Project: OLA Analysis 2025

- **About the Dataset:**

The dataset utilized in this project is entirely proprietary and not available from any public sources, as it was custom-generated by me using advanced AI techniques. Consisting of over 100,000 rows and 19 carefully selected columns, the dataset was meticulously designed to reflect real-world operational details from Ola's daily ride-hailing activities. This includes key aspects such as ride bookings, driver assignments, trip durations, pricing dynamics, customer ratings, peak-hour patterns, geographical distributions, and other relevant metrics that mirror the complexities of the industry.

After the synthetic dataset was created, a thorough exploratory data analysis (EDA) was conducted, followed by in-depth statistical and predictive modeling to uncover actionable insights, trends, and patterns that could inform business decisions in a ride-hailing context. This approach ensured the dataset's realism and practical applicability while maintaining complete control over data quality and relevance.

- **SQL Task:**
  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 Task:**
  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 Cleaning Process:**

The dataset for this project was generated using advanced AI techniques, providing a significant advantage in terms of minimal data cleaning requirements. The preprocessing involved was notably lightweight and straightforward, primarily consisting of negligible efforts to handle any minor inconsistencies.

Specifically, the key steps included removing duplicate rows to ensure data uniqueness and appropriately setting column-wise data types—for instance, converting numerical columns to numeric formats and date-related columns to proper datetime types. No extensive handling of missing values, outliers, or complex transformations was necessary, thanks to the high-quality synthetic generation process.

Once these minor preparations were completed, we proceeded directly to the core exploratory data analysis (EDA) and advanced analytical phases, allowing for a seamless and efficient transition into deriving meaningful insights and building predictive models. This streamlined workflow highlights the practical benefits of leveraging AI-generated datasets tailored to real-world scenarios in the ride-hailing domain.

- **Analysis & Insides:**

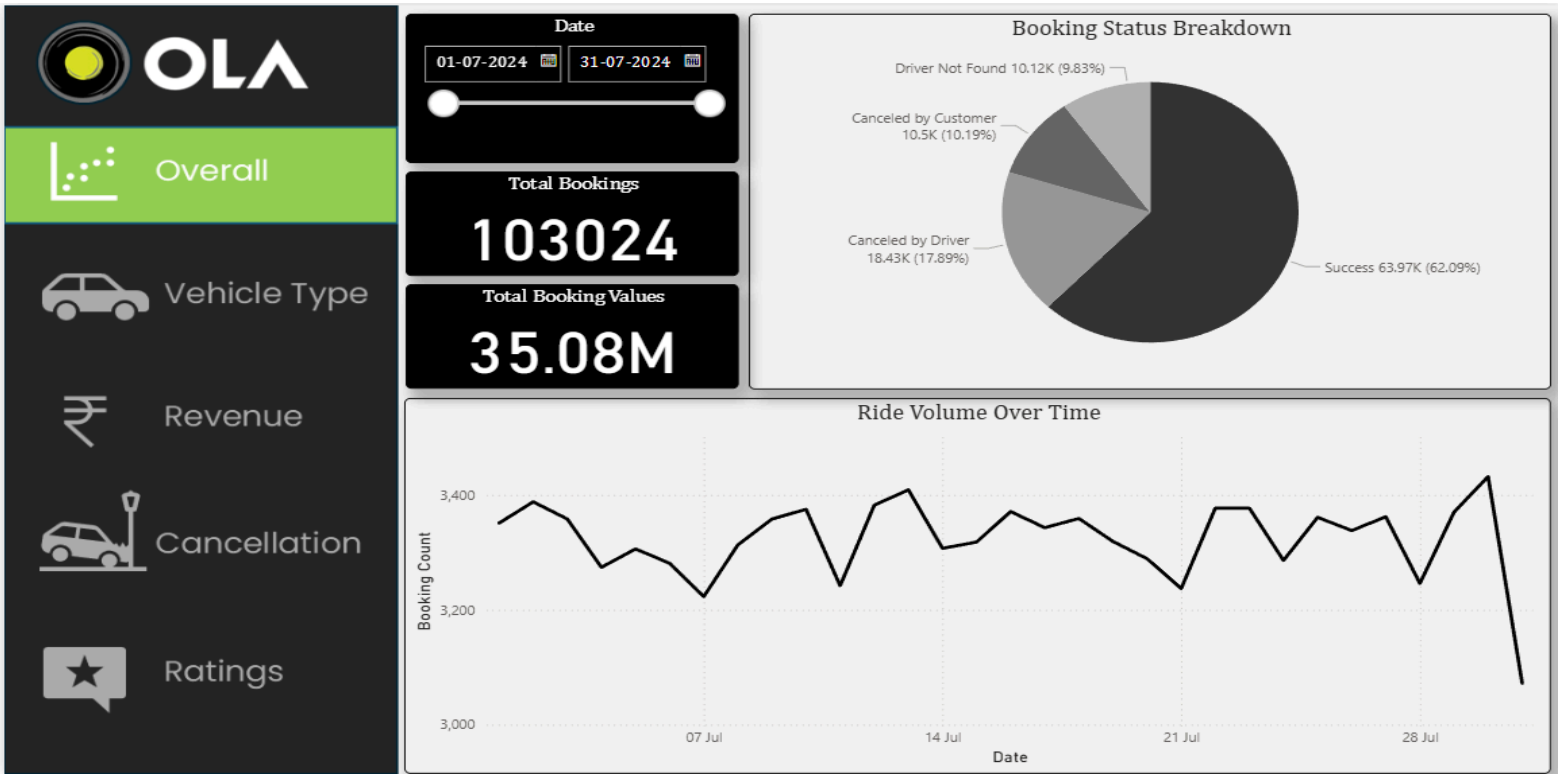
Actually, we are unable to manually type out the insights here, as they were derived directly from MySQL queries for the SQL-related tasks. Therefore, for all insights from these SQL tasks, I will paste the relevant queries below, which you can simply copy and paste as needed. Additionally, to make it easier for managers and stakeholders to review—even those without SQL knowledge—I have created a dedicated view for each task. This setup allows for quick and effortless access, enabling faster workflow and decision-making.

- **Power BI Dashboard:**

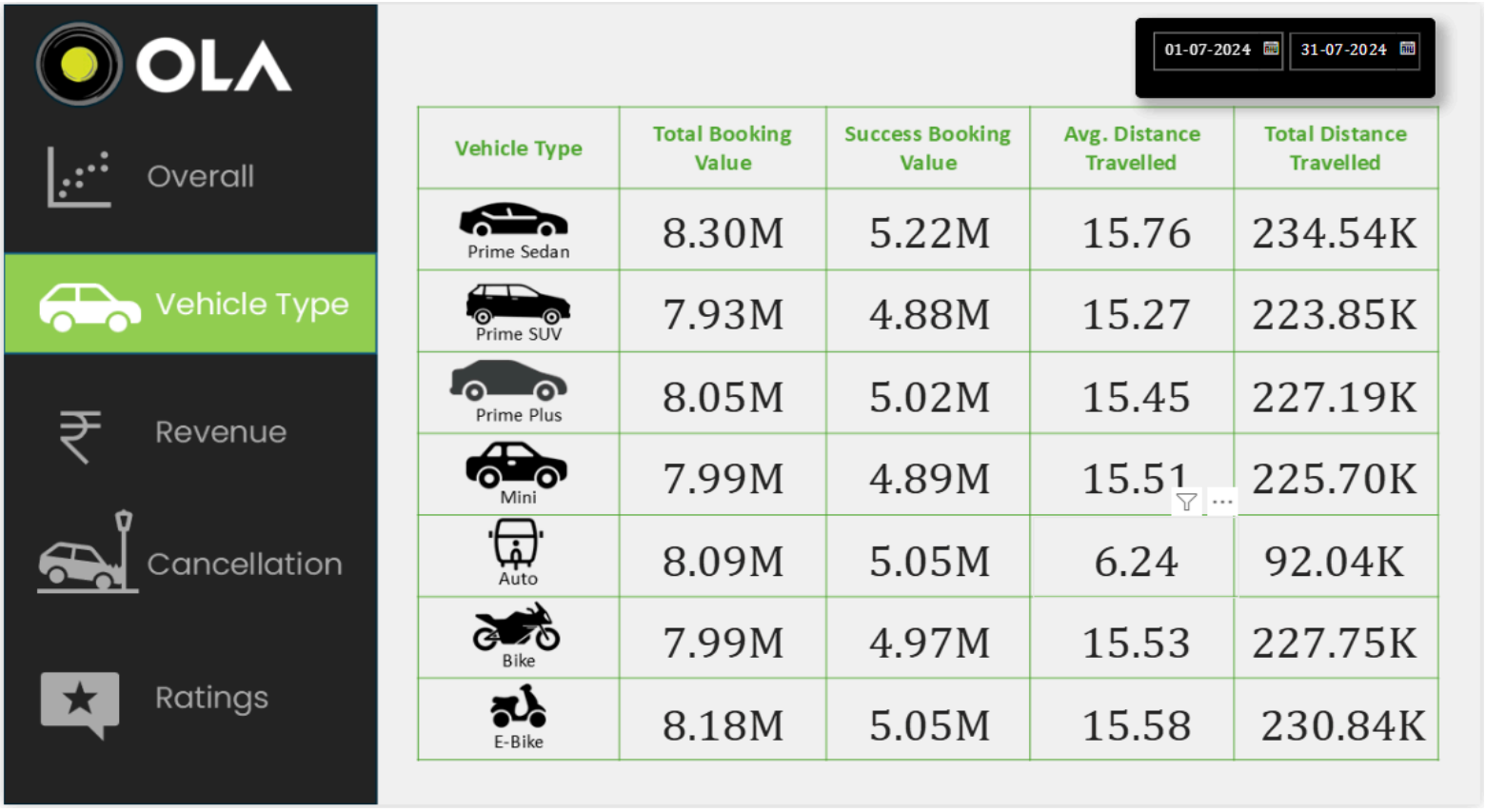
To present the findings from the Power BI tasks effectively, I have organized the analysis across five dedicated pages, each functioning as a focused dashboard. This multi-page structure ensures that insights are displayed clearly, logically, and in a visually engaging manner, enabling managers and stakeholders to easily understand and connect with the key outcomes of the analysis.

These dashboards highlight critical aspects of Ola's ride-hailing operations, including ride trends, revenue metrics, customer behavior, driver performance, and geographical patterns. By leveraging interactive visuals such as charts, maps, slicers, and KPIs, the dashboards facilitate quick comprehension and data-driven discussions, even for audiences with varying levels of technical expertise.

In this document, I am providing a comprehensive overview of each dashboard, including its purpose, key visuals, and the primary insights it conveys. This will serve as a reference guide to the Power BI report and demonstrate the depth and clarity of the analytical work performed.

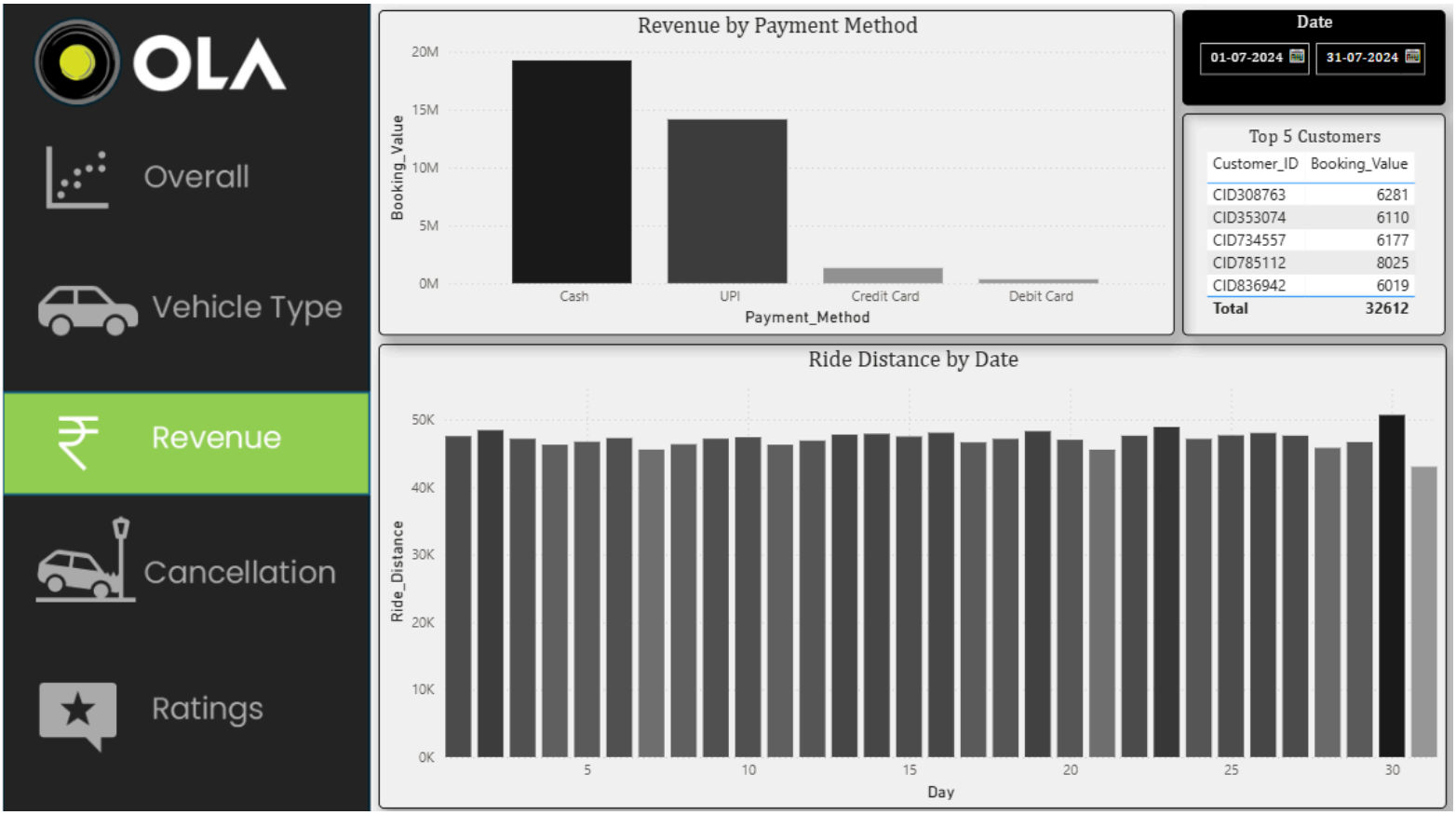


Page-1 : Overall

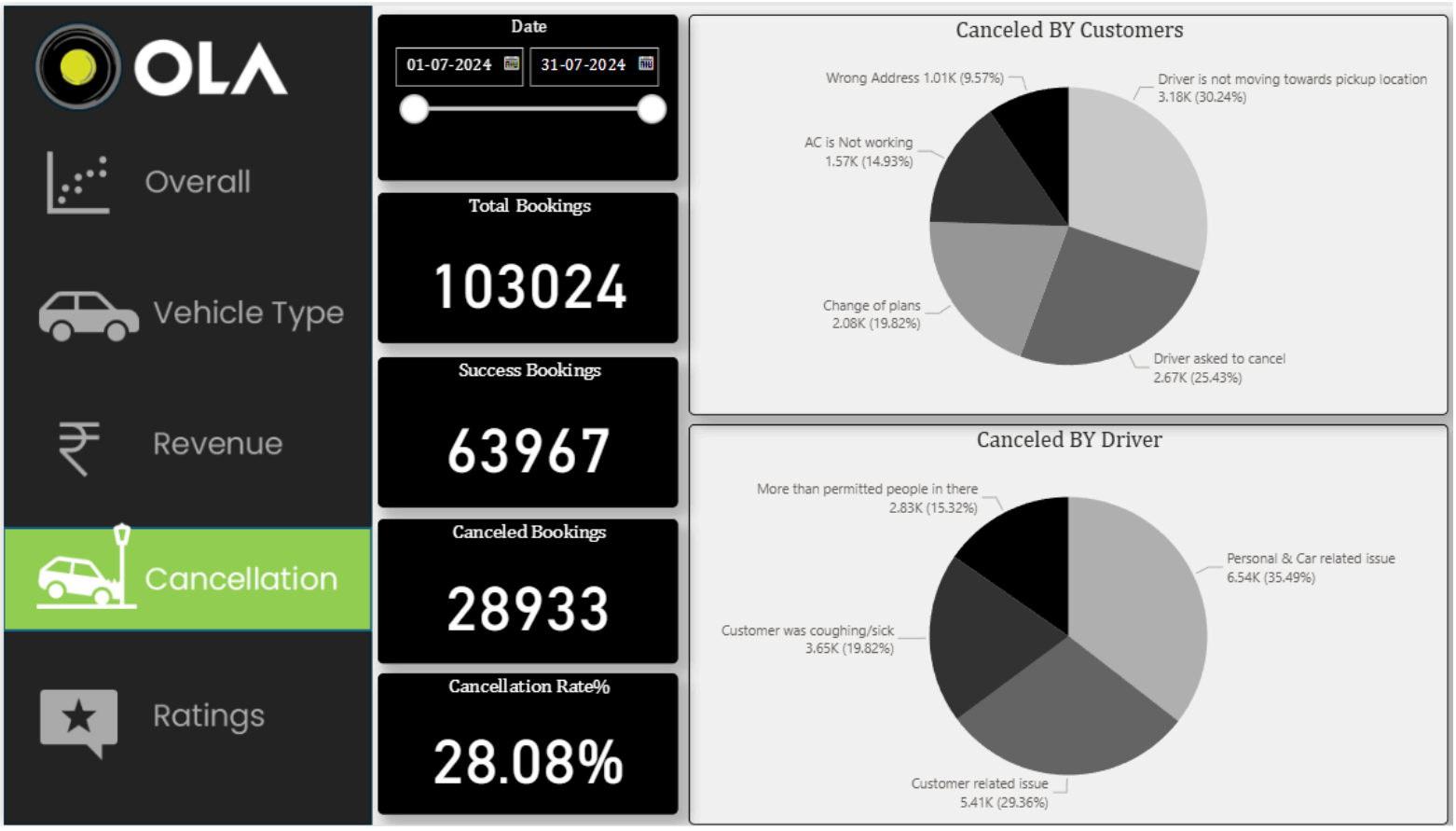


Vehicle Type	Total Booking Value	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
Prime Sedan	8.30M	5.22M	15.76	234.54K
Prime SUV	7.93M	4.88M	15.27	223.85K
Prime Plus	8.05M	5.02M	15.45	227.19K
Mini	7.99M	4.89M	15.51	225.70K
Auto	8.09M	5.05M	6.24	92.04K
Bike	7.99M	4.97M	15.53	227.75K
E-Bike	8.18M	5.05M	15.58	230.84K

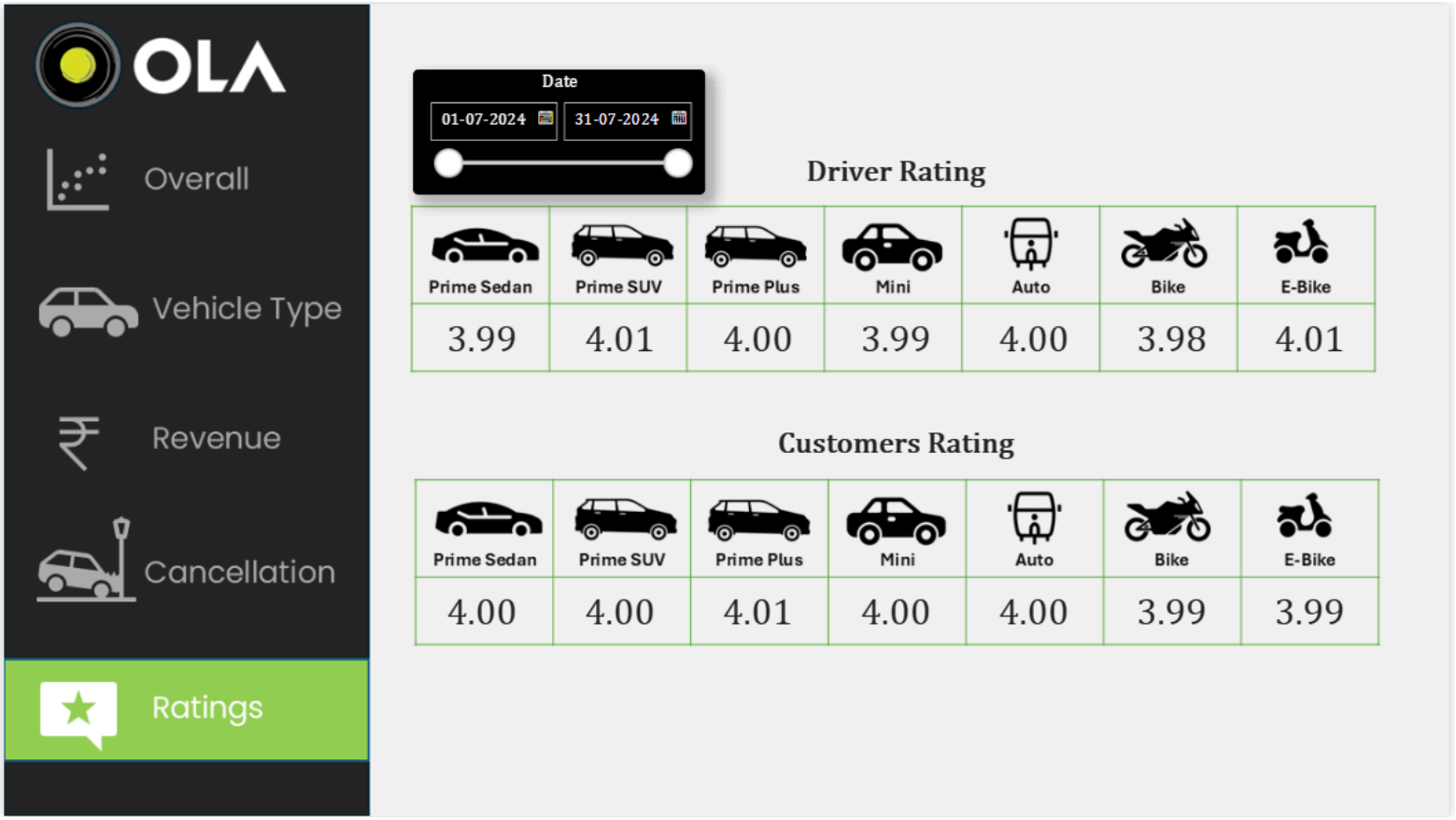
Page-2 : Vehicle Type



Page-3 : Revenue



Page-4 : Cancellation



Page-5 Rating

Conclusions:

1. **Strong Preference for Cash Payments:** Cash dominates as the primary payment method, contributing the highest revenue share (approximately twice that of UPI and significantly more than card payments). This indicates a continued reliance on cash transactions in the market, suggesting the need to maintain robust cash-handling processes while gradually promoting digital alternatives for efficiency.
2. **High Overall Booking Volume and Revenue:** Total bookings exceeded 1 million (1,030,024) in July 2024, generating ₹35.08 million in booking value. Successful bookings accounted for about 62% (639,967), highlighting a solid core business performance but also room to reduce the substantial cancellation rate.
3. **Significant Cancellation Challenge:** The overall cancellation rate stands at 28.08%, with 289,933 canceled bookings.

- **Primary reasons from customers:** Change of plans (19.82%), driver asked to cancel (25.43%), and AC not working (14.93%).
- **Primary reasons from drivers:** Personal/car-related issues (35.49%) and more than permitted people (15.32%). These insights point to opportunities for improving driver training, vehicle maintenance, and policy enforcement to minimize avoidable cancellations.

4. **Consistent Ride Patterns and Distance:** Daily ride distances show stable variation throughout the month, with average distances ranging from 15–16 km for most car and bike categories. Auto rides have noticeably shorter average distances (~6 km), aligning with their typical use case for short trips.
5. **Vehicle Type Performance:**

- E-Bike and Prime Sedan categories lead in total booking value (~₹8.3M each), followed closely by Auto and Prime Plus.
- Success rates are relatively consistent across vehicle types (around 60–62%), indicating balanced demand.
- Bike and E-Bike categories show strong total distances traveled, reflecting their popularity for longer or cost-effective rides.

6. **High Customer and Driver Satisfaction:** Both driver and customer ratings across all vehicle types are consistently high, ranging from 3.98 to 4.01 (out of 5). This reflects excellent service quality and positive experiences for both riders and drivers, serving as a strong competitive advantage.
7. **Top Customer Contribution:** A small group of frequent users drives significant value, with the top 5 customers contributing over ₹32,000 in booking value combined, underscoring the importance of loyalty programs or personalized offers to retain high-value users.

- **Overall Recommendation:**

While Ola demonstrates robust demand, healthy revenue generation, and excellent service ratings, the elevated cancellation rate represents the most critical area for improvement. Targeted interventions—such as enhanced vehicle checks, clearer passenger policies, and incentives for completing rides—could substantially boost successful bookings and overall profitability.

- **Next Steps / Future Action:**

1. The biggest priority should be bringing down that cancellation rate – 28% is way too high. Give drivers better training to ensure AC and vehicle maintenance are always spot-on, and get a bit stricter with drivers or customers who keep canceling rides. This would directly boost revenue.
2. Cash payments are dominating, but they're risky and costly for the company. Offer some discounts or cashback on UPI/wallet payments to slowly shift people toward digital options – cash usage will drop over time.
3. Those top customers who take tons of rides deserve special treatment. Create a simple loyalty program with free rides, priority bookings, or upgrades – it'll make them even more loyal.
4. E-Bikes and Prime Sedans are performing really well. Add more vehicles in these categories, especially in cities where people want eco-friendly or comfortable options.
5. With the data we have now, build a machine learning model that predicts when and where demand will spike. This way, we can allocate drivers more intelligently and reduce those "driver not found" issues.