

**PRÁCTICA 1****1.2. INTRODUCCIÓN DE DATOS**

**Ejemplo:** Introducir los siguientes datos, con el nombre `Datos_personas`, usando el editor de datos.

Nombre	Color_de_pelo	Altura
Luis	Moreno	1.75
María	Rubio	1.64
Pedro	Rubio	1.72
Alberto	Castaño	1.78
Isabel	Castaño	1.67
Manuel	Moreno	1.71
María	Castaño	1.73

Datos → Nuevo Conjunto de Datos

Podemos editar los datos si nos vamos al botón: Editar conjunto de datos. Y si queremos visualizarlo, Visualizar conjunto de datos.




Si queremos guardar el archivo con los datos introducidos, Datos → Conjunto de datos activo → Guardar el conjunto de datos activo. En el menú, seleccionaremos Archivos de datos de R.

**1.3. FUSIONAR CONJUNTOS DE DATOS**

Primero, creamos dos nuevos conjuntos de datos para realizar el ejercicio:




Nuevo conjunto de datos

Introducir el nombre del conjunto de datos:

 Ayuda  Aceptar  Cancelar

Nuevo conjunto de datos

Introducir el nombre del conjunto de datos:




 Ayuda  Aceptar  Cancelar

Editor de datos: Datos\_personas

Fichero Editar Ayuda

Añadir fila Añadir columna

	1	2	3
rowname	Nombre	Peso	Altura
1	Pedro	78	1.75
2	María	53	1.62
3	Luisa	56	1.64
4	Miguel	80	1.70
5	Ana	59	1.70




 Ayuda  Aceptar  Cancelar

Editor de datos: Datos\_persona...

Fichero Editar Ayuda

Añadir fila Añadir columna

	1	2
rowname	Nombre	Color_pelo
1	Pedro	Moreno
2	María	Rubio
3	Luisa	Castaño
4	Miguel	Rubio
5	Ana	Castaño

 Ayuda  Aceptar  Cancelar

Datos → Fusionar conjunto de datos

**Fusionar Conjuntos de Datos**

Nombre del conjunto de datos fusionado:




Primer conjunto de datos (seleccionar uno) Segundo conjunto de datos (seleccionar uno)

Conjunto\_Datos\_Perso  
Datos\_persona\_extra  
Datos\_personas

Dirección de la fusión

☐ Fusión de filas  
☒ Fusionar columnas

☐ Fusión únicamente de filas o columnas comunes

 Ayuda  Aceptar  Cancelar

**Conjunto\_Datos\_Perso**

	Nombre.x	Peso	Altura	Nombre.y	Color_pelo
1	Pedro	78	1.75	Pedro	Moreno
2	María	53	1.62	María	Rubio
3	Luisa	56	1.64	Luisa	Castaño
4	Miguel	80	1.70	Miguel	Rubio
5	Ana	59	1.70	Ana	Castaño

Vamos a eliminar las columnas de nombres ya repetidos...

Nos vamos a la columna que queramos eliminar, clic derecho, y le damos a Borrar columna actual.

Modificamos el nombre de "Nombre" y listo.

**Editor de datos: Conjunto\_Datos\_Perso**

Fichero Editar Ayuda

Añadir fila Añadir columna

	rowname	1	2	3	5
		Nombre	Peso	Altura	Color_pelo
1	1	Pedro	78	1.75	Moreno
2	2	María	53	1.62	Rubio
3	3	Luisa	56	1.64	Castaño
4	4	Miguel	80	1.7	Rubio
5	5	Ana	59	1.7	Castaño

 Ayuda  Aceptar  Cancelar



**Conjunto\_Dat...**

	Nombre	Peso	Altura	Color_pelo
1	Pedro	78	1.75	Moreno
2	María	53	1.62	Rubio
3	Luisa	56	1.64	Castaño
4	Miguel	80	1.70	Rubio
5	Ana	59	1.70	Castaño

#### 1.4. CONSTRUCCIÓN DE NUEVAS VARIABLES

Datos → Modificar variables del conjunto de datos activo → Calcular una nueva variable

OPERACIONES PARA EL CAMPO EXPRESIÓN A CALCULAR:

Funciones	
Raíz cuadrada: $\text{sqrt}(\cdot)$	Exponencial: $\text{exp}(\cdot)$
Logaritmo neperiano: $\text{log}(\cdot)$	Seno: $\text{sin}(\cdot)$
Coseno: $\text{cos}(\cdot)$	Tangente: $\text{tan}(\cdot)$

Operadores		
Aritméticos	De comparación	Logicos
Suma: +	Igualdad: ==	'Y' lógico: &
Diferencia: -	Distinto: !=	'No' lógico: !
Producto: *	Menor que: <	'O' lógico:
División: /	Mayor que: >	
Potencia: ^	Menor o igual que: <=	
	Mayor o igual que: >=	

### 1.5. FILTRADO DE DATOS

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo

**EJEMPLO: Queremos saber el consumo de los coches Europeos**

Filtrar el conjunto de datos

☐ Incluir todas las variables

Variables (seleccionar una o más)

- acel
- año
- cilindr
- consumo**
- cv
- motor

Expresión de selección

origen == 'Europa'

Nombre del nuevo conjunto de datos

Consumos\_Europa

Ayuda Aceptar Cancelar

Para ver los datos Estadísticos → Resúmenes → Conjunto de datos activo

```
consumo
Min.   : 5.000
1st Qu.: 8.000
Median : 9.000
Mean   : 8.886
3rd Qu.:10.000
Max.   :15.000
NA's   :3
```

Siendo, valor mínimo (Min.), valor máximo (Max.), y valor medio (Mean).

**Ejemplos:**

- Determinar el consumo medio de los coches japoneses que hay en la muestra.

Seleccionamos *RCars* como conjunto de datos activo. A continuación realizamos un filtrado del conjunto de datos como sigue: desmarcamos la casilla **Incluir todas las variables** y seleccionamos la variable *consumo*. Dado que sólo queremos seleccionar los casos en los que el origen del coche es japonés, como expresión de selección ponemos *origen=='Japón'* (cuidado con las mayúsculas y las tildes). Finalmente, pondremos como nombre al nuevo conjunto de datos *Consumos\_Japón*. Pulsamos **Aceptar** y ya tenemos listo el nuevo conjunto de datos. Finalmente elegimos la opción **Estadísticos**→**Resúmenes**→**Conjunto de datos activo**; en la salida podemos ver que el consumo medio de los coches japoneses es de 8.051 l/100Km.

- Calcular la potencia máxima de los coches que no tienen 8 cilindros.

Seleccionamos *RCars* como conjunto de datos activo y procedemos de manera similar al ejemplo anterior pero, en este caso, la variable a seleccionar es *cv* y la expresión de selección será *cilindr != 8*. Nótese que, al ser 8 un valor numérico no tenemos que usar comillas. Al nuevo conjunto de datos le pondremos como nombre *cv\_no\_8*. Finalmente, repitiendo el paso final del ejemplo anterior, encontramos que la potencia máxima de los coches que no tienen 8 cilindros es de 165cv.

**EJEMPLO 1:**

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo



```

consumo
Min.   : 5.000
1st Qu.: 7.000
Median : 7.000
Mean   : 8.051
3rd Qu.: 9.000
Max.   :13.000

```

**EJEMPLO 2:**

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo

**Filtrar el conjunto de datos**

☐ Incluir todas las variables

0

Variables (seleccionar una o más)

acel  
año  
cilindr  
consumo  
cv  
motor

Expresión de selección  
cilindr != 8

Nombre del nuevo conjunto de datos  
cv\_no\_8

Ayuda Aceptar Cancelar



```

cv
Min.    : 46.00
1st Qu.: 71.00
Median  : 86.00
Mean    : 85.34
3rd Qu.: 97.00
Max.    :165.00
NA's    :6

```

### Ejercicios:

- Determinar el valor máximo de la variable *acel* para los coches europeos que tienen cuatro cilindros (sol: 24.80 s.).
- Determinar la potencia media de los coches que tienen consumo no superior a 8 l/100Km (sol: 72.79cv).

### EJERCICIO 1:

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo

**Filtrar el conjunto de datos**

☐ Incluir todas las variables

0

Variables (seleccionar una o más)

acel  
año  
cilindr  
consumo  
cv  
motor

Expresión de selección  
cilindr == 4

Nombre del nuevo conjunto de datos  
acel\_cilindr\_europeos

Ayuda Aceptar Cancelar



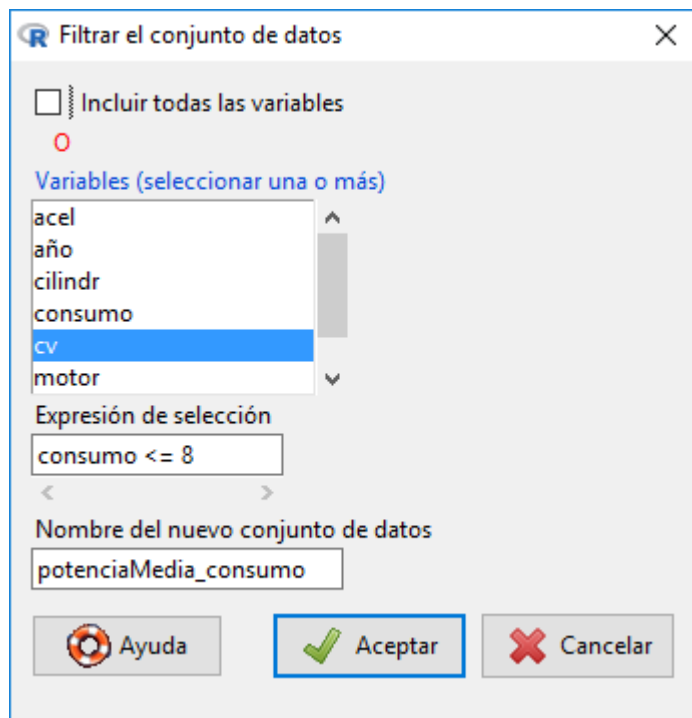
```

acel
Min.    :11.60
1st Qu.:14.85
Median  :16.20
Mean    :16.62
3rd Qu.:18.00
Max.    :24.80

```

**EJERCICIO 2:**

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo



```
      cv
Min.   : 48.00
1st Qu.: 65.00
Median : 70.00
Mean   : 72.79
3rd Qu.: 80.00
Max.   :132.00
NA's   :2
```