



Data Analyst Project

📌 Project Overview

This project simulates OLA ride data from **Bengaluru** for **July 2024** (~1,00,000 rows). It includes booking details, customer behavior, vehicle types, ratings, cancellation reasons, and revenue insights.

📊 Dataset Features

- Total Rows: 100,000
- Key Columns:
 - **Date, Time, Booking_ID, Booking_Status, Customer_ID**
 - **Vehicle_Type**: Auto, Prime Sedan, Prime SUV, Prime Plus, Mini, Bike, E-Bike
 - **Pickup/Drop Location** (50 areas in Bengaluru)
 - **VTAT, CTAT, Booking_Value, Ride_Distance**
 - **Driver_Ratings, Customer_Rating**
 - **Cancellation Reasons** (Driver/Customer)
 - **Incomplete Rides**

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:

SQL Answers:

Create Database Ola;

Use Ola;

-- 1. Successful bookings

```
CREATE VIEW Successful_Bookings AS
```

```
SELECT * FROM bookings WHERE Booking_Status = 'Success';
```

-- 2. Average ride distance by vehicle type

```
CREATE VIEW ride_distance_for_each_vehicle AS
```

```
SELECT Vehicle_Type, AVG(Ride_Distance) AS avg_distance FROM bookings GROUP BY Vehicle_Type;
```

-- 3. Total cancelled rides by customers

```
CREATE VIEW cancelled_rides_by_customers AS
```

```
SELECT COUNT(*) FROM bookings WHERE Booking_Status = 'cancelled by Customer';
```

-- 4. Top 5 customers by number of bookings

```
CREATE VIEW Top_5_Customers AS
```

```
SELECT Customer_ID, COUNT(Booking_ID) AS total_rides  
FROM bookings GROUP BY Customer_ID ORDER BY total_rides DESC LIMIT 5;
```

-- 5. Cancellations by driver due to personal/car issues

```
CREATE VIEW Rides_cancelled_by_Drivers_P_C_Issues AS
```

```
SELECT COUNT(*) FROM bookings WHERE cancelled_Rides_by_Driver = 'Personal & Car related issue';
```

-- 6. Max & Min driver ratings for Prime Sedan

```
CREATE VIEW Max_Min_Driver_Rating AS  
SELECT MAX(Driver_Ratings) AS max_rating, MIN(Driver_Ratings) AS min_rating  
FROM bookings WHERE Vehicle_Type = 'Prime Sedan';
```

-- 7. Rides paid with UPI

```
CREATE VIEW UPI_Payment AS  
SELECT * FROM bookings WHERE Payment_Method = 'UPI';
```

-- 8. Average customer rating per vehicle type

```
CREATE VIEW AVG_Cust_Rating AS  
SELECT Vehicle_Type, AVG(Customer_Rating) AS avg_customer_rating FROM bookings  
GROUP BY Vehicle_Type;
```

-- 9. Total revenue from successful rides





```
CREATE VIEW total_successful_ride_value AS  
SELECT SUM(Booking_Value) AS total_successful_ride_value FROM bookings WHERE  
Booking_Status = 'Success';
```




-- 10. Incomplete rides and their reasons

```
CREATE VIEW Incomplete_Rides_Reason AS  
SELECT Booking_ID, Incomplete_Rides_Reason FROM bookings WHERE Incomplete_Rides  
= 'Yes';
```

Power BI Insights

I also built an interactive dashboard in Power BI that visualizes the following:

-  **Ride Volume Over Time**
-  **Booking Status Breakdown**
-  **Top Vehicle Types by Distance**
-  **Customer & Driver Ratings**

-  Revenue by Payment Method
-  Top 5 Customers by Booking Value
-  Cancellation Reasons

Files Included

- [OLA_Dashboard.pbix](#) – Power BI File
- [SQL_Code.sql](#) – SQL Queries
- [Project_Screenshots/](#) – Dashboard screenshots
- [README.md](#) – Project summary (this file)

Contact

Chitransh Rahangdale

 chinturahangdale579@gmail.com

 [LinkedIn](#)