



**UNIVERSIDAD
DON BOSCO**

**Datawarehouse y Minería de Datos
DESAFÍO 1
ELABORADO POR:
Christian G. Crespin CL060107**

Docente:

Ing. Karen Medrano

Facultad de Ingeniería - UDB Virtual

Soyapango, Marzo 2025

INDICE

Introducción.....	1
Creación del programa y repositorio GIT.....	1
Creación de la base de datos, data Flow.....	3
Manejar posibles errores.....	11
Comprobando el funcionamiento.....	15
Obtener la data para el análisis de la campaña.....	16
Analisis de los datos.....	22
Conclusiones.....	27

Introducción

En un mercado cada vez más competitivo, la fidelización de clientes se ha convertido en una estrategia clave para garantizar el crecimiento y sostenibilidad de las empresas. En este contexto, el Spa "Diego" busca implementar una campaña de fidelización basada en la segmentación de sus clientes, con el fin de optimizar la oferta de servicios y mejorar la retención de usuarios.

Para ello, se ha llevado a cabo un análisis detallado de los datos de las tres sucursales del spa, utilizando herramientas de procesamiento y minería de datos. Este estudio permite identificar patrones de comportamiento en los clientes, diferenciándolos en función de su frecuencia de visitas, género, edad y preferencias de servicio.

A través del uso de SQL Server, SSIS y consultas analíticas, se han segmentado los clientes en dos grupos principales:

- Frecuentes: Aquellos que visitan el spa con una frecuencia igual o superior al promedio.
- Ocasionales: Aquellos que visitan el spa con una frecuencia inferior al promedio.

Además, se ha analizado la demanda de servicios como Masajes, Sauna, Hidroterapia y Yoga, desglosando el comportamiento de los clientes por género y rango de edad. También se ha comparado el desempeño de cada sucursal, identificando oportunidades para mejorar la oferta de servicios y aumentar la rentabilidad del negocio.

Este documento detalla el proceso de extracción y transformación de datos, la metodología empleada, los hallazgos obtenidos y las recomendaciones estratégicas basadas en el análisis. Con esta información, el Spa "Diego" podrá diseñar campañas personalizadas que impulsen la fidelización de clientes y aumenten la rentabilidad del negocio.

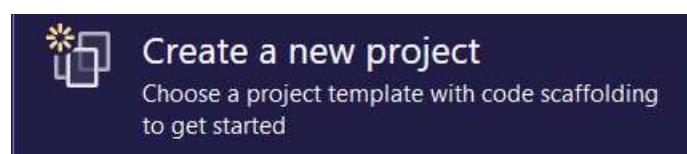
Creación del programa y repositorio GIT

El enlace público a esta solución creada es el siguiente:

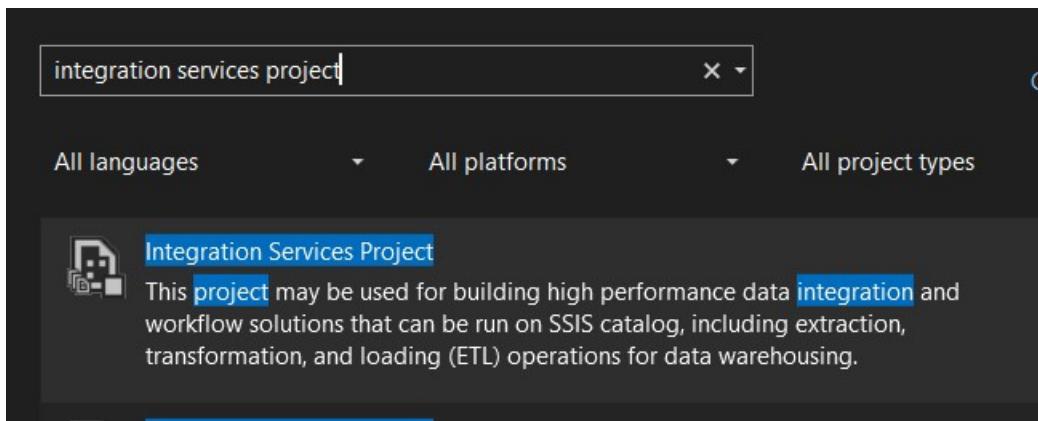
https://github.com/ChrisWooks/Desafio1_DMD941_CL060107.git

Paso a paso de la creación

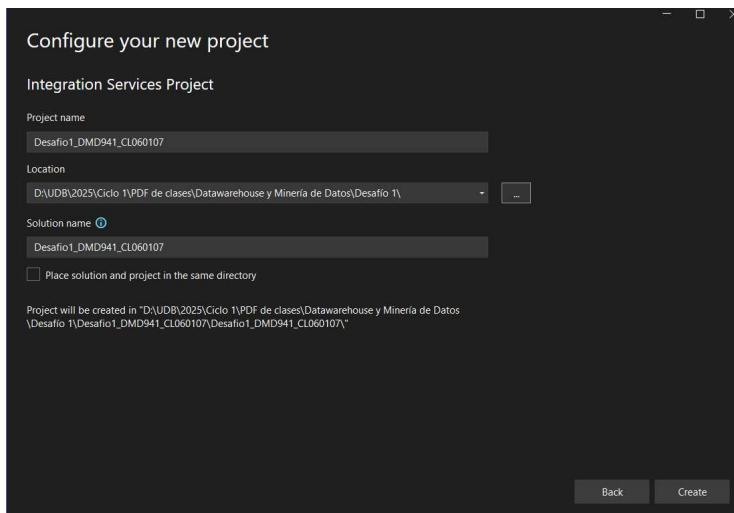
- Creo el proyecto en Visual Studio 2022
 - Creé el proyecto nuevo



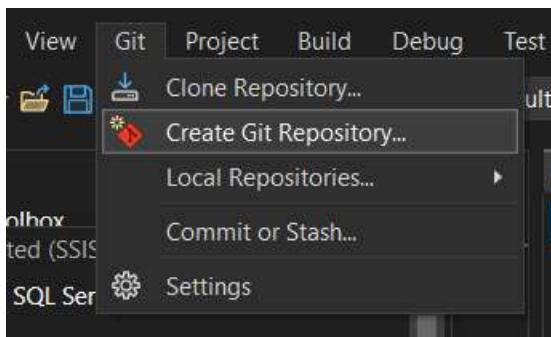
- Utilicé "integration services project"



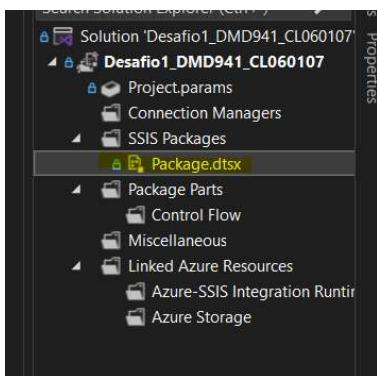
- Agregué el nombre y el directorio a guardar



- Creé el repositorio GIT, utilizando Visual Studio

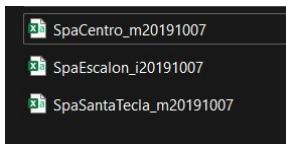


- Utilicé mi cuenta de GitHub para crear el repositorio, de igual manera le puse una descripción al repositorio.
- Después de confirmar que el repositorio fue creado correctamente, borré Package.dtsx, como vimos en la documentación de la clase

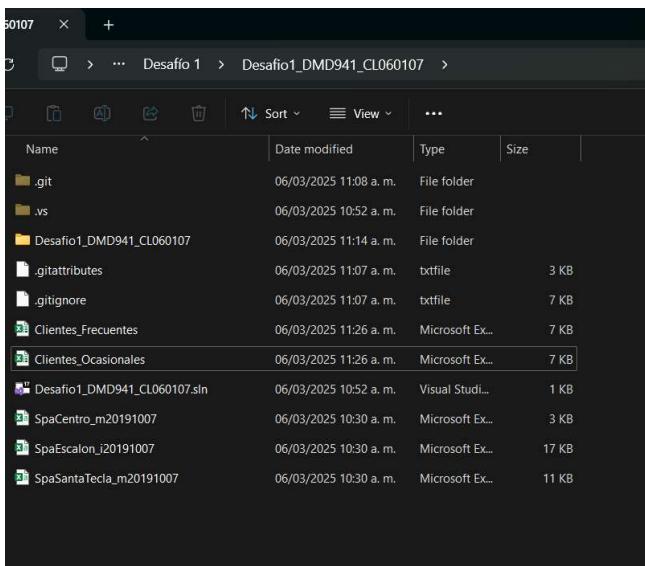


Creación de la base de datos, data Flow

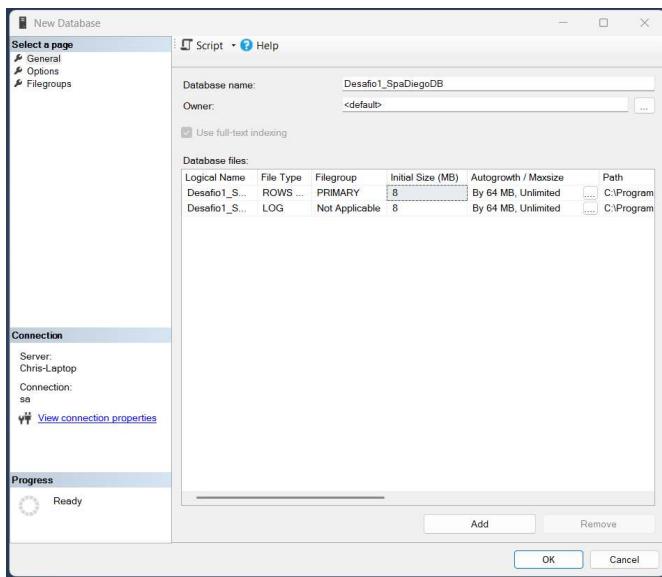
- Los datos ya los tenemos en los tres archivos Excel que nos dieron en la clase.



- Aproveché y los coloqué en el repositorio. De igual manera he creado dos Excel, Clientes_Frecuentes y Clientes_Ocasionales. Estos los utilizaré más adelante.

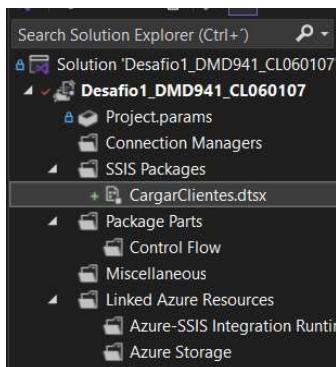


- Abrí Microsoft SQL Server Management Studio y cree la base de datos (vacía por el momento).

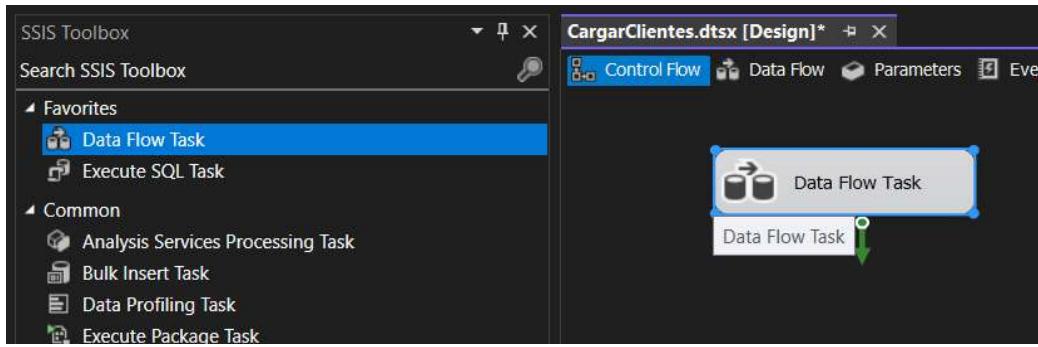


- Antes de hacer algún cambio, me aseguré de que los tres .CSV tienen la misma cantidad de columnas y que estas tienen el mismo nombre.

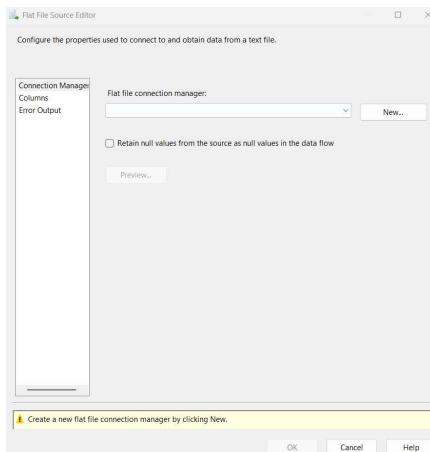
- Creé un paquete SSIS en blanco y lo nombré CargarClientes.dtsx



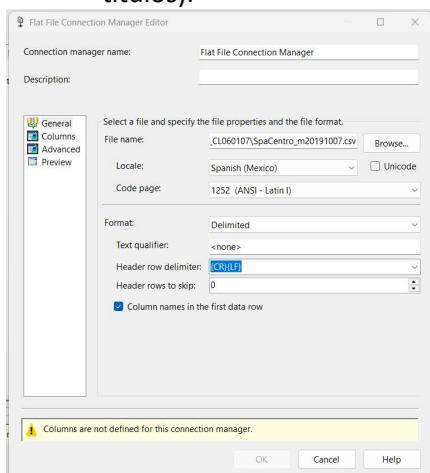
- Agregué una "Data Flow Task" al área de diseño del "Control Flow".



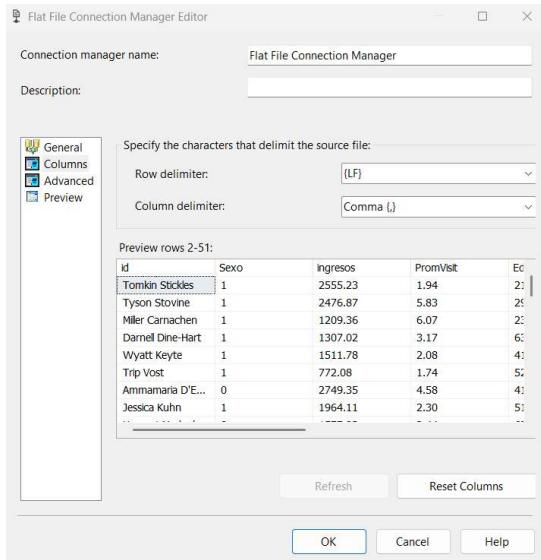
- Añadí un Flat File Source.
 - Hice doble clic en el Data Flow Task.
 - Agregué un "Flat File Source" al área de trabajo del data flow y le doy doble clic.
 - Luego le presioné "New"



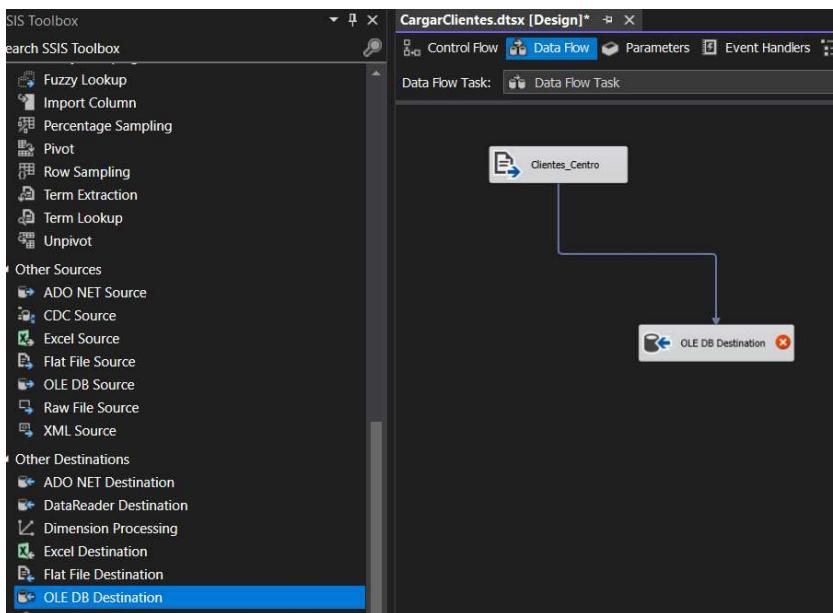
- Seleccioné el archivo SpaCentro_mm20191007.csv, dejé formato "Delimited" y me aseguré de que la opción "Column names in the first data row" estuviera marcada (para tener los mismos títulos).



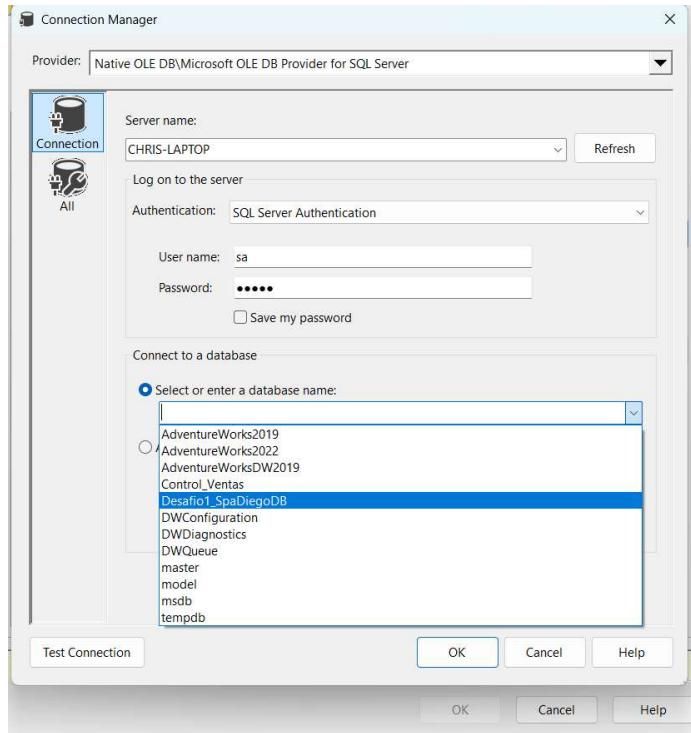
- Seleccioné la opción "Columns", para verificar que haya detectado todas las columnas.



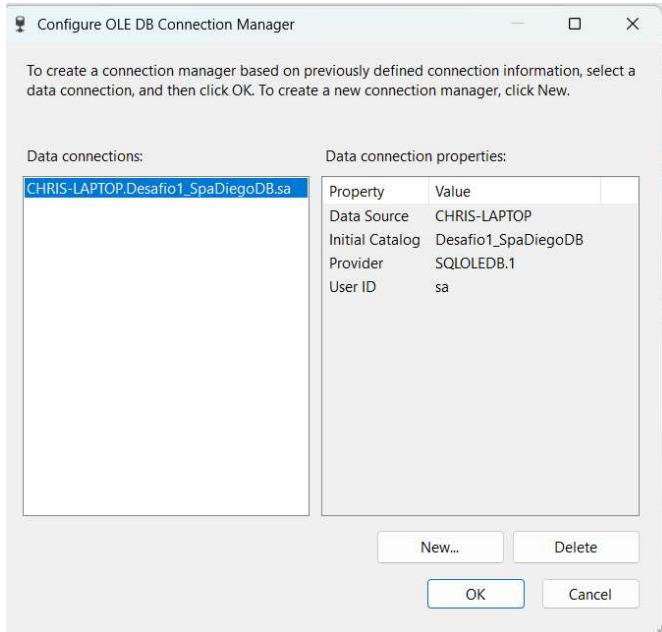
- Seleccioné "Ok" y luego cambie el nombre a Clientes_Centro.
- Luego creé una conexión OLE DB Destination y conecté el "source" con el "destination".



- Hice doble clic para configurar el destino.
 - Hice clic en "New" y en la nueva ventana "New" de nuevo.
 - Seleccioné Microsoft OLE DB Provider for SQL Server
 - Puse el nombre de mi server (en mi máquina es CHRIS-LAPTOP)
 - Seleccioné la base de datos Desafio1_SpaDiegoDB (creada antes)

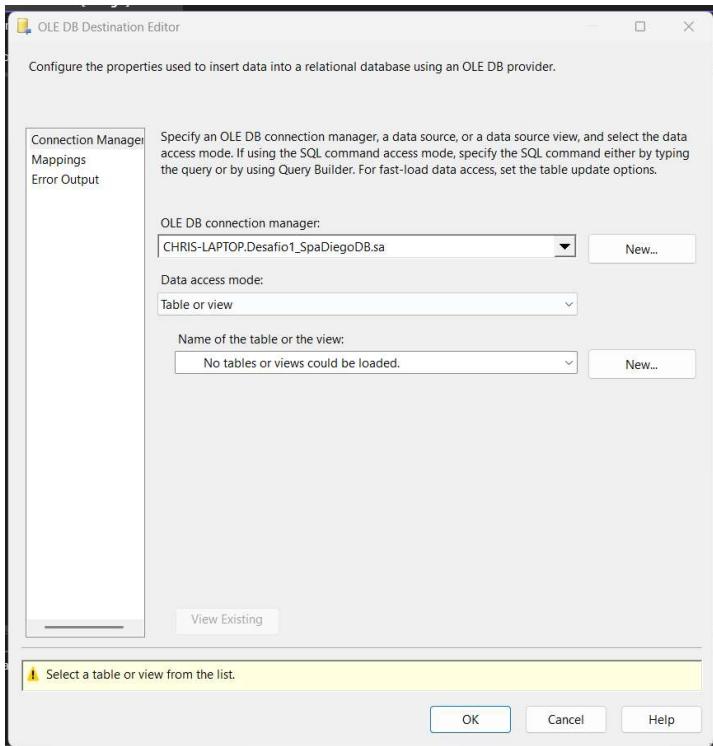


- Le doy clic en OK y OK de nuevo.

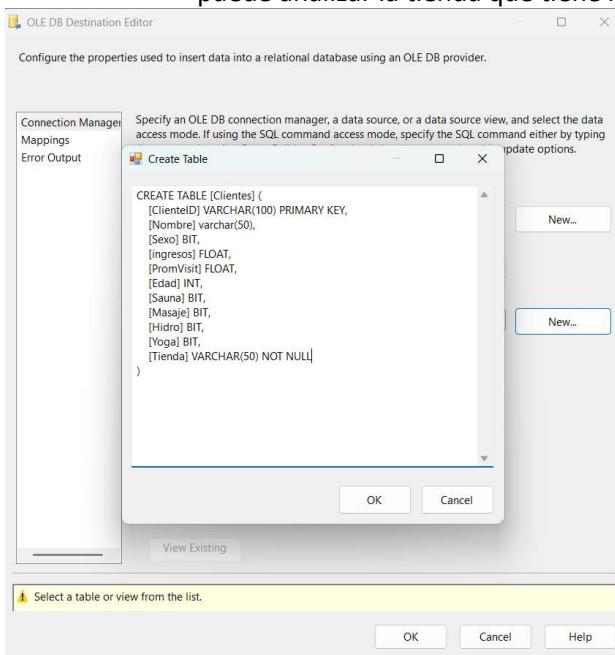


Ya teniendo la conexión de datos, hay que especificar la tabla a la que hay que mandar la información. Debido a que es datos de clientes, decidí crear la tabla "clientes".

- Seleccioné "New" en el nombre de la tabla.

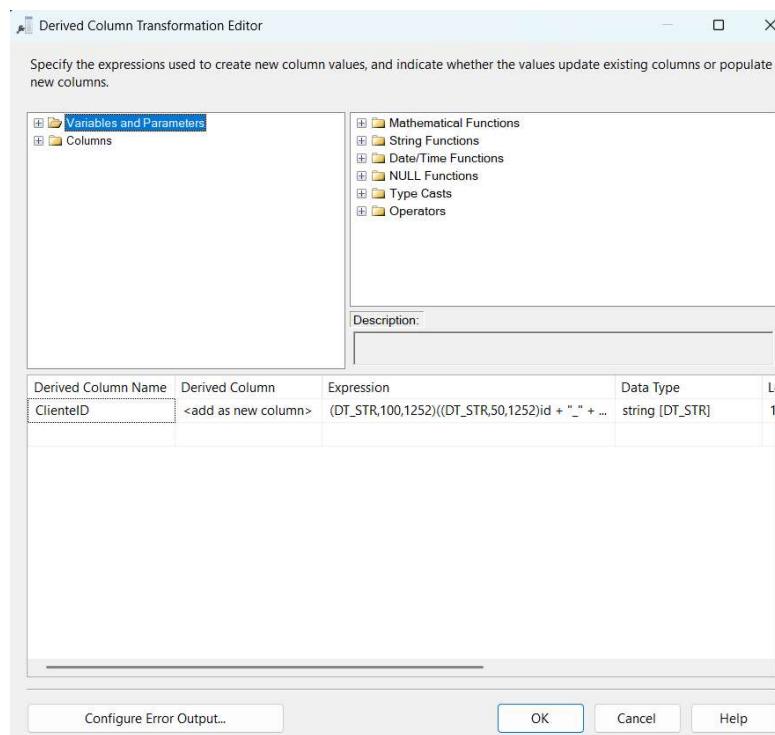
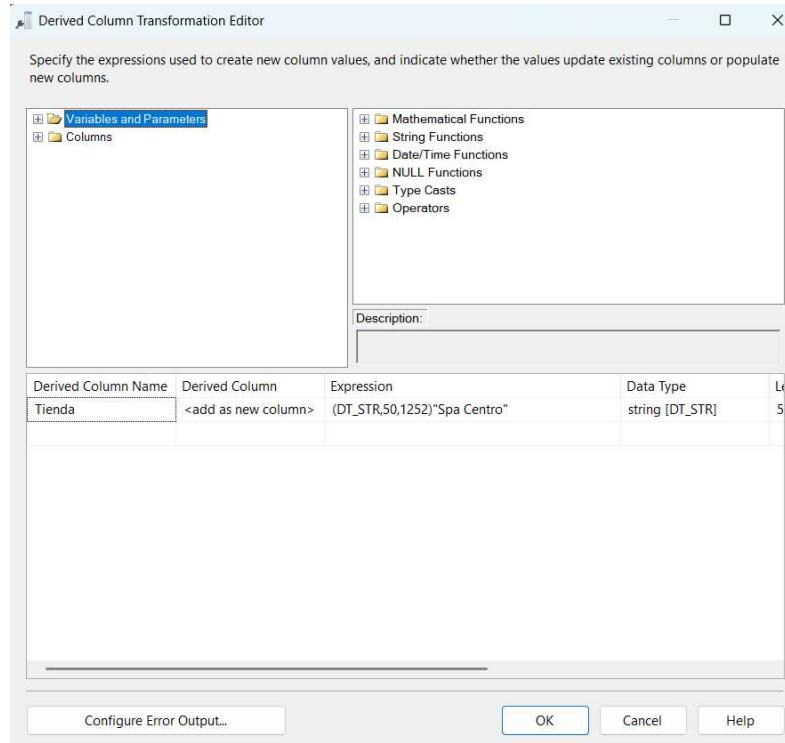


- Revisando la data, me hizo sentido crear la tabla de la manera que muestra la imagen mostrada abajo.
 - Hago notar que siguiendo las reglas de un Data Waterhouse, no he normalizado los campos.
 - Creé un campo que será calculado al que le llamé ClienteID, pues el nombre del cliente, aun si no se ha repetido, puede no ser la mejor llave primaria.
 - También he creado una columna llamada Tienda, para que cuando se cree la data, se puede analizar la tienda que tiene mayor cantidad de cada categoría de los clientes.

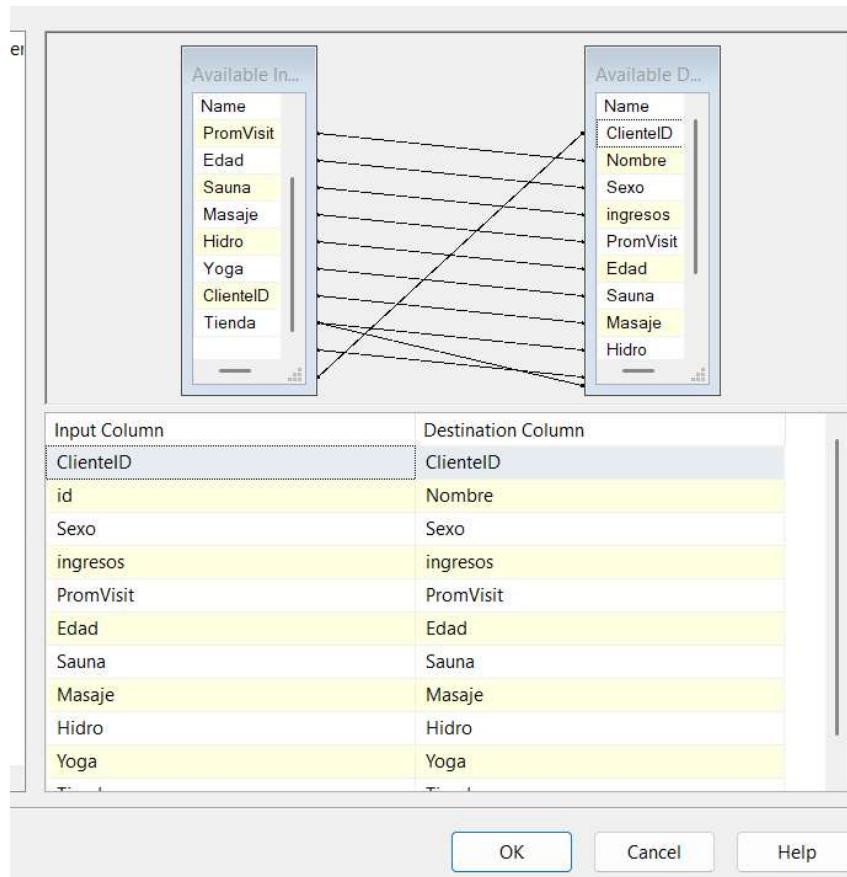


Le di OK y luego OK de nuevo. Regresaré después, para mapear bien los campos que voy a crear.

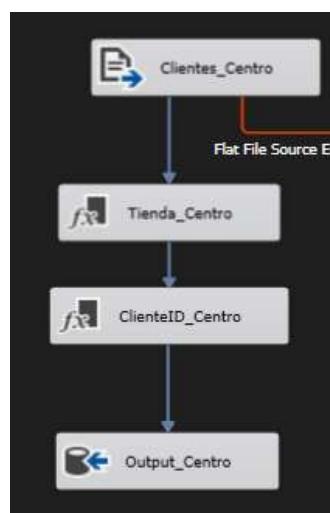
- Agregué una "Derived Column" en el "Data Flow", justo entre la conexión de la fuente y la base de datos.
 - Para hacer esto, quité la conexión de la source y la conecté con destination.
 - Junté el nombre (id) y la Tienda para hacer una llave primaria.



- Terminé de mapear los campos calculados.



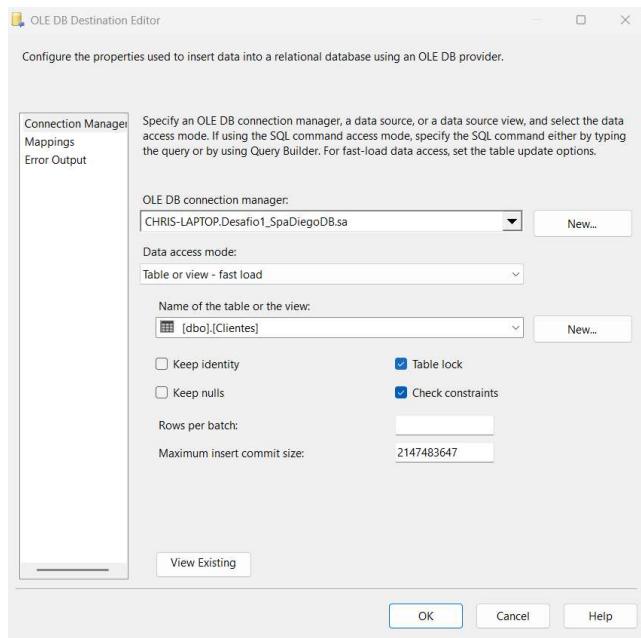
- Repetí el proceso de añadir clientes en el mismo "Data Flow" para los otros dos .CSV.



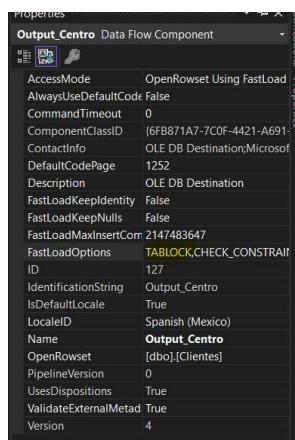
Manejar posibles errores

Evitar conflictos en la base de datos

- Cambié el Modo de Acceso a "Table or View - Fast Load"



- Verifiqué en propiedades que el TABLOCK esté activada.



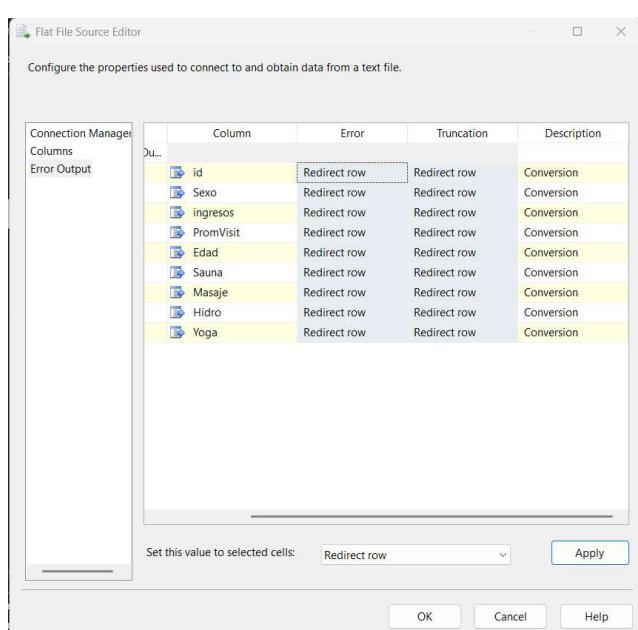
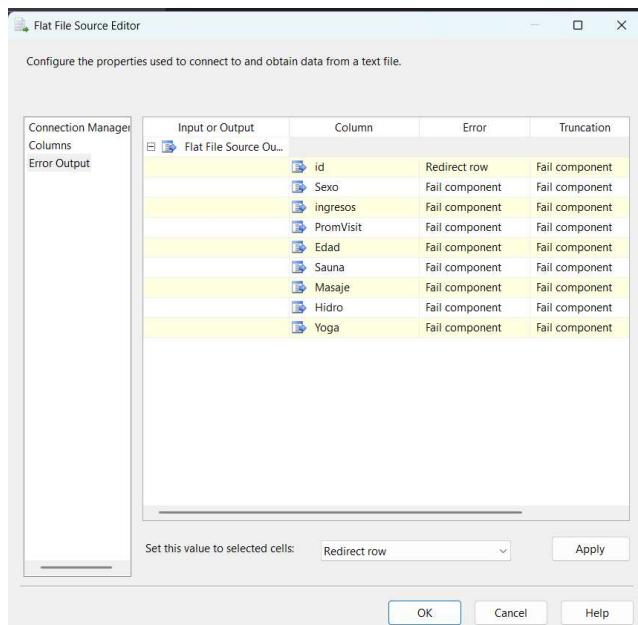
Esto lo hice para que los datos se carguen más rápido sin afectar otros procesos en SQL Server.

- "Table or View - Fast Load" usa transacciones eficientes y opciones avanzadas, como Table Lock, para acelerar la inserción de datos.
- TABLOCK: Habilita el bloqueo de tabla para mejorar el rendimiento al insertar datos en masa.
- CHECK_CONSTRAINTS: Verifica las restricciones de la tabla durante la inserción de datos.

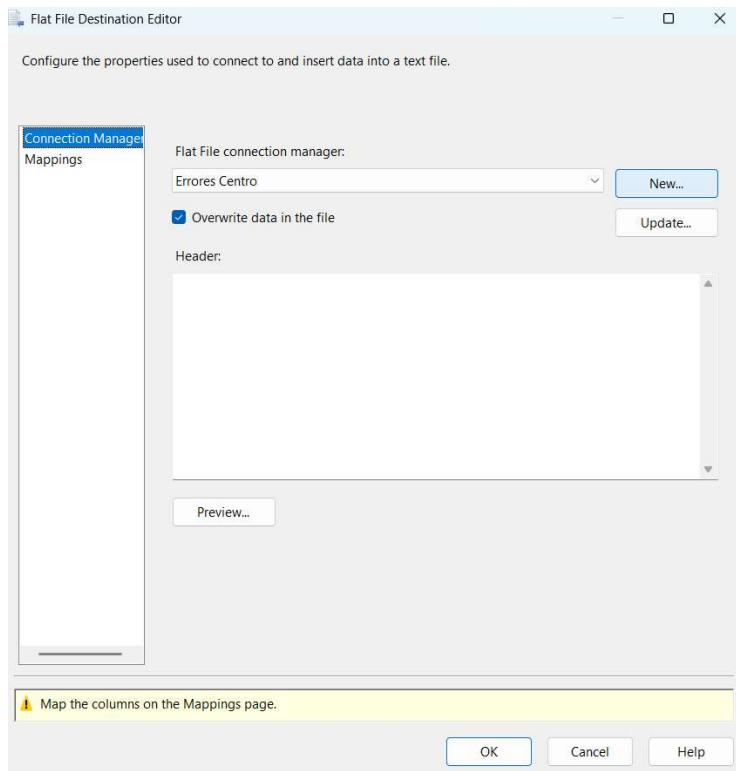
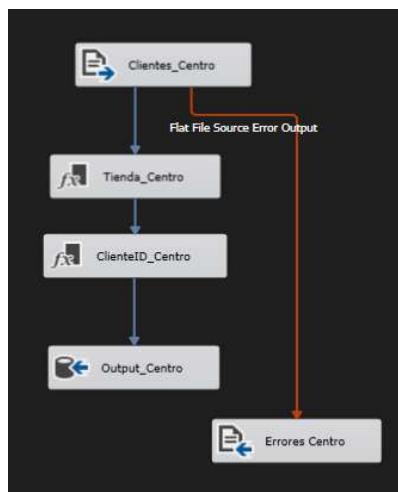
Evitar que errores individuales afecten todo el paquete (Error Output)

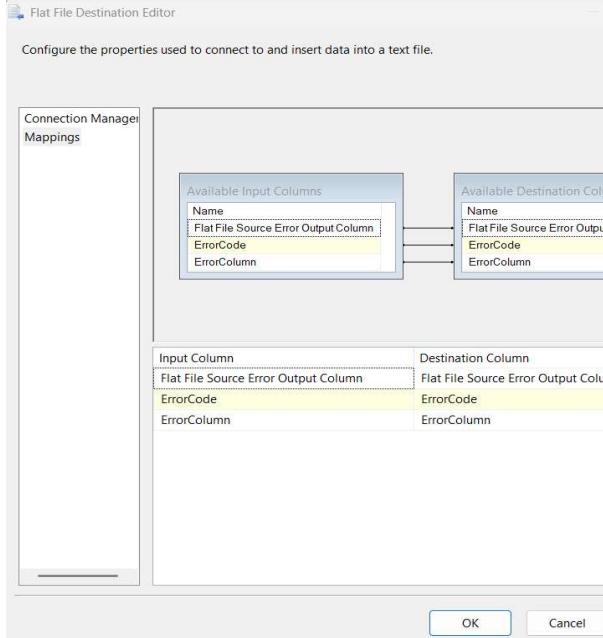
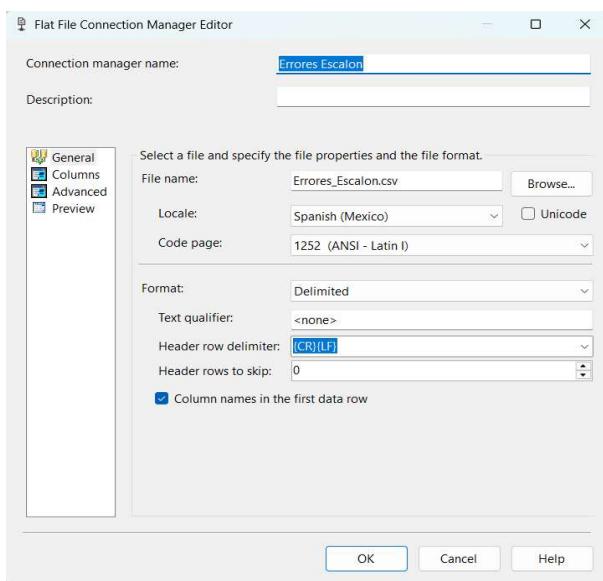
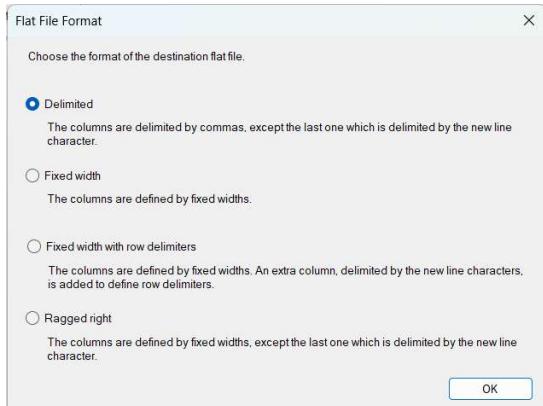
Si uno de los archivos CSV tiene errores (tipos de datos incorrectos, valores nulos, etc.), el paquete completo puede fallar. Para evitar esto, voy a configurar la salida de errores en SQL Server Integration Services.

- Hice doble clic en la fuente de la información (Clientes_Centro, Clientes_Escalón y Clientes_Santa_Tecla).
- Luego seleccioné la tab "Error Output" y cambié el error y truncation de "Fail component" a "Redirect Row".



- Luego añadí un "Flat File Destination" por cada .CSV, para tener una mejor idea de qué ha pasado.
 - Conecté el source y el flat file destination (con la flecha roja).
 - Me aseguré de que al darle doble clic, la ventana "Configure Error Output", todas las columnas tengan "Redirect Row" y presionar OK.
 - Al darle doble clic al destino de nuevo, en el Flat File Connection Manager seleccioné New, Delimited, Ok.
 - Le puse el nombre (por ejemplo Errores_Centro.CSV) en el editor y le di OK.
 - En la pestaña Mappings, pude verificar que las columnas estaban dirigidas de manera correcta.





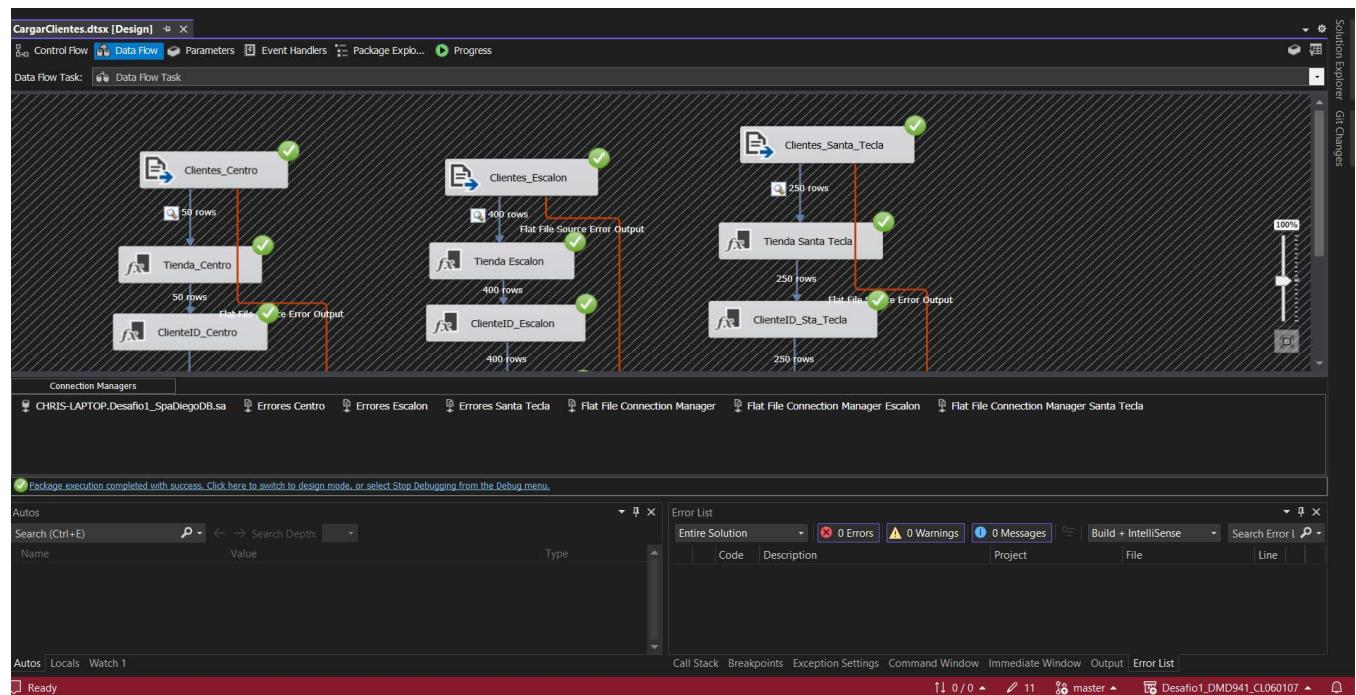
Este procedimiento lo repetí para los tres archivos.

Con esto conseguimos lo siguiente:

1. Si hay errores en los archivos CSV, los registros fallidos se guardarán en archivos separados.
2. No interrumpimos la ejecución del SSIS Package.
3. Nos facilita poder entrar a los archivos de error y corregir los datos defectuosos.

Comprobando el funcionamiento

Para comprobar el funcionamiento, ya que tengo todo listo, simplemente corrí el proceso y puse unos data viewers, para corroborar la cantidad de líneas que se procesaron.



Una vez el proceso termina, fui a ver la base de datos y confirmé que las 700 líneas se encontraban en ella.

Object Explorer

Connect ▾

CHRIS-LAPTOP (SQL Server 16.0.11135)

- Databases
 - System Database
 - Database Snapshots
 - AdventureWorks2019
 - AdventureWorks2022
 - AdventureWorksDW2019
 - Control_Ventas
 - Desafio1_SpaDiegoDB
- Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.Clientes
- Columns
- Keys
- Constraints
- Triggers
- Indexes
- Statistics
- Dropped Ledger Tables
- Views
- External Resources
- Synonyms

SQLQuery2.sql - CH...aDiegoDB (sa (90))*

```
select * from Clientes;
```

Results

	ClientID	Nombre	Sexo	ingresos	PromVisit	Edad	Seuna	Masaje	Hidro	Yoga	Tienda
1	Ab St Quenin_Spa Escalon	Ab St Quenin	1	2678.25	4.11	31	0	0	0	1	Spa Escalon
2	Abbi Boyd_Spa Escalon	Abbi Boyd	0	1803.34	3.63	62	1	1	0	1	Spa Escalon
3	Abelard Cassin_Spa Santa Tecla	Abelard Cassin	0	2077.93	6.85	38	1	1	0	1	Spa Santa Tecla
4	Ad Peebles_Spa Santa Tecla	Ad Peebles	0	1209.85	6.43	64	0	1	0	1	Spa Santa Tecla
5	Addy Dillington_Spa Escalon	Addy Dillington	0	1167.59	1.5	43	0	0	0	1	Spa Escalon
6	Adore Robotton_Spa Escalon	Adore Robotton	0	2723.01	4.76	61	1	1	0	0	Spa Escalon
7	Adriano Bellison_Spa Santa Tecla	Adriano Bellison	0	1874.17	2.44	55	1	1	0	0	Spa Santa Tecla
8	Aeriel Gligori_Spa Escalon	Aeriel Gligori	0	608.07	1.69	50	0	1	1	1	Spa Escalon
9	Aggi Cunningham_Spa Centro	Aggi Cunningham	0	2268.57	3.53	33	1	1	0	0	Spa Centro
10	Aguustin Spira_Spa Centro	Aguustin Spira	1	255.94	5.35	54	1	1	1	0	Spa Centro
11	Al Gibre_Spa Escalon	Al Gibre	0	428.3	5.85	46	0	0	0	0	Spa Escalon
12	Alexi Wildman_Spa Santa Tecla	Alexi Wildman	0	1734.3	3.92	62	0	0	0	1	Spa Santa Tecla
13	Alford Pieroni_Spa Santa Tecla	Alford Pieroni	1	1155.83	1.59	39	0	1	0	0	Spa Santa Tecla
14	Alfreda Pere_Spa Escalon	Alfreda Pere	1	2493.15	1.18	51	1	0	0	0	Spa Escalon
15	Alina Edson_Spa Escalon	Alina Edson	1	400.34	2.9	38	1	0	1	1	Spa Escalon
16	Alis Langton_Spa Escalon	Alis Langton	1	453.2	5.54	27	0	1	0	1	Spa Escalon

Query executed successfully.

CHRIS-LAPTOP (16.0 RTM) | sa (90) | Desafio1_SpaDiegoDB | 00:00:00 | 700 rows

| Desafio1_SpaDiegoDB | 00:00:00 | 700 rows

Como se puede observar, todo está llegando bien, lo cual cumple nuestro objetivo.

Obtener la data para el análisis de la campaña

Para hacer esto, cree otro "data flow task" en el área de control flow y la conecté con la original.

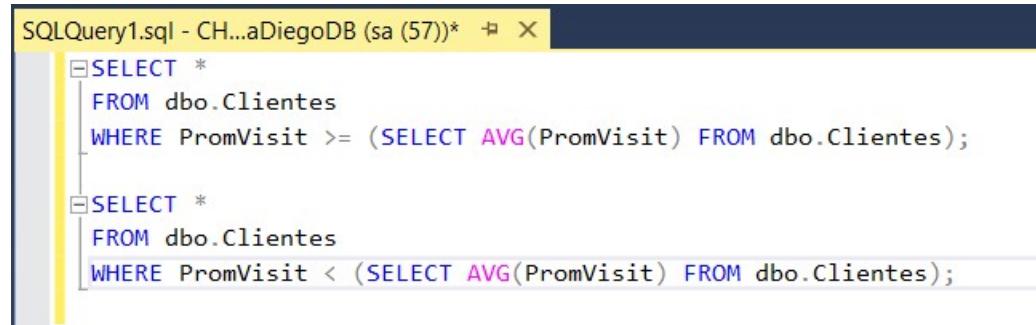


Abrí "Exportar datos de campaña" haciendo doble clic y procedí a crear el flujo para descargar en los .CSV de Clientes_Frecuentes.CSV y Clientes_Ocasionales.CSV.

Antes de escribir un Query o poner sources, me aseguré de que tenía Queries funcionales, así que me puse a hacer pruebas con SELECTS en la base de datos.

```
SELECT * FROM dbo.Clientes ORDER BY dbo.Clientes.PromVisit DESC;  
  
SELECT AVG(dbo.Clientes.PromVisit) FROM dbo.Clientes;  
  
SELECT *  
FROM dbo.Clientes  
WHERE PromVisit >= (SELECT AVG(PromVisit) FROM dbo.Clientes);
```

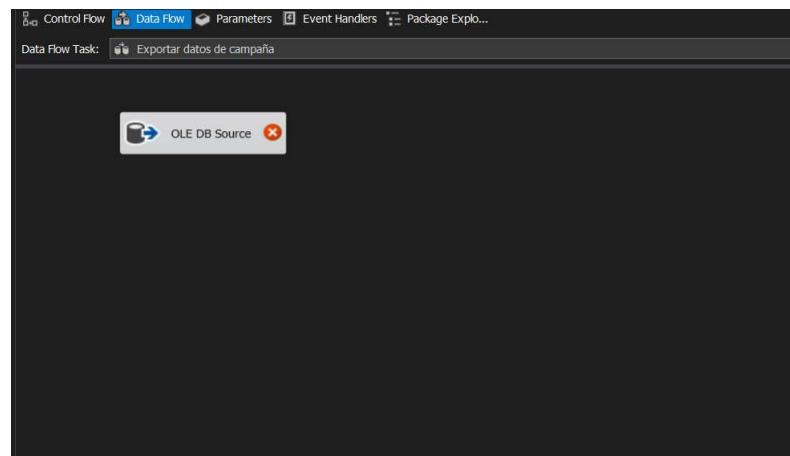
Al final me fui por los dos SELECTS mostrados abajo, el primero para los clientes que más nos visitan utilizando el promedio de visita (promVisit) y la segunda para los que menos nos visitan, usando lo mismo.



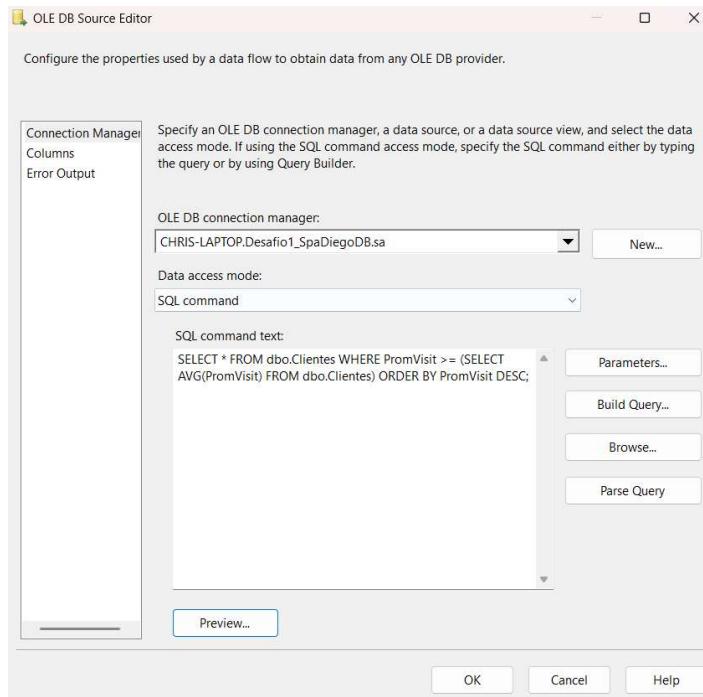
```
SQLQuery1.sql - CH...aDiegoDB (sa (57))*  ▶ X  
SELECT *  
FROM dbo.Clientes  
WHERE PromVisit >= (SELECT AVG(PromVisit) FROM dbo.Clientes);  
  
SELECT *  
FROM dbo.Clientes  
WHERE PromVisit < (SELECT AVG(PromVisit) FROM dbo.Clientes);
```

Hago notar que hice un promedio de PromVisit, para que si en un futuro este promedio cambia, podamos asegurarnos que las consultas siempre sean adecuadas.

- Agregué un OLE DB Source en el nuevo data flow.



- Lo configuré para que se conectara a la misma DB y utilicé la consulta que había creado anteriormente para los clientes frecuentes.
 - Aparte de esto, le puse un "ORDER BY PromVisit DESC" para que siempre se muestren los clientes con mayores valores.



Specify an OLE DB connection manager, a data source, or a data source view.

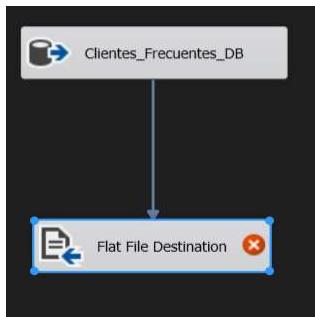
Preview Query Results

Query result (up to the first 200 rows):

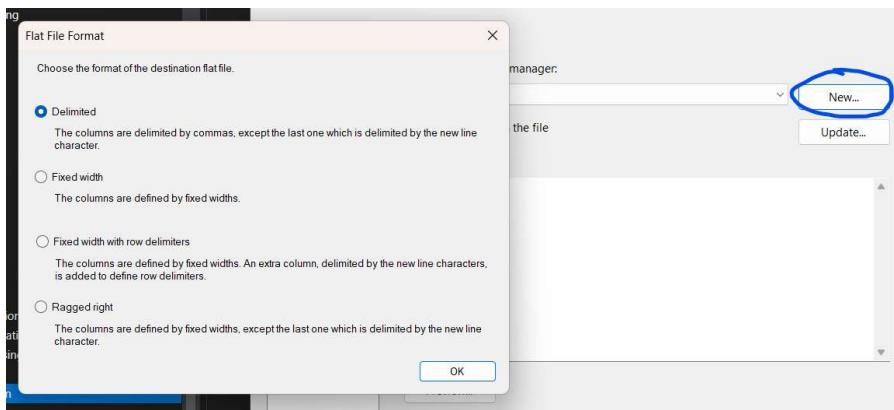
ClienteID	Nombre	Sexo	ingresos	PromVisit	Edad
Breena ...	Breena ...	False	1719.65	7	28
Fidel St...	Fidel St...	True	683.62	7	40
Duffy P...	Duffy P...	True	2213.12	6.99	62
Glory S...	Glory S...	False	293.78	6.96	32
Webste...	Webste...	False	1620.76	6.96	37
Leoline ...	Leoline ...	True	1934.92	6.94	36
Olivier B...	Olivier B...	False	1788.62	6.93	54
Putnam...	Putnam...	False	1446	6.92	52
Linnet ...	Linnet ...	False	2303.69	6.9	30
Robinia ...	Robinia ...	False	1285.23	6.89	26
Candi C...	Candi C...	False	259.06	6.89	33
Ana De...	Ana De...	False	705.04	6.88	21
Cynthy...	Cynthy...	False	1175.45	6.88	50
Abelard...	Abelard...	False	2077.93	6.85	38
Cointon...	Cointon...	True	1709.8	6.83	64

Close

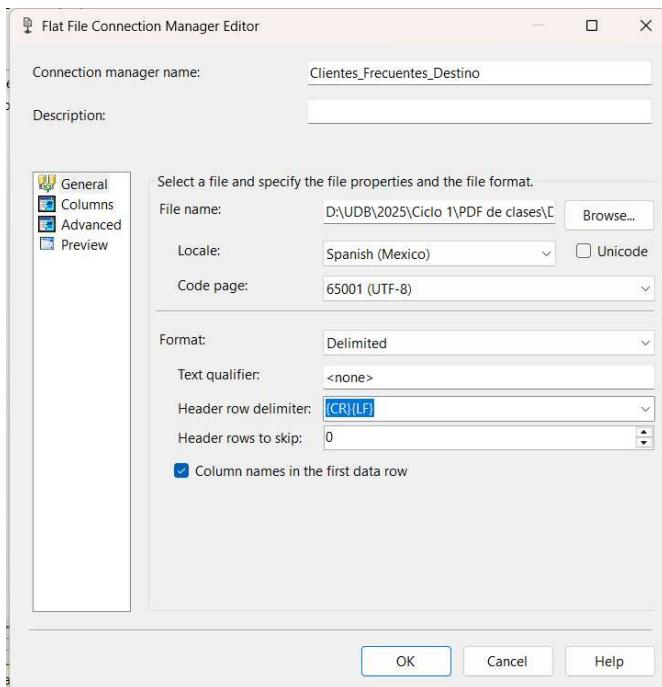
- Creé un nuevo "Flat File Destination" y lo conecté con la entrada de la DB.



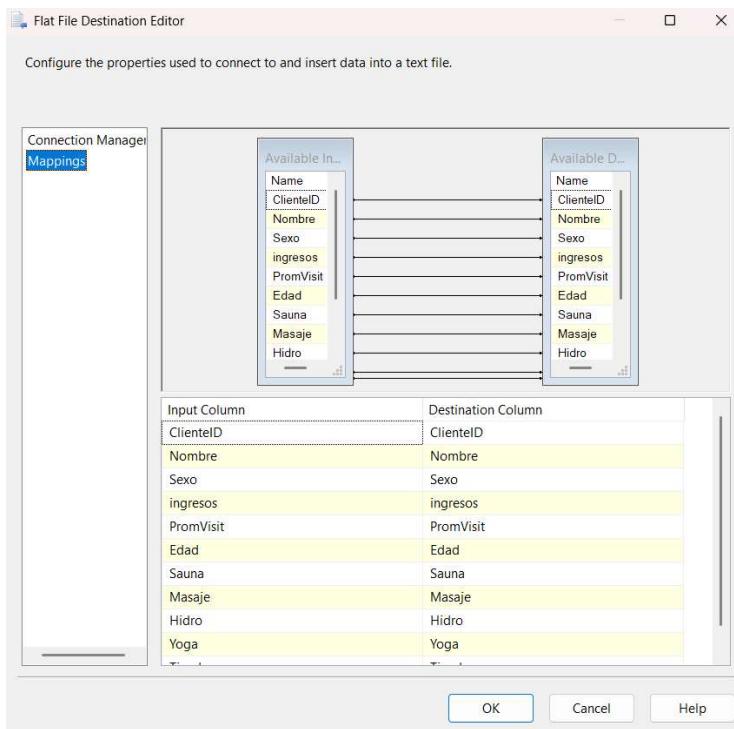
- Le di clic en "New" para poder conectarlo con Clientes_Frecuentes



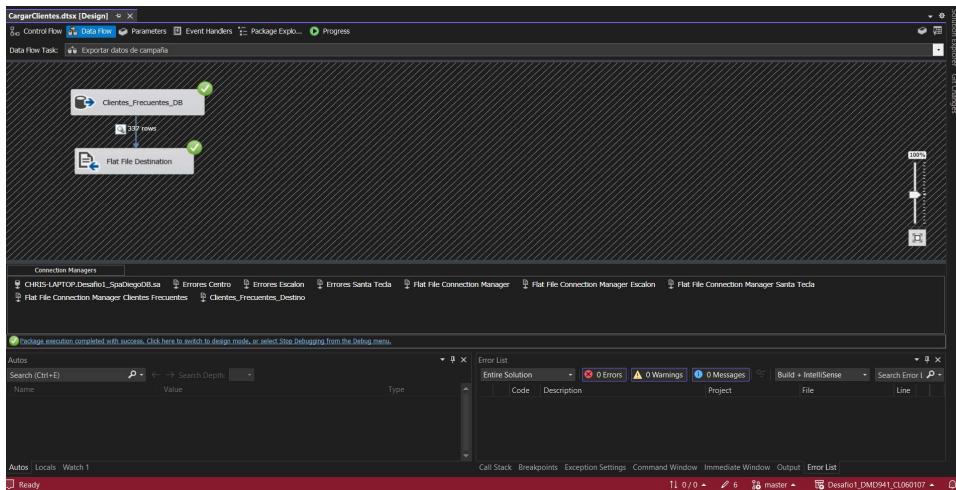
- Me aseguré de que esté conectado correctamente con el archivo.



- Luego en "Mappings" me aseguré de que las columnas estuvieran conectadas adecuadamente.



- Despues de esto, ejecuté el data flow y verifiqué que los datos llegaran al .CSV



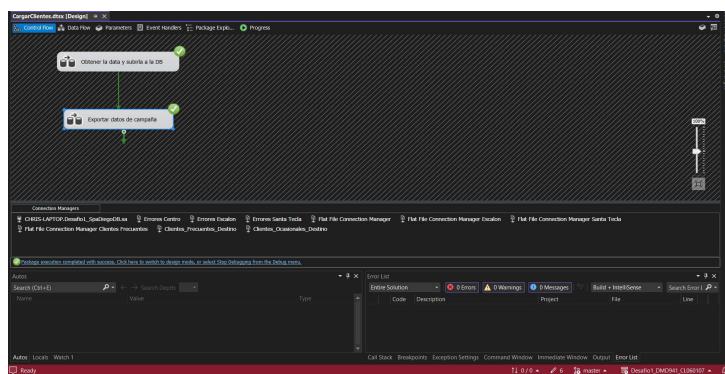
Clientes_Frecuentes

	A	B	C	D	E	F	G	H	I	J	K	L
1	ClientID	Nombre	Sexo	ingresos	PromVisit	Edad	Sauna	Masaje	Hidro	Yoga	Tienda	
2	Breena Wellb	Breena Wellb	False	1719.65	7	28	False	False	False	True	Spa Santa Tecla	
3	Fidel Staples	Fidel Staples	True	683.62	7	40	True	False	True	False	Spa Escalon	
4	Duffy Pallusch	Duffy Pallusch	True	2213.12	6.99	62	True	True	False	False	Spa Escalon	
5	Glory Shearst	Glory Shearst	False	293.78	6.96	32	False	True	False	True	Spa Escalon	
6	Webster Ivash	Webster Ivash	False	1620.76	6.96	37	False	False	False	True	Spa Santa Tecla	
7	Leoline Beeke	Leoline Beeke	True	1934.92	6.94	36	False	True	False	False	Spa Escalon	
8	Olivier Buzine	Olivier Buzine	False	1788.62	6.93	54	True	True	False	False	Spa Santa Tecla	
9	Putnam Healin	Putnam Healin	False	1446	6.92	52	False	False	False	False	Spa Escalon	
10	I Innet Reinher	I Innet Reinher	False	2303.69	6.9	30	False	False	True	False	Spa Santa Tecla	
11	Robinia Hinric	Robinia Hinric	False	1285.23	6.89	26	False	True	False	False	Spa Santa Tecla	
12	Candi Colefor	Candi Colefor	False	259.06	6.89	33	True	True	False	False	Spa Santa Tecla	
13	Ana Deppe	Ana Deppe	False	705.04	6.88	21	False	False	False	False	Spa Santa Tecla	
14	Cynthia Crich	Cynthia Crich	False	1175.45	6.88	50	True	True	False	True	Spa Santa Tecla	
15	Abelard Cassi	Abelard Cassi	False	2077.93	6.85	38	True	True	False	True	Spa Santa Tecla	
16	Cointon Tignei	Cointon Tignei	True	1709.8	6.83	64	True	True	False	False	Spa Centro	
17	Sisely Joscely	Sisely Joscely	False	1982.66	6.83	27	True	False	True	False	Spa Escalon	
18	Milo Burnhard	Milo Burnhard	False	1719.99	6.8	50	False	False	False	False	Spa Escalon	
19	Reid Henrichs	Reid Henrichs	True	2327.39	6.79	64	True	True	False	True	Spa Santa Tecla	
20	Lammond Yur	Lammond Yur	False	794.05	6.77	31	True	True	False	False	Spa Santa Tecla	
21	Winna Conklin	Winna Conklin	True	1992.07	6.77	49	False	False	True	False	Spa Santa Tecla	
22	Jock Slater	Jock Slater	False	2708.88	6.76	40	True	True	True	False	Spa Escalon	
23	Joella Abys	Joella Abys	False	2755.22	6.75	22	False	False	False	True	Spa Santa Tecla	
24	Deloria Gilleor	Deloria Gilleor	False	2996.03	6.73	34	True	True	True	False	Spa Escalon	
25	Harriett MacC	Harriett MacC	False	1144.77	6.73	49	False	True	True	False	Spa Escalon	
26	Rosemaria Mc	Rosemaria Mc	True	2309.15	6.72	50	True	True	False	False	Spa Centro	
27	Pepito Brumh	Pepito Brumh	False	1393.59	6.72	62	True	False	False	False	Spa Santa Tecla	
28	Trever Du Fray	Trever Du Fray	True	1937.06	6.7	45	True	True	True	True	Spa Escalon	
29	Pietro Gwaltei	Pietro Gwaltei	True	2511.42	6.69	39	True	False	True	True	Spa Escalon	
30	Julius Battell	Julius Battell	False	275.38	6.69	48	False	True	True	True	Spa Santa Tecla	

Repliqué esto para Clientes_Ocasionales.



Al final, procedí a correr todo para asegurarme que no hubiera conflicto a la hora de procesar todo.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Abbi Boyd_Sj	Abbi Boyd	False	1803.34	3.63	62	True	True	False	True	Spa Escalon			
2	Ady Dillinton	Ady Dillinton	False	1167.58	1.5	43	False	False	False	True	Spa Escalon			
3	Adriano Bellis	Adriano Bellis	False	1874.17	2.44	55	True	True	False	False	Spa Santa Tecla			
4	Aeriel Gligori	Aeriel Gligori	False	608.07	1.69	50	False	True	True	True	Spa Escalon			
5	Aggi Conningh	Aggi Conningh	False	2268.57	3.53	33	True	True	False	False	Spa Centro			
6	Alford Pieroni	Alford Pieroni	True	1155.83	1.59	39	False	True	False	False	Spa Santa Tecla			
7	Alfreda Pere	Alfreda Pere	True	2493.15	1.18	51	True	False	False	False	Spa Escalon			
8	Alina Edson	Alina Edson	True	400.34	2.9	38	True	False	True	True	Spa Escalon			
9	Almeria O'Kiel	Almeria O'Kiel	False	647.64	2.95	48	False	False	False	False	Spa Santa Tecla			
10	Amabel Halhe	Amabel Halhe	False	1487.82	2.26	26	False	False	False	True	Spa Santa Tecla			
11	Amanda Apfle	Amanda Apfle	False	913.26	1.67	39	True	False	False	False	Spa Santa Tecla			
12	Ame Crum	Ame Crum	True	2999.2	3.28	30	True	True	True	True	Spa Escalon			
13	Amil Issit	Spz Amil Issit	True	2172.6	1.47	27	False	True	False	True	Spa Escalon			
14	Anabella Bugg	Anabella Bugg	False	786.44	2.5	22	False	False	True	True	Spa Santa Tecla			
15	Angelita Jaque	Angelita Jaque	True	1458.84	2.82	36	True	True	False	True	Spa Escalon			
16	Angelle Ewenc	Angelle Ewenc	False	256.41	3.4	36	False	False	False	False	Spa Santa Tecla			
17	Annabell O'Liu	Annabell O'Liu	False	2354.55	3.15	33	False	True	True	False	Spa Escalon			
18	Anneliese Iwa	Anneliese Iwa	False	668.68	1.93	25	False	True	True	False	Spa Escalon			
19	Ann Slora	Ann Slora	True	2213.14	1.65	50	False	False	False	True	Spa Escalon			
20	Anristice Drach	Anristice Drach	False	2916.96	2.57	49	True	True	True	True	Spa Santa Tecla			
21	Antonin Rolfo	Antonin Rolfo	True	1862.99	2.9	28	False	False	False	True	Spa Escalon			
22	Arabelle Hard	Arabelle Hard	True	1626.07	2.64	23	True	False	True	True	Spa Escalon			
23	Arabelle Bale	Arabelle Bale	False	2945.95	3.49	30	True	False	True	True	Spa Escalon			
24	Arch Cawthn	Arch Cawthn	True	2031.06	2.32	34	True	True	False	True	Spa Escalon			

Pude confirmar que tanto los archivos Clientes_Ocasionales.CSV como Clientes_Frecuentes.CSV se exportan correctamente.

Lo último que he hecho es asegurarme que el folder GIT ya tiene todos los cambios y subirlo al “main”.

Análisis de los datos

Lo primero que he utilizado es un query que nos puede decir cuál es el servicio más utilizado por sucursal.

```
SELECT Tienda,
       'Sauna' AS Servicio,
       COUNT(*) AS Cantidad
  FROM Clientes
 WHERE Sauna = 1
 GROUP BY Tienda
```

UNION

```
SELECT Tienda,
       'Masaje' AS Servicio,
       COUNT(*) AS Cantidad
  FROM Clientes
 WHERE Masaje = 1
 GROUP BY Tienda
```

UNION

```

SELECT Tienda,
      'Hidro' AS Servicio,
      COUNT(*) AS Cantidad
  FROM Clientes
 WHERE Hidro = 1
 GROUP BY Tienda

```

UNION

```

SELECT Tienda,
      'Yoga' AS Servicio,
      COUNT(*) AS Cantidad
  FROM Clientes
 WHERE Yoga = 1
 GROUP BY Tienda;

```

The screenshot shows a Microsoft SQL Server Management Studio (SSMS) interface. At the top, there are tabs for 'Results' and 'Messages'. The results pane displays a table with three columns: 'Tienda' (Store), 'Servicio' (Service), and 'Cantidad' (Quantity). The data is ordered by 'Cantidad' in descending order. The table contains 12 rows, with the first row being highlighted.

	Tienda	Servicio	Cantidad
1	Spa Centro	Hidro	17
2	Spa Centro	Masaje	42
3	Spa Centro	Sauna	39
4	Spa Centro	Yoga	14
5	Spa Escalon	Hidro	199
6	Spa Escalon	Masaje	200
7	Spa Escalon	Sauna	194
8	Spa Escalon	Yoga	199
9	Spa Santa Tecla	Hidro	55
10	Spa Santa Tecla	Masaje	124
11	Spa Santa Tecla	Sauna	51
12	Spa Santa Tecla	Yoga	148

De esta manera se puede determinar que el servicio más utilizado en la sucursal del centro es de masajes, seguido de cerca por sauna, mientras que los menos utilizados son Yoga con 14 servicios brindados e hidro con 17 servicios brindados. Se recomienda utilizar promociones que incluyan los dos servicios más altos (masaje y sauna) y unirlos con servicios menos populares (clases de yoga e hidro), alguna promoción 2X1 o similar también puede incentivar a los clientes a utilizar Hidro y Yoga.

En el spa Escalón se puede ver que no hay tanta diferencia entre la cantidad de servicios brindados, en primer lugar, tenemos masajes con 200 contratados, seguidos por hidro y yoga con 199 servicios contratados cada uno, terminando con la sauna con 194 servicios. Se recomienda considerar crear promociones que incluyan a la sauna, para incentivar a los clientes que no lo han probado o no lo utilizan tan seguido a utilizar el servicio.

En la sucursal de Santa Tecla, se puede observar que Yoga es el servicio más popular por mucho, teniendo 148 servicios contratados, seguido por los masajes con 124 servicios. La diferencia es casi 3 veces menor para los otros dos servicios, hidro (55) y sauna (51).

Con la siguiente consulta se muestra el top producto por sucursal.

```
WITH Usos AS (
    SELECT Tienda, 'Sauna' AS Servicio, COUNT(*) AS Cantidad
    FROM Clientes
    WHERE Sauna = 1
    GROUP BY Tienda
    UNION
    SELECT Tienda, 'Masaje' AS Servicio, COUNT(*) AS Cantidad
    FROM Clientes
    WHERE Masaje = 1
    GROUP BY Tienda
    UNION
    SELECT Tienda, 'Hidro' AS Servicio, COUNT(*) AS Cantidad
    FROM Clientes
    WHERE Hidro = 1
    GROUP BY Tienda
    UNION
    SELECT Tienda, 'Yoga' AS Servicio, COUNT(*) AS Cantidad
    FROM Clientes
    WHERE Yoga = 1
    GROUP BY Tienda
)
SELECT Tienda, Servicio, Cantidad
FROM (
    SELECT *,
        ROW_NUMBER() OVER (PARTITION BY Tienda ORDER BY Cantidad DESC) AS rn
    FROM Usos
) AS Sub
WHERE rn = 1
ORDER BY Tienda;
```

	Tienda	Servicio	Cantidad
1	Spa Centro	Masaje	42
2	Spa Escalon	Masaje	200
3	Spa Santa Tecla	Yoga	148

Con lo que podemos decir que tanto en sucursal Centro como en Escalón se puede notar la preferencia por el servicio de masajes.

Luego me puse a corroborar por sus grupos (clientes frecuentes y clientes ocasionales) para poder ver si la tendencia es similar y obtuve resultados similares, teniendo que los clientes frecuentes han tenido más masajes (186) seguidos por Yoga (181). En el caso de clientes ocasionales, tanto masajes como yoga tienen la misma cantidad de servicios vendidos (180).

```
WITH CTE_Grupos AS (
    SELECT *,
        CASE
```

```

        WHEN PromVisit >= (SELECT AVG(PromVisit) FROM dbo.Clientes)
        THEN 'Frecuentes'
        ELSE 'Ocasionales'
    END AS GrupoVisita
    FROM dbo.Clientes
)
SELECT GrupoVisita,
    SUM(CAST(Sauna AS INT)) AS TotalSauna,
    SUM(CAST(Masaje AS INT)) AS TotalMasaje,
    SUM(CAST(Hidro AS INT)) AS TotalHidro,
    SUM(CAST(Yoga AS INT)) AS TotalYoga
FROM CTE_Grupos
GROUP BY GrupoVisita;

```

	GrupoVisita	TotalSauna	TotalMasaje	TotalHidro	TotalYoga
1	Frecuentes	146	186	133	181
2	Ocasionales	138	180	138	180

Por último, creé una consulta que nos muestra las tendencias de compra por tipo de cliente (frecuentes u ocasionales), genero, tres rangos de edad (menores de 30 años, de 30 a 59 y mayor a 59 años).

```

WITH CTE_Grupos AS (
    SELECT *,
        CASE
            WHEN PromVisit >= (SELECT AVG(PromVisit) FROM dbo.Clientes)
            THEN 'Frecuentes'
            ELSE 'Ocasionales'
        END AS GrupoVisita
    FROM dbo.Clientes
)
SELECT
    GrupoVisita,
    CASE WHEN Sexo = 1 THEN 'Hombre' ELSE 'Mujer' END AS Genero,
    CASE
        WHEN Edad < 30 THEN 'Menor de 30'
        WHEN Edad BETWEEN 30 AND 59 THEN '30-59'
        ELSE '60 o más'
    END AS RangoEdad,
    Tienda,
    SUM(CAST(Sauna AS INT)) AS TotalSauna,
    SUM(CAST(Masaje AS INT)) AS TotalMasaje,
    SUM(CAST(Hidro AS INT)) AS TotalHidro,
    SUM(CAST(Yoga AS INT)) AS TotalYoga
FROM CTE_Grupos
GROUP BY
    GrupoVisita,
    CASE WHEN Sexo = 1 THEN 'Hombre' ELSE 'Mujer' END,

```

CASE

WHEN Edad < 30 THEN 'Menor de 30'

WHEN Edad BETWEEN 30 AND 59 THEN '30-59'

ELSE '60 o más'

END,

Tienda

ORDER BY

GrupoVisita,

Genero,

RangoEdad,

Tienda;

	GrupoVisita	Genero	RangoEdad	Tienda	TotalSauna	TotalMasaje	TotalHidro	TotalYoga
1	Frecuentes	Hombre	30-59	Spa Centro	9	9	4	1
2	Frecuentes	Hombre	30-59	Spa Escalon	35	32	39	34
3	Frecuentes	Hombre	30-59	Spa Santa Tecla	3	7	4	10
4	Frecuentes	Hombre	60 o más	Spa Centro	1	2	1	0
5	Frecuentes	Hombre	60 o más	Spa Escalon	5	10	7	7
6	Frecuentes	Hombre	60 o más	Spa Santa Tecla	2	4	1	4
7	Frecuentes	Hombre	Menor de 30	Spa Centro	4	4	1	2
8	Frecuentes	Hombre	Menor de 30	Spa Escalon	9	9	9	10
9	Frecuentes	Hombre	Menor de 30	Spa Santa Tecla	1	3	2	6
10	Frecuentes	Mujer	30-59	Spa Centro	3	4	2	3
11	Frecuentes	Mujer	30-59	Spa Escalon	36	37	29	37
12	Frecuentes	Mujer	30-59	Spa Santa Tecla	16	34	13	35
13	Frecuentes	Mujer	60 o más	Spa Centro	1	1	0	0
14	Frecuentes	Mujer	60 o más	Spa Escalon	7	9	6	4
15	Frecuentes	Mujer	60 o más	Spa Santa Tecla	2	3	0	8
16	Frecuentes	Mujer	Menor de 30	Spa Centro	2	2	1	1
17	Frecuentes	Mujer	Menor de 30	Spa Escalon	8	10	9	9
18	Frecuentes	Mujer	Menor de 30	Spa Santa Tecla	2	6	5	10
19	Ocasionales	Hombre	30-59	Spa Centro	9	11	5	4
20	Ocasionales	Hombre	30-59	Spa Escalon	36	33	43	41
21	Ocasionales	Hombre	30-59	Spa Santa Tecla	2	13	3	7
22	Ocasionales	Hombre	60 o más	Spa Centro	3	2	2	2
23	Ocasionales	Hombre	60 o más	Spa Escalon	8	7	4	7
24	Ocasionales	Hombre	60 o más	Spa Santa Tecla	0	0	0	2
25	Ocasionales	Hombre	Menor de 30	Spa Centro	2	1	0	0
26	Ocasionales	Hombre	Menor de 30	Spa Escalon	11	9	5	10
27	Ocasionales	Hombre	Menor de 30	Spa Santa Tecla	0	0	0	3
28	Ocasionales	Mujer	30-59	Spa Centro	4	4	0	1
29	Ocasionales	Mujer	30-59	Spa Escalon	27	29	37	27
30	Ocasionales	Mujer	30-59	Spa Santa Tecla	15	39	16	37
31	Ocasionales	Mujer	60 o más	Spa Centro	1	1	1	0
32	Ocasionales	Mujer	60 o más	Spa Escalon	3	6	4	3
33	Ocasionales	Mujer	60 o más	Spa Santa Tecla	1	2	2	8
34	Ocasionales	Mujer	Menor de 30	Spa Centro	0	1	0	0
35	Ocasionales	Mujer	Menor de 30	Spa Escalon	9	9	7	10
36	Ocasionales	Mujer	Menor de 30	Spa Santa Tecla	7	13	9	18

Conclusiones

Segmentación por Frecuencia de Visitas

- Los **Clientes Frecuentes** (con visitas iguales o superiores al promedio) son los que más utilizan los servicios, lo que los convierte en el grupo ideal para programas de fidelización.
- Los **Clientes Ocasionales** (con visitas por debajo del promedio) representan una oportunidad para campañas de reactivación.

Recomendación: Implementar programas VIP para clientes frecuentes y promociones especiales para incentivar a los ocasionales a regresar.

Segmentación por Género

- Las **mujeres** prefieren **Masajes y Yoga**, mientras que los **hombres** utilizan más el **Sauna y la Hidroterapia**.
- Las mujeres en el grupo de **Clientes Frecuentes** representan una proporción significativa del consumo de servicios.

Recomendación: Diseñar promociones personalizadas según género y aprovechar programas de referidos dirigidos a mujeres.

Segmentación por Edad

- Los clientes de **30 a 59 años** son los más activos, especialmente en **Masajes y Sauna**.
- Los clientes **menores de 30 años** muestran una mayor preferencia por el **Yoga**.
- Los clientes **mayores de 60 años** tienen menor participación, pero pueden beneficiarse de servicios como **Hidroterapia y tratamientos relajantes**.

Recomendación: Crear paquetes especiales para clientes de 30-59 años, promociones de Yoga para jóvenes y programas de bienestar para adultos mayores.

Análisis por Sucursal

- **Spa Escalón** es la sucursal con mayor actividad y preferencia por **Masajes y Yoga**.
- **Spa Santa Tecla** tiene un buen desempeño, pero con oportunidades de crecimiento, especialmente en clientes ocasionales.
- **Spa Centro** muestra la menor participación y requiere estrategias para atraer más clientes.

Recomendación: Fortalecer la fidelización en **Spa Escalón**, atraer clientes ocasionales en **Spa Santa Tecla**, y aumentar la visibilidad y el "engagement" en **Spa Centro** mediante eventos exclusivos y promociones dirigidas.

Conclusión y Recomendaciones Estratégicas

- **Programas de Lealtad:** Recompensar a los clientes frecuentes con beneficios exclusivos.
- **Campañas Personalizadas:** Adaptar las promociones según género y edad.
- **Estrategias por Sucursal:** Ajustar las acciones de marketing según el perfil de clientes en cada ubicación.
- **Reactivación de Clientes:** Motivar a los clientes ocasionales con incentivos y descuentos dirigidos.

Con la implementación de estas **estrategias basadas en datos**, el Spa "Diego" podrá mejorar la retención de clientes, optimizar la demanda de servicios y aumentar la rentabilidad del negocio.