Customer Management Data Engineering Project Documentation

1. Entity-Relationship Diagram (ERD)

- Purpose: The ERD explains the relationships between the entities in the Customer Management system and clarifies the attributes of each entity.
- Tool: ERD created using SQL Server.
- Entities and Relationships: (Explain the characteristics of each entity and explain the relationships between entities such as 1:M, M:M, M:1.)

2. Database Creation

- Objective: Create the Customer Management database on SQL Server.
- SQL Code: (Use DDL, DML, Data Types)
- Tools: SQL Server.

3. Data Population with Faker

- Purpose: Populate the database with synthetic data for testing.
- Library Used: Python's faker library.
- Method: (import pyodbc, import faker.)

4. Data Warehouse Design (Star Schema)

- Objective: Create a Data Warehouse schema using the Star Schema model.
- SQL Code: (Use DDL, DML, Data Types)
- Design Overview: (Create a Fact table and create Dimensions tables.)
- Tools: SQL Server.

5. ETL Process

• Objective: Extract data from the Customer Management database, perform transformations, and load it into the Data Warehouse.

Steps:

- Extract: Extract data from the CustomerManagement database.
- Transform: Perform simple transformations.
- Load: Load data into the CustomerManagementDWH.
- Tools: Python.

6. Data Analysis

- Purpose: Analyze the data to answer business-related questions.
- Analysis Queries: (SELECT, FROM, WHERE, JOIN, ON, GROUB BY, HAVING, ORDER BY)

Business Questions Answered: Questions about the project's business.

7. Dashboard

• Objective: Visualize key metrics dynamically for business stakeholders.

• Tool: (POWER BI)

• Visualizations: (POWER BI)

8. Conclusion

This project demonstrates a complete Data Engineering pipeline from database creation to data warehousing, ETL, and business analysis.

Project Implementers

- · Ali Emad Abbas Mohammed
- Nour Eldeen Mohammed Ali
- Abdallah Housni
- Abdelrahman Mostafa