Uber Supply Demand Gap Assignment

UpGrad

Data Exploration

- There are six attributes associated with each request made by a customer:

1. Request id: A unique identifier of the request

2.Time of request: The date and time at which the customer made the trip request

3.Drop-off time: The drop-off date and time, in case the trip was completed

4.Pick-up point: The point from which the request was made

5.Driver id: The unique identification number of the driver

6.Status of the request: The final status of the trip, that can be either completed, cancelled by the driver or no cars available

- There are 6745 such requests

Data Cleaning and Manipulation

- In order to do analysis the data must be clean.
- To clean the data the following steps must be taken:

➢Rename the columns of the data frame properly

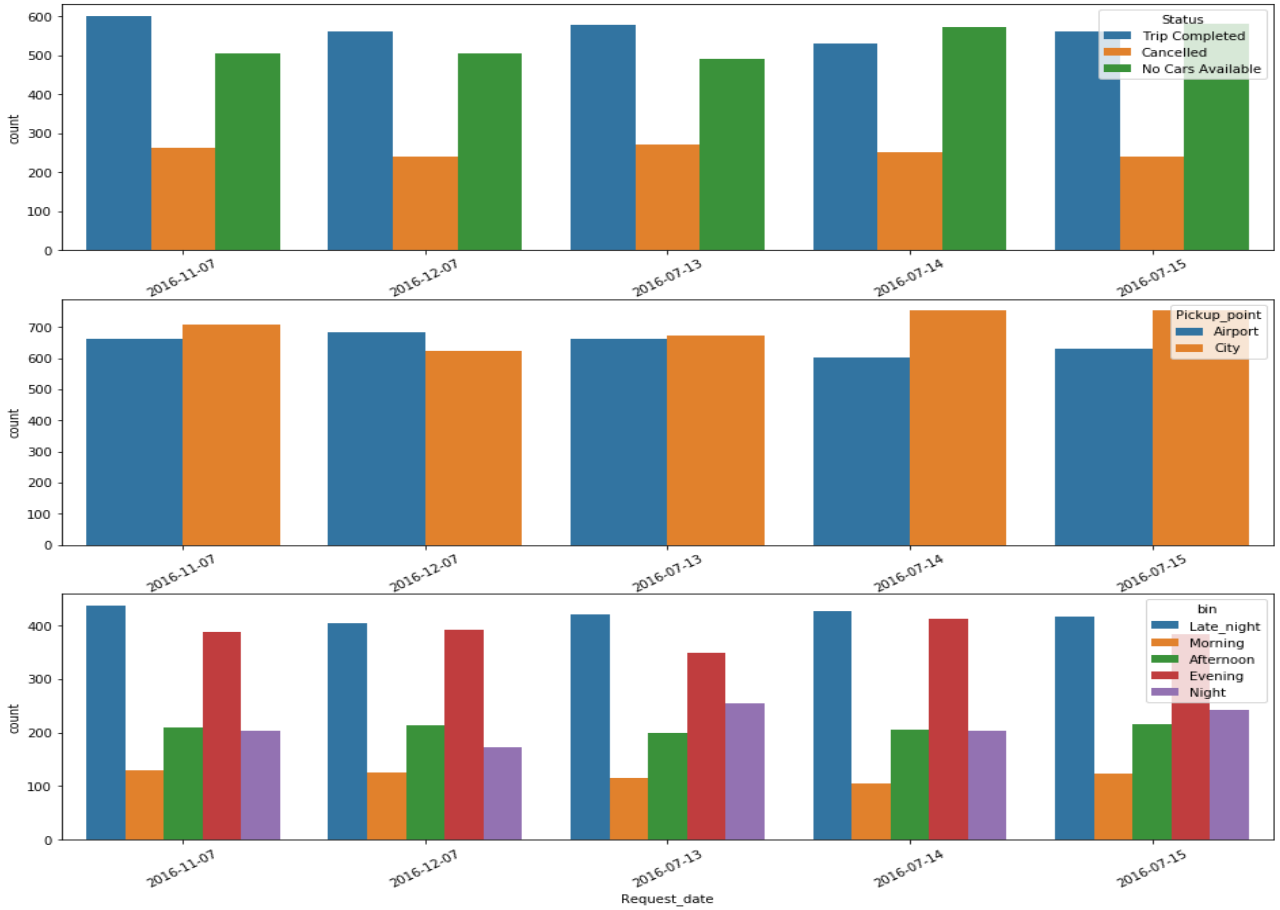
➢Find the attribute that is unique The number of unique attribute should be equivalent to the number of rows in the dataset.

➢Fill in the unnecessary missing values.

➢Change the type of Request timestamp and Drop timestamp

Analysing Trends For Each Day

- For given data, convert the Request timestamp into str format and split it on " " into Request_date and Request_time.
- Perform a group_by on Request_date and analyze the data.
- Visualise the data
- We see that data is given for the dates- 7-11-2016 , 7-12-2016 , 13-7-2016, 14-7-2016 , 15-7-2016.
- We see that majority of the requests are made on 15th of July.
- Let us see day by day-



Analysing Trends For Each Day

- 15-7-2016(1381 requests)
  - for 561 number of requests the trip was completed
  - for 580 number of requests there were no cars available(which is not an appreciable scenario)
  - and 240 number of requests were cancelled
- 7-11-2016(1367 requests)
  - for 601 number of requests the trip was completed
  - for 504 number of requests there were no cars available
  - and 262 number of requests were cancelled
- 14-7-2016(1353 requests)
  - for 530 number of requests the trip was completed

Analysing Trends For Each Day

- for 571 number of requests there were no cars available(which is again not appreciable)
- and 262 number of requests were cancelled

❑ 13-7-2016(1337 requests)

- for 577 number of requests the trip was completed
- for 490 number of requests there were no cars available
- and 270 number of requests were cancelled

❑ 7-12-2016(1307 requests)

- for 562 number of requests the trip was completed
- for 505 number of requests there were no cars available
- and 240 number of requests were cancelled

Combining Data For All Days

- When we look at the data as a whole

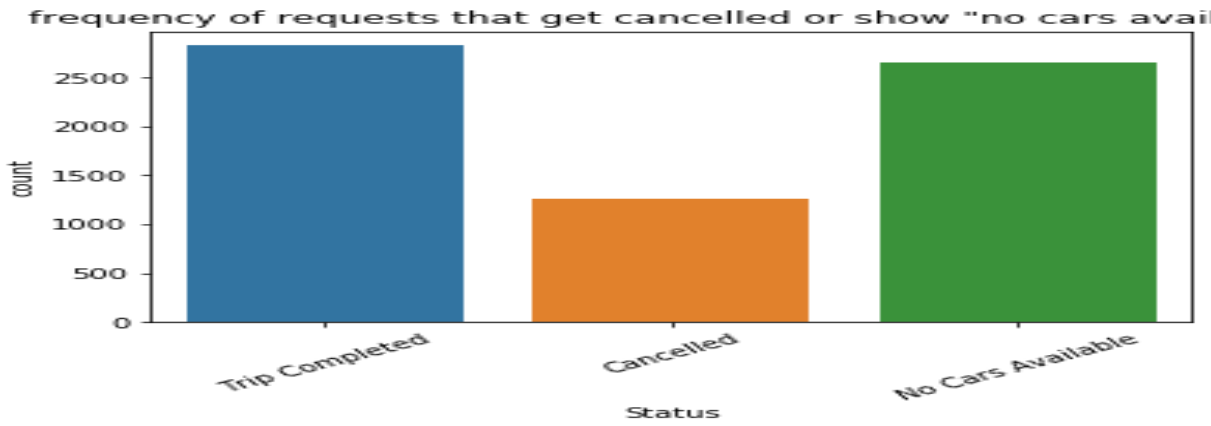
the following points can be seen.

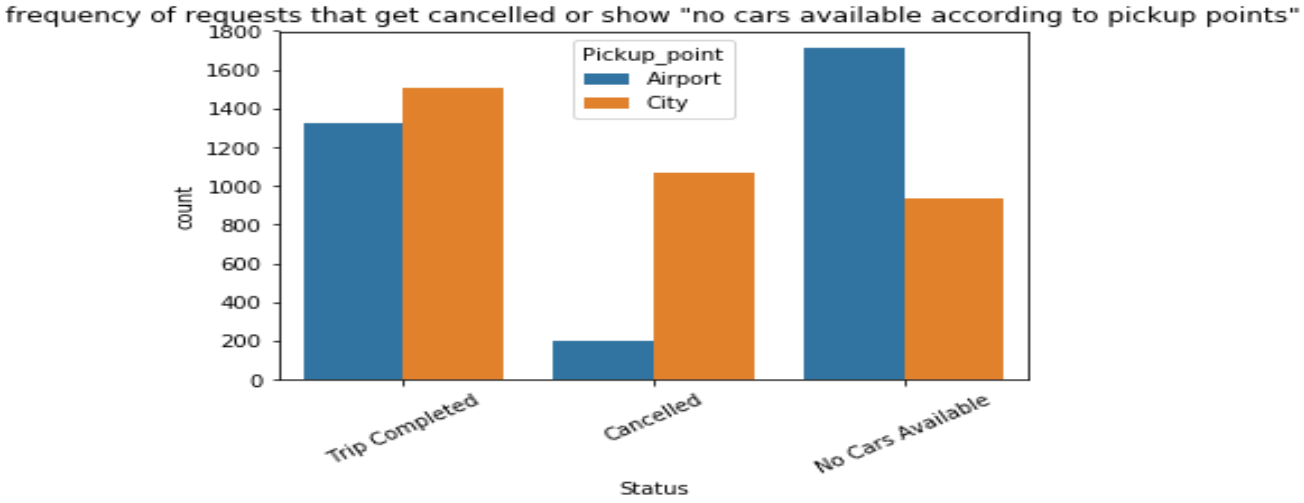
❖ for 2831 number of requests the trip was completed , 1327 from airport to city and 1504 from city to airport.

❖For 2650 number of requests there were no cars available

❖And 1264 number of requests were cancelled , 198 from airport to city and 1066 from city to airport.

- We can see that not many drivers are willing to go to airport and there are not many cars available at airport.



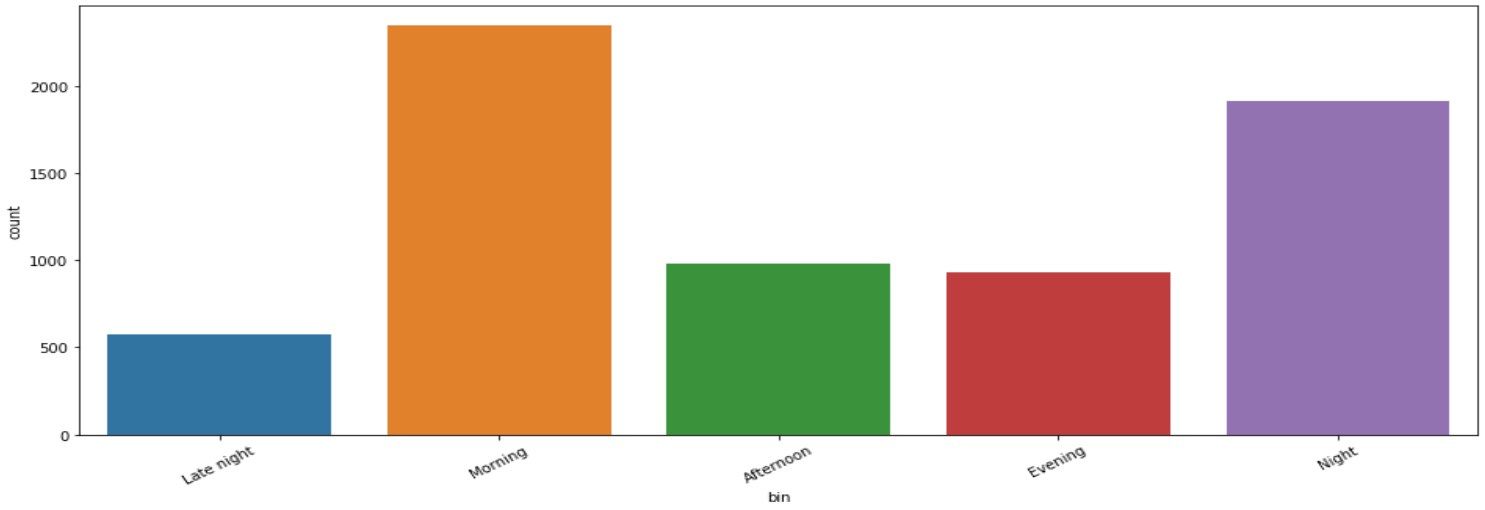


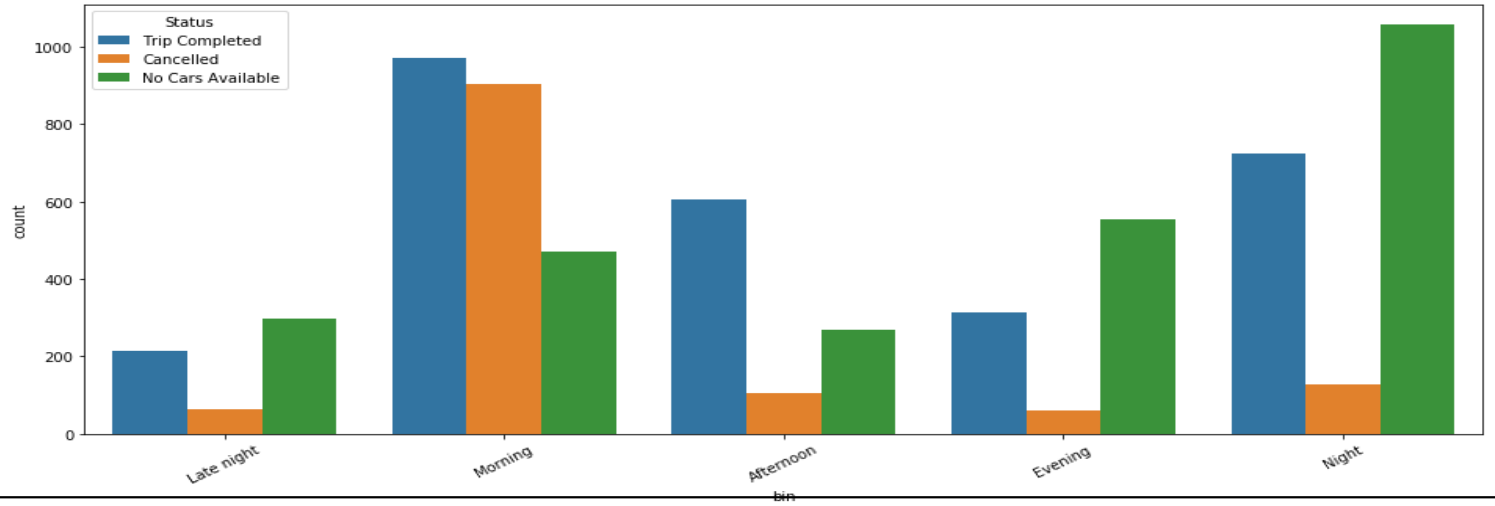
Binning Time Into 5 Categories

- Now divide the Requests according to the day times
- bins = [0, 5, 11, 17, 19, 24]
- labels = ['Late night’ , 'Morning’ , 'Afternoon’ , 'Evening’ , 'Night']
- 1.Morning(5 am -11am)
- 2.Afternoon(11 am – 5 pm)
- 3.Evening (5 pm – 7 pm)
- 4.Night (7 pm – 12 am)
- 5.Late night (12 am – 4 am)

Problem Identification – Morning and Evening

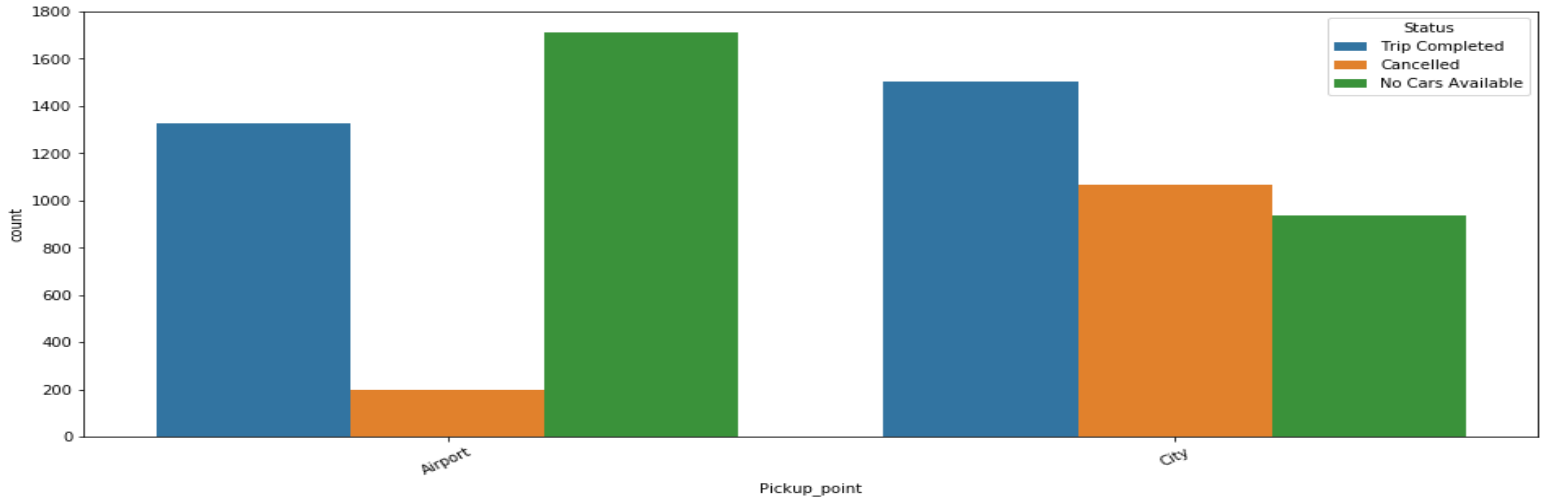
- We now know that most of the requests have been made in the morning.
- However we also see that for more number of requests the response is that there are ’no cars available’ in later half of the day starting from sunrise i.e., from 5 pm to 4 am.

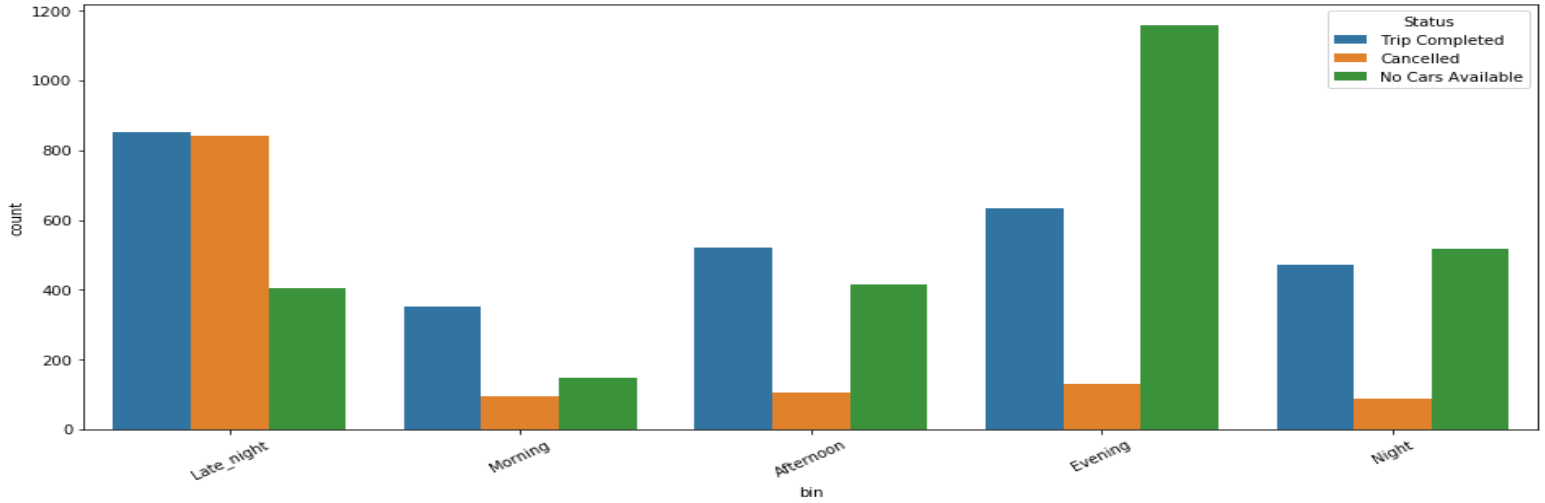




Problem 1 - Cancelled trips

- The majority of requests are cancelled from city to airport for they have to wait for customers who would book the car to go back to the city.
- The majority of requests are cancelled from 12 am to 4 am.

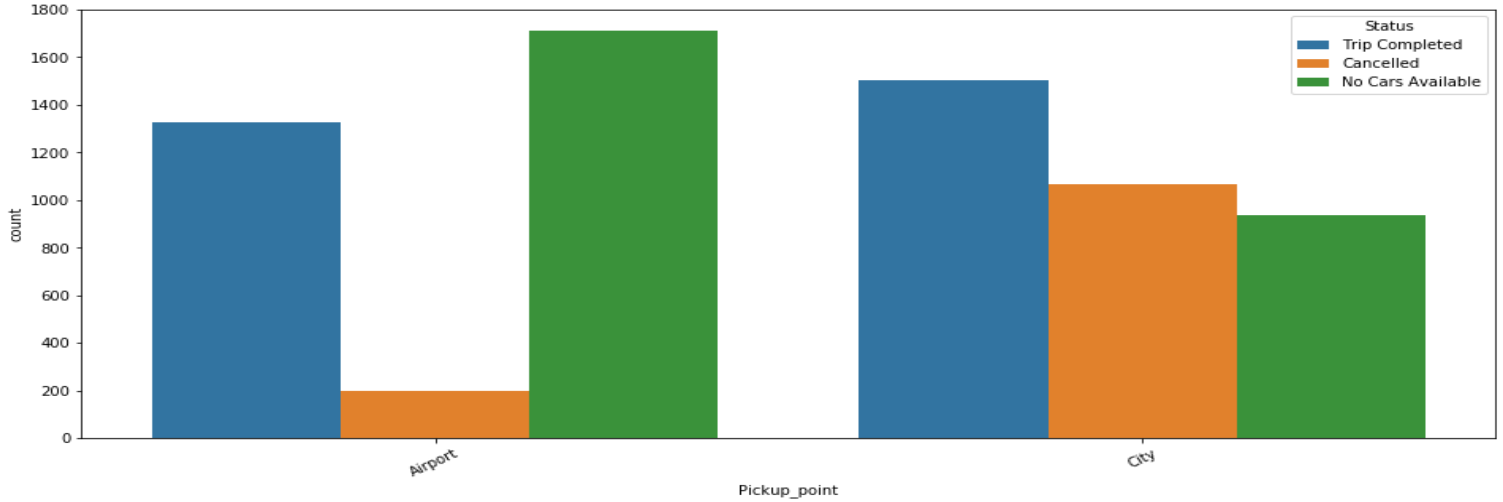


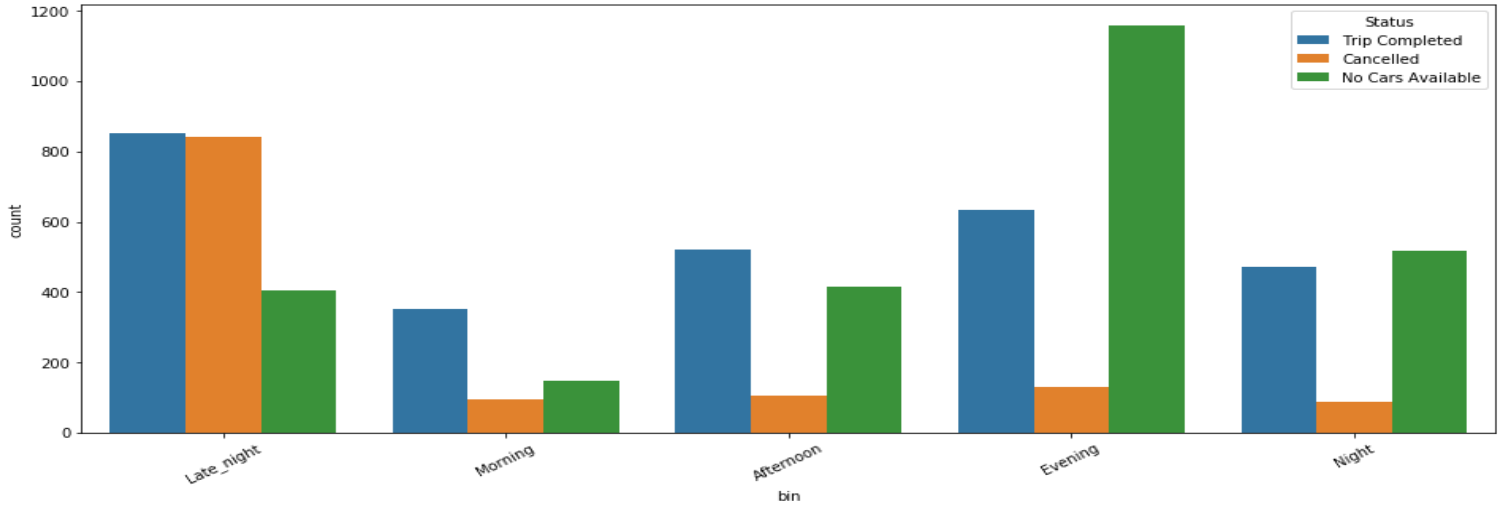


UpGrad

Problem 2 - No Cars Available

- For majority of requests from airport to city the response is ‘no cars available’ as not many are willing to wait for the customers to go back to the city.





UpGrad

Recommendations

- We know that the above two issues also impact the business of Uber.
- The following steps can be taken to improve the above situation :

✓ Collaborate with other uber drivers for a certain percentage of drivers to make a trip to the airport at a particular time of the day.

✓ Instead of cancelling the request refer to the other driver whose turn it is to make the trip.

✓ Introduce the credit system: In a particular day the driver with less number of cancellations are more likely to get the credits and will get an increase .

✓ The percentage can increase according to the occasions for eg during festivals , summer , winter when people are more likely to travel.