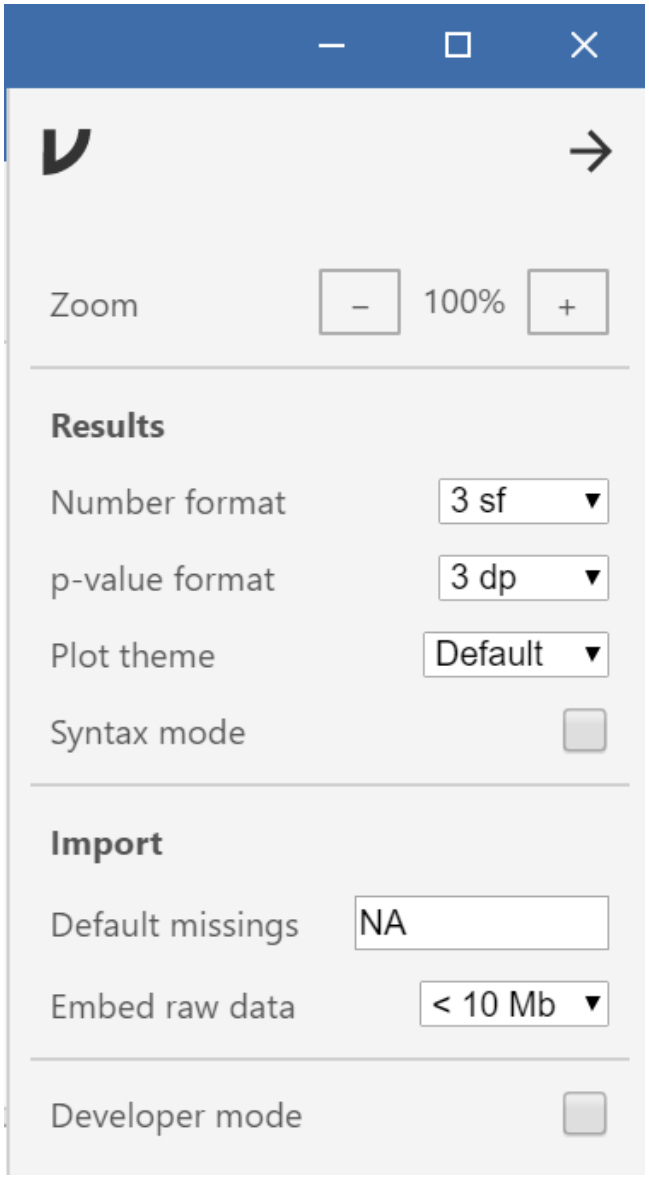
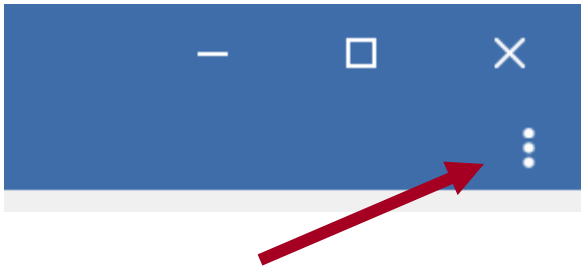
Análisis estadístico con Jamovi

Parte II: Estadística descriptiva

Juan R Gonzalez
Bioinformatics Research Group in Epidemiology,
Barcelona Institute for Global Health (ISGlobal)
and
Department of Mathematics,
Autonomous University of Barcelona (UAB)
http://brge.isglobal.org

Output settings





¿cómo resumimos las variables recogidas en nuestro estudio?

- Resumen de variables **categóricas**
  - Tablas de frecuencias
  - Medidas de frecuencia de la enfermedad
  - Gráficos de barras y sectores
- Resumen de variables **numéricas**
  - Medidas de tendencia central
  - Medidas de posición relativa
  - Medidas de dispersión
  - Medidas de forma
  - Gráficos: histogramas y cajas
- Resumen de la **relación entre 2 variables**
  - Una numérica y una categórica:
    - resumen de la numérica por categorías
    - comparación de cajas
  - Dos categóricas:
    - tabla de frecuencias (contingencia)
    - gráfico de barras
  - Dos numéricas
    - Otros cursos de estadística ...

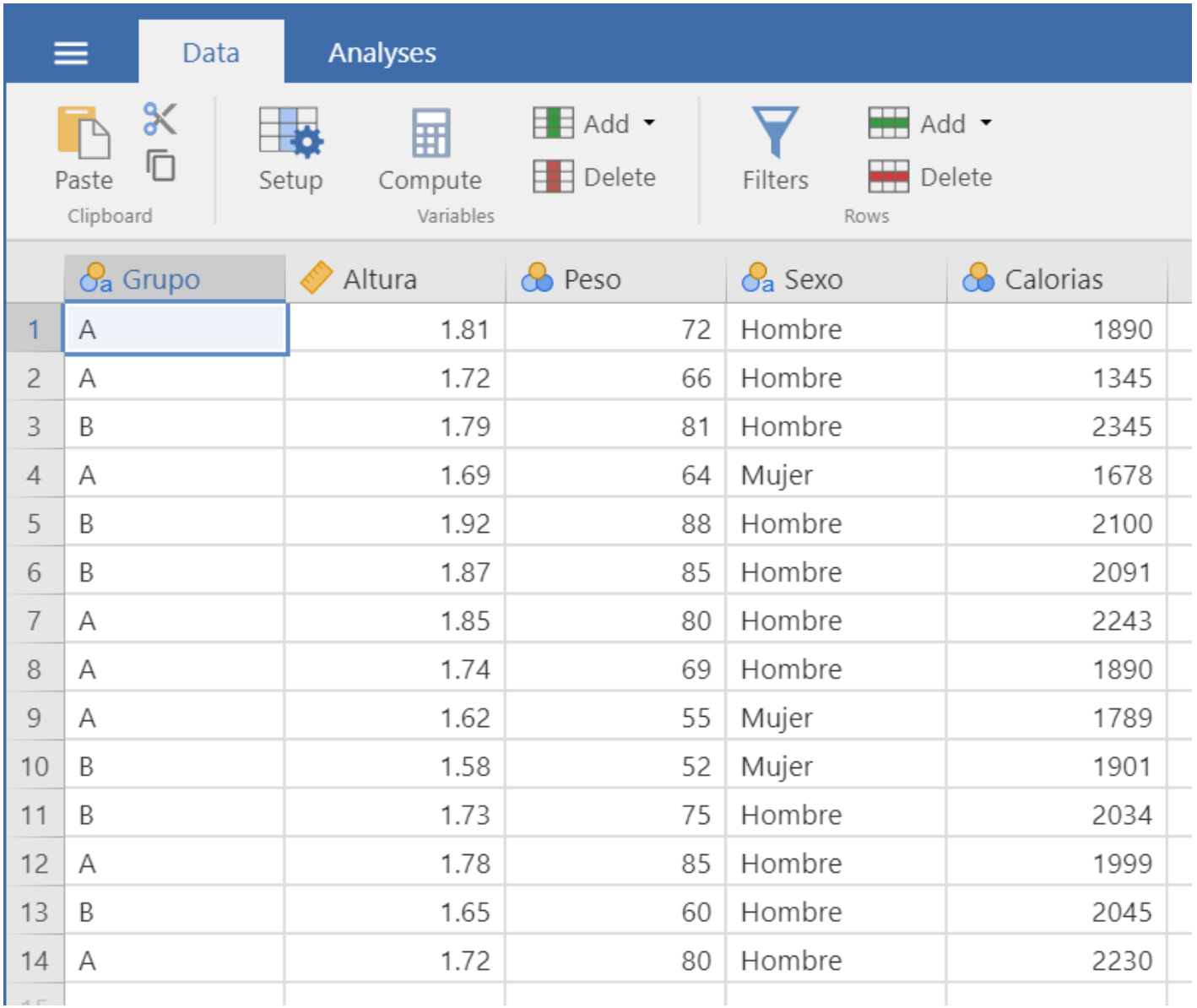
Diferenciar por tipo de variable

Categóricas
(Cualitativas)

Numéricas
(Cuantitativas)

Estadística descriptiva

ejemplo.txt



	Grupo	Altura	Peso	Sexo	Calorias
1	A	1.81	72	Hombre	1890
2	A	1.72	66	Hombre	1345
3	B	1.79	81	Hombre	2345
4	A	1.69	64	Mujer	1678
5	B	1.92	88	Hombre	2100
6	B	1.87	85	Hombre	2091
7	A	1.85	80	Hombre	2243
8	A	1.74	69	Hombre	1890
9	A	1.62	55	Mujer	1789
10	B	1.58	52	Mujer	1901
11	B	1.73	75	Hombre	2034
12	A	1.78	85	Hombre	1999
13	B	1.65	60	Hombre	2045
14	A	1.72	80	Hombre	2230

Resumen de variable continua

- **Tendencia central**
  - Media, mediana
- **Dispersión**
  - Desviación típica, varianza, rango, rango intercuartílico, coeficiente de variación
- **Posición**
  - Percentiles, terciles, cuartiles, quintiles
- **Forma**
  - Asimetría

Visualización de variable continua

- **Histograma**
- **Boxplot**
  - Media, mediana, RIQ, …
- **Violin plot**
  - Densidad de datos
- **Dotplot**

Boxplot

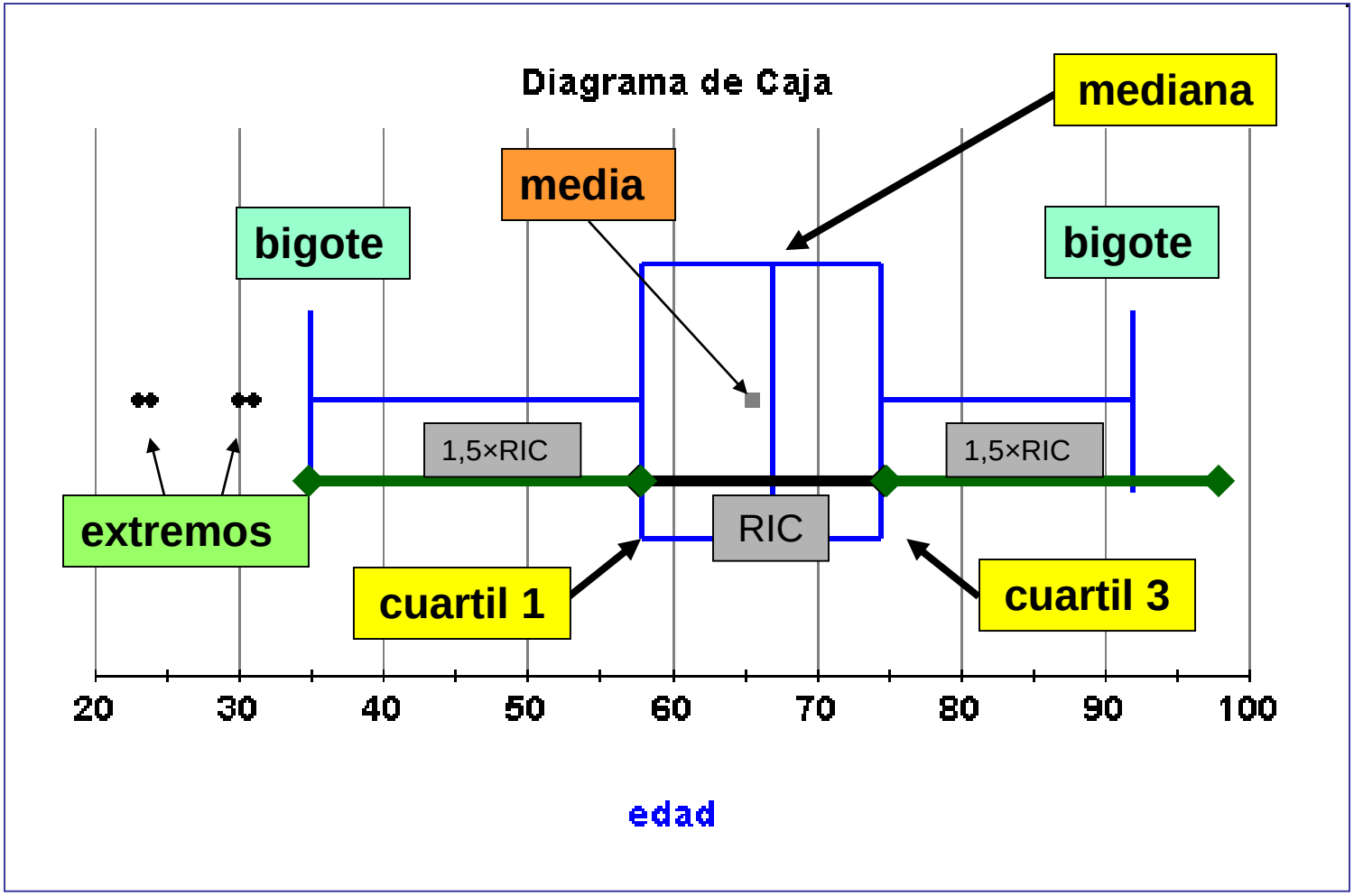
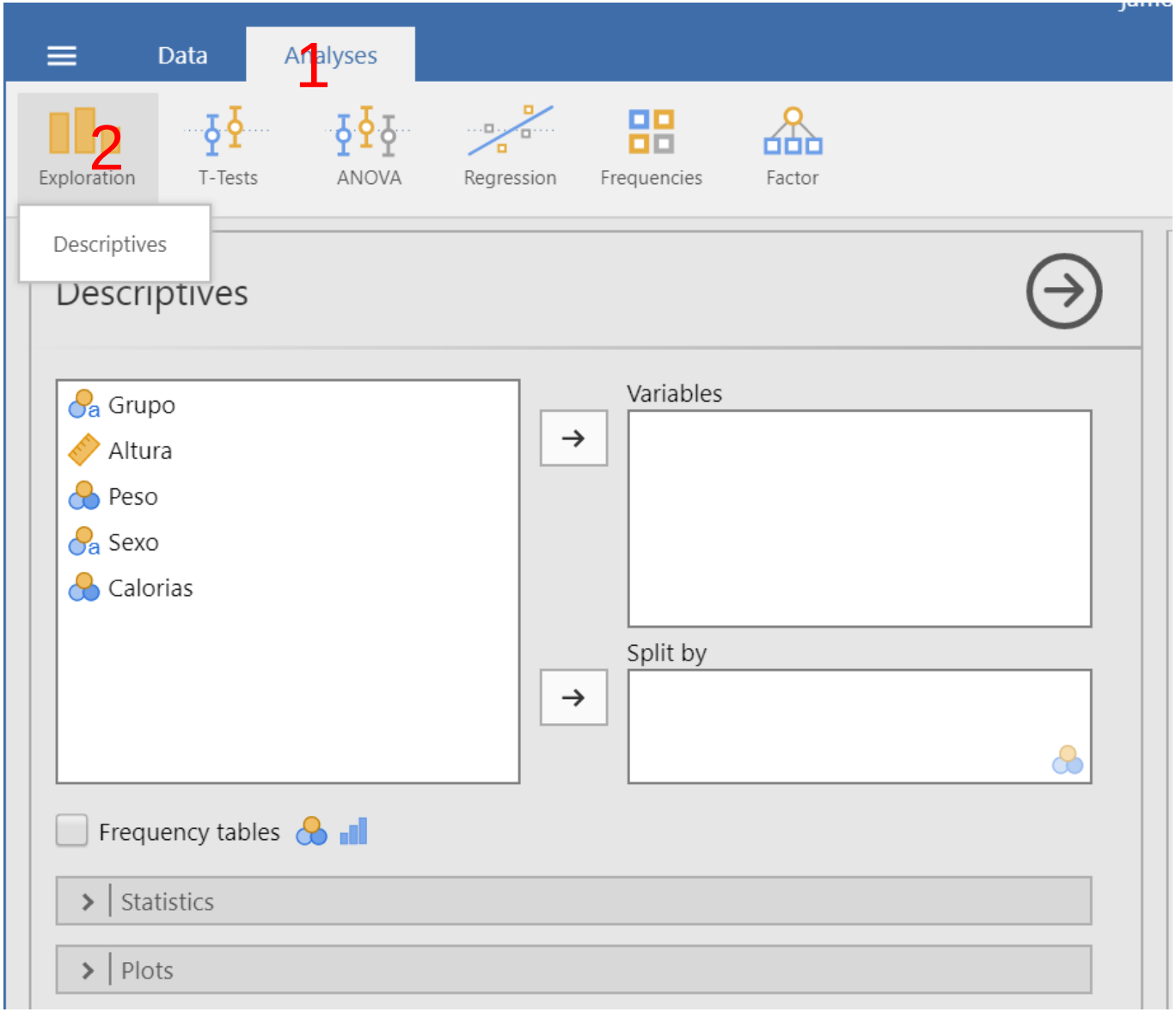
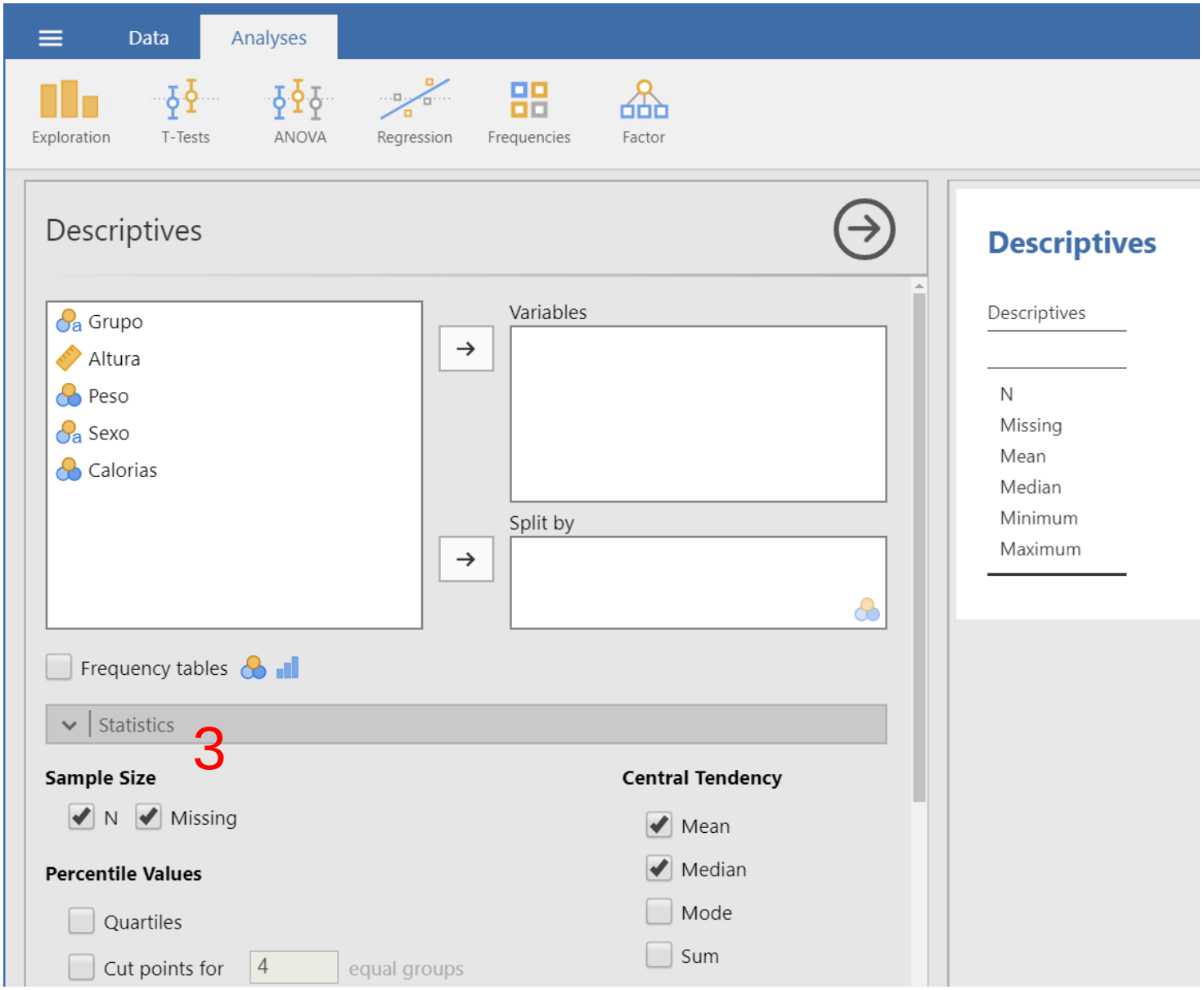


Diagrama de Caja
mediana
media
bigote
bigote
1,5×RIC
1,5×RIC
RIC
extremos
cuartil 1
cuartil 3
20
30
40
50
60
70
80
90
100
edad

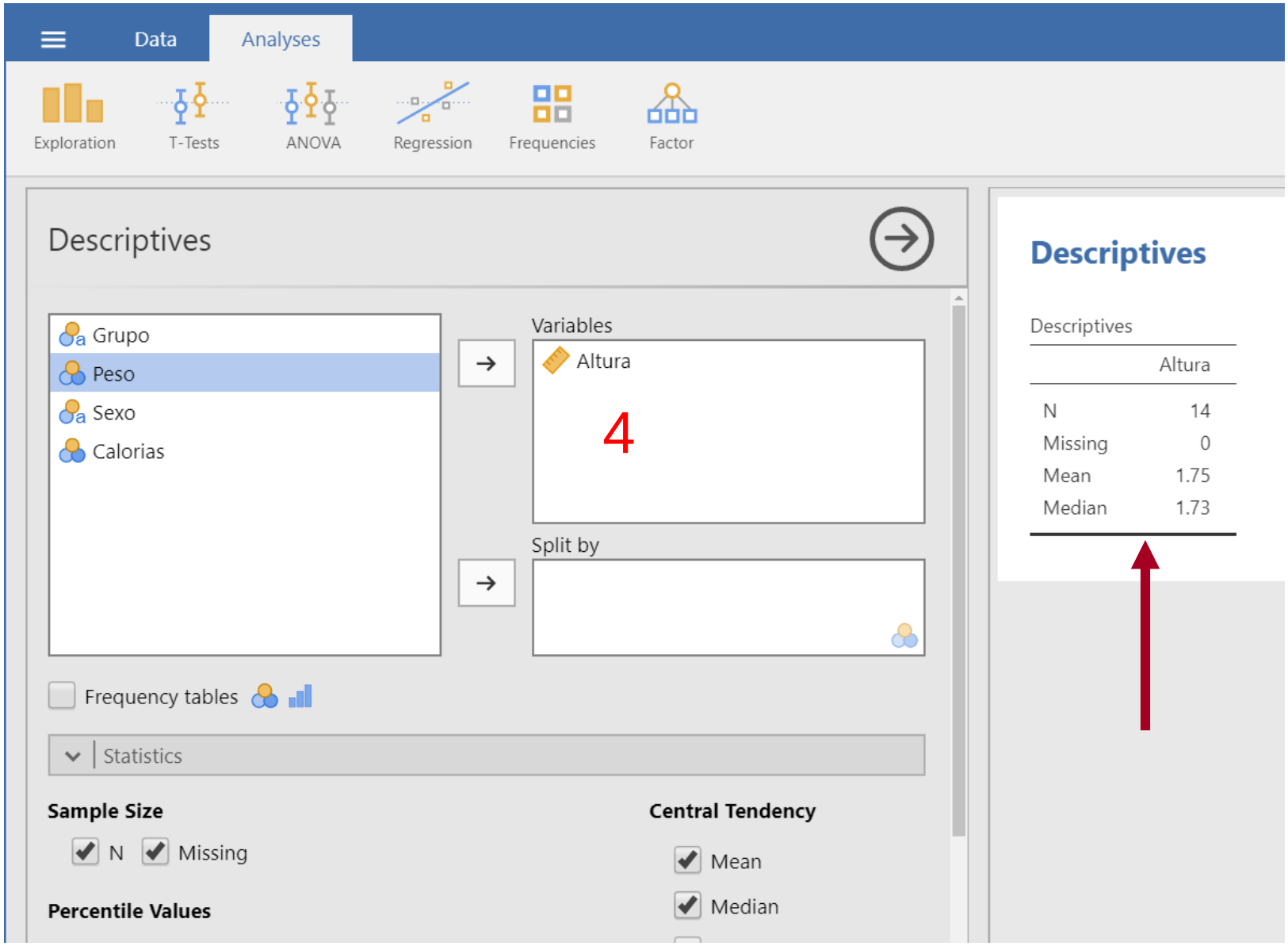
Estadística descriptiva



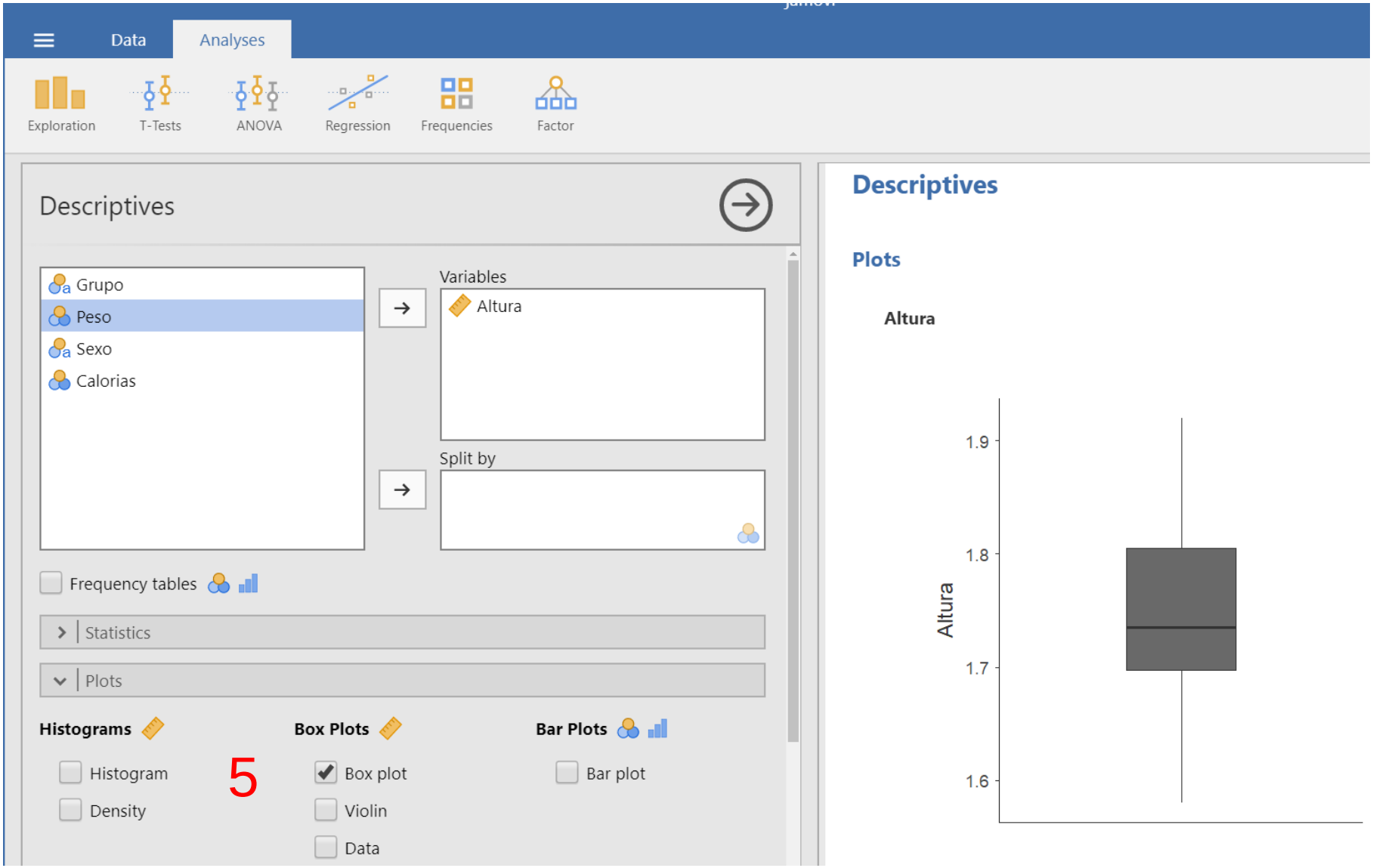
Variables continuas



Variables continuas



Variables continuas



Variables continuas

Data | Analyses

Exploration | T-Tests | ANOVA | Regression | Frequencies | Factor

Descriptives

Grupo
Peso
Sexo
Calorias

Variables: Altura

Split by

Frequency tables

Statistics

Plots

Histograms: Histogram, Density

Box Plots: Box plot, Violin, Data

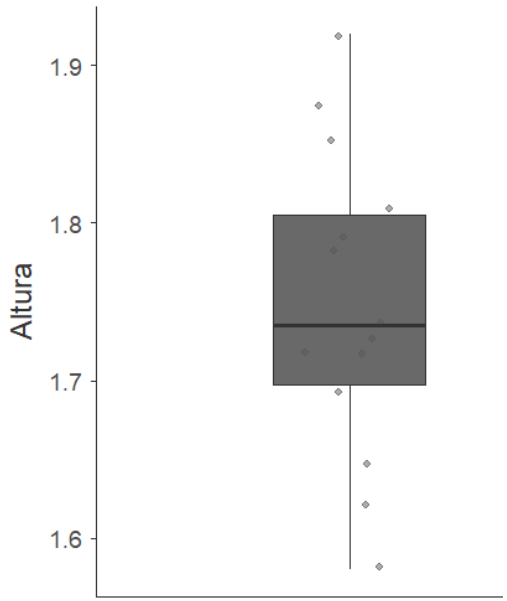
Bar Plots: Bar plot

5

Descriptives

Plots

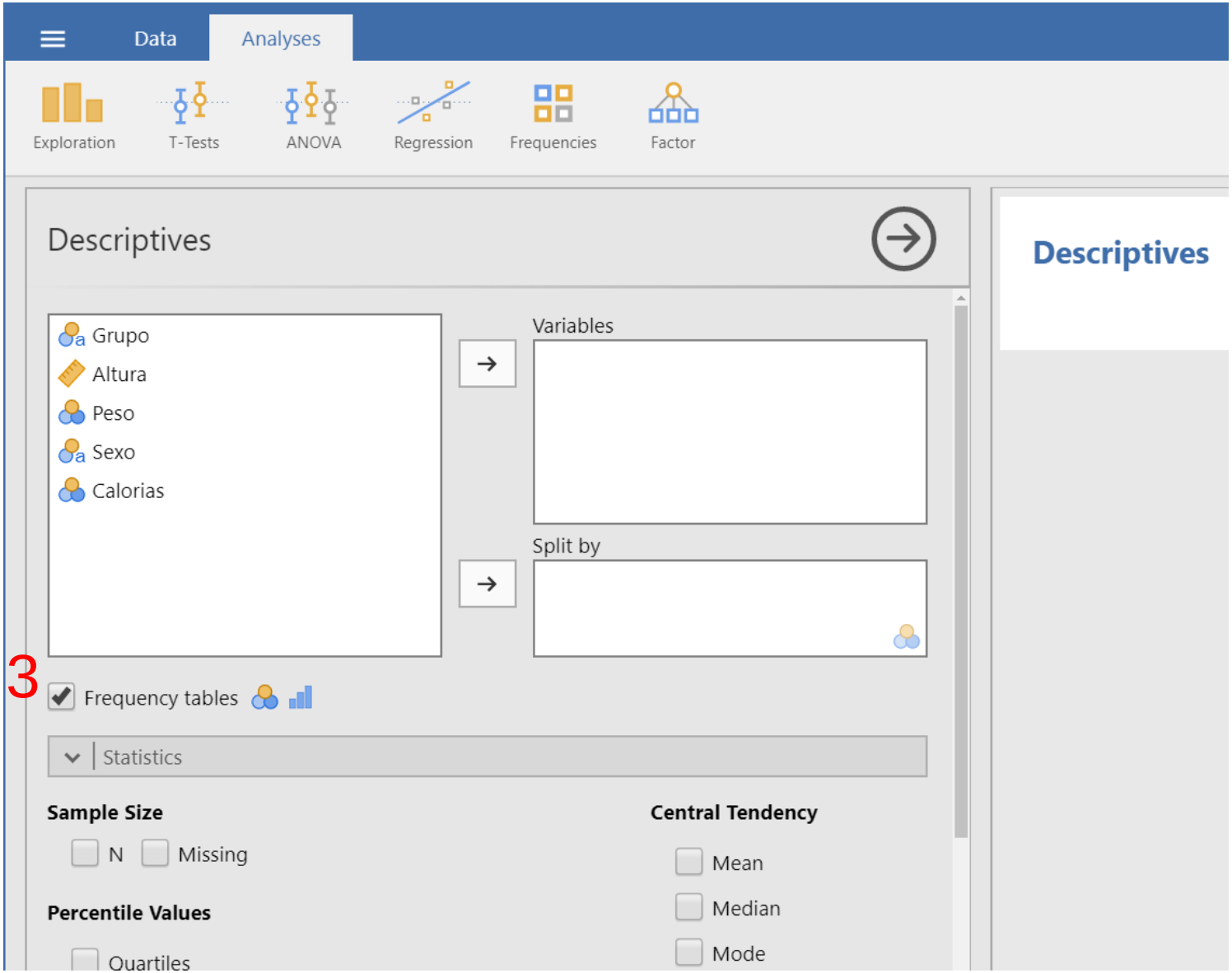
Altura



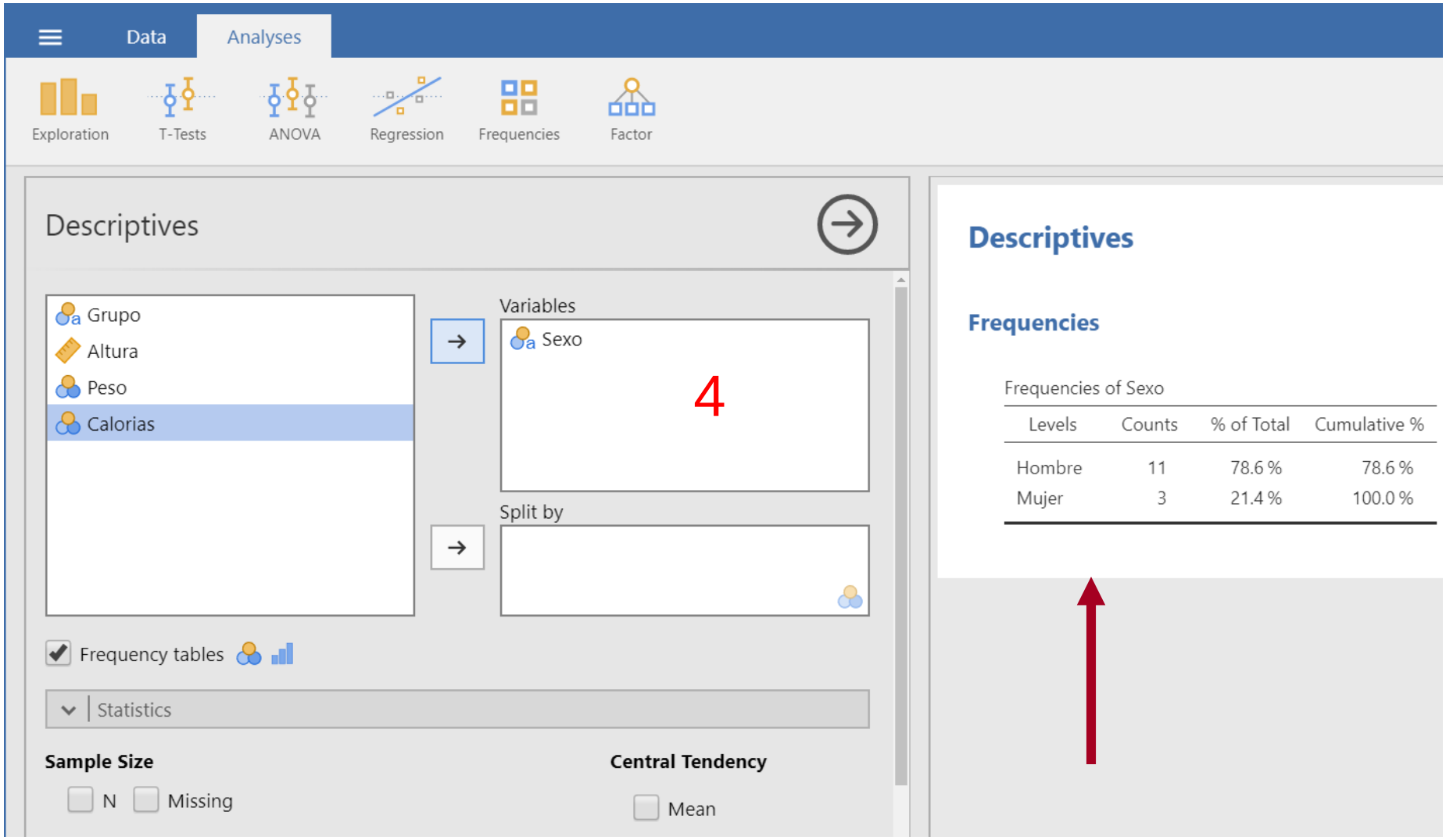
Variables categóricas

- **Resumen**
  - Frecuencias absolutas
  - Frecuencias relativas
- **Visualización**
  - Gráfico de barras
  - Gráfico de sectores (pie chart)

Categóricas



Categóricas



Categóricas

Data | Analyses

Exploration | T-Tests | ANOVA | Regression | Frequencies | Factor

Descriptives

Grupo
Altura
Peso
Calorias

Variables
Sexo

Split by

Frequency tables

Statistics

Plots

Histograms
Histogram
Density

Box Plots
Box plot
Violin
Data

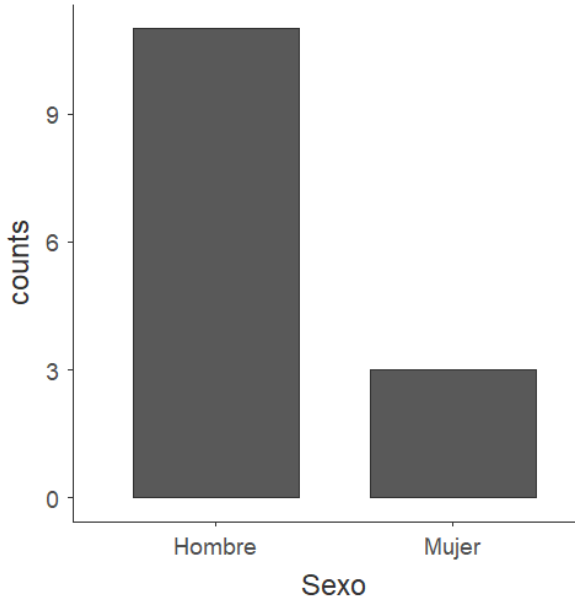
Bar Plots
Bar plot

5

Descriptives

Plots

Sexo



Descriptiva bivariante

Cómo estudiar la relación entre 2 variables

Diferenciar por tipo de variable

Categóricas
(Cualitativas)

Numéricas
(Cuantitativas)

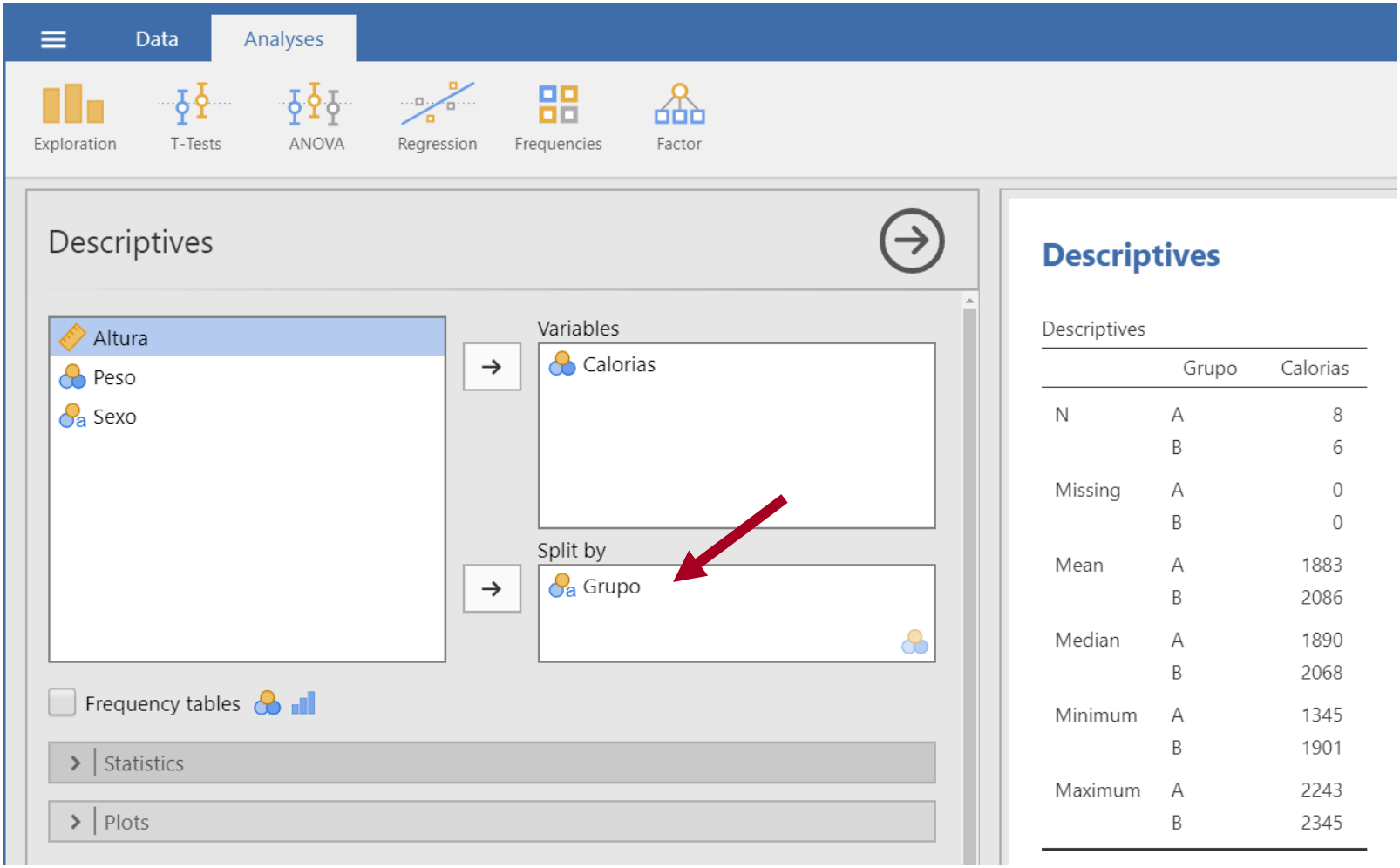
Definen grupos
a comparar

Variables categórica y numérica

¿Es cierto que los individuos del grupo A tienen un mayor consumo de calorías?

Variables categórica y numérica

Analyses -> Exploration -> Descriptives



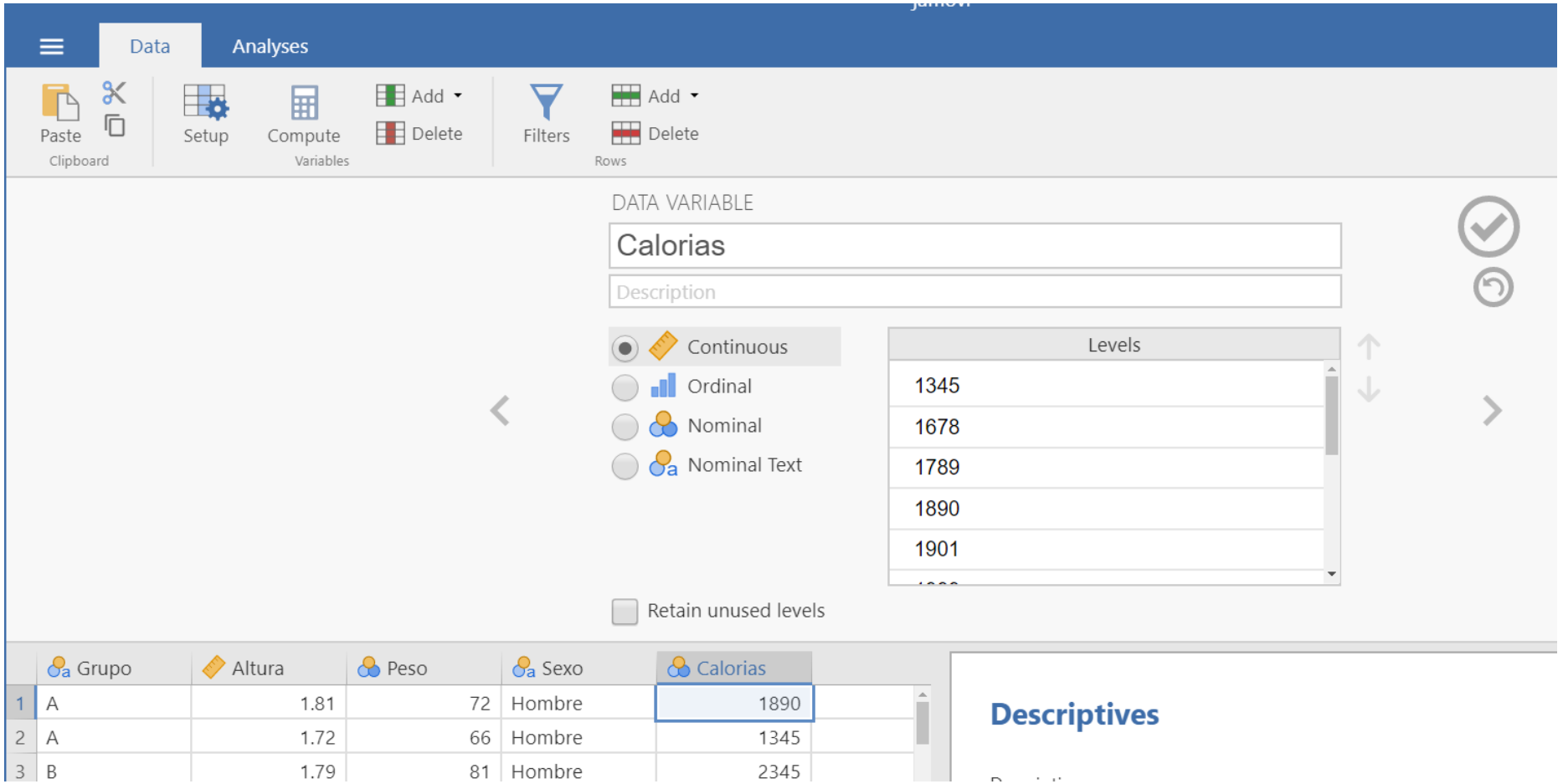
Descriptives

Descriptives

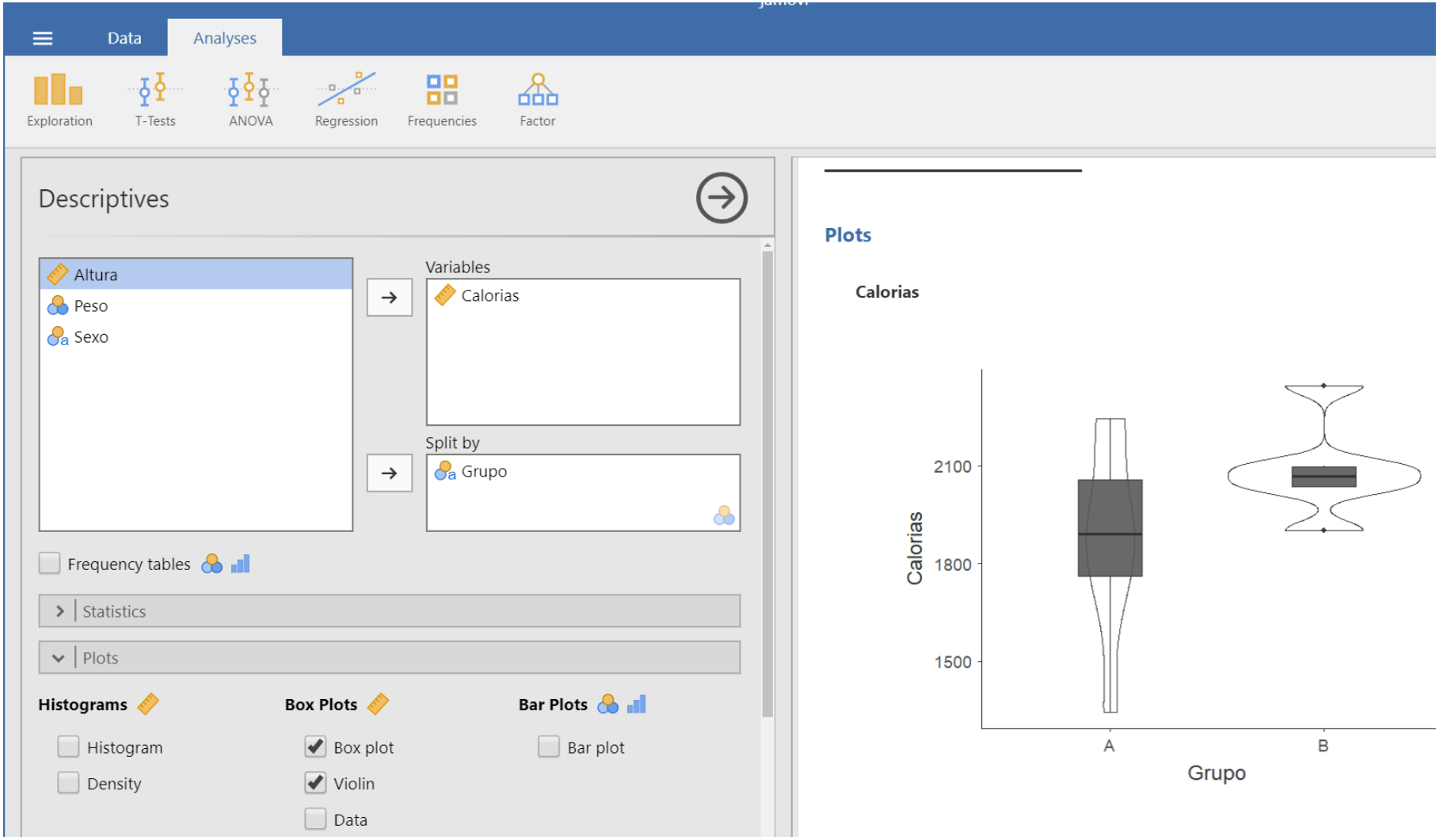
		Grupo	Calorias
N	A		8
	B		6
Missing	A		0
	B		0
Mean	A		1883
	B		2086
Median	A		1890
	B		2068
Minimum	A		1345
	B		1901
Maximum	A		2243
	B		2345

Variables categórica y numérica

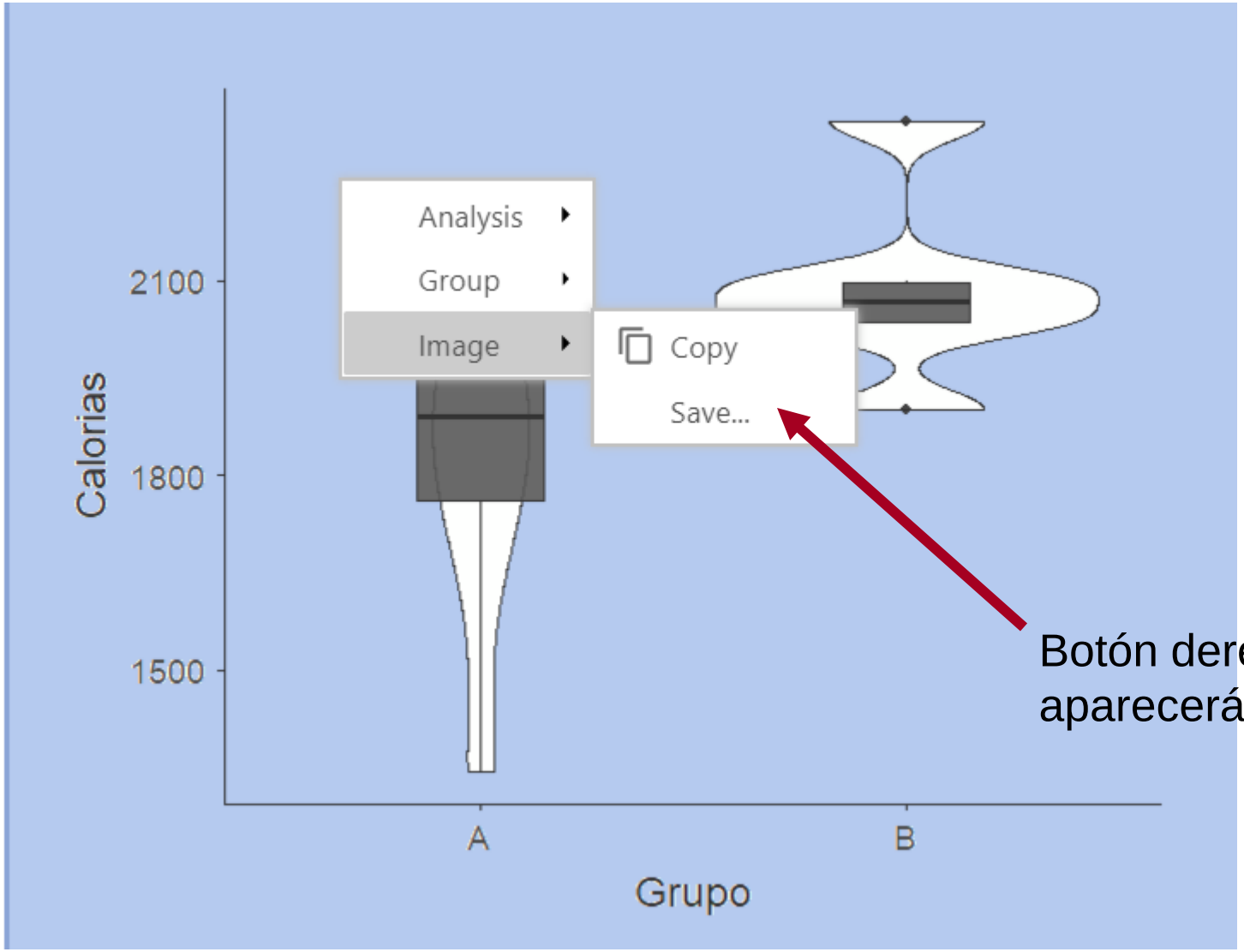
Sólo las variables continuas pueden visualizarse con boxplots



Variables categórica y numérica



Guardar/copiar plots



Botón derecho y aparecerá esto

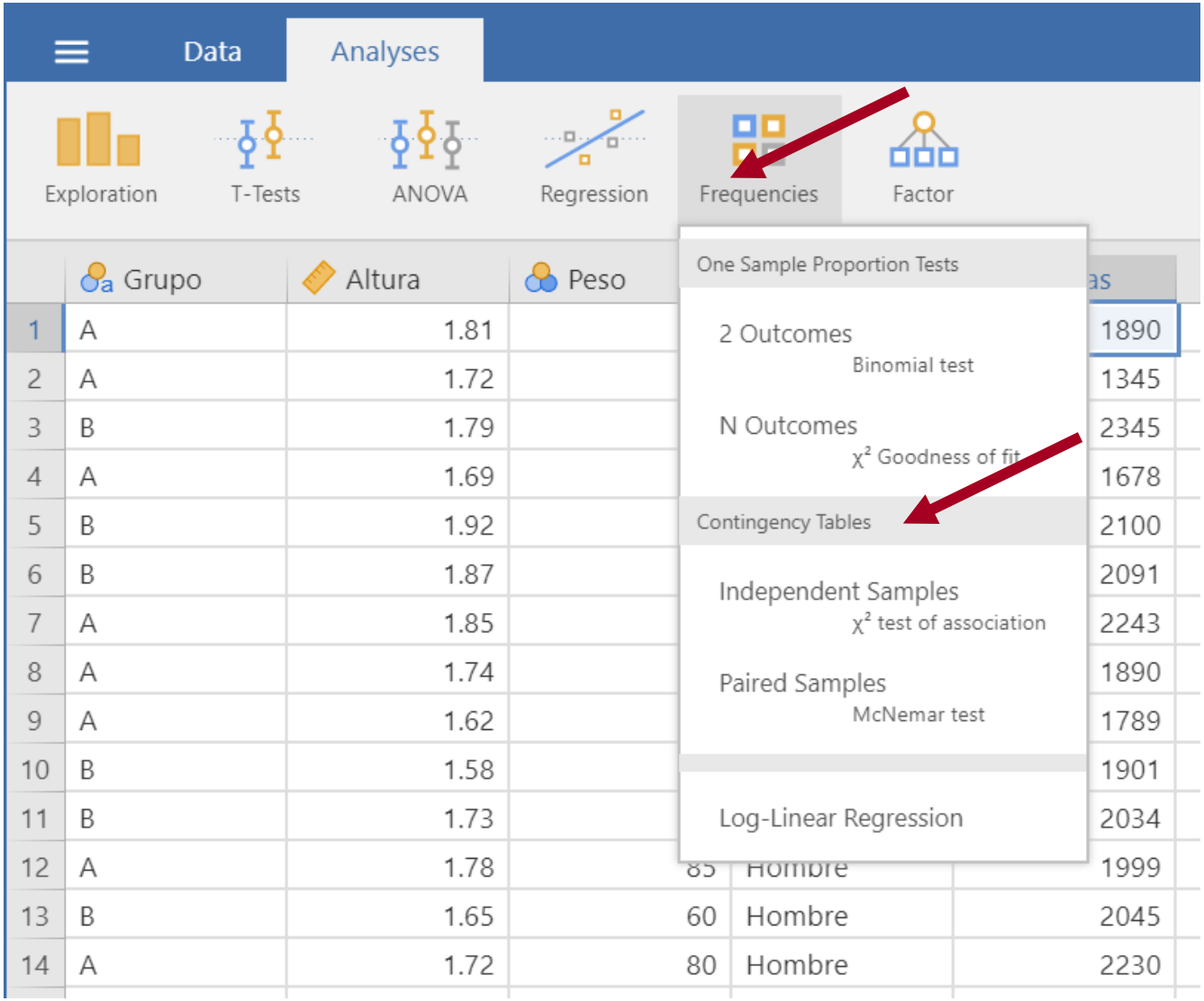
2 variables categóricas

- Como las 2 pueden definir grupos, hay que elegir una como más importante y describir la distribución de la 2ª en cada categoría de la 1ª
  - % de grupo A y B en hombres y mujeres

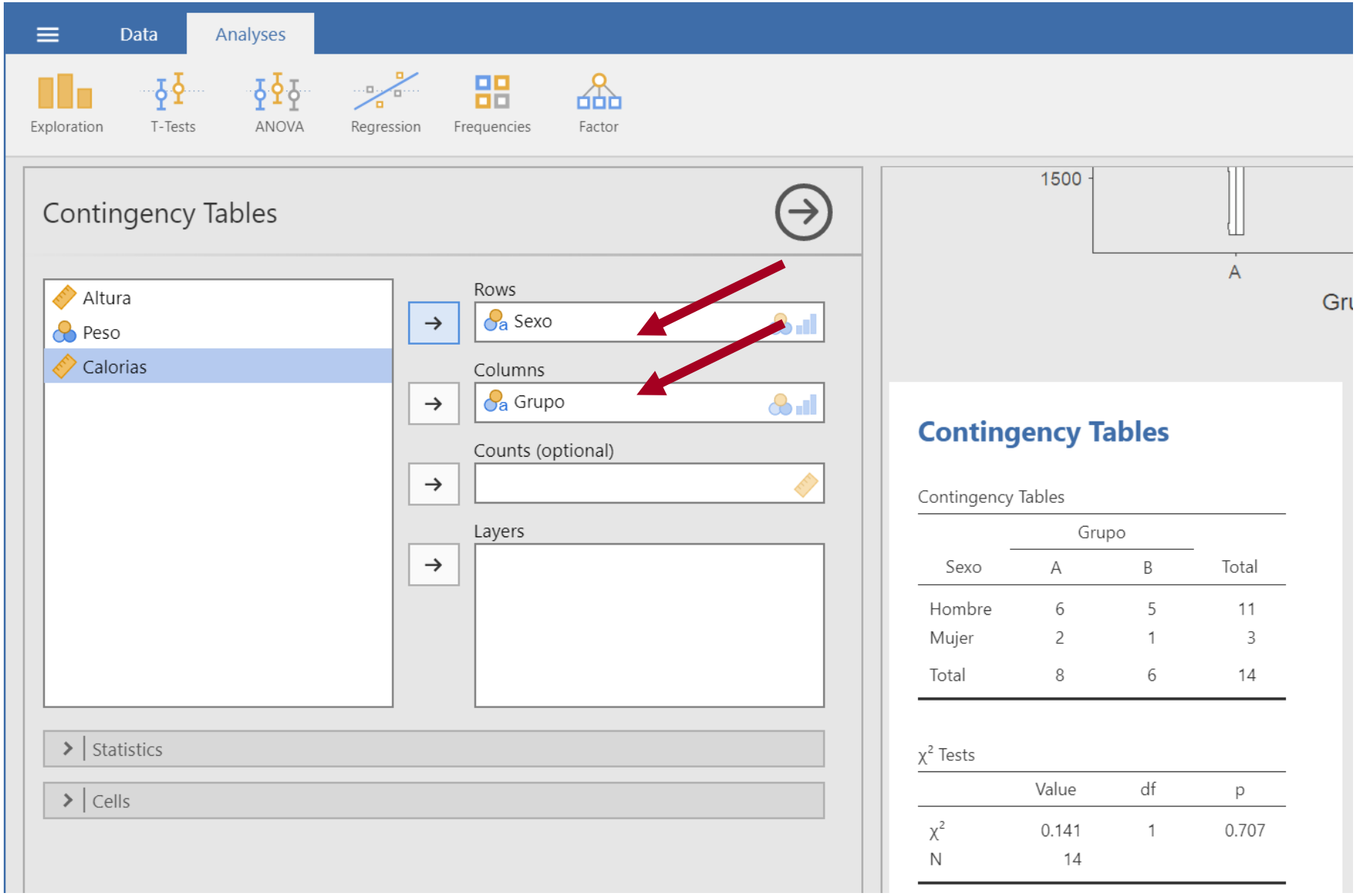
- **Tabla de frecuencias (o contingencia)**
  - Filas x Columnas

Es recomendable elegir la variable a comparar en filas

2 variables categóricas



2 variables categóricas

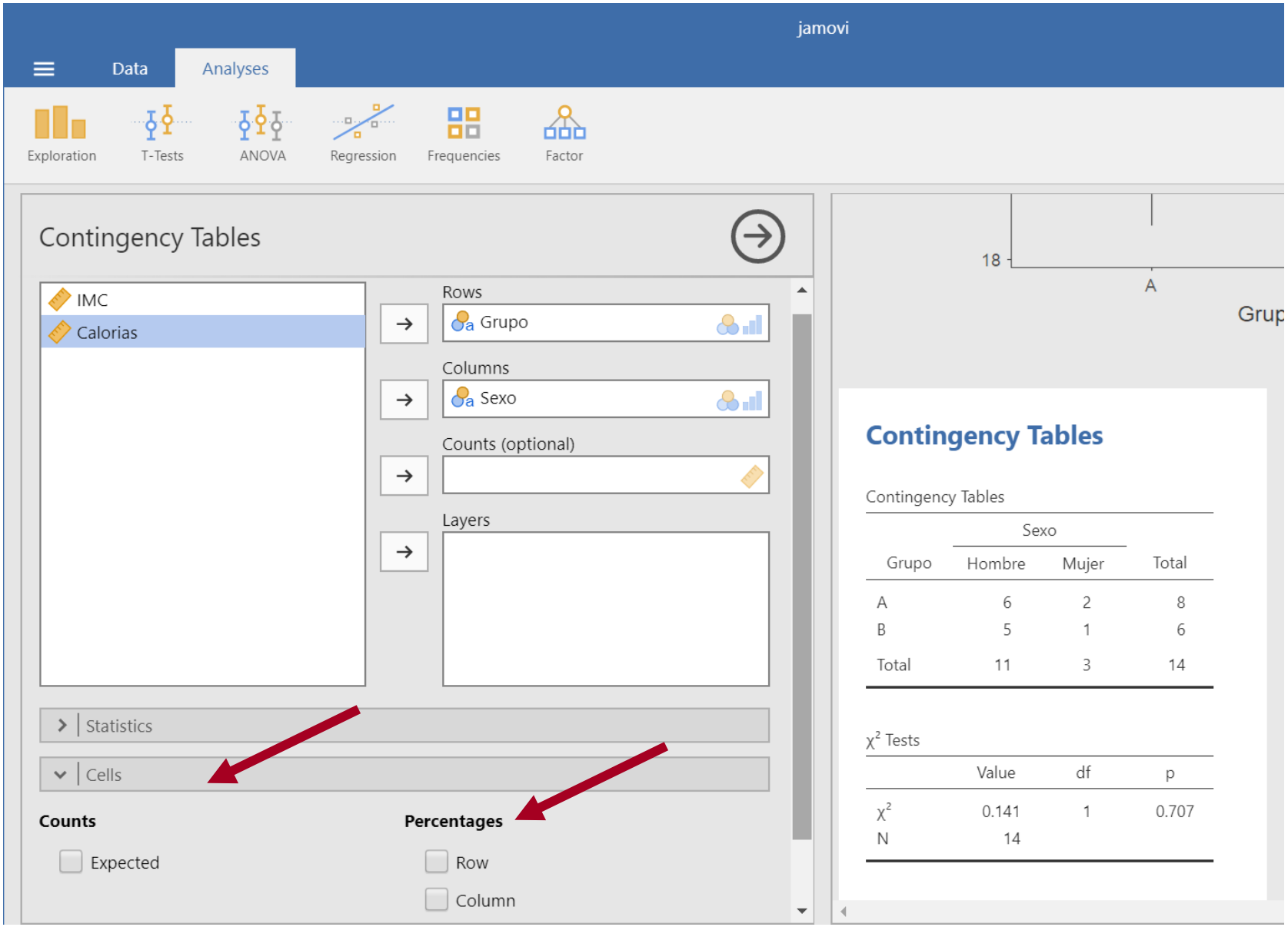


Porcentajes

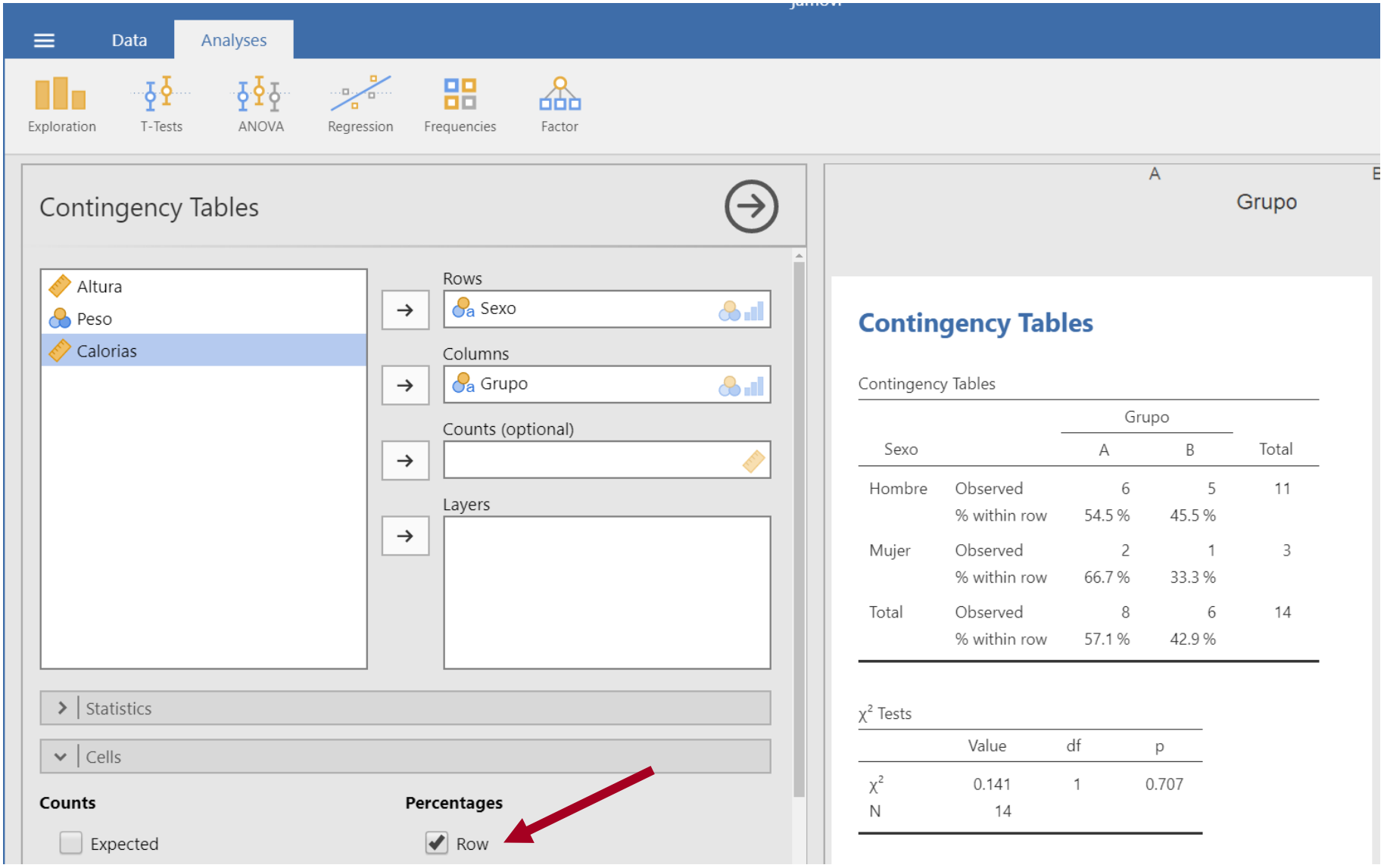
- Elegir los % de fila o columna, en función de cómo sea más fácil la interpretación
  - Sexo está en filas y Grupo en columnas
  - Si queremos **comparar grupo** (entre hombres y mujeres), elegir los porcentajes que suman **100 en cada sexo**→ filas

- Los porcentajes totales no sirven para comparar

2 variables categóricas



2 variables categóricas



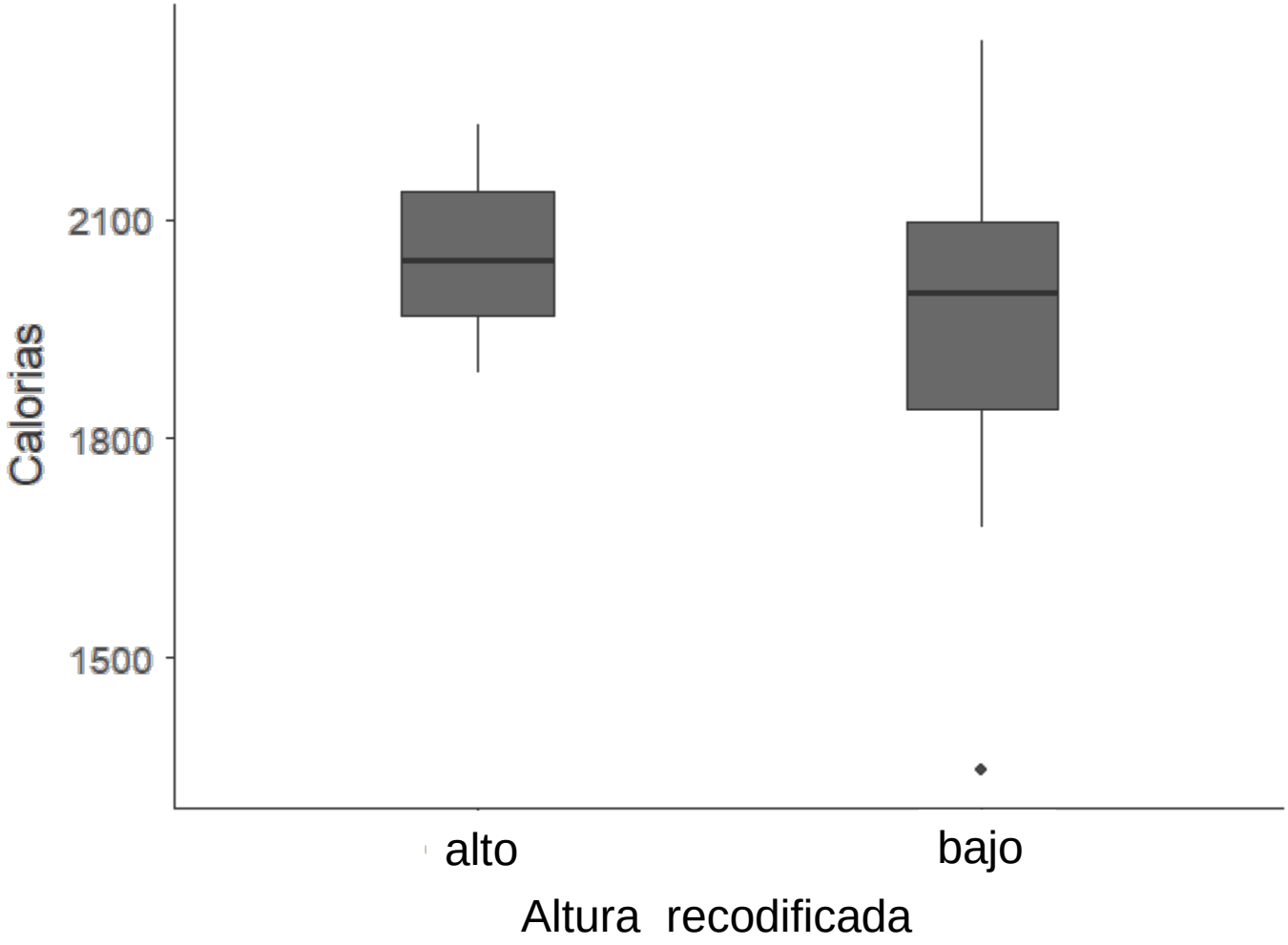
2 variables numéricas

¿Consumen más calorías las personas con mayor altura?

2 numéricas

- Necesitamos estudiar la correlación
- Se podría categorizar una de ellas y comparar las medias de la otra
  - Medias de calorías x obesidad

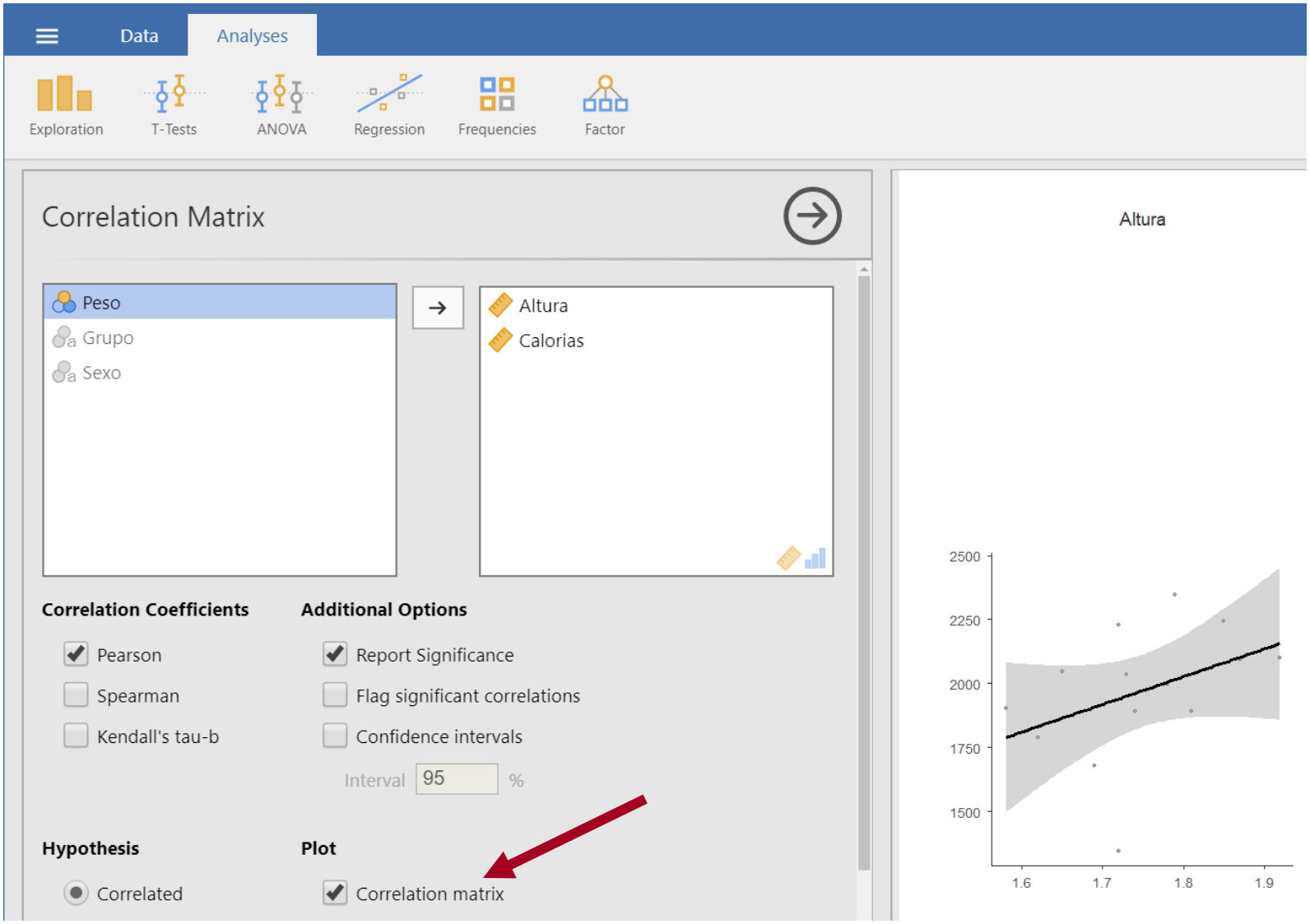
2 numéricas



Categorizar supone perder mucha información

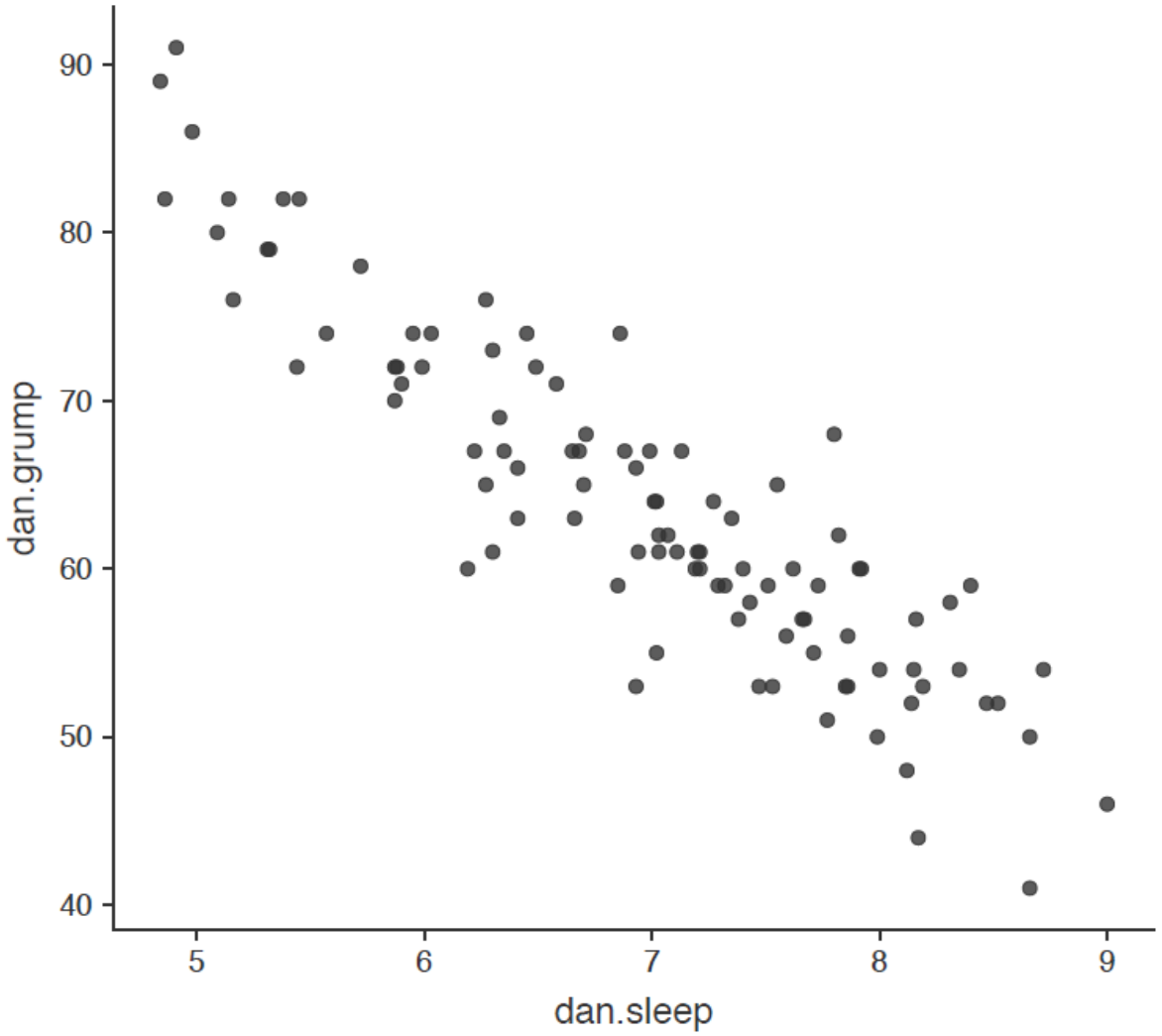
2 numéricas

Analyses -> regression -> correlation matrix



2 numéricas

La visualización se puede mejorar con el módulo ‘scatr’



Descriptiva bivariante

Tipo	**Categórica**	**Numérica**
Categórica	Tablas de frecuencia (barras)	Medias por grupo (cajas)
Numérica	Medias por grupo (cajas)	(Dispersión)