



Caso 111b ETL Integración carpeta II

01 ETL y Modelo de Datos. Integración de todos los ficheros actuales y futuros contenidos en una carpeta. Integración en una sola consulta de la información de los ficheros contenidos en una misma carpeta

Jose Ignacio González Gómez

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

www.jggomez.eu

V.2.3

Ejercicio Basado: [Importar y Combinar todos los Archivos de una Carpeta - Power BI para exceleros](#) Autor: Sergio Alejandro Campos en [Excel Blog - Sergio Alejandro Campos - EXCELEINFO](#)

Archivos fuentes: Datos

Resumen: Integración de todos los ficheros actuales y futuros contenidos en una carpeta. Integración en una sola consulta de la información de los ficheros contenidos en una misma carpeta

Técnicas y palabras clave: Crear conexiones a una carpeta, cambiar tipo a columna fecha de número a fechas, Recortar espacios en blanco

Enunciado

En este caso vamos a crear las bases de un cuadro de mando con archivos contenidos en una carpeta a la cual se añadirán posteriormente nuevos ficheros, todos ellos con la misma estructura, de tal forma que al actualizar el cuadro de mando la información contenida en estos últimos incorporados se añadirán automáticamente a la tabla, grafico o visualización dependiente de esta fuente.

En este caso disponemos de un archivo por cada trimestre, 1º y 2º en la carpeta trimestre y nos faltan aun por llegar los del 3º y 4º trimestre, por tanto vamos a trabajar con los disponibles en la carpeta trimestres que son los dos del 1º y 2º trimestre.

Trimestre

01Trimestre

02Trimestre

La información está relacionada con una empresa de renta car de Canarias y la aplicación de gestión nos ofrece la información tabulada con la siguiente estructura en formato tabular CSV donde destaca lo siguiente:

La información está relacionada con una empresa de renta car canaria y la aplicación de gestión nos ofrece la información tabulada con la siguiente estructura en formato CSV donde destaca lo siguiente:

- Producto hace referencia a la categoría a la que pertenece cada vehículo.
- Cantidad se refiere al número de días de alquiler para cada categoría gama.

	A	B	C	D	E
1	ID Producto	Producto	Mes	NMes	Cantidad
2	4525	Grupo A (Base)	Enero	1	92
3	6758	Grupo B (Medio)	Enero	1	92
4	8888	Grupo C (SUV)	Enero	1	111
5	9769	Grupo D (Lujo)	Enero	1	70
6	4525	Grupo A (Base)	Febrero	2	71
7	6758	Grupo B (Medio)	Febrero	2	34
8	8888	Grupo C (SUV)	Febrero	2	84
9	9769	Grupo D (Lujo)	Febrero	2	101
10	4525	Grupo A (Base)	Marzo	3	100
11	6758	Grupo B (Medio)	Marzo	3	55
12	8888	Grupo C (SUV)	Marzo	3	46
13	9769	Grupo D (Lujo)	Marzo	3	36

Recordar que en este tipo de caso de integración de datos tiene como condición que todos los ficheros deben contener la misma estructura, número de columnas, tipo, nombre y orden para poder consolidar la información.

Se pide (preguntas base del modelo)

Creacion del libro trabajo

Debemos trabajar con un archivo independiente del fichero de datos por tanto el primer paso es crearlo y que va a contener toda la analítica de datos desarrollada.

Sobre el analisis previo exploratorio

En el proceso previo de analisis, como es el analisis exploratorio debe tenerse en consideración los datos disponibles y su actualización asi como el proceso de conexión más adecuado.

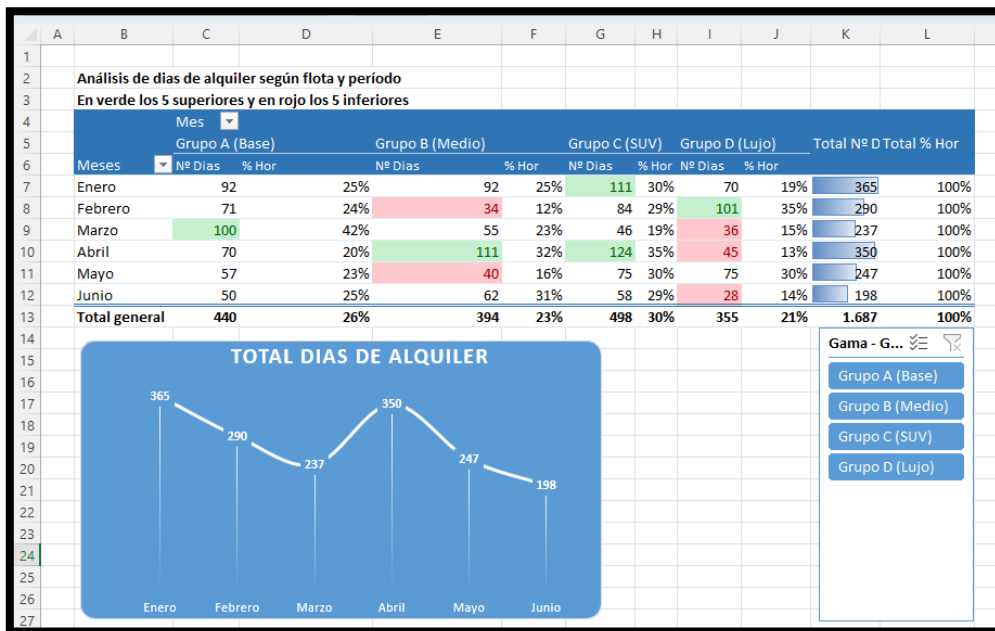
Sobre el proceso ETL

Atendiendo a las características del caso la conexión a las fuentes de información será a través de la carpeta "trimestre" que contiene los ficheros que serán actualizando.

Del analisis exploratorio previo es necesario eliminar (recortar) los espacios en blanco en especial de las columnas tipo texto y mes asi como garantizar la coherencia del tipo de campo con el contenido de la columna.

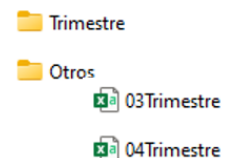
Cuadros de mando. Práctica con Tablas y Gráficos Dinámicos

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



Incorporar los dos trimestres que faltan

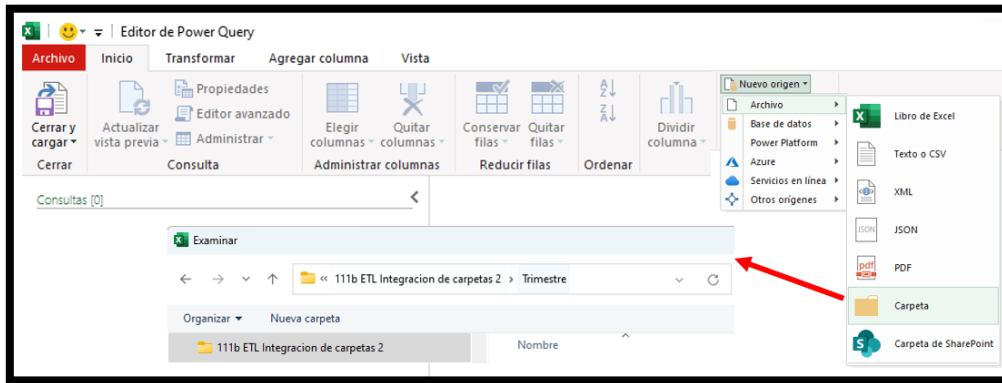
Con el fin de verificar que todo funciona correctamente debemos incorporar los dos ficheros que están en la carpeta Otros y que nos han enviado desde departamento de informática a la carpeta a la que estamos conectados "Trimestre" y comprobar que los cuadros de mando creados se han actualizado correctamente.



Solución- Pasos sugeridos

Como orientación, los pasos sugeridos son:

- Vamos a **crear una conexión a la carpeta** (Trimestre) y a todos los ficheros que contiene y para ello en Excel accedemos al editor de consultas y en la opción nuevo origen seleccionamos la carpeta a la que queremos conectar tal y como se muestra en la ilustración



Así seleccionada la carpeta se nos muestra un nuevo cuadro de diálogo en que vemos el tipo de archivo, nombre, extensión.. etc que contiene en estos momentos en la carpeta conectada.

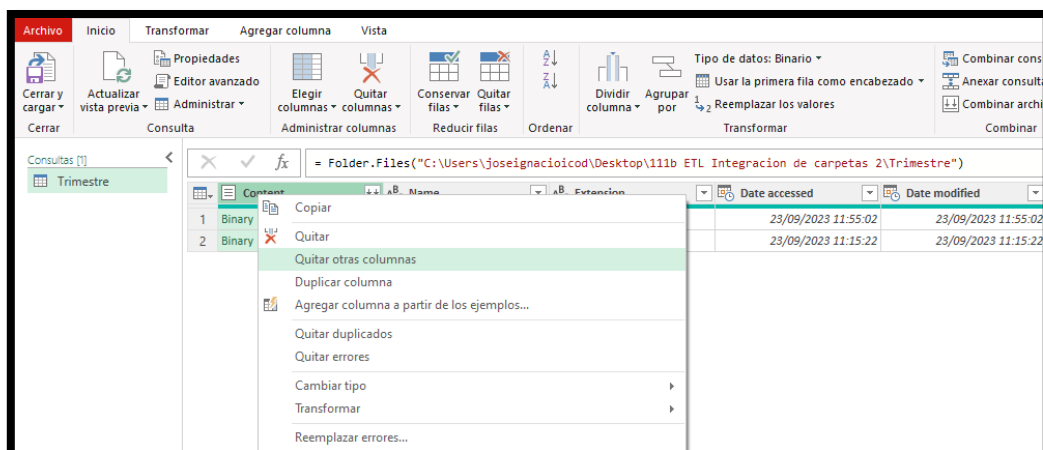
Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
Binary	01Trimestre.csv	.csv	23/09/2023 11:55:02	23/09/2023 11:55:02	23/09/2023 10:42:16	Record	C:\Users\joseignacioicod\Desktop\111b ETL Integracion de carpetas 2\Trimestre\01Trimestre.csv
Binary	02Trimestre.csv	.csv	23/09/2023 11:15:22	23/09/2023 11:15:22	23/09/2023 10:42:16	Record	C:\Users\joseignacioicod\Desktop\111b ETL Integracion de carpetas 2\Trimestre\02Trimestre.csv

Combinar ficheros y transformar datos

Quitar columnas no deseadas de la combinación de ficheros

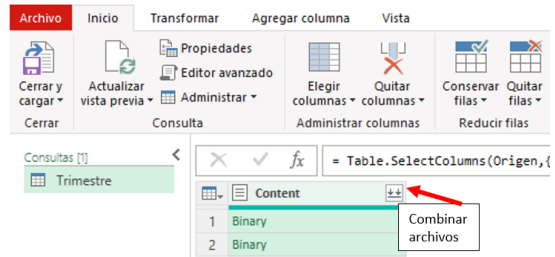
Con esta opción vamos a transformar los datos para quitar anomalías como espacios en blanco o recortar así como garantizar el contenido de las columnas se ajusta al tipo de dato.

De esta forma accedemos al siguiente cuadro de diálogo donde seleccionamos la columna Content y con el botón derecho quitamos el resto de columnas que no nos interesa

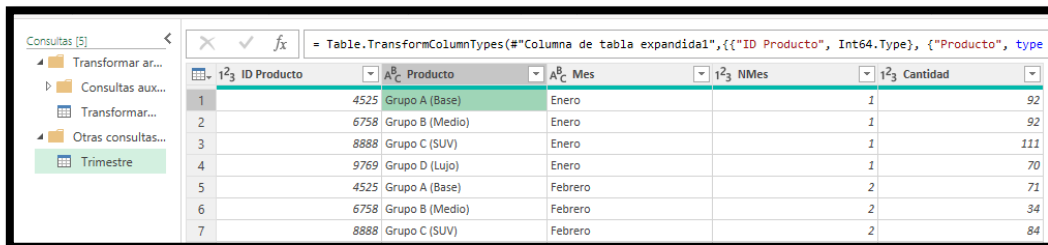
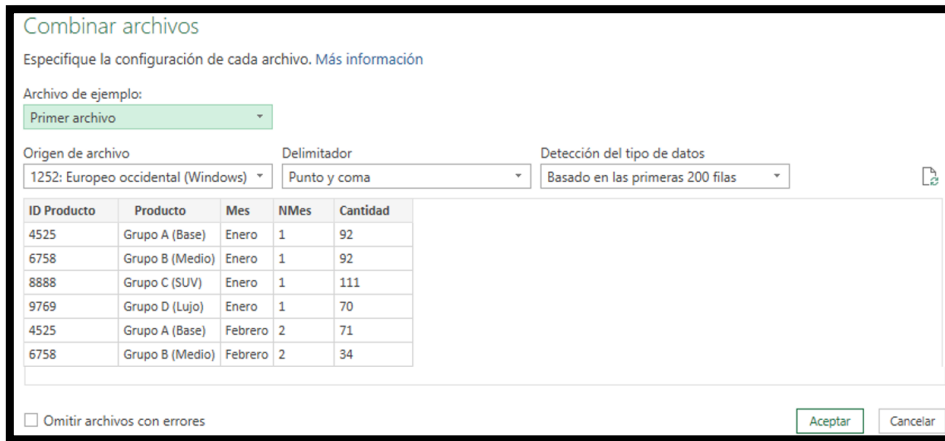


Combinar archivos

Expandimos la columna que nos queda y que contiene la información deseada y desplegando nos da la opción de combinar archivos, obteniendo así una consulta con la información combinada con los dos ficheros tal y como se muestra en la siguiente ilustración.

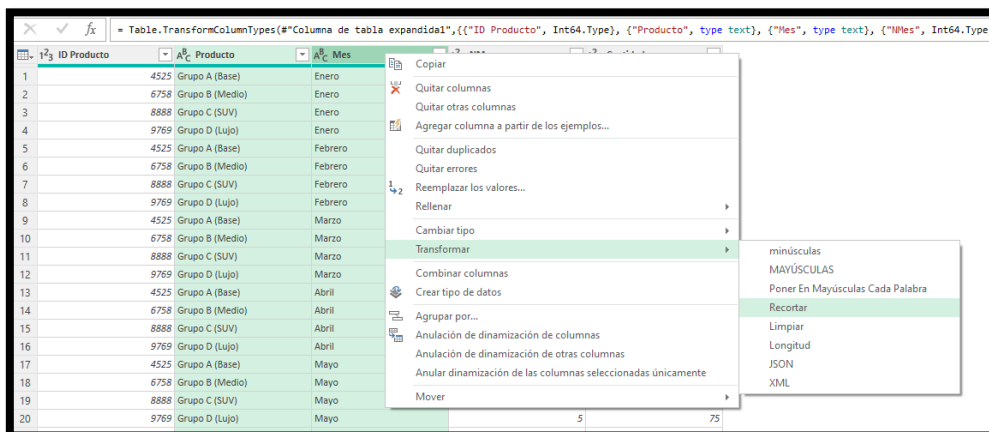


Pero antes de combinar los archivos tendremos que confirmar que el juego de caracteres es correcto así como los delimitadores de columnas. Confirmado se procede a combinar los archivos contenidos en la conexión a la carpeta seleccionada y de esta forma el resultado será una nueva consulta con los registros de los ficheros combinados



Recortar los espacios en blanco y revision de los tipos de campo.

Para quitar los espacios en blanco y depurara así las fuentes, seleccionamos las columnas con campo texto y con el botón derecho accedemos a la opción Transformar-Recortar.



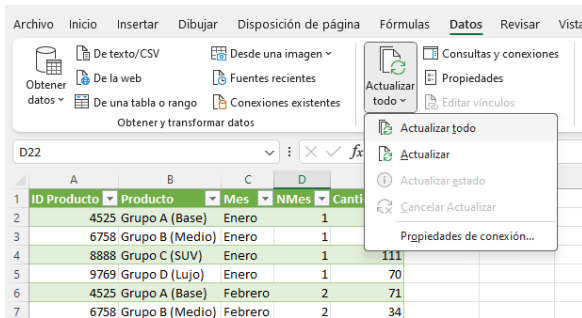
De esta forma queda la primera parte adaptada. El resultado de “Cerrar cargar” será una tabla en nuestro libro de trabajo con los datos conectadas a la consulta.

	A	B	C	D	E
1	ID Producto	Producto	Mes	NMes	Cantidad
2	4525	Grupo A (Base)	Enero	1	92
3	6758	Grupo B (Medio)	Enero	1	92
4	8888	Grupo C (SUV)	Enero	1	111
5	9769	Grupo D (Lujo)	Enero	1	70
6	4525	Grupo A (Base)	Febrero	2	71
7	6758	Grupo B (Medio)	Febrero	2	34
8	8888	Grupo C (SUV)	Febrero	2	84
9	9769	Grupo D (Lujo)	Febrero	2	101
10	4525	Grupo A (Base)	Marzo	3	100
11	6758	Grupo B (Medio)	Marzo	3	55
12	8888	Grupo C (SUV)	Marzo	3	46
13	9769	Grupo D (Lujo)	Marzo	3	36
14	4525	Grupo A (Base)	Abril	4	70
15	6758	Grupo B (Medio)	Abril	4	111
16	8888	Grupo C (SUV)	Abril	4	124
17	9769	Grupo D (Lujo)	Abril	4	45
18	4525	Grupo A (Base)	Mayo	5	57
19	6758	Grupo B (Medio)	Mayo	5	40
20	8888	Grupo C (SUV)	Mayo	5	75
21	9769	Grupo D (Lujo)	Mayo	5	75
22	4525	Grupo A (Base)	Junio	6	50
23	6758	Grupo B (Medio)	Junio	6	62
24	8888	Grupo C (SUV)	Junio	6	58
25	9769	Grupo D (Lujo)	Junio	6	28

Incorporar los dos trimestres que faltan y actualización de datos

Vamos a continuación a incorporar los dos ficheros que faltan (3º y 4º trimestre) a la carpeta de trabajo “Trimestre” para posterior actualizar la conexión de la tabla anterior lo que implica llevar a cabo todos los pasos realizados en el proceso ETL con el editor de consulta con todos los ficheros disponibles y el resultado es la tabla actualizad con todos los registros del año.

- Trimestre
- 01Trimestre
- 02Trimestre
- 03Trimestre
- 04Trimestre



	A	B	C	D	E
1	ID Producto	Producto	Mes	NMes	Cantidad
2	4525	Grupo A (Base)	Enero	1	92
3	6758	Grupo B (Medio)	Enero	1	92
4	8888	Grupo C (SUV)	Enero	1	111
5	9769	Grupo D (Lujo)	Enero	1	70
6	4525	Grupo A (Base)	Febrero	2	71
7	6758	Grupo B (Medio)	Febrero	2	34
8	8888	Grupo C (SUV)	Febrero	2	84
9	9769	Grupo D (Lujo)	Febrero	2	101
10	4525	Grupo A (Base)	Marzo	3	100
42	4525	Grupo A (Base)	Noviembre	11	16
43	6758	Grupo B (Medio)	Noviembre	11	16
44	8888	Grupo C (SUV)	Noviembre	11	52
45	9769	Grupo D (Lujo)	Noviembre	11	29
46	4525	Grupo A (Base)	Diciembre	12	58
47	6758	Grupo B (Medio)	Diciembre	12	63
48	8888	Grupo C (SUV)	Diciembre	12	35
49	9769	Grupo D (Lujo)	Diciembre	12	28

Actualizamos y comprobamos también que nuestros cuadros de mando se han actualizado.

Análisis de días de alquiler según flota y período											
En verde los 5 superiores y en rojo los 5 inferiores											
Mes ▼											
Grupo A (Base)			Grupo B (Medio)		Grupo C (SUV)		Grupo D (Lujo)		Total Nº Di: Total % Hor		
Meses	Nº Dias	% Hor	Nº Dias	% Hor	Nº Dias	% Hor	Nº Dias	% Hor			
Enero	92	25%	92	25%	111	30%	70	19%	365	100%	
Febrero	71	24%	34	12%	84	29%	101	35%	290	100%	
Marzo	100	42%	55	23%	46	19%	36	15%	237	100%	
Abril	70	20%	111	32%	124	35%	45	13%	350	100%	
Mayo	57	23%	40	16%	75	30%	75	30%	247	100%	
Junio	50	25%	62	31%	58	29%	28	14%	198	100%	
Julio	82	25%	107	32%	91	28%	50	15%	330	100%	
Agosto	29	15%	29	15%	88	46%	44	23%	190	100%	
Septiembre	59	31%	67	36%	39	21%	23	12%	188	100%	
Octubre	75	24%	125	40%	69	22%	41	13%	310	100%	
Noviembre	16	14%	16	14%	52	46%	29	26%	113	100%	
Diciembre	58	32%	63	34%	35	19%	28	15%	184	100%	
Total general	759	25%	801	27%	872	29%	570	19%	3.002	100%	

