



Power query

ANÁLISIS Y VISUALIZACIÓN DE DATOS

Power query

- ▶ Power query nos permite **obtener, combinar y refinar** datos de una amplia variedad de orígenes:
 - ▶ Desde la web
 - ▶ Desde un archivo (Excel, csv, xml, texto)
 - ▶ Desde una base de datos
 - ▶ Desde Hadoop
 - ▶ Desde Facebook
 - ▶ ...

Importar datos

- ▶ Ejemplos
 - ▶ Desde la web
 - ▶ Desde un archivo
 - ▶ Desde una base de datos
 - ▶ Desde una tabla de datos de Excel

Importar datos desde la web

- ▶ Vamos a importar datos sobre los premios Goya, obtenidos de la Wikipedia
- ▶ En un nuevo libro de Excel, seleccionamos la pestaña POWER QUERY
- ▶ En obtener datos externos, elegimos desde la web e introducimos la url
 - ▶ https://es.wikipedia.org/wiki/Premios_Goya
- ▶ Seleccionamos la tabla que nos interesa
 - ▶ Intérpretes galardonados con el premio Goya
- ▶ Podemos cargar directamente la tabla o **editarla**, para tener exactamente la información que nosotros queremos

Editor de consultas

- ▶ Yo solo quiero trabajar con las actrices galardonadas, por lo que puedo eliminar aquellas columnas referentes a los actores
- ▶ Si no quiero trabajar con todos los datos disponibles, puedo eliminar aquellas columnas y filas que no me interesen
 - ▶ Administrar columnas
 - ▶ Elegir columnas
 - ▶ Quitar columnas / Quitar otras columnas
 - ▶ Reducir filas
 - ▶ Conservar filas
 - ▶ Quitar filas

Editor de consultas

- ▶ Tengo varias columnas que incluyen dos informaciones en una misma columna: actriz y película. Quiero separarlos:
 - ▶ Transformar → columna de texto → dividir columna; por delimitador
 - ▶ Personalizado: (
 - ▶ En cada aparición del delimitador
 - ▶ Transformar → reemplazar los valores
 - ▶ Elimino números y paréntesis
 - ▶ Transformar → combinar columnas

Editor de consultas

- ▶ Renombrar columnas
 - ▶ Doble click sobre el enunciado de la columna
- ▶ En el panel izquierdo de configuración de la consulta puedo ver los pasos aplicados, además de modificarlos o eliminarlos
- ▶ También puedo cambiar el nombre de mi consulta:
Actrices de los Goya

Editor de consultas


- ▶ Una vez finalizada la consulta, debemos cargarla en nuestro libro de Excel

- ▶ Cerrar y cargar
- ▶ Cerrar y cargar en

Cargar en

Seleccione cómo desea ver estos datos en el libro.



 ☐ Crear solo conexión

Seleccione dónde se deben cargar los datos.

☒ Nueva hoja de cálculo

☐ Hoja de cálculo existente:



☐ Agregar estos datos al Modelo de datos

Cargar

Cancelar

- ▶ En el ejemplo, cargar en tabla y agregar al modelo de datos

Repaso. Modelo de datos y tablas dinámicas

- ▶ Mostrar en una tabla dinámica las ediciones en las que una misma película haya obtenido al menos dos de los tres premios seleccionados
 - ▶ Nueva columna: 2 pelis igual
 - ▶ =OR(EXACT([Película protagonista];[Película de reparto]);OR(EXACT([Película protagonista];[Película revelación]);EXACT([Película de reparto];[Película revelación])))
 - ▶ Tabla dinámica
 - ▶ Filas: edición
 - ▶ Filtros: 2 pelis igual (sólo nuestro TRUE)

Repaso. Modelo de datos y tablas dinámicas

- ▶ Además de la edición, mostrar también el título de la película
 - ▶ Nueva columna calculada
 - ▶ =IF(EXACT([Pelicula protagonista];[Pelicula de reparto]);[Pelicula protagonista];IF(EXACT([Pelicula protagonista];[Pelicula revelacion]);[Pelicula protagonista];IF(EXACT([Pelicula de reparto];[Pelicula revelacion]);[Pelicula de reparto];""))))
 - ▶ Tabla dinámica
 - ▶ Filas: edición y película dos premios
 - ▶ Filtros: 2 pelis igual (sólo nuestro TRUE)

Importar datos desde un libro de Excel

- ▶ En un nuevo libro de Excel, seleccionamos la pestaña POWER QUERY
- ▶ En obtener datos externos, elegimos desde archivo y, dentro de estas opciones, desde Excel
- ▶ Buscamos y elegimos el archivo de Excel que contiene los datos que queremos introducir (Products.xlsx)
- ▶ Podemos cargar directamente los datos, o preparar una nueva consulta a partir de ellos, mediante la opción **Editar**

Editor de consultas

- ▶ Si los nombres de las columnas de la vista previa no son las que deberían, podemos hacer que la primera fila de datos se convierta en encabezado:
 - ▶ Transformar → Usar primera fila como encabezado

Editor de consultas

- ▶ Si no quiero trabajar con todos los datos disponibles, puedo eliminar aquellas columnas y filas que no me interesen
 - ▶ Administrar columnas
 - ▶ Elegir columnas
 - ▶ Quitar columnas / Quitar otras columnas
 - ▶ Reducir filas
 - ▶ Conservar filas
 - ▶ Quitar filas
- ▶ En el ejemplo, nos quedamos con las columnas ProductID, ProductName, CategoryID, QuantityPerUnit

Editor de consultas

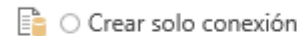
- ▶ En el panel izquierdo de configuración de la consulta puedo ver los pasos aplicados, además de modificarlos o eliminarlos
- ▶ En Vista → Editor avanzado también se pueden ver los pasos y añadir los que queramos en el lenguaje “M”

Editor de consultas

- ▶ Una vez finalizada la consulta, debemos cargarla en nuestro libro de Excel
 - ▶ Cerrar y cargar
 - ▶ Cerrar y cargar en

Cargar en

Seleccione cómo desea ver estos datos en el libro.



Seleccione dónde se deben cargar los datos.

☒ Nueva hoja de cálculo

☐ Hoja de cálculo existente:



☐ Agregar estos datos al Modelo de datos

Cargar

Cancelar

- ▶ En el ejemplo, creamos solo la conexión

Importar datos desde una fuente OData

- ▶ En la pestaña POWER QUERY, obtener datos externos, elegimos desde otros orígenes y, dentro de estas opciones, desde una fuente de OData
- ▶ Introducimos la url de donde están nuestros datos
 - ▶ <http://services.odata.org/Northwind/Northwind.svc>
- ▶ Elegimos la tabla con la que queremos trabajar y accedemos al editor de consultas mediante la opción **Editar**
- ▶ En nuestro ejemplo trabajamos con la tabla Orders

Editor de consultas

- ▶ En este ejemplo nos fijamos en las columnas Customer, Employee, Order_Details y Shipper. Los valores son en todos los casos Record o Table
 - ▶ Record: relacionado con una fila de otra tabla
 - ▶ Table: relacionado con una o varias filas de otra tabla
- ▶ Podemos expandir estas relaciones, añadiendo a los datos tantas columnas como tenga la tabla relacionada
 - ▶ Si es Table, los registros de la tabla se replican tantas veces como filas relacionadas
- ▶ En el ejemplo, de la tabla Order_Details expandimos solamente las columnas ProductID, UnitPrice y Quantity

Editor de consultas

- ▶ De toda la información disponible, yo quiero trabajar con las líneas de los pedidos. Por lo tanto, me quedo sólo con las columnas
 - ▶ OrderDate
 - ▶ ProductID
 - ▶ UnitPrice
 - ▶ Quantity

Editor de consultas

- ▶ Para cada línea de pedido, quiero calcular el precio total de ese elemento
 - ▶ Agregar columna personalizada
 - ▶ Nombre = Total de línea
 - ▶ Fórmula = `[Order_Details.UnitPrice]*[Order_Details.Quantity]`

Editor de consultas

- ▶ En lugar de la fecha del pedido completa, sólo me quiero quedar con el año
 - ▶ Botón derecho sobre el encabezado de la columna → transformar → año
 - ▶ Seleccionar columna. Pestaña transformar → columna de fecha y hora → año → año

Editor de consultas

- ▶ Podemos cambiar el nombre (encabezado) de nuestras columnas
 - ▶ Doble click en el encabezado
- ▶ OrderDate pasa a ser Año
- ▶ A las otras tres columnas, les quitamos el prefijo

Editor de consultas

- ▶ Podemos agrupar varias filas en una sola
 - ▶ Transformar → Agrupar por
- ▶ En el ejemplo queremos que sólo haya una línea por cada producto y cada año. En esta línea queremos calcular la suma del total de líneas. Es decir, queremos el total de dinero ingresado en cada año para cada producto

Agrupar por

Especifique las columnas por las que se realizará la agrupación.

Agrupar por	+
<input type="text" value="Año"/>	-
<input type="text" value="ProductID"/>	-

Nuevo nombre de columna	Operación	Columna	+
<input type="text" value="Ventas totales"/>	<input type="text" value="Suma"/>	<input type="text" value="Total de línea"/>	-

Aceptar

Cancelar

Editor de consultas

- ▶ Cambiamos el nombre de la consulta para que sea más intuitivo
 - ▶ Panel izquierdo de configuración de consulta
- ▶ Cargamos la consulta (sólo conexión)

Combinar y anexar consultas

- ▶ Cuando tenemos varias consultas, podemos unirlos para que formen una sola
 - ▶ Anexar
 - ▶ Unir las filas de una consulta a continuación de las filas de la otra
 - ▶ Combinar
 - ▶ Crear filas con las columnas de las dos consultas, teniendo en cuenta el campo por el que se pueden relacionar

Combinar y anexar consultas

- ▶ Vamos a combinar las dos consultas hechas previamente
 - ▶ Combinamos Products con Ventas
 - ▶ Es necesario marcar las columnas por las que se van a unir (valores coincidentes)
 - ▶ ProductID

Editor de consultas

- ▶ Nuestra consulta combinada tiene una columna expansible, del tipo Table
 - ▶ Expando la columna
 - ▶ Como la columna con el ID de producto ya la tengo, elijo sólo las otras dos
 - ▶ Año
 - ▶ Ventas totales
 - ▶ Renombro las nuevas columnas
- ▶ Renombro la consulta
- ▶ Cargo la consulta, sólo su conexión, en el modelo de datos

Repaso. Modelo de datos y tablas dinámicas

- ▶ Mostramos en una tabla dinámica, para cada producto, el importe total de ventas por año
- ▶ Insertar tabla dinámica
 - ▶ Filas: ProductName
 - ▶ Columnas: Año
 - ▶ Valores: suma de ventas totales