# Merge Transformation in SSIS

Dataset refer 33\_conditional split

## 1. Introduction

Merge Transformation in SSIS is used to combine two sorted datasets into a single dataset. It works similarly to the UNION operation in SQL but requires both inputs to be sorted.

## 2. When to Use Merge Transformation

Use Merge Transformation when:  
- You have two sorted datasets (e.g., CSV + SQL table) that need to be combined into one.  
- You need to perform incremental loads where new data from multiple sources is merged into a final table.  
- You want a single output stream from multiple sorted inputs.

## 3. Difference Between Merge and Union All

|  |  |  |
| --- | --- | --- |
| Feature | Merge Transformation | Union All Transformation |
| Sorting Requirement | Requires sorted inputs | Does not require sorting |
| Data Combination | Merges sorted datasets into one | Appends datasets directly |
| Performance | May be slower due to sorting | Faster, no sorting needed |
| Usage Scenario | When order matters in the merged dataset | When order doesn’t matter |

## 4. Real-World Example

Scenario:  
- Source 1 → CSV file (TestData\_1.csv) with 30 employee records  
- Source 2 → SQL Server table (males\_data) with 12 employee records  
- Goal → Merge both datasets into one table in SQL Server called MergeTransformationData

## 5. Step-by-Step Implementation in Visual Studio (SSIS)

### Step 1: Create a New SSIS Package

Open SQL Server Data Tools (SSDT) or Visual Studio.  
Drag Data Flow Task into the Control Flow area.  
Rename it 'Merge CSV and SQL Data'.

### Step 2: Configure CSV Source

Inside Data Flow, drag Flat File Source.  
Create a New Connection to C:\files\TestData\_1.csv.  
Verify columns in Preview and click OK.

### Step 3: Configure SQL Server Source

Drag OLE DB Source into Data Flow.  
Create a New OLE DB Connection to server 'developer' and database 'testing'.  
Choose table 'males\_data' and preview data.

### Step 4: Apply Sorting to Both Sources

Drag Sort Transformation after both sources.  
Sort by 'ID' in ascending order.

### Step 5: Configure Merge Transformation

Drag Merge Transformation into Data Flow.  
Connect both sorted outputs to Merge inputs.  
Verify sorted metadata.

### Step 6: Load Merged Data to SQL Server

Drag OLE DB Destination after Merge.  
Create a new table 'MergeTransformationData' and map columns.

### Step 7: Execute Package

Run the package and verify that 30 CSV + 12 SQL records = 42 total records are inserted.

## 6. Output Verification

In SQL Server Management Studio (SSMS):  
SELECT \* FROM MergeTransformationData;  
You should see 42 rows merged from both sources.

## 7. Error You May Face Without Sorting

If sorting is skipped, Merge Transformation shows:  
'The input Merge Input 1 must be sorted. The IsSorted property must be set to True.'  
Solution: Use Sort Transformation or ORDER BY in source query.

## 8. Key Points to Remember

- Requires Sorted Inputs (both datasets)  
- Works with two datasets only  
- Similar to SQL UNION but in SSIS environment  
- Can be used with different source types (CSV, Excel, SQL, Flat File)  
- Merge vs Merge Join: Merge → Combines datasets into one; Merge Join → Joins datasets based on a key

## 9. Real-World Use Cases

- Combining daily sales data from CSV and SQL for reporting  
- Merging incremental data from old and new systems  
- Combining data feeds from multiple vendors

## 10. Diagram of Merge Transformation Flow

CSV File (Sorted) ----\  
 \  
 --> Merge Transformation --> SQL Destination  
 /  
SQL Table (Sorted) ---/