



Caso 132 Limpieza de Datos

01 ETL y Modelo de Datos. Limpieza de Datos, dividir columnas, 1º fila como encabezado, eliminar filtrar, cambiar tipos. Menú vistas. Quitar espacios en blanco, formato y recortar texto

Jose Ignacio González Gómez

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

www.jggomez.eu

R24/25

Ejercicio Basado: **DATDATA** [/ Limpieza de datos con Power Query \(youtube.com\)](#)

Archivos fuentes: 132 Limpieza de Datos

Resumen: Contiene opciones de uso más frecuente en el Proceso ETL con Power Query.

Técnicas y palabras clave: *Dividir columnas, 1º fila como encabezado, eliminar filtrar, cambiar tipos. Menú vistas (mostrar espacios en blancos y otras propiedades, calidad de columnas, distribución de columnas y perfil de columnas). Quitar espacios en blanco, formato y recortar texto*

Contenido

Presentación del caso	2
Se pide (preguntas base del modelo)	2
Crear conexión y a la fuente de datos y llevar el proceso de depuración	2
Realizar los siguientes analisis básicos fundamentados en Tablas y Graficos Dinámicos ¡Error! Marcador no definido.	
Pasos sugeridos en el proceso ETL	3
Proceso ETL con Power Query Modelado de datos con Power Pivot	3
<i>Ajustes propuestos básicos propuestos</i>	3
<i>Quitar espacios en blanco: Reemplazar valores y Formato Recortar texto</i>	4
<i>Menú vista: Calidad de las columnas, distribución de las columnas y perfil de las columnas</i>	4
Referencias	5

Presentación del caso

Disponemos de un fichero con los datos de nuestro antiguo programa de gestión que presenta un formato poco convencional tal y como se muestra en la siguiente ilustración donde en una sola columna tenemos todos los datos separados por una barra

	A	B	C	D
1	Fecha	Categoría	Concepto	Monto
2	-----	-----	-----	-----
3	01/01/2023	Ingresos	Ventas Productos	729,41
4	02/01/2023	Ingresos	Servicios	628,91
5	03/01/2023	Salidas	Caja Chica	-219,75
6	04/01/2023	Salidas	Papelería	-305,23
7	05/01/2023	Salidas	Gastos Menores	-152,86
8	06/01/2023	Ingresos	Otros	482,10
9	07/01/2023	Ingresos	Ventas Productos	893,64
10	08/01/2023	Salidas	Caja Chica	-284,97
11	09/01/2023	Salidas	Papelería	-380,05
12	10/01/2023	Salidas	Gastos Menores	-115,80

Ilustración 1

Se pide (preguntas base del modelo)

Crear conexión y a la fuente de datos y llevar el proceso de depuración

Se solicita llevar a cabo el proceso ETL (Conexión-Transformación y Carga), con un resultado similar al siguiente Ilustración 2

	A	B	C	D
1	Fecha	Categoría	Concepto	Monto
2	01/01/2023	Ingresos	Ventas Productos	729,41
3	02/01/2023	Ingresos	Servicios	628,91
4	03/01/2023	Salidas	Caja Chica	-219,75
5	04/01/2023	Salidas	Papelería	-305,23
6	05/01/2023	Salidas	Gastos Menores	-152,86
7	06/01/2023	Ingresos	Otros	482,1
8	07/01/2023	Ingresos	Ventas Productos	893,64
9	08/01/2023	Salidas	Caja Chica	-284,97
10	09/01/2023	Salidas	Papelería	-380,05

Ilustración 2

Realizar los siguientes analisis básicos fundamentados en Tablas y Graficos Dinámicos

Ademas se propone que el alumno realice diferentes Tablas y Gráficos Dinámicos con el fin de familiarizarse con estas visualizaciones. A modo orientativo se proponen las siguientes:

Concepto	Imprte	%
Ingresos	206.786,11	100,00%
Otros	80.155,42	38,76%
Servicios	55.702,70	26,94%
Ventas Productos	70.927,99	34,30%
Salidas	-111.591,36	100,00%
Caja Chica	-37.654,62	33,74%
Gastos Menores	-19.945,93	17,87%
Papelería	-53.990,81	48,38%
Total general	95.194,75	

Ilustración 3

Fecha
Todos los períodos
MESES
2025
JUN JUL AGO SEP OCT NOV DIC

Pasos sugeridos en el proceso ETL

- Creamos nuestra hoja de solución al caso planteado

Proceso ETL con Power Query Modelado de datos con Power Pivot

Pasos propuestos correspondiente al proceso de Transformación. Este proceso los vamos a completar para la conexión y realizaremos los siguientes ajustes:

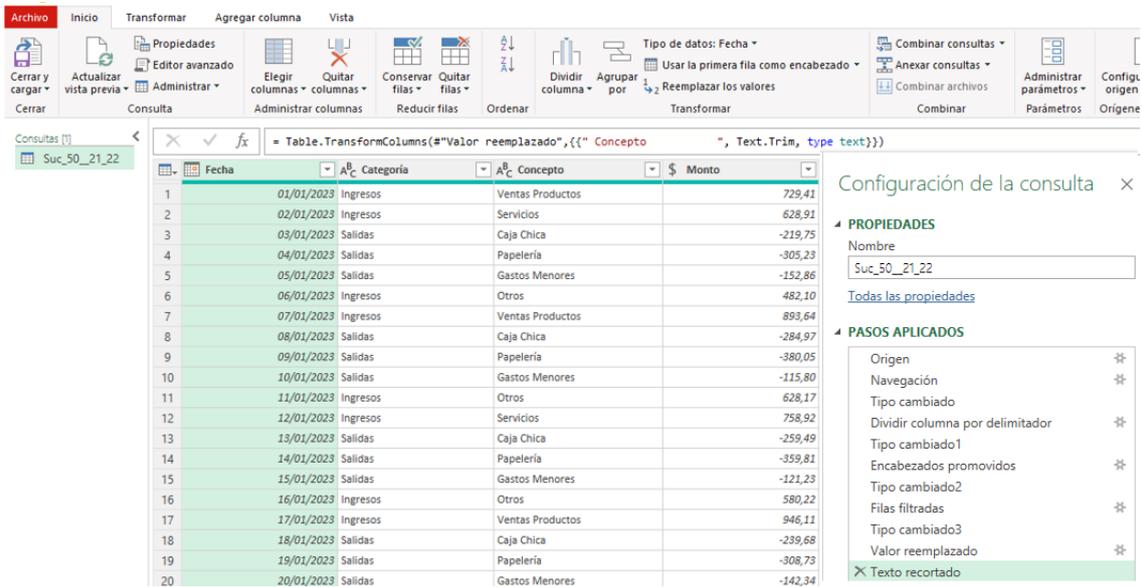


Ilustración 4

Ajustes propuestos básicos propuestos



Ilustración 5

Quitar espacios en blanco: Reemplazar valores y Formato Recortar texto

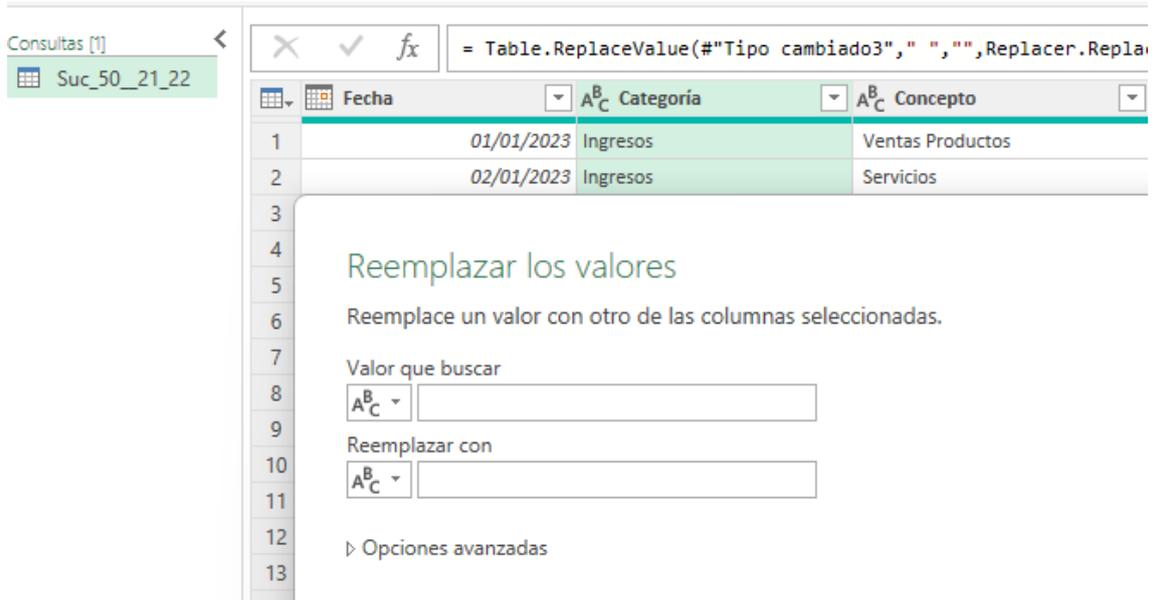


Ilustración 6

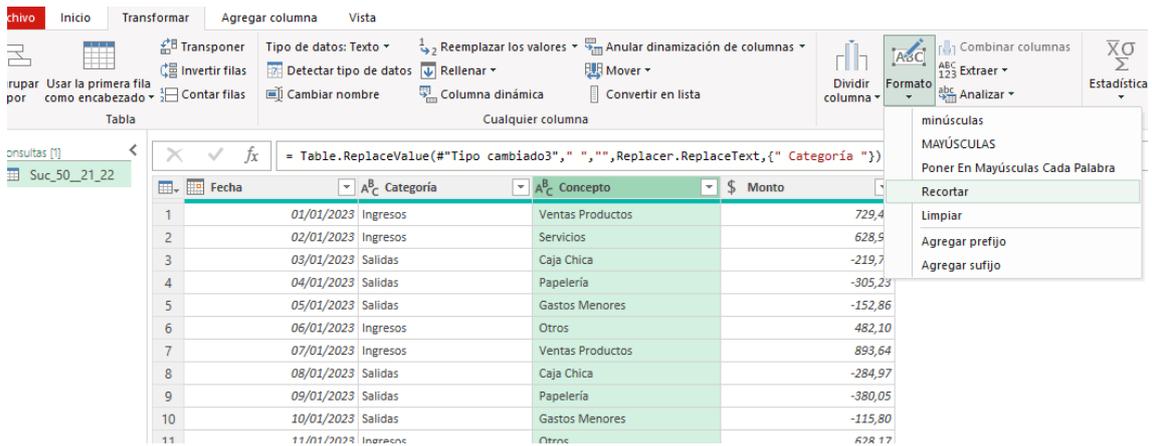


Ilustración 7

Menú vista: Calidad de las columnas, distribución de las columnas y perfil de las columnas

Estadísticas de columna

Recuento	730
Error	0
Vacío	0
Distintos	6
Únicos	0
Cadena vacía	0
Mín.	Caja Chi...
Máx.	Ventas P...

Distribución de valores

Caja Chica	729,41
Papelería	628,91
Gastos Menores	-219,75
Otros	-305,23
Servicios	-152,86
Ventas Productos	482,10

Ilustración 8

Referencias

[Limpieza de datos con Power Query \(youtube.com\)](https://www.youtube.com/watch?v=...)