**Uber Request Data Analysis Report**

1. **Project Overview**

Exploratory Data Analysis (EDA) of Uber ride request data is the goal of this study in order to find trends in demand, cancellations, and supply-related problems. Request timestamps, pickup locations, driver IDs, trip statuses, and drop timestamps are among the important details included in the collection. The three main stages of the investigation were Python-based wrangling and visualisation using Google Colab, SQL-based querying, and Excel-based cleaning and dashboarding.

1. **Problem Statement**

High cancellation rates and frequent car shortages during peak hours are just two of the service quality problems Uber has been dealing with. A bad customer experience is the result of these problems. To optimise operations and driver allocation, the organisation needs a comprehensive picture of when and where such challenges arise.

1. **Business Objective**

The business objective is to derive actionable insights from ride request data to:

* Understand hourly demand patterns
* Identify cancellation trends
* Analyze driver unavailability periods
* Compare airport vs city pickups
* Measure average trip duration

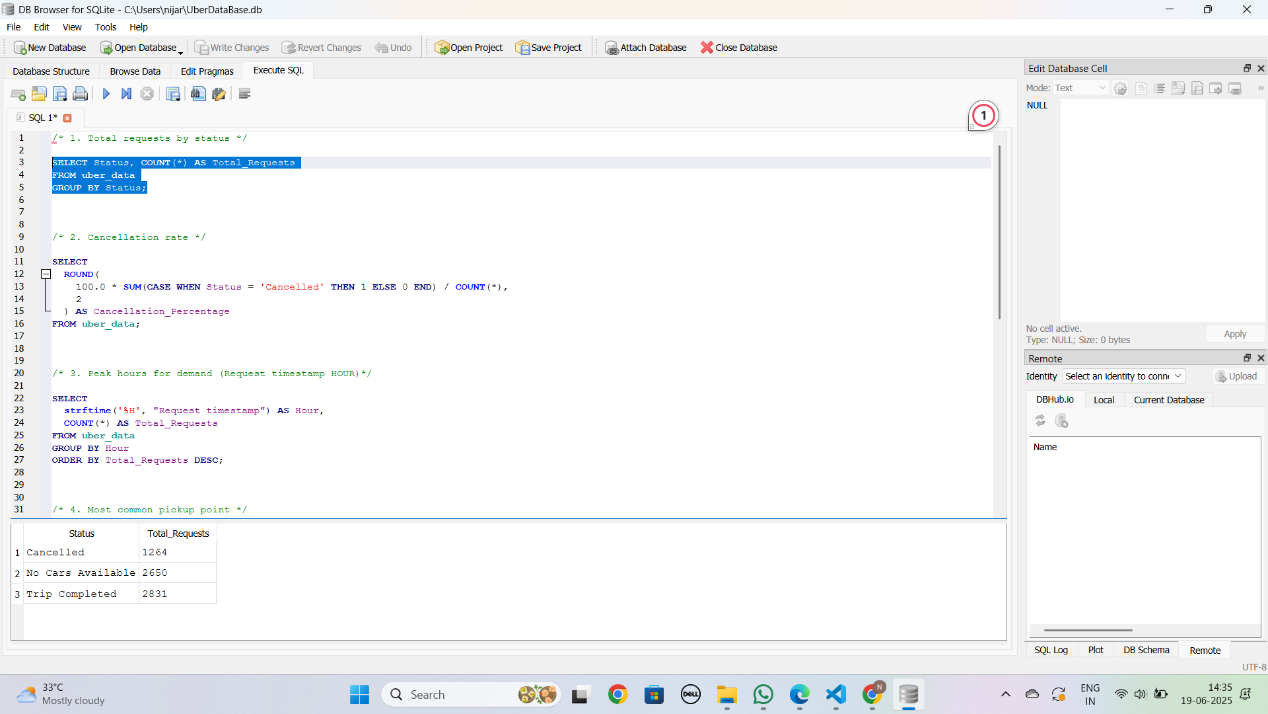
**Tools Used:** MS Excel, DB Browser for SQLite, Google Collab (Python)

1. **SQL Insights**

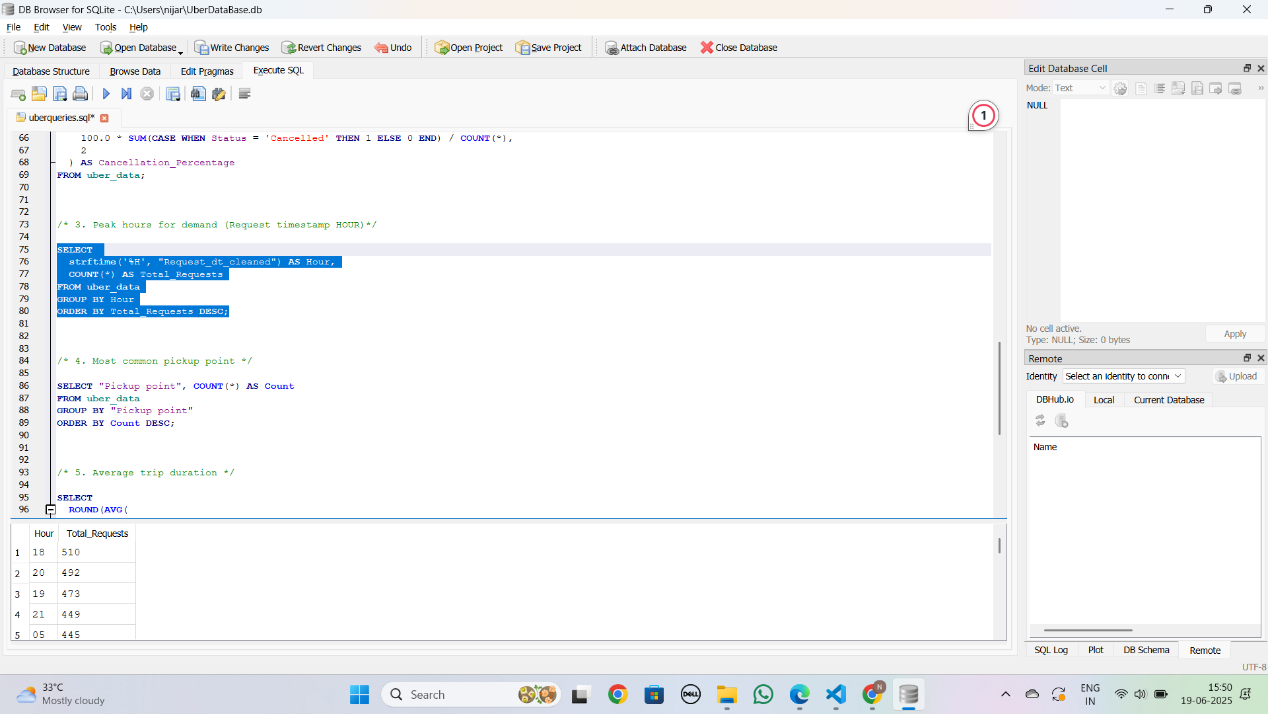
* Loaded the dataset into DB Browser and renamed the table to uber\_data
* Cleaned and added datetime columns using SQL.

**Key Queries & Outputs:**

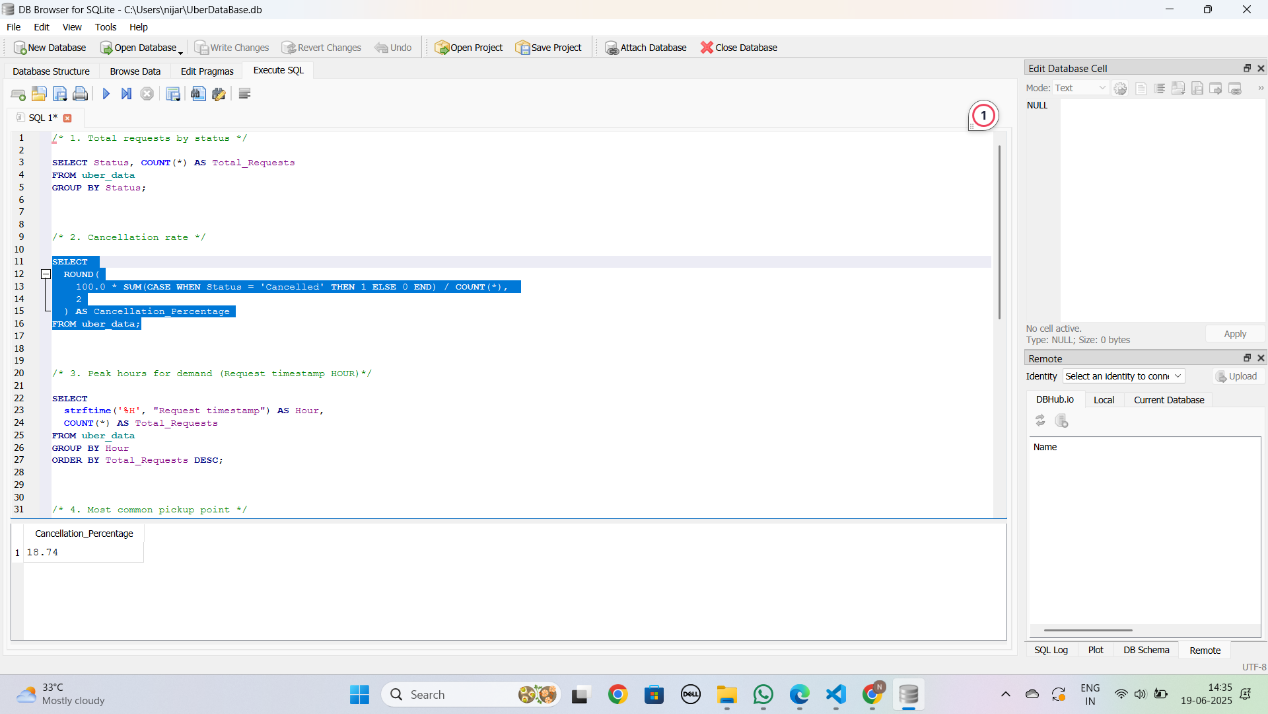
* Total requests by status



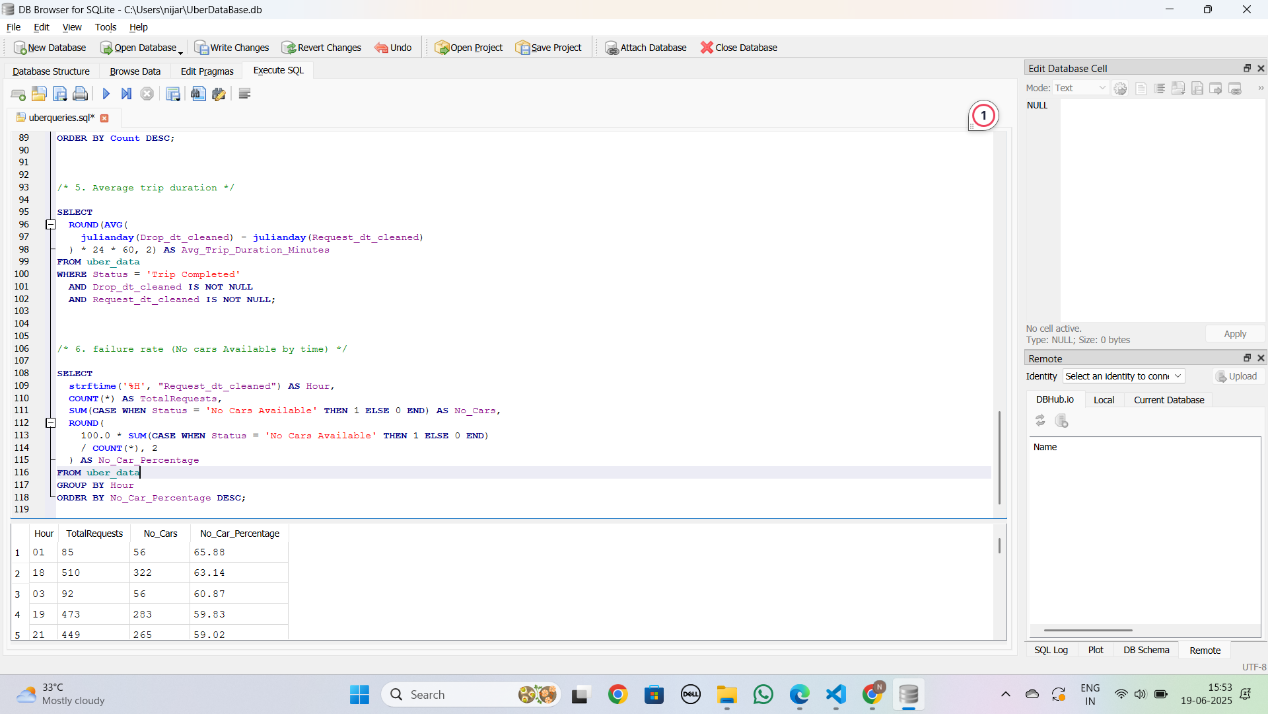
* Requests by hour (to find peak demand)



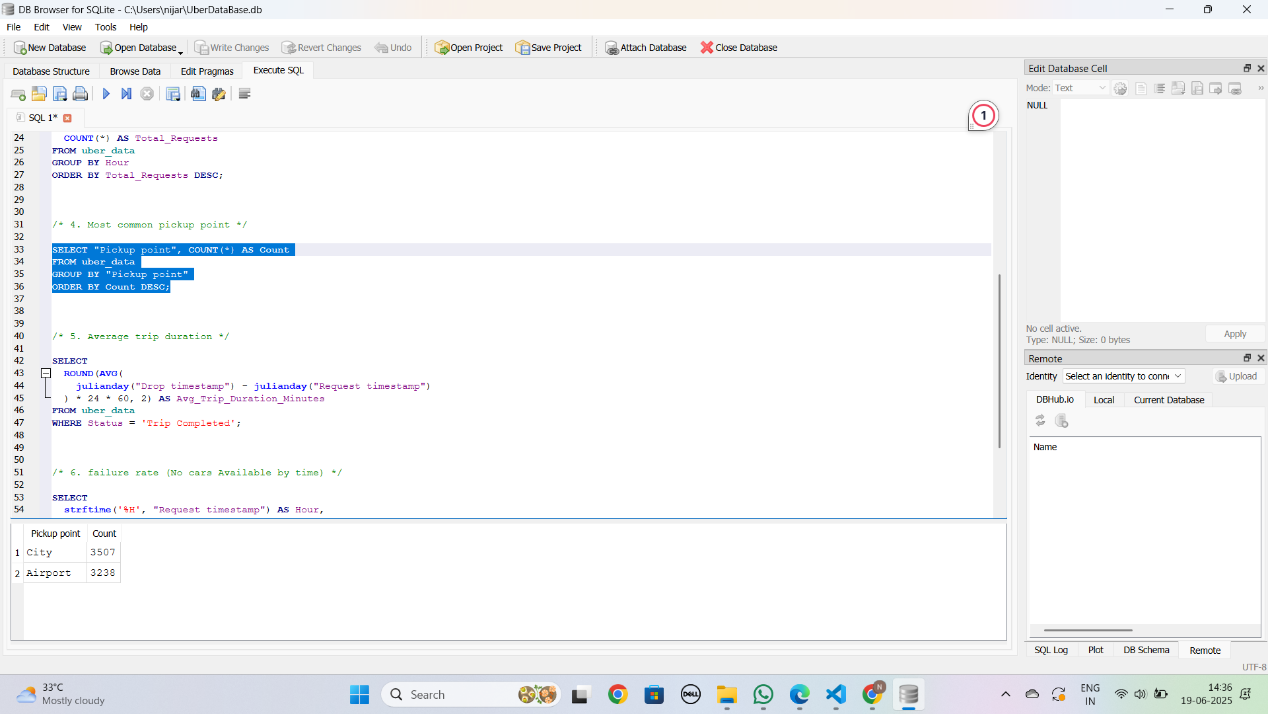
* Cancellation Percentage



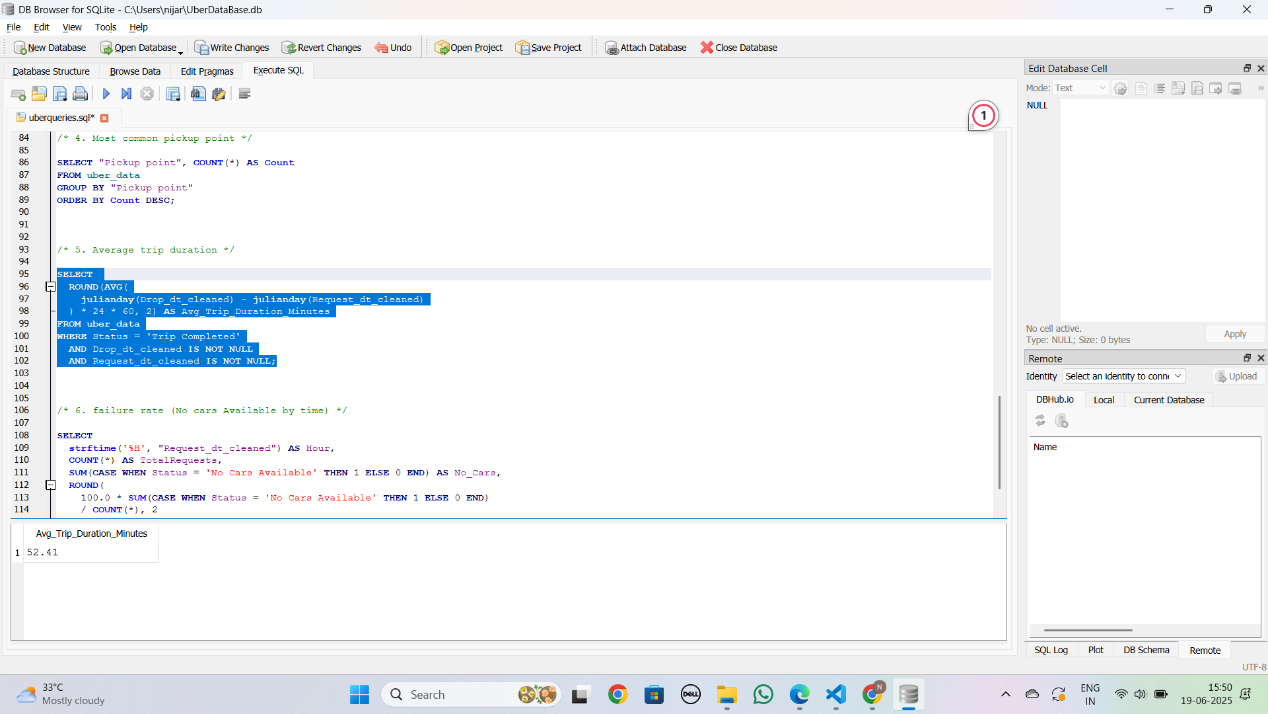
* Cancellations and No Car cases by hour



* Most frequent pickup point



* Average trip duration (for completed trips)

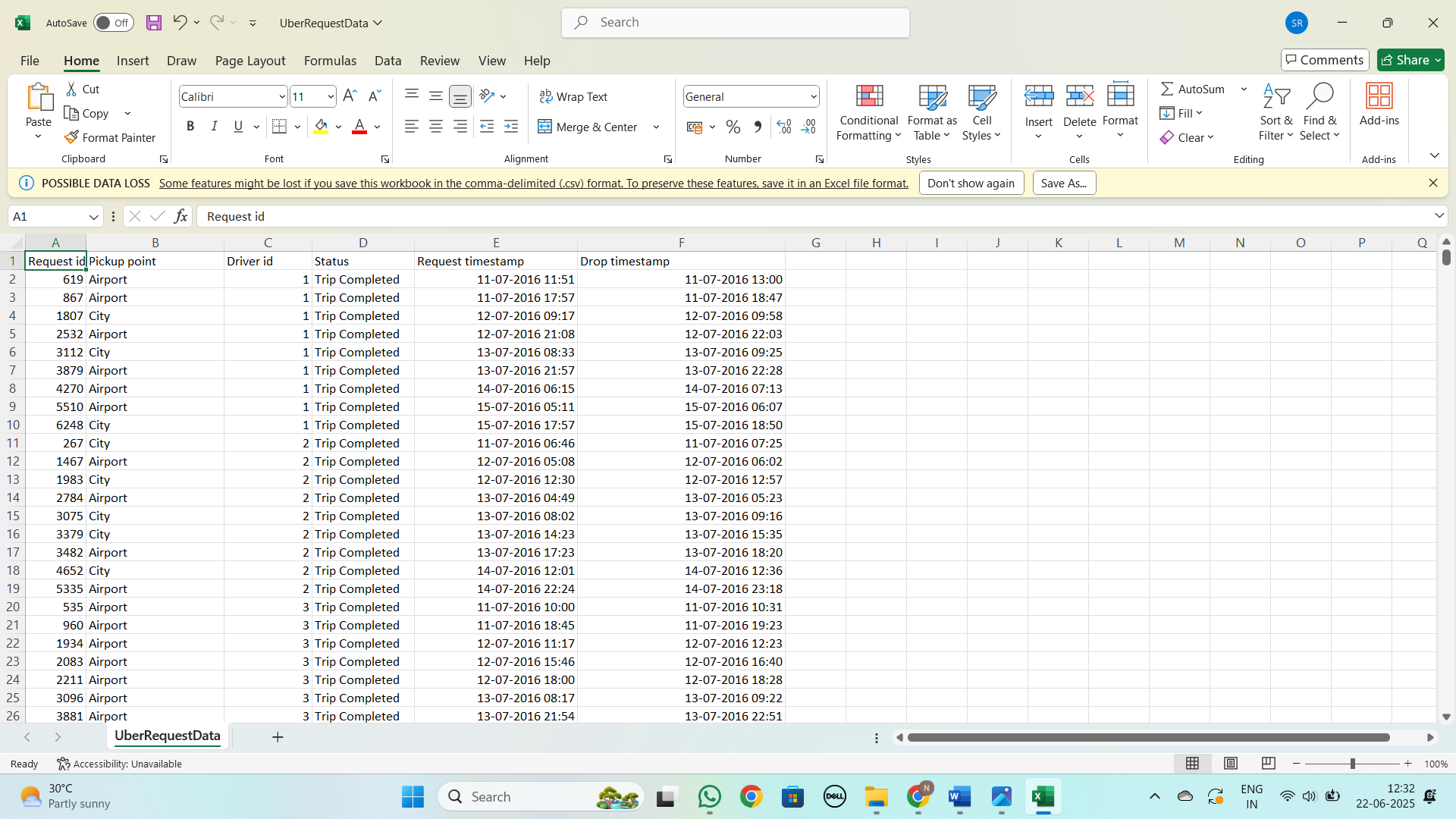


These queries provided foundational metrics later used in Excel dashboards.

1. **Excel Phase: Cleaning and Dashboarding**

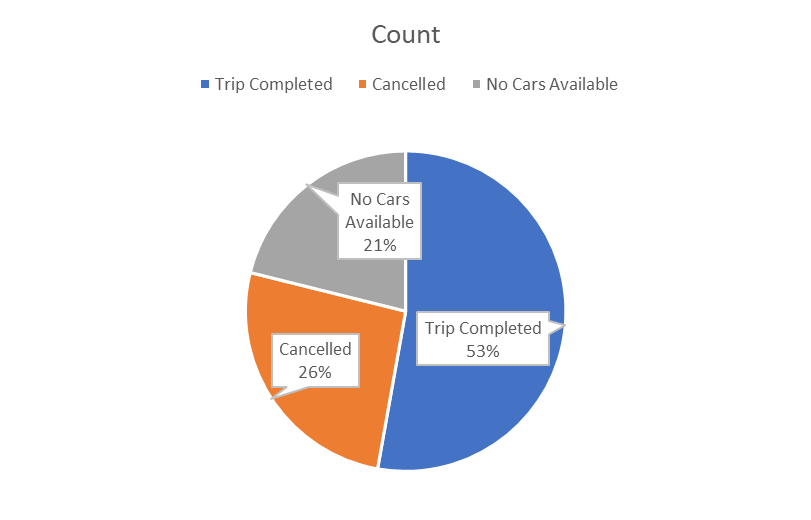
1) Data Cleaning:

* Verified presence of required columns: Request ID, Pickup Point, Driver ID, Status, Request Timestamp, Drop Timestamp
* Converted timestamps into uniform format (dd-mm-yyyy hh:mm)
* Replaced 'NA' with proper text where needed
* No empty columns or unnecessary fields

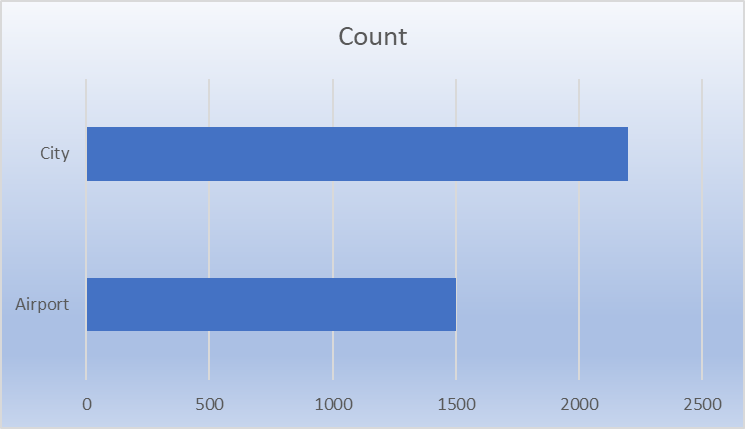


2) Dashboard Built (Excel):

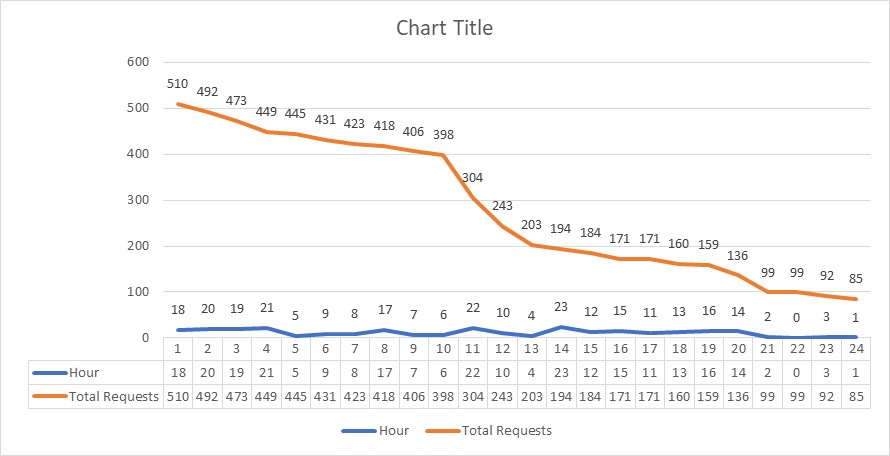
* Pie Chart: Distribution of request status (Completed, Cancelled, No Cars Available)



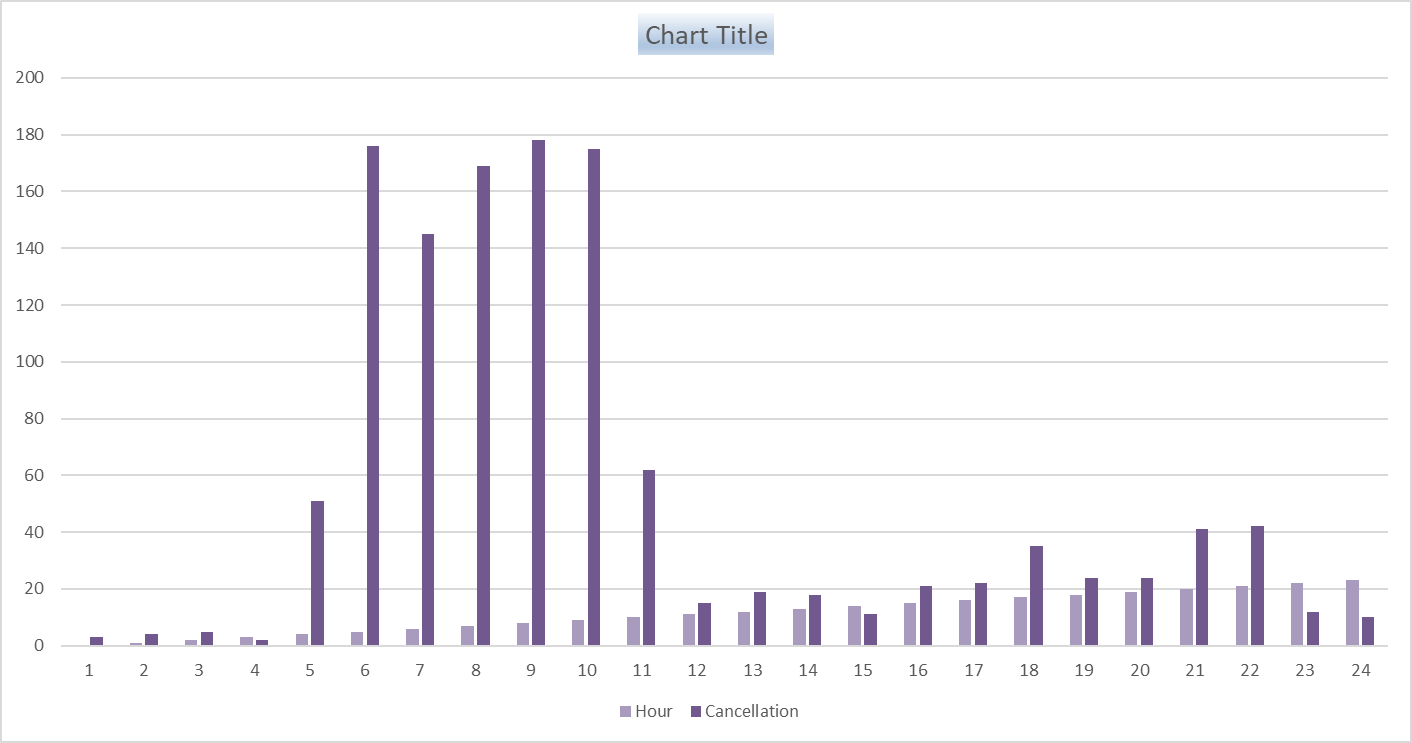
* Bar Chart: Frequency of pickup points (City vs Airport)



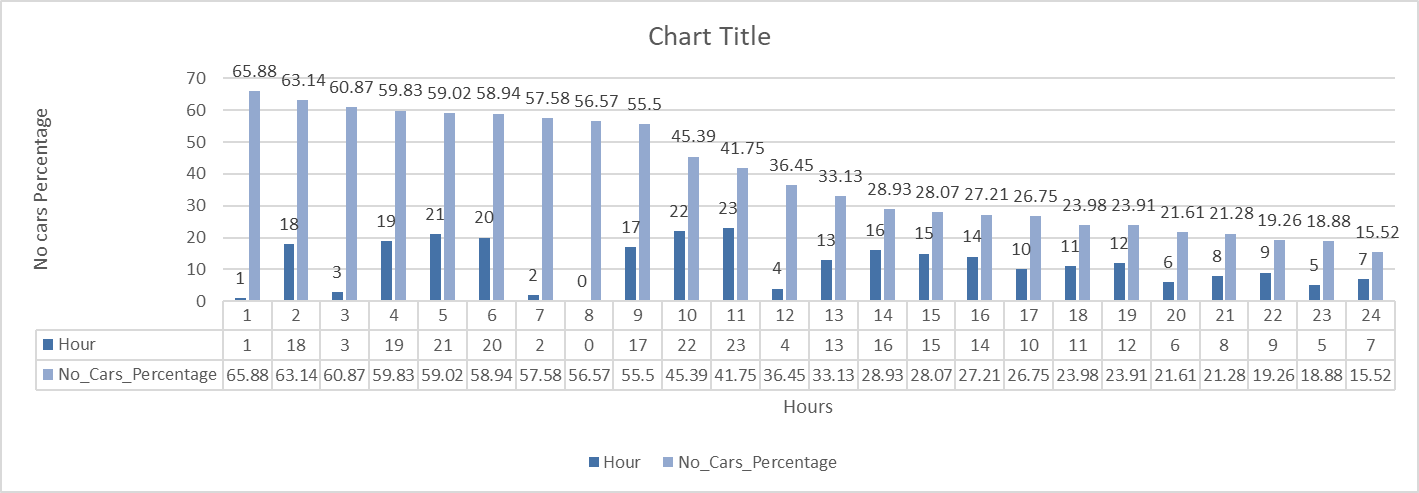
* Line Chart: Total requests per hour



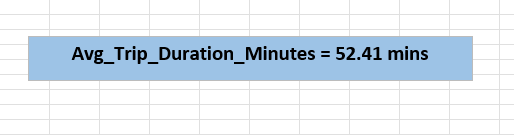
* Column Chart: Trip cancellations by hour



* Column Chart: No cars available % per hour



* Highlight Box: Average trip duration (in minutes)



**6. Python (Colab) Phase: Data Wrangling**

* Converted timestamp columns to datetime format
* Extracted the hour from the request time
* Calculated trip duration in minutes
* Created boolean flags for cancellation and no car availability
* Checked for and removed duplicates (none found)
* Preserved missing values in the Drop timestamp and Driver ID as they are informative

**7. Data Visualization & Storytelling**

Upcoming steps include:

* Creating Python-based charts using matplotlib/seaborn/plotly
* Exploring relationships between hour, status, pickup point, and duration
* Storytelling using visual patterns to recommend operational improvements.

**8. Conclusion**

This project demonstrates a full-cycle EDA approach, combining Excel, SQL, and Python tools. We identified patterns such as high cancellations in early mornings, peak demand during office hours, and major differences between city and airport rides. These insights will help Uber improve service reliability, driver deployment, and customer satisfaction.