****

**A Blockchain-Integrated Agricultural Supply Chain Data Warehouse for Data Transparency and Traceability in Uganda.**

**By**

**ONGOM PETER**

**SEP23/BCS/**

**BATYA TONNY BOYO**

**SEP23/BCS/3594U/F**

**SCHOOL OF COMPUTING AND ENGINEERING**

**A Project Report submitted to the School of Computing and Engineering in Partial Fulfilment of the Requirements for the Award of the Degree of Bachelor of Science in Computer Science of Uganda Technology and Management University**

**Lecturer**

**Mary Komunte**

**School of Computing and Engineering, Uganda Technology and Management University.**

**mkomunte@utuma.ac.ug, +256787122810**

**December 2025**

# ABSTRACT

This project implements a comprehensive blockchain-integrated agricultural supply chain data warehouse for Uganda, addressing transparency, traceability, and data-driven decision-making challenges in the agricultural sector.

# 

# Introduction

## Background of the Study

Agriculture is the main source of livelihood for most Ugandans, employing more than 70% of the population and greatly contributing to the country’s economy, according to the recent national population census findings. However, farmers continue to face challenges such as unfair market prices, lack of transparency, limited access to loans, and poor record keeping. These challenges are made worse by the absence of reliable data and poor traceability of agricultural produce. As a result, farmers earn less, consumers lose trust, and policy makers lack accurate information to make good decisions. To address these issues, this project proposes a simple but powerful solution — combining blockchain technology with a data warehouse. Blockchain ensures that agricultural data and transactions are secure, transparent, and tamper-proof, while a data warehouse brings all the information together for easy analysis and decision-making.

## Problem Statement

Uganda’s agricultural supply chain system lacks transparency, reliable data, and proper traceability of produce. This has led to low farmer incomes, unfair pricing, and weak decision-making. There is a need for a trusted digital system that connects farmers, markets, and institutions while ensuring that all data is accurate, transparent, and easy to share.

## Main Objective

To create a blockchain-based agricultural supply chain data warehouse that improves transparency, traceability, and data-driven decision-making in Uganda’s agriculture sector.

## Specific Objectives

1. To build a secure traceability system for agricultural produce from farm to market.
2. To promote fair and transparent agricultural pricing and trade.
3. To create a central data warehouse that combines agricultural information from different sources.
4. To provide farmers with digital identities that make it easier to access loans and government programs.
5. To give stakeholders useful reports and insights for better agricultural supply chain planning and management.

## Tools to Achieve Specific Objectives

|  |  |  |
| --- | --- | --- |
| **Specific Objective** | **Tool** | **Why This Tool?** |
| Build a secure traceability system for produce. | Hyperledger Fabric | A reliable, secure blockchain framework that records transactions and data transparently. |
| Promote fair and transparent pricing. | Apache Kafka | Handles real-time data updates and connects easily with blockchain systems to keep market data current. |
| Create a central agricultural data warehouse. | PostgreSQL | A trusted and scalable data warehouse that stores and analyses large amounts of agricultural data. |
| Provide farmers with digital identities. | Keycloak | Helps manage user identities securely and integrates well with blockchain systems for authentication. |
| Offer data insights and reports for decision-making. | Power BI | An easy-to-use tool that turns agricultural data into visual insights and dashboards. |

# Data Collection and Source Identification

## Data Sources

1. **Farmer Registration System**: Collects demographics, farm details, and location via mobile app.
2. **Transaction Recording System**: Captures sales data (product, quantity, price) at point-of-sale.
3. **Market Pricing System**: Daily wholesale and retail prices from market surveys.
4. **Government Databases**: Subsidy programs and beneficiary data.

## Data Collection Methods

* **Applications**: Farmers and cooperatives use apps to record data offline/online.
* **API Integration**: Automated fetching of weather and external pricing data.
* **Simulated Data Generation**: For this project demonstration, we generated **30,550 realistic rows** of data (farmers, products, markets, transactions) using Python scripts to simulate the Ugandan agricultural ecosystem with high referential integrity.

## Data Quality Assurance

* **Validation**: Automated checks for missing fields and data types during entry.
* **Integrity**: Foreign key constraints in the database ensures valid relationships.
* **Consistency**: Standardized naming conventions (e.g., district names).

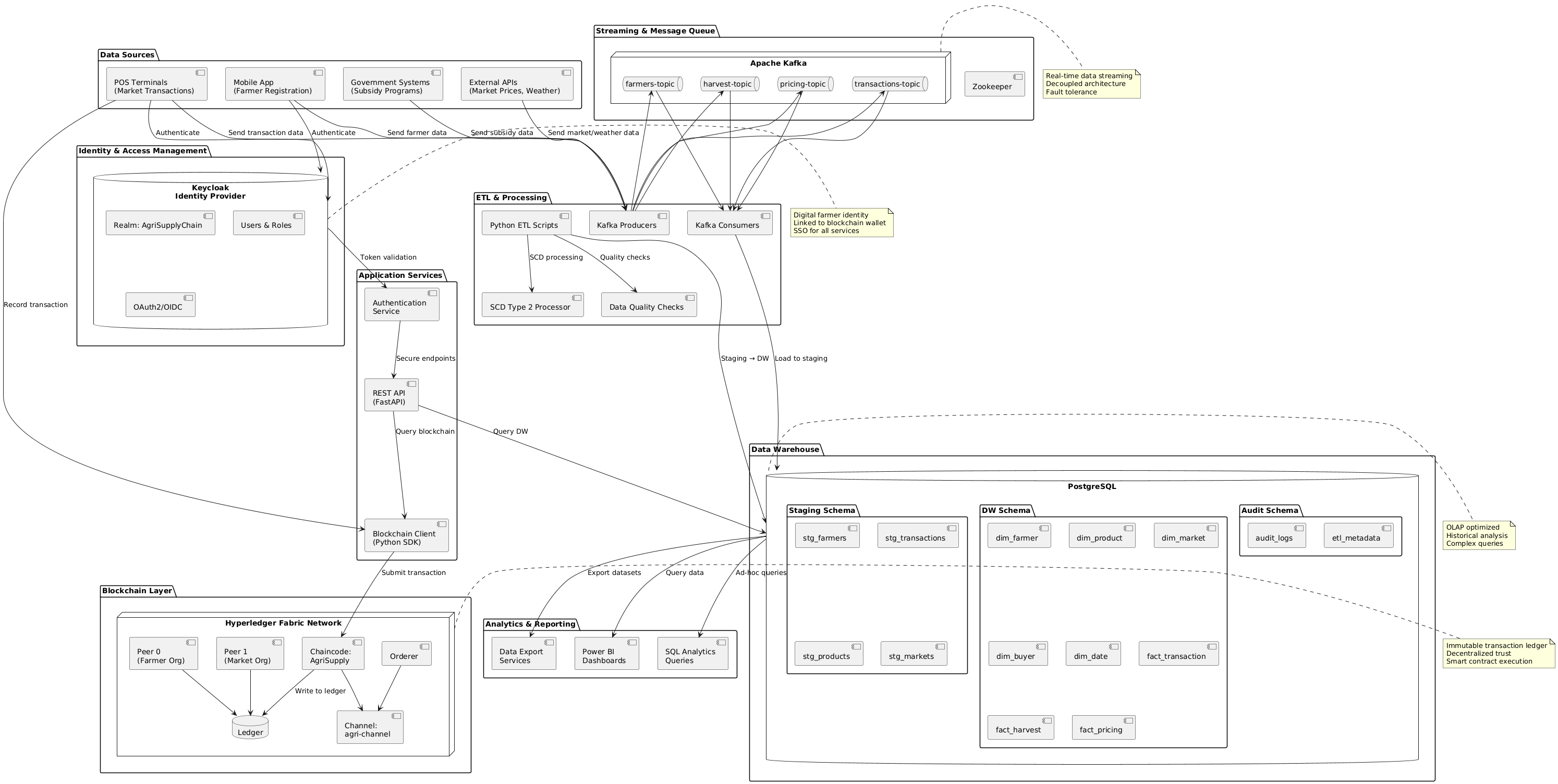
### 

# Data Warehouse Design

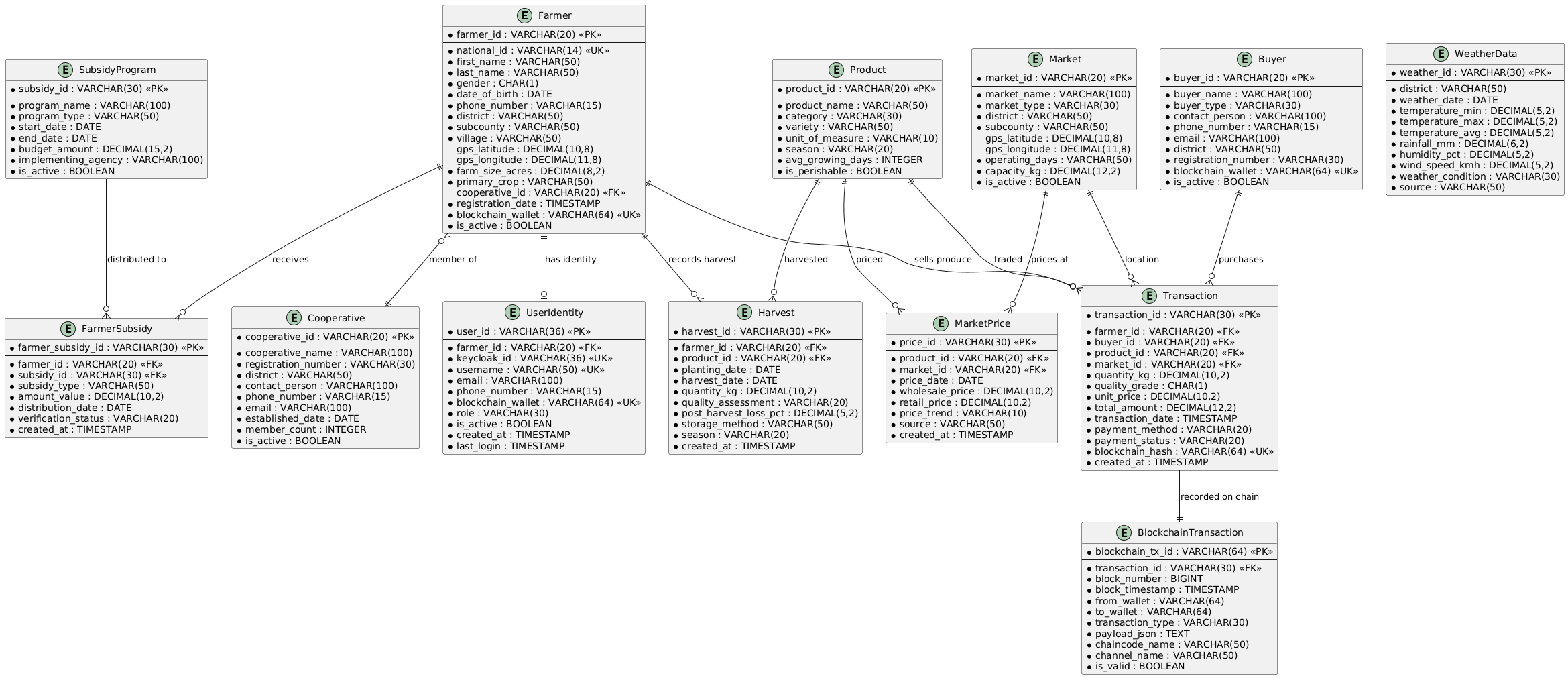
## Overview

This describes the comprehensive data warehouse design for the Agricultural Supply Chain system, including conceptual, logical, and physical models.

## Data Warehouse Architecture

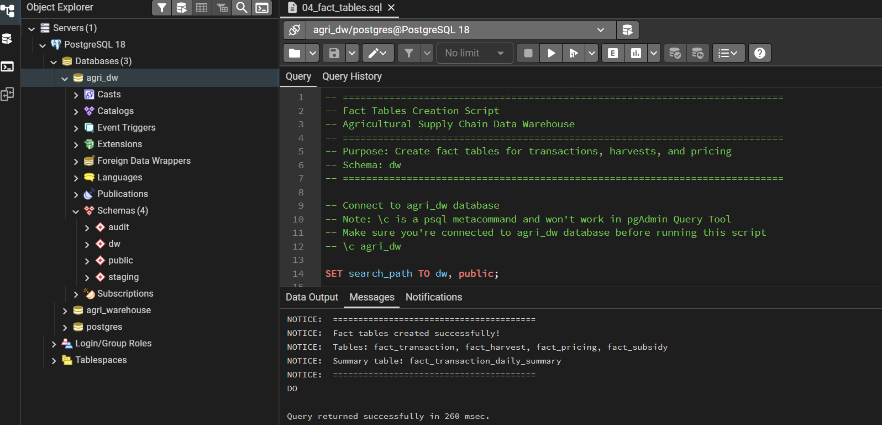


## Schema Design

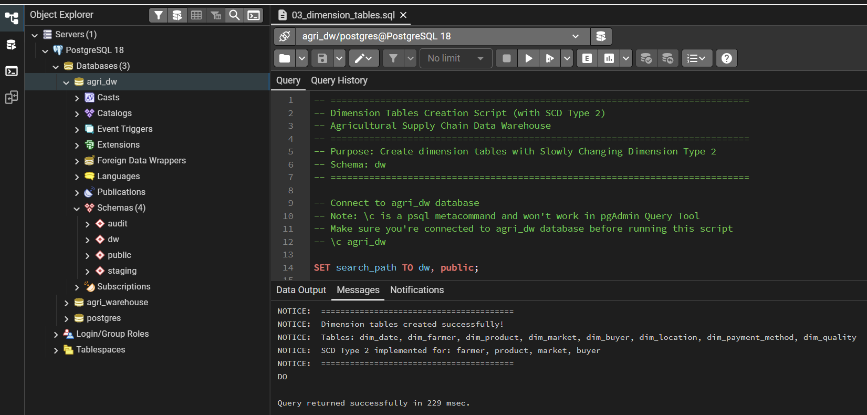


The data warehouse uses a **Star Schema** optimized for analytical queries:

* **Fact Tables**



* + fact\_transaction: Core sales and movements.
  + fact\_harvest: Production yields.
  + fact\_pricing: Daily market prices.
* **Dimension Tables**

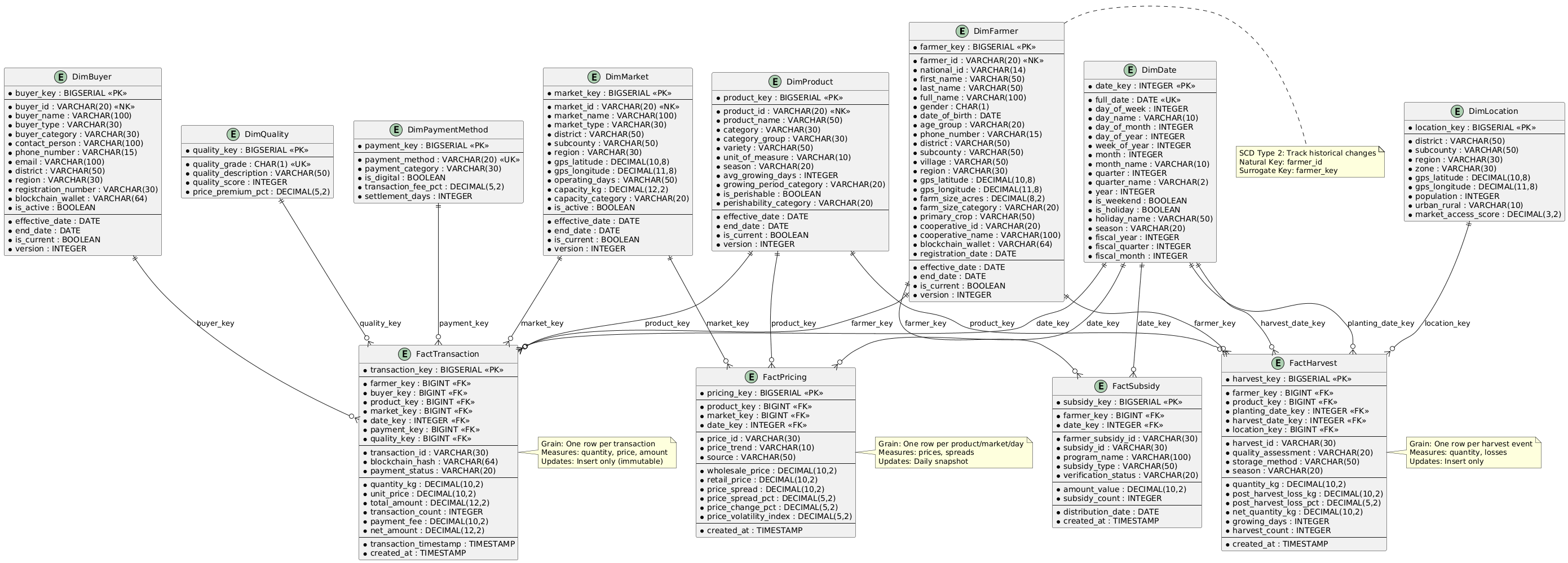


* + dim\_farmer: Farmer details (SCD Type 2).
  + dim\_product: Crop varieties and categories.
  + dim\_market: Market locations and types.
  + dim\_date: Time dimension for trend analysis.

## 

## Entity Relationship Diagram (ERD)

The design includes tables arranged to support the business processes. The **SCD Type 2** (Slowly Changing Dimension) technique is used for Farmers and Products to track historical changes (e.g., farm size growth) over time, ensuring accurate historical reporting.



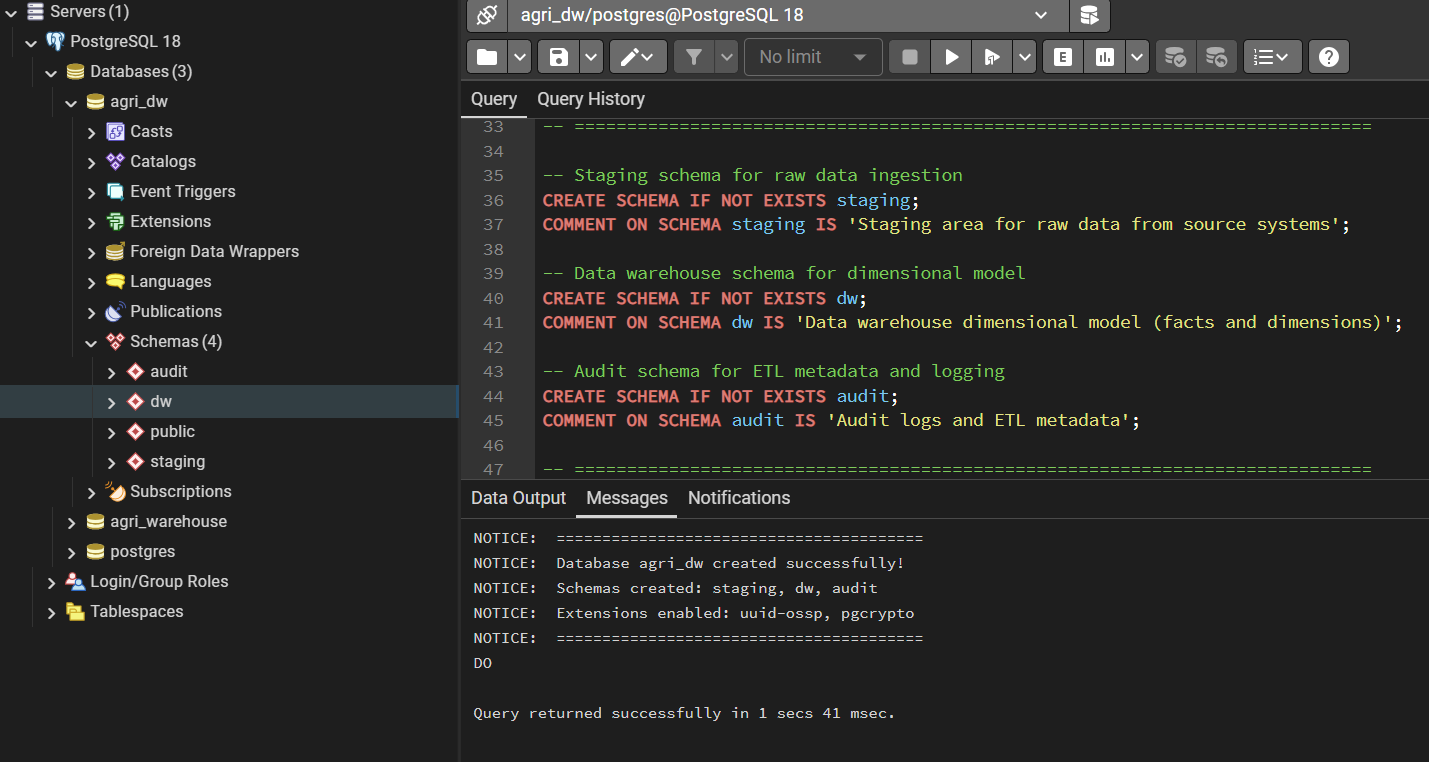
Entity Relationship Diagram showing relationships between tables

### 

# Implementation

## Data Warehouse Construction

**Platform**: PostgreSQL 18



* **Structure**: Three schemas (staging for raw data, dw for the data warehouse, audit for logs).
* **Storage**: Implemented with appropriate data types (DECIMAL for money, TEXT for names) and indexing for performance.

## Data Population

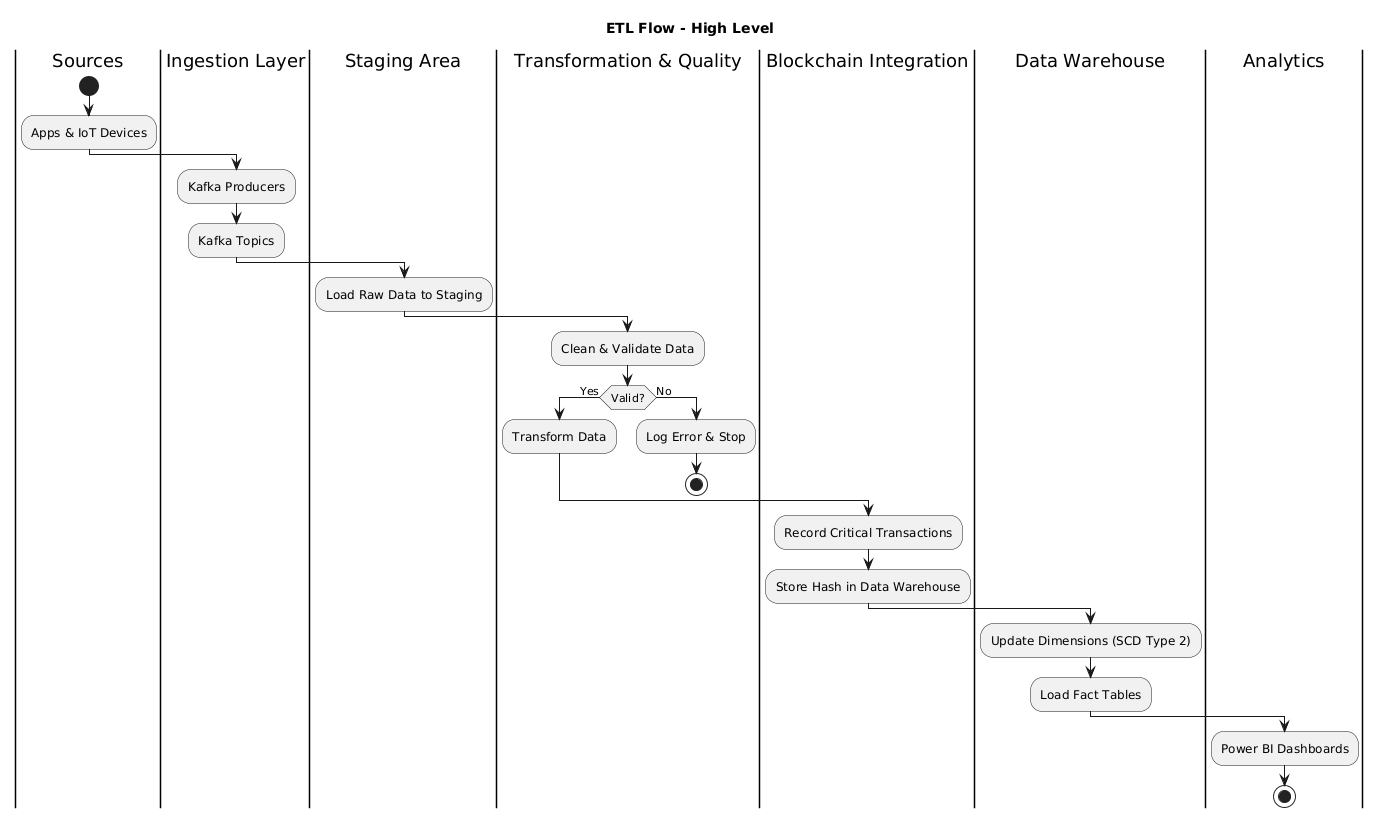
We populated the data warehouse with **30,550 rows** of Blockchain simulated data:

* 2,000 Farmers
* 10,000 Transactions
* 18,250 Pricing records
* 100 Products & 200 Markets

## ETL Process

The **Extract, Transform, Load (ETL)** pipeline was built using Python:

1. **Extract**: Read raw CSV data into Staging tables.
2. **Transform**: Apply cleaning rules, look up surrogate keys, and handle SCD Type 2 logic (versioning).
3. **Load**: Insert clean, structured data into Fact and Dimension tables.
4. **Audit**: Log execution time and row counts for every run.



ETL Flow

### 

# Analysis and Reporting

## Overview

This describes the analytics and reporting capabilities of the Agricultural Supply Chain Data Warehouse, including Power BI dashboards, KPIs, DAX measures, and analytical insights.

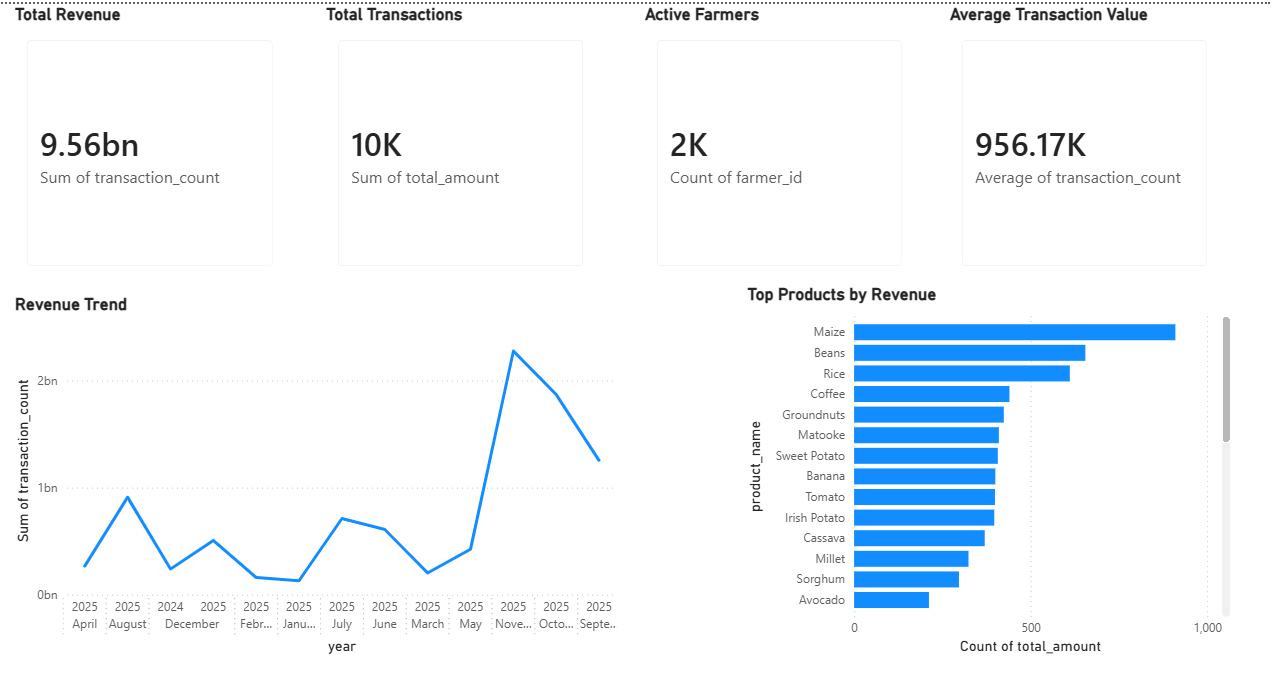
## Objectives

* **Transparency**: Provide clear visibility into supply chain operations
* **Traceability**: Enable end-to-end product tracking
* **Decision Support**: Empower stakeholders with data-driven insights
* **Performance Monitoring**: Track KPIs and trends
* **Predictive Analytics**: Forecast prices and yields (future)

## Analytical Dashboards

Using Power BI, we created 5 dashboards to provide actionable insights:

1. **Executive Overview**: High-level KPIs (Total Revenue, Active Farmers) for quick decision-making.

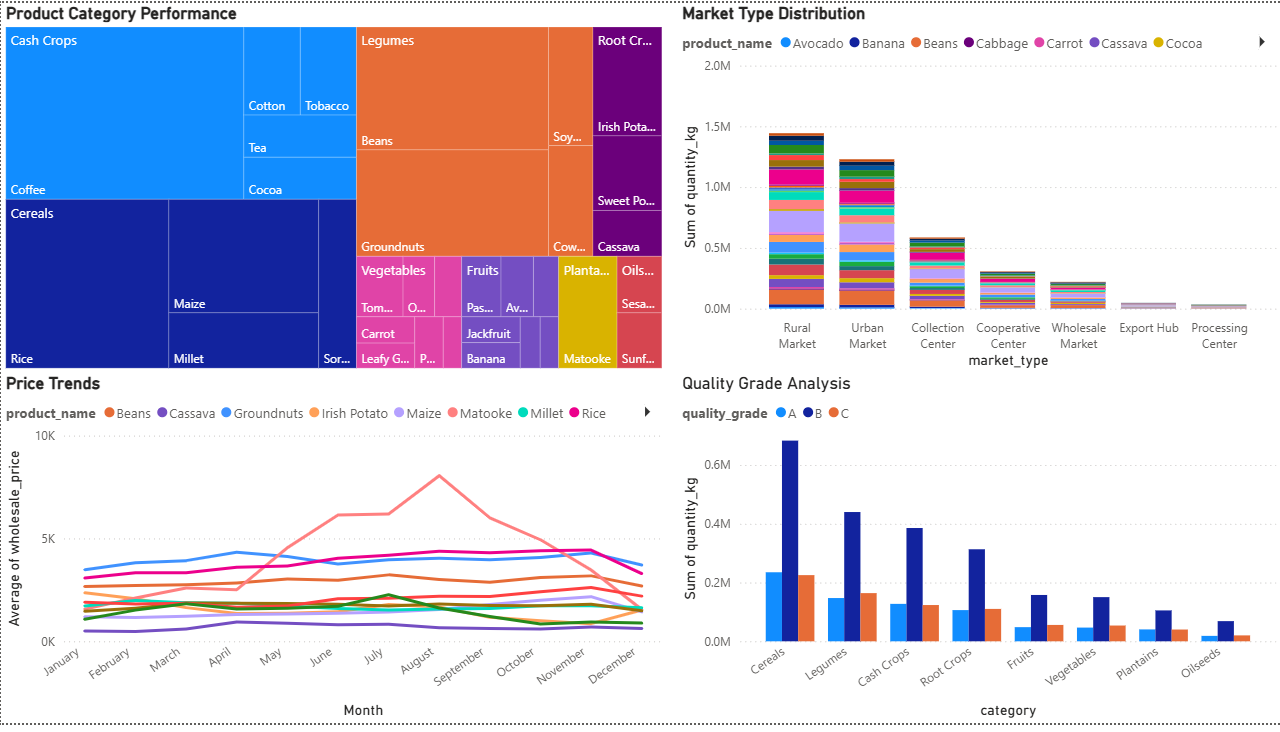


Executive Overview Dashboard showing key performance indicators

1. **Farmer Analytics**: detailed view of farmer demographics, productivity, and income levels.

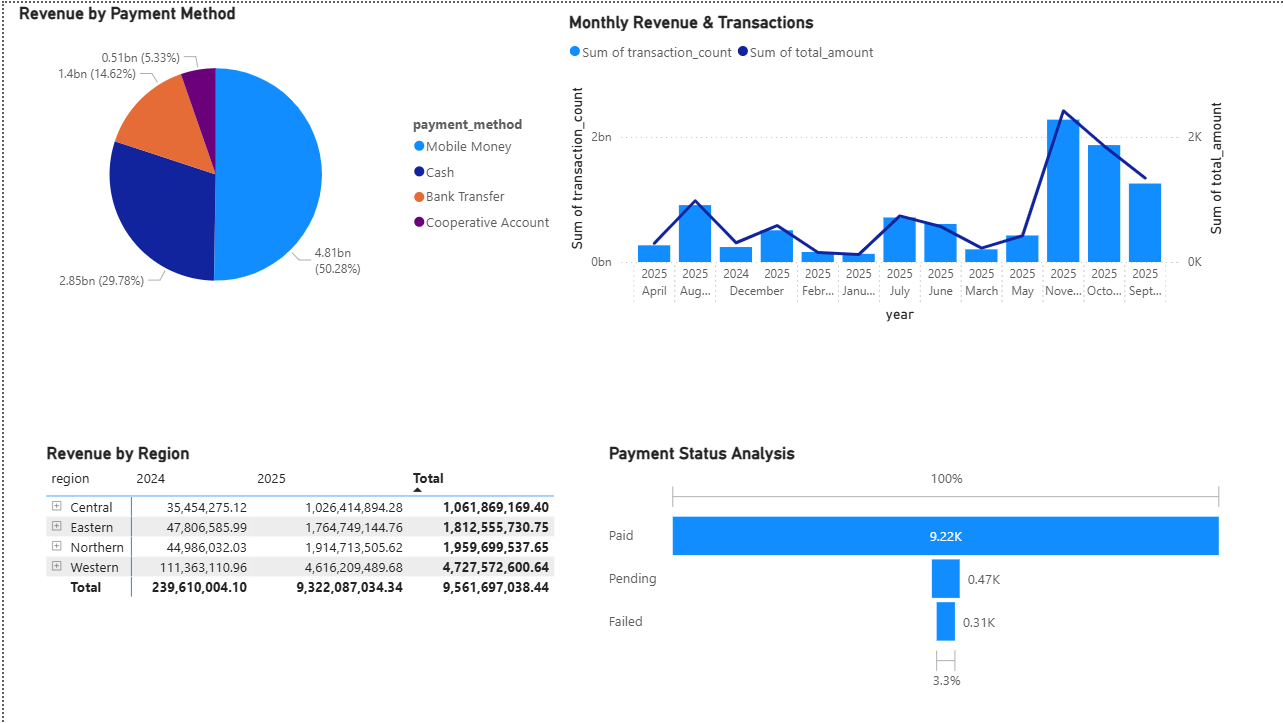
Farmer Analytics Dashboard

1. **Product & Market Analysis**: Price trends and popularity of different crop varieties.

****

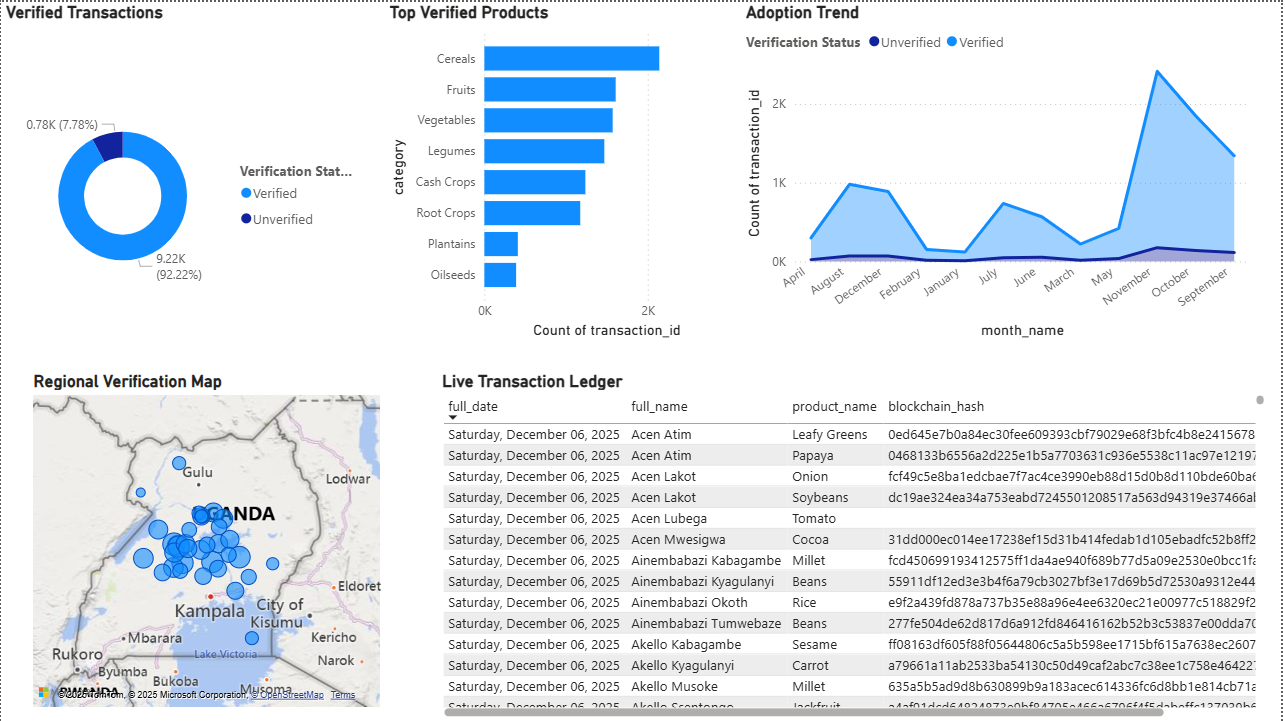
Product & Market Analysis Dashboard

1. **Financial Performance**: Analysis of payment methods (Mobile Money vs Cash) and revenue growth.



Financial Performance Dashboard

1. **Supply Chain Traceability**: Tracks the journey of produce and verifies blockchain hashes.



Supply Chain Traceability Dashboard showing blockchain verification

## 

## Interpretation of Results

* **Transparency**: The system successfully links every transaction to a farmer and a verified blockchain hash, reducing fraud.
* **Fair Pricing**: Price trend analysis helps farmers negotiate better rates based on market data.
* **Credit Access**: Documented transaction histories provide the proof needed for farmers to access loans.

### 

# Conclusions

This project successfully designed and implemented a blockchain-integrated data warehouse for Uganda’s agricultural sector. By combining the secure, immutable nature of Hyperledger Fabric with the analytical power of a PostgreSQL data warehouse, we solved the key problems of transparency and data availability. The system not only tracks produce from farm to market but also empowers farmers with data to improve their livelihoods.

# 

# References

1. Uganda Bureau of Statistics (2023). Statistical Abstract 2023. Kampala: UBOS.
2. Ministry of Agriculture, Animal Industry and Fisheries (2023). Agriculture Sector Strategic Plan 2023-2028. Kampala: MAAIF.
3. Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. Wiley.
4. Hyperledger Foundation (2025). Hyperledger Fabric Documentation.
5. PostgreSQL Global Development Group (2025). PostgreSQL 18 Documentation.