■■ E-Commerce Data Pipeline – End-to-End Flow

■ Data Sources

■ DummyJSON API (Customers, Products, Carts)

■ Mockaroo API (Orders)

■

▼

■ Bronze Layer – Raw Data

Stores unprocessed data in Delta format

■■ Customers • Products • Carts • Orders

■

▼

■ Silver Layer – Cleaned Data

Data cleansing, standardization, and deduplication

■ Slv_Customers • Slv_Products • Slv_Orders • Slv_Carts

■

▼

■ Gold Layer – Dimensional Model

Dimensional schema creation for analytics

■ Dim_Customers • Dim_Products • Fact_Orders • Fact_Carts

■

▼

■ Gold+ Aggregations

Business summary tables for BI reporting

■ Gold_Agg_Sales_Summary • Gold_Agg_Customer_Spending

■

▼

■ AI / ML Integration (Part C)

■ Hugging Face NLP – Sentiment Analysis on Product Reviews

■ ML_Sales_Prediction – Predict future daily sales trends

■ ML_Customer_Spending – Predict high-value customers

■

▼

■ Visualization & Insights

Power BI & Fabric Dashboards

■ Positive/Negative Sentiment • Sales Forecasts • Customer Value Trends

© 2025 E-Commerce Analytics Pipeline by Muhammad Saad