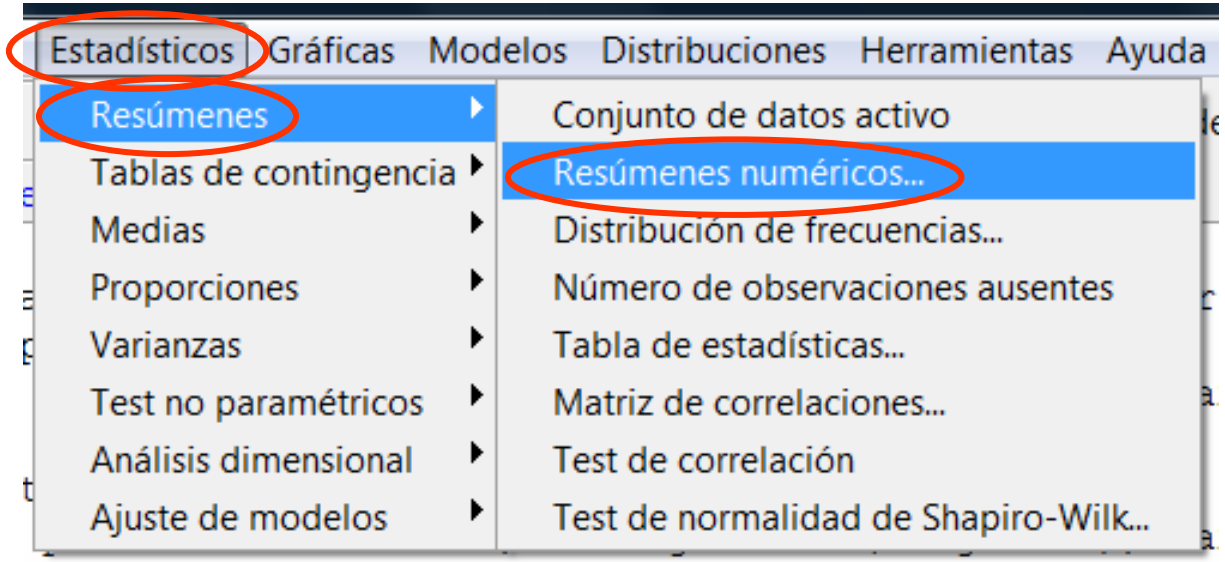
Laboratorio 1: Análisis de datos univariantes

Reproduce con la variable `region` el mismo análisis y observa las diferencias.

Se calculan los resúmenes de los datos o estadísticos descriptivos básicos:

Estadísticos → Resúmenes → Resúmenes Numéricos

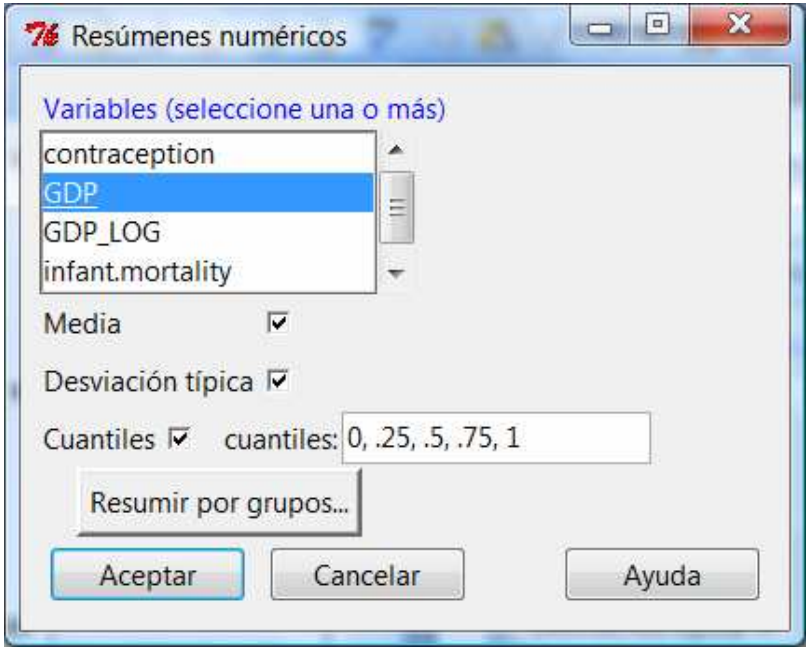


La variable "region" no es numérica

Datos

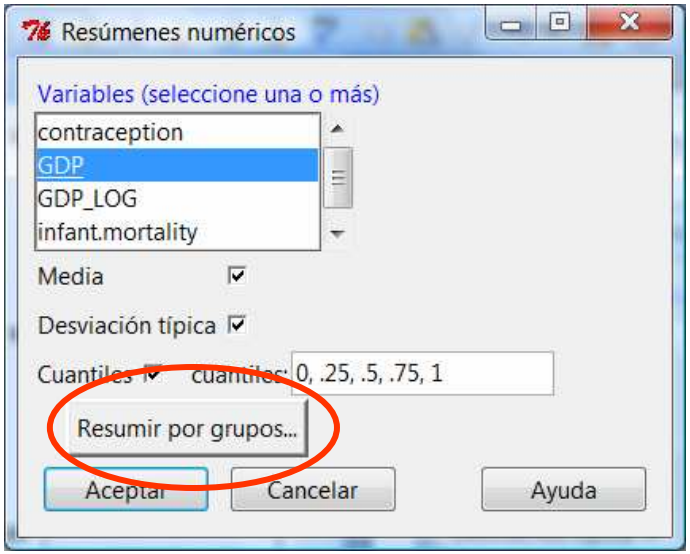
	TFR	contraception	infant.mortality	GDP	region	GDP_LOG
Algeria	3.81	52	44	1531	Africa	7.333676
Armenia	1.70	22	25	354	Europe	5.869297
Australia	1.89	76	6	20046	Oceania	9.905785
Austria	1.42	71	6	29006	Europe	10.275258
Azerbaijan	2.30	17	33	321	Asia	5.771441
Bahamas	1.95	62	14	12545	Americas	9.437077
Bahrain	2.97	53	18	9073	Asia	9.113058
Bangladesh	3.14	49	78	280	Asia	5.634790
Barbados	1.73	55	9	7173	Americas	8.878079
Belarus	1.40	50	15	994	Europe	6.901737
Belgium	1.62	79	7	26582	Europe	10.187990
Belize	3.66	47	30	2569	Americas	7.851272
Benin	5.83	16	84	391	Africa	5.968708
Bhutan	5.89	19	104	166	Asia	5.111988
Bolivia	4.36	45	66	909	Americas	6.812345
Botswana	4.45	33	56	3640	Africa	8.199739
Brazil	2.17	74	42	4510	Americas	8.414052
Burkina-Faso	6.57	8	97	165	Africa	5.105945
Burundi	6.28	9	114	205	Africa	5.323010
Cameroon	5.30	16	58	627	Africa	6.440947
Canada	1.61	66	6	18943	Americas	9.849190
Central-African-Rep	4.95	24	96	379	Africa	5.937536
Chile	2.44	30	13	4736	Americas	8.462948
China	1.80	83	38	582	Asia	6.366470
Colombia	2.69	72	24	2215	Americas	7.703008
Comoros	5.51	21	82	367	Africa	5.905362
Cook-Islands	3.50	38	26	5432	Oceania	8.600063
Costa-Rica	2.95	75	12	2696	Americas	7.899524
Cuba	1.55	70	9	1983	Americas	7.592366
Czech-Republic	1.40	69	9	4450	Europe	8.400659

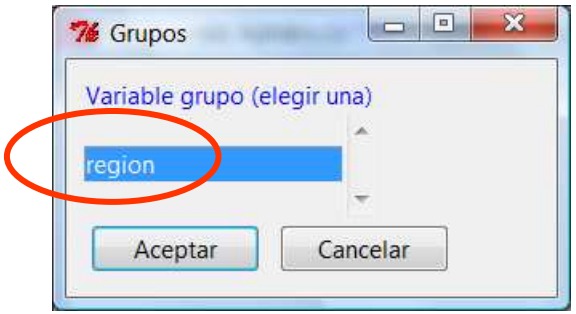
GDP en conjunto:



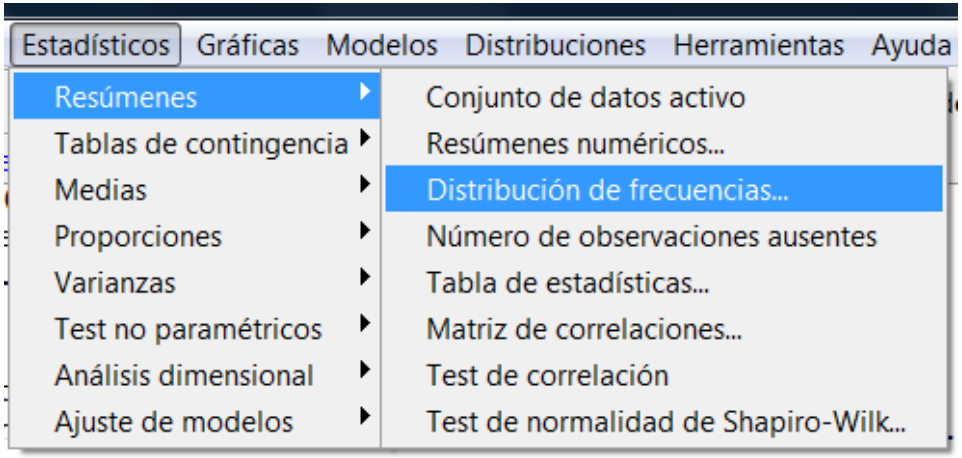

```
> numSummary(Datos[,"GDP"], statistics=c("mean", "sd", "quantiles"), quantiles=c(0,.25,.5,.75,1))
     mean       sd 0% 25%  50%  75%  100%   n
 5518.177 9114.24 36 391 1565 4450 42416 141
```

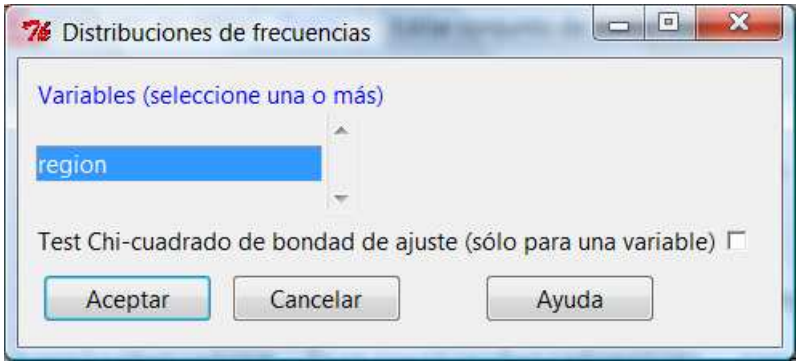
GDP en función de la *region* (Continente) al que pertenece el país





	mean	sd	0%	25%	50%	75%	100%	n
Africa	1166.122	2071.969	36	223.0	397	1124.0	11854	41
Americas	4615.759	5824.164	386	1660.0	2569	4510.0	26037	29
Asia	3575.226	5654.231	122	326.0	716	3234.5	22898	31
Europe	13638.680	13839.427	343	1595.0	4450	26444.0	42416	25
Oceania	9639.733	12339.857	654	1197.5	2593	18158.0	41718	15



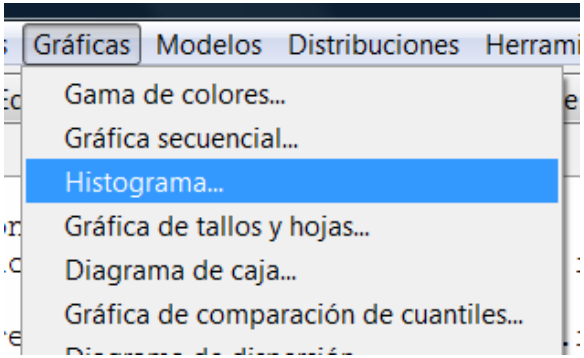


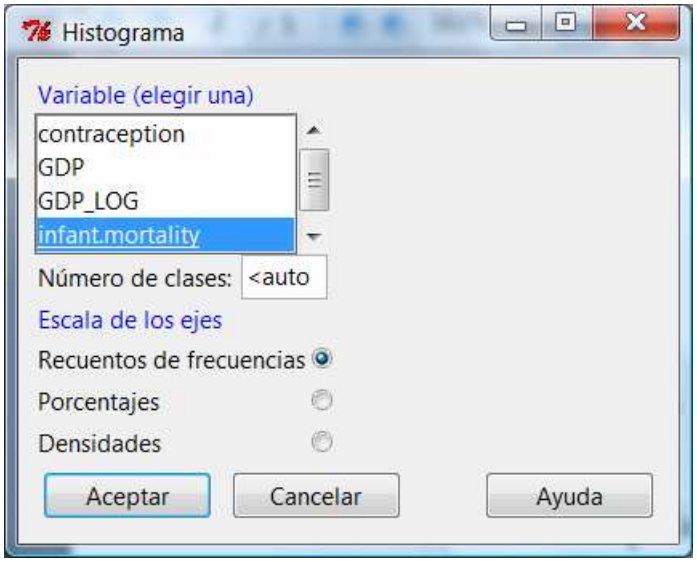

```
  Africa Americas     Asia   Europe  Oceania
29.07801 20.56738 21.98582 17.73050 10.63830
```

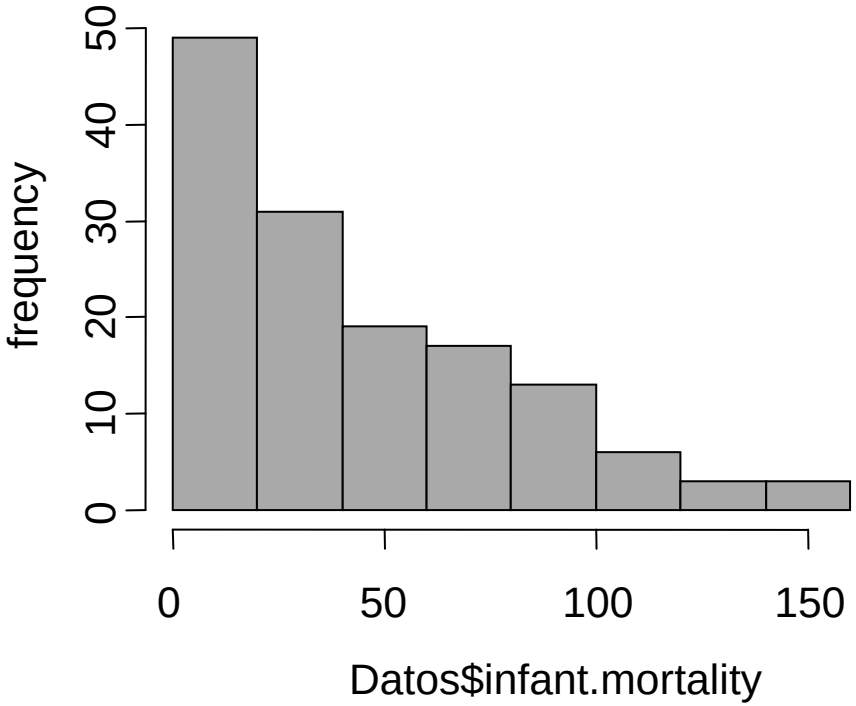
Histogramas

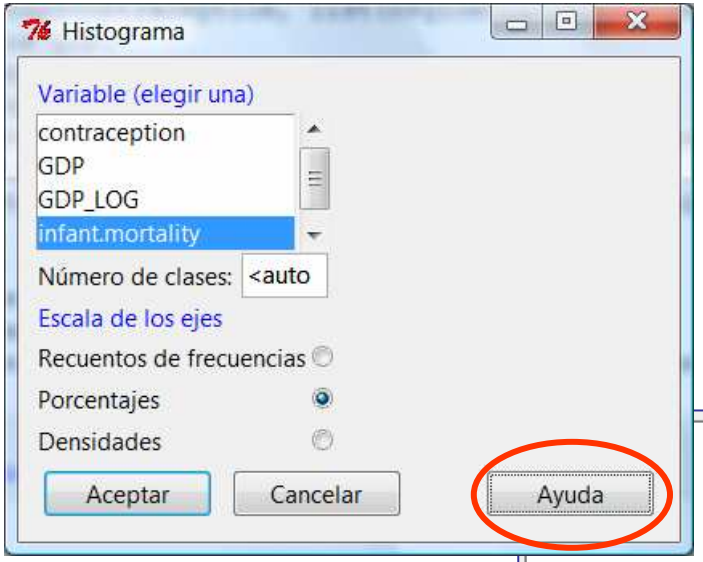
Gráficas → **Histograma** (e.g. de la variable `infant.mortality`)

Nota: Las opciones por defecto del gráfico se pueden cambiar modificando el comando que aparece en la *ventana de instrucciones*:









Plot a Histogram

Description

This function is a wrapper for the `hist` function in the `base` package, permitting percentage scaling of the vertica

Usage

```
Hist(x, scale=c("frequency", "percent", "density"), xlab=deparse(substitute(x)),
        ylab=scale, main="", ...)
```

Arguments

`x`	a vector of values for which a histogram is to be plotted.
`scale`	the scaling of the vertical axis: `"frequency"` (the default), `"percent"`, or `"density"`.
`xlab`	x-axis label, defaults to name of variable.
`ylab`	y-axis label, defaults to value of `scale`.
`main`	main title for graph, defaults to empty.
`...`	arguments to be passed to `hist`.

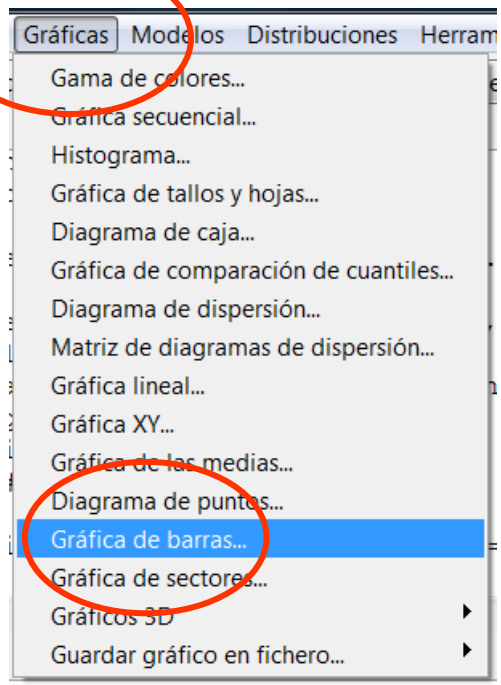
Value

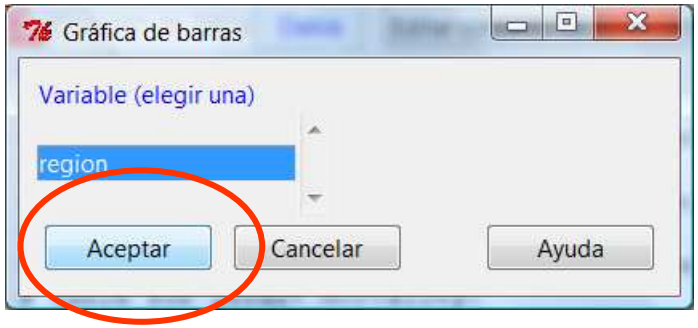
This function returns `NULL`, and is called for its side effect — plotting a histogram.

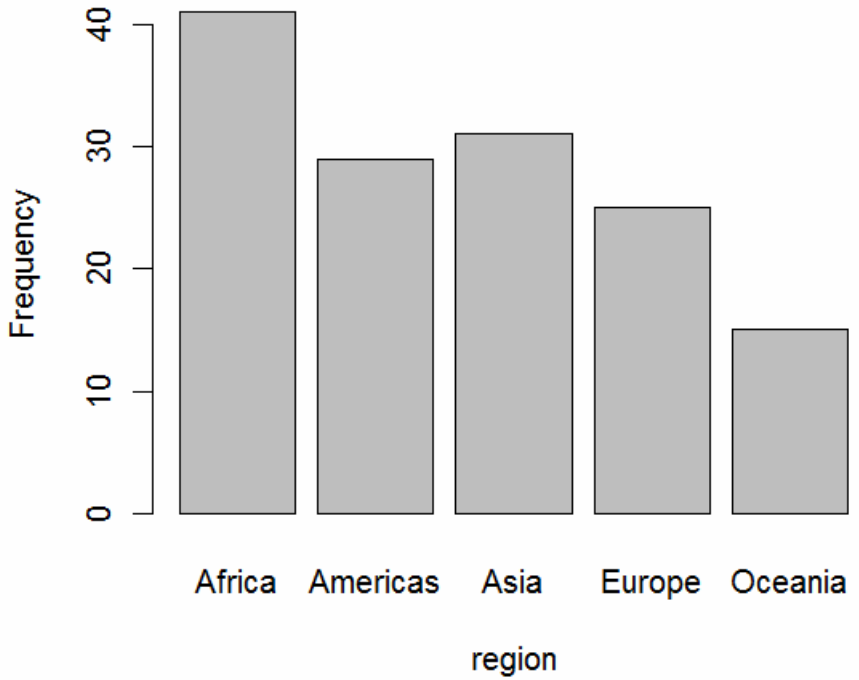
Author(s)

Diagrama de Barras

Gráficas → Gráfica de Barras (e.g. de la variable `region`)

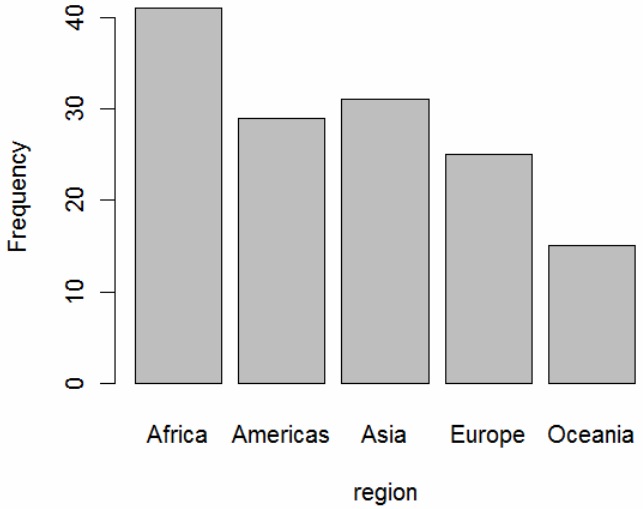




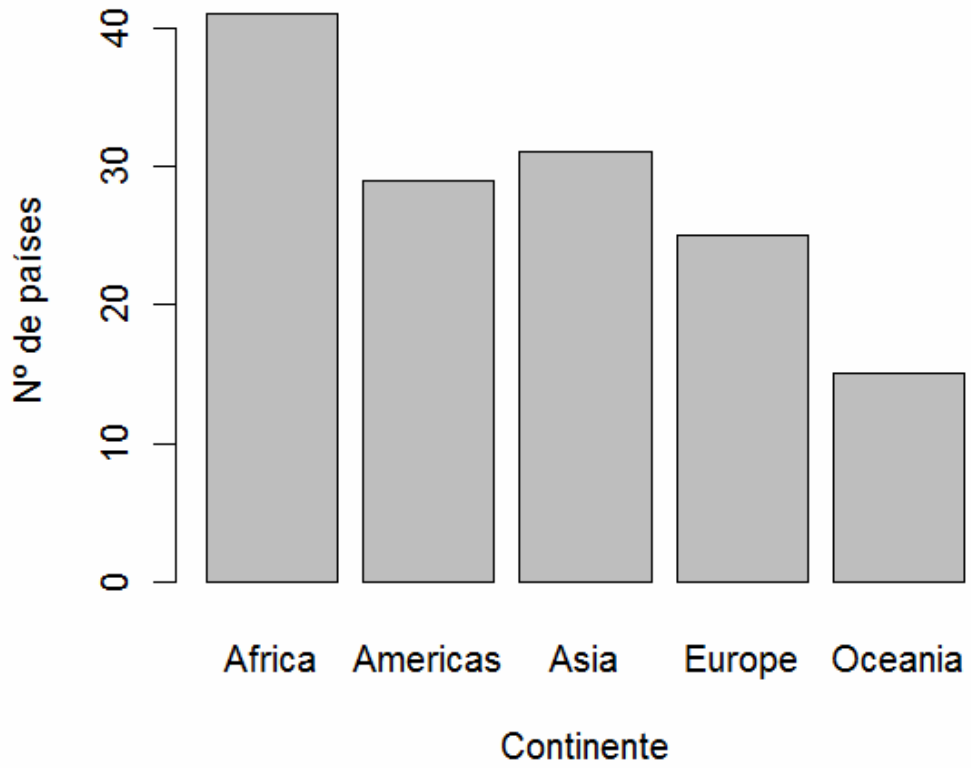



```
barplot(table(Datos$region), xlab="region", ylab="Frequency")
```

```
barplot(table(Datos$region), xlab="region", ylab="Frequency")
```

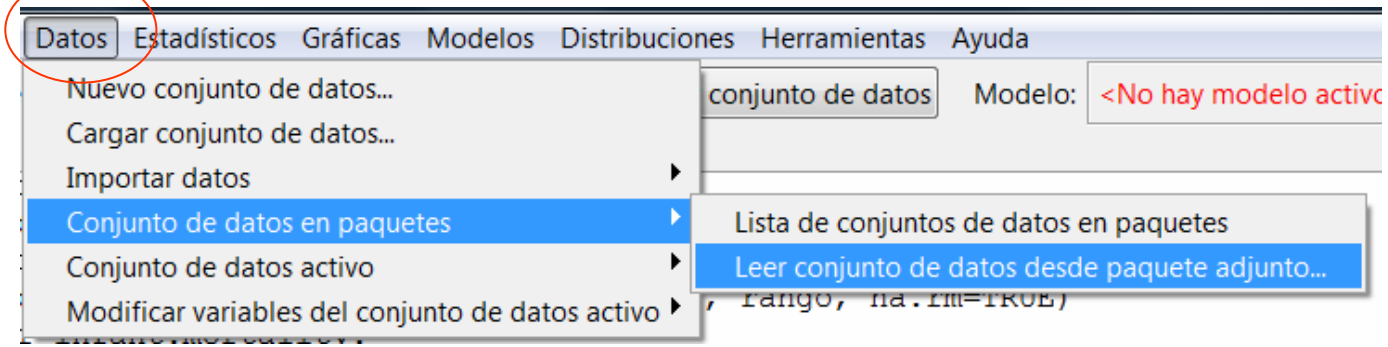


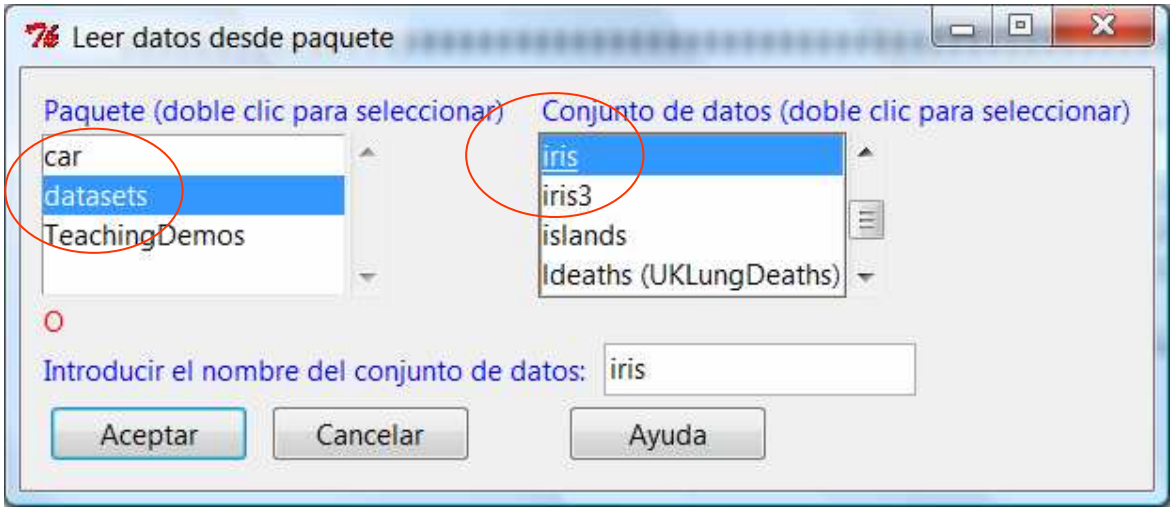

```
barplot(table(Datos$region), xlab="Continente", ylab="Nº de países")
```



Se consideran ahora los datos del ejemplo `iris` del paquete `datasets` de R.

Datos → Datos en paquetes → Leer datos desde paquete adjunto

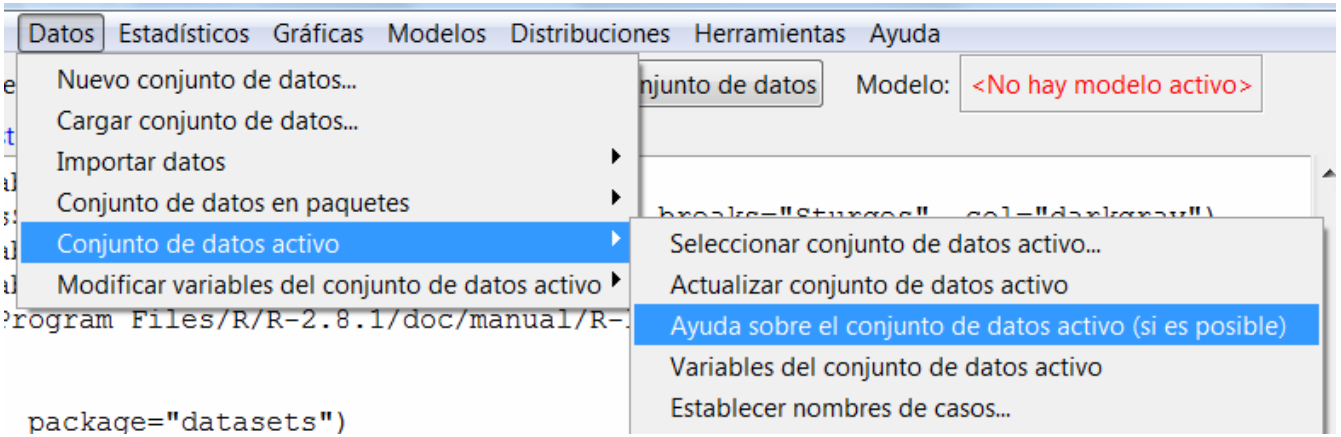




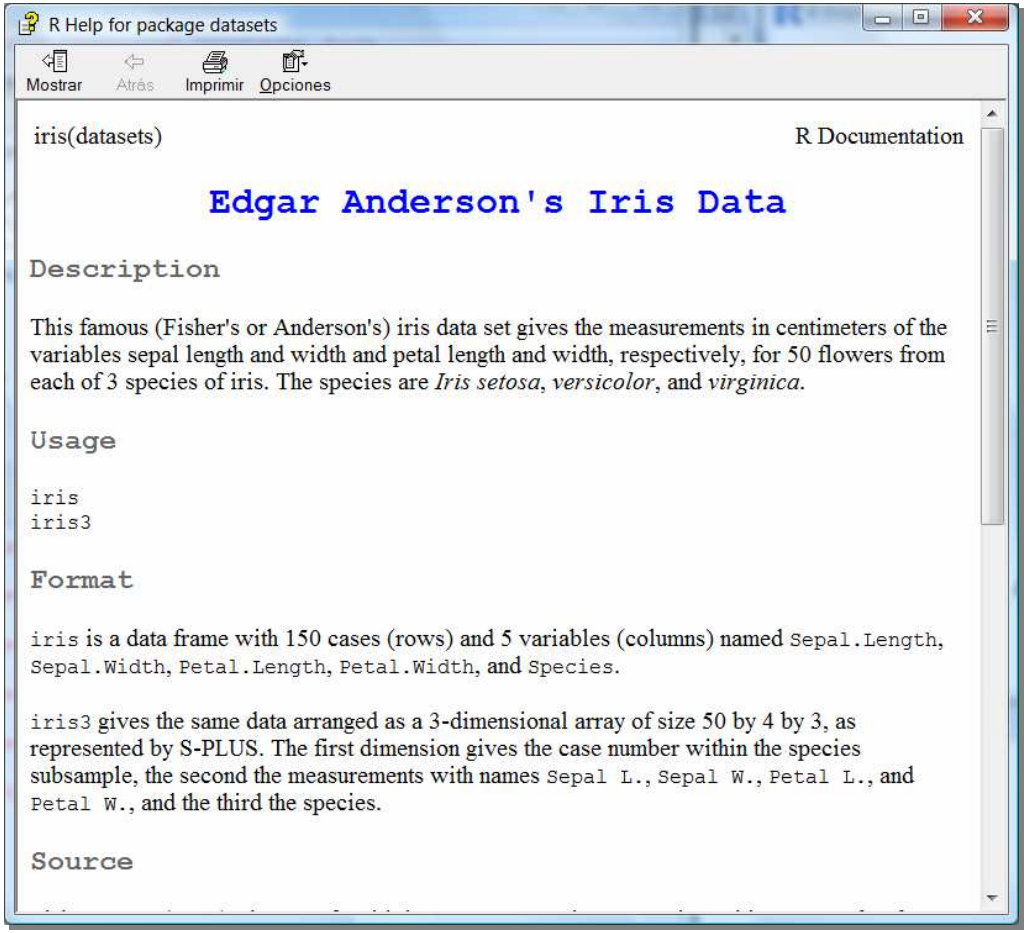
Este famosos conjunto de datos (Fisher's or Anderson's) corresponde a las medidas en centímetros de 4 variables (longitud y anchura de pétalos y sépalos) de 3 especies distinta de flores (*Iris setosa*, *Iris versicolor*, e *Iris virginica)*

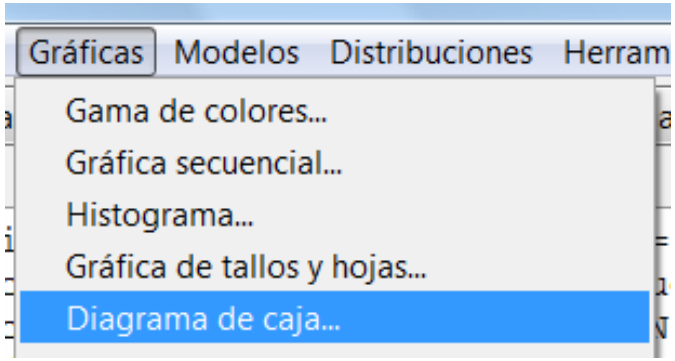
iris

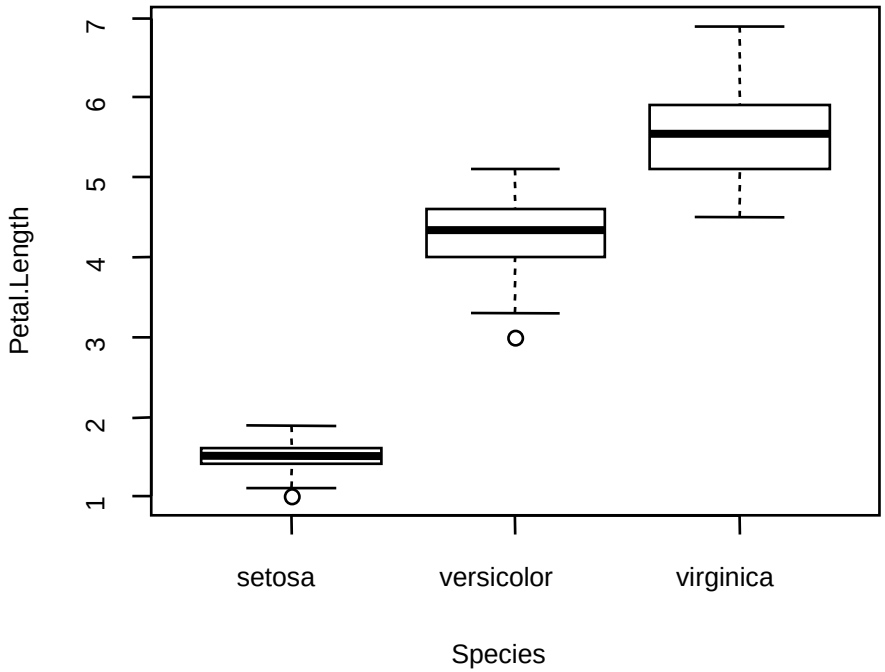
	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa
18	5.1	3.5	1.4	0.3	setosa
19	5.7	3.8	1.7	0.3	setosa
20	5.1	3.8	1.5	0.3	setosa
21	5.4	3.4	1.7	0.2	setosa
22	5.1	3.7	1.5	0.4	setosa
23	4.6	3.6	1.0	0.2	setosa
24	5.1	3.3	1.7	0.5	setosa
25	4.8	3.4	1.9	0.2	setosa
26	5.0	3.0	1.6	0.2	setosa
27	5.0	3.4	1.6	0.4	setosa
28	5.2	3.5	1.5	0.2	setosa
29	5.2	3.4	1.4	0.2	setosa
30	4.7	3.2	1.6	0.2	setosa



Es posible obtener una descripción detallada de los *datasets* disponibles







Diagramas de caja