Data Analysis

By Vaibhav Bari

🚖 Problem Statement

Ola generates vast amounts of ride-hailing data that can provide insights into customer behavior, ride demand, driver performance, and revenue trends. This analysis aims to improve operational efficiency, enhance customer satisfaction, and optimize pricing and driver allocation.

🎯 Objectives

1. **Analyze Ride Demand Trends** – Identify peak booking hours, popular routes, and seasonal trends.
2. **Customer Behavior Analysis** – Understand ride preferences, payment modes, and top customers.
3. **Driver Performance Evaluation** – Assess ratings, cancellations, and operational efficiency.
4. **Revenue Analysis** – Calculate total earnings, fare distribution, and surge pricing impact.
5. **Operational Efficiency** – Identify reasons for ride cancellations and optimize ride allocation.

By Vaibhav Bari

📊 Dataset Overview

The dataset consists of **100,000 ride records** from Ola, capturing details about ride bookings, customer and driver behavior, payment methods, and performance metrics.

🔹 Key Features of the Dataset

🚖 Booking Details

- Booking ID, Booking Status (Completed, Cancelled, Incomplete)
- Date & Time of the ride

👥 Customer & Ride Information

- Customer ID, Vehicle Type (Mini, Sedan, Prime, etc.)
- Pickup Location, Drop Location
- **Ride Distance**

📈 Performance Metrics

- Vehicle Turnaround Time (V_TAT) and Customer Turnaround Time (C_TAT)

❌ Cancellations & Incomplete Rides

- Cancelled Rides by Customer & Driver
- Incomplete Rides & Reasons

💰 Financial Data

- Booking Value (Total Ride Fare)
- Payment Methods (UPI, Cash, Card, etc.)

⭐ Ratings & Feedback

- Driver Ratings, Customer Ratings

By Vaibhav Bari

SQL Queries

OLA

Create Table

```sql
CREATE TABLE Ola_Rides (
    Date DATE,
    Time TIME,
    Booking_ID VARCHAR(50) PRIMARY KEY,
    Booking_Status VARCHAR(20),
    Customer_ID VARCHAR(50),
    Vehicle_Type VARCHAR(50),
    Pickup_Location VARCHAR(100),
    Drop_Location VARCHAR(100),
    V_TAT FLOAT,
    C_TAT FLOAT,
    Canceled_Rides_by_Customer TEXT,
    Canceled_Rides_by_Driver TEXT,
    Incomplete_Rides TEXT,
    Incomplete_Rides_Reason TEXT,
    Booking_Value DECIMAL(10,2),
    Payment_Method VARCHAR(50),
    Ride_Distance FLOAT,
    Driver_Ratings FLOAT,
    Customer_Rating FLOAT,
    Vehicle_Images TEXT
);
```

By Vaibhav Bari

OLA

1. Retrieve all successful bookings

```
Create View Successful_Bookings As
SELECT * FROM Ola_rides
WHERE Booking_Status = 'Success';

Select * From Successful_Bookings;
```

	date date	time time without time zone	booking_id character varying (50)	booking_status character varying (20)	customer_id character varying (50)
1	2024-07-25	22:20:00	CNR2940424040	Success	CID225428
2	2024-07-30	19:59:00	CNR2982357879	Success	CID270156
3	2024-07-02	09:02:00	CNR1797421769	Success	CID939555
4	2024-07-13	04:42:00	CNR8787177882	Success	CID802429
5	2024-07-23	09:51:00	CNR3612067560	Success	CID476071

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

```
Create View ride_distance_for_each_vehicle As
SELECT Vehicle_Type, AVG(Ride_Distance) as avg_distance
FROM Ola_rides
GROUP BY Vehicle_Type;

Select * from ride_distance_for_each_vehicle;
```

	vehicle_type character varying (50)	avg_distance double precision
1	eBike	15.580588552915767
2	Auto	6.238088783463232
3	Bike	15.533078706861273
4	Prime Sedan	15.764939167842979
5	Prime Plus	15.44747399197661

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

```
Create View Top_5_Customers As
SELECT Customer_ID, COUNT(Booking_ID) as total_rides
FROM Ola_rides
GROUP BY Customer_ID
ORDER BY total_rides DESC LIMIT 5;

SELECT * FROM Top_5_Customers;
```

	customer_id character varying (50)	total_rides bigint
1	CID954071	5
2	CID387617	4
3	CID356460	4
4	CID836942	4
5	CID309168	4

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

```
Create View Rides_cancelled_by_Drivers_P_C_Issues As
SELECT COUNT(*) FROM Ola_rides
WHERE Canceled_Rides_by_Driver = 'Personal & Car related issue';

SELECT * FROM Rides_cancelled_by_Drivers_P_C_Issues;
```

	count bigint
1	6542

By Vaibhav Bari

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

```
Create View Max_Min_Driver_Rating As
SELECT MAX(Driver_Ratings) as max_rating,
MIN(Driver_Ratings) as min_rating
FROM Ola_rides
WHERE Vehicle_Type = 'Prime Sedan';

SELECT * FROM Max_Min_Driver_Rating;
```

	max_rating double precision	min_rating double precision
1	5	3

6. Retrieve all rides where payment was made using UPI

```
Create View UPI_Payment As
SELECT * FROM Ola_rides
WHERE Payment_Method = 'UPI';

SELECT * FROM UPI_Payment;
```

	date date	time time without time zone	booking_id character varying (50)	booking_status character varying (20)
1	2024-07-30	19:59:00	CNR2982357879	Success
2	2024-07-13	04:42:00	CNR8787177882	Success
3	2024-07-27	13:18:00	CNR4524472111	Success
4	2024-07-16	09:54:00	CNR8181602032	Success
5	2024-07-02	10:25:00	CNR8090918544	Success

Total rows: 25881 Query complete 00:00:00.454

By Vaibhav Bari

7. Find the average customer rating per vehicle type

```
Create View AVG_Cust_Rating As
SELECT Vehicle_Type, AVG(Customer_Rating) as avg_customer_rating
FROM Ola_rides
GROUP BY Vehicle_Type;

SELECT * FROM AVG_Cust_Rating;
```

	vehicle_type character varying (50)	avg_customer_rating double precision
1	eBike	3.98785403050109
2	Auto	3.998810952329009
3	Bike	3.993376395883525
4	Prime Sedan	4.001588655506982
5	Prime Plus	4.009498622589555
6	Mini	3.9977312970341075
7	Prime SUV	3.999377501111586

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

```
Create View total_successful_ride_value As
SELECT SUM(Booking_Value) as total_successful_ride_value
FROM Ola_Rides
WHERE Booking_Status = 'Success';

SELECT * FROM total_successful_ride_value;
```

	total_successful_ride_value numeric
1	35080467.00

By Vaibhav Bari

10. List all incomplete rides along with the reason

```
Create View Incomplete_Rides_Reason As
SELECT Booking_ID, Incomplete_Rides_Reason
FROM Ola_Rides
WHERE Incomplete_Rides = 'Yes';

SELECT * FROM Incomplete_Rides_Reason;
```

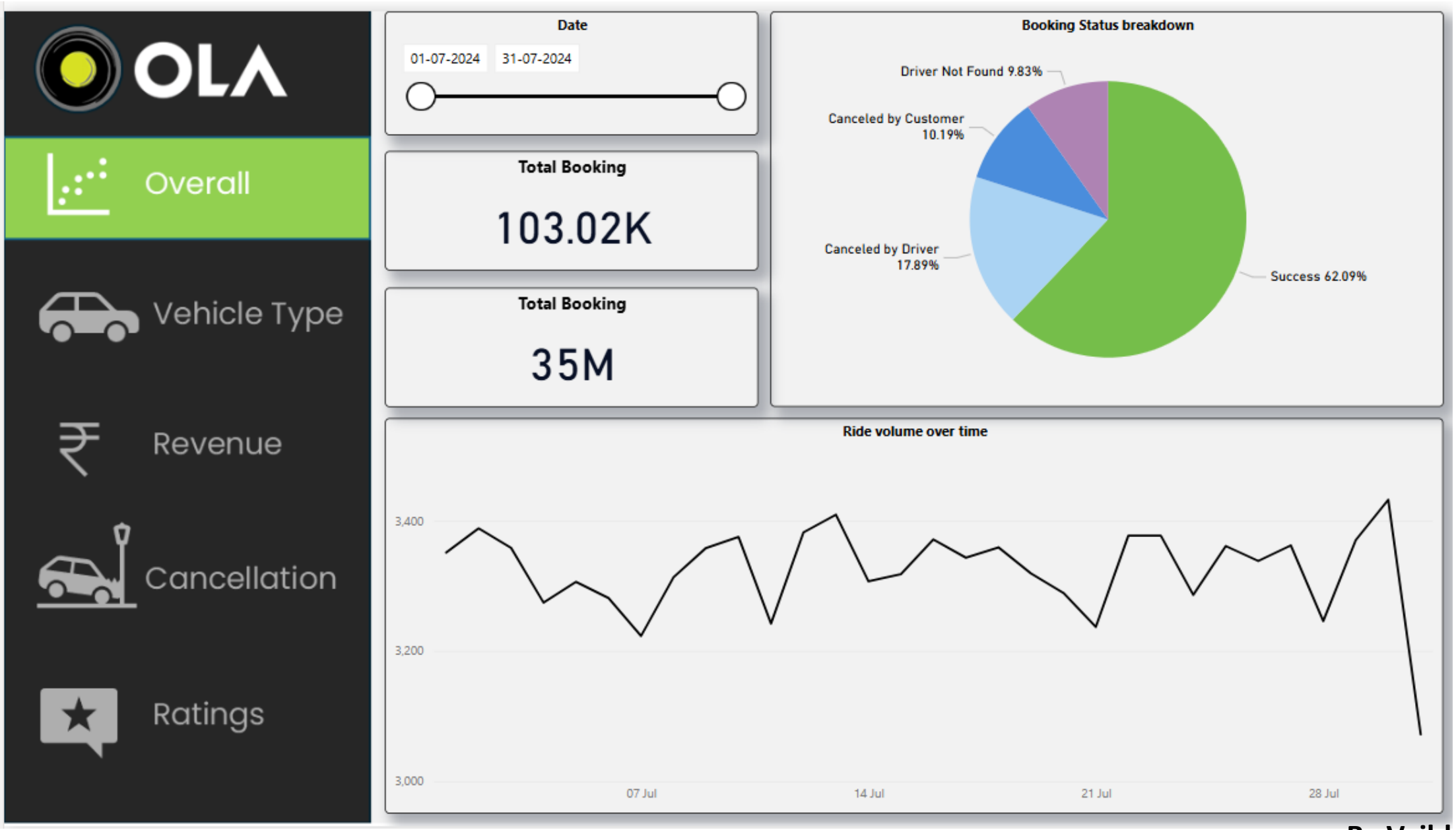
	booking_id character varying (50)	incomplete_rides_reason text
1	CNR5176704322	Customer Demand
2	CNR9312632867	Vehicle Breakdown
3	CNR7924302885	Customer Demand
4	CNR1640228587	Other Issue
5	CNR7623690602	Other Issue
6	CNR9590311980	Customer Demand
7	CNR5863244684	Customer Demand
8	CNR9526078867	Customer Demand

Total rows: 3926 Query complete 00:00:00.192

By Vaibhav Bari

Data Visualization

Overall



By Vaibhav Bari

Vehicle Type

OLA

Overall

Vehicle Type

Revenue

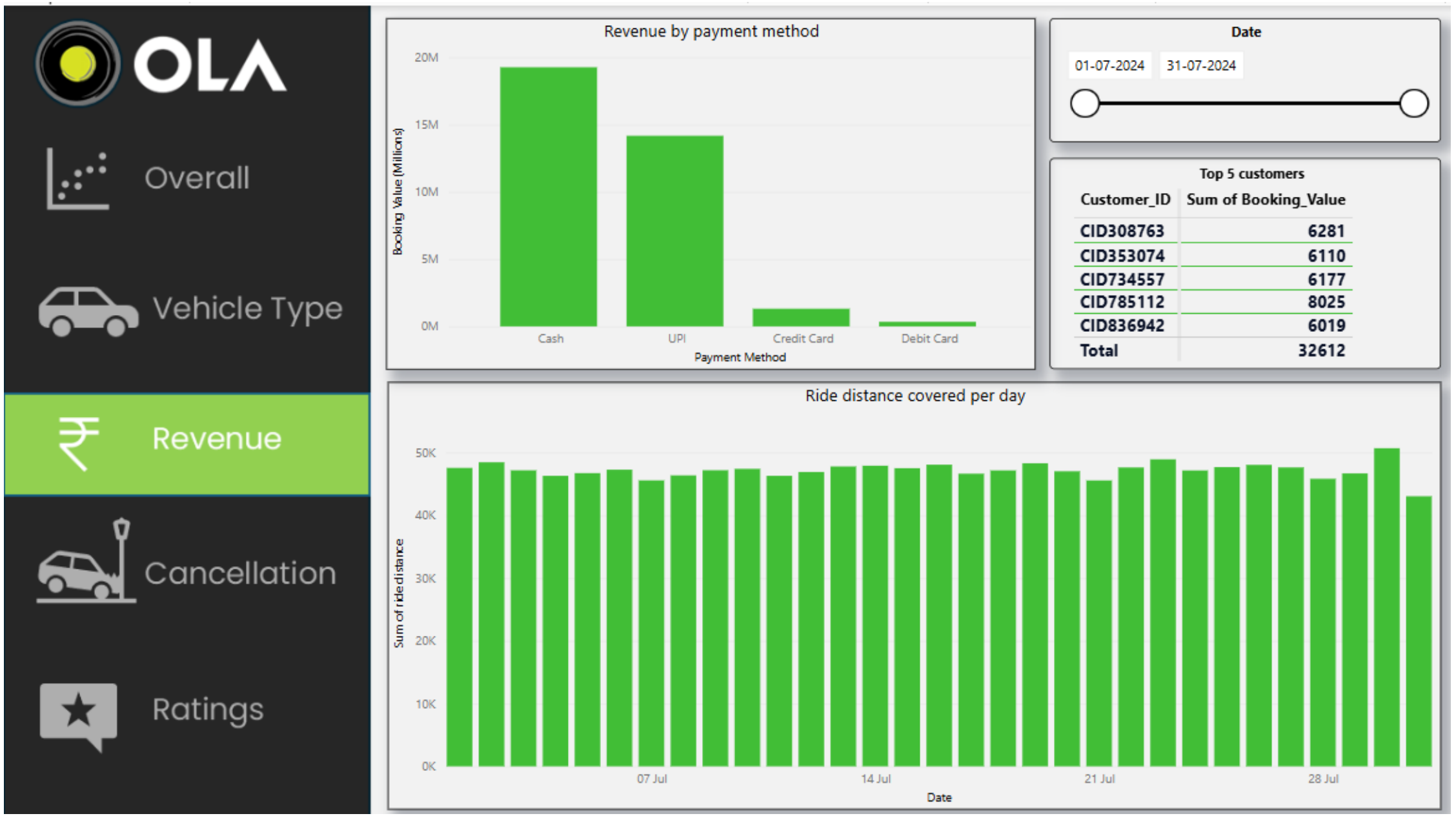
Cancellation

Ratings

Date: 01-07-2024 31-07-2024

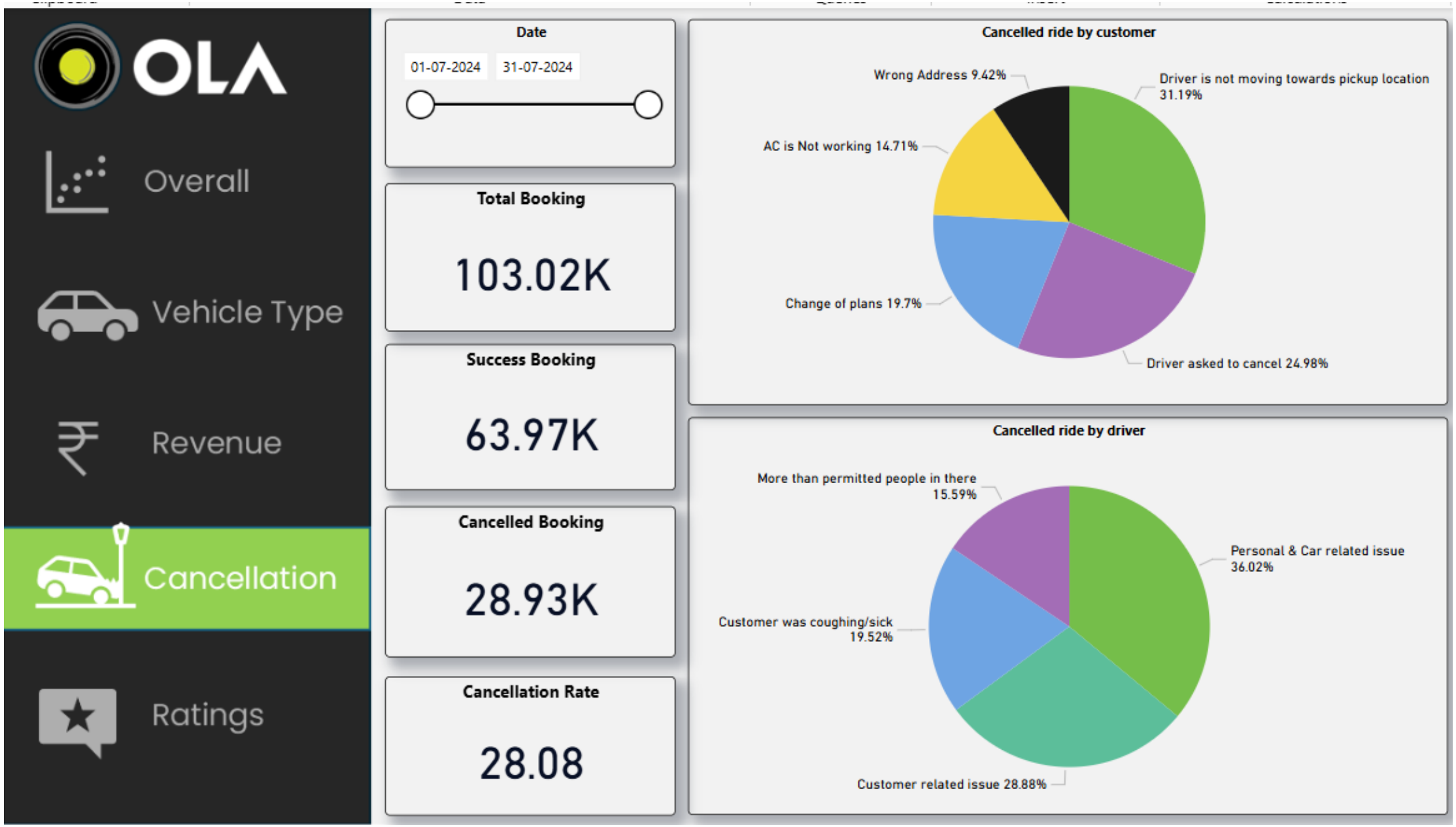
Vehicle Type	Total Booking Value	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
Prime Sedan	8.30M	5.22M	25.01	234.54K
Prime SUV	7.93M	4.88M	24.88	223.85K
Prime Plus	8.05M	5.02M	25.03	227.19K
Mini	7.99M	4.89M	24.98	225.70K
Auto	8.09M	5.05M	10.04	92.04K
Bike	7.99M	4.97M	24.93	227.75K
E-Bike	8.18M	5.05M	25.15	230.84K

Revenue



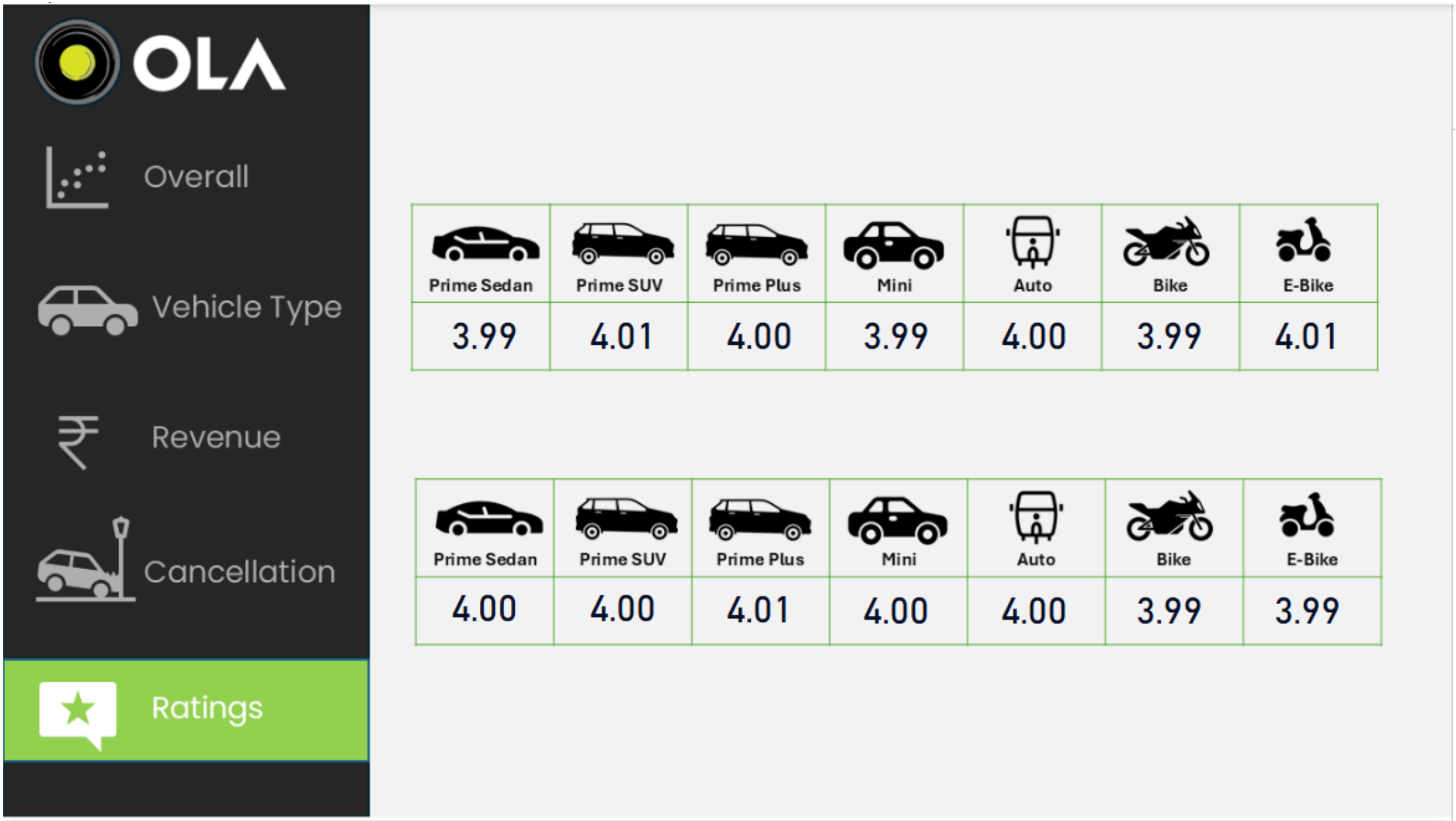
By Vaibhav Bari

Cancellation



By Vaibhav Bari

Ratings



Prime Sedan	Prime SUV	Prime Plus	Mini	Auto	Bike	E-Bike
3.99	4.01	4.00	3.99	4.00	3.99	4.01

Prime Sedan	Prime SUV	Prime Plus	Mini	Auto	Bike	E-Bike
4.00	4.00	4.01	4.00	4.00	3.99	3.99

By Vaibhav Bari

Insights & Actions

1. Booking Status
✓ **Successful Rides:** 73%
✓ **Cancellations:** Driver - 17%, Customer - 10%
🚀 **Action:** Reduce cancellations with driver incentives & penalties.

2. Vehicle Preference
✓ **Top Choice:** Mini & Prime Sedan (50%+ rides)
✓ **Least Used:** eBike & Prime SUV
🚀 **Action:** Increase availability of popular vehicles & promote eBikes.

3. Ride Trends
✓ **Short Rides (0-5 km):** 65%
✓ **Long Rides (10+ km):** 15%
✓ **High Demand Areas:** Metro stations, business hubs
🚀 **Action:** Optimize fleet & introduce short-ride pricing.

4. Payment Methods
✓ **Digital Payments:** 78%
✓ **Cash Transactions:** 22%
🚀 **Action:** Boost digital payments with cashback & ensure app stability.

5. Customer Ratings
✓ **Avg. Rating:** 4.2/5
✓ **Low Ratings (<3 stars):** 12%
🚀 **Action:** Improve driver service & collect feedback.

By Vaibhav Bari