Data Analytics Project

By Vinay.V

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. cancelled 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

Data Columns

1. Date
2. C_TAT
3. Time
4. Cancelled_Rides_by_Customer
5. Booking_ID
6. cancelled_Rides_by_Driver
7. Booking_Status
8. Incomplete_Rides
9. Customer_ID
10. Incomplete_Rides_Reason
11. Vehicle_Type
12. Booking_Value
13. Pickup_Location
14. Payment_Method
15. Drop_Location
16. Ride_Distance
17. V_TAT
18. Driver_Ratings
19. Customer_Rating

SQL Answers:

1. Retrieve all successful bookings:

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

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

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

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

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

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

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

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

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

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

```
MAX(Driver_Ratings) as max_rating, MIN(Driver_Ratings) as min_rating FROM bookings
WHERE Vehicle_Type = 'Prime Sedan';
```

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

```
SELECT * FROM bookings WHERE Payment_Method = 'UPI';
```

8. Find the average customer rating per vehicle type:

```
SELECT Vehicle_Type, AVG(Customer_Rating) as avg_customer_rating FROM bookings
GROUP BY Vehicle_Type;
```

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

```
SELECT SUM(Booking_Value) as total_successful_value FROM bookings WHERE
Booking_Status = 'Success';
```

10. List all incomplete rides along with the reason:

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

Power BI Answers:

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

Answers:

1. **Ride Volume Over Time:** A time-series chart showing the number of rides per day/week.
2. **Booking Status Breakdown:** A pie or doughnut chart displaying the proportion of different booking statuses (success, cancelled by the customer, cancelled by the driver, etc.).
3. **Top 5 Vehicle Types by Ride Distance:** A bar chart ranking vehicle types based on the total distance covered.
4. **Average Customer Ratings by Vehicle Type:** A column chart showing the average customer ratings for different vehicle types.
5. **cancelled Rides Reasons:** A bar chart that highlights the common reasons for ride cancellations by customers and drivers.
6. **Revenue by Payment Method:** A stacked bar chart displaying total revenue based on payment methods (Cash, UPI, Credit Card, etc.).
7. **Top 5 Customers by Total Booking Value:** A leaderboard visual listing customers who have spent the most on bookings.
8. **Ride Distance Distribution Per Day:** A histogram or scatter plot showing the distribution of ride distances for different Dates.
9. **Driver Rating Distribution:** A box plot visualizing the spread of driver ratings for different vehicle types.
10. **Customer vs. Driver Ratings:** A scatter plot comparing customer and driver ratings for each completed ride, analyzing correlations.

SQL Questions & Answers

```
Create Database Ola;
Use Ola;
```

#1. Retrieve all successful bookings:

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

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

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

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

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

#4. 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 bookings
GROUP BY Customer_ID
ORDER BY total_rides DESC LIMIT 5;
```

#5. 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 bookings
WHERE cancelled_Rides_by_Driver = 'Personal & Car related issue';
```

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

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

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

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

#8. Find the average customer rating per vehicle type:

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

#9. 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 bookings
WHERE Booking_Status = 'Success';
```

#10. List all incomplete rides along with the reason:

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

Retrieve All Answers

#1. Retrieve all successful bookings:

```
Select * From Successful_Bookings;
```

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

```
Select * from ride_distance_for_vehicle;
```

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

```
Select * from Cancelled_rides_by_customers;
```

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

```
Select * from Top_5_customers;
```

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

```
Select * from Rides_cancelled_by_drivers_P_C_issues;
```

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

```
Select * from Min_max_driver_ratings_for_Prime_Sedan;
```

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

```
Select * from UPI_Payment;
```

#8. Find the average customer rating per vehicle type:

```
Select * from Avg_customer_rating_per_vehicle;
```

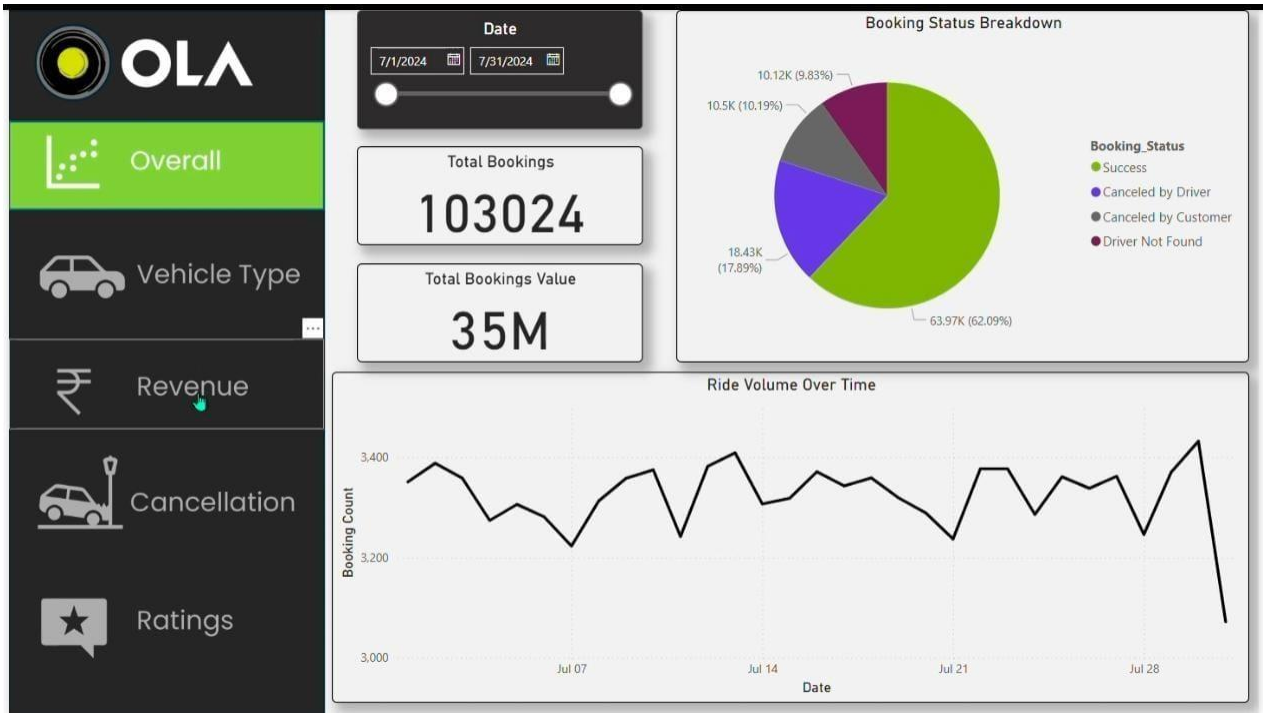
#9. Calculate the total booking value of rides completed successfully:

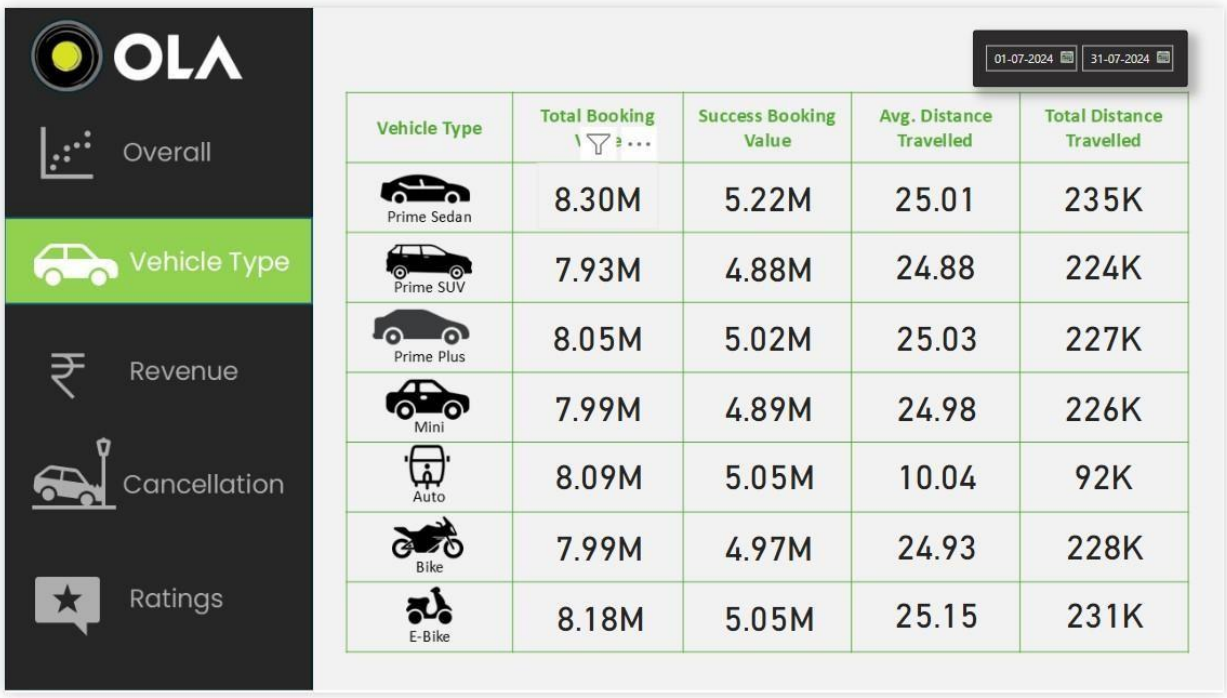
```
Select * from total_successful_ride_value;
```

#10. List all incomplete rides along with the reason:

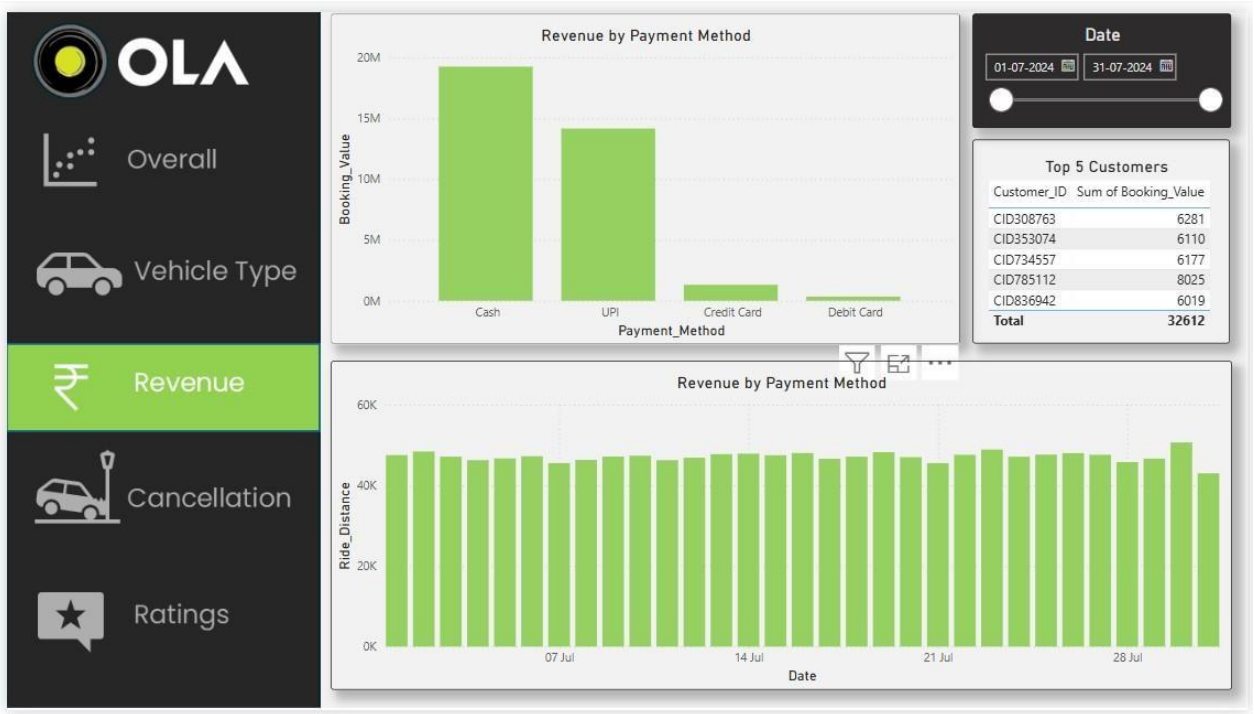
```
Select * from Incomplete_rides_with_reason;
```

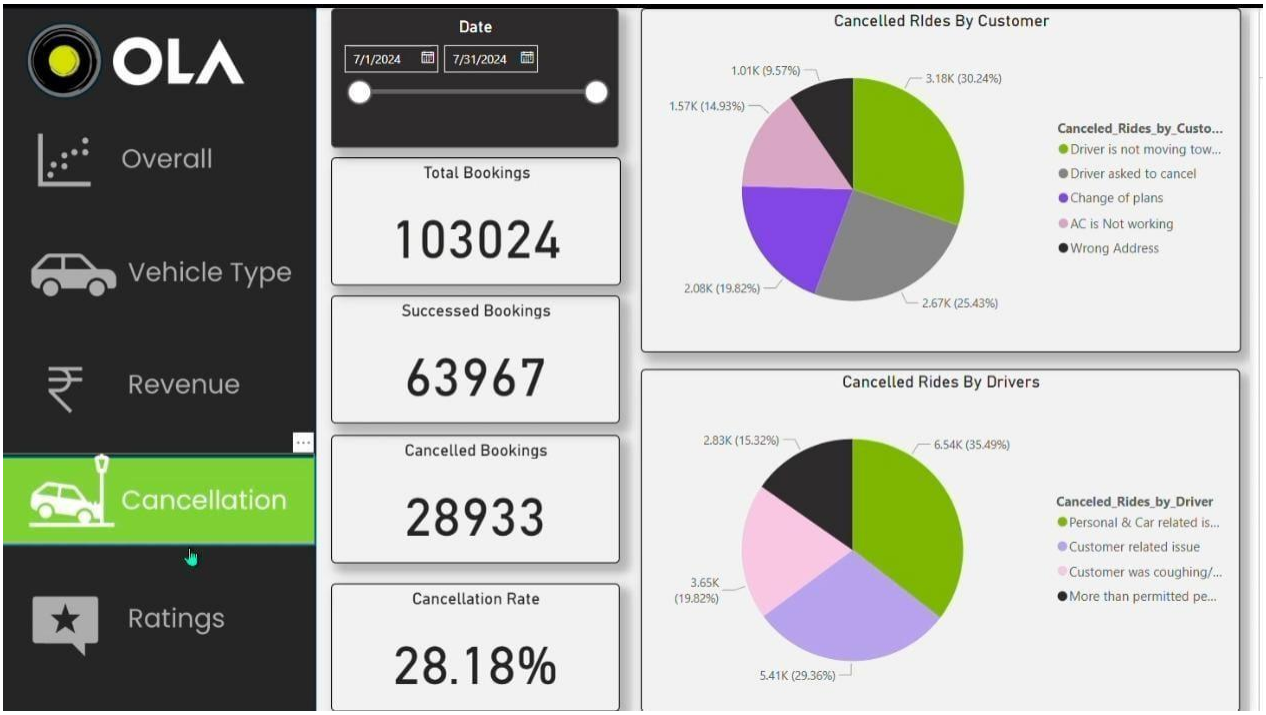
Dashboard Images:

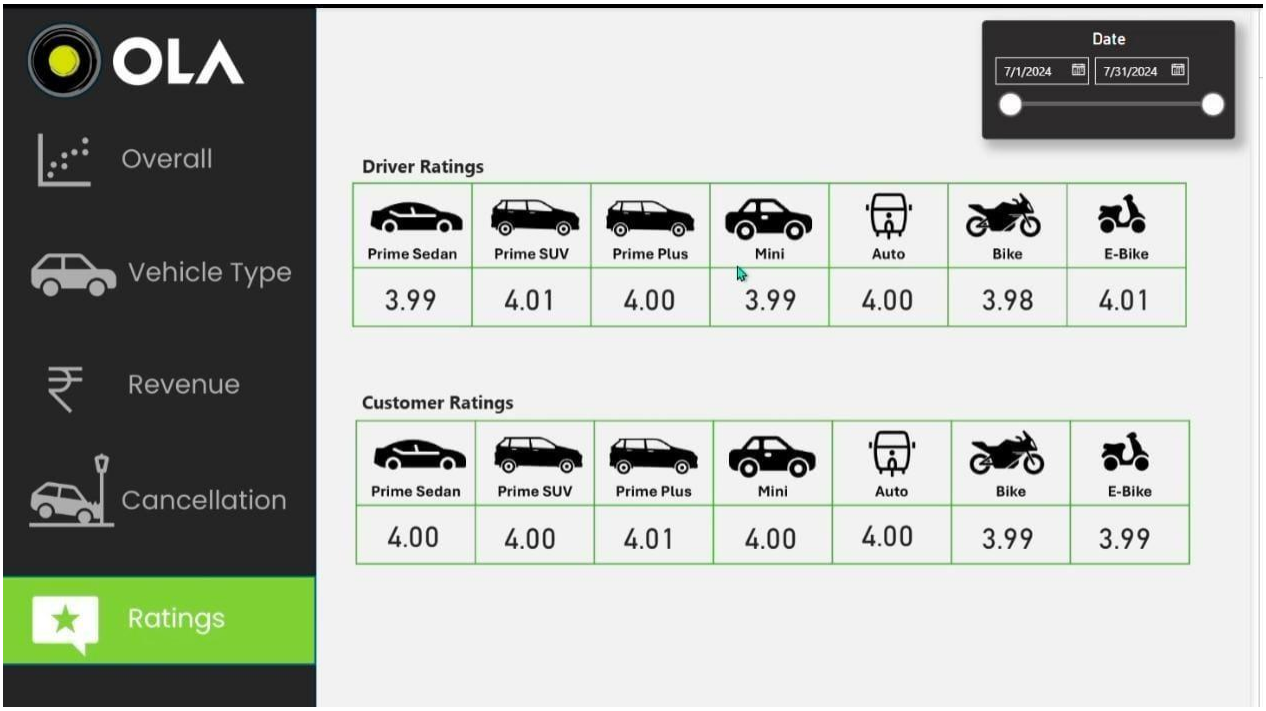




Vehicle Type	Total Booking	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
Prime Sedan	8.30M	5.22M	25.01	235K
Prime SUV	7.93M	4.88M	24.88	224K
Prime Plus	8.05M	5.02M	25.03	227K
Mini	7.99M	4.89M	24.98	226K
Auto	8.09M	5.05M	10.04	92K
Bike	7.99M	4.97M	24.93	228K
E-Bike	8.18M	5.05M	25.15	231K


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
Revenue by Payment Method
20M
15M
10M
5M
0M
Booking_Value
Cash
UPI
Credit Card
Debit Card
Payment_Method
Date
01-07-2024
31-07-2024
Top 5 Customers
Customer_ID Sum of Booking_Value
CID308763 6281
CID353074 6110
CID734557 6177
CID785112 8025
CID836942 6019
Total 32612
Revenue by Payment Method
60K
40K
20K
0K
Ride_Distance
07 Jul
14 Jul
21 Jul
28 Jul
Date


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
Date
7/1/2024
7/31/2024
Total Bookings
103024
Successed Bookings
63967
Cancelled Bookings
28933
Cancellation Rate
28.18%
Cancelled RIdes By Customer
1.01K (9.57%)
3.18K (30.24%)
1.57K (14.93%)
2.08K (19.82%)
2.67K (25.43%)
Canceled_Rides_by_Custo...
Driver is not moving tow...
Driver asked to cancel
Change of plans
AC is Not working
Wrong Address
Cancelled Rides By Drivers
2.83K (15.32%)
6.54K (35.49%)
3.65K (19.82%)
5.41K (29.36%)
Canceled_Rides_by_Driver
Personal & Car related is...
Customer related issue
Customer was coughing/...
More than permitted pe...


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
Date
7/1/2024
7/31/2024
Driver Ratings
Prime Sedan
Prime SUV
Prime Plus
Mini
Auto
Bike
E-Bike
3.99
4.01
4.00
3.99
4.00
3.98
4.01
Customer Ratings
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