EXPLORATORY DATA ANALYSIS OF UBER SUPPLY DEMAND GAP

By:
Abhirag R. Nagpure

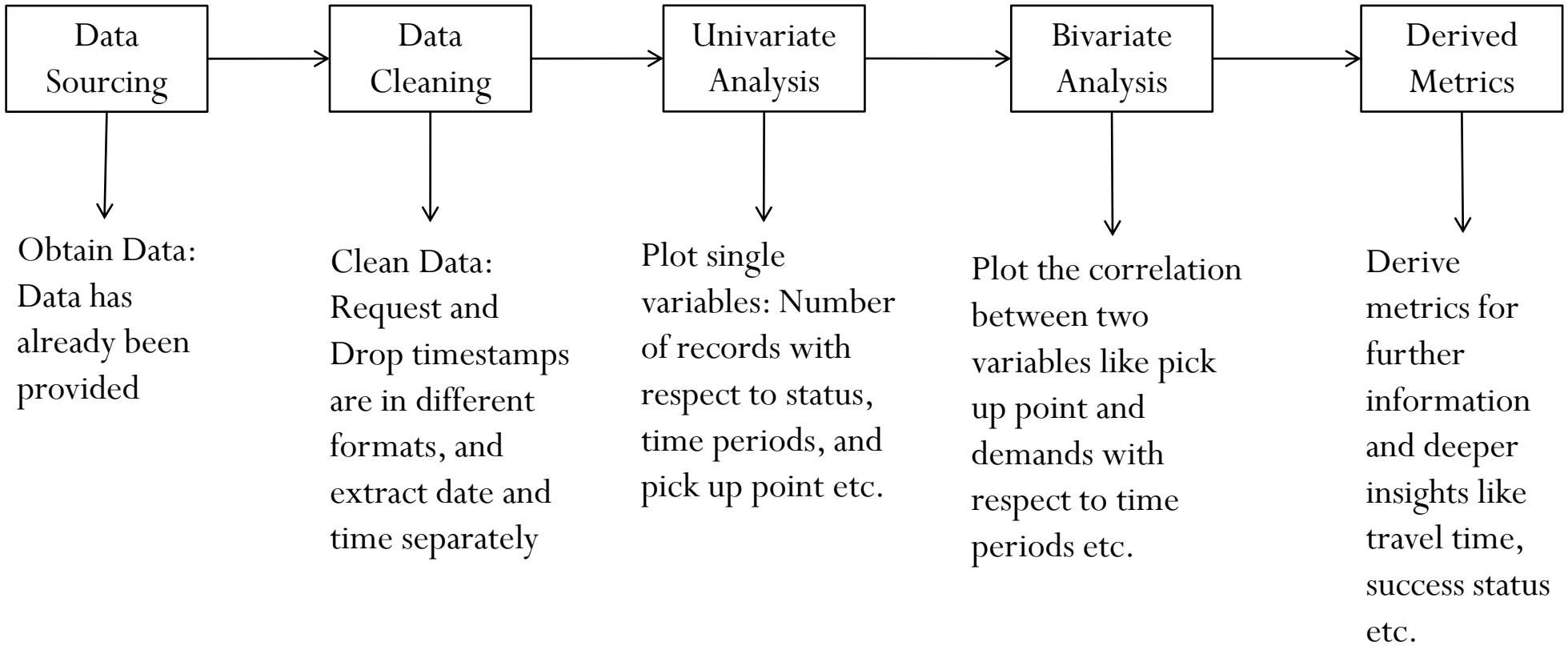
Problem Statement and Objectives.

The data is a subset of the original Uber dataset which shows the details of cab requests made by customers. The main problems faced by both the company and its customers are:

1. Non-availability of cabs
2. Cancellation of cab requests

Our task as an analyst is to identify the cause of these problems, with respect to the pick up points and also the time periods. We are also required to establish the supply-demand gap and recommend solutions based on our findings.

FLOW CHART ENCOMPASSING ENTIRE PROCESS.



Primary Univariate Analysis

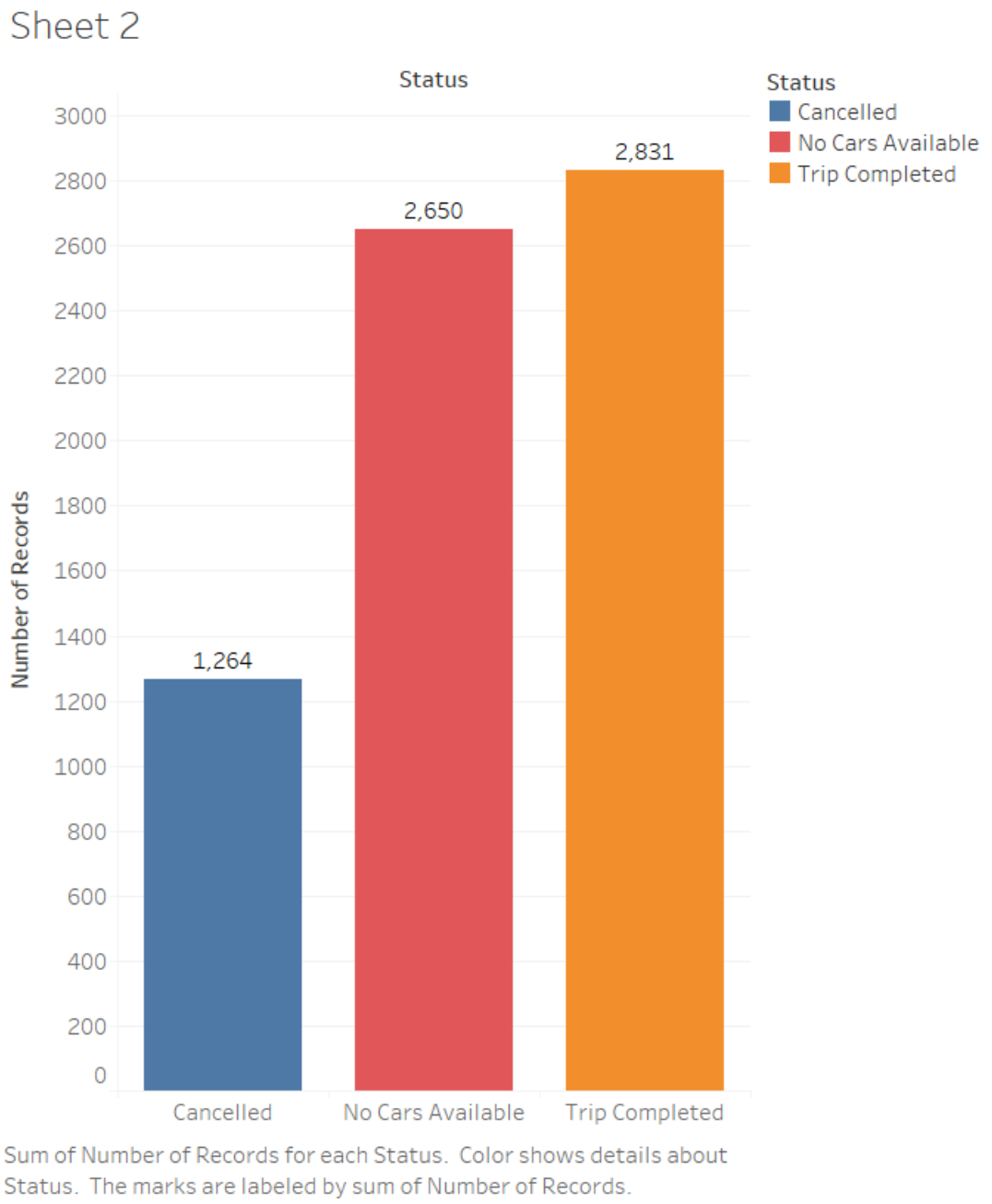
The number of cab requests according to their status:

The numbers are perfectly clear:

Trip Completed – 2831

No Cars Available – 2650

Cancelled – 1264



Derived Metrics
Univariate Analysis

The number of cab requests according to their request timestamp (hours extracted from TS):

Based on the hour of the day, time_period has been defined as:

12 AM to 4 AM – late night

4 AM to 8 AM – early morning

8 AM to 12 PM – morning

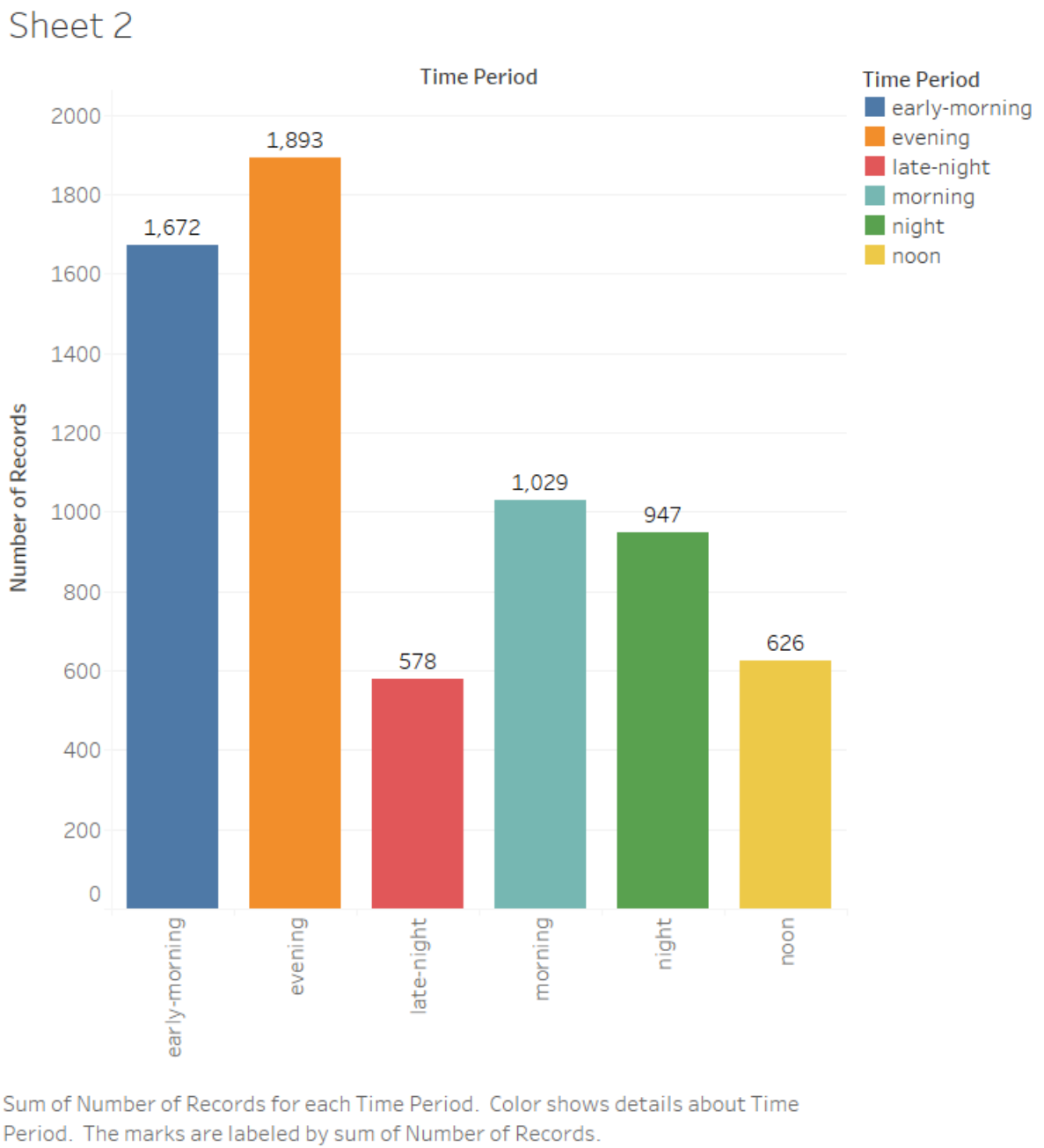
12 PM to 4 PM – noon

4 PM to 8 PM – evening

8 PM to 12 AM – night

Most requests – evening

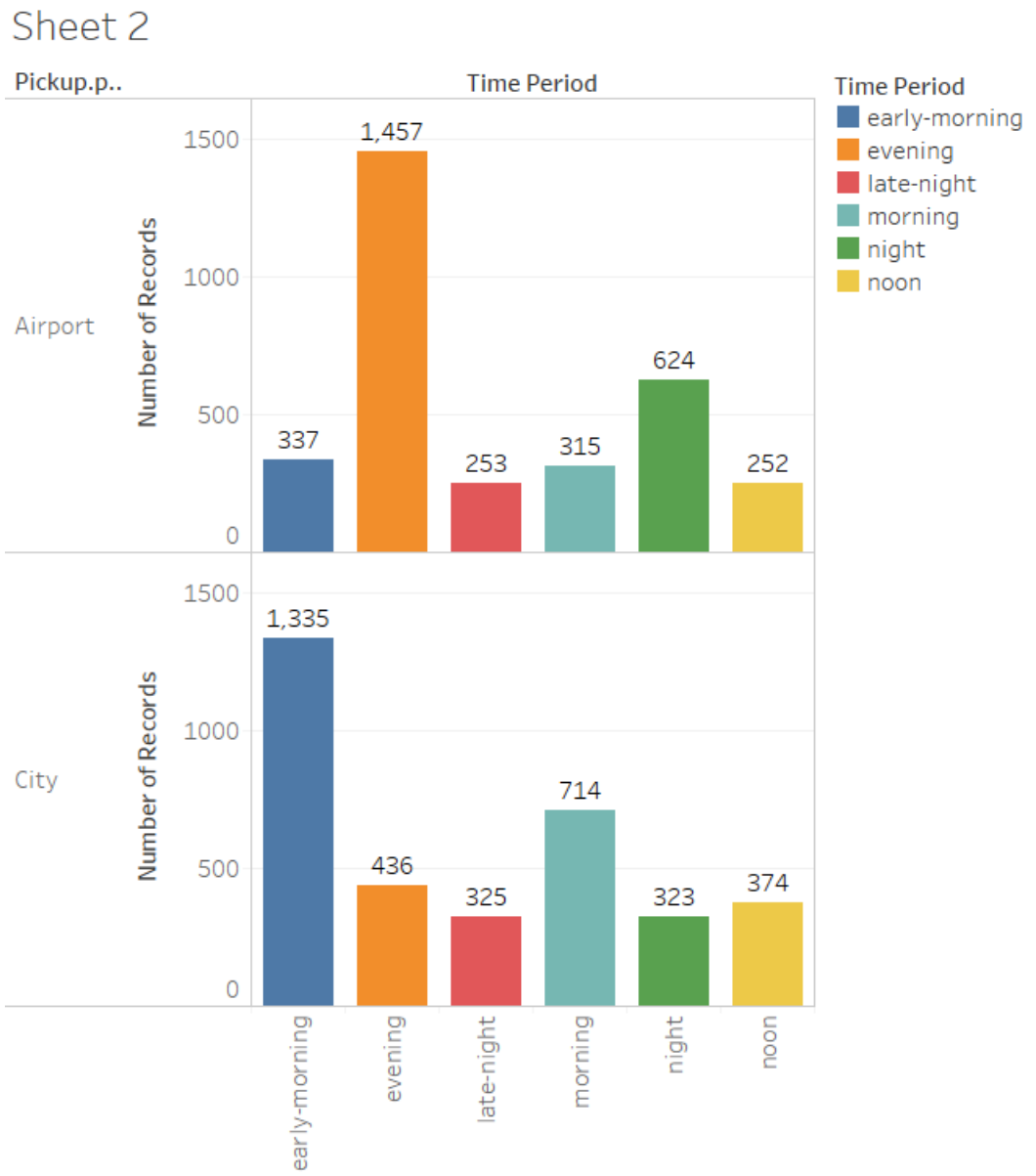
Least requests – late-night



Derived Metrics Primary Bivariate Analysis

The number of cab requests according to their pick up point and time period:

	Airport	City
Maximum	**Evening(1457)**	**Early morning(1335)**
Minimum	**Night(323)**	**Noon(252)**



Extensive Bivariate Analysis

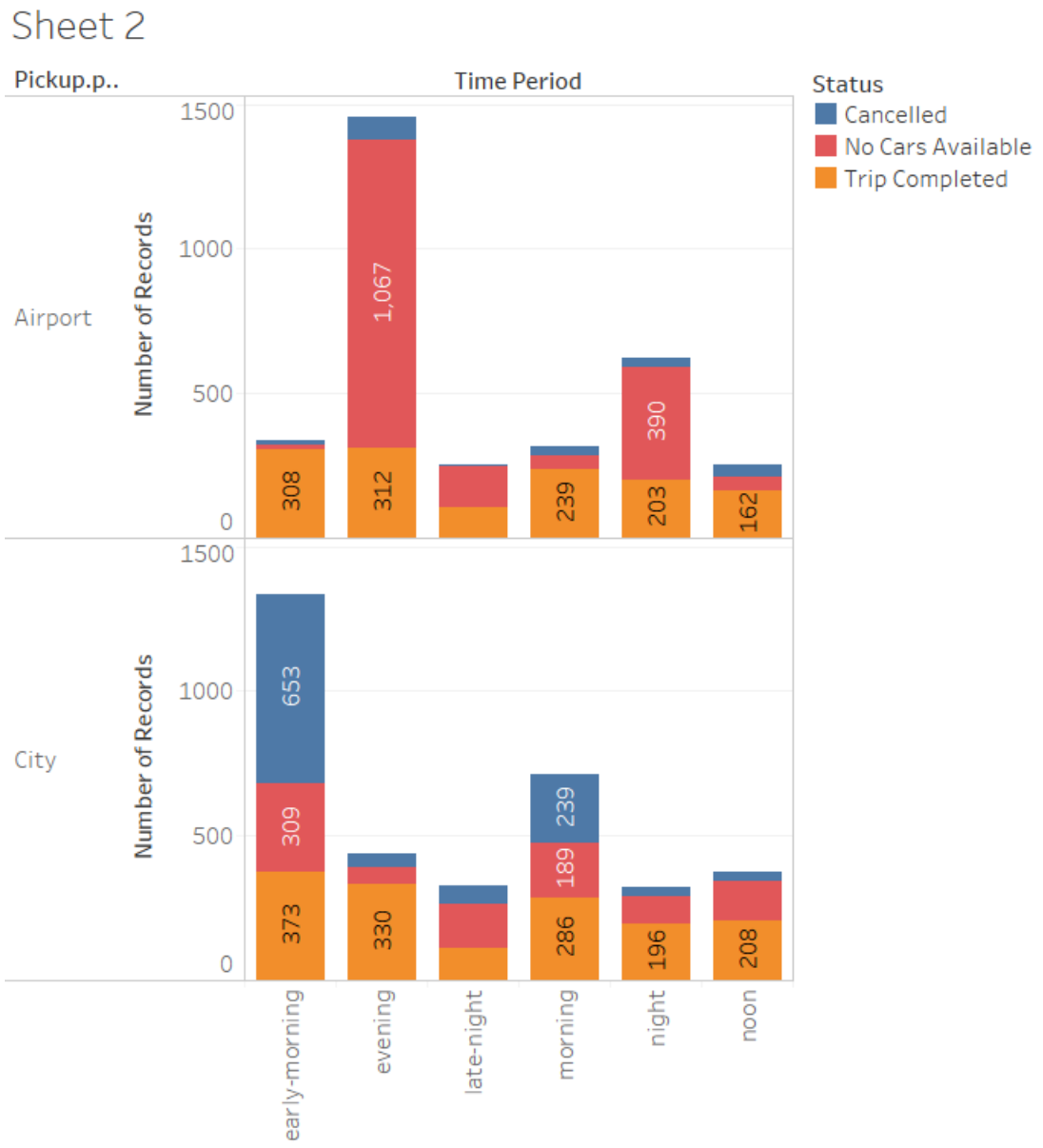
The distribution of Status in accordance with both the pick up point and the time period:

Maximum shortage – evening, at Airport(1067)

Maximum cancellations – early morning, at City(653)

Minimum shortage – early morning, at Airport(14)

Minimum cancellations – late night, at Airport(2)



Sum of Number of Records for each Time Period broken down by Pickup.point. Color shows details about Status. The marks are labeled by sum of Number of Records.

Supply Demand Analysis

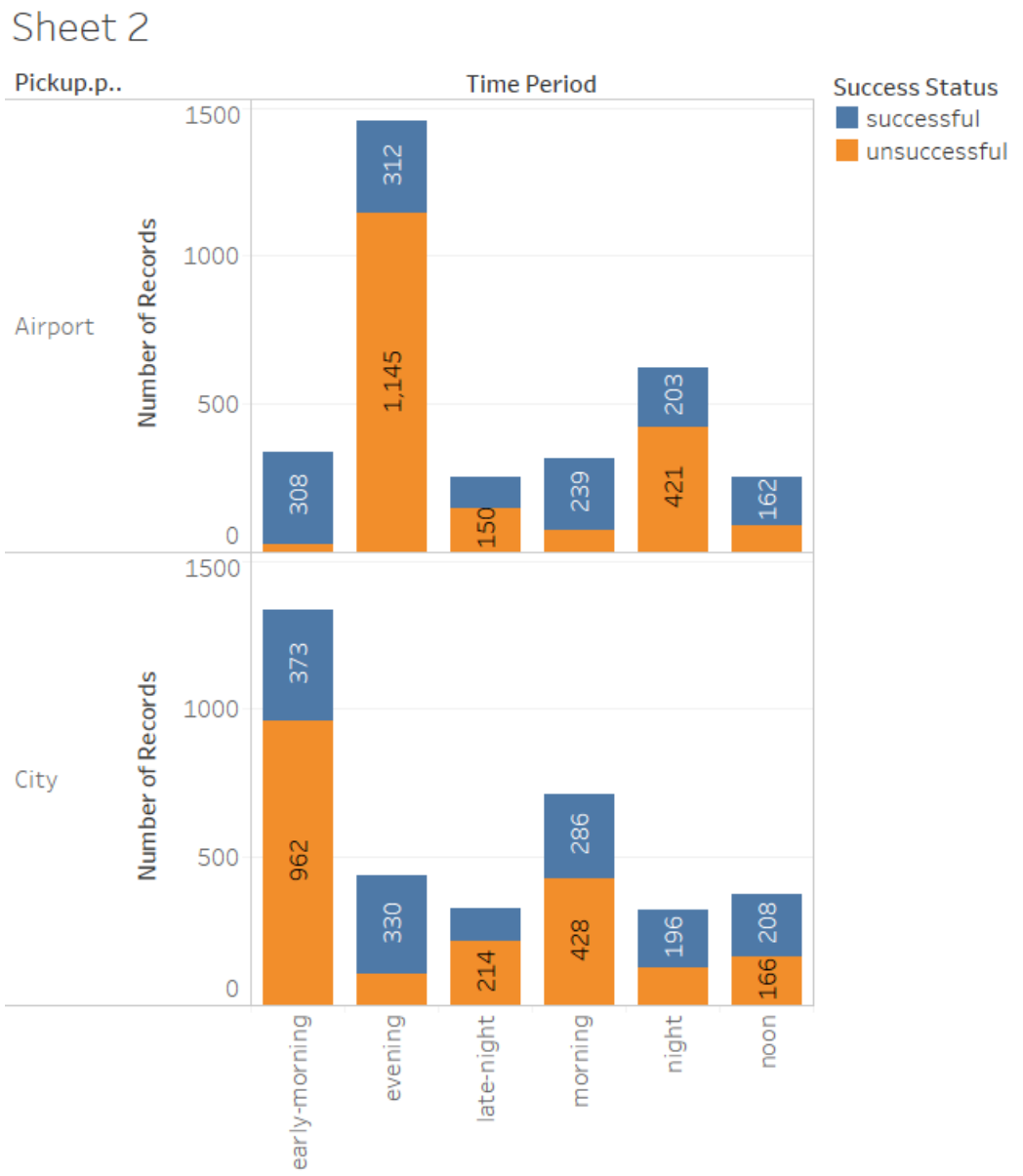
The trips have been separated into 2 categories depending on their status:

Successful –Trip Completed

Unsuccessful – Cancelled/No Cars Available

Maximum supply demand gap: Airport, in evening(1145)

Minimum supply demand gap: Airport, in early morning(29)



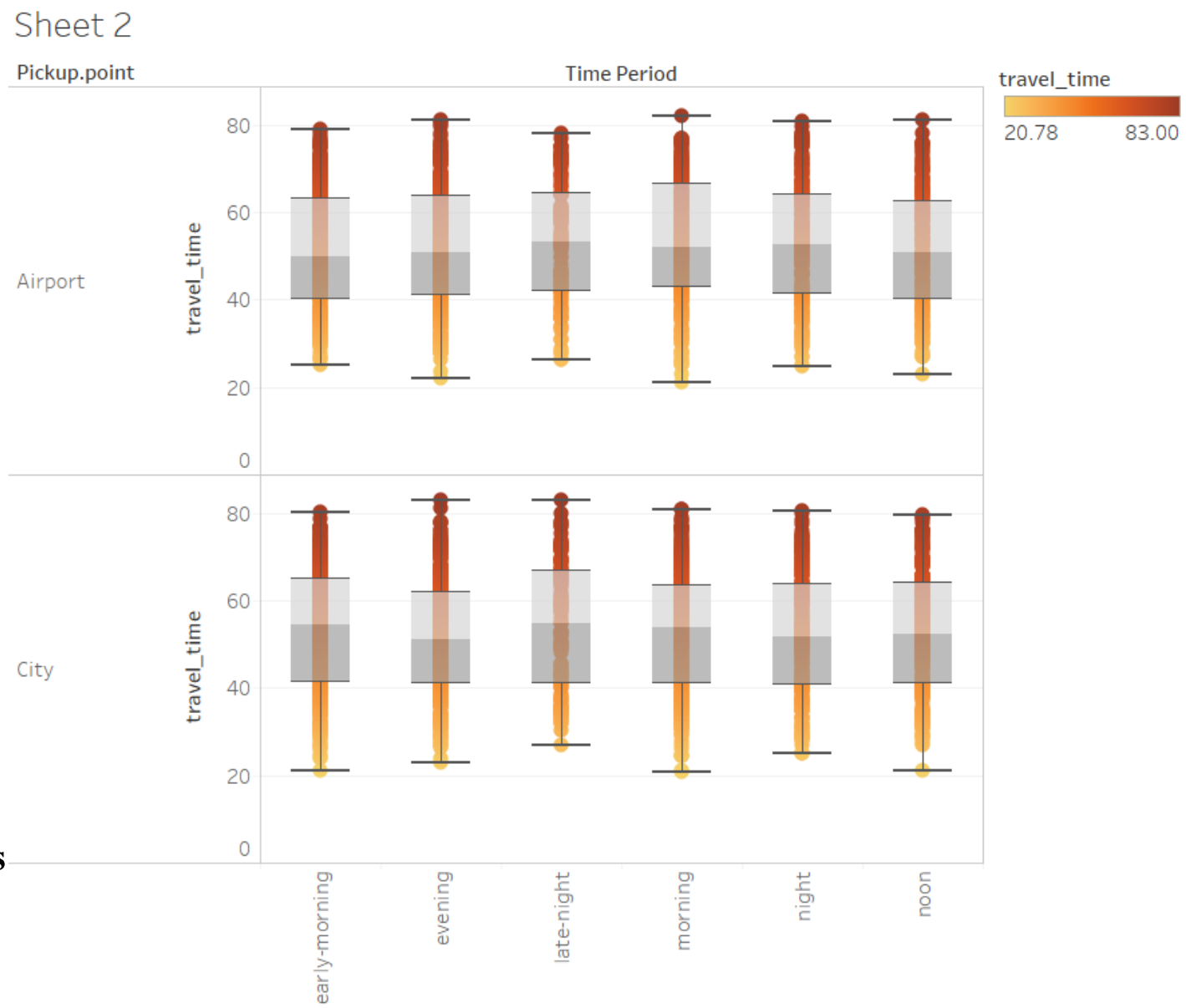
Sum of Number of Records for each Time Period broken down by Pickup.point. Color shows details about Success Status. The marks are labeled by sum of Number of Records.

Travel time Analysis

The travel time was obtained by taking the difference between the request timestamp and the drop timestamp.

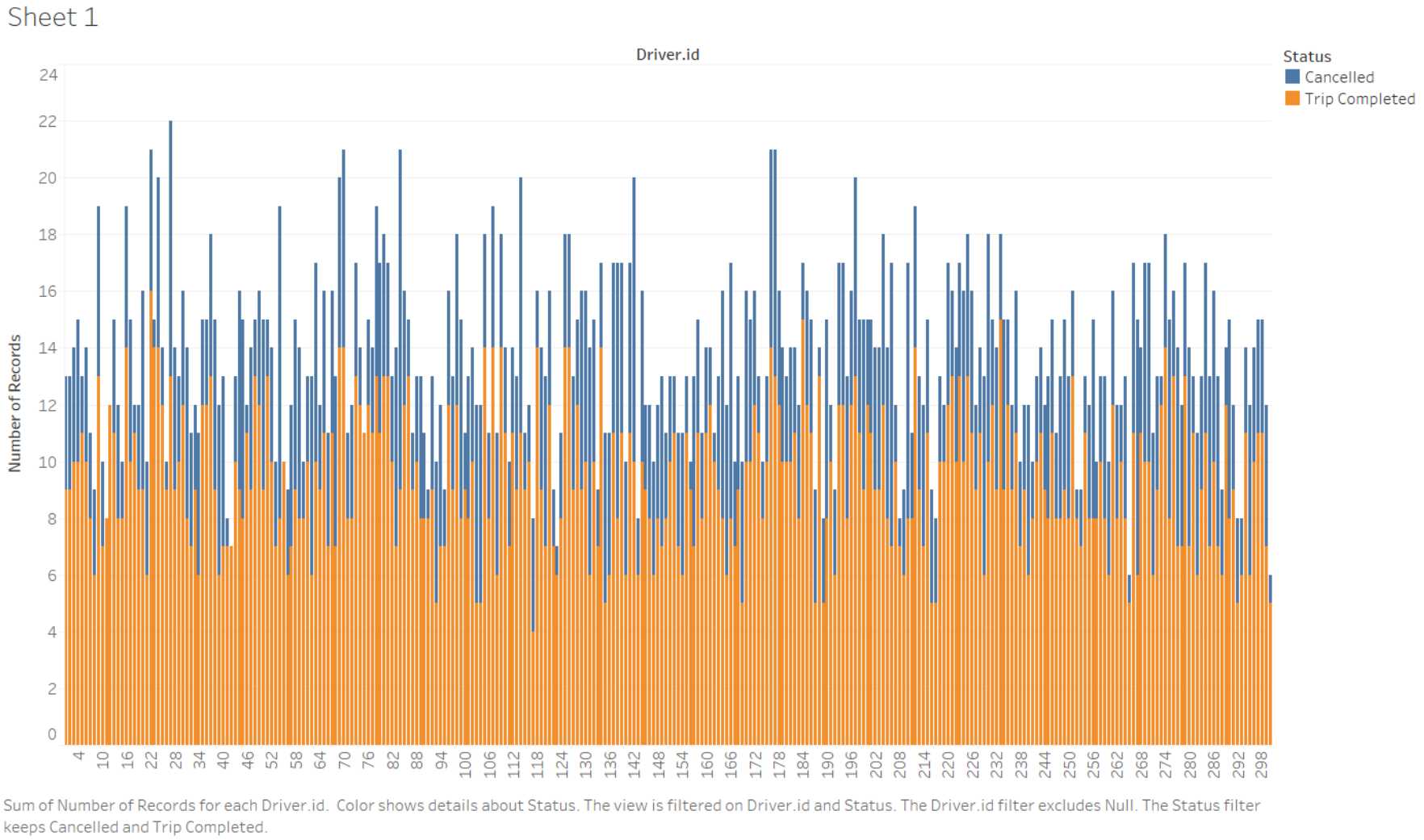
Obviously, travel time would be calculated only for successful trips.

There is not much of a difference between travel times due to pick up point or even the time period. Almost all trips will have a median travel time of 52.08 minutes.



Request Cancellation by Drivers Analysis

Some drivers (driver IDs) are more likely to cancel trips as compared to others.This is a potential loss of income to the company.



Inferences

There exists a **huge supply demand gap in the evening, at the airport and early morning, at the city**. These periods can be termed as peak periods i.e., demand is very high. The gap may exist due to **factors like long waiting times, unavailability of cars**, etc.

Also, from the exploratory data analysis which has been carried out, it is evident that **request cancellation and unavailability of cars are two major issues** impacting both the customers and the company revenues.

Also, it can be concluded that **traffic does not play a vital role in this analysis**, since all trips to and from the airport to the city have a uniform **median distribution of 52 minutes.**

Recommendations for reducing the gap:

Drivers who cancel requests very frequently **should be fined** so that they do not tend to cancel requests more often.

The **number of taxis waiting nearby each pick up point should be according to the demand during that time period**. For eg., if in the early morning period, airport requests are 1000 and city requests are 200, there should be more number of taxis available near the airport than the city. This would compensate for non-availability of cars.