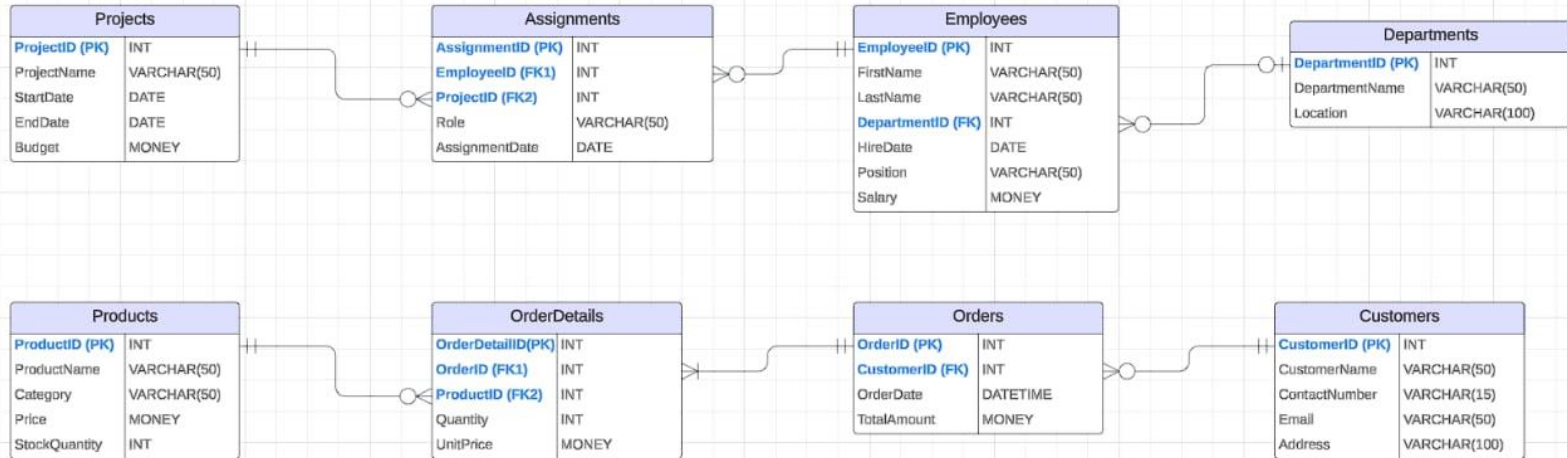


Data Warehousing Project Documentation



Script 1: Loading

Purpose:

- **Database and Schema Setup:** Initializes the DataWarehouse database and creates the loading_sources schema where temporary loading tables will reside.
- **Procedure Definition (loading_temp_tables):** Defines a stored procedure to facilitate the loading of data from source tables (CompanyDB and SourceDB) into temporary tables (loading_sources) within the DataWarehouse.
- **Dynamic SQL Execution:** Utilizes dynamic SQL (sp_executesql) to create temporary tables only if they do not already exist. This ensures that data loading operations do not overwrite existing tables inadvertently.

Execution:

- The script executes the loading_temp_tables procedure multiple times, each time specifying different source tables (Departments, Employees, Projects, Assignments) from CompanyDB and SourceDB. This populates the loading_sources schema with raw data from both databases.

Script 2: Surrogation

Purpose:

- **Schema Creation (surrogation):** Creates the surrogation schema where surrogate key tables will be stored.
- **Procedure Definition (surrogate_table):** Defines a stored procedure responsible for creating and populating surrogate key tables (employeeSurrogation, departmentSurrogation, projectSurrogation, assignmentSurrogation) within the surrogation schema.
- **Ensuring Data Integrity:** Checks for the existence of the surrogate key table before creation. Inserts data into these tables ensuring uniqueness based on the combination of dataSource and primary key values (EmployeeID, DepartmentID, ProjectID, AssignmentID).

Execution:

- Executes surrogate_table for each entity (Employees, Departments, Projects, Assignments) and from both CompanyDB and SourceDB. This establishes consistent surrogate keys across datasets, crucial for maintaining referential integrity in the data warehouse.
-

Script 3: Merging

Purpose:

- **View Creation (transformed_CompanyDB_*):** Creates views (transformed_CompanyDB_Employee, transformed_CompanyDB_Department, transformed_CompanyDB_Project, transformed_CompanyDB_Assignment) that merge data from CompanyDB into the DataWarehouse.
- **Integration of Surrogate Keys:** Utilizes surrogate keys (s_emp, s_dep, s_proj, s_assign) from the surrogation schema to establish relationships between entities (Employees, Departments, Projects, Assignments) across different schemas within the data warehouse.

Execution:

- These views consolidate and transform data from raw loading tables (loading_sources) into structured representations (transformed_CompanyDB_*) suitable for querying and reporting purposes. They ensure that data from CompanyDB is integrated and standardized within the data warehouse schema.
-

Script 4: Merging Sources

Purpose:

- **View Creation (transformed_SourceDB_*):** Creates views (transformed_SourceDB_Employee, transformed_SourceDB_Department,

transformed_SourceDB_Project, transformed_SourceDB_Assignment) that merge data from SourceDB into the DataWarehouse.

- **Utilization of Surrogate Keys:** Similar to Script 3, utilizes surrogate keys (s_emp, s_dep, s_proj, s_assign) to unify and integrate data from SourceDB across different entities (Employees, Departments, Projects, Assignments) within the data warehouse.

Execution:

- These views ensure that data from SourceDB is harmonized with CompanyDB data within the data warehouse schema. They provide a consistent view of information across multiple source databases, facilitating comprehensive data analysis and reporting.
-

Script 5: Base

Purpose:

- **Table Creation (base_*):** Creates base tables (base_departments, base_employees, base_projects, base_assignments) within the DataWarehouse schema.
- **Storage of Integrated Data:** Stores integrated and standardized data from both CompanyDB and SourceDB, utilizing the surrogate keys (s_emp, s_dep, s_proj, s_assign) for relational integrity and consistency.

Execution:

- Populates these tables by inserting data from the merged views (transformed_CompanyDB_*, transformed_SourceDB_*). These tables serve as the foundational data structures for querying and reporting, providing a single source of truth for analytics and decision-making processes.
-

Script 6: EmployeeProjectAssignmentsView

Purpose:

- **View Creation (EmployeeProjectAssignments):** Creates a view that consolidates employee assignments across projects within the DataWarehouse.
- **Data Aggregation:** Retrieves data from the base tables (base_employees, base_departments, base_projects, base_assignments) and aggregates it to provide insights into employee roles, departments, project assignments, and durations.

Execution:

- This view facilitates easy access to information about employee assignments, allowing stakeholders to analyze workforce allocation, project involvement, and resource utilization within the organization.

Script 7: Queries

Purpose:

- **Data Analysis Queries:** Contains a series of SQL queries designed to extract specific insights and analytics from the DataWarehouse.
- **Information Retrieval:** Each query addresses different aspects such as employee details, department budget summaries, project assignments, customer order analysis, product details extraction, and high-salary employees in specific departments.

Execution:

- These queries leverage the structured data stored in the base_* tables to generate meaningful reports and insights for decision-making purposes. They demonstrate the capability of the data warehouse to support complex analytical queries and provide actionable intelligence to stakeholders.