

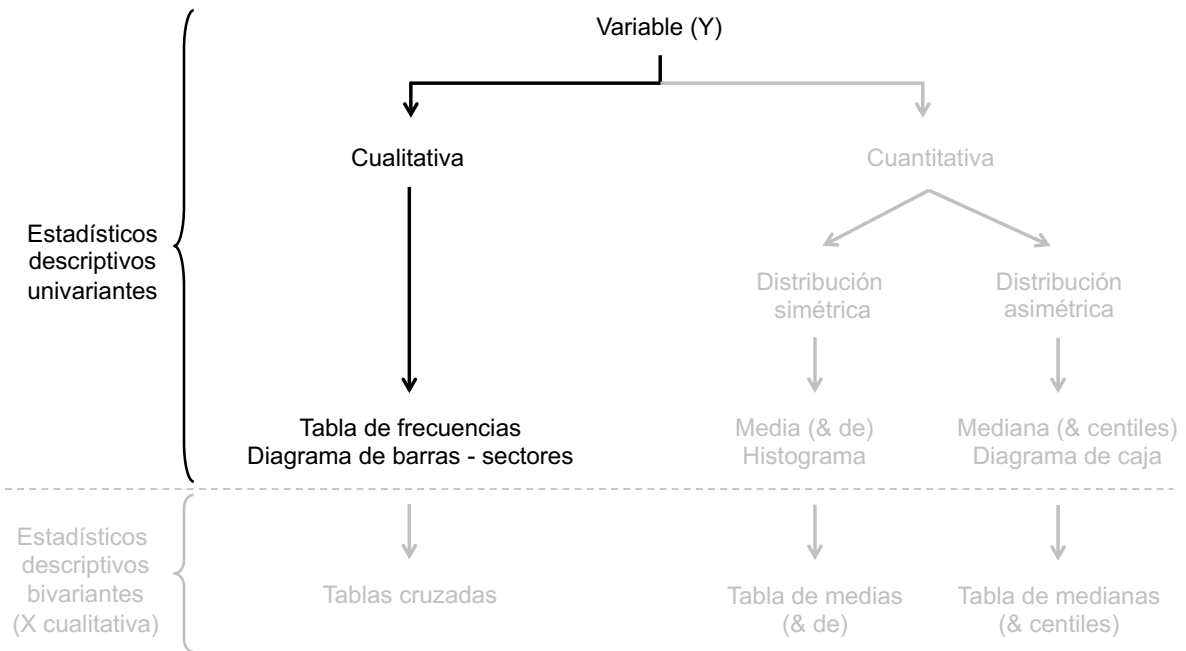
# Stata para usuarios de SPSS

**Aurelio Tobías**

Hospital de la Santa Creu i Sant Pau  
Barcelona, 16/01/2019

## 3. Análisis descriptivo y gráficos

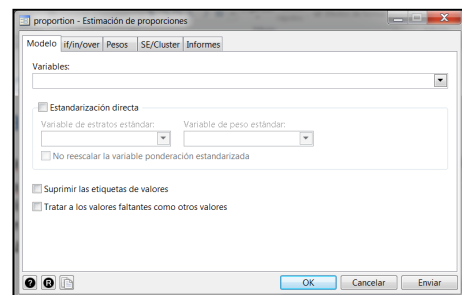
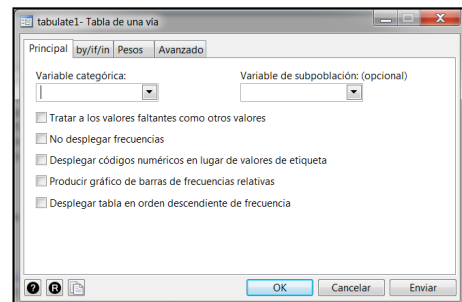
# Estadísticos descriptivos



3

# Descriptivo variables cualitativas

- Tabla de frecuencias
  - > Estadísticas > Sumarios, tablas y tests estadísticos > Tablas de frecuencias > Tabla de frecuencias
  - . **tabulate var**
- Intervalo de confianza para proporciones
  - > Estadísticas > Sumarios, tablas y tests estadísticos > Sumario y estadísticas descriptivas > Proporciones
  - . **proportion vars**



4

```
. use episer, clear
(EpiSER study dataset)
```

```
. tabulate art
```

osteoarthri tis	Freq.	Percent	Cum.
no	323	80.75	80.75
yes	77	19.25	100.00
Total	400	100.00	

```
. proportion art
```

```
Proportion estimation          Number of obs   =          400
-----
          |          Proportion   Std. Err.          [95% Conf. Interval]
-----+-----
art       |
   no    |          .8075   .0197132   .7657614   .8433229
   yes   |          .1925   .0197132   .1566771   .2342386
-----+-----
```

5

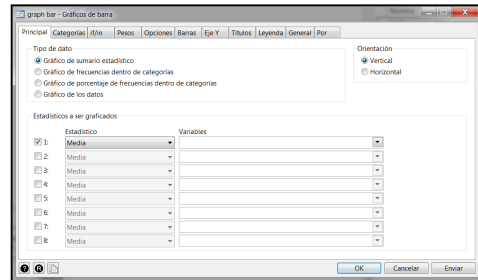
## Gráficos con Stata

- La mayoría de gráficos en **Stata** se realizan con el comando **graph** indicando el tipo de gráfico
  - . **graph bar**            diagrama de barras
  - . **graph pie**            diagrama de sectores
  - . **graph box**            diagrama de caja
  - . **graph twoway**        diagrama de dispersion
- Pero ... otros gráficos pueden tener su propio comando
  - . **histogram**           histograma

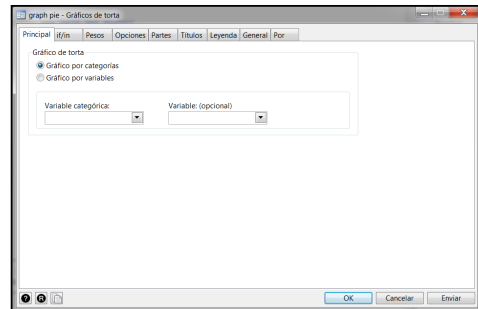
6

# Gráficos para variables cualitativas

- Diagrama de barras
  - > Gráficos > Gráficos de barra
  - . graph bar vars



- Diagrama de sectores
  - > Gráficos > Gráficos de torta
  - . graph pie vars

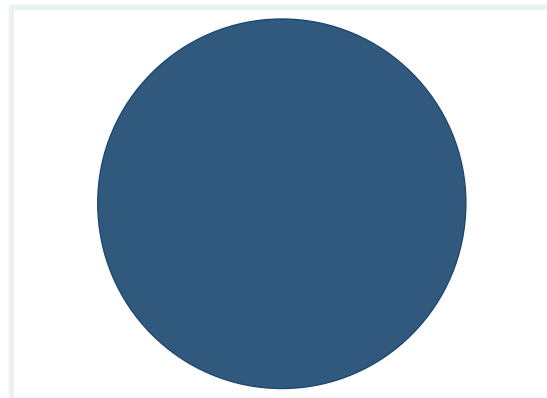


7

. graph bar art



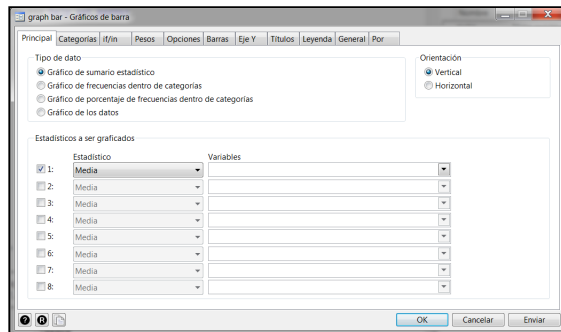
. graph pie art



8

# Diagramas de barras y sectores

- Para dibujar diagramas de barras y sectores en **Stata** es necesario definir tantas variables *dummy* (0/1) como categorías tenga la variable cualitativa
- **No es muy intuitivo para gráficos sencillos (!)** pero es muy eficaz para gráficos más avanzados



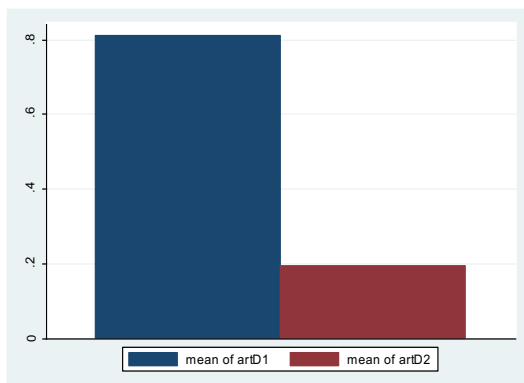
9

```
. tabulate art, generate(artD)  
...
```

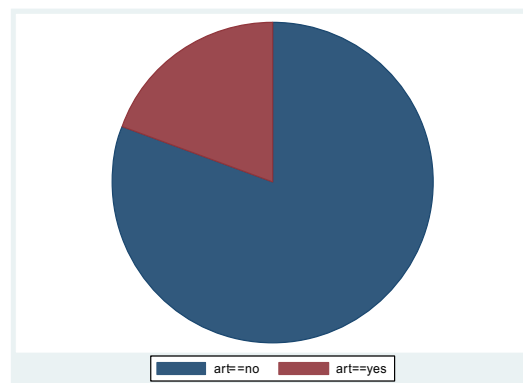
```
. describe artD*
```

variable name	storage type	display format	value label	variable label
artD1	byte	%8.0g		art==no
artD2	byte	%8.0g		art==yes

```
. graph bar artD1 artD2
```

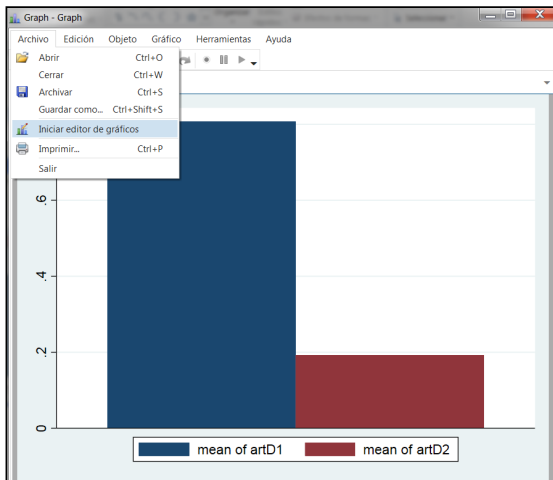


```
. graph pie artD1 artD2
```



10

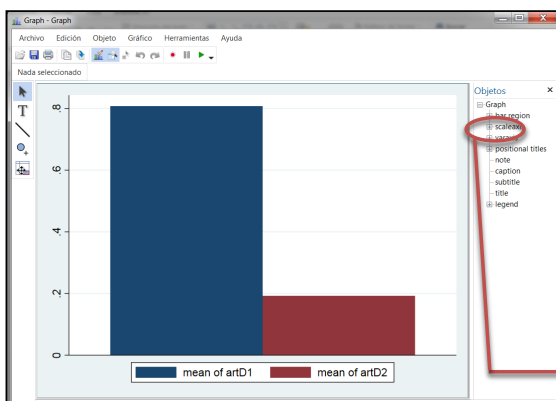
# Edición de gráficos



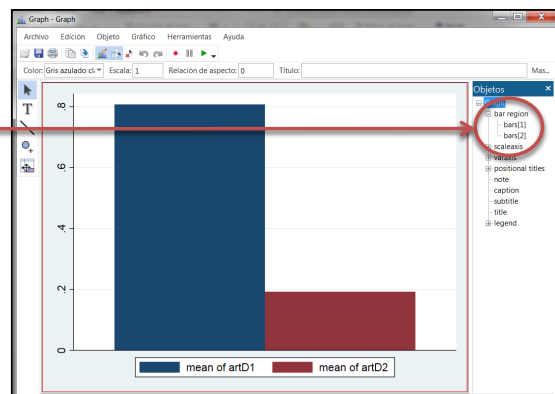
- Los gráficos pueden ser editados usando las opciones en la sintaxis, o de forma manual iniciando el **Editor de Gráficos**
- Todos los cambios que se realizan en el Editor de Gráficos **NO** serán guardados en la **Ventana de Revisión** como sintaxis de **Stata**

11

# Edición de gráficos

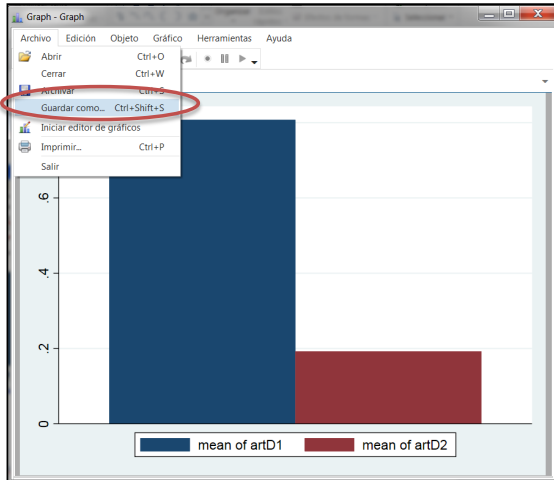


- Todos los atributos del gráfico pueden ser fácil e intuitivamente modificados en la ventana **Objetos**



12

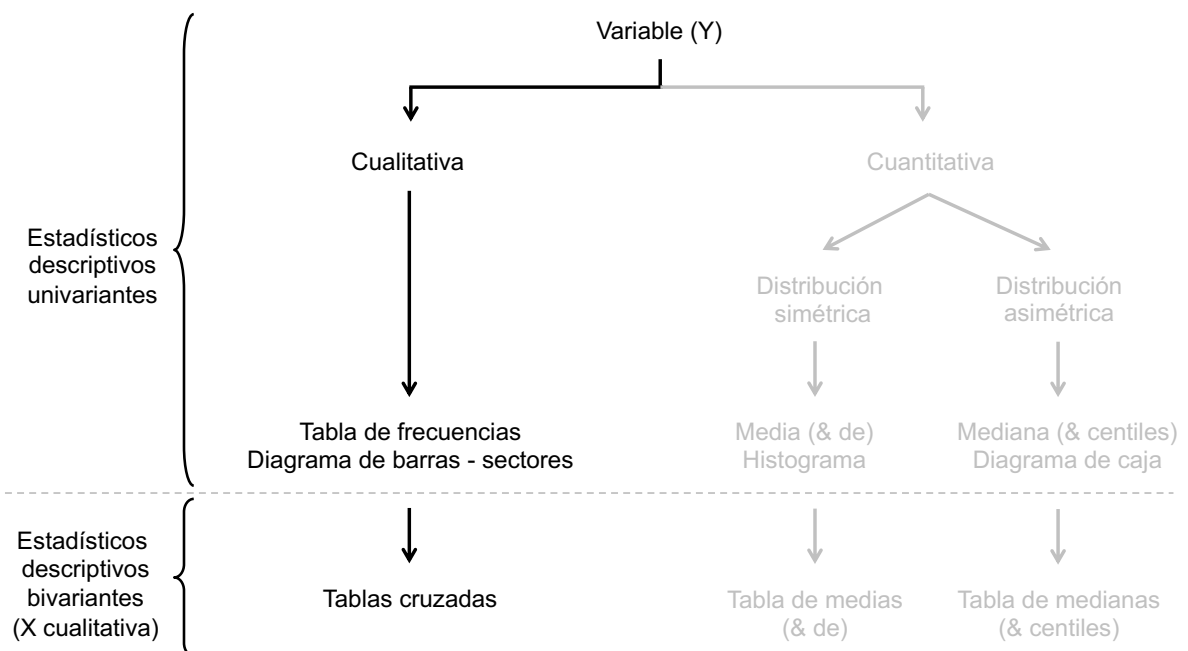
# Guardar gráficos



- Los gráficos pueden guardarse en diversos formatos
  - Gráfico de **Stata** (.gph)
  - Metarchivo Windows (.wmf)
  - PDF (.pdf)
  - PostScript (.ps, eps)
  - Imagen (.jpg, .png, .tif)

13

# Estadísticos descriptivos



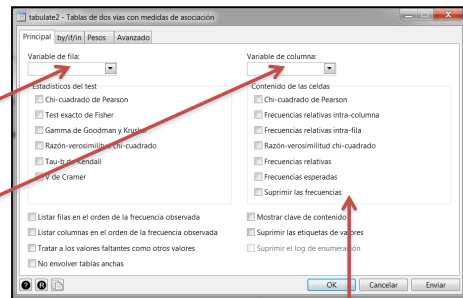
14

# Descriptivo para variables cualitativas

- Tablas cruzadas

> Estadísticas > Sumarios, tablas, y tests estadísticos > Tablas de frecuencias > Tablas de contingencia con medidas de asociación

. tabulate var1 var2



- En las celdas de la tabla cruzada se puede mostrar el % total (**cell**), % fila (**row**), % columna (**col**) y no mostrar frecuencias (**nofreq**)

15

```
. tabulate art sex
osteoarthr |      gender
  itis |    male    female |      Total
-----+-----+-----
    no |    156    167 |    323
    yes |     16     61 |     77
-----+-----+-----
 Total |    172    228 |    400
```

```
. tabulate art sex, col
osteoarthr |      gender
  itis |    male    female |      Total
-----+-----+-----
    no |    156    167 |    323
      |    90.70    73.25 |    80.75
-----+-----+-----
    yes |     16     61 |     77
      |     9.30    26.75 |    19.25
-----+-----+-----
 Total |    172    228 |    400
      |  100.00  100.00 |  100.00
```

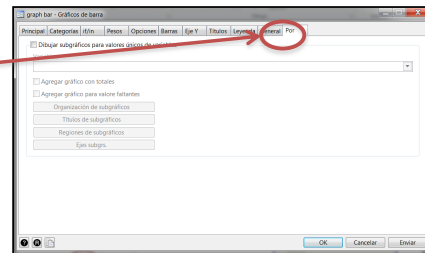
```
. tabulate art sex, col nofreq
osteoarthr |      gender
  itis |    male    female |      Total
-----+-----+-----
    no |    90.70    73.25 |    80.75
    yes |     9.30    26.75 |    19.25
-----+-----+-----
 Total |  100.00  100.00 |  100.00
```

16

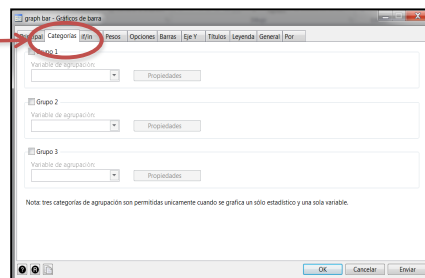


# Gráficos para variables cualitativas

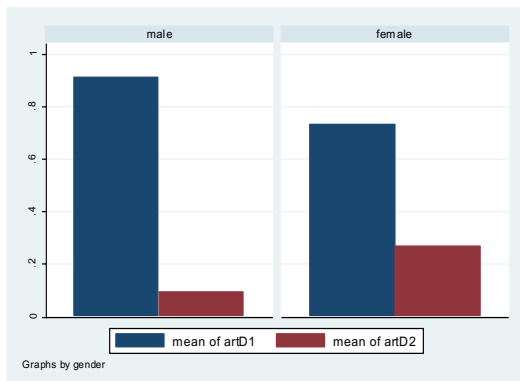
- Diagrama de barras y sectores
  - > Gráficos > Gráficos de barra
  - `graph bar vars, by(var) stack`
  - `by()` para gráficos separados en una ventana



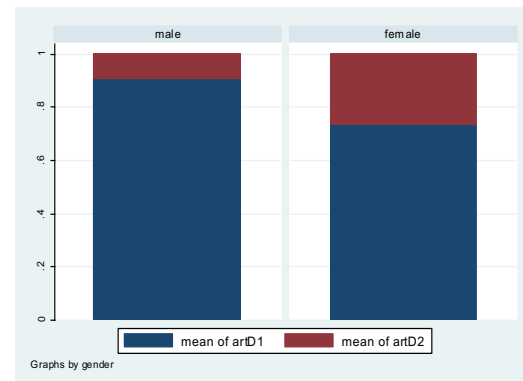
- Diagrama de barras
  - > Gráficos > Gráficos de barra
  - `graph bar vars, over(var)`
  - `over()` para gráficos repetidos en una ventana



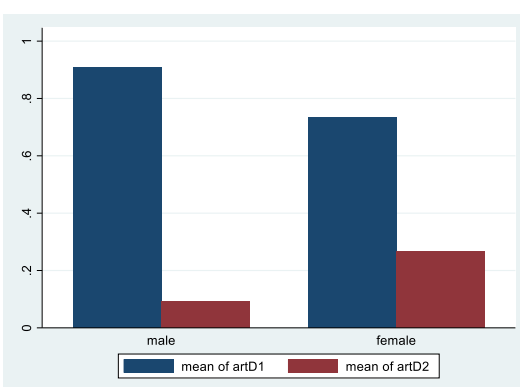
`. graph bar artD1 artD2, by(sex)`



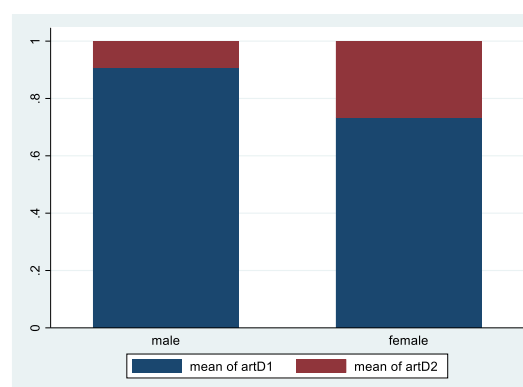
`. graph bar artD1 artD2, by(sex) stack`



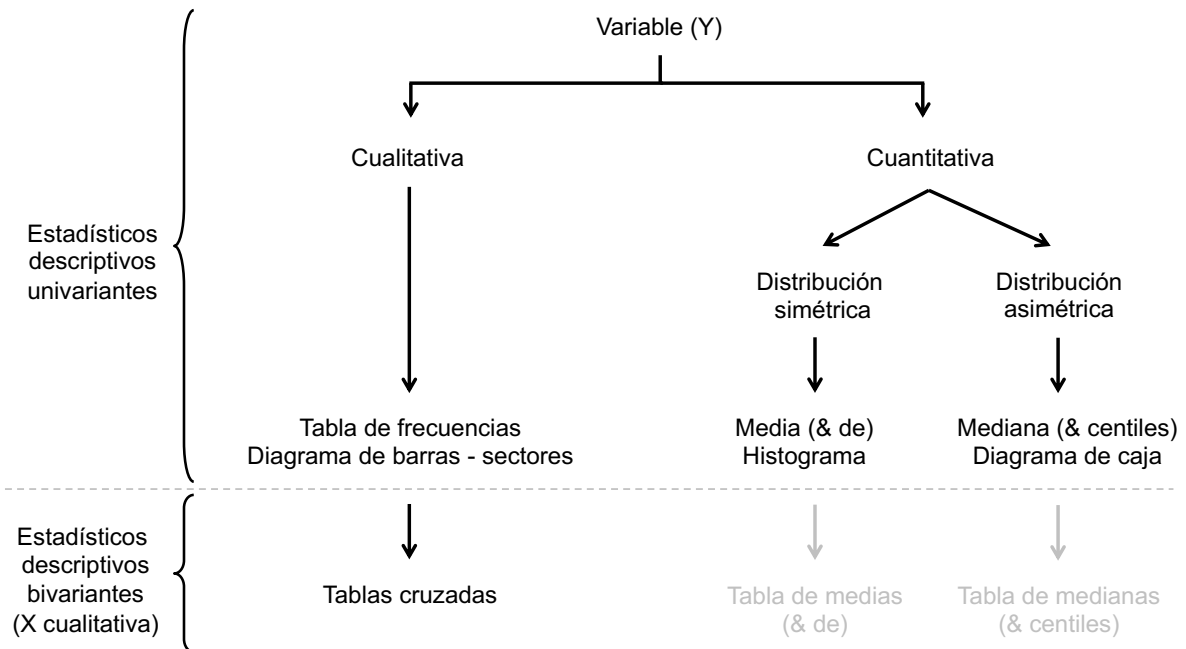
`. graph bar artD1 artD2, over(sex)`



`. graph bar artD1 artD2, over(sex) stack`



# Estadísticos descriptivos



19

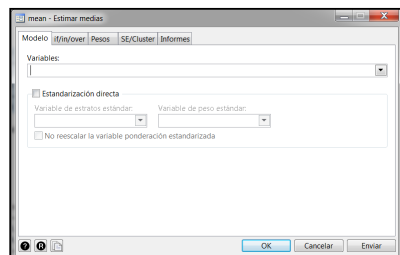
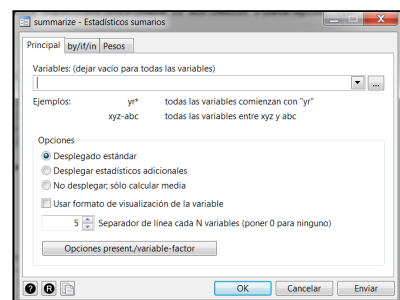
# Descriptivo para variables cuantitativas

- Media y desviación estándar  
 > Estadísticas > Sumarios, tablas, y tests estadísticos > Sumario y estadísticas descriptivas > Sumario de estadísticas

. summarize vars, detail

- Intervalo de confianza para medias  
 > Estadísticas > Sumarios, tablas, y tests estadísticos > Sumario y estadísticas descriptivas > Medias

. mean vars



20

```
. summarize weight
```

Variable	Obs	Mean	Std. Dev.	Min	Max
weight	400	70.8	13.10685	40	130

```
. summarize weight, detail
```

```
weight (kg)
```

Percentiles	Smallest		
1%	45.5	40	
5%	51.5	42	
10%	54.5	45	Obs 400
25%	62	45	Sum of Wgt. 400
50%	70		Mean 70.8
		Largest	Std. Dev. 13.10685
75%	78	116	
90%	87	120	Variance 171.7895
95%	94	120	Skewness .7235036
99%	111	130	Kurtosis 4.463082

```
. mean weight
```

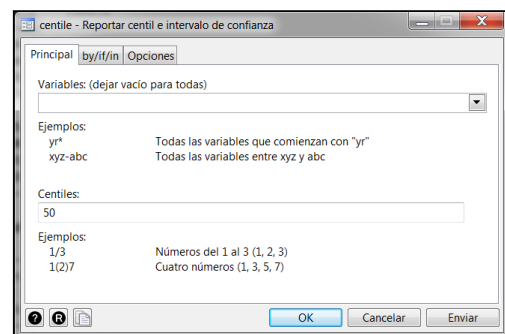
```
Mean estimation          Number of obs =          400
```

	Mean	Std. Err.	[95% Conf. Interval]
weight	70.8	.6553424	69.51164 72.08836

21

## Descriptivo para variables cuantitativas

- Centiles
  - > Estadísticas > Sumarios, tablas, y tests estadísticos > Sumario y estadísticas descriptivas > Centiles con ICs
  - . centile vars, centile()



22

**. centile weight**

Variable	Obs	Percentile	Centile	-- Binom. Interp. -- [95% Conf. Interval]	
weight	400	50	70	69	72

**. centile weight, c(33 66)**

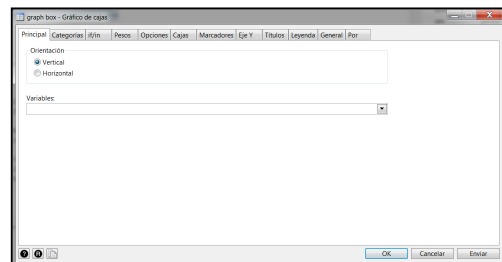
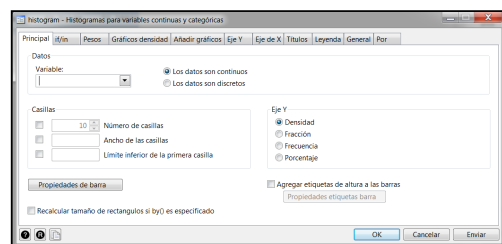
Variable	Obs	Percentile	Centile	-- Binom. Interp. -- [95% Conf. Interval]	
weight	400	33	65	63	65.08834
		66	75	74	76

**. centile weight, c(25 50 75)**

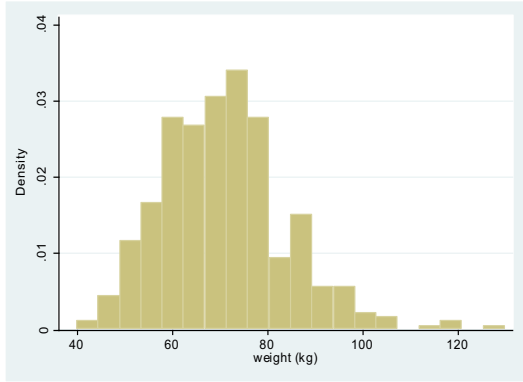
Variable	Obs	Percentile	Centile	-- Binom. Interp. -- [95% Conf. Interval]	
weight	400	25	62	60	63
		50	70	69	72
		75	78	76	80

## Gráficos para variables cuantitativas

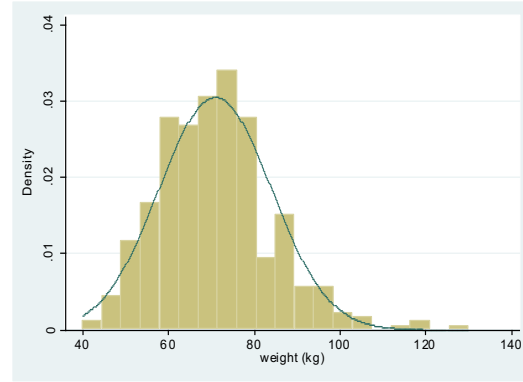
- Histograma
  - > *Gráficos > Histograma*
  - . histogram var, bin(#)**  
**normal**
- Diagrama de caja
  - > *Gráficos > Gráficos de caja*
  - . graph box vars**



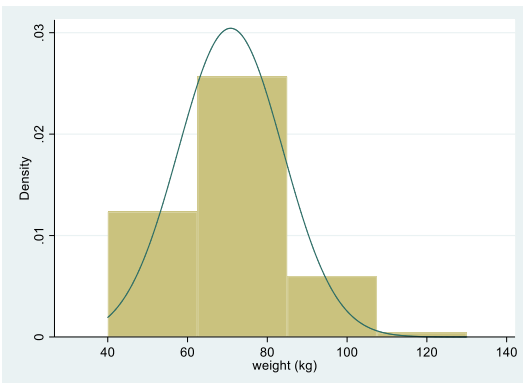
. histogram weight



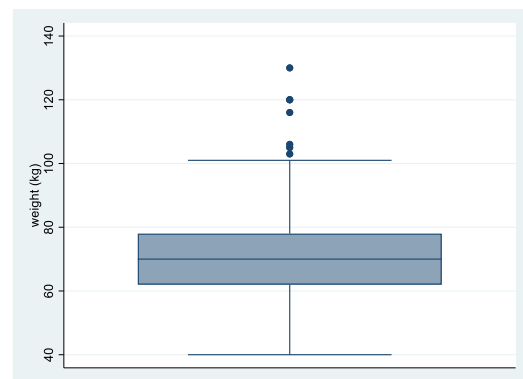
. histogram weight, normal



. histogram weight, normal bin(4)

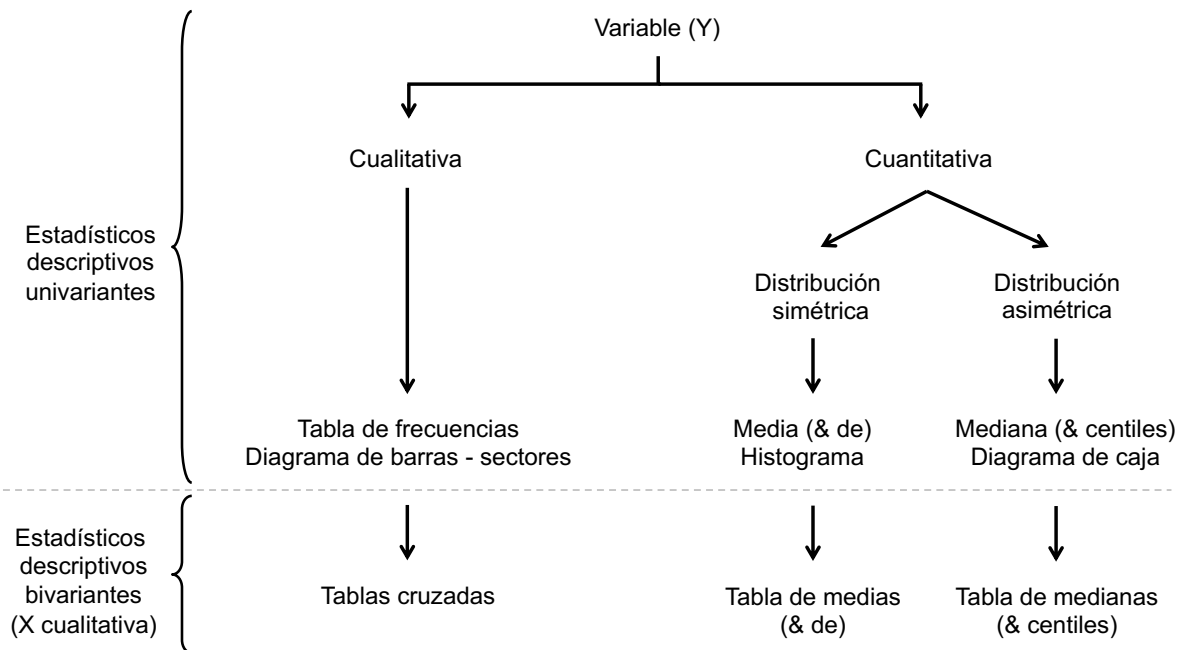


. graph box weight



25

# Estadísticos descriptivos



26

# Descriptivo para variables cuantitativas

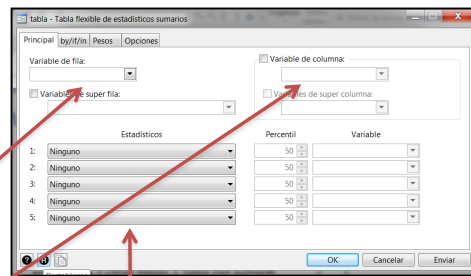
- Tablas cruzadas con estadísticos descriptivos
  - > Estadísticas > Sumarios, tablas, y tests estadísticos > Otras tablas > Tabla de medias, desviación standard y frecuencias
  - . **tabulate var1 var2, summarize(var)**

```
. tabulate art, summarize(weight)
osteoarthritis |      Summary of weight (kg)
      tis |      Mean   Std. Dev.   Freq.
-----+-----
      no |   70.619195   13.112765     323
      yes |   71.558442   13.140429      77
-----+-----
Total |      70.8   13.106848     400

. tabulate art sex, summarize(weight)
      Means, Standard Deviations and Frequencies of weight (kg)
osteoarthritis |      gender
      itis |      male   female |      Total
-----+-----+-----
      no |  77.666667  64.035928 |  70.619195
      |  11.790282  10.651598 |  13.112765
      |      156      167 |      323
-----+-----+-----
      yes |   77.0625  70.114754 |  71.558442
      |   6.7277411  14.042908 |  13.140429
      |      16      61 |      77
-----+-----+-----
Total |  77.610465  65.662281 |   70.8
      |  11.401985  11.931706 |  13.106848
      |      172      228 |      400
```

# Descriptivo para variables cuantitativas

- Tablas flexibles de estadísticos descriptivos
    - Son las tablas más flexibles y complejas. Permiten definir dos niveles en filas y columnas
    - Para cada celda definida permiten calcular hasta 5 estadísticos descriptivos de variables cuantitativas
- > Estadísticas > Sumarios, tablas, y tests estadísticos > Otras tablas > Tabla con sumario estadístico flexible
- `table var1 var2 var3, c(est1 var1 estad2 var2 ...) by (var4)`
  - Estadísticos: `mean sd count ...`



```
. table art, contents(mean weight sd weight)
```

```
-----
osteearth |
ritis     | mean(weight)    sd(weight)
-----+-----
          | no | 70.61919      13.11276
          | yes| 71.55844      13.14043
-----
```

```
. table art sex, c(mean weight sd weight)
```

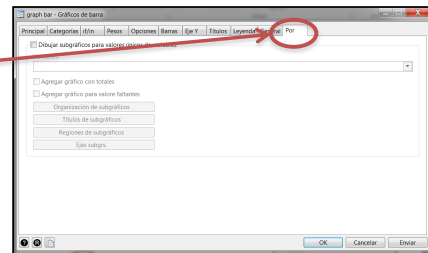
```
-----
osteearth |      gender
ritis     |      male      female
-----+-----
          | no | 77.66666  64.03593
          |   | 11.79028  10.6516
          | yes| 77.0625   70.11475
          |   | 6.727741  14.04291
-----
```

```
. table art area sex, contents(mean weight sd weight)
```

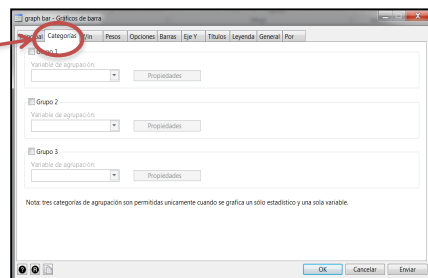
```
-----
osteearth |      gender and area of residence
ritis     | ----- male -----   ----- female -----
          |      urban      rural      urban      rural
-----+-----
          | no | 77.3271  78.40816  64.01667  64.08511
          |   | 11.71821 12.03453  10.48888  11.17223
          | yes| 78       75.5     71.13333  67.25
          |   | 7.803133 4.636809  14.81185  11.54989
-----
```

# Gráficos para variables cualitativas

- Diagrama de barras y sectores  
> Gráficos > Gráficos de barra  
• **graph bar vars, by(var) stack**  
**by ()** para gráficos separados en una ventana



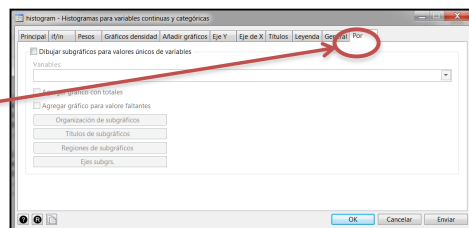
- Diagrama de barras  
> Gráficos > Gráficos de barra  
• **graph bar vars, over(var)**  
**over ()** para gráficos repetidos en una ventana



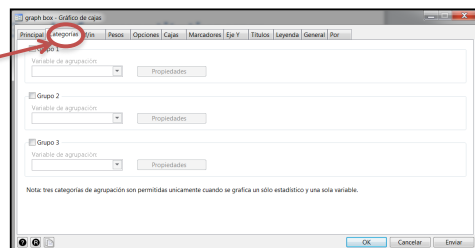
31

# Gráficos para variables cuantitativas

- Histograma y diagrama de cajas  
> Gráficos > Histograma  
• **histogram var, by(var) bin(#) normal**  
**by ()** para gráficos separados en una ventana



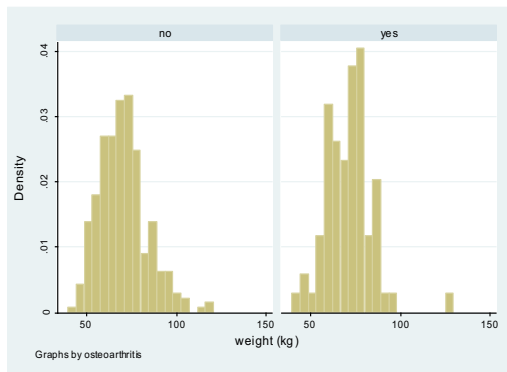
- Diagrama de caja  
> Gráficos > Gráficos de caja  
• **graph box vars, over(var)**  
**over ()** para gráficos repetidos en una ventana



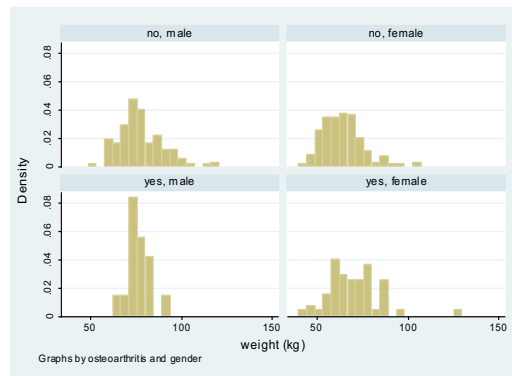
32



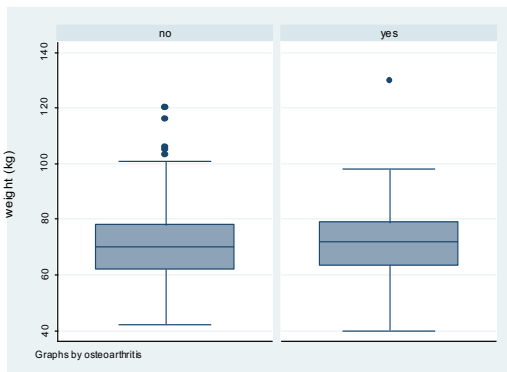
```
. histogram weight, by(art)
```



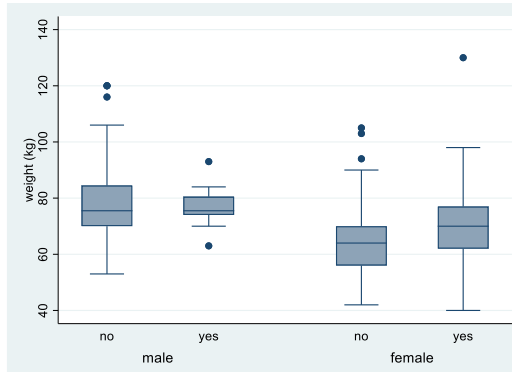
```
. histogram weight, by(art sex)
```



```
. graph box weight, by(art)
```



```
. graph box weight, over(art sex)
```



33

## Resumen

- Principales comandos para análisis descriptivo
  - . tabulate, proportion
  - . summarize, mean, centile
  - . table
- Principales comandos para gráficos
  - . graph [bar | pie | box]
  - . histogram

34