Set up data infrastructure and deployed machine learning models using Microsoft fabric infrastructure

Health and wellness products brand

Designed and developed a Data Warehouse Architecture in a robust, scalable and flexible platform (Microsoft Fabric) that meets the business requirements of the client.

Leveraged Microsoft Synapse Data Science (within the Microsoft Fabric ecosystem) to run a R-based regression model to generate sales projections.

Health and wellness product manufacturing company needs to set up data infrastructure

SET UP DATA INFRASTRUCTURE AND DEPLOYED MACHINE LEARNING MODELS USING MICROSOFT FABRIC INFRASTRUCTURE

Picture this…

You're looking to assess the business requirements and designed/deployed an enterprise data warehouse leveraging Microsoft Fabric. Also, to deploy predictive models (for sales projections, marketing-mix modeling, etc.) using Microsoft Synapse Data Science within the Fabric environment.

You turn to Accordion.

- We partner with your team to designed and developed a Data Warehouse Architecture in a robust, scalable and flexible platform (Microsoft Fabric) that meets the business requirements of the client. Leveraged Microsoft Synapse Data Science (within the Microsoft Fabric ecosystem) to run a R-based regression model to generate sales projections, including:

1) Developing a future state data architecture on Microsoft Fabric with optimized tools stack, providing enhanced cost control and making it more user friendly. The deployed enterprise data warehouse is based on medallion architecture (Raw, Clean, Aggregated & Semantic layers) and follows industry best practices by leveraging OneLake, Data Factory, Data Flows and Notebook (Microsoft Synapse Data Science).
2) Setting up end-to-end ETL flows (extract, transform, and load) by ingesting data from multiple sources (Navision ERP, Google Analytics, etc.) and standardized the data management processes to address the complexities of organizing and processing data, and implemented data quality measures to standardize the data types for data ingested from multiple systems.
3) Deploying predictive models using R-based algorithms for sales projections, marketing-mix modeling etc. leveraging Notebook (Microsoft Synapse Data Science) and visualized the results on a Power BI dashboard.

Your value is enhanced.

You have the executive team that has access to reliable performance reporting across the company. You have automated the ETL processes (extract, transform, load) and advanced analytics enabled the marketing function to take real-time decisions to improve the effectiveness of sales campaigns and allocate budget. You also have the shift from on-premises to cloud-based solution also enabled a more cost-effective data and reporting infrastructure and seamless monitoring/maintenance.

KEY RESULT

- ~80% of man-hours reduction
- 90% reduction in on-prem server maintenance and upkeep

VALUE LEVERS PULLED

- Designed and developed a Data Warehouse Architecture in a robust, scalable and flexible platform
- Leveraged Microsoft Synapse Data Science) to run a R-based regression model

Data architecture and reporting infrastructure assessment

Situation

- Client was leveraging disparate systems for each department such as Navision ERP for sales, BigQuery for Google Analytics, AWS Aurora for ecommerce, Snowflake for Amazon marketplace, and several manual Excel files. Due to the complexity in handling several systems and manual intervention required to process the data, the client was looking for an automated solution to replicate the current processes and reporting.
- Partnered with the client to assess the business requirements and designed/deployed an Enterprise Data Warehouse leveraging Microsoft Fabric. Also, we deployed predictive models (for sales projections, marketing-mix modeling, etc.) using Microsoft Synapse Data Science within the Fabric environment.

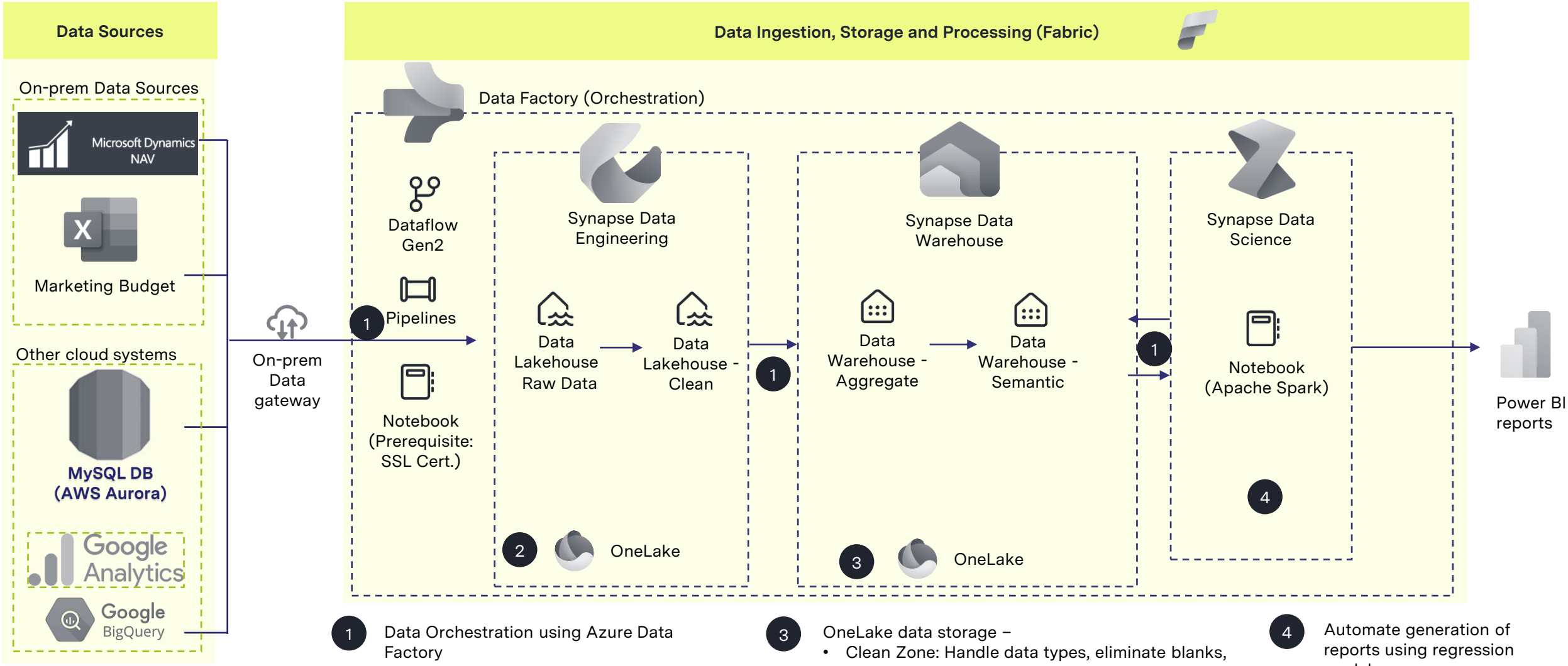
Accordion Value Add

- Developed a future state data architecture on Microsoft Fabric with optimized tools stack, providing enhanced cost control and making it more user friendly. The deployed Enterprise Data Warehouse is based on medallion architecture (Raw, Clean, Aggregate & Semantic layers) and follows industry best practices by leveraging OneLake, Data Factory, Data Flows and Notebook (Microsoft Synapse Data Science).
- Set up end-to-end ETL flows (extract, transform, and load) by ingesting data from multiple sources (Navision ERP, Google Analytics, etc.) and standardized the data management processes to address the complexities of organizing and processing data, and implemented data quality measures to standardize the data types for data ingested from multiple systems.
- Deployed predictive models using R-based algorithms for sales projections, marketing-mix modeling etc. leveraging Notebook (Microsoft Synapse Data Science) and visualized the results on a Power BI dashboard

Impact

- The new single source of truth enabled the executive team to have access to reliable performance reporting across the company
- Potential reduction in man-hours by up to 80%
- Potential reduction in man-hours by up to 90% due to reduction in on-prem server maintenance and upkeep

Implemented Microsoft Fabric achitecture



1. Data Orchestration using Azure Data Factory

2. OneLake data storage –
   - Raw Zone: Stores data from sources in parquet file format

3. OneLake data storage –
   - Clean Zone: Handle data types, eliminate blanks, remove duplicates
   - Aggregate Zone: Establish dimensions and facts
   - Semantic Zone: Utilize this data to perform analysis

4. Automate generation of reports using regression models