

Inteligencia de Negocios

Business Intelligence

Open Data y Limpieza de Datos con Excel y Power Query

1. Datos Abiertos (**Open Data**) características
2. Sector Infomediario. Valor del Open Data
3. Limpiar y reestructurar datos importados
4. Funciones y Herramientas de Excel que nos pueden ayudar
5. Power Query para limpiar y estructurar datos

Open Data

son datos que pueden ser

utilizados, reutilizados y redistribuidos libremente

por cualquier persona,

incluso para fines comerciales,

para **cualquier finalidad**

y que se encuentran sujetos, cuando más,

al **requerimiento de atribución**

y de compartirse

de la misma manera en que aparecen

(Open Data Handbook, 2010)



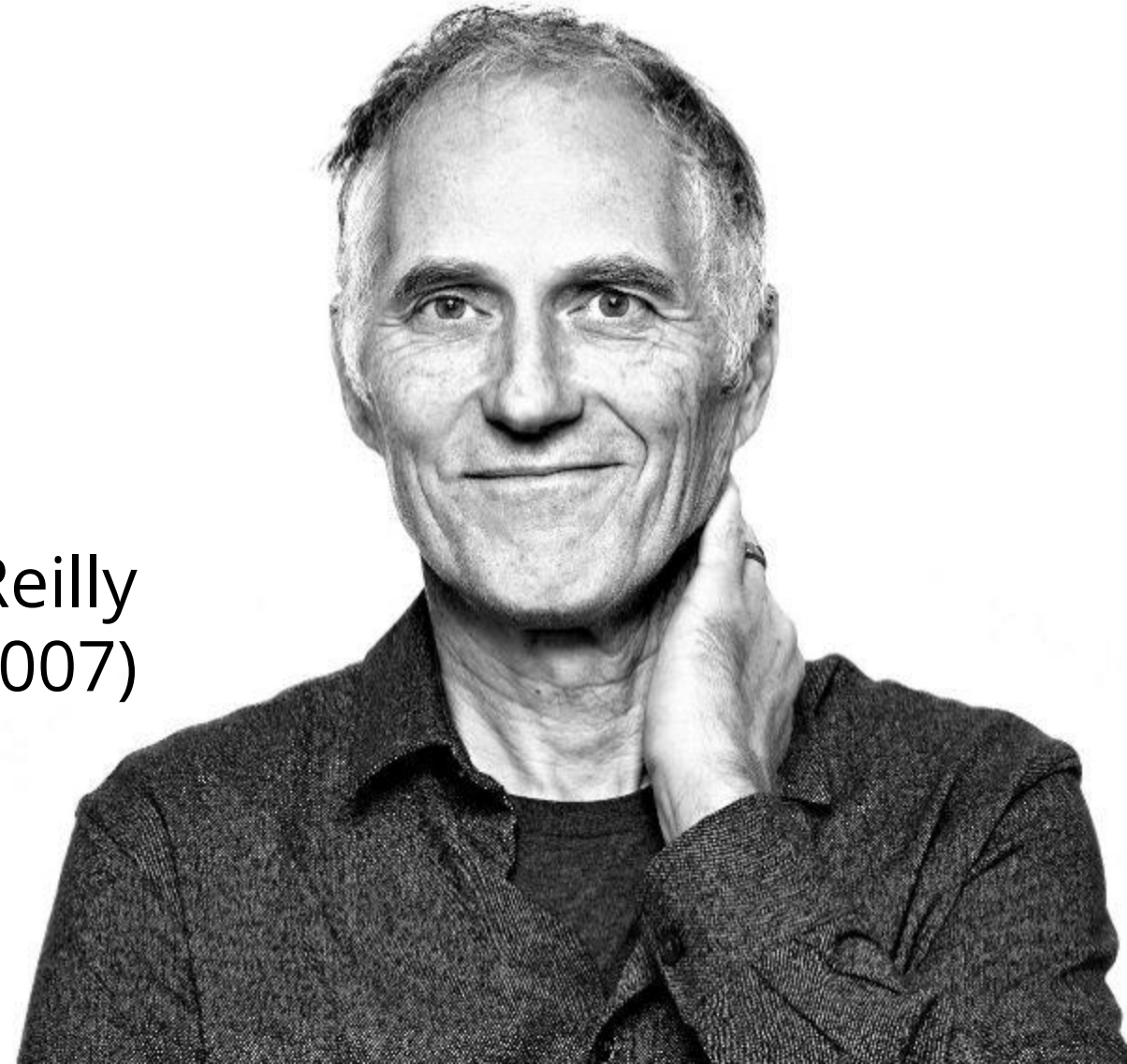
<https://okfn.org/>



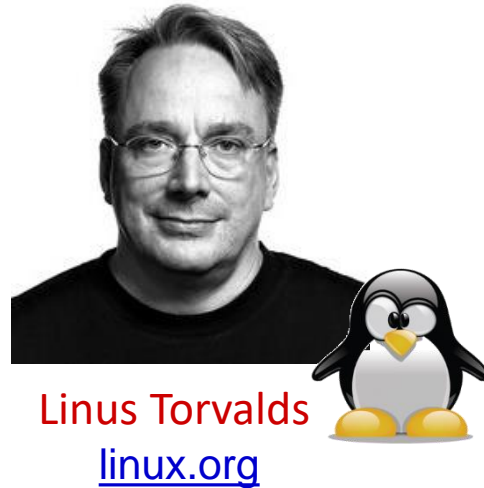
Los datos abiertos
en realidad
podrían ser
más importantes
que el código abierto

Tim O'Reilly
(2007)

<https://www.oreilly.com/>



Surge como una de las iniciativas que tienen origen en la filosofía abierta de internet.



Lawrence Lessig
creativecommons.org

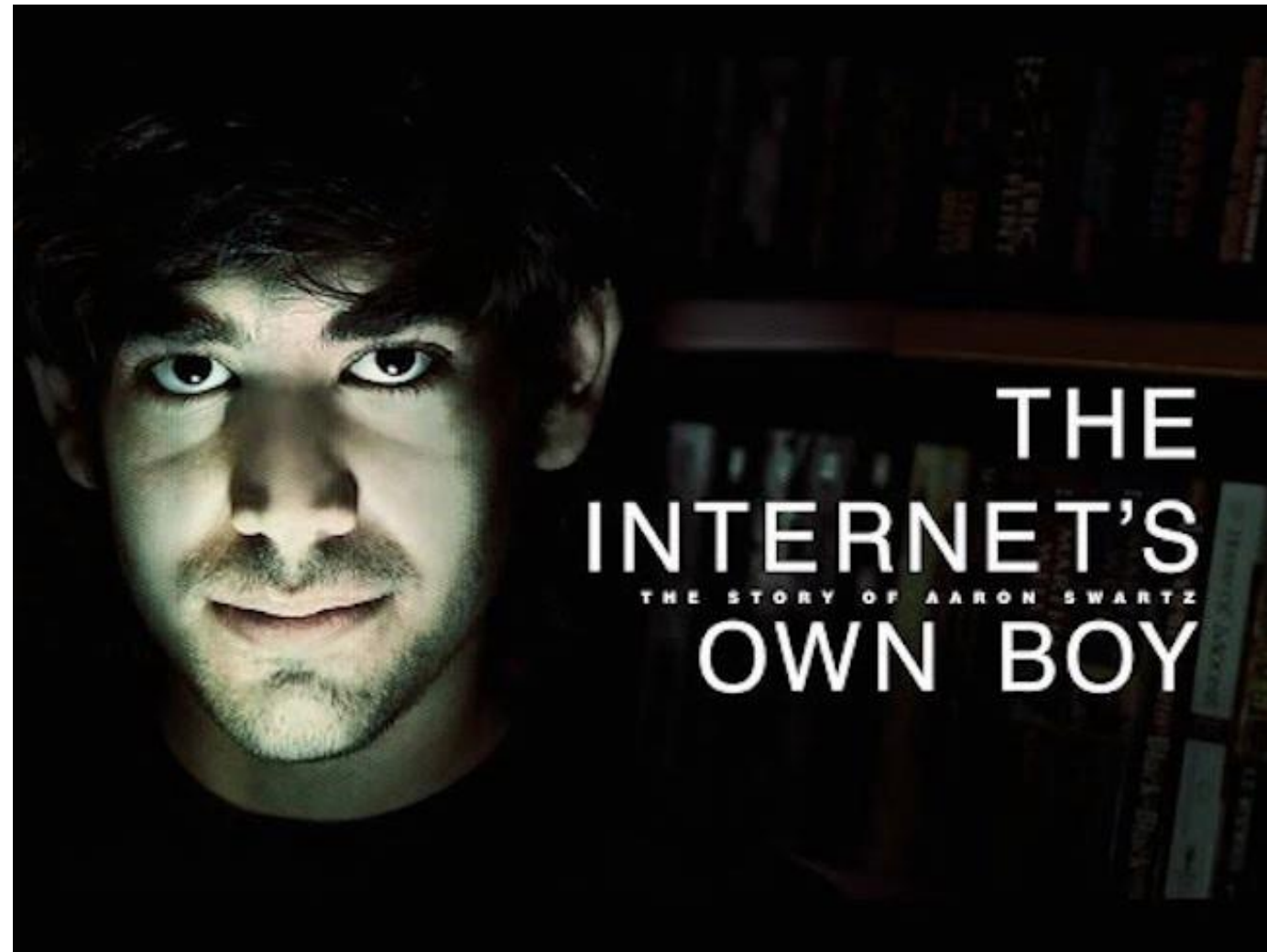


Sir Tim Berners Lee



www.w3c.es
w3schools.com

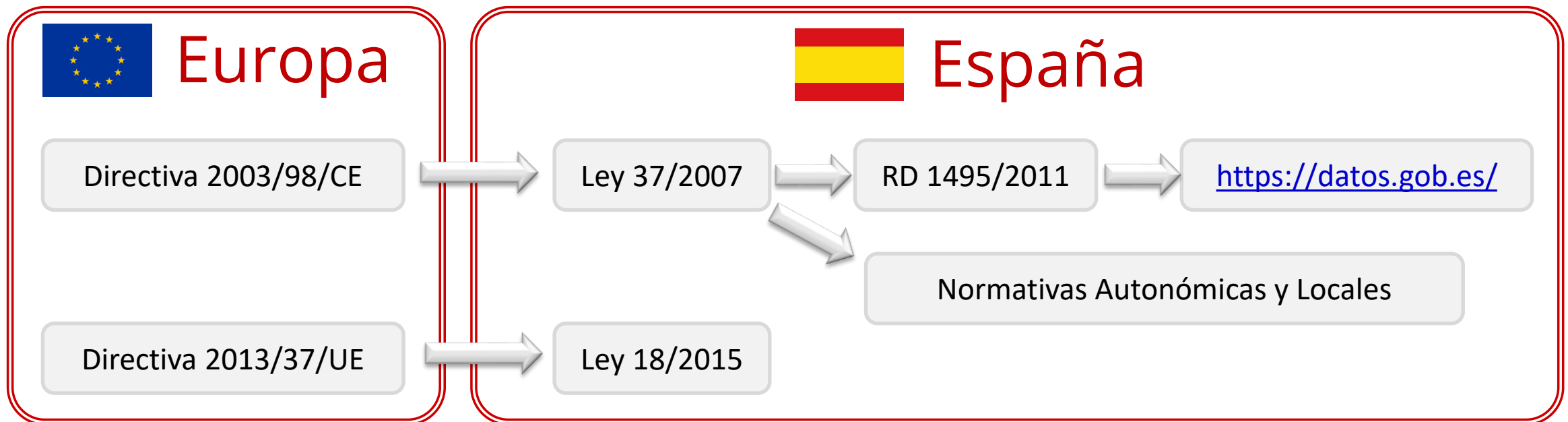




<https://www.youtube.com/watch?v=mT8Fjclx3HI>

Open Data juega un importante papel en el Open Government, ya que **garantiza** a los ciudadanos el **acceso a la información** generada por los **gobiernos** y las **administraciones públicas** en el transcurso de sus actividades.

El marco general para garantizar que las condiciones de reutilización de los documentos del sector público fueran equitativas, proporcionadas y no discriminatorias, se encuentra en la **DIRECTIVA 2003/98/CE** del Parlamento Europeo y del Consejo.

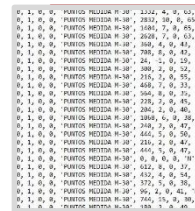


recomendaciones:

Los Datos deben ser



Públicos: No sujetos a limitaciones de privacidad, seguridad o privilegios



En Bruto: Se deben publicar tal y como se recogen de la fuente, manteniendo el mayor nivel de detalle posible.



Actualizados: En rápida disposición para preservar el valor de los mismos.



Accesibles: Disponibles para el mayor número de usuarios posible y para la más amplia gama de propósitos.



Estructurados: Que puedan ser procesados de manera automática por cualquier ordenador.



Sin Registro: (*acceso anónimo*) Disponibles para cualquier persona, sin necesidad de registro.



Abiertos: Disponibles en un formato tal, que ninguna entidad posea su control exclusivo.



Libres: ** No sujetos a ninguna norma de derechos de autor, patentes o copyright

Sistema de 5 estrellas de Tim Berners-Lee

<https://5stardata.info/es/>

				
Datos no estructurados	Datos estructurados	Datos estructurados Formato no propietario	Datos estructurados Formato no propietario URL única y estática	Datos estructurados Formato no propietario URL única y estática Bases de Datos interconectadas
				

Es importante destacar que aunque los datos estén en un determinado formato, esto no quiere decir que los mismos se encuentren estructurados de la forma que necesitamos



Los datos están disponibles en la web,
sea en el formato que sea,
con licencia abierta.

Sin embargo,
los datos están atrapados
en un documento.

Para su utilización,
deben transcribirse
o utilizar software para extraerlos.

*Para realizar un análisis sobre los mismos,
se requiere reestructurarlos.*

Tabla 1: Características de la población. Muestras Seguridad Social y funcionarios públicos. ENS 2003-2012

VARIABLES	Todas las bases ENS 2003, 2006 y 2012		
	Todas=72.135	SS=68.390	Sólo mutualidad pública=3.745
SEXO			
Mujeres	56.8%	56.9%	52.0%
Varones	43.2%	43.1%	48.0%
GRUPOS DE EDAD			
16-24	8.2%	8.2%	8.3%
25-54	50.2%	49.9%	55.7%
55-64	14.1%	14.0%	16.5%
65 o más	27.5%	28.0%	19.6%
ESTADO CIVIL			
Soltero	26.7%	26.7%	26.7%
Casado	55.3%	55.0%	59.2%
Viudo	12.8%	13.0%	8.6%
ESTUDIOS			
ESO o menos	61.5%	63.3%	28.6%
Secundaria segunda etapa	18.0%	17.9%	21.3%
Superior	20.5%	18.9%	50.0%
ESTADO DE SALUD			
Muy bueno	13.7%	13.4%	19.7%
Bueno	50.9%	50.6%	56.7%
Regular	25.4%	25.9%	18.0%
Malo	7.7%	7.9%	4.6%
Muy malo	2.2%	2.3%	1.0%
PESO			
Insuficiente	5.1%	5.1%	5.6%
Normal	53.2%	53.0%	56.4%
Sobrepeso	34.3%	34.4%	32.5%
Obesidad	7.0%	7.1%	5.3%
SITUACION LABORAL			
Jubilado (65 o más)	31.9%	32.4%	24.0%
Activo	49.7%	49.1%	59.8%
Inactivo	18.4%	18.5%	16.3%
COMUNIDAD AUTÓNOMA			
Andalucía	9.7%	9.7%	9.1%
Aragón	7.0%	6.9%	7.1%
Asturias	3.5%	3.5%	3.7%
Islas Baleares	4.3%	4.4%	3.2%
Canarias	4.2%	4.2%	3.7%
Cantabria	4.3%	4.4%	2.9%
Castilla La Mancha	9.6%	9.4%	12.8%
Castilla y León	4.2%	4.2%	3.8%
Cataluña	9.1%	9.2%	7.3%
Com. Valenciana	6.6%	6.7%	4.9%
Extremadura	3.5%	3.5%	3.4%
Galicia	7.9%	8.0%	7.3%
Com. de Madrid	7.7%	7.5%	10.2%
Murcia	4.9%	4.9%	5.2%
Navarra	4.3%	4.4%	2.7%
País Vasco	4.6%	4.6%	4.7%
La Rioja	2.8%	2.8%	2.1%
Ceuta	1.6%	1.4%	5.1%
Melilla	0.4%	0.3%	0.9%



Los datos están disponibles en la web,
en forma estructurada
(es decir, en formato para máquinas)
y con licencia abierta.



Sin embargo,
todavía **son solo datos**
publicados en la Web
y no datos
integrados a la Web.



*Es posible,
que para realizar un análisis
sobre los mismos,
se requiera reestructurarlos.*

The screenshot shows the INE (Instituto Nacional de Estadística) website. A dropdown menu titled 'Descargar' is open, listing various export formats: Excel (XLS and XLSx), CSV (separated by tabs or semicolons), Pc-Axis, Json, and Text (separated by tabs or semicolons). The background shows a table titled 'Principales series desde 1971' with columns for Sexo, Edad, Provincias, and Periodo. The table is filtered for 'Ambos sexos', 'Total', 'Total Nacional', and '1 de julio de 2020'.

Sexo	Edad	Provincias	Periodo
Ambos sexos	Total	Total Nacional	1 de julio de 2020
Hombres	0 años	02 Albacete	1 de enero de 2020
Mujeres	1 año	03 Alicante/Alacant	1 de julio de 2019
	2 años	04 Almería	1 de enero de 2019
	3 años	01 Araba/Álava	1 de julio de 2018



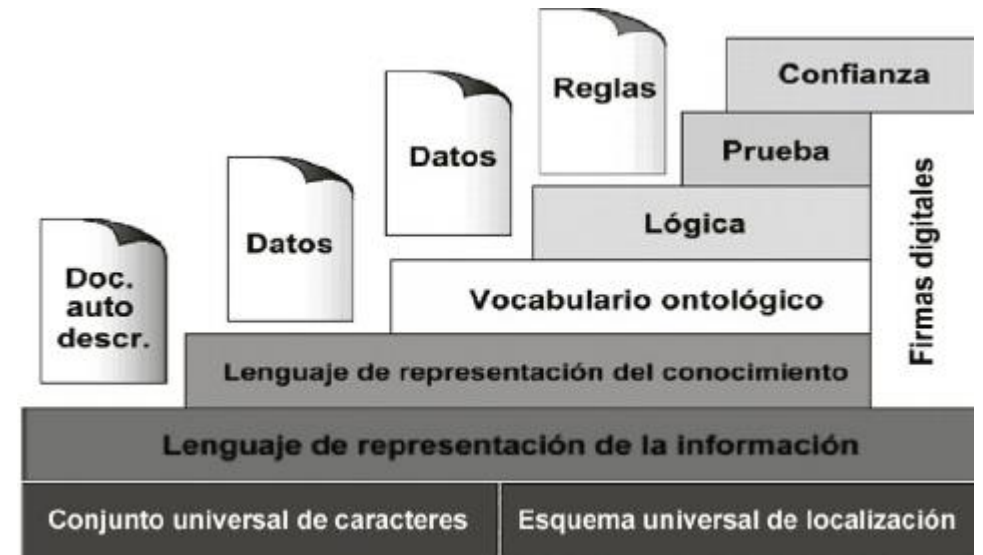
Los datos están disponibles en la web, en forma estructurada y con licencia abierta.



Se puede enlazar los datos desde otro sitio web, marcarlos como favoritos, reutilizar partes de los mismos, usar herramientas y librerías existentes o incluso combinarlos con otros datos.



Una forma de representar los datos es utilizar **RDF**, sin embargo otros formatos como **Atom** pueden ser convertidos/mapeados, si se requiere.





Las cinco estrellas se consiguen logrando los niveles anteriores y además, enlazándolos con otros datos.

Tanto el consumidor de los datos, como el editor de los mismos se benefician del efecto red.

Vincular los datos con los datos de otras organizaciones, proporciona contexto.

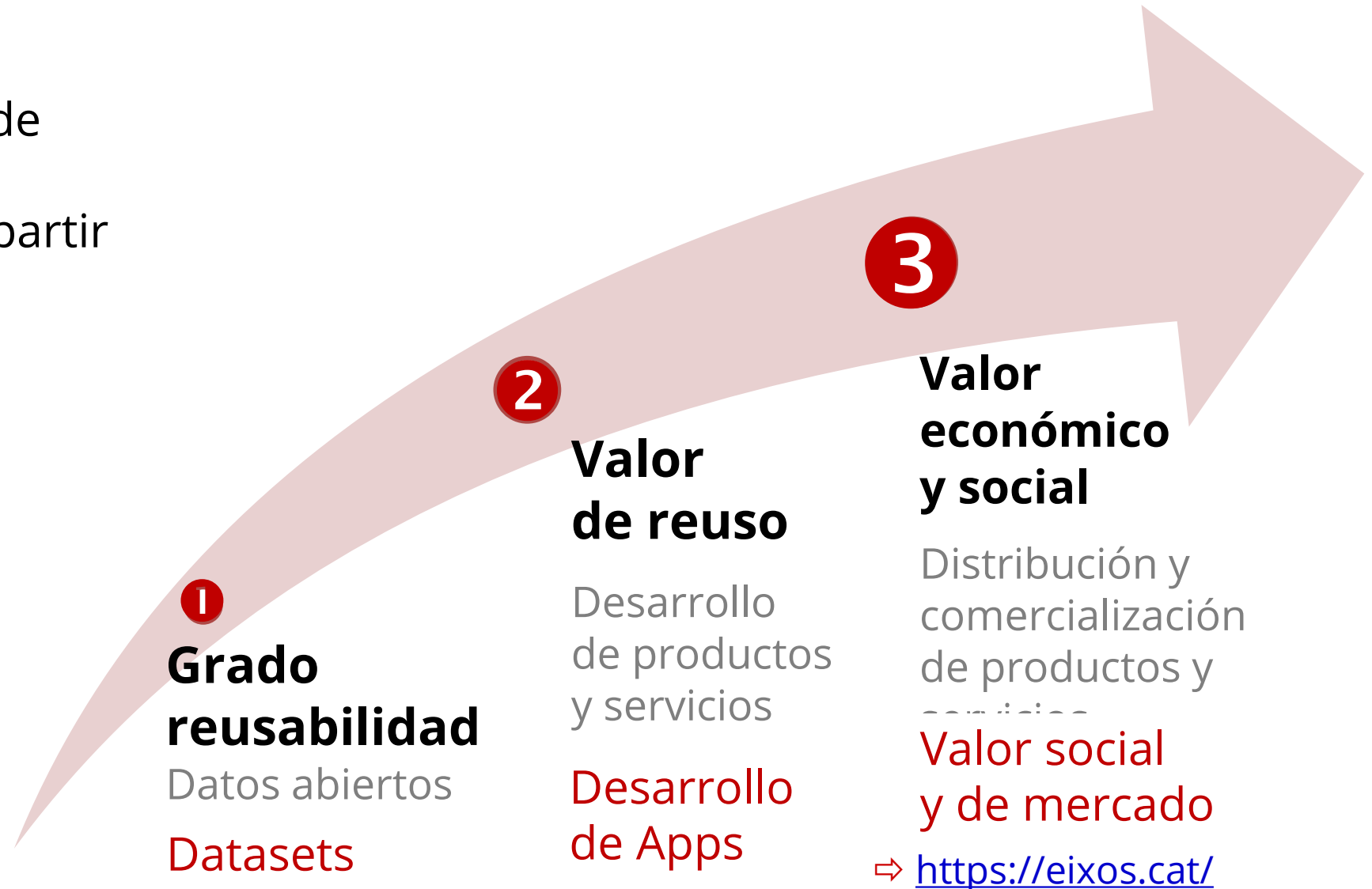
https://opendatabarometer.org/country-detail/?_year=2016&indicator=ODB&detail=ESP

https://data.europa.eu/sites/default/files/open_data_maturity_report_2019.pdf



El sector **infomediario** representa al conjunto de empresas que generan productos y servicios a partir del Open Data.

Incluye a las empresas que se han creado específicamente para este fin, como a aquellas que incorporan de alguna manera los datos abiertos.



	2016
Empresas Identificadas	636
Volumen de Negocio	1.704.829.680 €
Empleados	19.362
Capital Suscrito	366.023.315 €

El valor de los Datos en el Ecosistema Global

Encuentro Aporta. 7ª Edición. 2017

https://datos.gob.es/sites/default/files/aportameeting/evento_aporta-asedie_dionisio_torre_v2_1.pdf

Del sector infomediario a la economía del dato

Observatorio Nacional de Tecnología y Sociedad (ONTSI)

<https://www.ontsi.red.es/sites/ontsi/files/2020-06/EstudioInfomediarioEdicion2020.pdf>

Indicadores ONTSI

<https://www.ontsi.red.es/es/indicadores>

Oferta del Sector

56,8%

DATOS TRATADOS

55,6%

PUBLICACIONES

45,7%

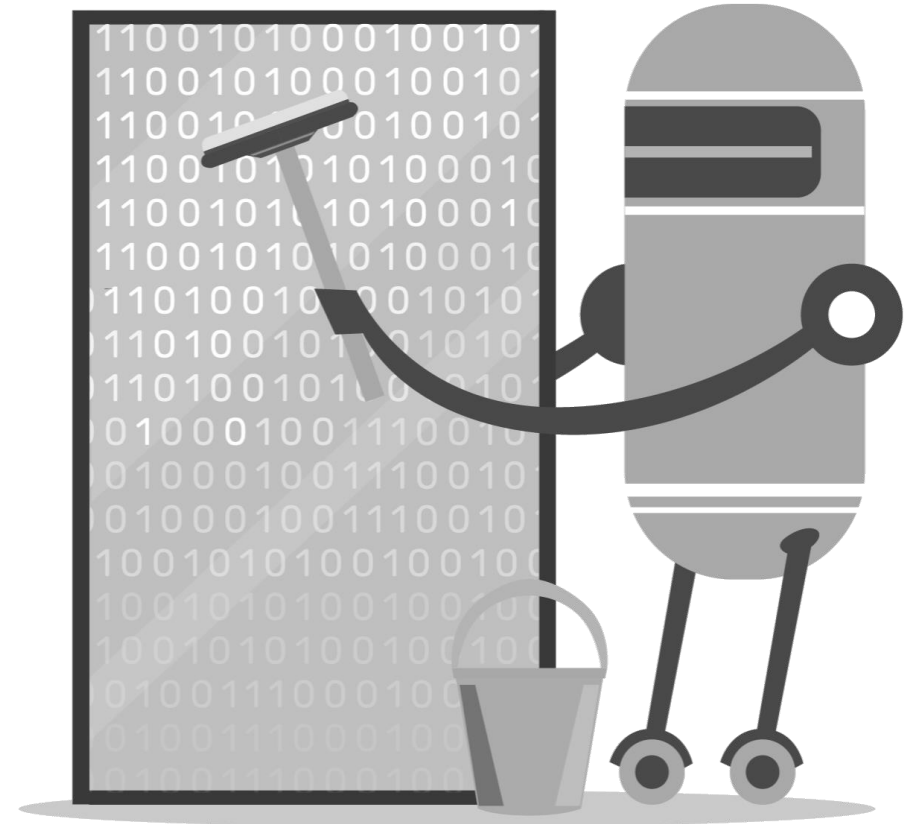
MAPAS Y
VISUALIZACIONES

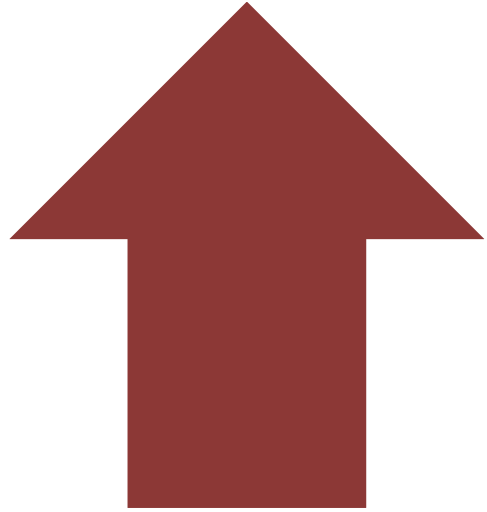
¿Cómo limpiar y reestructurar los datos?

*Es casi seguro,
que para realizar un análisis con datos importados,
se requiera reestructurarlos.*

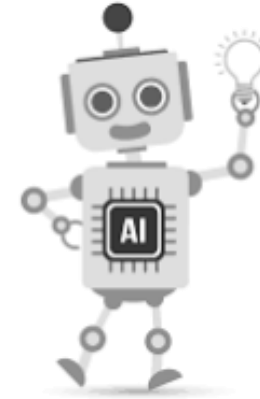
*La recogida,
limpieza
y transformación de datos
constituye un 80%
del trabajo de análisis*

<https://blog.sedic.es/2019/03/18/los-cientificos-de-datos-pasan-el-80-del-tiempo-limpiando-datos-y-el-otro-20-protestando-por-tener-que-limpiarlos/>



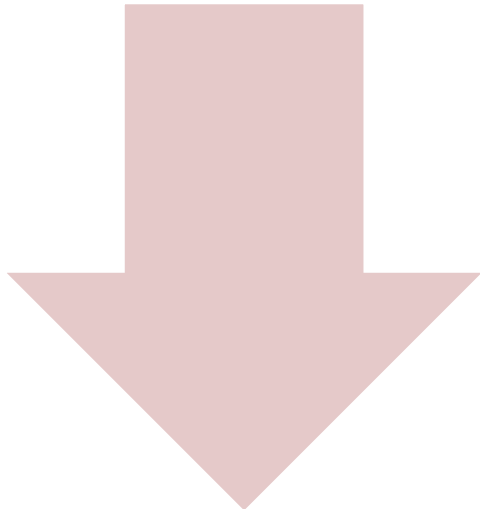


Análisis
recurrentes sobre
los
mismos datos



Data Cleaning Software
ETL Software
Macros

**Automatización de la
limpieza de los datos**

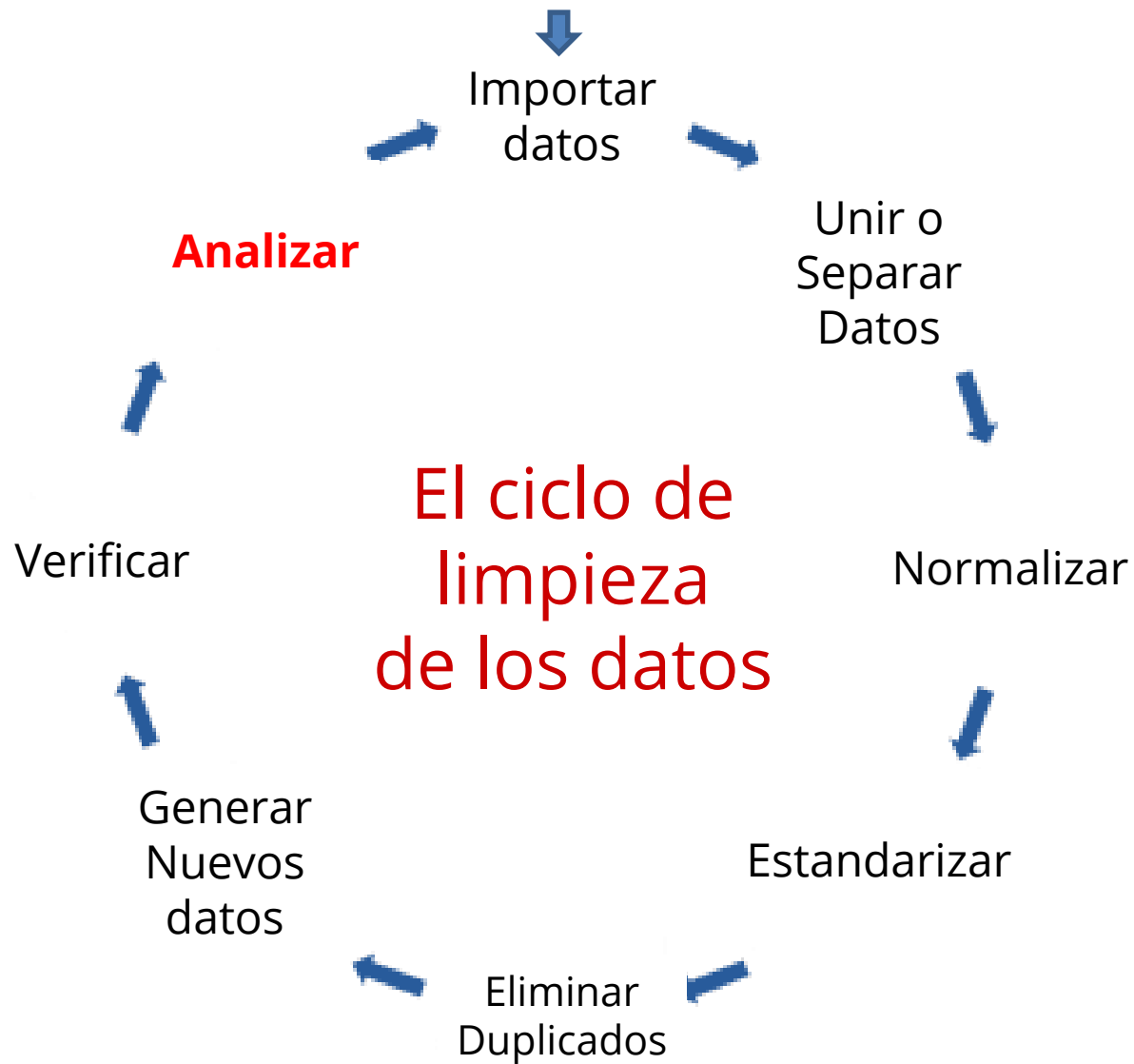


Análisis
circunstancial
sobre un tópico



Power Query

**Limpieza manual,
ayudado de un software**



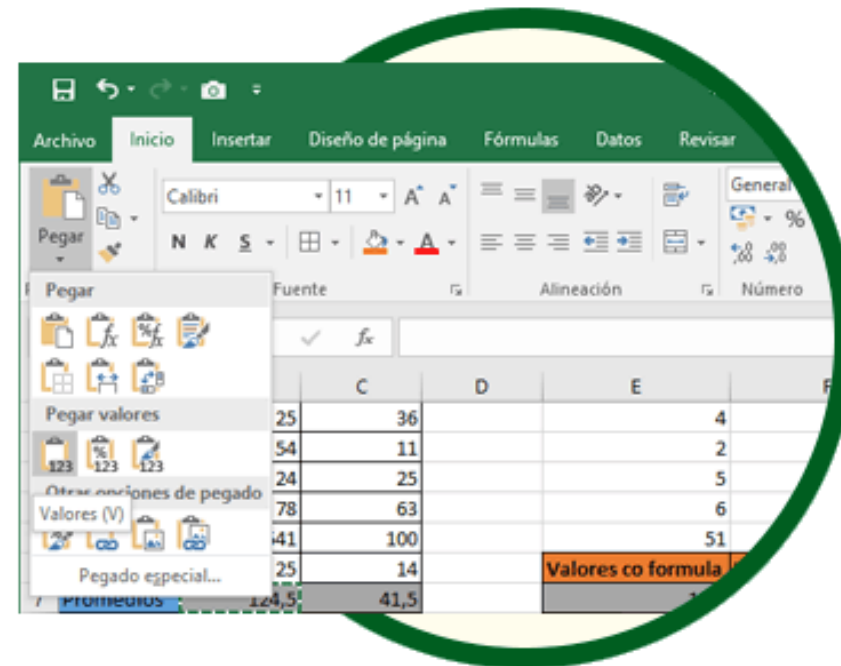
Los errores en el tratamiento de datos son tan comunes que existe una organización llamada *Grupo Europeo de Evaluación de Riesgos para Hojas de Cálculo* (**ESRIG**) que hace recomendaciones al respecto.



Funciones y Herramientas de Excel que nos pueden ayudar



Tomar en cuenta que una vez realizado los cambios con estas funciones, habrá que copiar y pegar valores en la columna que contiene la fórmula y eliminar la columna que está “sucia”.



El operador & y la Función CONCATENAR



El símbolo u operador **&** permite unir o concatenar una o más cadenas de texto con el fin de generar un único fragmento de texto.

La función CONCATENAR realiza la misma acción.

`=CONCATENAR (referencia1/texto1; referencia2/texto2;...)`

	A	B	C
1	Nombre	Apellido	Trato
2	Rita	de la Torre	Prof.
3	Paco	Pérez	Señor
4			

Prof. Rita de la Torre	<code>=CONCATENAR (C2;" ";A2;" ";B2)</code>	<code>=C2 & " " & A2 & " " & B2</code>
Señor Paco Pérez	<code>=CONCATENAR (C3;" ";A3;" ";B3)</code>	

En informática en general, los textos literales se escriben entre comillas.
En el ejemplo, el espacio entre comillas permite separar trato, nombre y apellidos.

Asistente de texto en columnas

Excel permite tomar el texto de una o más celdas y propagarlo en varias de ellas. Para ello usaremos la herramienta **Texto en Columnas** en la ficha **Datos** de la Cinta.



- Selecciona la celda o columna que contenga el texto a dividir en columnas. Asimismo, verifica que existen suficientes columnas vacías a la derecha de la celda o columna seleccionada
- Haz clic sobre el botón **Texto en columnas** en la ficha **Datos**.

Pantalla 1 Indica si los datos están **Delimitados** o tienen un **ancho fijo**

Pantalla 2 Indica el carácter empleado como separador de campos si has seleccionado **Delimitados** en la pantalla 1, o establece el ancho de los campos (saltos de columna).

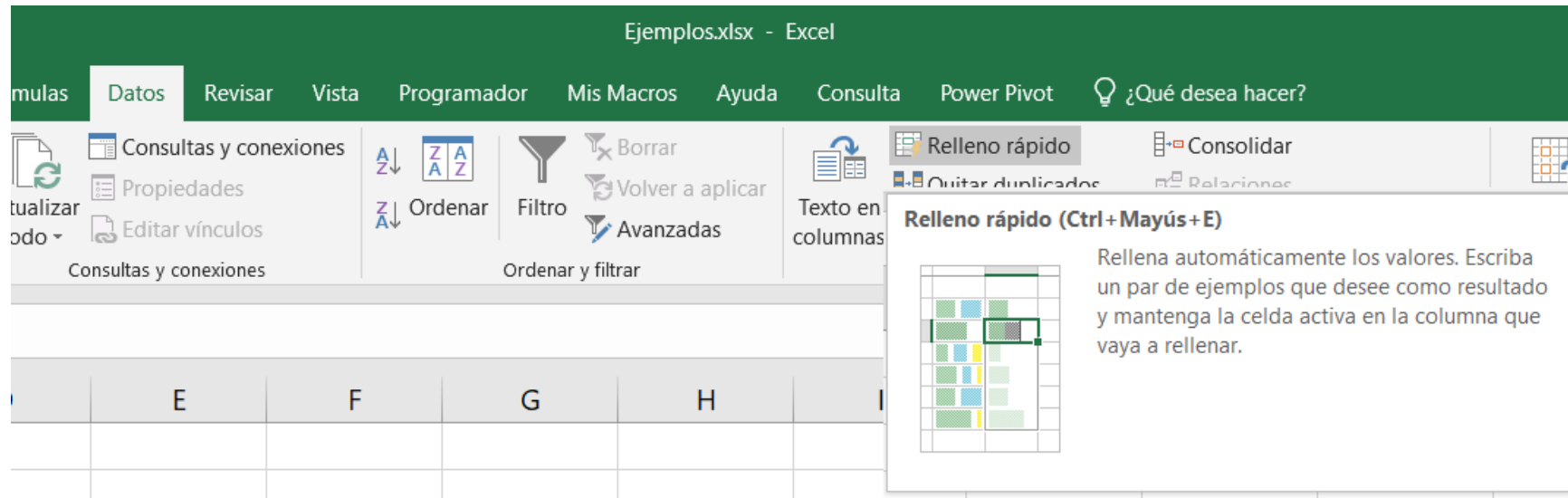
Pantalla 3 Selecciona las columnas que desees importar y establece el tipo de datos y/o formato de ser necesario. Adicionalmente, si pulsas el botón **Avanzadas...** podrás indicar cuál es el separador decimal y el separador de miles utilizados en los datos que estás importando.

Funciones para extraer partes de una fecha (o calcular fechas posteriores)

	A	B	C	D	E
3	Fecha:	22/04/2021			
4					
5	Función	Se usa para...	Sintaxis	Ejemplo	Resultado
6	DIA	Extraer el día de una fecha	=DIA (ref)	=DIA(B3)	22
7	MES	Extraer el mes de una fecha	=MES (ref)	=MES(B3)	4
8	AÑO	Extraer el año de una fecha	=AÑO (ref)	=AÑO(B3)	2021
9	DIASEM	Extraer el día de la semana de una fecha	=DIASEM (ref)	=DIASEM(B3)	5
10	NUM.DE.SEMANA	Extraer el número de la semana de una fecha	=NUM.DE.SEMANA (ref; [tipo])	=NUM.DE.SEMANA(B3)	17
11	FIN.MES	Indicar el último día del mes de una fecha, sin el argumento <i>meses</i> o de la fecha restando o sumando <i>meses</i> a la misma	=FIN.MES (ref; [meses])	=FIN.MES(B3)	30/04/2021
12	FECHA.MES	Suma o Resta <i>meses</i> a una fecha dada.	=FECHA.MES (ref; meses)	=FECHA.MES(B3;1)	22/05/2021

Funciones para extraer partes de un texto

	A	B	C	D	E
3	Texto	Barcelona·Ciutat·Condal			
4					
5	Función	Se usa para...	Sintaxis	Ejemplo	Resultado
6	IZQUIERDA	Extraer x caracteres a la izquierda de un texto	=IZQUIERDA (ref; caracteres)	=IZQUIERDA(B3;9)	Barcelona
7	EXTRAE	Extraer x caracteres a partir de una posición en un texto	=EXTRAE (ref; posición; caracteres)	=EXTRAE(B3;11;6)	Ciutat
8	DERECHA	Extraer x caracteres a la derecha de un texto	=DERECHA (ref; caracteres)	=DERECHA(B3;6)	Condal
9	LIMPIAR	Quita todos los caracteres no imprimibles del texto	=LIMPIAR (ref)	=LIMPIAR(B3)	Barcelona Ciutat Condal
10					



versión Excel 2013 y posteriores.

El botón **Relleno rápido** rellena una columna con datos, automáticamente, cuando detecta un patrón.

Esto permite, por ejemplo, usarlo para separar los nombres y apellidos en una sola columna, o viceversa, combinar datos de dos columnas diferentes.

Convertir números almacenados como textos a números

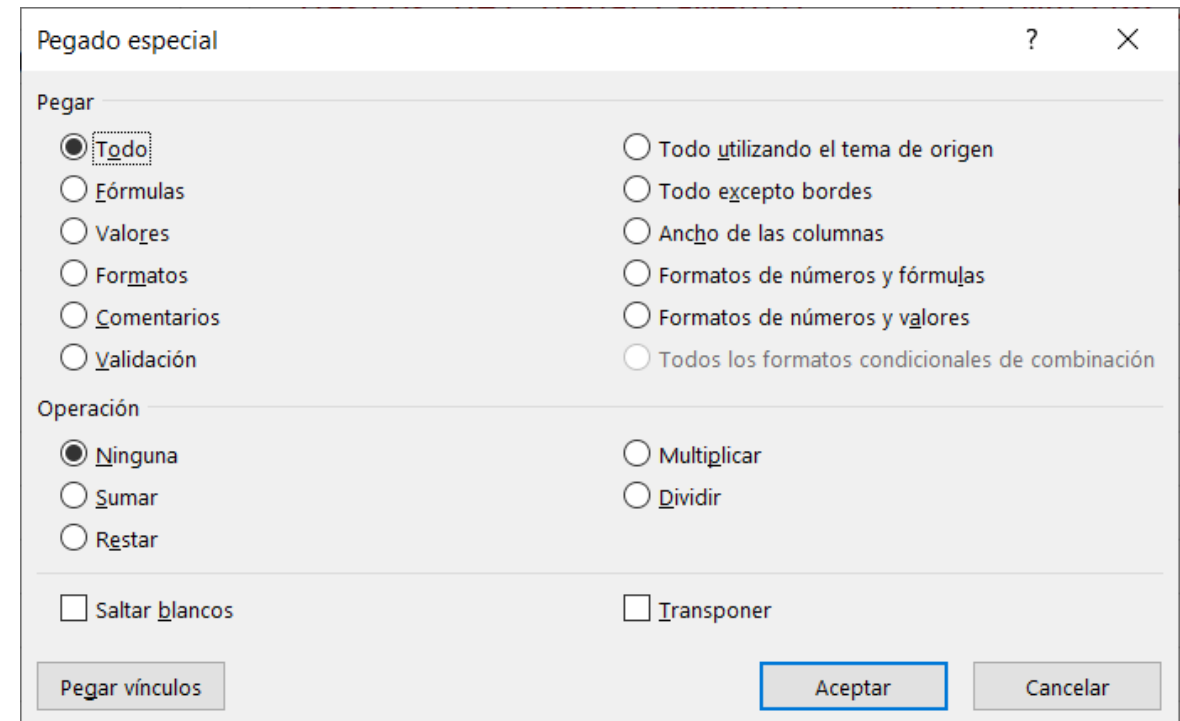
123

123

'123

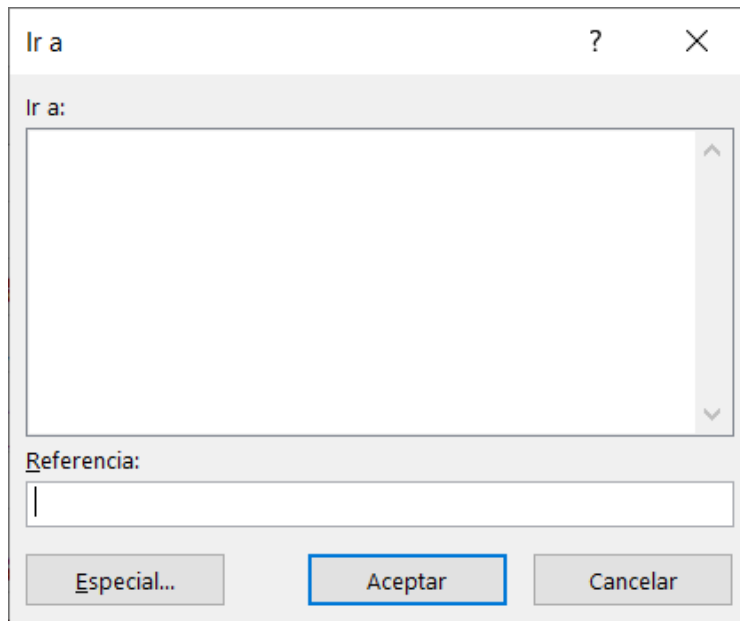
1. Escribir el número 1 en una celda en blanco.
2. Clic sobre el icono **Copiar**
3. Seleccionar los números almacenados como textos
4. Clic con el botón derecho y seleccionar **Pegado especial**.
5. Del cuadro de diálogo que aparece, seleccionar **Multiplicar**

Los números se pegarán como tales.

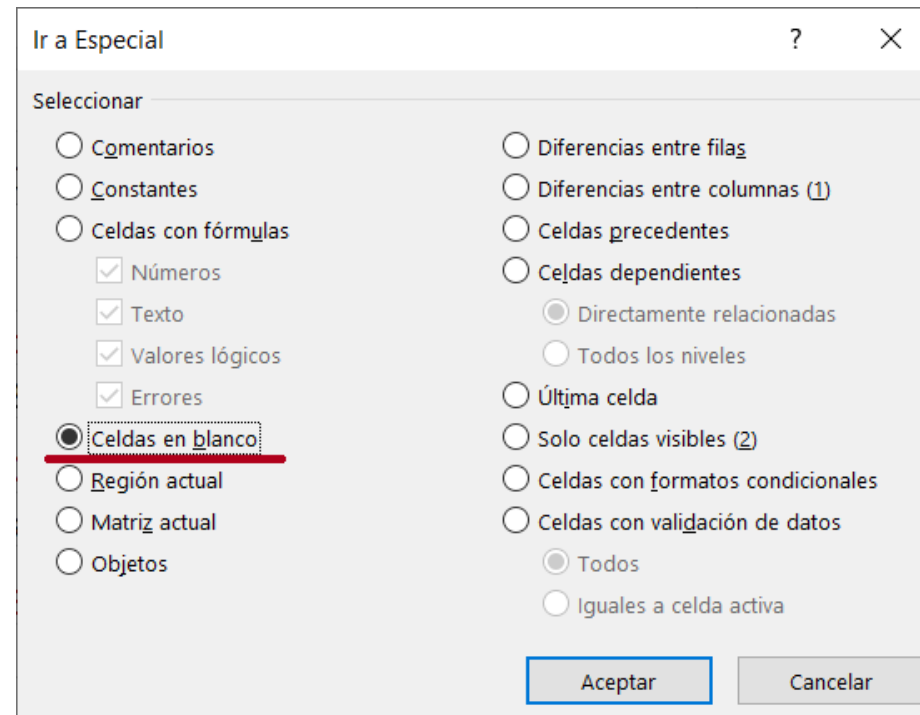


Seleccionar y rellenar todas las celdas en blanco

1. Presionar la tecla **F5**, y del cuadro de diálogo **Ir a...** seleccionar **Especial**.

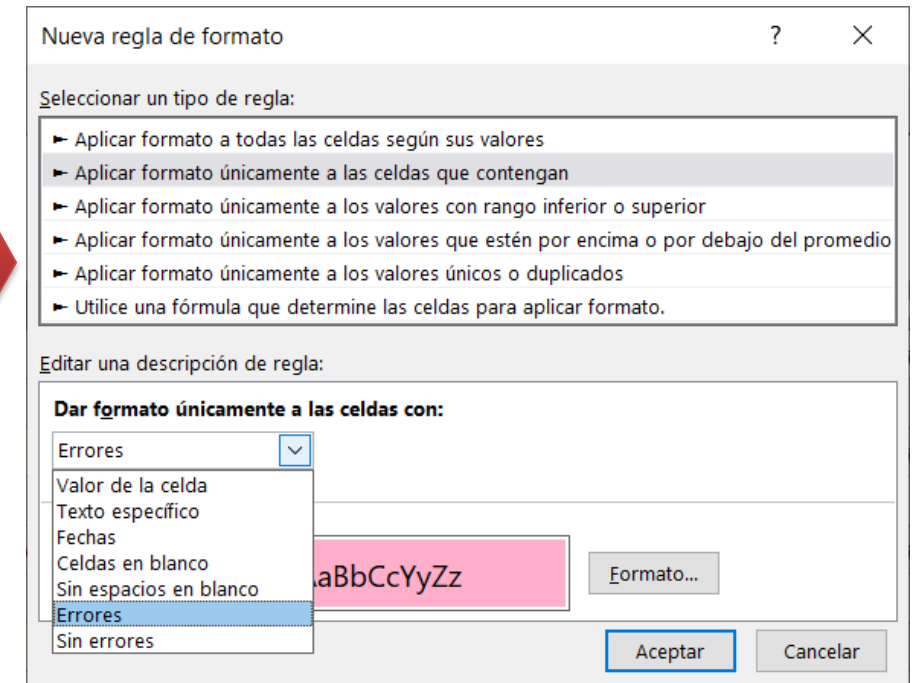
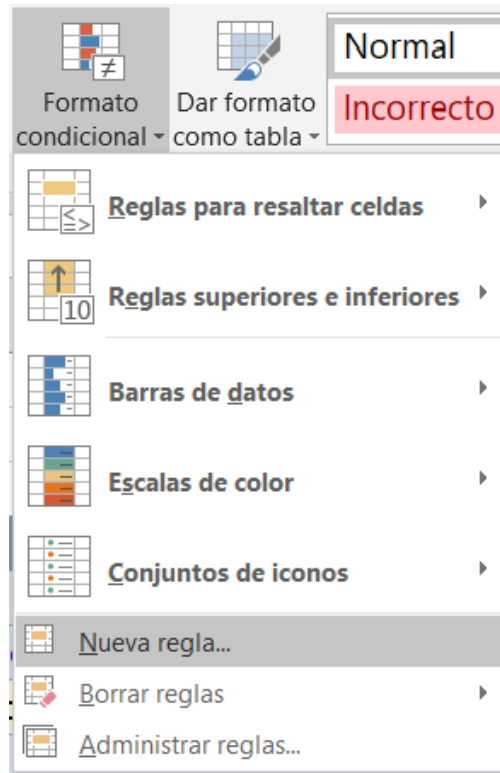


2. Seleccionar **Celdas en blanco**.



3. Escribir el texto o fórmula deseados y presionar simultáneamente las teclas **Ctrl Enter**.

Resaltar y corregir errores



Función SI.ERROR

Sintaxis:

`=SI.ERROR(valor;valor_si_error)`

Función TEXTO para formatear números complejos

(Ver documento ♦[Formatos personalizados en Excel](#)♦ para para mayor detalle)

Permite cambiar la apariencia de un número aplicándole formato con códigos de formato.

Es útil cuando se desea concatenar números con texto o símbolos en una misma celda.

Si el formato deseado es de tipo decimal o moneda, se pueden utilizar las funciones **DECIMAL** o **MONEDA** respectivamente. Esta última tomará la configuración de Windows.

Ejemplos:

	A	B	C	D
1				
2	Año:	2020		
3				= "Gastos de la empresa para el año " & B2 & ": " & MONEDA(B4)
4	Gastos de la Empresa:	120.000,00 €		Gastos de la empresa para el año 2020: 120.000,00 €
5				= "Gastos del departamento: " & DECIMAL(B6;2) & " euros"
6	Gastos del Departamento:	35.000,00 €		Gastos del departamento: 35.000,00 euros
7				= "El departamento ha consumido un " & TEXTO(B8;"00,00%") & " del presupuesto"
8	Porcentaje:	29,17%		El departamento ha consumido un 29,17% del presupuesto
9				

Eliminar espacios extras

```
welcome to digital vidya  
welcome  to digital vidya  
welcome   to digital vidya  
welcome           to digital vidya
```

Función:

=ESPACIOS (referencia)

Cambiar texto a

MAYÚSCULAS,
minúsculas,
Nombre Propio

Funciones:

=MAYUSC (referencia)

=MINUSC (referencia)

=NOMPROPIO (referencia)

Respectivamente.

Tomar en cuenta que una vez realizado los cambios con estas funciones, habrá que copiar y pegar valores en la columna que contiene la fórmula y eliminar la columna que está “sucia”.

Eliminar Duplicados

The screenshot shows the Microsoft Excel interface with the 'Datos' (Data) ribbon selected. The 'Quitar duplicados' (Remove Duplicates) button is highlighted in the ribbon. Below the ribbon, the 'Quitar duplicados' dialog box is open, displaying the following options:

- ☒ Seleccionar todo
- ☐ Anular selección
- ☒ Mis datos tienen encabezados

The 'Columnas' (Columns) list shows the following columns selected for removal of duplicates:

- ☒ Semana
- ☒ Fecha
- ☒ Día
- ☒ Tipo Sesión
- ☒ De

A handwritten note in the dialog box reads: *Selecciona las columnas para las cuales deseas eliminar los duplicados*.

Buttons at the bottom of the dialog box: Aceptar, Cancelar.

On the right side of the image, there is a text box with the following content:

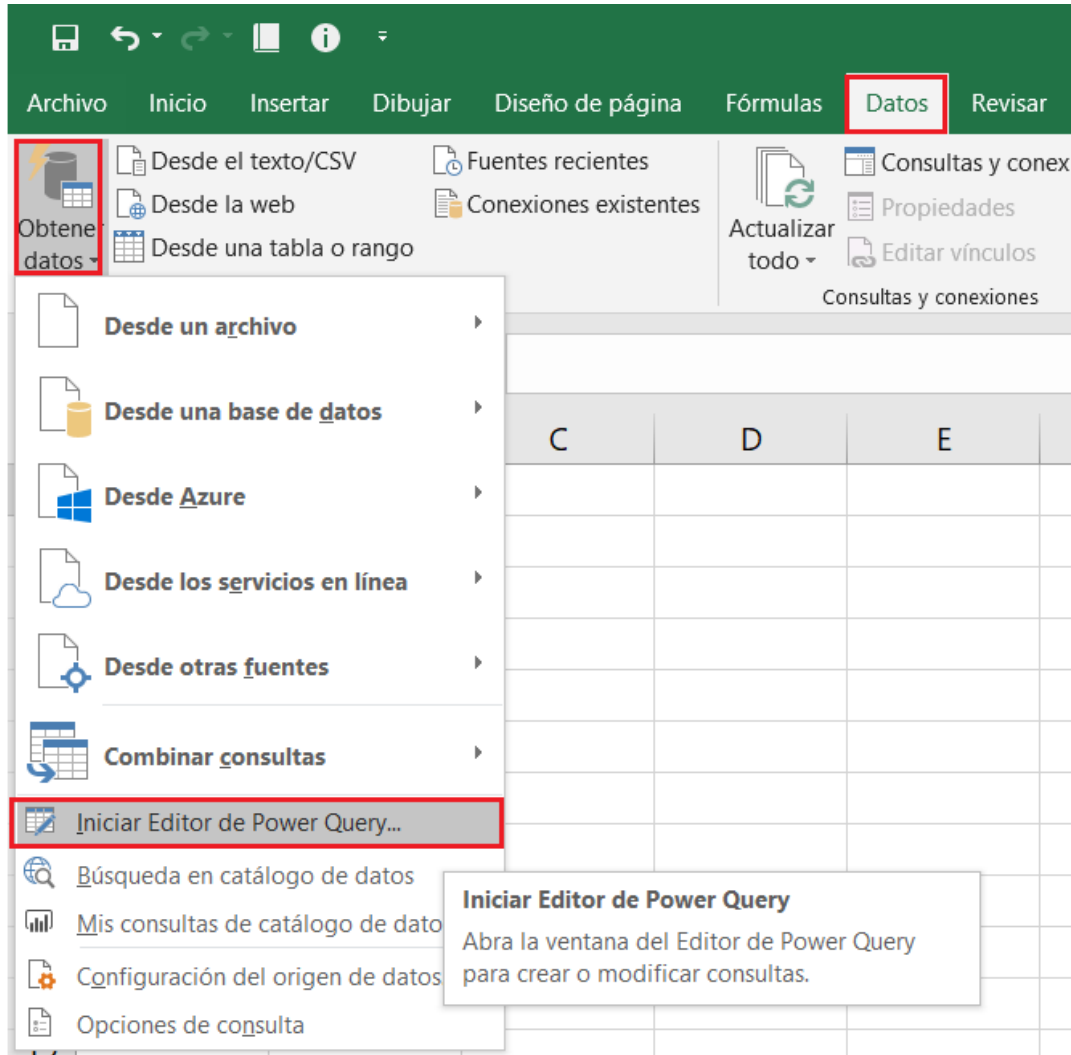
Quitar duplicados
Elimina las filas duplicadas de una hoja.
Puede especificar las columnas en las que se buscará información duplicada.



Power Query

Limpiar los Datos con Power Query

<https://www.howtoexcel.org/power-query/the-complete-guide-to-power-query/>



Power Query

es un motor de transformación y preparación de datos que viene en forma de complemento para Microsoft Excel y Power BI.

Power Query permite extraer información de distintas fuentes de datos, limpiarlos y transformarlos según necesites, para posteriormente cargarlos en algún sitio para su posterior uso; ya sea en una tabla de Excel o en un modelo de datos.

En Excel, podemos encontrarlo en la pestaña **Datos** de la cinta, en el primer icono (**Obtener Datos**)

de la sección **Obtener y transformar datos**: **Iniciar Editor de Power Query** tal y como se muestra en la imagen.



POWER QUERY

- Herramienta ETL
- Lenguaje M
- Obtener Datos
- Transformar Datos
- Combinar Tablas
- Fusionar Tablas
- Enriquecer Datos

Transformación de Datos con Power Query		
OPERACIONES CON COLUMNAS	GENERACIÓN DE COLUMNAS	LENGUAJE M
Filtrado Ordenación Cambiar valores Pivotar Despivotar Dividir Combinar Expandir	Duplicación Cálculos Básicos Cálculos Avanzados	



Power Query

- Descarga la carpeta **Documentos Excel**, que se encuentra dentro de la carpeta **Open Data**.
- En la subcarpeta **Documentos Excel** hallarás seis archivos **201801_ventas.xlsx... 201806_ventas.xlsx**.
Todos estos archivos guardan las ventas hipotéticas del mes de una empresa y presentan la misma estructura de datos.
- Combinaremos todos los archivos de Excel haciendo uso del Power Query.

1

Obtener datos ▾

- Desde un archivo
- Desde una base de datos
- Desde Azure
- Desde los servicios en línea
- Desde otras fuentes

2

Carpeta

Ruta a la carpeta

D:\@0Material\PowerBI\PowerBI_UPF\Documentos Excel Examinar...

3

Combinar ▾ Cargar ▾ Editar Cancelar

- Combinar y editar
- Combinar y cargar
- Combinar y cargar en...

4

Combinar archivos

Seleccione el objeto que quiera extraer de cada archivo. Más información

Archivo de ejemplo: Primer archivo ▾

Opciones de presentación ▾

- Parámetro del archivo de ejemplo2 [2]
- Tabla1
- ventas

Tabla1

Producto	Barcelona	Girona
0001 Armani (3,3 oz.) Agua de Colonia	17292	923
0002 Emporio (3,3 oz.) Perfume	5053	1653
0003 Gio (3,3 oz.) Perfume	17824	1483
0004 Mania (3,3 oz.) Perfume	19111	2163
0005 Ocean Dream (3 oz.) Agua de Colonia	17931	783
0006 Azzaro 9 (1 oz.) Extracto	20330	1013
0007 Chrome (3,3 oz.) Agua de Colonia	9449	1763
0008 Pure Lavender (5 oz.) Agua de Colonia	20988	1493
0009 Le Dix (1,65 oz.) Agua de Colonia	15284	1613
0010 Prelude (1 oz.) Agua de Colonia	15078	533
0011 Rumba (3,3 oz.) Agua de Colonia	14531	903
0012 Talisman (4,2 oz.) Perfume	8948	1313
0013 Colors (3,3 oz.) Colonia	8457	1523
0014 Funtastic (3,3 oz.) Agua de Colonia	9496	563
0015 Paradiso Inferno for man (3,3 oz.) Agua de Colonia	13458	1843
0016 Paradiso Inferno for woman (3,3 oz.) Agua de Colonia	16399	1013
0017 Tribu (3,3 oz.) Agua de Colonia	9734	1043
0018 Amazing (3,3 oz.) Agua de Colonia	20478	923
0019 Hot (1,65 oz.) Agua de Colonia	20823	1263
0020 Nude (1 oz.) Agua de Colonia	8749	623

☐ Omitir archivos con errores

Aceptar Cancelar

Selecciona la carpeta **Documentos Excel** que te has descargado anteriormente.



Importar datos ? X

Seleccione cómo desea ver estos datos en el libro.

☒ Tabla

☐ Informe de tabla dinámica

☐ Gráfico dinámico

☐ Crear solo conexión

¿Dónde desea situar los datos?

☐ Hoja de cálculo existente:

☒ Hoja de cálculo nueva

☐ Agregar estos datos al Modelo de datos

Propiedades... Aceptar Cancelar

Obtener datos

- Desde el texto/CSV
- Desde la web
- Desde una tabla o rango

Fuentes recientes

Conexiones existentes

Consultas y conexiones

Actualizar todo

Propiedades

Editar vínculos

Ordenar

Filtro

Borrar

Volver a aplicar

Avanzadas

Ordenar y filtrar

Texto en columnas

Relleno rápido

Quitar duplicados

Validación de datos

201801_ventas.xlsx

	B	C	D	E	F
		Barcelona	Girona	Tarragona	Lleida
ani (3,3 oz.) Agua de Colonia		17292	9234	12541	5271
orio (3,3 oz.) Perfume		5053	16533	17226	20551
(3,3 oz.) Perfume		17824	14870	9133	14413
ia (3,3 oz.) Perfume		19111	21684	11227	11005
an Dream (3 oz.) Agua de Colonia		17931	7897	16726	4151
ro 9 (1 oz.) Extracto		20330	10164	16130	19762
me (3,3 oz.) Agua de Colonia		9449	17680	6082	17158
oz.) Agua de Colonia		20988	14936	18633	15122
Agua de Colonia		15284	16134	10762	11838
de (1 oz.) Agua de Colonia		15078	5354	17744	11964
me (3,3 oz.) Agua de Colonia		14531	9063	4833	20551

Desde un archivo

Desde una base de datos

Desde Azure

Desde los servicios en línea

Desde otras fuentes

Combinar consultas

Iniciar Editor de Power Query...

Iniciar Editor de Power Query

Abra la ventana del Editor de Power Query para crear o modificar consultas.

Podrás apreciar cómo Power Query ha combinado los datos de todas las tablas en una sola...

Documentos Excel - Editor de Power Query

Archivo Inicio Transformar Agregar columna Vista

Cerrar y cargar Actualizar vista previa Consultas Administrar Elegir columnas Quitar columnas Conservar filas Quitar filas Ordenar Dividir columna Agrupar por Tipo de datos: Texto Usar la primera fila como encabezado Reemplazar los valores Combinar consultas Anexar consultas Combinar archivos Administrar parámetros Configuración de origen de datos Nuevo origen Orígenes recientes

Consultas [5]

- Transformar archivo de Documentos Excel [3]
- Consulta de ejemplo [2]
 - Parámetro del archivo de ejemplo1 (Archivo de ej...
 - Archivo de ejemplo
 - Transformar archivo de ejemplo de Documentos E...
 - Transformar archivo de Documentos Excel
- Otras consultas [1]
 - Documentos Excel

	Source.Name	Producto	Barcelona	Girona	Tar
1	201801_ventas.xlsx	0001 Armani (3,3 oz.) Agua de Colonia	17292	9234	
2	201801_ventas.xlsx	0002 Emporio (3,3 oz.) Perfume	5053	16533	
3	201801_ventas.xlsx	0003 Gio (3,3 oz.) Perfume	17824	14870	
4	201801_ventas.xlsx	0004 Mania (3,3 oz.) Perfume	19111	21684	
5	201801_ventas.xlsx	0005 Ocean Dream (3 oz.) Agua de Colonia	17931	7897	
6	201801_ventas.xlsx	0006 Azzaro 9 (1 oz.) Extracto	20330	10164	
7	201801_ventas.xlsx	0007 Chrome (3,3 oz.) Agua de Colonia	9449	17680	
8	201801_ventas.xlsx	0008 Pure Lavender (5 oz.) Agua de Colonia	20988	14936	
9	201801_ventas.xlsx	0009 Le Dix (1,65 oz.) Agua de Colonia	15284	16134	
10	201801_ventas.xlsx	0010 Prelude (1 oz.) Agua de Colonia	15078	5354	
11	201801_ventas.xlsx	0011 Rumba (3,3 oz.) Agua de Colonia	14531	9063	
12	201801_ventas.xlsx	0012 Talisman (4,2 oz.) Perfume	8948	13173	
13	201801_ventas.xlsx	0013 Colors (3,3 oz.) Colonia	8457	15215	
14	201801_ventas.xlsx	0014 Funtastic (3,3 oz.) Agua de Colonia	9496	5619	
15	201801_ventas.xlsx	0015 Paradiso Inferno for man (3,3 oz.) Agua de Colonia	13458	18494	
16	201801_ventas.xlsx	0016 Paradiso Inferno for woman (3,3 oz.) Agua de Colonia	16399	10105	
17	201801_ventas.xlsx	0017 Tribu (3,3 oz.) Agua de Colonia	9734	10439	
18	201801_ventas.xlsx	0018 Amazing (3,3 oz.) Agua de Colonia	20478	9274	
19	201801_ventas.xlsx	0019 Hot (1,65 oz.) Agua de Colonia	20823	12626	
20	201801_ventas.xlsx	0020 Nude (1 oz.) Agua de Colonia	8749	6252	
21	201801_ventas.xlsx	0021 Initial (3,3 oz.) Perfume	16463	16700	
22	201801_ventas.xlsx	0022 Injur (4 oz.) Perfume	12943	16211	

6 COLUMNAS, 834 FILAS

VISTA PREVIA DESCARGADA A LAS 18:08

Configuración de la co... x

PROPIEDADES

Nombre

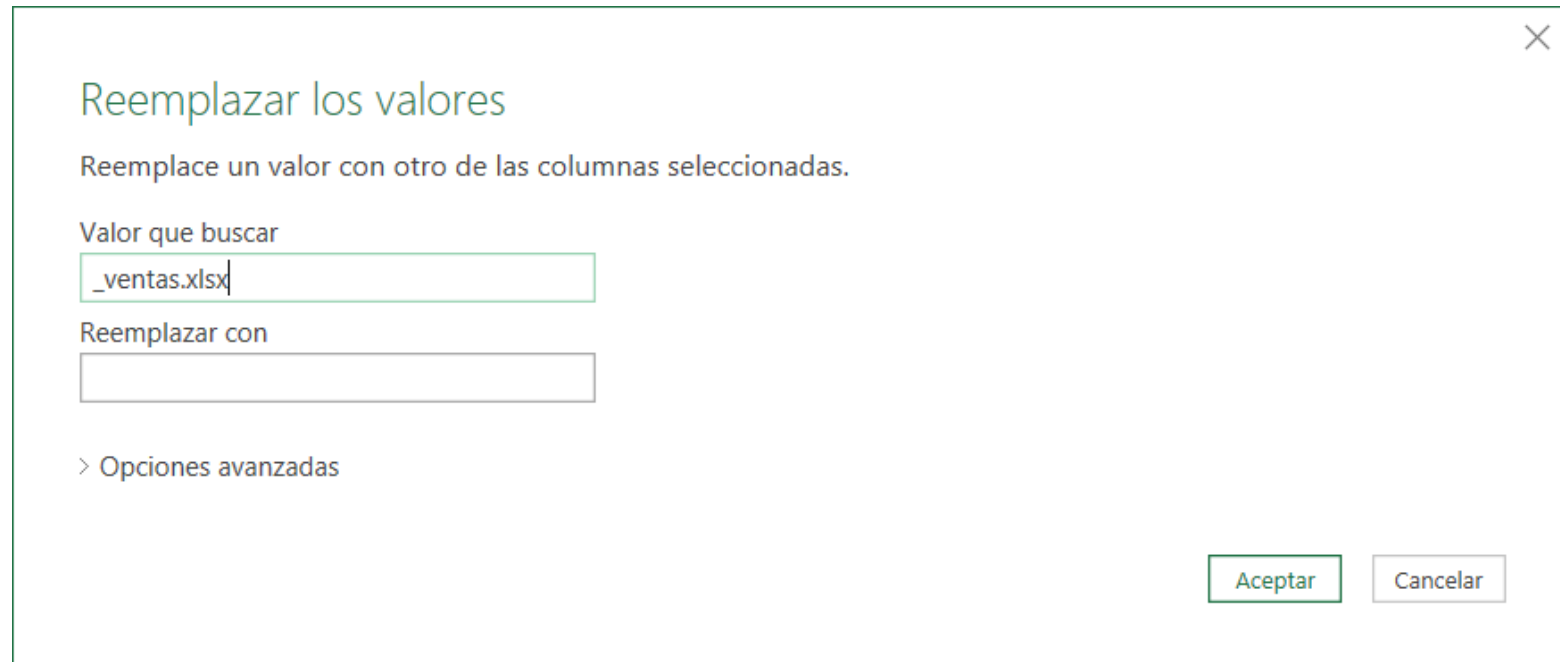
Documentos Excel

Todas las propiedades

PASOS APLICADOS

- Origen
- Archivos ocultos filtrados1
- Invocar función personalizada1
- Columnas con nombre cambi...
- Otras columnas quitadas1
- Columna de tabla expandida1
- X Tipo cambiado

- Selecciona la primera columna **Source.Name** y haz clic sobre la misma con el botón derecho.
- Selecciona ahora la opción **Reemplazar los Valores**.
- Indica que quieres reemplazar el valor **_ventas.xlsx** por un valor vacío, como se muestra en pantalla.
- Al presionar **Aceptar**, fíjate cómo a la derecha de la pantalla, en la lista **Configuración de la Consulta - Pasos Aplicados**, Power Query va apuntando las transformaciones realizadas.



Reemplazar los valores

Reemplace un valor con otro de las columnas seleccionadas.

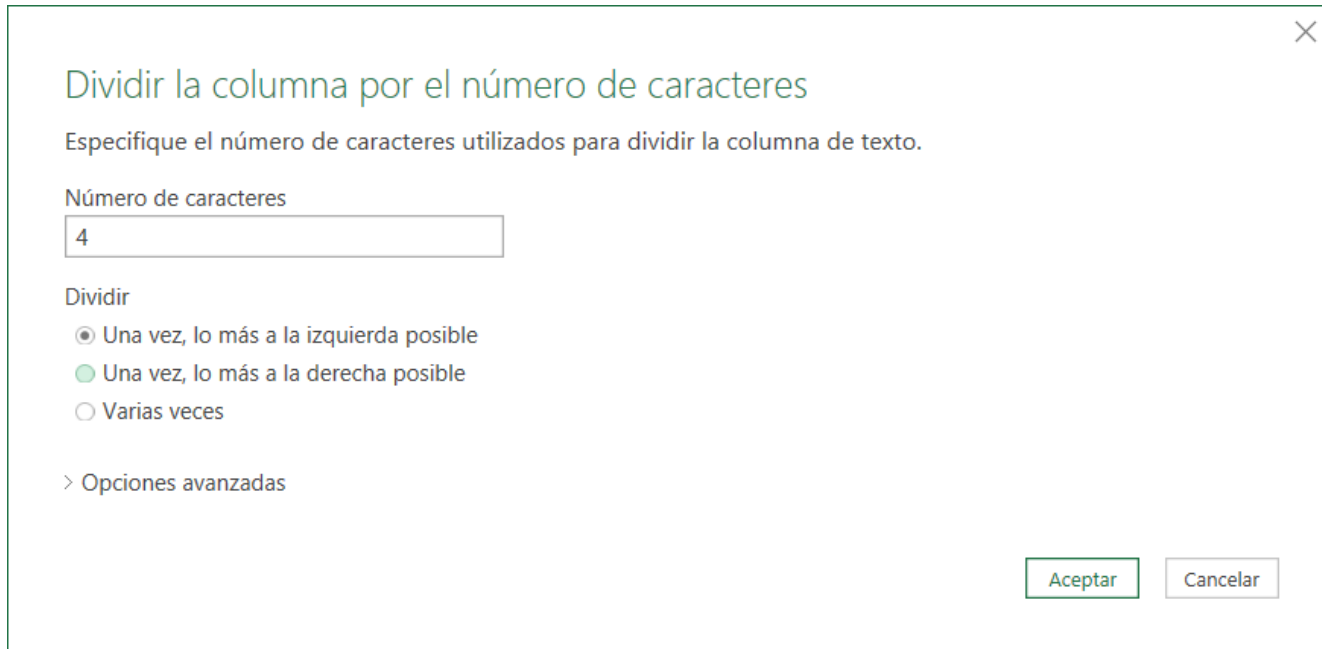
Valor que buscar

Reemplazar con

> Opciones avanzadas

Aceptar Cancelar

- Ahora dividiremos la columna resultante en año y mes.
Para ello selecciona del menú emergente (clic derecho en la columna) la opción **Dividir columna -> Por número de caracteres**
- Indica **4** en el número de caracteres y haz clic sobre OK



Dividir la columna por el número de caracteres

Especifique el número de caracteres utilizados para dividir la columna de texto.

Número de caracteres

4

Dividir

☒ Una vez, lo más a la izquierda posible

☐ Una vez, lo más a la derecha posible

☐ Varias veces

> Opciones avanzadas

Aceptar Cancelar

- Cambia el nombre a las dos primeras columnas por **Año** y **Mes** respectivamente haciendo doble clic sobre el encabezado de las mismas.
- Repite el proceso con la columna **Producto**. Indica **Id** y **Producto** como nombre de las columnas resultantes

- Ahora selecciona las cuatro primeras columnas de la tabla, haz clic con el botón derecho y selecciona la opción **Anulación de dinamización de otras columnas**.
- Cambia el nombre de las columnas **Atributo** por **Provincia** y **Valor** por **ventas**.
- Cierra y aplica los cambios



Cerrar y
cargar ▼

	1 ² ₃ Año	1 ² ₃ Mes	1 ² ₃ Id	A ^B _C Producto	A ^B _C Provincia	1.2 Ventas
1	2018	1	1	Armani (3,3 oz.) Agua de Colonia	Barcelona	17292
2	2018	1	1	Armani (3,3 oz.) Agua de Colonia	Girona	9234
3	2018	1	1	Armani (3,3 oz.) Agua de Colonia	Tarragona	12541
4	2018	1	1	Armani (3,3 oz.) Agua de Colonia	Lleida	5271
5	2018	1	2	Emporio (3,3 oz.) Perfume	Barcelona	5053
6	2018	1	2	Emporio (3,3 oz.) Perfume	Girona	16533
7	2018	1	2	Emporio (3,3 oz.) Perfume	Tarragona	17226
8	2018	1	2	Emporio (3,3 oz.) Perfume	Lleida	20551
9	2018	1	3	Gio (3,3 oz.) Perfume	Barcelona	17824
10	2018	1	3	Gio (3,3 oz.) Perfume	Girona	14870
11	2018	1	3	Gio (3,3 oz.) Perfume	Tarragona	9133
12	2018	1	3	Gio (3,3 oz.) Perfume	Lleida	14413

- ✓ <http://trucosycursos.es/todo-los-metodos-para-extraer-caracteres-de-una-celda-excel/>
- ✓ <https://www.exceltip.com/es/como-dividir-numeros-y-texto-de-cadena-en-excel.html>
- ✓ <https://opengovdata.org/>
- ✓ <https://ciberimaginario.es/2018/07/25/sobre-datos-abiertos/>
- ✓ <https://datos.gob.es/es/catalogo>
- ✓ <https://data.europa.eu/es>
- ✓ https://www.ospi.es/export/sites/ospi/documents/documentos/cotec_informe_reutilizacion-de-datos.pdf

