



Universidad
Norbert Wiener

Powered by Arizona State University

Manual de procesamiento estadístico para la investigación con SPSS

Oriana Rivera Lozada de Bonilla
Judith Soledad Yangali Vicente
José Luis Rodríguez López
Miguel Angel Ipanaqué Zapata

Month	Sales 1	Sales 2	Sales 3	Sales 4
Jan	32453	36754	50565	45674
Feb	34677	43456	53211	56333
Mar	56755	53422	58423	68211
Apr	54568	69424		





Universidad
Norbert Wiener

Powered by Arizona State University

Manual de procesamiento estadístico para la investigación con SPSS

Oriana Rivera Lozada de Bonilla
Judith Soledad Yangali Vicente
José Luis Rodríguez López
Miguel Angel Ipanaqué Zapata



**Universidad
Norbert Wiener**

Powered by **Arizona State University**

Manual de procesamiento estadístico para la investigación con SPSS

Oriana Rivera Lozada de Bonilla
Judith Soledad Yangali Vicente
José Luis Rodríguez López
Miguel Angel Ipanaqué Zapata

© Universidad Privada Norbert Wiener S. A.
Jr. Larrabure y Unanue 110, urb. Santa Beatriz. Lima, Perú
(01) 706 5555
www.uwiener.edu.pe
fondoeditorial@uwiener.edu.pe

Primera edición digital: febrero de 2023
Depósito Legal N.º: 2023-01749
ISBN 978-612-49230-0-5
DOI: <https://doi.org/10.37768/unw.vri.0011>

Edición general: Fondo Editorial de la Universidad Privada Norbert Wiener
Editor: Miguel Ruiz Effio
Asistente editorial: Maria Luz Ortiz
Diseño y diagramación: Rodolfo Loyola Mejía
Cuidado de textos: Juan Carlos Bondy

Se prohíbe la reproducción total o parcial de este libro sin autorización expresa de la Universidad Privada Norbert Wiener S. A.

Contenido

PRESENTACIÓN	7
SECCIÓN 1. INTRODUCCIÓN	9
¿Para qué sirve este manual?	9
¿Qué encontrará el lector?	10
SECCIÓN 2. CONOCIMIENTOS BÁSICOS EN ESTADÍSTICA	11
Estadística	11
Estadística descriptiva	11
Estadística inferencial	11
Estadística paramétrica	11
Estadística no paramétrica	11
Población	12
Muestra	12
Muestreo	12
Coefficiente de correlación	12
Significancia estadística	12
Variable	12
SECCIÓN 3. PROCESAMIENTO DE DATOS	15
Visualización de base de datos	15
Opciones de visualización de la información (vista de datos y vista de variables)	20
SECCIÓN 4. ANÁLISIS DE CONFIABILIDAD Y NORMALIDAD	25
Confiabilidad de un instrumento	25
Prueba de normalidad	28

SECCIÓN 5. ANÁLISIS DESCRIPTIVO	35
Escala valorativa de las variables	36
Tablas de frecuencia con SPSS	38
Tablas de contingencia o cruce de variables	46
SECCIÓN 6. ANÁLISIS INFERENCIAL	49
Análisis estadístico de estudios correlacionales	49
Análisis estadístico de estudios preexperimentales	62
Análisis estadístico de estudios cuasiexperimentales	71
Análisis estadístico de estudios comparativos: 2 grupos	79
Análisis estadístico de estudios comparativos: más de 2 grupos	84
Análisis estadístico de estudios correlacionales-causales	89
Análisis estadístico de estudios correlacionales-causales con variables cuantitativas a través de modelos de regresión	101
SECCIÓN 7. INTRODUCCIÓN A LOS MODELOS MULTINIVEL	113
Regresión lineal multinivel	113
SECCIÓN 8. GRÁFICO PARA VARIABLES CUALITATIVAS	121
Gráfico para variables cualitativas	121
REFERENCIAS BIBLIOGRÁFICAS	131
SOBRE LOS AUTORES	133

Presentación

La educación de los estudiantes universitarios comprende varias etapas de enseñanza teórica y práctica, con el objetivo de que se conviertan en profesionales con actitudes y habilidades para afrontar la realidad de sus especialidades. Al finalizar los semestres de educación, se solicita a los estudiantes un trabajo de investigación en formato de tesis o de artículo científico, para alcanzar el grado de bachiller o el título profesional.

Lo anterior representa una tarea muy complicada para muchos estudiantes: realizar dichos trabajos de investigación les toma más tiempo de lo establecido y, en algunos en casos, desertan del certificado profesional por no saber cómo afrontar este reto educativo. Dentro de la estructura de estas investigaciones, la sección con mayor complicación es la metodología de la investigación y el análisis estadístico, dado que demanda una idea básica sobre técnicas y plataformas estadísticas para responder a sus objetivos.

El presente manual abarca temáticas realizadas dentro del programa estadístico SPSS —de uso frecuente—, que, como se menciona en los próximos capítulos, es una buena plataforma para estudiantes o profesionales que se inician dentro del rubro de procesamiento y análisis de datos. Asimismo, dentro de cada capítulo se les ofrece una explicación general de las técnicas estadísticas, para conocer cuándo y cómo utilizarlas.

Por lo anteriormente expuesto y conociendo la importancia de la sección de metodología y estadística como el núcleo de cualquier trabajo de investigación, presentamos a la comunidad educativa este manual, con el propósito de que sirva como herramienta para generar sus resultados particulares.

Los autores

Sección 1

Introducción

La estadística en temas de investigación es fundamental dentro de la enseñanza en educación superior, prioritariamente para la elaboración de tesis de grado o trabajos de la especialidad, a fin de responder a un determinado problema específico. Para ello, los estudiantes necesitan una capacitación previa sobre el uso de técnicas estadísticas y la presentación correcta de resultados.

Al problema existente de la dificultad en el conocimiento sobre temas en estadística, se suma el complicado procesamiento de datos y la variedad de plataformas, *softwares* o programas de análisis estadísticos, como SPSS, Minitab, Stata, RStudio, entre otros, que exigen al investigador tener conocimientos de generación de comando para obtener los resultados esperados.

En contraste con estas plataformas, destaca, por su mayor facilidad de uso, el programa IBM SPSS Statistics, conocido comúnmente como SPSS, en sus distintas versiones. La principal ventaja de este programa es que brinda la opción de trabajo a nivel de usuario y programador. La primera opción es la más adecuada en una persona que se inicia en su uso, porque permite utilizar el cursor del *mouse* para ubicar las opciones de procesamiento y análisis de datos a través de un clic.

Por todo lo mencionado, el presente equipo de investigación busca abordar de forma introductoria, fácil y con casos prácticos, el procesamiento y análisis de datos usando SPSS.

¿Para qué sirve este manual?

El presente manual sirve de apoyo o guía al investigador para lograr sus resultados de forma autónoma y eficiente, con ejemplos de casos prácticos y bases de datos reales que usan el SPSS.

¿Qué encontrará el lector?

- El paso a paso para obtener resultados, con ejemplos replicables a sus trabajos de investigación.
- Capítulos o secciones con cada tipo estadístico, casos prácticos y un respaldo teórico conciso para orientar sobre la función y la utilidad del proceso estadístico.
- Casos prácticos provenientes de datos reales dentro de la diversidad de especialidad (por ejemplo, salud y educación).
- Conocimientos básicos de estadística en los dos primeros capítulos, y la guía para importar una base de datos.

Sección 2

Conocimientos básicos en estadística

Estadística

Disciplina que trata sobre recolección, clasificación, ordenamiento y organización de datos para su procesamiento, análisis, presentación e interpretación, lo que permitirá tomar decisiones respecto de algún fenómeno o tema de interés.

Estadística descriptiva

Es un conjunto de técnicas matemáticas mediante las cuales se obtienen, describen, organizan y presentan un conjunto de datos con el apoyo de tablas y gráficos. En la estadística descriptiva no se extraen inferencias sobre la población.

Estadística inferencial

Es la técnica mediante la cual se estima el comportamiento de una población a partir de una muestra representativa. La estadística inferencial permite la contrastación de hipótesis a través de la aplicación de pruebas estadísticas, que dependen del tipo de variables y de la naturaleza de los datos.

Estadística paramétrica

Herramienta de procesamiento de datos que se usa cuando los datos tienen una distribución normal y una homogeneidad de las varianzas. Entre las pruebas paramétricas más utilizadas tenemos a la prueba T, la regresión lineal, el coeficiente de Pearson, el ANOVA, etc. (Sánchez *et al.*, 2018).

Estadística no paramétrica

Herramienta de procesamiento de datos utilizada cuando los datos no siguen una distribución normal. Las pruebas no paramétricas más utilizadas son la U de Mann-Whitney, el chi cuadrado, el coeficiente de Spearman, la prueba de Wilcoxon, entre otras (Sánchez *et al.*, 2018).

Población

Está constituida por la totalidad de elementos o eventos con una característica en común, la cual es de interés para algún estudio o experimento que se desea realizar.

Muestra

Es una parte de la población que es extraída mediante técnicas o procedimientos específicos.

Muestreo

Es una técnica que se emplea para seleccionar un tamaño de muestra. El muestreo puede ser probabilístico (todos los elementos tienen la misma probabilidad de ser elegidos, es decir, al utilizar una técnica estadística) o no probabilístico (la selección es intencional, es decir, el tamaño lo elige el investigador según la experiencia o antecedente teórico).

Coefficiente de correlación

Valor que cuantifica la intensidad o el grado de relación entre las variables dentro de un análisis de correlación. Su símbolo es r (correlación de Pearson) o ρ (correlación de Spearman). La única diferencia es que el primero se emplea para una muestra con distribución normal, y el segundo, para una muestra con distribución no normal. El rango en que varía el coeficiente de correlación va desde -1 a 1 y es interpretable. Si el valor se acerca a -1, la relación es inversa entre las variables; si se acerca a 1, la relación es directa; y si el valor es igual a 0, no existe relación.

Significancia estadística

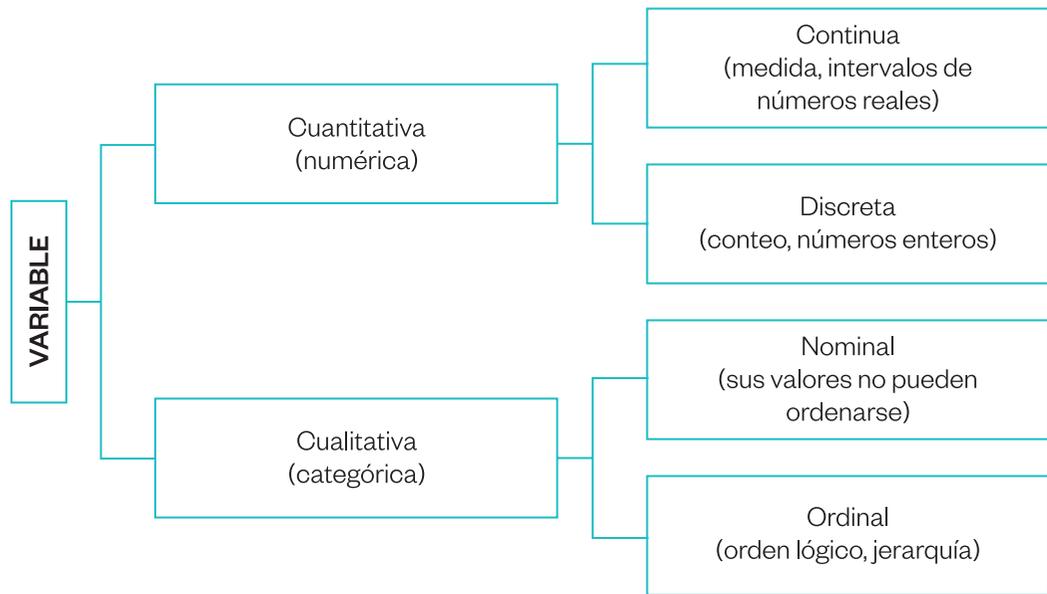
Es la medida de fiabilidad o probabilidad que consiste en cometer un error al rechazar una hipótesis nula cuando esta es cierta. La utilidad de este indicador estadístico se vincula para probar o evaluar asociaciones, contraste de grupos, obtención de muestra, entre otras situaciones. El símbolo de la significancia estadística es α . Los estándares de valores de significancia son 0,001, 0,05 y 0,10. El segundo es el más utilizado.

Variable

Es una característica o cualidad de un objeto de estudio que es susceptible de medición.

Las variables se pueden clasificar de la siguiente manera:

Según su naturaleza



VARIABLES CUANTITATIVAS: su característica principal es que son variables numéricas.

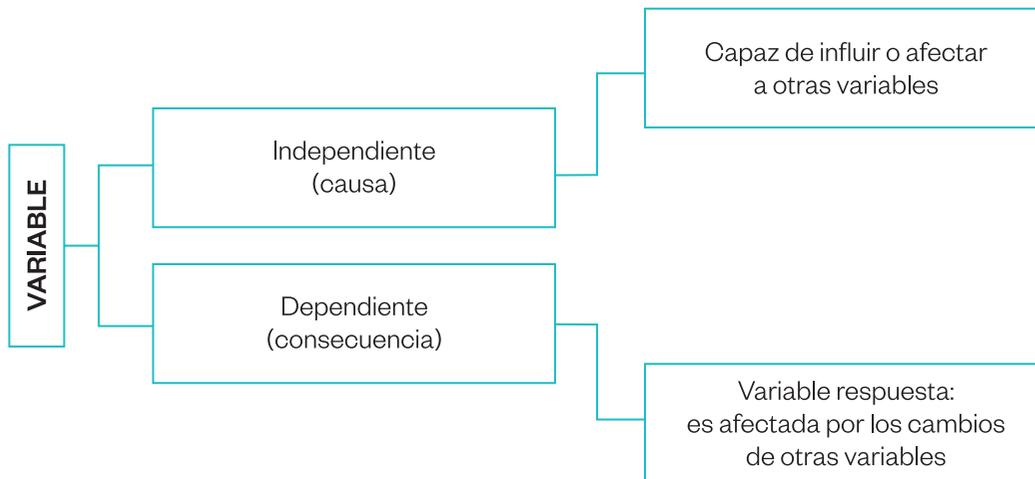
- **Continuas:** son variables con valores numéricos con decimales. Ejemplos: el peso y la talla de una persona, la latitud geográfica, etc.
- **Discretas:** son variables con valores numéricos y enteros. Ejemplos: la edad de una persona según su número de años, el número de equipos de laboratorio al mes, la cantidad de personas dentro del hogar, etc.

VARIABLES CUALITATIVAS: su característica principal es que responden a un determinado atributo de la unidad de análisis.

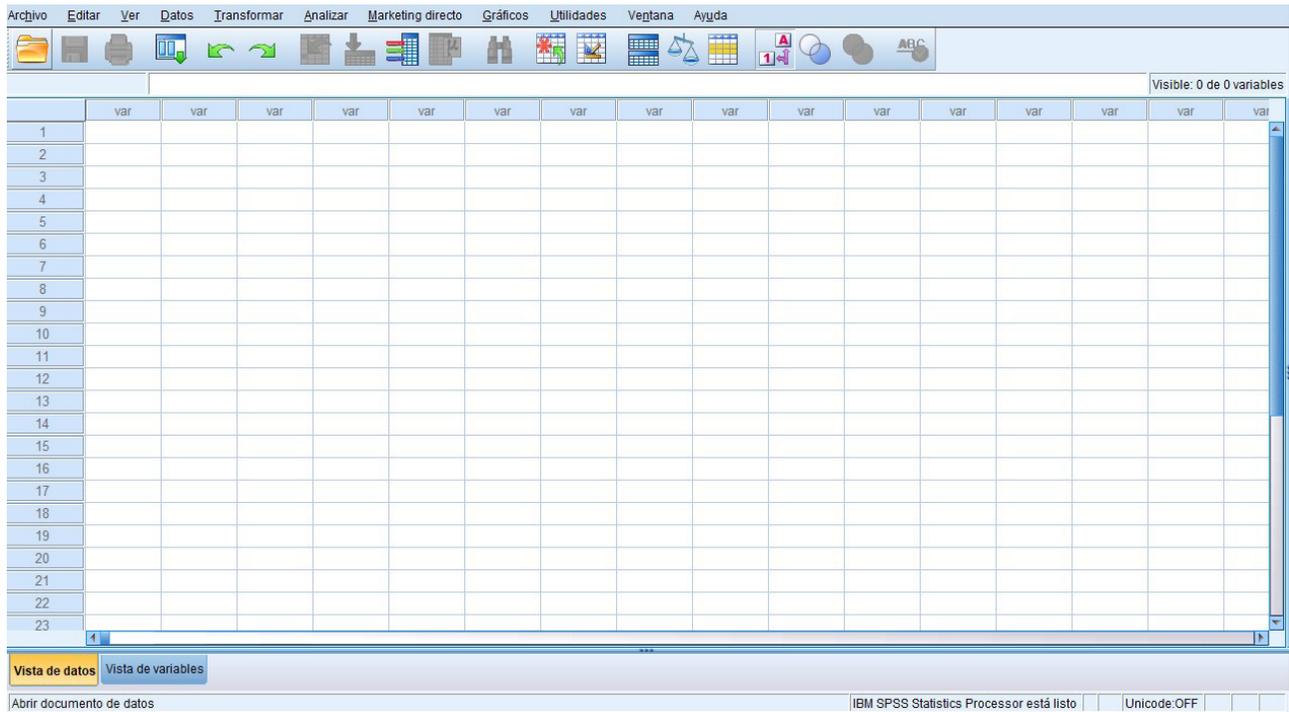
- **Nominales:** son variables que responden a un atributo sin un determinado orden. Ejemplos: el sexo de la persona, la presencia o ausencia de una determinada enfermedad o característica educativa, etc.
- **Ordinales:** son variables que responden a un atributo con un determinado orden. Ejemplos: el grado de instrucción, el nivel de conocimiento, el nivel socioeconómico, etc.

Según su posición respecto de otras variables

Este tipo de variables visualiza, en su mayoría de casos, relación entre dos o más variables, y el impacto de una variable sobre otra. Generalmente utilizan modelos de regresión (lineal, logística, multinivel, entre otros). Existe una gran diversidad de casos, según el tipo de variable dependiente. Ejemplos: la evaluación de la relación entre diversos factores que se asociaron al rendimiento académico, el impacto del bienestar psicológico en la depresión, etc.

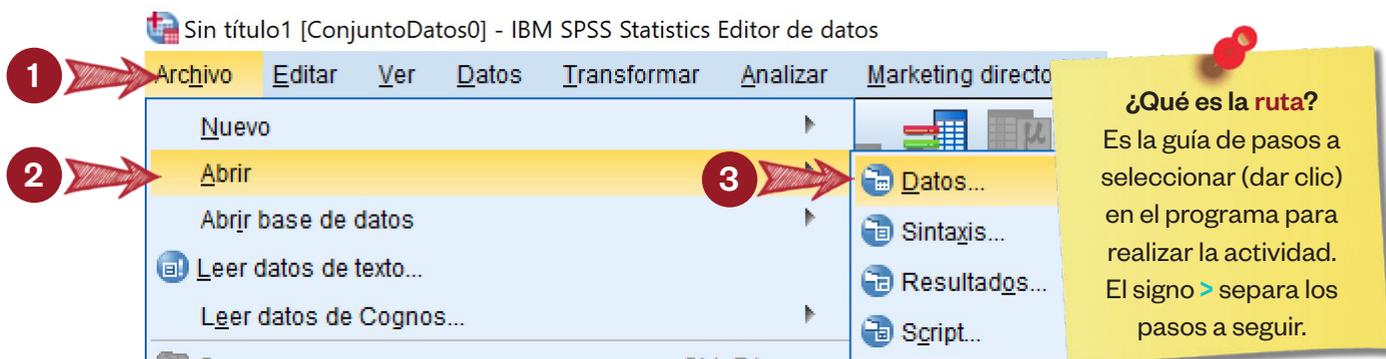


Segundo paso: abrir el programa SPSS. Podemos observar que su interfaz (plantilla) es muy parecida a la del Excel.



Ya dentro del programa SPSS, necesitamos visualizar los datos dentro de esta plantilla, por lo que debemos seguir la ruta.

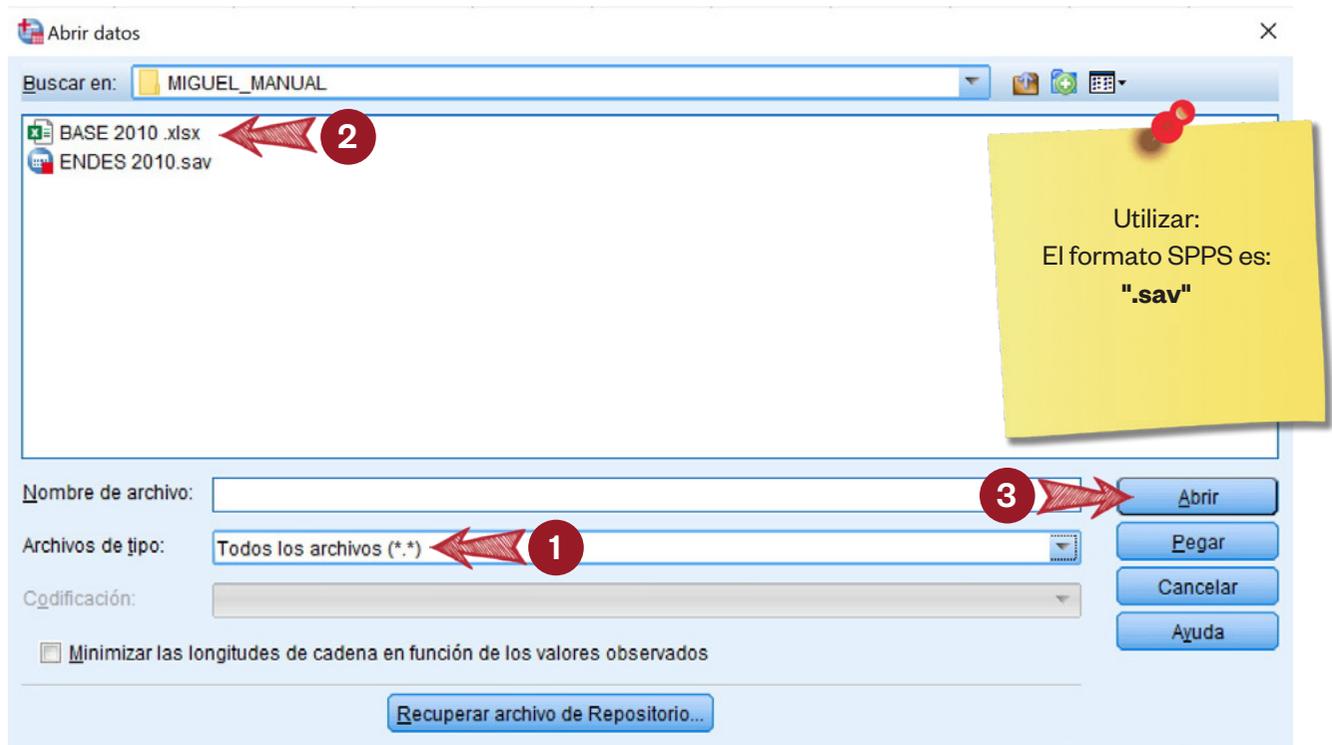
Ruta: Archivo > Abrir > Datos



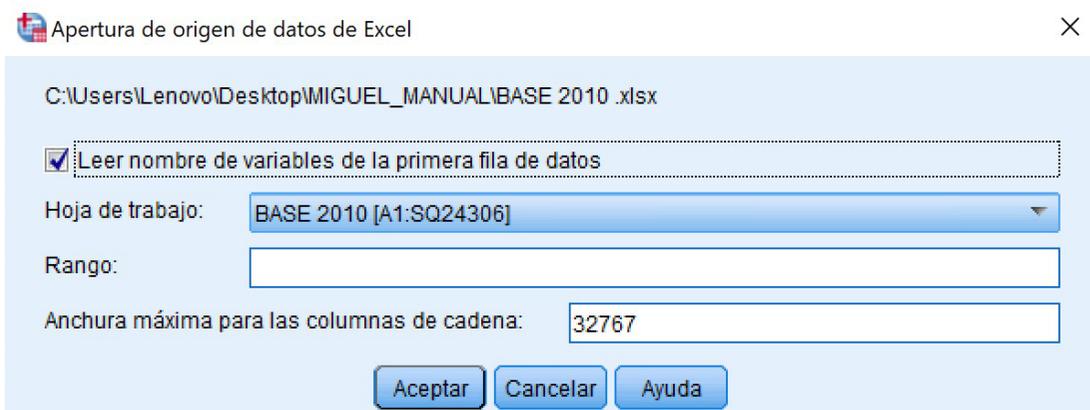
Luego de seguir la ruta mostrada, aparecerá la siguiente ventana por defecto en “Documentos”:



El archivo que buscamos no está en “Documentos”. Debemos ubicar dónde se guardó.



Seleccionamos el Excel que se quiere trabajar. Luego, aparecerá el rango de la base de datos.



En la hoja de trabajo aparecerán las hojas de Excel creadas. Debemos encontrar la hoja que estamos utilizando, que, en este ejemplo, es “BASE 2010”. Si deseamos seleccionar una muestra, hay que colocar el rango que se usará (es el espacio donde se ingresan los datos).

Recomendación: debemos seleccionar como primera fila el nombre de las variables, en la opción “Leer nombre de variables de la primera fila de datos”. A continuación, colocamos el rango desde el primer dato. En este caso, desde A2 (que inicia los nombres de variables), hasta F32 (si se quiere seleccionar hasta la muestra 32). Para seleccionar toda la muestra de la base de datos, la primera fila solo debe tener el nombre de variables. Luego, damos aceptar por *default*.

Resumen de conceptos y pasos:

- 1. Archivo tipo:** indica el formato de los archivos que deseamos que aparezcan. Para que aparezcan todos los archivos guardados en los distintos formatos debemos seleccionar “Todos los archivos”.
- 2. ENDES 2010.xlsx:** archivo en el formato compatible al Excel.
- 3. Abrir:** opción para abrir el archivo seleccionado. Aparecerá la ventana que determinará el rango que seleccionaremos nuestra muestra o por *default*. Finalizamos dándole aceptar.

Al mostrar la base de datos que utilizaremos ya hemos completado el primer paso.

*Sin título2 [ConjuntoDatos1] - IBM SPSS Statistics Editor de datos

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

Visible: 511 de 511 variables

	HHID	CASEID	HV002	V003	HV004	HV005	HV006	HV008	HV009	HV010
1	0002011	0002011 2	11,0	2,0	2,0	322759,0	4,0	.	3,0	.
2	0002015	0002015 2	15,0	2,0	2,0	322759,0	4,0	1324,0	4,0	.
3	0002019	0002019 2	19,0	2,0	2,0	322759,0	4,0	1324,0	5,0	.
4	0002036	0002036 2	36,0	2,0	2,0	322759,0	4,0	1324,0	2,0	.
5	0002040	0002040 1	40,0	1,0	2,0	322759,0	4,0	1324,0	4,0	.
6	0002040	0002040 3	40,0	3,0	2,0	322759,0	4,0	.	4,0	.
7	0002044	0002044 3	44,0	3,0	2,0	322759,0	4,0	.	5,0	.
8	0002044	0002044 2	44,0	2,0	2,0	322759,0	4,0	.	5,0	.
9	0002052	0002052 2	52,0	2,0	2,0	322759,0	4,0	1324,0	5,0	.
10	0002052	0002052 4	52,0	4,0	2,0	322759,0	4,0	.	5,0	.
11	0002052	0002052 3	52,0	3,0	2,0	322759,0	4,0	.	5,0	.
12	0002060	0002060 2	60,0	2,0	2,0	322759,0	4,0	1324,0	3,0	.
13	0002064	0002064 2	64,0	2,0	2,0	322759,0	4,0	1324,0	4,0	.
14	0002069	0002069 2	69,0	2,0	2,0	322759,0	4,0	1324,0	2,0	.
15	0002073	0002073 1	73,0	1,0	2,0	322759,0	5,0	1325,0	5,0	.
16	0002073	0002073 5	73,0	5,0	2,0	322759,0	5,0	.	5,0	.
17	0002077	0002077 2	77,0	2,0	2,0	322759,0	4,0	1324,0	4,0	.
18	0002081	0002081 2	81,0	2,0	2,0	322759,0	4,0	1324,0	3,0	.
19	0002089	0002089 1	89,0	1,0	2,0	322759,0	4,0	1324,0	2,0	.
20	0002105	0002105 2	105,0	2,0	2,0	322759,0	4,0	.	3,0	.
21	0002114	0002114 2	114,0	2,0	2,0	322759,0	4,0	1324,0	3,0	.
22	0003006	0003006 2	6,0	2,0	3,0	412671,0	4,0	.	4,0	.
23	0003014	0003014 2	14,0	2,0	3,0	412671,0	4,0	.	5,0	.
24	0003022	0003022 4	22,0	4,0	3,0	412671,0	4,0	.	5,0	.

Vista de datos Vista de variables

Opciones de visualización de la información (vista de datos y vista de variables)

Las opciones para la visualización y edición de la información dentro de la ventana del SPSS son de gran ayuda, ya que cada una tiene determinada función. Están enmarcadas en la parte inferior izquierda de la imagen y se seleccionan con un clic.

Vista de datos

Esta opción, que aparece como ventana principal al iniciar el SPSS, sirve para visualizar y editar los valores dentro de cada variable (por ejemplo, el valor de la variable HV002 en la fila 3 es 19). Esta ventana nos recuerda a la plantilla Excel, porque podemos ingresar, cambiar o eliminar datos.

The screenshot shows the IBM SPSS Statistics Editor de datos window. The title bar reads "ENDES 2010.sav [ConjuntoDatos1] - IBM SPSS Statistics Editor de datos". The menu bar includes Archivo, Editar, Ver, Datos, Transformar, Analizar, Marketing directo, Gráficos, Utilidades, Ventana, and Ayuda. The toolbar contains various icons for file operations, data manipulation, and analysis. The main area displays a data table with 25 rows and 14 columns. The columns are labeled HHID, V003, HV002, HV004, HV005, HV006, HV008, HV009, HV010, HV012, HV013, HV014, HV015, and HV017. The data values are numerical, with some missing values represented by dots. The status bar at the bottom right indicates "Visible: 506 de 506 variat". At the bottom left, there are two buttons: "Vista de datos" and "Vista de variables". A red arrow points to the "Vista de datos" button.

	HHID	V003	HV002	HV004	HV005	HV006	HV008	HV009	HV010	HV012	HV013	HV014	HV015	HV017
1	0002011	2	11	2	322759	4	.	3	1	3	3	0	Completa	
2	0002015	2	15	2	322759	4	1324	4	1	4	4	1	Completa	
3	0002019	2	19	2	322759	4	1324	5	1	5	5	1	Completa	
4	0002036	2	36	2	322759	4	1324	2	1	2	2	0	Completa	
5	0002040	1	40	2	322759	4	1324	4	2	2	4	0	Completa	
6	0002040	3	40	2	322759	4	.	4	2	2	4	0	Completa	
7	0002044	3	44	2	322759	4	.	5	2	5	5	1	Completa	
8	0002044	2	44	2	322759	4	.	5	2	5	5	1	Completa	
9	0002052	2	52	2	322759	4	1324	5	3	5	4	0	Completa	
10	0002052	4	52	2	322759	4	.	5	3	5	4	0	Completa	
11	0002052	3	52	2	322759	4	.	5	3	5	4	0	Completa	
12	0002060	2	60	2	322759	4	1324	3	1	3	2	0	Completa	
13	0002064	2	64	2	322759	4	1324	4	1	4	3	0	Completa	
14	0002069	2	69	2	322759	4	1324	2	1	2	2	0	Completa	
15	0002073	1	73	2	322759	5	1325	5	2	4	5	1	Completa	
16	0002073	5	73	2	322759	5	.	5	2	4	5	1	Completa	
17	0002077	2	77	2	322759	4	1324	4	1	4	3	1	Completa	
18	0002081	2	81	2	322759	4	1324	3	1	3	3	1	Completa	
19	0002089	1	89	2	322759	4	1324	2	1	2	2	0	Completa	
20	0002105	2	105	2	322759	4	.	3	1	3	3	1	Completa	
21	0002114	2	114	2	322759	4	1324	3	1	3	3	0	Completa	
22	0003006	2	6	3	412671	4	.	4	1	4	4	1	Completa	
23	0003014	2	14	3	412671	4	.	5	1	5	5	0	Completa	
24	0003022	4	22	3	412671	4	.	5	1	5	5	1	Completa	
25	0003034	1	34	3	412671	3	1323	3	1	3	3	1	Completa	

Vista de variables

Los nombres de cada columna (variables) dentro de “Vista de datos” aparecen como listado de variables y sus características. Al seleccionar la opción “Vista de variables” (ver imagen), se abre una ventana que nos brinda opciones para editar las características de las variables a utilizar.

	1 Nombre	2 Tipo	Anchura	Decimales	3 Etiqueta	4 Valores	5 perdidos	Columnas	Alineación	6 Medida	Rol
1	HVID	Cadena	12	0	Identificación d...	Ninguno	Ninguno	12	Izquierda	Nominal	Entrada
2	V003	Numérico	8	0	Número de líne...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
3	HV002	Numérico	8	0	Número de hogar	Ninguno	Ninguno	8	Derecha	Escala	Entrada
4	HV004	Numérico	8	0	Ultima unidad d...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
5	HV005	Numérico	8	0	Peso de la mue...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
6	HV006	Numérico	8	0	Mes de entrevista	Ninguno	Ninguno	8	Derecha	Escala	Entrada
7	HV008	Numérico	8	0	Fecha de entre...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
8	HV009	Numérico	8	0	Número de mie...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
9	HV010	Numérico	8	0	Número mujere...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
10	HV012	Numérico	8	0	Número de mie...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
11	HV013	Numérico	8	0	Número de mie...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
12	HV014	Numérico	8	0	Número de niño...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
13	HV015	Numérico	8	0	Resultado de la...	{1, Complet...	Ninguno	8	Derecha	Escala	Entrada
14	HV017	Numérico	8	0	Número de visitas	Ninguno	Ninguno	8	Derecha	Nominal	Entrada
15	HV021	Numérico	8	0	Unidad primaria...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
16	HV022	Numérico	8	0	Número de estr...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
17	HV024	Numérico	8	0	Región	{1, Amazon...	Ninguno	8	Derecha	Nominal	Entrada
18	HV025	Numérico	8	0	Tipo de lugar d...	{1, Urbano}	Ninguno	8	Derecha	Nominal	Entrada
19	HV026	Numérico	8	0	Lugares de resi...	{0, Capital, ...	Ninguno	8	Derecha	Nominal	Entrada
20	HV040	Numérico	8	0	Altitud del cong...	Ninguno	Ninguno	8	Derecha	Escala	Entrada
21	HV101	Numérico	8	0	(P3) Relación d...	{1, Jefe}	Ninguno	8	Derecha	Nominal	Entrada
22	HV102	Numérico	8	0	(P4) Residente ...	{0, No}	Ninguno	8	Derecha	Nominal	Entrada
23	HV103	Numérico	8	0	(P5) Durmió la ...	{0, No}	Ninguno	8	Derecha	Nominal	Entrada
24	HV104	Numérico	8	0	(P6) Sexo del ...	{1, Hombre}	Ninguno	8	Derecha	Nominal	Entrada
25	HV219	Numérico	8	0	(P6)Sexo del je...	{1, Hombre}	Ninguno	8	Derecha	Nominal	Entrada
26	HV105	Numérico	8	0	(P7) Edad del ...	{97, 97+}	Ninguno	8	Derecha	Escala	Entrada

Conceptos y pasos

1. **Nombre:** también es llamado “código” de variable. Se define como el nombre o identificador que se muestran en las columnas de “Vista de datos”. Esta opción en su edición no acepta espacios.

Consejo: podemos colocar guion bajo (_) para dar la idea de espacio entre palabras (por ejemplo, “Nivel_Educativo”).

2. **Tipo:** es el formato en el que se presentan los datos de la variable en “Vista de datos”. Tiene distintas opciones conforme se presenten: fecha, cadena (texto), numérico, etc.

Consejo: estos ejemplos nos ayudarán a elegir una característica dentro de esta opción, según los datos que debamos presentar: (i) 19/09/2020 = opción fecha, (ii) Rodríguez Rodríguez, Miguel = opción cadena, o (iii) 19 = opción numérica.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
22	HV102	Numérico	8	0	(P4) Residente habitual	{0, No}...	Ninguno	8	Derecha	Nominal	Entrada
23	HV103	Numérico	8	0	(P5) Durmió la noche anterior	{0, No}...	Ninguno	8	Derecha	Nominal	Entrada
24	HV104	Numérico	8	0	(P6) Sexo del miembro del hogar	{1, Hombre}...	Ninguno	8	Derecha	Nominal	Entrada
25	HV219	Numérico	8	0	(P9) Sexo del jefe del hogar	{1, Hombre}...	Ninguno	8	Derecha	Nominal	Entrada
26	HV105	Numérico	8	0					Derecha	Escala	Entrada
27	HV220	Numérico	8	0					Derecha	Escala	Entrada
28	HV115	Numérico	8	0					Derecha	Nominal	Entrada
29	HV116	Numérico	8	0					Derecha	Nominal	Entrada
30	SH11A	Numérico	8	0					Derecha	Nominal	Entrada
31	SH11B	Numérico	8	0					Derecha	Nominal	Entrada
32	SH11C	Numérico	8	0					Derecha	Nominal	Entrada
33	SH11D	Numérico	8	0					Derecha	Nominal	Entrada
34	SH11E	Numérico	8	0					Derecha	Nominal	Entrada
35	SH11Y	Numérico	8	0					Derecha	Nominal	Entrada
36	SH11Z	Numérico	8	0					Derecha	Nominal	Entrada
37	Grupo_Cob...	Numérico	8	0					Derecha	Nominal	Entrada
38	SH12	Numérico	8	0					Derecha	Nominal	Entrada
39	SH13	Numérico	8	0					Derecha	Nominal	Entrada
40	HV106	Numérico	8	0					Derecha	Ordinal	Entrada
41	NC_JH	Numérico	8	0	Nivel Educativo del Jefe del Hogar (Adicional)	{0, Sin nivel,...	Ninguno	8	Derecha	Ordinal	Entrada
42	HV107	Numérico	8	0	(P15) Mas alto año de educación alcanzado	{98, NS}...	Ninguno	8	Derecha	Escala	Entrada
43	HV110	Numérico	8	0	(P16) Miembros que estan aA^n en el colegio	{0, No}...	Ninguno	8	Derecha	Nominal	Entrada
44	HV121	Numérico	8	0	(P17) Miembros que asisten a la escuela du...	{0, No}...	Ninguno	8	Derecha	Nominal	Entrada
45	HV122	Numérico	8	0	(P18) Nivel de educación al que asiste en el...	{0, Sin nivel,...	Ninguno	8	Derecha	Ordinal	Entrada

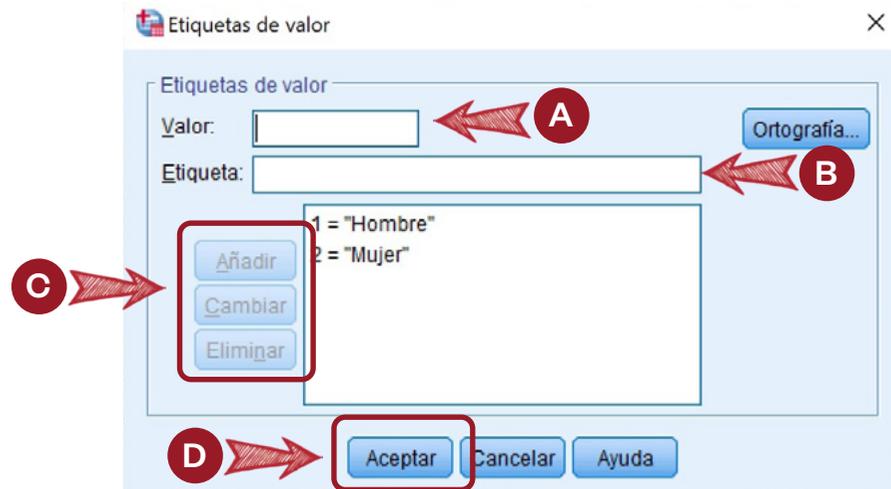
3. **Etiqueta:** esta característica otorga el nombre de las variables que deben aparecer en resultados. En la edición, esta opción sí acepta espacios entre cada palabra. Por ejemplo, “Nivel educativo de los jóvenes del Perú”.

4. **Valores:** esta característica permite visualizar y editar las categorías de las variables cualitativas, que en la vista de datos aparecen como valores numéricos. Por ejemplo, si tenemos la variable “sexo” y en vista de datos solo aparecen valores como 1 y 2, usamos estos números como una categoría de la variable. Así, 1 = Hombre y 2 = Mujer.

Consejo: es importante saber cuándo se trata de una variable cualitativa o cuantitativa (podemos revisar la definición en el glosario). Dentro de esta opción existen subsecciones que debemos conocer y llenar para generar las etiquetas (categorías), que presentamos a continuación.

A. Valor: colocamos valores numéricos (ver vista de datos) que referencian a cada categoría. En el ejemplo mencionado, los códigos a ingresar son los valores 1 y 2 para “Hombre” y “Mujer”, respectivamente.

B. Etiqueta: ingresamos el nombre de etiqueta o nombre de categoría. En el caso explicado, ingresamos 1 como valor numérico y en la etiqueta escribimos “Hombre”, y si el valor numérico es 2, escribimos “Mujer” en la etiqueta. Importante: tras ingresar el valor numérico, se coloca la etiqueta (el nombre de la categoría).



C.1. Añadir: nos ayuda a agregar el valor y la etiqueta ingresada. Por ejemplo, si ingresamos el valor 1 y la etiqueta “Hombre”, damos clic en “Añadir” para almacenar el vínculo generado entre el valor y la etiqueta, y así reportar la categoría. Así sucesivamente ingresamos todos los valores con sus respectivas etiquetas para generar las categorías de las variables.

C.2 Cambiar: sirve para corregir si hemos ingresado un valor errado o nos equivocamos en la redacción de la etiqueta, seleccionando la categoría generada.

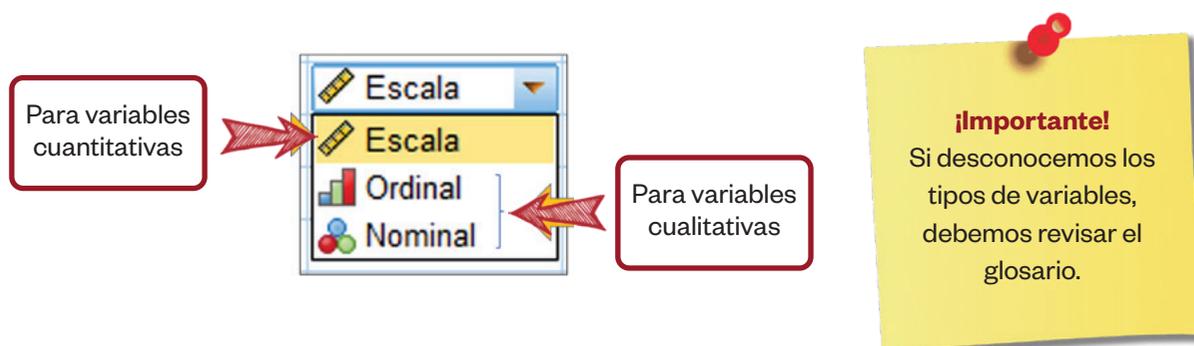
C.3 Borrar: permite eliminar categorías generadas.

D. Aceptar: sirve para confirmar las categorías generadas y la variable en “Vista de datos”. Ya no se visualizará como una variable cuantitativa, sino como una variable cualitativa (es decir, variable con categorías).

5. **Perdidos:** ayuda a identificar los valores faltantes o las celdas vacías de la base de datos en la ventana “Vista de datos”, al mostrarlos en la tabla de resultados con el código ingresado. Esa identificación también sirve para excluirlas o eliminarlas. Debemos considerar que los valores perdidos son respuestas que debieron hacerse, pero no se ingresaron por problemas con el encuestador o digitador.

Consejo: comúnmente se brinda un código de “999” o “9999” para identificar la celda vacía (llamada también valor perdido) a fin de excluirla o eliminarla.

6. **Medida:** permite identificar la variable según su tipo. Si es cuantitativa, la opción a elegir es “Escalar”, y si es cualitativa, existen dos opciones: (i) si es cualitativa sin regla de orden, la opción a elegir es “Nominal”, y (ii) si es cualitativa con regla de orden, la opción a elegir es “Ordinal”.



Llegado a este punto, podemos identificar o verificar que las variables cumplan con sus características básicas. Así, ya logramos la mitad del avance para familiarizarnos con el SPSS. Ahora podemos generar los análisis y resultados, que son lo más importante de nuestra investigación.

Organización de los datos

Antes de iniciar con el procesamiento de los datos, se recomienda organizarlos de manera clara y entendible. Por lo general, esta tarea se hace mediante la herramienta del Excel. La organización de datos consiste en la estructura físico-lógica que permite realizar operaciones computacionales como editar, actualizar, guardar, etc., respecto de un contenido de información.

Ejemplo 1

N.º	LIDERAZGO PEDAGÓGICO																	
	Desarrollo profesional docente						Gestión curricular								Objetivos educativos			
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18
1	2	1	1	3	1	1	2	1	1	3	1	2	1	3	3	1	1	2
2	1	3	1	1	3	2	1	3	2	1	3	1	3	1	1	3	1	1
3	1	2	1	1	2	2	1	2	2	1	2	1	2	3	1	2	1	1
4	2	3	1	2	3	1	2	3	2	2	3	2	3	1	2	3	3	2
5	3	3	1	3	3	3	3	3	2	3	3	3	3	1	3	3	1	3
6	1	3	1	1	3	1	1	3	2	1	3	1	3	3	1	3	2	1
7	2	3	1	2	3	1	2	3	2	2	3	2	3	1	2	3	1	2
8	3	2	1	3	2	1	3	2	2	3	2	3	2	1	3	2	1	3
9	2	2	1	2	2	3	2	2	3	2	2	2	2	2	2	2	2	2
10	1	2	1	1	2	1	1	2	3	1	2	1	2	1	1	2	1	1
11	2	2	1	2	2	1	2	2	3	2	2	2	2	1	2	2	1	2
12	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	3	1
13	2	2	1	2	2	3	2	2	1	2	2	2	2	1	2	2	2	2

Sección 4

Análisis de confiabilidad y normalidad

El análisis de confiabilidad es una técnica estadística importante para evaluar la consistencia interna del instrumento; es decir, si los ítems responden objetivamente a la variable de estudio.

Confiabilidad de un instrumento

La prueba de confiabilidad consiste en determinar el grado de precisión de un instrumento, esto es, si se aplica un instrumento repetidamente a un sujeto o produce iguales resultados. La confiabilidad permite “evaluar la homogeneidad de las preguntas; cuando las respuestas del cuestionario están formuladas en escala Likert o politómicas, los valores van desde (0) confiabilidad nula y (1) confiabilidad total” (Arispe *et al.*, 2020, p. 45).

Existen varios métodos para el cálculo de la confiabilidad de un instrumento. En esta ocasión utilizaremos el coeficiente del alfa de Cronbach.

Alfa de Cronbach

¿Qué es?

Es un índice utilizado para medir la consistencia interna de un instrumento.

¿Cómo se interpreta?

Cuanto más cerca se encuentre el valor de alfa a 1, mayor será la consistencia interna de los ítems analizados.

¿Que se recomienda?

Debemos considerar que tres es la cantidad mínima de ítems para una escala que explore un solo factor.

Alfa de Cronbach	
Rango	Escala
[0-0,2>	Muy baja
[0,2-0,4>	Baja
[0,4-0,6>	Moderada
[0,6-0,8>	Buena
[0,8-1>	Muy buena

Ruiz-Bolívar (2013)

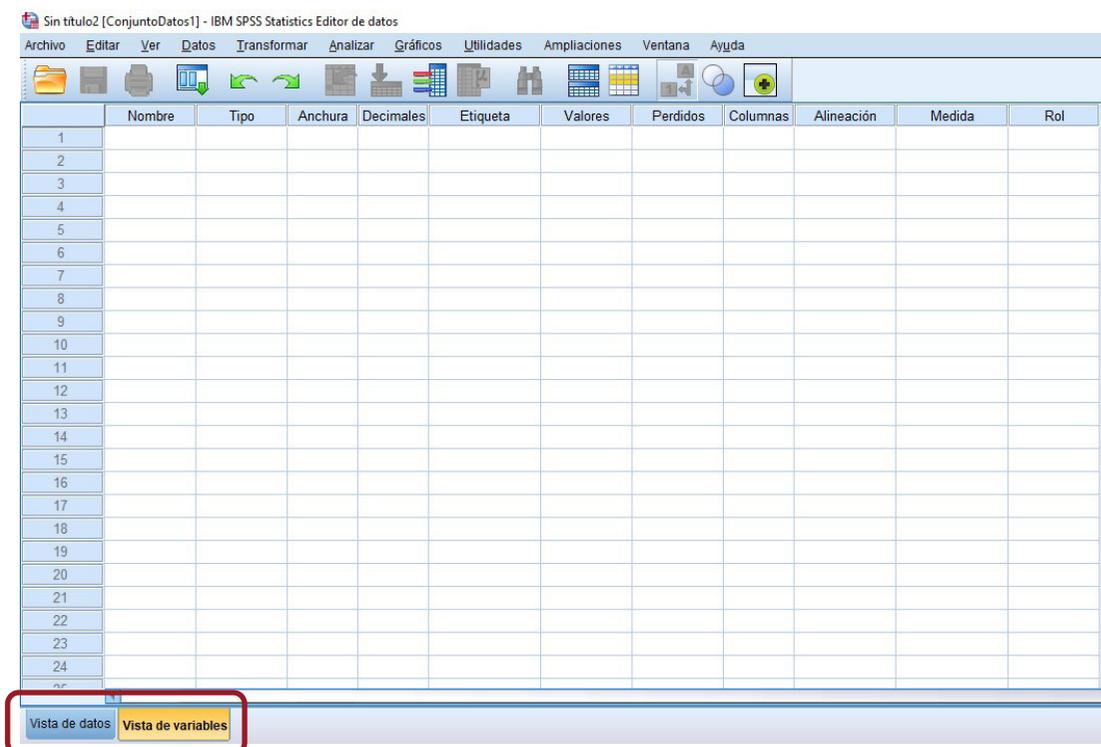
Análisis de confiabilidad con SPSS (alfa de Cronbach)

Ejemplo 2

Se pide determinar la confiabilidad del instrumento que mide el liderazgo pedagógico en los docentes de una institución educativa (se utilizarán los datos del ejemplo 1).

Procedimiento

Al ingresar al programa SPSS se observa el siguiente cuadro:



En la parte inferior izquierda disponemos de dos pestañas: vista de variables y vista de datos.

Para la prueba de confiabilidad ingresaremos los datos en la vista de datos, siguiendo los siguientes pasos:

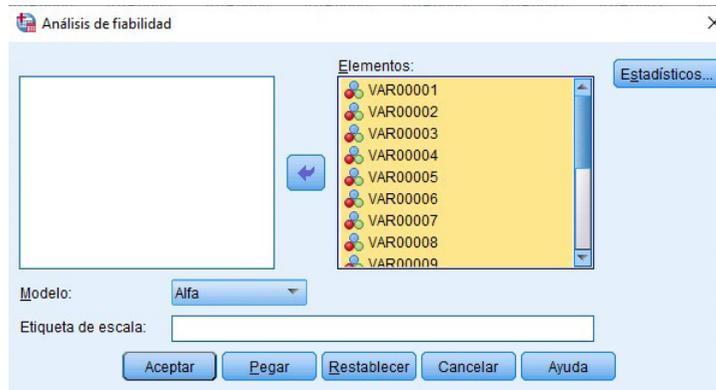
1. Ingreso de datos al SPSS.

	VAR00001	VAR00002	VAR00003	VAR00004	VAR00005	VAR00006	VAR00007	VAR00008	VAR00009	VAR00010	VAR00011	VAR00012	VAR00013	VAR00014	VAR00015	VAR00016	VAR00017	VAR00018	V
1	2,00	1,00	1,00	3,00	1,00	1,00	2,00	1,00	1,00	3,00	1,00	2,00	1,00	3,00	3,00	1,00	3,00	3,00	
2	1,00	3,00	1,00	1,00	3,00	2,00	1,00	3,00	2,00	1,00	3,00	1,00	3,00	1,00	3,00	1,00	3,00	1,00	
3	1,00	2,00	1,00	1,00	2,00	2,00	1,00	2,00	2,00	1,00	2,00	1,00	2,00	1,00	2,00	3,00	3,00	1,00	
4	2,00	3,00	1,00	2,00	3,00	1,00	2,00	3,00	2,00	2,00	3,00	2,00	3,00	3,00	1,00	2,00			
5	3,00	3,00	1,00	3,00	3,00	3,00	3,00	3,00	2,00	3,00	3,00	3,00	3,00	3,00	1,00	3,00			
6	1,00	3,00	1,00	1,00	3,00	1,00	1,00	3,00	2,00	1,00	3,00	1,00	3,00	3,00	3,00	1,00			
7	2,00	3,00	1,00	2,00	3,00	1,00	2,00	3,00	2,00	2,00	3,00	2,00	3,00	3,00	1,00	2,00			
8	3,00	2,00	1,00	3,00	2,00	1,00	3,00	2,00	2,00	3,00	2,00	3,00	2,00	2,00	1,00	3,00			
9	2,00	2,00	1,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00			
10	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	3,00	1,00	2,00	1,00	2,00	1,00	1,00	1,00			
11	2,00	2,00	1,00	2,00	2,00	1,00	2,00	2,00	3,00	2,00	2,00	2,00	2,00	2,00	1,00	2,00			
12	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	2,00	1,00			
13	2,00	2,00	1,00	2,00	2,00	3,00	2,00	2,00	1,00	2,00	2,00	2,00	2,00	2,00	1,00	2,00			
14	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	1,00	2,00	1,00	2,00	1,00			
15	2,00	3,00	1,00	2,00	3,00	1,00	2,00	3,00	2,00	2,00	3,00	2,00	3,00	3,00	3,00	2,00			
16	1,00	3,00	1,00	1,00	3,00	1,00	1,00	3,00	1,00	1,00	3,00	1,00	3,00	1,00	1,00	1,00			
17	2,00	2,00	1,00	2,00	2,00	2,00	2,00	2,00	1,00	2,00	2,00	2,00	2,00	2,00	1,00	2,00			
18	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	1,00	2,00	1,00	2,00	1,00			
19	1,00	2,00	1,00	1,00	2,00	2,00	1,00	2,00	1,00	1,00	2,00	1,00	2,00	1,00	2,00	1,00			
20	1,00	3,00	1,00	1,00	3,00	2,00	1,00	3,00	2,00	1,00	3,00	1,00	3,00	1,00	3,00	2,00			
21																			
22																			

2. Clic en “Analizar”, “Escala” y “Análisis de fiabilidad”.

	VAR00007	VAR00008	VAR00009	VAR00010
.00	2,00	1,00	1,00	3,00
.00	1,00	3,00	2,00	1,00
.00	1,00	2,00	2,00	1,00
.00	2,00	3,00	2,00	2,00
.00	3,00	3,00	2,00	3,00
.00	1,00	3,00	2,00	1,00
.00	2,00	3,00	2,00	2,00
.00	3,00	2,00	2,00	3,00
.00	2,00	2,00	3,00	2,00
.00	1,00	2,00	3,00	1,00

3. Trasladar las variables al cuadro de elementos.



4. Clic en aceptar.

Estadísticas de fiabilidad	
Alfa de Cronbach	N.º de elementos
,852	18

5. Interpretación: se observa que el coeficiente de alfa de Cronbach es 0,852, lo cual indica que el instrumento que mide el liderazgo pedagógico tiene muy buena confiabilidad.

Prueba de normalidad

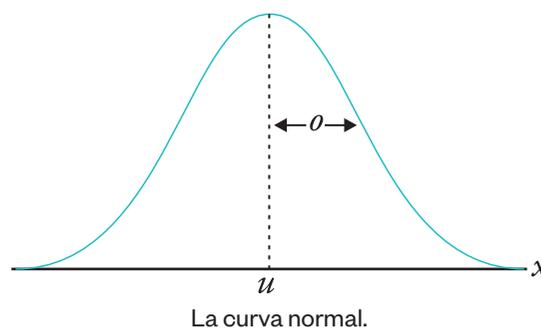
¿Qué es?

Consiste en determinar si los datos siguen una distribución normal.

Una de las características de la distribución normal es que la media es igual a la mediana y a la moda. Al hacer un histograma de frecuencia, la gráfica tendría la forma de una campana (campana de Gauss)

¿Por qué es importante?

Esta prueba resulta indispensable porque determinará el tipo de prueba estadística (paramétrica o no paramétrica) que se aplicará en el tratamiento de los datos.

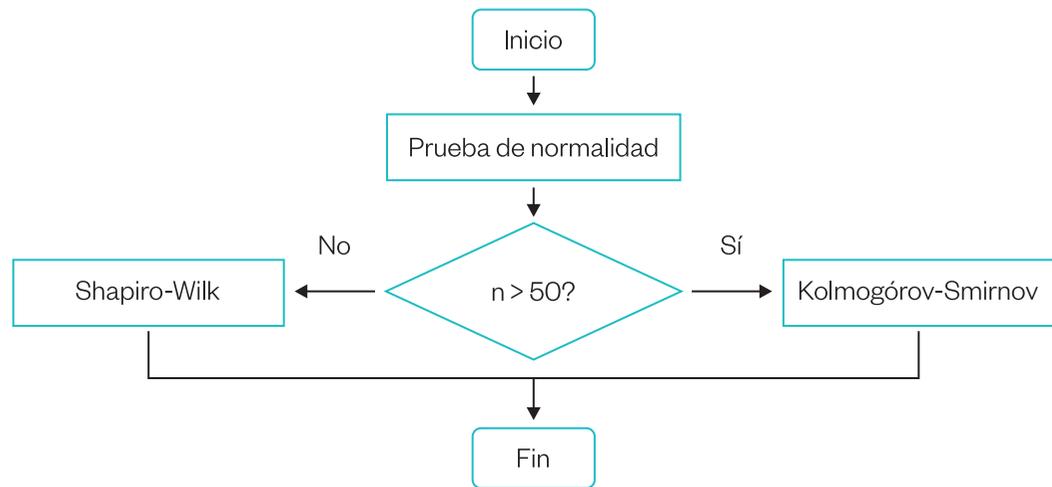


¿Qué estadística usar?

La estadística posee algunas pruebas para determinar la normalidad, como las de Kolmogórov-Smirnov y la de Shapiro-Wilk. Depende del tamaño de la muestra para considerar el uso de la prueba de Kolmogórov-Smirnov ($n > 50$) o la de Shapiro-Wilk ($n \leq 50$).

Prueba de normalidad

Es momento de explicar de forma gráfica y práctica cuándo y qué índice debemos utilizar para el test de normalidad.



Pasos a realizar en el test de normalidad

Paso 1. Plantear la hipótesis nula (H_0) y la hipótesis alternativa (H_1):

H_0 : Los datos provienen de una distribución normal

H_a : Los datos no provienen de una distribución normal

Paso 2. Seleccionar el nivel de significancia

Para efectos de la presente investigación se ha determinado que $\alpha = 0,05$

Paso 3. Escoger el valor estadístico de prueba

El valor estadístico de prueba que se ha considerado para la presente hipótesis es Kolmogórov-Smirnov.

Paso 4. Formulación de regla de decisión

Si $\alpha(\text{Sig}) > 0,05$, se acepta la hipótesis nula

Si $\alpha(\text{Sig}) < 0,05$, se rechaza la hipótesis nula

Paso 5. Toma de decisión

Procedimiento

1. Importar la base de datos del archivo Excel en el programa SPSS.
2. Declarar las variables y dimensiones en la vista de variables e ingresar las respectivas sumatorias dentro del programa SPSS.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	Numérico	8	0	Liderazgo peda...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	Numérico	8	0	Clima organiza...	Ninguno	Ninguno	8	Derecha	Escala
3	v1d1	Numérico	8	0	Desarrollo profe...	Ninguno	Ninguno	8	Derecha	Escala
4	v1d2	Numérico	8	0	Gestión curricular	Ninguno	Ninguno	8	Derecha	Escala
5	v1d3	Numérico	8	0	Objetivos educ...	Ninguno	Ninguno	8	Derecha	Escala
6	v2d1	Numérico	8	0	Estructura	Ninguno	Ninguno	8	Derecha	Escala
7	v2d2	Numérico	8	0	Conflicto y coo...	Ninguno	Ninguno	8	Derecha	Escala
8	v2d3	Numérico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
9										
10										
11										
12										

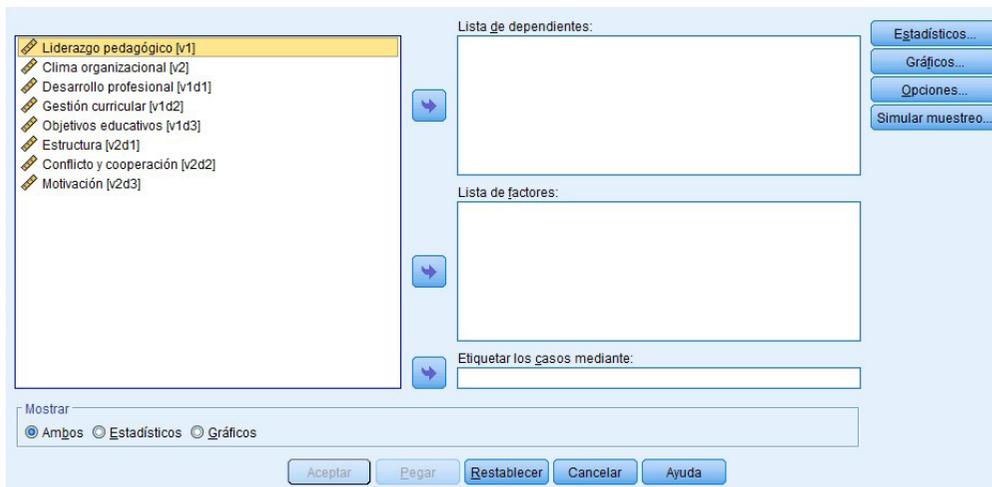
Visualización en “Vista de datos”

	v1	v2	v1d1	v1d2	v1d3	v2d1	v2d2	v2d3	var	var	var	var
1	30	29	9	14	7	10	11	8				
2	32	36	11	15	6	15	10	11				
3	28	30	9	14	5	12	10	8				
4	40	41	12	18	10	16	12	13				
5	47	50	16	21	10	20	15	15				
6	34	37	10	17	7	14	12	11				
7	38	41	12	18	8	16	12	13				
8	39	39	12	18	9	15	12	12				
9	37	37	12	17	8	16	11	10				
10	26	28	8	13	5	12	8	8				

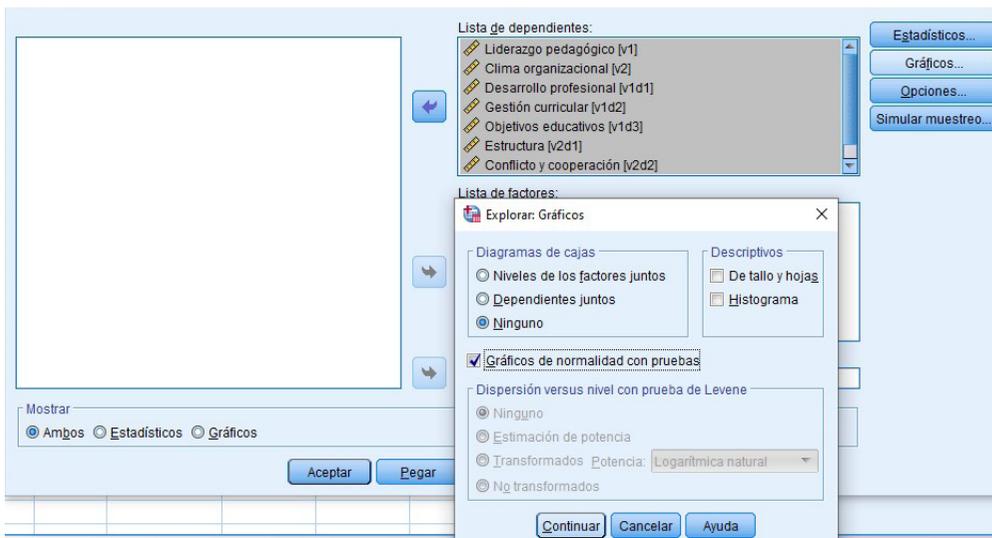
3. Para la obtención del test de normalidad: dar clic en “Analizar”, “Estadísticos descriptivos” y “Explorar”.

Nombre	Tipo	Anchura	
1	v1	Numérico	8
2	v2	Numérico	8
3	v1d1	Numérico	8
4	v1d2	Numérico	8
5	v1d3	Numérico	8
6	v2d1	Numérico	8
7	v2d2	Numérico	8
8	v2d3	Numérico	8
9			

4. Mover las variables a “Lista de dependientes”.



5. Clic en “Gráficos”, “Gráficos de normalidad con pruebas”, “Continuar” y “Aceptar”.



6. Obtener los resultados en la interface “Vista de resultados”.

Pruebas de normalidad

	Kolmogórov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Liderazgo pedagógico	,150	60	,002	,934	60	,003
Clima organizacional	,163	60	,000	,918	60	,001
Desarrollo profesional	,128	60	,016	,957	60	,032
Gestión curricular	,159	60	,001	,922	60	,001
Objetivos educativos	,178	60	,000	,933	60	,003
Estructura	,152	60	,002	,961	60	,050

Conflicto y cooperación	,156	60	,001	,949	60	,015
Motivación	,155	60	,001	,940	60	,005

a. Corrección de significación de Lilliefors

Interpretación

Observamos en la tabla que el valor de significancia calculada en el estadístico de Kolmogórov-Smirnov es menor a 0,05. Por consiguiente, se rechaza la hipótesis nula y se acepta la hipótesis alterna. Es decir, podemos afirmar que los datos no provienen de una distribución normal, y, por lo tanto, en las pruebas de hipótesis se aplicará una prueba no paramétrica.

Sección 5 Análisis descriptivo

Realizaremos el análisis descriptivo con la base de datos del ejemplo 2.

Liderazgo pedagógico			Clima organizacional		
Desarrollo profesional	Gestión curricular	Objetivos educativos	Estructura	Conflicto y cooperación	Motivación

Con ayuda del SPSS calcularemos los valores mínimos, máximos y percentiles.

Procedimiento

Estadísticos									
		Liderazgo pedagógico	Desarrollo profesional	Gestión curricular	Objetivos educativos	Clima organizacional	Estructura	Conflicto y cooperación	Motivación
N.º	Válidos	60	60	60	60	60	60	60	60
	Perdidos	0	0	0	0	0	0	0	0
Mínimo		21	6	9	4	21	7	6	5
Máximo		47	16	21	10	50	20	16	15

Con los datos de la tabla anterior (valores mínimos y máximos) elaboramos la escala valorativa de cada variable.

Escala valorativa de las variables

Variable / Dimensiones	Puntajes		Niveles		
	Mínimo	Máximo	Bajo	Medio	Alto
Liderazgo pedagógico	21	47			
Desarrollo profesional	6	16			
Gestión curricular	9	21			
Objetivos educativos	4	10			

Variable / Dimensiones	Puntajes		Niveles		
	Mínimo	Máximo	Malo	Regular	Bueno
Clima organizacional	21	50			
Estructura	7	20			
Conflicto y cooperación	6	16			
Motivación	5	15			

Para calcular la amplitud de los intervalos de cada nivel de la variable “Liderazgo pedagógico” se procede de la siguiente manera:

- a) Cálculo del rango: $R = \text{Máximo} - \text{Mínimo}$
 $R = 47 - 21 \rightarrow R = 26$
- b) Amplitud: $A = R/3$
 $A = 26/3 = 8,6667 \rightarrow A = 9$
 Es decir, cada nivel para la variable tendrá una amplitud de 9.

Para calcular la amplitud de los intervalos de los niveles en las dimensiones, se procede así:

Dimensión 1: desarrollo profesional

- a) Cálculo del rango: $R = \text{Máximo} - \text{Mínimo}$
 $R = 16 - 6 \rightarrow R = 10$
- b) Amplitud: $A = R/3$
 $A = 10/3 = 3,333 \rightarrow A = 4$
 Es decir, cada nivel para la variable tendrá una amplitud de 4.

Dimensión 2: gestión curricular

- a) Cálculo del rango: $R = \text{Máximo} - \text{Mínimo}$
 $R = 21 - 9 \rightarrow R = 12$
- b) Amplitud: $A = R/3$
 $A = 12/3 = 4 \rightarrow A = 4$
 Es decir, cada nivel para la variable tendrá una amplitud de 4.

Dimensión 3: Objetivos educativos

a) Cálculo del rango: $R = \text{Máximo} - \text{Mínimo}$

$$R = 10 - 4 \rightarrow R = 6$$

b) Amplitud: $A = R/3$

$$A = 6/3 = 2 \rightarrow A = 2$$

Es decir, cada nivel para la variable tendrá una amplitud de 2.

Así, la tabla de escala valorativa quedaría así:

Variable / Dimensiones	Puntajes		Niveles		
	Mínimo	Máximo	Bajo	Medio	Alto
Liderazgo pedagógico	21	47	21-29	30-38	39-47
Desarrollo profesional	6	16	6-9	10-13	14-16
Gestión curricular	9	21	9-12	13-16	17-21
Objetivos educativos	4	10	4-5	6-7	8-10

Realizamos el mismo procedimiento para la variable “Clima organizacional”.

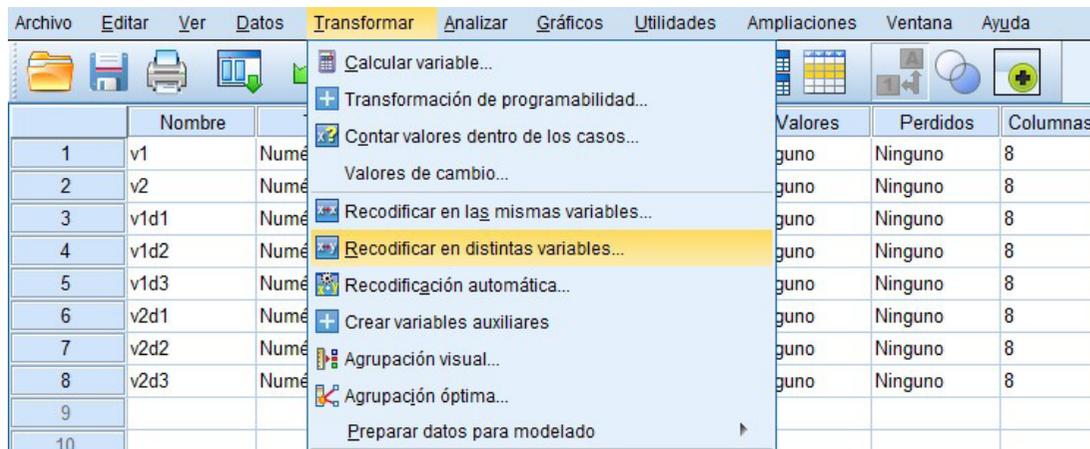
Variable / Dimensiones	Puntajes		Niveles		
	Mínimo	Máximo	Malo	Regular	Bueno
Clima organizacional	21	50	21-30	31-40	41-50
Estructura	7	20	7-11	12-16	17-20
Conflicto y cooperación	6	16	6-9	10-13	14-16
Motivación	5	15	5-8	9-12	13-15

Con los valores de la escala valorativa elaboramos con el SPSS las tablas de frecuencia y los gráficos estadísticos respectivos.

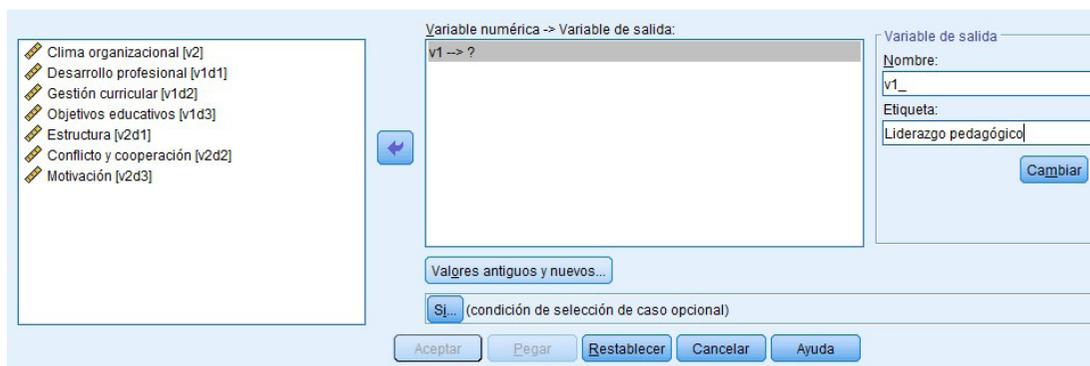
Tablas de frecuencia con SPSS

Procedimiento

1. Clic en “Transformar” y “Recodificar en distintas variables”.



2. Trasladar la primera variable al cuadro de la derecha. Declarar la variable de salida (v1_) y escribir la etiqueta (nombre) de la variable (“Liderazgo pedagógico”). Clic en “Valores antiguos y nuevos”.



3. Clic en “Rango”. Ingresar los límites del nivel 1 (21 y 29) y escribir “1” en el recuadro de valor. Con esto se indica al programa que el nivel “1” lo conformarán los valores comprendidos entre 21 y 29.

Valor antiguo

Valor:

Perdido del sistema

Perdido por el sistema o el usuario

Rango:

21

hasta

29

Rango, LOWEST hasta el valor:

Rango, valor hasta HIGHEST:

Todos los demás valores

Valor nuevo

Valor: 1

Perdido del sistema

Copiar valores antiguos

Antiguo -> Nuevo:

Añadir

Cambiar

Eliminar

Las variables de salida son series Anchura: 8

Convertir series numéricas a números (5'->5)

Continuar Cancelar Ayuda

4. Clic en “Añadir”.

Valor antiguo

Valor:

Perdido del sistema

Perdido por el sistema o el usuario

Rango:

Valor nuevo

Valor:

Perdido del sistema

Copiar valores antiguos

Antiguo -> Nuevo:

21 thru 29 --> 1

Añadir

Cambiar

Eliminar

Las variables de salida son series Anchura: 8

Convertir series numéricas a números (5'->5)

Continuar Cancelar Ayuda

5. Realizar el mismo procedimiento para las demás escalas, con el respectivo número de nivel. Luego hacer clic en “Continuar”.

Valor antiguo

Valor:

Perdido del sistema

Perdido por el sistema o el usuario

Rango:

hasta

Rango, LOWEST hasta el valor:

Rango, valor hasta HIGHEST:

Todos los demás valores

Valor nuevo

Valor:

Perdido del sistema

Copiar valores antiguos

Antiguo --> Nuevo:

21 thru 29 --> 1

30 thru 38 --> 2

39 thru 47 --> 3

Añadir

Cambiar

Eliminar

Las variables de salida son series Anchura: 8

Convertir series numéricas a números ('5'->5)

Continuar Cancelar Ayuda

6. Clic en “Cambiar” y “Aceptar”.

Variable numérica -> Variable de salida:

v1 -> v1_

Variable de salida

Nombre: v1_

Etiqueta: Liderazgo pedagógico

Cambiar

Valores antiguos y nuevos...

SI... (condición de selección de caso opcional)

Aceptar Pegar Restablecer Cancelar Ayuda

Con esto se habrá creado una nueva variable categórica del “Liderazgo pedagógico” (se convierte de escala a ordinal).

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	Númerico	8	0	Liderazgo peda...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	Númerico	8	0	Clima organiza...	Ninguno	Ninguno	8	Derecha	Escala
3	v1d1	Númerico	8	0	Desarrollo profe...	Ninguno	Ninguno	8	Derecha	Escala
4	v1d2	Númerico	8	0	Gestión curricular	Ninguno	Ninguno	8	Derecha	Escala
5	v1d3	Númerico	8	0	Objetivos educ...	Ninguno	Ninguno	8	Derecha	Escala
6	v2d1	Númerico	8	0	Estructura	Ninguno	Ninguno	8	Derecha	Escala
7	v2d2	Númerico	8	0	Conflicto y coo...	Ninguno	Ninguno	8	Derecha	Escala
8	v2d3	Númerico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
9	v1_	Númerico	8	2	Liderazgo peda...	Ninguno	Ninguno	10	Derecha	Ordinal
10										
11										
12										

Vista de datos Vista de variables

Se realiza el mismo procedimiento para las dimensiones y la otra variable.

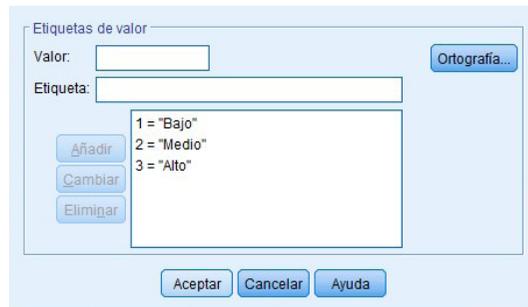
	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	Númerico	8	0	Liderazgo peda...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	Númerico	8	0	Clima organiza...	Ninguno	Ninguno	8	Derecha	Escala
3	v1d1	Númerico	8	0	Desarrollo profe...	Ninguno	Ninguno	8	Derecha	Escala
4	v1d2	Númerico	8	0	Gestión curricular	Ninguno	Ninguno	8	Derecha	Escala
5	v1d3	Númerico	8	0	Objetivos educ...	Ninguno	Ninguno	8	Derecha	Escala
6	v2d1	Númerico	8	0	Estructura	Ninguno	Ninguno	8	Derecha	Escala
7	v2d2	Númerico	8	0	Conflicto y coo...	Ninguno	Ninguno	8	Derecha	Escala
8	v2d3	Númerico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
9	v1_	Númerico	8	0	Liderazgo peda...	Ninguno	Ninguno	10	Derecha	Ordinal
10	v1d1_	Númerico	8	2	Desarrollo profe...	Ninguno	Ninguno	10	Derecha	Nominal
11	v1d2_	Númerico	8	2	Gestión curricular	Ninguno	Ninguno	10	Derecha	Nominal
12	v1d3_	Númerico	8	2	Objetivos educ...	Ninguno	Ninguno	10	Derecha	Nominal
13	v2_	Númerico	8	2	Clima organiza...	Ninguno	Ninguno	10	Derecha	Nominal
14	v2d1_	Númerico	8	2	Estructura	Ninguno	Ninguno	10	Derecha	Nominal
15	v2d2_	Númerico	8	2	Conflicto y coo...	Ninguno	Ninguno	10	Derecha	Nominal
16	v2d3_	Númerico	8	2	Motivación	Ninguno	Ninguno	10	Derecha	Nominal
17										

Posteriormente se procede a determinar los valores y el nombre de cada nueva escala para la variable “Liderazgo pedagógico” y sus dimensiones (bajo, medio, alto).

7. Clic en la casilla donde se cruza la fila de “v1_” y la columna de “Valores”.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	Númerico	8	0	Liderazgo peda...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	Númerico	8	0	Clima organiza...	Ninguno	Ninguno	8	Derecha	Escala
3	v1d1	Númerico	8	0	Desarrollo profe...	Ninguno	Ninguno	8	Derecha	Escala
4	v1d2	Númerico	8	0	Gestión curricular	Ninguno	Ninguno	8	Derecha	Escala
5	v1d3	Númerico	8	0	Objetivos educ...	Ninguno	Ninguno	8	Derecha	Escala
6	v2d1	Númerico	8	0	Estructura	Ninguno	Ninguno	8	Derecha	Escala
7	v2d2	Númerico	8	0	Conflicto y coo...	Ninguno	Ninguno	8	Derecha	Escala
8	v2d3	Númerico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
9	v1_	Númerico	8	0	Liderazgo peda...	Ninguno	Ninguno	10	Derecha	Ordinal
10	v1d1_	Númerico	8	2	Desarrollo profe...	Ninguno	Ninguno	10	Derecha	Nominal
11	v1d2_	Númerico	8	2	Gestión curricular	Ninguno	Ninguno	10	Derecha	Nominal
12	v1d3_	Númerico	8	2	Objetivos educ...	Ninguno	Ninguno	10	Derecha	Nominal
13	v2_	Númerico	8	2	Clima organiza...	Ninguno	Ninguno	10	Derecha	Nominal
14	v2d1_	Númerico	8	2	Estructura	Ninguno	Ninguno	10	Derecha	Nominal
15	v2d2_	Númerico	8	2	Conflicto y coo...	Ninguno	Ninguno	10	Derecha	Nominal
16	v2d3_	Númerico	8	2	Motivación	Ninguno	Ninguno	10	Derecha	Nominal
17										

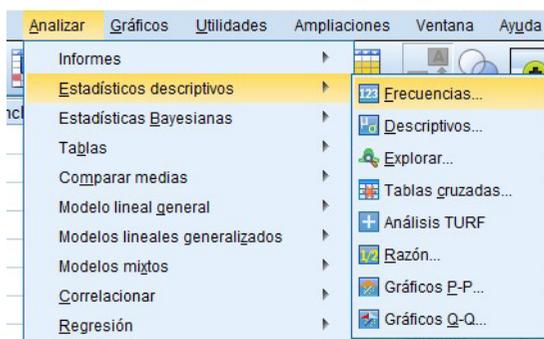
8. Clic en el recuadro azul. Ingresar los valores de cada escala y el nombre respectivo. Clic en “Añadir” (se repite el procedimiento para las dimensiones) y al finalizar “Aceptar”.



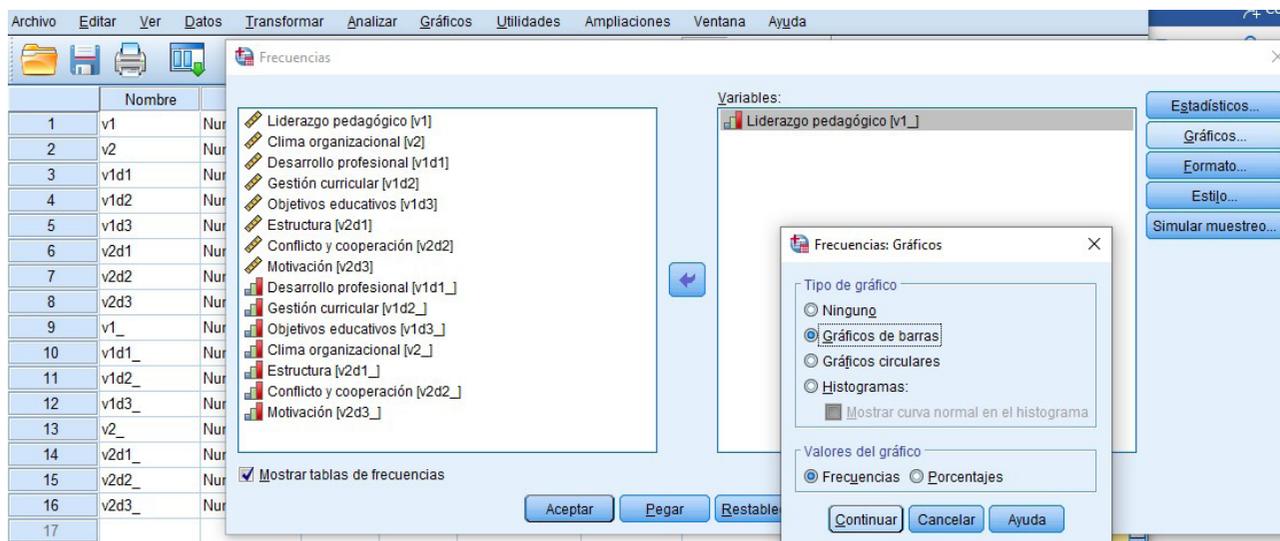
9. Se repite el proceso de recategorización en niveles para la segunda variable, “Clima organizacional”, y sus dimensiones. En la columna de valores convertir a “Ordinal” las nuevas medidas, con lo que quedaría de la siguiente manera:

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	N Numérico	8	0	Liderazgo peda...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	N Numérico	8	0	Clima organiza...	Ninguno	Ninguno	8	Derecha	Escala
3	v1d1	N Numérico	8	0	Desarrollo profe...	Ninguno	Ninguno	8	Derecha	Escala
4	v1d2	N Numérico	8	0	Gestión curricular	Ninguno	Ninguno	8	Derecha	Escala
5	v1d3	N Numérico	8	0	Objetivos educ...	Ninguno	Ninguno	8	Derecha	Escala
6	v2d1	N Numérico	8	0	Estructura	Ninguno	Ninguno	8	Derecha	Escala
7	v2d2	N Numérico	8	0	Conflicto y coo...	Ninguno	Ninguno	8	Derecha	Escala
8	v2d3	N Numérico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
9	v1_	N Numérico	8	0	Liderazgo peda...	{1, Bajo}...	Ninguno	10	Derecha	Ordinal
10	v1d1_	N Numérico	8	0	Desarrollo profe...	{1, Bajo}...	Ninguno	10	Derecha	Ordinal
11	v1d2_	N Numérico	8	0	Gestión curricular	{1, Bajo}...	Ninguno	10	Derecha	Ordinal
12	v1d3_	N Numérico	8	0	Objetivos educ...	{1, Bajo}...	Ninguno	10	Derecha	Ordinal
13	v2_	N Numérico	8	0	Clima organiza...	{1, Malo}...	Ninguno	10	Derecha	Ordinal
14	v2d1_	N Numérico	8	0	Estructura	{1, Malo}...	Ninguno	10	Derecha	Ordinal
15	v2d2_	N Numérico	8	0	Conflicto y coo...	{1, Malo}...	Ninguno	10	Derecha	Ordinal
16	v2d3_	N Numérico	8	0	Motivación	{1, Malo}...	Ninguno	10	Derecha	Ordinal

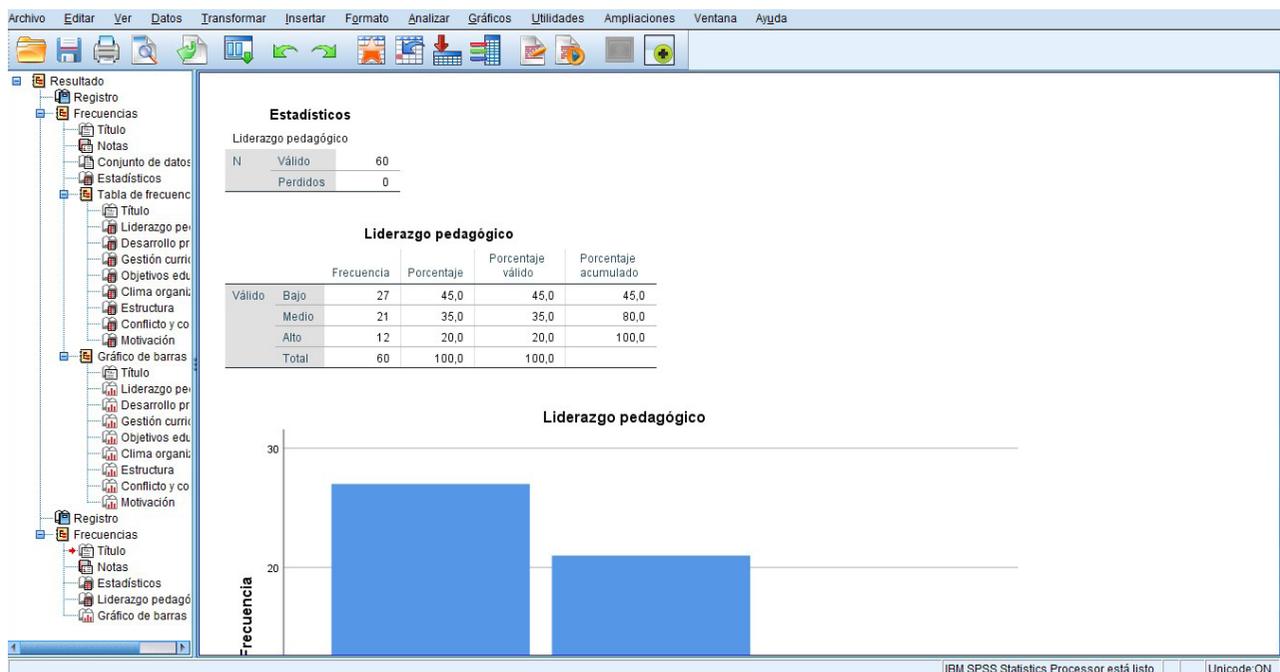
10. Clic en “Analizar”, “Estadísticos descriptivos” y “Frecuencias”.



11. Trasladar las nuevas variables al recuadro derecho. Clic en “Gráficos”, seleccionar el tipo de gráfico que se desea (por ejemplo, “Gráficos de barras”), clic en “Continuar” y “Aceptar”.

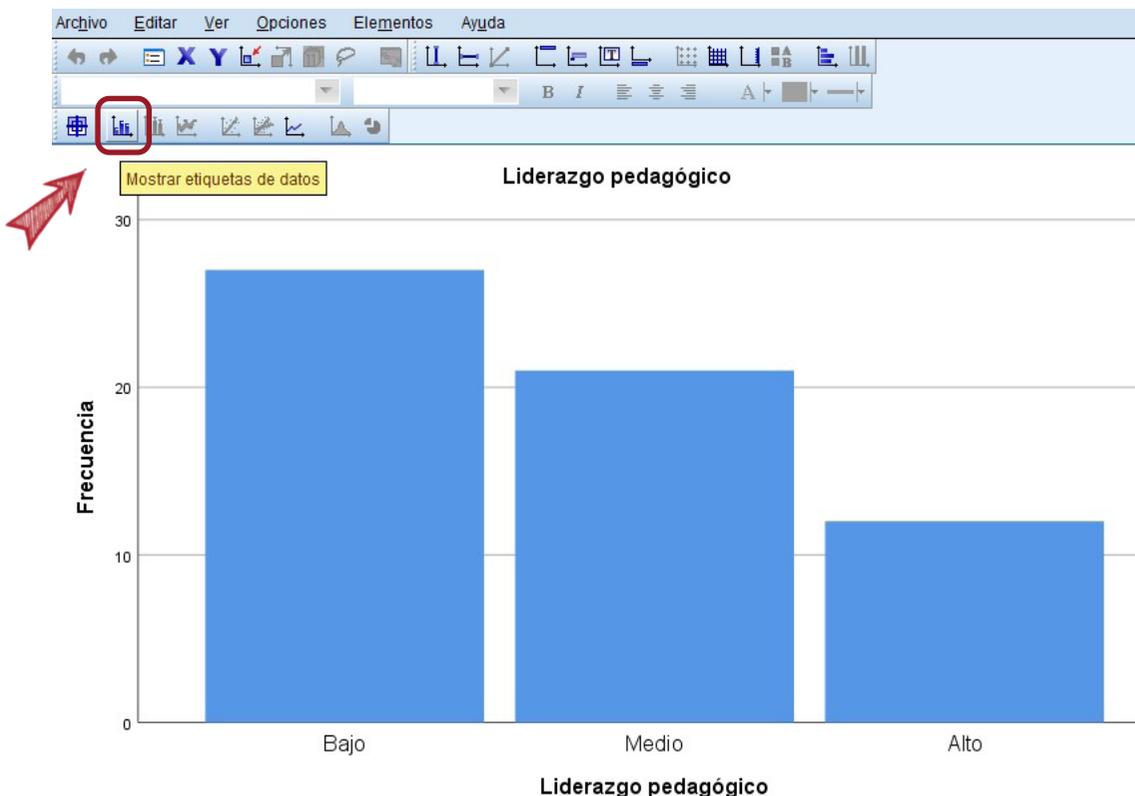


12. El resultado aparecerá en el visor de “Resultados”.

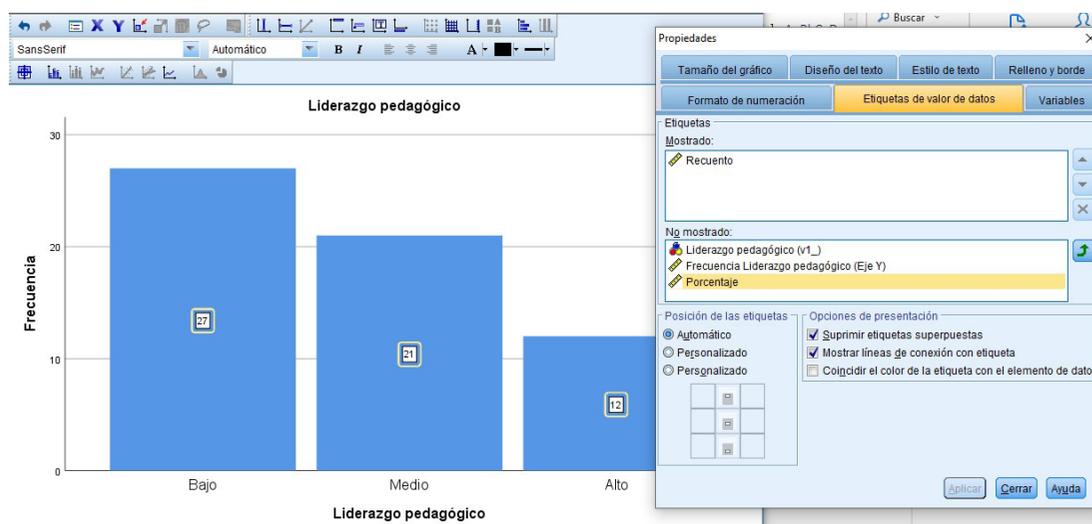


Liderazgo pedagógico					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Bajo	27	45,0	45,0	45,0
	Medio	21	35,0	35,0	80,0
	Alto	12	20,0	20,0	100,0
	Total	60	100,0	100,0	

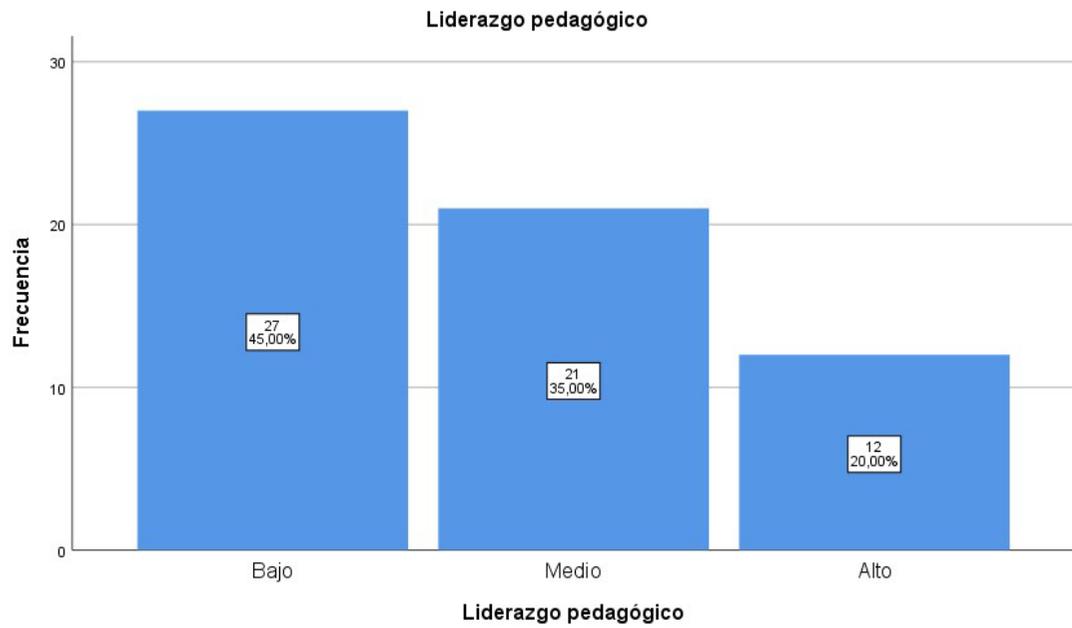
13. Para editar los gráficos se hace doble clic sobre el gráfico de barras. Aparecerá otra ventana. Hacer clic en el ícono de barras ubicado en la parte superior izquierda.



14. Aparecerá una nueva ventana. Seleccionar la opción “Porcentajes”, hacer clic en la flecha verde, “Aplicar” y “Cerrar”.



15. Aparecerá el gráfico con las frecuencias y los porcentajes. Queda listo para copiarlo y pegarlo en nuestro documento Word de tesis o artículo.



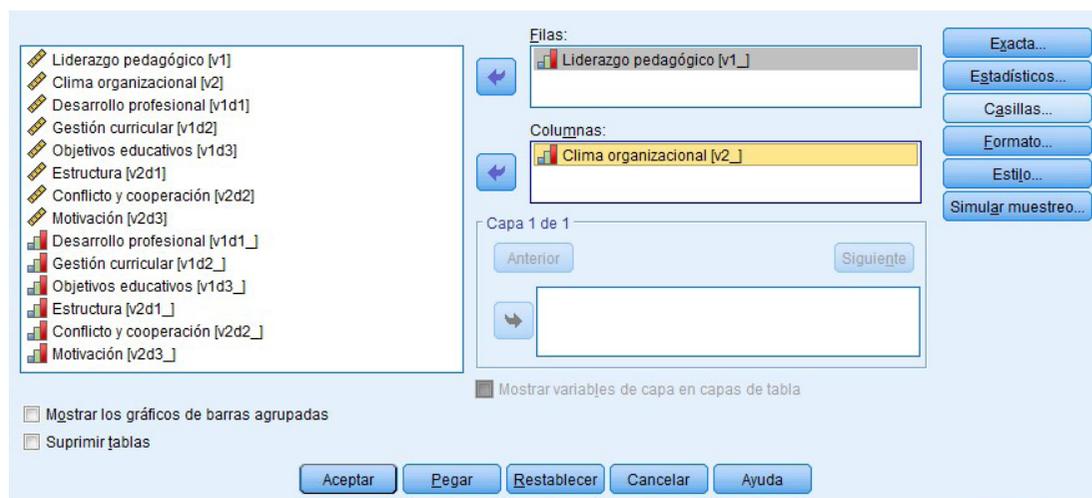
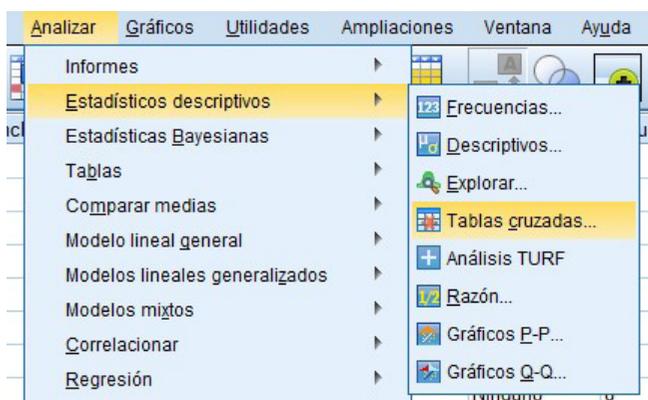
Repetir el proceso para las demás variables y dimensiones.

Tablas de contingencia o cruce de variables

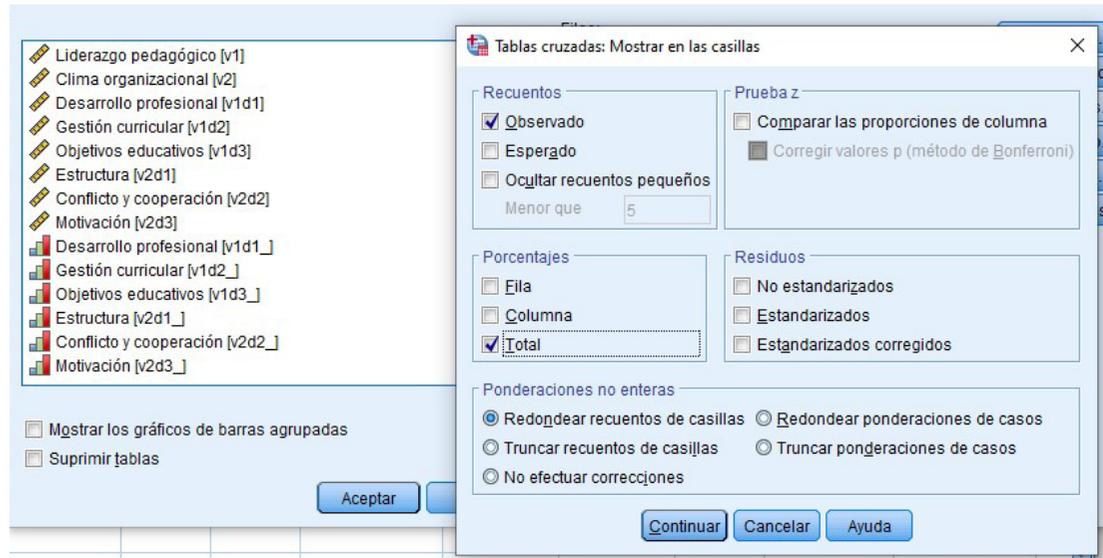
La tabla de contingencia permite visualizar la relación entre las variables.

Procedimiento

1. Clic en “Analizar”, “Estadísticos descriptivos” y “Tablas cruzadas”. Trasladar la variable 1 (independiente) al casillero de “Filas” y la variable 2 (dependiente), al casillero de “Columnas”.



2. Clic en “Casillas”. En la opción “Porcentajes” activar el recuadro de “Total”. Clic en “Continuar” y “Aceptar”.



3. La tabla aparecerá en el visor de “Resultados”.

Tabla cruzada Liderazgo pedagógico - Clima organizacional

			Clima organizacional			Total
			Malo	Regular	Bueno	
Liderazgo pedagógico	Bajo	Recuento	25	2	0	27
		% del total	41,7%	3,3%	0,0%	45,0%
	Medio	Recuento	6	14	1	21
		% del total	10,0%	23,3%	1,7%	35,0%
	Alto	Recuento	0	3	9	12
		% del total	0,0%	5,0%	15,0%	20,0%
Total	Recuento	31	19	10	60	
	% del total	51,7%	31,7%	16,7%	100,0%	

4. Interpretación

En la tabla 3 observamos que, del total de 60 sujetos encuestados, 25, que representan el 41,7 %, tienen un liderazgo pedagógico de nivel bajo y perciben que el clima organizacional es malo; 14 encuestados, que son el 23,3 %, tienen un liderazgo de nivel medio y perciben un clima organizacional regular; mientras que 9 encuestados, que constituyen el 15 %, tienen un alto liderazgo pedagógico y perciben que el clima organizacional es bueno.

Análisis estadístico de estudios correlacionales

¿Qué es?

Es un conjunto de técnicas estadísticas para evaluar las pruebas de hipótesis de relación entre dos variables. Sus valores oscilan entre -1 y 1.

¿Por qué es importante?

Es importante porque estos indicadores estadísticos nos informan si existe una relación significativa. A su vez, informan sobre qué tan fuerte es la magnitud de relación entre las variables estudiadas.

¿Qué estadística usar?

Las técnicas estadísticas de correlación varían según la prueba paramétrica:

Si es no paramétrica: coeficiente de correlación de Spearman

Si es paramétrica: coeficiente de correlación de Pearson

Pasos para evaluar las correlaciones:

1. Identificar la prueba de hipótesis estadística

Ha: existe relación significativa entre dos variables de estudio.

Ho: no existe una relación significativa entre dos variables de estudio.

2. Nivel de significancia

El nivel de significancia teórica es de $\alpha = 0,05$, que corresponde a un nivel de confiabilidad de 95 %.

Procedimiento

1. Ingresar datos (columnas de las sumatorias) y declarar las variables y dimensiones en el SPSS.

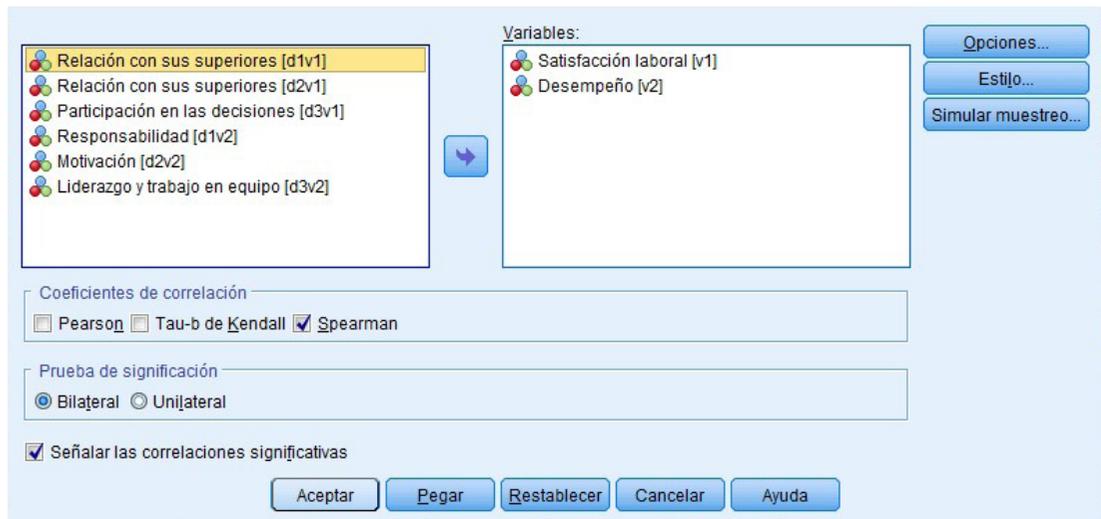
	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	v1	Numérico	8	0	Satisfacción la...	Ninguno	Ninguno	8	Derecha	Escala
2	v2	Numérico	8	0	Desempeño	Ninguno	Ninguno	8	Derecha	Escala
3	d1v1	Numérico	8	0	Relación con s...	Ninguno	Ninguno	8	Derecha	Escala
4	d2v1	Numérico	8	0	Relación con s...	Ninguno	Ninguno	8	Derecha	Escala
5	d3v1	Numérico	8	0	Participación e...	Ninguno	Ninguno	8	Derecha	Escala
6	d1v2	Numérico	8	0	Responsabilidad	Ninguno	Ninguno	8	Derecha	Escala
7	d2v2	Numérico	8	0	Motivación	Ninguno	Ninguno	8	Derecha	Escala
8	d3v2	Numérico	8	0	Liderazgo y tra...	Ninguno	Ninguno	8	Derecha	Escala
9										
10										
11										

	v1	v2	d1v1	d2v1	d3v1	d1v2	d2v2	d3v2	var	var	var
1	64	54	22	10	32	18	10	26			
2	66	63	21	12	33	21	8	34			
3	39	42	13	6	20	13	8	21			
4	72	67	24	12	36	21	11	35			
5	52	41	18	9	25	16	7	18			
6	62	64	20	10	32	24	10	30			
7	58	43	20	10	28	15	7	21			
8	63	56	20	11	32	21	5	30			
9	64	44	22	10	32	16	8	20			
10	23	25	7	4	12	11	3	11			
11	26	39	8	6	12	12	8	19			
12	65	64	22	10	33	22	12	30			
13	72	65	24	12	36	23	10	32			
14	28	37	8	4	16	14	5	18			

2. Clic en “Analizar”, “Correlacional” y “Bivariadas”.

Analizar	Gráficos	Utilidades	Ampliaciones	Ventana	Ayuda
Informes					
Estadísticos descriptivos					
Estadísticas Bayesianas					
Tablas					
Comparar medias					
Modelo lineal general					
Modelos lineales generalizados					
Modelos mixtos					
Correlacionar					
Regresión					
Loglineal					
Redes neuronales					
Clasificar					

- Mover las variables al cuadro derecho. Clic en “Coeficiente de correlación” y seleccionar Spearman, y clic en “Aceptar”.



- Visualización de resultados.

Correlaciones

			Satisfacción laboral	Desempeño
Rho de Spearman	Satisfacción laboral	Coeficiente de correlación	1,000	,784**
		Sig. (bilateral)	.	,000
		N	60	60
	Desempeño	Coeficiente de correlación	,784**	1,000
		Sig. (bilateral)	,000	.
		N	60	60

** La correlación es significativa en el nivel 0,01 (bilateral).

- Interpretación

Grado de relación entre las variables

En la tabla se observa que, según el valor del coeficiente de correlación de Spearman, el grado de relación entre la satisfacción laboral y el desempeño es $\rho = 0,784^*$, lo que indica que la relación entre las variables es buena y positiva.

Decisión estadística

Como el valor de significancia observada ($\text{sig.} = 0,000$) es menor al valor de la significancia teórica ($\alpha = 0,05$), se puede señalar que hay relación significativa entre las variables. Por lo tanto, se rechaza la hipótesis nula y se acepta la

hipótesis alternativa. Es decir, existe relación entre la satisfacción laboral y el desempeño de los trabajadores de una compañía industrial de Lima.

Ejercicio

Realizar la prueba de hipótesis específicas del problema anterior.

Satisfacción laboral	Desempeño
v1d1: relación con sus superiores	v2d1: responsabilidad
v1d2: condiciones de trabajo	v2d2: motivación
v1d3: participación en las decisiones	v2d3: liderazgo y trabajo en equipo

Coeficiente de correlación de Pearson

Consiste en una prueba paramétrica que mide el nivel o grado de relación de dos variables.

Ejemplo 5

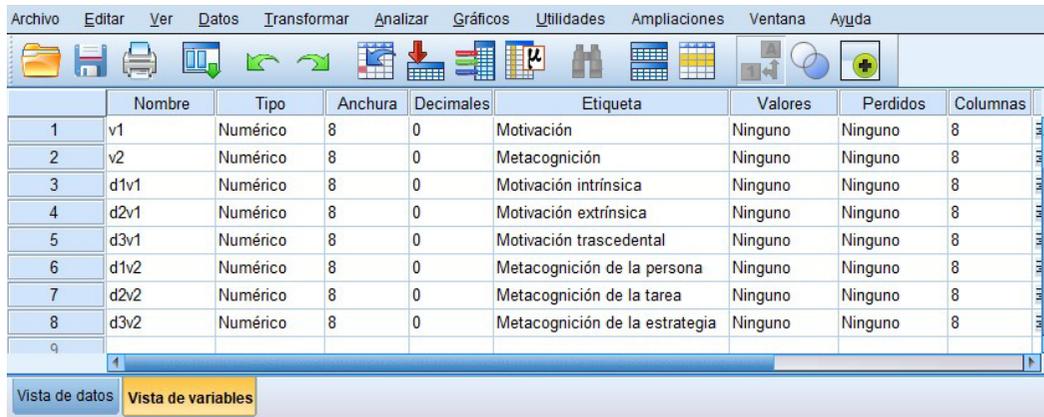
Se tiene el siguiente caso con su respectiva base de datos:

Problema	Objetivo	Hipótesis
¿Cuál es la relación entre la motivación y las estrategias metacognitivas de aprendizaje en estudiantes de maestría en Docencia Universitaria de una universidad privada?	Determinar la relación entre la motivación y las estrategias metacognitivas de aprendizaje en estudiantes de maestría en Docencia Universitaria de una universidad privada.	Existe relación significativa entre la motivación y las estrategias metacognitivas de aprendizaje en estudiantes de maestría en Docencia Universitaria de una universidad privada.

Motivación	Metacognición	Motivación intrínseca	Motivación extrínseca	Motivación trascendental	Metacognición de la persona	Metacognición de la tarea	Metacognición de la estrategia
v1	v2	d1v1	d1v1	d1v1	d1v2	d1v2	d1v2
100	93	36	44	20	36	27	30
106	91	38	48	20	35	26	30
95	83	36	42	17	34	24	25
107	98	39	48	20	38	30	30
85	94	38	29	18	39	27	28
85	83	35	32	18	28	29	26
75	80	32	27	16	32	24	24
96	80	40	37	19	32	24	24
72	74	32	26	14	30	24	20
77	81	34	37	6	32	28	21
82	89	35	28	19	34	28	27
46	50	20	17	9	25	13	12
98	89	36	42	20	33	26	30
80	81	31	33	16	33	22	26
102	98	35	48	19	39	29	30
79	88	30	32	17	33	26	29
53	47	19	24	10	23	12	12
110	96	40	50	20	37	29	30
83	80	31	35	17	31	25	24
82	83	35	30	17	31	24	28
95	82	38	37	20	33	23	26
100	96	38	44	18	39	27	30
88	80	32	40	16	32	24	24
104	96	40	44	20	38	29	29
78	80	34	28	16	32	24	24
89	82	32	40	17	33	24	25
95	79	39	37	19	30	24	25
92	78	33	43	16	33	25	20
54	55	18	21	15	21	15	19
86	82	34	36	16	35	24	23
89	90	34	38	17	35	26	29
84	73	36	34	14	28	21	24
88	75	34	35	19	28	23	24
82	85	32	34	16	32	29	24
65	50	30	31	4	21	15	14
100	96	40	40	20	37	29	30
75	85	33	26	16	35	26	24
93	78	36	41	16	33	23	22
77	74	33	28	16	31	21	22
83	75	32	36	15	27	24	24
99	83	40	40	19	31	28	24
92	82	35	39	18	33	24	25
64	65	36	21	7	35	15	15
110	96	40	50	20	37	29	30

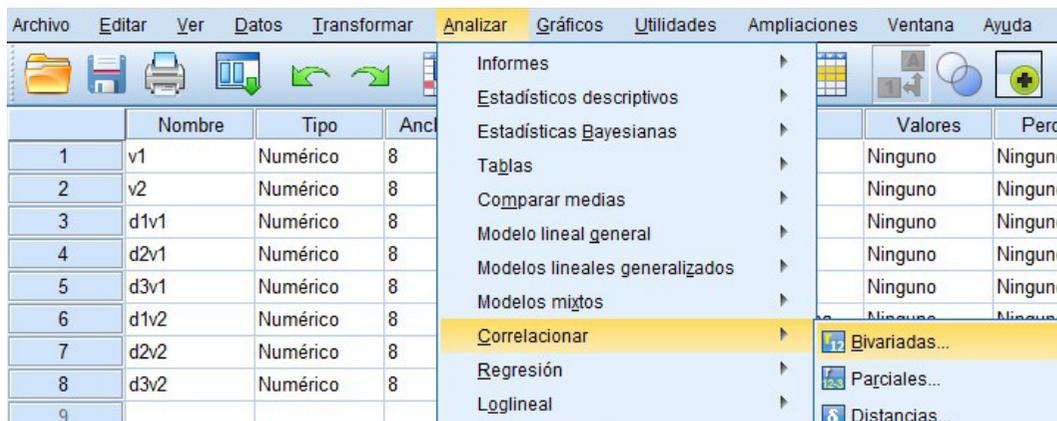
Procedimiento

1. Ingresar datos (columnas de las sumatorias) y declarar a las variables y dimensiones en el SPSS.

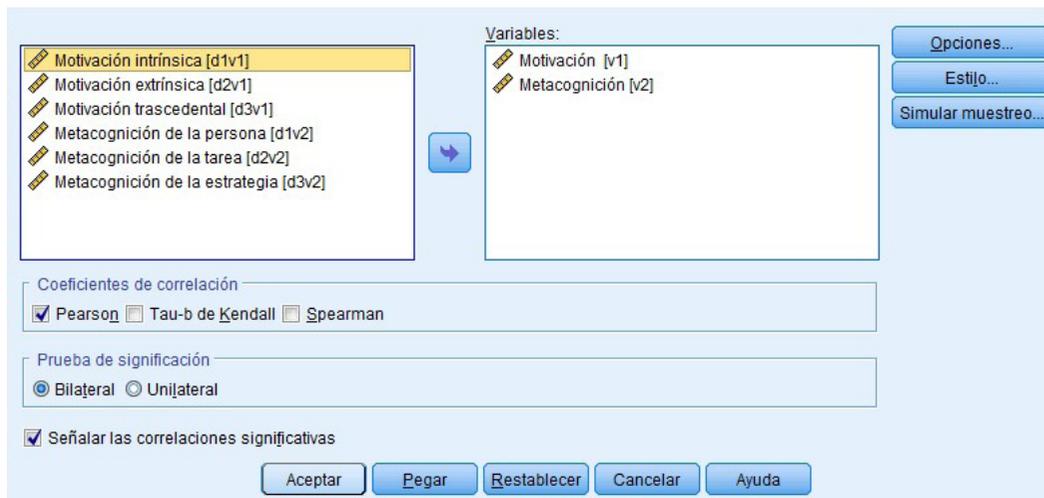


	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas
1	v1	Númerico	8	0	Motivación	Ninguno	Ninguno	8
2	v2	Númerico	8	0	Metacognición	Ninguno	Ninguno	8
3	d1v1	Númerico	8	0	Motivación intrínseca	Ninguno	Ninguno	8
4	d2v1	Númerico	8	0	Motivación extrínseca	Ninguno	Ninguno	8
5	d3v1	Númerico	8	0	Motivación trascendental	Ninguno	Ninguno	8
6	d1v2	Númerico	8	0	Metacognición de la persona	Ninguno	Ninguno	8
7	d2v2	Númerico	8	0	Metacognición de la tarea	Ninguno	Ninguno	8
8	d3v2	Númerico	8	0	Metacognición de la estrategia	Ninguno	Ninguno	8
9								

2. Clic en “Analizar”, “Correlacional” y “Bivariadas”.



3. Mover las variables al cuadro derecho. Clic en “Pearson” y “Aceptar”.



4. Visualización.

Correlaciones

		Motivación	Metacognición
Motivación	Correlación de Pearson	1	,742**
	Sig. (bilateral)		,000
	N.º	81	81
Metacognición	Correlación de Pearson	,742**	1
	Sig. (bilateral)	,000	
	N.º	81	81

** La correlación es significativa en el nivel 0,01 (bilateral).

5. Interpretación.

Grado de relación entre las variables

En la tabla se observa que, según el valor del coeficiente de correlación de Pearson, el grado de relación entre la motivación y la metacognición es $r_s = 0,742^*$, lo cual indica que la relación entre las variables es buena y positiva.

Decisión estadística

Debido a que el valor de significancia observada ($\text{sig.} = 0,000$) es menor al valor de la significancia teórica ($\alpha = 0,05$), se puede afirmar que existe una relación significativa entre las variables. Por lo tanto, se rechaza la hipótesis nula y se acepta la hipótesis alternativa. Es decir, existe relación significativa entre la motivación y las estrategias metacognitivas de aprendizaje en estudiantes de maestría en Docencia Universitaria de una universidad privada.

Ejercicio

Realizar la prueba de hipótesis específicas del problema anterior.

Motivación	Metacognición
v1d1: motivación intrínseca	v2d1: metacognición de la persona
v1d2: motivación extrínseca	v2d2: metacognición de la tarea
v1d3: motivación trascendental	v2d3: metacognición de la estrategia

Prueba de chi cuadrado

Es una prueba no paramétrica que permite reconocer la asociación entre dos variables categóricas (dicotómicas o politómicas).

Requisitos para su aplicación:

- i) Variables ordinales
- ii) Independencia de los datos
- iii) El valor de las medidas esperadas debe ser mayor a 5

Ejemplo 6

Problema	Objetivo	Hipótesis
¿Cuál es la relación entre el clima escolar y el logro académico en estudiantes de una institución educativa de Cañete?	Determinar la relación entre el clima escolar y el logro académico en estudiantes de una institución educativa de Cañete.	Existe relación significativa entre el clima escolar y el logro académico en estudiantes de una institución educativa de Cañete.

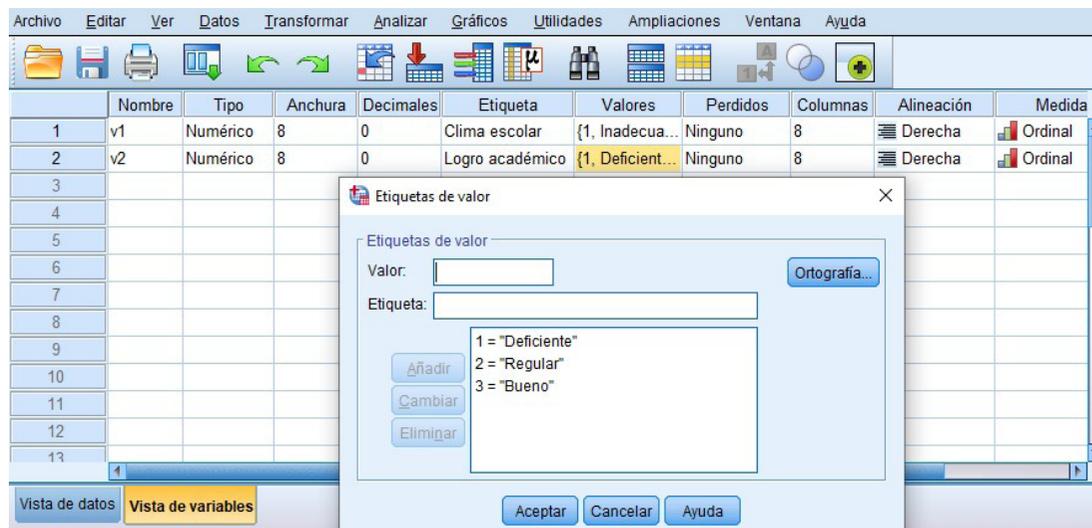
Se tiene el siguiente caso:

Clima académico		
Inadecuado	Regular	Adecuado
1	2	3

Logro académico		
Deficiente	Regular	Bueno
1	2	3

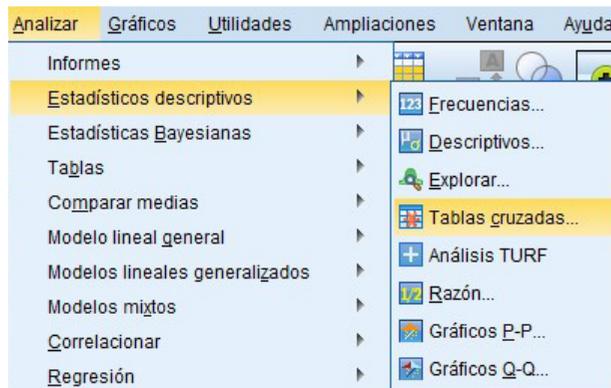
Procedimiento

- Ingresar los datos y declarar las variables.

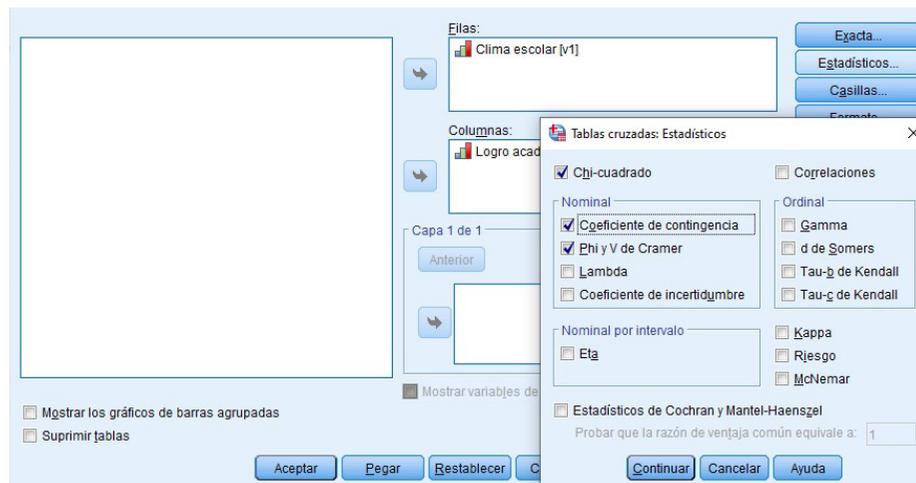


	v1	v2	var	var	var	var	var	var	v
1	3	2							
2	3	3							
3	1	1							
4	3	2							
5	2	1							
6	3	3							
7	2	2							
8	2	2							
9	3	3							
10	1	1							
11	3	3							
12	2	3							
13	2	3							
14	2	2							

2. Clic en “Analizar”, “Estadísticos descriptivos” y “Tablas cruzadas”.



3. Trasladar la variable 1 al casillero “Filas” y la variable 2 al casillero “Columnas”. Clic en “Estadísticos”. Seleccionar la opción “Chi cuadrado”, “Coeficiente de contingencia”, “Phi y V de Cramer” y “Continuar”.



4. Clic en “Casillas”. En “Recuentos” seleccionar “Observado” y “Esperado”, y en “Porcentajes” seleccionar “Total”. Clic en “Continuar”.

5. Visualizar.

Tabla cruzada “Clima escolar / Logro académico”

			Logro académico			Total
			Bueno	Regular	Deficiente	
Clima escolar	Inadecuado	Recuento	5	1	0	6
		Recuento esperado	1,0	2,2	2,8	6,0
		% del total	8,3 %	1,7 %	0,0 %	10,0 %
	Regular	Recuento	2	15	14	31
		Recuento esperado	5,2	11,4	14,5	31,0
		% del total	3,3 %	25,0 %	23,3 %	51,7 %
	Adecuado	Recuento	3	6	14	23
		Recuento esperado	3,8	8,4	10,7	23,0
		% del total	5,0 %	10,0 %	23,3 %	38,3 %
Total		Recuento	10	22	28	60
		Recuento esperado	10,0	22,0	28,0	60,0
		% del total	16,7 %	36,7 %	46,7 %	100,0 %

En la tabla anterior observamos que no se cumple el requisito 3, que indica que los valores esperados deben ser mayores a 5. Por consiguiente, los resultados que se deben reportar son los de la significancia bilateral de la prueba exacta de Fisher.

Pruebas de chi cuadrado

	Valor	df	Significación asintótica (bilateral)	Significación exacta (bilateral)	Significación exacta (unilateral)	Probabilidad en el punto
Chi cuadrado de Pearson	24,449 ^a	4	,000	,000		
Razón de verosimilitud	20,008	4	,000	,001		
Prueba exacta de Fisher	17,265			,001		
Asociación lineal por lineal	8,891 ^b	1	,003	,003	,002	,001
N.º de casos válidos	60					

En la tabla anterior observamos que la significancia calculada (de Fisher) es $0,001 < 0,05$. Esto permite afirmar que las variables tienen una relación significativa. La tabla siguiente nos muestra el nivel de asociación de las variables. Mientras más cerca esté de la unidad, mejor será el nivel de asociación.

Medidas simétricas

		Valor	Significación aproximada	Significación exacta
Nominal por nominal	Phi	,638	,000	,000
	V de Cramer	,451	,000	,000
	Coefficiente de contingencia	,538	,000	,000
N.º de casos válidos		60		

Como las variables tienen más de dos niveles, se considera V de Cramer para determinar el nivel de asociación. Según la tabla anterior, el coeficiente de V de Cramer (0,451) indica que la asociación o relación entre las variables es moderada.

Análisis estadístico de estudios preexperimentales

¿Qué es?

Es un conjunto de técnicas estadísticas que sirven para evaluar el impacto o la influencia de actor externo dentro de una muestra relacionada o independiente.

¿Por qué es importante?

Es importante porque permite evaluar el impacto de un determinado factor y su cambio en una muestra relacionada (pareada) o independiente.

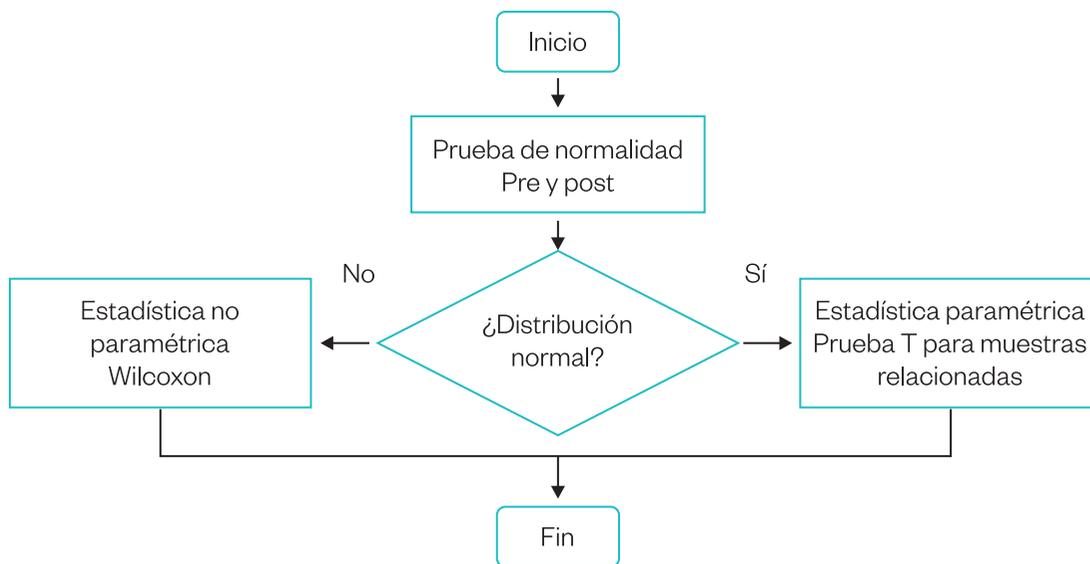
¿Qué estadística usar?

Las técnicas estadísticas de correlación varían según la prueba paramétrica:

Si es no paramétrica: Wilcoxon

Si es paramétrica: prueba T para muestras relacionadas

Pasos para evaluar las correlaciones:



1. Identificar la prueba de hipótesis estadística

Ha: la media de poblaciones son iguales.

Ho: la media de poblaciones son diferentes.

2. Nivel de significancia

El nivel de significancia teórica es $\alpha = 0,05$, que corresponde a un nivel de confiabilidad de 95 %.

3. Función de prueba

La prueba de hipótesis se realizará con el test de Wilcoxon o muestras relacionadas, según la normalidad de los datos.

4. Regla de decisión

Rechazar H_0 cuando la significancia observada “sig.” de los coeficientes del modelo es menor que α .

No rechazar H_0 cuando la significancia observada “sig.” de los coeficientes del modelo es mayor que α .

5. Interpretar los coeficientes de correlación

Se realiza luego de conocer si existe diferencia entre las dos muestras de estudios tomadas en diferentes tiempos o momentos, después de la aplicación de un factor externo.

Prueba de rangos de Wilcoxon

Ejemplo 7

Se tiene el siguiente caso con su respectiva base de datos: programa de capacitación en informática para mejorar las competencias digitales de los docentes de una universidad peruana.

Variable evaluada: competencias digitales docentes

Factor de impacto: programa de capacitación de informática

Unidad de análisis: docentes universitarios peruanos

Problema	Objetivo	Hipótesis
¿En qué medida la aplicación de un programa de capacitación en informática mejora las competencias digitales de los docentes de una universidad peruana?	Determinar la medida en que la aplicación de un programa de capacitación en informática mejora las competencias digitales de los docentes de una universidad peruana.	La aplicación de un programa de capacitación en informática mejora significativamente las competencias digitales de los docentes de una universidad peruana.

°. Z	Variable Competencias digitales		Dimensión 1 Instrumentales		Dimensión 2 Didáctico-metodológicas		Dimensión 3 cognitivas	
	Pretest	Postest	D1Pre	D1Post	D2Pre	D2Post	D3Pre	D3Post
1	43	51	6	20	12	16	17	15
2	30	36	5	13	9	11	11	12
3	50	47	7	15	15	14	19	18
4	64	56	10	17	17	14	26	25

5	54	52	9	17	15	15	21	20
6	45	53	8	19	13	16	17	18
7	48	55	10	19	14	16	18	20
8	51	60	8	21	14	18	20	21
9	42	48	7	17	13	15	15	16
10	53	53	9	17	16	14	19	22
11	45	52	7	18	12	16	18	18
12	46	50	9	17	14	14	16	19
13	57	57	9	19	16	16	22	22
14	43	49	8	17	13	15	16	17
15	47	57	9	20	12	18	19	19
16	45	53	8	19	13	16	17	18
17	40	49	6	18	11	14	16	17
18	60	51	11	15	19	12	21	24
19	46	50	6	17	12	15	19	18
20	33	42	9	15	11	12	10	15
21	54	62	12	21	17	19	19	22
22	47	46	5	16	12	13	20	17
23	48	56	9	19	14	18	18	19
24	50	52	12	17	16	15	17	20
25	32	47	5	18	9	16	12	13
26	72	65	12	20	20	17	28	28
27	54	61	9	21	17	18	19	22
28	51	57	9	20	14	17	20	20
29	44	52	7	18	12	16	17	18
30	46	58	7	21	13	18	18	19
31	39	52	7	19	9	17	17	16
32	42	46	7	17	13	14	15	15
33	36	52	6	20	10	17	14	15
34	40	57	7	22	12	18	15	17
35	49	61	8	22	14	19	19	20
36	42	48	7	17	12	15	17	16
37	36	52	4	20	9	16	15	16
38	47	53	7	20	14	16	18	17
39	48	50	9	17	14	15	19	18
40	72	65	12	21	20	17	28	27

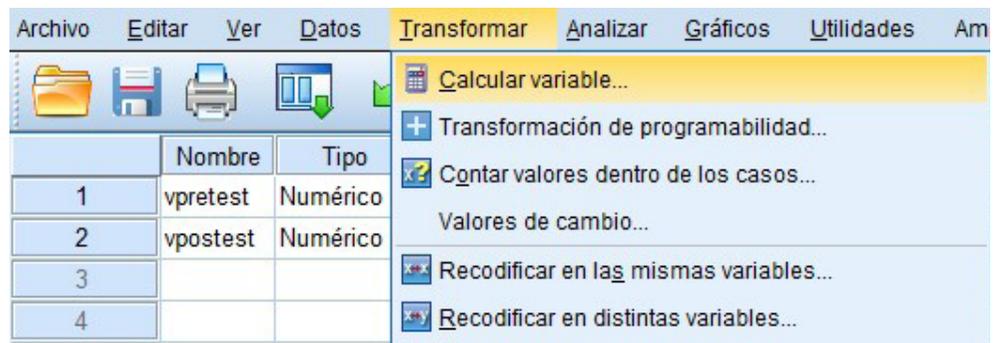
Normalidad de los datos

Procedimiento

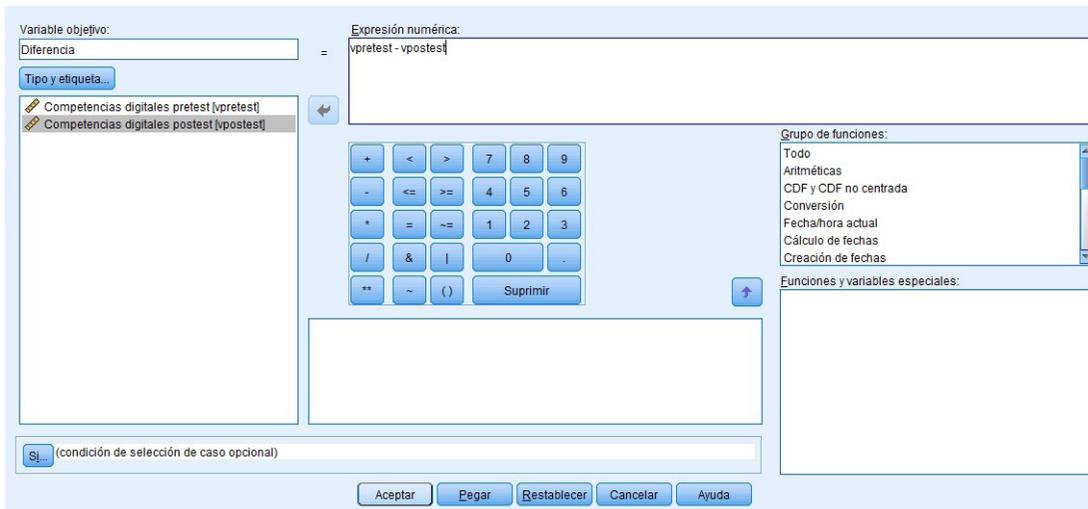
1. Ingresar los datos y declarar las variables en el SPSS.



2. Crear una nueva variable que denominaremos “Diferencia” ($vpretest - vpostest$) para calcular la normalidad de los datos. Clic en “Transformar” y “Calcular variable”.

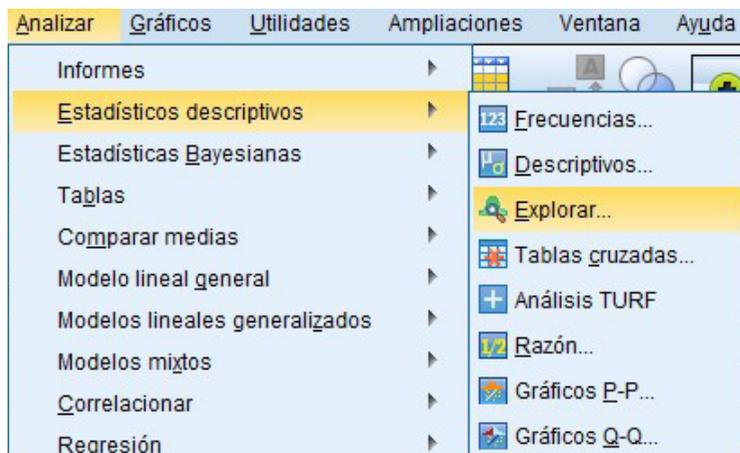


3. Clic en el recuadro “Variable objetivo”. Trasladar al casillero “Expresión numérica” la variable “Competencias digitales pretest”. Clic en el signo menos (-). Trasladar la variable “Competencia digitales postest”. Clic en “Aceptar”.

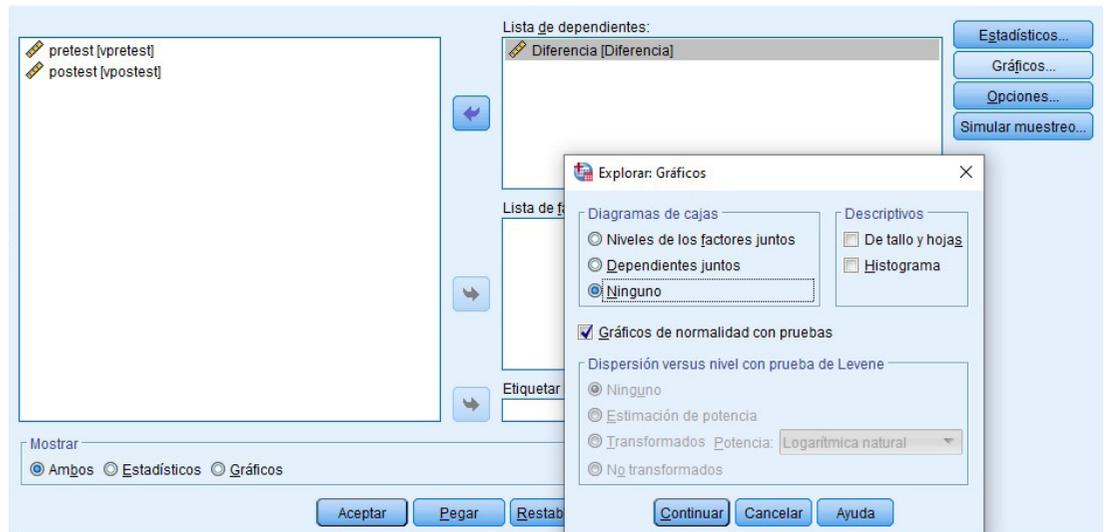


	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columna
1	vpretest	Numérico	8	0	pretest	Ninguno	Ninguno	8
2	vpostest	Numérico	8	0	postest	Ninguno	Ninguno	8
3	Diferen...	Numérico	8	2	Diferencia	Ninguno	Ninguno	12
4								
5								
6								
7								

4. Clic en “Analizar”, “Estadísticos descriptivos” y “Explorar”.



5. Trasladar la nueva variable “Diferencias” al casillero de “Lista de dependientes”. Clic en “Gráficos” y “Gráficos de normalidad con pruebas”. Luego, clic en “Continuar” y “Aceptar”.



6. Visualización e interpretación de resultados de normalidad de la variable.

Pruebas de normalidad

	Kolmogórov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	Gl	Sig.	Estadístico	gl	Sig.
Diferencia	,142	56	,007	,964	56	,097

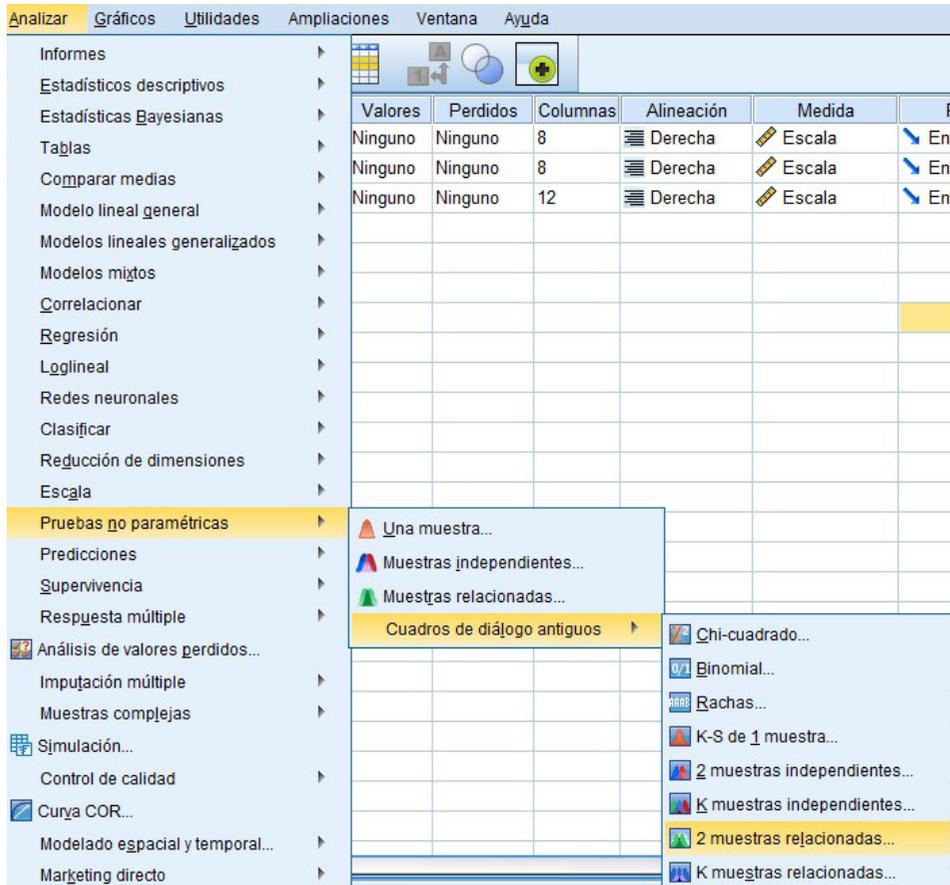
a. Corrección de significación de Lilliefors

En la tabla anterior observamos que el valor de la significancia según la prueba de Kolmogórov-Smirnov (sig.) es $0,007 < 0,05$. Se evidencia que la muestra no proviene de una distribución normal; por lo tanto, la prueba de hipótesis será con el estadístico de prueba de rangos de Wilcoxon.

Prueba de rangos de Wilcoxon

Procedimiento

1. Clic en “Analizar”, “Pruebas no paramétricas”, “Cuadros de diálogo antiguos” y “2 muestras relacionadas”.



2. Trasladar el pretest y postest al casillero “Contrastar pares”. En “Tipo de prueba” hacer clic en “Wilcoxon” y luego en “Aceptar”.



3. Visualización e interpretación.

Estadísticos de prueba^a

	Competencias digitales posttest - Competencias digitales pretest
Z	-4,477 ^b
Sig. asintótica(bilateral)	,000

a. Prueba de rangos con signo de Wilcoxon
b. Se basa en rangos negativos

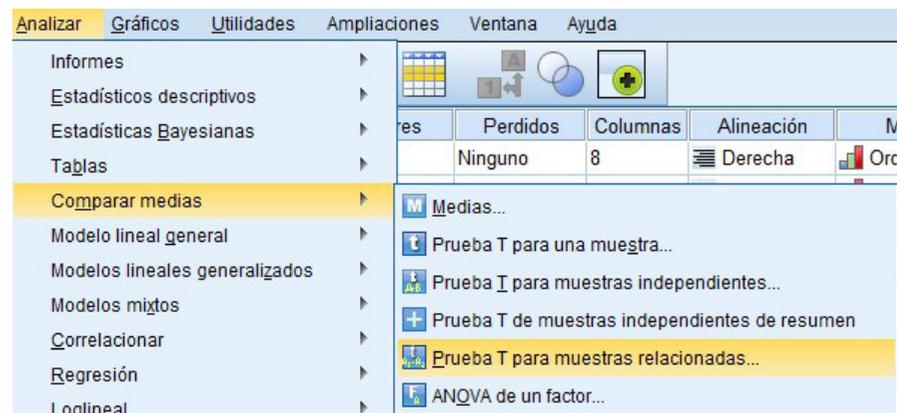
En la tabla anterior observamos que el valor de significancia es $0,000 < 0,05$, lo que confirma una diferencia significativa entre el pretest y posttest. En consecuencia, podemos afirmar que la aplicación de un programa de capacitación en informática mejora significativamente las competencias digitales de los docentes de una universidad peruana.

Prueba T para muestras relacionadas

Con la misma base de datos del ejemplo anterior, vamos a suponer que los datos no tienen distribución normal, por lo que aplicaremos la prueba T para muestras relacionadas.

Procedimiento

1. Clic en “Analizar”, “Comparar medias” y “Prueba T para muestras relacionadas”.



2. Trasladar las variables al casillero “Variables emparejadas”. Clic en “Aceptar”.

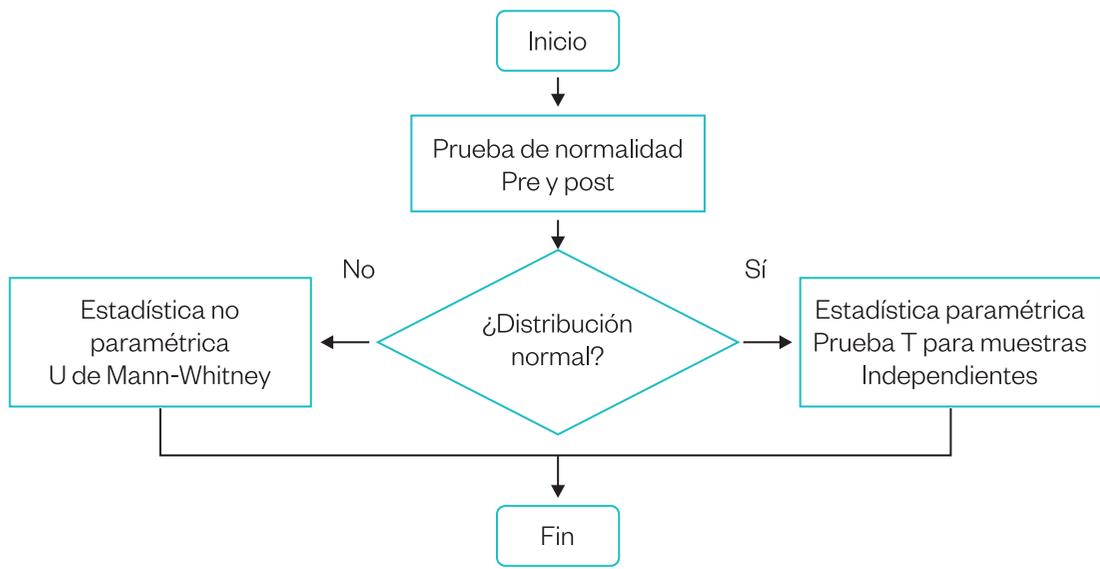


3. Visualización e interpretación.

Prueba de muestras emparejadas								
	Diferencias emparejadas					t	gl	Sig. (bilateral)
	Media	Desv. desviación	Desv. error promedio	95 % de intervalo de confianza de la diferencia				
				Inferior	Superior			
Competencias digitales pretest - Competencias digitales posttest	-4,839	6,508	,870	-6,582	-3,096	-5,564	55	,000

En la tabla se observa que el valor de significancia bilateral es $0,000 < 0,05$. Esto indica que existe una diferencia significativa entre las competencias digitales en el pretest y posttest.

Análisis estadístico de estudios cuasiexperimentales



Prueba U de Mann-Whitney

Ejemplo 8

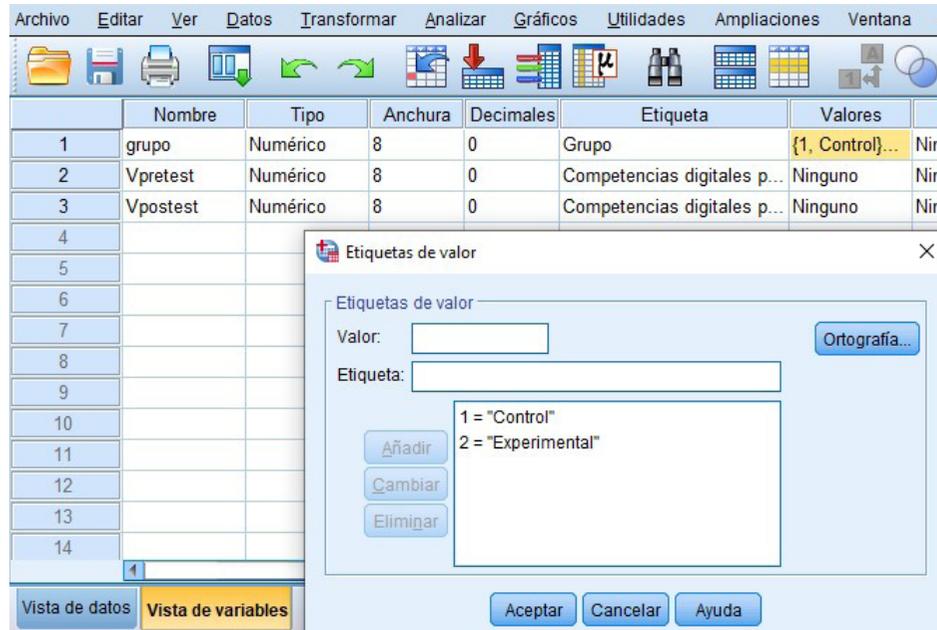
Se tiene el siguiente caso con su respectiva base de datos: aplicación del método Pólya para mejorar las capacidades matemáticas de los estudiantes de una universidad peruana.

Problema	Objetivo	Hipótesis
¿Cómo influye la aplicación del método Pólya en el desarrollo de las capacidades matemáticas de los estudiantes de una universidad peruana?	Determinar la influencia de la aplicación del método Pólya en el desarrollo de las capacidades matemáticas de los estudiantes de una universidad peruana.	La aplicación del método Pólya influye de manera significativa en el desarrollo de las capacidades matemáticas de los estudiantes de una universidad peruana.
VI: método Pólya		VD: capacidades matemáticas
D1: comprender el problema		D1: matematiza situaciones
D2: elaborar un plan		D2: comunica y representa
D3: ejecutar el plan		D3: elabora y usa estrategias
D4: examinar la solución		D4: razona y argumenta

Capacidades matemáticas			
G. Control		G. Experimental	
Pretest	Posttest	Pretest	Posttest
6	13	4	18
4	10	5	16
7	11	8	18
7	11	9	15
4	11	6	19
7	15	9	18
9	11	9	19
11	11	9	20
6	10	6	18
8	11	8	19
6	11	7	18
10	11	7	17
5	11	4	20
10	11	9	19
10	11	13	17
6	11	6	18
7	13	10	16
8	11	10	19
7	12	9	17
7	12	4	16
11	15	10	19
7	11	9	18
10	11	8	20
11	13	5	18
7	11	10	18
7	11	8	18
10	11	10	16
7	11	5	19
11	12	9	18
9	15	14	16

Procedimiento

1. Declarar las variables. Se creará una variable “Grupo” para definir el grupo control y el grupo experimental. Luego se definirá la variable “Competencias matemáticas” pretest y postest.

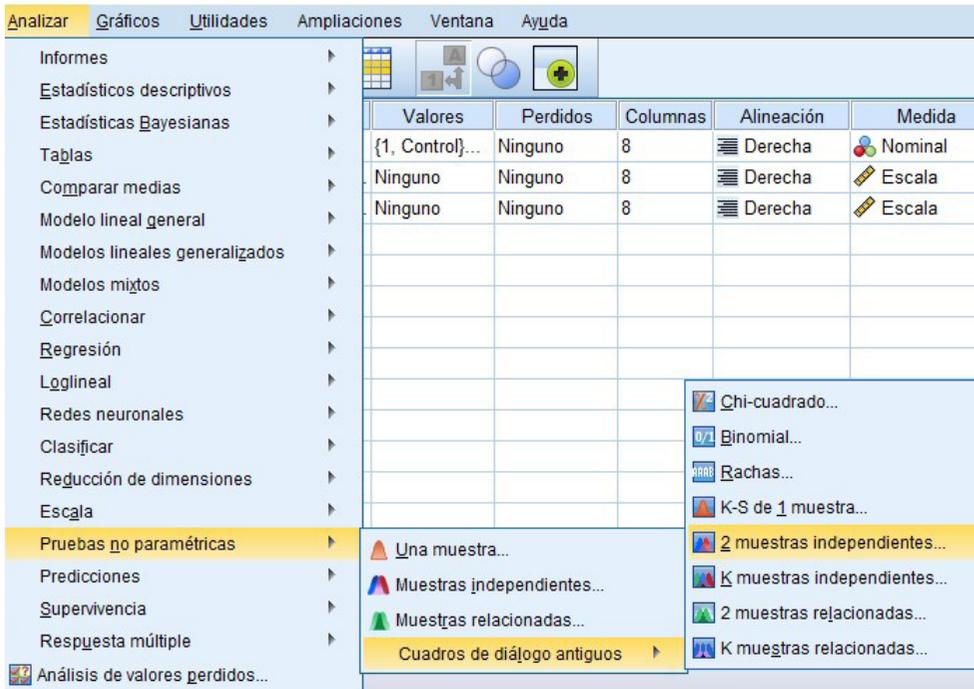


2. En “Vista de datos” ingresar el número “1” para los primeros 30 datos (grupo control). Del 31 al 60 ingresar el número “2” (grupo experimental). Ingresar los valores de la variable para el pretest y el postest, teniendo en cuenta que cada grupo está conformado por 30 estudiantes.

The screenshot shows the SPSS Data View window. The data is organized into columns: 'grupo', 'Vpretest', 'Vpostest', and several 'var' columns. The 'grupo' column contains the following values for rows 1 through 14:

Row	grupo	Vpretest	Vpostest
1	Control	6	16
2	Control	4	13
3	Control	7	14
4	Control	7	13
5	Control	4	14
6	Control	7	18
7	Control	9	14
8	Control	11	13
9	Control	6	13
10	Control	8	14
11	Control	6	14
12	Control	10	14
13	Control	5	14
14	Control	10	13

- Para realizar el análisis estadístico hacemos clic en “Analizar”, “Pruebas no paramétricas”, “Cuadros de diálogo antiguos” y “2 muestras independientes”.



- Trasladar las variables al casillero “Lista variables de prueba”. Luego, trasladar la variable “Grupo” al casillero “Variable de agrupación”. Activar la opción “U de Mann-Whitney”.





5. Clic en “Definir grupos”. Escribir “1” en la casilla “Grupo 1” y “2” en la casilla “Grupo 2”. Clic en “Continuar” y “Aceptar”.



6. Visualización e interpretación

Estadísticos de prueba^a

	Capacidades matemáticas pretest	Capacidades matemáticas posttest
U de Mann-Whitney	441,000	300,500
W de Wilcoxon	906,000	765,500
Z	-,134	-2,247
Sig. asintótica(bilateral)	,893	,025

a. Variable de agrupación: Grupo

En la tabla anterior observamos que en los resultados del postest el valor de la significancia calculada es $0,025 < 0,05$, lo cual indica que la diferencias entre grupo control y experimental es significativa. Por ello, se concluye que la variable independiente (“Método Pólya”) influye de manera significativa en la variable dependiente (“Capacidades matemáticas”).

Prueba T para muestras independientes

Ejemplo 9

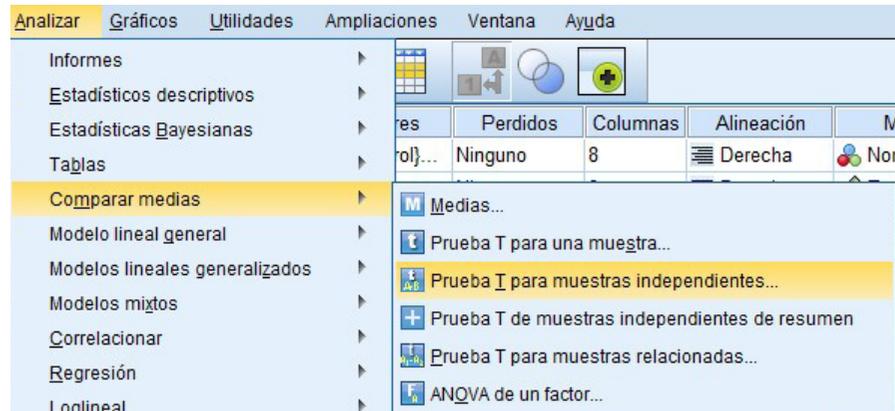
Se tiene el siguiente estudio: se desea aplicar un programa innovador para mejorar la atención en estudiantes de Medicina de una universidad privada de Cajamarca.

Problema	Objetivo	Hipótesis
¿Cómo la aplicación de un programa innovador mejora el nivel de atención de los estudiantes de Medicina de una universidad privada de Cajamarca?	Determinar cómo la aplicación de un programa innovador mejora el nivel de atención de los estudiantes de Medicina de una universidad privada de Cajamarca.	La aplicación de un programa innovador mejora de forma significativa el nivel de atención de los estudiantes de Medicina de una universidad privada de Cajamarca.
VI: programa innovador		VD: nivel de atención
		D1: atención focalizada
		D2: atención sostenida
		D3: atención selectiva

Nivel de atención			
G. Control		G. Experimental	
Pretest	Postest	Pretest	Postest
27	45	12	52
16	38	15	49
29	38	24	51
24	41	29	48
20	41	19	60
31	56	24	52
29	41	29	55
33	38	26	60
20	38	21	50
29	41	24	56
16	44	24	52
29	42	20	51
20	39	12	60
28	41	28	54
27	40	36	50
25	40	21	55
28	48	27	51
28	41	33	54
24	43	24	50
20	45	12	50
29	55	30	56
25	40	26	52
30	41	25	60
34	48	17	54
24	40	31	56

Procedimiento

1. Ingresar los datos y declarar las variables. Recordar que se debe crear la variable “Grupo” para definir el grupo control y el grupo experimental (igual que en el caso anterior).
2. Clic en “Analizar”, “Comparar medias” y “Prueba T para muestras independientes”.



3. Trasladar las variables pretest y postest al casillero “Variables de prueba”. Luego, trasladar la variable “Grupo” al casillero “Variable de agrupación”. Clic en “Definir grupos”. En la casilla “Grupo 1” asignar el valor “1” y en la casilla “Grupo 2” asignar el valor “2”. Clic en “Continuar” y “Aceptar”.



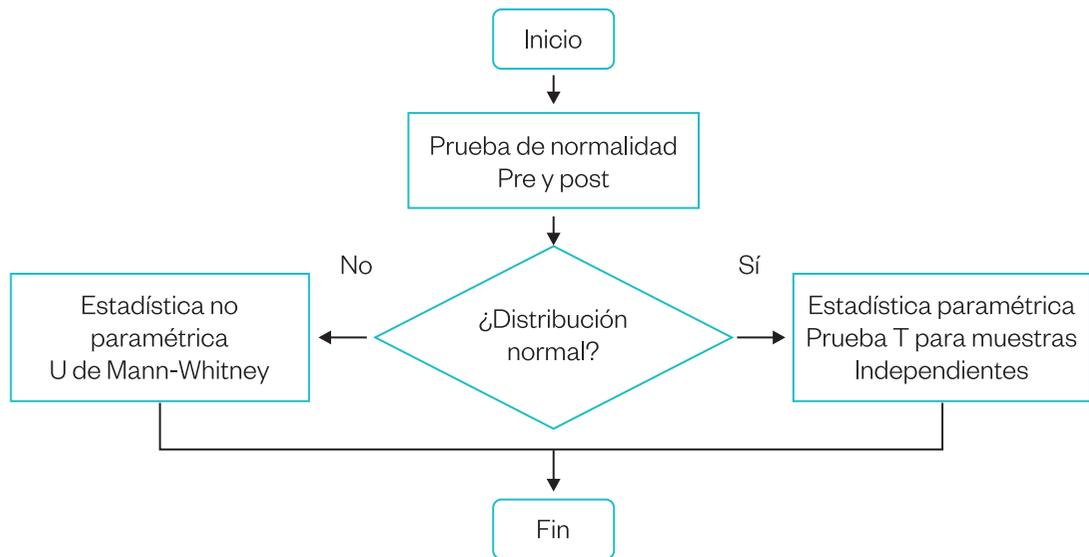
4. Visualización e interpretación.

Prueba de muestras independientes

		Prueba de Levene de igualdad 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
Atención pretest	Se asumen varianzas iguales	1,229	.273	1,380	48	.174	2,240	1,623	-1,023	5,503
	No se asumen varianzas iguales			1,380	44,544	.174	2,240	1,623	-1,029	5,509
Atención posttest	Se asumen varianzas iguales	.378	.542	-9,110	48	.000	-10,960	1,203	-13,379	-8,541
	No se asumen varianzas iguales			-9,110	44,907	.000	-10,960	1,203	-13,383	-8,537

En la tabla se observa que en el posttest el valor de significancia bilateral indica un valor de $0,000 < 0,05$. Por ello, se afirma que hay diferencia significativa entre el grupo control y experimental. Es decir, el programa innovador mejora significativamente el nivel de atención de los estudiantes de Medicina de una universidad privada de Cajamarca.

Análisis estadístico de estudios comparativos: 2 grupos



U de Mann-Whitney (2 muestras independientes)

Ejemplo 10

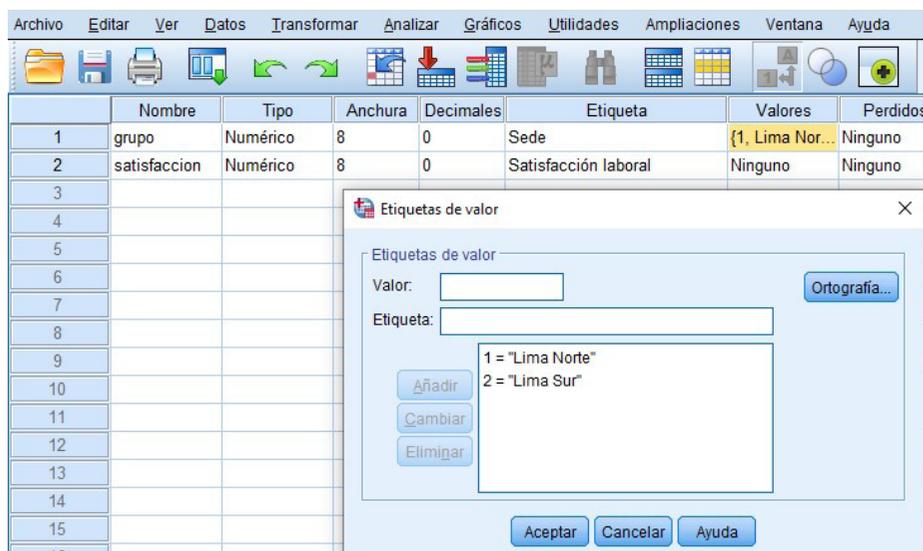
¿Cuál es la diferencia de la satisfacción laboral entre los trabajadores de dos centros de salud privados de Lima?

G1: sede Lima Norte	G2: sede Lima Sur
V: satisfacción laboral	V: satisfacción laboral
D1: relación con sus superiores	D1: relación con sus superiores
D2: condiciones de trabajo	D2: condiciones de trabajo
D3: participación en las decisiones	D3: participación en las decisiones

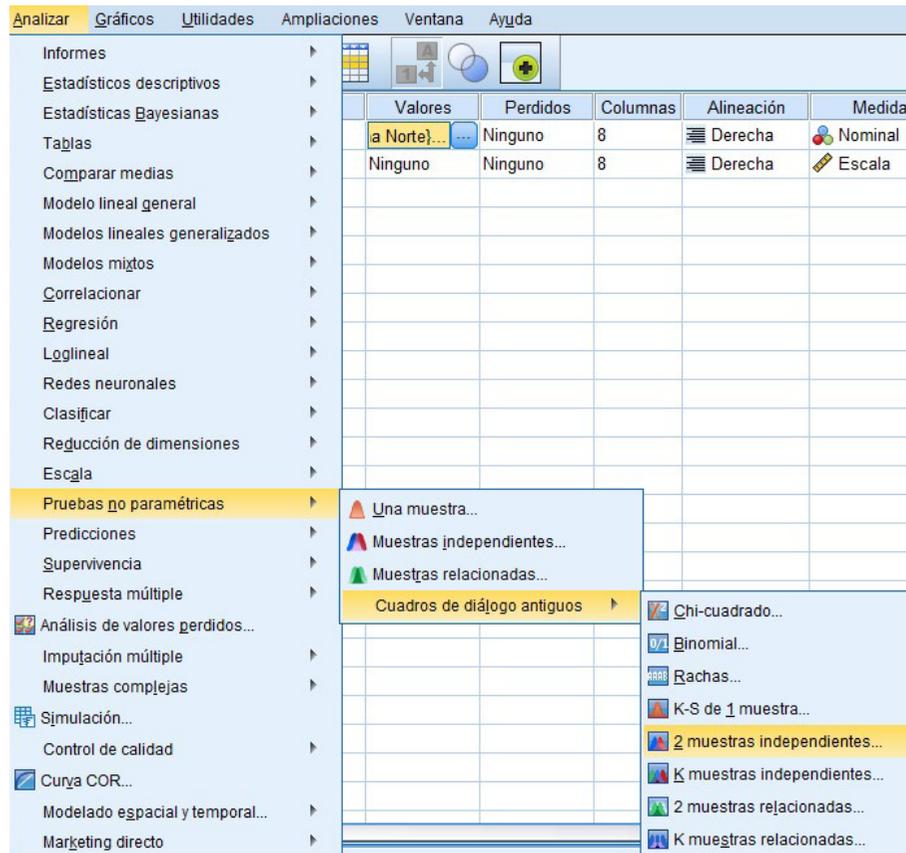
Satisfacción laboral	
Lima Norte	Lima Sur
64	41
66	42
62	42
72	42
52	41
62	41
58	43
63	56
64	44
42	40
55	39
64	48
72	43
67	37
66	43
66	43
63	43
65	39
61	49
65	39

Procedimiento

1. Declarar variables e ingresar datos (se creará una variable nueva para definir los grupos).



2. Clic en “Analizar”, “Pruebas no paramétricas”, “Cuadros de diálogo antiguos” y “2 muestras independientes”.



3. Trasladar la variable al casillero “Lista de variables”. Mover la variable “Grupo” al casillero “Variable de agrupación”. Activar la opción “U de Mann-Whitney”. Clic en “Definir grupos”. Designar con “1” y “2” los casilleros de “Grupo 1” y “Grupo 2”, respectivamente. Clic en “Continuar” y “Aceptar”.



4. Visualización e interpretación.

Estadísticos de prueba^a

	Satisfacción laboral
U de Mann-Whitney	12,500
W de Wilcoxon	202,500
Z	-4,999
Sig. asintótica (bilateral)	,000
Significación exacta [2*(sig. unilateral)]	,000 ^b

a. Variable de agrupación: Sede
b. No corregido para empates

Observamos que el valor de la significancia calculada es $0,000 < 0,05$. Por lo tanto, se comprueba que existe diferencia significativa en la satisfacción laboral entre ambos grupos (sedes) de trabajadores de salud.

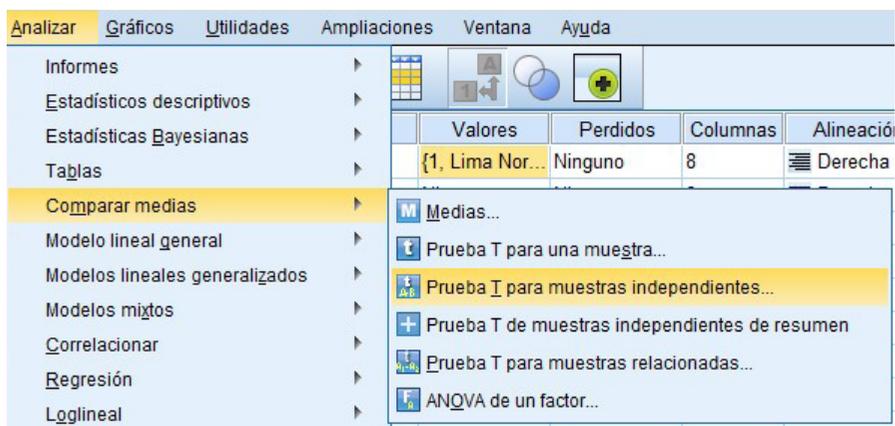
Prueba T (2 muestras independientes)

Ejemplo 11

Consideremos el mismo ejemplo para el análisis utilizando la prueba T.

Procedimiento

1. Clic en “Analizar”, “Comparar medias” y “Prueba T para muestras independientes”.



- Trasladar la variable al casillero de “Vista de variables”. Mover la variable “Grupo” al casillero “Variable de agrupación”. Clic en “Definir”. Designar con “1” y “2” los casilleros de “Grupo 1” y “Grupo 2”, respectivamente. Clic en “Continuar” y “Aceptar”.



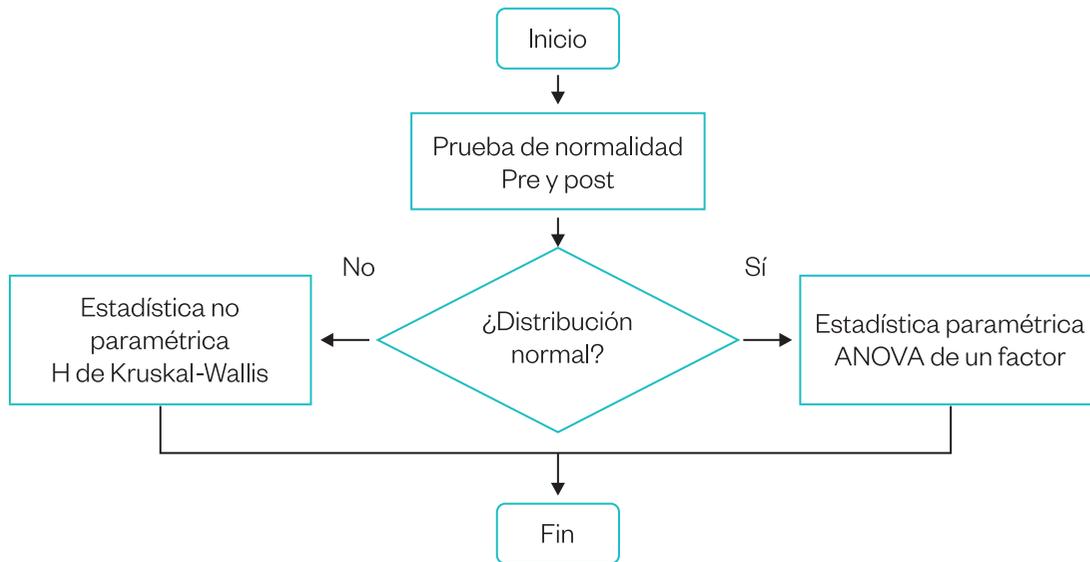
- Visualización e interpretación.

Prueba de muestras independientes

		Prueba de Levene de igualdad 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
Satisfacción laboral	Se asumen varianzas iguales	1,878	,179	11,064	38	,000	19,70000	1,78052	16,09552	23,30448
	No se asumen varianzas iguales			11,064	31,853	,000	19,70000	1,78052	16,07254	23,32746

El valor de la significancia calculada es $0,000 < 0,05$, lo que permite concluir que hay diferencias significativas entre ambos grupos.

Análisis estadístico de estudios comparativos: más de 2 grupos



Prueba H de Kruskal-Wallis

Ejemplo 12

Estrés Académico			
Ciudad			
Lima	Cuzco	Piura	Loreto
12	15	15	12
12	12	12	18
16	14	14	16
20	15	14	13
16	12	13	10
15	18	16	15
16	16	14	12
16	13	14	14
18	10	15	15
20	15	12	16
16	12	12	20
15	14	16	20
20	14	20	20
18	13	16	20
17	16	15	18
19	14	16	20
15	14	18	18
18	15	16	20
18	12	16	16
16	12	18	16

Se ha tomado un test para evaluar el estrés académico de un grupo de estudiantes de cuatro ciudades del Perú. Determinar si existen diferencias significativas entre los grupos.

Procedimiento

1. Ingresar los datos y declarar las variables.

The screenshot shows the SPSS interface with a data list and an open dialog box. The data list has the following columns: Nombre, Tipo, Anchura, Decimales, Etiqueta, Valores, and Perdidos. The first two rows are: 1 grupo Numérico 8 0 Ciudad {1, Lima}... Ningunc; 2 estres_acad... Numérico 8 0 Estrés académico Ninguno Ningunc. The 'Etiquetas de valor' dialog box is open, showing a list of values: 1 = "Lima", 2 = "Cuzco", 3 = "Piura", 4 = "Loreto".

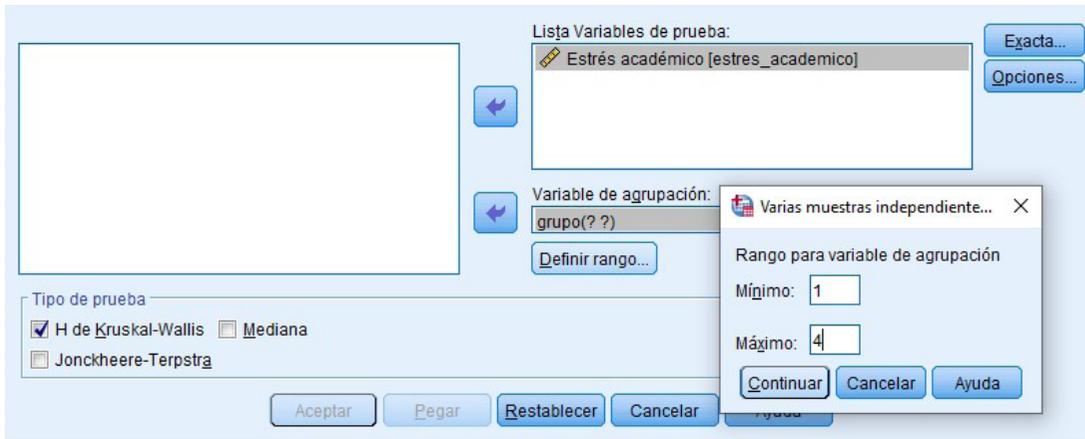
Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos
1 grupo	Numérico	8	0	Ciudad	{1, Lima}...	Ningunc
2 estres_acad...	Numérico	8	0	Estrés académico	Ninguno	Ningunc
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

2. Clic en “Analizar”, “Pruebas no paramétricas”, “Cuadros de diálogo antiguos” y “K muestras independientes”.

The screenshot shows the SPSS 'Analizar' menu open, with the path: Analizar > Pruebas no paramétricas > Cuadros de diálogo antiguos > K muestras independientes... highlighted. The background shows a data list with columns: Valores, Perdidos, Columnas, Alineación, and Medida. The first two rows are: {1, Lima}... Ninguno 8 Derecha Nominal; Ninguno Ninguno 8 Derecha Escala.

Valores	Perdidos	Columnas	Alineación	Medida
{1, Lima}...	Ninguno	8	Derecha	Nominal
Ninguno	Ninguno	8	Derecha	Escala

- Trasladar la variable al casillero “Lista variables de prueba”. Mover la variable “Grupo” al casillero “Variable de agrupación”. Clic en “Definir rango”. Designar con “1” y “4” en los casilleros “Mínimo” y “Máximo”, respectivamente (4 grupos). Clic en “Continuar” y “Aceptar”.



4. Visualización e interpretación

Rangos

	Ciudad	N.º	Rango promedio
Estrés académico	Lima	20	51,55
	Cusco	20	24,80
	Piura	20	37,05
	Loreto	20	48,60
	Total	80	

Estadísticos de prueba^{a,b}

	Estrés académico
H de Kruskal-Wallis	16,929
gl	3
Sig. asintótica	,001

a. Prueba de Kruskal-Wallis
b. Variable de agrupación: Ciudad

El valor de la significancia calculada es $0,001 < 0,05$, lo cual se interpreta como que al menos uno de los grupos es diferente a los demás.

Prueba ANOVA de un factor

Ejemplo 13

Se registró al azar los pesos (g) de recién nacidos en cuatro ciudades de la provincia de Lima. Determinar si existen diferencias significativas entre los grupos.

Peso de los recién nacidos (gr.)			
Ciudades			
Lima	Cañete	Chancay	Huaral
4100	2860	2100	3400
3450	3557	3400	2880
3890	2900	3420	2400
3557	3680	2002	2340
3200	3600	3245	3245
3680	3400	2190	3452
3600	2880	3542	3500
3700	2400	3523	3450
3120	2340	2200	2860
3480	3245	3456	3557
3990	3452	2456	2900
3127	3580	3700	3400
4000	4000	2059	3420
3560	2432	3522	2345
3980	3245	3340	3245
3990	3452	2456	2900
3127	3580	3700	3400
4000	4000	2059	3420
3560	2432	3522	2345
3980	3245	3340	3245

Procedimiento

1. Ingresar los datos y declarar las variables. En la casilla de valores, designar un valor (del 1 al 4) para cada ciudad.

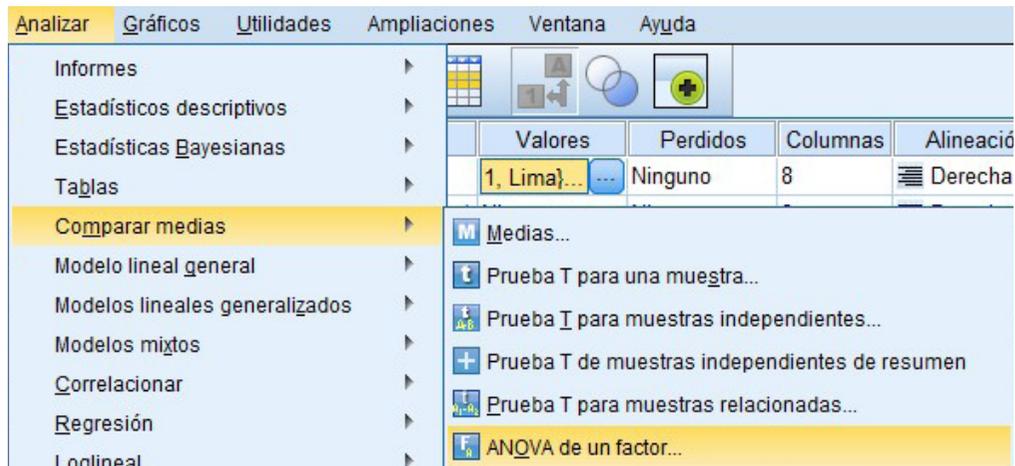
The screenshot shows the SPSS 'Variable View' window. The 'Variable View' table is as follows:

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores
1	grupo	Numérico	8	0	Ciudad	{1, Lima}...
2	pesos	Numérico	8	0	Pesos de recién nacidos (gr.)	Ninguno
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

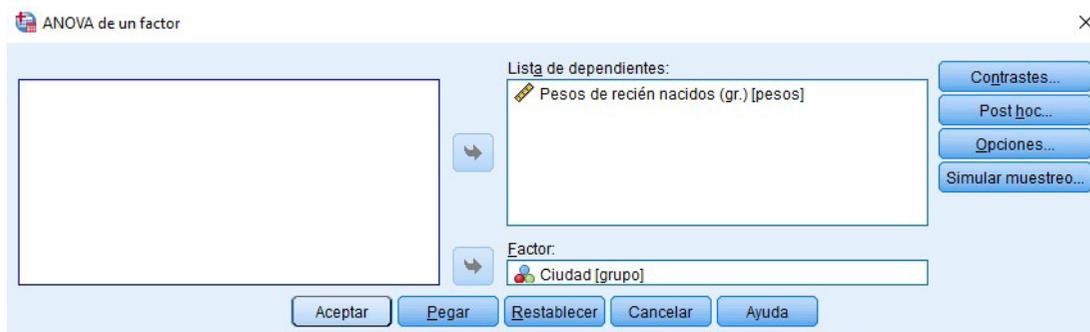
The 'Value Labels' dialog box is open, showing the following configuration:

- Valor:
- Etiqueta:
- Lista de valores: 1 = "Lima", 2 = "Cañete", 3 = "Chancay", 4 = "Huaral"
- Botones: Añadir, Cambiar, Eliminar, Aceptar, Cancelar, Ayuda, Ortografía...

2. Clic en “Analizar”, “Comparar medias” y “ANOVA de un factor”.



3. Trasladar la variable pesos al casillero “Lista de dependientes”. Mover la variable “Ciudad” al casillero “Factor”. Clic en “Aceptar”.



4. Visualización e interpretación.

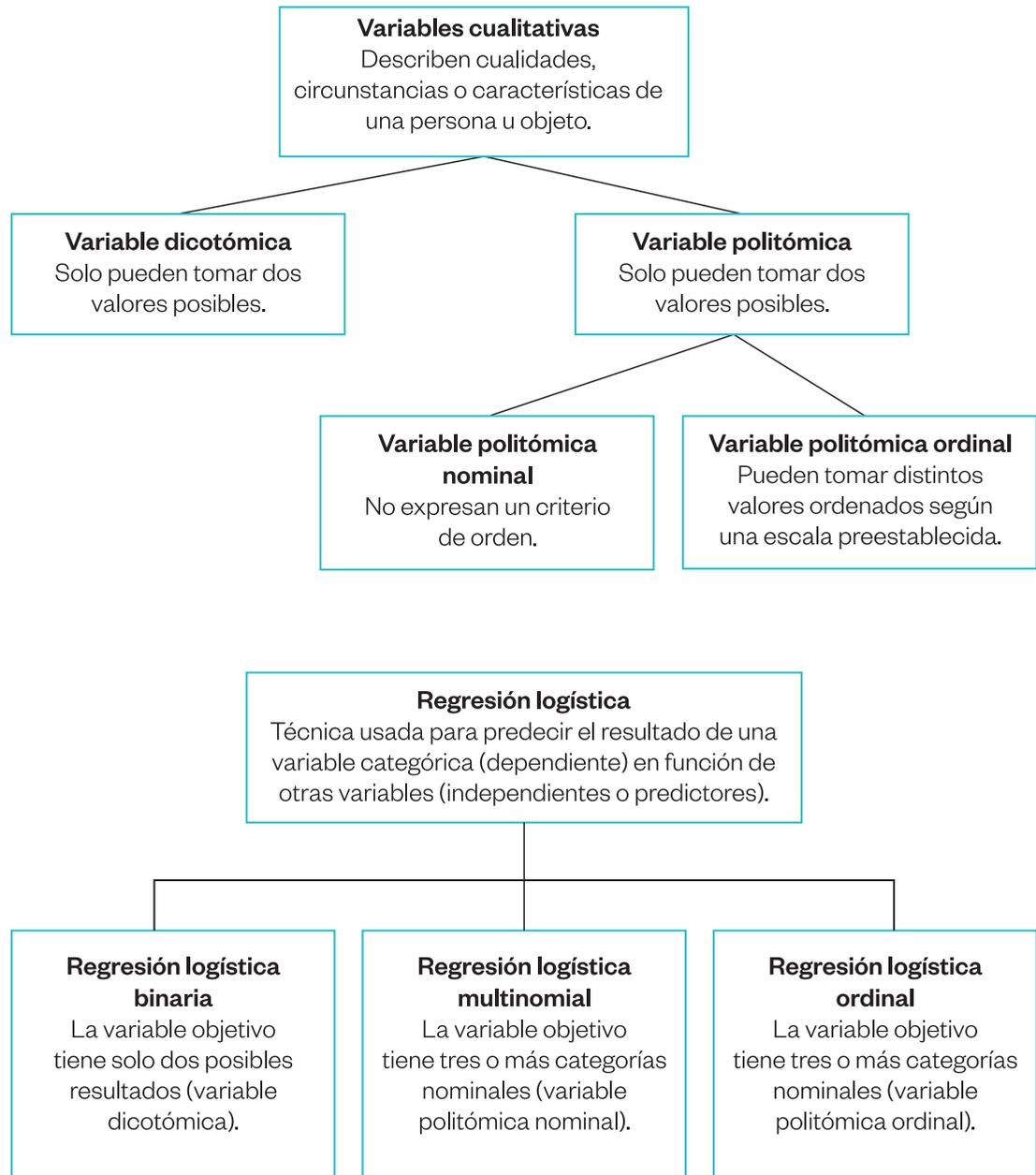
ANOVA

Pesos de recién nacidos (g)					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	5 469 977,938	3	1 823 325,979	7,250	,000
Dentro de grupos	19 112 518,950	76	251 480,512		
Total	24 582 496,888	79			

El valor de la significancia ($0,000 < 0,05$) indica que existe diferencia significativa del peso de los recién nacidos en al menos una de las ciudades.

Análisis estadístico de estudios correlacionales-causales

Análisis estadístico de estudios correlacionales-causales con variables cualitativas a través de modelos de regresión



Regresión logística binaria

Ejemplo 14

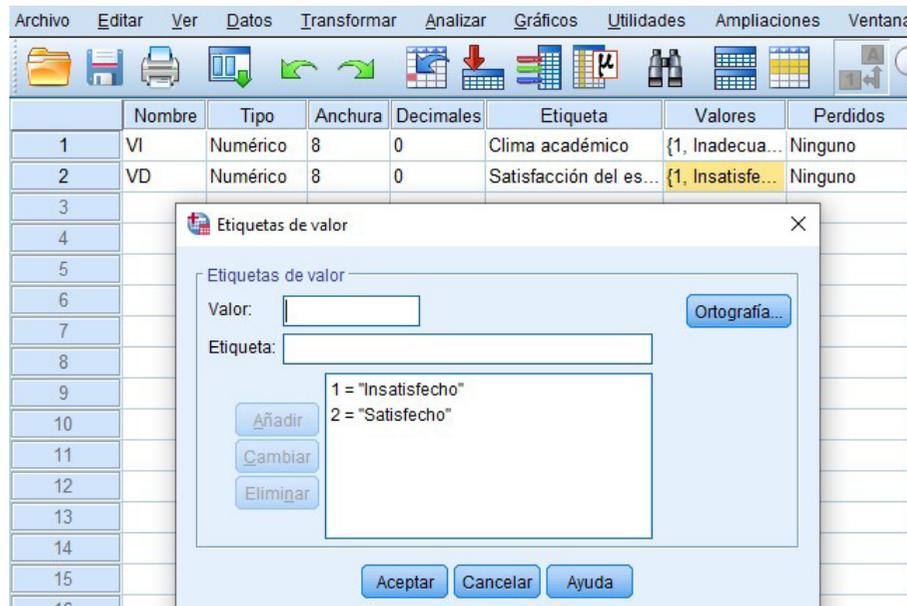
Caso: incidencia del clima académico en la satisfacción de los estudiantes de una universidad peruana.

Problema	Objetivos	Hipótesis
¿Cuál es la incidencia del clima académico en la satisfacción de los estudiantes de una universidad peruana?	Determinar la incidencia del clima académico en la satisfacción de los estudiantes de una universidad peruana.	El clima académico incide significativamente en la satisfacción de los estudiantes de una universidad peruana.

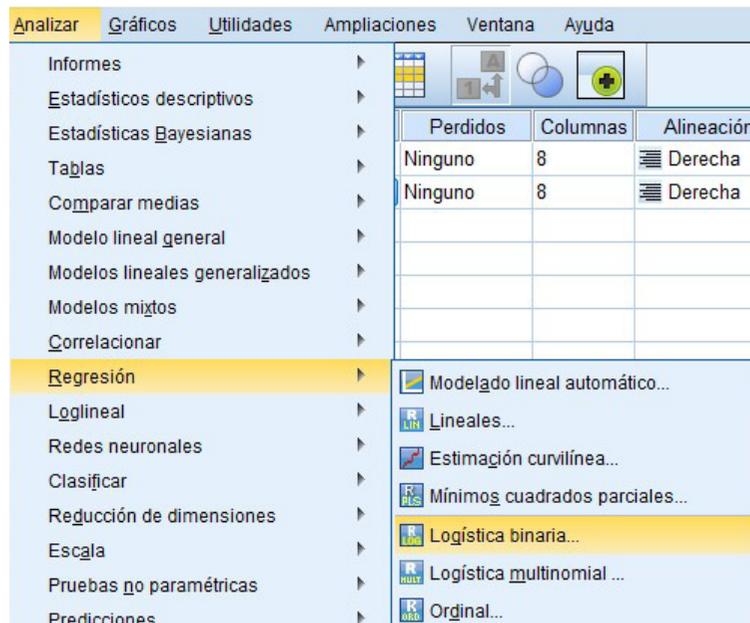
N°	Clima académico VI	Satisfacción del estudiante VD				
1	3	2		Clima académico		
2	3	2		Inadecuado	Regular	Adecuado
3	3	2		1	2	3
4	3	2				
5	3	2				
6	3	2		Satisfacción de estudiante		
7	2	1		Insatisfecho	Satisfecho	
8	2	2		1	2	
9	3	2				
10	3	2				
11	3	1				
12	2	1				
13	2	2				
14	2	1				
15	3	2				
16	3	2				
17	2	2				
18	2	1				
19	2	2				
20	3	2				

Procedimiento

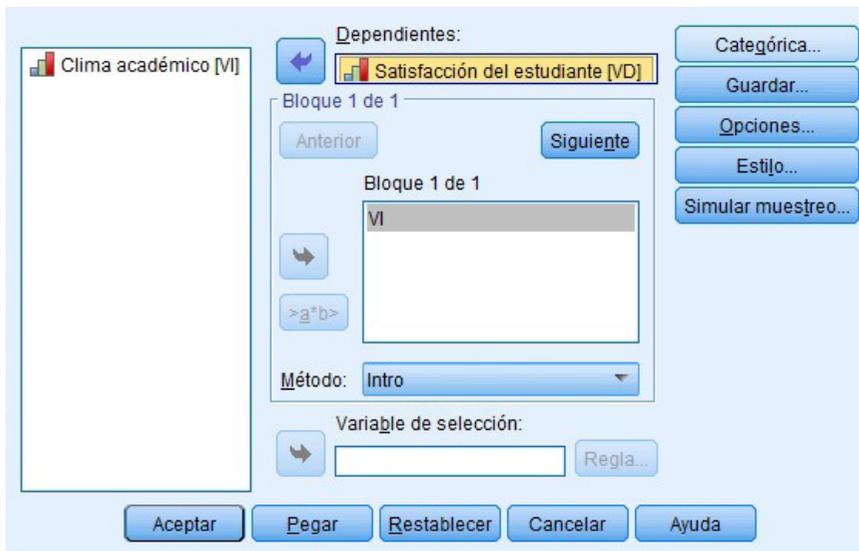
1. Declarar las variables e ingresar los datos.



2. Clic en "Analizar", "Regresión" y "Logística binaria".



3. Trasladar la variable dependiente al casillero de “Dependientes”. Trasladar la(s) variable(s) independiente(s) al casillero de “Bloque 1” de “1”. Clic en “Categoría”.



4. Trasladar la variable independiente al casillero “Covariables categóricas” (si hubiera más variables independientes que no son categóricas, se quedan en el casillero “Covariables”). Clic en “Continuar” y “Aceptar”.



5. Visualización e interpretación. Las tablas que nos ayudan a responder nuestro objetivo o probar la hipótesis son la “Prueba de ómnibus” y el “Resumen del modelo”.

Pruebas ómnibus de coeficientes de modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	4,519	1	,034
	Bloque	4,519	1	,034
	Modelo	4,519	1	,034

El valor de la significancia calculada es $0,034 < 0,05$, lo que indica que el modelo de regresión logística se ajusta al conjunto de datos. Es decir, hay evidencia estadística de una incidencia significativa de la variable independiente (“Clima académico”) sobre la variable dependiente (“Satisfacción del estudiante”).

Resumen del modelo

Paso	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	17,974 ^a	,202	,300

a. La estimación ha terminado en el número de iteración 5 porque las estimaciones de parámetro han cambiado en menos de ,001.

El coeficiente R cuadrado de Nagelkerke indica que la variable independiente explica el 30 % de la variable dependiente. Es decir, el clima académico influye en un 30 % sobre la satisfacción de los estudiantes.

Regresión logística multinomial

Ejemplo 15

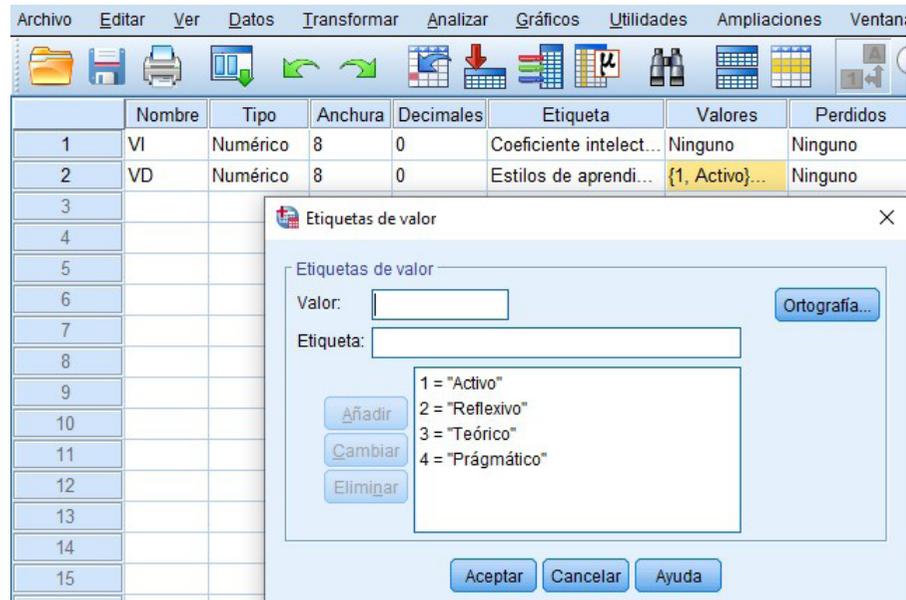
Caso: influencia del coeficiente intelectual en los estilos de aprendizaje de los alumnos de quinto año de una institución educativa de Loreto.

Problema	Objetivos	Hipótesis
¿Cuál es la influencia del coeficiente intelectual en los estilos de aprendizaje de los alumnos de quinto año de una institución educativa de Loreto?	Determinar la incidencia del coeficiente intelectual en los estilos de aprendizaje de los alumnos de quinto año de una institución educativa de Loreto.	El coeficiente intelectual incide significativamente en los estilos de aprendizaje de los alumnos de quinto año de una institución educativa de Loreto.

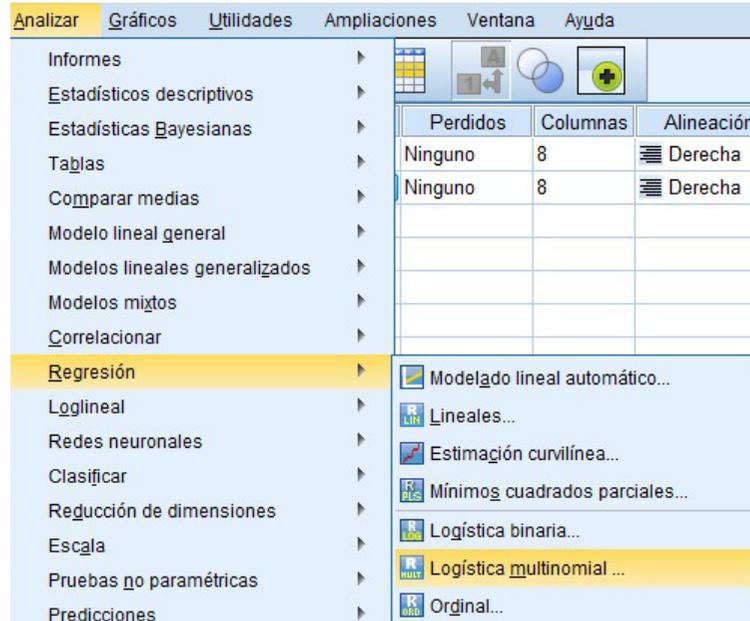
N°	Coeficiente intelectual VI	Estilos de aprendizaje VD	Estilos de aprendizaje			
			Activo	Reflexivo	Teórico	Pragmático
			1	2	3	4
1	120	1				
2	130	2				
3	109	1				
4	110	2				
5	110	2				
6	111	1				
7	122	3				
8	112	3				
9	130	3				
10	112	4				
11	111	4				
12	125	3				
13	124	4				
14	115	3				
15	129	2				
16	111	1				
17	130	2				
18	126	3				
19	130	4				
20	115	4				

Procedimiento

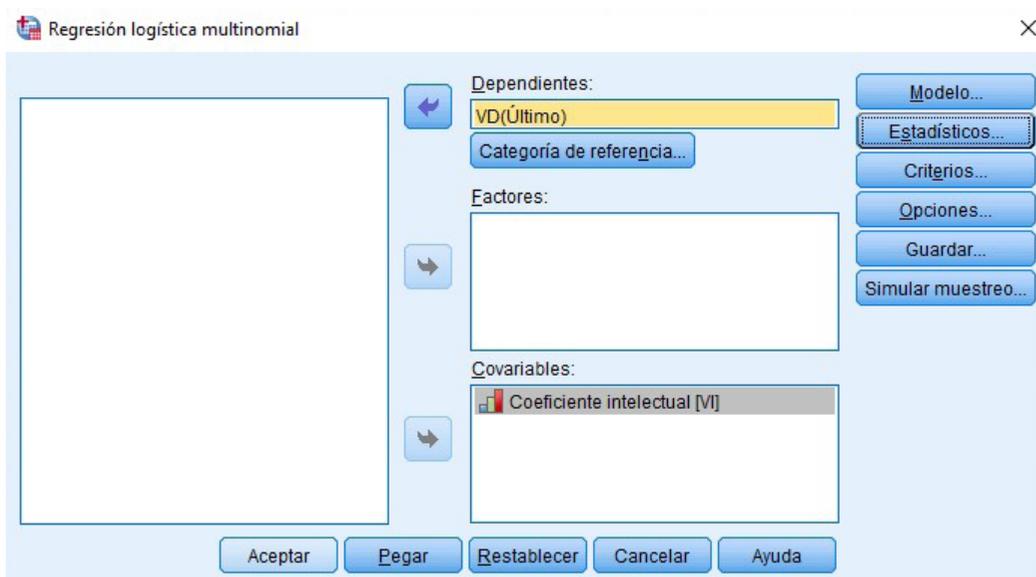
1. Ingreso de datos y declaración de variables.



2. Clic en “Analizar”, “Regresión” y “Logística multinomial”.



3. Trasladar la variable dependiente (“Estilos de aprendizaje”) al casillero “Dependientes”. Luego, trasladar la variable independiente (“Coeficiente intelectual”) al casillero “Covariables”. Finalmente, clic en “Aceptar”.



Nota: como la variable independiente es cuantitativa, se traslada al casillero “Covariables”. Si fuera cualitativa, se debía trasladar al casillero “Factores”.

4. Visualización e interpretación.

Las tablas que responden al objetivo y la prueba de hipótesis son las siguientes:

Información de ajuste de los modelos

Modelo	Crterios de ajuste de modelo	Pruebas de la razón de verosimilitud		
	Logaritmo de la verosimilitud -2	Chi cuadrado	gl	Sig.
Solo intersección	45,109			
Final	40,829	4,280	3	,233

La tabla anterior muestra el valor de la significancia igual a $0,233 > 0,05$, lo cual indica que el modelo no predice significativamente la variable dependiente. Es decir, la variable independiente (“Coeficiente intelectual”) no influye significativamente en la variable dependiente (“Estilos de aprendizaje”).

Pseudo R cuadrado

Cox y Snell	,193
Nagelkerke	,206
McFadden	,078

En la tabla “Pseudo R cuadrado” se muestra el valor del coeficiente de Nagelkerke igual a $0,206$, lo cual indica que el coeficiente intelectual influye en un 20,6 % del estilo de aprendizaje de los alumnos.

Regresión logística ordinal

Ejemplo 16

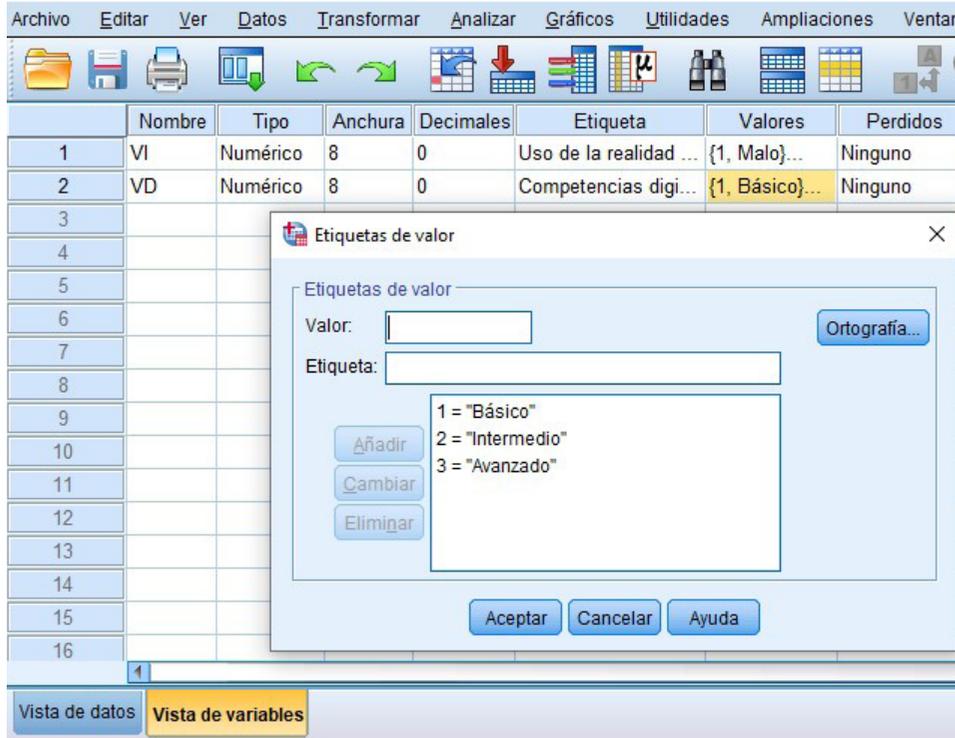
Caso: uso de la realidad aumentada y su incidencia en las competencias digitales de los docentes de la Facultad de Ingeniería Industrial de una universidad peruana.

Problema	Objetivos	Hipótesis
¿Cuál es la incidencia del uso de la realidad aumentada en el desarrollo de las competencias digitales de los docentes de la Facultad de Ingeniería Industrial de una universidad peruana?	Determinar la incidencia del uso de la realidad aumentada en el desarrollo de las competencias digitales de los docentes de la Facultad de Ingeniería Industrial de una universidad peruana.	El uso de la realidad aumentada incide de manera significativa en el desarrollo de las competencias digitales de los docentes de la Facultad de Ingeniería Industrial de una universidad peruana.

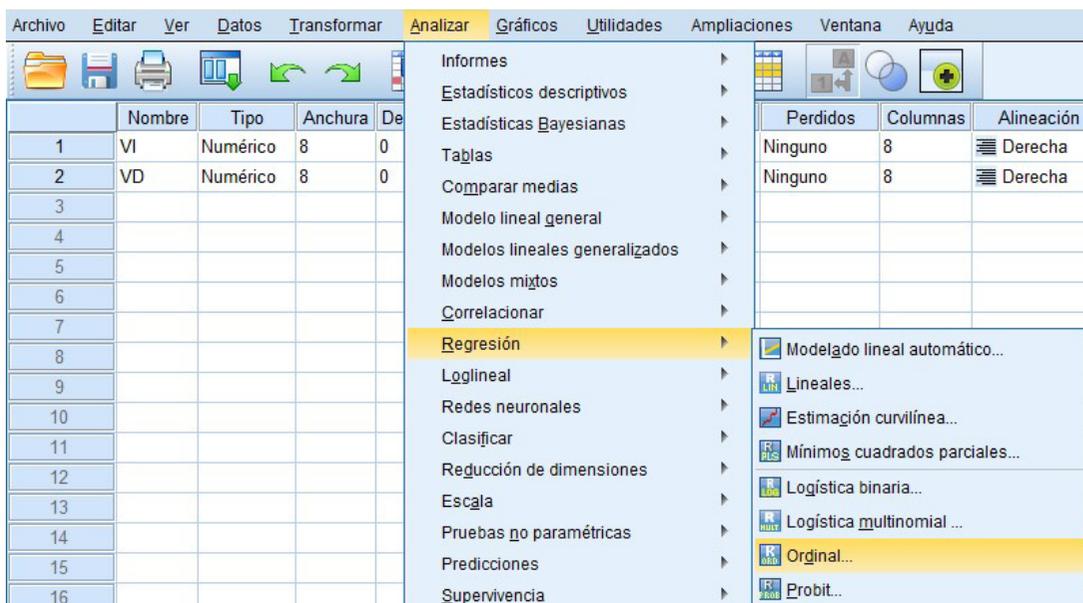
N°	Realidad Aumentada	Competencias Digitales	Uso de la Realidad Aumentada		Competencias Digitales (Variable Objetivo)	
	VI	VD				
1	1	1	Malo	1	Básico	1
2	2	3	Regular	2	Intermedio	2
3	3	3	Bueno	3	Avanzado	3
4	2	2				
5	3	3				
6	1	3				
7	2	2				
8	1	3				
9	1	3				
10	1	1				
11	1	3				
12	3	3				
13	1	2				
14	1	1				
15	3	3				
16	1	3				
17	2	3				
18	2	2				
19	2	3				
20	1	1				
21	2	3				
22	3	3				
23	2	3				
24	3	3				
25	3	3				
26	2	3				
27	2	2				
28	2	2				
29	1	2				
30	1	1				

Procedimiento

1. Ingresar los datos y declarar las variables.



2. Clic en "Analizar", "Regresión" y "Ordinal".

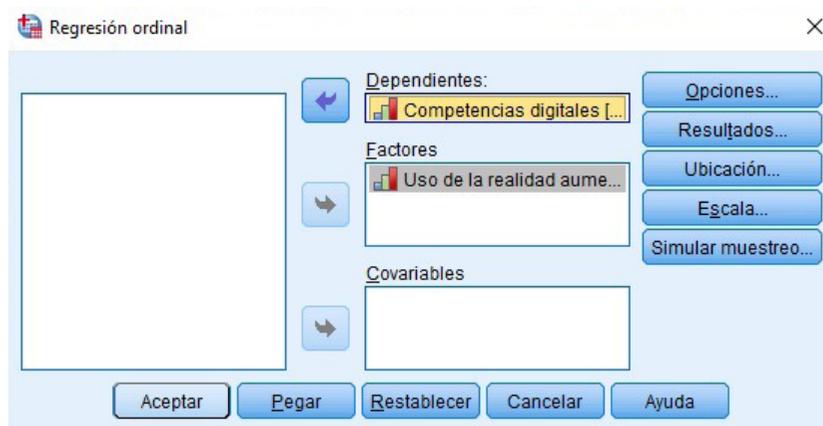


- Trasladar la variable dependiente (“Competencias digitales”) al casillero “Dependientes” y la variable independiente (“Uso de la realidad aumentada”) al casillero “Factores”.

Nota: si la variable independiente fuera numérica o cuantitativa, se moverá al casillero de covariables.



- Clic en “Aceptar”.



- Visualización e interpretación. De todas las tablas, estas dos son las que responderían al objetivo y la hipótesis.

Información de ajuste de los modelos

Modelo	Logaritmo de la verosimilitud -2	Chi cuadrado	gl	Sig.
Solo intersección	24,972			
Final	14,334	10,638	2	,005

Función de enlace: Logit

La tabla anterior muestra que el valor de significancia calculada es $0,005 < 0,05$, lo que indica que la variable independiente (“Uso de la realidad aumentada”) influye significativamente en la variable dependiente (“Competencias digitales”).

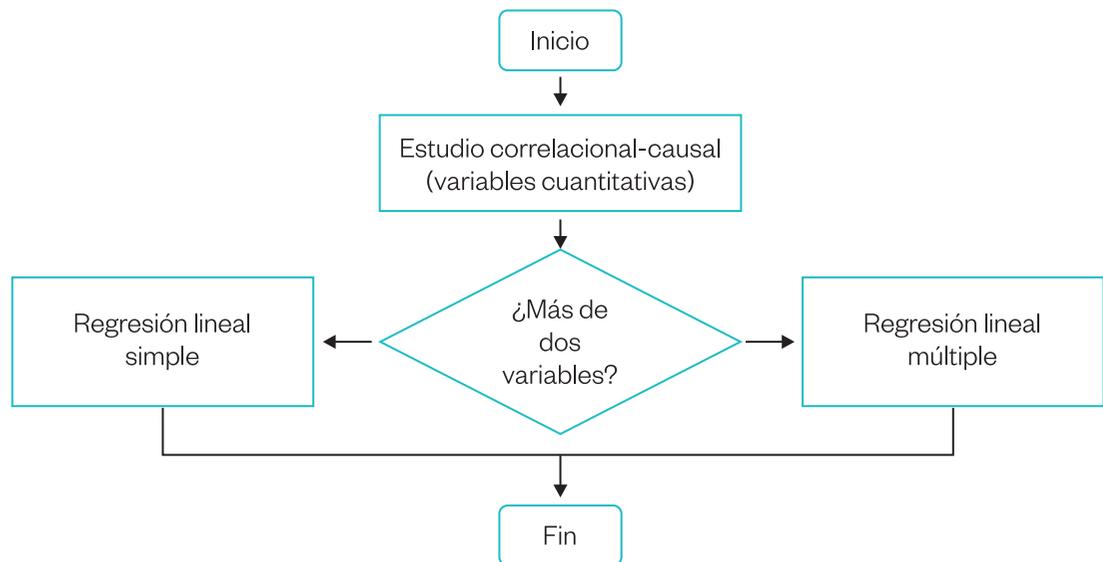
Nivel de incidencia de la variable independiente en la variable dependiente

Pseudo R cuadrado

Cox y Snell	,299
Nagelkerke	,352
McFadden	,188
Función de enlace: Logit	

En la tabla anterior se observa que el coeficiente de Nagelkerke es 0,352, lo cual indica que la variable independiente explica o influye en un 35,2 % sobre la variable dependiente. Es decir, el uso de la realidad aumentada incide en un 35,2 % de las competencias digitales.

Análisis estadístico de estudios correlacionales-causales con variables cuantitativas a través de modelos de regresión



Regresión lineal simple

Examina la relación entre una variable predictora (variable independiente) y una variable de respuesta (variable dependiente).



Un análisis de regresión lineal genera una ecuación (modelo) para describir la relación estadística entre una variable (y) según el valor de otra (x).

$$Y = B + Ax$$

- Y:** valor de la variable dependiente que se desea predecir
- B:** ordenada en el origen
- A:** pendiente o inclinación (razón de cambio)
- X:** valor que fijamos a la variable independiente

Ejemplo 17

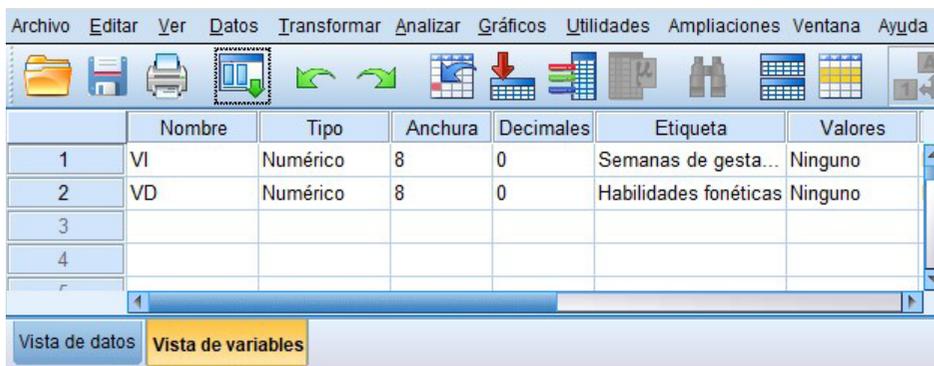
Caso: las semanas de gestación y las habilidades fonéticas en niños entre 6 y 12 años en un centro poblado de la Amazonía.

Problema	Objetivos	Hipótesis
¿Qué tanto influyen las semanas de gestación en las habilidades fonéticas en los niños entre 6 y 12 años de un centro poblado de la Amazonía?	Determinar la influencia de las semanas de gestación en las habilidades fonéticas en los niños entre 6 y 12 años de un centro poblado de la Amazonía.	Las semanas de gestación influyen en las habilidades fonéticas en los niños entre 6 y 12 años de un centro poblado de la Amazonía.

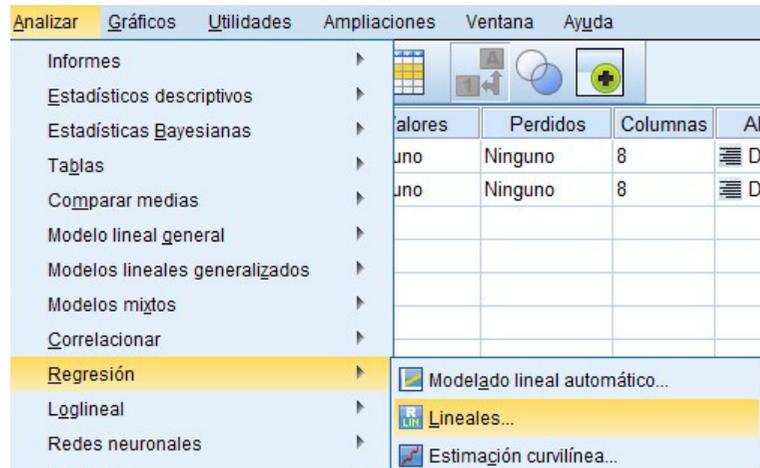
	Habilidades Fonéticas	Semanas de Gestación
1	28	38
2	20	33
3	25	36
4	24	37
5	16	29
6	17	29
7	26	37
8	21	34
9	25	35
10	18	30
11	14	28
12	27	38
13	23	35
14	19	29
15	20	33
16	15	29
17	26	37
18	23	33
19	20	34
20	17	28
21	24	33
22	13	28
23	17	32
24	26	35
25	30	40

Procedimiento

1. Declarar las variables e ingresar los datos.

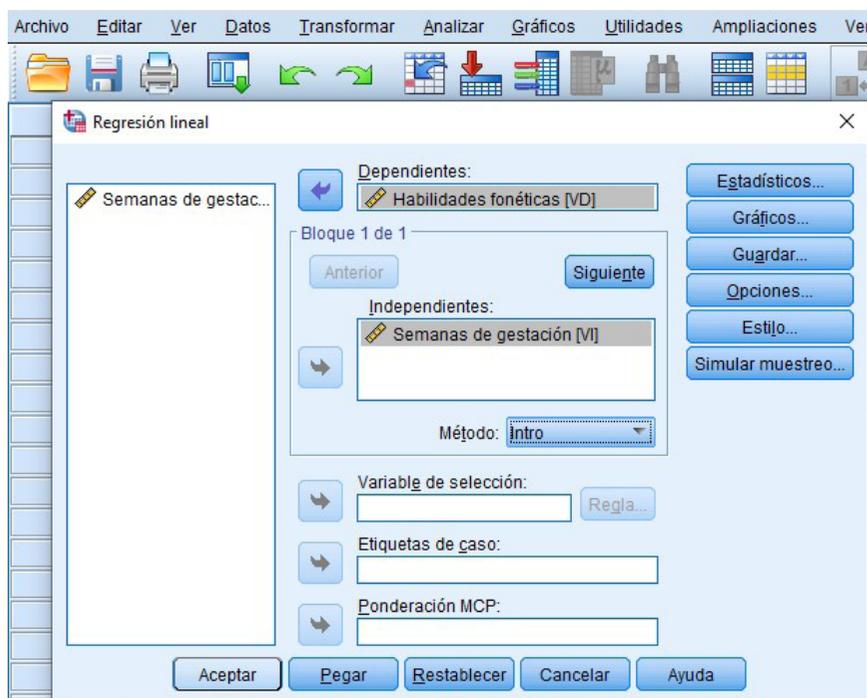


2. Clic en “Analizar”, “Regresión” y “Lineales”.



3. Trasladar la variable independiente (“Semanas de gestación”) al casillero “Independientes” y la variable dependiente (“Habilidades fonéticas”) al casillero “Dependientes”. Luego, clic en “Aceptar”.





4. Visualización e interpretación.

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	,940 ^a	,883	,878	1,270

a. Predictores: (Constante), Semanas de gestación

En la tabla anterior, $R = 0,940$ muestra el grado de relación entre las variables, y el coeficiente de correlación $R^2 = 0,883$ indica que la variable independiente explica o predice el 88,3 % de la variable dependiente.

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	280,882	1	280,882	174,050	,000 ^b
	Residuo	37,118	23	1,614		
	Total	318,000	24			

a. Variable dependiente: Habilidades fonéticas

b. Predictores: (Constante), Semanas de gestación

En la tabla anterior se observa que el valor de significancia calculada es $p = 0,000 < 0,05$, lo cual indica la efectividad que va a arrojar el modelo. Además, el valor de F es alto; es decir, la ecuación es estadísticamente significativa.

Coeficientes^a

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Desv. error	Beta		
1	(Constante)	17,705	1,202		14,734	,000
	Semanas de gestación	,725	,055	,940	13,193	,000

a. Variable dependiente: Habilidades fonéticas

Esta última tabla muestra los coeficientes que nos permitirán construir nuestro modelo (ecuación), donde el valor de la constante es $B = 17,705$ y el coeficiente ($A =$ pendiente de la recta) es $0,725$.

$$Y = B + Ax$$

$$Y = 17,705 + 0,725 X$$

Por ejemplo, calculemos la habilidad fonética (palabras por minuto) de un niño que nació con 37 semanas de gestación:

$$Y = 17,705 + 0,725(37)$$

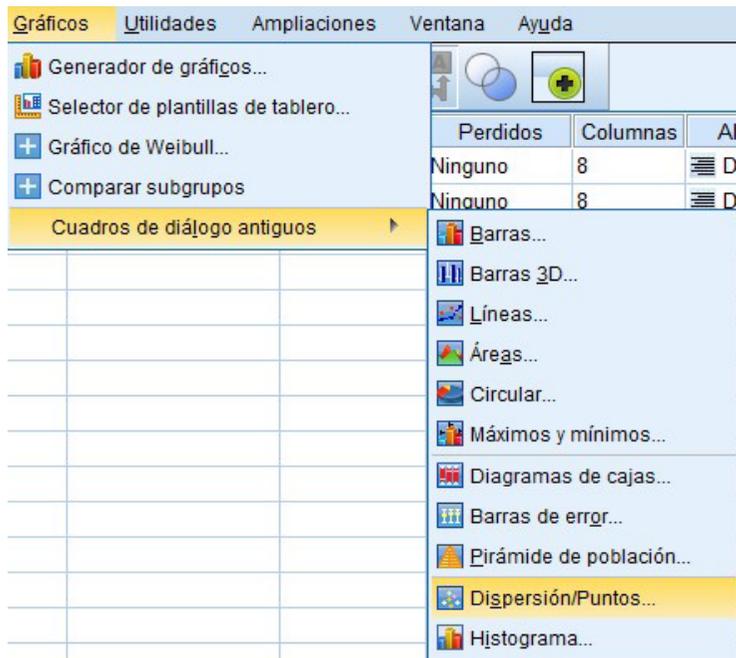
$$Y = 17,705 + 26,825$$

$$Y = 44,53$$

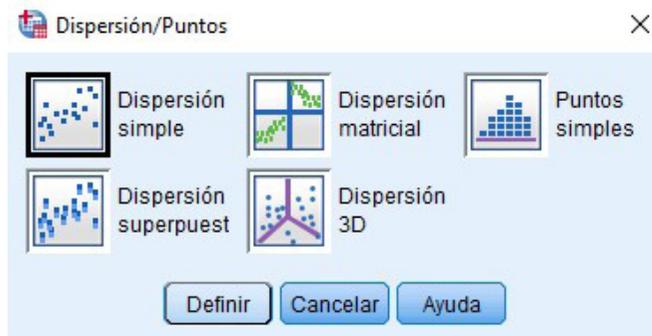
Es decir, un niño que nació con 37 semanas de gestación tendría una fluidez o habilidad fonética aproximada de 44,53 palabras por minuto.

También podemos hacer la gráfica.

5. Clic en “Gráficos”, “Cuadros de diálogo antiguos” y “Dispersión/Puntos”.



6. Clic en el recuadro “Dispersión simple” y “Definir”.



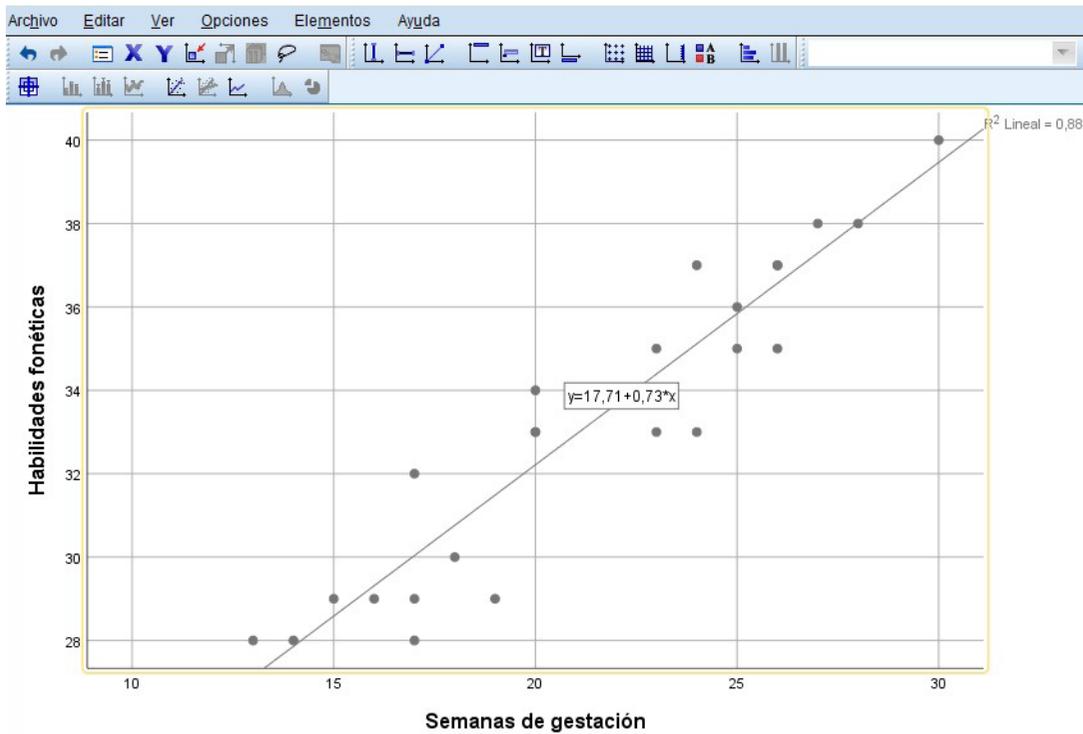
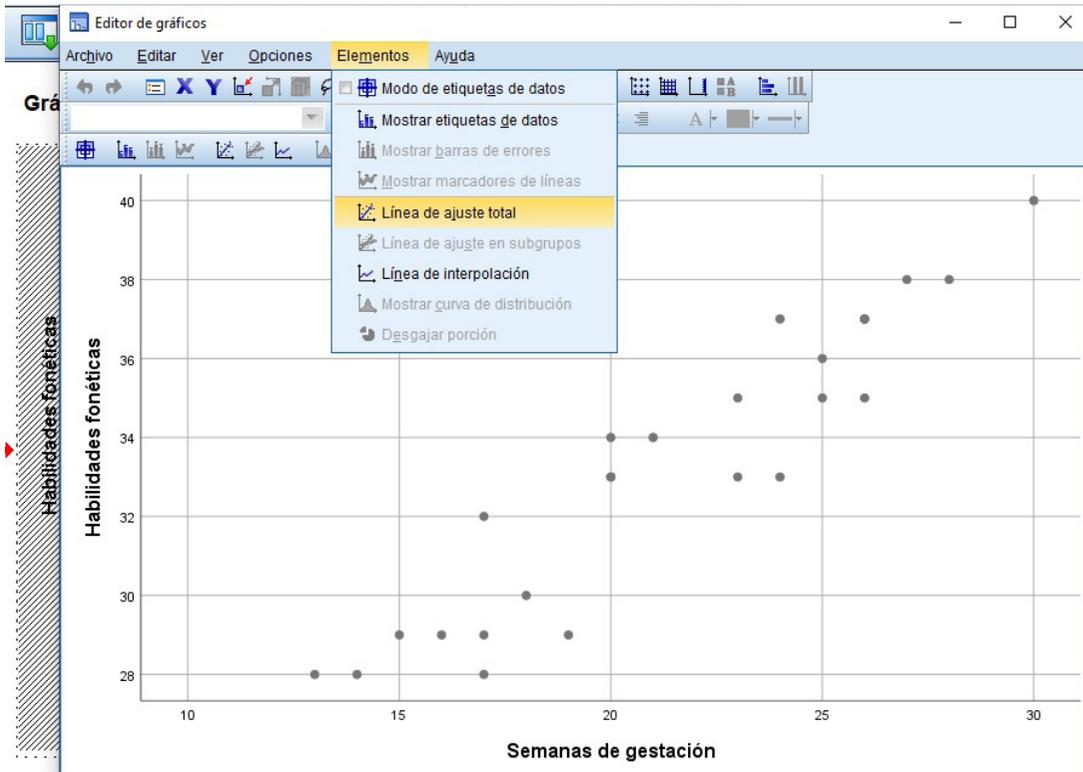
7. Trasladar la variable independiente (“Semanas de gestación”) al casillero del eje X y la variable dependiente (“Habilidades fonéticas”) al casillero del eje Y.

The screenshot shows the 'Eje Y' dialog box in SPSS. On the left, a list of variables includes 'Semanas de gestación' and 'Habilidades fonéticas'. The 'Eje Y:' field contains 'Habilidades fonéticas [VD]' and the 'Eje X:' field contains 'Semanas de gestación [VI]'. Below these are fields for 'Establecer marcas por:', 'Etiquetar los casos mediante:', and a 'Panel mediante' section with 'Filas:' and 'Columnas:' fields. At the bottom, there is a 'Plantilla' section with a checkbox 'Usar las especificaciones gráficas de:' and an 'Archivo...' button. The main buttons at the bottom are 'Aceptar', 'Pegar', 'Restablecer', 'Cancelar', and 'Ayuda'.

8. Clic en “Aceptar”.

This screenshot shows the same 'Eje Y' dialog box after the 'Aceptar' button has been clicked. The variable 'Habilidades fonéticas [VD]' is now highlighted in the 'Eje Y:' field, and 'Semanas de gestación [VI]' is highlighted in the 'Eje X:' field. The rest of the dialog box, including the 'Panel mediante' and 'Plantilla' sections, remains the same as in the previous screenshot.

9. Visualización e interpretación. Doble clic en el gráfico. Aparecerá otro cuadro. Clic en “Elementos” y “Línea de ajuste total”.

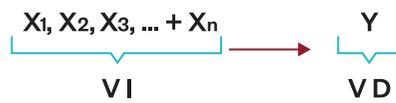


En la tabla anterior observamos, según el comportamiento de la gráfica, que se genera un patrón o tendencia a un modelo en particular. Aparece también la ecuación del modelo y el R cuadrado, que son los mismos que ya se habían observado en la tabla.

Regresión lineal múltiple

Examina la relación entre dos o más variables predictoras (variables independientes) y una variable de respuesta (variable dependiente).

Un análisis de regresión lineal múltiple genera una ecuación (modelo) para describir la relación estadística entre dos o más predictores (x_1, x_2, \dots) y la variable de respuesta (y)



VI: Variable independiente (X)

VD: Variable dependiente (Y)

$$Y = B + A_1X_1 + A_2X_2 + \dots + A_nX_n$$

Y: Valor de la variable dependiente que se desea predecir

B: Ordenada en el origen

A: Pendiente o inclinación (razón de cambio)

X: Valor que fijamos a las variables independientes

Ejemplo 18

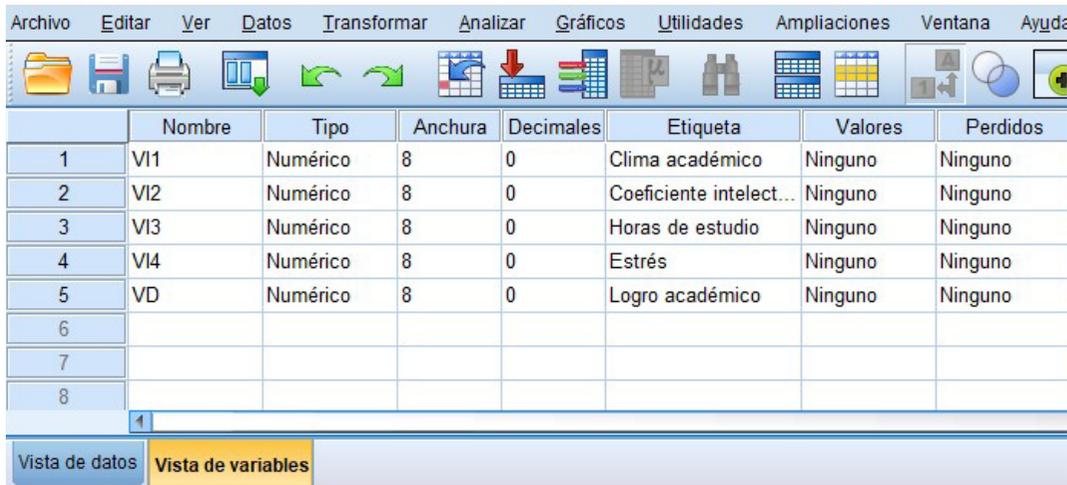
Caso: factores asociados que influyen en el rendimiento académico de los estudiantes de una universidad pública de Lima.

Problema	Objetivos	Hipótesis
¿Cómo influyen los factores asociados en el rendimiento académico de los estudiantes de una universidad pública de Lima?	Determinar la influencia de los factores asociados en el rendimiento académico de los estudiantes de una universidad pública de Lima.	Los factores asociados influyen significativamente en el rendimiento académico de los estudiantes de una universidad pública de Lima.

Variable dependiente (de respuesta)	Variable independiente (predictores)
Rendimiento académico	D1: clima académico
	D2: coeficiente intelectual
	D3: horas de estudio
	D4: estrés

Procedimientos

1. Declarar las variables e ingresar los datos.

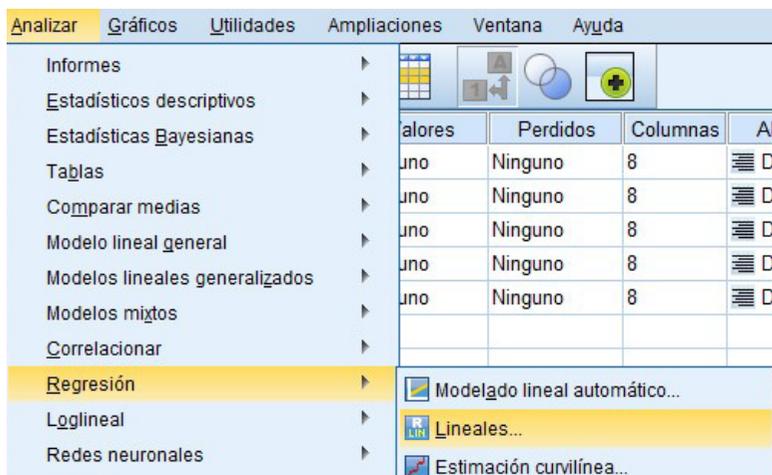


The screenshot shows the 'Variable View' window in SPSS. The menu bar includes Archivo, Editar, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, Ampliaciones, Ventana, and Ayuda. The toolbar contains icons for file operations, data manipulation, and analysis. The main area is a table with the following columns: Nombre, Tipo, Anchura, Decimales, Etiqueta, Valores, and Perdidos. The table contains five rows of variable definitions.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos
1	VI1	N Numérico	8	0	Clima académico	Ninguno	Ninguno
2	VI2	N Numérico	8	0	C Coeficiente intelect...	Ninguno	Ninguno
3	VI3	N Numérico	8	0	H Horas de estudio	Ninguno	Ninguno
4	VI4	N Numérico	8	0	E Estrés	Ninguno	Ninguno
5	VD	N Numérico	8	0	L Logro académico	Ninguno	Ninguno
6							
7							
8							

At the bottom, there are two buttons: 'Vista de datos' and 'Vista de variables', with 'Vista de variables' being the active view.

2. Clic en “Analizar”, “Regresión” y “Lineales”.



3. Trasladar la variable independiente (“Logro académico”) al casillero “Dependientes” y las variables independientes al casillero “Independientes”. Clic en “Aceptar”.



4. Visualización e interpretación.

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	,803 ^a	,644	,573	1,793

a. Predictores: (Constante), Estrés, Clima académico, Horas de estudio, Coeficiente intelectual

En la tabla anterior $R = 0,803$ muestra que el grado de relación entre las variables es muy bueno (cerca del 1). A su vez, el coeficiente de correlación R cuadrado = $0,803$ indica que las variables predictoras o independientes explican o predicen el 80,3 % de la variable dependiente.

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	116,372	4	29,093	9,054	,000 ^b
	Residuo	64,268	20	3,213		
	Total	180,640	24			

a. Variable dependiente: Logro académico

b. Predictores: (Constante), Estrés, Clima académico, Horas de estudio, Coeficiente intelectual

En la tabla ANOVA anterior se observa que el valor de significancia calculada es $p = 0,000 < 0,05$, lo cual indica la efectividad del modelo, aunque el valor de F es bajo.

Coeficientes^a

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Desv. error	Beta		
1	(Constante)	-8,858	7,608		-1,164	,258
	Clima académico	,364	,516	,098	,706	,488
	Coficiente intelectual	,199	,073	,591	2,723	,013
	Horas de estudio	,037	,164	,048	,228	,822
	Estrés	-,699	,328	-,304	-2,132	,046

a. Variable dependiente: Logro académico

La tabla muestra los parámetros para construir el modelo (ecuación), donde el valor de la constante es $B = -8,858$ y los valores de la primera columna indican la contribución de cada predictor al modelo. Es decir, indican la relación entre la variable de interés con los predictores: Clima académico $X_1 = 0,364$; Coeficiente intelectual $X_2 = 0,199$; Horas de estudio $X_3 = 0,037$; y Estrés $X_4 = -0,328$.

Esto se indica en que, por ejemplo, por cada unidad de aumento del “Clima académico”, el “Logro académico” aumentará en 0,364 unidades, siempre y cuando las demás variables se mantengan constantes; y así para cada variable o predictor.

También se observa que el “Coeficiente intelectual” ($\text{sig} = 0,013 < 0,05$) y el estrés ($\text{sig} = 0,046 < 0,05$) aportan significativamente al modelo; mientras que los demás predictores como “Clima académico” ($\text{sig} = 0,488 > 0,05$) y “Horas de estudio” ($\text{sig} = 0,822 > 0,05$) no tienen un aporte significativo.

Construimos el modelo.

$$Y = B + A_1X_1 + A_2X_2 + \dots + A_nX_n$$

$$Y = -8,858 + 0,364X_1 + 0,199X_2 + 0,037X_3 - 0,328X_4$$

Sección 7

Introducción a los modelos multinivel

Regresión lineal multinivel

¿Qué es?

Es un modelo para analizar datos cuando los casos están agrupados en macrovariables y variables a nivel más bajo (directamente de los casos).

¿Por qué es importante?

Permite diferenciar los distintos niveles jerárquicos independientes, separando la variabilidad existente en los grupos de los participantes del estudio.

Ejemplo 18

Caso: factores asociados en los síntomas depresivos de pacientes de centros de salud públicos y privados del Perú, teniendo en cuenta la influencia de la variabilidad de las variables de cada centro hospitalario.

Tamaño de muestra: 200 pacientes.

Se utilizó la base de datos de síntomas depresivos recolectada en el Perú empleando el instrumento Patient Health Questionnaire de 9 ítems (PHQ-9) en escala de Likert de cuatro niveles.

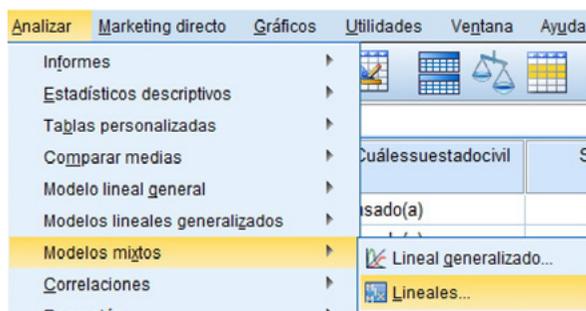
Problema	Objetivos	Hipótesis
¿Cómo influyen los factores asociados en los síntomas depresivos de pacientes de centros de salud públicos y privados del Perú, teniendo en cuenta la variabilidad de indicadores según centros de salud?	Determinar la influencia de los factores individuales en los síntomas depresivos de pacientes de centros de salud del Perú, teniendo en cuenta la influencia de variables del centro hospitalario.	Los factores individuales influyen significativamente en los síntomas depresivos de los pacientes de centros de salud del Perú.

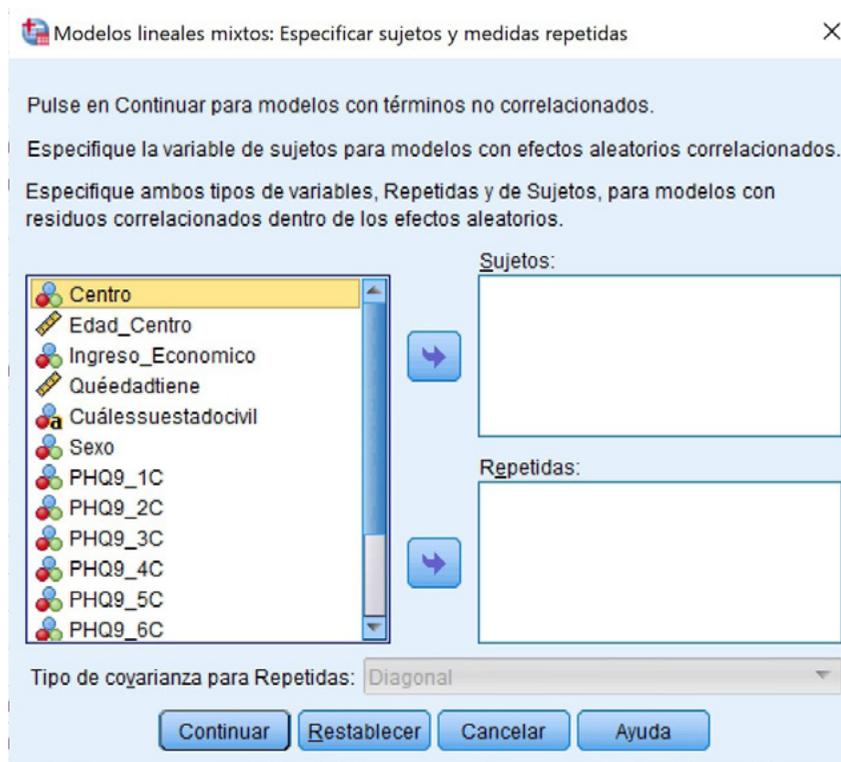
Variable dependiente (de respuesta)	Variable independiente (predictores)
Síntomas depresivos (puntajes)	Nivel 1: variables individuales
	Sexo
	Edad
	Ingreso económico
	Nivel 2: macrovariables del centro de salud
	Tipo de centro de salud
	Edad promedio de los pacientes del centro de salud

1. Edad_Centro	Centro	Edad_Centro	Ingreso_Economico	Quéadatiene	Cuálesuestadocivil	Sexo	PHQ9_1C	PHQ9_2C	PHQ9_3C	PHQ9_4C	PH
1	PUBLICO	28.3799991607666	950 - 2000	40	Casado(a)	Femenino	Más de la mita...	Más c...			
2	PUBLICO	28.38	950 - 2000	37	Casado(a)	Femenino	Más de la mita...	Más c...			
3	PUBLICO	28.38	950 - 2000	28	Soltero (a)	Masculino	Más de la mita...	Casi todos los ...	Casi todos los ...	Casi todos los ...	Más c...
4	PUBLICO	28.38	menor 950	37	Casado(a)	Femenino	Casi todos los ...	Casi todos los ...	Para nada	Casi todos los ...	Casi t...
5	PUBLICO	28.38	menor 950	52	Conviviente	Femenino	Para nada	Varios días	Para nada	Para nada	Casi t...
6	PUBLICO	28.38	2000 a más	45	Casado(a)	Femenino	Más de la mita...	Casi t...			
7	PUBLICO	28.38	950 - 2000	21	Soltero (a)	Femenino	Casi todos los ...	Casi todos los ...	Casi todos los ...	Más de la mita...	Casi t...
8	PUBLICO	28.38	2000 a más	30	Soltero (a)	Masculino	Más de la mita...	Más c...			
9	PUBLICO	28.38	2000 a más	32	Soltero (a)	Femenino	Más de la mita...	Casi todos los ...	Casi todos los ...	Varios días	Casi t...
10	PUBLICO	28.38	menor 950	20	Soltero (a)	Femenino	Casi todos los ...	Varios días	Para nada	Varios días	Casi t...
11	PUBLICO	28.38	950 - 2000	23	Soltero (a)	Femenino	Varios días	Casi todos los ...	Casi todos los ...	Varios días	Casi t...
12	PUBLICO	28.38	2000 a más	37	Casado(a)	Masculino	Más de la mita...	Casi t...			
13	PUBLICO	28.38	2000 a más	25	Conviviente	Femenino	Varios días	Casi todos los ...	Varios días	Casi todos los ...	Casi t...
14	PUBLICO	28.38	950 - 2000	27	Conviviente	Femenino	Varios días	Más de la mita...	Para nada	Más de la mita...	Más c...
15	PUBLICO	28.38	950 - 2000	40	Conviviente	Femenino	Varios días	Casi todos los ...	Varios días	Casi todos los ...	Casi t...
16	PUBLICO	28.38	2000 a más	22	Soltero (a)	Femenino	Varios días	Casi todos los ...	Casi todos los ...	Casi todos los ...	Casi t...
17	PUBLICO	28.38	950 - 2000	18	Soltero (a)	Femenino	Más de la mita...	Más de la mita...	Más de la mita...	Casi todos los ...	Más c...
18	PUBLICO	28.38	2000 a más	39	Casado(a)	Femenino	Casi todos los ...	Casi todos los ...	Más de la mita...	Casi todos los ...	Casi t...
19	PUBLICO	28.38	menor 950	19	Soltero (a)	Femenino	Varios días	Más de la mita...	Más de la mita...	Para nada	Más c...
20	PUBLICO	28.38	menor 950	36	Casado(a)	Femenino	Casi todos los ...	Varios días	Casi todos los ...	Casi todos los ...	Casi t...
21	PUBLICO	28.38	2000 a más	37	Casado(a)	Femenino	Casi todos los ...	Casi todos los ...	Casi todos los ...	Para nada	Casi t...
22	PUBLICO	28.38	950 - 2000	20	Soltero (a)	Femenino	Más de la mita...	Casi todos los ...	Varios días	Para nada	Casi t...
23	PUBLICO	28.38	2000 a más	23	Soltero (a)	Femenino	Para nada	Para nada	Varios días	Casi todos los ...	Casi t...
24	PUBLICO	28.38	950 - 2000	35	Conviviente	Femenino	Casi todos los ...	Varios días	Casi todos los ...	Casi todos los ...	Casi t...

Procedimiento

1. Utilizamos la base de datos de síntomas depresivos recolectada.
2. Primer paso: modelamos a nivel del centro de salud (Modelo Nulo) y seleccionamos la opción “Analizar”. Clic en “Modelos mixtos”. Seleccionamos la opción “Lineales”. Aparece una emergente. No seleccionamos ninguna variable y damos clic en “Continuar”.



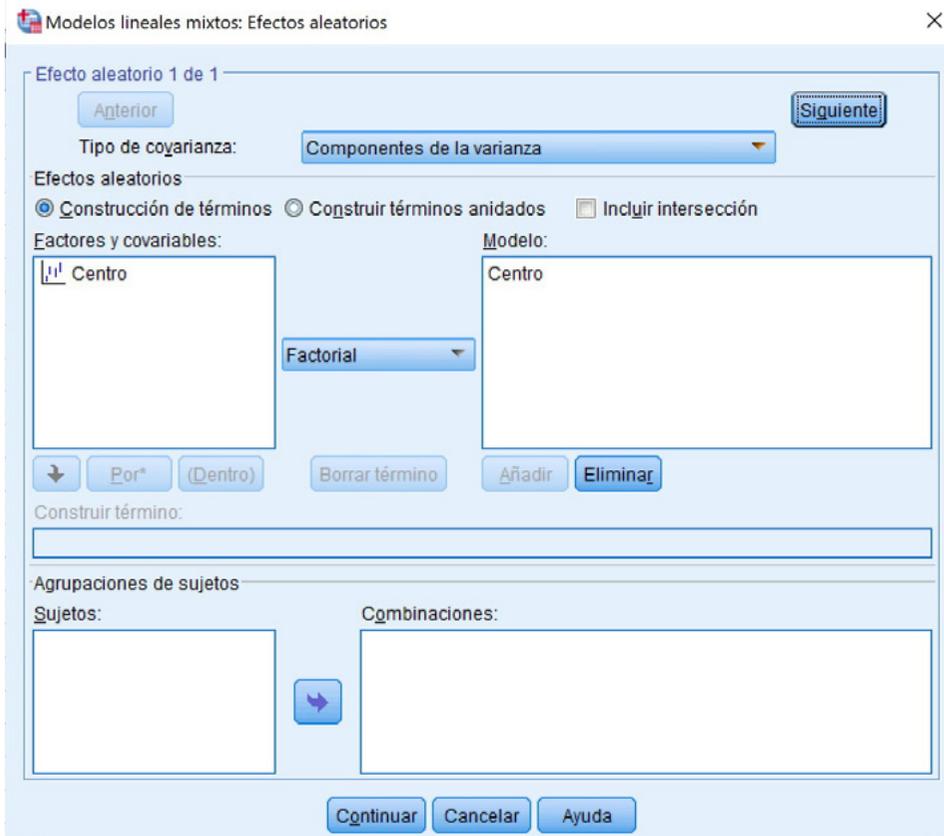


- Colocamos la variable “Puntaje de síntomas depresivos” en “Variable dependiente” (Código: Ptj_SD) y la variable de tipo de centro en “Factores” (Código: Centro).

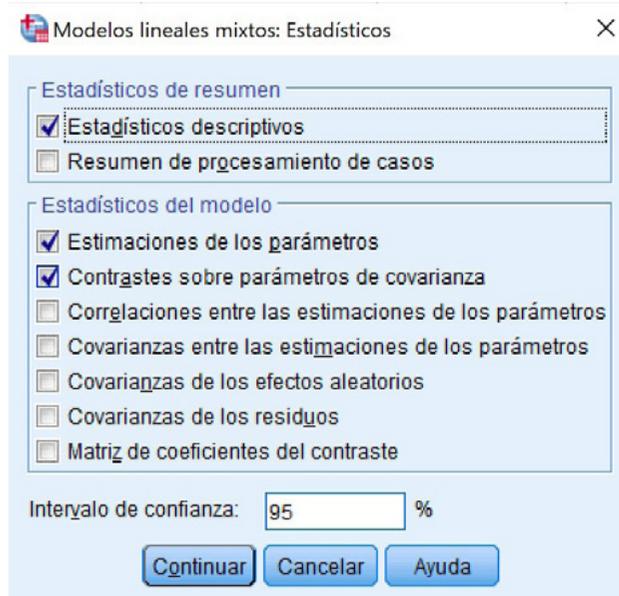


Adicionalmente:

Seleccionamos la opción “Aleatoria” para identificar la variable aleatoria, que en este caso son los centros. La agregamos en el modelo con la opción “Añadir”. Clic en “Continuar”.



Luego seleccionamos las siguientes opciones en “Estadísticos”:



Terminamos esta primera etapa dando los “Aceptar” y presentando los resultados:

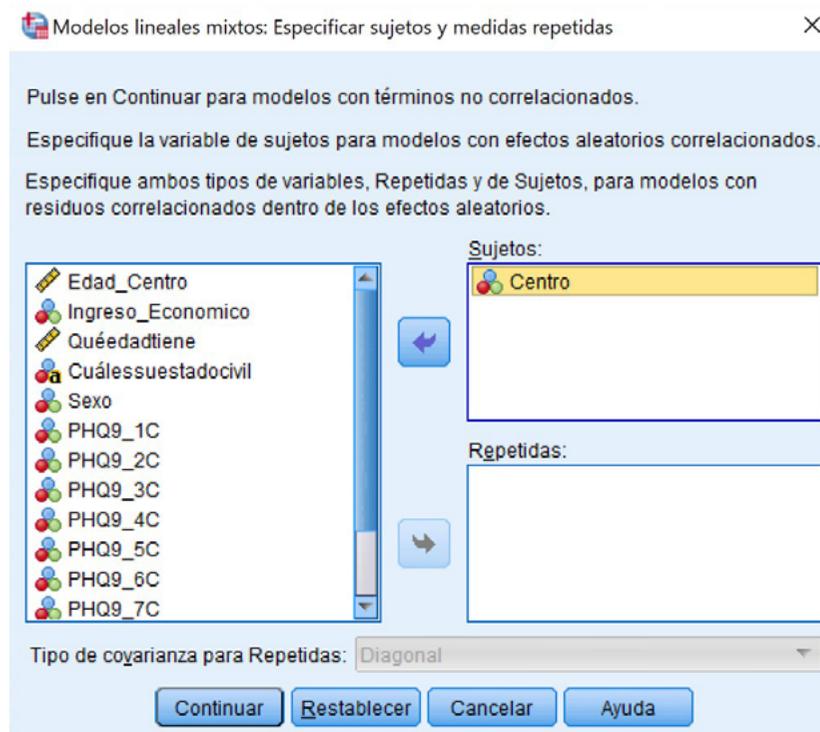
Pruebas de efectos fijos de tipo III^a

Origen	gl de numerador	gl de denominador	F	Sig.
Intersección	1	199	6786,613	,000

a. Variable dependiente: Ptj_SD.

Este modelo nulo (solo como variable independiente van los centros de salud) mostró que es un efecto fijo, ya que se rechaza la hipótesis nula (sig. = 0,000) de que el efecto del factor de tipo centro es nulo para el puntaje de síntomas depresivos.

4. Segundo paso: modelo con variables independientes individuales teniendo en cuenta la variable tipo de centro (modelo jerarquizado o multinivel de datos). Volvemos a ingresar a Analizar > Modelos lineales mixtos > Lineales. En la ventana emergente ingresamos como sujetos a la variable tipo de centros de salud. Clic en “Continuar”.

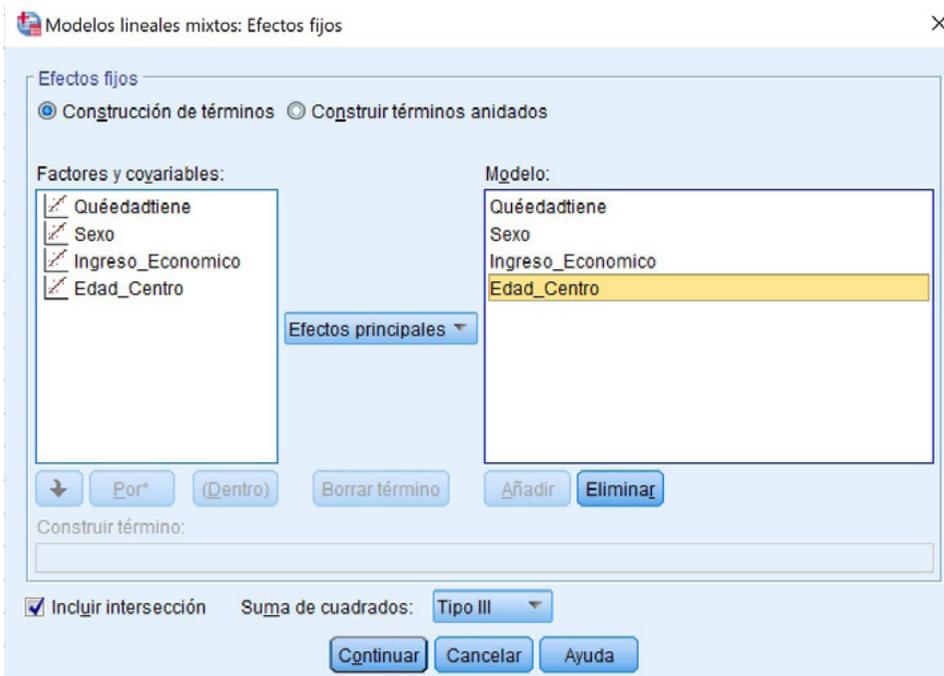


Ahora colocamos las variables independientes individuales dentro de la sección “Covariables” y en “Variable dependiente” la variable “Puntaje síntomas depresivos”.

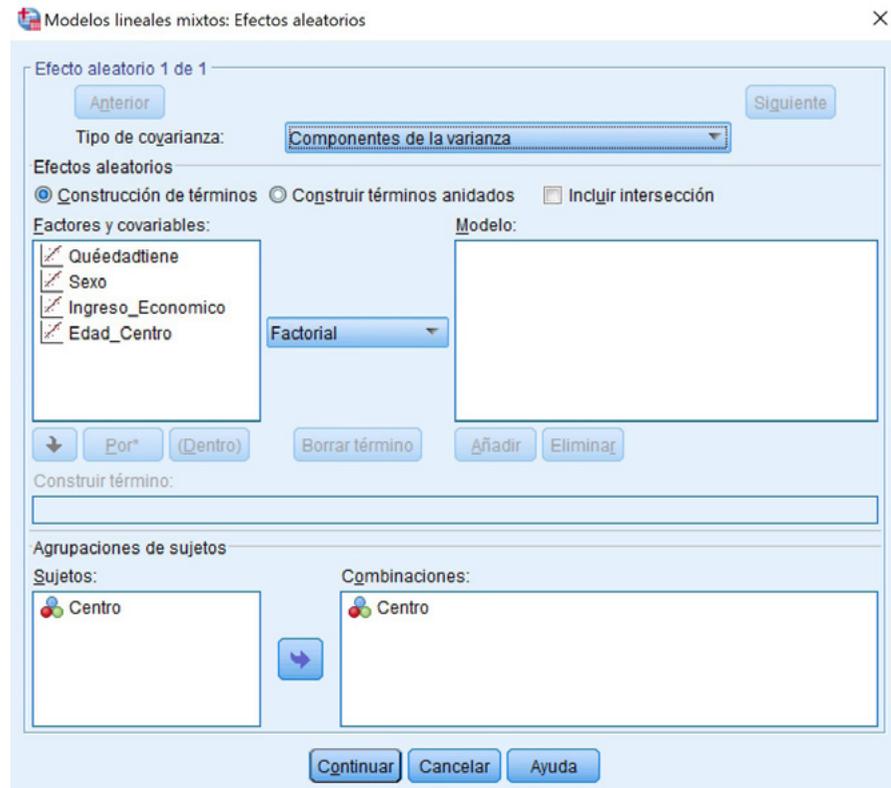


Adicionalmente:

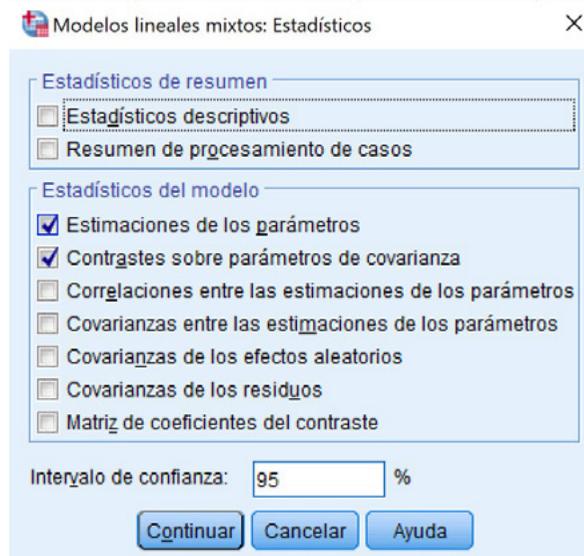
En la opción “Fija”, especificaremos cuáles son las variables individuales como factores fijos y los pasaremos al modelo como efectos principales (cambiar a esta opción antes de añadir). Terminamos dando “Continuar” para volver a la ventana principal.



Seleccionamos la opción “Aleatoria”, para identificar dentro del modelo las variables macro (variables “Intergrupos”, es decir, de los centros de salud). Nos ubicamos en la sección “Agrupación de sujetos”, seleccionamos la variable “Centros de salud” y la pasamos como “Combinaciones” al modelo. Terminamos dando clic en “Continuar” para volver a la ventana original.



En “Estadísticos” seleccionamos los siguientes parámetros y terminamos dando “Continuar”.



Una vez concluidos todos estos pasos, damos clic en “Aceptar” para mostrar los resultados obtenidos:

Pruebas de efectos fijos de tipo III^a

Origen	gl de numerador	gl de denominador	F	Sig.
Intersección	1	195	16,208	,000
Edad	1	195	6,279	,013
Sexo	1	195,000	1,047	,308
Ingreso_Economico	1	195	,717	,398
Edad_Centro	1	195	,668	,415

a. Variable dependiente: Ptj_SD.

En estos resultados la primera tabla refiere a las variables individuales y de macrovariable (“Edad promedio del paciente según centro de salud”) frente a los síntomas depresivos, y reporta que solo la edad del participante se asocia con el puntaje de síntomas depresivos (sig = 0,016), a diferencia del sexo e ingreso económico y la macrovariable “Edad promedio del paciente de cada centro de salud” (sig. > 0,05).

Estimaciones de efectos fijos^a

Parámetro	Estimación	Error estándar	gl	t	Sig.	Intervalo de confianza al 95 %	
						Límite inferior	Límite superior
Intercepción	28,877964	7,172995	195	4,026	,000	14,731353	43,024574
Edad	,077975	,031117	195	2,506	,013	,016606	,139344
Sexo	,756824	,739728	195,000	1,023	,308	-,702070	2,215718
Ingreso_Economico	,379283	,447836	195	,847	,398	-,503941	1,262507
Edad_Centro	-,192840	,235884	195	-,818	,415	-,658051	,272372

a. Variable dependiente: Ptj_SD

En esta segunda tabla también se presentan las variables independientes frente a los síntomas depresivos teniendo en cuenta la variabilidad según el tipo de centros de salud. También podemos identificar los factores. Adicionalmente, nos muestra otros indicadores como su estimador, error estándar e intervalos de confianza al 95 %.

Sección 8

Gráfico para variables cualitativas

Gráfico para variables cualitativas

Gráfico de barras y circular

¿Qué es?

Es una forma de presentación a través de una figura para las variables categóricas o cualitativas (nominal u ordinal).

¿Por qué es importante?

Esta forma de presentación de resultados de la información de variables categóricas tiene mayor efectividad de interpretación de valores porcentuales sobre una variable de estudio.

Ejemplo 19

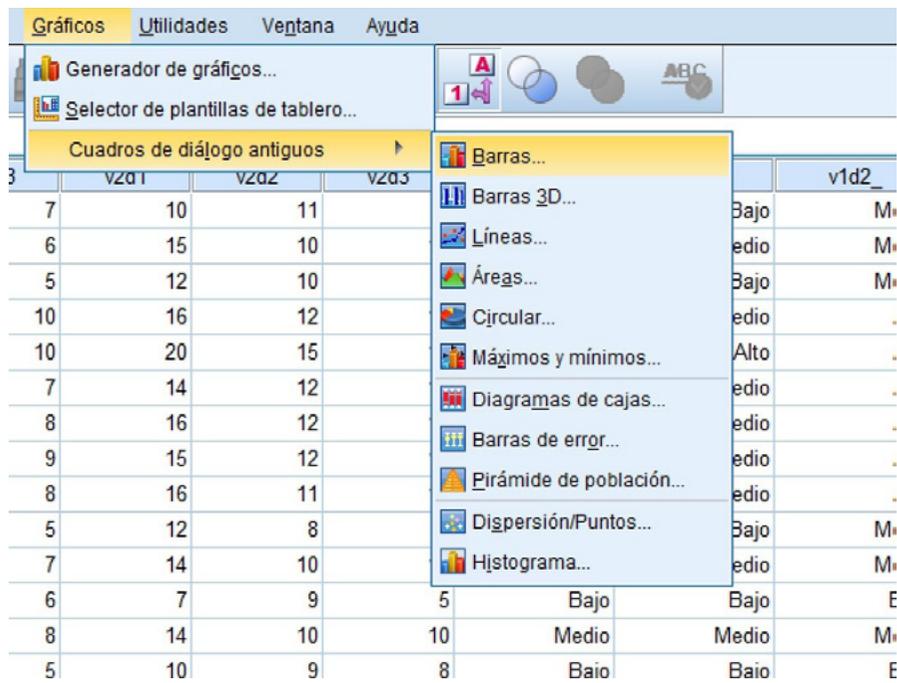
Se tiene el siguiente caso con su respectiva base de datos.

Variables: “Liderazgo pedagógico” y “Clima organizacional”.

Procedimientos

Gráfico de barras

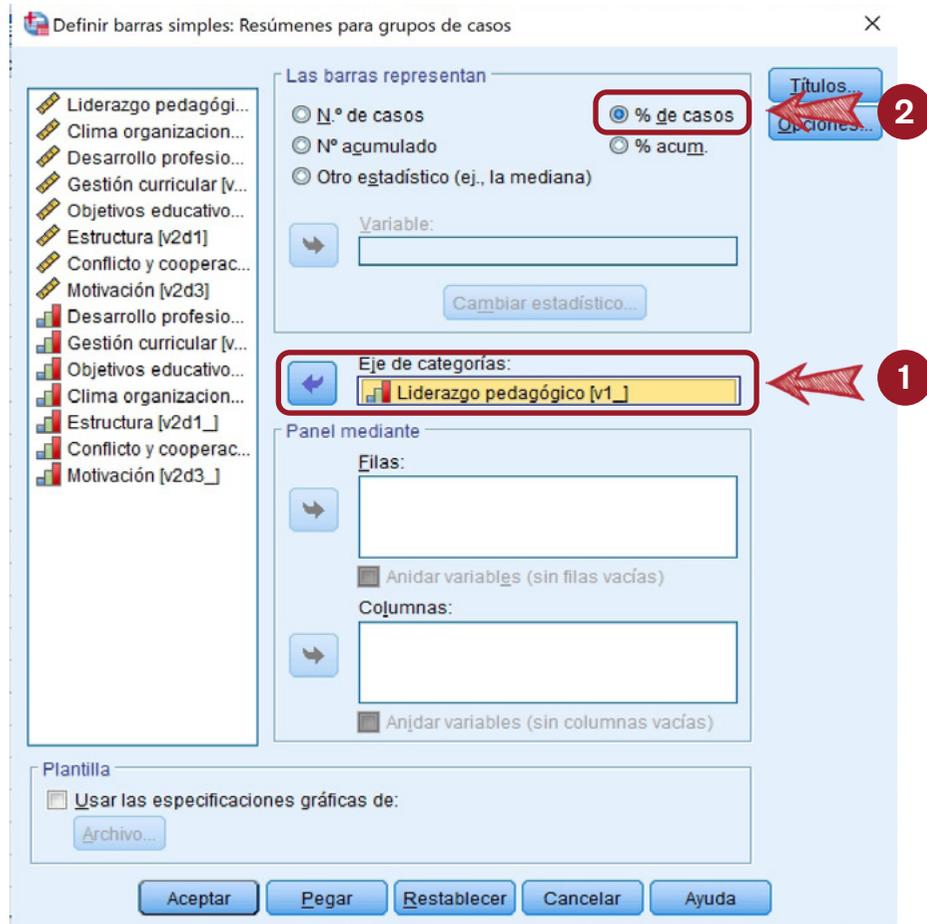
1. Utilizamos las variables categorizadas generadas en niveles. En este caso usaremos la variable satisfacción laboral (Código: v1_). Entramos en la opción “Gráficos”, hacemos clic en “Cuadros de diálogo antiguos” y seleccionamos la opción “Barras”.



2. Para una sola variable seleccionamos la opción de barras simples.



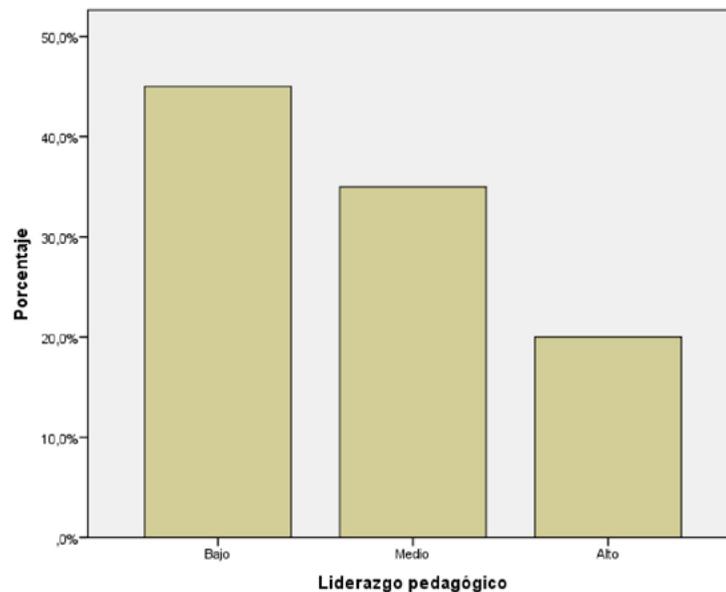
3. Primero seleccionamos la variable y la opción de cómo presentar la información (número de casos, números acumulados, porcentajes, porcentajes acumulados).



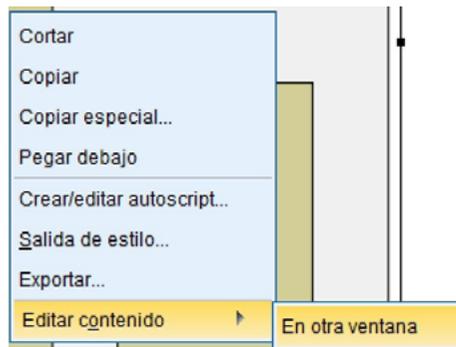
4. Automáticamente aparecerá un gráfico en la ventana de “Vista de resultados”.

Gráfico

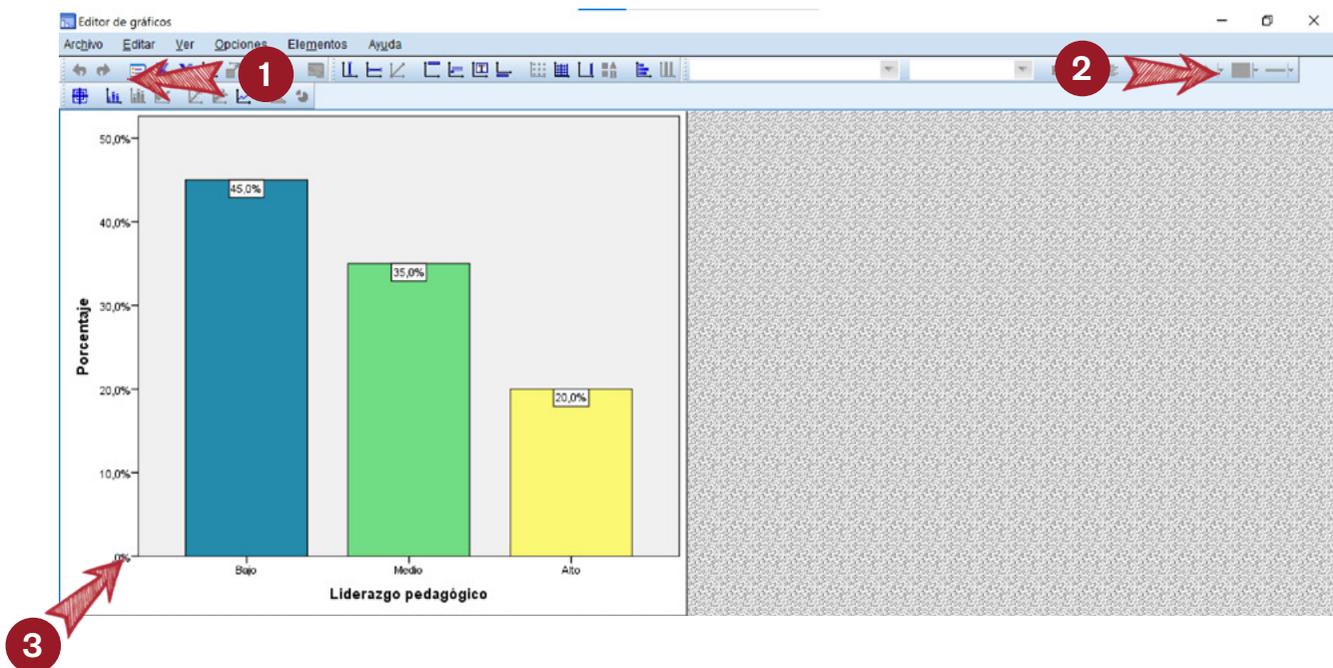
[ConjuntoDatos1] C:\Users\Lenovo\Downloads\BD 1.sav



5. Esta es la etapa final: diseñar esta gráfica. Para ello, damos clic derecho a la gráfica y seleccionamos “Editar contenido” y “En otra ventana”.



6. Aparecerá una nueva ventana donde se puede editar la gráfica: (1) seleccionamos la opción para porcentajes, (2) hacemos doble clic en cada barra y así podemos variar los colores.

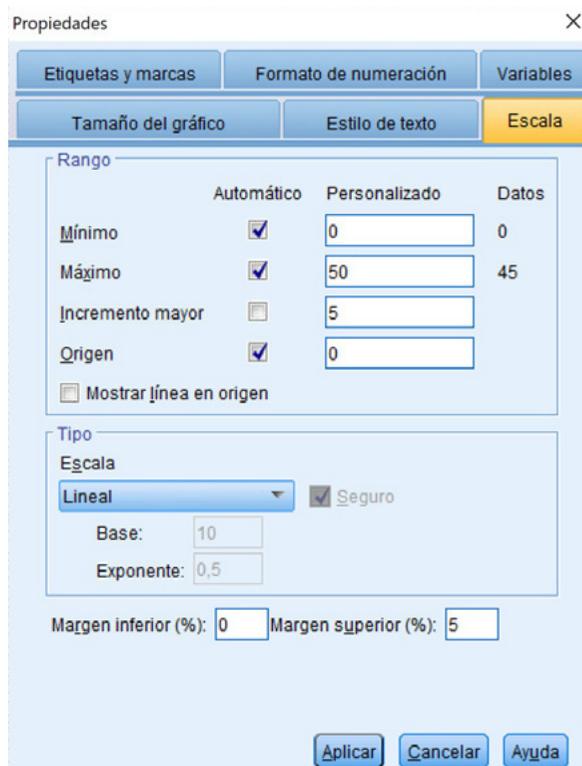


En el paso 1, damos clic en la opción “Presentación de etiquetas de barras” y llegamos a la ventana “Propiedades”. Para que las etiquetas aparezcan en la parte superior seleccionamos dentro de la posición de las etiquetas la opción “Personalizados” (escogemos la primera alternativa gráfica).



En el paso 2, como referimos, seleccionamos con doble clic cada barra. Se le puede cambiar el color en la opción mostrada.

En el paso 3, podemos dar doble clic en el eje superior para cambiar la cantidad de puntos de corte.



Finalmente, editada la gráfica, podemos copiarla y pegarla en cualquier documento (Word, Excel o PowerPoint) con solo un clic derecho al seleccionar “Copiar gráfico”.

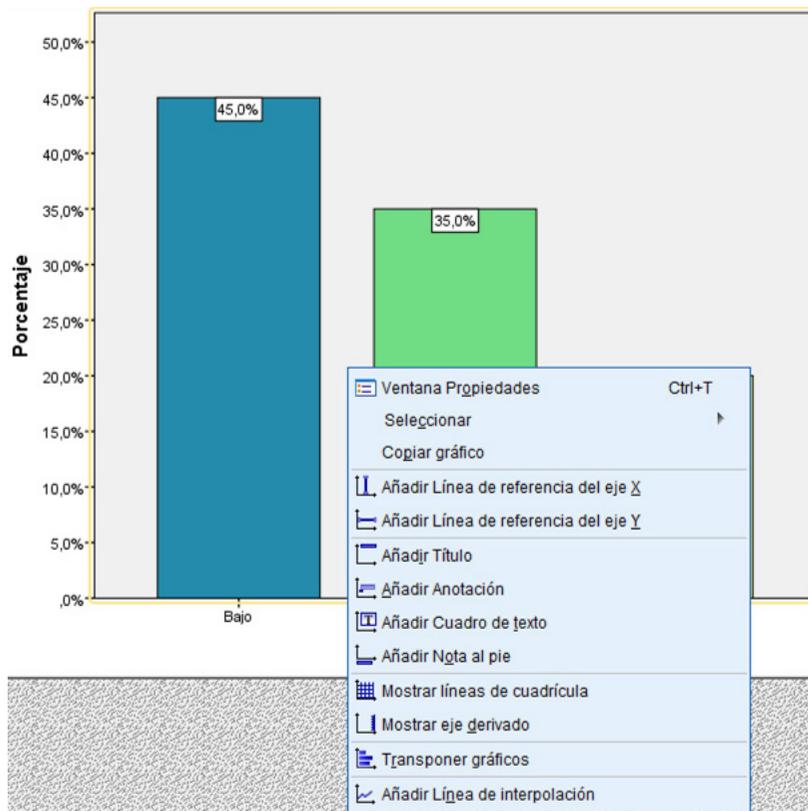
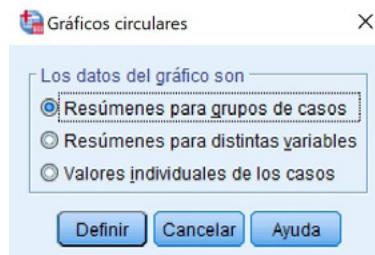


Gráfico circular

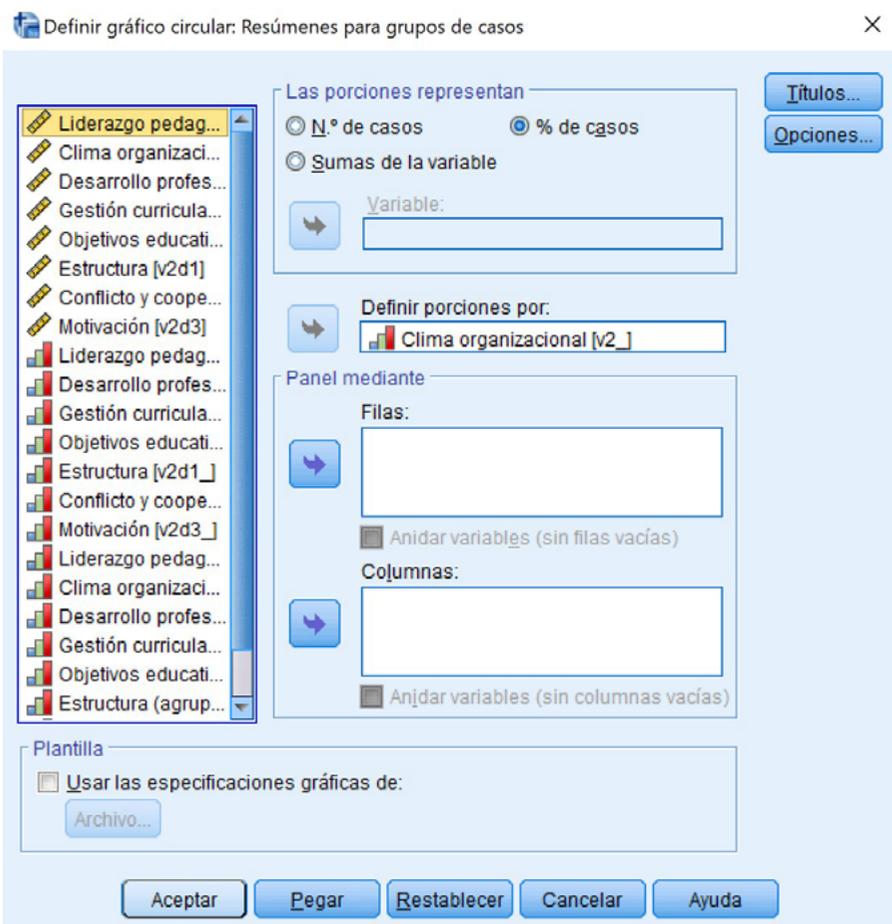
1. Utilizamos las variables categorizadas generadas en niveles. En este caso usamos la variable “Desempeño laboral” (Código: v2_). Entramos en la opción “Gráficos” y hacemos clic en “Cuadros de diálogo antiguos”. Seleccionamos la opción “Gráficos circulares”.



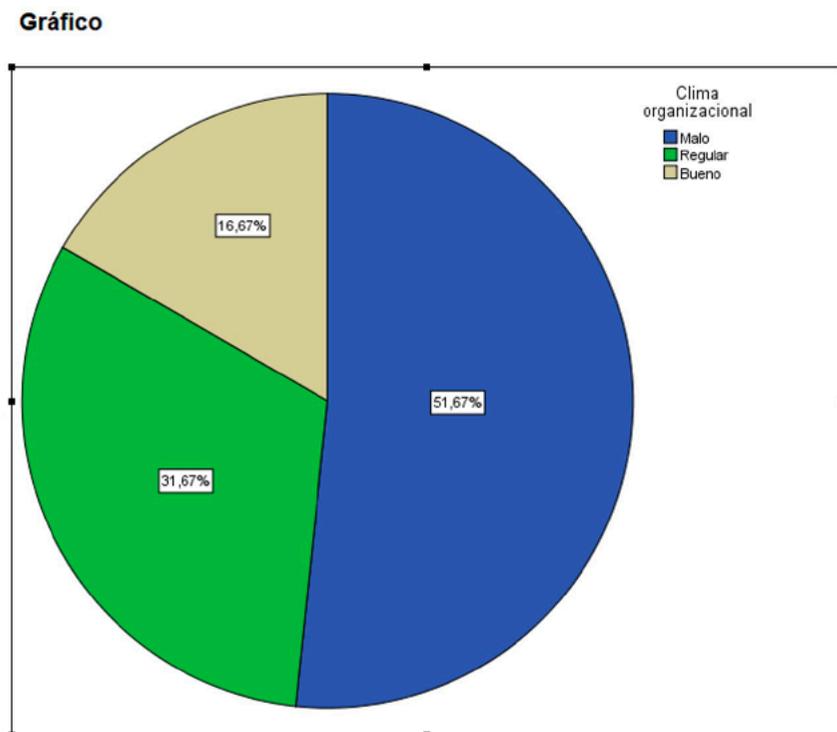
2. Seleccionamos la opción “Resúmenes para grupos de casos”. Clic en definir.



3. Seleccionamos la variable y la opción de cómo presentar la información (número de casos, números acumulados, porcentajes, porcentajes acumulados).



4. Automáticamente, así queda el gráfico:



A fin de editar colores para una mejor presentación, sigamos la ruta mencionada en el procedimiento de barras en los pasos 5 y 6.

Gráfico de variables cuantitativas

¿Qué es?

Es una forma de presentación a través de una figura para las variables numéricas o cuantitativas (continua o discreta).

¿Por qué es importante?

Esta forma de presentación de resultados de la información de variables numéricas tiene mayor efectividad de interpretación de su distribución de datos.

Ejemplo 20

Retomamos el caso utilizado en el ejemplo anterior.

Variables: “Liderazgo pedagógico” y “Clima organizacional”. Ambas variables son presentadas según sus puntajes.

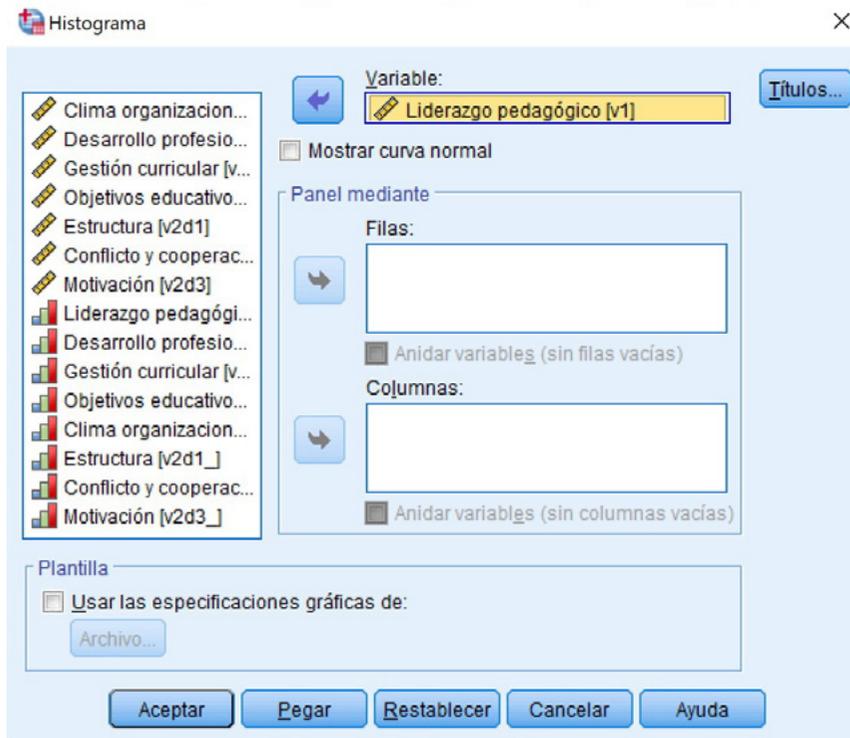
Histogramas

Procedimiento

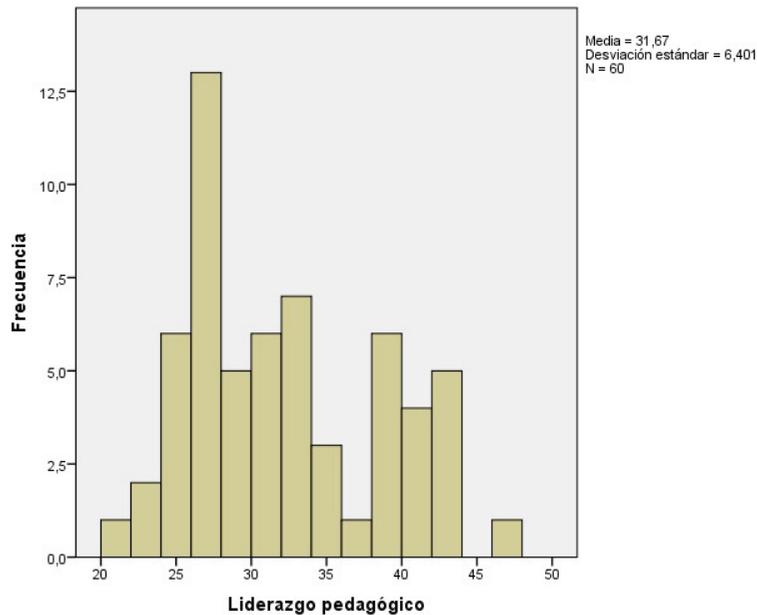
1. Utilizamos las variables categorizadas generadas en niveles. En este caso usamos la variable “Puntaje de liderazgo pedagógico” (Código: v1). Entramos en la opción “Gráficos”, hacemos clic en “Cuadros de diálogo antiguos” y seleccionamos la opción “Histogramas”.



2. Seleccionamos la variable a graficar.



3. En “Vista de resultados” aparecerá el histograma.



Adicionalmente, se presentan la media y desviación estándar de los datos. Toda esta información sirve para saber dónde se encuentra la acumulación de datos. Por ejemplo, en “Liderazgo pedagógico” se visualiza que la mayoría de puntaje obtenido está por debajo de la media.

Referencias bibliográficas

- Arispe, C., Yangali, J., Guerrero, M., Lozada, O., Acuña, L. y Arellano, C. (2020). *La investigación científica. Una aproximación para los estudios de posgrado*. Universidad Internacional del Ecuador. <https://repositorio.uide.edu.ec/bitstream/37000/4310/1/LA%20INVESTIGACION%20CIENTIFICA.pdf>
- Hernández-Sampieri, R. y Mendoza-Torres, C. (2018). *Metodología de la investigación: las rutas: cuantitativa y cualitativa y mixta*. McGraw Hill Educación. <http://repositorio.uasb.edu.bo:8080/handle/54000/1292>
- Ruiz-Bolívar, C. (2013). *Instrumentos y técnicas de investigación educativa*. DANAGA. https://www.academia.edu/37886948/Instrumentos_y_Tecnicas_de_Investigacion_Educativa_Carlos_Ruiz_Bolivar_pdf
- Sánchez, H., Reyes, C. y Mejía, K. (2018). *Manual de términos en investigación científica, tecnológica y humanística*. Universidad Ricardo Palma. <https://www.urp.edu.pe/pdf/id/13350/n/libro-manual-de-terminos-en-investigacion.pdf>

Sobre los autores

Dra. ORIANA RIVERA LOZADA DE BONILLA



Vicerrectora de Investigación de la Universidad Norbert Wiener. Investigadora en educación en salud, epidemiología y salud pública. Cuenta con la calificación de Investigadora Concytec y es miembro del Top 1000 Scientists AD Scientific Index. Reconocida por la Comisión de Salud y Población del Congreso de la República por su contribución al desarrollo de la ciencia y la difusión del conocimiento científico. Pertenece al top de inventoras peruanas de 2021 por sus patentes de invención. Entre su producción científica cuenta con libros, artículos científicos publicados en bases de datos como Scopus y Web Of Science (WoS). Certificada como evaluadora externa de programas de estudio de nivel superior universitario por el Sistema Nacional de Evaluación, Acreditación y Certificación de la Calidad Educativa (Sineace) (Resolución 000083-21-SINEACE).

Cuenta con títulos de doctorado en Salud Pública, Gestión Pública y Gobernabilidad, y Educación; maestrías en Epidemiología, Salud Pública con mención en Gestión en Servicios de Salud; y es especialista en Microbiología Clínica.

Su experiencia incluye la participación en consultorías, proyectos e investigaciones en enfermedades infecciosas, no transmisibles a nivel nacional e internacional, salud pública y epidemiología. Ha realizado consultorías a instituciones como la Organización Panamericana de la Salud (OPS), Partners In Health, PARSALUD (Ministerio de Salud), ADRA (Perú), CARE (Perú) y Tierra Viva (Colombia). Tiene 18 años de trayectoria como docente universitaria a nivel de pregrado y posgrado. En su experiencia en gestión de la investigación,

ha ocupado cargos como vicerrectora de investigación, directora de investigación y coordinadora de investigación. Es directora para el Perú de Joanna Briggs Institute (JBI) y coordinadora del Centro de Investigación South American Center for Education and Research in Public Health. Editora jefa de revistas de investigación, miembro de comités editoriales de revistas indexadas y miembro de la sociedad científica Sigma Xi.



Dra. JUDITH SOLEDAD YANGALI VICENTE

Docente principal en la Escuela de Posgrado de la Universidad Norbert Wiener e investigadora reconocida por Renacyt P0049918. Doctora en Educación y magíster en Administración de la Educación (Universidad César Vallejo), máster en Innovación Pedagógica y Gestión de Centros Educativos (EUCIM Business School, España), posdoctora en Educación (Universidad de Tijuana, México), doctoranda en Innovación, Responsabilidad Social y Sostenibilidad (Universidad Anáhuac, México), licenciada en Educación en la especialidad de Historia y Geografía (Universidad Nacional Mayor de San Marcos), segunda especialidad en Estadística e Investigación Científica (Universidad Nacional Federico Villarreal) y otra segunda especialidad en Inglés Educativo (Universidad José Carlos Mariátegui), especialista en Gestión de la Calidad, auditora de la Norma ISO 9001 y especialista en Investigación Científica (Universidad de Celaya, México).

Docente universitaria de posgrado en la Universidad Norbert Wiener y la Universidad Nacional Mayor de San Marcos, con experiencia en pregrado como docente de investigación. Desarrolla investigaciones en las líneas de investigación orientada a la calidad educativa, gestión investigativa, responsabilidad social, educación ambiental y pedagogía. Autora de artículos científicos en revistas indexadas. Ganadora de diversos proyectos de investigación docente con fondos concursables y autora de varios libros. Participa permanentemente en eventos académicos de investigación nacionales e internacionales como conferencista, ponente y par evaluador, y en proyectos de investigación internacional.

Es parte del comité científico de las revistas *INNOVA Research Journal* (UIDE-ECUADOR), *Ciencias Holguín* (Cuba), *Revista Varela* (Cuba) y la *Revista de Investigación de la Universidad Norbert Wiener* (Perú). Par evaluador en las revistas *Venezolana de Gerencia* (Universidad de Zulia, Venezuela), *Zona Próxima* (Universidad del Norte, Colombia), *Estudios en Educación* (Chile) y *Revista Científica UNTRM: Ciencias Sociales y Humanidades* (Perú). Es presidenta de la Red de Investigadores Latinoamericanos Kuélap, y miembro de la Red de Investigación

Anáhuac en Responsabilidad Social Universitaria (RIARSU), RedDolac, Emprende Sur, Rapre y Red Yammer.

Con gran experiencia profesional, actualmente es directora de Desarrollo de la Investigación en la Universidad Norbert Wiener. Se ha desempeñado como directora de Estudios Generales, coordinadora académica del Centro de Idiomas, directora de la Oficina de Tutoría Universitaria, coordinadora del Fondo Editorial, jefa de Calidad y Acreditación, coordinadora del Doctorado en Posgrado, así como docente investigadora en universidades públicas y privadas, y consultora externa de investigación y calidad.

Dr. JOSÉ LUIS RODRÍGUEZ LÓPEZ



Doctor en Educación y licenciado en Educación, con especialidad de Matemática y especialidad en Metodología y Estadística Aplicada a la Investigación. Autor y coautor de artículos científicos en revistas indexadas y libros, consultor de tesis y proyectos de investigación. Docente universitario en la Universidad del País Innova (México). Par evaluador de proyectos y artículos en revistas.

Ing. MIGUEL ANGEL IPANAQUÉ ZAPATA



Ingeniero estadístico y egresado de la maestría en dirección y gestión de proyectos.

Está registrado como investigador Renacyt por Concytec. Se desempeña como analista de datos e investigador para proyectos y artículos científicos de alcance nacional e internacional indexados, incluidos estudios psicométricos y educativos. Ha elaborado materiales educativos para facilitar la administración y el análisis de bases de datos secundarias, como en ENDES-PERÚ, y el uso de programas estadísticos. Es consultor estadístico para estudiantes y docentes universitarios en temas de investigación. Es becario del Taller Metodológico Estudio Niños del Milenio, organizado por el Grupo de Análisis para el Desarrollo (GRADE) para uso de bases de datos multicéntricas nacionales (Perú, India, Vietnam y Etiopía). Sus intereses en investigación apuntan a la salud pública, la educación, la validación de instrumentos y el *big data*.

Manual
de **procesamiento**
estadístico para la
investigación
con **SPSS**