# **📊 PhonePe Transaction Insights – Project Documentation**

## **1. 📌 Project Overview**

**PhonePe** is one of India's leading digital payments platforms. With growing adoption across the country, analyzing its usage trends can uncover valuable insights into user behavior, financial growth, and market opportunities.

This project is a comprehensive business analytics dashboard built on **PhonePe Pulse data**. It extracts data directly from GitHub, processes it using Python, stores it in a SQLite database, performs SQL analysis, and visualizes trends through charts and dashboards using Matplotlib, Seaborn, and Streamlit.

## **2. 🧠 Problem Statement**

To extract insights from India’s PhonePe Pulse dataset to identify transaction trends, visualize user behavior, evaluate regional performance, and build an interactive dashboard using **Python**, **SQL**, and **Streamlit**.

## **3. 🛠️ Tools & Technologies Used**

| **Tool/Library** | **Purpose** |
| --- | --- |
| Python | Data parsing, cleaning, analysis |
| Pandas | Data manipulation |
| SQLite | Database storage and querying |
| SQL | Business queries |
| Matplotlib/Seaborn | Visualization |
| Streamlit | Interactive dashboard |

## **4. 📂 Data Source & Format**

* **Source:** [PhonePe Pulse GitHub](https://github.com/PhonePe/pulse)
* **Data Format:** Nested JSON
* **Categories:**
  + Aggregated Transactions
  + Users
  + Insurance
  + Map Data (district/state-level)
  + Top Categories

## **5. 🗃️ Data Extraction & Cleaning**

* Data cloned from GitHub using git clone
* Extracted quarterly transaction info from JSON files
* Parsed fields:  
  + state, year, quarter, transaction\_type, count, amount
* Stored in a cleaned DataFrame (agg\_txn\_df)
* Handled duplicates and nulls
* Data saved in phonepe.db SQLite database

## **6. 📈 SQL Analysis & Visualizations**

### **Queries Performed:**

* Total amount per state
* Year-wise transaction trends
* Transaction type distribution
* Quarter-on-quarter growth

### **Basic Charts:**

* Bar chart: Top states by total amount
* Line chart: Year-wise growth
* Bar chart: Transaction type comparison
* Quarter-wise transaction trend chart

## **7. 📊 Advanced Charts & Trends**

### **Chart 5: Top 10 States by Transaction Amount (Q4 2023)**

* Karnataka, Telangana, and Maharashtra lead with ₹3T+ volume
* Indicates southern and western state dominance

### **Chart 6: Quarterly Trend for Top 5 States (2020–2023)**

* Consistent Q4 spikes
* All 5 states (Karnataka, Telangana, Maharashtra, Andhra Pradesh, UP) show stable growth

### **Chart 7: Top 10 Device Brands by Users (2022)**

* Xiaomi (88M), Vivo (77M), Samsung (65M)
* Android dominates PhonePe’s user base

### **Chart 8: Top 10 Districts by App Opens (Q4 2022)**

* Bengaluru Urban: 1.13B app opens
* Urban districts dominate usage

### **Chart 9: Top 10 States by Insurance Transactions (2022)**

* Karnataka & Maharashtra lead in both count and amount
* South India emerging as an insurance leader

### **Chart 10: Top 10 Districts by New Users (Q4 2022)**

* Bengaluru added 14.6M users
* Pune, Thane, and Hyderabad follow

## **8. 🎯 Key Business Insights (2022–2023)**

| **Area** | **Insight Summary** |
| --- | --- |
| 🧭 **Top States** | Karnataka, Telangana, and Maharashtra dominate |
| 📆 **Best Quarter** | Q4 shows seasonal transaction spikes |
| 💳 **Top Categories** | Recharge, Utility Bill, P2P Payments |
| 📍 **High App Usage** | Bengaluru Urban has 1.13B app opens |
| 📈 **New User Surge** | Bengaluru added 14.6M users in Q4 2022 |
| 📦 **Insurance Uptake** | Karnataka & Maharashtra show strong adoption |
| 📱 **Device Usage** | Xiaomi, Vivo, Samsung dominate PhonePe installs |

## **9. 💡 Future Scope: ML Possibilities**

Though the project is analytics-based, we can add ML features in the future:

| **Use Case** | **ML Technique** |
| --- | --- |
| Forecasting growth | Time Series (ARIMA, LSTM) |
| Fraud detection | Isolation Forest / SVM |
| User segmentation | Clustering (KMeans, DBSCAN) |
| Predicting churn | Classification (RF, XGB) |

## **10. 📦 Project Deliverables**

| **File** | **Description** |
| --- | --- |
| PhonePe\_Insights.ipynb | Main notebook with all code, visuals, logic |
| app.py | Streamlit dashboard |
| phonepe.db | SQLite database with cleaned data |
| README.md | GitHub project summary |
| charts/ | Visual outputs (optional) |