Bootcamp Project 4

Project Title: Incremental Data Loading and Automated Notifications using Microsoft Fabric

❖ **Problem Statement:**

In modern data ecosystems, organizations need to efficiently ingest, transform, and load data from various sources into centralized platforms for analytics, while also ensuring timely monitoring and notification upon successful data refreshes. This project addresses the challenge of incrementally loading data from on-premises sources to Microsoft Fabric Lakehouse, processing it through a structured transformation pipeline, and triggering automated notifications upon successful execution.

❖ **Project Objective:**

To build an end-to-end data pipeline on Microsoft Fabric that:

1. Ingests structured data from on-premises environments into a Fabric Lakehouse using the On-Prem Gateway
2. Utilizes the AI Bank Dataset as the source
3. Implements Dataflow Gen 1 to join tables, remove duplicates, and clean data
4. Loads the cleansed data into a Fabric Warehouse
5. Applies Slowly Changing Dimension (SCD) Type 1 logic using Fabric Notebooks and writes the results into separate warehouse tables
6. Schedules and monitors the pipeline, sending an automated email notification (via Outlook or Gmail) upon successful pipeline completion

❖ **Tools & Technologies:**

- Microsoft Fabric
- On-Premises Data Gateway
- Fabric Lakehouse and Warehouse
- Fabric Dataflow Gen 1
- Fabric Notebook
- Email Notification Task (in-built)
- Azure Key Vault (optional for secure credential management)
- Draw.io / Visio for architecture diagram

❖ **Architecture Components (Use Draw.io):**

A detailed architecture diagram will illustrate:

- Data movement from on-prem to Fabric Lakehouse
- Data transformation flow via Dataflow and Notebooks
- Loading patterns into the warehouse
- Notification triggers

❖ **Deliverables:**

- Comprehensive documentation detailing:
  - Data ingestion setup and configurations
  - Dataflow transformation logic
  - Notebook-based SCD Type 1 implementation
  - Pipeline scheduling and notification setup
- Architecture diagram (.drawio or image)
- Upload the project to GitHub