UPI TRANSACTION ANALYSIS REPORT

Objective

The main objective of this project is to analyze UPI transaction data to gain insights into user behavior, regional performance, transaction efficiency, and merchant engagement.The dashboard will help stakeholders understand transaction trends, identify operational bottlenecks, and improve decision-making through data-driven insights.

Purpose

To provide a business intelligence dashboard that enables:

- Monitoring of daily and monthly transaction performance.
- Identifying top-performing banks and merchants.
- Understanding user behavior by region, device, and age group.
- Detecting and analyzing fraudulent transactions.
- Supporting data-driven policy and infrastructure decisions.

Dataset Summary

The dataset represents UPI transaction details collected over a specific time period, containing both successful and failed transaction records.

- Total Records: Approximately 250000 transaction entries.
- Total Columns: 17 attributes after data cleaning and normalization.
- Type of Data: Transactional, time-based, categorical, and numerical.
- Format: CSV imported into PostgreSQL database for normalization.
- Source: Synthetic dataset prepared for analytical and academic use.

Key Cleaning and Transformation Steps:

- Removed duplicates and invalid transaction IDs.
- Standardized all column names to lowercase for database compatibility.
- Replaced missing or inconsistent values with NULLs or “Unknown.”
- Converted timestamps into date format (DD-MM-YYYY).
- Normalized categorical fields (state, bank, merchant, etc.) into dimension tables.
- Created calculated columns such as hour_of_day and is_weekend for trend analysis.

The final dataset is clean, relational, and structured for efficient querying and Power BI modeling.

Column-wise Assessment Summary

Column Name	Description	Cleaning/Transformation Applied
transaction_id	Unique identifier for each UPI transaction	Removed duplicates, filled missing IDs
transaction_date	Date of transaction	Standardized to DATE type
transaction_type	Type of transaction (e.g., P2P, P2M)	Text trimmed, categorized
merchant_category	Type of merchant (e.g., Retail, Food, Bills)	Normalized into merchant_categories table
amount_inr	Transaction amount in INR	Converted to numeric
transaction_status	Indicates Success, Failed, or Pending	Split into lookup table for normalization
state	State	Linked to states dimension

device_type	Device used (Android/iOS/Web)	Standardized text
network_type	Connection type (4G, Wi-Fi)	Cleaned for uniform values
fraud_flag	Indicates suspected fraudulent transaction	Converted to boolean (TRUE/FALSE)
sender_age_group	Categorical range (18–25, 26–35, etc.)	Normalized in age_groups table
hour_of_day	Extracted from timestamp	Derived using date-time split
day_of_week	Weekday derived from transaction_date	Used for time-based trend visuals

Data Model Overview

The project follows a Star Schema Model to optimize performance and ensure data scalability.

Central Table (Fact Table):

- fact_upi_transactions – Stores the main transactional details (amount, date, fraud_flag, transaction_status, etc.).

Dimension Tables:

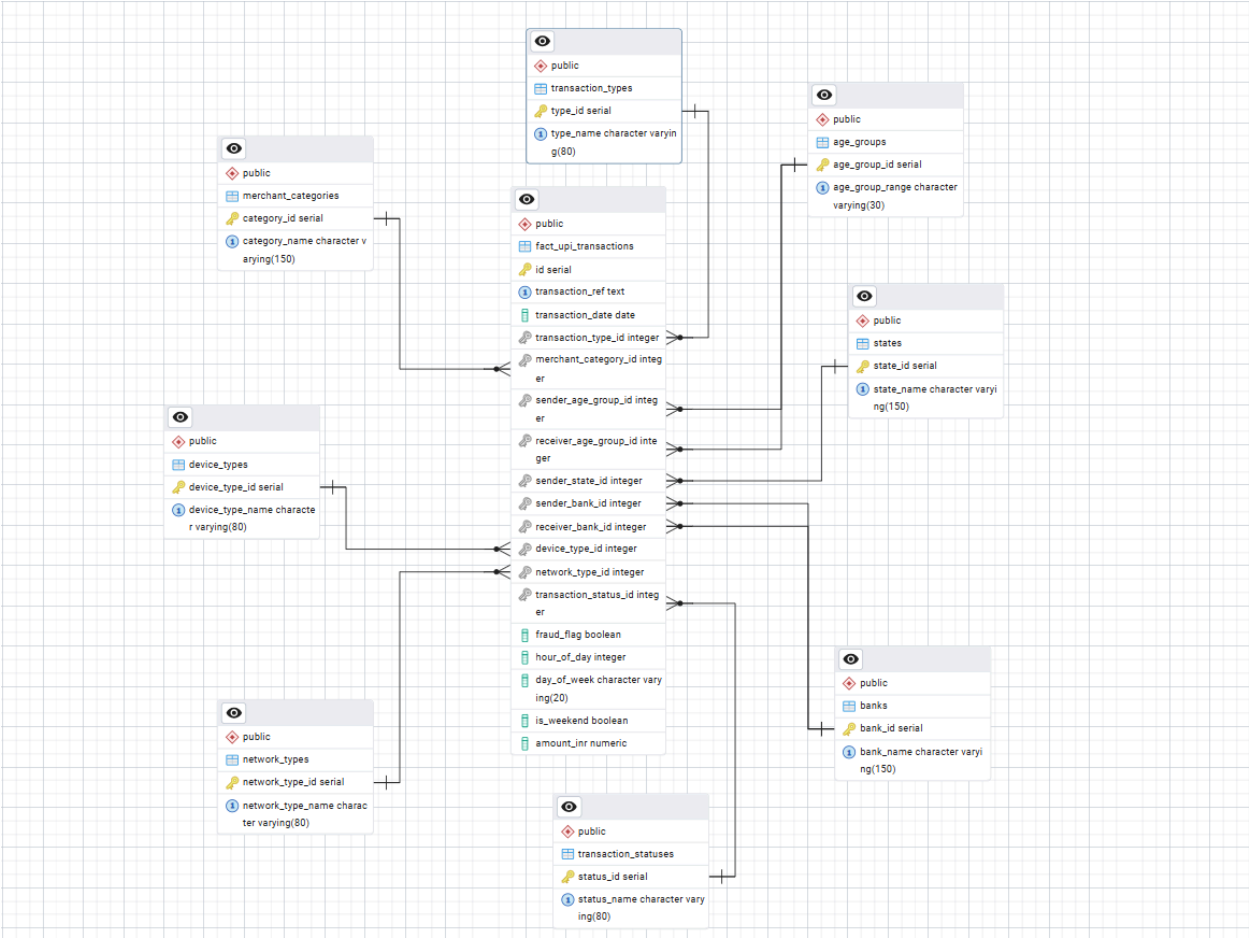
- banks: Contains bank names and IDs for both sender and receiver.
- merchant_categories: Stores merchant types.
- transaction_statuses: Defines transaction result categories (Success, Failed, Pending).
- states: Contains state names and regional data.
- age_groups: Defines sender and receiver age group ranges.
- network_types & devices: Store device and network categories.

Relationships:

- All dimension tables have a one-to-many relationship with the fact_upi_transactions table.
- Relationships are maintained through surrogate primary and foreign keys.

Calculated Fields & Measures:

- Measures like *Total Transactions*, *Total Amount*, *Success Rate %*, and *Fraud Rate %* are derived using DAX.
- Calculated columns include *hour_of_day*, *day_of_week* for temporal insights.



Analysis & Insights

1. Transaction Performance:

- The total transaction volume shows a steady upward trend over time.
- Success rates remain consistently high, averaging between 95–98%.
- Transaction failures are slightly more common during weekends and peak evening hours (6 PM – 9 PM).

3. Regional Performance:

- Maharashtra, Karnataka, and Tamil Nadu are the leading states in both transaction volume and amount.
- Smaller states like Goa and Sikkim have low activity but higher average transaction values.

4. User & Device Behavior:

- Users aged between 18–30 represent the largest transaction group.
- Android devices dominate UPI usage, accounting for over 70% of total volume.
- Users tend to perform higher-value transactions through Wi-Fi connections.

5. Fraud Trends:

- Fraudulent transactions are rare (~1.3%) but show concentration in certain states.
- Fraud cases are slightly higher in transactions between different banks (cross-bank transfers).

6. Temporal Patterns:

- Peak transaction hours are between 5 PM – 9 PM.
- Weekends see higher volume, especially Saturdays.
- End-of-month spikes suggest bill payments or salary-based transfers.

Conclusions

- UPI transactions demonstrate excellent reliability with very low failure rates.
- Fraud cases are minimal, but monitoring mechanisms should remain active, especially for inter-bank transactions.
- High engagement from younger users and Android devices highlights strong mobile adoption.
- The consistent performance across banks and merchants supports confidence in the UPI ecosystem.

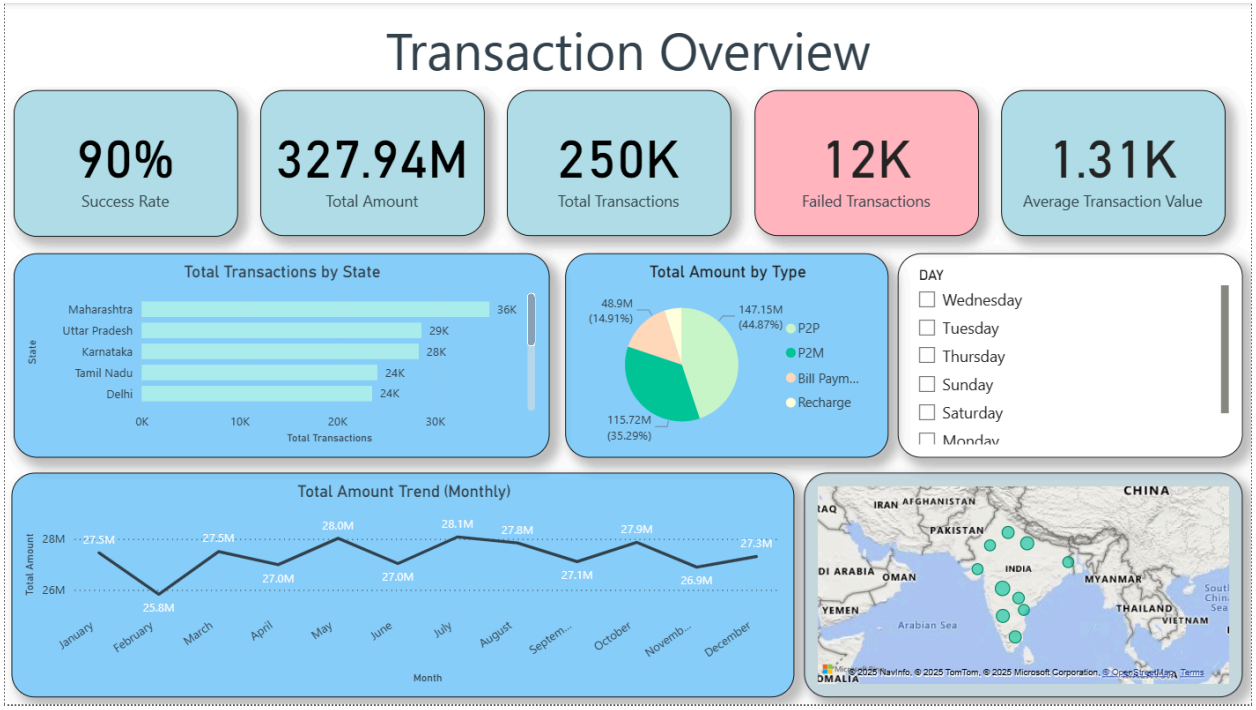
Overall, UPI continues to serve as a benchmark for digital payment infrastructure — scalable, secure, and widely adopted.

Recommendations

- Enhance Fraud Monitoring:
  Use predictive analytics or AI-based scoring to identify potential fraudulent patterns in real time.
- Improve Bank Coordination:
  Encourage collaboration between banks with slightly higher failure rates to optimize transaction reliability.
- Optimize Peak Load Handling:
  Expand infrastructure to handle high transaction load during weekends and evening peaks.
- Promote Awareness Among Users:
  Target campaigns toward older users or rural areas with low digital literacy.
- Expand Analytical Dashboards:
  Include predictive KPIs such as “Expected Success Rate” and “Transaction Growth Forecast.”

Dashboard Overview

Transaction Overview:



Purpose

Gives a high-level summary of total UPI activity, transaction growth, and key success indicators.

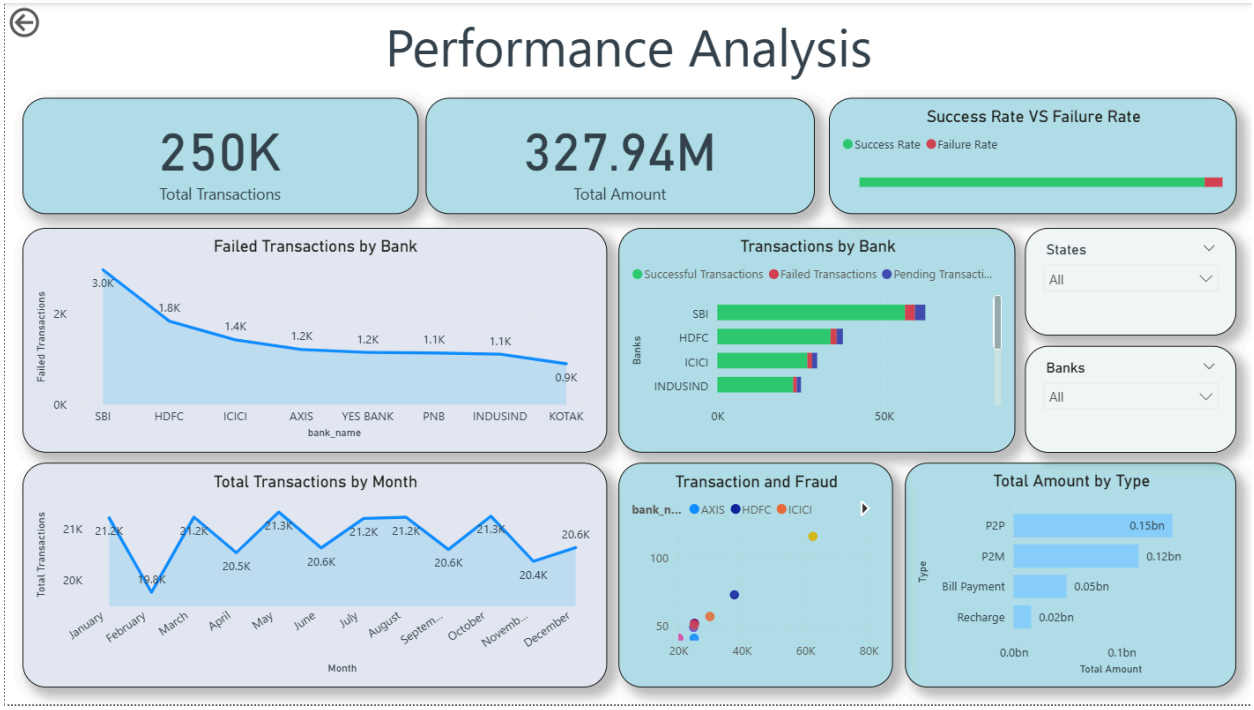
Visuals Used:

- KPI Cards: Total Transactions, Total Amount, Success Rate (%), Failed Transaction, Average Transaction Value
- Line Chart: Amount trend over time (Monthly).
- Pie Chart: Total Amount by types (P2P, P2M, etc.).
- Bar Chart: Total Transaction by state.
- Map: State.
- Slicer: Day.
- Tooltip: Amount trend over time(Day)

Insights Provided:

- Monitors transaction growth trends.
- Identifies days or months with unusual activity or spikes.
- Helps management understand the overall health of the UPI network.

Performance Analysis:



Purpose

Focuses on bank-level and system-level efficiency by analyzing success, failure, and pending transaction rates.

Visuals Used:

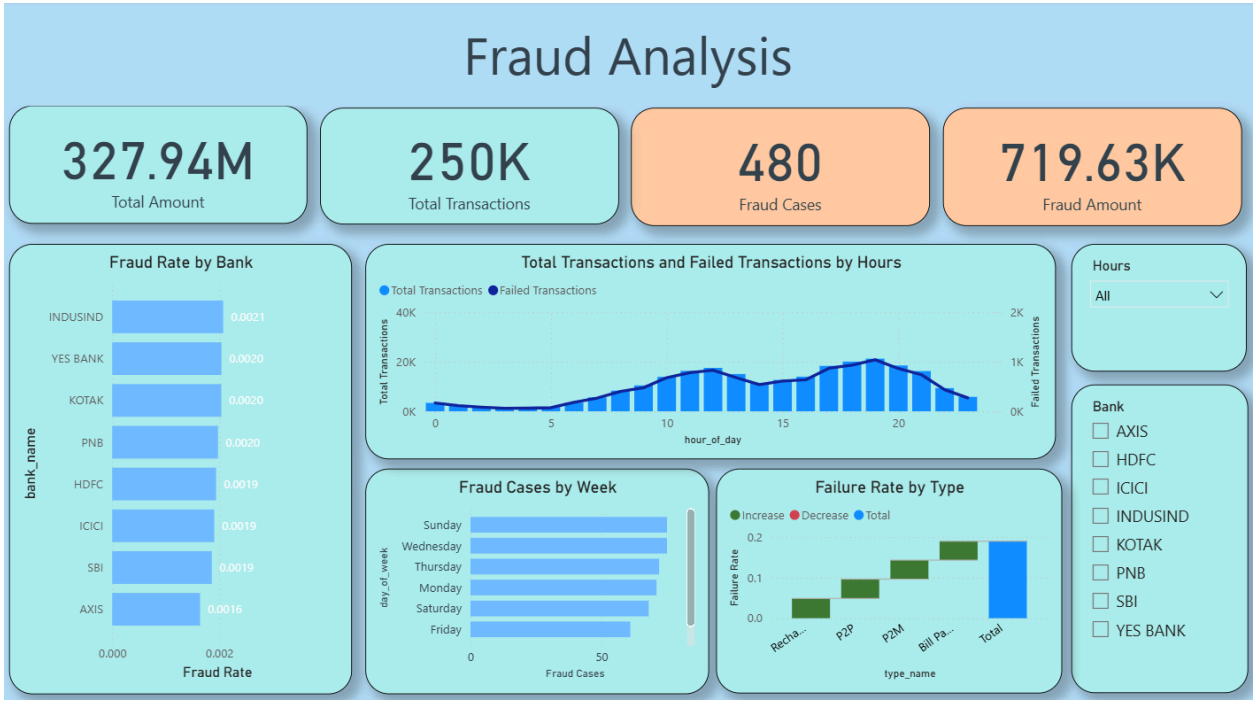
- KPI Card: Total Transaction, Total Amount
- Line Chart: Failed Transactions by Bank, Total Transactions by Month.
- Bar Chart: Transaction by Bank, Total Amount by Type.
- Scatter: Transaction and Fraud.

- Slicer: State, Bank.

Insights Provided:

- Identifies high-performing and low-performing banks.
- Reveals patterns in transaction failures.
- Helps improve service reliability and user satisfaction.

Fraud Analysis:



Purpose

Explores transaction behavior across different Bank to identify key contributors and business opportunities.

Visuals Used:

- KPI: Total Transaction, Total Amount, Fraud Cases, Fraud Amount.
- Bar Chart:Fraud Rate by bank, Fraud Cases by Week.
- Water fall chart: Failure rate by type.
- Line and Column Chart: Total transaction and Failed by hour.

- Slicer: Hours, Bank.

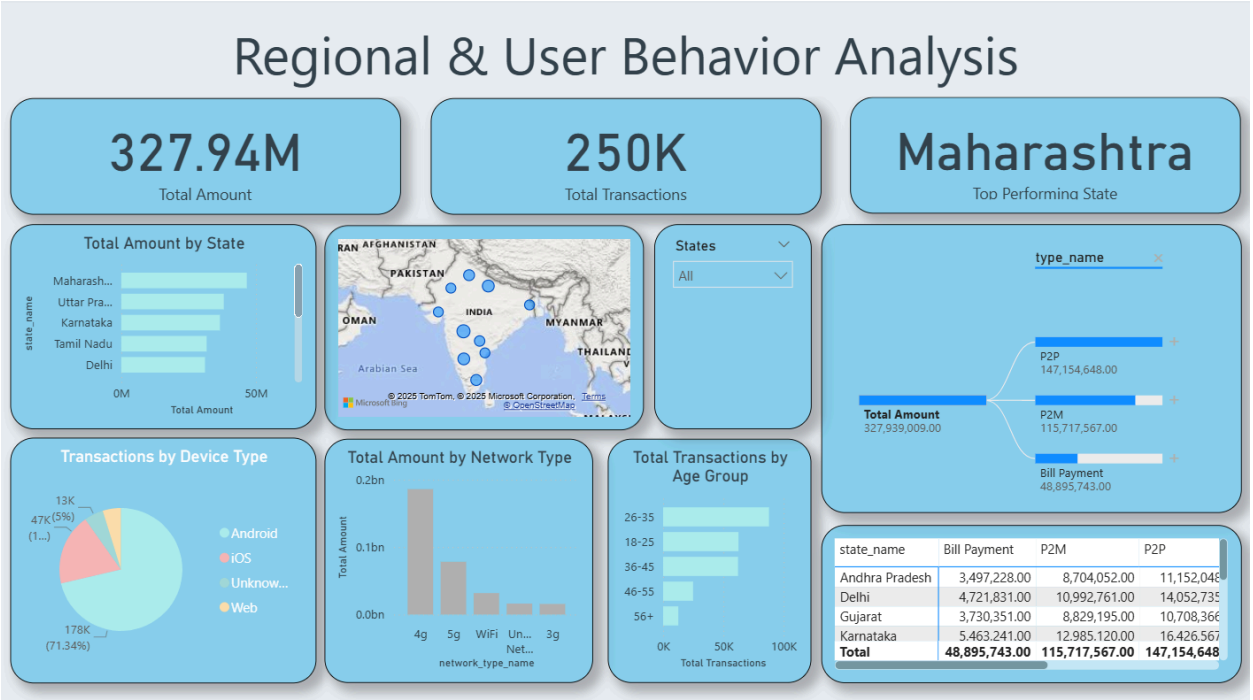
Insights Provided:

Fraud Risk is Highly Concentrated and Predictable by Time

Failure and Fraud are Distinct Problems that Require Separate Bank Strategies

Entertainment and Transport are High-Risk Categories with Low Volume

Regional & User Behavior Analysis:



Purpose

Analyzes transactions based on sender's state, device type, network, and age group for behavioral understanding.

Visuals Used:

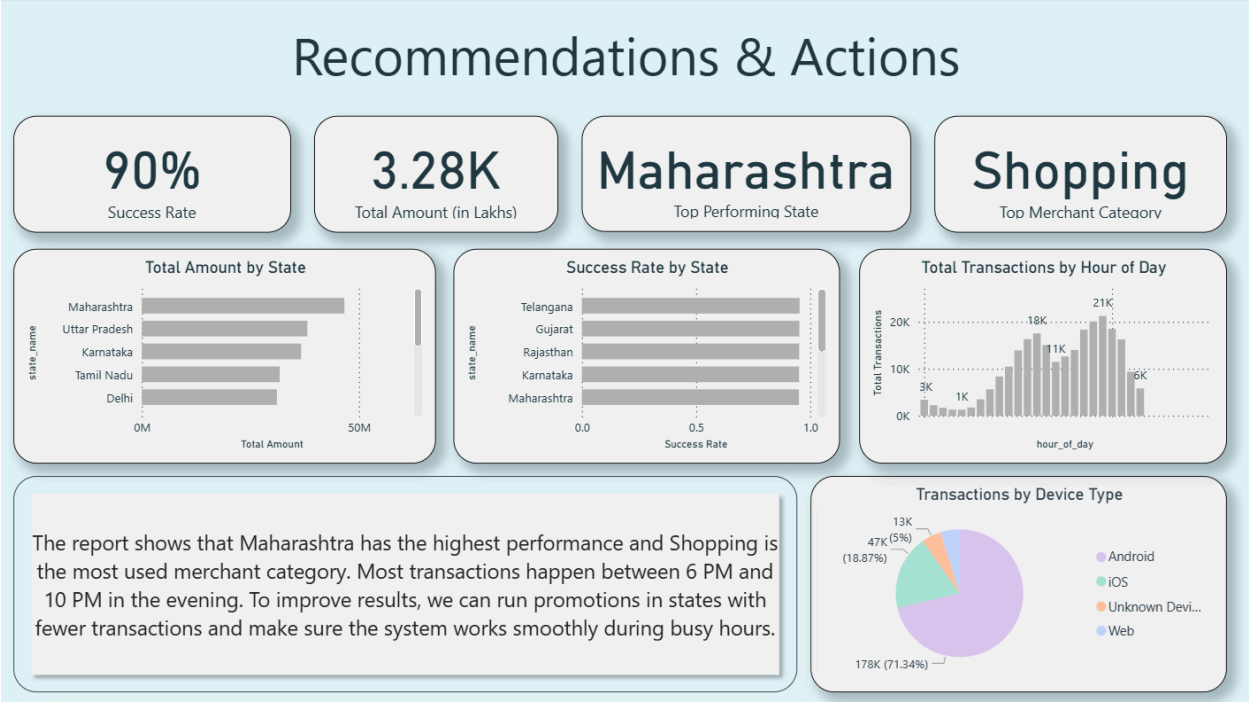
- Map Visual: State-wise transaction count.
- Bar Chart: Total amount by state, Total Transactions by Age Group.
- Pie Chart: Transactions by Device Type.
- Column Chart: Total Amount by Network Type.

- Matrics
- KPI: Total amount, Total transaction, Top performing state
- Decomposition chart
- Slicer: State.

Insights Provided:

- The Core Customer is the Young to Middle-Aged Adult.
- Maharashtra Dominates Both Volume and Value.
- Mobile Devices (Android & iOS) Drive the Vast Majority of Transactions.

Recommendation & Summary Page:



Purpose:
Summarizes insights, conclusions, and suggested business actions based on the overall data analysis.

Visuals Used:

- Card Visuals: Repeated key KPIs for quick recap.

- Text Box: Summary of key findings and recommendations.
- Infographic-Style Icons: Highlight fraud reduction, growth opportunities, etc.
- Navigation Buttons: Link back to Home Page or other dashboards.

Insights Provided:

- Offers actionable recommendations directly from data insights.
- Provides executives a concise view of what decisions should be made.
- Ideal for presentations and business meetings.

Filters & Slicers Used Across Dashboard

- Day – To analyze specific day periods.
- States – For regional comparison.
- Bank – analyze specific bank.
- Hours – analyze specific hours.

Notes / Limitations

- The dataset is sample-based and may not reflect complete real-world UPI data.
- Fraud flag is assumed based on probability, not confirmed by real investigation.
- Some smaller states and merchant categories have limited data representation.