TALLER TRANSFORMACION DE DATOS CON POWER BI

JUAN SEBASTIAN MORENO RODRIGUEZ NAREN ALEJANDRO BLANCO PORRAS

FACULTAD DE INGENIERIA, UNIVERSIDAD DE CUNDINAMARCA 596408 - ARQUITECTURA DE DATOS

JORGE ARMANDO JURADO PERALTA

1 de oct. de 25

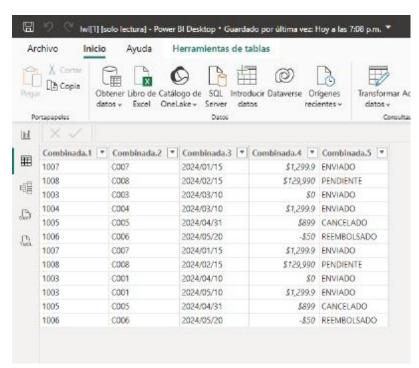
DESCRIPCION

En el presente documento se explicara que modificaciones se hicieron viendo la base de datos original, junto a la versión modificada explicando los cambios hechos

Base de datos

Ordenes de ventas

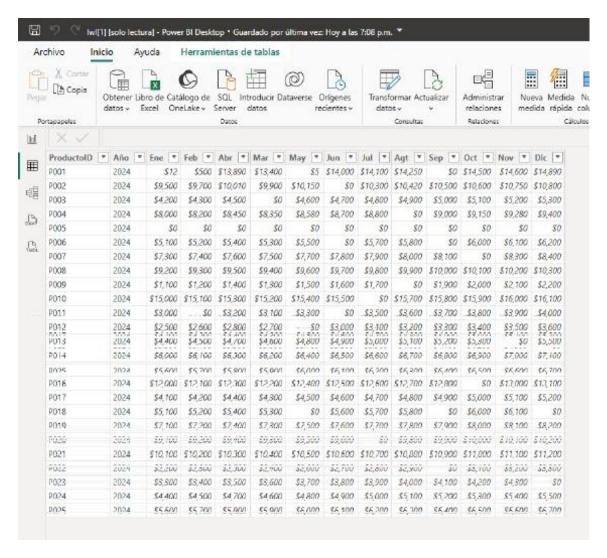




- Se separa todo por columnas
- Originalmente hay un conflicto por las comillas en ciertas entradas por lo que se ajusta para que se pueda leer el archivo correctamente
- El estado del envio tenia una diferencia entre masyusculas y minúsculas, se adecua todo a mayusuculas
- Para las entradas vacias y sin N/A, se establece que el valor predeterminado sea \$0
- La razón por la que las columnas aparecen con el nombre de Combinadas, es por las modificaciones realizadas para que el programa lea bien las entradas de informacion

Ventas mensuales

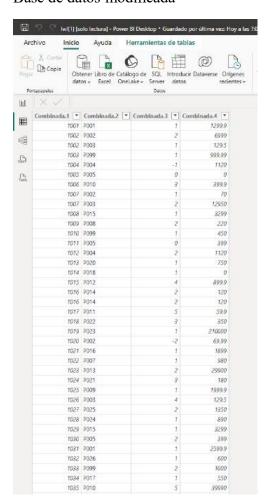
	A B C	D	Е	F	G
1	REPORTE DE VENTAS (BETA)				
2	Generado por: Área Comercial Formato	o no definitiv	/0		
3	ProductoID,Año,Ene,Feb,Mar,Abr,May,Ju			С	
4	P001,2024,\$12,500.00,13400,13890,5,140				0.14950
5	P002,2024,9500,9700,9900,10010,10150,,				,
6	P003,2024,4200,4300,NaN,4500,4600,470				
7	P004,24,8000,8200,8350,8450,8580,8700				
8	P005,2024,0,0,0,0,0,0,0,0,0,0,0,0				
9	P006,202X,5100,5200,5300,5400,5500,-,5	700,5800,,60	00,6100,6200		
10	P007,2024,7300,7400,7500,7600,7700,780	0,7900,8000	,8100,NaN,83	00,8400	
11	P008,2024,9200,9300,9400,9500,9600,970	0,9800,9900	,10000,10100,	10200,10300	
12	P009,202,1100,1200,1300,1400,1500,1600	,1700,,1900,	2000,2100,220	00	
13	P010,2024,15000,15100,15200,15300,1540	00,15500,N/A	A,15700,15800	,15900,16000	,16100
14	P011,2024,3000,0,3100,3200,3300,no,350	0,3600,3700	,3800,3900,40	00	
15	P012,2024,2500,2600,2700,2800,NaN,300	0,3100,3200	,3300,3400,35	00,3600	
16	P013,2024,4400,4500,4600,4700,4800,490	0,5000,5100	,5200,5300,,5	500	
17	P014,24,6000,6100,6200,6300,6400,6500,	6600,6700,68	800,6900,7000	,7100	
18	P015,2024,8700,8800,8900,9000,9100,920	0,9300,9400	,9500,9600,97	00,9800	
19	P016,2024,12000,12100,12200,12300,1240	00,12500,126	00,12700,128	00,,13000,131	00
20	P017,2024,4100,4200,4300,4400,4500,460	0,4700,4800	,4900,5000,51	00,5200	
21	P018,2024,5100,5200,5300,5400,-,5600,57	700,5800,-,60	000,6100,-		
22	P019,2024,7100,7200,7300,7400,7500,760	0,7700,7800	,7900,8000,81	00,8200	
23	P020,2024,9100,9200,9300,9400,9500,960	0,N/A,9800,9	9900,10000,10	100,10200	
24	P021,2024,10100,10200,10300,10400,10;5	00,10600,10	700,10800,109	000,11000,111	00,11200
25	P022,202,2200,2300,2400,2500,2600,2700	,2800,2900,,	3100,3200,33	00	
26	P023,2024,3300,3400,3500,3600,370;0,38	00,3900,4000	,4100,4200,4	300,NaN	
27	P024,2024,4400,4500,4600,4700,4800,490	0,5000,5100	,5200,5300,54	00,5500	
28	P025,2024,5;600,5700,5800,5900,6000,610	00,6;200,630	0,6400,6500,6	600,6700	
29					



- Se separa todo por columnas
- Se adecua las entradas de los meses para que se procesen como entradas de dinero, para que se incluya de manera automática el signo \$ y la coma numérica estandar

Detalle de ventas

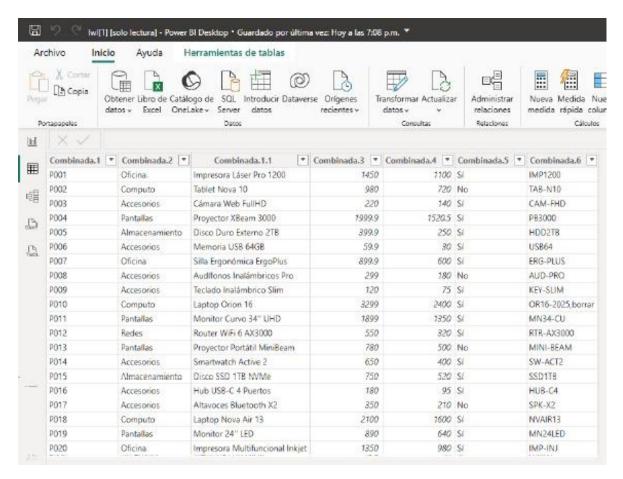
4	Α	В	С	D
1	OrderID,Prod	ductoID,Can	tidad,PrecioUr	nitario
2	1001,P001,1,	1299.90		
3	1002,P002,2,	"69,99"		
4	1002,P003,1,	129.5		
5	1003,P099,1,	999.99		
6	1004,P004,-1	,1120.00		
7	1005,P005,,			
8	1006,P010,3,	399.90		
9	1007,P002,1,	70.00		
10	1007,P003,2,	"129,50"		
11	1008,P015,1,	3299.0		
12	1009,P008,2,	220.0		
13	1010,P099,1,	450.0		
14	1011,P005,0,	399.0		
15	1012,P004,2,	1120,00		
16	1013,P020,1,	750.00		
17	1014,P018,1,			
18	1015,P012,4,	899.9		
19	1016,P014,2,	120.0		



- Se separa todo por columnas
- Originalmente hay un conflicto por las comillas en ciertas entradas por lo que se ajusta para que se pueda leer el archivo correctamente
- Para las entradas vacias y sin N/A, se establece que el valor predeterminado sea \$0
- La columna PrecioUnitario se adecua para ser numerico float para que incluya los valores decimales.
- La razón por la que las columnas aparecen con el nombre de Combinadas, es por las modificaciones realizadas para que el programa lea bien las entradas de informacion

Productos erp

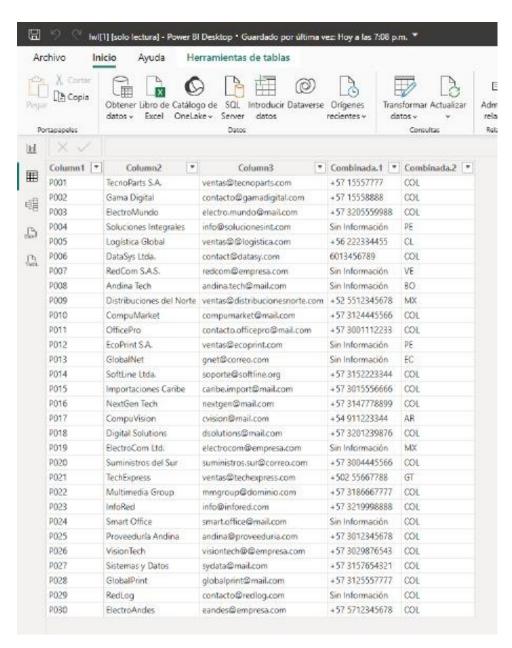
4	A B C D E F G
1	ProdcutID,Nombre,Categoria,PrecioLista,Costo,Activo,SKU,ColumnaIrrelevante
2	P006,Impresora Láser Pro 1200,Oficina,1450.0,1100.0,Sí,IMP1200,borrar
3	P007,Tablet Nova 10,Computo,980.0,720.0,NO,TAB-N10,borrar
4	P008,Cámara Web FullHD,Accesorios,220.0,140.0,True,CAM-FHD,borrar
5	P009,Proyector XBeam 3000,Pantallas,1999.9,1520.5,Si,PB3000,borrar
6	P010,Disco Duro Externo 2TB,Almacenamiento,399.9,250.0,sí,HDD2TB,borrar
7	P011,Memoria USB 64GB,Accesorios,59.9,30.0,Yes,USB64,borrar
8	P012,Silla Ergonómica ErgoPlus,Oficina,899.9,600.0,Sí,ERG-PLUS,borrar
9	P013,Audífonos Inalámbricos Pro,Accesorios,299.0,180.0,no,AUD-PRO,borrar
10	P014,Teclado Inalámbrico Slim,Accesorios,120.0,75.0,TRUE,KEY-SLIM,borrar
	P015,Laptop Orion 16"",Computo,3299.0,2400.0,si,OR16-2025,borrar
12	P016,Monitor Curvo 34" UHD,Pantallas,1899.0,1350.0,Sí,MN34-CU,borrar
	P017,Router WiFi 6 AX3000,Redes,550.0,320.0,sí,RTR-AX3000,borrar
	P018,Proyector Portátil MiniBeam,Pantallas,780.0,500.0,No,MINI-BEAM,borrar
	P019,Smartwatch Active 2,Accesorios,650.0,400.0,true,SW-ACT2,borrar
	P020,Disco SSD 1TB NVMe,Almacenamiento,750.0,520.0,SI,SSD1TB,borrar
	P021, Hub USB-C 4 Puertos, Accesorios, 180.0, 95.0, Sí, HUB-C4, borrar
	P022,Altavoces Bluetooth X2,Accesorios,350.0,210.0,no,SPK-X2,borrar
	P023,Laptop Nova Air 13"",Computo,2100.0,1600.0,TRUE,NVAIR13,borrar
	P024,Monitor 24" LED,Pantallas,890.0,640.0,Sí,MN24LED,borrar
21	P025,Impresora Multifuncional Inkjet,Oficina,1350.0,980.0,Si,IMP-INJ,borrar
22	
23	



- Se separa todo por columnas
- Se borro la ColumnaIrrelevante, se cambio el orden de las columnas, colocando primero que tipo de articulo se trata y luego la descripción, ejemplo Oficina – Impresora Laser Pro 1200 en lugar de ser Impresora Laser Pro 1200 – Oficina.
- Las columnas costo y PrecioLista se cambiaron a una entrada de tipo de Float para que procese los números decimales
- La razón por la que las columnas aparecen con el nombre de Combinadas, es por las modificaciones realizadas para que el programa lea bien las entradas de informacion

Contacto de proveedores

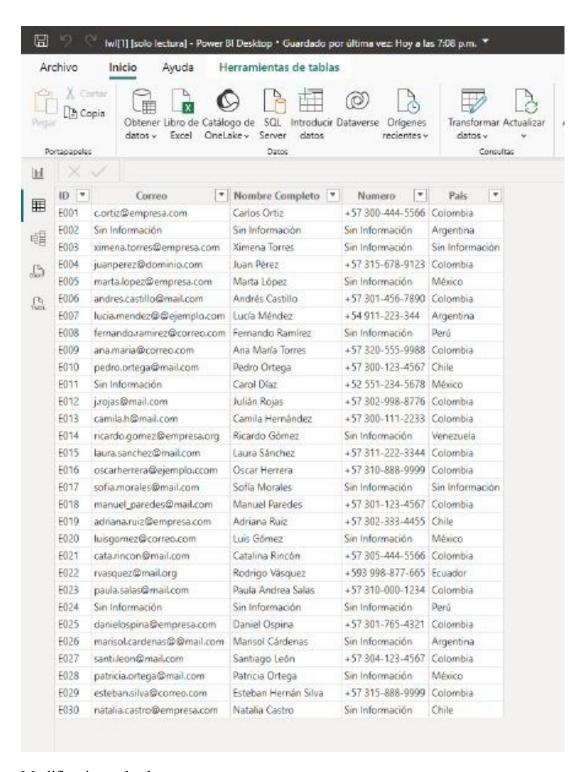
J							
4	Α	В	С	D	Е	F	G
1	Id,RazonSoci	al,Email,Tel	,Country				
2	P001,TecnoPa	arts S.A.,ven	itas@tecnopai	rts.com,+57-1	555-7777,CO)	
3	P002,Gama D	igital,conta	cto@gamadig	ital.co,(1)555	8888,Col		
4	P003,ElectroN	1undo,elect	ro.mundo@m	ail,com,320-5	555-9988,MX		
5	P004,Solucion	nes Integral	es,info@solud	cionesint.com	ı,,PE		
6	P005,Logístic	a Global,vei	ntas@@logisti	ica.com,+56 2	222334455,CL		
7	P006,DataSys	Ltda.,conta	ct@datasy.co	,601.345.6789	,co		
8	P007,RedCom	S.A.S.,redc	om@empresa	,,VE			
9	P008,Andina	Tech,andina	a.tech@mail,.	,BO			
10	P009,Distribu	ciones del l	Norte,ventas@	distribucion	esnorte.com,+	52 55123456	78,MX
11	P010,Compul	Market,com	pumarket@ma	ail.com,31244	145566,CO		
12	P011,OfficePr	o,contacto.	officepro@ma	il.,300.111.22	233,CL		
13	P012,EcoPrin	t S.A.,ventas	@ecoprint.co	"PE			
14	P013,GlobalN	let,gnet@co	rreo,,EC				
15	P014,SoftLine	Ltda.,sopor	rte@softline.o	rg,57 315-222	-3344,CO		
16	P015,Importa	ciones Cari	be,caribe.imp	ort@mail.cor	m,3015556666	,VE	
17	P016,NextGer	Tech,nextg	en@mail.co,3	14.777.8899,0	co		
18	P017,Compu\	/ision,cvisio	on@mailcom	1,+54 9112233	44,AR		
19	P018,Digital S	Solutions,ds	olutions@ma	il,co,3201239	876,CO		
20	P019,ElectroC	om Ltd.,ele	ctrocom@emp	oresa.com,,M	X		
21	P020,Suminis	tros del Sur	suministros.	sur@correo.,	3004445566,C	L	
22	P021,TechExp	ress,ventas	@techexpress	.co,+502 5566	67788,GT		
23	P022,Multime	edia Group,r	mmgroup@do	minio,org,318	8.666.7777,CO)	
24	P023,InfoRed	info@infore,	ed.com,57 321	-999-8888,PE			
25	P024,Smart O	ffice,smart.	office@mail.c	o,,CO			
26	P025,Proveed	uría Andina	andina@pro	veeduria.,301	.234.5678,BO		
27	P026,VisionTe	ech,visionte	ch@@empres	a.com,+57 30)29876543,CO		
28	P027,Sistema	s y Datos,sy	/data@mail.co	om,315-765-4	321,MX		
29	P028,GlobalP	rint,globalp	rint@mail.co,	312.555.7777	,CL		
30	P029,RedLog,	contacto@r	edlog,,CO				
31	P030,ElectroA	indes,eande	es@empresa.d	,57-1-234567	8,PE		
32							
	-						



- Se separa todo por columnas
- Originalmente en ciertas entradas los valores de Tel y Country se corrian una columna a la derecha, se hizo una modificación especial para corregir el problema
- Al usar diferentes formas de la nomenclatura, se uso una modificación especial para que el programa leyera bien los números de teléfono
- La razón por la que las columnas aparecen con el nombre de Combinadas, es por las modificaciones realizadas para que el programa lea bien las entradas de informacion

Contactos de empleados

4	A B C D E F G
1	Legajo,Nombre,e-mail,Celular,Pais
2	E001,Carlos Ortiz,c.ortiz@empresa.com,300.444.5566,Colombia
3	E002,na,n/a,,Argentina
4	E003,Ximena Torres,ximena.torres@empresa.com,,
5	E004, Juan Pérez, juanperez @dominio.com, 3156789123, Colombia
6	E005,Marta López,marta.lopez@empresa.,,México
7	E006,Andrés Castillo,andres.castillo@mail.com,301-456-7890,Colombia
8	E007,Lucía Méndez,lucia.mendez@@ejemplo.com,+54 911223344,Argentina
9	E008, Fernando Ramírez, fernando. ramirez@correo. com., Perú
	E009, Ana-María Torres, ana. maria@correo, co, 57 320-555-9988, Colombia
	E010, Pedro Ortega, pedro. ortega@mail.com, 3001234567, Chile
	E011, Carolina Díaz, ,+52 5512345678, México
	E012, Julián Rojas, j. rojas@mailcom, 3029988776, Colombia
	E013,Camila Hernández,camila.h@mail,co,300-111-2233,Colombia
	E014,Ricardo Gómez,ricardo.gomez@empresa.org,,Venezuela
	E015,Laura Sánchez,laura.sanchez@mail.com,311.222.3344,Colombia
	E016,Oscar Herrera,oscarherrera@ejemplo.c,3108889999,Colombia
	E017,Sofía Morales,sofia.morales@mail.com,,
	E018, Manuel Paredes, manuel_paredes@mail.com, +57 301-123-4567, Colombia
	E019,Adriana Ruiz,adriana.ruiz@empresa.com,302.333.4455,Chile
	E020,Luis Gómez,luisgomez@correo,,México
	E021,Catalina Rincón,cata.rincon@mail.com,3054445566,Colombia
	E022,Rodrigo Vásquez,rvasquez@mailorg,+593 998877665,Ecuador
	E023,Paula Andrea Salas,paula.salas@mail,310-000-1234,Colombia
	E024,na,n/a,,Perú
	E025, Daniel Ospina, danielospina @empresa.com, 301.765.4321, Colombia
	E026,Marisol Cárdenas,marisol.cardenas@@mail.com,,Argentina
	E027, Santiago León, santi.leon@mail.com, +57 3041234567, Colombia
	E028,Patricia Ortega,patricia.ortega@mail.co,,México
30	E029,Esteban-Hernán Silva,esteban.silva@correo.com,3158889999,Colombia
31	E030,Natalia Castro,natalia.castro@empresa.,,Chile
32	



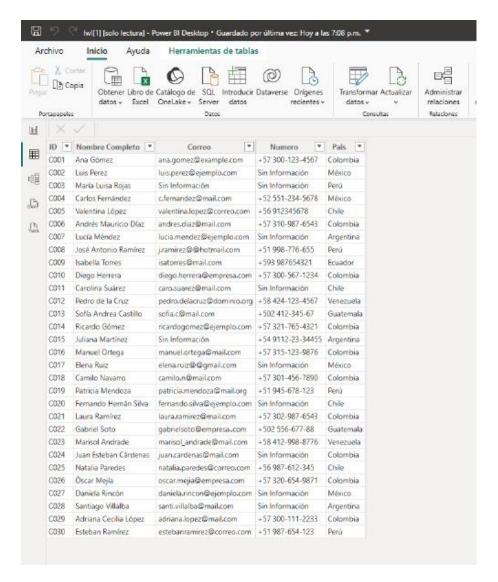
- Se separa todo por columnas
- Originalmente en ciertas entradas los valores de Tel y Country se corrian una columna a la derecha, se hizo una modificación especial para corregir el problema

- Al usar diferentes formas de la nomenclatura, se uso una modificación especial para que el programa leyera bien los números de teléfono
- Algunas entradas de Celular estaban vacias, se establecio que, ese caso, se introduzca "Sin Informacion" de manera automática
- Se cambiaron los nombres de las columnas para mayor comodidad.

Contacto clientes

4	A B C D	Е	F	G
1	ID,NombreCompleto,Correo,Telefono,Pais			
2	C001,Ana Gómez,ana.gomez@example.com,+57 30	01234567,Cold	ombia	
3	C002,Luis Perez,luis.perez@ejemplo,com,,México			
4	C003,María-Luisa Rojas,,—,Perú			
5	C004,Carlos Fernández,c.fernandez@mail.com,+52	5512345678,M	1éxico	
6	C005, Valentina López, valentina. lopez@correo, cl,+5	6 912345678,0	hile	
7	C006,Andrés-Mauricio Díaz,andres.diaz@mail.co,	57-310987654	3,Colombia	
8	C007,Lucía Méndez,lucia.mendez@ejemplo.com,,A	rgentina		
9	C008, José Antonio Ramírez, j. ramirez@@hotmail.co	m,+51 998776	655,Perú	
10	C009,Isabella Torres,isatorres@mail,com,+593 987	654321,Ecuad	or	
11	C010,Diego Herrera,diego.herrera@empresa.com,+	57 3005671234	,Colombia	
12	C011,Carolina Suárez,caro.suarez@mail.com,,Chil	e		
13	C012,Pedro de la Cruz,pedro.delacruz@dominio.or	g,+58 4241234	567,Venezuela	9
14	C013,Sofía-Andrea Castillo,sofia.c@mail.com,+502	41234567,Gua	atemala	
15	C014,Ricardo Gómez,ricardogomez@ejemplo.c,+57	3217654321,0	olombia	
16	C015, Juliana Martínez,,+54 91122334455, Argentina			
17	C016,Manuel Ortega,manuel.ortega@mail,+57-315	L239876,Colon	nbia	
18	C017,Elena Ruiz,elena.ruiz@@gmail.com,,México			
19	C018,Camilo Navarro,camilo.n@mail.com,+57 301	4567890,Colon	nbia	
20	C019,Patricia Mendoza,patricia.mendoza@mail.or	g,+51 94567812	23,Perú	
21	C020,Fernando-Hernán Silva,fernando.silva@ejem	plo,,Chile		
22	C021,Laura Ramírez,laura.ramirez@mail.com,+57	3029876543,Co	lombia	
23	C022,Gabriel Soto,gabrielsoto@empresacom,+50	2-55667788,Gu	atemala	
24	C023,Marisol Andrade,marisol_andrade@mail.com	n,+58 41299887	776,Venezuela	1
25	C024, Juan Esteban Cárdenas, juan. cardenas @mail	"Colombia		
26	C025,Natalia Paredes,natalia.paredes@correo,+56	987612345,Ch	ile	
27	C026,Óscar Mejía,oscar.mejia@empresa.com,+57	3206549871,Co	olombia	
28	C027,Daniela Rincón,daniela.rincon@ejemplo,.,M	éxico		
29	C028,Santiago Villalba,santi.villalba@mail.com,,A	rgentina		
30	C029,Adriana-Cecilia López,adriana.lopez@mail.c	o,+57 30011122	233,Colombia	
31	C030,Esteban Ramírez,estebanramirez@correo,+51	987654123,Pe	rú	

Base de datos modificada



Modificaciones hechas

- Se separa todo por columnas
- Originalmente en ciertas entradas los valores de Tel y Country se corrian una columna a la derecha, se hizo una modificación especial para corregir el problema
- Al usar diferentes formas de la nomenclatura, se uso una modificación especial para que el programa leyera bien los números de teléfono
- Algunas entradas de Celular estaban vacias, se establecio que, ese caso, se introduzca "Sin Informacion" de manera automática
- Se cambiaron los nombres de las columnas para mayor comodidad.