

Tema 2

Edición de datos

INTRODUCCIÓN A LA EDICIÓN Y
VISUALIZACIÓN DE DATOS INMEDIATOS



Proyecto
OpenCourseWare-UNIA
(ocw.unia.es)



Autora
Encarni Hinojosa



ÍNDICE

Introducción

Objetivos

Contenidos

2.1

- Programa: **Tabula PDF**
- Subida del archivo PDF y selección de la tabla que nos interesa
- Opciones de formato, dentro de Tabula PDF
- Extracción de la tabla resultante

2.2

- Programa: **Open Refine**
- Subida del archivo CSV y corregir la codificación del texto
- Crear el proyecto y refinar la tabla
- Exportar la tabla ya refinada

2.3

- Programa: **OpenOffice Calc**
- Copiar y pegar datos de la tabla de una web, de un PDF o de un texto
- Transformación y creación de datos (formatear celdas)
- Transformación y creación de datos (uso de fórmulas)
- Preguntas a los datos (acceso directo a preguntas generales)
- Preguntas a los datos (ordenar)
- Preguntas a los datos (filtros)

2.4

- Programa: **OpenOffice Calc**
- Abrir y editar el formato de archivos .csv
- Actualización automática de datos (vincular con datos externos)
- Actualización automática de datos (hoja desde archivo),
abrir archivos ASCII (datos electorales)
- Vincular una tabla editada a otra tabla sin editar, pero actualizable



INTRODUCCIÓN

Microsoft creó la primera versión de Excel (programa de hojas de cálculo) en 1985 y, sorprendentemente, para Macintosh. Dos años después, lo lanzó para Windows. Pronto se convirtió en el 'software' predilecto para gestionar hojas de cálculo y, hasta hoy, sigue estando en la primera posición.

Sin embargo, en el año 2000, la empresa Sun Microsystems lanzó la primera versión de StarOffice, que en unos años se convertiría en el OpenOffice que conocemos actualmente. Apache OpenOffice es un programa de ofimática libre con código abierto y distribución gratuita. Al igual que Microsoft Office, nos ofrece procesador de textos, hojas de cálculo, bases de datos, edición y diseño de presentaciones... Pero la diferencia es que OpenOffice es gratuito.

OpenOffice Calc es el procesador de hojas de cálculo y nada tiene que envidiarle al famoso Excel, ya que nos permite casi todas las opciones de edición de datos que ofrece el 'software' de Microsoft.

Tabula PDF es una sencilla pero importante herramienta que nos permite extraer de una manera fácil y, casi siempre, efectiva tablas de archivos PDF. Si seleccionamos una tabla de un PDF, copiamos y pegamos en una hoja de cálculo, algunas veces, no podremos pintar directamente la tabla que queremos, ya que el ordenador gestiona esos datos en el 'clipboard' como texto y no como tabla con separadores. Tabula PDF se encarga de asegurar que las tablas de un PDF se copien como tablas y nos respeta los separadores de las mismas. Esta herramienta también es de código abierto y gratuita y está creada por periodistas para periodistas.

Open Refine nació como un 'software' de código abierto en enero de 2010 de la mano de Metaweb. En julio del mismo año, Google adquirió esta empresa y renombró el 'software' a Google Refine. En 2012, el gigante tecnológico anunció que dejaría de desarrollar y ofrecer soporte de Google Refine, por lo que la herramienta se encuentra, desde entonces, en una regresión a sus orígenes 'open source' bajo el nombre de Open Refine. Este programa permite 'refinar' hojas de cálculo, es decir, editar en paquetes los datos de una tabla, agilizando muchísimo el trabajo de edición de datos a la hora de gestionar un gran 'data set'.



OBJETIVOS

Los objetivos de este tema se centrarán en los conocimientos que deberá adquirir el alumno para manejar, de manera básica, los programas tratados:

- Manejo básico de Tabula PDF
- Manejo básico de Open Refine
- Manejo básico de OpenOffice Calc
- Manejo básico de OpenOffice Calc (CSV, ASCII)



Para mostrar el uso de los diferentes programas, usaremos ejemplos básicos explicados paso por paso.

Enlaces de descarga de los programas tratados en este tema:

Tabula PDF: tabula.technology

Open Refine: openrefine.org/download

OpenOffice: openoffice.org/es/descargar

CONTENIDOS 2.1

Tabula PDF

1. Una vez abierto el programa, pulsamos para **buscar** el archivo PDF que queremos editar.

Tabula My Files About Help Source Code

Import one or more PDFs

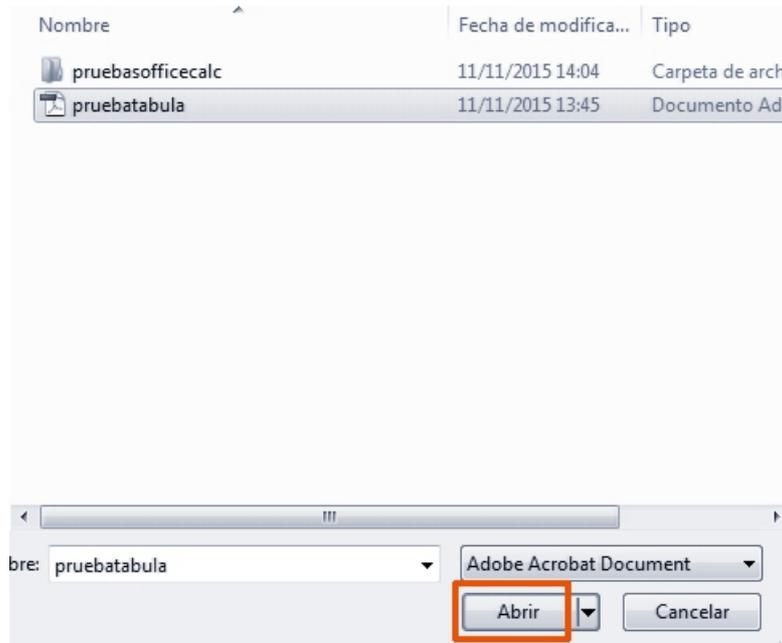
Browse... Import

Imported PDFs

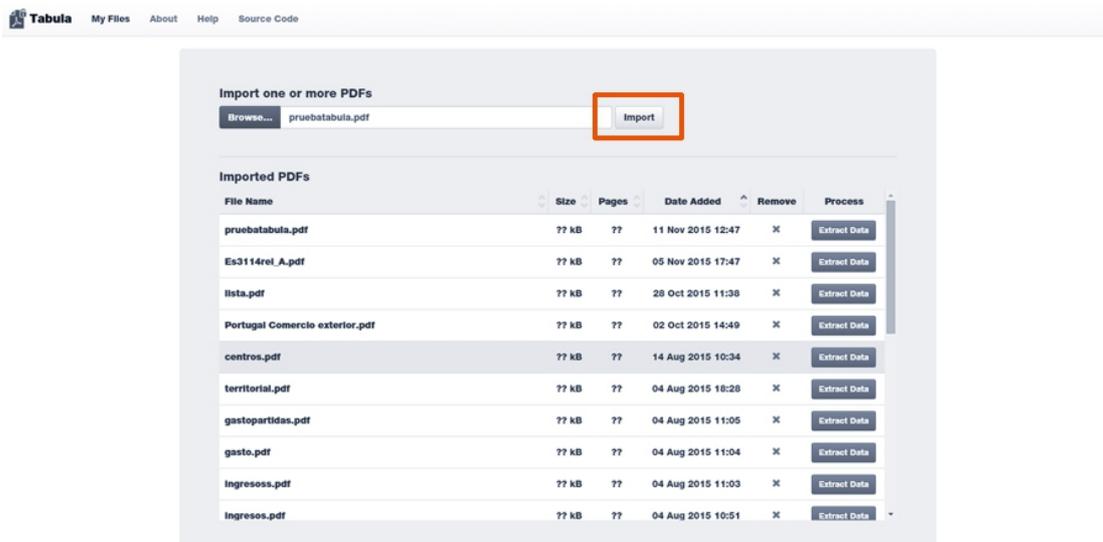
File Name	Size	Pages	Date Added	Remove	Process
pruebatabula.pdf	?? KB	??	11 Nov 2015 12:47	X	Extract Data
Es3114rel_A.pdf	?? KB	??	05 Nov 2015 17:47	X	Extract Data
ltsla.pdf	?? KB	??	20 Oct 2015 11:38	X	Extract Data
Portugal Comercio exterior.pdf	?? KB	??	02 Oct 2015 14:49	X	Extract Data
centros.pdf	?? KB	??	14 Aug 2015 10:34	X	Extract Data
territorial.pdf	?? KB	??	04 Aug 2015 18:28	X	Extract Data
gastopartidas.pdf	?? KB	??	04 Aug 2015 11:05	X	Extract Data
gasto.pdf	?? KB	??	04 Aug 2015 11:04	X	Extract Data
Ingresosa.pdf	?? KB	??	04 Aug 2015 11:03	X	Extract Data
Ingresos.pdf	?? KB	??	04 Aug 2015 10:51	X	Extract Data



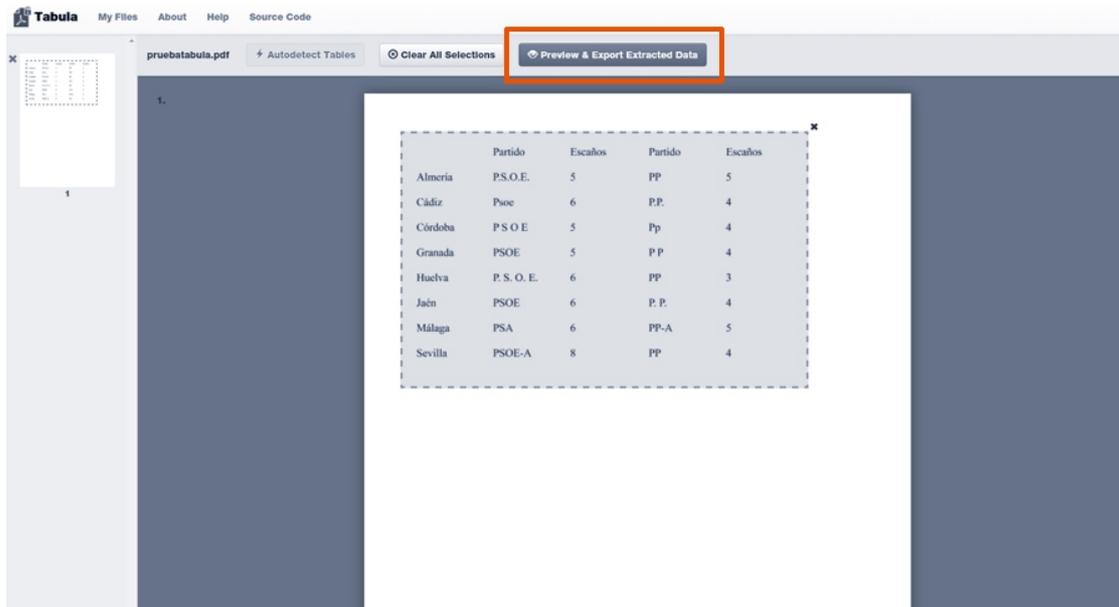
2. Seleccionamos un archivo PDF. Debe ser un archivo **PDF puro**, es decir, con texto real (seleccionable y con formato) y no un archivo PDF falso, simulando ser texto cuando, en realidad, es imagen (como un archivo JPG).



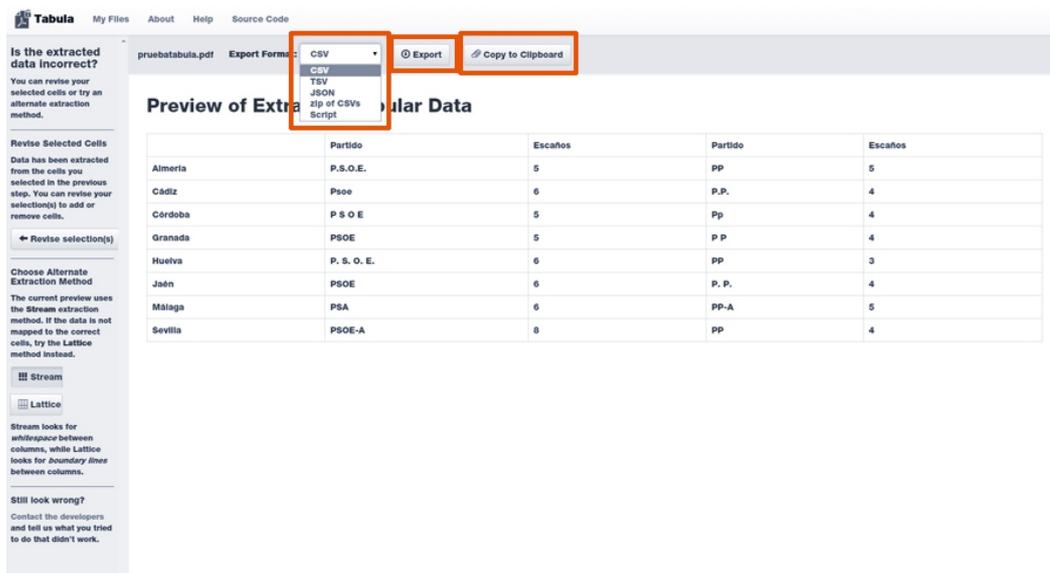
3. Después, pulsamos en **importar**.



4. Pinchamos en una esquina de la tabla que nos interesa extraer y, sin dejar de pulsar, arrastramos. Nos aparecerá un **área de selección**. Después, pulsamos en **Preview & Export...**



5. Seleccionamos el **formato** del archivo que queremos exportar con nuestra tabla. Podemos seleccionar **CSV** (recomendado), TSV, JSON, etc. Después, pulsaremos en **Exportar**. También, podemos copiar la tabla al portapapeles y pegarla en el programa de edición que queramos.





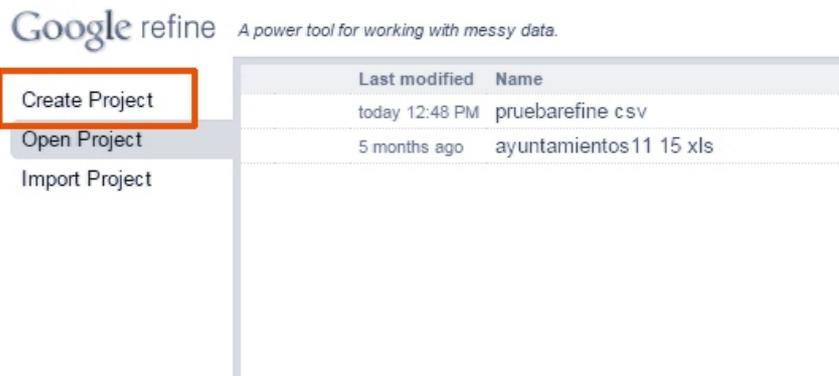
 **sabías que...**

CSV y TSV son los formatos más comunes de datos con **separadores de campo**. Los separadores de campo son elementos tipográficos que separan los datos de manera que, cuando se dispongan en una tabla, puedan estar separados en columnas. En los **CSV** los separadores de campo son **comas**. En los **TSV** los separadores de campo son **tabuladores**. Los más comunes son los CSV.

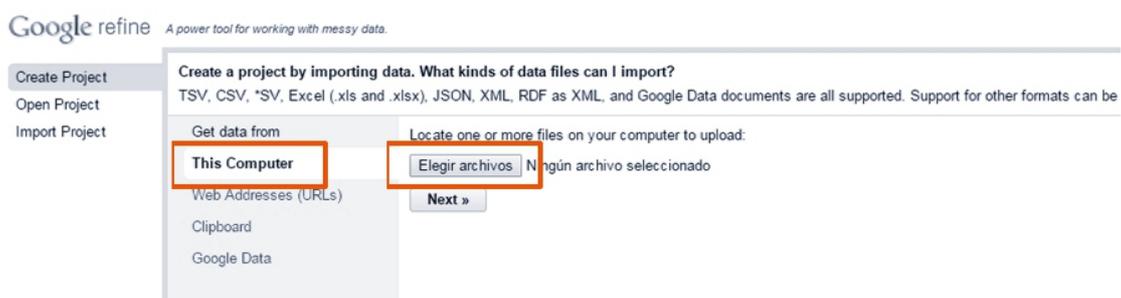
CONTENIDOS 2.2

Open Refine

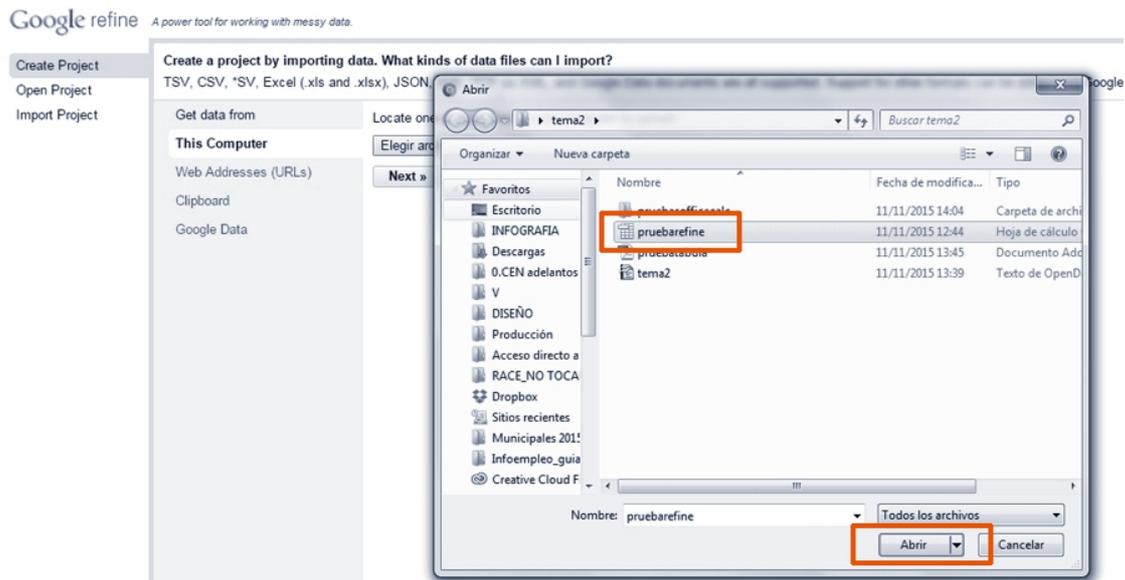
1. Pulsamos en **crear proyecto**.



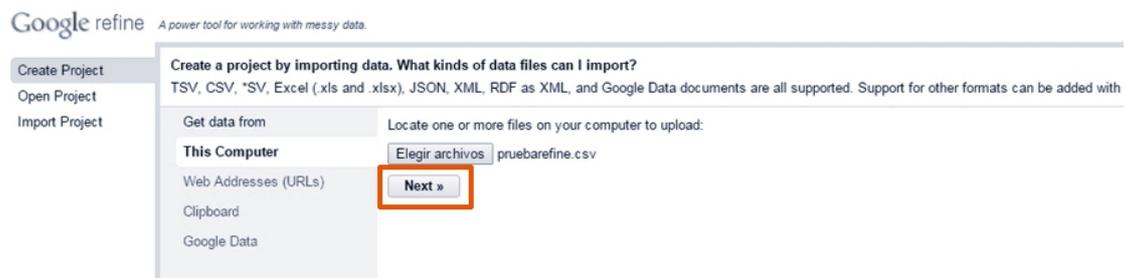
2. Seleccionamos **desde el ordenador** y pulsamos en **Elegir archivos**.



3. Seleccionamos un archivo de datos, por ejemplo, un **CSV**.

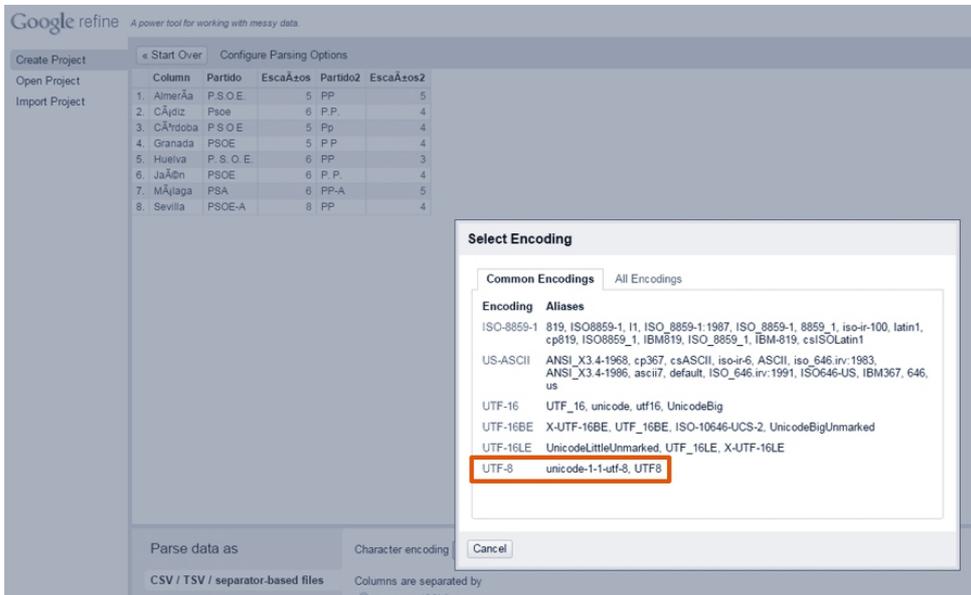


4. Después, pulsamos en **Next**.

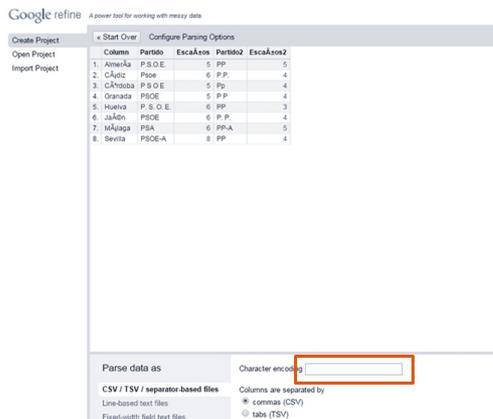




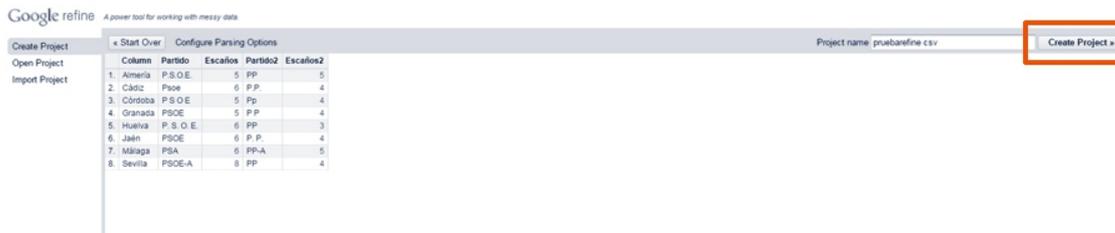
5. En el caso de que la tabla se importe con caracteres extraños (en vez de tildes, por ejemplo), hay que seleccionar **Character encoding**.



6. Probamos con las diferentes codificaciones, hasta dar con la que corrija nuestra tabla. Lo más probable, es que la que mejor nos funcione en Refine sea la **UTF8**.



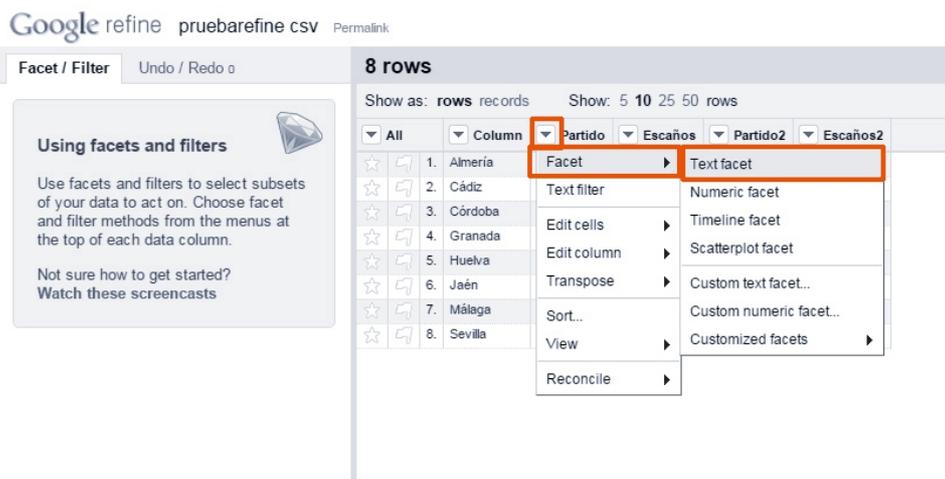
7. Después, pulsamos en **Crear proyecto**.



ej ejemplo

En este ejemplo, necesitamos editar los datos de la columna **Partido** y **Partido2**, para que todos los PSOE y PP se escriban así y no en las diferentes formas que vemos representadas. Esto suele pasar cuando se recopilan resultados electorales, ya que cada circunscripción registra los nombres de los partidos, aunque sean nacionales, y suelen producirse estas diferencias.

8. Para editar los datos de la columna **Partido**, que se refieren al PSOE, pulsamos en el icono de la flecha hacia abajo que está en el encabezamiento de esta columna. Después, de las diversas opciones que nos ofrece Open Refine, en este caso nos interesa **Facet > Text Facet**.





9. Se nos abrirá una ventana con un análisis de los datos de esa columna. Pulsaremos en **Cluster**.

Google refine pruebarefine csv Permalink

Facet / Filter Undo / Redo

Refresh Reset All Remove All

Partido change

7 choices Sort by: name count **Cluster**

P S O E 1
P. S. O. E. 1
P.S.O.E. 1
PSA 1
PSOE 2
Psoe 1
PSOE-A 1
Facet by choice counts

8 rows

Show as: rows records Show: 5 10 25 50 rows

All	Column	Partido	Escaños	Partido2	Escaños2
☆	1.	Almería P.S.O.E.	5	PP	5
☆	2.	Cádiz Psoe	6	P.P.	4
☆	3.	Córdoba P S O E	5	Pp	4
☆	4.	Granada PSOE	5	P P	4
☆	5.	Huelva P. S. O. E.	6	PP	3
☆	6.	Jaén PSOE	6	P. P.	4
☆	7.	Málaga PSA	6	PP-A	5
☆	8.	Sevilla PSOE-A	8	PP	4

10. Se nos abre una nueva ventana. En **Keying Function** seleccionamos **metaphone3** que, en este caso, es la opción que mejor resultado nos dará.

Cluster & Edit column "Partido"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. Find out more ...

Method key collision Keying Function **fingerprint** 2 clusters found

Cluster Size	Row Count	Values in Cluster	Merge
3	4	<ul style="list-style-type: none"> PSOE (2 rows) P.S.O.E. (1 rows) Psoe (1 rows) 	<input type="checkbox"/> fingerprint <input checked="" type="checkbox"/> ngram-fingerprint <input checked="" type="checkbox"/> metaphone3 <input type="checkbox"/> cologne-phonetic
2	2	<ul style="list-style-type: none"> P S O E (1 rows) P. S. O. E. (1 rows) 	<input type="checkbox"/> P S O E

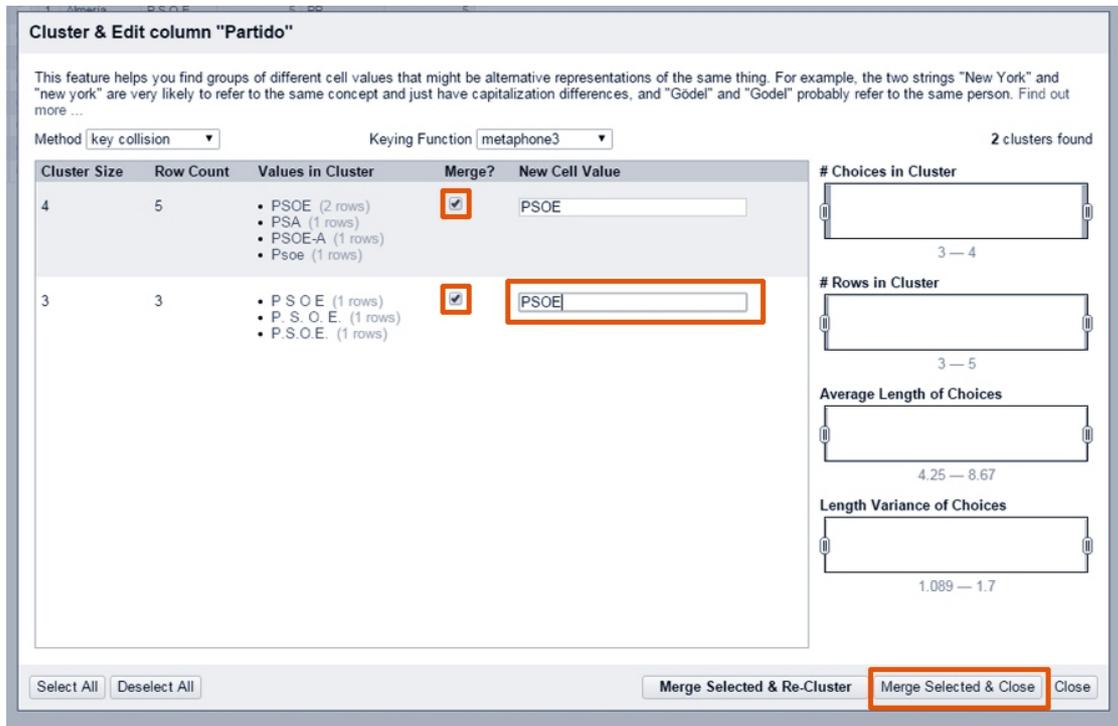
Choices in Cluster: 2 — 3

Rows in Cluster: 2 — 4

Average Length of Choices: 5.33 — 9

Length Variance of Choices

11. Debemos 'chequear' todas las opciones de **Merge?** que tengamos y en **New Cell Value** poner el dato correcto que queremos para todos, en este caso, PSOE. Después, seleccionamos **Merge Selected & Close**.



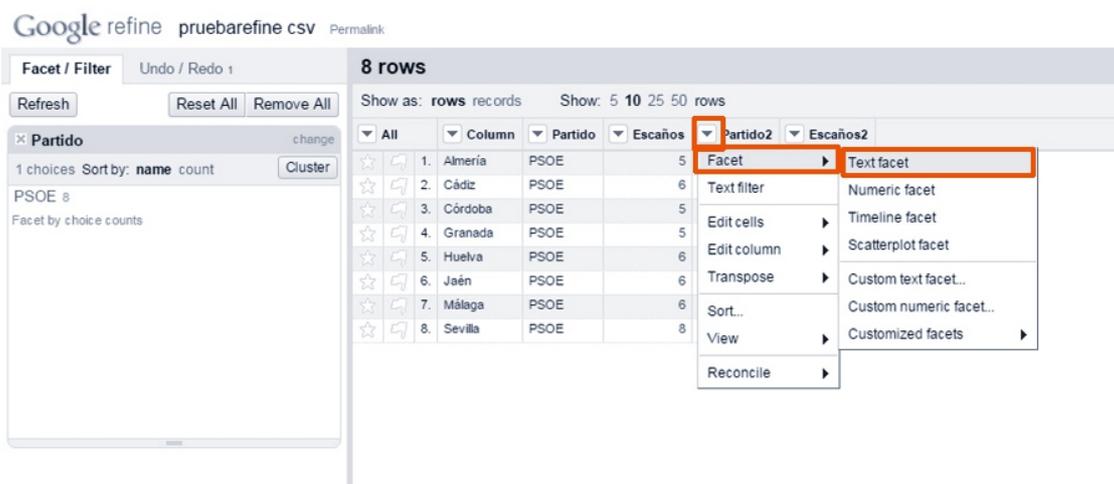
Cluster & Edit column "Partido"
 This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. Find out more ...
 Method: key collision Keying Function: metaphone3 2 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
4	5	<ul style="list-style-type: none"> PSOE (2 rows) PSA (1 rows) PSOE-A (1 rows) Psoe (1 rows) 	<input checked="" type="checkbox"/>	PSOE
3	3	<ul style="list-style-type: none"> P S O E (1 rows) P . S . O . E. (1 rows) P.S.O.E. (1 rows) 	<input checked="" type="checkbox"/>	PSOE

Choices in Cluster: 3 — 4
 # Rows in Cluster: 3 — 5
 Average Length of Choices: 4.25 — 8.67
 Length Variance of Choices: 1.089 — 1.7

Select All Deselect All Merge Selected & Re-Cluster **Merge Selected & Close** Close

12. Una vez que hemos editado toda la columna de PSOE, haremos lo mismo con la columna de **PP**.



Google refine pruebarefine.csv Permalink

Facet / Filter Undo / Redo 1

Refresh Reset All Remove All

Partido change Cluster

1 choices Sort by: name count

PSOE 8
 Facet by choice counts

8 rows Show as: rows records Show: 5 10 25 50 rows

All	Column	Partido	Esaños	Partido2	Esaños2
1	Almería	PSOE	5	Facet	Text facet
2	Cádiz	PSOE	6	Text filter	Numeric facet
3	Córdoba	PSOE	5	Edit cells	Timeline facet
4	Granada	PSOE	5	Edit column	Scatterplot facet
5	Huelva	PSOE	6	Transpose	Custom text facet...
6	Jaén	PSOE	6	Sort...	Custom numeric facet...
7	Málaga	PSOE	6	View	Customized facets
8	Sevilla	PSOE	8	Reconcile	



13.

Google refine [pruebarefine.csv](#) [Permalink](#)

Facet / Filter Undo / Redo 1

Refresh Reset All Remove All

Partido change
 1 choices Sort by: name count Cluster
 PSOE 8
 Facet by choice counts

Partido2 change
 6 choices Sort by: name count Cluster
 P P 1
 P. P. 1
 P.P. 1
 PP 3
 Pp 1
 PP-A 1
 Facet by choice counts

8 rows
 Show as: rows records Show: 5 10 25 50 rows

All	Column	Partido	Esaños	Partido2	Esaños2	
☆	1.	Almería	PSOE	5	PP	5
☆	2.	Cádiz	PSOE	6	P.P.	4
☆	3.	Córdoba	PSOE	5	Pp	4
☆	4.	Granada	PSOE	5	P P	4
☆	5.	Huelva	PSOE	6	PP	3
☆	6.	Jaén	PSOE	6	P. P.	4
☆	7.	Málaga	PSOE	6	PP-A	5
☆	8.	Sevilla	PSOE	8	PP	4

14.

Cluster & Edit column "Partido2"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. Find out more ...

Method Keying Function

Cluster Size	Row Count	Values in Cluster	Merge?
3	5	<ul style="list-style-type: none"> PP (3 rows) P.P. (1 rows) Pp (1 rows) 	<input type="checkbox"/>
2	2	<ul style="list-style-type: none"> P P (1 rows) P. P. (1 rows) 	<input type="checkbox"/>

Choices in Cluster:

Rows in Cluster:

Average Length of Choices:



15.

Cluster & Edit column "Partido2"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. Find out more ...

Method Keying Function 2 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
3	5	<ul style="list-style-type: none"> PP (3 rows) PP-A (1 rows) Pp (1 rows) 	<input checked="" type="checkbox"/>	<input type="text" value="PP"/>
3	3	<ul style="list-style-type: none"> P P (1 rows) P. P. (1 rows) P.P. (1 rows) 	<input checked="" type="checkbox"/>	<input type="text" value="PP"/>

Rows in Cluster:

Average Length of Choices:

Length Variance of Choices:

Select All Deselect All Merge Selected & Re-Cluster Merge Selected & Close Close

16. Cuando ya tengamos todos los datos editados, pulsamos en **Exportar** y seleccionamos el formato de archivo de datos que queramos (CSV).

Google refine pruebarefine.csv **Mass edit 8 cells in column Partido2** Undo

Open **Export** Help

Facet / Filter Undo / Redo 2 **8 rows** Show as: rows records Show: 5 10 25 50 rows

Refresh Reset All Remove All Extensions: Freebase

Partido 1 choices Sort by: name count Cluster

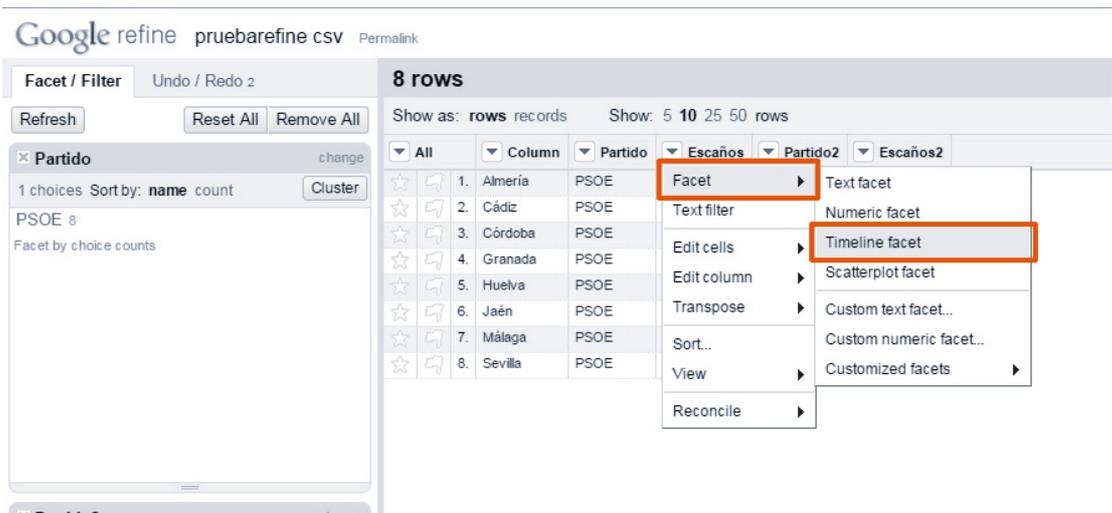
All	Column	Partido	Escala2	Partido2	Escala2
1	Almería	PSOE	5	PP	5
2	Cádiz	PSOE	6	PP	4
3	Córdoba	PSOE	5	PP	4
4	Granada	PSOE	5	PP	4
5	Huelva	PSOE	6	PP	3
6	Jáen	PSOE	6	PP	4
7	Málaga	PSOE	6	PP	5
8	Sevilla	PSOE	6	PP	4

Partido2 1 choices Sort by: name count Cluster

PP



Con Open Refine, también podemos editar un gran conjunto de datos de tiempo,
como fechas, horas, etc.



The screenshot shows the Google Refine interface with a table of 8 rows. The columns are 'All', 'Column', 'Partido', 'Esaños', 'Partido2', and 'Esaños2'. The 'Esaños' column is selected, and a context menu is open over it. The 'Facet' option is selected, and the 'Timeline facet' sub-option is highlighted. The table data is as follows:

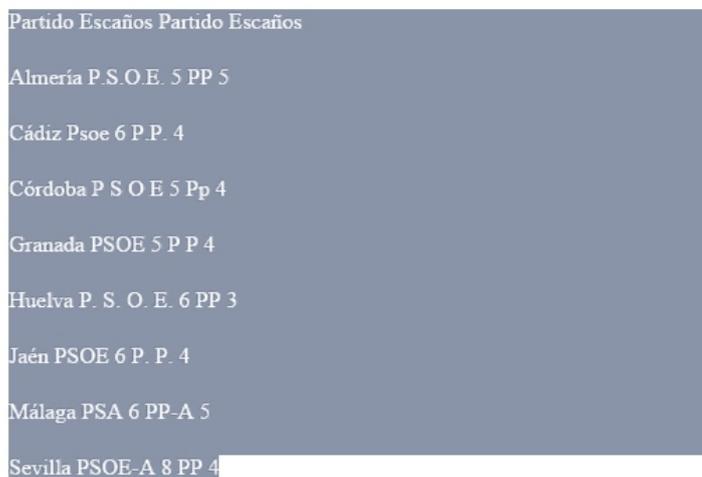
	All	Column	Partido	Esaños	Partido2	Esaños2
1.	Almería	PSOE				
2.	Cádiz	PSOE				
3.	Córdoba	PSOE				
4.	Granada	PSOE				
5.	Huelva	PSOE				
6.	Jaén	PSOE				
7.	Málaga	PSOE				
8.	Sevilla	PSOE				

CONTENIDOS 2.3

OpenOffice Calc

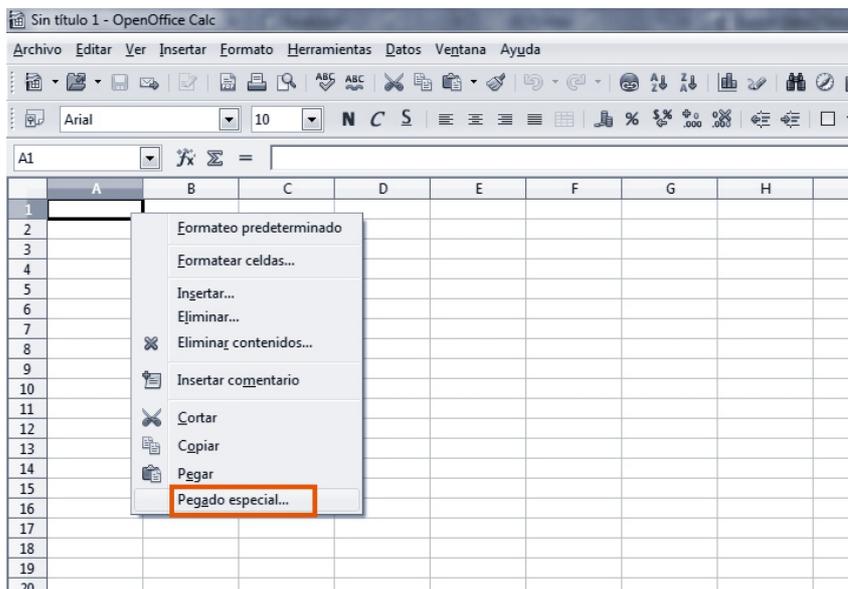
Copiar y pegar datos de una web

1. Seleccionamos el texto y le damos a **copiar**.

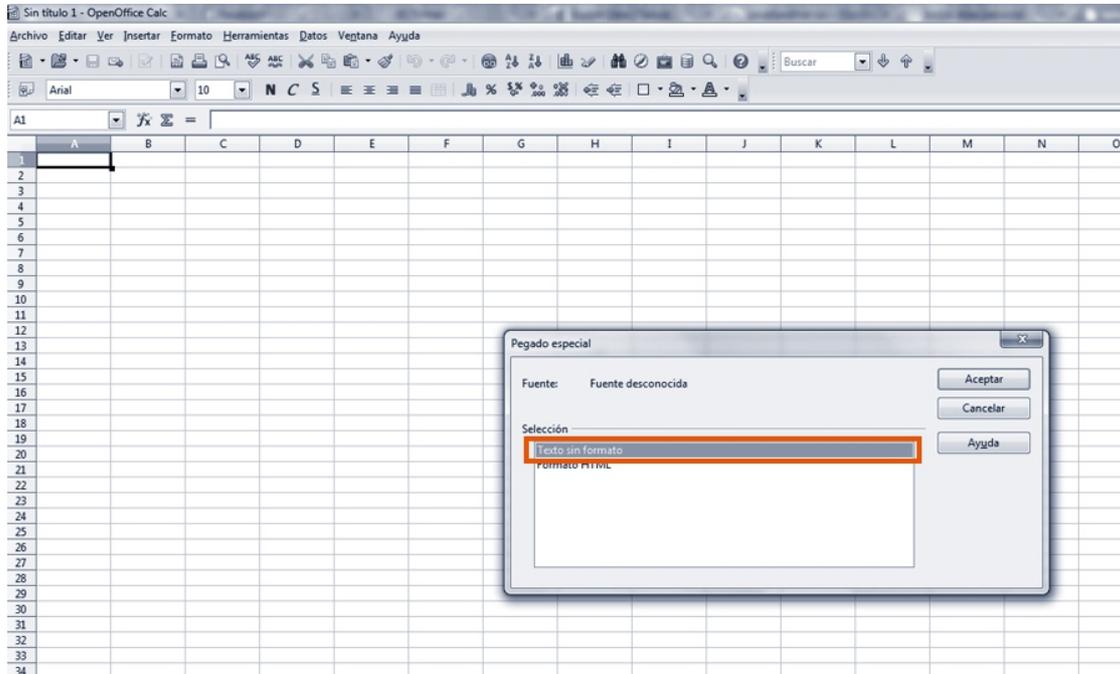


Partido Escaños Partido Escaños
Almería P.S.O.E. 5 PP 5
Cádiz Psoe 6 P.P. 4
Córdoba P S O E 5 Pp 4
Granada PSOE 5 P P 4
Huelva P. S. O. E. 6 PP 3
Jaén PSOE 6 P. P. 4
Málaga PSA 6 PP-A 5
Sevilla PSOE-A 8 PP 4

2. Creamos un **archivo nuevo** en OpenOffice Calc y, en la primera celda en blanco, le damos a botón derecho > **Pegado especial...** (**CONTROL+MAYÚSCULA+V**)



3. En las opciones de pegado, siempre hay que seleccionar **Texto sin formato**.

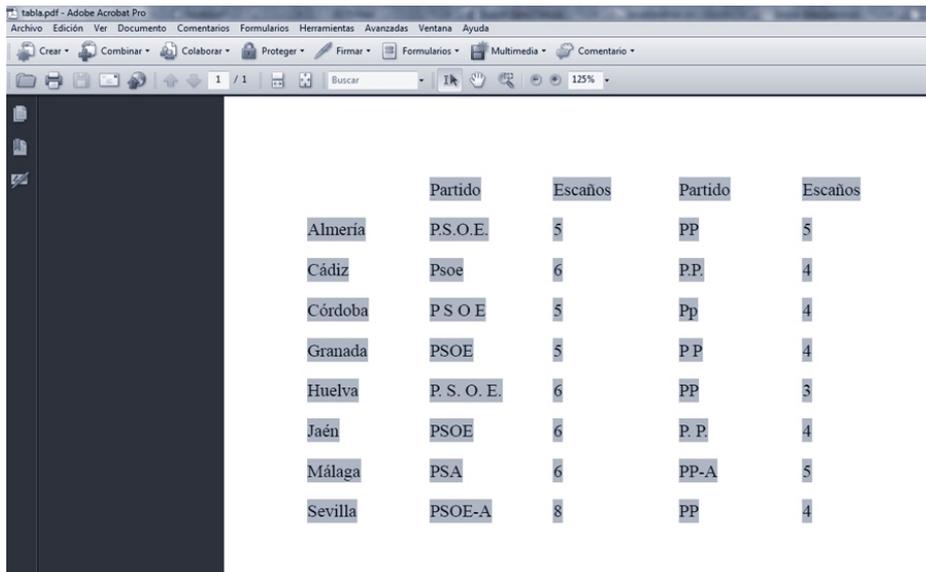


4. Y ya tendremos la tabla pegada en nuestra hoja de cálculo.

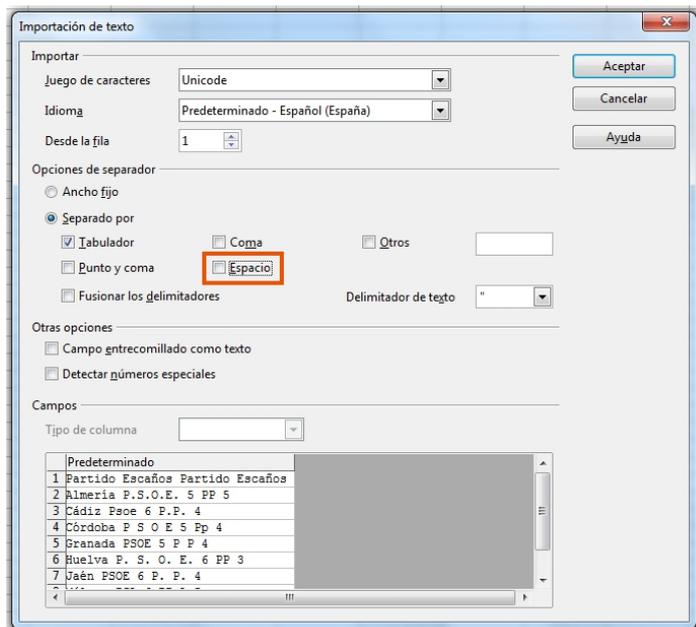
	A	B	C	D	E	F
1	Partido	Escaños	Partido	Escaños		
2	Almería	P.S.O.E.		5	PP	5
3	Cádiz	Psoe		6	P.P.	4
4	Córdoba	P S O E		5	Pp	4
5	Granada	PSOE		5	P P	4
6	Huelva	P. S. O. E.		6	PP	3
7	Jaén	PSOE		6	P. P.	4
8	Málaga	PSA		6	PP-A	5
9	Sevilla	PSOE-A		8	PP	4
10						
11						

Copiar y pegar datos de un PDF

1. Seleccionamos el texto y le damos a **copiar**.

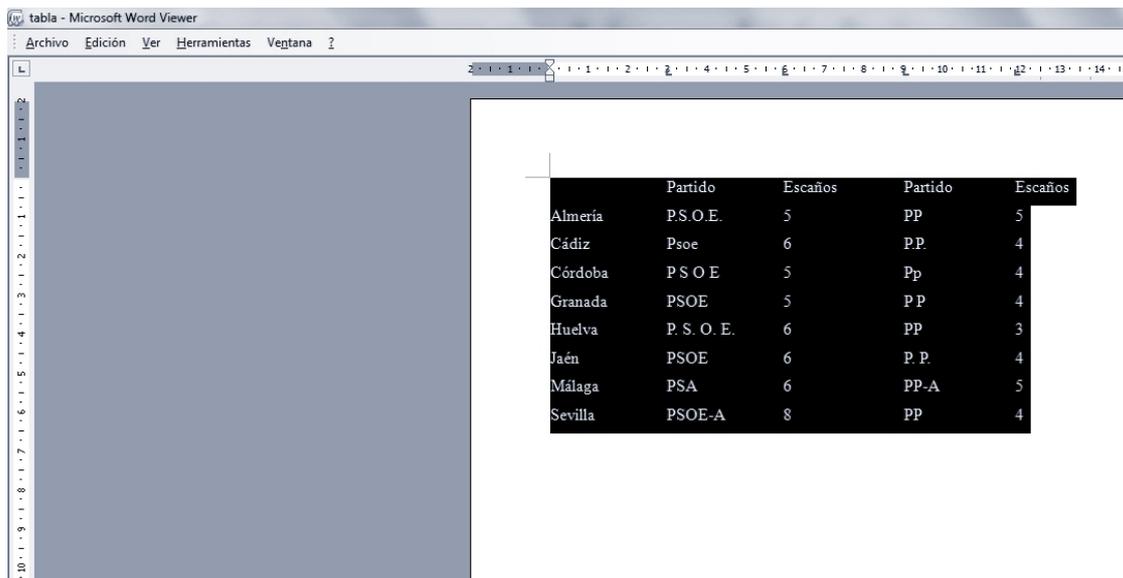


2. En OpenOffice Calc, le damos a **Pegado especial** y nos saltará las opciones de **Importar texto**. A la hora de seleccionar los **separadores de campo**, solo podríamos escoger **Espacio**, ya que el texto original no viene con ningún elemento tipográfico que pueda separar los elementos en columnas. Aquí es donde entraría en juego **Tabula PDF**, quien nos transformará este PDF sin separadores de campo en un CSV separado por comas.

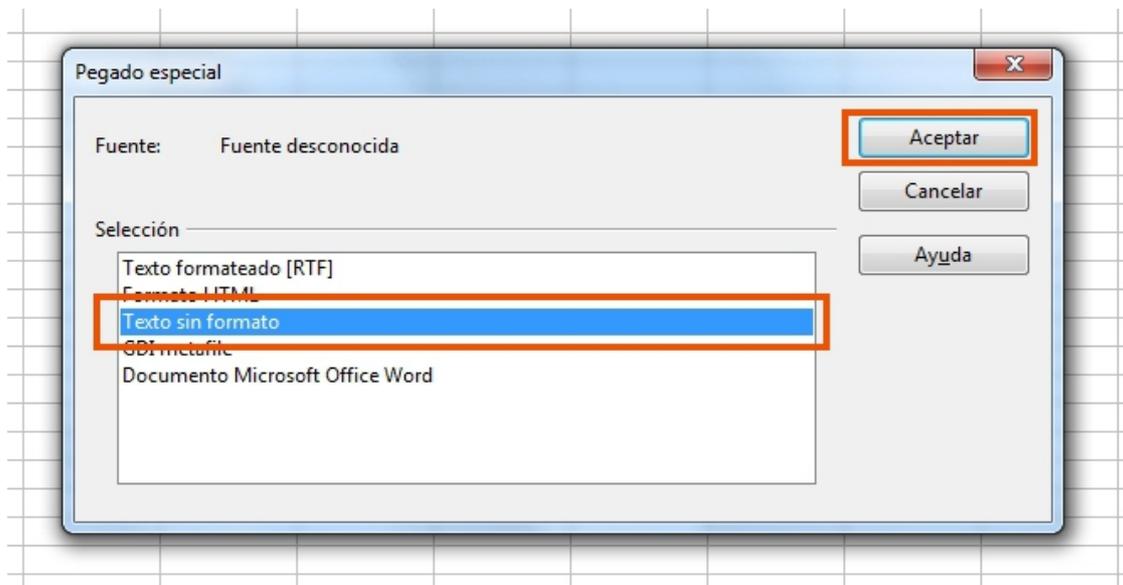


Copiar y pegar datos de un texto

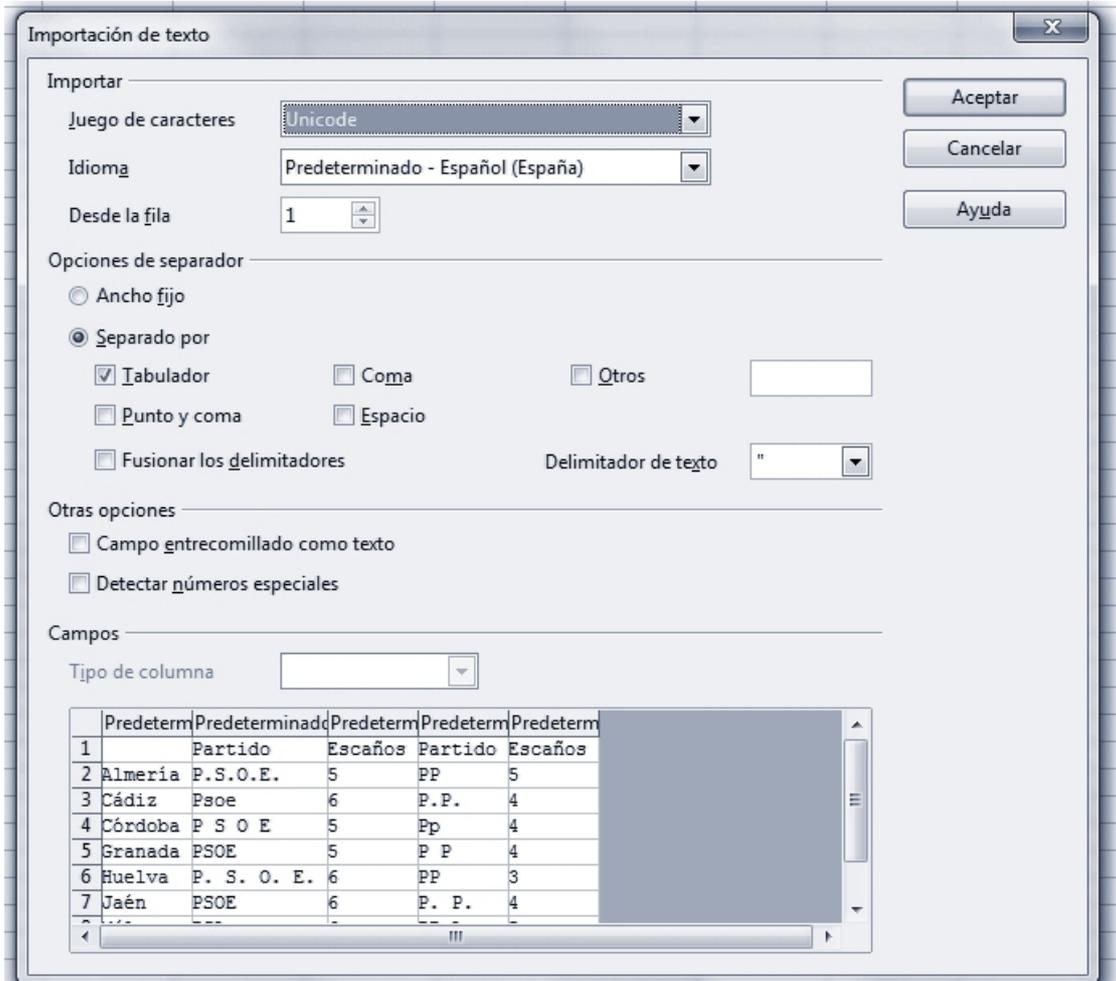
1. Seleccionamos el texto y le damos a **copiar**.



2. En OpenOffice Calc, le damos a **Pegado especial** y, después, **Texto sin formato**.



3. En las opciones de **Importación de texto** seleccionamos los **separadores de campo** que mejor nos funcionen, en este caso, los tabuladores.





Transformación de datos (formatear celdas)

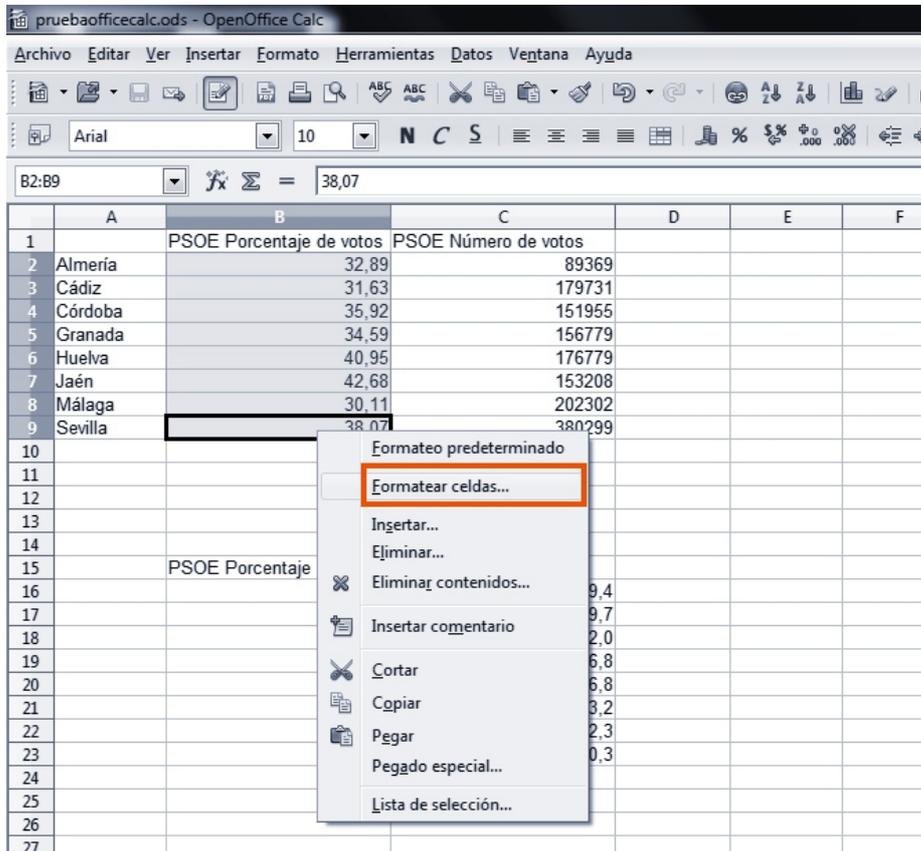
1. Pulsamos en la primera celda a formatear.

	A	B	C	D	E
1		PSOE Porcentaje de votos	PSOE Número de votos		
2	Almería	32,89	89369		
3	Cádiz	31,63	179731		
4	Córdoba	35,92	151955		
5	Granada	34,59	156779		
6	Huelva	40,95	176779		
7	Jaén	42,68	153208		
8	Málaga	30,11	202302		
9	Sevilla	38,07	380299		
10					
11					
12					

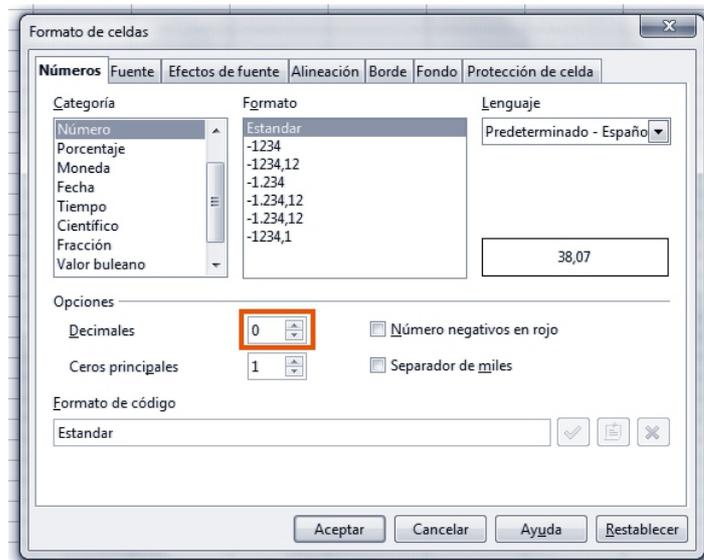
2. Con la **MAYÚSCULA** pulsada, seleccionamos la última celda con datos de esa columna.

	A	B	C	D	E
1		PSOE Porcentaje de votos	PSOE Número de votos		
2	Almería	32,89	89369		
3	Cádiz	31,63	179731		
4	Córdoba	35,92	151955		
5	Granada	34,59	156779		
6	Huelva	40,95	176779		
7	Jaén	42,68	153208		
8	Málaga	30,11	202302		
9	Sevilla	38,07	380299		
10					
11					
12					

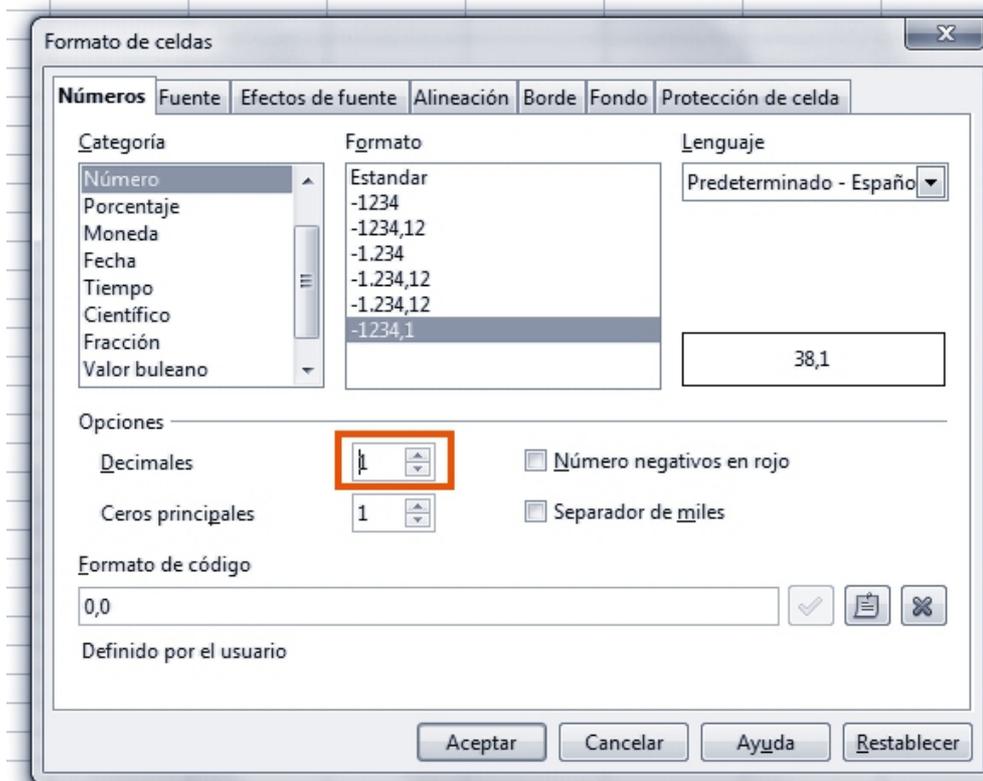
3. Le damos al botón derecho y seleccionamos **Formatear celdas**.



4. Por ejemplo, podemos cambiar el número de **decimales**.



5. Al poner 1, le decimos que todos los datos que hemos seleccionado tengan un solo decimal.



6. Al dar a **Aceptar**, quedaría ya formateado.

pruebaofficecalc.ods - OpenOffice Calc

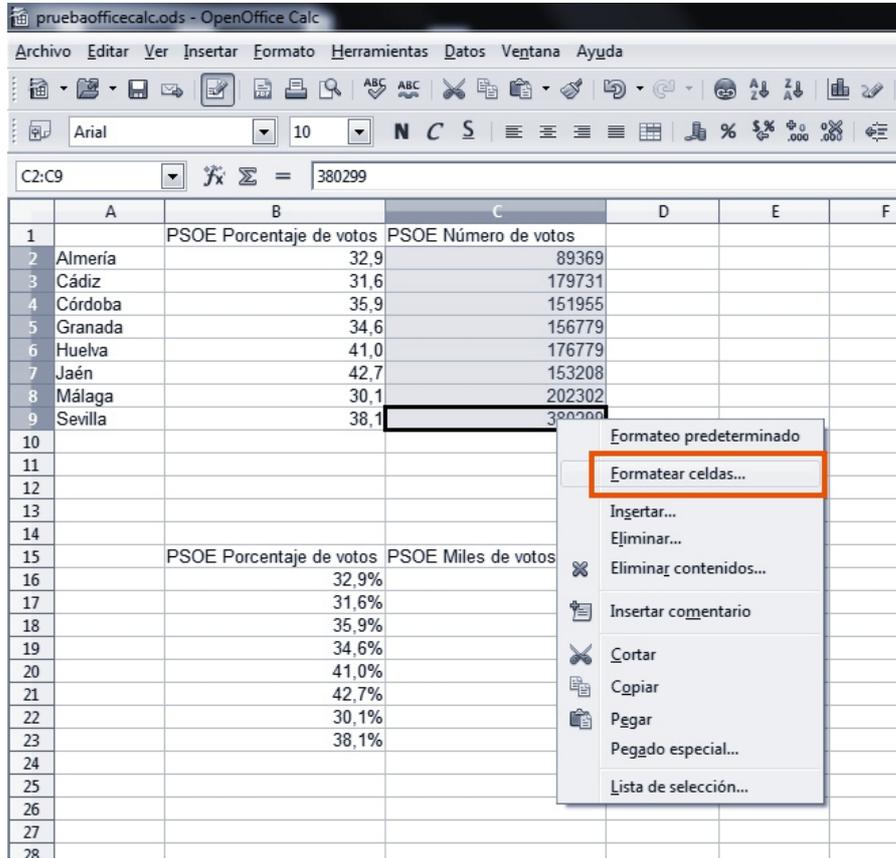
Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

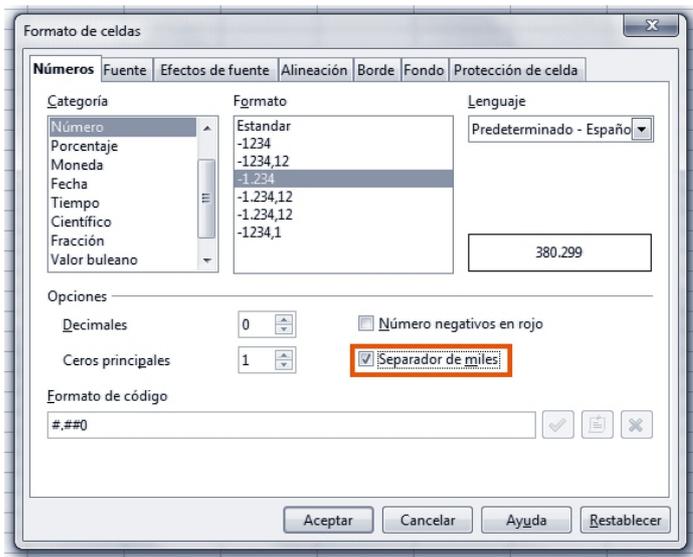
B2:B9 = 38,07

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,9	89369	
3	Cádiz	31,6	179731	
4	Córdoba	35,9	151955	
5	Granada	34,6	156779	
6	Huelva	41,0	176779	
7	Jaén	42,7	153208	
8	Málaga	30,1	202302	
9	Sevilla	38,1	380299	
10				
11				
12				

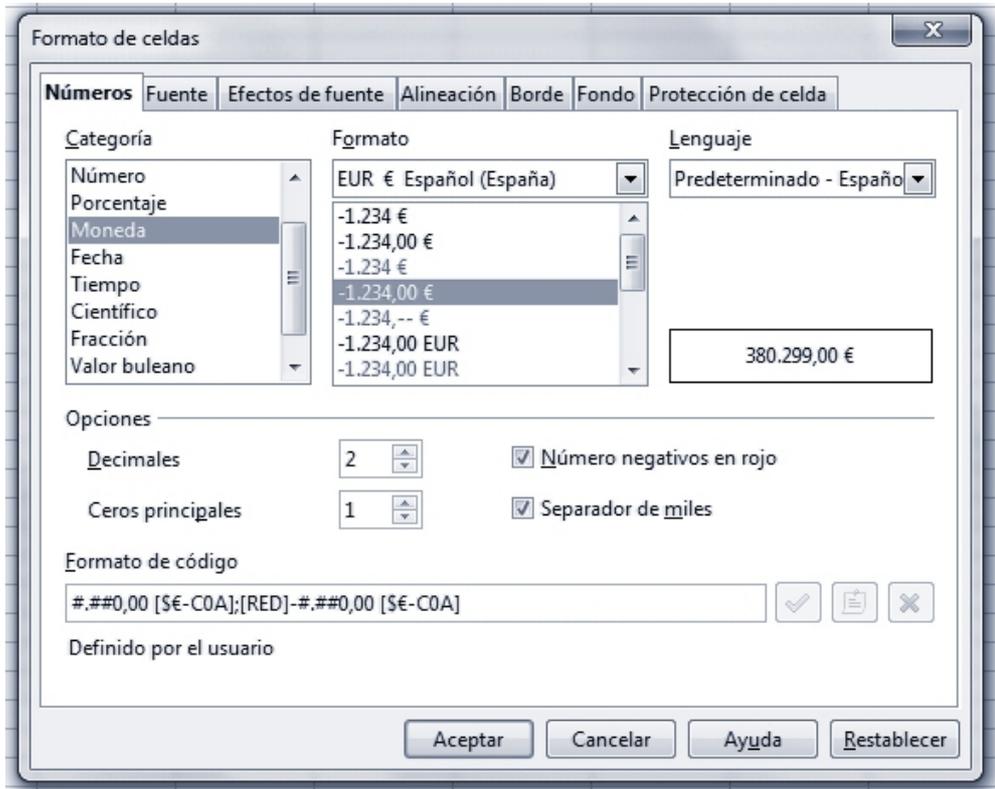
7. Si formateamos la segunda columna de datos...



8. Podremos, por ejemplo, ponerle a todos esos datos el punto que separa los millares.



9. Entre las diversas opciones que ofrece el Formato de celdas, tenemos también la opción de Moneda, Tiempo o Porcentaje.

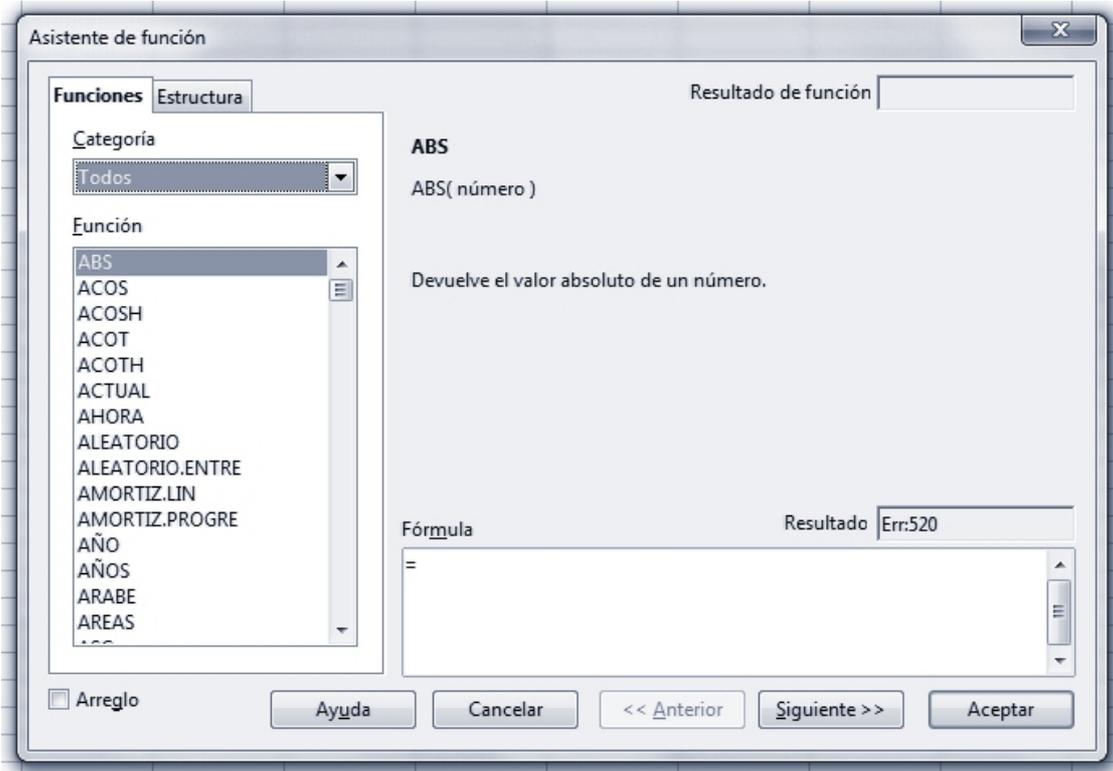


Transformación y creación de datos (uso de fórmulas)

1. Si pulsamos en el icono del **Asistente de Funciones**...

	A	B	C	D	E	F	G
1		PSOE Porcentaje de votos	PSOE Número de votos				
2	Almería	32,9	89.369				
3	Cádiz	31,6	179.731				
4	Córdoba	35,9	151.955				
5	Granada	34,6	156.779				
6	Huelva	41,0	176.779				
7	Jaén	42,7	153.208				
8	Málaga	30,1	202.302				
9	Sevilla	38,1	380.299				
10							

2. Nos saldrá el catálogo de fórmulas que tenemos a nuestra disposición.



3. Pero si es una fórmula sencilla, la podemos escribir nosotros mismos, antes seleccionando una **celda que esté vacía de datos**.

pruebaofficecalc.ods - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

C12

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12				
13				
14				

4. Por ejemplo, si queremos **sumar** los datos de dos celdas. Para ello, las referencias de cada celda que se va a sumar, C2 y C3, deben estar separadas por ;

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12			=SUMA(C2;C3)	
13				
14				



importante

Las fórmulas siempre deben ser iniciadas por un símbolo =, después, irán seguidas de la **operación** que se desee realizar y, a continuación entre paréntesis, **las celdas** a las que afecta con las operaciones complementarias que se quieran hacer:

+ suma - resta * multiplicar / dividir



5. Pero sería interesante que pudiéramos **sumar todas las celdas** que se encuentran entre una referencia de celda y otra. Para ello, repetiremos la fórmula, pero separando las referencias de celdas de inicio y fin por :

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12			=SUMA(C2:C9)	
13				

6. Al dar al **'check'** de la barra de fórmulas, veremos el resultado.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12			1490422	
13				
14				



7. También podríamos calcular el **promedio** de todas las celdas entre una referencia de inicio y otra de fin.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		=PROMEDIO(B2:B9)	1490422	
13				

8. También nos servirá para **realizar cálculos o aplicar formatos** de celdas a valores que se encuentran en otras celdas.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		=B2/100	89,4	
17			179,7	
18			152,0	
19			156,8	
20			176,8	
21			153,2	
22			202,3	
23			380,3	
24				



9. Una vez creada la primera fórmula en la primera celda de la nueva lista de datos, podemos pinchar en la **esquina inferior derecha de la celda** y, sin soltar, arrastrar hacia abajo. De esta manera, la fórmula se repetirá, pero el programa irá cambiando el valor de la fila. Del 2 inicial, pasará a la fila 3, luego a la fila 4...

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		0,3289	89,4	
17			179,7	
18			152,0	
19			156,8	
20			176,8	
21			153,2	
22			202,3	
23			380,3	
24				

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		0,3289	89,4	
17		0,3163	179,7	
18		0,3592	152,0	
19		0,3459	156,8	
20		0,4095	176,8	
21		0,4268	153,2	
22		0,3011	202,3	
23		0,3807	380,3	
24				



10. De esta manera, podríamos **reducir millares** de una manera muy sencilla y rápida y sin perder los datos originales.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		32,89%		
17		31,63%		
18		35,92%		
19		34,59%		
20		40,95%		
21		42,68%		
22		30,11%		
23		38,07%		
24				
25				

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89.369	
3	Cádiz	31,63	179.731	
4	Córdoba	35,92	151.955	
5	Granada	34,59	156.779	
6	Huelva	40,95	176.779	
7	Jaén	42,68	153.208	
8	Málaga	30,11	202.302	
9	Sevilla	38,07	380.299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		32,89%	=C2/1000	
17		31,63%		
18		35,92%		
19		34,59%		
20		40,95%		
21		42,68%		
22		30,11%		
23		38,07%		
24				
25				



pruebaofficecalc.ods - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

C16 =C2/1000

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89.369	
3	Cádiz	31,63	179.731	
4	Córdoba	35,92	151.955	
5	Granada	34,59	156.779	
6	Huelva	40,95	176.779	
7	Jaén	42,68	153.208	
8	Málaga	30,11	202.302	
9	Sevilla	38,07	380.299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		32,89%	89,37	
17		31,63%		
18		35,92%		
19		34,59%		
20		40,95%		
21		42,68%		
22		30,11%		
23		38,07%		
24				

pruebaofficecalc.ods - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

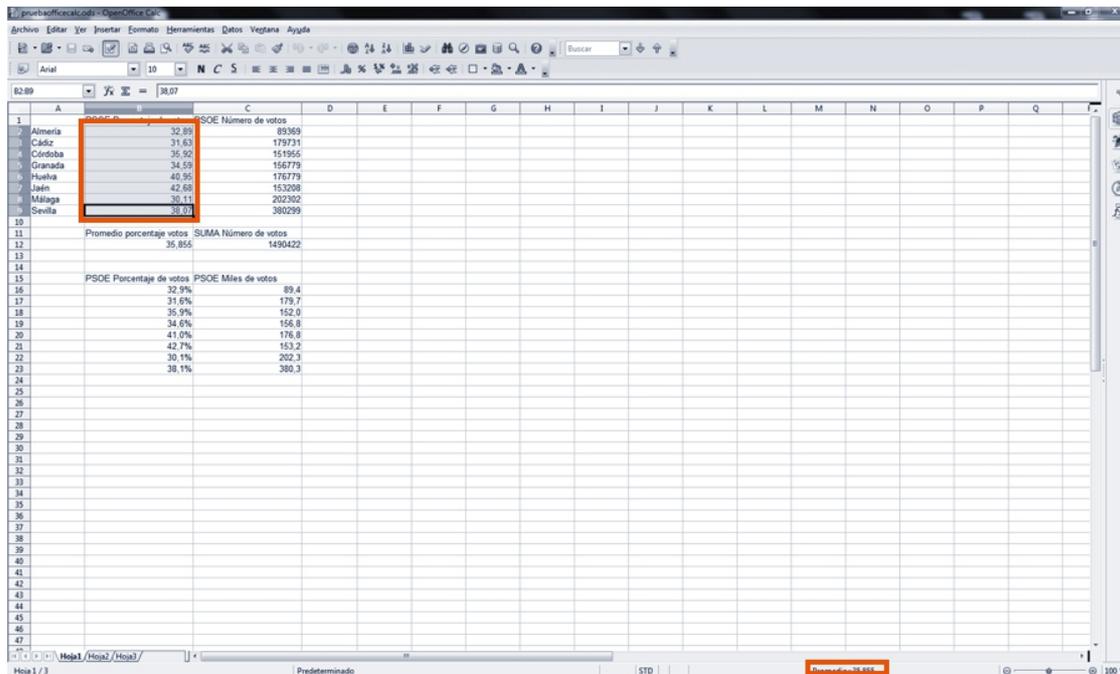
Arial 10 N C S

C16:C23 =C2/1000

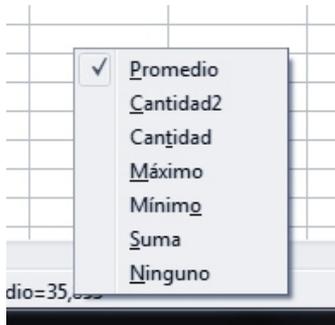
	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89.369	
3	Cádiz	31,63	179.731	
4	Córdoba	35,92	151.955	
5	Granada	34,59	156.779	
6	Huelva	40,95	176.779	
7	Jaén	42,68	153.208	
8	Málaga	30,11	202.302	
9	Sevilla	38,07	380.299	
10				
11		Promedio porcentaje votos	SUMA Número de votos	
12		35,855	1490422	
13				
14				
15		PSOE Porcentaje de votos	PSOE Miles de votos	
16		32,89%	89,37	
17		31,63%	179,73	
18		35,92%	151,96	
19		34,59%	156,78	
20		40,95%	176,78	
21		42,68%	153,21	
22		30,11%	202,30	
23		38,07%	380,30	
24				
25				

Preguntas a los datos (acceso directo a preguntas generales)

1. Si seleccionamos un grupo de datos, en la **esquina inferior**, parcialmente a la derecha, tenemos una pestaña con información general de esos datos.



2. Pulsando en esta pestaña con el **botón derecho** accederemos a las opciones: **Promedio**, **Cantidad2** y **Cantidad** (que se refieren al número de filas o de columnas de los datos seleccionados), **Máximo**, **Mínimo** y **Suma**. El resultado de cualquiera de estas opciones se indicará en la misma pestaña. Más que para editar datos, esta opción nos sirve como un atajo rápido a información básica sobre los datos.



Preguntas a los datos (ordenar)

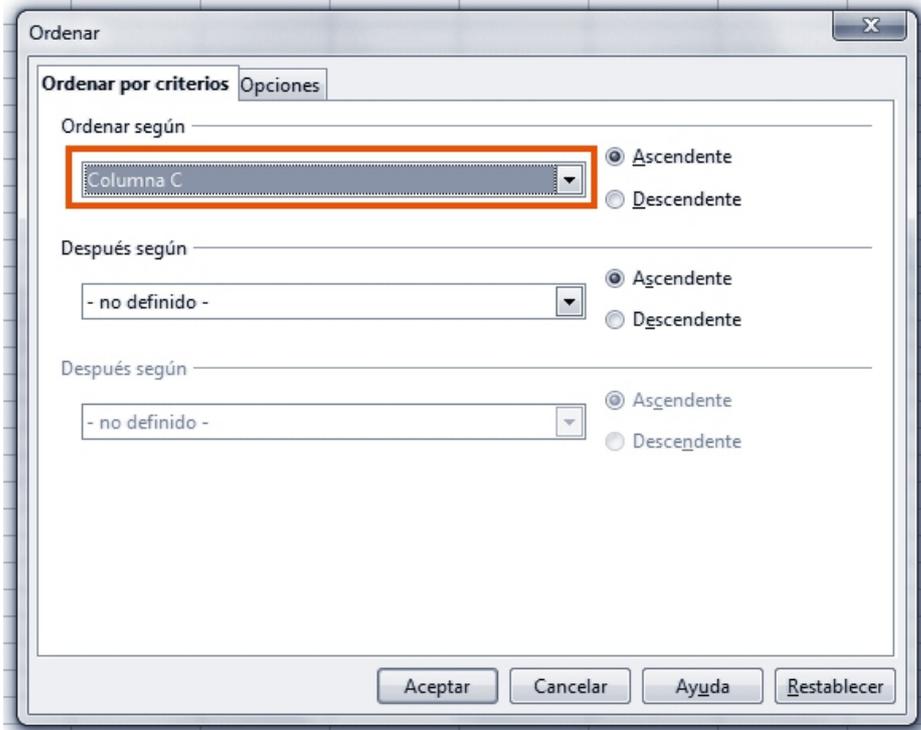
1. Seleccionamos todos los datos a ordenar, pulsando en la celda de la esquina superior izquierda y, con la **MAYÚSCULA** pulsada, en la celda de la esquina inferior derecha.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11				
12				
13				

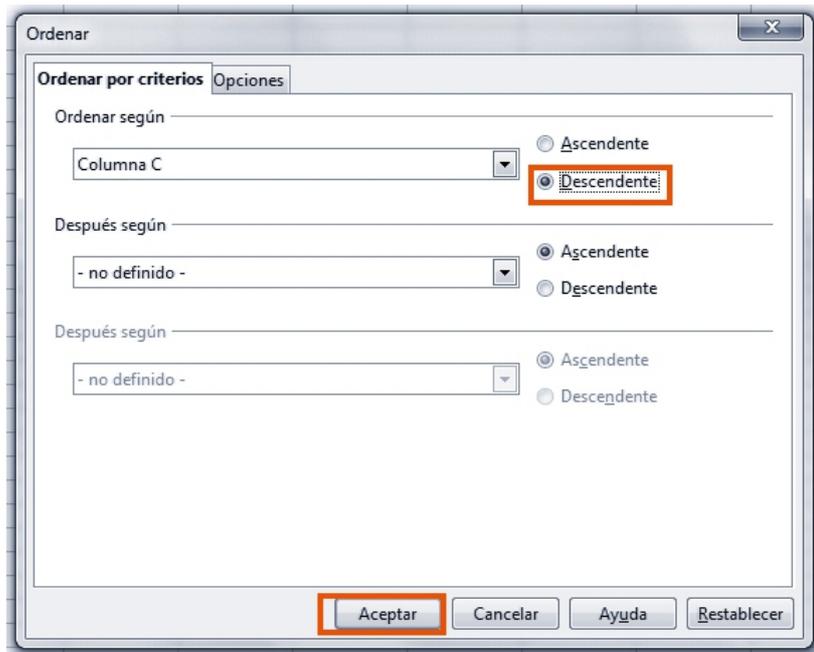
2. Pulsamos en **Datos > Ordenar**.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE	
2	Almería	32,89		
3	Cádiz	31,63		
4	Córdoba	35,92		
5	Granada	34,59		
6	Huelva	40,95		
7	Jaén	42,68		
8	Málaga	30,11		
9	Sevilla	38,07		
10				
11				

3. Elegimos la columna que servirá de **referencia para ordenar** los datos.



4. Elegimos si el orden será Ascendente o Descendente, en este caso, **descendente** porque nos interesa ordenar **de mayor a menor**.





5. Nos ordena todos los datos tomando como referencia el orden de mayor a menor de los datos de la columna C (Número de votos).

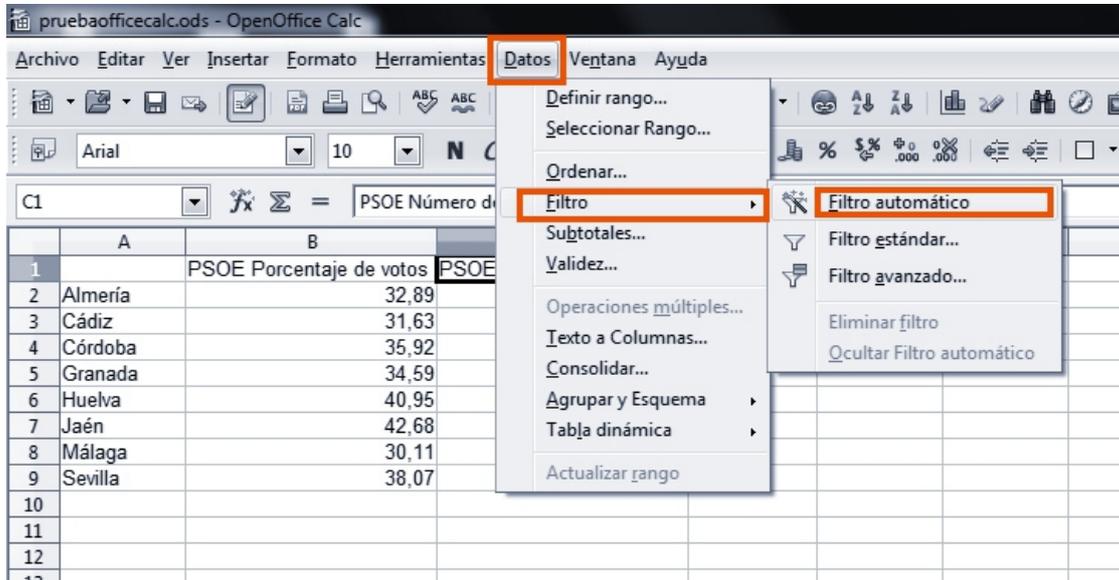
	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Sevilla	38,07	380299	
3	Málaga	30,11	202302	
4	Cádiz	31,63	179731	
5	Huelva	40,95	176779	
6	Granada	34,59	156779	
7	Jaén	42,68	153208	
8	Córdoba	35,92	151955	
9	Almería	32,89	89369	
10				

Preguntas a los datos (filtros)

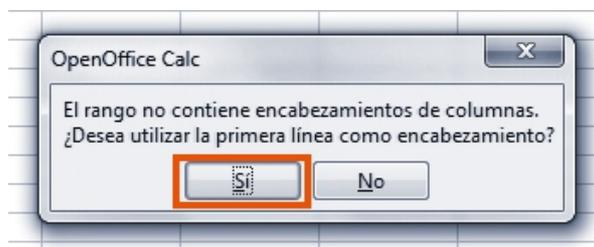
1. Pulsamos en el encabezamiento de una columna.

	A	B	C	D
1		PSOE Porcentaje de votos	PSOE Número de votos	
2	Almería	32,89	89369	
3	Cádiz	31,63	179731	
4	Córdoba	35,92	151955	
5	Granada	34,59	156779	
6	Huelva	40,95	176779	
7	Jaén	42,68	153208	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11				
12				

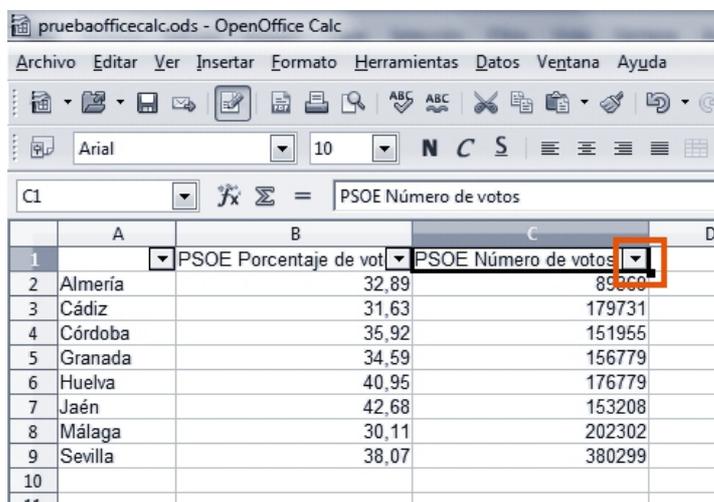
2. Seleccionamos **Filtro automático**.



3. Le decimos que **SÍ** deseamos usar la primera línea como encabezamiento.



4. Pulsamos en la **flecha hacia abajo** del encabezamiento de la columna sobre la que queremos aplicar el filtro.



5. Seleccionamos **Filtro predeterminado**.

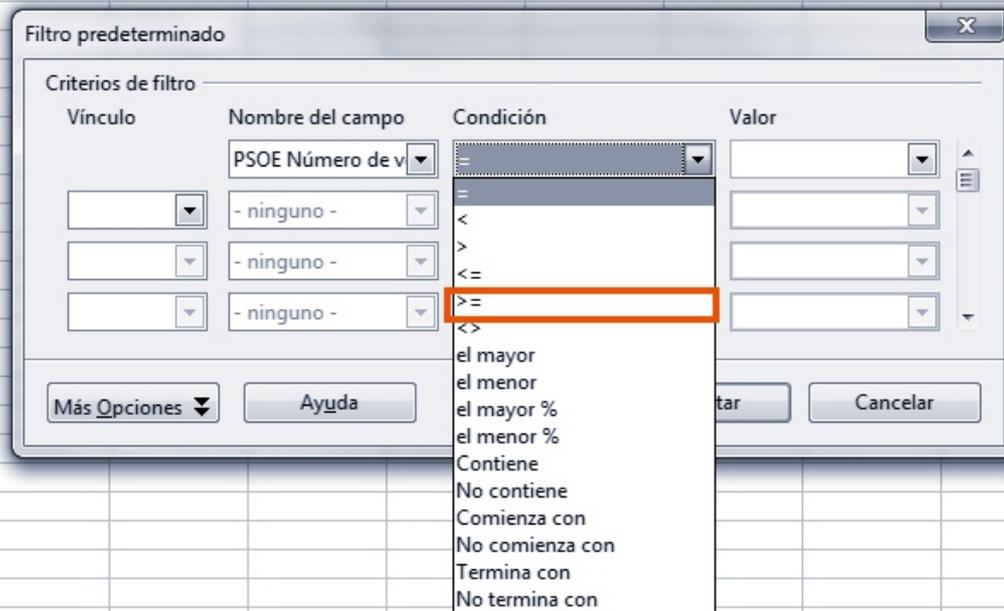
	A	B	C	D	E
1		PSOE Porcentaje de vot	PSOE Número de votos		
2	Almería	32,89	Todo		
3	Cádiz	31,63	Los 10 primeros		
4	Córdoba	35,92	Filtro predeterminado...		
5	Granada	34,59	89369		
6	Huelva	40,95	151955		
7	Jaén	42,68	153208		
8	Málaga	30,11	156779		
9	Sevilla	38,07	176779		
10			179731		
11			202302		
12			380299		
13					
14					

6. En la **Condición**, elegimos la que más nos convenga.

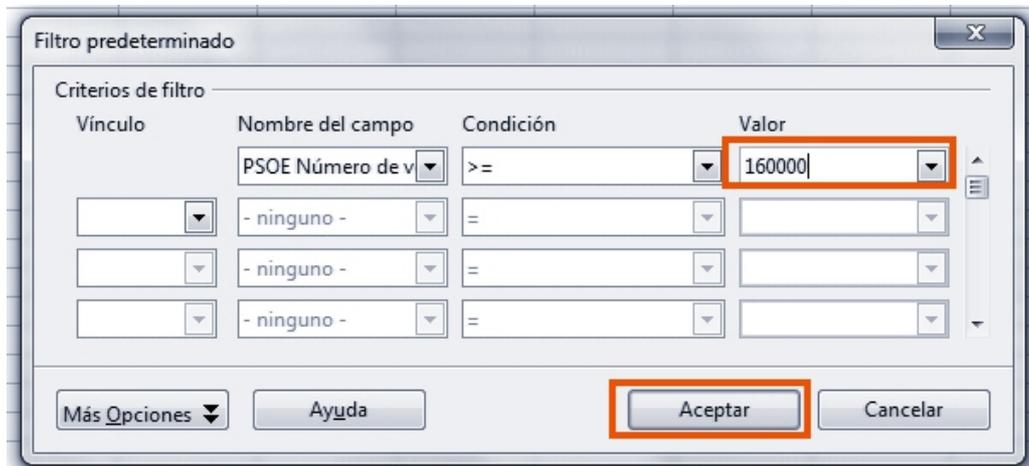
Filtro predeterminado			
Criterios de filtro			
Vínculo	Nombre del campo	Condición	Valor
	PSOE Número de v	=	
	- ninguno -	=	
	- ninguno -	=	
	- ninguno -	=	

Más Opciones | Ayuda | Aceptar | Cancelar

7. En este caso, escogeremos **mayor o igual**.



8. Y pondremos un **valor** de referencia.





9. De esta manera, si está aplicado el filtro, sólo veremos los datos de las provincias donde el PSOE ha conseguido 160.000 votos o más.

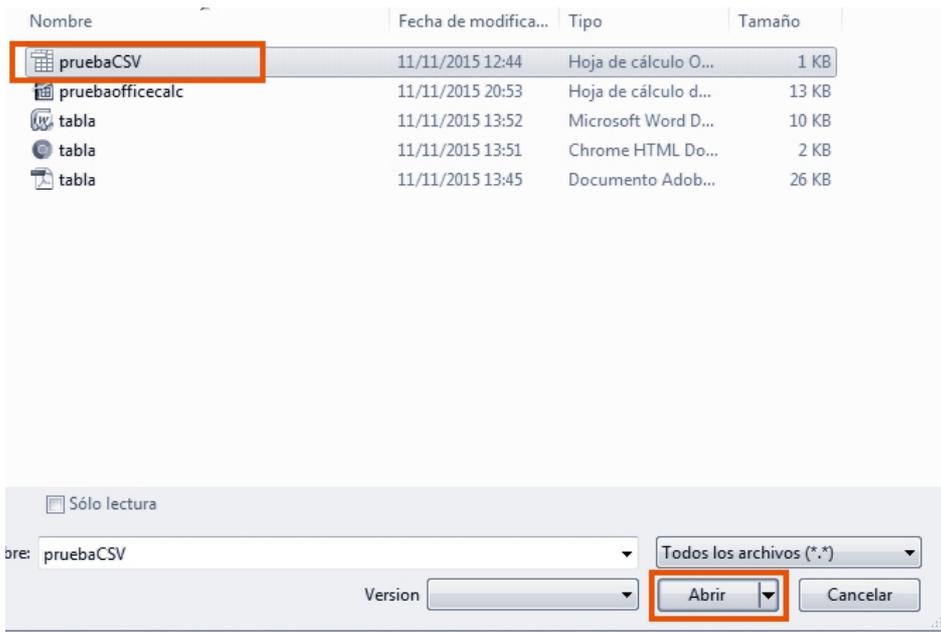
	A	B	C	D
1		PSOE Porcentaje de vot	PSOE Número de votos	
3	Cádiz	31,63	179731	
6	Huelva	40,95	176779	
8	Málaga	30,11	202302	
9	Sevilla	38,07	380299	
10				
11				

CONTENIDOS | 2.4

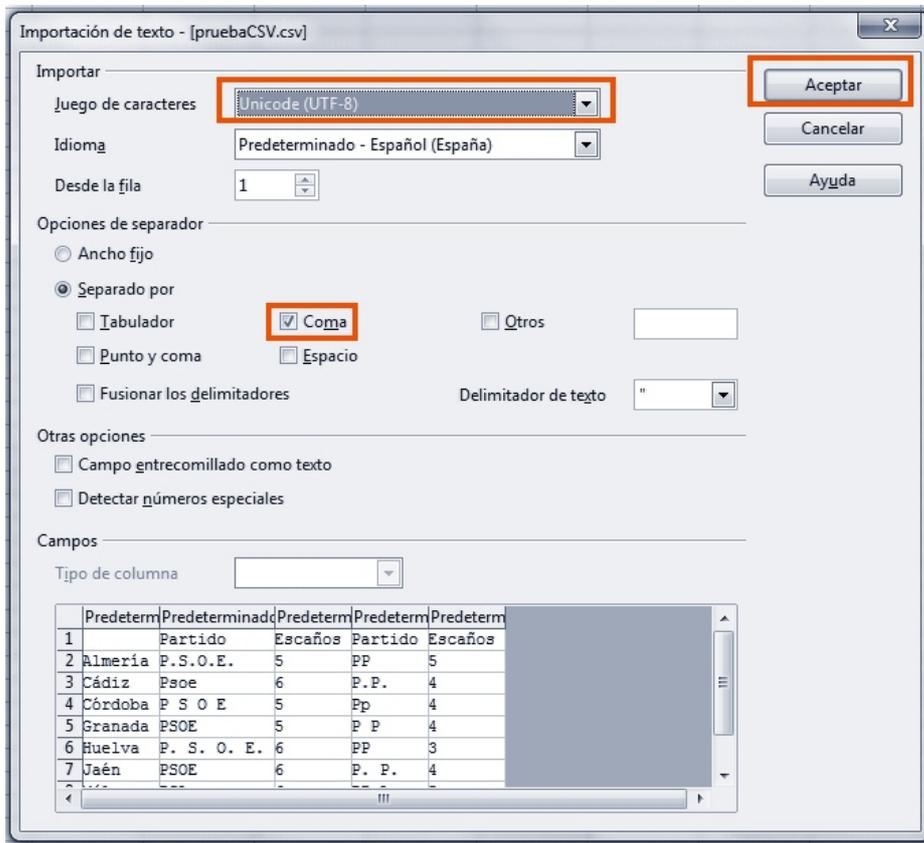
OpenOffice Calc

Abrir y editar el formato de archivos .csv

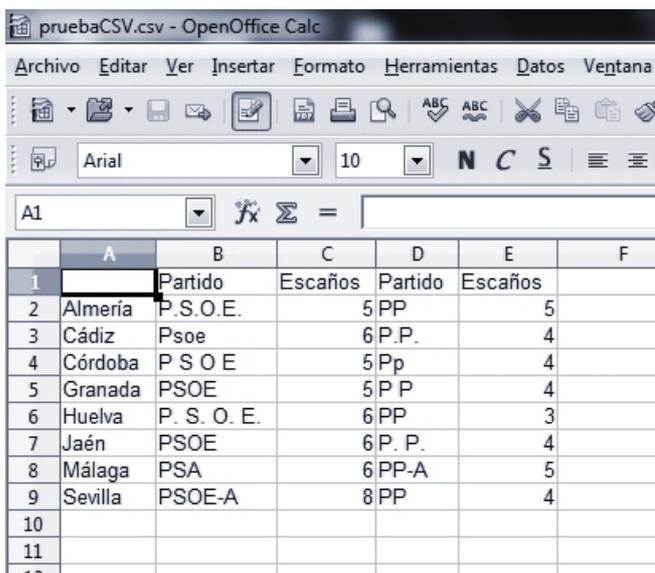
1. Abrimos un archivo con formato **CSV**.



2. Seleccionamos el **Juego de caracteres** que mejor resulte con los datos originales y, como **separador de campo**, elegiremos **Coma**.



3. Y ya tendremos el CSV en OpenOffice Calc.

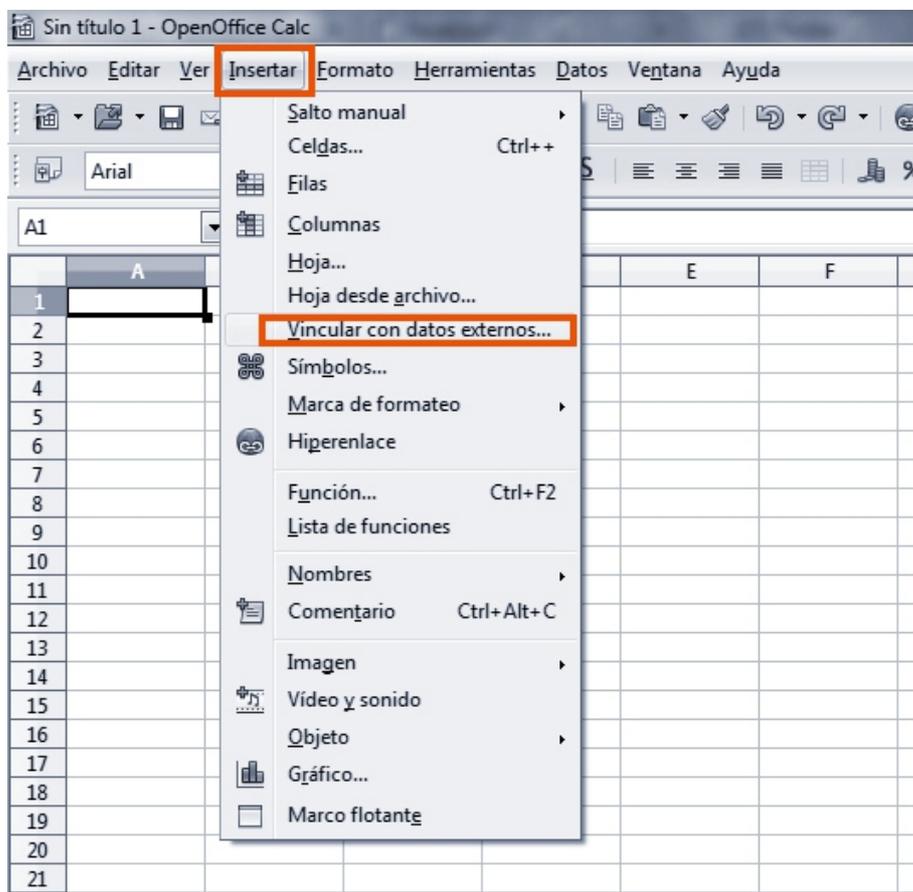


Actualización automática de datos (vincular con datos externos)

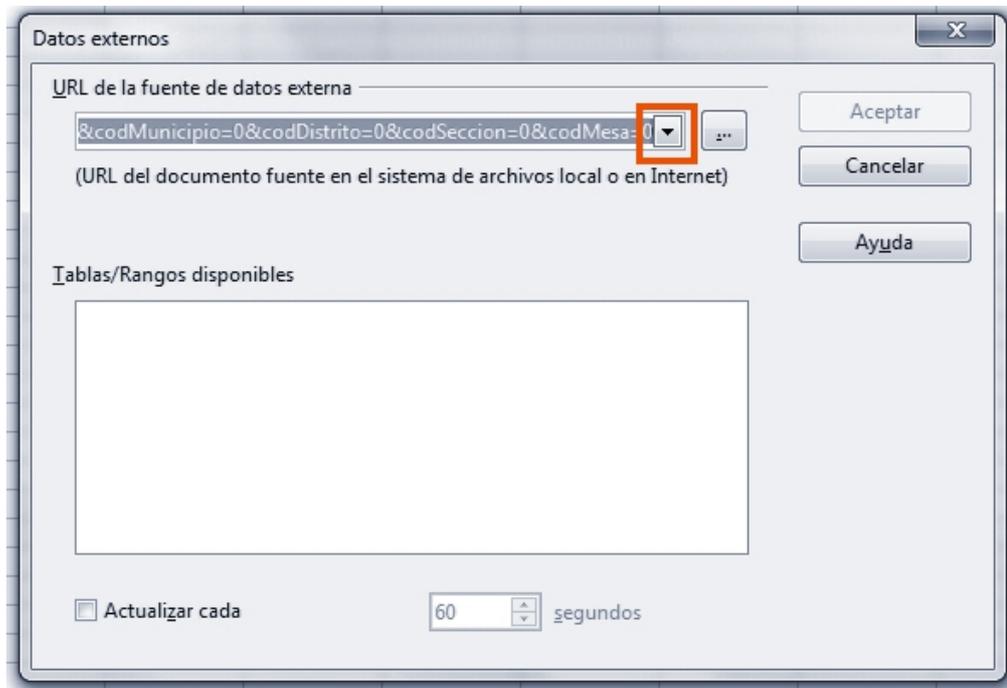
1. Como ejemplo, podemos copiar esta dirección web:

<http://www.infoelectoral.interior.es/min/busquedaAvanzadaAction.html?vuelta=1&codTipoEleccion=2&codPeriodo=201111&codEstado=99&codComunidad=0&codProvincia=0&codMunicipio=0&codDistrito=0&codSeccion=0&codMesa=0>

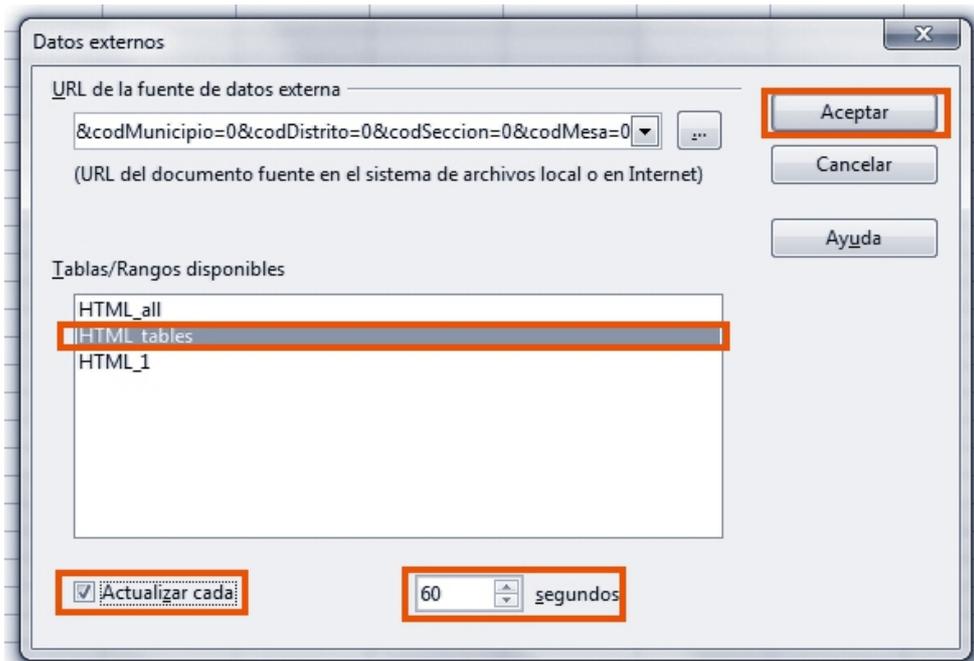
2. En OpenOffice Calc, seleccionamos Insertar > Vincular con datos externos.



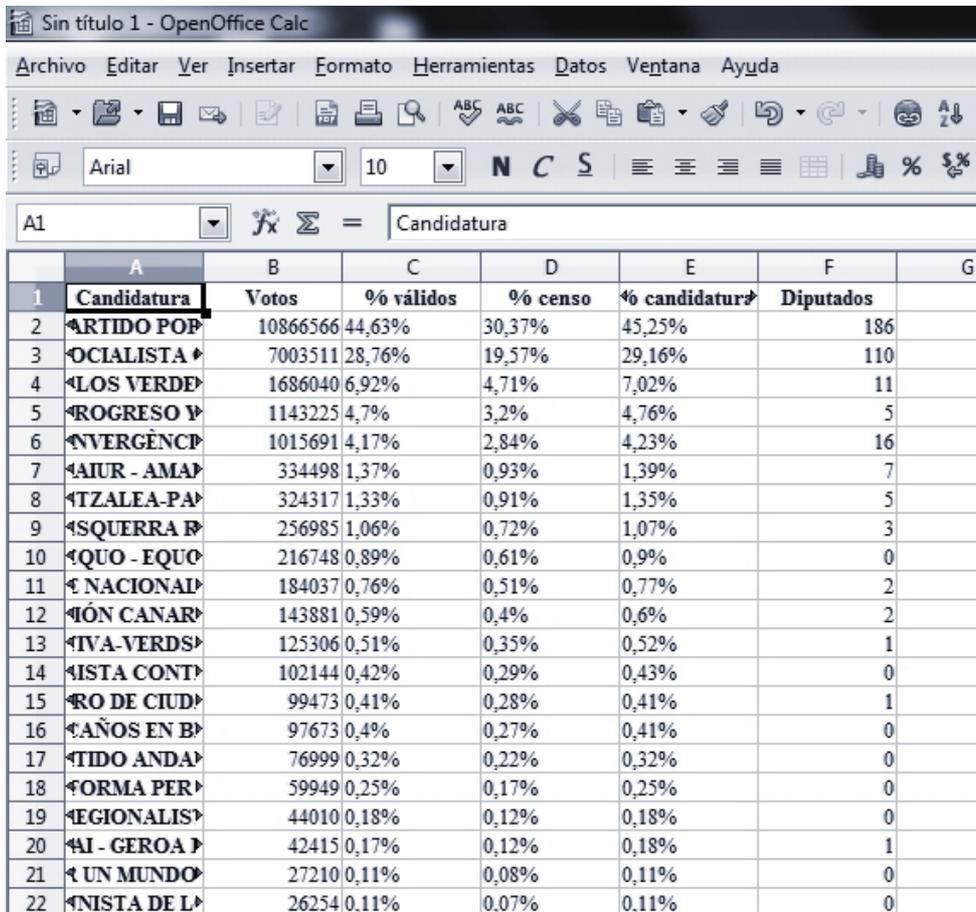
3. En **URL** pegamos la dirección de la página web a la que queremos vincularnos y pulsamos una vez en la flecha hacia abajo y, después, otra vez.



4. Seleccionamos **HTML_tables** y le damos a la opción de **Actualizar**, indicando el tiempo de refresco que nos interese.



5. Y ya tendríamos importada la tabla de la página web. Cada vez que en esa página web se **actualicen los datos**, nuestra tabla será actualizada automáticamente.

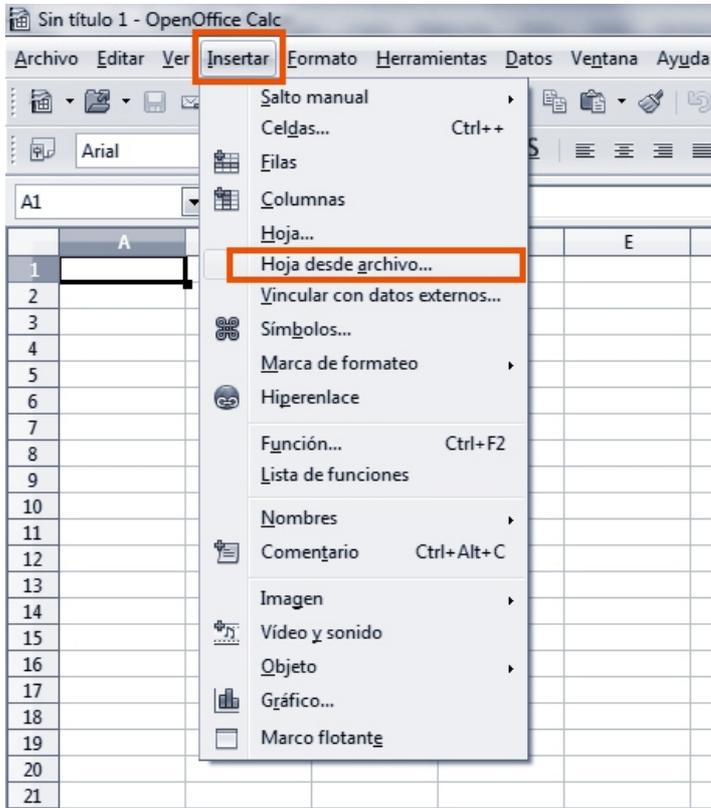


	A	B	C	D	E	F	G
1	Candidatura	Votos	% válidos	% censo	% candidatura	Diputados	
2	ARTIDO POP	10866566	44,63%	30,37%	45,25%	186	
3	OCIALISTA	7003511	28,76%	19,57%	29,16%	110	
4	LOS VERDE	1686040	6,92%	4,71%	7,02%	11	
5	ROGRESO Y	1143225	4,7%	3,2%	4,76%	5	
6	NVERGÊNC	1015691	4,17%	2,84%	4,23%	16	
7	AIUR - AM	334498	1,37%	0,93%	1,39%	7	
8	ITZALEA-P	324317	1,33%	0,91%	1,35%	5	
9	SQUERRA	256985	1,06%	0,72%	1,07%	3	
10	QUO - EQU	216748	0,89%	0,61%	0,9%	0	
11	NACIONAL	184037	0,76%	0,51%	0,77%	2	
12	IÓN CANAR	143881	0,59%	0,4%	0,6%	2	
13	IVA-VERDS	125306	0,51%	0,35%	0,52%	1	
14	ISTA CONT	102144	0,42%	0,29%	0,43%	0	
15	RO DE CIUD	99473	0,41%	0,28%	0,41%	1	
16	AÑOS EN B	97673	0,4%	0,27%	0,41%	0	
17	TIDO ANDA	76999	0,32%	0,22%	0,32%	0	
18	FORMA PER	59949	0,25%	0,17%	0,25%	0	
19	REGIONALIS	44010	0,18%	0,12%	0,18%	0	
20	AI - GEROA	42415	0,17%	0,12%	0,18%	1	
21	UN MUNDO	27210	0,11%	0,08%	0,11%	0	
22	NISTA DE L	26254	0,11%	0,07%	0,11%	0	

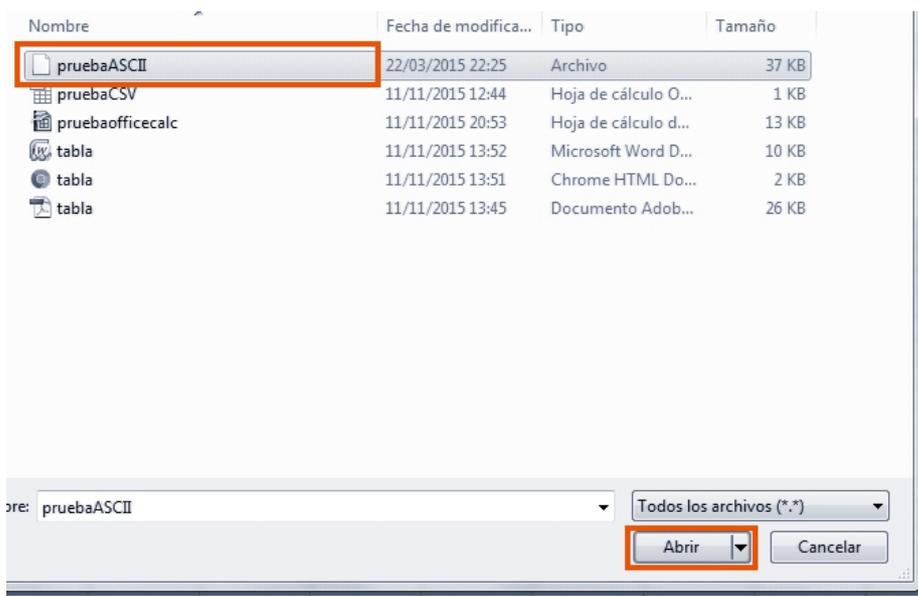
Actualización automática de datos (hoja desde archivo), abrir archivos ASCII (datos electorales)

Al igual que con la opción de Vincular con datos externos, podemos vincular una hoja de cálculo de OpenOffice Calc con un **archivo descargado** en nuestro ordenador. Cada vez que volvemos a descargar el archivo desde internet (porque la fuente lo haya actualizado) y reemplazemos el archivo anterior en el directorio correcto, nuestra hoja de cálculo se actualizará automáticamente.

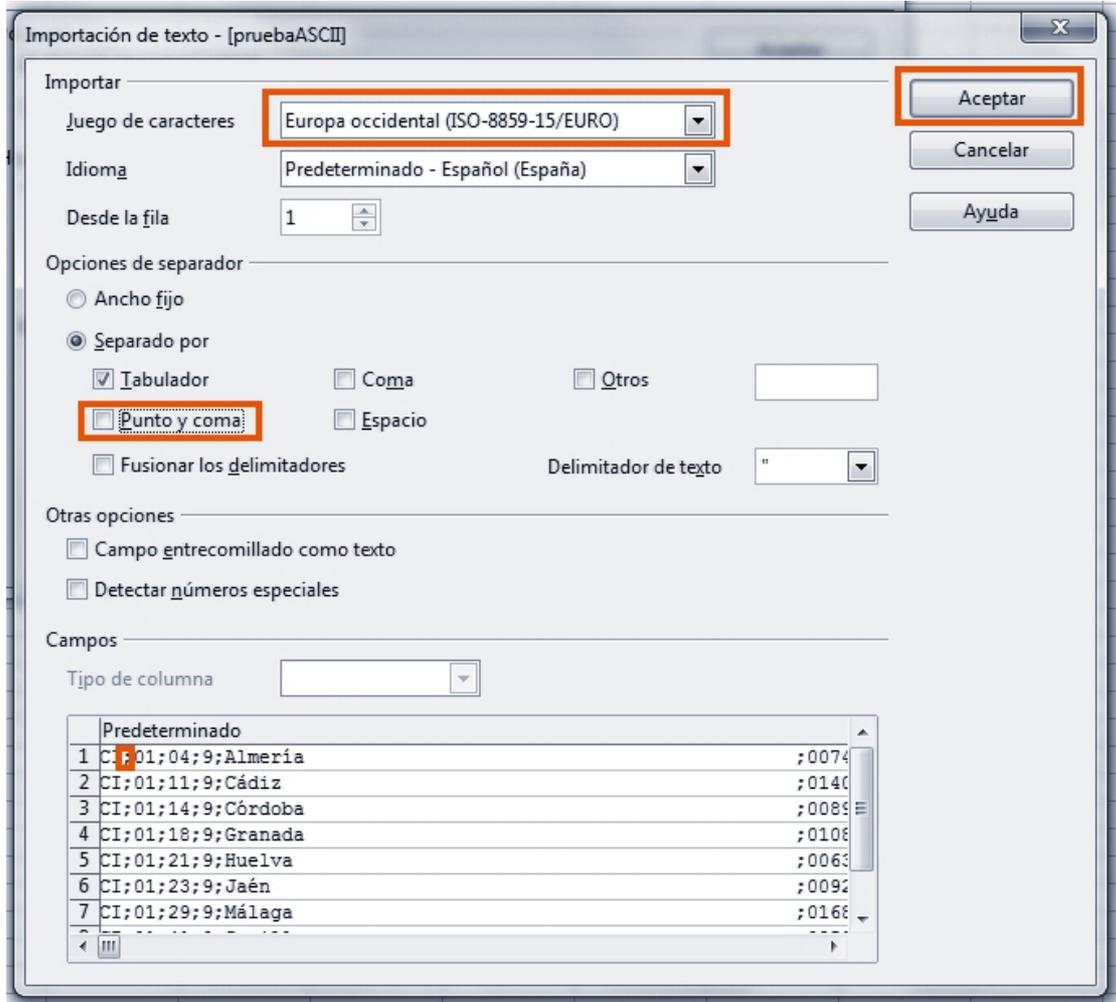
1.



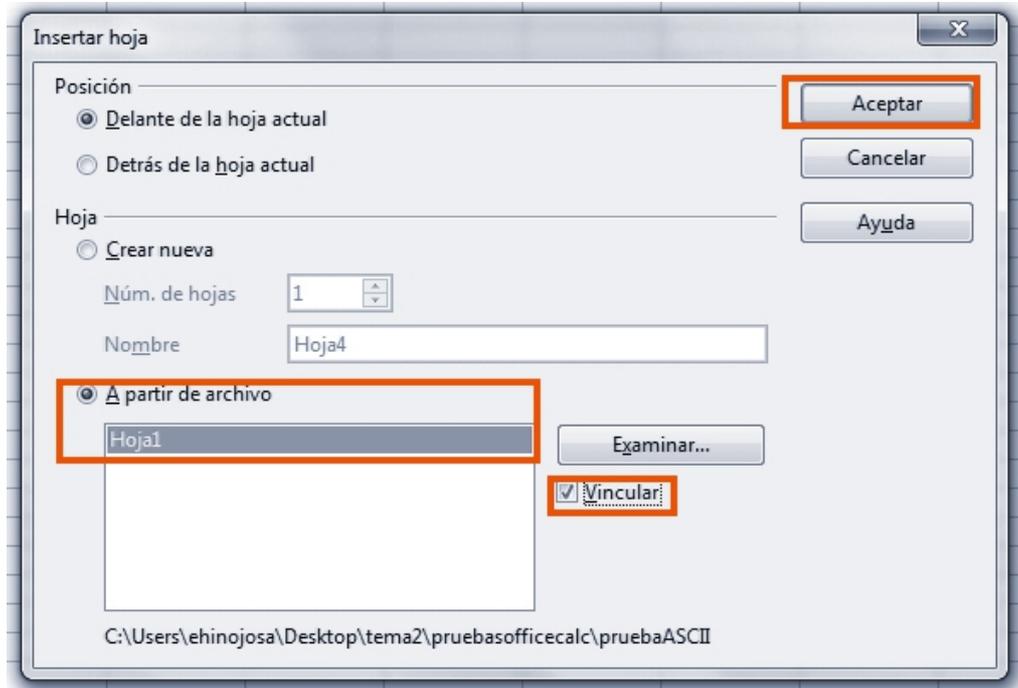
2. En este caso, abriremos un archivo **ASCII**, formato que usan las empresas que recopilan datos electorales para difundir los resultados.



3. Elegiremos el **Juego de caracteres** que nos funcione mejor, en este tipo de archivos, es mejor seleccionar el **ISO-8859-15/EURO**. También veremos qué **separadores de campo** usa el archivo original y lo seleccionaremos.



4. Escogeremos la opción **Vincular**.



5. Y ya tendríamos el archivo importado. Cada vez que reemplacemos este archivo en el mismo directorio donde está el original, todas las modificaciones de datos se actualizarán automáticamente en nuestra hoja de cálculo.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	CI	1	4	9	Almería	745	451037	449995	9977	272232	6050	177763	3950	3491	129	2146	79	12	20	PP
2	CI	1	11	9	Cádiz	1400	968858	966792	9979	571209	5908	395583	4092	9172	162	4598	80	15	23	PSOE-A
3	CI	1	14	9	Córdoba	897	637734	636882	9987	427687	6715	209195	3285	6936	164	5195	121	12	23	PSOE-A
4	CI	1	18	9	Granada	1083	704333	703629	9990	457159	6497	246470	3503	5244	116	4473	98	13	23	PSOE-A
5	CI	1	21	9	Huelva	634	389951	389211	9981	237638	6106	151573	3894	3452	147	3137	132	11	23	PSOE-A
6	CI	1	23	9	Jaén	920	524720	524720	10000	363625	6930	161095	3070	4232	118	4665	128	11	23	PSOE-A
7	CI	1	29	9	Málaga	1681	1109117	1109117	10000	677500	6108	431617	3892	8422	125	5706	84	17	23	PSOE-A
8	CI	1	41	9	Sevilla	2513	1501602	1501602	10000	1009444	6722	492158	3278	13839	139	11003	109	18	23	PSOE-A
9	CM	1	99	9	Andalucía	9873	6287352	6281948	9991	4016494	6394	2265454	3606	54788	138	40923	102	109	23	PSOE-A
10																				
11																				



Vincular una tabla editada a otra tabla sin editar, pero actualizable

1. Damos **doble click** a la pestaña de la **primera hoja** del documento.

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

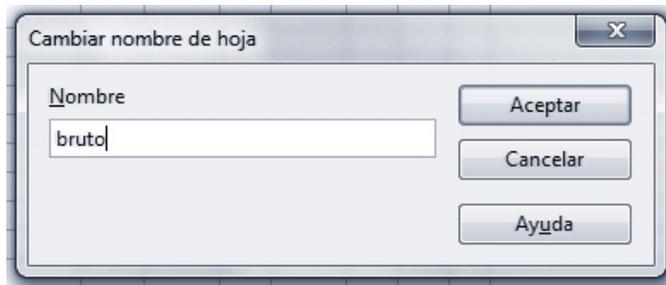
A1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	CI	1	4	9	Almería	745	451037	449995	9977	272232	6050	177763	3950	3491
2	CI	1	11	9	Cádiz	1400	968858	966792	9979	571209	5908	395583	4092	9172
3	CI	1	14	9	Córdoba	897	637734	636882	9987	427687	6715	209195	3285	6936
4	CI	1	18	9	Granada	1083	704333	703629	9990	457159	6497	246470	3503	5244
5	CI	1	21	9	Huelva	634	389951	389211	9981	237638	6106	151573	3894	3452
6	CI	1	23	9	Jaén	920	524720	524720	10000	363625	6930	161095	3070	4232
7	CI	1	29	9	Málaga	1681	1109117	1109117	10000	677500	6108	431617	3892	8422
8	CI	1	41	9	Sevilla	2513	1501602	1501602	10000	1009444	6722	492158	3278	13839
9	CM	1	99	9	Andalucía	9873	6287352	6281948	9991	4016494	6394	2265454	3606	54788
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42														
43														
44														
45														
46														
47														
48														

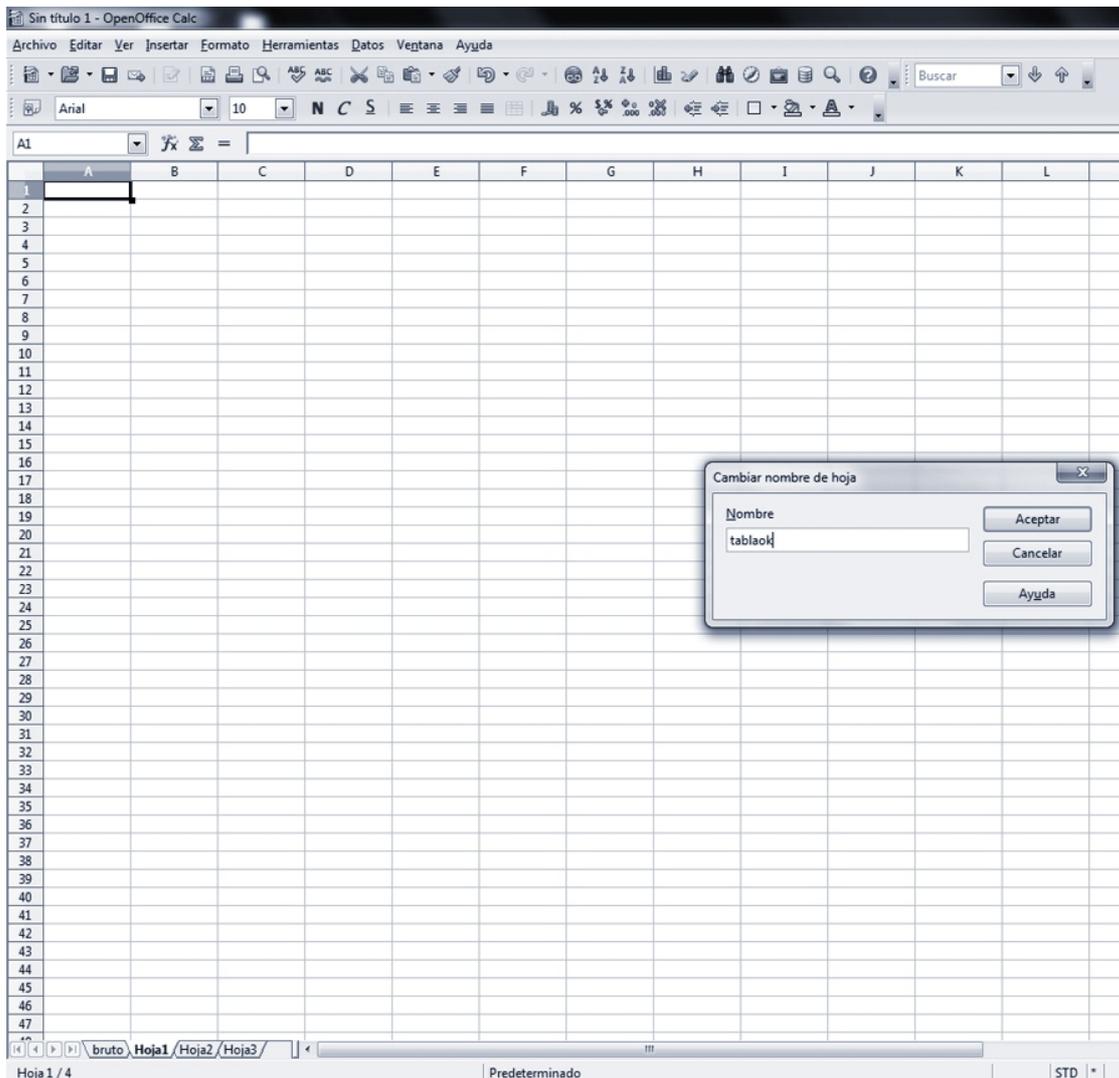
Hoja1_2 / Hoja1 / Hoja2 / Hoja3 /

Hoja 1 / 4 Predeterminado

2. Cambiamos el nombre. Por ejemplo, 'bruto', para que sepamos que ahí está el 'data set' sin editar.



3. Hacemos lo mismo con la segunda hoja, a la que podemos llamar 'tablaok', por ejemplo.





4. Así nos quedarían las hojas.



5. Vamos a tomar como fuente de datos la **columna K** de la hoja 'bruto'.

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

K1:K1048576 \sum = 6050

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	CI	1	4	9	Almería	745	451037	449995	9977	2722	26050	177763	395
2	CI	1	11	9	Cádiz	1400	968858	966792	9979	5712	95908	895583	409
3	CI	1	14	9	Córdoba	897	637734	636882	9987	4276	76715	209195	328
4	CI	1	18	9	Granada	1083	704333	703629	9990	4571	96497	246470	350
5	CI	1	21	9	Huelva	634	389951	389211	9981	2376	86106	151573	389
6	CI	1	23	9	Jaén	920	524720	524720	10000	3636	56930	161095	307
7	CI	1	29	9	Málaga	1681	1109117	1109117	10000	6775	06108	131617	389
8	CI	1	41	9	Sevilla	2513	1501602	1501602	10000	10094	46722	192158	327
9	CM	1	99	9	Andalucía	9873	6287352	6281948	9991	40164	46394	2265454	360
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													
34													
35													
36													
37													
38													
39													
40													
41													
42													
43													
44													
45													
46													
47													

bruto / tablaok

Hoja 1 / 2 Predeterminado



6. Primero, copiamos la lista de **provincias** de la hoja 'bruto'...

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

E1:E9 = Andalucía

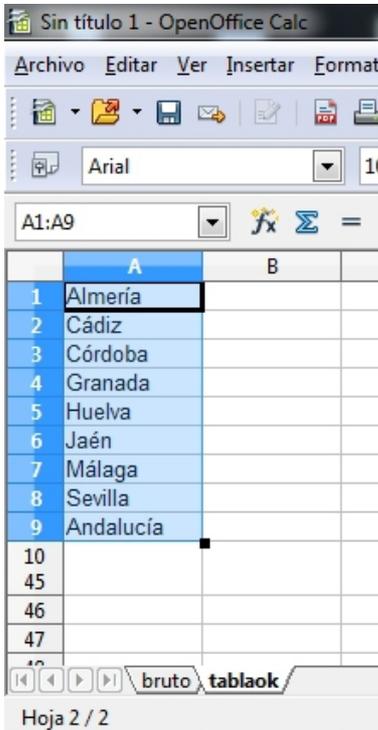
	A	B	C	D	F	G	H	I	J	K	L	M
1	CI	1	4	Almería	745	451037	449995	9977	272232	6050	177763	398
2	CI	1	11	Cádiz	1400	968858	966792	9979	571209	5908	395583	408
3	CI	1	14	Córdoba	897	637734	636882	9987	427687	6715	209195	328
4	CI	1	18	Granada	1083	704333	703629	9990	457159	6497	246470	350
5	CI	1	21	Huelva	634	389951	389211	9981	237638	6106	151573	388
6	CI	1	23	Jaén	920	524720	524720	10000	363625	6930	161095	308
7	CI	1	29	Málaga	1681	1109117	1109117	10000	677500	6108	431617	388
8	CI	1	41	Sevilla	2513	1501602	1501602	10000	1009444	6722	492158	328
9	CM	1	99	Andalucía	9873	6287352	6281948	9991	4016494	6394	2265454	368
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												
37												
38												
39												
40												
41												
42												
43												
44												
45												
46												
47												

bruto / tablaok /

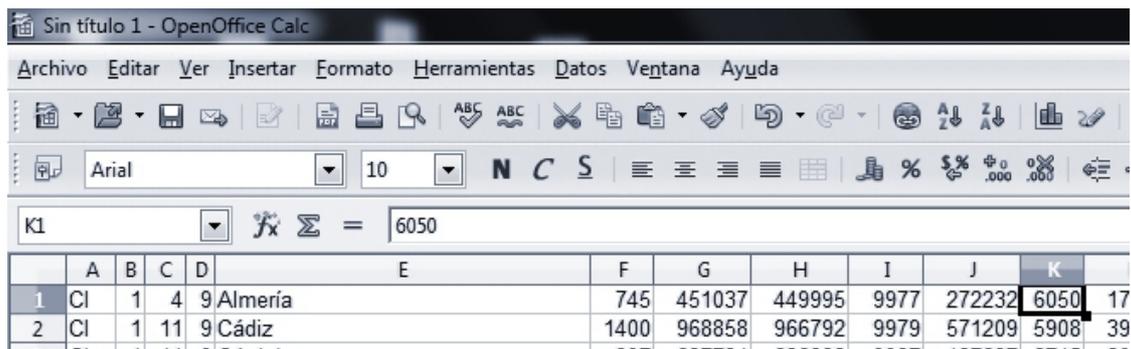
Hoja 1 / 2 Predeterminado



7. Y las pegamos en la primera columna de la hoja 'tablaok'.

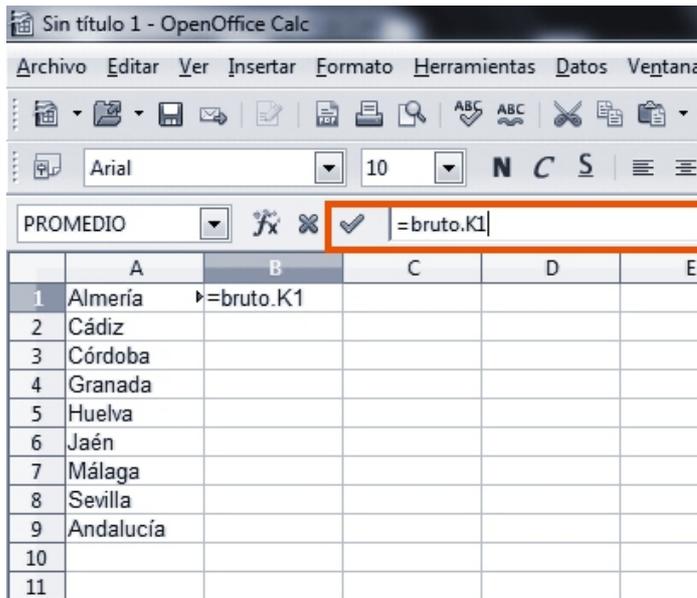


8. Después, volvemos a la hoja 'bruto' para tomar la referencia del primer dato de la lista que nos interesa, que en este caso sería **K1**.

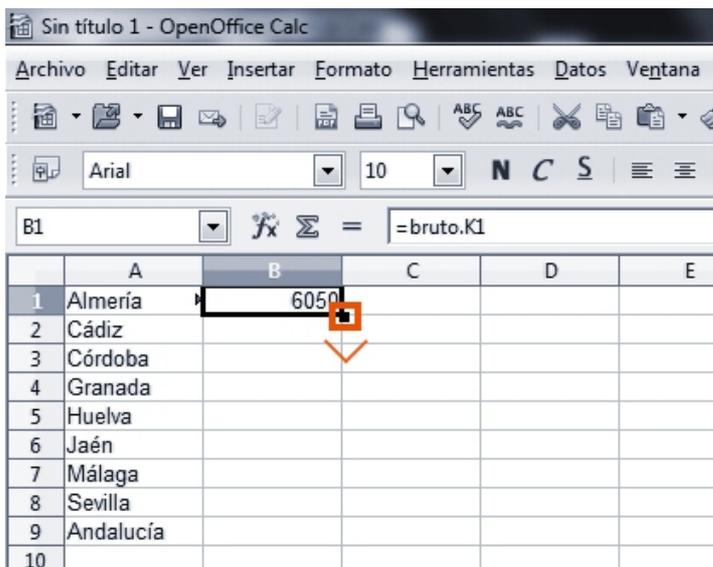




9. Nos vamos de nuevo a la hoja 'tablaok' y, en la primera celda de la segunda columna, escribimos la siguiente fórmula: **=bruto.K1** donde primero le decimos la hoja de origen, luego debe ir un punto y, a continuación, la coordenada del dato que queremos recoger de esa hoja de origen.



10. Aplicamos la fórmula al resto de celdas (pulsando y arrastrando hacia abajo).





11.

	A	B	C	D
1	Almería	6050		
2	Cádiz	5908		
3	Córdoba	6715		
4	Granada	6497		
5	Huelva	6106		
6	Jaén	6930		
7	Málaga	6108		
8	Sevilla	6722		
9	Andalucía	6394		
10				

12. También podemos aplicar **operaciones** a los datos de la hoja 'bruto' a los que llamemos.

	A	B	C	D
1	Almería	=bruto.K1/100		
2	Cádiz	5908		
3	Córdoba	6715		
4	Granada	6497		
5	Huelva	6106		
6	Jaén	6930		
7	Málaga	6108		
8	Sevilla	6722		
9	Andalucía	6394		
10				
11				

13. Y aplicar esa operación al resto de datos de la columna.

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas

Arial 10

B1:B9 =bruto

	A	B	C
1	Almería	60,5	
2	Cádiz	59,08	
3	Córdoba	67,15	
4	Granada	64,97	
5	Huelva	61,06	
6	Jaén	69,3	
7	Málaga	61,08	
8	Sevilla	67,22	
9	Andalucía	63,94	
10			
11			

14. E incluso, aplicarle un **formato nuevo** a los datos en esta hoja 'tablaok'.

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10

B1:B9 =bruto.K1/100

	A	B	C	D	E
1	Almería	60,5			
2	Cádiz	59,08			
3	Córdoba	67,15			
4	Granada	64,97			
5	Huelva	61,06			
6	Jaén	69,3			
7	Málaga	61,08			
8	Sevilla	67,22			
9	Andalucía	63,94			
10					
11					
12					
13					
14					
15					
16					
17					

- Formato predeterminado
- Formatar celdas...**
- Insertar...
- Eliminar...
- Eliminar contenidos...
- Insertar comentario
- Cortar
- Copiar
- Pegar
- Pegado especial...



15. Si en la hoja original, en este caso 'bruto', uno de los datos se modifica...

Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

K1 = 6050

	A	B	C	D	E	F	G	H	I	J	K	L
1	Cl	1	4	9	Almería	745	451037	449995	9977	272232	6050	1777
2	Cl	1	11	9	Cádiz	1400	968858	966792	9979	571209	5908	3955
3	Cl	1	14	9	Córdoba	897	637734	636882	9987	427687	6715	2091
4	Cl	1	18	9	Granada	1083	704333	703629	9990	457159	6497	2464
5	Cl	1	21	9	Huelva	634	389951	389211	9981	237638	6106	1515
6	Cl	1	23	9	Jaén	920	524720	524720	10000	363625	6930	1610
7	Cl	1	29	9	Málaga	1681	1109117	1109117	10000	677500	6108	4316
8	Cl	1	41	9	Sevilla	2513	1501602	1501602	10000	1009444	6722	4921
9	CM	1	99	9	Andalucía	9873	6287352	6281948	9991	4016494	6394	22654
47												

bruto / tablaok / Predeterminado

Hoja 1 / 2

Sin título 1 - OpenOffice Calc

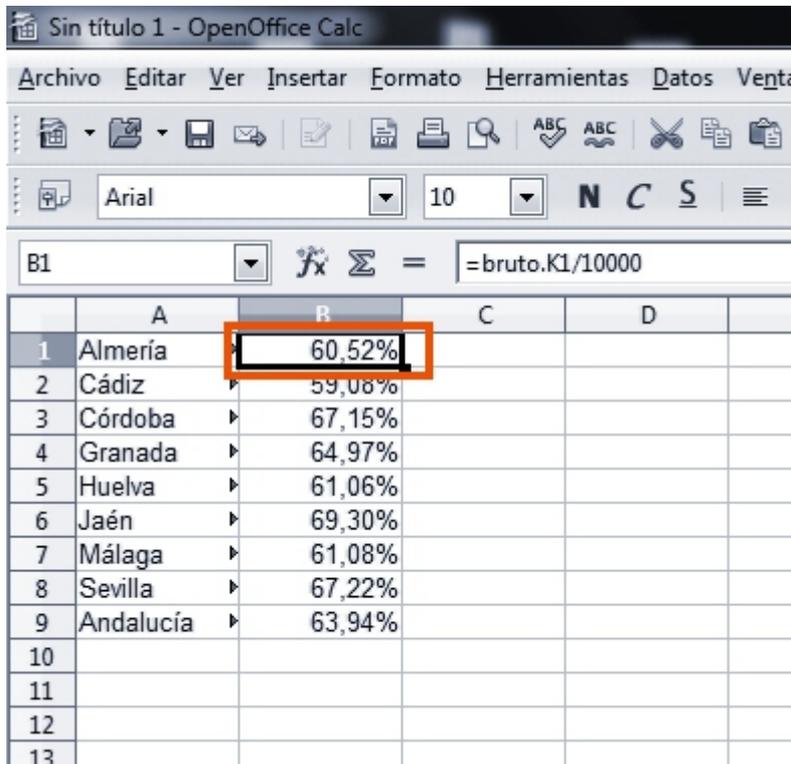
Archivo Editar Ver Insertar Formato Herramientas Datos Ventana Ayuda

Arial 10 N C S

K1 = 6052

	A	B	C	D	E	F	G	H	I	J	K	L
1	Cl	1	4	9	Almería	745	451037	449995	9977	272232	6052	177763
2	Cl	1	11	9	Cádiz	1400	968858	966792	9979	571209	5908	395583
3	Cl	1	14	9	Córdoba	897	637734	636882	9987	427687	6715	209195
4	Cl	1	18	9	Granada	1083	704333	703629	9990	457159	6497	246470
5	Cl	1	21	9	Huelva	634	389951	389211	9981	237638	6106	151573
6	Cl	1	23	9	Jaén	920	524720	524720	10000	363625	6930	161095
7	Cl	1	29	9	Málaga	1681	1109117	1109117	10000	677500	6108	431617
8	Cl	1	41	9	Sevilla	2513	1501602	1501602	10000	1009444	6722	492158
9	CM	1	99	9	Andalucía	9873	6287352	6281948	9991	4016494	6394	2265454
10												
11												

16. Esa modificación **se actualizará automáticamente** en la hoja 'tablaok', pero conservando la edición que hemos aplicado con respecto a las operaciones, fórmulas o formatos de los datos.



Sin título 1 - OpenOffice Calc

Archivo Editar Ver Insertar Formato Herramientas Datos Venta

Arial 10 N C S

B1 =bruto.K1/10000

	A	B	C	D
1	Almería	60,52%		
2	Cádiz	59,08%		
3	Córdoba	67,15%		
4	Granada	64,97%		
5	Huelva	61,06%		
6	Jaén	69,30%		
7	Málaga	61,08%		
8	Sevilla	67,22%		
9	Andalucía	63,94%		
10				
11				
12				
13				

IDEAS CLAVE

- **Tabula PDF** es un programa gratuito que nos permite extraer de documentos PDF (tienen que ser PDF con texto real) las tablas que estos contienen y exportar a un archivo de hoja de cálculo.
- Los archivos de tablas más comunes son **CSV** (comas como separadores de campo) y **TSV** (tabuladores como separadores de campo).
- **Open Refine** es un programa gratuito donde podemos refinar los datos de un archivo de hoja de cálculo (por ejemplo, CSV). Podemos editar un gran volumen de datos de una manera rápida y efectiva. El resultado puede ser otro archivo CSV.
- **OpenOffice Calc** es parecido al programa Microsoft Excel, pero gratuito. En él podemos abrir diferentes archivos de datos, realizar operaciones con esos datos, cambiar sus formatos e incluso crear vinculaciones con tablas alojadas en una web o con archivos ASCII.



referencias

TABULA PDF

<http://tabula.technology/>

OPEN REFINE

<http://openrefine.org/>

[Tutorial de Carolina Cristanchi sobre Open Refine, en español](#)

OPENOFFICE CALC

<https://www.openoffice.org/es/>

[Wiki Open Office](#)

[Versión más completa](#)

[Expresiones regulares Open Office](#)



tareas

- Visualizar los **videotutoriales** de Tabula PDF, Open Refine y OpenOffice Calc.
- **Probar** los programas tratados en este tema. Se pueden seguir los ejemplos aquí descritos o realizar otro tipo de ejercicios libremente.