***Employee Data Management System***

**Objective**:

Develop a simple Employee Data Management System that processes and stores employee records using AWS services (EC2, Lambda, Redshift, DynamoDB, Teradata). The system will handle data extraction, transformation, and loading (ETL) to enable quick retrieval, analysis, and archival of employee information.

**Business Case Overview**:

A company wants to manage employee data across multiple departments. Each department submits daily employee data in CSV format, containing new employee information. The company needs a system that can:

1. Extract the daily employee data.

2. Transform the data (validate, clean, and enrich it).

3. Load the data into an AWS Redshift warehouse for reporting and analytics.

4. Store the raw data in DynamoDB for quick lookups.

5. Archive the data in Teradata for long-term storage and backup.

**Source Requirements**:

Source Data (CSV Files):

The data is provided daily in CSV files containing the following fields:

Employee\_ID Name Department Join\_Date Salary Email

E001 Alice HR 2025-01-01 55000 [alice@example.com](mailto:alice@example.com)

E002 Bob IT 2025-01-02 60000 [bob@example.com](mailto:bob@example.com)

E003 Charlie Finance 2025-01-02 65000 [charlie@example.com](mailto:charlie@example.com)

E004 David IT 2025-01-03 70000 [david@example.com](mailto:david@example.com)

**Target Requirements**:

1. AWS S3:

* Store the raw CSV files temporarily in S3 for processing.
* S3 Path: /raw-data/employee/

2. AWS Redshift:

* + Store transformed and aggregated employee data for reporting and analytics.

Tables:

* + Employee\_Facts: Store employee data (ID, name, department, salary, etc.).
  + Department\_Summary: Aggregated data (total employees and average salary by department).

3. DynamoDB:

* + Store employee data for quick lookups by employee ID.
  + DynamoDB Table: Employee\_Lookup
  + Primary Key: Employee\_ID

4. Teradata:

* + Archive all raw employee data for long-term storage.
  + Teradata Table: Archived\_Employee\_Data
  + Fields: Employee\_ID, Name, Department, Join\_Date, Salary, Email