# Retail Sales Analysis

# using

# Microsoft Fabric

# 1. Project Overview

# **Objective** To design and implement an end-to-end data pipeline leveraging the Medallion architecture in Microsoft Fabric for retail sales analytics. This includes data ingestion, transformation, enrichment, modeling, and dashboarding.

# **Tools & Technologies Used**

# Microsoft Fabric (Lakehouse, Notebooks, Pipelines, Power BI)

# PySpark

# Delta Lake

# Power BI

# 2. Workspace & Lakehouse Setup

• Created Microsoft Fabric workspace

A screenshot of a computer

AI-generated content may be incorrect.

• Created a Lakehouse and added raw data via HTTP/CSV to the Bronze folder

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

# 3. Ingest Data (Bronze Layer)

• Used Dataflow Gen2 or Copy Data activity to load raw sales.csv into Lakehouse

A screenshot of a computer

AI-generated content may be incorrect.

• Data saved in Bronze layer as Delta table

A screenshot of a computer

AI-generated content may be incorrect.

# 4. Data Cleaning (Silver Layer via Notebook)

• Renamed columns to remove spaces

A screenshot of a computer

AI-generated content may be incorrect.

• Converted Order\_Date to date format

A screenshot of a computer

AI-generated content may be incorrect.

• Added calculated columns: Profit\_Margin, Order\_Year, Order\_Month

Notebook: Transform\_Silver\_Layer.ipynb

# 5. Customer Enrichment (Gold Layer)

• Created customers table from unique Customer\_IDs

• Added Age\_Group, Gender, Region, Loyalty\_Score columns

• Joined customers with superstore\_silver to create superstore\_gold

A screenshot of a computer

AI-generated content may be incorrect.

Notebook: Join\_Customers\_Gold\_Layer.ipynb

# 6. Semantic Model

• Created model from superstore\_gold table

A screenshot of a computer

AI-generated content may be incorrect.

• Added DAX measures:

Total Sales = SUM(superstore\_gold[Sales])

Total Orders = COUNT(superstore\_gold[Order\_ID])

Profit Margin % = AVERAGE(superstore\_gold[Profit\_Margin])

A screenshot of a computer

AI-generated content may be incorrect.

# 7. Power BI Dashboard

• Created visuals:

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

# 8. Lineage and Monitoring

• Viewed end-to-end lineage

A screenshot of a computer

AI-generated content may be incorrect.

• Monitored Copy Data and pipeline refresh status

A screenshot of a computer

AI-generated content may be incorrect.

# Final Notes

• Follows Medallion architecture (Bronze > Silver > Gold)

• End-to-end project demonstrates data engineering and analytics workflow in Microsoft Fabric