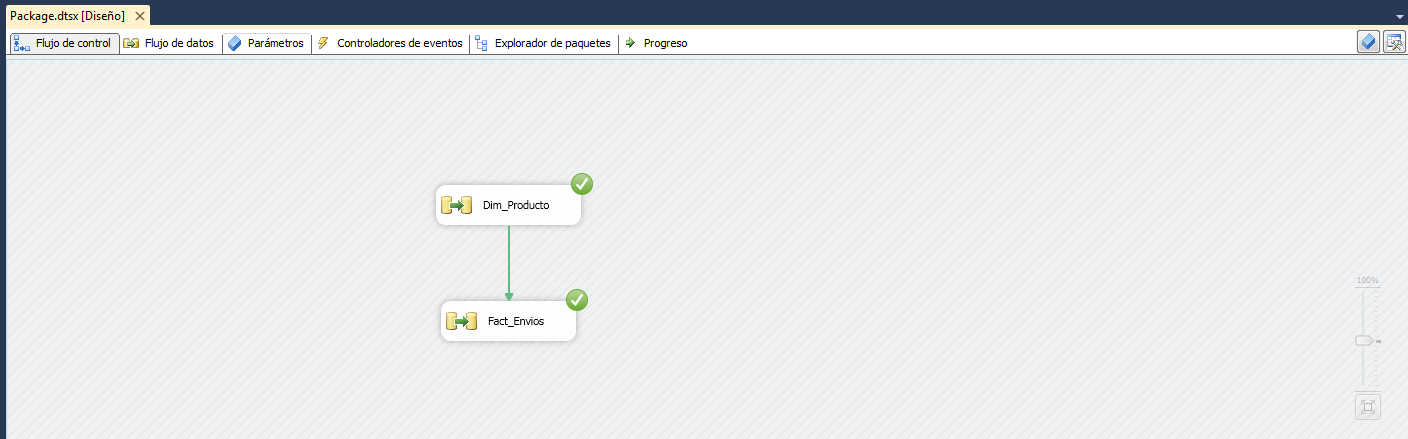
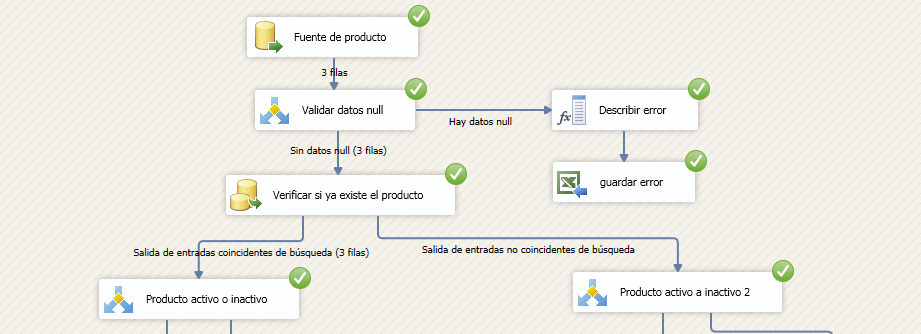
**Punto 2: Construcción del ETL**

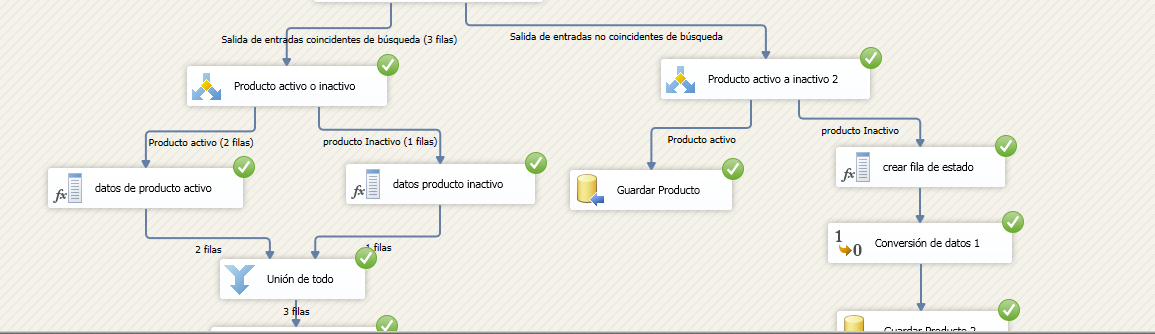
Para desarrollar la actividad inicialmente se crearon dos flujos de trabajos, uno para la dimensión producto y otro para la Fact de envíos que fueron las asignadas al grupo de trabajo, primero se debe ejecutar la dimensión para que al momento de ejecutarse la Fact esté las claves foráneas que necesita.



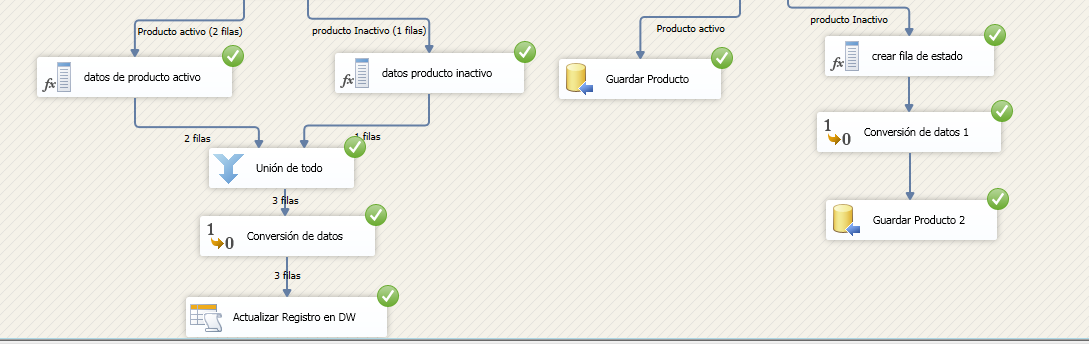
En el ETL de la dimensión producto se puede ver como inicialmente se traen todos los datos de la fuente , en este caso solo es una , de allí se hace una validación que no vengan datos vacíos o null ya que la bodega no los va a aceptar, allí pueden ocurrir dos cosas , que vengan datos vacíos o que no los haya , en caso de que hayan datos nulos crea una corta descripción del error y lo guarda en un archivo Excel destinado para esto, en caso de que no vengan datos nulos sigue con su flujo normal y verifica si el producto ya existe en la bodega de datos, en ambos casos valida si el producto que llega esta activo o inactivo (suspendido , en pausa etc.) , en caso de ser así crea la descripción de activo o inactivo según sea el caso ya que la bodega solo permitirá estos dos valores, activo o inactivo.



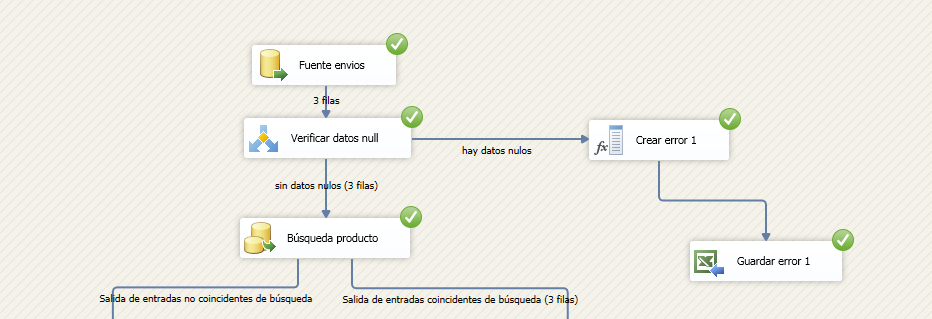
Una vez creo la descripción para el producto activo o inactivo, realiza una unión para que los datos queden en una sola tabla y sea más fácil manejarlos.



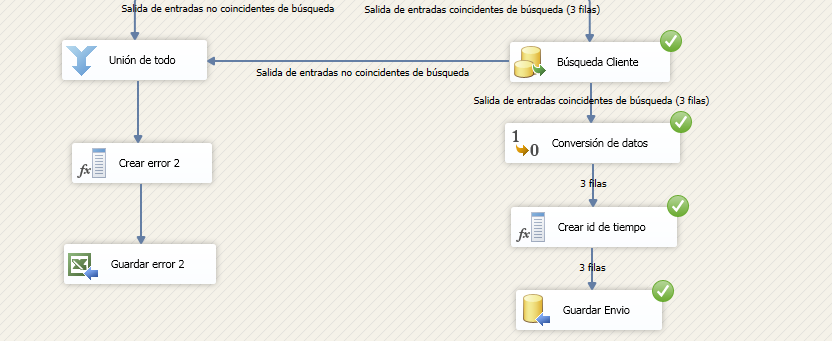
Más adelante realiza unas conversiones de datos para que los registros puedan ser ingresados a la bodega sin problemas, cuando termine este proceso se ingresan los datos a la fuente o si ya existe , los actualiza como se puede ver en las imágenes



Para el ETL de la Fact de envíos inicialmente se traen los datos de la fuente y posteriormente se verifica que ningún dato este nulo o vacío, si esto no se cumple, se crea un registro de error, en caso de que el dato no este vacío, continua el flujo normal del programa y busca el key del producto con el id que viene de la fuente.

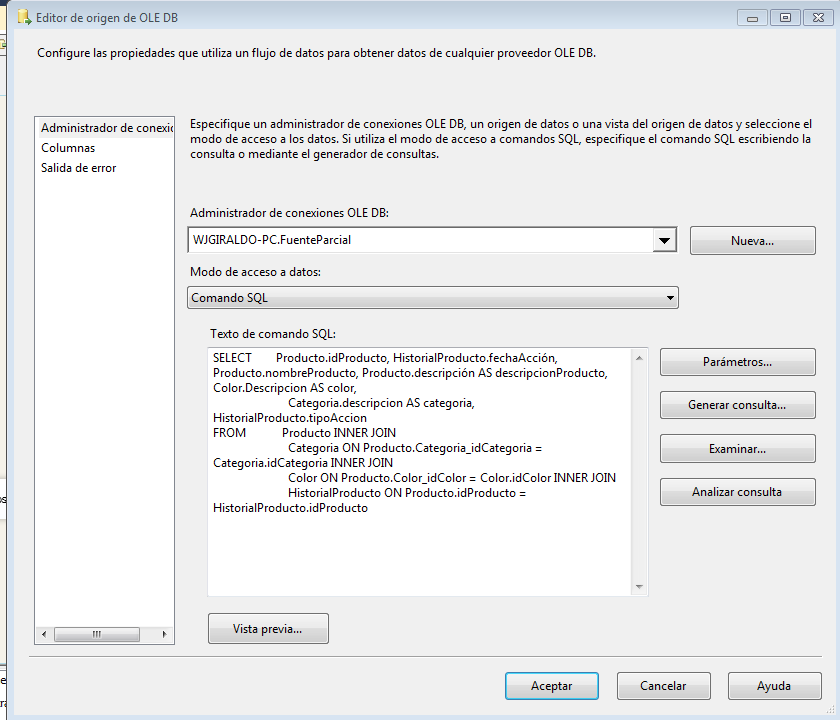


Si encontró el key del producto continua con el key del cliente , luego realiza unas conversiones de datos y crea el id de tiempo, luego guarda el envió en la bodega de datos.



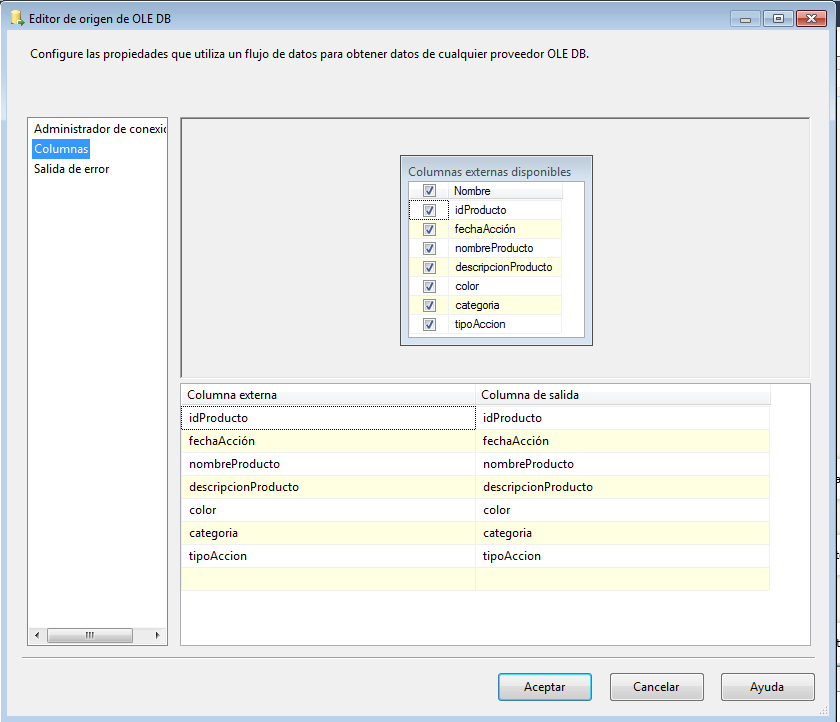
Para cargar traer el producto de la fuente se hace a través de una consulta SQL para no tener que hacer varias fuentes, en esta consulta se puede ver que se unen las siguientes tablas a través de JOIN

* Producto
* Categorías
* Color
* Historia de producto

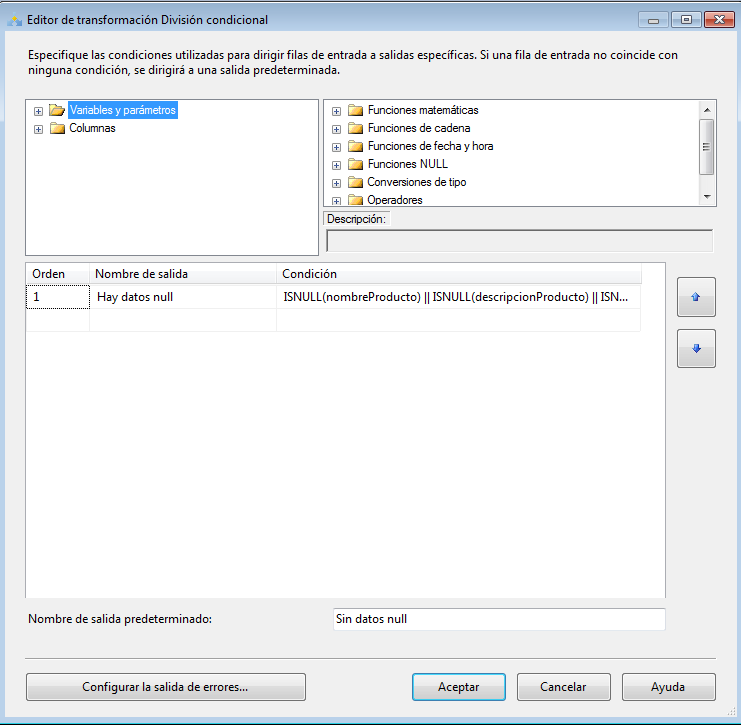


De esta manera solo es necesaria una fuente para traer todos los datos necesarios

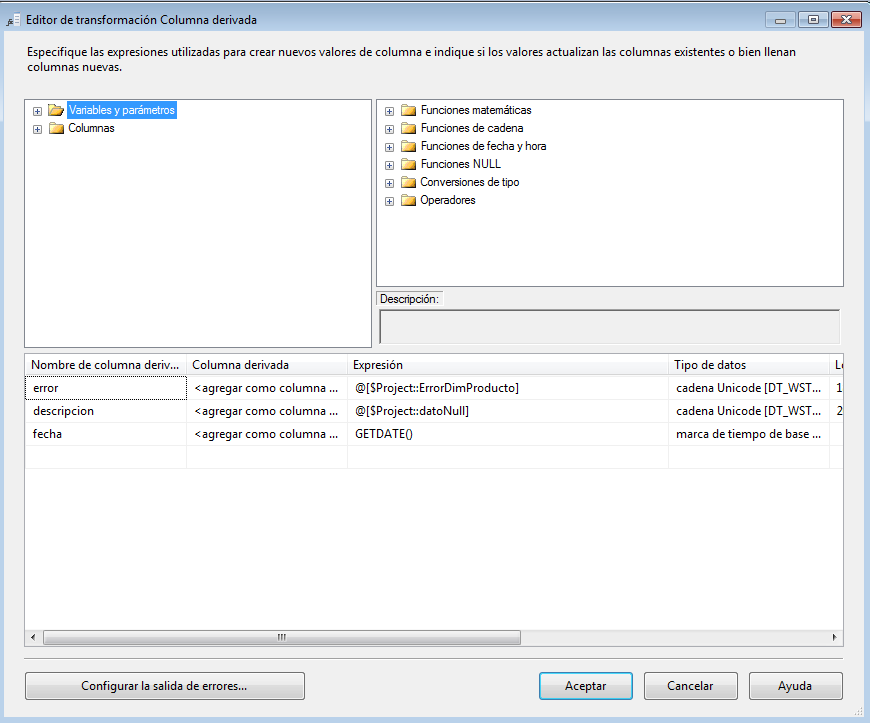
Se traen únicamente las columnas que necesitamos para ingresar a la bodega, por ejemplo la fila de id Producto y color



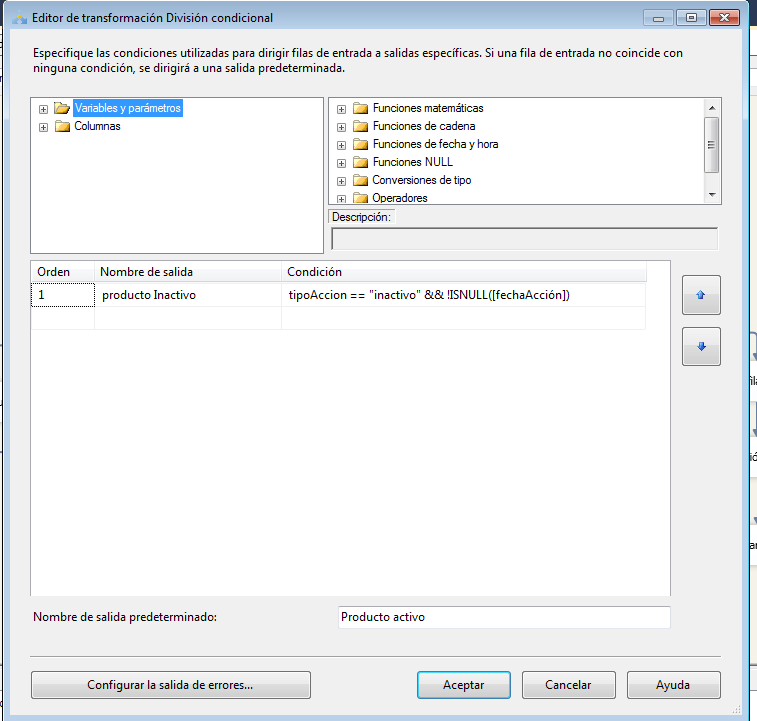
Se realiza la validación de que ningún dato sea null, en caso de que haya algún valor nulo se manda por la salida con el nombre Hay datos null



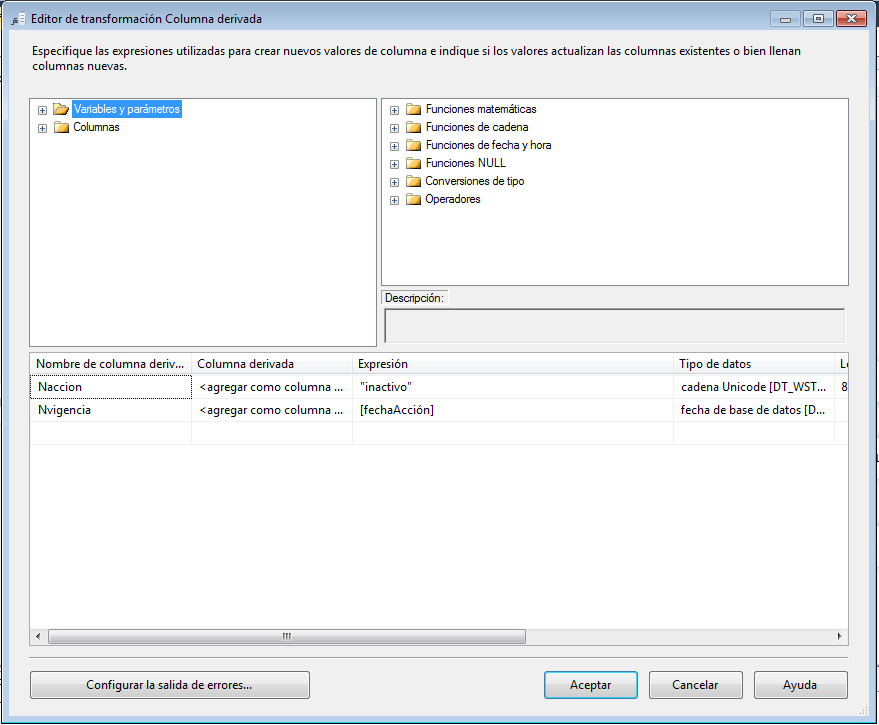
Esta columna derivada se realiza para que el etl guarde el nombre del error en un documento de Excel ,en este caso guarda el nombre ,la descripción y la tabla donde ocurrió el error



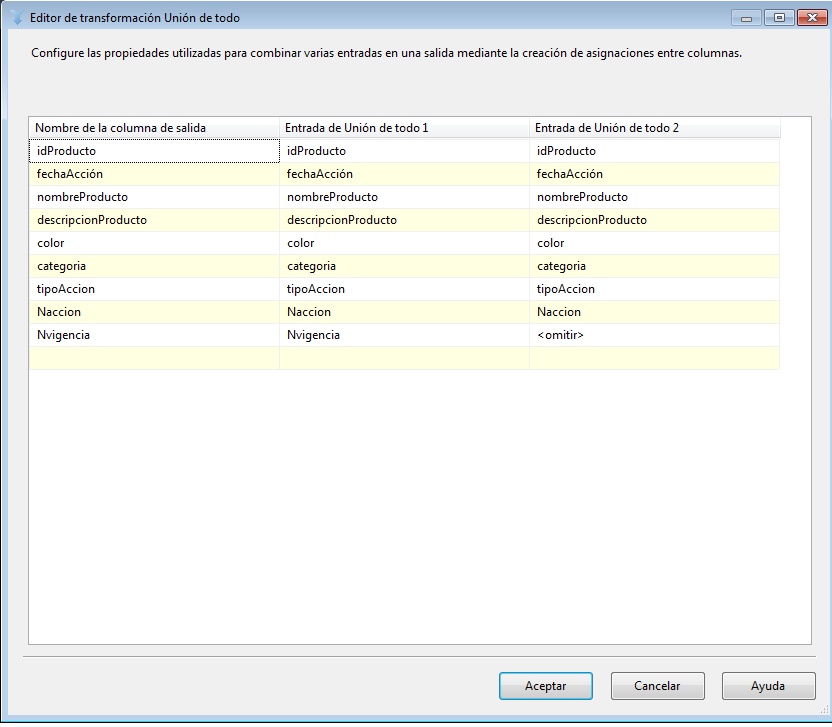
Se debe hacer una división condicional para saber si el producto está activo o inactivo , en caso de que el producto este activo , se envía por la salida predeterminada en caso de que no lo esté entonces se guarda la fecha de vigencia y se le coloca el nombre de inactivo.



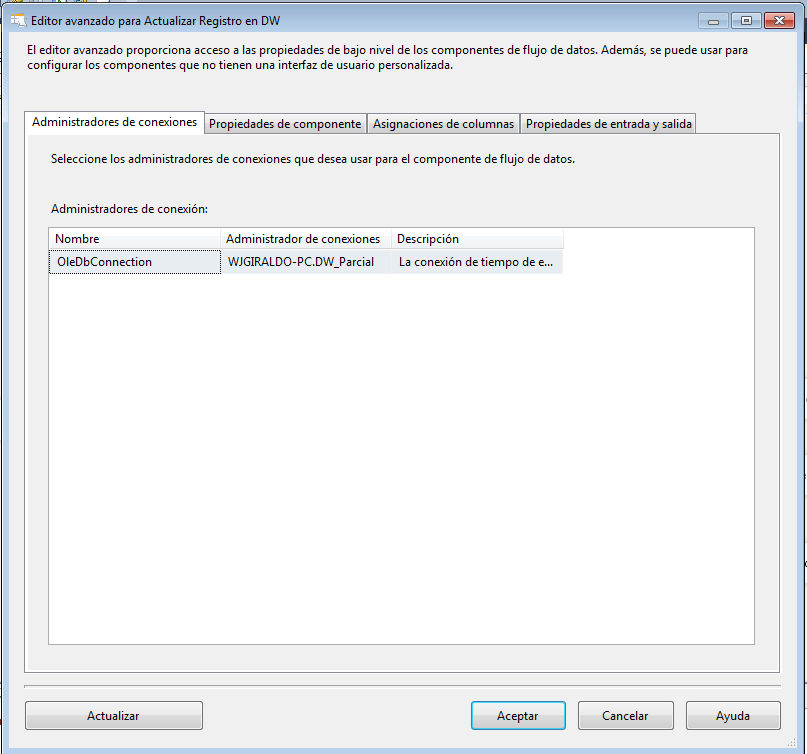
Si el producto está activo , no se guarda fecha de vigencia ya que este campo puede estar null y solo debe asignado cuando el producto este inactivo.



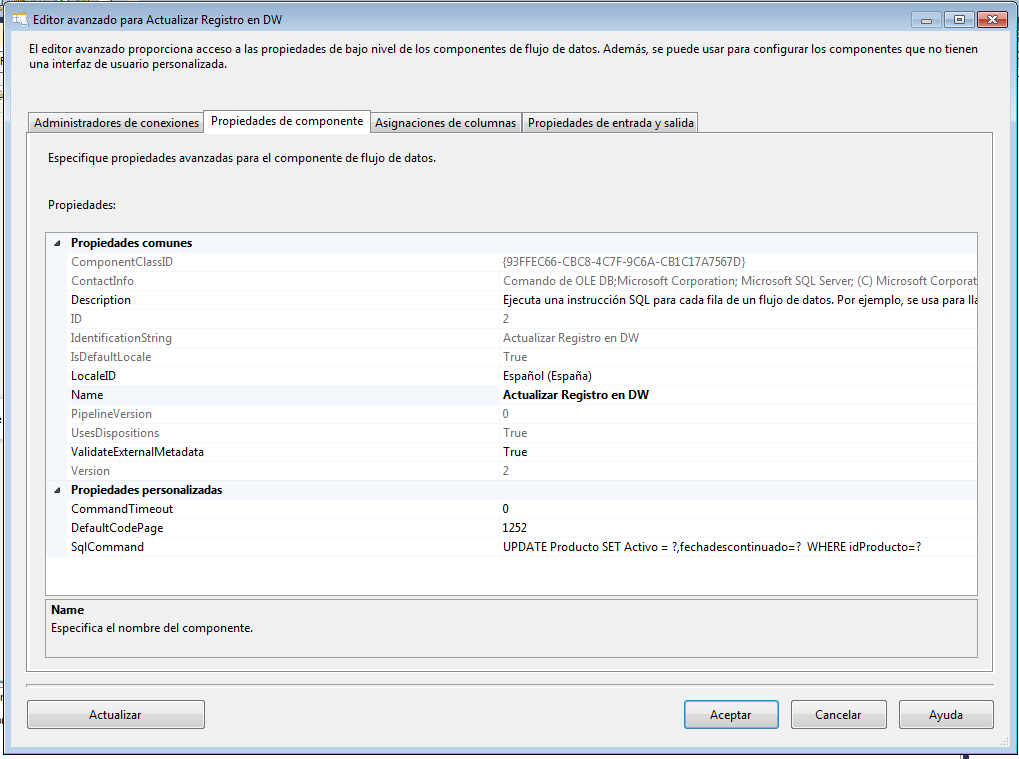
La unión para que los registros queden en la misma tabla.



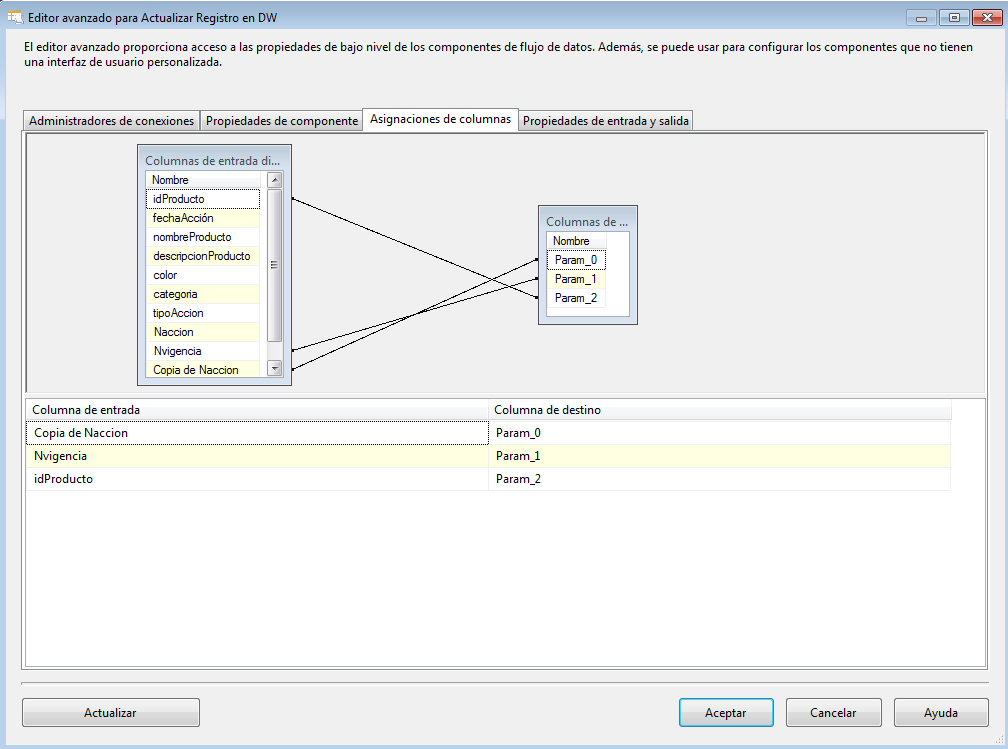
Por último se crea una consulta para actualizar el producto en caso de que ya existe en la bodega de datos, como se puede ver en la imagen se conecta con la bodega.



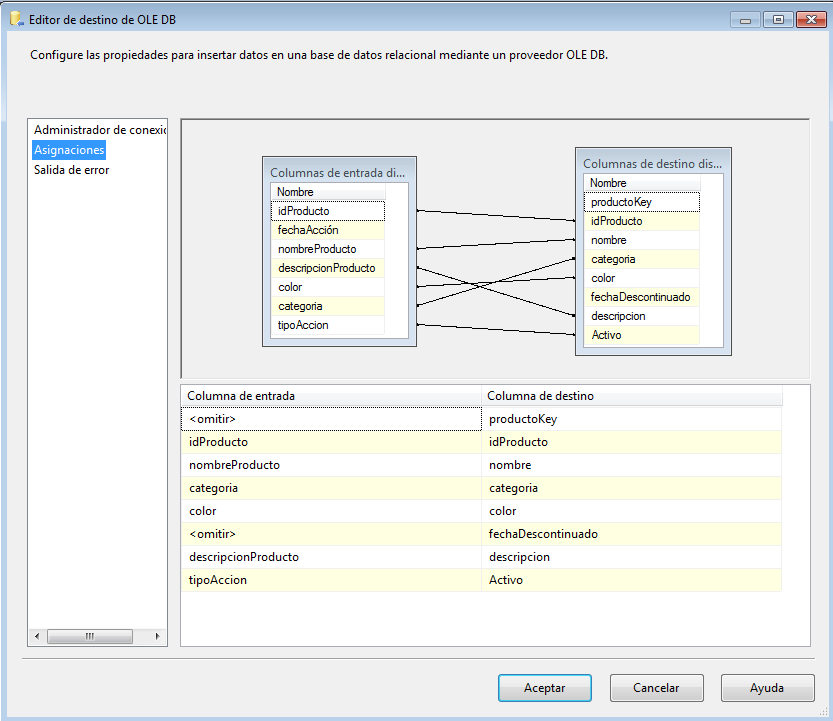
Posteriormente se realiza una consulta con la instrucción update para actualizar el producto , por ejemplo si el producto ya existe y está inactivo y llega un registro con el mismo producto pero activo entonces por medio de este comando se actualiza en la bodega



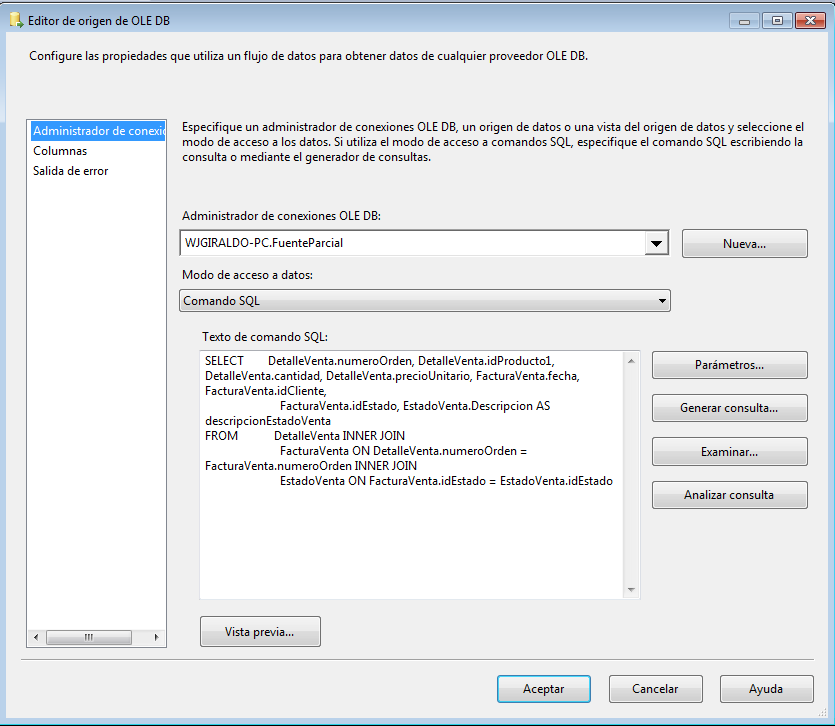
Se asignan los valores correspondientes para que pueda actualizar el producto en la bodega



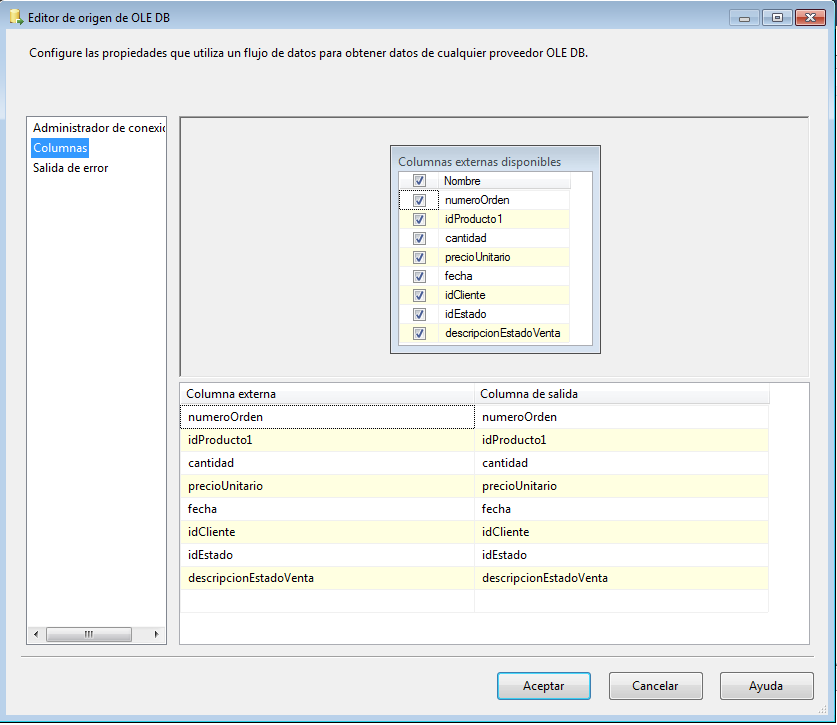
En caso de que sea la primer vez que se va a guardar un producto, entonces si se hace con el componente de destino, acá se realizan las relaciones entre las dos tablas y si el producto está activo, la fecha de vigencia será null



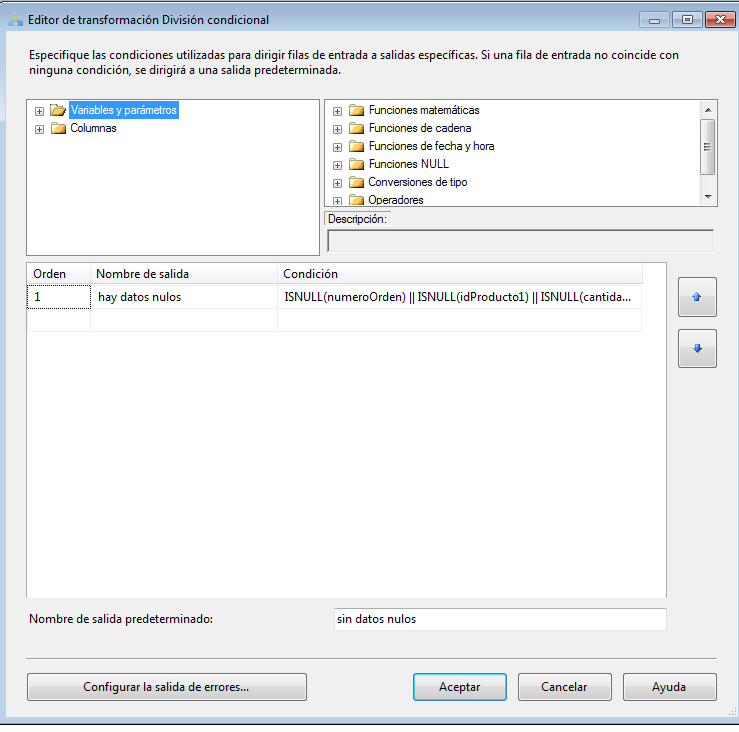
Para realizar la carga de la Fact de envíos también se realiza una consulta como la que se hizo para el producto para no tener que cargar cada tabla por separado, como se puede ver en la imagen se traen las tablas de detalle venta, factura de venta y estado venta.



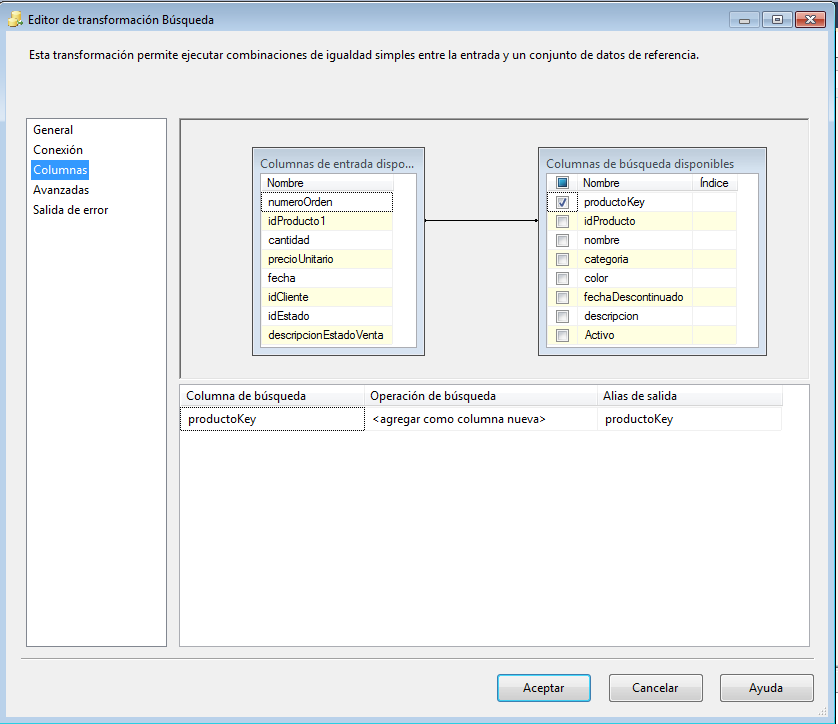
Solo se seleccionan las filas necesarias que van a ser ingresadas al Datamart y se le asigna sus nombres correspondientes



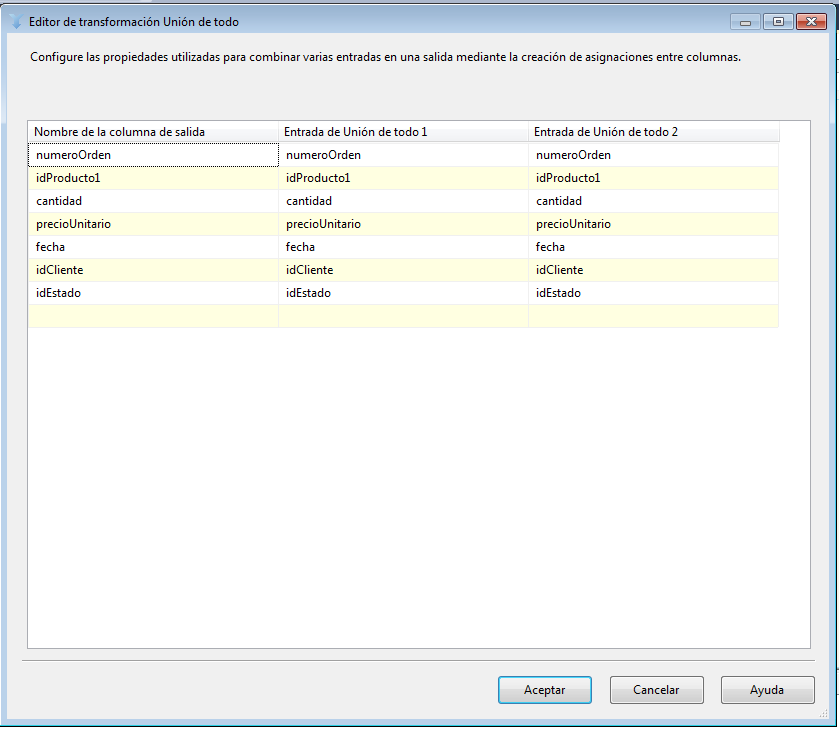
Se verifica que no vengan datos nulos desde las fuentes, para esto se realiza un división condicional donde se verifica si los datos que están llegando en realidad tienen datos o están vaciasen caso de que hayan datos vacíos se enviaran por la salida de hay datos nulos



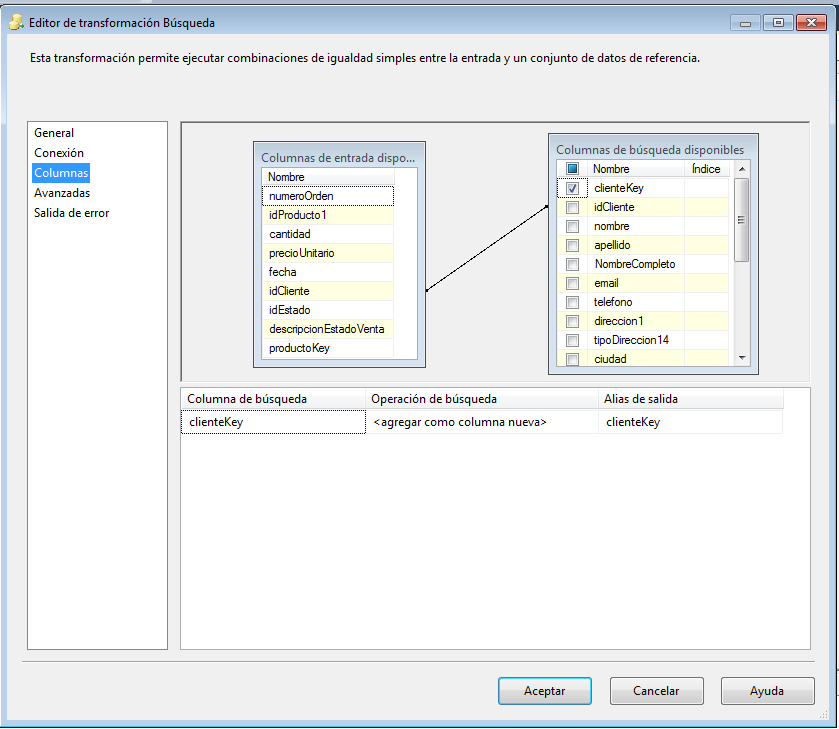
Se necesita traer la clave de la bodega de producto para esto se hace una búsqueda y se comparan los id de producto, se debe marcar la fila de productoKey ya que este es el valor que queremos que aparezca para ingresarlo a la Fact.



Se realiza una unión para seguir el flujo solo con una tabla

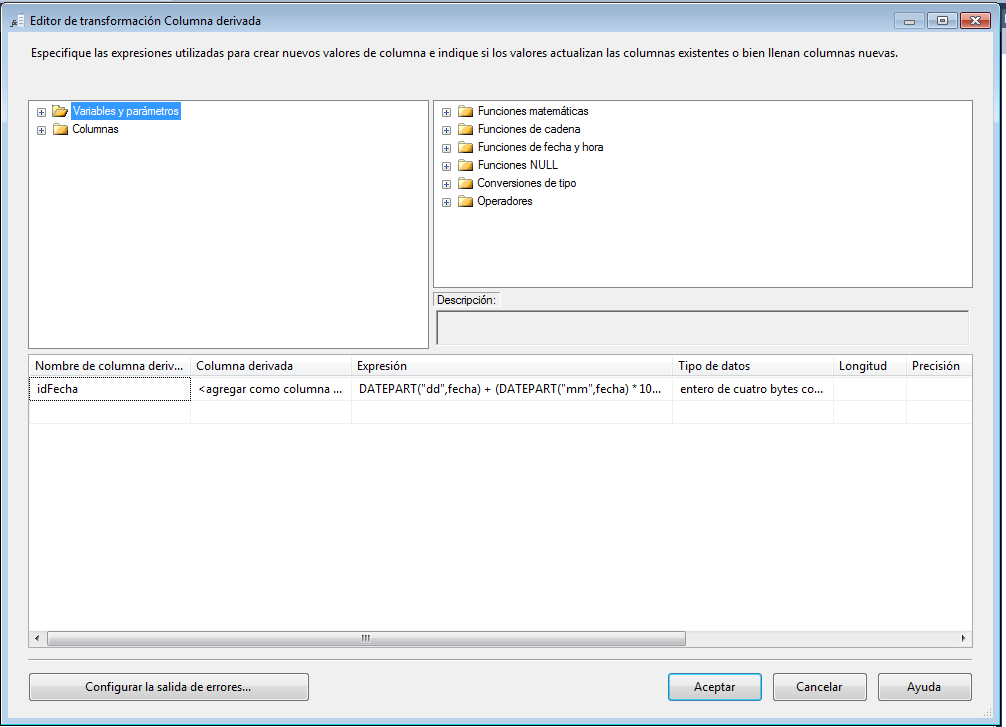


El mismo proceso que se realizó para traer el producto se realiza para traer el cliente de la fuente , también se hace comparación entre los id y se marca la fila de clienteKey para que este valor salga en los próximos componentes

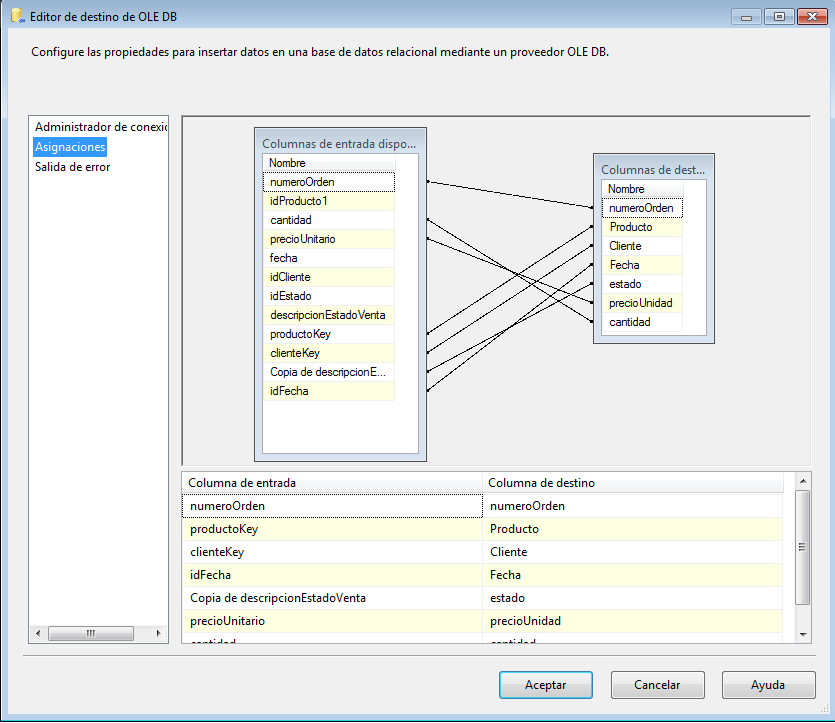


En este punto se debe crear el id de la fecha, ya que este id es una combinación sin símbolos de la fecha completa, esta debe ser generada para que sea única en un día por ejemplo.

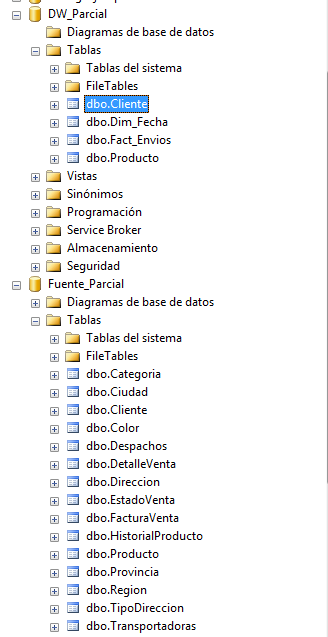
Para que cumpla que el tipo de dato sea entero entonces se realizan unas conversiones para extraer el día el mes y el año.



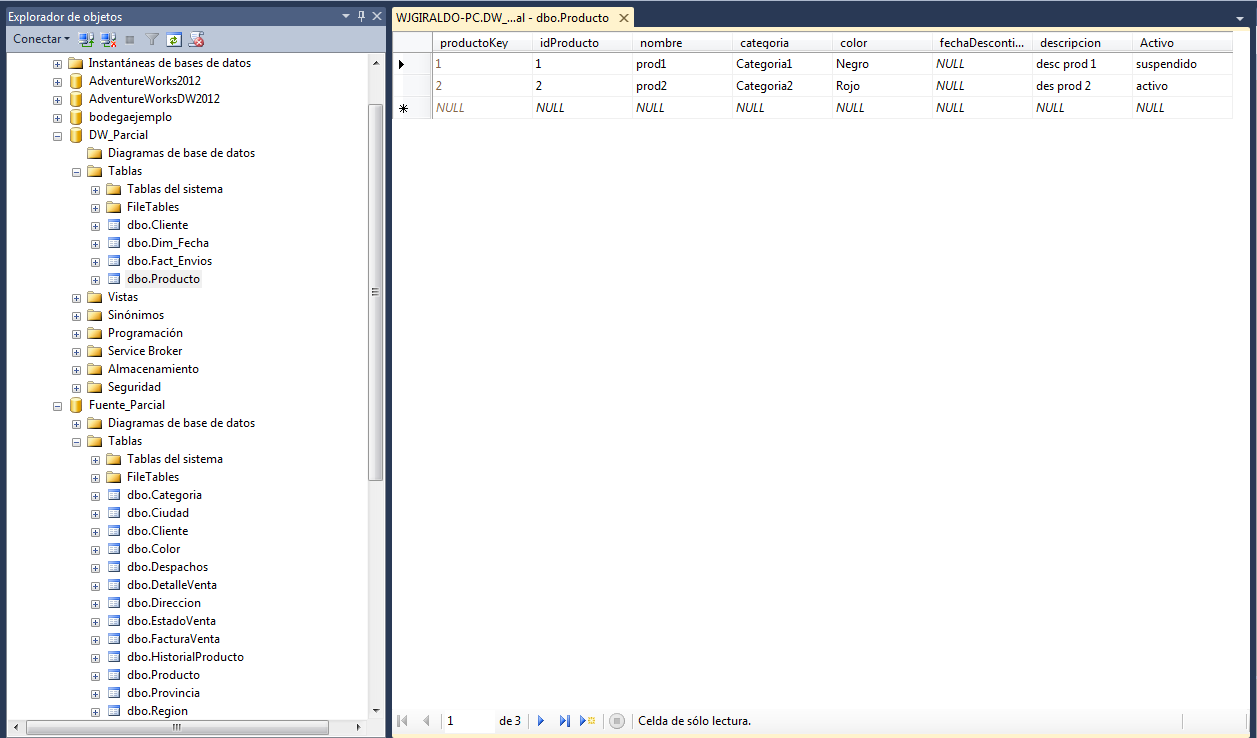
Se realizan las asignaciones para guardar los datos en la Fact de la bodega , en la imagen se puede ver como se relacionan las columnas , como ya se tienen los key del producto y del clientes solo se asignan a los valores correspondientes. Este proceso también se hace para las medidas



En la imagen se puede ver la estructura de la base de datos fuente y de la bodega de datos



En la imagen se puede observar como a pesar de que en la fuente hay dos productos con varias fechas de vigencia en la bodega solo pasa uno con su respectivo estado, este estado puede tomar dos valores , activo o inactivo , en caso de que el producto se encuentre en un estado diferente de activo , por ejemplo suspendido , se pasara a la bodega como producto inactivo



A continuación se muestran los dos ETL en ejecución con unos datos de prueba.

