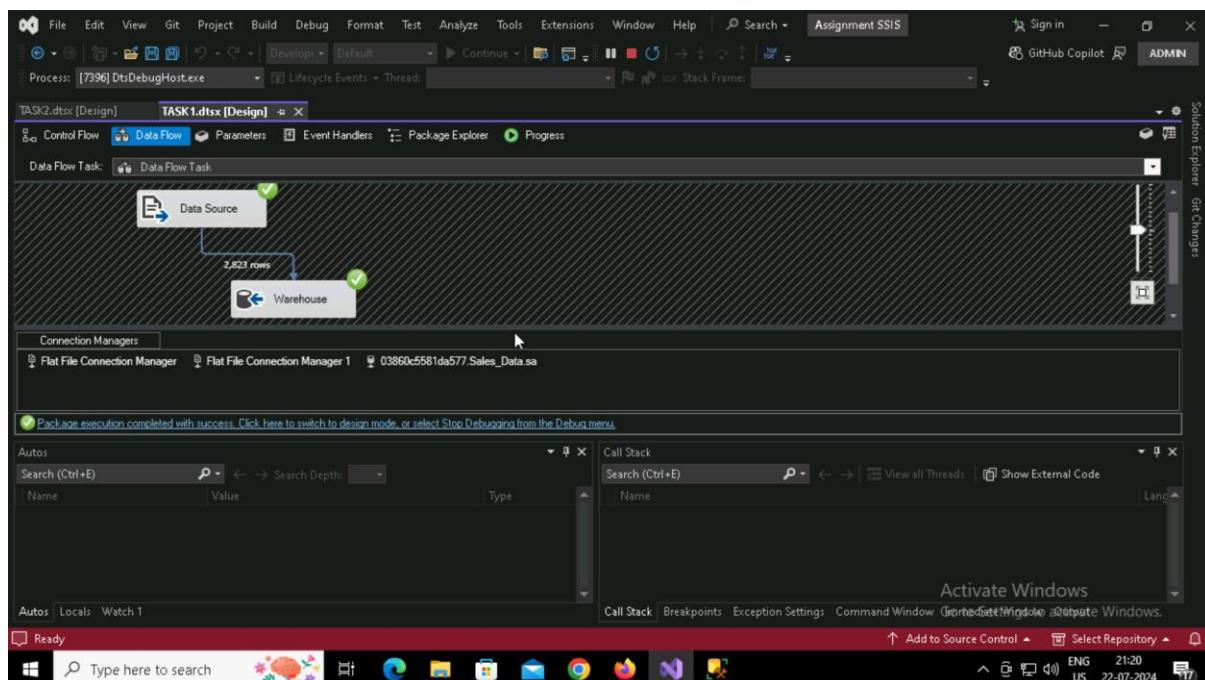
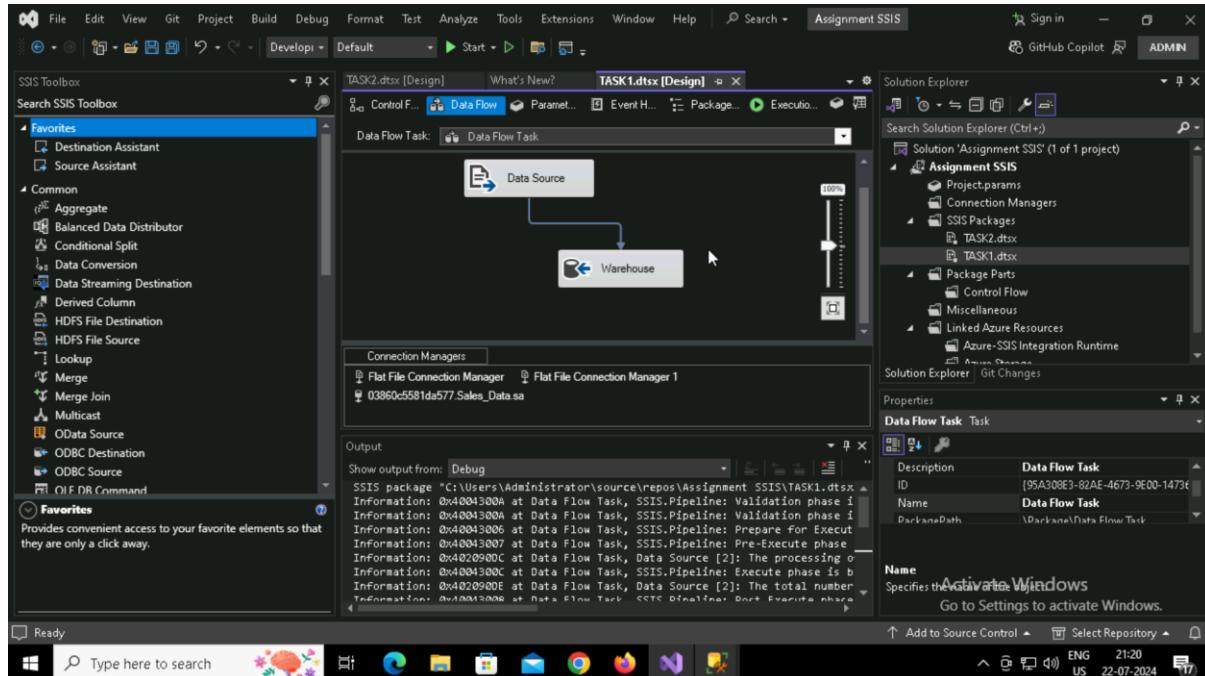


Task 1: Integration with ETL Data Warehouse (DWH)

Scenario: Your company has a data warehouse designed to consolidate data from various sources for analytical purposes. You need to create an SSIS package that extracts data from a transactional database and loads it into the data warehouse.



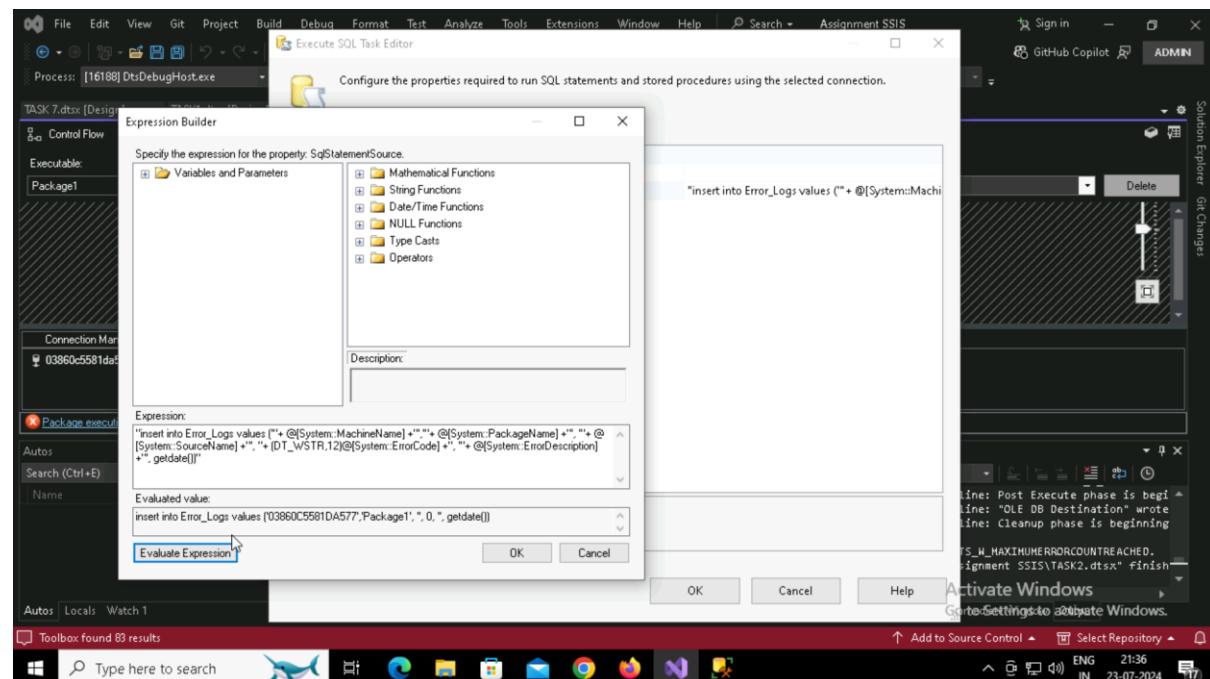
Task 2: Data Warehouse Migrations

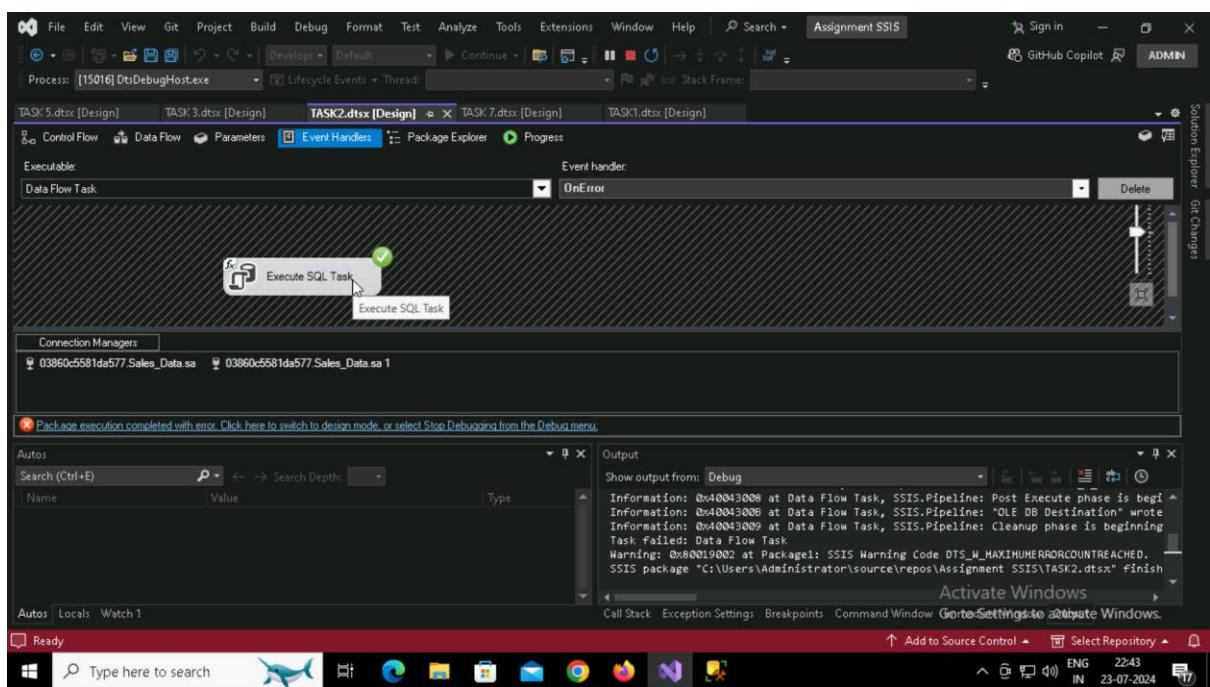
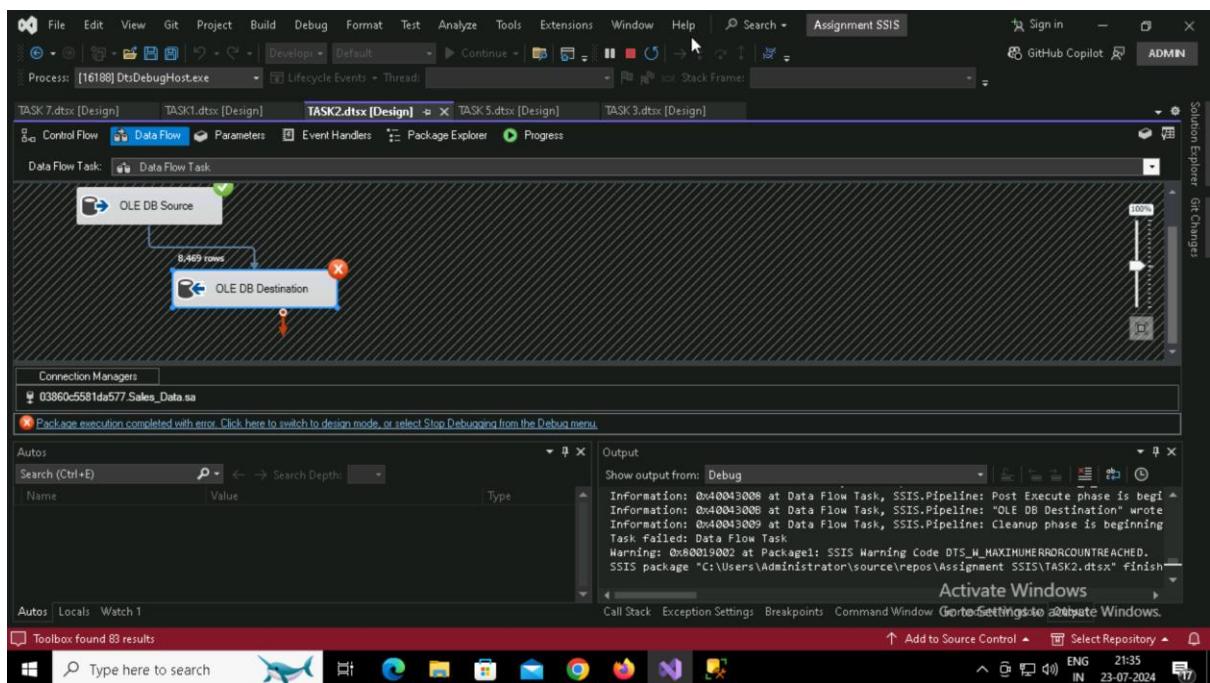
Scenario: Your organization is migrating its data warehouse from one server to another. You need to create an SSIS package that facilitates this migration.

Requirements:

1. Create Connection Managers for both the source and destination data warehouses.
2. Transfer Data from the source data warehouse to the destination using the Data Flow Task.
3. Ensure Data Integrity:
 - a) Include checks and balances to ensure data is correctly migrated.
 - b) Log the success or failure of the migration process.

```
IF NOT EXISTS(SELECT 1 FROM sysobjects with (nolock) WHERE ID = OBJECT_ID(N'Error_Logs') AND type = (N'U'))
drop table Error_Logs
go
CREATE TABLE Error_Logs(ID INT IDENTITY, MachineName varchar(200), PackageName varchar(200), TaskName varchar(200), ErrorCode int,
ErrorDescription varchar(max), Dated datetime)
go
IF EXISTS(SELECT 1 FROM sysobjects with (nolock) WHERE ID = OBJECT_ID(N'test') AND type = (N'U'))
drop table test
CREATE TABLE [dbo].[test](
[id] [varchar](50) NULL,
[first_name] [varchar](50) NULL,
[last_name] [varchar](50) NULL,
[Gender] [varchar](50) NULL,
[City] [varchar](50) NULL,
[Country] [varchar](50) NULL
) ON [PRIMARY]
```





The screenshot shows the Microsoft SQL Server Management Studio interface. In the Object Explorer, the database 'Sales_Data' is selected. In the center pane, a query window titled 'a\$SIGNMENT_ssis.sql...Sales_Data (sa (60))' displays the following SQL code:

```

select * from ERROR_LOGS

DROP TABLE Sales_New_tsk2

select * from Sales_pivot_table

```

The Results grid shows a table of errors from a Data Flow Task named 'Package1'. The columns are: ID, MachineName, PackageName, TaskName, ErrorCode, ErrorDescription, and Dated. The data is as follows:

ID	MachineName	PackageName	TaskName	ErrorCode	ErrorDescription	Dated
1	03860C5581DA577	Package1	Data Flow Task	-1071636471	SSIS Error Code DTS_E_OLEDBERROR. An OLE DB error has occurred. Column error.	2024-07-23 22:39:53.673
2	03860C5581DA577	Package1	Data Flow Task	-1071607780	There was an error with OLE DB Destination.Inputs[OLE DB Dest...]	2024-07-23 22:39:53.703
3	03860C5581DA577	Package1	Data Flow Task	-1071607767	SSIS Error Code DTS_E_INDUCEDTRANSFORMFAILUREONERRO...	2024-07-23 22:39:53.737
4	03860C5581DA577	Package1	Data Flow Task	-1073450974	SSIS Error Code DTS_E_PROCESSINPUTFAILED. The Process...	2024-07-23 22:39:53.770
5	03860C5581DA577	Package1	Data Flow Task	-1071636471	SSIS Error Code DTS_E_OLEDBERROR. An OLE DB error has o...	2024-07-23 22:40:33.073
6	03860C5581DA577	Package1	Data Flow Task	-1071607780	There was an error with OLE DB Destination.Inputs[OLE DB Dest...]	2024-07-23 22:40:33.103
7	03860C5581DA577	Package1	Data Flow Task	-1071607767	SSIS Error Code DTS_E_INDUCEDTRANSFORMFAILUREONERRO...	2024-07-23 22:40:33.130
8	03860C5581DA577	Package1	Data Flow Task	-1071636471	SSIS Error Code DTS_E_OLEDBERROR. An OLE DB error has o...	2024-07-23 22:42:47.643
9	03860C5581DA577	Package1	Data Flow Task	-1071607780	There was an error with OLE DB Destination.Inputs[OLE DB Dest...]	2024-07-23 22:42:47.673
10	03860C5581DA577	Package1	Data Flow Task	-1071607767	SSIS Error Code DTS_E_INDUCEDTRANSFORMFAILUREONERRO...	2024-07-23 22:42:47.690
11	03860C5581DA577	Package1	Data Flow Task	-1073450974	SSIS Error Code DTS_E_PROCESSINPUTFAILED. The Process...	2024-07-23 22:42:47.710
12	03860C5581DA577	Package1	Data Flow Task	-1071636471	SSIS Error Code DTS_E_OLEDBERROR. An OLE DB error has o...	2024-07-23 22:43:15.547

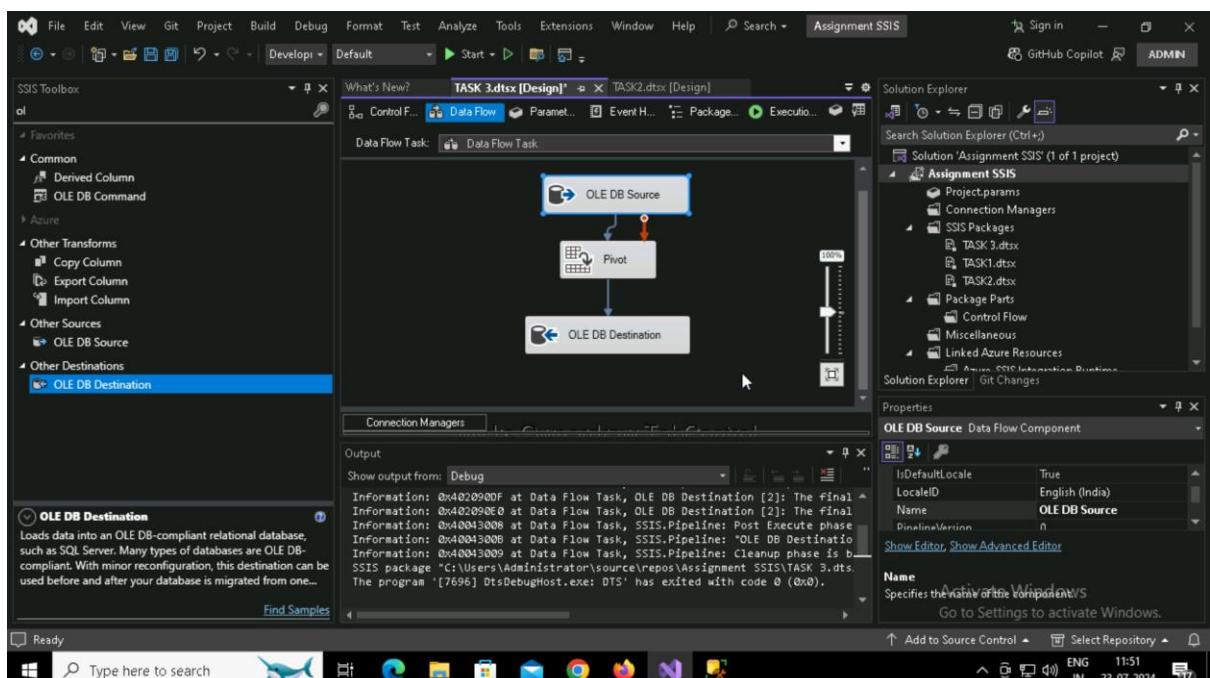
The status bar at the bottom right shows 'Query executed successfully.' and the date '2024-07-23 22:43:15.547'.

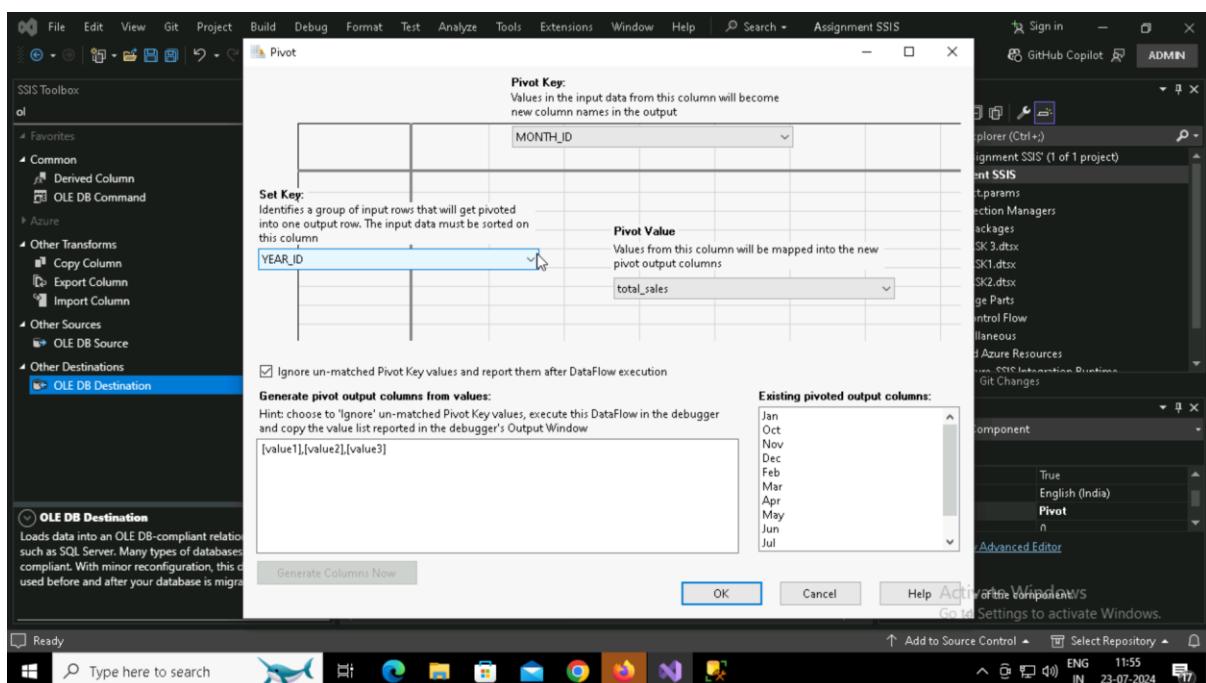
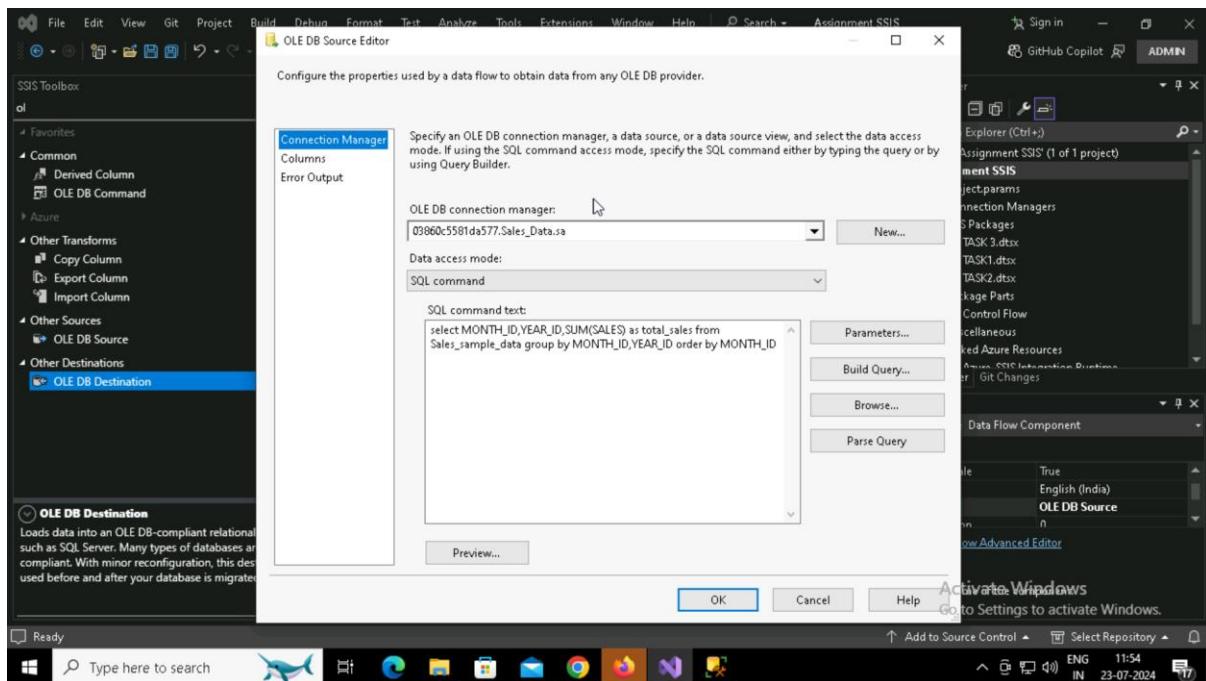
Task 3: Implementing a Pivot Transformation

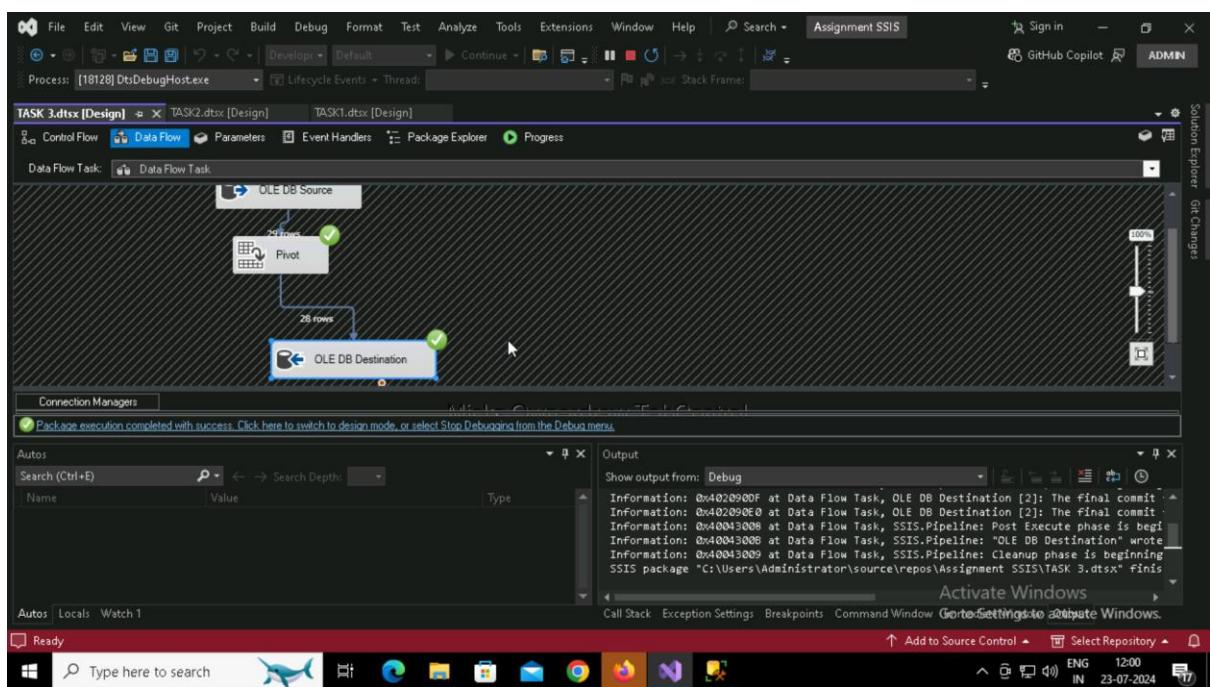
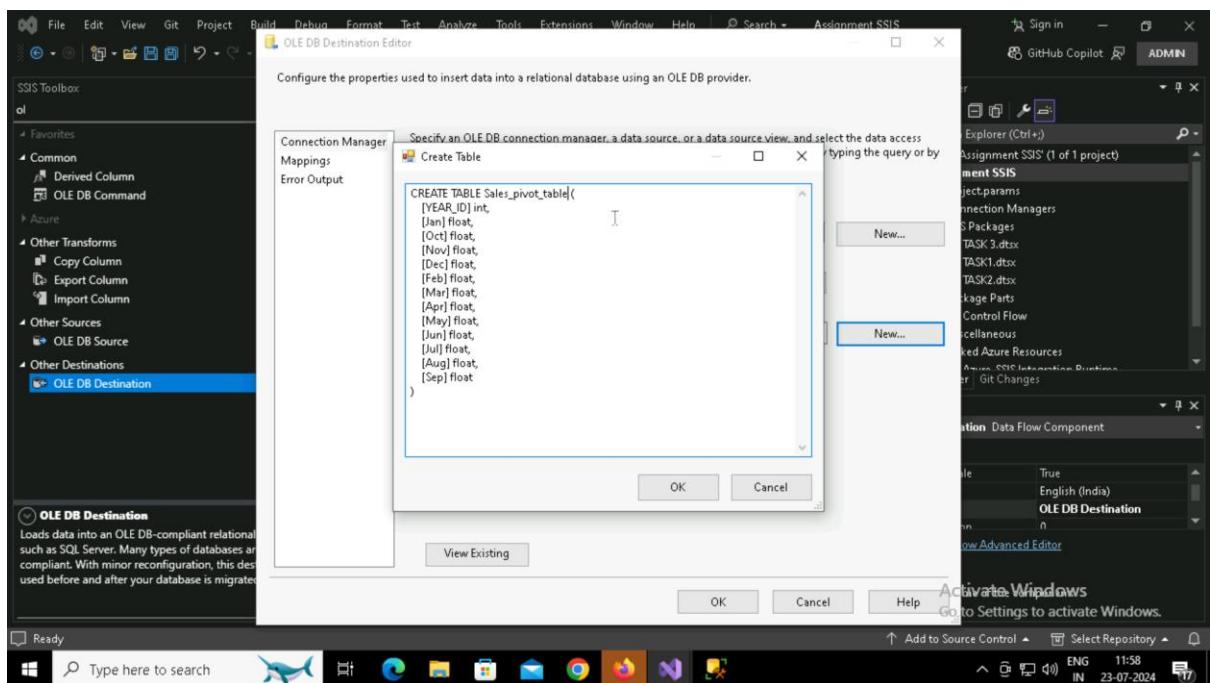
Scenario: You have data in a normalized format and need to pivot it for reporting purposes.

Requirements:

1. Extract Data from the source table using an OLE DB Source.
2. Apply a Pivot Transformation to transform the normalized data into a pivoted format.
3. Load the Pivoted Data into a destination table.







Assignment_CaseStudyEmployeeDB (sa (59))*

```
select * from Sales_pivot_table
```

100 %

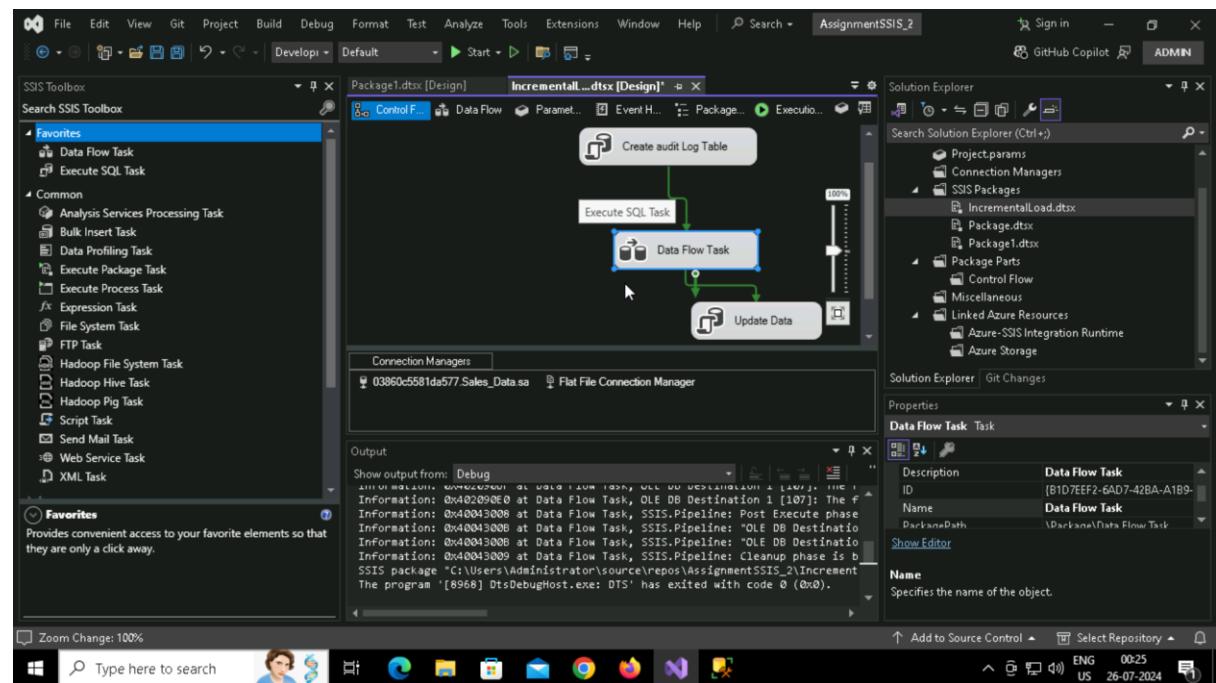
	YEAR_ID	Jan	Oct	Nov	Dec	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
1	2005	1018630.2588501	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	2003	389260.799194336	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	2004	949732.256286621	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	2003	NULL	NULL	NULL	NULL	422508.569824219	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	2005	NULL	NULL	NULL	NULL	1074558.53613281	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	2004	NULL	NULL	NULL	NULL	934258.590087891	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	2005	NULL	NULL	NULL	NULL	NULL	1122788.27929688	NULL	NULL	NULL	NULL	NULL	NULL
8	2003	NULL	NULL	NULL	NULL	NULL	523514.700256348	NULL	NULL	NULL	NULL	NULL	NULL
9	2004	NULL	NULL	NULL	NULL	NULL	617201.189575195	NULL	NULL	NULL	NULL	NULL	NULL
10	2005	NULL	NULL	NULL	NULL	NULL	NULL	784...	NULL	NULL	NULL	NULL	NULL
11	2003	NULL	NULL	NULL	NULL	NULL	NULL	604...	NULL	NULL	NULL	NULL	NULL
12	2004	NULL	NULL	NULL	NULL	NULL	NULL	618...	NULL	NULL	NULL	NULL	NULL
13	2005	NULL	NULL	NULL	NULL	NULL	NULL	NULL	137...	NULL	NULL	NULL	NULL
14	2003	NULL	NULL	NULL	NULL	NULL	NULL	NULL	578...	NULL	NULL	NULL	NULL
15	2004	NULL	NULL	NULL	NULL	NULL	NULL	NULL	820...	860...	NULL	NULL	NULL

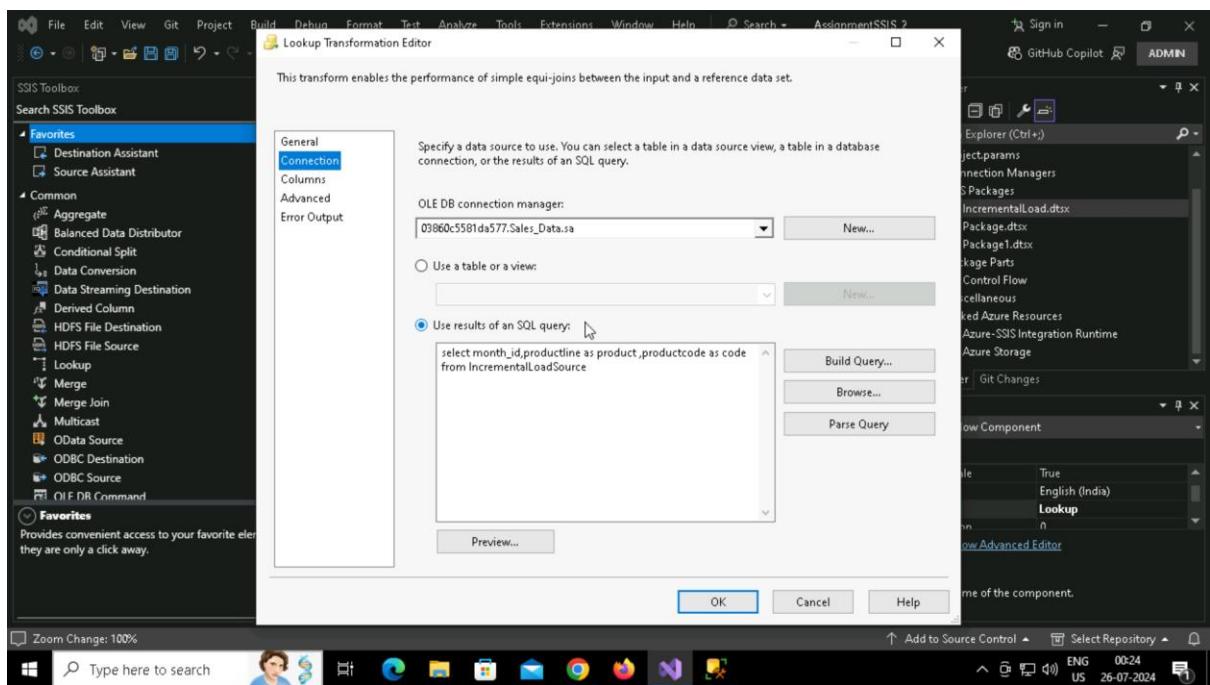
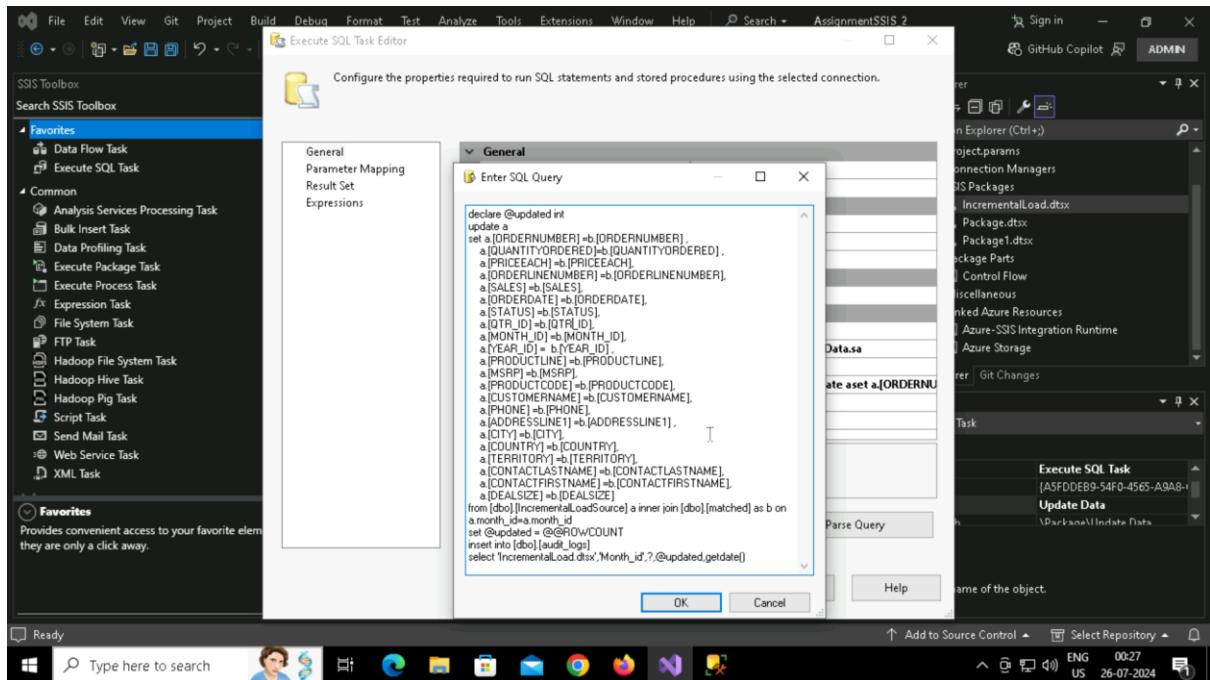
Task 4: Incremental Load

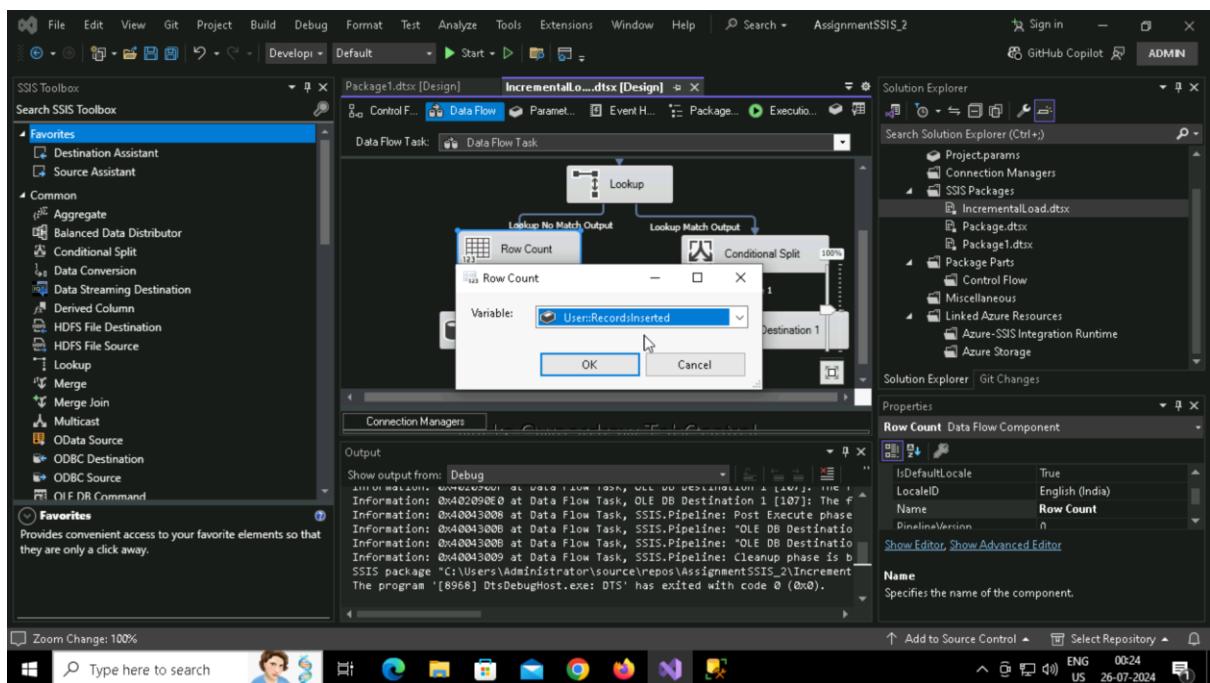
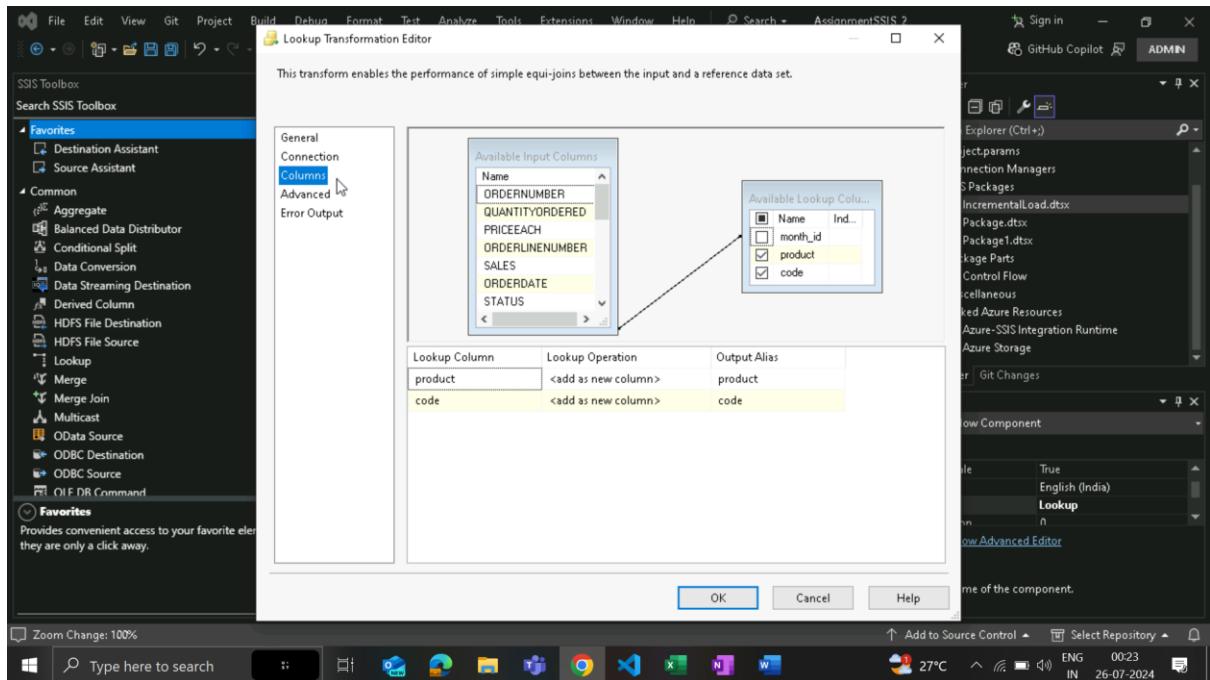
Scenario: To optimize ETL processes, you need to implement an incremental load to update only the changed data in the data warehouse.

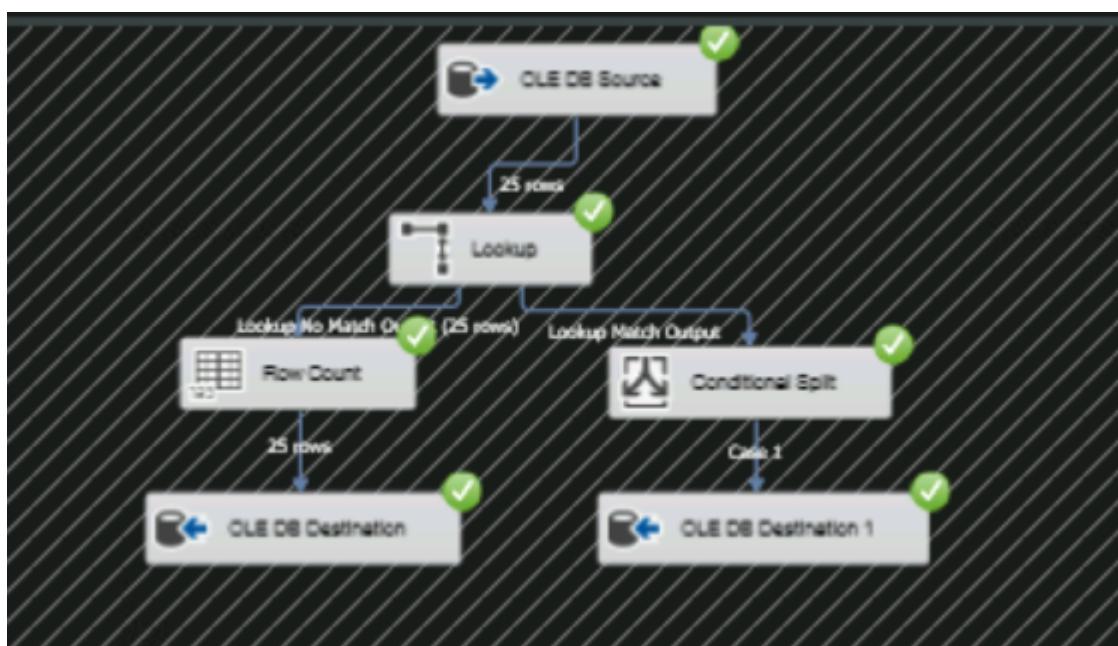
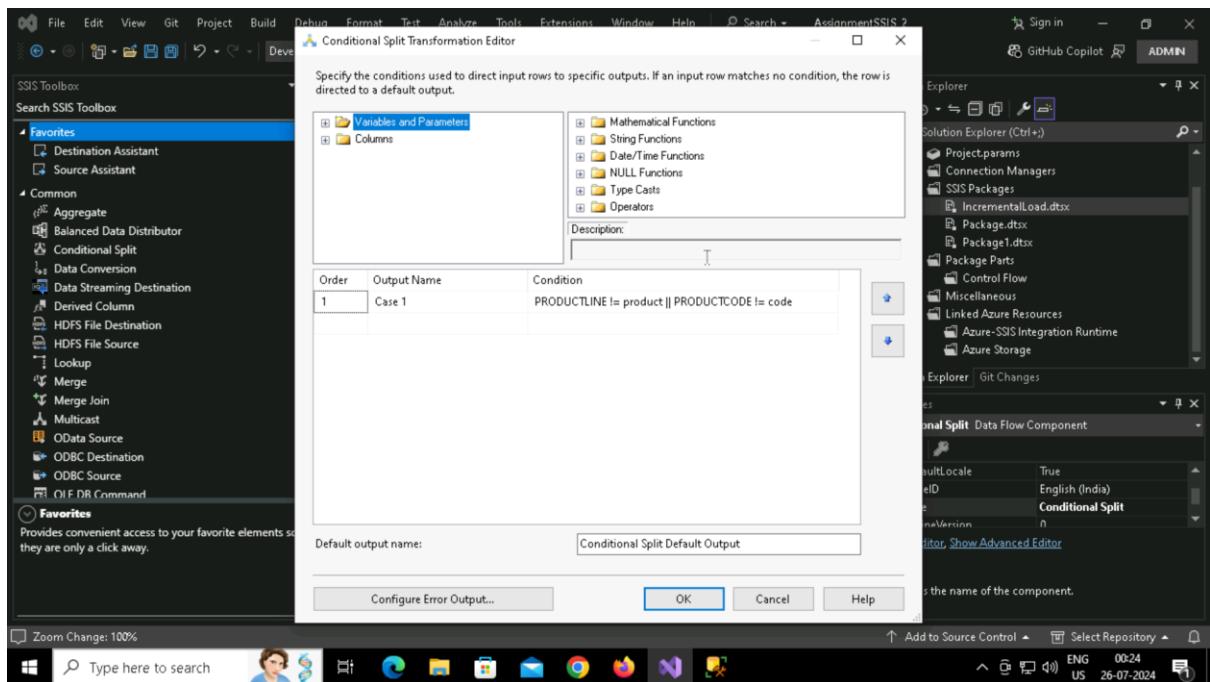
Requirements:

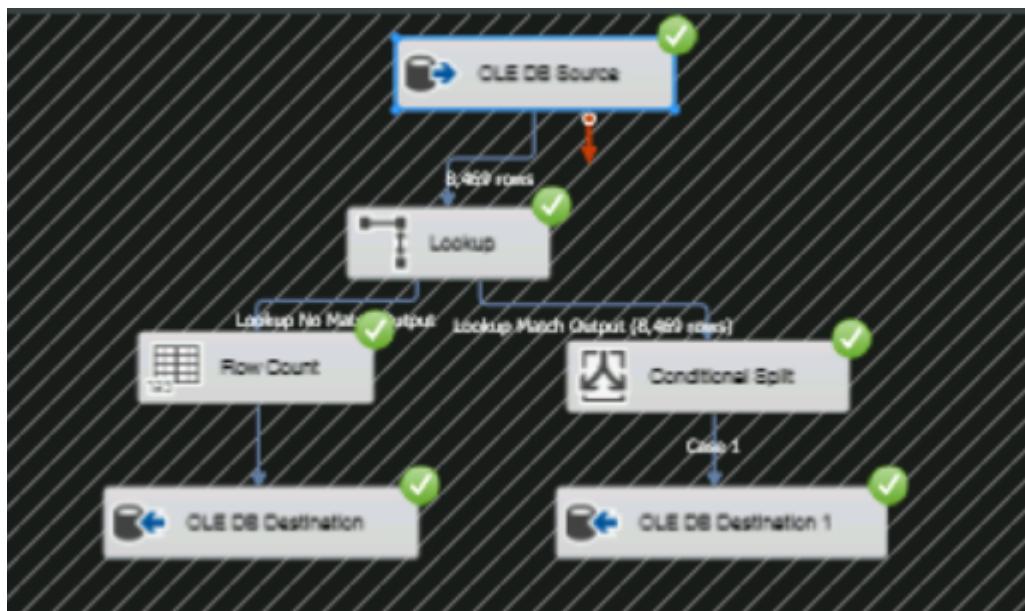
- Identify Changed Data: Use methods such as timestamps, change data capture using lookup, or checksums.
- Extract Only the Changed Data from the source.
- Update the Data Warehouse with the new and changed data only.











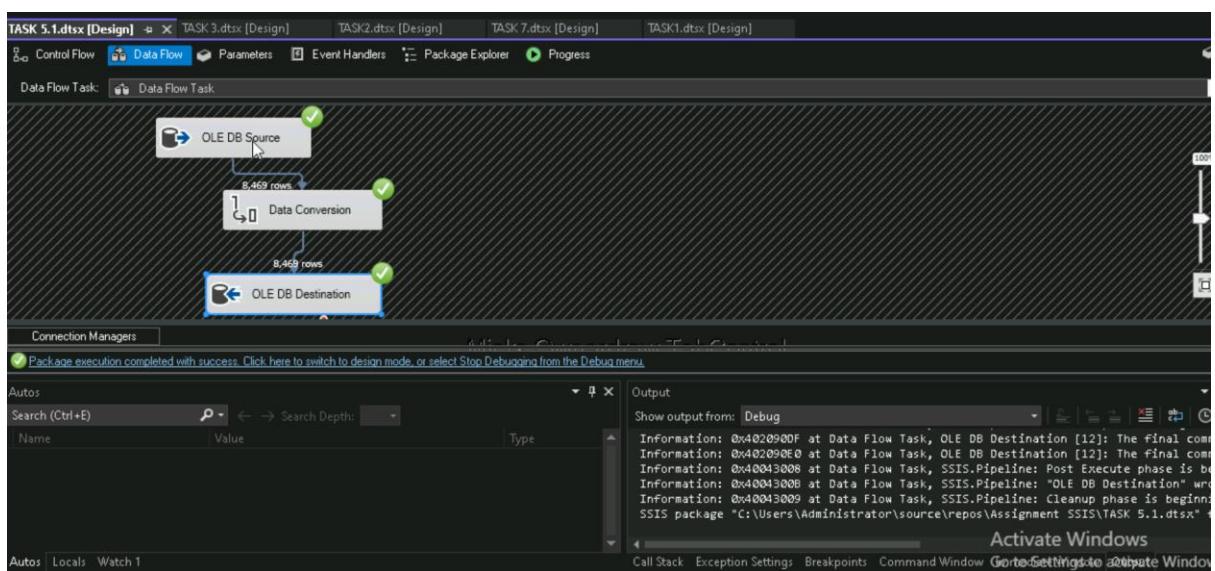
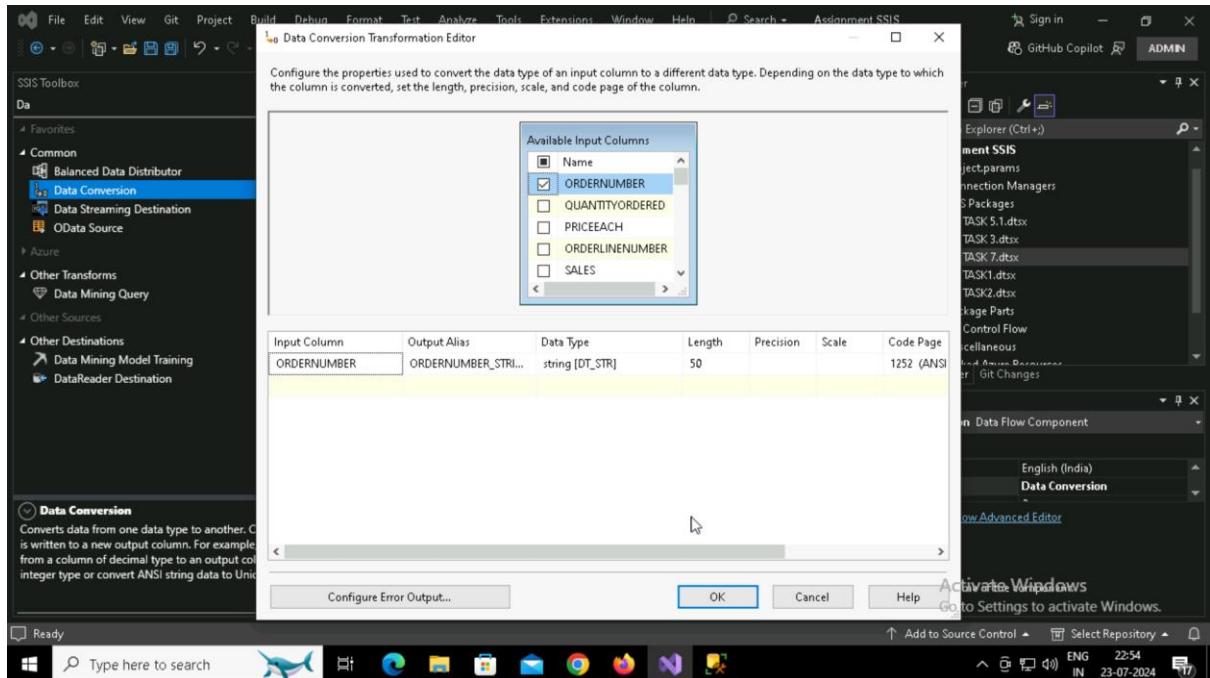
Task 5: Transformations

Scenario: Your company needs to transform raw data into a format suitable for reporting. You need to perform multiple transformations within an SSIS package.

Requirements:

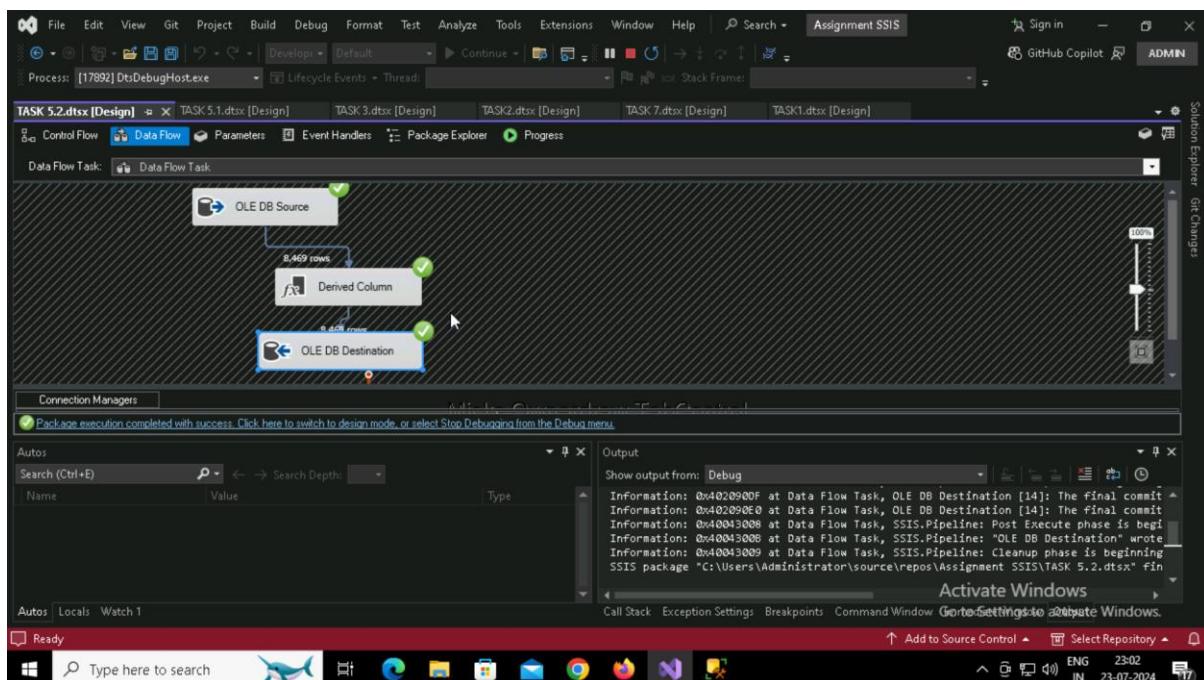
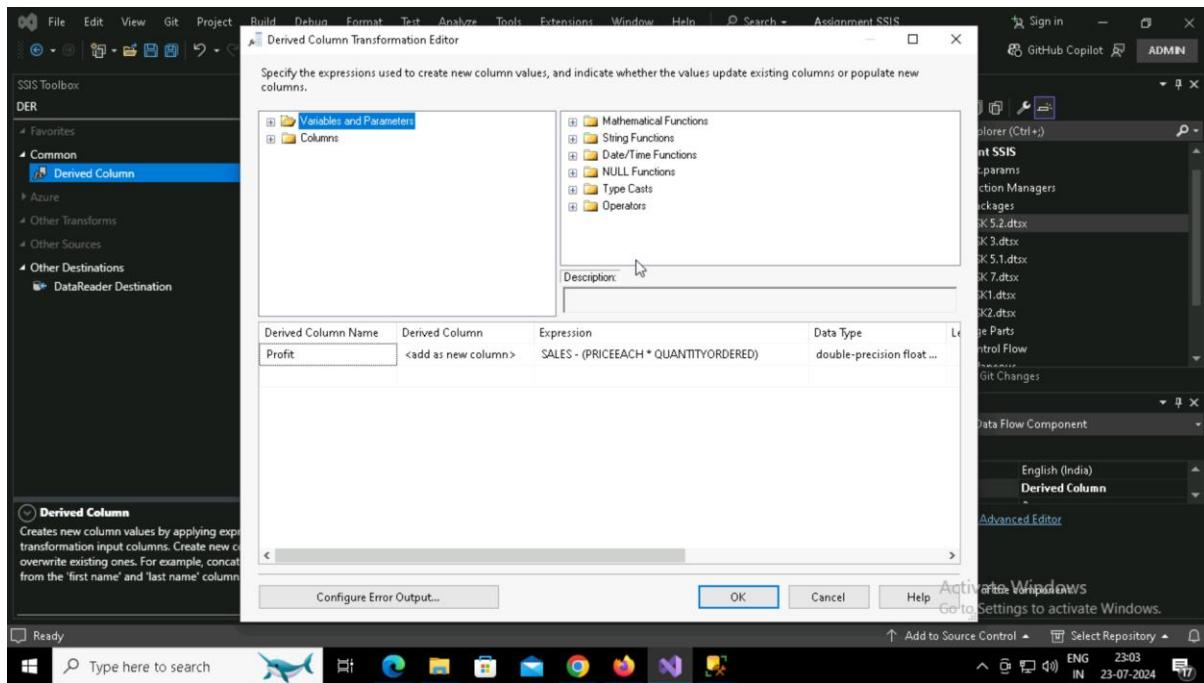
1. Extract Data from a source table using an OLE DB Source.
2. Apply Transformations such as:
 - Data Conversion
 - Derived Column
 - Conditional Split
 - Aggregate
3. Load Transformed Data into a destination table.

5.1 Data conversion



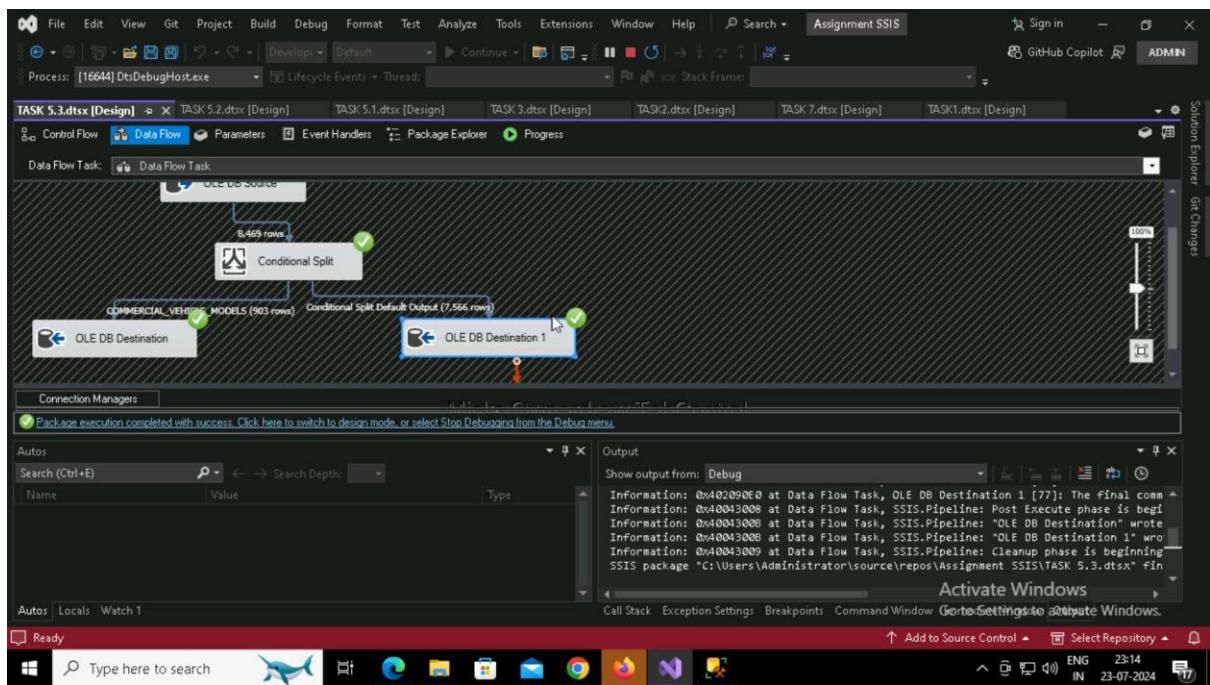
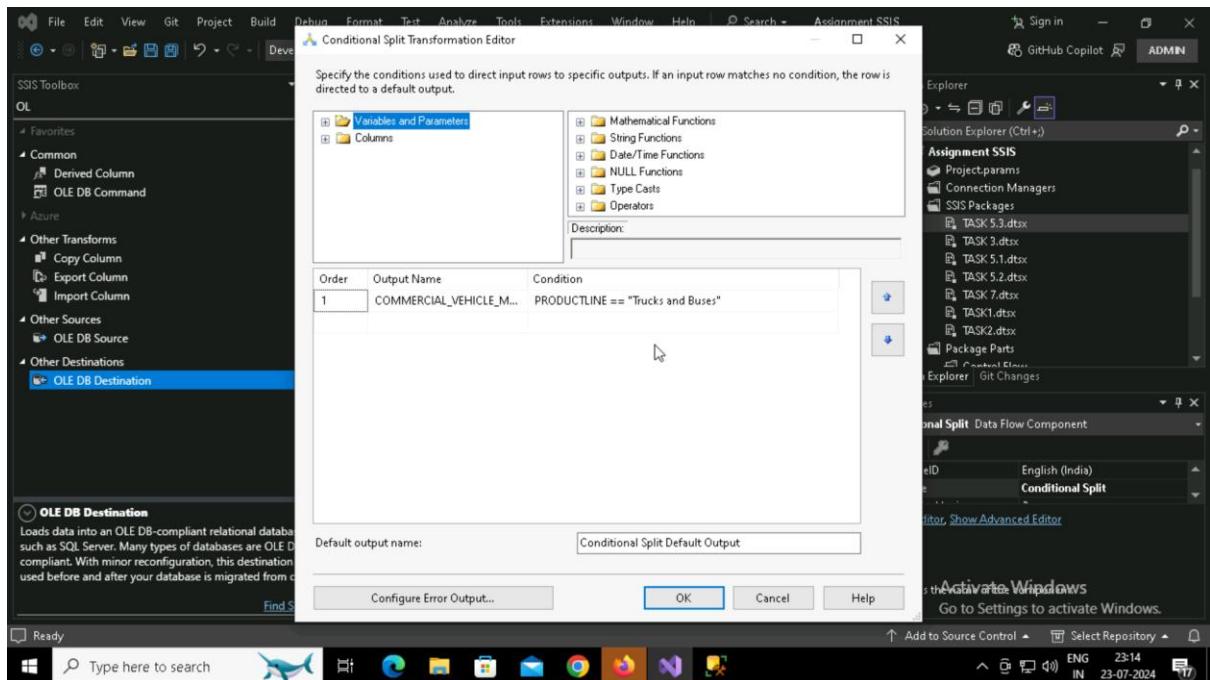
5.2

DataConversion

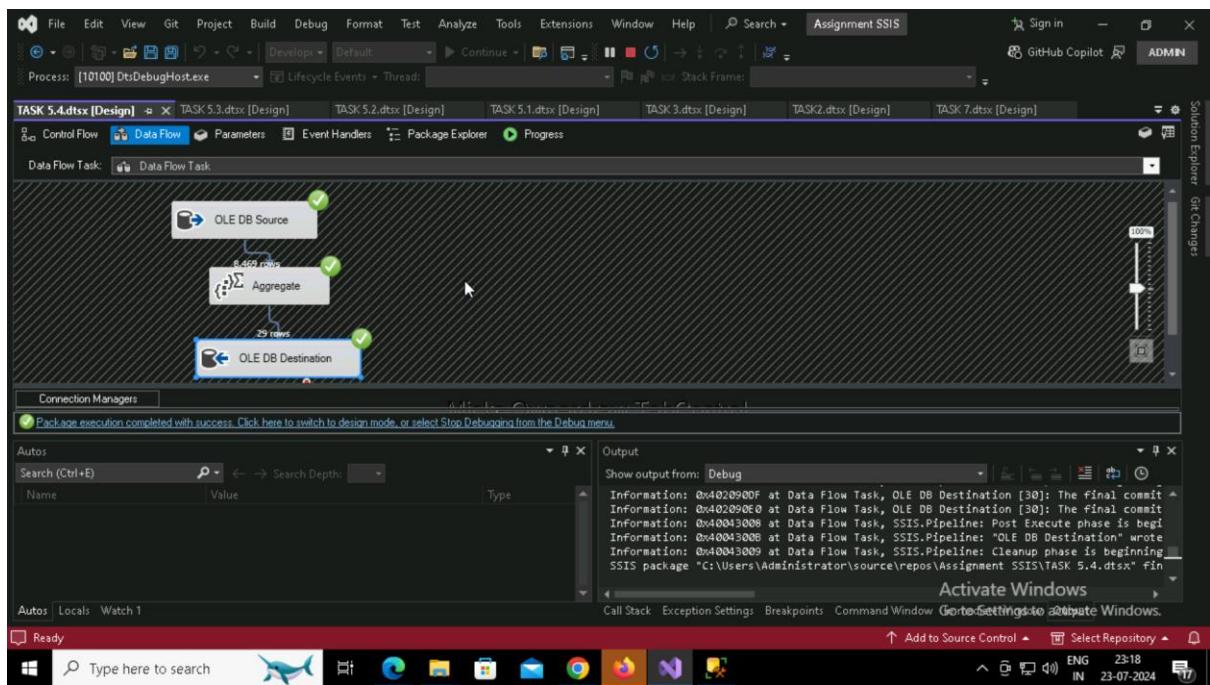
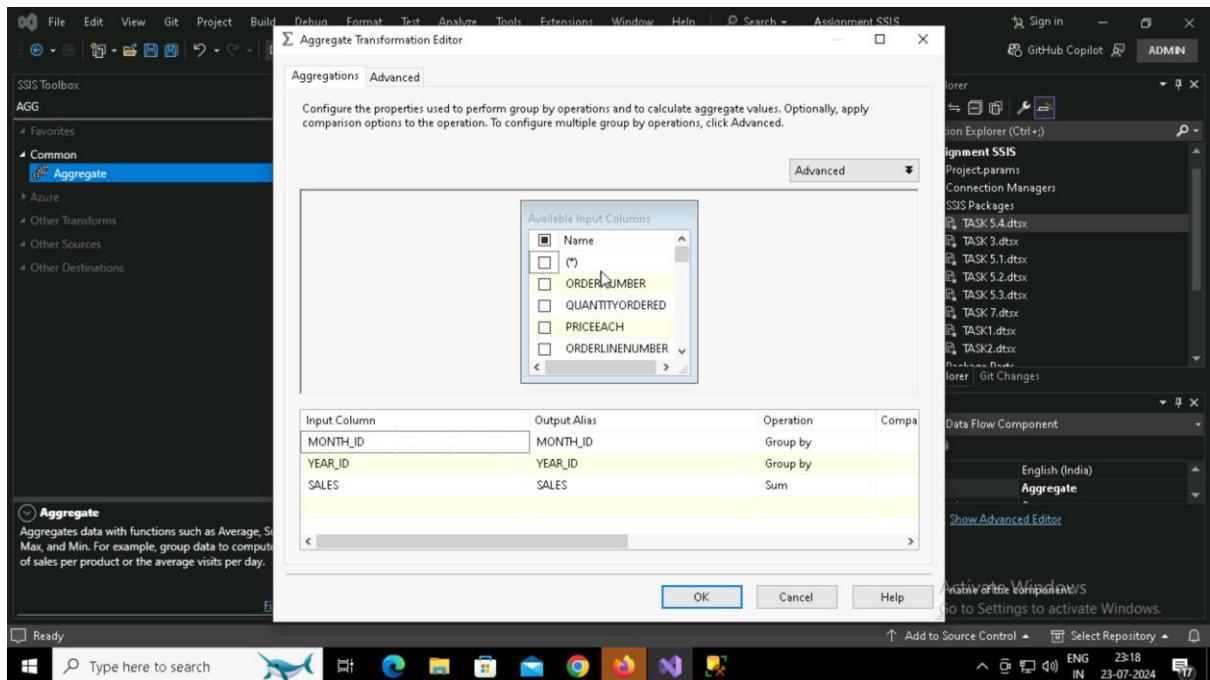


5.3

Conditional Split



5.4 Aggregate

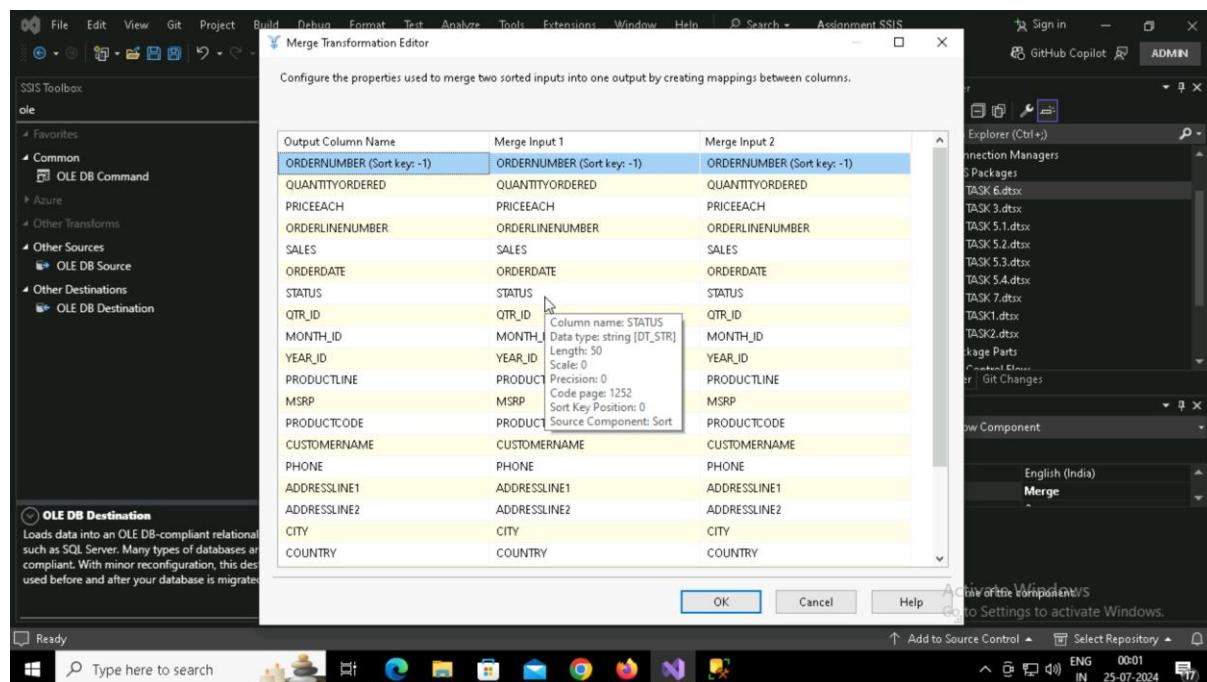
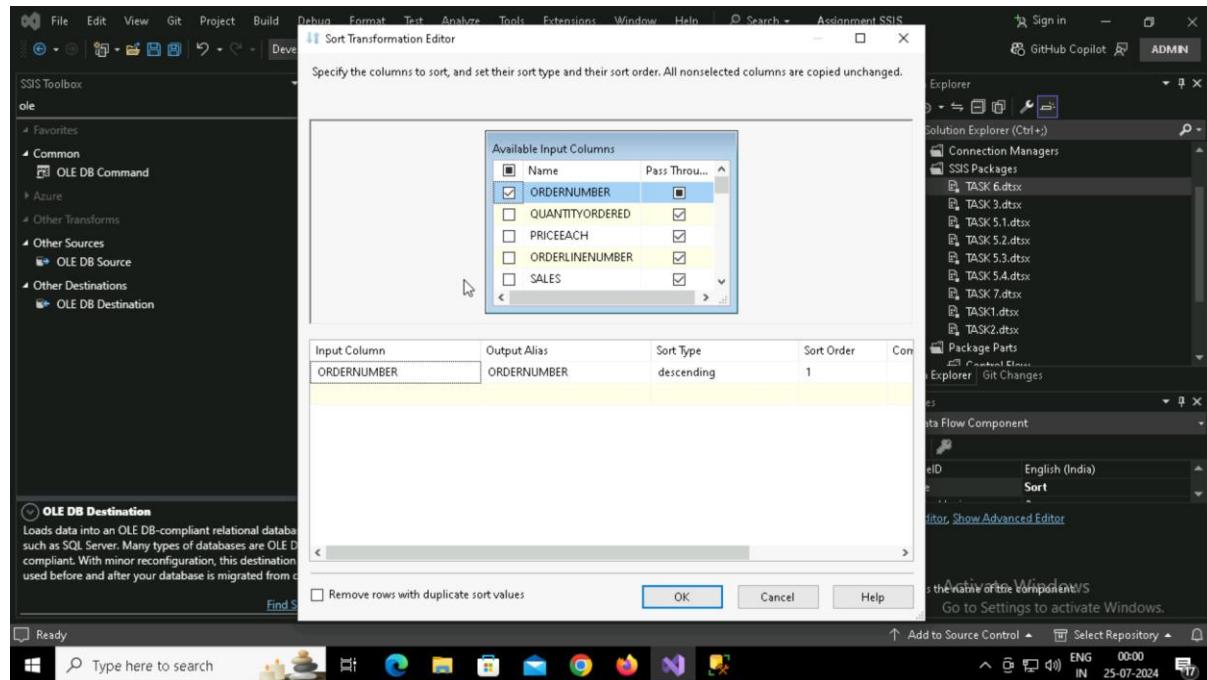


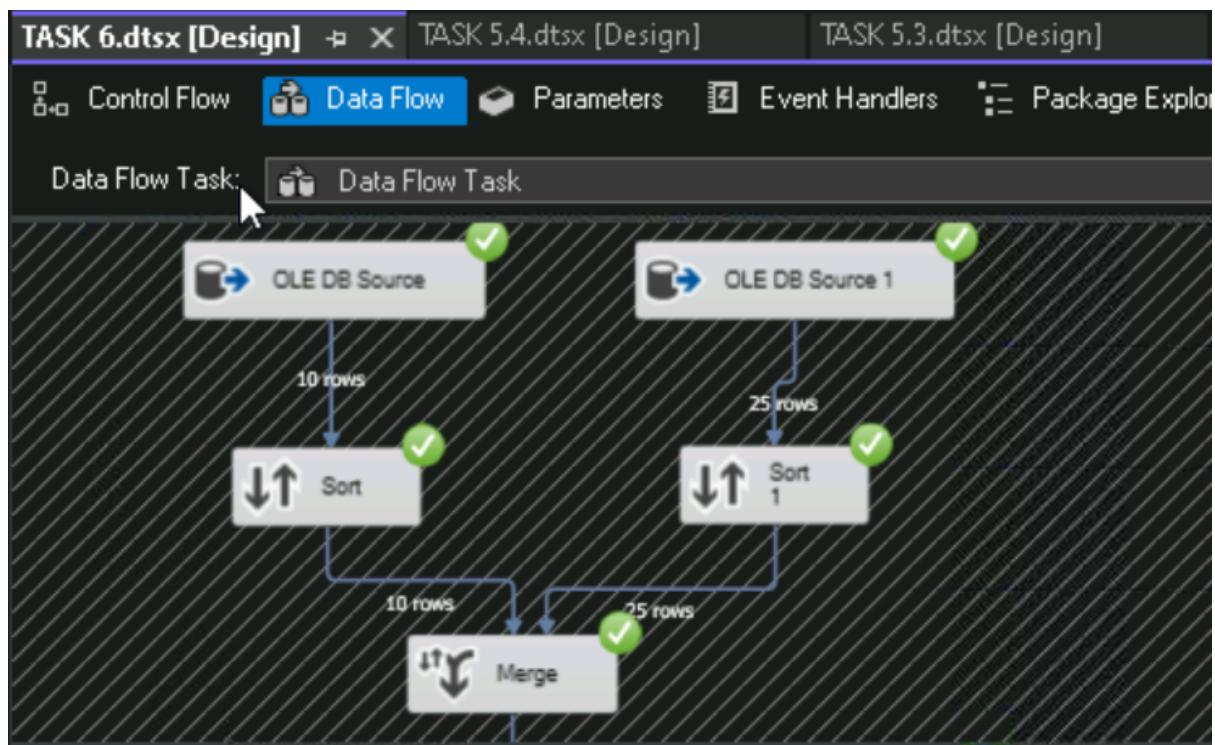
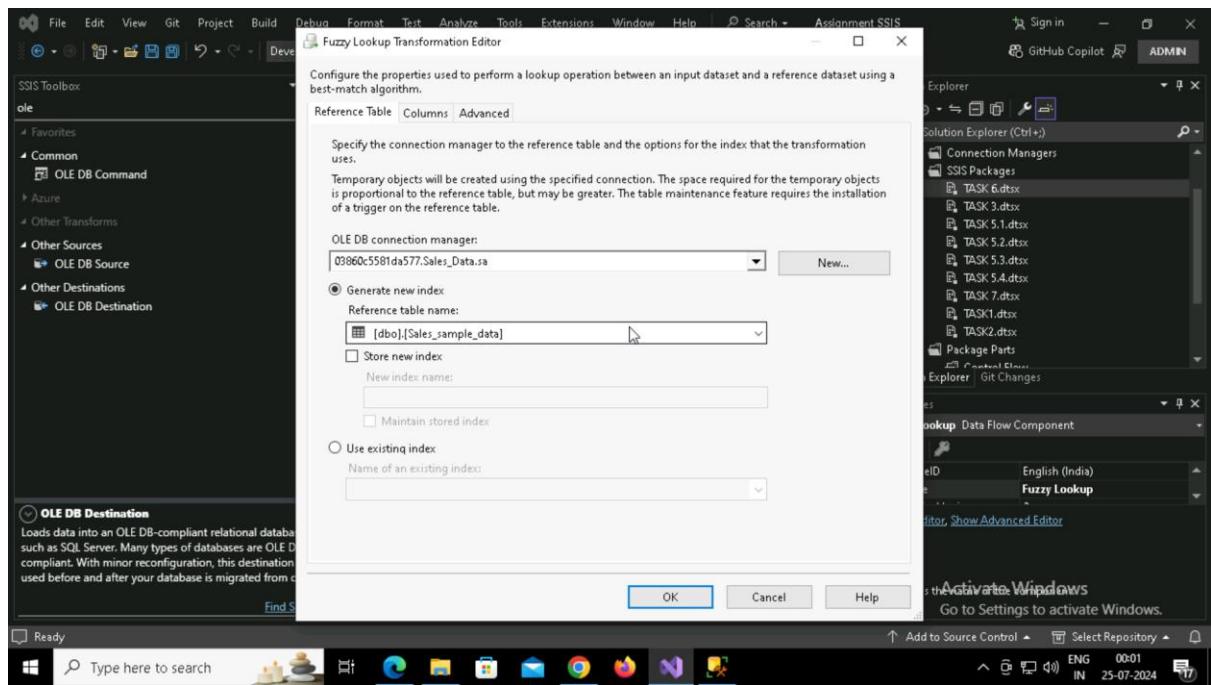
Task 6: MERGE & FUZZY LOOKUP

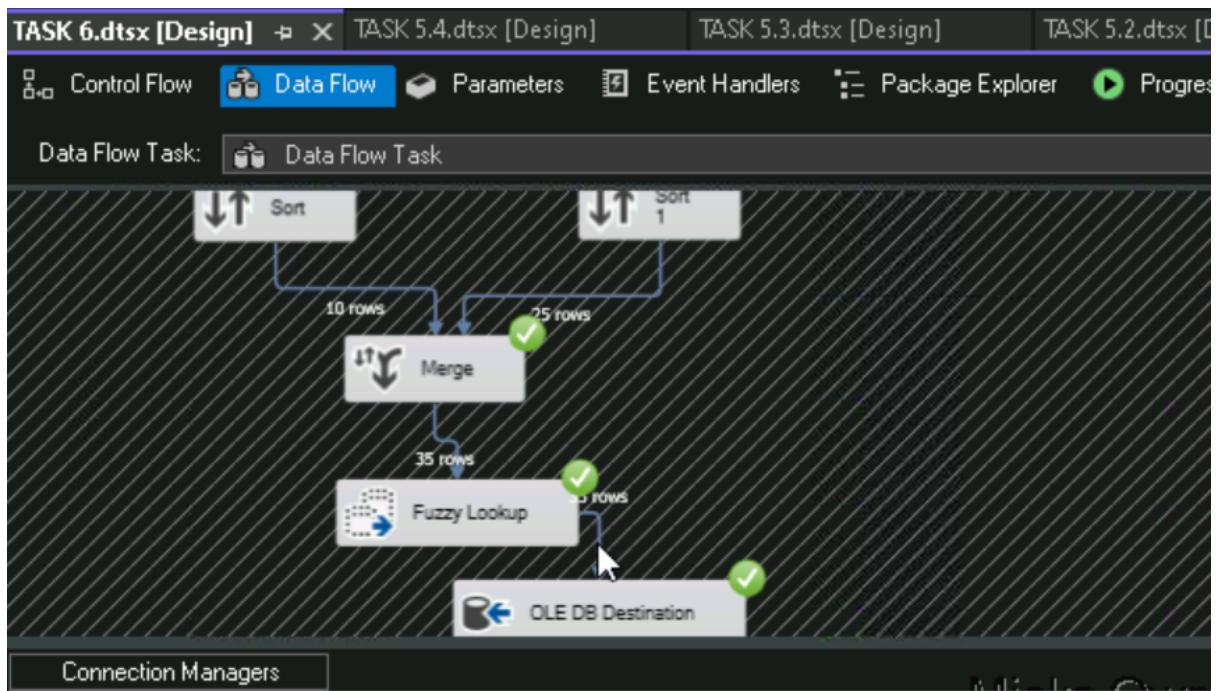
Scenario: You need to merge two datasets and use fuzzy matching to handle potential duplicates.

Requirements:

1. Extract Data from two source tables using OLE DB Sources.
2. Apply a Merge Join to combine the datasets based on a common key.
3. Use Fuzzy Lookup to identify and resolve duplicates in the merged data.
4. Load the Cleaned Data into a destination table.







Task 7: Using Script Task

Scenario: You need to perform a complex data transformation that is not supported by the standard SSIS components. A Script Task can be used to achieve this.

Requirements:

1. Add a Script Task to the Control Flow.
2. Write a Script: that performs the required transformation. e.g. Reading data from a file, processing it, and writing the results to a database table.
3. Execute the Script Task within an SSIS package.

C # CODE

```
public void Main()
{
    ConnectionManager cm = null;
    SqlConnection sqlConn = null;
    SqlCommand sqlComm = null;

    try
    {
        cm = Dts.Connections["sql_server_test"];
        sqlConn = (SqlConnection)cm.AcquireConnection(Dts.Transaction);

        string query = @"
            insert into script(sum_sales,productcode, qtr_id ,ranking)
            SELECT sum_sales, productcode, qtr_id, ranking
            FROM (
            SELECT productcode, sum_sales, qtr_id,
            RANK() OVER (PARTITION BY qtr_id ORDER BY sum_sales DESC) AS ranking
            FROM (
            SELECT SUM(sales) AS sum_sales, productcode, qtr_id
            FROM Sales_sample_data
            GROUP BY qtr_id, productcode
            ) AS ranked_data
            ) AS final_result where ranking<4;";

        sqlComm = new SqlCommand(query, sqlConn);
        sqlComm.ExecuteNonQuery();

        Dts.TaskResult = (int)ScriptResults.Success;
    }
    catch (Exception ex)
```

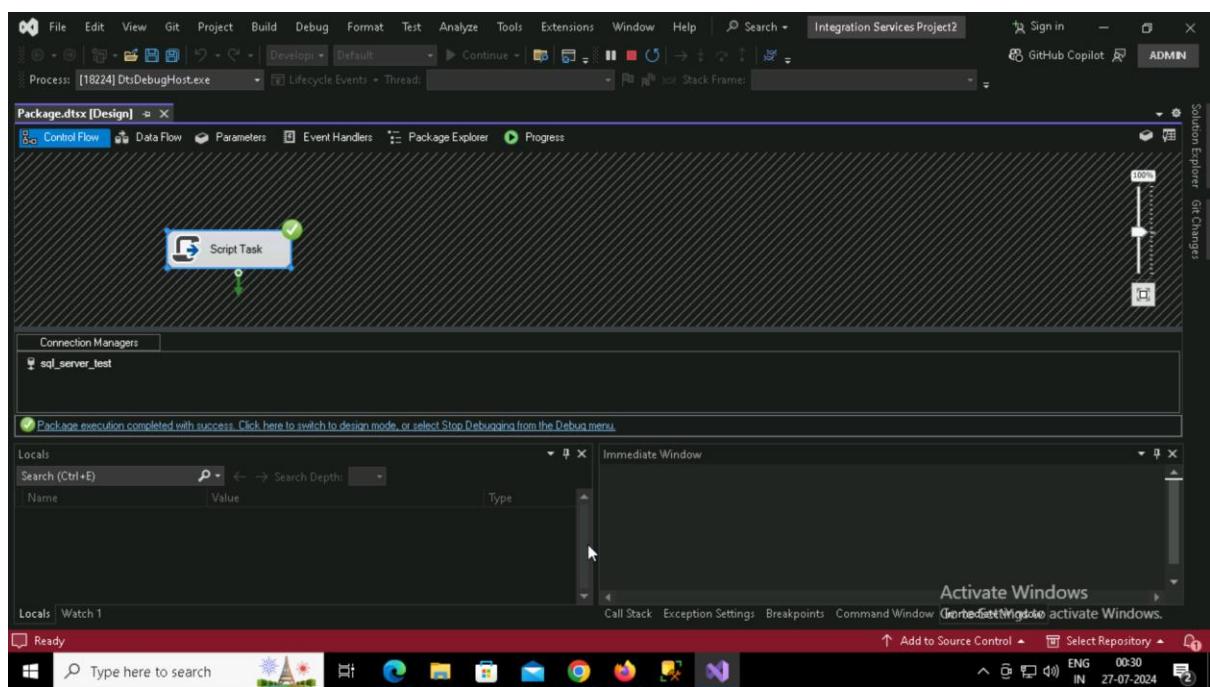
```

{
    Dts.Events.FireError(0, "Script Task Error", ex.Message, string.Empty, 0);

    Dts.TaskResult = (int)ScriptResults.Failure;
}

finally
{
    if (cm != null && sqlConn != null)
    {
        cm.ReleaseConnection(sqlConn);
    }
}
}
}

```



```
SELECT * FROM script
```

10 % ▾

Results Messages

	sum_sales	productcode	qtr_id	ranking
1	162738.119384766	S18_3232	1	1
2	127555.620117188	S12_1108	1	2
3	121015.618652344	S10_1949	1	3
4	210568.290527344	S18_3232	2	1
5	158426.071289063	S10_4698	2	2
6	116418.391113281	S18_2238	2	3
7	151056.031494141	S18_3232	3	1
8	118433.760498047	S12_1099	3	2
9	115272	S10_1949	3	3
10	340373.821289063	S18_3232	4	1
11	225117.658447266	S10_1949	4	2
12	185389.472167969	S18_2238	4	3