

# UNIVERSIDAD DON BOSCO

## DATAWAREHOUSE Y MINERÍA DE DATOS 02-2024



**Actividad: “DESAFIO 2”**

**Docente: Karens Medrano**

**Grupo:01T**

NOMBRE	CARNÉ
Erick Roberto Zavaleta Rivera	ZR171491
Josseline Beatriz Pérez Martínez	PM171434
Erick Iván Peña Rivas	PR170059
Jorge Marvin Peña Roque	PR243380
Hector José Marquez Chicas	MC233291
Ever Félix De León Mendoza	DM191820

**Fecha de entrega: 30 de noviembre del 2024.**

## ETL archivo Excel a SQL

### Configurar el flujo de datos

#### 1. Agregar un paquete de SSIS

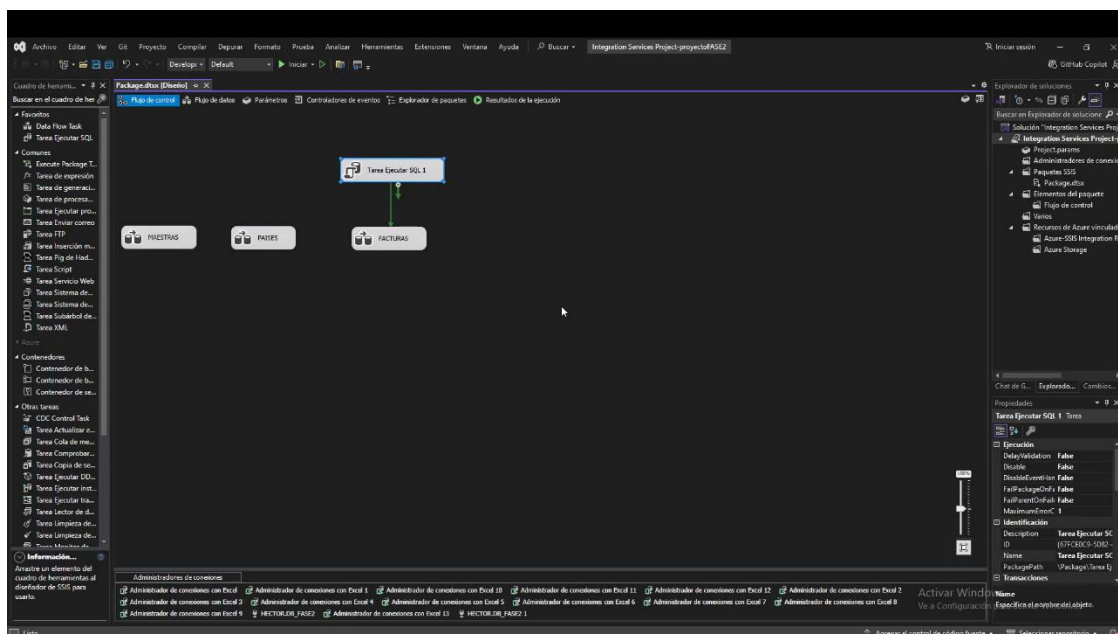
1. En el Explorador de Soluciones, abrir el archivo predeterminado.
2. Cambiar al panel de diseño de control (**Control Flow**).

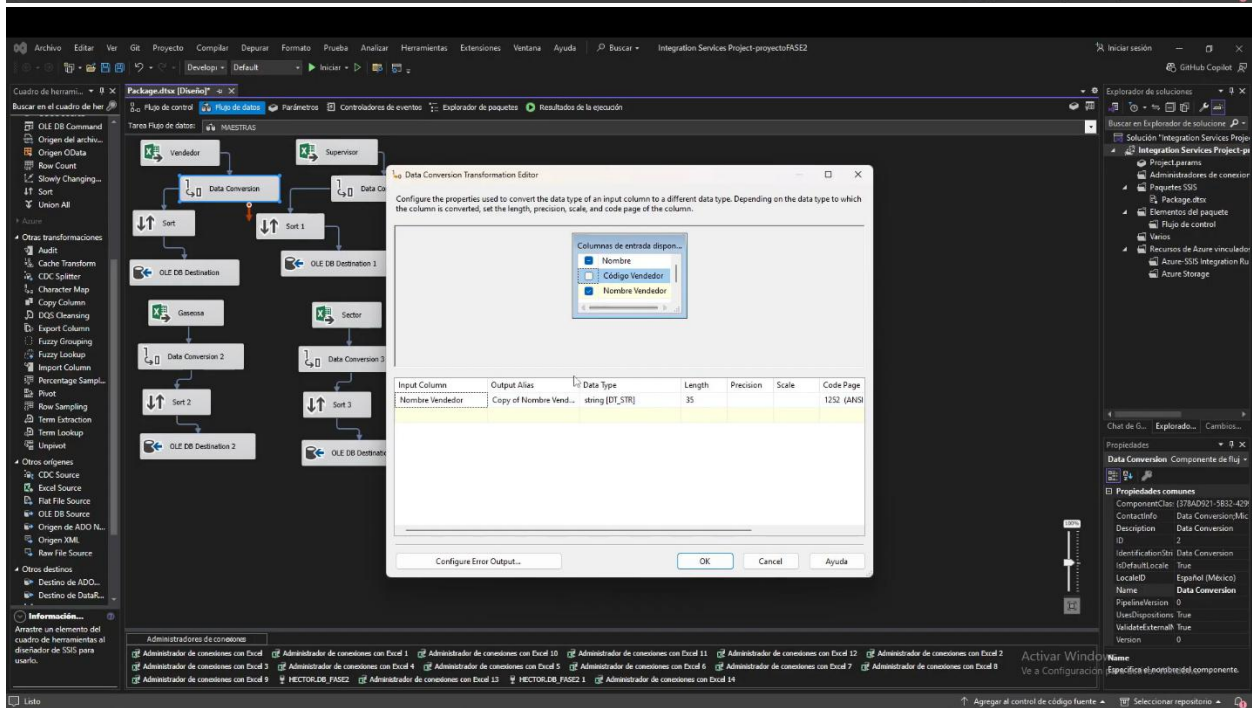
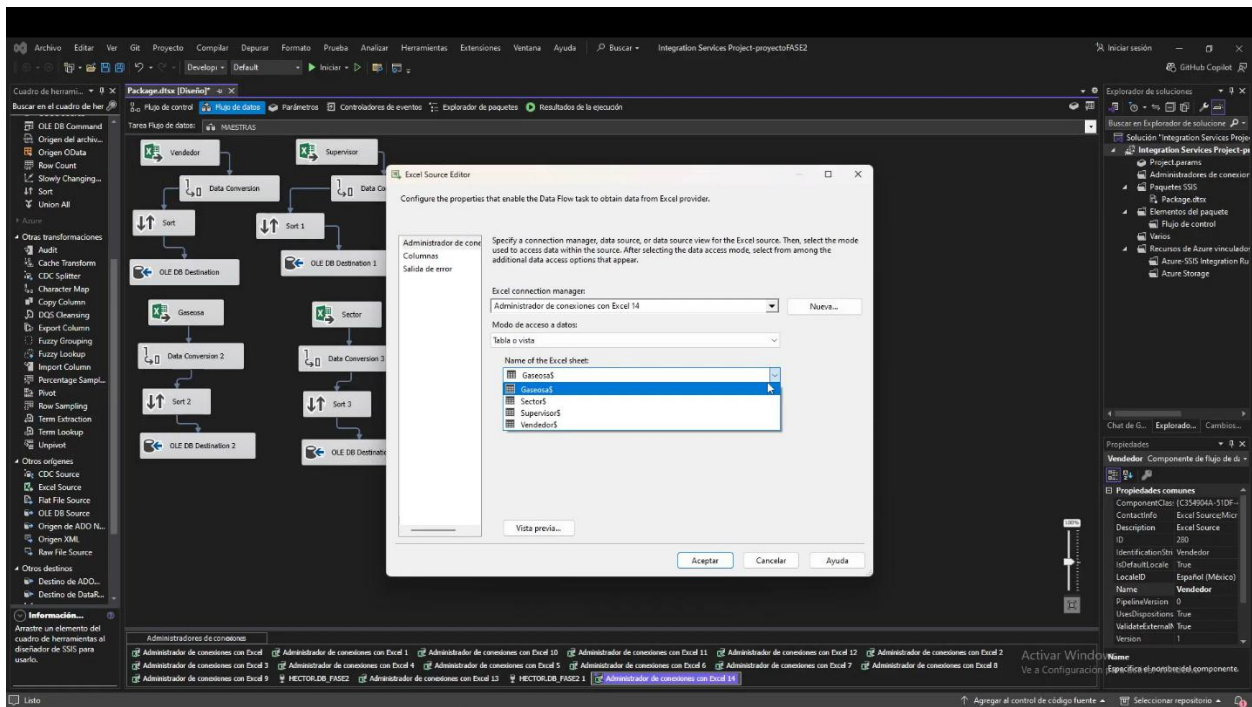
#### 2. Configurar la extracción de datos

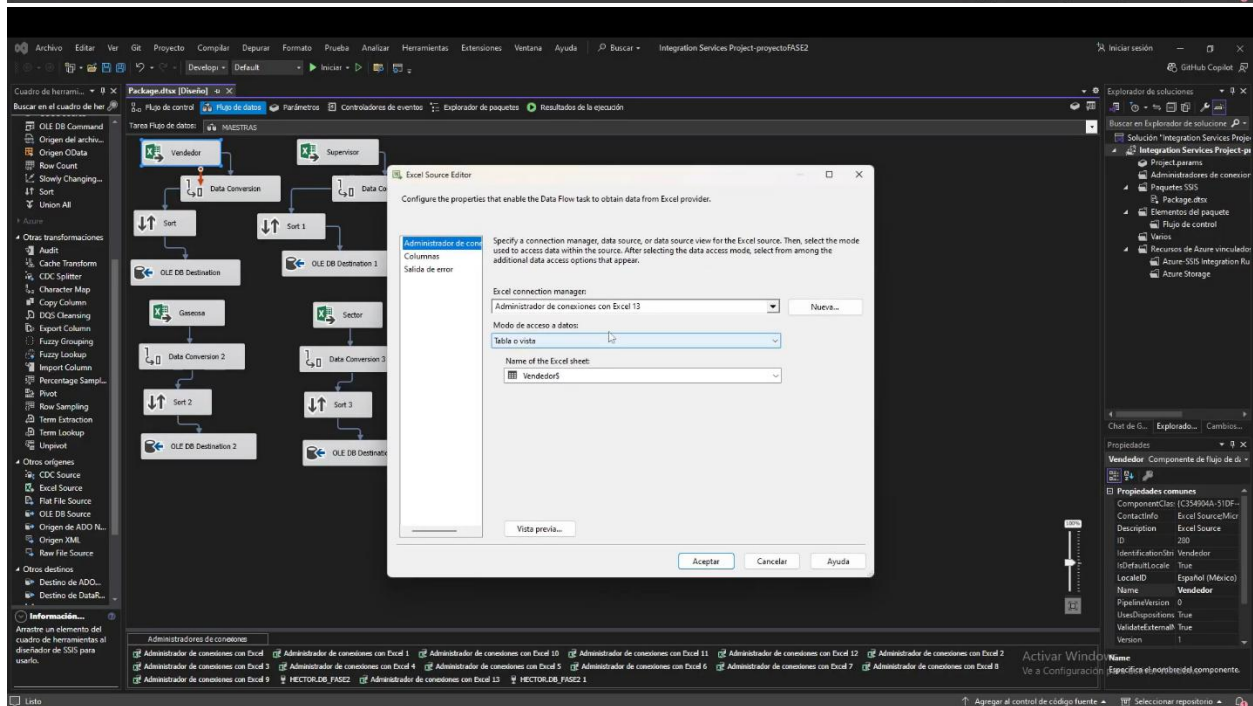
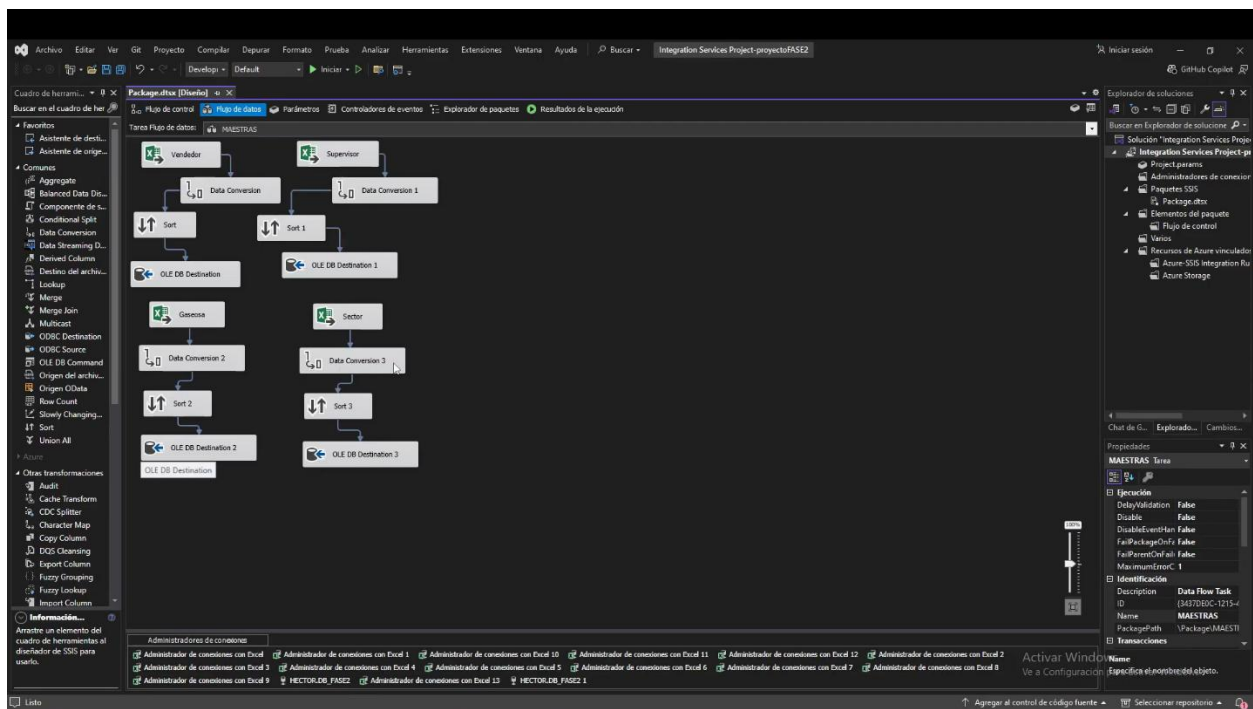
1. Arrastrar una tarea de **Data Flow Task** al lienzo.
2. Doble clic en la tarea para abrir el flujo de datos (**Data Flow**).
3. Agrega un **Source** adecuado:
  - **Archivo CSV o Excel:**
    - Usa un componente **Flat File Source** (para CSV) o **Excel Source**.
    - Configurar la conexión seleccionando el archivo fuente.
    - Verificar que las columnas del archivo se mapeen correctamente.

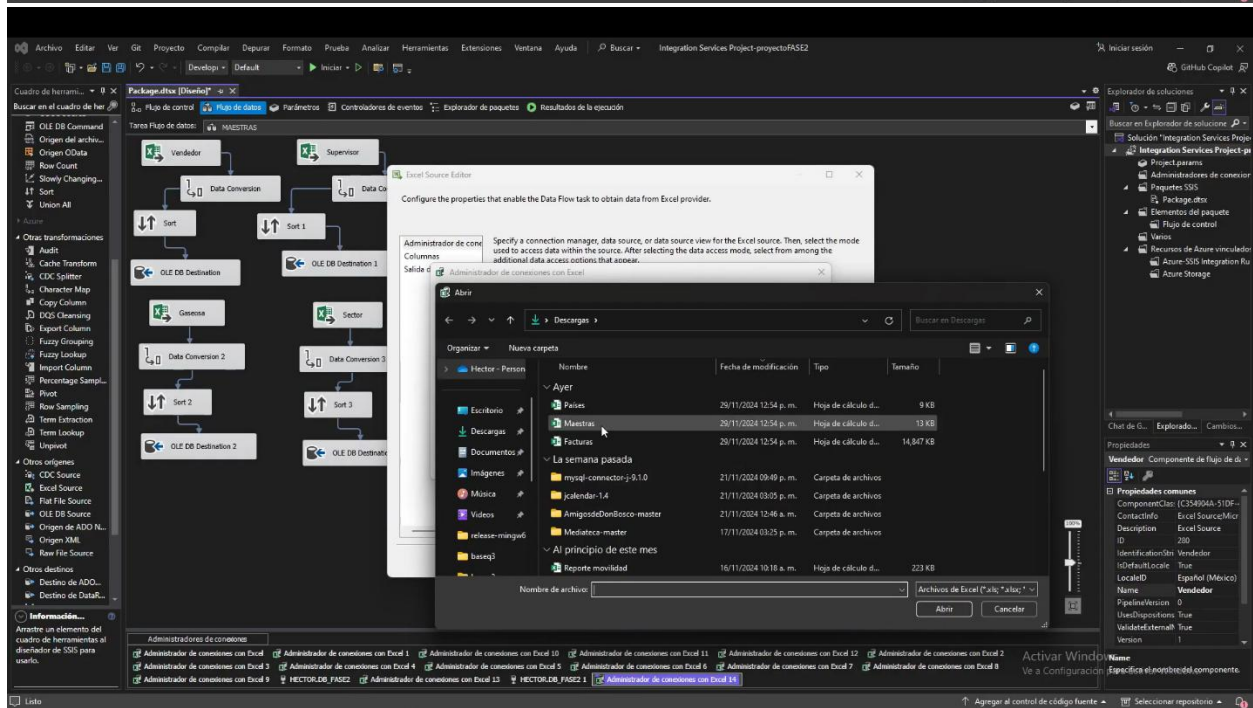
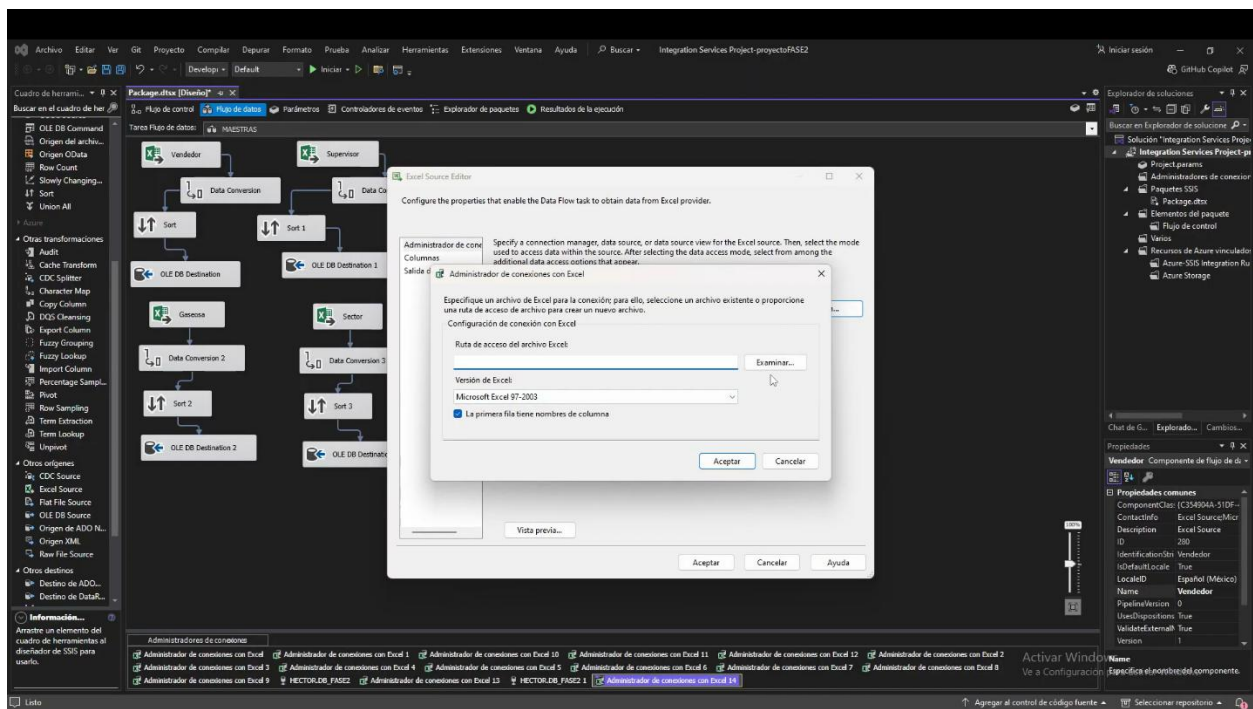
#### 3. Cargar los datos

1. Arrastrar un componente **OLE DB Destination** al lienzo.
2. Conectarlo a la salida del último componente de transformación.
3. Configurar la conexión:
  - Seleccionar el servidor y la base de datos de destino (por ejemplo, ETL\_Database).
  - Mapear las columnas de origen a las de la tabla de destino.

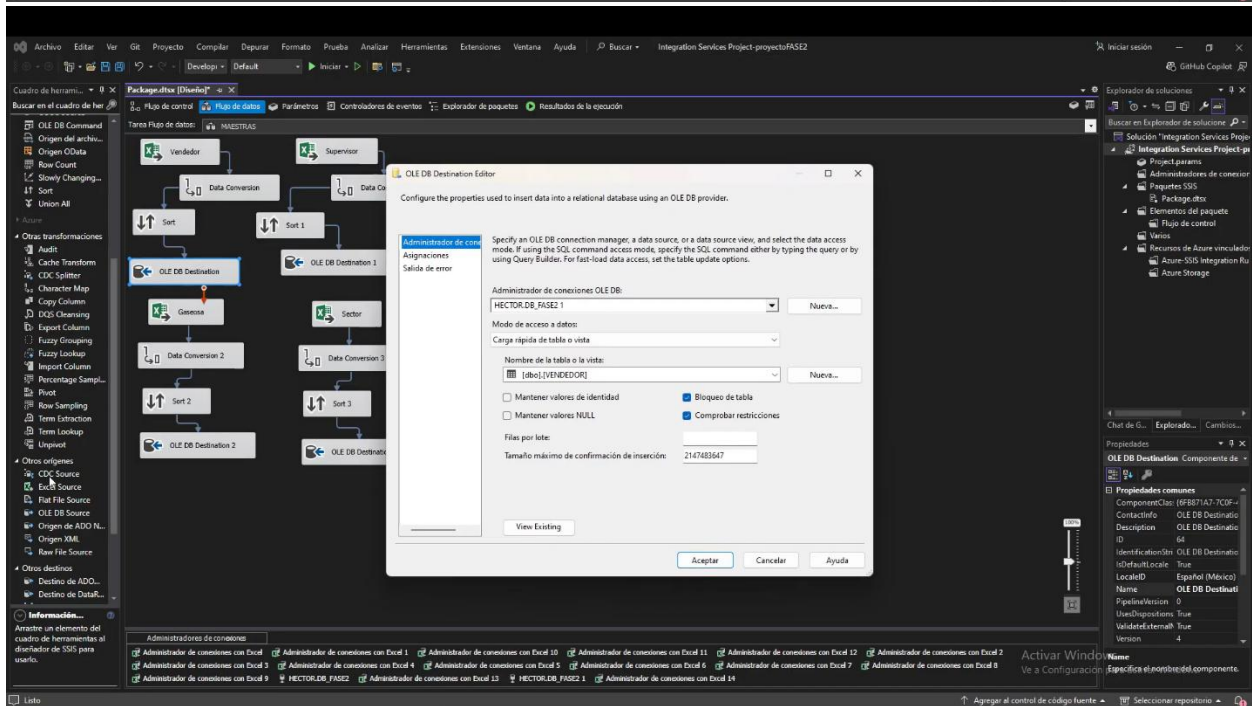
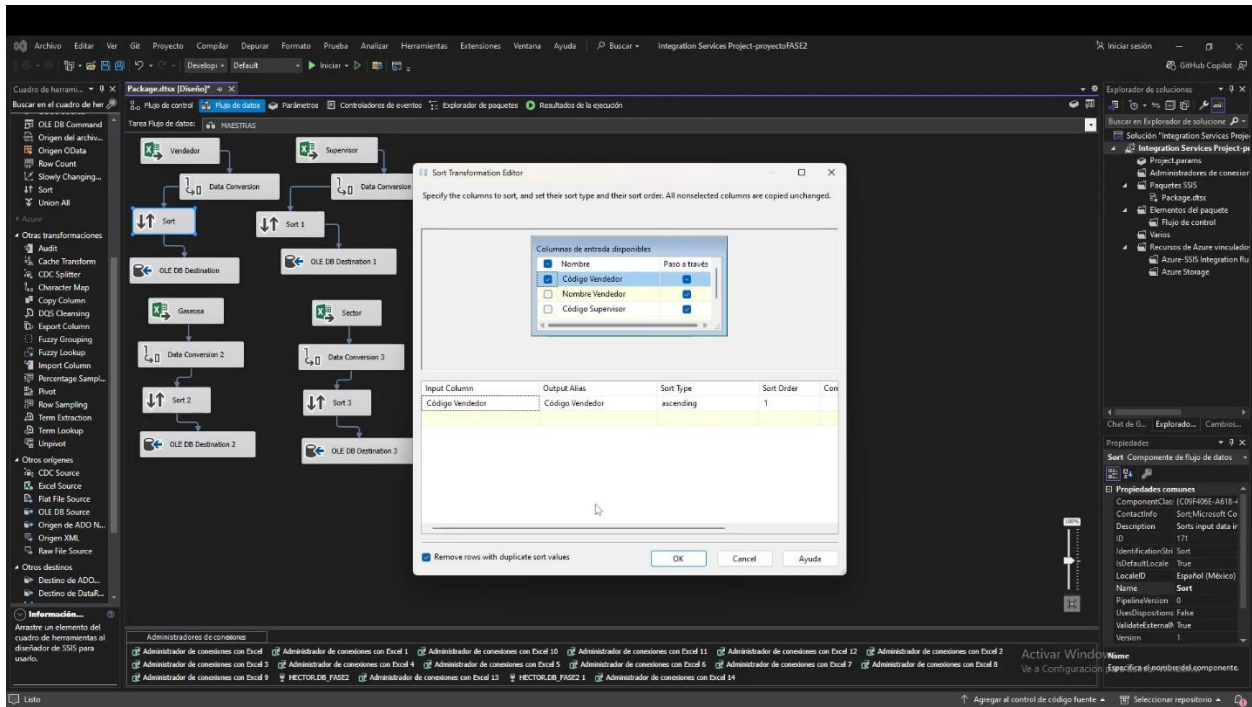




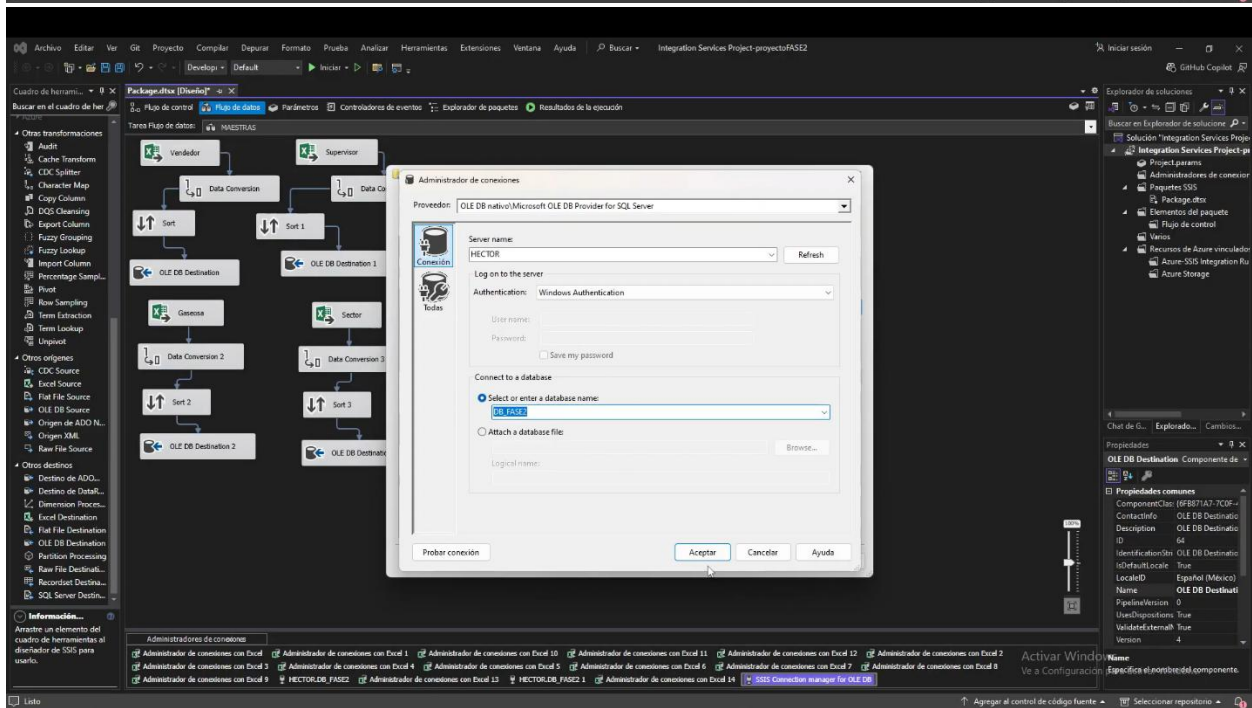
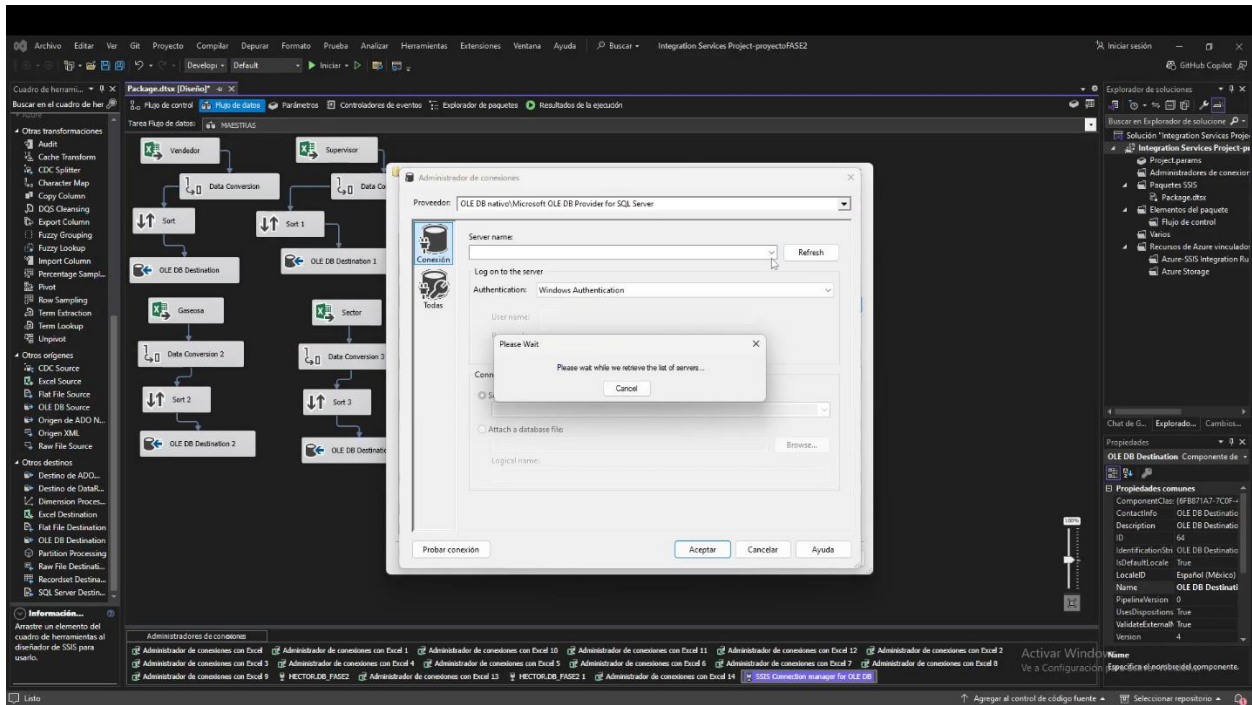




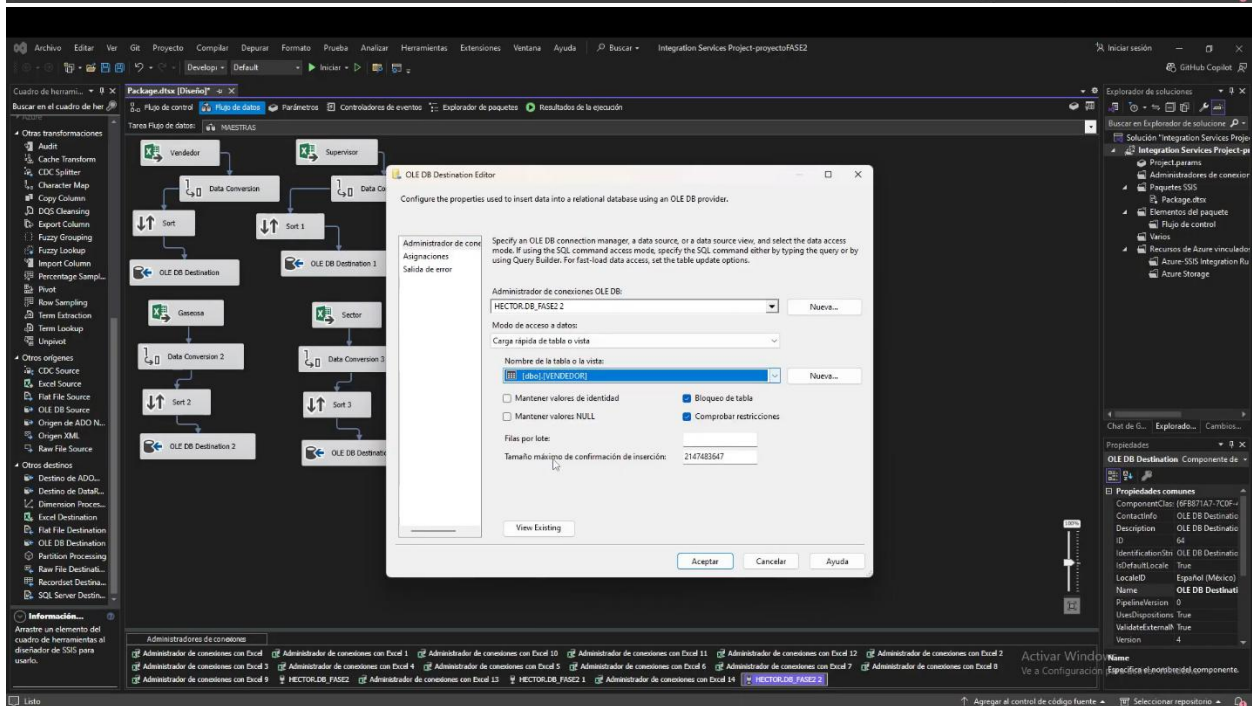
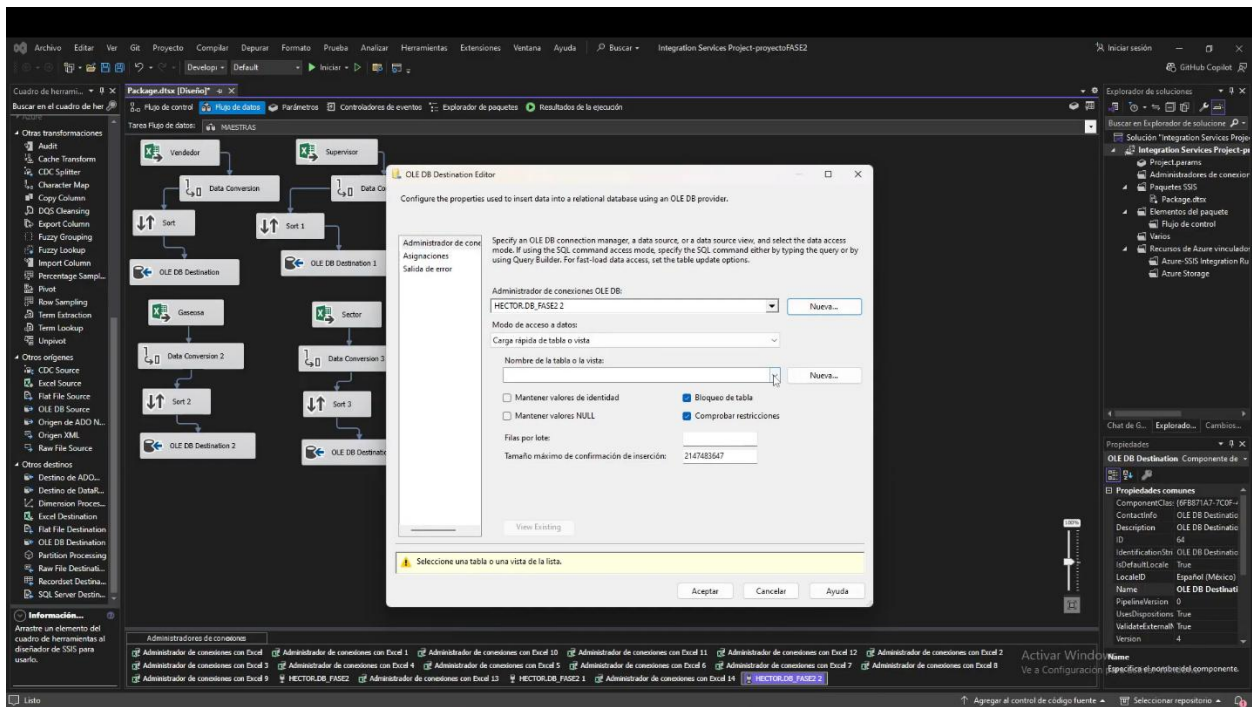


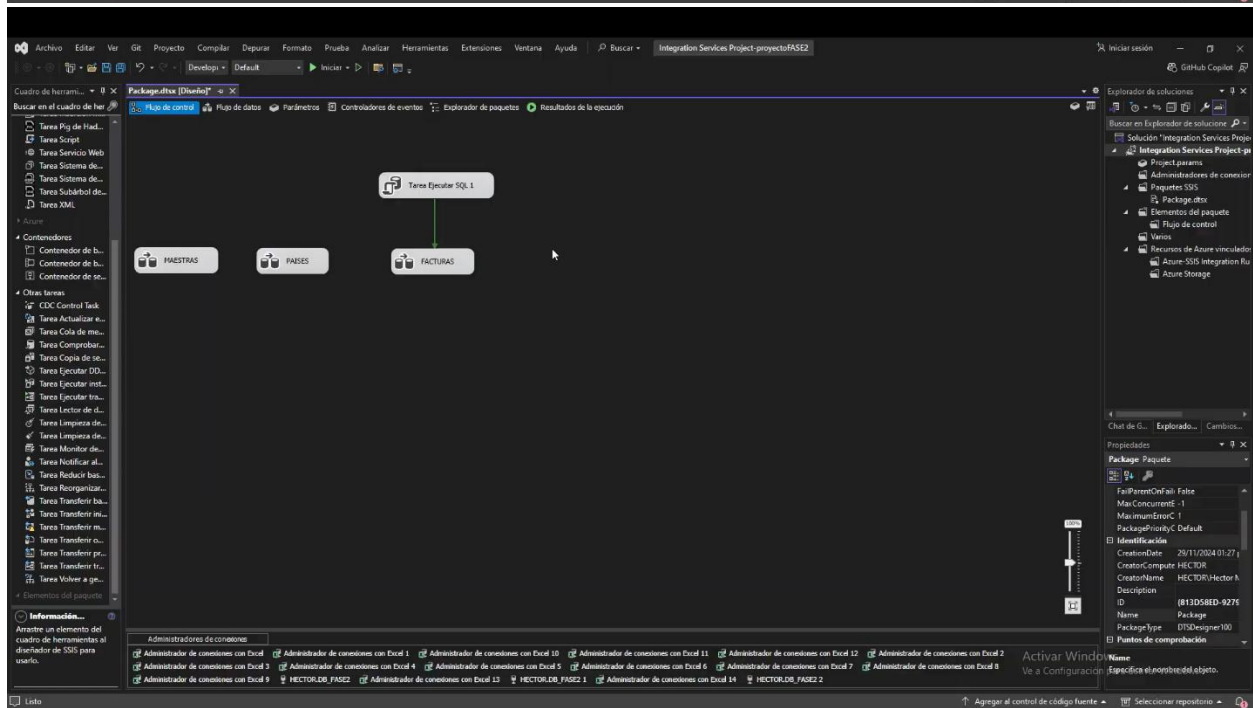
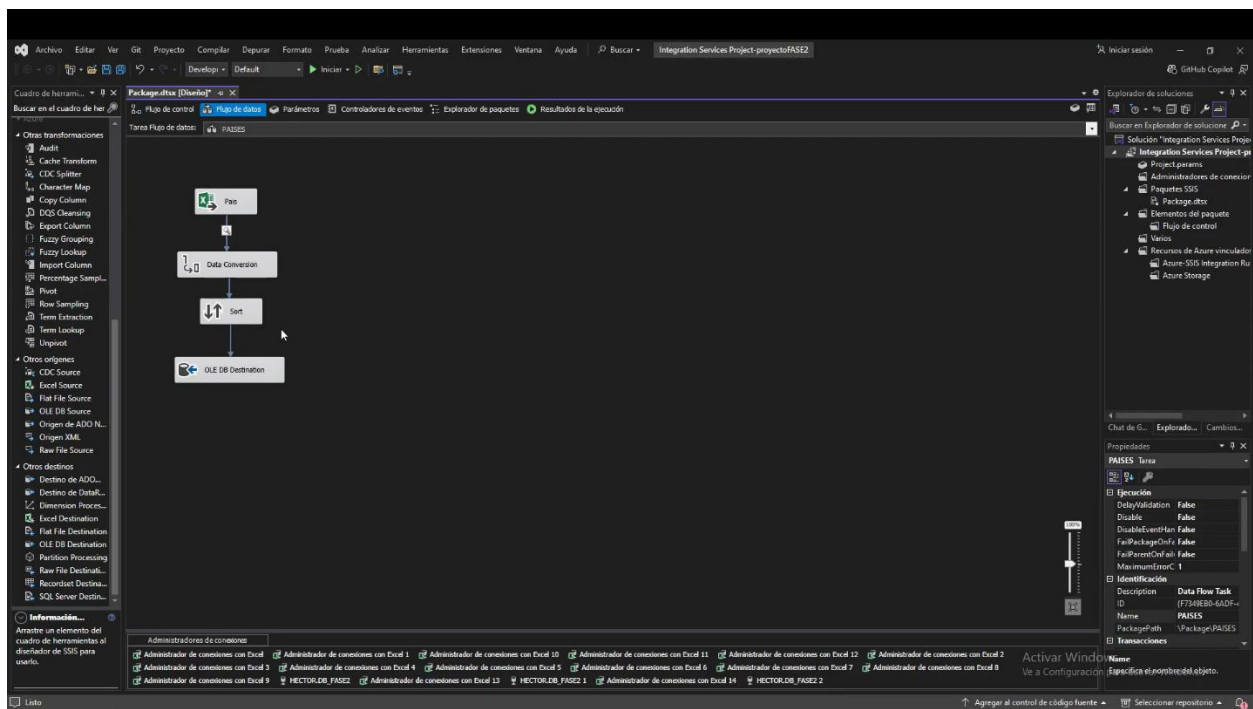


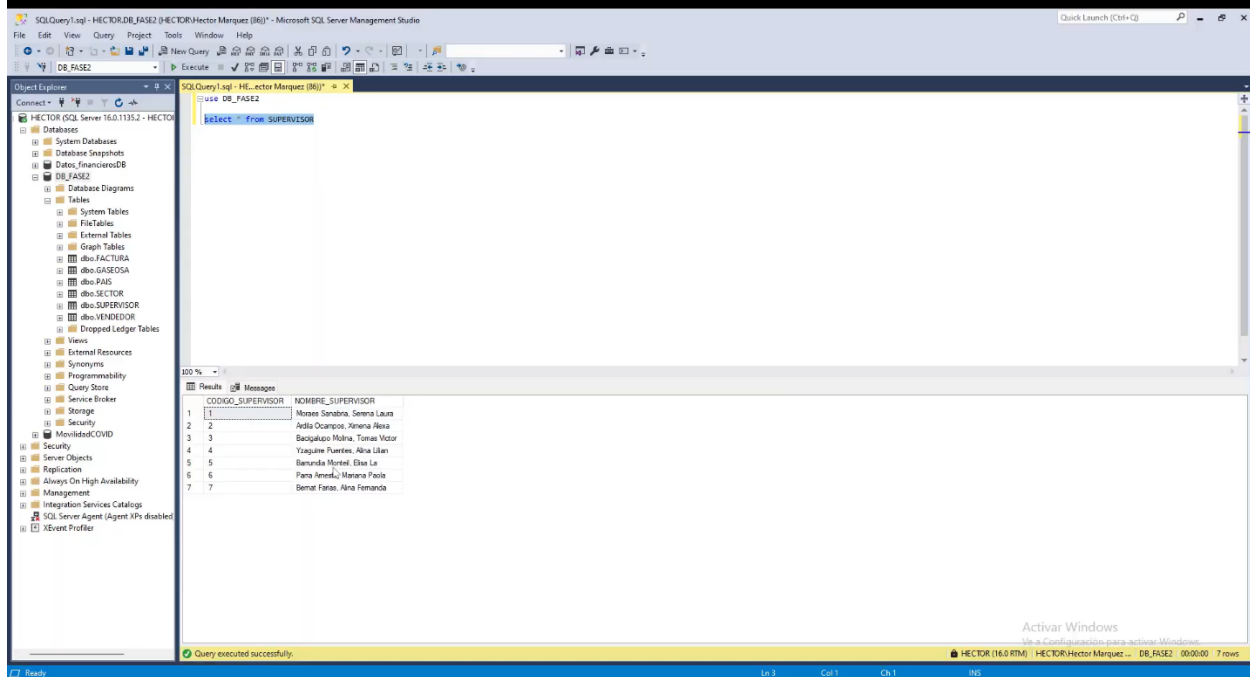
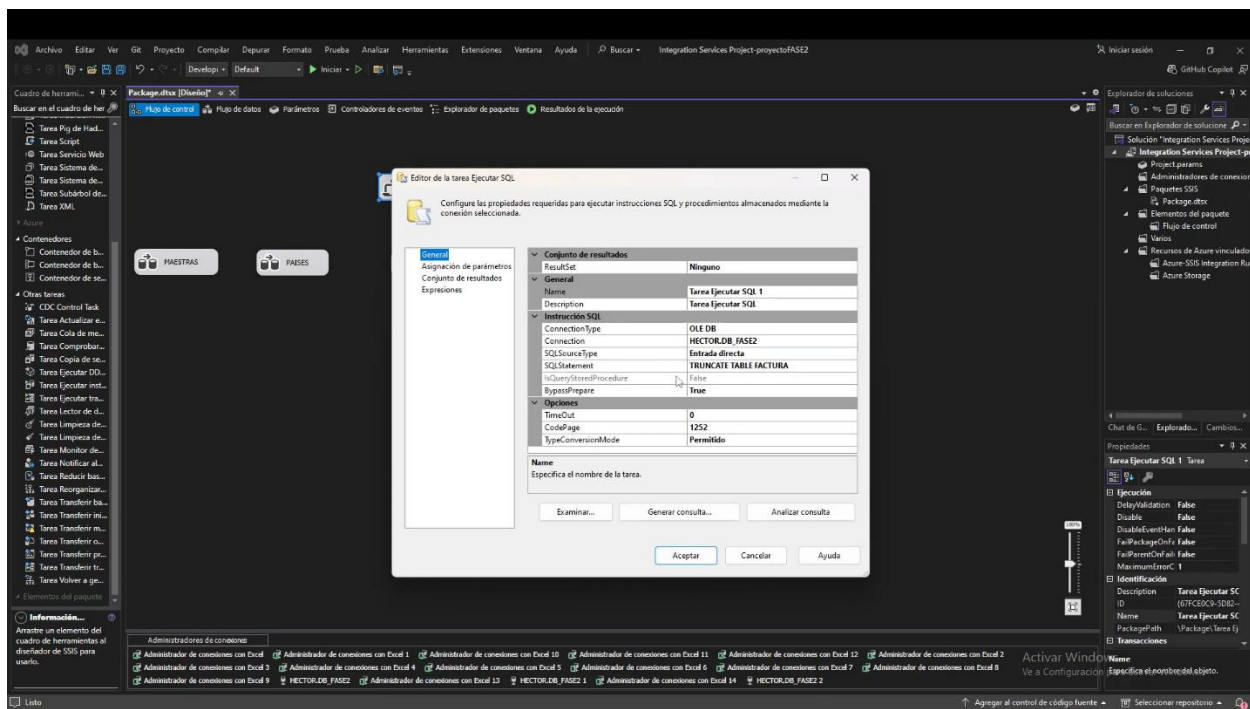












# MODELADO DE BASE DE DATOS

## SCRIPT DE CREACION DE BASE DE DATOS

```
----- GASEOSA
CREATE TABLE GASEOSA(
    CODIGO_GASEOSA INTEGER NOT NULL PRIMARY KEY
    ,NOMBRE_GASEOSA VARCHAR(10) NOT NULL
    ,PRECIO NUMERIC(4,2) NOT NULL
);

SELECT * FROM GASEOSA;

----- SECTOR
CREATE TABLE SECTOR(
    CODIGO_SECTOR INTEGER NOT NULL PRIMARY KEY
    ,DESCRIPCION_SECTOR VARCHAR(10) NOT NULL
);

SELECT * FROM SECTOR;

----- SUPERVISOR
CREATE TABLE SUPERVISOR(
    CODIGO_SUPERVISOR INTEGER NOT NULL PRIMARY KEY
    ,NOMBRE_SUPERVISOR VARCHAR(31) NOT NULL
);

SELECT * FROM SUPERVISOR;

----- VENDEDOR
CREATE TABLE VENDEDOR(
    CODIGO_VENDEDOR INTEGER NOT NULL PRIMARY KEY
    ,NOMBRE_VENDEDOR VARCHAR(35) NOT NULL
    ,CODIGO_SUPERVISOR INTEGER NOT NULL
);

----- PAIS
CREATE TABLE PAIS(
    CODIGO_PAIS INTEGER NOT NULL PRIMARY KEY
    ,NOMBRE VARCHAR(9) NOT NULL
);

----- FACTURA
CREATE TABLE [dbo].[FACTURA](
    [ID] [int] IDENTITY(1,1) NOT NULL,
    [SERIE] [int] NOT NULL,
    [NUMERO] [int] NOT NULL,
    [FECHA] [date] NOT NULL,
    [CODIGO_SECTOR] [int] NOT NULL,
```



```

        [CODIGO_PAIS] [int] NOT NULL,
        [CODIGO_VENDEDOR] [int] NOT NULL,
        [CODIGO_GASEOSA] [int] NOT NULL,
        [PPTO] [numeric](14, 7) NOT NULL,
        [CANTIDAD] [numeric](8, 1) NOT NULL,
        [MERMA] [int] NOT NULL,
        [SUBTOTAL] [numeric](16, 9) NOT NULL,
        [IGV] [numeric](13, 7) NOT NULL,
        [TOTAL] [numeric](13, 6) NOT NULL,
PRIMARY KEY CLUSTERED
(
    [ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO

```

```

DBCC CHECKIDENT ('FACTURA', RESEED, 0);
TRUNCATE TABLE FACTURA;
SELECT * FROM FACTURA;

```

--- Creacion de relaciones

```

ALTER TABLE FACTURA
ADD CONSTRAINT FK_FACTURA_GASEOSA
FOREIGN KEY (CODIGO_GASEOSA)
REFERENCES GASEOSA(CODIGO_GASEOSA);

```

```

ALTER TABLE FACTURA
ADD CONSTRAINT FK_FACTURA_SECTOR
FOREIGN KEY (CODIGO_SECTOR)
REFERENCES SECTOR(CODIGO_SECTOR);

```

```

ALTER TABLE FACTURA
ADD CONSTRAINT FK_FACTURA_PAIS
FOREIGN KEY (CODIGO_PAIS)
REFERENCES PAIS(CODIGO_PAIS);

```

```

ALTER TABLE FACTURA
ADD CONSTRAINT FK_FACTURA_VENDEDOR
FOREIGN KEY (CODIGO_VENDEDOR)
REFERENCES VENDEDOR(CODIGO_VENDEDOR);

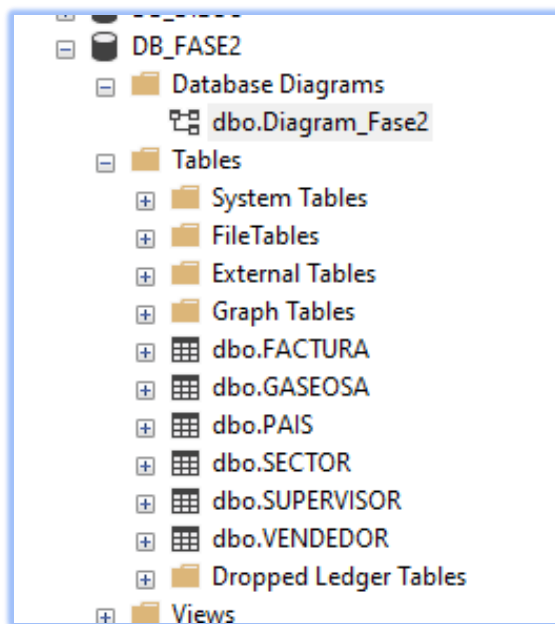
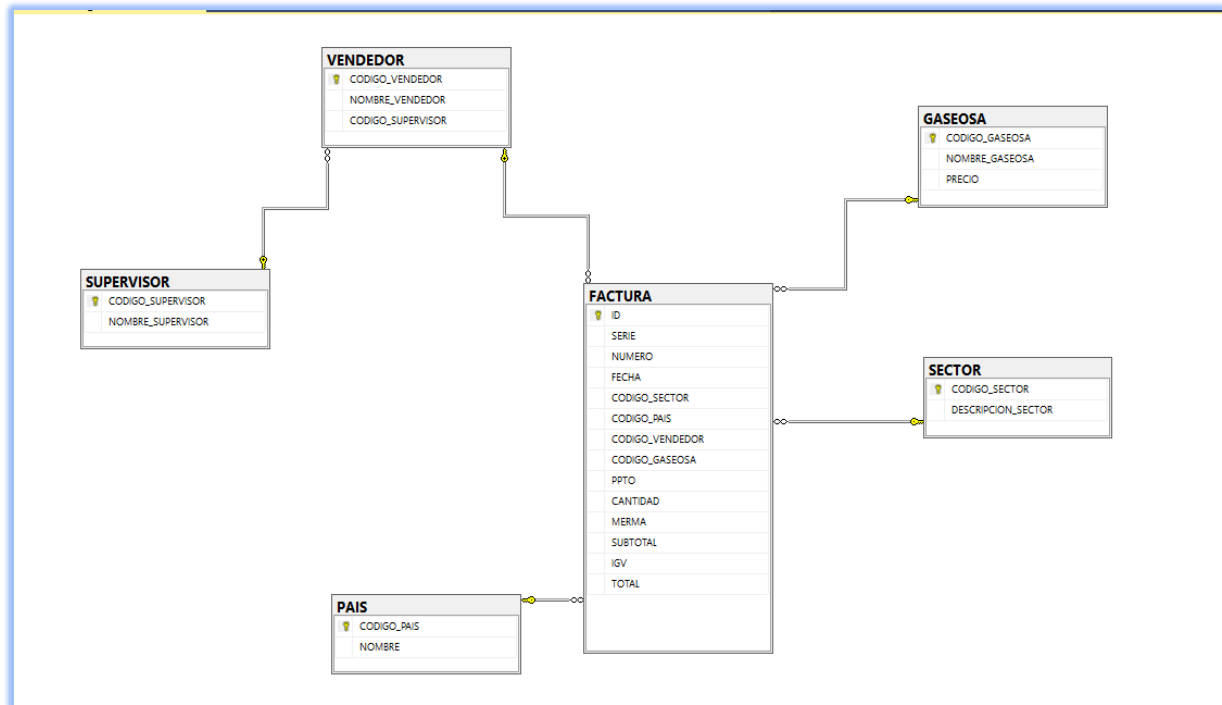
```

```

ALTER TABLE VENDEDOR
ADD CONSTRAINT FK_VENDEDOR_SUPERVISOR
FOREIGN KEY (CODIGO_SUPERVISOR)
REFERENCES SUPERVISOR(CODIGO_SUPERVISOR);

```

## DIAGRAMA DE BASE DE DATOS RELACIONAL



## KPI

### 1. Ventas

- **Ventas Totales:** Suma del campo TOTAL de la tabla FACTURA.
- **Promedio de Venta por Transacción:** Promedio del campo TOTAL por NUMERO de factura.
- **Cantidad Vendida:** Suma del campo CANTIDAD de la tabla FACTURA.
- **Merma Total:** Suma del campo MERMA.

### 2. Supervisión de Vendedores

- **Ventas por Vendedor:** Total vendido agrupado por CODIGO\_VENDEDOR y su respectivo supervisor (CODIGO\_SUPERVISOR).
- **Rendimiento por Supervisor:** Total vendido por cada supervisor (CODIGO\_SUPERVISOR).

### 3. Productos (Gaseosa)

- **Producto más Vendido:** Producto con la mayor suma en CANTIDAD, agrupado por CODIGO\_GASEOSA.
- **Margen por Producto:** (Suma de TOTAL - Costo de los productos vendidos, si esta información está disponible).

### 4. Geografía

- **Ventas por País:** Total vendido agrupado por CODIGO\_PAIS.
- **Ventas por Sector:** Total vendido agrupado por CODIGO\_SECTOR.

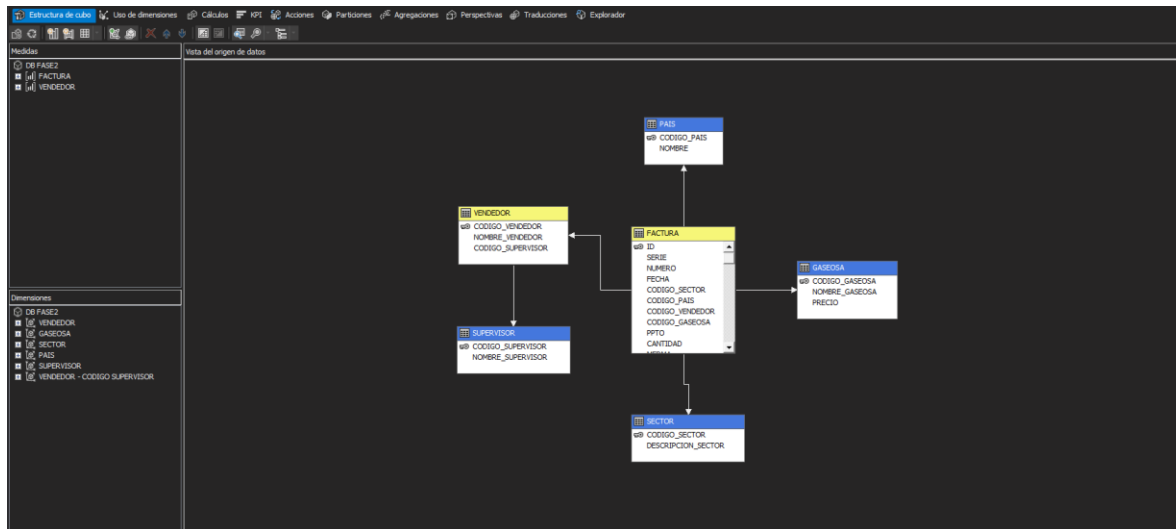
### 5. Cumplimiento de Presupuesto

- **Porcentaje de Cumplimiento:**  $(\text{Suma de TOTAL} / \text{Suma de PPTO}) * 100$ .
- **Desviación de Presupuesto:**  $\text{PPTO} - \text{TOTAL}$ .

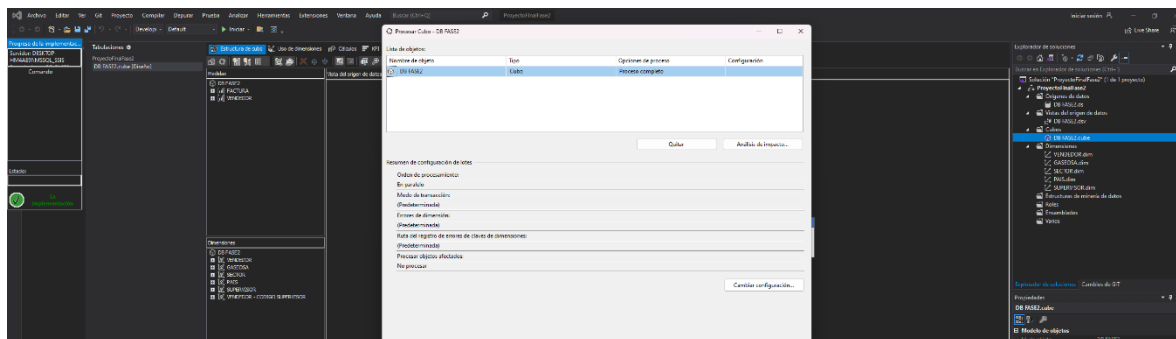
### 6. Impuestos

- **IGV Total:** Suma del campo IGV.
- **Porcentaje de IGV sobre Ventas Totales:**  $(\text{Suma de IGV} / \text{Suma de TOTAL}) * 100$ .

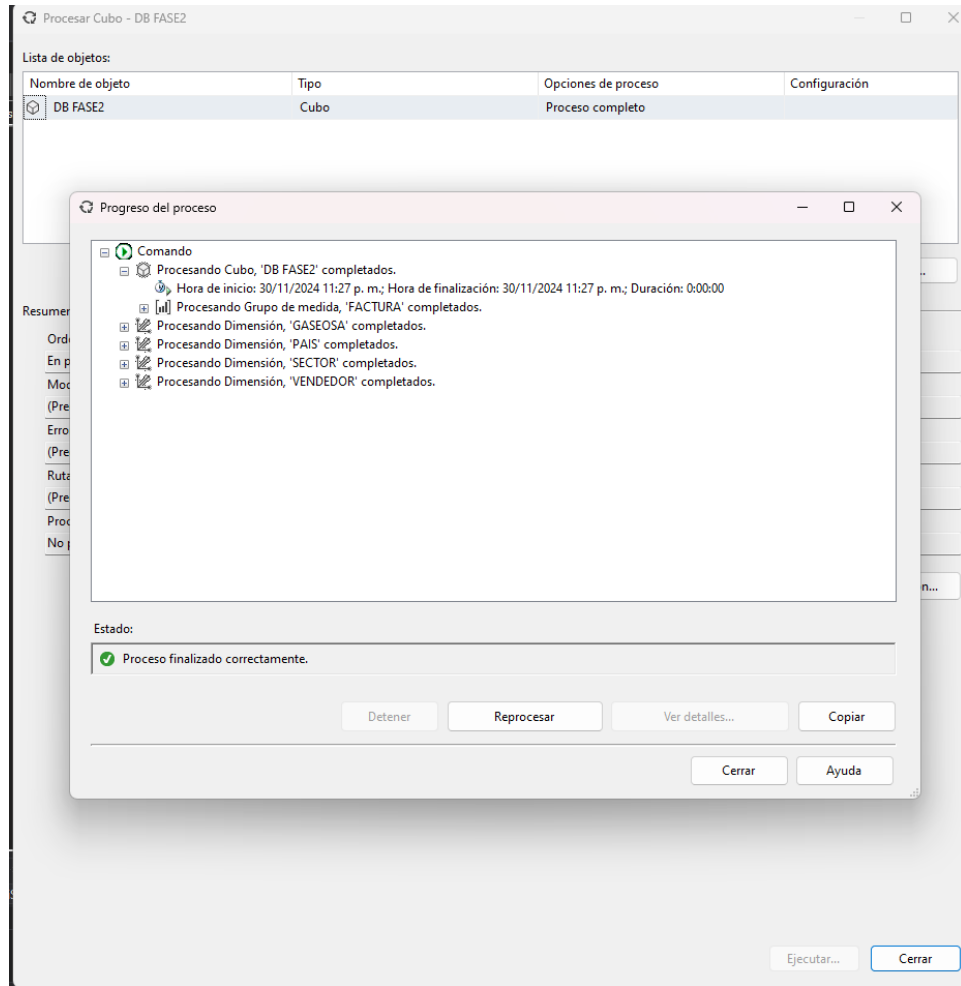
## CUBO OLAP



## PROCESAMIENTO DE CUBO

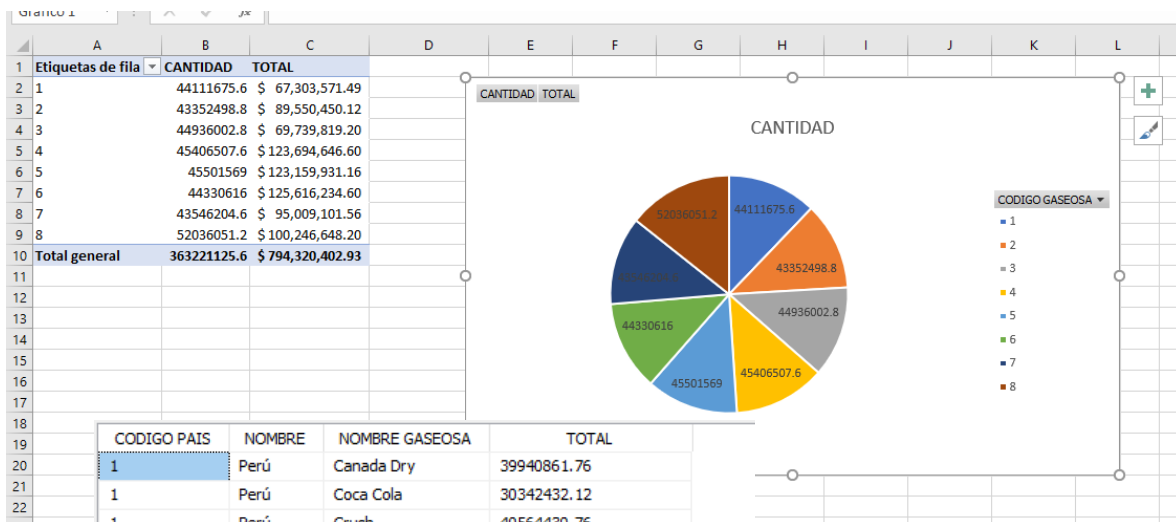






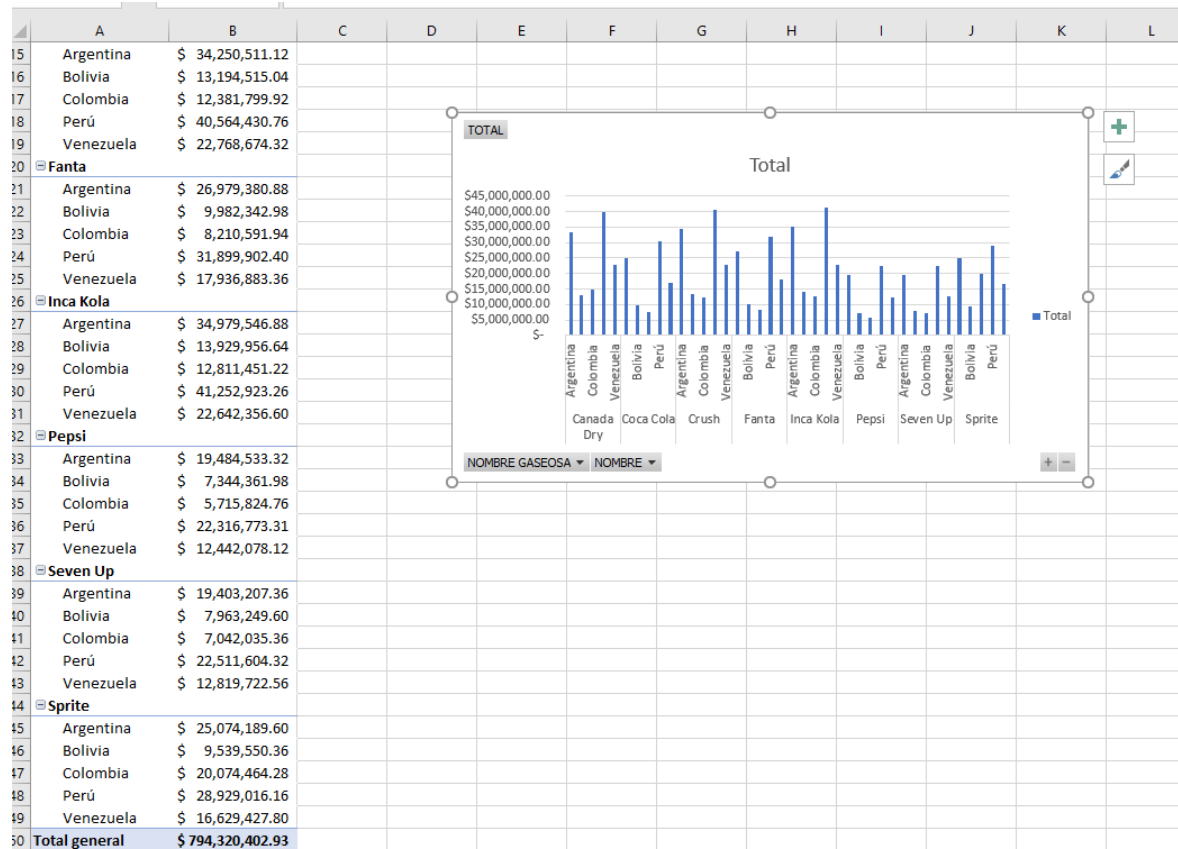
Total vendido por código de gaseosas:

CODIGO GASEOSA	CANTIDAD	SERIE	Recuento FACTURA	TOTAL
1	44111675.6	4418...	20409	67303571.48...
2	43352498.8	4350...	19994	89550450.11...
3	44936002.8	4408...	20331	69739819.19...
4	45406507.6	4383...	20204	123694646.6
5	45501569	4460...	20559	123159931.16
6	44330616	4348...	20001	125616234.6
7	43546204.6	4365...	20110	95009101.56...
8	52036051.2	4424...	20475	100246648.2



Totals  
in  
sales  
of  
sodas  
by  
countries

CODIGO PAIS	NOMBRE	NOMBRE GASEOSA	TOTAL
1	Perú	Canada Dry	39940861.76
1	Perú	Coca Cola	30342432.12
1	Perú	Crush	40564430.76
1	Perú	Fanta	31899902.4
1	Perú	Inca Kola	41252923.26
1	Perú	Pepsi	22316773.308981
1	Perú	Seven Up	22511604.32
1	Perú	Sprite	28929016.16
2	Argentina	Canada Dry	33219568.58
2	Argentina	Coca Cola	24971555.12
2	Argentina	Crush	34250511.12
2	Argentina	Fanta	26979380.88
2	Argentina	Inca Kola	34979546.88
2	Argentina	Pepsi	19484533.32
2	Argentina	Seven Up	19403207.36
2	Argentina	Sprite	25074189.6
3	Venezuela	Canada Dry	22791869.66
3	Venezuela	Coca Cola	16897102.04
3	Venezuela	Crush	22768674.32
3	Venezuela	Fanta	17936883.36
3	Venezuela	Inca Kola	22642356.6
3	Venezuela	Pepsi	12442078.12
3	Venezuela	Seven Up	12819722.56
3	Venezuela	Sprite	16629427.8
5	Bolivia	Canada Dry	12976941.98
5	Bolivia	Coca Cola	9745429.4
5	Bolivia	Crush	13194515.04
5	Bolivia	Fanta	9982342.98



[https://github.com/JorgeM21PR24/PROYECTO\\_FINAL\\_FASE2.git](https://github.com/JorgeM21PR24/PROYECTO_FINAL_FASE2.git)

<https://drive.google.com/drive/folders/1b3i6K8USauNIVul4Ni9lIHxjQqhhZf?usp=sharing>