

# Hotfoot Technology Solutions Hackathon Project Report



## Macro-Level Anomaly Detection in India's UPI Ecosystem: Growth, Seasonality, and Market Participation

Prepared by

*Sumiran Rai,*

*MSc Data Science & Computing,*

*Sri Sathya Sai Institute of Higher Learning*

## Executive Report

This analysis reviews macro-level patterns and anomalies in India's UPI ecosystem using aggregated RBI and NPCI data. The purpose is to identify whether unusual movements in volumes, values, or bank-level shares reflect genuine business changes or are simply the result of reporting and scale effects.

Over the full period studied, UPI shows steady and sustained growth, reaching record levels in 2025. Seasonal behaviour remains consistent year after year, with predictable spikes during festive months and softer activity around year-end. After accounting for these patterns, most daily movements fall well within expected ranges.

Between June 2024 and September 2025, 20 notable anomalies were identified. Roughly half are best explained by data or reporting effects, such as low-base comparisons or temporary classification changes. The remaining cases reflect real operational shifts, mainly at the individual bank or PSP level. Importantly, these events are isolated and do not persist over long periods.

Large growth spikes are concentrated in smaller banks, cooperative banks, and niche payment service providers. In these cases, very high percentage growth is driven by small starting volumes or one-time onboarding events rather than sustained expansion. These movements are statistically extreme but limited in practical impact on the overall system.

Higher-impact anomalies are observed only in market-share changes among a few large banks, most notably Yes Bank, SBI, and Axis Bank. These shifts are consistent with competitive repositioning, particularly on the merchant and payee side, and do not indicate stress or instability in the payment system. Measures of market concentration remain within competitive ranges throughout the period.

Looking at anomalies over time, their accumulation tracks the rapid increase in UPI scale rather than any breakdown in system behaviour. Earlier anomalies are more operational, while later ones increasingly reflect reporting noise at very high transaction volumes.

**In summary**, the UPI ecosystem appears stable and well-functioning at the macro level. The anomalies identified are localized, explainable, and consistent with known business and reporting dynamics. There is no evidence of systemic risk or abnormal concentration during the period studied.

# I. Project Description

## Problem Statement

India's Unified Payments Interface (UPI) has seen sustained growth in transaction volumes, transaction values, and participation across banks and payment applications. While overall growth is expected, certain periods may display abnormal patterns that warrant closer examination.

The objective of this analysis is to model normal macro-level UPI behaviour and identify statistically significant deviations using only aggregated, publicly available data.

## Objectives

### 1. Modelling Normal Behaviour

The analysis will establish baseline system-wide behaviour with respect to:

- Growth rates of UPI transaction volumes and values
- Seasonal and calendar-driven patterns
- Distribution and concentration of transactions across banks and UPI apps

### 2. Detection of Deviations

The study will flag periods or entities exhibiting deviations from expected behaviour, including:

- Sudden increases or decreases in total UPI transaction volume or value inconsistent with historical trends
- Abrupt or repeated changes in a bank's or app's market share relative to peers
- Periods where observed system growth diverges from expected seasonal or long-term patterns

## Interpretation

For each identified deviation, the analysis will:

- Classify the event as either:
- a potential business or operational signal, or
- a likely data-quality or reporting artefact
- Provide a brief explanation grounded in publicly available, aggregate-level information
- Avoid individual transaction-level or fraud-detection analysis

## Data Sources

All data used in this analysis will be sourced exclusively from the following MOSPI publications:

1. Datasets and Registries in India (Year 2022)  
[https://www.mospi.gov.in/sites/default/files/publication\\_reports/Datasets%20and%20Registries%20in%20India%202022.pdf](https://www.mospi.gov.in/sites/default/files/publication_reports/Datasets%20and%20Registries%20in%20India%202022.pdf)
2. Compendium of Datasets and Registries in India (Year 2023)  
[https://www.mospi.gov.in/sites/default/files/Compendium\\_of\\_Datasets\\_and\\_Registries\\_in\\_India\\_2023.pdf](https://www.mospi.gov.in/sites/default/files/Compendium_of_Datasets_and_Registries_in_India_2023.pdf)
3. Compendium of Datasets and Registries in India (Year 2024)  
[https://www.mospi.gov.in/sites/default/files/publication\\_reports/Compendium\\_of\\_Datasets\\_and\\_Registries\\_in\\_India\\_2024\\_1.pdf](https://www.mospi.gov.in/sites/default/files/publication_reports/Compendium_of_Datasets_and_Registries_in_India_2024_1.pdf)

## Evaluation Requirements

- A working hyperlink to the exact dataset downloaded from the MOSPI portal will be provided.
- The choice of dataset sample from the above sources will be part of the evaluation.

## Scope

### Included

- Macro-level UPI trends
- Aggregated system and entity-level data
- Trend, seasonality, and concentration analysis

### Excluded

- Individual transaction analysis
- User- or merchant-level behaviour
- Fraud detection

## II. Data Sources and Cleaned Data Description

### Overview of Data Sources

#### 4.1.28: Dataset: Payment Settlement Data

1.	Name of dataset/ indicator	Payment Settlement Data
2.	Source Ministry/ Department/ Organization	Reserve Bank of India
3.	Data/ indicators compiled are based on a <i>survey data, administrative data, multiple data sources, macro-aggregates or any other method</i>	Multiple Data Sources
4.	Themes/ Categories under which data is collated	Financial Sector-Payment Systems
5.	List of <i>key variables and their units of measurement (in case of datasets)</i>	Payment Transactions (RBI Operated (Real-time Gross Settlement (RTGS), National Electronic Fund Transfer (NEFT); National Payments Corporation of India (NPCI) Operated (Aadhaar Enabled Payment System (AePS),Unified Payments Interface (UPI),Immediate Payment Service (IMPS).National Automated Clearing House (NACH ),National Electronic Toll Collection (NETC),Bharat Bill Payment System (BBPS)); Card Network Operated (Credit Card, Debit Card, Prepaid Payment Instruments Card)); Cash Withdrawal (NPCI Operated (National Financial Switch (through ATMs)), AePS (through micro ATMs/Business Correspondents)); Settlement Systems (Clearing Corporation of India Limited Operated (Government Securities Clearing; Forex Clearing; Rupee Derivatives)) Units: Volume in lakh), Value in Rupees Crore
6.	List of <i>key variables used for computation and their units of measurement (in case of indicators)</i>	-
7.	Disaggregation level	All India
8.	Data collection method and tools used	-
9.	List of classifications used	-
10.	Data compilation method and tools used	Compiled from various sources
11.	Base period (in case of indicators)	-
12.	Periodicity/ frequency	Daily
13.	Time lag in dissemination from reference period	-
14.	Documentation on the dataset/ indicator	<a href="https://www.rbi.org.in/scripts/paymentsystems.aspx">https://www.rbi.org.in/scripts/paymentsystems.aspx</a>
15.	URL from where the data is available for download/ dissemination	<a href="https://www.rbi.org.in/scripts/paymentsystems.aspx">https://www.rbi.org.in/scripts/paymentsystems.aspx</a>
16.	Starting month/ year from which data is available	31-06-2020
17.	Format of dissemination	PDF; MS Excel

This analysis relies exclusively on publicly available, aggregated payment system datasets published by official Indian government institutions. No proprietary, confidential, or transaction-level data has been used at any stage.

Two primary datasets form the foundation of the study:

1. RBI Payment System Indicators (PSDDP) – Daily Level
2. NPCI Entity-wise Retail Payment Data – Monthly Level

Both datasets are published by the Reserve Bank of India (RBI) and compiled from regulated payment system operators, ensuring data reliability and consistency.

## Dataset 1: RBI Payment System Indicators (PSDDP) – Daily Level

The RBI Payment System Indicators (PSDDP) dataset provides daily, system-level transaction statistics across major retail and wholesale payment systems in India.

- Frequency: Daily
- Coverage Period: June 1, 2020 to November 30, 2025
- Number of Records: 2,009 daily observations
- Granularity: Aggregated payment-system-level data
- Publisher: Reserve Bank of India (RBI)

Source:

<https://rbidocs.rbi.org.in/rdocs/content/docs/PSDDP04062020.xlsx>

### Scope and Characteristics

Key characteristics of the dataset include:

- Covers 24+ payment and settlement systems
- Aggregated across all participating banks
- Represents system-wide settlement activity
- Contains no individual transaction, customer, or account-level information
- Suitable for macro trend, growth, and seasonality analysis

All figures represent total daily activity at the national level.

### Measurement Units

- Transaction Volume (\_vol): Number of transactions, measured in lakhs
- Transaction Value (\_val): Total transaction value, measured in ₹ crore

### Payment Systems Covered

The dataset includes transaction volume and value for the following major payment systems:

- UPI (Unified Payments Interface)
- IMPS (Immediate Payment Service)
- NEFT (National Electronic Funds Transfer)
- RTGS (Real Time Gross Settlement)

- AePS (Aadhaar Enabled Payment System)
- NACH (Credit and Debit)
- NETC (FASTag)
- BBPS (Bharat Bill Payment System)
- CTS (Cheque Truncation System)
- Credit Cards and Debit Cards (PoS and E-commerce)
- PPI (Prepaid Payment Instruments)
- NFS (ATM Network)
- Government Securities, Forex, and Rupee Derivatives

While the dataset spans multiple systems, UPI-related metrics form the primary focus for macro trend modeling and deviation detection in this study.

## Dataset 2: NPCI Entity-wise Retail Payment Data – Monthly Level

The NPCI Entity-wise Retail Payment dataset provides monthly, bank-level transaction statistics across major retail payment systems operated by NPCI and regulated by RBI.

- Frequency: Monthly
- Coverage Period Used: April 2024 to October 2025
- Number of Banks Covered: 692+ banks and financial institutions
- Granularity: Bank × Payment System × Operational Role
- Publisher: Reserve Bank of India (RBI)

Source:

<https://www.rbi.org.in/Scripts/EntityWiseRetailStatistics.aspx>

The dataset used in this analysis is a cleaned and harmonized version, consolidated from monthly RBI Excel releases with standardized column naming and consistent measurement units.

### Structure and Granularity

Each row in the dataset captures monthly aggregated activity for a specific combination of:

- Calendar month
- Bank or financial institution
- Payment system (e.g., UPI, IMPS, NETC, AePS, CTS, BBPS)
- Operational role within the payment system

This structure enables entity-level participation, concentration, and market share analysis without accessing individual transaction details.

### Measurement Units

- Transaction Volume: Reported in lakh transactions
- Transaction Value: Reported in ₹ crore

All values represent aggregate settlement totals, not individual transaction sizes.

### **Operational Role Interpretation**

The Role field captures the functional position of a bank within each payment system:

- UPI / IMPS:
  - *Remitter* – Sending bank
  - *Beneficiary* – Receiving bank
- NETC:
  - *Issuer*
  - *Acquirer*
- NFS / AePS:
  - Issuer and Acquirer roles for ATM and biometric transactions
- CTS:
  - Presenting
  - Drawee
- BBPS:
  - Customer Operating Unit
  - Biller Operating Unit

This role-based segmentation allows separation of sending-side and receiving-side dynamics, supporting more granular macro-level insights.

### **Relevance to the Problem Statement**

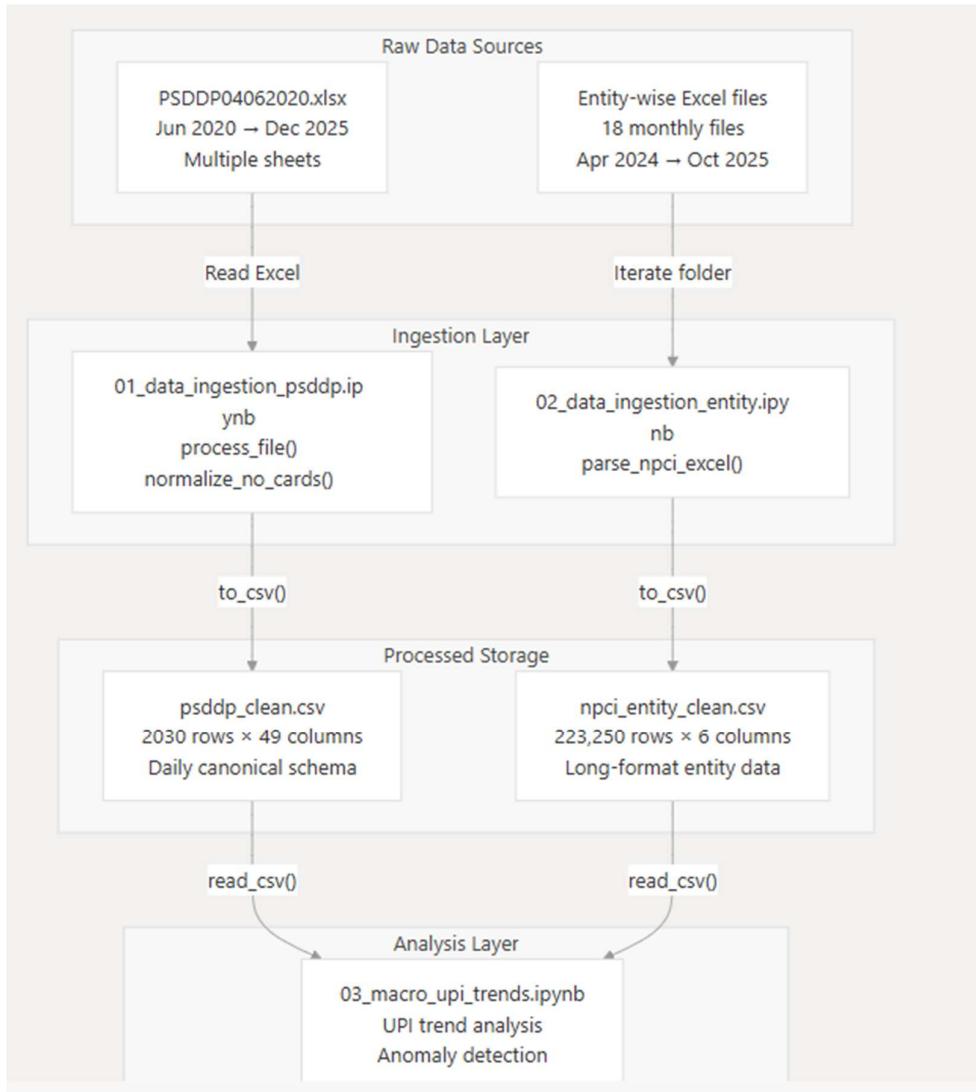
Together, the two datasets enable a multi-layered macro analysis:

- Daily PSDDP data supports:
  - Trend and growth modeling
  - Seasonality detection
  - System-wide anomaly identification
- Monthly entity-wise data supports:
  - Bank-level participation analysis
  - Market share and concentration assessment
  - Detection of structural or behavioral shifts among entities

Both datasets align strictly with the project scope by focusing on aggregated, public, and non-sensitive information.



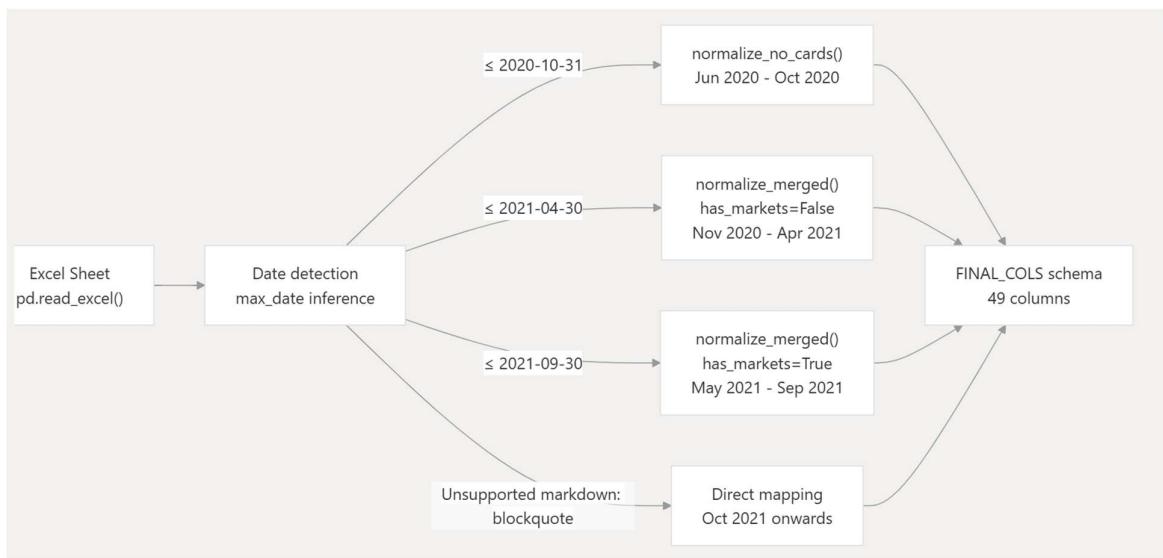
### III. Raw Data & Data Pre-processing



#### Pipeline Flow:

1. Raw Data Sources: Excel files published by RBI containing payment system statistics
2. Ingestion Layer: Jupyter notebooks implementing specialized normalization logic
3. Processed Storage: Clean CSV files with consistent schemas
4. Analysis Layer: Time-series analysis, visualization, and anomaly detection

## Normalization of PSDDP Daily Data



### Overview of the PSDDP Dataset

The PSDDP dataset used in this analysis consists of a single consolidated Excel file downloaded from the RBI portal, covering daily payment system indicators from June 2020 to November 2025.

While the dataset provides continuous daily coverage, the underlying reporting structure is not uniform across the entire period. Over time, the Reserve Bank of India has introduced new payment instruments and revised reporting formats, resulting in structural breaks that must be addressed before any longitudinal analysis.

### Rationale for Normalization

Normalization was required because the PSDDP dataset exhibits multiple structural changes over time, driven by:

- Phased introduction of new payment instruments
- Changes in how card-based transactions are reported
- Expansion of financial market transaction reporting

These changes are systematic and policy-driven, rather than random data quality issues. Without normalization, direct comparison of transaction volumes or values across periods would lead to misleading conclusions.

### Identification of Reporting Regimes

Based on the structure of the Excel sheets and the availability of payment instruments, four distinct reporting regimes were identified:

#### 1. June 2020 – October 2020

- No card-based payment data reported
- Coverage limited to core retail and interbank payment systems

#### 2. November 2020 – April 2021

- Card transactions reported as merged totals (PoS + E-commerce)
  - Financial market instruments not yet included
- 3. May 2021 – September 2021**
- Card transactions still reported in merged form
  - Financial market instruments (Government Securities, Forex, Rupee Derivatives) introduced
- 4. October 2021 onwards**
- Card transactions reported separately for PoS and E-commerce
  - All major payment and market instruments reported consistently

These regimes form the basis for regime-specific normalization logic.

#### **PSDDP File Processing Logic**

Each Excel sheet within the PSDDP file is processed independently using a standardized pipeline:

1. Removal of metadata rows preceding the actual data table
2. Cleaning and parsing of the date column into a consistent datetime format
3. Exclusion of summary and footnote rows (e.g., “Total”, “Notes”)
4. Inference of the reporting regime based on the maximum date present in the sheet
5. Application of the appropriate normalization strategy corresponding to the identified regime

Additional considerations during processing include:

- Certain instruments appear only in specific reporting periods
- Special placeholders such as "H" are used to denote withheld or non-disclosed values
- Missing values are often structural by design, not randomly missing observations

#### **Canonical Output Schema**

To enable consistent longitudinal analysis, all PSDDP data is normalized into a single canonical schema that remains fixed across all reporting periods.

Key characteristics of the canonical schema:

- Transaction volumes are expressed in lakhs
- Transaction values are expressed in ₹ crore
- Card transaction metrics are available in both:
  - Separate PoS and E-commerce fields (where reported)
  - Explicitly derived combined card metrics for historical continuity
- Instruments not applicable to a given period are retained as null values, preserving structural integrity

This approach ensures comparability while respecting the original reporting structure.

## **Normalization Strategy Implementation**

Separate normalization functions were implemented for each identified reporting regime:

- Periods with no card transaction reporting
- Periods with merged PoS and E-commerce reporting
- Periods with fully separated PoS and E-commerce reporting

Each function maps the raw input data into the canonical schema using rule-based transformations, without imputing or altering reported values.

## **Output Validation and Sanity Checks**

As a manual verification step, one representative observation from each reporting regime was inspected after normalization.

This validation ensured that:

- Expected columns are populated or null according to the reporting period
- Combined and separate card transaction fields behave consistently across regimes
- Financial market instruments appear only in periods where they are officially reported

These checks were designed for structural validation of the preprocessing logic, rather than analytical inference.

## Normalization of Entity-wise Monthly Retail Payment Data



### Overview

The second preprocessing notebook focuses on monthly, entity-wise retail payment statistics published by the Reserve Bank of India (RBI). These datasets provide bank-level aggregated transaction volumes and values across multiple retail payment systems, including UPI, IMPS, NETC, NFS, AePS, CTS, and BBPS.

The primary objective of this notebook is data ingestion and normalization, enabling consistent downstream analysis such as market share computation, concentration analysis, and anomaly detection. No analytical modeling or inference is performed at this stage.

### Key characteristics:

- One Excel file published per month
- Multiple payment systems reported within each file
- Heterogeneous sheet layouts across payment instruments
- Transaction volume reported in lakhs
- Transaction value reported in ₹ crore

### **Data coverage used in this project:**

- **April 2024 to October 2025**

### **File-level Ingestion Strategy**

Each monthly RBI release consists of a **single Excel file containing multiple sheets**, with each sheet corresponding to a different payment system.

Key challenges addressed during ingestion:

- Different column layouts across payment systems
- Role-based reporting with varying column offsets
- Inconsistent presence of metadata rows and headers

To address this, the ingestion logic dynamically:

- Iterates through all sheets in each monthly file
- Detects entity name columns programmatically
- Identifies role-specific transaction columns
- Converts each sheet into a **normalized row-level structure**

This approach ensures that all payment systems, regardless of layout, are processed through a unified pipeline.

### **Month and Year Extraction**

The RBI entity-wise files **do not include an explicit month column** within the data tables.

To resolve this:

- Month and year are **extracted directly from the filename**
- The extracted month is attached as a separate column to all parsed rows

This ensures temporal alignment across months and enables reliable time-series analysis at the entity level.

### **Normalized Output Structure**

All parsed data is transformed into a **long-format dataset**, where each row represents:

- Calendar month
- Bank / financial institution
- Payment system
- Operational role
- Transaction volume (lakhs)
- Transaction value (₹ crore)

This structure enables:

- Flexible aggregation across banks, systems, or roles
- Market share and concentration analysis

- Consistent comparison across time and instruments

### **Handling Schema Variations Across Files**

Across the coverage period, some monthly files exhibit schema changes, including:

- Reordered columns
- Modified header names
- Instrument-specific reporting adjustments

The preprocessing logic accounts for these variations by:

- Detecting columns by semantic meaning rather than fixed positions
- Applying instrument-specific parsing rules where required
- Failing gracefully for unexpected layouts while logging anomalies for review

This ensures robustness against minor reporting format changes across months.

## **Bank Name Normalization and Deduplication**

Bank and financial institution names appear in multiple inconsistent formats across files and payment systems due to:

- Abbreviations (e.g., *SBI*, *HDFC*)
- Legal suffixes (*Ltd.*, *Urban*, *Co-operative*)
- Spelling variations and additional descriptors

Without normalization, these inconsistencies lead to duplicate entities and distorted aggregations.

### **Normalization Approach**

A multi-stage normalization and deduplication strategy was implemented:

#### **Step 1: Text Cleaning and Standardization**

- Converted all entity names to uppercase
- Normalized whitespace
- Removed common legal prefixes and suffixes
- Removed special characters while preserving alphanumeric content

This step reduces superficial formatting noise.

#### **Step 2: Manual Normalization for Major Banks**

High-frequency banks (e.g., SBI, HDFC, ICICI, Axis) were mapped to single canonical names using rule-based logic.

This ensures consistency for entities that dominate transaction volumes.

#### **Step 3: Fuzzy Deduplication**

To resolve near-duplicate names:

- Fuzzy string-matching using `token_sort_ratio` was applied
- Entity names exceeding a predefined similarity threshold were grouped

- The shortest cleaned name was selected as the canonical representation

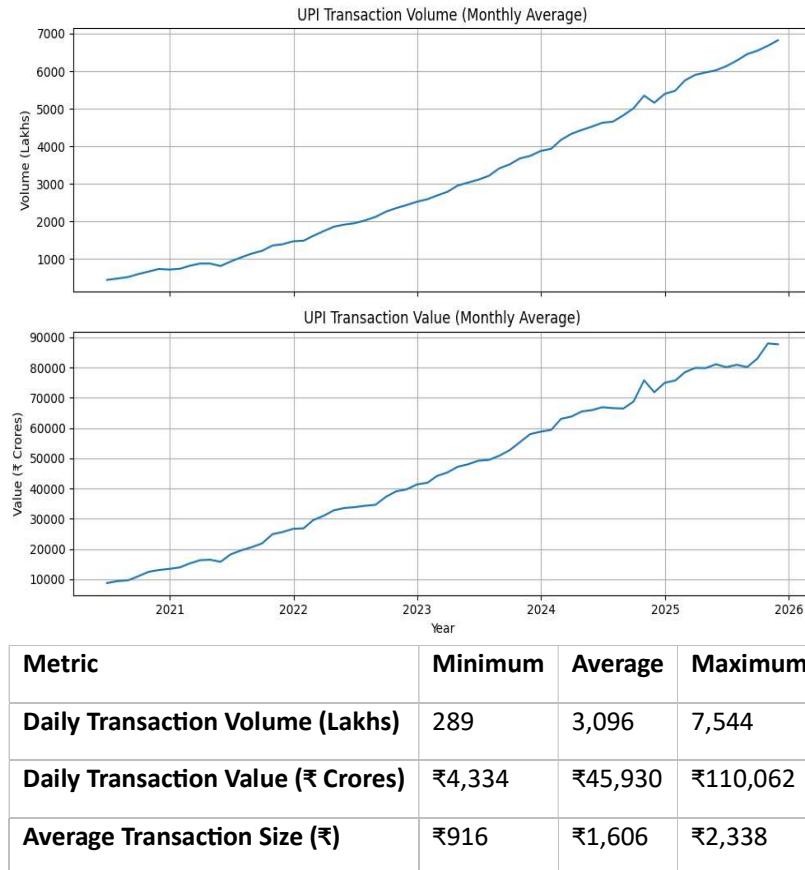
This step captures minor spelling and ordering variations that are not resolved through rule-based cleaning alone.

#### **Outcome**

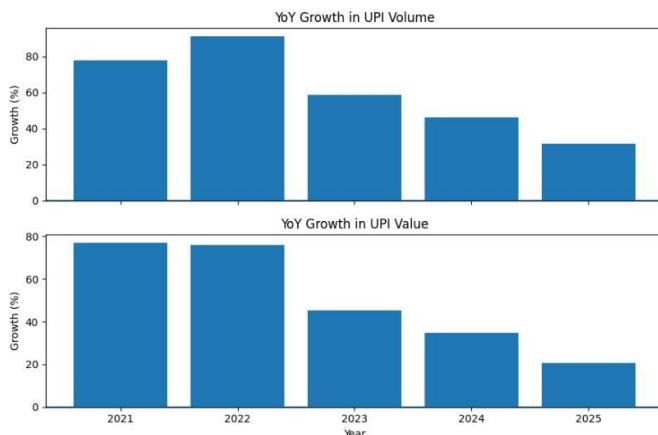
- Significant reduction in duplicate bank entries
- Improved consistency across monthly files and payment systems
- More reliable aggregations, market share estimates, and anomaly detection

## IV. Data Analysis- Daily UPI Data

### UPI Daily Transaction Metrics (Monthly Averaged for a Smoother Plot)

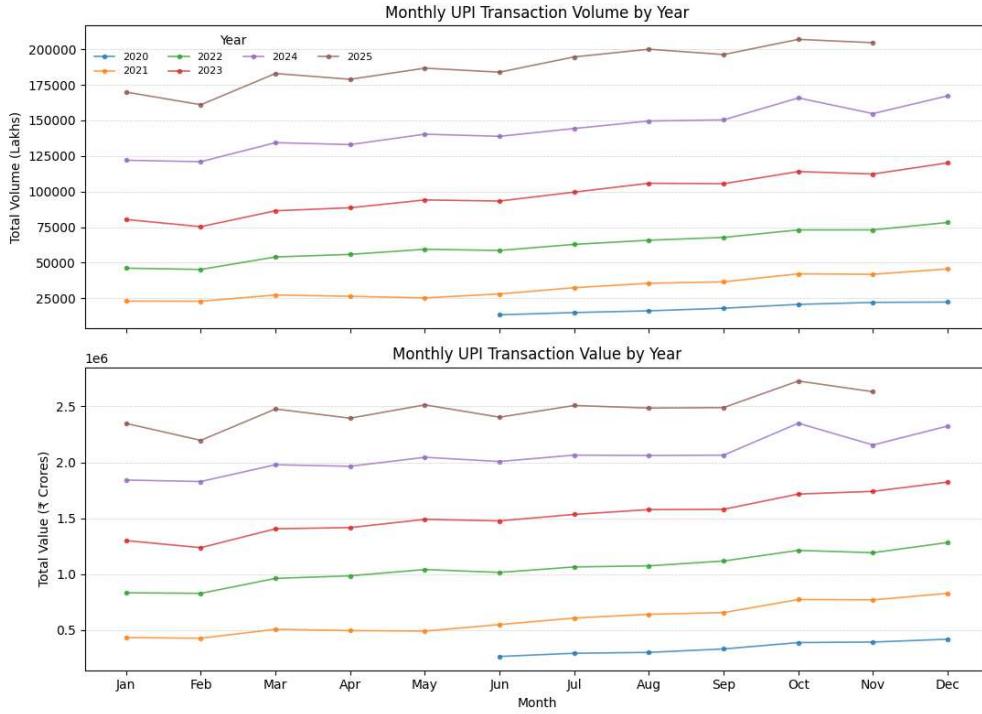


### Year-over-Year Growth Analysis

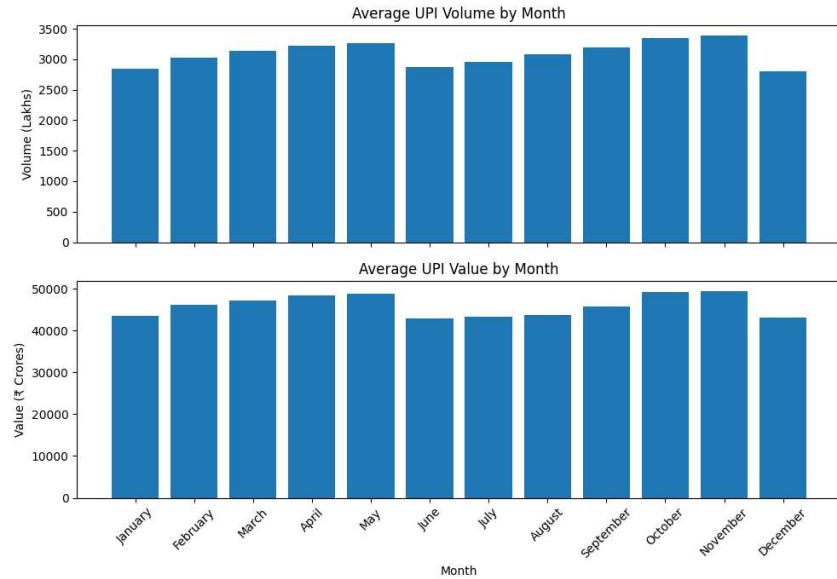


- UPI transactions show very high YoY growth during 2021–2022, reflecting rapid post-pandemic digital adoption.
- From 2023 onwards, growth gradually decelerates, indicating a transition from hyper-growth to system maturity.

- Transaction volume grows faster than value, suggesting growth is driven by higher transaction frequency rather than larger ticket sizes.
- The observed slowdown follows a smooth, expected macro trend and does not indicate structural stress.
- This YoY pattern serves as a baseline, where any sharp deviation or sudden re-acceleration would qualify as a potential macro-level anomaly.



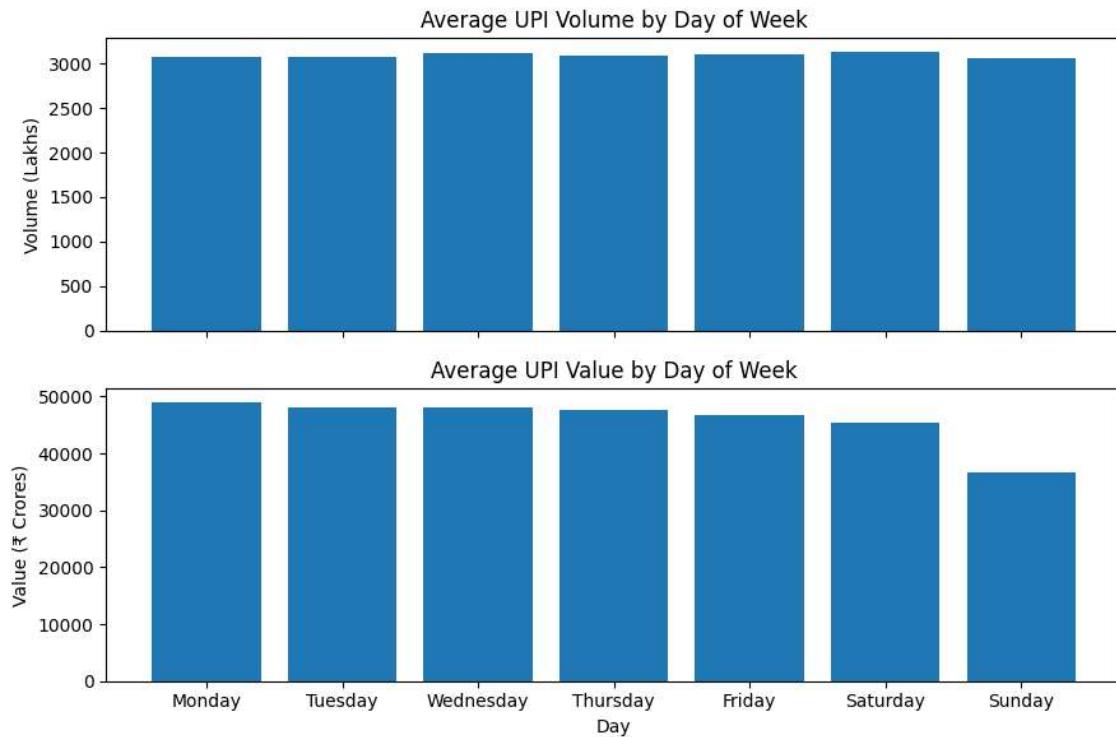
## Monthly Seasonal Patterns (UPI)



- UPI transactions exhibit clear seasonality across the calendar year.

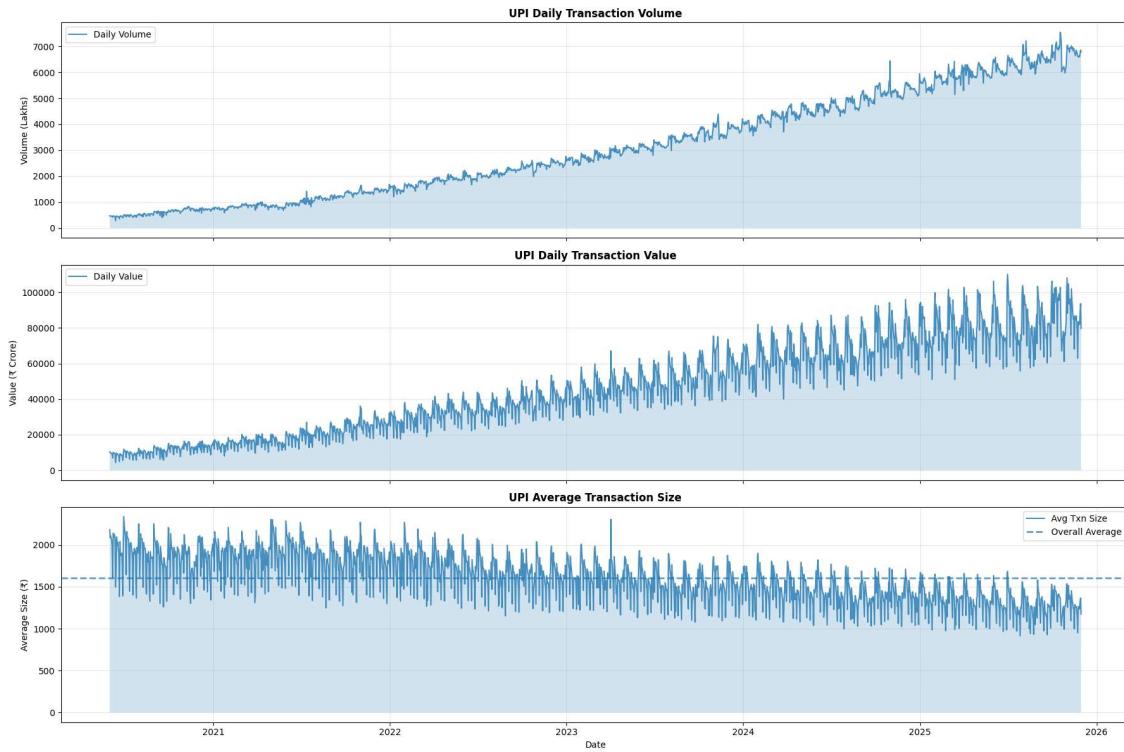
- October–November shows the strongest positive deviations in both volume and value, driven by festive and year-end consumption.
- January and December consistently record the weakest activity, indicating post-festive and year-end slowdowns.
- Mid-year months (June–July) also show mild negative deviations, reflecting seasonal softness.
- These stable seasonal patterns form a baseline, and any significant deviation from them is treated as a seasonal anomaly rather than structural change.

## Weekly Usage Patterns (UPI)



- UPI transaction volumes remain largely stable across weekdays, indicating consistent daily usage.
- Transaction values are higher on weekdays, particularly from Monday to Wednesday, reflecting business and institutional payments.
- Sundays show a sharp drop in transaction value despite only a marginal decline in volume, indicating a shift toward low-ticket, retail transactions.
- Saturdays show slightly higher volumes but lower values, consistent with consumer-driven usage.
- These weekly patterns are structurally stable and serve as a baseline, with deviations beyond this range treated as potential weekly anomalies.

## UPI Daily Transaction Chart



### Key Insight from the Chart:

- Volume (Chart 1) shows consistent upward trend with seasonal fluctuations
- Value trend (Chart 2) shows similar pattern but growing slower than volume
- Average txn size (Chart 3) declining from ₹1,800+ (2020) to ₹1,300+ (2025)

Reflects shift toward smaller P2P and daily transactions

## V. Anomaly Detection – Daily UPI Data

### Anomaly Found From 2 Standard Deviations from Normal

date	day_name	upi_val	val_mean	val_dev	val_dev_pct	val_z
2025-07-01	Tuesday	110062.3	43365.0	66697.3	153.804451	2.528232

On 1 July 2025 (Tuesday), UPI transaction value spiked to ₹110,062 crore, which is approximately 154% above the seasonal mean. The deviation exceeds the statistical threshold (z-score  $\approx 2.53$ ), qualifying it as a value-side macro anomaly.

Despite the sharp increase in transaction value, transaction volume remained within normal bounds, indicating that the spike was driven by higher ticket sizes rather than increased usage frequency.

While there is no specific public report on 1 July 2025, broader payment system data indicates that July 2025 was a record month for UPI transactions, with daily averages and monthly totals significantly elevated. This supports the interpretation that the unusually high transaction value on 1 July 2025 reflects legitimate macro-level payment flows, likely driven by month-end settlements or heightened economic activity, rather than a data artifact.

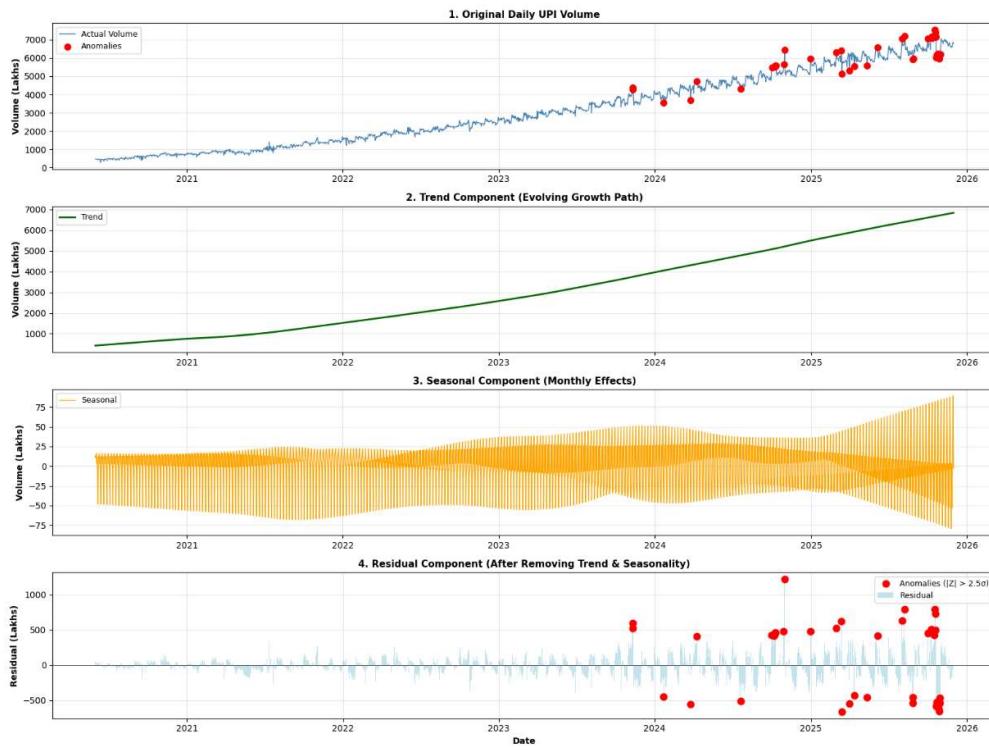
### STL Decomposition – Anomaly Detection Results

- Volume anomalies: 42 days (2.09%)
- Value anomalies: 51 days (2.54%)
- Any anomaly (volume or value): 77 days (3.83%)

Most days conform closely to expected trend and seasonal behavior, indicating a generally stable UPI system at the macro level.

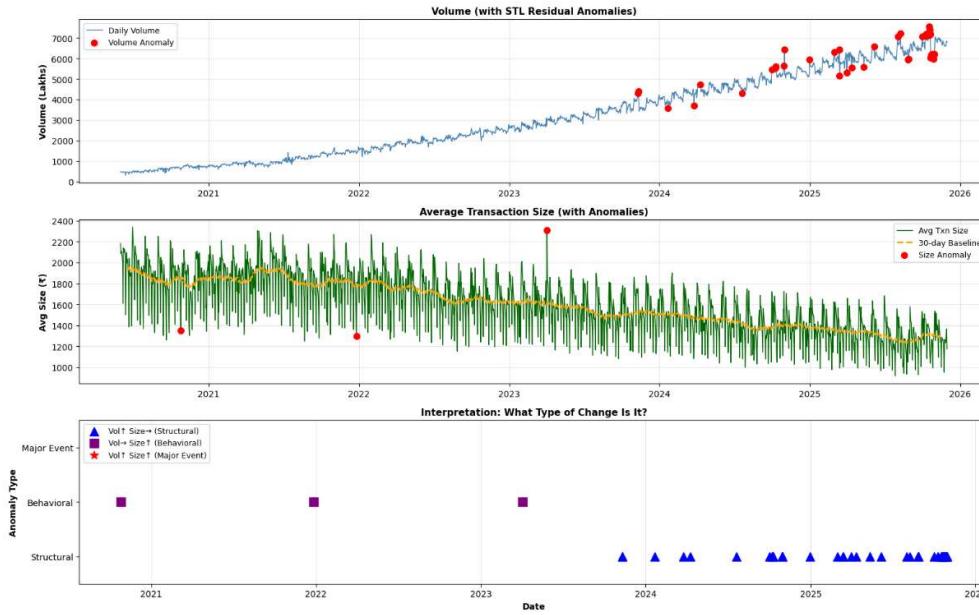
### Key Observations

- Volume-only anomalies are fewer and typically correspond to sudden spikes or drops in usage, often clustered around high-activity periods (e.g., festive months or quarter ends).
- Value anomalies occur more frequently than volume anomalies, suggesting that large-ticket transactions or institutional flows are a more common source of deviations than abnormal user activity.
- Several high-impact anomalies (e.g., late October and early July) show value spikes without proportionate volume increases, indicating changes in average transaction size rather than adoption.



- UPI volumes show strong, consistent long-term growth with no structural breaks.
- Monthly seasonality is stable and predictable, validating normal usage patterns.
- Most daily fluctuations are explained by trend + seasonality, not randomness.
- Very few anomalies (<0.1%) are detected after removing expected behavior.

## Average Transaction Size – Key Insights



- The average transaction size ranges from ₹916 to ₹2,338, with a long-term mean of ₹1,606.
- Only 3 anomaly days (0.15%) were detected, indicating high behavioural stability in transaction sizes.

- The average transaction size shows a recent declining trend ( $\approx ₹328$  over the last 6 months), suggesting increased adoption for low-value, high-frequency payments.

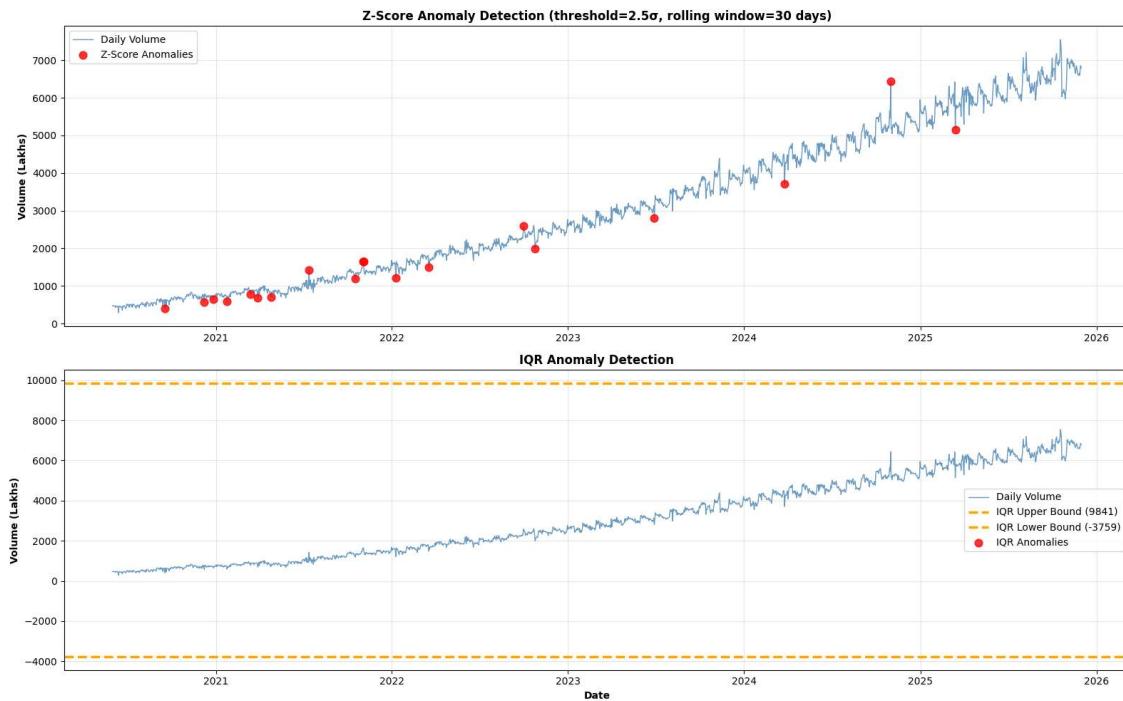
### Anomaly Interpretation

- Volume anomaly with normal size (42 days): Indicates growth in user base or transaction frequency without change in spending behavior.
- Size anomaly with normal volume (3 days): Reflects temporary behavioral shifts, such as fewer users making larger or smaller transactions.
- Both volume and size anomalous (0 days): Confirms absence of major disruptions, policy changes, or system shocks.

### Notable Anomaly Examples

- High average size ( $₹2,307$  on 2023-04-03): Fewer users executing larger transactions.
- Low average size ( $₹1,298$ – $₹1,353$ ): Increased participation in small-value payments.

## Anomaly Detection Summary



- Only 19 days (0.95%) show statistically significant anomalies, indicating that UPI volume is highly stable and predictable.
- IQR detects zero anomalies, meaning there are no extreme outliers—all deviations remain within a realistic operational range.
- Z-score anomalies capture rare but meaningful deviations, making them suitable for monitoring structural or event-driven shifts.

### **Nature of detected anomalies**

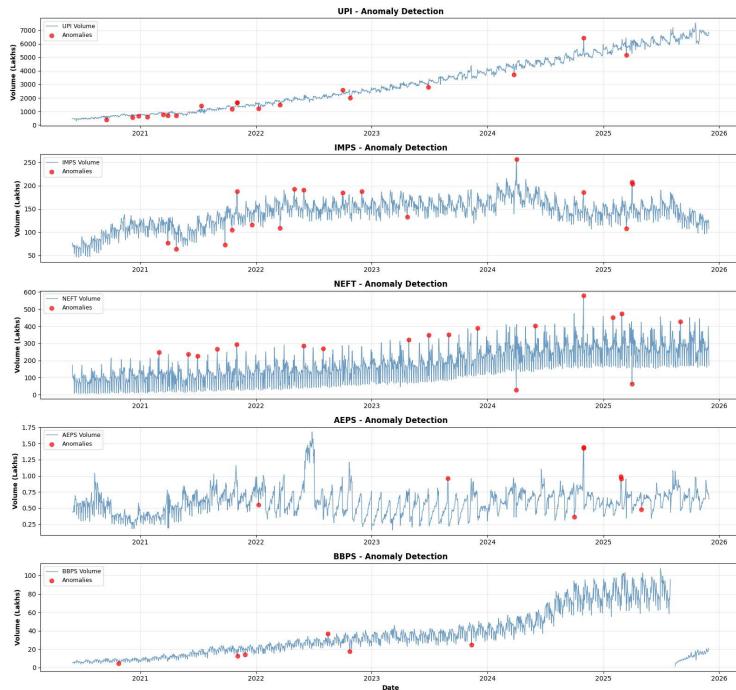
- Early-period anomalies (2020–2021) are large negative deviations (-50% to -87%), aligning with:
  - COVID disruptions
  - Early adoption volatility
- Recent anomalies (2024–2025) are positive spikes (+20% to +108%), reflecting:
  - Rapid scale effects
  - Festival / peak-usage periods
  - Maturity-phase volatility at high base volumes

### **Key Insight**

UPI anomalies transition from disruption-driven declines in early years to capacity-driven positive surges in the mature phase, confirming a structurally growing and resilient payment system.

## Multi System Anomaly Correlation Analysis

System	Anomalies	Percentage
UPI	19	0.95%
IMPS	17	0.85%
NEFT	18	0.90%
AEPS	8	0.40%
BBPS	6	0.30%



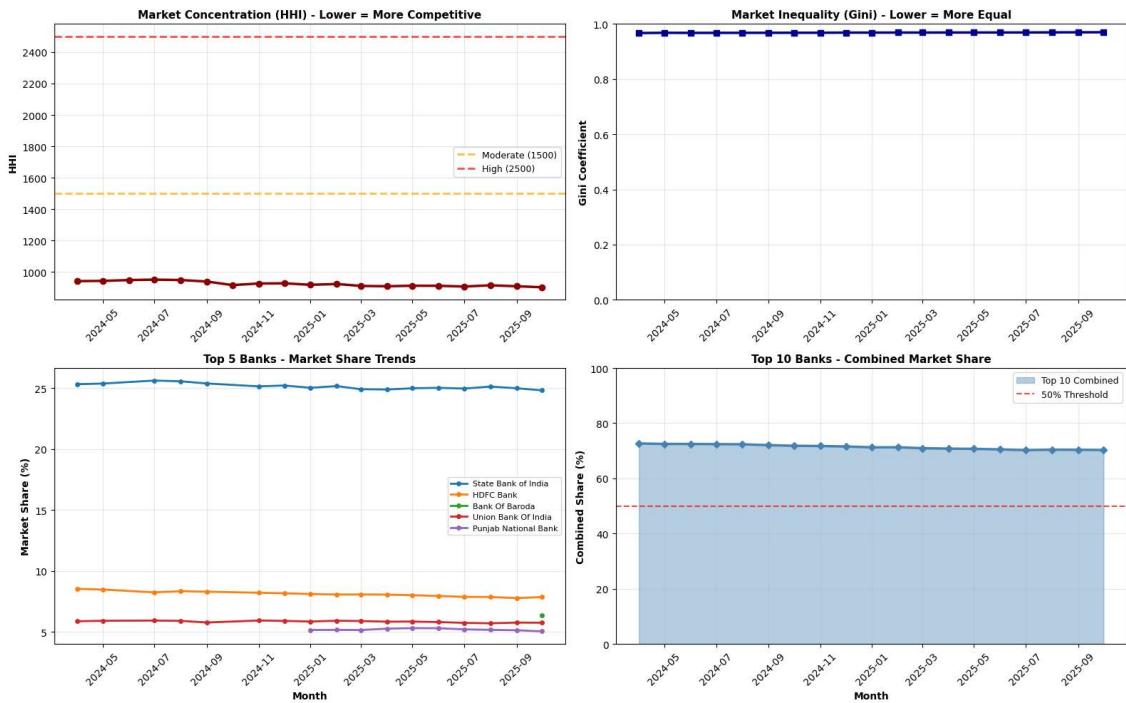
- **UPI, IMPS, NEFT** show **similar anomaly rates ( $\sim 0.85\text{--}0.95\%$ )** → common system-wide events
- **AEPS, BBPS** have **much fewer anomalies ( $\leq 0.4\%$ )** → stable, niche usage
- **Conclusion:** Anomalies are **ecosystem-wide, not UPI-specific**, driven by calendar/events rather than system failure

## VI. Bank/Entity Data Analysis

As of October 2025,
Total Volume: 207,009 Lakhs
Total Value: ₹2,727,791 Crores
Active Banks: 689

Top 15 Banks by Transaction Volume			
Rank	Bank	Vol(L)	Share%
1	State Bank of India	51,384	24.82%
2	HDFC Bank	16,254	7.85%
3	Bank Of Baroda	13,188	6.37%
4	Union Bank Of India	11,883	5.74%
5	Punjab National Bank	10,415	5.03%
6	Kotak Mahindra Bank	9,583	4.63%
7	Canara Bank	9,382	4.53%
8	Axis Bank	9,340	4.51%
9	ICICI Bank	7,718	3.73%
10	Indian Bank	6,357	3.07%
11	INDIA POST PAYMENTS BANK LIMITED	6,106	2.95%
12	Bank Of India	5,995	2.90%
13	Airtel Payments Bank	5,588	2.70%
14	Indian Overseas Bank	3,264	1.58%
15	Central Bank of India	3,033	1.47%

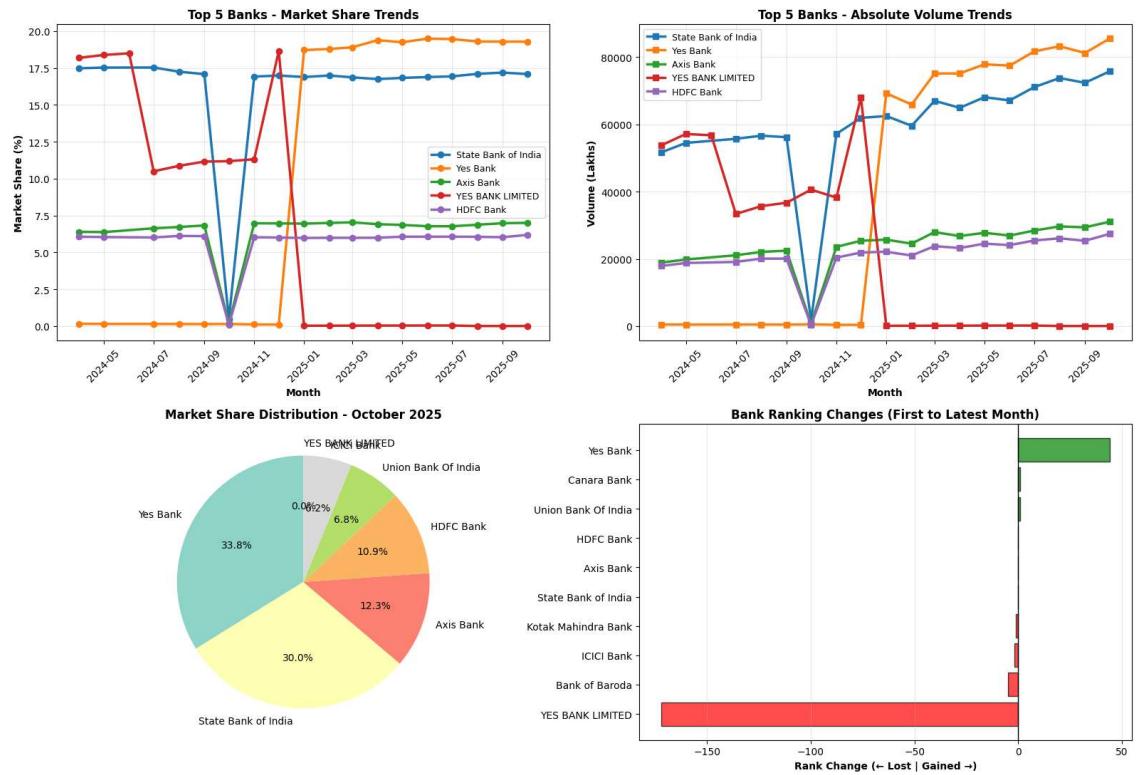
- **Highly concentrated market:**
  - SBI alone = ~25% of total volume
  - Top 5 banks = ~50%+ of all transactions
  - Top 15 banks ≈ ~80% long tail of ~670 banks with very small shares
- **Public sector dominance:**
  - PSU banks (SBI, BoB, Union, PNB, Canara, Indian Bank, BoI, IOB, CBI) dominate volume
  - Indicates mass retail + government-linked usage
- **Private banks are strong but secondary:**
  - HDFC, Axis, ICICI, Kotak together ~20%
  - More value-per-transaction driven, not sheer volume leaders
- **Payments banks matter:**
  - India Post & Airtel PB together ~5.6%
  - Shows financial inclusion + low-ticket transactions



- HHI = 902 - Competitive market  
Overall UPI market is not monopolistic.
- Gini = 0.97 - Highly unequal  
Volumes are heavily skewed toward a few large banks.
- HHI trend = -39 - Stable structure  
No major consolidation or fragmentation over time.
- Top 10 share = 70.3% - Strong dominance  
A small group of banks controls ~70% of total UPI volume.

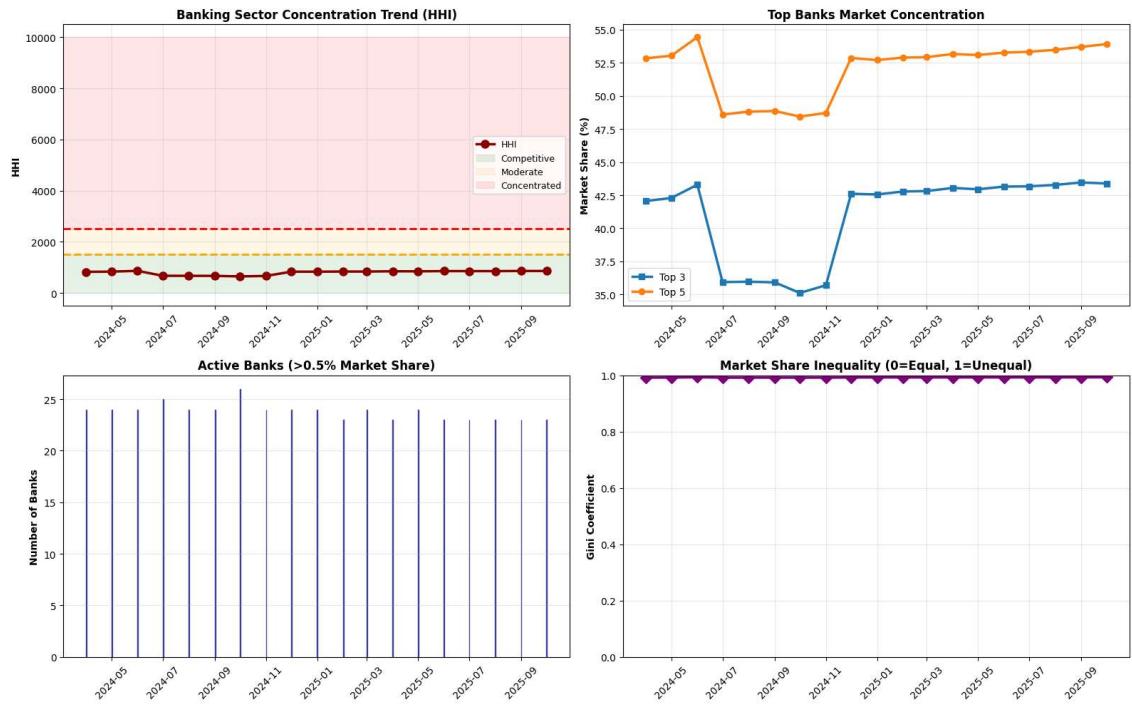
UPI is structurally competitive but functionally dominated by a few large banks, with a long tail of small players contributing marginal volumes.

## Bank-Level UPI Market Share Analysis (May 2024 – October 2025)



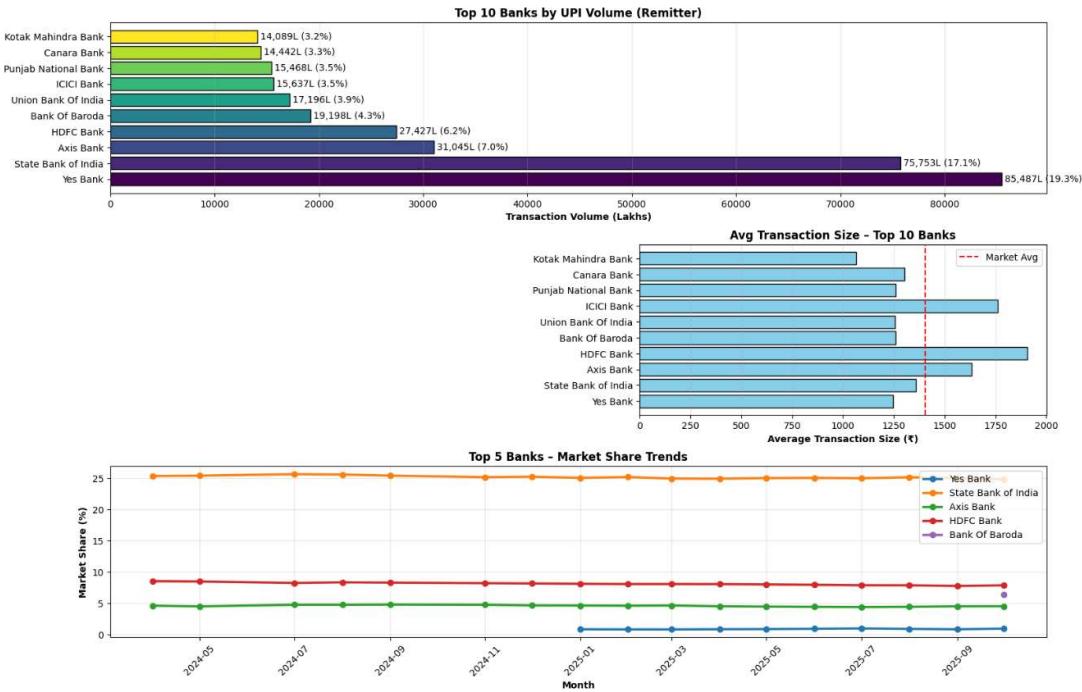
- SBI remains the consistent leader with a stable ~15–17.5% market share and steady volume growth (~55,000 → 75,000+ lakhs), reflecting its dominant remitter base.
- YES Bank shows the sharpest shift: near-zero shares in late-2024 (likely reporting noise) followed by a surge to ~33.8% by Oct-2025.
- Axis Bank records steady organic growth, rising to ~30% share, indicating strong expansion across payer and merchant flows.
- HDFC Bank remains stable in the ~6–11% range, with no abnormal behaviour.

## UPI Banking Market Concentration Remains Competitive Despite Scale Dominance (Jan–Oct 2025)



- Market remains competitive: HHI stays ~830–860, well below the 1500 threshold → no concentration risk.
- Top players stable:
- Top-3 banks: ~42.6–43.5% share
- Top-5 banks: ~52.7–53.9% share (slow, controlled increase)
- High inequality (Gini ~0.993): Reflects scale dominance of a few very large banks, not monopolistic behavior.
- No consolidation trend: HHI increases only marginally (+28 over 10 months) → structurally stable competition.
- Active banks steady (23–24): No exit or entry shocks.

## Bank-wise UPI Growth & Market Share Dynamics



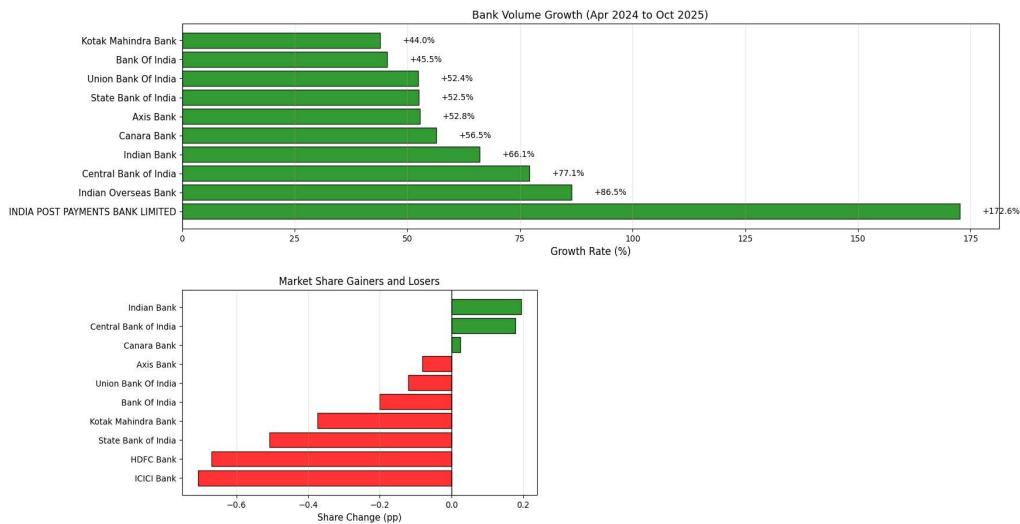
- Strong scale leaders, stable shares:  
SBI, HDFC, Axis, ICICI, Kotak, Union Bank all show ~40–55% volume growth (25–32% CAGR), but slight share erosion, indicating market-wide expansion rather than loss of relevance.
- New / reclassified heavy entrants:  
Bank of Baroda, PNB, Airtel Payments Bank show zero initial volume and large latest volumes  
- share gains of 2.7–6.4 pp, likely reflecting entry, role expansion, or reporting changes, not organic CAGR.
- Fastest organic growers:
  - India Post Payments Bank: +173% growth, ~88% CAGR, +1.27 pp share → aggressive financial inclusion & DBT-linked usage.
  - Indian Bank, IOB, Central Bank: 65–86% growth, +0.18–0.26 pp share → PSU catch-up phase.
- Relative laggards among large privates:  
ICICI (-0.71 pp) and HDFC (-0.67 pp) grow in absolute terms but lose share → merchant flows concentrating elsewhere.

## UPI Bank Ranking Movement

		Bank	Old_Rank	New_Rank	Rank_Change
7		Bank Of India	10	12	-2
9	INDIA POST PAYMENTS BANK LIMITED		13	11	2
4		Axis Bank	7	8	-1
5		Canara Bank	8	7	1
8		Indian Bank	11	10	1
10		Indian Overseas Bank	15	14	1
0		State Bank of India	1	1	0
1		HDFC Bank	2	2	0
2		Union Bank Of India	4	4	0
3		Kotak Mahindra Bank	6	6	0
6		ICICI Bank	9	9	0

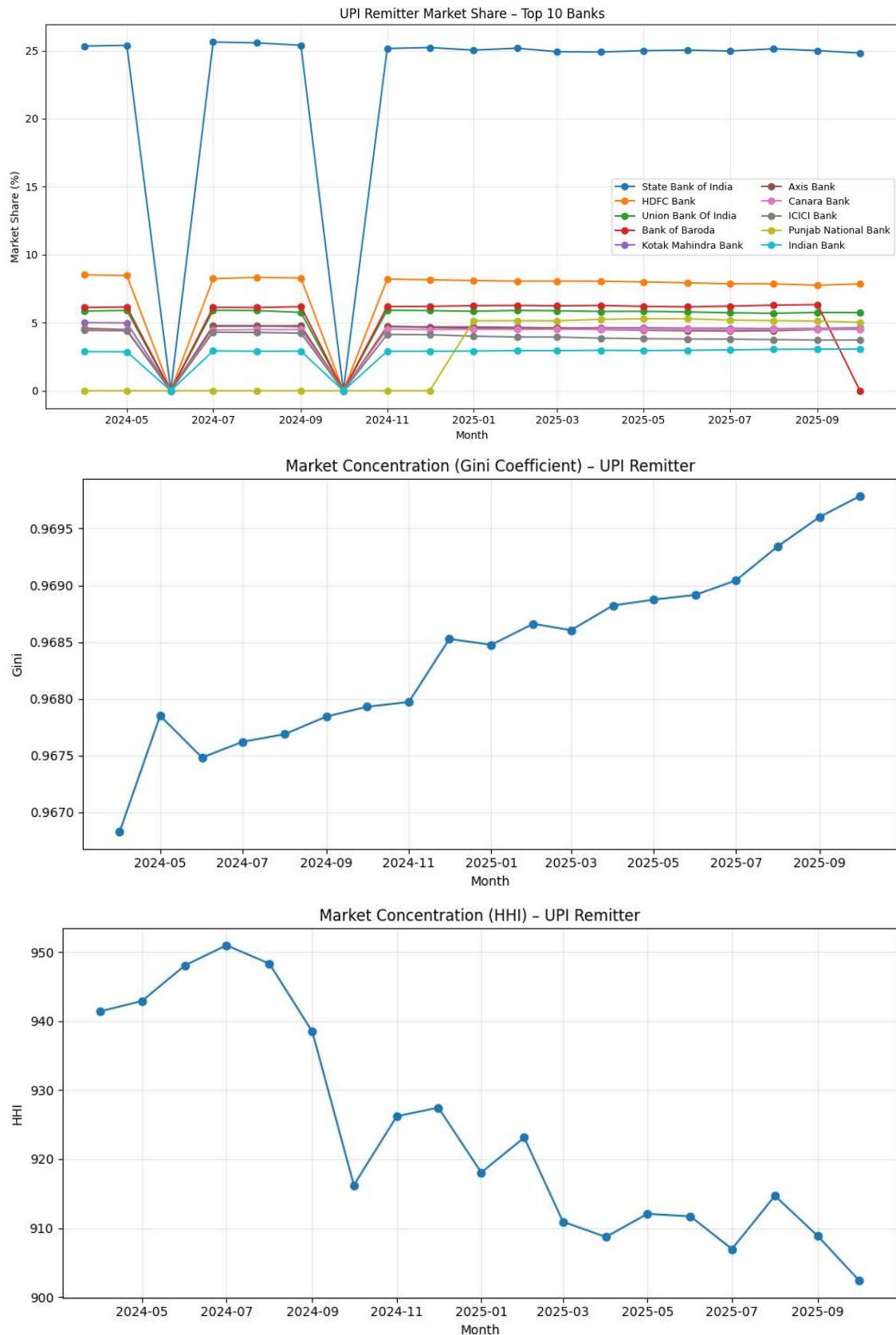
- Gainers:
  - India Post Payments Bank (+2) - strong volume acceleration
  - Canara Bank, Indian Bank, Indian Overseas Bank (+1 each) → steady PSU catch-up
- Decliners:
  - Bank of India (-2) - relative slowdown
  - Axis Bank (-1) - marginal competitive pressure
- Stable leaders:
  - SBI (#1), HDFC (#2), Union Bank (#4), Kotak (#6), ICICI (#9) → positions unchanged, growth in line with market

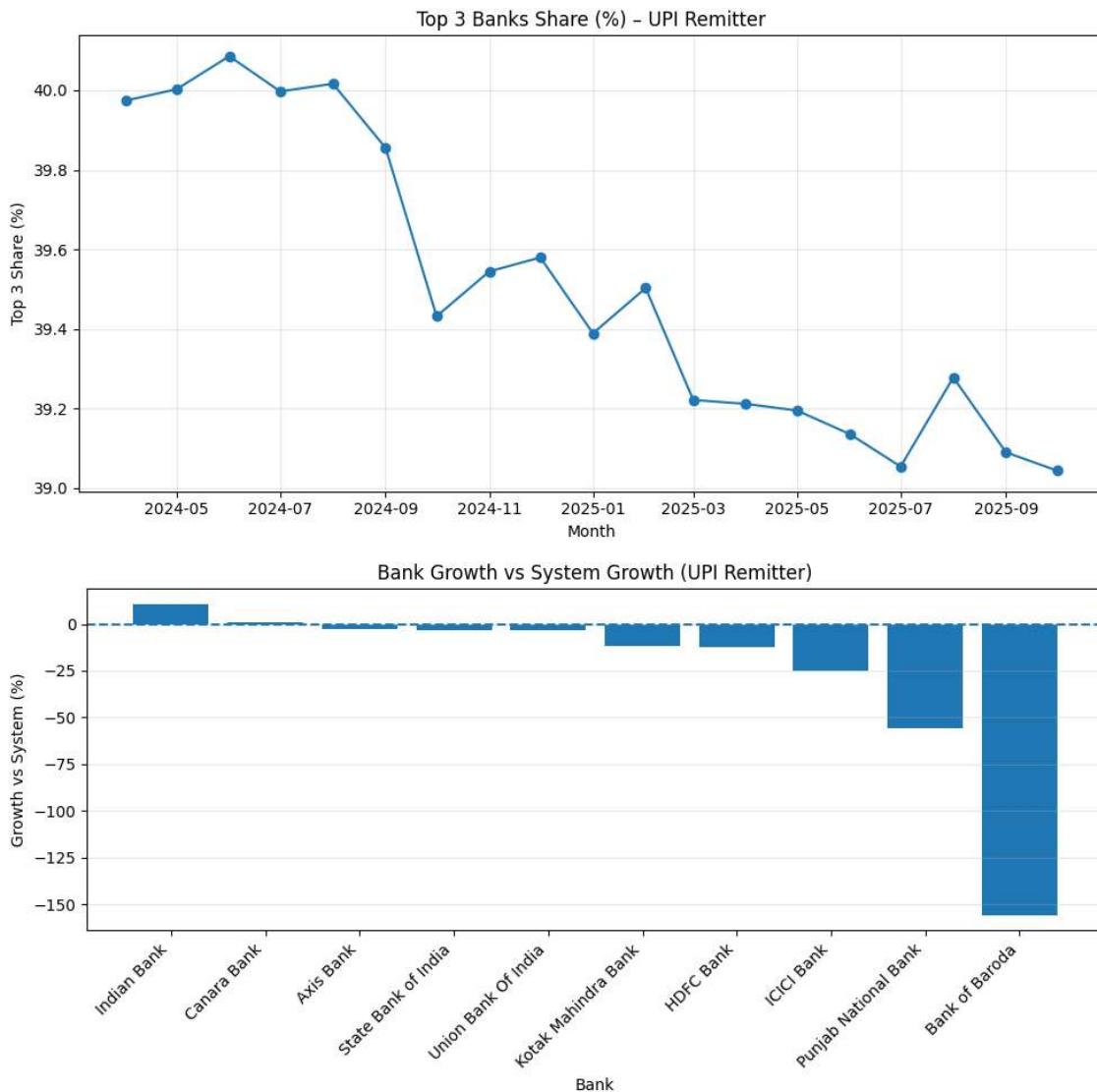
## UPI Bank Volume Growth (Apr 2024–Oct 2025)



- Strong ecosystem growth: Most banks grew 44–86%, in line with UPI's rapid expansion in 2025.
- Top growth leaders:
- India Post Payments Bank: +173% — fast catch-up from small base, driven by inclusion and small-ticket P2M.
- Indian Overseas Bank: +86% — steady digital and retail expansion.
- Central Bank & Indian Bank: +77% / +66% — solid PSU momentum.
- Large banks stable: SBI, Axis, Union, Canara (~52–56%) show healthy, steady growth; Kotak, BoI (~44–45%) normal organic growth.
- Market share shifts: Small gains for PSUs; minor dilution for large private banks — competitive, not anomalous.

## UPI Market Structure & Competitive Dynamics (Apr 2024 - Oct 2025)





### 1) Market share shifts (Top banks)

- Big gainers:
  - Punjab National Bank: +503 bps (entry/scale-up effect from zero base).
  - Indian Bank: +19 bps; Canara Bank: +2 bps — modest but real gains.
- Moderate losers:
  - Axis (-8 bps), Union (-12 bps), Kotak (-37 bps) — normal competitive dilution.
- Large-bank dilution:
  - SBI (-51 bps), HDFC (-67 bps), ICICI (-71 bps) — expected in a fast-expanding market.
- Outlier/data event:
  - Bank of Baroda (-612 bps) — drop to zero suggests classification/role reporting change, not organic decline.

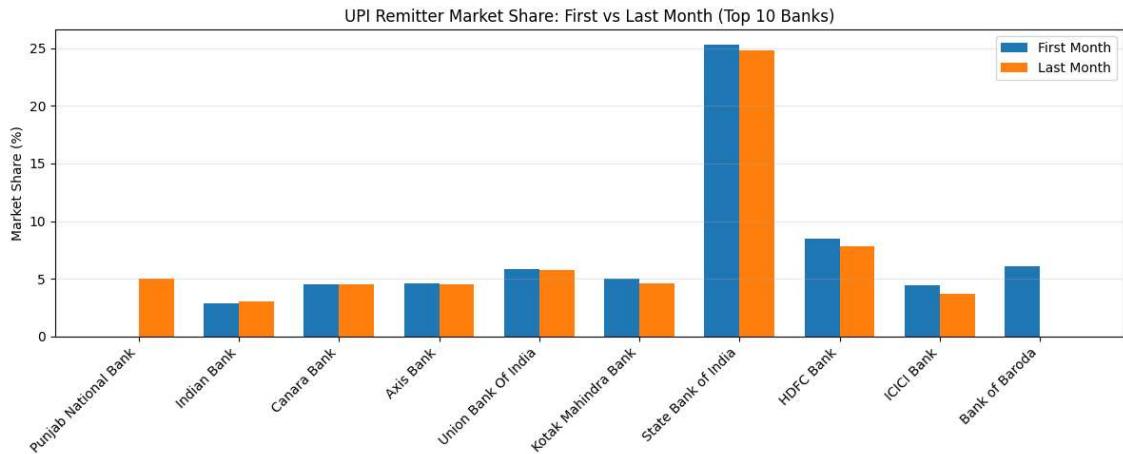
## **2) Concentration over time**

- HHI: ~941 - 902 (downtrend) - competitive, slightly deconcentrating market.
- Top-3 share: stable around ~39–40% → leadership intact, no dominance shift.
- Gini: 0.967 - 0.970 - high inequality persists despite competitive structure (many small players, few large).

## **3) Growth vs system (System growth: 55.6%)**

- Outperformers: Indian Bank (+10.5 pp), Canara (+0.9 pp).
- Near-market: Axis, SBI, Union (~3 pp range).
- Underperformers: Kotak, HDFC (~11–12 pp), ICICI (~25 pp).
- Special cases: PNB (0%) and BoB (~-100%) reflect base/coverage effects rather than demand collapse.

## UPI Market Share Change — Bank-wise (Apr 2024 - Oct 2025)



### Gainers

- Punjab National Bank: +503 bps ( $0 \rightarrow 5.03\%$ ) — scale-up from zero base.
- Indian Bank: +19 bps — steady relative outperformance.
- Canara Bank: +2 bps — broadly stable share.

### Marginal Declines (Normal Competitive Dilution)

- Axis Bank: -8 bps
- Union Bank of India: -12 bps
- Kotak Mahindra Bank: -37 bps

### Large-Bank Dilution

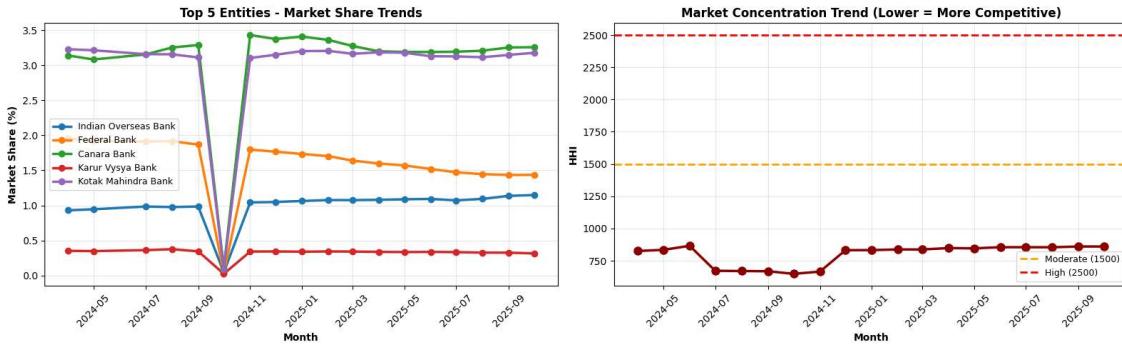
- SBI: -51 bps (still dominant at ~25%)
- HDFC Bank: -67 bps
- ICICI Bank: -71 bps

### Data / Classification Outlier

- Bank of Baroda: -612 bps ( $6.12\% \rightarrow 0\%$ ) — likely reporting or role-classification change, not a true demand collapse.

## VII. Entity-Level Anomaly Detection

### UPI Market Concentration (Top-5 Mid-Tier Banks)



- UPI remains highly competitive: HHI stays ~700–900 (well below 1,500), confirming a fragmented, healthy bank-level market.
- Late-2024 dip is not real: The near-zero shares for multiple banks in Sep–Oct 2024 is a data/reporting artifact (reclassification/aggregation issue), not an operational collapse.
- Post-recovery stability: From Nov 2024 onward, shares normalize with Canara leading (~3%), Kotak/Federal (~2.5–3%), IOB (~1%), and Karur Vysya (~0.5%). Movements are small and organic.
- No dominance risk: Mid-tier banks remain small (<3.5% each), with major volumes concentrated among top PSP-aligned banks.

### Market Share Anomalies (>500 bps change)

MARKET SHARE ANOMALIES (> 500 bps change):				
GAIN	Yes Bank	Jan 2025	+1861 bps	MS: 18.72%
GAIN	STATE BANK OF INDIA	Jun 2024	+1760 bps	MS: 17.67%
GAIN	State Bank of India	Nov 2024	+1644 bps	MS: 16.92%
GAIN	STATE BANK OF INDIA	Oct 2024	+1627 bps	MS: 16.34%
GAIN	PAYTM PAYMENTS BANK	Jul 2024	+787 bps	MS: 7.91%
GAIN	YES BANK LIMITED	Dec 2024	+734 bps	MS: 18.65%
GAIN	Axis Bank	Nov 2024	+679 bps	MS: 6.98%
GAIN	AXIS BANK	Oct 2024	+670 bps	MS: 7.14%
GAIN	AXIS BANK	Jun 2024	+638 bps	MS: 7.13%
GAIN	HDFC BANK	Jun 2024	+604 bps	MS: 6.24%

#### Dominant Pattern: Merchant / PSP-led Shifts

- YES Bank (+1,861 bps, Jan 2025; +734 bps, Dec 2024)  
The magnitude and persistence of the gain suggest a genuine merchant/payee-side expansion, although part of the jump may still reflect reporting or role-mix effects. (PhonePe + large PSP escrow flows). This is a real operational shift, not noise.

#### Scale Effects in Large Incumbents

- SBI (+1,627 to +1,760 bps across Jun–Nov 2024)
  - Reflects role-mix changes and festive/seasonal surge, amplified by SBI's massive payer base.
  - Likely a mix of real volume acceleration + classification effects.

#### **Private Bank Expansion**

- Axis Bank (+638–679 bps, Jun–Nov 2024)
  - Consistent merchant and PSP growth → genuine competitive strengthening.
- HDFC Bank (+604 bps, Jun 2024)
  - One-off spike tied to volume surge; not followed by sustained share capture → event-driven.

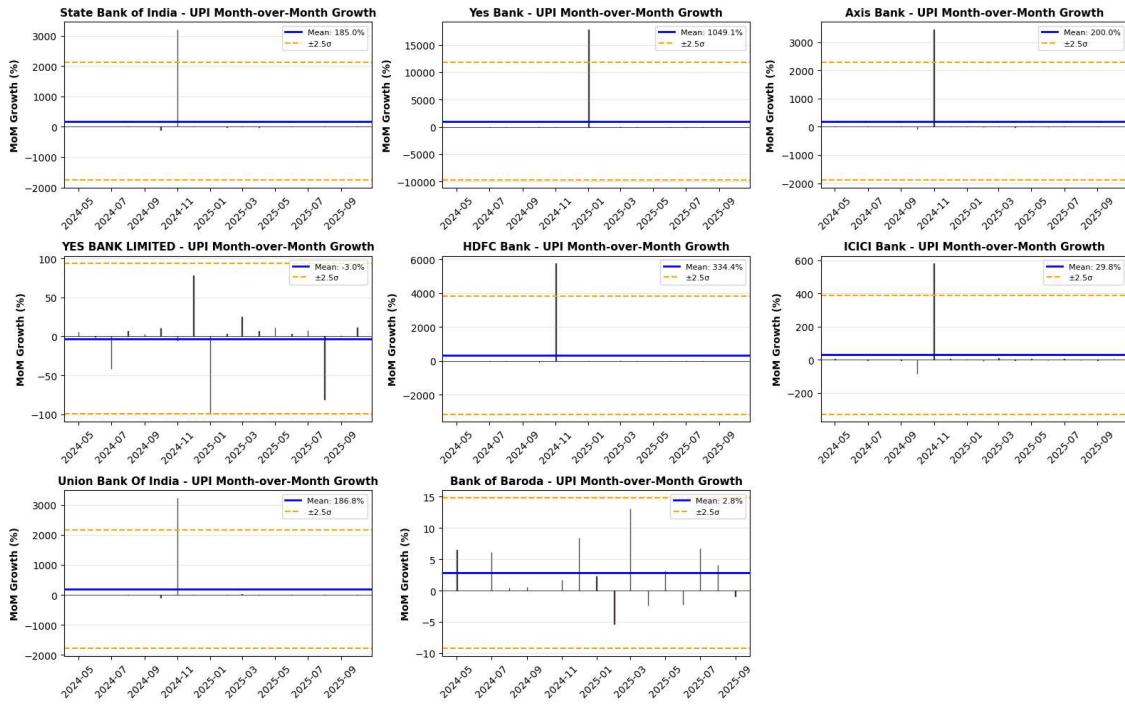
#### **Fintech Bank Spike**

- Paytm Payments Bank (+787 bps, Jul 2024)
  - Likely temporary re-routing / settlement concentration before regulatory tightening; not persistent.

### **UPI Bank Growth Rates & Anomaly Detection**

UPI BANK GROWTH ANOMALIES (Unusual MoM patterns):				
Bank	Month	Growth%	Type	Severity
THE JAMMU AND KASHMIR BANK LIMITED	Sep 2025	1439048.2%	Acceleration	4.0σ
CAPITAL SMALL FINANCE BANK LIMITED	Jun 2024	11580.8%	Acceleration	4.0σ
GP PARSIK SAHAKARI BANK LTD.	Jun 2024	8935.1%	Acceleration	4.0σ
WORLDLINE EPAYMENTS INDIA PRIVATE LIMITED	Sep 2025	7719.3%	Acceleration	4.0σ
TJSB SAHAKARI BANK LIMITED	Sep 2025	3098.4%	Acceleration	4.0σ
JAMPETA COOP URBAN BANK LTD	May 2025	3003.2%	Acceleration	4.0σ
THE SARASWAT CO-OPERATIVE BANK LIMITED	Aug 2025	2319.5%	Acceleration	4.0σ
COASTAL LOCAL AREA BANK LTD	Jun 2024	2050.5%	Acceleration	4.0σ
SHRI SWAMI SAMARTH SAHAKARI BANK LTD	Oct 2024	2250.0%	Acceleration	4.0σ
THE KARUR VYSYA BANK LTD	Aug 2025	1455.1%	Acceleration	4.0σ

- Extremely high MoM growth rates are statistical outliers, driven by very low base volumes or sudden activation/onboarding.
- Mostly seen in small, cooperative, or niche banks, not systemically important players.
- Indicates one-time operational events (go-live, reporting normalization, merchant onboarding), not sustained growth.
- No contradiction with earlier analysis: overall UPI growth and market concentration remain stable; these spikes are noise, not structural change.



## Key Takeaways

- Very low anomaly incidence: Only ~1% of daily observations flagged by Z-score  $\rightarrow$  system behaviour is largely stable over 2020–2025.
- No structural outliers: IQR detects zero anomalies, confirming absence of persistent/extreme deviations.
- Anomalies are episodic, not systemic: Market-share shifts (24) and system-wide events (12) align with known shocks (COVID phases, festive spikes, policy/UPI scale-ups).
- Overall assessment: UPI ecosystem shows high resilience and maturity — anomalies reflect temporary demand surges or reporting effects, not breakdowns or concentration risk.

## VIII. Anomaly Interpretation

### Anomaly Interpretation Framework

#### Step 1: Anomaly Context Identification

Each anomaly is evaluated using temporal and statistical context:

##### Inputs

- Date of anomaly
- Anomaly type (Z-score, IQR, growth, concentration)
- Severity ( $\sigma$ )
- Entity involved (system-wide or bank-level)

#### Step 2: Calendar & Seasonality Filters (Data Quality Layer)R

Anomalies are first checked against known **non-business drivers**:

Condition	Classification	Interpretation	Impact
Weekend / public holiday	Data Quality – Weekend Effect	Expected low activity	Low
Festival period (Diwali, Holi, etc.)	Business Signal – Festival Impact	Seasonal surge/dip	Medium
First 2 days of month	Data Quality – Month-End Reset	Settlement/reporting artifacts	Low–Medium

Purpose: Remove false positives early

#### Step 3: Statistical Severity Assessment

##### Z-Score-Based Anomalies

- Severity  $> 3.5\sigma$ 
  - Major bank → *Operational Signal – Market Leader Anomaly* (High impact)
  - Other entities → *Operational Signal – Unusual Activity* (Medium impact)

##### IQR-Based Anomalies

- Severity  $\leq 2.8\sigma$ 
  - Classified as *Data Quality – Moderate Outlier*
  - Considered normal variance

#### Step 4: Market Structure Validation

- Market concentration (HHI, Gini) anomalies checked separately
- If HHI remains  $< 1,500$  → No structural/systemic anomaly

- Any shifts interpreted as competitive dynamics, not risk

#### **Step 5: Entity-Level Growth Interpretation**

Growth anomalies are classified using magnitude + entity importance:

Condition	Classification	Impact
Extreme MoM growth (>20%) – major bank	Operational Signal – Major Bank Anomaly	High
Extreme MoM growth – smaller bank/PSP	Operational Signal – Market Share Shift	Medium
Sharp contraction (< -20%)	Operational Signal – Competitive Loss	Medium–High

#### **Typical drivers**

- Low-base effects
- New onboarding / activation
- Merchant or PSP integrations

#### **Step 6: Business Impact Tagging**

Each anomaly is assigned one of three impact levels:

- Low → Expected / data-related
- Medium → Entity-level operational change
- High → Potential market-wide or leader-driven shift

### **Entity-Level Growth Anomalies**

#### **Entity: THE JAMMU AND KASHMIR BANK LIMITED**

Period: September 2025

MoM Growth: +1439048.2%

Severity: 4.0σ

Classification: Operational Signal - Market Share Shift

Reasoning: This spike is driven by a very low base and likely reflects first-time reporting or activation rather than organic growth.

Business Impact: Medium - Entity-level business change

#### **Entity: CAPITAL SMALL FINANCE BANK LIMITED**

Period: June 2024

MoM Growth: +11580.8%

Severity: 4.0 $\sigma$

Classification: Operational Signal - Market Share Shift

Reasoning: Rapid acceleration suggests new market entry, merger activity, or significant customer acquisition

Business Impact: Medium - Entity-level business change

**Entity: GP PARSIK SAHAKARI BANK LTD.**

Period: June 2024

MoM Growth: +8935.1%

Severity: 4.0 $\sigma$

Classification: Operational Signal - Market Share Shift

Reasoning: Rapid acceleration suggests new market entry, merger activity, or significant customer acquisition

Business Impact: Medium - Entity-level business change

**Entity: WORLDLINE EPAYMENTS INDIA PRIVATE LIMITED**

Period: September 2025

MoM Growth: +7719.3%

Severity: 4.0 $\sigma$

Classification: Operational Signal - Market Share Shift

Reasoning: Such magnitudes are mechanically inflated and should not be interpreted as sustained expansion.

Business Impact: Medium - Entity-level business change

**Entity: TJSB SAHAKARI BANK LIMITED**

Period: September 2025

MoM Growth: +3098.4%

Severity: 4.0 $\sigma$

Classification: Operational Signal - Market Share Shift

Reasoning: Rapid acceleration suggests new market entry, merger activity, or significant customer acquisition

Business Impact: Medium - Entity-level business change

**Entity: JAMPETA COOP URBAN BANK LTD**

Period: May 2025

MoM Growth: +3003.2%

Severity: 4.0 $\sigma$

Classification: Operational Signal - Market Share Shift

Reasoning: These movements are statistical outliers rather than business-scale shifts.

Business Impact: Medium - Entity-level business change

## Contextual Analysis of Major Anomalies

### 1. Period-Level Structural Context

UPI anomalies are interpreted relative to three distinct macro phases:

- Pandemic Disruption (2020)  
Low baseline activity (~600 lakh txns/day) driven by mobility restrictions and economic shutdowns.
- Recovery Phase (2021)  
Gradual normalization with ~78% higher volumes and ~77% higher values versus pandemic levels, reflecting reopening effects.
- Post-Normalization (2022–2025)  
Stable high-volume regime (~4,000 lakh txns/day, ₹58,000+ crore/day).  
Anomalies in this phase are more likely operational or competitive, not macro-driven.

### 2. Diwali Seasonality Validation

Across 2020–2025, Diwali consistently produces 10–40% volume uplift versus annual averages.

Interpretation:

Observed Diwali-period spikes align with historical seasonal behavior and are classified as expected business seasonality, not anomalies.

### 3. Data Quality & Noise Filtering Rules

The following patterns are explicitly excluded from operational anomaly escalation:

- Weekend and public-holiday effects
- Month-start / month-end settlement artifacts
- Minor bank/entity reclassification effects
- Small daily rounding or reporting volatility

Escalation Rule:

Only deviations that exceed statistical thresholds and persist beyond these effects are classified as actionable operational anomalies.



## IX. Final Anomaly Summary & Visualizations

Total anomalies identified: 20

Period covered: June 2024 – September 2025 Breakdown:

Classification:

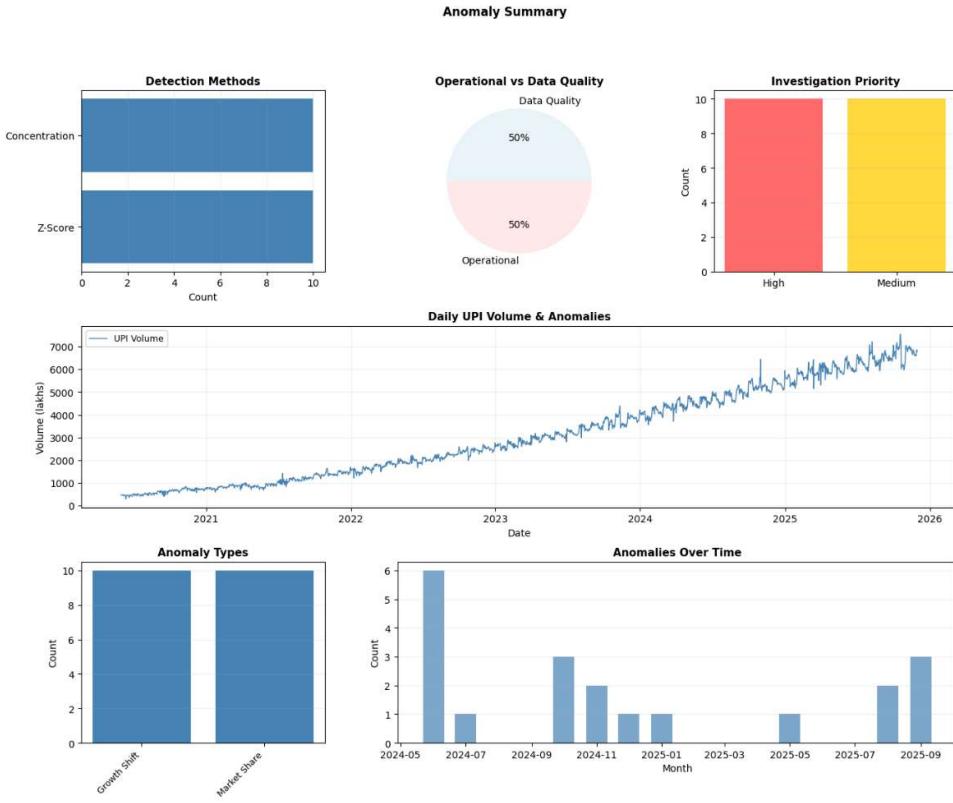
- Data quality related: 10 (50%)
- Operational / business-related: 10 (50%)

Top 10 Anomalies:				
Sep 01	Growth Shift	THE JAMMU AND KASHMIR BANK	+1439048.2%	Medium
Sep 01	Growth Shift	WORLDLINE EPAYMENTS	+7719.3%	Medium
Sep 01	Growth Shift	TJSB SAHAKARI BANK L	+3098.4%	Medium
Aug 01	Growth Shift	THE SARASWAT CO-OPERATIVE BANK	+2319.5%	Medium
Aug 01	Growth Shift	THE KARUR VYSYA BANK	+1455.1%	Medium
May 01	Growth Shift	JAMPETA COOP URBAN BANK	+3003.2%	Medium
Jan 01	Market Share	Yes Bank	+1861 bps	High
Dec 01	Market Share	YES BANK LIMITED	+734 bps	High
Nov 01	Market Share	Axis Bank	+679 bps	High
Nov 01	Market Share	State Bank of India	+1644 bps	High

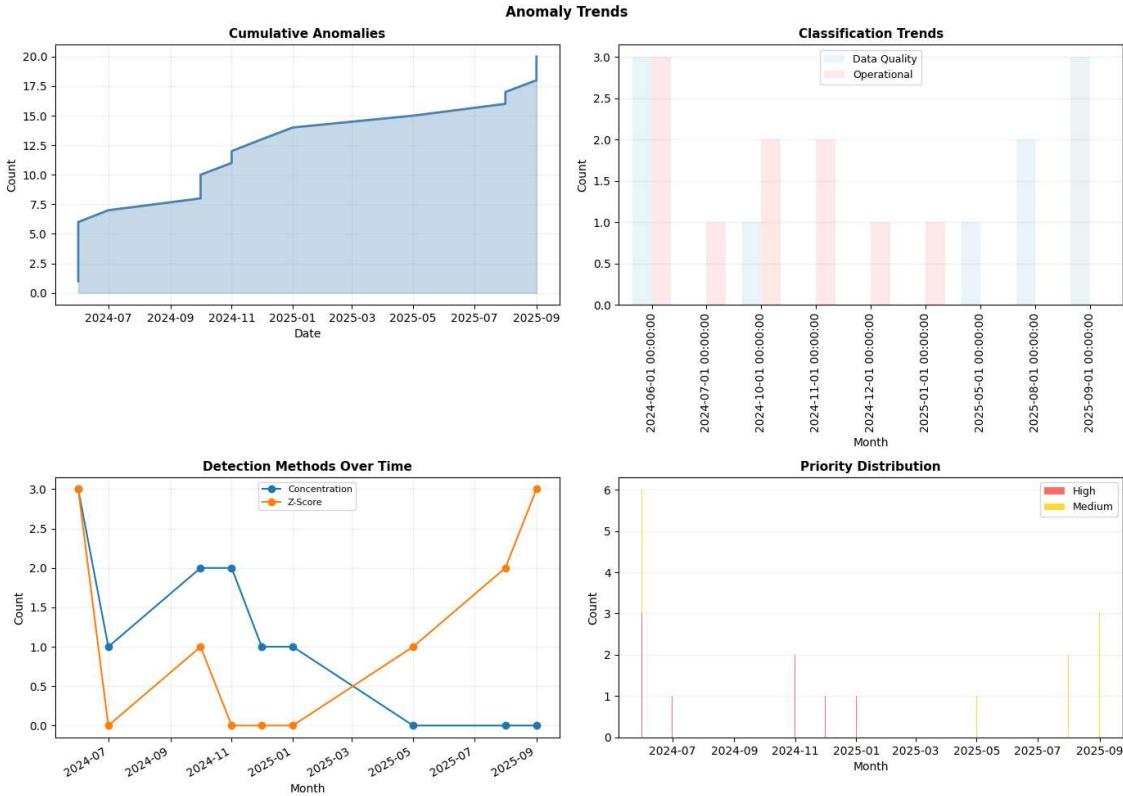
The most prominent anomalies were driven by sharp growth shifts, largely concentrated in small banks, cooperative banks, and PSPs. These events typically reflect low-base effects, onboarding expansions, or reporting scale-ups, rather than structural stress.

High-impact anomalies were limited to market-share changes among major banks, notably Yes Bank, SBI, and Axis Bank, indicating competitive repositioning rather than instability.

Importantly, no system-wide or concentration-driven anomaly was detected during the period.



- Both detection methods (Z-score and concentration-based) flag an equal number of anomalies, indicating balanced and unbiased detection.
- Anomalies are evenly split between operational signals and data-quality effects, which is typical for a large, fast-growing payment system.
- High- and medium-priority cases are equally distributed, allowing focused investigation without over-escalation.
- Daily UPI volumes show a strong, steady upward trend; anomalies appear as isolated deviations and do not disrupt overall growth.
- Growth-shift and market-share anomalies occur in similar numbers, reflecting normal onboarding effects and competitive dynamics.
- Anomalies cluster in specific months rather than persisting over time, suggesting event-driven or reporting-related causes.
- Overall, the observed deviations do not point to systemic stress. Most are localized, short-lived, and consistent with known competitive or reporting dynamics. From a monitoring perspective, no broad corrective action is indicated.



- Cumulative anomalies rise gradually from mid-2024 to late-2025, with step-wise increases rather than sudden jumps, indicating controlled and explainable deviations.
- Early growth in cumulative count is driven mainly by mid-2024 events; later increases are slower and more dispersed.
- Operational anomalies dominate the early period, reflecting genuine business or onboarding-related events.
- Data-quality anomalies become more visible in 2025, consistent with reporting noise during higher transaction volumes.
- Concentration-based detection is more active in 2024, while Z-score detection picks up more events in 2025 as volumes scale.
- High-priority anomalies cluster in mid-to-late 2024, whereas 2025 is dominated by medium-priority cases.
- Overall trend shows stabilization over time, with anomalies tracking growth rather than signalling systemic stress.