

# **APUNTES DE USO BÁSICO DE SPSS**

**Antonio Cervero, Ellián Tuero, Ana Bernardo y María Esteban**

## Índice de contenido

1.- Creación de variables e introducción de datos.....	5
1.1.- Creación de variables.....	5
1.2.- Introducción de datos.....	7
2.- Recodificación de variables y cálculo de variables derivadas.....	7
2.1.- Recodificación de variables.....	7
2.2.- Cálculo de variables derivadas.....	9
3.- Análisis descriptivo. Variabilidad de las respuestas.....	11
4.- Análisis factorial.....	15
5.- Fiabilidad.....	23
6.- Comparación de grupos: T de student y ANOVA.....	25
6.1.- T de Student.....	25
6.2.- ANOVA.....	28
7.- Correlación de variables.....	34
8. Análisis de regresión.....	35

## Índice de figuras

Fig. 1. Pestaña de variables .....	5
Fig. 2. Definición de variables.....	5
Fig. 3. Definición de valores .....	6
Fig. 4. Medida de las variables .....	6
Fig. 5. Introducción de datos.....	7
Fig. 6. Recodificar variables.....	8
Fig. 7. Asignación de nuevos valores.....	8
Fig. 8. Reasignación de valores.....	9
Fig. 9. Reasignación de la etiqueta al nuevo valor .....	9
Fig. 10. Finalización de la nueva variable .....	9
Fig. 11. Cálculo de nuevas variables.....	10
Fig. 12. Recalcular la variable .....	10
Fig. 13. Inclusión de la nueva variable .....	11
Fig. 14. Análisis de estadísticos descriptivos: frecuencias y porcentajes.....	11
Fig. 15. Selección de variables para análisis descriptivos .....	12
Fig. 16. Selección de análisis descriptivos.....	12
Fig. 17. Tablas de resultados de análisis descriptivos .....	13
Fig. 18. Amplitud intercuartil .....	14
Fig. 19. Análisis factorial.....	15
Fig. 20. Selección de variables para el análisis factorial.....	15
Fig. 21. Descriptivos del análisis factorial.....	16
Fig. 22. Resultados del análisis factorial.....	16
Fig. 23. Análisis factorial 2.....	17
Fig. 24. Extracción de factores .....	17
Fig. 25. Método de extracción de factores .....	18
Fig. 26. Otras opciones de visualización de factores.....	18
Fig. 27. Resultados: número de factores.....	19
Fig. 28. Análisis factorial 3.....	19
Fig. 29. Rotación de factores.....	20
Fig. 30. Métodos de rotación de factores .....	20
Fig. 31. Matriz de componentes rotados .....	21
Fig. 32. Opciones de análisis de factores .....	21
Fig. 33. Presentación de coeficientes en análisis factorial.....	22

Fig. 34. Matriz de componente rotado .....	22
Fig. 35. Análisis de fiabilidad .....	23
Fig. 36. Elementos del análisis de fiabilidad.....	23
Fig. 37. Descriptivos para el análisis de fiabilidad.....	24
Fig. 38. Resultado del análisis de fiabilidad.....	24
Fig. 39. Estadísticas del total del elemento del análisis de fiabilidad .....	24
Fig. 40. Pruebas t.....	25
Fig. 41. Selección de variables de las pruebas t .....	25
Fig. 42. Definición de variables en pruebas t .....	26
Fig. 43. Variable de agrupación en pruebas t.....	26
Fig. 44. Prueba de Levene .....	27
Fig. 45. Ejemplo de prueba t .....	27
Fig. 46. ANOVA .....	28
Fig. 47. Selección de variables en ANOVA.....	29
Fig. 48. Diferencias en ANOVA .....	29
Fig. 49. Resultados ANOVA.....	30
Fig. 50. Pestaña post hoc de ANOVA.....	30
Fig. 51. Asunción de varianzas en ANOVA .....	31
Fig. 52. Resultados ANOVA.....	31
Fig. 53. Opciones en análisis ANOVA .....	32
Fig. 54. Descriptivos en análisis ANOVA.....	33
Fig. 55. Tablas de resultados ANOVA .....	33
Fig. 56. Correlaciones .....	34
Fig. 57. Correlaciones bivariadas.....	34
Fig. 58. Resultados de las correlaciones.....	35
Fig. 59. Regresión lineal.....	35
Fig. 60. Selección de variables de regresión.....	36
Fig. 61. Cuadro ANOVA de regresión .....	36
Fig. 62. Valor $R^2$ ajustado del análisis de regresión.....	37
Fig. 63. Tabla de coeficientes .....	37
Fig. 64. Método por pasos para el análisis de regresión.....	38
Fig. 65. $R^2$ ajustado del método por pasos.....	39
Fig. 66. Tabla de coeficientes del modelo por pasos .....	39
Fig. 67. Valor F del modelo por método de pasos .....	40

## 1.- Creación de variables e introducción de datos

### 1.1.- Creación de variables

Una vez que abrimos un fichero nuevo en la base de datos lo primero que tenemos que hacer es crear las variables en la pestaña de variables (círculo rojo) [Fig. 1].

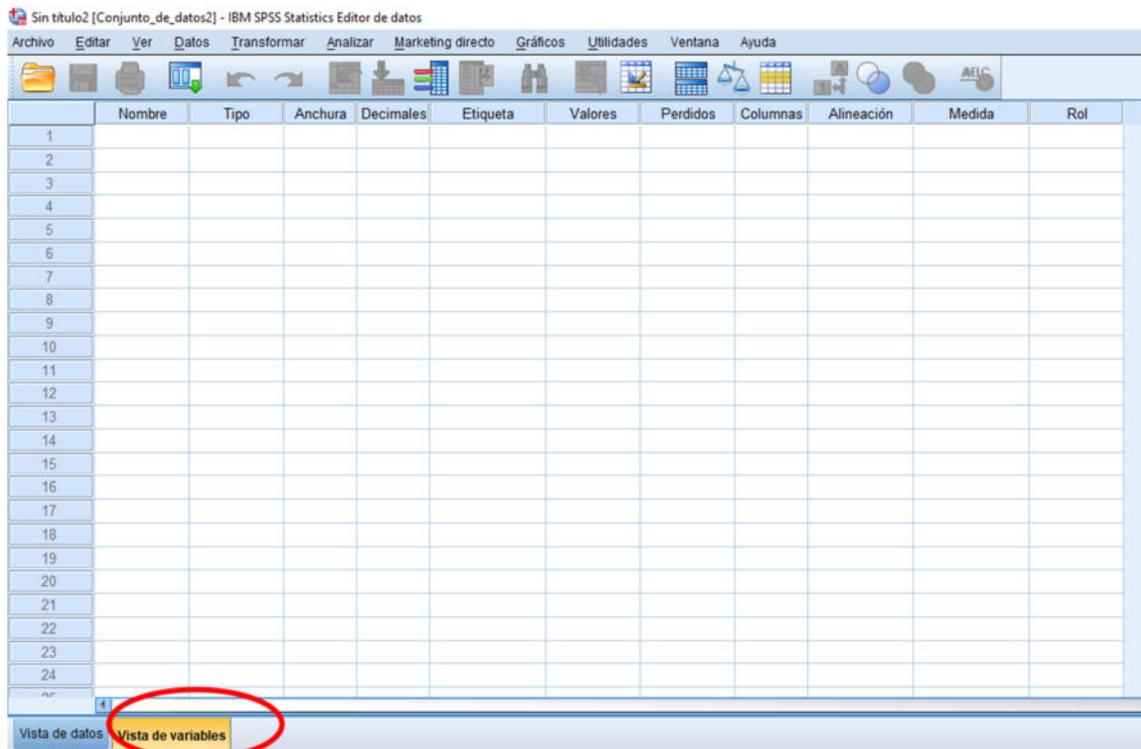


Fig. 1. Pestaña de variables

En esta ventana definimos cada una de las variables incluyendo los datos que nos solicita, como vemos en la imagen [Fig. 2].

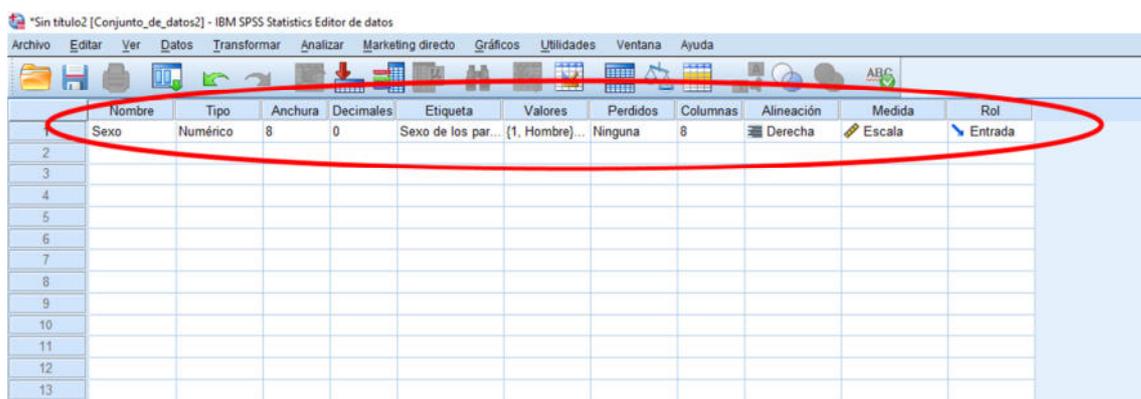


Fig. 2. Definición de variables

Los principales datos son:

- Nombre: Nombre a través del cual se identificará la variable.
- Tipo: Tipo de variable (en general numérico, o cadena si es de tipo cualitativo, aunque existen otras opciones como formato fecha, etc.).
- Anchura y decimales: Número máximo de caracteres que aceptará la variable y de decimales que contemplará.
- Etiqueta: Denominación de la variable de forma exhaustiva.
- Valores: Valores que asignamos a la variable. Así, en el caso del sexo por ejemplo, podemos asignar el valor 1 a los hombres y el 2 a las mujeres como vemos en la figura (círculo azul), añadiendo cada uno de los valores (círculo rojo) [Fig. 3].

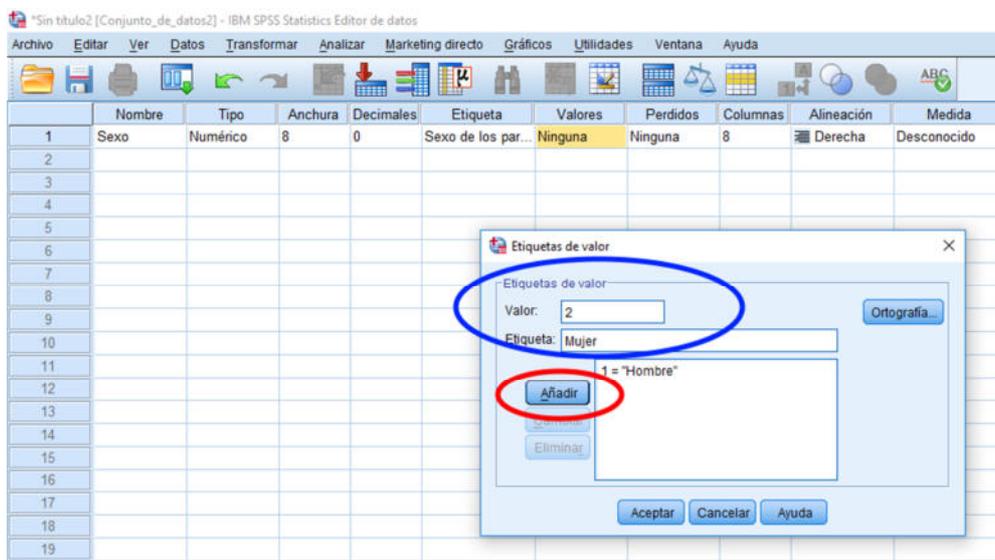


Fig. 3. Definición de valores

- Perdidos: Valores con que codificamos la falta de respuesta (por ejemplo ns/nc).
- Medida: Tipo de medida de la variable, que puede ser nominal, ordinal o escala [Fig. 4].

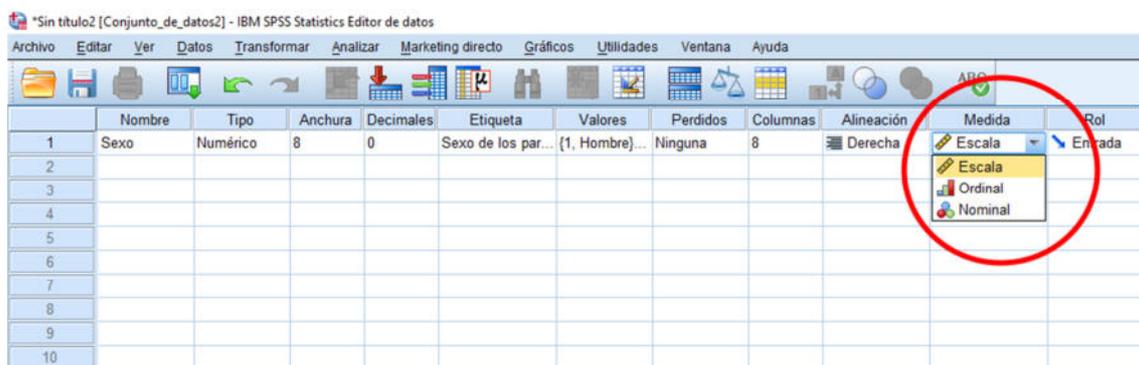


Fig. 4. Medida de las variables

## 1.2.- Introducción de datos

Tras definir todo el conjunto de variables pasamos a la pestaña de datos (círculo rojo), donde ya tendremos las variables definidas, e introducimos cada uno de los datos (círculo azul). [Fig. 5]

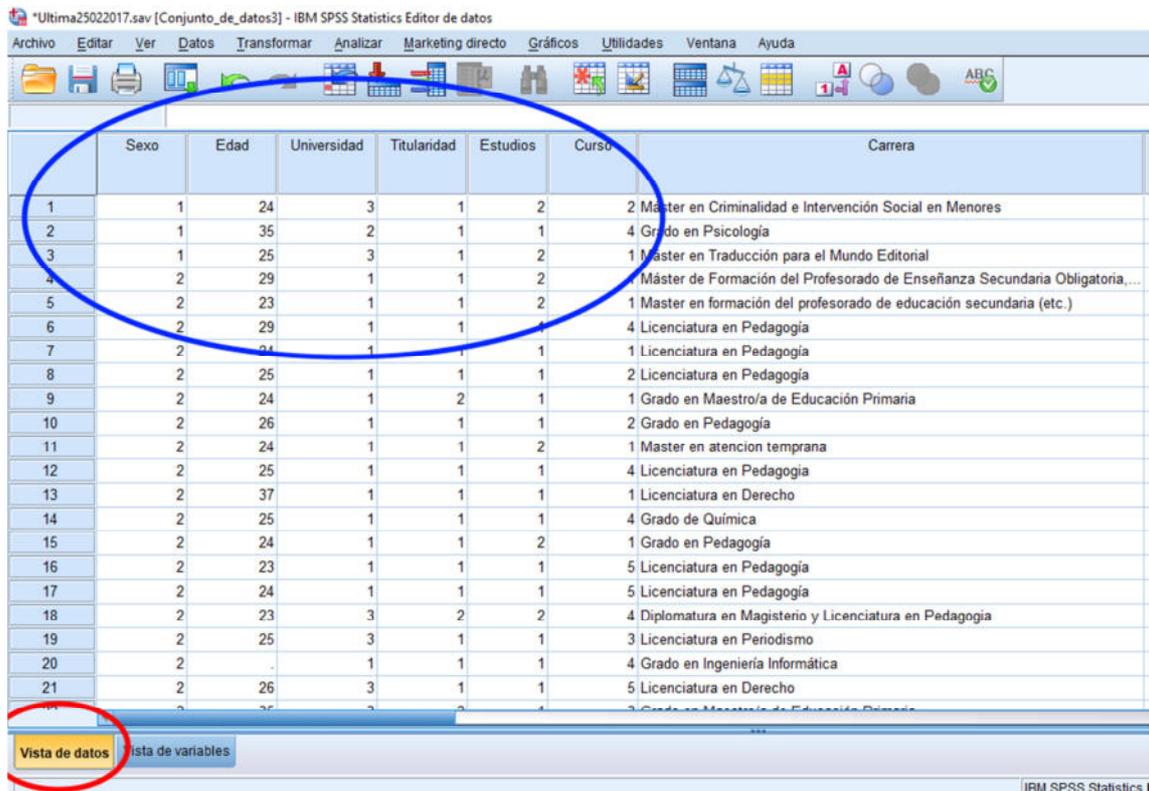


Fig. 5. Introducción de datos

## 2.- Recodificación de variables y cálculo de variables derivadas

### 2.1.- Recodificación de variables

Una vez que tenemos definidas las variables podemos transformarlas o crear nuevas variables calculadas a partir de las que ya tenemos.

Para recodificar variables, es decir, cambiar los valores de una variable o agruparlos, debemos seguir la secuencia: transformar-recodificar (en las mismas variables o en distintas variables, según nos interese mantener la columna de datos originales o sustituirla por una nueva) [Fig. 6].

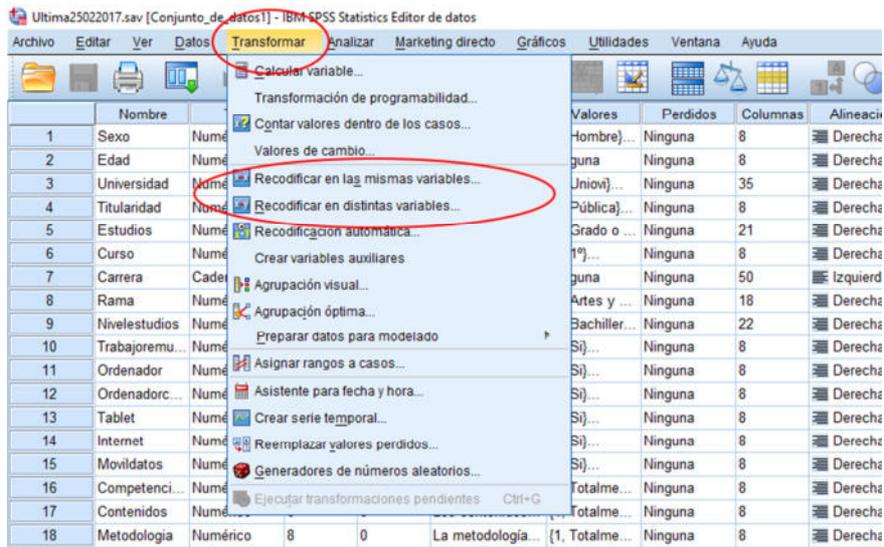


Fig. 6. Recodificar variables

Una vez en la ventana de recodificar variable seleccionamos la variable a recodificar (círculo azul) y seleccionamos valores antiguos y nuevos (círculo rojo) para asignar los nuevos valores [Fig. 7].

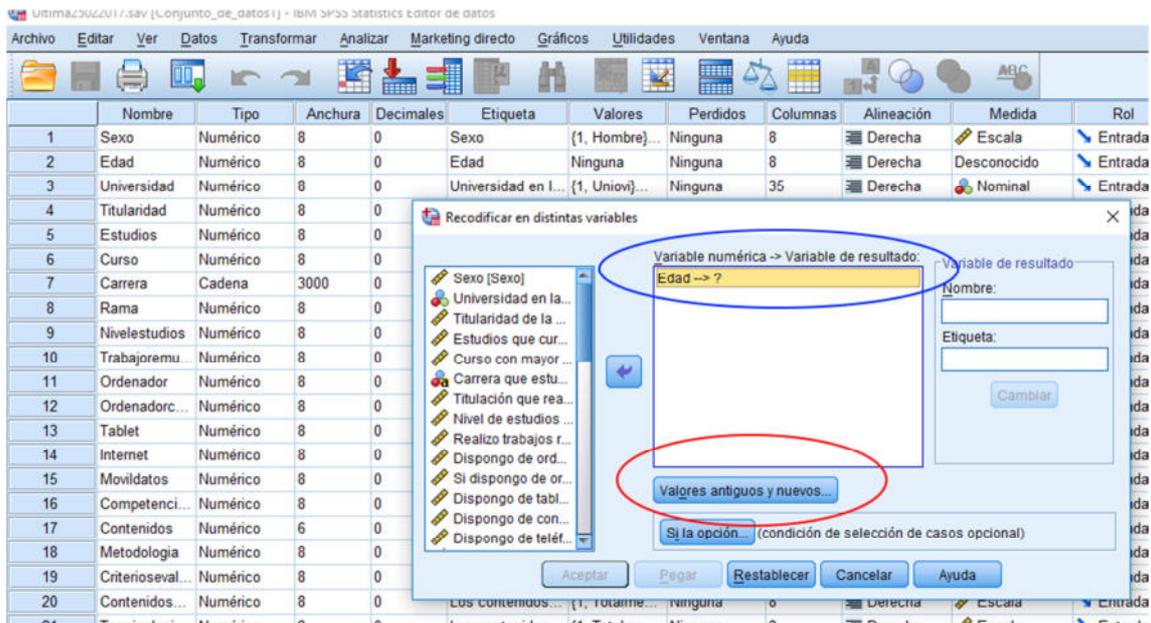


Fig. 7. Asignación de nuevos valores

Cambiamos los valores antiguos por los nuevos, pudiendo hacer esto en función de cada valor concreto (círculo azul) o bien por rangos (círculo rojo), y los añadimos (círculo verde), dándole a continuar [Fig. 8].

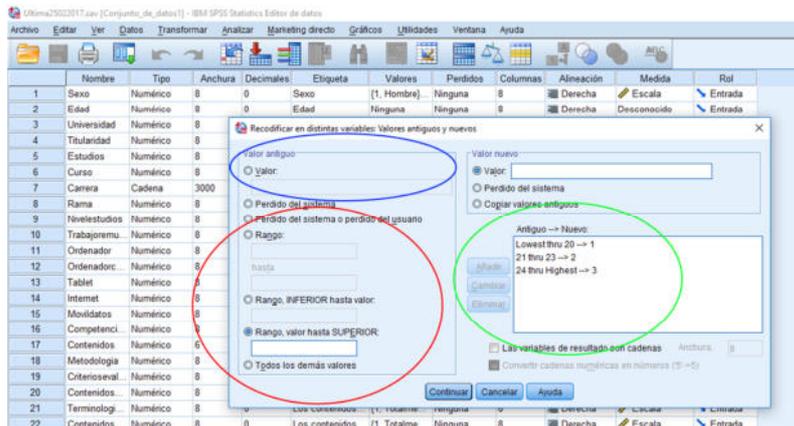


Fig. 8. Reasignación de valores

De nuevo en el cuadro principal, damos un nombre y etiqueta a la nueva variable, pinchando en cambiar (círculo azul), y vemos que el cambio se realizará cuando en el cuadro de variables nos señala la equivalencia del cambio (círculo rojo) [Fig. 9].

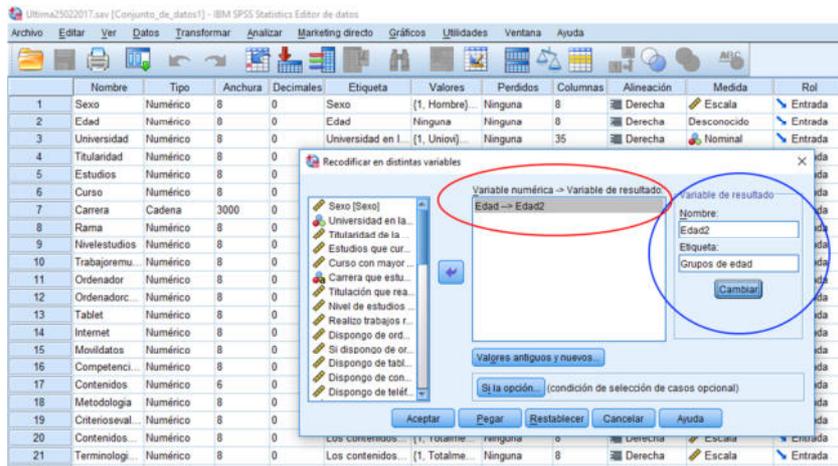


Fig. 9. Reasignación de la etiqueta al nuevo valor

Al aceptar, si hemos recodificado en la misma variable los nuevos valores de la variables serán los que hemos señalado, y si hemos recodificado en otra variable se creará esta columna en la fila de variables con los datos correspondientes [Fig. 10].

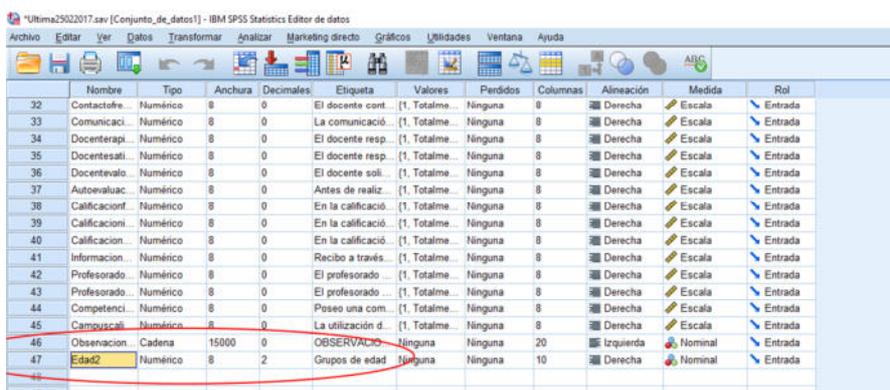


Fig. 10. Finalización de la nueva variable

## 2.2.- Cálculo de variables derivadas

Para calcular nuevas variables a partir de las ya existentes, la secuencia a seguir es: transformar-calcular variable [Fig. 11].

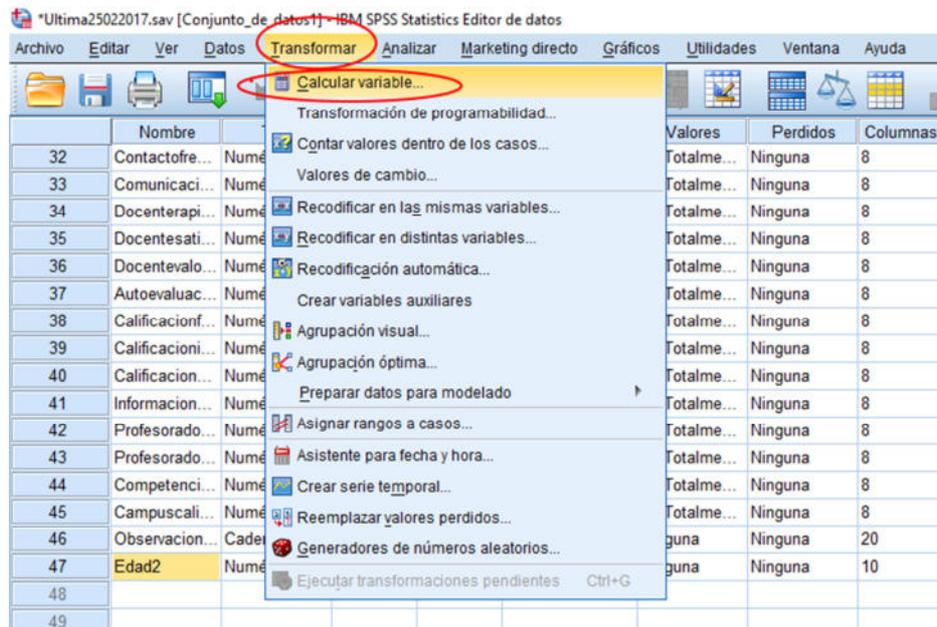


Fig. 11. Cálculo de nuevas variables

En la ventana de cálculo de variables incluimos en primer lugar el nombre de la nueva variable (círculo rojo), pasamos al cuadro de la derecha las variables con las que queremos operar introduciendo los signos matemáticos oportunos (círculo azul) y aceptamos [Fig. 12].

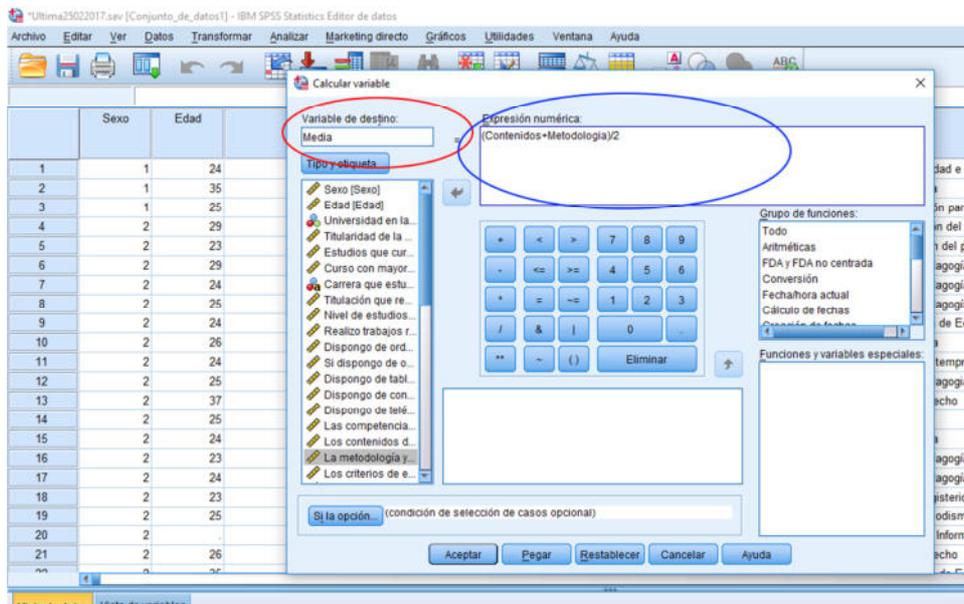


Fig. 12. Recalcular la variable

El resultado será la aparición de una nueva columna con la variable creada [Fig. 13].

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
41	Informacion...	Numérico	8	0	Recibo a través...	{1, Totalme...	Ninguna	8	Derecha	Escala	Entrada
42	Profesorado...	Numérico	8	0	El profesorado ...	{1, Totalme...	Ninguna	8	Derecha	Escala	Entrada
43	Profesorado...	Numérico	8	0	El profesorado ...	{1, Totalme...	Ninguna	8	Derecha	Escala	Entrada
44	Competenci...	Numérico	8	0	Poseo una com...	{1, Totalme...	Ninguna	8	Derecha	Escala	Entrada
45	Campuscali...	Numérico	8	0	La utilización d...	{1, Totalme...	Ninguna	8	Derecha	Escala	Entrada
46	Observacion...	Cadena	15000	0	OBSERVACIO...	Ninguna	Ninguna	20	Izquierda	Nominal	Entrada
47	Edad2	Numérico	8	2	Grupos de edad	Ninguna	Ninguna	10	Derecha	Nominal	Entrada
48	Media	Numérico	8	2		Ninguna	Ninguna	10	Derecha	Escala	Entrada
49											

Fig. 13. Inclusión de la nueva variable

### 3.- Análisis descriptivo. Variabilidad de las respuestas

Dentro de una muestra lo primero que debemos realizar es un análisis descriptivo de los datos. En este caso se suelen incorporar los siguientes valores: número de sujetos, valor mínimo y máximo de la escala (del cual podemos obtener el resultado de su amplitud realizando la resta), media, mediana (si es necesario y especialmente si los datos son ordinales) y desviación típica.

Para ello, seguimos la secuencia: analizar-estadísticos descriptivos-frecuencias [Fig. 14].

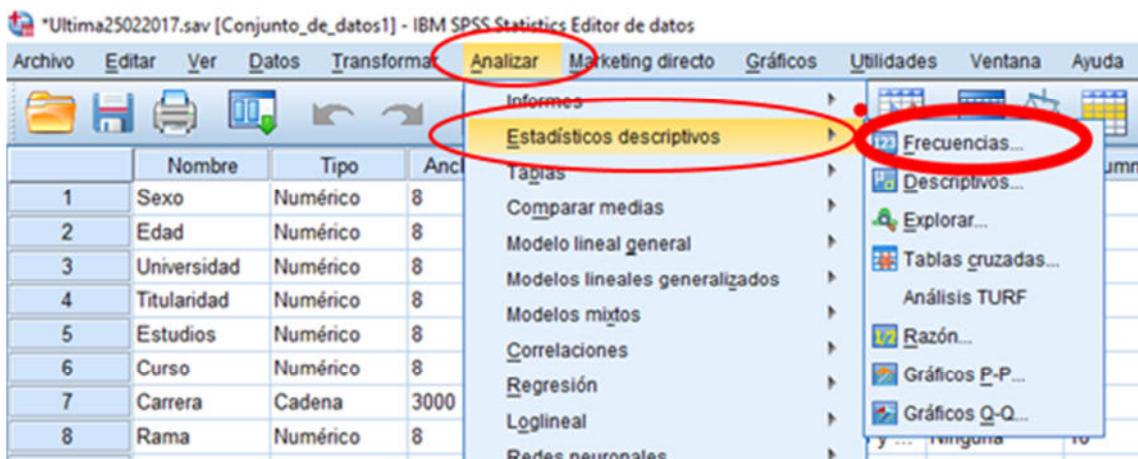


Fig. 14. Análisis de estadísticos descriptivos: frecuencias y porcentajes

A continuación, tenemos una ventana donde pasamos al lado derecho las variables que queremos analizar (círculo azul), y pinchamos el botón estadísticos para seleccionar los análisis a realizar (círculo rojo) [Fig. 15].

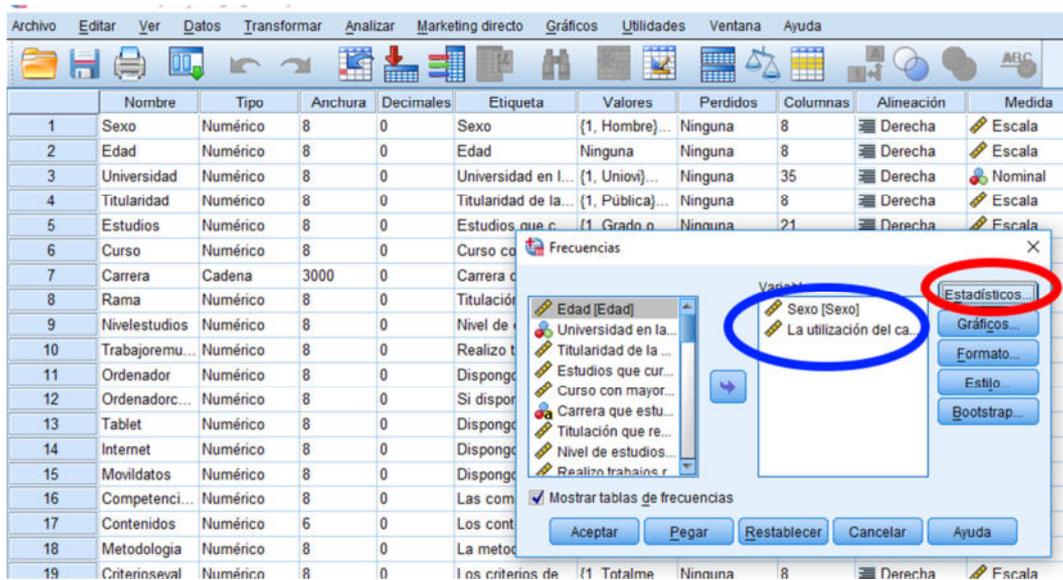


Fig. 15. Selección de variables para análisis descriptivos

En la nueva ventana marcamos los análisis que deseamos realizar, y presionamos en continuar y aceptar [Fig. 16].

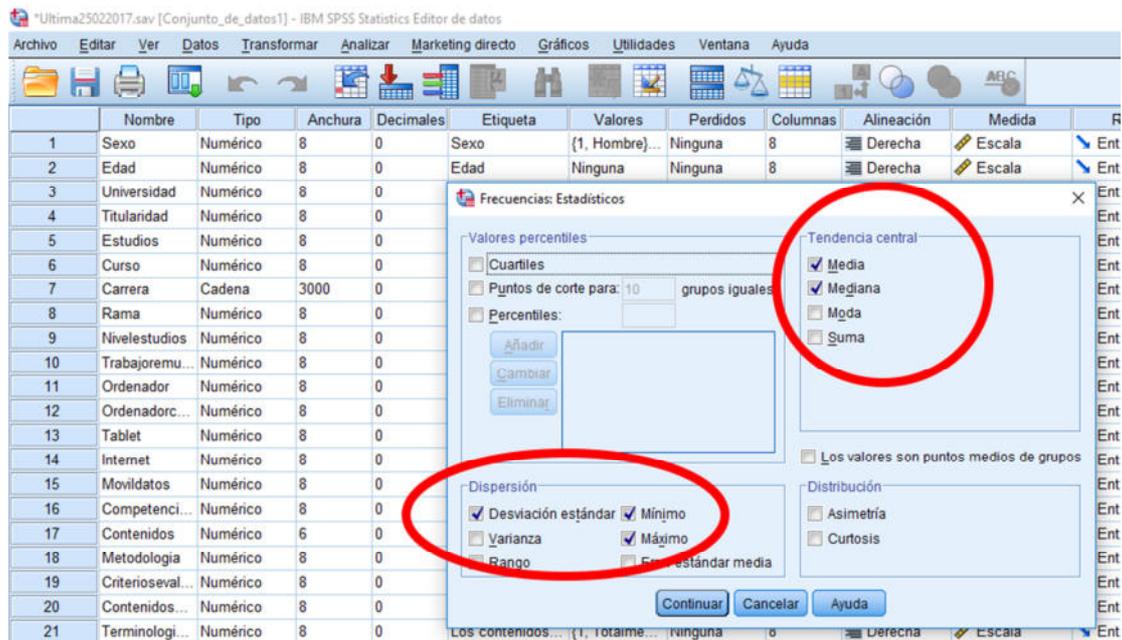


Fig. 16. Selección de análisis descriptivos

La salida de resultados sería como la siguiente [Fig. 17]:

**Estadísticos**

		Sexo		La utilización del campus virtual como recurso de apoyo en la enseñanza-aprendizaje universitaria otorga calidad a la educación recibida
N	Válido	311		311
	Perdidos	11		11
Media		1,70		2,96
Mediana		2,00		3,00
Desviación estándar		,460		,875
Mínimo		1		1
Máximo		2		4

**Tabla de frecuencia**

**Sexo**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Hombre	94	29,2	30,2	30,2
	Mujer	217	67,4	69,8	100,0
	Total	311	96,6	100,0	
Perdidos	Sistema	11	3,4		
Total		322	100,0		

Fig. 17. Tablas de resultados de análisis descriptivos

La representación en un artículo podría ser la que vemos en la Tabla 1.

Tabla 1. Representación de datos descriptivos

ÍTEM	N	Mín-Máx	$\bar{x}$	Med	s
Sexo	311	1-2	1,70	2	,460
Calidad del campus	311	1-4	2,96	3	,875
...	...	...	...	...	...

También podemos calcular la amplitud intercuartil, que es la diferencia entre el valor donde se encuentra el 25% de los encuestados y el 75% de los encuestados. Cuanto mayor es la amplitud mejor, pues mayor es la variabilidad de respuesta, lo que nos interesa al analizar grupos.

Así, en la figura, el 25% de la población se encontraría dentro del valor 3 y el 75% dentro del valor 4. Luego la amplitud intercuartil sería  $4-3= 2$  [Fig. 18]

**La utilización del campus virtual como recurso de apoyo en la enseñanza-aprendizaje universitaria otorga calidad a la educación recibida**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Totalmente en desacuerdo	27	8,4	8,7	8,7
	En desacuerdo	44	13,7	14,1	22,8
	De acuerdo	154	47,8	49,5	72,3
	Totalmente de acuerdo	86	26,7	27,7	100,0
	Total	311	96,6	100,0	
Perdidos	Sistema	11	3,4		
Total		322	100,0		

Fig. 18. Amplitud intercuartil

La representación en la tabla de datos descriptivos podemos hacerla como se representa en la Tabla 2.

Tabla 2. Datos descriptivos con la amplitud intercuartil

ÍTEM	N	Mín-Máx	$\bar{X}$	Med	s	Amplitud intercuartil
Sexo	311	1-2	1,70	2	,460	2-1=1
Calidad del campus	311	1-4	2,96	3	,875	4-3=1
...	...	...	...	...	...	

No es positivo que haya muchos valores extremos, ya que esto alejaría la distribución de la curva normal y de hecho podría implicar que la escala podría ampliarse. Si en una escala Likert de 1 a 4 por ejemplo, y hay muchos valores en 3 y 4, quizás podría ampliarse a una escala de 1 a 6.

Conceptos importantes son la media y la desviación típica porque aportan una información global y es especialmente relevante analizar la variabilidad de los datos, pudiendo ver la desviación típica o la varianza  $s^2$ .

Los valores óptimos de la desviación típica o de la amplitud dependen mucho de la escala, así por ejemplo para una escala tipo Likert de 1-6, un valor adecuado de la desviación típica sería mayor de 1, y para la amplitud intercuartil de 2. En el caso de variables ordinales también puede ser útil analizar los percentiles.

#### 4.- Análisis factorial

El análisis factorial es el análisis que nos permite determinar los puntos de convergencia de diferentes variables dentro de una misma categoría (factor). Por ejemplo, podemos tener en un mismo cuestionario un factor denominado análisis académico en que se incluyan variables relacionadas con la nota media de Bachillerato, la nota de acceso a la Universidad, la nota de primero de carrera, etc. Otro factor relacionado con el entorno social que incluya número de hermanos, relación con los amigos, presión de la familia para el estudio, etc. El análisis factorial nos indica las categorías (factores) de un cuestionario y los ítems pertenecientes a cada uno. **Para realizar un análisis factorial necesitamos una muestra mínima de 250 sujetos.**

Para ello, seguimos la secuencia: analizar-reducción de dimensiones-factor [Fig. 19].

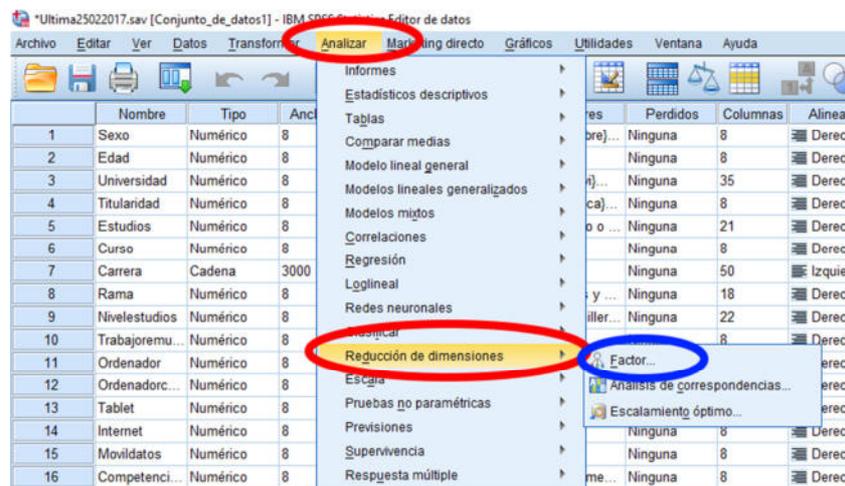


Fig. 19. Análisis factorial

En la siguiente ventana, pasamos las variables del cuestionario que queremos ver si podemos agrupar en factores [Fig. 20].

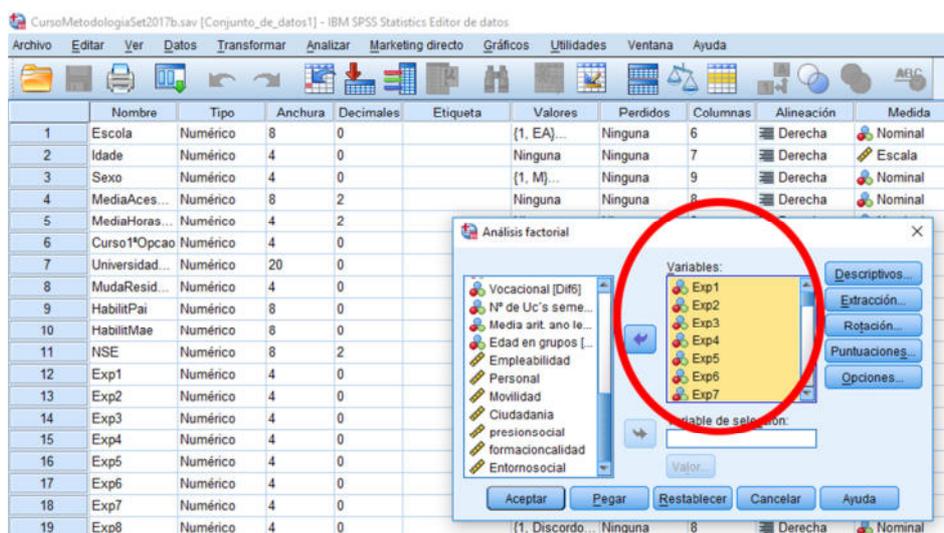


Fig. 20. Selección de variables para el análisis factorial

A continuación tenemos que seleccionar la pestaña de descriptivos y marcar dos valores importantes. En la sección de estadísticos, marcamos la opción solución inicial y en la opción de matriz de correlaciones marcamos la opción, KMO y prueba de esfericidad de Bartlett. Pinchamos en continuar [Fig. 21].

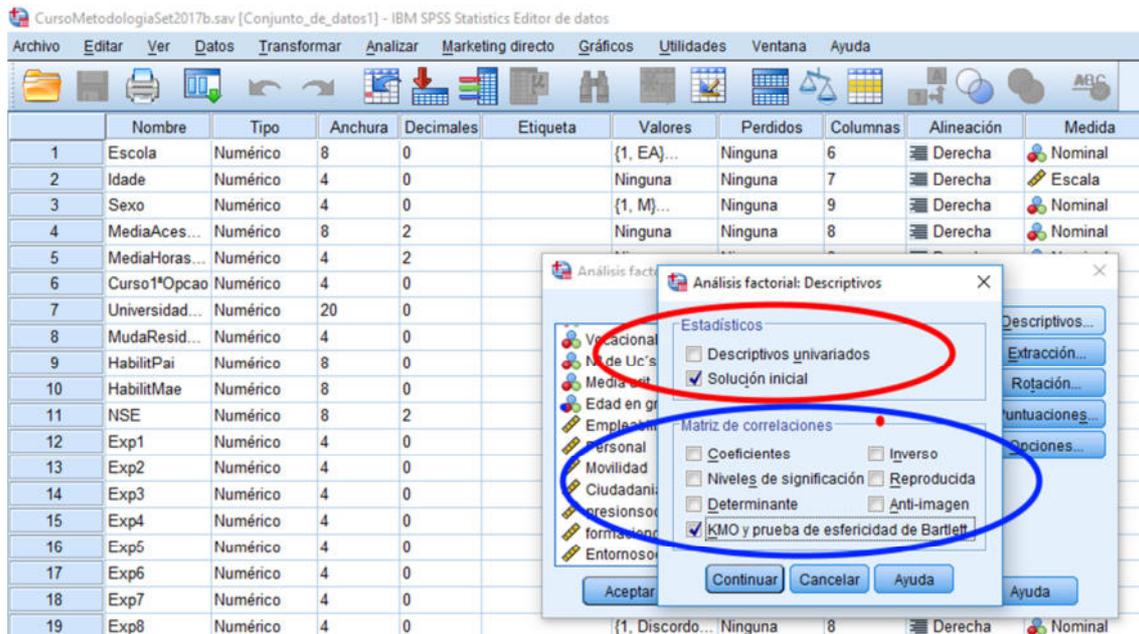


Fig. 21. Descriptivos del análisis factorial

El resultado que nos da nos indica si el análisis factorial es posible o significativo. Si el resultado de KMO es inferior a 0,70 o bien la prueba de esfericidad de Bartlett no es significativa, hay problemas y no se acepta la estructura factorial. Si el resultado de KMO es superior a 0,70 o la prueba de esfericidad de Bartlett es significativa, entonces se acepta la estructura factorial.

En nuestro caso, vemos los siguientes resultados. Como KMO es mayor que 0,70 y la prueba de esfericidad es significativa se acepta la estructura factorial [Fig. 22].

## ➔ Análisis factorial

[Conjunto\_de\_datos1] C:\Users\ancer\Desktop\CursoMetodologiaSet2017b.sav

Prueba de KMO y Bartlett		
Medida Kaiser-Meyer-Olkin de adecuación de muestreo		,953
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	43041,143
	gl	861
	Sig.	,000

Fig. 22. Resultados del análisis factorial

Una vez que vemos que se puede aceptar la estructura factorial volvemos al análisis factorial, para continuar con el análisis y ver el número de factores que podemos extraer. Repetimos la secuencia: analizar-reducción de dimensiones-factor [Fig. 23].

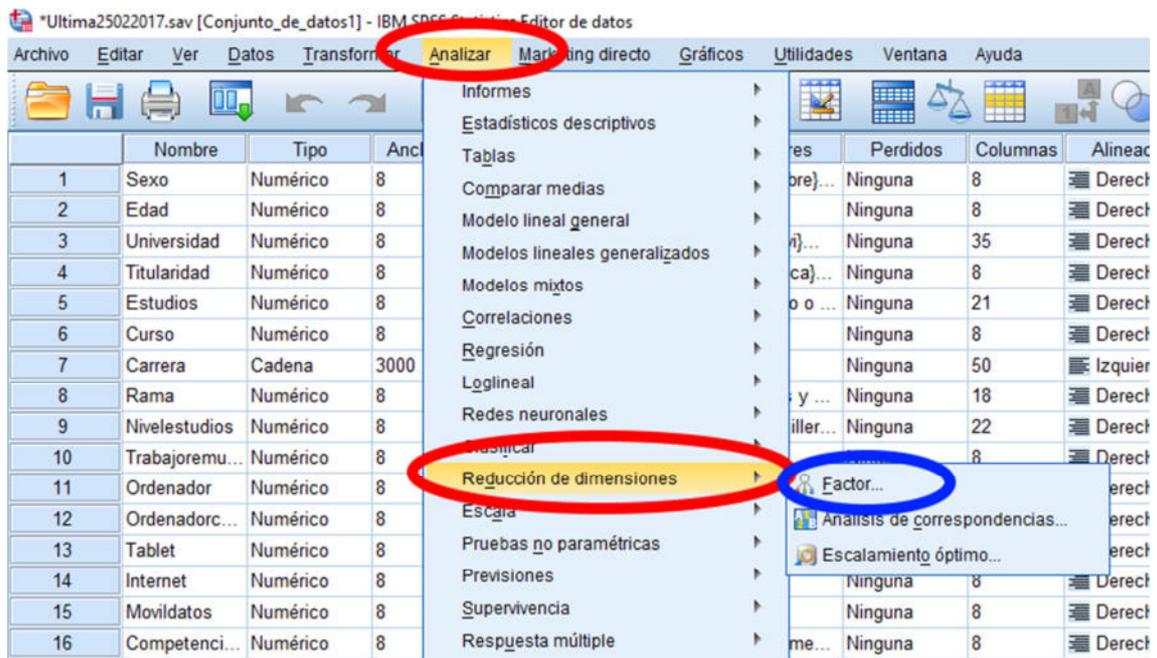


Fig. 23. Análisis factorial 2

Primero, pinchamos en la opción de extracción (círculo azul), para seleccionar el método [Fig. 24].

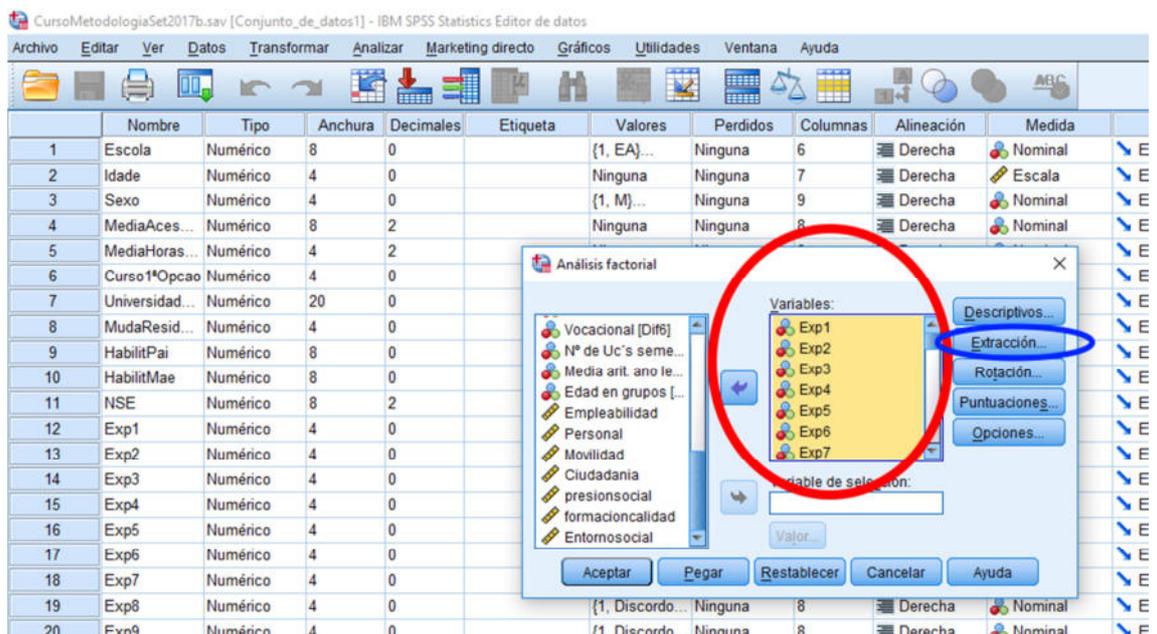


Fig. 24. Extracción de factores

En selección del método tenemos en cuenta que se suele utilizar el método de “Componentes Principales”, ya que es el que realiza análisis “más puros”. Pinchamos en continuar [Fig. 25].

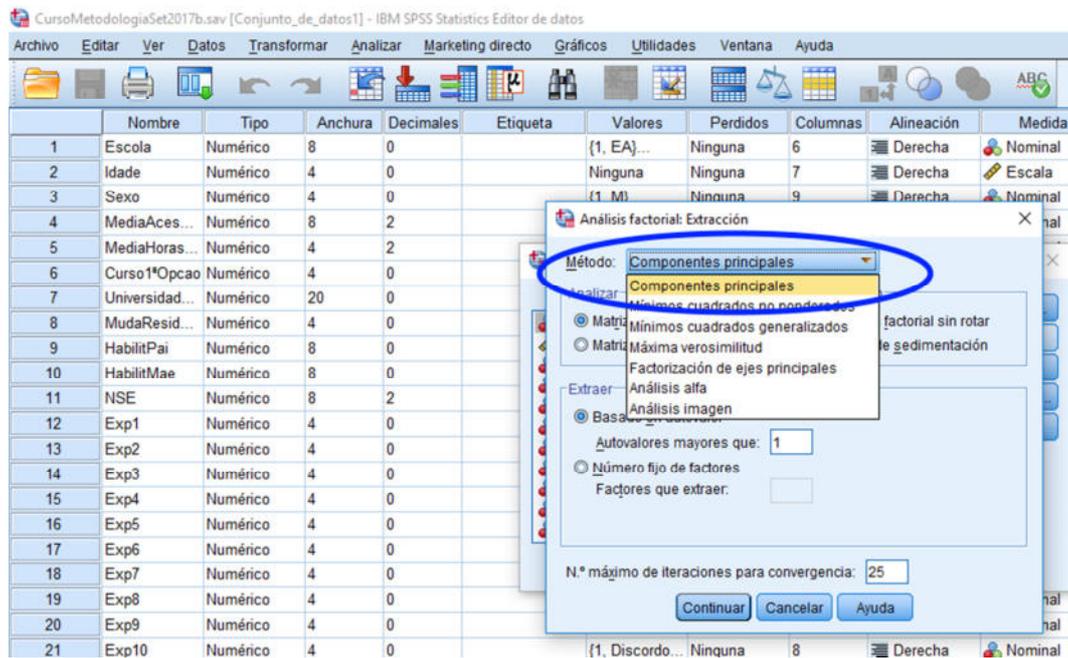


Fig. 25. Método de extracción de factores

También podemos pedir el gráfico de sedimentación para tener una visualización gráfica de los resultados y en el sistema de extracción, marcar la opción “basado en autovalores” (por defecto) o bien marcar un número fijo (en cuyo caso se refiere al número de factores que queremos que salgan) [Fig. 26].

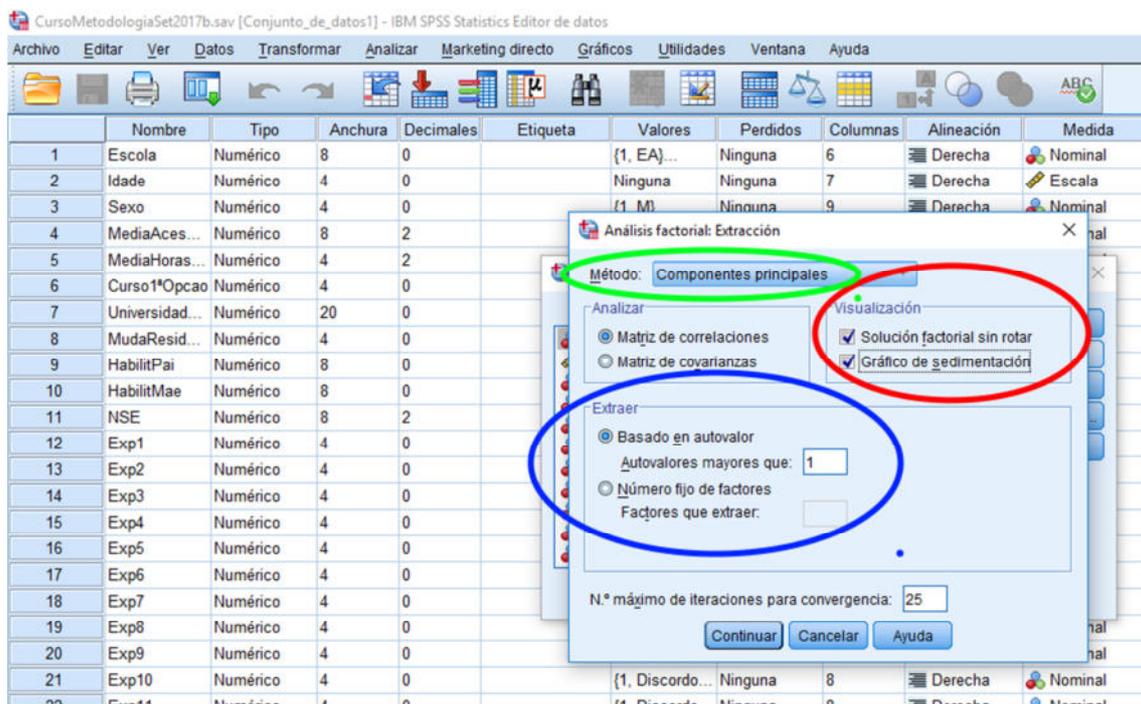


Fig. 26. Otras opciones de visualización de factores

Según el principio de Kaiser, un factor es válido cuando su valor es superior a la unidad en la tabla de sumas de extracción de cargas al cuadrado. Por tanto, en el caso de la figura de resultados siguiente, vemos que existen 7 factores, indicándose en cada uno el porcentaje de varianza que explica cada factor [Fig. 27].

Varianza total explicada						
Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	13,062	31,100	31,100	13,062	31,100	31,100
2	3,025	7,202	38,302	3,025	7,202	38,302
3	2,426	5,776	44,078	2,426	5,776	44,078
4	1,745	4,155	48,232	1,745	4,155	48,232
5	1,675	3,988	52,221	1,675	3,988	52,221
6	1,159	2,759	54,979	1,159	2,759	54,979
7	1,084	2,581	57,561	1,084	2,581	57,561
8	,918	2,185	59,745			
9	,838	1,996	61,742			
10	,784	1,867	63,608			
11	,765	1,822	65,430			
12	,741	1,765	67,195			
13	,728	1,733	68,928			
14	,674	1,604	70,533			
15	,665	1,584	72,117			
16	,659	1,568	73,685			

7 factores ya que su puntuación es > 1

El primer factor explica el 31% de la varianza, el segundo el 7%, etc.

Fig. 27. Resultados: número de factores

Finalmente, volvemos al análisis factorial, para continuar con el análisis y ver los ítems que pertenecen a cada factor. Repetimos la secuencia: analizar-reducción de dimensiones-factor [Fig. 28]].

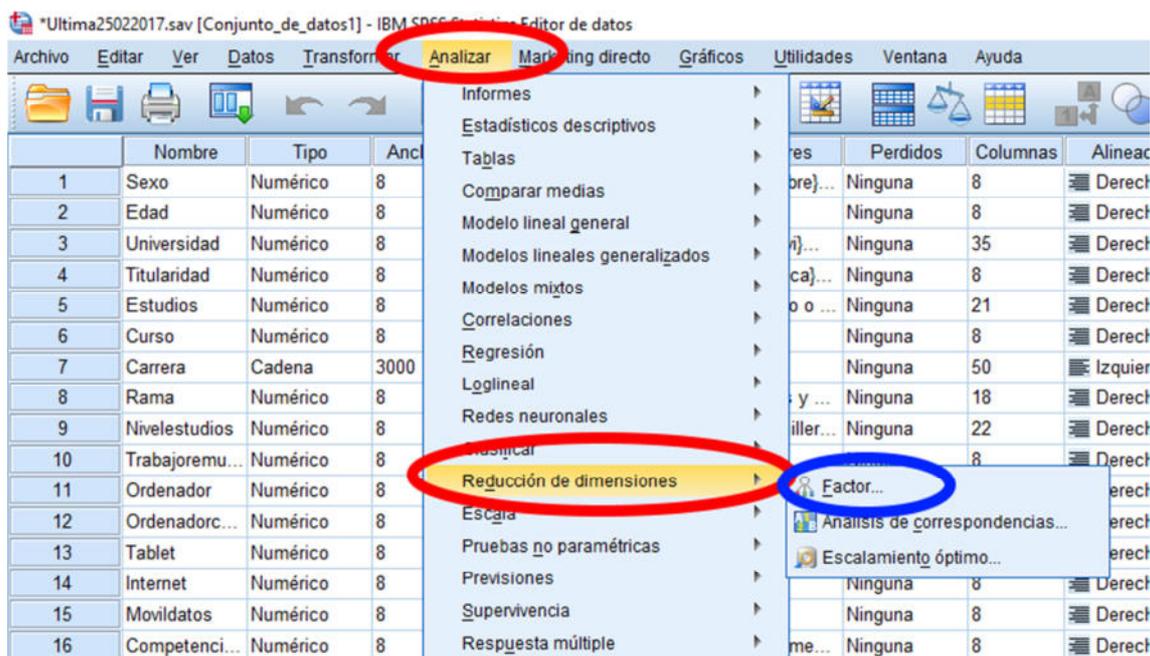


Fig. 28. Análisis factorial 3

Ahora pinchamos en la opción de rotación, donde tenemos diferentes métodos [Fig. 29].

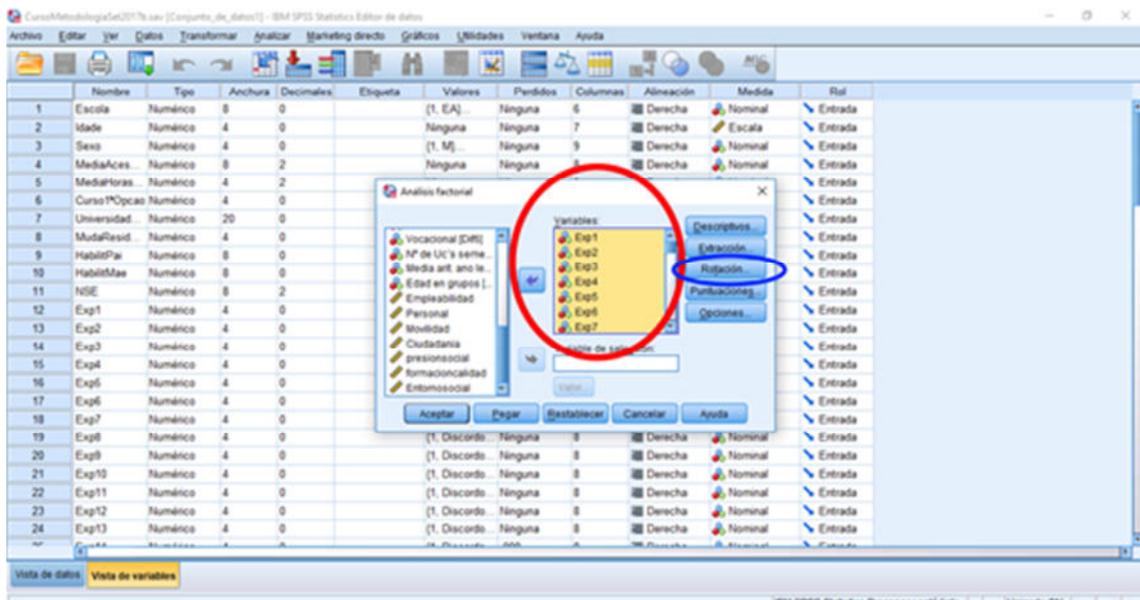


Fig. 29. Rotación de factores

Generalmente se utiliza el método “Varimax”, ya que permite mayor variabilidad y asume la independencia entre las variables seleccionadas. También se puede utilizar el método “Oblimin directo” que asume que los valores pueden estar cruzados entre sí [Fig. 30].

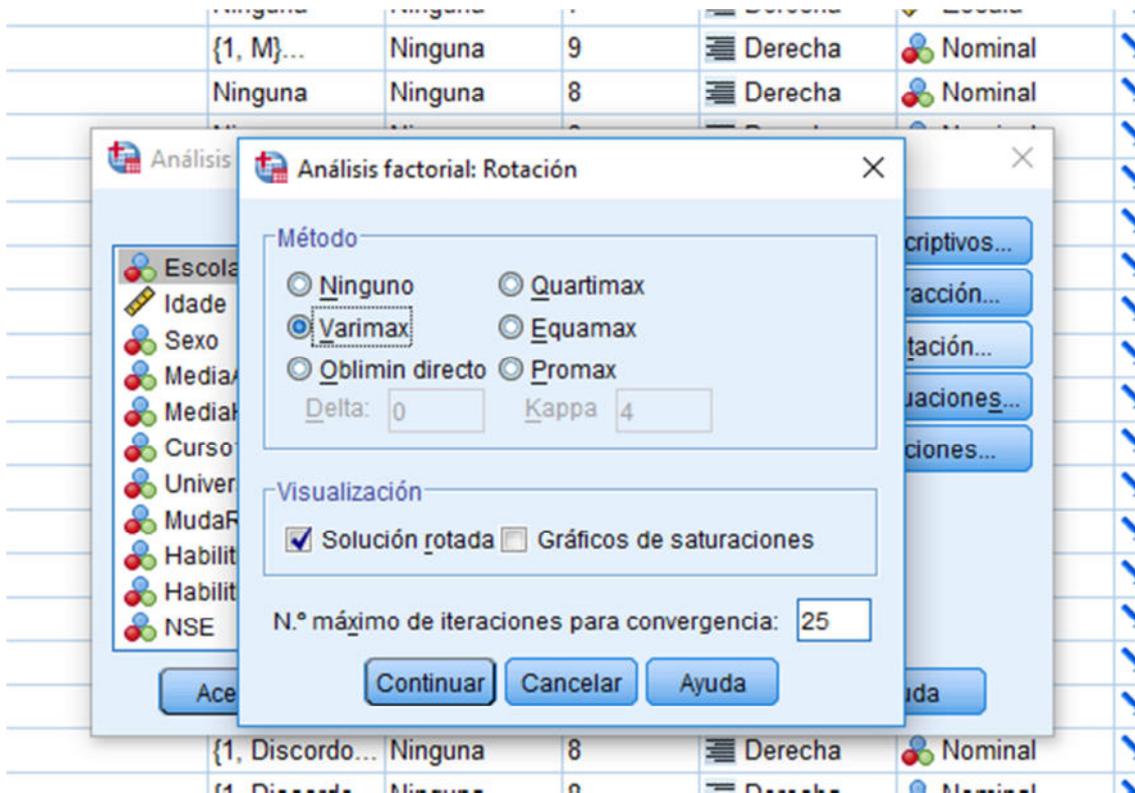


Fig. 30. Métodos de rotación de factores

En la tabla de resultados denominada “Matriz de componente rotado” vemos la vinculación de cada ítem con cada factor. El cuadrado de ese valor nos indica la varianza del factor que explica cada ítem. El ítem 1 pertenecería al factor 7, el ítem 2 al factor 1, el ítem 3 al factor 2, etc [Fig. 31].

**Matriz de componente<sup>a</sup>**

	Componente						
	1	2	3	4	5	6	7
Exp1	,413	-,280	,102	,197	-,033	,123	,450
Exp2	,545	-,050	-,138	-,195	-,070	,352	,323
Exp3	,414	,658	,088	,113	-,122	,066	,054
Exp4	,614	,132	-,154	-,275	,241	,040	,139
Exp5	,478	-,359	,361	,213	,226	,207	,173
Exp6	,455	,202	-,069	-,154	,229	-,366	,320
Exp7	,616	-,009	,168	-,301	-,376	-,015	,173
Exp8	,535	-,195	-,292	,192	-,203	-,012	,263
Exp9	,589	-,082	-,003	-,290	,008	,407	,101
Exp10	,458	,651	,126	,221	-,032	,161	,037

Fig. 31. Matriz de componentes rotados

No obstante en la pestaña de opciones del análisis de factor [Fig. 32], podemos ordenar por tamaño e incorporar un valor mínimo para que no se incluyan los valores que tengan una varianza menor de 0,30 o 0,40, que son los que se suelen utilizar en función de la fortaleza que queramos [Fig. 33].

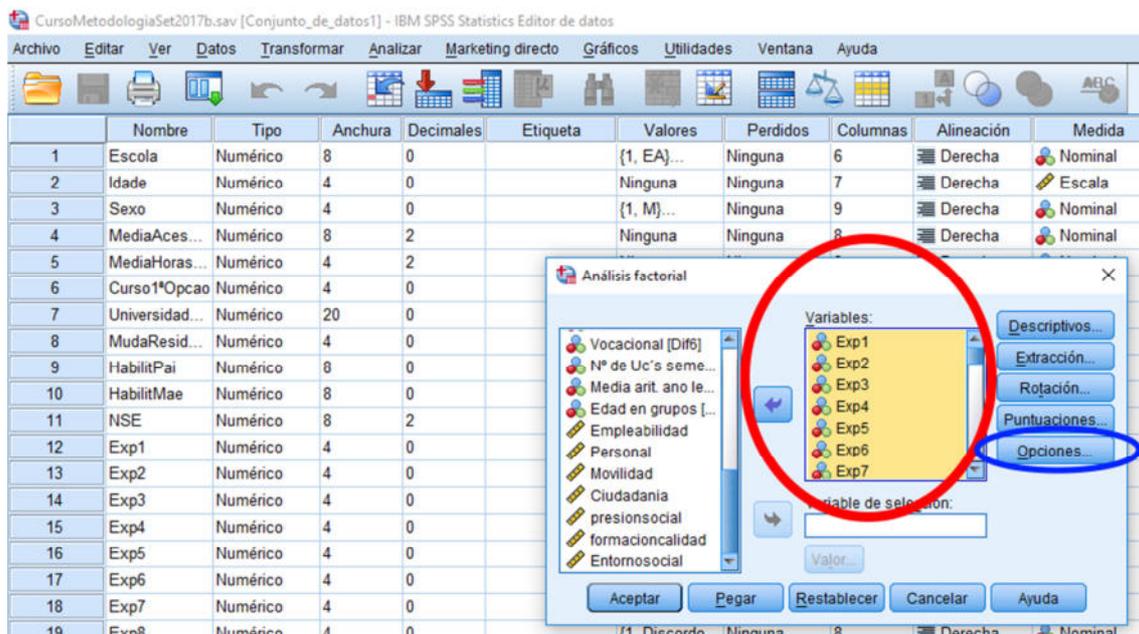


Fig. 32. Opciones de análisis de factores

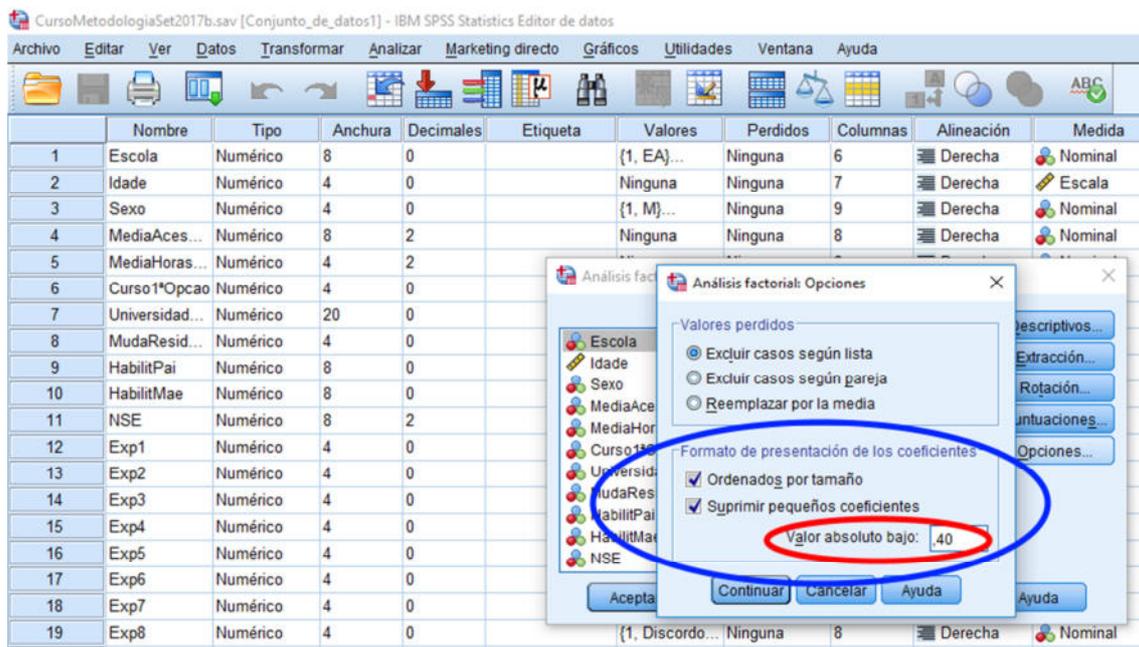


Fig. 33. Presentación de coeficientes en análisis factorial

En este caso, cuando pasamos a la matriz de componente rotado ya sale ordenada [Fig. 34].

Matriz de componente rotado <sup>a</sup>							
	Componente						
	1	2	3	4	5	6	7
Exp15	,753						
Exp22	,745						
Exp29	,669						
Exp8	,645						
Exp34	,632						
Exp36	,586						
Exp23	,534						
Exp24	,470		,457				
Exp13	,439					,407	
Exp26	,401						
Exp39		,661					
Exp32		,607					
Exp30		,605					
Exp37	,456	,601					
Exp18		,573					
Exp9		,547					,459
Exp4		,529				,418	
Exp11		,474				,416	
Exp16		,467					
Exp25		,456					

Fig. 34. Matriz de componente rotado

Un ítem que tenga dos resultados robustos en dos factores podemos situarlo en el factor conveniente según el modelo teórico.

El análisis factorial confirmatorio no trabaja con componentes principales y se puede subir el mínimo a partir del cual se considera a 0,6.

## 5.-Fiabilidad

La confianza o consistencia interna la calculamos mediante el  $\alpha$  de Cronbach, que nos indica en definitiva si los resultados son representativos del conjunto. Es decir, mide la fiabilidad de la escala de medida.

Por lo tanto podemos calcular la fiabilidad de un factor en un análisis factorial o bien del test en su conjunto.

La secuencia para analizar la fiabilidad es: analizar-escalas-análisis de fiabilidad [Fig. 35].

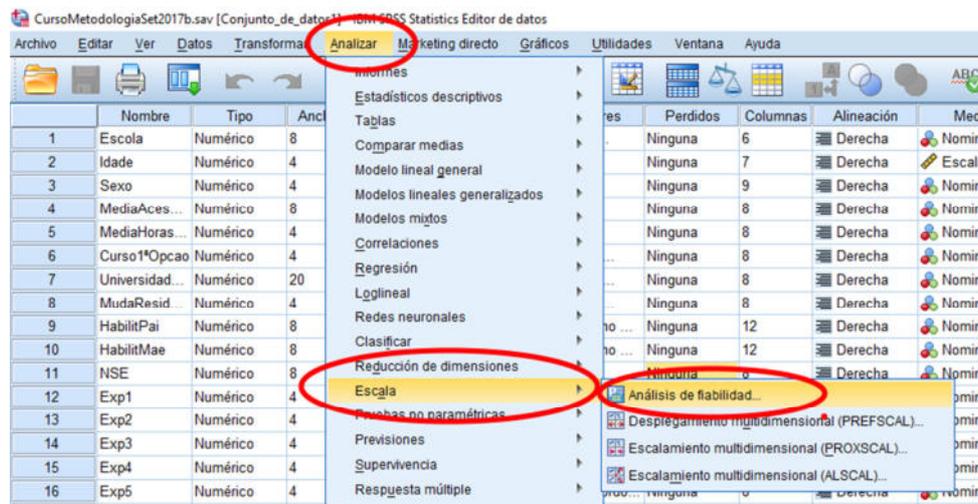


Fig. 35. Análisis de fiabilidad

Seleccionamos las variables del factor o del cuestionario para realizar el análisis y pinchamos en estadísticos [Fig. 36].

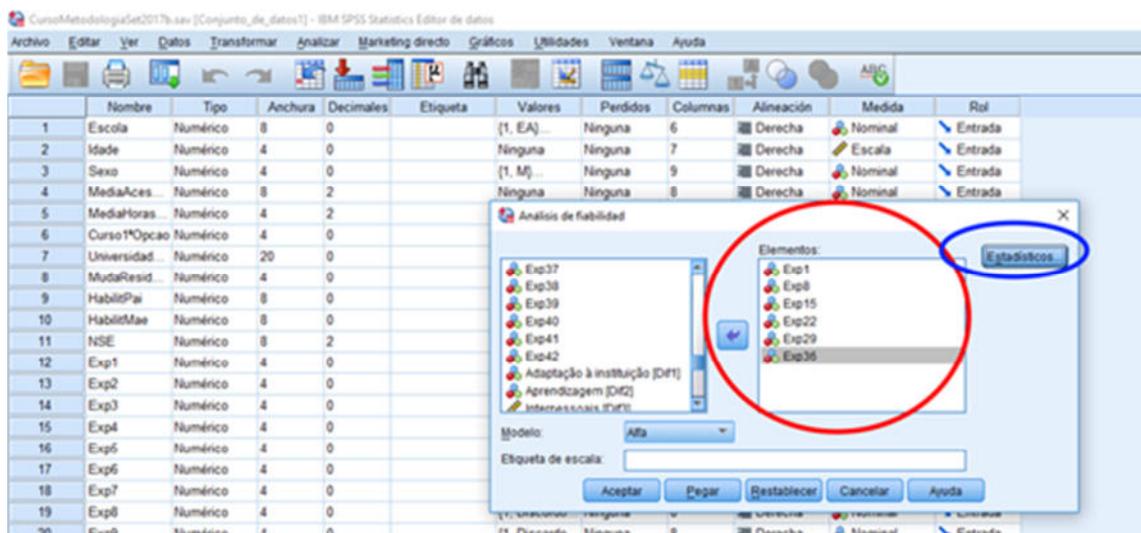


Fig. 36. Elementos del análisis de fiabilidad

Marcamos: elemento, escala y escala si se elimina elemento [Fig. 37].

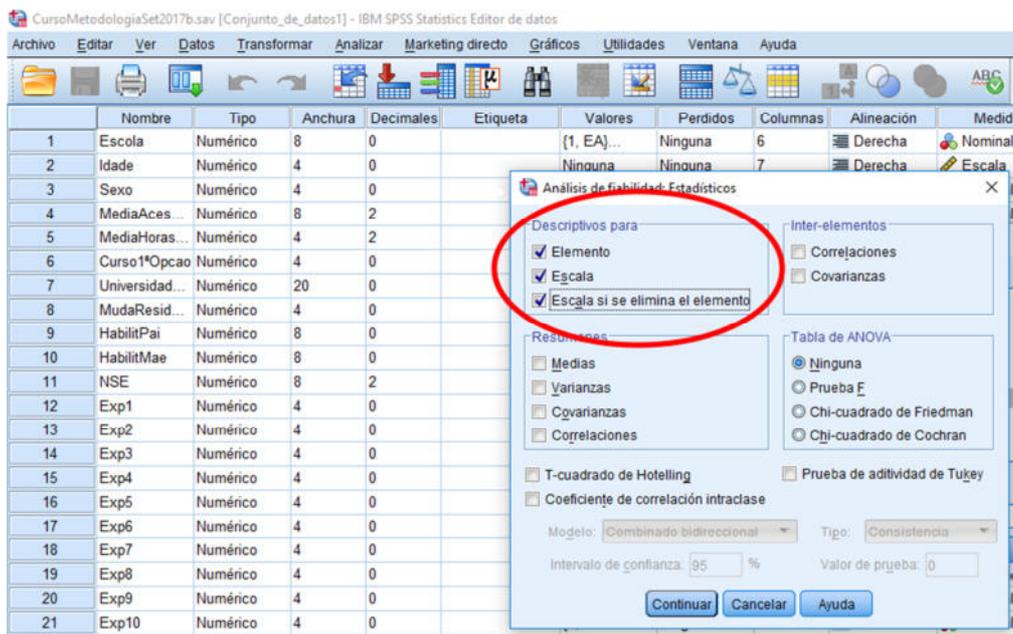


Fig. 37. Descriptivos para el análisis de fiabilidad

De esta forma obtenemos la tabla de resultados, viendo en el cuadro “Estadística de fiabilidad” el resultado del alfa [Fig. 38].

Estadísticas de fiabilidad	
Alfa de Cronbach	N de elementos
,777	6

Fig. 38. Resultado del análisis de fiabilidad

También podemos ver en las estadísticas del total del elemento, el valor del alfa si eliminamos cada uno de los elementos que hemos utilizado en el análisis. Así, si suprimimos el ítem exp1, el valor del alfa subirá a ,822 o lo que es lo mismo 82,2% [Fig. 39].

Estadísticas de total de elemento				
	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
Exp1	28,08	6,037	,367	,822
Exp8	27,41	6,912	,609	,731
Exp15	27,43	6,582	,658	,716
Exp22	27,48	6,719	,626	,724
Exp29	27,51	6,758	,532	,742
Exp36	27,61	6,565	,555	,736

Fig. 39. Estadísticas del total del elemento del análisis de fiabilidad

## 6.- Comparación de grupos: t de student y ANOVA

En general, si tenemos muchos datos utilizaremos pruebas paramétricas bajo el supuesto de que la muestra se ajusta a la curva normal. Si la muestra es pequeña utilizamos pruebas no paramétricas que son menos robustas. Por eso siempre es preferible utilizar pruebas paramétricas si los datos lo permiten.

La t de student nos permite comparar si existen diferencias significativas en una misma variable entre dos grupos diferentes o en dos momentos diferentes. En el caso de querer comparar las diferencias de una variable en tres o más grupos utilizaremos el análisis de varianzas (ANOVA).

### 6.1.- t de Student

La secuencia para realizar el análisis de t de student es: analizar-comparar medias-prueba T para muestras independientes (si queremos analizar una misma muestra en dos momentos diferentes elegiremos la opción Prueba T para muestras relacionadas) [Fig. 40].

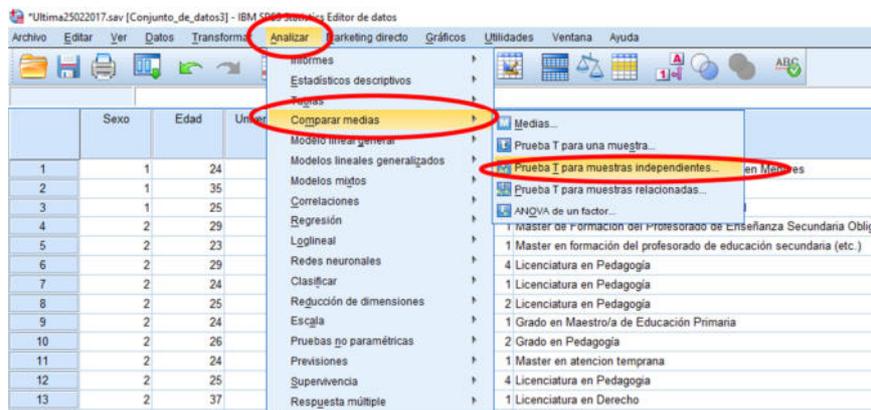


Fig. 40. Pruebas t

En la ventana de la prueba t seleccionamos las variables de prueba, que serían las variables dependientes y la variable de agrupación, que sería la variable independiente [Fig. 41].

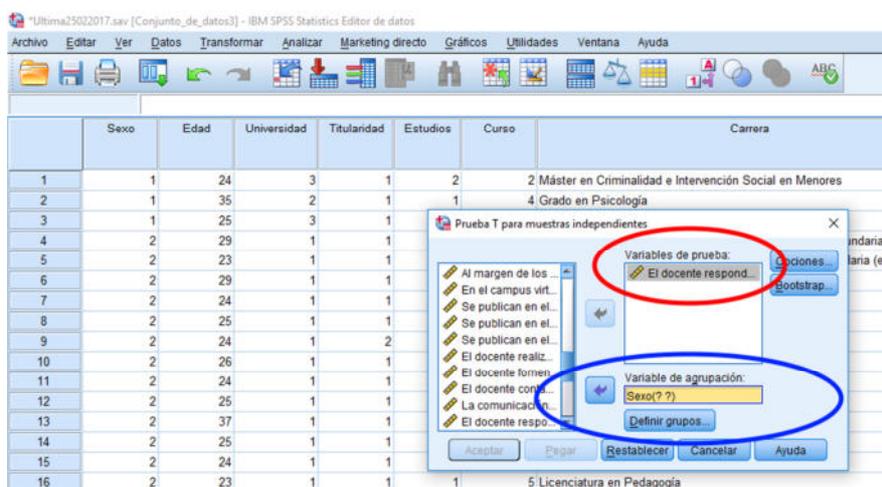


Fig. 41. Selección de variables de las pruebas t

Posteriormente debemos definir los grupos de la variable de agrupación, pinchando la opción definir grupos, donde concretaremos los dos grupos. En este caso el grupo uno será el valor 1 (hombres) y el grupo 2 el valor 2 (mujeres) [Fig. 42].

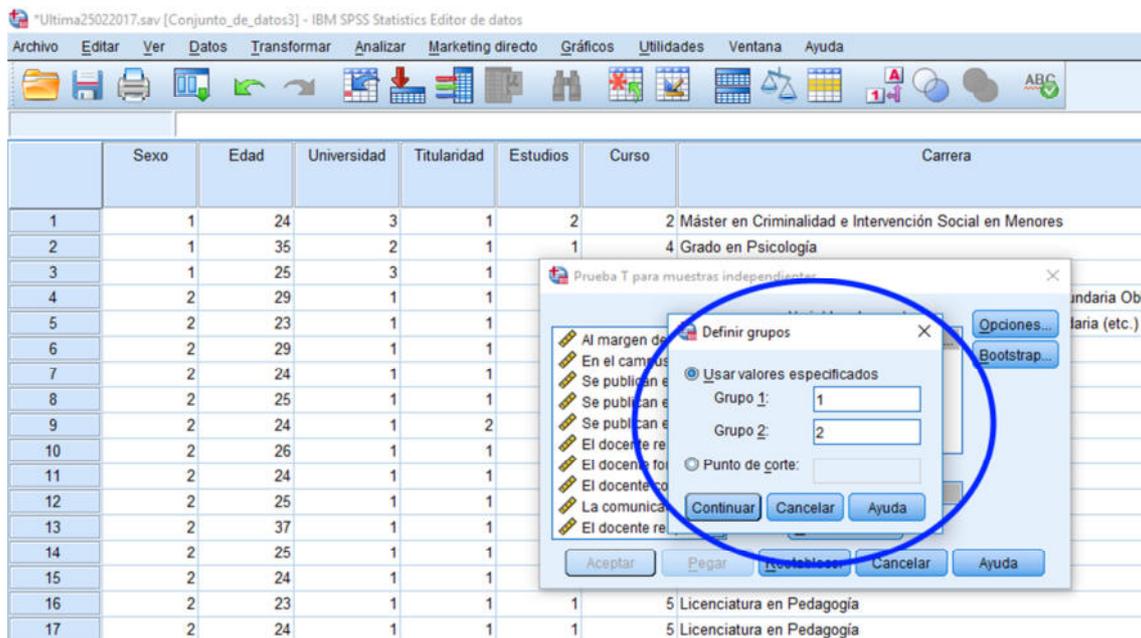


Fig. 42. Definición de variables en pruebas t

Al continuar los grupos ya aparecerán definidos en la pantalla general y podremos aceptar para ver los resultados [Fig. 43].

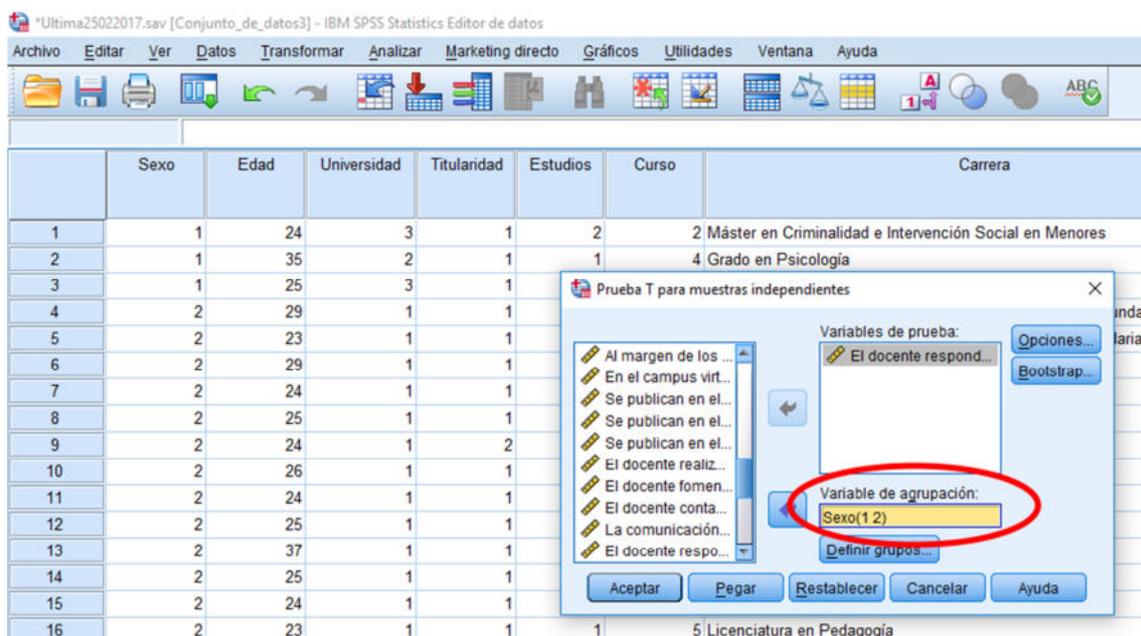


Fig. 43. Variable de agrupación en pruebas t

En las tablas de resultados, observaremos en primer lugar la significatividad de la Prueba de Levene. Esta nos dice si hemos de asumir en los resultados varianzas iguales (con lo cual miraremos los resultados de la fila superior) o si no se asumen varianzas iguales. Si la prueba de Levene es significativa NO asumiremos varianzas iguales. Si la prueba de Levene no es significativa asumiremos varianzas iguales [Fig. 44].

T-TEST GROUPS=Sexo (1 2)  
/MISSING=ANALYSIS  
/VARIABLES=Actividadesdebate  
/CRITERIA=CI (.95).

**Prueba T**

**Estadísticas de grupo**

	Sexo	N	Media	Desviación estándar	Media de error estándar
Se publican en el campus virtual actividades que fomentan la exposición de ideas, el debate...	Hombre	94	1,98	,776	,080
	Mujer	217	2,24	,786	,053

**Prueba de muestras independientes**

		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias						
		F	Sig.	t	gl.	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
								Inferior		Superior
Se publican en el campus virtual actividades que fomentan la exposición de ideas, el debate...	Se asumen varianzas iguales	2,457	,118	-2,698	309	,007	-.261	,097	-.451	-.071
	No se asumen varianzas iguales			-2,713	178,947	,007	-.261	,096	-.451	-.071

Fig. 44. Prueba de Levene

En este caso la prueba de Levene no es significativa por lo que asumiremos varianzas iguales y miraremos el resultado de la prueba t de la fila superior [Fig. 45].

T-TEST GROUPS=Sexo (1 2)  
/MISSING=ANALYSIS  
/VARIABLES=Actividadesdebate  
/CRITERIA=CI (.95).

**Prueba T**

**Estadísticas de grupo**

	Sexo	N	Media	Desviación estándar	Media de error estándar
Se publican en el campus virtual actividades que fomentan la exposición de ideas, el debate...	Hombre	94	1,98	,776	,080
	Mujer	217	2,24	,786	,053

**Prueba de muestras independientes**

		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias						
		F	Sig.	t	gl.	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
								Inferior		Superior
Se publican en el campus virtual actividades que fomentan la exposición de ideas, el debate...	Se asumen varianzas iguales	2,457	,118	-2,698	309	,007	-.261	,097	-.451	-.071
	No se asumen varianzas iguales			-2,713	178,947	,007	-.261	,096	-.451	-.071

Fig. 45. Ejemplo de prueba t

Al analizar el valor de la prueba T, vemos que efectivamente hay diferencias significativas, lo que nos indica que existen diferencias entre los dos grupos a la hora de valorar esa variable. Por otro lado el hecho de que el valor de T sea negativo indica que el valor del segundo grupo (mujeres) es mayor. Así, las mujeres están más de acuerdo en que los docentes publican en el

campus actividades que fomentan la exposición de ideas, que los hombres. Si el valor fuera positivo, la valoración del primer grupo sería mayor.

A la hora de exponer estos resultados, utilizaríamos la expresión:

$$t(309) = -2,698, p.007$$

siendo 309 el valor de los grados de libertad, 2,698 el valor de t y .007 la significatividad.

En el caso de una significatividad de .000 se utiliza la expresión  $p < .001$ , porque en realidad el valor nunca sería de .000 sino menor, es decir por ejemplo de .000001, etc.

## 6.2.- ANOVA

El análisis ANOVA permite comparar la misma variable en 3 o más grupos.

La secuencia de análisis sería la siguiente: Analizar-comparar medias-ANOVA de un factor [Fig. 46].

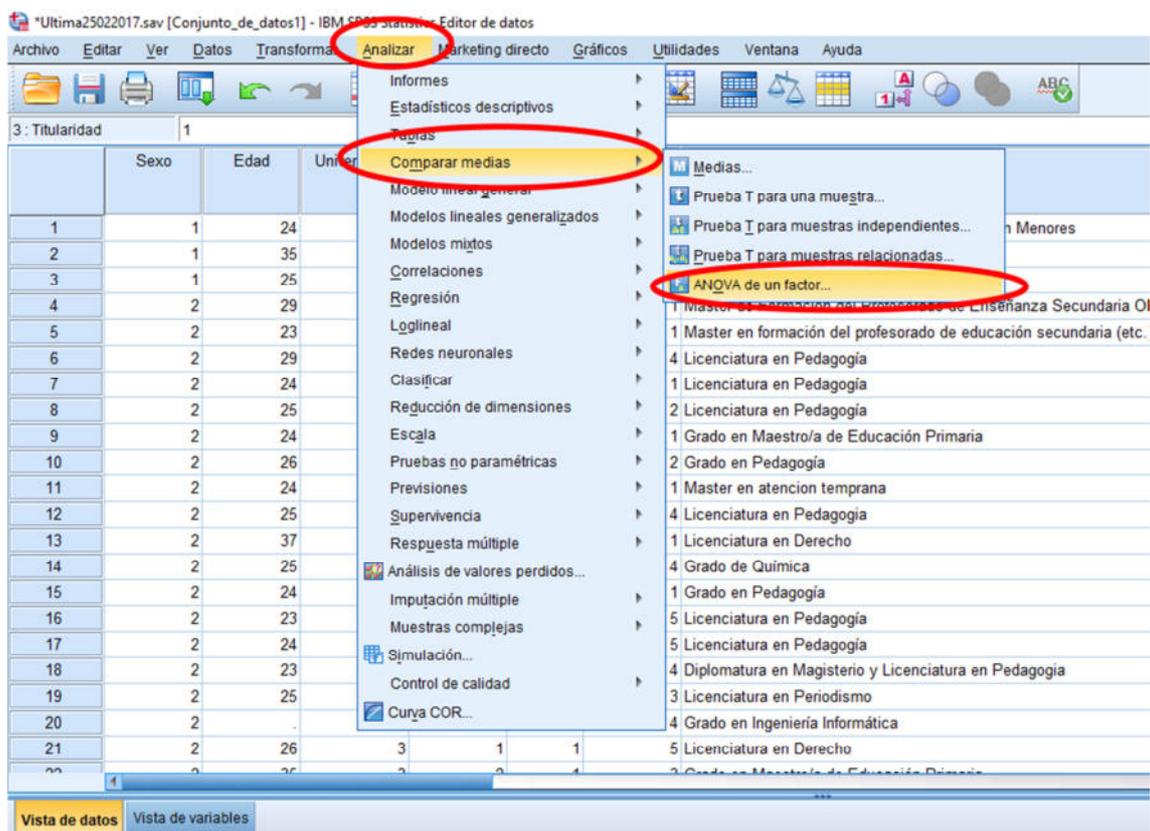


Fig. 46. ANOVA

Una vez en la ventana de análisis seleccionamos la lista de dependientes (variables dependientes) y el factor (que sería la variable independiente) [Fig. 47].

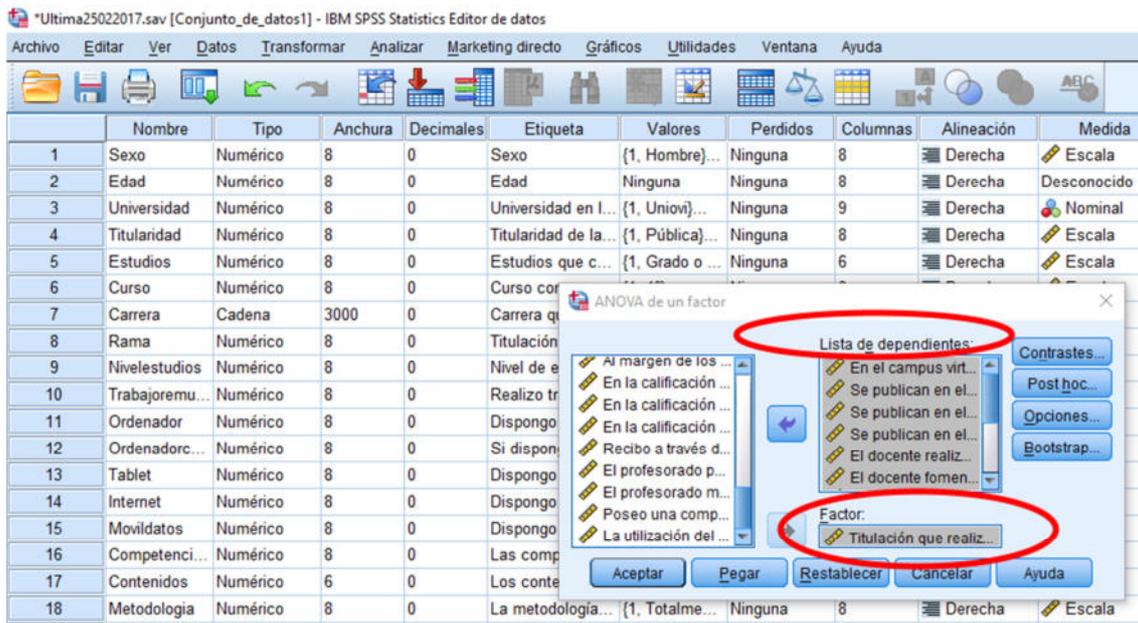


Fig. 47. Selección de variables en ANOVA

ANOVA así establecido nos indicará si existen diferencias estadísticamente significativas pero no el sentido de estas, que veremos posteriormente. Para saber si hay diferencias estadísticamente significativas entre grupos observamos si las diferencias son significativas (variables con el círculo rojo) [Fig. 48].

		ANOVA				
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
En el campus virtual acceso a los diferentes formatos de archivo (texto, audio, video...) publicados como complemento curricular	Entre grupos	7,832	4	1,958	3,208	.013
	Dentro de grupos	185,572	304	,610		
	Total	193,405	308			
Se publican en el campus virtual ejercicios y supuestos prácticos que hacen referencia a problemas y situaciones reales	Entre grupos	4,140	4	1,035	1,596	,175
	Dentro de grupos	197,148	304	,649		
	Total	201,288	308			
Se publican en el campus virtual actividades que fomentan el pensamiento crítico, la reflexión personal...	Entre grupos	7,823	4	1,956	3,029	.018
	Dentro de grupos	196,319	304	,646		
	Total	204,142	308			
Se publican en el campus virtual actividades que fomentan la exposición de ideas, el debate...	Entre grupos	6,875	4	1,719	2,824	.025
	Dentro de grupos	185,035	304	,609		
	Total	191,909	308			
El docente realiza labores de asesoramiento y orientación en sus interacciones a través del campus virtual	Entre grupos	4,146	4	1,036	1,709	,148
	Dentro de grupos	184,417	304	,607		
	Total	188,563	308			

Fig. 48. Diferencias en ANOVA

El resultado de la tabla que nos devuelve el SPSS tendría el formato de la siguiente [Fig. 49]:

**ANOVA**

Poseo una competencia digital adecuada: conocimientos, capacidades, habilidades, valores

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	6,362	4	1,590	3,082	,016
Dentro de grupos	156,887	304	,516		
Total	163,249	308			

Fig. 49. Resultados ANOVA

Y de cara a escribir un artículo, o mencionar un resultado utilizaríamos el siguiente formato:

$$F(4,304) = 3,082, p=0,16$$

Donde 4 es el valor menor de los grados de libertad; 304, el valor mayor de los grados de libertad; F, el valor de F y p la significatividad (recordando que si  $p=.000$  se nota de la siguiente manera:  $p < .001$ ).

Para analizar el sentido de estas diferencias volvemos a la ventana de selección de variables y pulsamos la casilla "Post hoc" (círculo azul). Aunque en realidad podemos hacerlo también en el paso anterior descartando posteriormente aquellas variables en que no haya diferencias significativas, quizás sea mejor hacerlo por pasos para eliminar previamente las variables donde no las hay, simplificando las tablas de resultados.

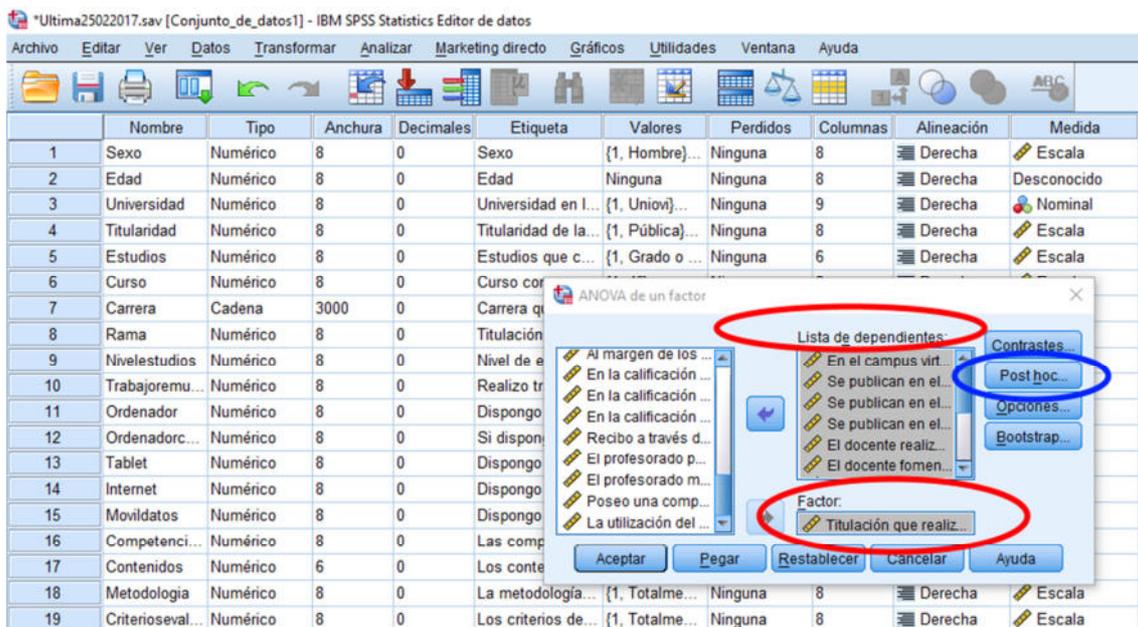


Fig. 50. Pestaña post hoc de ANOVA

En la nueva ventana, si asumimos varianzas iguales marcamos la opción de Scheffe (círculo rojo), aunque en algunos casos si bien menos habitualmente también se pueden utilizar las opciones de análisis de Bonferroni o Tukey (círculo azul) [Fig. 51].

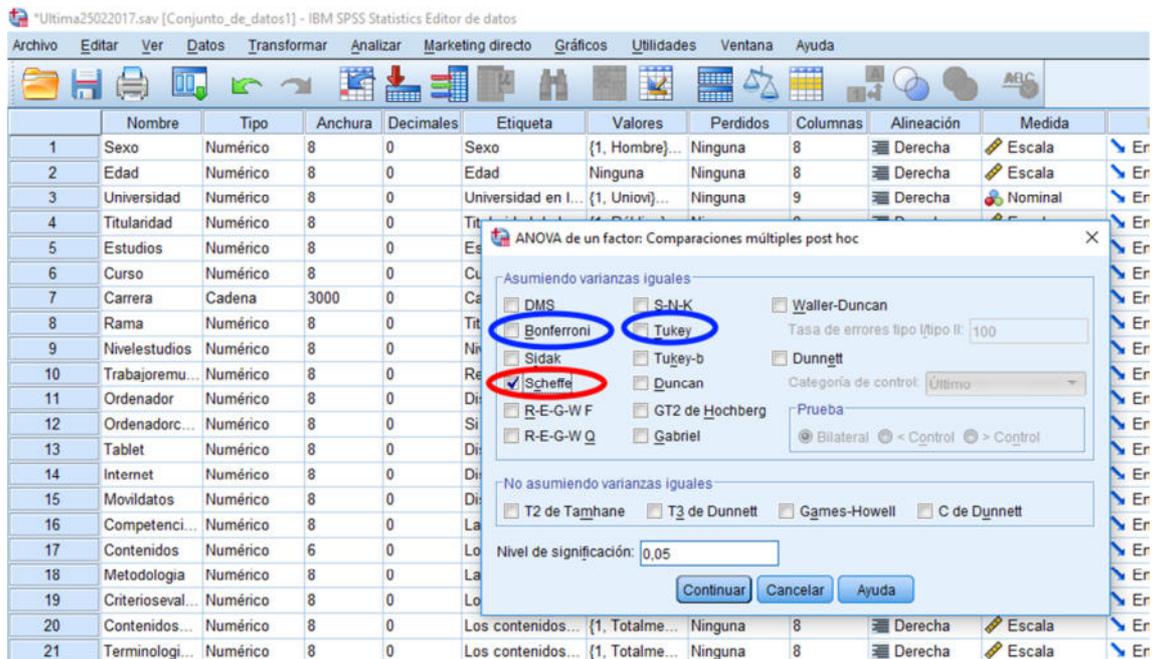


Fig. 51. Asunción de varianzas en ANOVA

El resultado de la tabla que nos devuelve el SPSS sería del formato de la siguiente [Fig. 52]:

**ANOVA**

Poseo una competencia digital adecuada: conocimientos, capacidades, habilidades, valores:

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	6,362	4	1,590	3,082	,016
Dentro de grupos	156,887	304	,516		
Total	163,249	308			

**Pruebas post hoc**

**Comparaciones múltiples**

Variable dependiente: Poseo una competencia digital adecuada: conocimientos, capacidades, habilidades, valores y actitudes relacionadas  
Scheffe

(I) Titulación que realizo (ejemplo: Grado en Psicología, Grado en Historia...)	(J) Titulación que realizo (ejemplo: Grado en Psicología, Grado en Historia...)	Diferencia de medias (I-J)	Error estándar	Sig.	95% de intervalo de confianza	
					Límite inferior	Límite superior
Artes y Humanidades	Ciencias de la Salud	,564	,206	,113	-,07	1,20
	Ciencias Jurídicas y Sociales	,577 <sup>a</sup>	,176	<b>,031</b>	,03	1,12
	Ingeniería y Arquitectura	,722	,268	,125	-,11	1,55
	Ciencias	,722	,261	,109	-,09	1,53
Ciencias de la Salud	Artes y Humanidades	-,564	,206	,113	-1,20	,07
	Ciencias Jurídicas y Sociales	,013	,126	1,000	-,38	,40
	Ingeniería y Arquitectura	,158	,238	,979	-,58	,90
	Ciencias	,158	,231	,976	-,56	,87

Fig. 52. Resultados ANOVA

Como se aprecia, nos vuelve a dar el valor de ANOVA y en las comparaciones múltiples podemos ver las diferencias entre los múltiples grupos. Así, en el ejemplo, podemos ver que para la variable “Poseo una competencia digital adecuada” las diferencias entre el grupo 1 (Artes y Humanidades) y el grupo 2 (Ciencias de la Salud) no son significativas ( $p=.113$ ), pero que sí lo son entre el grupo 1 (Artes y Humanidades) y el grupo 3 (Ciencias Sociales y Jurídicas), con una  $p=.031$ .

En este caso, vemos que hay diferencias estadísticamente significativas entre el grupo 1 y el 3. El signo de la diferencia de medias nos indica quién lo valora más positivamente y quién menos. En este caso la valoración del grupo de Artes y Humanidades es mayor, ya que la diferencia de medias es positiva y este es el primero grupo de la diferencia.

Para especificar los resultados utilizaremos la expresión:

$$\text{Diferencia de medias} = .577, p=0.31$$

Donde .577 es el valor de la diferencia de medias y .577 la significatividad estadística.

También podemos marcar los descriptivos en la pestaña de opciones para obtener estos datos. En ese caso primero pulsaremos la pestaña de opciones [Fig. 53] y posteriormente los estadísticos descriptivos [Fig. 54].

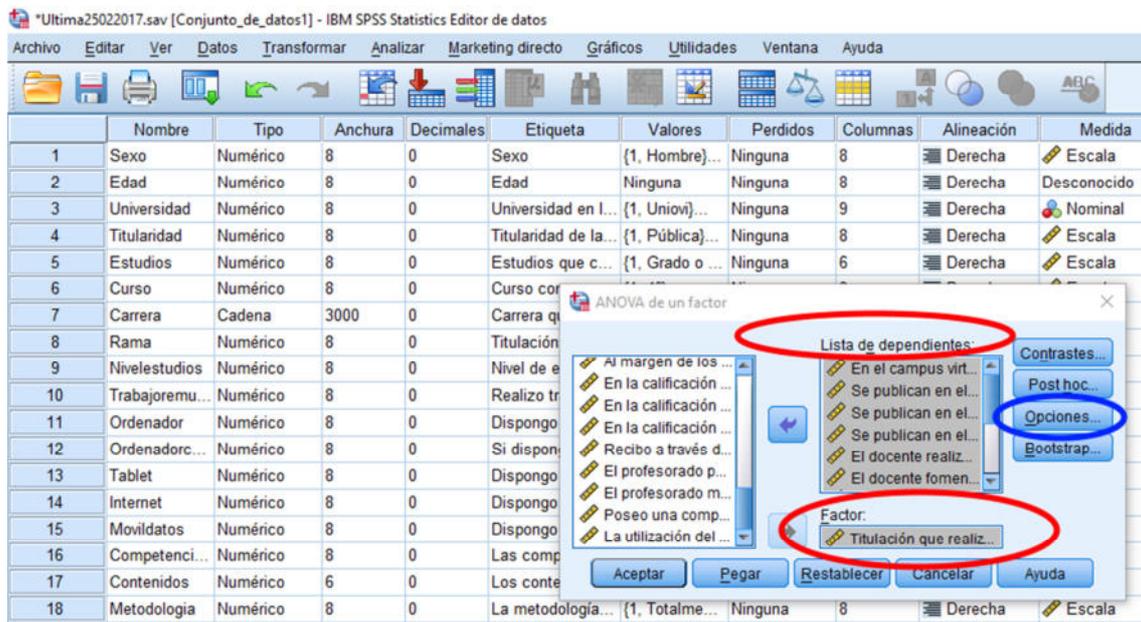


Fig. 53. Opciones en análisis ANOVA

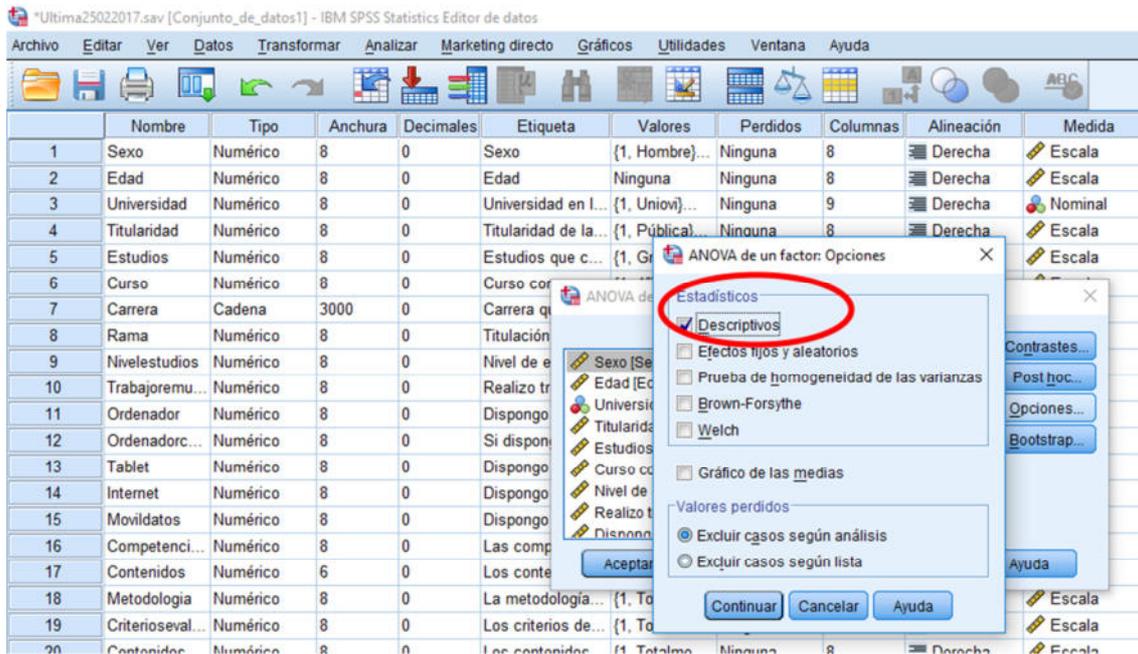


Fig. 54. Descriptivos en análisis ANOVA

En este caso la ventana de resultados nos dará primero la tabla de datos descriptivos (círculo rojo), posteriormente la tabla con los valores de ANOVA (círculo azul) y, finalmente, la tabla con las comparaciones múltiples (círculo verde) [Fig. 55].

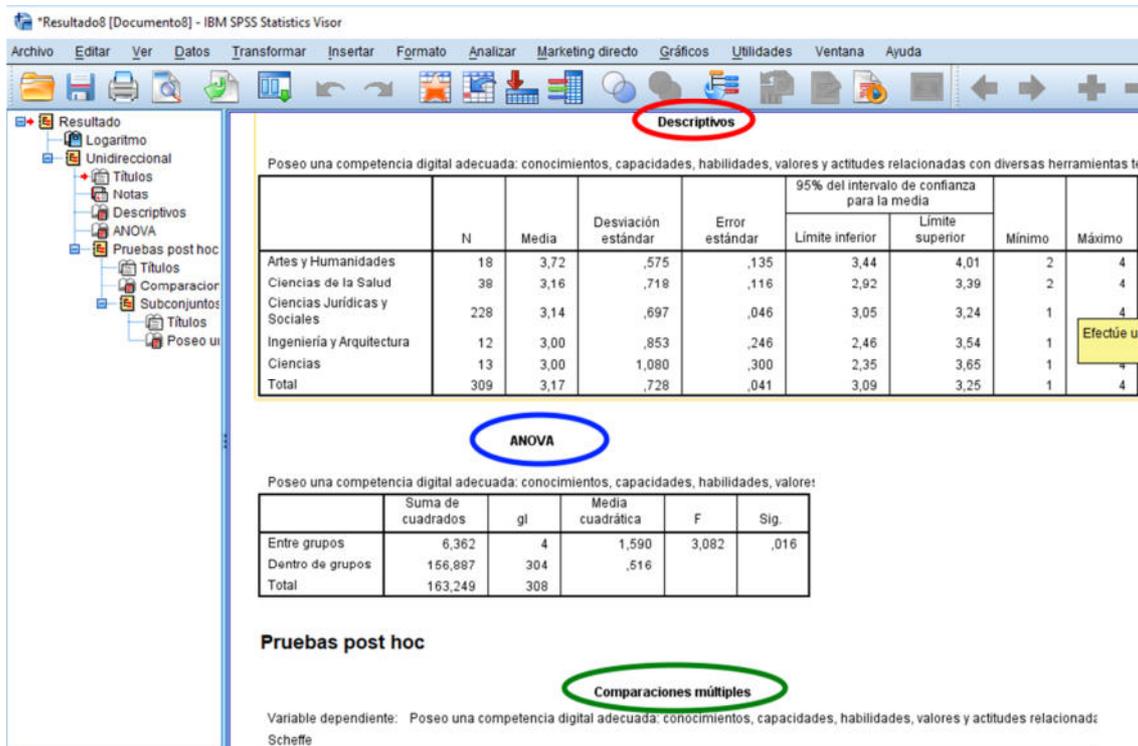


Fig. 55. Tablas de resultados ANOVA

## 7.- Correlación de variables

La prueba más utilizada para ver cómo correlacionan dos variables es el coeficiente de correlación de Pearson. La secuencia para calcularlo es: analizar-correlaciones-bivariadas [Fig. 56].

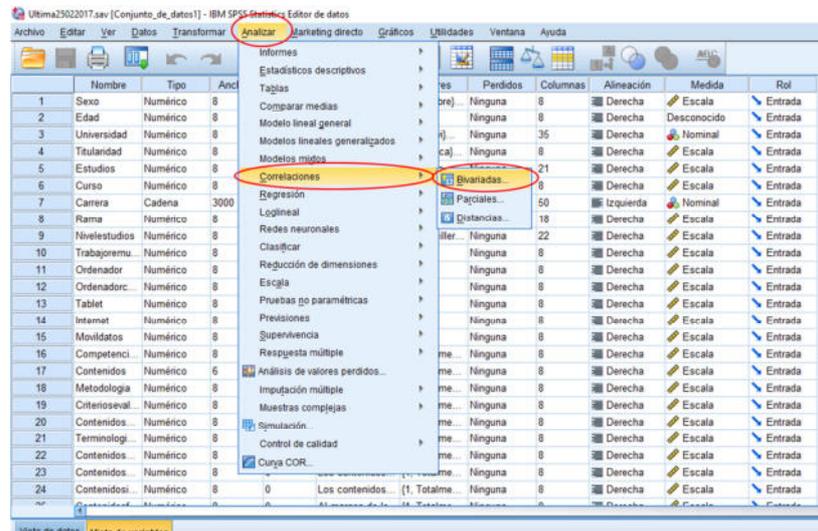


Fig. 56. Correlaciones

En la ventana correspondiente seleccionamos las variables, el coeficiente de correlación que queremos utilizar (círculo rojo) y la prueba de significación (círculo azul) [Fig. 57]. La prueba de significación más eficaz es la bilateral, donde la hipótesis nula indicaría que  $a=b$  y la hipótesis alternativa implicaría que hay diferencias, independientemente de que estas fueran " $a < b$ " o " $a > b$ ". Si la significación es unilateral solo contempla uno de los extremos de la curva normal (con una sig. .25). Si las variables son de escala utilizaremos el coeficiente de correlación de Pearson, para variables ordinales el de Kendall o Spearman, si bien Pearson se puede utilizar también con variables ordinales al ser más robusto.

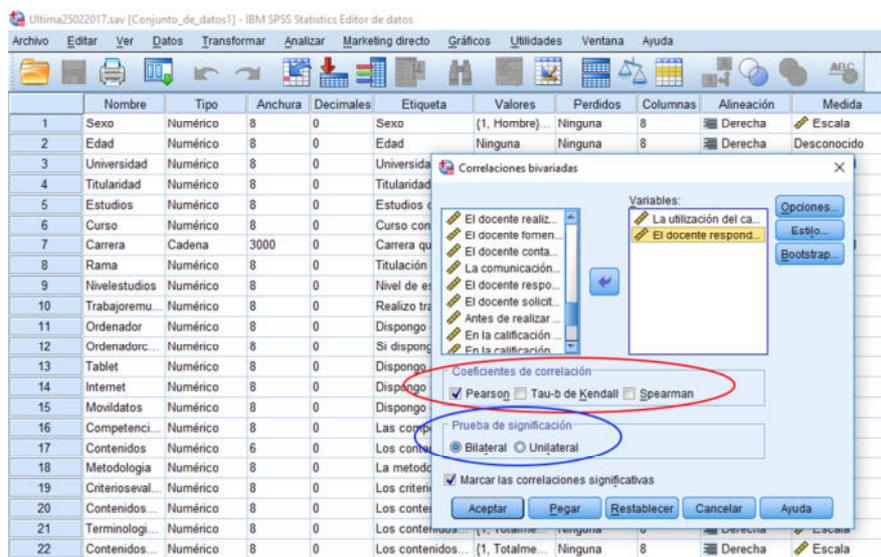


Fig. 57. Correlaciones bivariadas

En los resultados vemos la significatividad y si esta es  $<.05$  la relación es estadísticamente significativa (círculo rojo). El signo del coeficiente nos indica si la relación entre variables es directa o inversa (círculo azul) [Fig. 58].

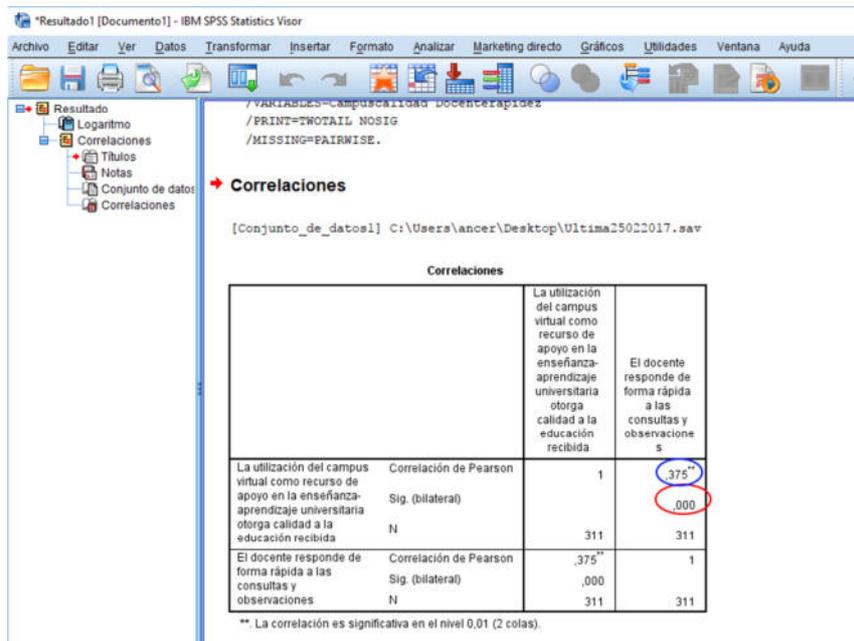


Fig. 58. Resultados de las correlaciones

## 8. Análisis de regresión

La regresión complementa la interdependencia entre variables. En lugar de ver cómo correlacionan dos variables (Pearson), lo que vemos es cómo un grupo de variables influyen en la que nosotros analizamos. En este caso se habla de variable criterio (el equivalente a la variable dependiente) y de variables predictoras (equivalentes a las variables independientes).

La secuencia a seguir es: analizar-regresión-lineal [Fig. 59].

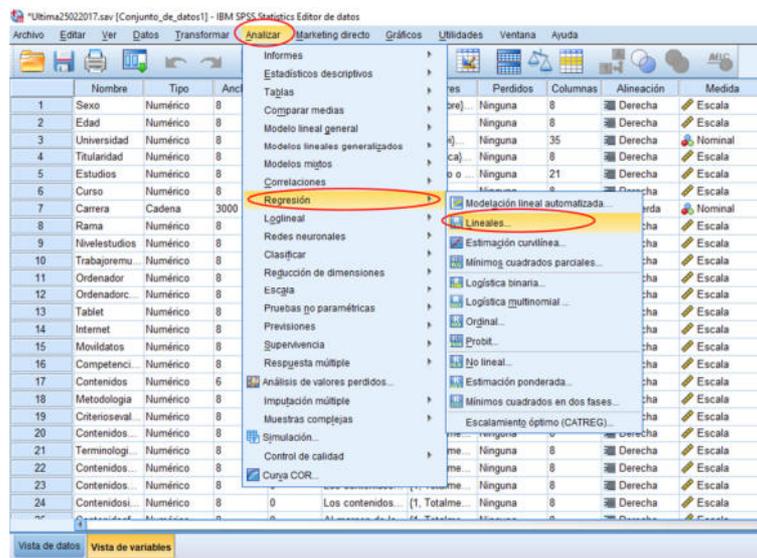


Fig. 59. Regresión lineal

Una vez en la ventana correspondiente, señalamos la variable criterio (dependiente), las variables predictoras (independientes) y vemos resultados [Fig. 60].

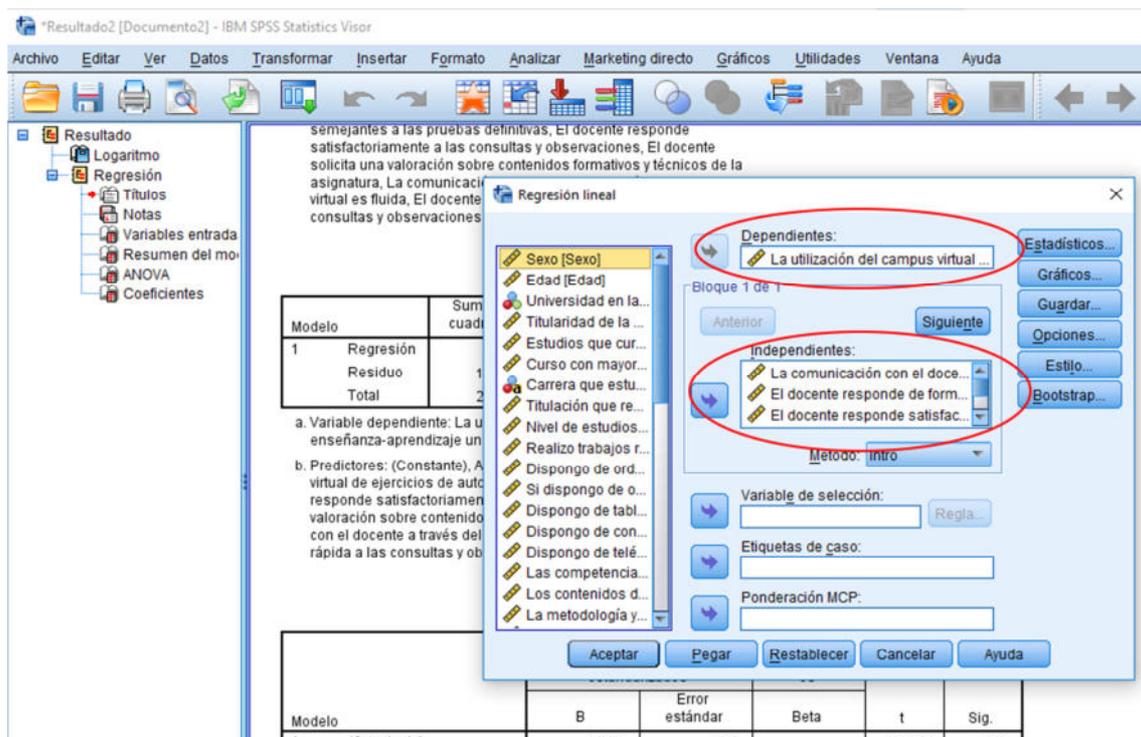


Fig. 60. Selección de variables de regresión

En el cuadro de resultados de ANOVA, vemos si hay significatividad y si la hay, vemos que hay una regresión aceptable [Fig. 61].

**ANOVA<sup>a</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	54,743	5	10,949	18,268	,000 <sup>b</sup>
	Residuo	182,794	305	,599		
	Total	237,537	310			

- a. Variable dependiente: La utilización del campus virtual como recurso de apoyo en la enseñanza-aprendizaje universitaria otorga calidad a la educación recibida
- b. Predictores: (Constante), Antes de realizar la evaluación final, dispongo en el campus virtual de ejercicios de autoevaluación semejantes a las pruebas definitivas, El docente responde satisfactoriamente a las consultas y observaciones, El docente solicita una valoración sobre contenidos formativos y técnicos de la asignatura, La comunicación con el docente a través del campus virtual es fluida, El docente responde de forma rápida a las consultas y observaciones

Fig. 61. Cuadro ANOVA de regresión

La forma de citarla en un artículo científico será:

$$F(5, 305)=18.268, p<.001$$

Si hay significatividad podemos seguir con el análisis. Vemos el cuadro del resumen del modelo y nos centramos en el valor de  $R^2$  ajustado [Fig. 62].

**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de a estimación
1	,480 <sup>a</sup>	,230	,218	,774

a. Predictores: (Constante), Antes de realizar la evaluación final, dispongo en el campus virtual de ejercicios de autoevaluación semejantes a las pruebas definitivas, El docente responde satisfactoriamente a las consultas y observaciones, El docente solicita una valoración sobre contenidos formativos y técnicos de la asignatura, La comunicación con el docente a través del campus virtual es fluida, El docente responde de forma rápida a las consultas y observaciones

Fig. 62. Valor  $R^2$  ajustado del análisis de regresión

El valor .218 implica que el conjunto de variables seleccionadas consigue explicar el 21,8% de la variable criterio. Posteriormente vamos al cuadro de coeficientes y podemos ver las variables por separado. Vemos las que son significativas y el valor Beta que nos dice que alguna variable es más importante que otra [Fig. 63].

\*Resultado2 [Documento2] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

Logaritmo Regresión Titulos Notas Variables entrada Resumen del mo ANOVA Coeficientes

rápida a las consultas y observaciones

**Coefficientes<sup>a</sup>**

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error estándar	Beta		
1	(Constante)	1,717	,160		10,763	,000
	La comunicación con el docente a través del campus virtual es fluida	,211	,076	,203	2,768	,006
	El docente responde de forma rápida a las consultas y observaciones	,027	,086	,025	,308	,758
	El docente responde satisfactoriamente a las consultas y observaciones	,344	,084	,325	4,118	,000
	El docente solicita una valoración sobre contenidos formativos y técnicos de la asignatura	,003	,066	,003	,044	,965
	Antes de realizar la evaluación final, dispongo en el campus virtual de ejercicios de autoevaluación semejantes a las pruebas definitivas	-,054	,058	-,053	-,932	,352

a. Variable dependiente: La utilización del campus virtual como recurso de apoyo en la enseñanza-aprendizaje universitaria otorga calidad a la educación recibida

Fig. 63. Tabla de coeficientes

En este caso podemos ver que el modelo queda establecido en dos variables. De este modo el uso del campus como elemento de calidad en la enseñanza, dependerá de dos variables fundamentalmente, si el docente responde satisfactoriamente a las consultas y observaciones y si la comunicación con el docente es fluida.

Otro método para hacerlo, es seleccionar en la ventana de regresión lineal el método por pasos, que nos va a ordenar ya las variables por importancia, eliminando aquellas que no tengan valor para el modelo [Fig. 64].

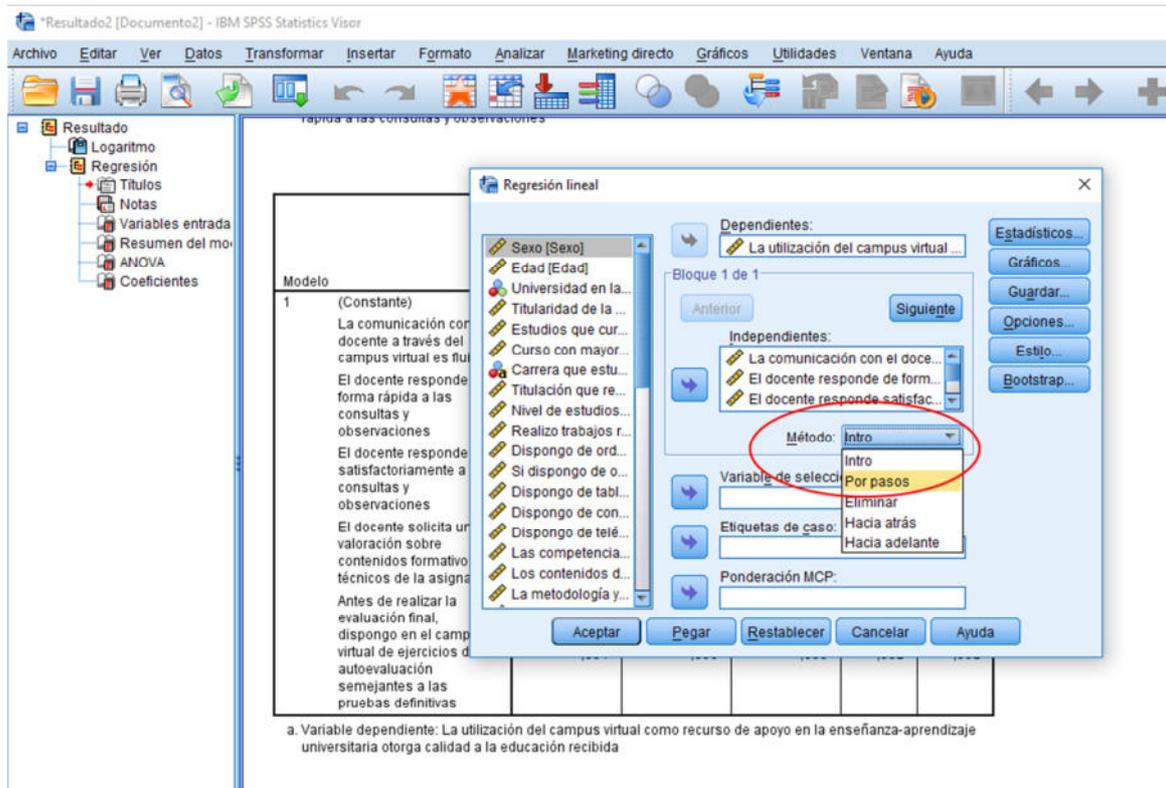


Fig. 64. Método por pasos para el análisis de regresión

Vemos que el resultado es el mismo, con las mismas variables predictoras significativas, la primera explicando el 20,3% y la segunda el 2% restante del 22,3% total que explica el modelo [Fig. 65].

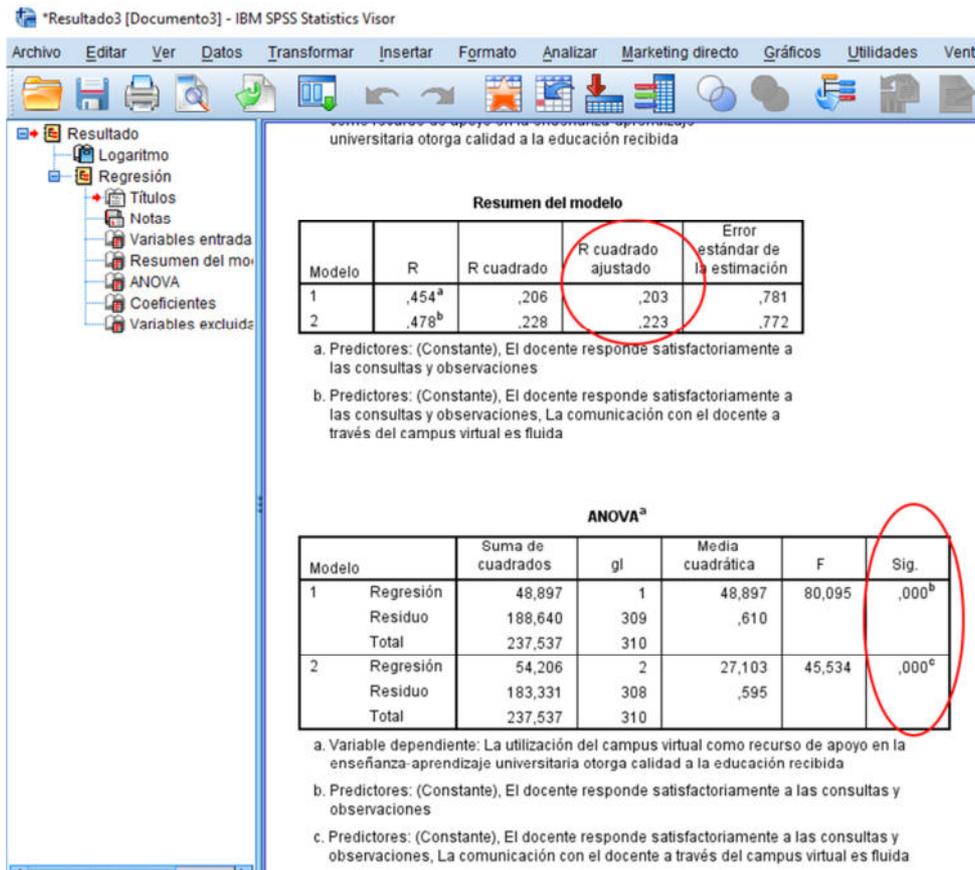


Fig. 65. R<sup>2</sup> ajustado del método por pasos

En la tabla de coeficientes podemos ver los valores para las dos variables [Fig. 66].

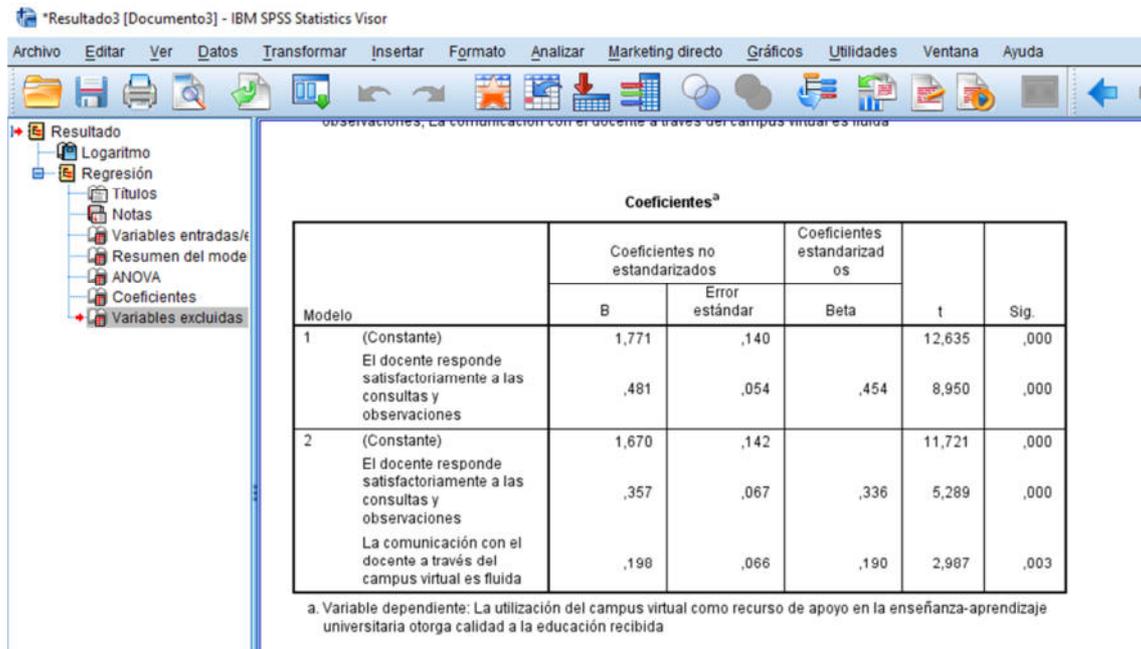


Fig. 66. Tabla de coeficientes del modelo por pasos

Si las primeras variables explican un amplio porcentaje del total, puede ser conveniente eliminar el resto ya que aportan muy poco. Cuando hacemos pasos sucesivos el valor de F es el último, que explica el total [Fig. 67].

**ANOVA<sup>a</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	48,897	1	48,897	80,095	,000 <sup>b</sup>
	Residuo	188,640	309	,610		
	Total	237,537	310			
2	Regresión	54,206	2	27,103	45,534	,000 <sup>c</sup>
	Residuo	183,331	308	,595		
	Total	237,537	310			

a. Variable dependiente: La utilización del campus virtual como recurso de apoyo en la enseñanza-aprendizaje universitaria otorga calidad a la educación recibida

b. Predictores: (Constante), El docente responde satisfactoriamente a las consultas y observaciones

Fig. 67. Valor F del modelo por método de pasos

La forma de exponerlo en un artículo científico sería la misma que anteriormente.

$$F(2,308)=45,534, p<.001.$$

A partir del 20% se puede aceptar el modelo de regresión como una buena explicación. En la regresión logística binaria la variable criterio no tiene por qué ser cuantitativa, sino que puede ser una cualitativa con dos valores, funcionando prácticamente igual que la lineal.