



U D B V I R T U A L

Universidad Don Bosco, El Salvador

Nombre del trabajo:

Desafío 1

Materia:

Datawarehouse y Minería de Datos

Docente:

Ingeniera Karens Medrano

Estudiantes:

Sermeño Zetino José Alexander SZ 202008

María Dolores Martínez de León ML202278

Fecha de entrega: 21-09-2021

Ejercicio 1

El Spa, "**Diego**", necesita segmentar sus clientes, para realizar una campaña de fidelización, y le pide a usted que efectué un análisis de sus tres sucursales, que defina cuantos grupos y que características tienen.

Desarrollo

Asistente para importación y exportación de SQL Server

Seleccionar un origen de datos
Seleccione el origen del que se copiarán los datos.

Origen de datos:

Nombre del servidor:

Autenticación

☒ Usar autenticación de Windows

☐ Usar autenticación de SQL Server

Nombre de usuario:

Contraseña:

Base de datos:

Asistente para importación y exportación de SQL Server

Seleccionar un destino
Especifique dónde desea copiar los datos.

Destino:

Seleccione un archivo y especifique sus propiedades y formato.

Nombre de archivo:

Configuración regional: ☐ Unicode

Página de códigos:

Formato:

Calificador de texto:

☒ Nombres de columna de la primera fila de datos

Asistente para importación y exportación de SQL Server

Proporcionar una consulta de origen

Escriba la instrucción SQL que seleccionará los datos de la base de datos de origen.

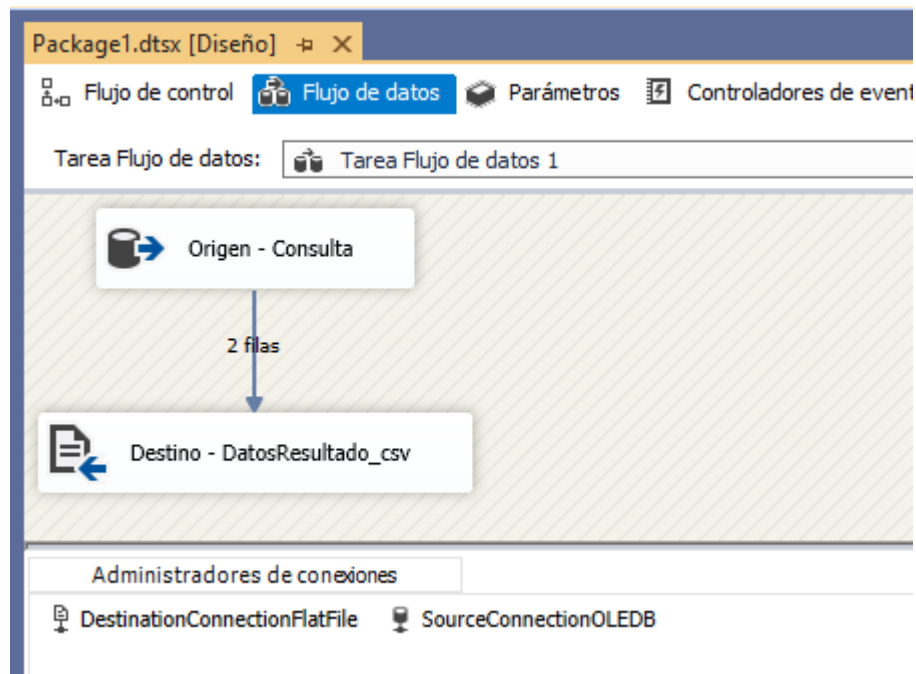
Instrucción SQL:

```
SELECT Sexo As Hombre_Mujer,
       COUNT(Sexo)
FROM SpaSucursales
GROUP BY Sexo;

SELECT
CASE
WHEN Edad BETWEEN 18 AND 25 THEN 'Cliente Joven'
WHEN Edad BETWEEN 26 AND 40 THEN 'Cliente Adulto Joven'
WHEN Edad BETWEEN 40 AND 55 THEN 'Cliente Adulto'
ELSE 'Adulto Mayor'
END AS Categoria,
COUNT(*) AS Cantidad
FROM SpaSucursales
GROUP BY
CASE
WHEN Edad BETWEEN 18 AND 25 THEN 'Cliente Joven'
WHEN Edad BETWEEN 26 AND 40 THEN 'Cliente Adulto Joven'
WHEN Edad BETWEEN 40 AND 55 THEN 'Cliente Adulto'
ELSE 'Adulto Mayor'
END
ORDER BY Categoria DESC;
```

Analizar Examinar...

Help < Back Next > Finish >> Cancel



A1			
	A	B	C
1			
2	Hombre_Mujer	Cantidades	
3	0	396	
4	1	304	
5			
6			

Editor de origen de OLE DB

Configure las propiedades que utiliza un flujo de datos para obtener datos de cualquier proveedor OLE DB.

Administrador de conexiones OLE DB: SourceConnectionOLEDB Nueva...

Modo de acceso a datos: Comando SQL

Texto de comando SQL:

```
SELECT
CASE
WHEN Edad BETWEEN 18 AND 25 THEN 'Cliente Joven'
WHEN Edad BETWEEN 26 AND 40 THEN 'Cliente Adulto Joven'
WHEN Edad BETWEEN 40 AND 55 THEN 'Cliente Adulto'
ELSE 'Adulto Mayor'
END AS Categoria,
COUNT(*) AS Cantidad
FROM SpaSucursales
GROUP BY
CASE
WHEN Edad BETWEEN 18 AND 25 THEN 'Cliente Joven'
WHEN Edad BETWEEN 26 AND 40 THEN 'Cliente Adulto Joven'
WHEN Edad BETWEEN 40 AND 55 THEN 'Cliente Adulto'
ELSE 'Adulto Mayor'
END
ORDER BY Categoria DESC;
```

Vista previa...

Parámetros... Generar consulta... Examinar... Analizar consulta

Aceptar Cancelar Ayuda

Categoría	Cantidad
Cliente Joven	79
Cliente Adulto Joven	251
Cliente Adulto	228
Adulto Mayor	142

	Categoría	Cantidad
1	Ocasional	121
2	Frecuente	229
3	Compro...	134
4	Casi fiel	216

	Categoría	Cantidad
1	1.Servicios Básicos	53
2	2.Servicios Convencionales	230
3	3.Servicios Especiales	283
4	4.Servicio Premium	134

	Oferta	Servicio
1	Hidro	64
2	Otros	70
3	Yoga	72
4	Mas...	210
5	Sauna	284

Conclusiones.

- A partir de la información recopilada por medio de las consultas realizadas se sabe que el número de hombre y mujeres que asisten al Spa puede considerarse equilibrado
- La mayoría de clientes son adultos de un nivel de ingresos considerable.
- El servicio requerido con mayor frecuencia es el Sauna seguido de los masajes.
- Se considera que el ejercicio fue resuelto al 100%

Introducción

El desafío consiste en la resolución de varios ejercicios utilizando las técnicas ELT presentadas en clases, los ejercicios hacen diversas solicitudes sobre el negocio que deberán ser resueltas solicitando las consultas más adecuadas y relevantes posibles.

Se incluirán imágenes como soporte del trabajo realizado para la resolución de cada ejercicio y conclusiones interpretando los datos obtenidos.

Objetivos.

- Interpretar las solicitudes del negocio en cada uno de los ejercicios planteados.
- Utilizar correctamente los recursos ELT disponibles en la plataforma de VISUAL STUDIO.
- Brindar respuestas certeras a partir de los datos extraídos de las muestras de datos provistas para los ejercicios.

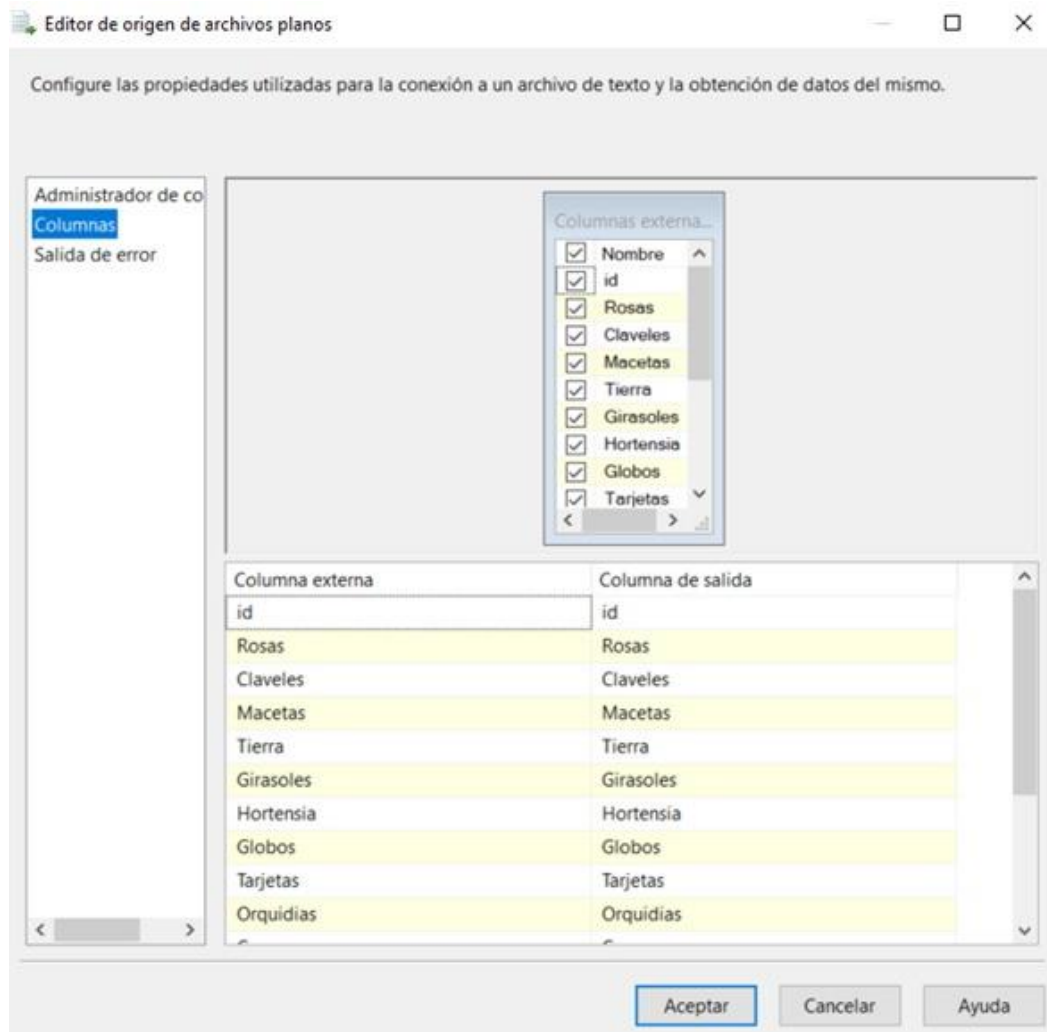
Ejercicio 1.

El Spa, "**Diego**", necesita segmentar sus clientes, para realizar una campaña de fidelización, y le pide a usted que efectué un análisis de sus tres sucursales, que defina cuantos grupos y que características tienen.

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.

Desarrollo del problema.



Consultas

```

1 SELECT COUNT(rosas) as 'rosas', COUNT(claveles) as 'claveles',COUNT(macetas)as 'macetas',COUNT(tierra)as 'tierra',
2 COUNT(Girasoles)as 'Girasoles',COUNT(Hortensia)as 'Hortensia',COUNT(Globos)as 'Globos',COUNT(Tarjetas)as 'Tarjetas',
3 COUNT(Orquidias)as 'Orquidias',COUNT(Carmes)as 'Carmes',COUNT(Lirios)as 'Lirios',COUNT(Aurora)as 'Aurora',
4 COUNT(Tulipanes)as 'Tulipanes',COUNT(Liston)as 'Liston'
5 from dbo.compra order by rosas;

```

The screenshot displays the SQL Server Enterprise Manager interface. On the left, the 'Object Explorer' shows the database structure for 'DESKTOP-FGEHH4G (SQL Server 15.0.2008.1)'. The 'dbo.compra' database is selected, showing tables, views, and other objects. On the right, the 'SQLQuery4.sql' window shows the execution results of the query. The results are presented in a grid format, with columns for 'departamento' and the count of each flower type. The query is ordered by the count of 'rosas'.

departamento	rosas
1 Santa Ana	176
2 San Miguel	157
3 San Salvador	612

departamento	Claveles
1 Santa Ana	246
2 San Miguel	137
3 San Salvador	350

departamento	macetas
1 Santa Ana	245
2 San Miguel	141
3 San Salvador	392

departamento	Tierra
1 Santa Ana	176
2 San Miguel	157
3 San Salvador	612

departamento	Girasoles
1 Santa Ana	246
2 San Miguel	137
3 San Salvador	350

departamento	Hortensia
1 Santa Ana	245
2 San Miguel	141
3 San Salvador	392

departamento	Globos
1 Santa Ana	176
2 San Miguel	157
3 San Salvador	612

departamento	Tarjetas
1 Santa Ana	246
2 San Miguel	137
3 San Salvador	350

departamento	Orquidias
1 Santa Ana	245
2 San Miguel	141
3 San Salvador	392

Query executed successfully.

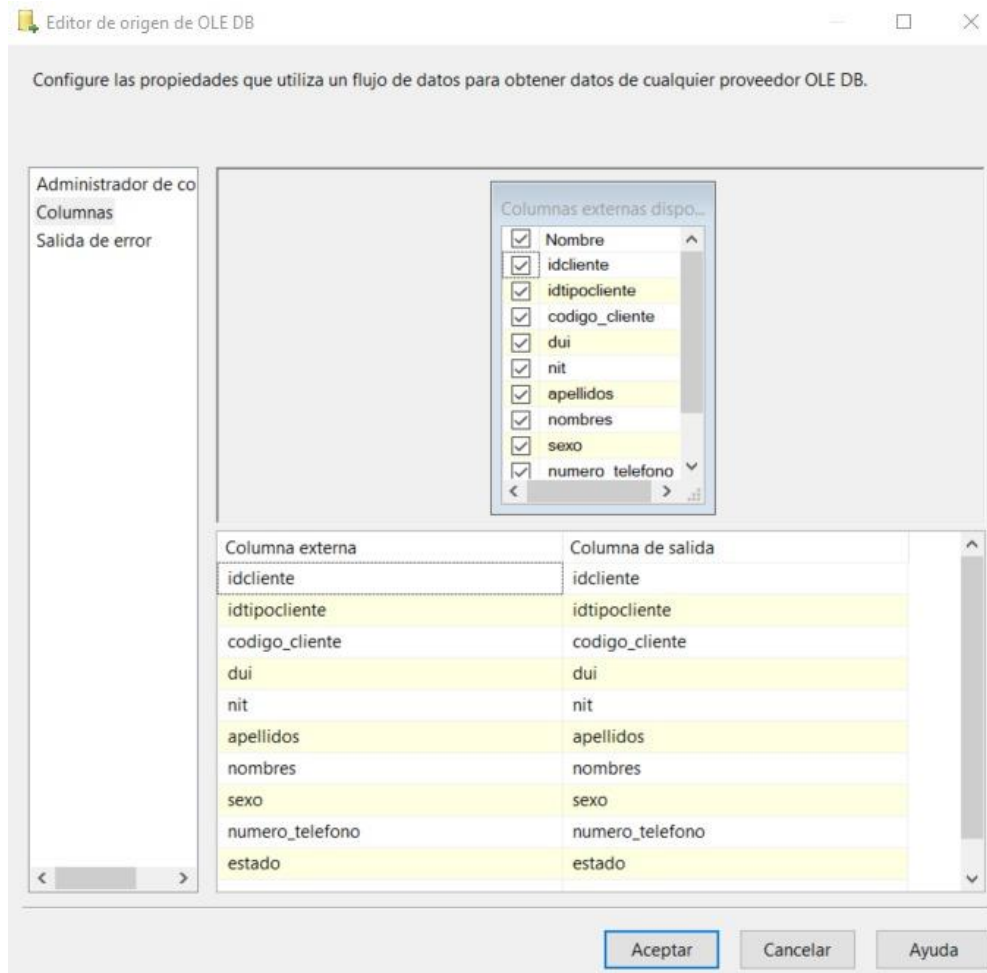
Conclusiones.

- Los productos más populares a nivel nacional son rosas, claveles, macetas, tierra, girasoles, hortensia, globos, tarjetas, orquídeas, carnes y las flores.
- El producto más popular de Santa Ana fue las tarjetas, de San Salvador, la Tierra y de San Miguel los globos.
- Consideramos que este problema se resolvió en un 100% de lo solicitado.

Ejercicio 3.

La telefonía “**FioDio**” solicita realizar un ETL que exporte una base de datos de Mysql y SQL Server, al final el destino serán dos archivos de Excel en donde en un archivo estarán los clientes preferenciales y ejecutivos y en el segundo los de gobierno y turista, adicional en los archivos de Excel se deberá crear un campo código de país, que se llenará como muestra e diagrama de la guía 1.

Desarrollo.



Configure las propiedades usadas por un flujo de datos para obtener datos de un proveedor de ADO.NET.

Administrador de co
Columnas
Salida de error

Especifique un administrador de conexiones de ADO.NET, un origen de datos o una vista del origen de datos, y seleccione el modo de acceso a los datos. Después de seleccionar el modo de acceso a los datos, seleccione entre las opciones adicionales de acceso a los datos que aparecerán.

Administrador de conexiones de ADO.NET:

127.0.0.1.no_auto_value_on_zero.root

Nueva...

Modo de acceso a datos:

Comando SQL

Texto de comando SQL:

select * from cliente

Generar consulta...

Examinar...

Vista previa...

Aceptar

Cancelar

Ayuda

Especifique las expresiones utilizadas para crear nuevos valores de columna e indique si los valores actualizan las columnas existentes o bien llenan columnas nuevas.

Variables y parámetros

Columnas

Funciones matemáticas

Funciones de cadena

Funciones de fecha y hora

Funciones NULL

Conversiones de tipo

Operadores

Descripción:

Nombre de columna ...	Columna derivada	Expresión	Tipo de datos	
codigo	<agregar como colu...	(DT_STR,15,1252)codigo_cliente	cadena [DT_STR]	1
dui1	<agregar como colu...	(DT_STR,9,1252)dui	cadena [DT_STR]	9
nit1	<agregar como colu...	(DT_STR,14,1252)nit	cadena [DT_STR]	1
apellido1	<agregar como colu...	(DT_STR,250,1252)apellidos	cadena [DT_STR]	2
nombre1	<agregar como colu...	(DT_STR,250,1252)nombres	cadena [DT_STR]	2
sexo1	<agregar como colu...	(DT_STR,1,1252)sexo	cadena [DT_STR]	1
numero	<agregar como colu...	(DT_STR,8,1252)numero_telefono	cadena [DT_STR]	8
estado1	<agregar como colu...	(DT_STR,1,1252)estado	cadena [DT_STR]	1

Configure las propiedades utilizadas para combinar varias entradas en una salida mediante la creación de asignaciones entre columnas.

Nombre de la columna de salida	Entrada de Unión de todo 1	Entrada de Unión de todo 2
idcliente	idcliente	idcliente
idtipocliente	idtipocliente	idtipocliente
codigo_cliente	codigo_cliente	codigo
dui	dui	dui1
nit	nit	nit1
apellidos	apellidos	apellido1
nombres	nombres	nombre1
sexo	sexo	sexo1
numero_telefono	numero_telefono	numero
estado	estado	estado1

Aceptar

Cancelar

Ayuda

Especifique las expresiones utilizadas para crear nuevos valores de columna e indique si los valores actualizan las columnas existentes o bien llenan columnas nuevas.

Variables y parámetros

Columnas

Funciones matemáticas

Funciones de cadena

Funciones de fecha y hora

Funciones NULL

Conversiones de tipo

Operadores

Descripción:

Nombre de columna ...	Columna derivada	Expresión	Tipo de datos	Le
codigopais	<agregar como colu...	SUBSTRING(codigo_cliente,1,2)	cadena Unicode [DT_...	2

< >

Configurar la salida de errores... Aceptar Cancelar Ayuda



Salida predeterminada de división condicional

Configure las propiedades que permiten la inserción de datos mediante un proveedor de Excel.

Administrador de co
Asignaciones
Salida de error

Especifique un administrador de conexiones, un origen de datos o una vista del origen de datos para el destino de Excel. A continuación, seleccione el modo utilizado para obtener acceso a los datos del destino. Tras seleccionar este modo, elija una de las opciones adicionales de acceso a datos que aparecen.

Administrador de conexiones de Excel:

Administrador de conexiones con Excel 6 ▼

Nueva...

Modo de acceso a datos:

Tabla o vista ▼

Nombre de la hoja de Excel:

Sheet1\$ ▼

Nueva...

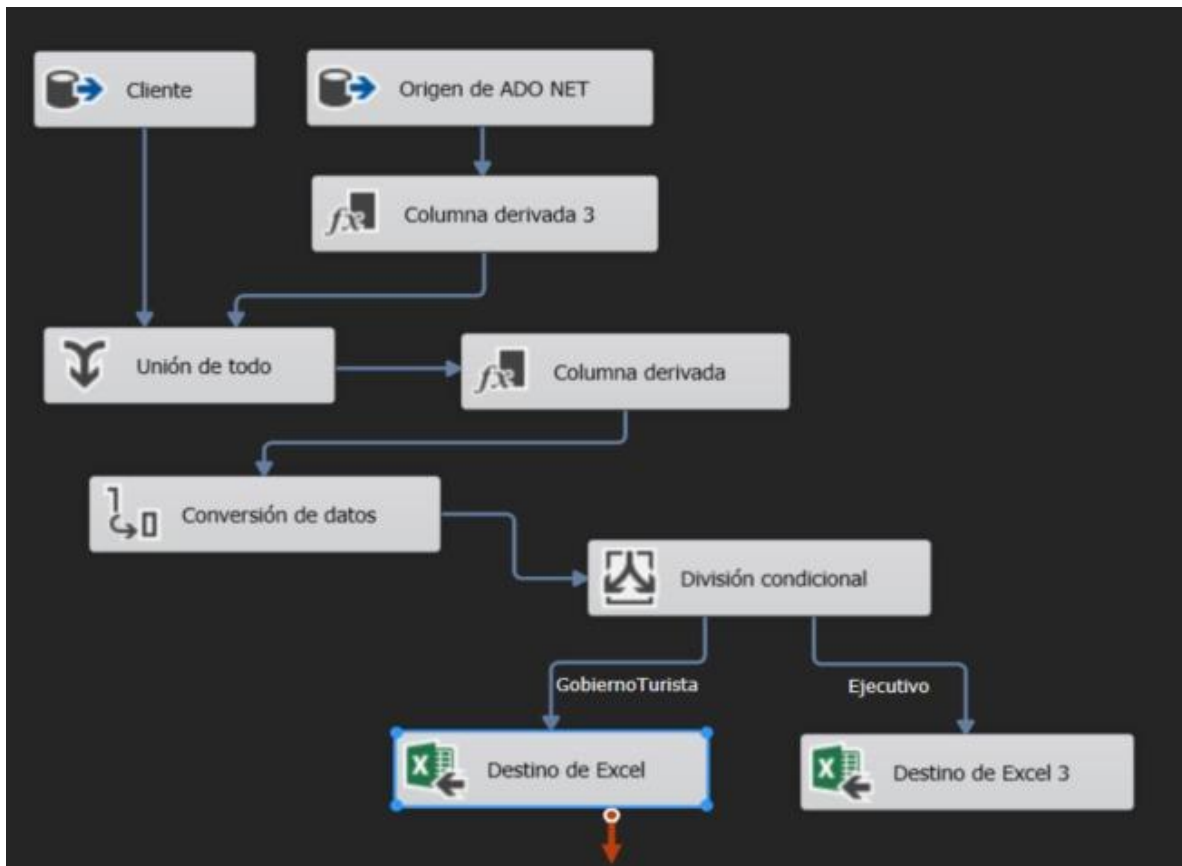
< >

Ver datos

Aceptar

Cancelar

Ayuda



Conclusiones. Se cumplió con lo solicitado con un 100%