****

**SALESIANOS UNIVERSIDAD DON BOSCO**

**FACULTAD DE INGENIERÍA**

**ESCUELA DE COMPUTACIÓN**

**CICLO 02-2020**

**“SEGUNDO DESAFÍO PRÁCTICO”**

**GRUPO DE LABORATORIO:**

01

**CARRERA:**

INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN.

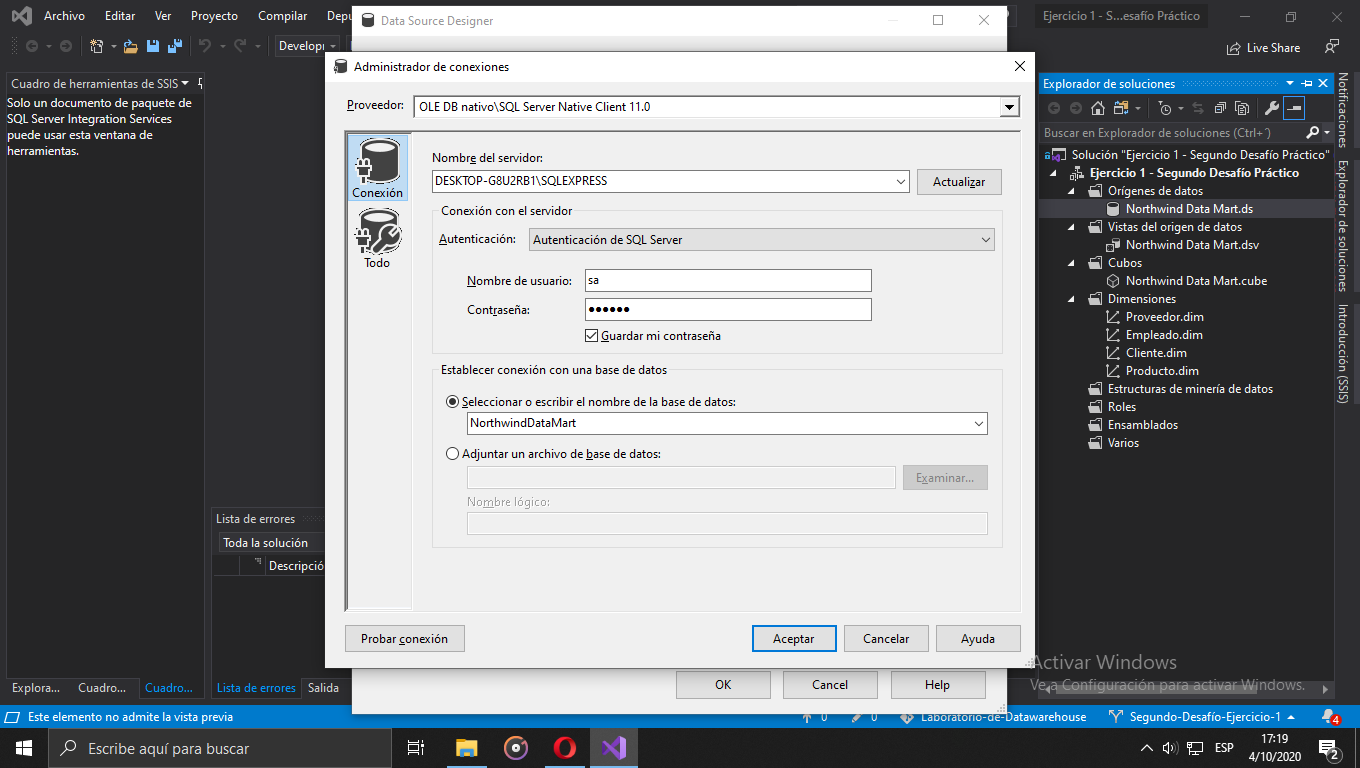
**PRESENTADO POR:**

|  |  |  |
| --- | --- | --- |
| **Carnet** | **Nombre** | **Apellido** |
| VC190544 | Francisco José | Valle Cornejo |
| AV190086 | César Adilson | Ayala Vásquez |

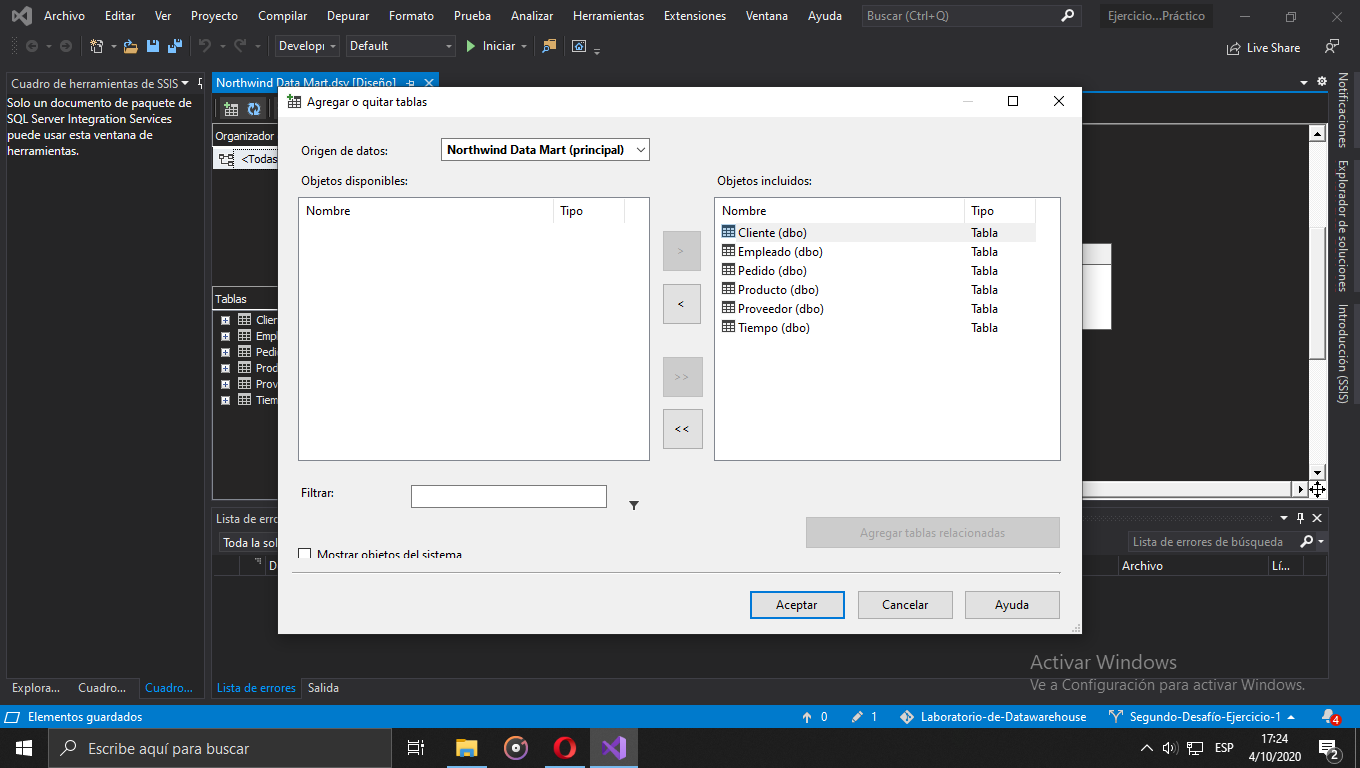
**DOCENTE:**

Alexander Alberto Sigüenza Campos

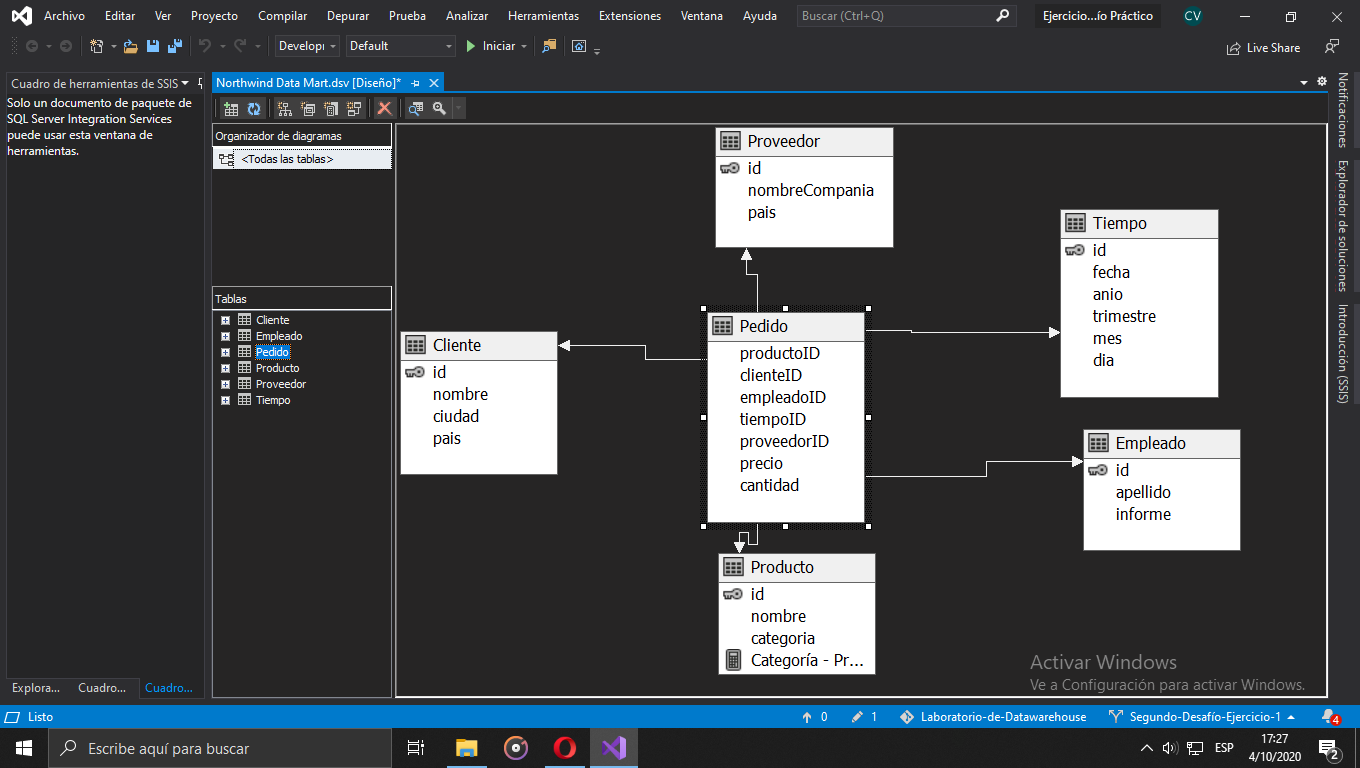
**EJERCICIO 1  
Paso 1:** Una vez iniciado Visual Studio, vamos a definir el origen de datos para nuestro procedimiento. En este caso, el origen será la base de datos Northwind Data Mart, la configuración quedará de la siguiente manera:



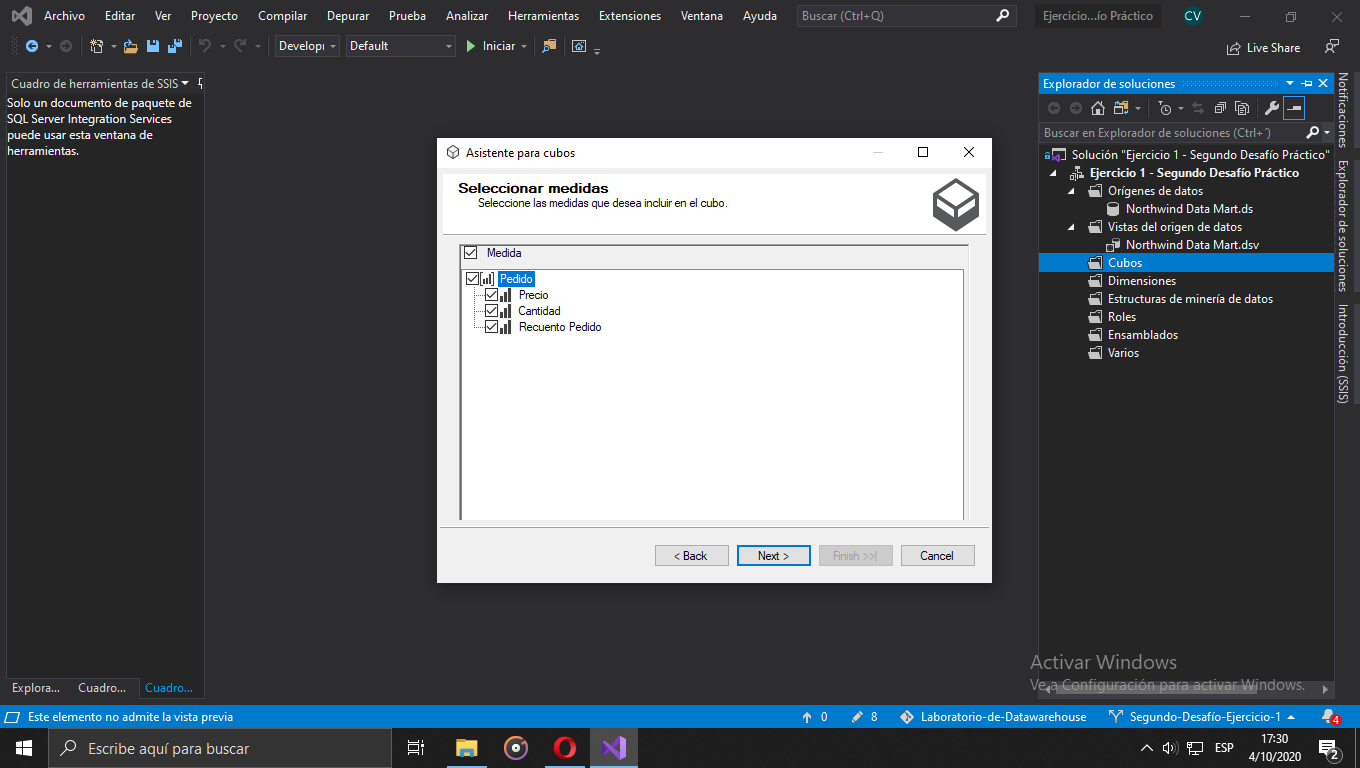
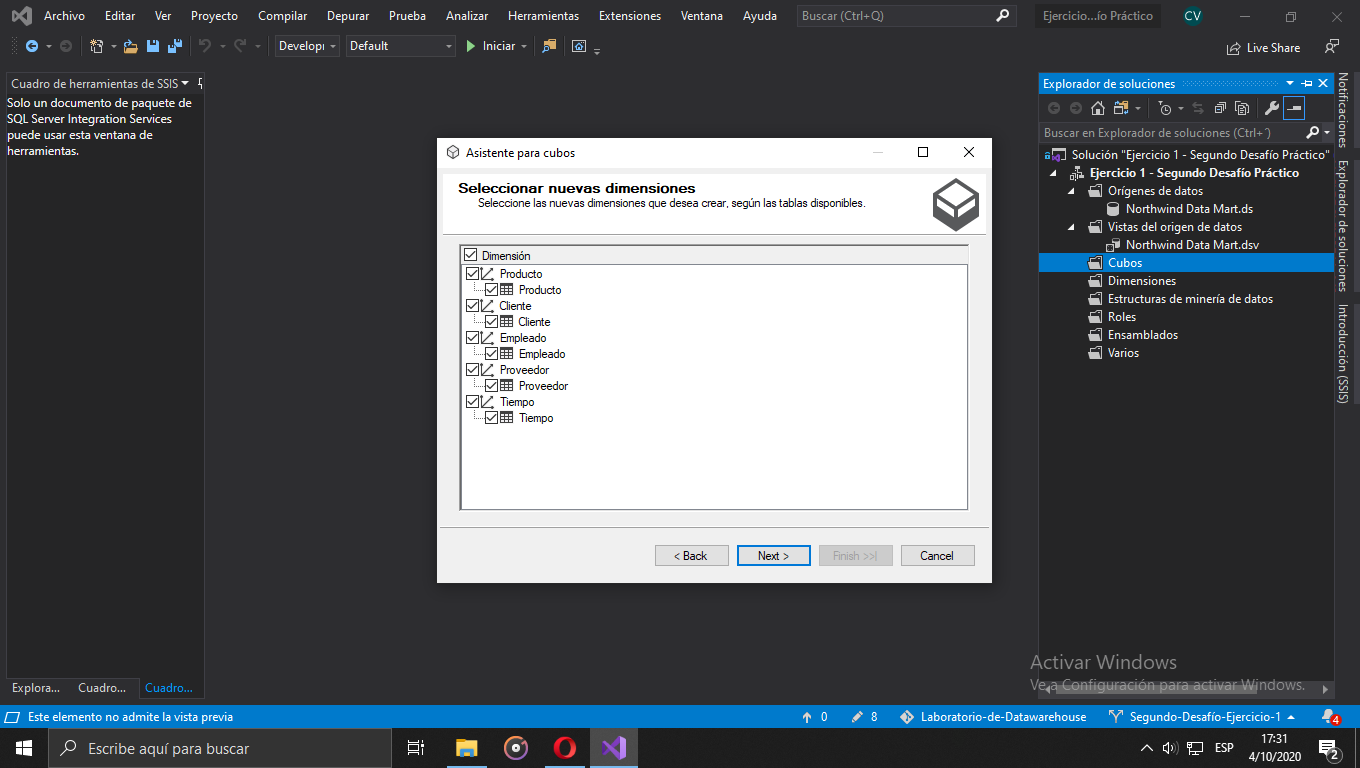
**Paso 2:** Definir una vista que incluya las tablas que deseamos analizar, en nuestro caso, escogeremos todas las tablas para un mayor detalle al diseñar e implementar el cubo. Nos quedará de la siguiente manera:



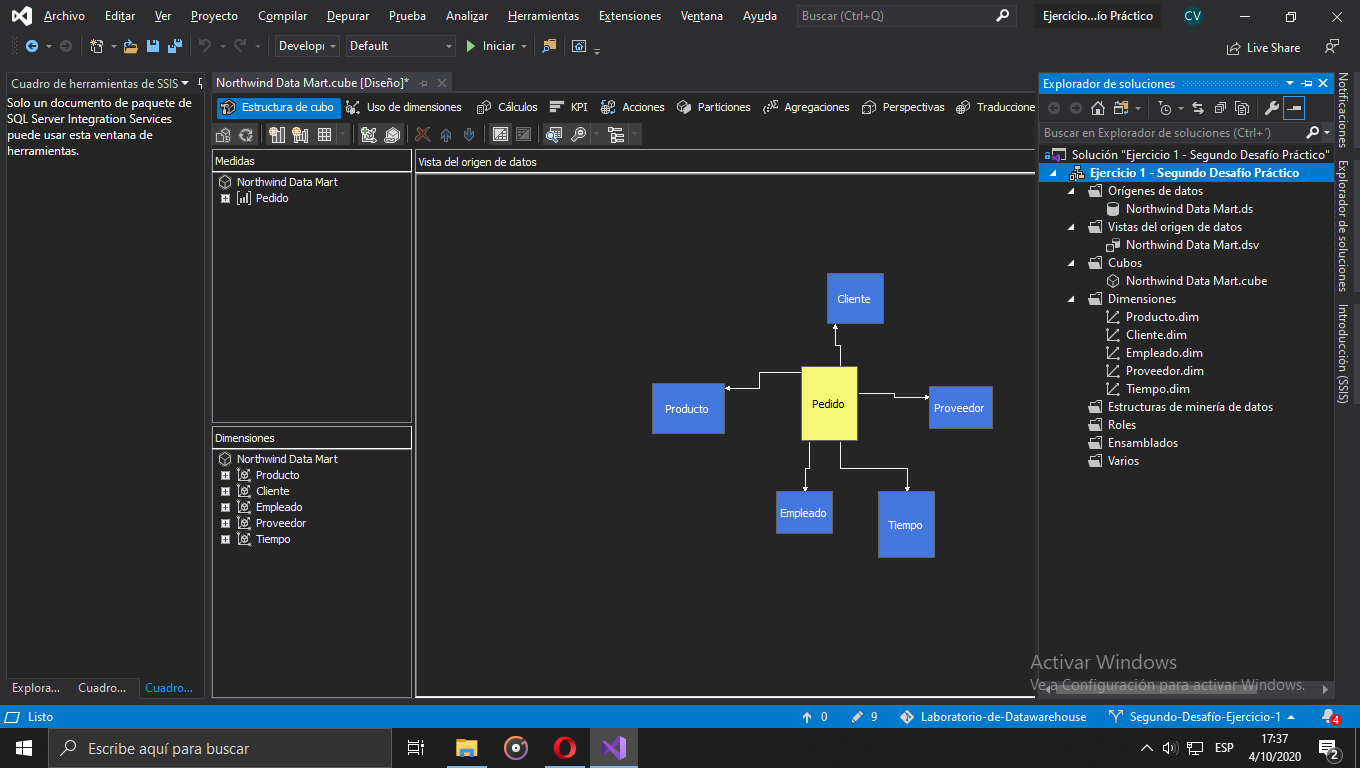
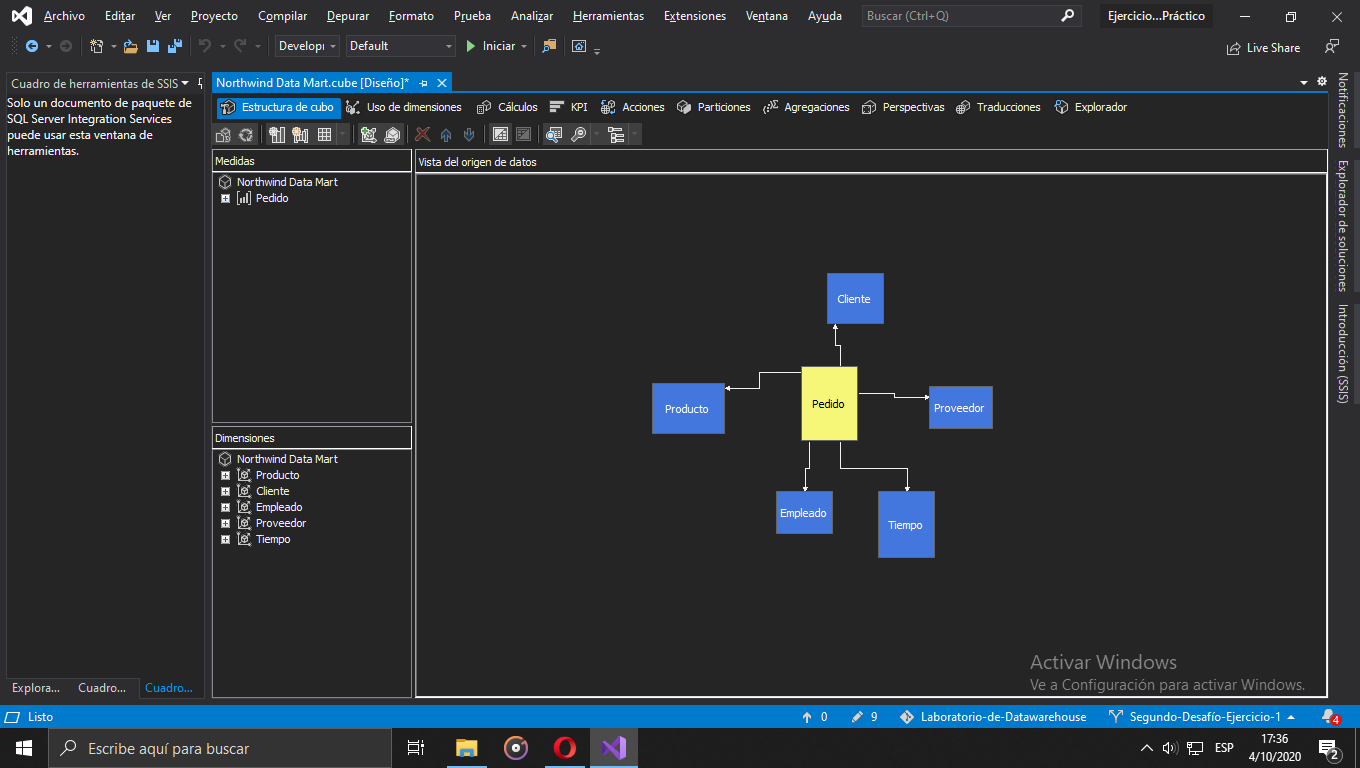
La vista puede puede ser visualizada de esta manera:



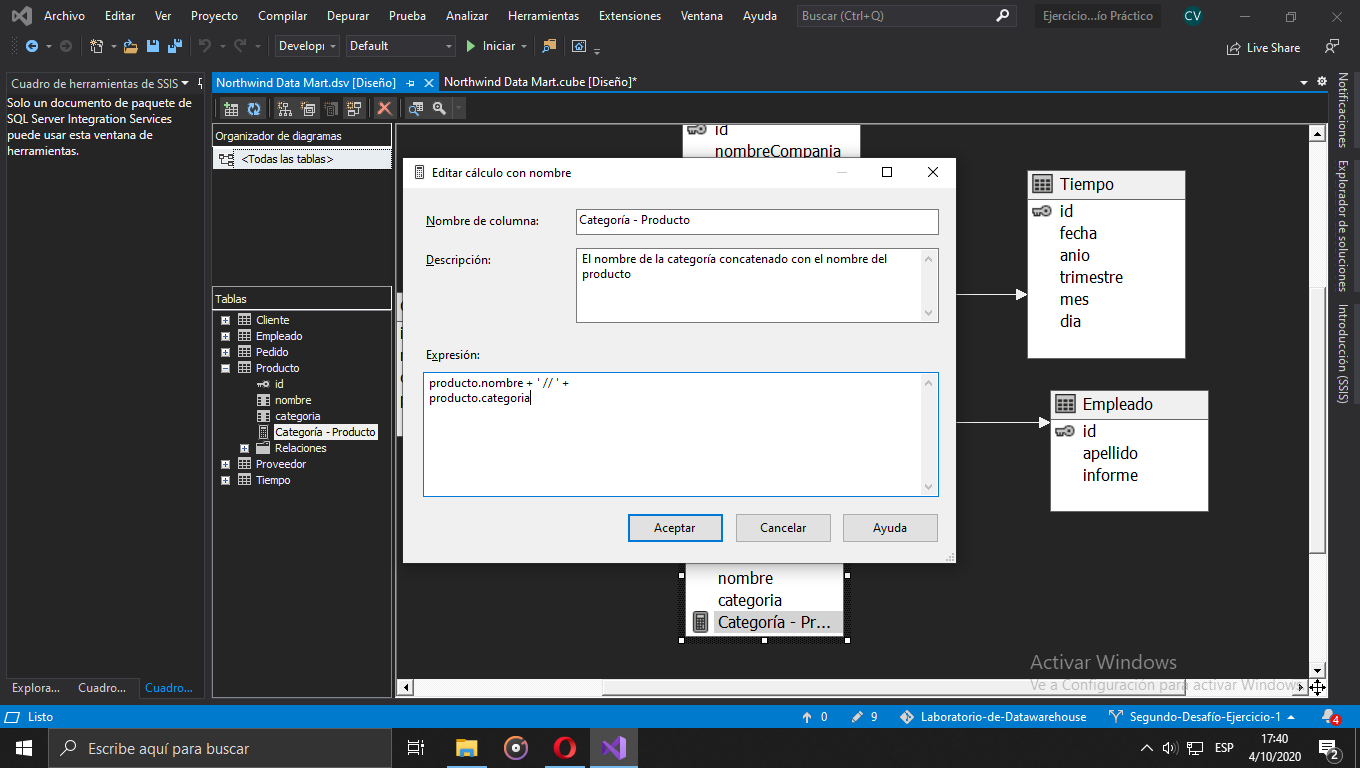
**Paso 3:** Crear un nuevo Cubo, y definiremos como tabla hechos a la tabla “pedidos”, posteriormente definir las dimensiones: Proveedor, tiempo, empleado, producto y cliente.

**Paso 4:** El nombre del cubo será el que el programa nos da por defecto, luego de construir el cubo podemos visualizarlo mejor con un zoom del 50%.  
Quedando nuestra tabla hechos con color amarillo y todas las dimensiones de color azul.



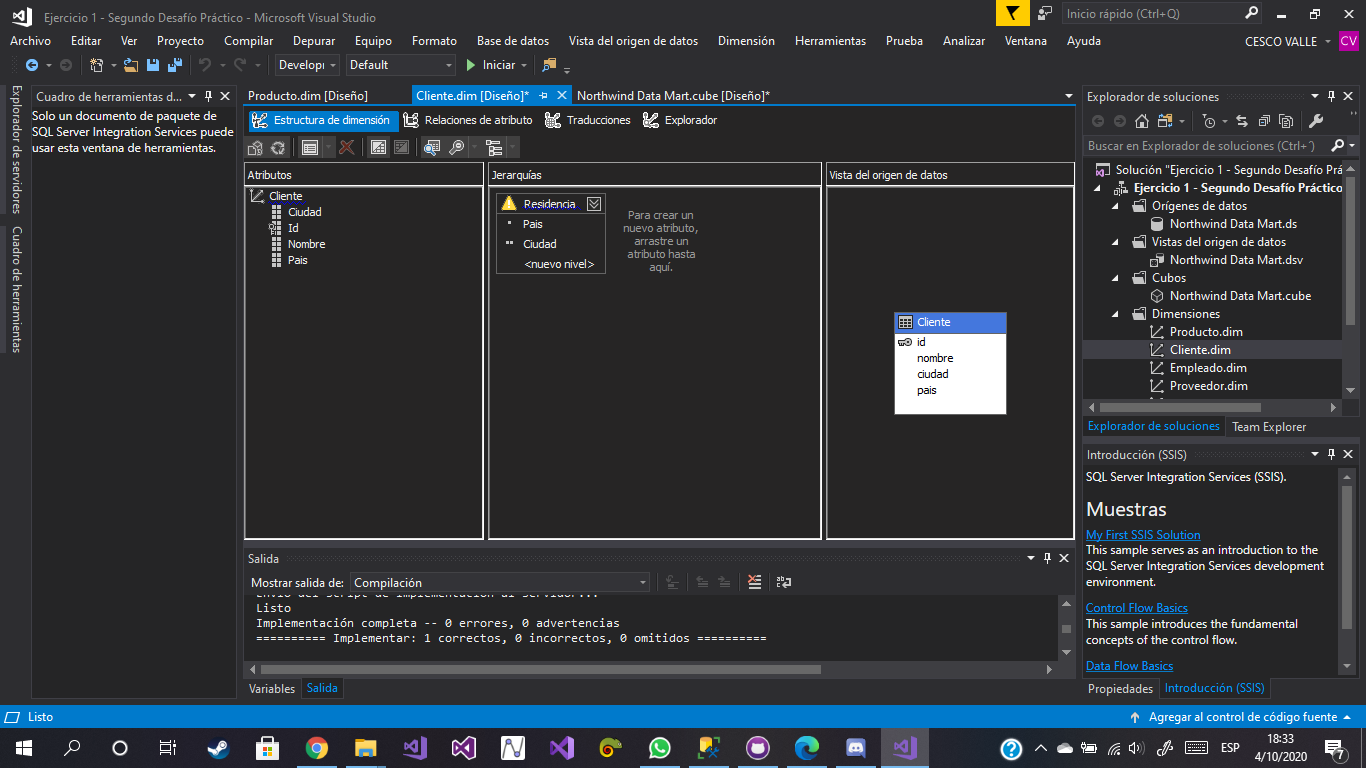
**Paso 5:** Antes de la implementación, crearemos un campo nuevo Nombre/Categoría de la tabla productos, para ello, nos dirigimos a la vista creada en el paso 2, click derecho en la tabla producto y “Nuevo cálculo con nombre…”. Configurar los campos como se muestra en la imagen:

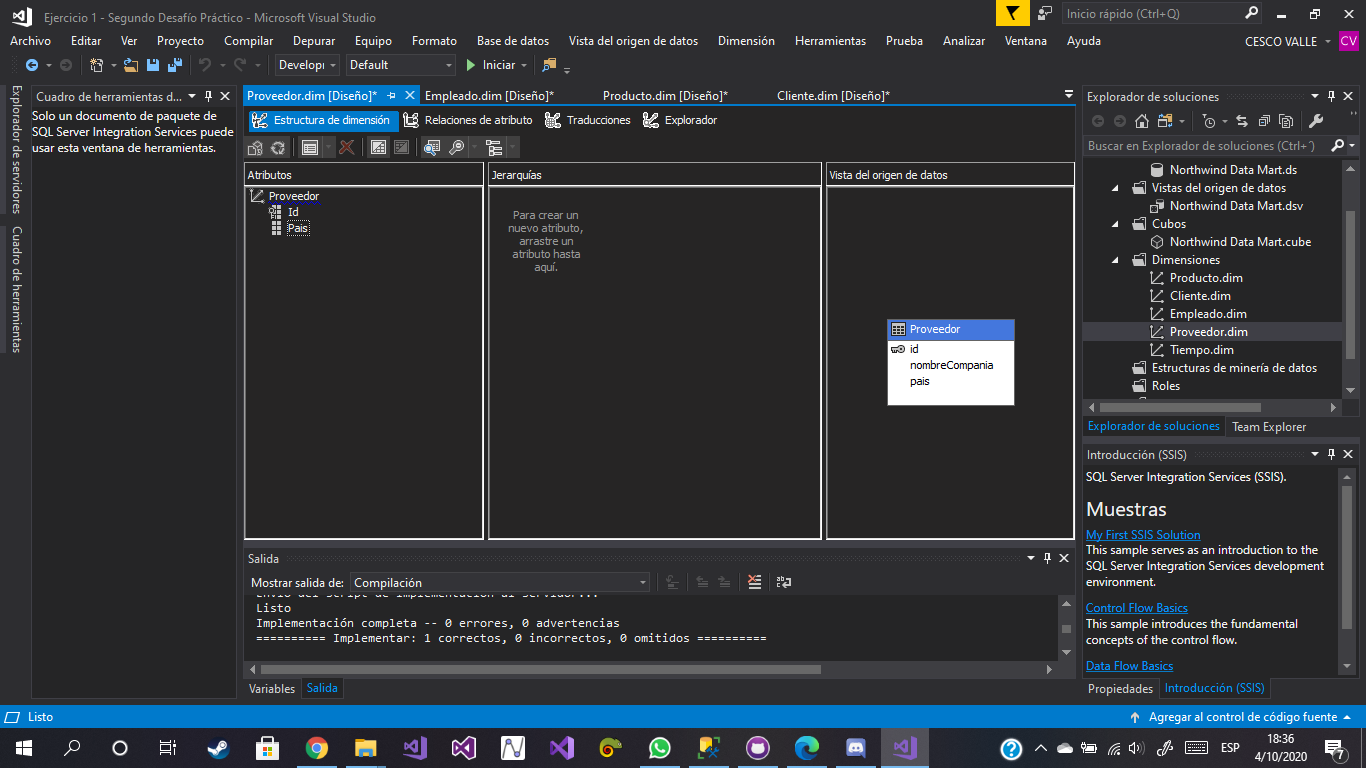
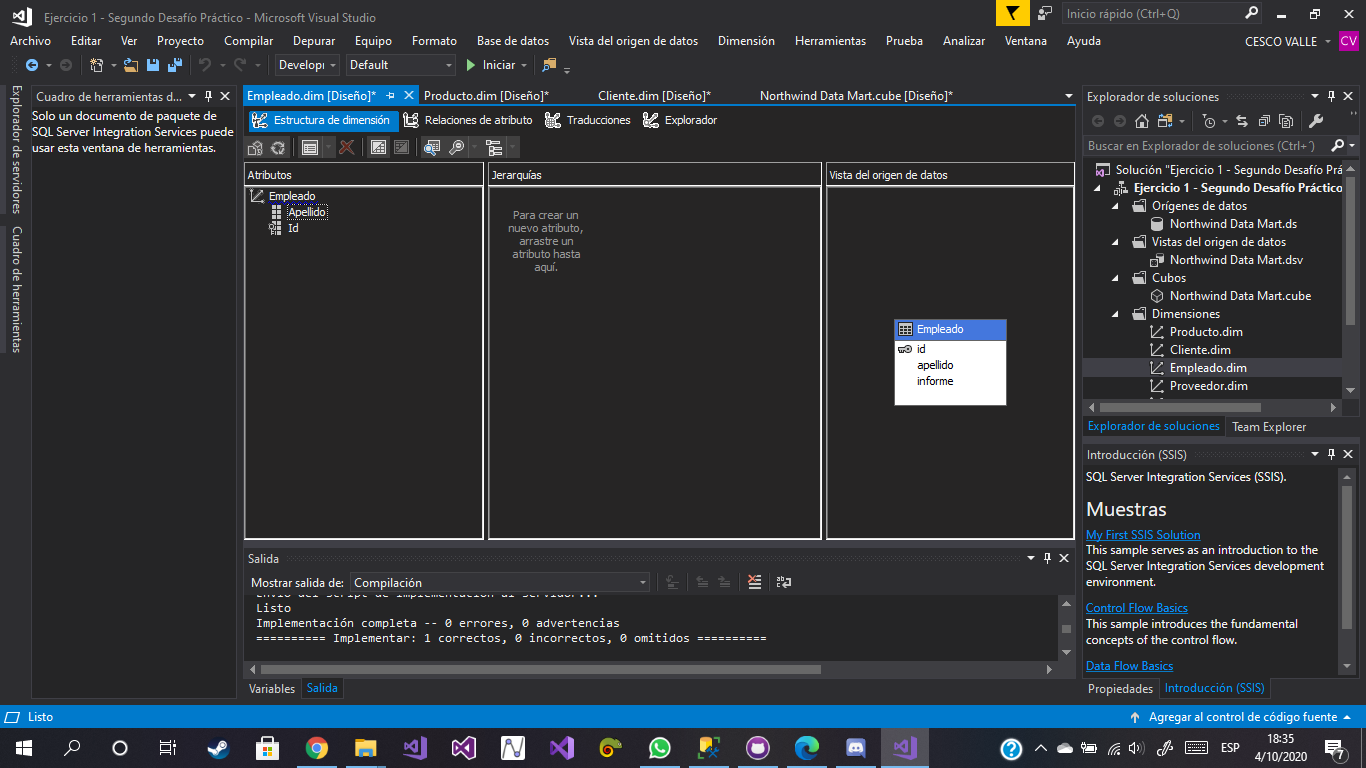
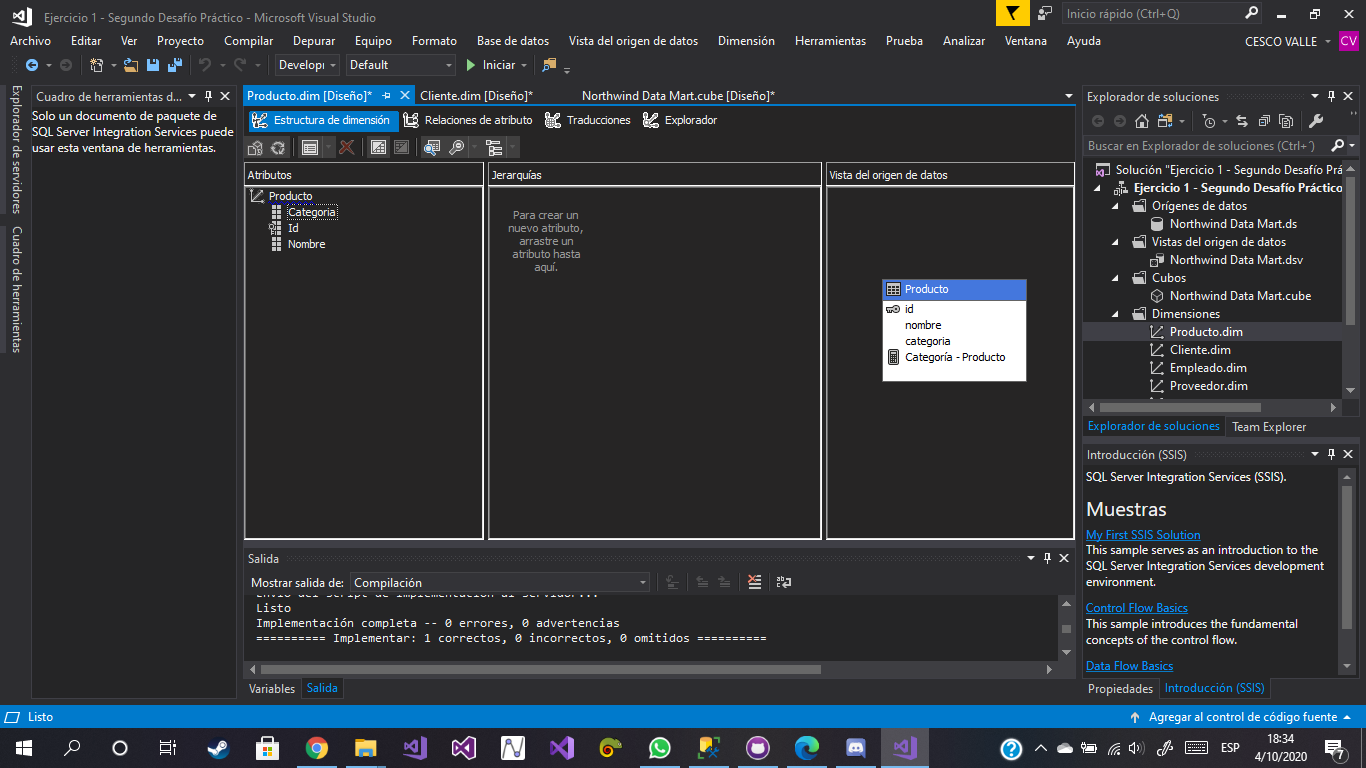


La expresión para concatenar los campos es la siguiente:

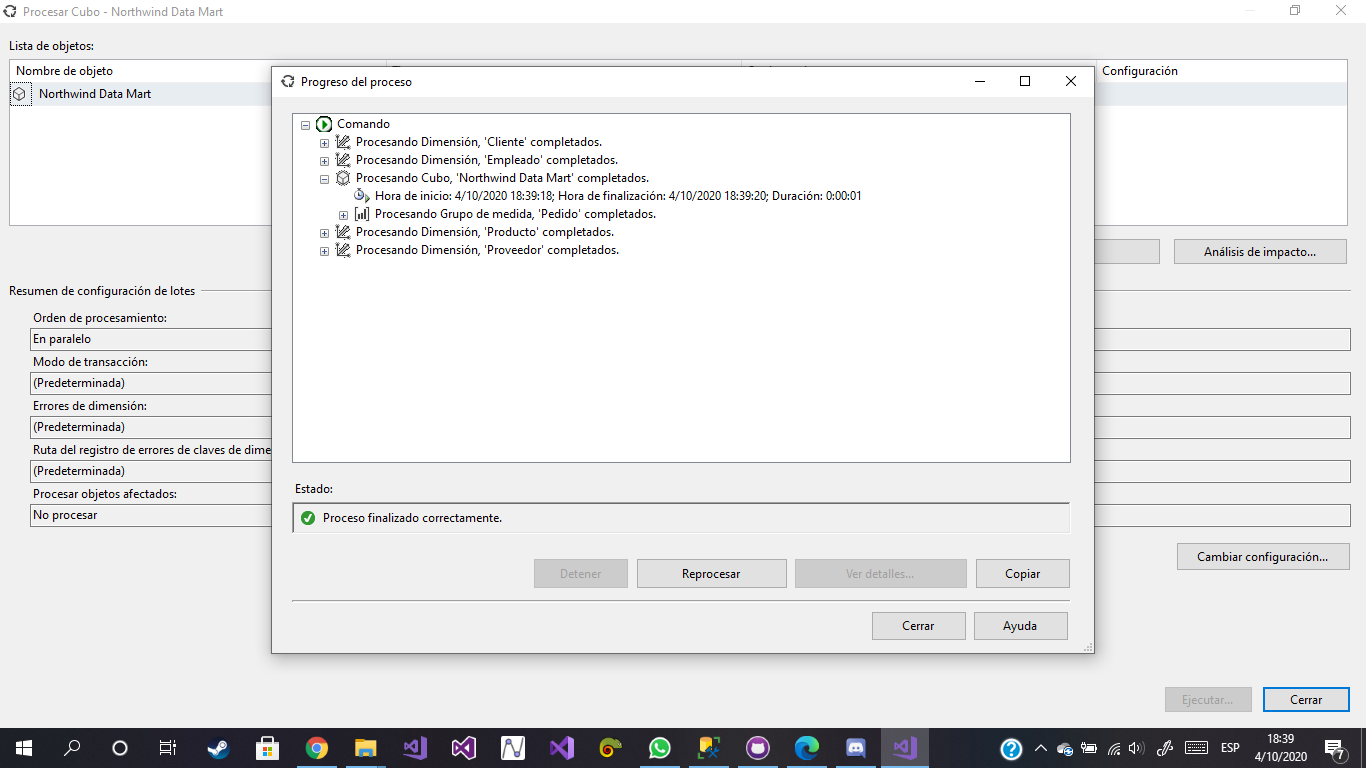
***producto.nombre + ' // ' + producto.categoria***

**Paso 6:** Cambiar las dimensiones para que las consultas se guien más allá de las llaves primarias

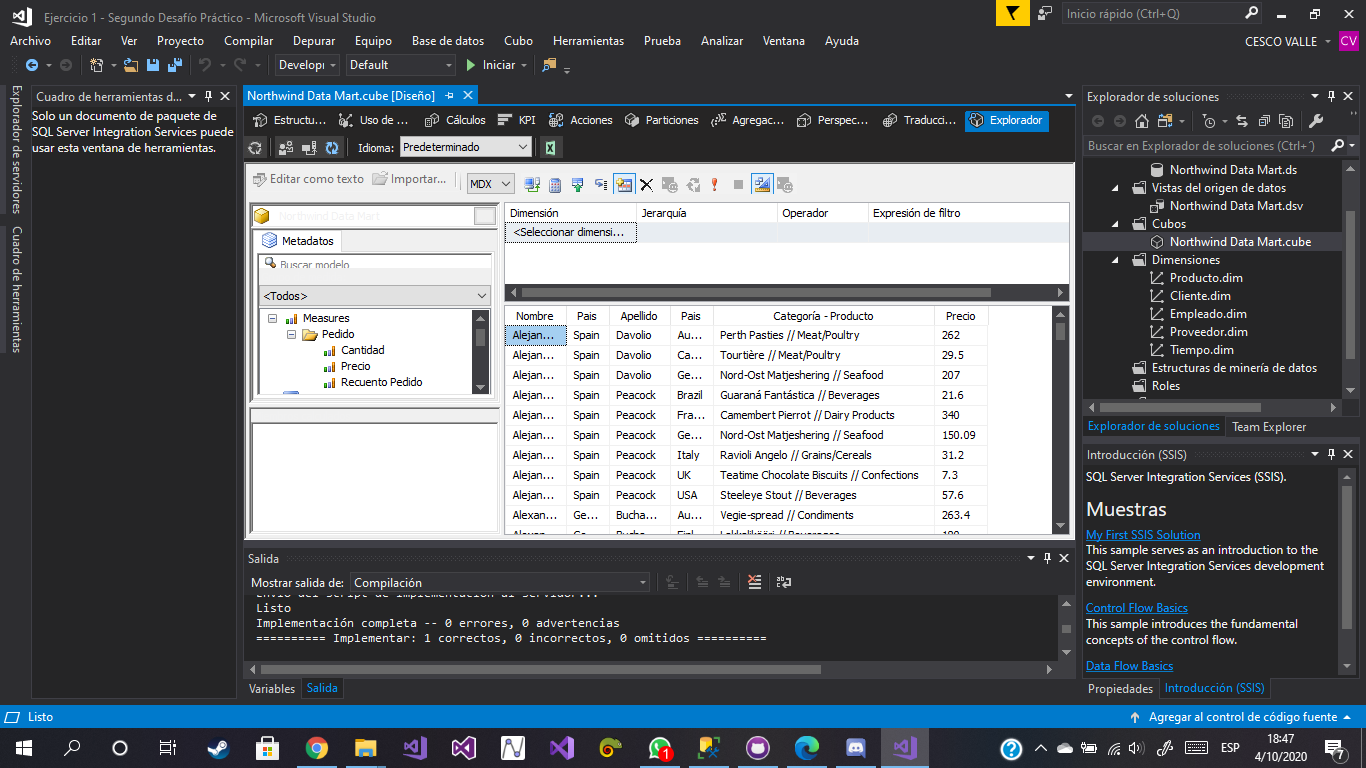




**Paso 7:** Procesar el cubo y examinar el cubo aplicando la consulta pedida

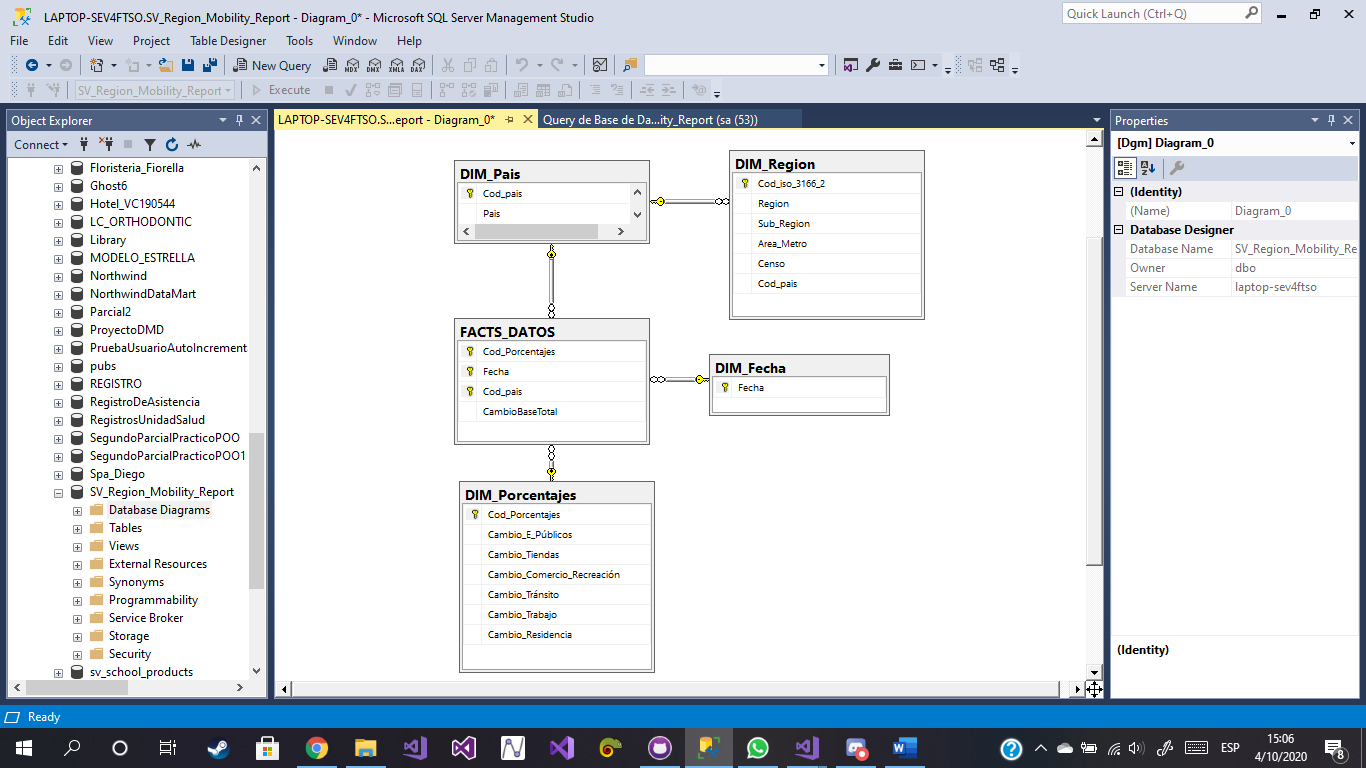


**Paso 8:** Ingresar la consulta pedida en el desafío un cubo en donde pueda visualizar nombre y país del cliente, apellido del empleado, país del proveedor, y nombre y categoría del producto. Estos dos últimos concatenados en una misma columna

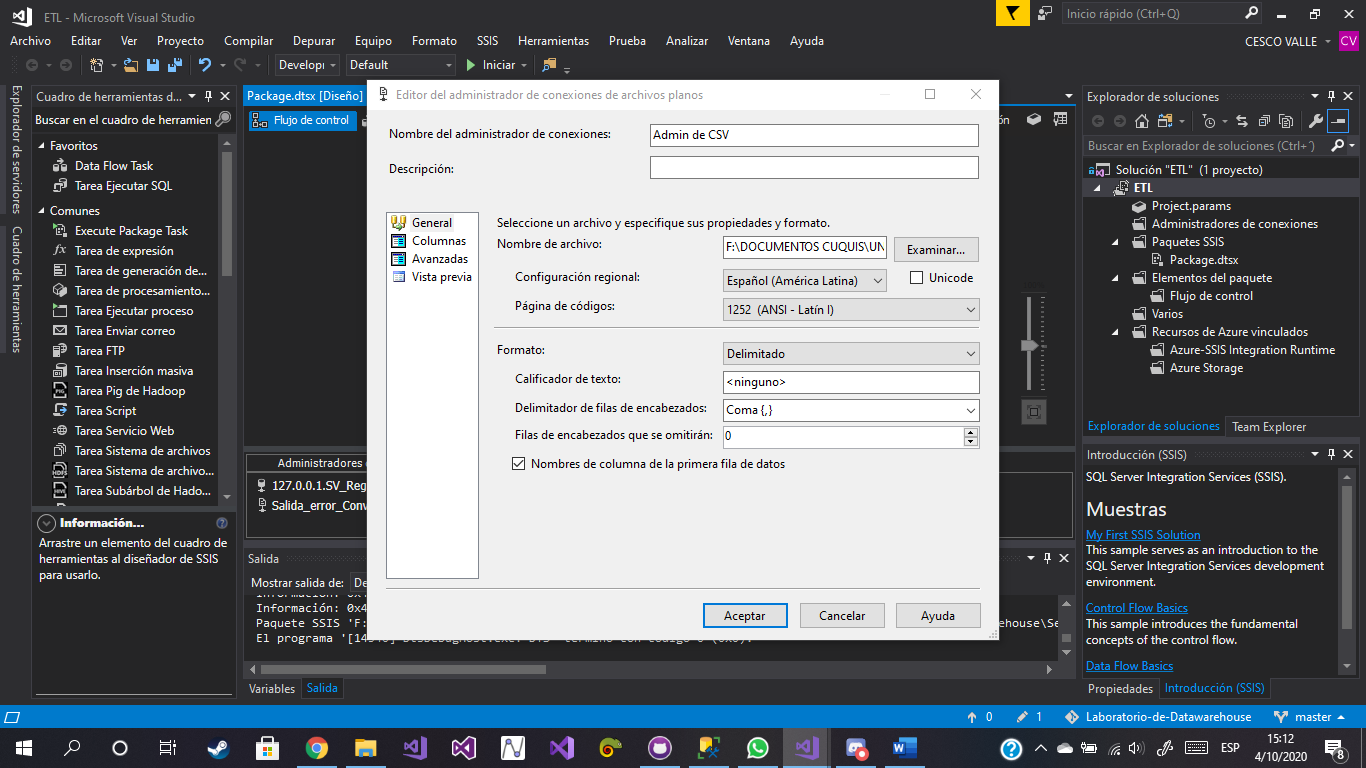


**EJERCICIO 2**

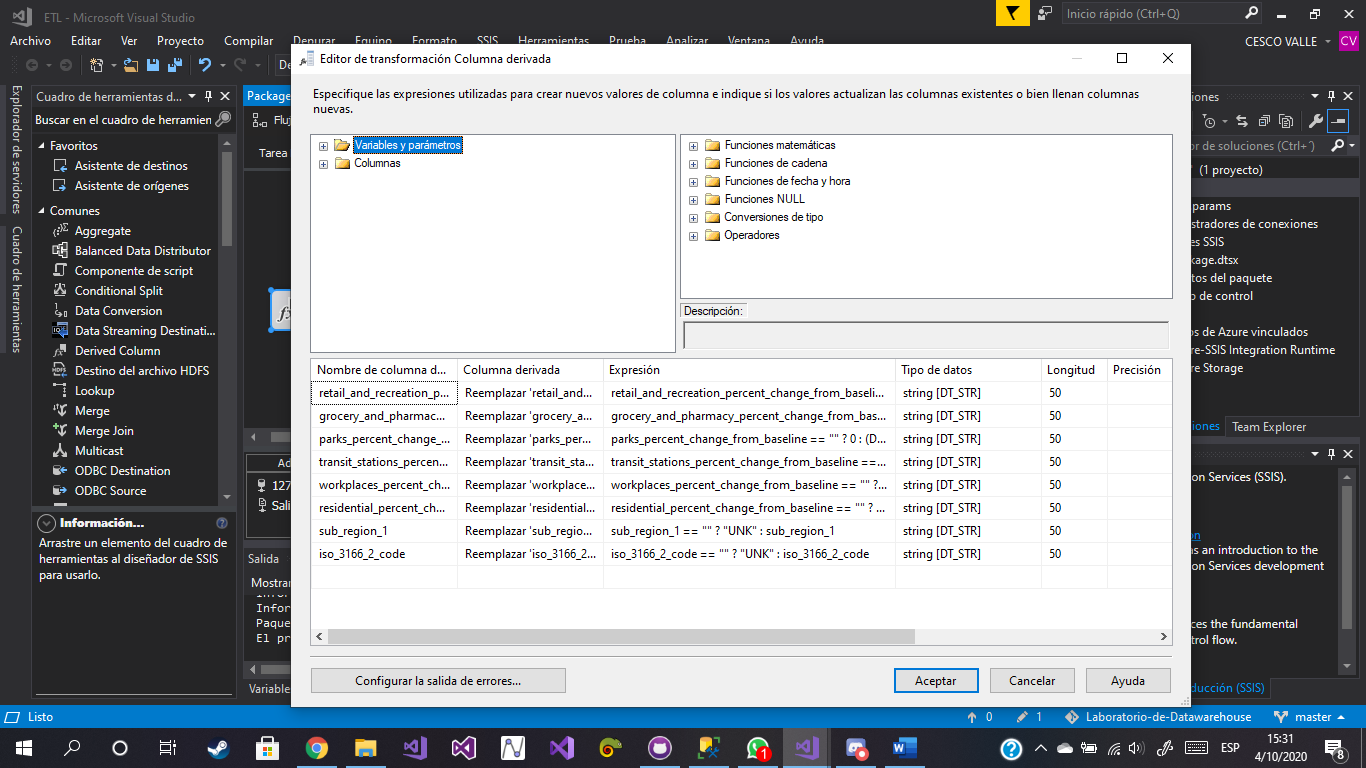
**Paso 1:** Tras analizar los documentos separados con comas brindados en los recursos se ha definido usar un modelo de copo de nieve para su base de datos, con 3 tablas dimensiones principales y una relaciona a la tabla País



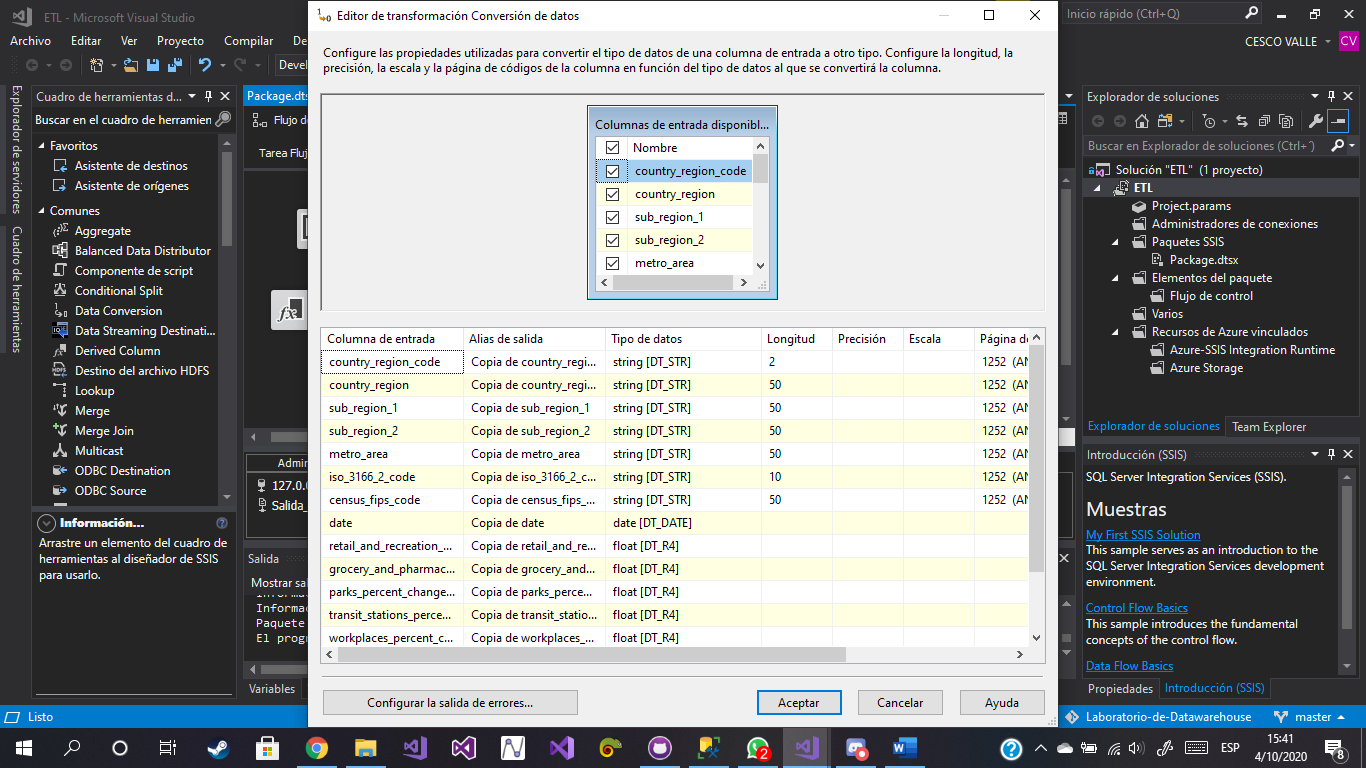
**Paso 2:** Se comienza con la creación del ETL, se usará un modelo de dos DataFlow para asegurar que las tablas Dimensiones ya tengan sus llaves primarias cuando la tabla Hechos comience a llenarse, primero se crea el origen a un archivo plano y se agrega que la separación de columnas sea por medio de la coma (,)



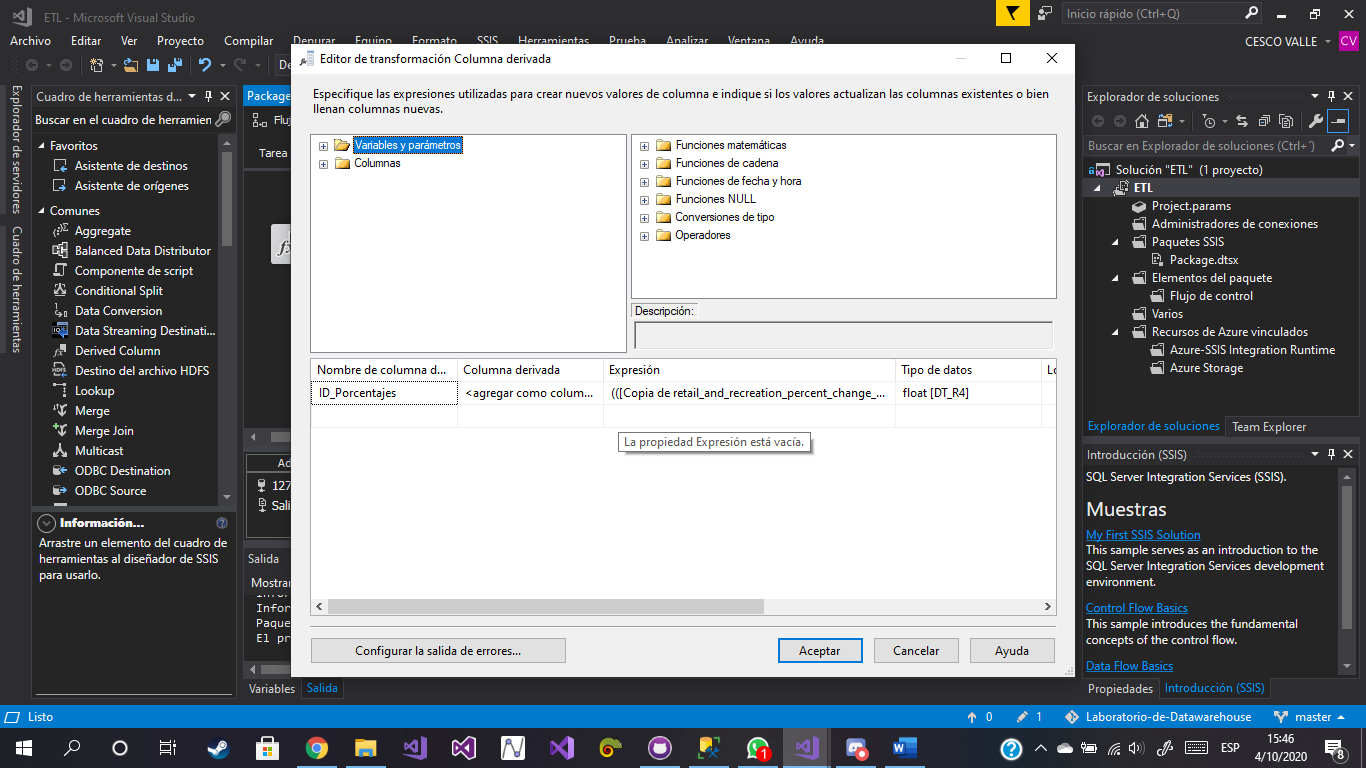
**Paso 3:** Crear una columna Derivada que lea cada fila y verifique que no hay NULL en sitios que no debería, cambiando los porcentajes base de NULL a “0” y la clave de Subregiones de NULL a “UNK”



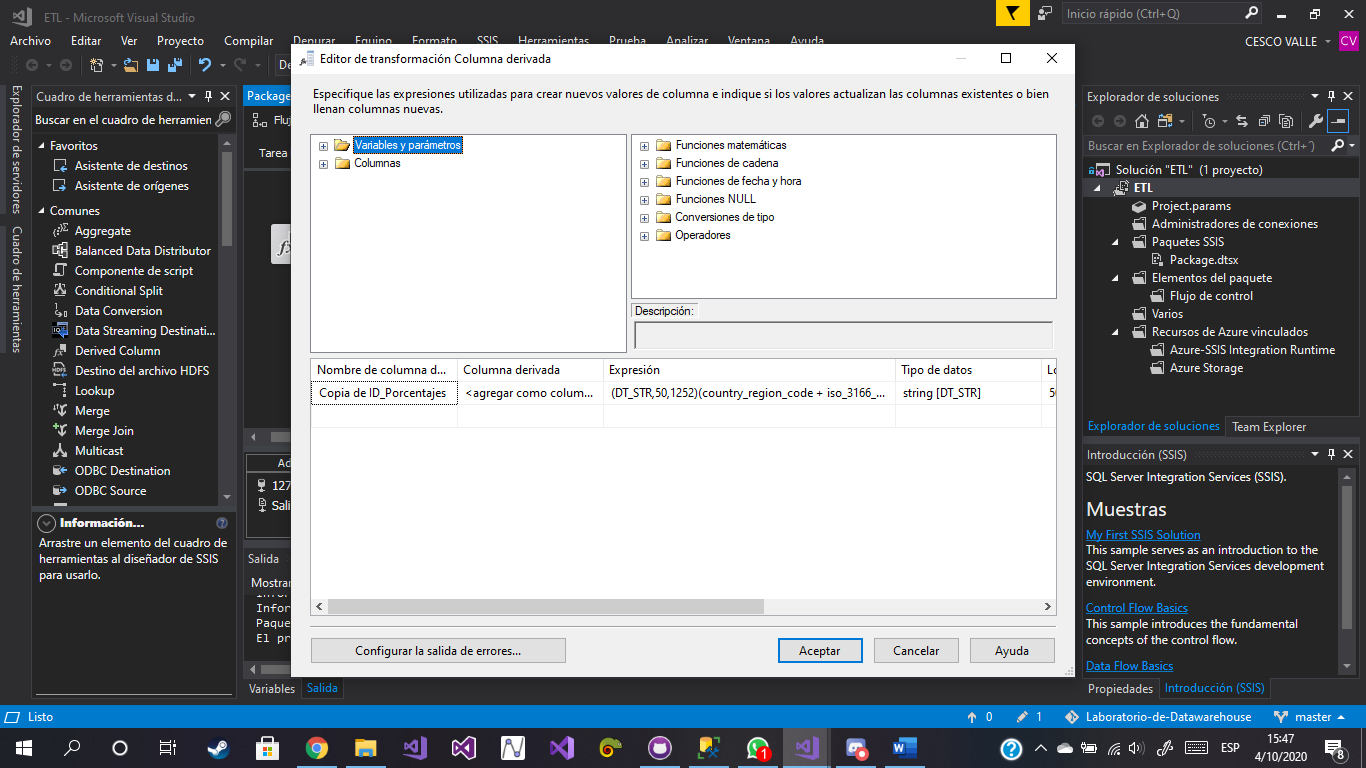
**Paso 4:** Convertir todos los datos posibles haciendo que concuerden con los de la base de datos



**Paso 5:** Los porcentajes son archivos del mismo tipo y siempre están enlazados entre sí, al pensar en separarlos en su propia Tabla se tiene que crear un dato que no existe que será su columna de ID, para esto usamos sus propios datos sumándose, multiplicándose y dividiéndose (El divisor se suma con constantes para que nunca sea cero) creando así una llave muy poco común, pero que aún es posible que se repita, por lo que se usa otra Derived Column para concatenarle su Fecha,Su país y su subRegión.

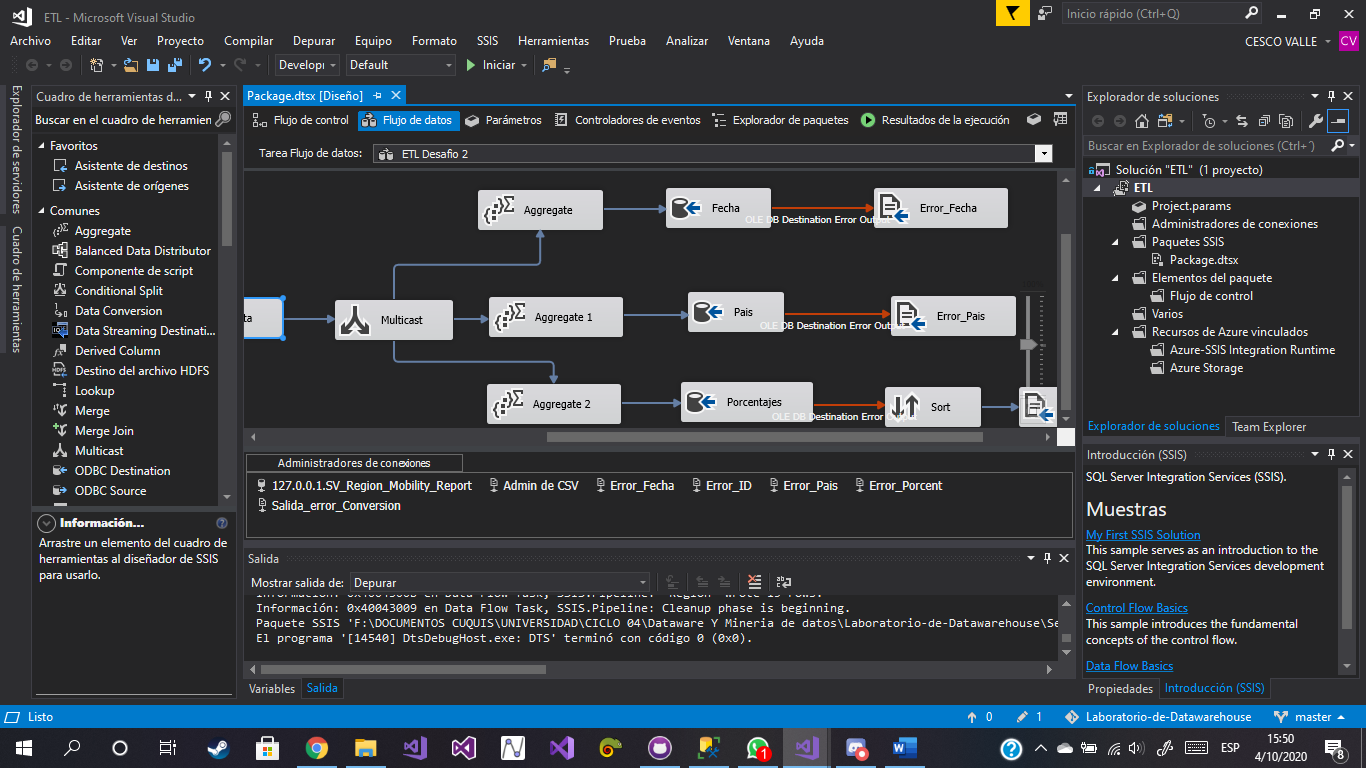


*(([Copia de retail\_and\_recreation\_percent\_change\_from\_baseline]+[Copia de grocery\_and\_pharmacy\_percent\_change\_from\_baseline])\*([Copia de parks\_percent\_change\_from\_baseline]+[Copia de transit\_stations\_percent\_change\_from\_baseline]))/ (1.265656+([Copia de workplaces\_percent\_change\_from\_baseline]+[Copia de residential\_percent\_change\_from\_baseline]))*

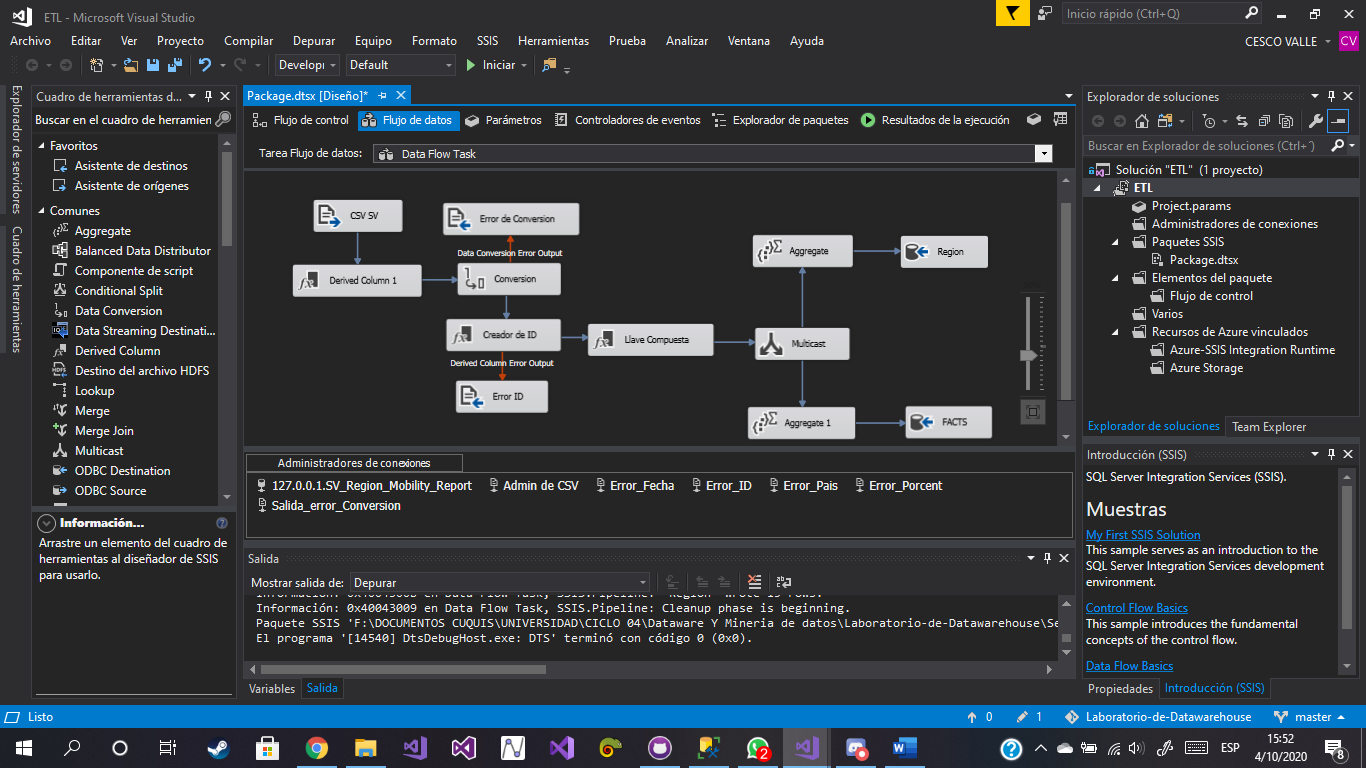


*(DT\_STR,50,1252)(country\_region\_code + iso\_3166\_2\_code + date + (DT\_STR,50,1252)ID\_Porcentajes)*

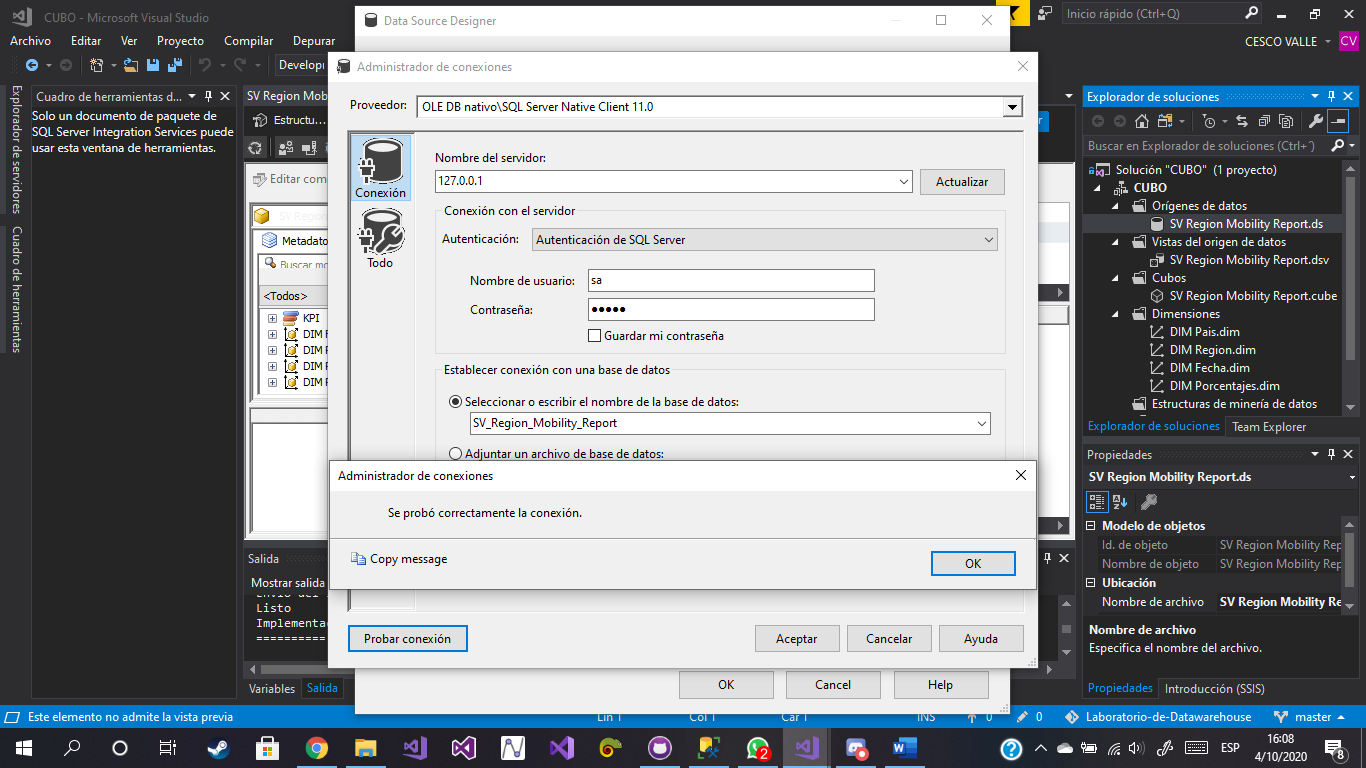
**Paso 6:** Con el uso de Multicast se pasa esta información a distintos Aggregate para que agrupen la información según la tabla lo necesite



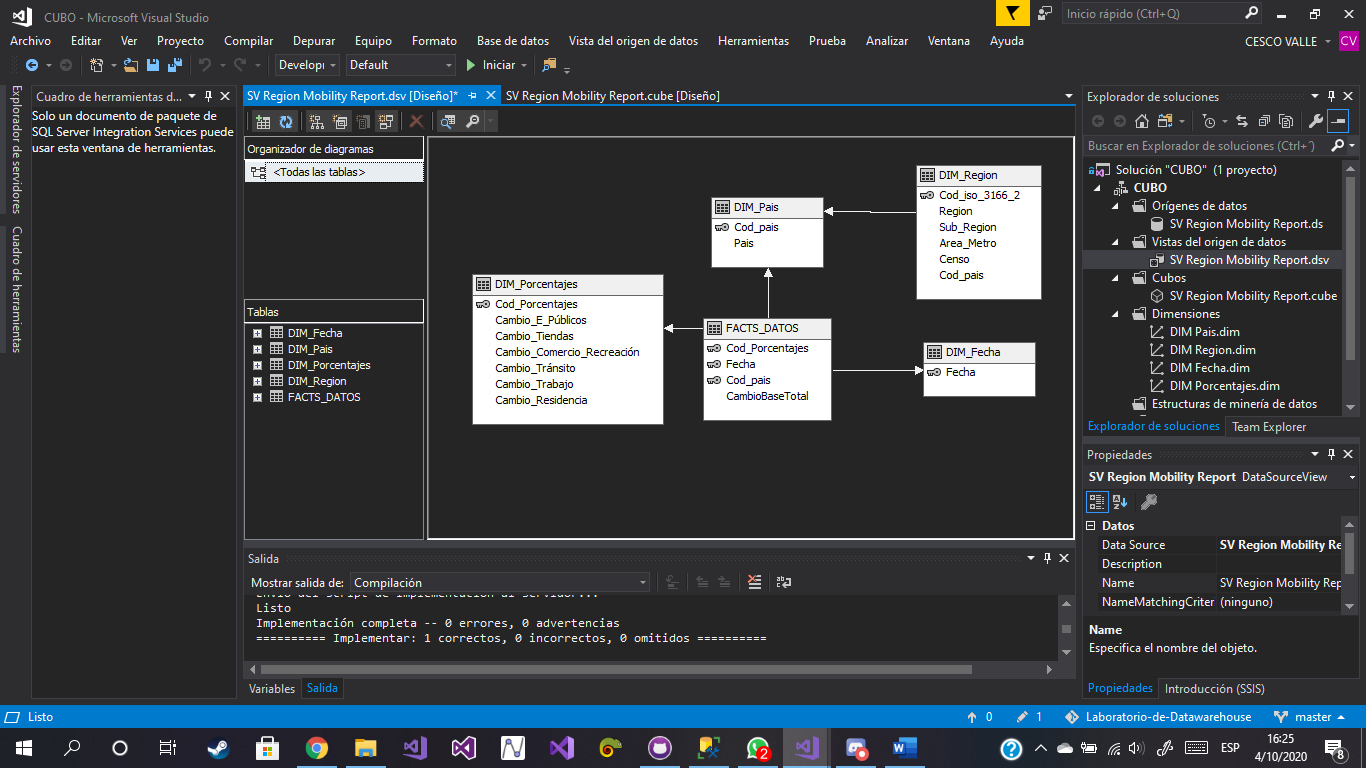
**Paso 7:** Se copia la estructura casi completa del ETL Anterior, pero a partir del multicast se cambian los aggregate para que nuevamente, este se adecue a las asignaciones que piden la Tabla Regiones y la Tabla Hechos



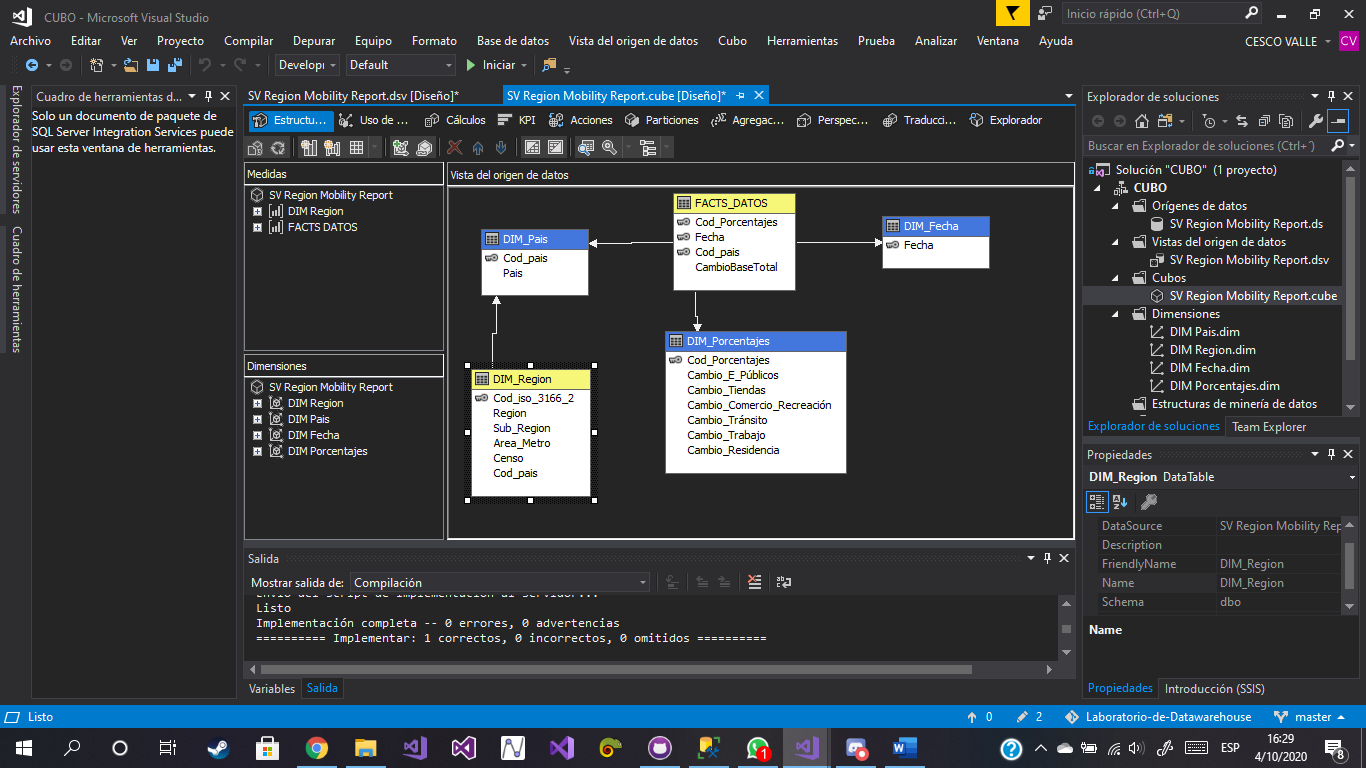
**Paso 8:** Se crea un proyecto de Analysis Services Multidimensional y se selecciona la base de datos creada como Origen de Datos, en este caso se uso la Autenticación de SQL Server y el servidor 127.0.0.1 (localhost)



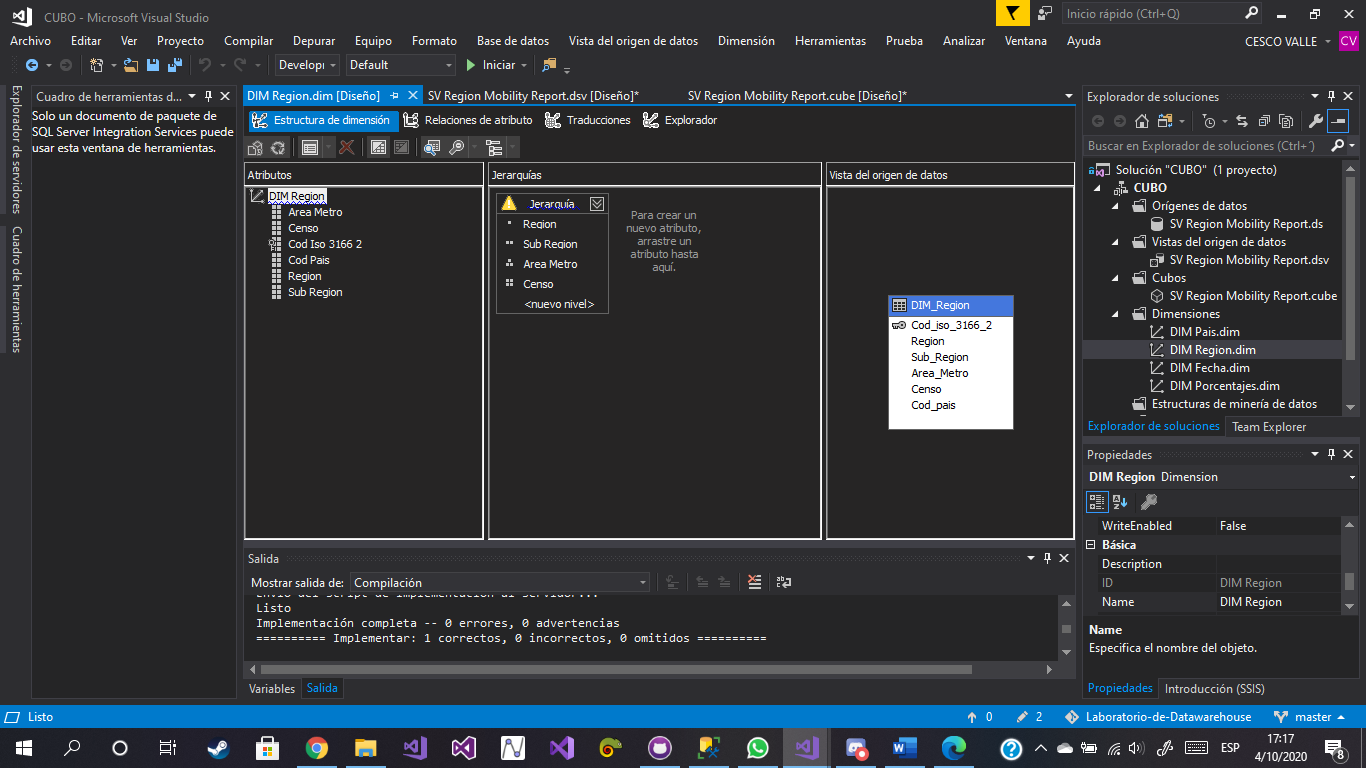
**Paso 9:** Crear la vista tomando el origen y todas las tablas contenidos en este



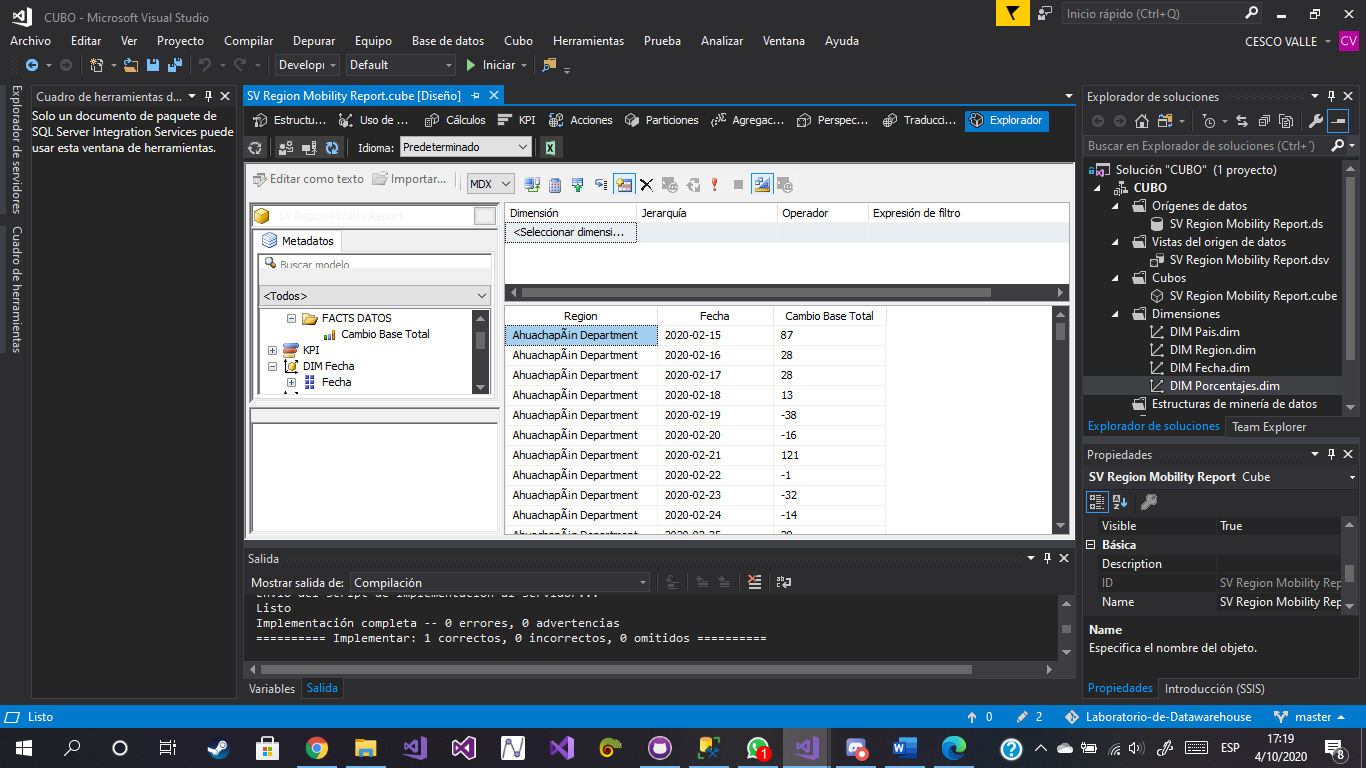
**Paso 10:** Crear el cubo tomando la tabla Fact y la DimRegion como puntos de medida, esta última agregada únicamente para hacer consultas con la dimensión de Región las cuales son más específicas que hacerlas únicamente por país



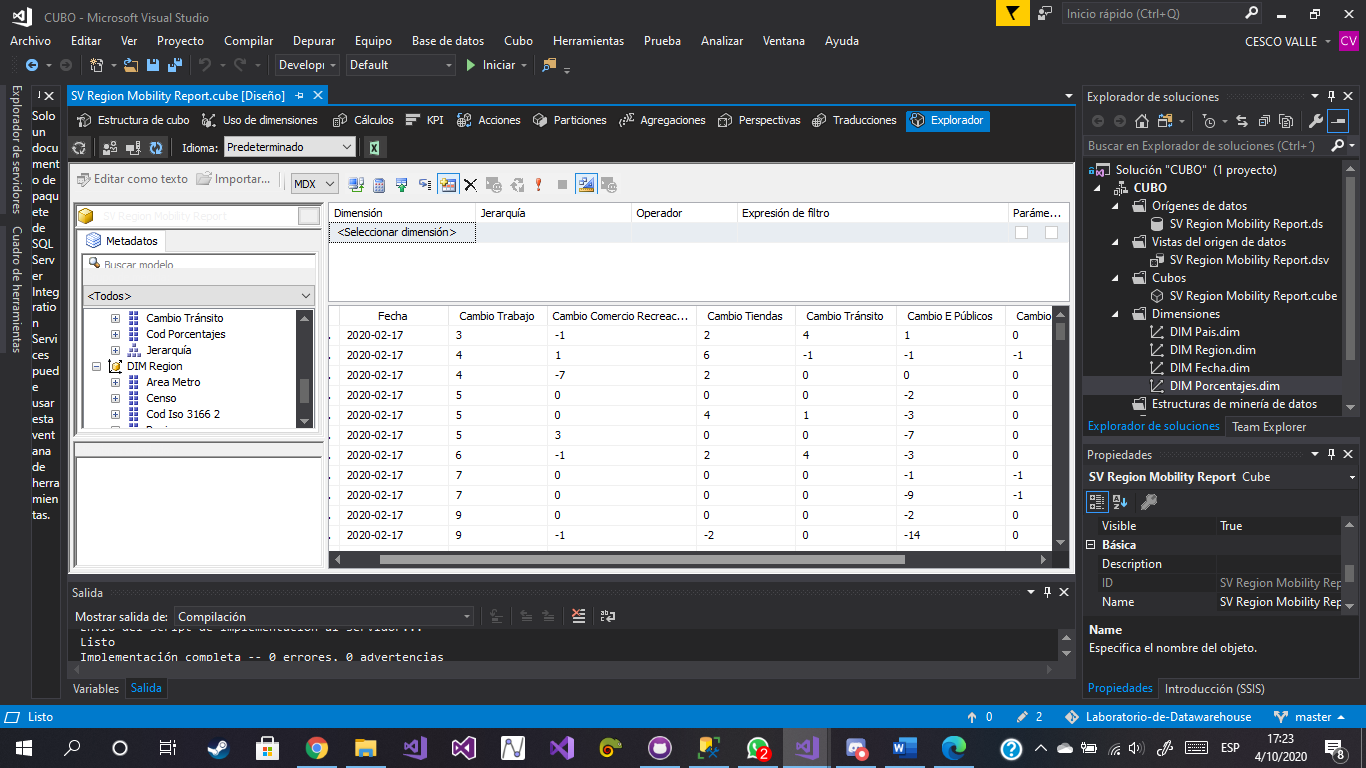
**Paso 11:** Se crean jerarquías para la Dimensión Región para que la Región,SubRegión,Metro y Censo sean juntos un mismo dato, y otra jerarquía en los porcentajes, están únicamente para agruparlos y hacer consultas de todos sin necesidad de traerlos uno por uno



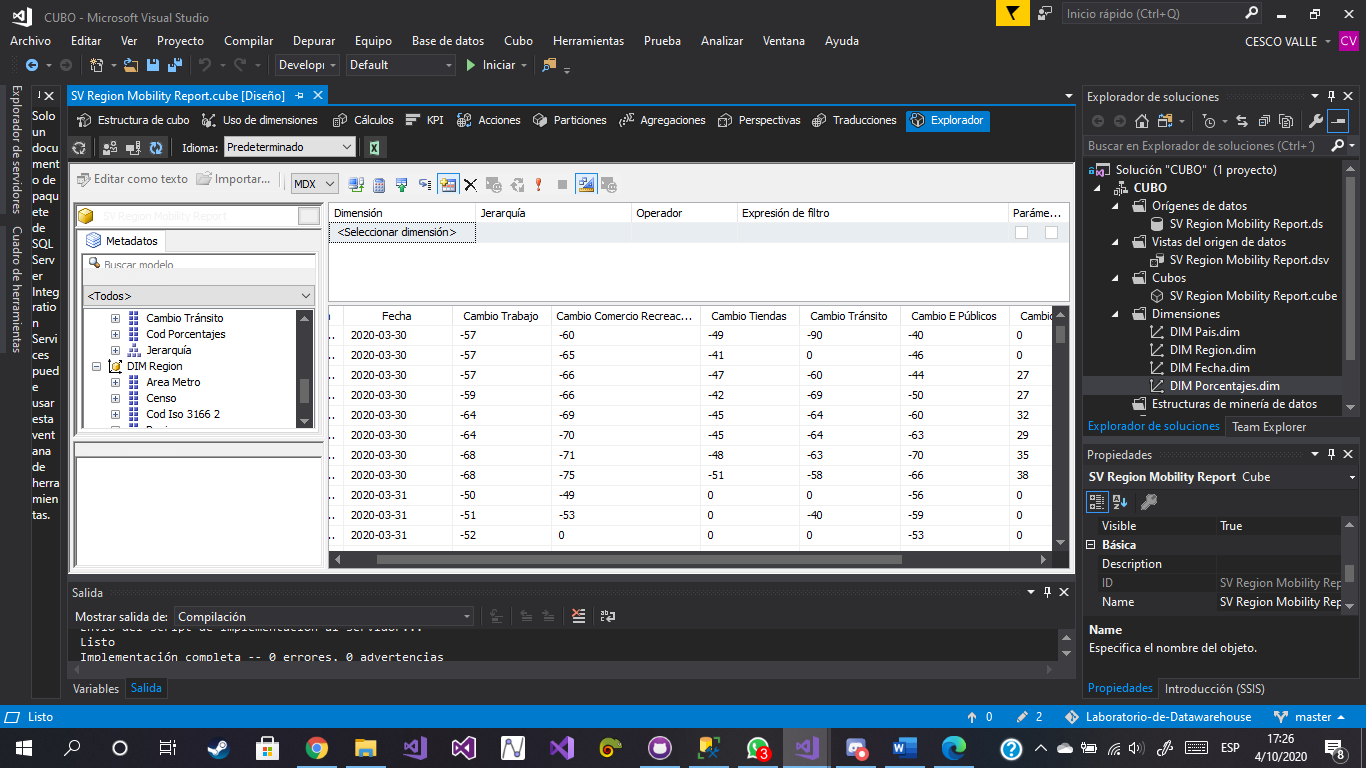
**Paso 12:** Procesar y Examinar el cubo para hacer consultas



Al hacer consultas como estas podemos ver detalles curiosos como que a partir de mediados de marzo el cambio base total daba picos negativos tan alto, y para mayores detalles podemos traer la jerarquía de porcentajes



Podemos ver cómo se mantenía una estabilidad en febrero que muchas veces se mantenía con números de cero o pequeños, siendo casi todos los cambios en tránsito, sitios públicos y el trabajo.



Esta tendencia de estabilidad cambia en finales de Marzo cuando podemos encontrar picos negativos en la mayoría de los rubros estudiados