



Caso 113 ETL Agrupar por dos criterios o mas

01 ETL y Modelo de Datos. Agrupar una consulta por dos o más criterios.

Jose Ignacio González Gómez

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

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/Agrupacion

Resumen: Usar PQ para agrupar datos por doble criterio, nº de factura y máxima fecha

Técnicas y palabras clave: Crear conexiones y utilizar las opciones de agrupar por avanzada

Enunciado

Contamos con un fichero que contienen dos campos Nº factura que se corresponde al código del cliente (cada cliente tiene una factura máxima por día) y la fecha de la factura correspondiente a todos los servicios prestados al cliente.

En la imagen adjunta muestra el formato de la tabla con las columnas disponibles resaltando las últimas fechas de facturación por cliente.

Esto nos permitirá, por ejemplo, calcular el tiempo transcurrido desde la última facturación utilizando medidas.

	A	B
1	Factura	Datos
2	1	01/01/2021
3	1	07/01/2021
4	1	04/01/2021
5	2	07/02/2021
6	2	09/02/2021
7	2	04/02/2021
8	2	03/02/2021
9	3	05/03/2021
10	3	12/03/2021
11	3	01/03/2021
12	5	01/04/2021
13	5	02/04/2021
14	5	03/04/2021
15	5	04/04/2021
16	5	05/04/2021

Se pide (preguntas base del modelo)

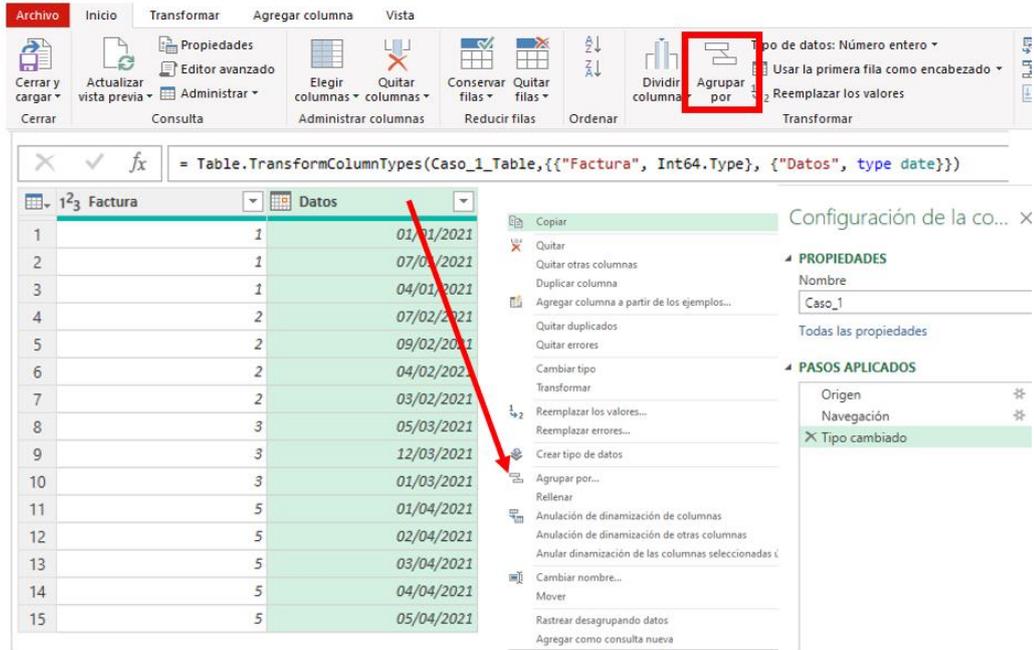
Deseamos tener una conexión que nos genere una consulta agrupada por número de factura (código cliente) y última fecha de facturación tal y como se muestra a continuación.

	Factura	Ultima Fecha Fact.
1	1	07/01/2021
2	2	09/02/2021
3	3	12/03/2021
4	5	05/04/2021

Pasos sugeridos en el proceso ETL

Proceso ETL con Power Query

- Cargamos el PQ en Excel, recordamos Datos – Obtener Datos – Iniciar el editor PQ.
- Vamos a **crear una conexión al fichero que contiene los datos** (nuevo origen).
- Añadimos un paso a la conexión anterior “Agrupar por” que puede ser a través del menú tal y como se muestra en la ilustración o con el botón derecho cuando tenemos seleccionado una columna de la conexión, es decir desde estas dos opciones podremos acceder al Agrupar por.



- Esto nos lleva a un asistente para facilitar esta agrupación y los parámetros básicos en el modo avanzado para nuestro caso son los que se muestran en la siguiente ilustración.

Agrupar por

Especifique las columnas por las que quiera realizar la agrupación y una o más salidas.

Básico Uso avanzado

Factura

Agregar agrupación

Nuevo nombre de columna

Ultima Fecha Fact.

Agregar agregación

Operación

Máx.

Columna

Datos

Aceptar

Cancelar

De esta forma se genera la consulta agrupada por los requerimientos demandados.