**Universidad Don Bosco**



**Ingeniería en computación**

**Tema:** Primer Desafío Practico

**Docente:** Ing. Alexander Alberto Sigüenza

**Materia:** Datawarehouse y minería de datos 01L

**Presentado por:**

Pineda Valle, Erick Alberto PV181965

Tobar Beltrán, Marvin Wilfredo TB181979

**Índice**

[Ejercicio 1 3](#_Toc49543508)

[Ejercicio 2 9](#_Toc49543509)

[Ejercicio 3 14](#_Toc49543510)

# Ejercicio 1

Creación de la base de datos SpaDiego

use master;

CREATE DATABASE SpaDiego

go

use SpaDiego

CREATE TABLE Sucursales

(

id\_cliente int IDENTITY(1,1) PRIMARY KEY,

nombre varchar(200) NOT NULL,

sexo varchar(25),

ingresos decimal,

promedio\_visitas decimal,

edad int,

sauna tinyint,

masaje tinyint,

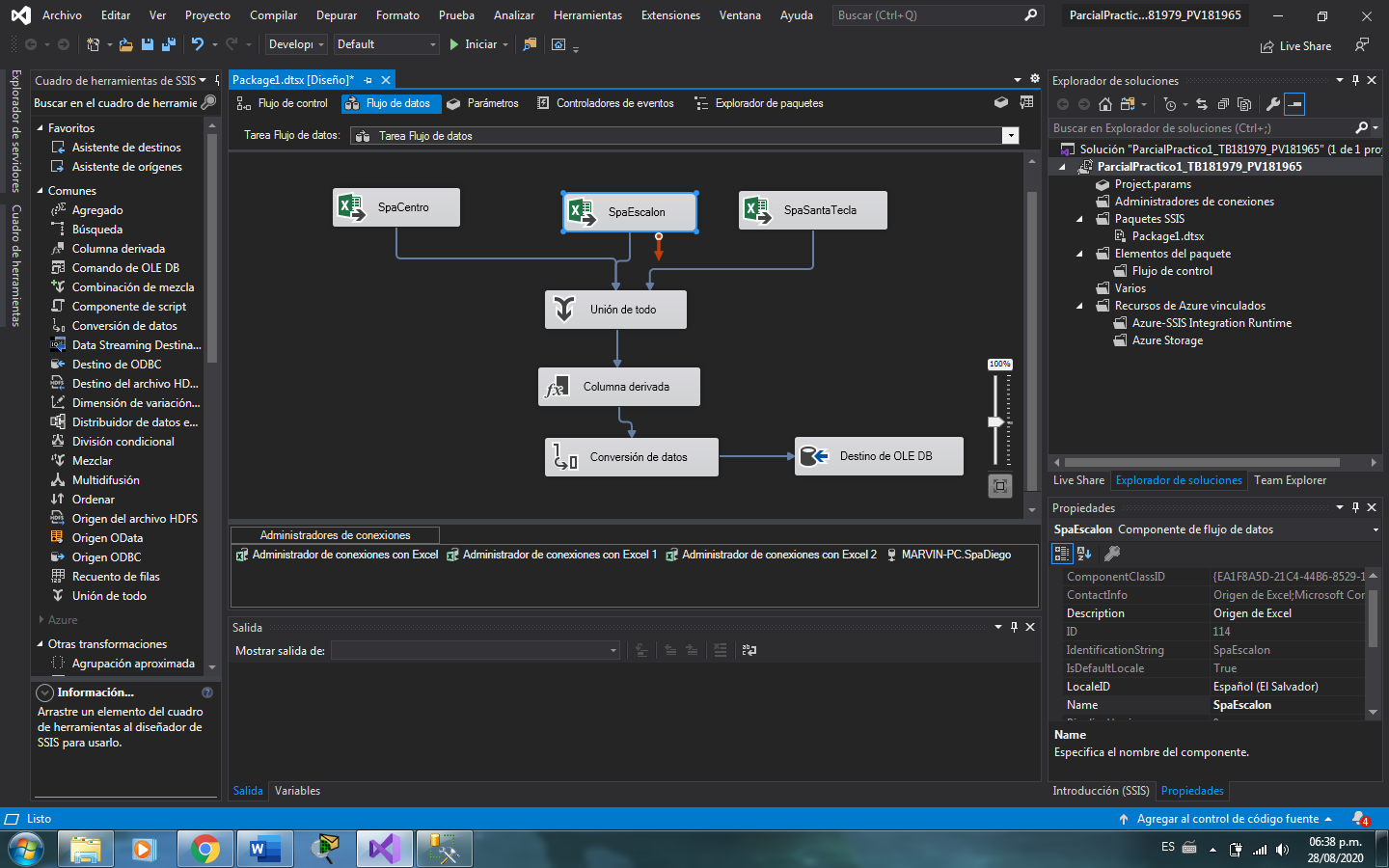
hidro tinyint,

yoga tinyint

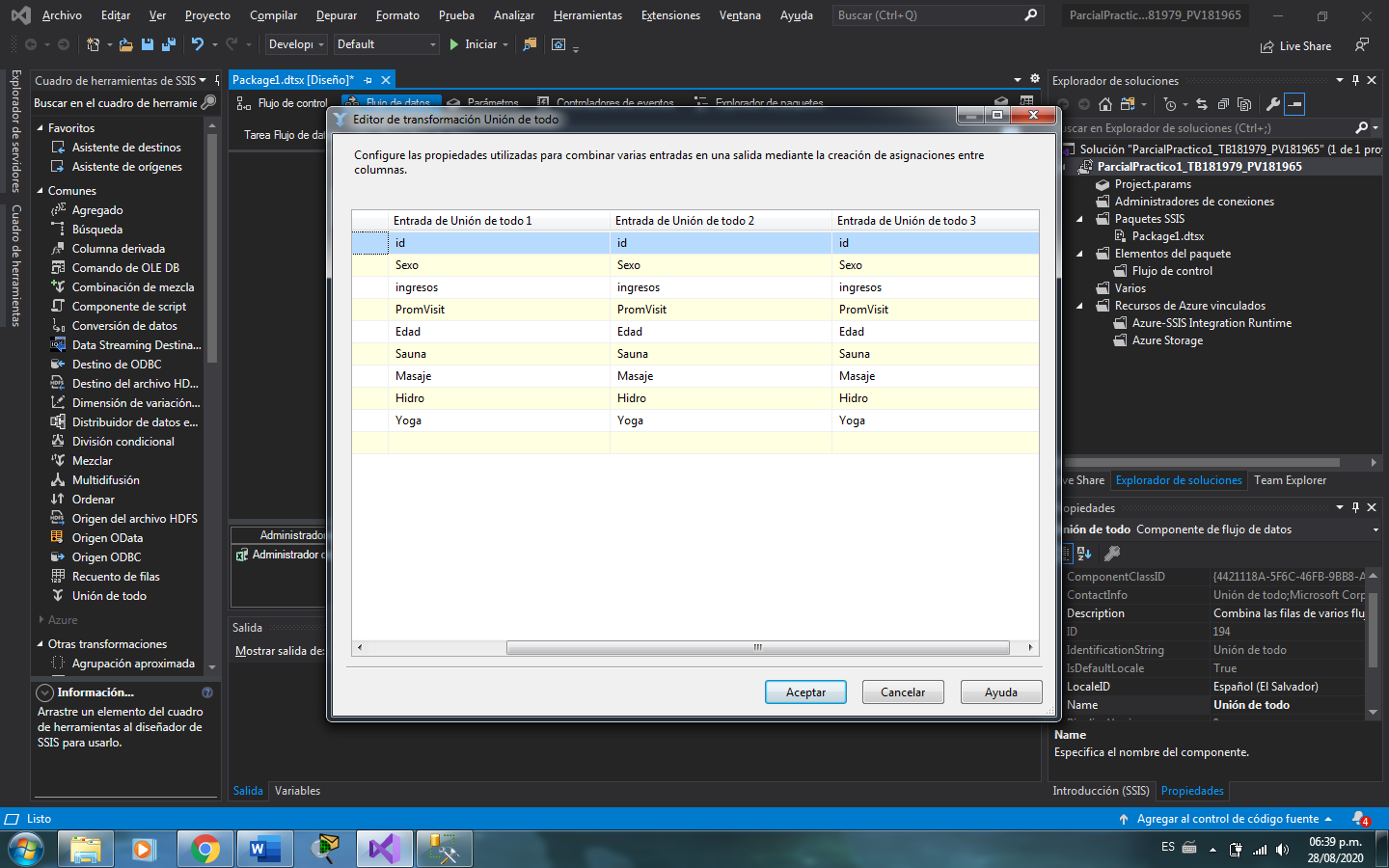
)

go

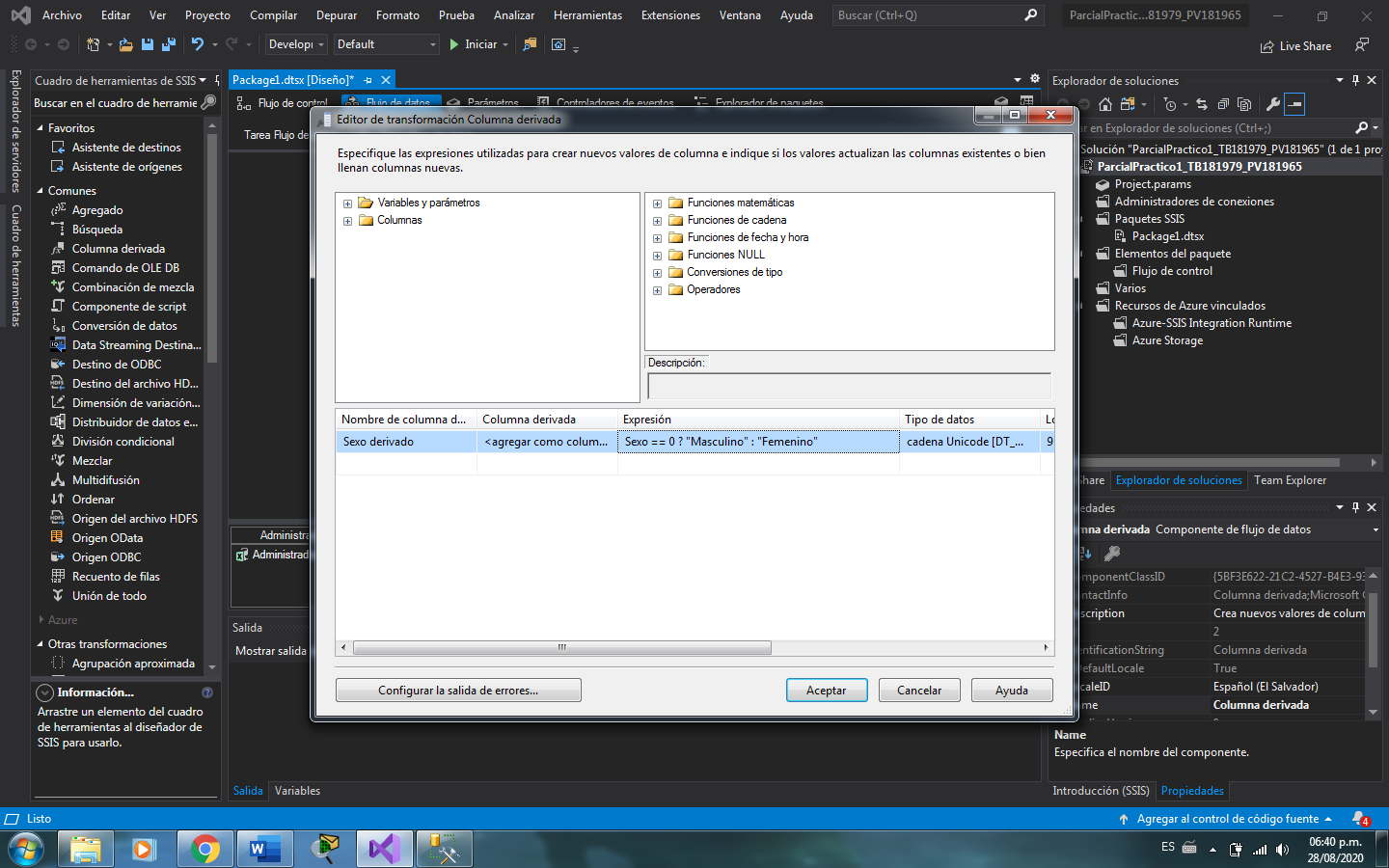
Flujo de datos



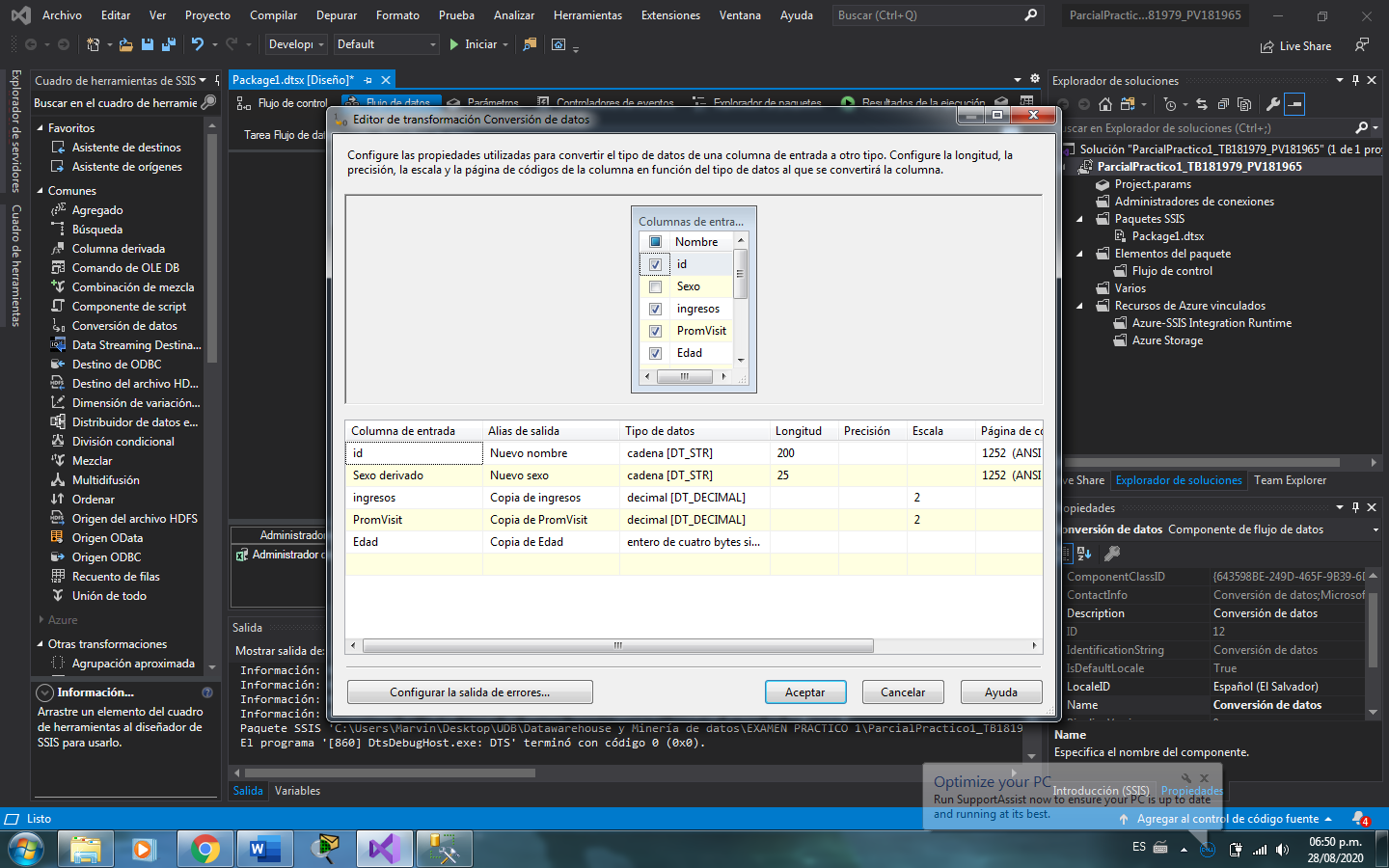
Unión de todo



Columna derivada



Conversión de datos



Consultas SQL

--PERSONAS CUYO PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON ALGUN SERVICIO EXTRA (SAUNA, MASAJE, HIDRO, YOGA)

SELECT \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna=1 or masaje=1 or hidro=1 or yoga=1)

--LOS 100 CLIENTES QUE CUMPLEN CON LOS REQUISITOS ANTERIORES

--Y ADEMAS ORDENADOS CON INGRESOS DE MAYOR A MENOR

SELECT TOP 100 \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna=1 or masaje=1 or hidro=1 or yoga=1)

ORDER BY ingresos DESC

--TODOS LOS REQUISITOS ANTERIORES

--Y ADEMAS QUE USEN AL MENOS A DOS SERVICIOS

SELECT TOP 100 \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

ORDER BY ingresos DESC

--CANTIDAD DE MUJERES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'MUJERES'

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Femenino'

--CANTIDAD DE HOMBRES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'HOMBRES'

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Masculino'

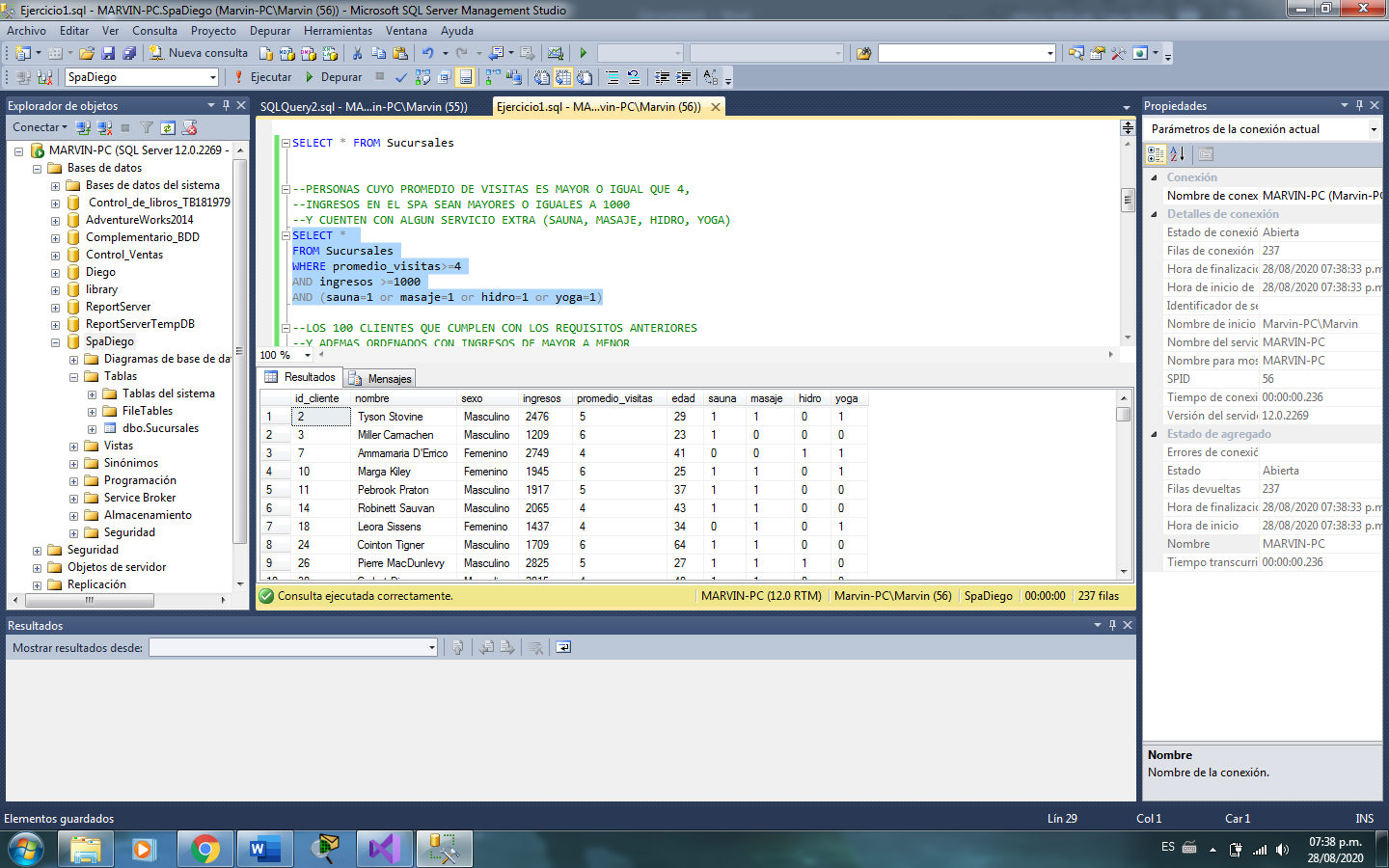
--SERVICIO QUE MAS UTILIZAN

SELECT COUNT(sauna) AS 'SAUNA' FROM Sucursales WHERE sauna=1 --284

SELECT COUNT(masaje) AS 'MASAJE' FROM Sucursales where masaje=1 --366

SELECT COUNT(hidro) AS 'HIDRO' FROM Sucursales where hidro=1 --271

SELECT COUNT(yoga) AS 'YOGA' FROM Sucursales where yoga=1 --361



--LOS 100 CLIENTES QUE CUMPLEN CON LOS REQUISITOS ANTERIORES

--Y ADEMAS ORDENADOS CON INGRESOS DE MAYOR A MENOR

SELECT TOP 100 \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna=1 or masaje=1 or hidro=1 or yoga=1)

ORDER BY ingresos DESC

--LOS 100 CLIENTES QUE CUMPLEN CON LOS REQUISITOS ANTERIORES

--Y ADEMAS ORDENADOS CON INGRESOS DE MAYOR A MENOR

SELECT TOP 100 \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna=1 or masaje=1 or hidro=1 or yoga=1)

ORDER BY ingresos DESC

--TODOS LOS REQUISITOS ANTERIORES

--Y ADEMAS QUE USEN AL MENOS A DOS SERVICIOS

SELECT TOP 100 \*

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

ORDER BY ingresos DESC

--CANTIDAD DE MUJERES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'MUJERES'

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Femenino'

--CANTIDAD DE HOMBRES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'HOMBRES'

FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Masculino'

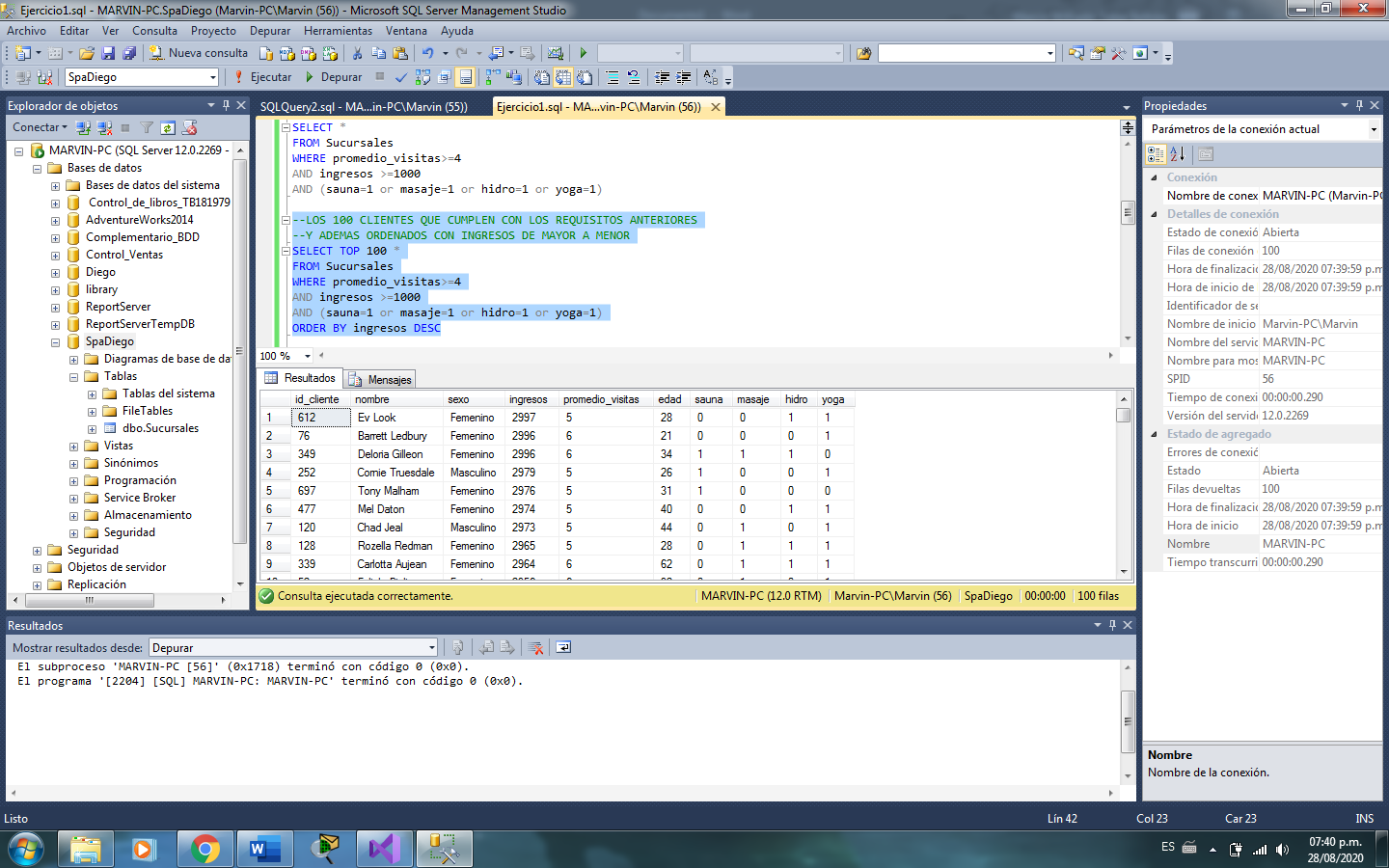
--SERVICIO QUE MAS UTILIZAN

SELECT COUNT(sauna) AS 'SAUNA' FROM Sucursales WHERE sauna=1 --284

SELECT COUNT(masaje) AS 'MASAJE' FROM Sucursales where masaje=1 --366

SELECT COUNT(hidro) AS 'HIDRO' FROM Sucursales where hidro=1 --271

SELECT COUNT(yoga) AS 'YOGA' FROM Sucursales where yoga=1 --361



--TODOS LOS REQUISITOS ANTERIORES

--Y ADEMAS QUE ASISTAN AL MENOS A DOS SERVICIOS

SELECT TOP 100 \*

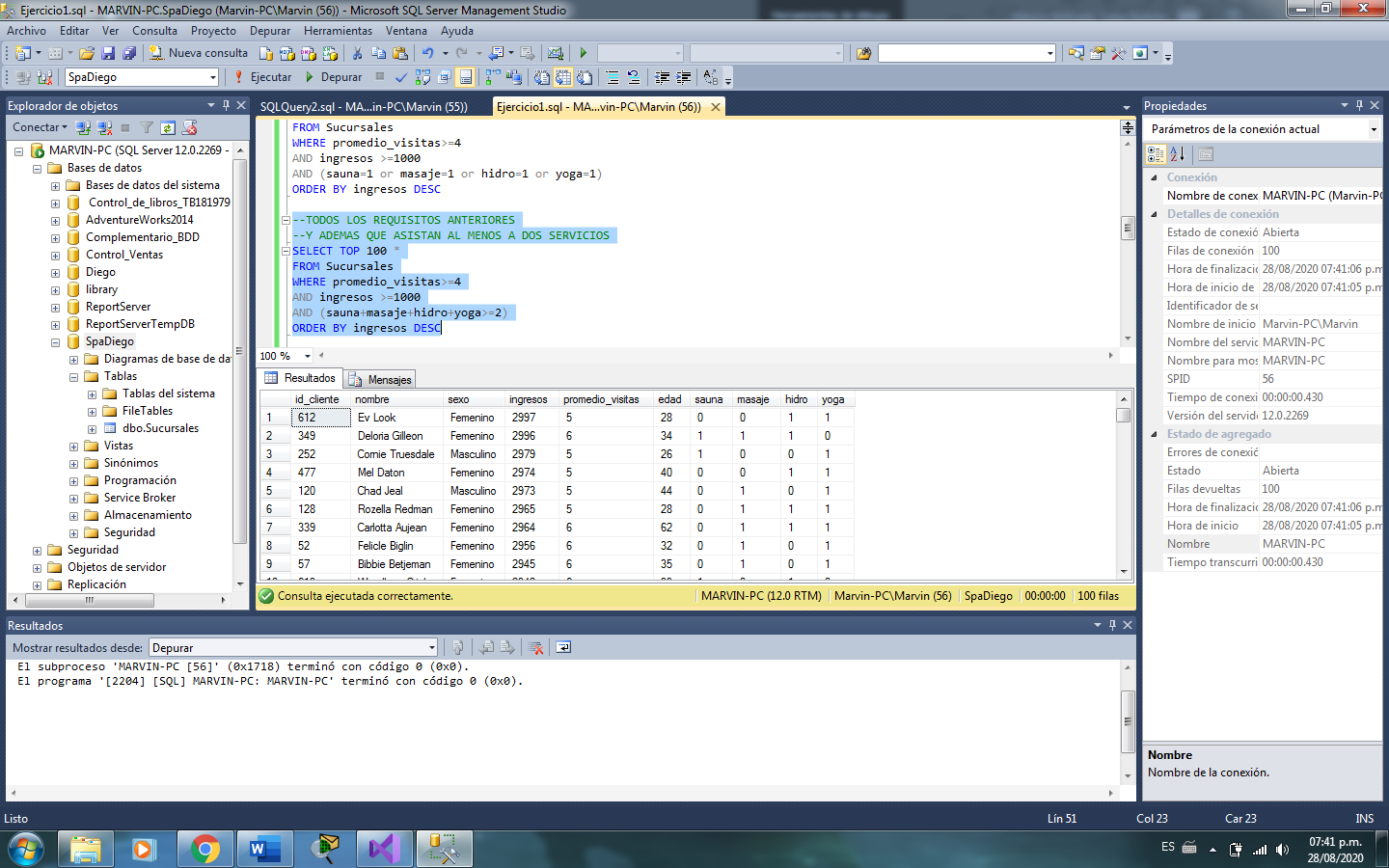
FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

ORDER BY ingresos DESC



--CANTIDAD DE MUJERES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'MUJERES'

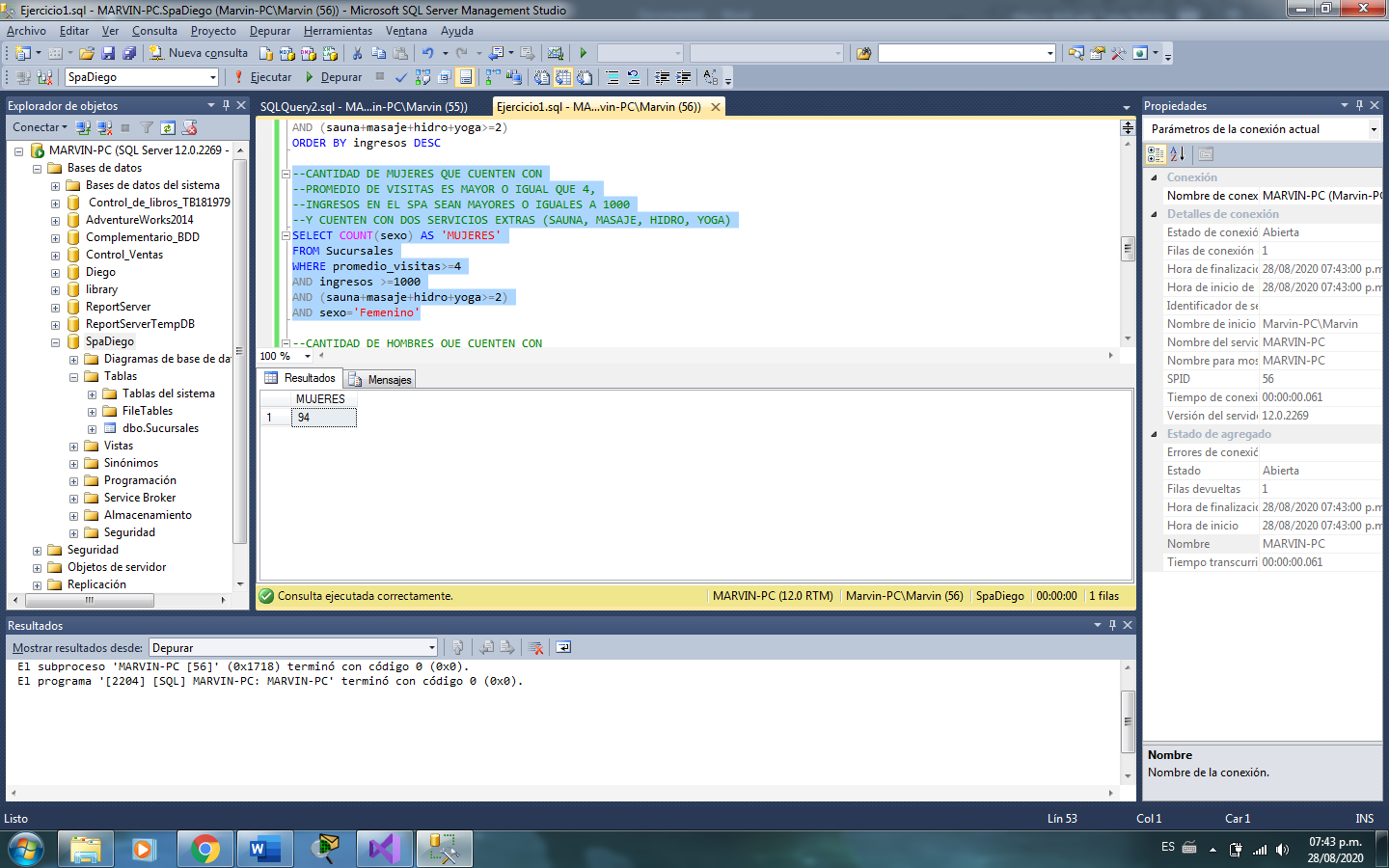
FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Femenino'



--CANTIDAD DE HOMBRES QUE CUENTEN CON

--PROMEDIO DE VISITAS ES MAYOR O IGUAL QUE 4,

--INGRESOS EN EL SPA SEAN MAYORES O IGUALES A 1000

--Y CUENTEN CON DOS SERVICIOS EXTRAS (SAUNA, MASAJE, HIDRO, YOGA)

SELECT COUNT(sexo) AS 'HOMBRES'

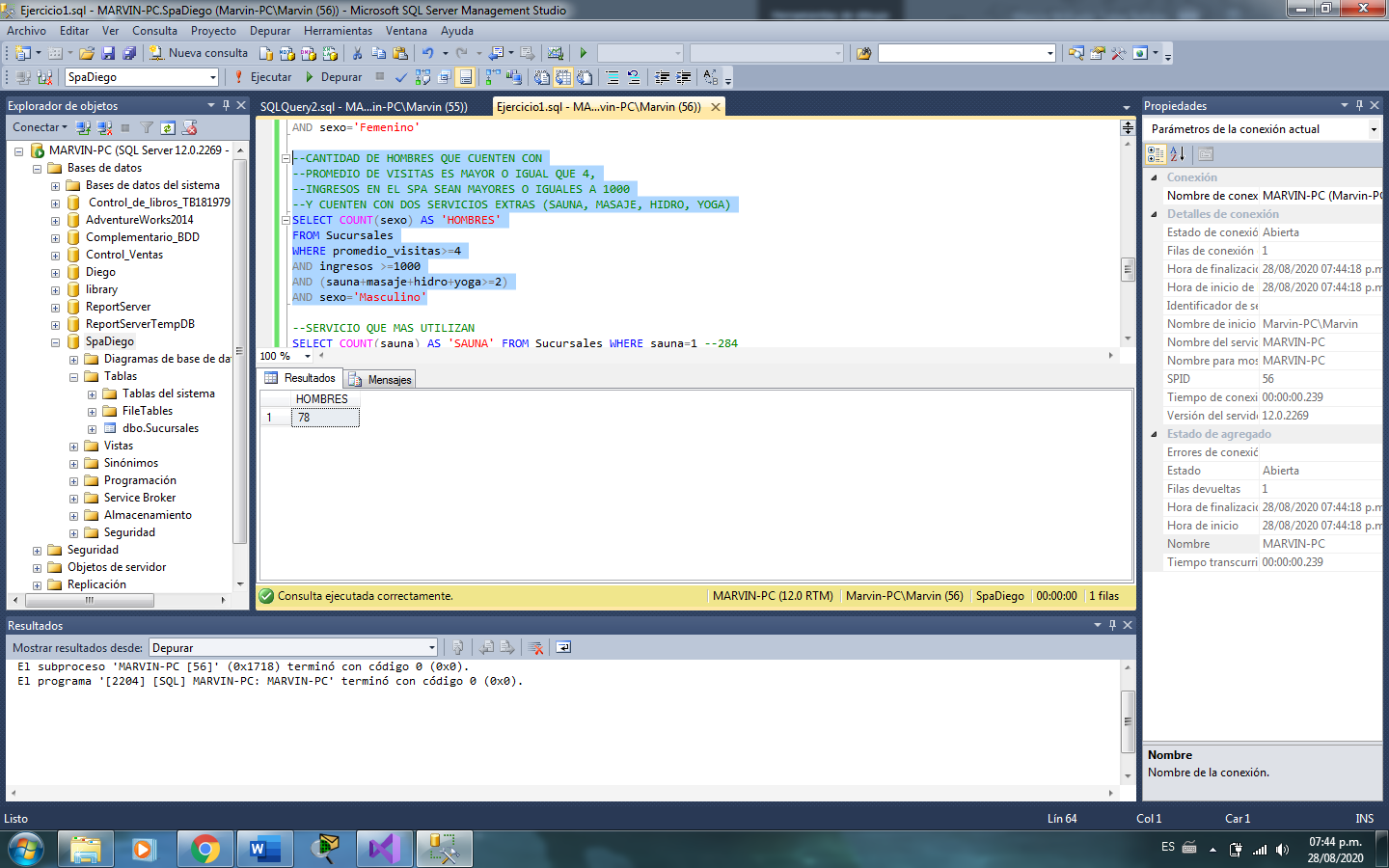
FROM Sucursales

WHERE promedio\_visitas>=4

AND ingresos >=1000

AND (sauna+masaje+hidro+yoga>=2)

AND sexo='Masculino'



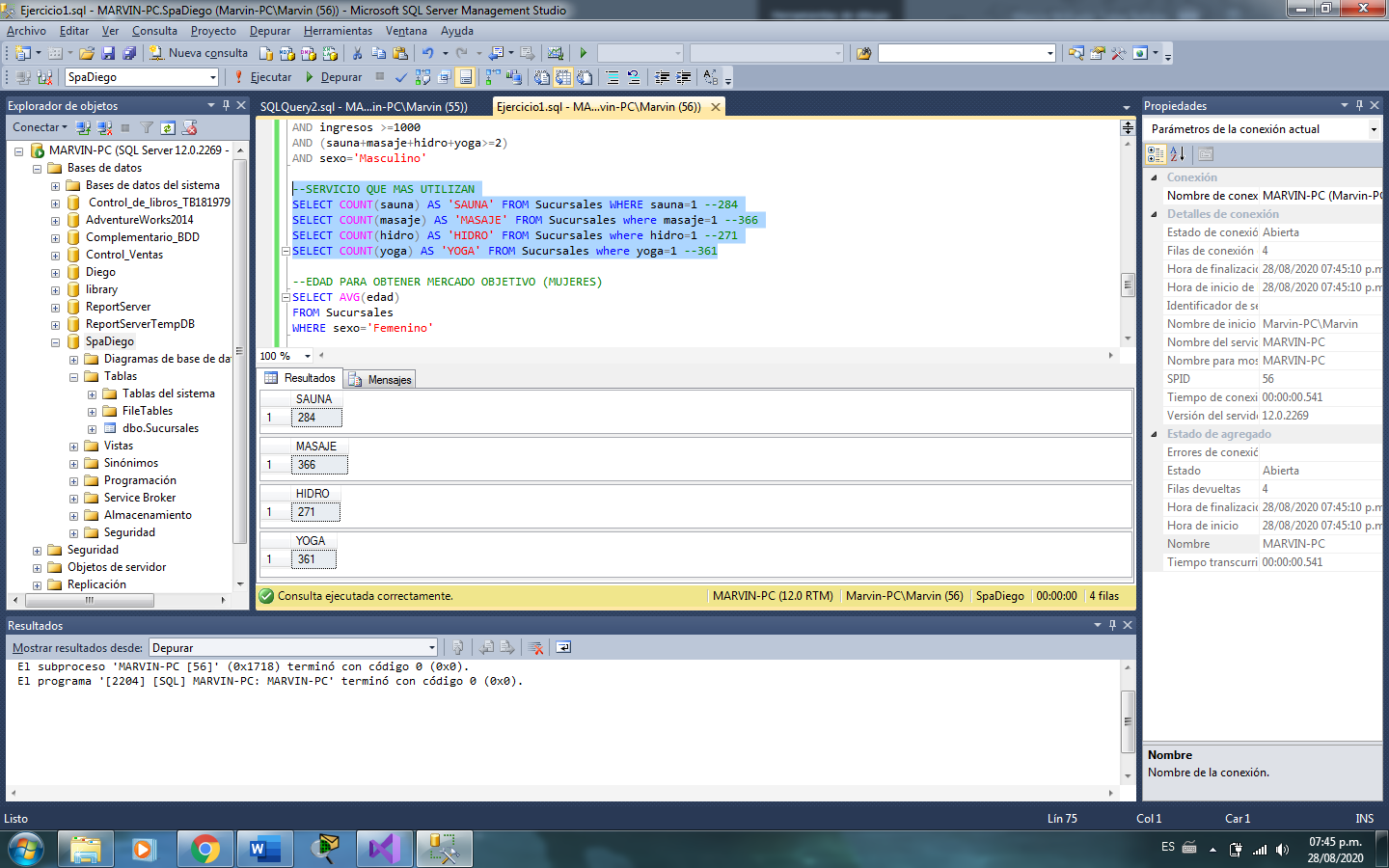
--SERVICIO QUE MAS UTILIZAN

SELECT COUNT(sauna) AS 'SAUNA' FROM Sucursales WHERE sauna=1 --284

SELECT COUNT(masaje) AS 'MASAJE' FROM Sucursales where masaje=1 --366

SELECT COUNT(hidro) AS 'HIDRO' FROM Sucursales where hidro=1 --271

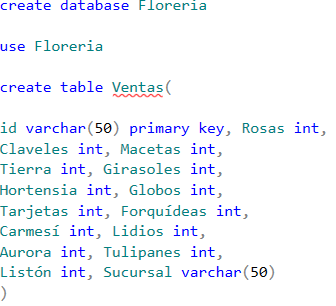
SELECT COUNT(yoga) AS 'YOGA' FROM Sucursales where yoga=1 --361



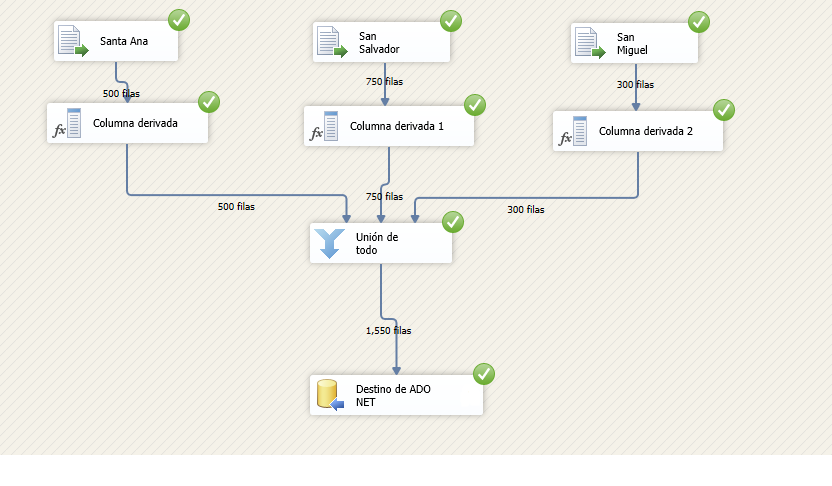
# Ejercicio 2

La Floristería "Fiorella" quiere saber cómo se compran sus productos, y tiene la data de tres departamentos del país, por lo cual les pide su opinión sobre qué productos sobresalen, que combinaciones son mejores y quieren este estudio por departamento y también por país.

Primero debemos crear una base de datos donde serán exportados los datos del archivo csv.



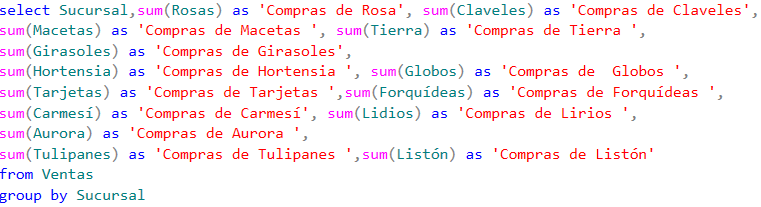
Ahora, para el ETL en SQL Data Tools quedará de la siguiente manera:



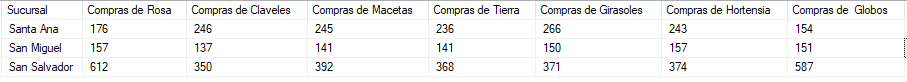
El diagrama muestra las tres entradas de datos (Santa Ana, San Salvador, San Miguel), un control para agregar una nueva columna a las que serán cargadas, un control para unir toda la información y finalmente, un control para enviar los datos completos a una tabla en SQL Server (BDD creada anteriormente).

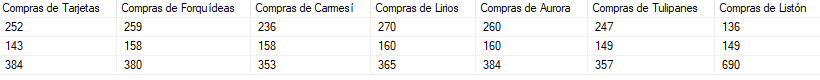
Una vez los datos se encuentren en la base de datos y tabla deseados, procedemos a realizar las consultas para obtener la siguiente información:

Productos que sobresalen en las sucursales:

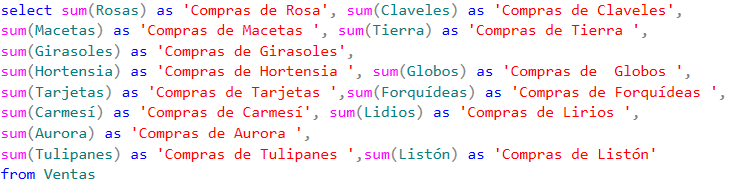


En la anterior consulta, se realiza el “conteo” de cuántas veces un producto fue comprado (un cliente decidió comprar de ese producto), el resultado se encuentra agrupado por las sucursales para presentar la información por departamento. El resultado:





Productos que sobresalen en todo el país:



En este apartado, la consulta queda de la misma forma que el punto “a” con la diferencia que no se agrupa los datos por departamento para presentar la información de todo el país.



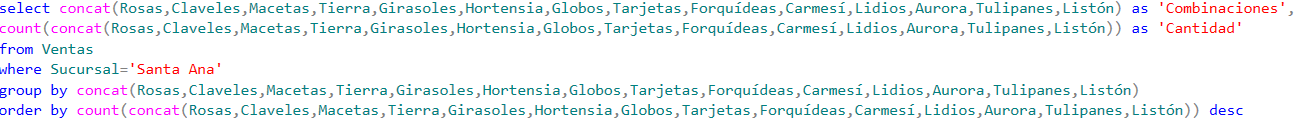
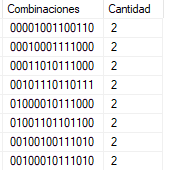


Las mejores combinaciones por sucursal.

En este apartado se realiza una búsqueda por combinación, se entiende “0” cuando el cliente no ha comprado dicho producto y “1” cuando si lo compra.

El orden de los productos es: Rosas, Claveles, Macetas, Tierra, Girasoles, Hortensia, Globos, Tarjetas, Orquídeas, Carmesí, Lidios, Aurora, Tulipanes y Listón.

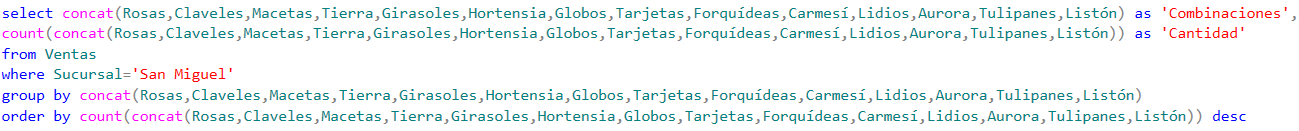
*Para Santa Ana:*

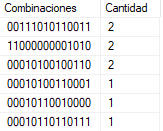


Las mejores combinaciones para Santa Ana son las que se repiten 2 veces. Por ejemplo:

* Girasoles + Tarjetas + Orquídeas + Aurora + Tulipanes.
* Tierra + Tarjetas + Orquídeas + Carmesí + Lidios.
* Tierra + Girasoles + Globos + Tarjetas + Orquídeas + Carmesí + Lidios.

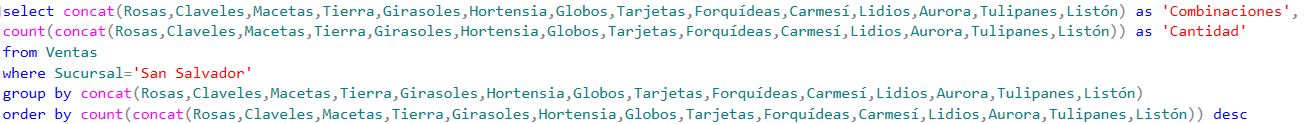
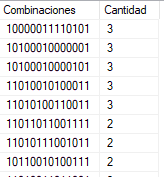
Para San Miguel:



Las mejores combinaciones para San Miguel son:

* Macetas + Tierra + Girasoles + Globos + Orquídeas + Carmesí + Tulipanes + Listón.
* Rosas + Claveles + Lidios + Tulipanes.
* Tierra + Hortensia + Orquídeas Aurora + Tulipanes.

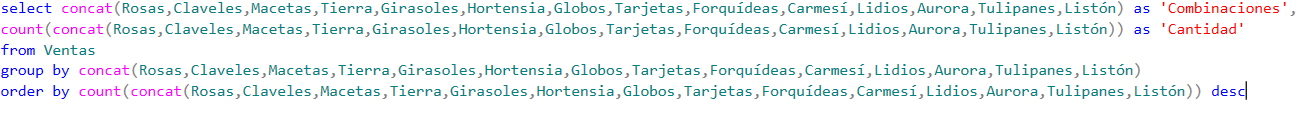
Para San Salvador:

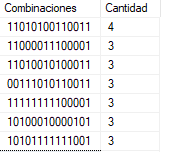


Las mejores combinaciones para San Salvador son, por ejemplo:

* Rosas + Globos + Tarjetas + Orquídeas + Carmesí + Aurora + Listón.
* Rosas + Macetas + Globos + Listón.
* Rosas + Macetas + Globos + Aurora +Listón.

Para todo el país:

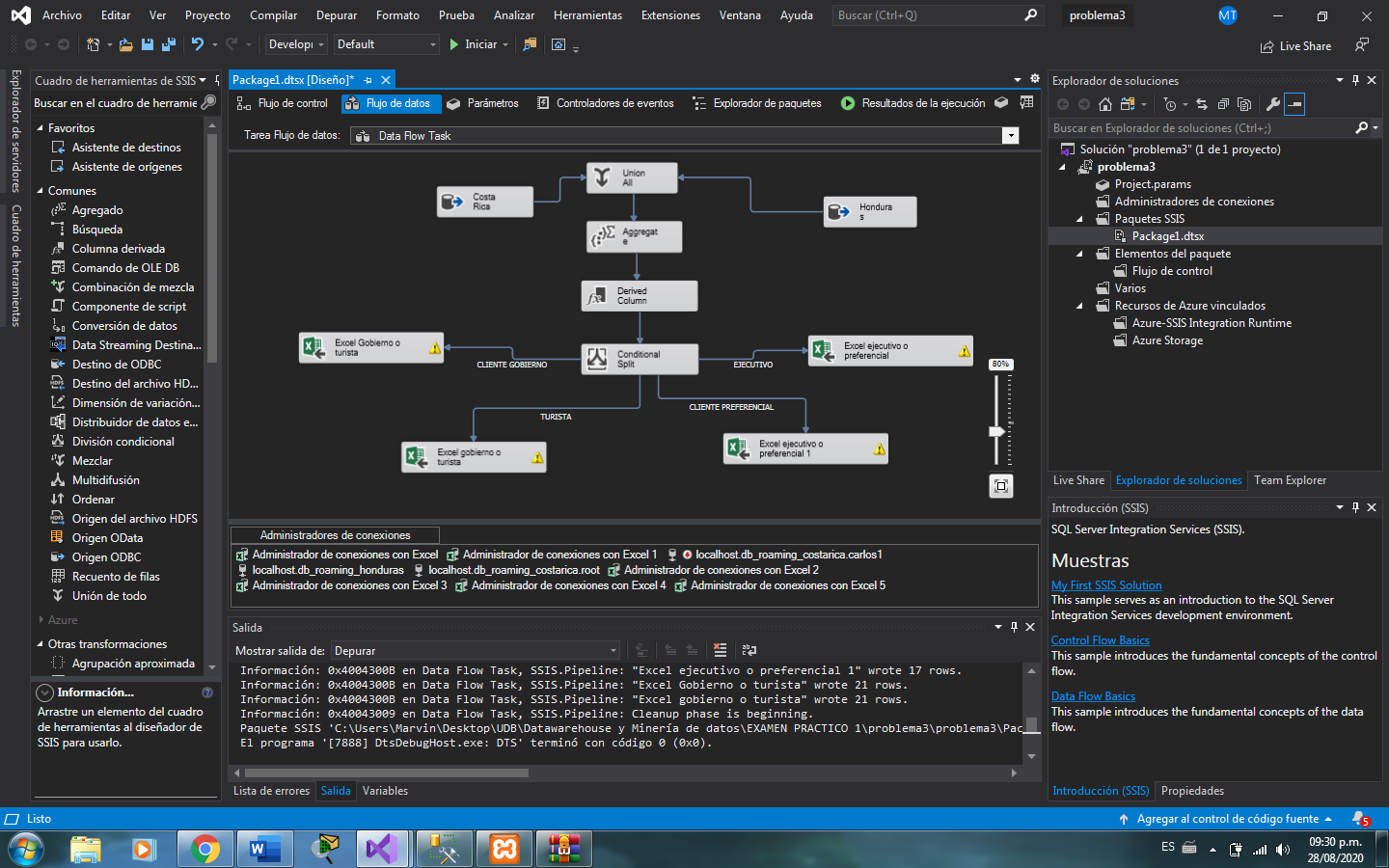


La mejor combinación para todo el país es:

Rosas + Claveles + Tierra + Hortensia + Orquídeas, Carmesí + Tulipanes + Listón.

# Ejercicio 3

Flujo de datos



ARCHIVOS EXCEL

