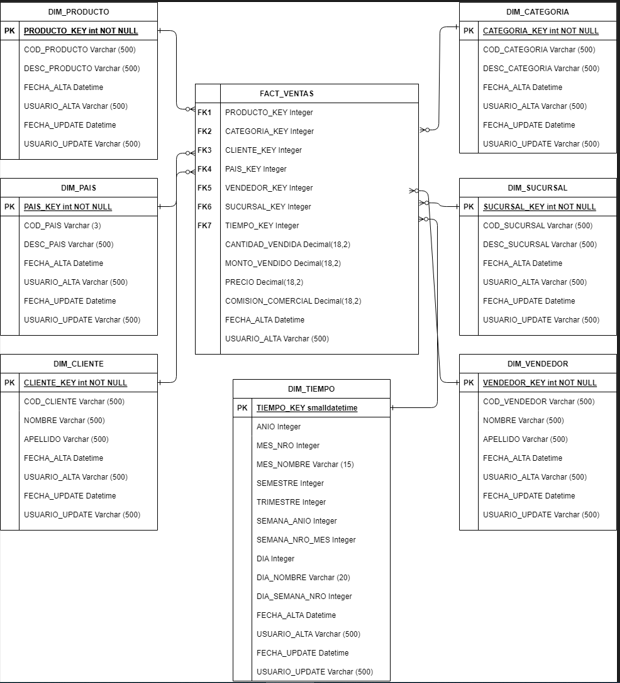
**Data Process I**

Ejercicio Integrador

Implementación de una Solución de Datos

Para implementar una solución de datos al ejercicio integrador, se optó por una serie de validaciones que argumenten la toma de decisiones realizada.

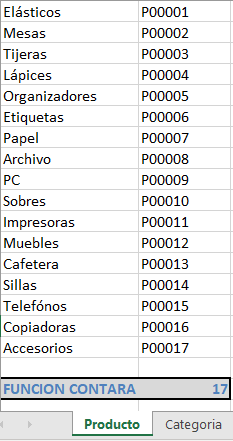
Para comenzar con las validaciones, es importante recordar que se escogió un Modelo Estrella, el por qué de su elección esta detallado en el archivo AXEL\_MORIENA\_DW\_COMERCIAL\_SUPUESTOS, pero para tenerlo a mano, también se transcribe a continuación:

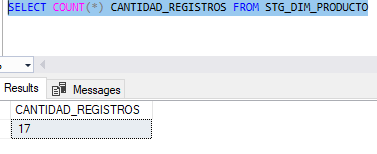


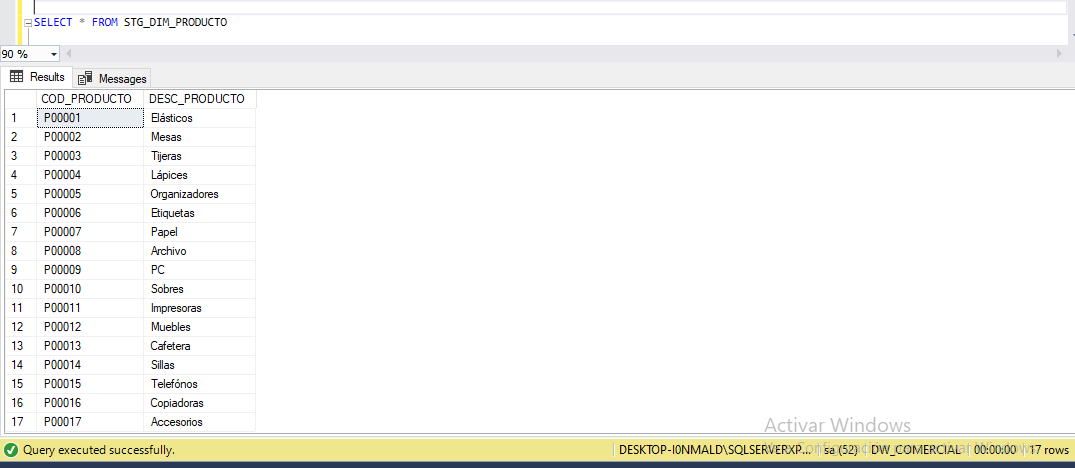
Se contó en un primer momento con un archivo Excel que contenía 7 (siete) pestañas con distinta información. Se cargó a través de la funcionalidad de importar datos de SQL Server. Primero fue importado a tablas de tipo Staging (STG), luego a Intermedias (INT) y por último a las definitivas, llamadas DIM, las de dimensiones o Fact, las de hecho.

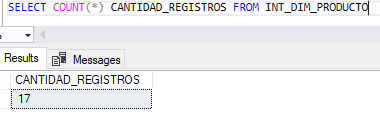
A continuación, se mostrarán los registros de cada tabla en Excel, en STG, en INT y FINAL, para validar que no se perdieron datos.

Tabla PRODUCTO.









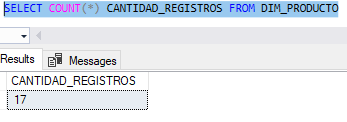
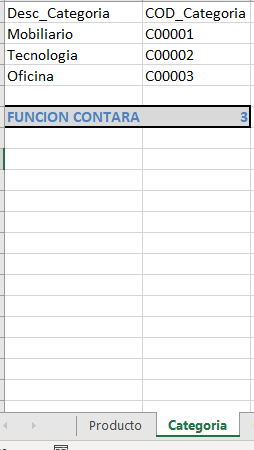
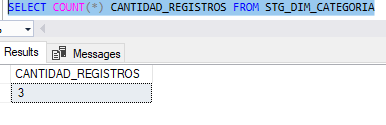
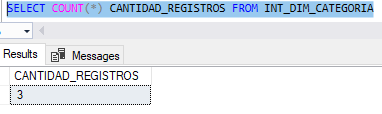


Tabla CATEGORIA.







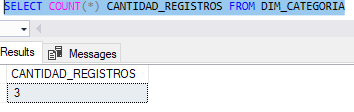
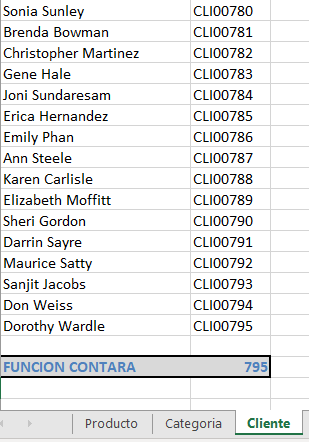
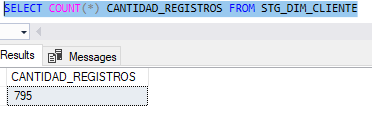
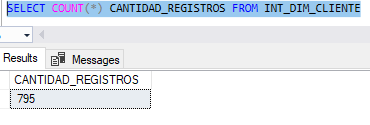


Tabla CLIENTE.







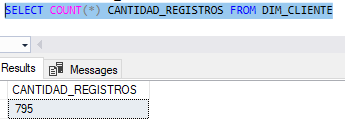
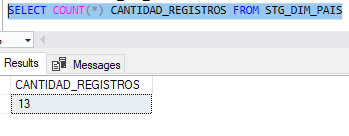
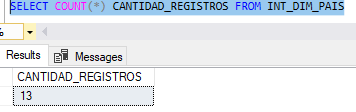


Tabla PAIS.







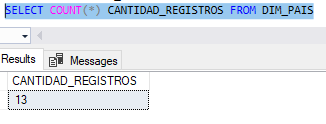
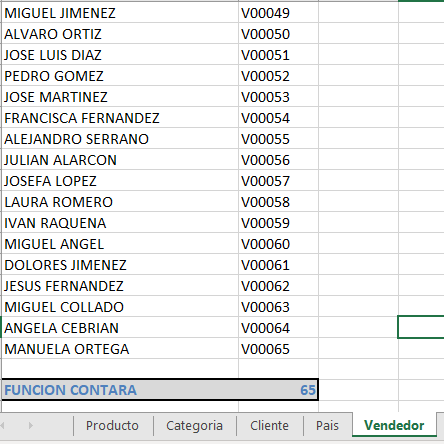
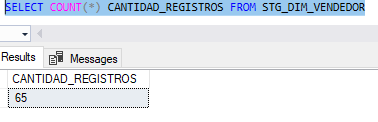
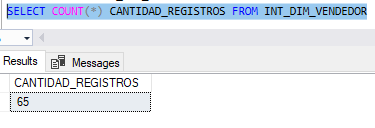


Tabla VENDEDOR.







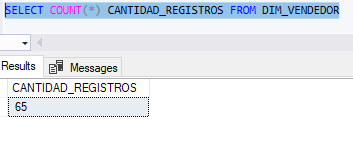
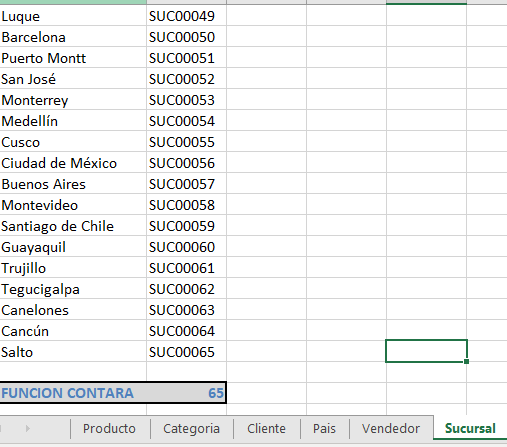
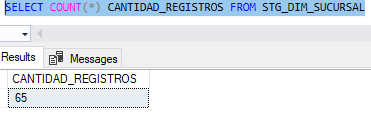
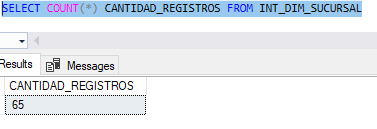


Tabla SUCURSAL.







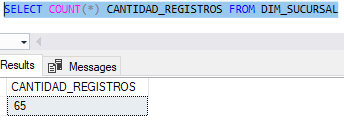
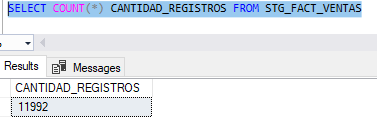
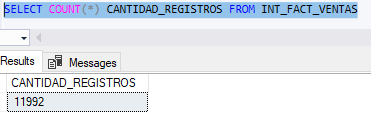
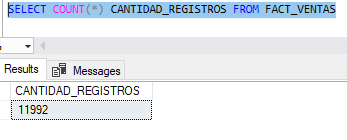


Tabla VENTAS.

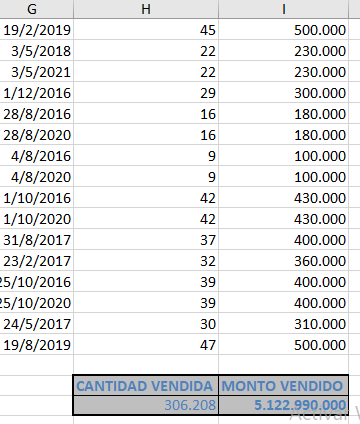


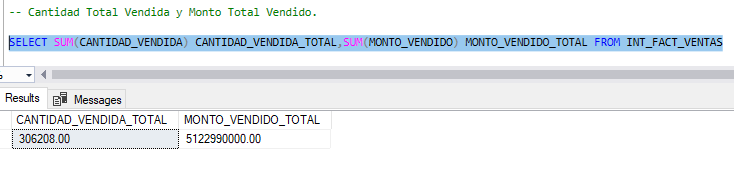




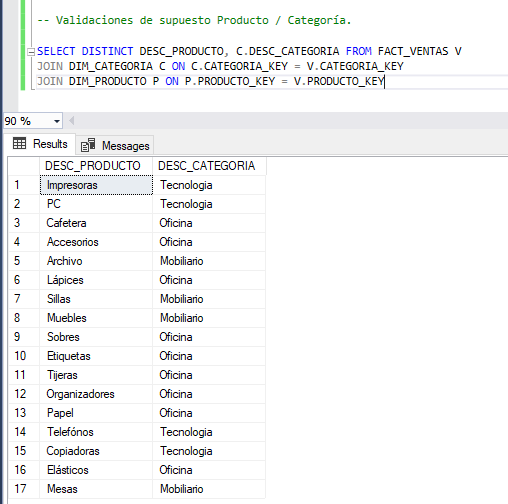


Cantidad Vendida Total y Monto Vendido Total. (Sumatoria de la Cantidad y del Monto Vendido en la Tabla de Ventas).





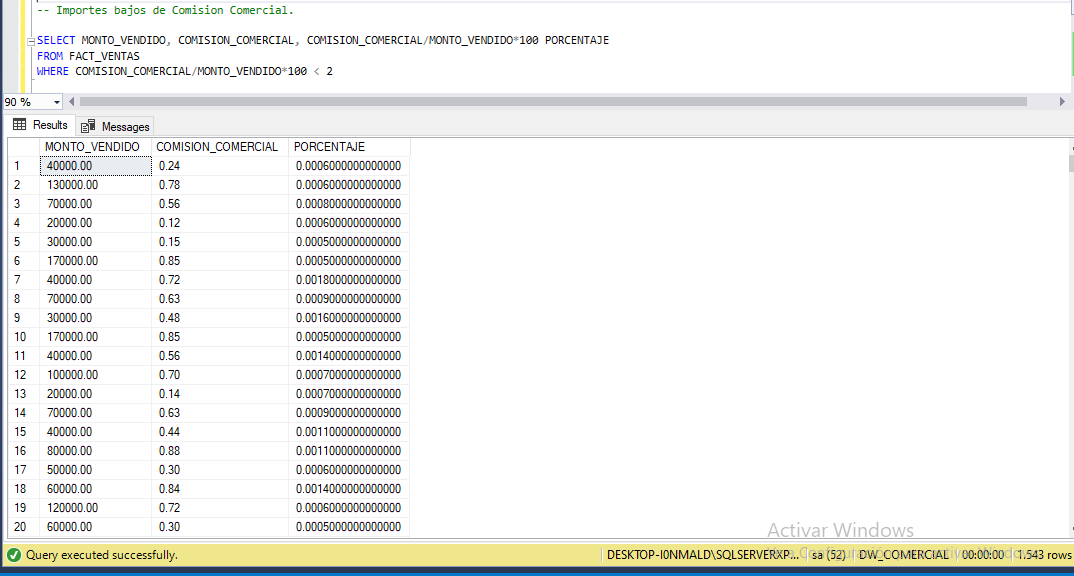
Validaciones de Producto / Categoría, mencionada en el archivo de Supuestos.



Esto indica que para cada producto, corresponde solo y una sola categoría (siempre la misma).

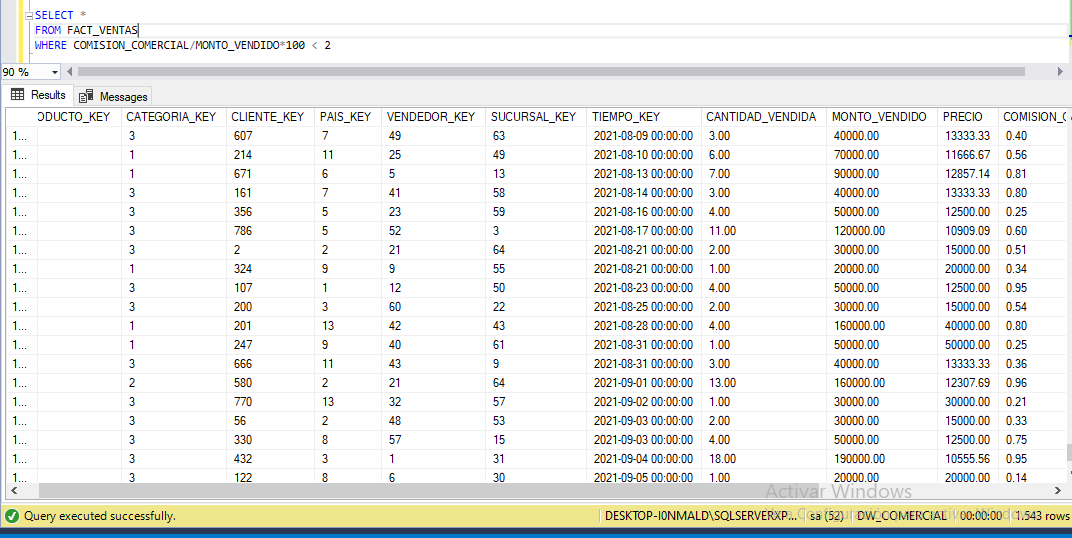
**Validaciones punto 2 (Dos).**

Respecto a comisiones muy bajas, presentadas como supuestos.



Se presentan 1.543 registros con un porcentaje menor al 2% de la venta, coincide en que todas las comisiones son de “0, tanto”, es decir, menores a 1.

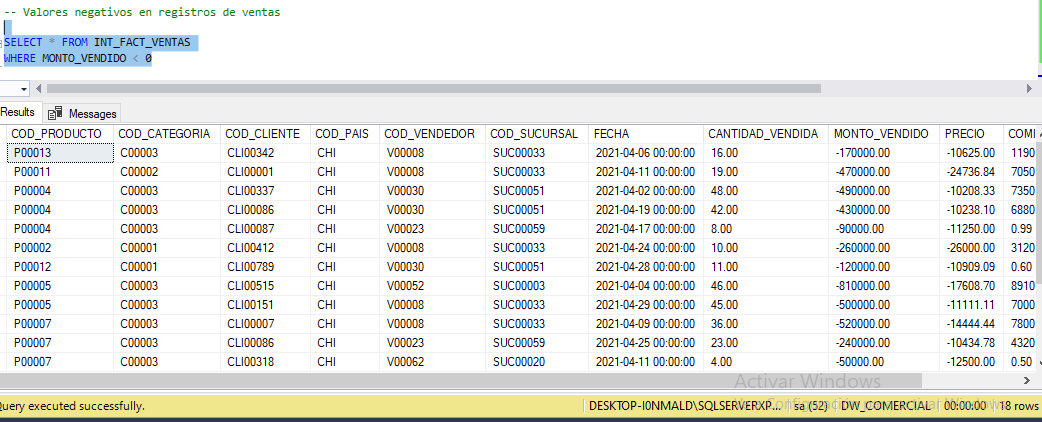
Obteniendo los campos totales, no se puede encontrar relaciones entre las mismas que indiquen otra razón. Por ejemplo que hayan sido de la misma fecha (hay de todos los meses y todos los años) de un cierto producto, categoría, cliente, sucursal. Es todo aparentemente “aleatorio”.

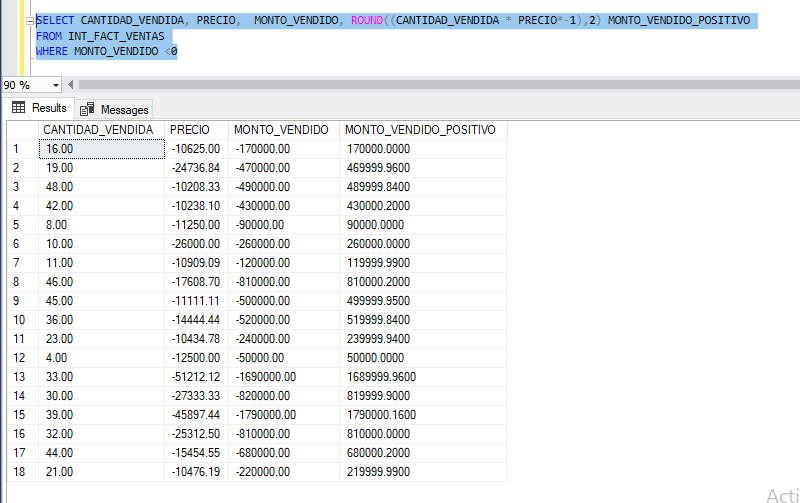


Por lo tanto, se puede enviar esta consulta a la empresa para consultar, pero es muy probable que esto indique que no tenga comisiones, pero que por sistema u otra razón, no se pueda presentar en “0” las comisiones.

Ventas con valores en negativo.

En un primer momento se creyó que los valores negativos en 18 registros de ventas (monto vendido y precio) podrían ser por devoluciones, pero en vistas de que tanto la cantidad vendida como la comisión comercial eran valores positivos, y que además la multiplicación entre cantidad y precio (siendo positivo), en todos los casos da el resultado del monto (también positivo), se decidió que podía ser un error de tipeo y que tanto precios como los montos correspondían a valores positivos y ventas efectivamente realizadas, por lo que cambiando el signo a dichos registros se solucionaría.



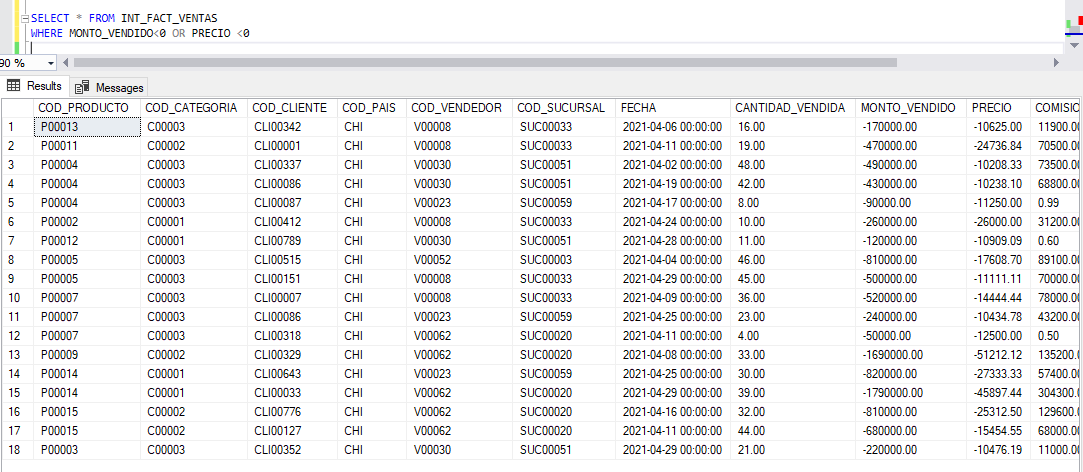


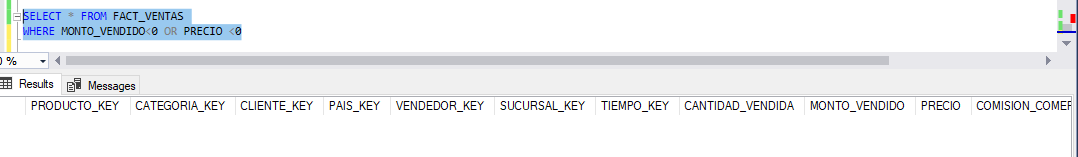
Por este motivo y porque las fechas están desordenadas es que se utilizan vistas para generar la tabla FACT\_VENTAS. En primer lugar, al cargar en un Insert las tablas (en este caso la FACT\_VENTAS) no se pueden utilizar en la cláusula un ORDER BY, y en razón de que las ventas no venían ordenadas por fecha, se decidió ordenarlas por fecha a través de una VISTA (VIEW).

Por lo que en primer lugar se genera una vista de la tabla INT\_FACT\_VENTAS, con un top de 12.000 en este caso, ya que sin TOP no funciona el INSERT con ORDER BY. Consecuentemente, se insertan los datos en la tabla final FACT\_VENTAS desde esta vista. Esto se hace esta sola vez, ya que los próximos registros vendrán ordenados por fecha y al cargarlos a través del sp, quedarán ordenados en la tabla final, debajo de los ya cargados.

Cabe aclarar, que además de esta vista, también se realiza otra intermedia que considere las ventas negativas consideradas anteriormente (los 18 registros), para que finalmente queden todas en positivas. De esta manera, primero se cargan el total de registros de la vista que vienen de la tabla INT\_FACT\_VENTAS, luego se eliminan las negativas y luego se cargan las positivas (que se transforman a través de otra vista, en este caso llamada VENTASNEG.

De esta manera, pasamos de los 18 registros con números en negativo de la tabla INT\_FACT\_VENTAS, a ningún registros en negativo de la tabla FACT\_VENTAS final.





Integridad de los Datos.

En cuanto a la integridad de datos a las “objeciones” ya las consideramos, también se puede sumar que al cargar los datos a las tablas STG a través de la herramienta de SQL los demás campos figuraban nulos, por lo que tuve que eliminarlos. Sumado a las ya mencionadas, como valores negativos, que los campos de la tabla de Ventas no tienen el mismo formato, por ejemplo Monto\_Vendido, con Precio y Comision\_Comercial, también las comisiones casi nulas. Todo esto se podría modificar fácilmente o realizar una documentación al respecto de qué y por qué.

Otro aspecto llamativo, fue que no tengan ordenadas las ventas por fecha.

Pero más allá de eso, los datos parecen confiables, no tenían duplicados de ningún tipo y estaban relacionados a la actividad de manera clara, aunque sí hizo falta analizarlos bastante en un comienzo.

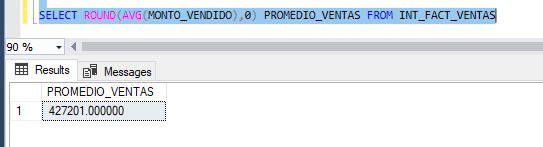
Una vez que ya se encuentra la carga realizada en el DW, hay muchos de estos tipos de inconsistencias que se pueden subsanar, por ejemplo temas de formato de datos, ventas ordenadas por fecha y que deban cargarse de determinada manera.

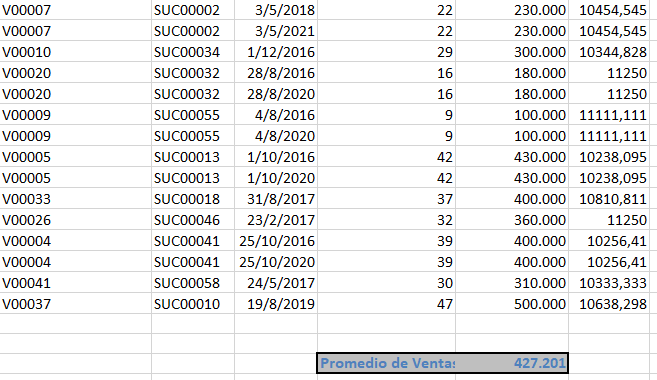
Métricas de Ventas.

Obtener:

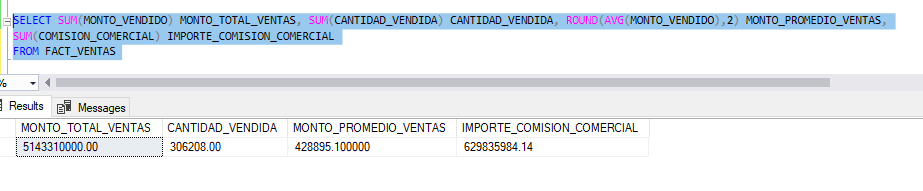
* Monto Total de Ventas ($).
* Cantidad Vendida (#).
* Monto Promedio de Ventas. ($).
* Importe Comisión Comercial ($).
* Cantidad de Clientes (#).

Varias de éstas métricas ya están validadas anteriormente en este documento cuando se validó la carga comparada con el archivo de EXCEL, solamente nos quedó el Promedio de Ventas.

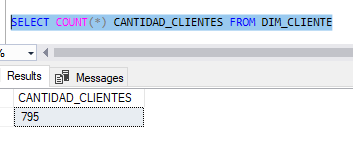


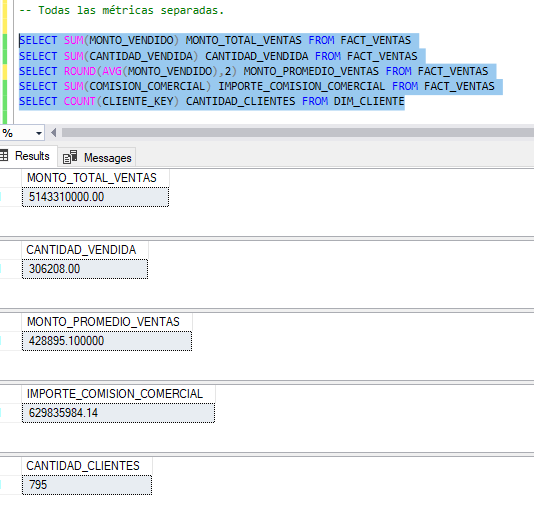


En este caso, se observa que la coincidencia se da en la Tabla INT\_FACT\_VENTAS, ya que luego se cambió el signo de las ventas, por lo que cada uno de estos datos (relacionados a la venta) se cambiaron y por ende se mostrarán los nuevos, todos juntos, a continuación:



El caso de cantidad de clientes, obviamente no cambia y sigue coincidiendo con el Excel.





A continuación y a modo de cierre para que le sirva a la empresa, se dejarán algunas consultas que sirven para demostrar el funcionamiento del DataWarehouse y que pueda también servir de ayuda para la empresa y tal vez la toma de decisiones de los altos manos de la misma.

