



# Caso 111 ETL Integración en carpeta

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](http://www.jggomez.eu) **V.2.2**

Ejercicio Basado: Master Class “Limpieza de datos de cero a 100” en [Excel Free Blog - Power Platform](#), Autores: [Miguel Caballero Sierra](#) y [Fabian Torres](#)

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

## Enunciado

Hemos solicitado a los informáticos de nuestra empresa la relación del inventario de los últimos años (periodo 1999-2014) y por limitaciones técnicas nos ha facilitado un fichero por año y los hemos agrupado en la misma carpeta.

Esta situación técnica relativamente frecuente podría darse también por ejemplo con las nóminas de los empleados que se nos facilitarían desde recursos humanos un fichero por mes.

En nuestro caso, estos ficheros se caracterizan por tener todos ellos la misma estructura solo tres columnas, una con el id o código del producto (SKU), una fecha en formato numérico y por tanto tendremos que convertirla (cambiar formato) y la cantidad en inventario.

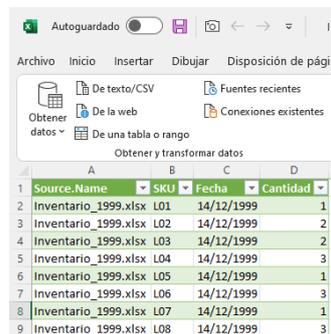


	A	B	C
1	SKU	Fecha	Cantidad
2	L01	36508	1
3	L02	36508	2
4	L03	36508	2
5	L04	36508	3

Es importante tener presente 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)

Aplicar el proceso limpieza, transformación y consolidación, es decir generar una sola consulta integrando todos los ficheros actuales y futuros que contenga la carpeta origen en una sola consulta de tal forma que cualquier nuevo fichero añadido a la carpeta quedara vinculado a la consulta.



## Pasos sugeridos en el proceso ETL

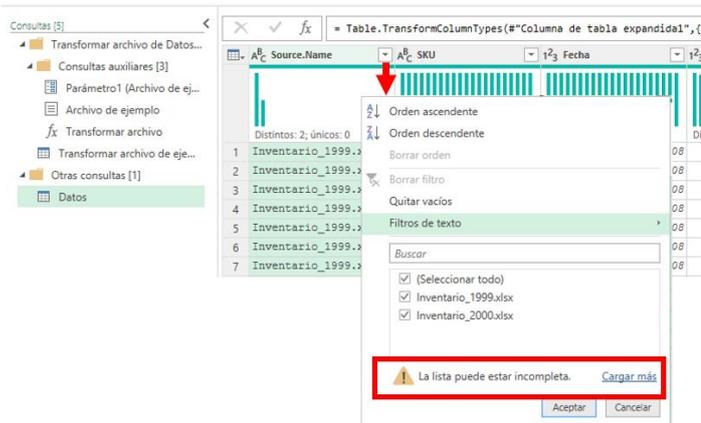
De un primer análisis de la información disponible nos encontramos que esta se encuentra concentrada en una carpeta llamada “Datos” observando que son ficheros normalizados y por tanto no deberíamos tener problemas para su conexión, en todo caso realizaremos una serie de transformaciones como son:

### Proceso ETL con Power Query

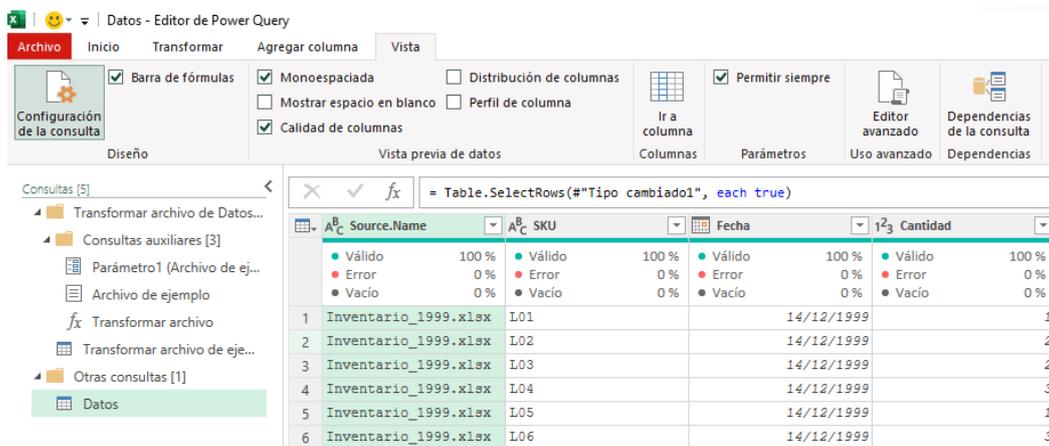
- Vamos a **crear una conexión a la carpeta** (Datos) y a todos los ficheros que contiene y para ello en Excel vamos a menú Datos-Obtener Datos – De un archivo – De una carpeta y simplemente seleccionamos la ruta donde esta nuestra carpeta datos. Recordar que si más adelante modificamos la ubicación del fichero o bien el nombre de la carpeta deberemos actualizar esta referencia.

En este primer paso nos ofrece los metadatos de los archivos y debemos acceder a la opción “Combinar y transformar datos” donde debemos informarle que los datos siempre se localizan en la pestaña “Hoja1” si fuera de distintas pestañas u orígenes tendríamos que cambiar el procedimiento creando la conexión fichero por fichero y después consolidando.

Podemos ver el editor de Power Query y activando la columna Source Name tendremos a disposición todos los ficheros consolidados



- Antes de cerrar el editor PQ aprovechamos para **transformar la columna fecha** y pasarla de tipo numérico a fecha.



- Opcionalmente podríamos quitar la columna “Source.Name” que carece de interés para nuestros análisis.

- Finalmente damos a cerrar y cargar y esta conexión queda resuelta en nuestro libro de Excel

De esta forma facilitamos la conexión y transformación. Si posteriormente incorporamos nuevos ficheros a la carpeta, cada vez que carguemos la hoja de cálculo se volverán a realizar las conexiones y transformaciones actualizando el fichero de Excel con las nuevas incorporaciones.

