

Análisis estadístico con Jamovi aplicado a Enfermería y CAFE

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>

¿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

<

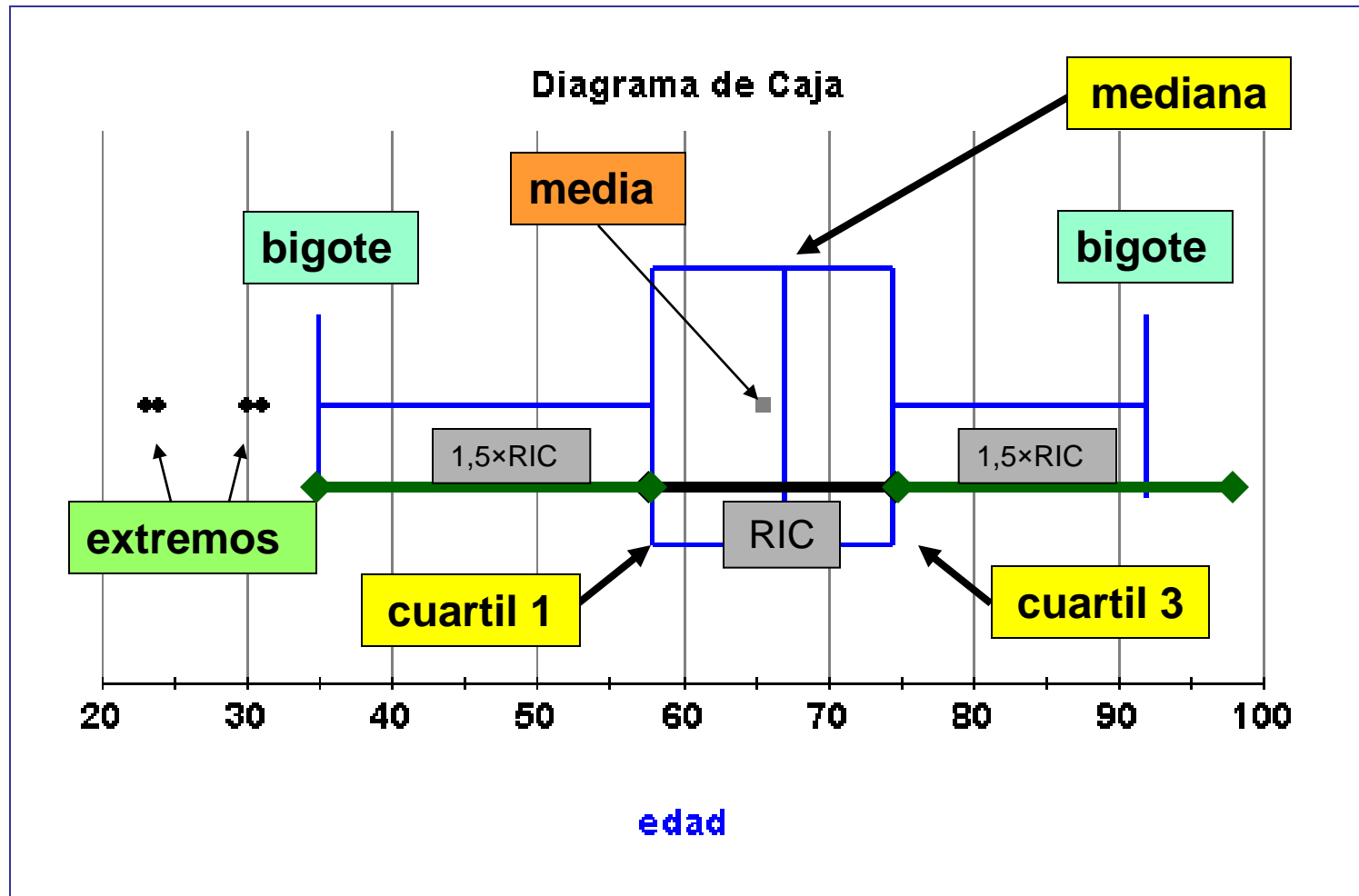
Resumen de variable continua

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

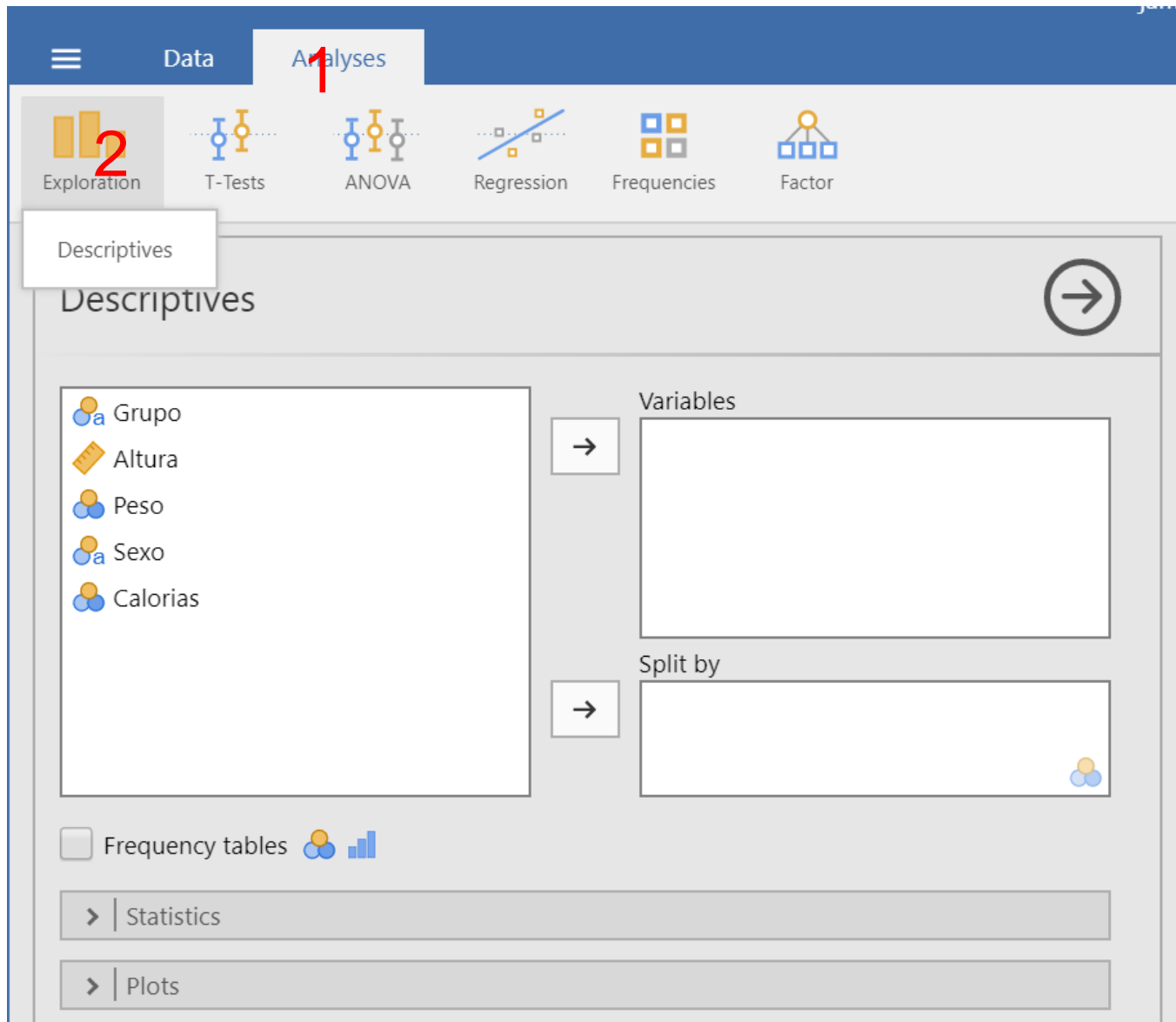
Visualización de variable continua

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

Boxplot



Estadística descriptiva



The image shows the SPSS 'Descriptives' dialog box. At the top, the 'Data' menu is open, and the 'Analyses' option is highlighted with a red '1'. Below this, the 'Exploration' icon is highlighted with a red '2'. The 'Descriptives' dialog box is open, showing a list of variables on the left: 'Grupo', 'Altura', 'Peso', 'Sexo', and 'Calorias'. The 'Variables' box on the right is empty. The 'Split by' box is also empty. At the bottom, there are checkboxes for 'Frequency tables', 'Statistics', and 'Plots'. The 'Statistics' and 'Plots' sections are currently collapsed.

Analyses

Exploration

T-Tests

ANOVA

Regression

Frequencies

Factor

Descriptives

Descriptives

Grupo

Altura

Peso

Sexo

Calorias

Variables

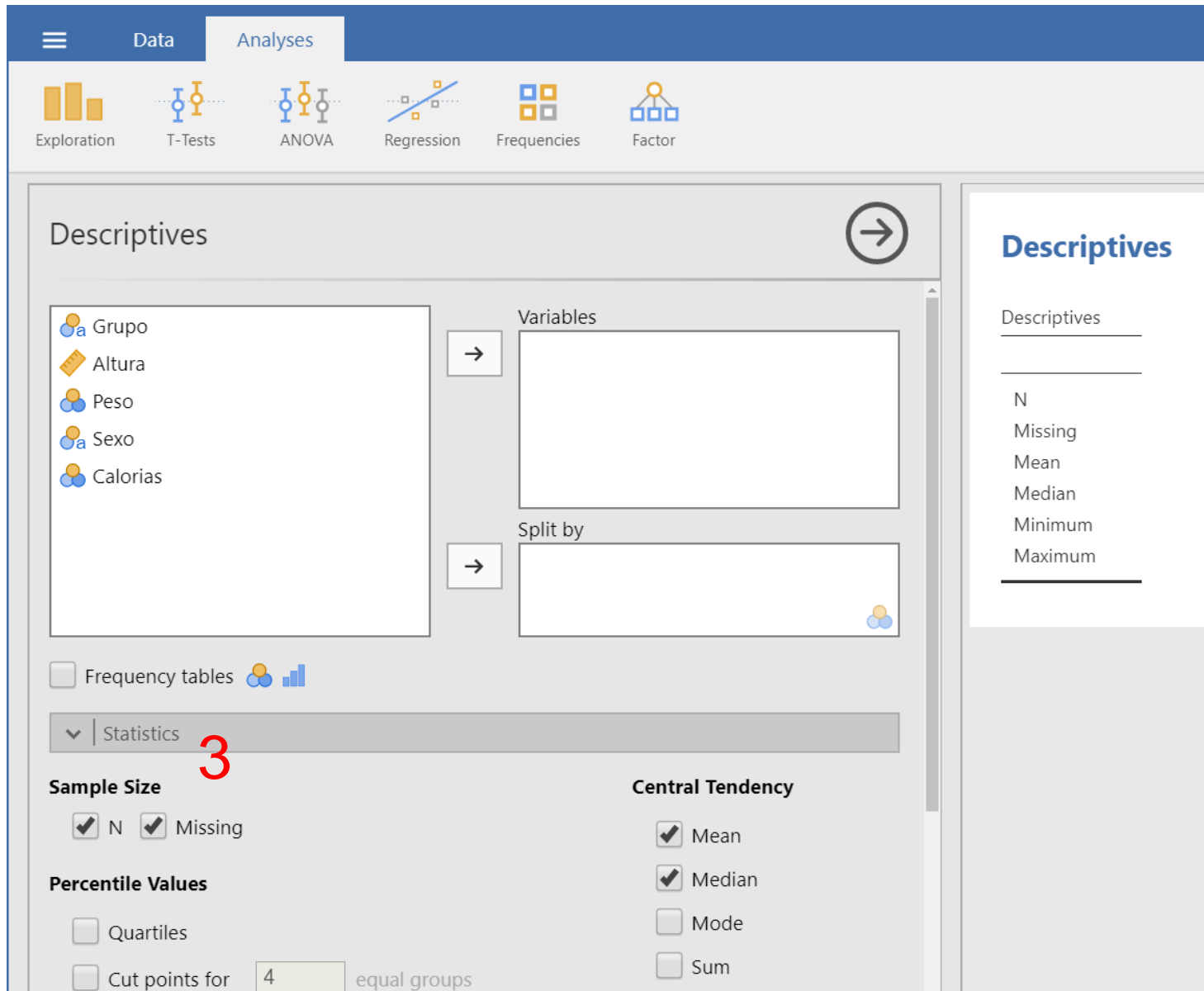
Split by

☐ Frequency tables

> | Statistics

> | Plots

Variables continuas



The screenshot shows the SPSS 'Descriptives' dialog box. The 'Data' tab is selected. The 'Analyses' section includes icons for Exploration, T-Tests, ANOVA, Regression, Frequencies, and Factor. The 'Descriptives' section is active, showing a list of variables on the left: Grupo, Altura, Peso, Sexo, and Calorias. The 'Variables' box is empty, and the 'Split by' box is also empty. Below the variable list, there is a checkbox for 'Frequency tables' and a 'Statistics' dropdown menu. The 'Statistics' dropdown is expanded, showing a list of statistical measures: N, Missing, Mean, Median, Minimum, and Maximum. The 'Sample Size' section has checkboxes for 'N' and 'Missing'. The 'Percentile Values' section has a checkbox for 'Quartiles' and a checkbox for 'Cut points for' with a value of 4 and 'equal groups'. The 'Central Tendency' section has checkboxes for 'Mean', 'Median', 'Mode', and 'Sum'. A red number '3' is overlaid on the 'Statistics' dropdown menu.

Descriptives

Variables

Split by

☐ Frequency tables

Statistics

Sample Size

☒ N ☒ Missing

Percentile Values

☐ Quartiles

☐ Cut points for 4 equal groups

Central Tendency

☒ Mean

☒ Median

☐ Mode

☐ Sum

Descriptives

Descriptives

N

Missing

Mean

Median

Minimum


Maximum


Variables continuas


≡


Data


Analyses


Exploration

T-Tests


ANOVA


Regression


Frequencies


Factor

Descriptives

Grupo


Peso

Sexo

Calorias

→


Variables

Altura

4

Split by

→

☐ Frequency tables 

Statistics

Sample Size

☒ N ☒ Missing

Percentile Values

Central Tendency

☒ Mean ☒ Median

Descriptives

Descriptives

	Altura
N	14
Missing	0
Mean	1.75
Median	1.73

Variables continuas

Software interface for data analysis, showing the "Analyses" tab and the "Descriptives" sub-tab.

Analyses Tab:

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

Descriptives Sub-tab:

Variables:

- Grupo
- Peso
- Sexo
- Calorias

Split by:

-

Frequency tables: ☐

Statistics:

- > | Statistics

Plots:

- ▼ | Plots

Histograms:

- ☐ Histogram
- ☐ Density

Box Plots:

- ☒ Box plot
- ☐ Violin
- ☐ Data

Bar Plots:

- ☐ Bar plot

Descriptives Plots:

Altura

The box plot for 'Altura' shows the distribution of height values. The y-axis is labeled 'Altura' and ranges from 1.6 to 1.9. The box plot indicates a median around 1.73, with the interquartile range (IQR) spanning from approximately 1.70 to 1.81. The whiskers extend from about 1.58 to 1.92.


Statistic	Value (approx.)
Minimum	1.58
First Quartile (Q1)	1.70
Median	1.73
Third Quartile (Q3)	1.81
Maximum	1.92


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

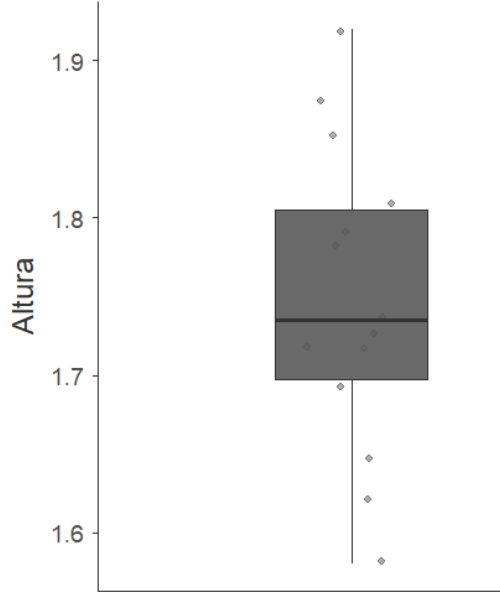
Bar Plots 

☐ Bar plot

Descriptives

Plots

Altura

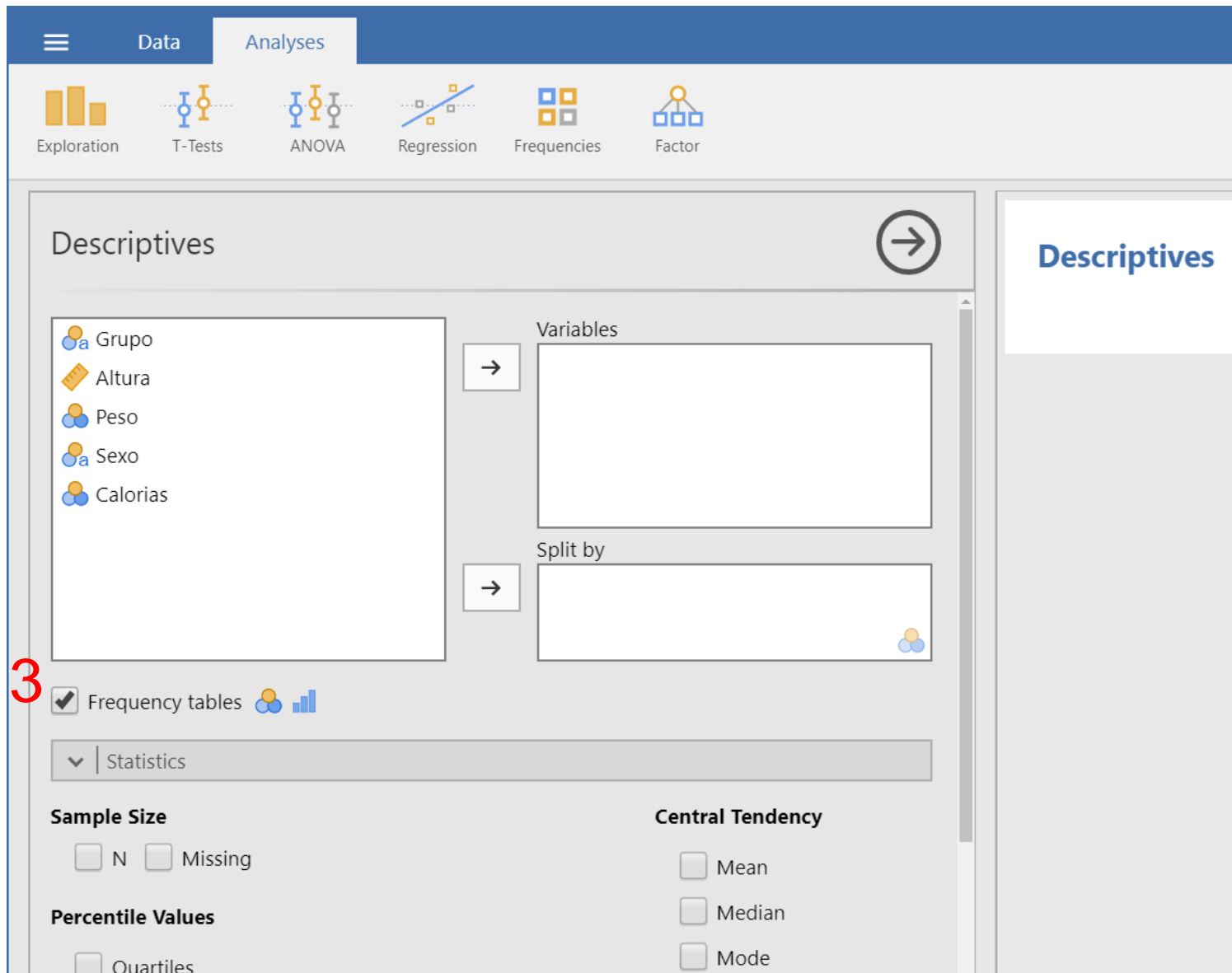


5

Variables categóricas

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

Categóricas



The image shows the SPSS 'Descriptives' dialog box. The 'Data' tab is selected. The 'Variables' list on the left contains 'Grupo', 'Altura', 'Peso', 'Sexo', and 'Calorias'. The 'Statistics' section is expanded, showing options for 'Frequency tables', 'Sample Size', 'Percentile Values', 'Central Tendency', and 'Dispersion'. The 'Frequency tables' checkbox is checked. The 'Sample Size' section has 'N' and 'Missing' checkboxes. The 'Percentile Values' section has 'Quartiles' checkbox. The 'Central Tendency' section has 'Mean', 'Median', and 'Mode' checkboxes. The 'Dispersion' section has 'Variance' and 'Std. Deviation' checkboxes. The 'Display' section has 'Ordered by' dropdown set to 'Ascending' and 'Display' checkbox checked. The 'Format' section has 'Display' checkbox checked. The 'Save' section has 'Save as type' dropdown set to 'Descriptives' and 'Save' checkbox checked. The 'OK' button is highlighted.

Descriptives

Variables

Split by

3 ☒ Frequency tables

Statistics

Sample Size

☐ N ☐ Missing

Percentile Values

☐ Quartiles

Central Tendency

☐ Mean

☐ Median

☐ Mode

Dispersion

☐ Variance ☐ Std. Deviation

Display

☒ Display

Format

☒ Display

Save

☒ Save as type Descriptives

☒ Save


OK


Categóricas


≡


Data


Analyses


 Exploration

 T-Tests


 ANOVA


 Regression


 Frequencies


 Factor

Descriptives


 Grupo

 Altura

 Peso

 Calorias


→

 Sexo

4

→

Split by

☒ Frequency tables 

Statistics

Sample Size

☐ N ☐ Missing

Central Tendency

☐ Mean

Descriptives

Frequencies

Frequencies of Sexo

Levels	Counts	% of Total	Cumulative %
Hombre	11	78.6 %	78.6 %
Mujer	3	21.4 %	100.0 %

Categóricas

☰

Data

Analyses

Exploration

T-Tests

ANOVA

Regression

Frequencies

Factor

Descriptives

→

Grupo

Altura

Peso

Calorias

→

Variables

Sexo

Split by

☐ Frequency tables

Statistics

Plots

Histograms

Box Plots

Bar Plots

☐ Histogram

☐ Density

☐ Box plot

☐ Violin

☐ Data

☒ Bar plot

Descriptives

Plots

Sexo

Sexo	counts
Hombre	10
Mujer	3

5

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

The screenshot shows the SPSS 'Descriptives' dialog box. On the left, a list of variables includes 'Altura', 'Peso', and 'Sexo'. The 'Variables' box on the right contains 'Calorias'. The 'Split by' box contains 'Grupo', which is highlighted by a red arrow. At the bottom, there are checkboxes for 'Frequency tables', 'Statistics', and 'Plots'. To the right of the dialog box, a preview of the output table is shown.

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

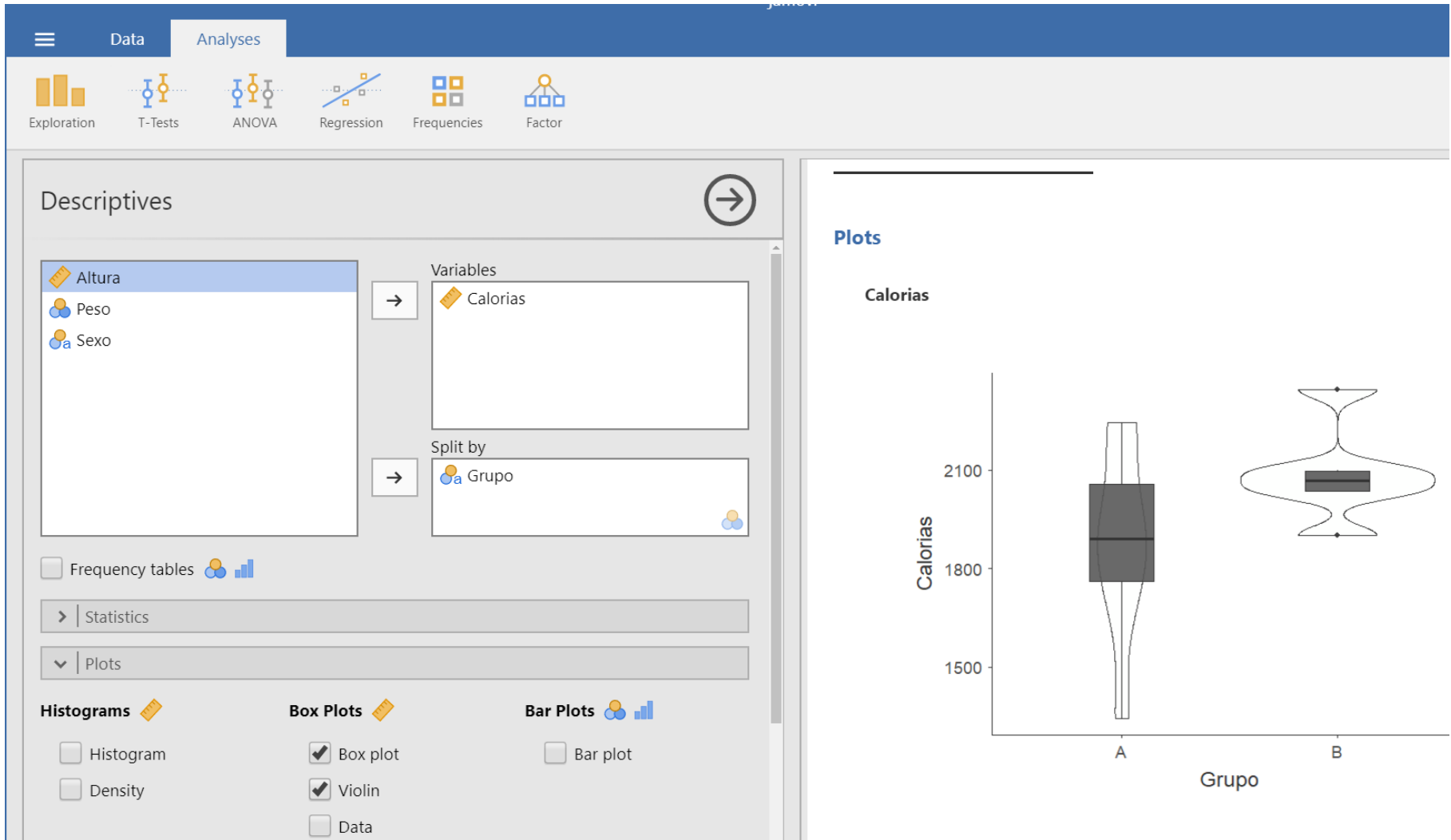
Variables categórica y numérica

Sólo las variables continuas pueden visualizarse con boxplots

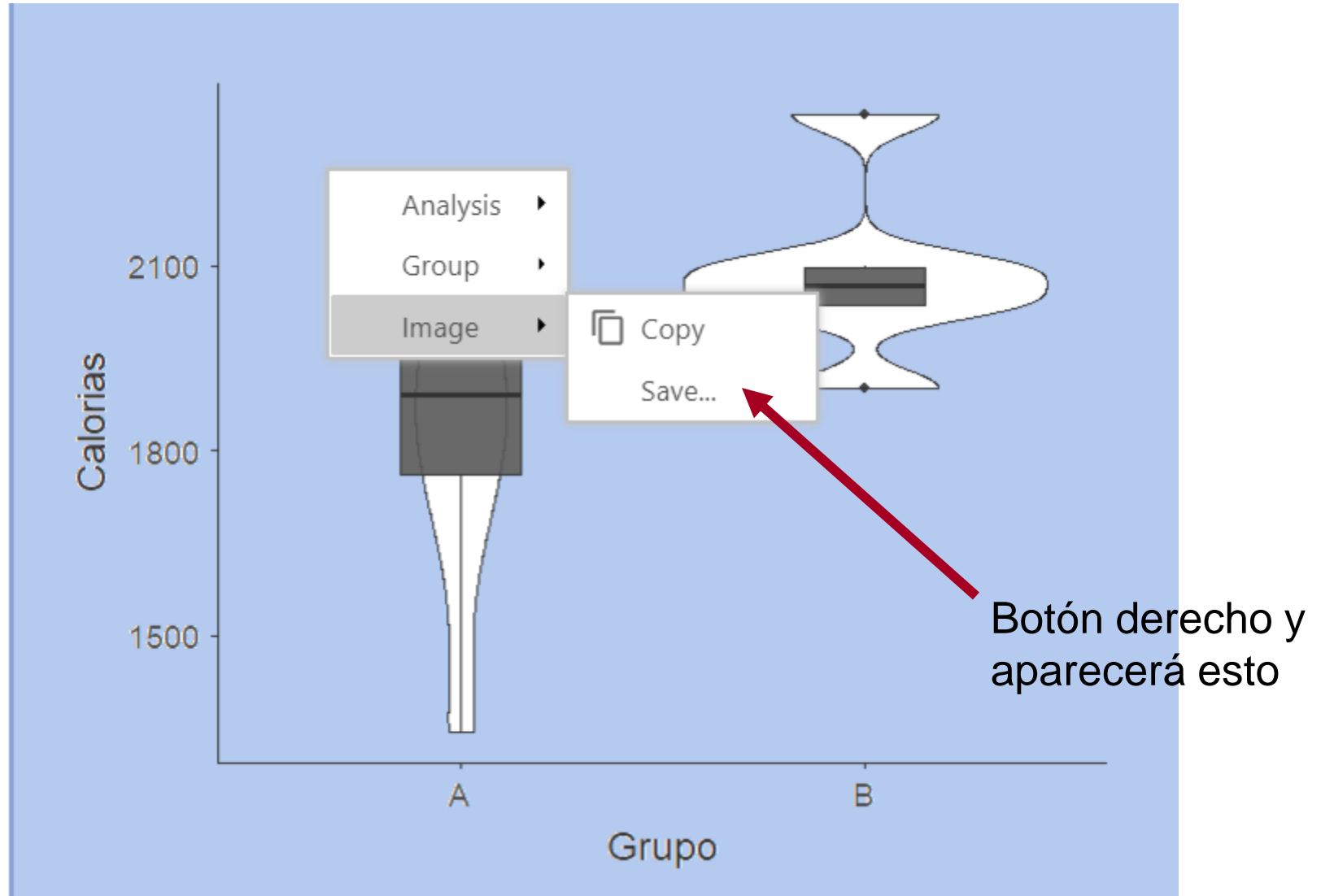
The screenshot shows a software interface with a top navigation bar containing 'Data' and 'Analyses' tabs. Below this is a toolbar with icons for 'Paste', 'Clipboard', 'Setup', 'Compute', 'Variables', 'Filters', and 'Rows'. The main workspace is titled 'DATA VARIABLE' and shows the variable 'Calorias' selected. Below the variable name is a 'Description' field. To the right of the variable name is a 'Levels' list containing the values: 1345, 1678, 1789, 1890, and 1901. To the left of the 'Levels' list are radio buttons for variable types: 'Continuous' (selected), 'Ordinal', 'Nominal', and 'Nominal Text'. Below these is a checkbox labeled 'Retain unused levels'. At the bottom of the interface is a data table with columns: 'Grupo', 'Altura', 'Peso', 'Sexo', and 'Calorias'. The table contains three rows of data. The 'Calorias' column is highlighted with a blue border. To the right of the table is a 'Descriptives' panel.

	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

Variables categórica y numérica



Guardar/copiar plots



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

The screenshot shows a statistical software interface with a blue header bar containing 'Data' and 'Analyses' tabs. Below the header, there are icons for various analyses: Exploration, T-Tests, ANOVA, Regression, Frequencies, and Factor. The 'Frequencies' icon is highlighted with a red arrow. A dropdown menu is open from the 'Frequencies' icon, showing the following options:

- One Sample Proportion Tests
 - 2 Outcomes
 - Binomial test
 - N Outcomes
 - χ^2 Goodness of fit
- Contingency Tables (highlighted with a red arrow)
- Independent Samples
 - χ^2 test of association
- Paired Samples
 - McNemar test
- Log-Linear Regression

The background shows a data table with columns: Grupo, Altura, and Peso. The data is as follows:


	Grupo	Altura	Peso
1	A	1.81	
2	A	1.72	
3	B	1.79	
4	A	1.69	
5	B	1.92	
6	B	1.87	
7	A	1.85	
8	A	1.74	
9	A	1.62	
10	B	1.58	
11	B	1.73	
12	A	1.78	
13	B	1.65	
14	A	1.72	


2 variables categóricas


≡


Data


Analyses


Exploration

T-Tests


ANOVA


Regression


Frequencies

Factor

Contingency Tables

Altura

Peso

Calorias

→

Sexo

→

Grupo

→

Counts (optional)


→

Layers

> | Statistics

> | Cells

1500



A

Gr

Contingency Tables

Contingency Tables

Sexo	Grupo		Total
	A	B	
Hombre	6	5	11
Mujer	2	1	3
Total	8	6	14

χ^2 Tests

	Value	df	p
χ^2	0.141	1	0.707
N	14		

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

jamovi

Menu: Data, Analyses

Tools: Exploration, T-Tests, ANOVA, Regression, Frequencies, Factor

Contingency Tables

Variables: IMC, Calorias

Rows: Grupo

Columns: Sexo

Counts (optional):

Layers:

Statistics: > | Statistics

Cells: ▾ | Cells

Counts: ☐ Expected

Percentages: ☐ Row, ☐ Column

Contingency Tables

Contingency Tables

Grupo	Sexo		Total
	Hombre	Mujer	
A	6	2	8
B	5	1	6
Total	11	3	14

χ^2 Tests


	Value	df	p
χ^2	0.141	1	0.707
N	14		


2 variables categóricas


≡


Data


Analyses


 Exploration

 T-Tests


 ANOVA


 Regression


 Frequencies

 Factor

Contingency Tables


 Altura

 Peso

 Calorias


→

Rows

 Sexo

→

Columns

 Grupo

→

Counts (optional)

→

Layers

> | Statistics

▼ | Cells

Counts

☐ Expected

Percentages

☒ Row

A

Grupo

Contingency Tables

Contingency Tables

		Grupo		Total
		A	B	
Hombre	Observed	6	5	11
	% within row	54.5 %	45.5 %	
Mujer	Observed	2	1	3
	% within row	66.7 %	33.3 %	
Total	Observed	8	6	14
	% within row	57.1 %	42.9 %	

χ² Tests

	Value	df	p
χ ²	0.141	1	0.707
N	14		

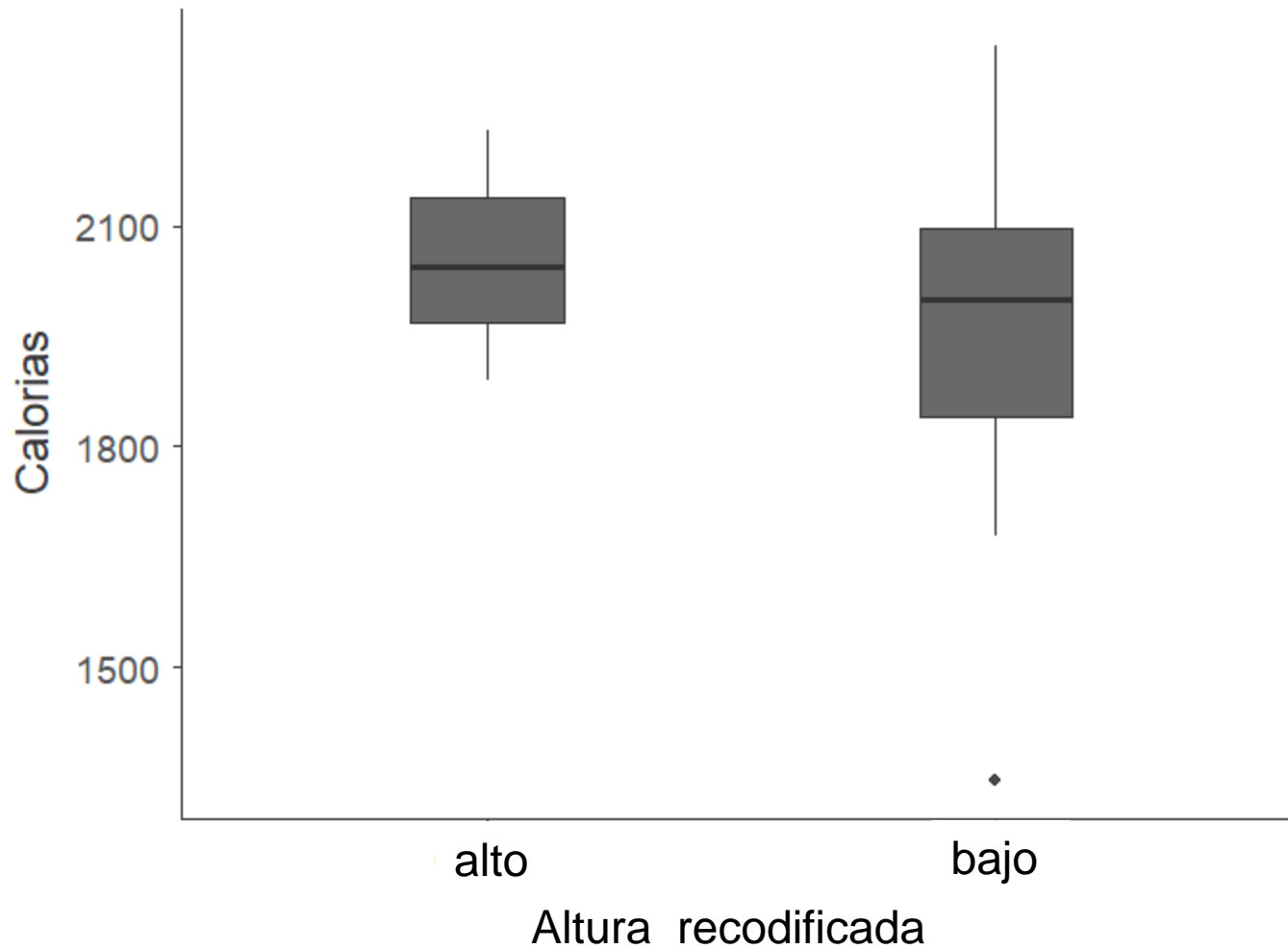
2 variables numéricas

¿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

The screenshot displays a statistical software interface with a top navigation bar containing 'Data' and 'Analyses' tabs. Below the tabs is a toolbar with icons for 'Exploration', 'T-Tests', 'ANOVA', 'Regression', 'Frequencies', and 'Factor'. The main window is titled 'Correlation Matrix' and features a list of variables on the left: 'Peso', 'Grupo', and 'Sexo'. A right arrow points to a list of selected variables: 'Altura' and 'Calorias'. Below the variable lists are two sections: 'Correlation Coefficients' and 'Additional Options'. The 'Correlation Coefficients' section includes checkboxes for 'Pearson' (checked), 'Spearman', and 'Kendall's tau-b'. The 'Additional Options' section includes checkboxes for 'Report Significance' (checked), 'Flag significant correlations', and 'Confidence intervals'. Below these is an 'Interval' field set to '95 %'. At the bottom left, the 'Hypothesis' section has a radio button for 'Correlated'. The 'Plot' section has a checked checkbox for 'Correlation matrix', which is highlighted by a red arrow. To the right of the dialog box is a scatter plot titled 'Altura' showing the relationship between 'Altura' (Height) on the x-axis and 'Calorias' (Calories) on the y-axis. The plot includes a regression line and a shaded confidence interval.

Correlation Matrix

Variables: **Peso**, **Grupo**, **Sexo**

Selected Variables: **Altura**, **Calorias**

Correlation Coefficients

- ☒ Pearson
- ☐ Spearman
- ☐ Kendall's tau-b

Additional Options

- ☒ Report Significance
- ☐ Flag significant correlations
- ☐ Confidence intervals

Interval: **95** %

Hypothesis

- ☐ Correlated

Plot

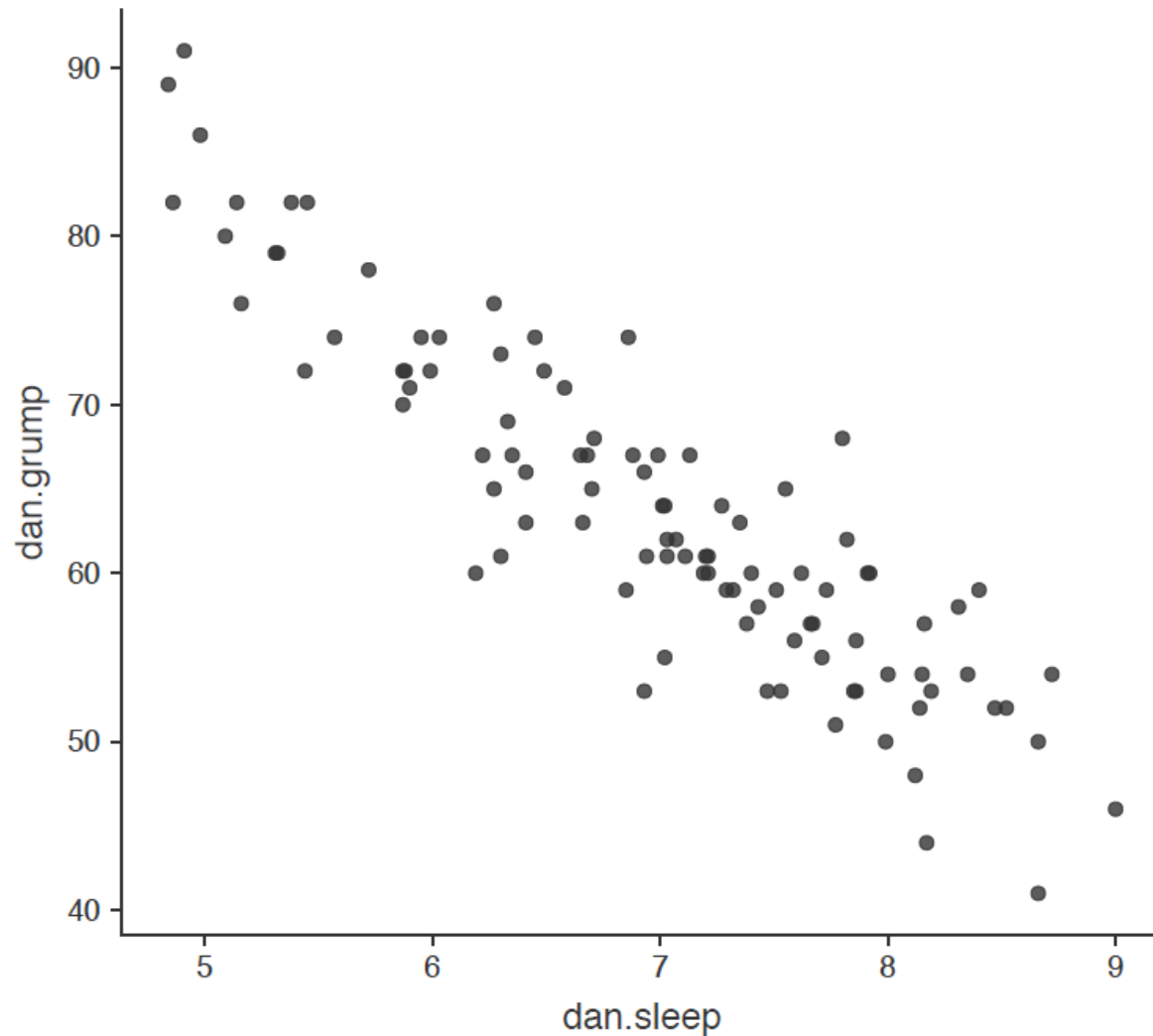
- ☒ Correlation matrix

Altura

Scatter plot showing the relationship between **Altura** (Height) and **Calorias** (Calories). The plot includes a regression line and a shaded confidence interval.

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)