

Tahoe – Data Lake and Demand Forecasting

Executive Summary

BUSINESS NEEDS

A Modern Data Platform with a modular, flexible and secure data lake which consolidates multiple data silos to equip the user to forecast demand to take informed decisions on the stocking levels in the store at a product level to keep in alignment with the business and growth needs which ends stagnated inventory and over purchasing. Using Predictive Models, the stores can set appropriate inventory management plans which will optimize future replenishment and customer demand, while realigning inventory with sales growth

A modern data platform in Azure, which ingests the data from heterogeneous sources, prepare, process and persist the data for,

- Business Insights
- Inventory Optimization
- Downstream Consumptions

FEATURES

| Data Ingestion | Data Preparation & Processing | Data Storage | Orchestration | Data Consumption | CI/CD |
|--|---|--|--|--|--|
| Ingest the data from RDBMS, Flat files, Json files, APIs | Parse, curate and process the data for different consumptions | Store data at different levels of data processing - Raw, Prepared data, Processed data | Automate and Orchestrate the end-end pipelines | Consume the data for Visualization, Downstream applications, Predictive models | Automate the deployment with CI/CD pipelines |

BENEFITS

Unified view of data through Cloud data warehouse and Data Lake. Faster than traditional DWs

Achieve a Data and Analytics platform with industry standards that is agile towards futuristic needs

Versatile services available for varying needs of data management to solve the different business use cases and leverages elasticity and automation

Optimize Inventory for visibility on stocks for better planning at a SKU level taking a 360-degree approach for accurate forecasted numbers

Minimal effort required in maintenance of solution

Cost advantage on scale and pay-as-you-go service

Advanced visualization tool for integrated dashboards in client portal

Source Systems

- Order Management System
- Cost Savings Application
- PeopleSoft
- Solutions System

Ingestion

- Data Factory to ingest the data from On-Prem SQL Server, CSV & json files from SFTP location into Azure Data Lake raw storage
- Kafka to ingest the real-time transaction events into Azure Data Lake raw storage

Storage

- Azure Data Lake to store the raw and curated data, with appropriate partitions and file formats
- Azure Synapse stores the processed and aggregated data

Data Preparation

- Databricks prepares the data with necessary curation and persist the data as conformed datasets into Azure Data Lake

Data Aggregation

- Azure Synapse processes the curated data with necessary aggregations and business logics and persists the data for consumption

Orchestration

- Azure Data Factory orchestrates the end-end data pipelines
- Azure Logic Apps for notification

Data Consumption

Predictive Models

- o Models built and hosted in Databricks and outputs consumed by the web application

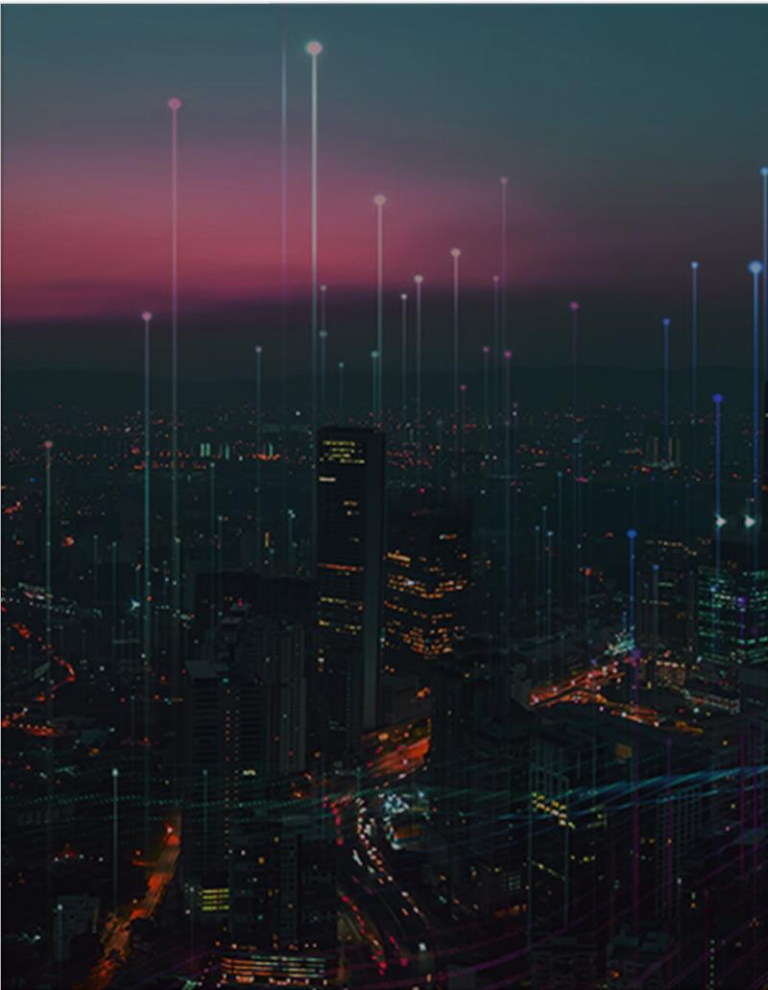
Visualization

- o Azure Analysis Services for Tabular model built on top of Synapse
- o Power BI for Visualization and Reporting

CI/CD

- Azure Repos as code repository along with Azure DevOps services helps for Continuous Integration and Deployment

Architecture Variations



| Solution | Description |
|----------------------|--|
| Real Time Dashboards | <p>Get insights in real-time using Power BI real-time Dashboards</p> <ul style="list-style-type: none">▪ Azure Stream Analytics Job will process the streaming data and feed to Power BI real-time datasets for visualization |
| Real Time Analytics | <p>Process the live streaming data from IoT devices, website click streams, applications in real-time</p> <ul style="list-style-type: none">▪ Ingestion options<ul style="list-style-type: none">• HDInsight Kafka, IoT Hub, Event Hub, API Management▪ Stream Processing options<ul style="list-style-type: none">• Stream Analytics Job, Azure Databricks |
| Advanced Analytics | <ul style="list-style-type: none">▪ Consume the data insights from real-time web and mobile apps using Cosmos DB▪ Derive deeper insights using machine learning / statistical techniques with Azure Databricks |

Happiest People · Happiest Customers

