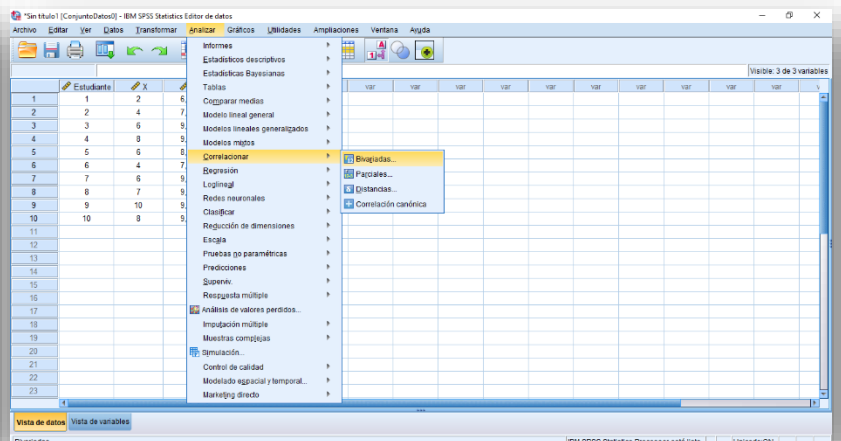
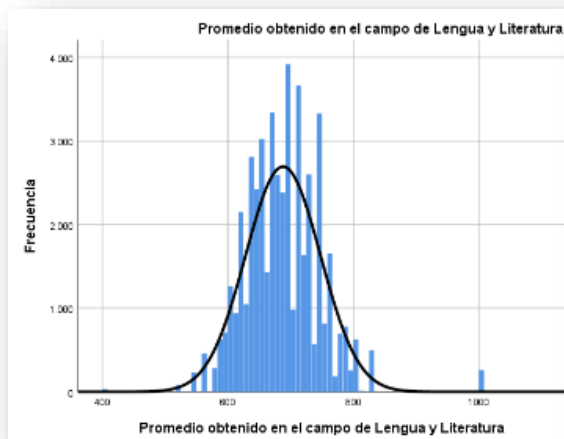


ESTADÍSTICA DESCRIPTIVA PARA TODOS: FUNDAMENTOS Y APLICACIONES - VOLUMEN III



Autor:
Mario Orlando Suárez Ibujés

IMBABURA - ECUADOR

**ESTADÍSTICA DESCRIPTIVA PARA TODOS:
FUNDAMENTOS Y APLICACIONES - VOLUMEN III**

IMBABURA – ECUADOR

2025

AUTOR

Mgs. Mario Orlando Suárez Ibujés

Docente ocasional en la Facultad de Posgrado de la UTN

Correos electrónicos: mosuarez@utn.edu.ec; mariosuarezibujes@gmail.com

ORCID: <https://orcid.org/0000-0002-3962-5433>

PARES REVISORES

Mgs. Liliana Chamorro Hernández

Docente en la Universidad Politécnica Estatal del Carchi

Correo electrónico: lilianam.chomorro@upec.edu.ec

Mgs. Víctor Mario García Mora

Docente en la Escuela Superior Politécnica de Chimborazo

Correo electrónico: victor.garcia@espoch.edu.ec

Mgs. Gandy Patricio Rivadeneira Martínez

Asesor Educativo en la Coordinación Zonal 1- Educación

Correo electrónico: gandy.rivadeneira@educacion.gob.ec

PhD. Cristian Eduardo Terán Montalvo

Funcionario de la Coordinación Zonal 1- Educación

Correo electrónico: cristiane.teran@educacion.gob.ec

ISBN: Primera Edición 2025

Con el Auspicio de

ISBN: 978-9942-51-798-2



Esta obra no puede ser reproducida total ni parcialmente por ningún medio sin expreso consentimiento previo y por escrito del autor.

DEDICATORIA

Con amor infinito en expansión,
para mi esposa, Dyana Rivera, el amor de mi vida y de todas mis vidas;
para mis hijos, Emily Monserrath y Mathías Josué, la continuación de mi ser,
mi inspiración más profunda y mi más anhelado sueño hecho realidad.

A mis padres, Bertha Ibujés y Segundo Suárez,
por su ejemplo incansable de perseverancia y amor.

AGRADECIMIENTO

Agradezco profundamente a los docentes pares revisores por sus valiosas aportaciones y sugerencias, indispensables para el perfeccionamiento de esta obra. Cada observación fue recibida con gratitud y consideración, demostrando que el conocimiento se construye en comunidad.

De igual manera, expreso mi profundo agradecimiento a las instituciones que respaldaron esta obra y permitieron incluir su logo institucional como muestra de su apoyo: Equipo de Asesores Educativos de la Coordinación Zonal 1-Educación, Colegio Profesional de Asesores Educativos del Ecuador, Centro Cultural Antonio Ante, y Asociación de Maestros de Excelencia Educativa, instituciones de las cuales tengo el honor de ser integrante.

CONTENIDOS

CONTRAPORTADA	1
DEDICATORIA	3
AGRADECIMIENTO	4
CONTENIDOS	5
PRESENTACIÓN	7
EVALUACIÓN DIAGNÓSTICA	8
 CAPÍTULO I	
ASIMETRÍA	10
1.1 Definición	
1.2 Propiedades	
1.3 Coeficiente de asimetría basado en momentos	11
A Fórmula	
B Tabla de interpretación del coeficiente de asimetría (g_1)	
1.4 Error estándar de asimetría	12
A Fórmula	
B Regla práctica	
1.5 Ejemplos de aplicación	13
1.6 Actividades de aplicación	33
 CAPÍTULO II	
CURTOSIS	36
2.1 Definición	
2.2 Propiedades	
2.3 Coeficiente de curtosis	37
A Fórmula	
B Tabla de interpretación del coeficiente de curtosis (g_2)	
2.4 Error estándar de curtosis	38
A Fórmula	
B Regla práctica	39
2.5 Ejemplos de aplicación	
2.6 Actividades de aplicación	50
 CAPÍTULO III	
CORRELACIÓN	53
3.1 Definición	
3.2 Clasificación	54
3.3 Coeficiente de correlación de Pearson	56
A Definición	
B Fórmula	
C Supuestos para aplicar	
D Interpretación del p – valor	57
E Ejemplos de aplicación	58
3.4) Coeficiente de correlación de Spearman	83
A Definición	
B Características	

C Diferencia con la correlación lineal (Pearson)	84
D Fórmula	
E Ejemplos de aplicación	85
3.5 Coeficiente de correlación de Tau-b de Kendall	93
A Definición	
B Características	
C Diferencias con Spearman y Pearson	94
D Fórmula	
E Ejemplos de aplicación	
3.6 Actividades de aplicación	102
FORMULARIO	111
SOLUCIONARIO DE LA EVALUACIÓN DIAGNÓSTICA	125
BIBLIOGRAFÍA	132
DATOS BIOGRÁFICOS DEL AUTOR	134

PRESENTACIÓN

Este libro, *Estadística Descriptiva para Todos: Fundamentos y Aplicaciones-Volumen III* está dedicado a profundizar en conceptos avanzados de estadística descriptiva, con un enfoque práctico y accesible para estudiantes, profesionales y cualquier persona interesada en el análisis de datos. El libro combina teoría fundamentada con ejemplos aplicados, utilizando herramientas como Excel y SPSS, lo que facilita la comprensión y la implementación de los conceptos en contextos reales.

El contenido se organiza en tres capítulos fundamentales: Capítulo I: Asimetría, donde se profundiza en la medición e interpretación de la asimetría en distribuciones de datos, integrando fórmulas, análisis prácticos y ejemplos aplicados; Capítulo II: Curtosis, dedicado al estudio del "pico" y las "colas" de las distribuciones, su comparación con la curva normal y ejercicios para afianzar los conceptos; y Capítulo III: Correlación, que presenta métodos clave para analizar relaciones entre variables, como los coeficientes de Pearson, Spearman y Tau-b de Kendall, con énfasis en su aplicación e interpretación en contextos reales. Cada sección combina teoría sólida con recursos didácticos que facilitan el aprendizaje progresivo.

Lo que hace especial a esta obra es su equilibrio único entre rigor académico y accesibilidad. Combina un enfoque didáctico con explicaciones claras, tablas resumen y gráficos ilustrativos que facilitan la comprensión de conceptos estadísticos complejos. Además, incluye herramientas prácticas con instrucciones paso a paso para aplicar técnicas en Excel y SPSS, lo que permite al lector llevar la teoría a la práctica de inmediato. El libro refuerza el aprendizaje con ejercicios y soluciones detalladas, diseñados para consolidar los conocimientos adquiridos. Finalmente, su enfoque en aplicaciones reales, como el análisis de datos de evaluaciones educativas como la prueba Ser Estudiante y otros contextos, asegura que los contenidos sean relevantes y útiles en escenarios profesionales y académicos.

Este libro está diseñado para estudiantes, profesionales y entusiastas que busquen desarrollar habilidades fundamentales en estadística descriptiva. Ya sea que requieras analizar datos en tu ámbito laboral, interpretar investigaciones académicas o tomar decisiones basadas en evidencia, esta obra les proporcionará las herramientas necesarias para lograrlo con claridad y confianza.

Debemos creer en nosotros mismos sin importar nuestro origen, dinero y tiempo que tengamos, recordando que los sueños sin acciones son simples sueños, las acciones sin sueños carecen de sentido, pero un sueño puesto en acción puede cambiar el mundo.

EVALUACIÓN DIAGNÓSTICA

OBJETIVO: Verificar los resultados de aprendizaje previos adquiridos por los lectores a través del presente cuestionario para emitir juicios de valor y tomar decisiones.

INSTRUCCIONES:

Estimado lector:

- ✓ La evaluación tiene una duración de 2 horas.
- ✓ Cada pregunta tiene una valoración de dos puntos.
- ✓ No se otorgará valoración a una respuesta correcta que no esté acompañada de un proceso de solución escrito.
- ✓ Emplee hojas adicionales y un esferográfico para resolver el presente cuestionario.
- ✓ Lea cuidadosamente el cuestionario y conteste empleando sus conocimientos previos.

¡Éxito!

CUESTIONARIO

1) En una institución educativa se va a realizar una encuesta a estudiantes de 4to a 10mo sobre temas para construir el Modelo Institucional de Evaluación Educativa. Los datos se muestran en la siguiente tabla:

Curso	Nº de estudiantes
4to	35
5to	34
6to	38
7mo	40
8vo	36
9no	39
10mo	40

Realice los cálculos empleando Excel.

a) ¿Cuántas encuestas se debe aplicar si se emplea el cálculo del tamaño de la muestra al 95% de confianza con un error de muestreo del 5%?

- b) ¿Cuántas encuestas se deben aplicar en cada curso?
- c) Realice la interpretación de los resultados.

2) Con los siguientes datos sobre las calificaciones trimestrales en la Asignatura de Matemática evaluadas sobre 10 obtenidas de 40 estudiantes, empleando Excel calcular la media, error típico, mediana, moda, desviación estándar, varianza de la muestra, curtosis, coeficiente de asimetría, rango, mínimo, máximo, suma y cuenta. Interpretar los resultados.

8,83	8,85	8,74	7,70
4,48	8,40	8,70	8,99
6,99	8,79	7,90	7,69
7,00	8,70	8,93	9,11
8,84	9,05	8,85	9,83
9,83	8,33	8,88	8,90
8,78	9,01	8,96	8,38
8,71	8,85	8,55	8,11
8,84	6,24	8,99	4,00
7,33	8,26	9,23	9,31

- 3) Elabore un diagrama de barras, un histograma y un diagrama de sectores considerando los intervalos empleados en el sistema educativo del Ministerio de Educación del Ecuador. Realice la interpretación.
- 4) Calcule los cuartiles empleando Excel. Realice la interpretación.
- 5) Elabore un diagrama de caja y bigotes. Realice la interpretación.

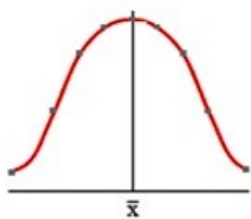
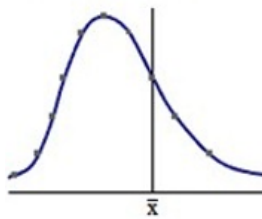
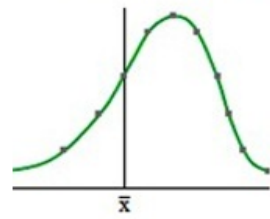
CAPÍTULO I ASIMETRÍA

1.1) DEFINICIÓN

La asimetría (skewness = oblicuidad) es una medida estadística que describe el grado de desviación de la simetría en una distribución de datos. Indica si los datos están sesgados hacia la izquierda, hacia la derecha o si son simétricos. La asimetría es clave para entender la forma de la distribución de datos.

1.2) PROPIEDADES

Tabla: Propiedades de la Asimetría

Característica	Distribución Simétrica	Asimetría Positiva (Derecha)	Asimetría Negativa (Izquierda)
Forma de la distribución	Simétrica (espejo en la media)	Cola larga hacia la derecha	Cola larga hacia la izquierda
Relación entre medidas	Media = Mediana = Moda	Media > Mediana > Moda	Media < Mediana < Moda
Coefficiente de asimetría	= 0	> 0	< 0
Ejemplo práctico	Altura de adultos en una población sana.	Ingresos económicos (muchos bajos, pocos altos).	Edad de jubilación (la mayoría se jubila tarde, pocos temprano).
Gráfico típico	Campana perfecta (Normal).  $\bar{x} = Me = Mo$	Pico a la izquierda, cola derecha alargada.  $\bar{x} > Me > Mo$	Pico a la derecha, cola izquierda alargada.  $\bar{x} < Me < Mo$
Efecto en análisis	La media es representativa.	La media es arrastrada por valores altos (usar mediana).	La media es arrastrada por valores bajos (usar mediana).

1.3) COEFICIENTE DE ASIMETRÍA BASADO EN MOMENTOS

A) FÓRMULA

La fórmula del coeficiente de asimetría basado en momentos, **Fisher-Pearson**, (g_1) es una medida de asimetría que se calcula utilizando los momentos estadísticos de la distribución, específicamente el tercer momento estandarizado.

Los momentos estadísticos son medidas que describen la forma de una distribución.

El primer momento (media, $\bar{x}$) indica la tendencia central.

El segundo momento (varianza, s^2) mide la dispersión.

El tercer momento estandarizado (m_3), basado en la media y la desviación estándar, capta la asimetría.

La fórmula del coeficiente de asimetría basado en momentos (Fisher-Pearson) con ajuste para una muestra es:

$$g_1 = \frac{n}{(n-1)(n-2)} \cdot \frac{m_3}{s^3} = \frac{n}{(n-1)(n-2)} \cdot \frac{\sum (x_i - \bar{x})^3}{s^3} = \frac{n}{(n-1)(n-2)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^3$$

Donde:

x_i = cada uno de los valores

n = tamaño de la muestra

$\bar{x}$ = media aritmética

s = desviación estándar de la muestra

B) TABLA DE INTERPRETACIÓN DEL COEFICIENTE DE ASIMETRÍA (g_1)

Valor de g_1	Tipo de Asimetría	Forma de la Distribución	Interpretación
$g_1 = 0$	Simétrica	Colas equilibradas (distribución normal).	$\bar{x} = Me = Mo$
$-0,5 < g_1 < 0$	Leve asimetría negativa	Cola izquierda ligeramente más larga.	Presencia de valores atípicos bajos. $\bar{x} < Me$
$g_1 \leq -0,5$	Moderada/Fuerte negativa	Cola izquierda pronunciada.	Valores bajos distorsionan la distribución. $\bar{x} \ll Me$
$0 < g_1 < 0,5$	Leve asimetría positiva	Cola derecha ligeramente más larga.	Presencia de valores atípicos altos. $\bar{x} > Me$.
$g_1 \geq 0,5$	Moderada/Fuerte positiva	Cola derecha pronunciada.	Valores altos distorsionan la distribución. $\bar{x} \gg Me$

1.4) ERROR ESTÁNDAR DE ASIMETRÍA

A) FÓRMULA

El error estándar de asimetría (o *standard error of skewness*, $SE_{asimetría}$) es una medida estadística que cuantifica la variabilidad esperada del coeficiente de asimetría si repitiéramos el muestreo múltiples veces bajo las mismas condiciones. En otras palabras, estima cuánto podría fluctuar el valor de asimetría debido al azar en una muestra dada.

La fórmula para calcular el error estándar de asimetría es:

$$SE_{asimetría} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

Donde:

n = Número total de observaciones válidas de la muestra

Esta fórmula deriva de la distribución asintótica del coeficiente de asimetría bajo la hipótesis nula de normalidad (es decir, cuánto variaría el valor de asimetría si los datos fueran perfectamente normales).

Para muestras grandes ($n > 150$), se aproxima a:

$$SE_{asimetría} = \sqrt{\frac{6}{n}}$$

B) REGLA PRÁCTICA:

La regla práctica para determinar si la asimetría de una distribución es estadísticamente significativa (es decir, no aleatoria) se expresa matemáticamente como:

$$|g_1| > 2 \cdot SE_{asimetría}$$

Donde:

g_1 = coeficiente de asimetría

$SE_{asimetría}$ = error estándar de asimetría

Por lo tanto, si el valor absoluto del coeficiente de asimetría es mayor que 2 veces su error estándar, se considera que la distribución es significativamente asimétrica (no es simétrica y tiene una cola derecha más larga con un 95% de confianza aproximado).

Nota:

El valor **2** surge de la aproximación a una distribución normal estándar (nivel de confianza del 95% para muestras grandes).

1.5) EJEMPLOS DE APLICACIÓN

1) Calcular la asimetría de las siguientes calificaciones 6, 7, 7, 7, 8, 9, 9, 10, 10, 10

Solución:

Se calcula la media y la desviación estándar

$$\bar{x} = \frac{\sum x_i}{n} = \frac{6 + 7 + 7 + 7 + 8 + 9 + 9 + 10 + 10 + 10}{10} = 8,3$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{(6 - 8,3)^2 + (7 - 8,3)^2 + \dots + (10 - 8,3)^2}{10 - 1}} = 1,494$$

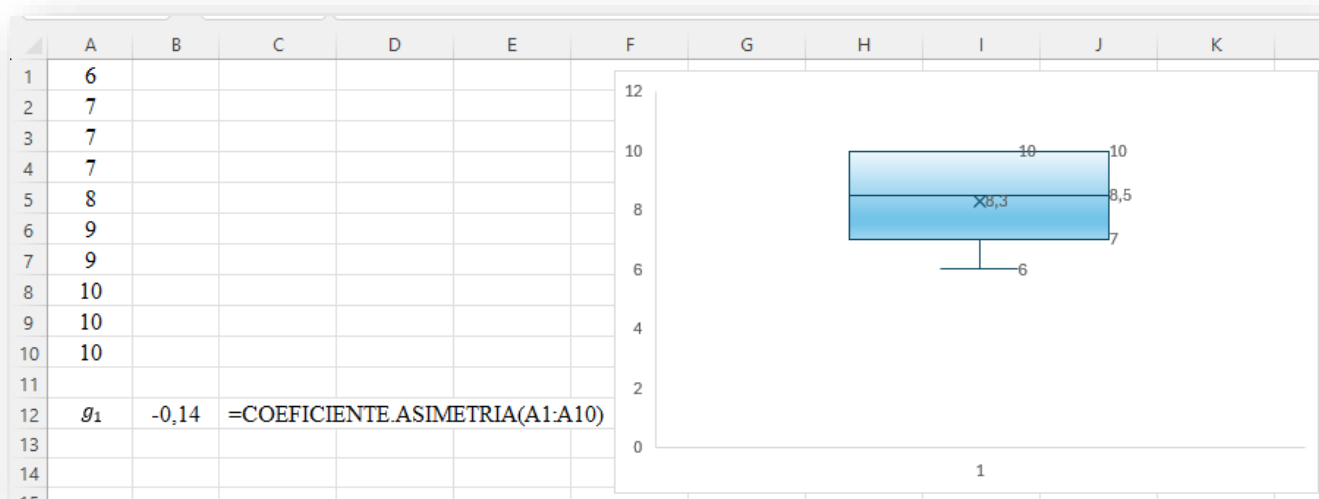
Se llena la siguiente tabla:

x_i	$\left(\frac{x_i - \bar{x}}{s}\right)^3$
6	-3,645
7	-0,658
7	-0,658
7	-0,658
8	-0,008
9	0,103
9	0,103
10	1,472
10	1,472
10	1,472
Total	-1,007

Se calcula el coeficiente de asimetría

$$g_1 = \frac{n}{(n-1)(n-2)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^3 = \frac{10}{(10-1)(10-2)} \cdot (-1,007) = -0,14$$

Empleando Excel



Interpretación:

La media (8,3) es menor que la mediana (8,5) y menor que una de las modas (10), lo que confirma el sesgo negativo ($\bar{x} < Me < Mo$)

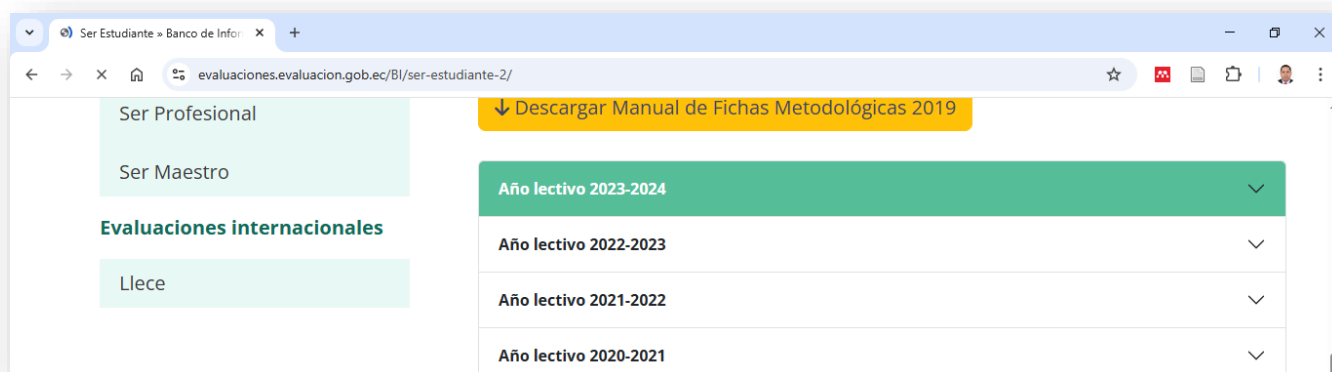
Existe una leve asimetría negativa ($-0,5 < g_1 < 0$).

La media es arrastrada por el valor bajo (6), por lo que se podría decir que mediana es más representativa que la media.

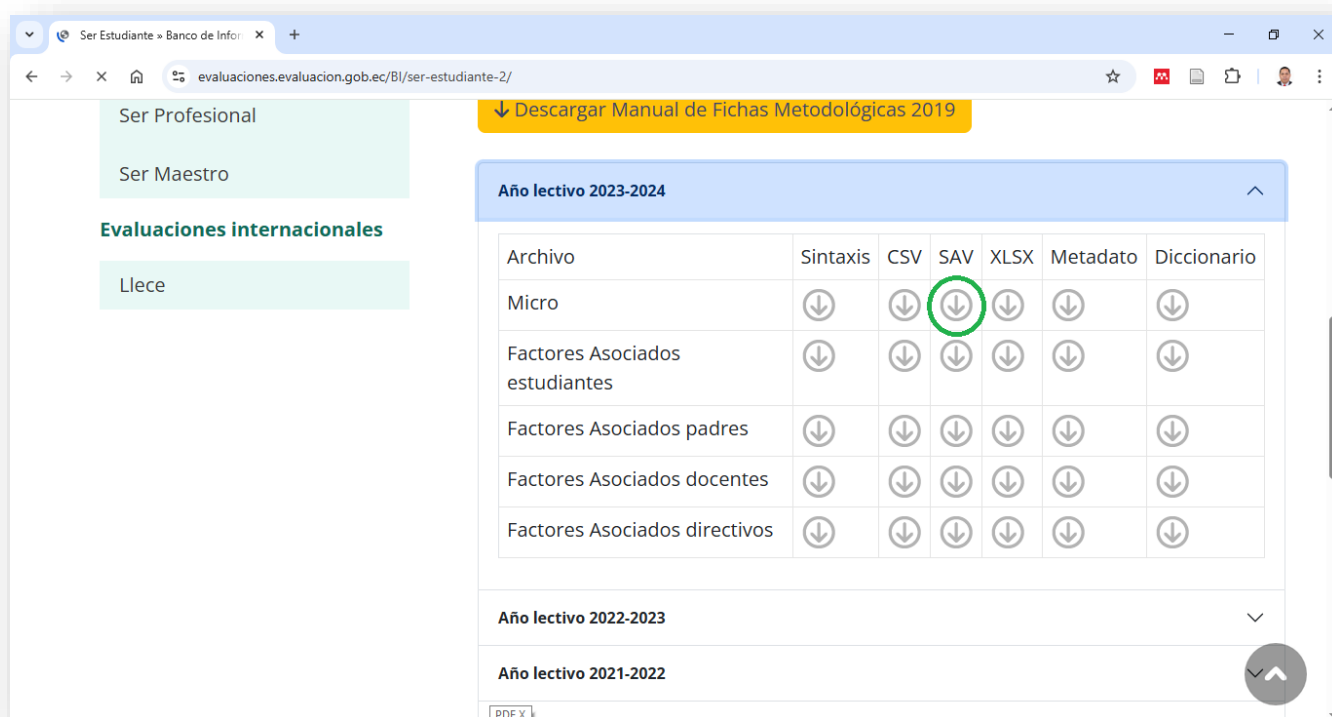
2) Calcular la asimetría del Promedio global (Nota global obtenida por el sustentante) de la prueba Ser Estudiante 2023-2024 empleando SPSS.

Solución:

Se descarga la base de datos de la Prueba Ser Estudiante 2023-2024 (Micro SAV) disponible en el portal oficial: <https://evaluaciones.evaluacion.gob.ec/BI/ser-estudiante-2/>



Descargue Micro SAV del Año lectivo 2023-2024



Empleando la base de datos de la Prueba Ser Estudiante 2023-2024 se procede de la siguiente manera:

Visible: 78 de 78 variables

	ciclo	grado	estado_eval	codigo	a
1	2023-2024	Subnivel Básica Superior	Evaluado	0BY3566978	UNIDAD EDUCATIVA A
2	2023-2024	Subnivel Básica Superior	Evaluado	0EYB576428	UNIDAD EDUCATIVA A
3	2023-2024	Subnivel Básica Superior	Evaluado	0LXC578175	UNIDAD EDUCATIVA A
4	2023-2024	Subnivel Básica Superior	Evaluado	1EEF574289	UNIDAD EDUCATIVA A
5	2023-2024	Subnivel Básica Superior	Evaluado	223Y549256	UNIDAD EDUCATIVA A
6	2023-2024	Subnivel Básica Superior	Evaluado	2H6D566318	UNIDAD EDUCATIVA A
7	2023-2024	Subnivel Básica Superior	Evaluado	2OA0534114	UNIDAD EDUCATIVA A
8	2023-2024	Subnivel Básica Superior	Evaluado	2UN7556301	UNIDAD EDUCATIVA A
9	2023-2024	Subnivel Básica Superior	Evaluado	40IF541082	UNIDAD EDUCATIVA A
10	2023-2024	Subnivel Básica Superior	Evaluado	4DT7565045	UNIDAD EDUCATIVA A
11	2023-2024	Subnivel Básica Superior	Evaluado	57TA553757	UNIDAD EDUCATIVA A
12	2023-2024	Subnivel Básica Superior	Evaluado	64A5539493	UNIDAD EDUCATIVA A
13	2023-2024	Subnivel Básica Superior	Evaluado	7DDD543579	UNIDAD EDUCATIVA A
14	2023-2024	Subnivel Básica Superior	Evaluado	7VMK538939	UNIDAD EDUCATIVA A
15	2023-2024	Subnivel Básica Superior	Evaluado	8RGA544078	UNIDAD EDUCATIVA A
16	2023-2024	Subnivel Básica Superior	Evaluado	9ERQ535115	UNIDAD EDUCATIVA A
17	2023-2024	Subnivel Básica Superior	Evaluado	9MC1548020	UNIDAD EDUCATIVA A
18	2023-2024	Subnivel Básica Superior	Evaluado	9MV7559121	UNIDAD EDUCATIVA A
19	2023-2024	Subnivel Básica Superior	Evaluado	9OZY570785	UNIDAD EDUCATIVA A
20	2023-2024	Subnivel Básica Superior	Evaluado	9WK5568297	UNIDAD EDUCATIVA A
21	2023-2024	Subnivel Básica Superior	Evaluado	C677561799	UNIDAD EDUCATIVA A
22	2023-2024	Subnivel Básica Superior	Evaluado	CFHB562186	UNIDAD EDUCATIVA A
23	2023-2024	Subnivel Básica Superior	Evaluado	D7T6578751	UNIDAD EDUCATIVA A

Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo Unicode:ON

Clic en Analizar, Estadísticos descriptivos, Frecuencias

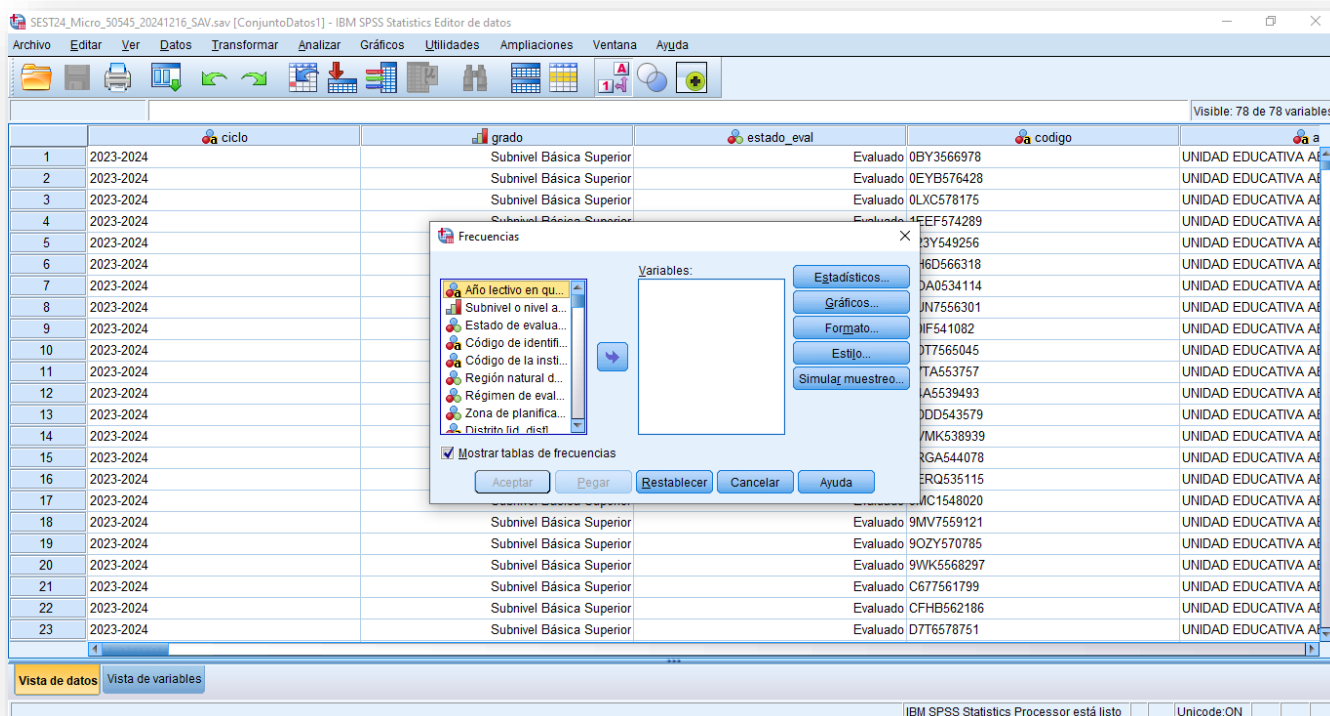
Visible: 78 de 78 variables

	ciclo	grado	estado_eval	codigo	a
1	2023-2024	Subnivel Básica Superior	Evaluado	0BY3566978	UNIDAD EDUCATIVA A
2	2023-2024	Subnivel Básica Superior	Evaluado	0EYB576428	UNIDAD EDUCATIVA A
3	2023-2024	Subnivel Básica Superior	Evaluado	0LXC578175	UNIDAD EDUCATIVA A
4	2023-2024	Subnivel Básica Superior	Evaluado	1EEF574289	UNIDAD EDUCATIVA A
5	2023-2024	Subnivel Básica Superior	Evaluado	223Y549256	UNIDAD EDUCATIVA A
6	2023-2024	Subnivel Básica Superior	Evaluado	2H6D566318	UNIDAD EDUCATIVA A
7	2023-2024	Subnivel Básica Superior	Evaluado	2OA0534114	UNIDAD EDUCATIVA A
8	2023-2024	Subnivel Básica Superior	Evaluado	2UN7556301	UNIDAD EDUCATIVA A
9	2023-2024	Subnivel Básica Superior	Evaluado	40IF541082	UNIDAD EDUCATIVA A
10	2023-2024	Subnivel Básica Superior	Evaluado	4DT7565045	UNIDAD EDUCATIVA A
11	2023-2024	Subnivel Básica Superior	Evaluado	57TA553757	UNIDAD EDUCATIVA A
12	2023-2024	Subnivel Básica Superior	Evaluado	64A5539493	UNIDAD EDUCATIVA A
13	2023-2024	Subnivel Básica Superior	Evaluado	7DDD543579	UNIDAD EDUCATIVA A
14	2023-2024	Subnivel Básica Superior	Evaluado	7VMK538939	UNIDAD EDUCATIVA A
15	2023-2024	Subnivel Básica Superior	Evaluado	8RGA544078	UNIDAD EDUCATIVA A
16	2023-2024	Subnivel Básica Superior	Evaluado	9ERQ535115	UNIDAD EDUCATIVA A
17	2023-2024	Subnivel Básica Superior	Evaluado	9MC1548020	UNIDAD EDUCATIVA A
18	2023-2024	Subnivel Básica Superior	Evaluado	9MV7559121	UNIDAD EDUCATIVA A
19	2023-2024	Subnivel Básica Superior	Evaluado	9OZY570785	UNIDAD EDUCATIVA A
20	2023-2024	Subnivel Básica Superior	Evaluado	9WK5568297	UNIDAD EDUCATIVA A
21	2023-2024	Subnivel Básica Superior	Evaluado	C677561799	UNIDAD EDUCATIVA A
22	2023-2024	Subnivel Básica Superior	Evaluado	CFHB562186	UNIDAD EDUCATIVA A
23	2023-2024	Subnivel Básica Superior	Evaluado	D7T6578751	UNIDAD EDUCATIVA A

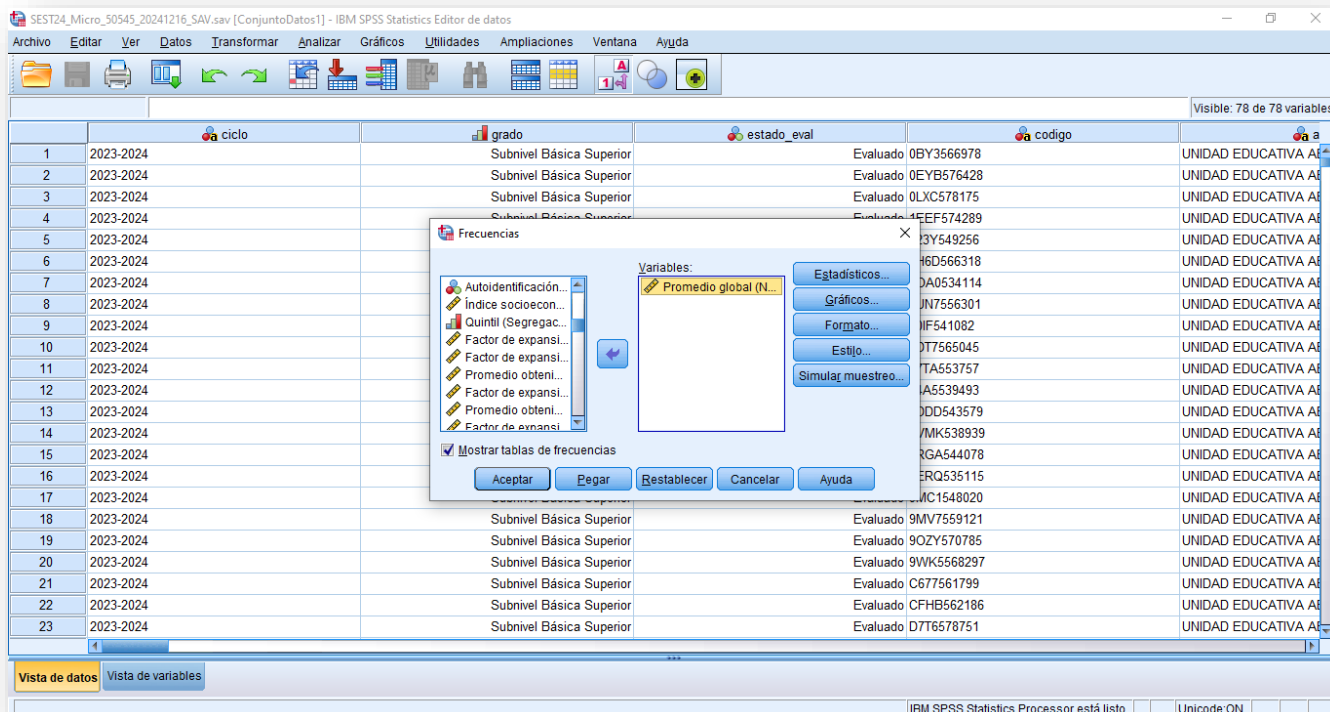
Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo Unicode:ON

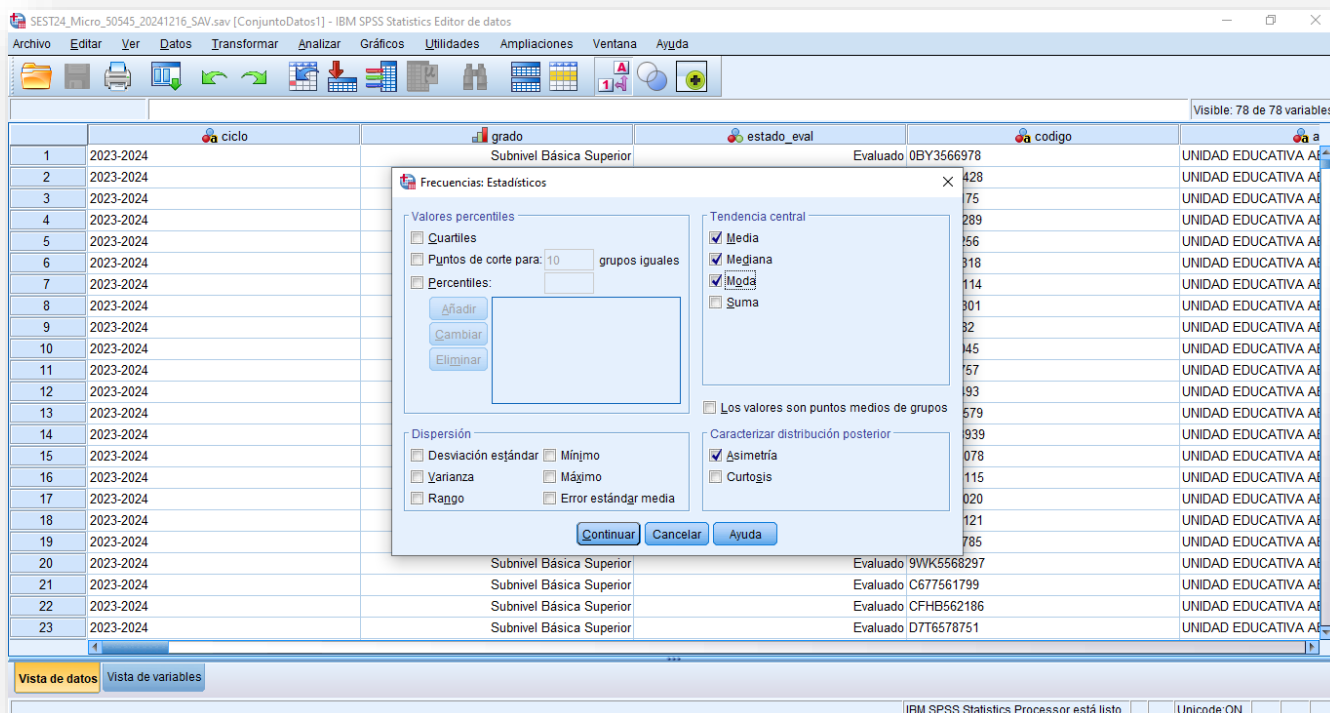
Clic en Frecuencias



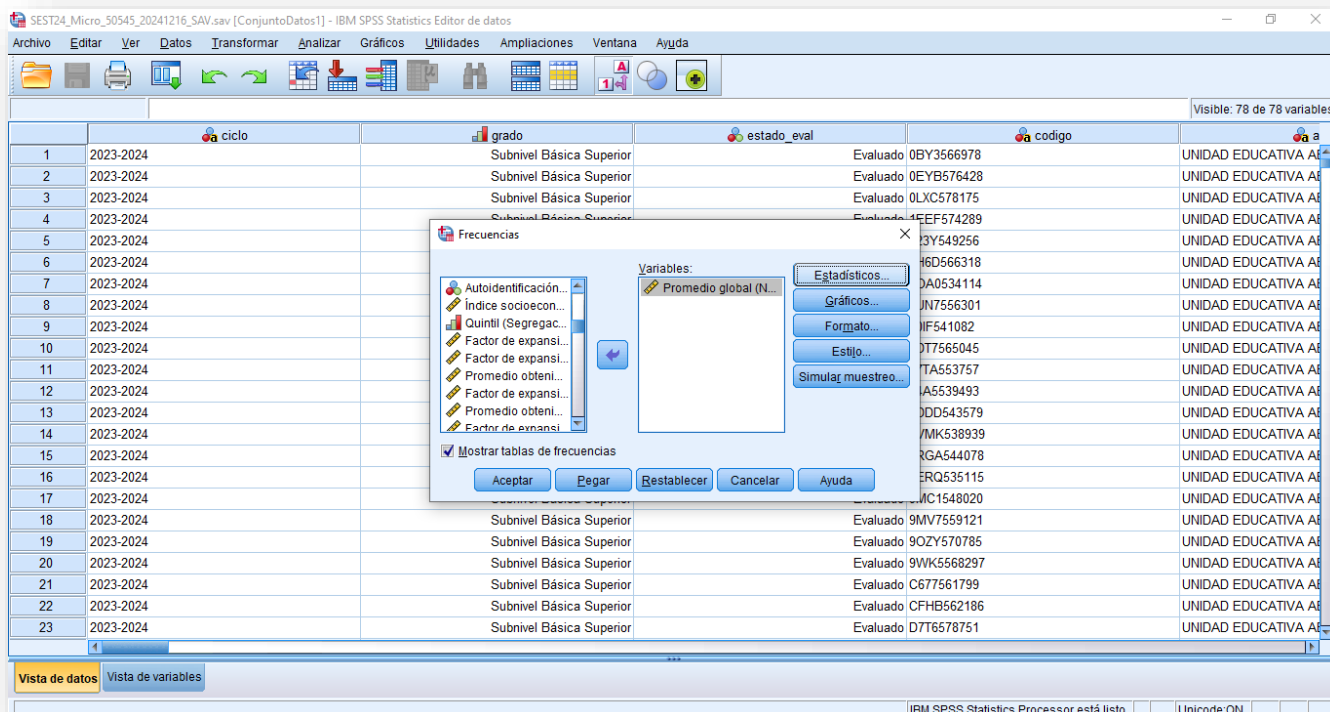
Seleccionar la variable Promedio global



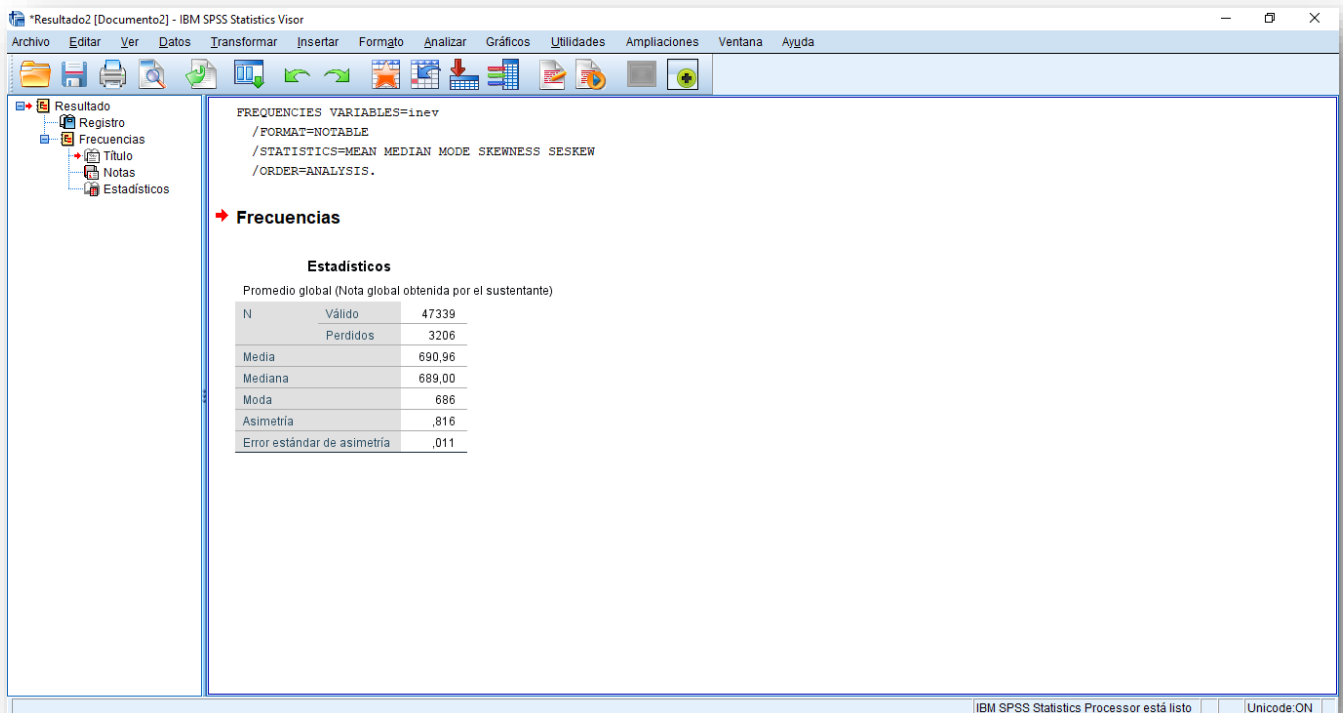
Clic en Estadísticos. Clic en Media, Mediana, Moda y Asimetría



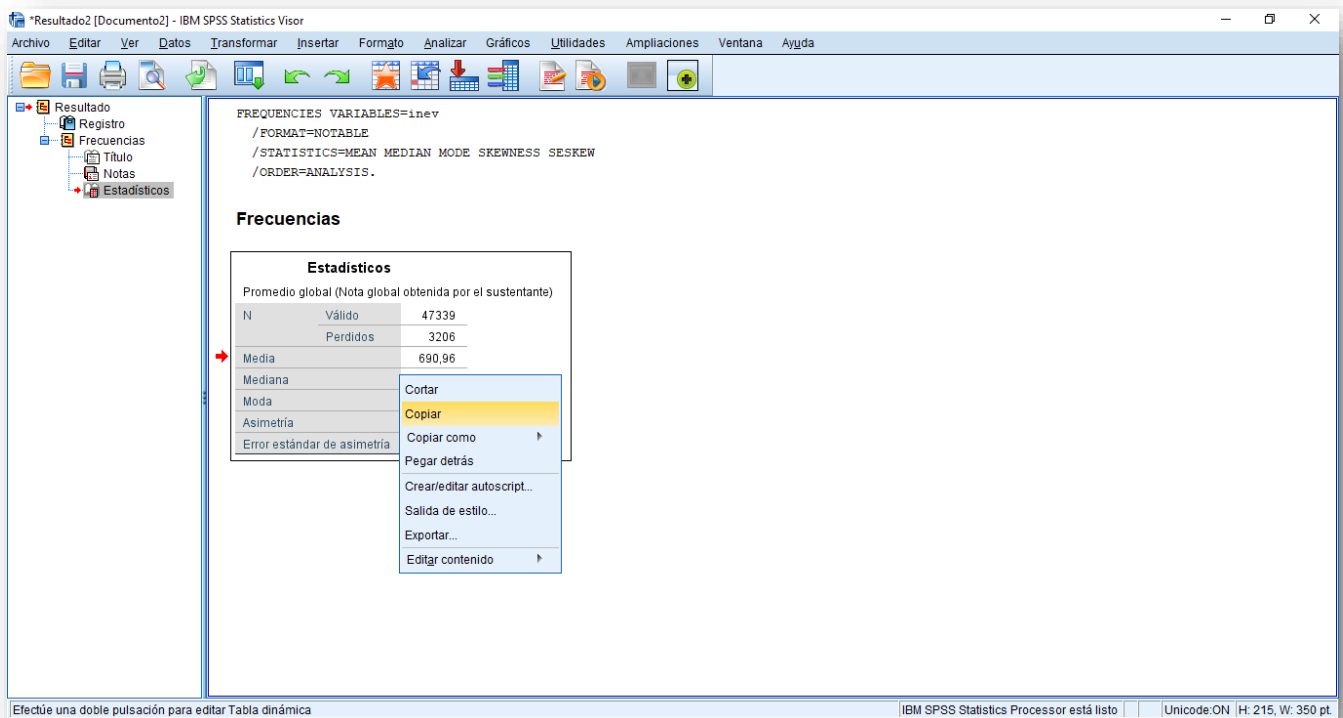
Clic en continuar



Clic en Mostrar tablas de frecuencias (se desactiva esta opción). Clic en Aceptar



Clic derecho en la tabla. Copiar



Pegar

Estadísticos

Promedio global (Nota global obtenida por el sustentante)

N	Válido	47339
	Perdidos	3206
Media		690,96
Mediana		689,00
Moda		686
Asimetría		,816
Error estándar de asimetría		,011

Interpretación:**Tamaño de la muestra:**

Válidos: 47 339 sustentantes (casos con datos disponibles).

Perdidos: 3 206 sustentantes (casos sin datos, ya sea por no presentarse o por problemas en el registro).

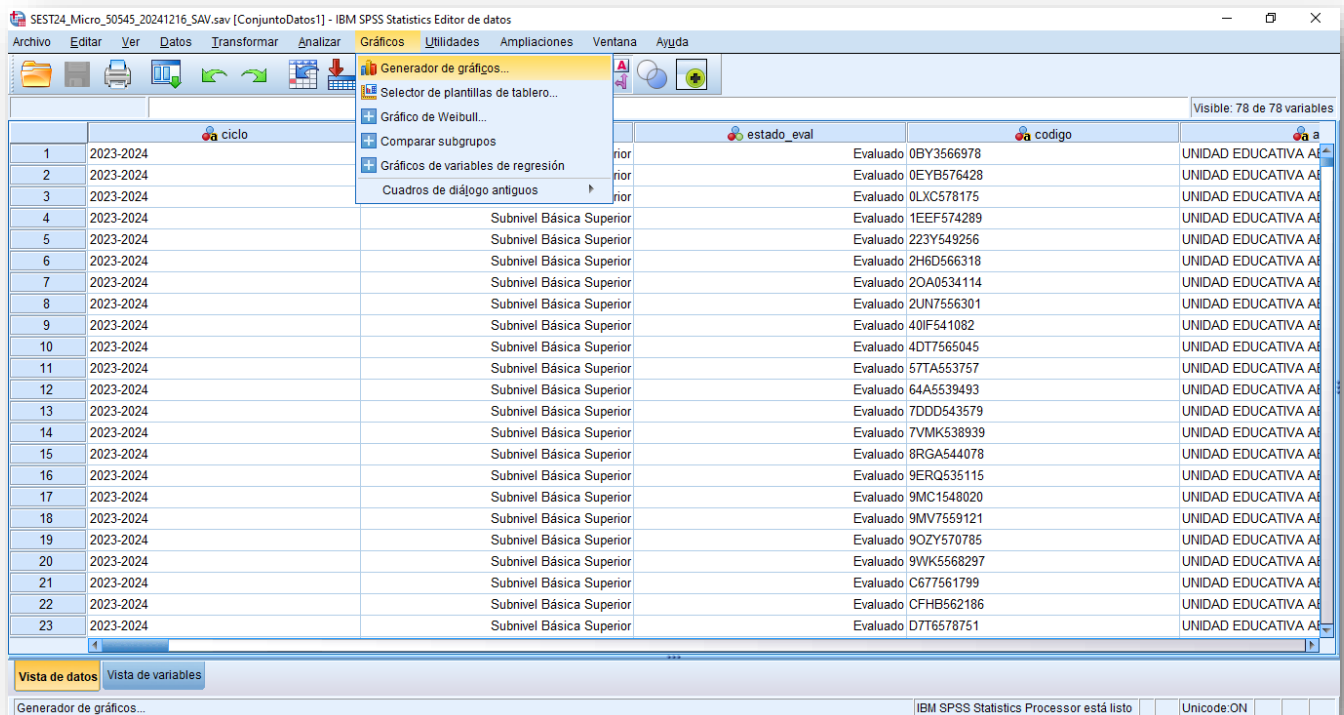
Total de sustentantes: $47,339 + 3,206 = 50,545$ (siendo el 6.3% de datos perdidos, lo que sugiere que la mayoría de los datos están completos).

Como $\bar{x} > Me > Mo$ ($690,96 > 689,00 > 686,00$), la distribución del Promedio global tiene una cola más larga hacia la derecha (asimetría positiva). Es decir:

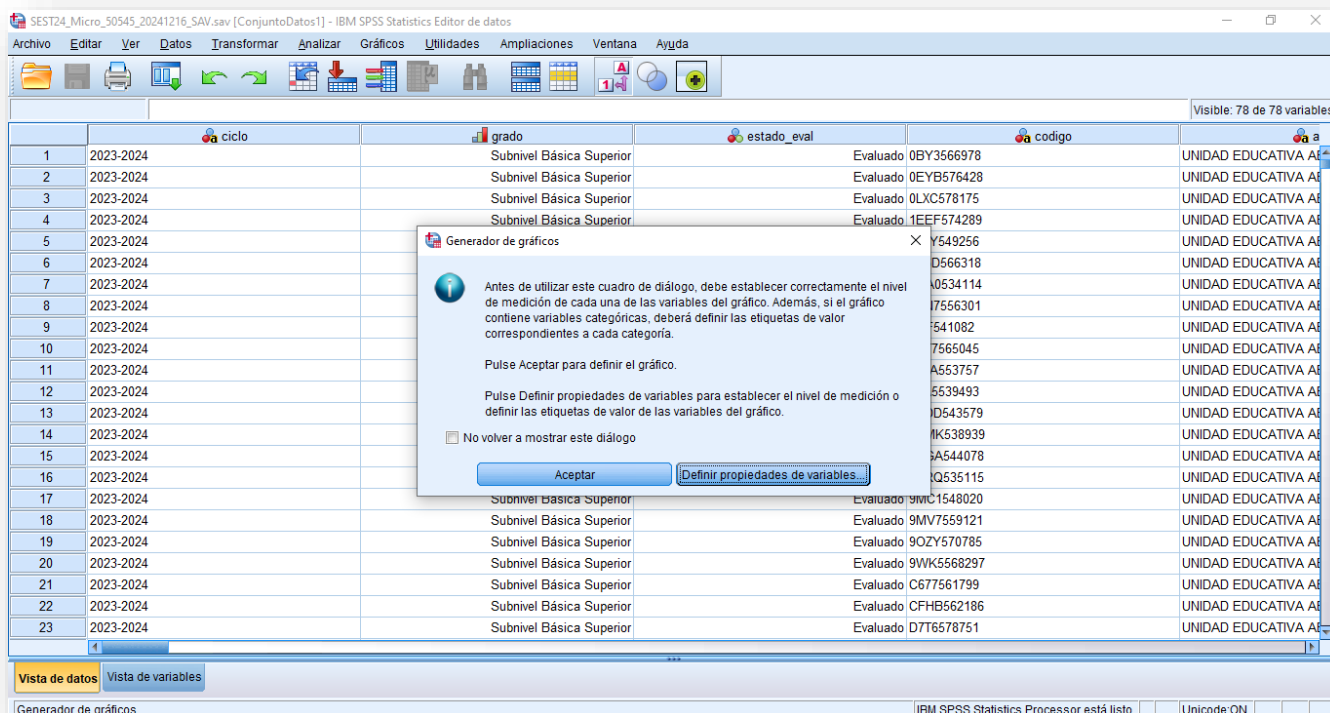
- ✓ La mayoría de los sustentantes obtuvieron puntuaciones por debajo de la media.
- ✓ Pocos sustentantes obtuvieron puntuaciones excepcionalmente altas, lo que "arrastra" la media hacia la derecha.
- ✓ La media es arrastrada por valores altos, por lo que es recomendable usar la mediana como medida de tendencia central.

Para observar gráficamente esta distribución empleando SPSS se procede de la siguiente manera:

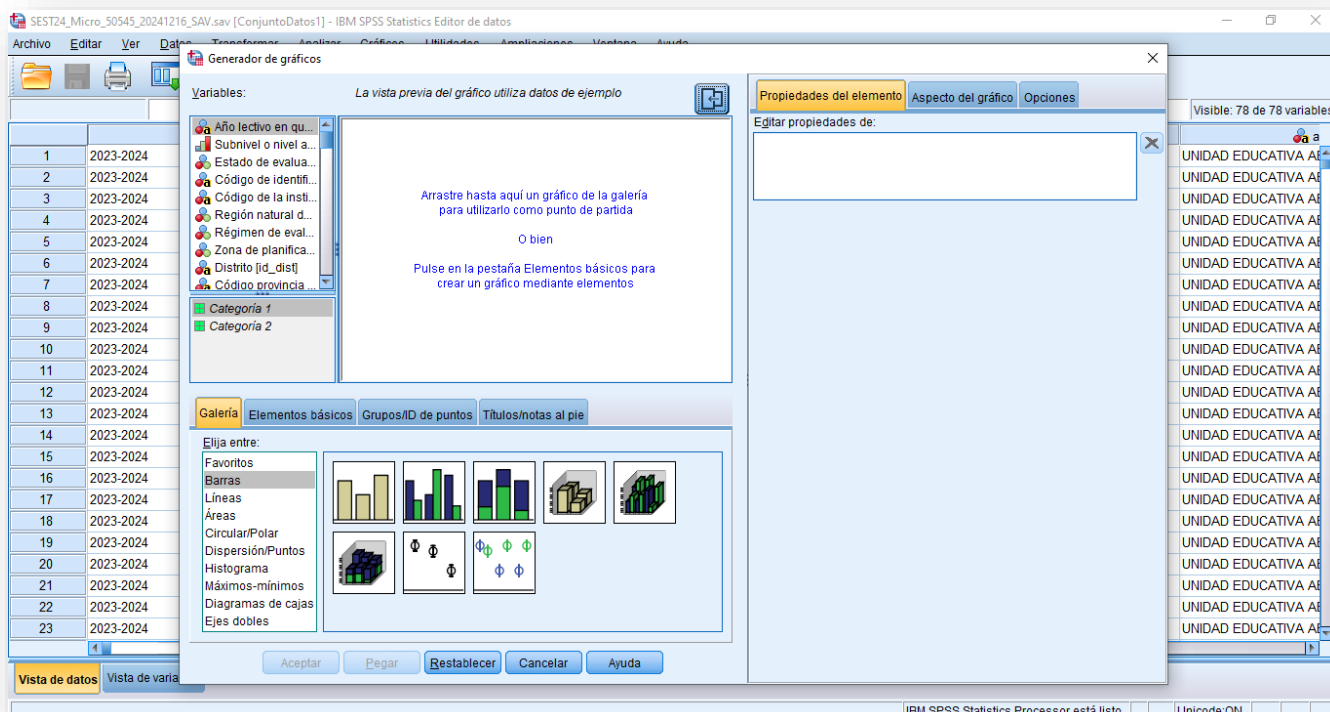
Clic en Gráficos, Generador de gráficos



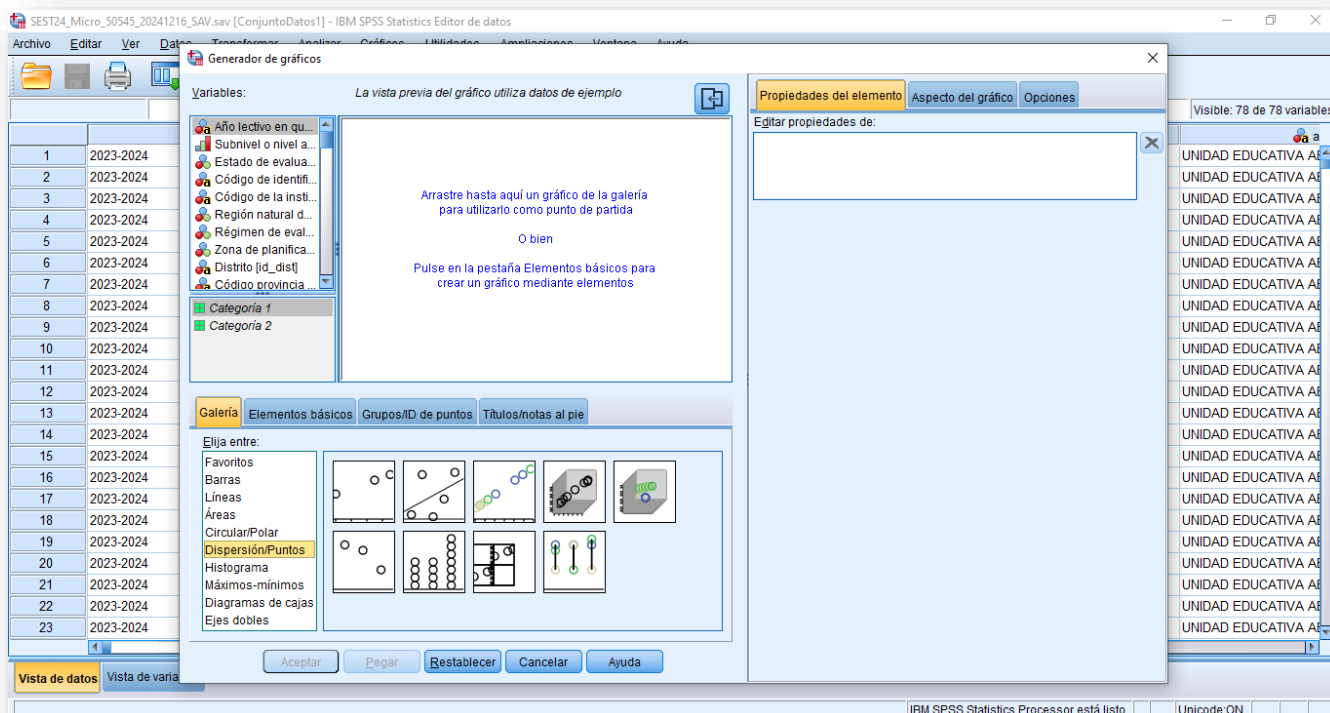
Clic en Generador de gráficos



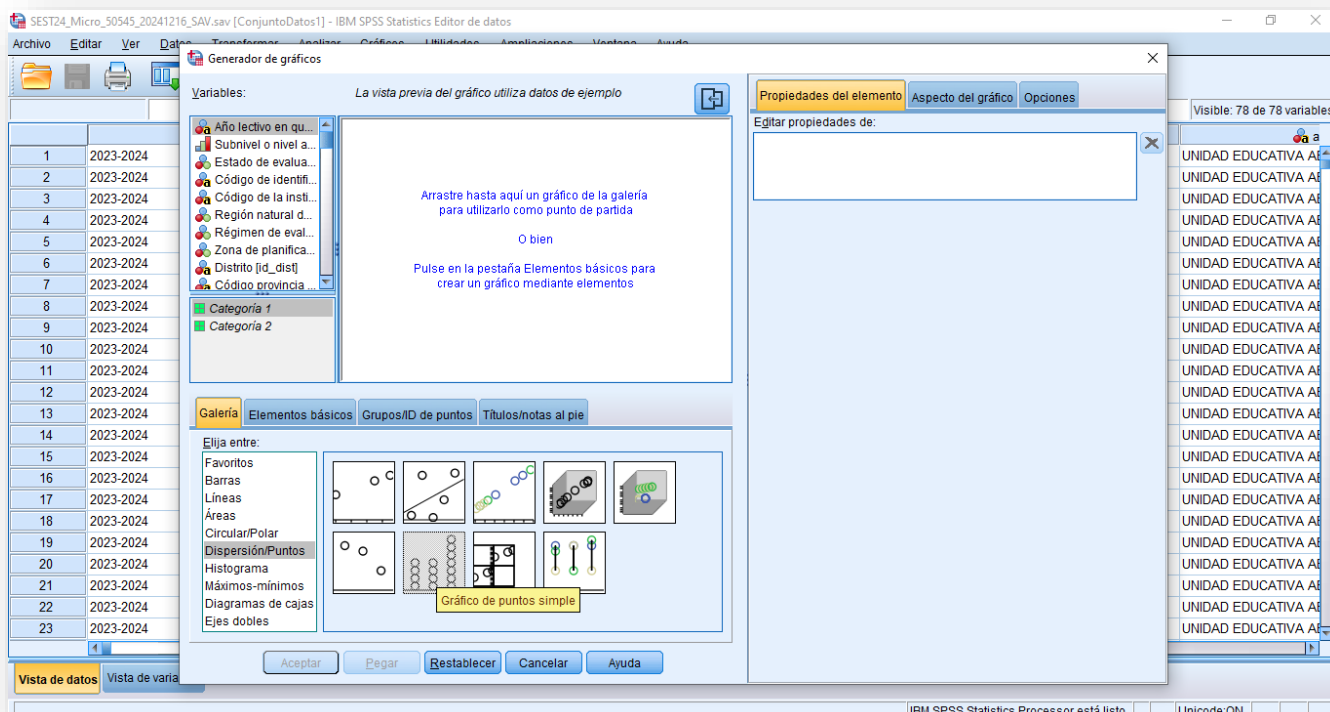
Clic en Aceptar



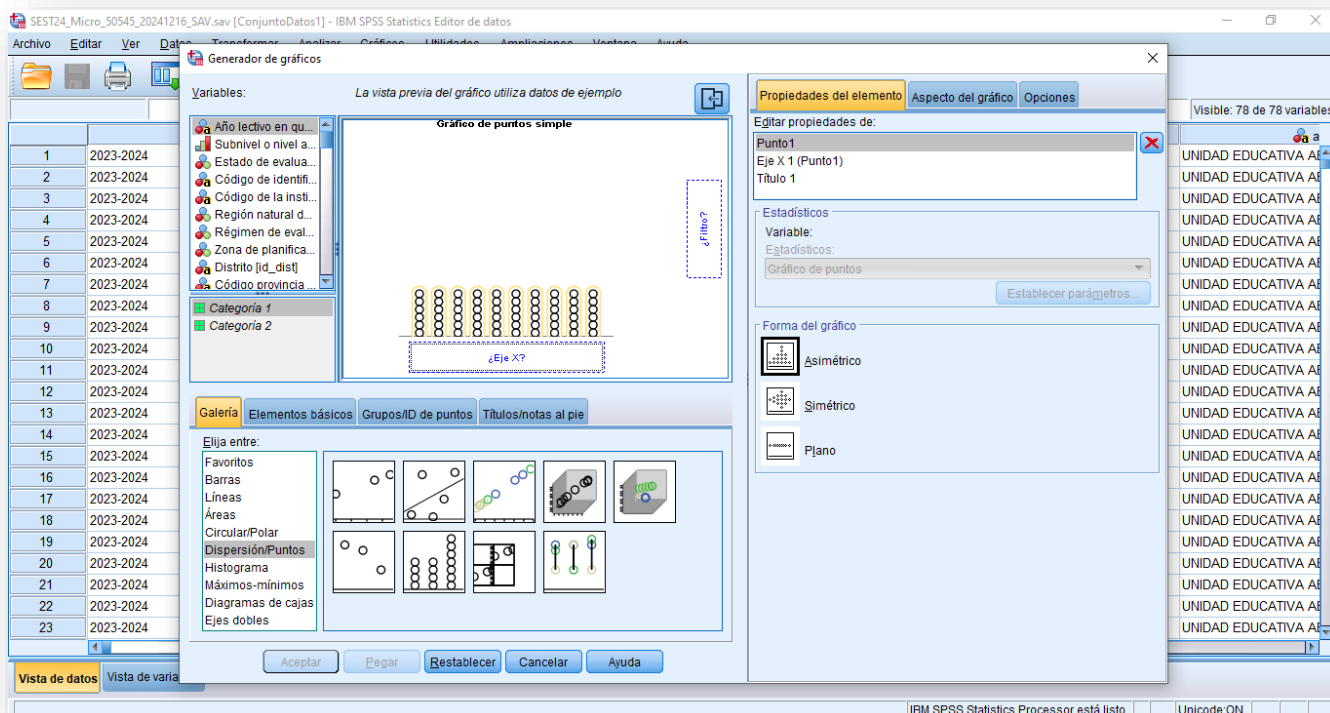
Clic en Dispersión Puntos



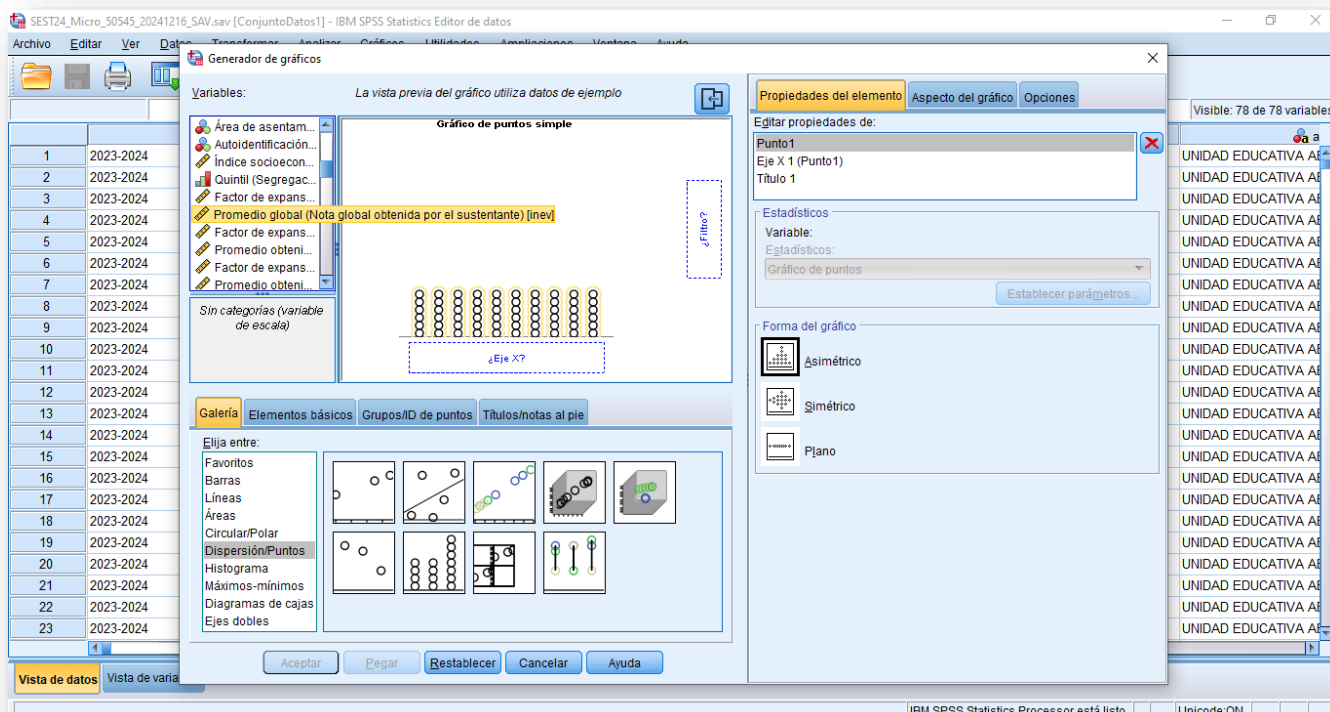
Clic en Gráfico de puntos simple



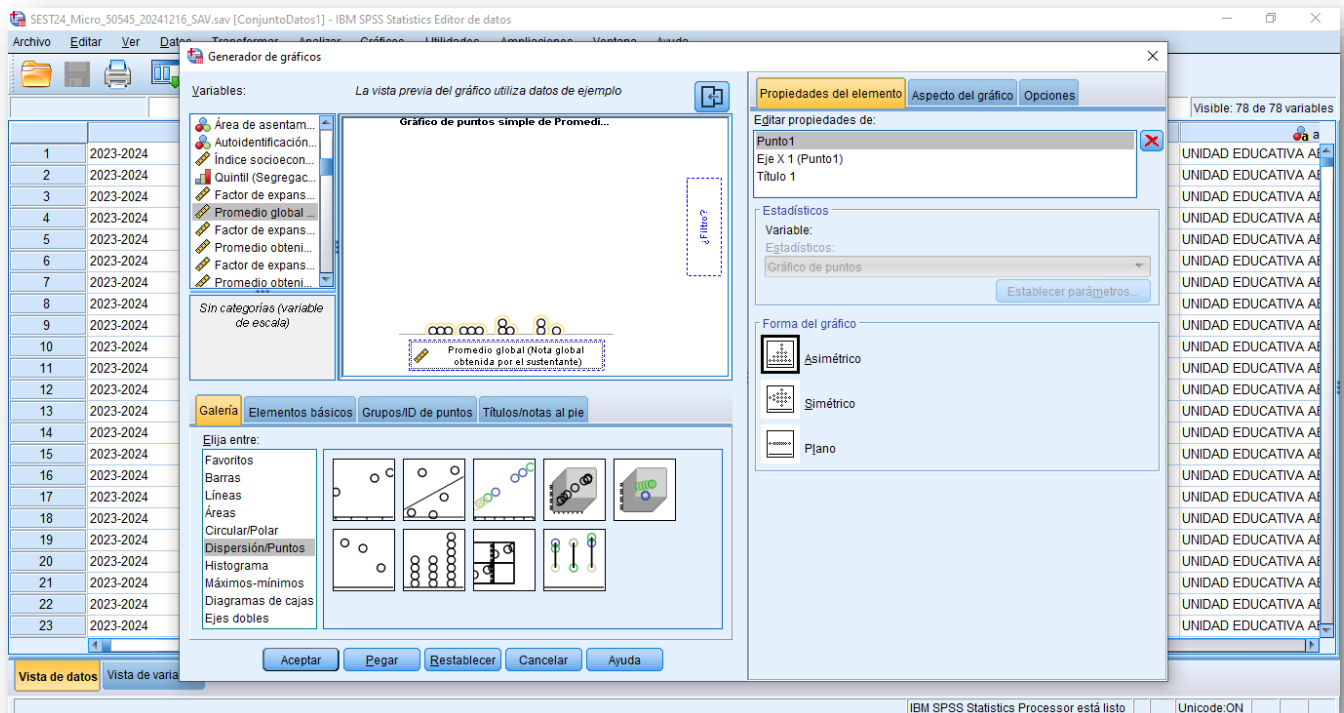
Arrastrar Gráfico de puntos simple



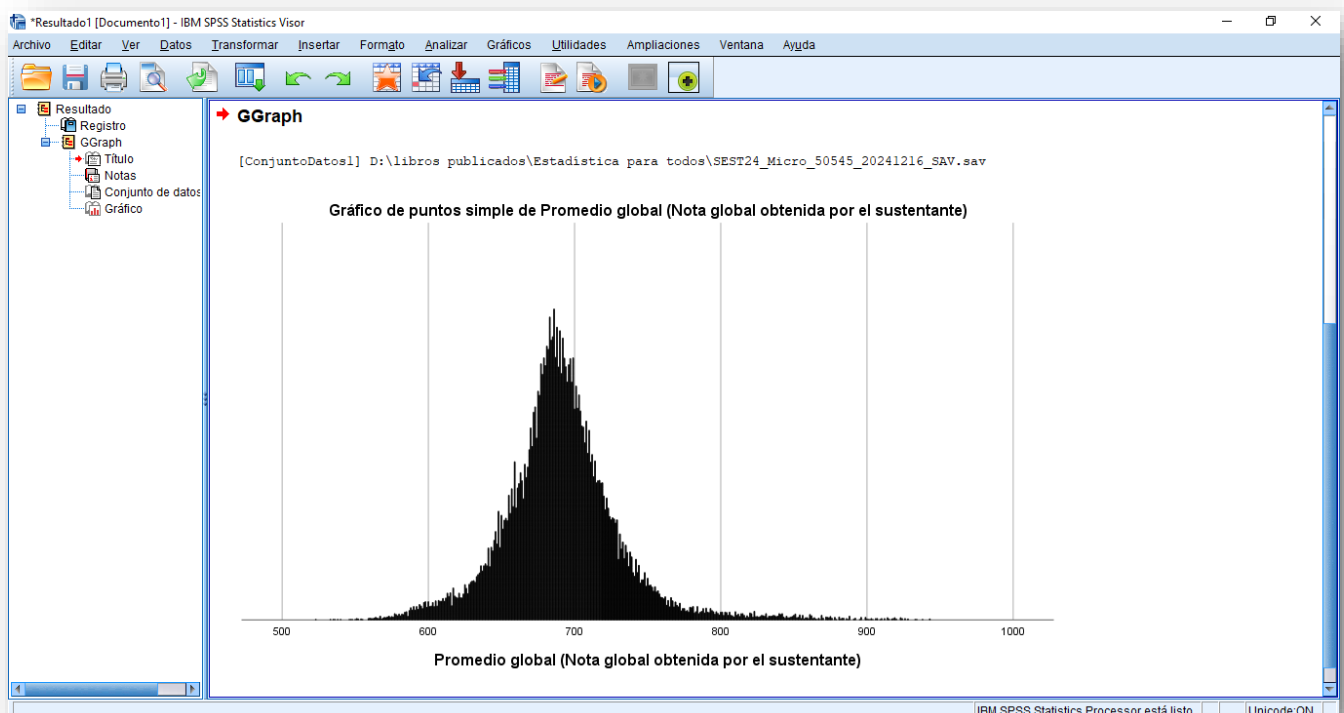
Seleccionar la variable promedio global



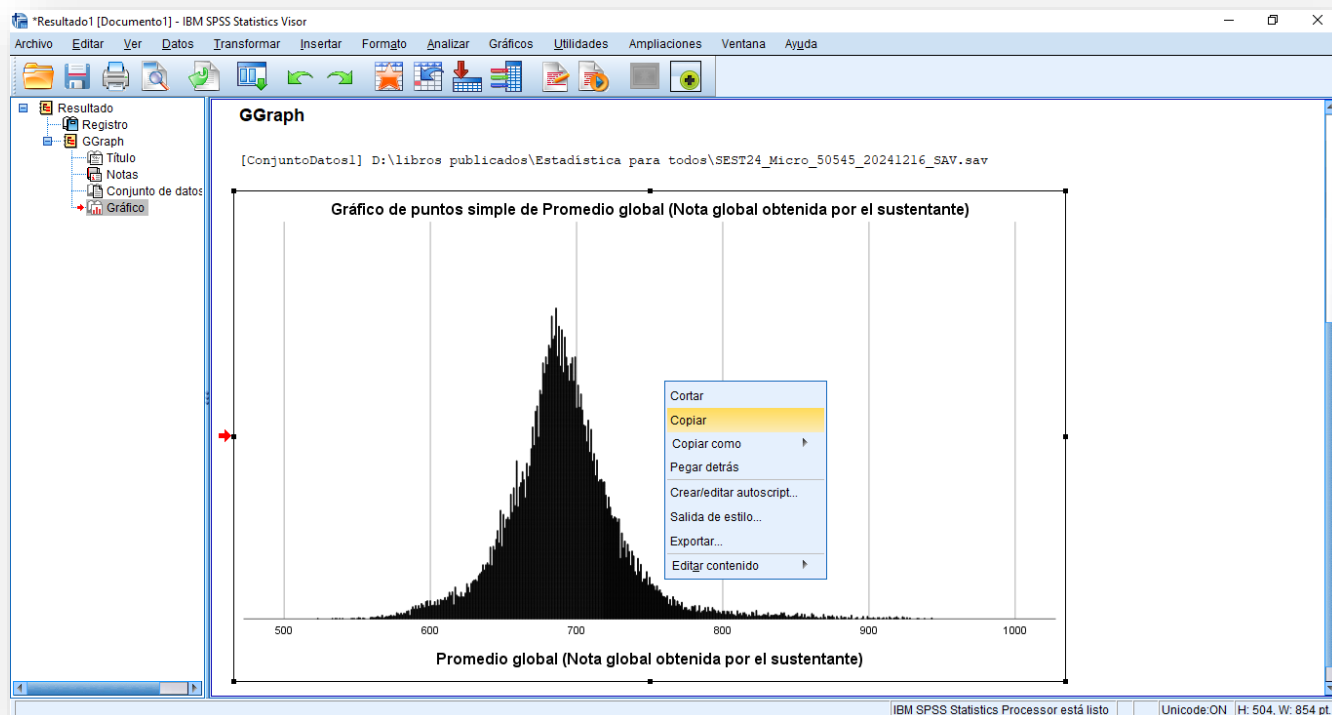
Arrastre la variable promedio global en el eje X del gráfico



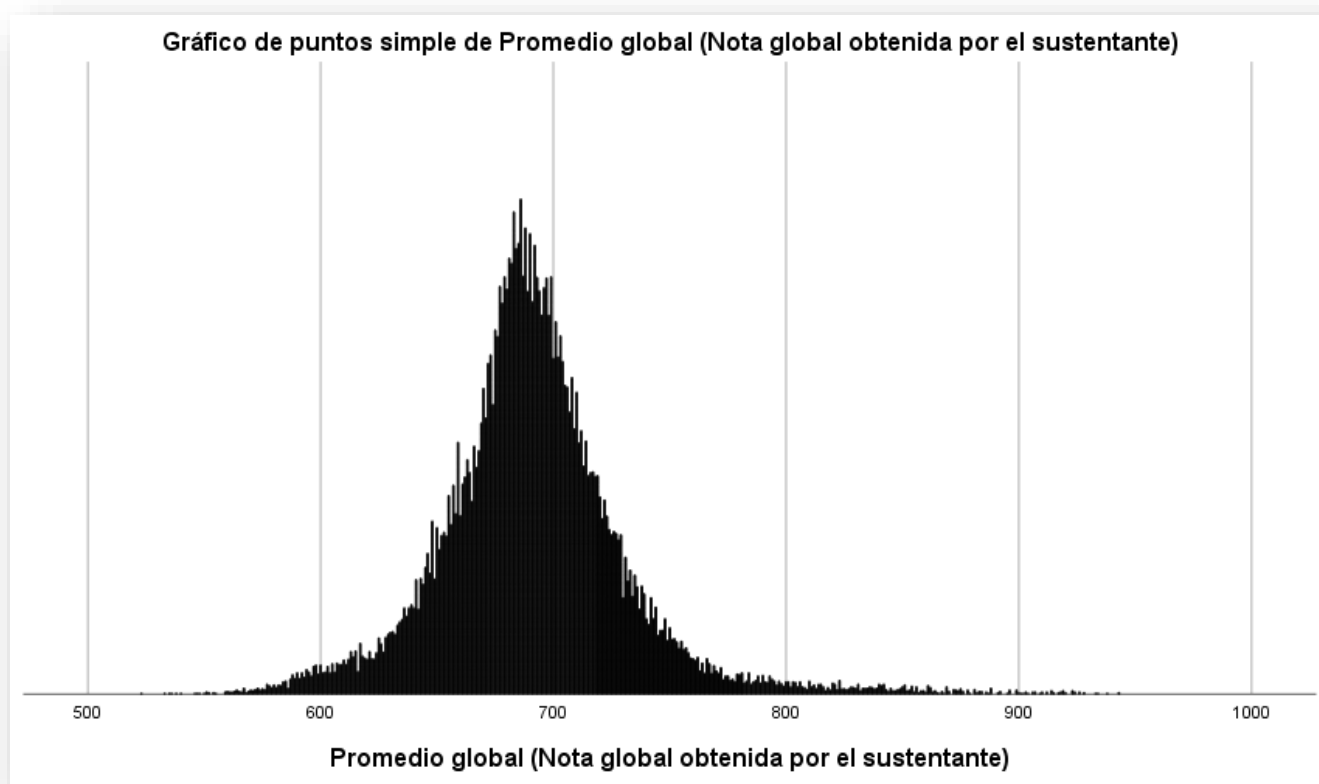
Clic en Aceptar



Clic derecho en el gráfico. Copiar

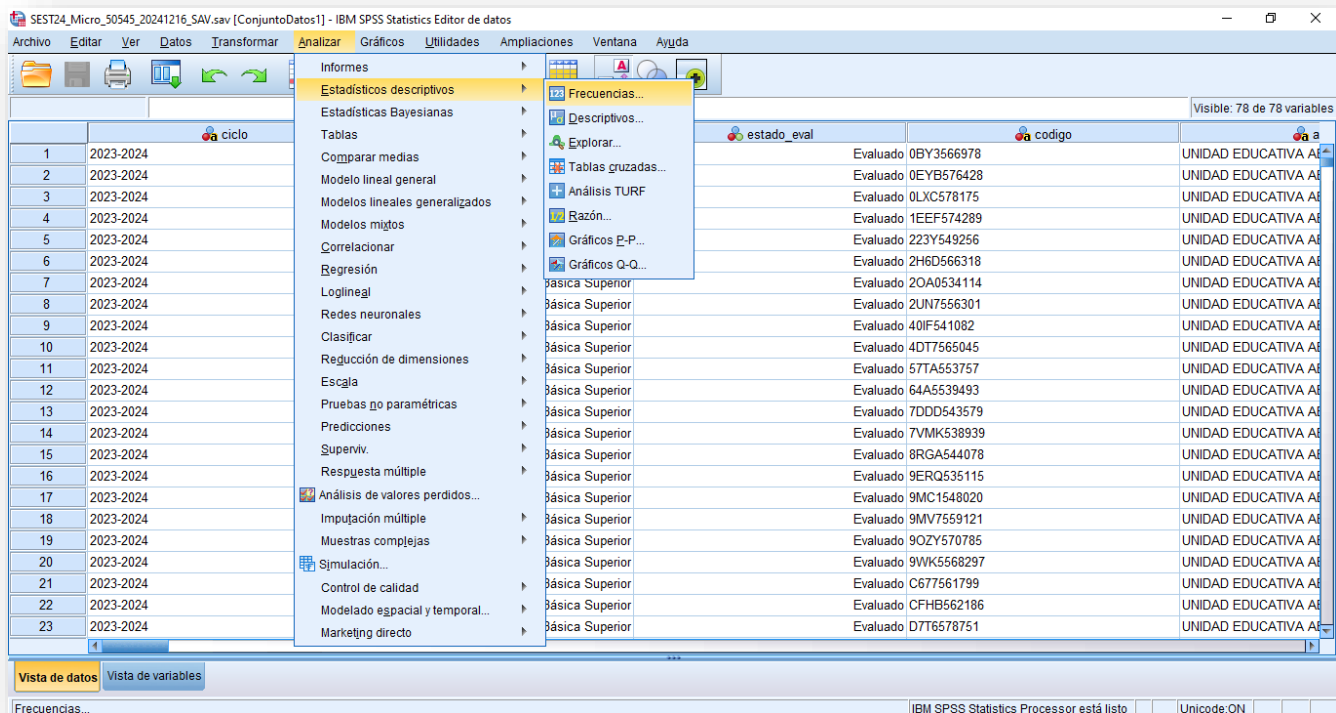


Pegar

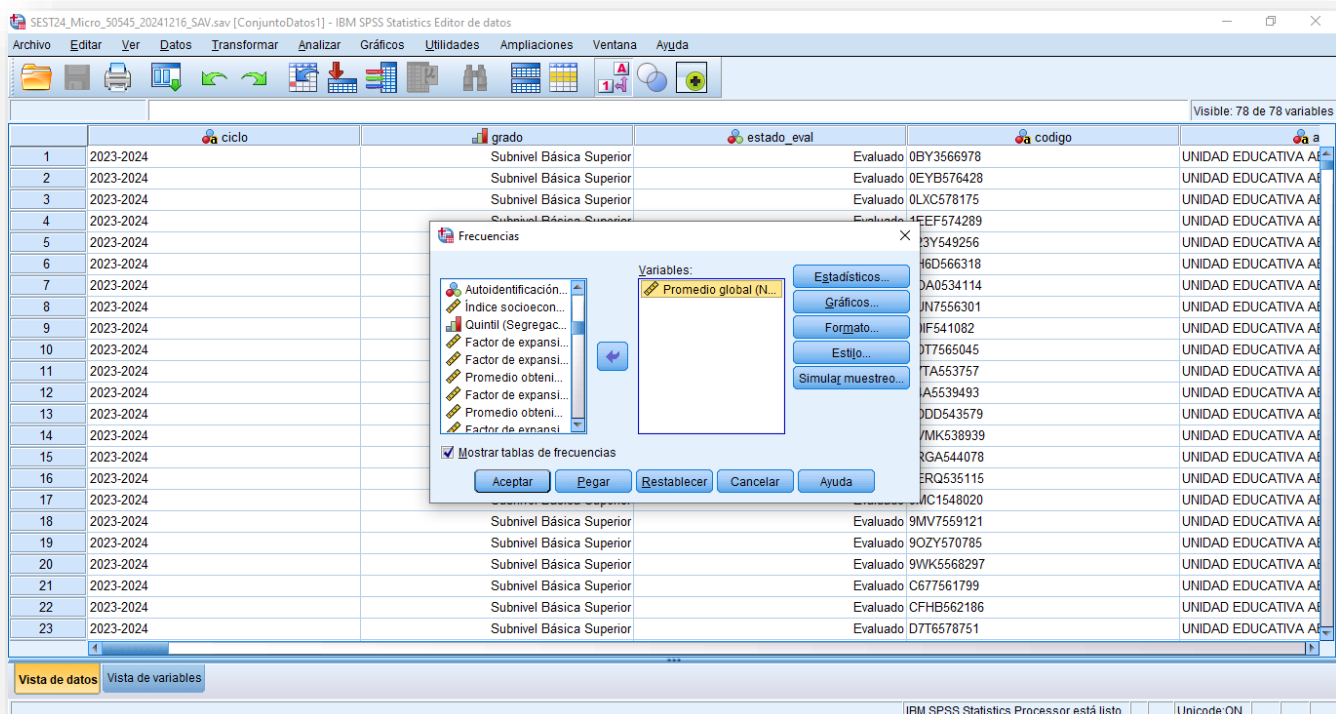


Otra forma de graficar es:

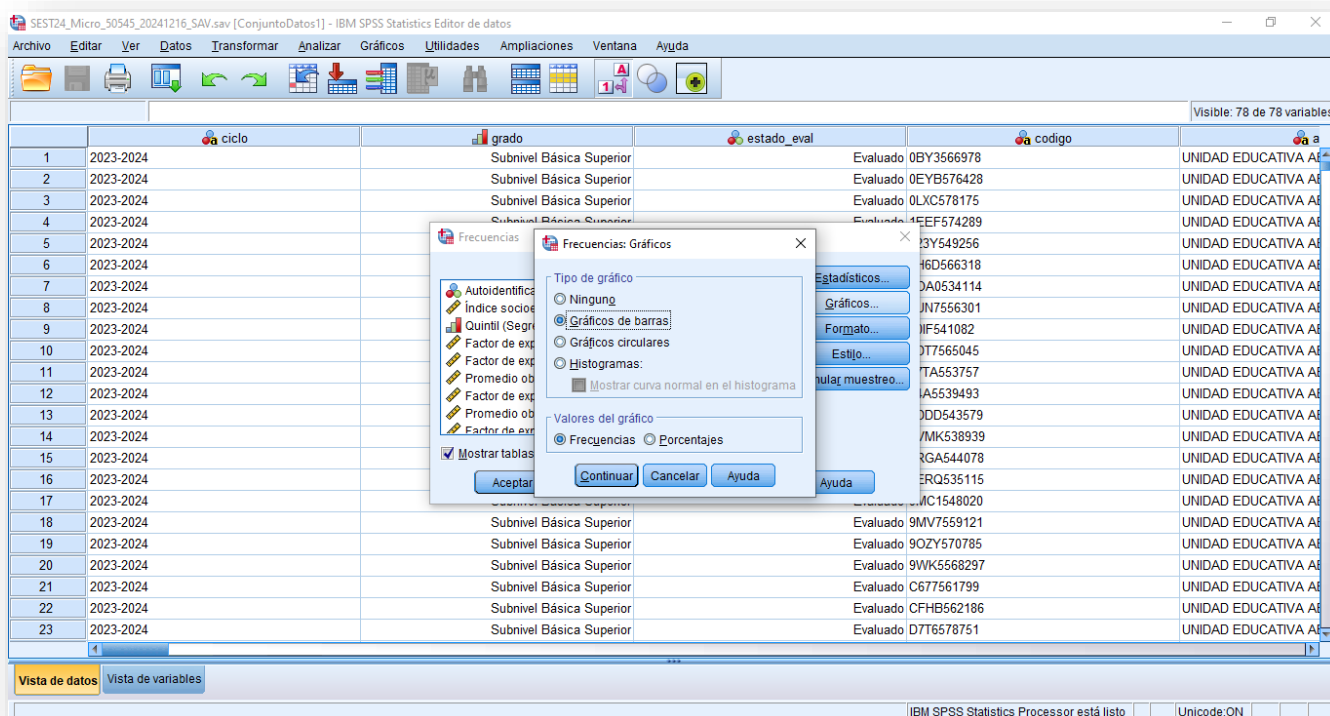
Clic en Analizar, Estadísticos descriptivos, Frecuencias



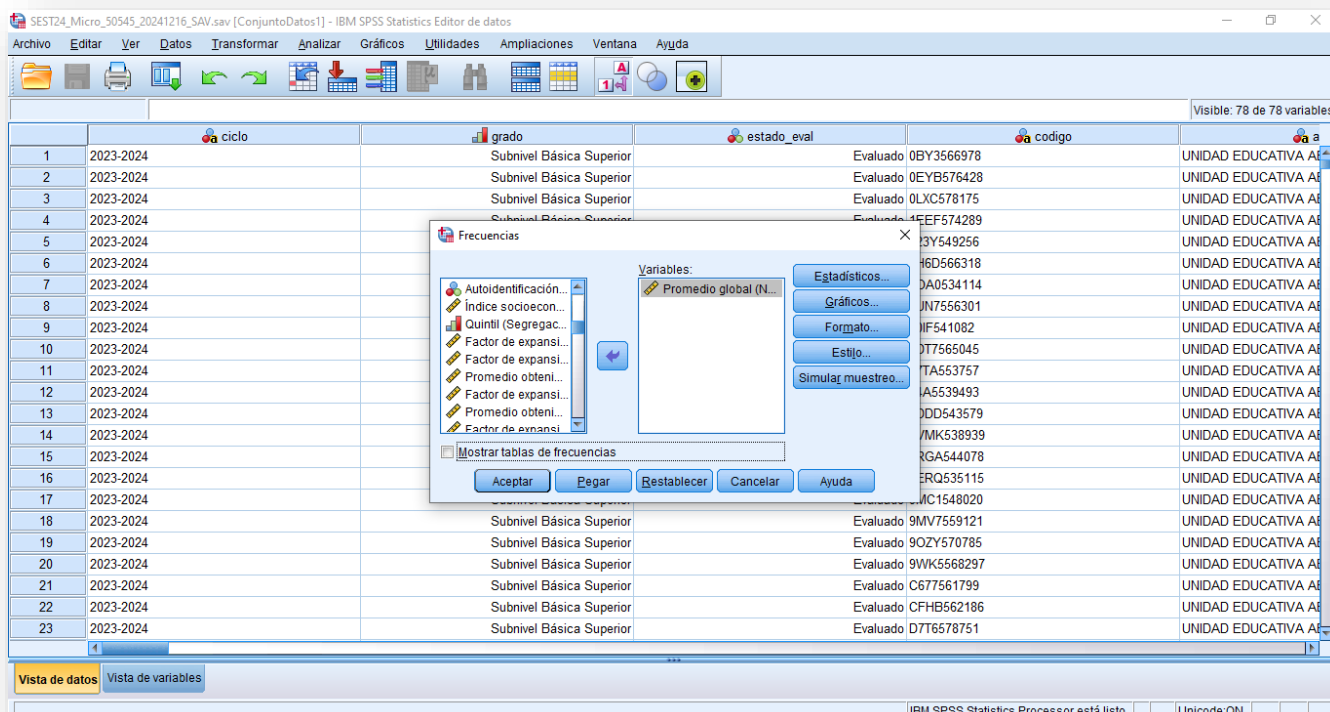
Clic en frecuencias. Seleccionar Promedio global



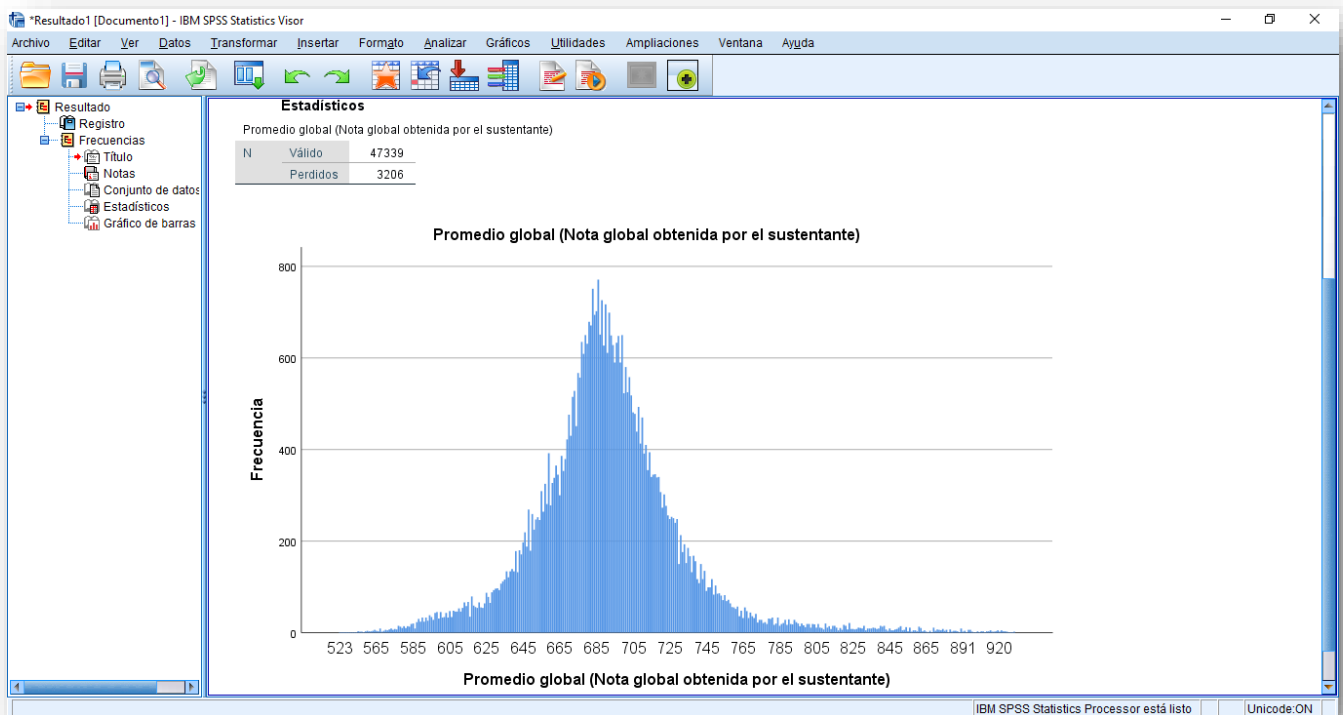
Clic en gráficos. Clic en diagrama de barras



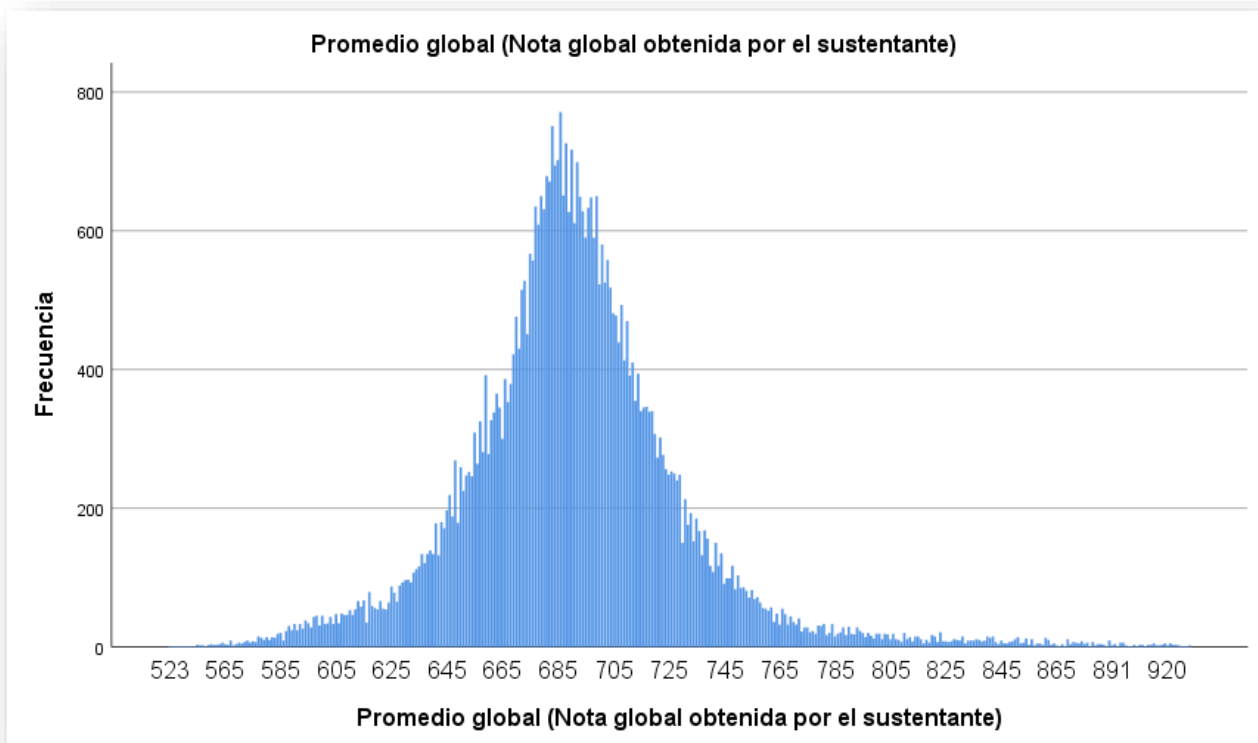
Clic en Continuar. Desactivar Mostrar tablas de frecuencias



Clic en Aceptar



Clic derecho en el gráfico. Copiar. Pegar



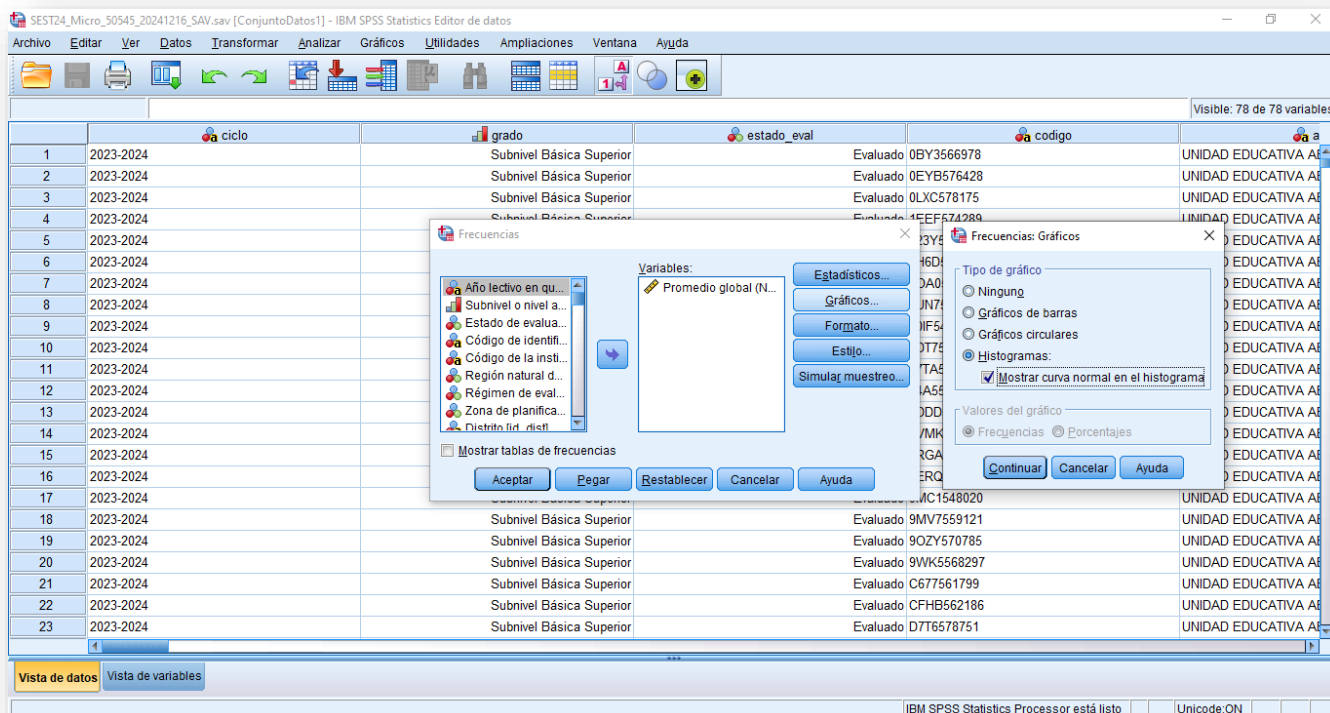
La distribución del Promedio global tiene una cola más larga hacia la derecha (asimetría positiva).

$\bar{x} > Me > Mo$ (690,96 > 689,00 > 686,00).

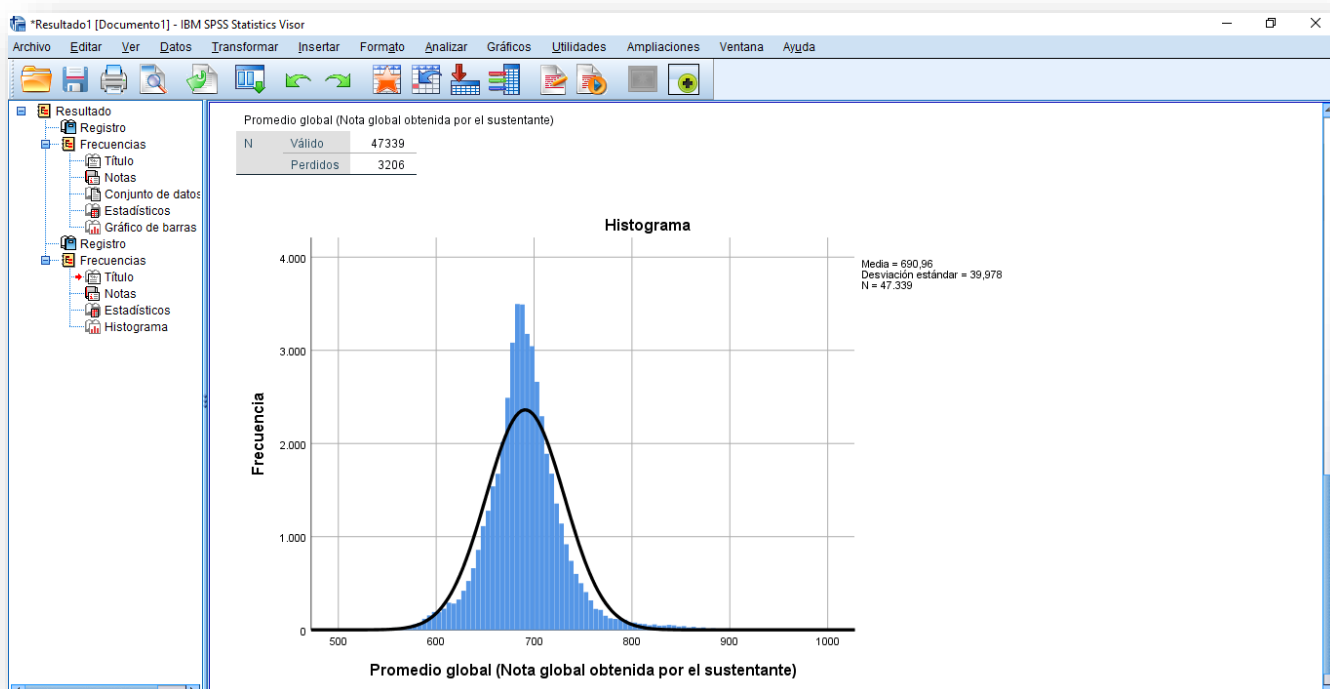
La distribución es sesgada a la izquierda (Pico a la izquierda, cola derecha alargada), es decir, la mayoría de los estudiantes tuvieron un desempeño modesto, con una minoría alcanzando puntajes altos.

Otro gráfico es el Histograma con Curva Normal

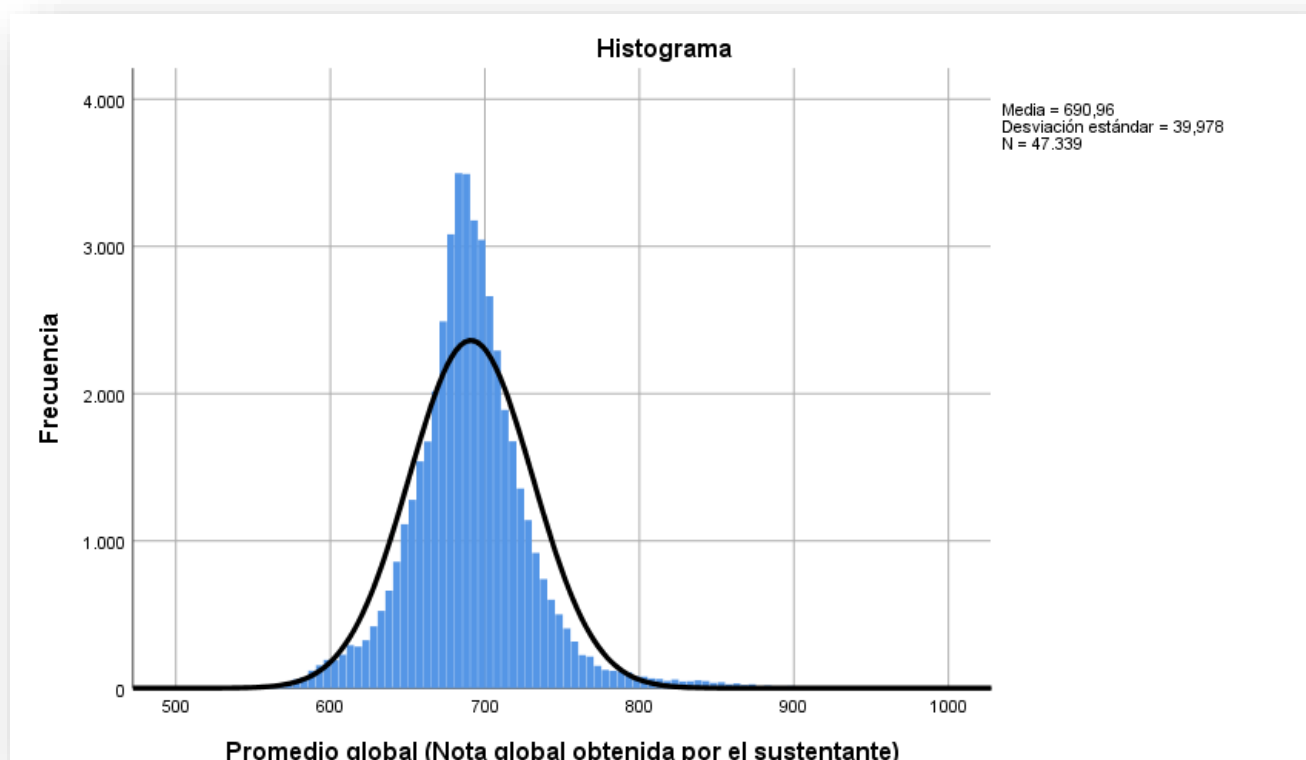
Repetir los pasos anteriores. Clic en Histogramas. Clic en Mostrar curva normal en el histograma



Clic en Continuar. Clic en Aceptar



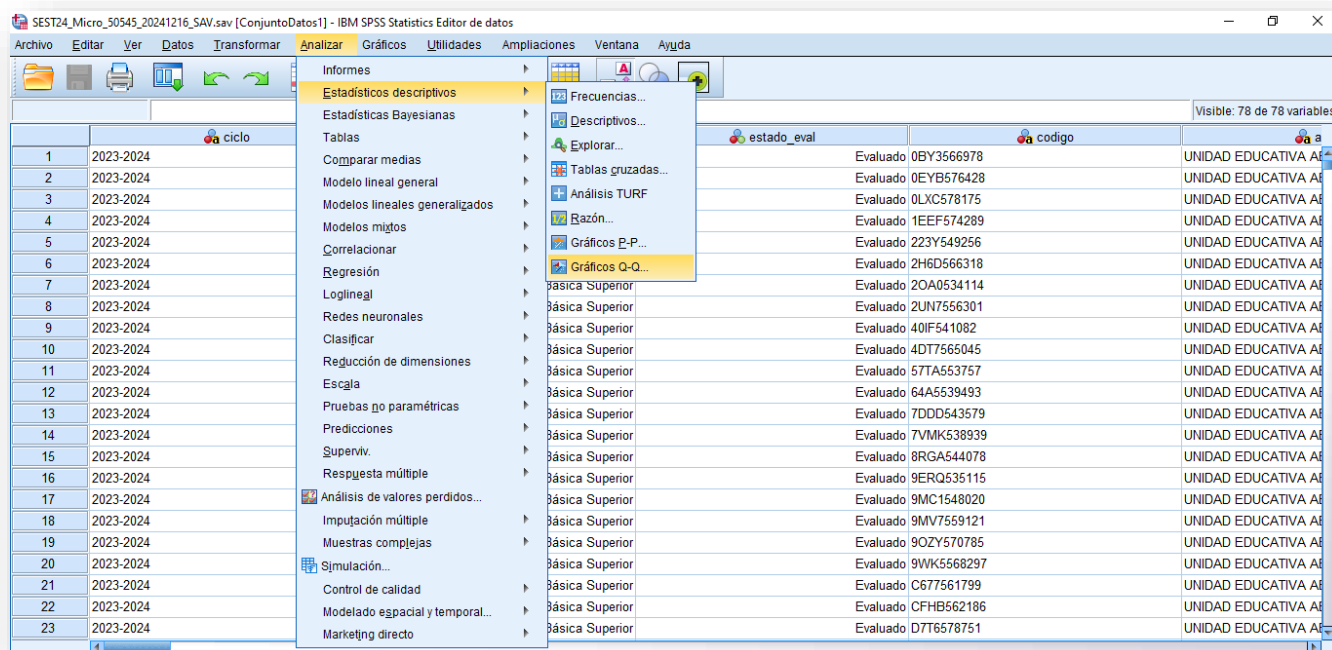
Clic derecho en el gráfico. Copiar. Pegar



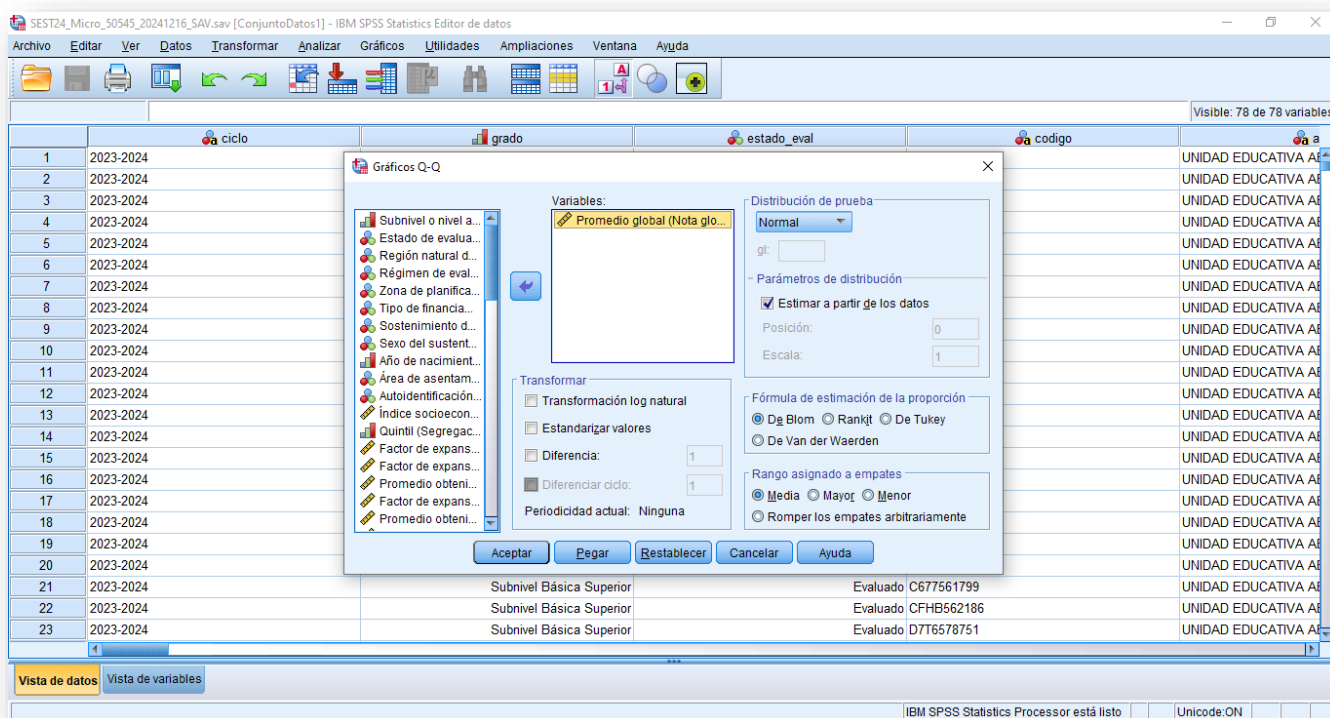
Interpretación:

Los datos tienen una cola derecha (asimetría positiva), con mayoría de notas concentradas a la izquierda. La curva normal (línea negra) no se ajusta bien.

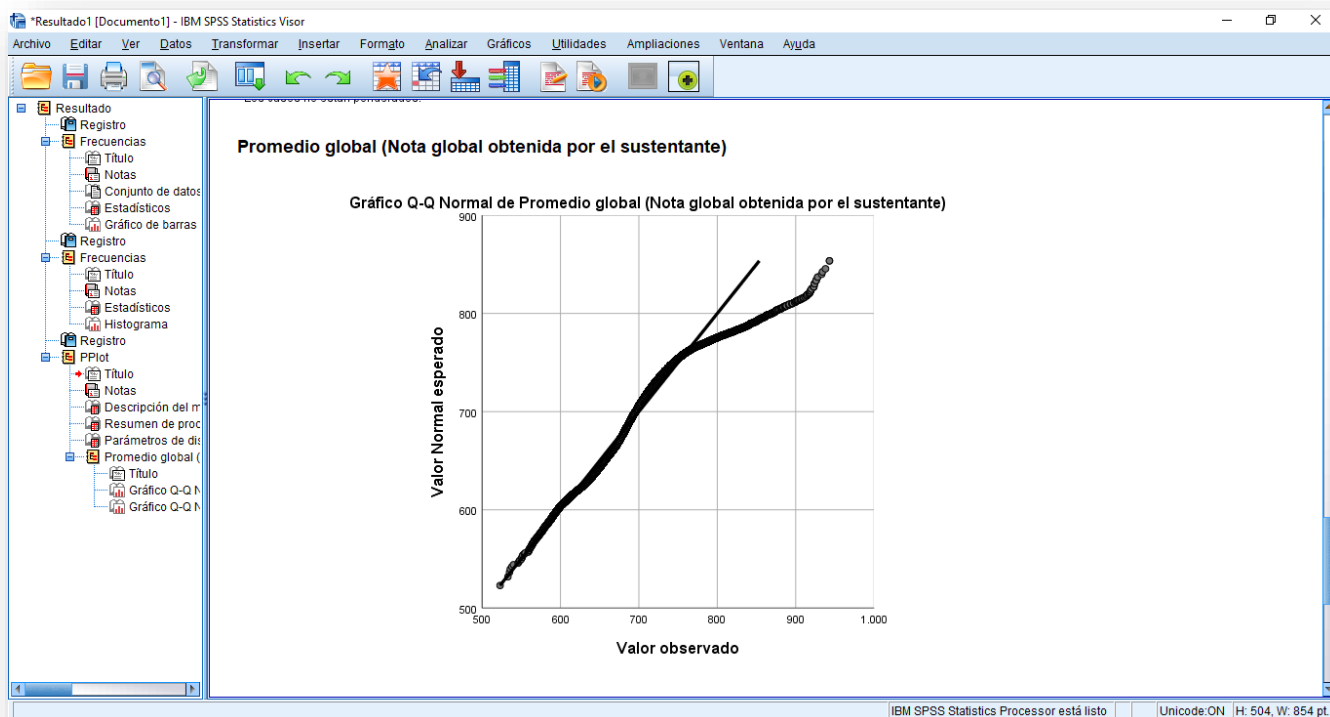
Otro gráfico es el Q-Q Plot. Clic en Analizar, Estadísticos descriptivos, Gráficos Q-Q



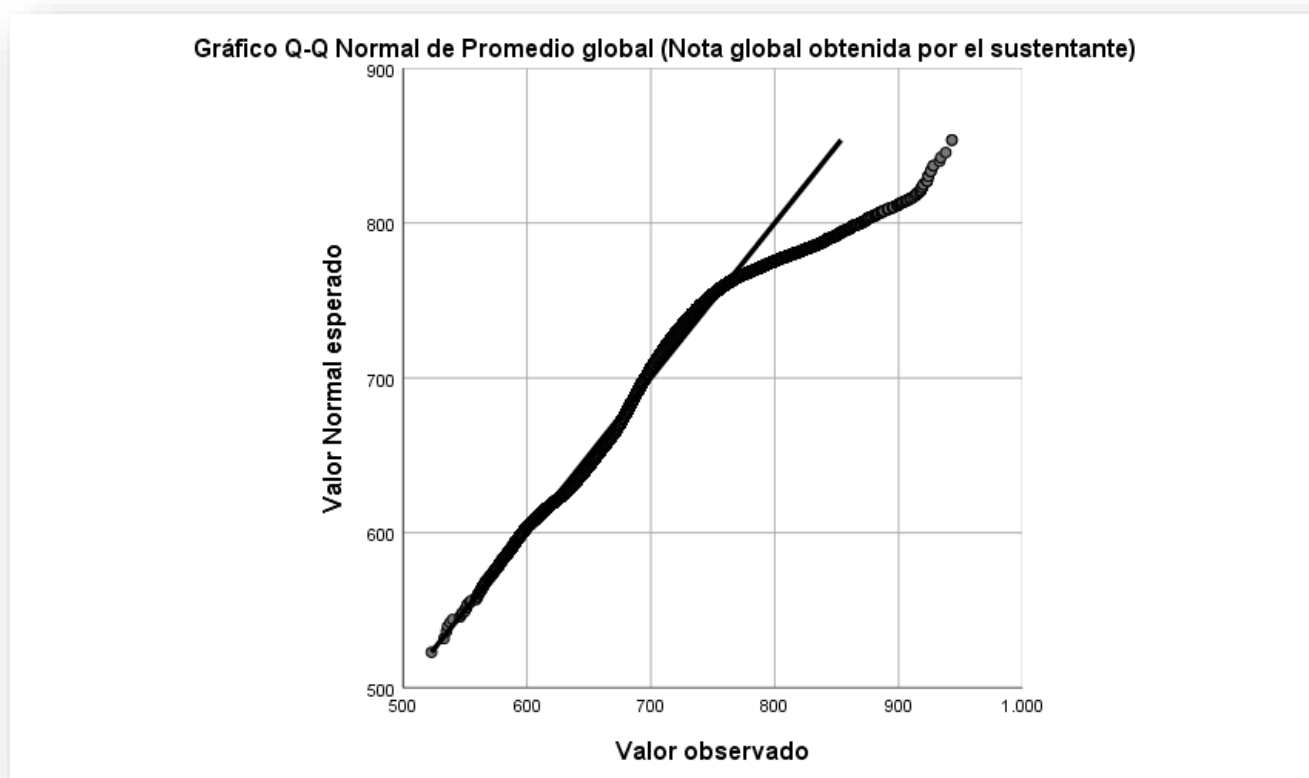
Clic en Gráficos Q-Q. Seleccionar la variable Promedio global



Clic en Aceptar



Clic derecho en el gráfico. Copiar. Pegar



Interpretación:

El Q-Q Plot (Quantile-Quantile Plot) es una herramienta gráfica utilizada en estadística para comparar la distribución de un conjunto de datos con una distribución teórica (Ejemplo, distribución normal). Su objetivo es determinar si los datos siguen una distribución específica o si existen desviaciones sistemáticas.

En este caso, los puntos se desvían de la línea diagonal (45°) en los extremos (especialmente hacia arriba en la cola derecha), confirmando no normalidad (si los datos siguieran una distribución normal, los puntos estarían alineados).

Asimetría:

Valor: 0,816 (positivo).

Valor de g_1	Tipo de Asimetría	Forma de la Distribución	Interpretación
$g_1 \geq 0,5$	Moderada/Fuerte positiva	Cola derecha pronunciada.	Valores altos distorsionan la distribución. $\bar{x} \gg Me$

Error estándar de asimetría:

Valor de SPSS: 0,011 (muy bajo).

Calculando el error estándar de asimetría empleando la fórmula, para nuestro ejemplo con $n = 47339$ sería:

$$SE_{asimetría} = \sqrt{\frac{6}{n}} = \sqrt{\frac{6}{47339}} = 0,011$$

Lo que coincide con el valor calculado con SPSS

Aplicando la regla práctica

$$|g_1| > 2 \cdot SE_{asimetría}$$

$$|0,816| > 2 \cdot 0,011$$

$$|0,816| > 0,022 \text{ (verdadero)}$$

Como se cumple la desigualdad, se concluye que la asimetría observada (0,816) es estadísticamente significativa (la asimetría no se debe al azar, es no aleatorio).

El valor de asimetría 0,816 es significativamente mayor que su error estándar (0,011), lo que indica un patrón no aleatorio. Una posible causa podría ser la dificultad de la prueba o desigualdad en la preparación de los estudiantes. Si fuera aleatorio, la asimetría estaría cerca de cero ($|g_1| \leq 2 \cdot SE_{asimetría}$), sugiriendo que las variaciones se deben solo al azar, pero no es el caso en este ejemplo.

Por lo tanto, la distribución del Promedio global en la evaluación Ser Estudiante del año lectivo 2023-2024 es significativamente asimétrica (no es simétrica y tiene una cola derecha más larga con un 95% de confianza aproximado).

1.6) ACTIVIDADES DE APLICACIÓN

- 1) Realice un organizador gráfico del capítulo I
- 2) Empleando Excel calcule la media, error típico, mediana, moda, desviación estándar, varianza de la muestra, coeficiente de asimetría, rango, mínimo, máximo, suma y cuenta de las siguientes calificaciones 4, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8, 8, 8, 8, 8, 9, 9, 10, 10, 10, 10. Interprete los resultados.

	A	B
1	Columna1	
2		
3	Media	7,55
4	Error típico	0,407010927
5	Mediana	8
6	Moda	8
7	Desviación estándar	1,820208201
8	Varianza de la muestra	3,313157895
9	Curtosis	-0,8137418
10	Coefficiente de asimetría	-0,234620694
11	Rango	6
12	Mínimo	4
13	Máximo	10
14	Suma	151
15	Cuenta	20

- 3) Con los datos del ejercicio anterior calcule el coeficiente de variación. Interprete los resultados.
- 4) Con los datos del ejercicio 2 calcule la regla práctica de la asimetría. Interprete los resultados.
- 5) Descargue la base de datos de la Prueba Ser Estudiante 2023-2024 (Micro SAV) disponible en el portal oficial: <https://evaluaciones.evaluacion.gob.ec/BI/ser-estudiante-2/>

liante-2/						
↓ Descargar Manual de Fichas Metodológicas 2019						
Año lectivo 2023-2024						
Archivo	Sintaxis	CSV	SAV	XLSX	Metadato	Diccionario
Micro	↓	↓	↓	↓	↓	↓
Factores Asociados estudiantes	↓	↓	↓	↓	↓	↓
Factores Asociados padres	↓	↓	↓	↓	↓	↓
Factores Asociados docentes	↓	↓	↓	↓	↓	↓
Factores Asociados directivos	↓	↓	↓	↓	↓	↓

a) Empleando SPSS calcule la media, error estándar de la media, error estándar de la media, mediana, moda, desviación estándar, varianza, rango, mínimo, máximo, suma, cuartiles, asimetría, error estándar de asimetría del promedio obtenido en el campo de Matemática (imat) y del promedio obtenido en el campo de Lengua y Literatura (ilyl). Copie y pegue la tabla. Interprete los resultados.

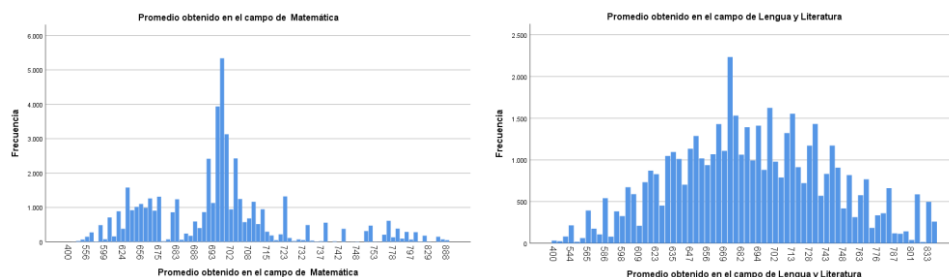
Estadísticos

N	Promedio obtenido en el campo de Matemática		Promedio obtenido en el campo de Lengua y Literatura
	Válidos	Perdidos	
		48297	48330
		2248	2215
Media		691,54	687,95
Error estándar de la media		,183	,271
Mediana		694,00	684,00
Moda		697	670
Desv. Desviación		40,178	59,635
Varianza		1614,247	3556,356
Asimetría		,145	,702
Error estándar de asimetría		,011	,011
Curtosis		3,139	3,445
Error estándar de curtosis		,022	,022
Rango		488	600
Mínimo		400	400
Máximo		888	1000
Suma		33399191	33248716
Percentiles	25	675,00	649,00
	50	694,00	684,00
	75	706,00	728,00

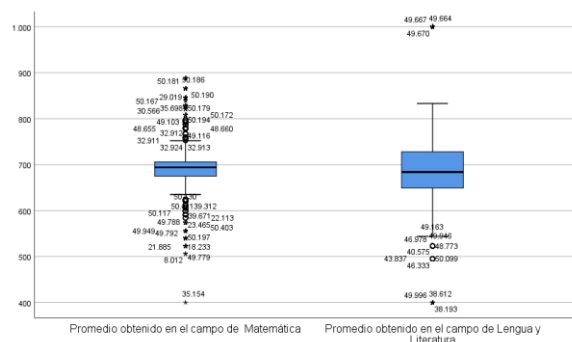
b) Calcule el coeficiente de variación para imat y para ilyl. Interprete los resultados.

c) Calcule la regla práctica de la asimetría para imat y para ilyl. Interprete los resultados.

d) Elabore diagramas de barras para imat y para ilyl. Interprete los resultados.



e) Calcule el IQR, límites inferior y superior de los bigotes y los valores atípicos para imat y para ilyl. Empleando SPSS elabore un diagrama de Caja y Bigotes. Interprete los resultados.



6) Cree y resuelva un ejercicio similar al anterior empleando la misma base de datos.

CAPÍTULO II

CURTOSIS

2.1) DEFINICIÓN

Es una medida estadística que cuantifica el grado de concentración de los datos alrededor de la media, específicamente evaluando el "peso" de las colas de la distribución en comparación con una distribución normal (mesocúrtica), es decir, mide cuán puntiaguda es una distribución.

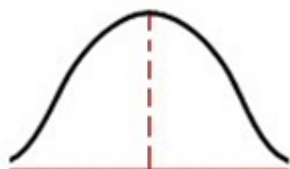
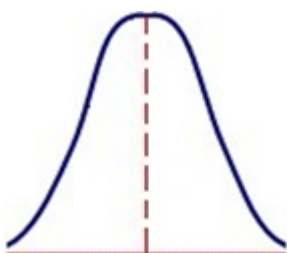
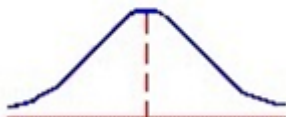
Enfoque clave:

Mide si los datos tienen colas más pesadas (más extremos) o más ligeras (menos extremos) que una distribución normal.

No confundir con asimetría: la curtosis se centra en las colas, no en el sesgo.

2.2) PROPIEDADES

Tabla: Propiedades de la Curtosis

Característica	Mesocúrtica	Leptocúrtica	Platicúrtica
Concentración de los valores en la región central de la distribución	Existe una concentración normal. Distribución normal.	Existe una gran concentración. Colas pesadas y pico alto (mayor concentración central).	Existe una baja concentración. Colas ligeras y pico bajo (dispersión uniforme).
Coefficiente de asimetría	= 0	> 0	< 0
Gráfico típico	 Mesocúrtica	 Leptocúrtica	 Platicúrtica
Ejemplo práctico	Altura de adultos en una población sana.	Riesgo de valores extremos (Ejemplo, finanzas).	Datos más "seguros" (Ejemplo, mediciones controladas).

2.3) COEFICIENTE DE CURTOSIS

A) FÓRMULA

La fórmula del coeficiente de curtosis basado en momentos, **Fisher-Pearson**, (g_2) es una medida de curtosis que se calcula utilizando los momentos estadísticos de la distribución, específicamente el cuarto momento estandarizado.

Los momentos estadísticos son medidas que describen la forma de una distribución.

El primer momento (media, $\bar{x}$) indica la tendencia central.

El segundo momento (varianza, s^2) mide la dispersión.

El tercer momento estandarizado (m_3), basado en la media y la desviación estándar, capta la asimetría.

El cuarto momento estandarizado (m_4), basado en la media y la desviación estándar, capta la curtosis.

La fórmula del coeficiente de asimetría basado en momentos (Fisher-Pearson) con ajuste para una muestra es:

$$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \frac{m_4}{s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \frac{\sum (x_i - \bar{x})^4}{s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Donde:

x_i = cada uno de los valores

n = tamaño de la muestra

$\bar{x}$ = media aritmética

s = desviación estándar de la muestra

B) TABLA DE INTERPRETACIÓN DEL COEFICIENTE DE CURTOSIS g_2

Valor de g_2	Tipo de Curtosis	Forma de la Distribución	Interpretación
$g_2 = 0$	Mesocúrtica	Similar a la distribución normal.	La distribución tiene colas y pico comparables a una normal (referencia).
$g_2 < 0$	Platicúrtica	Pico más bajo y colas más ligeras.	Los datos están menos concentrados alrededor de la media, con menor probabilidad de valores extremos.
$g_2 > 0$	Leptocúrtica	Pico más alto y colas más pesadas.	Mayor concentración de datos cerca de la media y presencia de valores atípicos (outliers).
$g_2 \ll 0$	Platicúrtica fuerte	Pico muy aplanado y colas extremadamente ligeras.	Distribución es mucho más plana que una normal con datos dispersos de manera uniforme, sin concentración cerca de la media y con ausencia de outliers.
$g_2 \gg 0$	Leptocúrtica Fuerte	Pico muy agudo y colas extremadamente pesadas.	Distribución con alta concentración central y outliers significativos.

Nota: $\ll$ se lee "mucho menor que" ; $\gg$ se lee "mucho mayor que"

2.4) ERROR ESTÁNDAR DE CURTOSIS

A) FÓRMULA

El **error estándar de curtosis** (o *standard error of kurtosis*, $SE_{curtosis}$) es una medida que cuantifica la variabilidad esperada en la estimación de la curtosis muestral debido al azar en el muestreo. Indica cuánto puede diferir la curtosis calculada en una muestra de la curtosis verdadera de la población.

La fórmula para calcular el error estándar de curtosis es:

$$SE_{curtosis} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$$

Donde:

n = Número total de observaciones válidas de la muestra

Para muestras grandes ($n > 150$), se aproxima a:

$$SE_{curtosis} = \sqrt{\frac{24}{n}}$$

B) REGLA PRÁCTICA:

Si el valor de curtosis dividido entre su error estándar cae fuera del rango ± 2 , se considera una desviación estadísticamente significativa de la normalidad.

$$\frac{g_2}{SE_{curtosis}} = \pm 2$$

Donde:

g_2 = coeficiente de curtosis

$SE_{curtosis}$ = error estándar de curtosis

2.5) EJEMPLOS DE APLICACIÓN

1) Calcular la curtosis de las siguientes calificaciones 6, 7, 7, 7, 8, 9, 9, 10, 10, 10

Solución:

Se calcula la media y la desviación estándar

$$\bar{x} = \frac{\sum x_i}{n} = \frac{6 + 7 + 7 + 7 + 8 + 9 + 9 + 10 + 10 + 10}{10} = 8,3$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{(6 - 8,3)^2 + (7 - 8,3)^2 + \dots + (10 - 8,3)^2}{10 - 1}} = 1,494$$

Se llena la siguiente tabla:

x_i	$\left(\frac{x_i - \bar{x}}{s}\right)^4$
6	5,611
7	0,573
7	0,573
7	0,573
8	0,002
9	0,048
9	0,048
10	1,675
10	1,675
10	1,675
Total	12,450

Se calcula el coeficiente de asimetría

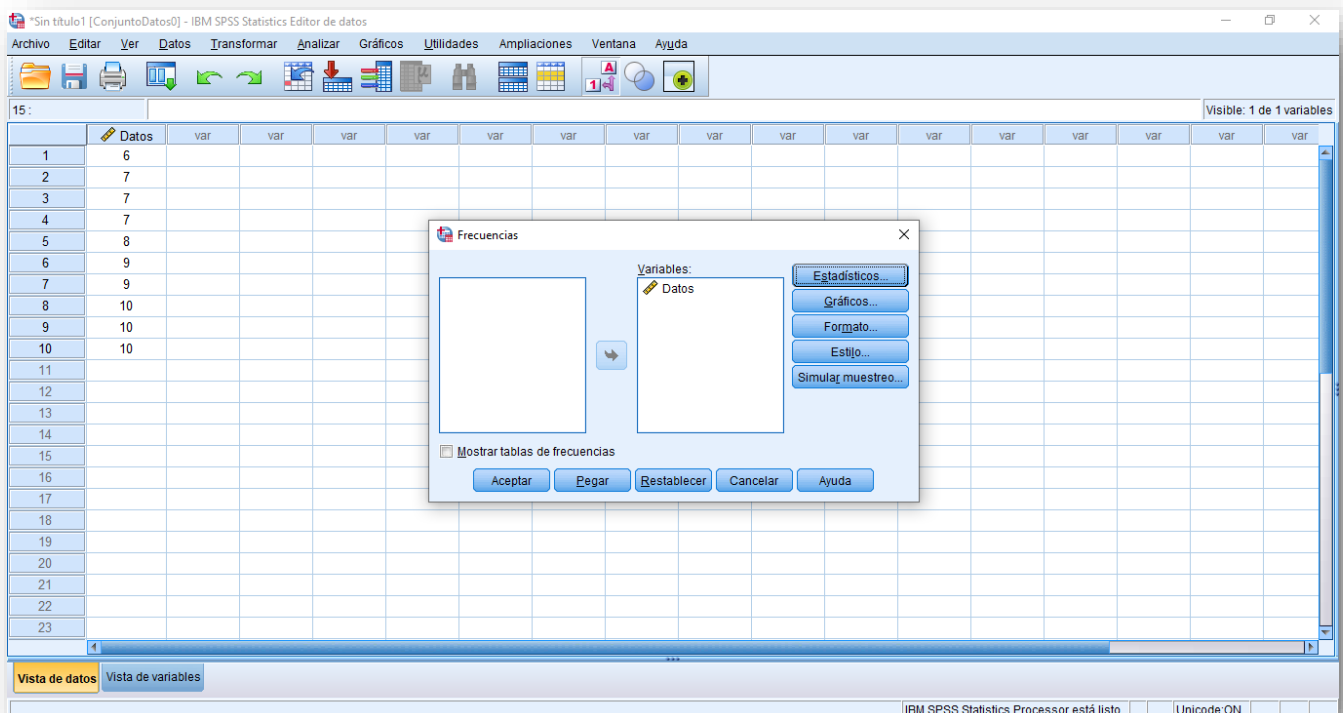
$$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$g_2 = \frac{10(10+1)}{(10-1)(10-2)(10-3)} \cdot 12,450 \cdot \frac{3(10-1)^2}{(10-2)(10-3)} = -1,62$$

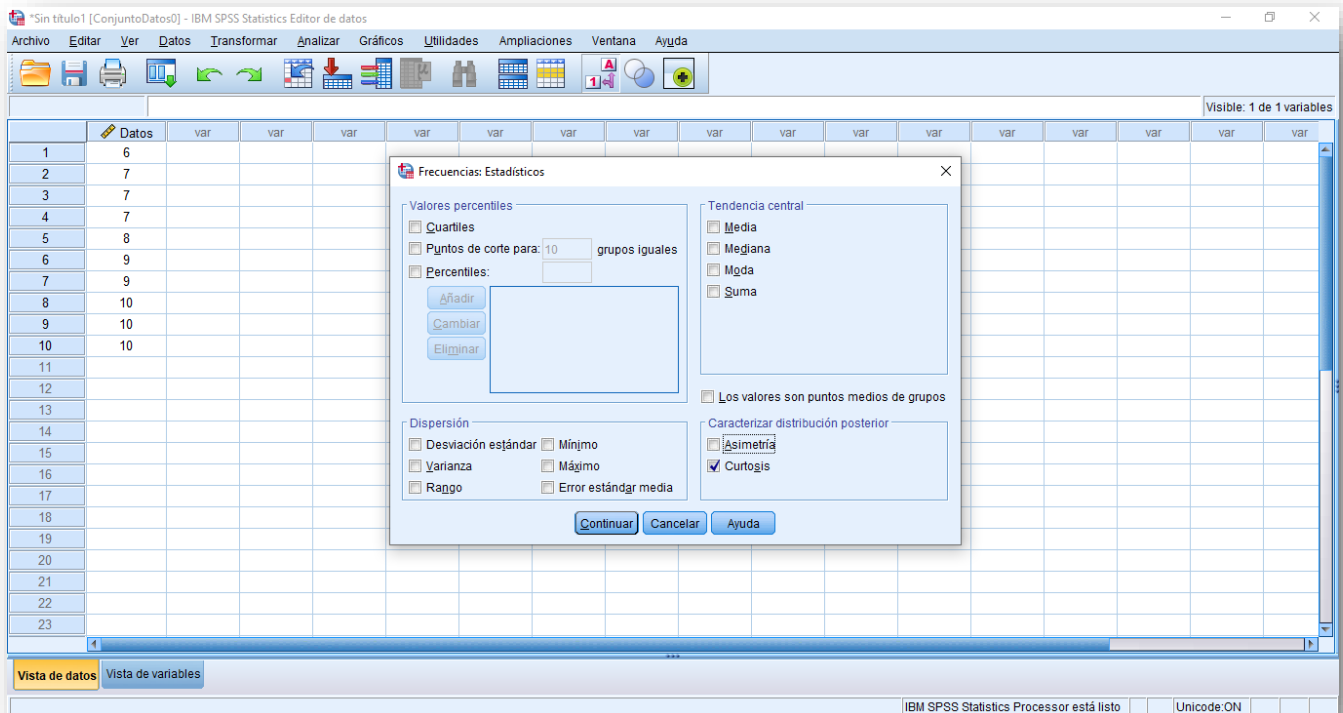
Empleando Excel

	A	B	C	D	E
1	6				
2	7				
3	7				
4	7				
5	8				
6	9				
7	9				
8	10				
9	10				
10	10				
11					
12	g_2	-1,62	=CURTOSIS(A1:A10)		
13					

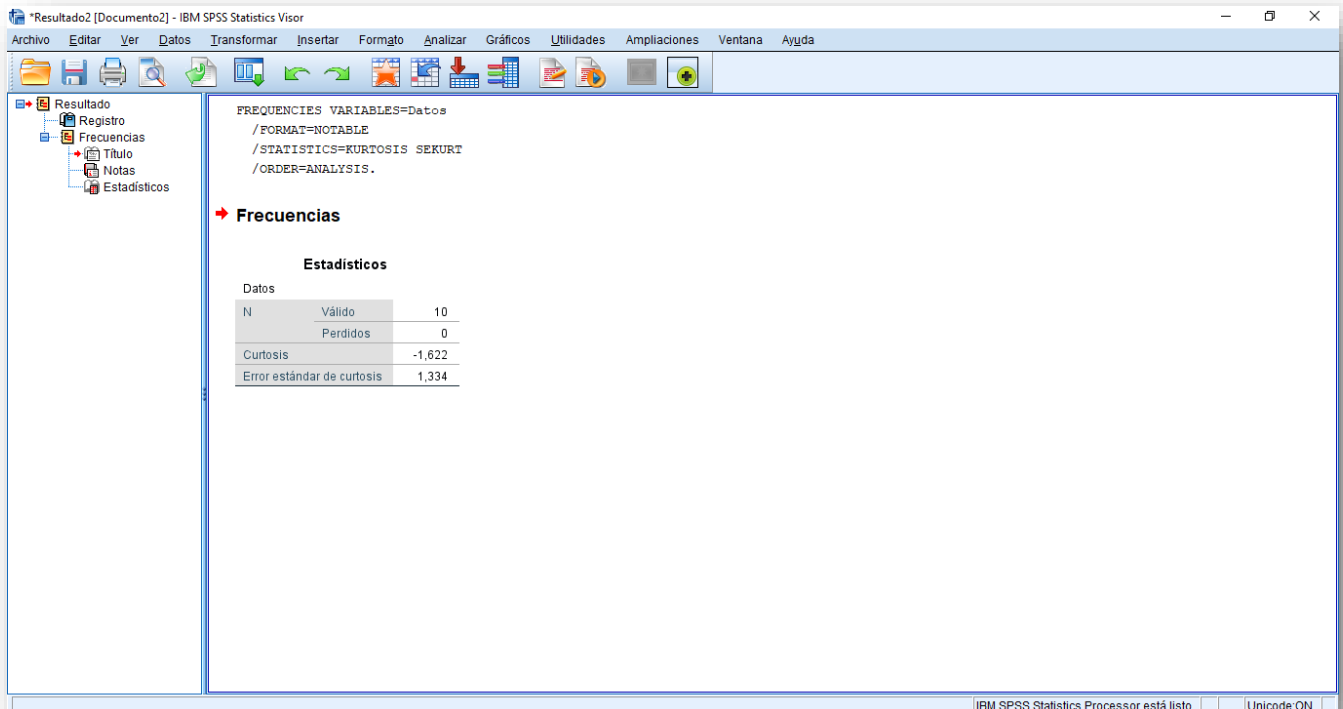
Empleando SPSS. Clic en Analizar. Estadísticos descriptivos. Frecuencias. Seleccionar la variable de los datos.



Clic en Estadísticos. Clic en Curtosis



Clic en Continuar. Clic en Aceptar



Clic derecho en la tabla. Copiar. Pegar.

Estadísticos

Datos

N	Válido	10
	Perdidos	0
Curtosis		-1,622
Error estándar de curtosis		1,334

Interpretación:

La curtosis mide el "pico" y la "cola" de la distribución de los datos en comparación con una distribución normal (curtosis = 0).

Como $g_2 = -1,62$ es < 0 , indica que la distribución es más platicúrtica (achatada) que la normal, es decir, menor concentración de datos alrededor de la media (pico más bajo) y colas menos pesadas (menos valores extremos) que en una distribución normal. Lo que implica que los datos están menos concentrados en el centro y tienen menos outliers de lo esperado en una distribución normal.

Reemplazando valores en la fórmula para calcular el error estándar de curtosis se obtiene:

$$SE_{curtosis} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}} = \sqrt{\frac{24 \cdot 10 \cdot (10-1)^2}{(10-3)(10-2)(10+3)(10+5)}} = 1,334$$

Coincide con el valor calculado con SPSS

Si el valor de curtosis es -1,622 con un SE = 1,334, el valor estandarizado empleando la regla práctica es

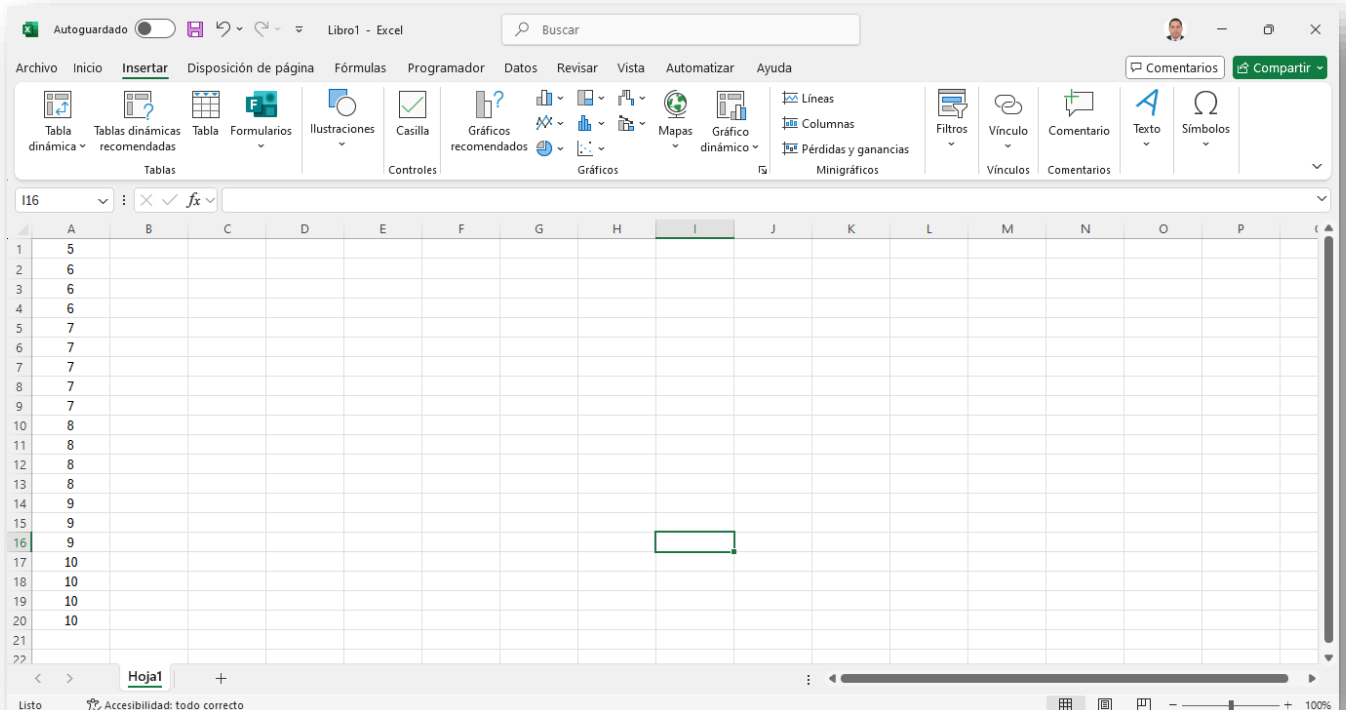
$$\frac{g_2}{SE_{curtosis}} = \pm 2$$

$$\frac{-1,622}{1,334} = -1,22$$

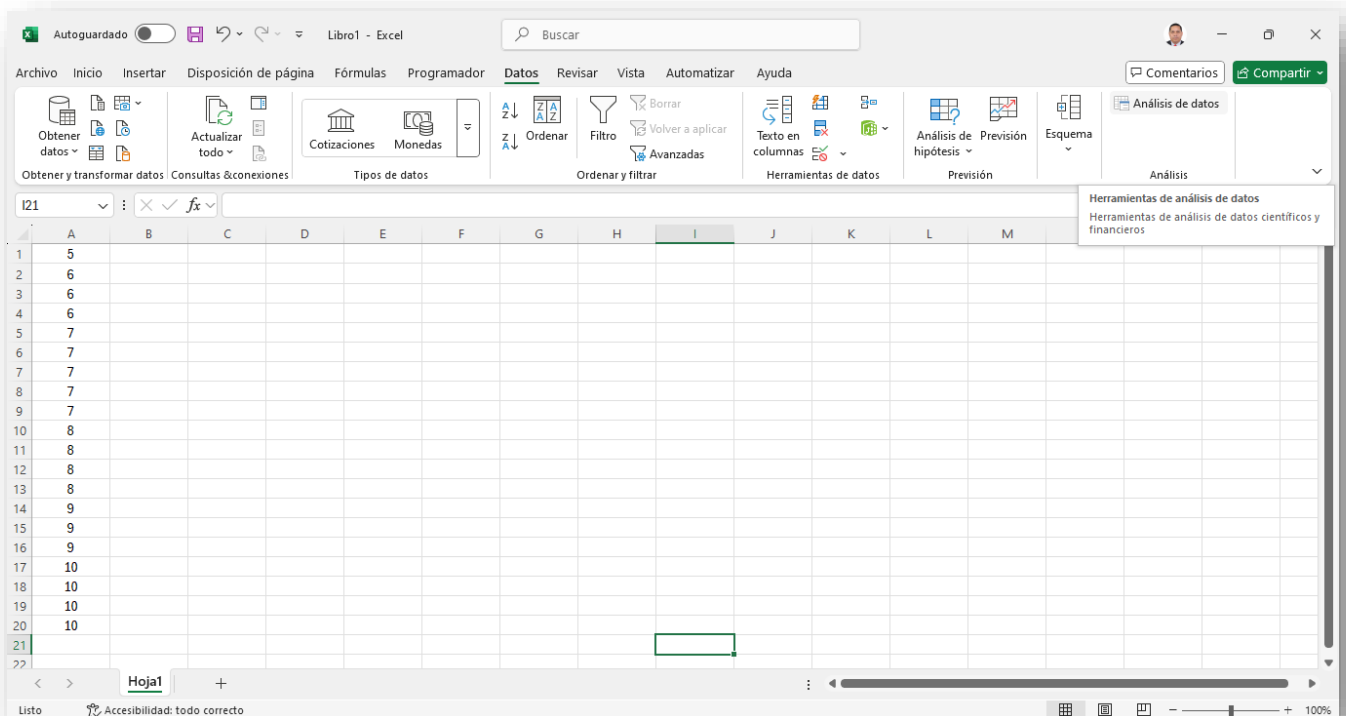
Indica que la desviación de la normalidad no es estadísticamente significativa (la distribución no es significativamente más plana que la normal), ya que -1,22 cae dentro del rango ± 2 , es decir, el valor de -1,22 sugiere que la distribución es ligeramente más plana (platicúrtica) que una normal, pero esta desviación podría deberse al azar (variabilidad muestral) en lugar de ser una característica real.

2) Empleando Excel, calcular la media, error típico, mediana, moda, desviación estándar, varianza de la muestra, curtosis, coeficiente de asimetría, rango, mínimo, máximo, suma y cuenta de las siguientes calificaciones 5, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 8, 8, 9, 9, 9, 10, 10, 10, 10

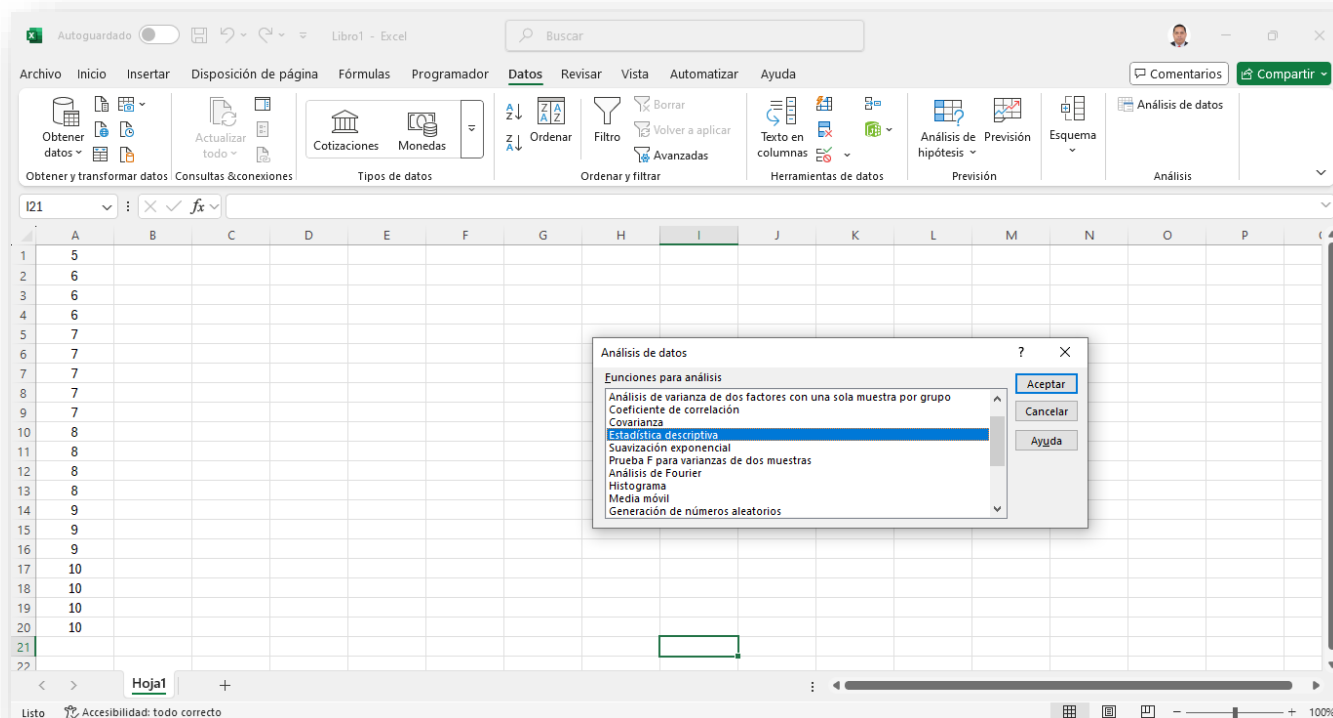
Solución: Ingresar los datos en Excel



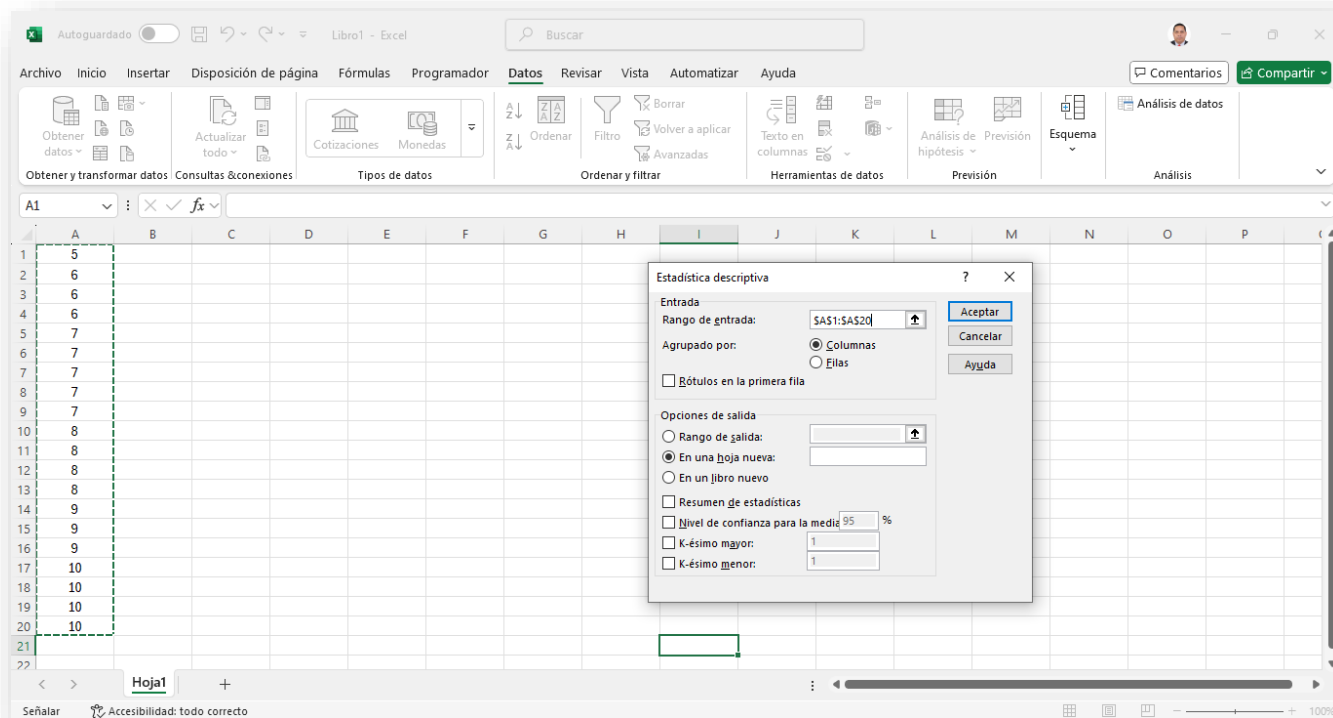
Clic en datos. Análisis de datos



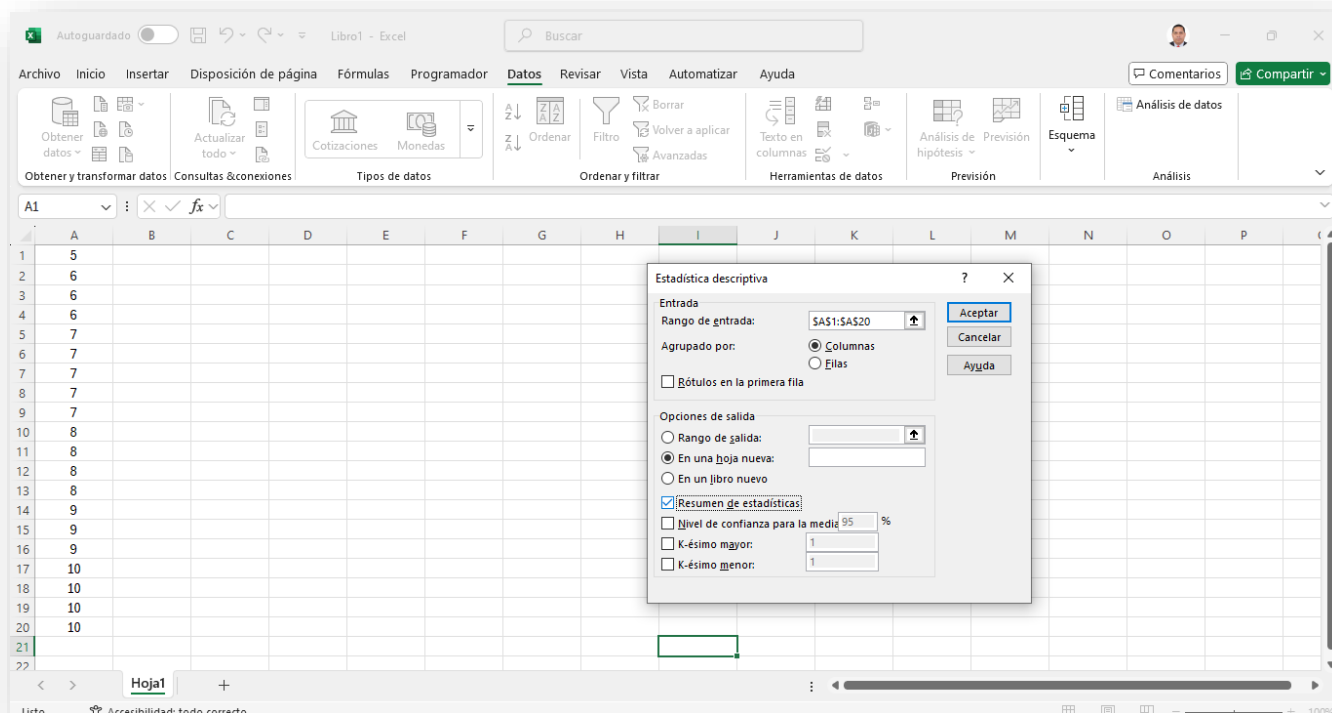
Clic en Analizar datos. Clic en Estadística Descriptiva



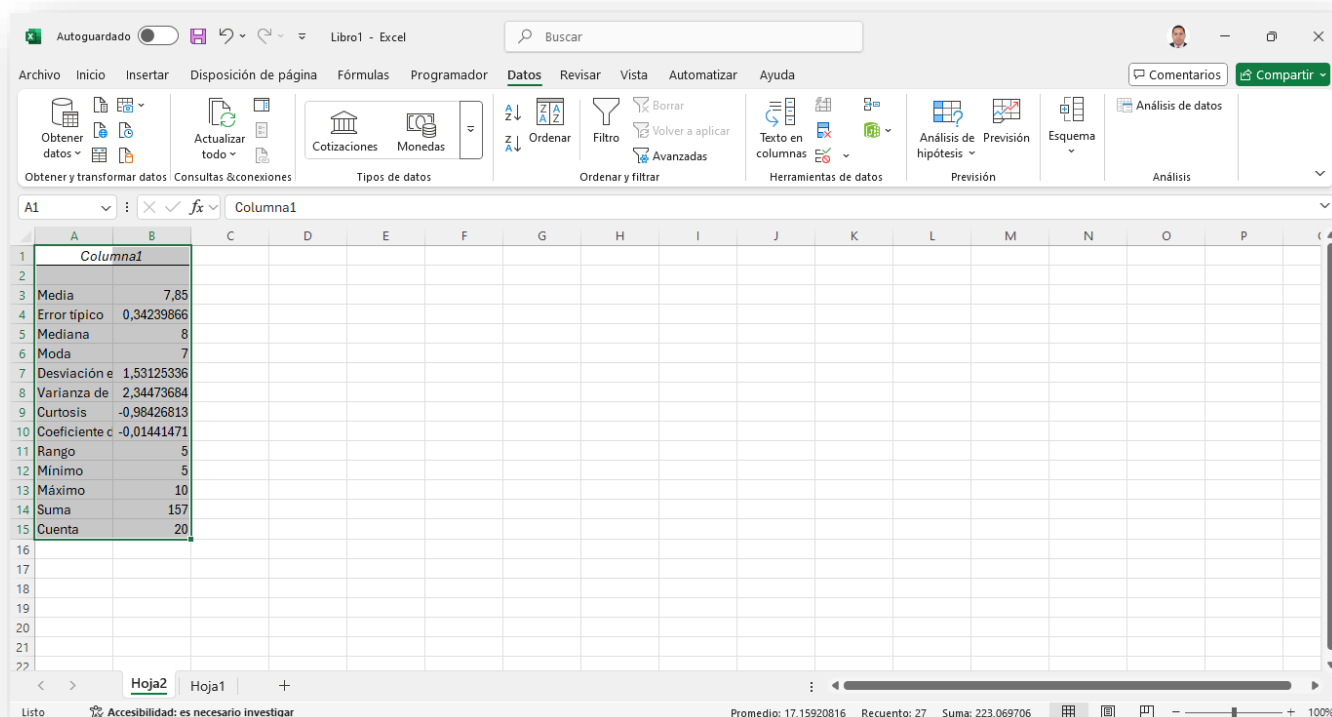
Clic en Aceptar. En Rango de entrada, seleccionar los datos.



Clic en Resúmenes de estadísticas



Clic en Aceptar



Copiar y pegar la tabla

<i>Columnal</i>	
Media	7,85
Error típico	0,34239866
Mediana	8
Moda	7
Desviación estándar	1,531253357
Varianza de la muestra	2,344736842
Curtosis	-0,984268127
Coeficiente de asimetría	-0,014414713
Rango	5
Mínimo	5
Máximo	10
Suma	157
Cuenta	20

Interpretación:

Tendencia Central

Media (7,85): El promedio de las calificaciones es 7,85, lo que sugiere un rendimiento académico relativamente bueno en general.

Mediana (8): El 50% de los estudiantes obtuvo una nota igual o inferior a 8, y el otro 50% igual o superior. Como la media (7,85) < mediana (8), hay una ligera asimetría negativa (confirmada más adelante por el coeficiente de asimetría cercano a 0).

Moda (7): La calificación más frecuente fue 7, lo que indica que varios estudiantes se agrupan en este valor.

Error Típico de la media

Mide cuánto podría variar la media muestral (7,85) si repitiéramos el estudio.

El intervalo de confianza al 95% es $\bar{x} \pm 1,96 \cdot SE \rightarrow 7,85 \pm 1,96 \cdot 0,342$

$7,85 - 1,96 \cdot 0,342 = 7,18$; $7,85 + 1,96 \cdot 0,342 = 8,52$

Esto significa que podemos estar 95% seguros de que si repitiéramos la evaluación con otros 20 alumnos la media estaría entre 7,18 y 8,52.

Valores Extremos y Resumen

Mínimo (5) y Máximo (10): La calificación más baja fue 5 y la más alta 10 (ningún estudiante estuvo por debajo de 5 o arriba de 10).

Suma (157) y Cuenta (20): Corroborar que la media es correcta: $157/20 = 7,85$

Dispersión de los Datos

Desviación estándar (1,53): Las calificaciones se desvían en promedio 1,53 puntos de la media (7,85).

Interpretación: Si la distribución es normal (68% de los datos estarían entre 6,32 y 9,38 ($7,85 \pm 1,53$)).

Varianza (2,34): Es el cuadrado de la desviación estándar

Rango (5): La diferencia entre la nota máxima (10) y mínima (5) es de 5 puntos, lo que indica una variabilidad moderada.

Forma de la Distribución

Asimetría (-0,014): El valor cercano a 0 indica que la distribución es casi simétrica. El ligero signo negativo sugiere una cola izquierda muy leve (pero insignificante en la práctica).

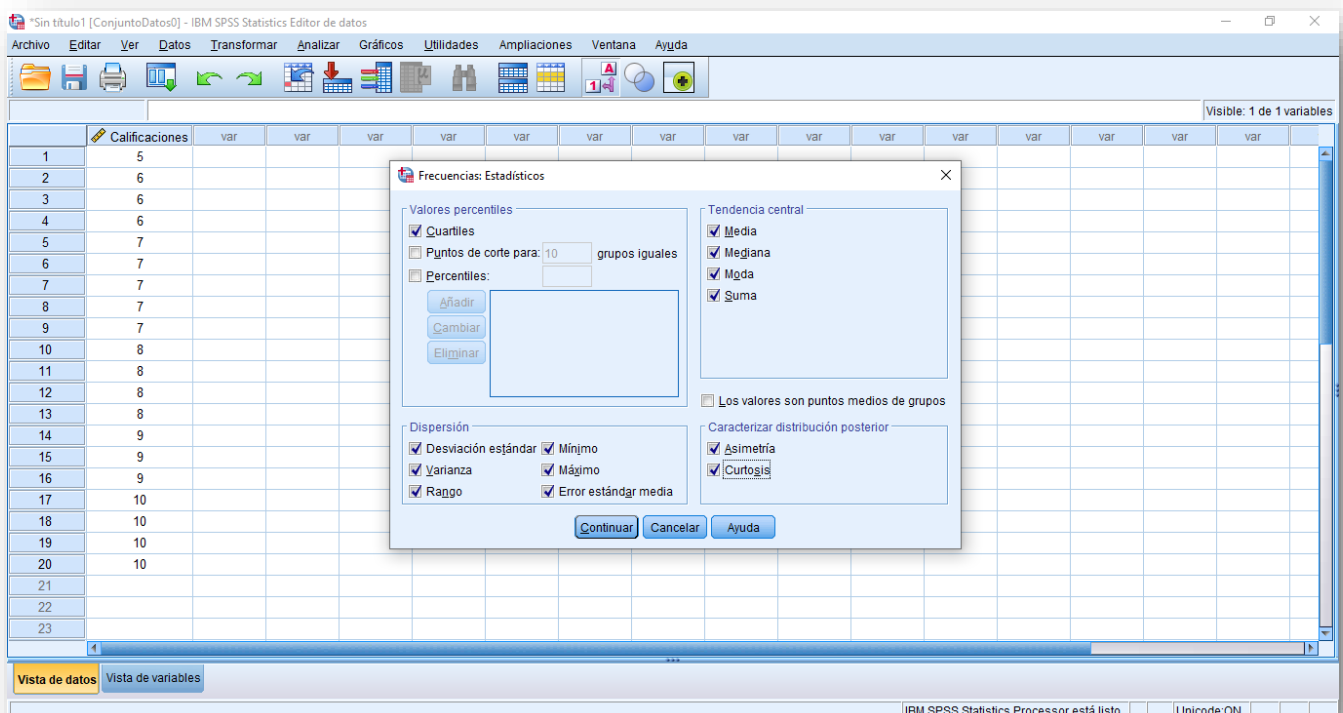
Curtosis (-0,984): El valor negativo indica una distribución más plana (platicúrtica) que la normal.

Implicación: Hay menos datos extremos (outliers) y una menor concentración alrededor de la media.

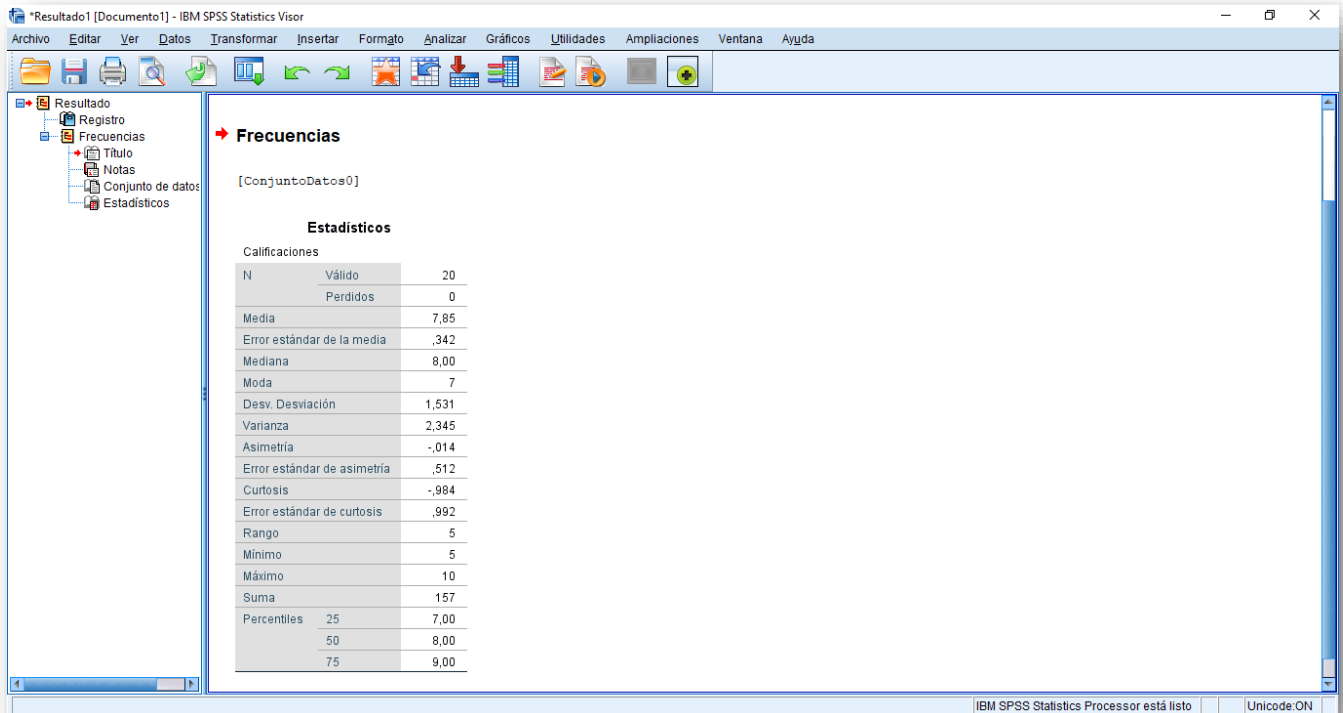
3) Resolver el ejercicio anterior empleando SPSS. También calcular los cuartiles.

Solución:

Ingresa los datos en SPSS. Clic en Analizar. Estadísticos descriptivos. Frecuencias. Clic en los estadísticos a calcular.



Clic en continuar. Clic en Aceptar



Clic derecho en la tabla. Copiar. Pegar

Estadísticos
Calificaciones

N	Válido		20
	Perdidos		0
Media			7,85
Error estándar de la media			,342
Mediana			8,00
Moda			7
Desv. Desviación			1,531
Varianza			2,345
Asimetría			-,014
Error estándar de asimetría			,512
Curtosis			-,984
Error estándar de curtosis			,992
Rango			5
Mínimo			5
Máximo			10
Suma			157
Percentiles	25		7,00
	50		8,00
	75		9,00

Interpretación:

1. Datos Básicos

N (Válido): Hay 20 observaciones válidas (sin datos perdidos).

Suma: La suma total de todas las calificaciones es 157.

2. Medidas de Tendencia Central

Media: El promedio de las calificaciones es 7,85.

Mediana (Percentil 50): El 50% de las calificaciones son ≤ 8 (valor central de la distribución).

Moda: La calificación más frecuente es 7 (aparece 5 veces).

3. Medidas de Dispersión

Desviación Estándar: Las calificaciones se desvían, en promedio, 1,531 puntos de la media (7,85).

Varianza: Es el cuadrado de la desviación estándar ($1,531^2 = 2,345$). Indica la dispersión de los datos.

Rango: 5 (Mínimo = 5, Máximo = 10). La diferencia entre la calificación más alta (10) y la más baja (5) es de 5 puntos.

4. Posición (Cuartiles)

Percentil 25 (Q_1): El 25% de las calificaciones son $\leq 7,00$

Percentil 50 (Q_2 = Mediana): 8,00

Percentil 75 (Q_3): El 75% de las calificaciones son $\leq 9,00$

5. Forma de la Distribución

Asimetría (Skewness): -0,014. El valor cercano a 0 sugiere una distribución simétrica (ligeramente inclinada a la izquierda, pero casi neutra).

Error estándar de asimetría: 0,512 $\rightarrow$ El valor (-0,014) está dentro de ± 1 error estándar, lo que confirma que la asimetría no es significativa.

Curtosis: -0,984. Valor negativo indica una distribución **platicúrtica** (menos apuntada y con colas más ligeras que una normal).

Error estándar de curtosis: 0,992 $\rightarrow$ La curtosis es significativamente diferente de una distribución normal (ya que $-0,984/0,992 \approx -1$, supera ± 1).

6. Error Estándar de la Media: Indica la precisión de la media calculada. Un valor bajo (0,342) sugiere que la media (7,85) es una buena estimación de la población.

Conclusión:

El promedio de calificaciones fue 7,85, con la mitad de los estudiantes obteniendo 8 o más. La distribución es equilibrada (simétrica), aunque con una frecuencia notable de calificaciones bajas (5-6) y altas (10), lo que explica su forma aplanada (curtosis negativa).

2.6) ACTIVIDADES DE APLICACIÓN

- 1) Realice un organizador gráfico del capítulo II
- 2) Empleando Excel calcule la media, error típico, mediana, moda, desviación estándar, varianza de la muestra, curtosis, coeficiente de asimetría, rango, mínimo, máximo, suma y cuenta de las siguientes calificaciones 4, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8, 8, 8, 8, 9, 9, 10, 10, 10, 10. Interprete los resultados.

	A	B
1	Columna1	
2		
3	Media	7,55
4	Error típico	0,407010927
5	Mediana	8
6	Moda	8
7	Desviación estándar	1,820208201
8	Varianza de la muestra	3,313157895
9	Curtosis	-0,8137418
10	Coeficiente de asimetría	-0,234620694
11	Rango	6
12	Mínimo	4
13	Máximo	10
14	Suma	151
15	Cuenta	20

- 3) Con los datos del ejercicio anterior calcule la regla práctica de la curtosis. Interprete los resultados.
- 5) Descargue la base de datos de la Prueba Ser Estudiante 2023-2024 (Micro SAV) disponible en el portal oficial: <https://evaluaciones.evaluacion.gob.ec/BI/ser-estudiante-2/>

liante-2/

☆ 📄 📁 🗑️

[↓ Descargar Manual de Fichas Metodológicas 2019](#)

Año lectivo 2023-2024

Archivo	Sintaxis	CSV	SAV	XLSX	Metadato	Diccionario
Micro	↓	↓	↓	↓	↓	↓
Factores Asociados estudiantes	↓	↓	↓	↓	↓	↓
Factores Asociados padres	↓	↓	↓	↓	↓	↓
Factores Asociados docentes	↓	↓	↓	↓	↓	↓
Factores Asociados directivos	↓	↓	↓	↓	↓	↓

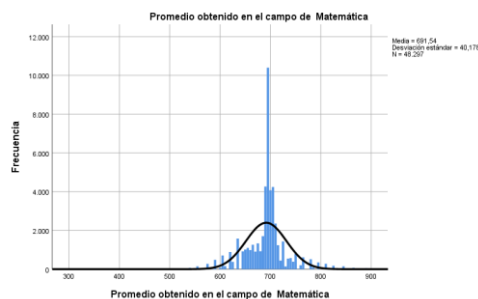
a) Empleando SPSS calcule la media, error estándar de la media, error estándar de la media, mediana, moda, desviación estándar, varianza, rango, mínimo, máximo, suma, cuartiles, asimetría, error estándar de asimetría, curtosis y error estándar de curtosis, del promedio obtenido en el campo de Matemática (imat) y del promedio obtenido en el campo de Lengua y Literatura (ilyl). Copie y pegue la tabla. Interprete los resultados.

Estadísticos

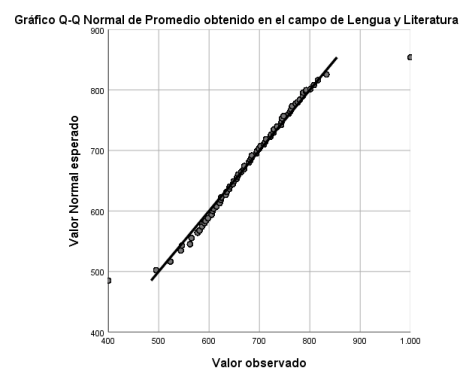
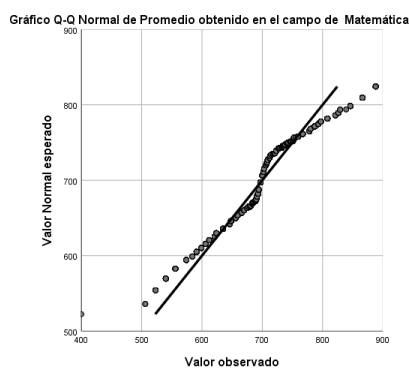
		Promedio obtenido en el campo de Matemática	Promedio obtenido en el campo de Lengua y Literatura
N	Válido	48297	48330
	Perdidos	2248	2215
Media		691,54	687,95
Error estándar de la media		,183	,271
Mediana		694,00	684,00
Moda		697	670
Desv. Desviación		40,178	59,635
Varianza		1614,247	3556,356
Asimetría		,145	,702
Error estándar de asimetría		,011	,011
Curtosis		3,139	3,445
Error estándar de curtosis		,022	,022
Rango		488	600
Mínimo		400	400
Máximo		888	1000
Suma		33399191	33248716
Percentiles	25	675,00	649,00
	50	694,00	684,00
	75	706,00	728,00

b) Calcule la regala práctica de la curtosis para imat y para ilyl. Interprete los resultados.

c) Empleando SPSS elabore Histogramas con Curva Normal para imat y para ilyl. Interprete los resultados.



d) Empleando SPSS elabore gráficos Q-Q Plot para imat y para ilyl. Interprete los resultados.



6) Cree y resuelva un ejercicio similar al anterior empleando la misma base de datos.

CAPÍTULO III CORRELACIÓN

3.1) DEFINICIÓN

La correlación es una medida estadística que describe la relación lineal entre dos variables cuantitativas. Indica la dirección (positiva o negativa), la fuerza (débil, moderada o fuerte) y el grado de asociación entre ellas, pero no implica causalidad.

Valor	Significado de la correlación
-1	Negativa perfecta
-0,7 a -0,99	Negativa fuerte
-0,4 a -0,69	Negativa moderada
-0,01 a -0,39	Negativa débil
0	Nula
0,01 a 0,39	Positiva débil
0,4 a 0,69	Positiva moderada
0,7 a 0,99	Positiva fuerte
1	Positiva perfecta

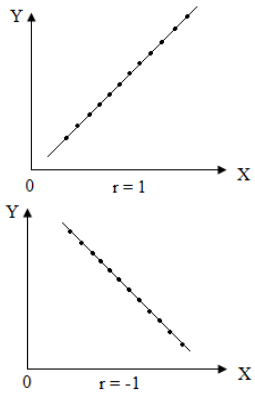
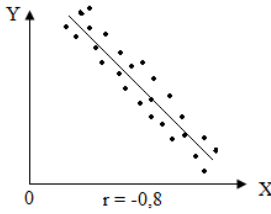
Correlación $\neq$ Causalidad: La correlación no implica causalidad, siempre considerar variables ocultas o factores externos. Una correlación fuerte no significa que una variable cause la otra.

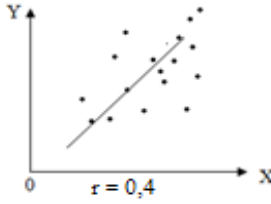
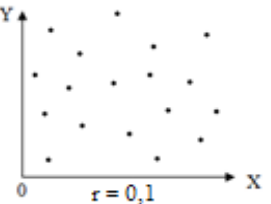
Además, Pearson fue el primero en alertar sobre las “correlaciones espurias”, donde explica que dos variables pueden estar fuertemente correlacionadas entre sí sin que entre ambas exista una relación causa-efecto o ni si siquiera una causa común (Ejemplo: relación entre el consumo de helados y el número de ahogamientos en piscinas). La correlación no implica causalidad. En este caso, una tercera variable oculta (la temperatura ambiental) explica la relación:

1. En verano hace más calor $\rightarrow$ la gente consume más helados.
2. En verano más personas van a piscinas $\rightarrow$ aumentan los ahogamientos.

Por lo tanto, aunque helados y ahogamientos parecen relacionados, en realidad no hay una conexión directa entre ellos. Es una correlación espuria, ya que la temperatura actúa como factor de confusión.

3.2) CLASIFICACIÓN

Criterio de Clasificación	Tipo de Correlación	Descripción	Ejemplo	Coefficiente Usual
Dirección	Positiva	Ambas variables aumentan o disminuyen juntas.	Horas de estudio vs. Calificaciones	Pearson ($r > 0$)
	Negativa	Una variable aumenta mientras la otra disminuye.	Horas de TV vs. Rendimiento académico	Pearson ($r < 0$)
	Nula (o neutra)	No hay relación lineal discernible.	Altura vs. Calificaciones	Pearson ($r = 0$)
Fuerza	Perfecta ($r = \pm 1$)	Relación lineal exacta (todos los puntos caen en una línea recta). 	Peso en kg vs. Peso en libras (si no hay error)	Pearson ($r = 1$ o $r = -1$)
	Fuerte ($r = \pm 0,7$ a $\pm 0,99$)	Relación lineal evidente, pero con dispersión. 	Ingresos vs. Años de educación	Pearson o Spearman
	Moderada ($r = \pm 0,4$ a $\pm 0,69$)	Relación observable, pero con mayor dispersión.	Edad vs. Presión arterial	

Criterio de Clasificación	Tipo de Correlación	Descripción	Ejemplo	Coefficiente Usual
				
	Débil ($r = \pm 0,1$ a $\pm 0,39$)	Relación mínima, casi imperceptible. 	Talla de zapato vs. Inteligencia	
Forma de la Relación	Lineal	La relación se ajusta a una línea recta.	Temperatura vs. Consumo de energía	Pearson (r)
	No lineal (monótona)	La relación sigue una tendencia (creciente o decreciente), pero no es recta.	Edad vs. Preferencia por tecnología	Spearman (ρ) o Kendall (τ)
	No monótona	La relación tiene patrones complejos (Ejemplo: en forma de U).	Ansiedad vs. Rendimiento (ley de Yerkes-Dodson)	No aplica
Tipo de Datos	Entre variables cuantitativas	Ambas variables son numéricas y continuas.	Peso vs. Altura	Pearson (r)
	Entre variables ordinales y no cumple el supuesto de normalidad	Variables con categorías ordenables (Ejemplo: escalas Likert).	Nivel socioeconómico vs. Satisfacción	Spearman (ρ) o Kendall (τ)
	Entre una cuantitativa y una dicotómica	Una variable es numérica y la otra binaria (Ejemplo. sí/no).	Género (0/1) vs. Salario	Punto-biserial

3.3) COEFICIENTE DE CORRELACIÓN DE PEARSON

A) DEFINICIÓN

El coeficiente de correlación de Pearson (r) mide la relación lineal entre dos variables cuantitativas, con valores entre -1 (correlación negativa perfecta) y 1 (correlación positiva perfecta).

Ejemplos:

Estudio de la relación entre el tiempo de ejercicio semanal (horas) y el nivel de colesterol HDL (mg/dL) en pacientes. Resultado posible: Si $r = -0,75$, indica una fuerte correlación negativa: a más ejercicio, menor colesterol.

Estudio de la relación entre las horas de estudio semanales y las calificaciones finales en un grupo de estudiantes. Resultado posible: Si $r = 0,60$, existe una correlación positiva moderada.

Estudio de la relación entre la temperatura ambiental ($^{\circ}\text{C}$) y el consumo de energía eléctrica (kW/h) en una ciudad. Resultado posible: Si $r = 0,80$ en verano, sugiere que a mayor temperatura, mayor uso de aire acondicionado.

Consideraciones importantes:

Correlación $\neq$ Causación: Una correlación alta no implica que una variable cause la otra (puede haber factores externos), es decir, dos variables pueden estar linealmente relacionadas, sin embargo, esto no quiere decir que una variable sea causa de la otra.

Linealidad: Pearson solo capta relaciones lineales. Para patrones no lineales, se usan otros métodos (Ejemplo: Spearman).

Outliers: Datos extremos pueden distorsionar r .

B) FÓRMULA

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

Donde:

n = número de observaciones; X, Y = valores de las variables

C) SUPUESTOS PARA APLICAR

El coeficiente de correlación de Pearson (r) mide la relación lineal entre dos variables cuantitativas. Para que sus resultados sean válidos, deben cumplirse los siguientes supuestos:

1) Relación Lineal

La asociación entre las variables debe ser lineal (una línea recta puede describirla).

¿Cómo verificarlo?:

Con gráfico de dispersión (scatter plot): Los puntos deben seguir una tendencia recta (no curva). Si la relación es no lineal (Ejemplo: exponencial), se debe emplear Spearman.

2) Normalidad

Ambas variables deben distribuirse normalmente en la población.

¿Cómo verificarlo?:

Con pruebas de normalidad: Shapiro-Wilk, Kolmogorov-Smirnov

Prueba	Uso Recomendado	Hipótesis	p-valor
Shapiro-Wilk	$n \leq 50$	Hipótesis nula (H_0): Los datos provienen de una distribución normal	$p < 0,05$ Rechazamos H_0
Kolmogorov-Smirnov	$n > 50$	Hipótesis alternativa (H_1): Los datos no provienen de una distribución normal	$p \geq 0,05$ Aceptamos H_0 No existe evidencia significativa para rechazar H_0

D) INTERPRETACIÓN DEL P-VALOR

El p-valor en el coeficiente de Pearson indica si la relación lineal entre dos variables es estadísticamente significativa o podría deberse al azar. Se contrastan las hipótesis:

Hipótesis nula (H_0): No hay correlación lineal ($r = 0$)

Hipótesis alternativa (H_1): Existe correlación lineal ($r \neq 0$)

Regla de decisión clave

p-valor	Interpretación	Decisión
$p < 0,05$	Correlación significativa: Rechazamos H_0 . Hay evidencia de relación lineal.	Ejemplo: $p = 0,01 \rightarrow$ Significativa.
$p \geq 0,05$	No significativa: No hay suficiente evidencia para rechazar H_0	Ejemplo: $p = 0,20 \rightarrow$ No significativa.

E) EJEMPLOS DE APLICACIÓN

1) Calcular el coeficiente de correlación de Pearson sobre la relación entre horas de estudio y calificaciones empleando los siguientes datos:

Estudiante	Horas de estudio (X)	Calificación (Y)
1	2	6,0
2	4	7,5
3	6	9,1
4	8	9,2
5	6	8,5
6	4	7,5
7	6	9,3
8	7	9,4
9	10	9,6
10	8	9,5

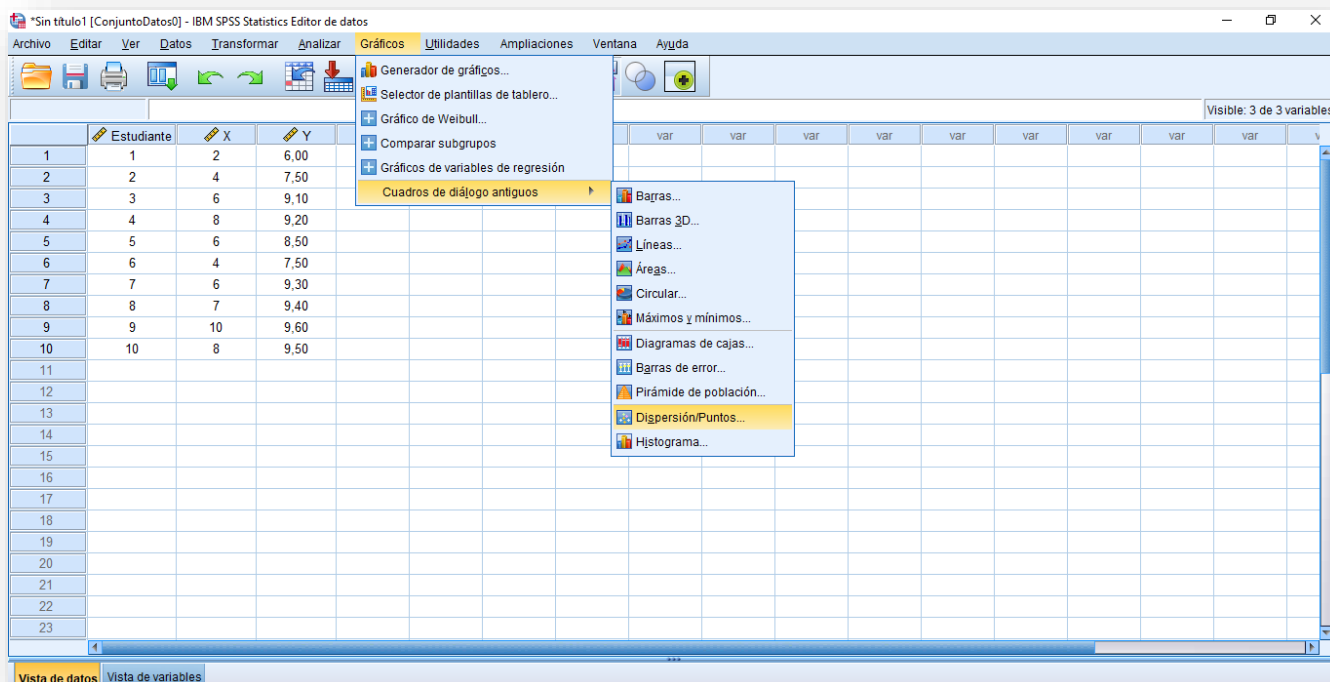
Solución:

A continuación, se comprueban los supuestos necesarios para aplicar el coeficiente de correlación de Pearson a los datos de horas de estudio (X) y calificaciones (Y).

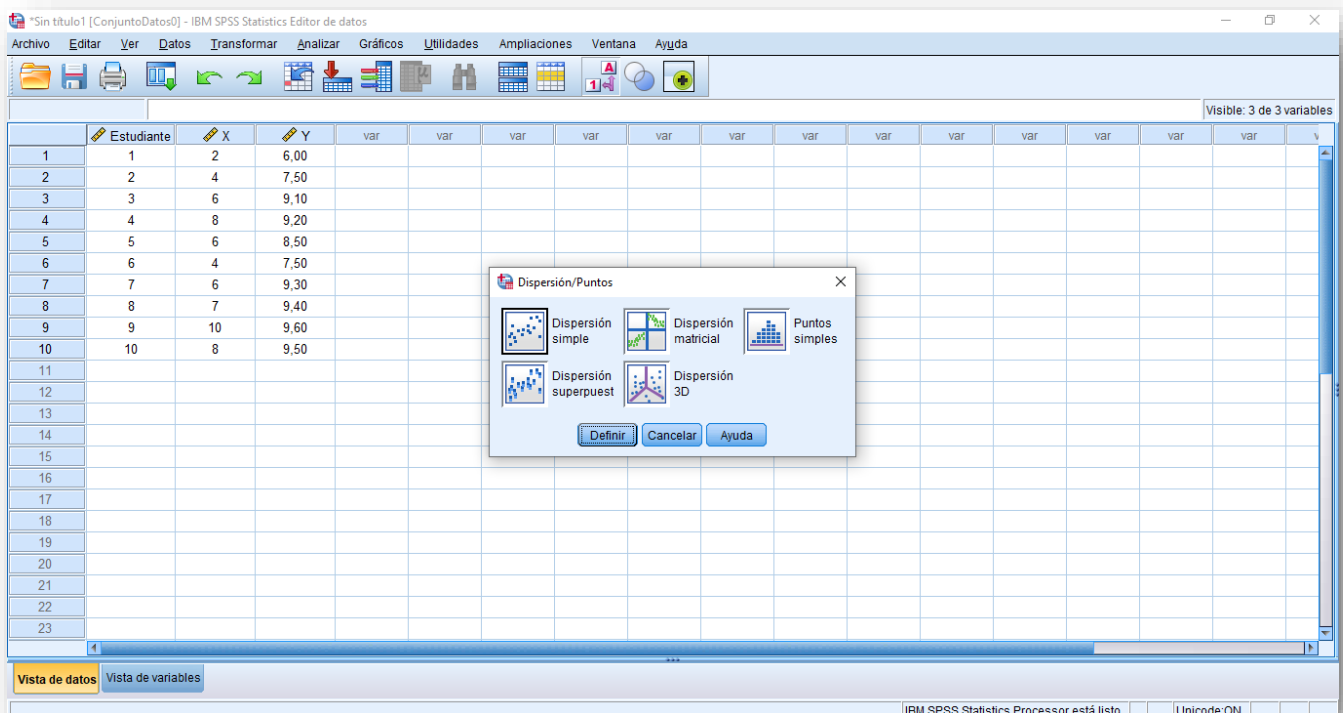
a) Relación Lineal

Método: Gráfico de dispersión (*Scatter Plot*).

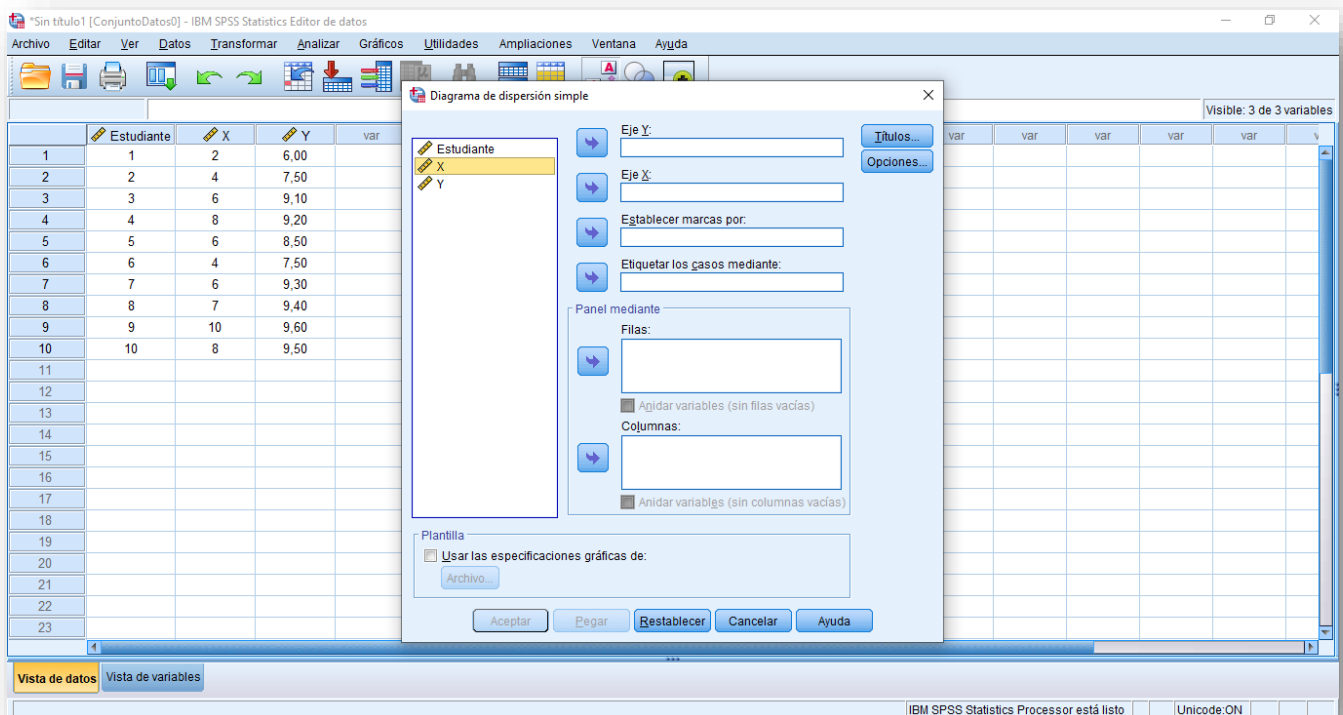
Para elaborar un diagrama de dispersión en SPSS. Clic en Gráficos. Cuadros de diálogo antiguos. DispersiónPuntos.



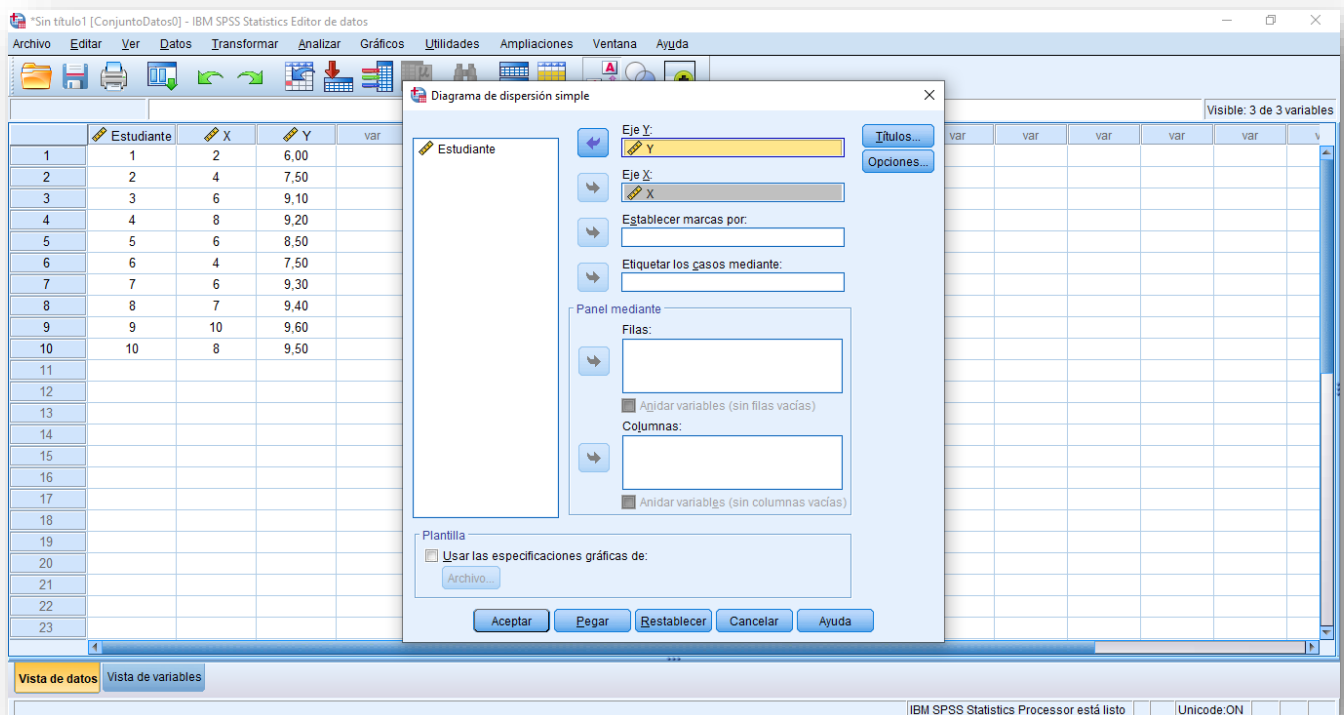
Clic en DispersiónPuntos.



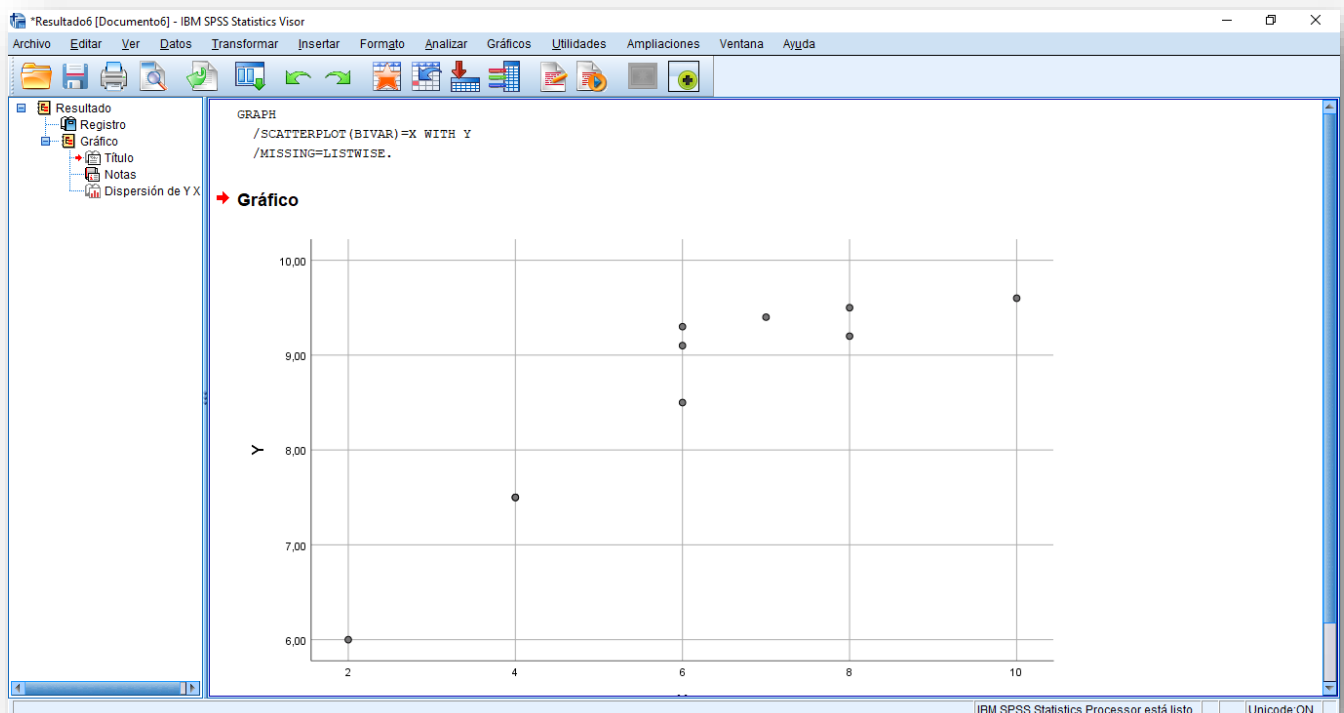
Clic en Dispersión simple. Clic en Definir



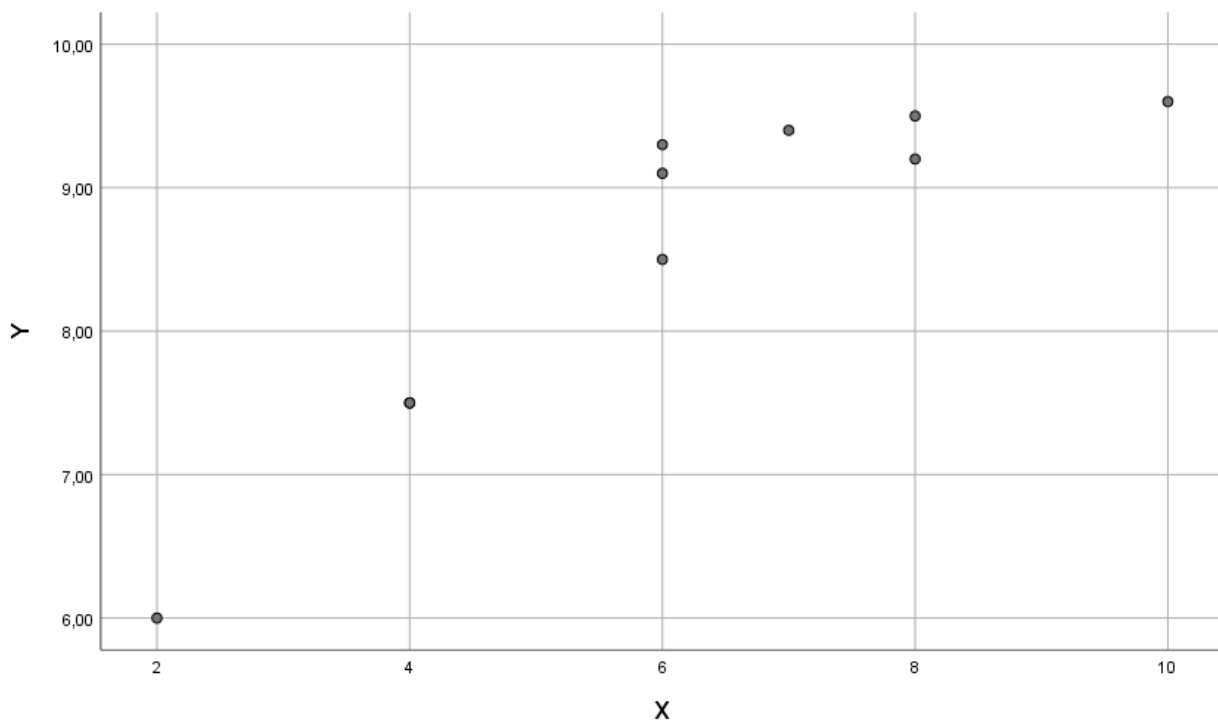
Seleccionar las variables X, Y. Clic en la fecha



Clic en Aceptar

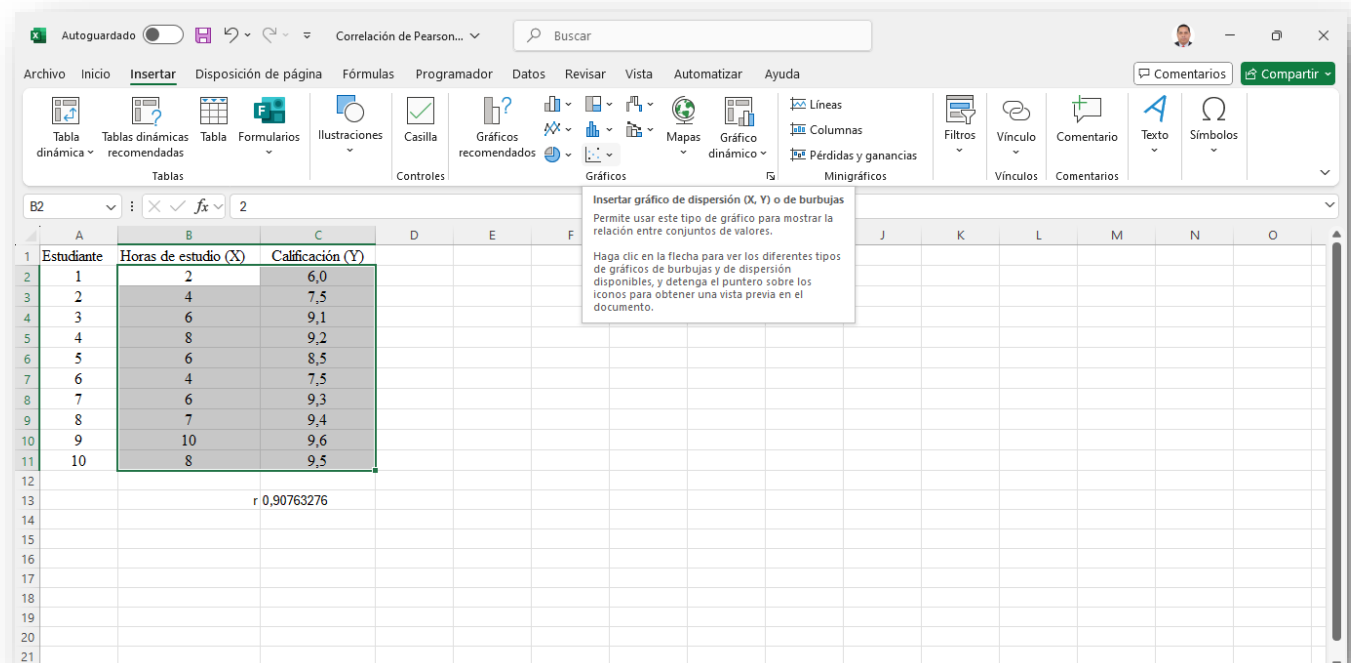


Clic derecho en el gráfico. Copiar. Pegar

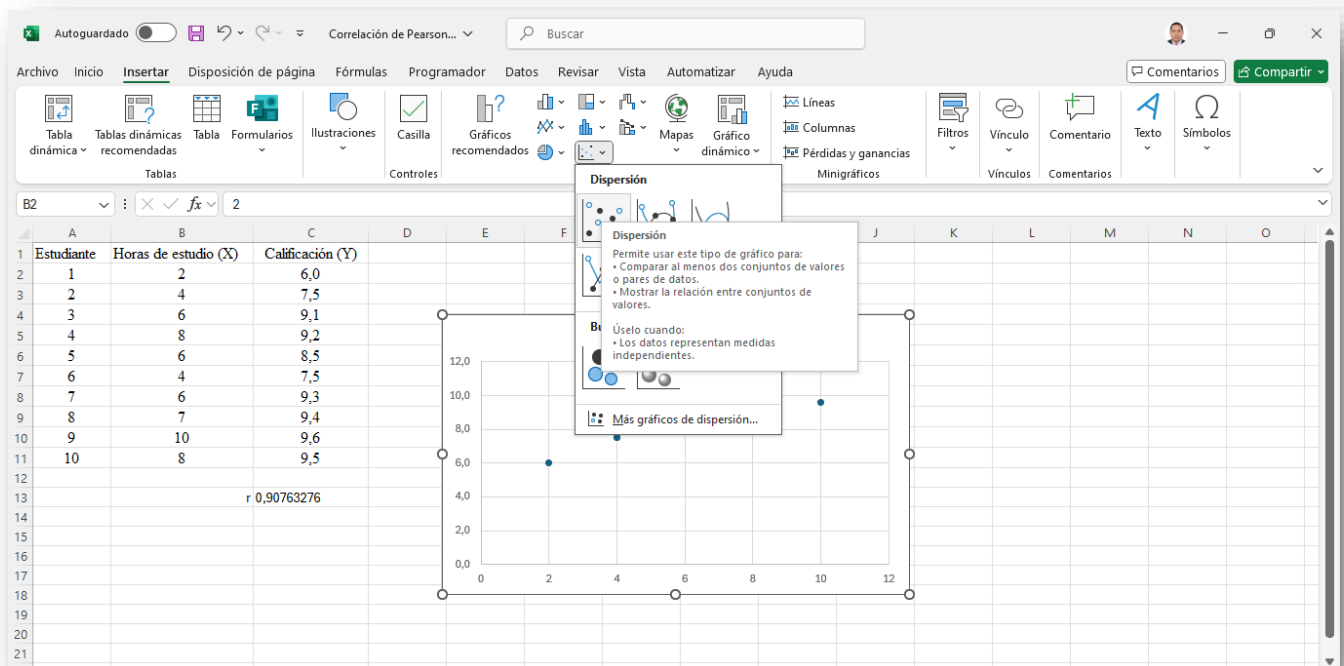


Empleando Excel

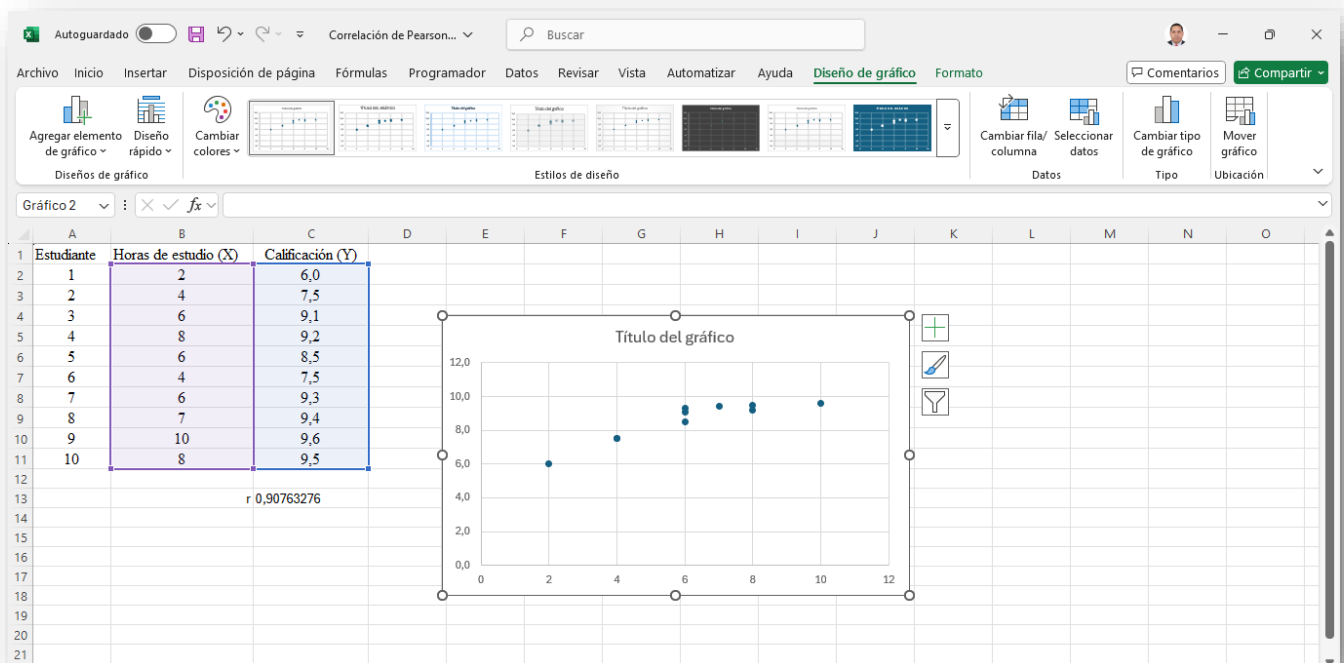
Seleccionar los datos de X, Y. Clic en Insertar. Gráficos



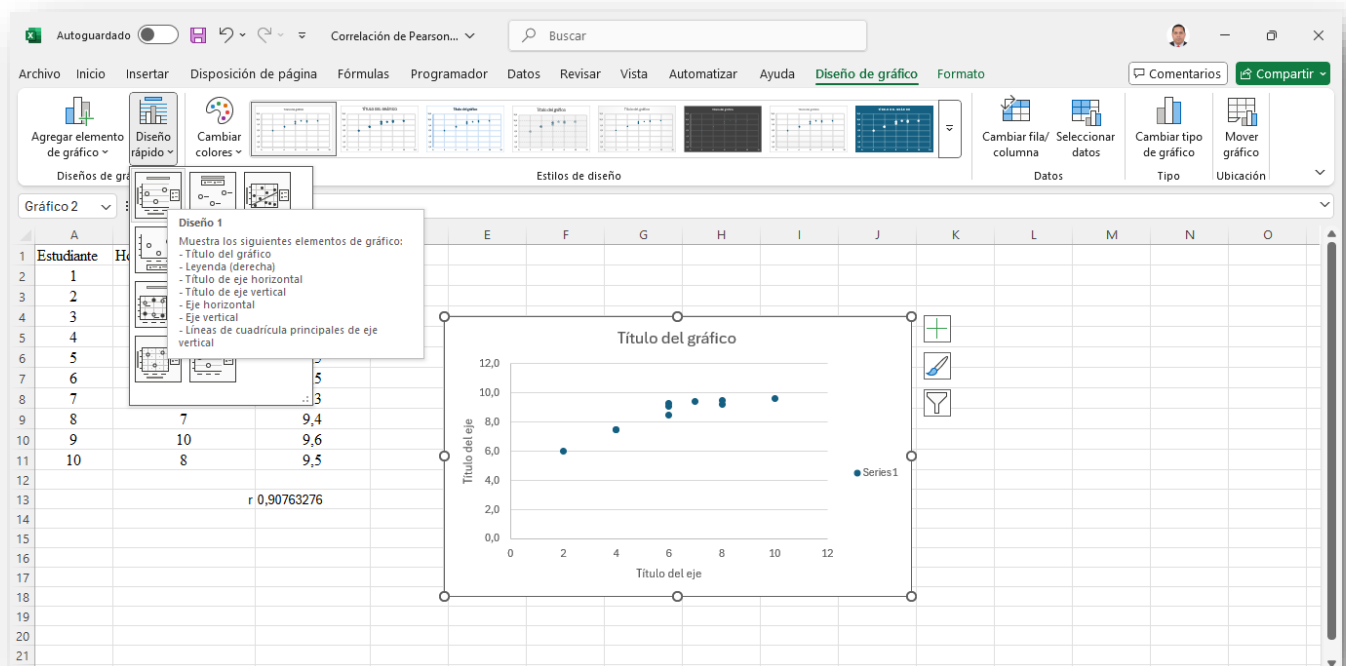
Clic en Dispersión



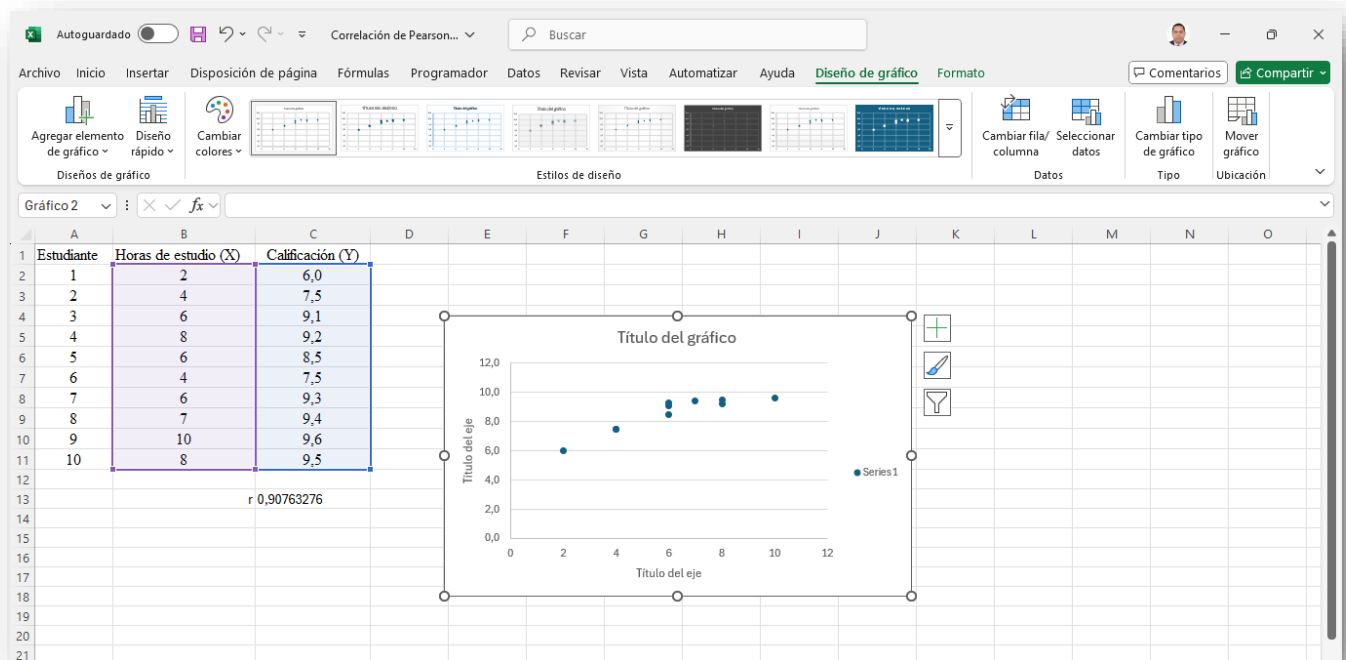
Clic en Dispersión



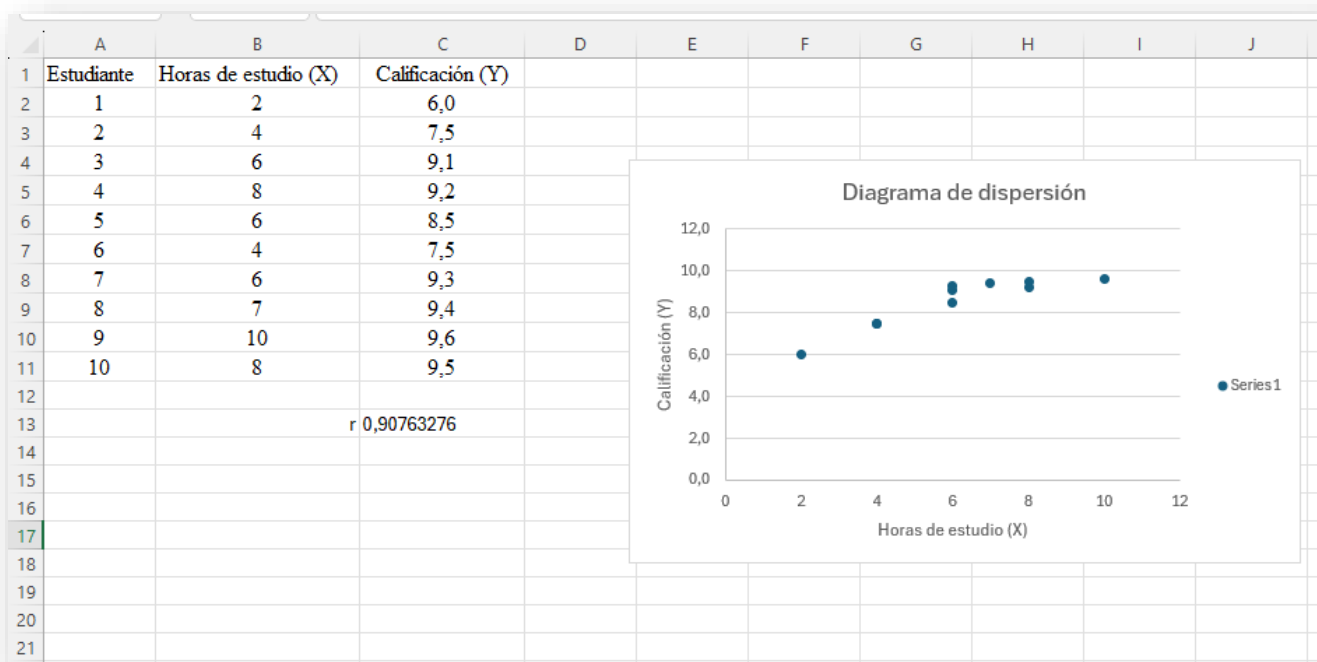
Clic en Diseño Rápido.



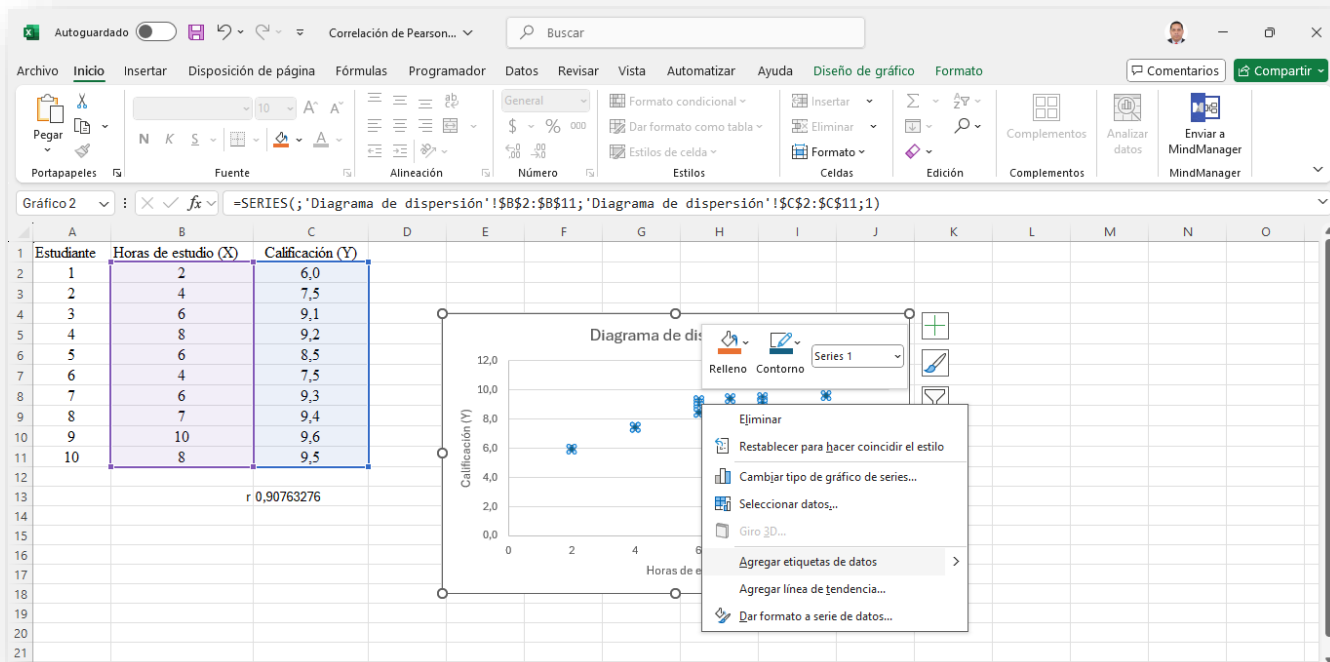
Clic en Diseño 1



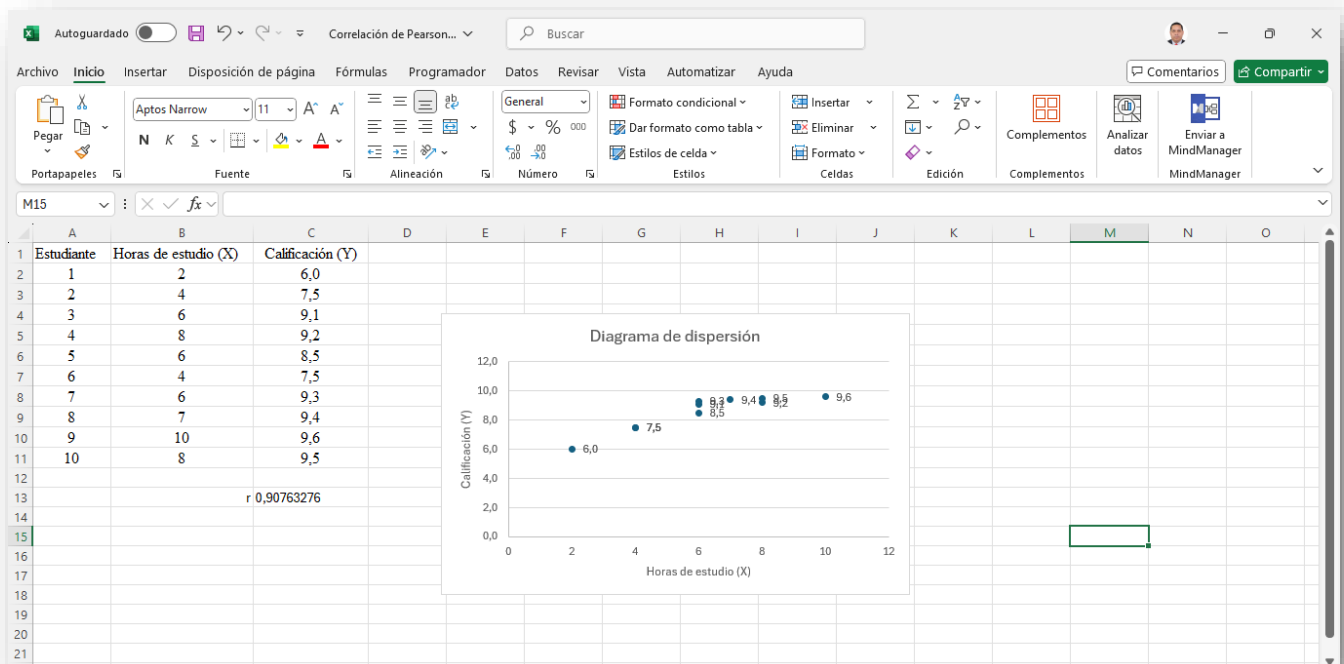
Editar el título y los ejes



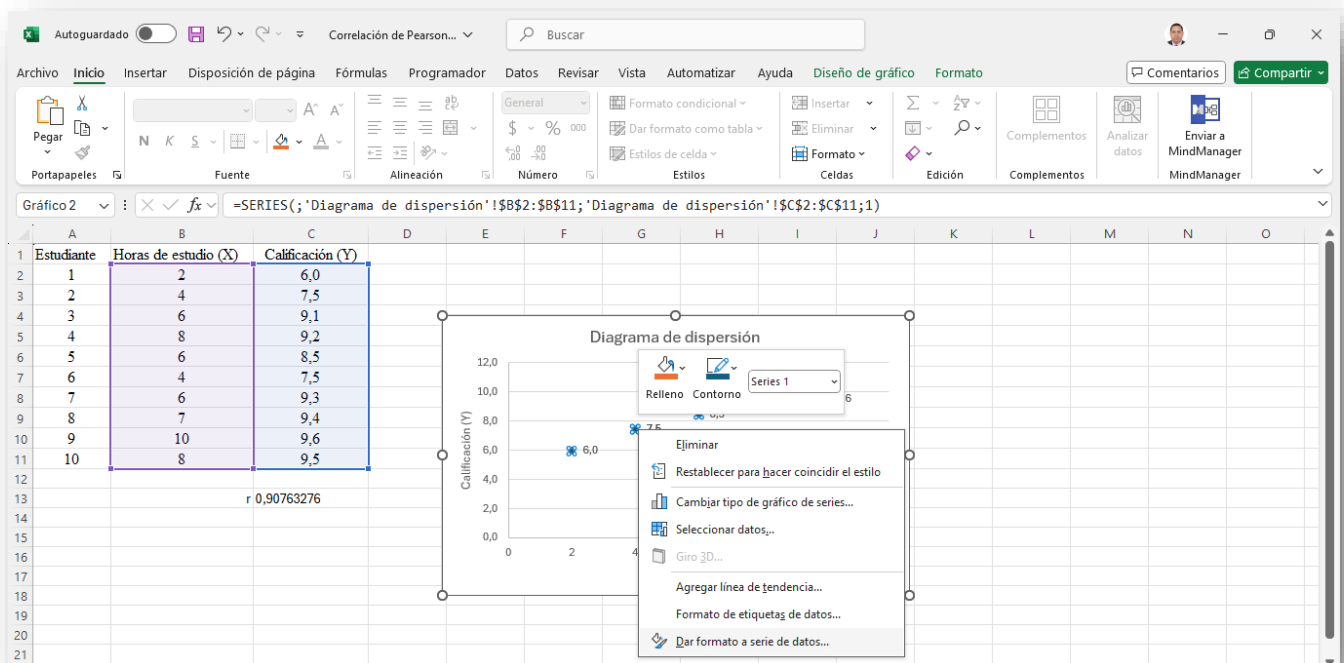
Borrar Series1. Clic derecho en el gráfico. Agregar etiquetas de datos



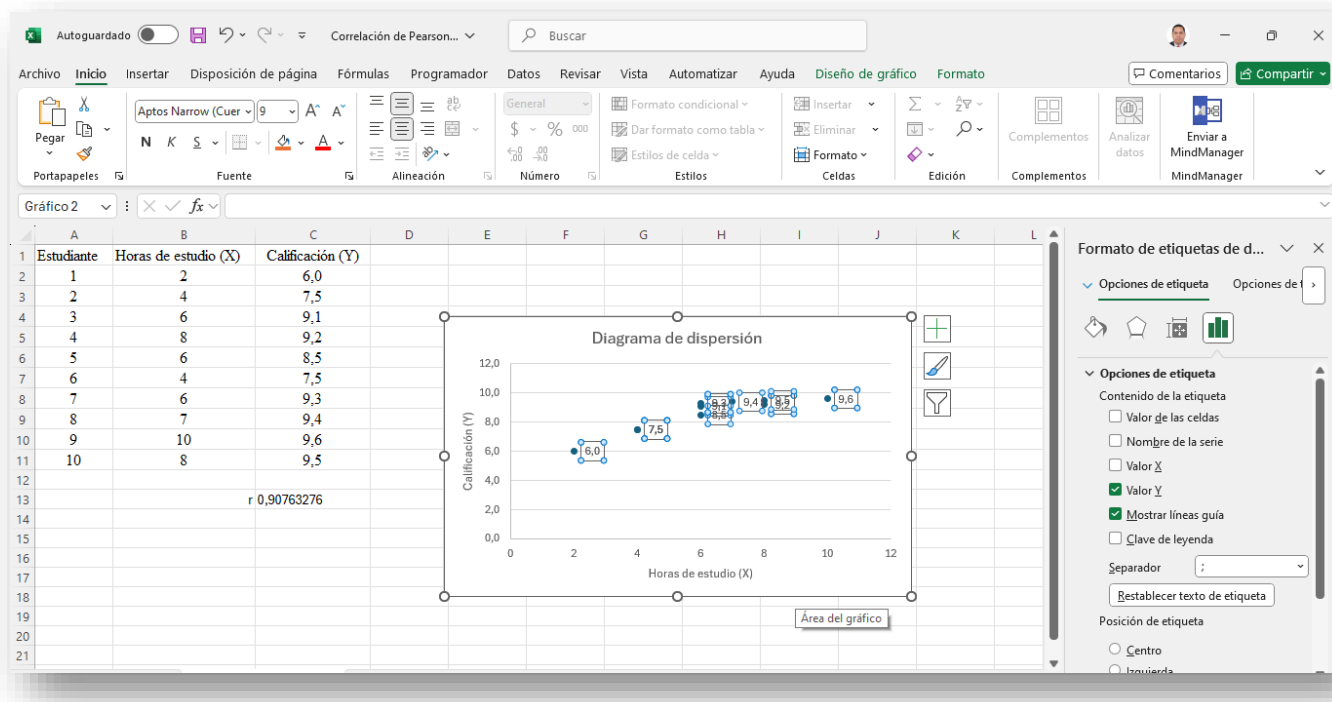
Clic en Agregar etiquetas de datos



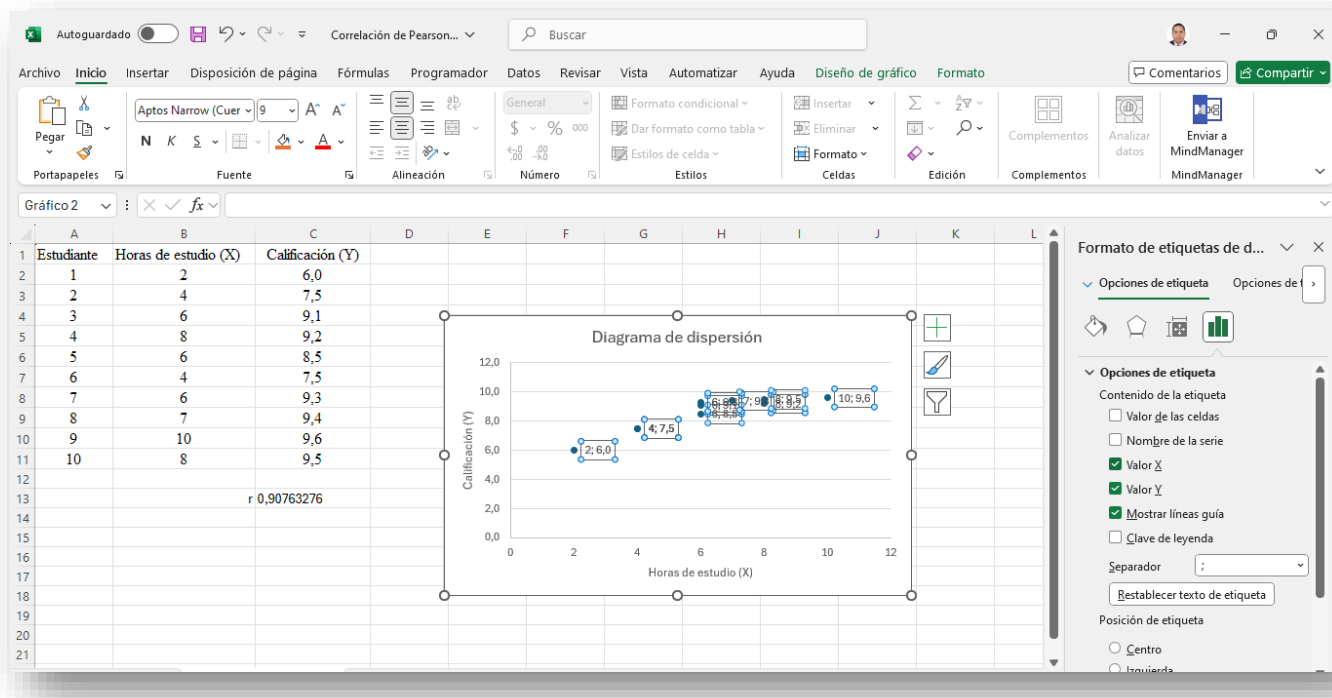
Clic derecho en etiquetas de datos



Clic Formato de etiquetas de datos



Clic en Valor X



Interpretación:

Los puntos se alinean o siguen una tendencia recta en una dirección ascendente, lo que sugiere que a medida que aumenta X, también aumenta Y. El patrón de relación es casi perfectamente lineal.

b) Normalidad

Como los datos son menores a 50 se aplica la prueba de Shapiro-Wilk (para cada variable).

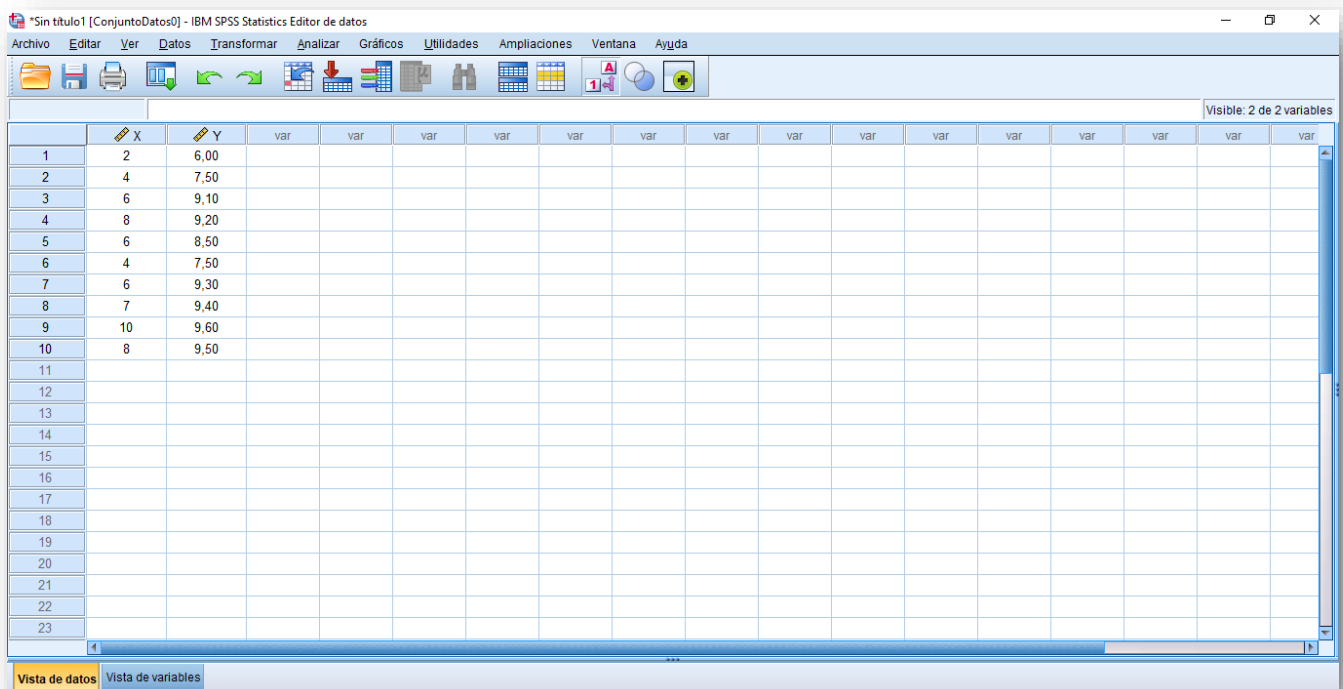
Las hipótesis son:

Hipótesis nula (H_0): Los datos provienen de una distribución normal

Hipótesis alternativa (H_1): Los datos no provienen de una distribución normal

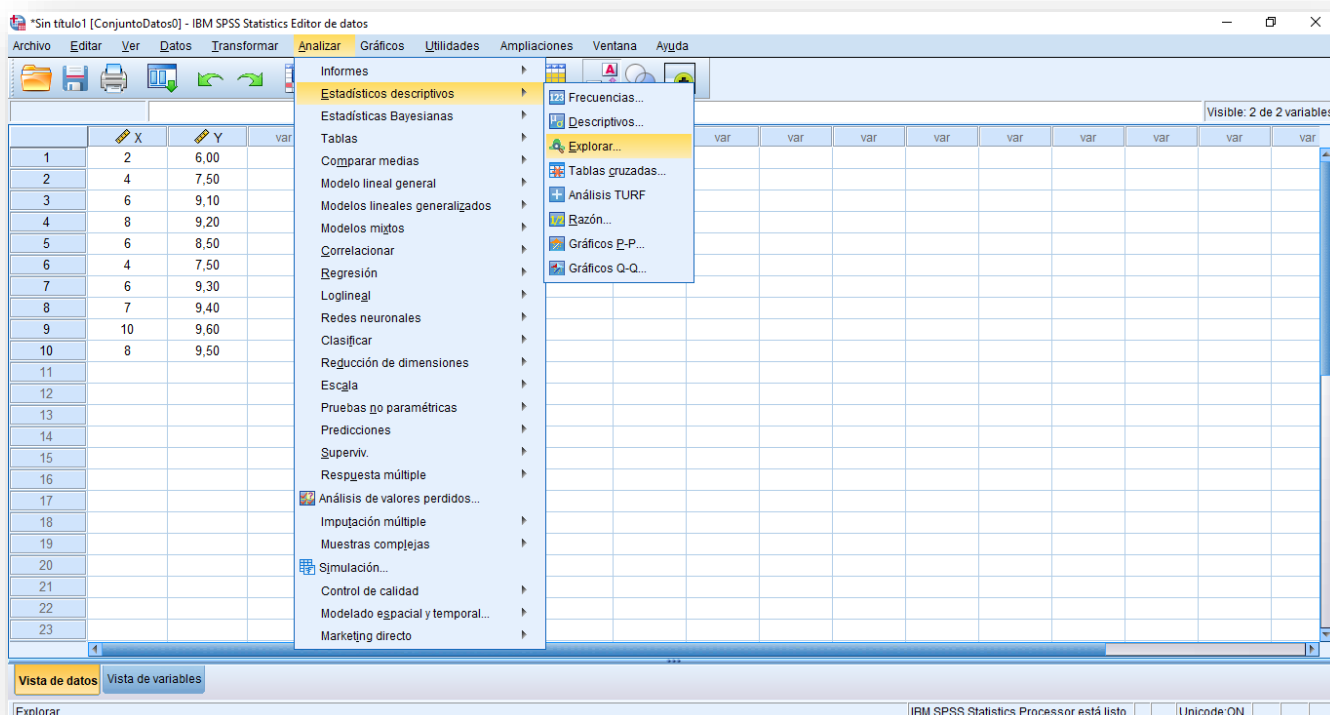
Vamos a calcular empleando SPSS

Copiar los datos en SPSS

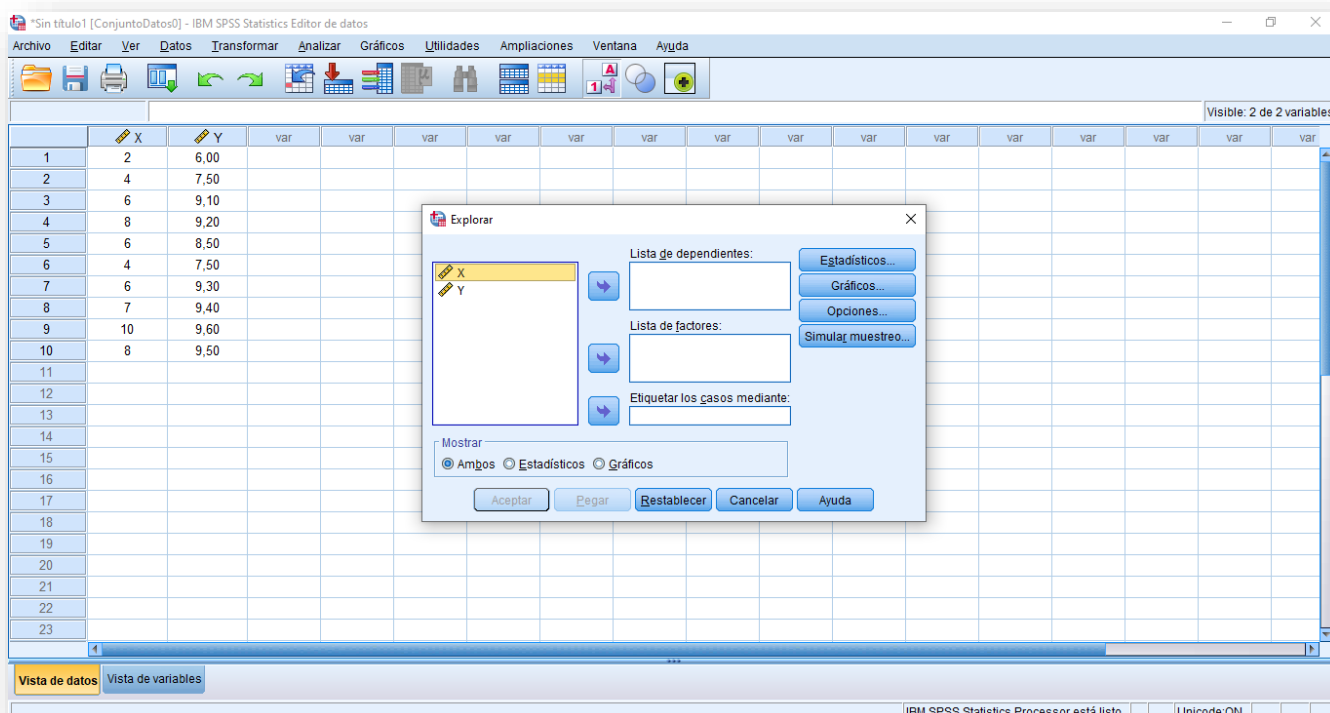


	X	Y	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	2	6,00															
2	4	7,50															
3	6	9,10															
4	8	9,20															
5	6	8,50															
6	4	7,50															
7	6	9,30															
8	7	9,40															
9	10	9,60															
10	8	9,50															
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	

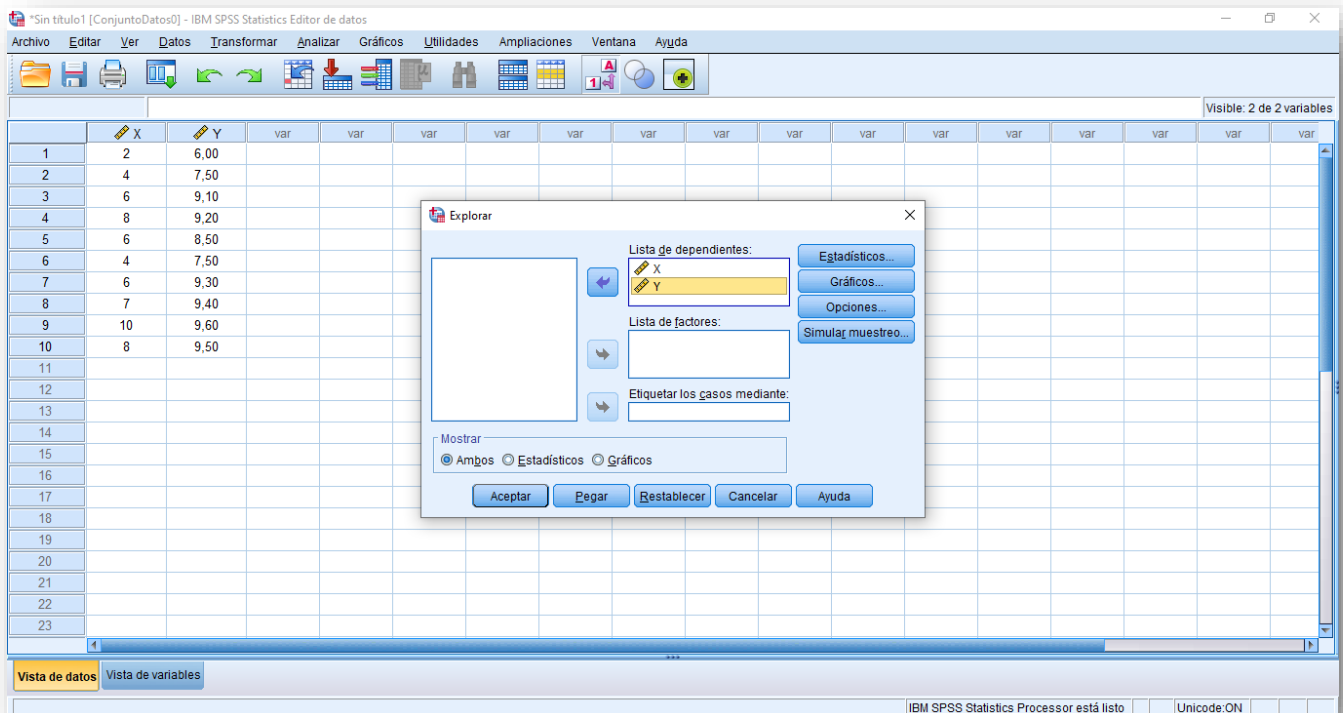
Clic en Analizar. Estadísticos descriptivos. Explorar



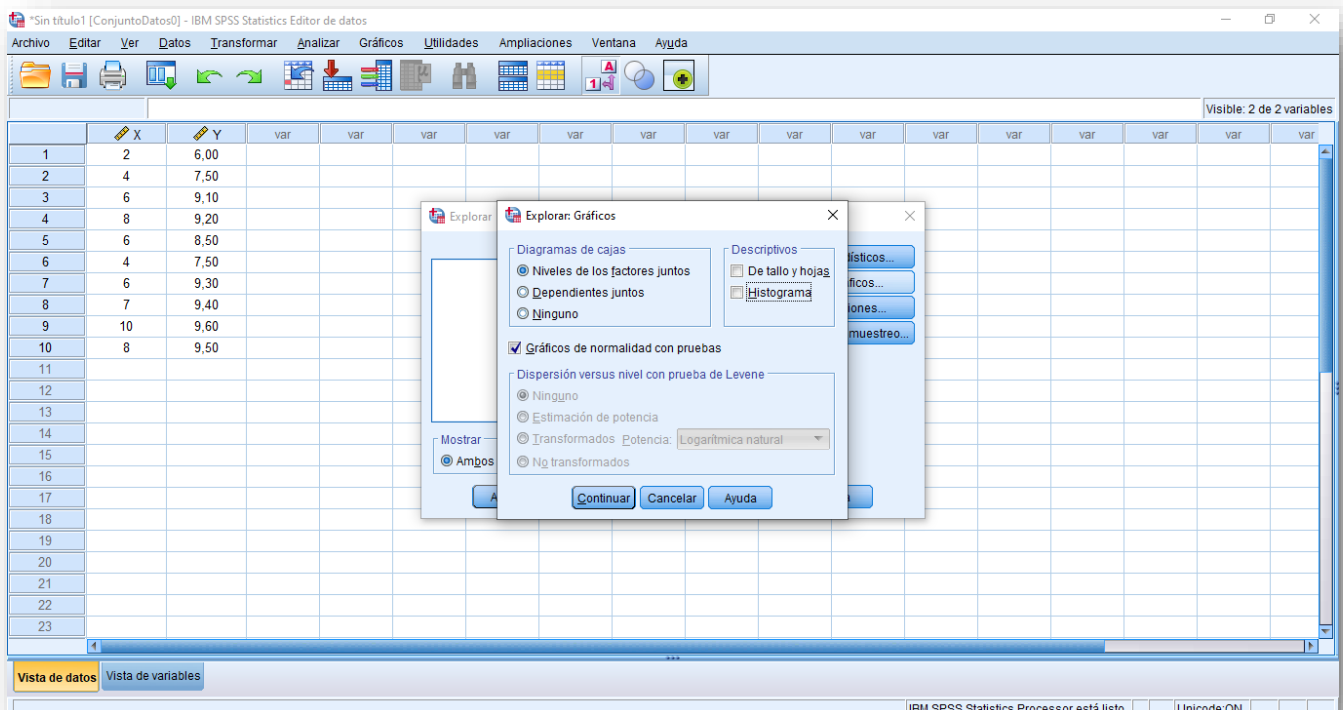
Clic en Explorar



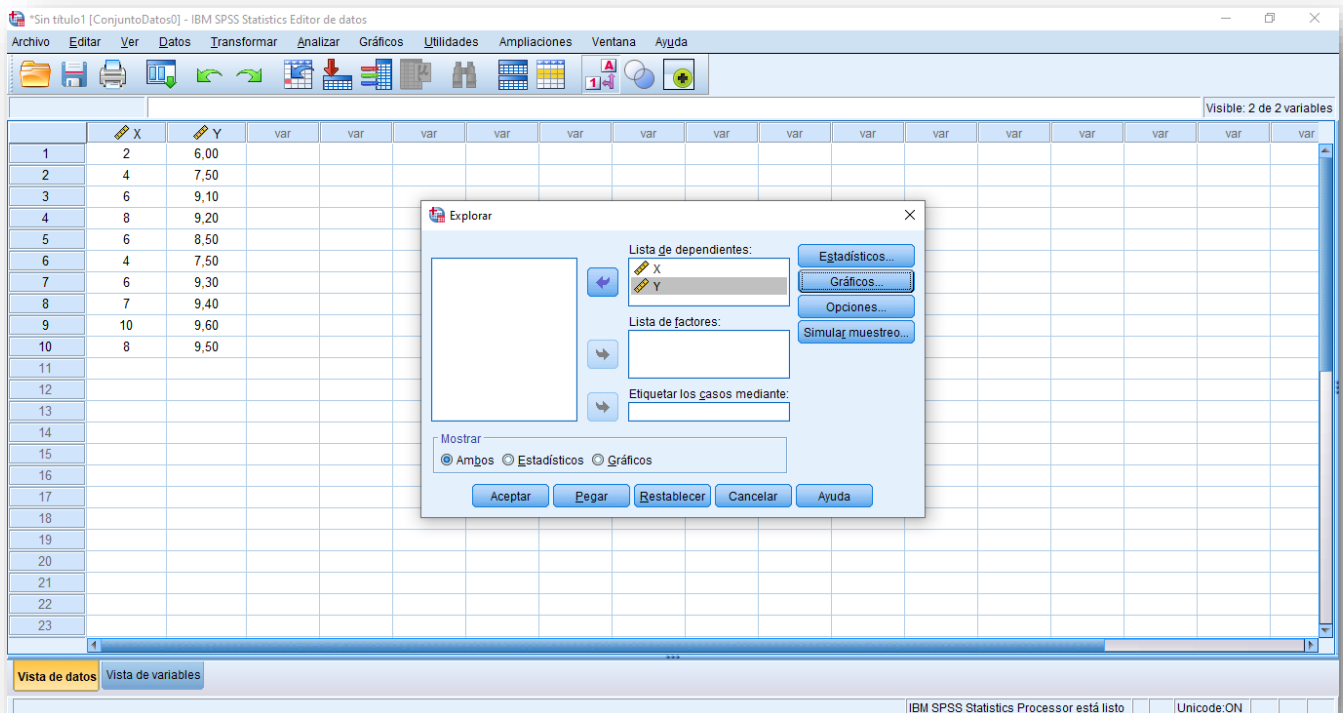
Seleccionar X, Y. Pasar a Lista de dependientes



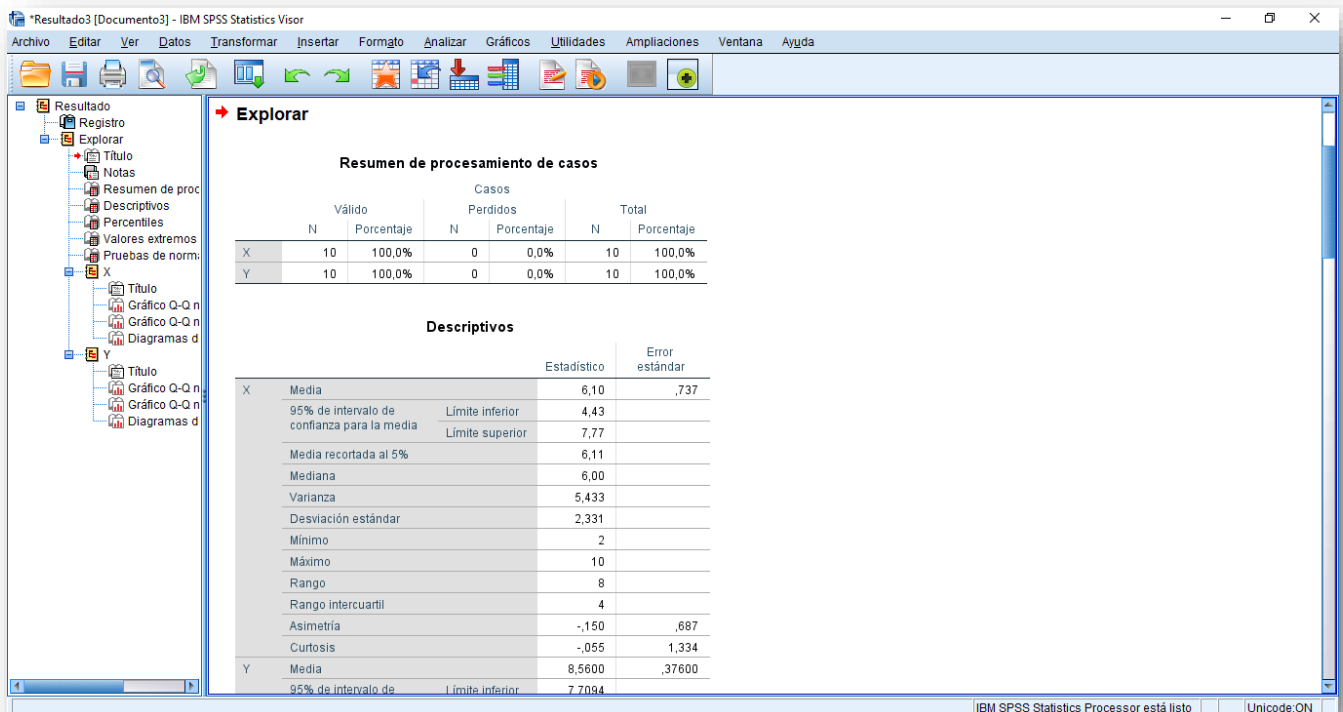
Clic en Gráficos. Clic en Gráficos de normalidad con pruebas.



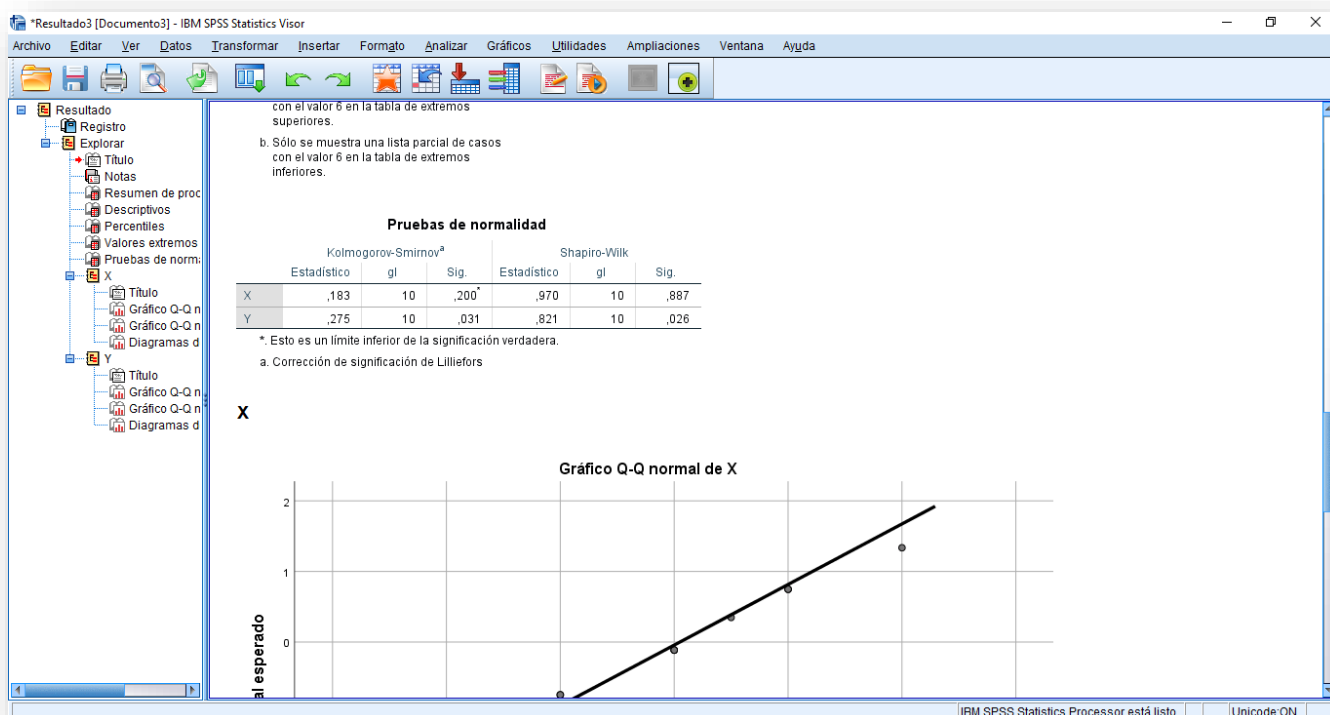
Clic en Continuar



Clic en Aceptar



Desplazamos la barra vertical de navegación hacia abajo para ver los demás resultados



Clic derecho en la tabla de Pruebas de normalidad. Copiar. Pegar

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	Gl	Sig.	Estadístico	gl	Sig.
X	,183	10	,200*	,970	10	,887
Y	,275	10	,031	,821	10	,026

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Interpretación:

En Horas de estudio (X):

Como Sig. 0,887 ($p = 0,887$) es mayor que 0,05 aceptamos H_0 , es decir, no existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos de las horas de estudio provienen de una distribución normal.

En Calificaciones (Y)

Como Sig. 0,026 ($p = 0,026$) es menor que 0,05 no aceptamos H_0 , es decir, existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos de las horas de estudio no provienen de una distribución normal.

Si los datos hubieran cumplido con los supuestos de linealidad y normalidad se hubiera procedido de la siguiente manera:

Se llena la siguiente tabla:

Estudiante	Horas de estudio (X)	Calificación (Y)	X^2	Y^2	XY
1	2	6,0	4	36	12
2	4	7,5	16	56,25	30
3	6	9,1	36	82,81	54,6
4	8	9,2	64	84,64	73,6
5	6	8,5	36	72,25	51
6	4	7,5	16	56,25	30
7	6	9,3	36	86,49	55,8
8	7	9,4	49	88,36	65,8
9	10	9,6	100	92,16	96
10	8	9,5	64	90,25	76
Total	61	85,6	421	745,46	544,8

Reemplazando valores en la fórmula

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{10(544,8) - (61)(85,6)}{\sqrt{[10 \cdot 421 - (61)^2][10 \cdot 745,46 - (85,6)^2]}} = 0,9076 = 0,91$$

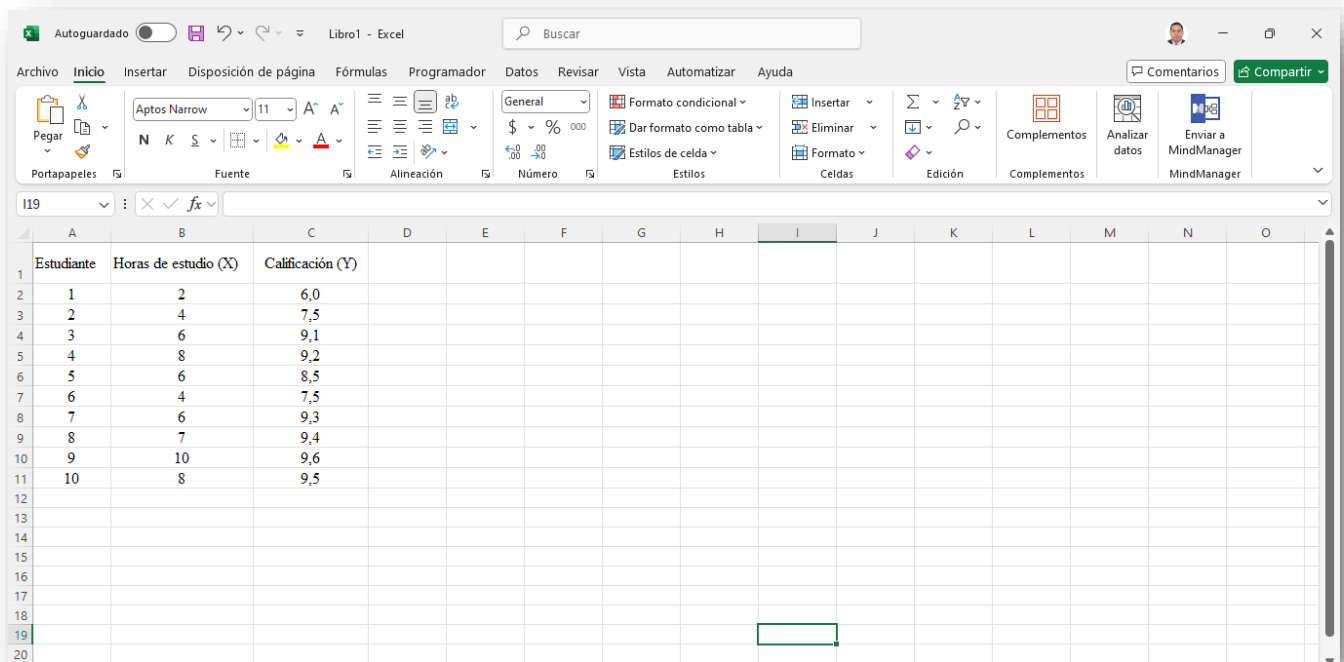
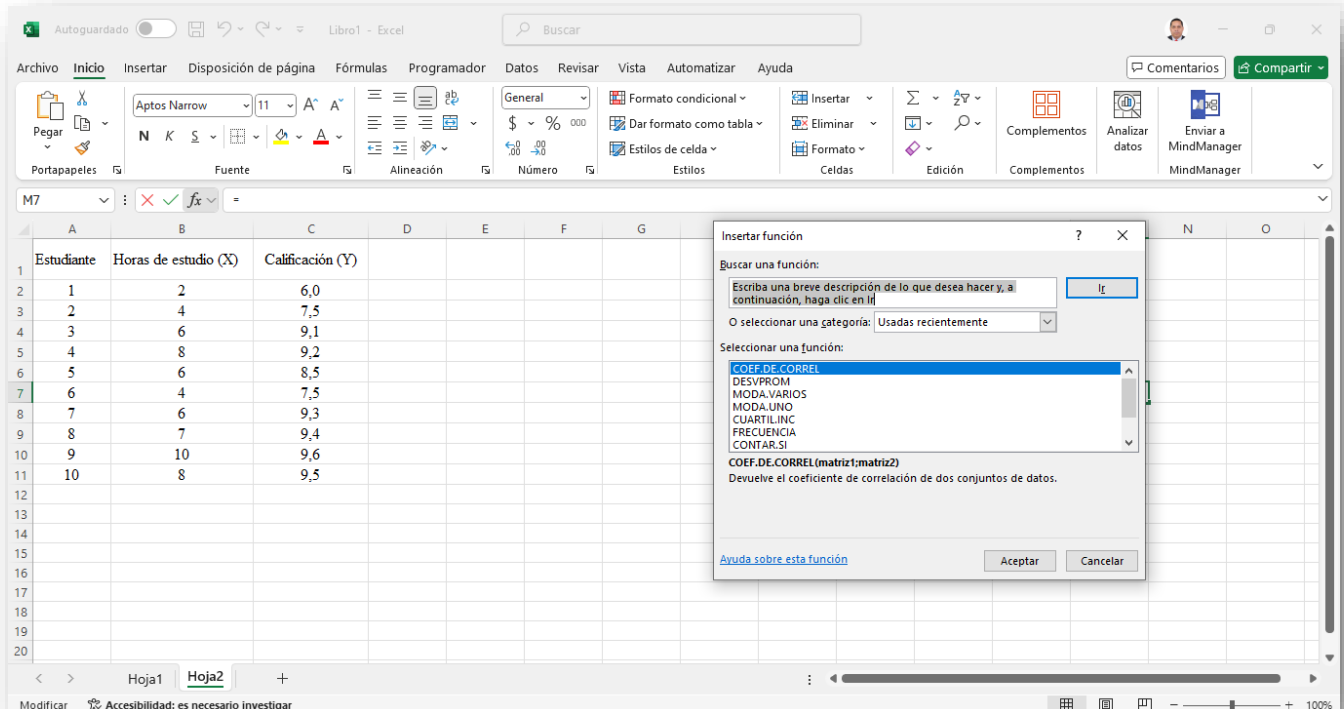
Interpretación:

$r = 0,91$ indica una correlación positiva fuerte entre horas de estudio y calificaciones.

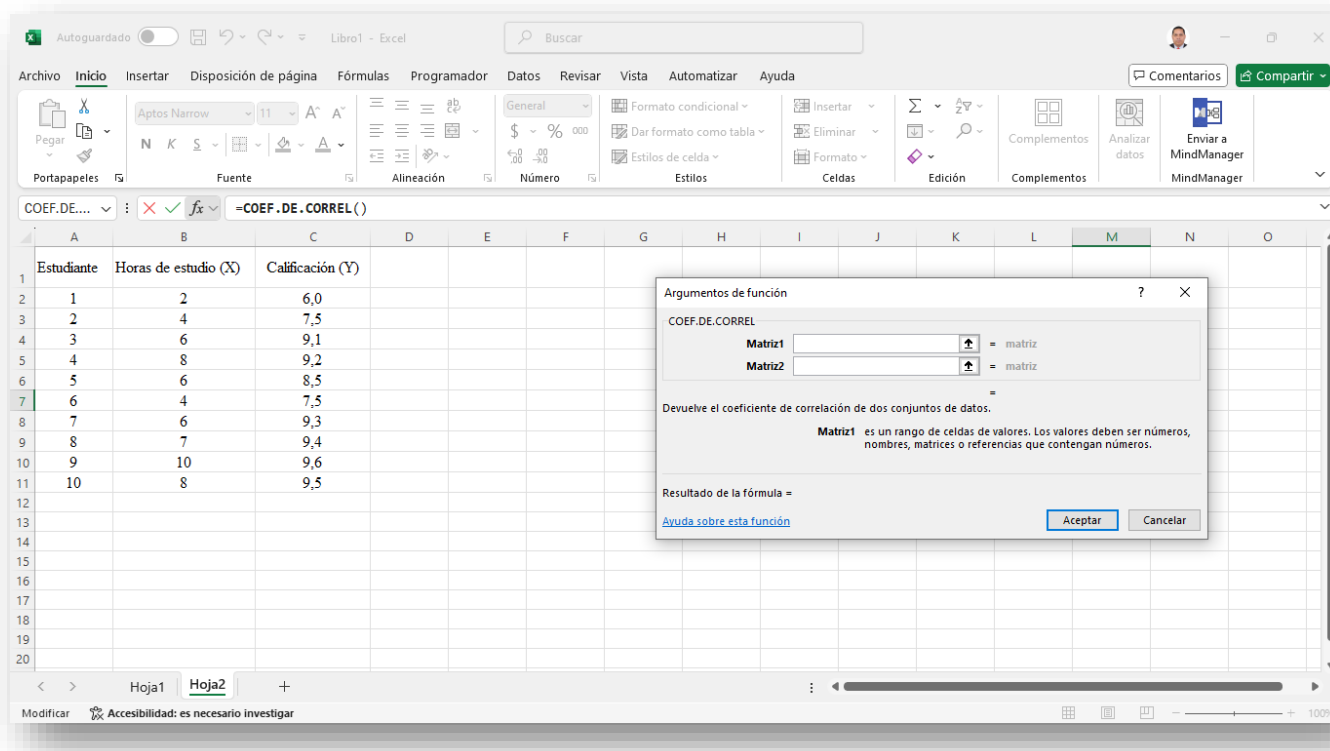
Precaución: Correlación no implica causalidad. Otros factores (Ejemplo: calidad del estudio) podrían influir.

Empleando Excel se procede de la siguiente manera:

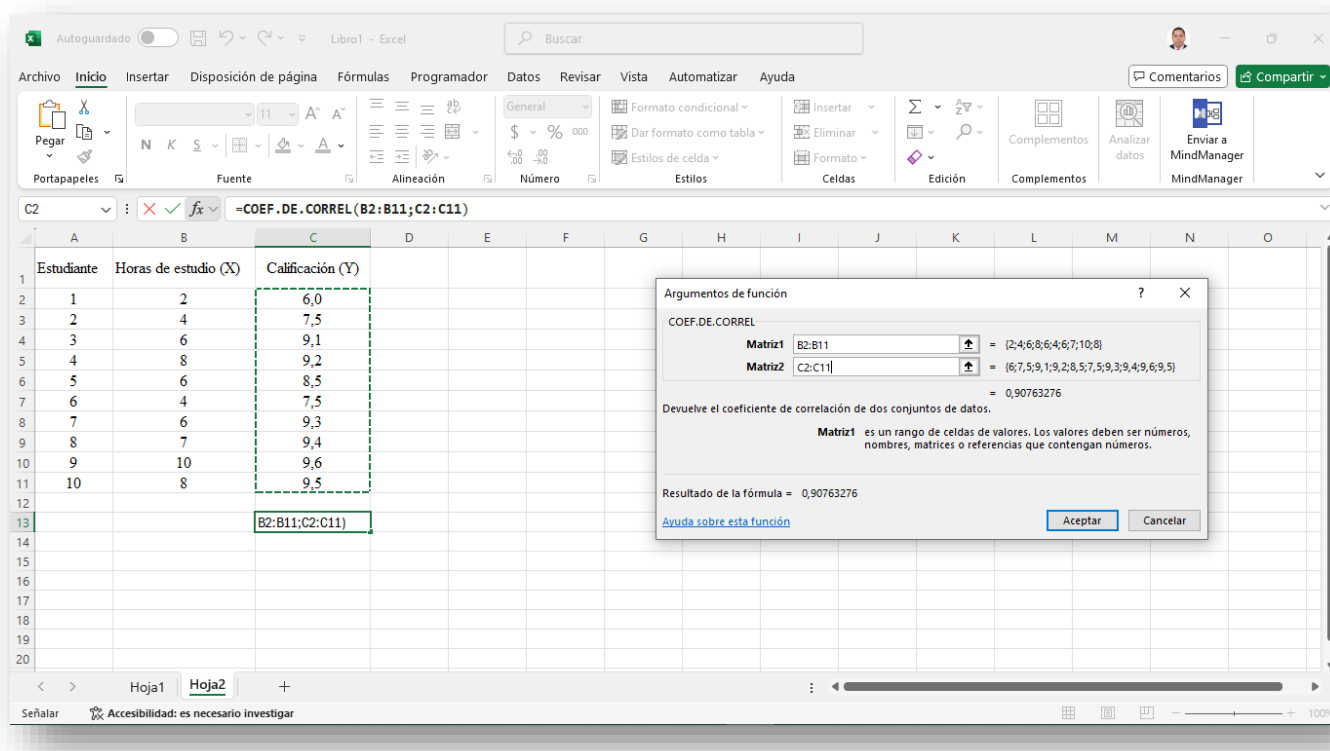
Copiar los datos en Excel

Clic en $f(x)$ para insertar función

En seleccionar una función. Seleccionar COEF.DE.CORREL. Clic en Aceptar



En Matriz 1 seleccionar con el ratón la columna de los datos de X. En Matriz 2 seleccionar la columna de los datos de Y



Clic en Aceptar

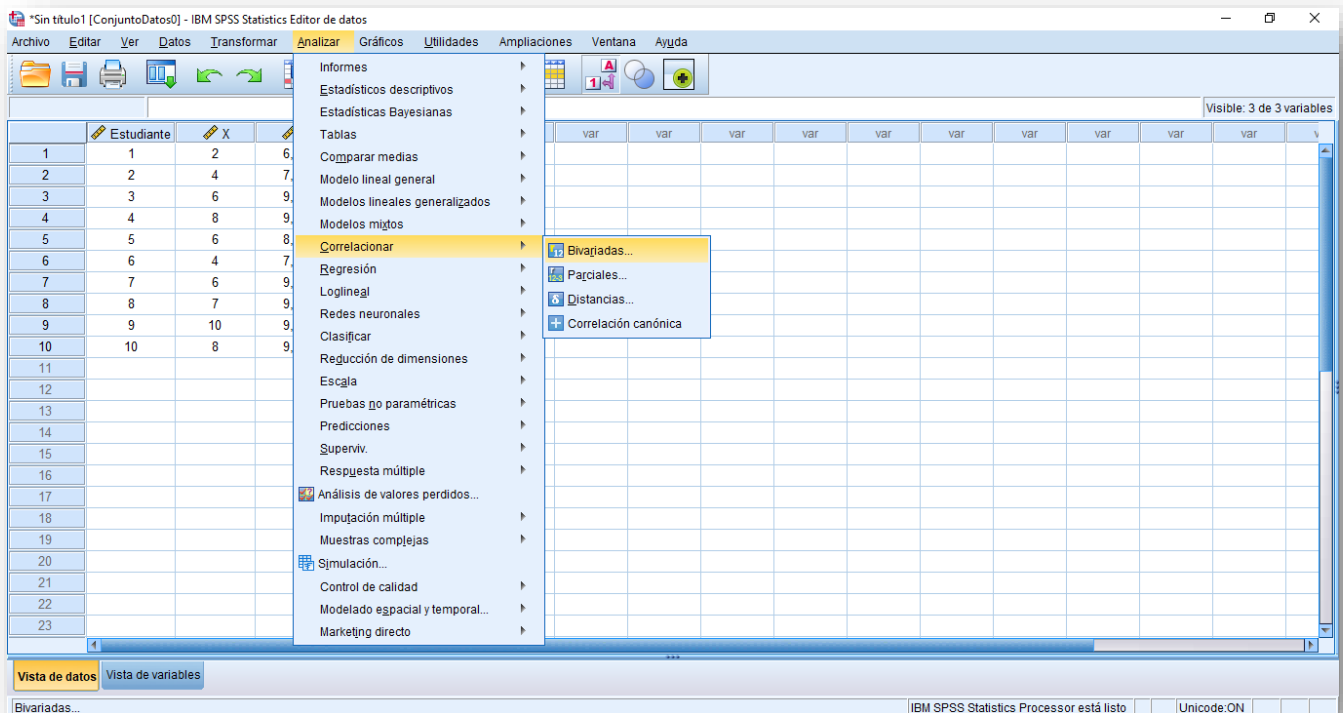
C13    f_x  =COEF.DE.CORREL(B2:B11;C2:C11)					
	A	B	C	D	E
	Estudiante	Horas de estudio (X)	Calificación (Y)		
1					
2	1	2	6,0		
3	2	4	7,5		
4	3	6	9,1		
5	4	8	9,2		
6	5	6	8,5		
7	6	4	7,5		
8	7	6	9,3		
9	8	7	9,4		
10	9	10	9,6		
11	10	8	9,5		
12					
13			0,90763276		

En SPSS

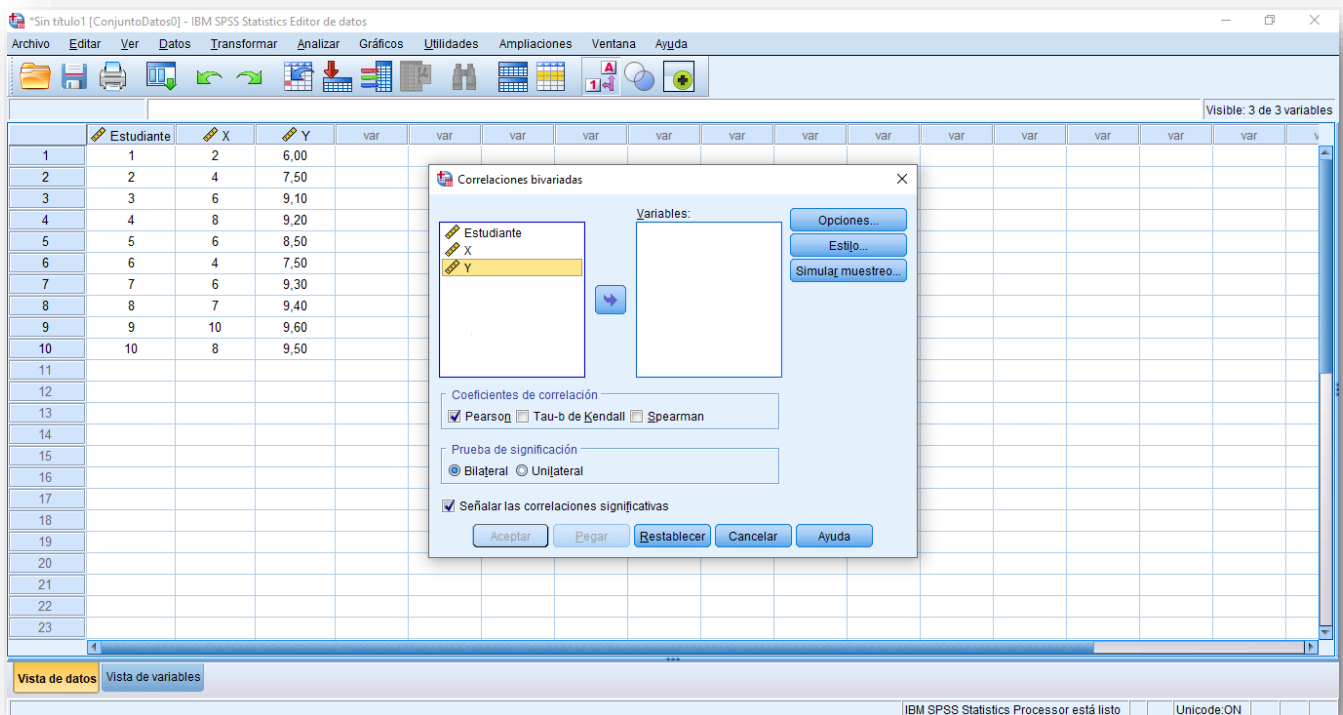
Copiamos los datos en SPSS

*Sin título1 [ConjuntoDatos0] - IBM SPSS Statistics Editor de datos																
Archivo Editor Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda																
Visible: 3 de 3 variables																
	Estudiante	X	Y	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1	2	6,00													
2	2	4	7,50													
3	3	6	9,10													
4	4	8	9,20													
5	5	6	8,50													
6	6	4	7,50													
7	7	6	9,30													
8	8	7	9,40													
9	9	10	9,60													
10	10	8	9,50													
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																

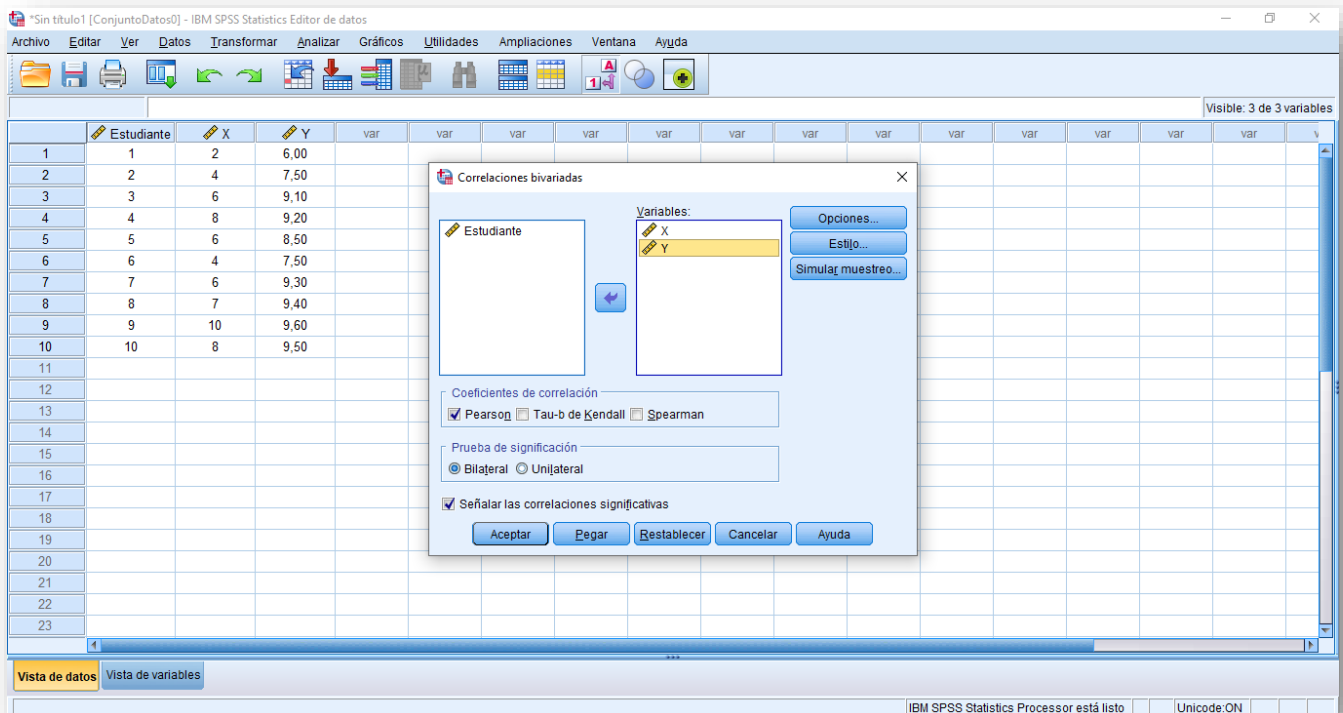
Clic en Analizar. Correlacionar. Bivariadas



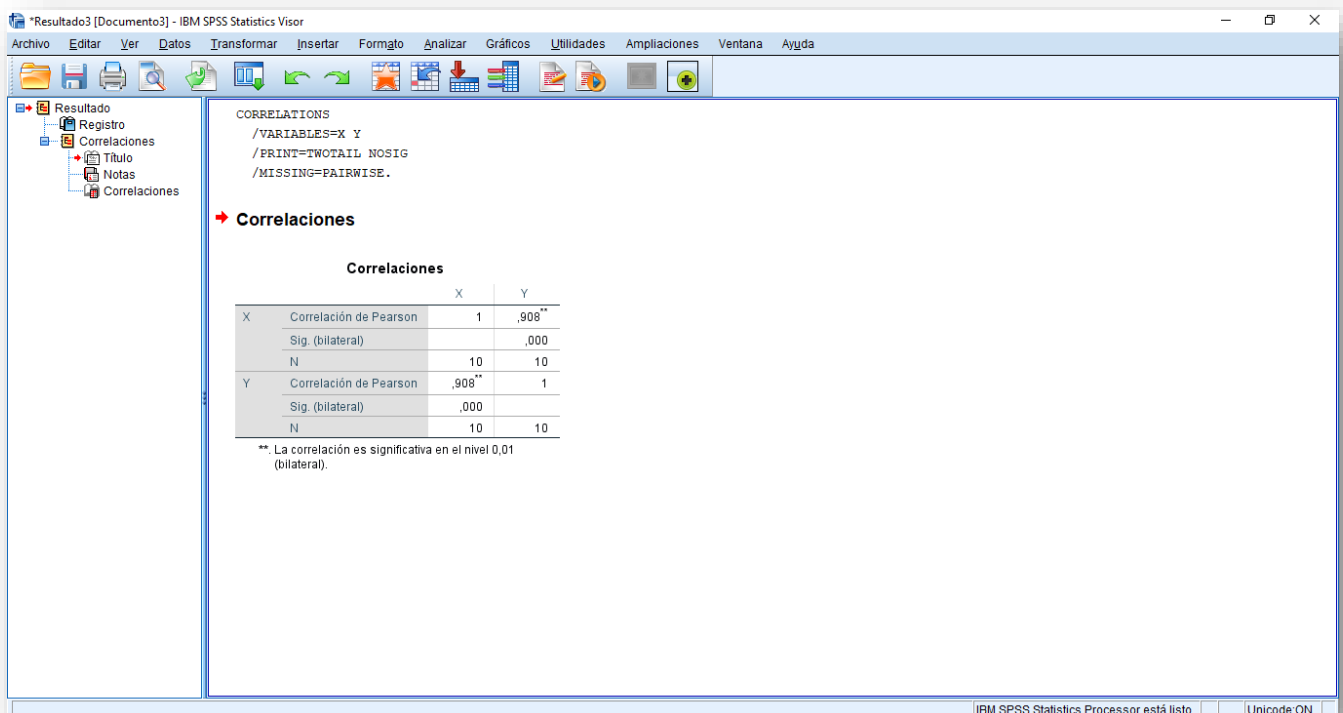
Clic en Bivariadas



Seleccionar las variables X, Y, Clic en la fecha



Clic en Aceptar



Clic derecho en la tabla. Copiar. Pegar

Correlaciones

		X	Y
X	Correlación de Pearson	1	,908**
	Sig. (bilateral)		,000
	N	10	10
Y	Correlación de Pearson	,908**	1
	Sig. (bilateral)	,000	
	N	10	10

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

Interpretación:

$r = 0,908 = 0,91$ indica una correlación positiva fuerte entre horas de estudio y calificaciones.

Nota: Sin embargo, los datos de Y no cumplieron con el supuesto de normalidad, por lo que esta interpretación no estaría correcta. Habría que calcular el coeficiente de correlación de Spearman para obtener una interpretación más apegada a la realidad (Ver Ejemplo de Aplicación 3 de la Correlación de Spearman).

2) Calcular el coeficiente de la correlación de Pearson de la relación entre precio y demanda de un producto empleando los siguientes datos:

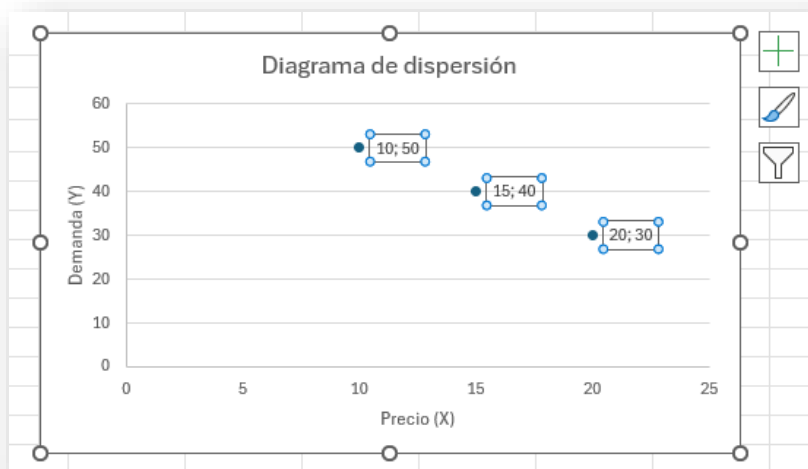
Producto	Precio (X)	Demanda (Y)
A	10	50
B	15	40
C	20	30

Solución:

A continuación, se comprueban los supuestos necesarios para aplicar el coeficiente de correlación de Pearson a los datos de horas de estudio (X) y calificaciones (Y).

a) Relación Lineal

Método: Gráfico de dispersión (*Scatter Plot*). Empleando Excel



Interpretación:

Los puntos se alinean o siguen una tendencia recta en una dirección descendente, lo que sugiere que a medida que aumenta X, disminuye Y. El patrón de relación es perfectamente lineal.

b) Normalidad

Como los datos son menores a 50 se aplica la prueba de Shapiro-Wilk (para cada variable).

Las hipótesis son:

Hipótesis nula (H_0): Los datos provienen de una distribución normal

Hipótesis alternativa (H_1): Los datos no provienen de una distribución normal

Vamos a calcular empleando SPSS

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	Gl	Sig.	Estadístico	gl	Sig.
X	,175	3	.	1,000	3	1,000
Y	,175	3	.	1,000	3	1,000

a. Corrección de significación de Lilliefors

Interpretación:

Precio (X): Como Sig. 1,000 ($p = 1,000$) es mayor que 0,05 aceptamos H_0 , es decir, no existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos del precio siguen una distribución normal.

Demanda (Y): Como Sig. 1,000 ($p = 1,000$) es mayor que 0,05 aceptamos H_0 , es decir, no existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos de la demanda siguen una distribución normal.

Como cumple con los supuestos, se procede a calcular el coeficiente de correlación de Pearson.

Se contrastan las hipótesis:

Hipótesis nula (H_0): No hay correlación lineal ($r = 0$)

Hipótesis alternativa (H_1): Existe correlación lineal ($r \neq 0$)

Regla de decisión clave

p-valor	Interpretación
$p < 0,05$	Correlación significativa: Rechazamos H_0 . Hay evidencia de relación lineal.
$p \geq 0,05$	No significativa: No hay suficiente evidencia para rechazar H_0

Empleando SPSS:

Correlaciones

		X	Y
X	Correlación de Pearson	1	-1,000**
	Sig. (bilateral)		,000
	N	3	3
Y	Correlación de Pearson	-1,000**	1
	Sig. (bilateral)	,000	
	N	3	3

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

Interpretación:

Coeficiente ($r = -1,000$): → Correlación negativa perfecta. Significa que por cada aumento unitario en el precio (X), la demanda (Y) disminuye de manera exactamente proporcional.

p-valor (Sig. bilateral): 0,000 ($< 0,01$) → La correlación es altamente significativa. Rechazamos la hipótesis nula y concluimos que existe una correlación lineal y esta relación no es aleatoria.

3) Calcular el coeficiente de la correlación de Pearson de la relación entre la edad y la presión arterial empleando los siguientes datos de 4 pacientes:

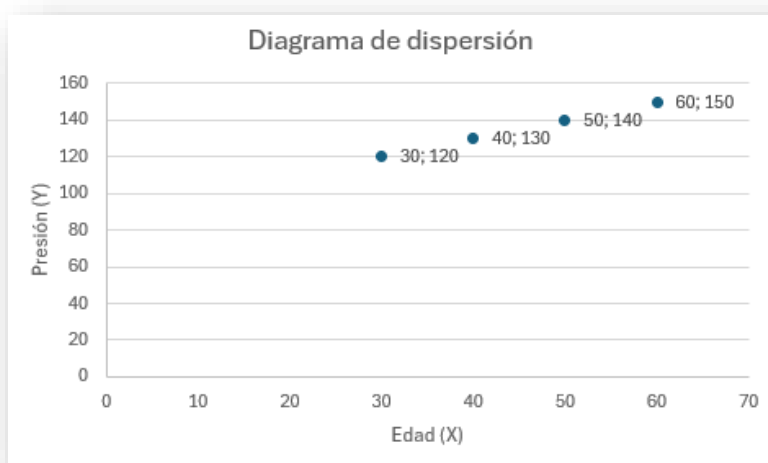
Paciente	Edad (X)	Presión (Y)
1	30	120
2	40	130
3	50	140
4	60	150

Solución:

A continuación, se comprueban los supuestos necesarios para aplicar el coeficiente de correlación de Pearson a los datos de horas de estudio (X) y calificaciones (Y).

a) Relación Lineal

Método: Gráfico de dispersión (*Scatter Plot*). Empleando Excel



Interpretación:

Los puntos se alinean o siguen una tendencia recta en una dirección ascendente, lo que sugiere que a medida que aumenta X, aumenta Y. El patrón de relación es perfectamente lineal.

b) Normalidad

Como los datos son menores a 50 se aplica la prueba de Shapiro-Wilk (para cada variable).

Las hipótesis son:

Hipótesis nula (H_0): Los datos provienen de una distribución normal

Hipótesis alternativa (H_1): Los datos no provienen de una distribución normal

Vamos a calcular empleando SPSS

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	Gl	Sig.	Estadístico	gl	Sig.
Edad_X	,151	4	.	,993	4	,972
Presión_Y	,151	4	.	,993	4	,972

a. Corrección de significación de Lilliefors

Interpretación:

Precio (X): Como Sig. 0,972 ($p = 0,972$) es mayor que 0,05 aceptamos H_0 , es decir, no existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos de la Edad siguen una distribución normal.

Demanda (Y): Como Sig. 0,972 ($p = 0,972$) es mayor que 0,05 aceptamos H_0 , es decir, no existe evidencia significativa para rechazar H_0 , por tanto, se concluye que los datos de la Presión siguen una distribución normal.

Como cumple con los supuestos, se procede a calcular el coeficiente de correlación de Pearson.

Se contrastan las hipótesis:

Hipótesis nula (H_0): No hay correlación lineal ($r = 0$)

Hipótesis alternativa (H_1): Existe correlación lineal ($r \neq 0$)

Regla de decisión clave

p-valor	Interpretación
$p < 0,05$	Correlación significativa: Rechazamos H_0 . Hay evidencia de relación lineal.
$p \geq 0,05$	No significativa: No hay suficiente evidencia para rechazar H_0

Calculando r con SPSS:

Correlaciones

		Edad_X	Presión_Y
Edad_X	Correlación de Pearson	1	1,000**
	Sig. (bilateral)		,000
	N	4	4
Presión_Y	Correlación de Pearson	1,000**	1
	Sig. (bilateral)	,000	
	N	4	4

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

Interpretación:

Coeficiente ($r = 1,000$): → Correlación positiva perfecta. Significa que por cada aumento unitario en la Edad (X), la Presión (Y) aumenta de manera exactamente proporcional.

Valor p (Sig. bilateral): 0,000 ($< 0,01$) → La correlación es altamente significativa. Rechazamos la hipótesis nula y concluimos que existe una correlación lineal y esta relación no es aleatoria.

3.4) COEFICIENTE DE CORRELACIÓN DE SPEARMAN

A) DEFINICIÓN

El coeficiente de correlación rho de Spearman (ρ o r_s) es una medida no paramétrica de la correlación monótona entre dos variables. A diferencia del coeficiente de Pearson, que mide relaciones lineales, Spearman evalúa si una variable tiende a aumentar (o disminuir) cuando la otra lo hace, sin necesidad de que la relación sea lineal, es decir, captura tendencias generales, incluso si son curvas o escalonadas. Se basa en los rangos de los datos en lugar de sus valores brutos, es decir, es aplicado después de transformar los datos originales en rangos, lo que lo hace robusto frente a outliers y adecuado para variables ordinales o no normales.

B) CARACTERÍSTICAS

No paramétrico: No asume normalidad en los datos

Rango basado: Una la posición relativa (orden) de los datos

Mide correlación monotónica: Detecta la relación entre dos variables en la que la dirección (crecientes o decrecientes) de la asociación es consistente, pero no necesariamente lineal. Esto significa que:

- Si una variable aumenta, la otra tiende a aumentar (monótona creciente).
- Si una variable aumenta, la otra tiende a disminuir (monótona decreciente).

Rango de valores: Varía de -1 a 1

$\rho = 1$: Correlación monotónica perfecta positiva.

$\rho = -1$: Correlación monotónica perfecta negativa.

$\rho = 0$: No hay correlación monotónica.

Robusto ante outliers (valores atípicos): Es resistente a valores extremos porque trabaja con órdenes (rangos) en lugar de los valores originales. Esto lo hace ideal para datos con distribuciones no normales o posibles errores de medición.

C) DIFERENCIA CON LA CORRELACIÓN LINEAL (PEARSON)

Aspecto	Pearson (Lineal)	Spearman (Monotónica)
Tipo de relación	Solo rectas	Cualquier tendencia creciente/decreciente
Sensibilidad	Sensible a no linealidad y outliers	Robustez frente a outliers y no linealidad
Datos requeridos	Continuos y normales	Ordinales, continuos o no normales

D) FÓRMULA

$$\rho = r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Donde:

d_i = Diferencia entre los rangos de las dos variables para el i-ésimo caso

n = Número de observaciones

E) EJEMPLOS DE APLICACIÓN

1) Datos sin empates: Calcular el coeficiente de Spearman para las siguientes calificaciones de 5 estudiantes en Matemática (X) y Física (Y):

Estudiante	Matemática (X)	Física (Y)
A	8,0	7,5
B	6,0	5,5
C	9,0	8,5
D	5,0	6,0
E	7,0	6,5

Solución:

a) Se asigna rangos

Se ordena X e Y por separado y asignamos rangos (1 al más bajo):

Rangos para X :

X : 5,0(1); 6,0(2); 7,0(3); 8,0(4); 9,0(5)

Rangos para Y :

Y : 5,5(1); 6,0(2); 6,5(3); 7,5(4); 8,5(5)

b) Tabla con rangos

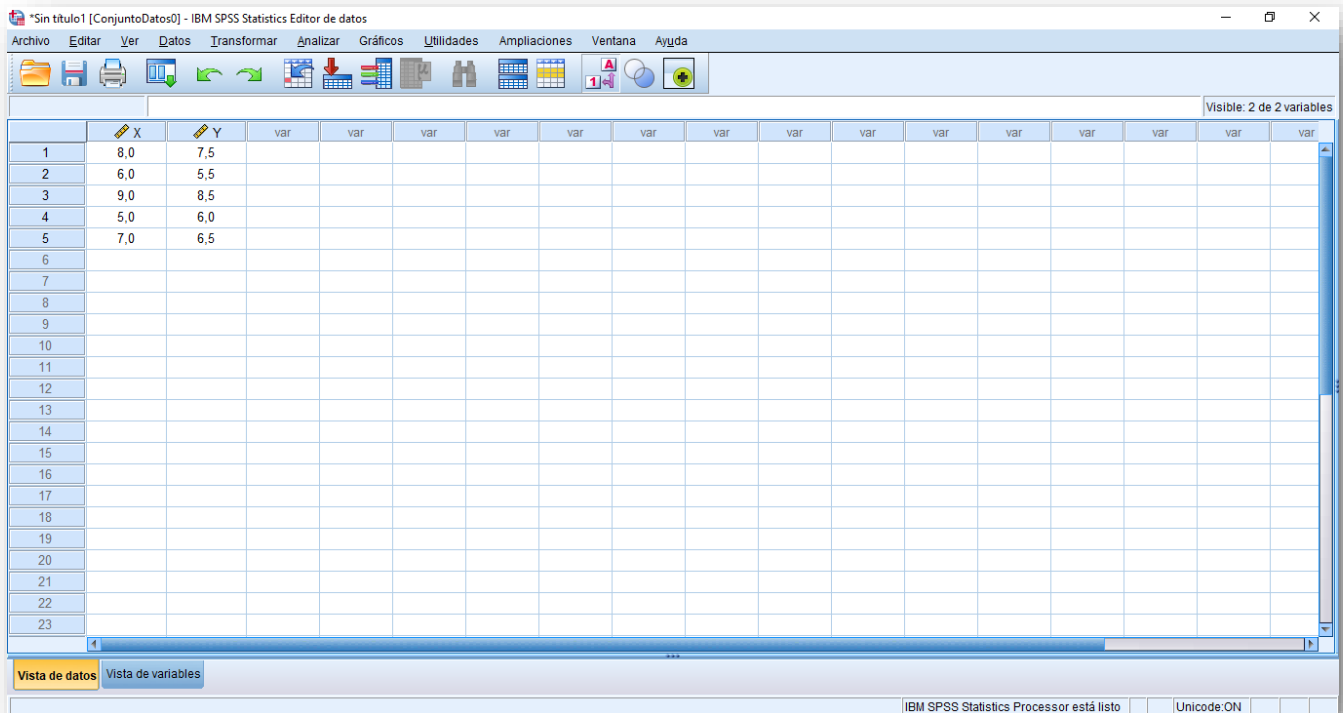
Estudiante	Matemática (X)	Física (Y)	Rango X (R_x)	Rango Y (R_y)	$d_i = R_x - R_y$	d_i^2
A	8,0	7,5	4	4	0	0
B	6,0	5,5	2	1	1	1
C	9,0	8,5	5	5	0	0
D	5,0	6,0	1	2	-1	1
E	7,0	6,5	3	3	0	0
Total						2

c) Reemplazando valores en la fórmula

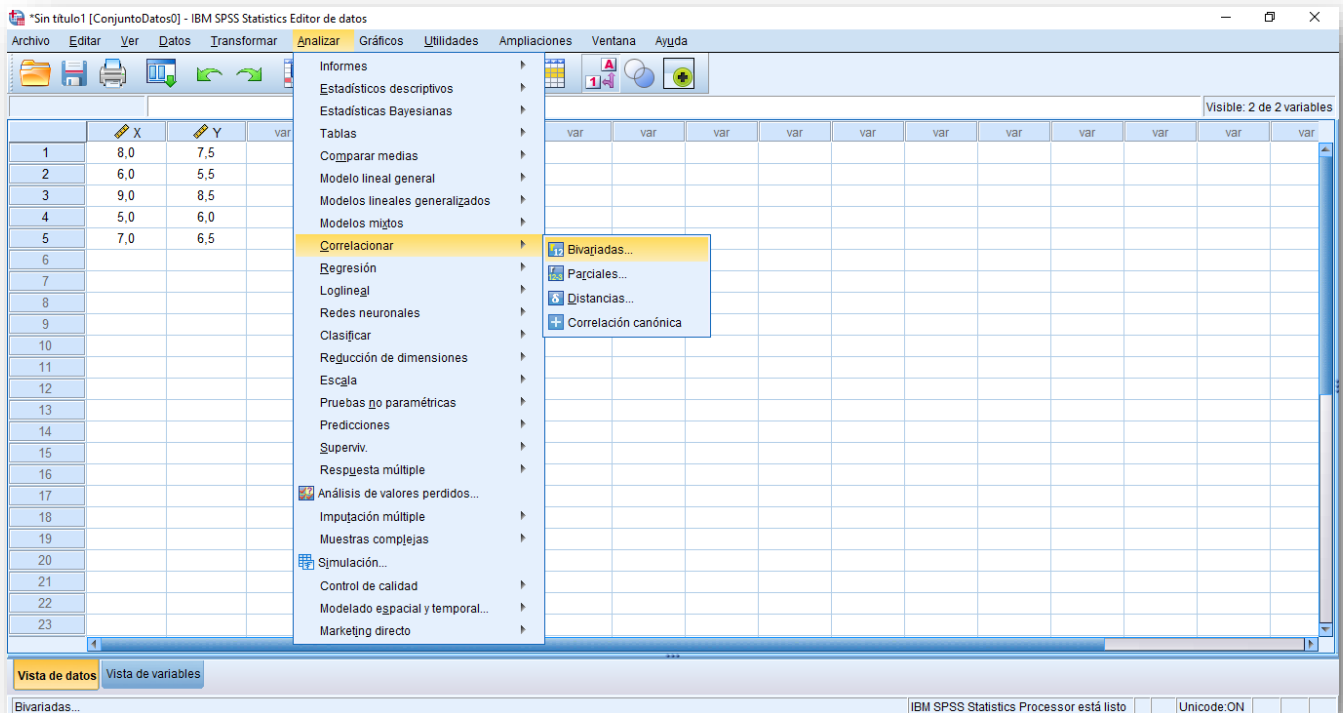
$$\rho = r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6(2)}{5(5^2 - 1)} = 0,9$$

d) En SPSS

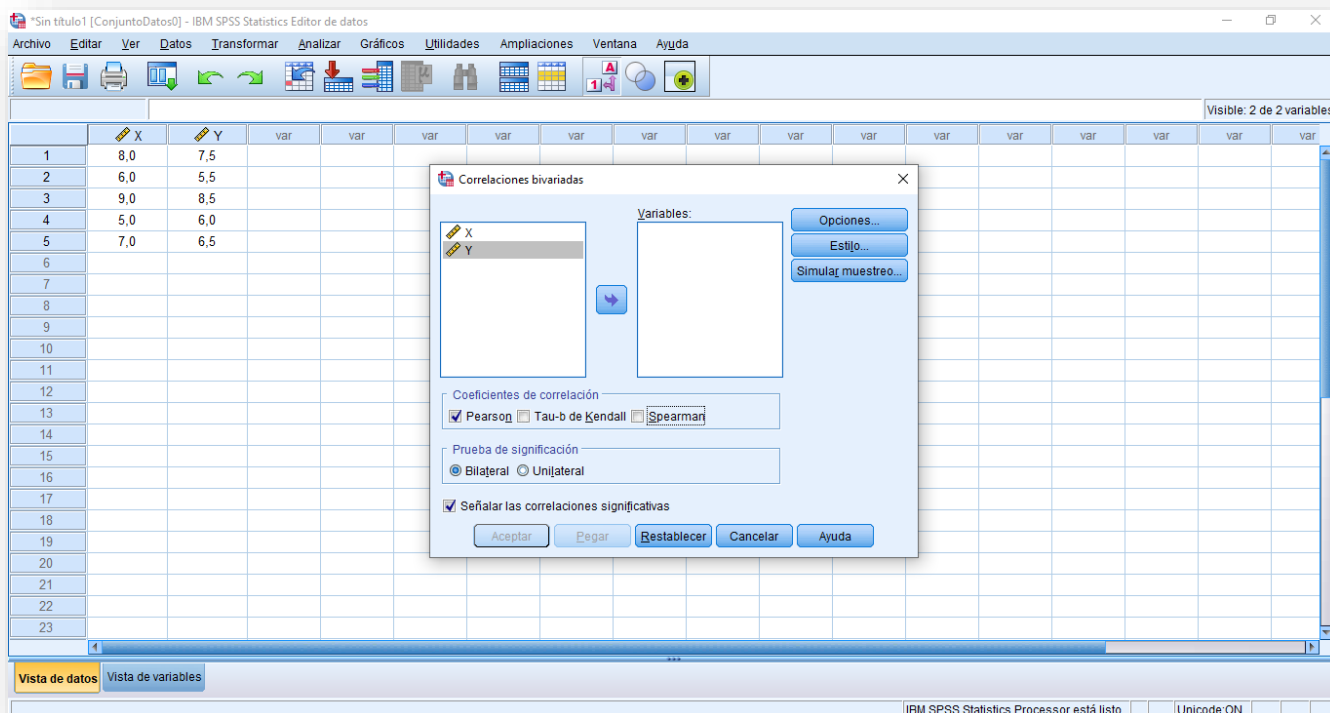
Ingresar los datos



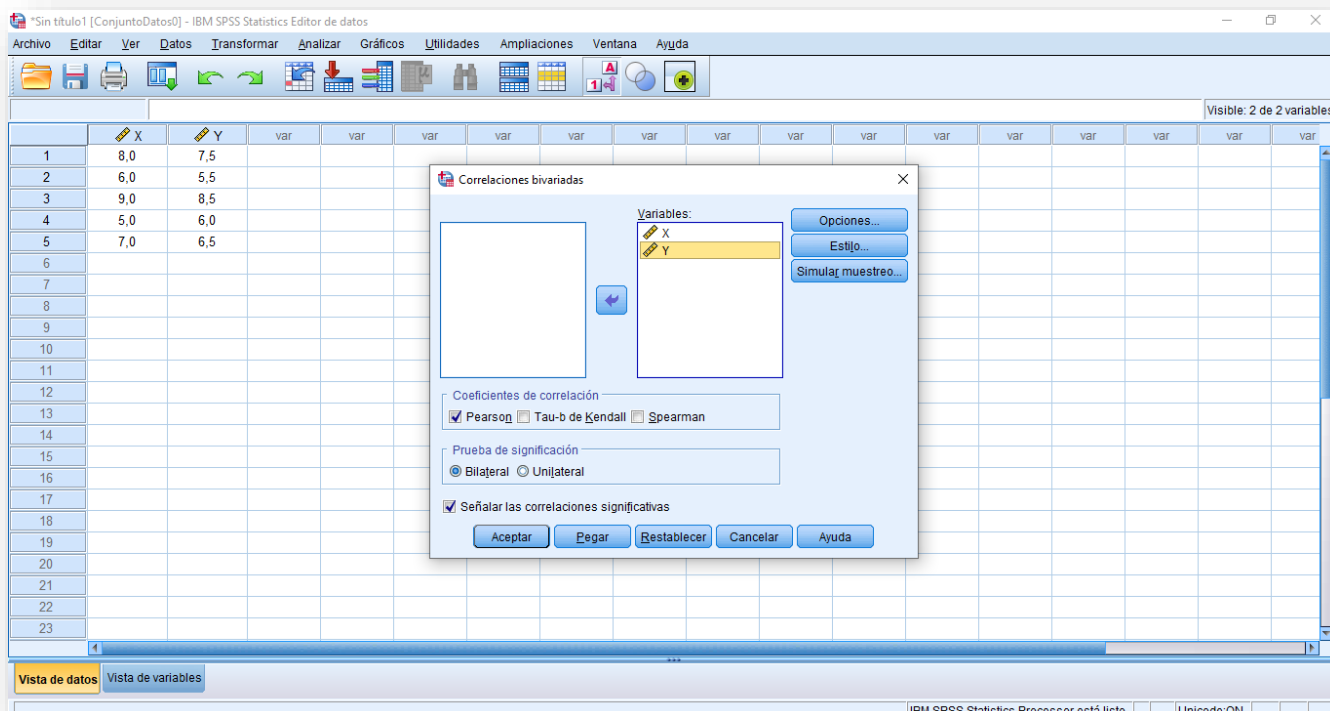
Clic en Analizar, Correlacionar, Bivariadas



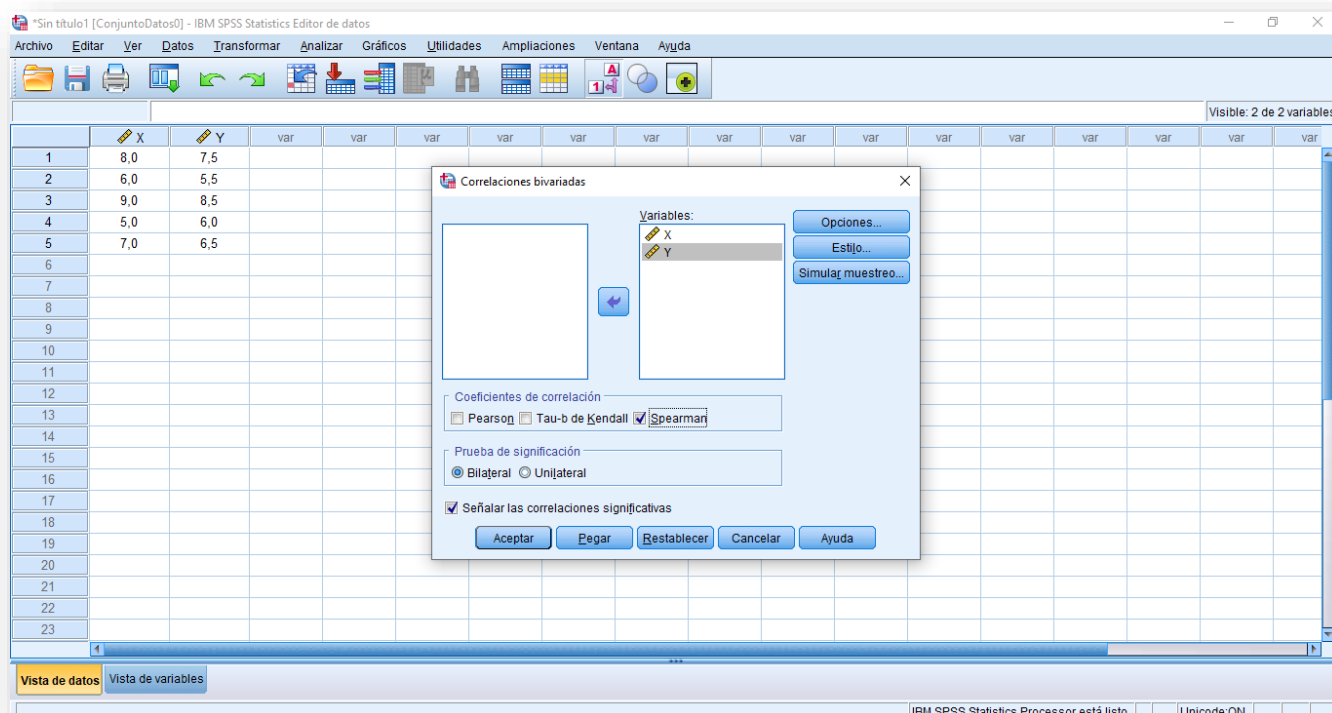
Clic en Bivariadas



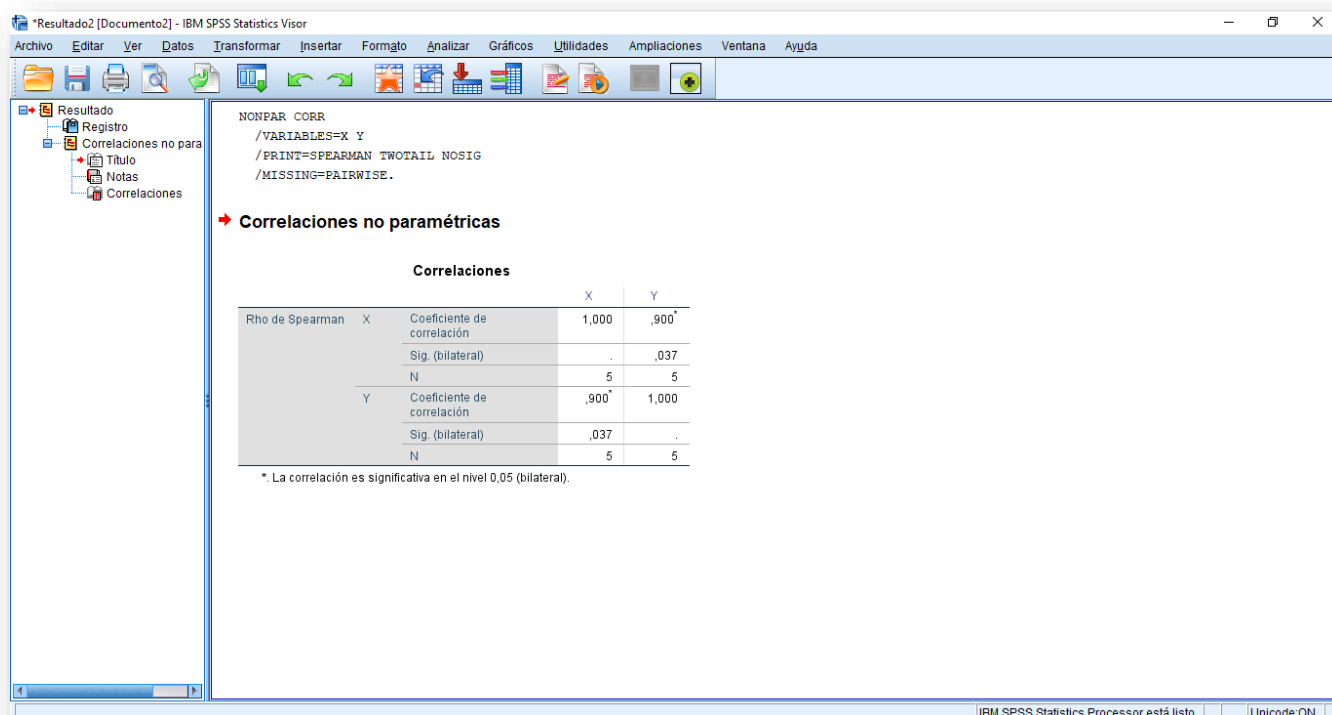
Pasar las variables X e Y a la casilla de Variables.



Clic en Pearson (para desactivar). Clic en Spearman (para activar)



Clic en Aceptar



Clic derecho en la tabla. Copiar y Pegar

Correlaciones

			X	Y
Rho de Spearman	X	Coefficiente de correlación	de 1,000	,900*
		Sig. (bilateral)	.	,037
		N	5	5
	Y	Coefficiente de correlación	de ,900*	1,000
		Sig. (bilateral)	,037	.
		N	5	5

*. La correlación es significativa en el nivel 0,05 (bilateral).

Interpretación:

Existe una correlación positiva fuerte ($\rho = 0,9$) entre las calificaciones de Matemática y Física.

2) Datos con empates: Calcular el coeficiente de Spearman sobre la relación entre horas de estudio y calificaciones empleando los siguientes datos:

Estudiante	Horas de estudio (X)	Calificaciones (Y)
A	3	5
B	3	6
C	5	7
D	6	8
E	6	8

Solución:

a) Se asigna rangos

Se ordena X e Y por separado y asignamos rangos (1 al más bajo):

Rangos para X:

Valores 3 (empate en posiciones 1 y 2):

$$\frac{1 + 2}{2} = 1,5$$

Valor 5: 3

Valores 6 (empate en posiciones 4 y 5):

$$\frac{4+5}{2} = 4,5$$

Rangos para Y:

5(1), 6(2), 7(3)

8 (empate en posiciones 4 y 5):

$$\frac{4+5}{2} = 4,5$$

b) Tabla con rangos

Estudiante	Horas de estudio (X)	Calificaciones (Y)	Rango X (R_x)	Rango Y (R_y)	$d_i = R_x - R_y$	d_i^2
A	3	5	1,5	1	0,5	0,25
B	3	6	1,5	2	-0,5	0,25
C	5	7	3	3	0	0
D	6	8	4,5	4,5	0	0
E	6	8	4,5	4,5	0	0
Total						0,5

c) Reemplazando valores en la fórmula

$$\rho = r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6(0,5)}{5(5^2 - 1)} = 0,975$$

Empleando SPSS

NONPAR CORR
 /VARIABLES=X Y
 /PRINT=SPEARMAN TWOTAIL NOSIG
 /MISSING=PAIRWISE.

→ **Correlaciones no paramétricas**

Correlaciones

			X	Y
Rho de Spearman	X	Coefficiente de correlación	1,000	,973**
		Sig. (bilateral)	.	,005
		N	5	5
		Y	Coefficiente de correlación	,973**
Sig. (bilateral)	,005		.	
N	5		5	

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

Correlaciones

			X	Y
Rho de Spearman	X	Coeficiente de correlación	de 1,000	,973**
		Sig. (bilateral)	.	,005
		N	5	5
	Y	Coeficiente de correlación	de ,973**	1,000
		Sig. (bilateral)	,005	.
		N	5	5

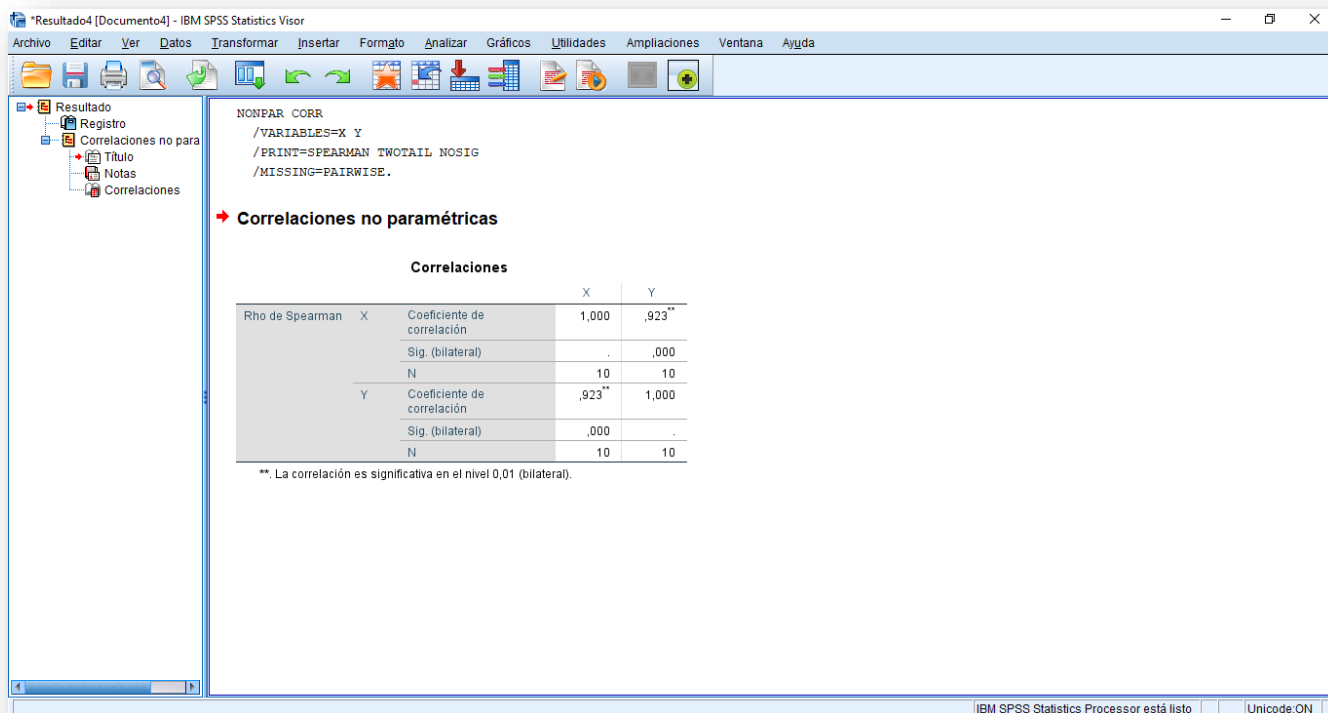
** . La correlación es significativa en el nivel 0,01 (bilateral).

Interpretación:

Existe una correlación positiva fuerte ($\rho = 0,97$) entre horas de estudio y calificaciones.

3) Calcular el coeficiente de Spearman sobre la relación entre horas de estudio y calificaciones empleando los siguientes datos:

Estudiante	Horas de estudio (X)	Calificación (Y)
1	2	6,0
2	4	7,5
3	6	9,1
4	8	9,2
5	6	8,5
6	4	7,5
7	6	9,3
8	7	9,4
9	10	9,6
10	8	9,5

Solución:**Empleando SPSS****Correlaciones**

			X	Y
Rho de Spearman	X	Coefficiente de correlación	1,000	,923**
		Sig. (bilateral)	.	,000
		N	10	10
	Y	Coefficiente de correlación	,923**	1,000
		Sig. (bilateral)	,000	.
		N	10	10

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

Interpretación:

Un coeficiente de correlación $\rho = 0,923$ indica una correlación positiva fuerte entre horas de estudio (X) y calificaciones (Y). El valor cercano a 1 significa que, cuando una variable aumenta, la otra tiende a

aumentar también. Sin embargo, correlación no implica causalidad: podría deberse a una relación real, a una variable oculta o incluso ser espuria.

La Significancia estadística (Sig. = 0,000) es menor que 0,01, lo que confirma que la correlación es estadísticamente significativa al nivel del 1% (muy baja probabilidad de que sea por azar).

Los asteriscos (**) refuerzan que es significativa al 99% de confianza.

Nota: La correlación de Spearman mide la relación monótona (no necesariamente lineal), útil cuando los datos no siguen una distribución normal o tienen outliers.

3.5) COEFICIENTE DE CORRELACIÓN DE TAU-B DE KENDALL

A) DEFINICIÓN

El Tau-b de Kendall (τ_b) es una medida no paramétrica de asociación monótonica entre dos variables ordinales o rangueadas. A diferencia de Pearson (lineal) y Spearman (monótona basada en rangos), Tau-b es especialmente útil cuando hay muchos empates (valores repetidos) en los datos. Su interpretación es similar a Spearman, pero con mejor ajuste para datos "repetidos".

Al ser una medida no paramétrica no depende de supuestos rígidos sobre la población, haciéndolo ideal para variables ordinales cuando los datos no cumplen los supuestos de normalidad y linealidad, en muestras pequeñas o con outliers influyentes.

B) CARACTERÍSTICAS

No paramétrico: No asume los supuestos de normalidad y linealidad en los datos

Rango basado: Usa la posición relativa (orden) de los datos

Mide correlación monótonica: Detecta la relación entre dos variables en la que la dirección (crecientes o decrecientes) de la asociación es consistente, pero no necesariamente lineal. Esto significa que:

- Si una variable aumenta, la otra tiende a aumentar (monótona creciente).
- Si una variable aumenta, la otra tiende a disminuir (monótona decreciente).

Rango de valores: Varía de -1 a 1

$\tau_b = 1$: Correlación monótonica perfecta positiva.

$\tau_b = -1$: Correlación monótonica perfecta negativa.

$\tau_b = 0$: No hay correlación monótonica.

Robusto ante outliers (valores atípicos): Es resistente a valores extremos porque trabaja con órdenes (rangos) en lugar de los valores originales. Esto lo hace ideal para datos con distribuciones no normales.

C) DIFERENCIAS CON SPERAMAN Y PEARSON

Aspecto	Tau-b de Kendall (τ_b)	Spearman (ρ)	Pearson (r)
Tipo de relación	Monotónica	Monotónica	Lineal
Datos ideales	Ordinales con empates	Ordinales	Continuos/Normales
Robustez a outliers	Alta	Alta	Baja
Interpretación	Más conservadora	Similar a Pearson	Directa

D) FÓRMULA

$$\tau_b = \frac{C - D}{\sqrt{(C + D + T_x)(C + D + T_y)}}$$

Donde:

C = Número de pares concordantes (mismo orden en X e Y).

D = Número de pares discordantes (orden opuesto en X e Y).

T_x = Empates en X; T_y = Empates en Y

E) EJEMPLOS DE APLICACIÓN

1) Calcular el coeficiente de correlación de Tau-b de Kendall entre Experiencia Docente y Satisfacción Laboral de la evaluación a 8 profesores en dos variables ordinales:

Experiencia (X): 1 = Baja (1-5 años); 2 = Media (6-10 años); 3 = Alta (11+ años)

Satisfacción (Y): 1 = Baja; 2 = Media; 3 = Alta

Los resultados se muestran en la siguiente tabla:

Profesor	Experiencia (X)	Satisfacción (Y)
1	1	1
2	2	2
3	3	3
4	1	1
5	2	2
6	3	2
7	2	3
8	3	3

Solución:

Clasificando cada par correctamente en Pares Concordantes (C) y Discordantes (D)

Pares Concordantes (C): Si un docente tiene X e Y mayores que otro.

Pares Discordantes (D): Si un docente tiene X mayor pero Y menor (o viceversa).

Simbología:

X aumenta: $X \uparrow$

Y aumenta: $Y \uparrow$

Mismo X: $X =$

Mismo Y: $Y =$

X disminuye: $X \downarrow$

Y disminuye: $Y \downarrow$

Comparaciones:

1) Profesor1 vs Profesor 2: Profesor 1 ($X=1, Y=1$); Profesor 2 ($X=2, Y=2$):

X aumenta ($X \uparrow$), Y aumenta ($Y \uparrow$) $\rightarrow$ Concordante (C)

2) Profesor1 vs Profesor 3: Profesor 1 ($X=1, Y=1$); Profesor 3 ($X=3, Y=3$):

$X \uparrow, Y \uparrow \rightarrow C$

3) Profesor1 vs Profesor 4: Profesor 1 ($X=1, Y=1$); Profesor 4 ($X=1, Y=1$):

Empate X ($X =$), Empate Y ($Y =$) $\rightarrow$ No cuenta

4) Profesor1 vs Profesor 5: Profesor 1 ($X=1, Y=1$); Profesor 5 ($X=2, Y=2$):

$X \uparrow, Y \uparrow \rightarrow C$

5) Profesor1 vs Profesor 6: Profesor 1 ($X=1, Y=1$); Profesor 6 ($X=3, Y=2$):

$X \uparrow, Y \uparrow \rightarrow C$

6) Profesor1 vs Profesor 7: Profesor 1 ($X=1, Y=1$); Profesor 7 ($X=2, Y=3$):

$X \uparrow, Y \uparrow \rightarrow C$

7) Profesor1 vs Profesor 8: Profesor 1 ($X=1, Y=1$); Profesor 8 ($X=3, Y=3$):

$X \uparrow, Y \uparrow \rightarrow C$

8) Profesor 2 vs Profesor 3: Profesor 2 ($X=2, Y=2$); Profesor 3 ($X=3, Y=3$):

$X \uparrow, Y \uparrow \rightarrow C$

9) Profesor 2 vs Profesor 4: Profesor 2 ($X=2, Y=2$); Profesor 4 ($X=1, Y=1$):

$X \downarrow, Y \downarrow \rightarrow C$

- 10) Profesor 2 vs Profesor 5: Profesor 2 ($X=2, Y=2$); Profesor 5 ($X=2, Y=2$)
 $X=, Y= \rightarrow$ **No cuenta**
- 11) Profesor 2 vs Profesor 6: Profesor 2 ($X=2, Y=2$); Profesor 6 ($X=3, Y=2$):
 $X\uparrow, Y= \rightarrow$ **Empate Y**
- 12) Profesor 2 vs Profesor 7: Profesor 2 ($X=2, Y=2$); Profesor 7 ($X=2, Y=3$)
 $X=, Y\uparrow \rightarrow$ **Empate X**
- 13) Profesor 2 vs Profesor 8: Profesor 2 ($X=2, Y=2$); Profesor 8 ($X=3, Y=3$):
 $X\uparrow, Y\uparrow \rightarrow$ **C**
- 14) Profesor 3 vs Profesor 4: Profesor 3 ($X=3, Y=3$); Profesor 4 ($X=1, Y=1$):
 $X\downarrow, Y\downarrow \rightarrow$ **C**
- 15) Profesor 3 vs Profesor 5: Profesor 3 ($X=3, Y=3$); Profesor 5 ($X=2, Y=2$):
 $X\downarrow, Y\downarrow \rightarrow$ **C**
- 16) Profesor 3 vs Profesor 6: Profesor 3 ($X=3, Y=3$); Profesor 6 ($X=3, Y=2$):
 $X=, Y\downarrow \rightarrow$ **Empate X**
- 17) Profesor 3 vs Profesor 7: Profesor 3 ($X=3, Y=3$); Profesor 7 ($X=2, Y=3$):
 $X\downarrow, Y= \rightarrow$ **Empate Y**
- 18) Profesor 3 vs Profesor 8: Profesor 3 ($X=3, Y=3$); Profesor 8 ($X=3, Y=3$):
 $X=, Y= \rightarrow$ **No cuenta**
- 19) Profesor 4 vs Profesor 5: Profesor 4 ($X=1, Y=1$); Profesor 5 ($X=2, Y=2$):
 $X\uparrow, Y\uparrow \rightarrow$ **C**
- 20) Profesor 4 vs Profesor 6: Profesor 4 ($X=1, Y=1$); Profesor 6 ($X=3, Y=2$):
 $X\uparrow, Y\uparrow \rightarrow$ **C**
- 21) Profesor 4 vs Profesor 7: Profesor 4 ($X=1, Y=1$); Profesor 7 ($X=2, Y=3$):
 $X\uparrow, Y\uparrow \rightarrow$ **C**
- 22) Profesor 4 vs Profesor 8: Profesor 4 ($X=1, Y=1$); Profesor 8 ($X=3, Y=3$):
 $X\uparrow, Y\uparrow \rightarrow$ **C**
- 23) Profesor 5 vs Profesor 6: Profesor 5 ($X=2, Y=2$); Profesor 6 ($X=3, Y=2$):
 $X\uparrow, Y= \rightarrow$ **Empate Y**
- 24) Profesor 5 vs Profesor 7: Profesor 5 ($X=2, Y=2$); Profesor 7 ($X=2, Y=3$):
 $X=, Y\uparrow \rightarrow$ **Empate X**

25) Profesor 5 vs Profesor 8: Profesor 5 (X=2, Y=2); Profesor 8 (X=3, Y=3):

X↑, Y↑ → C

26) Profesor 6 vs Profesor 7: Profesor 6 (X=3, Y=2); Profesor 7 (X=2, Y=3):

X↓, Y↑ → D

27) Profesor 6 vs Profesor 8: Profesor 6 (X=3, Y=2); Profesor 8 (X=3, Y=3):

X=, Y↑ → Empate X

28) Profesor 7 vs Profesor 8: Profesor 7 (X=2, Y=3); Profesor 8 (X=3, Y=3):

X↑, Y= → Empate Y

Conteo Final:

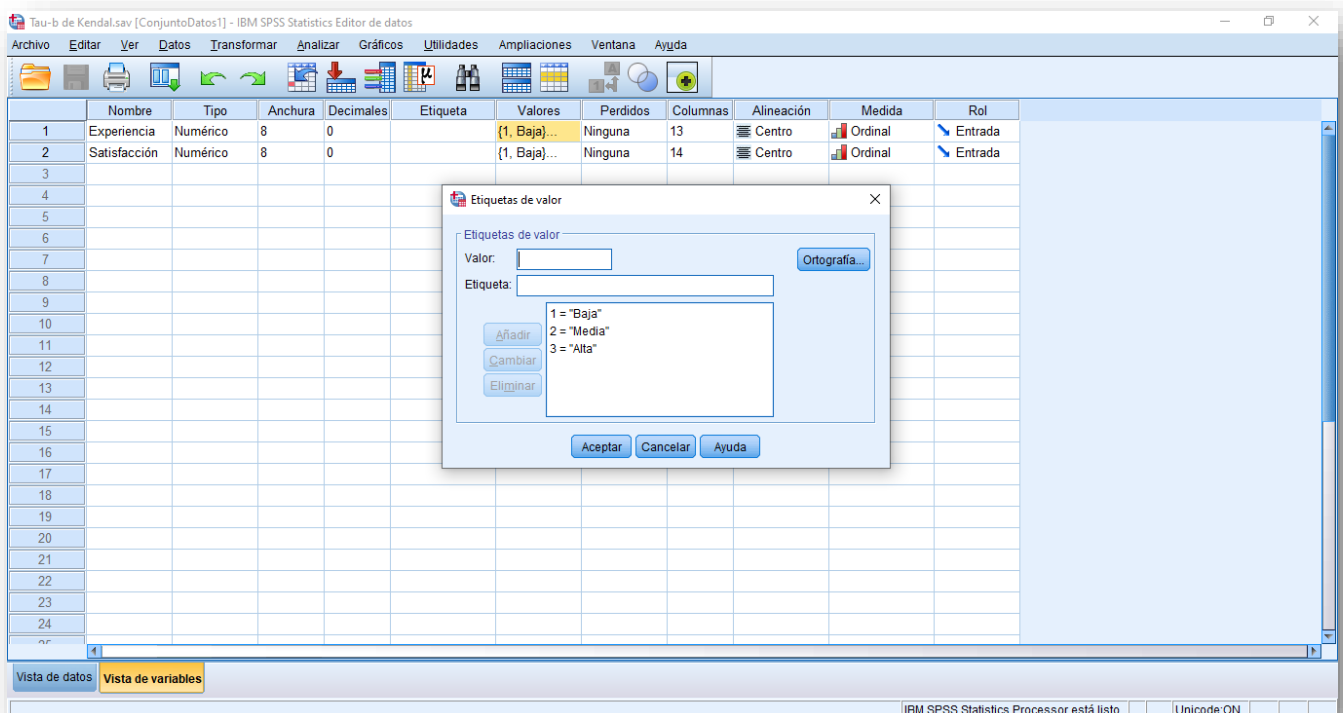
Pares Concordantes (C): 16; Pares Discordantes (D): 1; Empates X (T_x): 4; Empates Y (T_y): 4

Reemplazando valores en la fórmula

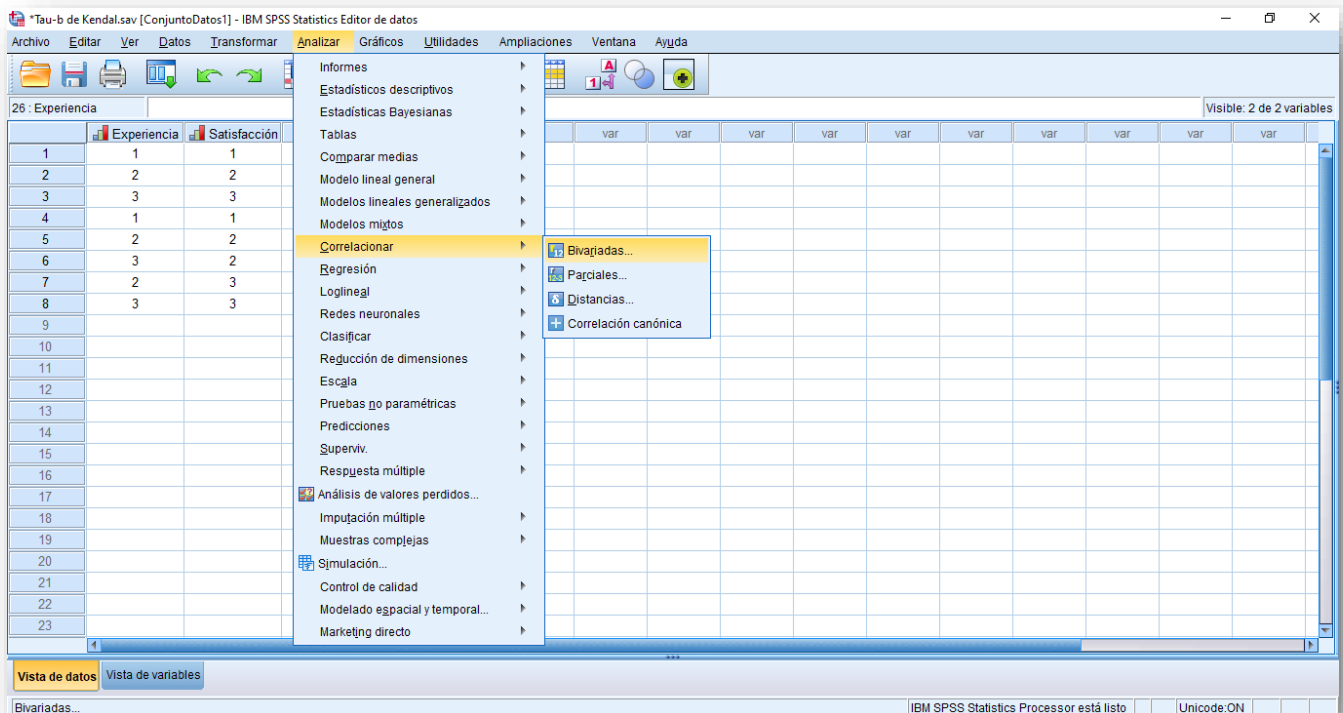
$$\tau_b = \frac{C - D}{\sqrt{(C + D + T_x)(C + D + T_y)}} = \frac{16 - 1}{\sqrt{(16 + 1 + 4)(16 + 1 + 4)}} = \frac{15}{21} = 0,714$$

Empleando SPSS

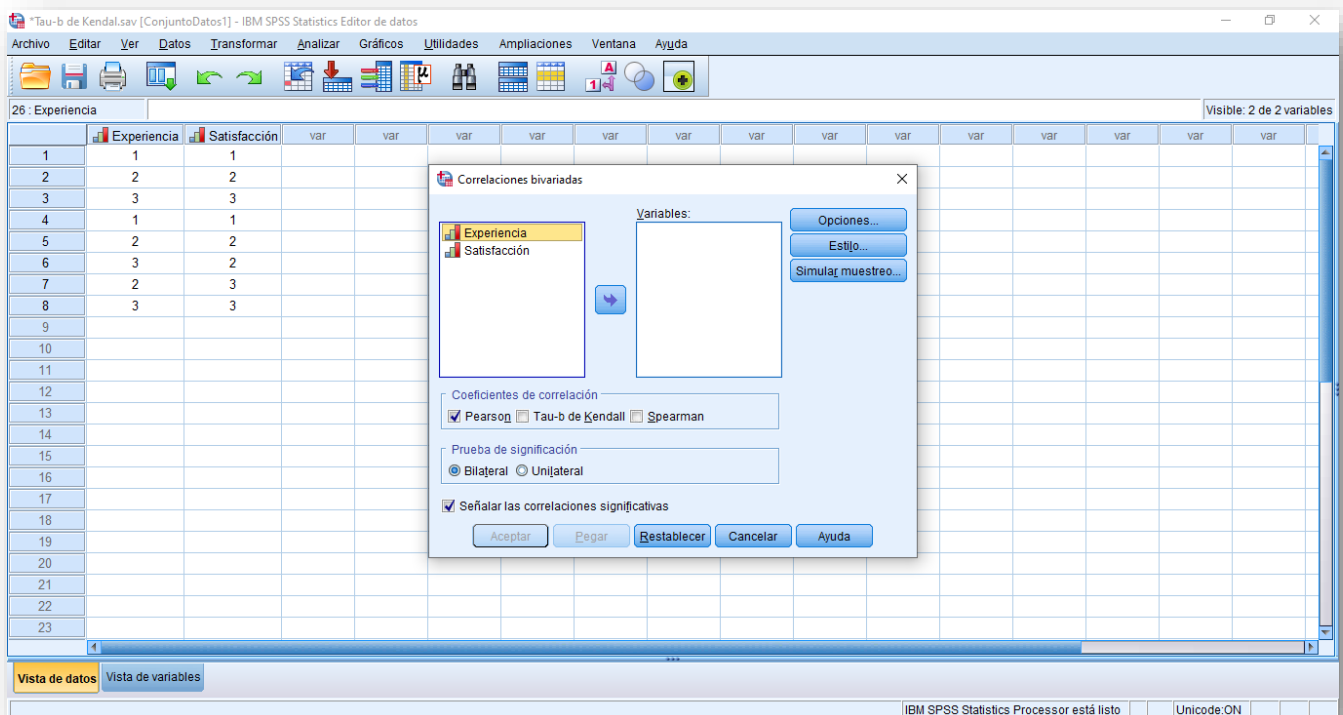
Ingresando los datos en SPSS



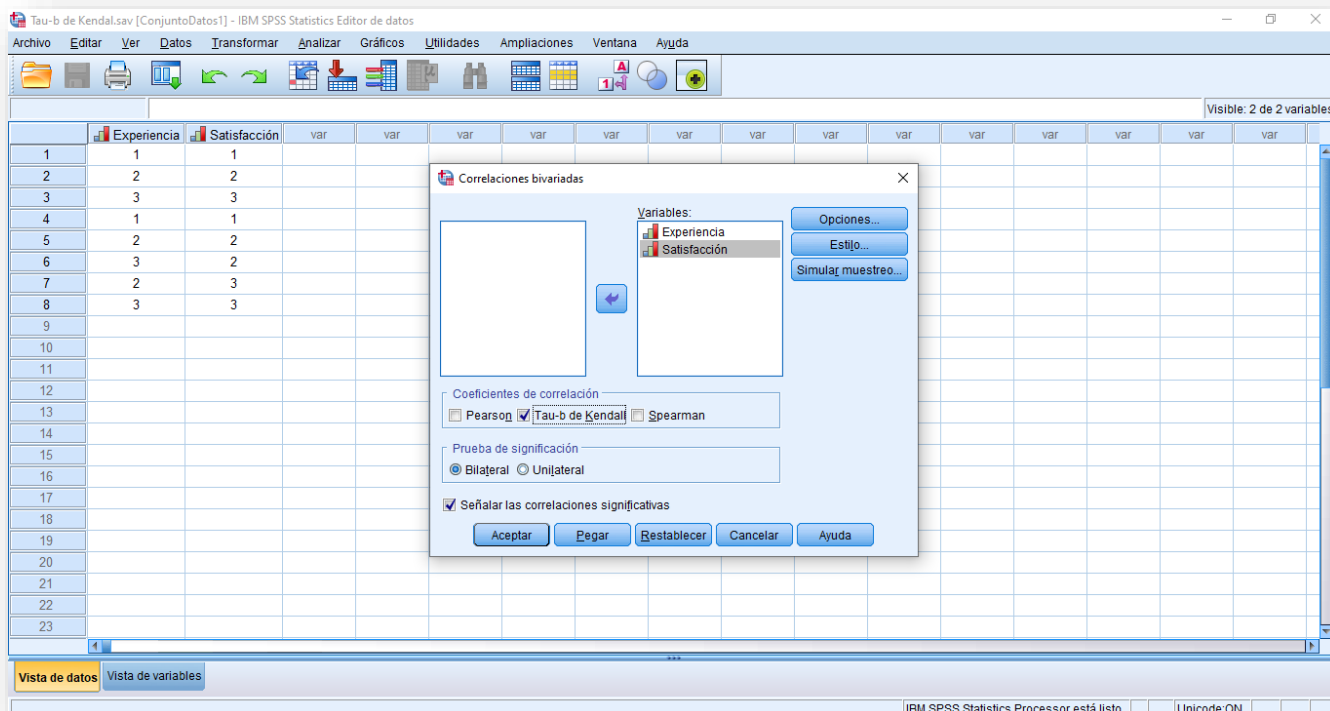
Clic en Analizar, Correlacionar, Bivariadas



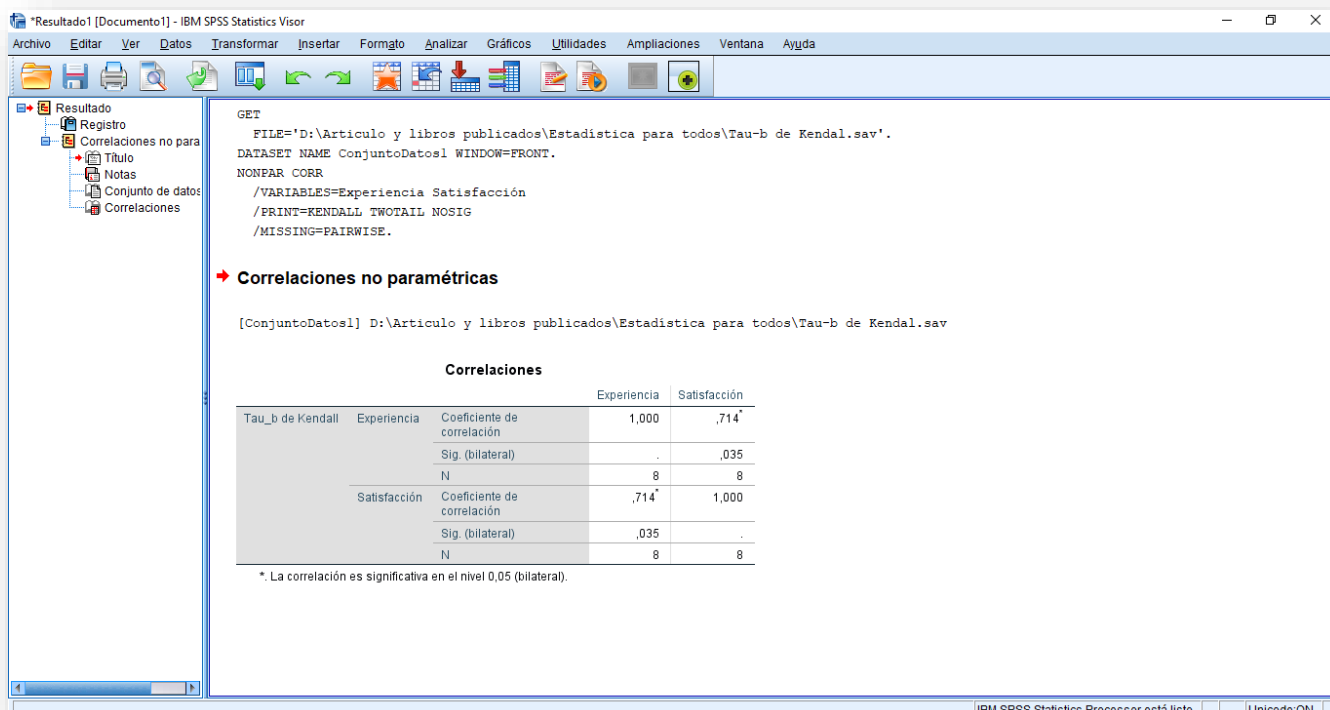
Clic en Bivariadas



Pasar las variables de Experiencia y Satisfacción a la casilla de Variables. Clic en Pearson (para desactivar). Clic en Tau-b de Kendall (para activar)



Clic en Aceptar



Clic derecho en la tabla. Copiar. Pegar

Correlaciones

			Experiencia	Satisfacción
Tau_b de Kendall	Experiencia	Coefficiente de correlación	1,000	,714*
		Sig. (bilateral)	.	,035
		N	8	8
	Satisfacción	Coefficiente de correlación	,714*	1,000
		Sig. (bilateral)	,035	.
		N	8	8

*. La correlación es significativa en el nivel 0,05 (bilateral).

Interpretación:

El valor $\tau_b = 0,714$ indica una correlación positiva fuerte entre Experiencia Docente y Satisfacción Laboral. Esto significa que, en estos datos, los profesores con más experiencia tienden a tener mayor satisfacción laboral, y viceversa.

El p-valor = 0,035 es menor que 0,05, lo que indica que la correlación es estadísticamente significativa (al nivel del 5%). Por lo tanto, se concluye que existe una asociación significativa entre ambas variables.

2) Medicina: Relación entre nivel de dolor (X) y dosis de analgésico (Y). Evaluar si existe correlación entre el nivel de dolor reportado por pacientes (en escala ordinal: 1 = leve, 2 = moderado, 3 = severo) y la dosis de analgésico administrada (1 = baja, 2 = media, 3 = alta).

Paciente	Nivel de dolor (X)	Dosis de analgésico (Y)
1	1	1
2	1	2
3	2	2
4	2	3
5	3	3

Solución:

Empleando SPSS

Correlaciones

			X	Y
Tau_b de Kendall	X	Coefficiente de correlación	1,000	,750
		Sig. (bilateral)	.	,097
		N	5	5
	Y	Coefficiente de correlación	,750	1,000
		Sig. (bilateral)	,097	.
		N	5	5

Interpretación:

El valor $\tau_b = 0,75$ indica una correlación positiva fuerte entre el nivel de dolor y la dosis de analgésico. A mayor nivel de dolor reportado, mayor tendencia a recibir una dosis más alta de analgésico (y viceversa).

El p-valor = 0,097 es mayor que 0,05, lo que sugiere que la correlación no es estadísticamente significativa al nivel tradicional del 5%. Aunque existe una asociación fuerte en los datos, el pequeño tamaño de la muestra ($n = 5$) limita la potencia estadística para confirmar que esta relación no sea producto del azar. En estudios con muestras más grandes, este resultado podría alcanzar significancia.

3) Psicología: Relación entre estrés laboral (X) y satisfacción (Y). Analizar si hay asociación entre el nivel de estrés laboral (1 = bajo, 2 = medio, 3 = alto) y la satisfacción laboral (1 = insatisfecho, 2 = neutral, 3 = satisfecho).

Empleado	Estrés (X)	Satisfacción (Y)
1	1	3
2	2	2
3	2	1
4	3	1
5	3	2
6	1	3
7	2	2
8	2	1
9	3	1
10	3	2
11	1	3
12	2	2
13	2	1
14	3	1
15	3	2
16	1	3
17	2	2
18	2	1
19	3	1
20	3	2

Solución:

Empleando SPSS

Correlaciones

		X	Y
Tau_b de Kendall	X	Coefficiente de correlación	1,000
		Sig. (bilateral)	,016
		N	20
	Y	Coefficiente de correlación	-,500*
		Sig. (bilateral)	,016
		N	20

*. La correlación es significativa en el nivel 0,05 (bilateral).

Interpretación:

El valor $\tau_b = -0,5$ indica una correlación negativa moderada entre el estrés laboral y la satisfacción laboral. A mayor nivel de estrés laboral (X), menor satisfacción laboral (Y), y viceversa.

El p-valor = 0,016 es menor que 0,05, lo que indica que la correlación es estadísticamente significativa al nivel del 5%. La asociación observada no es producto del azar; existe evidencia sólida para afirmar que hay una relación inversa entre estrés y satisfacción laboral en esta muestra. El asterisco (*) en la tabla de SPSS confirma que esta correlación es significativa ($p < 0.05$).

3.6) ACTIVIDADES DE APLICACIÓN

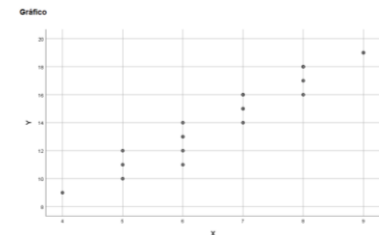
- 1) Realice un organizador gráfico del capítulo II
- 2) Con los siguientes datos de Ciencias de la Salud. X: Horas de sueño por noche (promedio semanal).
Y: Promedio académico (escala 0-20).

Estudiante	X	Y
1	7	14
2	7	16
3	5	12
4	8	18
5	6	13
6	8	17
7	5	10
8	9	19
9	6	11
10	7	15
11	6	14
12	8	16
13	6	12
14	8	18
15	7	15
16	4	9
17	7	16
18	5	11
19	9	19
20	6	13

a) Calcule la asimetría y la curtosis. Interpretar los resultados

Estadísticos			
		X	Y
N	Válido	20	20
	Perdidos	0	0
Asimetría		-,064	-,079
Error estándar de asimetría		,512	,512
Curtosis		-,617	-,979
Error estándar de curtosis		,992	,992

b) Realice un diagrama de dispersión e interprete el resultado.



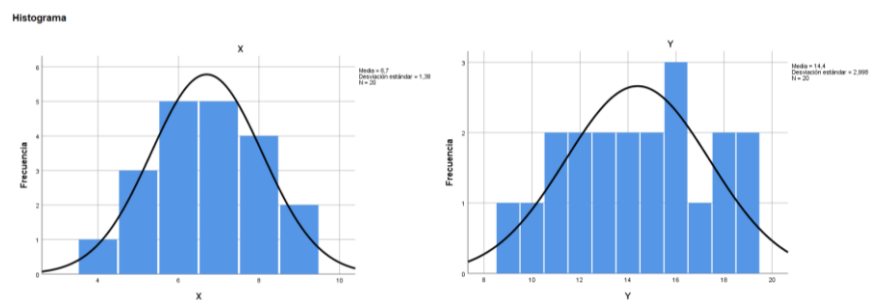
c) Realice la prueba de normalidad e interprete los resultados.

Pruebas de normalidad						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
X	,144	20	,200 [*]	,950	20	,372
Y	,103	20	,200 [*]	,964	20	,619

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

d) Realice un histograma con una curva normal para X y para Y. Interprete los resultados.



d) Calcule el coeficiente de correlación de Pearson e interprete los resultados

Correlaciones			
		X	Y
X	Correlación de Pearson	1	,959**
	Sig. (bilateral)		,000
	N	20	20
Y	Correlación de Pearson	,959**	1
	Sig. (bilateral)	,000	
	N	20	20

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

3) Con los siguientes datos de Psicología. X: Horas diarias en redes sociales. Y: Puntaje de ansiedad (escala GAD-7, 0-21).

Participante	X	Y
1	1,0	5
2	2,5	8
3	3,0	10
4	4,5	14
5	2,0	7
6	5,0	16
7	1,5	6
8	6,0	18
9	3,5	12
10	4,0	13
11	0,5	4
12	5,5	17
13	2,5	9
14	4,5	15
15	3,0	11

Nota:

Escala GAD-7 (Generalized Anxiety Disorder 7-item Scale)

La GAD-7 es una herramienta de autoevaluación ampliamente utilizada en psicología y psiquiatría para medir la severidad de los síntomas de ansiedad generalizada. Fue diseñada para ser breve, fácil de aplicar y con alta validez clínica.

Características Clave:

7 ítems: Evalúa la frecuencia de síntomas de ansiedad en las últimas 2 semanas.

Puntuación total: 0–21 puntos, donde:

0–4: Ansiedad mínima o ausente.

5–9: Ansiedad leve.

10–14: Ansiedad moderada.

15–21: Ansiedad severa.

Uso clínico:

Detección rápida de trastorno de ansiedad generalizada (TAG).

Monitoreo de la respuesta al tratamiento.

Ejemplo de Ítems (Preguntas):

Cada ítem se puntúa del 0 al 3 según la frecuencia: 0 = Nunca; 1 = Varios días; 2 = Más de la mitad de los días; 3 = Casi todos los días.

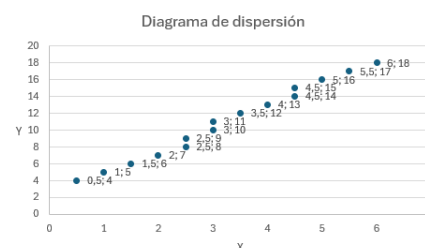
1. "Me siento nervioso/a, ansioso/a o con los nervios de punta."
2. "No puedo controlar mis preocupaciones."
3. "Me preocupo demasiado por diferentes cosas."
4. "Tengo dificultad para relajarme."
5. "Estoy tan inquieto/a que no puedo quedarme quieto/a."
6. "Me irrito o molesto con facilidad."
7. "Siento miedo como si algo terrible fuera a pasar."

Interpretación de la Puntuación (0–21):

Puntuación	Nivel de Ansiedad	Recomendación
0–4	Mínima	No requiere intervención.
5–9	Leve	Vigilancia o psicoeducación.
10–14	Moderada	Considerar terapia psicológica.
15–21	Severa	Evaluación por psiquiatría + terapia.

Aunque es útil, no sustituye un diagnóstico clínico. Siempre debe ser interpretada por un profesional de la salud mental.

a) Realice un diagrama de dispersión e interprete el resultado.



b) Realice la prueba de normalidad e interprete los resultados.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
X	,106	15	,200 [*]	,977	15	,944
Y	,082	15	,200 [*]	,964	15	,755

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

b) Calcule el coeficiente de correlación de Pearson e interprete los resultados

Correlaciones

		X	Y
X	Correlación de Pearson	1	,995**
	Sig. (bilateral)		,000
	N	15	15
Y	Correlación de Pearson	,995**	1
	Sig. (bilateral)	,000	
	N	15	15

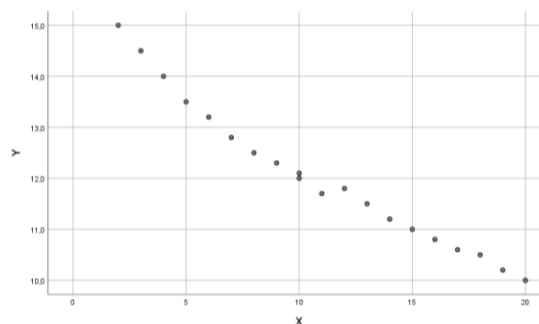
**.. La correlación es significativa en el nivel 0,01 (bilateral).

	A	B	C	D	E
1	1	5			
2	2,5	8			
3	3	10			
4	4,5	14			
5	2	7			
6	5	16			
7	1,5	6			
8	6	18			
9	3,5	12			
10	4	13			
11	0,5	4			
12	5,5	17			
13	2,5	9			
14	4,5	15			
15	3	11			
16					
17	r	0,99456769	=COEF.DE.CORREL(A1:A15;B1:B15)		

4) Con los siguientes datos de Deportes. X: Horas de entrenamiento semanal. Y: Tiempo en 100 m planos (segundos).

Atleta	X	Y
1	8	12,5
2	12	11,8
3	5	13,5
4	15	11,0
5	10	12,0
6	14	11,2
7	4	14,0
8	18	10,5
9	9	12,3
10	13	11,5
11	3	14,5
12	16	10,8
13	6	13,2
14	17	10,6
15	11	11,7
16	2	15,0
17	20	10,0
18	7	12,8
19	19	10,2
20	10	12,1

a) Realice un diagrama de dispersión e interprete el resultado.



b) Realice la prueba de normalidad e interprete los resultados.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
X	,072	20	,200*	,967	20	,687
Y	,089	20	,200*	,961	20	,560

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

c) Calcule el coeficiente de correlación de Pearson e interprete los resultados

Correlaciones

		X	Y
X	Correlación de Pearson	1	-,983**
	Sig. (bilateral)		,000
	N	20	20
Y	Correlación de Pearson	-,983**	1
	Sig. (bilateral)	,000	
	N	20	20

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

5) Descargue la base de datos de la Prueba Ser Estudiante 2023-2024 (Micro SAV) disponible en el portal oficial: <https://evaluaciones.evaluacion.gob.ec/BI/ser-estudiante-2/>

Año lectivo 2023-2024						
Archivo	Sintaxis	CSV	SAV	XLSX	Metadato	Diccionario
Micro	↓	↓	↓	↓	↓	↓
Factores Asociados estudiantes	↓	↓	↓	↓	↓	↓
Factores Asociados padres	↓	↓	↓	↓	↓	↓
Factores Asociados docentes	↓	↓	↓	↓	↓	↓
Factores Asociados directivos	↓	↓	↓	↓	↓	↓

a) Realice la prueba de normalidad para imat y para ilyl. Interprete los resultados.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a		
	Estadístico	gl	Sig.
Promedio obtenido en el campo de Matemática	,156	47643	,000
Promedio obtenido en el campo de Lengua y Literatura	,055	47643	,000

a. Corrección de significación de Lilliefors

b) Con los datos para imat y para ilyl es correcto calcular el coeficiente de correlación de Pearson. Justifique su respuesta.

c) Calcule el coeficiente de correlación de Spearman. Interprete los resultados.

Correlaciones

			Promedio obtenido en el campo de Matemática	Promedio obtenido en el campo de Lengua y Literatura
Rho de Spearman	Promedio obtenido en el campo de Matemática	Coeficiente de correlación	1,000	,414**
		Sig. (bilateral)		,000
		N	48297	47643
	Promedio obtenido en el campo de Lengua y Literatura	Coeficiente de correlación	,414**	1,000
		Sig. (bilateral)	,000	
		N	47643	48330

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

6) Calcular el coeficiente de correlación de Spearman e interpretar el resultado para la Contaminación vs. hospitalizaciones por asma en ciudades, siendo:

X: Nivel de contaminación

1= Baja

2= Moderada

3 = Alta

Y: Categorías de hospitalizaciones

1 = 0 a 5 casos

2= 6 a 10 casos

3= 11 a más casos

Ciudad	Contaminación (X)	Hospitalizaciones (Y)
A	1	1
B	1	1
C	2	2
D	2	2
E	3	3
F	3	3
G	2	2
H	3	3
I	1	1
J	3	3

Correlaciones				
Rho de Spearman	X		X	Y
		Coeficiente de correlación	1,000	1,000**
		Sig. (bilateral)	.	.
		N	10	10
	Y	Coeficiente de correlación	1,000**	1,000
		Sig. (bilateral)	.	.
		N	10	10

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

7) Calcular el coeficiente de correlación de Spearman e interpretar el resultado para el Ejercicio vs. bienestar, siendo:

X: Días de ejercicio a la semana

Y: Bienestar en la escala 1 a 10. Siendo 1 lo más bajo y 10 lo más alto.

Paciente	Días ejercicio (X)	Bienestar (Y)
1	0	3
2	1	5
3	2	6
4	3	7
5	4	8
6	5	9
7	0	2
8	3	7
9	1	4
10	5	9

Correlaciones				
Rho de Spearman	X		X	Y
		Coefficiente de correlación	1,000	,994**
		Sig. (bilateral)	.	,000
		N	10	10
	Y	Coefficiente de correlación	,994**	1,000
		Sig. (bilateral)	,000	.
		N	10	10

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

8) Calcular el coeficiente de correlación de Spearman e interpretar el resultado para Ranking de marcas vs. lealtad, siendo:

X: Posición en ranking: 1 (mejor) - 10 (Peor)

Y: Lealtad de recompra: 1 (mínimo) - 5 (máximo)

Consumidor	Ranking marca (X)	Recompra (Y)
1	1	5
2	2	4
3	3	4
4	4	3
5	5	3
6	6	2
7	7	2
8	8	1
9	9	1
10	10	1

Correlaciones				
Rho de Spearman	X		X	Y
		Coefficiente de correlación	1,000	-,979**
		Sig. (bilateral)	.	,000
		N	10	10
	Y	Coefficiente de correlación	-,979**	1,000
		Sig. (bilateral)	,000	.
		N	10	10

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

9) Calcular el coeficiente de correlación de Tau-b de Kendall e interpretar el resultado de la relación entre nivel de depresión y adherencia a terapia, siendo:

X: Depresión: 1 (Leve), 2 (Moderada), 3 (Severa)

Y: Adherencia: 1 (Baja), 2 (Media), 3 (Alta)

Paciente	Depresión (X)	Adherencia (Y)
1	1	3
2	2	2
3	1	2
4	2	1
5	3	1
6	2	3
7	1	3
8	3	1
9	2	2
10	3	1

Correlaciones				
			X	Y
Tau_b de Kendall	X	Coefficiente de correlación	1,000	-,727*
		Sig. (bilateral)	.	,015
		N	10	10
	Y	Coefficiente de correlación	-,727*	1,000
		Sig. (bilateral)	,015	.
		N	10	10

*. La correlación es significativa en el nivel 0,05 (bilateral).

10) Calcular el coeficiente de correlación de Tau-b de Kendall e interpretar el resultado de la relación entre ideología política vs. apoyo a políticas ambientales, siendo:

Ideología (X): 1=Izquierda, 2=Centro, 3=Derecha

Apoyo ambiental (Y): 1=Bajo, 3=Alto

Encuestado	Ideología (X)	Apoyo ambiental (Y)
1	1	3
2	1	3
3	2	2
4	2	1
5	3	1
6	3	1
7	1	3
8	2	2
9	3	1
10	2	2

Correlaciones				
			X	Y
Tau_b de Kendall	X	Coefficiente de correlación	1,000	-,909**
		Sig. (bilateral)	.	,002
		N	10	10
	Y	Coefficiente de correlación	-,909**	1,000
		Sig. (bilateral)	,002	.
		N	10	10

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

FORMULARIO

Frecuencia Relativa ($fr = h$)

$$fr = h = \frac{f}{n}$$

Donde:

f = frecuencia absoluta

n = es el total de observaciones

Puede expresarse como fracción, decimal o porcentaje.

Frecuencia Porcentual ($f\% = p$)

Llamada también frecuencia relativa porcentual, es decir, es la frecuencia relativa expresada en porcentaje.

$$f\% = fr \cdot 100\% = p = h \cdot 100\%$$

Donde:

$fr = h$ = frecuencia relativa

Frecuencia Relativa Acumulada ($fra = H$)

$$H = \frac{F}{n}$$

Donde:

F = frecuencia absoluta acumulada. Es la suma de las frecuencias absolutas de todos los valores menores o iguales a un valor dado, es decir, es la suma de la frecuencia absoluta primera con la segunda, este valor con la tercera, y así sucesivamente.

n = es el total de observaciones

Frecuencia Relativa Acumulada Porcentual ($fra\%$) o Frecuencia Porcentual Acumulada (P)

Se obtiene multiplicando la frecuencia relativa acumulada por 100%. Se calcula así:

$$fra\% = fra \cdot 100\% = P = H \cdot 100\%$$

Tamaño de la muestra para proporciones en poblaciones infinitas o desconocidas. Cuando el tamaño de la población es muy grande o desconocido, se utiliza la siguiente fórmula:

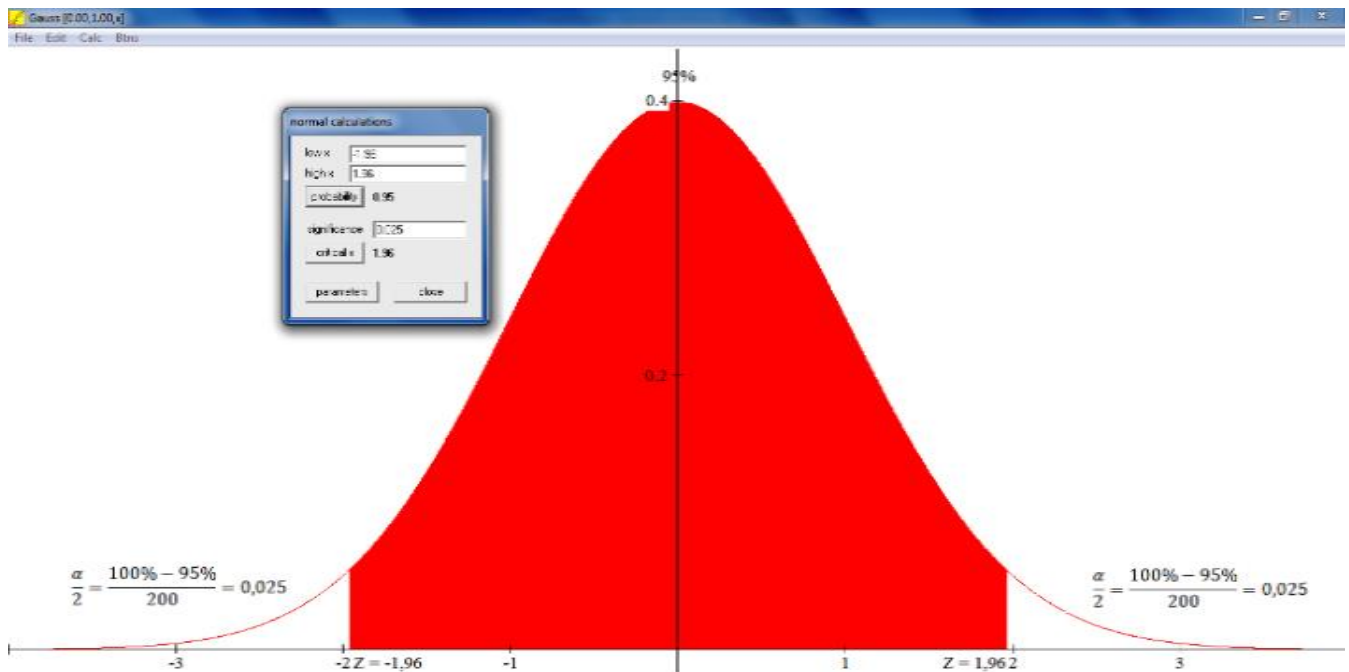
$$n = \frac{pqZ^2}{e^2}$$

Donde:

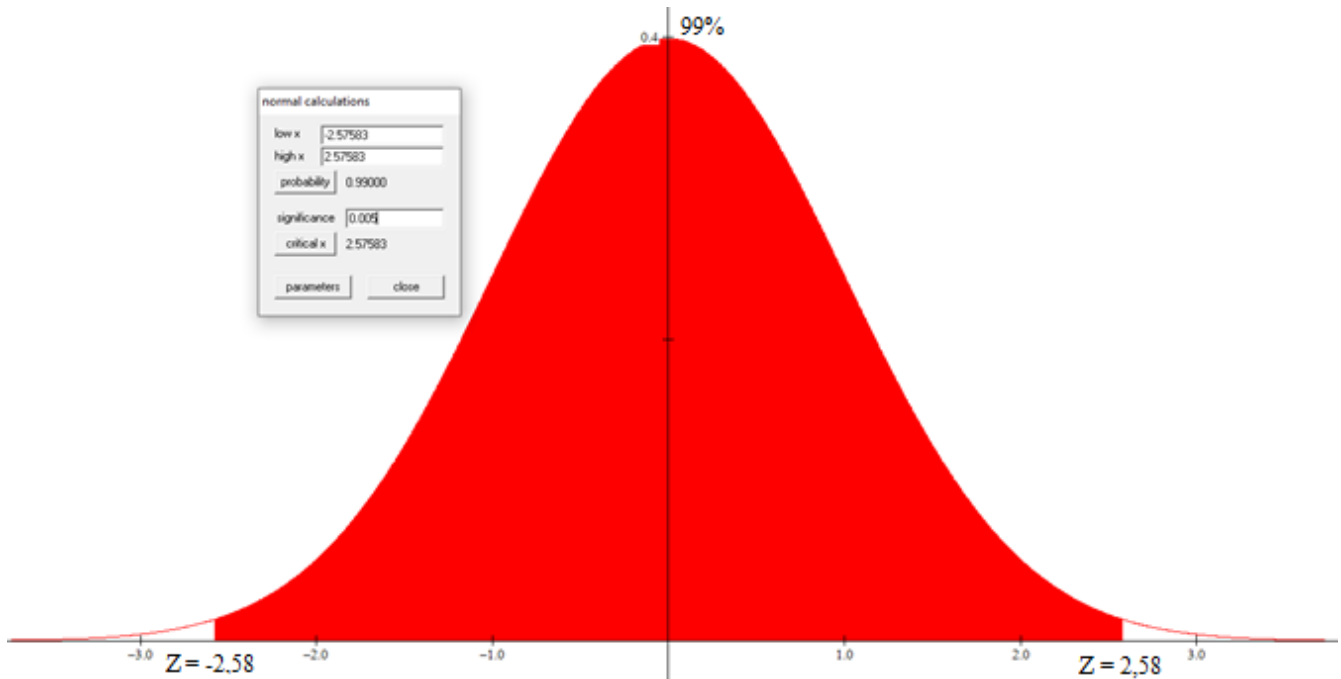
n : Tamaño de la muestra.

Z : Valor crítico de la distribución normal estándar, que depende del nivel de confianza, así por ejemplo:

$Z = 1,96$ para un nivel de confianza del 95%



$Z = 2,58$ para un nivel de confianza del 99%



p : Proporción estimada de la población que posee la característica de interés. Si no se conoce, se usa $p = 0,5$ para maximizar el tamaño de la muestra.

q : Complemento de p , es decir, $q = 1 - p$.

e : Margen de error expresado en proporción, límite aceptable de error muestral o error máximo admisible (por ejemplo, 0,05 para un margen de error del 5%).

Tamaño de la muestra para proporciones en poblaciones finitas

$$n = \frac{NpqZ^2}{(N - 1)e^2 + pqZ^2}$$

Donde:

N : Tamaño de la población.

Los demás términos tienen el mismo significado de la fórmula anterior.

Tamaño de la muestra para medias en poblaciones infinitas

$$n = \frac{\sigma^2 Z^2}{e^2}$$

Donde:

σ : Desviación estándar de la población (si no se conoce, se puede estimar a partir de estudios previos o una muestra piloto).

e : Margen de error expresado en las mismas unidades que la media.

Tamaño de la muestra para medias en poblaciones finitas

$$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2}$$

Media aritmética simple

a) Para datos sin tabular

La media de una población es el parámetro μ (que se lee “miu”). Si hay N observaciones en el conjunto de datos de la población, la media se calcula así:

$$\mu = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N} = \frac{\sum x_i}{N}$$

La media de una muestra es un estadístico $\bar{x}$ (que se lee “x barra”). Con n observaciones en el conjunto de datos de la muestra ($x_1, x_2, \dots$), la media se determina así:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

b) Para datos ordenados en tablas de frecuencias

$$\bar{x} = \frac{\sum f \cdot x_i}{n}$$

Donde:

f = frecuencia absoluta

x_i = Observaciones individuales de los datos

n = número total de datos

c) Para datos agrupados en intervalos

$$\bar{x} = \frac{\sum f \cdot xm}{n}$$

Donde:

f = frecuencia absoluta

xm = marca de clase

n = número total de datos

Media aritmética ponderada

$$\bar{x} = \frac{p_1 \cdot x_1 + p_2 \cdot x_2 + p_3 \cdot x_3 + \cdots p_k \cdot x_k}{p_1 + p_2 + p_3 + \cdots p_k} = \frac{\sum p \cdot x_i}{\sum p}$$

Donde:

p = peso o ponderación

x_i = Observaciones individuales de los datos

La Mediana**a) Para datos no agrupados**

Se aplica la fórmula para encontrar la ubicación:

$$Me = x_{\frac{n+1}{2}}$$

Donde:

$x = x_i$ = posición del dato (ejemplo: $Me = x_5$)

n = número total de datos

b) Para datos en tablas de frecuencia

Se aplica la fórmula para encontrar la ubicación o posición:

$$Me = x_{\frac{n+1}{2}}$$

Donde:

$x = x_i$ = posición del dato (ejemplo: $Me = x_8$)

n = número total de datos

c) Para datos agrupados en intervalos

Se emplea la ecuación

$$Me = L_i + \left(\frac{\frac{n}{2} - F_{i-1}}{f_i} \right) \cdot A_i$$

Donde:

L_i = Límite inferior del intervalo de clase de la mediana

n = Número total de datos

F_{i-1} = Frecuencia acumulada del intervalo de clase que antecede al intervalo de la mediana.

f_i = Frecuencia absoluta del intervalo de la mediana

A_i = Amplitud o Ancho del intervalo

La Moda

Para datos agrupados en intervalos se emplea la siguiente fórmula:

$$Mo = L_i + \left(\frac{d_1}{d_1 + d_2} \right) \cdot A_i$$

Donde:

L_i = Límite inferior de la clase modal

d_1 = Diferencia entre la frecuencia de la clase modal y la frecuencia del intervalo anterior.

d_2 = Diferencia entre la frecuencia de la clase modal y la frecuencia del intervalo posterior.

A_1 = Amplitud o ancho de la clase modal.

Cuantiles (cuartiles, quintiles, deciles, percentiles)**a) Para datos no agrupados**

Ordenar los datos de menor a mayor y aplicar la fórmula de posición

$$Posición = x_{\left[\frac{k(n+1)}{m} \right]}$$

Donde:

x = posición del dato (ejemplo: x_4)

k = Número del cuantil (Ejemplo: 1 para Q_1 , 25 para P_{25})

n = Número total de datos

m = Número de divisiones (4 para cuartiles, 5 para quintiles, 10 para deciles, 100 para percentiles)

b) Para datos en tablas de frecuencia

Ordenar de menor a mayor y aplicar la fórmula anterior de posición para datos no tabulados

$$\text{Posición} = x_{\left[\frac{k(n+1)}{m}\right]}$$

$$\text{Cuantil} = L_i + \left(\frac{\frac{nk}{m} - F_{i-1}}{f_i} \right) \cdot A_i$$

Donde:

Cuantil: Cuartiles Q_k , Quintiles K_k , Deciles D_k , Percentiles P_k

L_i = Límite inferior del intervalo del cuantil

n = Número total de datos

k = Número del cuantil (Ejemplo: 1 para Q_1 , 25 para P_{25})

m = Número de divisiones (4 para cuartiles, 5 para quintiles, 10 para deciles, 100 para percentiles)

F_{i-1} = Frecuencia acumulada del intervalo que antecede al intervalo del cuantil

f_i = Frecuencia absoluta del intervalo del cuantil

A_i = Amplitud o Ancho del intervalo

A continuación, se presenta una tabla sobre las medidas de posición.

Medida	Símbolo	División	Ejemplo de Uso
Cuartiles	Q_1, Q_2, Q_3	Divide los datos en 4 partes iguales (25%, 50%, 75%).	$Q_1 = P_{25}$ $Q_2 = \text{Mediana (50\% de los datos)}$ $Q_3 = P_{75}$
Quintiles	$K_1 \text{ a } K_4$	Divide los datos en 5 partes iguales (20%, 40%, 60%, 80%).	Usado en economía para analizar desigualdad de ingresos.
Deciles	$D_1 \text{ a } D_9$	Divide los datos en 10 partes iguales (10%, 20%, ..., 90%).	$D_5 = \text{Mediana (50\% de los datos)}$
Percentiles	$P_1 \text{ a } P_{99}$	Divide los datos en 100 partes iguales (1%, 2%, ..., 99%).	$P_{90} = \text{Umbral alto (solo el 10\% supera este valor)}$

Nota:

$$D_1 = P_{10} ; D_2 = P_{20} = K_1 ; D_3 = P_{30} ; D_4 = P_{40} = K_2$$

$$Md = Q_2 = D_5 = P_{50}$$

$$D_6 = P_{60} = K_3 ; D_7 = P_{70} ; D_8 = P_{80} = K_4 ; D_9 = P_{90}$$

Desviación media o desviación promedio**Para datos no agrupados**

$$DM = \frac{\sum |x_i - \bar{x}|}{n}$$

Para datos ordenados en tablas de frecuencia

$$DM = \frac{\sum f |x_i - \bar{x}|}{n}$$

Para datos agrupados en intervalos

$$DM = \frac{\sum f |xm - \bar{x}|}{n}$$

Donde:

x_i = observaciones individuales de los datos

xm = marca de clase.

Varianza y desviación estándar**Para datos no agrupados**

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

Donde:

x_i = observaciones individuales de la población

μ = media aritmética poblacional

N = número de observaciones de la población

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Donde:

x_i = observaciones individuales de la muestra

$\bar{x}$ = media aritmética de la muestra

n = número de observaciones de la muestra

La desviación estándar de una muestra:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Para datos ordenados en tablas de frecuencias

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum f(x_i - \mu)^2}{N}$$

Donde:

f = frecuencia absoluta.

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum f(x_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum f(x_i - \bar{x})^2}{n - 1}$$

La desviación estándar de una muestra se calcula con:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum f(x_i - \bar{x})^2}{n - 1}}$$

Para datos agrupados en intervalos

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum f(xm_i - \mu)^2}{N}$$

Donde:

f = frecuencia absoluta

xm = marca de clase

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum f(xm_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum f(xm_i - \bar{x})^2}{n - 1}$$

La desviación estándar de una muestra se calcula con:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum f(xm_i - \bar{x})^2}{n - 1}}$$

Relación empírica entre desviación estándar y rango

$$s \approx \frac{\text{Rango}}{4}$$

A continuación, se presenta una tabla sobre la relación empírica entre desviación estándar y el rango

Relación	Interpretación	Implicación en los Datos
$s \ll \frac{\text{Rango}}{4}$	Dispersión baja (datos muy concentrados alrededor de la media).	- Poca variabilidad. - La mayoría de los valores están cerca del promedio.
$s \approx \frac{\text{Rango}}{4}$	Dispersión moderada (variabilidad esperada, sin valores extremos dominantes).	- Distribución equilibrada. Ejemplo: $s = 7,052$ $\frac{\text{Rango}}{4} = 7,5$
$s \gg \frac{\text{Rango}}{4}$	Dispersión alta (datos muy dispersos o con <i>outliers</i>).	- Presencia de valores atípicos. - Distribución posiblemente sesgada o con colas pesadas.

Nota: << se lee "mucho menor que" ; $\approx$ se lee "casi igual a" ; >> se lee "mucho mayor que"

Error típico de la media

$$SE = \frac{s}{\sqrt{n}}$$

Donde:

s = Desviación estándar muestral (mide la dispersión de los datos).

n = Tamaño de la muestra.

Rango intercuartílico

Amplitud intercuartílica o rango intercuartílico = tercer cuartil - primer cuartil

$$IQR = Q_3 - Q_1$$

Límites para bigotes

Límite inferior: $Q_1 - 1,5 \cdot IQR$

Límite superior: $Q_3 + 1,5 \cdot IQR$

Dispersión relativa o coeficiente de variación

Para una población

$$CV = \frac{\sigma}{\mu} \cdot 100\%$$

Donde:

CV = Coeficiente de variación; σ = desviación estándar de la población

μ = media aritmética de la población

Para una muestra

$$CV = \frac{s}{\bar{x}} \cdot 100\%$$

Donde:

CV = Coeficiente de variación; s = desviación estándar de la muestra

$\bar{x}$ = media aritmética de la muestra

En la siguiente tabla se muestra la interpretación del CV según su valor

Valor del CV (%)	Interpretación	Implicaciones
CV < 15%	Dispersión baja (datos homogéneos).	- Los datos están muy concentrados alrededor de la media. - Poca variabilidad. Ideal en controles de calidad. Ejemplo: Tiempos de producción en una fábrica eficiente (CV=10%).
15% ≤ CV ≤ 30%	Dispersión moderada (datos relativamente homogéneos).	- Variabilidad aceptable. - Común en estudios sociales o biológicos. Ejemplo: Edades en un grupo de adultos (CV=20%).
CV > 30%	Dispersión alta (datos heterogéneos o con outliers).	- Alta variabilidad. - Puede indicar inconsistencia o grupos mixtos (ej.: ingresos económicos muy desiguales). Ejemplo: Ventas mensuales de productos con algunos éxitos y muchos fracasos (CV=50%).

Coeficiente de asimetría basado en momentos (Fisher-Pearson) con ajuste para una muestra es:

$$g_1 = \frac{n}{(n-1)(n-2)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^3$$

Donde:

x_i = cada uno de los valores

n = tamaño de la muestra

$\bar{x}$ = media aritmética

s = desviación estándar de la muestra

Tabla de interpretación del coeficiente de asimetría (g_1)

Valor de g_1	Tipo de Asimetría	Forma de la Distribución	Interpretación
$g_1 = 0$	Simétrica	Colas equilibradas (distribución normal).	$\bar{x} = Me = Mo$
$-0,5 < g_1 < 0$	Leve asimetría negativa	Cola izquierda ligeramente más larga.	Presencia de valores atípicos bajos. $\bar{x} < Me$
$g_1 \leq -0,5$	Moderada/Fuerte negativa	Cola izquierda pronunciada.	Valores bajos distorsionan la distribución. $\bar{x} \ll Me$
$0 < g_1 < 0,5$	Leve asimetría positiva	Cola derecha ligeramente más larga.	Presencia de valores atípicos altos. $\bar{x} > Me$.
$g_1 \geq 0,5$	Moderada/Fuerte positiva	Cola derecha pronunciada.	Valores altos distorsionan la distribución. $\bar{x} \gg Me$

Error estándar de asimetría

$$SE_{asimetría} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

Donde:

n = Número total de observaciones válidas de la muestra

Para muestras grandes ($n > 150$), se aproxima a:

$$SE_{asimetría} = \sqrt{\frac{6}{n}}$$

Regla práctica:

La regla práctica para determinar si la asimetría de una distribución es estadísticamente significativa (es decir, no aleatoria) se expresa matemáticamente como:

$$|g_1| > 2 \cdot SE_{asimetría}$$

Donde:

g_1 = coeficiente de asimetría

$SE_{asimetría}$ = error estándar de asimetría

Coeficiente de curtosis

La fórmula del coeficiente de asimetría basado en momentos (Fisher-Pearson) con ajuste para una muestra es:

$$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \sum \left(\frac{x_i - \bar{x}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Donde:

x_i = cada uno de los valores; n = tamaño de la muestra

$\bar{x}$ = media aritmética; s = desviación estándar de la muestra

Tabla de Interpretación de g_2

Valor de g_2	Tipo de Curtosis	Forma de la Distribución	Interpretación
$g_2 = 0$	Mesocúrtica	Similar a la distribución normal.	La distribución tiene colas y pico comparables a una normal (referencia).
$g_2 < 0$	Platicúrtica	Pico más bajo y colas más ligeras.	Los datos están menos concentrados alrededor de la media, con menor probabilidad de valores extremos.
$g_2 > 0$	Leptocúrtica	Pico más alto y colas más pesadas.	Mayor concentración de datos cerca de la media y presencia de valores atípicos (outliers).
$g_2 \ll 0$	Platicúrtica fuerte	Pico muy aplanado y colas extremadamente ligeras.	Distribución es mucho más plana que una normal con datos dispersos de manera uniforme, sin concentración cerca de la media y con ausencia de outliers.
$g_2 \gg 0$	Leptocúrtica Fuerte	Pico muy agudo y colas extremadamente pesadas.	Distribución con alta concentración central y outliers significativos.

Nota: $\ll$ se lee "mucho menor que" ; $\gg$ se lee "mucho mayor que"

Error estándar de curtosis

$$SE_{curtosis} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$$

Donde:

n = Número total de observaciones válidas de la muestra

Para muestras grandes ($n > 150$), se aproxima a:

$$SE_{curtosis} = \sqrt{\frac{24}{n}}$$

Regla práctica:

Si el valor de curtosis dividido entre su error estándar cae fuera del rango ± 2 , se considera una desviación estadísticamente significativa de la normalidad.

$$\frac{g_2}{SE_{curtosis}} = \pm 2$$

Donde:

g_2 = coeficiente de curtosis

$SE_{curtosis}$ = error estándar de curtosis

Coeficiente de correlación de Pearson

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

Donde:

n = número de observaciones; X, Y = valores de las variables

Coeficiente de Correlación de Spearman

$$\rho = r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Donde:

d_i = Diferencia entre los rangos de las dos variables para el i -ésimo caso

n = Número de observaciones

Coeficiente de Correlación de Tau-b de Kendall

$$\tau_b = \frac{C - D}{\sqrt{(C + D + T_x)(C + D + T_y)}}$$

Donde:

C = Número de pares concordantes (mismo orden en X e Y).

D = Número de pares discordantes (orden opuesto en X e Y).

T_x = Empates en X ; T_y = Empates en Y

SOLUCIONARIO DE LA EVALUACIÓN DIAGNÓSTICA

CUESTIONARIO

1) En una institución educativa se va a realizar una encuesta a estudiantes de 4to a 10mo sobre temas para construir el Modelo Institucional de Evaluación Educativa. Los datos se muestran en la siguiente tabla:

Curso	Nº de estudiantes
4to	35
5to	34
6to	38
7mo	40
8vo	36
9no	39
10mo	40

Realice los cálculos empleando Excel.

- ¿Cuántas encuestas se debe aplicar si se emplea el cálculo del tamaño de la muestra al 95% de confianza con un error de muestreo del 5%?
- ¿Cuántas encuestas se deben aplicar en cada curso?
- Realice la interpretación de los resultados.

Solución:

	A	B	C	D	E	F
1	N	262	=B20			
2	p	0,5				
3	q	0,5	=1-B2			
4	Confianza	95				
5	Área a la izquierda de -Z	0,025	=(100-B4)/200			
6	-Z	-1,96	=INV.NORM.ESTAND(B5)			
7	Z	1,96	=B6*-1			
8	e = error de muestreo	5	%			
9	e	0,05	=B8/100			
10	$n = \frac{NpqZ^2}{(N-1)e^2 + pqZ^2}$	156,01	=(B1*B2*B3*B7^2)/((B1-1)*B9^2+B2*B3*B7^2)			
11		156	=REDONDEAR(B10;0)			
12	Curso	Nº de estudiantes	Encuestas			
13	4to	35	20,84	=B13/\$B\$20)*\$B\$11	21	=REDONDEAR(C13;0)
14	5to	34	20,24	=B14/\$B\$20)*\$B\$11	20	=REDONDEAR(C14;0)
15	6to	38	22,63	=B15/\$B\$20)*\$B\$11	23	=REDONDEAR(C15;0)
16	7mo	40	23,82	=B16/\$B\$20)*\$B\$11	24	=REDONDEAR(C16;0)
17	8vo	36	21,44	=B17/\$B\$20)*\$B\$11	21	=REDONDEAR(C17;0)
18	9no	39	23,22	=B18/\$B\$20)*\$B\$11	23	=REDONDEAR(C18;0)
19	10mo	40	23,82	=B19/\$B\$20)*\$B\$11	24	=REDONDEAR(C19;0)
20		262	=SUMA(B13:B19)		156	=SUMA(E13:E19)

Interpretación:

a) Se requiere un mínimo de 156 encuestas, con un nivel de confianza del 95% y un error máximo del 5%.

b) La distribución proporcional asegura que cada curso esté representado de acuerdo con su tamaño poblacional:

4to (35 estudiantes): 21 encuestas

5to (34 estudiantes): 20 encuestas

6to (38 estudiantes): 23 encuestas

7mo (40 estudiantes): 24 encuestas

8vo (36 estudiantes): 21 encuestas

9no (39 estudiantes): 23 encuestas

10mo (40 estudiantes): 24 encuestas

c) Esto garantiza que los resultados del estudio sean estadísticamente representativos.

2) Con los siguientes datos sobre las calificaciones trimestrales en la Asignatura de Matemática evaluadas sobre 10 obtenidas de 40 estudiantes, empleando Excel calcular la media, error típico, mediana, moda, desviación estándar, varianza de la muestra, curtosis, coeficiente de asimetría, rango, mínimo, máximo, suma y cuenta. Interpretar los resultados.

8,83	8,85	8,74	7,70
4,48	8,40	8,70	8,99
6,99	8,79	7,90	7,69
7,00	8,70	8,93	9,11
8,84	9,05	8,85	9,83
9,83	8,33	8,88	8,90
8,78	9,01	8,96	8,38
8,71	8,85	8,55	8,11
8,84	6,24	8,99	4,00
7,33	8,26	9,23	9,31

Solución:

	A	B	
1	<i>Columna1</i>		
2			
3	Media	8,3465	
4	Error típico	0,189957	
5	Mediana	8,785	
6	Moda	8,84	
7	Desviación estándar	1,20139353	
8	Varianza de la muestra	1,44334641	
9	Curtosis	5,67680998	
10	Coeficiente de asimetría	-2,24629577	
11	Rango	5,83	
12	Mínimo	4	
13	Máximo	9,83	
14	Suma	333,86	
15	Cuenta	40	
16			

Interpretación:

1. Media (8,3465): El promedio de las calificaciones de los 40 estudiantes en Matemáticas es 8.35 sobre 10, lo que indica un rendimiento general relativamente alto.
2. Error típico (0,190): El error estándar de la media sugiere que, si se tomaran múltiples muestras de 40 estudiantes, la media variaría en aproximadamente $\pm 0,19$ puntos. Esto refleja una precisión moderada en la estimación de la media poblacional.
3. Mediana (8.785): El 50% de los estudiantes obtuvo una calificación igual o inferior a 8,79, y el otro 50% igual o superior. La mediana es mayor que la media, lo que indica una posible asimetría negativa (cola izquierda).
4. Moda (8,84): La calificación más frecuente fue 8,84, que aparece 3 veces (verificado en los datos). Esto refuerza la concentración de notas altas.
5. Desviación estándar (1,201): Las calificaciones se desvían en promedio 1,20 puntos de la media. Una desviación estándar relativamente baja sugiere que la mayoría de los datos están cercanos a la media, aunque hay algunos valores atípicos.
6. Varianza (1,443): La dispersión de los datos alrededor de la media es 1,44 puntos, coherente con la desviación estándar.

7. Curtosis (5,677): Un valor positivo elevado indica una distribución leptocúrtica, con colas más pesadas y un pico más agudo que la normal. Esto se debe a la presencia de valores atípicos (como 4,00 y 4,48) que aumentan la dispersión en las colas.
8. Coeficiente de asimetría (-2,246): El signo negativo revela una asimetría izquierda: la cola izquierda es más larga, y hay concentración de datos en las notas altas. Esto coincide con la mediana $>$ media.
9. Rango (5,83): La diferencia entre la nota máxima (9,83) y mínima (4,00) es 5,83 puntos, mostrando una amplitud significativa, aunque la mayoría de las notas están agrupadas en el rango alto.
10. Mínimo (4,00) y Máximo (9,83): La calificación más baja fue 4.00, un posible atípico que podría deberse a casos excepcionales. La nota más alta fue 9,83, cercana al límite superior (10).
11. Suma (333,86) y Cuenta (40): Suma total de todas las calificaciones: 333,86. Número de estudiantes: 40, lo que confirma que no hay datos faltantes.

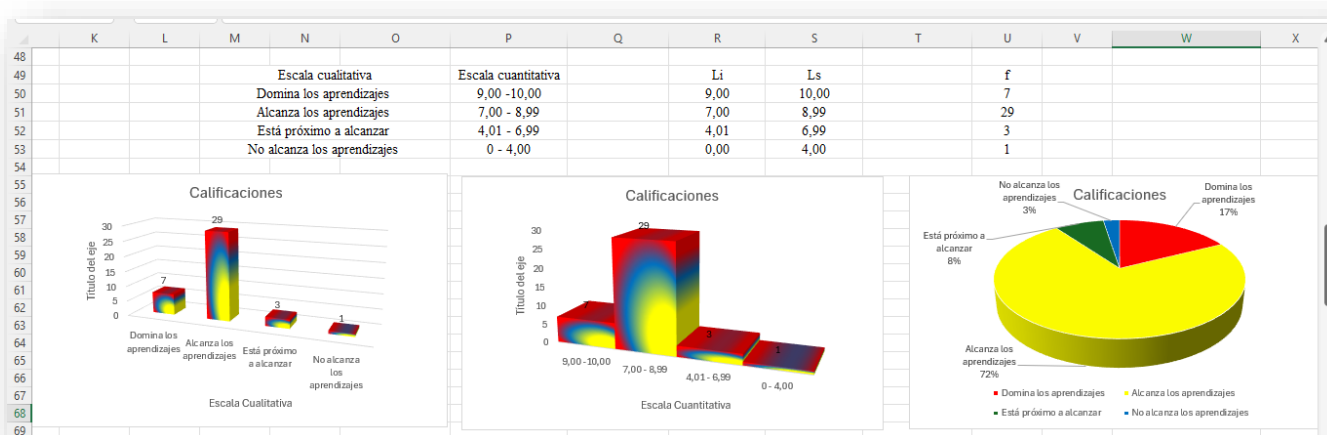
Conclusiones clave:

- Distribución sesgada a la izquierda: La mayoría de los estudiantes tuvo buenas calificaciones (media = 8,35, mediana = 8,79), pero unos pocos obtuvieron notas muy bajas (4,00; 4,48), lo que arrastró la media hacia abajo.
- Presencia de atípicos: Los valores extremos bajos explican la alta curtosis y la asimetría negativa.
- Consistencia en altas notas: La moda (8,84) y la mediana (8,79) sugieren que el grupo tuvo un rendimiento generalmente alto, con una dispersión limitada (desviación estándar de 1,20).

3) Elabore un diagrama de barras, un histograma y un diagrama de sectores considerando los intervalos empleados en el sistema educativo del Ministerio de Educación del Ecuador. Realice la interpretación.

Solución:

Empleando Excel



Interpretación:

La Mayoría está en "Alcanza los aprendizajes" (29 estudiantes): La gran parte del grupo se ubica en un nivel satisfactorio (7,00 – 8,99), lo que indica un buen desempeño general.

Segundo lugar: "Domina los aprendizajes" (7 estudiantes): Un grupo menor pero significativo demuestra un excelente dominio (9,00 – 10,00), destacándose por su alto rendimiento.

Baja frecuencia en niveles inferiores: Solo 3 estudiantes están "próximos a alcanzar" (4,01 – 6,99), y 1 estudiante no logra los aprendizajes (0 – 4,00). Esto sugiere que la mayoría supera los estándares mínimos.

4) Calcule los cuartiles empleando Excel. Realice la interpretación.

Solución:

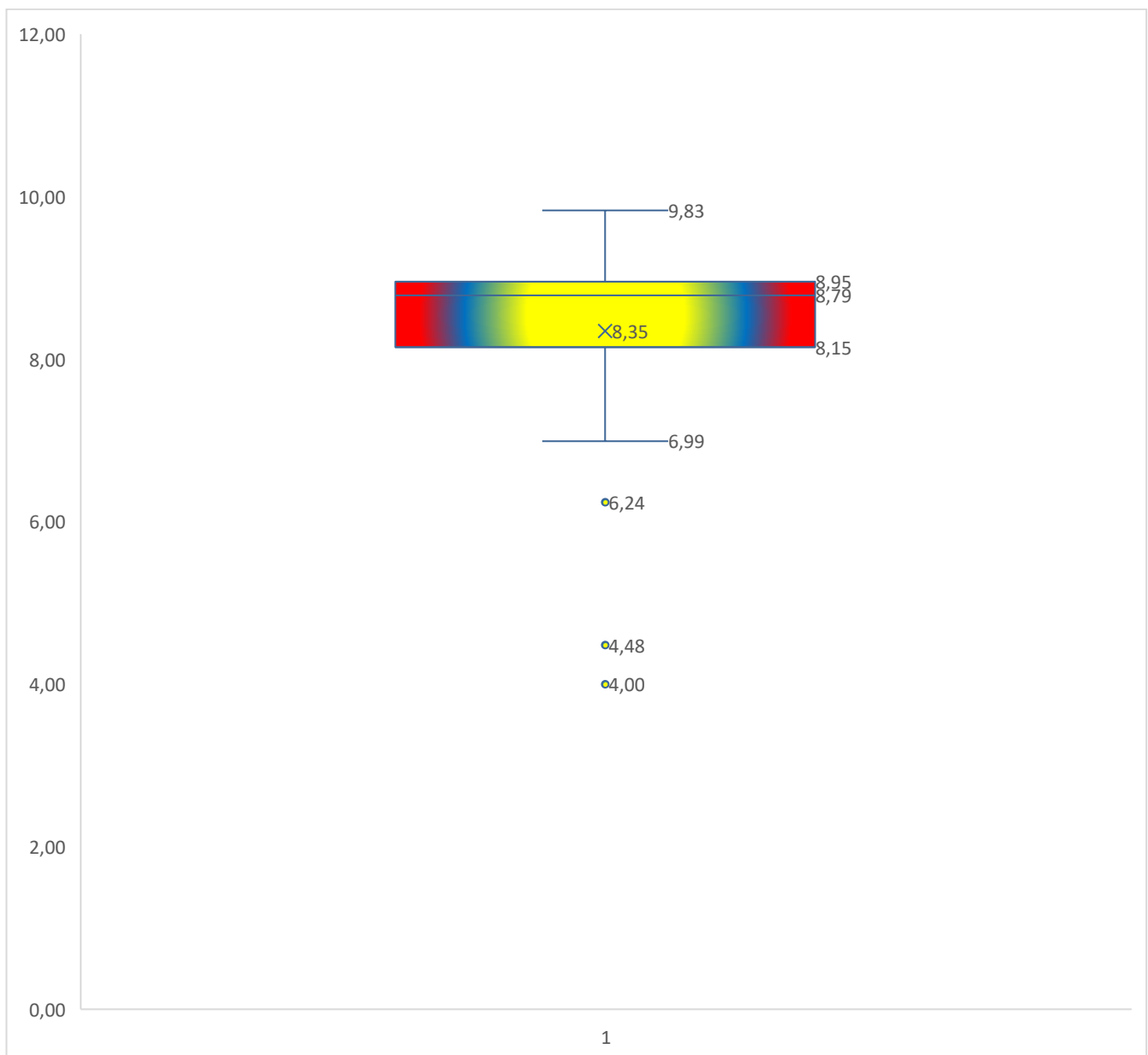
	A	B	C	D	E	F	G
1	8,83						
2	4,48						
3	6,99						
4	7,00		Q_1	8,15	=CUARTIL.EXC(A1:A40;1)		
5	8,84		Q_2	8,79	=CUARTIL.EXC(A1:A40;2)		
6	9,83		Q_3	8,95	=CUARTIL.EXC(A1:A40;3)		
7	8,78						
8	8,71						
9	8,84						

Interpretación:

El $Q_1 = 8,15$ significa que el 25% de los estudiantes obtuvo una calificación $\leq 8,15$, y el 75% restante obtuvo una nota superior.

El $Q_2 = 8,79$ quiere decir que el 50% de los estudiantes está por debajo de 8,79 (coincide con la mediana).

El $Q_3 = 8,95$ se interpreta que el 75% de los estudiantes obtuvo una calificación $\leq 8,95$, y el 25% más alto superó este valor.

5) Elabore un diagrama de caja y bigotes. Realice la interpretación.**Solución:**

Cálculo del Rango Intercuartílico (IQR)

$$IQR = Q_3 - Q_1 = 8,95 - 8,15 = 0,8$$

Límites para los Bigotes

Ordenando los datos se obtiene:

4,00, 4,48, 6,24, 6,99, 7,00, 7,33, 7,69, 7,70, 7,90, 8,11, 8,26, 8,33, 8,38, 8,40, 8,55, 8,70, 8,70, 8,71, 8,74, 8,78, 8,79, 8,83, 8,84, 8,84, 8,85, 8,85, 8,85, 8,88, 8,90, 8,93, 8,96, 8,99, 8,99, 9,01, 9,05, 9,11, 9,23, 9,31, 9,83, 9,83

Límite Inferior (LI):

$$LI = Q_1 - 1,5 \cdot IQR = 8,15 - 1,5 \cdot 0,8 = 6,95$$

Valor mínimo dentro del límite: **Primer dato** $\geq 6,95 \rightarrow 6,99$.

Límite Superior (LS):

$$LS = Q_3 + 1,5 \cdot IQR = 8,95 + 1,5 \cdot 0,8 = 10,15$$

Valor máximo dentro del límite: **Último dato** $\leq 10,15 \rightarrow 9,83$.

Valores Atípicos

Atípicos inferiores: Valores $< 6,95 \rightarrow 4,00; 4,48; 6,24$

Atípicos superiores: Valores $> 10,15 \rightarrow$ Ninguno.

Interpretación:**a) Tendencia Central**

Mediana (8,79): El 50% de los estudiantes obtuvo una nota $\leq 8,79$.

Caja (8,95–8,15): El 50% central de las notas está en este rango estrecho, lo que indica baja dispersión en la mayoría de los datos.

b) Dispersión y Variabilidad

IQR (0,8): Rango intercuartílico pequeño, confirmando concentración de datos.

c) Valores Atípicos

4,00; 4,48; 6,24: Notas significativamente más bajas que el resto.

BIBLIOGRAFÍA

- Alvarado Barzallo, S. M., Ahrango Yacelga, I. A., Asinc Benites, H. E., Caiza Guevara, J. D., Coronel González, D. B., Cortez Solano, J. P., Curay Correa, X. P., Escorcía Romero, T. G., López Carrera, Y. E., López Quintero, H. J., López Realpe, G. G., Mejía Molina, D. Y., Mestanza Barragán, C. A., Molerio Rosa, L. D. I. C., Muñoz Armijos, J. A., Obaco Jaramillo, J. L., Quille Quille, M. U., Quishpi Ojeda, Á. G., Sánchez Enríquez, J. P., Suarez Ijujés, M. O., Suárez Santillán, L. J., Toledo Calle, S. V., & Torres Sarmiento, M. E. (2024). *Docentes en Iberoamérica: Reflexiones hacia la excelencia educativa*. Editorial Red Ecuatoriana de Pedagogía Entremaestros. <https://unirep.edu.ec/repositorioeditorial/>
- Cachuput, J., Suárez, M., Salguero, S., y Reyes, E., (2024). *Estrategias pedagógicas basadas en el enfoque constructivista para mejorar la comprensión de las matemáticas*. Reincisol, 3(6), pp. 4718-4742. [https://doi.org/10.59282/reincisol.V3\(6\)4718-4742](https://doi.org/10.59282/reincisol.V3(6)4718-4742)
- Guangasi Gómez, E. M., Valencia Nuñez, E. R., Montoya Puglla, S. D., & Suárez Ijujés, M. O. (2024). *Exploring the impact of indebtedness on the financial profitability of the automotive sector: a statistical analysis and prediction of bankruptcy*. Runas. Journal of Education and Culture, 5(9), e240168. <https://doi.org/10.46652/runas.v5i9.168>
- Moreira Parrales ML, Mejía Carrillo M de J, Suarez Ijujés MO, Torres Penafiel JS. *Gamification for learning mathematics in secondary school: Most effective gamified strategies to motivate students and improve their performance in mathematics*. Salud, Ciencia y Tecnología. 2024; 4:1016. <https://doi.org/10.56294/saludcyt20241016>
- Núñez-De Luca, J., Gallegos-Robles, J., Suárez-Ijujés, M., y Santillán-Tasigchana, M. (2025). Optimización del aprendizaje conceptual y práctico en matemáticas, física y química mediante la implementación de tecnologías digitales y estrategias de gamificación en la educación superior. Revista Social Fronteriza; 5(3):e707. [https://doi.org/10.59814/resofro.2025.5\(3\)707](https://doi.org/10.59814/resofro.2025.5(3)707)
- Pearson, K. (1895). *Contributions to the Mathematical Theory of Evolution*. Philosophical Transactions of the Royal Society A.
- Olivo Solis, J. E. , Murillo García, M. A., Suárez Ijujés, M. O., & Rizzo Orellana, E. B. (2025). *Una educación más innovadora y de mayor impacto a través de la inteligencia artificial, mediante el aprendizaje personalizado: transformando las estrategias de enseñanza en el nivel superior*. Revista Social Fronteriza, 5(2), e-637. [https://doi.org/10.59814/resofro.2025.5\(2\)637](https://doi.org/10.59814/resofro.2025.5(2)637)
- Pearson, K. (1896). *Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia*. Philosophical Transactions of the Royal Society of London, 187, 253-318.

- Shao, S. (1980). *Estadística para economistas y administradores de empresas*. Herrero Hermanos.
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). McGraw-Hill.
- Suárez, M. y Tapia, F. (2018). *Interaprendizaje de Estadística Básica*. <http://repositorio.utn.edu.ec/handle/123456789/8696>
- Suárez-Ibujés, M. O., & Sánchez-Pozo, N. N. (2023). *Análisis de la influencia geográfica en la evaluación Ser Estudiante a través de estadística multivariante*. Prometeo Conocimiento Científico, 3(1), e27. <https://doi.org/10.55204/pcc.v3i1.e27>
- Suárez-Ibujés, M. O., Hernández-Dávila, C. A., Peñafiel, E. J. A., & Villena-Atoche, C. A. (2024). *Utilización de juegos de razonamiento lógico para potenciar competencias matemáticas en estudiantes de bachillerato*. MQRInvestigar, 8(2), 2931–2950. <https://doi.org/10.56048/MQR20225.8.2.2024.2931-2950>
- Suárez, M. (2011). *Coficiente de correlación de Karl Pearson*. <https://repositorio.utn.edu.ec/handle/123456789/766>
- Suárez, M. (2011). *Coficiente de correlación por rangos de Spearman*. <https://repositorio.utn.edu.ec/handle/123456789/768>
- Suárez, M. (2018). *Probabilidades y estadística empleando las TIC*. <https://repositorio.utn.edu.ec/handle/123456789/8698>
- Suárez, M. (2018). *Los poliprismas y su aplicación en la enseñanza de la matemática*. <https://repositorio.utn.edu.ec/handle/123456789/8697>
- Suárez, M. (2025). *Estadística Descriptiva para Todos: Fundamentos y Aplicaciones-Volumen I*. https://www.researchgate.net/publication/391908519_ESTADISTICA_DESCRIPTIVA_PARA_TODOS_FUNDAMENTOS_Y_APLICACIONES_VOLUMEN_I
- Suárez, M. (2025). *Estadística Descriptiva para Todos: Fundamentos y Aplicaciones-Volumen II*. https://www.researchgate.net/publication/392165180_Estadistica_Descriptiva_para_Todos_Fundamentos_y_Aplicaciones-Volumen_II

DATOS BIOGRÁFICOS DEL AUTOR

Mario Orlando Suárez Ibujés nació el 24 de marzo de 1978 en el barrio La Florida de la ciudad de Ibarra, provincia de Imbabura, Ecuador. Hijo de Segundo Suárez y Bertha Ibujés. Realizó sus estudios primarios en la Escuela Fiscal Mixta "Alejandro Pasquel Monge" de su localidad, donde destacó como Abanderado del Estandarte de la Escuela, Abanderado del Pabellón Nacional y Mejor Egresado. Posteriormente, cursó la educación secundaria en el Colegio "Teodoro Gómez de la Torre" de Ibarra, donde sobresalió académicamente: fue Mejor Alumno en Matemática durante sus tres últimos años, Abanderado del Estandarte del Colegio y Mejor Egresado. A nivel superior, se graduó como Licenciado en Ciencias de la Educación, con especialización en Física y Matemática, en la Universidad Técnica del Norte (UTN) de Ibarra, donde obtuvo el reconocimiento de Mejor Egresado. Su formación de cuarto nivel incluye:

- Magíster en Gerencia de Proyectos Educativos y Sociales (UTN).
- Magíster en Estadística Aplicada (Universidad Politécnica Estatal del Carchi), graduado con Mención Honorífica.

Formación Doctoral en progreso:

- Doctorado PhD en Educación (Universidad Benito Juárez, México). Egresado.
- Doctorado PhD en Educación e Innovación (Universidad de Investigación e Innovación de México).

Actualmente, es Asesor Educativo en la Coordinación Zonal 1-Educación de Ecuador y docente ocasional de posgrado en la UTN. Tiene cinco Doctorados Honoris Causa. Ha publicado 16 libros, 8 artículos en revistas indexadas, es autor de 5 rompecabezas llamados Poliprismas, un Juego Matemático en la Chakana y de 405 recursos educativos en repositorios digitales.

Trayectoria profesional:

- Docente de Matemática del Bachillerato Internacional (BI) y (EGB) en la Unidad Educativa. Teodoro Gómez de la Torre, en donde inicia su experiencia profesional docente a los 20 años de edad.
- Docente de Matemática en la Escuela Alejandro Pasquel Monge.
- Docente de Matemática de Educación General Básica (EGB) y Bachillerato General Unificado (BGU) en la Academia Militar San Diego.
- Docente de Matemática de EGB en la Unidad Educativa Mariano Suárez Veintimilla.
- Docente de Matemática de BGU y BI en la Unidad Educativa Ibarra.
- Director Distrital de Educación 10D02 Antonio Ante-Otavaló.
- Docente de la UTN en la Facultad de Ciencias Administrativas y Económicas (FACAE).
- Actual Docente ocasional en la Facultad de Posgrado de la UTN y Asesor Educativo en la Coordinación Zonal 1-Educación.

Producción Académica

Libros publicados (16 obras):

- Unidades para Producir Medios Instruccionales en Educación (coautor a los 24 años).
- Interaprendizaje Holístico de Matemática (autor).
- Hacia un Interaprendizaje Holístico de Álgebra y Geometría (autor).
- Matemática Recreativa (coautor).
- Interaprendizaje de Probabilidades y Estadística Inferencial Empleando Excel, Winstats y Graph (autor).
- Interaprendizaje de Estadística Básica (coautor).
- Probabilidades y Estadística empleando las TIC (autor).
- Matemática y sus aplicaciones empleando las TIC (coautor).
- Los Poliprismas y su aplicación en la enseñanza de la Matemática (autor).
- El PAPT en Cotacachi (coautor).
- Docentes en Iberoamérica: Reflexiones hacia la Excelencia Educativa (coautor).
- Hacia un interaprendizaje de Matemática Financiera (autor).
- Maestros de Excelencia Transformando la Educación en Iberoamérica Reflexiones y Desafíos (coautor).
- Estadística Descriptiva para Todos. Fundamentos y Aplicaciones-Volumen I (autor).
- Estadística Descriptiva para Todos. Fundamentos y Aplicaciones-Volumen II (autor).
- Estadística Descriptiva para Todos. Fundamentos y Aplicaciones-Volumen III (autor).

https://isbnecuador.com/catalogo.php?mode=busqueda_menu&id_autor=6289

Artículos científicos (8 en revistas indexadas):

- Guía didáctica para el interaprendizaje de Trigonometría Básica empleando el Poliprisma. Revista el Investigador N° 4 de la Universidad Técnica del Norte.
<https://issuu.com/utnuniversity/docs/el-investigador-n04>
- Análisis de la influencia geográfica en la evaluación Ser Estudiante a través de estadística multivariante. Prometeo Conocimiento Científico, 3(1), e27.
<https://doi.org/10.55204/pcc.v3i1.e27>
- Gamification for learning mathematics in secondary school: Most effective gamified strategies to motivate students and improve their performance in mathematics. Salud, Ciencia y Tecnología. 2024; 4:1016. <https://doi.org/10.56294/saludcyt20241016>
- Exploring the impact of indebtedness on the financial profitability of the automotive sector: a statistical analysis and prediction of bankruptcy. Runas. Journal of Education and Culture, 5(9), e240168. <https://doi.org/10.46652/runas.v5i9.168>
- Utilización de juegos de razonamiento lógico para potenciar competencias matemáticas en estudiantes de bachillerato. MQRInvestigar, 8(2), 2931–2950.
<https://doi.org/10.56048/MQR20225.8.2.2024.2931-2950>
- Estrategias pedagógicas basadas en el enfoque constructivista para mejorar la comprensión de las matemáticas. Reincisol, 3(6), pp. 4718-4742. [https://doi.org/10.59282/reincisol.V3\(6\)4718-4742](https://doi.org/10.59282/reincisol.V3(6)4718-4742)

- Una educación más innovadora y de mayor impacto a través de la inteligencia artificial, mediante el aprendizaje personalizado: transformando las estrategias de enseñanza en el nivel superior. *Revista Social Fronteriza*, 5(2), e-637. [https://doi.org/10.59814/resofro.2025.5\(2\)637](https://doi.org/10.59814/resofro.2025.5(2)637)
- Optimización del aprendizaje conceptual y práctico en matemáticas, física y química mediante la implementación de tecnologías digitales y estrategias de gamificación en la educación superior. *Revista Social Fronteriza*; 5(3): e707. [https://doi.org/10.59814/resofro.2025.5\(3\)707](https://doi.org/10.59814/resofro.2025.5(3)707)

Obras literarias inéditas y material didáctico:

- Poliprisma 3.0, Poliprisma 4.0, Poliprisma 7.0, Poliprisma 9.0, Poliprisma 9.1.
- Juego Matemático en la Chakana.

Temas publicados en repositorios digitales

405 temas sobre temas educativos se encuentran publicados en:

<https://www.researchgate.net/profile/Mario-Suarez-Ibujes>

<http://es.scribd.com/mariosuarezibujes>

<https://amee.ec/author/mario-amee/>

<https://repositorio.utn.edu.ec/simple->

[search?filterquery=Su%C3%A1rez+Ibuj%C3%A9s%2C+Mario+Orlando&filtername=author&filtertype>equals](https://repositorio.utn.edu.ec/simple-search?filterquery=Su%C3%A1rez+Ibuj%C3%A9s%2C+Mario+Orlando&filtername=author&filtertype>equals)

<https://www.youtube.com/@marioorlandosuarezibujes4651/videos>

<https://scholar.google.com/citations?user=FUoyU1cAAAAJ&hl=e>

Reconocimientos y Premios Nacionales e Internacionales

Doctorados Honoris Causa

- Doctor Honoris Causa en Investigación otorgado por la Organización Internacional para la Inclusión y Calidad Educativa (OIICE). Perú-Cusco, año 2025.
<https://youtube.com/shorts/shJ-AZUZ0Lk?feature=share>
<https://youtu.be/F7BJCvlu-iE>
- Doctor Honoris Causa otorgado por Remzion University of Jerusalem. Cartagena de Indias, Colombia, año 2025
<https://es.scribd.com/presentation/839031232/Doctor-Honoris-Causa-en-Remzion-University-of-Jerusalem>

- Doctor Honoris Causa en Andragogía otorgado por OIICE. Perú-Cusco, año 2024.
<https://youtu.be/-KZ9pK44hBw>
<https://www.scribd.com/document/839033529/Doctor-Honoris-Causa-en-Andragogia>
- Doctor Honoris Causa otorgado por la Universidad del Norte de Tamaulipas. Perú-Lima, año 2024.
<https://es.scribd.com/document/708966545/Mario-Orlando-Suarez-Ibujes-Doctor-Honoris-Causa-UNT>
<https://www.youtube.com/shorts/NA-hLFXaxqo>
<https://www.youtube.com/watch?v=5cREEjBLBZw>
- Doctor Honoris Causa otorgado por OIICE. Perú- Cusco, año 2023.
<https://es.scribd.com/document/649433171/Galardon-a-La-Excelencia-Educativa-OIICE-Cusco-2023>
https://www.imbaburaenlinea.com/2023/06/08/docente-imbabureno-recibe-reconocimiento-internacional/?fbclid=IwAR0ls92m6pLJ_rPo6d5WQkibkLO7U65w87WJSn7gHKx-ZwaoMnJIKkLTRXo
<https://youtu.be/JJGzdiSlu6c>

Premios Internacionales

- Galardón a la Excelencia Educativa Cusco 2025 otorgado por la Organización Internacional para la Inclusión y Calidad Educativa OIICE. Perú- Cusco, año 2025.
<https://youtu.be/F7BJCvlu-iE>
- Premio a la Excelencia Gerencial otorgado por la Sociedad Internacional de Gerencia. Cartagena de Indias, Colombia, año 2025
<https://es.scribd.com/presentation/839031232/Doctor-Honoris-Causa-en-Remzion-University-of-Jerusalem>
- Education Leadership Awards (Premio al Liderazgo en Educación) 2024 otorgado por Christian Chambers Entrepreneurs. Perú-Lima, año 2024.
<https://youtu.be/Smsfu20MAoE>
- Global Education Prize (Premio Global de Educación) 2024 otorgado por la Asociación de Acreditación Internacional y Certificación de Entidades Privadas (AAICEP). Perú-Lima, año 2024.
<https://www.youtube.com/watch?v=B-qqUbNMCoQ>

- Global Learning Awards (Premio Global de Aprendizaje) en la Categoría Educación Innovadora otorgado por la Cámara Iberoamericana de Educación (IBEROCAM). Lima-Perú, año 2024.
<https://www.youtube.com/watch?v=tenrIEKFpXY>
- Premio Educa Latinoamérica 2024 en la categoría Educación de Excelencia otorgado por la Cámara Peruana de Desarrollo y Educación. Lima-Perú, año 2024.
<https://www.youtube.com/watch?v=Ok6alM5GXiU&t=15s>
<https://www.scribd.com/document/717050277/Premio-Educa-Latinoamerica-2024>
- Orden al Mérito Educativo y Cultural Magnus Docentis otorgado (OIICE). Perú-Cusco, año 2024.
<https://youtu.be/-KZ9pK44hBw>
<https://www.scribd.com/document/839033529/Doctor-Honoris-Causa-en-Andragogia>
- Galardón a la Excelencia Educativa Edición Cusco 2024 otorgado por OIICE. Perú-Cusco, año 2024.
<https://youtu.be/-KZ9pK44hBw>
<https://www.scribd.com/document/839033529/Doctor-Honoris-Causa-en-Andragogia>
- Colegiatura Oficial Internacional de Doctorado Honoris Causa otorgado por el Colegio Internacional de Doctores. Costa Rica-San José, año 2024
<https://es.scribd.com/document/708961975/Mario-Orlando-Suarez-Ibujes-Colegiatura-Internacional-de-Doctor-Honoris-Causa>
- Educador de Eminencia otorgado por Universidad Ricardo Palma-Escuela de Marketing y Administración Comercial. Costa Rica-San José, año 2024
<https://es.scribd.com/document/708961975/Mario-Orlando-Suarez-Ibujes-Colegiatura-Internacional-de-Doctor-Honoris-Causa>
- Orden Dorada Magisterial otorgado por OIICE. Perú- Cusco, año 2023.
<https://es.scribd.com/document/649433171/Galardon-a-La-Excelencia-Educativa-OIICE-Cusco-2023>
<https://youtu.be/JJGzdiSlu6c>
- Galardón a la Excelencia Educativa Cusco 2023 otorgado por OIICE. Perú- Cusco, año 2023.
<https://es.scribd.com/document/649433171/Galardon-a-La-Excelencia-Educativa-OIICE-Cusco-2023>
<https://youtu.be/JJGzdiSlu6c>

Distinciones Nacionales

- Condecoración al Mérito Educativo "Alfredo Pérez Guerrero" otorgado por el Gobierno Autónomo Descentralizado Municipal de San Miguel de Ibarra. Ibarra-Ecuador, año 2025.
<https://youtu.be/KNLH6qQeYQ8>
<https://www.facebook.com/share/195kUYAsPF/>
- Condecoración “Medalla Julio Miguel Aguinaga” al mérito educativo otorgado por el Gobierno Autónomo Descentralizado Municipal de Antonio Ante. Atuntaqui-Ecuador, año 2024.
<https://www.youtube.com/watch?v=3VqYG7lLhjQ>
<https://www.youtube.com/watch?v=99t2paNDI6I>
<https://fb.watch/qA4Z5YGhnR/>
- Diploma y Placa de Reconocimiento por trayectoria docente y aporte invaluable al Magisterio Fiscal otorgado por la Coordinación Zonal 1 del Ministerio de Educación del Ecuador. Antonio Ante-Ecuador, año 2024.
<https://es.scribd.com/document/839014751/Diploma-y-Placa-de-Reconocimiento-CZ1#logout>
- Diploma de Reconocimiento por la destacada contribución a la innovación educativa otorgado por la Dirección Distrital 10D02 Antonio Ante-Otavalo-Educación. Otavalo-Ecuador, año 2024.
<https://www.facebook.com/share/v/RNti5UtrqkDQC8QC/?mibextid=WC7FNe>
- Diploma de Reconocimiento al mérito profesional por el aporte a la educación otorgado por la Unidad Educativa “Sarance”. Otavalo-Ecuador, año 2024.
<https://www.facebook.com/share/p/YTyLodqDJpflHahv/?mibextid=xfxF2i>
- Placa de reconocimiento por excelente trayectoria académica y profesional otorgado por el área de Matemática y Física de la Unidad Educativa Teodoro Gómez de la Torre. Ecuador-Ibarra, año 2024
<https://es.scribd.com/document/839071664/Reconocimientos-Profesionales-Hasta-Marzo-de-2025>
- Placa de Homenaje de Gratitud y Reconocimiento, otorgado por la 1ra Cohorte de la Maestría en Pedagogía Mención Currículo Modalidad en Línea de la Universidad Técnica del Norte. Ecuador-Ibarra, año 2023.
- Diploma de reconocimiento por el aporte a la investigación científica y tecnológica al haber contribuido con publicaciones científicas durante el año 2017 otorgado por la Universidad Técnica del Norte. Ecuador- Ibarra, año 2018.
- Premio Nacional Galardón Nacional Estatuilla “Nöus” por ser el ganador del VI Concurso Nacional de Excelencia Educativa, otorgando por la Fundación para la Integración y Desarrollo

de América Latina (FIDAL) y la Revista Edu@news. Ecuador-Quito, año 2014. Se encuentra publicado en

<https://www.youtube.com/watch?v=hiIX-jZUM8g>

https://www.youtube.com/watch?v=l-H_rkSZdbS

- Diploma por el valioso aporte al cumplimiento de los objetivos institucionales, otorgado por la Unidad Educativa Ibarra. Ecuador- Ibarra, año 2014.
- Premio Nacional a la Excelencia Docente “Rita Lecumberri” en la categoría Educador Innovador otorgado por el Ministerio de Educación del Ecuador, año 2013.
<https://educacion.gob.ec/cinco-docentes-ecuatorianos-fueron-reconocidos-por-su-excelencia-docente/>
https://www.youtube.com/watch?v=fN614do_3II
<https://es.scribd.com/doc/135847484/Premio-Rita-Lecumberri>
- Diploma y placa de reconocimiento por la excelente trayectoria como docente investigador y destacado profesor universitario otorgado por la Universidad Técnica del Norte. Asociación de Profesores de la Facultad de Ciencias Administrativas y Económicas. Ecuador- Ibarra, año 2013.
- Estatuilla “El Pensador” al Mérito Académico otorgado por la Asociación General de Profesores de la Universidad Técnica del Norte. Ecuador-Ibarra, año 2013.
- Diploma como Profesor tutor de estudiantes ganadores de Concursos Intercolegiales de Matemática. Academia Militar “San Diego”. Ecuador-Ibarra, año 2008.
- Diploma de Honor por haber aportado positivamente al desarrollo académico de Academia Militar “San Diego”. Academia Militar “San Diego”. Ecuador-Ibarra, año 2005.
- Diploma como Asesor de proyectos ganadores en la Primera Feria Binacional de Ciencia y Tecnología Ecuador Colombia. Unidad Educativa Experimental “Teodoro Gómez de la Torre”. Ecuador-Ibarra, año 2005.
- Mejor Trabajo de Investigación. Certificado de la UTN-Centro Universitario de Investigación Científica y Tecnológica, por haber presentado la Tesis “Interaprendizaje de poliedros irregulares de bases paralelas empleando al Multiprisma” en la Casa Abierta. Ecuador- Ibarra, año 2003.
- Diploma por haber asesorado satisfactoriamente en el VII Concurso Provincial de Matemática, otorgado por el Colegio Nacional Ibarra. Ecuador- Ibarra, año 2003
- Mención Especial en Ciencias Básicas (Matemática), Premio Nacional otorgado por la VI Feria de Ciencia, Tecnología e Innovación por haber triunfado con el Proyecto Multiprisma (Un rompecabezas tridimensional bicolor integrado por partes prismáticas). Ecuador-Quito, año 2001.
<https://es.scribd.com/document/839071664/Reconocimientos-Profesionales-Hasta-Marzo-de-2025>

MÉRITOS ESTUDIANTILES



PRODUCCIÓN ACADÉMICA



RECONOCIMIENTOS DESTACADOS

