البنك التجاري اليمني
Yemen Commercial Bank

Data Bank Project Results Report

Executive Summary

The Data Bank Project has made significant progress in establishing a robust data infrastructure to unify banking system data. To date, we have successfully implemented key components including database setup, data extraction pipelines, transformation processes, and staging for a unified data model. This foundation will enable advanced analytics and AI-driven applications such as YCB BI Analytics, YCB Agents, and YCB Bots. This report outlines the achievements and sets the stage for next steps.

Project Overview

The Data Bank Project aims to create a centralized, unified data model that integrates data from various banking systems. This model will serve as a single source of truth, facilitating efficient data processing, analysis, and utilization across multiple applications. The project leverages modern tools like Oracle, Mage AI, and PostgreSQL to handle data extraction, transformation, loading (ETL), and storage.

Key objectives include:

- Automating data fetching and processing to reduce manual efforts.
- Ensuring data quality and consistency through transformation.
- Preparing data for downstream consumption in analytics and AI tools.

Achieved Results

As of August 11, 2025, the team has accomplished the following milestones:

1. **Oracle Database Setup**
   We have successfully configured and deployed a local Oracle database instance. This will serve as the primary repository for raw data from banking systems, providing a scalable and secure foundation for data storage.
2. **Mage AI Pipeline Implementation**
   A local Mage AI-based ELT pipeline has been developed . This pipeline performs the following functions:
   - **Batch and Scheduled Data Extraction:** Data is extracted in batches from the banking systems on a scheduled basis, ensuring timely and consistent data ingestion without disrupting operational systems.
   - **Data Preprocessing and Transformation:** Within Mage AI, the extracted data undergoes preprocessing steps, including cleaning, normalization, and transformation to align with standardized formats. This step addresses inconsistencies across different banking sources.
3. **Export to PostgreSQL Staging Area**
   The transformed data is exported to a dedicated staging area in PostgreSQL. This intermediate layer allows for validation and further refinement before final integration, minimizing risks to production environments.
4. **Unified Data Model Formation**
   Using the staged data in PostgreSQL, we are analyzing and planning to form a unified data model. This model consolidates data from all banking systems into a cohesive structure, eliminating silos and enabling holistic views.

The model is designed to support:

- **YCB BI Analytics:** For real-time business intelligence and reporting.
- **YCB Agents:** To power intelligent agents for customer interactions and automation.
- **YCB Bots:** For chatbot-driven services and operational efficiencies.
-

These achievements have been tested in a local development environment, with initial performance metrics showing reliable data flow and minimal latency in processing batches.

Benefits and Impact

The completed phases position the project for immediate value delivery:

- **Improved Data Accessibility:** The unified model will reduce the query complexity and enhances data usability across departments.
- **Scalability:** Scheduled pipelines ensure the system can handle growing data volumes.
- **Foundation for AI and Analytics:** Direct integration with YCB tools will accelerate insights and innovation.

Potential risks mitigated include data inconsistencies and manual errors, with built-in monitoring in Mage AI for ongoing oversight.

Conclusion

The Data Bank Project has reached a critical milestone, with the core infrastructure now in place to support a unified data ecosystem. These accomplishments demonstrate the team's expertise and commitment, paving the way for enhanced banking operations through data-driven technologies. We look forward to your feedback and continued support.