**Universidad Don Bosco**



**Docente:** Ing. Alexander Alberto Siguenza

**Materia**: Datawarehouse y Minería de Datos

**Presentado por**:

Ruiz Saz, Carlos José RS181977

Rosales Mendoza, Patrick Ernesto RM181976

**Ciclo:** 02-2020

Soyapango, viernes 29 de agosto del 2020

Contenido

[Capturas desafio I 3](#_Toc49418322)

[Problema I 3](#_Toc49418323)

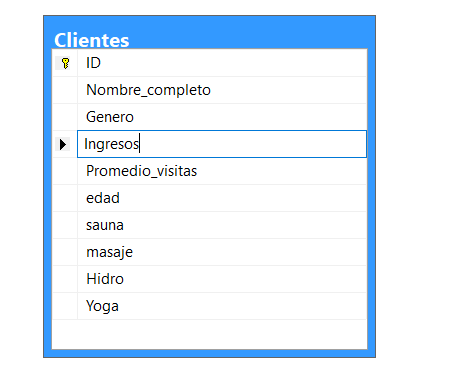
[Problema II 5](#_Toc49418324)

[Problema III 8](#_Toc49418325)

# Capturas desafio I

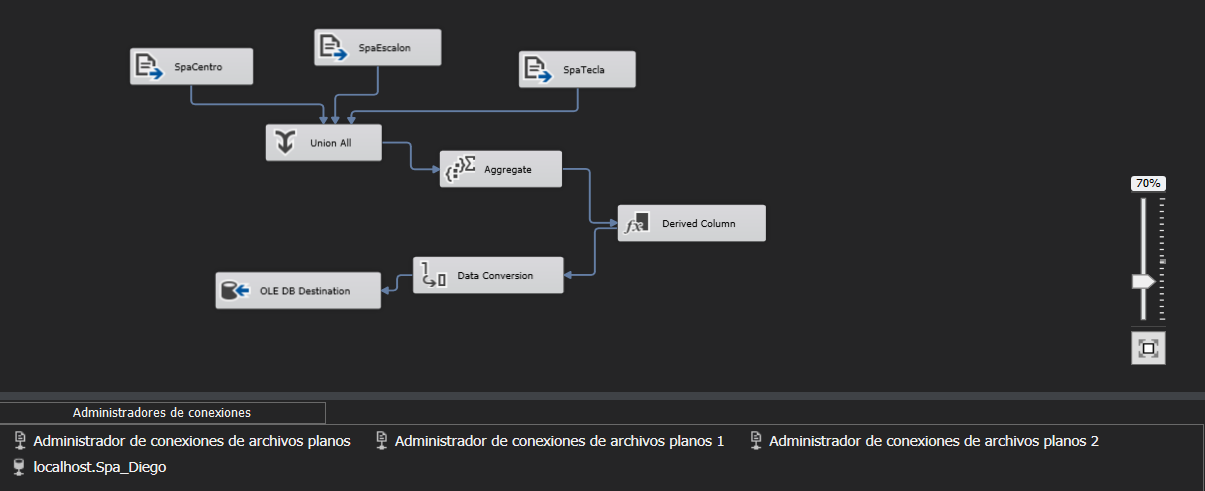
## Problema I

**Diagrama de la base de datos**

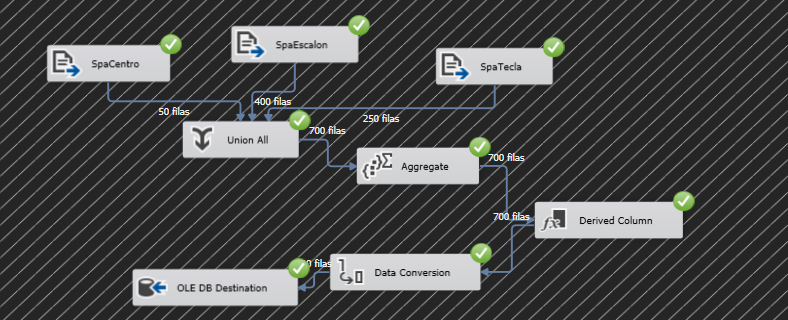


**Flujo de datos**

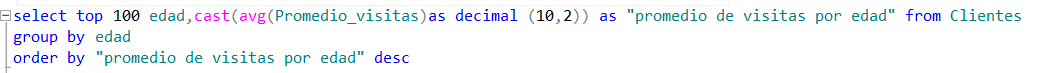
Se ha utilizado derived column para cambiar la columna de sexo, en la cual el 1 ahora será hombre y 0 será mujer.



**Prueba de ejecución de paquete**

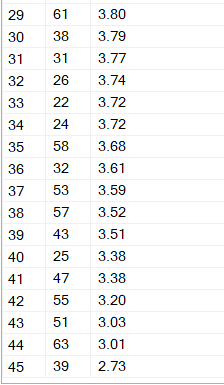
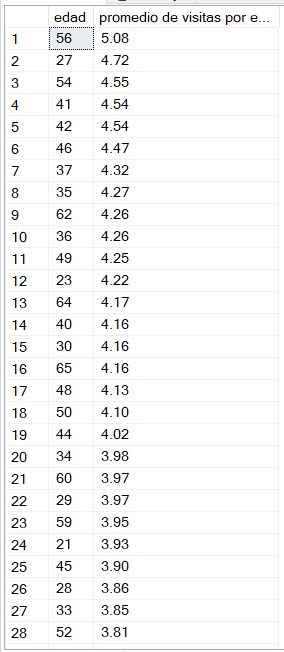


**Primera consulta (promedio de visitas según edad)**



Para identificar a los grupos y características que se tienen necesitamos saber tanto las edades y el sexo que tiene un mayor promedio de visita, con el fin de que el spa pueda dar su campaña de fidelidad.

Al ejecutar la consulta anterior tenemos lo siguiente:

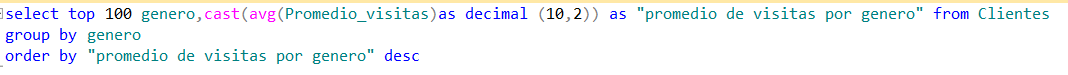


Si analizamos la siguiente información las edades que tienen un mayor promedio de visita del spa son personas de:

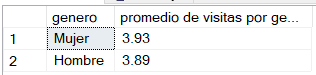
* 56 años
* 27 años
* 54 años
* 41 años
* 42 años

Son personas con una considerable edad, en la cual pueden tener una vida muy ocupada, ya sea por sus trabajos, responsabilidades personales o familiares, por lo que este grupo de personas tienen un nivel de estrés elevado, por lo cual asisten al spa con el objetivo de relajarse.

**Segunda consulta (promedio de visitas según genero)**



Al ejecutar esta consulta tenemos el siguiente resultado:

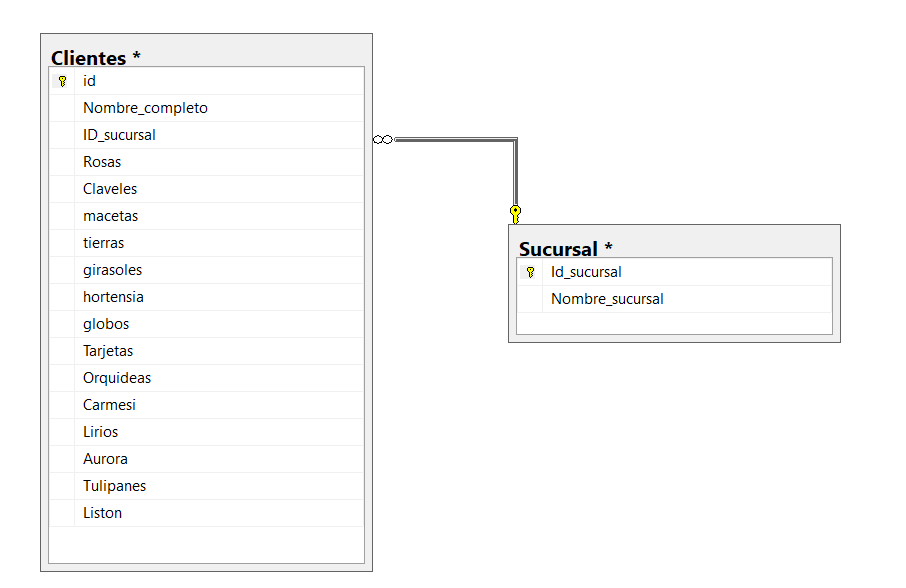


De estos promedios podemos afirmar que van más mujeres que hombres al spa, sin embargo, el promedio de los hombres no es tan inferior al de las mujeres.

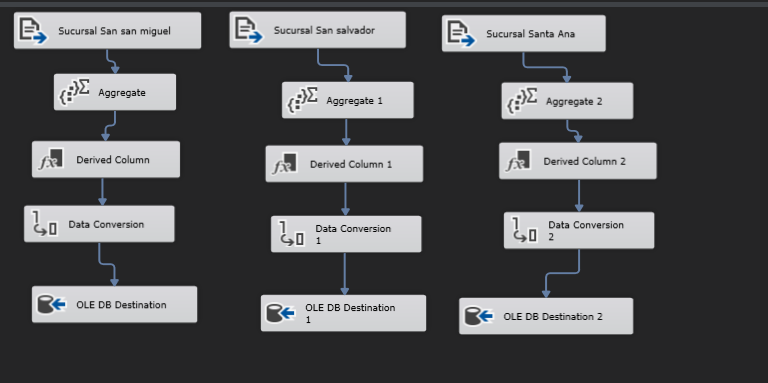
Con estas 2 conclusiones podemos afirmar: que en el spa para su campaña de fidelización debería enfocarse en personas con un estilo de vida ocupada, es decir tengan altos niveles de estrés ya sea hombre o mujer, ya que sus 2 promedios son casi similares.

## Problema II

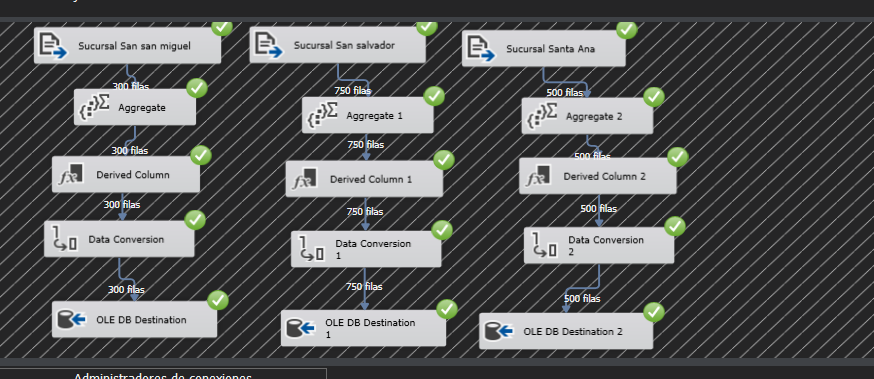
Diagrama de la base de datos



**Flujo de datos**

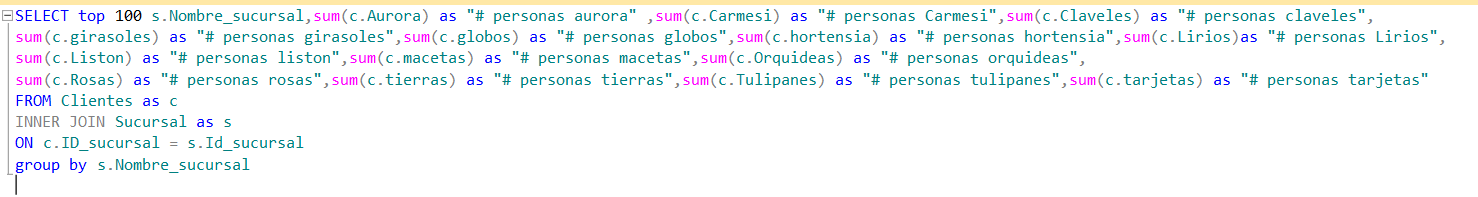


**Ejecución del paquete**

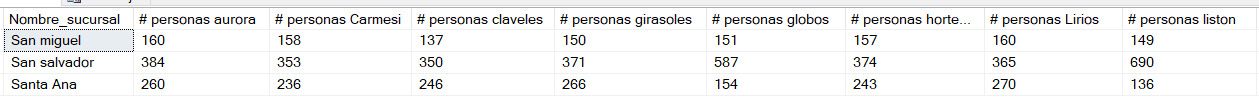


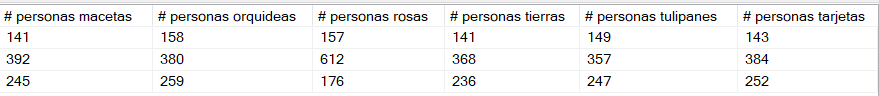
**Estudio por departamento**

Para saber los productos que más personas han comprado ejecutamos la siguiente consulta:



Al ejecutarla da los siguientes resultados



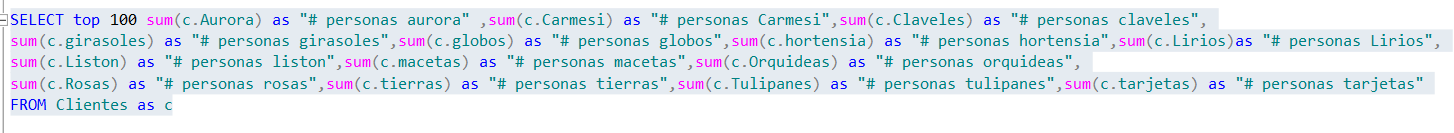


Como podemos observar en la tabla anterior nos muestra el número de personas que compran cada uno de los productos en las diferentes sucursales.

En la tabla anterior podemos observar que la sucursal de San salvador es la que más ventas ha tenido, en todos los productos.

Estudio por país

Para saber los productos que más se venden, usamos la siguiente consulta:



Y nos mostraría el siguiente resultado:



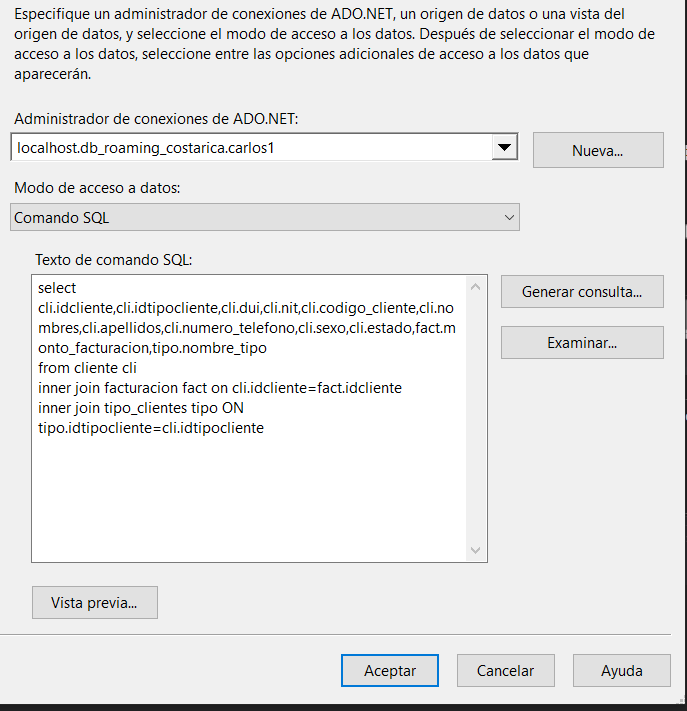


En este caso a nivel de país, los 3 productos que más se venden serían:

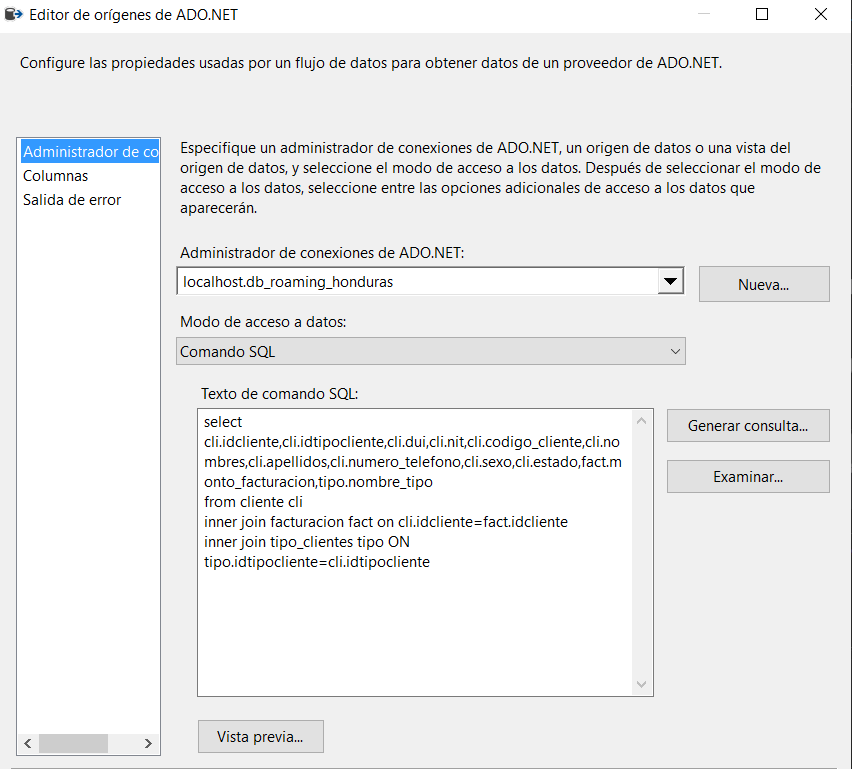
* Listones
* Rosas
* Globos

## Problema III

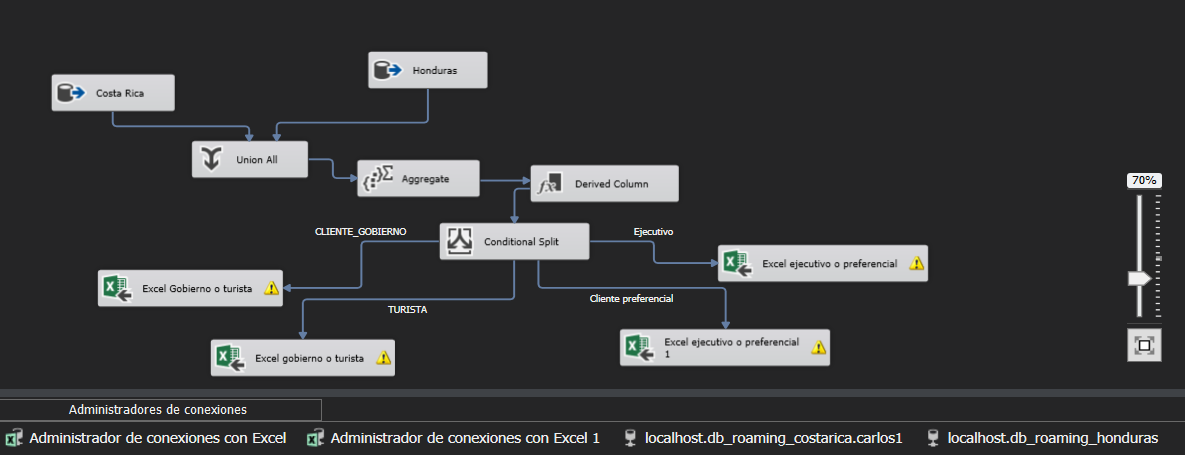
**Conexión a base de datos de costa rica**



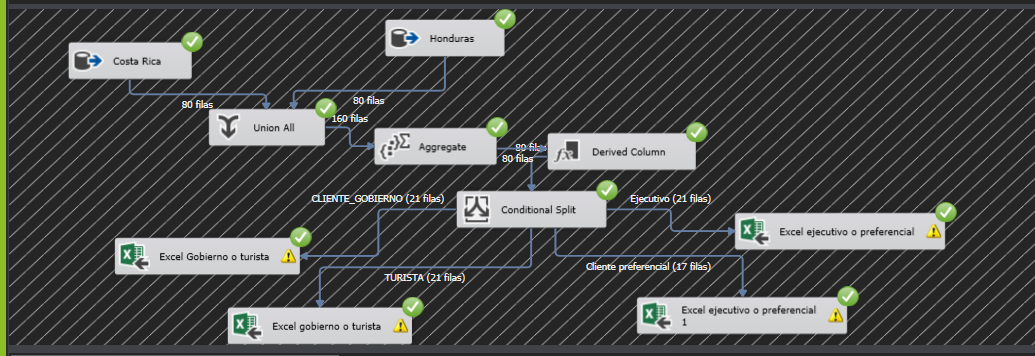
**Conexión a base de datos de honduras**



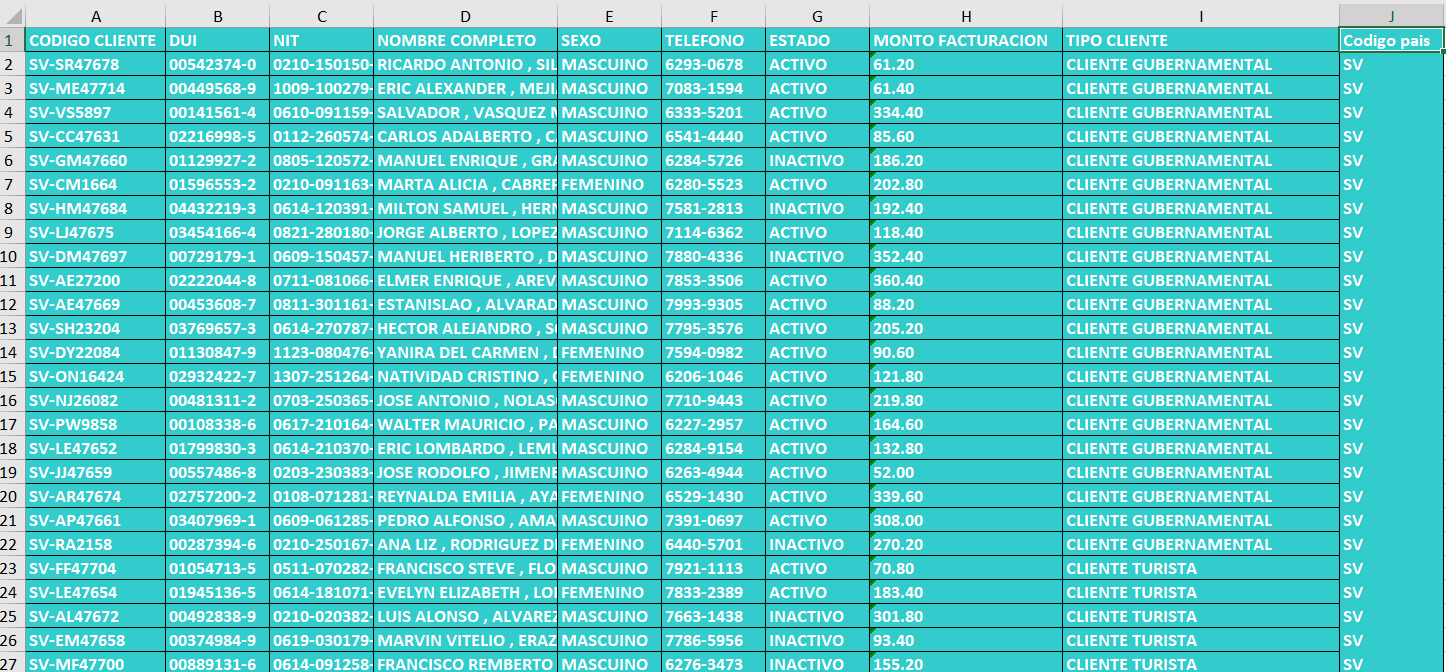
**Flujo de datos**



**Prueba de ejecución**



**Capturas de los archivos Excel obtenidos**

**Excel clientes de gobierno y turista**

**Excel clientes ejecutivos y preferenciales**

