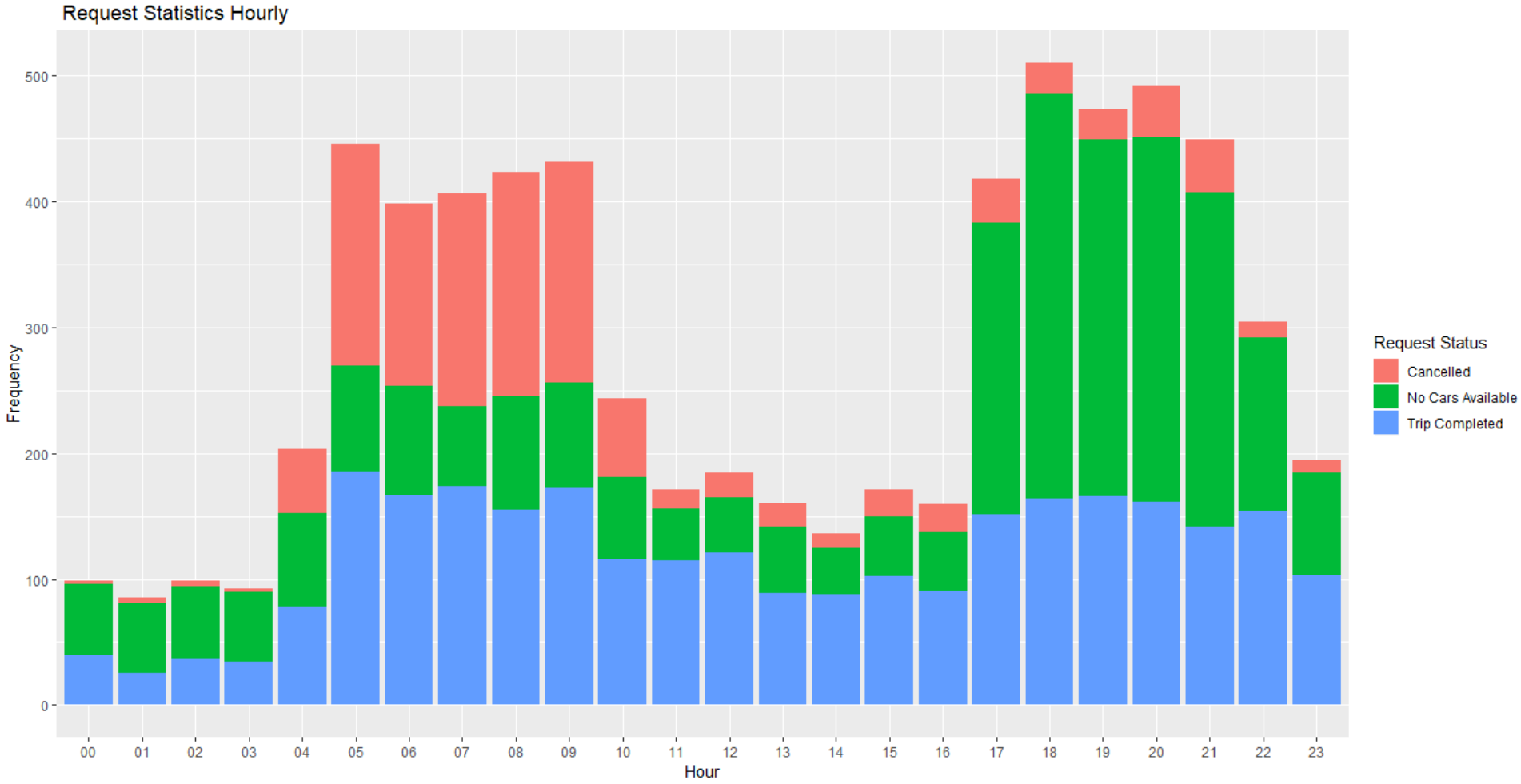
Request Statistics Hourly by Pickup Point



Observations

- Morning Time(4AM~10AM)Ride Requests from the City to Airport outnumber those from Airport to City.
- Evening Time(5PM~10PM) Ride Requests from the Airport to City outnumber those from City to Airport.

Request Statistics Hourly by Status



Observations

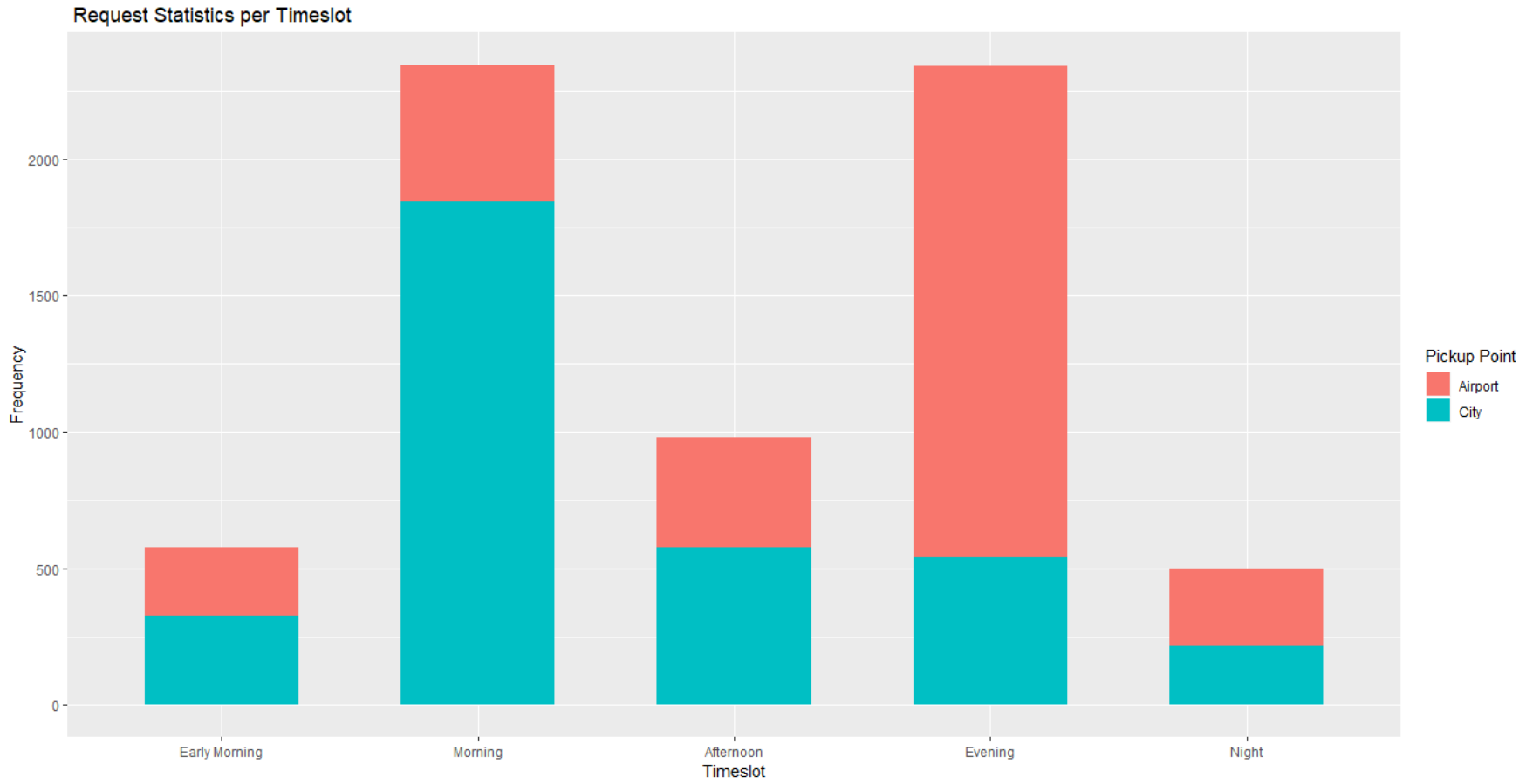
- Morning Time(5AM~10AM) has a huge number of cancellations by drivers.
- Evening Time(5PM~11PM),Ride Requests from the Airport to City are met with non-availability of cars.

Division of the day into Timeslots

For Better analysis for the results shown by the hour in previous slides, the day is divided into Timeslots as per following

TimeSlot	Name
Midnight till 5 AM	**Early Morning**
5AM till 11AM	**Morning**
11AM till 5PM	**Afternoon**
5PM till 10PM	**Evening**
10PM till Midnight	**Night**

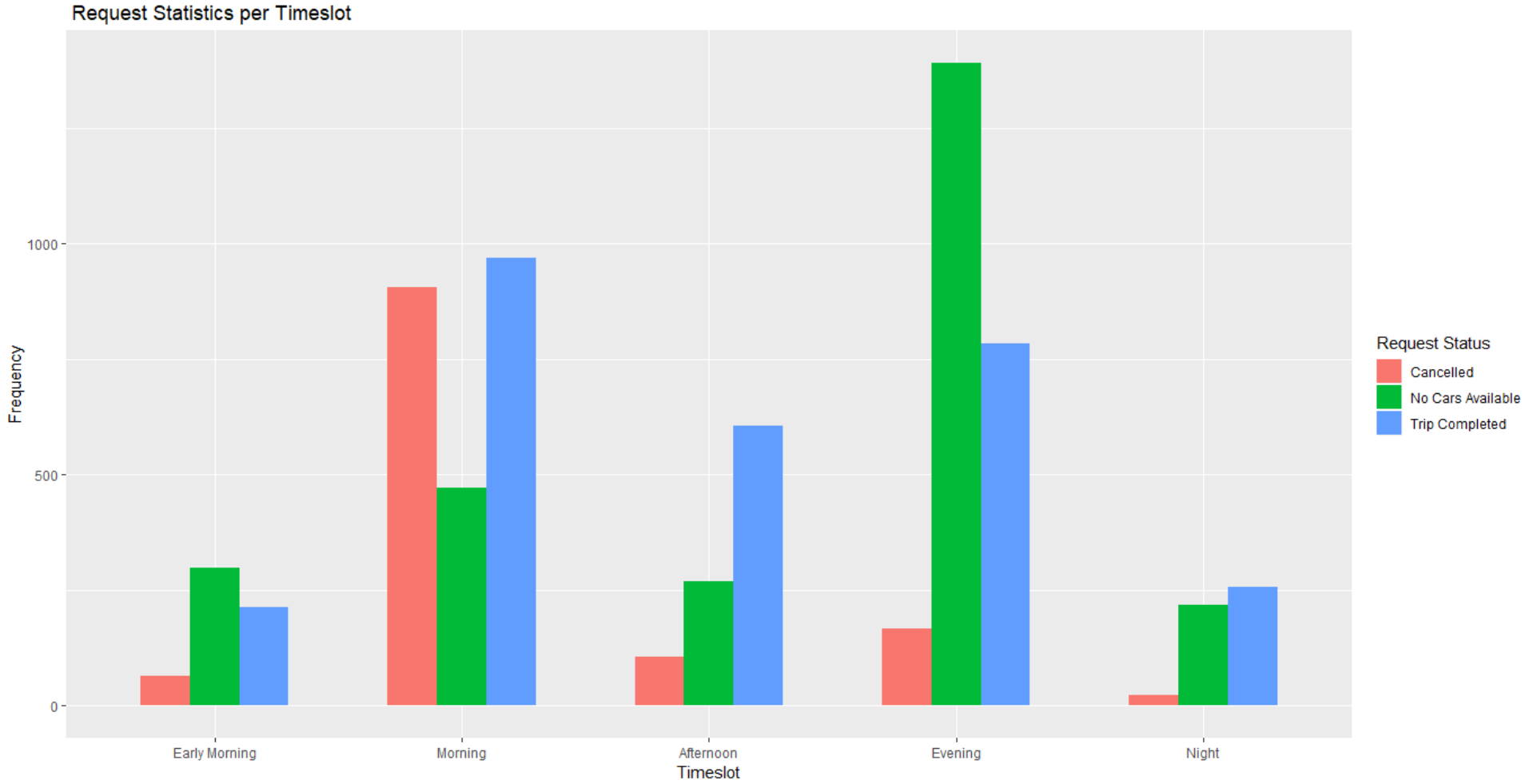
Request Statistics per Timeslot by Pickup Point



Observations

- In the Morning Timeslot, Ride Requests from the City to Airport outnumber those from Airport to City.
- In the Evening Timeslot, Ride Requests from the Airport to City outnumber those from City to Airport.

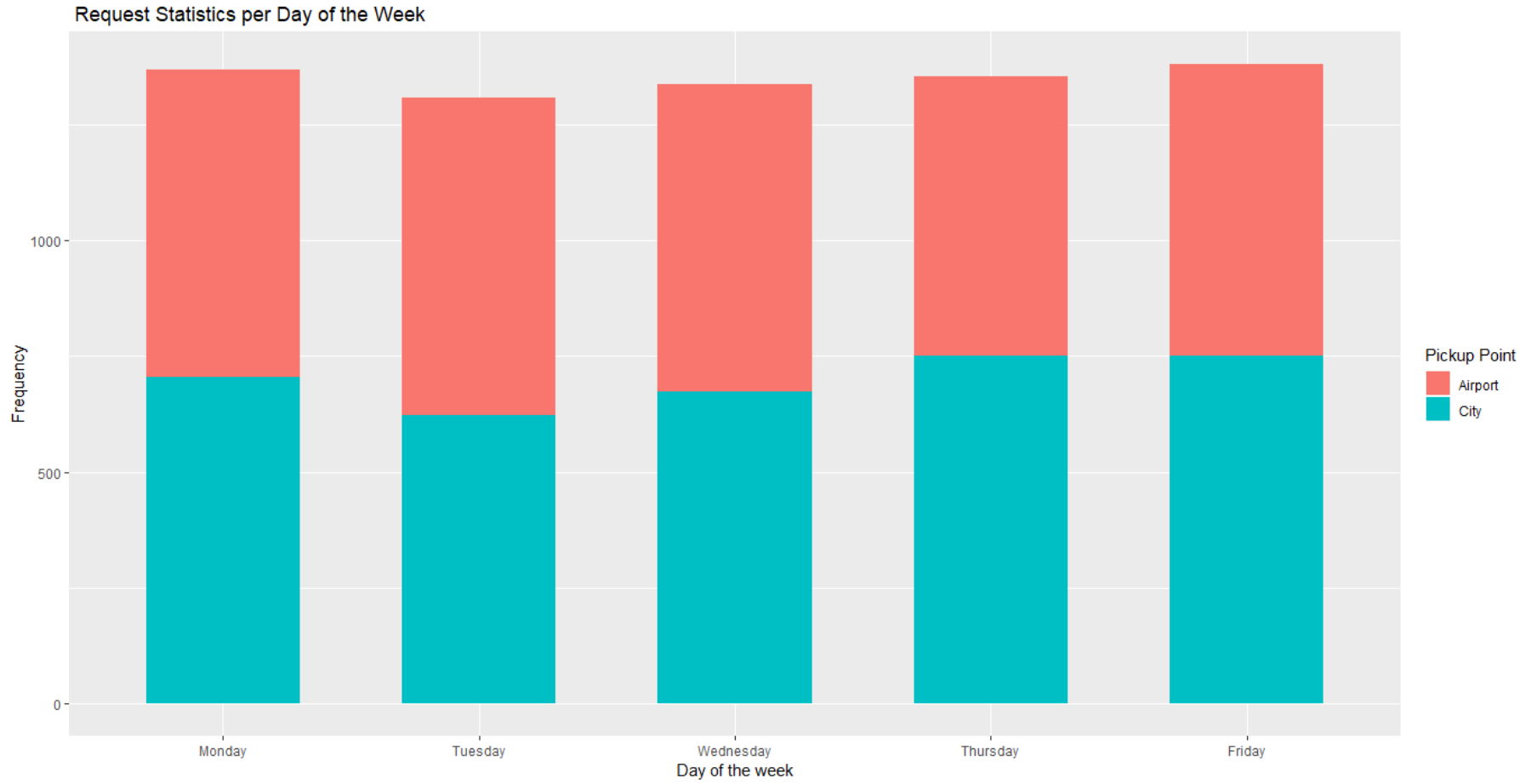
Request Statistics per Timeslot by Status



Observations

- In the Morning Timeslot, trips completed and cancellations are higher as compared to other timeslots.
- In the Evening Timeslot, non-availability of cars is the highest.

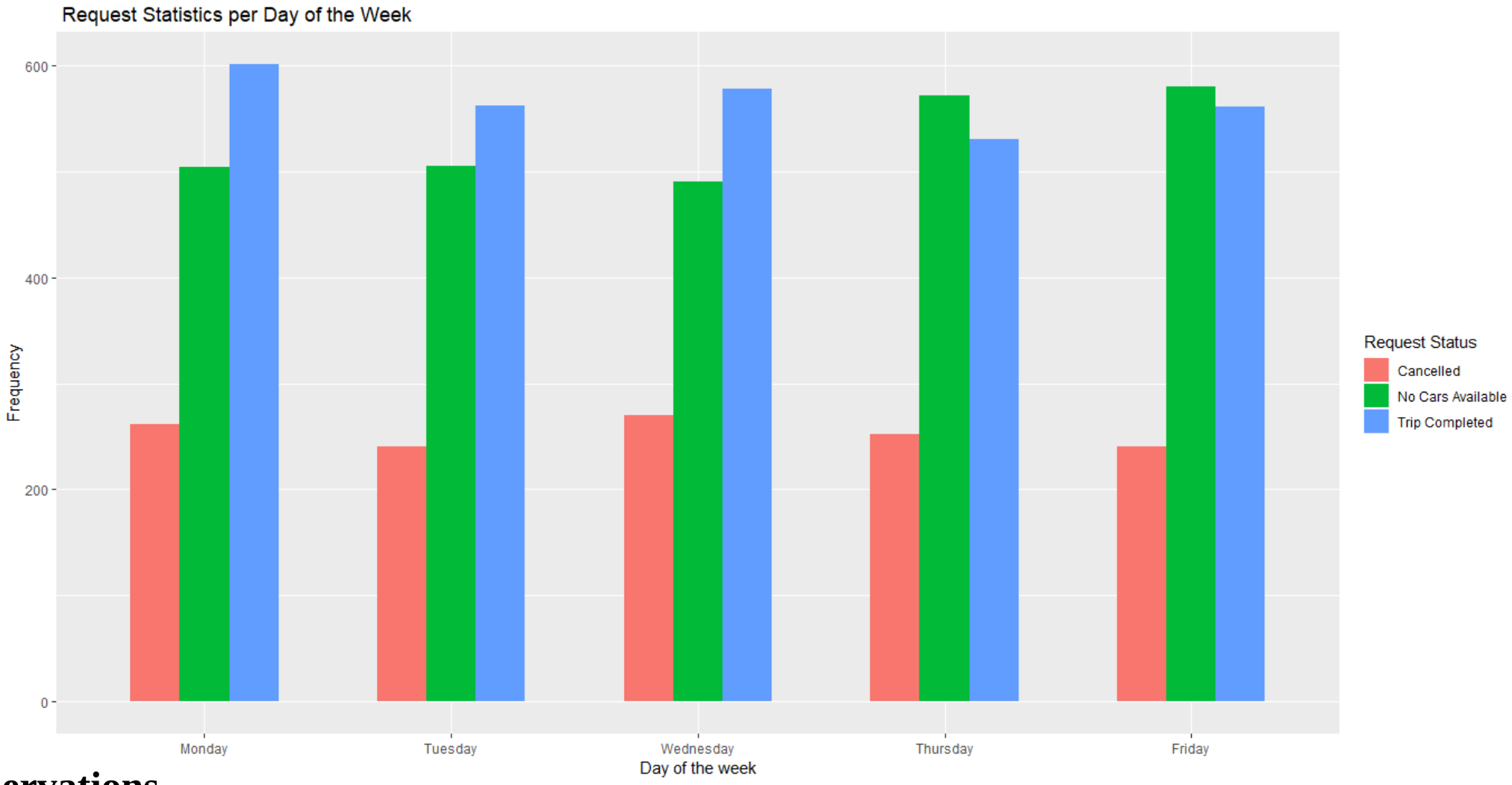
Request Statistics per DOW by Pickup Point



Observations

- Ride requests are highest on Mondays and Fridays when people enter and leave the city at the start and end of the work week.

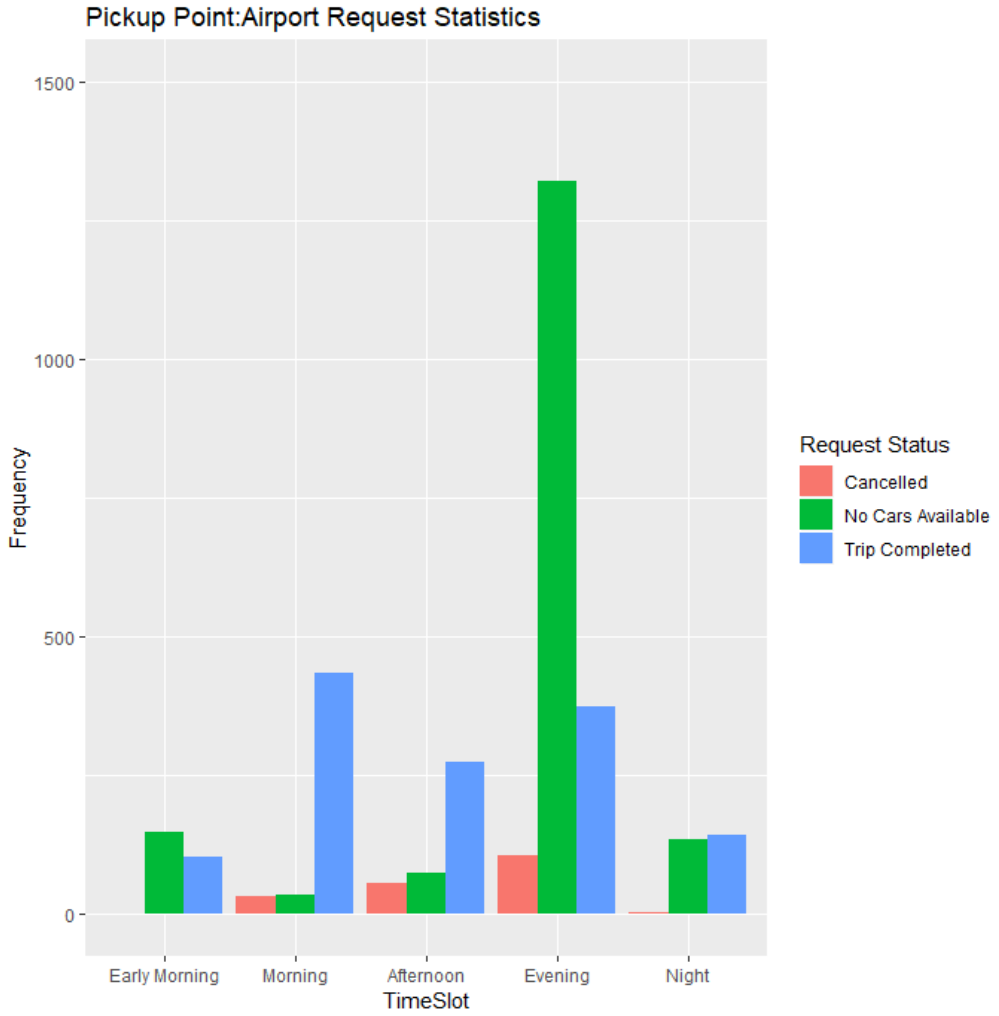
Request Statistics per DOW by Status

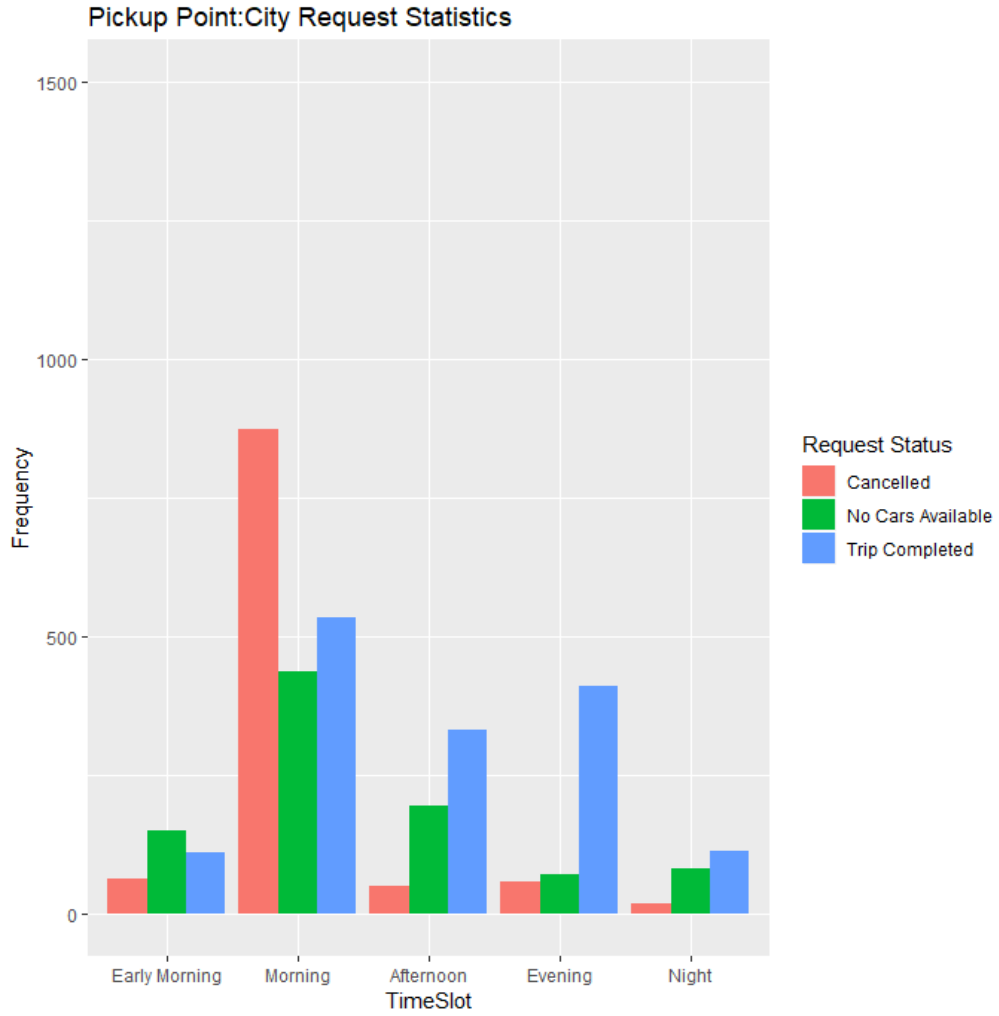


Observations

- Non-availability of cars is highest at the end of work week(Thursday, Friday)

Request Statistics by Timeslots comparing Pickup Points

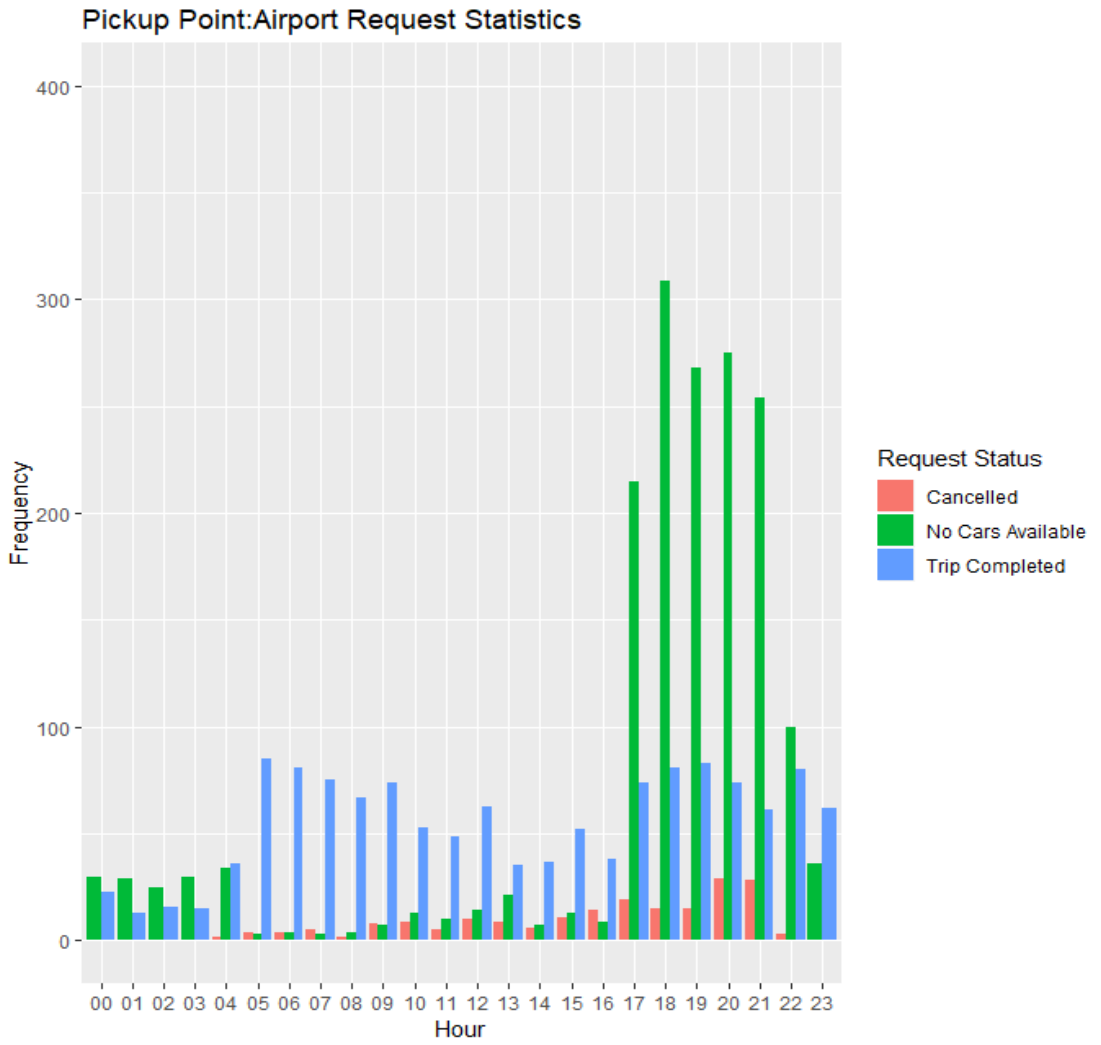


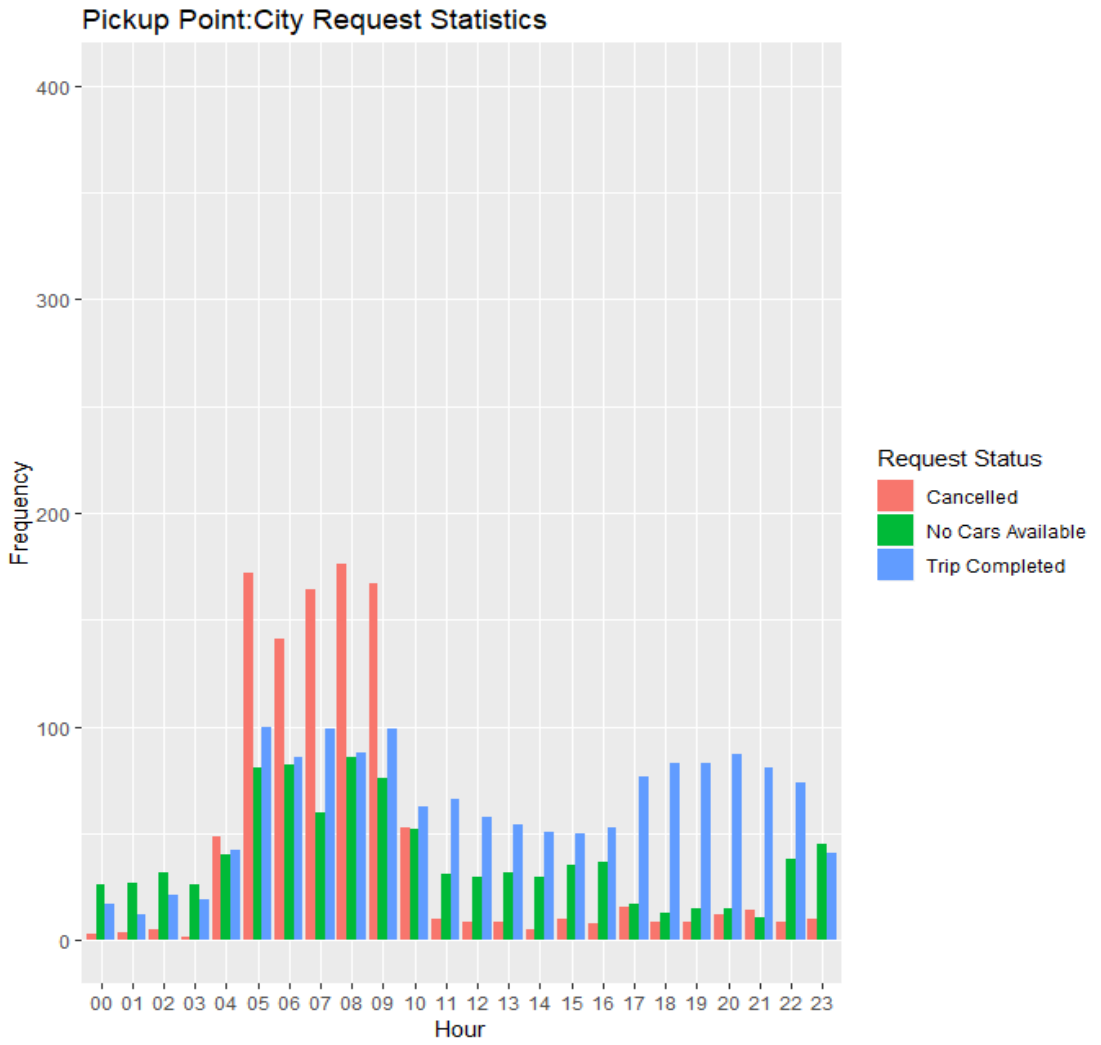


Observations

- In the Morning Timeslot, cancellations and trips completed from the City are the highest.
- In the Evening Timeslot, non-availability of cars from the Airport is the highest.

Request Statistics by Hour comparing Pickup Points

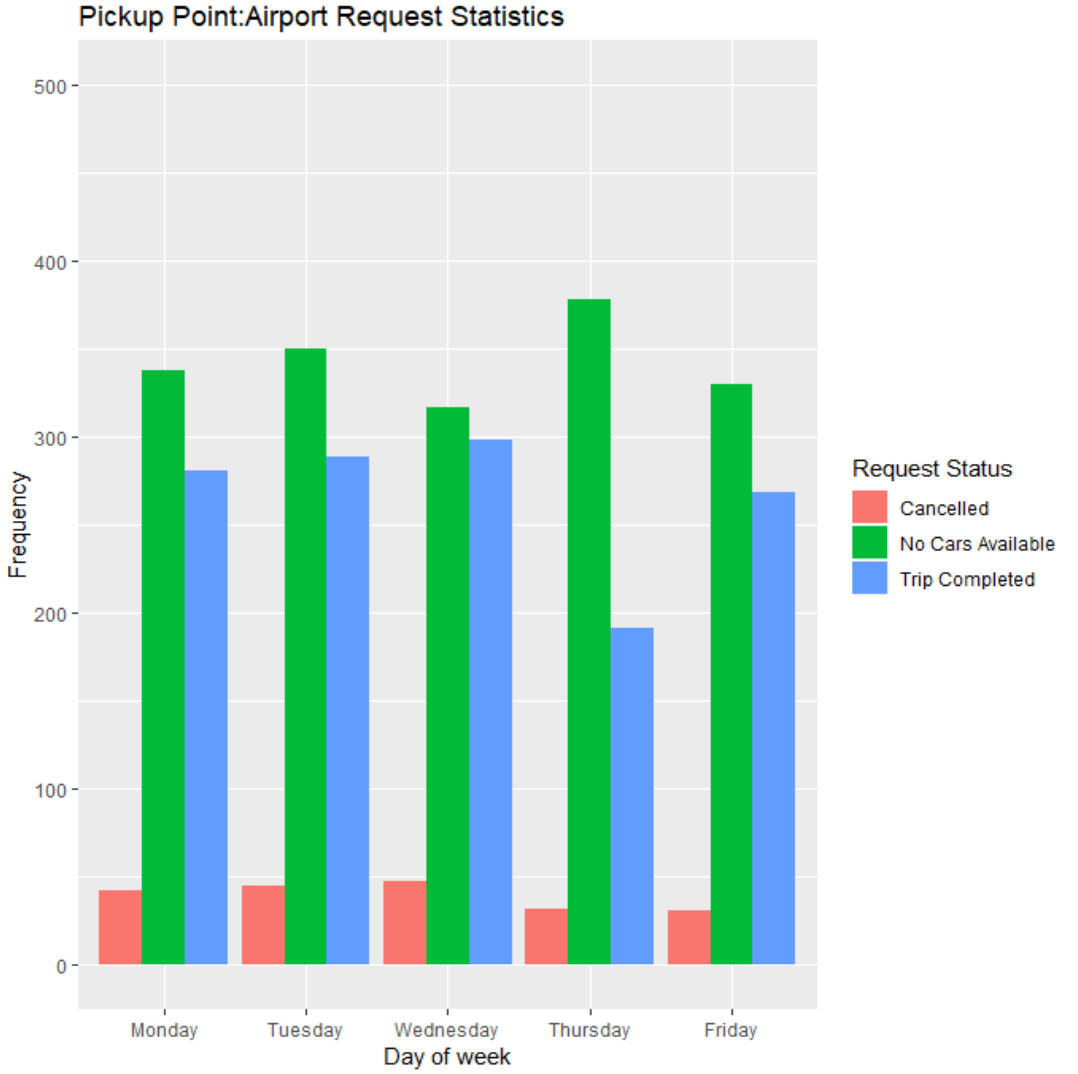


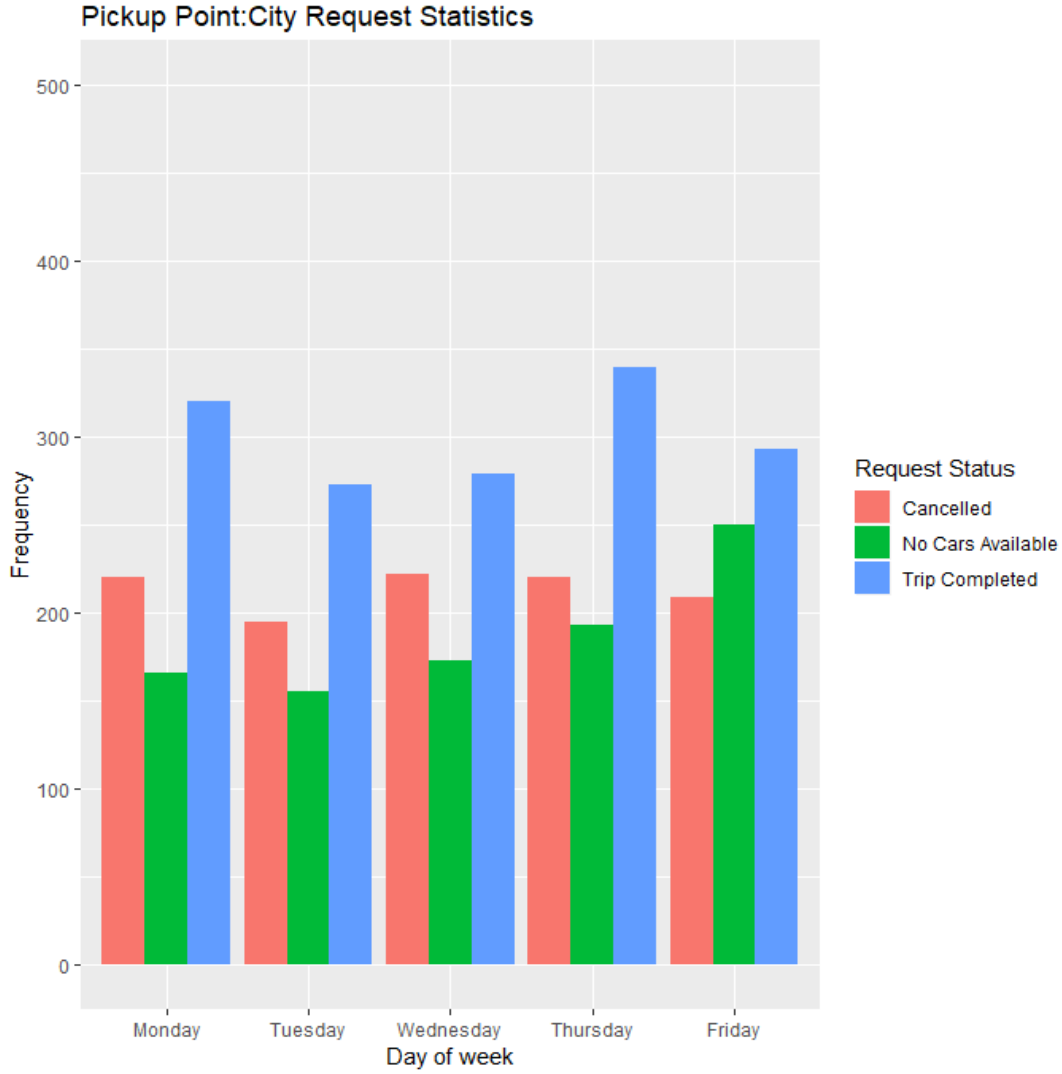


Observations

- Morning Time(5AM~10AM) has a huge number of cancellations by drivers at the City.
- Evening Time(5PM~10PM),Ride Requests from the Airport to City are met with non-availability of cars.

Request Statistics by DOW comparing Pickup Points

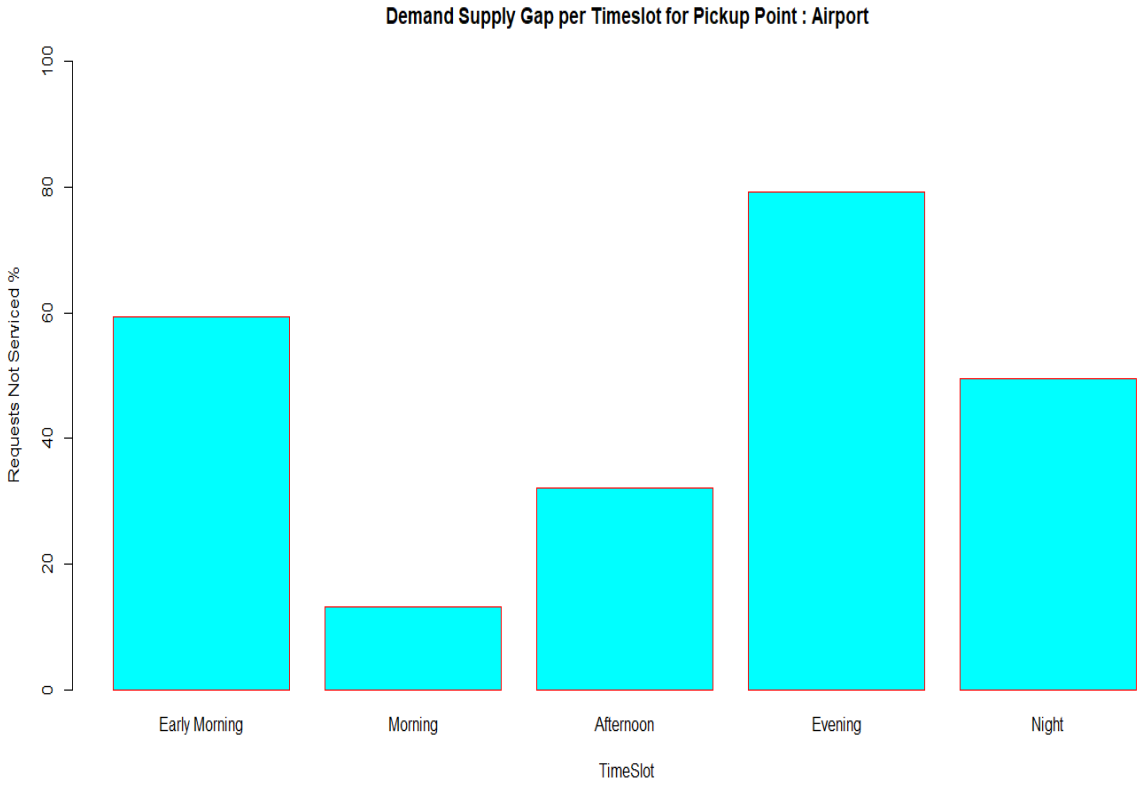


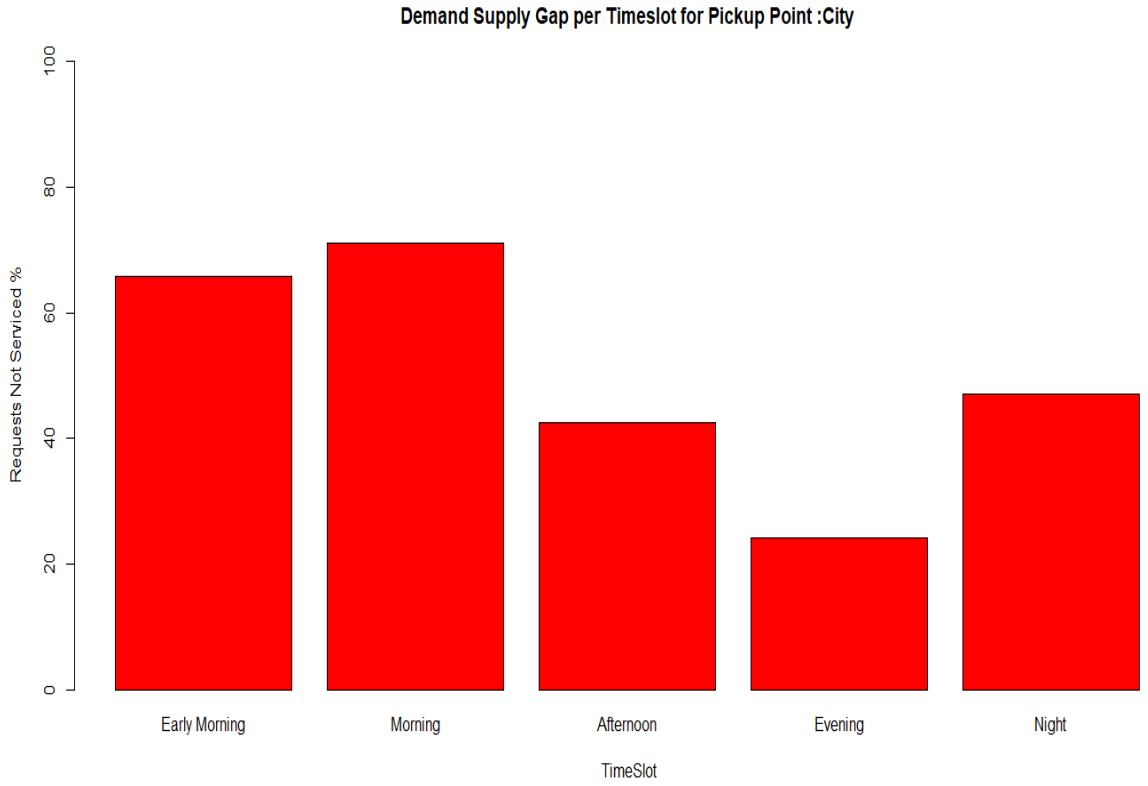


Observations

- Requests starting airport face non-availability of cars throughout the week. This is larger than trips completed

Demand Supply Gap% by TimeSlot

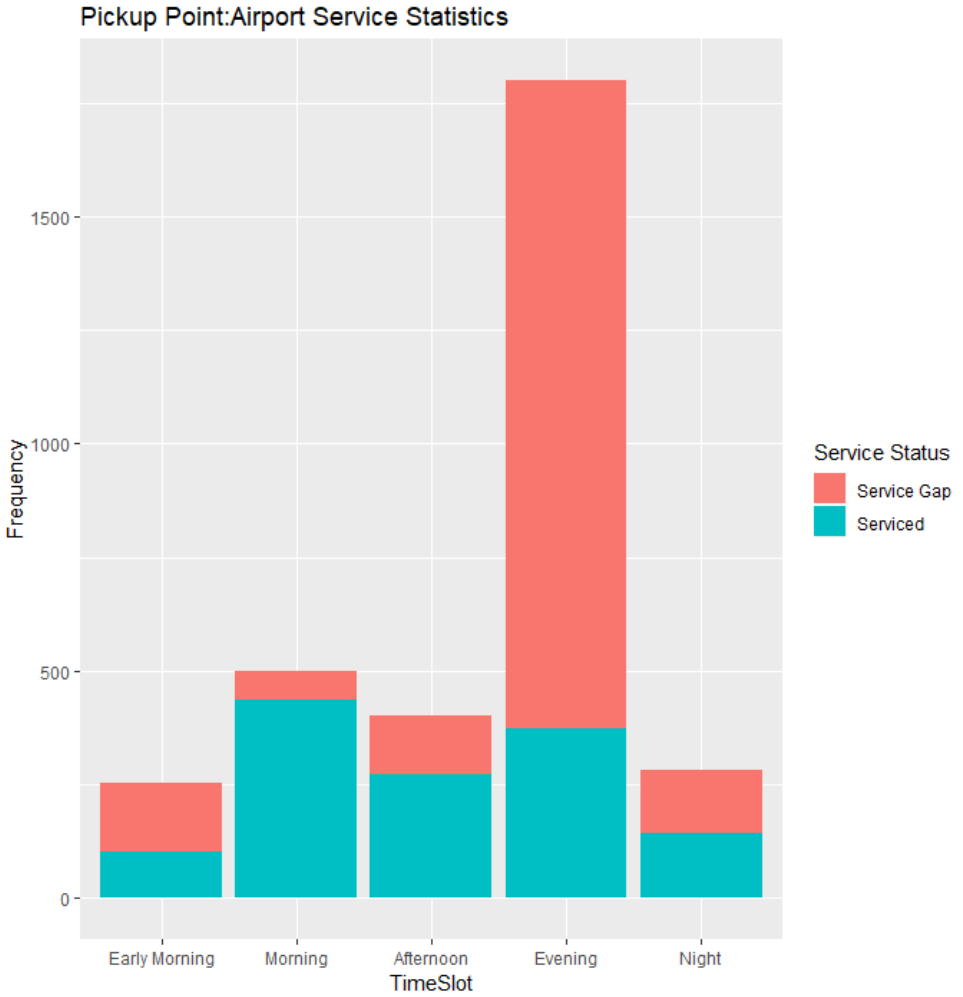


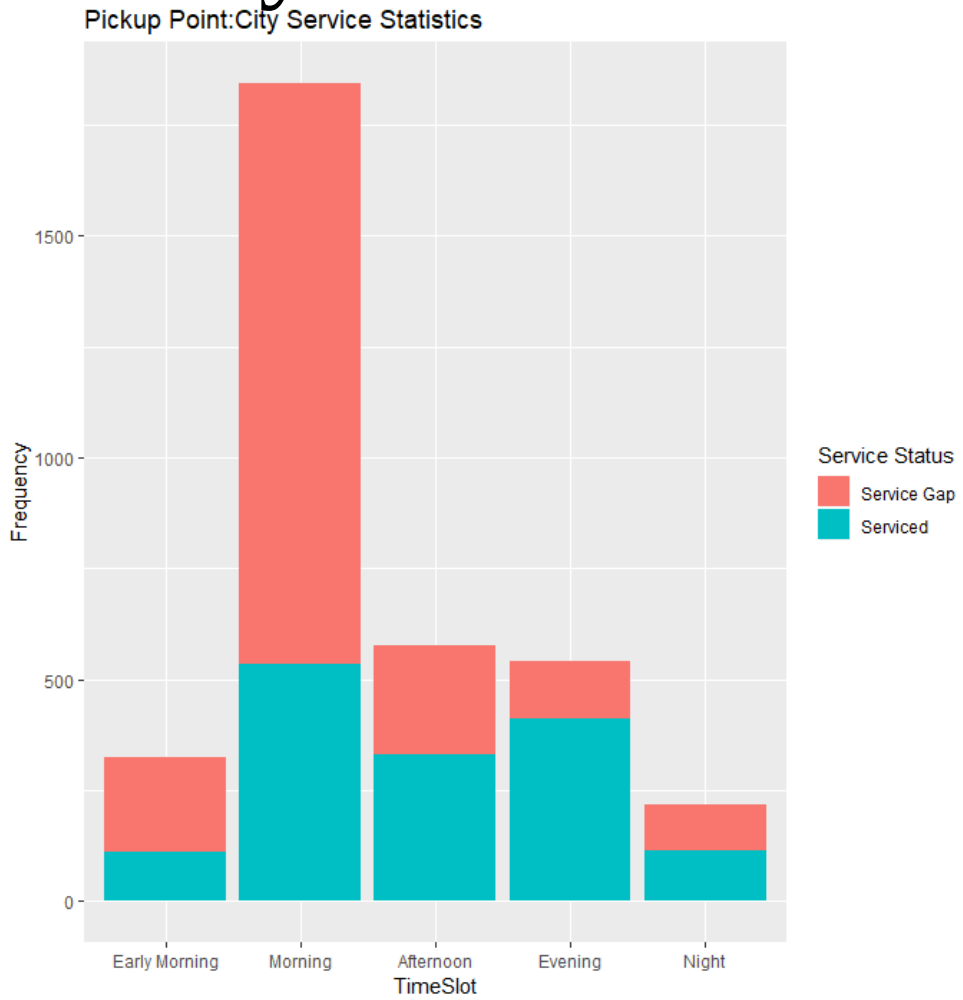


Observations

- Demand Supply Gap percentage at Airport is highest during Evening TimeSlot. It is also substantially high during Early Morning TimeSlot.
- Demand Supply Gap percentage at City is highest during Morning TimeSlot. It is also substantially high from Early Morning TimeSlot.

Service Statistics by TimeSlot

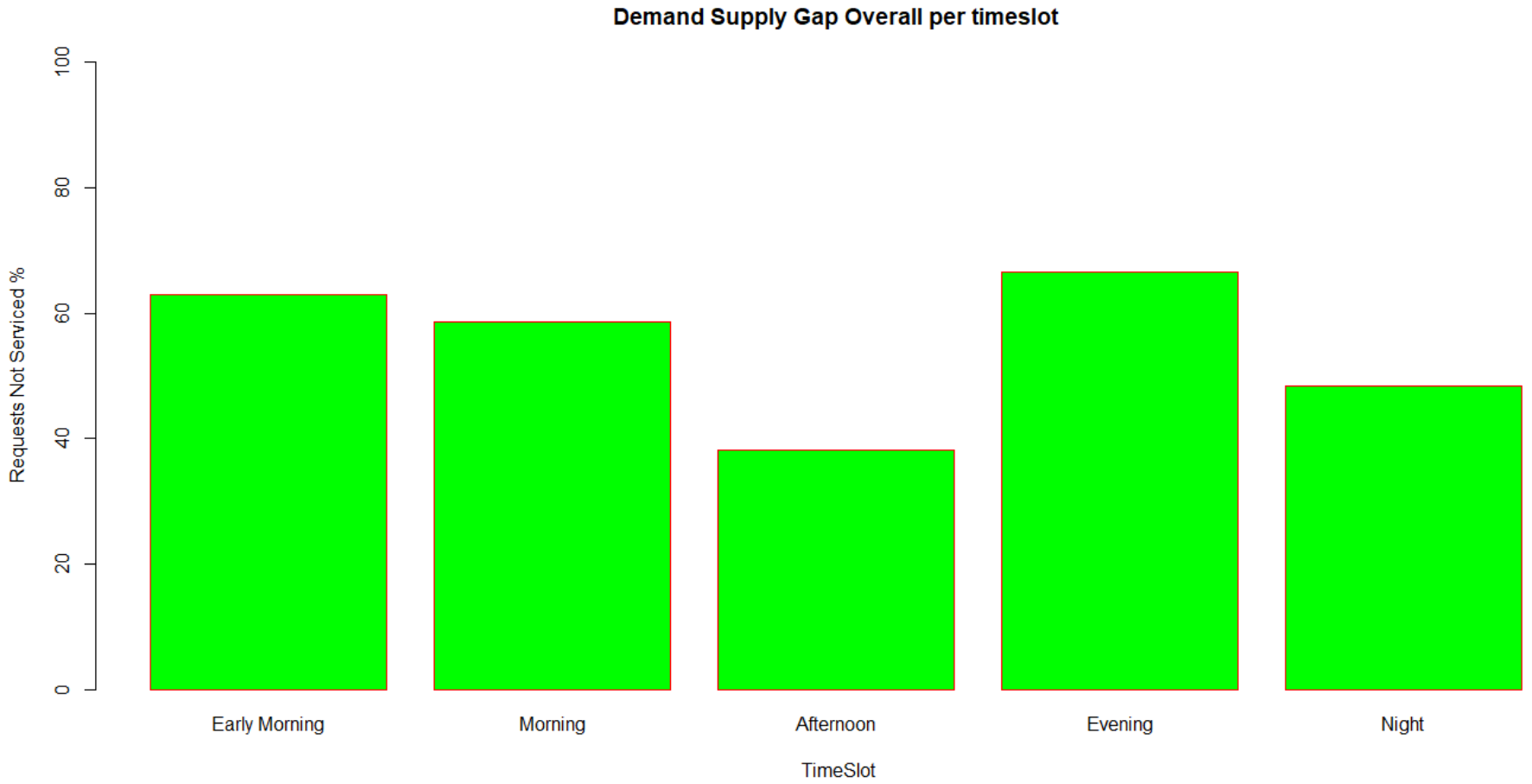




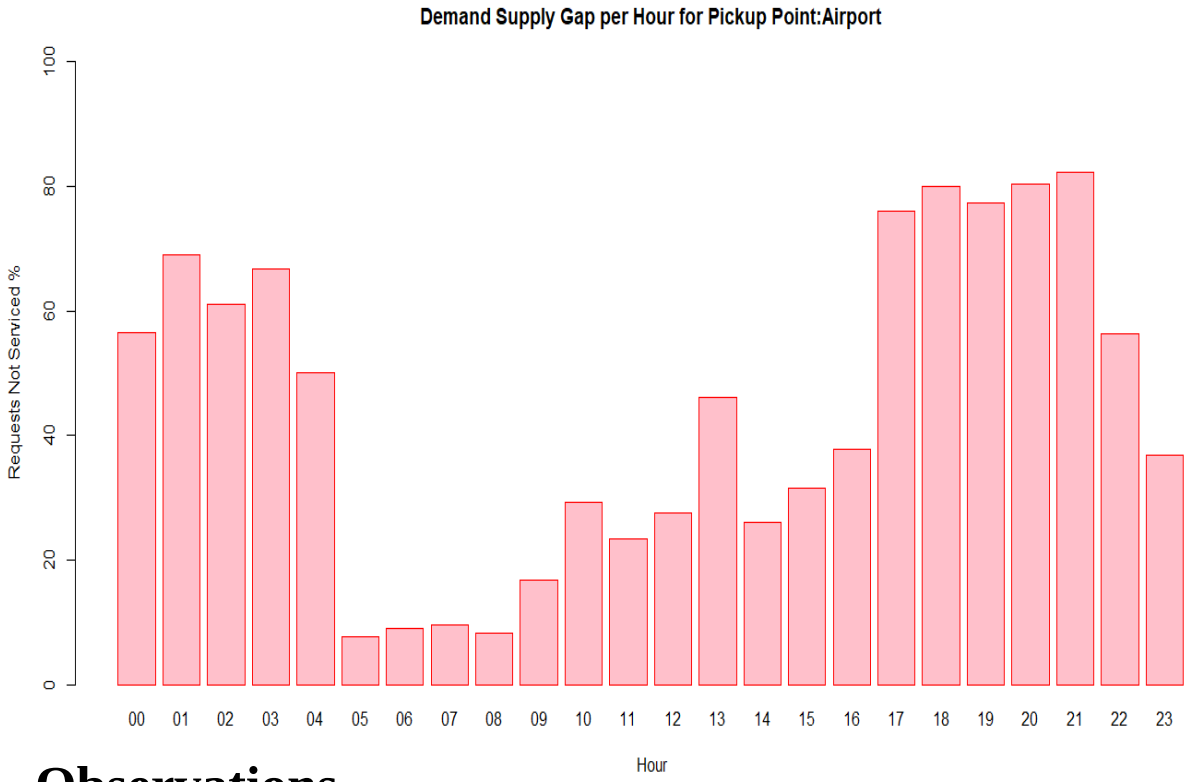
Observations

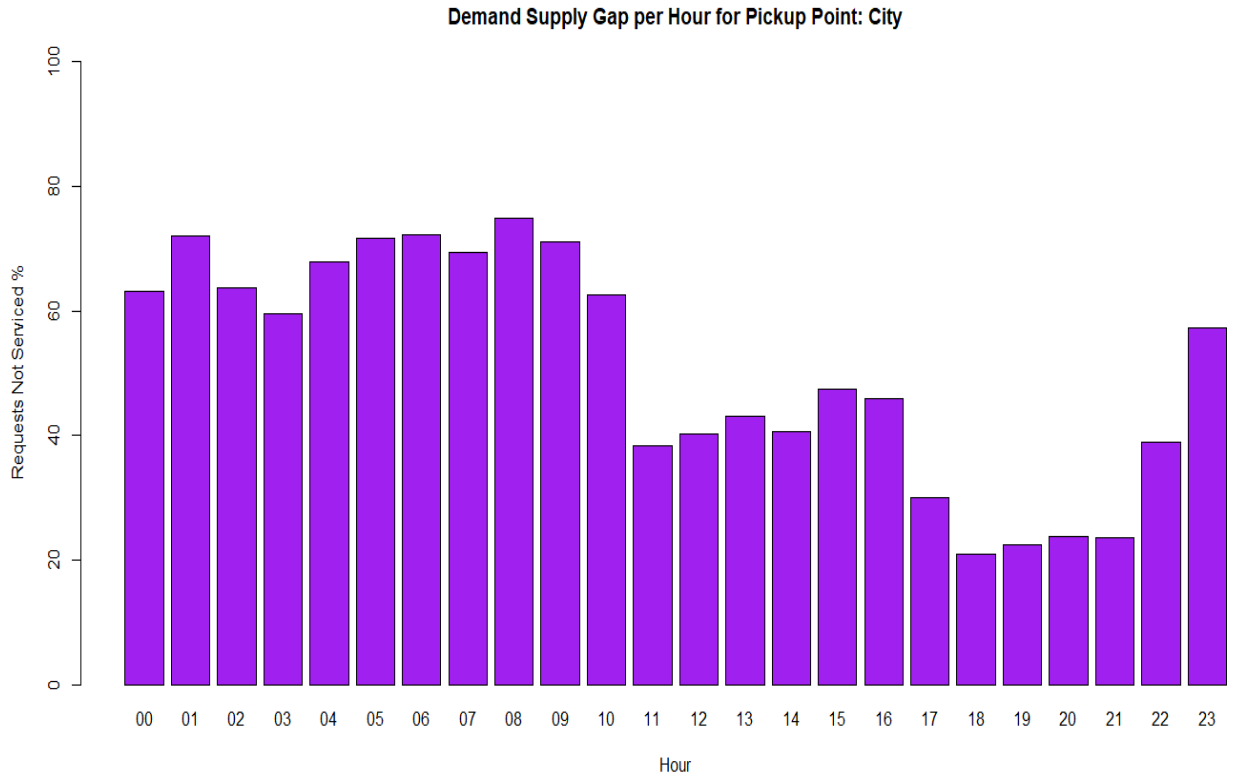
- Service Gap in terms of Requests at Airport is highest in the Evening Timeslot.
- Service Gap in terms of Requests at City is highest in the Morning Timeslot.

Demand Supply Gap% Overall by TimeSlot



Demand Supply Gap% by Hour of the day

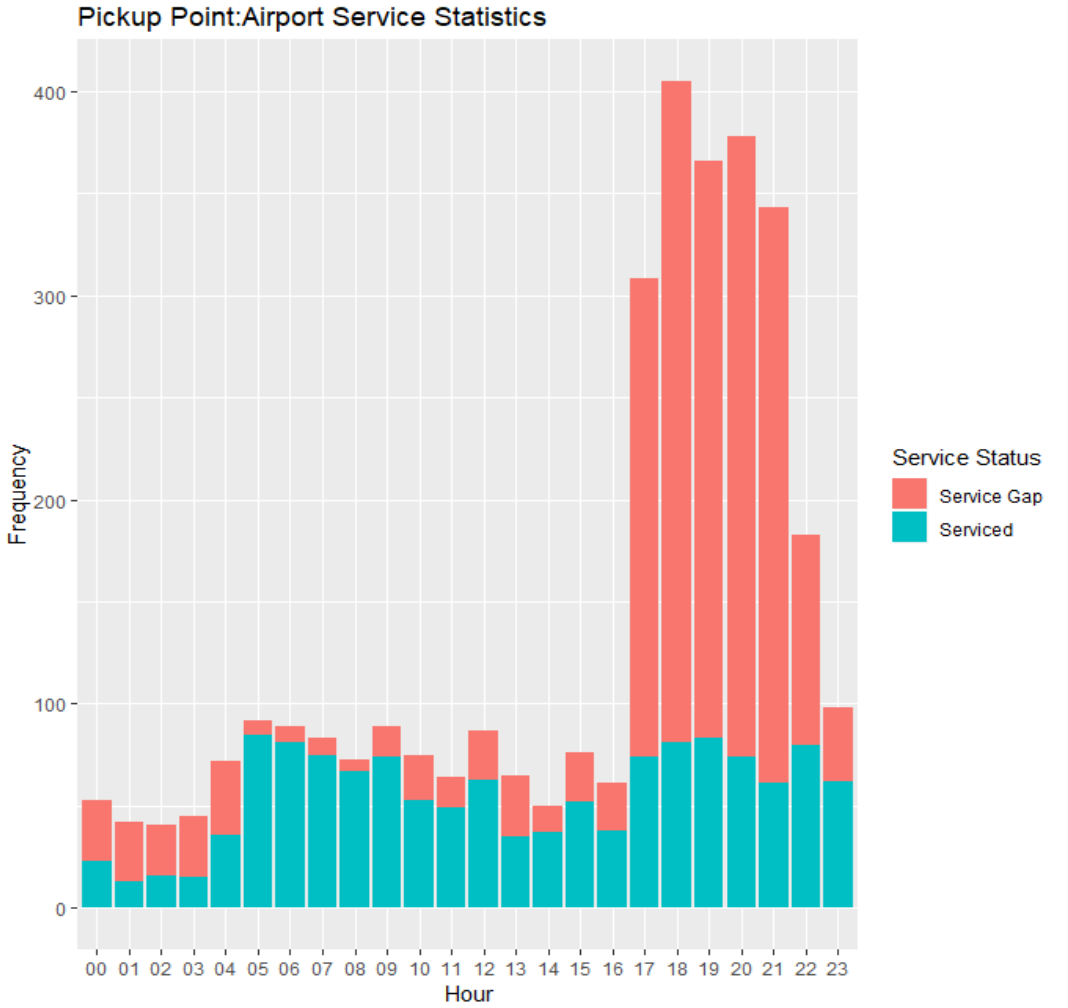


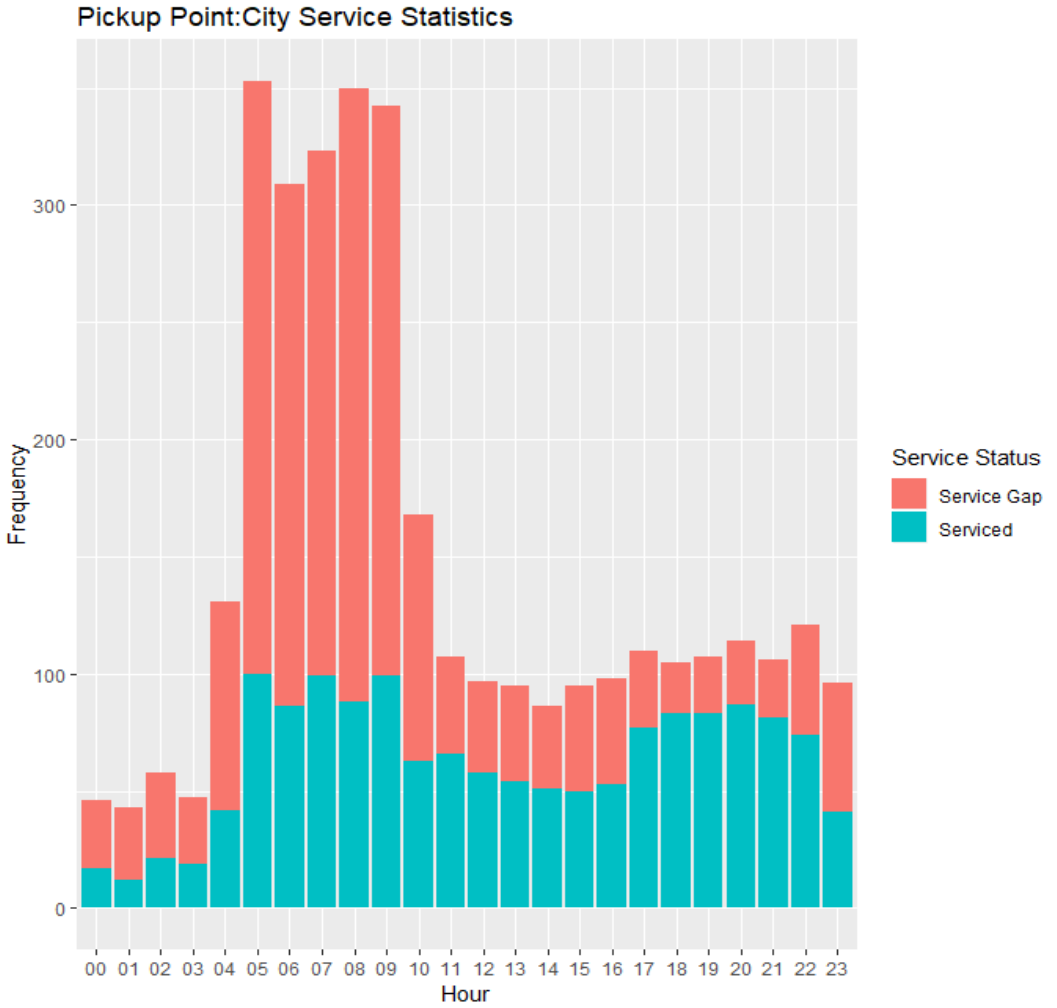


Observations

- Demand Supply Gap percentage at Airport is highest around 9 PM. It is also substantially high from Midnight~5AM
- Demand Supply Gap percentage at City is highest around 8AM. It is also substantially high from Midnight~5AM

Service Statistics by Hour of the day

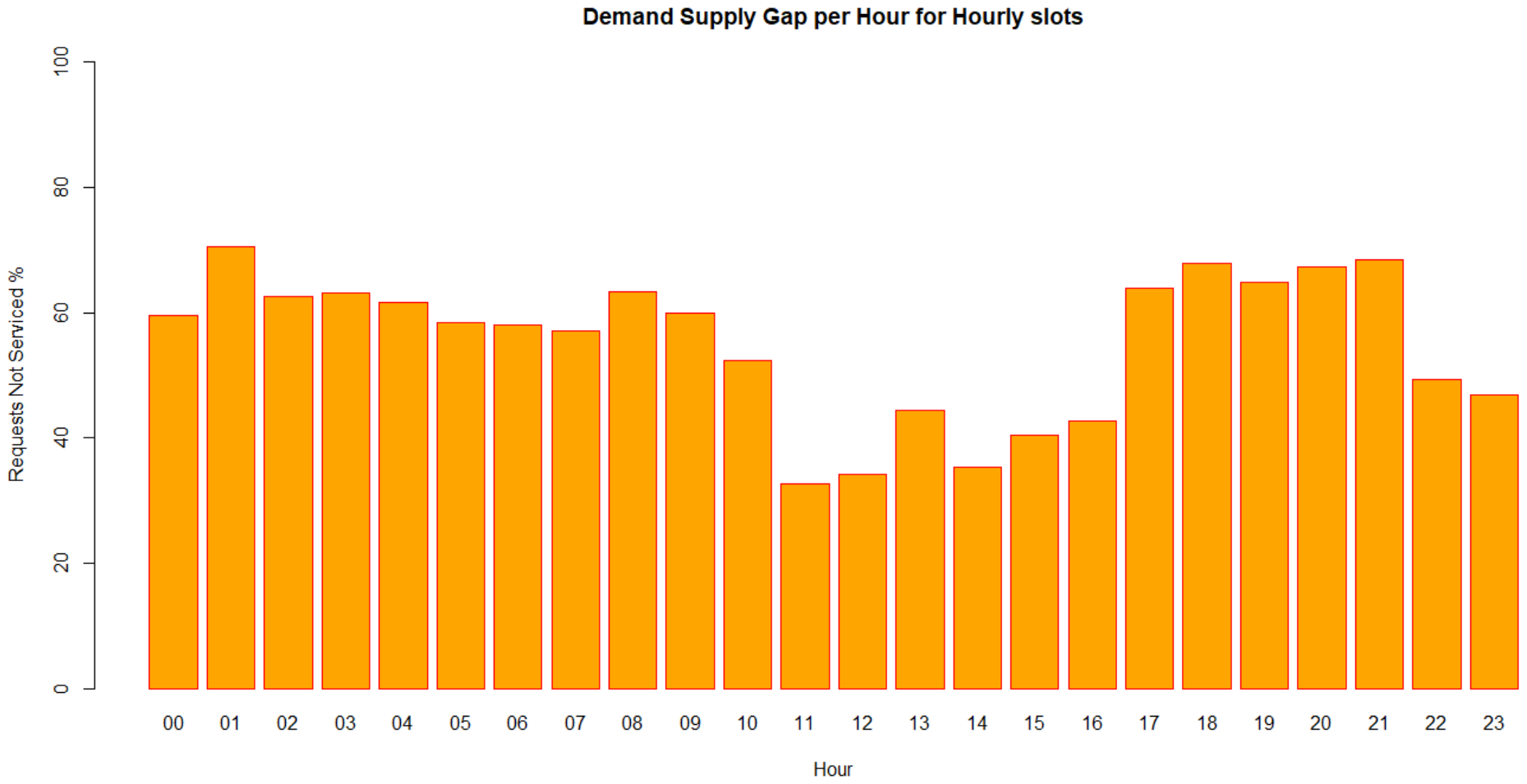




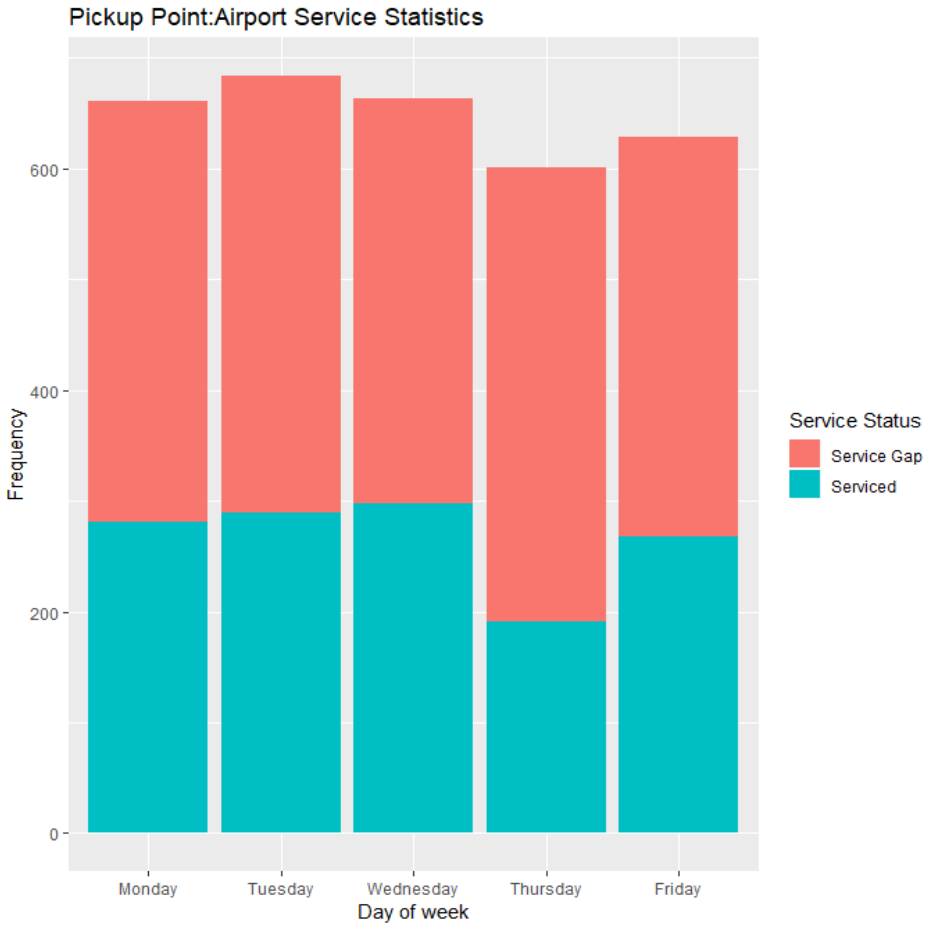
Observations

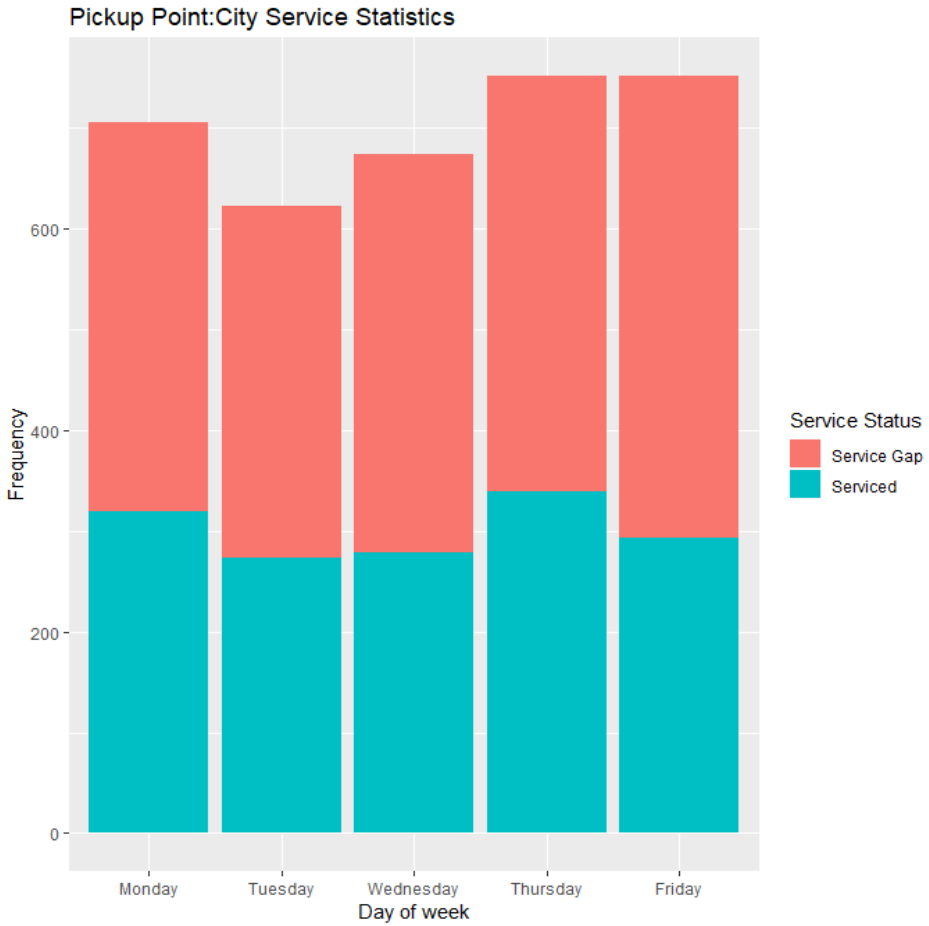
- Service Gap in terms of Requests at Airport is highest in the Evening(5PM~10PM)
- Service Gap in terms of Requests at City is highest in the Morning(5AM~10AM)

Demand Supply Gap% Overall by Hour of the day



Service Statistics by Day of Week

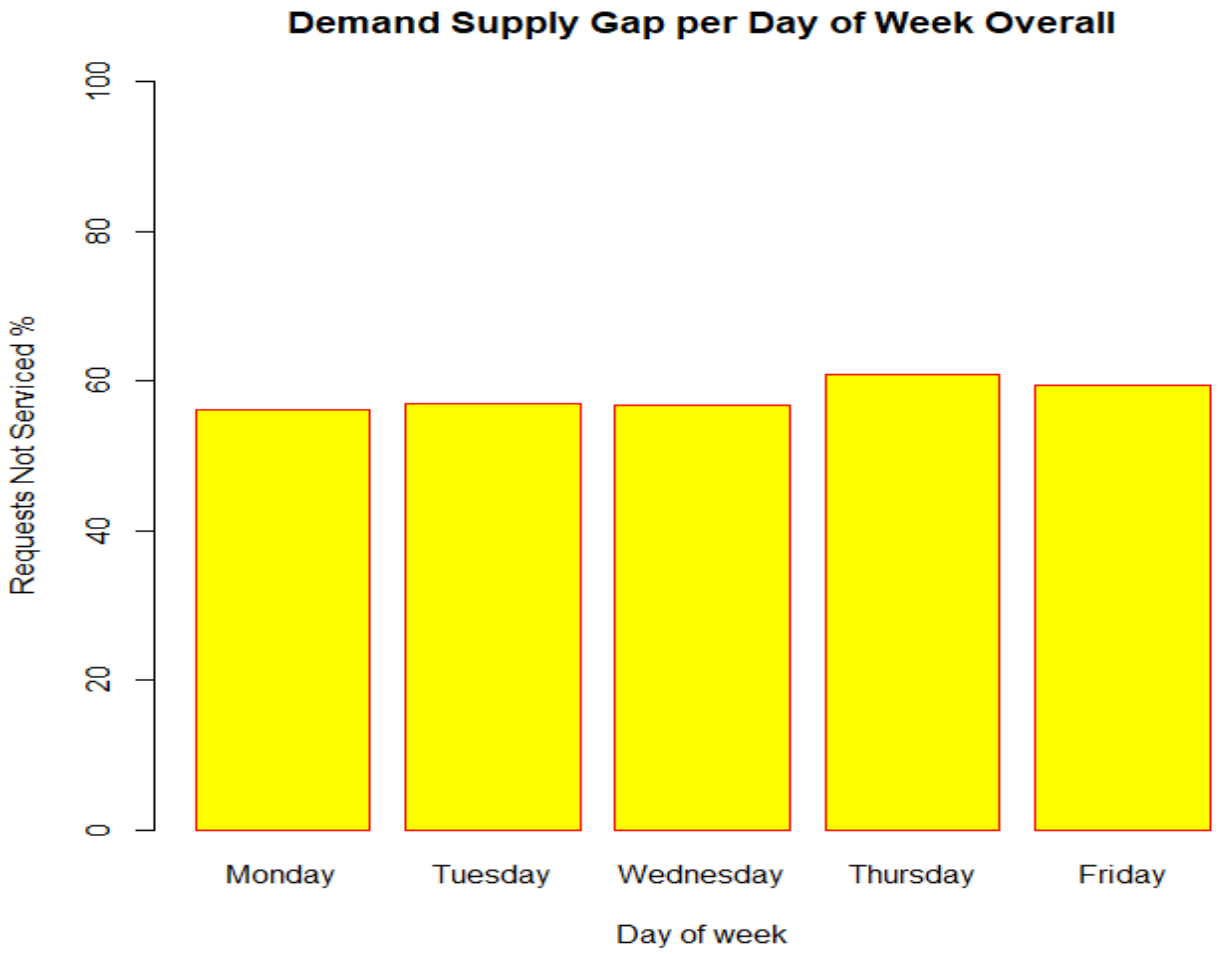




Observations:

- No of Service Requests from Airport decline at end of work week and rise during start of the week.
- No of Service Requests from City rise at end of work week and decline during start of the week.

Demand Supply Gap % Overall by Day of Week



Observations:

Not much change is observed during the week.

3.Reasons for the Demand-Supply Gap

1. Huge Demand-Supply Gap exists for requests in Morning Timeslot from the City because of cancellations by drivers. This might be because they do not get a customer on their way back to the city.

2. Huge Demand-Supply Gap exists for requests in Evening Timeslot from the Airport because of non-availability of cars. This might be because the number of incoming flights and the existing number of cars cannot match.

3. Huge Demand-Supply Gap exists for requests in Early Morning Timeslot due to the fact that the drivers are winding up for the day and do not want to entertain new trips.

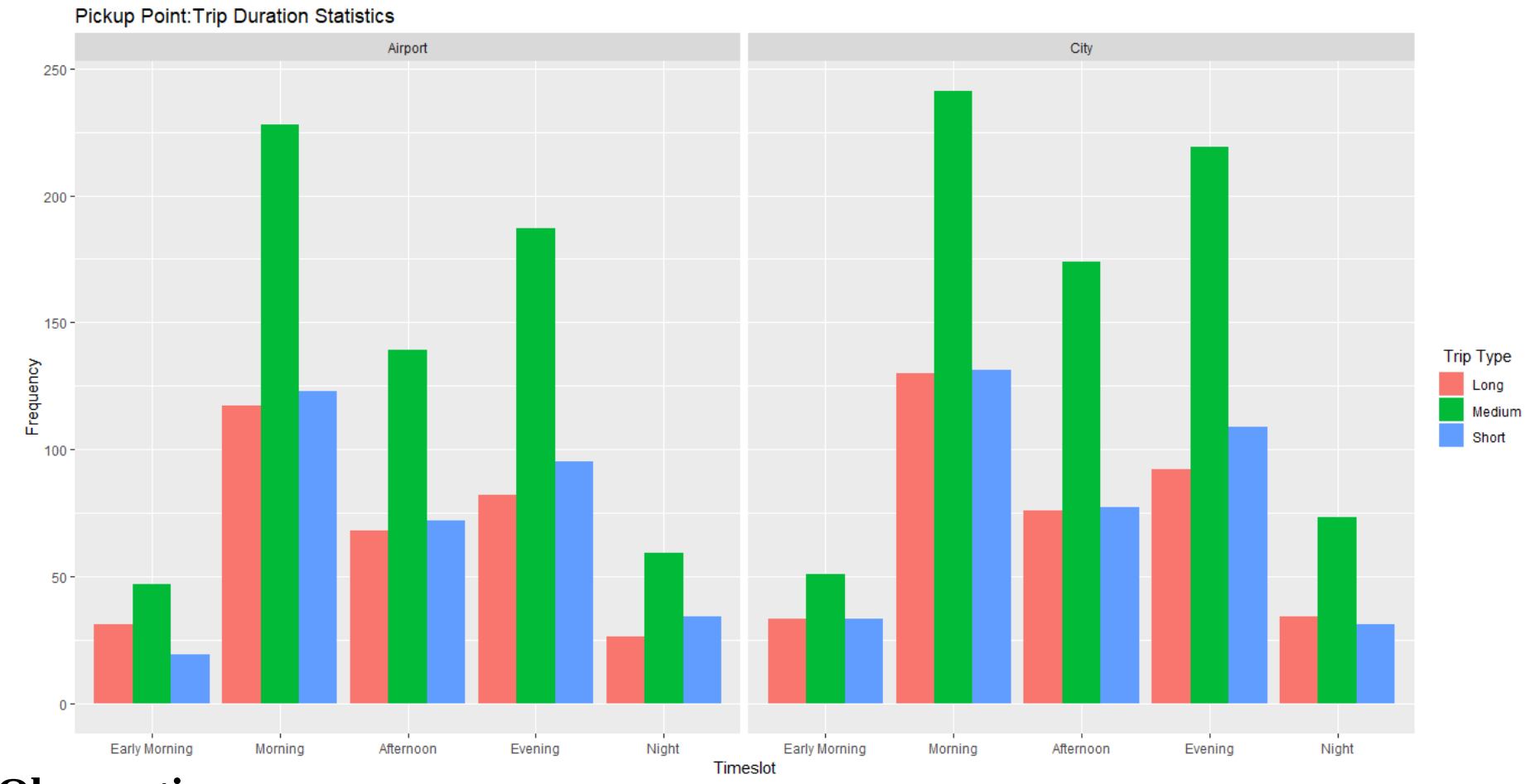
Analysis of Trip Duration

Measure Of Trip Duration	Overall(in Minutes)	Trips from Airport (In Minutes)	Trips from City (In Minutes)
Average	52	52	53
Median	52	52	53
First Quartile	41	41	41
Third Quartile	64	64	64

Classifications of Trips into " Short","Medium","Long" as per following criteria

Trip Duration(In Minutes)	Trip Type
<=41	Short
>41 & <=64	Medium
>64	Long

Trip Duration Statistics by Timeslot



Observations:

- Substantial number of “Short” trips exist during the Morning and Evening Timeslots.
- “Medium” trips are the highest in any timeslot.

4.Ways to resolve the Demand-Supply Gap

Incentive to driver for “Short” trips

Give incentive to drivers who have to undertake "Short" trips(<= 41 min) to/from airport rather than cancelling. This is during the "Morning“, "Evening" slots. Since distance is not much incentive will drive more drivers towards the airport/city means lesser "No Cars Available". We can assume a 25% decrease in "No Cars Available" & "Cancelled"resolve the supply-demand gap.

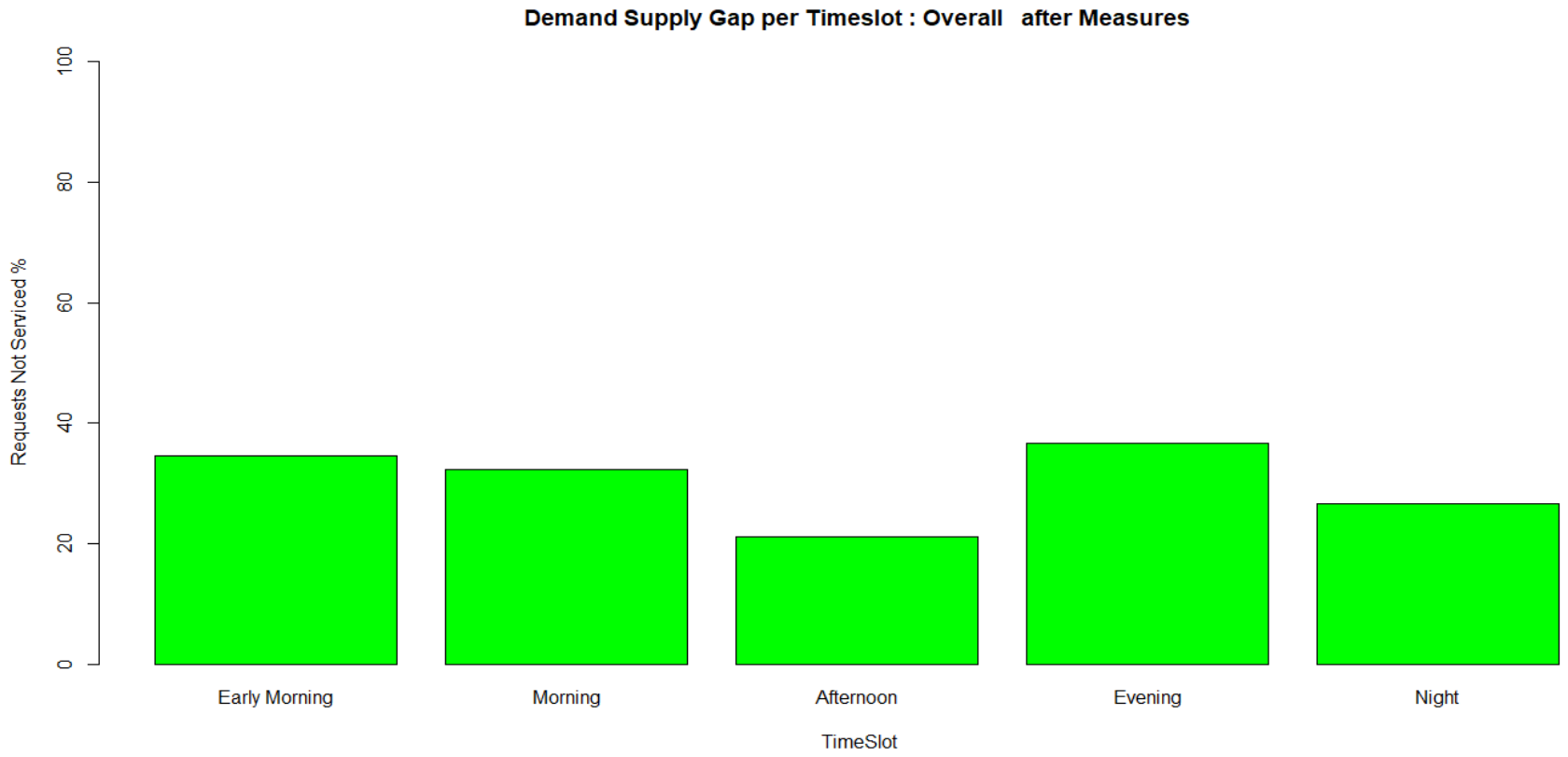
Increase in fleet by 25%

Uber increases fleet by 25%.Increase in drivers has probability for more cancellations and Increase in overhead costs. Assuming a 20 % decrease in "No Cars Available" & "Cancelled“.

Assuming the Impact of both measures would be 45% decrease in "No Cars Available" & "Cancelled“.

Demand-Supply Gap percentage after the measures is on following slide.

Demand Supply Gap% by TimeSlot after Measures



It is seen that the