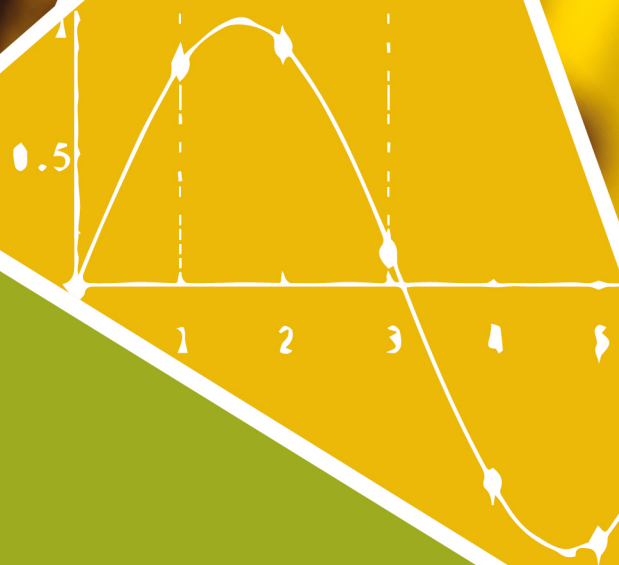


Manual de análisis estadístico y uso de bases de datos



Jaime Eduardo Guzmán Pantoja • Elsa Armida Gutiérrez Román
Gloria Angélica Hernández Obledo • Jesús Aarón Curiel Beltrán
Camilo Patiño García • José María Jiménez Ávila • Diana Guzmán Pantoja

Universidad de Guadalajara

Manual de análisis estadístico y uso de bases de datos

Manual de análisis estadístico y uso de bases de datos

JAIME EDUARDO GUZMÁN PANTOJA
ELSA ARMIDA GUTIÉRREZ ROMÁN
GLORIA ANGÉLICA HERNÁNDEZ OBLEDO
JESÚS AARÓN CURIEL BELTRÁN
CAMILO PATIÑO GARCÍA
JOSÉ MARÍA JIMÉNEZ ÁVILA
DIANA GUZMÁN PANTOJA

UNIVERSIDAD DE GUADALAJARA
2015

Primera edición, 2015

D.R. © 2015, Universidad de Guadalajara
Centro Universitario de la Costa
Av. Universidad 203, Delegación Ixtapa
48280 Puerto Vallarta, Jalisco

ISBN: 978-607-742-216-7

Impreso y hecho en México
Printed and made in Mexico

Contenido

Introducción	9
Diseño de estudios de investigación	11
Variables	15
Estadística descriptiva	19
Estadística inferencial	23
Representación gráfica del análisis estadístico	27
Base de datos	29
Pruebas de normalidad	45
Media	55
Mediana	61
Moda	65
Rango	69
Varianza	73
Desviación típica	77
T de Student	81
Anova	91
Correlación de Pearson	97
Odds ratio	103
Chi ²	109
McNemar	115

Wilcoxon	121
U de Mann-Whitney	127
Kruskal-Wallis	133
Correlación de Spearman	139
Regresión lineal	145
Breviario estadístico	153
Bibliografía	155

Introducción

La ciencia actual se basa en el inductivismo, es decir, todo conocimiento está sustentado en la evidencia obtenida mediante el método experimental.

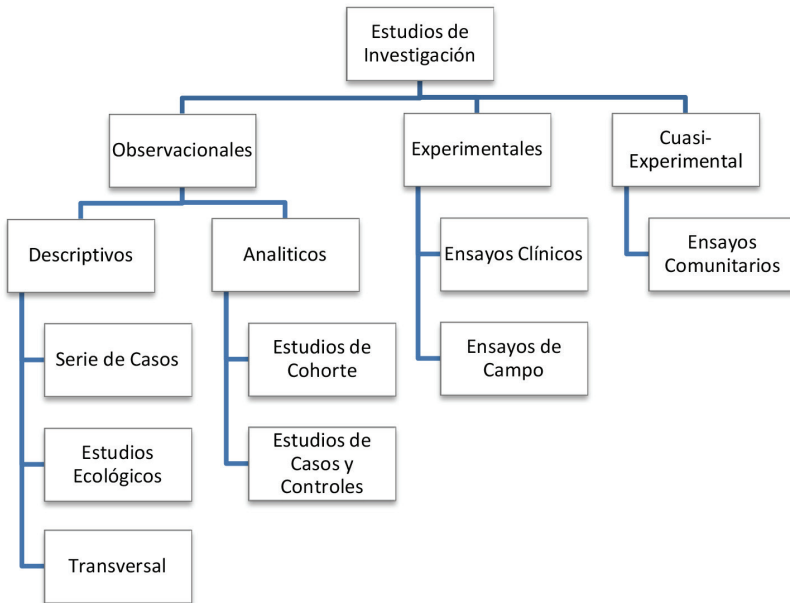
La ciencia médica en particular obtiene sus conocimientos a partir del inductivismo, por lo que es fundamental contar con los conocimientos del análisis estadístico para realizar una correcta investigación en salud.

Este manual tiene como objetivo facilitar el uso del programa SPSS versión 20.0 con el propósito de analizar las variables de estudio.

Diseño de estudios de investigación

La investigación biomédica tiene como objetivo describir la distribución de las enfermedades y eventos de salud en poblaciones humanas, así como contribuir al descubrimiento de los elementos que influyen en estas condiciones.

Desde el punto de vista metodológico, los estudios realizados en poblaciones humanas se clasifican de acuerdo con su diseño de la siguiente forma:



Estudios observacionales

Los estudios observacionales se diferencian, según el objetivo que persigan, en:

- Descriptivos: el investigador se limita a describir el fenómeno objeto de estudio.
- Analíticos: el investigador analiza las causas del fenómeno en estudio.

Estudios descriptivos

Como su nombre lo dice, estos estudios tienen como objetivo describir los factores determinantes de la enfermedad en una población específica sin establecer la relación causa-efecto. Comúnmente describen la frecuencia de presentación de una enfermedad, la población vulnerable, los factores de riesgo, así como las características propias de la enfermedad.

En la *serie de casos* se describen las características observadas de una patología en un paciente o un grupo de pacientes con un diagnóstico similar, con el propósito de generar nuevas hipótesis. Frecuentemente describe la presencia de nuevas enfermedades o efectos adversos, contribuyendo al mantenimiento de la vigilancia epidemiológica.

- Ventaja: permite generar nuevas hipótesis, mantiene una vigilancia epidemiológica.
- Desventaja: ausencia de un grupo control.

Los *estudios ecológicos* centran su unidad de estudio en la población, describiendo patologías en relación con una exposición determinada o variables de interés como edad, sexo, factores de riesgo, etcétera.

- Ventajas: se realizan con facilidad y rapidez ya que la información está disponible. Son de bajo o nulo costo económico. Permite generar nuevas hipótesis.
- Desventajas: no permite obtener conclusiones a nivel individual. Incapacidad para controlar por variables confusoras.

El *estudio transversal* es el mejor para evaluar prevalencia; estudia de forma simultánea tanto exposición como enfermedad en un momento determinado, buscando hipótesis etiológicas que se tendrán que confirmar con estudios analíticos.

- Ventajas: se realiza con facilidad y rapidez, es de bajo costo económico. Permite estudiar simultáneamente múltiples variables.

- Desventajas: no permite determinar si la exposición precedió a la enfermedad o viceversa, ocasionando muchos sesgos.

Estudios analíticos

Tienen el objetivo de analizar la relación entre la enfermedad y otras variables, por lo que son los que mejor evalúan los factores de riesgo. En los estudios con diseño analítico el investigador observa, analiza y concluye.

Los *de cohorte o de seguimiento* son estudios longitudinales que parten de sujetos sanos, evaluando la frecuencia con que aparece la enfermedad en relación de la presencia o ausencia de exposición a un determinado factor.

Estos pueden ser *prospectivos* cuando el estudio se inicia en el presente y se continúa hacia el futuro o *retrospectivos o de cohorte histórico* cuando tanto la exposición como la enfermedad ya se presentaron al momento de iniciar el estudio.

- Ventajas: estima incidencia de la enfermedad. Presenta mínimos sesgos.
- Desventajas: costo económico elevado. No son útiles para enfermedades poco comunes o de largo periodo de latencia. Necesitan grandes tamaños de muestra, ya que puede haber pérdidas durante el seguimiento.

Los estudios de *casos y controles* son los más utilizados en la investigación no experimental, con sentido longitudinal retrospectivo, donde se evalúa la frecuencia de exposición a un factor determinado entre un grupo de individuos enfermos (casos) y un grupo de individuos sanos (control).

- Ventajas: fácil de realizar y de bajo costo económico. Útil para enfermedades poco comunes, crónicas o de largo periodo de latencia. Sirve para farmacovigilancia.
- Desventajas: sólo valora prevalencia de exposición en enfermos y sanos. Alta probabilidad de presentar sesgos.

Estudios experimentales

Los estudios experimentales evalúan el impacto que se produce por la intervención del investigador al determinar la exposición del grupo experimental en relación con la evolución del grupo control. Son de utilidad para valorar la eficacia de terapias farmacológicas, de actividades preventivas y de programación sanitarios.

El *ensayo clínico* es el estudio experimental más frecuente y el mejor para evaluar la asociación causal. Su principal característica es que los individuos son asignados de forma aleatoria al grupo de intervención y al grupo control.

- Ventaja: mejor control de diseño con menos posibilidad de sesgos. Es reproducible y comparable con otros ensayos clínicos.
- Desventaja: costo económico elevado. Limitaciones éticas.

El *ensayo de campo* centra su unidad de estudio en la población sana o en riesgo de enfermedad con el fin de evaluar la eficacia de medidas preventivas.

- Ventajas: determina la eficacia de las medidas preventivas en la población.
- Desventajas: alto costo económico. Necesitan grandes tamaños de muestra (unidad de estudio es la población). Limitaciones éticas.

Estudios cuasi-experimentales

En los estudios cuasi-experimentales o estudios de “antes y después” existe manipulación por parte del investigador pero no aleatorización, determinando la incidencia con intervención y sin intervención.

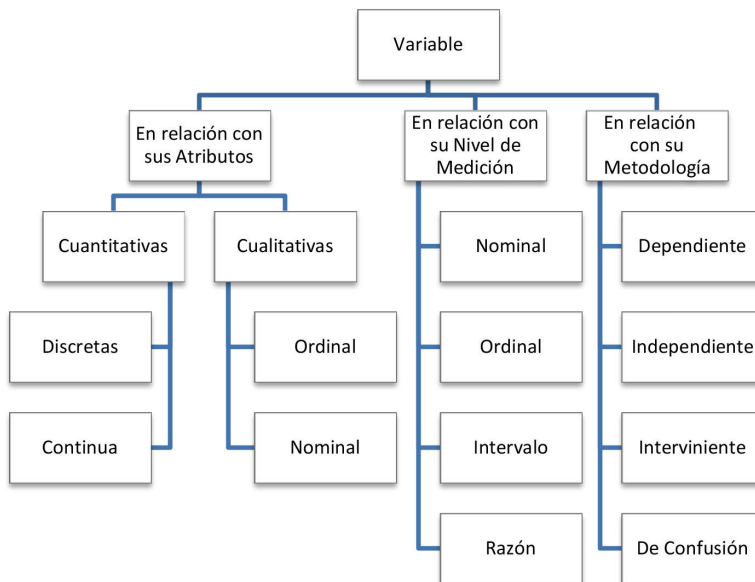
El *ensayo comunitario* centra su unidad de estudio en la población, en la que una o varias comunidades no aleatorizadas recibirán intervención y otras servirán como control.

- Ventaja: determinan incidencia en grandes poblaciones.
- Desventajas: no aleatorización. Limitaciones éticas.

Variables

Se entiende como *variable* toda propiedad o atributo de los fenómenos en estudio con susceptibilidad de medición.

Las variables de estudio se clasifican de acuerdo con sus características en:



En relación con sus atributos

Las *variables cuantitativas* son aquellas que expresan una característica que se encuentra definida por valores numéricos, como por ejemplo el peso, la edad, la altura, el número de hijos, etcétera.

Las variables cuantitativas se clasifican de acuerdo al valor numérico a cuantificar. De esta forma, se definen como:

- *Discreta o discontinua*: cuando se puede cuantificar sólo en valores enteros, como en el caso de la edad, el número de hijos, etcétera.
- *Continua*: cuando se puede cuantificar en valores fraccionados, como en el caso del peso (67.5 kg), la estatura (1.68 m), la temperatura (37.4 °C), nivel de glucosa (323 mg/dl), etcétera.

Las *variables cualitativas* son aquellas que expresan una característica que es imposible cuantificar, como por ejemplo el sexo, el estado civil, la profesión, la nacionalidad, etcétera.

Las variables cualitativas se clasifican de acuerdo con sus características específicas:

- *Nominal*: cuando entre las variables a estudiar no existe punto de comparación u orden específico, como en el caso de la nacionalidad (mexicano, argentino, chileno, etc.), el estado civil (soltero, casado, viudo). La variable nominal puede ser *dicotómica* cuando se admiten sólo dos categorías dentro de la variable (como por ejemplo masculino o femenino).
- *Ordinal*: cuando las características de la variable siguen cierto orden, como en el caso de las clases sociales (alta, media y baja), el orden de nacimiento (primero, segundo, tercero, etc.) o el nivel educativo (primaria, secundaria, preparatoria, etcétera).

En relación con su nivel de medición

Las variables nominal y ordinal poseen las características cualitativas ya mencionadas, mientras que las variables de intervalo y de razón poseen características cuantitativas, en donde:

La *variable de intervalo* utiliza al cero como unidad de medición de forma arbitraria con la posibilidad de cuantificar entre positivos y negativos, como en el caso de la temperatura, en donde los grados bajo cero no indican ausencia de temperatura.

La *variable de razón*, por el contrario, utiliza al cero como escala absoluta. Si el objeto que se cuantifica obtiene el valor de cero, entonces no posee la propiedad que se está midiendo (como en el caso del peso, en donde cero gramos indica ausencia de peso).

En relación con su metodología

La *variable dependiente o de criterio* es la variable principal que se está estudiando y que se encuentra en el objetivo general o en el título de la investigación y que, por lo tanto, representa el *efecto* posible.

La *variable independiente* es aquella que explica el porqué de la variable dependiente y, por lo tanto, representa la *causa* que origina cambios en la variable dependiente.

La *variable interviniente* es aquella que facilita la comprensión entre la variable dependiente y la independiente, sin modificar el resultado.

La *variable de confusión o confusora* es el factor que distorsiona la asociación entre la variable dependiente e independiente.

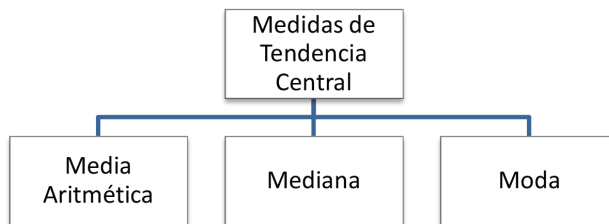
Por ejemplo si estudiamos si la disfunción conyugal es consecuente de la depresión postparto, la disfunción conyugal es la variable dependiente, la depresión postparto es la variable independiente, el nivel educativo o económico de la madre es la variable interviniente y la ansiedad o autoestima de la madre es la variable confusora.

Estadística descriptiva

La estadística descriptiva tiene como objetivo principal realizar un análisis descriptivo del comportamiento de los datos que se obtienen de las variables en estudio mediante frecuencias y proporciones con el uso de las medidas de tendencia central y de dispersión.

Medidas de tendencia central

Las *medidas de tendencia central* son los valores estadísticos representativos de la distribución del conjunto de datos en estudio. Las principales medidas de tendencia central son:



La *media aritmética o promedio* se define como el valor obtenido de la suma de todos los valores observados, dividido entre el número total de observaciones. Es la medida más utilizada para variables cuantitativas. Por ejemplo, en la siguiente secuencia de números, la media es de:

8	6	7	6
6	9	9	7
9	9	5	5
Media = $\frac{\text{Suma de los valores observados}}{\text{Número total de observaciones}}$ Media = $\frac{8+6+7+6+6+9+9+7+9+9+5+5}{12} = 7.16$			

La *mediana* se define como el valor que se encuentra exactamente a la mitad de los datos observados y los separa en dos partes iguales. En el caso de que la distribución de los datos sea simétrica la mediana coincidirá con la media. Por ejemplo en la siguiente secuencia de números, la mediana es de:

8	6	7	6
6	9	9	7
9	9	5	5
5, 5, 6, 6, 6, 7, 7, 8, 9, 9, 9, 9 Posición de la mediana = $\frac{n + 1}{2}$ Mediana = 7			

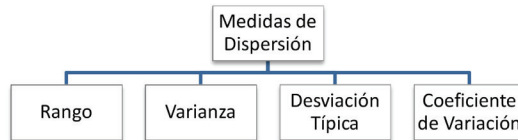
La *moda* se define como el valor que más se repite dentro del conjunto de datos observados. Por ejemplo en la siguiente secuencia de números, la moda es de:

8	6	7	6
6	9	9	7
9	9	5	5
Moda = 9			

En el caso de que la distribución de los datos observados sea normal, los valores de la media, la mediana y la moda coincidirán.

Medidas de dispersión

Las *medidas de dispersión* indican qué tanto se encuentran agrupados o dispersos los valores en relación con la media. Las principales medidas de dispersión son:



El *rango o recorrido* es la diferencia entre el valor máximo y mínimo de la distribución de las variables. Por lo tanto mientras mayor sea el rango, mayor será la dispersión de las variables. Por ejemplo, en la siguiente secuencia de números, el rango es de:

8	6	7	6
6	9	9	7
9	9	5	5
$X_M - X_m$ Rango = 9 - 5 Rango = 4			

La *varianza* es la media del cuadrado de las desviaciones en relación con la media aritmética. Entre mayor sea la varianza, mayor será la dispersión. Por lo tanto expresa la dispersión en la unidad de medida al cuadrado de la variable estudiada.

$$S^2 = \frac{\Sigma ((X - \mu))^2}{n}$$

La *desviación típica o desviación estándar* es la medida de dispersión más utilizada. Es la raíz cuadrada de la varianza. Expresa la dispersión de la distribución en las mismas unidades de medida de la variable en estudio.

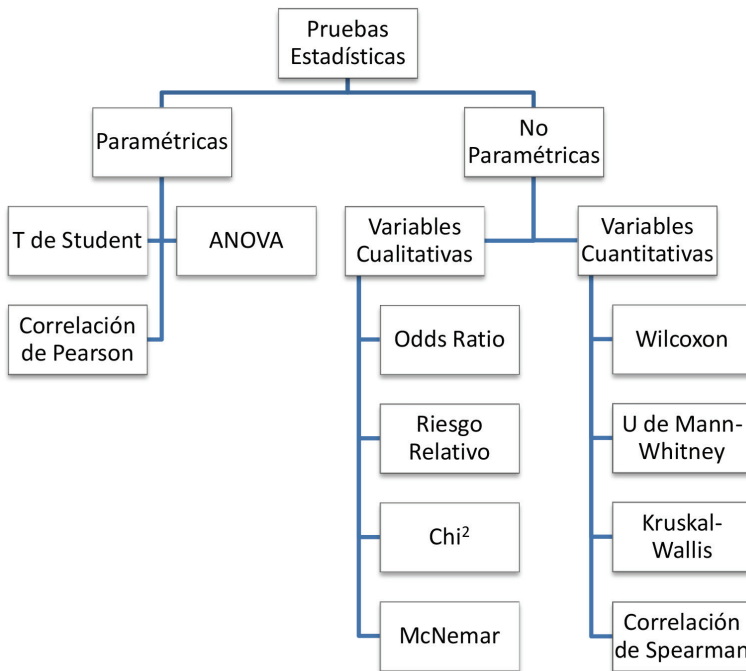
$$S = \sqrt{S^2}$$

El *coeficiente de variación* no posee unidades de medida por lo que es muy útil para comparar la dispersión de variables con distintas unidades de medida. Se calcula al dividir la desviación típica de la muestra entre la media y multiplicando el resultado por 100.

$$CV = \frac{s}{x} (100)$$

Estadística inferencial

La inferencia estadística permite establecer conclusiones a través de los resultados del análisis estadístico mediante los cuales se aceptarán o rechazarán las hipótesis.



Pruebas paramétricas

Las *pruebas paramétricas* son las de mayor potencia estadística y las que más información aportan, ya que permiten realizar una estimación de los parámetros de la población a través de muestras estadísticas.

Para el uso de las pruebas paramétricas se debe contar con ciertas características:

- Variables con *distribución normal*, la cual se determina mediante la prueba de Kolmogorov-Smirnov y la prueba de Shapiro-Wilk.
- No se pueden utilizar para variables ordinales.
- Tamaño de muestra mínimo de 30 personas por grupo, ya que mientras más grande sea la muestra más exacta será la estimación probabilística.
- La hipótesis se realiza sobre valores numéricos, especialmente el promedio o media de la población.

Pruebas no paramétricas

Las *pruebas no paramétricas o de distribución libre* son de fácil aplicación y se utilizan para analizar las variables ordinal o nominal cuando no se requiere plantear inferencias sobre los parámetros de la población.

Para el uso de las pruebas no paramétricas se debe contar con ciertas características:

- Variables de tipo ordinal o nominal.
- Se utiliza cuando dos series de observaciones provienen de distintas poblaciones.
- El tamaño de la muestra puede ser ≤ 20 personas.
- La hipótesis se realiza sobre rangos, mediana o frecuencia de los datos.

La selección de la prueba estadística depende del tipo de variables, de la presencia o ausencia de normalidad en la distribución de los grupos a analizar y de la pregunta de investigación que se pretende responder.

De acuerdo con lo anterior, se formuló la siguiente tabla para seleccionar la prueba estadística más adecuada:

<i>Variable</i>		<i>Prueba Paramétrica</i>	<i>Prueba No Paramétrica</i>
<i>Muestras Independientes ó No Relacionadas</i>			
Dependiente e Independiente de tipo Cualitativa			Chi ²
Dependiente de tipo Cualitativa Independiente de tipo Cuantitativa	Dicotómica	T de Student	Wilcoxon U de Mann-Whitney
	>2 Variables	ANOVA	KrusKal-Wallis
Dependiente de tipo Cuantitativa Independiente de tipo Cuantitativa		Correlación de Pearson	Correlación de Spearman
<i>Muestras Dependientes, Relacionadas ó Pareadas</i>			
Dependiente e Independiente de tipo Cualitativa			McNeman
Dependiente e Independiente de tipo Cuantitativa	2 Medidas	T de Student Pareada	Wilcoxon Pareado
	>2 Medidas	ANOVA Pareada	

Representación gráfica del análisis estadístico

La calidad de la presentación de los resultados del análisis estadístico es de suma importancia para su adecuada comprensión e interpretación.

Diagrama de barras

El diagrama de barras es adecuado para *variables cuantitativas discretas y cualitativas en escala nominal*.

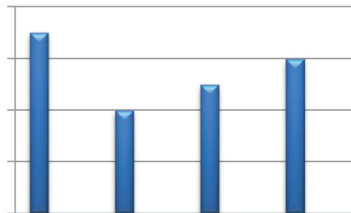
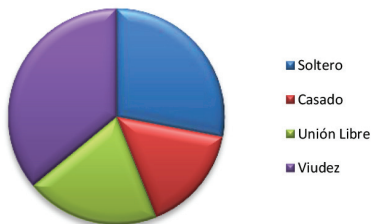


Diagrama de pastel

El diagrama de pastel es adecuado para *variables cualitativas en escala nominal*.



Histograma

El histograma es adecuado para *variables cuantitativas continuas* debido a que agrupa a la frecuencia de presentación de las variables en intervalos.

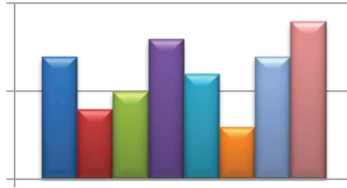


Diagrama de caja

El diagrama de caja es adecuado para *variables cuantitativas* con desviación de la *distribución normal*. Para su interpretación, el extremo superior de la caja corresponde al percentil 75, el extremo inferior corresponde al percentil 25 y la línea horizontal que divide la caja corresponde a la media.

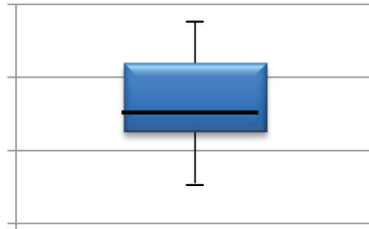


Diagrama de tallo y hojas

El diagrama de tallo y hojas es adecuado para *variables cuantitativas* de pocos valores, debido a que en este diagrama cada uno de los valores de las variables de estudio se agrupan en relación con su distribución.


Tallo	Hoja
1	3 3 4 6
2	1 1 7 7 7 8
3	1 4 4 6 8
4	1 2 2 5 5 7 7 8

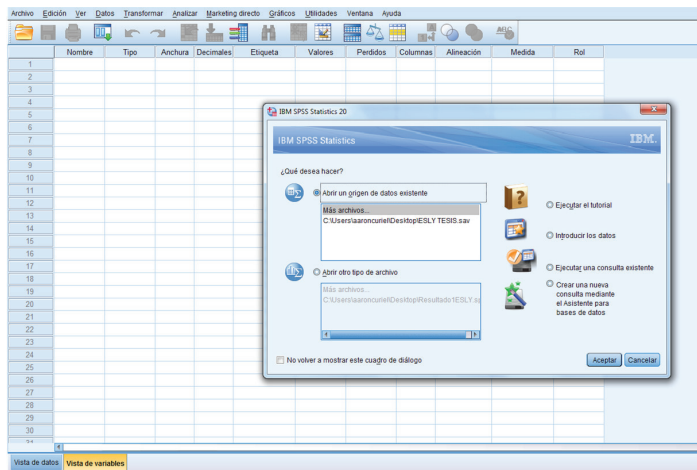
Base de datos

Se entiende como *base de datos* al conjunto de información perteneciente a un mismo contexto dentro de nuestra investigación.

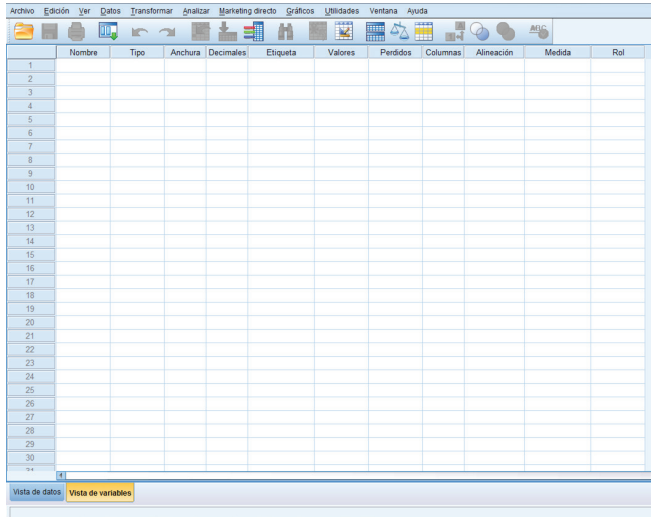
Para crear la base de datos en el programa SPSS versión 20.0 se realizarán los siguientes pasos:

Paso 1

Abra el programa  IBM SPSS Statistics 20 y aparecerá la ventana IBM SPSS Statistics 20, donde podrá identificar las bases de datos existentes. En caso de no haber trabajado previamente alguna base de datos, dé *click* en *Cerrar* para continuar.

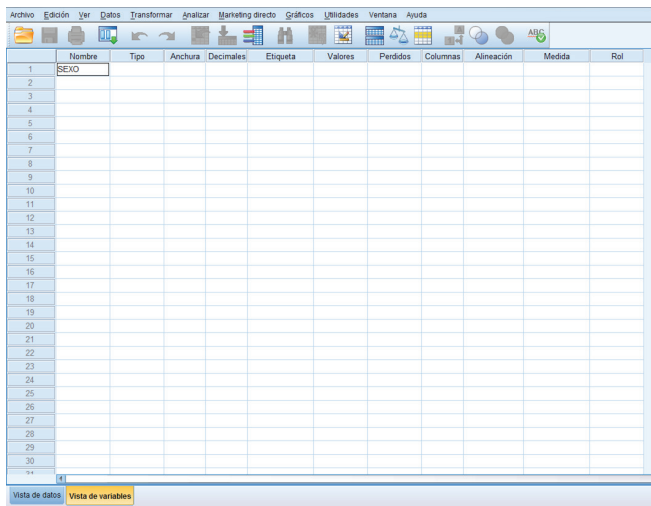


Una vez que haya cerrado la ventana anterior, podrá visualizar la pantalla de esta forma (sin base de datos).

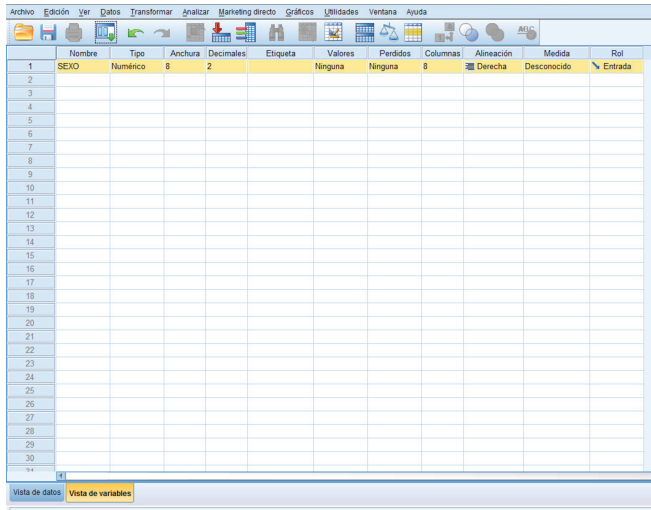


Paso 2

Para crear su base de datos dé *click* en la pestaña de *Vista de variables* y escriba su primera variable en la *Celda de nombre*, por ejemplo *Sexo*. Recuerde que en la *Columna de nombre* se escribe una variable por fila, de forma abreviada y sin espacio entre los caracteres.

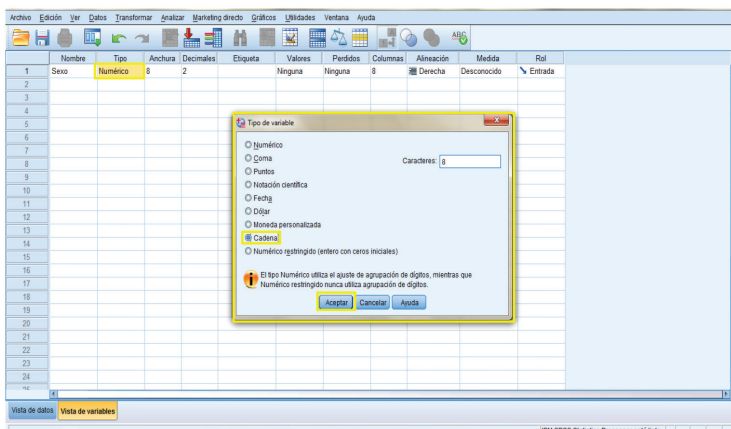


Se completa automáticamente el resto de la fila.



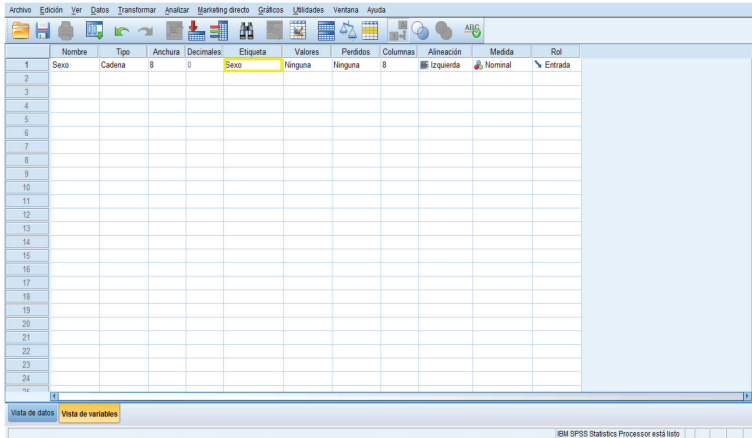
Paso 3

Dé click en la Celda de tipo y aparecerá la ventana de Tipo de variable. Siguiendo con el ejemplo de la variable de Sexo, al ser de tipo cualitativa tipifique como Cadena y posteriormente dé click en Aceptar.



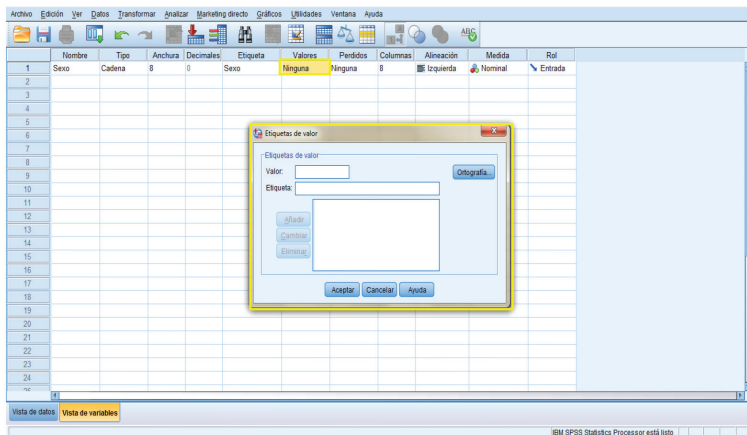
Paso 4

En la *Celda de etiquetas* escriba como quiera que aparezca la variable en los gráficos y tablas de resultados.

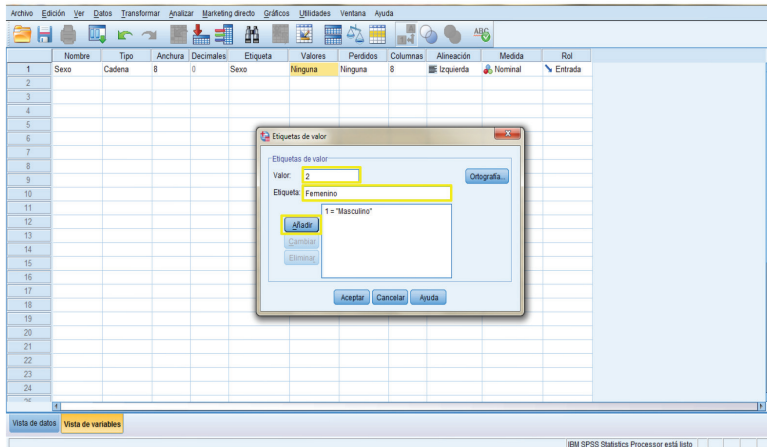


Paso 5

Dé click en la *Celda de valores* y aparecerá la ventana de *Etiquetas de valor*.

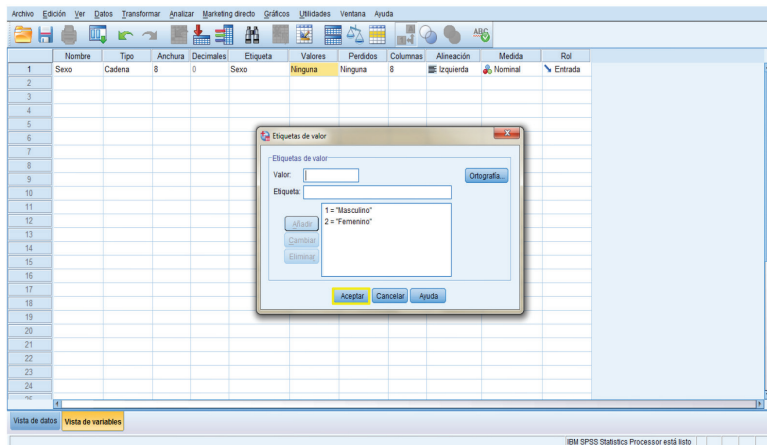


Agregue los valores y las etiquetas correspondientes a la variable de estudio y dé *click* en *Añadir* entre cada valor de etiqueta agregado. Por ejemplo 1 para masculino y 2 para femenino.



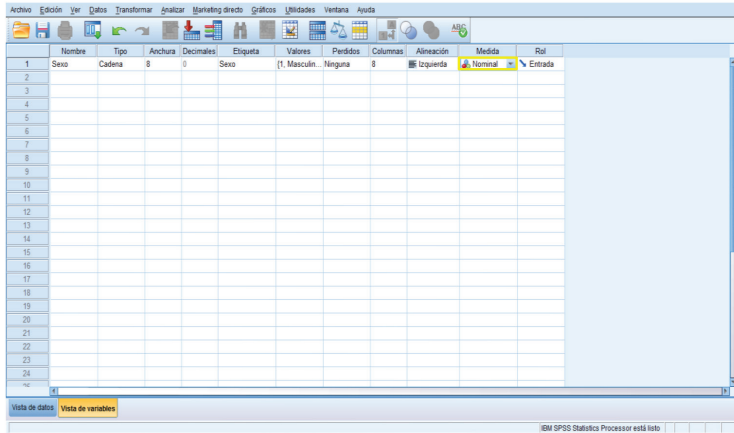
Paso 6

Cuando tenga todas las etiquetas de valor dé *click* en *Aceptar*.



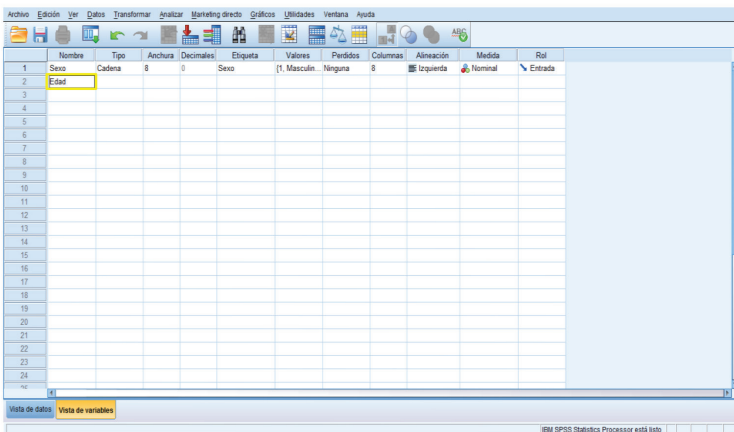
Paso 7

En relación con las características de la variable tipifique la medida. En este caso la variable de *Sexo* al ser de tipo cualitativa dicotómica se tipifica como *Nominal*.

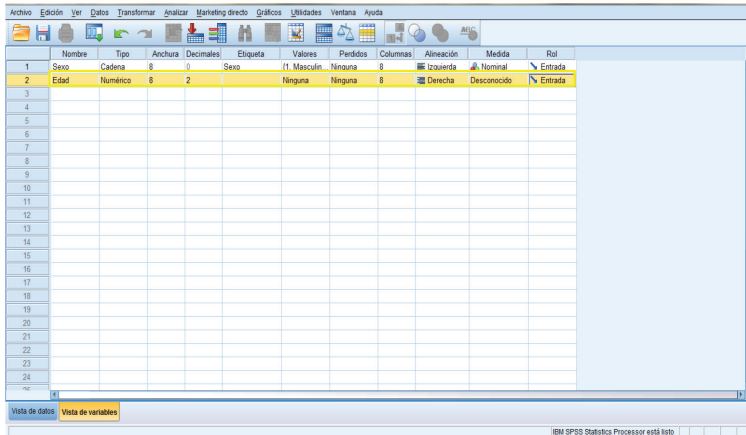


Paso 8

En la segunda fila de la *Celda de nombre* escriba la siguiente variable, por ejemplo *Edad*.

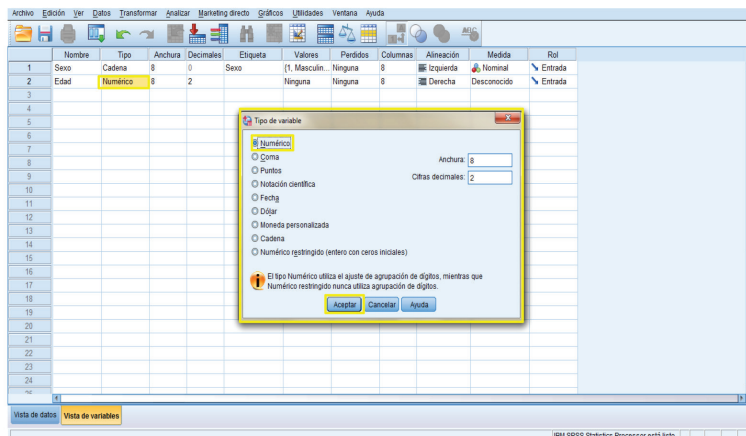


Se completa automáticamente el resto de la fila.



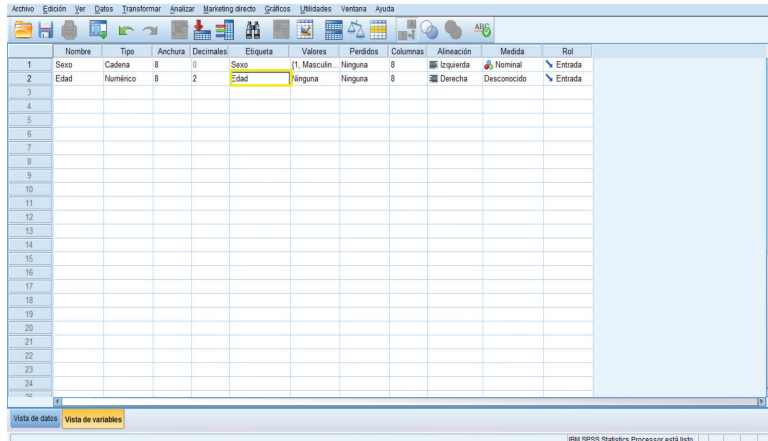
Paso 9

Dé click en la Celda de tipo y aparecerá la ventana de *Tipo de variable*. Siguiendo con el ejemplo de la variable de *Edad*, al ser de tipo cuantitativa tipifique como *Numérico* y posteriormente dé click en *Aceptar*.



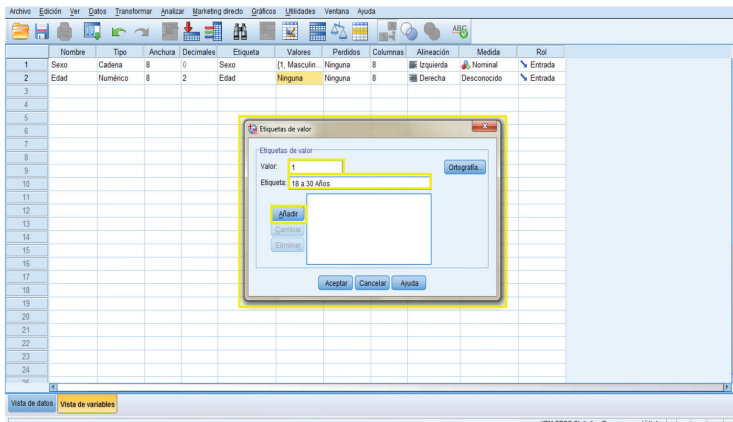
Paso 10

En la *Celda de etiquetas* escriba como quiera que aparezca la variable en los gráficos y tablas de resultados.

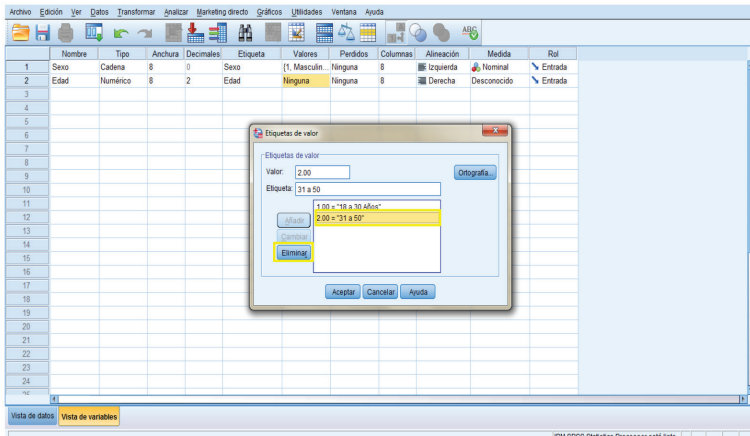


Paso 11

Dé click en *Celda de valores* para que aparezca la ventana de *Etiquetas de valor* y agregue los valores y las etiquetas correspondientes a la variable de estudio y dé click en *Añadir* entre cada valor de etiqueta agregado.

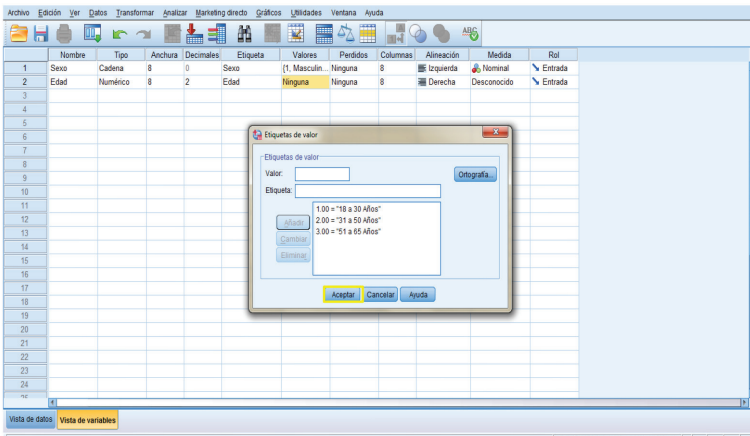


En el caso de equivocarse en alguna etiqueta de valor selecciónela y dé *click* en *Eliminar*.



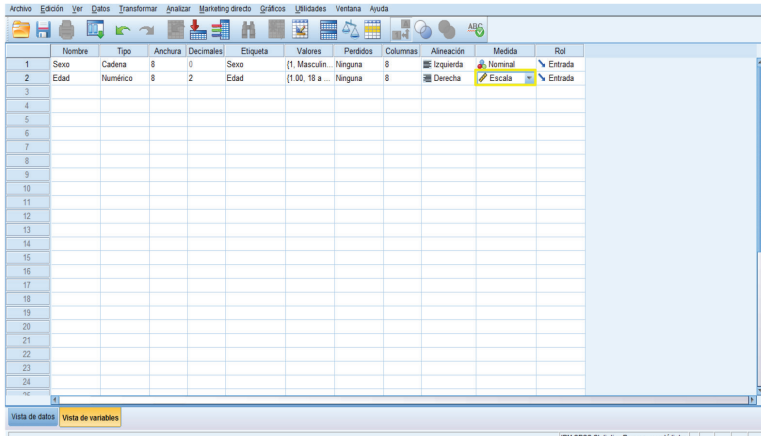
Paso 12

Cuando tenga todas las etiquetas de valor dé *click* en *Aceptar*.



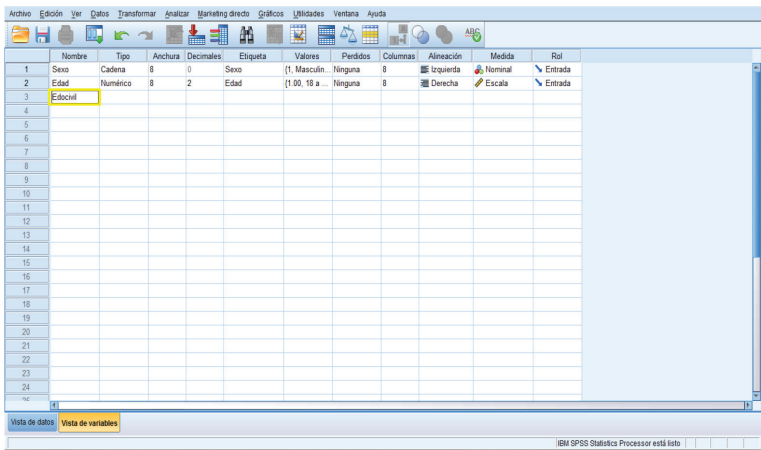
Paso 13

En relación con las características de la variable, tipifique la medida. En este caso la variable de *Edad* al ser de tipo cuantitativa se tipifica como *Escala*.

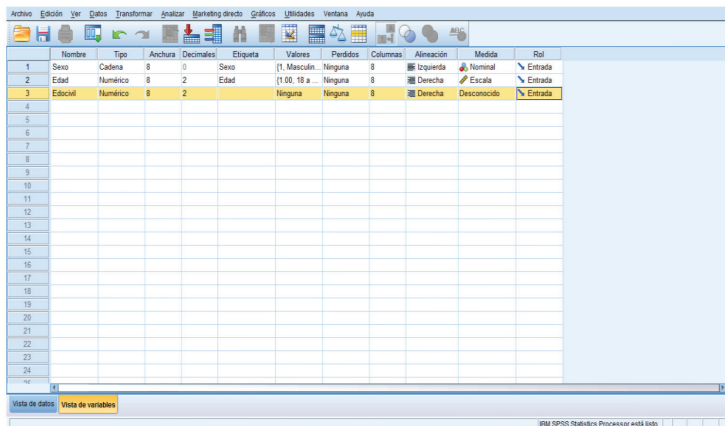


Paso 14

En la tercera fila de la *Celda de nombre* escriba la siguiente variable, por ejemplo *Estado Civil (edocivil)*, recuerde que en la *Columna de nombre* se escribe una variable por fila, de forma abreviada y sin espacio entre los caracteres.

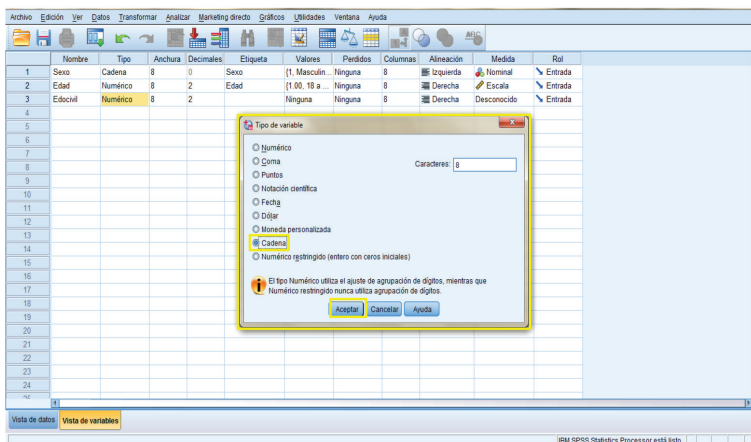


Se completa automáticamente el resto de la fila.



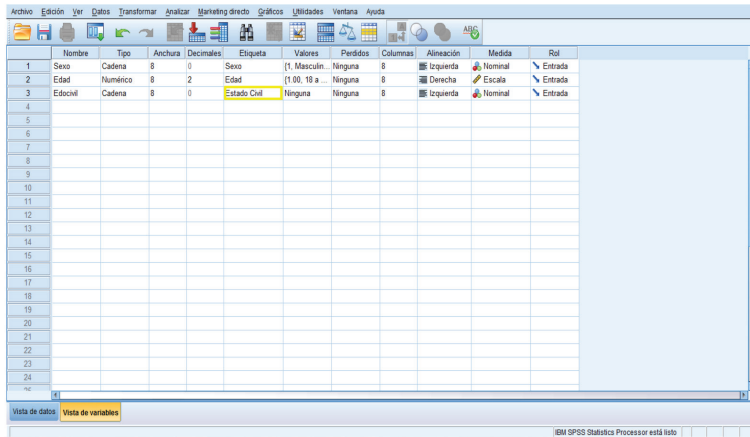
Paso 15

Dé click en la Celda de tipo y aparecerá la ventana de Tipo de variable. Siguiendo con el ejemplo de la variable de estado civil, al ser de tipo cualitativa tipifique como Cadena y posteriormente dé click en Aceptar.



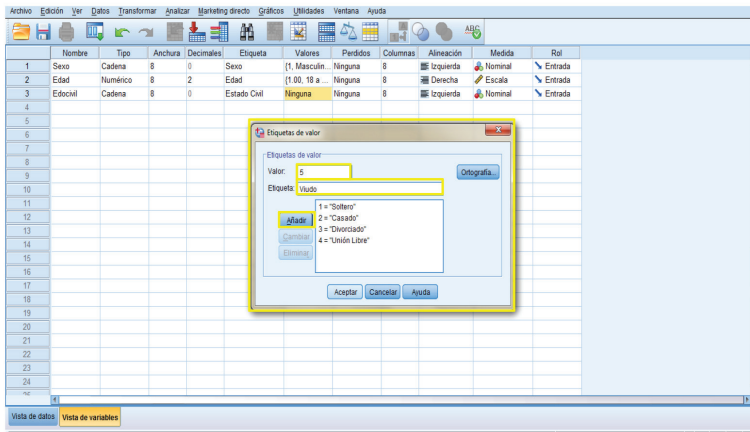
Paso 16

En la *Celda de etiquetas* escriba como quiera que aparezca la variable en los gráficos y tablas de resultados.



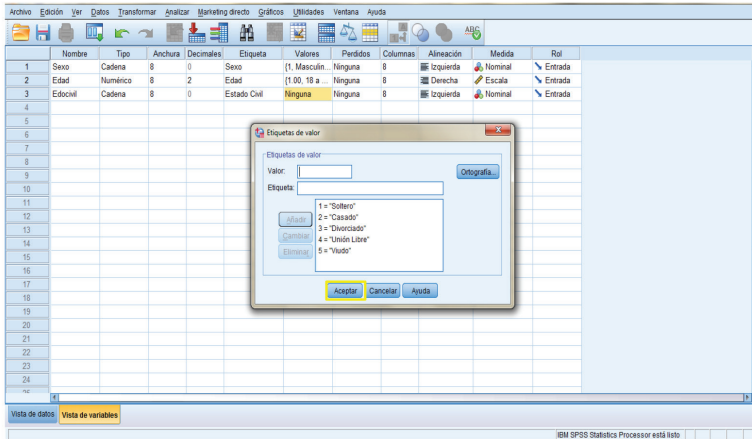
Paso 17

Dé *click* en *Celda de valores* para que aparezca la ventana de *Etiquetas de valor* y agregue los valores y las etiquetas correspondientes a la variable de estudio y dé *click* en *Añadir* entre cada valor de etiqueta agregado.



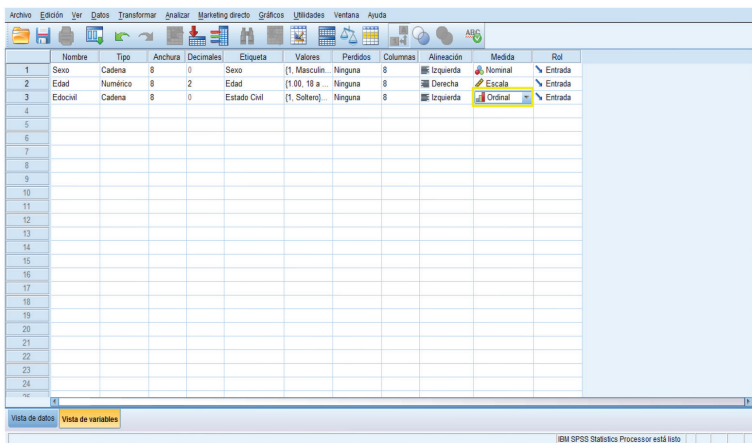
Paso 18

Cuando tenga todas las etiquetas de valor, dé *click* en *Aceptar*.



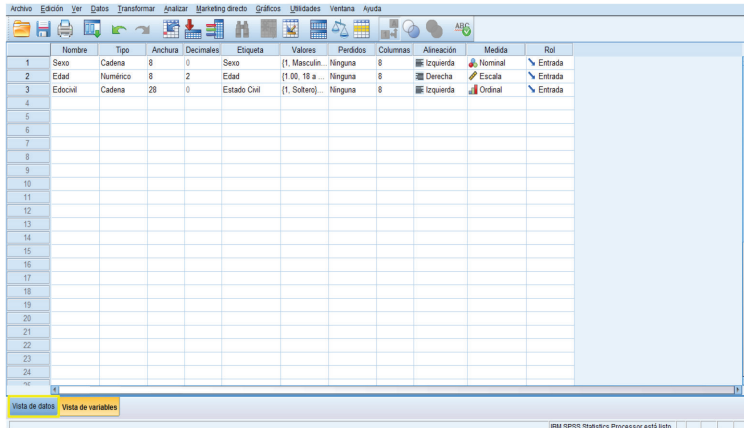
Paso 19

En relación con las características de la variable, tipifique la medida. En este caso la variable de estado civil, al ser de tipo cualitativa, se tipifica como *Ordinal*.



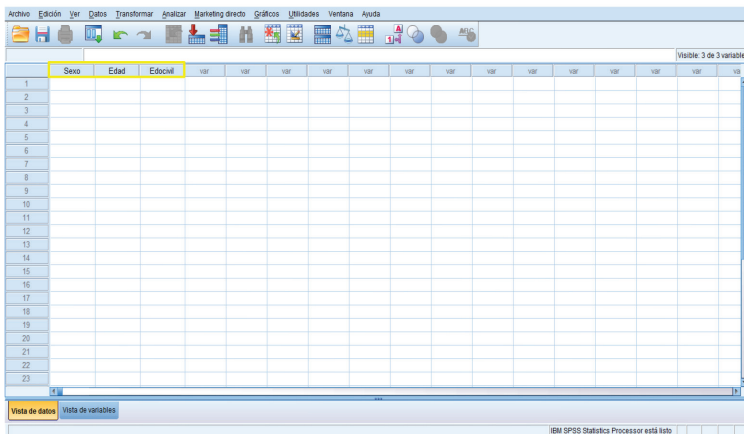
Paso 20

Una vez que haya capturado todas las variables de estudio que le interesen, cambie la visualización de la pantalla a *Vista de datos*.



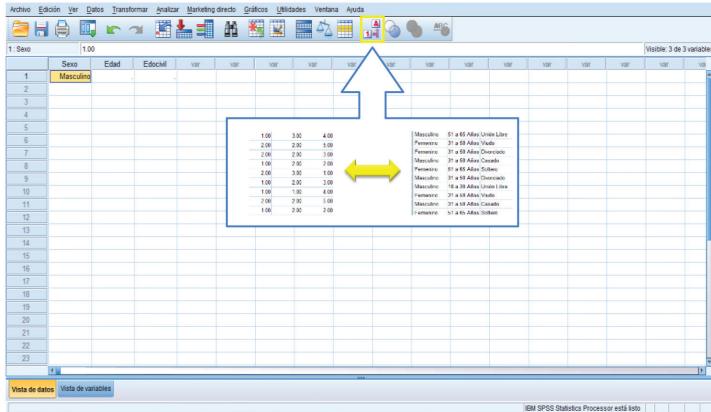
Paso 21

En la pantalla a *Vista de datos* se observan las variables que capturó previamente.



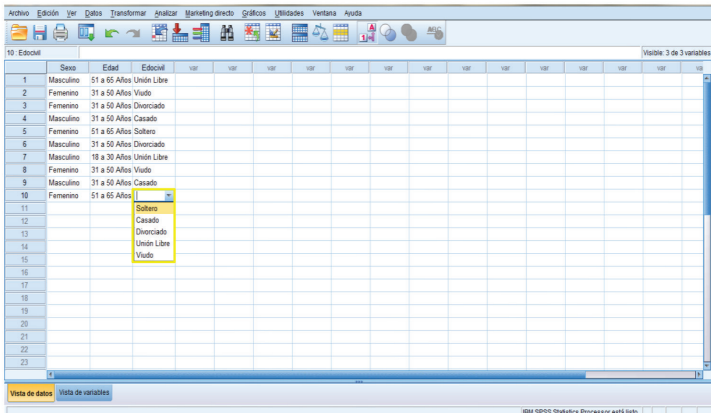
Paso 22

Capture su información. En caso que se le dificulte la captura con el valor numérico de la etiqueta, dé *click* en el icono de *Etiquetas de valor* para cambiar la captura de datos.



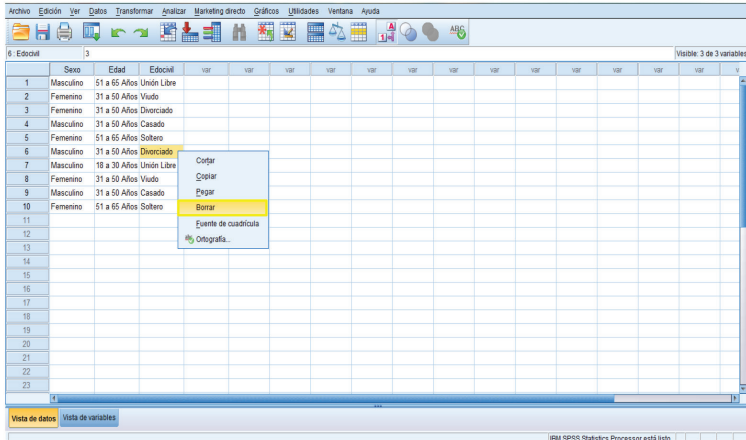
Paso 23

Si da *click* en alguna celda podrá observar las opciones de etiquetas correspondientes a la variable de la columna seleccionada.



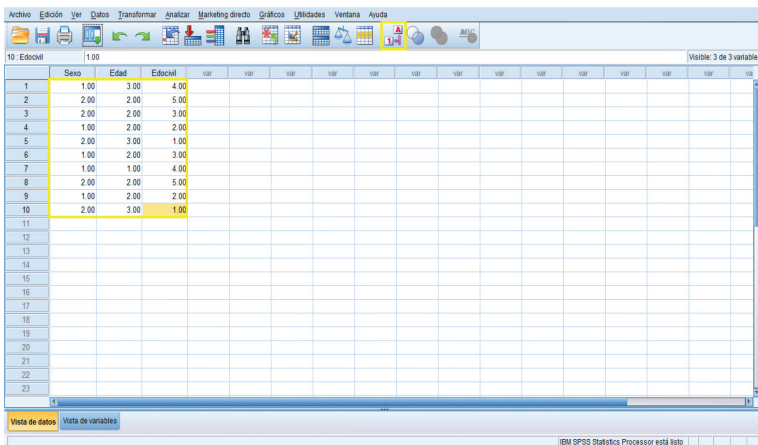
Paso 24

En caso que se equivoque, seleccione la etiqueta errónea, presione el botón izquierdo del mouse y aparecerá un cuadro de texto donde debe dar *click* en *Borrar*.



Paso 25

Si desea observar el valor numérico de la etiqueta, dé *click* en el icono de *Etiquetas de valor*.



Pruebas de normalidad

La normalidad de la distribución de la muestra es de utilidad para discernir entre el uso de media, mediana, desviación estándar o rango para el análisis estadístico descriptivo y el uso de pruebas paramétricas o no paramétricas para el análisis estadístico inferencial.

La prueba de Kolmogorov-Smirnov o K-S es la más utilizada y la más potente para evaluar la normalidad de distribución de tamaños de muestra mayores de 30 ($n > 30$).

$$Dn = |Fn((x)) - F((x))|$$

En donde:

$F_n(x)$ = Frecuencia o distribución teórica

$F(x)$ = Frecuencia o distribución observada

Cuando el tamaño de la muestra es menor a 30 ($n < 30$), la prueba de Shapiro-Wilks es la indicada para el análisis de la normalidad de distribución.

$$W = \frac{(\sum_{i=1}^n \alpha_i x(i))^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

En donde:

Σ = Sumatoria

$x(i)$ = Número que ocupa la i -ésima posición en la muestra

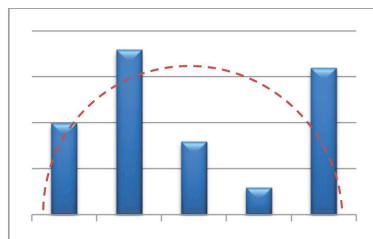
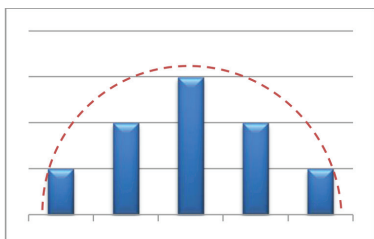
\bar{x} = Media muestral

En relación con la distribución de la muestra se aceptará o rechazará la hipótesis:

- Hipótesis Nula (H_0): establece que la muestra sigue una distribución normal
- Hipótesis Alternativa (H_1): establece que la muestra no sigue una distribución normal.

En relación con la presencia o ausencia de normalidad será el tipo de análisis estadístico descriptiva e inferencial a realizar:

- Un valor de $p > 0.05$ indica una muestra con *distribución normal* y por lo tanto el uso de *media y desviación estándar* para su análisis descriptivo y se sugiere el uso de *pruebas paramétricas* para su análisis inferencial.
- Un valor de $p \leq 0.05$ indica una muestra con *ausencia de normalidad en la distribución* y por lo tanto el uso de *mediana y rango* para su análisis descriptivo y se sugiere el uso de *pruebas no paramétricas* para su análisis inferencial.



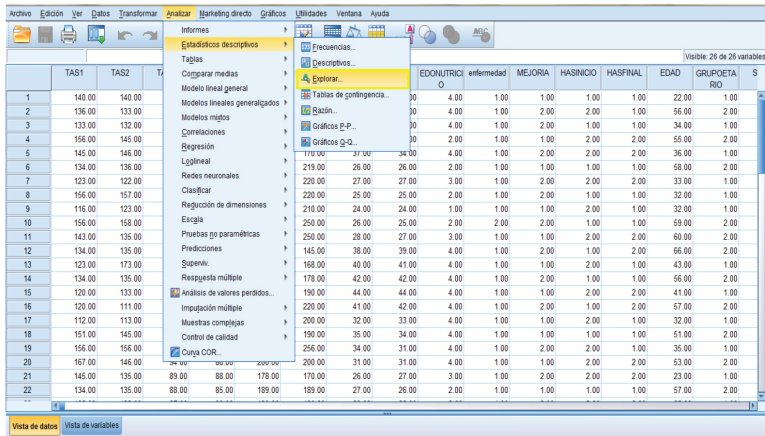
Valor de $p > 0.05$	Valor de $p \leq 0.05$
<i>Estadística Descriptiva</i>	
Media / Desviación Estándar	Mediana / Rango
<i>Estadística Inferencial</i>	
Pruebas Paramétricas	Pruebas No Paramétricas

Cuando el tamaño de muestra es menor a 50, el programa SPSS calcula automáticamente la prueba de Kolmogorov-Smirnov y la prueba de Shapiro-Wilks.

Para realizar el análisis de la normalidad de la distribución de la muestra mediante la prueba de Kolmogorov-Smirnov en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

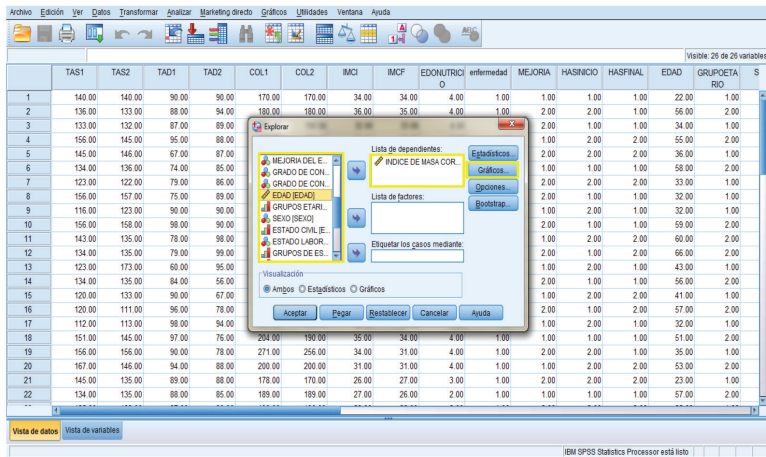
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Explorar*.



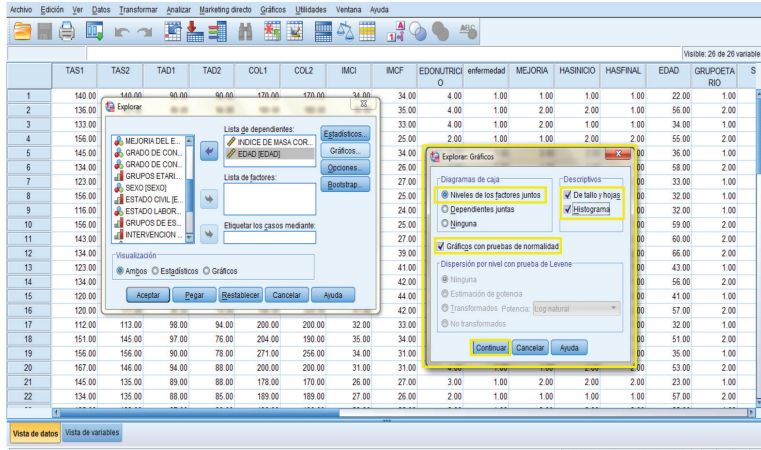
Paso 2

Aparecerá la ventana de *Explorar*, seleccione e introduzca las variables cuantitativas que le interesen en el recuadro de *Variable* y posteriormente dé *click* en *Gráficos*.



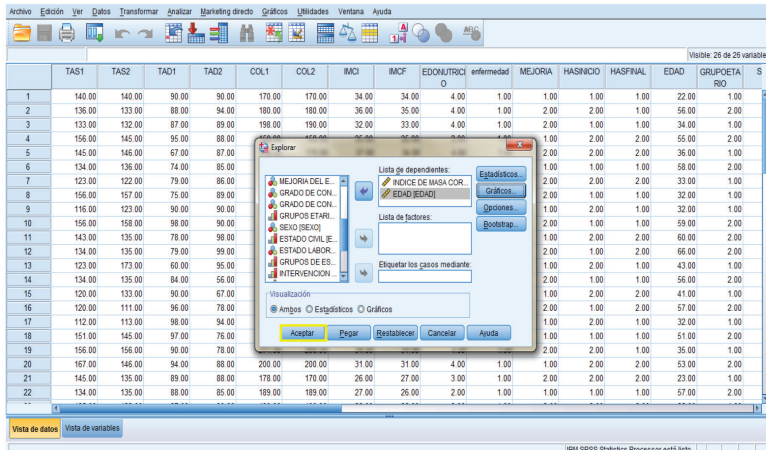
Paso 3

Aparecerá la ventana de *Explorar: Gráficos*, dé click en *Niveles de factores juntos*, en *Gráficos con pruebas de normalidad*, en *De tallo y hoja*, en *Histograma* y por último en *Continuar*.



Paso 4

Aparecerá nuevamente la ventana de *Explorar*, donde deberá dar click en *Aceptar*.



Paso 5

De esta forma se obtiene el análisis principal de las variables analizadas.

Resumen del procesamiento de los casos

	Casos				Total	Porcentaje
	Válidos	Perdidos	N	Porcentaje		
INDICE DE MASA CORPORAL INICIAL	50	0	50	100.0%	50	100.0%
EDAD	50	0	50	100.0%	50	100.0%

Descriptivos

		Estadístico	Error tp.
INDICE DE MASA CORPORAL INICIAL	Media	30.8200	.85365
	Intervalo de confianza para la media al 95%	29.1045	
		Limite inferior	32.5355
		Limite superior	
	Media recortada al 5%	30.6444	
	Mediana	30.0000	
	Varianza	36.436	
	Desv. tp.	6.03625	
	Mínimo	21.00	
	Máximo	44.00	
EDAD	Media	45.5800	2.05160
	Intervalo de confianza para la media al 95%	41.4572	
		Limite inferior	49.7028
		Limite superior	
	Media recortada al 5%	45.4333	
	Mediana	48.0000	
	Varianza	210.453	
	Desv. tp.	14.50699	
	Mínimo	22.00	
	Máximo	78.00	

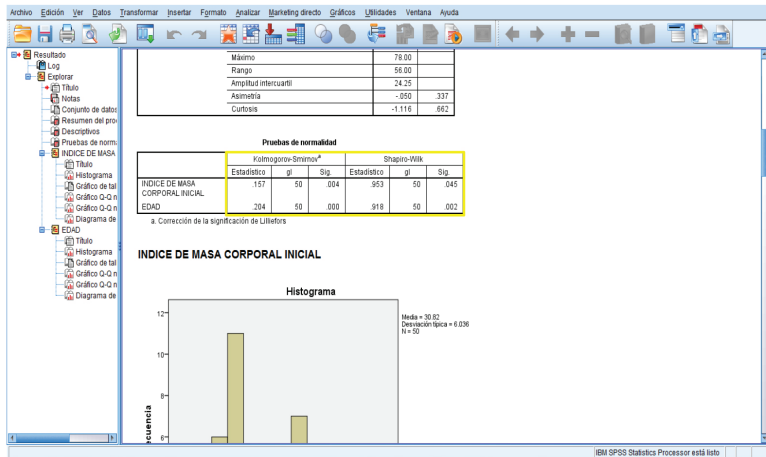
Al bajar el cursor, se observa la *Media* y el *Intervalo de Confianza a 95%* de las variables seleccionadas.

Descriptivos

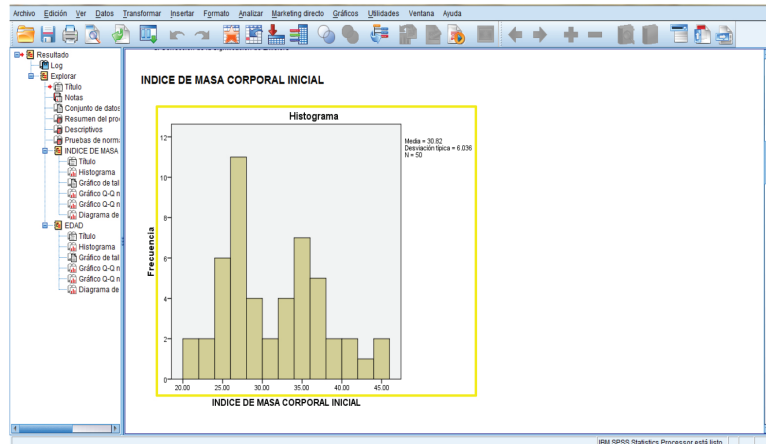
		Estadístico	Error tp.
INDICE DE MASA CORPORAL INICIAL	Media	30.8200	.85365
	Intervalo de confianza para la media al 95%	29.1045	
		Limite inferior	32.5355
		Limite superior	
	Media recortada al 5%	30.6444	
	Mediana	30.0000	
	Varianza	36.436	
	Desv. tp.	6.03625	
	Mínimo	21.00	
	Máximo	44.00	
EDAD	Media	45.5800	2.05160
	Intervalo de confianza para la media al 95%	41.4572	
		Limite inferior	49.7028
		Limite superior	
	Media recortada al 5%	45.4333	
	Mediana	48.0000	
	Varianza	210.453	
	Desv. tp.	14.50699	
	Mínimo	22.00	
	Máximo	78.00	

Así como el resultado de la prueba de normalidad de Kolmogorov-Smirnov y de Shapiro-Wilk.

Recuerde que la *significancia* marcará la pauta para el uso entre *pruebas paramétricas* y *pruebas no paramétricas*.

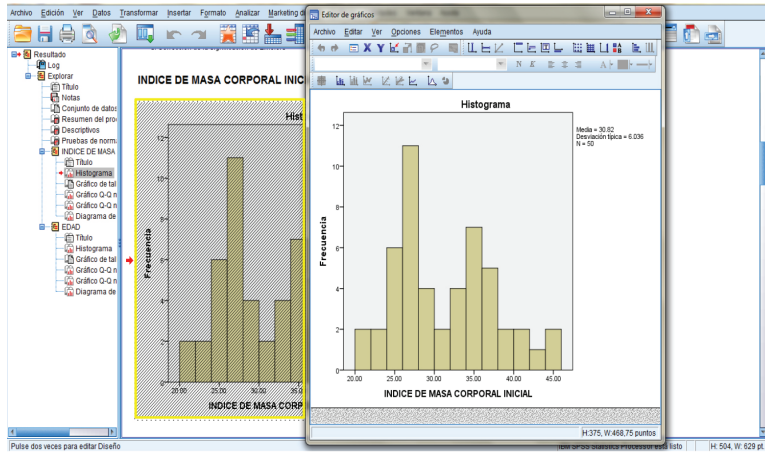


Se observan los *Histogramas* de las variables seleccionadas



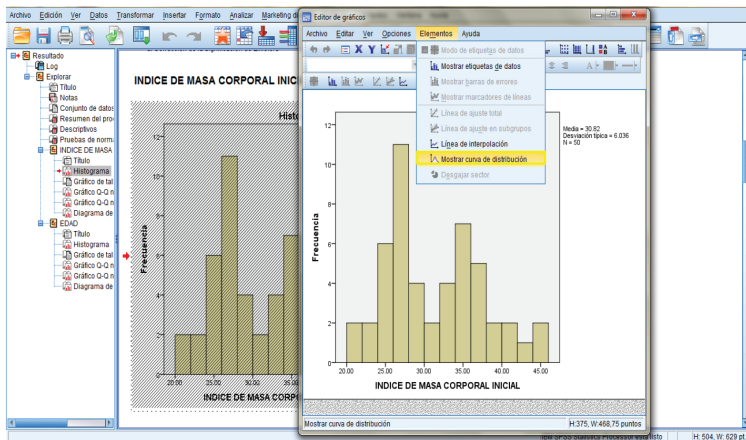
Paso 6

Para editar el *Histograma* y valorar su normalidad, dé *doble click* sobre la imagen del *Histograma* y aparecerá la ventana de *Editor de gráficos*.



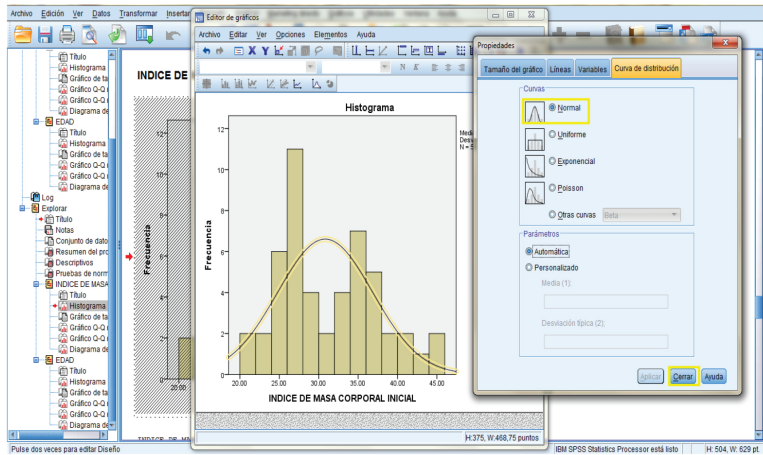
Paso 7

Dé *click* en *Elementos* y posteriormente en *Mostrar curva de distribución*.



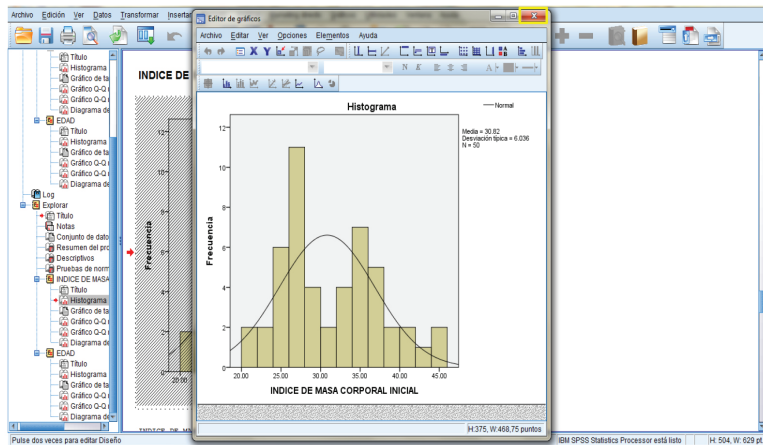
Paso 8

Automáticamente aparecerá la curva de normalidad sobre el *Histograma*, dé click en *Cerrar*.



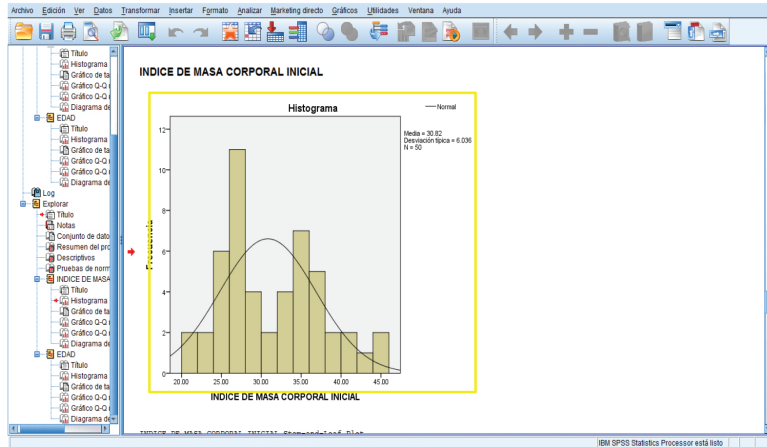
Paso 9

Cierre la ventana de *Editor de gráficos*.



Paso 10

De esta forma obtendrá la *Curva de normalidad* del *Histograma* seleccionado.



Interpretación de resultados

Para valorar la presencia o ausencia de normalidad de la distribución de las variables analizadas y discernir el tipo de prueba estadística a utilizar, debe tomar en cuenta la significancia de las variables analizadas.

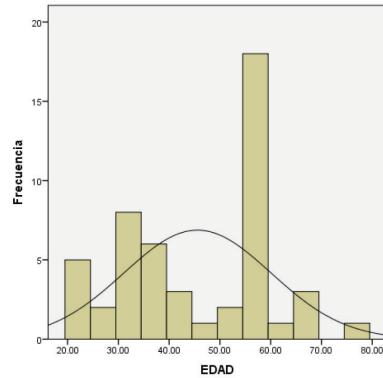
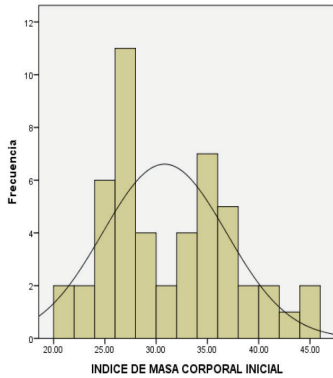
Una significancia con valor de $p > 0.05$ nos indica una muestra con distribución normal, por lo tanto el uso de *media* y *desviación estándar* para su análisis descriptivo y el uso de *pruebas paramétricas* para su análisis inferencial, mientras que un valor de $p \leq 0.05$ indica ausencia de normalidad en la distribución de la muestra, por lo tanto el uso de *mediana* y *rango* para su análisis descriptivo y el uso de *pruebas no paramétricas* para su análisis inferencial.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
INDICE DE MASA CORPORAL INICIAL	.157	50	.004	.953	50	.045
EDAD	.204	50	.000	.918	50	.002

a. Corrección de la significación de Lilliefors

Por ejemplo, en el análisis realizado a través de la prueba de Kolmogorov-Smirnov se obtuvo una significancia de 0.004 para la variable de índice de masa corporal inicial y una significancia de 0.000 para la variable de edad, por lo tanto hablamos de variables con distribución no normal, donde utilizaremos para su análisis descriptivo mediana y rango, mientras que para su análisis inferencial pruebas no paramétricas. Lo cual se confirma al observar la curva de distribución de los histogramas.



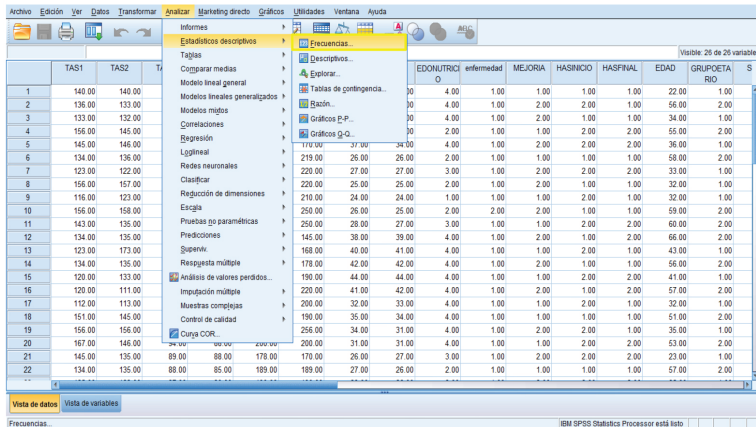
Media

La *media aritmética* o también llamada *promedio* es la medida más utilizada para variables cuantitativas y se define como el valor obtenido de la suma de todos los valores observados, dividido entre el número total de observaciones.

Para obtener la *media* en el programa SPSS versión 20.0 se realizarán los siguientes pasos:

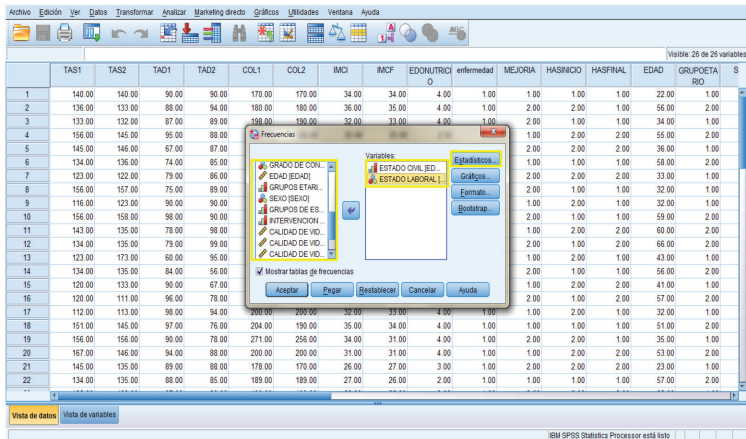
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



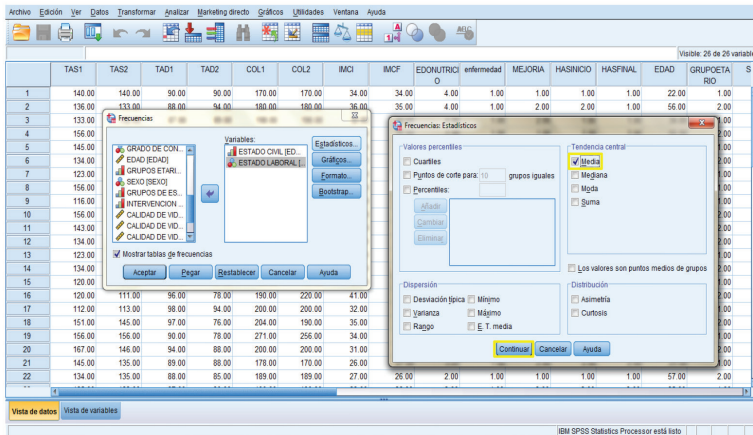
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



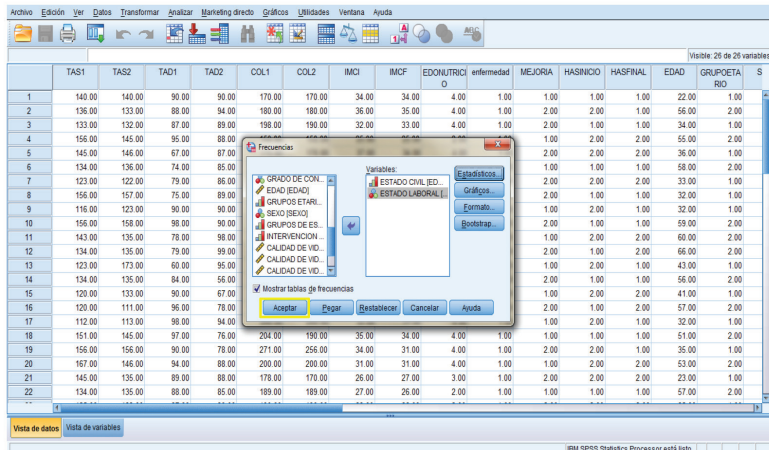
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Media* y posteriormente en *Continuar*.



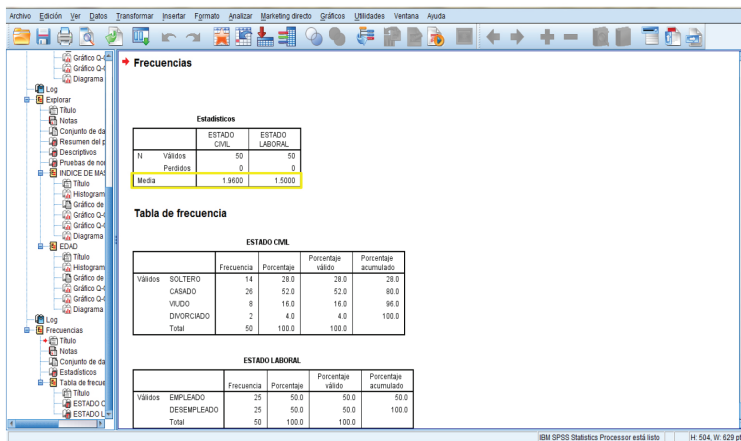
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene la *Media* de las variables analizadas.

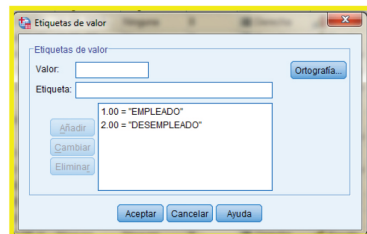
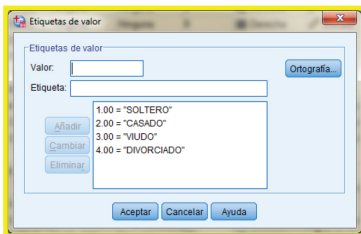


Interpretación de resultados

Para interpretar el resultado de la *Media* que se obtiene mediante el análisis en el programa SPSS versión 20.0, debe tomar en cuenta:

- Los valores que asignó en las etiquetas de valor de sus variables analizadas.

Por ejemplo, para la variable de *Estado civil* se etiquetó del 1 al 4 para soltero, casado, viudo y divorciado, respectivamente. Para la variable de *Estado laboral* se etiquetó del 1 al 2 para *empleado* y *desempleado*, respectivamente.



Estadísticos

		ESTADO CIVIL	ESTADO LABORAL
N	Válidos	50	50
	Perdidos	0	0
Media		1.9600	1.5000

Tabla de frecuencia

ESTADO CIVIL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	SOLTERO	14	28.0	28.0	28.0
	CASADO	26	52.0	52.0	80.0
	VIUDO	8	16.0	16.0	96.0
	DIVORCIADO	2	4.0	4.0	100.0
	Total	50	100.0	100.0	

ESTADO LABORAL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	EMPLEADO	25	50.0	50.0	50.0
	DESEMPLEADO	25	50.0	50.0	100.0
	Total	50	100.0	100.0	



Por ejemplo, en el análisis realizado a través de la prueba T de Student para muestras dependientes o relacionadas se obtuvo una *Media* de 30.82 para la variable de índice de masa corporal inicial, 29.98 para la variable de índice de masa corporal final y una significancia de 0.000 al comparar las medias de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que existe un cambio en el *índice de masa corporal* posterior a la implementación de un régimen alimenticio adecuado, concluimos que:

- Entre el índice de masa corporal inicial y el final hubo una diferencia, debido a que se obtuvo una media de 30.82 para la variable de índice de masa corporal inicial y una media de 29.98 para la variable de índice de masa corporal final.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que hubo una disminución en el índice de masa corporal final o posterior a la intervención.
- Y que esta diferencia fue estadísticamente significativa, debido a que se obtuvo una significancia ≤ 0.05 . Por lo tanto, se acepta la hipótesis alterna.

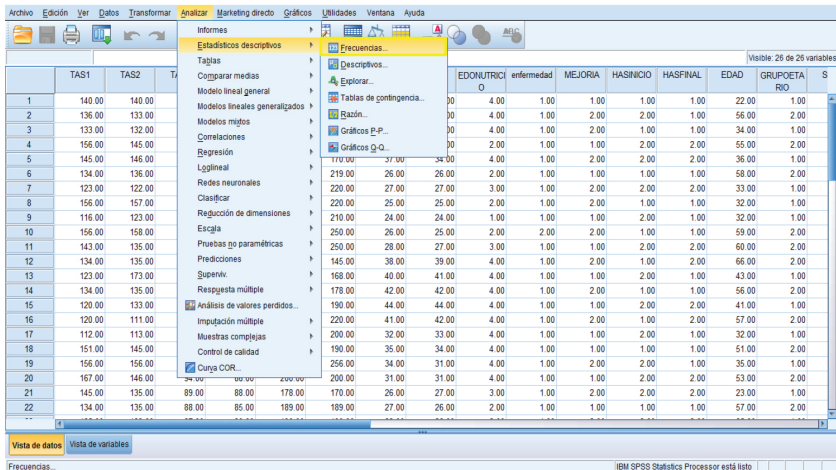
Mediana

Se define a la *mediana* como el valor que se encuentra exactamente a la mitad y separa en dos partes iguales a los datos observados. Cuando la distribución de los datos observados sea simétrica, la mediana coincidirá con la media.

Para obtener la *mediana* en el programa SPSS versión 20.0 se realizarán los siguientes pasos:

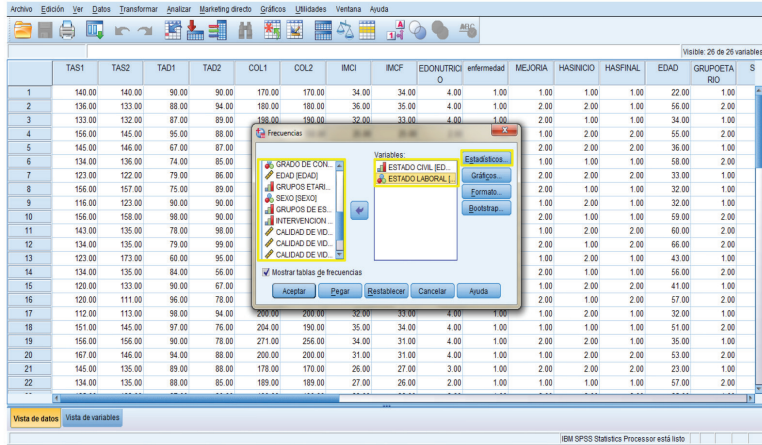
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



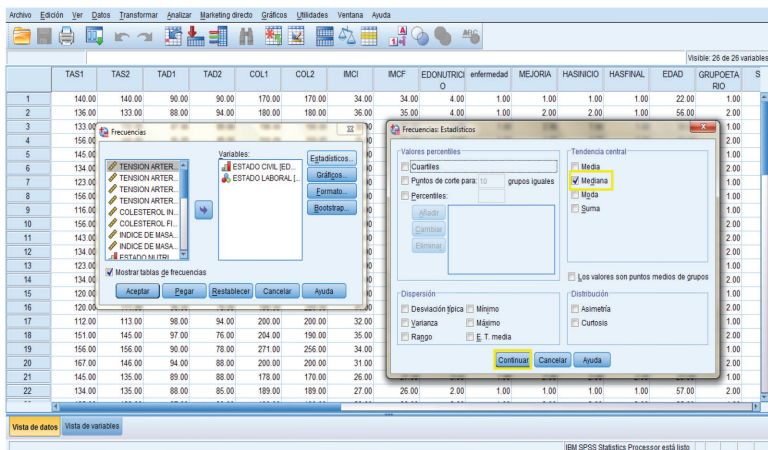
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



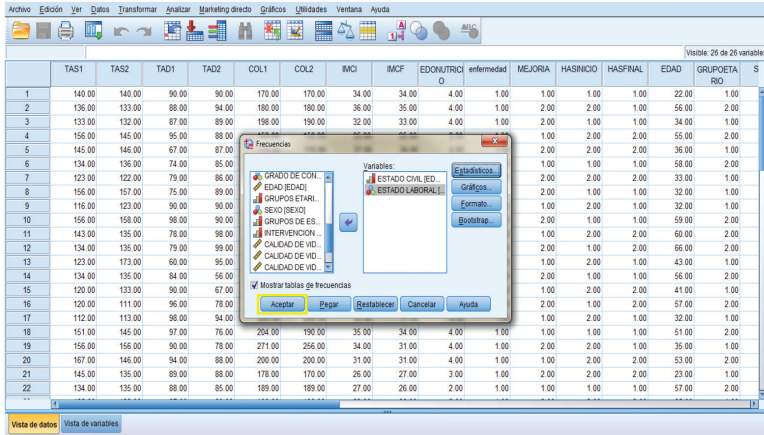
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Mediana* y posteriormente en *Continuar*.



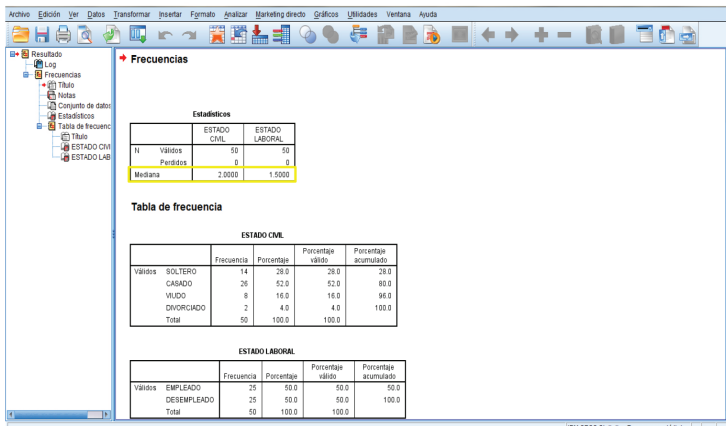
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene la mediana de las variables analizadas.



Interpretación de resultados

- Conjunto de datos
- Estadísticos
- Tabla de frecuenc
- Título
- ESTADO CIVI
- ESTADO LAB

Estadísticos

		ESTADO CIVIL	ESTADO LABORAL
N	Válidos	50	50
	Perdidos	0	0
	Mediana	2.0000	1.5000

Tabla de frecuencia

ESTADO CIVIL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	SOLTERO	14	28.0	28.0	28.0
	CASADO	26	52.0	52.0	80.0
	VIUDO	8	16.0	16.0	96.0
	DIVORCIADO	2	4.0	4.0	100.0
	Total	50	100.0	100.0	

ESTADO LABORAL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	EMPLEADO	25	50.0	50.0	50.0
	DESEMPLEADO	25	50.0	50.0	100.0
	Total	50	100.0	100.0	

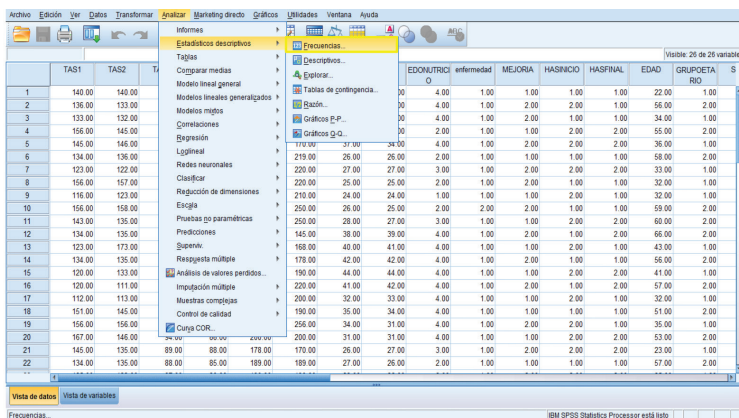
Moda

La *moda* se define como el valor que más se repite dentro del conjunto de datos observados.

Para obtener la *moda* en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

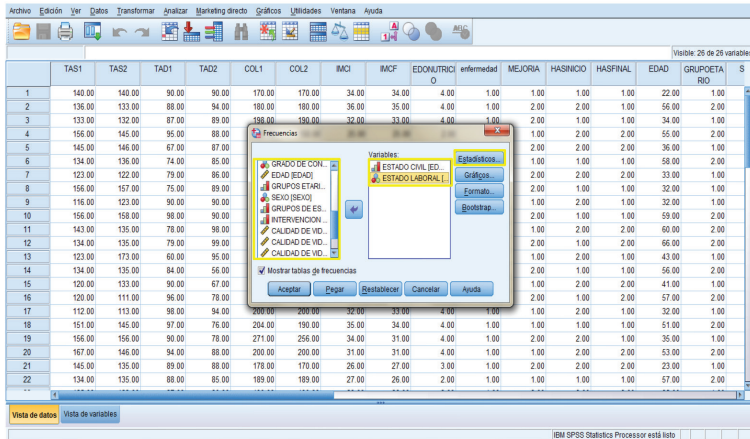
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



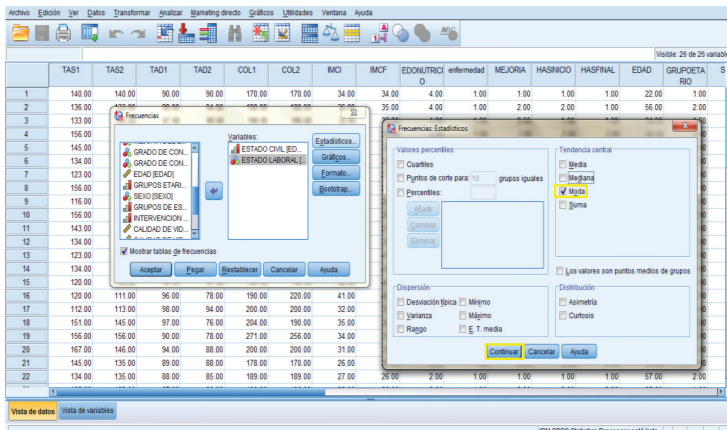
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



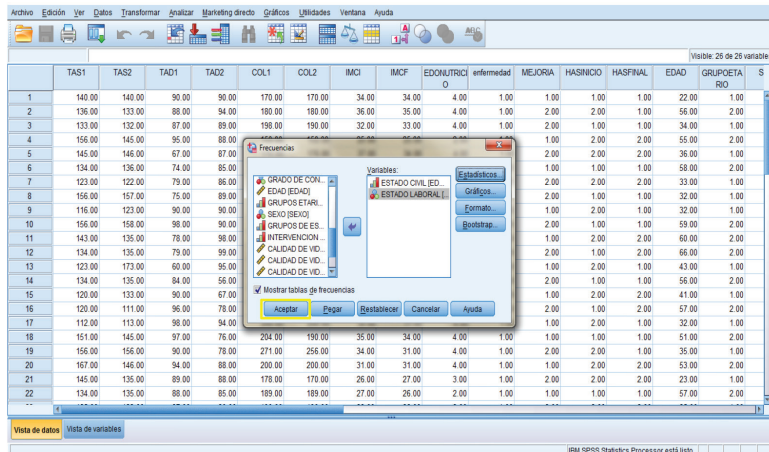
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Moda* y posteriormente en *Continuar*.



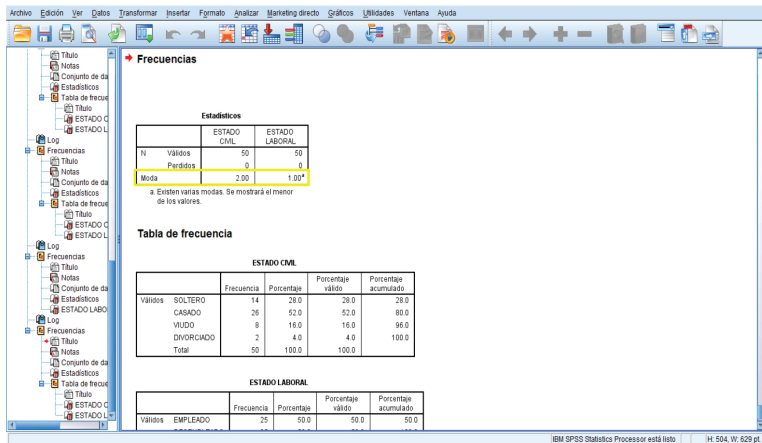
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.

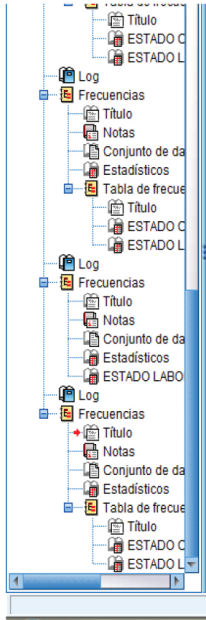


Paso 5

De esta forma se obtiene la *Moda* de las variables analizadas.



Interpretación de resultados



Estadísticos

		ESTADO CIVIL	ESTADO LABORAL
N	Válidos	50	50
	Perdidos	0	0
	Moda	2.00	1.00 ^a

a. Existen varias modas. Se mostrará el menor de los valores.

Tabla de frecuencia

ESTADO CIVIL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	SOLTERO	14	28.0	28.0	28.0
	CASADO	26	52.0	52.0	80.0
	VIUDO	8	16.0	16.0	96.0
	DIVORCIADO	2	4.0	4.0	100.0
	Total	50	100.0	100.0	

ESTADO LABORAL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	EMPLEADO	25	50.0	50.0	50.0

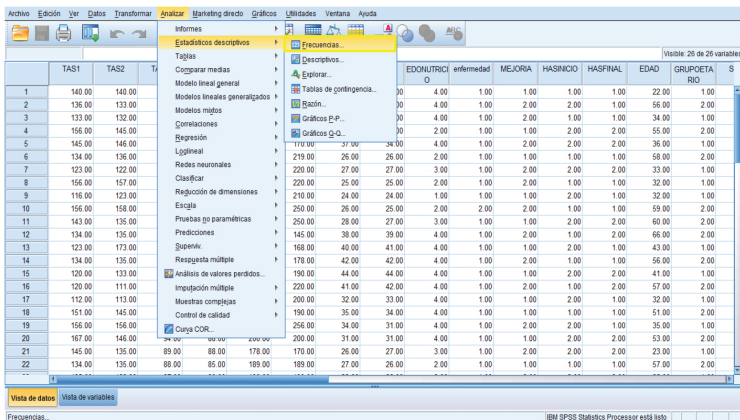
Rango

El *rango* o recorrido es la diferencia entre el valor máximo y mínimo de la distribución de las variables.

Para obtener el *rango* en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

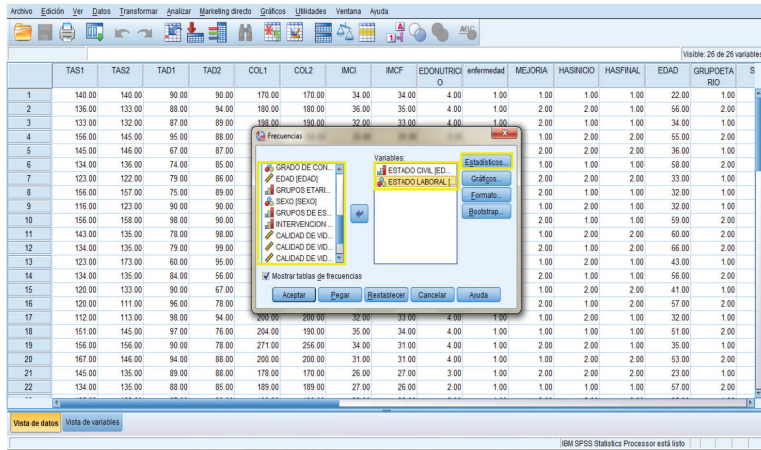
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



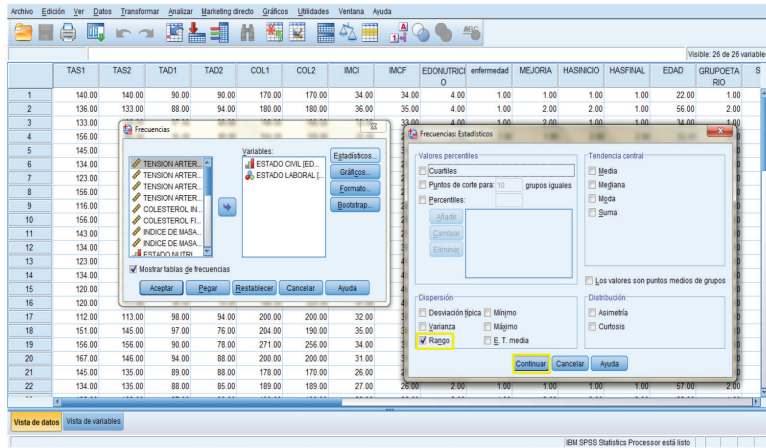
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



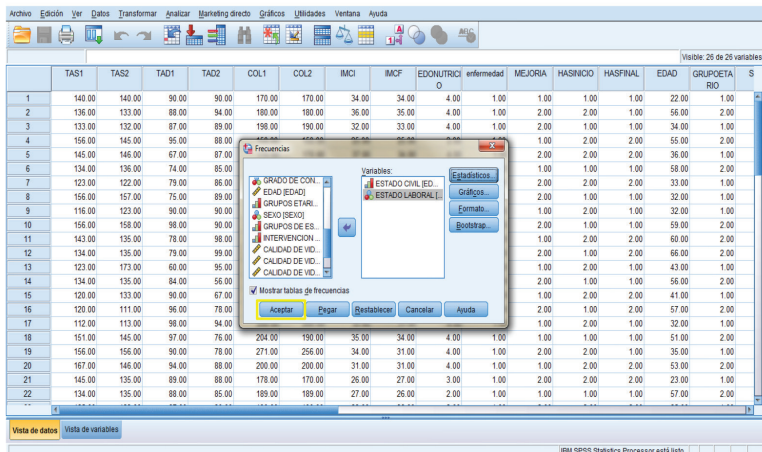
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Moda* y posteriormente en *Continuar*.



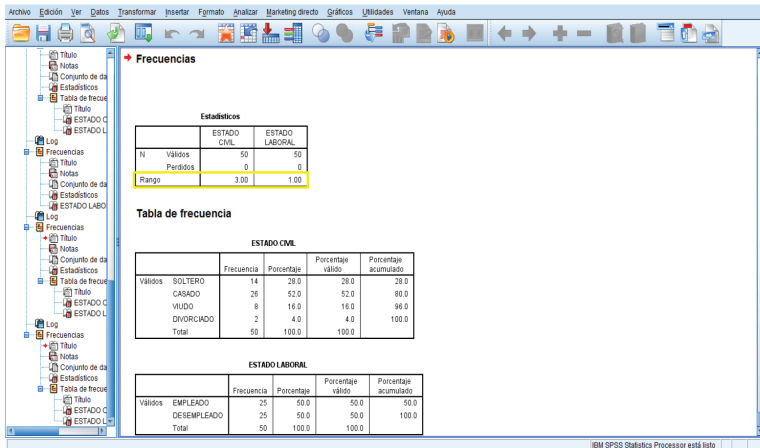
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene el *Rango* de las variables analizadas.



Interpretación de resultados

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

Frecuencias

[Conjunto_de_datos1] C:\Users\TOSHIBA\Desktop\CIEP\Proyecto SPSS\BAS

Estadísticos

		ESTADO CIVIL	ESTADO LABORAL
N	Válidos	50	50
	Perdidos	0	0
Rango		3.00	1.00

Tabla de frecuencia

ESTADO CIVIL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	SOLTERO	14	28.0	28.0	28.0
	CASADO	26	52.0	52.0	80.0
	VIUDO	8	16.0	16.0	96.0
	DIVORCIADO	2	4.0	4.0	100.0
	Total	50	100.0	100.0	

ESTADO LABORAL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	EMPLEADO	25	50.0	50.0	50.0
	DESEMPLEADO	25	50.0	50.0	100.0
	Total	50	100.0	100.0	

Varianza

Se entiende como *varianza* a la media del cuadrado de las desviaciones en relación con la media aritmética, y el resultado de la dispersión se expresa en la unidad de medida de la variable estudiada al cuadrado.

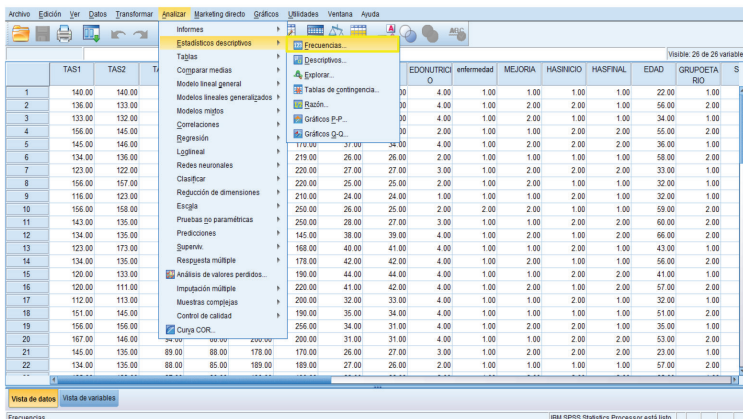
Entre mayor sea la varianza mayor será la dispersión.

$$S^2 = \frac{\Sigma ((X - \mu))^2}{n}$$

Para obtener la *varianza* en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

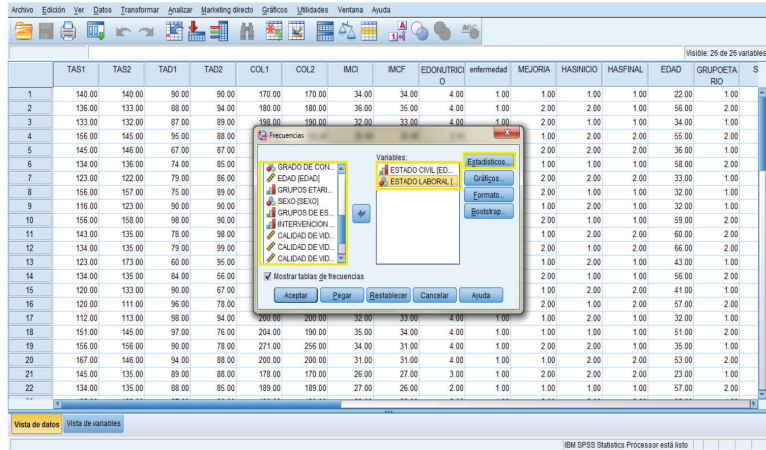
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



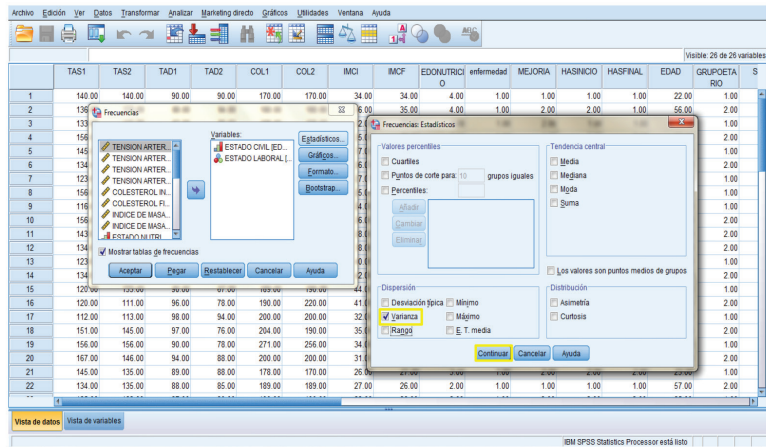
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



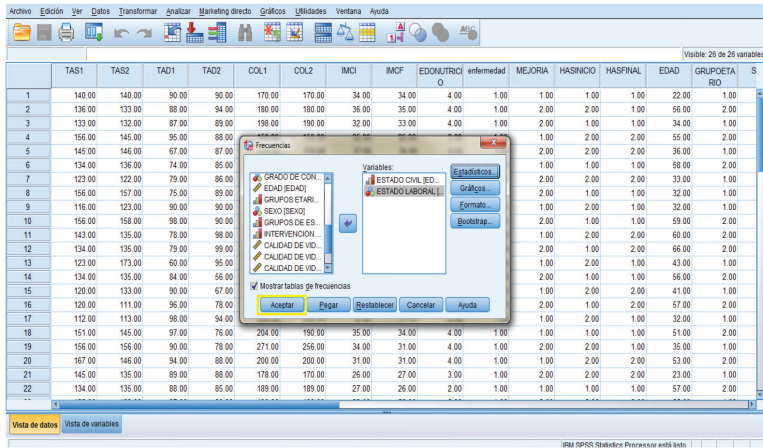
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Moda* y posteriormente en *Continuar*.



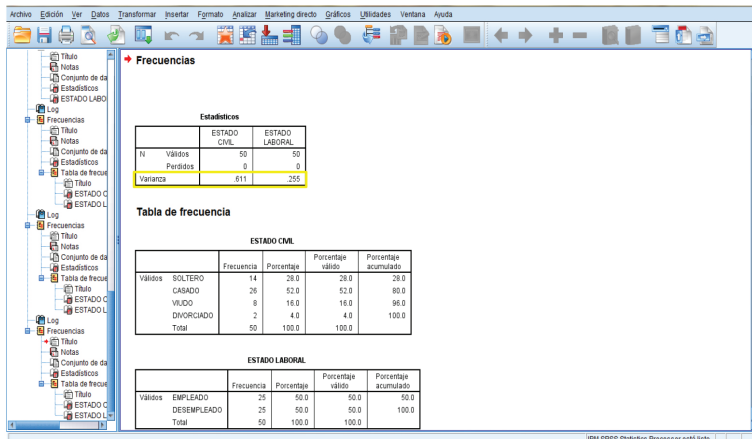
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene la *Varianza* de las variables analizadas.



Interpretación de resultados

Archivo Edición Ver Datos Transformar Insertar Formato Analizar Marketing directo Gráficos Utilidades Ventana Ay

Frecuencias

[Conjunto_de_datos1] C:\Users\TOSHIBA\Desktop\CIEF\Proyecto SPSS\BASE

Estadísticos

		ESTADO CIVIL	ESTADO LABORAL
N	Válidos	50	50
	Perdidos	0	0
	Varianza	.611	.255

Tabla de frecuencia

ESTADO CIVIL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	SOLTERO	14	28.0	28.0	28.0
	CASADO	26	52.0	52.0	80.0
	VIUDO	8	16.0	16.0	96.0
	DIVORCIADO	2	4.0	4.0	100.0
	Total	50	100.0	100.0	

ESTADO LABORAL

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	EMPLEADO	25	50.0	50.0	50.0
	DESEMPLEADO	25	50.0	50.0	100.0
	Total	50	100.0	100.0	

Desviación típica

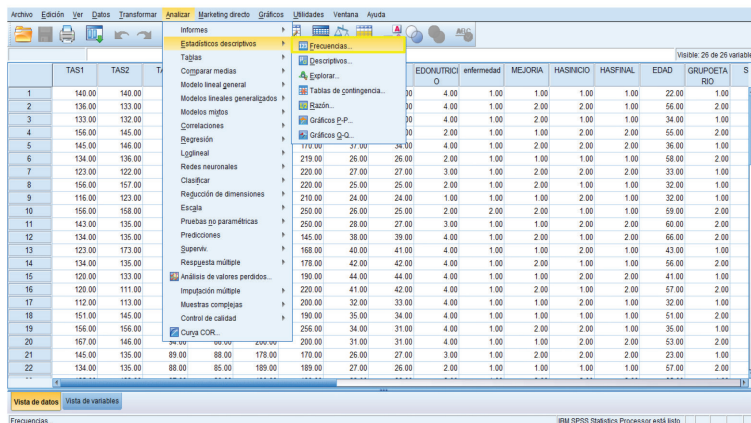
La desviación típica o desviación estándar es la medida de dispersión más utilizada y se obtiene mediante la raíz cuadrada de la varianza. Expresa la dispersión de la distribución en las mismas unidades de medida de la variable en estudio.

$$s = \sqrt{s^2}$$

Para obtener la desviación típica en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

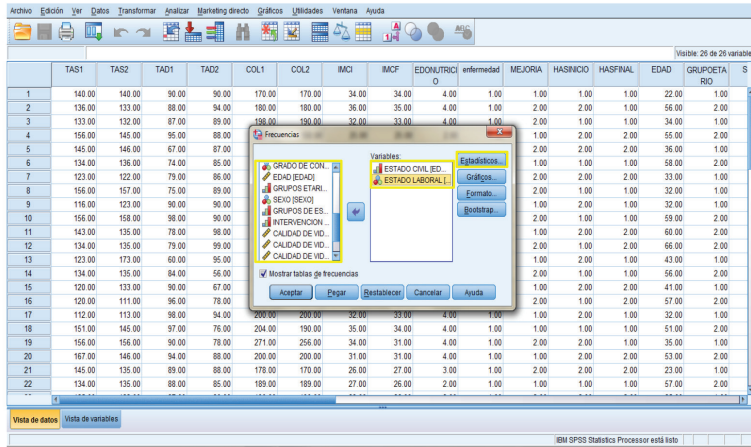
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Frecuencias*.



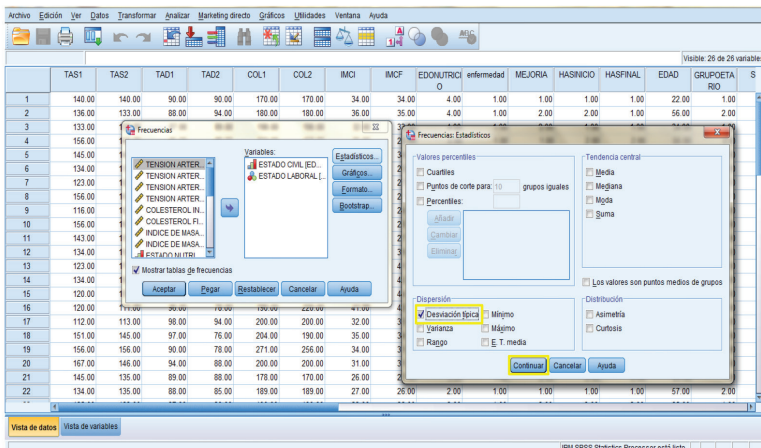
Paso 2

Aparecerá la ventana de *Frecuencias*, seleccione e introduzca las variables cualitativas que le interesen en el recuadro de *Variable* y posteriormente dé click en *Estadísticos*.



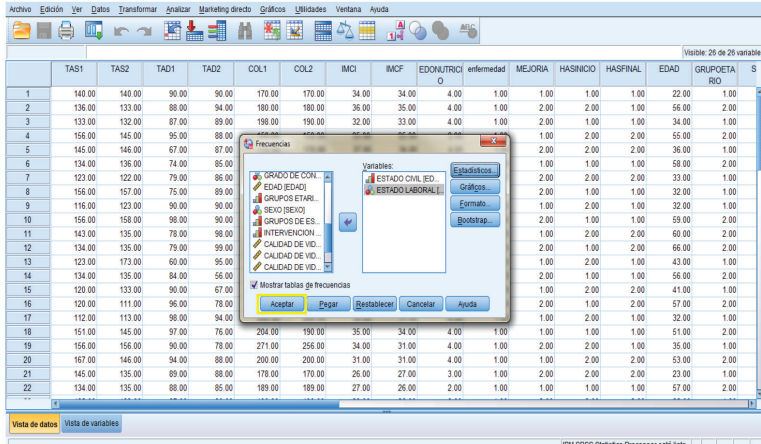
Paso 3

Aparecerá la ventana de *Frecuencias: Estadísticos*, dé click en *Moda* y posteriormente en *Continuar*.



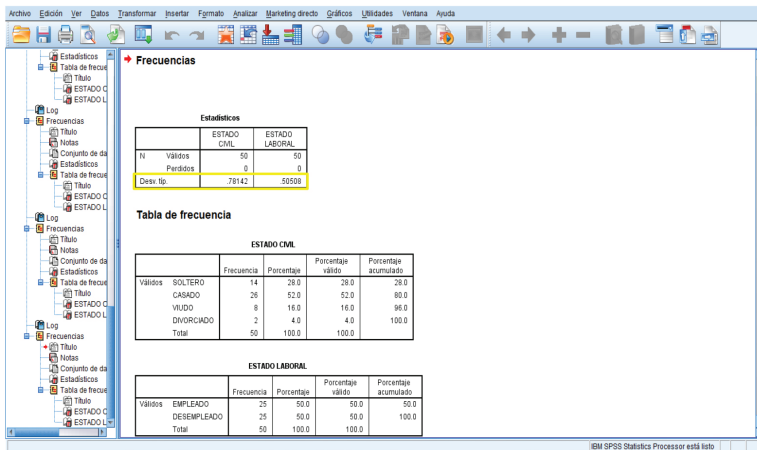
Paso 4

Aparecerá nuevamente la ventana de *Frecuencias* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene la desviación típica de las variables analizadas.



Interpretación de resultados

The screenshot shows the SPSS 'Frecuencias' (Frequencies) dialog box and its output. The data source is 'C:\Users\TOSHIBA\Desktop\CIEP\Proyecto SPSS\BASE'. The output includes summary statistics, a cross-tabulation for 'ESTADO CIVIL', and another for 'ESTADO LABORAL'.

Estadísticos

	ESTADO CIVIL	ESTADO LABORAL
N Válidos	50	50
Perdidos	0	0
Desv. tip.	.78142	.50508

Tabla de frecuencia

ESTADO CIVIL

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos SOLTERO	14	28.0	28.0	28.0
CASADO	26	52.0	52.0	80.0
VIUDO	8	16.0	16.0	96.0
DIVORCIADO	2	4.0	4.0	100.0
Total	50	100.0	100.0	

ESTADO LABORAL

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos EMPLEADO	25	50.0	50.0	50.0
DESEMPLEADO	25	50.0	50.0	100.0
Total	50	100.0	100.0	

T de Student

La prueba de T de Student o también conocida como t-Student o Test-T, es una prueba paramétrica para variables con distribución normal. Su cálculo estima si la diferencia de la comparación de las medias entre dos grupos es estadísticamente significativa.

En relación con la hipótesis, la hipótesis alterna establece que la diferencia entre los dos grupos es significativa, mientras que la hipótesis nula establece que no existe diferencia significativa entre los dos grupos.

La prueba T de Student en relación con la dependencia o independencia de la muestra, se clasifica en:

- T de Student para muestras dependientes, relacionadas o pareadas.
- T de Student para muestras independientes o no relacionadas.

T de Student para muestras dependientes o relacionadas

La prueba T de Student para muestras dependientes determina la diferencia de medias en un mismo grupo en diferente tiempo. Por ejemplo un estudio que evalúe en el mismo grupo los niveles de glucosa antes y después del tratamiento.

Se obtiene con la siguiente fórmula:

$$t = \frac{\bar{d}}{\frac{\sigma d}{\sqrt{N}}}$$

En donde:

d = Media de las diferencias pre y post evento

σd = Desviación estándar pre y post evento

N = Tamaño de la muestra

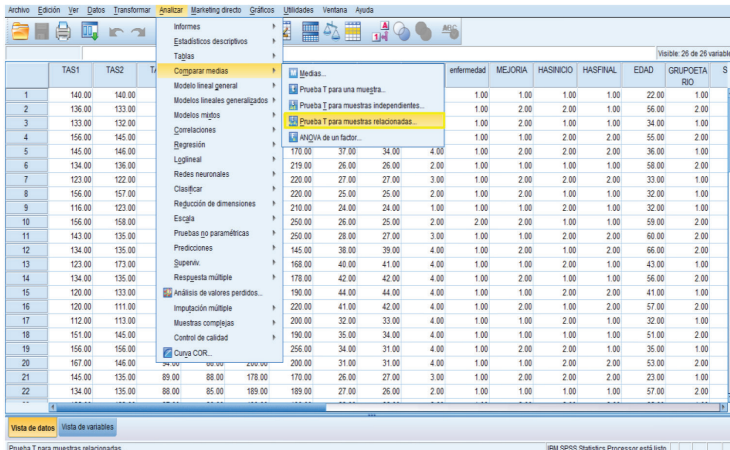
La significancia estadística se considera con el valor de “p”:

- < 0.05: es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05: no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de T de Student para muestras dependientes o relacionadas en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

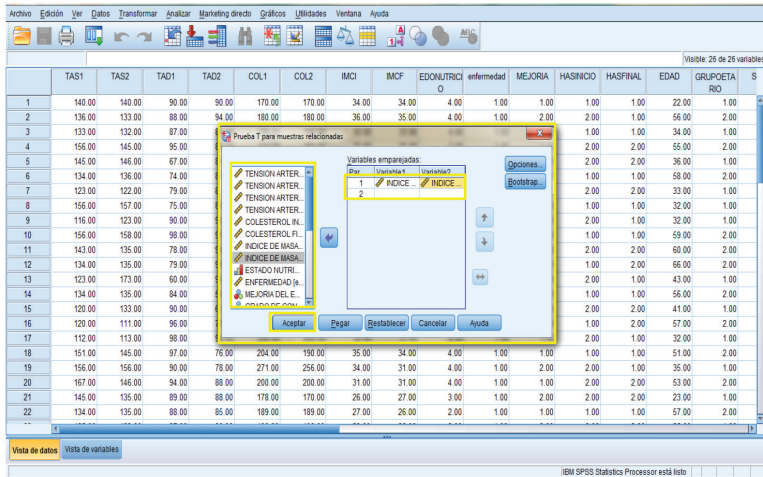
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Comparar medias* y posteriormente dé *click* en *Pruebas T para muestras relacionadas*.



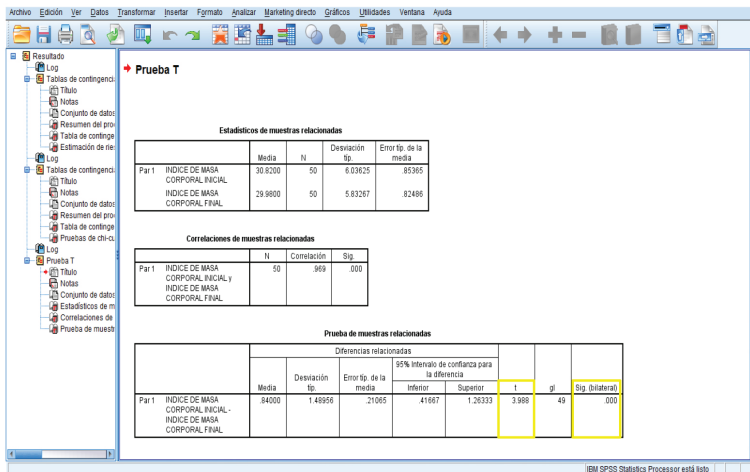
Paso 2

Aparecerá la ventana de *Pruebas T para muestras relacionadas*, seleccione e introduzca la variable dependiente de tipo cuantitativa previa a intervención que le interese y la variable independiente de tipo cuantitativa post intervención que le corresponda (por ejemplo, índice de masa corporal inicial e índice de masa corporal final o después de la intervención) y posteriormente dé *click* en *Aceptar*.



Paso 3

De esta forma se obtiene el resultado de la prueba T para muestras dependientes o relacionadas y la significancia de las variables analizadas.



Interpretación de resultados

Para valorar si la diferencia entre las medias de un mismo grupo en diferente tiempo es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe diferencia entre el grupo posterior a la intervención.
- Si el cambio que existió fue bueno o malo, positivo o negativo.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Estadísticos de muestras relacionadas

		Media	N	Desviación tip.	Error tip. de la media
Par 1	INDICE DE MASA CORPORAL INICIAL	30.8200	50	6.03625	.85365
	INDICE DE MASA CORPORAL FINAL	29.9800	50	5.83267	.82486

Correlaciones de muestras relacionadas

	N	Correlación	Sig.
Par 1 INDICE DE MASA CORPORAL INICIAL y INDICE DE MASA CORPORAL FINAL	50	.969	.000

Prueba de muestras relacionadas

	Diferencias relacionadas					t	gl	Sig. (bilateral)
	Media	Desviación tip.	Error tip. de la media	95% Intervalo de confianza para la diferencia				
				Inferior	Superior			
Par 1 INDICE DE MASA CORPORAL INICIAL - INDICE DE MASA CORPORAL FINAL	.84000	1.48956	.21065	.41667	1.26333	3.988	49	.000

Por ejemplo, en el análisis realizado a través de la prueba T de Student para muestras dependientes o relacionadas se obtuvo una media de 30.82 para la variable de índice de masa corporal inicial, 29.98 para la variable de índice de masa corporal final y una significancia de 0.000 al comparar las medias de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que existe un cambio en el índice de masa corporal posterior a la implementación de un régimen alimenticio adecuado, concluimos que:

- Entre el índice de masa corporal inicial y el final sí hubo una diferencia, debido a que se obtuvo una media de 30.82 para la variable de

índice de masa corporal inicial y una media de 29.98 para la variable de índice de masa corporal final.

- Que la diferencia que se presentó entre las variables fue positiva, debido a que hubo una disminución en el índice de masa corporal final o posterior a la intervención.
- Y que esta diferencia fue estadísticamente significativa debido a que se obtuvo una significancia ≤ 0.05 , por lo tanto se acepta la hipótesis alterna.

T de Student para muestras independientes

La prueba T de Student para muestras independientes determina la diferencia de medias entre dos grupos en el mismo periodo de tiempo. Por ejemplo, en un estudio que evalúe la efectividad de un tratamiento, se estima la diferencia de las medias entre el grupo que recibió el tratamiento y el grupo control.

Se obtiene con la siguiente fórmula:

$$t = \frac{\overline{x1} - \overline{x2}}{\sigma p \sqrt{\frac{1}{N1} + \frac{1}{N2}}}$$

En donde:

$x1$ = Valor promedio del grupo 1

$x2$ = Valor promedio del grupo 2

σp = Desviación estándar ponderada de ambos grupos

$N1$ = Tamaño de la muestra del grupo 1

$N2$ = Tamaño de la muestra del grupo 2

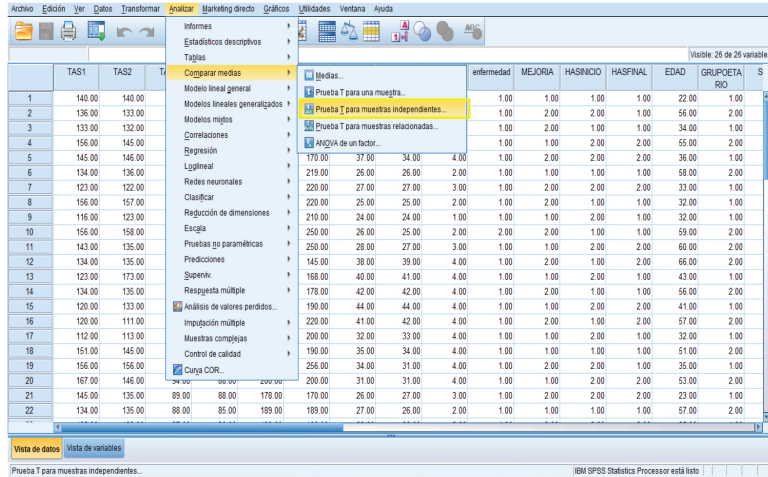
La significancia estadística se considera con el valor de “p”

- < 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de T de Student para muestras independientes o no relacionadas en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

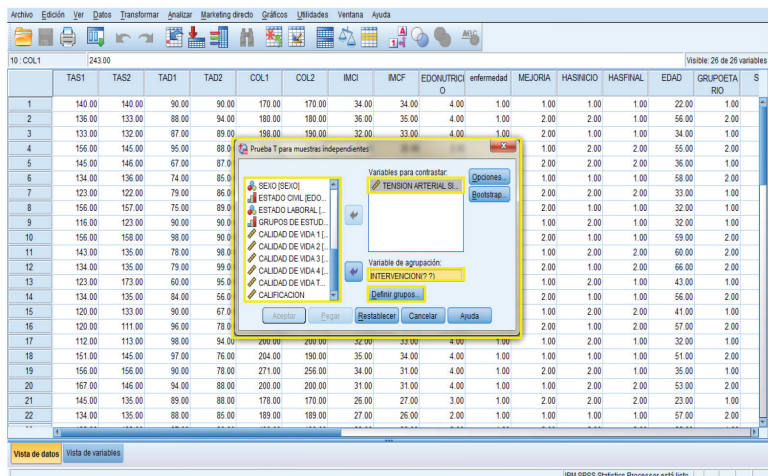
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Comparar medias* y posteriormente dé *click* en *Pruebas T para muestras independientes*.



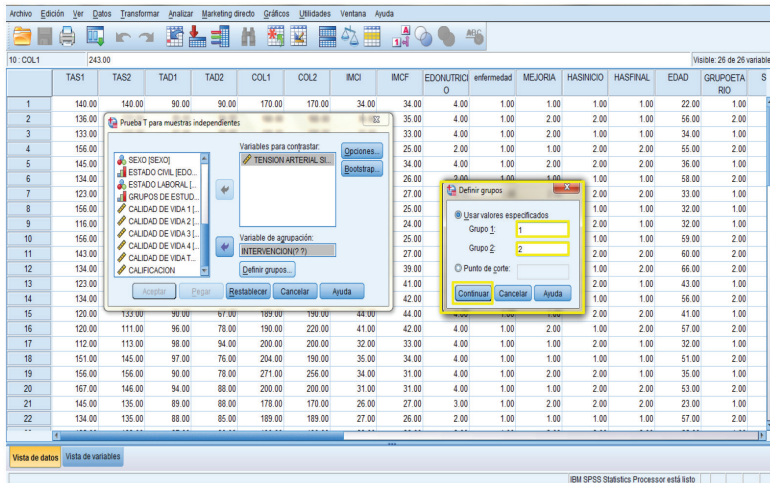
Paso 2

Aparecerá la ventana de *Pruebas T para muestras independientes*, seleccione e introduzca la variable dependiente de tipo cuantitativa que le interese en *Variables para contrastar* y la variable independiente de tipo cualitativa-dicotómica correspondiente en *Variable de agrupación* y posteriormente dé *click* en *Definir grupos*.



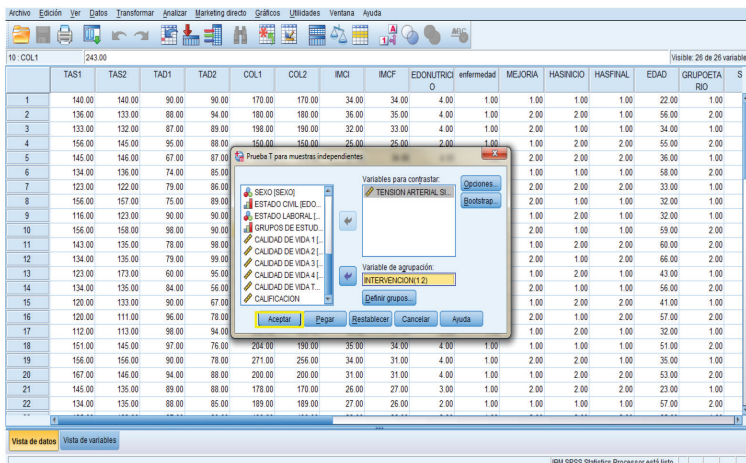
Paso 3

Aparecerá la ventana de *Definir grupos*, introduzca los grupos que dese incluir en la comparación de medias según el número de valor de etiqueta que asignó cuando creó su base de datos y posteriormente dé *click* en *Continuar*.



Paso 4

Aparecerá nuevamente la ventana de *Prueba T para muestras independientes* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene el resultado de la prueba T para muestras independientes o no relacionadas y la significancia de las variables analizadas, así como la media y la desviación estándar de los grupos analizados.

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

Prueba T

Estadísticos de grupo

	INTERVENCION O ESTUDIO	N	Media	Desviación tp.	Error tp. de la media
TENSION ARTERIAL	GRUPO CONTROL	25	135.7200	13.514493	3.10297
SISTOLICA INICIAL	GRUPO EXPERIMENTAL	25	143.3600	22.53309	4.50662

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% intervalo de confianza para la diferencia	
									Inferior	Superior
TENSION ARTERIAL	Se han asumido varianzas iguales	4.423	.041	-1.396	48	.169	-7.64000	5.47156	-18.64131	3.36131
SISTOLICA INICIAL	No se han asumido varianzas iguales			-1.396	42.580	.170	-7.64000	5.47156	-18.67761	3.39761

Interpretación de resultados

Para valorar si la diferencia entre las medias de dos grupos en el mismo periodo de tiempo es significativa, se debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe diferencia entre los dos grupos.
- Si el cambio que existió fue bueno o malo, positivo o negativo.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Estadísticos de grupo

INTERVENCIÓN ESTUDIO	N	Media	Desviación tip.	Error tip. de la media
TENSION ARTERIAL SISTOLICA INICIAL GRUPO CONTROL	25	135.7200	15.51483	3.10297
TENSION ARTERIAL SISTOLICA INICIAL GRUPO EXPERIMENTAL	25	143.3600	22.53309	4.50662

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
TENSION ARTERIAL SISTOLICA INICIAL	Se han asumido varianzas iguales	4.423	.041	-1.396	48	.169	-7.64000	5.47156	-18.64131	3.36131
	No se han asumido varianzas iguales			-1.396	42.580	.170	-7.64000	5.47156	-18.67761	3.39761

Por ejemplo, en el análisis realizado a través de la prueba T de Student para muestras independientes o no relacionadas se obtuvo una media de 135.72 para la variable de tensión arterial sistólica inicial del grupo control, 143.36 para la variable de tensión arterial sistólica inicial del grupo experimental y una significancia de 0.169 al comparar las medias de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia en la tensión arterial sistólica inicial entre el grupo control y el grupo experimental, concluimos que:

- Entre el grupo control y el grupo experimental si hubo una diferencia, debido a que se obtuvo una media de 135.72 para la variable de tensión arterial sistólica inicial del grupo control y una media de 143.36 para la variable de tensión arterial sistólica inicial del grupo experimental.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna.

Anova

La prueba de ANOVA también conocida como análisis de la varianza, análisis de varianza unidireccional, ANOVA de Fisher, prueba F, o análisis de la varianza de Fisher, es una prueba paramétrica para muestras con distribución normal.

Su cálculo estima la diferencia de las medias entre tres o más grupos en el mismo periodo de tiempo. Por ejemplo, evaluar la efectividad antipirética de los distintos AINE (grupo 1 con paracetamol, grupo 2 con ibuprofeno, grupo 3 con metamizol, etcétera).

En relación con la hipótesis, la hipótesis alterna establece que la diferencia entre los grupos es significativa, mientras que la hipótesis nula establece que no existe diferencia significativa entre los grupos.

Su obtiene con la siguiente fórmula:

$$SC\ Total = \Sigma x^2 - \frac{(\Sigma x)^2}{N}$$

En donde:

Σ = Sumatoria

N = Número de datos de todos los grupos

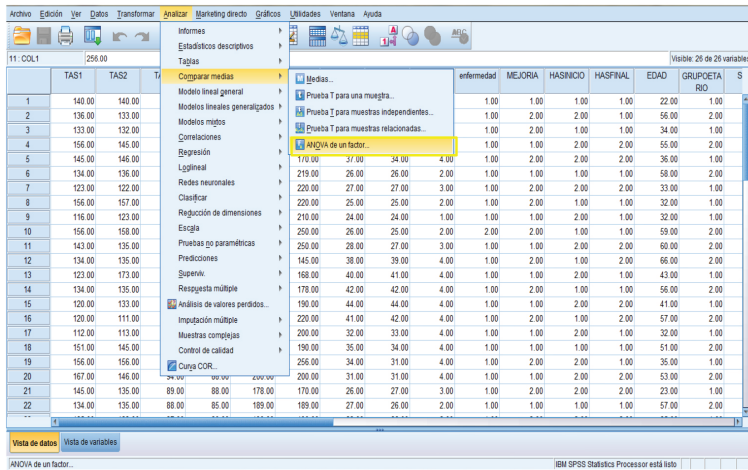
La significancia estadística se considera con el valor de “p”:

- < 0.05: es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05: no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de ANOVA en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

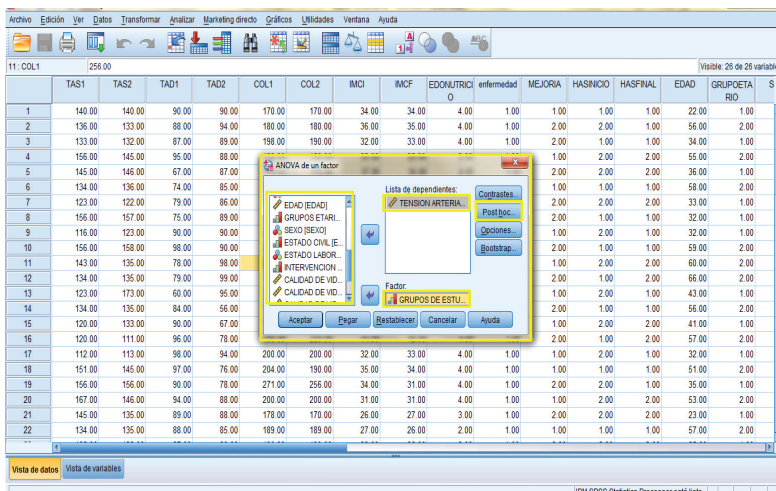
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Comparar medias* y posteriormente dé *click* en *ANOVA de un factor*.



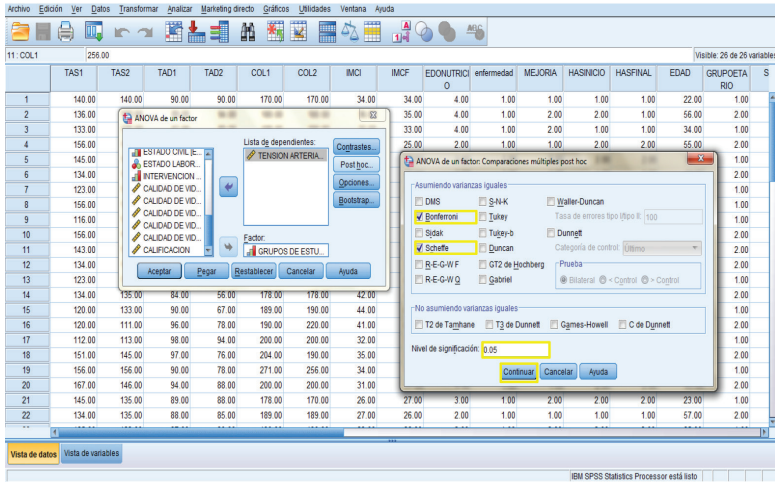
Paso 2

Aparecerá la ventana de *ANOVA de un factor*, seleccione e introduzca la variable dependiente de tipo cuantitativa que le interese en *Lista de dependientes* y la variable independiente de tipo cualitativa-ordinal a comparar en *Factor* y posteriormente dé *click* en *Post hoc*.



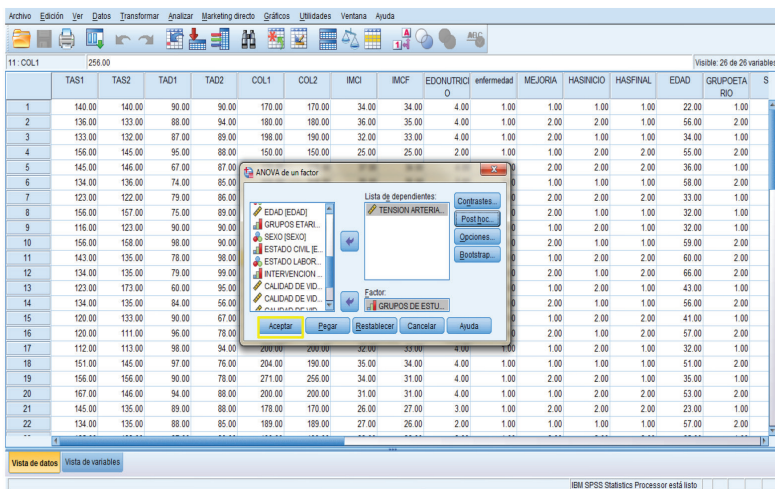
Paso 3

Aparecerá la ventana de ANOVA de un factor: comparaciones múltiples post hoc, dé click en Bonferroni y Scheffe, escriba 0.05 en Nivel de significación y posteriormente en Continuar.



Paso 4

Aparecerá nuevamente la ventana de ANOVA de un factor donde deberá dar click en Aceptar.



Paso 5

De esta forma se obtiene el resultado de la *significancia* de la prueba de ANOVA y la significancia de cada uno de los grupos de las variables analizadas.

ANOVA de un factor

	Suma de cuadrados	g	Medio cuadrática	F	Sig.
Entre grupos	452,714	2	423,357	1,132	,331
Intra-grupos	17833,756	47	379,441		
Total	18692,420	49			

Pruebas post hoc

Comparaciones múltiples

Variable dependiente: TENSION ARTERIAL SISTOLICA INICIAL

	I (GRUPOS DE ESTUDIO)	J (GRUPOS DE ESTUDIO)	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
Scheffé	GRUPO CONTROL	GRUPO SODHI	-6,29412	6,68132	,644	-23,1836	10,5954
		GRUPO EXPERIMENTAL	-10,07353	6,78491	,341	-27,2249	7,0778
	GRUPO SODHI	GRUPO CONTROL	6,29412	6,68132	,644	-10,5954	23,1836
Bonferroni	GRUPO EXPERIMENTAL	GRUPO CONTROL	10,07353	6,78491	,341	-7,0778	27,2249
		GRUPO SODHI	3,77941	6,78491	,857	-13,3720	20,9308
	GRUPO CONTROL	GRUPO SODHI	-6,29412	6,68132	1,000	-22,8818	10,2936
Scheffé	GRUPO CONTROL	GRUPO EXPERIMENTAL	-10,07353	6,78491	,433	-26,9184	6,7713
		GRUPO SODHI	6,29412	6,68132	1,000	-10,2936	22,8818
	GRUPO EXPERIMENTAL	GRUPO CONTROL	-3,77941	6,78491	1,000	-20,6243	13,0655
Bonferroni	GRUPO EXPERIMENTAL	GRUPO CONTROL	10,07353	6,78491	,433	-6,7713	26,9184
		GRUPO SODHI	3,77941	6,78491	1,000	-13,0655	20,6243

Y la significancia de la prueba de ANOVA por Scheffe.

Variable dependiente: TENSION ARTERIAL SISTOLICA INICIAL

	I (GRUPOS DE ESTUDIO)	J (GRUPOS DE ESTUDIO)	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
Scheffé	GRUPO CONTROL	GRUPO SODHI	-6,29412	6,68132	,644	-23,1836	10,5954
		GRUPO EXPERIMENTAL	-10,07353	6,78491	,341	-27,2249	7,0778
	GRUPO SODHI	GRUPO CONTROL	6,29412	6,68132	,644	-10,5954	23,1836
Bonferroni	GRUPO EXPERIMENTAL	GRUPO CONTROL	10,07353	6,78491	,341	-7,0778	27,2249
		GRUPO SODHI	3,77941	6,78491	,857	-13,3720	20,9308
	GRUPO CONTROL	GRUPO SODHI	-6,29412	6,68132	1,000	-22,8818	10,2936
Scheffé	GRUPO CONTROL	GRUPO EXPERIMENTAL	-10,07353	6,78491	,433	-26,9184	6,7713
		GRUPO SODHI	6,29412	6,68132	1,000	-10,2936	22,8818
	GRUPO EXPERIMENTAL	GRUPO CONTROL	-3,77941	6,78491	1,000	-20,6243	13,0655
Bonferroni	GRUPO EXPERIMENTAL	GRUPO CONTROL	10,07353	6,78491	,433	-6,7713	26,9184
		GRUPO SODHI	3,77941	6,78491	1,000	-13,0655	20,6243

Subconjuntos homogéneos

TENSION ARTERIAL SISTOLICA INICIAL

GRUPOS DE ESTUDIO	N	Subconjunto para	Haga clic para
GRUPO CONTROL	17	134,1765	
GRUPO SODHI	17	140,4796	
GRUPO EXPERIMENTAL	16	144,9500	
Sig.			,337

Se muestran las medias para los grupos en los subconjuntos homogéneos.

Interpretación de resultados

Para valorar si la diferencia de las medias entre tres o más grupos en el mismo periodo de tiempo es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe diferencia entre tres o más grupos en el mismo periodo de tiempo.
- Si el cambio es estadísticamente significativo.

Recordemos que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

ANOVA de un factor

TENSION ARTERIAL SISTOLICA INICIAL					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	858.714	2	429.357	1.132	.331
Intra-grupos	17833.706	47	379.441		
Total	18692.420	49			

Pruebas post hoc

Comparaciones múltiples

Variable dependiente: TENSION ARTERIAL SISTOLICA INICIAL							
	(I) GRUPOS DE ESTUDIO	(J) GRUPOS DE ESTUDIO	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
Scheffé	GRUPO CONTROL	GRUPO SODHI	-6.29412	6.68132	.644	-23.1836	10.5954
		GRUPO EXPERIMENTAL	-10.07353	6.78491	.341	-27.2249	7.0778
	GRUPO SODHI	GRUPO CONTROL	6.29412	6.68132	.644	-10.5954	23.1836
		GRUPO EXPERIMENTAL	-3.77941	6.78491	.857	-20.9308	13.3720
	GRUPO EXPERIMENTAL	GRUPO CONTROL	10.07353	6.78491	.341	-7.0778	27.2249
		GRUPO SODHI	3.77941	6.78491	.857	-13.3720	20.9308
Bonferroni	GRUPO CONTROL	GRUPO SODHI	-6.29412	6.68132	1.000	-22.8818	10.2936
		GRUPO EXPERIMENTAL	-10.07353	6.78491	.433	-26.9184	6.7713
	GRUPO SODHI	GRUPO CONTROL	6.29412	6.68132	1.000	-10.2936	22.8818
		GRUPO EXPERIMENTAL	-3.77941	6.78491	1.000	-20.6243	13.0655
	GRUPO EXPERIMENTAL	GRUPO CONTROL	10.07353	6.78491	.433	-6.7713	26.9184
		GRUPO SODHI	3.77941	6.78491	1.000	-13.0655	20.6243

Subconjuntos homogéneos

TENSION ARTERIAL SISTOLICA INICIAL			
	GRUPOS DE ESTUDIO	N	Subconjunto para alfa = 0.05
			1
Scheffé ^{a,b}	GRUPO CONTROL	17	134.1765
	GRUPO SODHI	17	140.4706

Por ejemplo, en el análisis realizado a través de la prueba ANOVA se obtuvo una significancia de 0.331 y según Scheffe una significancia de 0.337 al comparar las medias de las variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia en la tensión arterial sistólica inicial entre el grupo control, el grupo SODHI y el grupo experimental, concluimos que:

- Entre el grupo control, el grupo SODHI y el grupo experimental sí hubo una diferencia para la variable de tensión arterial sistólica inicial.
- Esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna.

Correlación de Pearson

La correlación de Pearson es una prueba paramétrica para muestras independientes con distribución normal.

Su cálculo estima la asociación entre la variable dependiente e independiente determinando en qué proporción una variable interviene para que otra se modifique.

Se obtiene con la siguiente fórmula:

$$r = \frac{N \Sigma xy - \Sigma x \Sigma y}{\sqrt{[N \Sigma x^2 - (\Sigma x)^2][N \Sigma y^2 - (\Sigma y)^2]}}$$

En donde:

Σxy = Sumatoria de los productos de la variable dependiente e independiente

Σx = Sumatoria de los valores de la variable dependiente

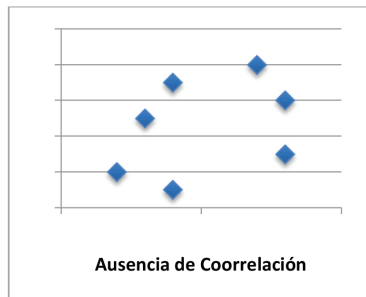
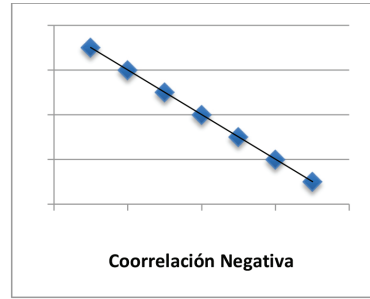
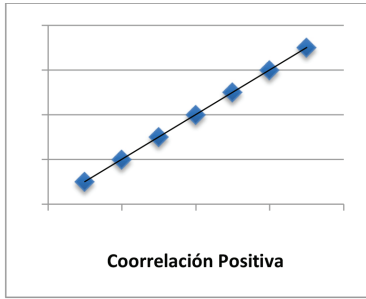
Σy = Sumatoria de los valores de la variable independiente

Σx^2 = Sumatoria de los valores de la variable dependiente al cuadrado

Σy^2 = Sumatoria de los valores de la variable independiente al cuadrado

N = Tamaño de la muestra

Se grafica a través de coordenadas cartesianas, en donde la variable dependiente corresponde al eje de la “y”, mientras que la variable independiente corresponde al eje de la “x”.



Sus resultados oscilan entre +1 a -1 y se interpretan como:

- $r = +1$ o -1 : correlación perfecta (positiva o negativa).
- $r = 0$: ausencia de correlación.
- $r = 0.80 - 0.99$: correlación fuerte.
- $r = 0.60 - 0.79$: correlación moderada.
- $r = 0.40 - 0.59$: correlación parcial.
- $r = 0.20 - 0.39$: correlación ligera.
- $r = 0.00 - 0.19$: correlación fortuita o insignificante.

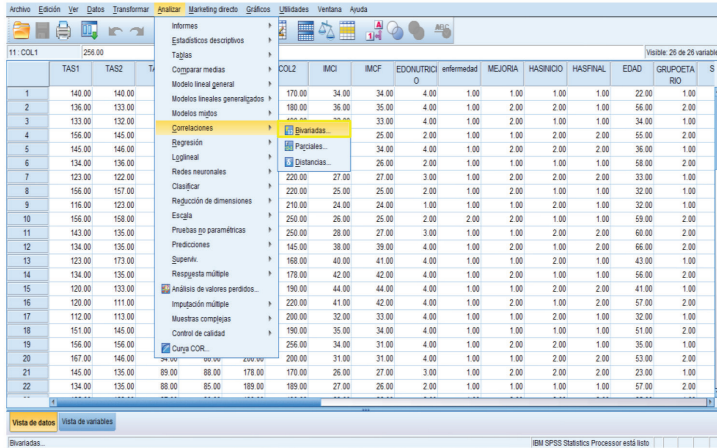
La significancia estadística se considera con el valor de “p”.

- ≤ 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la correlación de Pearson en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

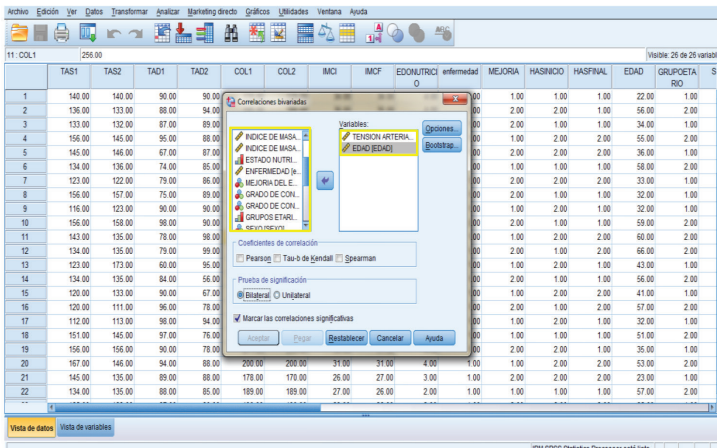
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Correlaciones* y posteriormente dé *click* en *Bivariadas*.



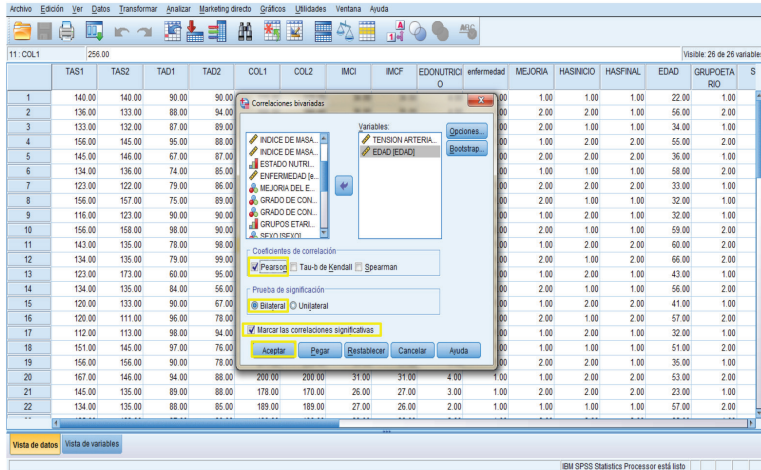
Paso 2

Aparecerá la ventana de *Correlaciones bivariadas*, seleccione e introduzca la variable dependiente e independiente de tipo cuantitativo que le interese correlacionar.



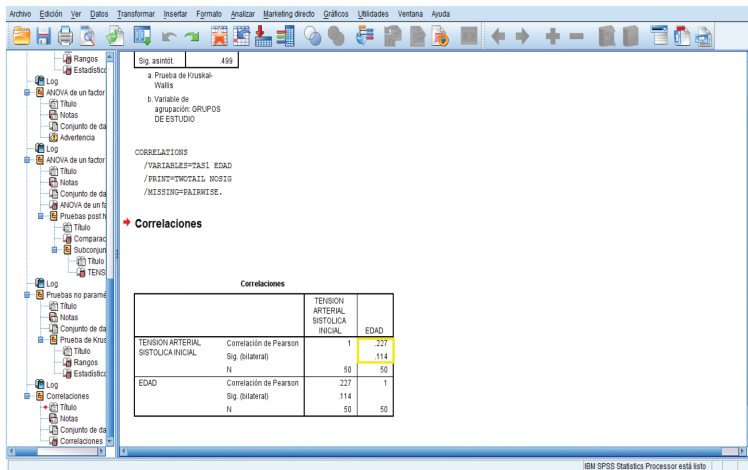
Paso 3

Dé click en *Pearson*, en *Bilateral*, en *Marcar las correlaciones significativas* y posteriormente en *Aceptar*.



Paso 4

De esta forma se obtiene el resultado de la correlación de Pearson y la significancia de las variables correlacionadas.



Interpretación de resultados

Para valorar si la proporción en que la variable independiente interviene para que la variable dependiente se modifique es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece si la proporción en que una variable interviene para que otra variable se modifique es estadísticamente significativa.

Recordemos que los resultados de la correlación de Spearman oscilan entre +1 a -1 y se interpretan como correlación perfecta (ya sea positiva o negativa) cuando el valor de r es de +1 o -1, ausencia de correlación cuando el valor de r es de 0, correlación fuerte cuando el valor de r es de 0.80 a 0.99, correlación moderada cuando el valor de r es de 0.60-0.79, correlación parcial cuando el valor de r es de 0.40-0.59, correlación ligera cuando el valor de r es de 0.20-0.39 y correlación fortuita o insignificante cuando el valor de r es de 0.01 a 0.19.

Y una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Correlaciones

		TENSION ARTERIAL SISTOLICA INICIAL	EDAD
TENSION ARTERIAL SISTOLICA INICIAL	Correlación de Pearson	1	.227
	Sig. (bilateral)		.114
	N	50	50
EDAD	Correlación de Pearson	.227	1
	Sig. (bilateral)	.114	
	N	50	50

Por ejemplo, en el análisis realizado a través de la correlación de Pearson se obtuvo una correlación de 0.227 y una significancia de 0.114 al valorar la asociación entre la variable de tensión arterial sistólica inicial y la variable de edad.

Al suponer que nuestra hipótesis alterna establezca que la edad interviene para que la tensión arterial sistólica inicial se modifique, concluimos que:

- La relación entre la tensión arterial sistólica inicial y la edad no fue estadísticamente significativa, es decir, la edad no modifica la tensión arterial sistólica inicial, debido a que se obtuvo una correlación

ligera y una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna.

Odds ratio

También conocida como razón de momios, razón de chances, razón de productos cruzados, razón de disparidad, razón de predominio, proporción de desigualdades, razón de oposiciones, oposición de probabilidades contrarias, cociente de probabilidades relativas, oportunidad relativa.

Permite cuantificar la asociación entre la causa y el efecto, es decir, mide qué tan frecuente es la enfermedad en el grupo de los expuestos en relación con los no expuestos, por lo que es la medida de asociación de los estudios de casos y controles.

Se obtiene con la siguiente fórmula:

	<i>Casos</i>	<i>Controles</i>
Expuestos	a	b
No Expuestos	c	d
$OR = \frac{ad}{bc}$		

Sus resultados oscilan entre 0 e infinito y se interpretan como:

- OR <1: factor de protección.
- OR =1: no es factor de riesgo ni de protección.
- OR >1: factor de riesgo.

La significancia estadística depende del intervalo de confianza (IC 95%), por lo que:

- Si incluye 1, no es estadísticamente significativo.
- Si no incluye 1, es estadísticamente significativo.

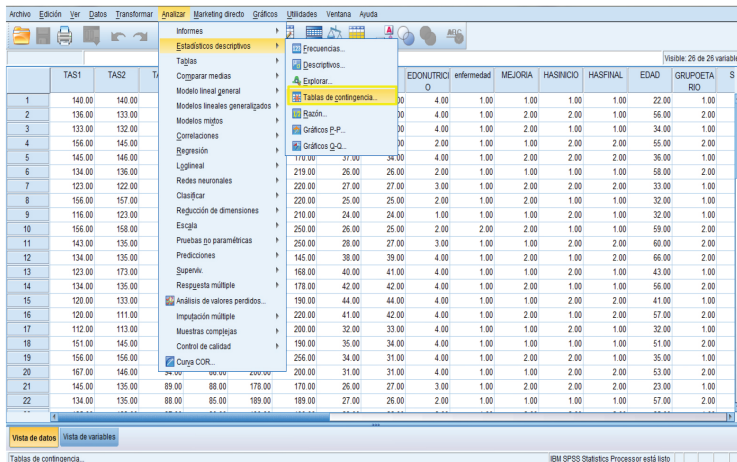
Si el valor del OR fuese menor a 1 habrá que convertirlo para su interpretación, esto en caso de que no se pueda leer como un factor de protección.

Por ejemplo si el valor de OR es .63 se deberá realizar la siguiente ecuación para su lectura $1/.63 = 1.58$.

Para realizar el análisis estadístico mediante la prueba de Odds Ratio en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

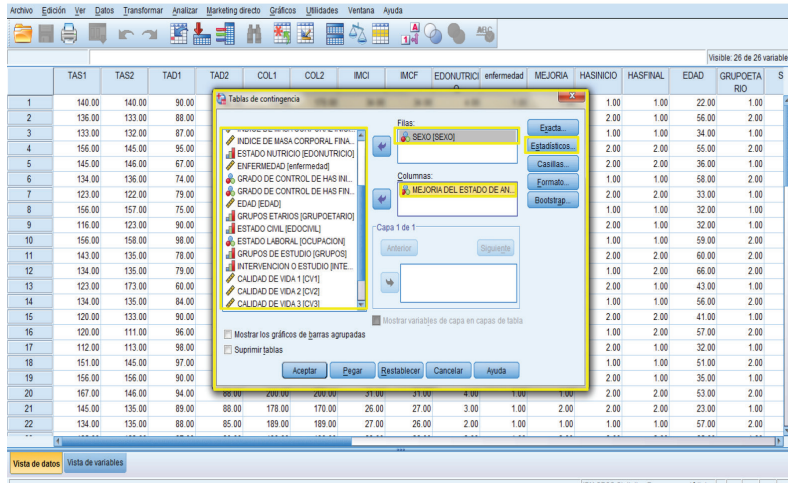
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Tablas de contingencia*.



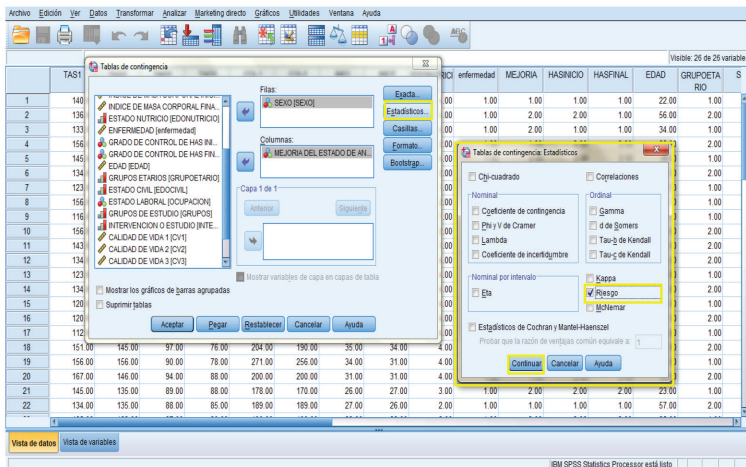
Paso 2

Aparecerá la ventana de *Tablas de contingencia*, seleccione e introduzca la variable dependiente de tipo cualitativo que le interese en *Fila* y la variable independiente de tipo cualitativo que le interese en *Columna* y posteriormente dé *click* en *Estadísticos*.



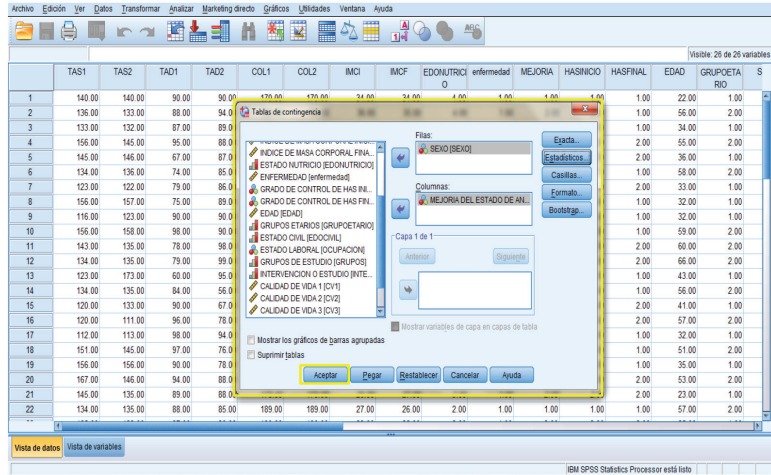
Paso 3

Aparecerá la ventana de *Tablas de contingencia: estadísticos*, dé *click* en *Riesgo* y posteriormente en *Continuar*.



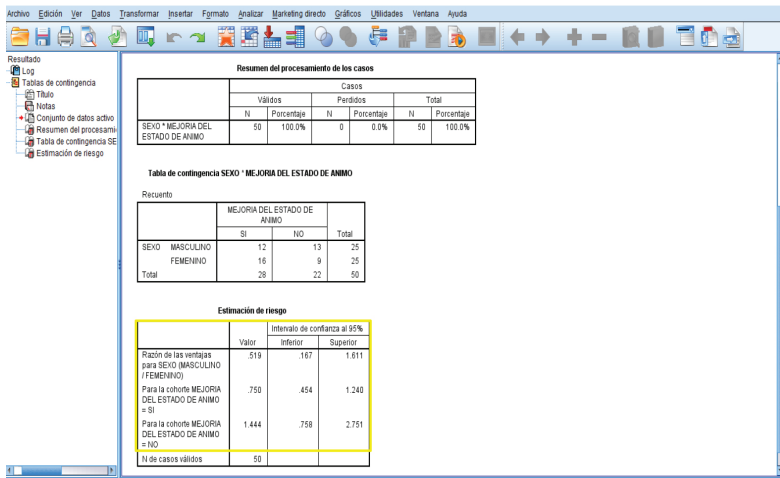
Paso 4

Aparecerá nuevamente la ventana de *Tablas de contingencia* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene el resultado de OR y los intervalos de confianza 95% de las variables analizadas.



Interpretación de resultados

Para valorar si la asociación entre la causa y el efecto es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe asociación entre la causa y el efecto.
- Si la asociación es estadísticamente significativa.

Recordando que sus resultados oscilan entre 0 e infinito y se interpretan como factor de protección cuando el OR es <1 , factor de riesgo cuando el OR es >1 y como que no es factor de riesgo ni de protección cuando el OR es igual a 1.

Y la significancia estadística depende del intervalo de confianza a 95%, si incluye 1 indica que la asociación es estadísticamente significativa, mientras que si no incluye 1 indica que la asociación no es estadísticamente significativa.

Tabla de contingencia SEXO * MEJORIA DEL ESTADO DE ANIMO

Recuento		MEJORIA DEL ESTADO DE ANIMO		Total
		SI	NO	
SEXO	MASCULINO	12	13	25
	FEMENINO	16	9	25
Total		28	22	50

Estimación de riesgo

	Valor	Intervalo de confianza al 95%	
		Inferior	Superior
Razón de las ventajas para SEXO (MASCULINO / FEMENINO)	.519	.167	1.611
Para la cohorte MEJORIA DEL ESTADO DE ANIMO = SI	.750	.454	1.240
Para la cohorte MEJORIA DEL ESTADO DE ANIMO = NO	1.444	.758	2.751
N de casos válidos	50		

Por ejemplo, en el análisis realizado a través de la prueba de Odds Ratio se obtuvo una razón de las ventajas para sexo de 0.519 con un IC 95% de 0.161 a 1.611, para la cohorte de mejoría del estado de ánimo de 0.750 con un IC 95% de 0.454 a 1.240, para la cohorte de no mejoría del estado de ánimo de 1.444 con un IC 95% de 0.758 a 2.751.

Al suponer que nuestra hipótesis alterna establezca que sí existe asociación entre la mejoría del estado de ánimo y el sexo, concluimos que:

- Entre el grupo control y el grupo experimental sí hubo una diferencia, debido a que se obtuvo una media de 135.72 para la variable de tensión arterial sistólica inicial del grupo control y una media de 143.36 para la variable de tensión arterial sistólica inicial del grupo experimental.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 .

Chi²

La prueba de Chi² también conocida como X² o ji cuadrada, es una prueba no paramétrica.

Su cálculo estima la asociación entre dos variables en relación con la existencia de diferencia entre la frecuencia observada y la frecuencia esperada en la distribución de las variables. Por ejemplo, al comparar la eficacia del uso de un medicamento de nueva generación (pantoprazol) en relación con otro del mismo grupo farmacológico (omeprazol).

Se obtiene con la siguiente fórmula:

$$X^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

En donde:

Σ = Sumatoria

f_o = Frecuencia observada

f_e = Frecuencia esperada

Al medir la diferencia entre la frecuencia observada y la esperada, cuanto mayor sea el valor de la diferencia de la Chi² mayor será la relación entre ambas variables, es decir, existirá una diferencia significativa.

Grados de Libertad	Probabilidad											
	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001	
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83	
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82	
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27	
4	0.71	1.06	1.65	2.20	3.36	4.35	5.99	7.78	9.49	13.28	18.47	
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52	
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46	
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32	
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12	
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88	
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59	
	No Significativo								Significativo			

La hipótesis nula (H₀) determina que no existe asociación entre las variables y la hipótesis alterna (H_A) determina que sí existe asociación entre las variables.

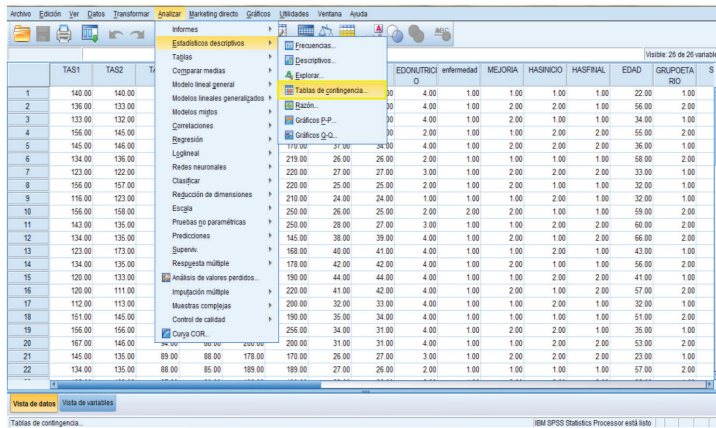
La significancia estadística se considera con el valor de “p”:

- < 0.05: es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05: no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de **Chi²** en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

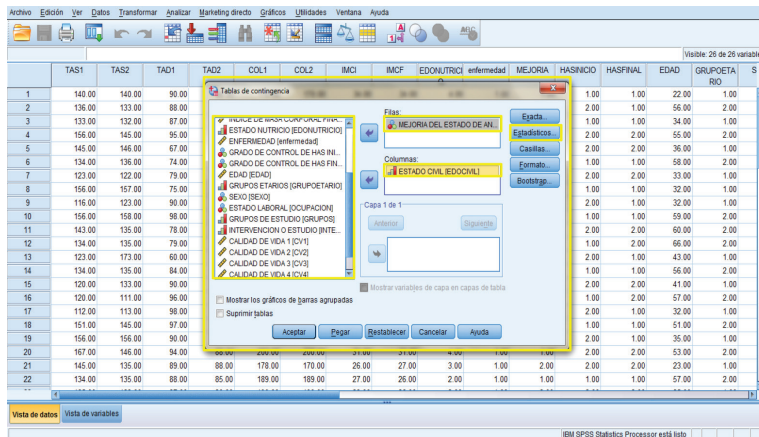
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Tablas de contingencia*.



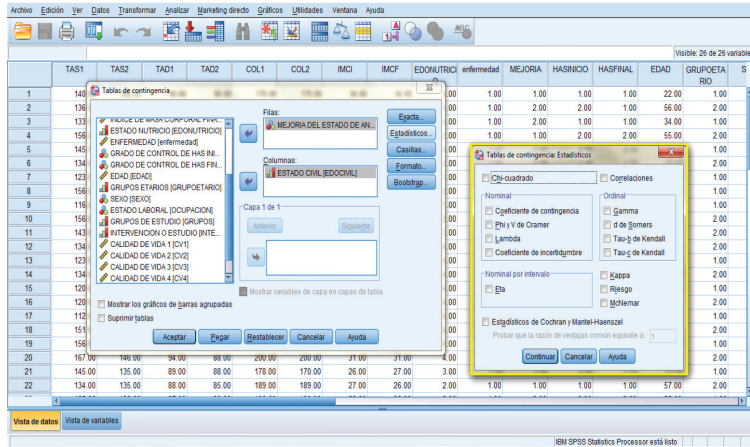
Paso 2

Aparecerá la ventana de *Tablas de contingencia*, seleccione e introduzca la variable dependiente de tipo cualitativa que le interese en *Fila* y la variable independiente de tipo cualitativa correspondiente en *Columna* y posteriormente dé *click* en *Estadísticos*.



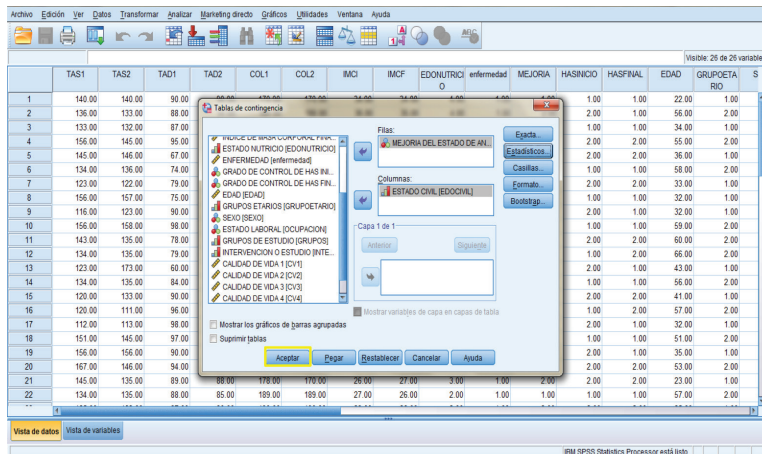
Paso 3

Aparecerá la ventana de *Tablas de contingencia: estadísticos*, dé click en *Chi cuadrado* y posteriormente en *Continuar*.



Paso 4

Aparecerá nuevamente la ventana de *Tablas de contingencia* donde deberá dar click en *Aceptar*.



Paso 5

De esta forma se obtiene el resultado de la prueba χ^2 y la *Significancia* de las variables analizadas.

The screenshot displays the SPSS interface with the following data:

Resumen del procesamiento de los casos

	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
MEJORA DEL ESTADO DE ANIMO * ESTADO CIVIL	50	100.0%	0	0.0%	50	100.0%

Tabla de contingencia MEJORA DEL ESTADO DE ANIMO * ESTADO CIVIL

Recuento

		ESTADO CIVIL				Total
		SOLTERO	CASADO	VIUDO	DIVORCIADO	
MEJORA DEL ESTADO DE ANIMO	SI	10	10	8	2	28
	NO	4	16	2	0	22
	Total	14	26	8	2	50

Pruebas de chi cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	7.342 ^a	3	.063
Razón de verosimilitudes	8.198	3	.042
Asociación lineal por índex	.167	1	.683
N de casos válidos	50		

Interpretación de resultados

Para valorar si la asociación entre dos variables es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que si existe asociación entre las dos variables.
- Si la asociación es estadísticamente significativa.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Tabla de contingencia MEJORIA DEL ESTADO DE ANIMO * ESTADO CIVIL

Recuento		ESTADO CIVIL				Total
		SOLTERO	CASADO	VIUDO	DIVORCIADO	
MEJORIA DEL ESTADO DE ANIMO	SI	10	10	6	2	28
	NO	4	16	2	0	22
Total		14	26	8	2	50

Pruebas de chi-cuadrado

	Valor	gl.	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	7.342 ^a	3	.062
Razón de verosimilitudes	8.198	3	.042
Asociación lineal por lineal	.167	1	.683
N de casos válidos	50		

Por ejemplo, en el análisis realizado a través de la prueba de Chi² se obtuvo una significancia de 0.062 al valorar la asociación de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una asociación entre la mejoría del estado de ánimo y el estado civil, concluimos que:

- Entre la mejoría del estado de ánimo y el estado civil la asociación no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05, por lo tanto se rechaza la hipótesis alterna.

McNemar

La prueba de McNemar es una prueba no paramétrica con distribución no normal de variables dicotómicas.

Su cálculo estima el cambio de respuesta de una variable dicotómica después de un evento o intervención específica en el mismo sujeto.

Las variables de estudio en la prueba de McNemar deben contar con las siguientes características:

- Ser dicotómicas.
- Que se represente con el mismo valor los dos periodos (previo y posterior al evento o intervención).
- Que representen frecuencias.
- La respuesta pre y post evento pertenezca al mismo sujeto.

La hipótesis del estudio se encuentra en relación con la ausencia o presencia de cambio, por lo que la hipótesis nula establece que no existe cambio de respuesta, mientras que la hipótesis alterna postula que sí existe cambio de respuesta. Cuando el cambio de respuesta es mayor de lo esperado, se rechaza la hipótesis nula.

Se obtiene con la siguiente *fórmula*:

		Después	
		Sin Cambio	Cambio
Antes	Cambio	a	b
	Sin Cambio	c	d
$X^2 = \frac{((a-d) - 1)^2}{a+b}$			

En donde:

S = Sumatoria

f_o = Frecuencia observada

f_e = Frecuencia esperada

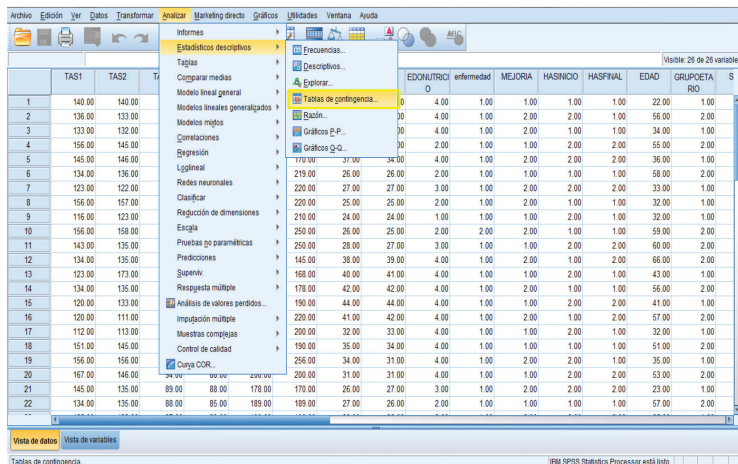
La significancia estadística se considera con el valor de “p”:

- < 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de McNemar en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

Paso 1

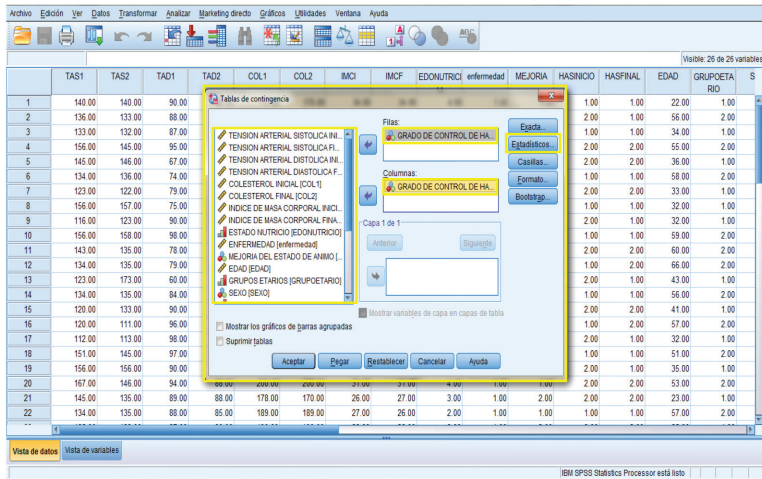
En la barra de herramientas dé *click* en *Analizar*, seleccione *Estadísticos descriptivos* y posteriormente dé *click* en *Tablas de contingencia*.



Paso 2

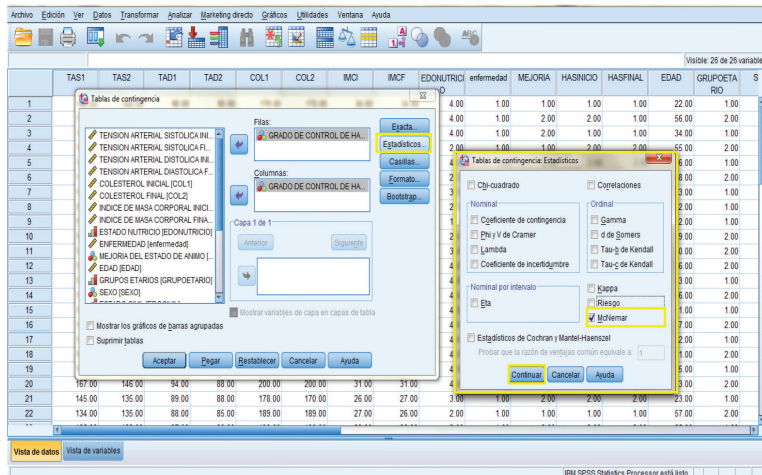
Aparecerá la ventana de *Tablas de contingencia*, seleccione e introduzca la variable dependiente de tipo cualitativa-dicotómica previa a la interacción que le interese en *Fila* y la variable independiente de tipo cua-

litativa-dicotómica post intervención que le corresponda en *Columna*, posteriormente dé *click* en *Estadísticos*. (Por ejemplo, grado de control de hipertensión arterial inicial y grado de control de hipertensión arterial final o después de la intervención.)



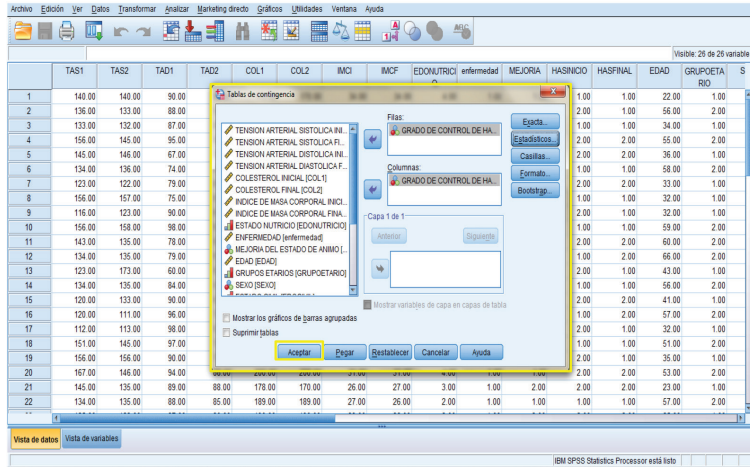
Paso 3

Aparecerá la ventana de *Tablas de contingencia: estadísticos*, dé *click* en *McNemar* y posteriormente en *Continuar*.



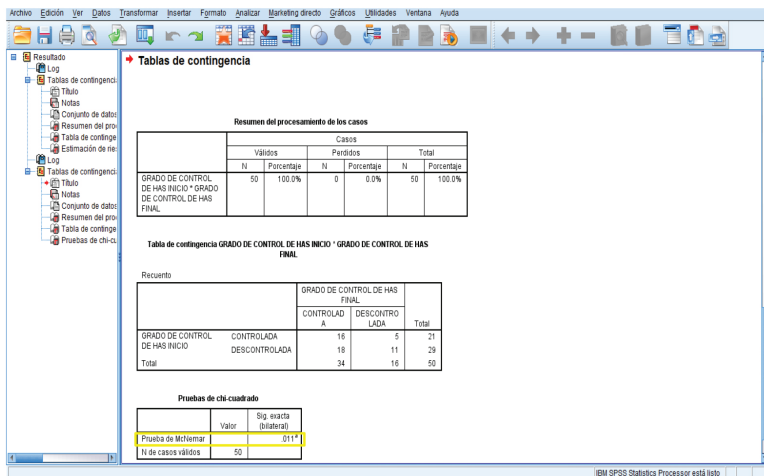
Paso 4

Aparecerá nuevamente la ventana de *Tablas de contingencia* donde deberá dar *click* en *Aceptar*.



Paso 5

De esta forma se obtiene el resultado de la significancia de la prueba de McNemar de las variables analizadas.



Interpretación de resultados

Para valorar si el cambio de respuesta de una variable dicotómica después de un evento o intervención específica en el mismo sujeto es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí hay cambio en la variable después de un evento o intervención específica.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Tabla de contingencia GRADO DE CONTROL DE HAS INICIO * GRADO DE CONTROL DE HAS FINAL

Recuento		GRADO DE CONTROL DE HAS FINAL		Total
		CONTROLAD A	DESCONTR O LADA	
GRADO DE CONTROL DE HAS INICIO	CONTROLADA	16	5	21
	DESCONTR O LADA	18	11	29
Total		34	16	50

Pruebas de chi-cuadrado

	Valor	Sig. exacta (bilateral)
Prueba de McNemar		.011 ^a
N de casos válidos	50	

Por ejemplo, en el análisis realizado a través de la prueba de McNemar se obtuvo una significancia de 0.011 al comparar las variables en la tabla de contingencia.

Al suponer que nuestra hipótesis alterna establezca que sí existe cambio en el grado de control de HAS posterior a la implementación de un cambio de estilo de vida saludable, concluimos que:

- Entre el grado de control de HAS inicio y el grado de control de HAS final sí hubo un cambio estadísticamente significativo, debido a que se obtuvo una significancia ≤ 0.05 , por lo tanto se acepta la hipótesis alterna.

Wilcoxon

La prueba de Wilcoxon también conocida como la prueba de los signos de Wilcoxon, es una prueba no paramétrica para muestras independientes con distribución no normal, utilizada como alternativa de la prueba T de Student para muestras relacionadas.

Su cálculo estima si la diferencia de rangos o medianas en un mismo grupo en diferente tiempo es estadísticamente significativa.

Se utiliza en muestras pequeñas con un tamaño de 6 a 25. Se obtiene con la siguiente fórmula:

$$W^+ = \sum_{Z1>0} R_1$$

En donde:

Σ = Sumatoria

R = Rango

Con tamaños de muestra mayores a 25, se debe calcular el valor Z para buscar normalidad de distribución.

$$Z_T = \frac{T - \bar{X}_T}{\sigma_T}$$

En donde:

T = Valor estadístico de Wilcoxon

\bar{X}_T = Promedio de la T de Wilcoxon

σ_T = Desviación estándar de la T de Wilcoxon

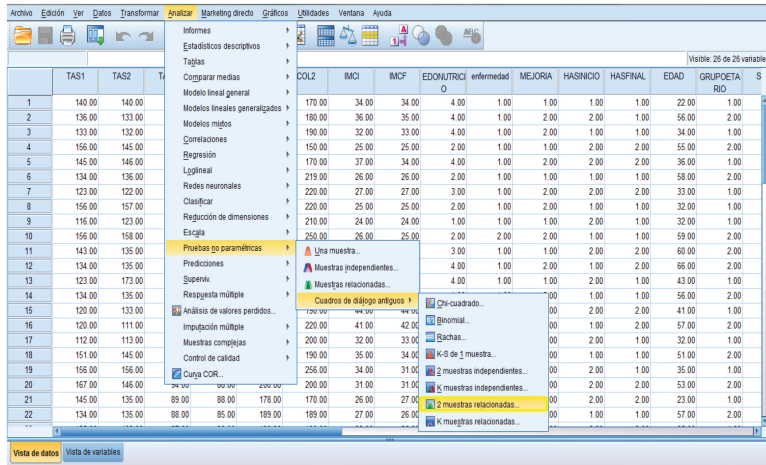
La significancia estadística se considera con el valor de “p”

- < 0.05: es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05: no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de Wilcoxon en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

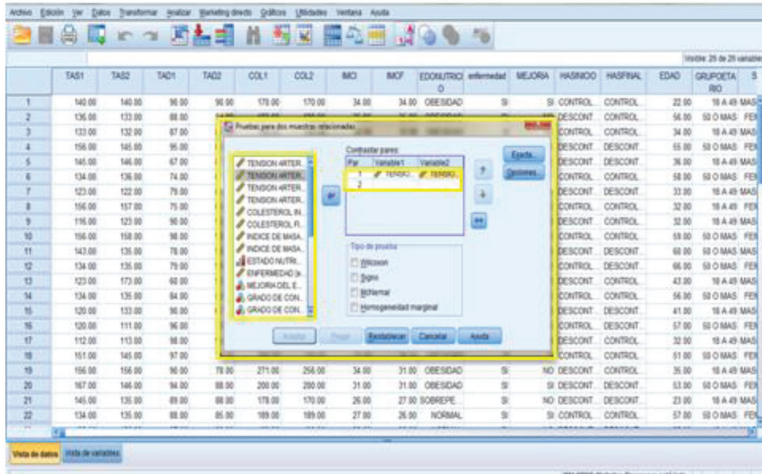
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Pruebas no paramétricas*, posteriormente *Cuadros de diálogo antiguos* y dé *click* en *2 muestras relacionadas*.



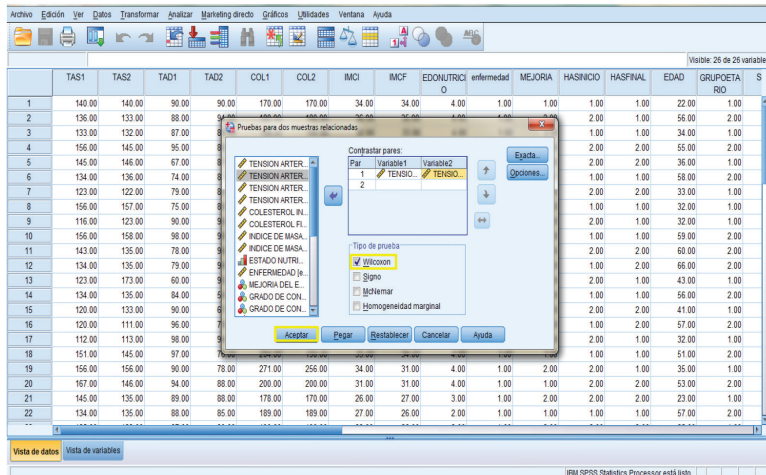
Paso 2

Aparecerá la ventana de *Pruebas para dos muestras relacionadas*, seleccione e introduzca la variable dependiente de tipo cuantitativa previa a la intervención que le interese y la variable independiente de tipo cuantitativa post intervención que le corresponda. (Por ejemplo, tensión arterial inicial y tensión arterial final o después de la intervención.)



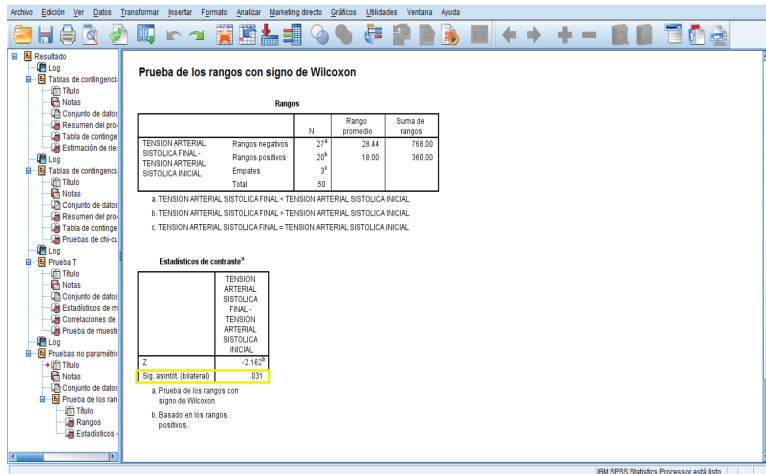
Paso 3

Dé click en Wilcoxon y posteriormente en Aceptar.



Paso 4

De esta forma se obtiene el resultado de la significancia de la prueba de Wilcoxon de las variables analizadas.



Interpretación de resultados

Para valorar si la diferencia entre los rangos o medianas en un mismo grupo en diferente tiempo es significativa, se debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe diferencia entre el grupo posterior a la intervención.
- Si el cambio que existió fue bueno o malo, positivo o negativo.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Prueba de los rangos con signo de Wilcoxon

Rangos

		N	Rango promedio	Suma de rangos
TENSION ARTERIAL SISTOLICA FINAL - TENSION ARTERIAL SISTOLICA INICIAL	Rangos negativos	27 ^a	28.44	768.00
TENSION ARTERIAL SISTOLICA INICIAL - TENSION ARTERIAL SISTOLICA FINAL	Rangos positivos	20 ^b	18.00	360.00
	Empates	3 ^c		
	Total	50		

a. TENSION ARTERIAL SISTOLICA FINAL < TENSION ARTERIAL SISTOLICA INICIAL

b. TENSION ARTERIAL SISTOLICA FINAL > TENSION ARTERIAL SISTOLICA INICIAL

c. TENSION ARTERIAL SISTOLICA FINAL = TENSION ARTERIAL SISTOLICA INICIAL

Estadísticos de contraste^a

	TENSION ARTERIAL SISTOLICA FINAL - TENSION ARTERIAL SISTOLICA INICIAL
Z	-2.162 ^b
Sig. asintót. (bilateral)	.031

Por ejemplo, en el análisis realizado a través de la prueba de Wilcoxon se obtuvo un rango de 18.0 para la variable de tensión arterial sistólica inicial, 28.44 para la variable de tensión arterial sistólica final y una significancia de 0.031 al comparar los rangos de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia entre la tensión arterial sistólica inicial y la tensión arterial final en un mismo grupo posterior a intervención, concluimos que:

- Entre la tensión arterial sistólica inicial y la tensión arterial final sí hubo una diferencia, debido a que se obtuvo un rango de 18.0 para la variable de tensión arterial sistólica inicial y un rango de 28.44 para la variable de tensión arterial sistólica final.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia fue estadísticamente significativa, debido a que se obtuvo una significancia ≤ 0.05 , por lo que se acepta la hipótesis alterna.

U de Mann-Whitney

La prueba de U de Mann Whitney también conocida como Mann-Whitney-Wilcoxon, prueba de suma de rangos Wilcoxon o prueba de Wilcoxon-Mann-Whitney, es considerada la prueba no paramétrica más potente para muestras independientes con distribución no normal, utilizada como alternativa de la prueba T de Student para muestras independientes.

Su cálculo estima la diferencia de medianas o rangos entre dos grupos en el mismo periodo de tiempo. Por ejemplo evaluar en un grupo el uso citología cervicovaginal como detección oportuna de cáncer cervicouterino en relación con un grupo control.

Se obtiene con la siguiente fórmula:

$$U_1 = ((n_1 \cdot n_2)) + \frac{n_1(n_1 + 1)}{2} - \Sigma R_1$$

$$U_2 = ((n_2 \cdot n_1)) + \frac{n_2(n_2 + 1)}{2} - \Sigma R_2$$

En donde:

n_1 = Tamaño de la muestra del grupo 1

n_2 = Tamaño de la muestra del grupo 2

ΣR_1 = Sumatoria de los rangos del grupo 1

ΣR_2 = Sumatoria de los rangos del grupo 2

La prueba de U de Mann-Whitney se utiliza con tamaños de muestra menores de 20, en caso de muestras mayores se debe calcular el valor Z para buscar normalidad de distribución.

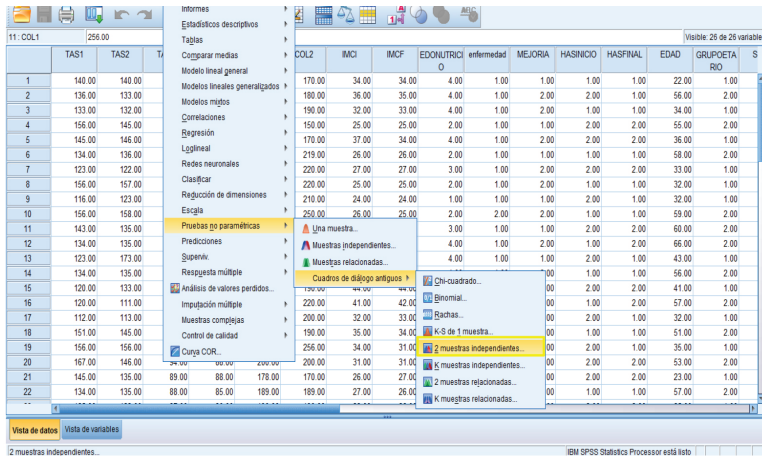
La significancia estadística se considera con el valor de “p”

- < 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de U de Mann-Whitney en el programa SPSS versión 20.0 se realizarán los siguientes pasos:

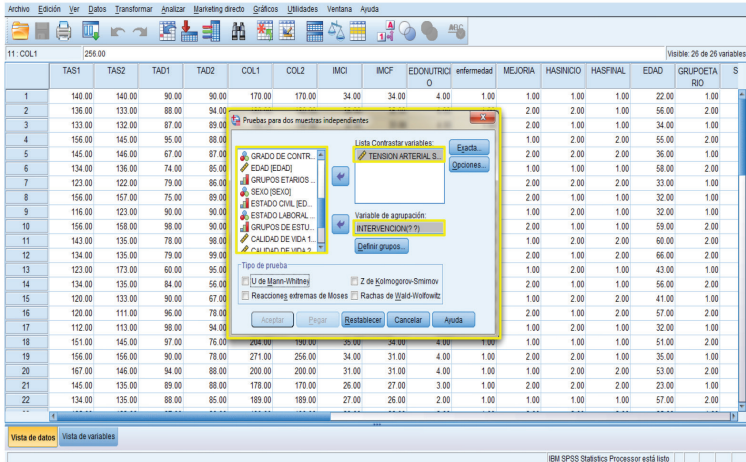
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Pruebas no paramétricas*, posteriormente *Cuadros de diálogo antiguos* y dé *click* en *2 muestras independientes*.



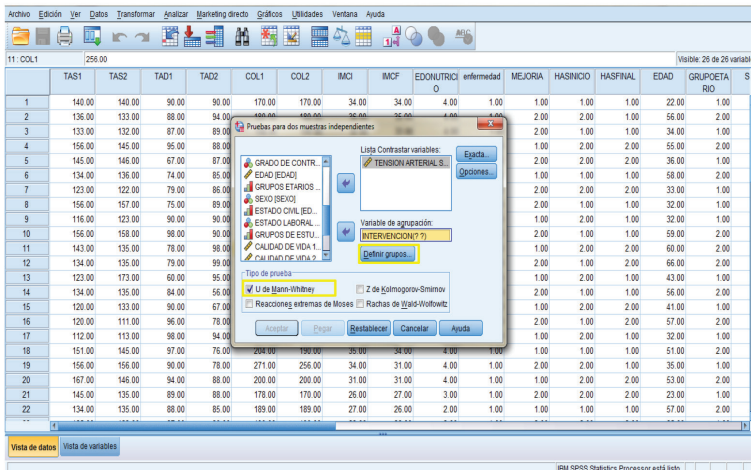
Paso 2

Aparecerá la ventana de *Pruebas para dos muestras independientes*, seleccione e introduzca la variable dependiente de tipo cuantitativa que le interese en *Lista contrastar variables* y la variable independiente de tipo cualitativa-dicotómica que le corresponda en *Variable de agrupación*.



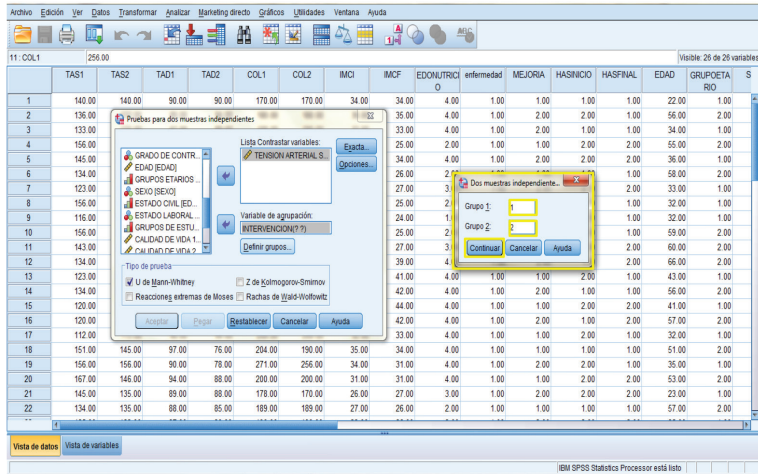
Paso 3

Dé click en U de Mann-Whitney y posteriormente en *Definir grupos*.



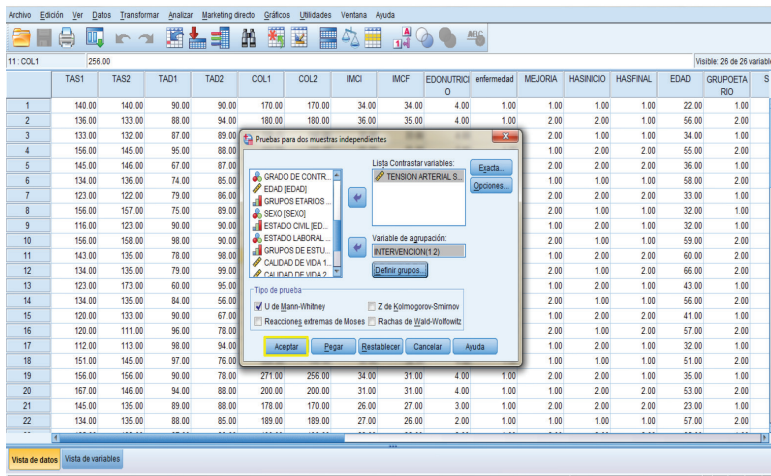
Paso 4

Aparecerá la ventana de *Dos muestras independientes*, donde deberá introducir los grupos que desee incluir en la comparación según el número de valor de etiqueta que asigno en la base de datos y posteriormente dé *click* en *Continuar*.



Paso 5

Aparecerá nuevamente la ventana de *Pruebas para dos muestras independientes* y dé *click* en *Aceptar*.



Paso 6

De esta forma se obtiene el resultado de la prueba de U de Mann-Whitney y la significancia de las variables analizadas.

The screenshot shows the IBM SPSS Statistics interface. The main window displays the results of a Mann-Whitney U test. The window title is "MISSING ANALYSIS". The main area is titled "Pruebas no paramétricas" and "Prueba de Mann-Whitney".

Rangos

	INTERVENCIÓN O ESTUDIO	N	Rango promedio	Suma de rangos
TENSION ARTERIAL	GRUPO CONTROL	25	23.30	582.50
SISTOLICA INICIAL	GRUPO EXPERIMENTAL	25	27.70	692.50
	Total	50		

Estadísticos de contraste^a

	TENSION ARTERIAL SISTOLICA INICIAL
U de Mann-Whitney	257.500
W de Wilcoxon	682.500
Z	-1.073
Sig. asintótico (bilateral)	.283

^a Variable de agrupación: INTERVENCIÓN O ESTUDIO

Interpretación de resultados

Para valorar si la diferencia entre los rangos de dos grupos en el mismo periodo de tiempo es significativa, se debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna la cual establece que sí existe diferencia entre los dos grupos es significativa.
- Si el cambio que existió fue bueno o malo, positivo o negativo.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Prueba de Mann-Whitney

Rangos

INTERVENCIÓN O ESTUDIO		N	Rango promedio	Suma de rangos
TENSIÓN ARTERIAL SISTÓLICA INICIAL	GRUPO CONTROL	25	23.30	582.50
	GRUPO EXPERIMENTAL	25	27.70	692.50
Total		50		

Estadísticos de contraste^a

	TENSIÓN ARTERIAL SISTÓLICA INICIAL
U de Mann-Whitney	257.500
W de Wilcoxon	582.500
Z	-1.073
Sig. asintót. (bilateral)	.283

Por ejemplo, en el análisis realizado a través de la prueba de U de Mann-Whitney se obtuvo un rango de 23.30 para la variable de tensión arterial sistólica inicial del grupo control, 27.70 para la variable de tensión arterial sistólica inicial del grupo experimental y una significancia de 0.283 al comparar los rangos de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia en la tensión arterial sistólica inicial entre el grupo control y el grupo experimental, concluimos que:

- Entre el grupo control y el grupo experimental sí hubo una diferencia, debido a que se obtuvo un rango de 23.30 para la variable de tensión arterial sistólica inicial del grupo control y un rango de 27.70 para la variable de tensión arterial sistólica inicial del grupo experimental.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna

Kruskal-Wallis

La prueba de Kruskal-Wallis es una extensión de la prueba de U de Mann-Whitney, por lo que es considerada la prueba con mayor potencia no paramétrica para muestras independientes con distribución no normal, utilizada como alternativa de la prueba ANOVA.

Su cálculo estima la diferencia de las medianas o rangos entre tres o más grupos en el mismo periodo de tiempo.

Se obtiene con la siguiente fórmula:

$$H = \left[\left[\frac{12}{N \text{ total} \cdot (N \text{ total} + 1)} \right] \right] + \left[\left[\frac{(\Sigma R_1)^2}{n_1} + \frac{(\Sigma R_2)^2}{n_2} + \dots + \frac{(\Sigma R_k)^2}{n_k} \right] \right] - [3 \cdot ((N \text{ total} + 1))]$$

En donde:

$N \text{ total}$ = Tamaño de la muestra

$\Sigma R_1, \Sigma R_2, \dots, \Sigma R_k$ = Sumatoria de los rangos al cuadrado de cada grupo

n_1, n_2, \dots, n_k = Tamaño de la muestra de cada grupo

La significancia estadística se considera con el valor de “p”:

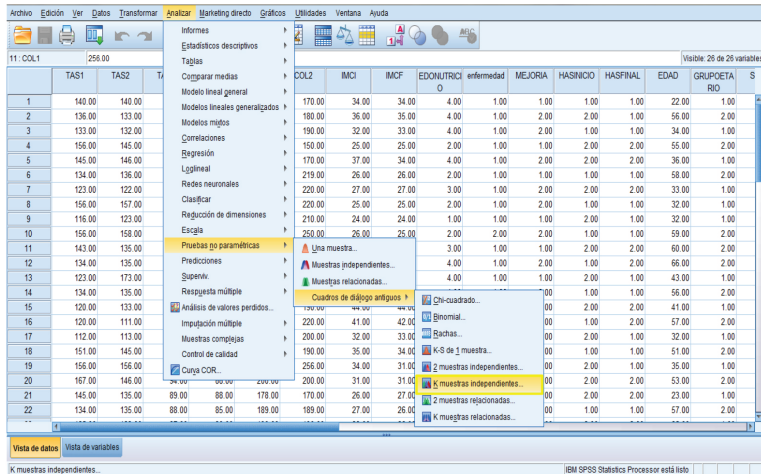
- < 0.05: es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05: no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para determinar cuál es grupo que presenta la diferencia de la mediana o rango, se utiliza la prueba de U de Mann-Whitney para analizar los grupos de dos en dos.

Para realizar el análisis estadístico mediante la prueba de Kruskal-Wallis en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

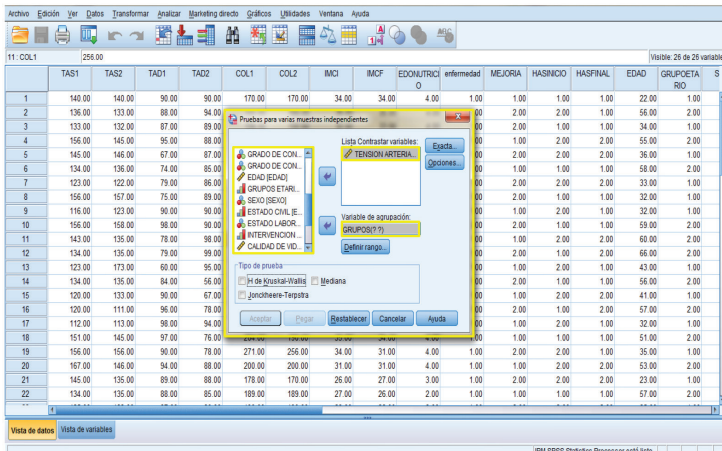
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Pruebas no paramétricas*, posteriormente *Cuadros de diálogo antiguos* y dé *click* en *K muestras independientes*.



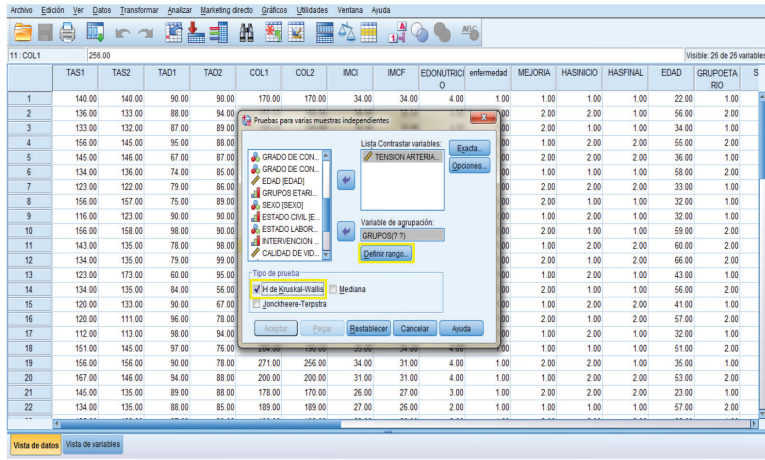
Paso 2

Aparecerá la ventana de *Pruebas para dos muestras independientes*, seleccione e introduzca la variable dependiente de tipo cuantitativa que le interese en *Lista contrastar variables* y la variable independiente de tipo cualitativa – ordinal que le corresponda en *Variable de agrupación*.



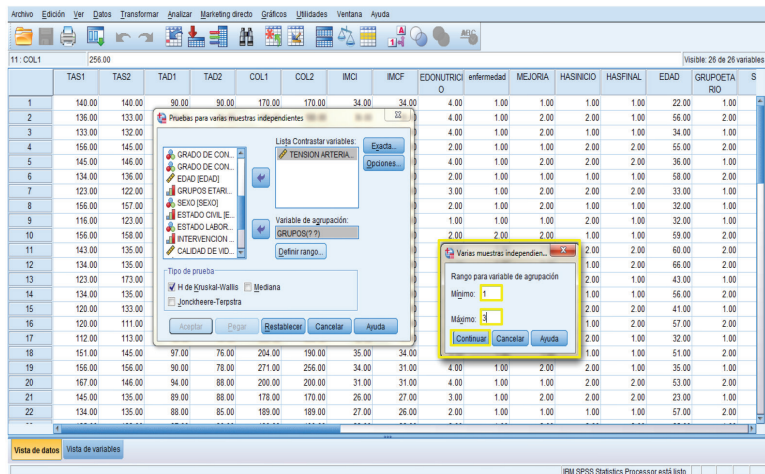
Paso 3

Dé click en *H de Kruskal-Wallis* y posteriormente en *Definir rango*.



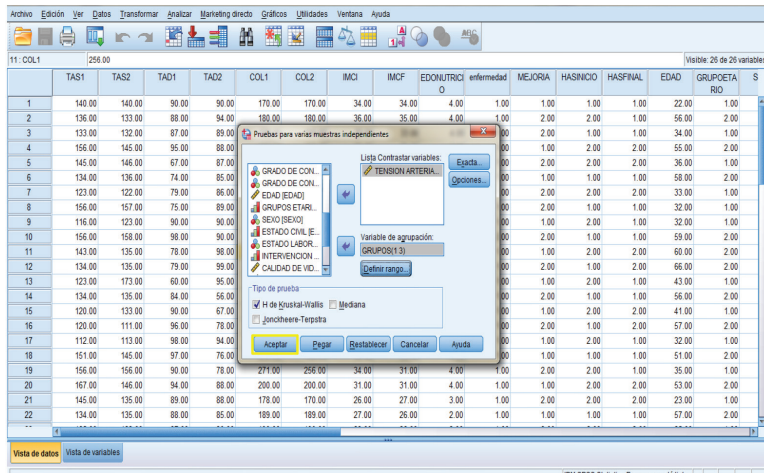
Paso 4

Aparecerá la ventana de *Varias muestras independientes* donde deberá introducir el mínimo y el máximo para la agrupación de las variables, según el número de valor de etiqueta que asigno cuando creo su base de datos, posteriormente dé click en *Continuar*.



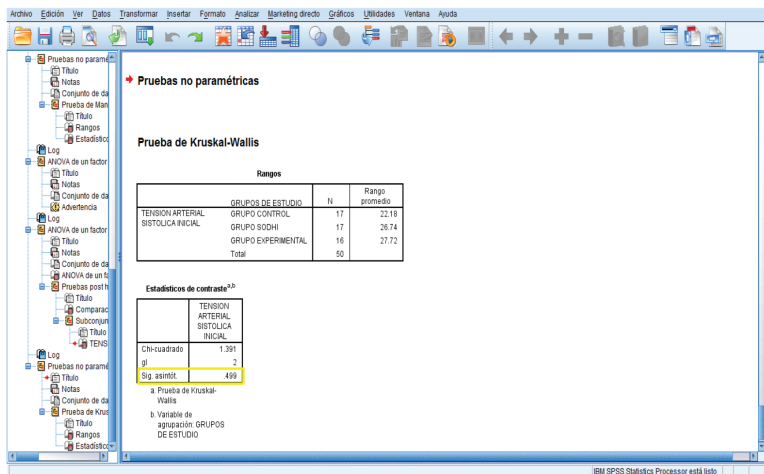
Paso 5

Aparecerá nuevamente la ventana de *Pruebas para varias muestras independientes* y dé click en *Aceptar*.



Paso 6

De esta forma se obtiene el resultado de la significancia de la prueba de Kruskal-Wallis de las variables analizadas.



Interpretación de resultados

Para valorar si la diferencia entre la diferencia de las medianas o rangos entre tres o más grupos en el mismo periodo de tiempo es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que sí existe diferencia entre tres o más grupos en el mismo periodo de tiempo es significativa.
- Si el cambio es estadísticamente significativo.

Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Prueba de Kruskal-Wallis

Rangos			
GRUPOS DE ESTUDIO		N	Rango promedio
TENSION ARTERIAL SISTOLICA INICIAL	GRUPO CONTROL	17	22.18
	GRUPO SODHI	17	26.74
	GRUPO EXPERIMENTAL	16	27.72
Total		50	

Estadísticos de contraste^{a,b}

	TENSION ARTERIAL SISTOLICA INICIAL
Chi-cuadrado	1.391
gl	2
Sig. asintót	.499

Por ejemplo, en el análisis realizado a través de la prueba de Kruskal-Wallis se obtuvo un rango de 22.18 para la variable de tensión arterial sistólica inicial del grupo control, 26.74 para la variable de tensión arterial sistólica inicial del grupo SODHI, 27.72 para la variable de tensión arterial sistólica inicial del grupo experimental y una significancia de 0.499 al comparar los rangos de las variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia en la tensión arterial sistólica inicial entre el grupo control, el grupo SODHI y el grupo experimental, concluimos que:

- Entre el grupo control, el grupo SODHI, y el grupo experimental sí hubo una diferencia, debido a que se obtuvo un rango de 22.18 para la variable de tensión arterial sistólica inicial del grupo control, 26.74 para el grupo SODHI, 27.72 para el grupo experimental.
- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna.

Correlación de Spearman

El coeficiente de correlación de Spearman también conocida como Rho Spearman, es una prueba no paramétrica para muestras independientes con distribución no normal, utilizada como alternativa de la Correlación de Pearson.

Su cálculo estima la asociación entre la variable dependiente e independiente determinando en qué proporción una variable interviene para que otra se modifique.

Se obtiene con la siguiente fórmula:

$$rs = 1 - \frac{6 \Sigma d^2}{n^3 - n}$$

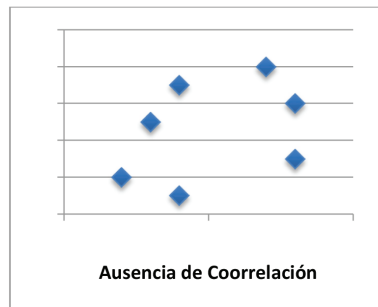
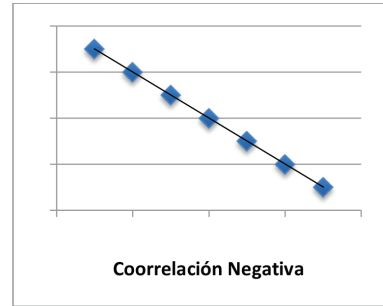
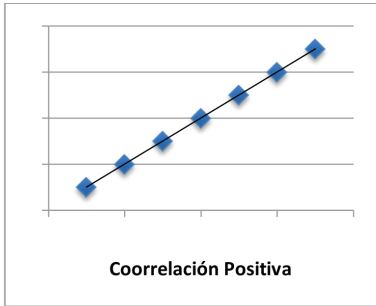
En donde:

Σ = Sumatoria

n = Tamaño de la muestra en rangos de las variables

d^2 = Diferencia entre los rangos de las dos variables al cuadrado

Se grafica a través de coordenadas cartesianas, en donde la variable dependiente corresponde al eje de la “y” mientras que la variable independiente corresponde al eje de la “x”.



Sus resultados oscilan entre +1 a -1 y se interpretan como:

- $r = +1$ o -1 : correlación perfecta (positiva o negativa).
- $r = 0$: ausencia de correlación.
- $r = 0$: ausencia de correlación.
- $r = 0.80 - 0.99$: correlación fuerte.
- $r = 0.60 - 0.79$: correlación moderada.
- $r = 0.40 - 0.59$: correlación parcial.
- $r = 0.20 - 0.39$: correlación ligera.
- $r = 0.00 - 0.19$: correlación fortuita o insignificante.

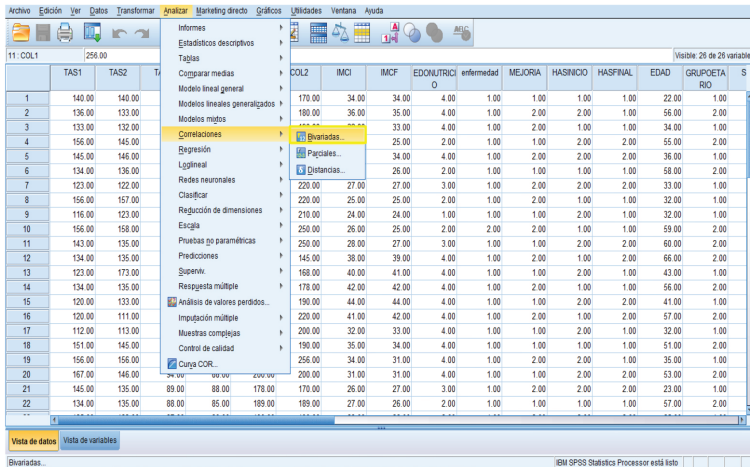
La significancia estadística se considera con el valor de “p”:

- ≤ 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la correlación de Spearman en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

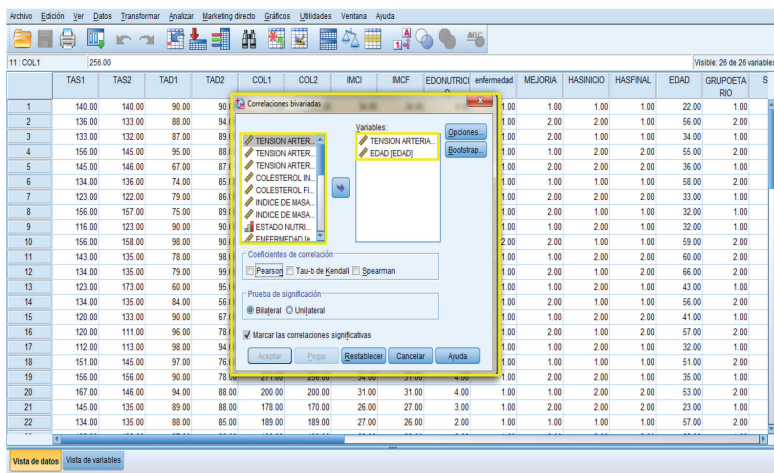
Paso 1

En la barra de herramientas dé *click* en *Analizar*, seleccione *Correlaciones* y posteriormente dé *click* en *Bivariadas*.



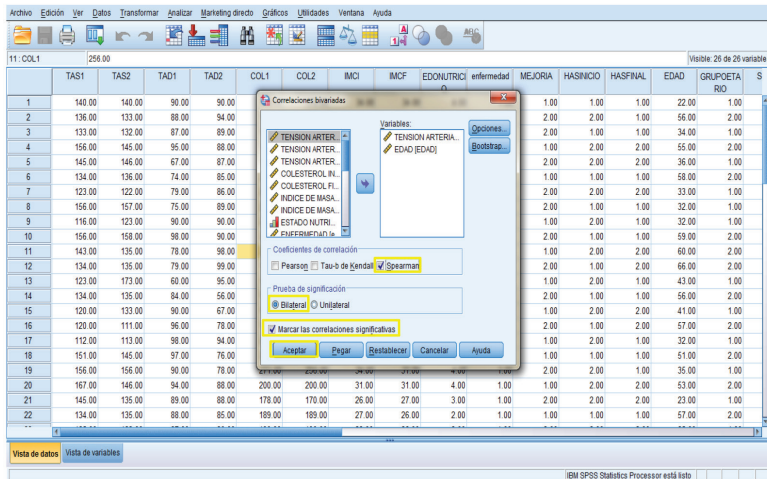
Paso 2

Aparecerá la ventana de *Correlaciones bivariadas*, seleccione e introduzca la variable dependiente e independiente de tipo cuantitativo que le interese correlacionar en *Variables*.



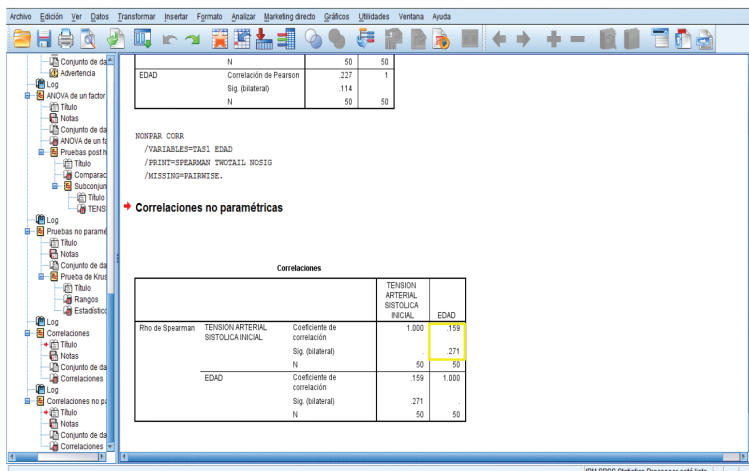
Paso 3

De click en Spearman, en Bilateral, en Marcar las correlaciones significativas y posteriormente en Aceptar.



Paso 4

De esta forma se obtiene el resultado de la correlación de Spearman y la significancia de las variables correlacionadas.



Interpretación de resultados

Para valorar si la proporción en que la variable independiente interviene para que la variable dependiente se modifique es significativa, debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece si la proporción en que una variable interviene para que otra se modifique es estadísticamente significativo.

Recordando que los resultados de la correlación de Spearman oscilan entre +1 a -1 y se interpretan como correlación perfecta (ya sea positiva o negativa) cuando el valor de r es de +1 o -1, ausencia de correlación cuando el valor de r es de 0, correlación fuerte cuando el valor de r es de 0.80 a 0.99, correlación moderada cuando el valor de r es de 0.60-0.79, correlación parcial cuando el valor de r es de 0.40-0.59, correlación ligera cuando el valor de r es de 0.20-0.39 y correlación fortuita o insignificante cuando el valor de r es de 0.01 a 0.19.

Y una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.

Correlaciones

			TENSION ARTERIAL SISTOLICA INICIAL	EDAD
Rho de Spearman	TENSION ARTERIAL SISTOLICA INICIAL	Coefficiente de correlación	1.000	.159
		Sig. (bilateral)	.	.271
		N	50	50
	EDAD	Coefficiente de correlación	.159	1.000
		Sig. (bilateral)	.271	.
		N	50	50

Por ejemplo, en el análisis realizado a través de la correlación de Spearman se obtuvo un coeficiente de correlación de 0.159 y una significancia de 0.271 al valorar la asociación entre la variable de tensión arterial sistólica inicial y la variable de edad.

Al suponer que nuestra hipótesis alterna establezca que la edad interviene para que la tensión arterial sistólica inicial se modifique concluimos que:

- La relación entre la tensión arterial sistólica inicial y la edad no fue estadísticamente significativa, es decir, la edad no modifica la ten-

sión arterial sistólica inicial, debido a que se obtuvo una correlación insignificante y una significancia > 0.05 , por lo tanto se rechaza la hipótesis alterna.

Regresión lineal

La **regresión lineal** también conocida como **ajuste lineal** permite representar gráficamente la relación de las variables.

Su cálculo estima la asociación entre la variable dependiente e independiente en la regresión lineal simple o entre la variable dependiente y varias variables independientes en la regresión lineal múltiple.

Se obtiene con la siguiente fórmula:

- Regresión lineal simple

$$Y = B_0 + B_1 \cdot x + u$$

En donde:

Y = Variable dependiente

B_0, B_1 = Parámetros a estimar

x = Variable independiente

u = Residual o término de error, que en el caso de poblaciones es 0

- Regresión lineal múltiple

$$Y = B_0 + B_1 \cdot X_1 + B_2 \cdot X_2 + \dots + B_k \cdot X_k + u$$

En donde:

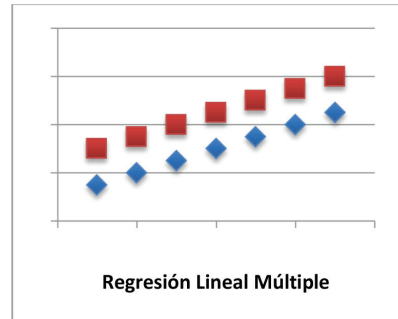
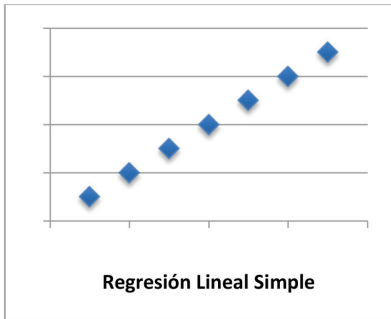
Y = Variable dependiente

$B_0, B_1, B_2, \dots, B_k$ = Parámetros a estimar

x = Variable independiente

u = Residual o término de error, que en el caso de poblaciones es 0

Se grafica a través de coordenadas cartesianas, en donde la variable dependiente corresponde al eje de la “y” mientras que la variable independiente corresponde al eje de la “x”.



Sus resultados oscilan entre +1 a -1 y se interpretan como:

- $r = 1$: correlación perfecta.
- $r = 0$: ausencia de correlación.

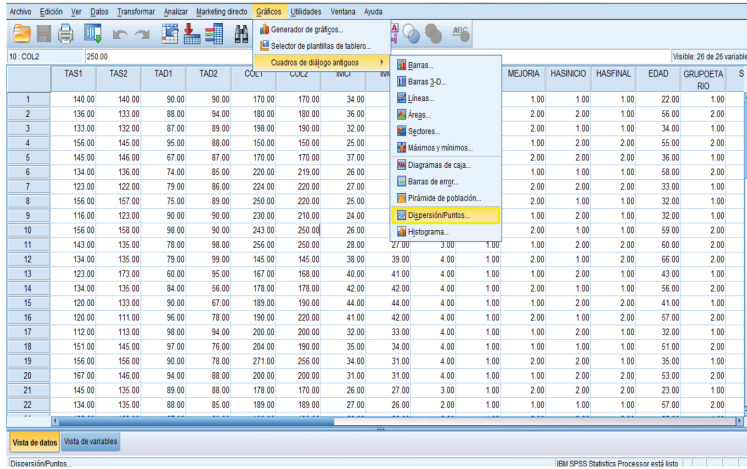
La significancia estadística se considera con el valor de “p”.

- ≤ 0.05 : es estadísticamente significativo, se acepta la hipótesis alterna y se rechaza la hipótesis nula.
- > 0.05 : no es estadísticamente significativo, se acepta la hipótesis nula y se rechaza la hipótesis alterna.

Para realizar el análisis estadístico mediante la prueba de regresión lineal en el programa SPSS versión 20.0 se realizarán los siguientes pasos.

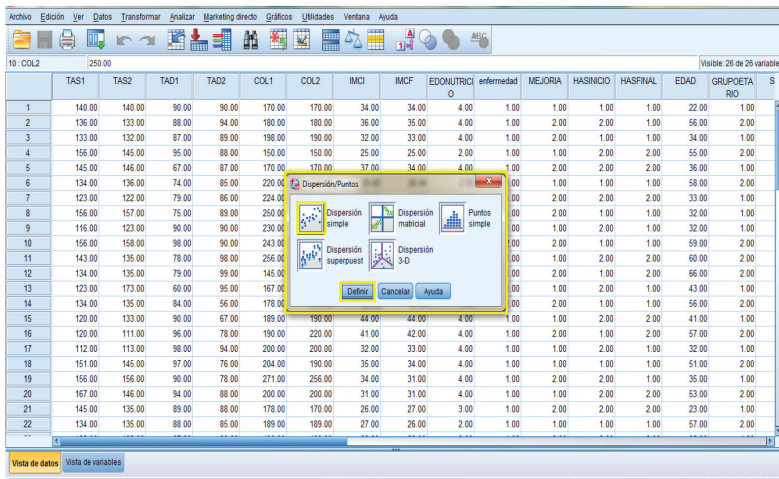
Paso 1

En la barra de herramientas dé *click* en *Gráficos*, seleccione *Cuadros de diálogo antiguos* y posteriormente dé *click* en *Dispersión/Puntos*.



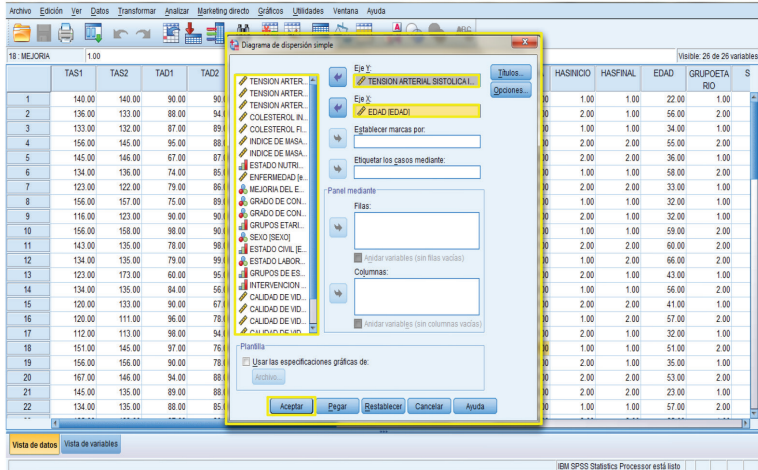
Paso 2

Aparecerá la ventana de *Dispersión/Puntos*, seleccione *Dispersión simple* y dé *click* en *Definir*.



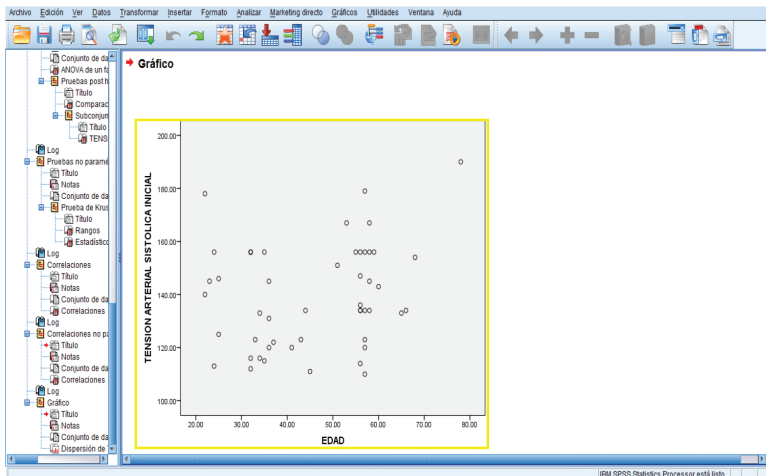
Paso 3

Aparecerá la ventana de *Diagrama de dispersión simple*, seleccione e introduzca la variable dependiente que le interese en Eje Y, la variable independiente que le interese relacionar en Eje X y posteriormente en *Aceptar*.



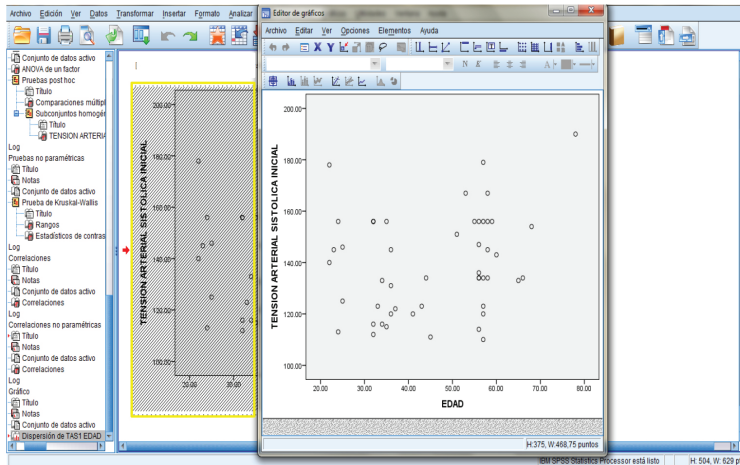
Paso 4

De esta forma se obtiene la gráfica de correlación.



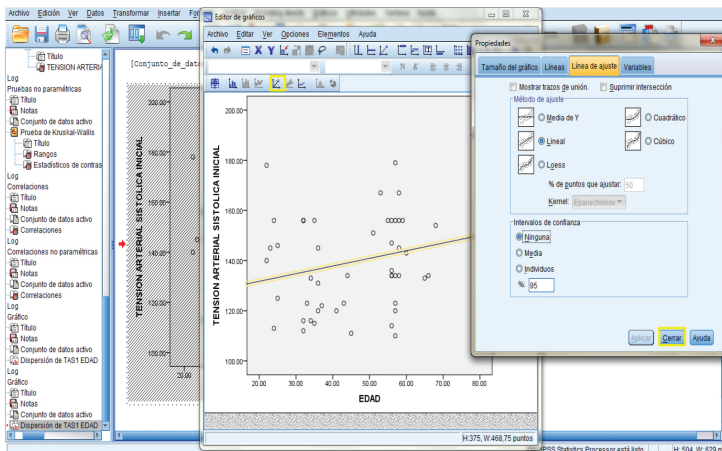
Paso 5

Para editar la gráfica de correlación y obtener la línea de correlación y el valor de la regresión lineal o R^2 lineal, dé doble *click* sobre la imagen del *Gráfico de correlación* y aparecerá la ventana de *Editor de gráficos*.



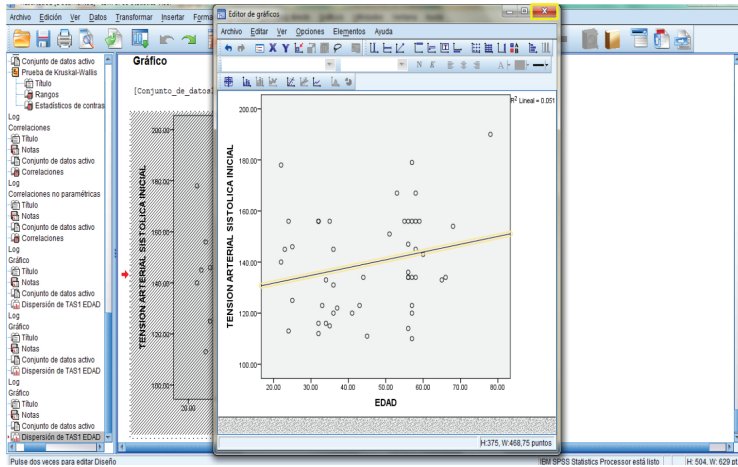
Paso 6

Dé *click* en *Añadir línea de ajuste total* y automáticamente aparecerá la *Línea de ajuste total* sobre la *Gráfica de correlación*, posteriormente dé *click* en *Cerrar*.



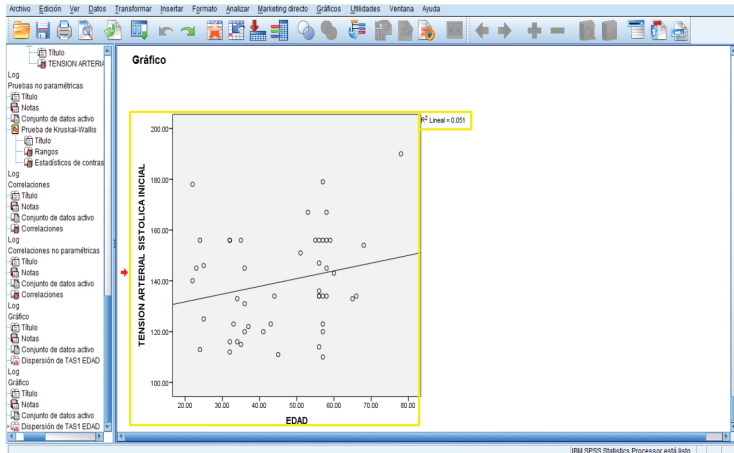
Paso 7

Cierre la ventana de *Editor de gráficos*.



Paso 8

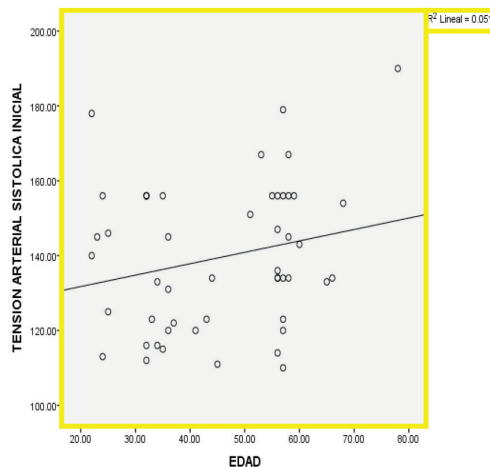
De esta forma se obtiene la línea de correlación y el valor de la regresión lineal o R^2 lineal del gráfico de correlación seleccionado.



Interpretación de resultados

Para valorar si la diferencia entre las medias de dos grupos en el mismo periodo de tiempo es significativa, se debe tomar en cuenta:

- Si hubo cambio en relación con la hipótesis alterna, la cual establece que la diferencia entre los dos grupos es significativa.
- Si el cambio que existió fue bueno o malo, positivo o negativo.
- Si el cambio es estadísticamente significativo.
- Recordando que una significancia con valor de $p \leq 0.05$ indica que la diferencia es estadísticamente significativa, mientras que un valor de $p > 0.05$ indica que la diferencia no es estadísticamente significativa.



Por ejemplo, en el análisis realizado a través de la regresión lineal se obtuvo una R^2 lineal de 0.051 para la variable de tensión arterial sistólica inicial del grupo control, 143.36 para la variable de tensión arterial sistólica inicial del grupo experimental y una significancia de 0.169 al comparar las medias de ambas variables.

Al suponer que nuestra hipótesis alterna establezca que sí existe una diferencia en la tensión arterial sistólica inicial entre el grupo control y el grupo experimental, concluimos que:

- Entre el grupo control y el grupo experimental sí hubo una diferencia, debido a que se obtuvo una media de 135.72 para la variable de tensión arterial sistólica inicial del grupo control y una media de 143.36 para la variable de tensión arterial sistólica inicial del grupo experimental.

- Que la diferencia que se presentó entre las variables fue positiva, debido a que sí existe una diferencia entre el grupo control y el grupo experimental.
- Y que esta diferencia no fue estadísticamente significativa, debido a que se obtuvo una significancia > 0.05 .

Breviario estadístico

H0: Hipótesis nula
H1: Hipótesis alterna
Ho: Hipótesis nula
Ha: Hipótesis alterna
K-S: Kolmogorov-Smirnov
OR: Odds Ratio

Bibliografía

- AMIR Medicina (2010). Academia de Estudios MIR, S.L. Marbán Libros. España.
- Arguedas Arguedas, O. (2010). Tipos de diseño en estudios de investigación. *Acta Médica Costarricense*, 52, 1, enero-marzo.
- Castro Jiménez, M. A., Díaz Martínez, L. A. (2009). Las variables en el proceso de investigación en salud: Importancia, clasificación y forma de presentación en protocolos de investigación. *Revista de la Facultad de Ciencias de la Salud*, Universidad Autónoma de Bucaramanga, Colombia, 12, 3, diciembre.
- Estadística. Recuperado de: http://www.ray-design.com.mx/psicoparaest/index.php?option=com_content&view=section&id=8&Itemid=55. Fecha de consulta: 11 de abril de 2012.
- Gómez Gómez, M., Danglot Banck, C., Velásquez Jones, L. (2001). Bases para la revisión crítica de artículos médicos. *Revista Mexicana de Pediatría*, 68, 4, julio-agosto.
- (2003). Sinopsis de pruebas estadísticas no paramétricas. Cuándo usarlas. *Revista Mexicana de Pediatría*, 70, 2, marzo-abril.
- Hernández Ávila, M., Garrido Latorre, F., López Moreno, S. (2000). Diseño de estudios epidemiológicos. *Salud Pública de México*, 42, 2, marzo-abril.
- Pita Fernández, S. (2001). Tipos de estudios clínico epidemiológicos. Atención Primaria en la Red. Recuperado de www.fisterra.com. Consultado el 01 de abril de 2012.
- , Pértega Díaz, S. (s. f.). Asociación de variables cualitativas: Test de Chi-cuadrado. Recuperado de www.fisterra.com. Consultado el 11 de abril de 2012.
- , Pértega Díaz, S. (2001). Estadística descriptiva de los datos. Atención Primaria en la Red. Recuperado de www.fisterra.com. Consultado el 01 de abril de 2012.
- Pruebas no paramétricas. Recuperado de http://www.uclm.es/actividades0708/cursos/estadistica/pdf/descargas/spss_PruebasNoParametricas.pdf. Consultado el 11 de abril de 2012.

Tabla de distribución de Chi². Recuperado de: http://www.uam.es/personal_pdi/ciencias/gpepe/g-poblaciones/clases/Tema_04/TablaChi2.htm. Consultado el 11 de abril de 2012.

Talavera, J. O., Rivas Ruiz, R. (2011). Pertinencia de la prueba estadística seleccionada. *Revista Médica del Instituto Mexicano del Seguro Social*, 49, 2.

Veiga de Cabo, J., De la Fuente Díez, E., Zimmermann Verdejo, M. (2008). Modelos de estudio en investigación aplicada: Conceptos y criterios para el diseño. *Medicina y Seguridad del Trabajo*, 54, 210, marzo.

Manual de análisis estadístico y uso de bases de datos
se terminó de imprimir en junio de 2015
en los talleres de Ediciones de la Noche
Madero #687, Zona Centro
Guadalajara, Jalisco
El tiraje fue de 500 ejemplares.

www.edicionesdelanoche.com

Elaboración de tablas y gráficas
Carlos Alfonso Yeo Barba
Sergio Christopher Rodríguez González
Yicoaldo Camacho Ruiz
Karen Beatriz Romero Pérez
Iván Alejandro Gardiel Mireles

La ciencia actual se basa en el inductivismo, es decir, todo conocimiento está sustentado en la evidencia obtenida mediante el método experimental.

La ciencia médica en particular obtiene sus conocimientos a partir del inductivismo, por lo que es fundamental contar con los conocimientos del análisis estadístico para realizar una correcta investigación en salud.

Este manual tiene como objetivo facilitar el uso del programa SPSS versión 20.0, con el propósito de analizar las variables de estudio.



UNIVERSIDAD DE GUADALAJARA
CENTRO UNIVERSITARIO DE LA COSTA

ISBN: 978-607-742-216-7



9 786077 422167