To Chief Data Scientist

UBER CASE STUDY

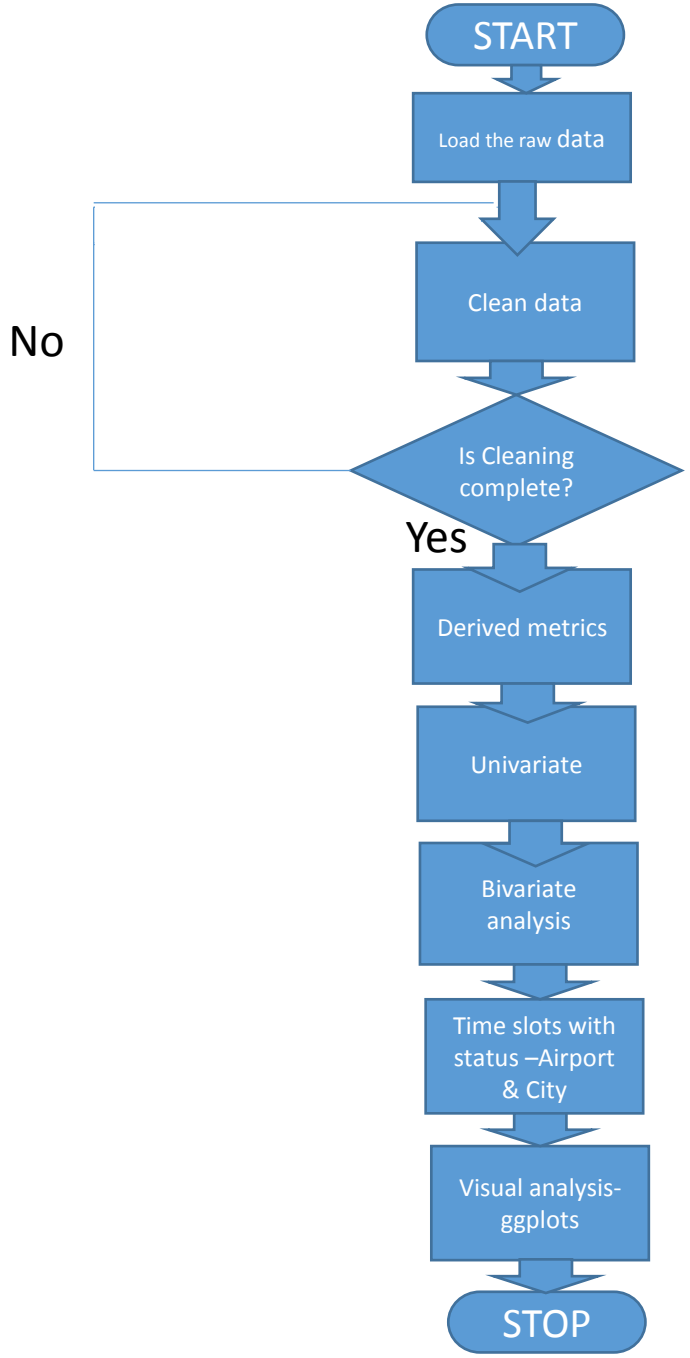
BY

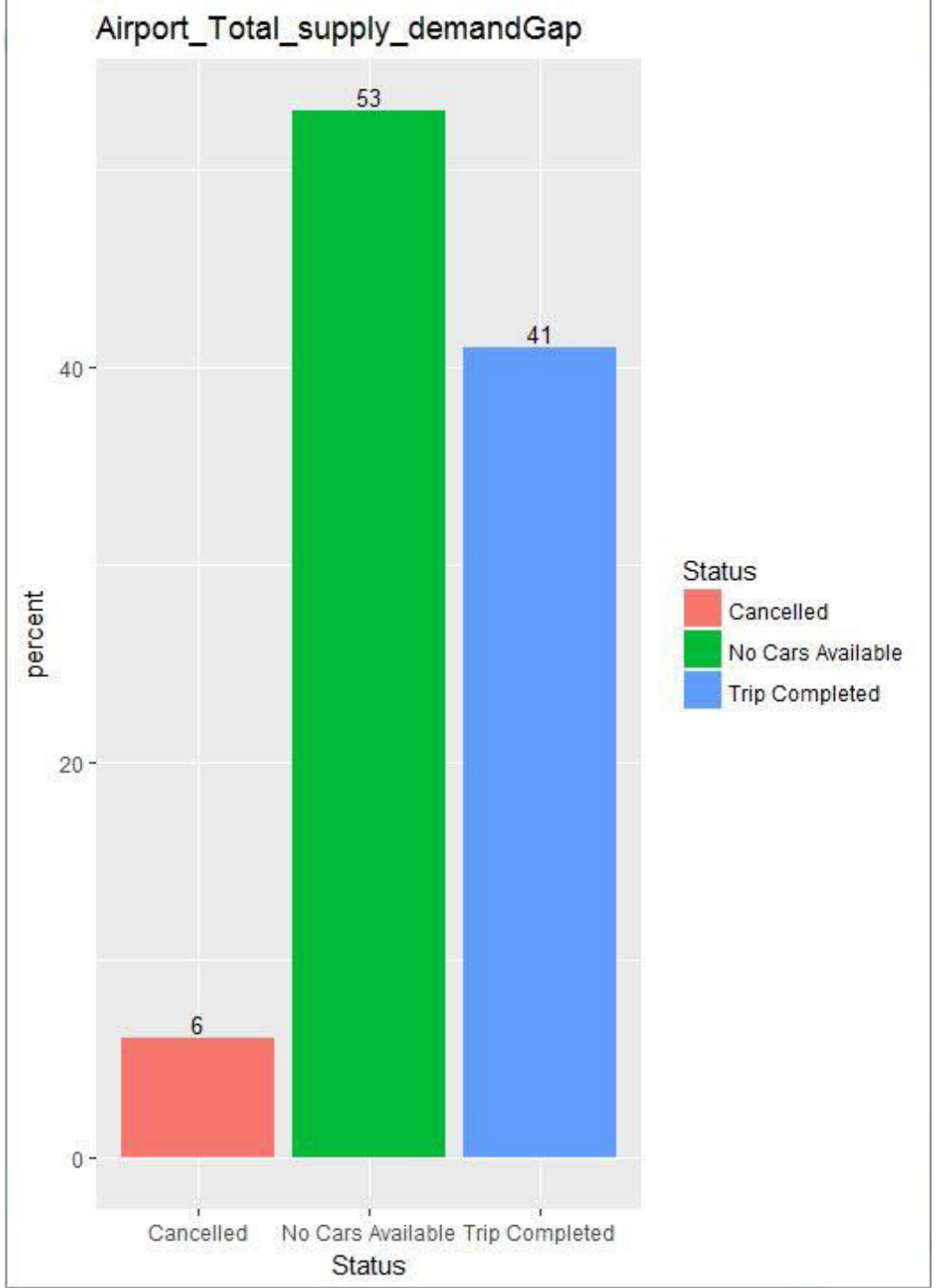
ASHWIN SURESH

Uber is cab service company.

Business Objectives

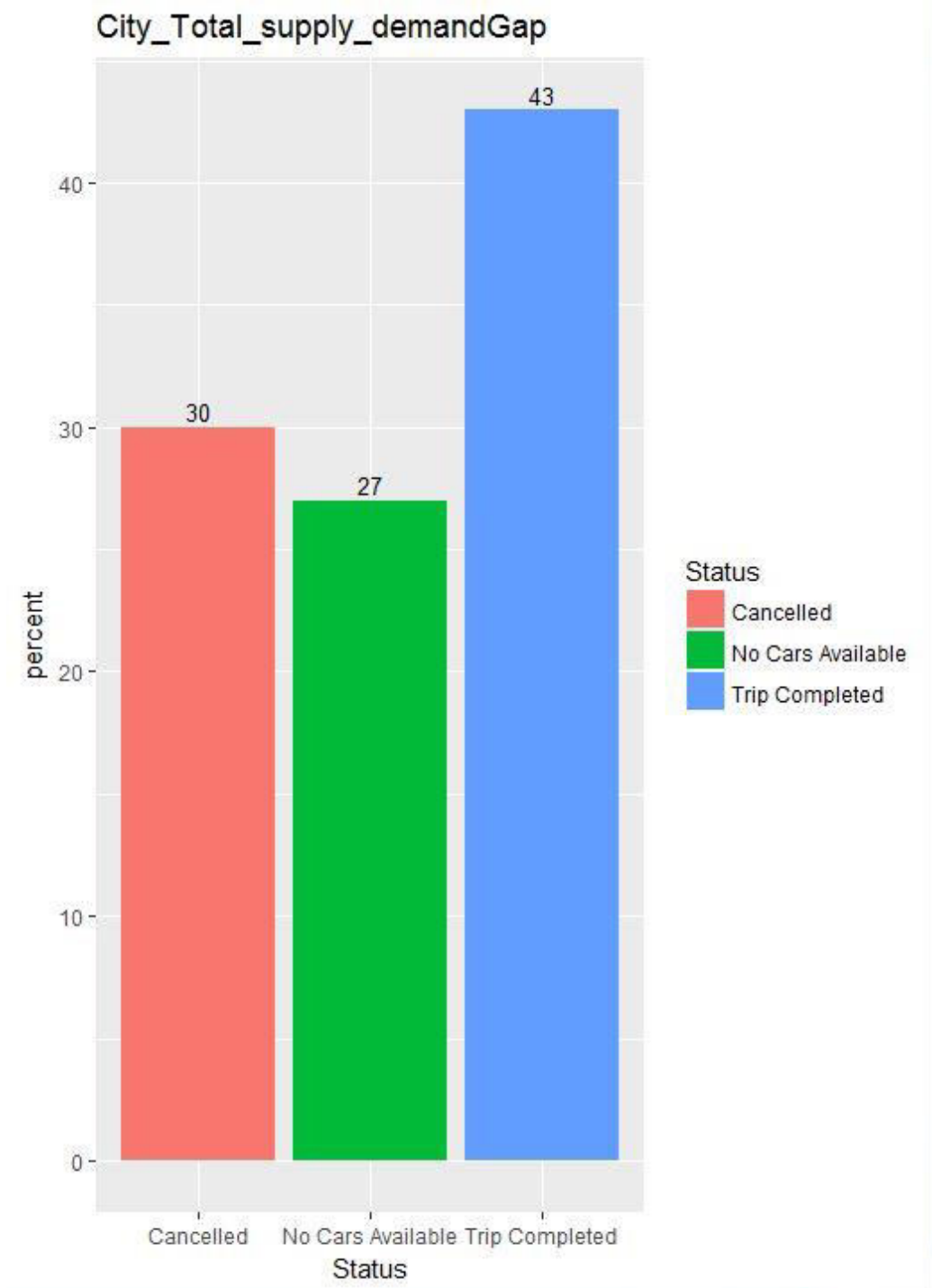
The aim of analysis is to identify the root cause of the problem (i.e. cancellation and non-availability of cars) and recommend ways to improve the situation. As a result of your analysis, you should be able to present to the client the root cause(s) and possible hypotheses of the problem(s) and recommend ways to improve them.


START
Load the raw data
Clean data
No
Is Cleaning complete?
Yes
Derived metrics
Univariate
Bivariate analysis
Time slots with status –Airport & City
Visual analysis- ggplots
STOP

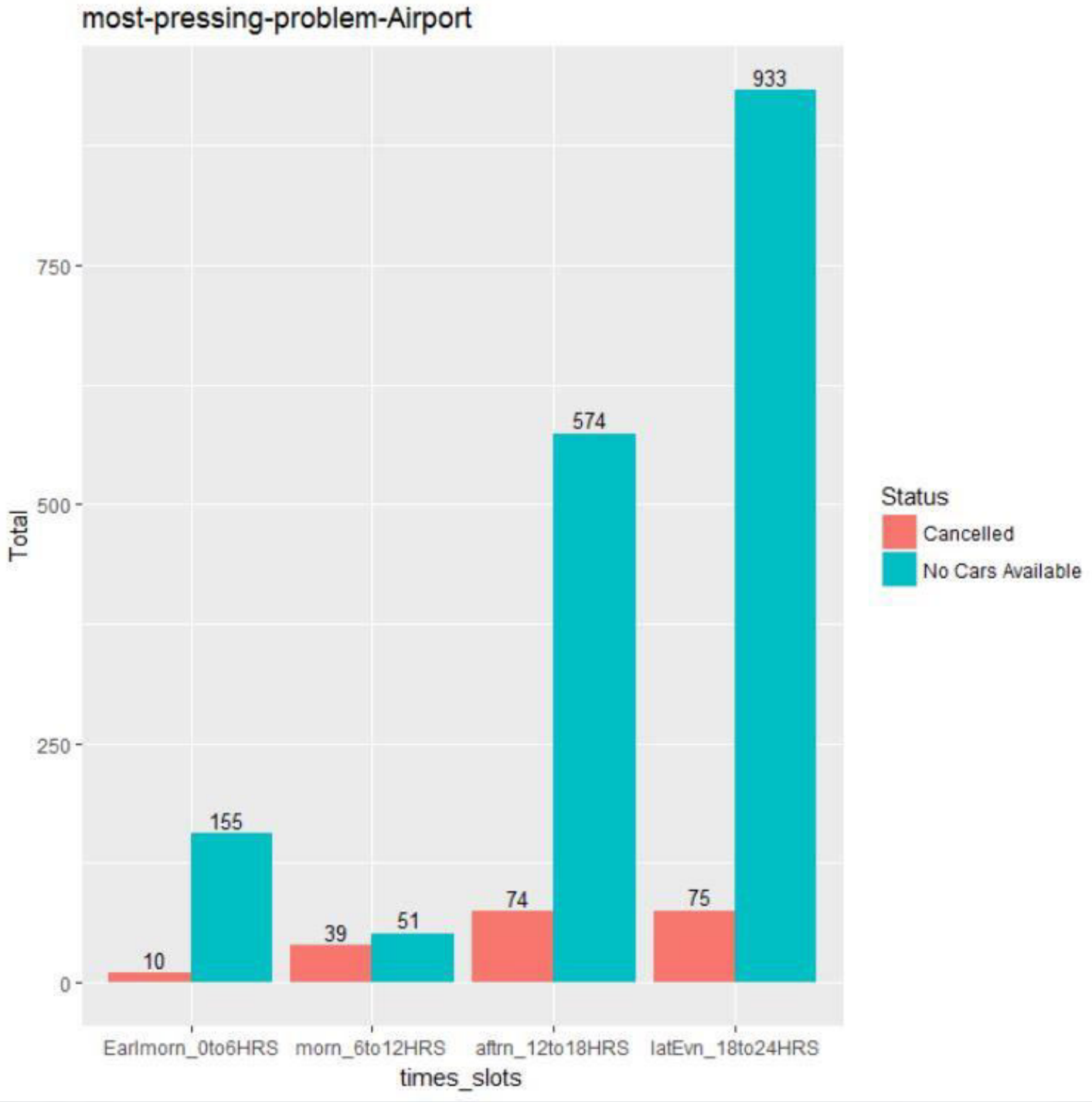


("Trips completed","No Cars Available","Cancelled“) = (1327,1713,198)

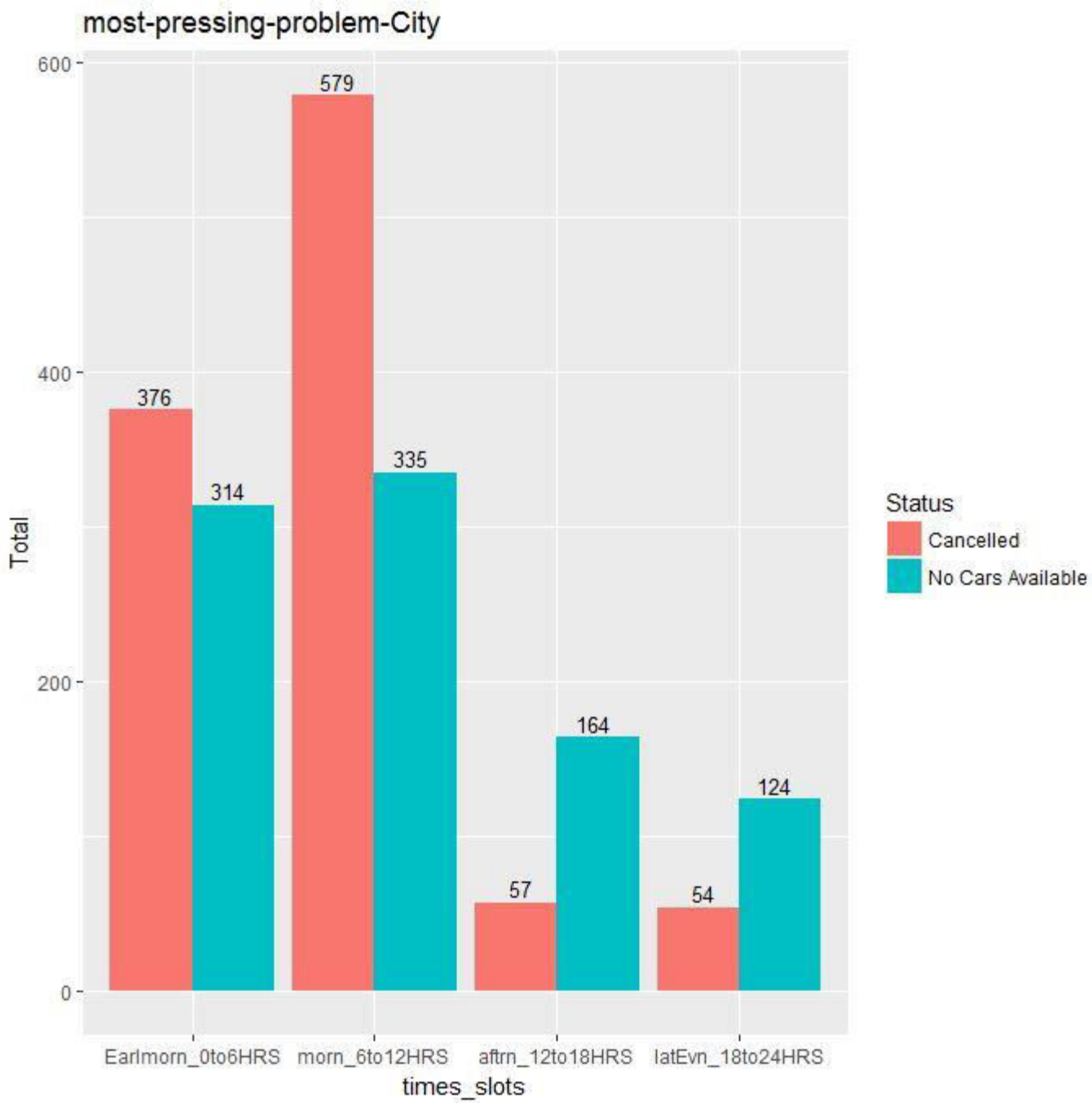
The supply demand gap is nearly 59% critical area =" No Cars Available"



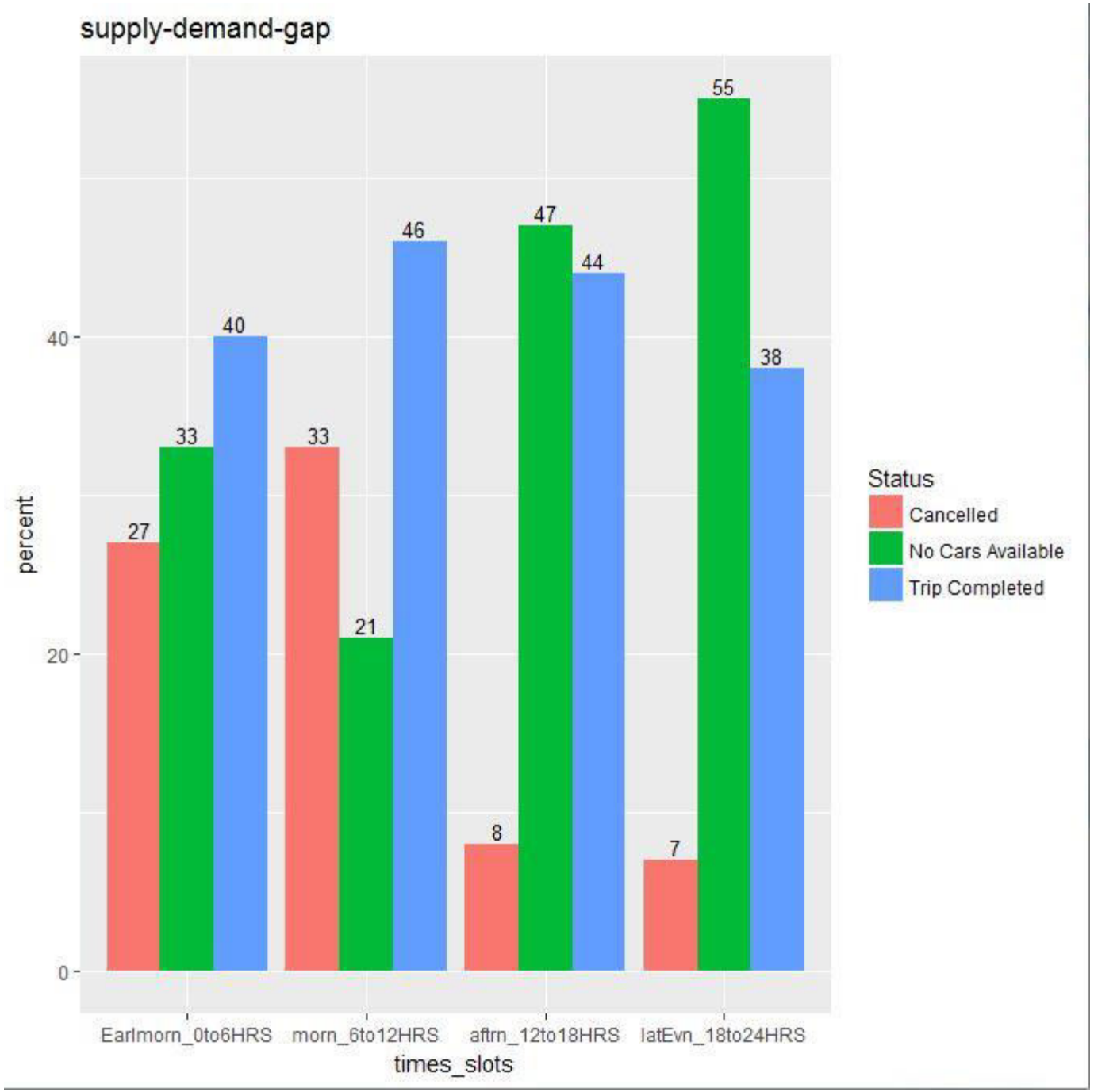
("Trips completed","No Cars Available","Cancelled") = (1504,937,1066)
The supply demand gap is nearly 57%
critical area =" Cancelled"



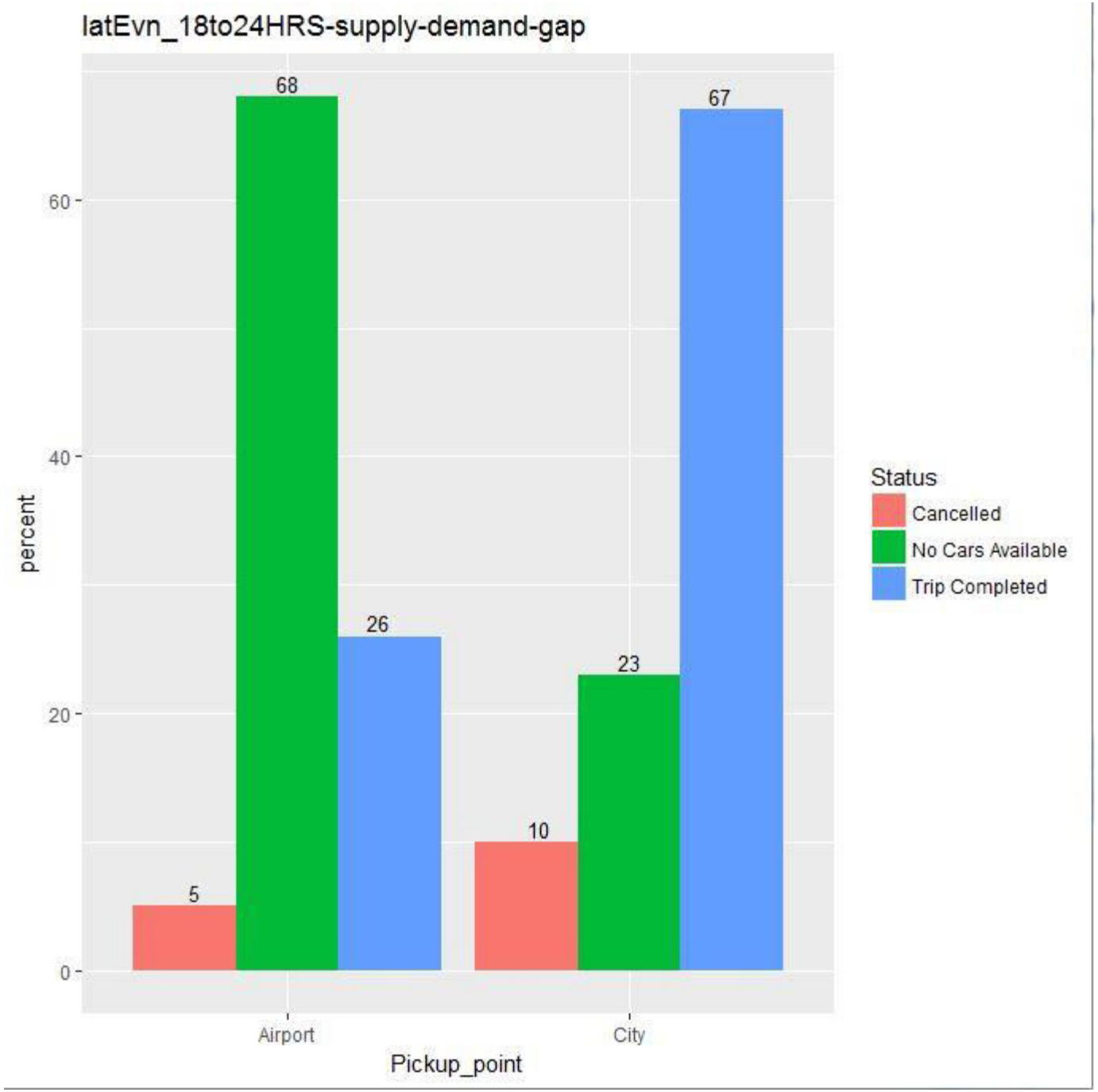
Most pressing problem at Airport is "No cars available" between latEvn_18to24HRS



Most pressing problem at City is "Cancelled" between morn_6to12HRS



The supply demand gap is the highest when percent of "Trip completed" is the least.
In the plot the height of the "trip completed"in each of the times_slot examined gives the least in "latEvn18to-24HRS"



The supply demand gap is the highest when percent of "Trip completed" is the least.
In the plot the height of the "Trip completed"in "Airport" and "City" examined gives the least at "Airport" in "latEvn18to-24HRS"

Reasons for the supply-demand gap.

Airport: In time slots 0-6HRS,6-12HRS,12-18HRS,18-24HRS, the supply demand gap are respectively 60%, 56%,56%, 62%. The critical issue in each of the slots is “No Cars Available”. Further the most pressing problem occurs during 18-24 HRS.

HYPTOTHESIS OF THE PROBLEMS:

1) Large amount of down time for the drivers at the airport.
2) Insufficient number of drivers/cars operating leading to no cars available
3) Most drivers don’t schedule themselves.

City: In time slots 0-6HRS,6-12HRS,12-18HRS,18-24HRS, the supply demand gap are respectively 70%, 66%, 38%, 33%. The critical issues in each of the slots are “Cancelled”, “Cancelled”, “No Cars Available”, “No Cars Available”. The most pressing problem occurs during 0-6HRS slot.

HYPTOTHESIS OF THE PROBLEMS:

1) Cancellations may be because of large downtime resulting at the airport and hence the drivers are reluctant to provide service.
2) Insufficient number of drivers/cars operating leading to no cars available
3) Most drivers don’t schedule themselves.

Ways to resolve the supply-demand gap.

Uber is the top most company among taxi service providers and will see a lot of competition from other's like OLA and others.
The companies objective will be to stay at the top of their game and take steps that will lead to higher growth and profitability in long run. To achieve this here are some solutions:

1) The company should recruit more number of drivers to beat the "No Cars Available" issue.
2) The company should give extra incentive to drivers to work in Time slots where the problem is most pressing.
3) The company should think of providing the different segment of cars like "Mini", "Micro", Luxury" etc., to beat the competitors and beneficial to the consumers.
4) The company also can think of suitable pricing in a limited manner especially in the slots where the problem is most pressing.(The most pressing issue at the Airport is "No cars available" between latEvn_18to24HRS.The most pressing issue at the City is "Cancelled" between morn_6to12HRS)
5) Prior booking may be allowed by the consumers so that the drivers commit to the trip.