Database Setup, Data Integration, and Relationship Establishment

Description

This step focused on creating the **Northwind database** in SQL Server, populating it with data from 11 CSV files, and establishing relationships between tables using primary and foreign key constraints. The goal of this step was to build a fully functional, normalized database ready for querying, analysis, and integration with other systems.

Implementation Details

1. Database Creation

- **Database Design:**
  - Created the **Northwind database** in SQL Server Management Studio (SSMS).
  - Designed the schema to include 11 tables, ensuring normalization to eliminate redundancy and improve data integrity.
  - Tables included:
    - **Transactional Tables:** Orders , Order Details.
    - **Dimensional Tables:** Employees, Products, Customers, Shippers, Suppliers, Categories, Territories, etc.

- **Schema Definition:**
  - Defined appropriate data types for each column .
  - Added necessary constraints (e.g., `NOT NULL`, `UNIQUE`) to ensure data quality.

```sql
/*******************Creating Table Territories************************/

CREATE TABLE [dbo].[Territories](
    [TerritotyID] [nvarchar](20) NOT NULL,
    [TerritoryDescription] [nchar](50) NULL,
    [RegionID] [int] NULL,
    CONSTRAINT [PK_Territories] PRIMARY KEY ([TerritoryID])
)


/*******************Creating Table Suppliers************************/

CREATE TABLE [dbo].[Suppliers](
    [SupplierID] [int] NOT NULL,
    [CompanyName] [nvarchar](40) NULL,
    [ContactName] [nvarchar](30) NULL,
    [ContactTitle] [nvarchar](30) NULL,
    [Address] [nvarchar](60) NULL,
    [City] [nvarchar](15) NULL,
    [Region] [nvarchar](15) ,
    [PostalCode] [nvarchar](10) NULL,
    [Country] [nvarchar](15) NULL,
    [Phone] [nvarchar](24) NULL,
    [Fax] [nvarchar](24) NULL,
    [HomePage] [ntext] NULL,
    CONSTRAINT [PK_Suppliers] PRIMARY KEY ( [SupplierID] )
)

/*******************Creating Table Shippers************************/
```

```sql
/*******************Creating Table Shippers************************/
CREATE TABLE [dbo].[Shippers](
    [ShipperID] [int] NOT NULL,
    [CompanyName] [varchar](40) NULL,
    [Phone] [varchar](20) NULL,
    CONSTRAINT [PK_Shippers] PRIMARY KEY ( [ShipperID])
)


/*******************Creating Table Regions************************/
CREATE TABLE [dbo].[Regions](
    [RegionID] [int] NOT NULL,
    [regiondescription] [nchar](50) NULL
)


/*******************Creating Table Products************************/

CREATE TABLE [dbo].[Products](
    [ProductID] [int] NOT NULL,
    [ProductName] [nvarchar](40) NULL,
    [SupplierID] [int] NULL,
    [CategoryID] [int] NULL,
    [QuantityPerUnit] [nvarchar](20) NULL,
    [UnitPrice] [money] NULL,
    [UnitsInStock] [smallint] NULL,
    [UnitsOnOrder] [smallint] NULL,
    [ReorderLevel] [smallint] NULL,
    [Discontinued] [bit] NULL
)
```

```sql
/*******************Creating Table Orders*************************/

CREATE TABLE [dbo].[Orders](
    [OrderID] [int] NOT NULL,
    [CustomerID] [nchar](5) NULL,
    [EmployeeID] [int] NULL,
    [ShipVia] [int] NULL,
    [OrderDate] [datetime] NULL,
    [ShipDate] [datetime] NULL,
    [RequiredDate] [datetime] NULL,
    [Freight] [money] NULL,
    [ShipName] [nvarchar](40),
    [ShipAddress] [nvarchar](60),
    [ShipRegion] [varchar](15) NULL,
    [ShipCountry] [varchar](15) NULL,
    [ShipCity] [varchar](15) NULL,
    [ShipPostalCode] [varchar](10) NULL,
    [ModifiedDate] [datetime]
)

/*******************Creating Table OrderDetails************************/
CREATE TABLE [dbo].[OrderDetails](
    [OrderID] [int] NOT NULL,
    [ProductID] [int] NOT NULL,
    [UnitPrice] [money] NULL,
    [Quantity] [smallint] NULL,
    [Discount] [real] NULL
)
```
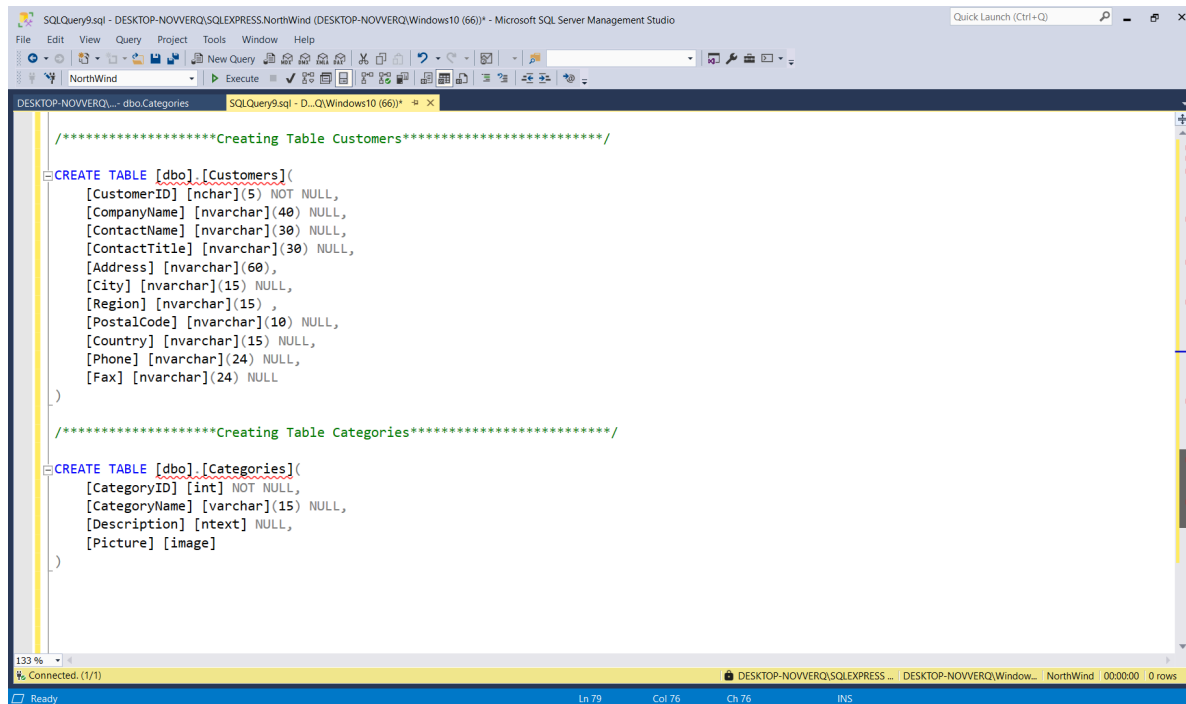
```sql
/*******************Creating Table Employees_Territories************************/

CREATE TABLE [dbo].[Employees_Territories](
    [EmployeeID] [int] NULL,
    [TerritoryID] [nvarchar](20) NULL
)

/*******************Creating Table Employees************************/

CREATE TABLE [dbo].[Employees](
    [EmployeeID] [int] NOT NULL,
    [LastName] [nvarchar](20) NULL,
    [FirstName] [nvarchar](10) NULL,
    [Title] [nvarchar](30) NULL,
    [TitleOfCourtesy] [nvarchar](25) NULL,
    [BirthDate] [datetime] NULL,
    [HireDate] [datetime] NULL,
    [Address] [nvarchar](60) NULL,
    [City] [nvarchar](15) NULL,
    [Region] [nvarchar](15) NULL,
    [PostalCode] [nvarchar](10) NULL,
    [Country] [nvarchar](15) NULL,
    [HomePhone] [nvarchar](24) NULL,
    [Extension] [nvarchar](4) NULL,
    [Notes] [ntext] NULL,
    [ReportsTo] [int] NULL,
    [PhotoPath] [nvarchar](255) NULL
)
```

```
/*******************Creating Table Customers*************************/

CREATE TABLE [dbo].[Customers](
    [CustomerID] [nchar](5) NOT NULL,
    [CompanyName] [nvarchar](40) NULL,
    [ContactName] [nvarchar](30) NULL,
    [ContactTitle] [nvarchar](30) NULL,
    [Address] [nvarchar](60),
    [City] [nvarchar](15) NULL,
    [Region] [nvarchar](15) ,
    [PostalCode] [nvarchar](10) NULL,
    [Country] [nvarchar](15) NULL,
    [Phone] [nvarchar](24) NULL,
    [Fax] [nvarchar](24) NULL
)

/*******************Creating Table Categories*************************/

CREATE TABLE [dbo].[Categories](
    [CategoryID] [int] NOT NULL,
    [CategoryName] [varchar](15) NULL,
    [Description] [ntext] NULL,
    [Picture] [image]
)
```
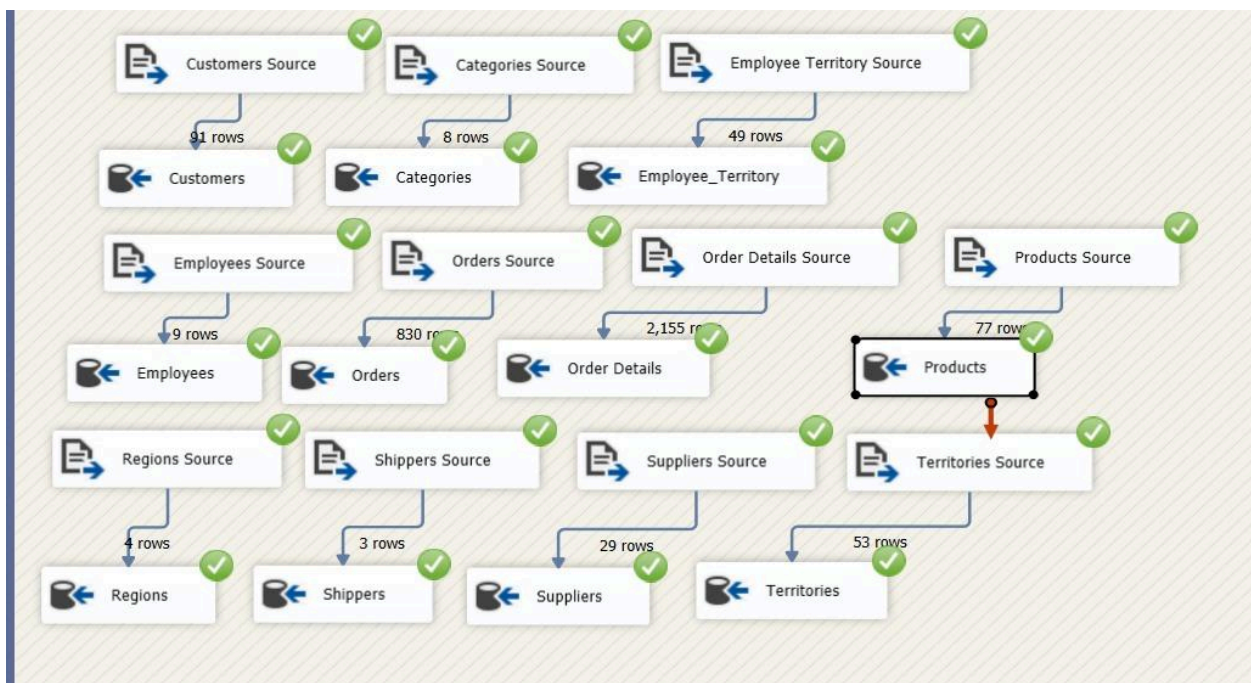
2. Data Population Using SSIS
- **ETL Process Overview:**
  - Developed an **SSIS (SQL Server Integration Services)** package to handle the ETL (Extract, Transform, Load) process.
  - The package was designed to extract data from 11 CSV files, transform it as needed, and load it into the corresponding tables in the Northwind database.
- **Data Extraction:**
  - Used SSIS to read data from CSV files, which included:
    - Orders (831 Records)
    - Order Details (2156 Records)
    - Employees (9 Records)
    - Products (77 Records)
    - Customers (91 Records)
    - Shippers (6 Records)
    - Suppliers (29 Records)
    - Categories (8 Records)
    - Territories (53 Records)
    - Regions (4 Records)
    - Employees Territoires (49 Records)

- **Data Transformation:**
    - Performed data cleaning and transformation tasks, such as:
        - Formatting dates , numeric and string fields to match the database schema.
        - Validating data to ensure consistency (e.g., checking for invalid customer IDs in the Orders table).
- **Data Loading:**
    - Loaded the transformed data into the respective tables in the Northwind database.
    - Handled large datasets efficiently:
        - Orders table: 800+ records.
        - Order Details table:2000+ records.
    - Ensured data integrity during the loading process .

3. Establishing Relationships
- **Primary Keys:**
    - Defined **primary keys** for each table to ensure uniqueness and enable efficient querying.
    - Examples:

- ■ `Orders.OrderID`
- ■ `Customers.CustomerID`
- ■ `Products.ProductID`
- ● **Foreign Key Constraints:**
  - ○ Created **foreign key constraints** to enforce relationships between tables and maintain referential integrity.
  - ○ Examples:
    - ■ `Orders.CustomerID` references `Customers.CustomerID`.
    - ■ `OrderDetails.OrderID` references `Orders.OrderID`.
    - ■ `OrderDetails.ProductID` references `Products.ProductID`.

- ● **Data Integrity Checks:**
  - ○ Verified that all relationships were correctly established by running test queries and checking for orphaned records.

```
-- ###################  EmployeesTerritories>>PK_Constraint  ################

ALTER TABLE EmployeesTerritories
 add constraint d_PK PRIMARY KEY (EmployeeID,TerritoryId)

--#####################  OrderDetails>> PK & Not Null Constraint  #########################

ALTER TABLE OrderDetails
 alter column orderid INT  Not null

ALTER TABLE OrderDetails
 alter column PRODUCTid INT  Not null

ALTER TABLE OrderDetails
 add constraint dou_PK PRIMARY KEY (OrderId,ProductId)

--######################  Region>>> Pk Consstraint  ##############################

ALTER TABLE Regions
 alter column RegionId INT  Not null

ALTER TABLE Regions
 add constraint reg_PK PRIMARY KEY (RegionId)

--####################  Shippers>>>> PK Constraint  #######################

ALTER TABLE Shippers
 alter column ShipperId INT  Not null
```

```sql
--####################  Suppliers>> PK Constraint  ####################

ALTER TABLE Suppliers
 alter column SupplierId INT  Not null

ALTER TABLE Suppliers
 add constraint sup_PK PRIMARY KEY (SupplierId)


 --################################  Territories>>> PK & FK Constraint  ########################

ALTER TABLE Territories
 alter column TerritoryId INT  Not null

ALTER TABLE Territories
 add constraint Ter_PK PRIMARY KEY (TerritoryId)


ALTER TABLE Territories
 add constraint REG_FK Foreign KEY (RegionId) REFERENCES Regions (RegionId)

 --###########################  Products>>> PK & FK Constraints  ########################

ALTER TABLE Products
 alter column ProductId INT  Not null

ALTER TABLE Products
 add constraint P_PK PRIMARY KEY (ProductId)
```

```sql
--############################  Orders>>>> PK & FK Constraint  ###############################

ALTER TABLE Orders
 alter column OrderId INT  Not null



ALTER TABLE Orders
 add constraint Ord_PK PRIMARY KEY (OrderId)

ALTER TABLE Orders
 add constraint Cust_FK Foreign KEY (CustomerId) REFERENCES Customers (CustomerId)

ALTER TABLE Orders
 add constraint Emp_FK Foreign KEY (EmployeeId) REFERENCES Employees (EmployeeId)

ALTER TABLE Orders
 add constraint Ship_FK Foreign KEY (ShipVia) REFERENCES Shippers (ShipperId)


 --############################  Category>>> PK Constraint ############################

ALTER TABLE Categories
 alter column CategoryId INT  Not null



ALTER TABLE Categories
 add constraint Cat_PK PRIMARY KEY (CategoryId)

 --##################################  Customers>>> PK Constraint  #################################
```

```sql
--##################################  Customers>>> PK Constraint  #################################

ALTER TABLE Customers
 alter column Customers nvarchar(50)  Not null

ALTER TABLE Customers
 add constraint Cust_PK PRIMARY KEY (Customers)

 --##################################  Employees>>> PK & FK  Constraint  ##################################

ALTER TABLE Employees
 alter column EmployeeId INT  Not null


ALTER TABLE Employees
 add constraint Emp_PK PRIMARY KEY (EmployeeId)

ALTER TABLE Employees
 add constraint EmpSupV_FK Foreign KEY (ReportSto) REFERENCES Employees (EmployeeId)
```