

01a: Data Analysis Report

A comprehensive analysis using Python, SQL (MySQL), Power BI, and Excel to drive operational excellence





Project Objective

Comprehensive analysis of booking activity, customer behavior, driver performance, payment trends, and cancellation patterns to optimize operations and revenue.



Identify Cancellation Drivers

Analyze root causes of high cancellation rates impacting customer experience



Booking & Vehicle Trends

Understand demand patterns and vehicle performance metrics



Operational Bottlenecks

Pinpoint "Driver Not Found" incidents and supply gaps



Revenue Optimization

Evaluate ratings, payment preferences, and revenue opportunities

Dataset Overview

Dataset Scale

103K

Total Records

Comprehensive booking data

21

Data Columns


Detailed attributes

16.5MB

Memory Size


Efficient dataset

Key Variables Captured




Booking Details

Date, Time, ID, Status tracking




Customer Metrics

IDs, Ratings, Behavior patterns




Driver Performance

Ratings, Cancellations, Availability



Ride Intelligence

Vehicle type, Distance, Locations



Financial Data

Value, Payment methods, Revenue

Unique Data Points

- 4 Booking Status types
- 7 Vehicle Types
- 50 Pickup & Drop Locations each
- 4 Payment Methods
- 21 distinct Driver & Customer Rating levels

Data Quality Challenges

Significant missing values identified in:

- V_TAT, C_TAT metrics
- Driver & Customer Ratings
- Cancellation Reasons
- Incomplete Ride explanations

Addressed through systematic cleaning and logical filtering



Data Cleaning & Preprocessing

Rigorous preparation using Python (Pandas) and Power BI to ensure analytical accuracy

1

Structure Cleanup

Removed blank columns, validated ID uniqueness, checked nulls

2

Format Standardization

Unified date/time formats, cleaned text with TRIM/CLEAN

3

Data Type Conversion

Converted numeric fields, normalized categorical labels

4

Validation Complete

Removed spaces, ensured consistency for accurate insights



Result: Clean, consistent dataset ready for accurate SQL queries, Python analysis, and Power BI visualizations



Analytical Methodology

Python: Exploratory Analysis

Imported dataset and reviewed structure using `.info()`, `.describe()`, `.nunique()`

Evaluated missing data distribution and performed initial transformations

SQL: Deep-Dive Queries

Executed advanced queries to answer critical business questions:

- Successful booking patterns
- Average ride distance by vehicle
- Top customer identification
- Driver cancellation analysis
- Revenue from completed rides
- Incomplete ride root causes

Power BI: Interactive Dashboard

Visualizations created: Line charts (daily trends), Donut charts (status & payments), Column charts (vehicle performance, ratings), Tables (revenue, incomplete rides), Date slicers

DAX measures: Cancellation rate, Total bookings, Success vs. cancelled comparison

This multi-tool approach delivered a complete 360° understanding of platform performance



Key Insights

Data-driven findings that reveal opportunities for operational improvement and revenue growth

Made with **GAMMA**

A. Booking Trends: Stable Demand



Predictable Demand

Consistent booking patterns indicate stable market demand without major volatility or unexpected fluctuations

Strategic Advantage

Predictability enables better resource planning and driver allocation strategies

July Performance Overview

Daily bookings demonstrated remarkable stability throughout July, ranging between 3,200–3,400 bookings per day

3,432

Peak Day

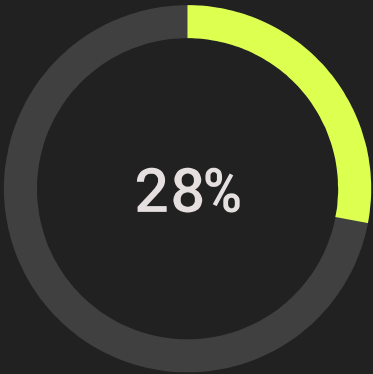
July 30 highest volume

3,072

Lowest Day

July 31 minimum

B. Booking Status: Critical Cancellation Challenge



Cancellation Rate

Nearly 3× industry benchmark of 10%

Cancellation Breakdown

- Driver Cancellations: 18K
Highest contributor to cancellation rate
- Driver Not Found: 10K
Strong indicator of supply-demand imbalance requiring immediate attention

C. Vehicle Type Performance Analysis

Prime Sedan

Top Performer

- Highest total booking value
- Highest successful trip value
- Greatest total distance traveled

E-Bike

Distance Leader

Highest average ride distance per trip, ideal for longer commutes

Prime Plus

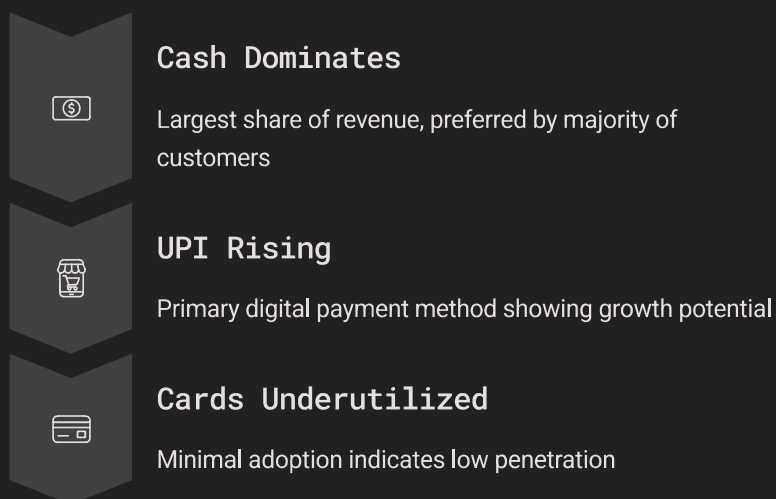
Customer Favorite

Highest customer rating score, demonstrating superior service quality

D. Payment Method Insights



Payment Preferences



Strategic Implication

Low digital payment adoption suggests opportunity for:

- Enhanced digital payment incentives
- User education campaigns
- Simplified card linking process
- UPI promotion strategies

