

ETL Data Pipeline Development using Streamlit on GCP



Client Overview

The client is an African retail supply chain transformation company enabling small shops and retailers to adopt digital economy through mobile ordering & distribution systems, and enterprise services.

Business Requirements

- Build and manage the entire ETL (Extraction, Transformation & Loading) data pipeline and ETL integration with their existing ERP system, FMCG process and Loan application process.
- Develop Reports and Analytics to monitor Delivery agent's performance, Vehicle Utilization, Credit services and Sales activities.
- Conversion of existing environment (Java & Postgres) to Big data platform for better efficiencies.

Indium Solution

- Developed and implemented an ETL pipeline to manage the data using Big Query and Google Cloud Platform (Google Data Flow and Apache Beam)
- Developed Reports and Dashboards using Tableau covering end-to-end user activities.
- Integrated QuickBooks ERP, Loan Management System and their Existing Customer APP/Backend Applications into the ETL Pipeline (Big Query - GCP) for Analytics & Reporting
- Designed and Implemented Report automation (Self automated Reports for Daily/Weekly/Monthly) using Python & Streamlit.

Business Impact

- Runtime Reduction: 60 % reduction in ETL process runtime for Transition to a Cloud Big Data ecosystem
- Performance Optimization: 20 % Increase in performance efficiency with the developed solution that created visibility into business activities and performance metrics for each category
- Cost Reduction: Nearly 20 % reduction in delivery costs after implementation of delivery planning methods and visualization through Reports and insights



Tools



Google BigQuery



Tableau



Streamlit