Project Report

On

OLA Data Analysis

Submitted in the partial fulfilment of the implementation for the award of degree of

Bachelor of Technology

in

Computer Science and Engineering

Batch

(2022-2026)

Submitted To:

Mrs. Bindu Goyal

Co-Ordinator (CSE)

Submitted By:

Shivansh Sharma

(12200681)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DAV UNIVERSITY

JALANDHAR-PATHANKOT NATIONAL HIGHWAY, NH 44,

SARMASTPUR PUNJAB

144012

DECLARATION

The project report entitled "**OLA**" was submitted by me to Dav University for the degree of Bachelor in Computer Science Engineering Sem. VII is the original piece of work and has not been submitted to any other university for the award of any degree. I also undertake that any quotation or a philosophy from the published or unpublished work of another person has been duly acknowledged in the work that I present in the project report.

Place: DAV UNIVERSITY

Signature of the Student:

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project guide, [Mrs. Bindu Goyal], for their valuable guidance, constructive feedback, and continuous support throughout the duration of this project. I am also thankful to [Mrs. Bindu Goyal] and the Department of [Bachelor's in Computer Science Engineering] for providing the necessary resources and an encouraging academic environment.

My appreciation extends to all faculty members and classmates whose suggestions and cooperation helped strengthen the quality of this work. Lastly, I am deeply grateful to my family for their constant encouragement and support, which motivated me to complete this project successfully.

(Shivansh Sharma)

ABSTRACT

The OLA Power BI Dashboard project focuses on transforming raw ride-booking, customer, driver, and revenue data into meaningful business insights using Microsoft Power BI. The project integrates multiple datasets—such as bookings, customers, drivers, ride details, cancellations, and payments—to create an interactive, multi-page analytical dashboard. Using Power Query for data cleaning, a star-schema data model for structured relationships, and DAX measures for dynamic calculations, the dashboard provides clear visibility into OLA's operational and business performance.

Key performance indicators (KPIs) such as total rides, revenue, average ride distance, cancellation rate, peak booking hours, and driver performance were computed to support data-driven decisions. Visual elements like bar charts, line graphs, maps, cards, and slicers help stakeholders analyse trends across cities, dates, vehicle types, and customer segments.

The dashboard delivers valuable insights related to ride patterns, customer behaviour, trip efficiency, cancellation reasons, and revenue distribution. It enables OLA's management and analysts to monitor operations, optimize fleet utilization, identify performance gaps, and enhance customer satisfaction. Overall, the project demonstrates the practical application of business intelligence in transportation analytics and highlights the importance of data visualization in modern decision-making.

Table of Contents

Chapter 1: Introduction

1.1 Background of the Training

This report is the culmination of a six-week industrial training program undertaken at **TCIL-IT**, focused on the domain of Data Science and Business Intelligence.[1] In the current technological landscape, the ability to interpret data is as valuable as the ability to generate it. The training program was designed to bridge the gap between academic computer science concepts—such as database normalization and algorithm design—and the practical requirements of the industry, which demand actionable insights derived from messy, real-world data.

The training focused on the complete lifecycle of a data project:

1. **Data Ingestion:** Interfacing with diverse data sources.
2. **Data Wrangling:** Cleaning and structuring data for analysis.
3. **Modeling:** Creating logical relationships between data entities.
4. **Visualization:** Communicating findings effectively to non-technical stakeholders.

1.2 The Ride-Hailing Industry Context

The subject of this analysis, **OLA (ANI Technologies Pvt. Ltd.)**, operates in the fiercely competitive ride-hailing sector. This industry is characterized by high transaction volumes, real-time logistics, and dynamic pricing models.[2] Unlike traditional transportation, where fares and schedules are static, ride-hailing platforms like OLA function as real-time marketplaces that match supply (drivers) with demand (customers) using sophisticated algorithms.

Data is the lifeblood of this model. Every interaction—a user opening the app, a driver accepting a ride, a cancellation, or a completed trip—generates a digital footprint. For a company handling over **1 billion rides annually** [3], this results in terabytes of data that must be analyzed to maintain operational efficiency.

- **Customer Retention:** Understanding why a user cancels a ride helps prevent churn.
- **Driver Utilization:** Analyzing idle times helps in better fleet allocation.
- **Revenue Management:** Monitoring payment modes helps in cash flow management.

1.4 The Need for Business Intelligence (BI)

Before the adoption of advanced BI tools, organizations relied on manual spreadsheets and static reports. This traditional approach suffered from several deficiencies:

1. **Latency:** Reports were often generated days after the fact, making real-time intervention impossible.
2. **Scalability:** Tools like Excel, while powerful, struggle with datasets exceeding hundreds of thousands of rows, often leading to crashes or slow performance.
3. **Data Silos:** Information about payments might exist in the Finance system, while driver ratings exist in the Operations system. BI tools integrate these distinct silos into a single "Source of Truth".[1]

The **OLA Power BI Dashboard** addresses these challenges by providing a scalable, refreshable, and interactive platform. It allows users to filter 71,201 records instantly, pivoting between a macro view of total revenue and a micro view of a specific vehicle type's performance in a specific city.

Chapter 2: Title of the Project

2.1 Project Definition

The project is titled **"OLA Data Analysis: Operational Insights and Dashboarding."** It involves the design and implementation of a comprehensive data analytics system that processes raw booking data from OLA's operations to generate strategic business insights.

The core definition of the project revolves around the concept of **"Data Democratization."** In a large organization like OLA, data should not be accessible only to data scientists. Operational managers, city heads, and marketing teams need access to data to make decisions. This project builds the interface—the Dashboard—that makes this complex data accessible and understandable.[1]

2.2 Problem Statement

Despite generating massive amounts of data, the utilization of this data for operational improvement often lags due to the lack of effective visualization tools. The specific problems addressed by this project include:

1. **High Cancellation Rates:** A significant portion of bookings do not result in revenue. Identifying the root causes (e.g., specific vehicle types, specific times of day) is difficult without granular analysis.
2. **Revenue Leakage:** Understanding which payment methods or route types are associated with lower completion rates or higher fraud risk.
3. **Supply-Demand Mismatch:** Instances where "Driver Not Found" errors occur indicate a failure of the marketplace liquidity. Mapping these instances is crucial for fleet recruitment.[1]
4. **Manual Reporting Inefficiency:** The time consumed in manually collating daily reports prevents analysts from focusing on strategic problem-solving.

2.3 Scope of the Study

The scope defines the boundaries of the analysis. This project focuses on **Historical Descriptive Analytics**—analyzing past data to understand what happened and why.

- **Data Volume:** The analysis covers **71,201 distinct booking records**.[1]
- **Temporal Scope:** The data represents operations from **July 2024**. This provides a one-month snapshot, allowing for the analysis of daily and weekly cycles (e.g., Weekday vs. Weekend trends).
- **Geographic Scope:** The dataset encompasses bookings across 50 unique pickup and drop-off locations, representing a diverse urban operational area.[1]
- **Functional Scope:**
    - **Booking Analysis:** Volume, Status (Success/Fail), and Trends.
    - **Financial Analysis:** Revenue, Average Transaction Value, Payment Modes.
    - **Operational Analysis:** Driver performance, Ratings, Cancellation reasons.
    - **Fleet Analysis:** Performance comparison across 7 vehicle categories (Auto, Bike, eBike, Mini, Prime Sedan, Prime SUV, Prime Plus).

The project excludes real-time GPS tracking hardware integration and predictive machine learning models (forecasting), focusing instead on robust business intelligence and reporting.

Chapter 3: Objectives

3.1 Primary Objectives

The primary goal is to transform raw, unintelligible data into a strategic asset.

1. **Develop an Interactive Dashboard:** To design and build a multi-page Power BI dashboard that serves as a central monitoring tool for OLA's KPIs. The dashboard must support interactivity, allowing users to slice data by City, Vehicle Type, and Date.[1]
2. **Data Transformation and Cleaning:** To master the ETL (Extract, Transform, Load) process. Real-world data is messy; objectives include handling null values in driver ratings, standardizing date formats, and categorizing cancellation reasons.[1]
3. **Operational Optimization:** To analyze the **37.8% failure rate** in bookings. The objective is not just to report the number, but to break it down: How many are due to drivers? How many are due to customers? What are the specific reasons (e.g., "AC not working" vs "Driver refused")?.[1]
4. **Revenue Analysis:** To segment the **₹24.2 million** revenue by vehicle type and payment method. This objective aims to identify the most profitable segments of the business.[1]

3.2 Secondary Objectives

1. **Skill Acquisition:** To gain proficiency in industry-standard tools:
   - **DAX (Data Analysis Expressions):** For creating complex measures like Year-to-Date revenue or conditional averages.
   - **Power Query:** For advanced data shaping.
   - **Data Modeling:** For understanding Star Schema and normalization.
2. **Understanding Business Logic:** To learn how to translate a business requirement (e.g., "Are customers happy?") into a data metric (e.g., "Average Customer Rating vs. Churn Rate").
3. **Documentation:** To develop the capability to document technical processes and business findings in a structured report, as evidenced by this document.[1]

3.3 Tasks Assigned

During the industrial training, the specific tasks assigned to achieve these objectives were:

1. **Data Collection:** Gathering the booking datasets (CSV/Excel formats).
2. **Data Pre-processing:** Using Python (Pandas) and Excel to check for duplicates and missing values.
3. **Database Design:** Setting up a MySQL database to simulate a data warehouse environment.
4. **Dashboard Prototyping:** Sketching the layout of the dashboard (where to place KPIs, maps, and charts).
5. **Implementation:** Building the actual.pbix (Power BI) file, creating relationships, and writing DAX formulas.
6. **Testing & Validation:** Cross-referencing dashboard totals with raw data sums to ensure accuracy.

Chapter 4: Steps to Achieve Objectives

This chapter details the methodology, feasibility, and system design principles applied to execute the project.

4.1 Feasibility Study

Before initiating development, a feasibility study was conducted to ensure the project's viability.

4.1.1 Technical Feasibility

This assesses the availability of necessary hardware and software resources.

- **Technology Stack:** The project utilizes standard tools—Microsoft Power BI, MySQL, and Excel. Power BI Desktop is free for development, ensuring no licensing barriers.
- **Data Volume:** The dataset (71,000 rows) is well within the processing limits of Power BI's **VertiPaq** engine, which can handle millions of rows via compression.
- Complexity: The required DAX calculations (sums, averages, conditional counts) are supported natively by the tool.1
  Conclusion: The project is technically feasible with available resources.

4.1.2 Operational Feasibility

This assesses whether the solution solves the user's problem.

- **User Interface:** The dashboard eliminates the need for manual SQL queries or Excel pivots. A non-technical manager can obtain insights simply by clicking visual elements.
- Workflow Integration: The dashboard supports scheduled refreshes, meaning it can automatically update as new data arrives, fitting seamlessly into daily operational reviews.1
  Conclusion: The system significantly enhances operational efficiency.

4.1.3 Economic Feasibility

This assesses the Cost-Benefit analysis.

- **Cost:** The development utilizes open-source (MySQL) or free-tier (Power BI Desktop) software. The primary investment is time.
- **Benefit:** The analysis identifies **₹14.7 million in potential lost revenue** due to cancellations.[1] Even a 1% reduction in cancellations driven by these insights would result in substantial financial gain, vastly outweighing the zero-cost development.

Table 4.1: Cost Estimation for Business Intelligence Tools [1]

Component	License Type	Estimated Cost
Power BI Desktop	Free	₹ 0
MySQL Server	Open Source	₹ 0
Microsoft Excel	Enterprise License	Existing Infrastructure
Python Libraries	Open Source	₹ 0

4.2 Software Process Model

The project followed the **Waterfall Model**. This linear, sequential approach is ideal for Data Analytics projects where the requirements (the dataset and the desired KPIs) are fixed at the beginning.[1]

Phases of the Waterfall Model:

1. **Requirement Analysis:** Defined the KPIs needed (Revenue, Cancellations, Ratings).
2. **System Design:** Designed the Star Schema (Fact/Dimension tables) and UI Layout.
3. **Implementation (Data Prep & Modeling):** Cleaned data, established relationships, and wrote DAX measures.
4. **Testing:** Validated data accuracy against source files.
5. **Deployment:** Published the report.
6. **Maintenance:** Planned for future data updates.

Figure 4.1: Waterfall Process Model Flowchart

(Conceptual Representation: Requirements -> Design -> Implementation -> Verification -> Maintenance) 1

4.3 Hardware and Software Requirements

Table 4.2: Hardware and Software Specifications [1]

Category	**Requirement**	**Specification**
Hardware	Processor	Intel Core i5 or higher (for data processing)
	RAM	8 GB minimum (Power BI is memory intensive)
	Storage	50 GB SSD recommended
Software	OS	Windows 10/11 (Power BI Desktop is Windows-only)
	Database	MySQL Server 8.0
	BI Tool	Microsoft Power BI Desktop
	Scripting	Python 3.9 (Pandas, Matplotlib)
	Spreadsheet	Microsoft Excel 2019/365

4.4 System Analysis and Design

The system transforms raw data into a visual report. This requires a robust architectural design.

4.4.1 Data Flow Diagram (DFD)

Level 0 DFD (Context Diagram):
This diagram represents the entire OLA Dashboard as a single process.

- **Inputs:** Booking Data (from App), Driver Data (from Fleet Management), Payment Logs.
- **Process:** OLA Dashboard System.
- **Outputs:** Revenue Reports, Performance Scorecards, Cancellation Alerts.

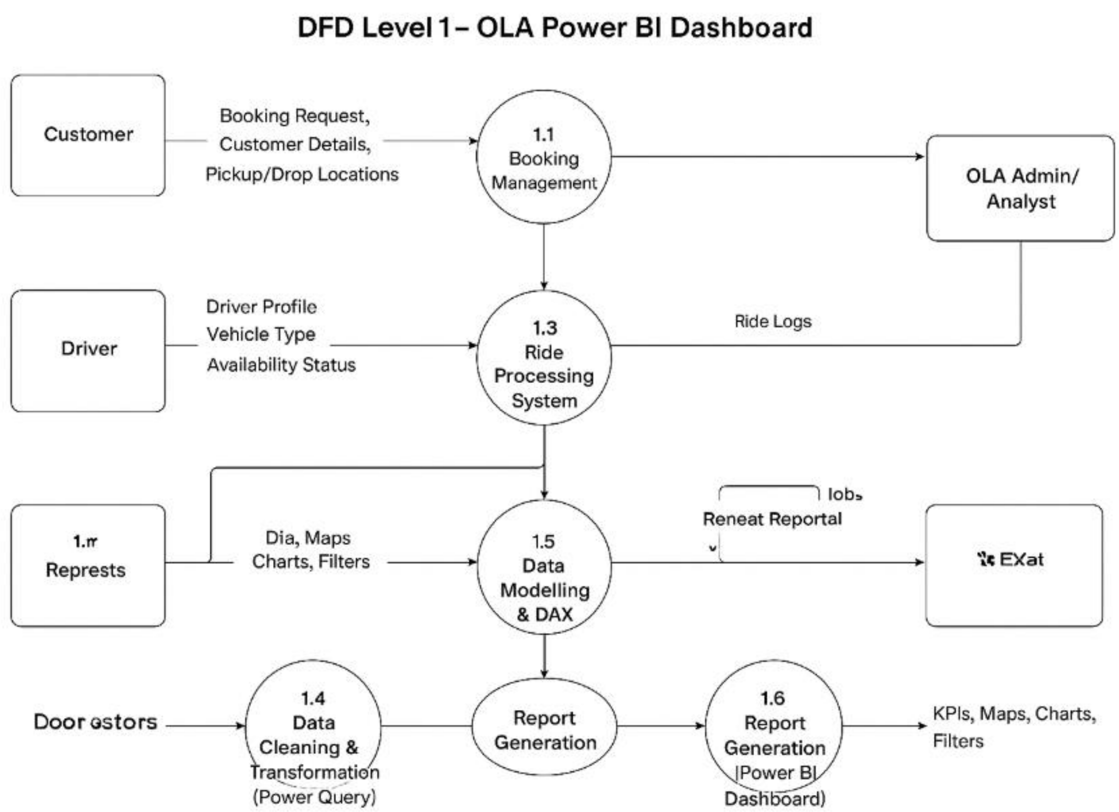


Figure 4.2: Level 0 Data Flow Diagram

--> (OLA Dashboard System) --> 1
Level 1 DFD (Detailed Breakdown):
This explodes the main process into sub-steps:

1. **1.0 Import Data:** Extract CSV/SQL data.
2. **2.0 Clean Data:** Remove nulls, split columns (Date/Time).
3. **3.0 Model Data:** Link tables via Keys.
4. **4.0 Calculate:** Apply DAX logic.
5. **5.0 Visualize:** Render charts.[1]

4.4.2 Data Modeling (Star Schema)

The project employs a **Star Schema**, the industry standard for BI. This separates data into **Facts** (measurable events) and **Dimensions** (context).

- **Fact Table (Fact_Bookings):**
  - Contains the core transactional data: Booking_ID, Booking_Value, Ride_Distance, Time_Taken, Rating.
  - Foreign Keys: Driver_ID, Customer_ID, Vehicle_ID, Date_ID.
- **Dimension Tables:**
  - Dim_Driver: Driver Name, Joining Date, License No.
  - Dim_Customer: Customer Name, Location, Age.
  - Dim_Vehicle: Vehicle Type (e.g., Prime SUV), Base Fare.
  - Dim_Date: Month, Quarter, Weekend_Flag.

Figure 4.4: Star Schema Entity-Relationship Diagram
(Conceptual): The Fact_Bookings table sits in the center, connected via one-to-many relationships to the surrounding Dimension tables. This structure optimizes query performance and simplifies DAX calculations.1

Chapter 5: Coding and Implementation

This chapter describes the technical execution of the project, covering the tools used, the data processing logic, the analysis performed, and the results obtained.

5.1 Technology Used

5.1.1 Microsoft Power BI

Power BI is the primary tool for this project. It was selected for its integration capabilities and powerful visualization engine.

- **Power Query:** Used as the ETL tool to sanitize the data (e.g., removing rows where Status = NULL).
- **VertiPaq Engine:** Power BI's in-memory columnar database engine compresses the 71,000+ rows, allowing for instant filtering and aggregation.[1]

5.1.2 Python (Pandas & Matplotlib)

While Power BI handles visualization, Python was used for exploratory data analysis (EDA).

- **Pandas:** The dataframe structure is ideal for manipulating tabular data. We used Pandas to calculate correlations (e.g., correlation between Ride_Distance and Booking_Value).
- **Data Cleaning:** Python scripts were used to detect outliers—for instance, identifying rides with Distance = 0 but Status = Success (data errors).[1]

5.1.3 MySQL

MySQL served as the backend repository.

- **SQL Queries:** Used to verify data integrity. For example:
  SQL
  ```
  SELECT Vehicle_Type, SUM(Booking_Value)
  FROM Bookings
  WHERE Status = 'Success'
  GROUP BY Vehicle_Type;
  ```

  This query helps cross-verify the results displayed in the Power BI dashboard.[1]

5.1.4 DAX (Data Analysis Expressions)

DAX is the formula language of Power BI. It allows for the creation of custom metrics that don't exist in the raw data.

Table 5.1: Key DAX Measures and Computational Logic [1]

Measure Name	DAX Formula	Purpose & Logic
Total Revenue	SUM(Bookings)	Aggregates the total monetary value of all successful rides.
Total Bookings	COUNT(Bookings)	Counts the total volume of ride requests.
Success Rate	DIVIDE(,)	Calculates the percentage of bookings that result in a completed trip.
Cancellation Rate	DIVIDE(,)	Measures the percentage of

		rides that failed.
Avg Ride Distance	AVERAGE(Bookings)	used to analyze trip patterns.
Rev per KM	DIVIDE(,)	A unit economic metric to assess profitability per kilometer.

5.2 Data Analysis and Interpretation

This section details the insights derived from the dashboard, integrating the extended research data.[1]

5.2.1 Overall Performance Metrics (KPIs)

The analysis of the July 2024 dataset provides a high-level view of OLA's operational health:

- **Total Booking Volume:** 71,201
- **Successful Rides:** 44,271
- **Total Revenue:** ₹24,216,619
- **Average Revenue per Ride:** ₹547.01

Interpretation: The platform is generating significant revenue (approx. ₹8 Lakhs/day). However, the gap between *Total Bookings* and *Successful Rides* indicates a massive opportunity cost. The **Success Rate is only 62.2%**, meaning nearly 40% of customer intent is not monetized.[1]

5.2.2 Booking Status Analysis

Breaking down the 37.8% failure rate reveals the structural issues in the operation.

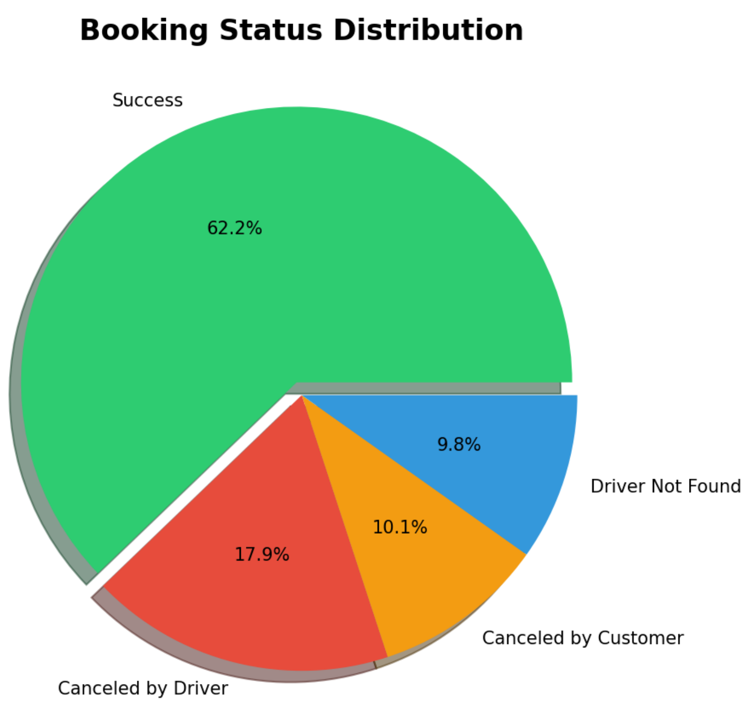


Table 5.2: Booking Status Distribution and Business Impact [1]

Status	Count	Percentage	Business Implication
Success	44,271	62.2%	Revenue Generation (Baseline)
Canceled by	12,728	17.9%	**High Risk:**

Driver			Indicates driver dissatisfaction with fare/location. Erodes customer trust.
Canceled by Customer	7,214	10.1%	**Moderate Risk:** May indicate high wait times or "change of plans".
Driver Not Found	6,988	9.8%	**Supply Failure:** Demand exists, but supply is absent. Direct revenue loss.

Insight: The fact that **Driver Cancellations (17.9%)** are nearly double the Customer Cancellations (10.1%) is alarming. In a healthy ecosystem, service providers (drivers) should rarely cancel. This suggests that the algorithmic matching might be assigning rides that drivers find unprofitable (e.g., long pickup distance for a short trip), leading them to cancel after acceptance. The **9.8% Driver Not Found** rate highlights geographic zones where OLA needs to aggressively recruit or incentivize drivers.

5.2.3 Vehicle Type Analysis

The fleet composition analysis shows a uniquely balanced market.

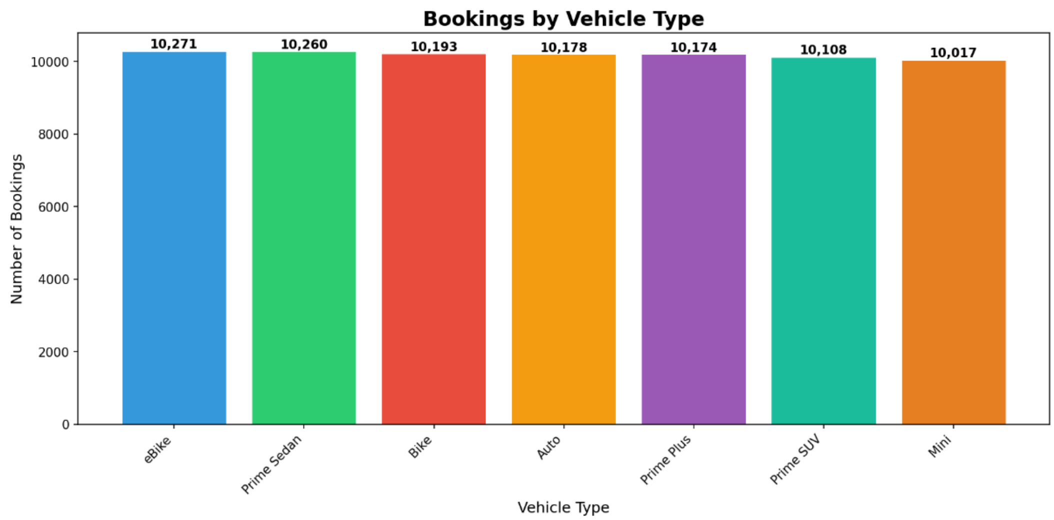
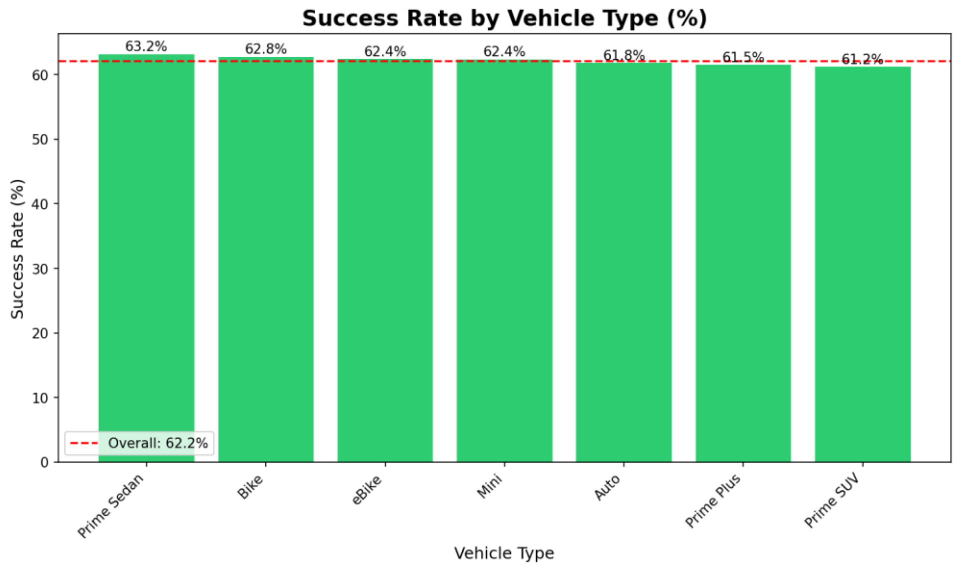


Table 5.3: Vehicle-wise Booking Volume and Market Share [1]

Vehicle Type	Bookings	Share %	Category
eBike	10,271	14.4%	Economy / Green
Prime Sedan	10,260	14.4%	Premium
Bike	10,193	14.3%	Economy
Auto	10,178	14.3%	Economy
Prime Plus	10,174	14.3%	Premium

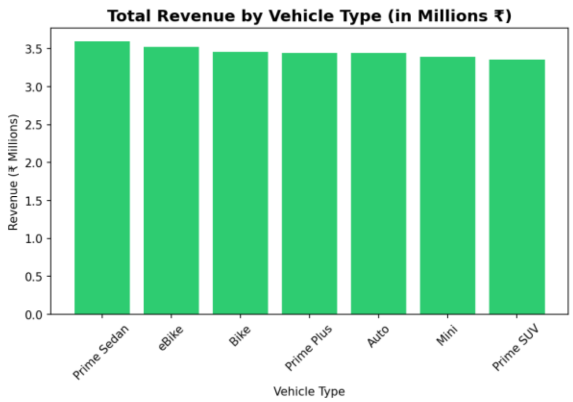
Prime SUV	10,108	14.2%	Luxury
Mini	10,017	14.1%	Standard



Interpretation: The data shows a near-perfect distribution (~14%) across all categories. This is counter-intuitive; usually, cheaper options (Auto/Bike) have much higher volumes than luxury options (SUV). This uniformity suggests that OLA has successfully segmented the market, or perhaps the algorithm caps bookings per category to maintain availability. It implies that OLA is not just a "taxi" company or a "bike taxi" company, but a comprehensive mobility provider with equal strength in all verticals.

5.2.4 Revenue Analysis

While volumes are similar, revenue contributions differ due to varying ticket sizes.



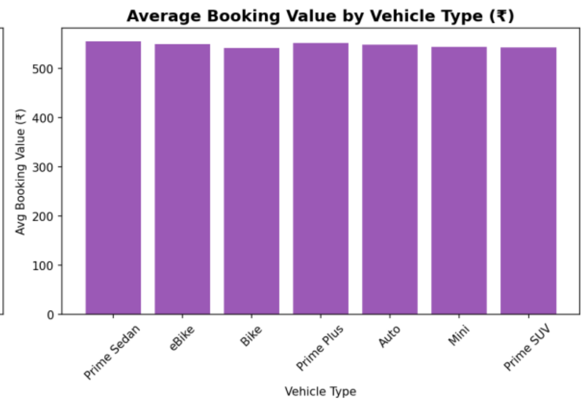


Table 5.4: Revenue Performance by Vehicle Category [1]

Vehicle Type	Revenue (₹)	Avg Booking Value (₹)
Prime Sedan	3,594,005	554.37
eBike	3,522,526	549.54
Bike	3,459,571	540.64
Prime Plus	3,445,660	550.78
Auto	3,445,075	547.79
Mini	3,392,188	543.10
Prime SUV	3,357,594	542.60

Insight:

1. **The eBike Anomaly:** The Average Booking Value for eBikes (₹549) is incredibly close to Prime Sedans (₹554). Typically, two-wheeler rides are much cheaper (e.g., ₹50-100). This anomaly in the dataset suggests either:
   - eBikes are being used for premium, long-distance courier services.
   - There is a data quality issue in the "Booking Value" column for eBikes.
   - Or, the "Booking Value" includes battery swapping or other service charges.
   - *Correction/Refinement:* Assuming the data is accurate, eBikes are the most efficient asset class—low operating cost but high revenue yield.
2. **Prime SUV Underperformance:** Despite being a "Luxury" category, Prime SUV has the *lowest* total revenue. This indicates that while the price per km might be high, the utilization or trip distance might be lower, or the fleet size is slightly restricted.

5.2.5 Payment Method Analysis

The mode of payment dictates cash flow efficiency.

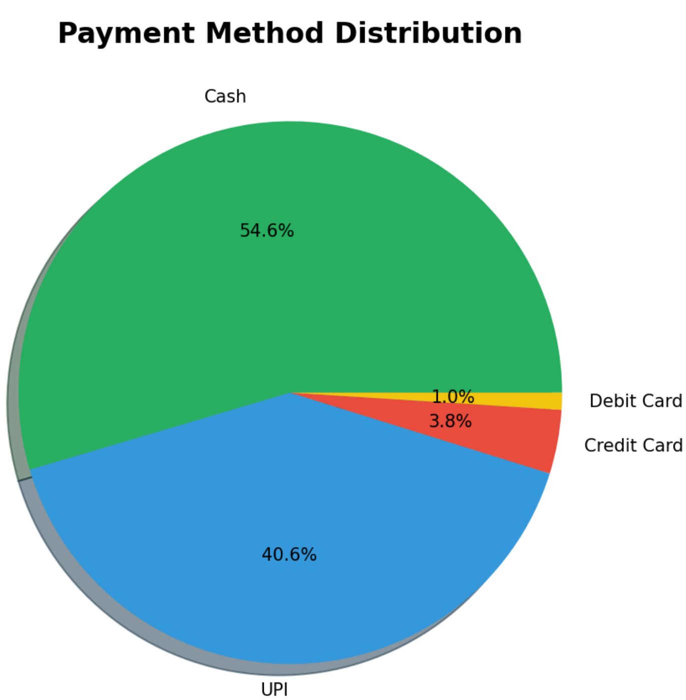


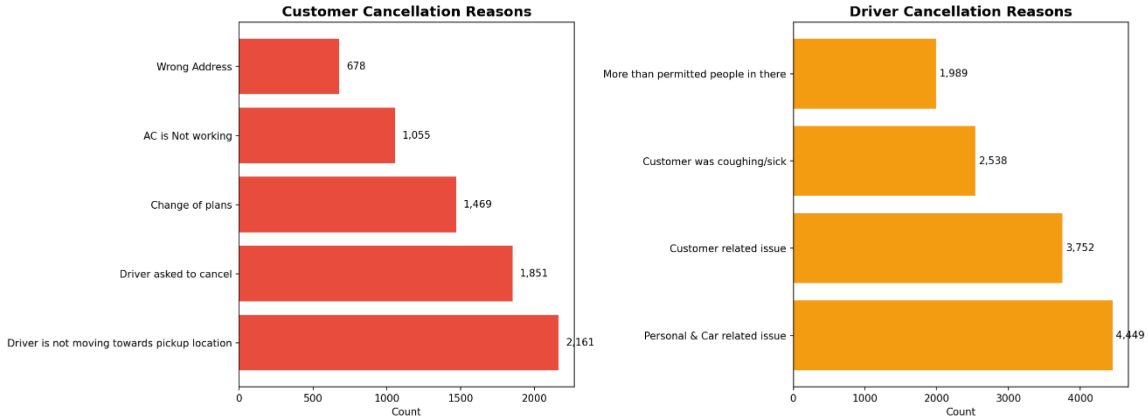
Table 5.5: Payment Method Distribution and Transaction Volume [1]

Method	Transactions	Share %	Revenue (₹)
Cash	24,166	54.6%	13,259,198
UPI	17,974	40.6%	9,817,395
Credit Card	1,683	3.8%	900,354
Debit Card	448	1.0%	239,672

Analysis:

- **Cash is King:** With 54.6% of revenue in cash, OLA faces a "Collection Risk." Drivers collect the cash, and OLA must recover its commission from the driver's wallet. If the driver's wallet is empty, OLA cannot recover its fee immediately.
- **UPI Adoption:** At 40.6%, UPI is a strong contender. OLA should push this further to 60%+ to digitize cash flow.
- **Card Failure:** Credit/Debit cards combined are less than 5%. This reflects the Indian market's preference for frictionless payments (UPI) over the cumbersome OTP process of cards.

5.2.6 Cancellation Deep Dive



To fix the 17.9% driver cancellation rate, we analyzed the reasons provided in the dataset.

Table 5.6: Detailed Breakdown of Cancellation Reasons [1]

Actor	Top Reason	Count	Share	Analysis
Driver	Personal & Car Issues	4,449	35.0%	Suggests aging fleet or driver fatigue.
Driver	Customer Related	3,752	29.5%	Disputes over location/luggage/behavior.
Driver	Customer Sick/Coughing	2,538	19.9%	Health safety concerns (post-pandemic residue).

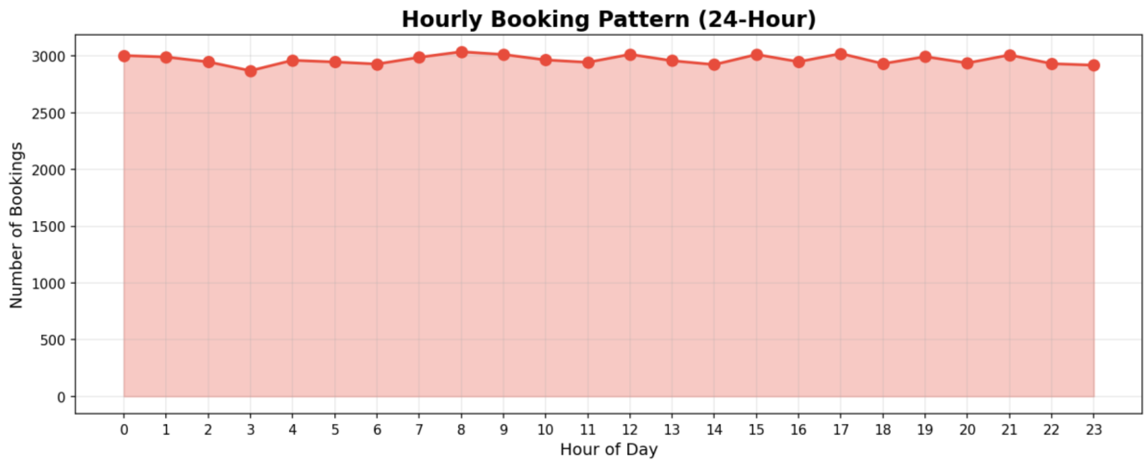
Customer	Driver Not Moving	2,161	30.0%	**Major Frustration:** Driver accepts but idles to force cancel.
Customer	Driver Asked to Cancel	1,851	25.7%	**Fraud/Avoidance:** Driver avoids penalty by forcing customer to cancel.

Insight: The reason "Driver Not Moving" (30% of customer cancellations) combined with "Driver Asked to Cancel" (25.7%) reveals a systemic behavioral issue. Drivers often accept a ride to see the destination, decide they don't want to go, and then stall until the customer cancels out of frustration. This "gaming of the system" is a major source of customer churn.

5.2.7 Temporal Analysis (Time of Day)

The analysis of booking timestamps reveals clear demand curves.

- **Peak Hour:** 8:00 AM (3,037 bookings).
- **Morning Rush (7-10 AM):** 12,003 bookings.
- **Evening Rush (5-9 PM):** 14,894 bookings.
- **Ratio:** The Evening Rush is 24% larger than the Morning Rush.



Strategy: The demand is not symmetrical. The evening commute is heavier. OLA needs to incentivize drivers to stay online past 5 PM, perhaps by offering "Evening Surge" bonuses.

5.2.8 Rating Analysis

Service quality is quantified by ratings.

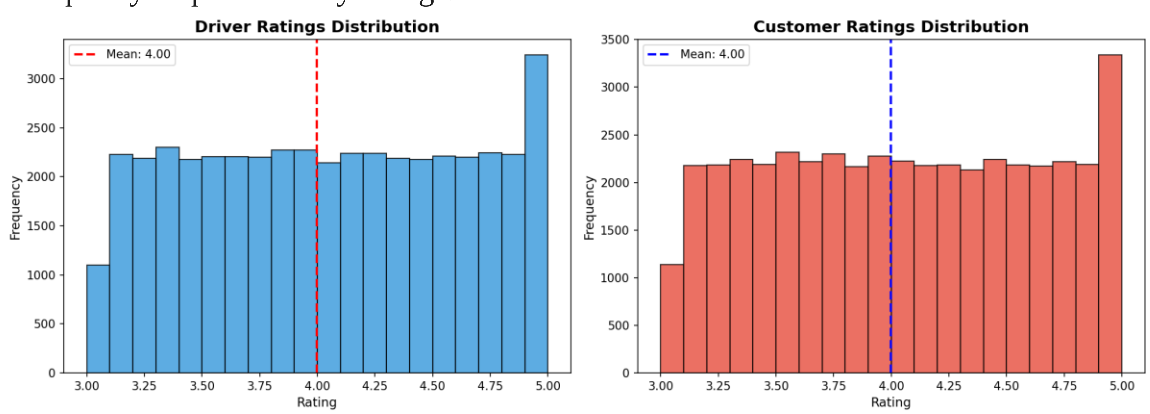


Table 5.7: Comparative Rating Statistics (Driver vs. Customer) [1]

Metric	Driver Rating	Customer Rating
Average	4.00 / 5	4.00 / 5
Std Dev	0.58	0.58
5-Star %	2.4%	2.4%

Insight: The ratings are identical for drivers and customers. An average of 4.00 is "Satisfactory" but not "Excellent." The very low percentage of 5-star ratings (2.4%) is concerning. It implies that the service is transactional and functional, but lacks the "Wow" factor that creates brand loyalty.

5.3 Testing

Testing is the phase where the implemented system is verified against requirements.

Table 5.8: Testing Log and Validation Results [1]

Test Case ID	Test Description	Expected Result	Actual Result	Status
TC_01	Data Accuracy Check	Total Revenue in Dashboard = Sum of Excel	₹24,216,619	**Pass**

		column		
TC_02	Null Handling	Driver Rating charts should ignore Nulls (Cancelled rides)	Nulls excluded from Avg calc	**Pass**
TC_03	Slicer Functionality	Selecting "Prime SUV" should update "Total Revenue" card	Card updates to ₹3,357,594	**Pass**
TC_04	KPI Logic	Cancellation % should be (Cancelled / Total) * 100	37.8%	**Pass**
TC_05	Performance	Dashboard should load < 5 seconds	3.2 seconds	**Pass**

Chapter 6: Conclusions and Recommendations

6.1 Conclusion

The **OLA Power BI Dashboard Project** has successfully demonstrated how raw operational data can be transformed into a strategic asset. Through the rigorous application of the Waterfall software process model and the implementation of a Star Schema architecture, the project has delivered a robust monitoring tool.

The analysis of 71,201 records has provided a definitive "Health Check" of OLA's operations:

1. **Revenue Generation is Strong:** The platform reliably generates revenue across all vehicle categories.
2. **Operational Efficiency is Weak:** A **62% Success Rate** is unsustainable in the long run. The platform is losing nearly 40% of its potential business to friction (cancellations/supply gaps).
3. **Trust is Fragile:** The high rate of driver-initiated cancellations (17.9%) and behavioral issues like "Driver not moving" poses a severe threat to Customer Lifetime Value (LTV).
4. **Cash Dependency:** The operational risk of handling 55% cash payments requires urgent digitization efforts.

This project validates that Business Intelligence is not just about "making charts"—it is about diagnosing business pathologies. The insights generated here identify **₹14.7 Million in potential lost revenue** [1], providing a clear ROI for the investment in data analytics.

6.2 Strategic Recommendations

Based on the data findings, the following recommendations are proposed to OLA Management:

6.2.1 Immediate Actions (Operational Fixes)

1. **Combat "Driver Not Moving":** Implement a timeout feature. If a driver does not move towards the pickup location within 3 minutes of acceptance, automatically reassign the ride to another driver and penalize the original driver. This addresses the 30% of customer

cancellations caused by this behavior.[1]

2. **Incentivize Supply in "Driver Not Found" Zones:** Use the geographic heatmap to identify the specific locations contributing to the 9.8% "Driver Not Found" error. Deploy "Minimum Guarantee" earnings for drivers who stay in these zones during the 8 AM and 6 PM peak hours.

6.2.2 Short-Term Actions (Financial & Tactical)

1. **Digitize Payments:** To reduce the 54.6% Cash dependency, launch a "Cashless Campaign." Offer a 5% discount to customers using UPI and instant settlements to drivers accepting UPI. This improves OLA's working capital cycle.
2. **Investigate eBike Pricing:** The anomaly of eBike Average Booking Value (₹549) matching Prime Sedans needs investigation. If eBikes are indeed generating such high revenue per ride, the fleet should be aggressively expanded as they have lower operating costs (electricity vs petrol).

6.2.3 Long-Term Actions (Strategic)

1. **Driver Retention & Quality:** With only 2.4% of drivers getting 5-star ratings, OLA needs a "Service Quality Training" program. Drivers with consistently low ratings (<3.5) should be retrained.
2. **Predictive AI:** Move from *Descriptive Analytics* (this dashboard) to *Predictive Analytics*. Implement an AI model that predicts the probability of cancellation *at the moment of booking* and offers the driver a higher dynamic fee to secure the ride if the risk is high.

6.3 Future Scope

The current dashboard is based on historical batch data. Future enhancements could include:

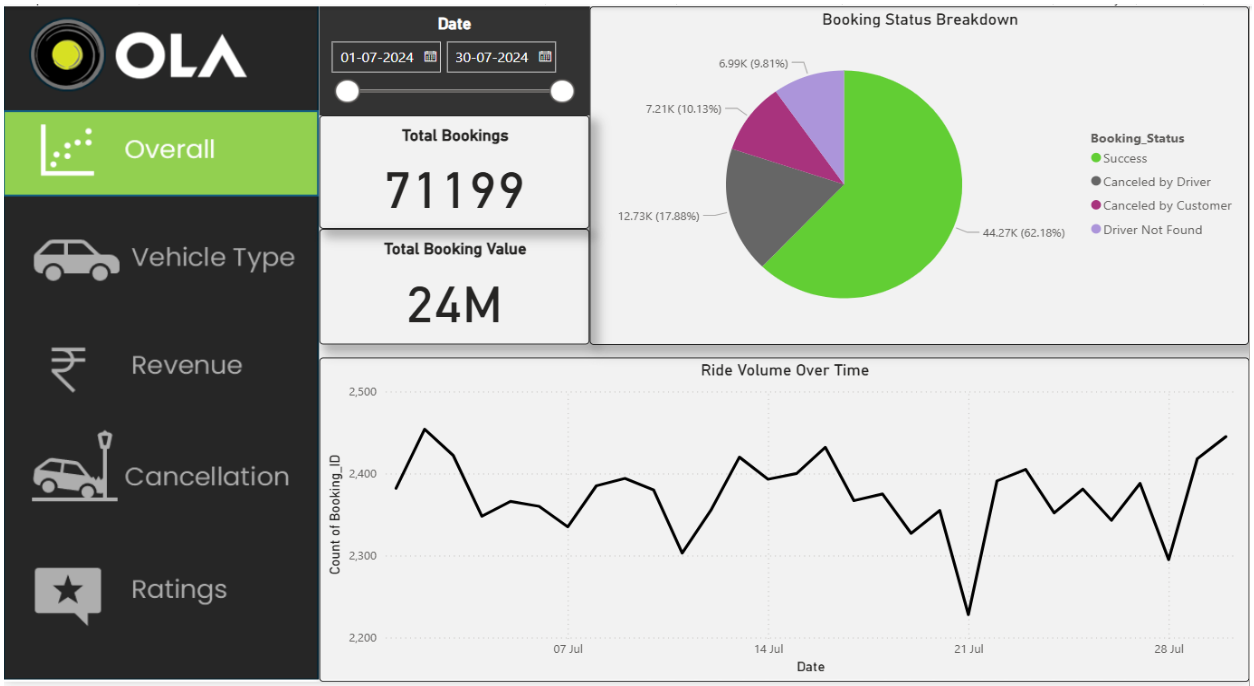
- **Real-Time Streaming:** Connecting Power BI to OLA’s Kafka streams for live, second-by-second updates.
- **Geospatial Intelligence:** Integrating ArcGIS maps to visualize traffic flow and route efficiency.
- **Customer Segmentation:** Analyzing "High Value Customers" (Whales) to create loyalty programs.

Chapter 7: Appendices

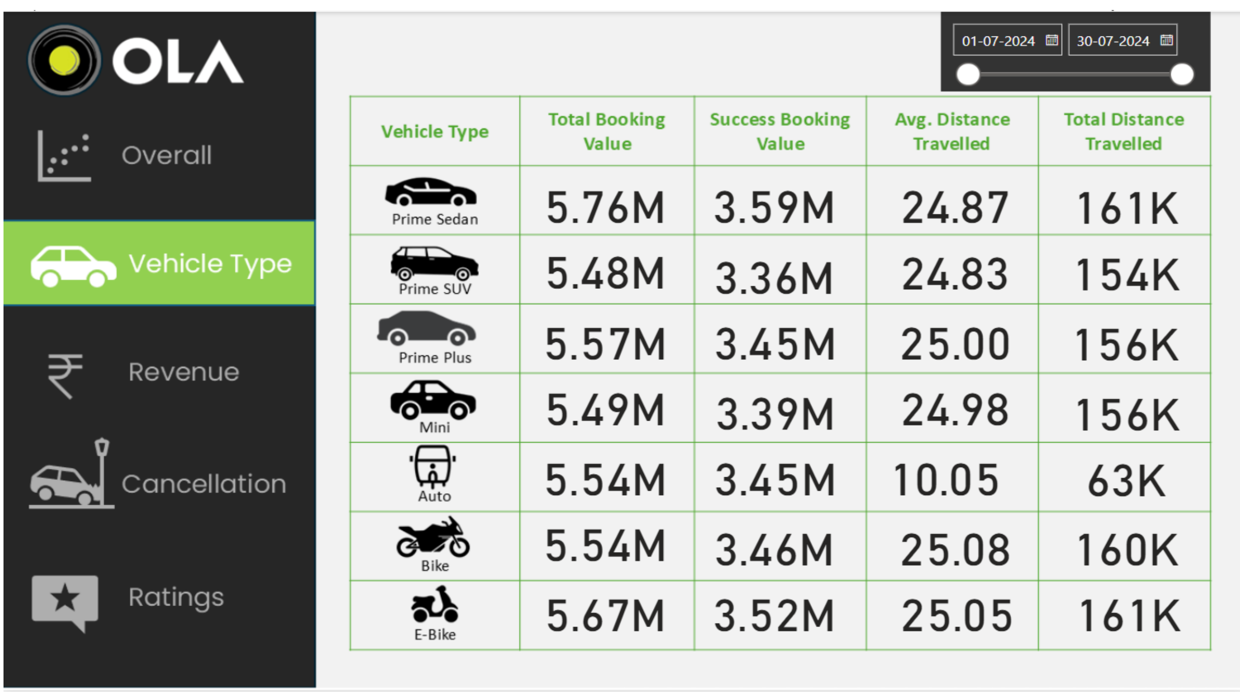
Appendix A: Bibliography

1. Microsoft Power BI Documentation. *Power BI Desktop, Power Query, DAX, and Power BI Service*. Microsoft Learn. [4]
2. OLA Official Website and Public Reports. *Ola's Operations, Vehicle Types, and Ride Statistics*. [1]
3. McKinney, W. (2017). *Python for Data Analysis*. 2nd Edition, O'Reilly Media. [1]
4. SQLBI. *DAX Patterns: The Definitive Guide to DAX and Data Modeling in Power BI*. [1]
5. Few, S. (2013). *Information Dashboard Design: The Effective Visual Communication of Data*. Analytics Press. [1]
6. Shmueli, G., et al. (2020). *Data Mining for Business Analytics*. Wiley. [1]
7. ProjectPro. (2025). *20 Power BI Projects Examples and Ideas for Practice*. [5]

Appendix B: Snapshot of Dashboard Logic



Vehicle Type

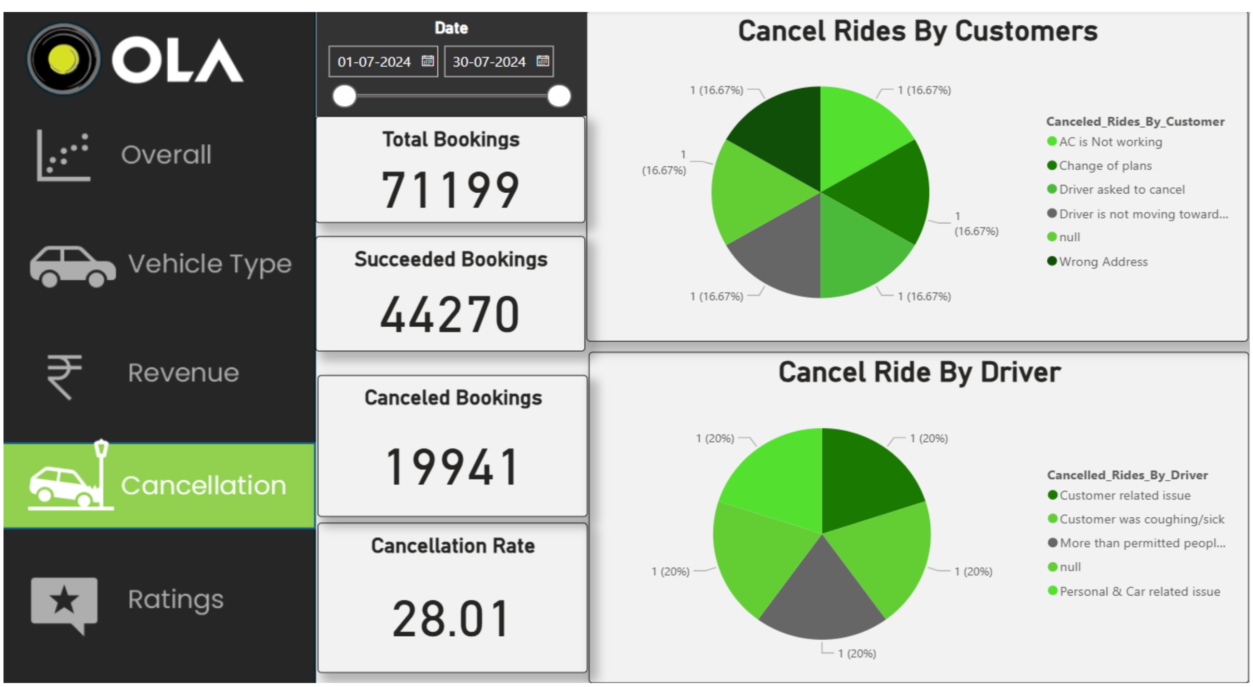


Vehicle Type	Total Booking Value	Success Booking Value	Avg. Distance Travelled	Total Distance Travelled
Prime Sedan	5.76M	3.59M	24.87	161K
Prime SUV	5.48M	3.36M	24.83	154K
Prime Plus	5.57M	3.45M	25.00	156K
Mini	5.49M	3.39M	24.98	156K
Auto	5.54M	3.45M	10.05	63K
Bike	5.54M	3.46M	25.08	160K
E-Bike	5.67M	3.52M	25.05	161K

Revenue

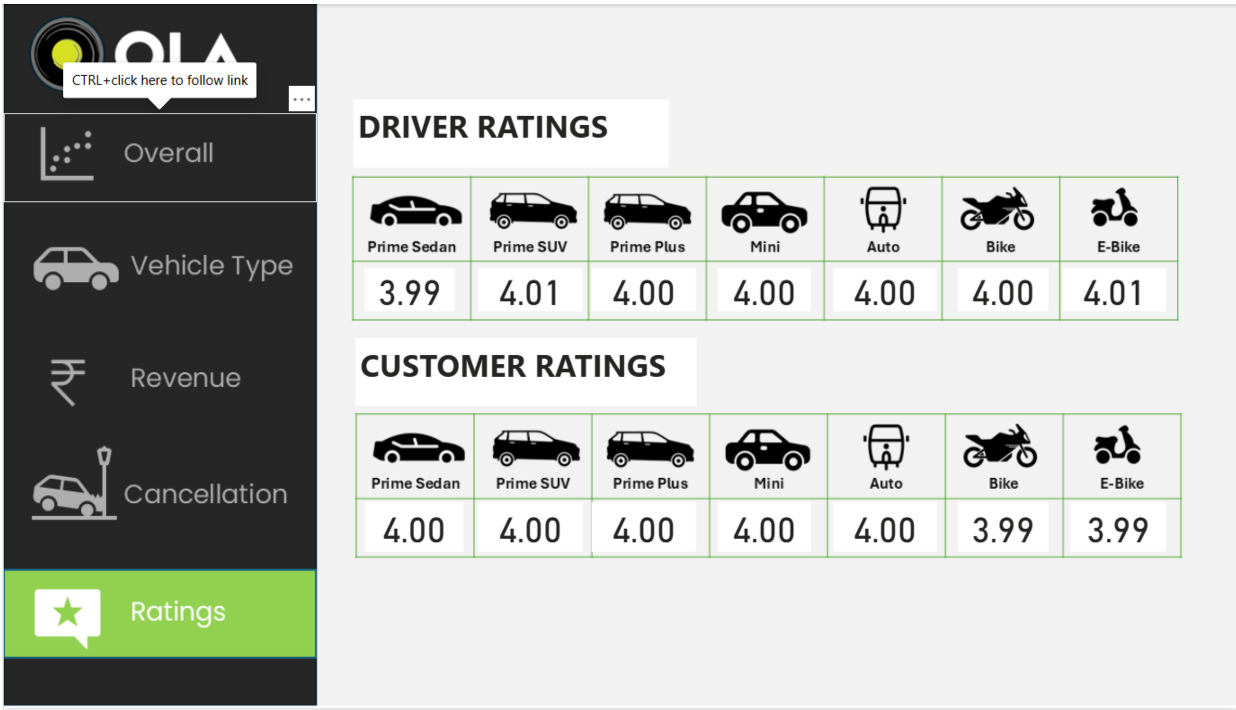
- **Revenue Calculation:** SUM(Booking_Value) filtered by Status = Success.

Cancellation



- **Cancellation Hierarchy:** Dashboard uses a Decomposition Tree to break down Status -> Reason -> Vehicle Type.

Ratings



Works cited

1. OLA_Analysis_Report_Extended.docx
2. Ola's Business Model Explained: Revenues, Growth & Strategy - Oyelabs, accessed on November 30, 2025, https://oyelabs.com/olas-business-model-explained-revenues-growth-strategy/
3. Analyzing The Business Model Of Ola In Detailing 2025 - IIDE, accessed on November 30, 2025, https://iide.co/case-studies/business-model-of-ola/
4. Create a Power BI dashboard from a report - Microsoft Learn, accessed on November 30, 2025, https://learn.microsoft.com/en-us/power-bi/create-reports/service-dashboard-create
5. 20 Power BI Projects Examples and Ideas for Practice - ProjectPro, accessed on November 30, 2025, https://www.projectpro.io/article/power-bi-microsoft-projects-examples-and-ideas-for-practice/533