

DOCUMENTACIÓN DEL PROCESO ETL

Equipo: La computadora cuántica

Este documento describe el proceso de Extracción, Transformación y Carga (ETL) que alimenta el Data Warehouse (DW) representado con la base de datos HotelDW a partir de la base OLTP HotelOLTP ambas almacenadas en un contenedor de SQL Server en Docker. El flujo se implementa en Visual Studio 2022 mediante proyectos Integration Services (SSIS).

Arquitectura General

La solución ETL sigue una arquitectura modular de paquetes por entidad. Cada paquete consta de fases de Extracción, Transformación y Carga hacia tablas destino en el DW con el esquema dbo.Entidad.

Entidades Incluidas

- **Cliente:** Permite segmentar clientes, analizar nacionalidades, género y cohortes para campañas de marketing y lealtad.
- **Habitacion:** Soporta dashboards de ocupación, tarifa promedio y análisis de capacidad.
- **Reserva:** Base para métricas de booking window, duración de estadías y forecasting de demanda.
- **Pago:** Fuente principal de ingresos habilita análisis de revenue por método y fecha.
- **Empleado:** Facilita KPIs de desempeño y planeación de recursos humanos.
- **Tarea:** Mide productividad y eficiencia operativa del personal.
- **Mantenimiento:** Permite controlar costos, tiempos fuera de servicio y calidad de instalaciones.
- **Producto:** Catálogo necesario para análisis de costos unitarios y categorías de artículos.
- **Inventario:** Alimenta reportes de rotación de stock y previsiones de abastecimiento.

- Encuesta: Mide satisfacción del cliente y detecta áreas de mejora.
- Visita: Ayuda a estimar tráfico, estancia promedio y promociones efectivas.

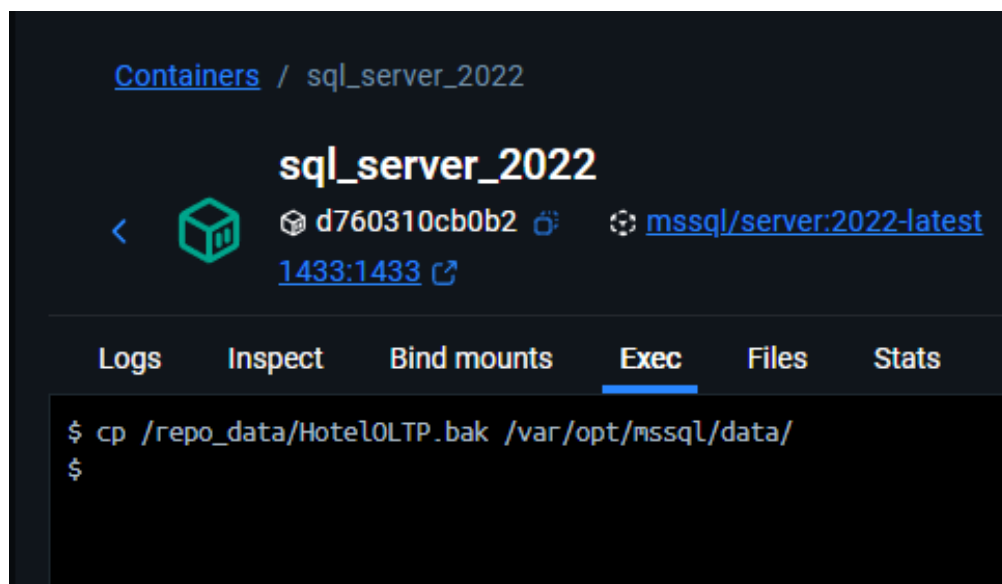
Requisitos previos

- Visual Studio 2022 con extensión SQL Server Integration Services Projects.
- Docker con container de SQL Server 2022
- Azure Data Studio como IDE para SQL Server
- Usuario con permisos de lectura sobre HotelOLTP y escritura sobre HotelDW.
- Base de datos origen: HotelOLTP (con la data cargada).
- Base de datos destino: HotelDW.
- Usuario con permisos de lectura en HotelOLTP y escritura en HotelDW.

Creación de las bases de datos

Para la creación de la base de datos HotelOLTP aplicamos el proceso utilizado para la importación de las bases de datos AdventureWorks importando el archivo de backup .bak dentro de la carpeta C:\sqlserver_data posteriormente ejecutando el siguiente comando en la consola (Exec) de Docker

```
cp /repo_data/HotelOLTP.bak /var/opt/mssql/data/
```



Posteriormente se restaura la base de datos desde Azure Data Studio

```
Run Cancel Disconnect Change Database: master Estimated Plan
1 RESTORE DATABASE HotelOLTP
2 FROM DISK = '/var/opt/mssql/data/HotelOLTP.bak'
3 WITH
4     MOVE 'PROYECTO_BI' TO '/var/opt/mssql/data/HotelOLTP.mdf',
5     MOVE 'PROYECTO_BI_Log' TO '/var/opt/mssql/data/HotelOLTP_log.ldf',
6     FILE = 1,
7     NOUNLOAD,
8     STATS = 5;
9
10 RESTORE FILELISTONLY FROM DISK = '/var/opt/mssql/data/HotelOLTP.bak';
11
```

Podemos corroborar que todas las tablas fueron creadas seleccionando la base de datos y ejecutando el siguiente comando.

```
Run Cancel Disconnect Change Database: HotelOLTP
1 SELECT TABLE_SCHEMA, TABLE_NAME
2 FROM INFORMATION_SCHEMA.TABLES
3
```

Results		Messages	
	TABLE_SCHEMA	TABLE_NAME	
1	Tarea	Asignacion	
2	Tarea	Horario	
3	Mantenimiento	Registro	
4	Mantenimiento	Costo	
5	Producto	Unidad	
6	Producto	Producto	
7	Producto	Stock	
8	Inventario	TipoMovimiento	
9	Inventario	Movimiento	
10	Encuesta	Fuente	
11	Encuesta	Satisfaccion	
12	Visita	Registro	
13	Cliente	Nacionalidad	
14	Cliente	Cliente	
15	Cliente	Telefono	
16	Habitacion	Tipo	
17	Habitacion	Estado	
18	Habitacion	Habitacion	
19	Habitacion	EstadoActual	
20	Reserva	Fuente	
21	Reserva	Estado	
22	Reserva	Reserva	
23	Reserva	EstadoActual	
24	Pago	Estado	
25	Pago	Transaccion	
26	Pago	EstadoActual	
27	Empleado	Puesto	
28	Empleado	Turno	
29	Empleado	Empleado	
30	Empleado	TurnoAsignado	

Creación de la base de datos DW

Vamos a crear las tablas dbo para guardar la información procedente de la base de datos OLTP ejecutamos el siguiente script tras crear la base de datos

```
Run Cancel Disconnect Change Database: HotelDW
1 CREATE DATABASE HotelDW
2 USE HotelDW
3 GO
4
5 -- Tabla dbo.Cliente
6 CREATE TABLE dbo.Cliente (
7     id_cliente INT PRIMARY KEY,
8     nombre NVARCHAR(100),
9     apellido NVARCHAR(100),
10    email NVARCHAR(150),
11    fecha_nacimiento DATE,
12    genero VARCHAR(20),
13    nacionalidad NVARCHAR(50),
14    telefonos NVARCHAR(MAX)
15 );
16
17 -- Tabla dbo.Habitacion
18 CREATE TABLE dbo.Habitacion (
19     id_habitacion INT PRIMARY KEY,
20     numero NVARCHAR(10),
21     tipo_habitacion NVARCHAR(50),
22     capacidad_base INT,
23     capacidad INT,
24     precio_base DECIMAL(10,2),
25     estado_actual NVARCHAR(20),
26     fecha_estado DATETIME
27 ); -- .....
```

Proceso de creación de la Solución SSIS

- 1- Crear el proyecto SSIS: •Abrir Visual Studio 2022 y crear un nuevo proyecto tipo Integration Services Project y nombrado como "Hotel_ETL".

Integration Services Project

Nombre del proyecto

Hotel_ETL

Ubicación

C:\Users\Diego\Documents\GitHub\DiegoCastro_00117322_ETL

Nombre de la solución ⓘ

Hotel_ETL

2- Configurar conexiones: En la pestaña Connection Managers agregar dos conexiones tipo OLE DB.

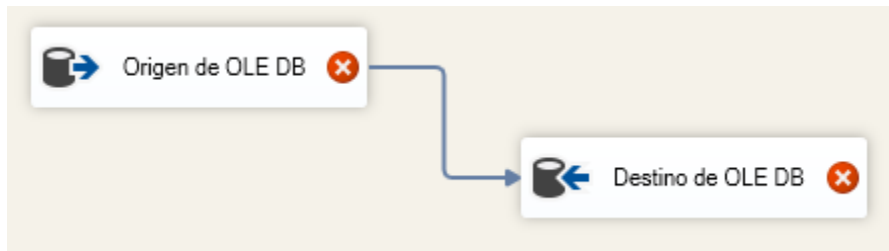
- Una conexión apuntando al origen de datos HotelOLTP.

The screenshot shows the 'OLE DB Provider' configuration window. The 'Proveedor' dropdown at the top is set to 'OLE DB nativo\Microsoft OLE DB Driver for SQL Server'. On the left sidebar, the 'Conexión' icon is selected. The main area is titled 'OLE DB Provider:' and shows 'Microsoft OLE DB Driver for SQL Server' selected in the dropdown, with a 'Data Links...' button to its right. Below this, the 'Enter a server or file name' section has 'Server or file name:' set to '192.168.0.8' and an empty 'Location:' field. The 'Log on to the server' section has two radio buttons: 'Use Windows NT Integrated Security' (unselected) and 'Use a specific user name and password:' (selected). Under the selected option, 'User name:' is 'sa' and 'Password:' is masked with dots. There are checkboxes for 'Blank password' and 'Allow saving password', both of which are unchecked. At the bottom, the 'Initial catalog:' dropdown is set to 'HotelOLTP'.

- Una conexión apuntando al destino de datos HotelDW.

This screenshot is identical in layout and most settings to the one above, but the 'Initial catalog:' dropdown at the bottom is set to 'HotelDW' instead of 'HotelOLTP'. All other fields, including the server name '192.168.0.8' and the 'sa' user name, remain the same.

3. Conexión entre el origen y el destino: Debemos conectar el origen con el destino para que desaparezca el error.



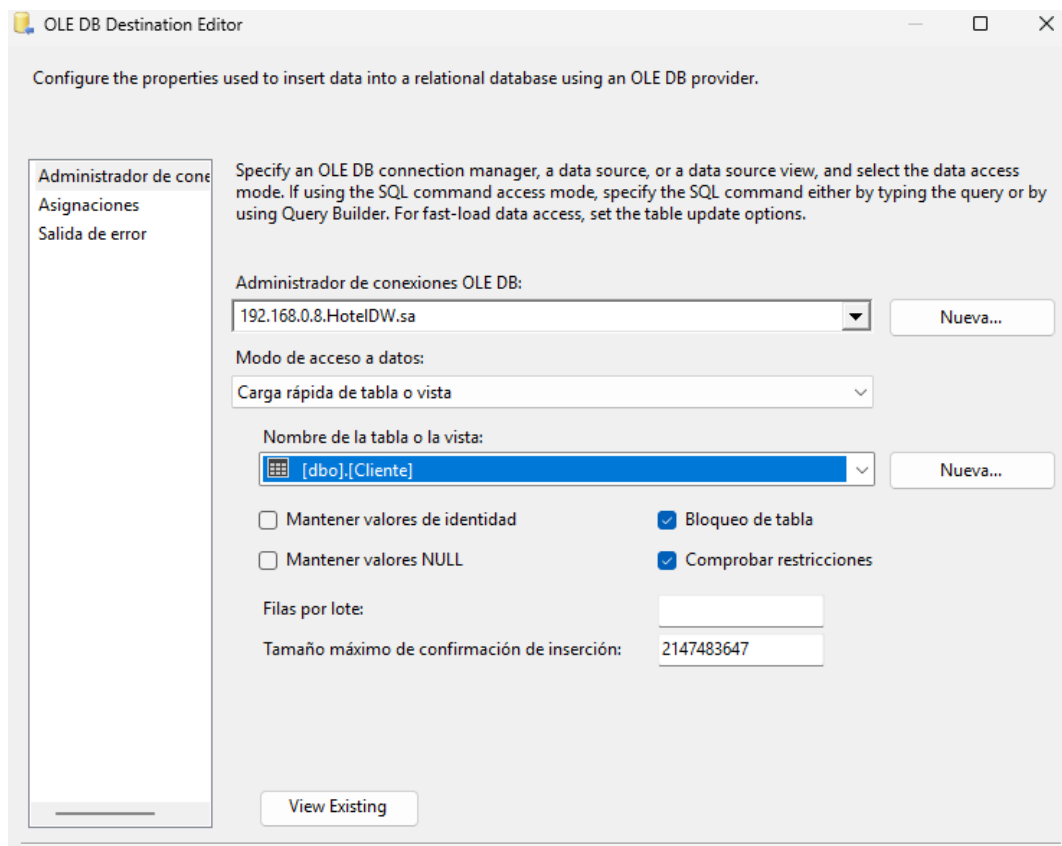
4. Configuración del ETL: Ahora tenemos que indicar que data es la que vamos a enviar desde el origen utilizando la consulta SQL.
- OLE DB Source: Ejecutar la consulta SQL para obtener datos desde HotelOLTP en este caso información de los clientes.

The screenshot shows the 'OLE DB Source Editor' window. The title bar reads 'OLE DB Source Editor'. The main text says: 'Configure the properties used by a data flow to obtain data from any OLE DB provider.' On the left is a sidebar with 'Administrador de conexiones', 'Columnas', and 'Salida de error'. The main area has a description: 'Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.' Below this, 'Administrador de conexiones OLE DB:' has a dropdown menu showing '192.168.0.8.HotelOLTP.sa' and a 'Nueva...' button. 'Modo de acceso a datos:' has a dropdown menu showing 'Comando SQL'. 'Texto de comando SQL:' has a text area containing a SQL query:

```
SELECT
  c.id_cliente,
  c.nombre,
  c.apellido,
  c.email,
  c.fecha_nacimiento,
  CASE c.genero
    WHEN 'M' THEN 'Masculino'
    WHEN 'F' THEN 'Femenino'
    WHEN 'O' THEN 'No respondió'
    ELSE 'Desconocido'
  END AS genero,
  n.nombre AS nacionalidad,
  STRING_AGG(t.telefono, ',' ) AS telefonos
```

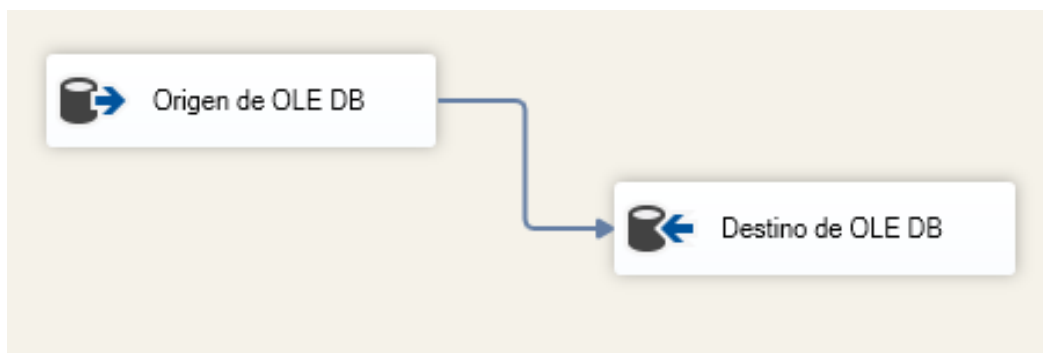
 To the right of the text area are buttons: 'Parámetros...', 'Generar consulta...', 'Examinar...', and 'Analizar consulta'. At the bottom left of the main area is a 'Vista previa...' button. At the bottom right are 'Aceptar', 'Cancelar', and 'Ayuda' buttons.

- OLE DB Destination: Establecer la conexión al DW y seleccionar la tabla destino dbo.Cliente donde se almacenara el resultado de la consulta.



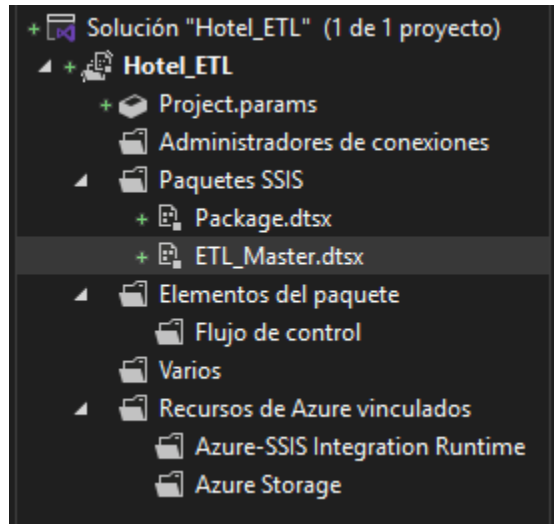
5. Validar los paquetes:

- Ahora ya podemos ejecutar el proyecto de SSIS para realizar el proceso del ETL y comprobar si es exitoso pero como manejamos diferentes consultas realizaremos el siguiente paso para controlar varios subprocesos a la vez.

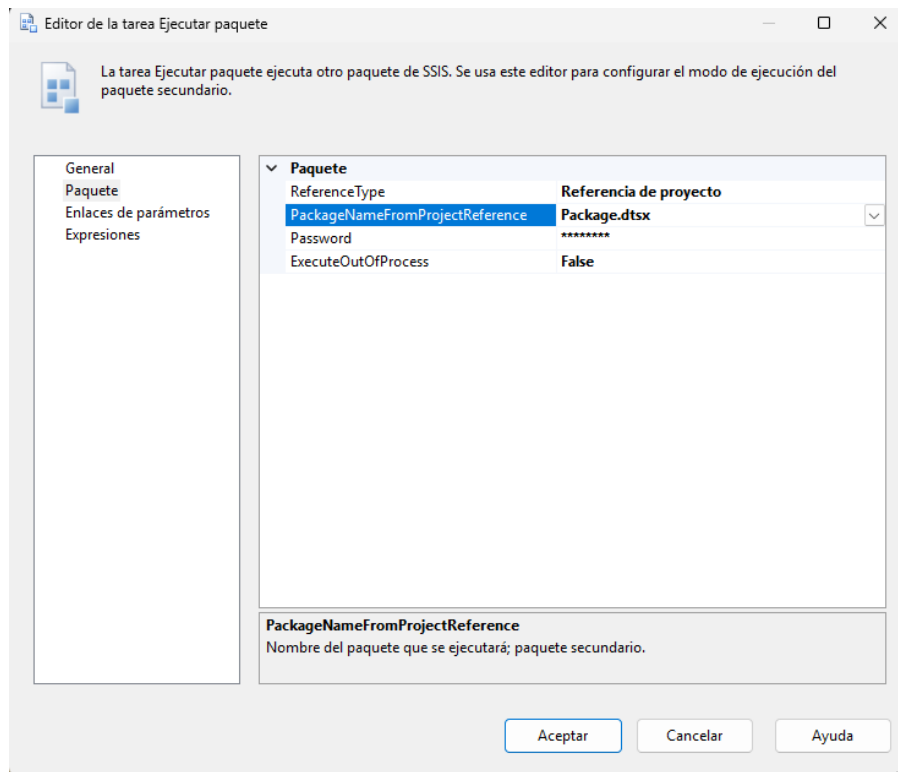


6. Crear Paquete Maestro:

- Agregar un paquete maestro (Master_ETL.dtsx).

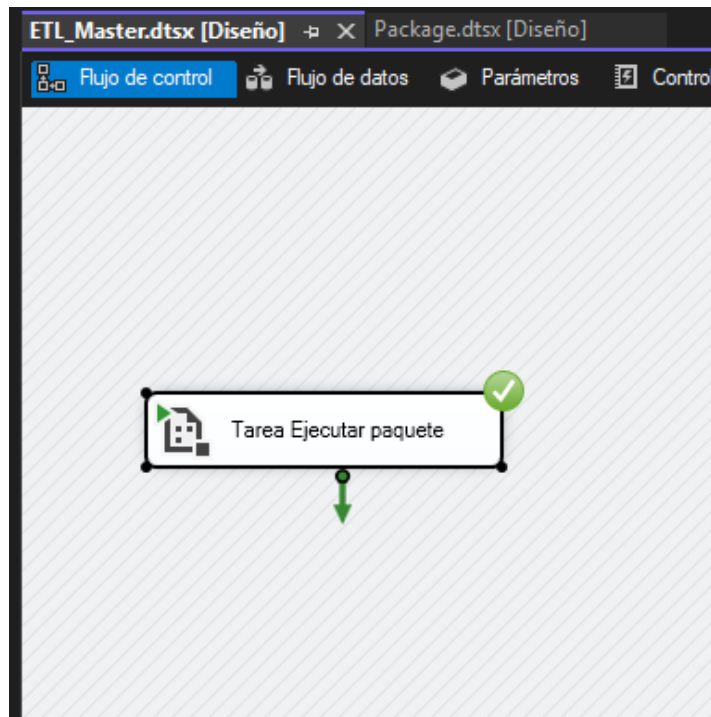


- Utilizar Execute Package Task para ejecutar secuencialmente cada paquete ETL creado anteriormente.

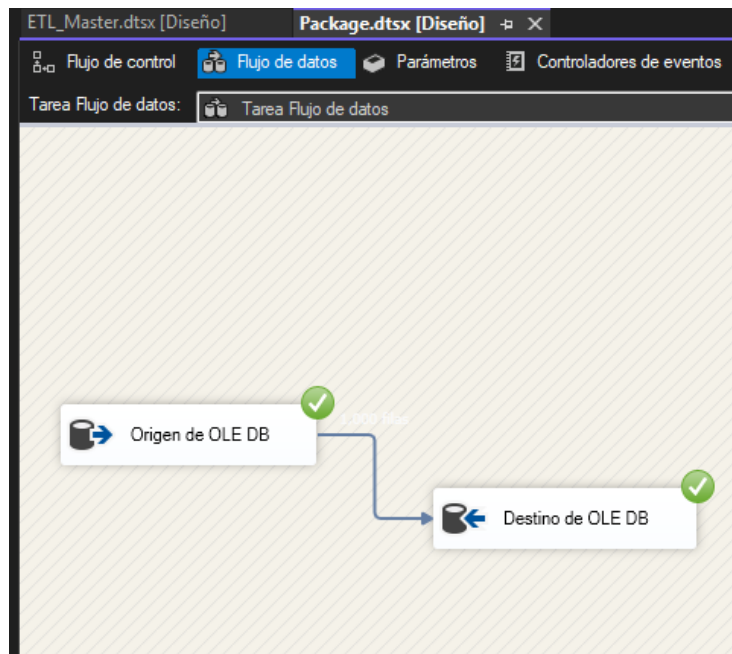


7. Ejecución Completa del proceso ETL:

- Ejecutar el paquete maestro desde Visual Studio.



- Lo que hace que automáticamente se ejecute el proceso de ETL creado anteriormente



8. Comprobación de la transferencia de datos en Azure Data Studio:

- Realizar una consulta simple para comprobar la carga exitosa de los datos desde HotelOLTP.

17 SELECT * FROM dbo.Cliente

Results Messages

	id_cliente	nombre	apellido	email	fecha_nacimiento	genero	nacionalidad	telefonos
1	1	Ardine	Strickland	astrickland0@cnbc.com	2001-06-07	Masculino	Reino Unido	9088220450
2	2	Clo	Foucar	cfoucar1@nhs.uk	1993-10-27	Femenino	Colombia	7626326948
3	3	Enriqueta	Addeycott	eaddeycott2@narod.ru	1996-07-15	Masculino	República Dominicana	9463811331
4	4	Nananne	Orrin	norrin3@discuz.net	1997-12-09	No respondió	Reino Unido	1122129921
5	5	Koressa	Cuthbertson	kcuthbertson4@e-rech...	1965-10-04	No respondió	Portugal	8753876276
6	6	Jeanne	Rochell	jrochell15@scribd.com	2001-11-20	Femenino	Guatemala	2443230776
7	7	Linet	Livesley	llivesley6@addthis.c...	1971-08-20	Masculino	El Salvador	6858980985
8	8	Kellyann	Eakin	keakin7@tiny.cc	1980-05-06	Femenino	Sudáfrica	7604483728
9	9	Lewiss	Medhurst	lmedhurst8@google.cn	1994-01-26	No respondió	Costa Rica	5905440983
10	10	Leroy	Jezard	ljezard9@statcounter...	1980-07-20	No respondió	Ecuador	6657394274
11	11	Wayland	Stenson	wstenson@4shared.com	1992-03-25	No respondió	Cuba	5713458048
12	12	Warde	Renault	wrenaultb@google.ru	1966-08-02	Masculino	República Dominicana	4118053072
13	13	Tina	Boodell	tboodellc@utexas.edu	1990-06-09	Masculino	Jamaica	5706284073
14	14	Brynna	Wurst	bwurst@forbes.com	1970-04-22	No respondió	Portugal	7263830003
15	15	Lorelle	Poser	lposer@deli.com	1975-01-08	Femenino	Guatemala	2176940804
16	16	Aurore	Boulds	abouldsf@usnews.com	1966-05-10	No respondió	Corea del Sur	3407868173
17	17	Major	Sabey	msabey@wp.com	1968-11-15	No respondió	España	4747591796
18	18	Aurie	Stockford	astockfordh@umn.edu	1968-10-31	Femenino	Corea del Sur	9426054576
19	19	Tiffi	Warde	twardei@flickr.com	2001-11-29	No respondió	España	4794547673
20	20	Hannis	Lassey	hlasseyj@cbsnews.com	1963-01-24	Femenino	China	3376218087
21	21	Giralda	Houseley	ghouseleyk@hao123.com	1995-04-17	No respondió	Venezuela	9822362436
22	22	Carilyn	Hilary	chilaryl@github.com	1992-11-19	No respondió	Portugal	6983448871
23	23	Jenica	MacAleese	jmacaleese@sohu.com	1988-03-30	Masculino	Uruguay	9106502489
24	24	Fanni	Straun	fstraunn@imgur.com	1986-02-04	No respondió	Canadá	1223768373

Ahora realizamos desde el paso 2 para realizar el ETL de cada una de las entidades que necesitemos en nuestro DW.

Conclusión

El proceso ETL descrito, implementado mediante Visual Studio y SSIS, permite una transferencia eficiente y confiable de los datos operativos desde la base transaccional HotelOLTP hacia el HotelDW alojado en SQL Server con Docker. La validación a través de SSIS en Visual Studio asegura la correcta carga y consistencia de los datos, estableciendo una base robusta y confiable para los análisis estratégicos, dashboards interactivos y la generación continua de reportes de inteligencia de negocios que apoyan decisiones informadas en la gestión del hotel.