

Uber Supply Demand Gap Analysis

Presented By : Madhuri Pingle

Project Summary

Uber is an online car aggregate service. Uber loses out on its revenue due to the supply and demand gap of the cars. This is mainly due to cancellation of requests by drivers and non-availability of cars. In this project, we will study, visualize and solve Uber's Supply Demand Gap Problem.

In Uber Dataset we have total 6745 record. All the trips are related to only Airport-City and City Airports.

Among that there are three types of Trips:

1. Trips completed
2. Cancelled
3. No cars available

The aim of analysis is to identify the root cause of Uber's problem (i.e. cancellation and non-availability of cars) and recommend ways to improve the situation. The approach for this project has been to divide the entire case study into various checkpoints to meet each of the sub-goals using EDA and Visualization.

For this project, we will follow the following strategy:

1. Identify the problematic trip status for trips originating from airport and city.
2. Analyze if the day of week has any impact on the trips originating from airport and city for the problematic trip status.
3. Determine the pressing problems by analyzing the request frequency of trips from airport and city for problematic trip statuses.
4. Analyze the demand for cabs for each timeslot for trips from airport and city.
5. Identify the timeslot that has the highest demand and supply gap, and find the request type (airport to city or city to airport) that is hugely effected.
6. Identify the reasons for the gap.
7. Recommend ways to resolve the gap.

Business Problems

We avail Uber or any other cab services most of the times for travelling to and from the airport.

Many times, we face the problem of cancellation by the driver or non-availability of cars.

These problems faced by customers also impact the business of Uber as Uber loses out on its revenue .

Uber would like to address these problems of driver cancellation and non-availability of cars leading to loss of potential revenue

Business Objective

- 1. Identify patterns and trends in trip requests by time of day, day of the week, and pickup location.
- 2. Detect peak demand hours and areas where the supply of drivers consistently fails to meet user demand.
- 3. Analyze the reasons behind unfulfilled requests, such as high cancellation rates or no car availability.
- 4. Generate actionable insights that can help Uber optimize driver distribution and improve customer satisfaction.

Business Data

- ▶ This dataset captures **ride request information** from **Uber** in **Bangalore** over a period of **several days in July 2016**. It is intended for analyzing operational challenges and service performance, particularly at two major locations: **Airport** and **City**.
- ▶ Total number of Trip Requests = 6745
- ▶ Total Number of attributes for each request = 6
- ▶ **Request ID**: Unique identifier for each ride request
- ▶ **Pickup point**: Location from where the ride was requested - either **City** or **Airport**
- ▶ **Driver ID**: Unique identifier for the driver (may be blank if no driver was assigned)
- ▶ **Status**: Final status of the ride - **Trip Completed**, **Cancelled**, or **No Cars Available**
- ▶ **Request Timestamp**: Time when the ride was requested
- ▶ **Drop Timestamp**: Time when the trip ended (only present if the trip was completed)

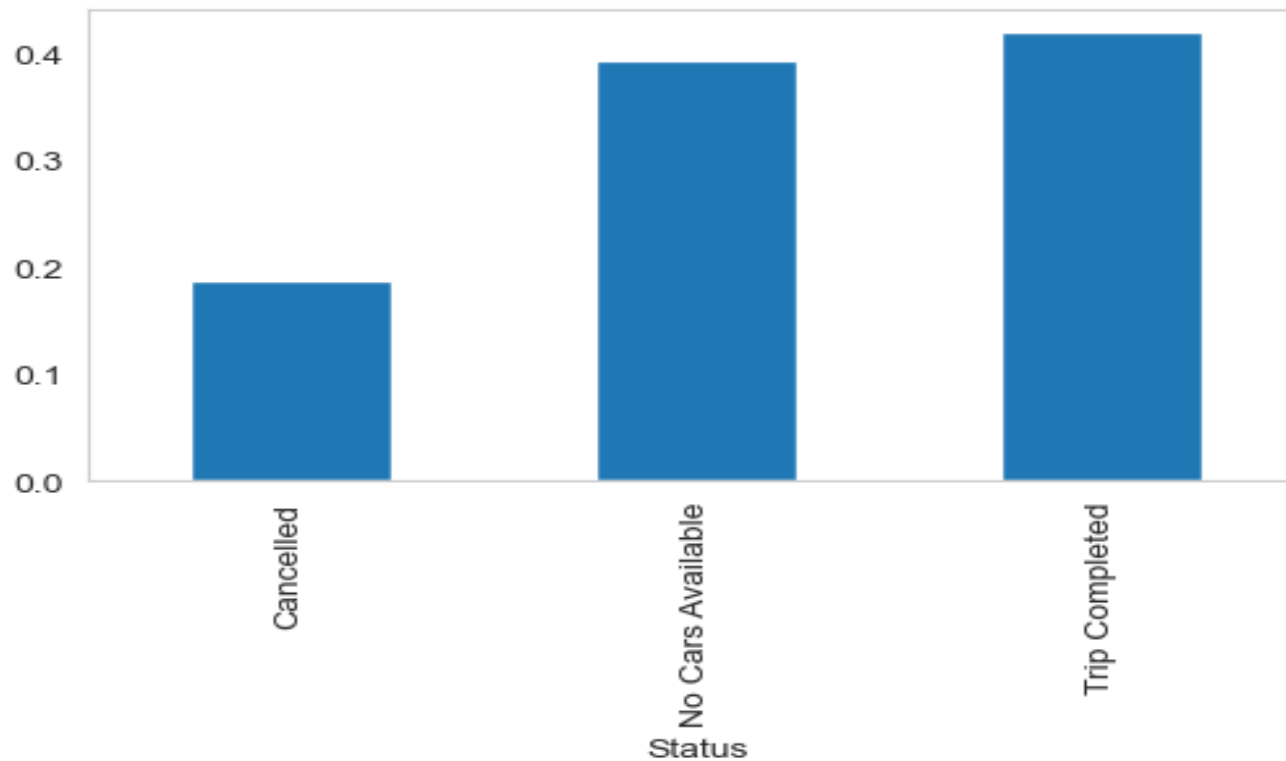
Primary Objectives & Analysis Goals

- ▶ Identify **supply-demand** gaps.
- ▶ Analyze **driver availability** patterns.
- ▶ Understand **ride request behavior** at different times of the day.
- ▶ Study cancellation trends and “No cars available” scenarios.
- ▶ Derive **actionable insights** for operations improvement.

Root Causes : Frequency the Request Status Distribution

The number of trips marked as Cancelled or No Cars Available is significantly high, indicating operational inefficiencies or supply-demand mismatch.

If Trip Completed is significantly lower than the other statuses, it suggests that many customers are unable to get rides, possibly due to driver unavailability.

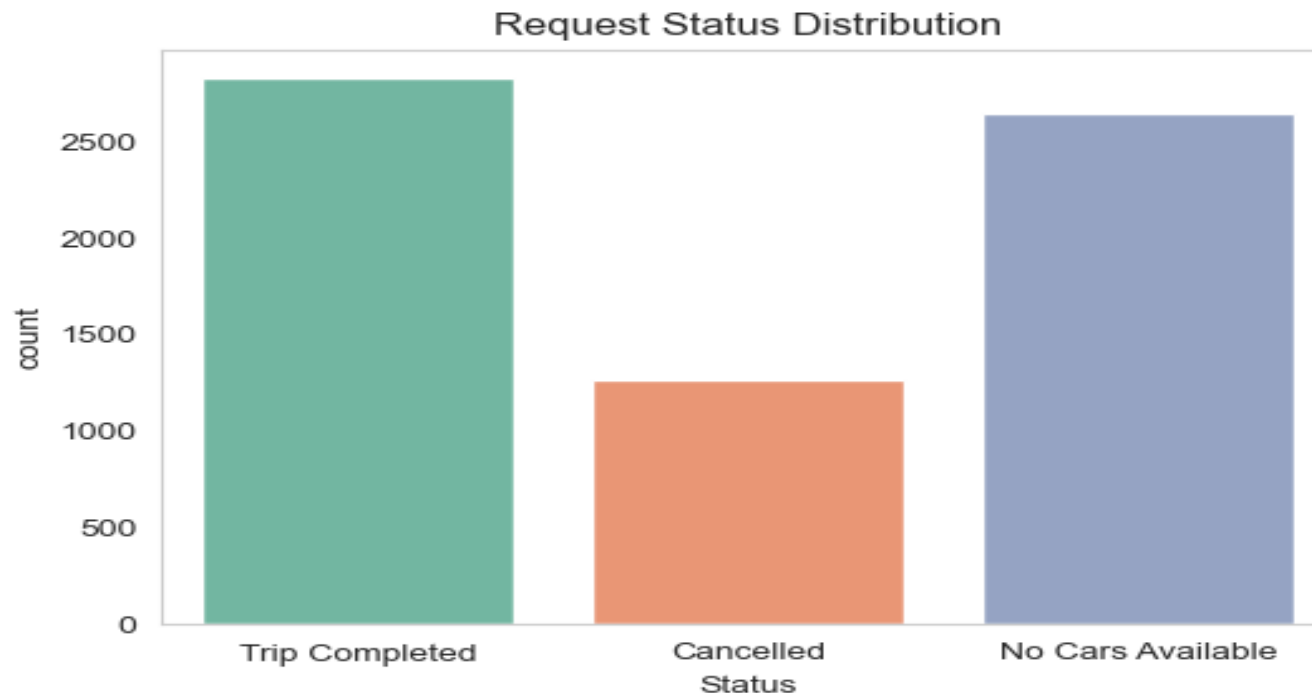


Root Causes : Count of Request Status Distribution

A high number of Cancelled trips could imply rider impatience or driver behavior.

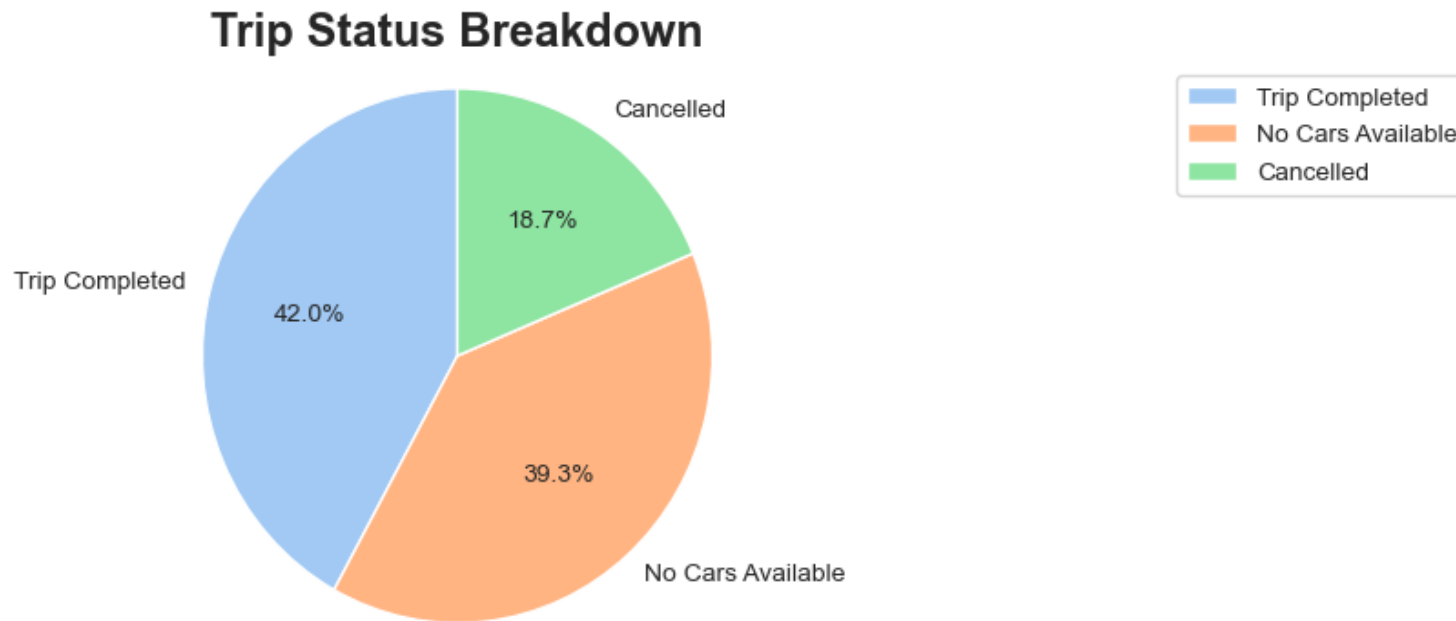
A high number of No Cars Available trips may reflect a lack of available drivers during peak hours or in certain locations.

If Trip Completed is significantly lower than the other statuses, it suggests that many customers are unable to get rides, possibly due to driver unavailability.



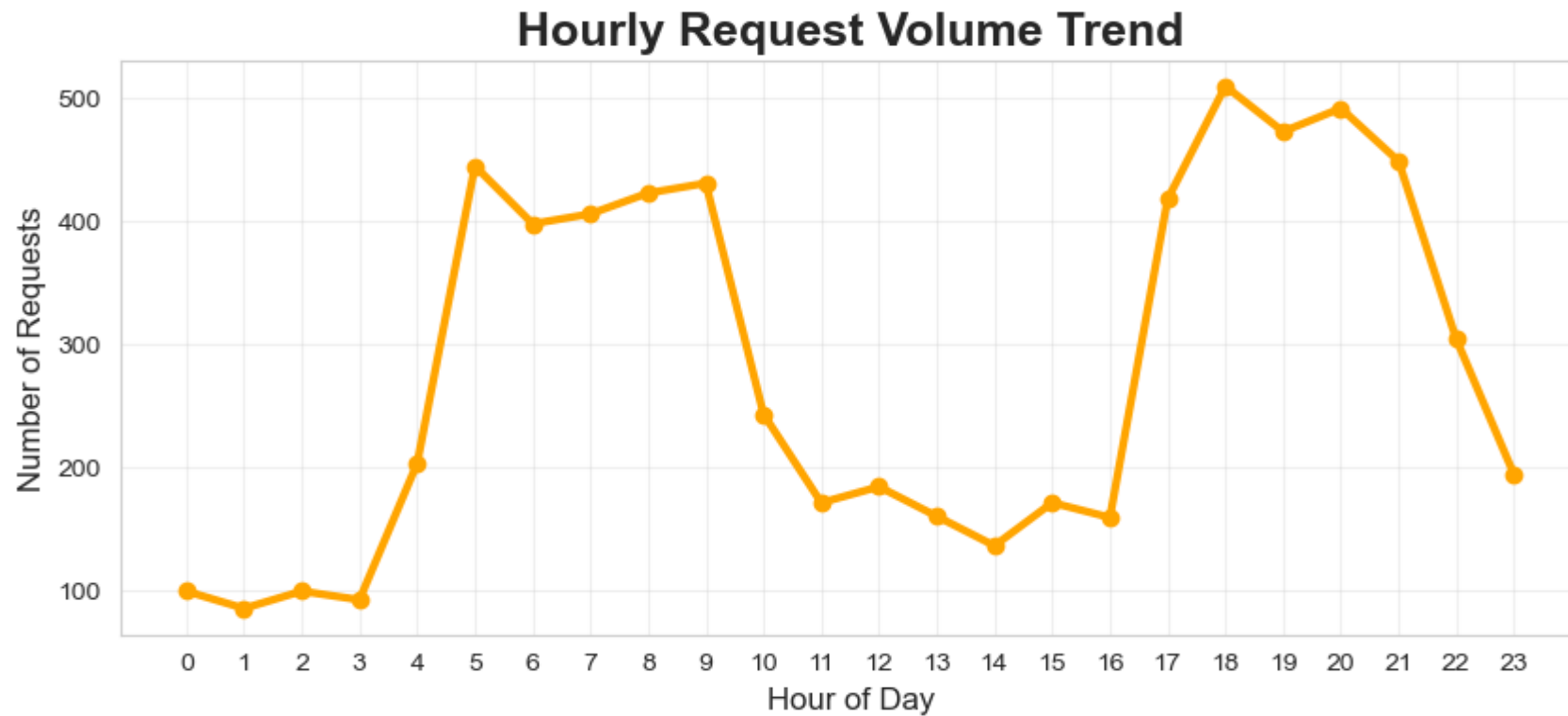
Root Causes : Proportions of categorical variables and determine which trip statuses predominate

Most of the trips are cancelled or unavailable cars, and hence most of the requests are not fulfilled.



Root Causes : Hourly Request Volume Trend

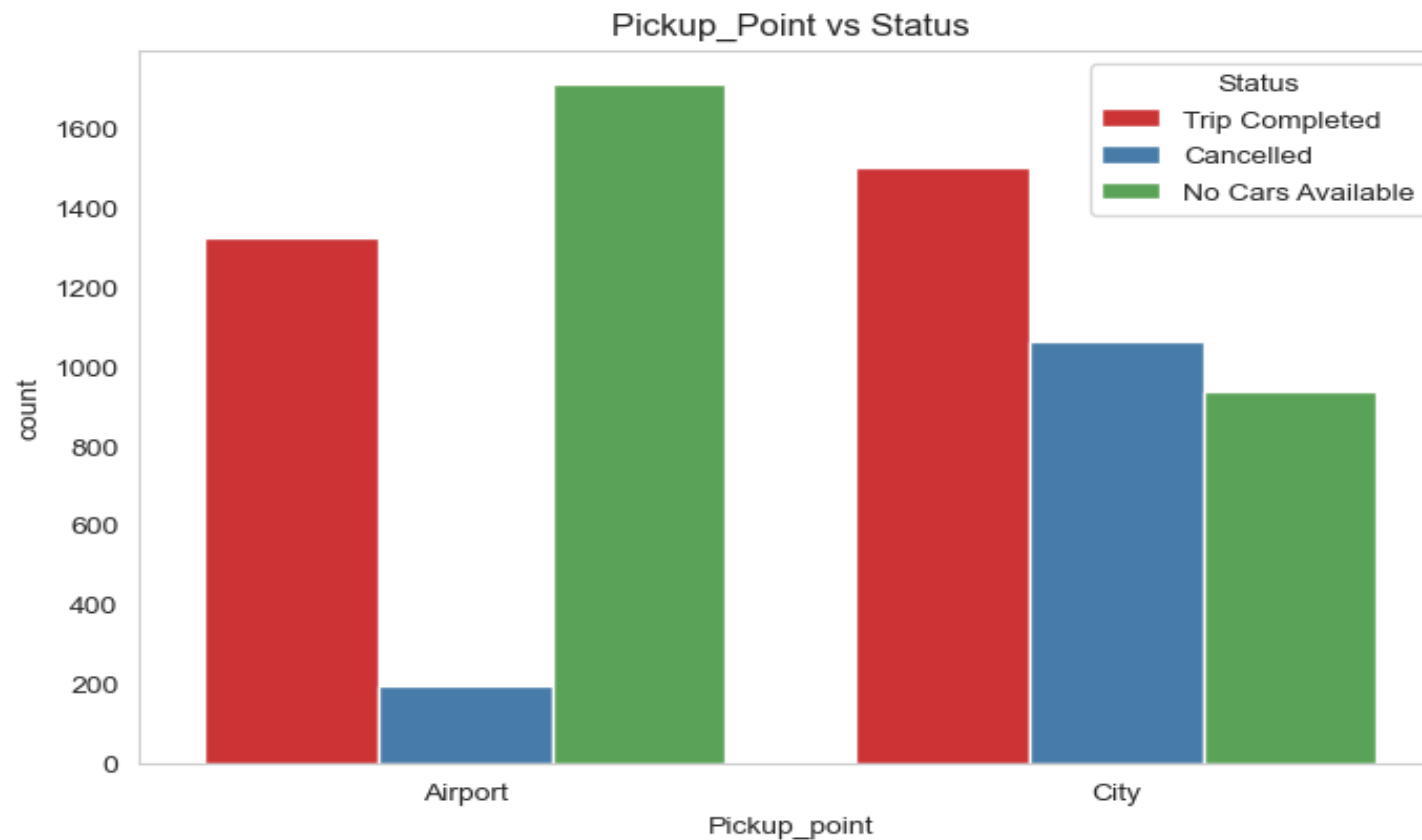
Peak demand occurs during morning (5-9 AM) and evening (5-9 PM) hours, matching the customer behavior.



Root Causes : Pickup Point Vs Status analysis

There are significantly more requests originating from the Airport compared to the City.

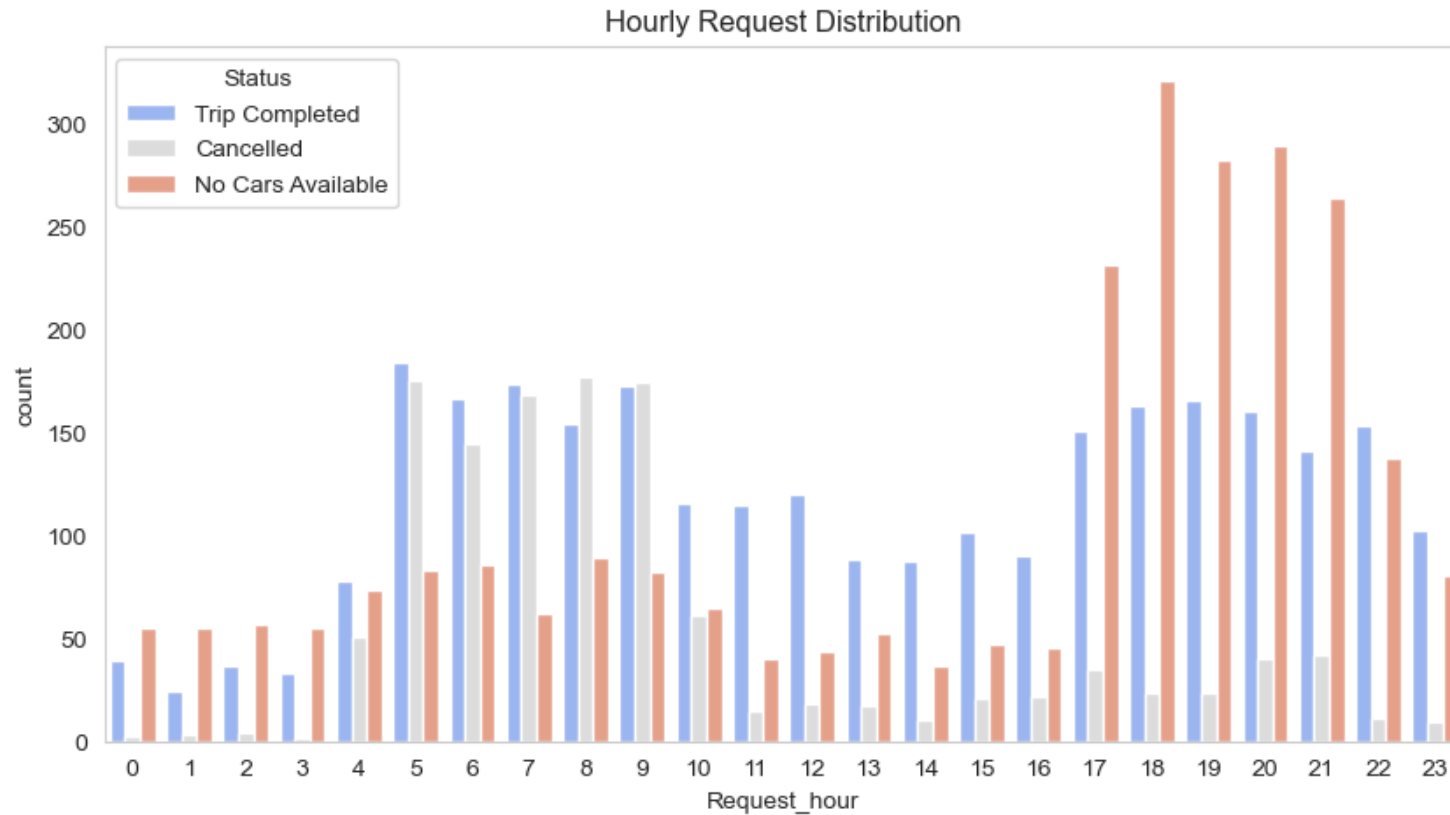
This suggests that passengers arriving at the airport are frequently booking cabs.



Root Causes : Hourly Request distribution analysis

There's a noticeable spike in requests. A large portion of these are "No Cars Available", especially from the City pickup point. Suggests supply-demand mismatch during office commute time.

Another major demand spike. Significant number of "Cancelled" rides. Indicates issues like driver cancellations, possibly due to traffic or unprofitable routes.

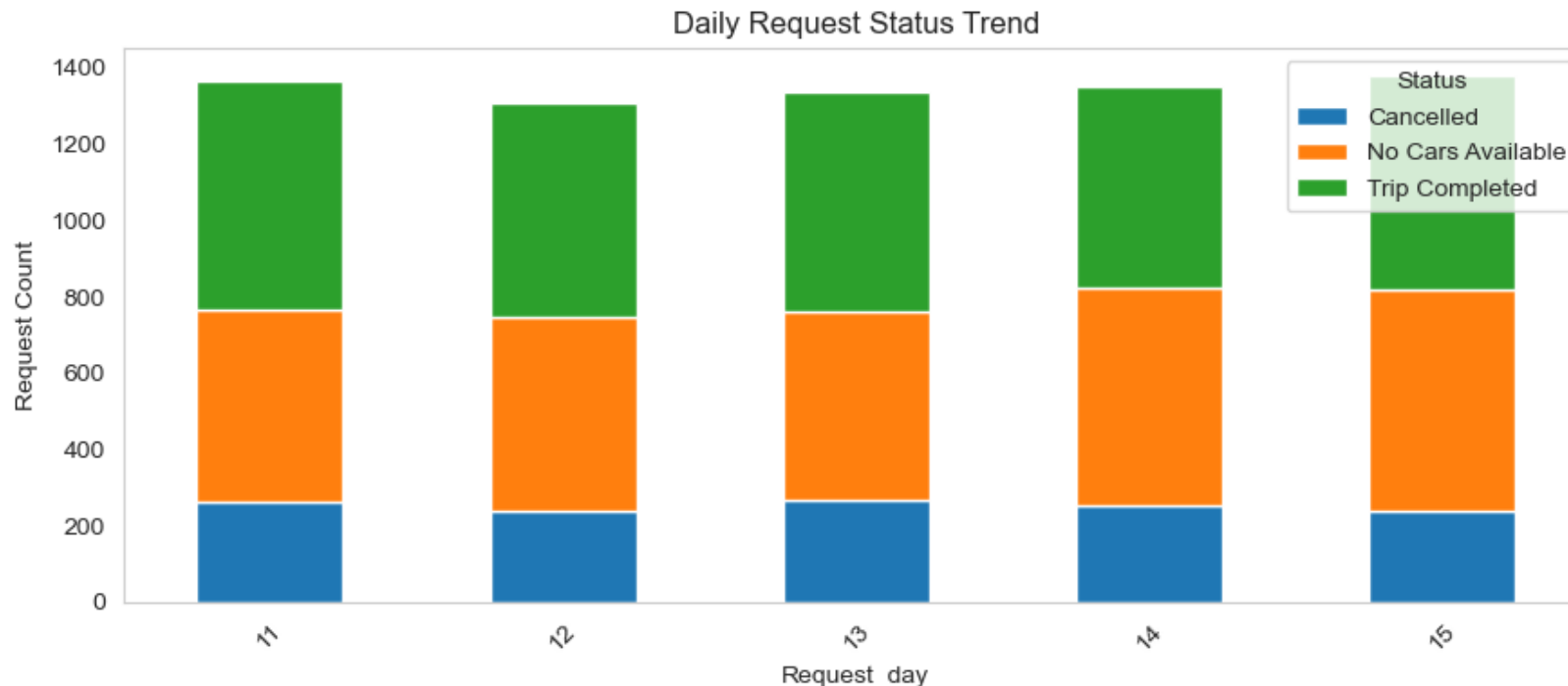


Root Causes : Daily Request Status Trend analysis

Days with taller bars represent high request volumes. You can spot which days had the highest demand, useful for resource planning (e.g., more drivers needed).

Larger green portions (Trip Completed) indicate better service performance on those days. If the green section is consistently small, it suggests low fulfillment rates.

Days where "No Cars Available" dominates are opportunities to deploy more drivers or improve coverage. Cancellations clustering on specific dates may need policy changes or incentives for drivers.



Root Causes : Request Status By Time slot analysis

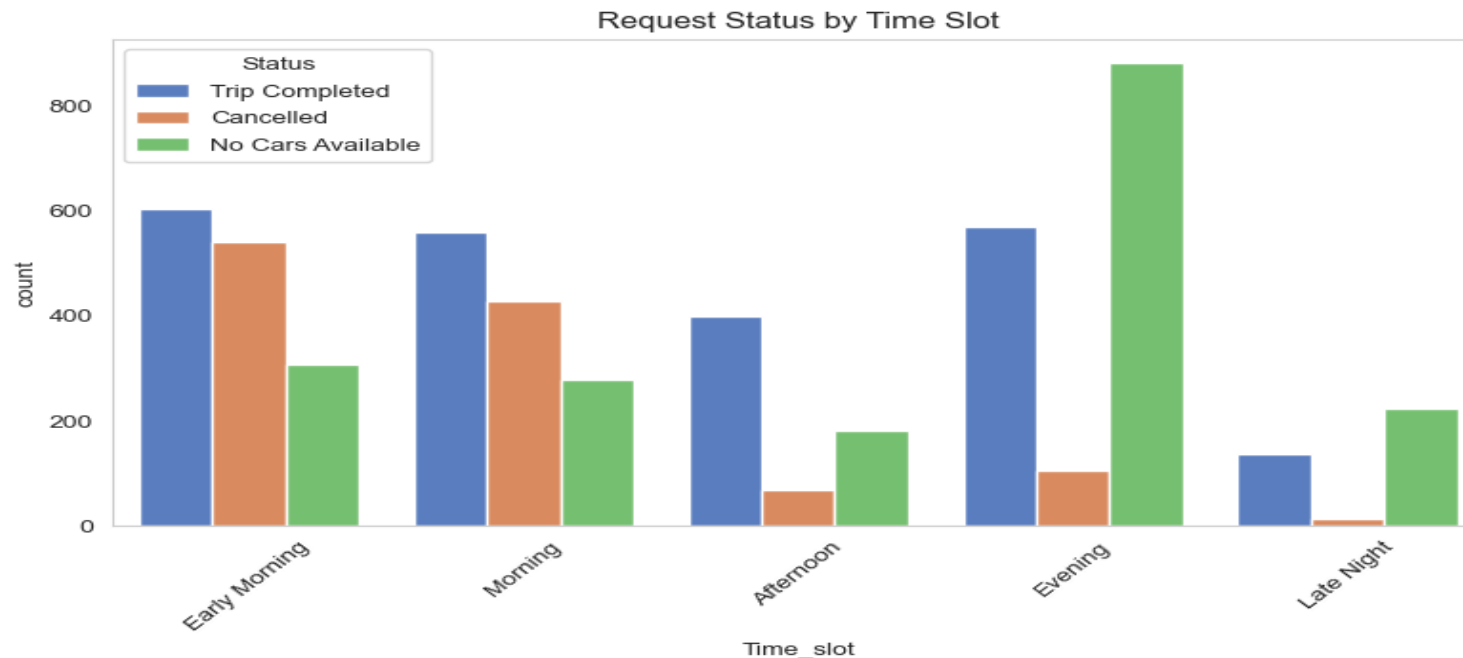
1.Early Morning Rush (5 AM - 8 AM): High number of "No Cars Available": Most requests are not fulfilled due to unavailability of cars.

2.Morning Rush (8 AM - 12 AM): High number of "No Cars Available": Most requests are not fulfilled due to unavailability of cars. Suggests demand > supply in the morning.

3.Evening Rush (5 PM - 10 PM): High number of "Cancelled" requests: Drivers cancel more frequently during this time. Indicates possible driver-side issues like traffic or selective acceptance.

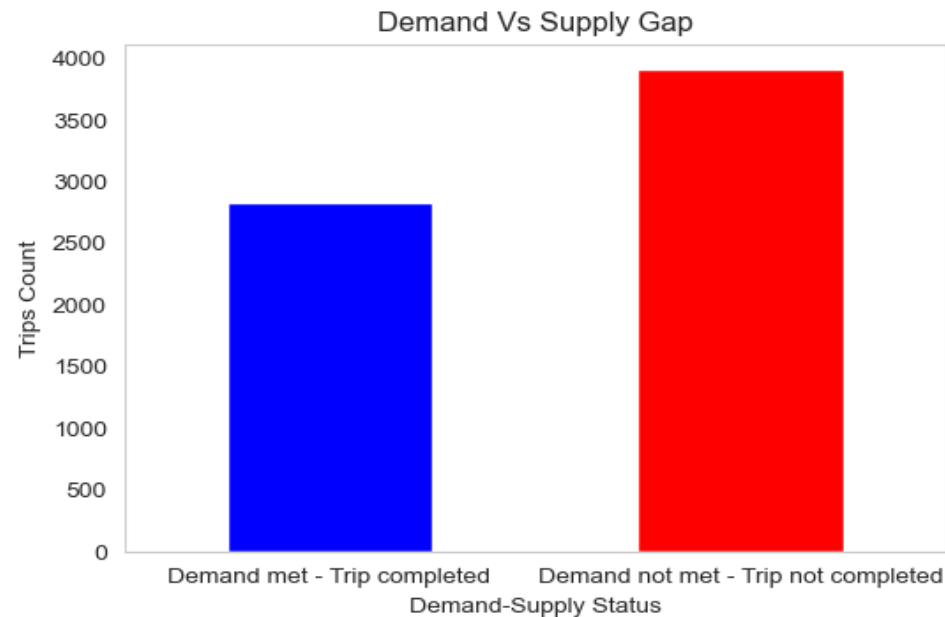
4.Night and Early Morning (10 PM - 5 AM): Fewer total requests, but a higher proportion of "No Cars Available". Could reflect limited driver availability during off-peak hours.

5.Afternoon (11 AM - 4 PM): Mostly completed trips, low issues. Indicates a relatively balanced demand-supply during this time.



Root Causes :Supply Vs Demand Gap analysis

If trip is "cancelled" or "No cars available" is only When supply is not met to demand. If trip is said to be "Trip Completed" when Demand and supply is met. Demand not met is more than Demand met, means highest gap is available in Demand and supply •

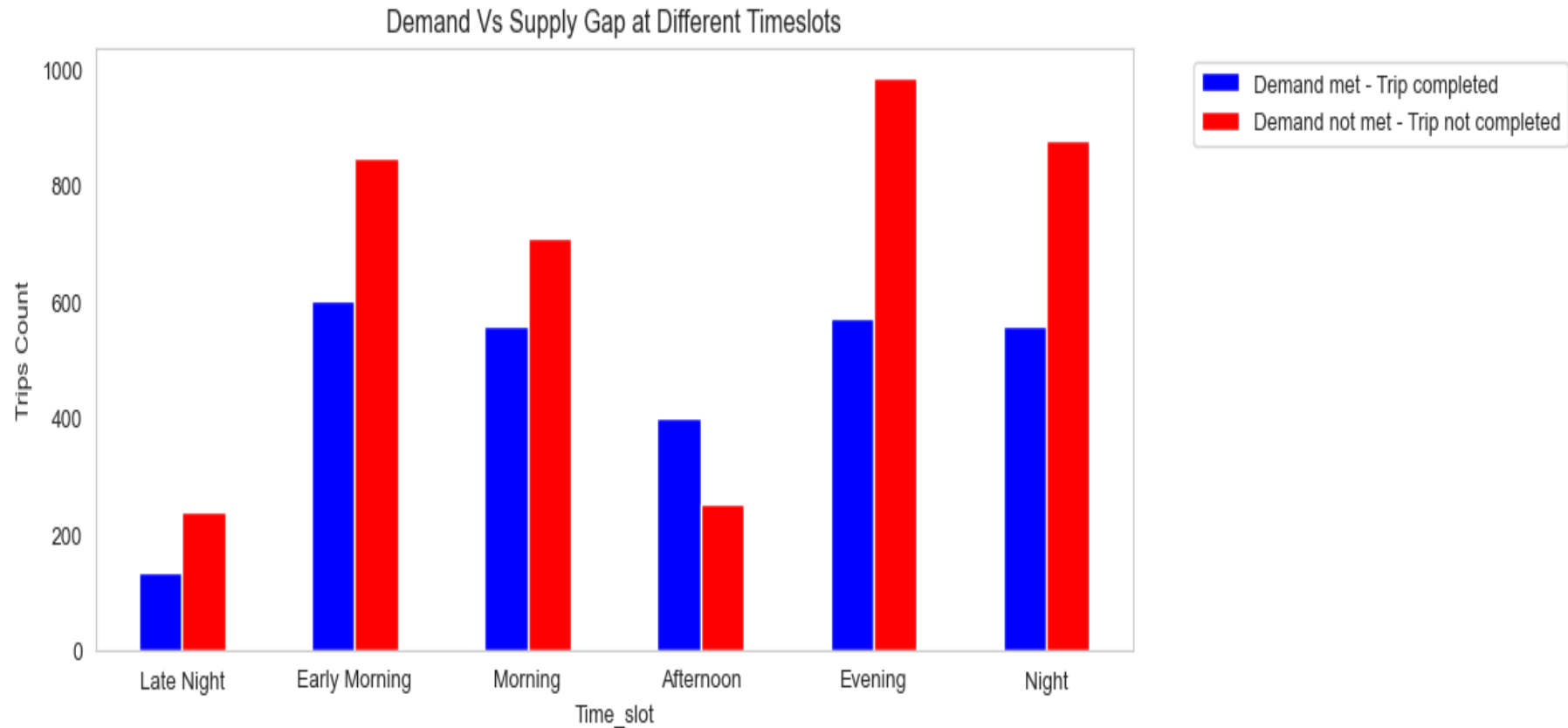


Root Causes :Supply Vs Demand Gap at Different Time slot analysis

Demand-Supply Gap exists highest at the following Time Slots

At Early Morning (5 am -8 am)

At Evening (17 pm - 20 pm)



Solution to Business Objective

- ▶ Based on the data analysis performed, following recommendation can be used by Uber to bridge the gap between supply and demand:-
- ▶ For bridging the demand supply gap from airport to city, making a permanent stand in the airport itself where the cabs will be available at all times and the incomplete requests can come down significantly.
- ▶ Uber can provide some incentives to the driver who complete the trip from city to airport in the morning part. This might result the driver to not cancel the request from city to airport trips.
- ▶ Last but sure solution to bring down the gap is to increase the numbers of cab in its fleet.

Suggested Strategy to Achieve Business Objective

- ▶ Balance Demand and Supply - Especially During Peak Hours
- ▶ Morning (5 AM - 10 AM):
 - ▶ Issue: High demand but many “No Cars Available”. Action: Increase driver incentives to work during morning peak hours. Benefit: Meet demand, reduce lost rides, increase revenue.
 - ▶ Evening (5 PM - 10 PM):
 - ▶ Issue: High rate of driver cancellations. Action: Investigate cancellation reasons (traffic, selective pickups), improve app-based penalty/reward systems, and optimize matching algorithms. Benefit: Fewer cancellations = better customer satisfaction.
- ▶ Driver Shift Optimization
 - ▶ Use historical data to predict high-demand periods. Encourage shift-based planning so more drivers are available during peak times and fewer during off-peak. May involve driver scheduling tools or offering flexible incentives.
- ▶ Improve Real-time Allocation Algorithm
 - ▶ Modify the matching algorithm to:
 - ▶ Prioritize nearest available driver
 - ▶ Avoid assigning requests to drivers likely to cancel
 - ▶ Use machine learning models to predict driver behavior and adjust allocation accordingly.
- ▶ Customer Communication Enhancements
 - ▶ During high-demand hours, send push notifications:
 - ▶ Informing users about wait times or surge pricing
 - ▶ Suggesting alternate time slots for cheaper or faster rides
 - ▶ 5. Expand Driver Pool in Underserved Areas
 - ▶ If cancellations or “no cars” cluster around certain pickup locations (e.g., airports, outskirts), consider:
 - ▶ Targeted driver recruitment campaigns
 - ▶ Temporary location-based incentives

Conclusion

- ▶ Exploratory Data Analysis (EDA) of the Uber request data provided clear supply-demand imbalances, which were mainly due to time-of-day and pickup location effects.
- ▶ Demand peaks sharply in the early morning and evening, particularly for airport pickups, while demand hours are fairly short, leaving a gap rate over 70% in peak timeslots.
- ▶ City pickups experience relatively better fulfillment, especially in the evening.
- ▶ Visualizations also highlighted that trip duration remain consistent, and consequently, inefficiencies arise not due to ride size but missed timing of driver availability.
- ▶ The multivariate analysis further confirmed that gap rates are highest during high-demand time slots and that there are no strong correlations between numeric variables, affirming the significance of categorical and time based patterns.
- ▶ Therefore, these findings are critical to guiding targeted interventions such as driver incentives, real-time pricing, or fleet repositioning in order to boost compliance and customer satisfaction, ultimately generating positive business outcomes by responding to supply shortages at the time and location it is most valuable.

**Hurrah! You have successfully completed
your EDA Capstone Project**

Thank You

