



PhonePe Transaction Insights: From Data to Dashboard

Reporter

Contents

01

Introduction: Project Overview

02

Data Extraction and Transformation (ETL)

03

Dashboard Development and Features

04

Insights and Analysis

05

Development Best Practices

06

Future Enhancements





01.

Introduction: Project Overview

Project Goals and Objectives

01

Project Aim

Develop a comprehensive data analytics pipeline, from initial data extraction to a live, interactive dashboard for PhonePe transaction insights.

02

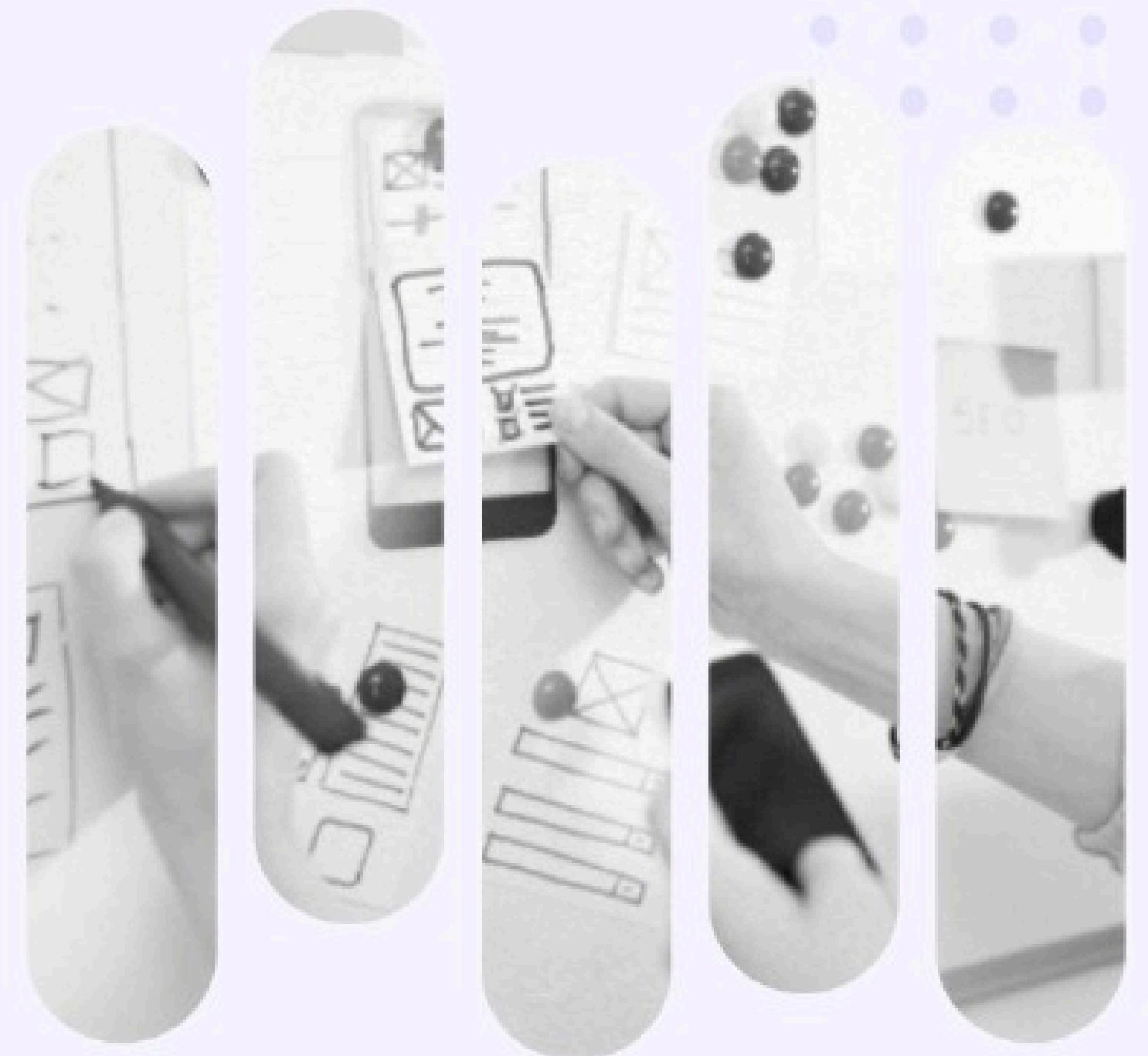
Key Objectives.

Enable granular analysis of transaction behavior across states, categories, and time periods; uncover key patterns, and ultimately deliver real-world value for analysts.

03

Deliverables Highlight

The project delivers Python source code, SQL scripts, and a Streamlit application, along with comprehensive documentation.



Value Proposition and Target Audience.

Key Value

The dashboard reveals important patterns such as the rapid growth of UPI transactions and varying insurance adoption.

Target Stakeholders

The project will primarily benefit analysts, fintech strategists, and policy planners interested in India's digital financial landscape.

Project Usefulness

This project demonstrates how open-source data can be transformed into meaningful business intelligence; it provides a practical tool for understanding digital finance trends.



02.

Data Extraction and Transformation (ETL)

Nested JSON Data Extraction.

01

Extraction Tools

Python was used for extracting data from nested JSON formats related to PhonePe transactions.



02

Data Understanding

Nested JSON data was meticulously studied to understand structure and identify key transactional elements.



03

Handling Complexities

Techniques were employed to flatten nested structures into tabular formats suitable for database storage and analysis.



Database Schema Design and Implementation



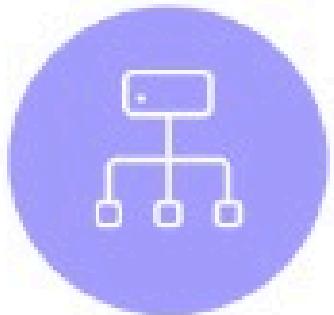
Database Choice

SQLite was selected for its ease of use and suitability for initial project scope, supporting robust analytical capabilities.



Schema Design

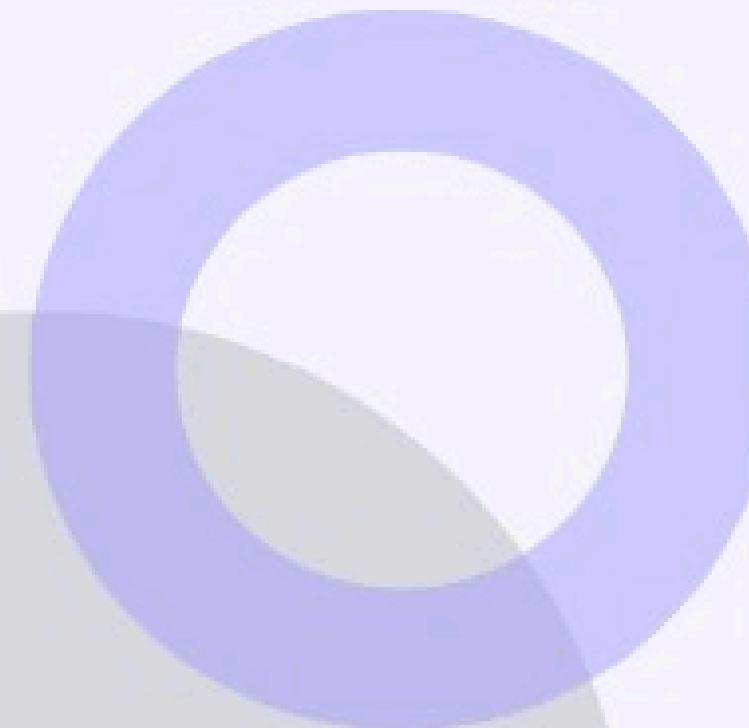
Schemas involved designing tables to store transaction details effectively, including user demographics, transaction amounts, and timestamps.



SQL Scripts

SQL scripts were created to define database schema and create tables for structured data storage.

Data Validation and Cleaning



01

Data Integrity Focus

Data validation processes ensure accuracy and reliability of transaction data throughout the ETL pipeline.

02

Validation Methods

Clean and modular Python code supports data transformation, which ensures that the system is maintainable.

03

Pytest Integration

Pytest was integrated to validate all data transformations and ensure accuracy across the entire pipeline process.



03.

Dashboard Development and Features

Streamlit Application Overview

01

Dashboard Purpose

Dashboard built using Streamlit to provide data visualization and exploration capabilities for analyzing PhonePe transactions.

02

Interactivity Goal

The Streamlit application offers real-time data visualization, enabling users to explore trends and patterns.

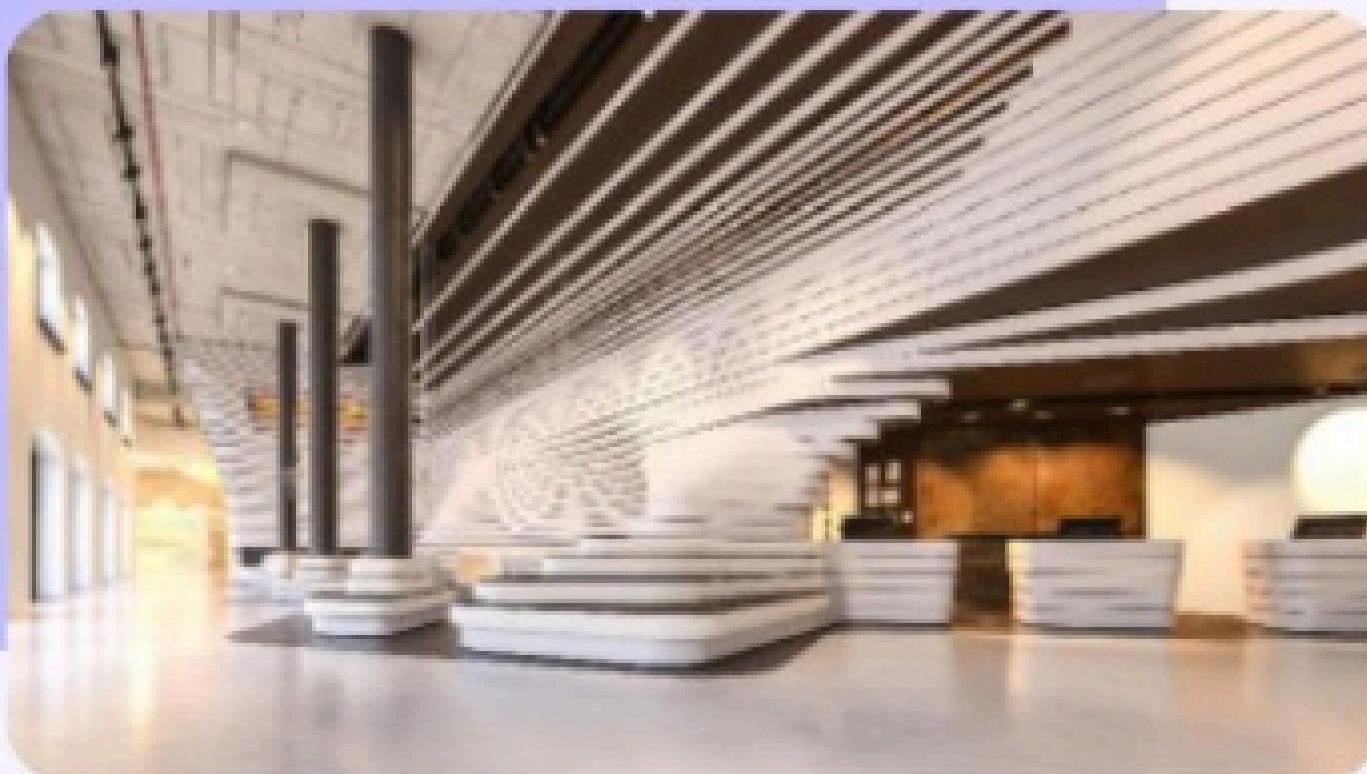
03

User Experience

The dashboard is designed to enhance user experience through interactive charts and filters to facilitate analysis.



Interactive Visualizations



Visualization Tools

Interactive charts were created to support easy exploration of transaction data across various states, transaction categories and time periods.

Visual Insights

Visualizations reveal important trends in digital financial transactions, such as the growth of UPI payments in Tier-2 and Tier-3 cities.

Regional Analysis

The dashboard enables comparative analysis of transaction behavior across different regions in India.

Key Performance Indicators (KPIs)

01

KPI Tracking

KPIs for monitoring transaction volumes, user engagement, and growth metrics are built in the dashboard.

02

Data Interpretation

The dashboard allows users to track and interpret KPIs, providing quick insights into trends such as insurance product adoption.

03

KPI Usefulness

These KPIs deliver insight for different stakeholders, particularly fintech strategists who need a deep dive in the Indian market.



04.

Insights and Analysis

Transaction Growth Trends

UPI Revolution

The dashboard highlights the rapid growth of UPI transactions, especially in Tier-2 and Tier-3 cities; this showcases India's digital payment revolution.

Market Expansion

The data reveals the significance of regional markets and their contribution to the overall transaction growth.

Data Validation

Transaction volumes have been analyzed per region and benchmarked to validate the growth trends.

User Behavior Patterns



User Engagement.

Patterns in how users engage with PhonePe for different transaction types can be unveiled through cohort analysis.



Transaction Frequency

Insights into the frequency and size of transactions performed by different user segments are delivered from the data.



Driving Factors

Demographic factors influencing transaction adoption, such as age, income, and location are analyzed, yielding useful insight.

Regional Disparities



Insurance Adoption

Uneven adoption of insurance products are discovered based on the regional adoption differences.

Regional Needs

The project uncovers regional disparities influencing digital financial behavior to better serve customer interests.

Data Validation

The team has benchmarked insurance adoption data against national averages to validate uneven adoption statistics.



05.

Development Best Practices

Coding Standards

Code Modularity

Coding best practices implemented in Python and SQL to modularise code and ensure maintainability.

Documentation

Well-documented code explaining key functions and data flows, with comprehensive readability and accessibility.

Version Control

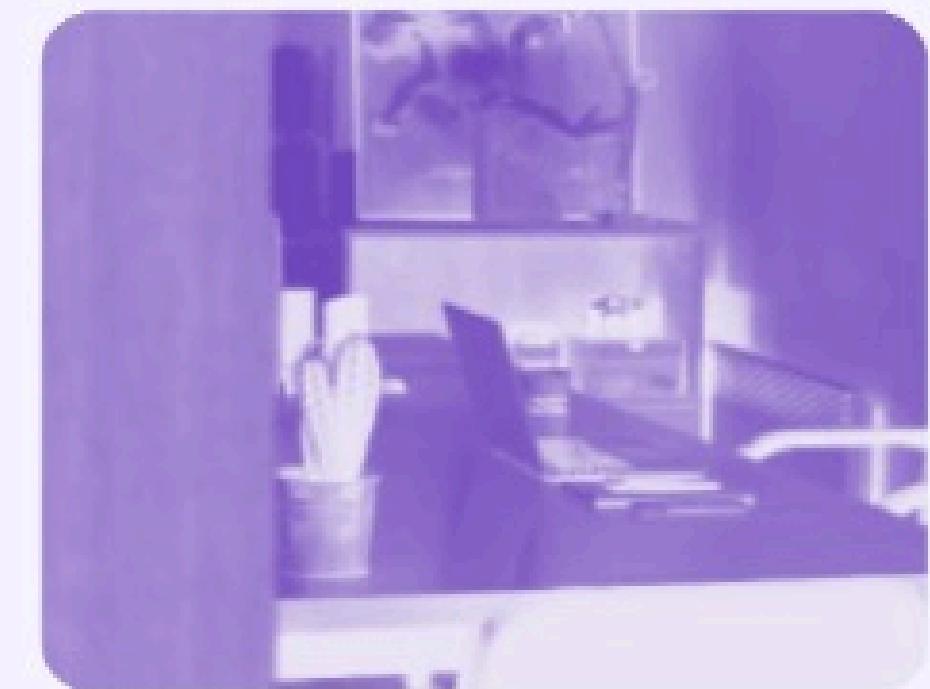
Git Utilization

Git was used for version control to track changes and manage collaboration effectively with other team members.



Branching Strategy

Development followed a structured branching strategy for feature development, bug fixes, and releases.



Testing and Validation



Regression Testing

Regression testing implemented to ensure continuous accuracy of the data pipeline after updates is carried out.



Testing Framework

Pytest framework integrated for automated testing and data validation.



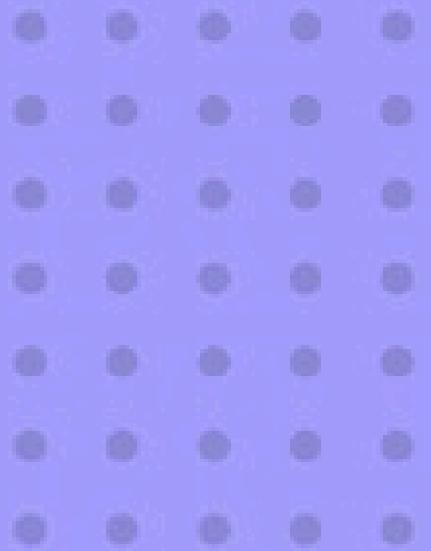
Testing and Validation

Data transformation testing ensures the project's reliability and stability.

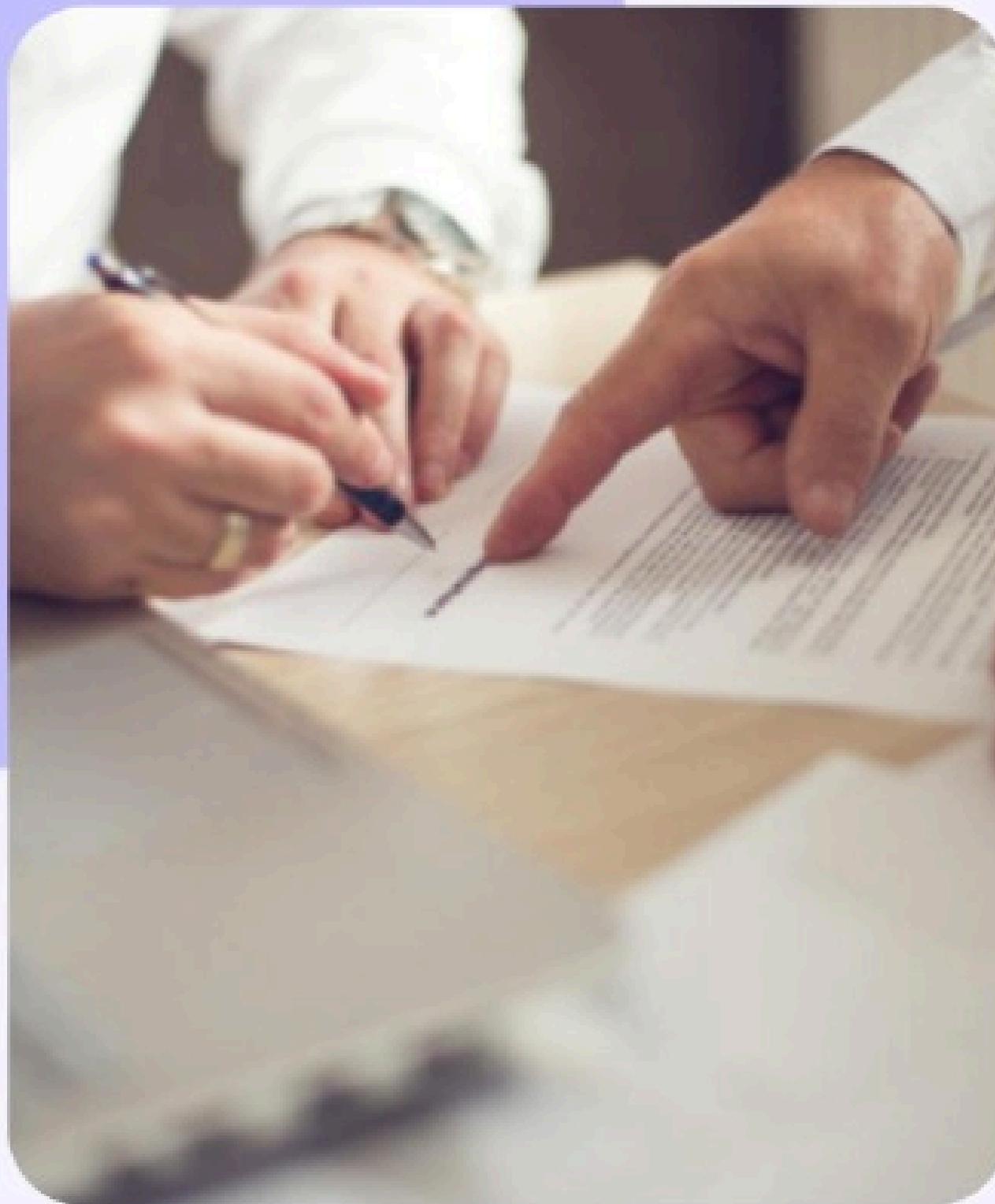


06.

Future Enhancements



Machine Learning Integration



● Predictive Modeling

Machine is to be integrated for predictive insights to forecast transaction patterns and user behavior.

● Anomaly Detection

Anomaly detection model to identify unusual transaction activities for risk management and fraud prevention.

Scalable Cloud Databases

Transition Plan

Migration to scalable cloud databases is being planned (PostgreSQL or Supabase) to handle larger volumes of data and higher traffic.

Cloud Benefits

Benefits associated with faster data processing and enhanced scalability will be delivered through the transition.

Live Data Ingestion

API Development

APIs are being developed now to enable live data ingestion to ensure continuous analysis of transactional data.



Real-Time Insights

The live data ingestion ensures real-time monitoring and proactive response to emerging trends.



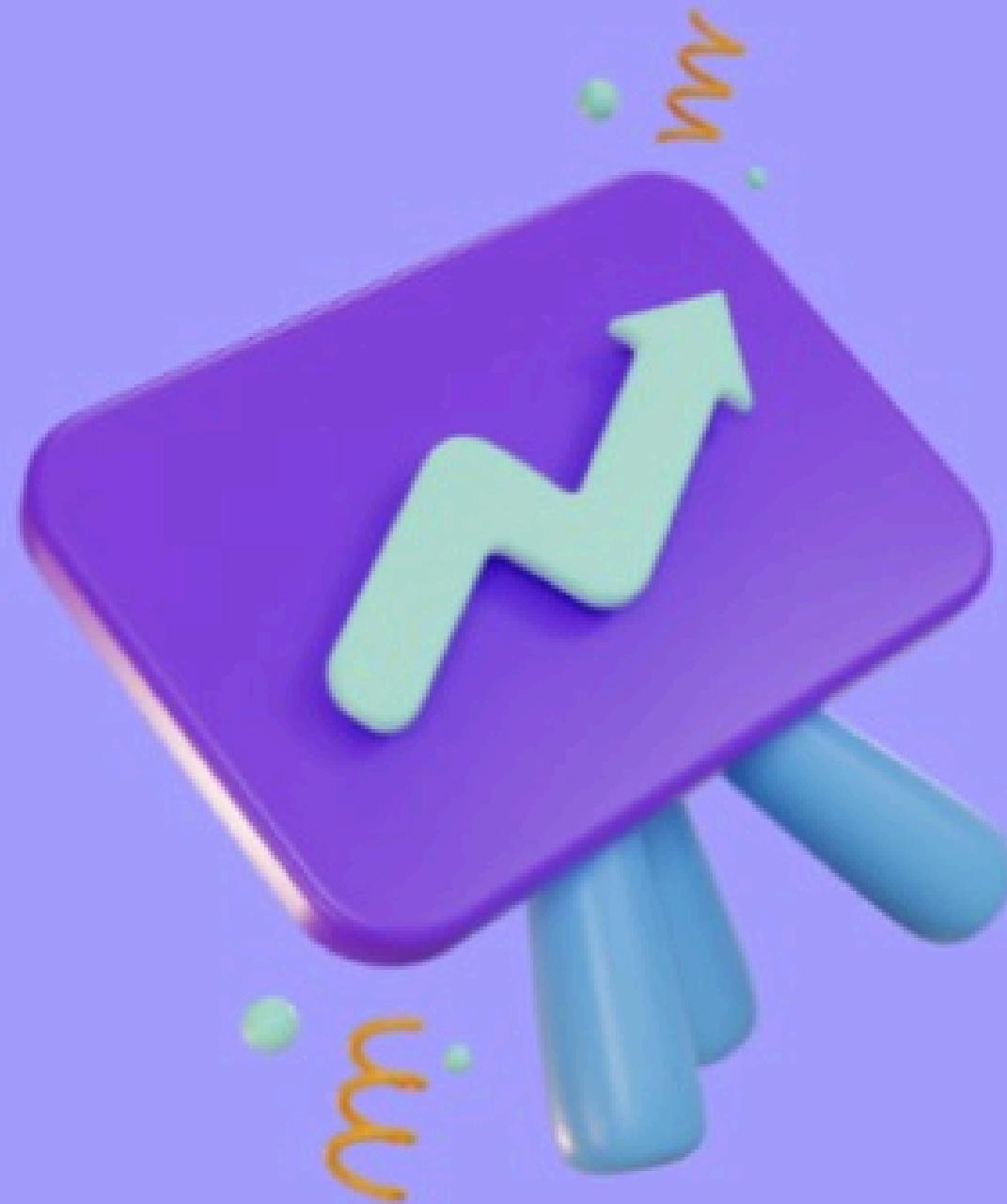
Docker Deployment

Cloud Platform

The entire pipeline will be deployed using Docker on cloud platforms to streamline deployment and management.

Easy Deployment

The cloud deployment benefits will lead to easier scaling and maintenance of the dashboard.



07.

Conclusion & Final Deliverables

Project Summary

01

Recap of Work

The project presents a complete data analytics pipeline—from extracting nested JSON data to building a live, interactive dashboard.

02

Key Deliverables

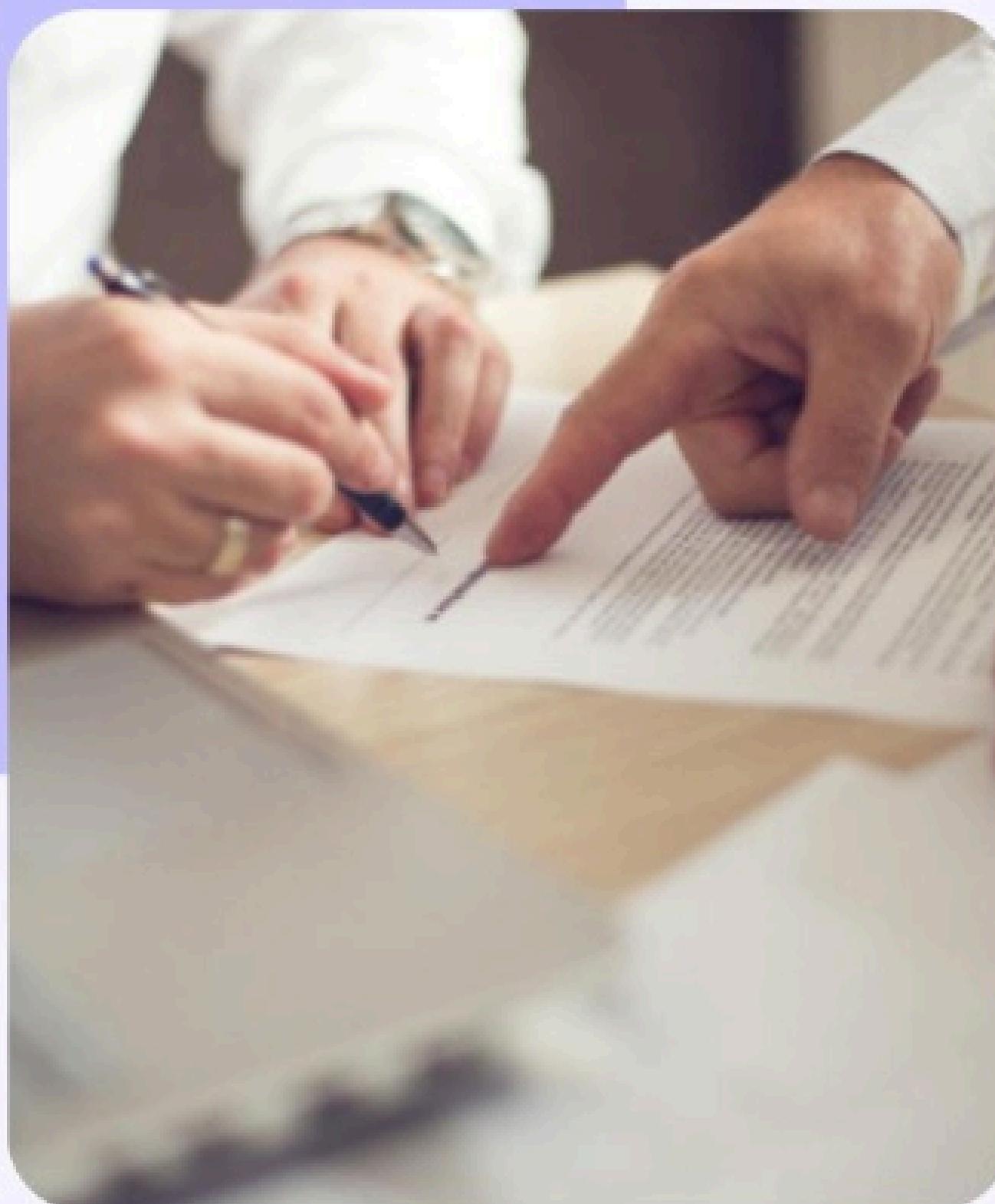
Deliverables include Python source code for ETL, SQL scripts for database schema and analytical queries, and a Streamlit application.

03

Documentation

Comprehensive documentation is provided with the project, which includes user guides, technical documentation, and presentation slides.

Key Takeaways



Understanding Trends

The dashboard reveals important patterns such as the rapid growth of UPI transactions in Tier-2 and Tier-3 cities and uneven adoption of insurance products.

Demonstrating Data Conversion

This project demonstrates how open-source data can be transformed into meaningful business intelligence with thorough statistical analysis.

Impact and Value

01

Stakeholders Benefit

The project delivers real-world value for analysts, fintech strategists, and policy planners interested in India's digital financial landscape.



02

Useful Platform

The dashboard will be leveraged in decision-making and strategic planning.



03

Further Research

The project lays the groundwork for further research and development in fintech innovation and data analytics.



Thank you
for watching.

Reporter

