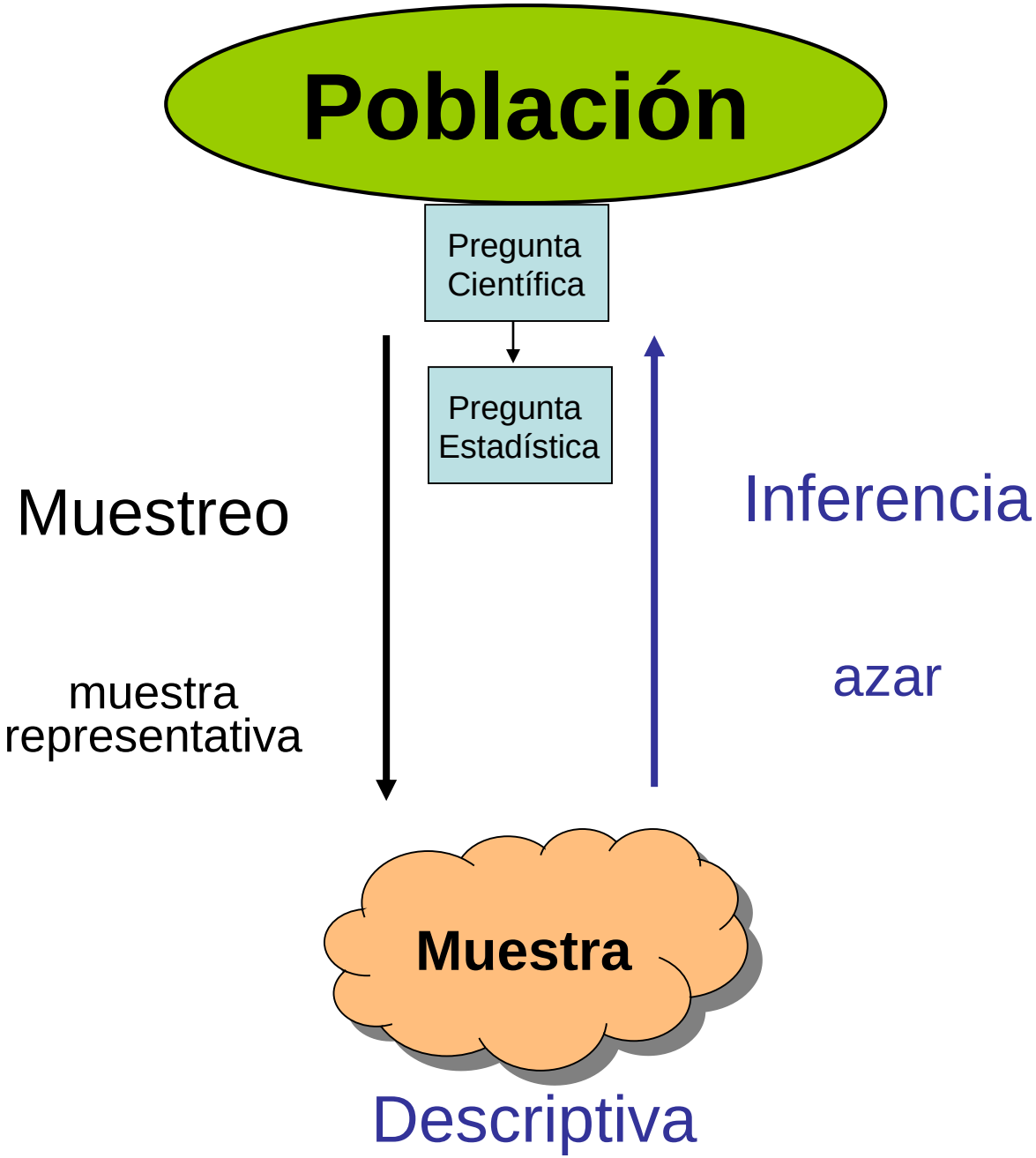
Análisis estadístico con Jamovi

Parte III: Pruebas de Hipótesis

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


Población
Pregunta Científica
Pregunta Estadística
Muestreo
muestra representativa
Inferencia
azar
Muestra
Descriptiva

Tests para datos categóricos (Chi-cuadrado)

Pulso, ejercicio físico y fármaco

Factores que se asocian al aumento de la frecuencia cardiaca (pulso.txt)

- IB: Número de identificación
- Sexo: 1=Hombre; 2=Mujer
- Fumador: 1=sí; 2=no
- Edad: Edad en años
- FC1: Frecuencia cardiaca antes del ejercicio
- FC2: Frecuencia cardiaca después del ejercicio
- FC2FC1 : Incremento de la frecuencia cardiaca
- Status : Nivel de entrenamiento físico 1, 2 ó 3
- Farmaco : 1=Fármaco1; 2=Fármaco2

Pulso, ejercicio físico y fármaco

- HIPÓTESIS: Sospecho que en mi población la prevalencia de fumadores es menor que en la población general (π=0.3)
- PREGUNTA CIENTÍFICA: ¿La prevalencia de fumadores es de 0.3?
- PREGUNTA ESTADÍSTICA:

H_0: π=0.3

- MUESTRA: la que hay en ‘pulso.txt’

Estadístico: Una proporción

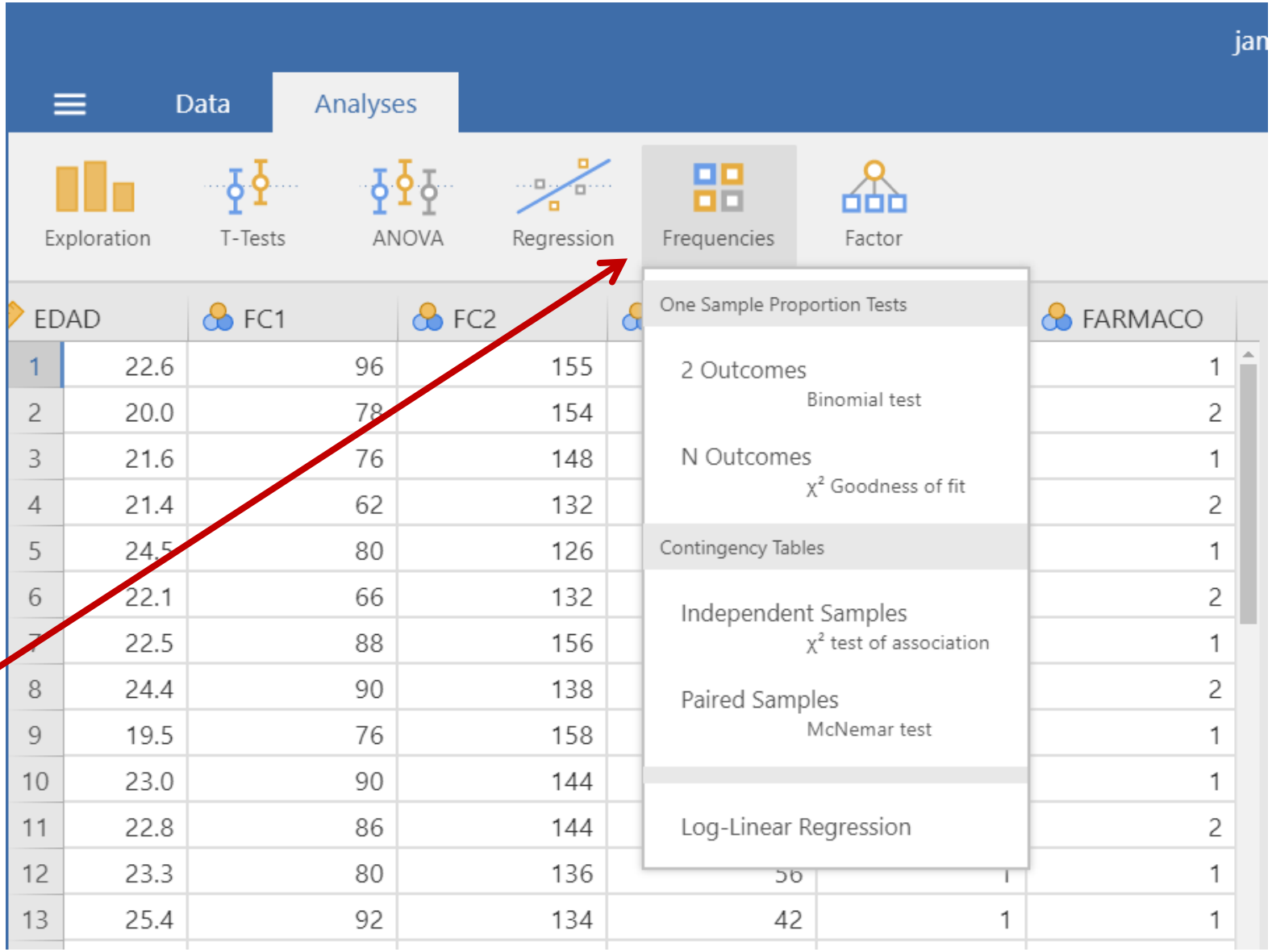
Contraste de hipótesis (1 proporción)

- Un 30% de la muestra es fumador:

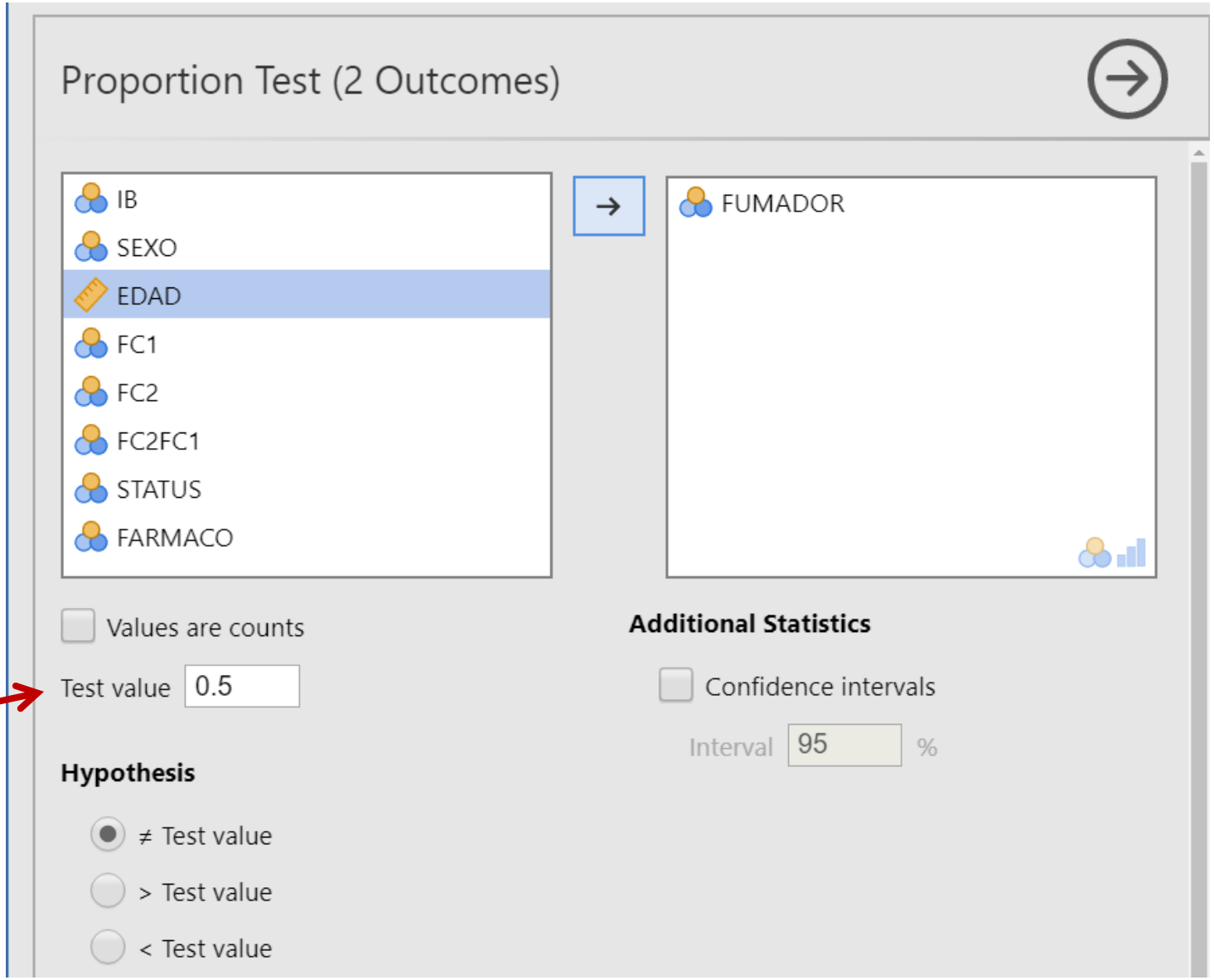
$$H_0: \pi = 0,3$$

- Veamos cómo obtener el p-valor del test

Cálculo del P-valor



Cálculo del P-valor



Cálculo del P-valor

Proportion Test (2 Outcomes)

Binomial Test

	Level	Count	Total	Proportion	p
FUMADOR	1	16	40	0.400	0.268
	2	24	40	0.600	0.268

Note. H_a is proportion ≠ 0.5

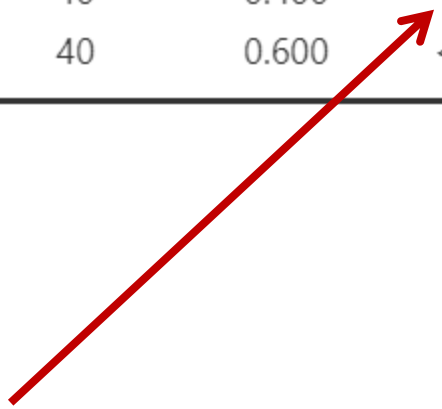
Cálculo del P-valor

Proportion Test (2 Outcomes)

Binomial Test

	Level	Count	Total	Proportion	p
FUMADOR	1	16	40	0.400	0.170
	2	24	40	0.600	< .001

Note. H_a is proportion ≠ 0.3

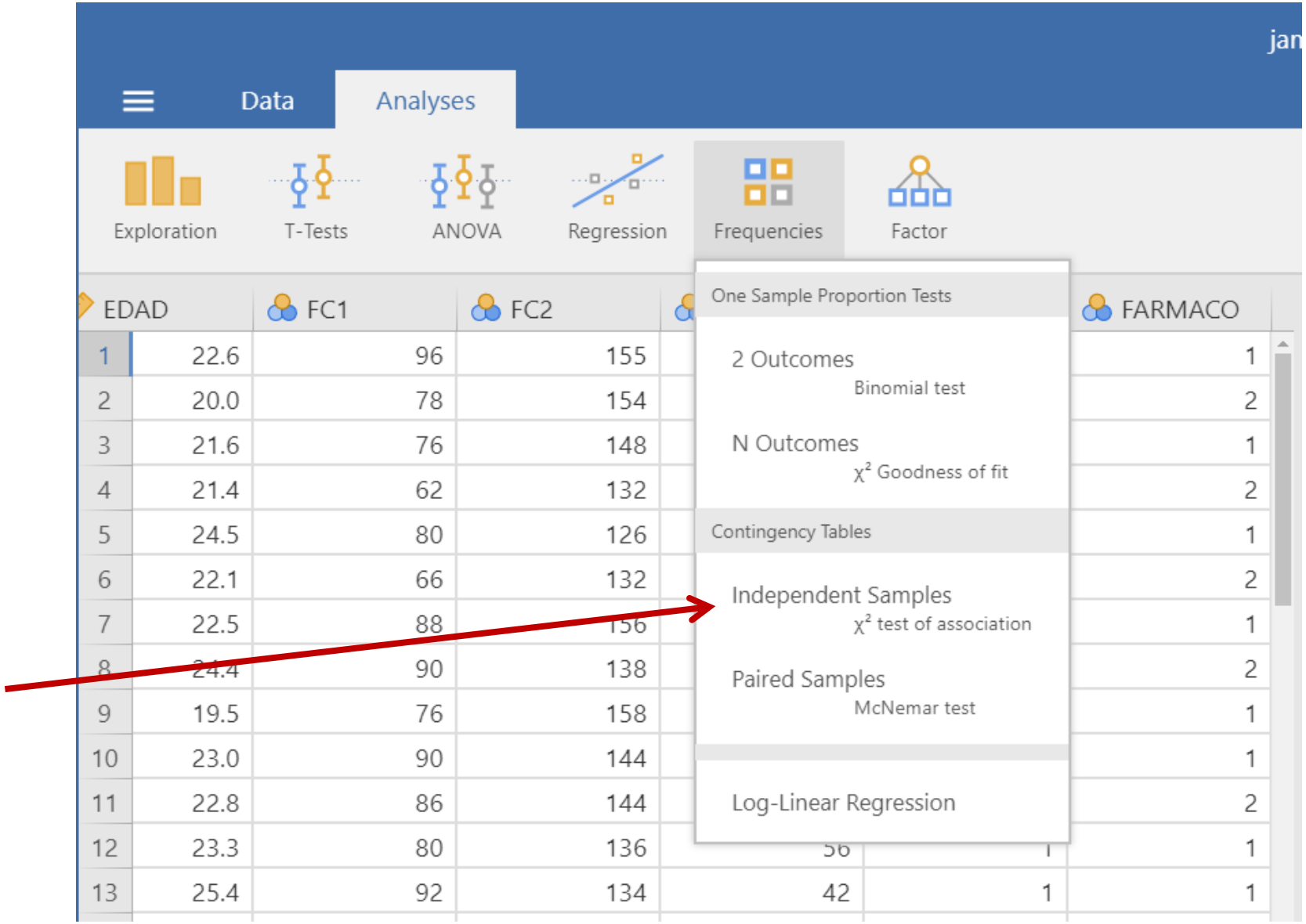


Dos variables categóricas

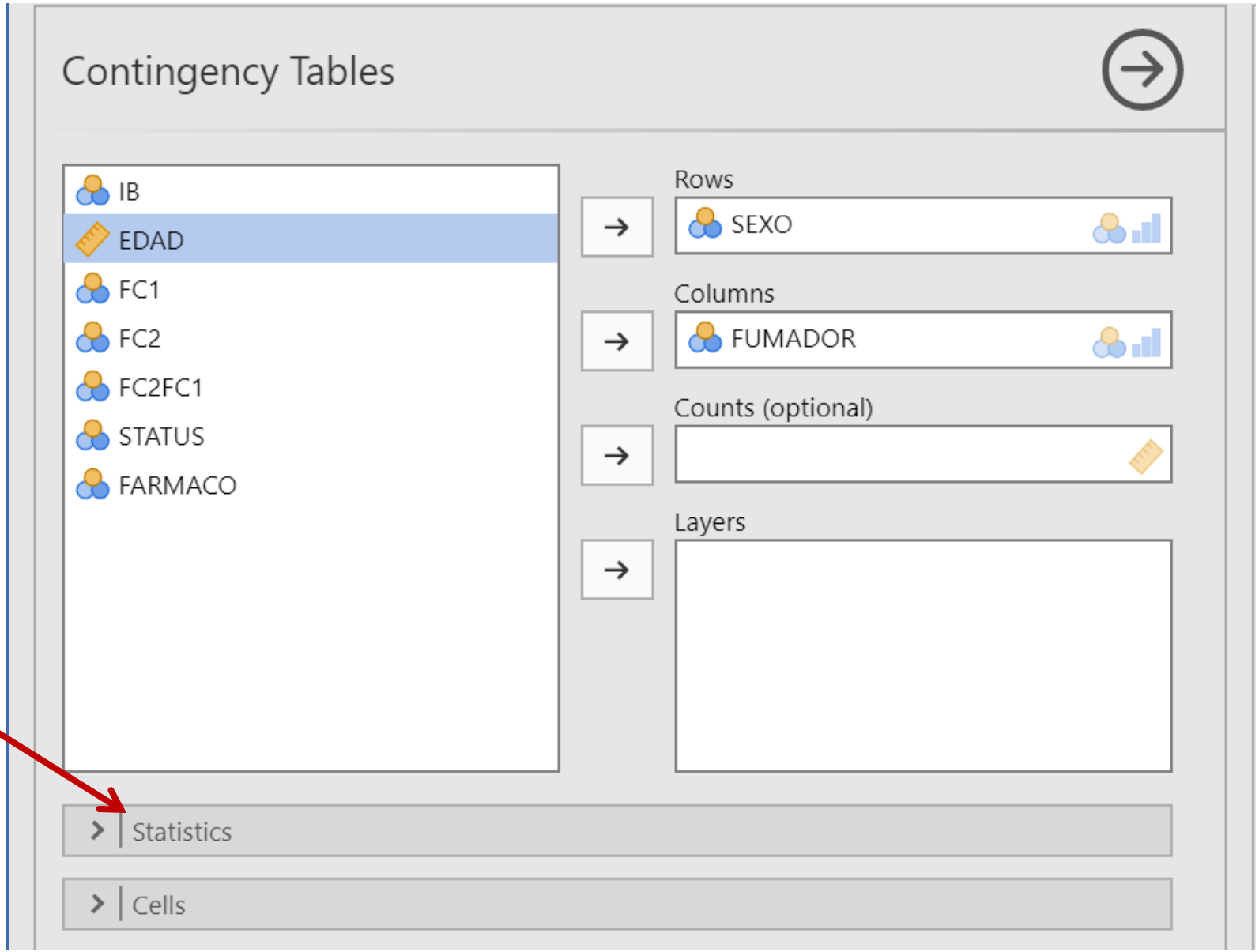
- ¿La proporción de fumadores es igual en hombres que en mujeres?

- Dos variables categóricas:
  - Fumadores (‘Si’ y ‘No’)
  - Sexo (Hombre y Mujer)

Cálculo del P-valor



Cálculo del P-valor



OR, RR

Cálculo del P-valor

Contingency Tables

Contingency Tables

	FUMADOR		
SEXO	1	2	Total
1	10	12	22
2	6	12	18
Total	16	24	40

χ^2 Tests

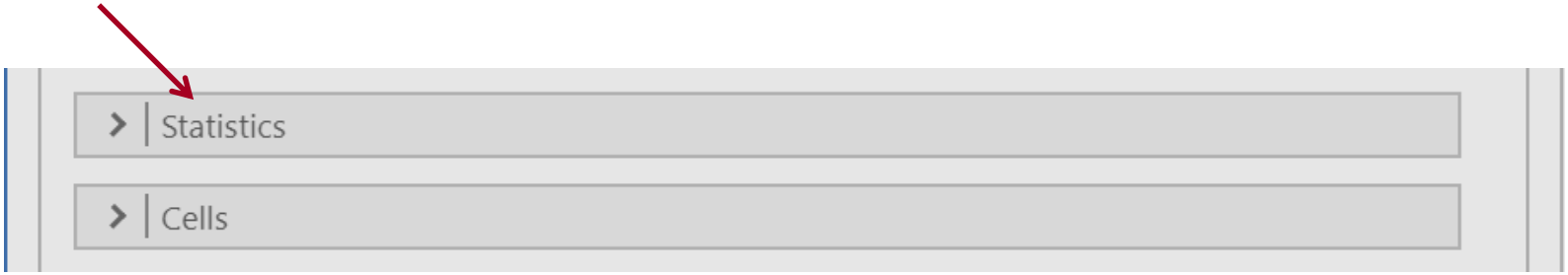
	Value	df	p
χ^2	0.606	1	0.436
N	40		

Test exacto de Fisher

- Cuando en alguna casilla la frecuencia esperada es <5, se usa este test y no el de χ^2

- Con Jamovi:

  Statistics -> Fisher's exact test



Cálculo del P-valor (Jamovi)

Contingency Tables

Contingency Tables

	FUMADOR		
SEXO	1	2	Total
1	10	12	22
2	6	12	18
Total	16	24	40

χ^2 Tests

	Value	df	p
χ^2	0.606	1	0.436
Fisher's exact test	1.65		0.526
N	40		

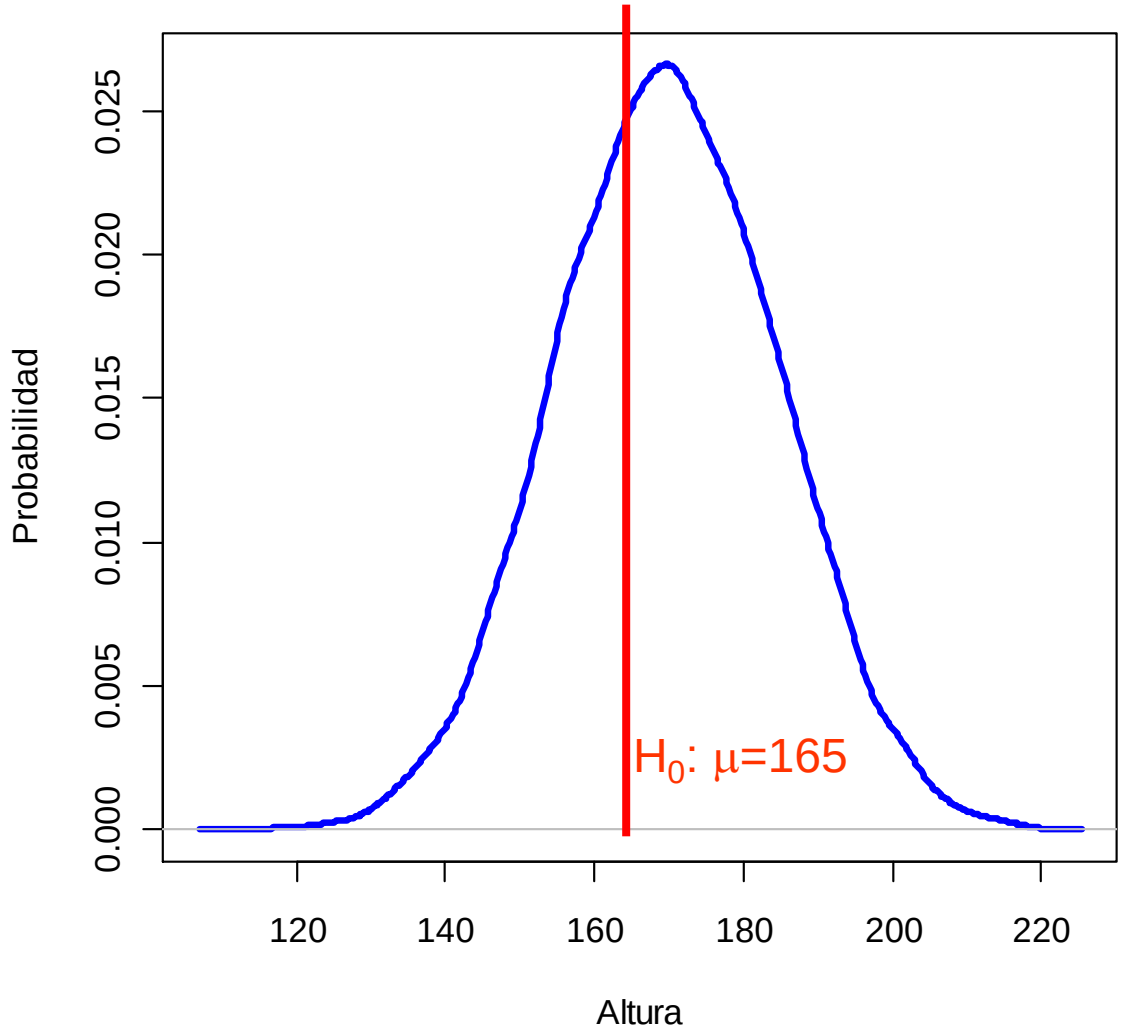
Tests para datos continuos (t-Student)

Objetivos: Preguntas

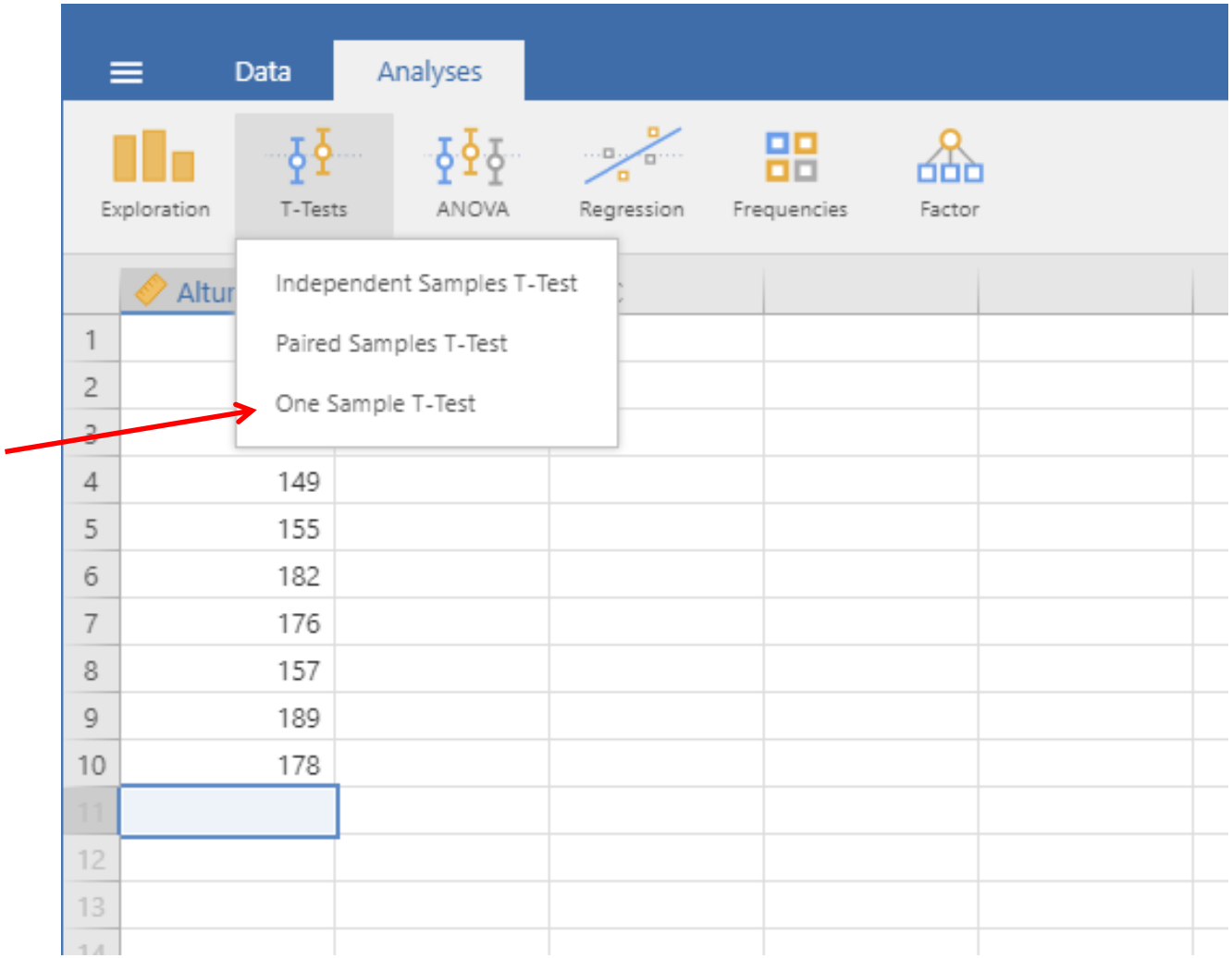
- Pregunta científica:
  - 1 muestra
  - ¿La altura de la población española es 1,65?
  - ¿El nivel de colesterol en España es 200?

  - 2 muestras
  - ¿La altura en España es igual que en Italia?
  - ¿El nivel de colesterol es igual en los individuos que hacen deporte que en los sedentarios?

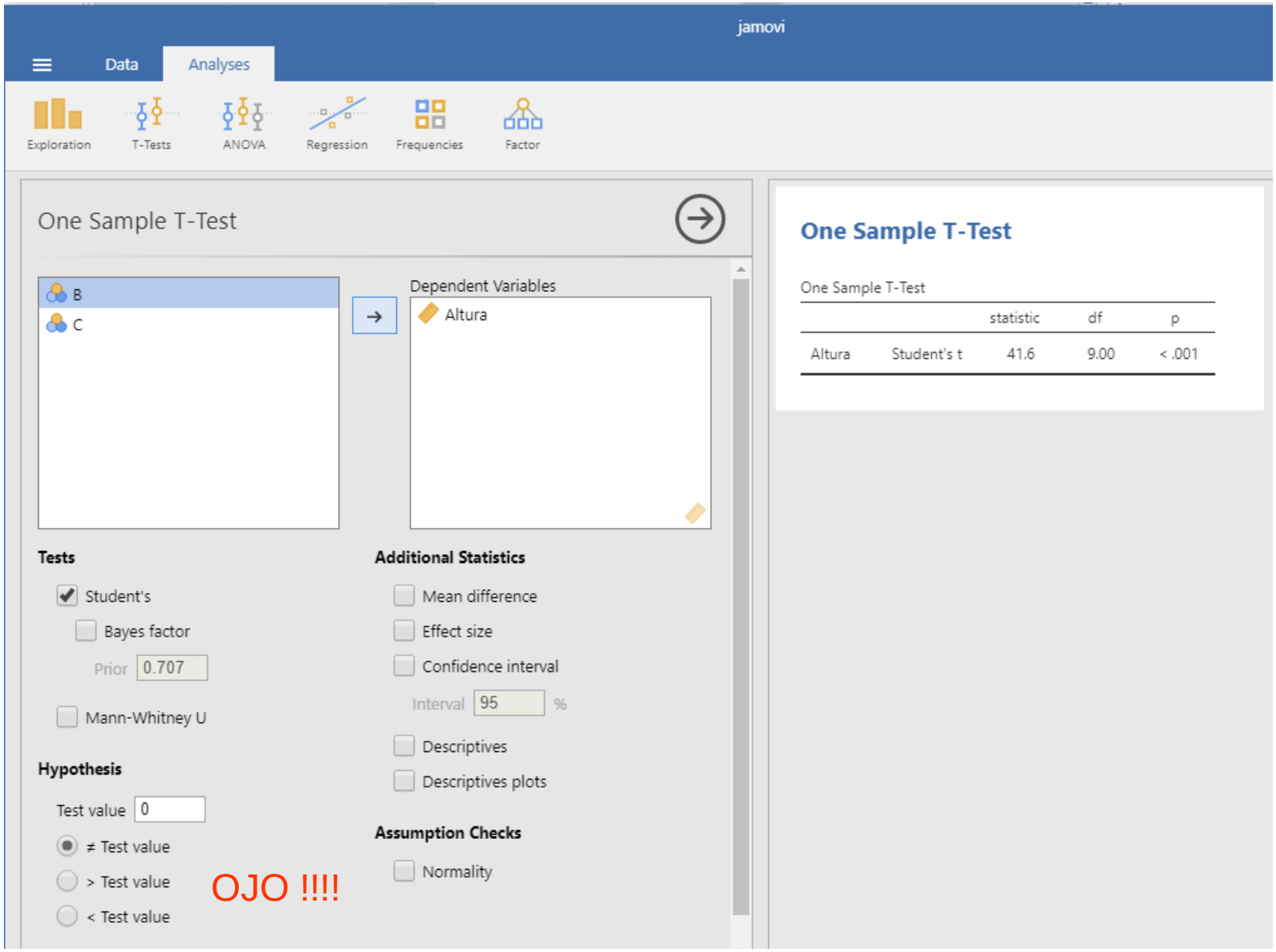
Pregunta: ¿La altura media es 165 cm?



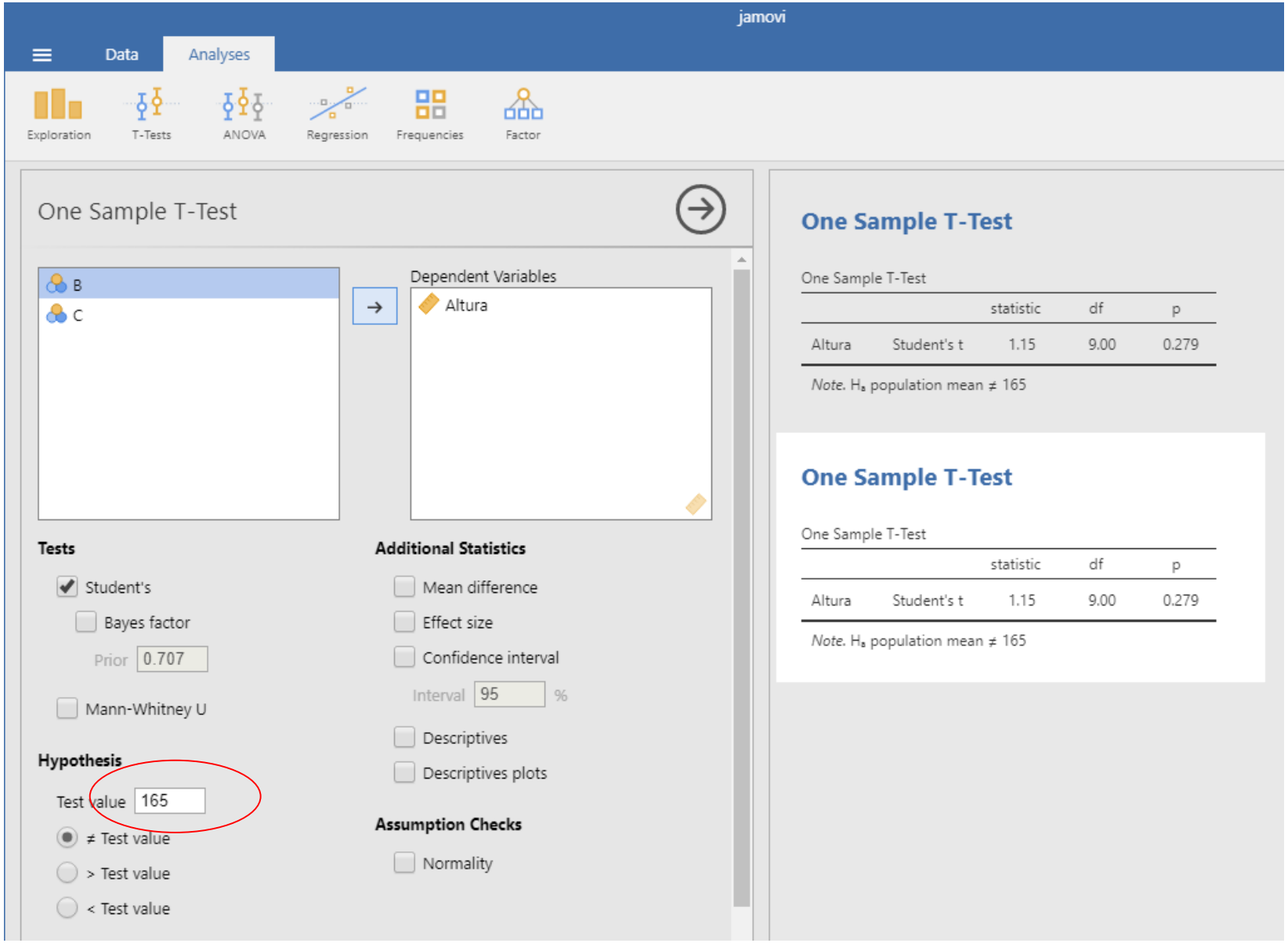
¿Altura = 165?


Data
Analyses
Exploration
T-Tests
ANOVA
Regression
Frequencies
Factor
Independent Samples T-Test
Paired Samples T-Test
One Sample T-Test
149
155
182
176
157
189
178

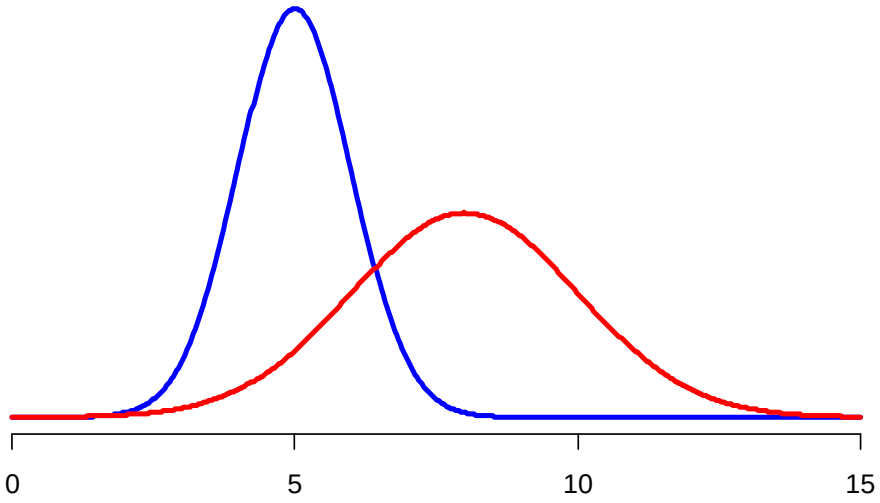
¿Altura = 165?


jamovi
Data
Analyses
Exploration
T-Tests
ANOVA
Regression
Frequencies
Factor
One Sample T-Test
B
C
Dependent Variables
Altura
Tests
Student's
Bayes factor
Prior 0.707
Mann-Whitney U
Hypothesis
Test value 0
≠ Test value
> Test value
< Test value
OJO !!!!
Additional Statistics
Mean difference
Effect size
Confidence interval
Interval 95 %
Descriptives
Descriptives plots
Assumption Checks
Normality
One Sample T-Test
One Sample T-Test
statistic df p
Altura Student's t 41.6 9.00 < .001

¿Altura = 165?


jamovi
Data
Analyses
Exploration
T-Tests
ANOVA
Regression
Frequencies
Factor
One Sample T-Test
B
C
Dependent Variables
Altura
Tests
Student's
Bayes factor
Prior 0.707
Mann-Whitney U
Hypothesis
Test value 165
≠ Test value
> Test value
< Test value
Additional Statistics
Mean difference
Effect size
Confidence interval
Interval 95 %
Descriptives
Descriptives plots
Assumption Checks
Normality
One Sample T-Test
One Sample T-Test
statistic df p
Altura Student's t 1.15 9.00 0.279
Note. Hₐ population mean ≠ 165
One Sample T-Test
One Sample T-Test
statistic df p
Altura Student's t 1.15 9.00 0.279
Note. Hₐ population mean ≠ 165

Comparación de grupos



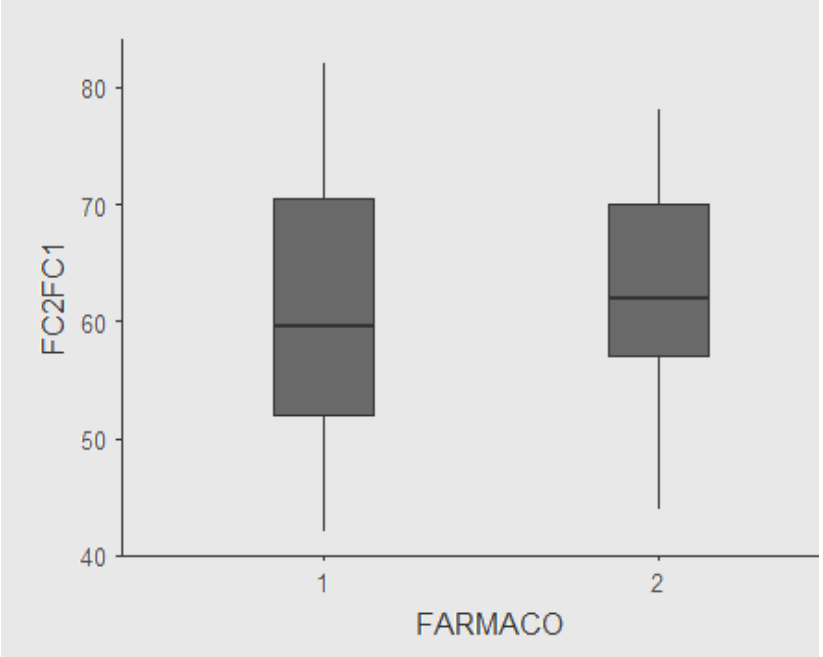
Comparación de grupos independientes

- ¿La diferencia de FC antes y despues se asocia al consumo de cierto fármaco?

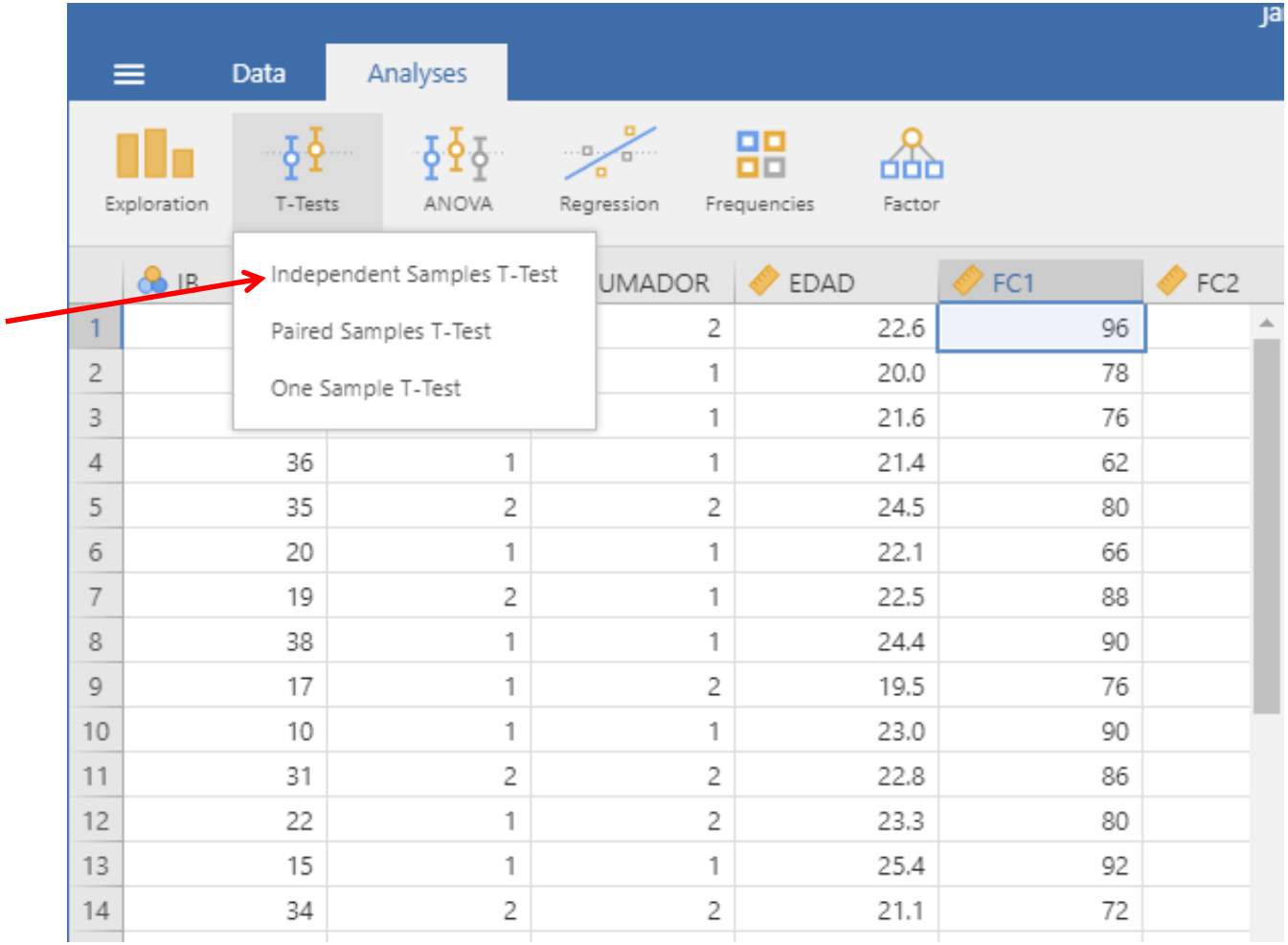
Descriptives

Descriptives

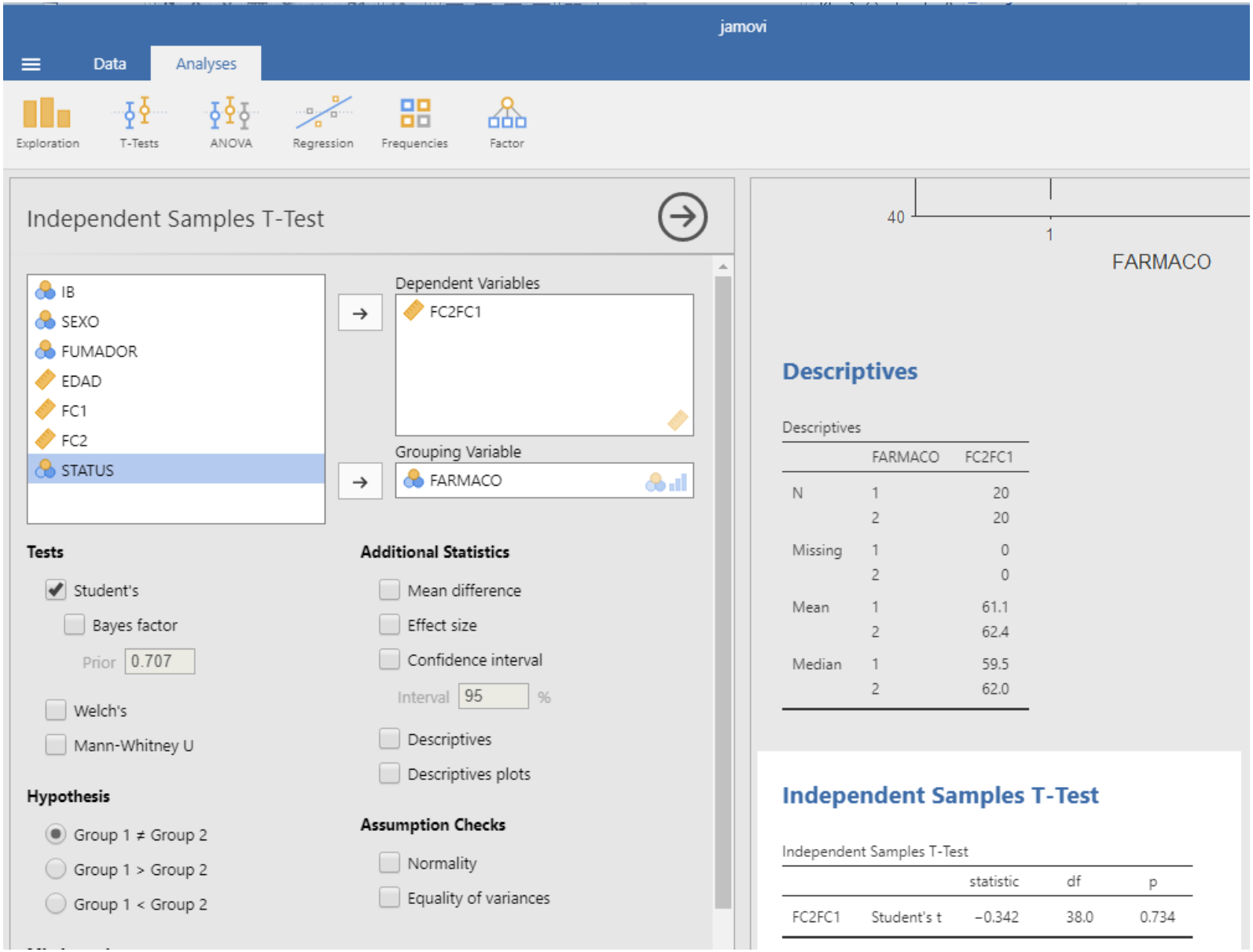
	FARMACO	FC2FC1
N	1	20
	2	20
Missing	1	0
	2	0
Mean	1	61.1
	2	62.4
Median	1	59.5
	2	62.0



Datos independientes



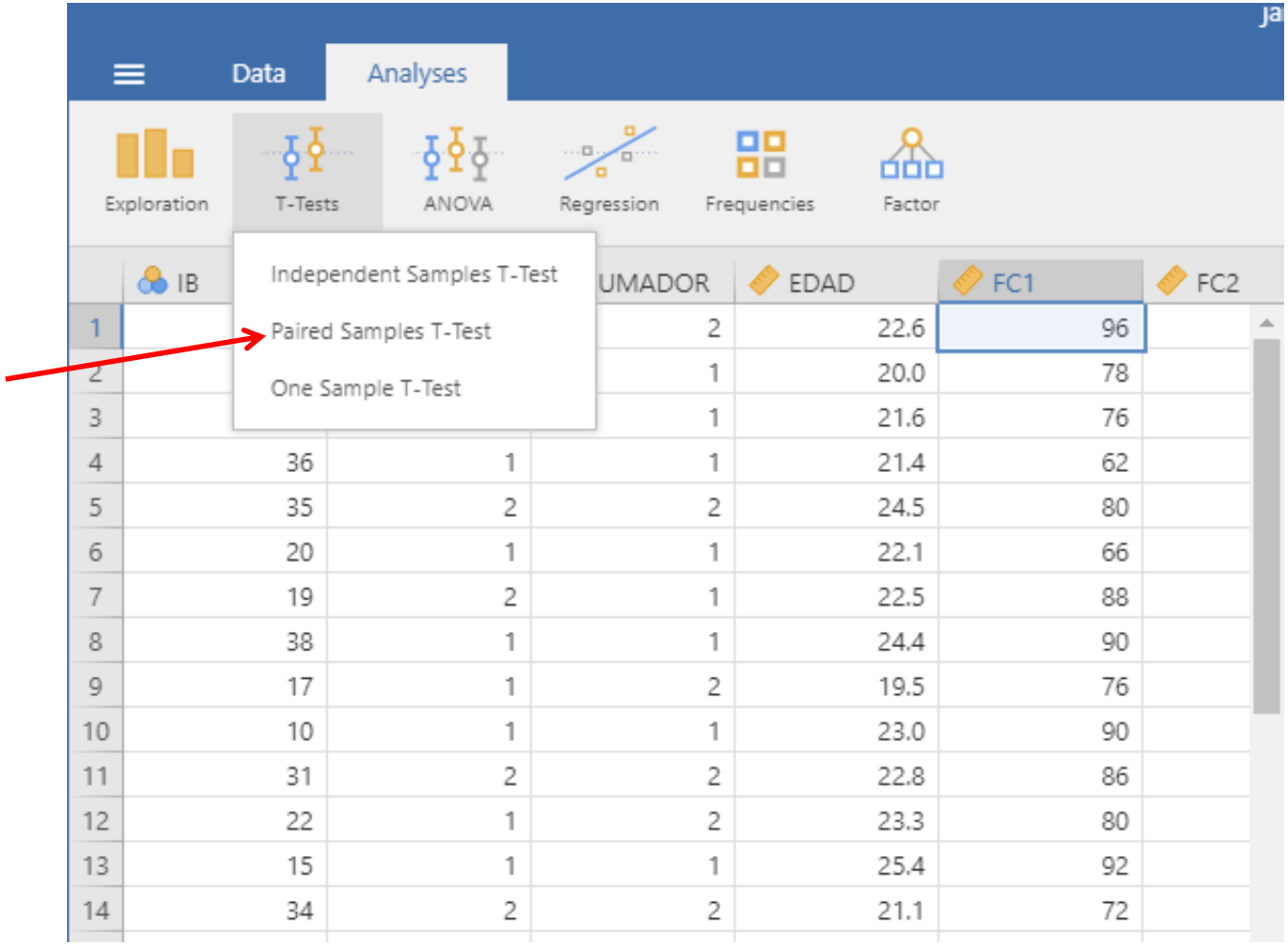
Datos independientes



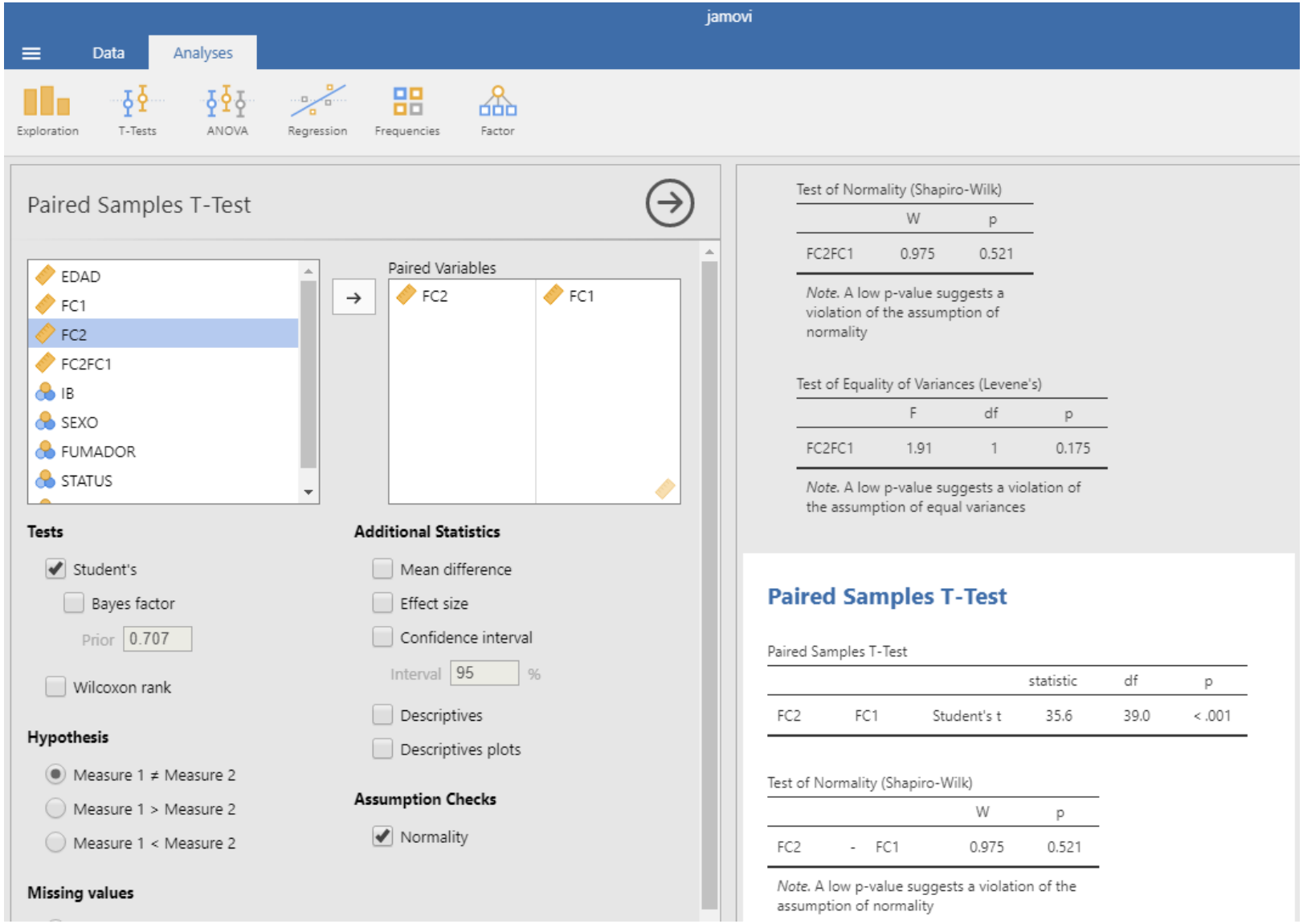
Datos apareados: FC2 vs FC1

- En el estudio hemos recogido sobre los mismos individuos la FC antes y después de tomar el fármaco (o hacer ejercicio, o tomar un recuperador, …)
- Estamos ante un diseño apareado
- Se calculan las diferencias entre mediciones y se compara la media de la diferencia con la diferencia teórica esperada (generalmente $\delta=0$)
- Desde el punto de vista de análisis estadístico es idéntico el análisis de 1 población

Datos apareados



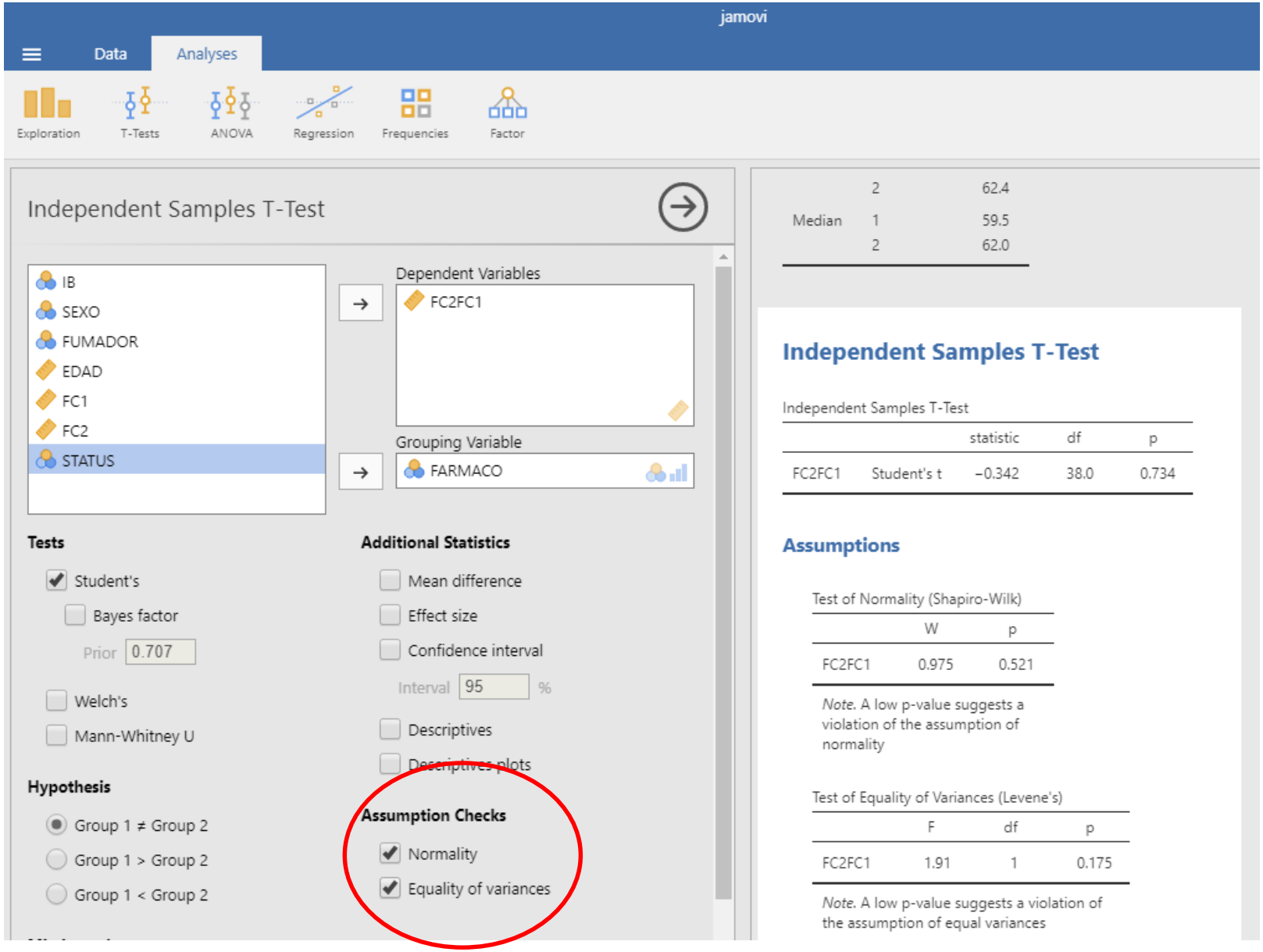
Datos apareados



Condiciones

- Normalidad de la variable numérica
  - Si el tamaño de muestra no es muy pequeño, esta condición no es muy importante
  - Se pueden transformar los datos
- Homogeneidad de varianzas (homocedasticidad)
  - Esta condición sí afecta cuando los tamaños de grupo son diferentes
  - No es muy importante con grupos del mismo tamaño

Asunciones


jamovi
Data
Analyses
Exploration
T-Tests
ANOVA
Regression
Frequencies
Factor
Independent Samples T-Test
IB
SEXO
FUMADOR
EDAD
FC1
FC2
STATUS
Dependent Variables
FC2FC1
Grouping Variable
FARMACO
Tests
Student's
Bayes factor
Prior 0.707
Welch's
Mann-Whitney U
Hypothesis
Group 1 ≠ Group 2
Group 1 > Group 2
Group 1 < Group 2
Additional Statistics
Mean difference
Effect size
Confidence interval
Interval 95 %
Descriptives
Descriptives plots
Assumption Checks
Normality
Equality of variances
2 62.4
Median 1 59.5
2 62.0
Independent Samples T-Test
Independent Samples T-Test
statistic df p
FC2FC1 Student's t −0.342 38.0 0.734
Assumptions
Test of Normality (Shapiro-Wilk)
W p
FC2FC1 0.975 0.521
Note. A low p-value suggests a violation of the assumption of normality
Test of Equality of Variances (Levene's)
F df p
FC2FC1 1.91 1 0.175
Note. A low p-value suggests a violation of the assumption of equal variances

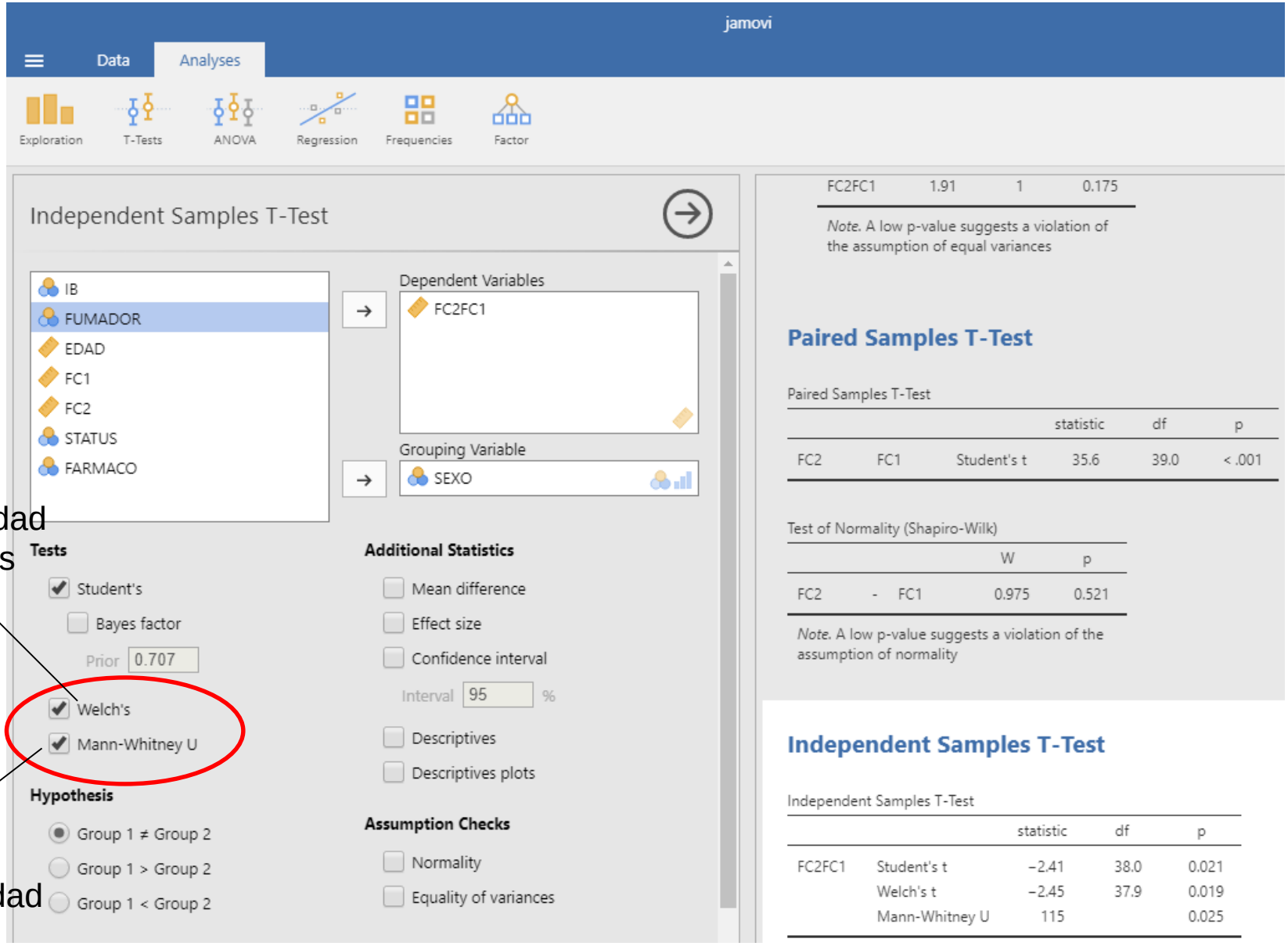
No se cumplen las condiciones

- ¿Qué podemos hacer cuando nuestra variable de interés no cumple las condiciones de la t-Student?
  - No normalidad
  - No homocedasticidad

Transformaciones (logaritmo) / Welch's

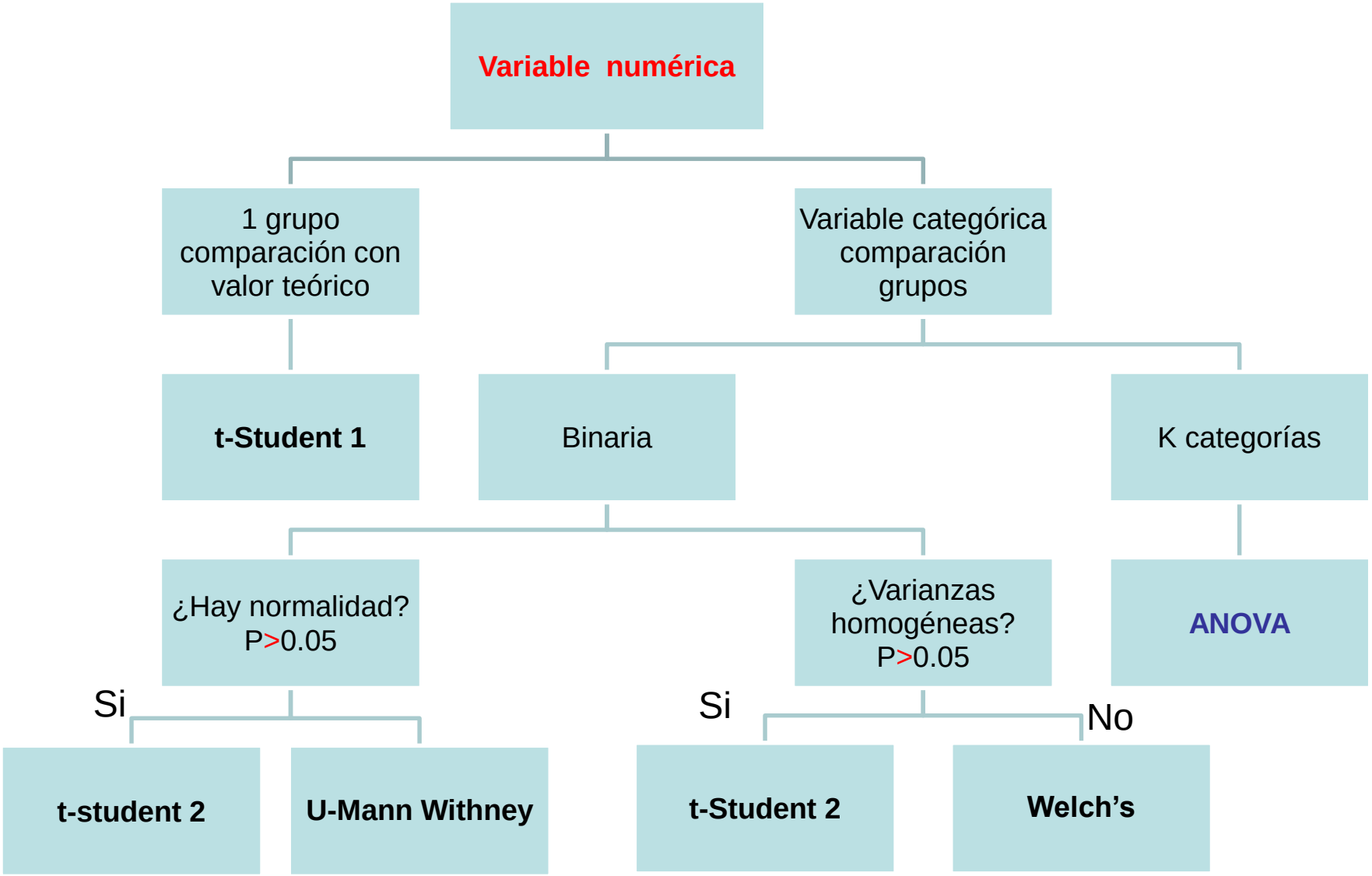
Métodos no paramétricos

Test con Jamovi



No homogeneidad de varianzas

No normalidad


Variable numérica
1 grupo comparación con valor teórico
Variable categórica comparación grupos
t-Student 1
Binaria
K categorías
¿Hay normalidad? P>0.05
¿Varianzas homogéneas? P>0.05
ANOVA
Si
Si
No
t-student 2
U-Mann Withney
t-Student 2
Welch's