UBER CASE STUDY SUBMISSION

Submitted by:

Sumanth Sundaramurthy - DDA1730138

Business Objective – Identify the cause and suggest service improvements for cab availability for Uber

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 service for Uber.

As a result of your analysis, we would be able to present Uber some of the root cause(s) and possible hypotheses of the problem(s) and recommend ways to improve service to customers to and from the airport

We shall also share any trends/patterns we identify as part of our analysis

Uber service issues analysis strategy:

The strategy of the analysis is to identify the issues with requests which get cancelled or show “no cars available”; also assess gap of supply vs demand

- Goals of data analysis:
  - Visually identify the most pressing problems for Uber where cabs are unavailable or rides get cancelled during various timeslots of the day
  - Identify and visually show the gap between supply and demand
  - Identify the time slots when the highest gap exists between supply and demand
  - Find the types of requests (city-airport or airport-city) for which the gap is the most severe in the identified time slots
  - Provide insight on the reason for this issue for the supply-demand gap
  - Provide recommendation for ways to resolve the supply-demand gap

Problem Solving Methodology

Data Collection

- Uber request raw data "Uber Request Data.csv" was downloaded
- Raw data was loaded into R Studio

Data Cleaning

- Necessary data cleaning techniques were followed such as:
  - Formatting of Request and Drop dates using parse_date_time method from lubridate package
  - The remainder of data did not require any additional cleaning

Data Analysis

- Univariate analysis:
  - Analyzed the dataset based on Request time
- Segmented analysis:
  - Analyzed the dataset based on
    - Timeslot of day
    - Origin (Airport/City) of request

Data Visualization

- Plotted graphs in R and Tableau to visualize
  - Requests by day
  - Requests in timeslots by location of request (Airport/City)
  - Plotted Demand vs. Supply vs. Gap to identify the problematic areas for service

Request Date analysis

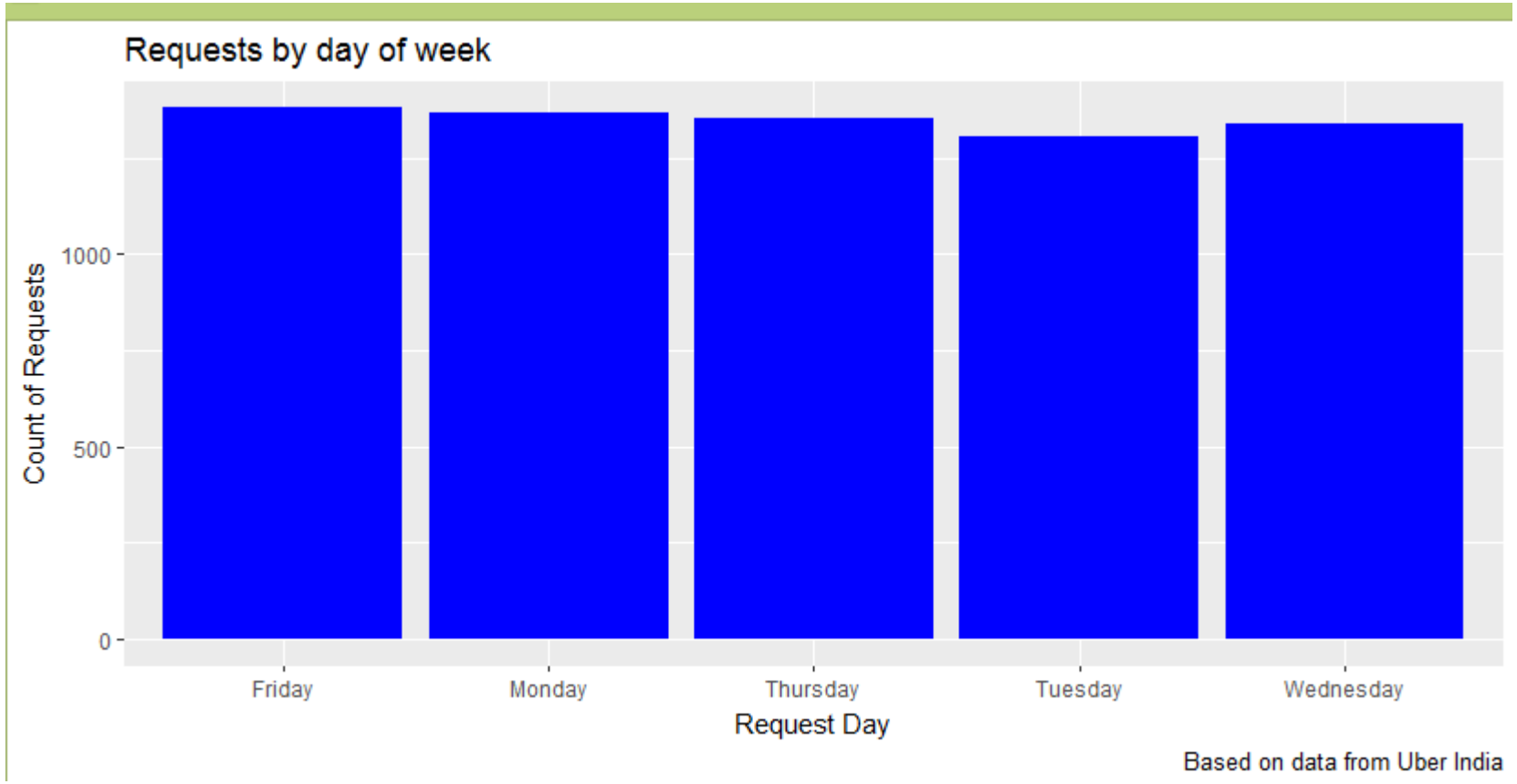
Univariate type analysis in R:

Derived the day of week - to see if there are any patterns in request frequency by day

Analysis result:

a) The Uber requests seem to be more or less constant during the week
b) There is no weekend data in the dataset provided

Funding type results



Based on our analysis in R there is no specific trend in Request by day

Request Date analysis

Segmented analysis in R:

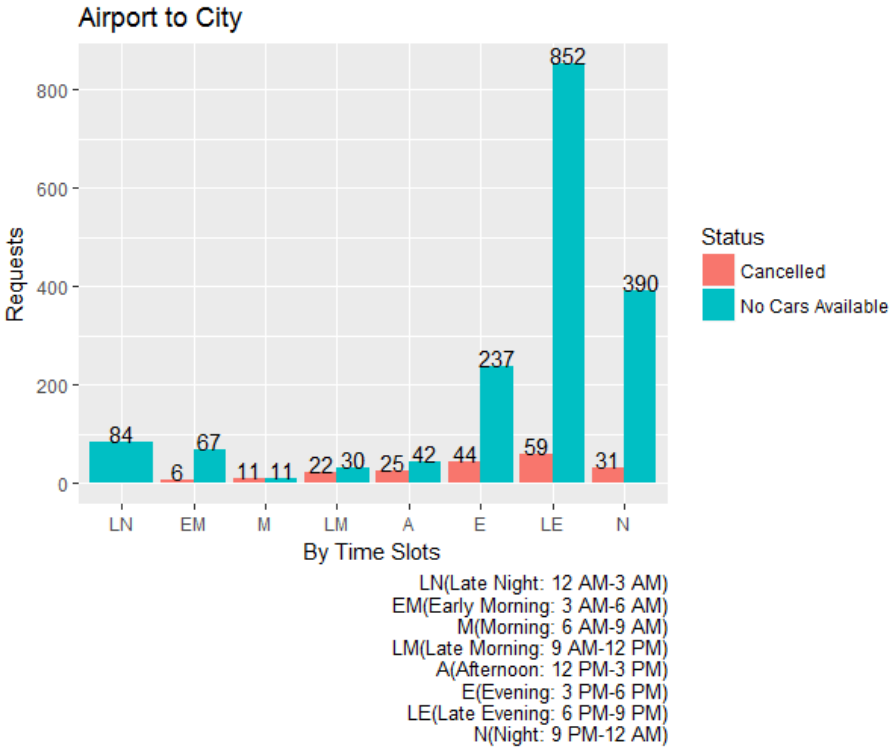
Segmented the data into timeslots and origin of request (Airport/City) - to see if there are any patterns in request frequency by day

Analysis result:

Based on our analysis in R we can see that there is a HIGH number of cancellation and unavailability of cabs at

1. Airport – in the Late evening (6 PM – 9 PM) slot
2. City – in the Morning (6 AM – 9 AM) slot

Request time and Origin of request results



Based on our analysis in R we can see that there is a HIGH number of cancellation and unavailability of cabs at

Airport – in the Late evening (6 PM – 9 PM) slot

City – in the Morning (6 AM – 9 AM) slot

Supply vs. Demand analysis

Supply vs. Demand analysis in R:

Segmented the data into timeslots and calculated the Demand (all kind of requests), Supply (where status of trip is completed) and Gap (Supply – Demand) - to see if there are any patterns in request frequency by day

Analysis result:

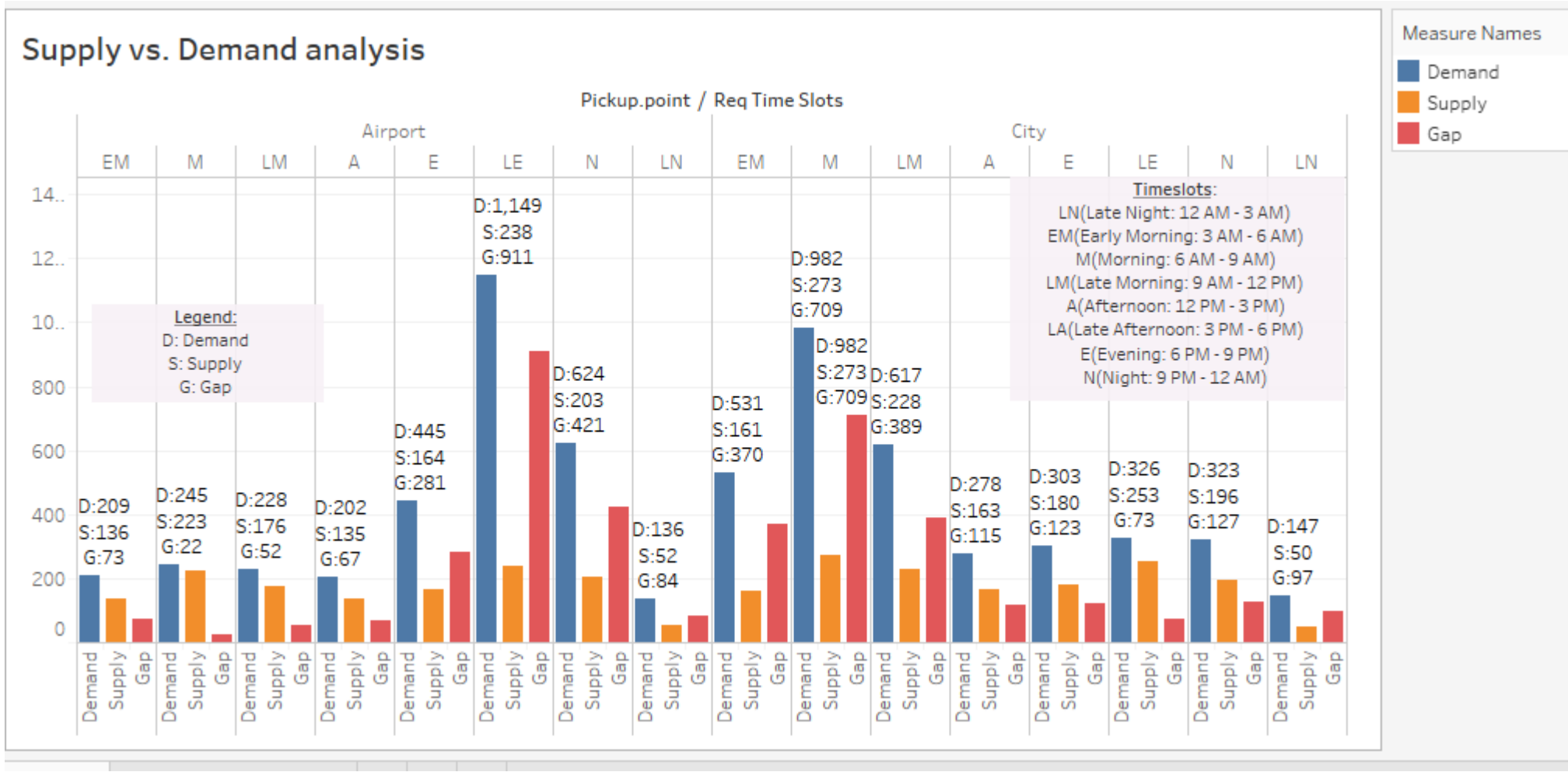
Based on our analysis in R we can see that there is a VERY HIGH number of cancellation and unavailability of cabs

1. From Airport – in the Late evening (6 PM – 9 PM) slot
2. From City – in the Morning (6 AM – 9 AM) slot

Based on our analysis in R we can see that there is a HIGH number of cancellation and unavailability of cabs

1. From City – in the Early Morning (3 AM – 6 AM) slot
2. From City – in the Late Morning (9 AM – 12 PM) slot

Supply vs. Demand results



Analysis recommendations

- Based on the Uber data analysis our recommendation is for Uber to enable more cabs to be available during the below mentioned time slots at the mentioned location of requests

- Based on our analysis in R we can see that there is a VERY HIGH number of cancellation and unavailability of cabs
    1. From Airport – in the Late evening (6 PM – 9 PM) slot
    2. From City – in the Morning (6 AM – 9 AM) slot

- Based on our analysis in R we can see that there is a HIGH number of cancellation and unavailability of cabs
    1. From City – in the Early Morning (3 AM – 6 AM) slot
    2. From City – in the Late Morning (9 AM – 12 PM) slot