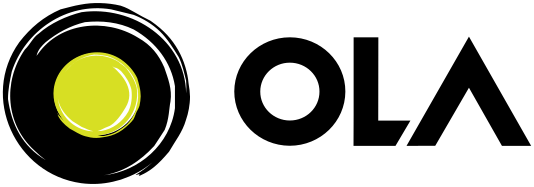
****

|  |  |
| --- | --- |
| **Project Title** | **Ola Ride Insights** |
| **Skills take away From This Project** | **SQL querying, data preprocessing, Power BI visualization, Streamlit app development, and business intelligence insights.** |
| **Domain** | **Ride-Sharing & Mobility Analytics** |

**Problem Statement:**

The rise of ride-sharing platforms has transformed urban mobility, offering convenience and affordability to millions of users. OLA, a leading ride-hailing service, generates vast amounts of data related to ride bookings, driver availability, fare calculations, and customer preferences. However, deriving actionable insights from this data remains a challenge. To enhance operational efficiency, improve customer satisfaction, and optimize business strategies, this project focuses on analyzing OLA’s ride-sharing data. By leveraging data analytics, visualization techniques, and interactive applications, the goal is to extract meaningful insights that can drive data-informed decisions. The project will involve cleaning and processing raw ride data, performing exploratory data analysis (EDA), developing a dynamic Power BI dashboard, and creating a Streamlit-based web application to present key findings in an interactive and user-friendly manner.

**Business Use Cases:**

* Identifying peak demand hours and optimizing driver allocation.
* Analyzing customer behavior for personalized marketing strategies.
* Understanding pricing patterns and surge pricing effectiveness.
* Detecting anomalies or fraudulent activities in ride data.

**Approach:**

**Data Understanding & Exploration**

* Load and examine the dataset structure.
* Identify key variables like ride status, payment method, and ratings.
* Perform initial exploratory data analysis (EDA).

**Data Cleaning & Preprocessing**

* Handle missing or inconsistent values.
* Convert data types and standardize formats.
* Create derived features if necessary for better insights.

**SQL Query Development**

* Write queries to extract insights (e.g., ride trends, cancellations, ratings).
* Optimize queries for performance and accuracy.
* Validate results against the dataset.

**Power BI Dashboard Creation**

* Design interactive visualizations for ride trends, revenue, and cancellations.
* Use filters and slicers for dynamic data exploration.
* Integrate KPIs and metrics for business insights.

**Streamlit Application Development**

* Create a user-friendly UI to display SQL query results.
* Implement interactive filters and search options.
* Embed Power BI visuals into the Streamlit app for a complete analytics experience.

**Project Documentation & Deployment**

* Document insights, queries, and dashboard explanations.
* Ensure the Streamlit app is deployed and accessible.
* Present findings with business-oriented storytelling.

**SQL Questions**

1. Retrieve all successful bookings:

2. Find the average ride distance for each vehicle type:

3. Get the total number of cancelled rides by customers:

4. List the top 5 customers who booked the highest number of rides:

5. Get the number of rides cancelled by drivers due to personal and car-related issues:

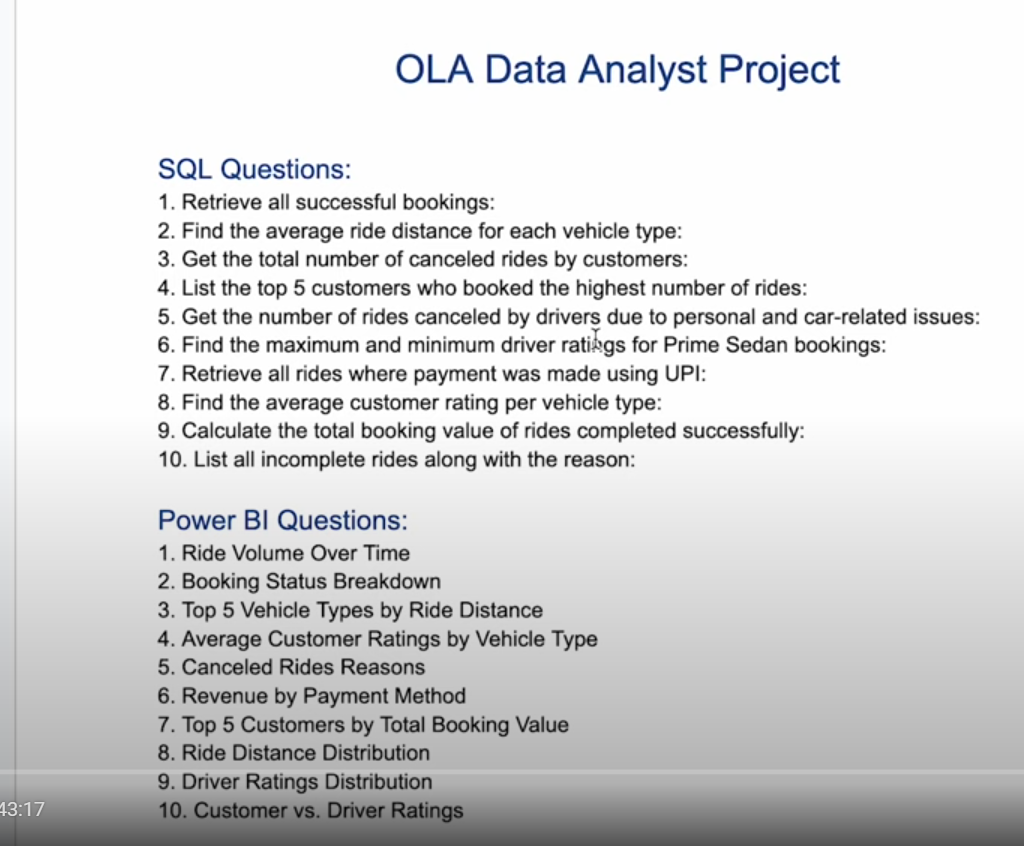
6. Find the maximum and minimum driver ratings for Prime Sedan bookings:

7. Retrieve all rides where payment was made using UPI:

8. Find the average customer rating per vehicle type:

9. Calculate the total booking value of rides completed successfully:

10. List all incomplete rides along with the reason



**Power BI Questions**

1. Ride Volume Over Time

2. Booking Status Breakdown

3. Top 5 Vehicle Types by Ride Distance

4. Average Customer Ratings by Vehicle Type

5. Canceled Rides Reasons

6. Revenue by Payment Method

7. Top 5 Customers by Total Booking Value

8. Ride Distance Distribution Per Day

9. Driver Ratings Distribution

10. Customer vs. Driver Ratings

**Segregation of the views**

1. Overall

- Ride Volume Over Time

- Booking Status Breakdown

2. Vehicle Type

- Top 5 Vehicle Types by Ride Distance

3. Revenue

- Revenue by Payment Method

- Top 5 Customers by Total Booking Value

- Ride Distance Distribution Per Day

4. Cancellation

- Cancelled Rides Reasons (Customer)

- cancelled Rides Reasons(Drivers)

5. Ratings

- Driver Ratings

- Customer Ratings

**Results:**

* Interactive dashboard and application showcasing key insights.
* Streamlined access to booking trends, ratings, and revenue analysis.
* Actionable insights for improving ride experience and service efficiency.

**Project Evaluation metrics:**

* Accuracy of SQL queries and insights.
* Effectiveness of Power BI visualizations.
* Usability and responsiveness of the Streamlit application.

**Technical Tags:**

Python, Pandas, NumPy, Matplotlib, Seaborn, SQL, Data Visualization, Power BI, Streamlit, Data Cleaning, Feature Engineering, EDA.

**Data Set:**

Data set: [**OLA PROJECT**](https://drive.google.com/drive/folders/12onyQF6Zh_M8voCPSdSpt4BSfw6T3tg3?usp=sharing)

**Project Deliverables:**

* Maintain clean and optimized SQL queries.
* Ensure Power BI dashboards are interactive and user-friendly.
* Follow coding best practices in Streamlit development.

**Project Guidelines:**

* **Clean the dataset** – Handle missing values and format data properly.
* **Write SQL queries** – Extract insights using joins, aggregations, and filters.
* **Create Power BI dashboard** – Build interactive and clear visualizations.
* **Develop Streamlit app** – Display SQL results and embed Power BI visuals.
* **Follow coding standards** – Write clean, well-documented code with Git version control.
* **Submit project files** – Provide SQL queries, Power BI dashboard, and Streamlit app.
* **Present findings** – Explain insights with a business-focused approach.

**Timeline:**

The project must be completed and submitted **within 7 days from the assigned date**

**Reference**

| **Project Live Evaluation** | [**Project Live Evaluation**](https://docs.google.com/document/u/0/d/1QisLD2kqDWFZJG2oDknKn2eMGi-Xq8oFPgA7UWSbcIQ/edit) |
| --- | --- |
| **EDA Guide** | [**Exploratory Data Analysis (EDA) Guide**](https://docs.google.com/document/d/1tHiTU1X9UwXSLySpJ-FVCohlf_8xpXwa75vlK9S6wl8/edit?usp=sharing) |
| **EDA Detail Document** | [**EDA\_INDEPTH**](https://docs.google.com/document/d/1bzW_w-9m5WfyjVWvid2x8GR7C5mB4ONiMaqWzXTnQf0/edit?usp=sharing) |
| **Capstone Explanation Guideline** | [**Capstone Explanation Guideline**](https://docs.google.com/document/d/1gbhLvJYY7J73lu1g9c6C9LRJvYemiDOdRDAEMe632w8/edit) |
| **GitHub Reference** | [**How to Use GitHub.pptx**](https://docs.google.com/presentation/d/1XHCbgUOqbcXNUyQ87vTlKdKRgAbBxtkA/edit?usp=sharing&ouid=109735616107417446342&rtpof=true&sd=true) |
| **Streamlit recording (English)** | [**Special session for STREAMLIT(11/08/2024)**](https://docs.google.com/document/d/1aR3pUZFlCi8gicpF6aPHPESeFdOtGMlfob5PckresZk/edit?usp=sharing) |
| **PowerBi and SQL connection** | [**Connect MySQL to PowerBI**](https://docs.google.com/document/d/10UH4dMidSXX2kSEqhevlK5GahJfvE2wd77FJkvSPZog/edit?usp=sharing) |
| **Steps to Create a Free Power BI Work Account** | [**Steps to Create a Free Power BI Work Account**](https://docs.google.com/document/d/1t1VVd584dl6Xdp_HKVbx4xR2BtTqjV1McHt0R7_PTFQ/edit?usp=sharing) |
| **Power Bi Slides** | [**Ola-Slides.pptx**](https://docs.google.com/presentation/d/19Nz8bf0BP1VyfQyNLndOMnJcHrHFvGQi/edit?usp=sharing&ouid=109735616107417446342&rtpof=true&sd=true) |
| **DATA CLEANING & PREPROCESSING EDA (Tamil)** | [**Project Excellence Series: Guided Learning & Problem Solving [EDA](Tamil)**](https://docs.google.com/document/d/120-XJMD8ol-WBt3GaoSsdkdQD--i-TcS5t3N4Ur9LhQ/edit?usp=sharing) |
| **DATA CLEANING & PREPROCESSING EDA (English)** | [**Project Excellence Series: Guided Learning & Problem Solving [EDA](English)**](https://docs.google.com/document/d/1kbAZAtGhABWK5S7IAWXs3qiQ8PZybqCpv67S0DwppJ0/edit?usp=sharing) |
| **Project Orientation (English)** | [**Ola Ride Insights Session Recordings[DS-C-WD-E-B48](08/05/2025):**](https://docs.google.com/document/d/1wIPKDDLlH1C4EfhhpB_r1wlKwm9ysM7Zb7NfTobpxRA/edit?usp=sharing) |