Name: *Shivali Muthukumar*
Date: *October 2025*
Tools Used: Excel, SQL (SQLite), Python (Pandas, Matplotlib, Seaborn)

1. Introduction

The purpose of this project is to analyze Uber request data and identify the major causes of the supply–demand gap.
Using **Excel, SQL, and Python**, data was cleaned, visualized, and analyzed to uncover trends in trip completions, cancellations, and car unavailability.

The dataset contains **6,745 Uber ride requests**, including details such as pickup point, request timestamp, drop timestamp, driver ID, and status of each trip.

2. Objectives

- To understand the **supply–demand gap** in Uber ride requests.
- To identify **time periods** and **pickup points** with the largest number of unfulfilled requests.
- To provide **data-driven recommendations** for improving driver availability and reducing cancellations.

3. Dataset Overview

Attribute	Description
Request id	Unique identifier for each trip request.
Pickup point	Either "City" or "Airport".
Driver id	Unique identifier for assigned driver (may be null).
Status	Trip Completed, Cancelled, or No Cars Available.
Request timestamp	Time the customer requested a ride.
Drop timestamp	Time the trip was completed (if applicable).

4. Data Cleaning and Preparation

- **Excel** was used for initial cleaning:
  - Removed blank cells and trimmed extra spaces.
  - Standardized text (capitalization and spelling).
  - Converted timestamps into date/time format.

- **SQL (SQLite)** was used to summarize trip statuses and compute supply–demand ratios.
- **Python (Pandas)** was used for deeper exploratory data analysis (EDA) and visualization.

5. Visual Analysis and Findings
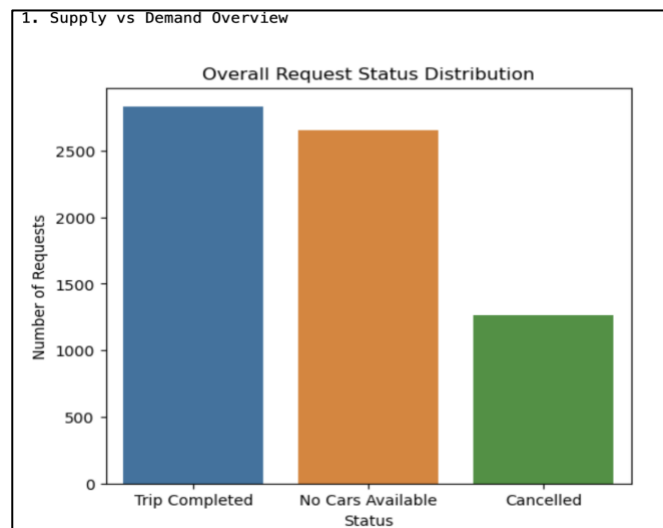
5.1 Overall Request Status Distribution

Fig. 1: Overall Requests by Status

Insight:
Out of 6,745 total requests:

- **2,831** were completed
- **1,264** were cancelled
- **2,650** had no cars available
  This indicates that **over 58% of total requests were not fulfilled**, showing a clear supply–demand imbalance.
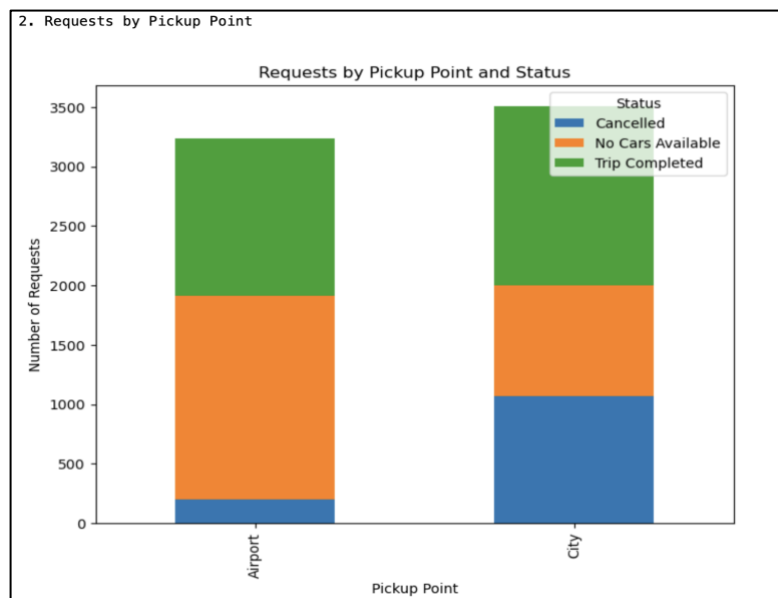
5.2 Requests by Pickup Point

Fig. 2: Requests by Pickup Point

Insight:

- At the **Airport**, more requests failed due to *No Cars Available* (≈59%).
- In the **City**, most unfulfilled requests were *Cancelled* by drivers (≈57%).
  This suggests that **car availability is a bigger issue at the Airport**, while **driver cancellations** dominate in the City.
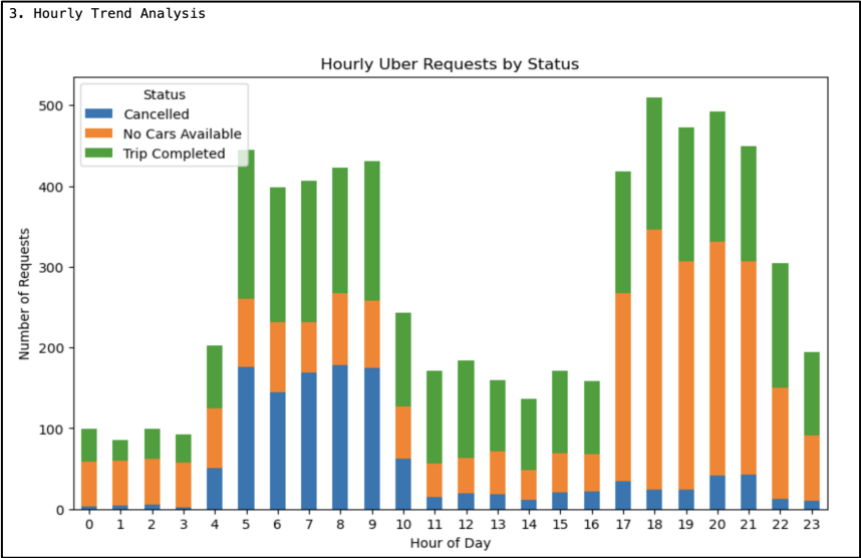
5.3 Hourly Trend Analysis

Fig. 3 : Hourly Supply–Demand Pattern

Insight:

- Demand spikes between **5 AM–9 AM** (Airport pickups to City).
- Another rise is seen around **10 PM–12 AM** (City to Airport).
- During these times, the number of completed trips drops significantly, showing **severe driver shortage during early morning and night hours.**

6. SQL Analysis Summary

Queries Run:

1. Total Requests by Status
2. Requests by Pickup Point
3. Hourly Supply–Demand Summary
4. Supply–Demand Gap Percentage.

Key SQL Results:

Pickup Point	Total Requests	Completed	Gap (%)
Airport	3,238	1,327	59.02%
City	3,507	1,504	57.11%

```
Part 1: SQL Analysis

1. Total Requests by Status:
                Status  Total_Requests
0             Cancelled            1264
1     No Cars Available            2650
2        Trip Completed            2831

2. Hourly Supply vs Demand Gap:
   Hour  Supply  Demand_Gap
0  None    2831        3914

3. Requests by Pickup Point:
   Pickup point  Completed  Cancelled  No_Cars
0       Airport       1327        198     1713
1          City       1504       1066      937

4. Supply—Demand Gap Percentage:
   Pickup point  Total_Requests  Completed  Gap_Percentage
0       Airport            3238       1327           59.02
1          City            3507       1504           57.11
```

Fig. 4: Results of SQL Analysis

<u>**Insight:**</u>
The Airport shows a slightly higher supply–demand gap compared to the City, mainly due to lower driver availability during peak hours.

7. Python EDA Summary

Techniques Used:

- pandas - data manipulation.
- matplotlib & seaborn – visualization.
- datetime features - analyze hourly trends.

Core Insights:

- Supply–demand gap ≈ **58–60%**
- Early morning and late-night requests show highest unavailability
- Airport requests face driver shortage; City requests face higher cancellations

8. Recommendations

1. **Increase driver incentives** for early morning (4–8 AM) and late-night (10 PM–1 AM) shifts.
2. **Improve coordination** for Airport pickups, possibly with pre-scheduled driver allocation.
3. **Introduce penalty or warning system** for frequent driver cancellations.
4. **Monitor demand patterns** weekly to adjust supply dynamically.

9. Conclusion

The analysis clearly highlights a **major supply–demand gap** in Uber operations, especially during off-peak hours and Airport pickups.
By optimizing driver allocation and encouraging participation during low-supply periods, Uber can significantly reduce unfulfilled requests and improve customer satisfaction.