**Customer 360 Data Integration**

**Objective:**

The goal of this project is to integrate data from multiple sources to create a unified Customer 360 view, providing insights into customer behavior, interactions, and preferences. The integration includes online transactions, in-store purchases, customer service interactions, and loyalty program data.

**Scope:**

* **Data Sources**: Online transactions, in-store purchases, customer service interactions, loyalty programs.
* **Technologies Used**: Azure Synapse Analytics, Azure Data Lake Storage (ADLS), Azure SQL Database, Power BI.
* **Architecture**: The solution involves using Azure Data Factory for data ingestion, Azure Synapse Analytics for data curation, Azure SQL Database for analytics, and Power BI for visualization.

**Architecture Diagram:**

A diagram of a company

AI-generated content may be incorrect.

**Role:**

**Resources:**

Necessary resources for this project have been create under resource group – customer\_360\_workpsace.

A screenshot of a computer

AI-generated content may be incorrect.

**Data Ingestion:**

As part of data ingestion process, necessary tables are created in SQL pool of synapse - cus\_db\_curated by leveraging copy activity in azure data factory. For this linked service for source data adls and synapse has been created using service principal authentication for safer connection.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**DDL Scripts:**

****

**Data Cleaning:**

We performed data cleaning activities including removing duplicate records using ROW\_NUMBER() for partitioning by unique identifiers like CustomerID, ProductID, OrderID, etc. We addressed missing or inconsistent values, especially in transactions, by ensuring data consistency and correctness. Additionally, we standardized date and time formats across tables and applied necessary transformations to ensure accurate data for analysis. All these steps were aimed at preparing the dataset for reliable reporting and analytics.

A screenshot of a computer

AI-generated content may be incorrect.

**Data Cleaning Scripts:**

****

**Dara Warehouse:**

Azure SQL server is leveraged for data warehousing,(generally in synapse dedicated SQL pool). Necessary final analytical tables are created in sql database - customer\_db\_golden. Business transformations are implemented leveraging SQL to produce necessary business daatsets. Copy activity in data factory is leveraged to copy business data from curated to gold layer.

A screenshot of a computer

AI-generated content may be incorrect.

**SQL Scripts:**



**Analytics Table:**

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.

A screenshot of a computer

AI-generated content may be incorrect.

**Report:**

Data warehouse is connected to Power Bi to pull business data and create needful interactive dashboards as below.

**A screenshot of a computer

AI-generated content may be incorrect.**

**GitHub Link: https://github.com/Arunkumar-Senthilkumar/customer-ETL**