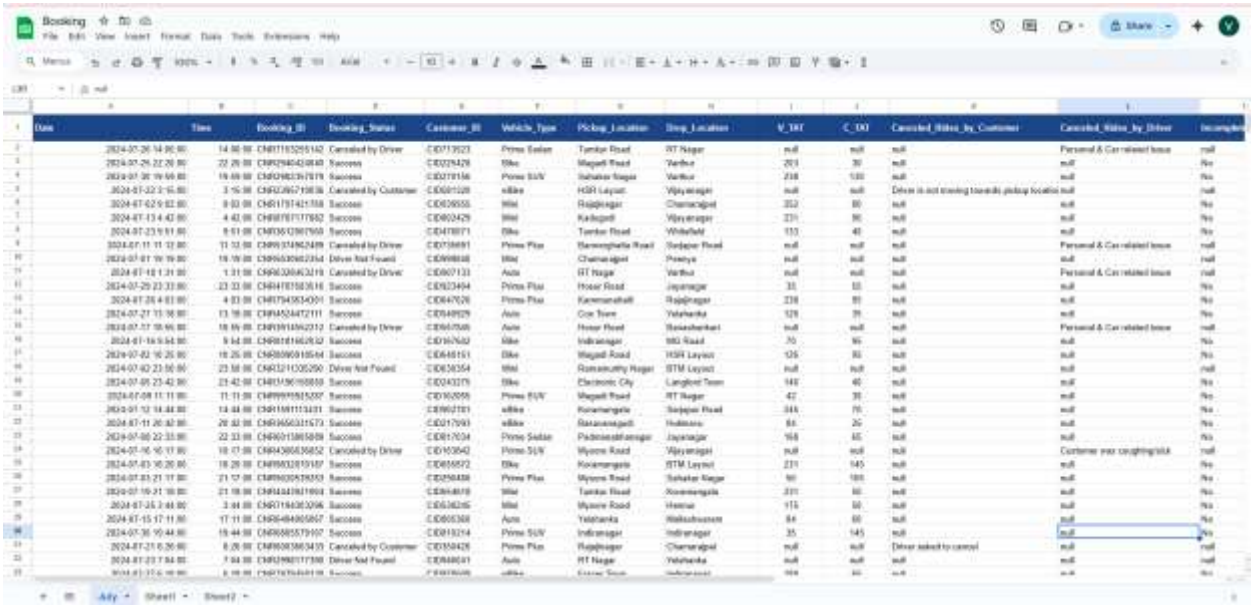
Data Analyst Project

July OLA Dataset

Contain 40,000+ rows



Data Columns

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

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:

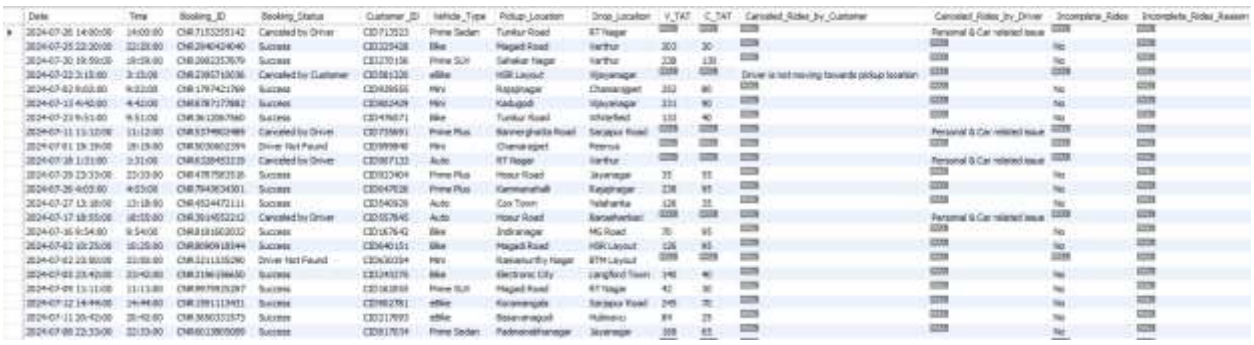
10. List all incomplete rides along with the reason:

Solution with output :

Create database ola;

use ola;

select * from bookings



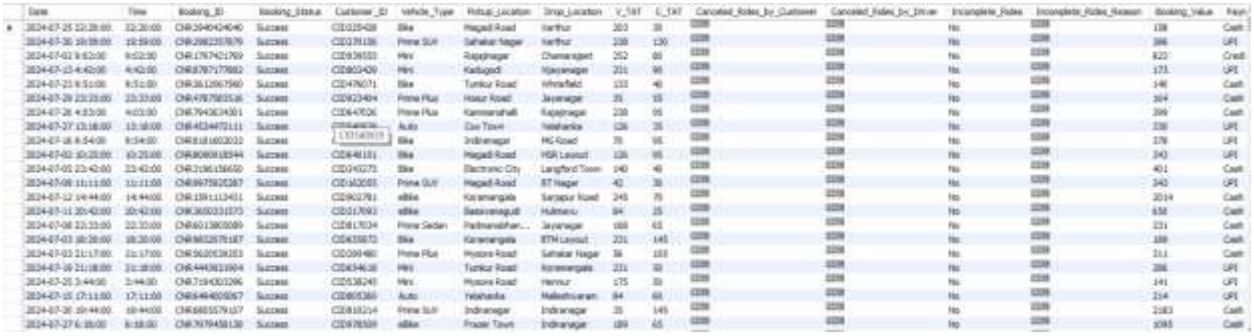
 # 1. Retrieve all successful bookings:

```
create view Successful_Bookings as

select * from bookings

where Booking_status  = 'Success';


select * from Successful_Bookings;
```



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

```
create view avg_ride_distance as

select Vehicle_Type , avg(Ride_Distance) as Average_Distance

from Bookings

group by Vehicle_Type;


Select * from avg_ride_distance;
```

Vehicle_Type	Average_Distance
Prime Sedan	15.5386
Bike	16.1128
Prime SUV	14.7179
eBike	16.2155
Mini	15.6942
Prime Plus	15.0391
Auto	6.3034

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

```
create view Total_canceled_rides_by_customers as
select count(*) as Total_Canceled_rides from Bookings
where Booking_Status = 'Canceled by Customer';

select * from Total_canceled_rides_by_customers;
```

	Total_Canceled_rides
▶	1103

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

```
create view Top5_Customers as
select Customer_ID , count(Booking_ID) as Total_Rides
from bookings
group by Customer_ID
order by Total_Rides desc
limit 5;

select * from Top5_Customers;
```

	Customer_ID	Total_Rides
▶	CID120985	2
	CID806714	2
	CID634243	2
	CID852504	2
	CID991911	2

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

```
create view Total_canceled_rides_by_driver as

select count(*) as Canceled_rides

from Bookings

where Booking_Status = 'Canceled by Driver' and Canceled_Rides_by_Driver = 'Personal & Car related issue';


select * from Total_canceled_rides_by_driver;
```

	Canceled_rides
▶	671

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

```
create view Prime_Sedan_rating as

select max(Driver_Ratings) as Maximum_Rating , min(Driver_ratings) as Minimum_Rating

from Bookings

where vehicle_type = 'Prime Sedan';


select * from Prime_Sedan_rating;
```

	Maximum_Rating	Minimum_Rating
▶	5	3

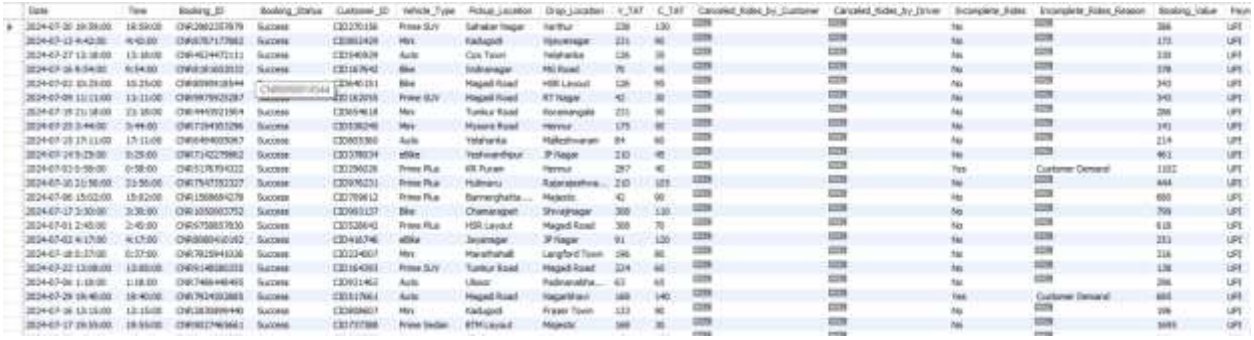
 # 7. Retrieve all rides where payment was made using UPI

```
create view UPI_Rides as

select * from Bookings

where Payment_Method = 'UPI';


select * from UPI_Rides;
```



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

```
create view  Customer_rating as

select Vehicle_Type , avg(Customer_Rating) as Ratings

from Bookings

group by Vehicle_Type;


select * from Customer_rating;
```

	Vehicle_Type	Ratings
▶	Prime Sedan	3.992450879007239
	Bike	3.961749745676499
	Prime SUV	3.9874863982589686
	eBike	3.9949743589743587
	Mini	4.009307135470526
	Prime Plus	4.000455580865606
	Auto	4.027310061601642

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

```
create view Completed_rides as
select sum(Booking_value) as Total_Booking_value
from Bookings
where Booking_Status ='Success';

select * from Completed_rides;
```

	Total_Booking_value
▶	3602017

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

```
create view Incomplete_rides as
Select Booking_ID , Incomplete_Rides_Reason
from Bookings
where Incomplete_Rides = 'Yes';

select * from Incomplete_rides ;
```

Booking_ID	Incomplete_Rides_Reason
CNR5176704322	Customer Demand
CNR9312632867	Vehicle Breakdown
CNR7924302885	Customer Demand
CNR1640228587	Other Issue
CNR7623690602	Other Issue
CNR9590311980	Customer Demand
CNR5863244684	Customer Demand
CNR9526078867	Customer Demand
CNR7154043084	Customer Demand
CNR3193710797	Other Issue
CNR7073850950	Customer Demand
CNR9952584604	Customer Demand
CNR5433575259	Vehicle Breakdown
CNR3575066041	Vehicle Breakdown
CNR7537935962	Customer Demand
CNR2722435581	Vehicle Breakdown
CNR5495479048	Other Issue
CNR6585703917	Customer Demand
CNR3178581771	Customer Demand
CNR7582255360	Vehicle Breakdown
CNR8535011317	Customer Demand
CNR5106611023	Other Issue
CNR7127343811	Vehicle Breakdown

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

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

Explanation:

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.).

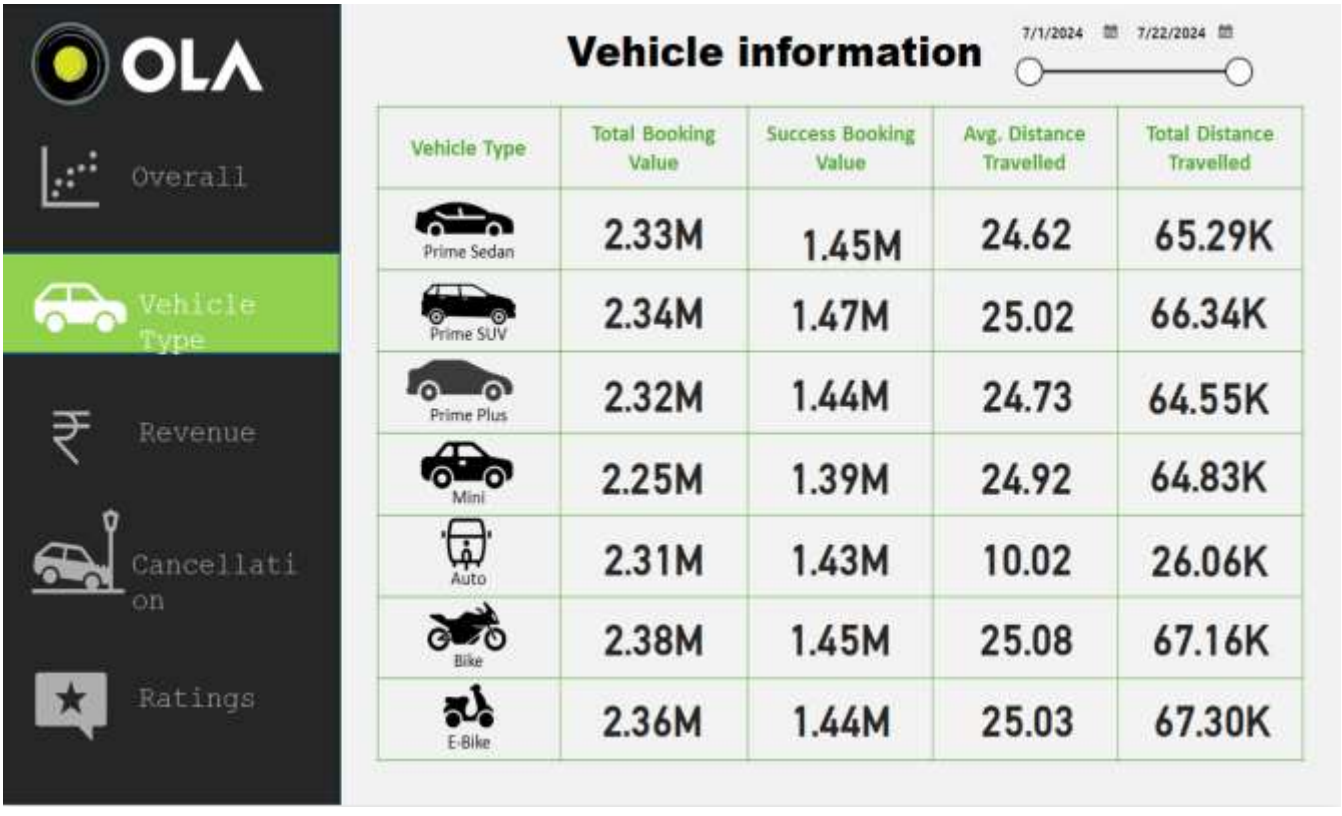
7. Top 5 Customers by Total Booking Value: A leaderboard visual listing customers who have spent the most on bookings.

8. Ride Distance Distribution Per Day: A histogram or scatter plot showing the distribution of ride distances for different Dates.

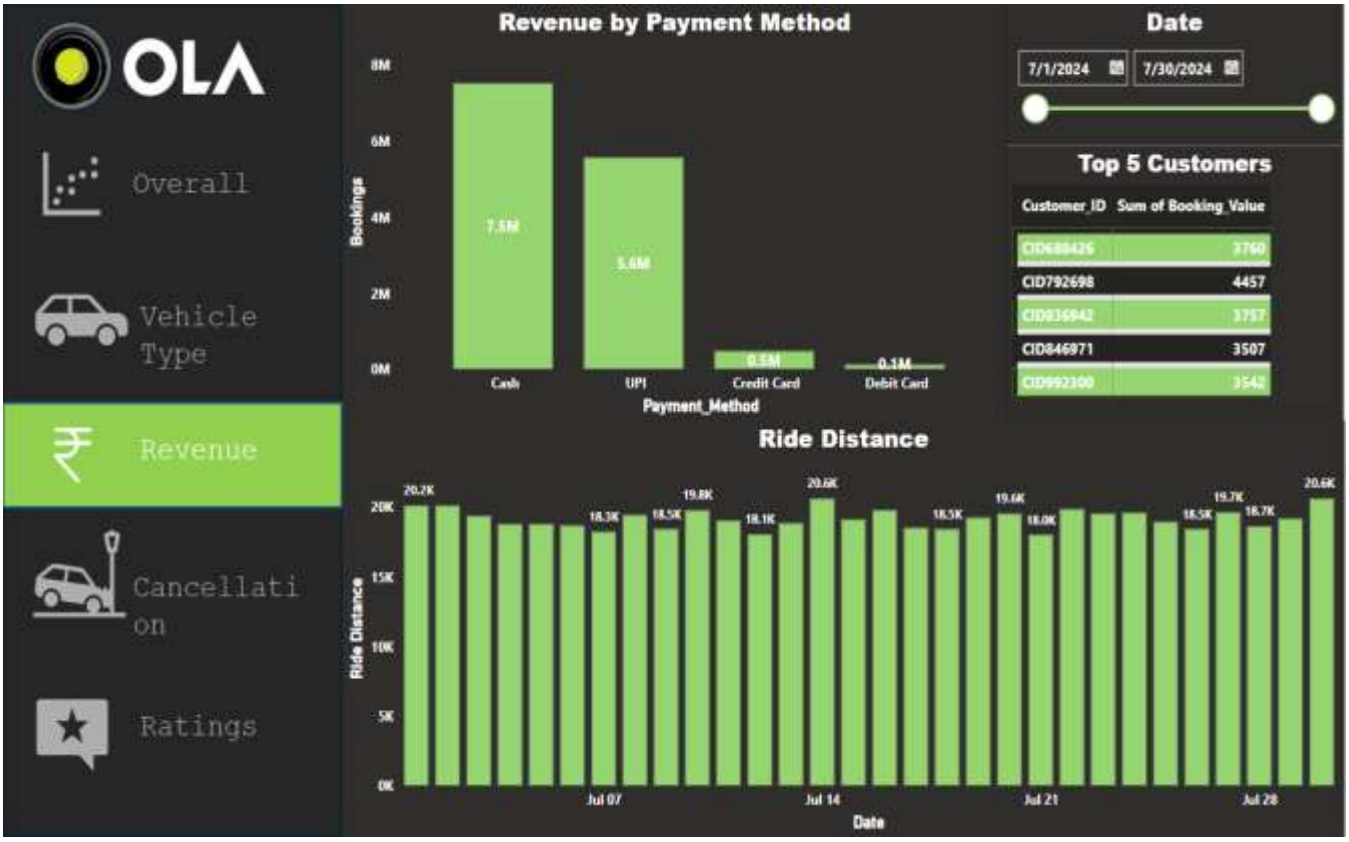
9. Driver Rating Distribution: A box plot visualizing the spread of driver ratings for different vehicle types.

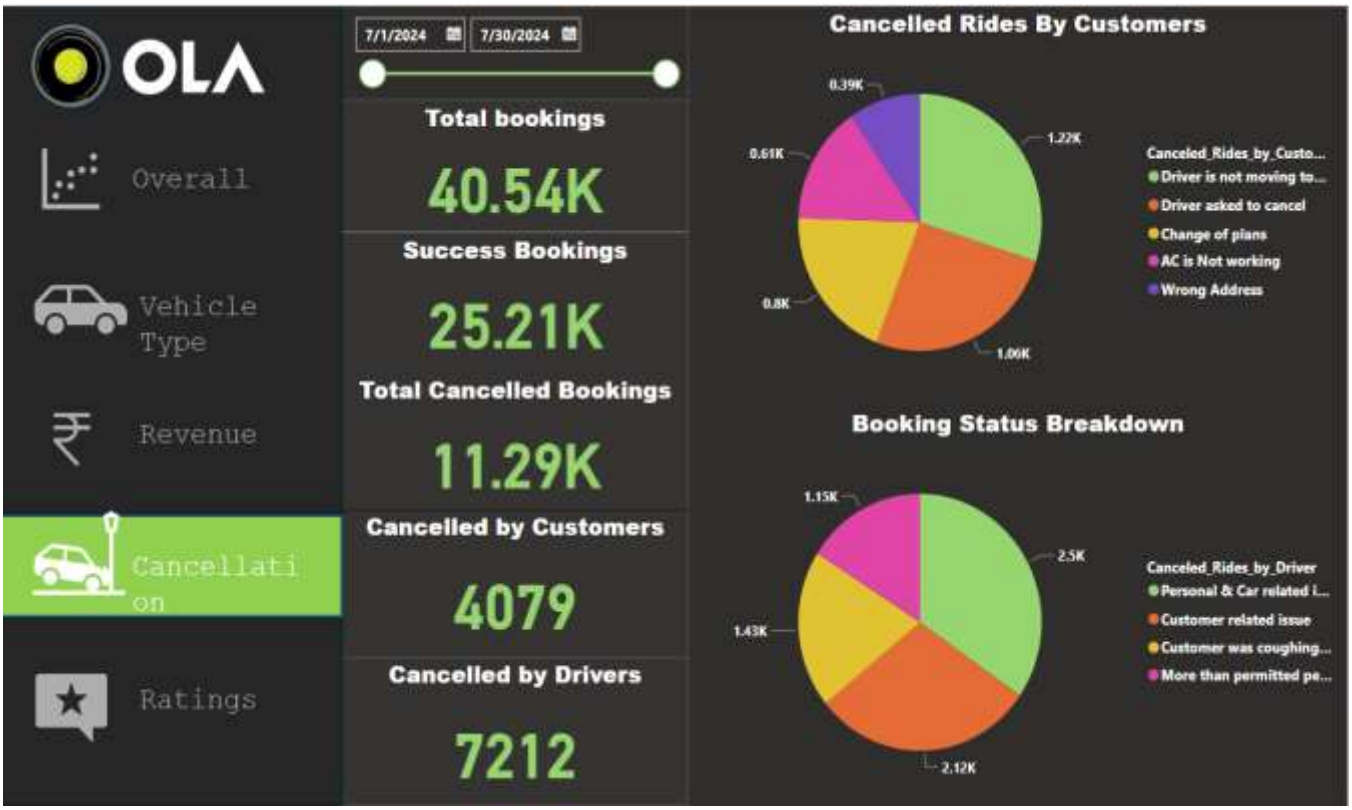
10. Customer vs. Driver Ratings: A scatter plot comparing customer and driver ratings for each completed ride, analyzing correlations

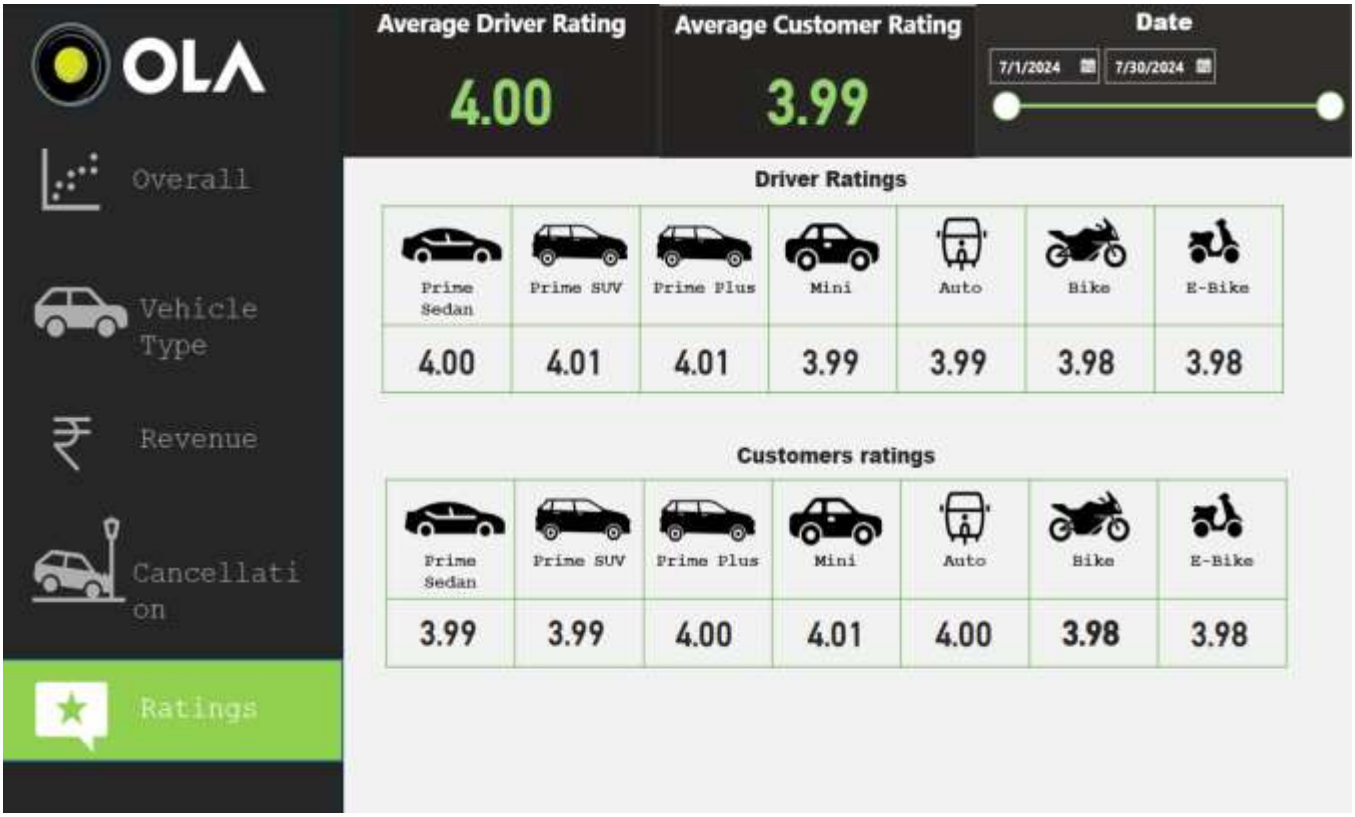
Power BI Report



Vehicle Type	Total Booking Value	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
Prime Sedan	2.33M	1.45M	24.62	65.29K
Prime SUV	2.34M	1.47M	25.02	66.34K
Prime Plus	2.32M	1.44M	24.73	64.55K
Mini	2.25M	1.39M	24.92	64.83K
Auto	2.31M	1.43M	10.02	26.06K
Bike	2.38M	1.45M	25.08	67.16K
E-Bike	2.36M	1.44M	25.03	67.30K


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
Revenue by Payment Method
Bookings
8M
6M
4M
2M
0M
7.5M
5.6M
0.5M
0.1M
Cash
UPI
Credit Card
Debit Card
Payment_Method
Date
7/1/2024
7/30/2024
Top 5 Customers
Customer_ID
Sum of Booking_Value
CID792698
4457
CID846971
3507
Ride Distance
Ride Distance
20K
15K
10K
5K
0K
20.2K
18.3K
18.5K
19.8K
18.1K
20.6K
18.5K
19.6K
18.0K
18.5K
19.7K
18.7K
20.6K
Jul 07
Jul 14
Jul 21
Jul 28
Date


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
7/1/2024
7/30/2024
Total bookings
40.54K
Success Bookings
25.21K
Total Cancelled Bookings
11.29K
Cancelled by Customers
4079
Cancelled by Drivers
7212
Cancelled Rides By Customers
0.39K
1.22K
0.61K
0.8K
1.06K
Canceled_Rides_by_Custo...
Driver is not moving to...
Driver asked to cancel
Change of plans
AC is Not working
Wrong Address
Booking Status Breakdown
1.15K
2.5K
1.43K
2.12K
Canceled_Rides_by_Driver
Personal & Car related i...
Customer related issue
Customer was coughing...
More than permitted pe...


OLA
Overall
Vehicle Type
Revenue
Cancellation
Ratings
Average Driver Rating
4.00
Average Customer Rating
3.99
Date
7/1/2024
7/30/2024
Driver Ratings
Prime Sedan
Prime SUV
Prime Plus
Mini
Auto
Bike
E-Bike
4.00
4.01
4.01
3.99
3.99
3.98
3.98
Customers ratings
Prime Sedan
Prime SUV
Prime Plus
Mini
Auto
Bike
E-Bike
3.99
3.99
4.00
4.01
4.00
3.98
3.98