### **Use Case: Self-Service BI for a Retail Company**

**Scenario:**A retail company has implemented Azure Data Lake Storage as its central data repository. The data lake stores information about sales, inventory, customer behavior, and supply chain activities. The company now wants to empower its business users, such as store managers, regional managers, and marketing teams, to analyze and visualize data using Power BI without relying on IT for every report or insight.

### **Current Structure:**

#### **1. Data Organization in Azure Data Lake**

* **Raw Data Layer:**Ingest data from various sources such as POS systems, e-commerce platforms, ERP systems, and customer databases.
  + Example: Store daily sales data in JSON files and customer purchase history in Parquet.
* **Transformed Data Layer:**Use tools like Azure Synapse or Databricks to clean, transform, and integrate data.
  + Example: Aggregate sales data at daily, weekly, and monthly levels for easier analysis.
* **Curated Data Layer:**Create business-friendly datasets with clear structures.
  + Example: A dataset containing product-level sales, inventory, and profit margins for each store.

### **Goals**

1. **Enable Self-Service Analytics:** Provide business users access to clean, curated datasets.
2. **Improve Decision-Making:** Allow teams to derive actionable insights from sales, inventory, and customer data.
3. **Centralize Metrics:** Create consistent KPIs across teams to avoid discrepancies in reporting.
4. **Ensure Data Governance:** Maintain data security and compliance while enabling easy access.

### 

### 

### 

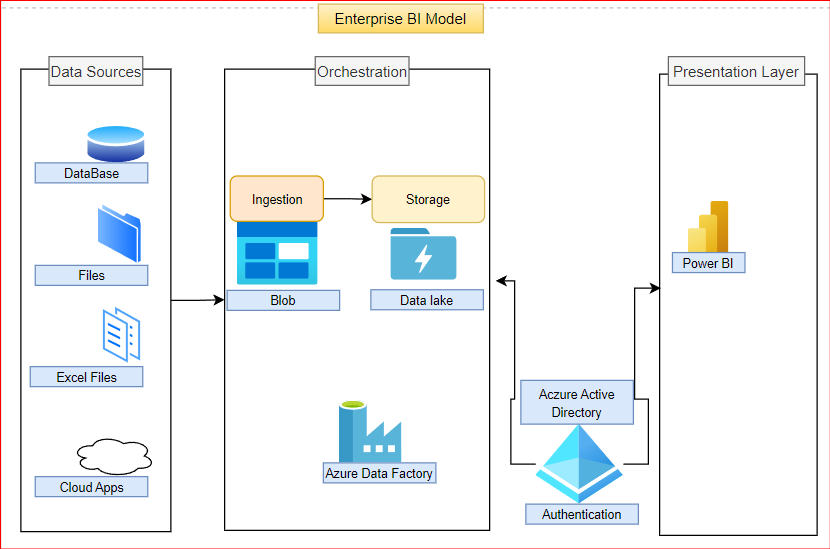
### 

### 

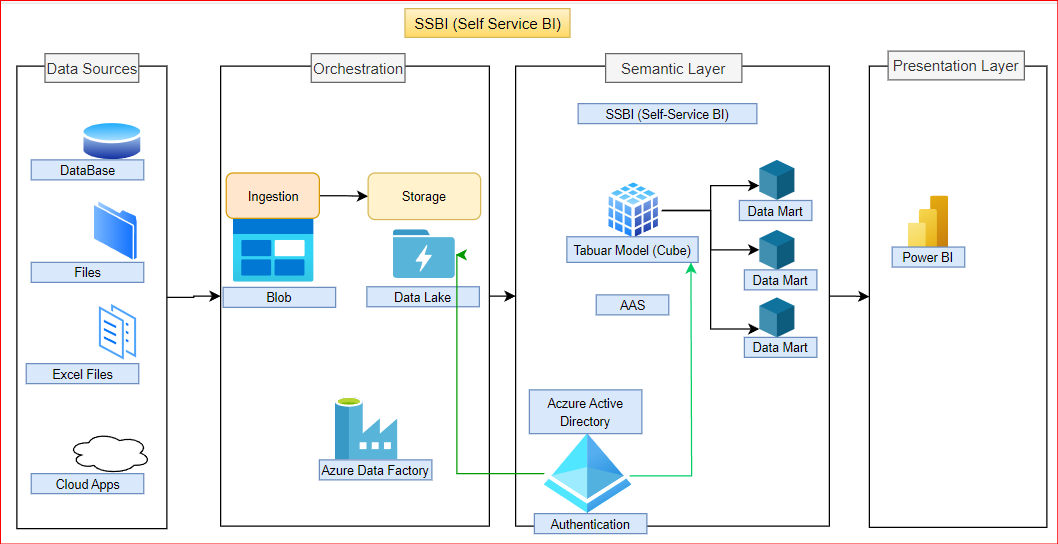
### 

#### **Architecture Diagram:**

Current Architecture: (Enterprise Data Model)



Proposed Architecture: (SSBI)



### **Architecture Layers Explanation**

1. **Data Sources**
   * **Purpose:** Raw data ingestion from multiple sources.
   * **Examples:**
     + **Databases:** Transactional systems like SQL Server, Oracle, or MySQL.
     + **Files:** CSV, Parquet, JSON, or XML files.
     + **Excel Files:** Ad-hoc data managed by business users.
     + **Cloud Apps:** SaaS platforms like Salesforce, Dynamics 365, or Google Analytics.
2. **Orchestration Layer**
   * **Components:**
     + **Ingestion:** Data is ingested into storage from various sources using tools like:
       - **Azure Data Factory (ADF):** Orchestrates the movement of data.
       - **Blob Storage:** Temporary or intermediate storage for ingested data.
     + **Storage:**
       - **Azure Data Lake Storage:** Acts as the central repository for structured and unstructured data.
   * **Key Actions:**
     + Data pipelines extract and load data into the lake.
     + Transformation tasks (e.g., cleansing, deduplication) are managed through ADF or external tools.
3. **Semantic Layer**
   * **Key Components:**
     + **Tabular Model (Cube):**
       - Built in **Azure Analysis Services (AAS)** or Power BI's premium dataset.
       - Provides a semantic layer with business-friendly entities, measures, and relationships.
       - Pre-calculates and organizes data for faster query performance.
     + **Data Marts:**
       - Specialized datasets tailored for business domains like Sales, Marketing, or Inventory.
       - Provide simplified access to domain-specific insights.
   * **Authentication:**
     + Secured by **Azure Active Directory (AAD)**, ensuring role-based access to the tabular model and data marts.
4. **Presentation Layer**
   * **Tool:** **Power BI**
     + Connects to the tabular models or data marts for data visualization and analysis.
     + Enables business users to build dashboards, reports, and conduct ad-hoc queries.

#### 

#### 

### **Implementation Steps**

#### **Creating Semantic Models and Dataflows in Power BI**

* **Reusable Datasets:**Publish datasets for key business areas such as:
  + **Sales Data:** Daily, weekly, and monthly sales trends.
  + **Customer Analytics:** Segmentation based on purchase behavior and demographics.
  + **Inventory Management:** Current stock levels, reorder points, and stock turnover rates.
* **Dataflows for Automation:**Set up Power BI Dataflows to automate recurring data preparation tasks, such as:
  + Merging daily sales data from multiple stores.
  + Calculating metrics like Average Transaction Value (ATV) and Conversion Rates.

#### **Empowering Business Users with Power BI**

* **Prebuilt Dashboards and Templates:**Provide starter dashboards for different user groups:
  + **Store Managers:** Monitor daily sales performance, top-selling products, and inventory alerts.
  + **Regional Managers:** Compare sales across stores and regions, track revenue targets, and identify underperforming stores.
  + **Marketing Team:** Analyze campaign effectiveness, customer lifetime value (CLV), and trends in customer preferences.
* **Self-Service Capabilities:**
  + Enable users to create ad-hoc reports by connecting to curated datasets.
  + Offer drag-and-drop functionality with prebuilt fields like "Total Sales," "Customer Segments," and "Profit Margin."

### **How This Architecture Benefitted the Retail company**

#### **Sales and Customer Insights**

1. **Data Sources:**
   * POS systems, CRM, and e-commerce platforms provide sales and customer data.
2. **Orchestration Layer:**
   * Azure Data Factory loads raw data into Azure Data Lake daily.
   * Transformations clean and organize data into curated datasets.
3. **Semantic Layer:**
   * Tabular models define key metrics like **Total Sales**, **Profit Margin**, and **Customer Retention Rate**.
   * Data marts for Sales, Customer Analytics, and Inventory simplify access.
4. **Presentation Layer:**
   * Power BI dashboards visualize:
     + Daily sales trends for store managers.
     + Customer segmentation and buying patterns for marketing teams.