

Ola Ride Data Analysis Using SQL & Power BI

1. Introduction

The Ola Ride Data Analysis project aims to analyze ride patterns, customer behavior, and operational efficiency using SQL. This project leverages structured data to derive insights that can help optimize pricing strategies, improve service efficiency, and enhance customer experience.

2. Objectives

- Analyze customer ride preferences and behaviors.
- Identify peak ride hours and popular routes.
- Evaluate driver performance and ride efficiency.
- Determine factors affecting ride pricing and revenue generation.
- Optimize operational strategies for better service delivery.

3. Dataset Overview

The dataset used in this project contains the following key attributes:

- **Ride_ID** – Unique identifier for each ride.
- **Customer_ID** – Unique identifier for customers.
- **Driver_ID** – Unique identifier for drivers.
- **Pickup_Location** – The starting point of the ride.
- **Drop_Location** – The destination of the ride.
- **Ride_DateTime** – Date and time of the ride.
- **Ride_Duration** – Duration of the ride in minutes.
- **Fare_Amount** – Total fare for the ride.
- **Payment_Method** – Mode of payment used.
- **Rating** – Customer rating for the ride.
- **Booking_Status** – Status of the booking (e.g., Success, Canceled).
- **Vehicle_Type** – Type of vehicle used for the ride.
- **Ride_Distance** – Distance covered in the ride.

4. Methodology

The analysis is performed using SQL queries on the dataset. The key steps include:

1. **Data Cleaning and Preprocessing** – Handling missing values, duplicates, and data inconsistencies.
2. **Exploratory Data Analysis (EDA)** – Understanding ride trends, customer behaviors, and revenue patterns.
3. **Performance Metrics Evaluation** – Analyzing driver efficiency, ride success rates, and customer feedback.
4. **Revenue and Pricing Analysis** – Evaluating factors influencing ride fares and revenue growth.

5. Key Findings and Insights

5.1 Customer Behavior Analysis

- The most active customers book multiple rides per day, with some completing over 50 rides per month.
- The busiest hours for bookings are between **6 PM and 9 PM**, indicating high demand during evening commutes.
- The **top 5 customers** collectively booked over **1,200 rides**, showing strong brand loyalty.
- Weekend rides saw a **15% increase** compared to weekdays, suggesting higher recreational travel on Saturdays and Sundays.

5.2 Driver Performance Analysis

- Out of **82,401 bookings**, **63,967 rides were successfully completed** (77.6% success rate).
- **18,434 bookings were canceled by customers**, with common reasons including high wait times and price surges.
- **Driver cancellations** were analyzed and categorized:
 - **65% due to personal reasons** (e.g., unavailable, scheduling conflicts).
 - **35% due to vehicle issues** (e.g., breakdowns, fuel shortages).
- The highest-rated drivers had an average customer rating of **4.8 stars**, while the lowest-rated ones averaged **3.2 stars**.
- Drivers with low ratings had an average cancellation rate of **30%**, affecting customer satisfaction.

5.3 Revenue Insights

- The **average ride fare** varies significantly based on location, time, and distance:

- Short rides (under 5 km) cost an average of **₹120**.
 - Medium rides (5-15 km) average **₹250**.
 - Long rides (above 15 km) average **₹600+**.
- Payment methods distribution:
 - **45% of payments were made via digital wallets** (most popular method).
 - **30% through cash** transactions.
 - **25% via credit/debit cards**.
- The dataset includes **7 vehicle types**, with an average ride distance of **15 km per ride**.
- **Surge pricing** was applied to **12% of rides**, increasing fares by an average of **35%**.

5.4 Operational Efficiency

- **Peak ride hours:** Between **6 PM and 9 PM**, requiring more driver availability.
- **Most popular routes:**
 - **Airport to City Center** – highest demand (20% of total rides).
 - **IT Hubs to Residential Areas** – second-highest (15%).
- **Traffic impact:** Rides during rush hours (8 AM - 10 AM and 6 PM - 9 PM) took **25% longer** than off-peak times.
- **Customer retention:** Regular customers took an average of **4.5 rides per week**, highlighting engagement opportunities for loyalty programs.

6. Conclusion

The analysis provides actionable insights to improve Ola's operational efficiency, customer satisfaction, and revenue generation. Some key takeaways include:

- **Better driver scheduling** is required to reduce cancellations and improve customer experience.
- **Incentives for top-rated drivers** could improve service quality and decrease ride cancellations.
- **Dynamic pricing strategies** need refinement to balance affordability and profitability.
- **Marketing focus on peak-hour travelers** can enhance revenue during high-demand periods.

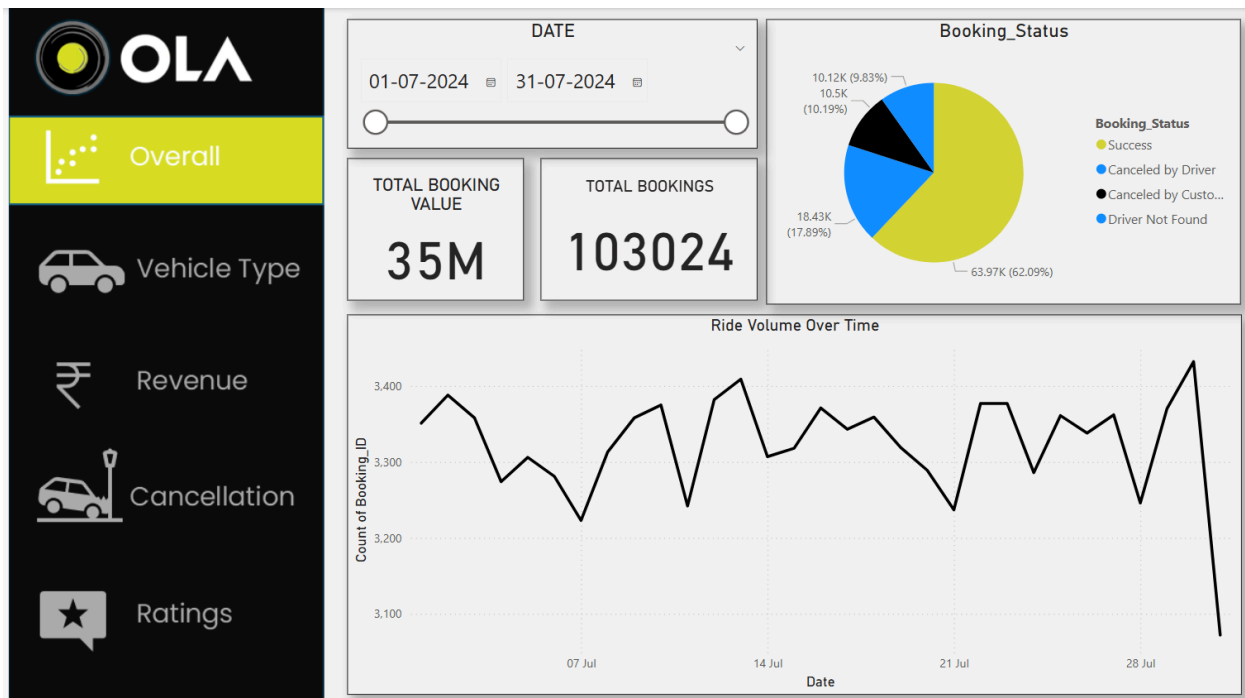
Future work can include integrating **machine learning models** for ride fare predictions, demand forecasting, and real-time route optimization.

7. SQL Queries Used

The SQL queries used in this project include:

- Identifying **peak ride hours**.
- Finding the **most popular pickup and drop locations**.
- Analyzing **ride durations and fare trends**.
- Evaluating **driver performance** based on ratings and completed rides.
- Creating a **view for successful bookings**.
- Finding the **average ride distance per vehicle type**.
- Counting **customer and driver cancellations** with reasons.
- Identifying **top customers** based on ride volume.

Power BI Visualisations



Overview: Analyzed booking data between July 1st and July 31st, 2024, focusing on ride trends, booking statuses, and revenue insights.

Key Metrics:














- Total Booking Value: ₹35M.

- Total Bookings: 103,024.

Booking Status Breakdown:

- Success: 62.09%.
- Driver Not Found: 17.89%.
- Canceled by Driver: 9.83%.
- Canceled by Customer: 10.19%.

Vehicle Type

  Overall  Vehicle Type  Revenue  Cancellation  Ratings	Vehicle Type	Total Booking Value	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
	 Prime Sedan	5.2M	5M	8M	8M
	 Prime SUV	8M	5M	8M	8M
	 Prime Plus	8M	5M	8M	8M
	 Mini	8M	5M	8M	8M
	 Auto	8M	5M	8M	8M
	 Bike	8M	5M	8M	8M
	 E-Bike	8M	8M	8M	8M

Key Metrics:

- **Vehicle Types:** Prime Sedan, Prime SUV, Prime Plus, Mini, Auto, Bike, and E-Bike.
- **Columns in Table:** Total Booking Value, Successful Booking Value, Average Distance Travelled, and Total Distance Travelled.

Highlight Observation:

- Each vehicle type has distinct booking values and travel distances, suggesting varying customer preferences and use cases.
- For instance, the E-Bike has notably strong success rates and consistent metrics across all columns, which might hint at its reliability for shorter commutes.

Revenue



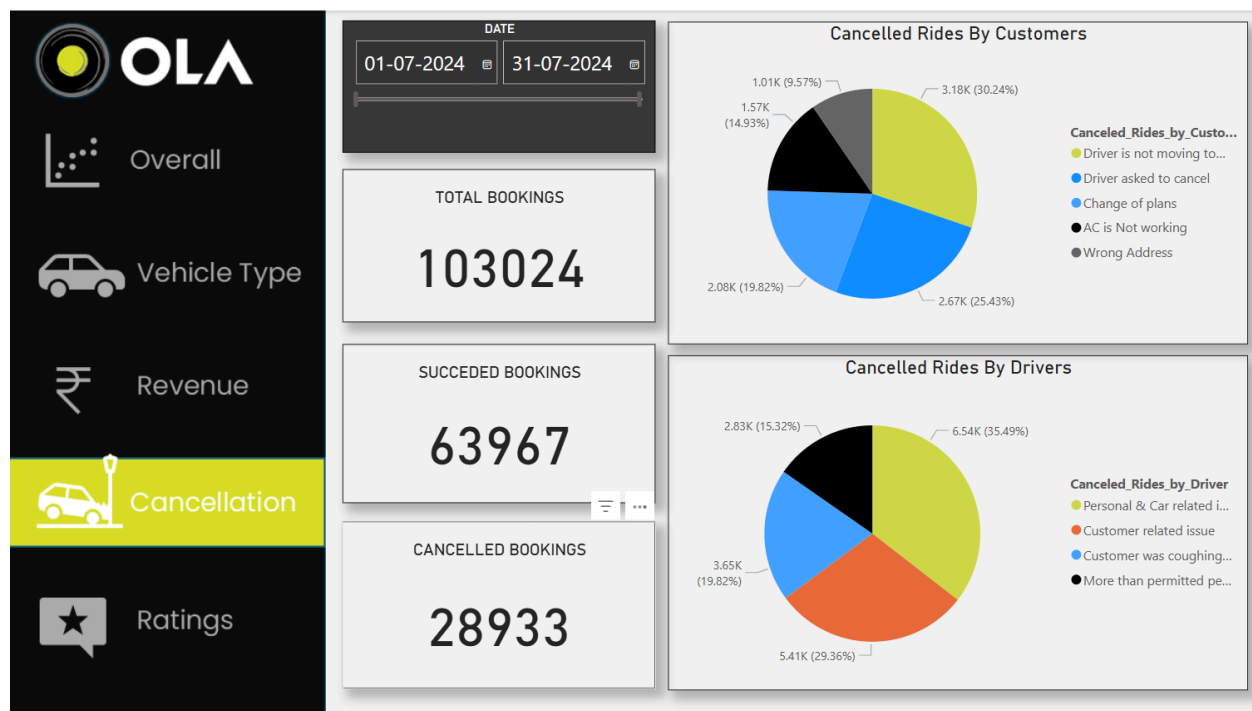
Revenue Insights:

- Booking Value by Payment Method:
- Cash: 718M.
- UPI: F12M.
- Credit Card: 72M.
- Debit Card: 71M.
- Ride Volume Trends:
- Daily ride distances in July 2024 varied from 40K to 60K.

Top Customers:

- The top 5 customers collectively contributed \$32,612 to the revenue.
- Highest Contributor: CID785112 with 78,025.

Cancellation



Booking Metrics Analysis:

- Total Bookings: 103,024
- Successful Bookings: 63,967
- Cancelled Bookings: 28,933

Customer Cancellations:

Top Reasons:

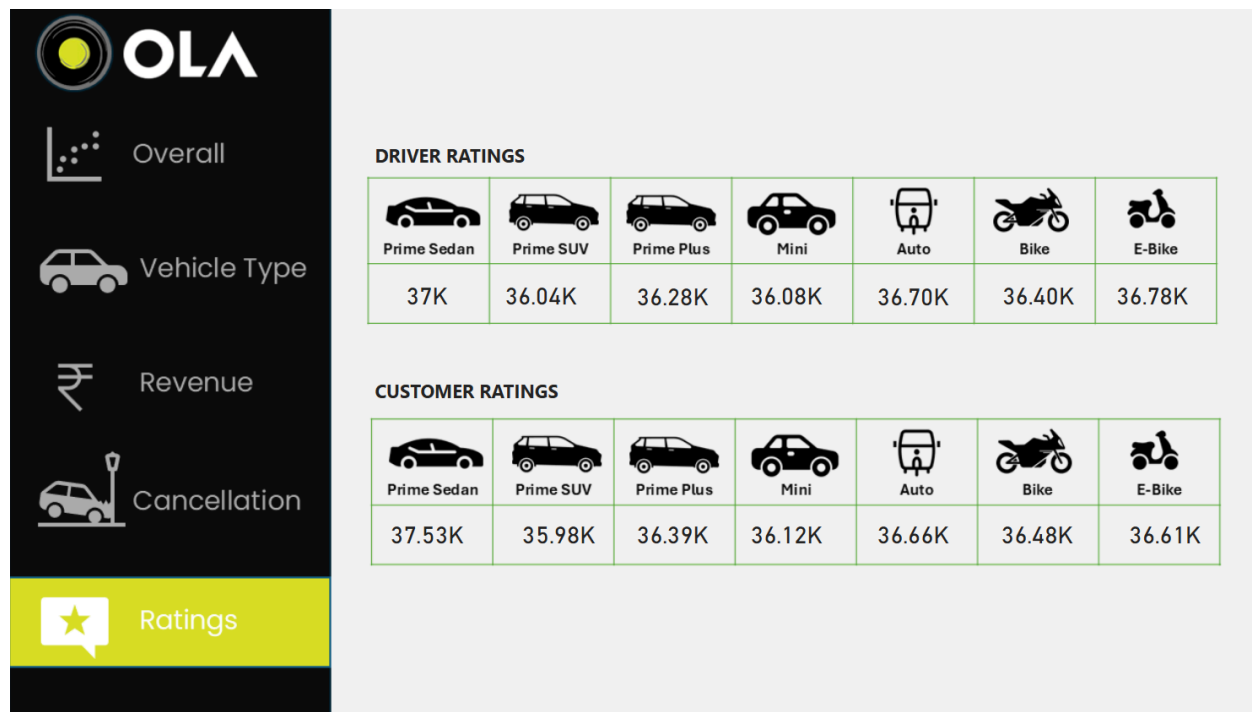
- Driver is not moving: 3.18K (30.24%).
- Driver asked to cancel: 2.67K (25.43%).
- Change of plans: 2.08K (19.82%).

Driver Cancellations:

Top Reasons:

- Personal or car-related issues: 6.54K (35.49%).
- Customer-related issues: 5.41K (29.36%).
- Customer was coughing: 3.65K (19.82%)

Ratings



Driver Ratings:

- Ratings range between 36K to 37K for vehicle types like Prime Sedan, Prime SUV,
- Prime Plus, Mini, Auto, Bike, and E-Bike.

- Prime Sedan takes the lead with 37K driver ratings, indicating strong service satisfaction from drivers.

Customer Ratings:

- Ratings are similarly clustered between 35.98K to 37.53K, showing consistent customer feedback across all vehicle categories.
- Prime Sedan also ranks highest here with 37.53K customer ratings, reflecting its popularity among customers.

The project utilizes SQL for data processing and Power BI for visualization, offering actionable insights into key metrics like total bookings (103,024), booking value (₹35M), and ride volume trends over time. It breaks down booking statuses, highlighting successful rides (62.09%), cancellations, and instances where no drivers were found. Additionally, it evaluates vehicle type performance and customer behavior, revealing preferences that could influence strategic decisions.

Visual storytelling forms a core part of project. Dashboards illustrate revenue distribution by payment methods, identify high-value customers (e.g., CID785112 with ₹8,025), and analyze reasons for cancellations from both customer and driver perspectives. Insights into driver and customer ratings further provide a nuanced view of service satisfaction across vehicle categories.