

Universidad Don Bosco



Docente: Ing. Alexander Alberto Sigüenza

Materia: Datawarehouse y Minería de Datos

Presentado por:

Ruiz Saz, Carlos José RS181977

Rosales Mendoza, Patrick Ernesto RM181976

Porcentaje realizado: 100%

Ciclo: 02-2020

Soyapango, viernes 28 de agosto del 2020

Contenido

Capturas desafio I.....	3
Problema I	3
Problema II	5
Problema III	9

Capturas desafio I

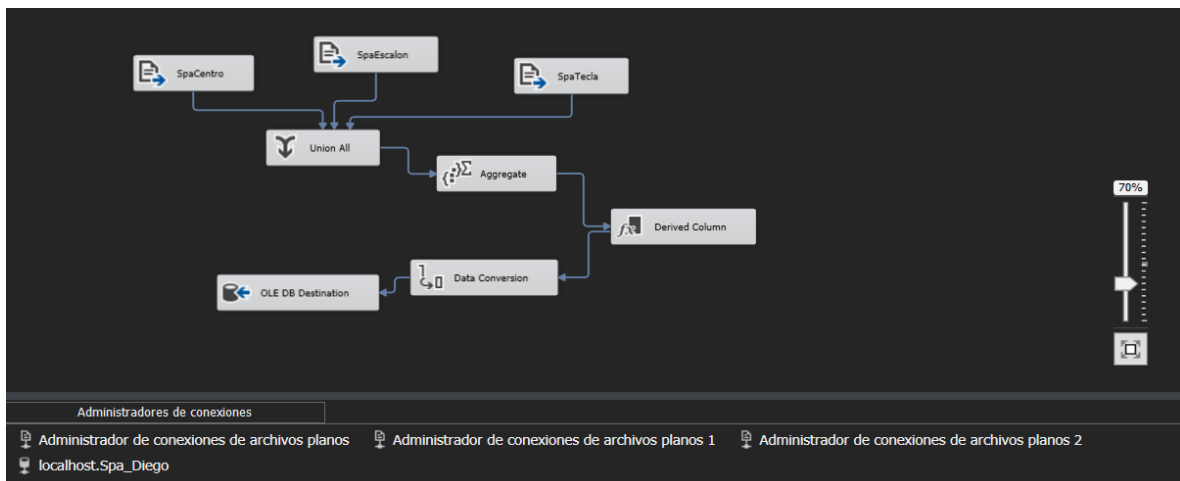
Problema I

Diagrama de la base de datos

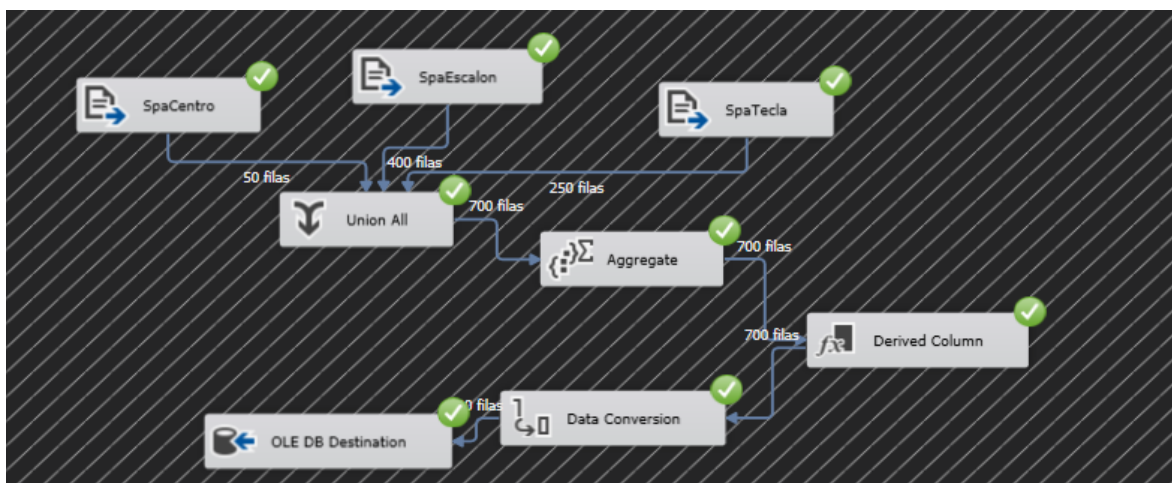
Clientes	
ID	
Nombre_completo	
Genero	
Ingresos	
Promedio_visitas	
edad	
sauna	
masaje	
Hidro	
Yoga	

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)

```
select top 100 edad, cast(avg(Promedio_visitas) as decimal (10,2)) as "promedio de visitas por edad" from Clientes
group by edad
order by "promedio de visitas por edad" desc
```

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:

	edad	promedio de visitas por e...
1	56	5.08
2	27	4.72
3	54	4.55
4	41	4.54
5	42	4.54
6	46	4.47
7	37	4.32
8	35	4.27
9	62	4.26
10	36	4.26
11	49	4.25
12	23	4.22
13	64	4.17
14	40	4.16
15	30	4.16
16	65	4.16
17	48	4.13
18	50	4.10
19	44	4.02
20	34	3.98
21	60	3.97
22	29	3.97
23	59	3.95
24	21	3.93
25	45	3.90
26	28	3.86
27	33	3.85
28	52	3.81

	edad	promedio de visitas por e...
29	61	3.80
30	38	3.79
31	31	3.77
32	26	3.74
33	22	3.72
34	24	3.72
35	58	3.68
36	32	3.61
37	53	3.59
38	57	3.52
39	43	3.51
40	25	3.38
41	47	3.38
42	55	3.20
43	51	3.03
44	63	3.01
45	39	2.73

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 género)

```
select top 100 genero, cast(avg(Promedio_visitas) as decimal (10,2)) as "promedio de visitas por genero" from Clientes  
group by genero  
order by "promedio de visitas por genero" desc
```

Al ejecutar esta consulta tenemos el siguiente resultado:

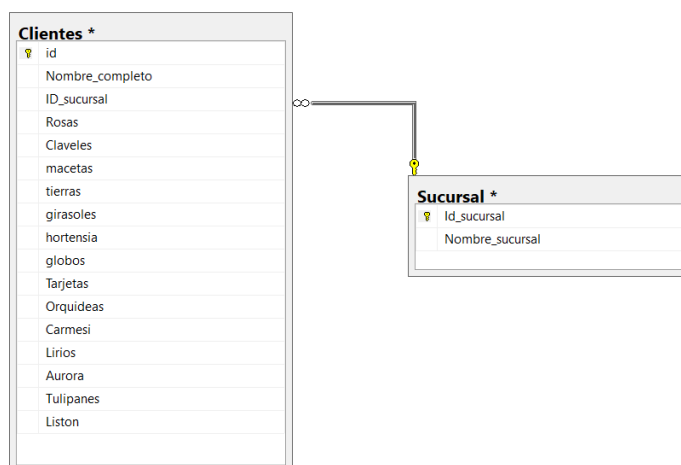
	genero	promedio de visitas por ge...
1	Mujer	3.93
2	Hombre	3.89

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

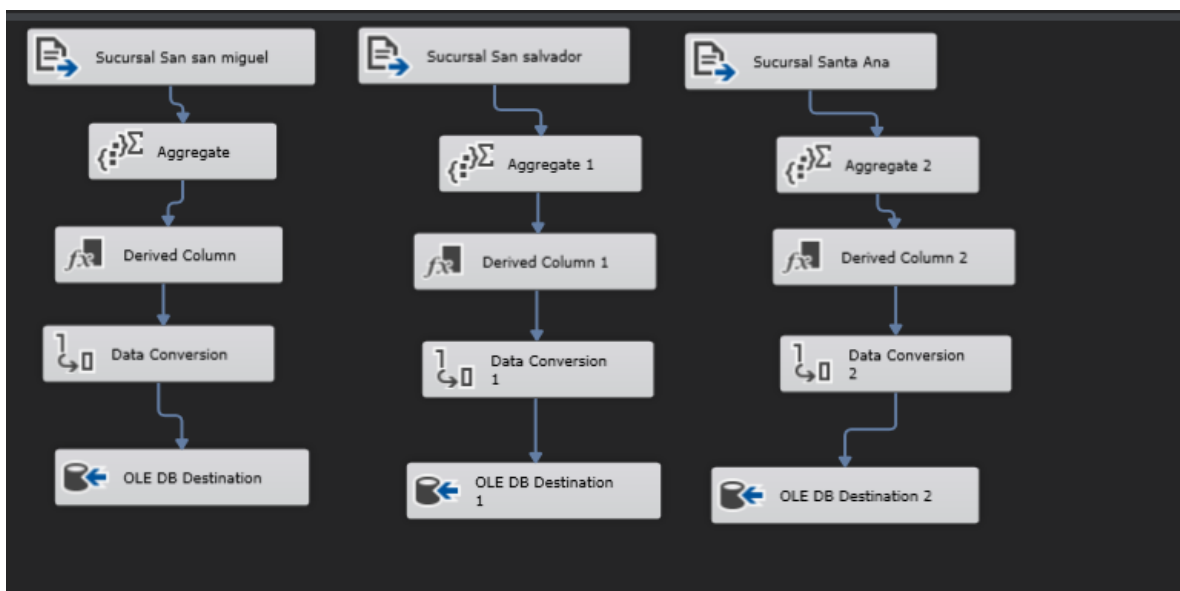
En este caso, se trabajan los archivos de texto plano por separado, ya que en tabla sucursales tienen los siguientes id:

1-San Miguel

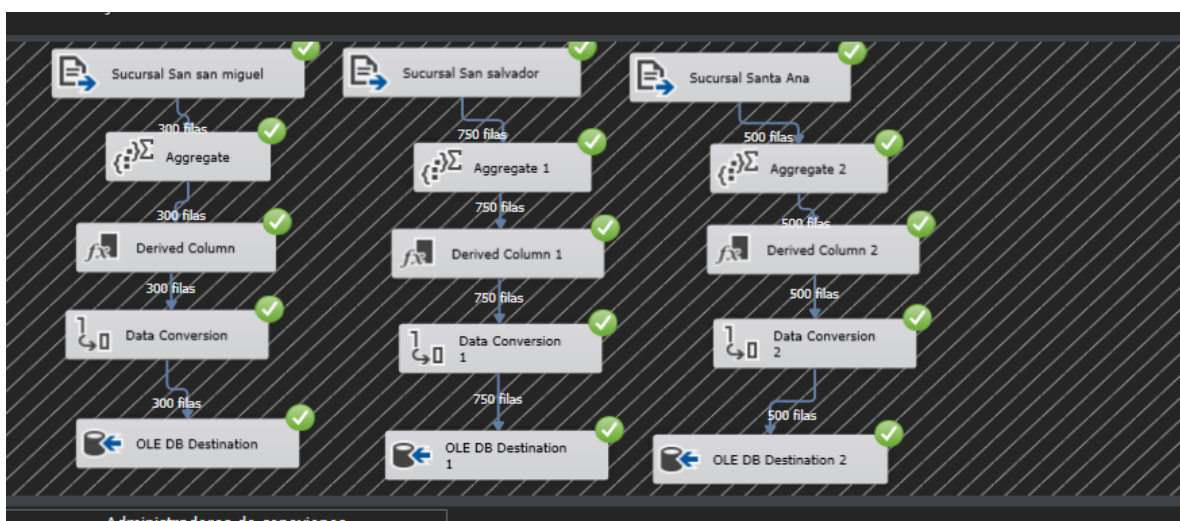
2-San Salvador

3-Santa Ana

Por lo que, en el derived column de san miguel se agrega un campo llamado id_sucursal=1, en san salvador se agrega id_sucursal=2 y en santa Ana se agrega id_sucursal=3, ya que estos datos se enviarán a la tabla clientes, entonces cada registro debe llevar el id de su sucursal, para establecer la relación entre tablas.



Ejecución del paquete



Estudio por departamento

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

```
SELECT top 100 s.Nombre_sucursal, sum(c.Aurora) as "# personas aurora", sum(c.Carmesi) as "# personas Carmesi", sum(c.Claveles) as "# personas claveles",  
sum(c.girasoles) as "# personas girasoles", sum(c.globos) as "# personas globos", sum(c.hortensia) as "# personas hortensia", sum(c.Lirios) as "# personas Lirios",  
sum(c.Liston) as "# personas liston", sum(c.macetas) as "# personas macetas", sum(c.Orquideas) as "# personas orquideas",  
sum(c.Rosas) as "# personas rosas", sum(c.tierras) as "# personas tierras", sum(c.Tulipanes) as "# personas tulipanes", sum(c.tarjetas) as "# personas tarjetas"  
FROM Clientes as c  
INNER JOIN Sucursal as s  
ON c.ID_sucursal = s.Id_sucursal  
group by s.Nombre_sucursal
```

Al ejecutarla da los siguientes resultados

Nombre_sucursal	# personas aurora	# personas Carmesi	# personas claveles	# personas girasoles	# personas globos	# personas hortensia	# personas Lirios	# personas liston
San miguel	160	158	137	150	151	157	160	149
San salvador	384	353	350	371	587	374	365	690
Santa Ana	260	236	246	266	154	243	270	136

# personas macetas	# personas orquideas	# personas rosas	# personas tierras	# personas tulipanes	# personas tarjetas
141	158	157	141	149	143
392	380	612	368	357	384
245	259	176	236	247	252

El top 6 de productos más vendidos por departamentos:

San Miguel

- Lirios
- Aurora
- Carmesí
- Orquídeas
- Hortensias
- Rosas

San salvador

- Listones
- Rosas
- Globos
- Macetas
- Aurora
- Tarjetas

Santa Ana

- Lirios
- Girasoles
- Aurora
- Orquídeas
- Tarjetas
- Tulipanes

Top 3 de las mejores combinaciones por departamento:

San Miguel

- Lirios y Aurora
- Lirios y Carmesí
- Aurora y Carmesí

San Salvador

- Listones y Rosas
- Listones y Globos
- Rosas y Globos

Santa Ana

- Lirios y Girasoles
- Lirios y Aurora
- Girasoles y Aurora

Estudio por país

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

```
SELECT top 100 sum(c.Aurora) as "# personas aurora",sum(c.Carmesi) as "# personas Carmesi",sum(c.Claveles) as "# personas claveles",  
sum(c.girasoles) as "# personas girasoles",sum(c.globos) as "# personas globos",sum(c.hortensia) as "# personas hortensia",sum(c.Lirios)as "# personas Lirios",  
sum(c.liston) as "# personas liston",sum(c.macetas) as "# personas macetas",sum(c.Orquideas) as "# personas orquideas",  
sum(c.Rosas) as "# personas rosas",sum(c.tierras) as "# personas tierras",sum(c.Tulipanes) as "# personas tulipanes",sum(c.tarjetas) as "# personas tarjetas"  
FROM Clientes as c
```

Y nos mostraría el siguiente resultado:

# personas aurora	# personas Carmesi	# personas claveles	# personas girasoles	# personas globos	# personas hortensia	# personas Lirios	# personas liston	# personas macetas
804	747	733	787	892	774	795	975	778

# personas orquideas	# personas rosas	# personas tierras	# personas tulipanes	# personas tarjetas
797	945	745	753	779

En este caso a nivel de país, los 6 productos que más se venden serían:

- Listones
- Rosas
- Globos
- Aurora
- Orquídea
- Girasoles

Las 3 mejores combinaciones a nivel de país serían:

- Listones y Rosas
- Listones y Globos
- Rosas y Globos

Problema III

Conexión a base de datos de costa rica

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:
localhost.db_roaming_costarica.carlos1 Nueva...

Modo de acceso a datos:
Comando SQL

Texto de comando SQL:

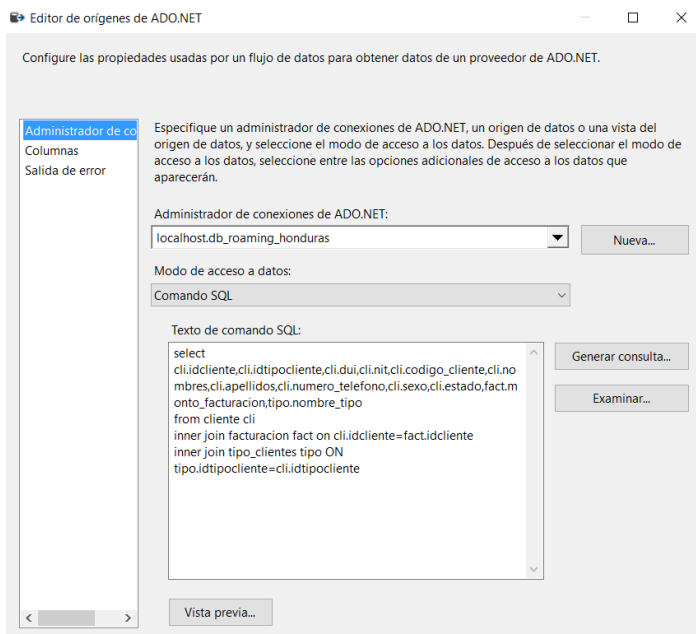
```
select  
cli.idcliente,cli.idtipopaciente,cli.dui,cli.nit,cli.codigo_cliente,cli.no  
mbres,cli.apellidos,cli.numero_telefono,cli.sexo,cli.estado,fact.m  
onto_facturacion,tipo.nombre_tipo  
from cliente cli  
inner join facturacion fact on cli.idcliente=fact.idcliente  
inner join tipo_clientes tipo ON  
tipo.idtipopaciente=cli.idtipopaciente
```

Generar consulta...
Examinar...

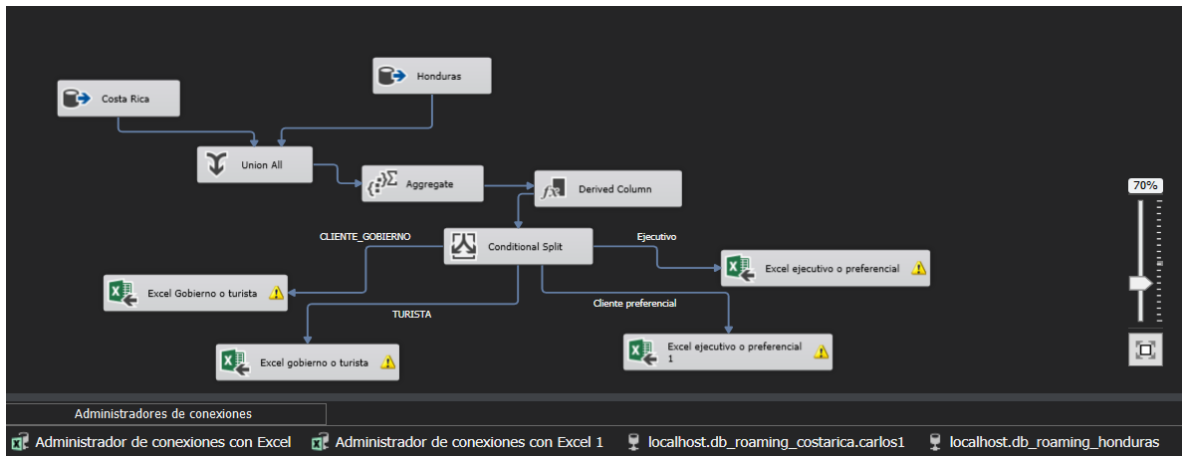
Vista previa...

Aceptar Cancelar Ayuda

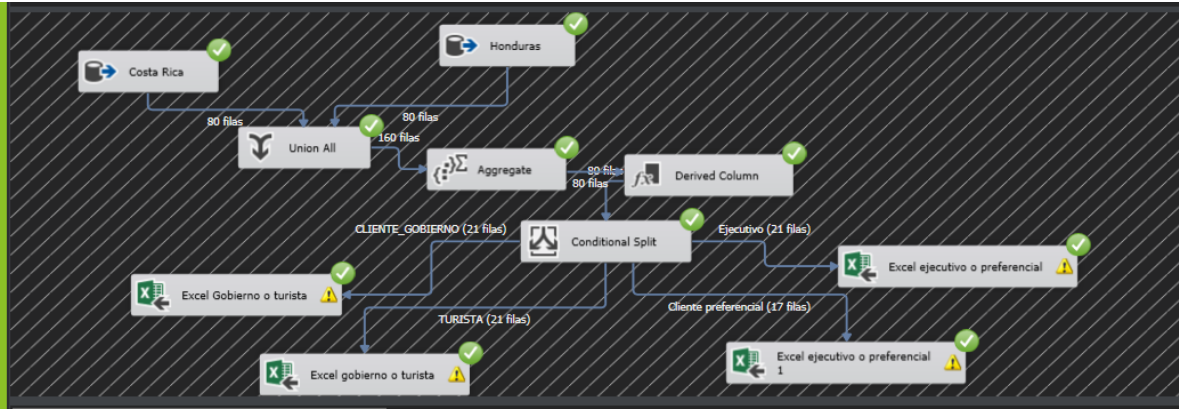
Conexión a base de datos de honduras



Flujo de datos



Prueba de ejecución



Capturas de los archivos Excel obtenidos

Excel clientes de gobierno y turista

	A	B	C	D	E	F	G	H	I	J
1	CODIGO CLIENTE	DUI	NIT	NOMBRE COMPLETO	SEXO	TELEFONO	ESTADO	MONTO FACTURACION	TIPO CLIENTE	Codigo pais
2	SV-SR47678	00542374-0	0210-150150	RICARDO ANTONIO, SILVA	MASCUINO	6293-0678	ACTIVO	61.20	CLIENTE GUBERNAMENTAL	SV
3	SV-ME47714	00449568-9	1009-100279	ERIC ALEXANDER, MEJIA	MASCUINO	7083-1594	ACTIVO	61.40	CLIENTE GUBERNAMENTAL	SV
4	SV-VS5897	00141561-4	0610-091159	SALVADOR, VASQUEZ M	MASCUINO	6333-5201	ACTIVO	334.40	CLIENTE GUBERNAMENTAL	SV
5	SV-CC47631	02216998-5	0112-260574	CARLOS ADALBERTO, C	MASCUINO	6541-4440	ACTIVO	85.60	CLIENTE GUBERNAMENTAL	SV
6	SV-GM47660	01129927-2	0805-120572	MANUEL ENRIQUE, GR	MASCUINO	6284-5726	INACTIVO	186.20	CLIENTE GUBERNAMENTAL	SV
7	SV-CM1664	01596559-2	0210-091163	MARTA ALICIA, CABRE	FEMENINO	6280-5523	ACTIVO	202.80	CLIENTE GUBERNAMENTAL	SV
8	SV-HM47684	04432219-3	0614-120391	MILTON SAMUEL, HER	MASCUINO	7581-2813	INACTIVO	192.40	CLIENTE GUBERNAMENTAL	SV
9	SV-LI47675	03454166-4	0821-280180	JORGE ALBERTO, LOPEZ	MASCUINO	7114-6362	ACTIVO	118.40	CLIENTE GUBERNAMENTAL	SV
10	SV-DM47697	00729179-1	0609-150457	MANUEL HERIBERTO, D	MASCUINO	7880-4336	INACTIVO	352.40	CLIENTE GUBERNAMENTAL	SV
11	SV-AE27200	02222044-8	0711-081066	ELMER ENRIQUE, AREV	MASCUINO	7853-3506	ACTIVO	360.40	CLIENTE GUBERNAMENTAL	SV
12	SV-AE47669	00453608-7	0811-301161	ESTANISLAO, ALVARAD	MASCUINO	7993-9305	ACTIVO	88.20	CLIENTE GUBERNAMENTAL	SV
13	SV-SH23204	03769657-3	0614-270787	HECTOR ALEJANDRO, SI	MASCUINO	7795-3576	ACTIVO	205.20	CLIENTE GUBERNAMENTAL	SV
14	SV-DY22084	01130847-9	1123-080476	YANIRA DEL CARMEN, I	FEMENINO	7594-0982	ACTIVO	90.60	CLIENTE GUBERNAMENTAL	SV
15	SV-ON16424	02932422-7	1307-251264	NATIVIDAD CRISTINO, C	FEMENINO	6206-1046	ACTIVO	121.80	CLIENTE GUBERNAMENTAL	SV
16	SV-NJ26082	00481311-2	0703-250365	JOSE ANTONIO, NOLAS	MASCUINO	7710-9443	ACTIVO	219.80	CLIENTE GUBERNAMENTAL	SV
17	SV-PW9858	00108338-6	0617-210164	WALTER MAURICIO, PA	MASCUINO	6227-2957	ACTIVO	164.60	CLIENTE GUBERNAMENTAL	SV
18	SV-LE47652	01799830-3	0614-210370	ERIC LOMBARDO, LEM	MASCUINO	6284-9154	ACTIVO	132.80	CLIENTE GUBERNAMENTAL	SV
19	SV-JI47659	00557486-8	0203-230383	JOSE RODOLFO, JIMENE	MASCUINO	6263-4944	ACTIVO	52.00	CLIENTE GUBERNAMENTAL	SV
20	SV-AR47674	02757200-2	0108-071281	REYNALDA EMILIA, AYA	FEMENINO	6529-1430	ACTIVO	339.60	CLIENTE GUBERNAMENTAL	SV
21	SV-AP47661	03407969-1	0609-061285	PEDRO ALFONSO, AMA	MASCUINO	7391-0697	ACTIVO	308.00	CLIENTE GUBERNAMENTAL	SV
22	SV-RA2158	00287394-6	0210-250167	ANA LIZ, RODRIGUEZ D	FEMENINO	6440-5701	INACTIVO	270.20	CLIENTE GUBERNAMENTAL	SV
23	SV-FF47704	01054713-5	0511-070282	FRANCISCO STEVE, FLO	MASCUINO	7921-1113	ACTIVO	70.80	CLIENTE TURISTA	SV
24	SV-LE47654	01945136-5	0614-181071	EVELYN ELIZABETH, LO	FEMENINO	7833-2389	ACTIVO	183.40	CLIENTE TURISTA	SV
25	SV-AL47672	00492838-9	0210-020382	LUIS ALONSO, ALVARE	MASCUINO	7663-1438	INACTIVO	301.80	CLIENTE TURISTA	SV
26	SV-EM47658	00374984-9	0619-030179	MARVIN VITELIO, ERAZ	MASCUINO	7786-5956	INACTIVO	93.40	CLIENTE TURISTA	SV
27	SV-MF47700	00889131-6	0614-091258	FRANCISCO REMBERTO	MASCUINO	6276-3473	INACTIVO	155.20	CLIENTE TURISTA	SV

Excel clientes ejecutivos y preferenciales

	A	B	C	D	E	F	G	H	I	J
1	CODIGO CLIENTE	DUI	NIT	NOMBRE COMPLETO	SEXO	TELEFONO	ESTADO	MONTO FACTURACION	TIPO CLIENTE	Codigo pais
2	SV-HE47712	01771937-7	0614-020980	ERICK STEFFAN , HERNANDEZ	MASCUINO	7695-5228	ACTIVO	364.60	CLIENTE EJECUTIVO	SV
3	SV-RJ47679	00088387-2	1416-271278	JUAN DAVID , REYES SALAS	MASCUINO	7071-0453	INACTIVO	292.40	CLIENTE EJECUTIVO	SV
4	SV-DF22100	02377745-6	0614-180648	FRANCISCO ENRIQUE , FLORES	MASCUINO	6523-6600	ACTIVO	210.40	CLIENTE EJECUTIVO	SV
5	SV-CH47710	01695478-8	1324-220868	HECTOR MANUEL , CHAN	MASCUINO	7778-1096	ACTIVO	82.60	CLIENTE EJECUTIVO	SV
6	SV-MS47711	04326581-9	1123-011290	SAMUEL GERARDO , MEJIA	MASCUINO	7869-2068	ACTIVO	219.40	CLIENTE EJECUTIVO	SV
7	SV-PG47677	00927645-4	0522-200472	GUILLERMO , POCASAN	MASCUINO	7789-4991	ACTIVO	272.80	CLIENTE EJECUTIVO	SV
8	SV-ME47627	02454429-6	0607-270265	EVELYN DE LOS ANGELES	FEMENINO	7070-4632	ACTIVO	395.80	CLIENTE EJECUTIVO	SV
9	SV-PR20278	00374800-4	0821-220252	RAUL ALBERTO , PINEDA	MASCUINO	6278-7729	ACTIVO	302.80	CLIENTE EJECUTIVO	SV
10	SV-BJ15868	01167192-9	1217-210658	JOSE ATILIO , BENITEZ P	MASCUINO	6207-4477	ACTIVO	308.60	CLIENTE EJECUTIVO	SV
11	SV-CB47683	01062229-4	1109-250583	BLANCA MARIBEL , CRUZ	FEMENINO	7886-6026	ACTIVO	335.40	CLIENTE EJECUTIVO	SV
12	SV-CR47598	02593311-3	0614-200854	ROSA MARGARITA , CORTES	FEMENINO	6243-7362	ACTIVO	363.40	CLIENTE EJECUTIVO	SV
13	SV-LC6418	01421695-6	0614-031062	CESAR MAURICIO , LOPEZ	MASCUINO	6305-5568	ACTIVO	86.60	CLIENTE EJECUTIVO	SV
14	SV-PJ19861	02726756-8	0210-170964	JOSE MAURICIO , PINEDA	MASCUINO	6232-1846	ACTIVO	369.80	CLIENTE EJECUTIVO	SV
15	SV-GM47595	01890496-7	0502-250251	MARIO , GUARDADO	MASCUINO	7928-9424	ACTIVO	318.80	CLIENTE EJECUTIVO	SV
16	SV-MJ25233	02435337-8	0614-161264	JORGE ALBERTO , MIRANDA	MASCUINO	6272-9067	INACTIVO	214.40	CLIENTE EJECUTIVO	SV
17	SV-NM9278	00090088-6	0614-271260	MANUEL VICENTE , NUÑEZ	MASCUINO	6131-7133	INACTIVO	35.20	CLIENTE EJECUTIVO	SV
18	SV-OL19455	02921225-0	0103-240966	LUIS WALTER , ORTIZ M	MASCUINO	6294-7980	ACTIVO	251.80	CLIENTE EJECUTIVO	SV
19	SV-GS47519	00829876-2	1010-030362	SARA LORENA , GUZMAN	FEMENINO	6284-6340	ACTIVO	384.20	CLIENTE EJECUTIVO	SV
20	SV-RM20831	00196268-7	0315-180165	MAURICIO ERNESTO , RAMOS	MASCUINO	6261-0010	ACTIVO	149.20	CLIENTE EJECUTIVO	SV
21	SV-RJ47676	00881659-4	0512-030470	JACOBO , RAMOS PEREZ	MASCUINO	7165-1054	INACTIVO	184.40	CLIENTE EJECUTIVO	SV
22	SV-GE25309	02575654-5	1407-201281	EDWIN JOVANY , GONZALEZ	MASCUINO	7410-0034	ACTIVO	163.80	CLIENTE EJECUTIVO	SV
23	SV-MM47693	00310459-4	0702-240253	MANUEL ANTONIO , MEJIA	MASCUINO	6235-7596	INACTIVO	189.80	CLIENTE PREFERENCIAL	SV
24	SV-LR47673	02054051-7	0614-190981	RICARDO ERNESTO , LAZAR	MASCUINO	7985-7656	INACTIVO	223.40	CLIENTE PREFERENCIAL	SV
25	SV-GC42266	01564617-0	0511-121181	CARLOS RICARDO , GHIFFI	FEMENINO	7247-1706	INACTIVO	214.20	CLIENTE PREFERENCIAL	SV
26	SV-RR5909	03365212-9	0610-150982	ROBERTO , RAMIREZ DE	MASCUINO	7536-9733	ACTIVO	143.60	CLIENTE PREFERENCIAL	SV
27	SV-QA47695	00619610-0	0614-290358	ANA DORA ALICIA , QUIROGA	FEMENINO	6271-6439	ACTIVO	234.20	CLIENTE PREFERENCIAL	SV