



Ola Rides Insights Project Report

1. Overview

The ride-sharing industry has revolutionized urban transportation by offering affordable, convenient, and efficient mobility solutions. The **Ola Rides Insights Project** aims to transform raw datasets into meaningful business intelligence by applying data cleaning, exploratory data analysis (EDA), and visualization techniques. A Power BI dashboard and a Streamlit-based web application are developed to present ride trends, customer behaviors, and operational metrics in an interactive manner. This project enables OLA to optimize resource allocation, enhance customer satisfaction, and make data-driven strategic decisions.

2. Problem Statement

By leveraging data analytics, visualization, and interactive applications, the project aims to:

- Identify peak demand hours to optimize driver allocation.
 - Analyze customer behavior to design personalized marketing strategies.
 - Understand pricing and surge patterns to evaluate effectiveness.
 - Detect irregularities or potential fraudulent activities in ride data.
-

3. Dataset Description

Column Name	Description
Date	Date of the ride booking.
Time	Time of the booking request.
Booking_ID	Unique identifier for each ride booking.
Booking_Status	Status of the ride (e.g., Success, Canceled by Customer, Canceled by Driver).
Customer_ID	Unique identifier for each customer.

Vehicle_Type	Type of vehicle booked (e.g., Bike, Mini, Prime Sedan, SUV, eBike).
Pickup_Location	Pickup point location of the ride.
Drop_Location	Drop-off location of the ride.
V_TAT	Vehicle Turnaround Time – time taken by the driver to reach pickup location (in seconds).
C_TAT	Customer Turnaround Time – time waited by customer for the ride (in seconds).
Canceled_Rides_by_Customer	Reason for cancellation if canceled by the customer.
Canceled_Rides_by_Driver	Reason for cancellation if canceled by the driver.
Incomplete_Rides	Whether the ride was incomplete (Yes/No).
Incomplete_Rides_Reason	Reason for incomplete ride (if applicable).
Booking_Value	Fare amount charged for the booking (in INR).
Payment_Method	Mode of payment (e.g., Cash, UPI, Credit Card).
Ride_Distance	Total distance of the ride (in kilometers).
Driver_Ratings	Rating given by the customer to the driver.
Customer_Rating	Rating given by the driver to the customer.

4. Tools & Technologies Used

- **Data Cleaning & Analysis:** Python (Pandas, NumPy, Jupyter Notebook)
 - **Database:** SQLite
 - **Backend:** Python, SQLAlchemy
 - **Frontend & Visualization:** Power BI, Streamlit, Plotly
 - **Deployment:** GitHub, Streamlit Cloud
-

5. Methodology

5.1 Data Understanding & Exploration

- Loaded the Ola Rides dataset into **Google Colab** for initial inspection.
- Examined dataset structure and identified key variables such as **Booking Status**, **Payment Method**, **Customer Rating**, **Vehicle Type**, and **Booking Value**.

5.2 Data Cleaning & Preprocessing

- Handled missing values.
- Removed duplicate entries to ensure data consistency.
- Standardized data formats (e.g., datetime, numeric conversions).

5.3 SQL Query Development (SQLite)

- Stored the cleaned dataset in a **SQLite database (ola_rides.db)** for structured querying.
- Developed optimized SQL queries to analyze ride trends, cancellations, customer ratings, and revenue distribution.
- Validated query outputs against raw dataset statistics to maintain accuracy.

5.4 Power BI Dashboard Creation

- Designed an interactive **Power BI dashboard** to visualize KPIs and trends.
- Integrated visual elements such as:
 - Ride volume trends over time.
 - Cancellation and payment method distribution.
 - Customer ratings by vehicle type.
 - Revenue performance by city.
- Added **filters and slicers** for dynamic user interaction.

5.5 Streamlit Application Development

- Built a **Streamlit app (app.py)** for real-time interaction with the dataset.
- Connected the app to the SQLite database for dynamic querying.
- Added user-friendly filters and visualized results using **Plotly**.
- Embedded Power BI visuals within the app for a unified analytics platform.

5.6 Project Documentation & Deployment

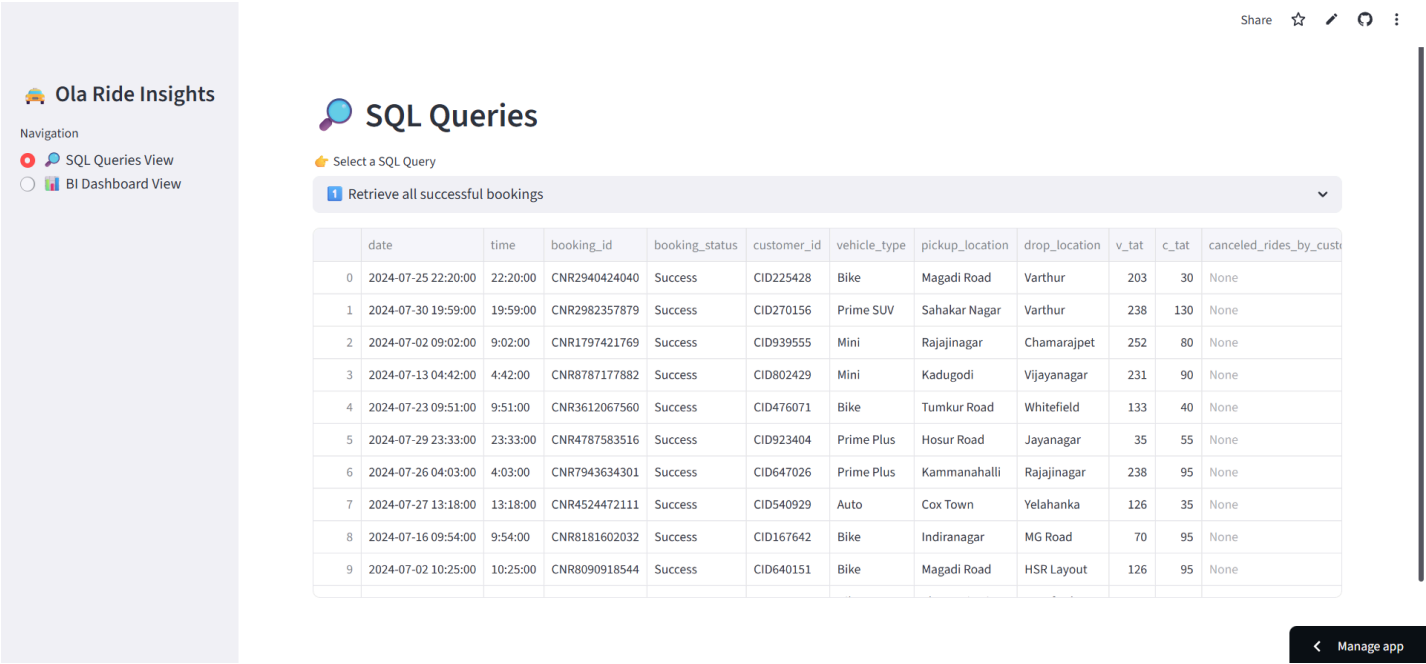
- Documented the entire process including data preparation, SQL queries, and dashboard design.
 - Pushed the project files to **GitHub** for version control and accessibility.
 - Deployed the Streamlit application on **Streamlit Cloud**, ensuring public accessibility for stakeholder
-

6. Ola Ride Insights Streamlit App Interface

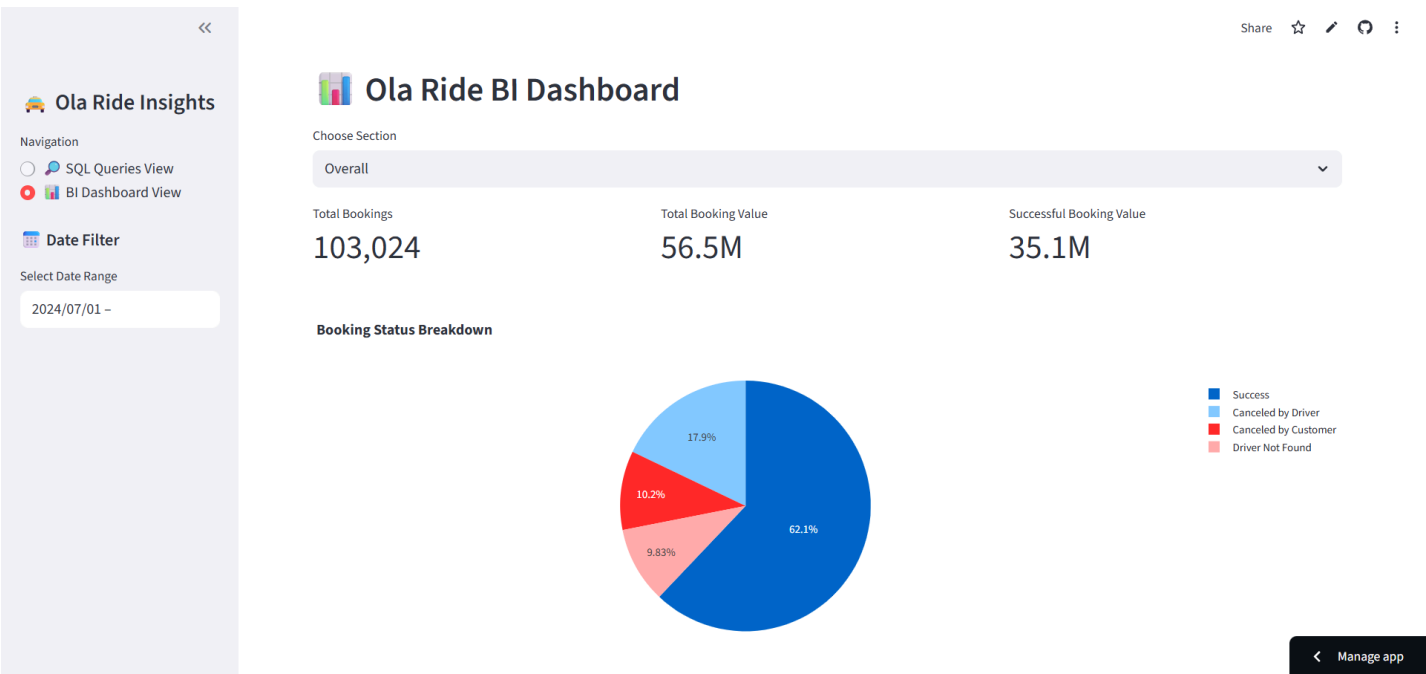
Streamlite App (Ola Ride Insights) consists of 2 Navigations - SQL Queries View and BI Dashboard View with 5 sections - Overall, Vehicle Type, Revenue, Cancellation, and Ratings.

App Link : <https://ola-ride-insights.streamlit.app/>

NAVIGATION 1 (🔍 SQL Queries View)



NAVIGATION 2 (📊 BI Dashboard View)



SECTION 1 (Overall)

Share ☆ ✎ ↺ ⋮

Ola Ride BI Dashboard

Choose Section

Overall

Total Bookings

103,024

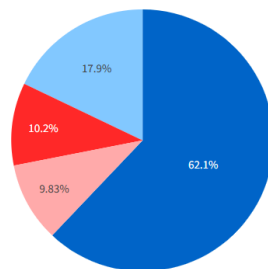
Total Booking Value

56.5M

Successful Booking Value

35.1M

Booking Status Breakdown



■ Success
■ Canceled by Driver
■ Canceled by Customer
■ Driver Not Found

< Manage app

SECTION 2 (Vehicle Type)

Share ☆ ✎ ↺ ⋮

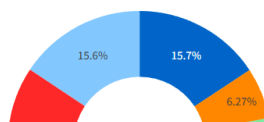
Ola Ride BI Dashboard

Choose Section

Vehicle Type

	Vehicle_Type	total_value	success_value	avg_distance	total_distance
0	Auto	8092210	5051846	10.0407	92043
1	Bike	7987054	4971968	24.9339	227746
2	Mini	7990996	4885961	24.9782	225703
3	Prime Plus	8050889	5015165	25.0343	227186
4	Prime SUV	7933330	4876815	24.8831	223848
5	Prime Sedan	8298612	5224050	25.0064	234535
6	eBike	8181523	5054662	25.1462	230842

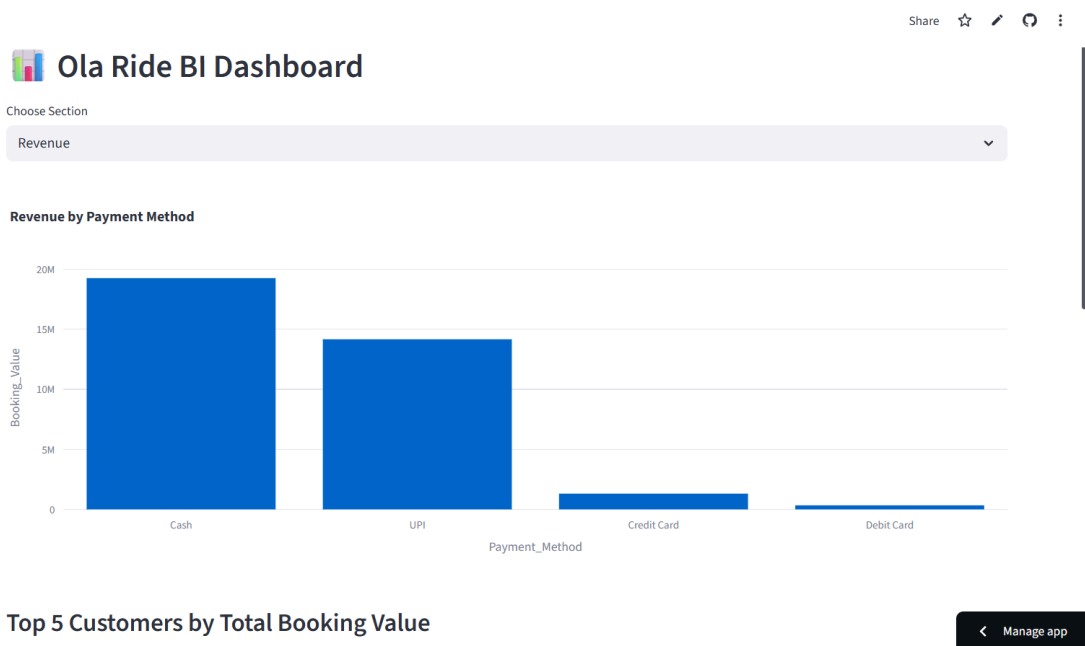
Avg Distance Travelled by Vehicle Type (Success Only)



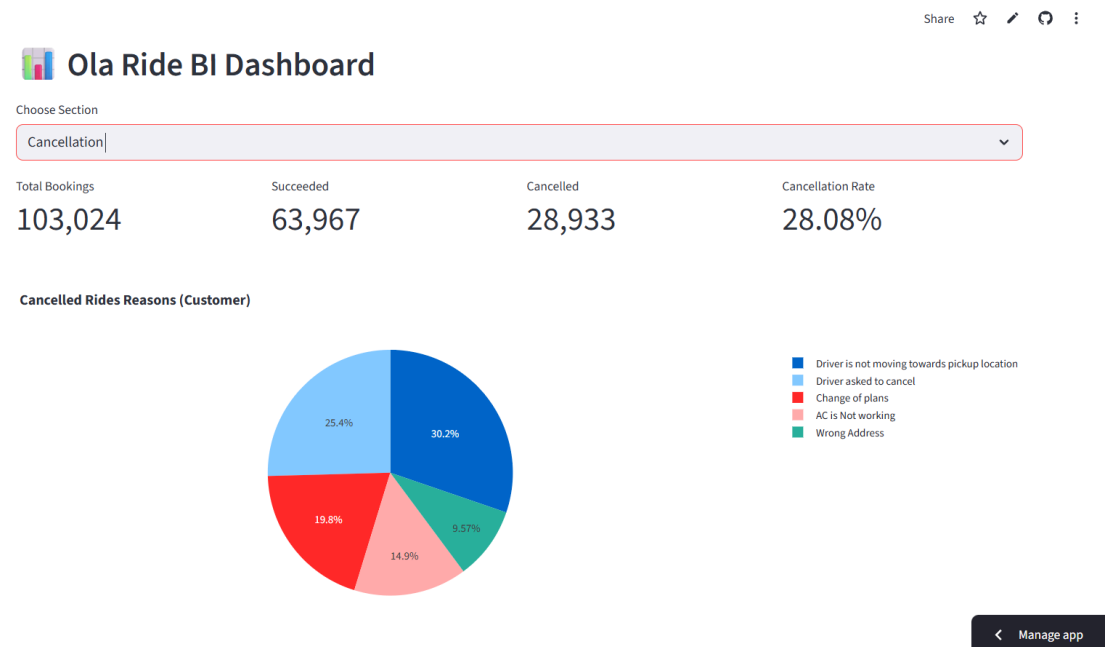
■ eBike
■ Prime Plus
■ Prime Sedan
■ Mini
■ Bike
■ Prime

< Manage app

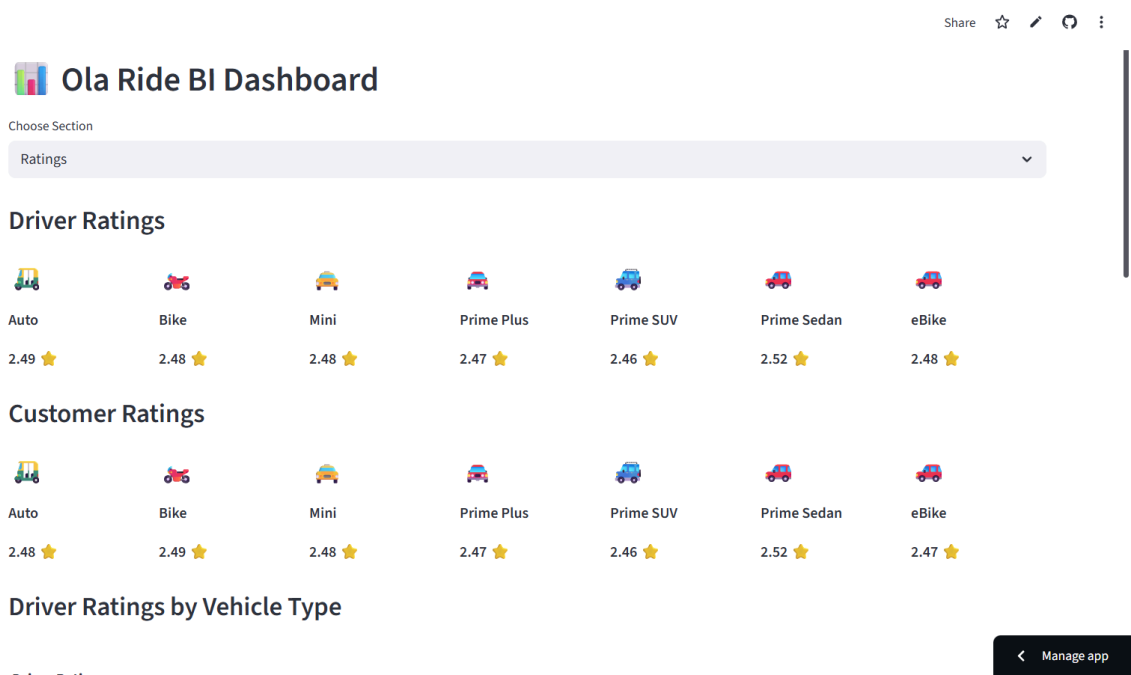
SECTION 3 (Revenue)



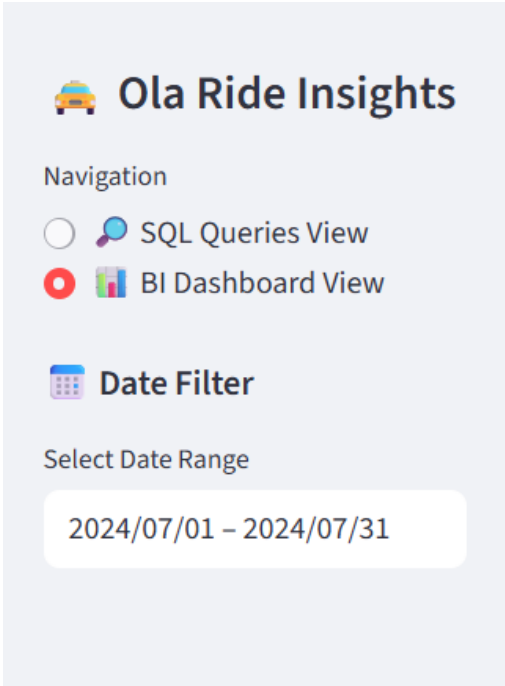
SECTION 4 (Cancellation)



SECTION 5 (Ratings)



NAVIGATIONS & FILTERS



Date Filter

Select Date Range

2024/07/01 – 2024/07/31

7. Analysis Insights

7.1 SQL Queries

I. Retrieve all successful bookings:

1 Retrieve all successful bookings											
	date	time	booking_id	booking_status	customer_id	vehicle_type	pickup_location	drop_location	v_tat	c_tat	canceled_rides_by_cust
0	2024-07-25 22:20:00	22:20:00	CNR2940424040	Success	CID225428	Bike	Magadi Road	Varthur	203	30	None
1	2024-07-30 19:59:00	19:59:00	CNR2982357879	Success	CID270156	Prime SUV	Sahakar Nagar	Varthur	238	130	None
2	2024-07-02 09:02:00	9:02:00	CNR1797421769	Success	CID939555	Mini	Rajajinagar	Chamarajpet	252	80	None
3	2024-07-13 04:42:00	4:42:00	CNR8787177882	Success	CID802429	Mini	Kadugodi	Vijayanagar	231	90	None
4	2024-07-23 09:51:00	9:51:00	CNR3612067560	Success	CID476071	Bike	Tumkur Road	Whitefield	133	40	None
5	2024-07-29 23:33:00	23:33:00	CNR4787583516	Success	CID923404	Prime Plus	Hosur Road	Jayanagar	35	55	None
6	2024-07-26 04:03:00	4:03:00	CNR7943634301	Success	CID647026	Prime Plus	Kammanahalli	Rajajinagar	238	95	None
7	2024-07-27 13:18:00	13:18:00	CNR4524472111	Success	CID540929	Auto	Cox Town	Yelahanka	126	35	None
8	2024-07-16 09:54:00	9:54:00	CNR8181602032	Success	CID167642	Bike	Indiranagar	MG Road	70	95	None
9	2024-07-02 10:25:00	10:25:00	CNR8090918544	Success	CID640151	Bike	Magadi Road	HSR Layout	126	95	None

Insights:

This Table is showcasing all the bookings having Booking status as success.

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

2 Find the average ride distance for each vehicle type		
	vehicle_type	avg_distance
0	Auto	10.04
1	Bike	24.93
2	Mini	24.98
3	Prime Plus	25.03
4	Prime SUV	24.88
5	Prime Sedan	25.01
6	eBike	25.15

Insights:

Based on this Table, it appears that the average ride distance for most vehicle types is consistently around **25 kilometers**. The **Auto**, with a much shorter average ride distance of just over **10 kilometers**.

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

3 Get the total number of cancelled rides by customers

	count(*)
0	10499

Insights:

The total number of rides canceled by customers is **10,499**. This is a direct count of all canceled rides within the dataset.

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

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

	customer_id	Total_Rides
0	CID954071	5
1	CID980727	4
2	CID969725	4
3	CID966929	4
4	CID952434	4

Insights:

The customer with the most bookings is **CID954071**, with a total of **5 rides**. The remaining top customers, **CID980727**, **CID9669725**, **CID9669629**, and **CID952434**, have each booked **4 rides**. This indicates that the most frequent customer booked only one more ride than the next tier of frequent customers.

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

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

	count(*)
0	6542

Insights:

The total number of rides canceled by drivers due to personal and car-related issues are **6,542**. This is a direct count from the dataset, highlighting a significant number of driver-initiated cancellations.

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

6 Find the maximum and minimum driver ratings for Prime Sedan bookings			▼
	Maximum_rating	min_rating	
0		5	0

Insights:

The maximum driver rating for Prime Sedan bookings is **5**, and the minimum is **0**. This indicates the full range of possible ratings from the lowest to the highest.

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

7

Retrieve all rides where payment was made using UPI|

▼

	date	time	booking_id	booking_status	customer_id	vehicle_type	pickup_location	drop_location	v_tat	c_tat	canceled_rides_by_cust
0	2024-07-30 19:59:00	19:59:00	CNR2982357879	Success	CID270156	Prime SUV	Sahakar Nagar	Varthur	238	130	None
1	2024-07-13 04:42:00	4:42:00	CNR8787177882	Success	CID802429	Mini	Kadugodi	Vijayanagar	231	90	None
2	2024-07-27 13:18:00	13:18:00	CNR4524472111	Success	CID540929	Auto	Cox Town	Yelahanka	126	35	None
3	2024-07-16 09:54:00	9:54:00	CNR8181602032	Success	CID167642	Bike	Indiranagar	MG Road	70	95	None
4	2024-07-02 10:25:00	10:25:00	CNR8090918544	Success	CID640151	Bike	Magadi Road	HSR Layout	126	95	None
5	2024-07-09 11:11:00	11:11:00	CNR9975925287	Success	CID162055	Prime SUV	Magadi Road	RT Nagar	42	30	None
6	2024-07-19 21:18:00	21:18:00	CNR4443921904	Success	CID654618	Mini	Tumkur Road	Koramangala	231	50	None
7	2024-07-25 03:44:00	3:44:00	CNR7194303296	Success	CID538245	Mini	Mysore Road	Hennur	175	50	None
8	2024-07-15 17:11:00	17:11:00	CNR6494005067	Success	CID805360	Auto	Yelahanka	Malleshwaram	84	60	None
9	2024-07-14 05:25:00	5:25:00	CNR7142279862	Success	CID378034	eBike	Yeshwanthpur	JP Nagar	210	45	None

Insights:

The displayed rides were successfully completed using UPI as the payment method. The data shows a variety of vehicle types and customers, with pickups and drop-offs at different locations and times throughout July 2024.

VIII. Find the average customer rating per vehicle type:

8 Find the average customer rating per vehicle type|

	vehicle_type	avg_customer_rating
0	Auto	2.4844
1	Bike	2.4878
2	Mini	2.4824
3	Prime Plus	2.4741
4	Prime SUV	2.455
5	Prime Sedan	2.5227
6	eBike	2.4709

Insights:

The average customer ratings are very similar across all vehicle types, clustering around **2.5**. **Prime Sedans** have the highest average rating at **2.5227**, while the **Prime SUV** has the lowest at **2.455**. This suggests there is little variation in customer satisfaction ratings based on the type of vehicle.

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

9 Calculate the total booking value of rides completed successfully		
	total_successful_value	
0		35080467

Insights:

The total booking value for all successfully completed rides is **35,080,467**. This is a direct sum of the revenue generated from all successful bookings in the dataset.

X. List all incomplete rides along with the reason:

10 List all incomplete rides along with the reason		
	incomplete_rides	incomplete_rides_reason
0	Yes	Customer Demand
1	Yes	Vehicle Breakdown
2	Yes	Customer Demand
3	Yes	Other Issue
4	Yes	Other Issue
5	Yes	Customer Demand
6	Yes	Customer Demand
7	Yes	Customer Demand
8	Yes	Customer Demand
9	Yes	Other Issue

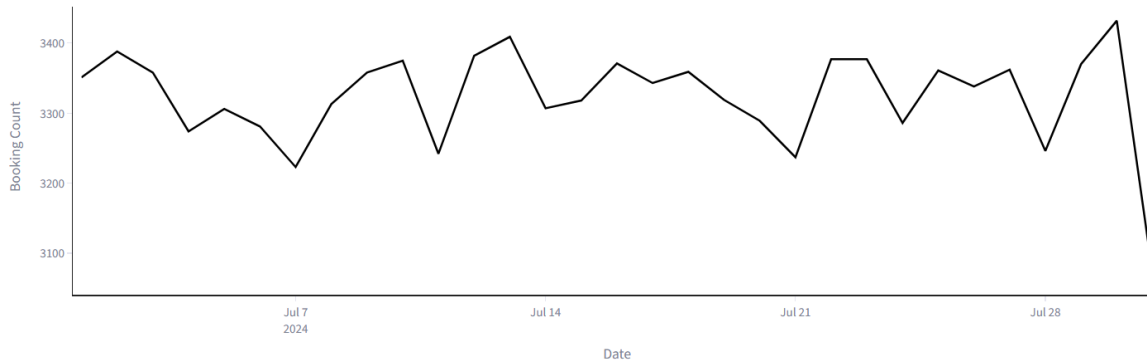
Insights:

The reasons for incomplete rides are primarily attributed to **Customer Demand** and **Other Issues**. There is also few instances of an incomplete ride due to **Vehicle Breakdown**. This suggests that customer-related factors are a major cause of ride cancellations.

7.2 BI Dashboard

I. Ride Volume Over Time:

Ride Volume Over Time



Insights

Based on this line chart, the **booking count** fluctuates daily but generally stays between 3,200 and 3,400 throughout the month. A noticeable **sharp decline** occurred at the very end of the period, around July 30, 2024.

II. Booking Status Breakdown:

Booking Status Breakdown



Insights

Based on this pie chart, a majority of bookings, at 62.1%, are successful. Cancellations account for a significant portion of the unsuccessful bookings, with customers canceling at a rate of 10.2% and drivers canceling at 17.9%. The "Driver Not Found" status represents a smaller but notable portion at 9.83%.

III. Top 5 Vehicle Types by Ride Distance:

	Vehicle_Type	total_value	success_value	avg_distance	total_distance
0	Auto	8092210	5051846	10.0407	92043
1	Bike	7987054	4971968	24.9339	227746
2	Mini	7990996	4885961	24.9782	225703
3	Prime Plus	8050889	5015165	25.0343	227186
4	Prime SUV	7933330	4876815	24.8831	223848
5	Prime Sedan	8298612	5224050	25.0064	234535
6	eBike	8181523	5054662	25.1462	230842

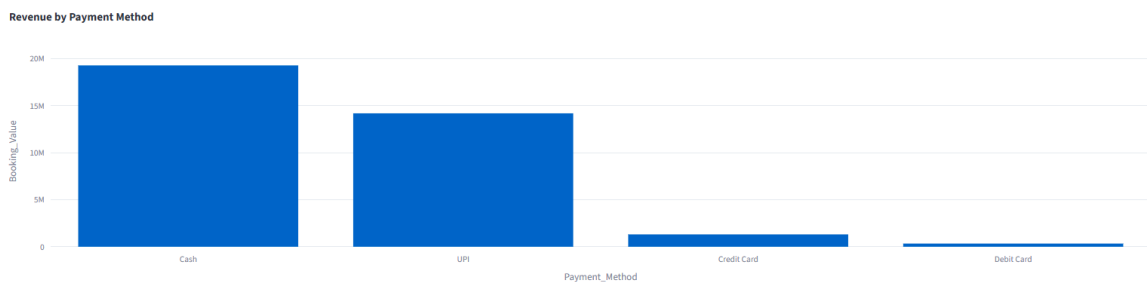
Avg Distance Travelled by Vehicle Type (Success Only)



Insights

Based on this Table, **Auto** rides have the shortest average distance at 10.04 km, while all other vehicle types (Bike, Mini, Prime Plus, Prime SUV, Prime Sedan, eBike) have significantly longer average distances, ranging from 24.93 km to 25.14 km. The donut chart also shows that the total distance for successful rides is fairly evenly distributed across all vehicle types except for **Auto**, which accounts for a very small fraction (0.2%) of the total distance.

IV. Revenue by Payment Method:



Insights

Based on this bar chart, the majority of the booking revenue comes from **Cash** and **UPI** payments, which are the two dominant methods by a large margin. Payments made via **Credit Card** and **Debit Card** contribute a very small portion of the total revenue, each being less than 1/10th of the revenue generated by UPI.

V. Top 5 Customers by Total Booking Value:

Top 5 Customers by Total Booking Value

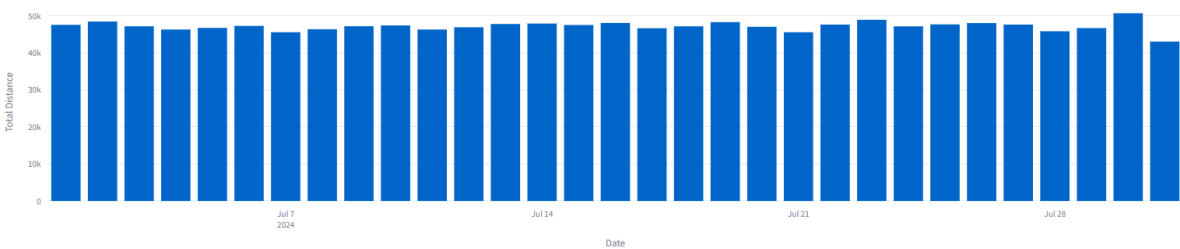
	Customer_ID	Booking_Value
71883	CID785112	8025
21846	CID308763	6281
66511	CID734557	6177
26527	CID353074	6110
77406	CID836942	6019

Insights

Based on the table, the customer with ID **CI0705112** has the highest total booking value at 8,025. The top five customers all have a booking value over 6,000, indicating a small group of high-value customers. The values for the top five customers are relatively close to each other, with a difference of about 2,000 between the first and fifth-ranked customers.

VI. Ride Distance Distribution Per Day:

Ride Distance Distribution Per Day



Insights

Based on the bar chart, the **total distance** traveled each day is quite consistent, generally fluctuating between 45,000 and 50,000 units. There is a notable **increase** around July 29, 2024, followed by a **sharp decrease** at the very end of the period, around July 30, 2024.

VII. Cancelled Rides Reasons (Customer) :

Cancelled Rides Reasons (Customer)



Insights

Based on the pie chart, the most common reason for customer cancellations is the **driver not moving towards the pickup location** (30.2%), followed by the **driver asking the customer to cancel** (25.4%). The least common reasons are **AC not working** (14.9%) and **Wrong Address** (9.57%).

VIII. Cancelled Rides Reasons (Drivers) :

Cancelled Rides Reasons (Driver)

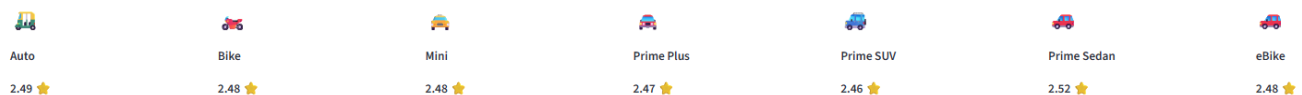


Insights

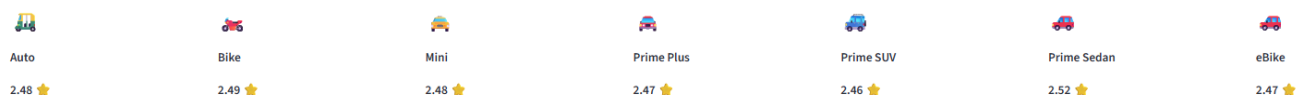
Based on the pie chart, the most common reason for a driver to cancel is a **personal or car-related issue**, accounting for 35.5% of cancellations. The second largest reason is a **customer-related issue**, at 29.4%. Reasons like a customer being sick (19.8%) and more than the permitted number of people (15.3%) make up smaller but significant portions of cancellations.

IX. Customer vs. Driver Ratings:

Driver Ratings



Customer Ratings



Insights

Based on the provided data, the **driver ratings** are very similar across all vehicle types, with all ratings falling between 2.46 and 2.52 out of a possible 5 stars. The **customer ratings**

also show a similar pattern, with most vehicle types having ratings between 2.46 and 2.52. **Prime Sedan** has the highest ratings from both drivers and customers, whereas **Prime SUV** consistently has the lowest ratings from both groups.

8. Results

- Cleaned and preprocessed Ola Rides dataset for accurate analysis.
 - Built **SQLite database** and executed optimized SQL queries.
 - Created an interactive **Power BI dashboard** for ride insights.
 - Developed a **Streamlit app** with SQL view and BI dashboard integration.
 - Deployed app on **Streamlit Cloud** for public accessibility.
-

9. Business Recommendations

1. **Reduce Customer Cancellations** – Implement proactive notifications and track driver punctuality to minimize delays.
 2. **Minimize Driver Cancellations** – Provide flexible shifts, support programs, and backup drivers to ensure ride fulfillment.
 3. **Optimize Fleet Allocation** – Assign vehicles based on ride distances and demand patterns to improve efficiency.
 4. **Upgrade Vehicle Services** – Enhance Prime SUV and Prime Sedan offerings to boost customer satisfaction and ratings.
 5. **Promote Digital Payments** – Encourage UPI and card payments via rewards, reducing reliance on cash.
 6. **Improve Customer & Driver Ratings** – Conduct regular training, monitor low ratings, and address complaints promptly.
 7. **Target High-Value Customers** – Introduce loyalty programs, personalized offers, and premium service options.
 8. **Enhance Operational Efficiency** – Use predictive analytics for driver allocation and manage ride volume fluctuations.
-

10. Conclusion

The Ola Rides Insights Project successfully transformed raw ride data into actionable business intelligence through SQL analysis, Power BI visualizations, and an interactive Streamlit app. Key insights into ride patterns, cancellations, payment methods, and customer-driver ratings were uncovered, enabling data-driven decisions. Recommendations focused on reducing cancellations, optimizing fleet allocation, enhancing payment adoption, improving customer experience, and targeting high-value customers. By leveraging these insights, Ola can increase operational efficiency, boost revenue, and elevate customer and driver satisfaction. The deployed dashboard ensures stakeholders have real-time access to critical performance metrics for strategic planning and continuous improvement.