

STATA para Ciencias Sociales y Gestión Pública

Sesión 3: Visualización de datos

César Mora Ruiz

QLAB - PUCP

Julio de 2022

Temas a abordar

Esta sesión dará a conocer los procedimientos para elaborar gráficos informativos sistematizados de información, teniendo en cuenta su naturaleza discreta o continua.

Las aplicaciones se realizarán usando la Encuesta Nacional de Docentes (ENDO) del año 2018.

Los temas a tratar son:

- Gráficos de una variable
- Gráficos de dos variables (continua y discreta)
- Gráficos de dos variables continuas
- Gráficos combinados
- Gráficos de tres variables
- Uso del graph editor/recorder
- Mapas coropléticos en Stata

La Encuesta Nacional de Docentes (ENDO)

La Encuesta Nacional de Docentes

Ingresa a: <http://www.minedu.gob.pe/politicas/docencia/encuesta-nacional-a-docentes-endo.php>

Ministerio de Educación

 APRENDIZAJES

 **DOCENCIA** -

Bienestar Docente +

Reforma Magisterial

Encuesta Nacional a Docentes

Formación y capacitación +

Reconocimiento +



ENDO

ENCUESTA
NACIONAL A DOCENTES

DE INSTITUCIONES EDUCATIVAS PÚBLICAS
DE EDUCACIÓN BÁSICA REGULAR

DOCENCIA | **ENCUESTA NACIONAL A DOCENTES**

La Encuesta Nacional a Docentes (ENDO) constituye un esfuerzo muy importante del sector educación para obtener información pertinente y oportuna de los docentes de instituciones educativas de

La Encuesta Nacional de Docentes

- Esta encuesta recoge información de los docentes de instituciones educativas de Educación Básica Regular públicas y privadas a nivel nacional, sobre sus condiciones de vida, empleo, entre otras.
- Esta información permitirá conocer mejor los perfiles de los docentes, así como sus necesidades con el objetivo de diseñar programas y políticas a su favor.
- Se puede revisar mayor detalle en su respectiva Ficha técnica, disponible en:

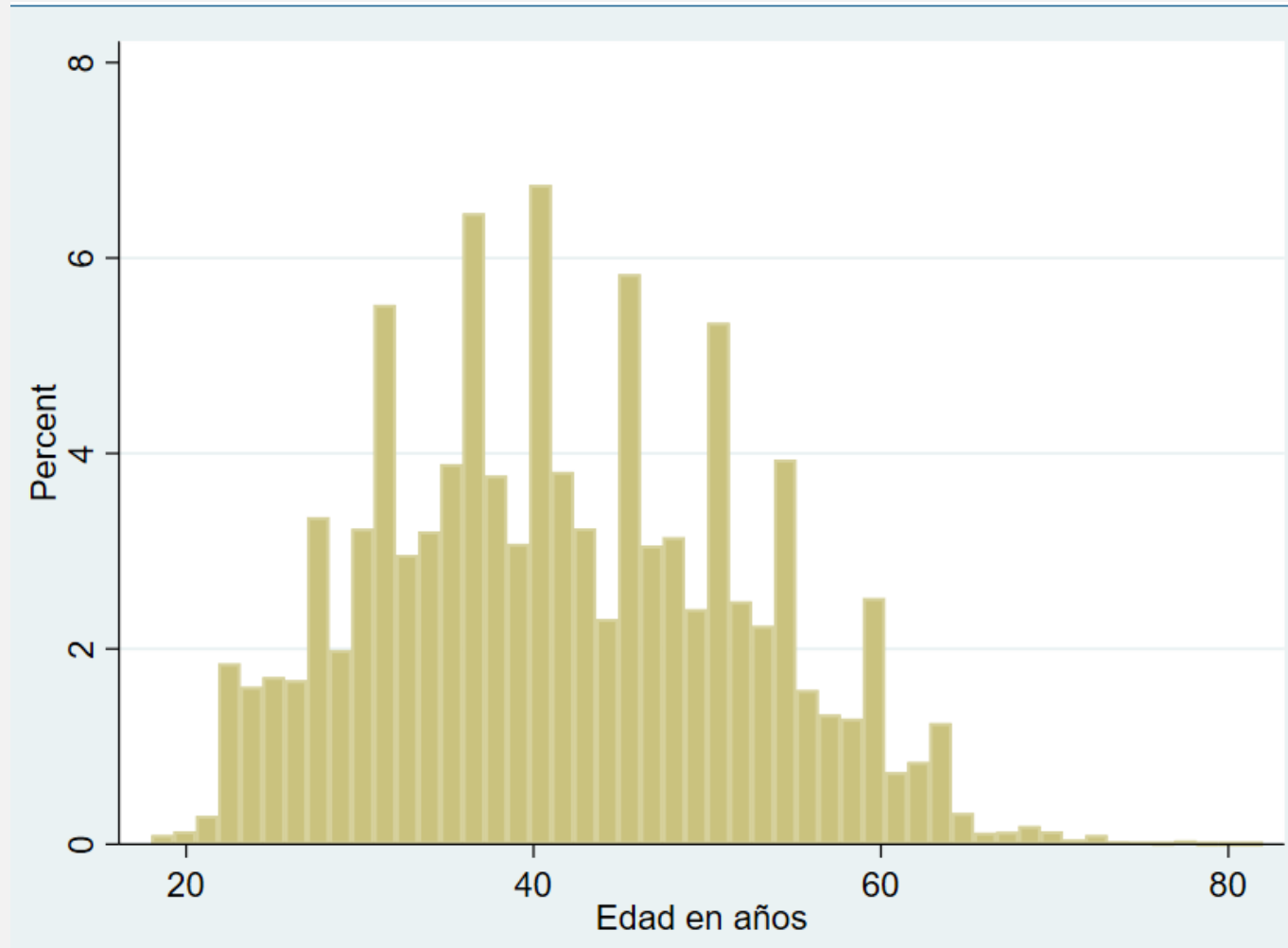
<http://escale.minedu.gob.pe/documents/10156/5384052/FICHA+TECNICA+ENDO+2018.pdf>

Gráficos de una variable

Gráficos de una variable

Histograma

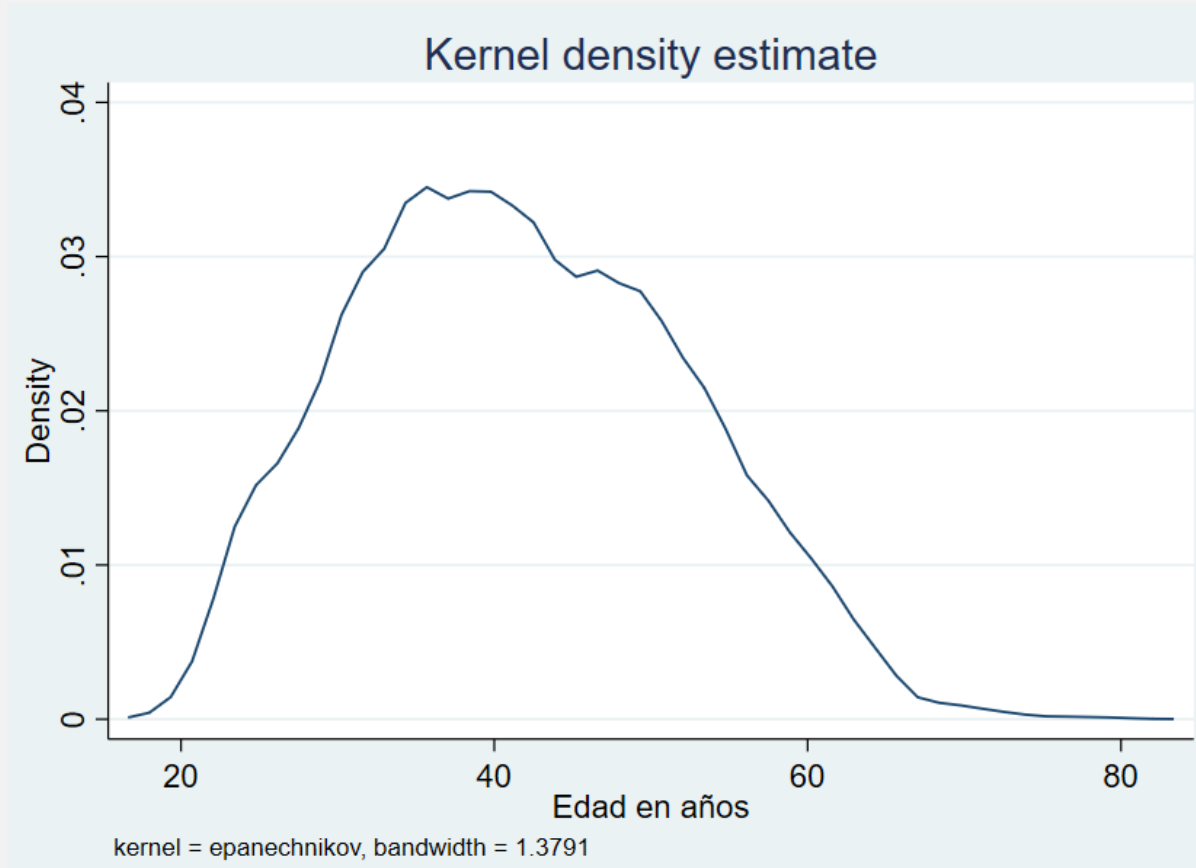
histogram edad_docente, bin(50) percent



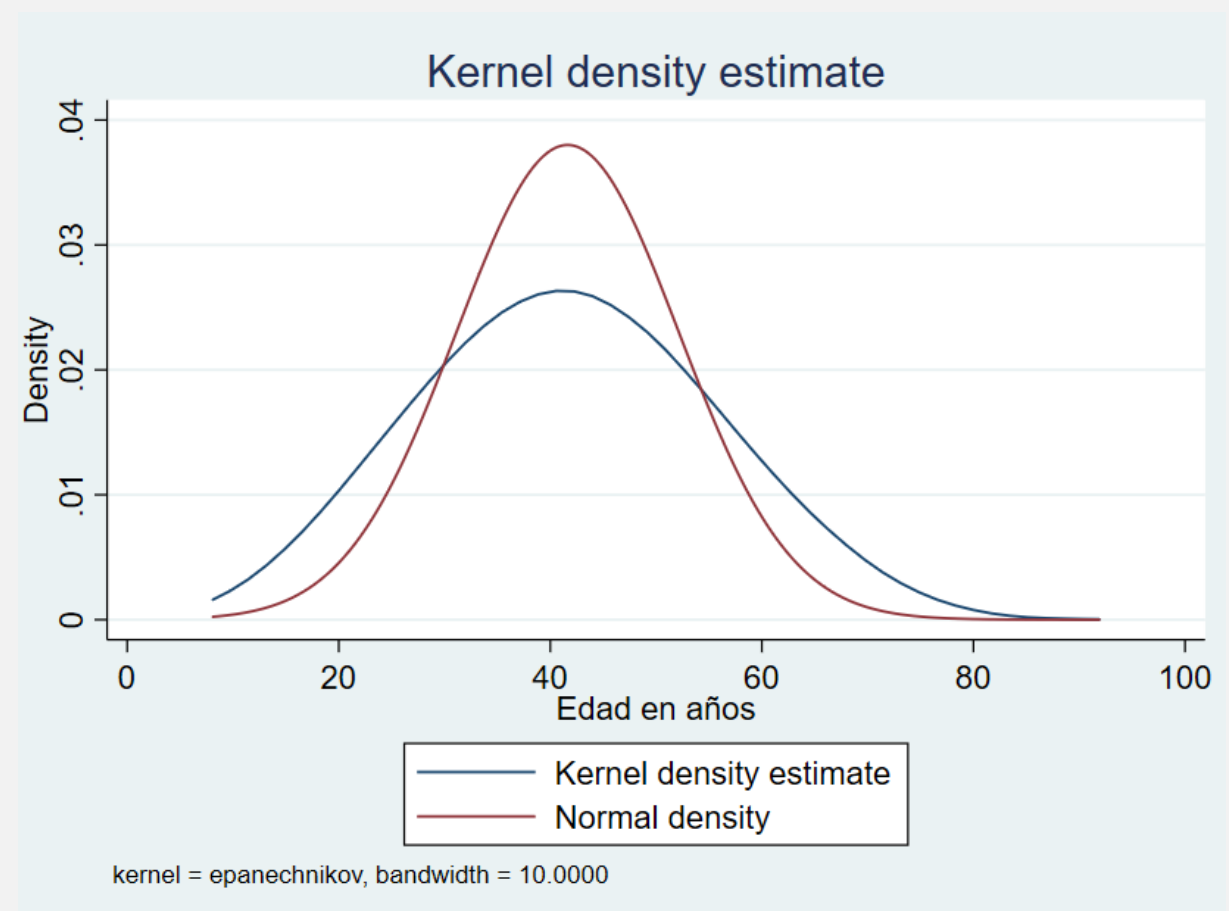
Gráficos de una variable

Densidad Kernel (Epanechnikov)

kdensity edad_docente



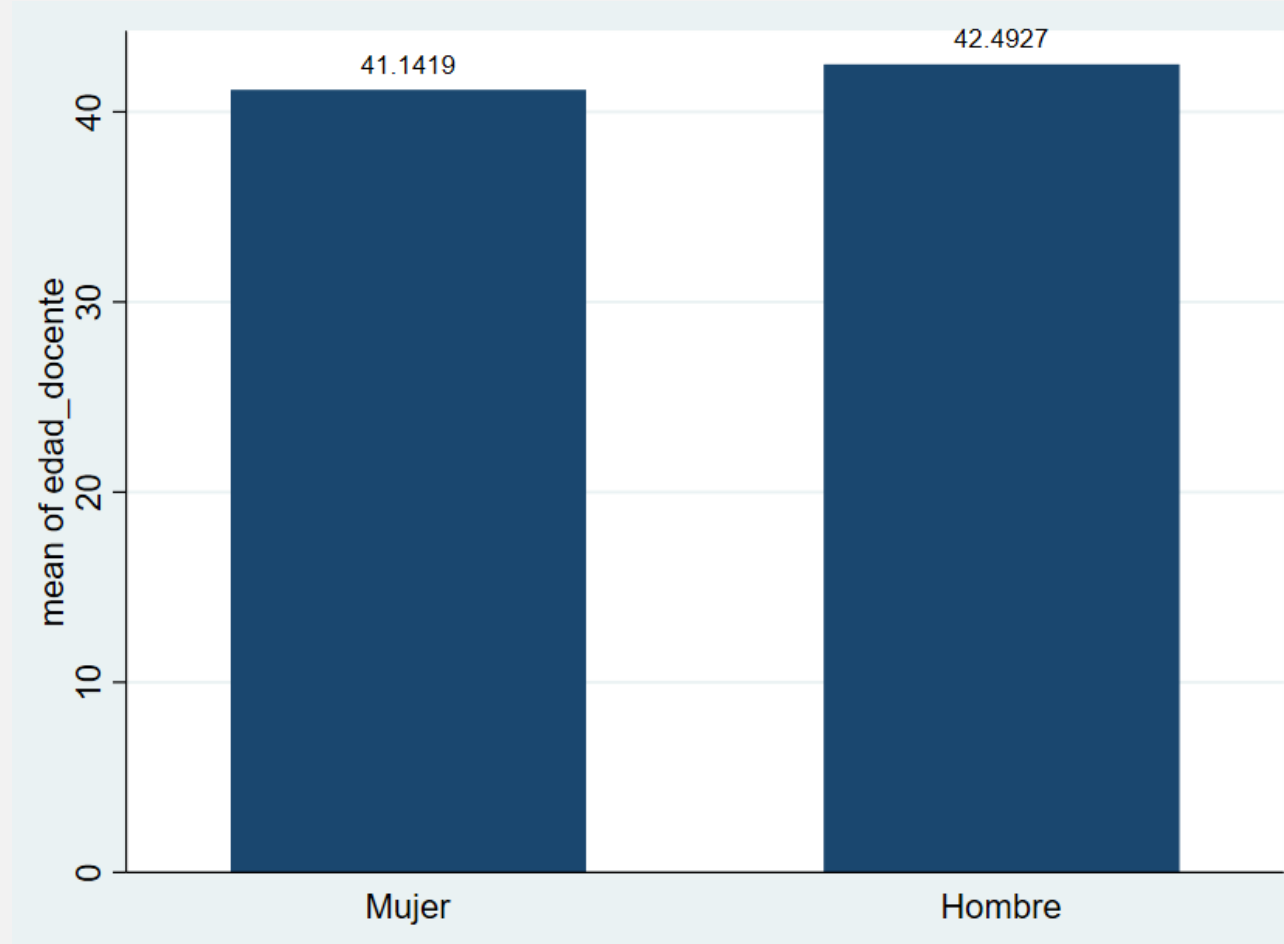
kdensity edad_docente, bwidth(10) normal



Gráficos de una variable

De barra, sobre categorías:

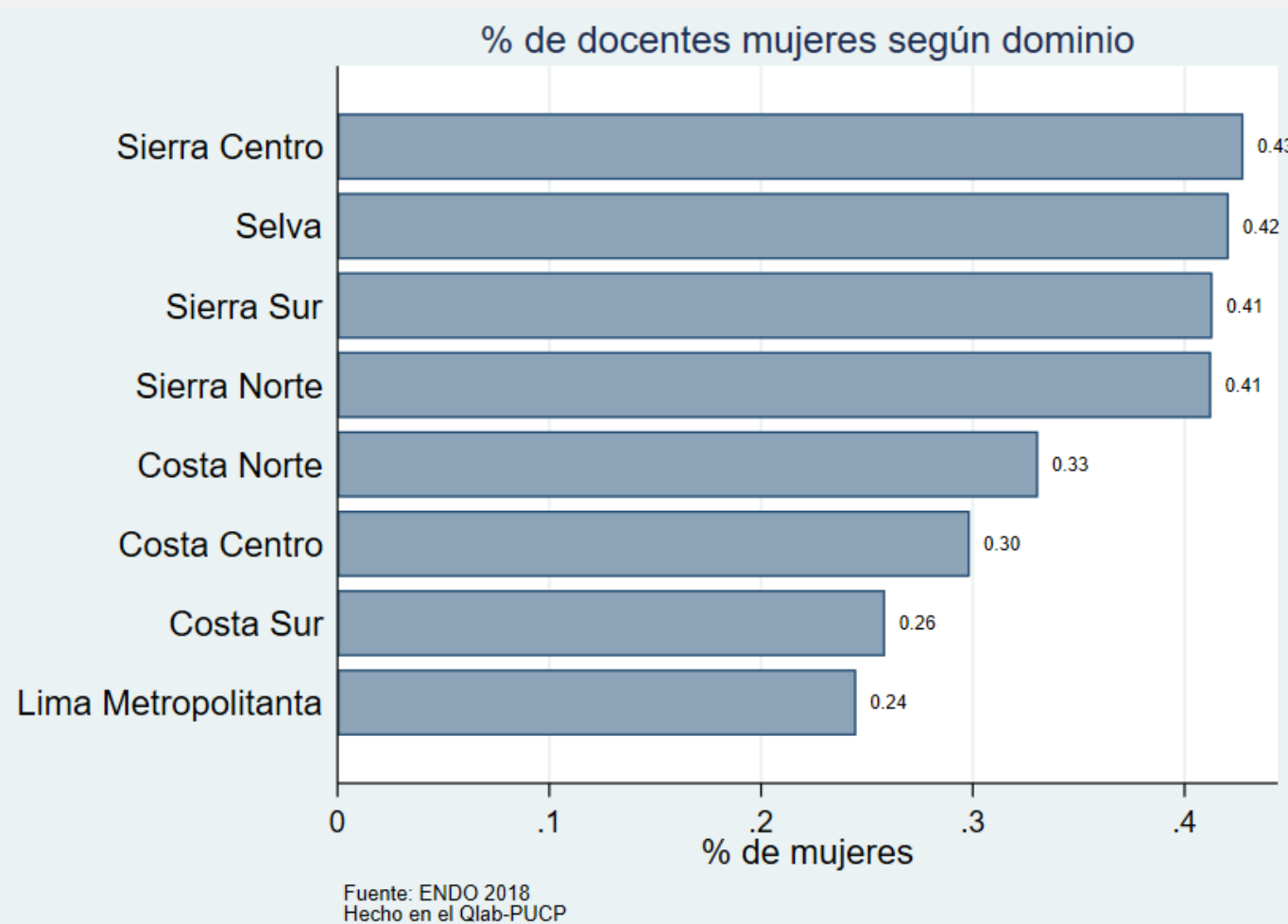
General simple: `graph bar (stat) var1, over(var2) blabel(bar)`



Gráficos de una variable

De barra horizontales, sobre categorías:

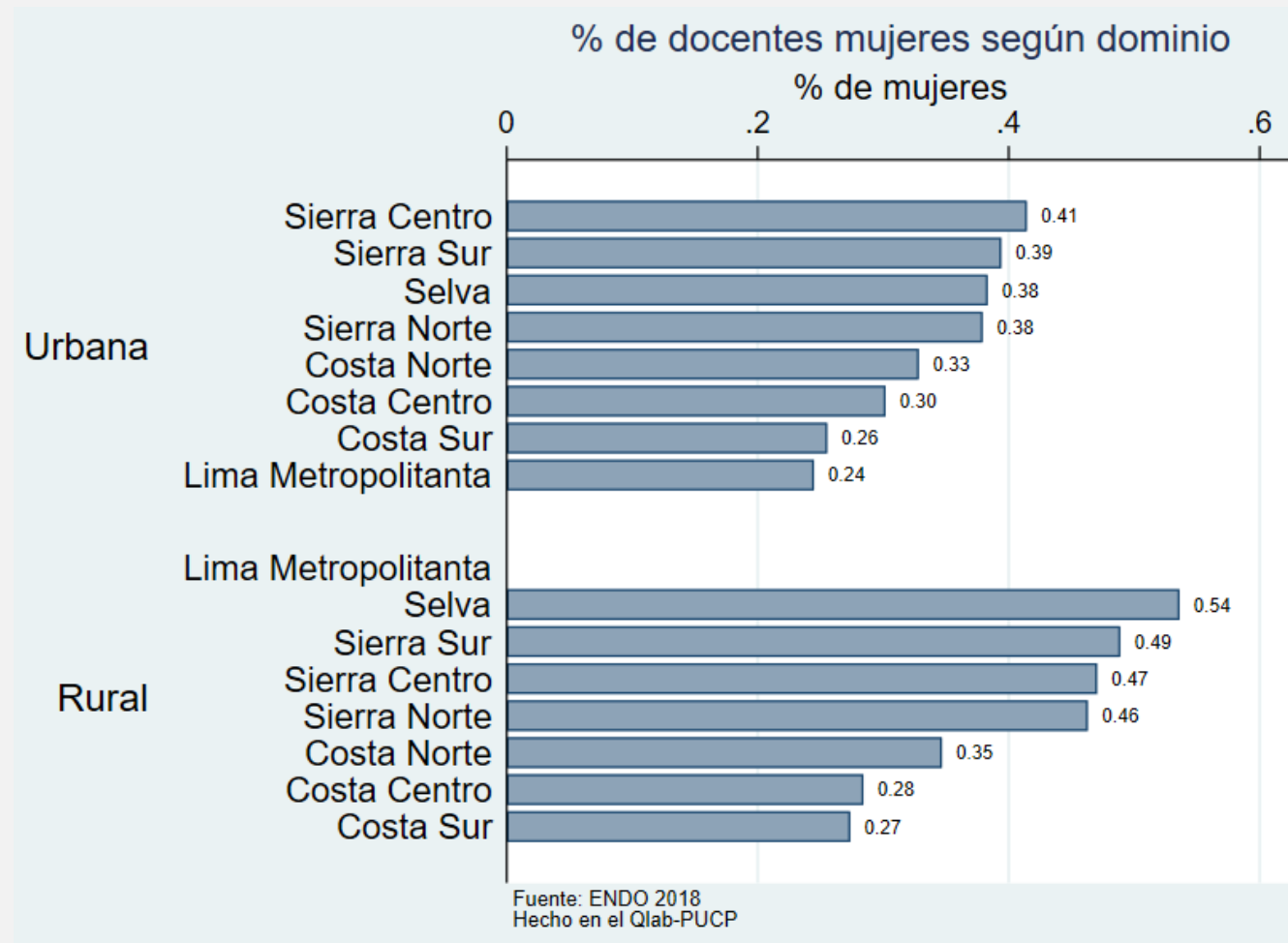
General simple: `graph hbar (stat) var1, over(var2) blabel(bar)`



Gráficos de una variable

De barra horizontales, sobre 2 categorías:

General simple: `graph hbar (stat) var1, over(var2) over (var3)`

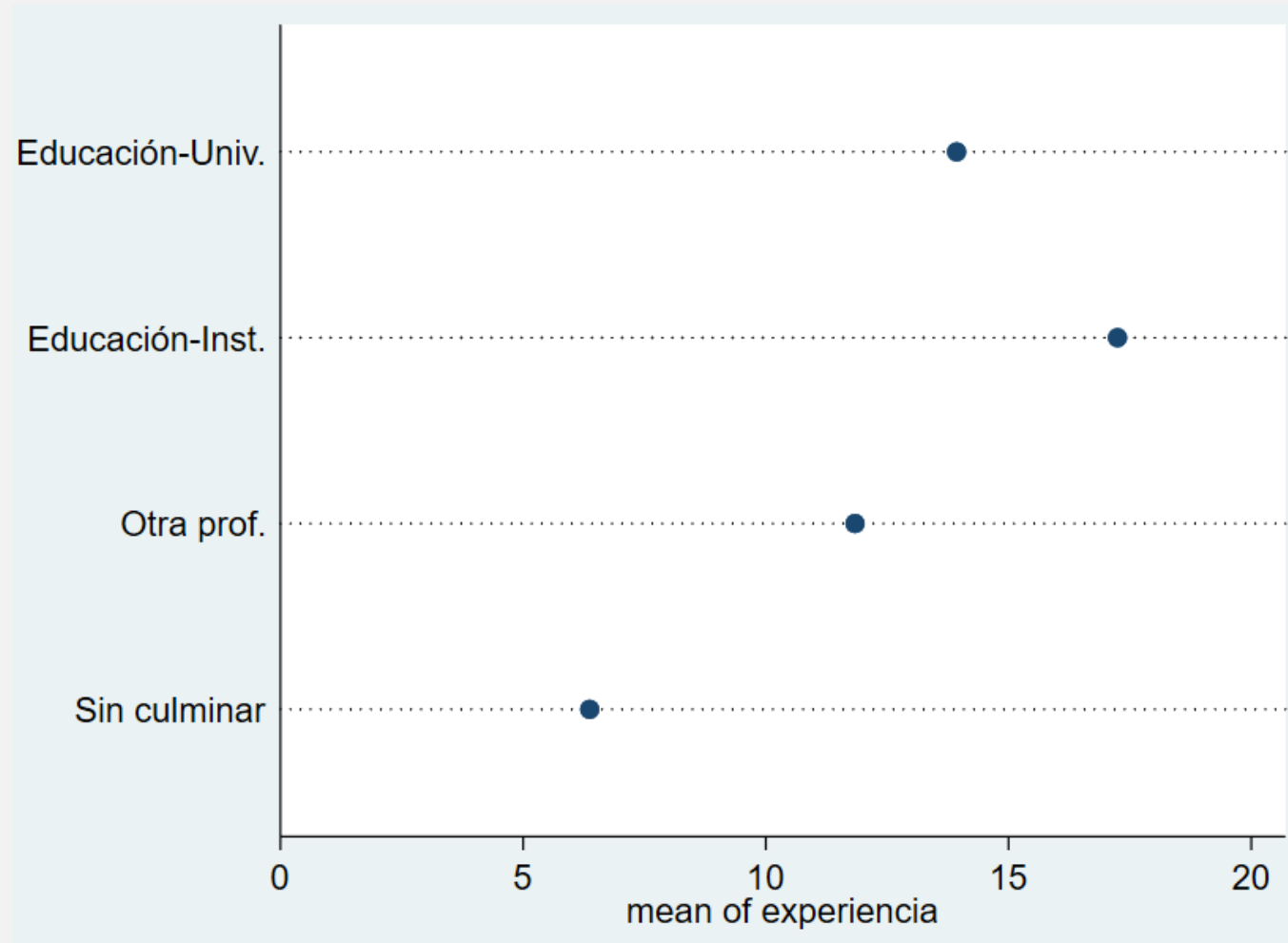


Gráficos de dos variables (continua y discreta)

Gráficos dos variables (continua y discreta)

De puntos:

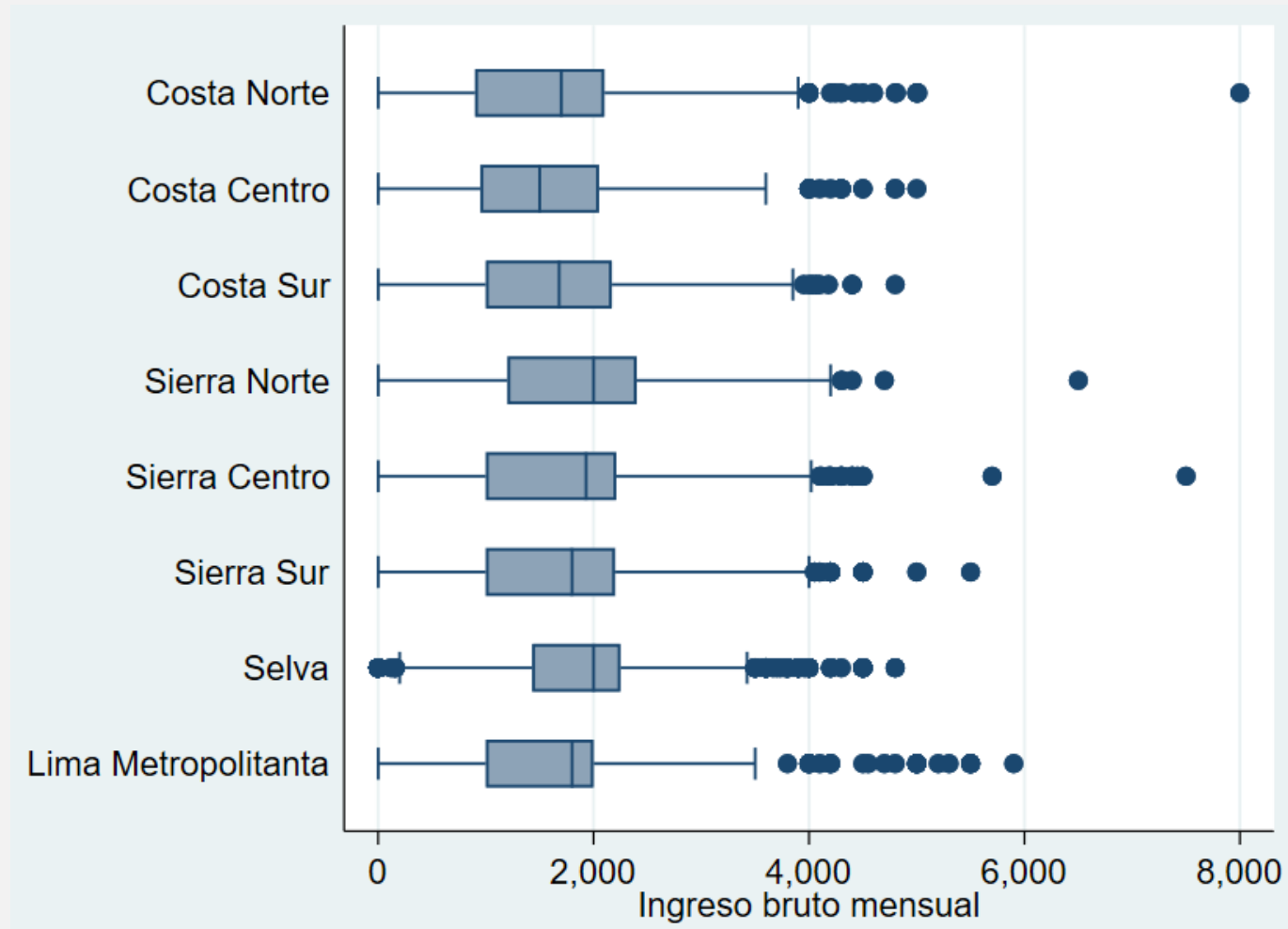
General simple: `graph dot (stat) var1,over(var2)`



Gráficos dos variables (continua y discreta)

De cajas:

General simple: `graph hbox var1,over(var2)`

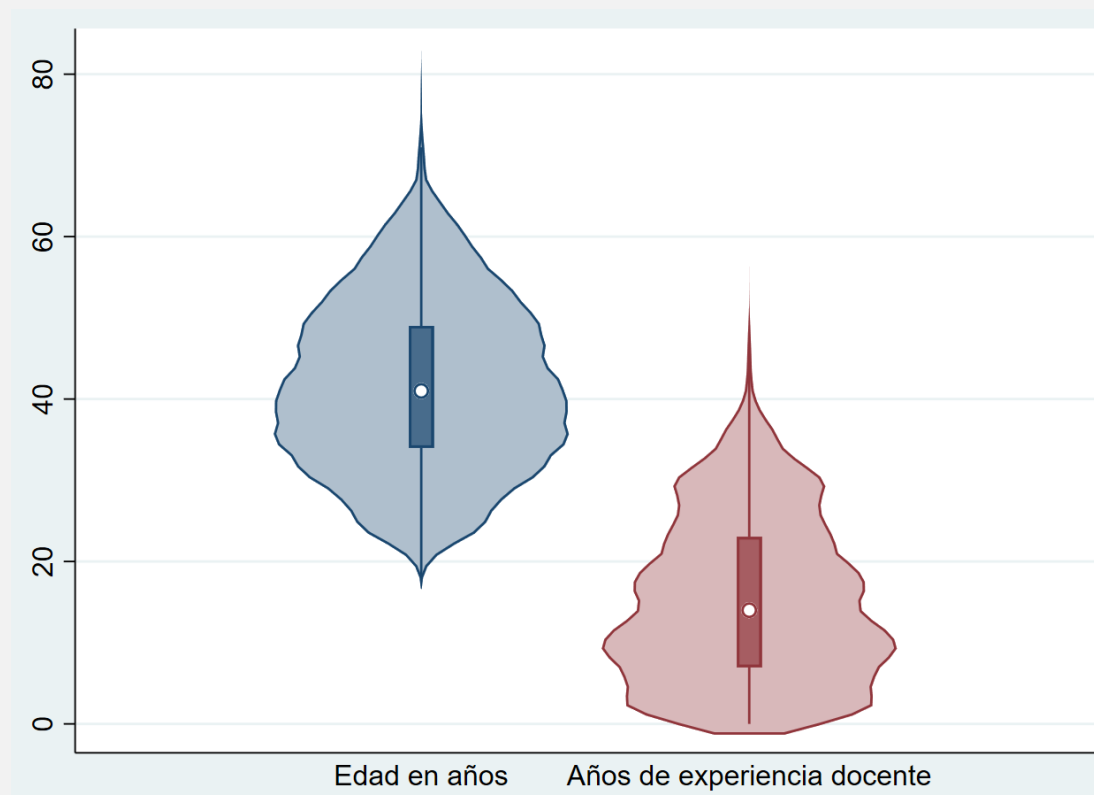


Gráficos dos variables (continua y discreta)

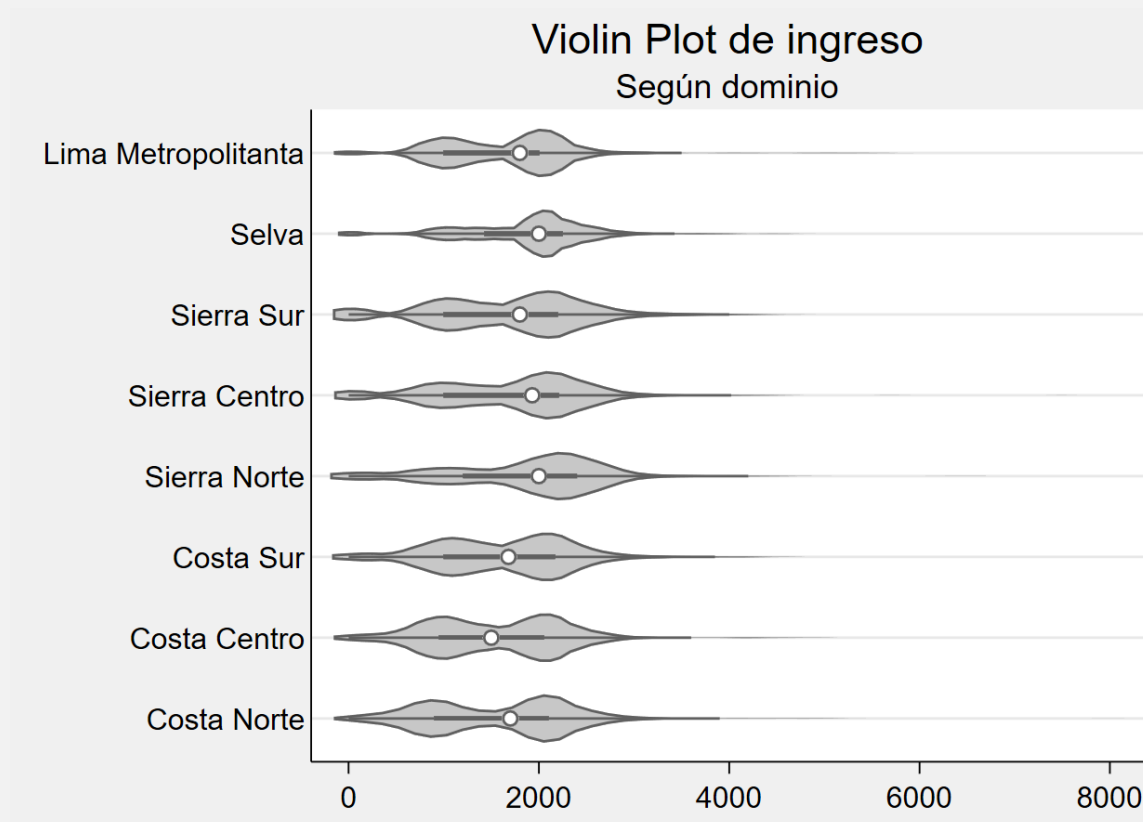
De violín:

General simple: `vioplot var1 var2`

Violín de dos variables



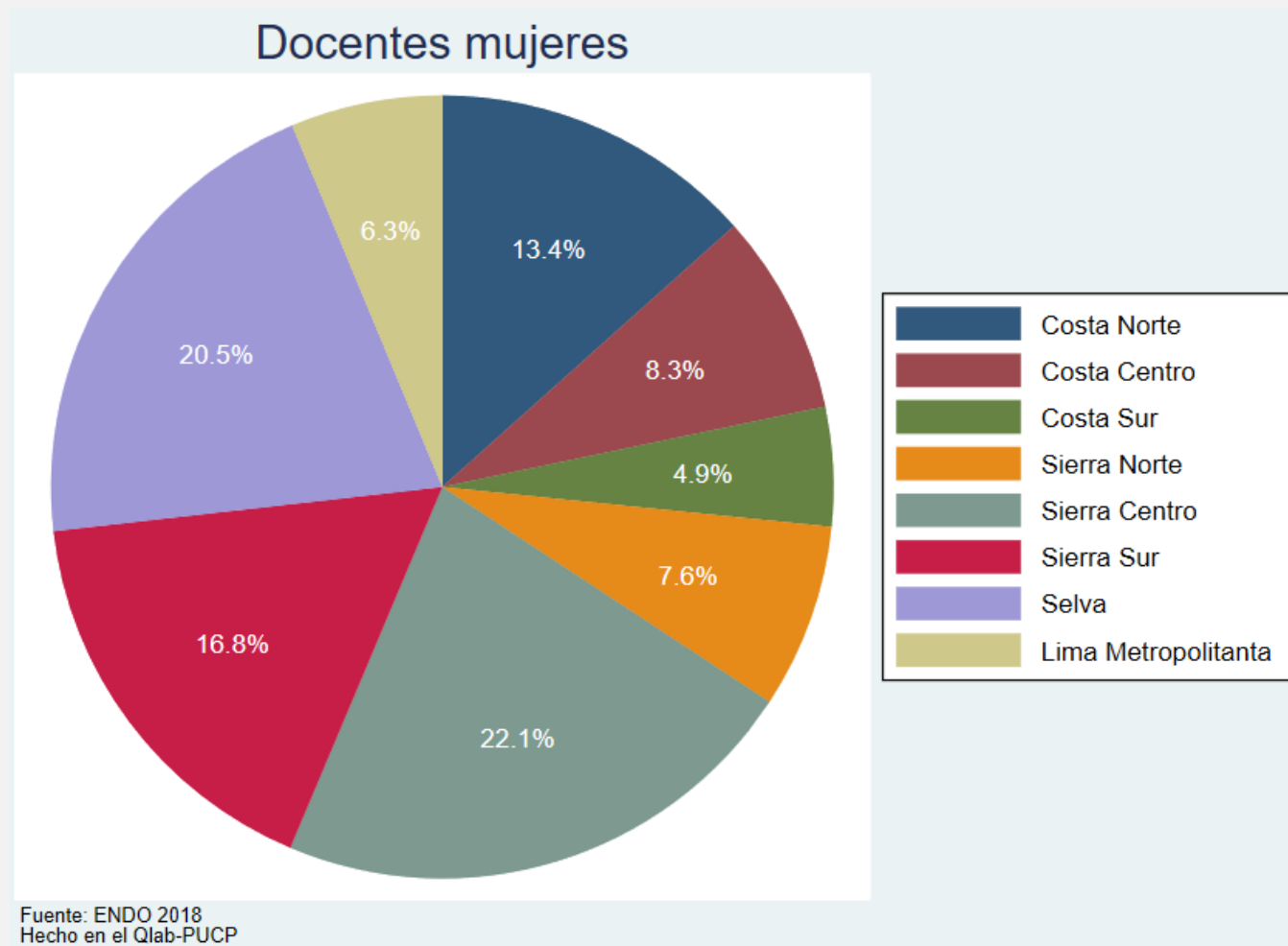
Violín de una var. según categoría



Gráficos dos variables (continua y discreta)

De Pye:

General simple: `graph pie var1, over(var2)`



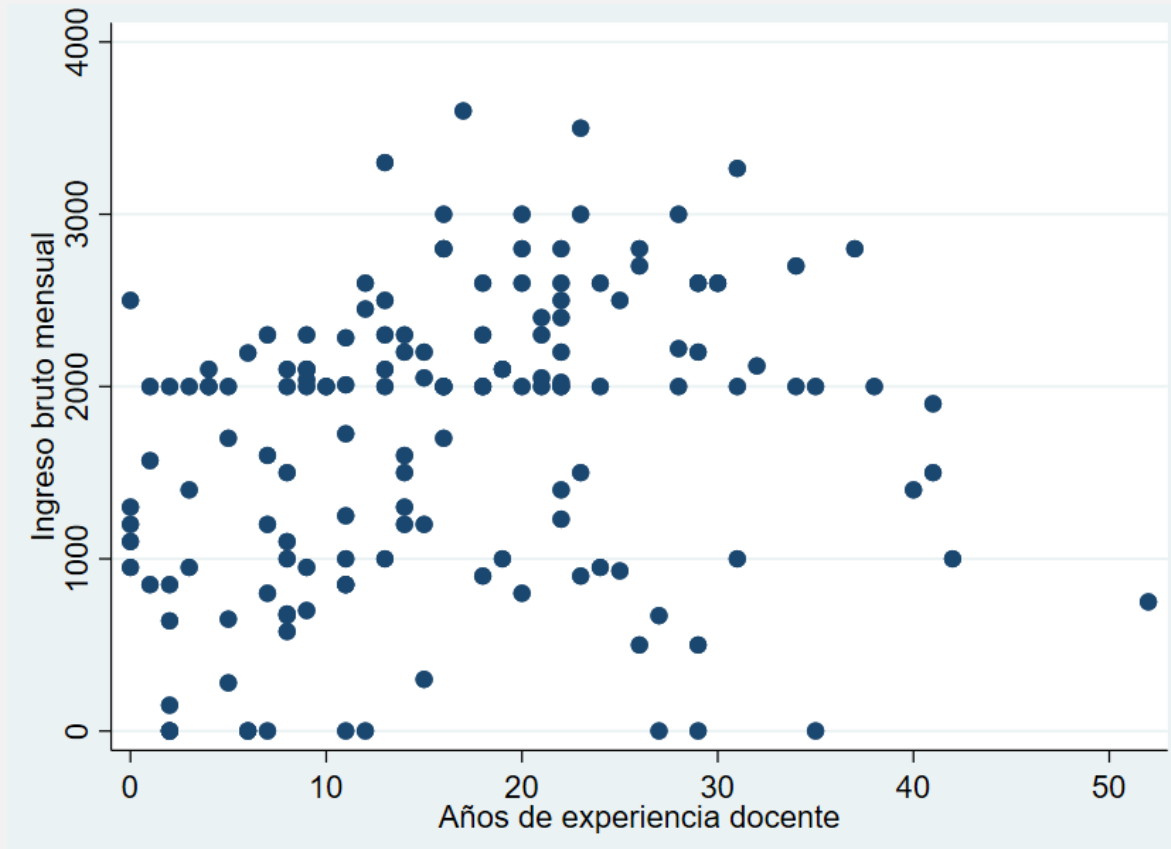
Gráficos de dos variables continuas

Gráficos dos variables continuas

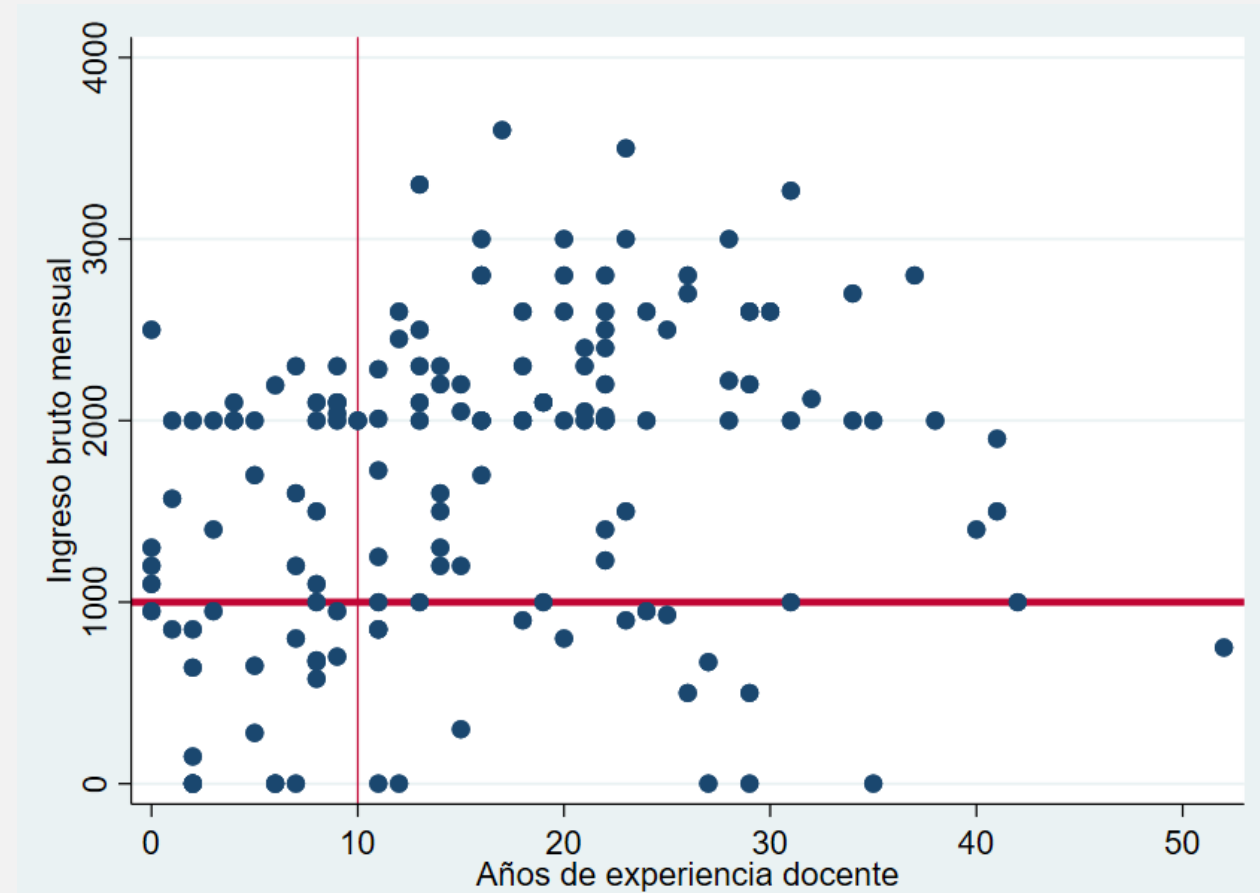
De dispersión:

General simple: `scatter var1 var2`

Sencillo



Con detalles

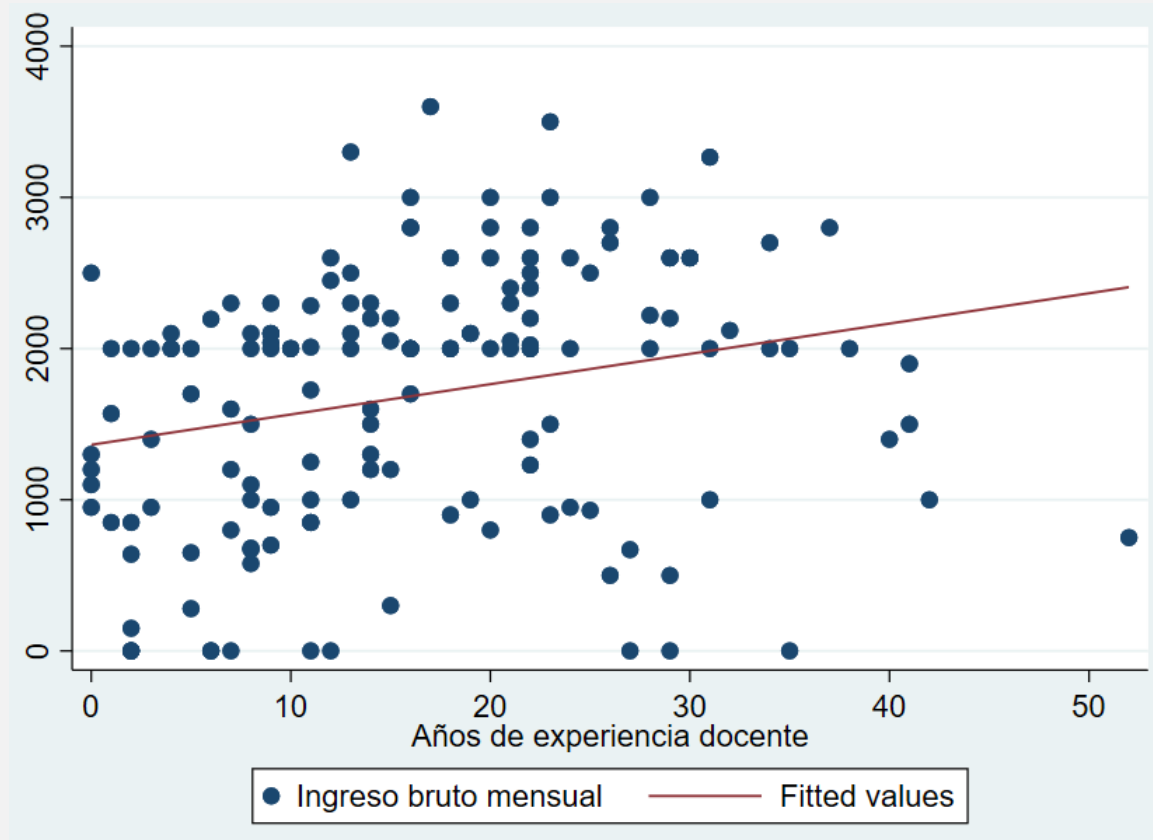


Gráficos dos variables continuas

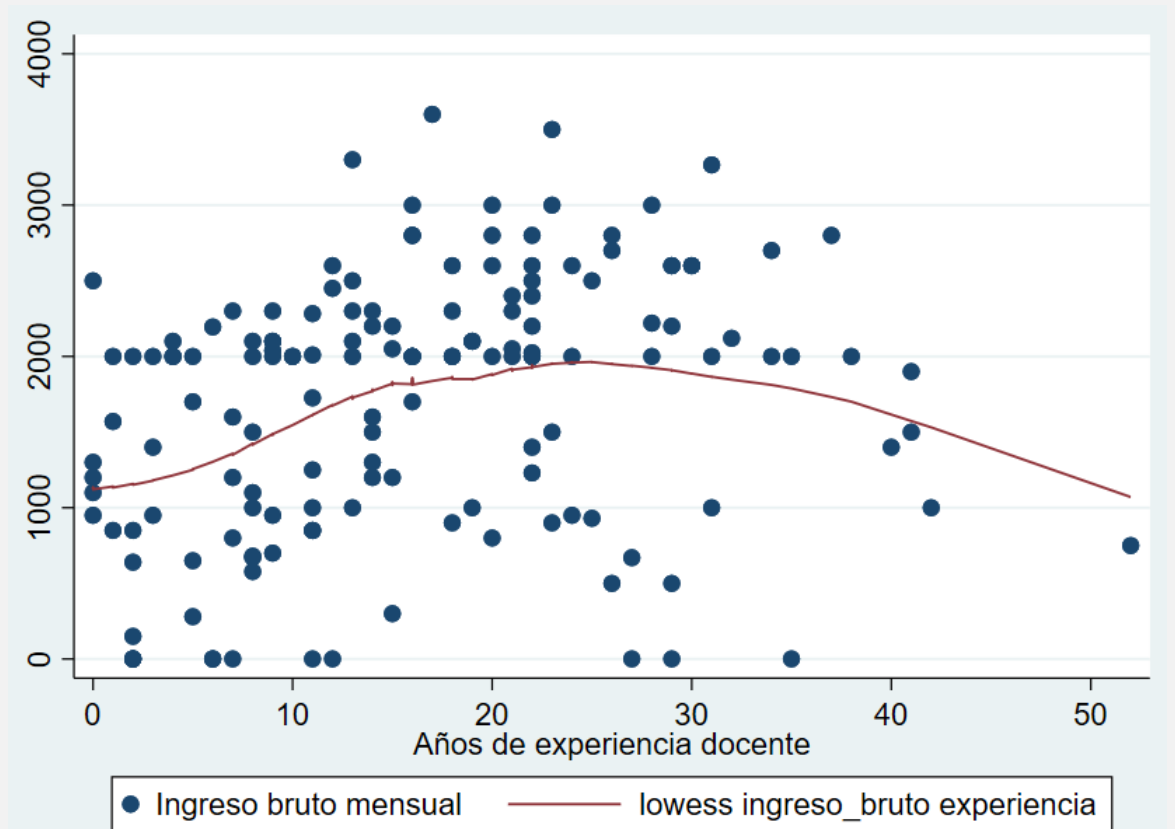
De dispersión con línea de ajuste:

General simple: *lfit/qfit/lowess* var1 var2

Con ajuste lineal



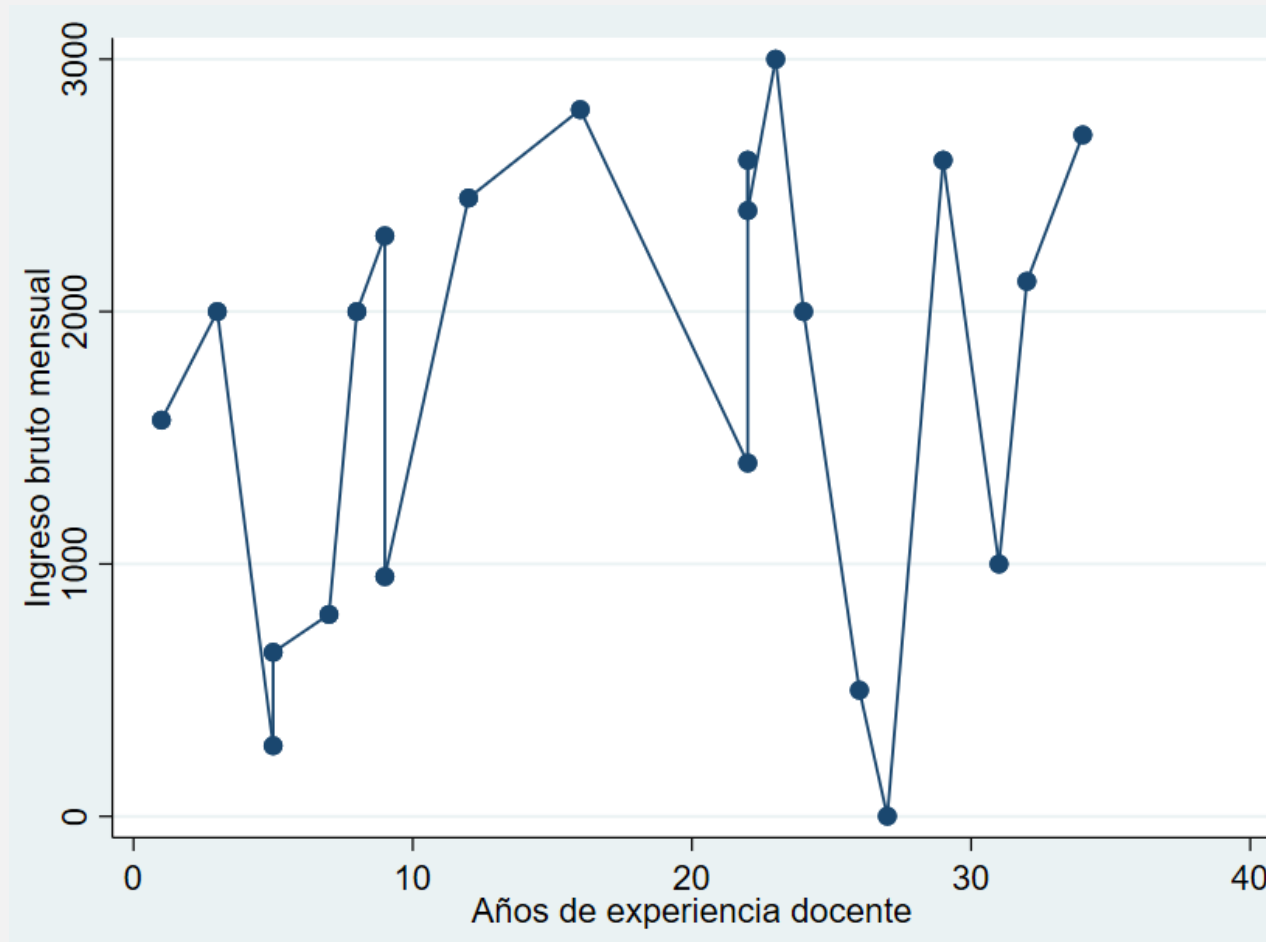
Con ajuste suavizado



Gráficos dos variables continuas

De puntos unidos:

General simple: *twoway connected* var1 var2, sort(var1/var2)

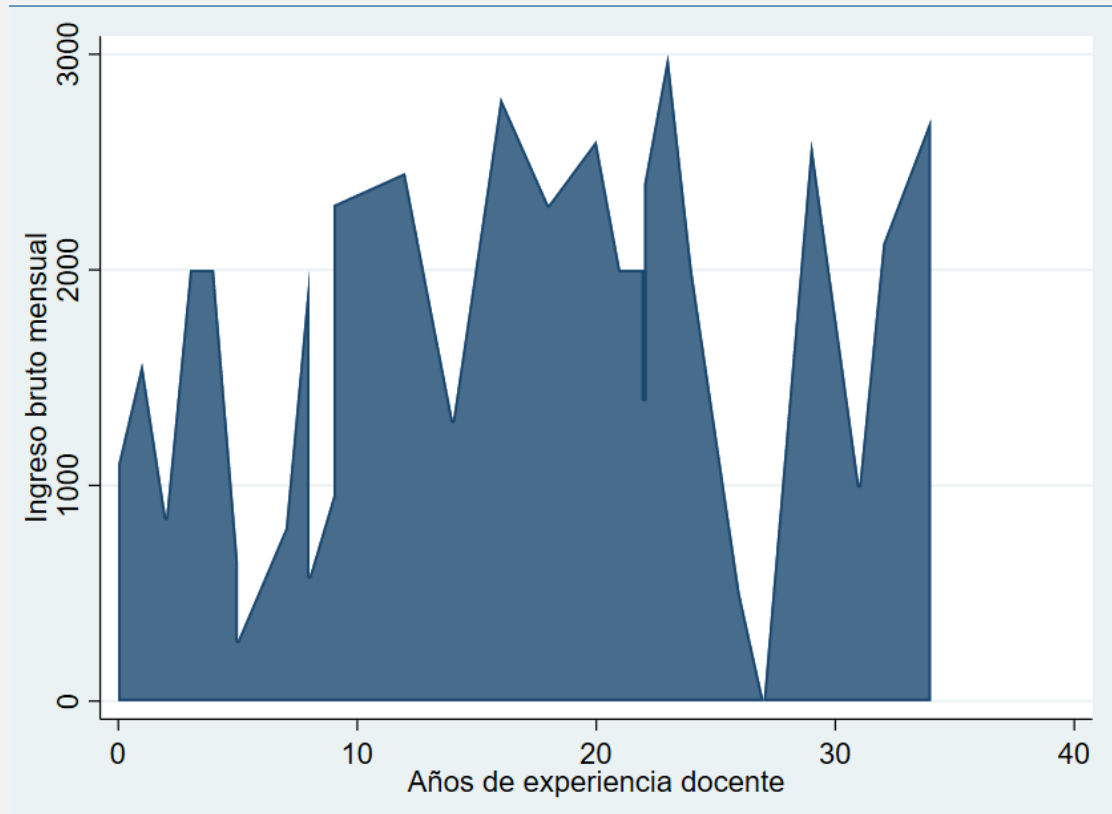


Gráficos dos variables continuas

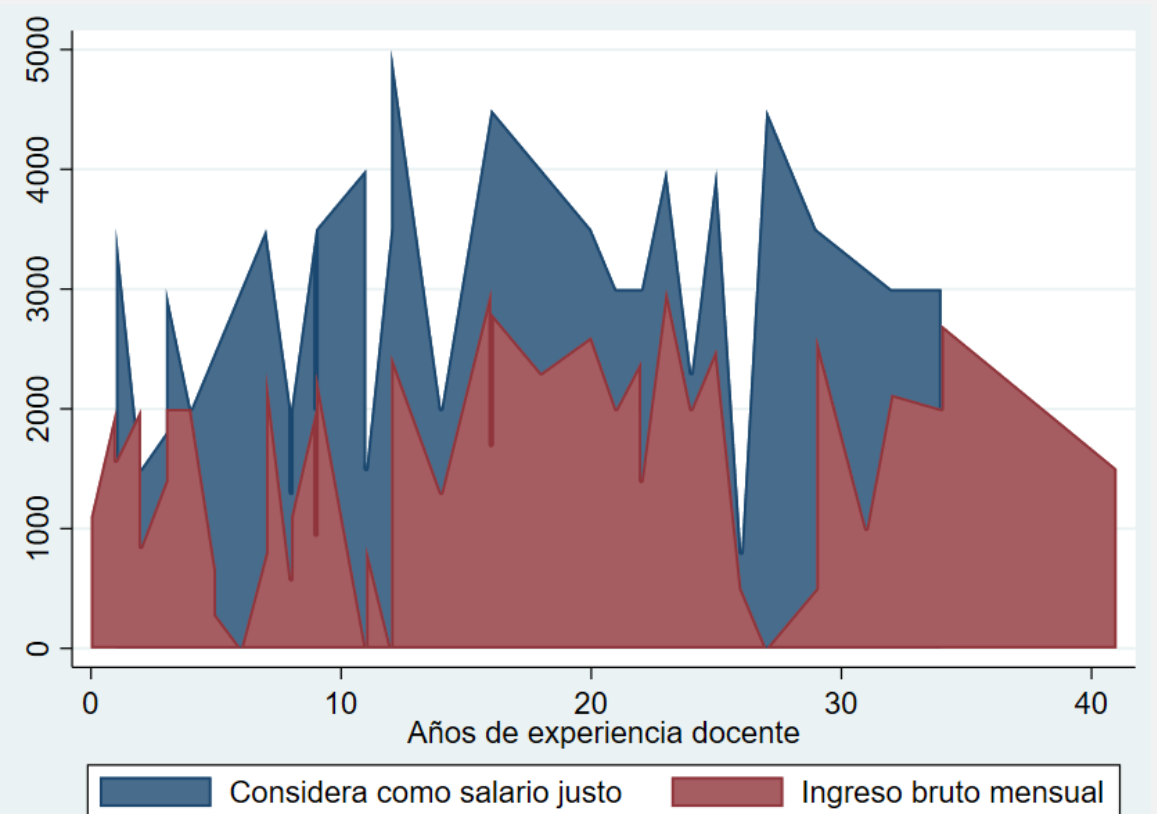
De área:

General simple: `twoway area var1 var2, sort(var2)`

De una variable



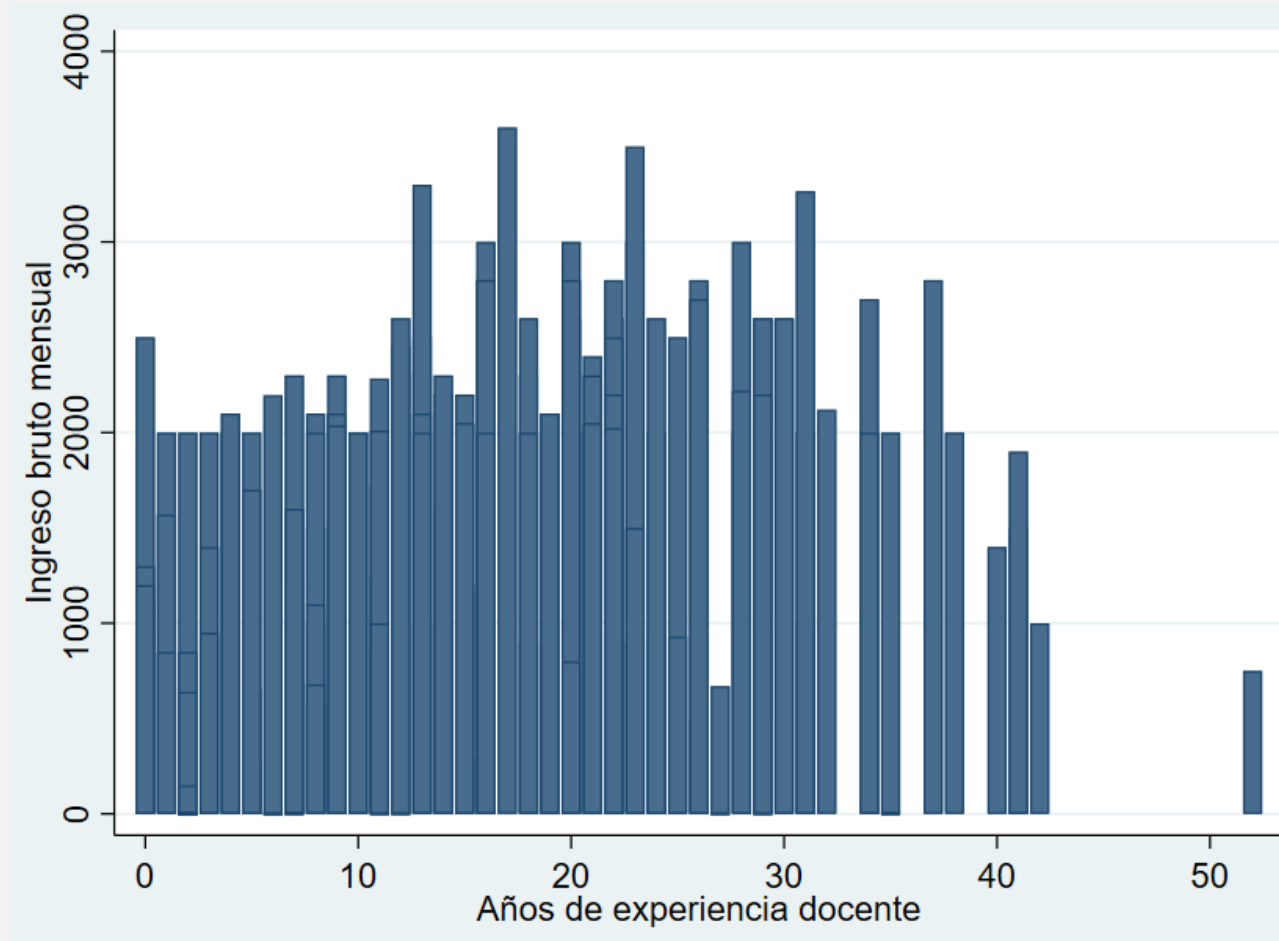
De dos variables



Gráficos dos variables continuas

De barras dobles:

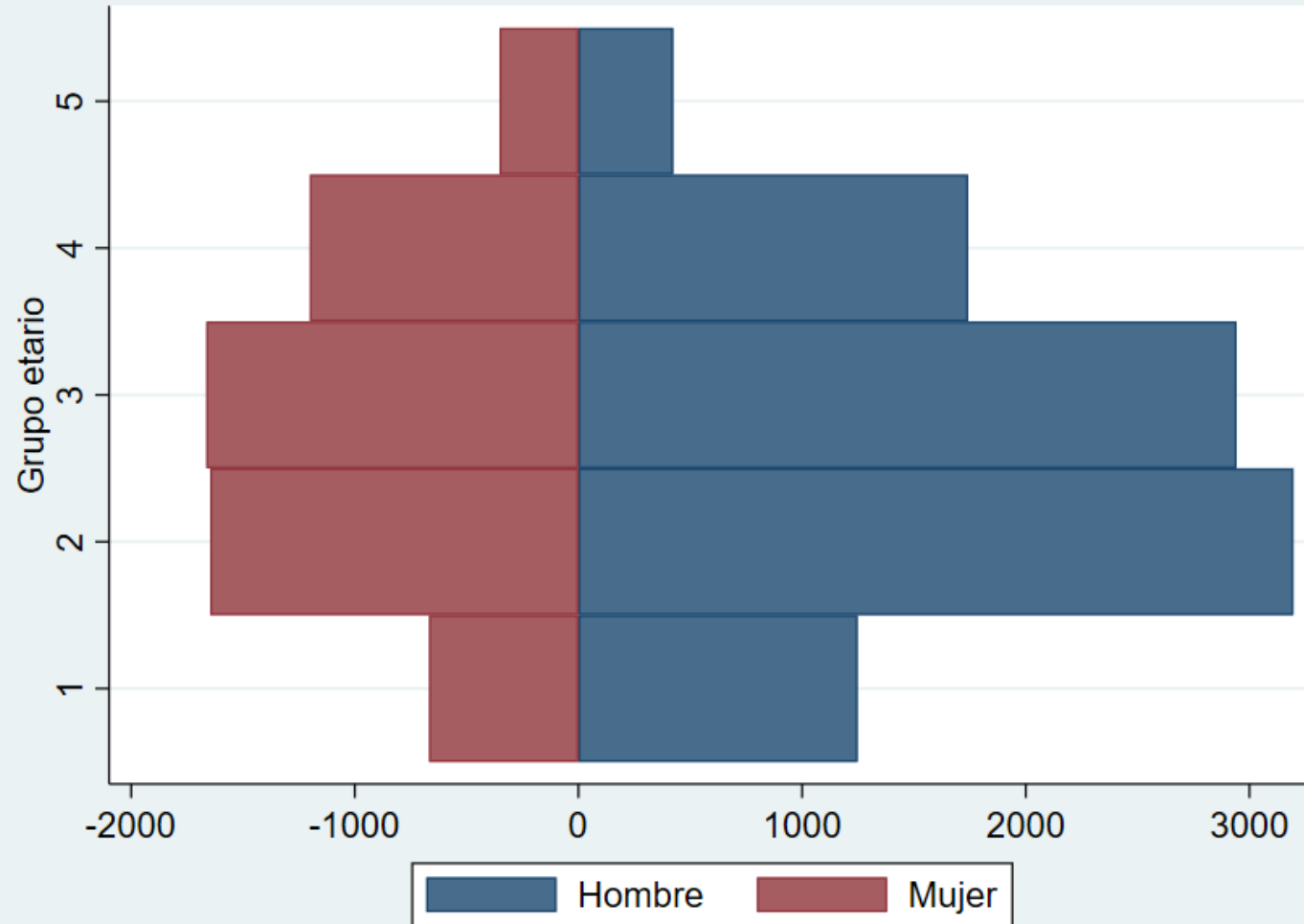
General simple: *twoway bar var1 var2*



Gráficos dos variables continuas

De barras dobles:

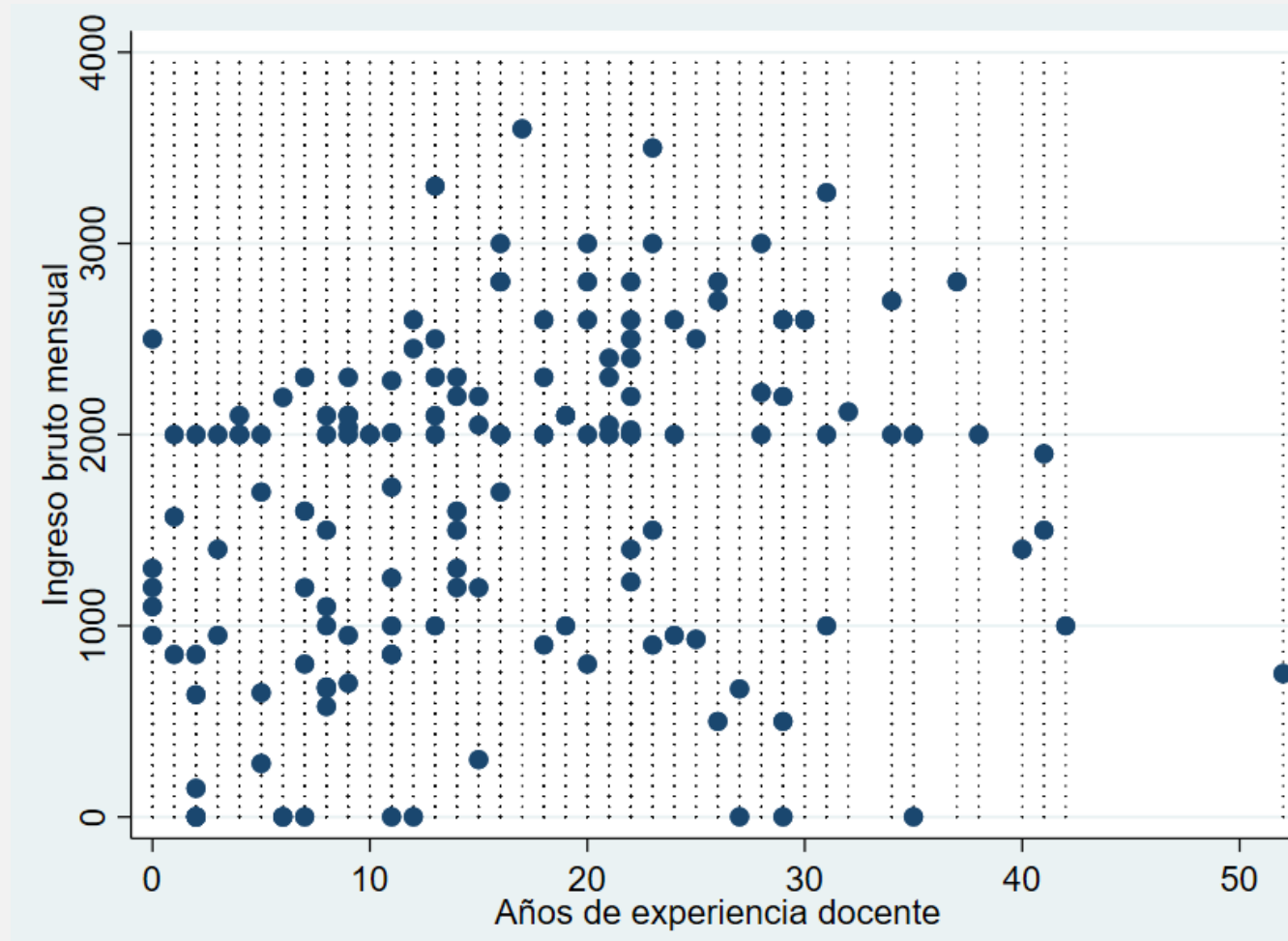
Aplicación para construir pirámide de edades



Gráficos dos variables continuas

De puntos dobles:

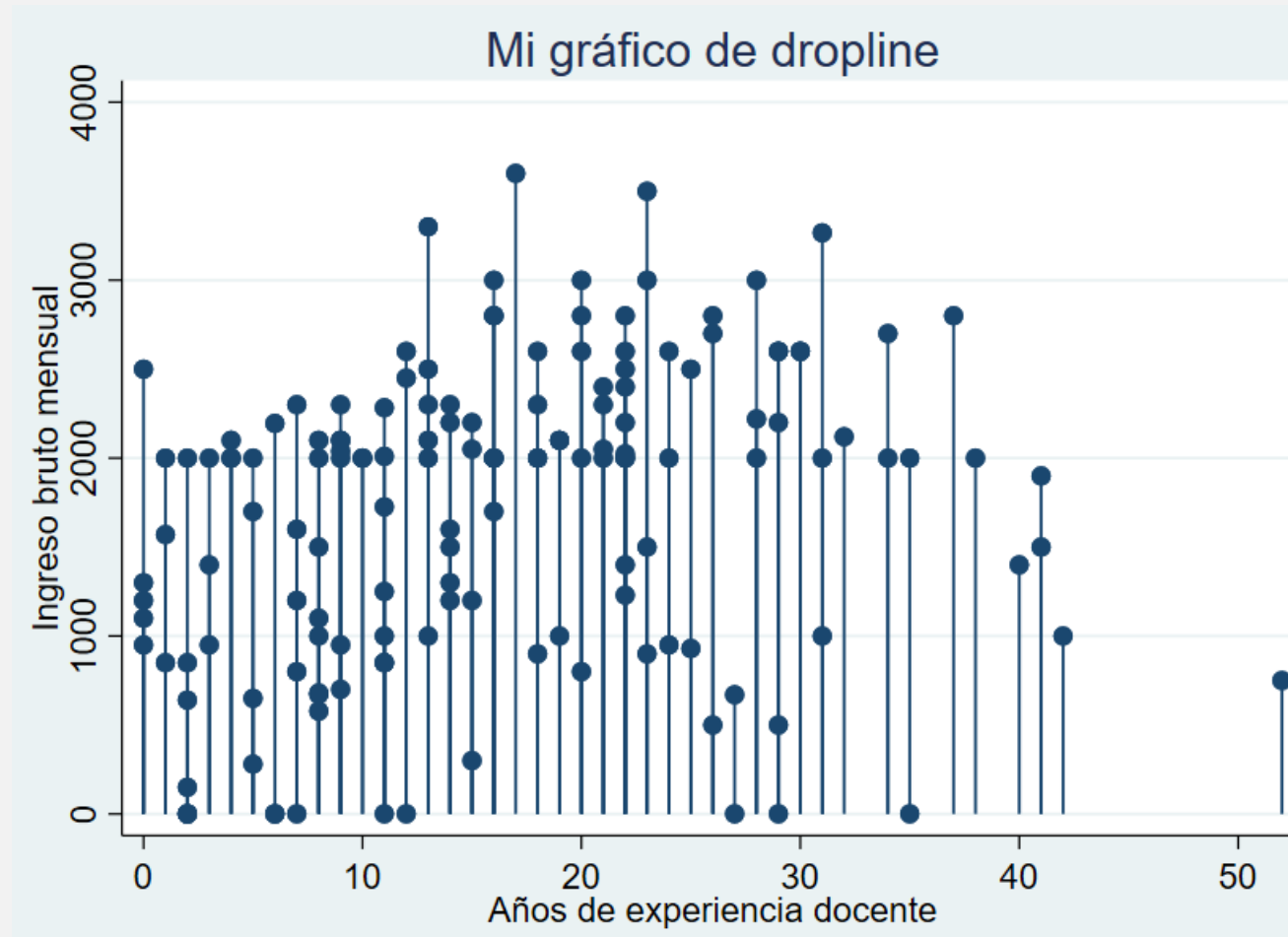
General simple: `twoway dot var1 var2`



Gráficos dos variables continuas

De puntos y líneas:

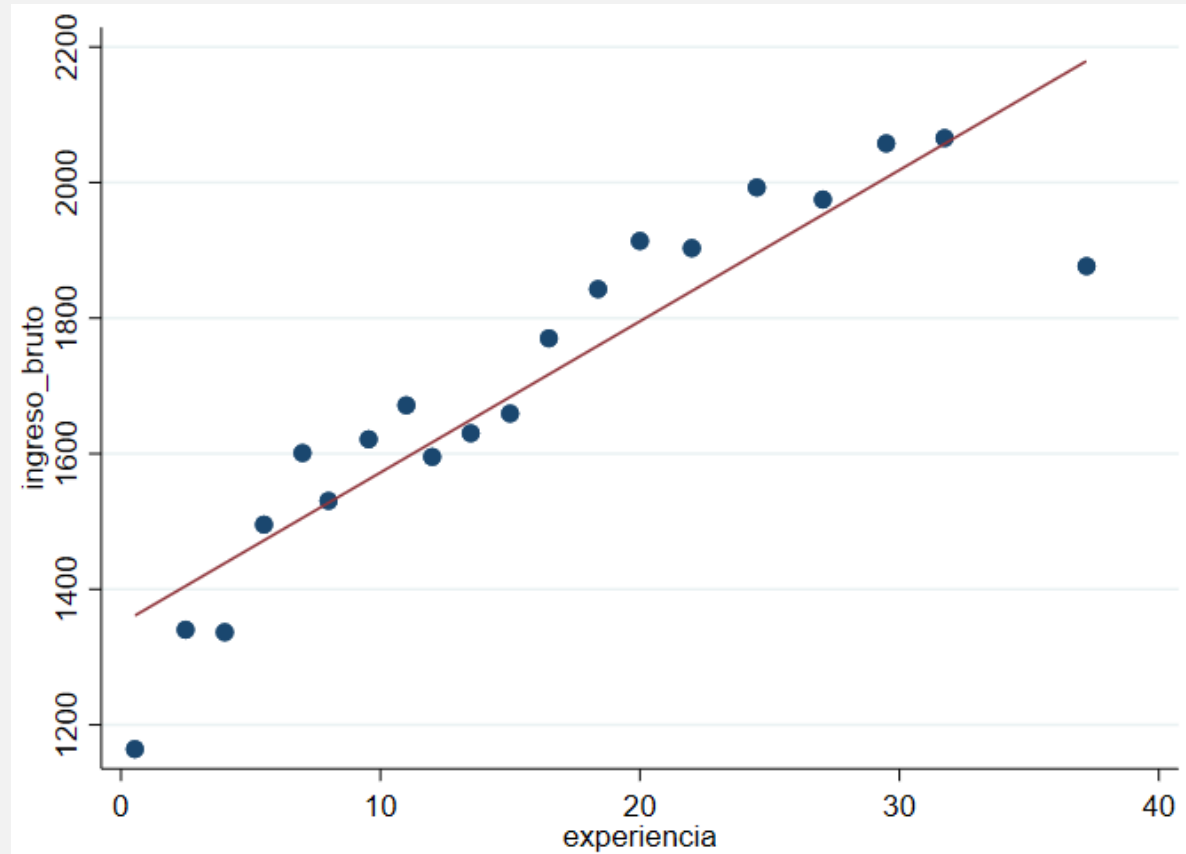
General simple: *twoway dropline var1 var2*



Gráficos dos variables continuas

Binscatter: media o mediana de variable “y” para cada “x”

General simple: `binscatter var1 var2`



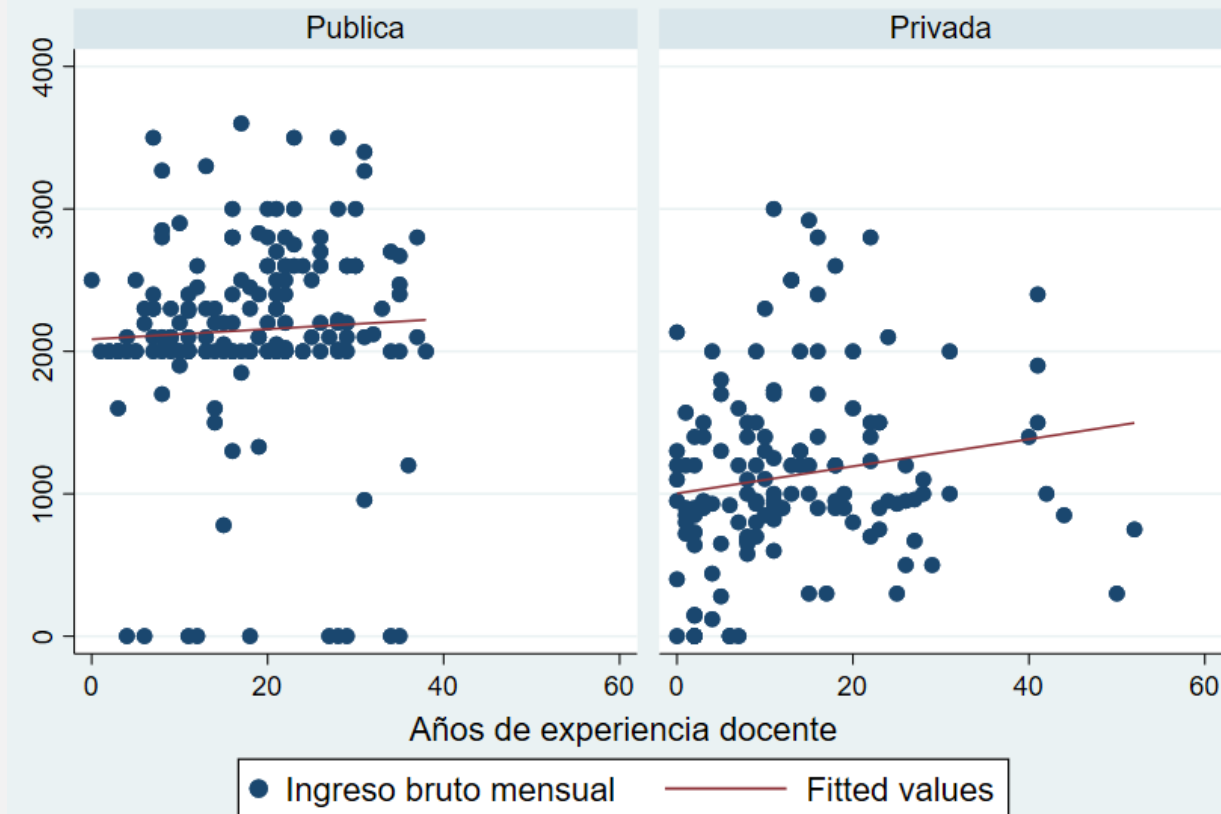
Gráficos combinados

Gráficos combinados

Diferentes categorías en un panel:

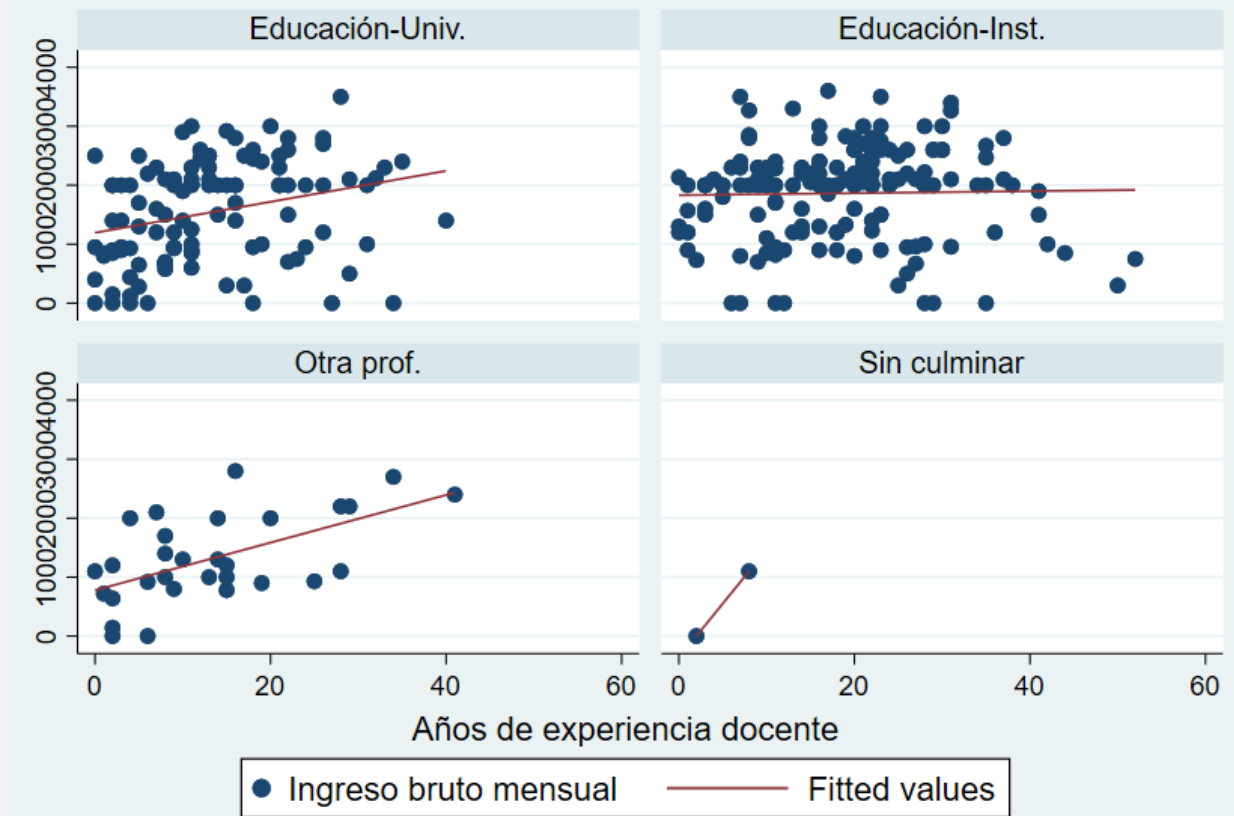
General simple: `twoway (graph1 var1 var2) (graph2 var1 var2), by(var_categ)`

Por gestión de la IE



Graphs by Gestión de su IE

Por estudios

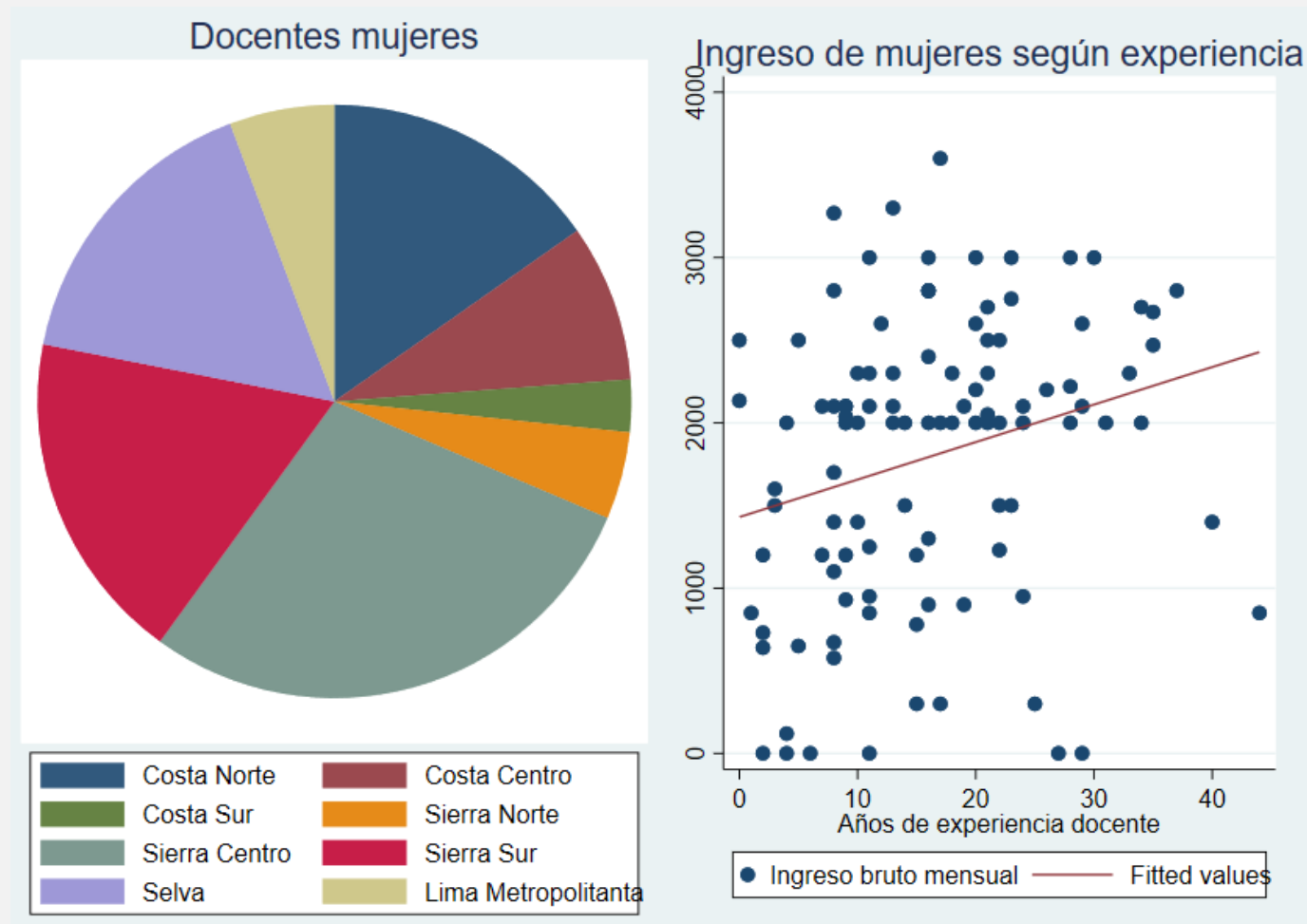


Graphs by Primera carrera que estudió

Gráficos combinados

Combinando gráficos elaborados por separado

General simple: `graph combine graph1 graph2, col(xx) row(yy)`



Gráficos combinados

Matriz de gráficos:

General simple: *graph matrix* var1 var2 var3



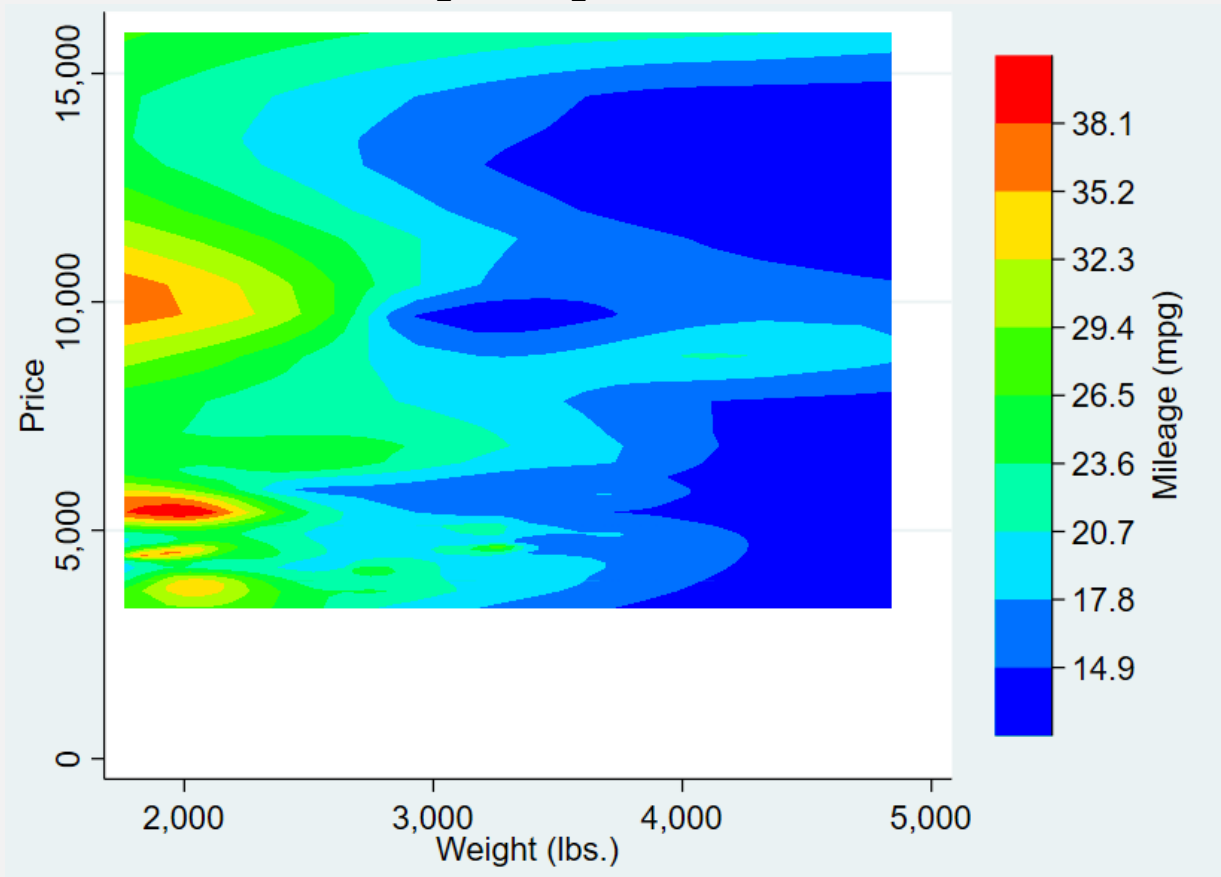
Gráficos de tres variables

Gráficos de tres variables

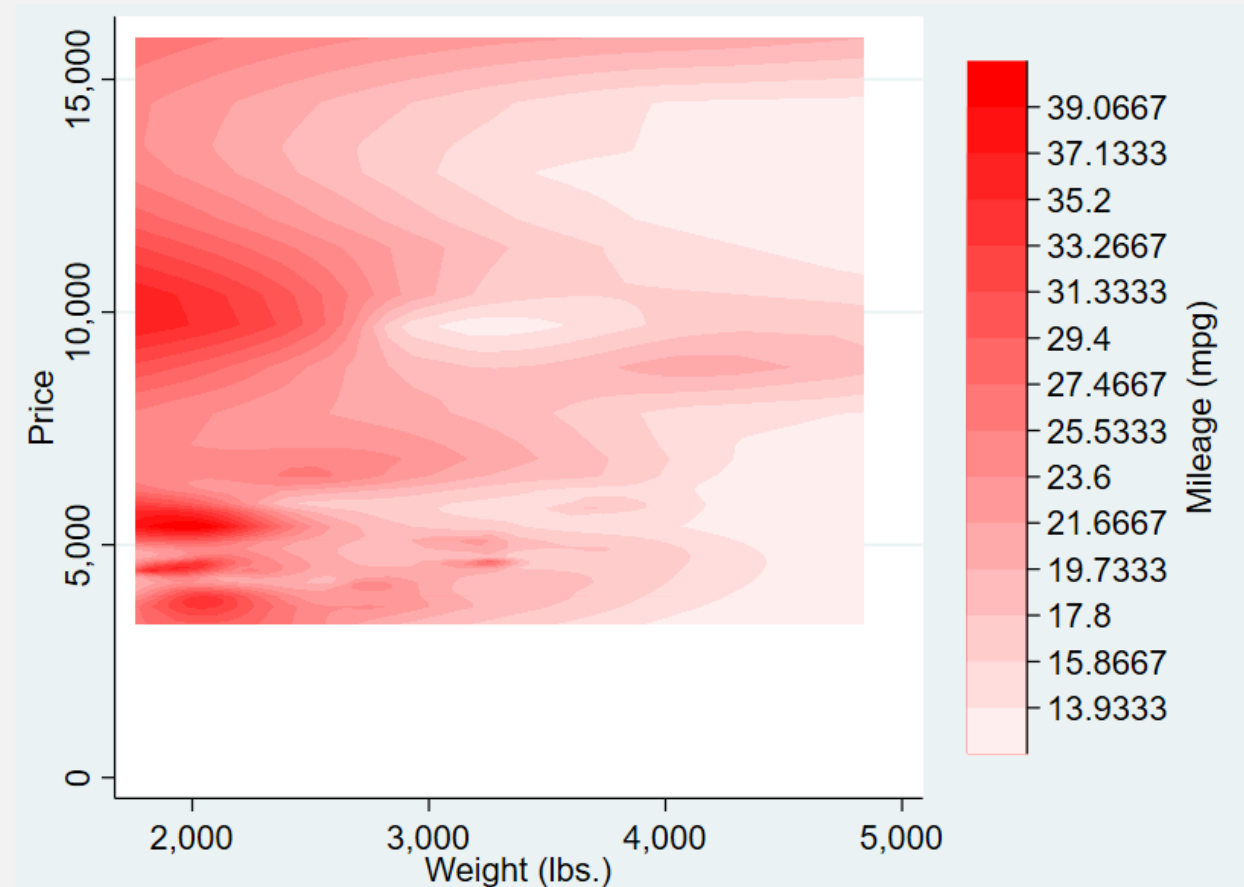
Gráfico de contorno con sombra de área

General simple: *twoway contour* var1 var2 var3

Tipo mapa de calor



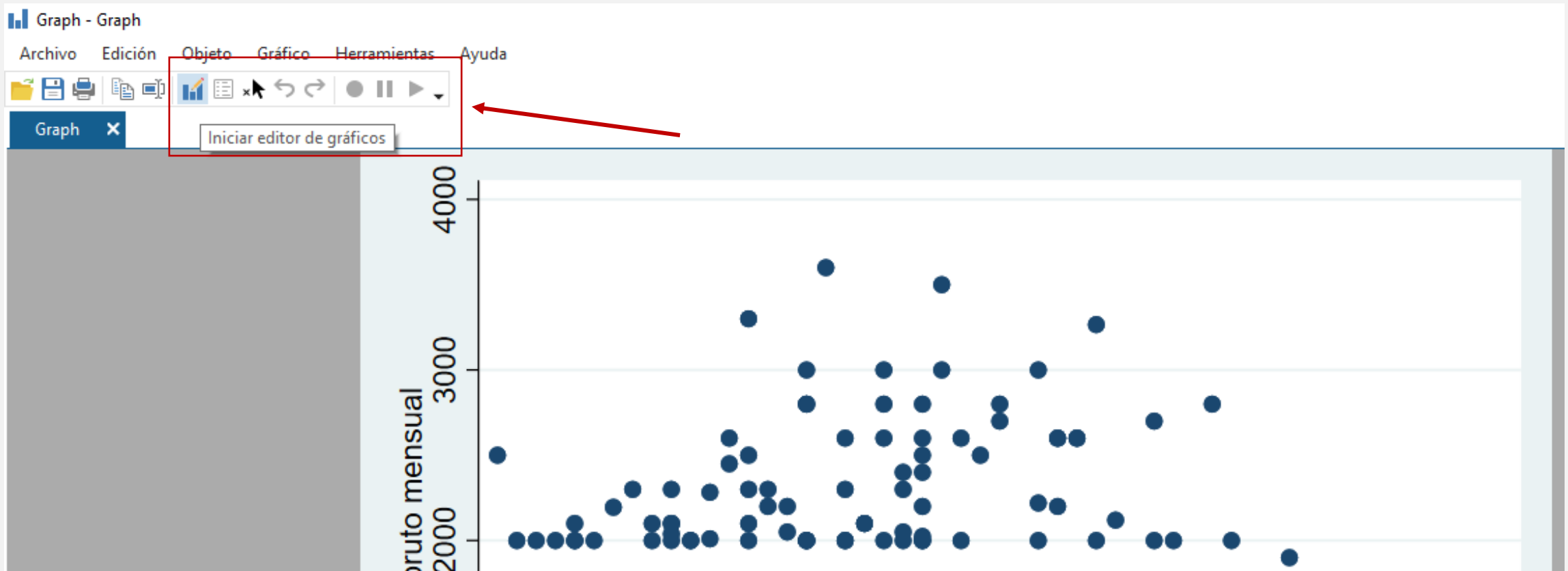
Por intensidad de color



Uso del graph editor/recorder

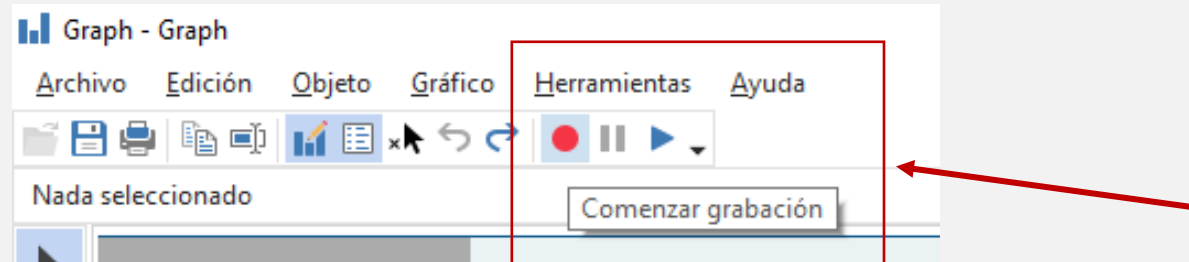
Uso del graph editor/recorder

- El graph editor/recorder es una herramienta que nos permite editar los gráficos elaborados, además de realizar cambios de manera manual y grabarlos para replicarlos posteriormente, ya sea con las mismas variables o con otras.
- El primer paso consiste en elaborar un gráfico y luego hacer click en el **ícono de edición**



Uso del graph editor/recorder

- Posteriormente, antes de realizar los cambios manuales, hay que presionar el botón de grabación:

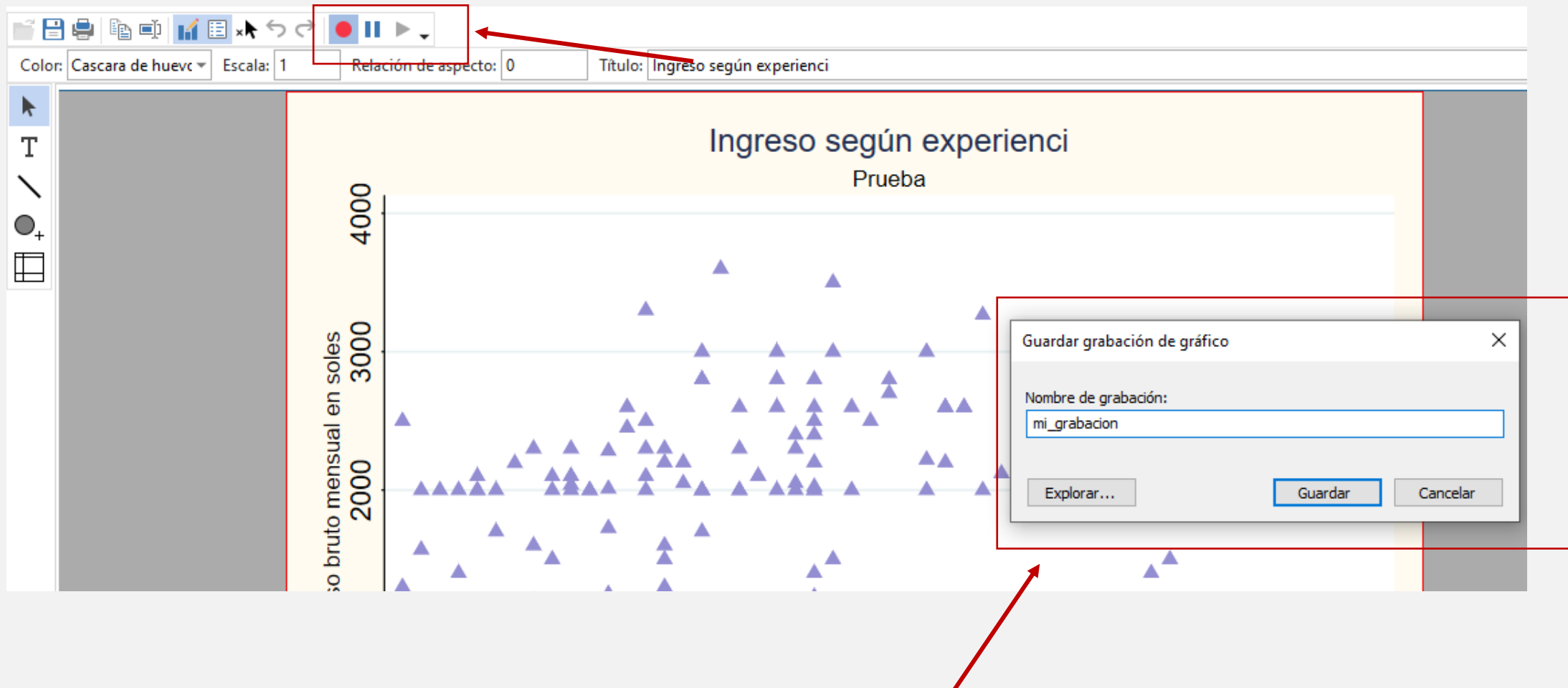


- Existe una serie de opciones que podemos llevar a cabo: cambio/incorporación de títulos, subtítulos, modificación de tamaños de letras, colores, formas, entre otros.



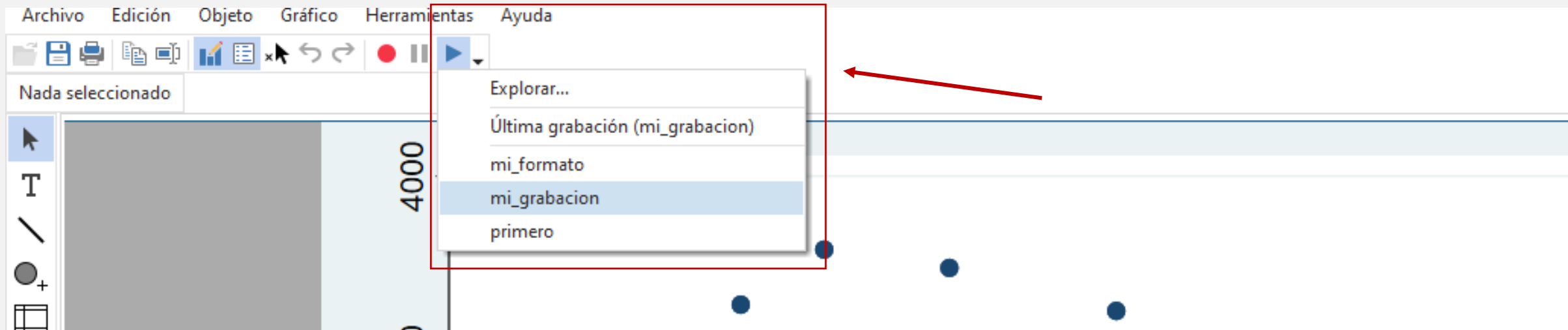
Uso del graph editor/recorder

- Una vez que hayamos culminado de realizar los cambios, presionamos nuevamente el botón de grabación para detener el proceso, y le colocamos un nombre para guardarlo en su memoria y podamos replicarlo en el futuro.



Uso del graph editor/recorder

- Para replicar el formato, solo hace falta usar la opción, *play(nombre_grabacion)*, donde nombre_grabación, es el nombre que colocamos al proceso.
- También se puede ejecutar manualmente desde el editor de gráficos, haciendo click en el botón de play, y seleccionando el nombre del proceso.
- Ejemplo: *scatter ingreso_bruto edad_docente,play(mi_formato)*



Mapas coropléticos en STATA

Mapas coropléticos en STATA

- Para elaborar mapas en STATA, usaremos una base de datos a nivel distrital, procedente de la web del Centro Nacional de Planeamiento Estratégico (CEPLAN), denominada “Información para el planeamiento a nivel departamental, provincial y distrital”.
- El acceso se da a través del siguiente link: [Información para el planeamiento a nivel departamental, provincial y distrital | CEPLAN](#)
- Descargar la base disponible en el link “Datos útiles para el planeamiento”



Mapas coropléticos en STATA

La elaboración de mapas en STATA requiere contar con los siguientes insumos adicionales:

- Archivos de polígonos que definen los límites espaciales (shapefiles)
- Instalación de los comandos “shp2dta” (para convertir el shapefile a formato .dta de STATA) y “spmap” (para elaborar el gráfico de mapa).
- Los archivos shapefiles están siendo proporcionados como parte del material de esta sesión

Uso del comando shp2dta:

```
shp2dta using "...\\xxxx.shp", database(dist_dta) coordinates(dist_coord)   genid(dist_id)
```

Nombre del
archivo shapefile

Nombre de la base
a guardar como dta

Información de
coordenadas

Generador de ámbito
geográfico generado

Mapas coropléticos en STATA

Uso del comando spmap: mapa nacional distrital

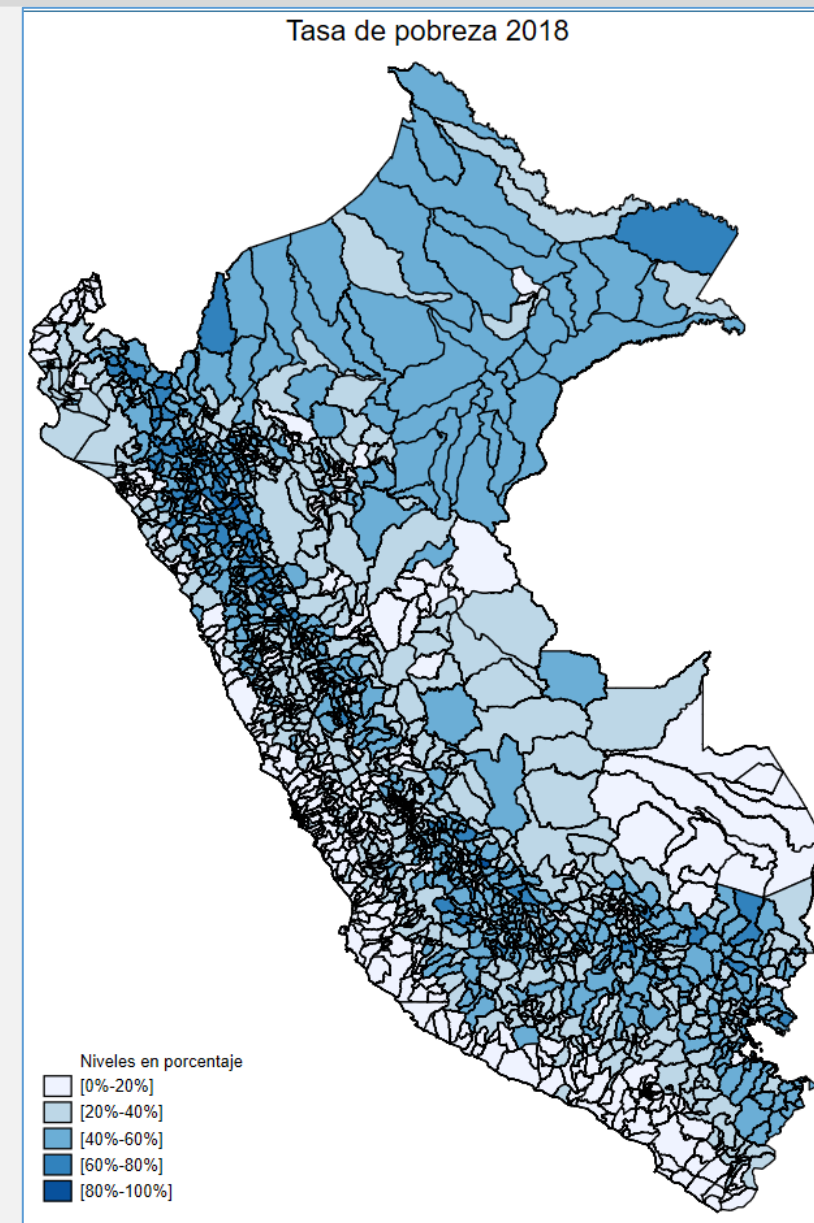
Variable a
graficar

coordenadas

color
identificador

color de
missing

```
spmap categ_pobre using dist_coord, id(dist_id) ocolor(black) fcolor(Blues) ndfcolor (white) ///  
clmethod(unique) title("Tasa de pobreza 2018", size(*0.7)) legenda(on) legend(size(small)) ///  
legtitle("Niveles en porcentaje")
```



Mapas coropléticos en STATA

Uso del comando spmap: mapa de región distrital

condicionante

```
spmap categ_IDH using dist_coord if NOMBDEP=="AYACUCHO", id(dist_id) ocolor(black)
fcolor(Blues) ndfcolor (white) ///
clmethod(unique) title("Índice de Desarrollo Humano" "Ayacucho", size(*0.7)) legenda(on)
legend(size(small)) ///
legtitle("Niveles de IDH")
```

