



Caso 137 Dos Niveles de Encabezados Columna

01 ETL y Modelo de Datos. PQ Limpieza de datos de 1º y 2º grado, técnicas de depuración. Anulación de la dinamización. Unpivot: Dos nivel de Encabezado de Columna

Jose Ignacio González Gómez

Departamento de Economía, Contabilidad y Finanzas - Universidad de La Laguna

www.jggomez.eu

R24/25

Ejercicio Basado: *EL ADN de Power Query – Miguel Caballero y Fabian Torres*

Contenido

Resumen y objetivos	1
1 Presentación.....	2
2 Se Pide.....	2
3 Técnicas de depuración, principales ajustes propuestos.....	3
3.1 Configuración de la consulta.....	3
3.2 Detalles del proceso de depuración.....	3

Resumen y objetivos

Power Query. Creando parámetros de conexión

Proceso Transformación: Aplicación de técnicas de depuración sobre filas y columnas basicas como formateo, encabezado promovidos, tipos cambiado, quitar filas en blanco, filtros (quitar totales, null..), quitar columnas vacías, rellenar, crear columna combinada, dividir columnas, etc.

Técnicas avanzadas: Transponer, anulamos dinamización de columna, etc..

1 Presentación

En este caso nos encontramos con una tabla con dos niveles de encabezado de columna, es decir, formato con dos niveles de encabezados en columnas, por tanto al tener otro nivel existirá una categoría que agrupa otros campos, en concreto Mes y Medidas (Ingreso, Presupuesto y Variación).

El segundo nivel de encabezado vemos que agrupa varias columnas, es decir, enero cubija las columnas *Ingresos, Presupuesto y Variación*.

El primer nivel de encabezado será siempre la etiqueta natural de la columna o nombre del campo, este caso únicamente agrupa una única columna, por ejemplo: *Ingresos*

	Enero			Febrero		
	Ingresos	Presupuesto	Variación	Ingresos	Presupuesto	Variación
Argentina	\$28.918	\$4.820	\$24.098	\$27.176	\$3.882	\$23.294
Brasil	\$33.260	\$5.543	\$27.717	\$27.422	\$3.917	\$23.505
Chile	\$29.707	\$4.951	\$24.755	\$27.838	\$3.977	\$23.861
Colombia	\$29.799	\$4.967	\$24.833	\$32.557	\$4.651	\$27.906
Ecuador	\$34.322	\$5.720	\$28.602	\$34.322	\$4.903	\$29.419
Paraguay	\$32.176	\$5.363	\$26.813	\$30.138	\$4.305	\$25.832
Perú	\$39.924	\$6.654	\$33.270	\$28.669	\$4.096	\$24.574
Uruguay	\$29.621	\$4.937	\$24.684	\$27.088	\$3.870	\$23.219
Venezuela	\$32.036	\$5.339	\$26.697	\$33.409	\$4.773	\$28.637
TOTAL	\$289.763	\$48.294	\$241.469	\$268.620	\$38.374	\$230.246

Ilustración 1

2 Se Pide

El formato de los datos está relacionado con **tabla (Unpivot) con dos nivel de encabezado de columna** y que el resto de datos no exigen una limpieza de datos de primer grado amplia, por tanto se deben hacer otros ajustes de segundo grado para lograr un resultado final como el adjunto.

	A	B	C	D
1	Pais	Mes	Medida	Valor
2	Argentina	Enero	Ingresos	28917,7235
3	Argentina	Enero	Presupuesto	4819,620583
4	Argentina	Enero	Variación	24098,10292
5	Argentina	Febrero	Ingresos	27176,2425
6	Argentina	Febrero	Presupuesto	3882,320357
7	Argentina	Febrero	Variación	23293,92214
8	Brasil	Enero	Ingresos	33260,2145
9	Brasil	Enero	Presupuesto	5543,369083

Ilustración 2

3 Técnicas de depuración, principales ajustes propuestos

3.1 Configuración de la consulta

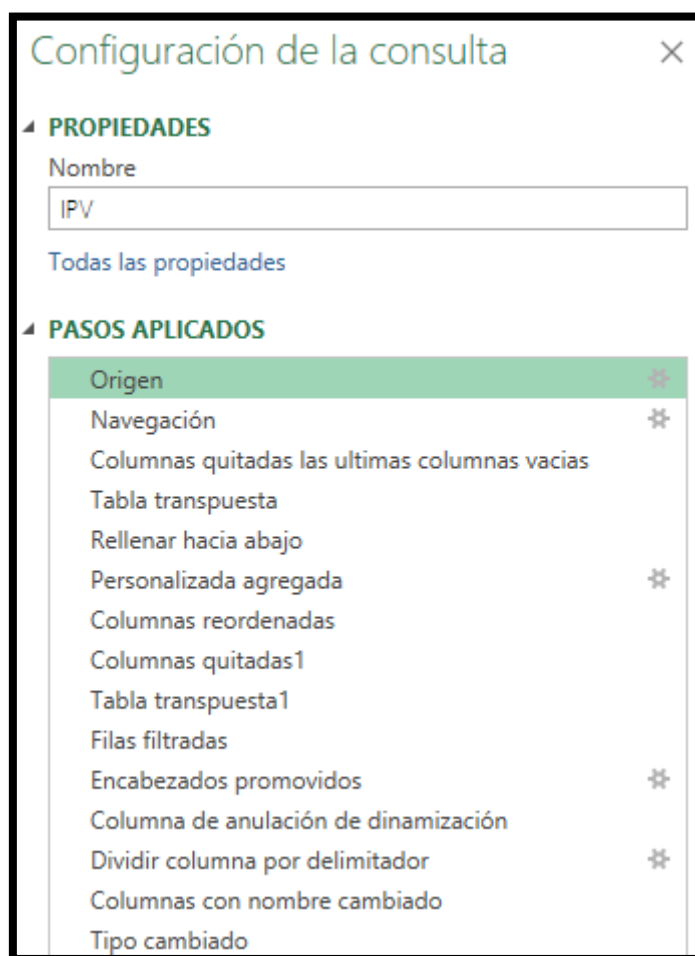


Ilustración 3

3.2 Detalles del proceso de depuración

En estos casos una de las propuestas es combinar las dos primeras filas de la tabla siguiendo los siguientes pasos.

- Accedemos a la opción transponer la tabla (ver Ilustración 4). En este momento podemos llevar cabo o como paso previo, el proceso de limpieza básica como:
 - Seleccionamos la primera columna procedemos a *Quitar filas en blanco*.
 - En la Columna 1 desplegamos los filtros y deshabilitamos TOTAL.
 - Buscamos las columnas en blanco y/o vacías y seleccionamos *Quitar Columnas*.
- Resultado dos columnas nuevas meses y variable y los meses rellenamos hacia abajo
- Creamos una columna combinada de mes-variable. Una vez realizado podemos quitar la Columna 1 y 2.

Table A: A pivot table with 'Column1' as the row field and 'Column2', 'Column3', 'Column4', 'Column5', 'Column6', and 'Column7' as column fields. The data is organized by month (Enero, Febrero) and variable (Ingresos, Presupuesto, Variación). A red box highlights the first three rows.

	Column1	Column2	Column3	Column4	Column5	Column6	Column7
1	null	Enero	Presupuesto	Argentina	Brasil	Chile	
2	null	Ingresos	Variación	28917,7235	33260,2145	29706,549	
3	Argentina	Ingresos	Presupuesto	4819,620583	5543,369083	4951,0915	
4	Brasil	Ingresos	Variación	24098,10292	27716,84542	24755,4575	
5	Chile	Ingresos	Presupuesto	27176,2425	27422,3415	27838,1715	
6	Colombia	Ingresos	Variación	32557,1	34221,885	4651,014286	
7	Ecuador	Ingresos	Presupuesto	28601,57083	34221,885	4903,126429	
8	Paraguay	Ingresos	Variación	26813,18125	30137,5395	4305,362786	
9	Perú	Ingresos	Presupuesto	33270,36	28669,3375	4095,619643	
10	Uruguay	Ingresos	Variación	24684,28042	27088,4145	3869,7735	
11	Venezuela	Ingresos	Presupuesto	26696,9225	33409,442	4772,777429	
12	TOTAL	Ingresos	Variación	241469,2854	268620,474	38374,35343	
13	null	Presupuesto	Variación	23293,92214	23504,86414	23861,28986	
14	null	Variación	Presupuesto	23293,92214	23504,86414	23861,28986	

Table B: A pivot table with 'Column1' as the row field and 'Column2', 'Column3', 'Column4', and 'Column5' as column fields. The data is organized by month (Enero, Febrero) and variable (Ingresos, Presupuesto, Variación). A red box highlights the first three rows, and blue arrows point to the 'Enero' and 'Febrero' rows.

	Column1	Column2	Column3	Column4	Column5
1	null	Argentina	Brasil	Chile	
2	Enero	Ingresos	28917,7235	33260,2145	29706,549
3	null	Presupuesto	4819,620583	5543,369083	4951,0915
4	null	Variación	24098,10292	24755,4575	24755,4575
5	Febrero	Ingresos	27176,2425	27422,3415	27838,1715
6	null	Presupuesto	3882,320357	3917,477357	3976,881643
7	null	Variación	23293,92214	23504,86414	23861,28986

Table C: A pivot table with 'Combinada' as the row field and 'Column1', 'Column2', 'Column3', and 'Column4' as column fields. The data is organized by month (Enero, Febrero) and variable (Ingresos, Presupuesto, Variación). A red box highlights the first three rows.

	Combinada	Column1	Column2	Column3	Column4
1	null	Argentina	Brasil		
2	Enero-Ingresos	Enero	Ingresos	28917,7235	33260,2145
3	Enero-Presupuesto	Enero	Presupuesto	4819,620583	5543,369083
4	Enero-Variación	Enero	Variación	24098,10292	27716,84542
5	Febrero-Ingresos	Febrero	Ingresos	27176,2425	27422,3415
6	Febrero-Presupuesto	Febrero	Presupuesto	3882,320357	3917,477357
7	Febrero-Variación	Febrero	Variación	23293,92214	23504,86414

Ilustración 4

- D. Volvemos a transponer la tabla (y filtramos quitando Total y vacíos), Ilustración 5
- E. Promovemos encabezados
- F. Seleccionamos las columnas con Mes-Variable y aplicamos anulación de dinamicación.
- G. Dividimos la columna por delimitar y terminamos de aplicar ajustes básicos como cambio de nombre a las columnas, etc.. ver Ilustración 5

ABC 123	Column1	ABC 123	Column2	ABC 123	Column3	ABC 123	Column4	ABC 123	Column5
1	D	null	Enero-Ingresos	Enero-Presupuesto	Enero-Variación	Febrero-Ingresos			
2	Argentina		28917,7235	4819,620583	24098,10292	27176,2425			
3	Brasil		33260,2145	5543,369083	27716,84542	27422,3415			
4	Chile		29706,549	4951,0915	24755,4575	27838,1715			
5	Colombia		29799,0775	4966,512917	24832,56458	32557,1			
6	Ecuador		34321,885	5720,314167	28601,57083	34321,885			
7	Paraguay		32175,8175	5362,63625	26813,18125	30137,5395			
8	Perú		39924,432	6654,072	33270,36	28669,3375			
9	Uruguay		29621,1365	4936,856083	24684,28042	27088,4145			
10	Venezuela		32036,307	5339,3845	26696,9225	33409,442			
11	TOTAL		289763,1425	48293,85708	241469,2854	268620,474			
12	null		null	null	null	null			
13	null		null	null	null	null			

ABC 123	Column1	ABC 123	Enero-Ingresos	ABC 123	Enero-Presupuesto	ABC 123	Enero-Variación	ABC 123	Febrero-Ingresos
1	Argentina		28917,7235		4819,620583		24098,10292		27176,2425
2	Brasil		33260,2145		5543,369083		27716,84542		27422,3415
3	Chile		29706,549		4951,0915		24755,4575		27838,1715
4	Colombia		29799,0775		4966,512917		24832,56458		32557,1

ABC 123	Column1	A ^B _C	Atributo	ABC 123	Valor
1	Argentina		Enero-Ingresos		28917,7235
2	Argentina		Enero-Presupuesto		4819,620583
3	Argentina		Enero-Variación		24098,10292
4	Argentina		Febrero-Ingresos		27176,2425

ABC 123	Column1	A ^B _C	Atributo.1	A ^B _C	Atributo.2	ABC 123	Valor
1	Argentina		Enero		Ingresos		28917,7235
2	Argentina		Enero		Presupuesto		4819,620583
3	Argentina		Enero		Variación		24098,10292
4	Argentina		Febrero		Ingresos		27176,2425
5	Argentina		Febrero		Presupuesto		3882,320357

Ilustración 5

De esta forma el resultado final sería similar al siguiente:

ABC 123	A ^B _C	Pais	A ^B _C	Mes	A ^B _C	Medida	1.2	Valor
1		Argentina		Enero		Ingresos		28917,7235
2		Argentina		Enero		Presupuesto		4819,620583
3		Argentina		Enero		Variación		24098,10292
4		Argentina		Febrero		Ingresos		27176,2425
5		Argentina		Febrero		Presupuesto		3882,320357

Ilustración 6