# **OLA Cabs Data Analysis Project Report**

**Duration Analyzed:** 1 July 2024 – 31 July 2024  
**Rows Processed:** 100,000+  
**Analyst:** Aakash Vishwakarma

# **1. Project Objective**

# The primary objective of this project was to generate actionable business insights for OLA Cabs through a comprehensive analysis of one month's large-scale ride booking data. By examining key operational metrics, identifying booking and cancellation patterns, and segmenting performance by vehicle type, the analysis aims to support enhanced business decision-making, operational optimization, and customer satisfaction.

# **2. Data Description**

# The dataset comprised over 100,000 ride booking records from July 2024. Each record contained variables such as:

* **Booking Information:**

Booking ID, date/time, status (success, cancelled by driver/customer, driver not found)

* **Customer & Driver Details:**

Ratings, IDs

* **Vehicle Type:**

Auto, Bike, eBike, Mini, Prime Plus, Prime Sedan, Prime SUV

* **Ride Metrics:**

Distance, ratings, revenue, payment method

* **Cancellation Reasons:**

Segregated for customers and drivers

**3. Data Processing & Methodology**

# The analytical methodology involved the following steps:

* **Data Extraction:**

Raw data was imported into SQL, ensuring data integrity and addressing missing or invalid entries.

* **Aggregation & Transformation:**

Multiple aggregate SQL views were developed to facilitate granular insights, including ride distance per vehicle, cancellation rates, and rating distributions.

* **KPI Identification:**

Focus was placed on impactful business Key Performance Indicators (KPIs), such as booking volume, total booking value, ride success/cancellation rates, payment preferences, and customer/driver sentiment.

* **Dashboard Development:**

Analyses and visualizations were integrated into a unified dashboard, reflecting both temporal trends and categorical performance breakdowns.

**4. Results & Key Insights**

# **A. Ride Volume & Revenue**

* **Total Bookings:**

103,024

* **Total Booking Value:**

35 million INR

* **Success Rate:**

Approximately 62% of rides were completed successfully.

* **Cancellation Rate:**

Approximately 28% (attributed to customer, driver, or driver not found).

* **Vehicle Type Performance:**

Prime Sedan and eBike categories demonstrated the highest volumes and revenue generation.

**B. Ride Distance & Customer Engagement**

* **Average Ride Distance:**

Approximately 14.19 km per ride.

* **Total Ride Distance:**

208,843.3 km, with Prime Sedan leading in vehicle-wise distance.

* **Top 5 Customers:**

These customers accounted for a substantial portion of rides and revenue.

| **Customer ID** | **Value (Rs)** |
| --- | --- |
| CID785112 | 8030 |
| CID308763 | 6280 |
| CID734557 | 6180 |
| CID353074 | 6110 |
| CID836942 | 6020 |

**C. Cancellations – Trends & Root Causes**

* **Predominant Customer Cancellation Reasons:**

“Change of plans,” “Driver is not moving,” and “Driver asked to cancel.”

* **Predominant Driver Cancellation Reasons:**

“Personal & Car related issue” and "Customer related issue."

* **Highest Cancellations:**

Observed in eBikes and Autos.

**D. Payment Preferences**

* **UPI:**

The most popular online payment method , generating 14.2 million INR in booking value.

* **Other Methods:**

Credit/Debit Card, Cash, and Other Mobile (OM) accounted for a smaller share.

**E. Ratings & Service Quality**

* **Customer Ratings:**

Averaged at 4.0 across all vehicle types.

* **Driver Ratings:**

Displayed moderate variance, with a max-min analysis performed for Prime Sedan.

**5. SQL Analysis Snapshots**

# SQL Views created to automate and scale insights include:

* Successful\_Bookings:

Extracts all successful bookings.

* Ride\_Distance\_For\_Each\_Vehicle:

Segments average ride distance by vehicle type.

* Cancelled\_Rides\_By\_Customers:

Counts all rides cancelled by customers.

* Top\_5\_Customers:

Identifies the highest-frequency riders.

* Rides\_Cancelled\_By\_Drivers\_P\_C\_Issues:

Focuses analysis on driver-initiated cancellations due to personal/car reasons.

* Max\_Min\_Driver\_Rating:

Identifies best/worst driver ratings among Prime Sedan bookings.

* UPI\_Payment:

Segments all bookings paid via UPI.

* Avg\_Cust\_Rating:

Averages customer ratings by vehicle segment.

* Total\_Successful\_Ride\_Value:

Provides insight into total revenue from completed rides.

**6. Business Recommendations**

# Based on the analysis, the following recommendations are proposed:

* **Reduce Cancellations:**

Implement targeted interventions for frequent customer and driver cancellation reasons through incentives, in-app messaging, and route optimization.

* **Loyalty Programs:**

Engage top customers with reward schemes to increase repeat rides and lifetime value.

* **Payment Innovations:**

Expand and further promote UPI/online payments for operational efficiency and reduced cashless handling.

* **Driver Training:**

Focus on driver behavior and support to uplift ratings and minimize "personal/car related" cancellations.

* **Dynamic Fleet Management:**

Allocate more Prime Sedans and Minis during high-demand periods to boost utilization and revenue.

**7. Conclusion**

# This analytical dashboard provides OLA's business and operations teams with real-time intelligence on ride patterns, revenue streams, cancellation pain points, and customer/driver experiences. The modular, SQL-driven methodology ensures scalability for ongoing monitoring and supports continuous business improvement.

# 

# **OLA Cabs Data Analysis Project Report**

Duration Analyzed: 1 July 2024 – 31 July 2024  
Rows Processed: 100,000+  
Analyst: Aakash Vishwakarma

## **1. Project Objective**

The primary goal of this project was to deliver actionable business insights for OLA Cabs by thoroughly analyzing one month of large-scale ride booking data. By examining key operational metrics, identifying booking and cancellation patterns, and segmenting performance by vehicle type, the analysis aims to support improved business decision-making, operational optimization, and customer satisfaction.

## **2. Data Description**

The dataset covered over 100,000 ride booking records from July 2024. Each record encapsulated variables such as:

* Booking Information: Booking ID, date/time, status (success, cancelled by driver/customer, driver not found)
* Customer & Driver Details: Ratings, IDs
* Vehicle Type: Auto, Bike, eBike, Mini, Prime Plus, Prime Sedan, Prime SUV
* Ride Metrics: Distance, ratings, revenue, payment method
* Cancellation Reasons (segregated for customers/drivers)

## **3. Data Processing & Methodology**

* Data Extraction: Imported raw data into SQL, ensuring data integrity and cleaning missing/invalid entries.
* Aggregation & Transformation: Developed multiple aggregate SQL views for granular insight (e.g., ride distance per vehicle, cancellation rates, rating distributions).
* KPI Identification: Focused on the most impactful business KPIs: booking volume, total booking value, ride success/cancellation, payment preferences, customer/driver sentiment.
* Dashboard Development: Appended analyses and visualizations to a unified dashboard reflecting both temporal trends and categorical performance breakdowns.

## **4. Results & Key Insights**

## **A. Ride Volume & Revenue**

* Total Bookings: 103,024
* Total Booking Value: 35 million INR
* Success Rate: ~62% rides completed (Success)
* Cancellation Rate: ~28% (Cancelled by Customer/Driver/Not Found)
* Vehicle Type Performance: Prime Sedan and Mini had the highest volumes and revenue generation.

## **B. Ride Distance & Customer Engagement**

* Average Ride Distance: ~15.53 km per ride
* Total Ride Distance: 227,746 km (Prime Sedan led vehicle-wise distance)
* Top 5 customers by booking value: Accounted for a substantial portion of rides and revenue.

## **C. Cancellations – Trends & Root Causes**

* Predominant reasons (Customer): “Change of plans”, “Driver is not moving”, “Driver asked to cancel”
* Predominant reasons (Driver): “Personal & Car related issue”, "Customer related issue"
* Highest Cancellations: Observed in eBikes and Autos.

## **D. Payment Preferences**

* UPI: Most popular, driving 14.2M INR of booking value
* Other Methods: Credit/Debit Card, Cash, OM accounted for a smaller share

## **E. Ratings & Service Quality**

* Customer Ratings: Averaged at 4.0 across all vehicle types
* Driver Ratings: Displayed moderate variance, max-min analysis performed for Prime Sedan

## **5. SQL Analysis Snapshots**

SQL Views created to automate and scale insights include:

* Successful\_Bookings: Extracted all successful bookings
* Ride\_Distance\_For\_Each\_Vehicle: Average ride distance segmented by vehicle type
* Cancelled\_Rides\_By\_Customers: Counted all rides cancelled by customers
* Top\_5\_Customers: Identified highest-frequency riders
* Rides\_Cancelled\_By\_Drivers\_P\_C\_Issues: Focused analysis on driver-initiated cancellations due to personal/car reasons
* Max\_Min\_Driver\_Rating: Found best/worst driver ratings among Prime Sedan bookings
* UPI\_Payment: Segmented all bookings paid via UPI
* Avg\_Cust\_Rating: Averaged customer ratings by vehicle segment
* Total\_Successful\_Ride\_Value: Fast-tracked insight on total revenue from completed rides

## **6. Business Recommendations**

* Reduce Cancellations: Targeted interventions for frequent customer/driver cancellation reasons via incentives, app messaging, and route optimization.
* Loyalty Programs: Engage top customers with reward schemes to increase repeat rides and lifetime value.
* Payment Innovations: Expand and further promote UPI/online payments for operational efficiency and cashless handling.
* Driver Training: Focus on driver behavior and support to uplift ratings and minimize "personal/car related" cancellations.
* Dynamic Fleet Management: Allocate more Prime Sedans and Minis during high-demand slots to boost utilization and revenue.

## **7. Conclusion**

This analytical dashboard equips OLA’s business and operations teams with real-time intelligence on ride patterns, revenue streams, cancellation pain points, and customer/driver experiences. The modular, SQL-driven methodology ensures scalability for ongoing monitoring and supports continuous business improvement.

Appendix:

* [Dashboard Visuals – Bookings Trend, Revenue by Payment, Vehicle Performance, Top Customers]
* [SQL Script: Complete set of analytical queries/views]