

Experience with Global Food & Beverage Retailer

Data Engineering & Advanced Data Analytics Ecosystem Implementation

Retail & CPG

Data Engg. & Visualization

ADF, ADL, ADB, SQL, PBI

Customer Profile

Customer is a **Global multinational Food and Beverage Manufacturing and Retail conglomerate** with headquarters in Vevey, Switzerland. With CHF 92Bn revenue and 352 thousands employee base, it is the largest food company in the world, since 2014. Its product lines include baby food, medical food, bottled water, breakfast cereals, coffee and tea, confectionery, dairy products, ice cream, frozen food, pet foods, and snacks.

Problem Statement

- Client is currently handling the data analytics work load and data in SAP BW
- They have considered Azure Data lake Gen 2 as the storage account to handle the data
- So the raw data will be dumped into Azure data lake and from there, data has to be processed into subsequent layers
- This data includes SAP Master data and also transactional data
- As the data load is heavy and processing takes time but also few points like optimization, performance and accuracy needs to be considered
- Considering the use case it is required to load the raw compressed data to a target location with some filters and schema
- For every object in source has to be filtered to a particular site and loaded into target object folder

Implemented Solution

- Databricks is used to implement the solution over the source data to load the target with required business requirements
- Overwrite mode for daily refresh data
- Databricks delta for transactional Data, where in we can do an upsert
- Delta features like Vacuum to retain required history, Optimize and Z Order
- Target specific to market sites as parent folder and will contain all its objects to the specific site as child items

Value Delivery

- Better query performance and also enables block level compression which is considered as an advantage for data storage
- Easy optimized cluster management