

## PRÁCTICA DE LABORATORIO Nº1

La última versión de este documento está disponible en

<ftp://carleos.epv.uniovi.es/peritos/217>

### 1. OBJETIVOS GENERALES

Con las prácticas de laboratorio se pretende poner al alcance de los alumnos las prestaciones básicas de la hoja de cálculo CALC del paquete ofimático OPENOFFICE.ORG para el análisis de datos estadísticos y el cálculo de probabilidades.

Mediante las guías de las prácticas se intenta que el alumno sea capaz de seguir por sí solo el desarrollo de cada práctica, de forma que el profesor le sirva de apoyo cuando se encuentre con cualquier dificultad. Las guías no son, ni pretenden ser, una ayuda o introducción a CALC, ni siquiera un texto de estadística con CALC. Por lo tanto, para cualquier ampliación acerca del funcionamiento de CALC o de sus funciones estadísticas, remitimos a los textos que en la bibliografía se citan al respecto.

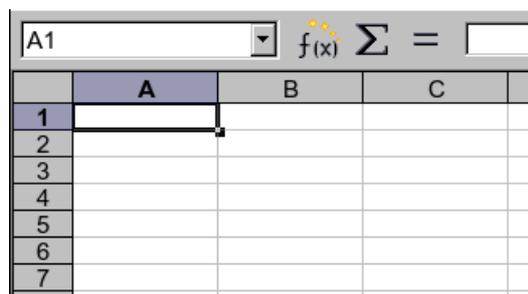
### 2. OBJETIVOS DE LA PRÁCTICA Nº1

Con la primera práctica se pretende enseñar al alumno las nociones básicas de trabajo con CALC y el uso de las funciones estadísticas de CALC para la descripción de una variable y para el trabajo con variables condicionadas.

### 3. ENTORNO DE TRABAJO EN CALC

CONCEPTOS: celda; referencia relativa; copiar y pegar; fórmula.

Los documentos de CALC (cuya extensión es *ods*, de Open Document Spreadsheet, un formato público y abierto) se denominan *libros* y pueden constar de varias *hojas (de cálculo)*. Se puede navegar por las hojas del libro con los botones y solapas  *Hoja2* / *Hoja3*, que aparecen en la parte baja de la ventana. Los nombres de las hojas se pueden cambiar pinchando dos veces sobre el nombre actual, y el resto de opciones de hoja se obtiene con el botón derecho del ratón. Las hojas están divididas en celdas o casillas, a las que se referencia mediante las coordenadas de la columna y de la fila. Cada columna se indica mediante una letra y cada renglón mediante un número. En cada una de las hojas veremos la *celda activa*, enmarcada por un recuadro más oscuro, que indica la celda donde actualmente se están introduciendo los datos. La referencia de la celda activa se muestra en un cuadro en la parte superior de la hoja, como se puede apreciar en la ilustración 1, donde la celda activa es la A1 (fila 1, columna A).



*Ilustración 1*

*Grosso modo*, en las celdas se puede introducir contenido textual (etiquetas) o bien contenido matemático (una *fórmula*). El usuario puede definir fórmulas mediante operaciones entre números, referencias a otras celdas y funciones predefinidas de CALC. Por omisión, CALC alinea las etiquetas por la izquierda y las fórmulas por la derecha.

Tenga en cuenta que para forzar que un contenido (por ejemplo: 001) sea almacenado tal cual (y no sea interpretado como un número) se ha de introducir previamente un apóstrofo: '001. Para forzar la interpretación matemática de un contenido (por ejemplo: 3+5) se ha de teclear previamente un signo igual: =3+5.

EJERCICIO 1: Intente reproducir la siguiente tabla en CALC:

Artículo	Precio €/unidad	Unidades	Total	
001	0,04	10	0,40	
002	0,15	4	0,60	
010	1,20	5	6,00	
			7,00	=0,4+0,6+6

De hecho, para introducir cualquier fórmula manualmente, sitúese en la celda en la cual se quiere el resultado y escriba un signo igual (=); tras escribir la fórmula se pulsa *Intro* y en la celda aparecerá el valor de la fórmula, mientras que en la *barra de fórmulas* (encima de la rejilla de las celdas) aparecerá la expresión que hemos tecleado.

Para escribir las fórmulas se pueden utilizar los operadores aritméticos habituales: suma (+), resta (-), multiplicación (\*), división (/) y potencia (^); también se pueden utilizar paréntesis para agrupar expresiones; además, CALC ofrece una colección de funciones predefinidas, de las que hablaremos más adelante. Para utilizar en la fórmula el valor de otra celda, basta teclear la referencia de la celda en la fórmula (o situarse sobre la celda con el ratón mientras se compone la fórmula). En la ilustración 2, se ve cómo se realiza una operación entre varias celdas, y en la ilustración 3 se observa el resultado final.

	A	B	C	D
1	45			
2	4			
3	2	=A1+A2*A3		
4				

*Ilustración 2*

	A	B	C	D
1	45			
2	4			
3	2	53		
4				

*Ilustración 3*

EJERCICIO 2: Escriba la “tabla de multiplicar” del número 13 en una hoja de CALC. (Empiece escribiendo los números del 1 al 10 en una columna.)

## 4. FUNCIONES EN CALC

CONCEPTOS: rango de celdas; referencia absoluta; desbordamiento; pegar sólo número

Como señalábamos arriba, CALC dispone de una gran cantidad de funciones predefinidas. Nos referiremos más adelante a las específicamente estadísticas y a algunas de las funciones matemáticas como LN (logaritmo neperiano) o EXP (exponencial). Cada fórmula tiene sus propias características (respecto a los datos que necesita para ejecutarse) así que es recomendable consultar la propia ayuda de CALC para utilizar las fórmulas correctamente.

Para usar cierta función, se puede teclear su nombre directamente en la fórmula, o con el ratón se puede pinchar bien en *Insertar > Función* o bien sobre el icono  y, a continuación, seguir las instrucciones de la función. En general, se nos pedirá el rango de datos (celda o celdas) sobre los cuales se va a calcular la función y el resto de valores necesarios para cada función.

Veamos un ejemplo, utilizando la función SUMA para calcular la suma de unas celdas. Sumaremos las celdas desde A1 hasta la A6, así que nos situamos donde deseemos el resultado y pinchamos en el botón de insertar función, seleccionando a continuación la función SUMA de la lista que se nos ofrece (las funciones están clasificadas por categorías; ilustración 4).

El siguiente paso es, ya dentro de la función, seleccionar el rango de celdas que queremos sumar, indicándolo en la casilla correspondiente de la función. Cualquier rango de celdas contiguas formando un rectángulo se denota mediante las coordenadas de la primera celda seguidas de dos puntos (:) y las coordenadas de la última celda; así, las celdas A1, A2, A3, A4, A5 y A6 serán A1:A6, como se aprecia en la ilustración 5.

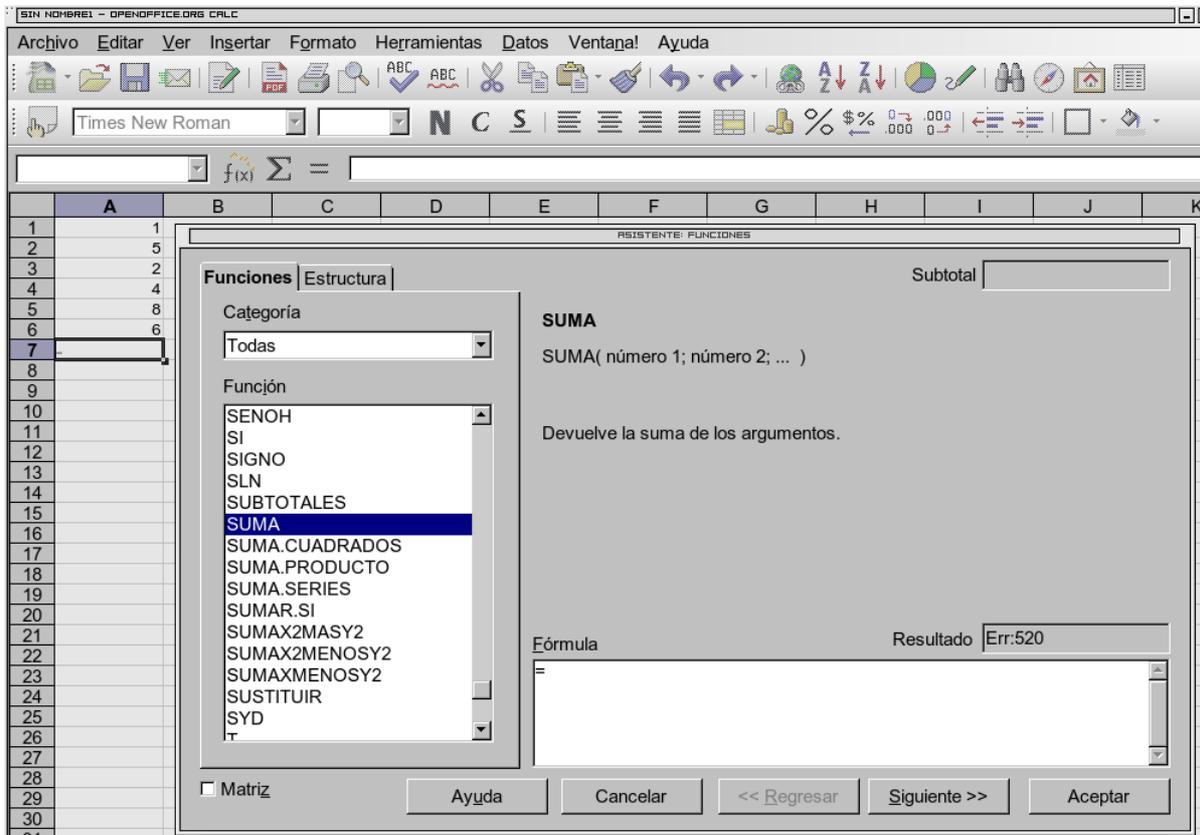


Ilustración 4

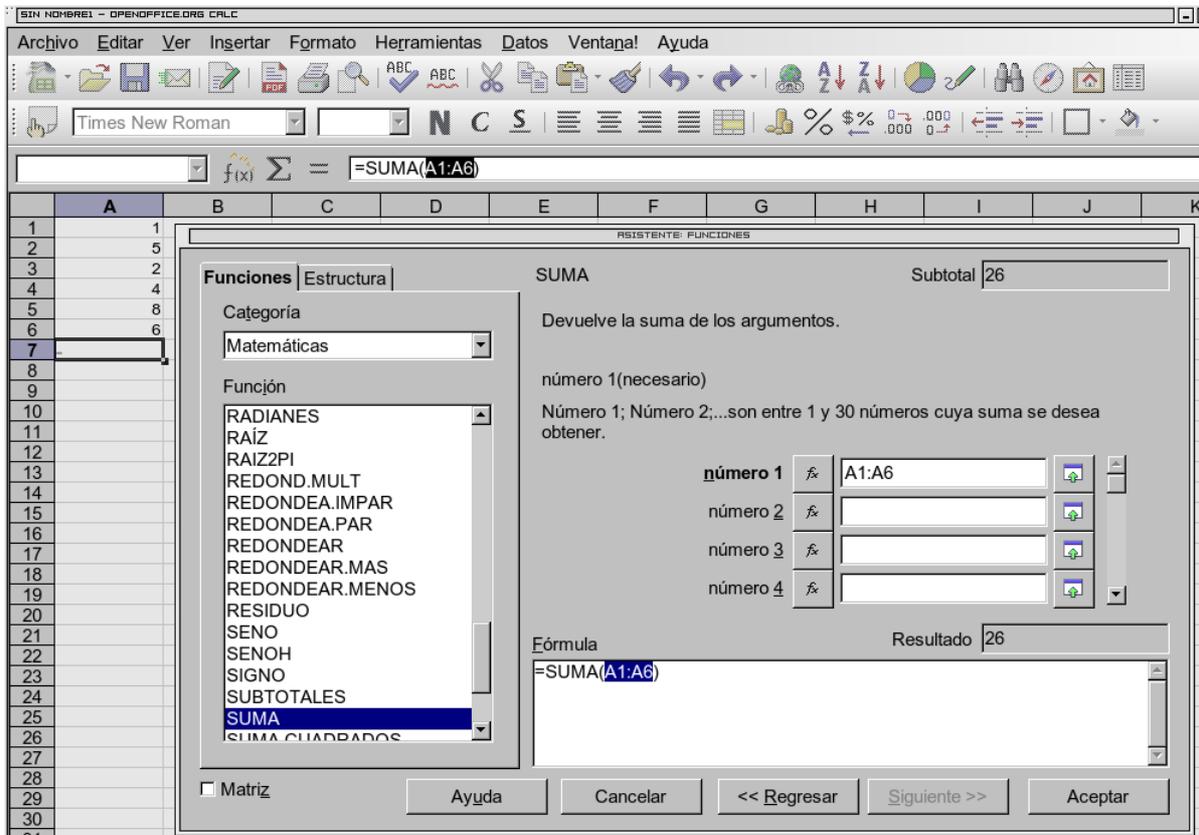


Ilustración 5

Finalmente, ya sólo queda aceptar y la celda nos mostrará el resultado de la fórmula, 26 en este caso (ilustración 6).

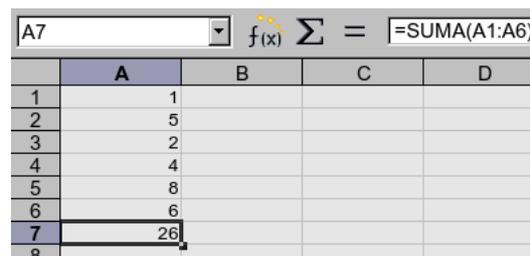


Ilustración 6

**EJERCICIO 3:** Utilice la función PRODUCTO sobre un rango de celdas con los números 1, 5, 2, 4, 8 y 6, para calcular el resultado de multiplicar tales números.

Si el resultado de una fórmula es un número con demasiadas cifras para la anchura de la columna, aparecerán tres almohadillas (###) para indicar el desbordamiento. No se trata de ningún error: basta con ensanchar la columna manualmente o con la opción de menú *Formato > Columna > Anchura óptima*.

Si copiamos una celda que incluye una fórmula, al pegarla reproduciremos no su valor sino la expresión de la fórmula. Esto resulta muy útil cuando se reproduce la misma operación sobre un conjunto de datos. Por ejemplo, si queremos calcular el logaritmo neperiano de cada uno de los valores de la columna A, basta con realizar la operación para el primer valor (ilustración 7) y luego simplemente copiar y pegar la misma (lo que equivale a *extenderla* pinchando en el recuadro situado en la esquina inferior derecha de la celda).

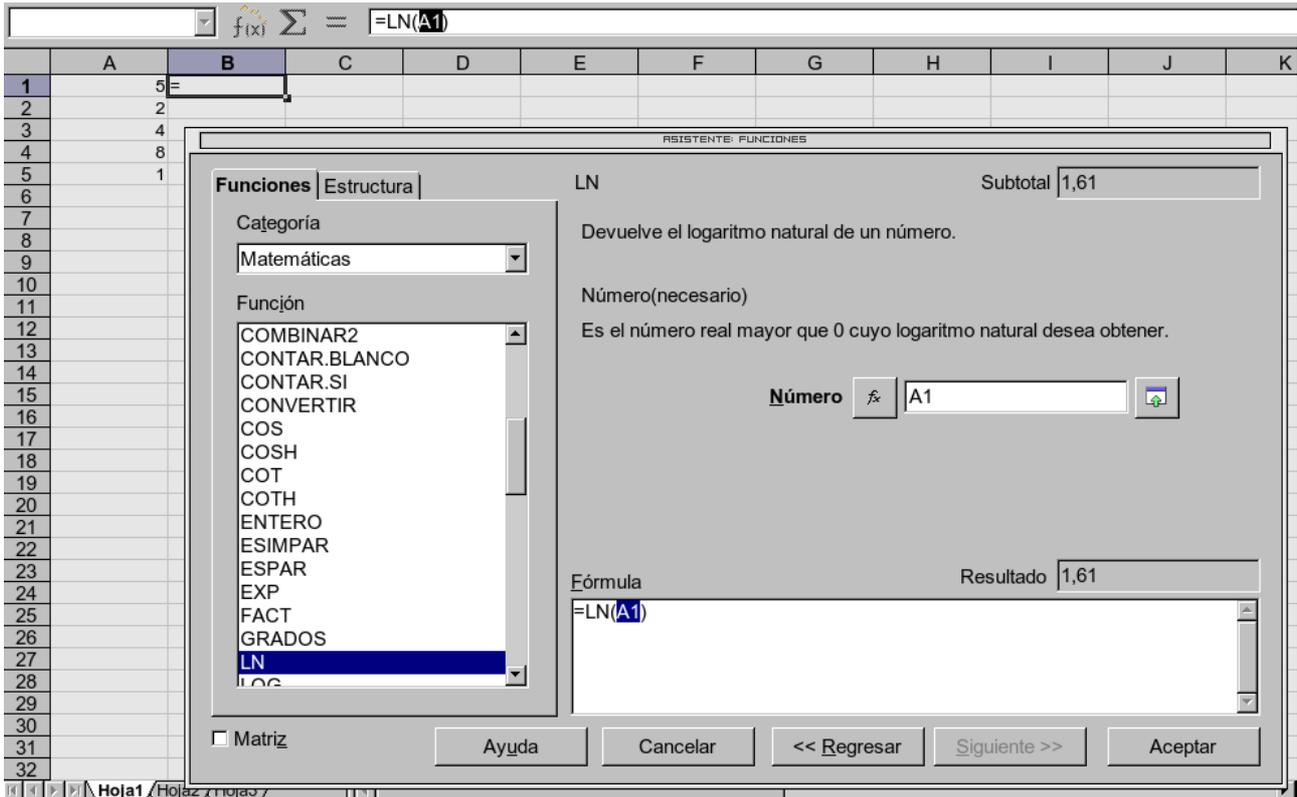


Ilustración 7

Así, la ilustración 8 muestra la celda copiada; la 9, el rango de celdas sobre el que se producirá el pegado; y la 10, el resultado tras pegar. Nótese que no se pega el valor de B1, sino la fórmula. Nótese también que la fórmula pegada modifica las referencias a las celdas incluidas en la fórmula; en el ejemplo, si la fórmula original en B1 consiste en la expresión LN(A1), entonces la fórmula copiada en B2 consistirá en la expresión LN(A2).

	A	B
1	5	1,61
2	2	
3	4	
4	8	
5	1	
6		

Ilustración 8

	A	B
1	5	1,61
2	2	
3	4	
4	8	
5	1	
6		

Ilustración 9

	A	B
1	5	1,61
2	2	0,69
3	4	1,39
4	8	2,08
5	1	0
6		

Ilustración 10



EJERCICIO 4: Escriba los números 256, 196, 169, 525 y 625 en una columna (o fila) de celdas. Mediante la función RAÍZ, calcule la raíz cuadrada de cada número.

La modificación de las referencias permite que, al copiar y pegar una fórmula, obtengamos en general el resultado esperado, intuitivo. Sin embargo, en ocasiones necesitamos que cierta referencia a una celda permanezca constante al desplazar la fórmula; en tal caso, dicha referencia ha de hacerse de modo “absoluto” (es decir, como \$A\$1 en vez de A1).

Por ejemplo, véanse las ilustraciones 11 a 16. Se trata de incrementar una serie de precios mediante un porcentaje de IVA; dicho porcentaje está guardado en la celda C7. En las ilustraciones 11 a 13 se muestra el resultado (incorrecto) de copiar y pegar la fórmula original tal como describimos; véase que en D5 la referencia a C7 se convierte en C9. En las ilustraciones 14 a 16 se muestra el resultado correcto, al sustituir la referencia relativa C7 por la referencia absoluta \$C\$7, que no se modifica al copiar y pegar la fórmula.

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			=A3*(1+C7/100)
4	175			
5	200			
6				
7			IVA = 16%	

Ilustración 11

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			116
4	175			175
5	200			200
6				
7			IVA = 16%	

Ilustración 12

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			116
4	175			175
5	200			=A5*(1+C9/100)
6				
7			IVA = 16%	
8				
9				

Ilustración 13

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			=A3*(1+\$C\$7/100)
4	175			
5	200			
6				
7			IVA = 16%	

Ilustración 14

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			116
4	175			203
5	200			232
6				
7			IVA = 16%	

Ilustración 15

	A	B	C	D
1	Precio sin IVA			Precio con IVA
2				
3	100			116
4	175			203
5	200			=A5*(1+\$C\$7/100)
6				
7			IVA = 16%	

Ilustración 16

También existen las referencias parcialmente absolutas (como \$A1 y A\$1) que permiten fijar solamente la columna (\$A1) o solamente la fila (A\$1).

EJERCICIO 5:

- a) En una hoja en blanco, escriba en la columna A los valores 1100, 1200, 1875, 595 y 725, que representen precios de artículos. En la casilla D2 escriba el número 20, que representará un porcentaje de rebaja. En la columna B calcule los precios rebajados, aplicando el porcentaje contenido en la celda D2. Luego cambie el porcentaje de rebaja a un 25%.
- b) Escriba la tabla de multiplicar  $10 \times 10$  completa [nivel *avanzado*].

Si en alguna ocasión deseamos pegar el contenido de las celdas, y no las funciones correspondientes, después de *Editar > Copiar* debemos escoger *Editar > Pegado especial...*, desactivar las opciones *Pegar todo* y *Fórmulas* y activar la de *Números* (ilustración 17).

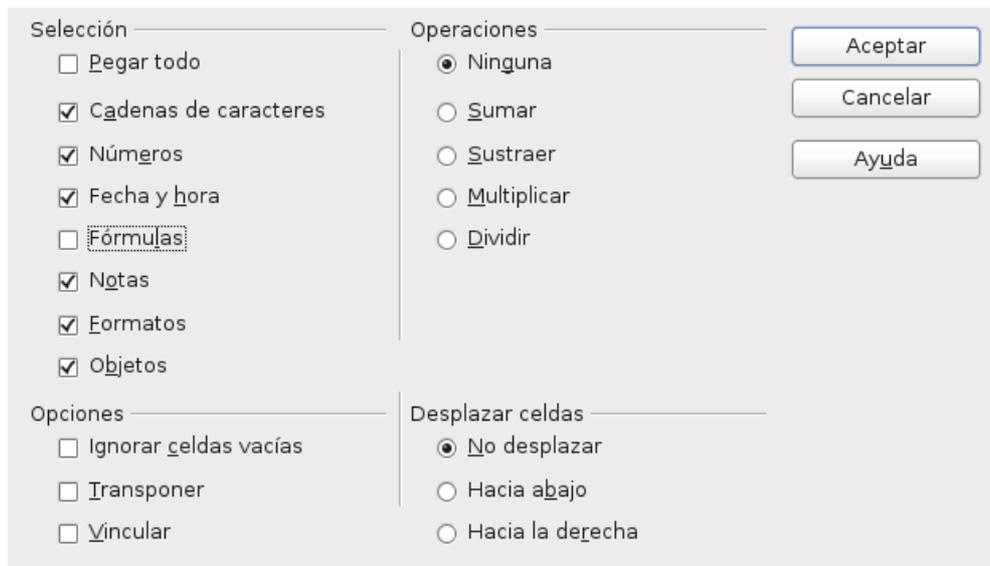


Ilustración 17

EJERCICIO 6: Copia sólo los valores de la columna B del ejercicio anterior (los precios rebajados) en una hoja nueva.

## 5. DESCRIPCIÓN DE UNA VARIABLE CON CALC

CONCEPTOS: promedio; mediana; moda; varp; desvestp; mín; máx; percentil; rango.percentil.

Para la primera práctica utilizaremos un fichero llamado `países.ods`, en el cual se recogen los datos relativos a una serie de variables medidas sobre un conjunto de países en el año 1995. Las variables estudiadas son las siguientes: población (medida en miles de habitantes),

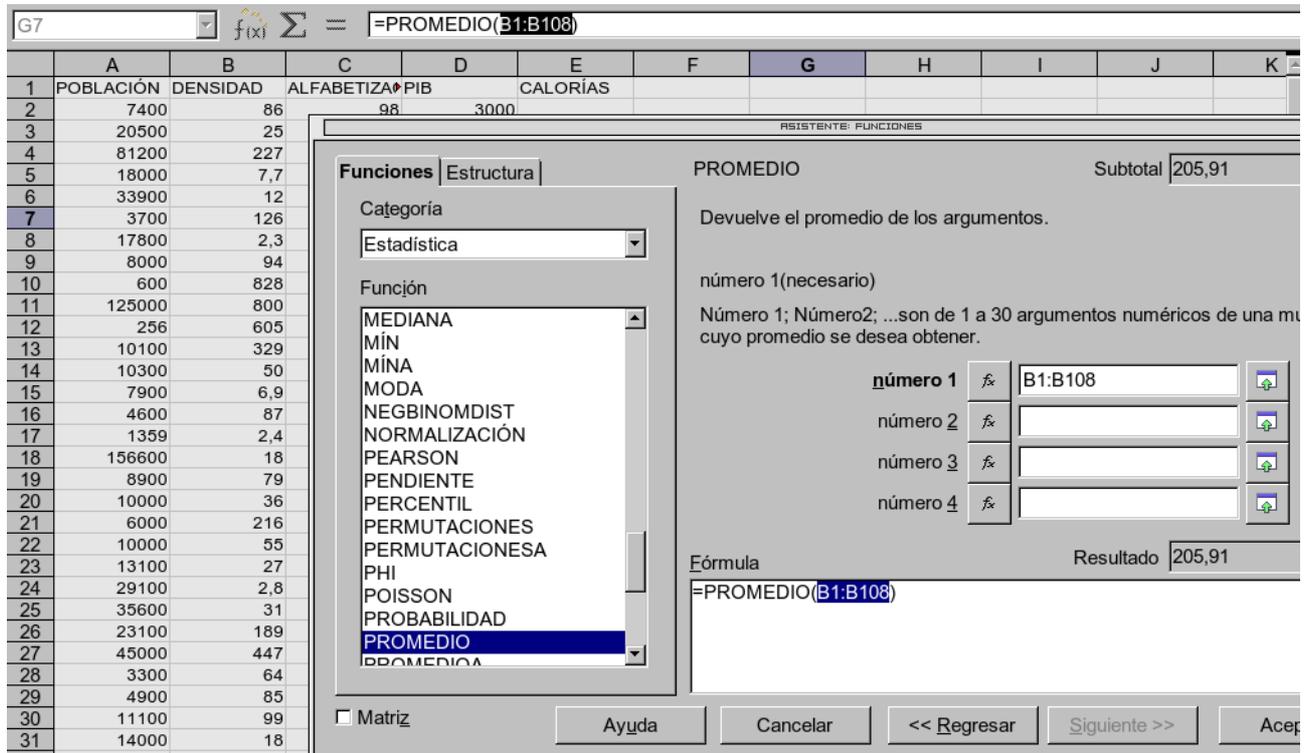
densidad de población (habitantes/km<sup>2</sup>), tasa de alfabetización (porcentaje de población alfabetizada sobre el total), P.I.B. per cápita (en dólares EE.UU.), ingesta diaria de calorías (cantidad media diaria de calorías por habitante). Las variables están ordenadas por columnas, con una fila inicial que contiene unos encabezamientos.

EJERCICIO 7: Carga el fichero y calcula:

1. La suma total de habitantes.
2. La suma total de habitantes analfabetos.
3. La proporción total de analfabetos.
4. La suma total de km<sup>2</sup> de las extensiones de los países.

## Medidas de centralización

En CALC disponemos de sendas funciones para calcular media, mediana y moda. La media se calcula utilizando la función PROMEDIO, la mediana mediante la función MEDIANA y la moda mediante la función MODA. En cualquiera de los tres casos, únicamente tenemos que indicar el rango de las celdas cuya medida queremos calcular. Así, por ejemplo, para calcular la media de la variable Densidad, tendremos simplemente que insertar en la celda deseada la función PROMEDIO e indicar en la casilla *Número 1* el rango de celdas correspondiente a los valores de la variable Densidad (se puede incluir o no la fila de encabezamiento, no produce ningún error en este tipo de funciones). La casilla *Número 2* (y las sucesivas que pueden aparecer) están disponibles por si los datos sobre los que queremos aplicar la función están fraccionados en varias partes de la hoja de cálculo. El resultado de estas operaciones, así como el valor final de la media (205,91) se recoge en la ilustración 18. Análogamente, se procede con la mediana y con la moda.



	A	B	C	D	E	F	G	H	I	J	K
1	POBLACIÓN	DENSIDAD	ALFABETIZACIÓN	PIB	CALORÍAS						
2	7400	86	98	3000							
3	20500	25									
4	81200	227									
5	18000	7,7									
6	33900	12									
7	3700	126									
8	17800	2,3									
9	8000	94									
10	600	828									
11	125000	800									
12	256	605									
13	10100	329									
14	10300	50									
15	7900	6,9									
16	4600	87									
17	1359	2,4									
18	156600	18									
19	8900	79									
20	10000	36									
21	6000	216									
22	10000	55									
23	13100	27									
24	29100	2,8									
25	35600	31									
26	23100	189									
27	45000	447									
28	3300	64									
29	4900	85									
30	11100	99									
31	14000	18									

Ilustración 18

EJERCICIO 8:

Halla la población y densidad de los hipotéticos *país medio* y *país mediano*.  
 ¿Cómo es España, en comparación?  
 ¿Tiene sentido calcular la moda de esas variables?

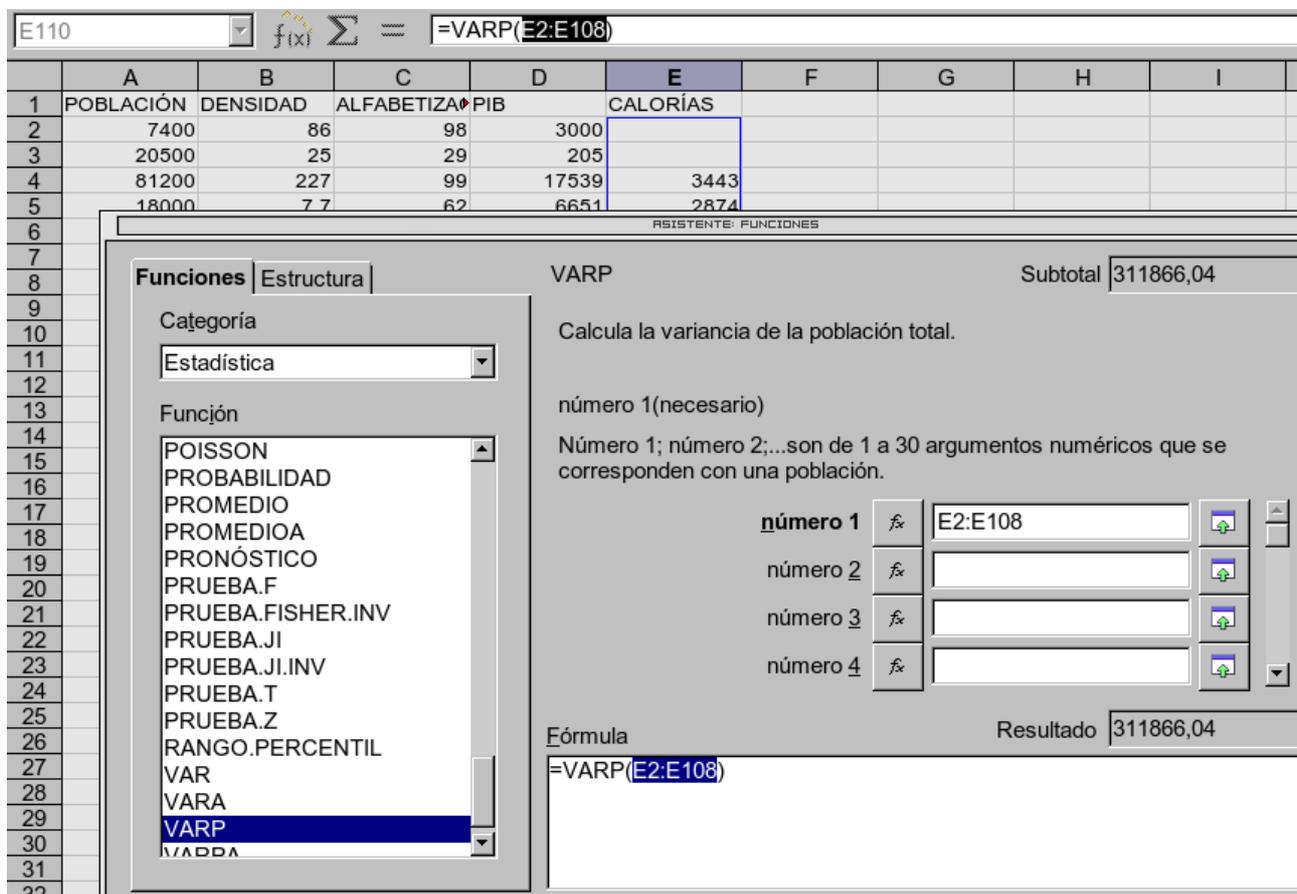
## Medidas de dispersión

De las medidas estudiadas en clase, CALC incorpora funciones para calcular automáticamente la desviación típica y la varianza, así como el recorrido intercuartílico. No existen funciones para calcular de forma inmediata el rango o recorrido ni el coeficiente de variación, pero se pueden calcular de forma muy sencilla. Así, el rango o recorrido es la diferencia entre el mayor valor de la variable (que se calcula con la función MÁX) y el menor (función MÍN). Por su parte, el coeficiente de variación se calcula como el cociente entre la desviación típica y la media en valor absoluto, lo que se puede llevar a cabo mediante la introducción de una sencilla fórmula en una celda.

En cuanto a la varianza, se calcula utilizando la función VARP; es importante notar la “p” final de esta función, pues CALC incorpora otra función denominada simplemente VAR que no calcula la varianza, sino una modificación denominada cuasivarianza. De la misma forma, la desviación típica se calcula con la función DESVESTP, también terminada en “p” para diferenciarla de la DESVEST, llamada cuasidesviación típica.

Veámoslo con el ejemplo del fichero `países.ods`. Si queremos calcular la varianza de la variable CALORÍAS debemos, como siempre, ubicarnos en la celda donde deseamos el resultado (en este caso los haremos en una celda bajo la columna de datos de la variable) y, a continuación, insertar la función VARP, teniendo en cuenta que en *Número 1* indicaremos el rango de valores de la variable CALORÍAS (en este caso, E2:E108, como se ve en la ilustración 19). Aceptando en el cuadro de diálogo de la función VARP llegaremos al resultado final (311866,04). Análogamente se procedería con la desviación típica mediante la función DESVESTP.

EJERCICIO 9: Calcula la desviación típica de la variable CALORÍAS.



	A	B	C	D	E	F	G	H	I
1	POBLACIÓN	DENSIDAD	ALFABETIZA	PIB	CALORÍAS				
2	7400	86	98	3000					
3	20500	25	29	205					
4	81200	227	99	17539	3443				
5	18000	7.7	62	6651	2874				

**Funciones** | Estructura | **VARP** | Subtotal: 311866,04

Categoría: Estadística

Función: **VARP**

Calcula la varianza de la población total.

número 1(necesario)

Número 1; número 2;...son de 1 a 30 argumentos numéricos que se corresponden con una población.

número 1: E2:E108

número 2: [ ]

número 3: [ ]

número 4: [ ]

Fórmula: =VARP(E2:E108) | Resultado: 311866,04

Ilustración 19

Podemos intentar, siguiendo los pasos descritos en las páginas 6 y 6, repetir el cálculo de la varianza para el resto de las variables. Como hemos insertado el resultado bajo una columna de datos, únicamente tenemos que copiar y pegar (o extender) la celda donde está la varianza de las CALORÍAS hacia la izquierda, y repetiremos la operación de la varianza, pero para cada columna.

A110:D110 $f(x)$ $\Sigma$ = =VARP(A2:A108)						
	A	B	C	D	E	
106	20600	22	88	2829	2582	
107	73100	218	88	230	2233	
108	9100	11	73	573	2077	
109						
110	21696580141,96	460433,08	518,75	42336057,77	311866,04	

Ilustración 20

Como señalábamos arriba, para calcular el recorrido o el coeficiente de variación, debemos realizar la operación utilizando el mínimo y el máximo o la media y la desviación típica, respectivamente. Por ejemplo, para calcular el coeficiente de variación de la variable DENSIDAD, únicamente hay que considerar su media y su desviación típica y dividir la segunda entre la primera mediante una fórmula de CALC, como se observa en la ilustración 21.

DESVESTP $f(x)$ $\times$ $\checkmark$ =G7/G6							
	A	B	C	D	E	F	G
1	POBLACIÓN	DENSIDAD	ALFABETIZACIÓN	PIB	CALORÍAS		
2	7400	86	98	3000			
3	20500	25	29	205			
4	81200	227	99	17539	3443		
5	18000	7,7	62	6651	2874	Densidad	
6	33900	12	95	3408	3113	Media	205,91
7	3700	126	98	5000		DT	678,55
8	17800	2,3	100	16848	3216	CV	=G7/G6
9	8000	94	99	18396	3495		

Ilustración 21

EJERCICIO 10: Calcula el recorrido y el coeficiente de variación de la variable PIB.

## Medidas de posición

En cuanto a las medidas de posición, CALC incorpora una función para calcular percentiles (PERCENTIL) que en realidad es válida para calcular cualquier cuantil, ya que no se limita a los 99 percentiles sino que permite calcular posiciones intermedias. La única variación respecto a las funciones anteriores radica en que para la función PERCENTIL hay que indicar no solamente el rango de datos sino también la posición del percentil que se desea calcular. En el ejemplo, si se desea calcular el percentil 65,5 de la variable PIB, será necesario insertar la función PERCENTIL indicando en el campo *Datos* el rango de datos de la variable (D2:D108, como se aprecia en la ilustración 22) y en el campo *k* la posición de percentil buscada, expresada como tanto por uno, es decir, como un número entre 0 y 1; así, para el percentil 65,5, tendremos que teclear 0,655. El resultado será el valor 5948,7.

EJERCICIO 11: Calcula la mediana, el percentil 5, el percentil 50 y el rango intercuartílico de la variable ALFABETIZACIÓN.

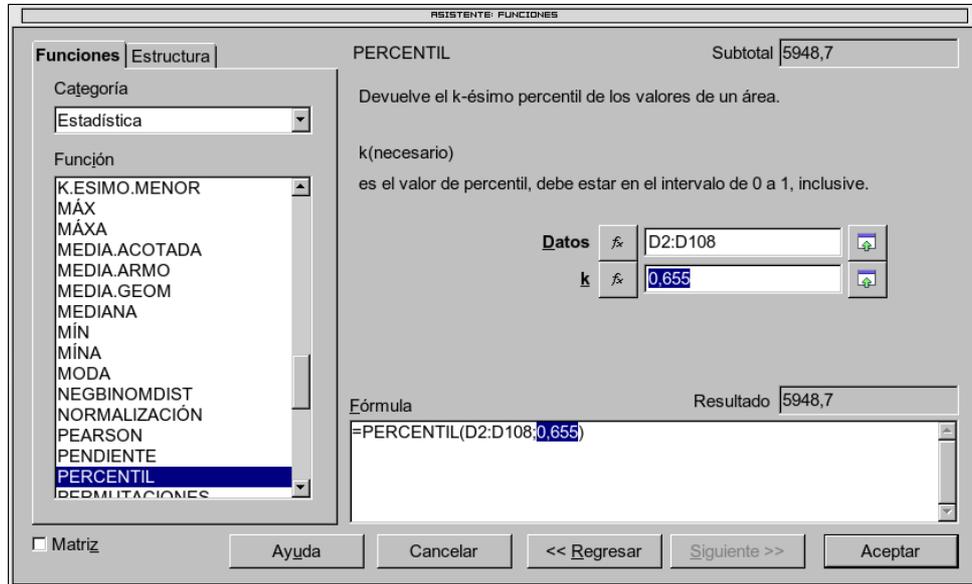


Ilustración 22

CALC incorpora la función RANGO.PERCENTIL que permite hacer la operación inversa al cálculo de un percentil; es decir, dado un determinado valor, se calcula la proporción de observaciones a su izquierda. Propiamente hablando, se trata de una función que calcula la función de distribución empírica o valor F de frecuencia relativa acumulada. En el ejemplo, si queremos calcular qué porcentaje de países tienen una densidad de población inferior o igual a 250, utilizaremos la función RANGO.PERCENTIL aplicada a la variable DENSIDAD (su rango de datos se indica en el campo *Datos*) y al valor 250 (en el campo *x*). La operación se aprecia en la ilustración 23. Para responder en términos de porcentaje, únicamente habría que multiplicar por cien la solución obtenida con la función.

EJERCICIO 12: Sea un país con cuarenta y cinco megahabitantes. ¿Qué porcentaje de países tendría una POBLACIÓN mayor?

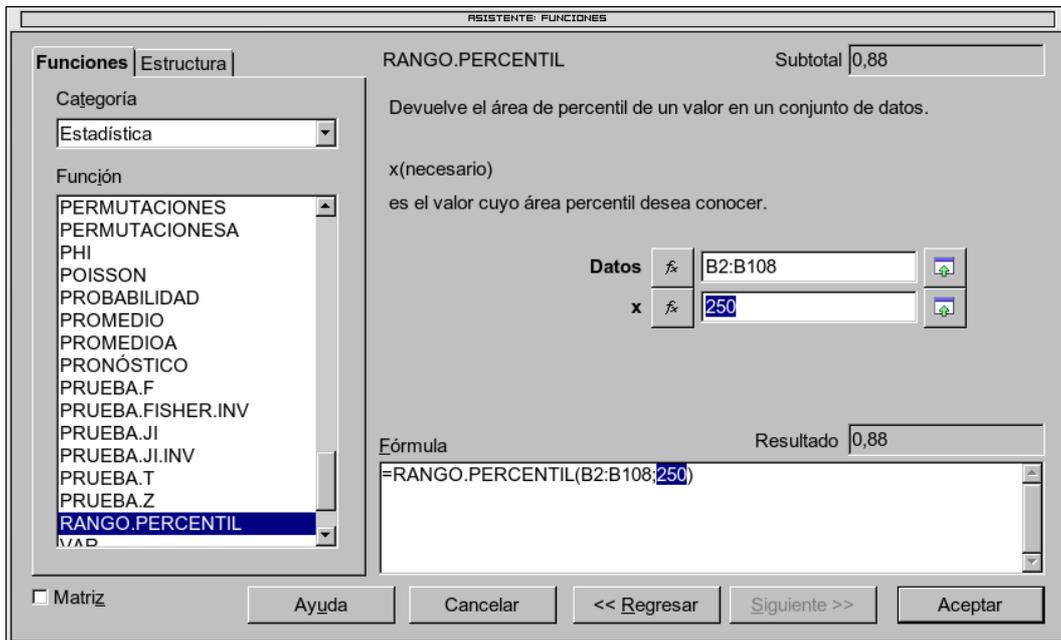


Ilustración 23

## 6. VARIABLES CONDICIONADAS CON CALC

CONCEPTOS: filtro; copiar y pegar antes de calcular.

Hay varios métodos igualmente válidos para trabajar con distribuciones condicionadas en CALC. Nosotros optaremos por una de ellas porque creemos que es la más sencilla, la que menos problemas puede producir para un usuario no habituado al programa y la que tiene un carácter más general.

Para utilizar una distribución condicionada, seguiremos los mismos pasos que al trabajar con los ejercicios realizados sobre el papel. Dada una variable estadística, y una vez aislados los datos que verifican la condición, estudiaremos sobre ellos las medidas que sean necesarias. Para especificar la condición y seleccionar las observaciones que la cumplen utilizaremos un *filtro*. Iremos viéndolo con un ejemplo.

Imaginemos que se nos pide calcular el valor máximo de P.I.B. para el 35% de los países con menor P.I.B., siempre que la ingesta de calorías no supere las 2000. Lo primero que debemos hacer es determinar lo que hay que calcular. En este caso se trata del percentil 35 de la variable P.I.B. | Calorías  $\leq$  2000.

Para establecer la condición, nos situamos sobre la fila de encabezamientos y elegimos *Datos > Filtro > Filtro predeterminado* o *automático* en el menú (ilustración 24).

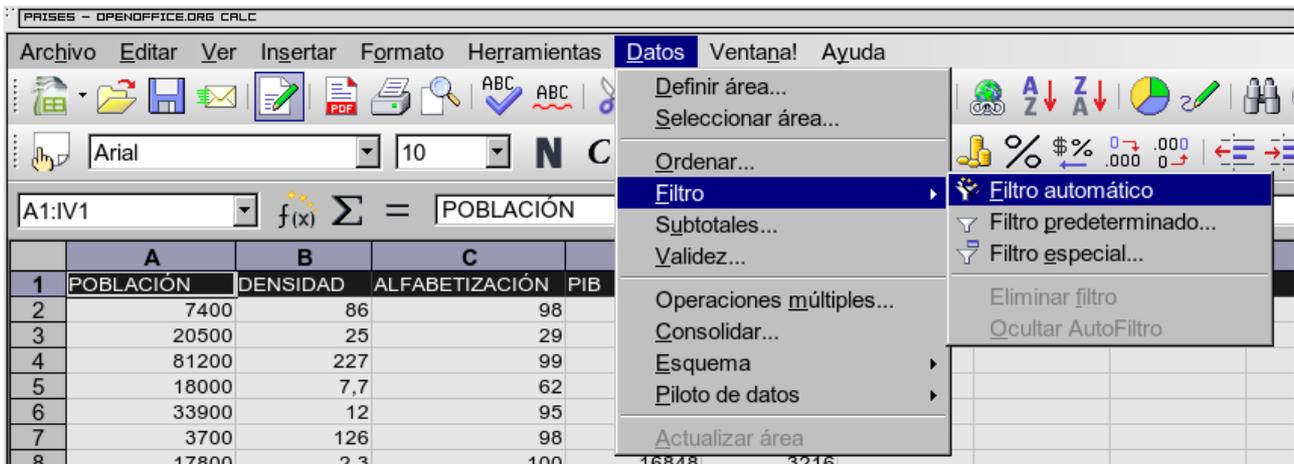


Ilustración 24

En el caso de haber elegido un filtro automático, aparecerán junto a los nombres de las variables unas listas desplegables como ésta , que nos permitirán seleccionar la condición o las condiciones. Por ejemplo, desplegando la lista de CALORÍAS, vemos que aparecen varias alternativas (ilustración 25): *todo* (que no realiza ninguna selección o anula las anteriores). *Los 10 primeros* (permite seleccionar los 10 mayores), *Predeterminado* (sirve para expresiones más complicadas) y *vacío* y *no vacío* (en el caso de que haya alguna celda en blanco; si no es así, estas dos últimas opciones no aparecerán).

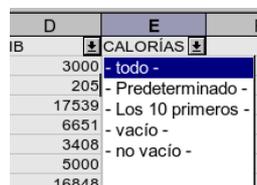


Ilustración 25

A la hora de poner condiciones, seleccionar un valor de la lista de valores de la variable nos permite expresar condiciones de igualdad como CALORÍAS = 1916 (por ejemplo). Para imponer condiciones de desigualdad (como el CALORÍAS ≤ 2000 que necesitamos ahora) o más complejas (combinando igualdades o desigualdades en varias variables) debemos ir a *Predeterminado*. En el cuadro de diálogo que se abre seleccionamos la condición (u otras condiciones sobre esta misma variable), como se ve en la ilustración 26. Fíjese a la izquierda, en la columna *Vínculo*: permite simultanear varias condiciones utilizando un operador *Y* o un operador *O*.

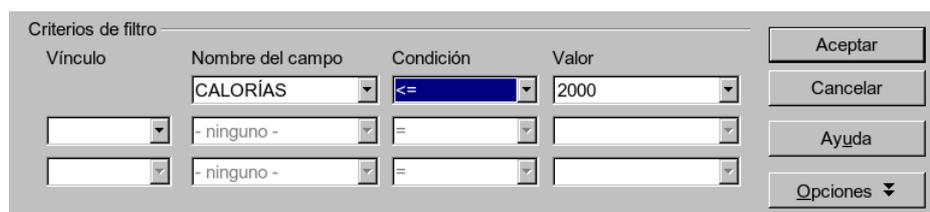


Ilustración 26

Una vez realizada esta operación en la lista de datos observaremos que ya no aparecerán todos los números de fila, sino únicamente los correspondientes a aquellos individuos en los que se verifica la condición  $CALORÍAS \leq 2000$  (ilustración 27).

	A	B	C	D	E
1	POBLACIÓN	DENSIDAD	ALFABETIZACIÓN	PIB	CALORÍAS
15	7900	6,9	78	730	1916
21	6000	216	50	208	1932
41	55200	47	24	122	1667
88	8400	311	50	292	1971
94	6667	10	24	2126	1906
109					

Ilustración 27

A continuación, lo que debemos hacer es seleccionar de entre estos datos los de la variable que queremos estudiar, y en este ejemplo se trata del P.I.B. (recordemos que buscamos el percentil 35 de la variable P.I.B. condicionada a  $CALORÍAS \leq 2000$ ).

Para trabajar con esta variable debemos dar un último paso, muy importante, antes de calcular el percentil. Como apreciábamos en la ilustración 27, de los datos originales nos hemos quedado con los que verificaban la condición, pero el resto, a pesar de no mostrarse, sigue estando en la hoja de cálculo (de ahí que los números de fila no sean correlativos). Por ese motivo, si realizamos cualquier cálculo sobre esta hoja utilizaremos tanto los datos que se ven como los que no se ven. Por consiguiente, para continuar debemos copiar y pegar en una hoja de cálculo nueva los datos que necesitamos. En este ejemplo son los del P.I.B., por lo tanto, los seleccionamos y copiamos (se puede hacer pinchando con el botón derecho, ilustración 28).

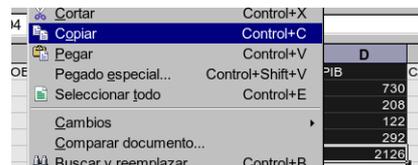


Ilustración 28

A continuación, pegamos estos datos en una nueva hoja. Para ello, o bien seleccionamos *Insertar > Hoja de cálculo* en el menú principal (ilustración 29) o bien pinchamos con el botón derecho en las solapas inferiores y elegimos *Insertar > Hoja de cálculo* (ilustración 30).

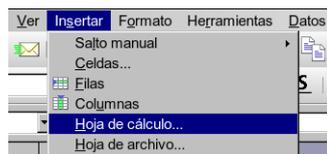


Ilustración 29

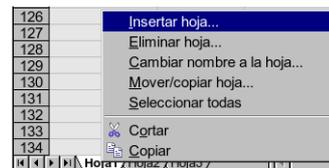


Ilustración 30

En la nueva hoja de cálculo (cuyo nombre podemos cambiar, tal y como indicamos al

comienzo de la guía), pegamos los datos. Así, en este caso, ya habremos construido la variable P.I.B. | Calorías  $\leq$  2000 (ilustración 31).

	A
1	PIB
2	730
3	208
4	122
5	292
6	2126
7	

Ilustración 31

Sobre esta nueva variable (P.I.B. | Calorías  $\leq$  2000) realizamos los cálculos necesarios; en este ejemplo, el percentil 35 (ilustración 32).

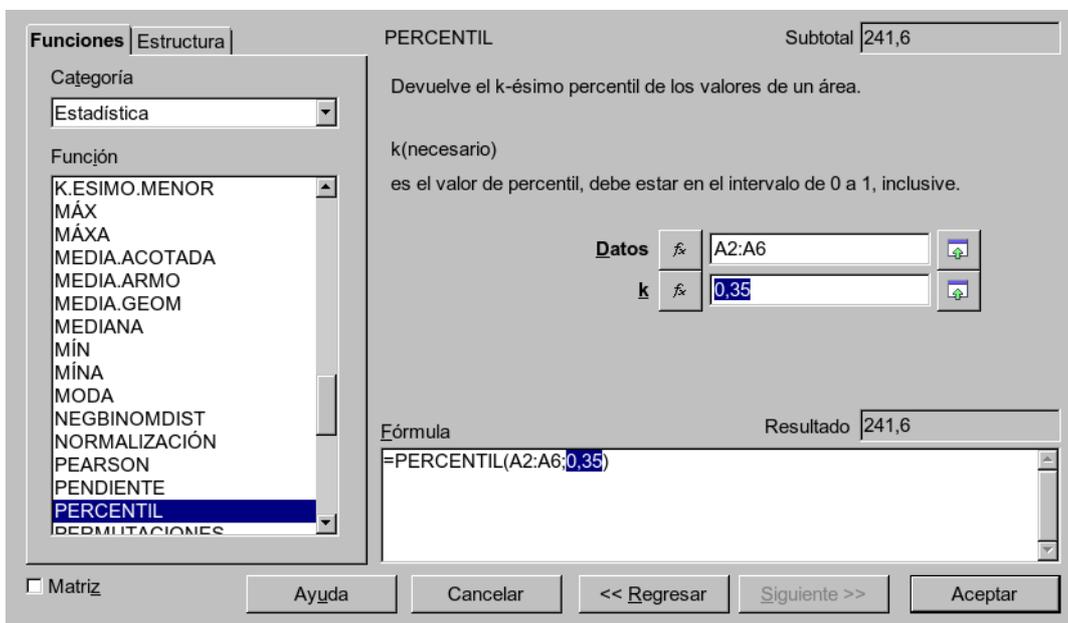


Ilustración 32

Es muy importante tener en cuenta que, una vez que ha finalizado el trabajo con una variable condicionada, construida a partir de un autofiltro, para comenzar a construir otra variable condicionada es indispensable volver a la situación inicial. Es decir, hay que tener cuidado para no imponer condiciones nuevas sobre la selección que se había hecho para una variable anterior. Para ello hay que desactivar las condiciones correspondientes a la variable que ya no se necesitan. Esto se hace seleccionando la opción *Todo* dentro de aquellos desplegados del Autofiltro que estén activos (ilustración 33) o bien seleccionando *Datos > Filtro > Eliminar filtro* (ilustración 34).

	A	B	C	D	E
1	POBLACIÓN	DENSIDAD	ALFABETIZACIÓN	PIB	CALORÍAS
15	7900	todo -	78	730	1916
21	6000	- Predeterminado -	50	208	1932
41	55200	- Los 10 primeros -	24	122	1667
88	8400	- vacío -	50	292	1971
94	6667	- no vacío -	24	2126	1906

Ilustración 33

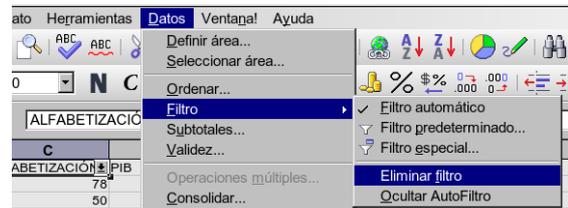


Ilustración 34

### EJERCICIOS:

- 13) ¿Cuál es la tasa de alfabetización más frecuente entre los países con una densidad de población no inferior a 100 hab/km<sup>2</sup> y con una ingesta de, a lo sumo, 3000 calorías?
- 14) ¿Una tasa de alfabetización del 84% en un país con ingesta de calorías superior a 3000 es relativamente mayor o menor que una tasa de alfabetización del 25% en un país de P.I.B. no mayor de 3000?
- 15) ¿Cuál es la población mínima del 40% de los países con mayor población?
- 16) ¿Cuál es el P.I.B. medio per cápita en países cuya tasa de alfabetización es a lo sumo del 60%? ¿Y en los que la tasa de alfabetización es al menos del 65%? ¿Cuál de ellos representa mejor su grupo de países?

## 7. Soluciones:

- 1) Tenga en cuenta que para algunas celdas será necesario teclear antes un apóstrofo. Considere también el formato de las celdas para establecer el número de decimales.
- 2) Por ejemplo, si ha escrito los números del 1 al 10 en el rango A1:A10, puede escribir la fórmula =A1\*13 en la celda B1; luego, extienda esa fórmula en vertical hasta B10.  
(Tenga en cuenta que puede “automatizar” la tarea de escribir los números del 1 al 10: puede escribir un 1 en A1 y, luego, escribir =A1+1 en la celda A2 y extender dicha fórmula hasta A10. O más fácil aún: si tiene un 1 escrito en A1, puede arrastrarlo directamente en vertical y CALC genera automáticamente los siguientes enteros correlativos; esto funciona con cualquier número entero de partida, no sólo con el uno.)
- 3) La fórmula sería algo parecido a PRODUCTO(A1:A6) y el resultado es 1920.
- 4) Los resultados son: 16, 14, 13, 22'91 y 25.
- 5) a) Si los precios están en el rango A1:A5, ponga en B1 la fórmula =A1-A1\*\$D\$2/100. Los resultados son 880, 960, 1500, 476 y 580. Al cambiar D2 a 25, los resultados son 825, 900, 1406'25, 446'25 y 543'75.  
b) En una hoja en blanco, escriba 1 en A1. Extiéndalo en vertical hasta A10, para tener en columna los diez primeros números naturales. Vuelva a seleccionar A1, y extienda hasta J1, para tener en fila los diez primeros números naturales. Escriba en A2 la fórmula

= $A2*B\$1$ . Copie el contenido de A2 y péguelo en el rango A2:J10. Alternativamente, podría utilizar las funciones FILA y COLUMNA; por ejemplo, en una hoja en blanco ponga en A1 la fórmula =FILA(A1)\*COLUMNA(A1), luego copie A1 y pegue en el rango A1:J10.

- 6) Recuerda hacer *Pegado especial*, desactivando *Pegar todo* y *Fórmulas* y activando *Números*.
- 7) La suma total de kilohabitantes =SUMA(A2:A108) y da 5.189.603 kilohabitantes, es decir, 5.189.603.000 habitantes, luego algo más de 5 gigahabitantes.

Como conocemos la alfabetización, para cada país podemos calcular el número de analfabetos. Por ejemplo, en F2 podemos insertar la fórmula = $A2*(100-C2)/100$  y extender la fórmula hasta F108. El resultado final lo daría la fórmula =SUMA(F2:F108), y sería 1.352.288,97 kilohabitantes.

La proporción total de analfabetos es  $1352288,97 / 5189603$ , aproximadamente igual a 0'26, es decir, un 26%.

Podemos calcular la extensión de un país a partir de su población y su densidad. Como densidad = población / extensión, tenemos que extensión = población / densidad. Para obtener directamente el resultado en kilómetros cuadrados, podemos poner en G2 la fórmula = $A2*1000/B2$  y extenderla hasta G108. El resultado final lo da la fórmula =SUMA(G2:G108) y resulta aproximadamente 109.621.215,82 km<sup>2</sup>.

- 8) País medio: 48.500.963 habitantes, 205'9 habitantes por kilómetro cuadrado.

País mediano: 10.400.000 hab, 64 hab/km<sup>2</sup>.

Según la Wikipedia España tiene (tenía registrados el uno de enero de 2008) 46.157.822 hab y 504.645 km<sup>2</sup>, luego 91'5 km<sup>2</sup> aproximadamente. Luego España es un país “medio” en población. En cuanto a densidad, estaría entre la media y la mediana; en todo caso, en un lugar “central”.

El cálculo de la moda es posible, pero no tiene sentido ya que se trata de variables con muchas modalidades (toman muchos valores posibles) y el valor más frecuente tendrá frecuencia uno o dos, lo cual es poco representativo.

- 9) DESVESTP(E2:E108) da 558,45.

- 10) El recorrido es el máximo menos el mínimo:  $23474 - 122 = 23352$ .

El coeficiente de variación es la desviación típica entre la media:  $6506'62 / 5831'4 = 1'12$ .

- 11) La mediana equivale al percentil 50.

El percentil 5 se calcula con PERCENTIL(C2:C108;0,05) y es igual a 30,8.

El percentil 50 es PERCENTIL(C2:C108;0,5) y vale 88.

El rango intencuartílico en la diferencia entre el percentil 75 y el percentil 25:  $98 - 63 = 35$ .

- 12) Dicho país tiene una población de 45000 khab. El rango percentil de ese valor es RANGO.PERCENTIL(A2:A108;45000) y vale 0,78, es decir, un 78% de países tienen una

población menor. Por tanto, un 22% de países tendrán una población mayor.

- 13) Se trata de calcular la moda (valor más frecuente) de la variable Tasa de alfabetización | Densidad  $\geq 100$  y Calorías  $\leq 3000$ . Para ello, con un Autofiltro, imponemos las condiciones simultáneamente en cada una de las variables (Densidad  $\geq 100$  y Calorías  $\leq 3000$ ) y seleccionamos la variable de interés (Tasa de alfabetización) para copiarla y pegarla en una hoja nueva. Con estos datos ya podemos calcular su moda (mediante la función estadística MODA), que resultará 50%.
- 14) Para responder debemos tipificar ambos valores con el objeto de compararlos. El primer valor (84%) corresponde a la variable Alfabetización | Calorías  $> 3000$ . Por consiguiente, con el autofiltro realizamos la selección Calorías  $> 3000$ , y seleccionamos la Alfabetización, pasándola mediante copia a una hoja nueva. Una vez construida la variable Alfabetización | Calorías  $> 3000$  debemos calcular su media (función PROMEDIO) y su desviación típica (DESVESTP) para tipificar posteriormente el valor 84%, con una operación realizada “a mano” en la celda: (valor-media) / desviación típica, obteniendo el valor -0'4838. Quitando la condición anterior e imponiendo ahora P.I.B.  $\leq 3000$ , construimos de forma análoga la variable Alfabetización | P.I.B.  $\leq 3000$ . Calculando su media y su desviación típica, tipificamos 25%, y obtenemos -1,73. Por lo tanto, la primera tasa (84%) es relativamente mayor
- 15) Se nos pide el percentil 60 de la variable población, que dejará por encima el 40% de los valores. Por lo tanto, sin necesidad de filtros, mediante la función PERCENTIL (indicando 0,6 en la casilla correspondiente), obtenemos el valor 18900 (medido en miles de habitantes).
- 16) Hay que comparar la representatividad de dos medias, lo cual se realiza mediante la comparación de los coeficientes de variación de las dos variables. La primera de ellas es el P.I.B. | Alfabetización  $\leq 60$ . Mediante un filtro construimos la variable, y calculamos su media (628,56) y su desviación típica (547,887) con las funciones de CALC, y “a mano” el coeficiente de variación (0,8716). La segunda variable es P.I.B. | Alfabetización  $\geq 65$ . Análogamente, construimos la variable con un filtro, calculamos su media (7550,85), su desviación típica (6794,14) y su coeficiente de variación (0,8998). Por consiguiente, es más representativa la primera media (la variable tiene menos variación relativa).