

**UNIVERSIDAD DON BOSCO**



# Datawarehouse y Minería de Datos

## DMD941 G01T

### DESAFÍO 1

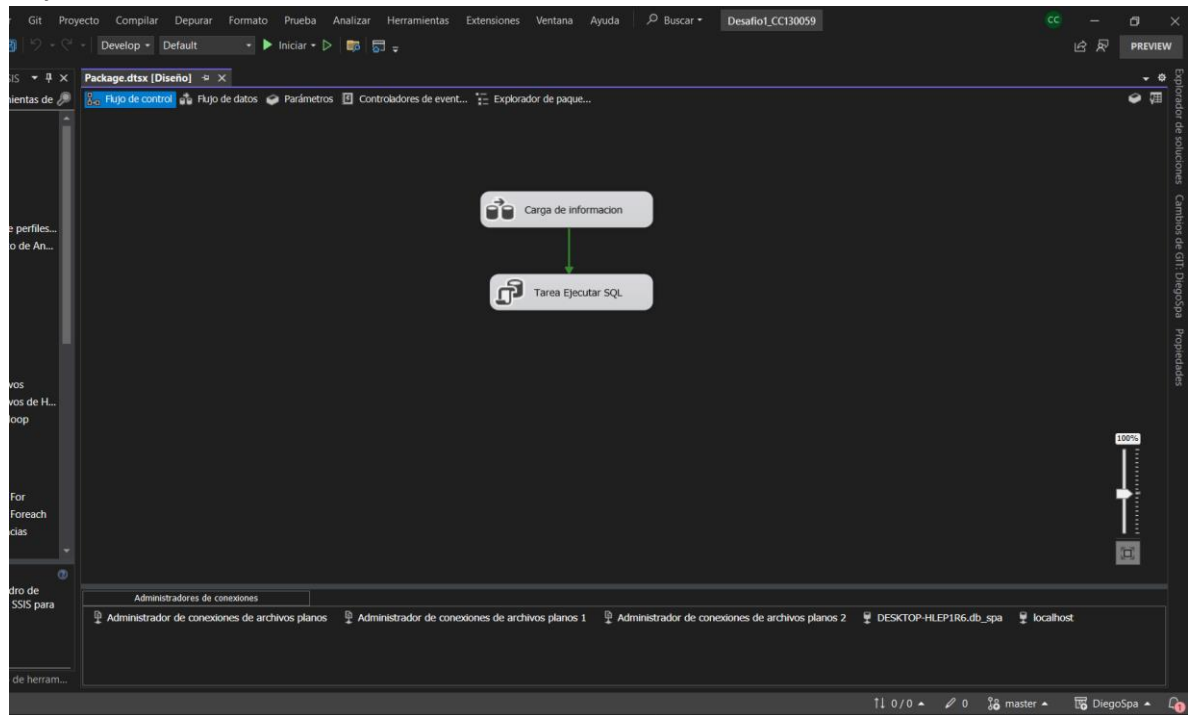
#### INTEGRANTES

<b>APELLIDOS</b>	<b>NOMBRES</b>	<b>CARNÉ</b>
CRUZ CALDERÓN	CARLOS DAVID	CC130059

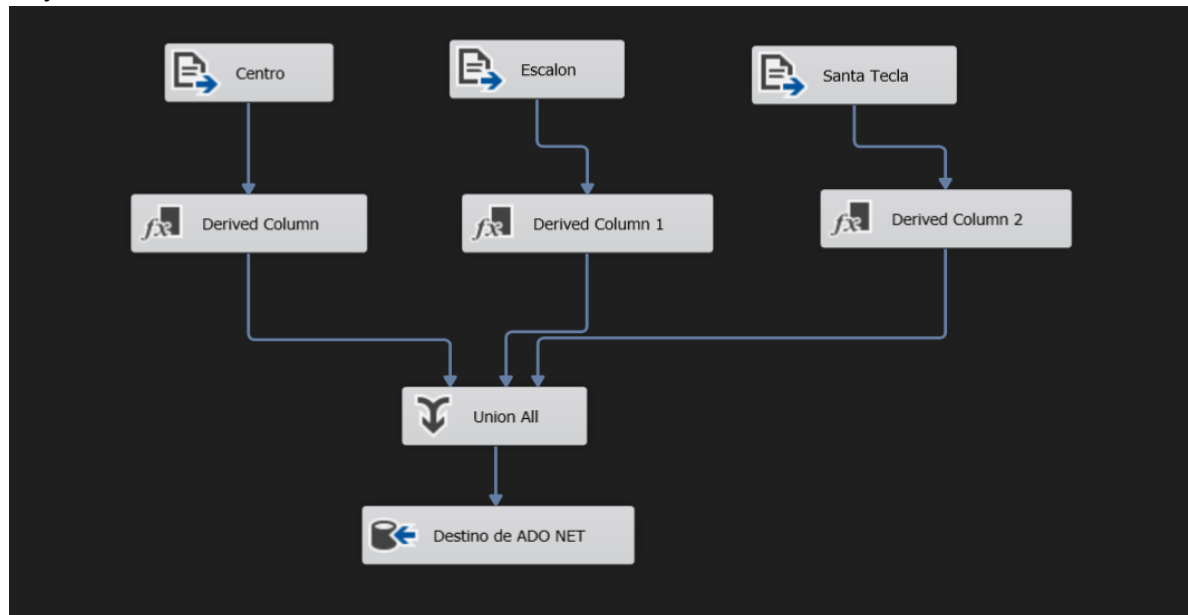
## Ejercicio #1

El Spa, "Diego", necesita segmentar sus clientes, para realizar una campaña de fidelización, y le pide a usted que efectúe un análisis de sus tres sucursales, que defina cuantos grupos y que características tienen.

1. Se crea el ETL para migrar la información de los 3 archivos a una base de datos en sql server
  - a. Flujo de control



b. Flujo de datos



c. Ejecutar SQL para generar un campo Grupo el cual realiza una segmentación llamando a un Procedimiento almacenado siguiente

```
USE [db_spa]
GO
/***** Object: StoredProcedure [dbo].[sp_calcular_segmento]    Script Date: 17/3/2024 14:33:11 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[sp_calcular_segmento] AS
-- Paso 2: Definir rangos para cada característica
DECLARE @EdadMin INT, @EdadMax INT;
SELECT @EdadMin = MIN(CAST(Edad AS INT)), @EdadMax = MAX(CAST(Edad AS INT)) FROM Spa;

DECLARE @IngresosMin FLOAT, @IngresosMax FLOAT;
SELECT @IngresosMin = MIN(CAST(ingresos AS FLOAT)), @IngresosMax = MAX(CAST(ingresos AS FLOAT)) FROM Spa;

DECLARE @PromVisitasMin FLOAT, @PromVisitasMax FLOAT;
SELECT @PromVisitasMin = MIN(CAST(PromVisit AS FLOAT)), @PromVisitasMax = MAX(CAST(PromVisit AS FLOAT)) FROM Spa;

-- Paso 3: Segmentar los clientes en grupos basados en los rangos definidos y actualizar la columna 'Grupo'
UPDATE Spa
SET Grupo =
CASE
WHEN CAST(Edad AS INT) BETWEEN @EdadMin AND (@EdadMax / 3) THEN 'Joven'
WHEN CAST(Edad AS INT) BETWEEN ((@EdadMax / 3) + 1) AND ((@EdadMax / 3) * 2) THEN 'Adulto'
ELSE 'Mayor'
END +
CASE
WHEN CAST(ingresos AS FLOAT) BETWEEN @IngresosMin AND ((@IngresosMax / 3) * 2) THEN '_Bajo'
ELSE '_Alto'
END +
CASE
WHEN CAST(PromVisit AS FLOAT) BETWEEN @PromVisitasMin AND ((@PromVisitasMax / 3) * 2) THEN '_Bajo'
ELSE '_Alto'
END +
'_' + Sexo;
```

d. Prueba de funcionamiento

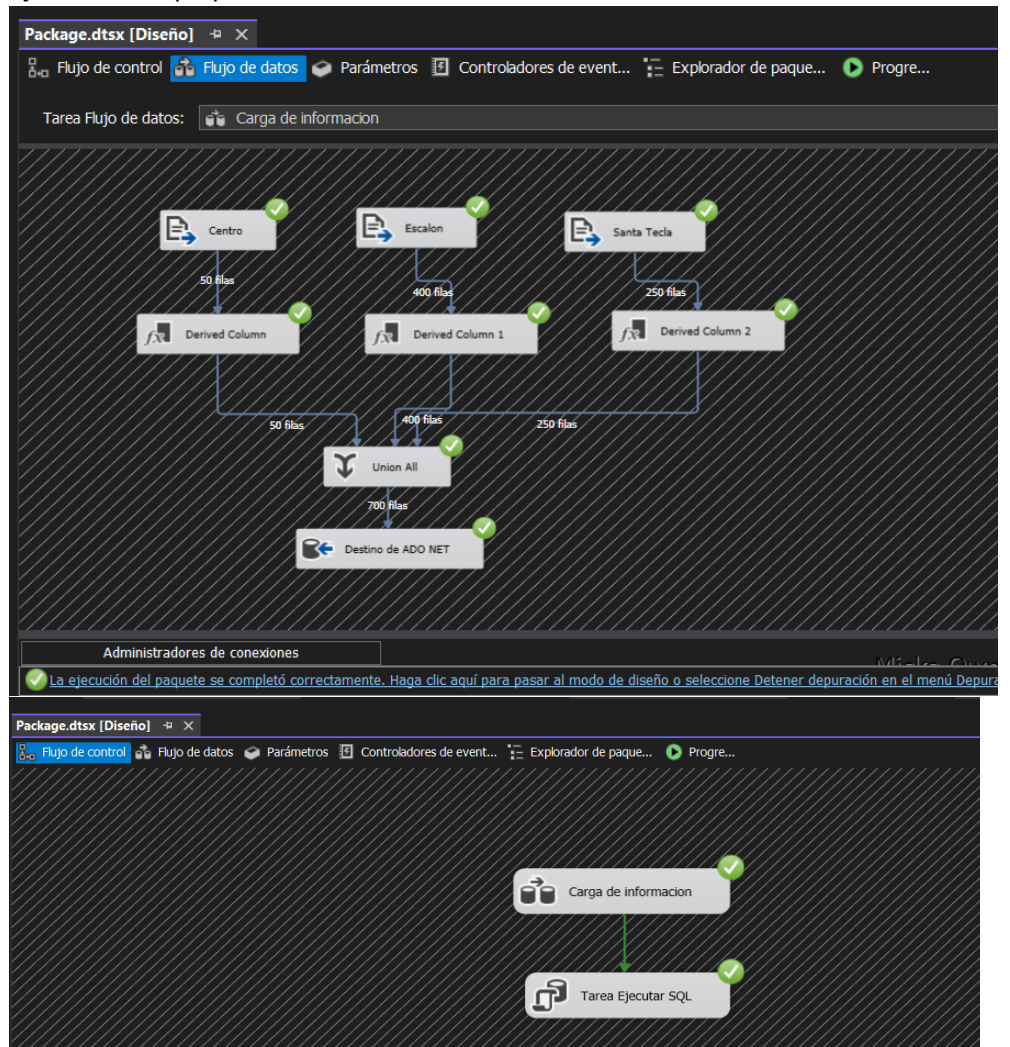
i. Tabla vacía

The screenshot shows the SQL Server Enterprise Manager interface. The top toolbar includes buttons for 'Execute', 'Save', 'Print', 'Format', 'Copy', 'Paste', 'Find', 'Find Next', 'Find Previous', 'Go To', 'Stop', 'Help', and 'Refresh'. The main window displays a query in the 'SQLQuery7.sql' file, which is part of a project named 'D:\...R6\ROG STRIX (83)'. The query is a SELECT statement with the following structure:

```
SELECT TOP (1000) [id]
      ,[cliente]
      ,[Sexo]
      ,[ingresos]
      ,[PromVisit]
      ,[Edad]
      ,[Sauna]
      ,[Masaje]
      ,[Hidro]
      ,[Yoga]
      ,[Sucursal]
      ,[Grupo]
FROM [db_spa].[dbo].[Spa]
```

Below the query editor, the 'Results' tab is active, showing an empty grid with 12 columns: id, cliente, Sexo, ingresos, PromVisit, Edad, Sauna, Masaje, Hidro, Yoga, Sucursal, and Grupo. The 'Messages' tab is also visible but empty. The status bar at the bottom indicates 'es' and 'ito'.

## ii. Ejecucion del paquete



### 2. Analisis de los segmentos de fidelización generados

El campo Grupo es generado para segmentar a los clientes de la siguiente manera:

**Edad** = Mayor, Joven, Adulto

**Ingresos** = Bajo y Alto

**Promedio de Visitas** = Bajo y Alto

Y se obtiene el siguiente grupo: edad\_ingresos\_promvisitas\_sexo

```
--top 10 de segmento de clientes por miembros
SELECT TOP 10 Sucursal, Grupo, COUNT(*) AS TotalClientes
FROM Spa
GROUP BY Sucursal, Grupo
order by 3 desc;
```

	Sucursal	Grupo	TotalClientes
1	ESCALON	Mayor_Bajo_Bajo_1	44
2	ESCALON	Adulto_Bajo_Bajo_0	43
3	SANTA TECLA	Adulto_Bajo_Bajo_0	43
4	SANTA TECLA	Mayor_Bajo_Bajo_0	39
5	ESCALON	Adulto_Bajo_Bajo_1	38
6	ESCALON	Mayor_Bajo_Bajo_0	35
7	ESCALON	Adulto_Alto_Bajo_1	32
8	ESCALON	Mayor_Alto_Bajo_1	26
9	ESCALON	Adulto_Bajo_Alto_0	23
10	SANTA TECLA	Mayor_Bajo_Alto_0	22

```
--Segmentos por tipo de Sauna
select Sucursal, Grupo, COUNT(*) as TotalClientes
from Spa
Where Sauna = 1
group by Sucursal, Grupo
Order by 3 desc;
```

100 %

	Sucursal	Grupo	TotalClientes
1	ESCALON	Mayor_Bajo_Bajo_1	23
2	ESCALON	Adulto_Bajo_Bajo_0	22
3	ESCALON	Adulto_Bajo_Bajo_1	20
4	ESCALON	Adulto_Alto_Bajo_1	16
5	ESCALON	Mayor_Bajo_Bajo_0	15
6	SANTA TECLA	Adulto_Bajo_Bajo_0	13
7	ESCALON	Mayor_Bajo_Alto_0	13
8	ESCALON	Mayor_Alto_Bajo_1	11
9	ESCALON	Adulto_Bajo_Alto_0	11
10	ESCALON	Adulto_Bajo_Alto_1	10

```
--Segmentos por tipo de Masaje
select Sucursal, Grupo, COUNT(*) as TotalClientes
from Spa
Where Masaje = 1
group by Sucursal, Grupo
Order by 3 desc;
```

	Sucursal	Grupo	TotalClientes
1	ESCALON	Mayor_Bajo_Bajo_0	24
2	SANTA TECLA	Adulto_Bajo_Bajo_0	20
3	ESCALON	Mayor_Bajo_Bajo_1	19
4	SANTA TECLA	Mayor_Bajo_Bajo_0	18
5	ESCALON	Adulto_Alto_Bajo_1	18
6	ESCALON	Adulto_Bajo_Alto_0	18
7	ESCALON	Adulto_Bajo_Bajo_0	17
8	SANTA TECLA	Mayor_Bajo_Alto_0	14
9	ESCALON	Mayor_Bajo_Alto_1	13
10	ESCALON	Adulto_Bajo_Bajo_1	13

```
--Segmentos por tipo de Hidro
select Sucursal, Grupo, COUNT(*) as TotalClientes
from Spa
Where Hidro = 1
group by Sucursal, Grupo
Order by 3 desc;
```

	Sucursal	Grupo	TotalClientes
1	ESCALON	Mayor_Bajo_Bajo_1	24
2	ESCALON	Adulto_Bajo_Bajo_0	19
3	ESCALON	Mayor_Bajo_Bajo_0	18
4	ESCALON	Adulto_Bajo_Bajo_1	17
5	ESCALON	Adulto_Alto_Bajo_1	14
6	SANTA TECLA	Adulto_Bajo_Bajo_0	14
7	ESCALON	Mayor_Alto_Bajo_1	12
8	ESCALON	Mayor_Bajo_Alto_1	12
9	ESCALON	Adulto_Alto_Bajo_0	12
10	ESCALON	Adulto_Bajo_Alto_1	12

```
--Segmentos por tipo de Yoga
select Sucursal, Grupo, COUNT(*) as TotalClientes
from Spa
Where Yoga = 1
group by Sucursal, Grupo
Order by 3 desc;
```

100 %

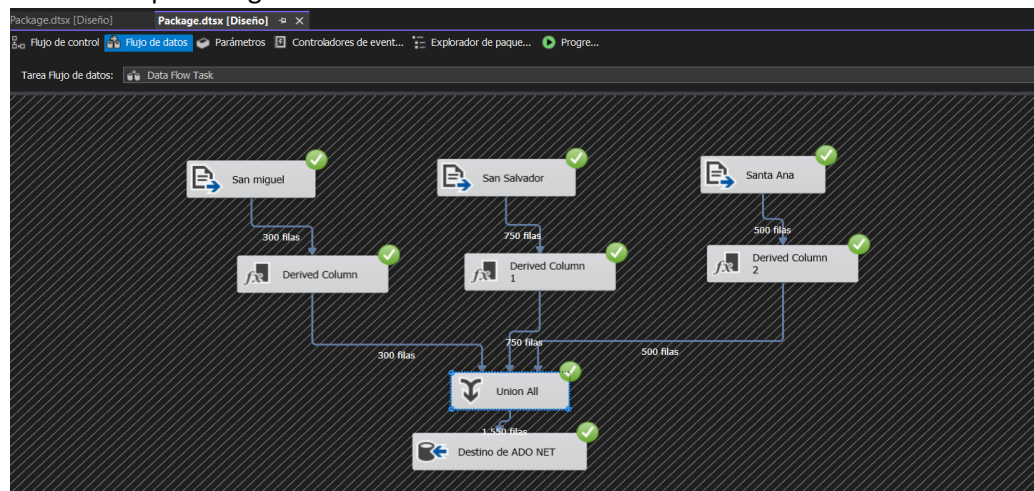
Results Messages

	Sucursal	Grupo	TotalClientes
1	SANTA TECLA	Adulto_Bajo_Bajo_0	25
2	SANTA TECLA	Mayor_Bajo_Bajo_0	24
3	ESCALON	Adulto_Bajo_Bajo_1	22
4	ESCALON	Adulto_Bajo_Bajo_0	21
5	ESCALON	Mayor_Bajo_Bajo_1	21
6	ESCALON	Adulto_Alto_Bajo_1	19
7	ESCALON	Mayor_Bajo_Bajo_0	16
8	ESCALON	Adulto_Bajo_Alto_0	13
9	SANTA TECLA	Adulto_Bajo_Alto_0	13
10	SANTA TECLA	Mayor_Bajo_Alto_0	13

## Ejercicio #2

La Floristería "Fiorella" quiere saber cómo se compran sus productos, y tiene la data de tres departamentos del país, por lo cual les pide su opinión sobre qué productos sobresalen, que combinaciones son mejores y quieren este estudio por departamento y también por país.

1. Se crea ETL para cargar los datos de los archivos





## 2. ANALISIS

### a. Ventas de cada producto

```
--Analizar ventas de cada producto
SELECT
    ProductName,
    SUM(CAST(ProductValue AS INT)) AS TotalSales
FROM
    (SELECT
        'Rosas' AS ProductName, [Rosas] AS ProductValue FROM Fiorella UNION ALL
        SELECT 'Claveles', [Claveles] FROM Fiorella UNION ALL
        SELECT 'Macetas', [Macetas] FROM Fiorella UNION ALL
        SELECT 'Tierra', [Tierra] FROM Fiorella UNION ALL
        SELECT 'Girasoles', [Girasoles] FROM Fiorella UNION ALL
        SELECT 'Hortensia', [Hortensia] FROM Fiorella UNION ALL
        SELECT 'Globos', [Globos] FROM Fiorella UNION ALL
        SELECT 'Tarjetas', [Tarjetas] FROM Fiorella UNION ALL
        SELECT 'Orquideas', [Orquideas] FROM Fiorella UNION ALL
        SELECT 'Carmesi', [Carmesi] FROM Fiorella UNION ALL
        SELECT 'Lirios', [Lirios] FROM Fiorella UNION ALL
        SELECT 'Aurora', [Aurora] FROM Fiorella UNION ALL
        SELECT 'Tulipanes', [Tulipanes] FROM Fiorella
    ) AS Products
GROUP BY ProductName
ORDER BY TotalSales DESC;
```

Results		Messages	
	ProductName	TotalSales	
1	Rosas	945	
2	Globos	892	
3	Aurora	804	
4	Orquideas	797	
5	Lirios	795	
6	Girasoles	787	
7	Tarjetas	779	
8	Macetas	778	
9	Hortensia	774	
10	Tulipanes	753	
11	Carmesi	747	
12	Tierra	745	
13	Claveles	733	

b. Combinación de productos

--Analizar combinaciones de productos

```
SELECT
    Combination,
    COUNT(*) AS TotalSales
FROM
    (SELECT
        CONCAT_WS(',',
            CASE WHEN [Rosas] = '1' THEN 'Rosas' ELSE NULL END,
            CASE WHEN [Claveles] = '1' THEN 'Claveles' ELSE NULL END,
            CASE WHEN [Macetas] = '1' THEN 'Macetas' ELSE NULL END,
            CASE WHEN [Tierra] = '1' THEN 'Tierra' ELSE NULL END,
            CASE WHEN [Girasoles] = '1' THEN 'Girasoles' ELSE NULL END,
            CASE WHEN [Hortensia] = '1' THEN 'Hortensia' ELSE NULL END,
            CASE WHEN [Globos] = '1' THEN 'Globos' ELSE NULL END,
            CASE WHEN [Tarjetas] = '1' THEN 'Tarjetas' ELSE NULL END,
            CASE WHEN [Orquideas] = '1' THEN 'Orquideas' ELSE NULL END,
            CASE WHEN [Carmesi] = '1' THEN 'Carmesi' ELSE NULL END,
            CASE WHEN [Lirios] = '1' THEN 'Lirios' ELSE NULL END,
            CASE WHEN [Aurora] = '1' THEN 'Aurora' ELSE NULL END,
            CASE WHEN [Tulipanes] = '1' THEN 'Tulipanes' ELSE NULL END
        ) AS Combinacion
    FROM Fiorella) AS Combinations
WHERE Combination IS NOT NULL
GROUP BY Combination
ORDER BY TotalSales DESC;
```

	Combinacion	TotalSales
1	Rosas,Claveles,Tierra,Hortensia,Orquideas,Carmesi...	4
2	Rosas,Globos,Tarjetas,Orquideas,Carmesi,Aurora	4
3	Rosas,Macetas,Tierra,Hortensia,Globos,Tarjetas,Liri...	4
4	Rosas,Macetas,Globos	3
5	Rosas,Macetas,Globos,Aurora	3
6	Rosas,Macetas,Girasoles,Hortensia,Globos,Tarjetas...	3
7	Rosas,Globos,Orquideas,Carmesi,Lirios,Tulipanes	3
8	Rosas,Macetas,Girasoles,Globos,Tarjetas,Orquidea...	3
9	Rosas,Claveles,Tierra,Globos,Orquideas,Tulipanes	3
10	Rosas,Claveles,Macetas,Tierra,Girasoles,Hortensia,...	3
11	Claveles,Girasoles,Hortensia,Tarjetas,Orquideas,Liri...	3
12	Macetas,Tierra,Girasoles,Globos,Orquideas,Carmes...	3
13	Rosas,Claveles,Globos,Tarjetas,Orquideas	3
14	Rosas,Claveles,Lirios,Tulipanes	3
15	Rosas,Claveles,Macetas,Tierra,Girasoles,Aurora,Tul...	3
16	Rosas,Claveles,Macetas,Tierra,Girasoles,Globos,Ta...	2
17	Rosas,Claveles,Macetas,Girasoles,Hortensia,Globo...	2
18	Rosas,Claveles,Macetas,Girasoles,Hortensia,Tarjet...	2
19	Rosas,Claveles,Macetas,Girasoles,Hortensia,Tarjet...	2
20	Rosas,Claveles,Macetas,Girasoles,Hortensia,Globo...	2
21	Rosas,Claveles,Macetas,Girasoles,Tarjetas,Orquide...	2
22	Rosas,Claveles,Macetas,Girasoles,Tarjetas,Orquide...	2
23	Rosas,Claveles,Macetas,Globos,Lirios,Tulipanes	2
24	Rosas,Claveles,Hortensia,Globos,Lirios,Tulipanes	2
25	Rosas,Claveles,Globos,Lirios,Aurora	2
26	Rosas,Claveles,Hortensia,Globos,Tarjetas,Orquidea...	2
27	Rosas,Claveles,Macetas,Girasoles,Globos,Carmesi	2
28	Rosas,Claveles,Globos,Tarjetas,Lirios	2
29	Rosas,Claveles,Macetas,Girasoles,Globos,Orquidea...	2
30	Rosas,Claveles,Macetas,Girasoles,Globos,Tarjetas,...	2
31	Rosas,Claveles,Macetas,Girasoles,Globos,Orquidea...	2

c. Ventas por departamento

```

-- Analizar ventas por departamento
SELECT
    Departamento,
    SUM(CAST(ProductValue AS INT)) AS TotalSales
FROM
    (SELECT
        Departamento,
        'Rosas' AS ProductName, [Rosas] AS ProductValue FROM Fiorella UNION ALL
        SELECT Departamento, 'Claveles', [Claveles] FROM Fiorella UNION ALL
        SELECT Departamento, 'Macetas', [Macetas] FROM Fiorella UNION ALL
        SELECT Departamento, 'Tierra', [Tierra] FROM Fiorella UNION ALL
        SELECT Departamento, 'Girasoles', [Girasoles] FROM Fiorella UNION ALL
        SELECT Departamento, 'Hortensia', [Hortensia] FROM Fiorella UNION ALL
        SELECT Departamento, 'Globos', [Globos] FROM Fiorella UNION ALL
        SELECT Departamento, 'Tarjetas', [Tarjetas] FROM Fiorella UNION ALL
        SELECT Departamento, 'Orquideas', [Orquideas] FROM Fiorella UNION ALL
        SELECT Departamento, 'Carmesi', [Carmesi] FROM Fiorella UNION ALL
        SELECT Departamento, 'Lirios', [Lirios] FROM Fiorella UNION ALL
        SELECT Departamento, 'Aurora', [Aurora] FROM Fiorella UNION ALL
        SELECT Departamento, 'Tulipanes', [Tulipanes] FROM Fiorella
    ) AS SalesByDepartment
GROUP BY Departamento
ORDER BY Departamento, TotalSales DESC;

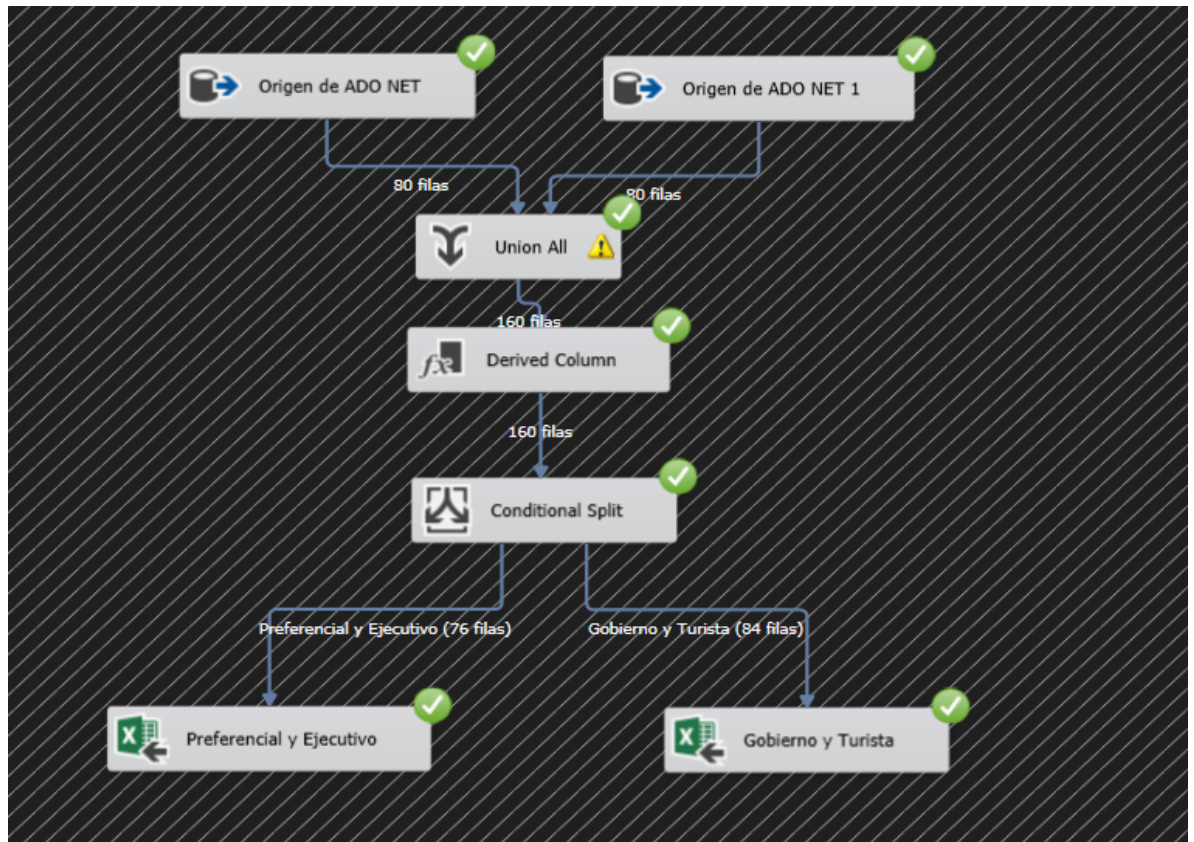
```

	Departamento	TotalSales
1	San Miguel	1962
2	San Salvador	5277
3	Santa Ana	3090

### Ejercicio #3

La telefonía “FioDio” solicita realizar un ETL que exporte una base de datos de Mysql y SQL Server, al final el destino serán dos archivos de Excel en donde en un archivo estarán los clientes preferenciales y ejecutivos y en el segundo los de gobierno y turista, adicional en los archivos de Excel se deberá crear un campo código de país, que se llenará sustraendo los dos primeros caracteres de código cliente.

## 1. Creacion de ETL



## 2. Creacion de la columna derivada para la creación del código de país

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables y parámetros

Columns

Funciones matemáticas

Funciones de cadena

Funciones de fecha y hora

Funciones NULL

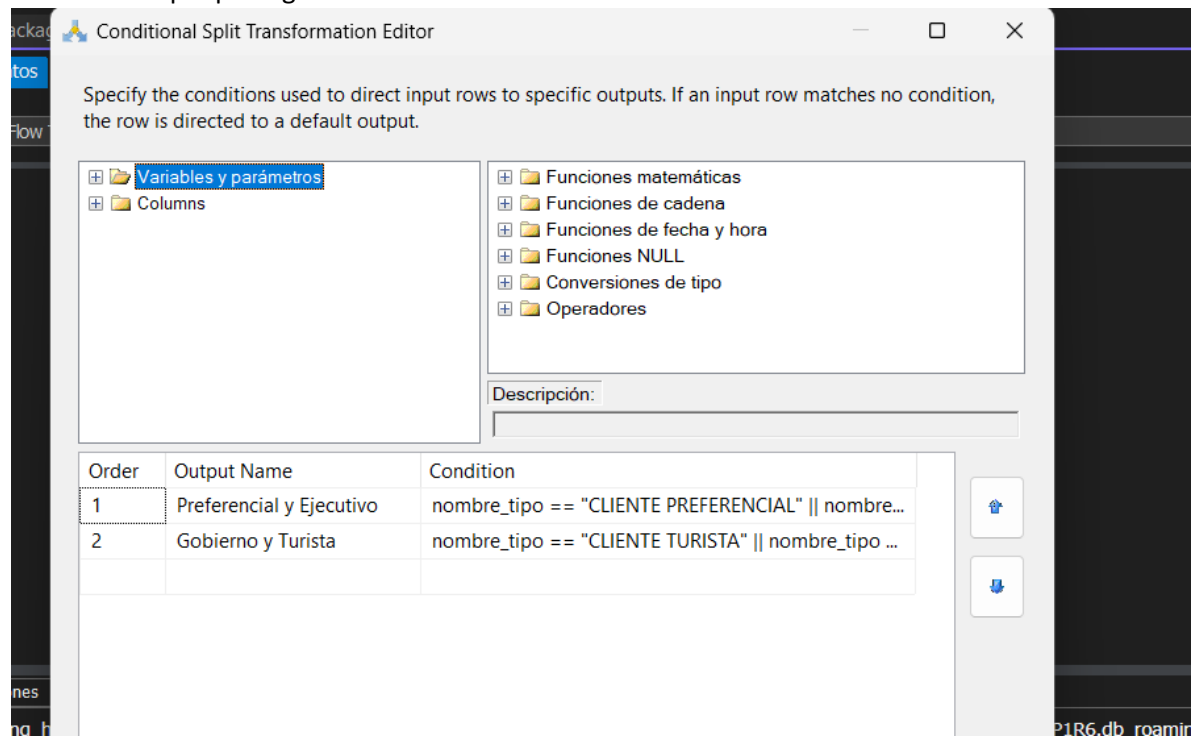
Conversiones de tipo

Operadores

Descripción:

Derived Column Name	Derived Column	Expression	Data Type	
codigo_pais	<add as new column>	LEFT(codigo_cliente,2)	Unicode string [DT_...	2

### 3. Condicional split para generar 2 excel



### 4. Excel generados

Nombre	Fecha de modificación	Tipo	Tamaño
Gobierno y Turista.xls	18/3/2024 21:18	Hoja de cálculo d...	20 KB
Preferencial y Ejecutivo.xls	18/3/2024 21:18	Hoja de cálculo d...	15 KB

a. Ventas por Pais

```
--Pais venta
SELECT
    Departamento,
    SUM(CAST(ProductValue AS INT)) AS TotalSales
FROM
    (SELECT
        'PAIS' AS Departamento,
        'Rosas' AS ProductName, [Rosas] AS ProductValue FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Claveles', [Claveles] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Macetas', [Macetas] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Tierra', [Tierra] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Girasoles', [Girasoles] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Hortensia', [Hortensia] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Globos', [Globos] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Tarjetas', [Tarjetas] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Orquideas', [Orquideas] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Carmesi', [Carmesi] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Lirios', [Lirios] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Aurora', [Aurora] FROM Fiorella UNION ALL
        SELECT 'PAIS' AS Departamento, 'Tulipanes', [Tulipanes] FROM Fiorella
    ) AS SalesByDepartment
GROUP BY Departamento
ORDER BY Departamento, TotalSales DESC;
```

Results		
	Departamento	TotalSales
1	PAIS	10329