Project Title: Cloud Based Data Engineering for Loonima Retail Store

Author: Mitul Makwana

Project Context: Data Engineering Cloud Architecture Project

Table of Contents

Overview

Loonima Retail Store is transitioning to a cloud-based data architecture to integrate diverse data sources, enhance operations, and enable real-time insights. This scalable solution combines data from POS systems, website tracking, and ecommerce platforms. It streamlines data processing and analysis using Azure services to drive better decision-making and improve business growth.

Objectives

1. Enhance Operational Efficiency: Streamline inventory management and sales processes.

2. Enable Data Driven Decision Making: Leverage analytics for better customer targeting and forecasting.

3. Implement RealTime Insights: Ensure timely actions based on sales trends and stock levels.

Cloud Architecture Design

Key Components:

Data Sources: POS systems, website tracking, and ecommerce platforms.

Ingestion Layer: Azure Data Factory for batch processing and Azure Event Hubs for streaming data.

Storage Layer:

Bronze Layer: Raw data storage.

Silver Layer: Cleaned and curated data.

Gold Layer: Aggregated and mature data.

Transformation Layer:

Azure Synapse Analytics: Handles data cleaning and formatting in the Silver Layer.

Databricks: Aggregates data and applies business rules in the Gold Layer.

Consumption Layer: Tableau, Power BI, machine learning models, and Azure Data Share for insights and reporting.

Data Pipeline Design

1. Source Layer:

Collects batch and streaming data from multiple sources.

Scheduled triggers (batch) and tumbling window triggers (streaming) feed data into ingestion services.

2. Ingestion Layer:

Azure Data Factory: Manages batch ingestion.

Azure Event Hubs: Handles real-time data streaming.

Linkedin

3. Storage Layer:

Data stored hierarchically in Bronze (raw), Silver (cleaned), and Gold (aggregated) layers.

4. Transformation Layer:

Silver Layer Transformation (Synapse Analytics): Deduplication, handling missing values, and formatting.

Gold Layer Transformation (Databricks): Aggregation, applying business rules, and consistency checks.

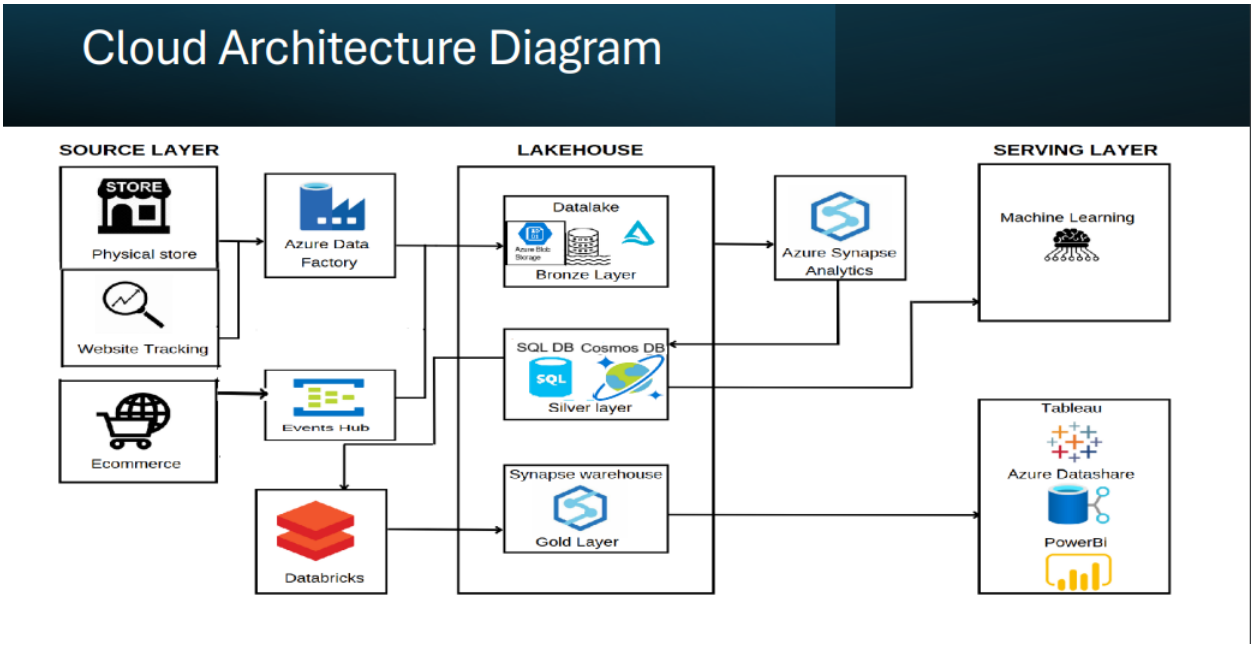
5. Consumption Layer:

Data visualizations and reports via Tableau and Power BI.

Machine learning models for predictive insights.

Diagram:

Cloud Architecture Diagram for Retail Store



Linkedin

Conclusion

The cloud-based architecture for Loonima Retail Store integrates diverse data sources to deliver actionable insights, enhancing operational efficiency and enabling data driven decision making. The use of Azure Data Factory, Synapse Analytics, and Databricks ensures seamless data processing from ingestion to transformation and visualization. This scalable solution empowers Loonima Retail Store to adapt to market demands and drive sustainable growth.

GitHub Repository Link: [Your Repository Link]

Linkedin