

**UNIVERSIDAD SAN PEDRO
FACULTAD DE INGENIERÍA**

GUÍA DE SPSS

Statistical Package for the Social Sciences

Autor. Dr. Antonio Alva Santos

2008.

INTRODUCCIÓN ESTADÍSTICA DESCRIPTIVA (CONCEPTOS BÁSICOS)

La estadística tiene como finalidad estudiar detalladamente las características de una muestra de datos. Este estudio se realiza a través de variables. Variables: Son cantidades que pueden tomar distintos valores de un momento a otro, y son propiedades de los casos que pueden tomar distintos valores. Ej. En la investigación de ejemplo se estudian las características de una muestra de 150 sujetos. Se estudian siete variables: Edad, sexo, estatus, región, ci, estatura y peso. La edad es una propiedad (tiempo vivido) de los casos (sujetos en este caso) que pueden tomar distintos valores. Sexo es una variable porque puede tomar distintos valores: hombre o mujer.

Variables y Casos: Todo conjunto de datos puede ser visto como una matriz. Las filas se corresponden con los distintos sujetos o unidades de investigación, sobre los que se han realizado las distintas observaciones. Ésta últimas se corresponden con las columnas. Los sujetos (filas) suelen denominarse casos y las observaciones (columnas) suelen denominarse variables. A modo de ejemplo, por caso, dependiendo del estudio, puede entenderse un paciente, o bien una muestra de laboratorio, o una visita médica o un episodio de recurrencia de una enfermedad. Por variables, asociadas a los casos anteriores, puede entenderse el sexo y la edad del paciente, la fórmula leucocitaria de la muestra de laboratorio, la fecha y presión arterial medidas en una determinada visita médica. Es importante notar que, con frecuencia se confunde, en este último caso, cuál es el sujeto de análisis, si el paciente con varios episodios de recurrencia o cada episodio de recurrencia es un caso en sí, pudiendo haber varios episodios para un mismo paciente. La distinción de la unidad de análisis en cada situación la marcará el objetivo de la investigación.

Tipo de variables		
Cualitativas		Cuantitativas
Contienen información cualitativa de los sujetos denotando la cualidad o la categoría a la que pertenece un sujeto.		Miden la magnitud de una medida sobre un individuo.
Nominales	Ordinales	Escala
Las observaciones se clasifican en categorías (mutuamente excluyentes y colectivamente exhaustivas) sin que éstas tengan un orden específico (por ejemplo los grupos sanguíneos). Un caso particular de datos nominales es el caso de los datos dicotómicos en los que existen sólo dos categorías (como por ejemplo en el sexo).	En esta escala de medición no tiene sentido medir la distancia entre posibles valores de la variable ni se pueden realizar operaciones aritméticas con ellos pues no toman valores numéricos específicos ni existe proporcionalidad entre categorías vecinas. Sí pueden sin embargo ser ordenados pues existe una relación de orden entre las distintas clases o categorías. El pronóstico de una enfermedad (<i>menos grave, grave, más grave</i>) es un dato medido en escala ordinal.	Estas variables miden una magnitud de los sujetos resultando en un valor numérico. Utilizan escalas numéricas y sí tiene sentido medir la distancia que hay entre posibles valores de los datos. En estas variables cabe distinguir aquellos casos donde el punto cero es arbitrario (medición a nivel de intervalo) como por ejemplo la temperatura, o datos donde el cero es un punto fijo (nivel de razón) como por ejemplo el peso o la glucemia. Si bien el SPSS no hace distinción entre variables del tipo escala, los procedimientos estadísticos distinguen entre variables discretas y continuas . En el primer caso el conjunto de posibles valores de la variable es finito siendo en el segundo continuo. Un ejemplo de variable discreta es el número de días de ingreso en un Hospital, y un ejemplo de variable continua es la glucemia.

Tipo de Variables			
<p>Cualitativas o Nominales</p> <ul style="list-style-type: none"> Sus valores son distintos. <p>Ej. Sexo: hombre, mujer Región: norte, este, oeste, sur. Se estudian mediante:</p> <ul style="list-style-type: none"> La distribución de frecuencias La gráfica de barras o de sectores Un estadístico descriptivo: La MODA. 	<p>Cuantitativas</p> <ul style="list-style-type: none"> Sus valores son distintos y se pueden ordenar (de mayor a menor). 		
	<p>Ordinales</p> <ul style="list-style-type: none"> Valores distintos. Se pueden ordenar. <p>Ej- estatus socioeconómico: Alto, Medio, Bajo Se estudian con:</p> <ul style="list-style-type: none"> Distribución de frecuencias. Gráfica de sectores si hay pocos valores distintos. Descriptivos: mediana, máximo, mínimo, rango o amplitud. Con más de 10 valores distintos se analizan como variables de intervalo. En el caso de depuración de ítems de pruebas y encuestas se utilizan los índices de asimetría y apuntamiento 	<p>De Intervalo</p> <ul style="list-style-type: none"> Valores distintos. Se pueden ordenar. Existe una unidad común. <p>Ej- CI</p>	<p>De Razón</p> <ul style="list-style-type: none"> Valores distintos. Se pueden ordenar. Existe una unidad común. Existe un cero real. <p>Ej. Peso, edad, estatura</p>
	<p>Se estudian con los mismos estadísticos:</p> <ul style="list-style-type: none"> Histograma Descriptivos: media, desviación típica, máximo, mínimo, amplitud, los coeficientes de asimetría y apuntamiento. 		
<ul style="list-style-type: none"> Relaciones de las variables: Sexo, estatus; región <p>Mediante: Tablas de contingencia (Crosstabs), Chi cuadrado, Coef. De contingencia, Phi, V de Cramer, Lambda. Gráfica de Barras. Porcentaje por celdilla.</p>	<ul style="list-style-type: none"> Relaciones bivariadas: ci, peso y estatura <p>Mediante: Coeficiente de correlación de Pearson. Gráficas de dispersión (Scatter). Recta de regresión lineal.</p>		

APLICANDO UN CUESTIONARIO

UNIVERSIDAD PRIVADA SAN PEDRO

Sr(a)(ta): Se solicita su valiosa colaboración que nos permita recolectar datos para analizarlos e interpretarlos de un Trabajo de Investigación que favorecerá a la ciudad de Chimbote.

Escriba en el espacio en blanco o con la letra "X" según corresponda.

GRACIAS POR SU COLABORACIÓN

- 1.- N° de Identificación de la encuesta: (Valores del 001 al 150)

- 2.- Edad (en años cumplidos):

- 3.- Sexo: Hombre () Mujer ()

- 4.- Escriba solamente su primer o único nombre:

- 5.- En forma relativa indique el estatus socioeconómico en el que se ubicaría:
Alto () Medio () Bajo ()

- 6.- De las alternativa dadas ubique, la región o zona en que se ubica su domicilio:
Norte () Este () Oeste () Sur ()

- 7.- Escriba el valor relativo de su cociente intelectual. (Escriba valores como mínimo desde 52 hasta un valor máximo de 159):

- 8.- Estatura (en centímetros) :

- 9.- Peso en kilogramos (Con aproximación a un decimal):

PRACTICA N° 01: DEFINICIÓN DE VARIABLES DEL CUESTIONARIO AL PAQUETE ESTADÍSTICO

1.- Ingrese al programa estadístico SPSS. Active la ventana Variable New. Nombre del archivo a Guardar: (Escriba su Apellido)

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	iden	Numérico	3	0	Número de ide	Ninguno	-9	8	Derecha	Escala
2	edad	Numérico	2	0		Ninguno	-9	8	Derecha	Ordinal
3	sexo	Numérico	1	0		{1, Hombre}...	9	8	Derecha	Ordinal
4	nombre	Cadena	10	0		Ninguno	Ninguno	10	Izquierda	Nominal
5	estatus	Numérico	1	0	Estatus socioe	{1, Alto}...	9	8	Derecha	Ordinal
6	región	Numérico	1	0		{1, Norte}...	9	8	Derecha	Ordinal
7	ci	Numérico	3	1	Cociente intele	Ninguno	-9,0	8	Derecha	Escala
8	estatura	Numérico	3	0		Ninguno	-9	8	Derecha	Escala
9	peso	Numérico	5	1		Ninguno	-9,0	8	Derecha	Escala

Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
Iden	Numerico	3	0	Número de identificación del sujeto	Ninguno	-9	8	Derecha	Escala
edad	Numerico	2	0		Ninguno	-9	8	Derecha	Ordinal
Sexo	Numerico	1	0		1 → Hombre	9	8	Derecha	Ordinal
					2 → Mujer				
nombre	Cadena	10			Ninguno		10	Derecha	Nominal
Estatus	Numerico	1	0	Estatus socioeconómico	1 → Alto	9	8	Derecha	Ordinal
					2 → Medio				
					3 → Bajo				
región	Numerico	1	0		1 → Norte	9	8	Derecha	Ordinal
					2 → Este				
					3 → Oeste				
					4 → Sur				
ci	Numerico	3	1	Cociente Intelectual	Ninguno	-9.0	8	Derecha	Escala
estatura	Numerico	3	0		Ninguno	-9	8	Derecha	Escala
peso	Numerico	5	1		Ninguno	-9.0	8	Derecha	Escala

PRACTICA N° 02: INGRESO DE DATOS

1.- Mediante trabajo en equipo (5 personas), ingresará al archivo que contiene la definición de las variables, deberá dividir el trabajo para ingresar los datos de las 150 encuestas. Para lo cual copie el archivo (Ej. ALVA.SAV), en cada una de las computadoras de los integrantes de su equipo. Luego cada uno, abrirá el archivo ALVA y podrá ingresar, por ejemplo los datos de 30 casos. Cada alumno grabará el archivo por ejemplo, como: ALVA-(1-30); ALVA-(31-60); ALVA-(61-90); ALVA-(91-120); ALVA-(121-150).

2.- Una vez ingresados todos los datos. Unir todos los archivos en uno solo. Una (Merge File). Utilizará el menú: DATA---Merge Files --- Add Cases.

Guarde todos los datos ingresados en un solo archivo con el nombre, por ejemplo: ALVA – (1-150). Cada integrante cambiará el nombre del archivo con su propio apellido, quedará: MIAPELLIDO-(1-150)

DATOS RECOLECTADOS A PARTIR DEL CUESTIONARIO APLICADO PARA UN TRABAJO DE INVESTIGACIÓN.

IDEN	EDAD	SEXO	NOMBRE	ESTATUS	REGIÓN	CI	ESTATURA	PESO
1	15	2	Abigail	1	3	124	154	57.9
2	16	1	Carlos	1	3	95	154	56.7
3	31	1	Alberto	1	2	101	156	52.8
4	24	2	Adela	2	3	116	184	104.2
5	24	1	Luis	1	1	102	173	87.5
6	32	2	Adelina	2	1	102	170	71.5
7	26	1	Antonio	1	1	81	176	93.3
8	21	1	Juan	1	4	109	188	99.2
9	22	2	Adriana	2	4	61	169	76.6
10	18	2	Agata	2	1	93	175	87.9
11	25	1	Cosme	9	9	119	177	80.3
12	29	2	Aida	3	1	52	181	82.3
13	25	2	Alameda	1	1	128	164	61.8
14	19	1	Juan	2	1	65	160	75.0
15	30	1	Carlos	1	1	116	162	60.3
16	29	2	Alana	1	1	65	156	53.9
17	29	1	Jose	3	4	147	172	78.6
18	16	2	Albertina	1	1	113	175	85.7
19	31	1	Armando	1	3	99	168	75.4
20	18	1	Constante	2	4	117	159	73.0
21	21	1	Toribio	1	2	105	178	87.6
22	29	1	Borja	1	2	92	174	74.2
23	23	2	Alejandra	3	4	78	178	100.9
24	28	2	Alicia	1	4	97	163	77.1
25	19	1	Manuel	3	4	77	180	74.6
26	31	1	Carlos	3	4	81	174	70.2
27	33	1	Antonio	2	1	126	137	99.3
28	17	2	Alma	1	3	95	153	62.1
29	34	2	Almeda	2	1	122	168	77.4
30	20	2	Almeria	1	2	85	161	67.9

IDEN	EDAD	SEXO	NOMBRE	ESTATUS	REGIÓN	CI	ESTATURA	PESO
31	26	2	Almira	1	3	123	270	86.8
32	25	1	Jose	2	2	94	167	70.5
33	15	2	Alvina	1	4	101	173	85.7
34	23	2	Amalia	1	3	80	187	95.6
35	20	1	Luis	1	3	95	161	65.5
36	34	1	Hugo	1	1	110	172	81.4
37	23	2	Amanda	2	3	-9	162	57.2
38	16	1	Jose	3	3	88	166	64.1
39	19	2	Ambar	2	2	75	170	78.5
40	19	1	Julian	2	2	85	164	65.4
41	35	1	Jeremias	3	2	75	176	82.6
42	30	1	Manuel	3	1	90	177	82.1
43	29	2	Anabel	2	3	70	164	69.9
44	32	2	Anastasia	2	2	111	147	59.4
45	34	1	Sergio	2	1	70	167	84.2
46	33	2	Andrea	3	2	116	169	79.4
47	32	2	Ana	2	3	159	155	55.1
48	18	1	Sergio	2	3	145	163	71.3
49	23	1	Norberto	3	2	65	164	72.8
50	28	1	Gustavo	1	3	76	176	75.2
51	17	1	Gustavo	1	1	106	161	75.9
52	27	1	Manuel	1	3	74	174	82.4
53	22	2	Angela	2	1	90	163	76.7
54	23	2	Angélica	2	1	101	162	71.9
55	21	1	Alberto	1	2	68	188	95.7
56	22	2	Anselma	2	3	77	155	68.5
57	19	2	Antonieta	1	4	95	141	54.4
58	22	1	Antonio	1	3	95	184	90.1
59	30	1	Jeremias	3	4	63	161	80.4
60	26	2	Antonina	1	2	124	171	72.9
61	23	2	Arabela	3	4	118	160	65.3
62	23	2	Ariadna	1	1	101	163	67.8
63	24	1	Gabriel	3	4	92	148	58.8
64	24	2	Ariana	2	2	117	174	79.5
65	31	2	Arlena	1	4	140	180	77.0
66	34	1	Diego	1	1	102	191	99.5
67	20	2	Arlinda	2	2	87	174	84.6
68	18	2	Astra	1	2	88	158	68.5
69	21	1	Diego	1	1	95	159	61.8
70	27	2	Astrid	1	2	86	151	63.6
71	16	1	Gabriel	1	4	87	159	54.0
72	33	2	Atalanta	1	2	113	156	71.7
73	23	2	Augusta	1	3	68	153	57.1
74	32	2	Aurea	1	1	102	175	90.0

IDEN	EDAD	SEXO	NOMBRE	ESTATUS	REGIÓN	CI	ESTATURA	PESO
75	20	2	Aurora	2	1	74	173	73.1
76	15	1	Nicomedo	3	1	94	154	68.8
77	26	1	Nicolas	3	3	144	180	94.3
78	26	1	Juan	1	3	130	167	78.9
79	22	1	Andrés	2	3	114	176	84.3
80	16	2	Bárbara	3	1	118	160	64.0
81	21	1	Millán	2	4	118	186	86.7
82	28	2	Beatriz	2	2	128	151	62.6
83	17	1	Chema	1	3	102	161	59.8
84	16	2	Benedicta	2	2	88	173	96.9
85	27	1	Humberto	1	2	115	169	80.4
86	20	1	Jose	1	2	86	161	64.7
87	20	2	Betina	1	4	111	160	66.2
88	-9	2	Bianca	3	2	96	157	53.5
89	28	2	Bibi	1	3	146	165	74.0
90	20	1	Felipe	1	1	99	191	94.1
91	15	1	Antón	1	1	113	173	76.0
92	21	2	Brigida	1	4	123	152	58.2
93	28	1	Federico	1	1	76	131	93.7
94	28	2	Camila	1	4	118	167	70.5
95	24	2	Carlota	2	3	129	176	79.4
96	21	2	Carola	2	3	116	161	69.1
97	16	2	Carolina	1	3	107	175	88.7
98	20	1	Luis	1	2	130	183	83.3
99	26	2	Carula	1	4	98	162	75.5
100	24	1	Juan	1	2	111	168	72.8
101	33	1	Juan	1	2	60	172	84.7
102	17	1	Jose	1	1	125	183	86.2
103	29	1	Fadrique	2	1	69	179	95.1
104	21	2	Carolona	1	4	69	160	78.7
105	30	1	Roberto	1	1	100	184	96.7
106	18	1	Jose	2	3	121	187	94.8
107	33	1	Alberto	1	2	140	175	84.5
108	33	1	Emilio	1	1	110	172	79.0
109	21	2	Caroline	1	3	140	157	71.6
110	20	2	Carmela	3	4	141	188	92.0
111	21	2	Casandra	1	2	108	176	81.4
112	20	2	Catia	3	1	106	173	79.6
113	26	2	Celeste	1	1	135	166	77.6
114	25	1	Hugo	2	1	100	179	95.4
115	30	2	Caridad	3	4	99	165	87.8
116	19	2	Charo	3	2	154	168	72.9
117	15	1	Juan	3	3	136	164	75.5
118	26	2	Clara	3	2	86	159	57.7
119	27	2	Clarisa	1	4	93	158	59.4

IDEN	EDAD	SEXO	NOMBRE	ESTATUS	REGIÓN	CI	ESTATURA	PESO
120	19	2	Claudia	2	3	112	180	97.1
121	28	1	Niobio	1	3	101	175	82.4
122	22	2	Conchita	1	3	62	162	77.0
123	30	1	Nobel	2	1	59	174	88.4
124	33	2	Consuelo	1	3	62	162	77.0
125	23	1	Julian	3	3	75	161	69.9
127	34	2	Dalida	1	4	93	170	74.0
128	34	1	Juan	3	2	147	179	87.5
129	18	1	Jose	2	4	97	152	50.4
130	33	1	Alberto	1	4	118	184	94.5
131	24	2	Daniela	2	3	85	170	62.9
132	16	1	Jonas	2	3	116	172	81.3
133	19	2	Debora	1	2	127	168	79.3
134	25	2	Diana	3	2	109	164	73.5
135	27	1	Julian	1	1	61	182	93.4
136	32	2	Dominga	3	4	87	164	67.7
137	24	2	Dorotea	2	1	93	171	81.6
138	26	1	Julio	1	2	89	172	72.6
139	16	2	Dulcinea	3	3	100	166	72.2
140	33	1	Martín	2	3	111	171	85.0
141	23	1	Manuel	3	1	104	160	75.2
142	17	2	Elena	2	2	127	165	65.8
143	21	1	Manel	1	1	124	174	77.9
144	22	1	Jordi	1	1	109	166	75.0
145	32	2	Eloisa	3	3	110	157	59.0
146	27	1	Aitor	1	3	100	179	91.5
147	29	1	Lumbago	1	1	97	190	92.7
148	20	1	Yul	2	2	80	174	79.3
149	27	1	Honorio	3	1	107	190	90.7
150	-9	1	Roberto	9	2	128	160	-9.0

NOTA:

- Deberá tener una copia de seguridad de su archivo que contiene los 150 registros.
- Cualquier error cometido por alguno de los integrantes de su grupo, se reflejará en los resultados al trabajar en el programa SPSS.

PRACTICA N° 03 CREACIÓN Y TRANSFORMACIÓN DE VARIABLES

Abra el archivo MIAPELLIDO.sav: File → Open. Busca el archivo: MIAPELLIDO.sav. Luego Abrir.

FUNDAMENTO: Los datos directos necesitan en muchas ocasiones ser transformados antes de llevar a cabo los análisis estadísticos. Se pueden realizar con las opciones del menú Transformar (Transform) o con lenguaje de comandos SPSS.

RECODIFICACION DE VARIABLES (RECODE).

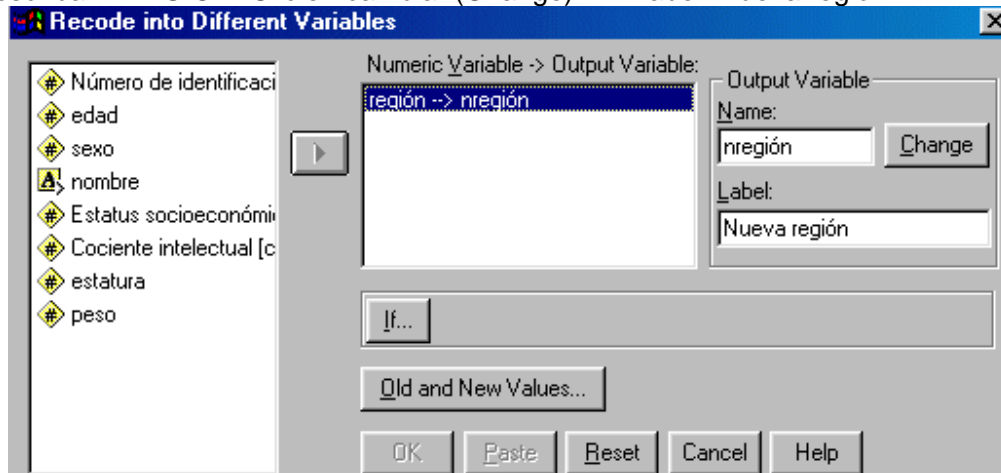
Elija el menú Transformar → Recodificar (Transform → Recode). Existen dos posibilidades: “En distintas variables” recodifica la variable y crea una variable nueva con esta recodificación; y la opción “En las mismas variables” recodifica la variable, es decir, los datos originales de la variable se sustituyen por los nuevos.

1.- RECODIFICAR de la variable REGION sus alternativas ESTE y OESTE a una sola con el nombre CENTRO creando una nueva variable con el nombre NREGION con esta recodificación.

Variable: <u>REGION</u>		Valores recodificados en la Variable: <u>NREGION</u>	
Value (Valor)	Value Label (Etiquetas)	Antiguos → Nuevos (Old → New)	Value Label (Etiquetas)
1	Norte	1 → 1	1 → Norte
2	Este	2 → 2	2 → Centro
3	Oeste	3 → 2	3 → Sur
4	Sur	4 → 3	

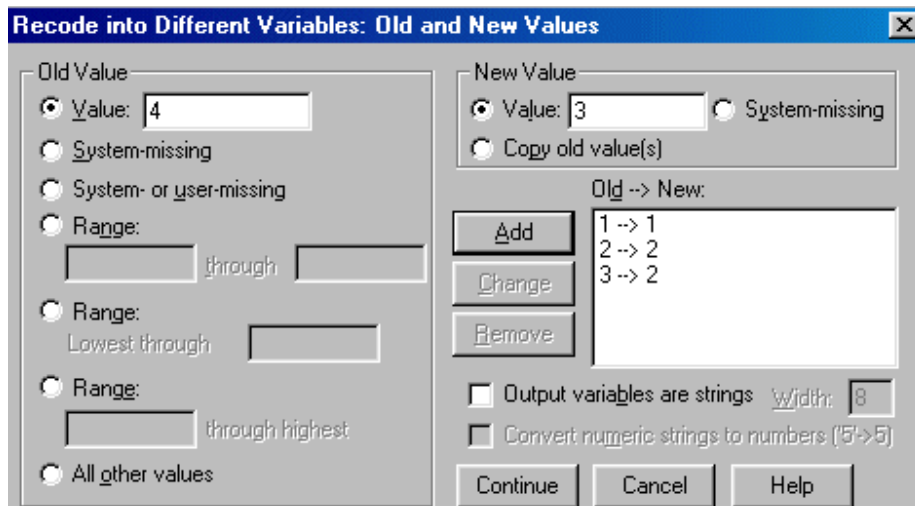
PASOS A SEGUIR:

- a) Transformar → Recodificar → En distintas variables. (Transform-Recode-Into Different Variables...).
- b) Clic en REGION y en el botón central de Transferencia. En el recuadro Nombre (Name) escriba: NREGION. Clic en cambiar (Change). En Label: Nueva región

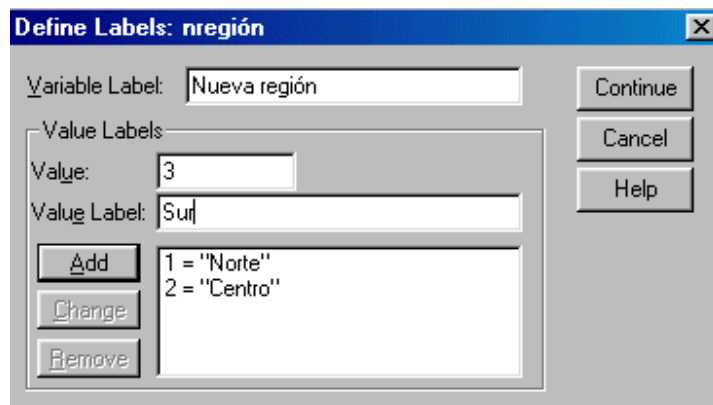


- c) Clic en valores antiguos y nuevos (Old and New Values)
- d) En la sección Valor antiguo (Old Value) situarse en la casilla Valor (value) y escribir el valor antiguo: 1
- e) En la sección Valor nuevo (New Value) situarse en la casilla Valor (Value) y escribir el nuevo valor: 1

- f) Haga clic en el botón añadir (Add). La recodificación se habrá escrito en la sección Antiguo → Nuevo (Old → New).
- g) Repita los pasos d, e y f para los valores antiguos 2 y 3. Escriba como valor nuevo (New value) 2 para ambos casos.
- h) Repita los pasos d, e y f para el valor antiguo 4 y escriba 3 como nuevo valor.
- i) Pulse el botón continuar (Continue) y aceptar (OK).



- j) Finalmente en la ventana editor de datos, haga doble clic en la nueva variable NREGION.
- k) Edita Type: Width:2 y Decimal places: 0;
- l) Para las etiquetas, clic en la opción LABEL (Etiqueta) y cambie los valores de esta nueva variable así: Variable label: Nueva Región. Value:1, Value label:Norte; Value:2, Value label:Centro, Value:3, Value label:Sur. Clic en continuar y aceptar (OK).



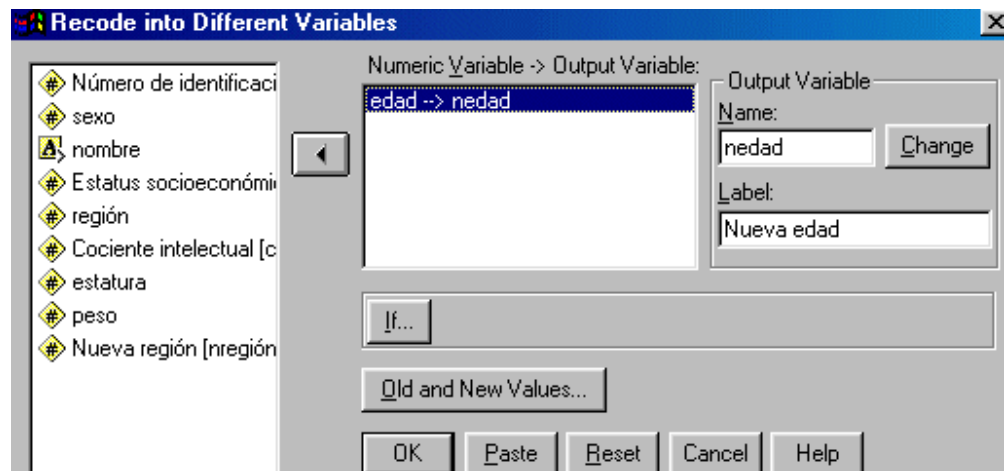
- m) Para missing values. Escriba: 9 dentro del recuadro de Discrete missing values. Corrija en la tabla de datos, en caso de haber algún valor perdido, escribiendo: 9
- n) Para hacer otra recodificación diferente con otra variable. Clic en el botón Reset
- o) File → Save as: MIAPELLIDO.SAV.

RECODIFICACIÓN DE RANGO DE VALORES.

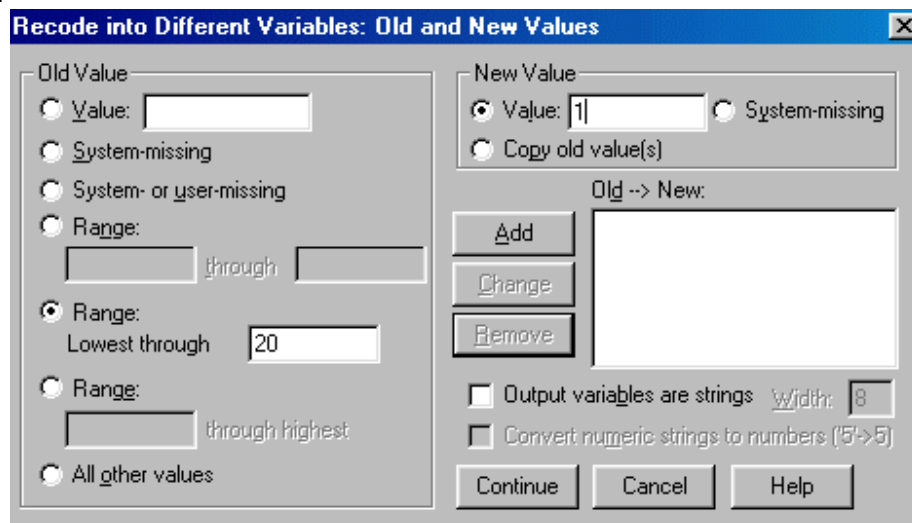
2.- Recodificar la variable EDAD (continua) en grupos de edad del archivo: ROSA- (1-150).SAV, de acuerdo a la siguiente tabla:

EDAD	Del menor a 20	De 20 a 25	De 25 a 30	De 30 al mayor
NEDAD	1	2	3	4

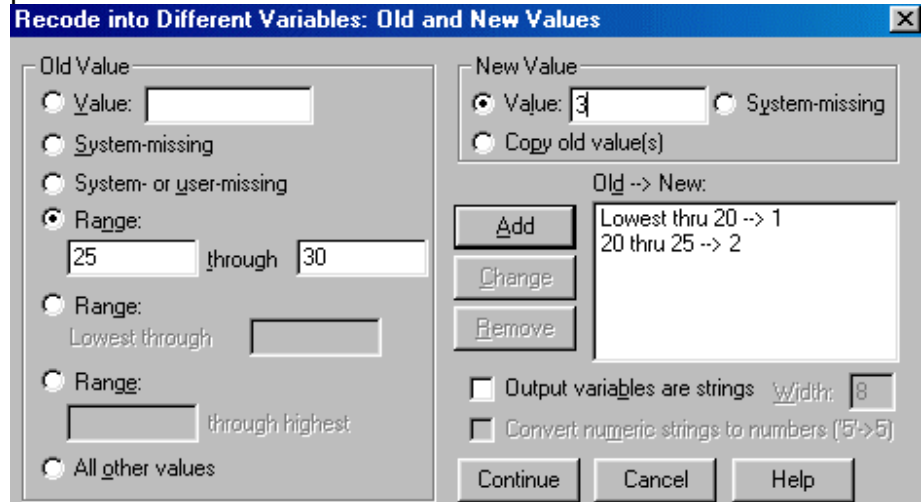
- Transformar → Recodificar → En distintas variables. (Transform-Recode-Into Different Variables...).
- Para hacer otra recodificación diferente con otra variable. Clic en el botón Reset
- Clic en EDAD y en el botón central de Transferencia. En el recuadro Nombre escriba: NEDAD. Clic en Cambiar (Change). En Label: Nueva edad.



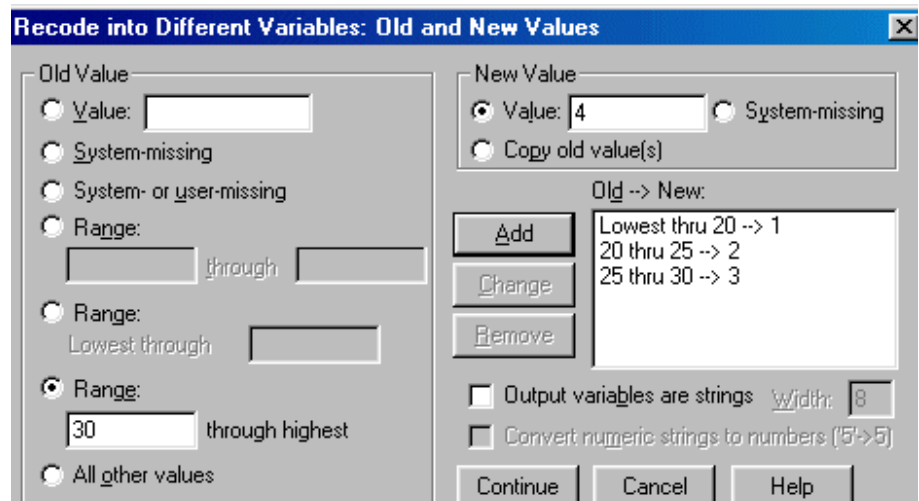
- Clic en valores antiguos y nuevos. (Old and New Values)
- En la sección Valor antiguo (Old Value) marcar el círculo al lado de Rango (Range) (ubicada en la parte central lado izquierdo) y en la celdilla al lado de “Del menor hasta” (Lowest through) escribir: 20. En la sección Valor nuevo (New value) escribir en Value:1 y elegir el botón añadir (Add). Aparece: Lowest thru 20 → 1



- Marcar el círculo al lado de Rango (Range) (ubicada encima de la anterior) y en la celdilla a la izquierda de “hasta” (through) escribir: 20 y en la de la derecha escribir: 25, en Valor nuevo escribir en Value: 2 y pulsar el botón añadir (Add). (Se recodificarán como 2 los valores de edad mayores que 20 y menores o iguales a 25.) Aparece: 20 thru 25 → 2



- Marcar el círculo al lado de Rango (Range) (ubicada encima de la anterior) y en la celdilla a la izquierda de “hasta” (through) escribir: 25 y en la de la derecha escribir: 30, en Valor nuevo escribir en Value: 3 y pulsar el botón añadir (Add). (Se recodificarán como 3 los valores de edad mayores que 25 y menores o iguales a 30.) Aparece: 25 thru 30 → 3
- Elegir el círculo al lado de Rango (Range) (ubicada al final de los anteriores) y en la celdilla a la izquierda de “hasta el mayor” (Through Highest) escribir: 30, en Valor nuevo(New Value) escribir en Value: 4 y elegir el botón añadir (Add). Aparece: 30 thru Highest → 4



- Pulsar el botón Continuar (Continue) y después el botón Aceptar (OK).
- Editar la variable NEDAD. Los datos deben estar sin decimales. Las etiquetas (Labels) deben ser: Value:1; Value Label:Del menor a 20; 2:De 20 a 25; 3:De 25 a 30 y 4:De 30 al mayor.

CREACIÓN O TRANSFORMACIÓN DE VARIABLES (COMPUTE)

Si se elige Transformar → Calcular se pueden crear variables nuevas o transformar las ya existentes mediante operadores aritméticos, funciones aritméticas, funciones estadísticas, funciones lógicas y otras (Comando COMPUTE).

Con esta primera opción dentro del menú **Transformar** podemos crear nuevas variables asignando a éstas los valores resultantes de operaciones algebraicas.

3.- A partir de los datos de peso (Kg.) y altura (en cm) de un individuo, calcular su índice de masa corporal de acuerdo a la expresión algebraica:

$$\text{índice de masa corporal} = \text{peso} / (\text{estatura} / 100)^2.$$

- Transformar → Calcular. (Transform –Compute)
- Escribir: IMC en el casillero Variable de destino(Target Variable)
- En la casilla Expresión numérica se construye la definición de esta variable (Se puede emplear la lista de las variables, los operadores aritméticos, los números y las funciones):
- Transferir la variable Peso a la ventana Expresión numérica (Numeric Expresión). Clic en el operador: /. Luego clic en los paréntesis. Introducir la variable Estatura. Fuera del paréntesis, clic en **** para elevar a la potencia**. Luego clic en el número: 2 para elevarlo al cuadrado y clic en OK.
- Editar la variable de ser necesario.

d	sexo	nombre	estatus	región	ci	estatura	peso	imc
15	2	Abigail	1	3	124	154	57.9	24.41
16	1	Carlos	1	3	95	154	56.7	23.91
31	1	Alberto	1	2	101	156	52.8	21.70
24	2	Adela	2	3	116	184	104.2	30.78
								29.24
								24.74
								30.12
								28.07
								26.82
								28.70
								25.63
								25.12
								22.98
								29.30

- **Grabar como: File → Save as.** Nombre de archivo: MIAPELLIDO
Clic en Guardar.
- **Cerrar el archivo MIAPELLIDO.SAV**

PRACTICA N° 04 CÁLCULO CON VARIABLES TIPO FECHA

FUNDAMENTO: Para el cálculo de la edad en años de un individuo sabiendo la fecha actual y su fecha de nacimiento, debemos recurrir a la función denominada

CTIME.DAYS(fecha).

Supongamos la fecha actual 15 de marzo del 2003, y la variable f_nacim conteniendo la fecha de nacimiento de los casos. Introduciendo en la ventana de las expresiones la siguiente:

(CTIME.DAYS(CTIME.DAYS(03,15,2003))-CTIME.DAYS(f_nacim))/365

Nota: El formato de fecha puede variar.

Como resultado obtendremos los años transcurridos entre ambas fechas. Si queremos un mayor grado de pulcritud podemos dividir entre 365.25 días para tomar en cuenta los bisieptos o usar la función **RND(exp_numérica)** para redondear el resultado al entero más próximo.

La función **DATE.DMY(día,mes,año)** transforma los argumentos día, mes y año a un dato tipo fecha para poder ser empleado, al igual que f_nacim, como argumento de la función **CTIME.DAYS(valorfecha)**

1.- Crear el archivo: FECHAS.SAV.

- **File → New → Data.**

a) Definir las siguientes variables. Doble clic sobre Var. Luego llenar según Tabla:

Variable Name	FNACIM	FINGRESO	FALTA
Type	Date dd.mm.yyyy	Date dd.mm.yyyy	Date dd.mm.yyyy
Variable Label	Fecha de nacimiento	Fecha de ingreso	Fecha de alta

b) Ingresar los siguientes 5 registros :

	Fnacim	Fingreso	Falta	
1	07.12.1959	10.03.2003	14.03.2003	
2	01.04.1956	01.03.2003	14.03.2003	
3	01.01.1990	01.01.2002	01.01.2003	
4	31.12.1980	31.12.1999	31.12.2002	
5	28.02.1990	30.12.1999	05.05.2003	

c) Calcular la variable Edad, expresada en años cumplidos a partir de la fecha actual.

(Ej. Supongamos que la fecha actual es: 15 de marzo del 2003).

$$EDAD = FECHA \text{ ACTUAL } (15,03,03) - FECHA \text{ DE NACIMIENTO}$$

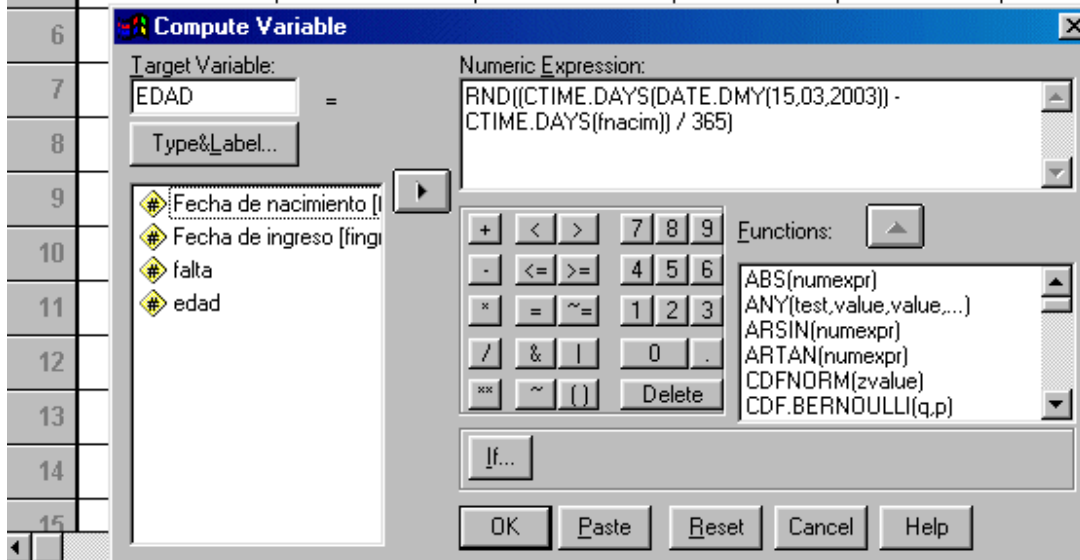
Es decir, la EDAD se calculará ingresando las variables y fechas:

Transform → Compute

$$RND((CTIME.DAYS(DATE.DMY(15,03,2003))-CTIME.DAYS(fnacim))/365)$$

Resultados:

	fnacim	fingreso	falta	edad	var	va
1	07.12.1959	10.03.2003	14.03.2003	43		
2	01.04.1956	01.03.2003	14.03.2003	47		
3	01.01.1990	01.01.2002	01.01.2003	13		
4	31.12.1980	31.12.1999	31.12.2002	22		
5	28.02.1990	30.12.1999	05.05.2003	13		



d) CALCULA LAS NUEVAS VARIABLES:

- SEGUNDOS = Los segundos permanecidos en condición de internados comprendidos entre la fecha de ingreso y la fecha de alta.
- DÍAS = Los días comprendidos entre la fecha de ingreso y la fecha de alta.
- DÍAS1 = Los días internados incluyendo el día de ingreso.
- AÑOS = Los años permanecidos internados.

Resultados:

	fnacim	Fingreso	Falta	SEGUNDOS	DÍAS	DÍAS1	AÑOS
1	07.12.1959	10.03.2003	14.03.2003	345600	4	5	0
2	01.04.1956	01.03.2003	14.03.2003	1123200	13	14	0
3	01.01.1990	01.01.2002	01.01.2003	31536000	365	366	1
4	31.12.1980	31.12.1999	31.12.2002	94694400	1096	1097	3
5	28.02.1990	30.12.1999	05.05.2003	1.06+08	1222	1223	3

Solución:

SEGUNDOS = falta – fingreso

DÍAS = CTIME.DAYS(falta) – CTIME.DAYS (fingreso)

DÍAS1 = (CTIME.DAYS(falta) – CTIME.DAYS (fingreso) + 1

AÑOS = RND((CTIME.DAYS(falta) – CTIME.DAYS (fingreso)) / 365)

OTRA FORMA DE CÁLCULO: Utilizando comandos en la ventana de sintaxis

- (FILE → Data → Syntax). Escribir:

COMPUTE SEGUNDOS = falta - fingreso .
EXECUTE .

COMPUTE DÍAS = CTIME.DAYS(falta) - CTIME.DAYS(fingreso) .
EXECUTE .

COMPUTE DÍAS1 = RND(CTIME.DAYS(falta) - CTIME.DAYS(fingreso)) + 1 .
EXECUTE .

COMPUTE AÑOS = RND((CTIME.DAYS(falta) - CTIME.DAYS(fingreso)) / 365) .
EXECUTE .

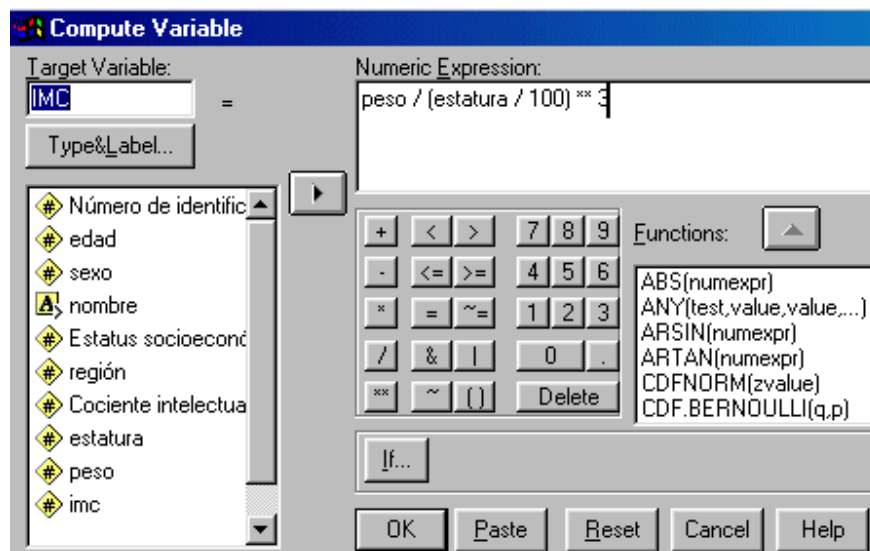
- Guardar como: File → Save as. Escribir nombre del archivo: Fechas
- Clic en Guardar.

PRACTICA N° 05 CREACIÓN O TRANSFORMACIÓN CONDICIONAL DE VARIABLES (IF)

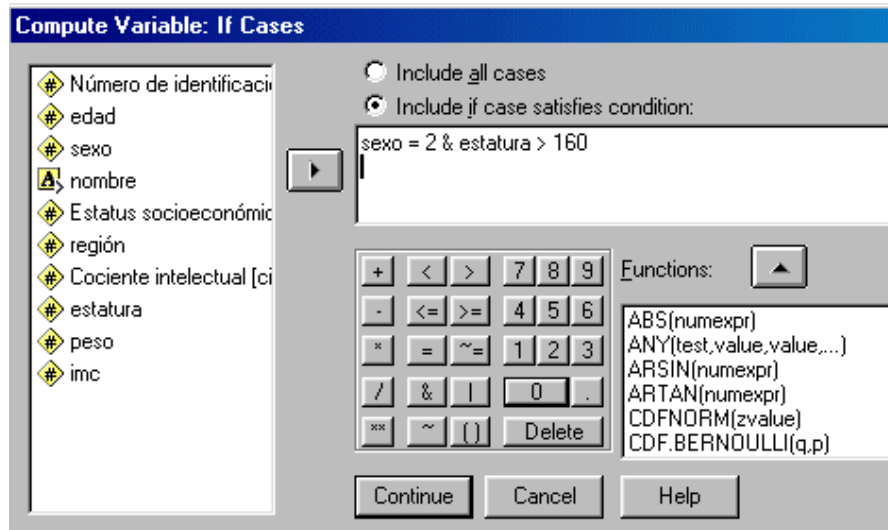
File → Open. Abrir el archivo: MIAPELLIDO.SAV

1.- Se quiere hallar la variable IMC que tiene la fórmula $\text{peso}/(\text{estatura}/100)^2$ pero sólo para las mujeres que midan más de 1.60 m. El proceso es:

- Elegir Transformar → Calcular (Transform—Compute). Clic en Reset.
- Escribir el nombre de la nueva variable: IMC en la celdilla debajo de Variable de destino (Target Variable).
- Escribir la expresión $\text{peso}/(\text{estatura}/100)^2$ en la celdilla debajo de Expresión numérica.
- Peso; /; (); estatura; /; 100; **; 2



- Haga clic en el botón IF (Si)... y aparece el Cuadro “Calcular variable: Si los casos” Compute variables: If Cases).
- En este cuadro marcar el círculo al lado de “Incluir si el caso satisface la condición” (“Include if case satisfies condition”).
- Escribir en la casilla la expresión $\text{sexo} = 2 \ \& \ \text{estatura} > 160$ o ayudándose de los elementos del cuadro de diálogo.



- Clic en Continuar(Continue) y Aceptar (OK).
- Guardar el archivo como: TRANSFORMACIÓN o pulsar las teclas Control + S si se quiere grabar las últimas modificaciones o cálculos realizados.

NOTA: Si la variable IMC ya existe se calculará la expresión para las mujeres de más de 160 cm; para el resto de las mujeres y todos los hombres la variable IMC tendrá el valor que tenía antes de hacer la transformación.

Resultados comparativos:

Resultados del ejercicio No. 3 sobre el cálculo del IMC. Se muestran 10 de los 150 casos.				
Sexo	estatura	peso	IMC	IMC
Mujer	154	57.9	24.41	24.41
Hombre	154	56.7	23.91	23.91
Hombre	156	52.8	21.70	21.70
Mujer	184	104.2	30.78	16.73
Hombre	173	87.5	29.24	29.24
Mujer	170	71.5	24.74	14.55
Hombre	176	93.3	30.12	30.12
Hombre	188	99.2	28.07	28.07
Mujer	169	76.6	26.82	15.87
Mujer	175	87.9	28.70	16.40

Utilizando comandos en la ventana de sintaxis (File → New→ Syntaxis)

Escribir:

```
IF (sexo = 2 & estatura > 160)
IMC = peso / (estatura / 100) ** 3 .
EXECUTE .
```

- Guardar como: File→ Save as. Escribir nombre del archivo: Condiciones
- Clic en Guardar.

PRACTICA N° 06: SELECCIÓN DE CASOS O GRUPOS

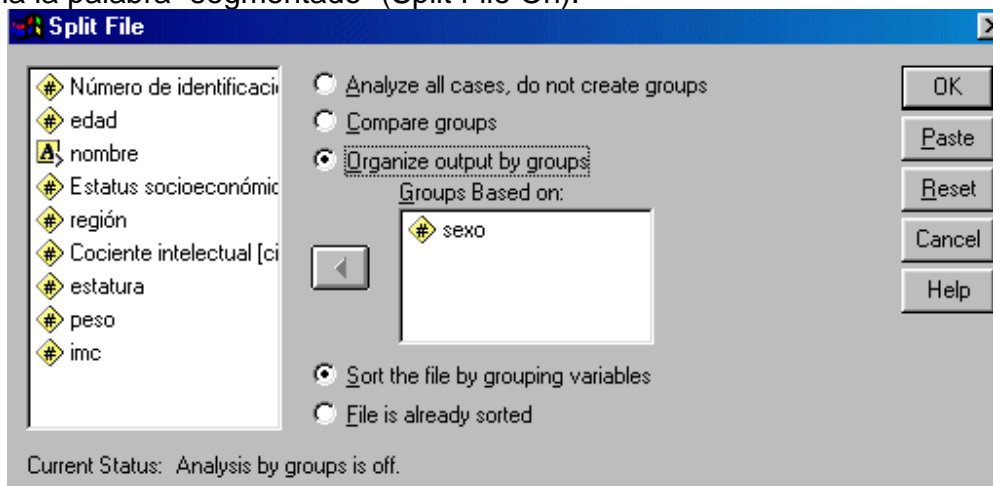
SEGMENTAR ARCHIVOS (SPLIT FILE)	SELECCION DE CASOS
Describe cómo realizar el análisis estadísticos para cada grupo de sujetos definidos por una variables. Datos-Segmentar archivos	Para realizar el análisis de datos con una parte de los sujetos que cumplen una o más condiciones. Datos—Seleccionar casos Se pueden seleccionar los hombres que pesan más de 80 kg. Restringir el análisis a un rango de casos (del 5 al 30). Analizar un porcentaje aleatorio de sujetos. Analizar desde los sujetos del 1 al n.

- Abrir el archivo: **MIAPELIIDO.SAV**

SEGMENTAR ARCHIVOS (SPLIT FILE)

1.- Hallar la correlación de Pearson entre peso y estatura por separado para hombres y mujeres.

- Clic en Datos—Segmentar archivos. (Data—Split File). Entonces se marca el círculo al lado de “Comparar los grupos” (“Compare groups”).
- Transferir la variable sexo y clic en OK. En la barra de estado aparecerá en la última casilla la palabra “segmentado” (Split File On).



Utilizando comandos en la ventana de sintaxis sería:

Sort cases by sexo .

split file

separate by sexo .

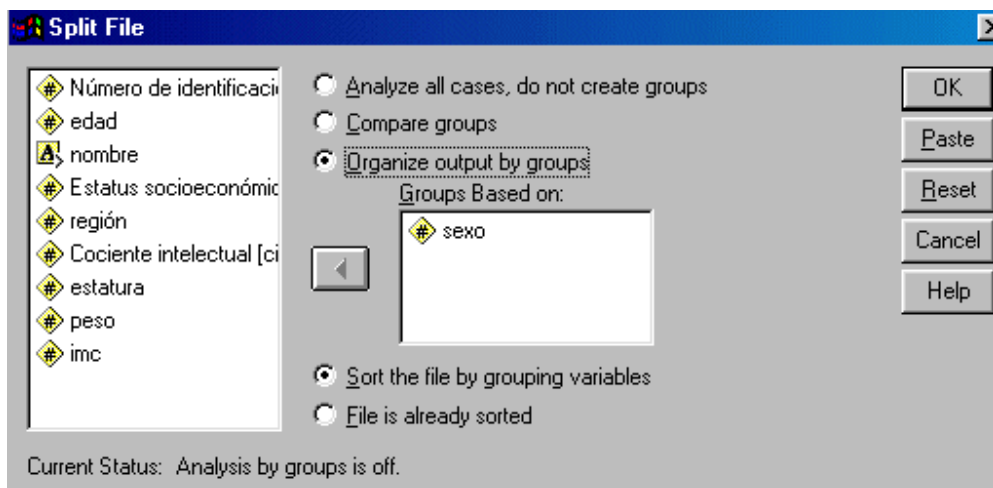
- Luego se elige Analizar—Correlaciones—Bivariadas (Analyze—Correlate—Bivariate) se transfieren las dos variables y clic en OK.

Resultados del análisis:

SEXO			PESO	ESTATURA
Hombre	PESO	Pearson Correlation	1.000	.595
		Sig. (2-tailed)	.	.000
		N	75	75
	ESTATURA	Pearson Correlation	.595	1.000
		Sig. (2-tailed)	.000	.
		N	75	76
Mujer	PESO	Pearson Correlation	1.000	.594
		Sig. (2-tailed)	.	.000
		N	74	74
	ESTATURA	Pearson Correlation	.594	1.000
		Sig. (2-tailed)	.000	.
		N	74	74

** Correlation is significant at the 0.01 level (2-tailed).

- Si se quiere volver al estado original hay que acceder al cuadro de diálogo “segmentar archivo” (Data- Split file) y marcar el círculo al lado de “Analizar todos los casos, no crear grupos” (“Analyze all cases, do not create groups”), y clic en OK. También puede hacer clic en el botón RESET (Restablecer) y en OK.

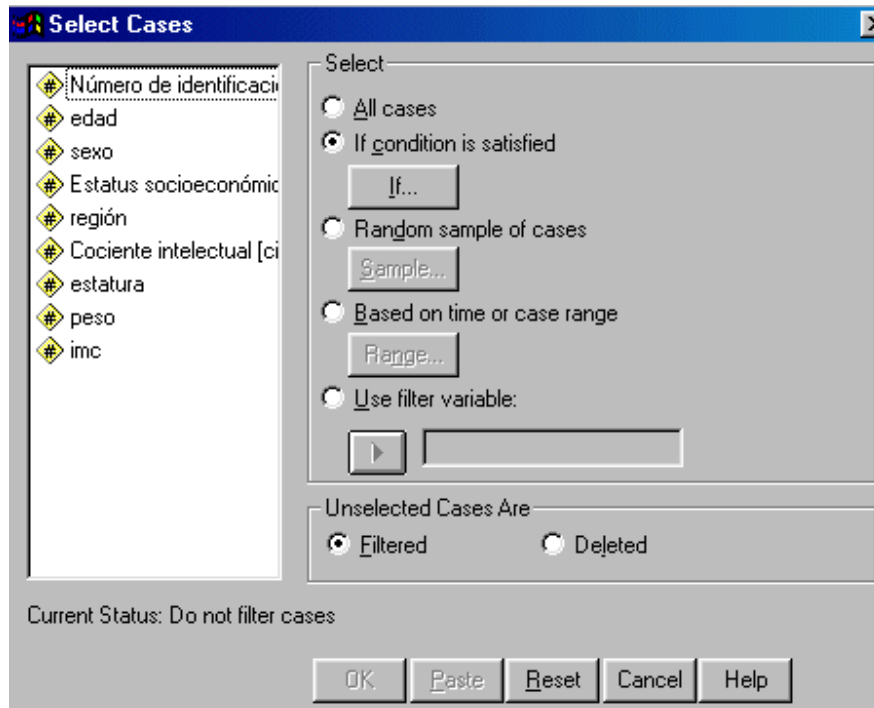


- De la barra de estado desaparecerá el mensaje “Segmentado” (Split File On). Y proceder a realizar otras acciones, caso contrario solamente analizará los datos filtrados.

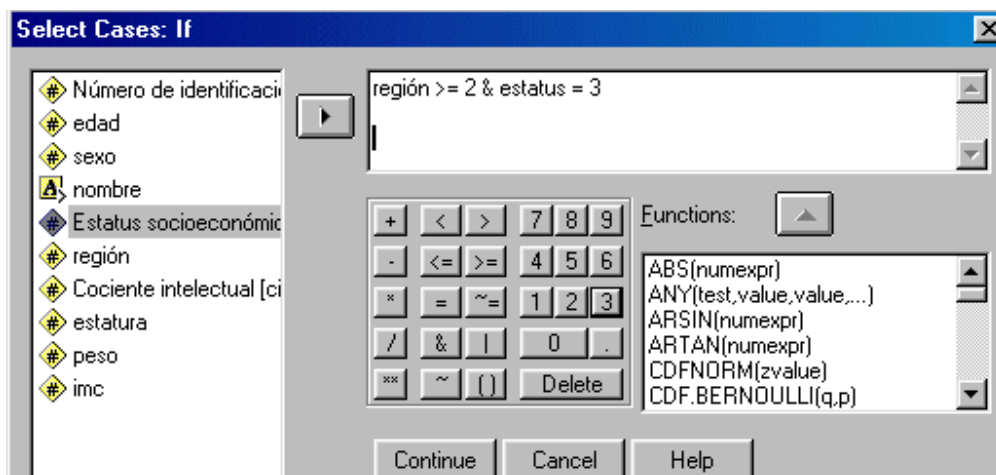
SELECCIÓN DE CASOS

2.- Analizar sólo los casos de estatus bajo (valor 3) que no sean del Norte (valores 2,3 y 4 de región).

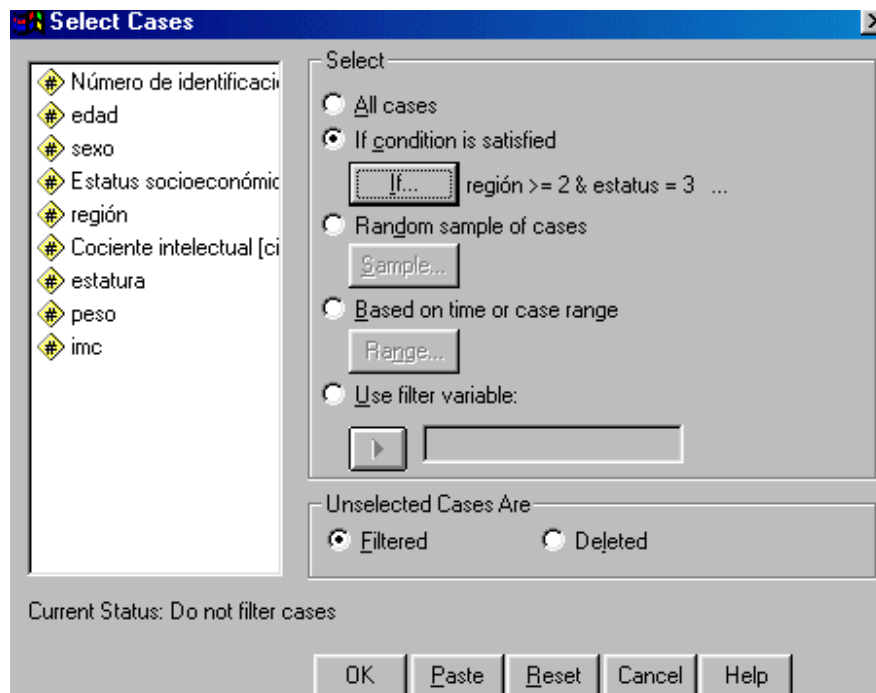
- Clic en Datos – Seleccionar casos (Data—Select cases)
- Marcar el círculo al lado de “Si se satisface la condición” (“If condition is satisfied”).



- Pulsar el botón “Si...” (“If...”). Se accede al cuadro de diálogo “Seleccionar casos: si”. (Select Cases: IF)
- Escribir o componer la expresión: `región >= 2 & estatus = 3`. Transferir la variables y utilizar los operadores necesarios.



- Clic en Continue

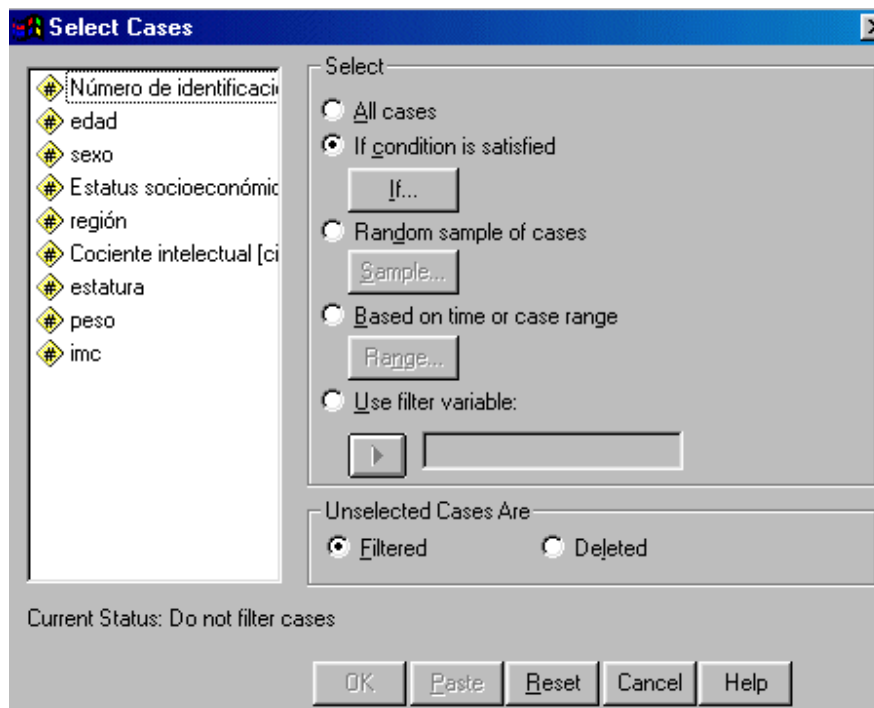


- Clic en OK. En la barra de estado aparecerá el mensaje “filtrado” (Filter On).
- A partir de este momento sólo se procesarán los casos que cumplan la condición.
- Observe en el editor de datos cómo se marcan los casos no seleccionados.
- En la ventana editor de datos, en la última columna se crea la variable filter_\$ que tiene valor 1 si se cumple la condición (región sea mayor que 2 y el estatus 3), y tiene valor 0 en el resto de los casos.
- Clic en el icono Value labels (Etiquetas de los valores) de la barra de herramientas.
- Observará que 1 equivale a caso seleccionado (Select) y los casos con valores perdidos o que no cumplen la condición (Valor 2) equivale a Not Select (No seleccionado).

16	16	29	Mujer	Alana	Alto	Norte	65	156	53.9	22.15	Not Selecte
17	17	29	Homb	Jose	Bajo	Sur	147	172	78.6	26.57	Selected
18	18	16	Mujer	Albertina	Alto	Norte	113	175	85.7	15.99	Not Selecte
19	19	31	Homb	Armando	Alto	Oeste	99	168	75.4	26.71	Not Selecte
20	20	18	Homb	Constante	Medio	Sur	117	159	73.0	28.88	Not Selecte
21	21	21	Homb	Toribio	Alto	Este	105	178	87.6	27.65	Not Selecte
22	22	29	Homb	Borja	Alto	Este	92	174	74.2	24.51	Not Selecte
23	23	23	Mujer	Alejandra	Bajo	Sur	78	178	100.9	17.89	Selected
24	24	28	Mujer	Alicia	Alto	Sur	97	163	77.1	17.80	Not Selecte
25	25	19	Homb	Manuel	Bajo	Sur	77	180	74.6	23.02	Selected
26	26	31	Homb	Carlos	Bajo	Sur	81	174	70.2	23.19	Selected

Si se quiere volver a analizar todos los casos: Data—Select cases:

- Marcar el círculo al lado de “Todos los casos” (“All cases).
- Luego clic en OK (Aceptar).
- Observe que debe desaparecer el mensaje de datos filtrados (Filter On) de la barra de estado de la pantalla.



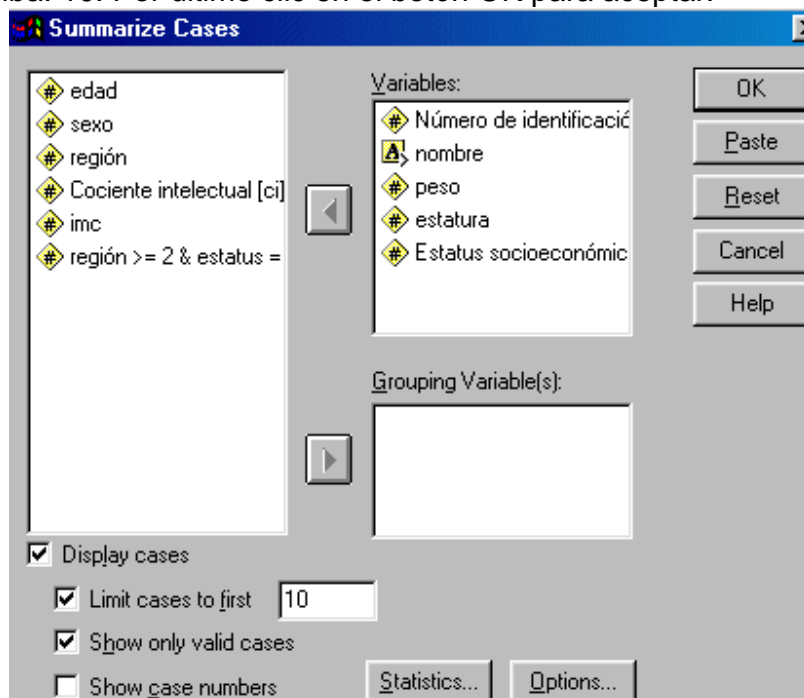
- Guardar como: File→ Save as. Nombre del archivo: Casos-Grupos.
- Clic en Guardar.

PRACTICA N° 07: DESCRIPCIÓN Y MANIPULACIÓN BÁSICA

▪ Listado más simple de los casos: El listado de los valores en las variables

1.- Producir un listado de las variables: iden, nombre, peso, estatura y estatus para los 10 primeros casos de la muestra de datos del archivo de la práctica N° 04.

- Clic en el Menú: Analyze ---- Reports --- Case summaries (Analizar –Informes – Resúmenes de casos.)
- Se seleccionan y transfieren a la sección “Variables” las variables: iden (número de identificación), nombre, peso, estatura y estatus (estatus socioeconómico).
- En la casillero, al lado de Limit cases to list (limitar los casos a los primeros) escriba: 10. Por último clic en el botón OK para aceptar.



Resultados: (Tabla editada)

Case Summaries^a

		Número de identificación del sujeto	NOMBRE	PESO	ESTATUR A	Estatus socioeconómico
1		1	Abigail	57.9	154	Alto
2		2	Carlos	56.7	154	Alto
3		3	Alberto	52.8	156	Alto
4		4	Adela	104.2	184	Medio
5		5	Luis	87.5	173	Alto
6		6	Adelina	71.5	170	Medio
7		7	Antonio	93.3	176	Alto
8		8	Juan	99.2	188	Alto
9		9	Adriana	76.6	169	Medio
10		10	Agata	87.9	175	Medio
Total	N	10	10	10	10	10

^a Limited to first 10 cases.

- **El listado de casos se puede fraccionar por los valores de variables (categóricas usualmente)**

2.- Listar por separado a los hombres y a las mujeres.

- Para ello en el cuadro de diálogo, hay que seleccionar la variable sexo y transferirla a la sección Grouping variables (Variables de agrupación).
- En el caso de fraccionamiento del listado puede ser relevante describir los subgrupos (mujeres, hombres) con estadísticos descriptivos: media (mean), desviación típica (Standard Deviation), etc) haciendo clic en el botón Statistics (Estadísticos) y seleccionando dichos estadísticos. Continúe.
- Clic en Options (Opciones) cambie el título: Case summaries por Resúmenes de casos. Finalmente, si todo está correcto clic en OK.

▪ **ORDENAR CASOS (SORT)**

Se pueden ordenar los datos en orden ascendente (por defecto) o descendente en función de una o más variables.

3.- Ordenar los datos por el ci(cociente intelectual).

- Clic en el menú: Data – Sort Cases (Datos – Ordenar casos).
- Seleccionar la variable ci (cociente intelectual) y transferirla a la sección Sort by (Ordenar por). Clic en la casilla de ascendente o descendente según sea necesario. Finalmente clic en OK para aceptar.
- Observe los datos ordenados en la pantalla para comprobarlo.
- Nota: Se puede ordenar por más de una variable, situada en el cuadro de diálogo, se seleccionan y transfieren a la sección Sort by (Ordenar por). Cada variable de ordenación puede ir en orden ascendente o descendente independientemente. Si después de realizar una ordenación se guardan los datos, se mantiene la ordenación por los criterios elegidos.

IMPORTANTE:

▪ **DESCRIPCIÓN RÁPIDA DE VARIABLES**

Sólo disponible con lenguaje de comandos. Empleando el comando DESCRIPTIVES se puede conseguir una descripción condensada (media, desviación típica, mínimo, máximo y número de sujetos o cuestionarios aplicados) y rápida de las variables. El comando es: DESCRIPTIVES ALL.

- Clic en el menú: File – New – Syntax
- Escriba:

DESCRIPTIVES ALL.

- Clic en el menú: Run – All **(¿Algún comentario sobre la Tabla?)**

PRACTICA N° 08: DESCRIPCIÓN DE VARIABLES NOMINALES

Las variables nominales se estudian mediante la distribución de frecuencias, la gráfica de barras o la gráfica de sectores y un estadístico descriptivo: la moda. Variables nominales en el estudio que venimos tomando como ejemplo son el sexo y la región.

1.- DESCRIBIR LA VARIABLE REGIÓN.

- Analyze—Descriptive Statistics –Frecuencias (Analizar – Estadísticos descriptivos – Frecuencias)
- Seleccione región y transferirlo a la sección Variables.
- Clic en el botón Chart (Gráficos) de este cuadro, marcar la opción Bar Charts (Gráficos de barras) y hacer clic en botón Continue (Continuar). Alternativamente se puede marcar la opción Pie Chart (Gráfico de sectores).
- Para hallar la moda. Clic en el botón statistics (Estadísticos) del cuadro de diálogo, marcar el estadístico Mode (Moda) en la sección de Central tendency (tendencia central) y se pulsa el botón Continue (Continuar).
- Por último clic en el botón OK para aceptar.

Nota: Puede realizar la descripción de las dos variables región y sexo en forma simultánea si lo prefiere.

Comentarios:

- En la Tabla pivote se puede ver que hay 43 personas del Norte que suponen un 28,7 % de la muestra (43 de 150). La diferencia entre porcentaje válido y porcentaje es que en porcentaje válido no se tiene en cuenta los casos perdidos (no se contabilizan los casos perdidos), en esta variable existe un caso perdido (valor 9) y por lo tanto hay 149 sujetos válidos; lo que supone un 28,9% (43 de 149). El porcentaje acumulado va sumando la frecuencia de cada valor con las frecuencias de los valores menores a él; aunque con variables nominales no tiene mucho sentido.
- En la gráfica de barras se observa que la barra de menor altura corresponde a la categoría SUR con menor número de casos. Y que la barra de mayor altura es la del Norte, que es la que incluye mayor número de casos.
- Si desea cambiar el aspecto de la gráfica o añadir nuevos elementos hay que editarla. Para editar una gráfica pulse clic dos veces sobre ella. Por ejemplo, puede añadir el número de casos a cada barra pulsando el icono (Bar Label style) en forma de dos pequeñas barras de color amarillo, ubicada en dicha barra de herramientas. Clic en estándar y en Apply all.
- Trate de cambiar el color de las barras a AZUL y el color blanco para indicar el número de casos. El título del gráfico debe decir: Distribución de casos según Región.

PRACTICA N° 09: DESCRIPCIÓN DE VARIABLES ORDINALES

Las variables ordinales se estudian mediante la distribución de frecuencias, la gráfica de barras (o la gráfica de sectores si hay pocos valores distintos) y los estadísticos descriptivos: mediana, amplitud (Rango), máximo y mínimo. Cuando existen una gran cantidad de valores distintos (por ejemplo, más de 10) se analizan como variables de intervalo. También en el caso de depuración de items de pruebas y encuestas se utilizan los índices de asimetría y apuntamiento. Variables ordinal en este estudio es el estatus.

1.- Describir la variable ordinal estatus socioeconómico

- Para describir la mediana, rango, mínimo y máximo: Analizar –Estadísticos Descriptivos – Frecuencias (Analyze – Descriptive Statistics – Frequencies).
- Transferir la variable estatus socioeconómico a la sección Variables.
- Luego clic en statistics. Marcar solamente la mediana (median) para la tendencia central (Central tendency) y, en dispersión, marcar: rango (range), mínimo (minimum) y máximo (maximum). Clic en Continue.
- Para graficar clic en Chart y activar Pie Chart y en continue. Clic en OK.
- Editar el gráfico: Clic en el elemento título debe decir: Estatus socioeconómico gráfico de sectores. Para que se tenga los valores y porcentajes en cada tipo de región. Clic en el nombre de cualquier región y activar las casillas Value (Valor) y Percents (Porcentajes).

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Alto	76	50.7	51.4	51.4
	Medio	41	27.3	27.7	79.1
	Bajo	31	20.7	20.9	100.0
	Total	148	98.7	100.0	
Missing	9	2	1.3		
Total		150	100.0		

Comentario:

De estatus alto hay 76 casos. Son el 50,7% de todos los casos (76 / 150). Y suponen un porcentaje válido del 51,4% (76 /148).

PRACTICA N° 10: DESCRIPCIÓN DE VARIABLES DE INTERVALO

Las variables de intervalo se estudian mediante el histograma, y los estadísticos descriptivos: media (mean), desviación típica (Standard Deviation), range (range), máximo (maximum) y mínimo (minimum), y los coeficientes de asimetría y apuntamiento. No conviene obtener la distribución de frecuencias porque al tener usualmente un gran número de valores distintos la variable de intervalo, resulta demasiado extensa la distribución. Para ello elimine la marca de la opción “Mostrar tabla de frecuencias” (“Display frequency tables”) en el cuadro de diálogo de Frecuencias (Frequencies). Variables de intervalo o razón en este estudio son: la edad, el peso, la estatura y el ci (cociente intelectual).

1.- Describir las variables edad, peso, estatura y ci

- Elegir Analizar-Estadísticos Descriptivos–Frecuencias (Analyze-Descriptives Statistics-Frequencies). Seleccionar y transferir las variables. Desmarcar la casilla de “Mostrar la tabla de frecuencias” (Display frequency tables).
- Pulsar el botón gráficos (Charts) y marcar Histogramas y la opción “Con curva normal” para que sobreimprima la curva de la distribución normal. Pulsar el botón Continue.
- Pulsar el botón Estadísticos (Statistics) y marcar la media, la desviación standard, el rango, el mínimo y el máximo. La simetría y la curtosis o apuntamiento. Automáticamente se hallan también los errores típicos de la simetría (S.E. Skew) y del apuntamiento (S.E. Kurt). Para obtener el error típico de la media marcar “E.T. media”. Clic en continue y en OK.

Statistics

		EDAD	PESO	ESTATURA	Cociente intelectual
N	Valid	148	149	150	149
	Missing	2	1	0	1
Mean		24.25	76.774	168.85	101.658
Std. Error of Mean		.46	1.004	1.12	1.857
Std. Deviation		5.60	12.255	13.73	22.672
Skewness		.169	-.030	2.505	.100
Std. Error of Skewness		.199	.199	.198	.199
Kurtosis		-1.076	-.658	18.934	-.410
Std. Error of Kurtosis		.396	.395	.394	.395
Range		20	53.8	139	107.0
Minimum		15	50.4	131	52.0
Maximum		35	104.2	270	159.0

Comentarios:

La media del ci es 101,6, la desviación típica 22,6, el coeficiente de apuntamiento (curtosis) vale –0,41 (menos apuntada que la distribución normal), el coeficiente de asimetría es 0,1 (ligeramente asimétrica positiva), el valor mínimo es 52 y el máximo 159, el rango de 107. Hay 149 casos válidos y 1 caso perdido. De estos estadísticos el que puede, de entrada, ofrecer menos información es la desviación típica (medida de dispersión). Sin embargo, si se recuerda una de las propiedades de la distribución normal: que el 95% se encuentra entre dos desviaciones típicas a cada lado de la media, se obtiene una idea mejor de la dispersión del grupo: el 95% de los casos estará entre 101,6 +- 45,3. La desviación típica da una información más intuitiva cuando se comparan las dispersiones de varios grupos.

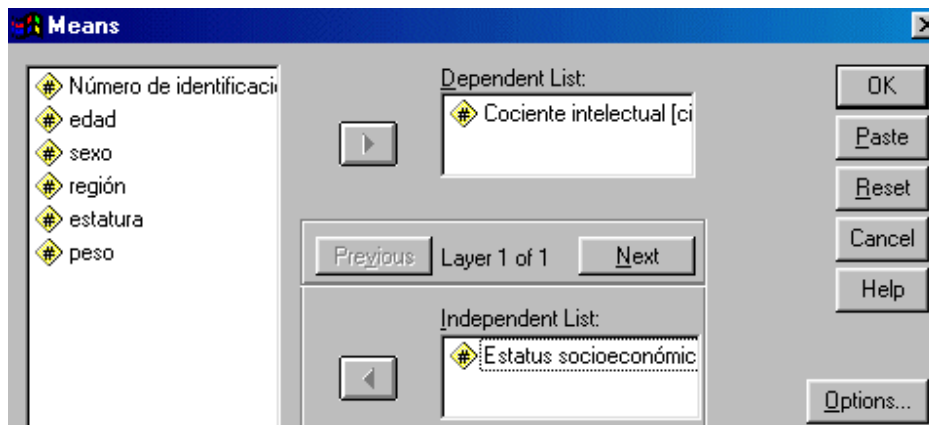
Para construir el Histograma, el programa agrupa las puntuaciones en intervalo, el punto medio del intervalo es su valor representativo. La anchura del intervalo, en esta gráfica, es de 5 puntos y el valor debajo de cada rectángulo del histograma es el punto medio del intervalo dado (sólo se imprimen los puntos medios de intervalos alternos).

PRACTICA N° 11: DESCRIPCIÓN DE VARIABLES POR GRUPOS (MEANS)

Esta acción se emplea para describir variables cuantitativas en función de una o más variables cualitativas (categóricas).

1.- Describir las variable ci (cociente intelectual), en función del estatus.

- Analizar—Comparar medias—Medias (Analyze-Compare means-Means).
- En el cuadro de diálogo Means hay que transferir la variable ci (cociente intelectual) a “Dependientes” (Dependent List) y después transfiera la variable estatus socioeconómico a la ventana Independientes (Independent List) y se pulsa el botón OK.



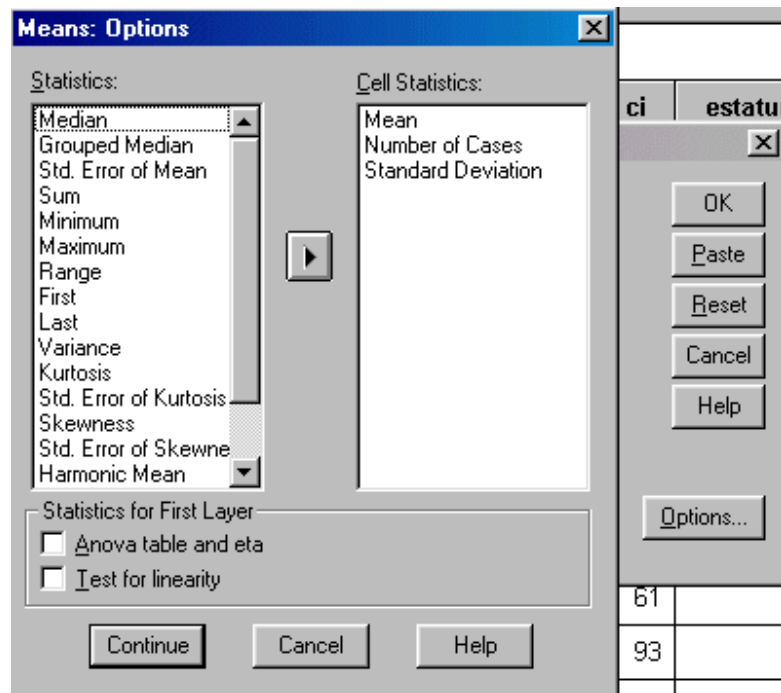
- Resultados:

Cociente intelectual			
Estatus socioeconómico	Mean	N	Std. Deviation
Alto	101.645	76	20.433
Medio	100.500	40	23.822
Bajo	101.774	31	26.823
Total	101.361	147	22.675

Comentario:

Se ofrece para cada grupo de estatus la media, desviación típica y el número de casos. Se ofrecen estos estadísticos, también para la muestra total. El ci medio de los de estatus alto es 101,64. Las desviaciones típicas son muy similares.

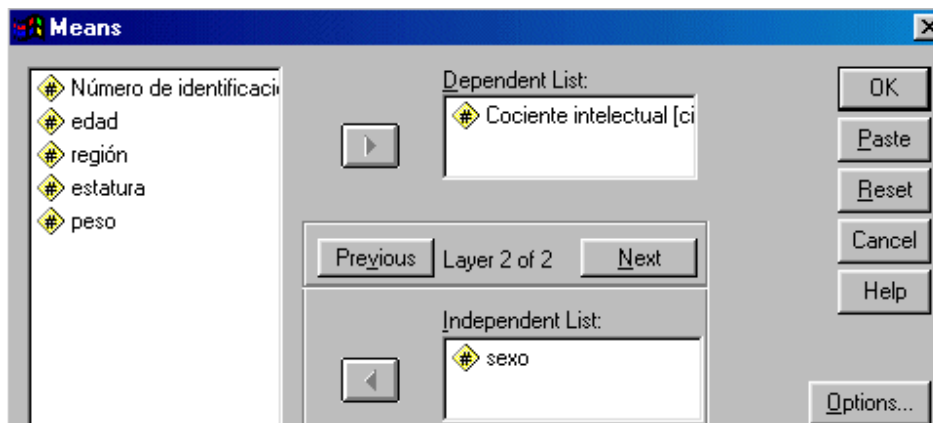
- Nota: con el botón opciones (Options) se pueden hallar otros estadísticos además de la media, desviación típica y número de sujetos, que son los que se hallan por defecto, y, también realizar un análisis de varianza y la prueba de linealidad.



DESCRIPCIÓN JERARQUIZADA: Se puede describir la variable cuantitativa en función de varias variables de clasificación de una manera jerarquizada.

2.- Diferenciar por grupos de sexo dentro de los grupos de estatus.

- Para ello, en el cuadro de diálogo anterior, hay que pulsar el botón central siguiente: Layer 1 of 1 (Next), transferir la variable sexo a “Independientes” (Independent list) de Layer 2 of 2. Pulsar el botón OK.



Aparecen los resultados mostrados en la siguiente Tabla:

Cociente intelectual

Estatus socioeconómico	SEXO	Mean	N	Std. Deviation
Alto	Hombre	100.098	41	18.116
	Mujer	103.457	35	22.991
	Total	101.645	76	20.433
Medio	Hombre	99.235	17	24.692
	Mujer	101.435	23	23.674
	Total	100.500	40	23.822
Bajo	Hombre	99.063	16	29.188
	Mujer	104.667	15	24.729
	Total	101.774	31	26.823
Total	Hombre	99.676	74	22.108
	Mujer	103.068	73	23.264
	Total	101.361	147	22.675

Las mujeres de estatus alto tienen un ci medio de 103,45 y una desviación típica de 22,9.

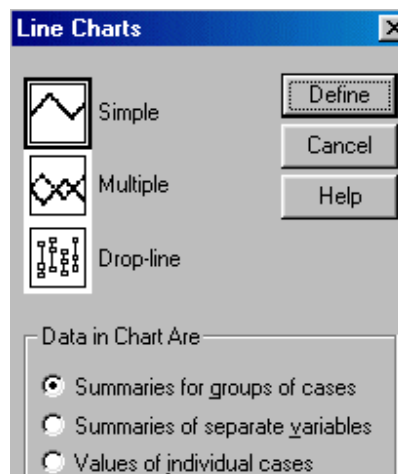
- **DESCRIPCION DE MÁS DE UNA VARIABLE:** Si se desean describir por grupos varias variables cuantitativas éstas se seleccionan y se transfieren a la ventana “Dependientes” (Dependent list).

GRAFICAS

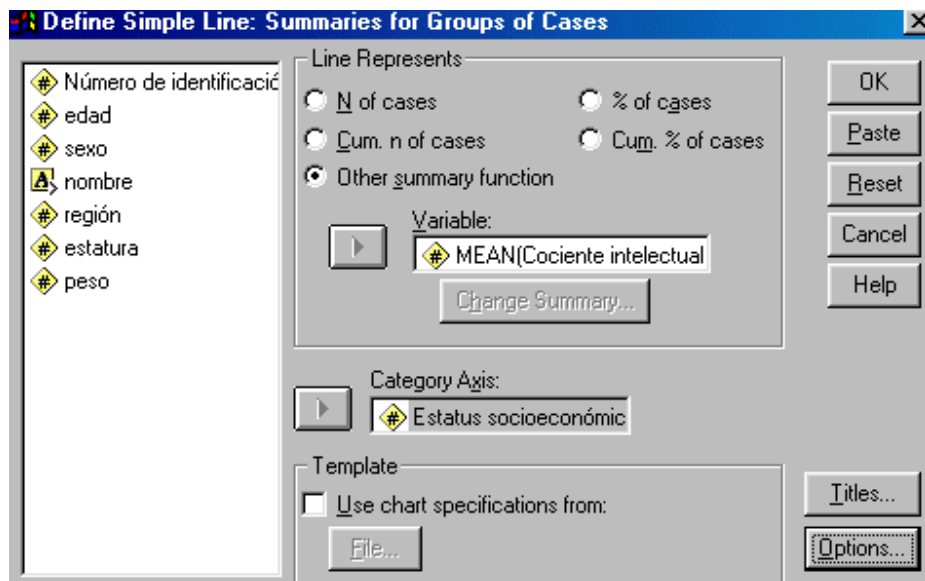
SIMPLES: Otra manera de describir es realizar gráficas de los estadísticos de los subgrupos.

3.- Hallar una gráfica con las medias en ci (cociente intelectual) de los tres grupos de estatus).

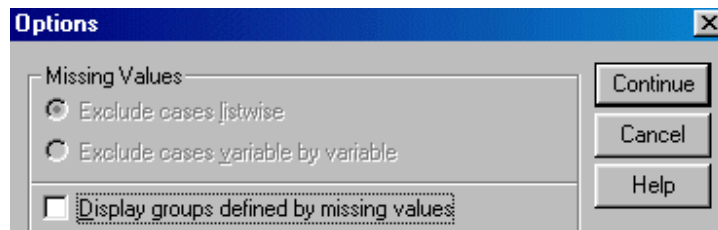
- Gráficos-Líneas (Graphs – Line). Tipo: Simple. Summaries for groups cases. Clic en Define.



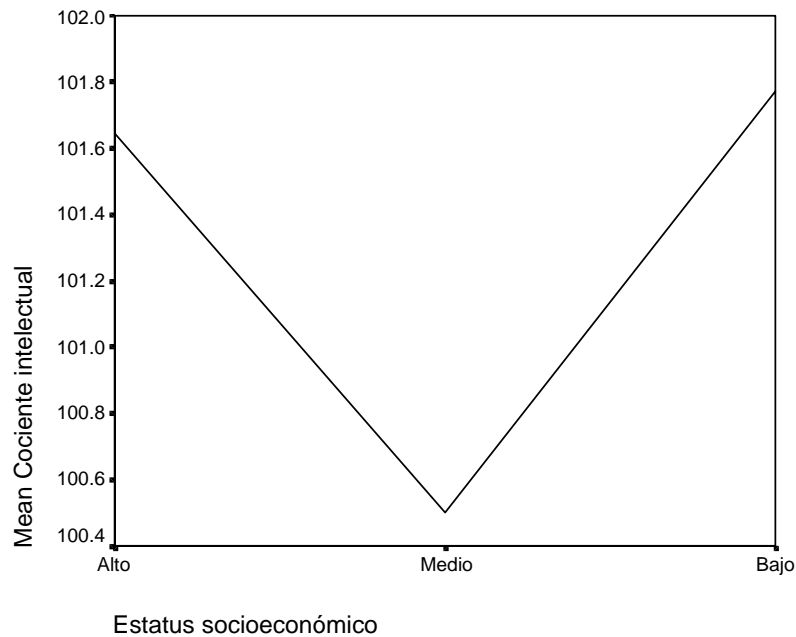
- En la ventana o cuadro de diálogo siguiente: Clic en la casilla Other summary function.
- Transferir ci (cociente intelectual) al recuadro de la opción “Variable” y Estatus socioeconómico al recuadro “Category Axis” (Eje de categorías). Pulse en OK y observe que tipo de gráfico obtiene.



- Repita los pasos anteriores. Pero antes de pulsar el botón OK. Haga clic en el botón Opciones (Options) y desactive la casilla de "display groups defined by missing values" (Mostrar grupos definidos por valores perdidos o datos en blanco) para no incluirlos en el análisis de los datos.



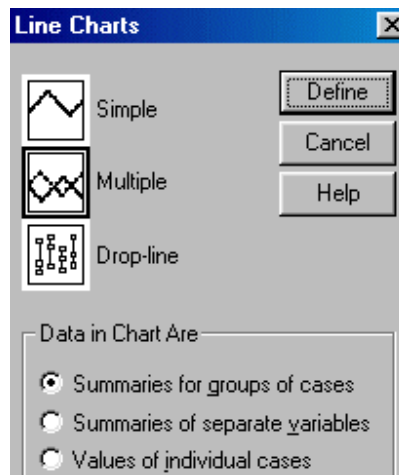
- Clic en Continue y en OK. Obtendrá el gráfico:



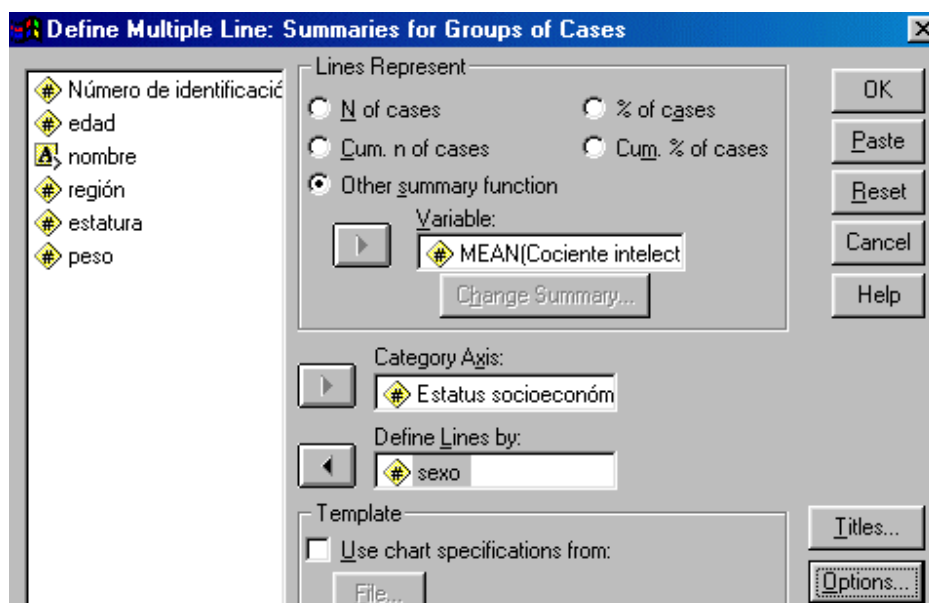
GRAFICAS DE DOBLE CLASIFICACION (MÚLTIPLES)

4.- Obtener una gráfica de líneas (múltiple) de las medias de ci por grupos de estatus y separadas por sexo.

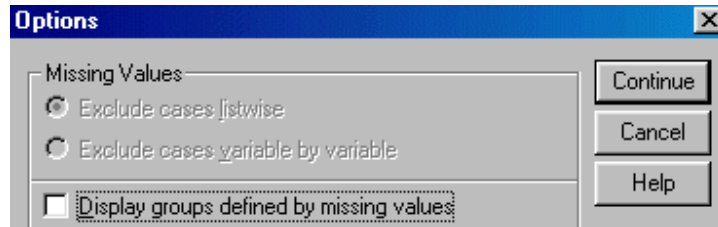
- Gráficos-Líneas (Graphs – Line). Tipo de Line Charts: Multiple. Debe estar seleccionado summaries for groups cases. Clic en Define.



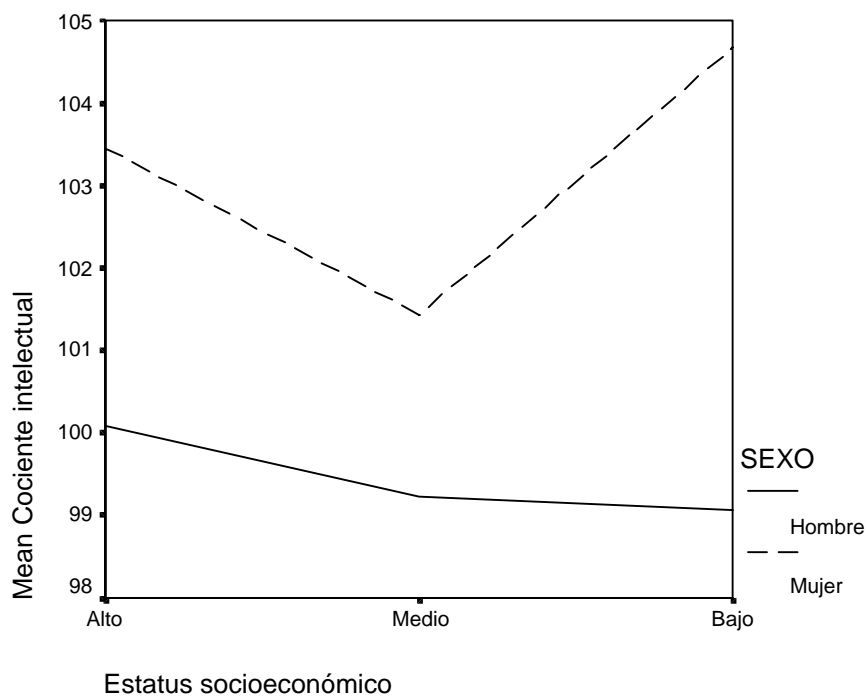
- Luego. Clic en la casilla Other summary function.
- Transferir ci (cociente intelectual) al recuadro: Variable y Estatus socioeconómico al recuadro: Category Axis (eje de categoría). Luego la variable sexo hacia: Defined Lines by (Definir líneas por).



- OJO: No olvidar hacer Clic en Opciones (Options) y desactivar la casilla de “display groups defined by missing values” para no incluir los datos definidos como valor perdidos o datos en blanco. Clic en Continue y en OK.



- Editar el gráfico: doble clic sobre el gráfico.
- Clic sobre cada una de las líneas y cambiarlas a color negro.
- Para el caso de la variable sexo (Mujeres): Modificar el estilo de línea (Line style).

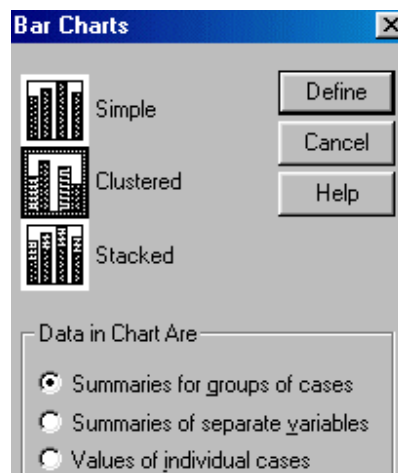


GRÁFICAS DE DOBLE CLASIFICACION CON BARRAS

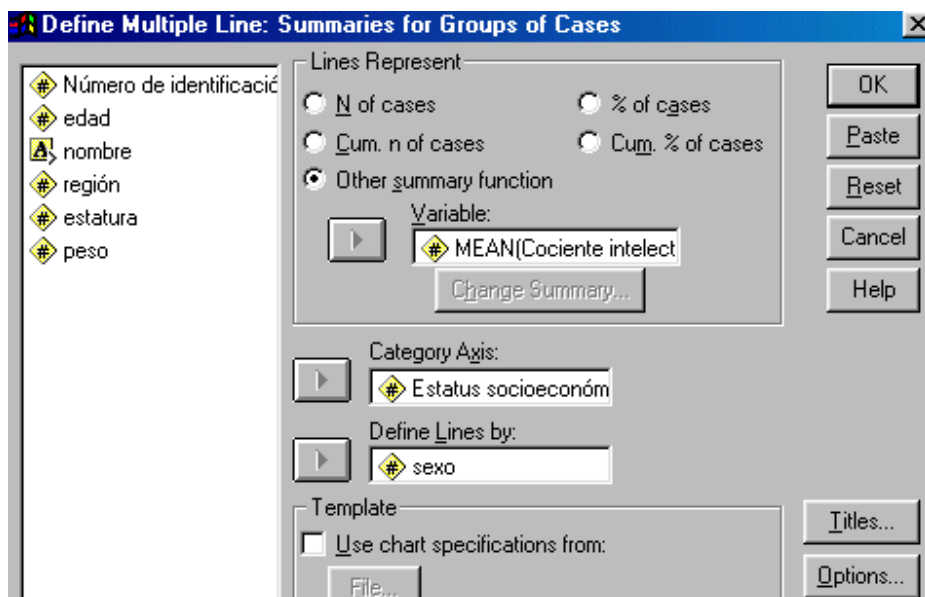
Tanto en el caso de una variable de clasificación como en el de doble clasificación se pueden realizar gráficas con barras en vez de líneas.

5.- Obtener la gráfica de barras de las medias de ci por cada estatus y separado por sexos. Editar el gráfico.

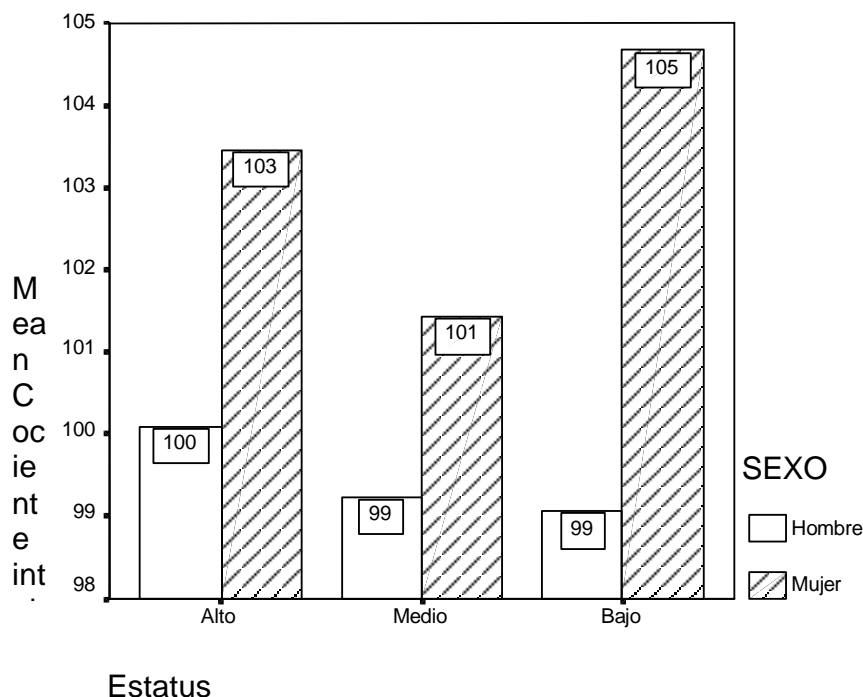
- Gráficos-Líneas (Graphs – Bar). Tipo de Bar Charts: Clustered (Agrupadas). Debe estar seleccionado summaries for groups cases. Clic en Define.



- Luego. Clic en la casilla Other summary function.
- Transferir ci (cociente intelectual) al recuadro: Variable y Estatus socioeconómico al eje de categorías (Category Axis). Luego la variable sexo a Defined Clustered by (Definir agrupación por).



- Clic en Opciones (Options) y desactive la casilla de “display groups defined by missing values” para no incluir los datos definidos como valor perdidos. Clic en Continue y en OK.
- Editar el gráfico: Color de relleno: blanco. Con framed (marcos con los valores)
- Para el color: Seleccionar la barra correspondiente. Clic en icono:Color. Fill: Seleccionar el color blanco. Apply (Aplicar). Proceder con la barra que corresponda. Para finalizar haga Clic en Close (cerrar ventanita).
- Para el framed (Marco con valores). Clic en el icono: Bar Label Style. Seleccionar: framed. Clic en Apply All. Clic en close (cerrar ventanita).
- Para relleno de cada barra: Seleccionar la barra correspondiente. Clic en el icono Fill pattern (patrón de relleno). Seleccionar el relleno correspondiente. Clic en Apply. Proceder igual con la otra barra. Al concluir, clic en close (cerrar ventanita)
- Cerrar la ventana de edición de gráfico.



PRACTICA N° 12: RELACIÓN ENTRE VARIABLES

Los índices estadísticos adecuados para el estudio de la relación entre variables dependen del nivel de medida de las variables. Estos índices indican la magnitud de la relación. Existen pruebas estadísticas asociadas para comprobar si la relación es significativa.

Con variables nominales se emplean, por ejemplo los estadísticos chi cuadrado (χ^2) o el coeficiente de contingencia (Crosstabs).

Con variables de intervalo se emplea el coeficiente de correlación de Pearson.

Con variables ordinales se puede emplear el coeficiente ro de Spearman o la tau de Kendall.

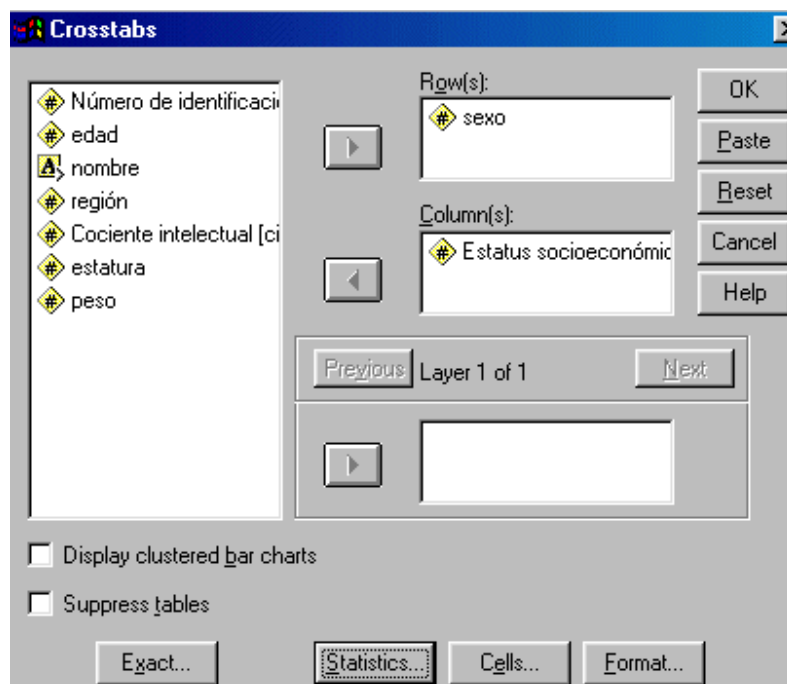
▪ **TABLAS DE CONTINGENCIA: RELACION ENTRE VARIABLES NOMINALES (CROOSTABS)**

Se emplea para hallar simultáneamente distribuciones de frecuencias de dos variables y para hallar índices estadísticos que midan la fuerza de la asociación entre las variables.

En el estudio de ejemplo existen dos variables nominales: sexo y región y una variable ordinal: estatus socioeconómico, que para todos los efectos se va a emplear en los ejemplos como nominal.

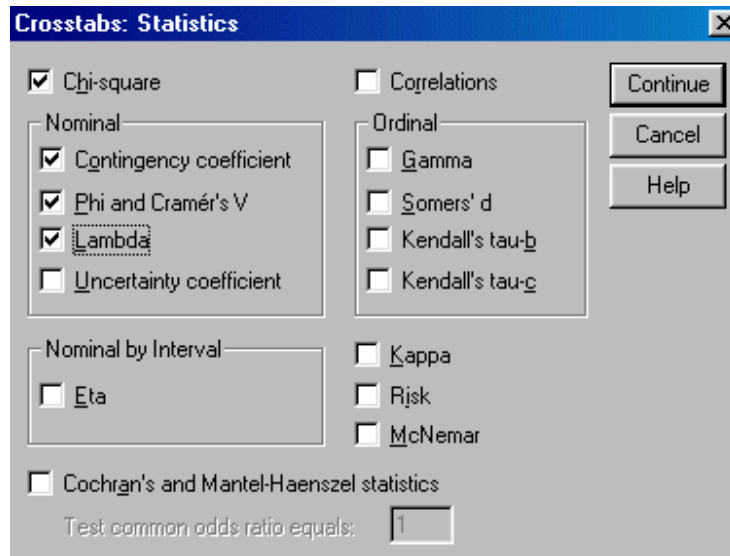
1.- Estudiar la relación entre las variables sexo y estatus.

- Para hallar la distribución de frecuencias cruzadas (tabla de contingencia) del sexo y el estatus elegir Analizar-Estadísticos Descriptivos-Tablas de contingencia (Analyze-Descriptive Statistics-Crosstabs), y en el cuadro de diálogo:
 - Seleccionar la variable sexo y transferirla a la sección “filas” (Rows)
 - Seleccionar la variable estatus y transferirla a “Columnas” (Columns)



- Pulsar el botón “Estadísticos” (Statistics)

- Marcar los estadísticos Chi cuadrado (Chi-square), coeficiente de contingencia (Contingency coefficient), Phi y V de Cramer (Phi and Cramer's V) y Lambda (Lambda). Pulsar el botón continuar (Continue) y Aceptar (OK).



Resultados:

Tabla de contingencia SEXO * Estatus socioeconómico

		Estatus socioeconómico			Total
		Alto	Medio	Bajo	
SEXO	Hombre	41	17	16	74
	Mujer	35	24	15	74
Total		76	41	31	148

En la tabla de Chi-cuadrado la pregunta que se plantea es: ¿existe relación entre el estatus y el sexo?. El contraste de chi cuadrado de Pearson comprueba si existe independencia o no entre ambas variables. $X^2(2)=1,701$ $p \leq 0,427$ indica que no existe relación entre ambas variables (probabilidad mayor que 0,05, es decir, no hay asociación sistemática de determinados valores de estatus con determinados valores de sexo).

Pruebas de Chi-cuadrado

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1,701	2	,427
Likelihood Ratio	1,707	2	,426
Linear-by-Linear Association	,266	1	,606
N of Valid Cases	148		

- a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15.50.
- a. 0 Casillas (.0%) tienen una frecuencia esperada inferior a 5.

La frecuencia mínima esperada es 15,50.

Medidas Simétricas (Symmetric Measures)

		Value	Approx. Sig.
Nominal by Nominal	Phi	,107	,427
	Cramer's V	,107	,427
	Contingency Coefficient	,107	,427
N of Valid Cases		148	

- a Not assuming the null hypothesis. (No asumiendo la hipótesis nula)
- b Using the asymptotic standard error assuming the null hypothesis.
(Empleando el error asintótico basado en la hipótesis nula)

Tabla editada para mostrar sólo los estadísticos Lambda:

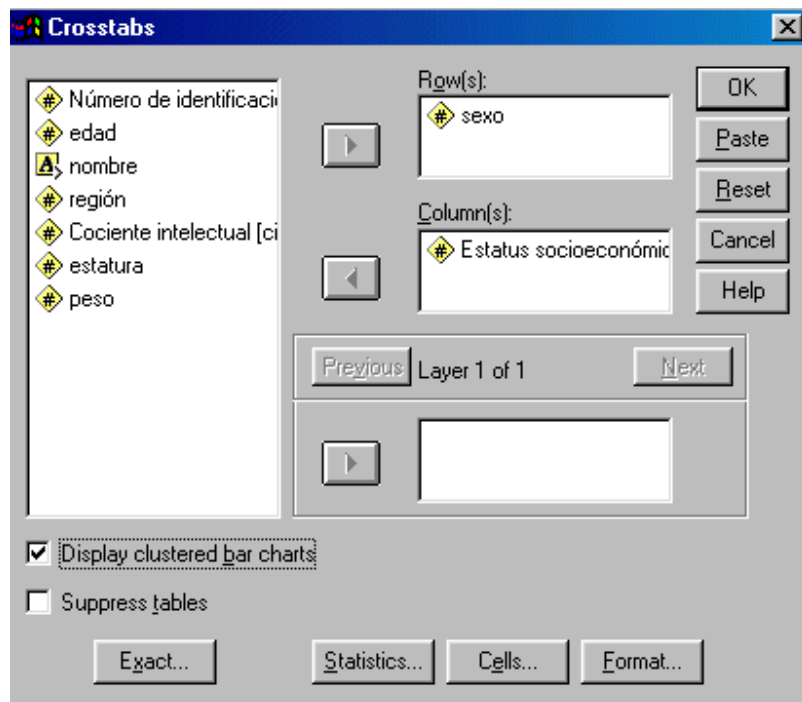
Directional Measures (Medidas direccionales)
Nominal por nominal Lambda

	Value	Asymp. Std. Error	Approx. T	Approx. Sig.
Symmetric	.048	.043	1.098	.272
SEXO Dependent	.095	.082	1.098	.272
Estatus socioeconómico Dependent	.000	.000	.	.

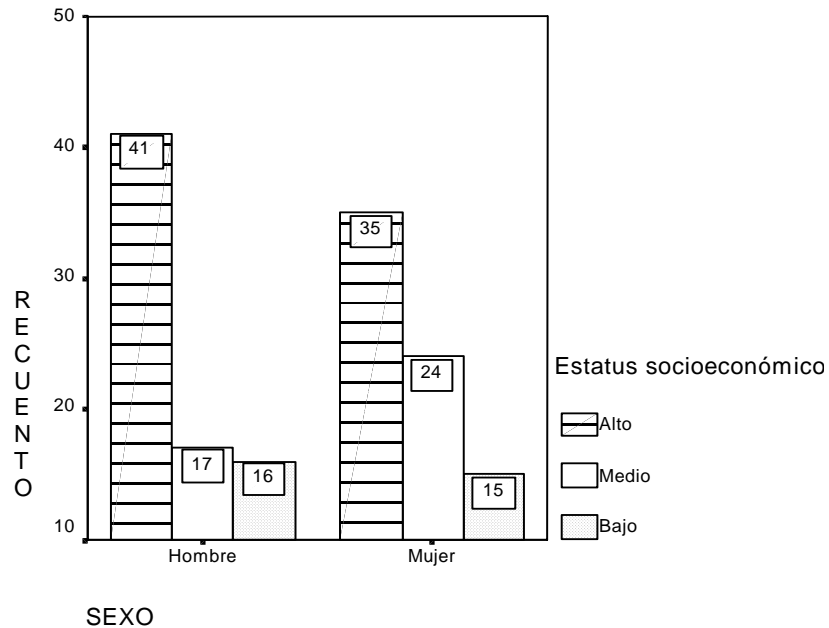
- a Not assuming the null hypothesis. (No asumiendo la hipótesis nula)
- b Using the asymptotic standard error assuming the null hypothesis.
(Empleando el error típico asintótico basado en la hipótesis nula)
- c Cannot be computed because the asymptotic standard error equals zero.
No se puede efectuar el cálculo porque el error típico asintótico es igual a cero.

GRÁFICA DE LA RELACIÓN

Se puede obtener una gráfica de barras si en el cuadro de diálogo de Tablas de contingencia se marca el casillero "Mostrar los gráficos de barras agrupadas" (Display clustered bar charts").



Al editar la gráfica, se obtiene:



PORCENTAJES POR CELDILLA

La tabla cruzada entre región y estatus se puede enriquecer hallando las frecuencias relativas por filas, por columnas y por el total.

En el cuadro de diálogo de Tablas de contingencia pulsar el botón "Casillas" (Cell) y marcar en la sección Porcentajes (Percentages): Fila (Row), Columna (Column) y Total (Total). Clic en continue y luego en OK.

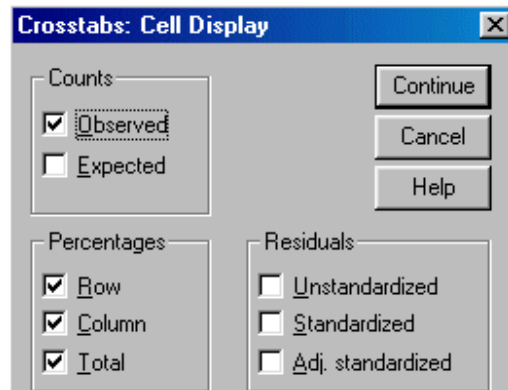


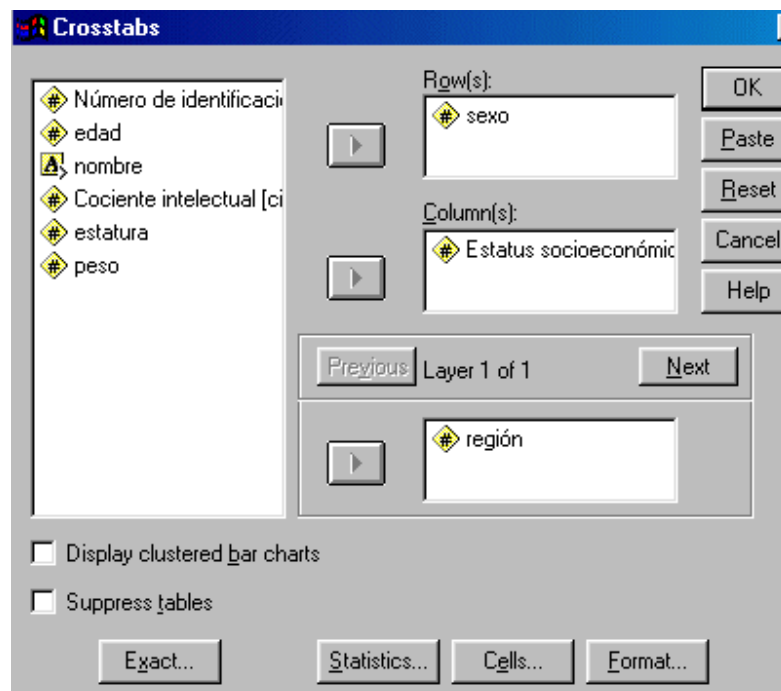
Tabla de contingencia SEXO * Estatus Socioeconómico

		Estatus socioeconómico			Total	
		Alto	Medio	Bajo		
SEXO	Hombre	Recuento	41	17	16	74
		% de SEXO	55.4%	23.0%	21.6%	100.0%
		% de Estatus socioeconómico	53.9%	41.5%	51.6%	50.0%
		% del Total	27.7%	11.5%	10.8%	50.0%
Mujer		Recuento	35	24	15	74
		% de SEXO	47.3%	32.4%	20.3%	100.0%
		% de Estatus socioeconómico	46.1%	58.5%	48.4%	50.0%
		% del Total	23.6%	16.2%	10.1%	50.0%
Total		Count	76	41	31	148
		% de SEXO	51.4%	27.7%	20.9%	100.0%
		% de Estatus socioeconómico	100.0%	100.0%	100.0%	100.0%
		% del Total	51.4%	27.7%	20.9%	100.0%

Más de dos variables de clasificación

También se puede describir relaciones bivariadas en base a más variables de clasificación, por ejemplo, tablas de la relación entre el estatus y el sexo para cada región.

- En el cuadro de diálogo de Tablas de contingencia seleccionar la variable región y transferirla en la ventana debajo de “Capa 1 de 1” (Layer 1 of 1).
- Desactivar la casilla: Display clustered bar charts.
- Desactivar las opciones de Statistics (estadísticas)
- De la opción Cell, sólo debe estar activado la casilla: Observed.
- Clic en Continue y luego en OK.



Resultados:

REGIÓN			Estatus socioeconómico			Total
			Alto	Medio	Bajo	
Norte	SEXO	Hombre	17	6	4	27
		Mujer	6	7	3	16
	Total		23	13	7	43
Este	SEXO	Hombre	11	3	3	17
		Mujer	8	7	5	20
	Total		19	10	8	37
Oeste	SEXO	Hombre	10	5	4	19
		Mujer	10	9	2	21
	Total		20	14	6	40
Sur	SEXO	Hombre	3	3	5	11
		Mujer	11	1	5	17
	Total		14	4	10	28

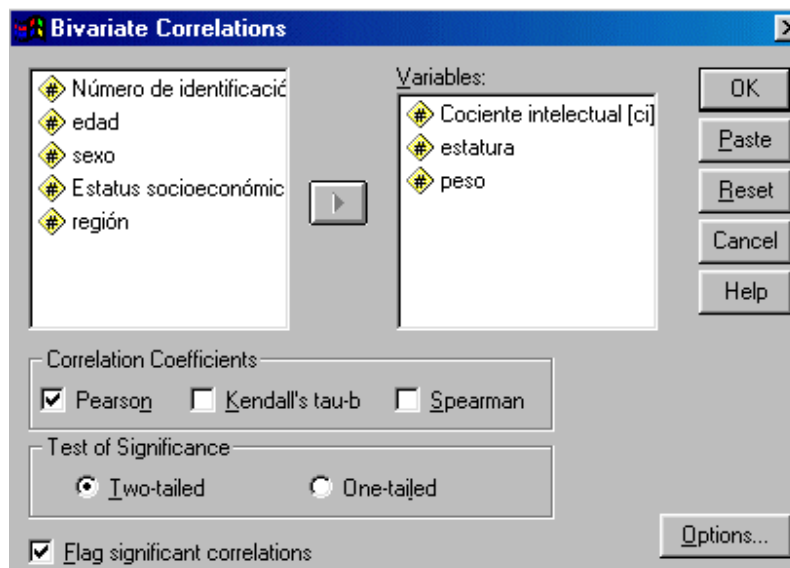
PRACTICA N° 13: RELACIÓN ENTRE VARIABLES DE INTERVALO O RAZÓN, RELACION ENTRE VARIABLES ORDINALES (CORRELATION)

El coeficiente de correlación r de Pearson es adecuado para medir la magnitud y el signo de la relación (lineal) entre dos variables cuantitativas (de intervalo o razón). El coeficiente r varía entre -1 y 1 . La fuerza de la asociación (o magnitud de la relación) es una función directa del valor absoluto del coeficiente r , cuánto más se acerca a 0 menor es la relación entre las variables, cuanto más se acerca a 1 (en valor absoluto) mayor es la relación entre las variables. La relación perfecta se da tanto cuando el coeficiente vale 1 como cuando vale -1 . Y no existe ninguna relación entre las variables cuando r es igual a 0 . El signo de r indica el signo de la relación, si es positivo indica: a mayor valor en una variable mayor valor en la otra variable, si es negativo indica: a mayor valor en una variable menor valor en la otra variable. Una relación de signo positivo se da entre el peso y la altura de las personas, en general cuanto mayor la altura, mayor el peso. Una relación de signo negativo se da entre el número de cigarrillos fumados y la esperanza de vida, en general cuantos más cigarrillos se fumen menos esperanza de vida se tiene.

El coeficiente r se interpreta normalmente como el porcentaje de varianza común entre las dos variables. Así un coeficiente $r=0,6$ quiere decir que el 36% de la variabilidad (varianza) de ambas variables es común. Este porcentaje se halla elevando al cuadrado r y multiplicando por cien: $0,6^2=0,36$; $0,36 \times 100= 36$.

1.- Hallar las relaciones bivariadas entre el peso, la estatura y el ci

Se elige Analizar—Correlaciones—Bivariadas. (Analyze—Correlation—Bivariate). Seleccionar y transferir cada una de las tres variables y pulsar el botón OK (Aceptar). Está premarcada la correlación de Pearson.



La tabla de con los coeficientes de correlación de Pearson aparece en la tabla abajo mostrada.

El coeficiente de correlación de Pearson entre peso y ci no es significativo: $r=0,001$ $p<0,998$.

El coeficiente entre estatura y ci tampoco es significativo: $r=0,081$ $p<0,325$.

El coeficiente de correlación entre peso y estatura sí es significativo: $r=0,6$ $p<0,000$.

Para comprobar la significación del coeficiente se realiza un contraste que sigue la distribución t de Student con n-2 grados de libertad, datos que el SPSS no proporciona; lo que sí ofrece es la probabilidad del estadístico t.

		Cociente intelectual	ESTATURA	PESO
Cociente intelectual	Correlación de Pearson	1.000	.081	.001
	Sig. (bilateral)	.	.325	.988
	N	149	149	148
ESTATURA	Correlación de Pearson	.081	1.000	.600
	Sig. (bilateral)	.325	.	.000
	N	149	150	149
PESO	Correlación de Pearson	.001	.600	1.000
	Sig. (bilateral)	.988	.000	.
	N	148	149	149

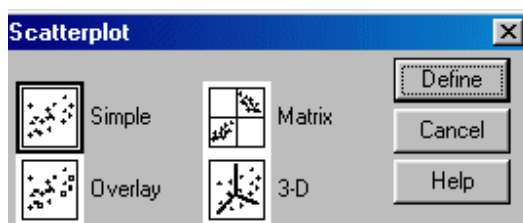
** Correlation is significant at the 0.01 level (2-tailed).
(La correlación es significativa al nivel 0,01(bilateral))

GRAFICA DE LA RELACION BIVARIADA (GRAPH SCATTERPLOT)

Se emplea para obtener una representación gráfica (la nube de puntos) de la relación entre dos variables. La nube de puntos consiste en un aglomerado de puntos en el que cada uno representa las puntuaciones obtenidas en las dos variables por cada caso. Se llama centroide al punto en el que coinciden las medias de las dos variables.

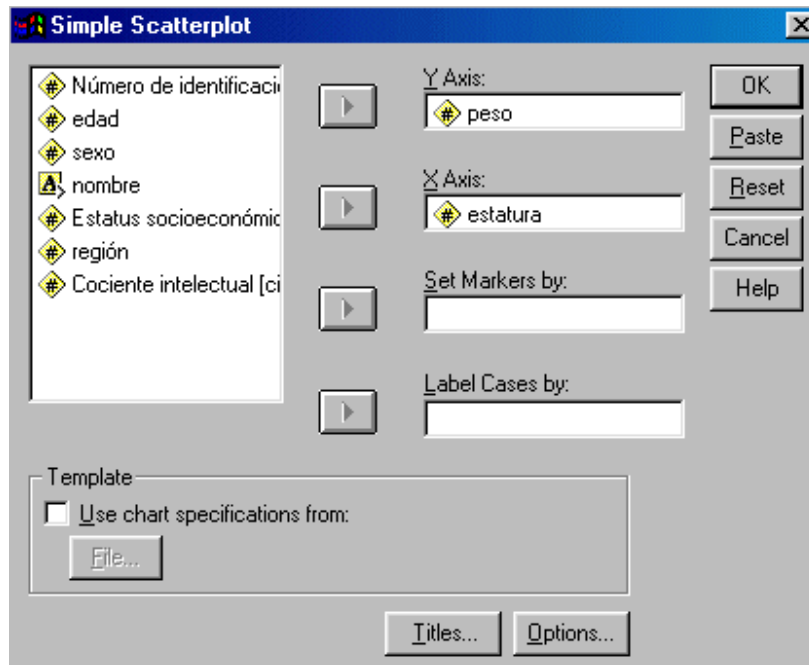
2.- Hallar la relación entre peso y estatura.

- Gráficos---Dispersión (Graph—Scatter). Definir. Está previamente marcada la opción "simple". La opción "Superpuesto" se emplea para representar varias relaciones bivariadas en la misma gráfica. La opción "Matricial" para representar más de una nube de puntos por separado. Y la opción 3-D para representar la relación entre variables.

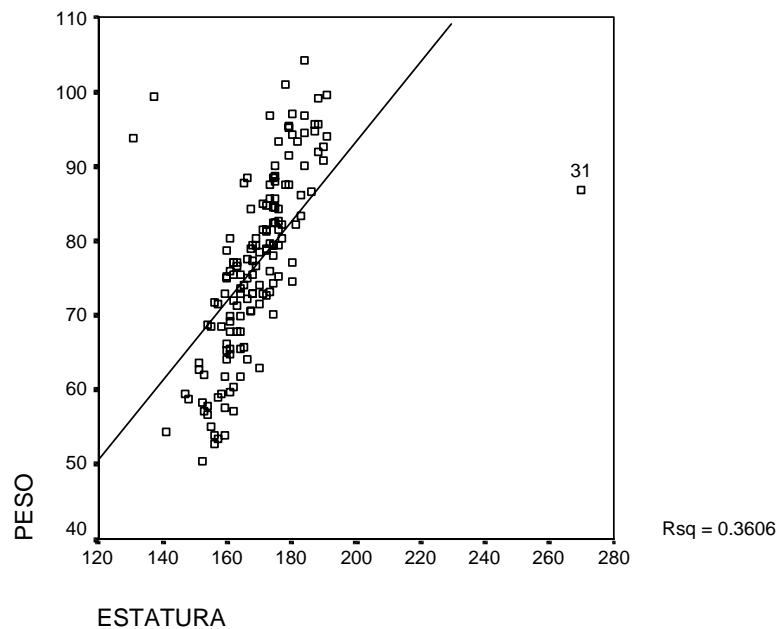


- Seleccionar y transferir la variable peso al eje Y (eje vertical). Seleccionar y transferir al eje X (eje horizontal) la variable estatura.

- OJO: No olvidar hacer Clic en Opciones (Options) y desactivar la casilla de “display groups defined by missing values” para no incluir los datos definidos como valor perdidos o datos en blanco. Clic en Continue y en OK.

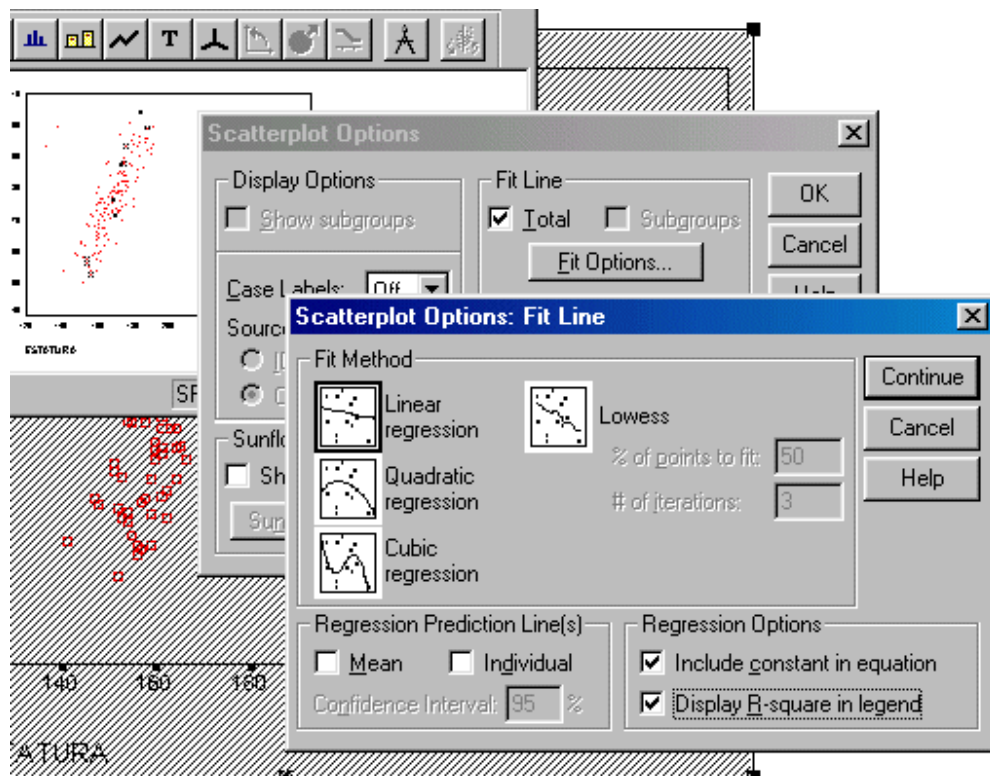


Resultado (luego de editar la gráfica):



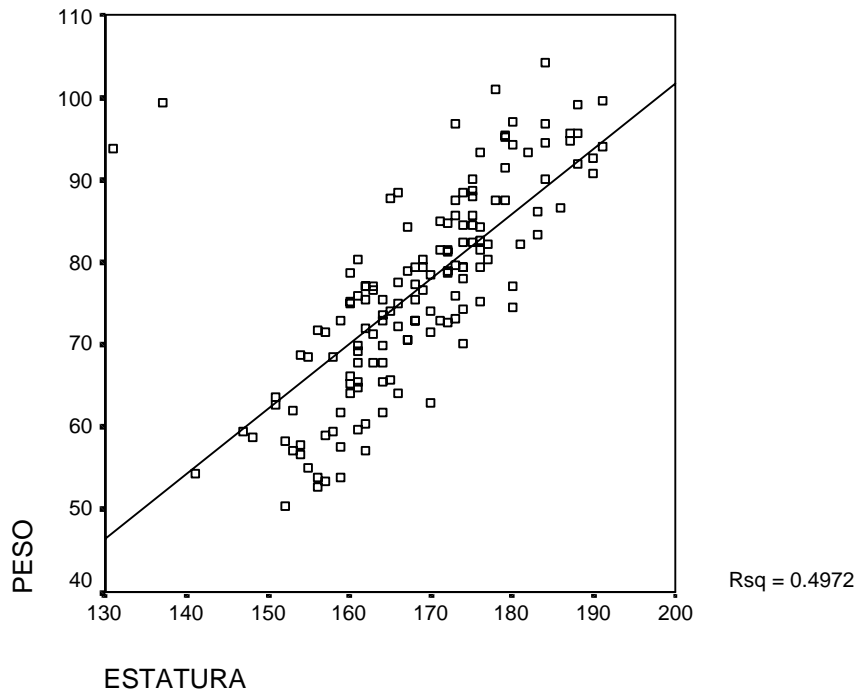
Editar la gráfica:

- El botón título se puede emplear para añadir títulos (hasta dos líneas), subtítulo, y notas al pie (hasta dos líneas) a la gráfica.
- La relación es claramente rectilínea y r es positivo. Lo que es evidente es la presencia de algunos casos extremos, como el situado en la parte superior derecha, el de un sujeto que mide más de 2,70 m de altura. (Sujeto 31).
- Para identificar el valor del caso extremo: Editar el gráfico. Hacer clic en el icono "Point identification" el cursor cambia de forma, mover el ratón hasta el sujeto y hacer clic sobre él, automáticamente se insertará el número de sujeto a su lado. Si ahora se sitúa en la ventana de datos se comprobará que tiene una estatura de 270 cm.
- Para incluir la recta de regresión: editar la gráfica. Clic en el icono Chart Options (Opciones de gráfico). En la sección "Ajustar línea" (Fit line) marcar Total.
- Si además se desea que figure el valor del coeficiente de correlación elevado al cuadrado (R^2) hacer clic en el botón "Opciones de ajuste" (Fit Options) y en el siguiente cuadro de diálogo marcar "Mostrar R-cuadrado en la leyenda" (Display R-square in legend) en la sección opciones de la regresión.



3.- Hallar la recta de regresión y el valor de R^2 tal como se hizo antes. El ajuste mejora notablemente hasta 0,4972 y también ha aumentado el valor de R^2 . Para hallar los parámetros de la recta de regresión (constante y pendiente).

- Menú: File –New – Sintaxis. Escribir:
IF (\$CASENUM EQ 31) estatura = 0.
FILTER BY estatura.
GRAPH SCATTERPLOT = estatura WITH peso.
- Menu: Run – All. Se obtiene el siguiente gráfico:



Determinación de los parámetros de la recta de regresión:

- Para hallar los parámetros de la recta de regresión (constante y pendiente) de la ecuación: $Y = b + m X$
- Clic sobre el Gráfico para seleccionarlo. Luego clic en Analizar—Regresión – Lineal (Analyze—Regression—Linear).
- Transferir peso como variable dependiente y estatura como variable independiente. Clic en OK.
- Los coeficientes se muestran en la tabla:

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	-55.831	11.054		-5.051	.000
	ESTATURA	.788	.066	.705	12.015	.000

a Dependent Variable: PESO

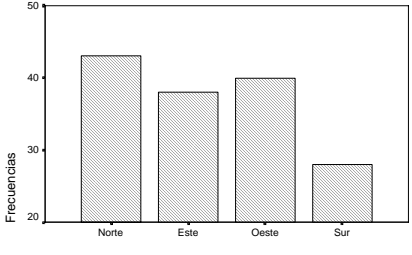
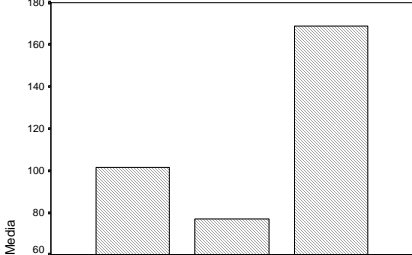
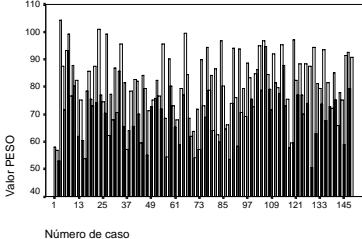
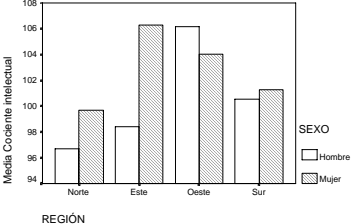
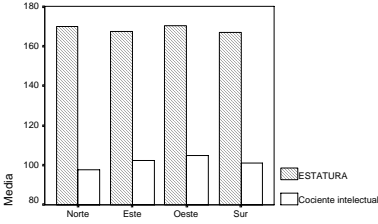
La ecuación es: $PESO = -55,83 + (0,78 * ESTATURA)$

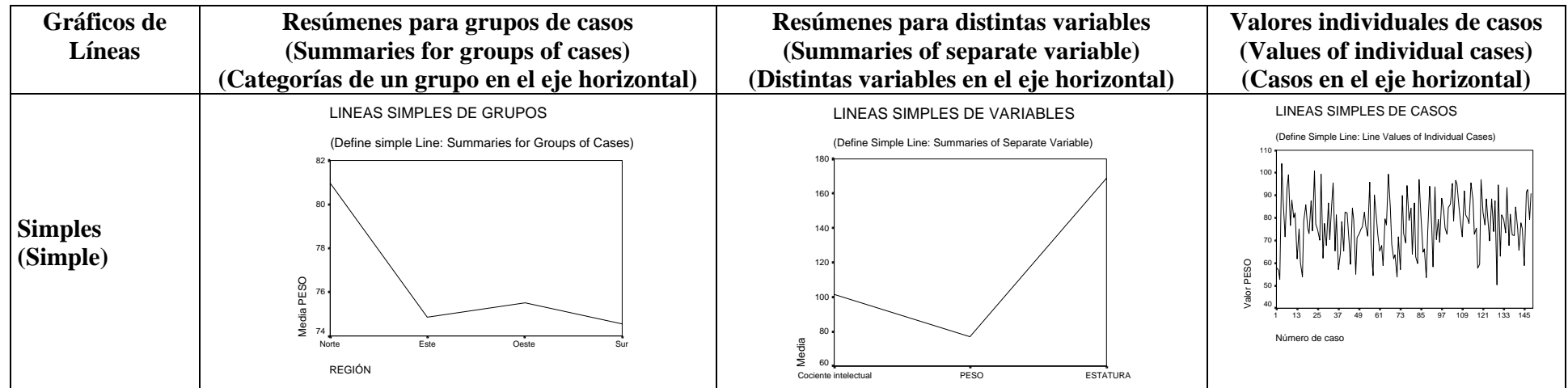
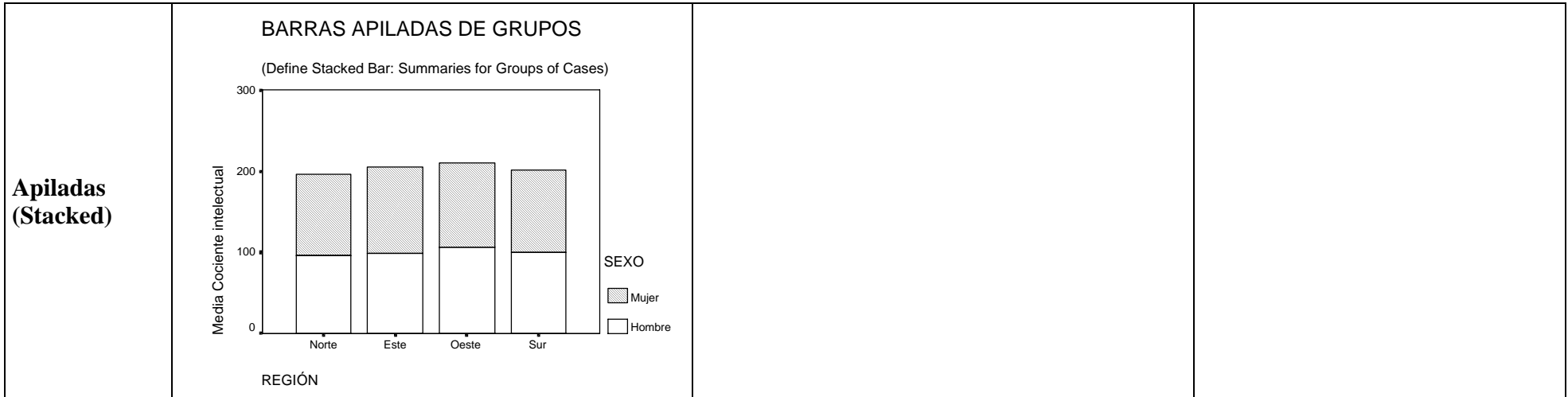
El punto de corte con el eje vertical es $-55,83$ y la pendiente es $0,78$.

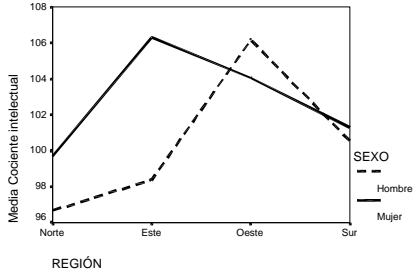
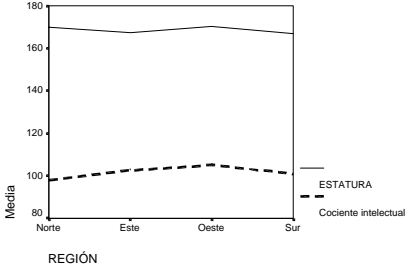
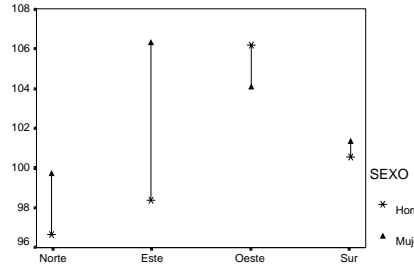
Nota: Hay dos casos que pueden eliminarse aún para aumentar R^2 hasta $0,7182$

PRACTICA N° 14: GRÁFICAS


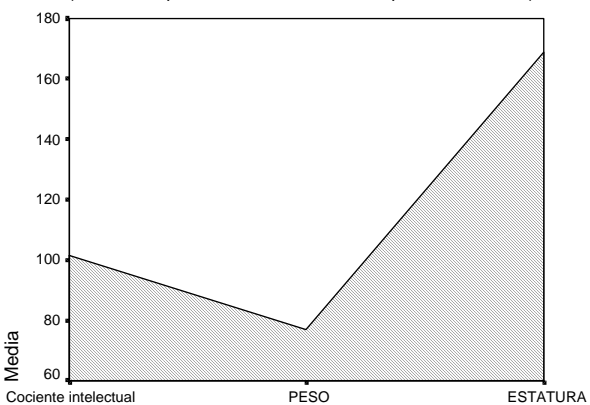
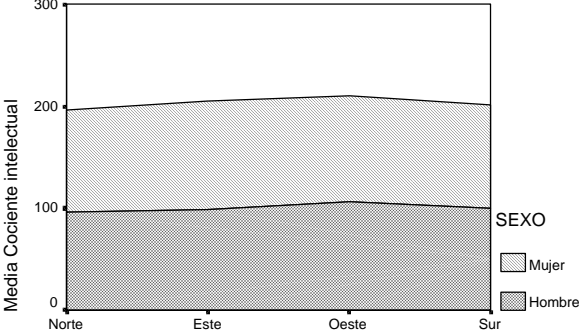
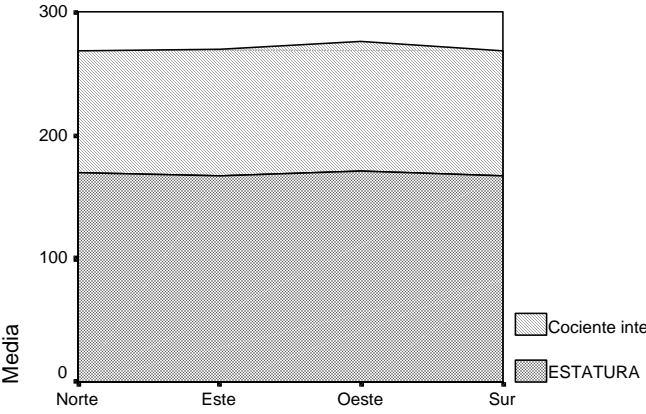
Obtener las siguientes Gráficas a partir del archivo de la Práctica N° 04.

Gráficos de Barras	Resúmenes para grupos de casos (Summaries for groups of cases) (Categorías de un grupo en el eje horizontal)	Resúmenes para distintas variables (Summaries of separate variable) (Distintas variables en el eje horizontal)	Valores individuales de casos (Values of individual cases) (Casos en el eje horizontal)
<p>Simples (Simple)</p>	<p>BARRAS SIMPLES AGRUPADAS (Define simple Bar: Summaries for Groups of Cases)</p>  <p>REGIÓN</p>	<p>BARRAS SIMPLES DE VARIABLES (Define Simple Bar: Summaries of Separate Variables)</p>  <p>Media</p>	<p>BARRAS SIMPLES DE CASOS (Define Simple Bar: Values of Individual Cases)</p>  <p>Valor PESO</p> <p>Número de caso</p>
<p>Agrupadas (Clustered)</p>	<p>BARRAS AGRUPADAS DE GRUPOS (Define Clustered Bar: Summaries for Groups of Cases)</p>  <p>Media Cociente intelectual</p> <p>REGIÓN</p> <p>SEXO Hombre Mujer</p>	<p>BARRAS AGRUPADAS DE VARIABLES (Define Clustered Bar: Summaries of Separate Variables)</p>  <p>Media</p> <p>REGIÓN</p> <p>ESTATURA Cociente intelectual</p>	



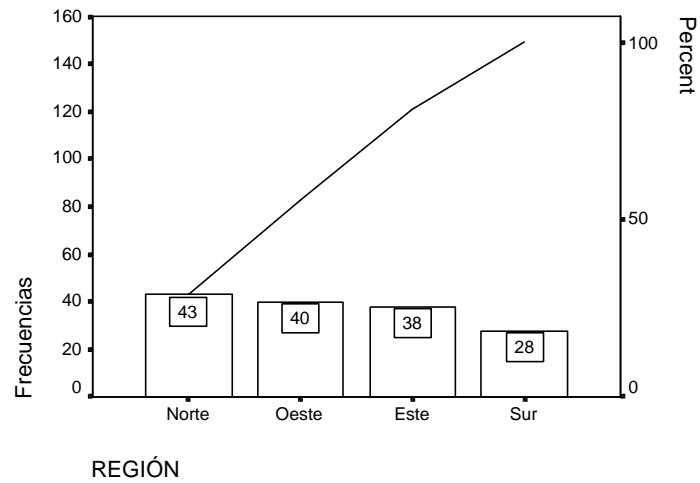
<p>Múltiples (Multiple)</p>	<p>LINEAS MULTIPLES DE GRUPOS (Define Multiple Line: Summaries for Groups of Cases)</p>  <p>Y-axis: Media Cociente intelectual (96-108) X-axis: REGIÓN (Norte, Este, Oeste, Sur) Legend: SEXO (Hombre: solid line, Mujer: dashed line)</p>	<p>LINEAS MULTIPLES DE VARIABLES (Define Multiple Line: summaries of Separate Variable)</p>  <p>Y-axis: Media (80-180) X-axis: REGIÓN (Norte, Este, Oeste, Sur) Legend: ESTATURA (solid line), Cociente intelectual (dashed line)</p>	
<p>Líneas verticales (Drop-Line)</p>	<p>LINEAS VERTICALES ENTRE GRUPOS (Define Drop-Line: Summaries for Groups of Cases)</p>  <p>Y-axis: Media Cociente intelectual (96-108) X-axis: REGIÓN (Norte, Este, Oeste, Sur) Legend: SEXO (Hombre: *, Mujer: ^)</p>	<p>* Ejemplos de Gráficas de Sectores. *File → New→ Sintaxis. Luego: Escribir los siguientes comandos: * 1 Simple de grupos. GRAPH /PIE = COUNT BY región /MISSING=exclude. * 2 Estadísticos de grupos. GRAPH /PIE = MEAN(ci) BY región /MISSING=exclude. * 3 Simple de variables. GRAPH /PIE = MEAN(peso estatura)/ MISSING=LISTWISE . Run → All</p>	

<p>Gráficos de Áreas</p>	<p>Resúmenes para grupos de casos (Summaries for groups of cases) (Categorías de un grupo en el eje horizontal)</p>	<p>Resúmenes para distintas variables (Summaries of separate variable) (Distintas variables en el eje horizontal)</p>	<p>Valores individuales de casos (Values of individual cases) (Casos en el eje horizontal)</p>
---------------------------------	--	--	---

<p>Simple (Simple)</p>	<p>AREAS DE GRUPOS (Define Simple Area: Summaries for Groups of Cases)</p>  <p>Media PESO</p> <p>REGIÓN</p> <p>* Otro ejemplo: 1.-Simple de grupos. GRAPH /LINE (AREA)=COUNT BY región /MISSING=exclude.</p>	<p>AREAS DE VARIABLES (Define Simple Area: Summaries of Separate Variables)</p>  <p>Media</p> <p>Cociente intelectual PESO ESTATURA</p>	<p>* 2.-Simple de sujetos. GRAPH /LINE (AREA)=VALUE (peso) .</p> <p>* 3.-Agrupada de sujetos. GRAPH /LINE (AREA)=VALUE (peso) by sexo .</p>
<p>Apiladas (Stacked)</p>	<p>AREAS APILADAS DE GRUPOS (Define Stacked Area: Summaries for Groups of Cases)</p>  <p>Media Cociente intelectual</p> <p>REGIÓN</p> <p>SEXO Mujer Hombre</p>	<p>AREAS DE VARIABLES AGRUPADAS (Define Stacked Area: Summaries of Separate Variables)</p>  <p>Media</p> <p>REGIÓN</p> <p>Cociente intel ESTATURA</p>	<p>*DIFERENCIAS .</p> <p>* 4.- Areas de diferencias de grupos agrupada. GRAPH /LINE (DIFFEREN)=MEAN (ci) BY región BY sexo /MISSING=exclude.</p> <p>* 5.- Areas de diferencias de variables agrupada. GRAPH /LINE (DIFFEREN)=MEAN (peso estatura) BY región/ MISSING=LISTWISE .</p> <p>* 6.- Diferencias de sujetos. GRAPH /LINE (DIFFEREN)=N (peso estatura) by sexo .</p>

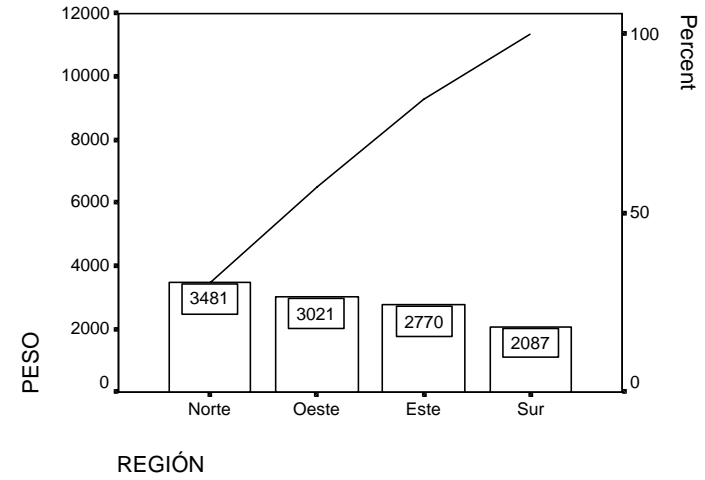
PARETO DE GRUPOS, FRECUENCIAS

(Define Simple Pareto: Counts or Sums for Groups of Cases)



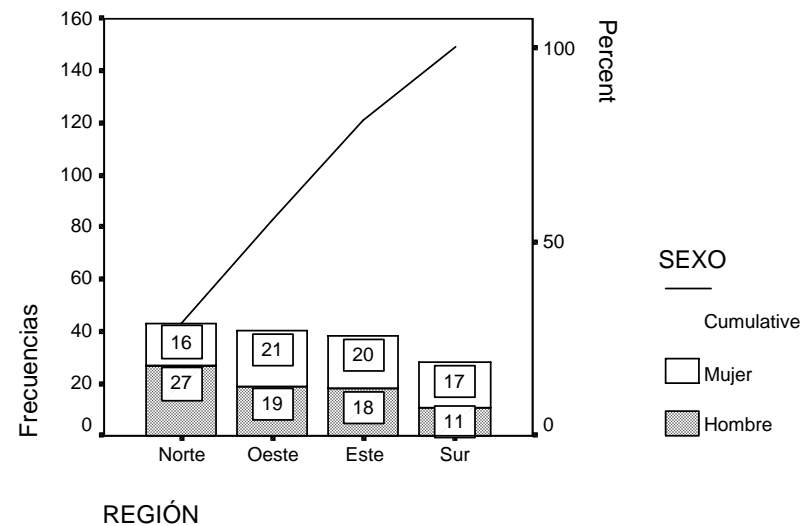
PARETO DE GRUPOS, SUMA

(Define Simple Pareto: Counts or Sums for Groups of Cases)

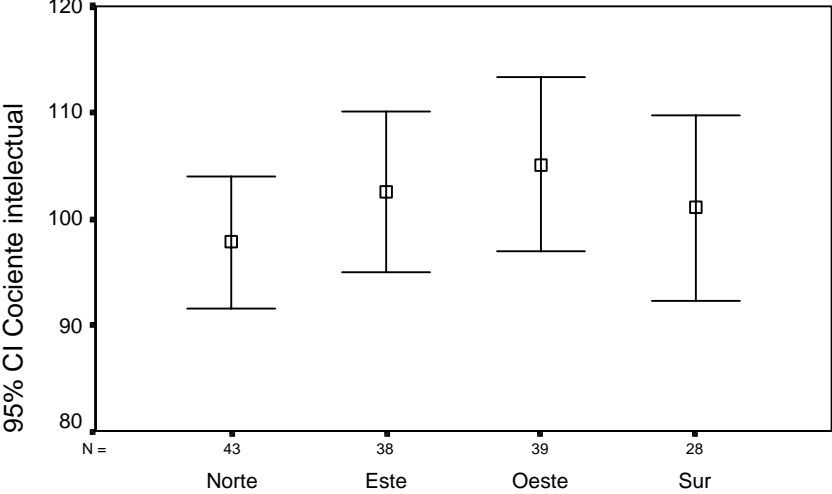


PARETO APILADA, FRECUENCIAS

(Define Stacked Pareto: Counts or Sums for Groups of Cases)



GRAFICAS DE CAJAS	Resúmenes para grupos de casos (Summaries for groups of cases) (Categorías de un grupo en el eje horizontal)	Resúmenes para distintas variables (Summaries of separate variable) (Distintas variables en el eje horizontal)
Simple (Simples)	<p>Box plot showing 'Cociente intelectual' for four regions: Norte, Este, Oeste, and Sur. The y-axis ranges from 40 to 180. Sample sizes are N=43, 38, 39, and 28 respectively. Outliers are present for Norte (cases 23 and 2).</p>	<p>Box plot showing 'Cociente intelectual', 'ESTATURA', and 'PESO' for three groups. The y-axis ranges from 0 to 300. Sample sizes are N=148 for each. Outliers are present for ESTATURA (cases 31, 3, 5).</p>
Clustered (Agrupadas)	<p>Clustered box plot showing 'Cociente intelectual' by region and sex. The y-axis ranges from 40 to 180. Regions are Norte, Este, Oeste, and Sur. Sexes are Hombre (white) and Mujer (shaded). Sample sizes are N=27, 16, 18, 20, 19, 20, 11, 17. Outlier for Norte (case 2).</p>	<p>Clustered box plot showing 'Cociente intelectual', 'ESTATURA', and 'PESO' by socioeconomic status. The y-axis ranges from 0 to 300. Statuses are Alto, Medio, and Bajo. Legend indicates Cociente intelectual (shaded), ESTATURA (diagonal lines), and PESO (white). Sample sizes are N=76, 76, 40, 40, 40, 31, 31, 31. Outliers for Alto (case 31) and Medio (case 27).</p>

GRAFICAS DE BARRAS DE ERROR	Resúmenes para grupos de casos (Summaries for groups of cases) (Categorías de un grupo en el eje horizontal)	Resúmenes para distintas variables (Summaries of separate variable) (Distintas variables en el eje horizontal)																									
Simple (Simples)	<p style="text-align: center;">BARRAS DE ERROR SIMPLES</p> <p style="text-align: center;">(Define Simple Error Bar: Summaries for Groups of Cases)</p>  <table border="1" style="margin-left: auto; margin-right: auto;"> <caption>Data for Simple Error Bars</caption> <thead> <tr> <th>REGIÓN</th> <th>N</th> <th>Mean (CI)</th> <th>95% CI Lower</th> <th>95% CI Upper</th> </tr> </thead> <tbody> <tr> <td>Norte</td> <td>43</td> <td>~98</td> <td>~92</td> <td>~104</td> </tr> <tr> <td>Este</td> <td>38</td> <td>~102</td> <td>~95</td> <td>~110</td> </tr> <tr> <td>Oeste</td> <td>39</td> <td>~105</td> <td>~97</td> <td>~113</td> </tr> <tr> <td>Sur</td> <td>28</td> <td>~101</td> <td>~93</td> <td>~110</td> </tr> </tbody> </table>	REGIÓN	N	Mean (CI)	95% CI Lower	95% CI Upper	Norte	43	~98	~92	~104	Este	38	~102	~95	~110	Oeste	39	~105	~97	~113	Sur	28	~101	~93	~110	<p>*BARRA DE ERROR DE VARIABLES</p> <p>GRAPH ERRORBAR(STDDEV 2)= ci estatura peso.</p>
REGIÓN	N	Mean (CI)	95% CI Lower	95% CI Upper																							
Norte	43	~98	~92	~104																							
Este	38	~102	~95	~110																							
Oeste	39	~105	~97	~113																							
Sur	28	~101	~93	~110																							
Clustered (Agrupadas)	<p>*BARRAS DE ERROR AGRUPADAS</p> <p>GRAPH ERRORBAR(STDDEV 2)= ci BY región BY sexo</p>	<p>*BARRAS DE ERROR DE VARIABLES AGRUPADAS</p> <p>GRAPH ERRORBAR(STDDEV 2)= ci estatura peso BY región.</p>																									

PRACTICA N° 15: PRÁCTICA CALIFICADA PRESIÓN ARTERIAL

Se ha diseñado un ensayo clínico aleatorizado, doble-ciego y controlado con placebo, para estudiar el efecto de dos alternativas terapéuticas en el control de la hipertensión arterial. Se han reclutado 100 pacientes hipertensos y estos han sido distribuidos aleatoriamente en tres grupos de tratamiento. A uno de los grupos (control) se le administró un placebo, a otro grupo se le administró un inhibidor de la enzima convertidora de la angiotensina (IECA) y al otro un tratamiento combinado de un diurético y un Antagonista del Calcio. Las variables respuesta final fueron las presiones arteriales sistólica y diastólica.

Los datos con las claves de aleatorización han sido introducidos en una base de datos que reside en la central de aleatorización, mientras que los datos clínicos han sido archivados en dos archivos distintos, uno para cada uno de los dos centros participantes en el estudio.

Las variables almacenadas en estos archivos clínicos son las siguientes:

Tabla 1.1 Variables en los ficheros con los datos del estudio. ("hipertens – Datos")

Define variable		Type			Column Format		Labels
Variable Name (Nombre)	Measurement (medida)	Define variable (Define variable)	Width (ancho)	Decimal places (Decimales)	Column width (Ancho)	Text alignment (alineación)	Variable Labels (Rótulo)
clave	Scale	numeric	4	0	5	Right	Clave de aleatorización
nombre	nominal	string	Characters: 3		11	Center	Iniciales
f_nacim	scale	Date	dd.mm.yyyy		11	Right	Fecha de nacimiento
f_inclus	scale	Date	dd.mm.yyyy		11	Right	Fecha de inclusión
sexo	nominal	Numeric	2	0	5	Right	0:Varón 1:Mujer
altura	scale	Numeric	8	0	8	Right	Altura (cm)
peso	Scale	Numeric	6	0	8	Right	Peso en kg.
pad_ini	Scale	Numeric	3	0	8	Right	Presión diastólica inicial
pad_fin	Scale	Numeric	3	0	8	Right	Presión diastólica final
pas_ini	Scale	Numeric	3	0	8	Right	Presión sistólica inicial
pas_fin	scale	Numeric	3	0	8	Right	Presión sistólica final

Nota: En la variable sexo. Para Labels. Value: 0 Value label:Varón; Value: 1 Value label:Mujer

El archivo de claves de aleatorización contiene sólo dos variables.

Tabla 1.2. Variables en el fichero de claves “hipertens – Claves”.

Define variable		Type			Column Format		Labels
Variable Name (Nombre)	Measurement (medida)	Define variable (Define variable)	Width (ancho)	Decimal places (Decimales)	Column width (Ancho)	Text alignment (alineación)	Variable Labels (Rótulo)
clave	Scale	numeric	4	0	5	Right	Clave de aleatorización
FARMACO	nominal	numeric	1	0	5	Right	Tratamiento aplicado Value / Value label 0 / Placebo 1 / IECA 2 / Ca Antagonista+diurético

La Tabla con los datos relativos a los 10 de los pacientes enrolados en uno de los centros de investigación es la siguiente:

Tabla 1.3. Datos de los 10 pacientes reclutados en el centro con 10 pacientes

clave	Nombre	F_macim	F_inclus	Sexo	Altura	Peso	Pad_ini	Pad_fin	Pas_ini	Pas_fin
1	SGL	8-9-1941	13-7-1998	1	165	78	78	104	176	175
2	JCG	10-7-1957	9-5-1998	1	154	74	95	114	162	160
3	APZ	18-8-1967	1-4-2000	0	156	81	93	102	141	150
4	NDG	8-5-1956	13-11-1998	0	181	82	86	91	162	161
5	CLO	2-11-1958	24-2-1999	1	184	78	89	94	165	162
6	LFZ	13-6-1953	16-3-2000	0	179	80	74	99	141	148
7	OAR	20-7-1961	21-8-1999	1	159	60	94	95	168	152
8	SGH	6-3-1970	16-12-1999	0	153	70	94	79	133	143
9	ZLZ	6-2-1962	21-6-2000	1	172	57	74	64	149	135
10	YSJ	5-4-1968	15-10-1998	0	154	54	89	87	147	155

Operaciones a realizar con los archivos de datos

Guardar todos los archivos o ficheros dentro de una Carpeta de trabajo que tenga como nombre su primer nombre y su apellido paterno.

1. Crear el fichero de datos (Tabla 1.1: Variables) definida con once variables. Luego introducir los datos de la Tabla 1.3. Grabar este archivo asignándole el nombre "Apellidos - Datos (1-10)". Revisar bien las 11 variables definidas
 - Unir o Fusionar (Merge Files) el archivo: "hipertens - Datos (1-100)": Activando el menú DATA → MERGE FILES → ADD CASES. Seleccionar el archivo: "hipertens – Datos (11-100)" y clic en Abrir. Observe si se unieron correctamente los datos.
 - Borrar los casos o filas en blanco del 11 al 21: Seleccionar la columna numerada en la barra izquierda. Luego Edit → Clear.
 - Grabar (File → Save as) los datos fusionados (100 pacientes) a un archivo con el nombre: "Apellidos - Datos (1-100)". Observe los resultados.
2. Con archivo: Apellidos-Datos (1-100) abierto o activo agregar el archivo "hipertens - claves aleatorizacion" que contiene la variable Fármaco:
 - Activando Data → Merge Files → Add Variables. Seleccionar el archivo con las claves denominado: "hipertens - claves aleatorizacion" y clic en abrir.
 - Grabar el archivo resultante con el nombre: "Apellidos - Datos+Claves" .
3. Grabar (o exportar) el archivo "Apellidos-Datos+Claves" que contiene los datos de todo el estudio (100 pacientes) a diferentes formatos:
 - Formato Excel con el nombre: "Apellidos - excel.xls": FILE → SAVE AS.
En el cuadro de texto desplegable: "Guardar como archivo de tipo" seleccionar Excel (*.xls).
 - Formato con Texto delimitado por tabuladores con el nombre: "Apellidos - tab.dat" .
En el cuadro de texto desplegable: "Guardar como archivo de tipo" seleccionar: Tab-delimited (*.dat).
 - Formato Texto de ancho fijo con el nombre "Apellidos - fijo.dat"
En el cuadro de texto desplegable: "Guardar como archivo de tipo" seleccionar : Fixed ASCII (*.dat).
4. Crear un archivo para cada uno de los grupos de tratamiento. Denominar a estos archivos "Apellidos - placebo", "Apellidos - IECA" e "Apellidos - Ca", respectivamente. (Sugerencia: Utilice la opción:
 - Data→Select CasesSeleccionar casos)
 - Recuerde que la variable Tratamiento aplicado (fármaco) corresponde placebo = 0; IECA =1 y Ca Antagonista+diurético =2

Operaciones a realizar con las variables del archivo de datos.

Una vez introducidos, grabados y fusionados los correspondientes archivos de datos, sobre el archivo completo (con 100 pacientes con los datos clínicos y los fármacos administrados a cada sujeto), realizar las siguientes operaciones con el archivo: **“Apellidos-Datos+Claves.SAV”**:

5. Calcular, para cada paciente al momento del estudio, la edad en años (redondeando al entero más próximo). Denominar la nueva variable "Edad" y etiquetarla correspondientemente. Releer ejemplos de prácticas anteriores sobre cálculo con fechas para utilizar: `CTIME.DAYS(fecha)` y `RDN (EXPRESIÓN NUMÉRICA)`. Ej.

$$\text{RND} ((\text{CTIME.DAYS}(\text{F_INCLUS})-\text{CTIME.DAYS}(\text{F_NACIM})) / 365)$$

6.- Recodificar dicha edad de forma que la nueva variable, de nombre `clasedad`, tome los siguientes valores. Luego colocar las etiquetas de valor.

Edad en años	Clasedad	Etiquetas (Labels)
Menores de 36	1	"Menores de 36 años"
De 37 a 44	2	"De 37 a 44"
De 45 a 52	3	"De 45 a 52"
Mayores de 52	4	"Mayores de 52"

7. Calcular, para cada paciente, el índice de masa corporal (IMC):

$$\text{PESO} / (\text{ALTURA} / 100)^2$$

Nota:

- Altura está dada en centímetros. En la fórmula se divide entre 100 para obtener metros.
- Utilice los botones que se muestran en la opción del menú `Transform→Compute`. El símbolo de potencia es `**`

8. Recodificar dicho índice de masa corporal de forma que la nueva variable, de nombre obesidad, tome los siguientes valores y etiquetas de valor según el sexo del paciente.
- (Sugerencia: Data→ Select Cases. Si la condición satisface. Si es sexo= 1 (para el caso de Mujer). Luego: Transform → Recode→ Into different Variables.
 - Grabar como: “Apellidos-Obesidad-Mujer”
 - Volver a deseleccionar la opción activada. Data→Select cases→ Reset. OK

Sexo	IMC	ObeMujer	Etiqueta
1	Menor de 19	1	“Desnutrida”
	19,01 a 24,94	2	“Normal”
	24,95 a 29,94	3	“Sobre peso”
	29,95 a 39,94	4	“Obesa”
	Mayor de 39,95	5	“Muy obesa”

- Abrir el archivo de trabajo: “Apellidos-Datos+Claves”
- Sugerencia: Data→ Select Cases. Si la condición satisface. Si es sexo= 0 (para el caso de varón). Luego: Transform → Recode→ Into different Variables.
- Grabar como: “Apellidos-Obesidad-Varon”
- Volver a deseleccionar la opción activada. Data→Select cases→ Reset. OK

Sexo	IMC	ObeVaron	Etiqueta
0	Menor de 21	1	“Desnutrido”
	21,01 a 26,94	2	“Normal”
	26,95 a 32,94	3	“Sobre peso”
	32,95 a 43,94	4	“Obeso”
	Mayor de 43,95	5	“Muy obeso”

Estadística descriptiva

Realizar una descripción de los pacientes enrolados en el estudio. Analizar sus características en el momento de inclusión en el estudio y preparar un visor de resultados que contenga exclusivamente los siguientes elementos. Realizar este análisis primero para el total de la muestra y posteriormente estratificando los resultados por tratamiento recibido.

09.- Frecuencia de varones y mujeres

10.- Para la altura.

- Determinar las Frecuencias.
- Determinar los estadísticos: Media, Desviación estándar. Mostrar además los percentiles 25, 50 y 75.
- Gráfico: Histograma (Editar el Histograma)

11.- Mediante procedimiento descriptivo determinar para la Altura: N, rango, mínimo, máximo, media, desviación típica.

12.- Mediante el procedimiento de estadística descriptiva: Explorar. Estratificar los resultados para la variable altura estratificando según los *tratamientos* recibidos (Placebo, IECA, Ca Antagonista+diurético) y, generar gráficos como el histograma para Fármaco = Ca Antagonista+Diurético y el diagrama de cajas y patillas (Altura vs. Tratamiento aplicado). Editar los gráficos.

13.- Determinar los resultados según la Prueba de Kolmogorov-Smirnov (Sample K-S), donde, La prueba de normalidad para la variable índice de masa corporal.

14.- Determinar los resultados según el Procedimiento descriptivo Medias de la variable *índice de masa corporal*, estratificando por *tratamiento* recibido y por *sexo* de los pacientes.

15.- Según la Prueba T de Student para muestras independientes

Mostrar los resultados de la comparación de las medias de la variable *índice de masa corporal* según el *sexo* de los pacientes.

16.- Según la Prueba T de Student para muestras relacionadas determinar los resultados siguientes se refieren al contraste de la *presión arterial sistólica inicial y final* del conjunto de los pacientes

17.- Resultados sobre Características basales referidos a las variables Edad en años, Índice de masa corporal (IMC), presión arterial media inicial; según tratamiento del conjunto de los pacientes.

Nota: Presion diferencial = presion maxima o sistolica – presion minima o diastolica

Presion arterial media = presion diastolica + (1/3) x presion diferencial

18.- Resultados sobre Características cualitativas basales según sexo y tratamiento recibido.

RESPUESTAS:

Tras realizar las operaciones planteadas anteriormente, se obtienen los resultados relativos a las características de los pacientes en el momento de la inclusión en el estudio.

Pregunta 9: Resultados Procedimiento Frecuencias

A continuación se muestran los resultados del análisis de las variables *sexo* y *altura*, obtenidos mediante este procedimiento, de acuerdo a las siguientes opciones:

- Analyze → Descriptive Statistics → Frequencies. OK

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos Varón	54	54,0	54,0	54,0
Mujer	46	46,0	46,0	100,0
Total	100	100,0	100,0	

Pregunta 10: Resultados para la Variable ALTURA:

- Analyze → Descriptive Statistics → Frequencies. Pasar la variable: Altura
- Statistics. Seleccionar: Mean, Std. Deviation
- Percentile(s): 25 (add), 50 (Add), 75 (Add). Continue
- Charts. Histogram. With normal curve. Continue. OK.

Estadístico			Altura (cm)				
Altura			Válidos	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
N	Válido	100	146	1	1,0	1,0	1,0
	Perdido	0	150	3	3,0	3,0	4,0
Medi		167,7	152	1	1,0	1,0	5,0
Desv.		10,38	153	1	1,0	1,0	6,0
Percentile	25	159,0	154	3	3,0	3,0	9,0
	50	167,0	155	2	2,0	2,0	11,0
	75	175,0	156	7	7,0	7,0	18,0
			157	2	2,0	2,0	20,0
			158	3	3,0	3,0	23,0
			159	4	4,0	4,0	27,0
			160	1	1,0	1,0	28,0
			161	1	1,0	1,0	29,0
			162	3	3,0	3,0	32,0
			163	3	3,0	3,0	35,0
			164	5	5,0	5,0	40,0
			165	4	4,0	4,0	44,0
			166	1	1,0	1,0	45,0
			167	7	7,0	7,0	52,0
			168	2	2,0	2,0	54,0
			169	3	3,0	3,0	57,0
			170	4	4,0	4,0	61,0
			171	5	5,0	5,0	66,0
			172	2	2,0	2,0	68,0
			173	4	4,0	4,0	72,0
			174	2	2,0	2,0	74,0
			175	3	3,0	3,0	77,0
			176	1	1,0	1,0	78,0
			177	3	3,0	3,0	81,0
			178	2	2,0	2,0	83,0
			179	3	3,0	3,0	86,0
			180	1	1,0	1,0	87,0
			181	1	1,0	1,0	88,0
			182	1	1,0	1,0	89,0
			183	1	1,0	1,0	90,0
			184	2	2,0	2,0	92,0
			185	3	3,0	3,0	95,0
			186	1	1,0	1,0	96,0
			187	2	2,0	2,0	98,0
			188	1	1,0	1,0	99,0
			194	1	1,0	1,0	100,0
			Total	100	100,0	100,0	

Luego de editar el eje Y y el eje X:

- Doble clic en Eje X y en el eje X de la ventanita de edición.
- Labels. Orientation: Horizontal. Activar Every. Continue. OK

11.- Resultados Procedimiento Descriptivos

La tabla siguiente ha sido producida por este procedimiento de acuerdo a la sintaxis del cuadro. Se analiza la variable *altura*.

Estadísticos descriptivos

	N	Rango	Mínimo	Máximo	Media	Desv. típ.
Altura (cm)	100	48	146	194	167,78	10,38
N válido (según lista)	100					

- Analyze → Reports → Case Summaries
- Variables: Altura (cm) [altura]
- Desactivar: Display cases
- Statistics: Activar Number of cases, Range, Minimum, Maximum, Mean, Standard Deviation
- Options: Escriba "Estadísticos Descriptivos" y activar: subheading for totals y exclude cases with missing values listwise
- Continue. OK

12.- Resultados Procedimiento Explorar

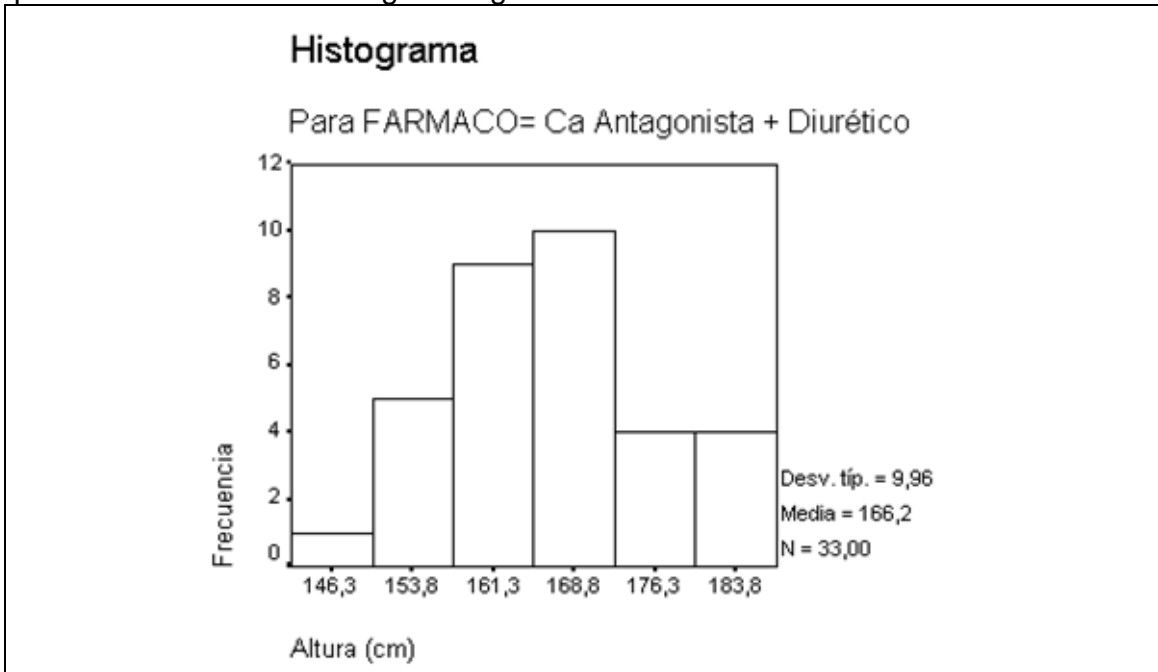
Las tablas siguientes han sido producidas por este procedimiento de acuerdo a la sintaxis del cuadro. Obsérvese que los resultados se han estratificado según los *tratamientos* recibidos.

- Analyze → Descriptive Statistics → Explore
- Dependent List: Altura (cm) [altura]. En Factor List: Tratamiento aplicado [fármaco]
- Plots. Activar Histograma y desactivar Stem-and-Leaf plots. Continue.
- Options: Exclude cases pairwise. Continue y OK.

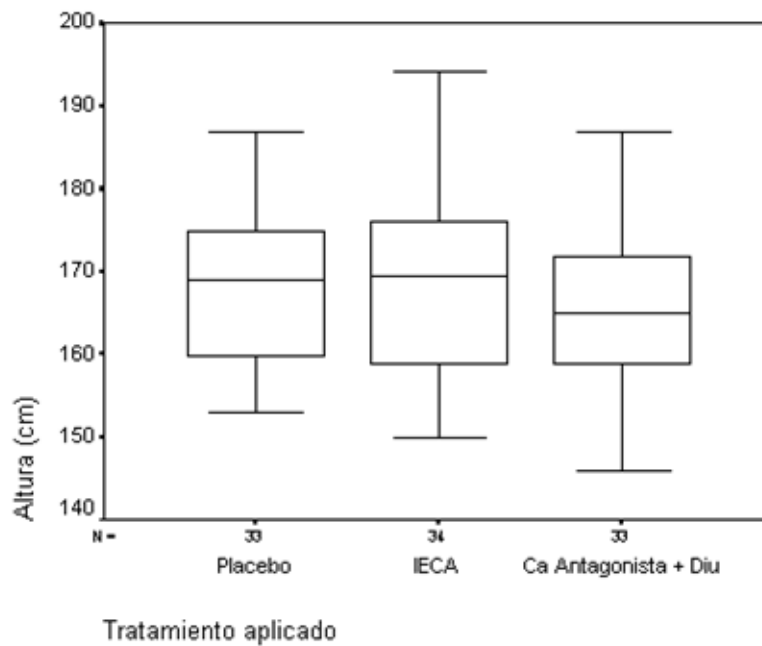
Descriptivos

Tratamiento			Estadístico	Error típ.		
Altura (cm)	Placebo	Media	168,18	1,62		
		Intervalo de confianza para la media al 95%	Límite inferior	164,88		
			Límite superior	171,48		
			Media recortada al 5%	168,00		
		Mediana	169,00			
		Varianza	86,716			
		Desv. típ.	9,31			
		Mínimo	153			
		Máximo	187			
		Rango	34			
		Amplitud intercuartil	16,50			
		Asimetría	,062	,409		
		Curtosis	-,837	,798		
		IECA		Media	168,97	2,02
				Intervalo de confianza para la media al 95%	Límite inferior	164,86
Límite superior	173,08					
Media recortada al 5%	168,77					
Mediana	169,50					
Varianza	138,878					
Desv. típ.	11,78					
Mínimo	150					
Máximo	194					
Rango	44					
Amplitud intercuartil	18,25					
Asimetría	,154			,403		
Curtosis	-,721			,788		
Ca Antagonista + Diurético				Media	166,15	1,73
				Intervalo de confianza para la media al 95%	Límite inferior	162,62
		Límite superior	169,68			
		Media recortada al 5%	166,03			
		Mediana	165,00			
		Varianza	99,258			
		Desv. típ.	9,96			
		Mínimo	146			
		Máximo	187			
		Rango	41			
		Amplitud intercuartil	14,00			
		Asimetría	,384	,409		
		Curtosis	-,279	,798		

Por último, este procedimiento genera gráficos como el histograma y el diagrama de cajas y patillas mostrados en las figuras siguientes.



- Para editar el gráfico anterior: Doble clic en el eje X, luego doble clic en el eje X de la ventana de edición.
- Labels. Orientation: Horizontal. Continue
- Custom → Define . Cambiar: # of interval de 9 a 6. Interval width de 5 a 7.5.
- Continue. OK.



13.-Resultados Prueba de Kolmogorov-Smirnov (K-S)

La prueba de normalidad para la variable *índice de masa corporal* se muestra en la siguiente tabla.

Prueba de Kolmogorov-Smirnov para una muestra

		Índice de masa corporal
N		100
Parámetros normales ^{a,b}	Media	25,9685
	Desviación típica	4,9237
Diferencias más extremas	Absoluta	,077
	Positiva	,067
	Negativa	-,077
Z de Kolmogorov-Smirnov		,773
Sig. asintót. (bilateral)		,588

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

- Analyze → Nonparametrics Tests → Sample K-S
- Test variable List: imc. OK

14.-Resultados Procedimiento Medias

A continuación se muestran los resultados del procedimiento descriptivo Medias de la variable *índice de masa corporal*, estratificando por *tratamiento* recibido y por *sexo* de los pacientes.

Informe

Índice de masa corporal

Tratamiento aplicado	Sexo	Media	N	Desv. típ.
Placebo	Varón	26,5189	19	5,0819
	Mujer	24,5163	14	4,4598
	Total	25,6693	33	4,8598
IECA	Varón	25,5215	19	5,4966
	Mujer	25,0789	15	5,2713
	Total	25,3262	34	5,3214
Ca Antagonista + Diurético	Varón	26,6190	16	4,8124
	Mujer	27,2217	17	4,4157
	Total	26,9295	33	4,5496
Total	Varón	26,1976	54	5,0834
	Mujer	25,6996	46	4,7710
	Total	25,9685	100	4,9237

- Analyze → Compare Means → Means
- Dependent List: imc Independent List: tratamiento aplicado [farmaco].
- Next. Independent List: sexo. OK

15.-Resultados Prueba T de Student para muestras independientes

En las tablas siguientes se muestran los resultados de la comparación de las medias de la variable *Índice de masa corporal* según el *sexo* de los pacientes.

Estadísticos de grupo					
	Sexo	N	Media	Desviación típ.	Error típ. de la media
Índice de masa corporal	Varón	54	26,1976	5,0834	,6918
	Mujer	46	25,6996	4,7710	,7035

Prueba de muestras independientes											
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Inferior	Superior	
Índice de masa corporal	Se han asumido varianzas iguales	,550	,460	,502	98	,617	,4980	,9917	-1,4699	2,4659	
	No se han asumido varianzas iguales			,505	97,054	,615	,4980	,9866	-1,4601	2,4661	

- Analyze → Compare Means → Independent-Samples T Test
- Test Variable(s): imc
- Grouping variable(s): sexo (??)
- Define groups: Grupo 1: 0 Grupo 2: 1. Continue. OK

16.-Resultados Prueba T de Student para muestras relacionadas

Los resultados siguientes se refieren al contraste de la *presión arterial sistólica inicial y final* del conjunto de los pacientes

Estadísticos de muestras relacionadas					
		Media	N	Desviación típ.	Error típ. de la media
Par 1	Presión sistólica inicial	152,52	100	15,13	1,51
	Presión sistólica final	149,86	100	16,63	1,66

Correlaciones de muestras relacionadas				
		N	Correlación	Sig.
Par 1	Presión sistólica inicial y Presión sistólica final	100	,843	,000

Prueba de muestras relacionadas									
		Diferencias relacionadas				t	gl	Sig. (bilateral)	
		Media	Desviación típ.	Error típ. de la media	95% Intervalo de confianza para la diferencia				
					Inferior	Superior			
Par 1	Presión sistólica inicial - Presión sistólica final	2,66	9,02	,90	,87	4,45	2,948	99	,004

- Analyze → Compare Means → Paired Samples T Test
- Seleccionar las variables Presion sistólica inicial y presión sistólica final y transferirlas amabas a la vez al cuadro de diálogo Paired variables. Luego OK.

17.-Resultados sobre Características basales

Los resultados siguientes se refieren a las variables Edad en años y Tratamiento del conjunto de los pacientes

Que luego de editarla, se debe obtener:

Características basales									
	Placebo		IECA		Ca Antagonista + Diurético		Total de grupo		
	Media	SD	Media	SD	Media	SD	Media	SD	
Edad en años	42,00	(9,76)	44,82	(9,63)	44,03	(9,71)	43,63	(9,67)	
Índice de masa corporal	25,67	(4,86)	25,33	(5,32)	26,93	(4,55)	25,97	(4,92)	
Presión media inicial	109,36	(7,94)	110,79	(7,62)	109,17	(7,26)	109,79	(7,57)	

- Analyze → Compare Means → Means
- Dependent List: edad imc
- Independent List: tratamiento aplicado [farmaco]
- Options: En Cell Statistics activar: Mean, Standard deviation.
- Continue. OK.
- Para editar la Tabla: Doble clic sobre la Tabla. Clic en Pívor → Transpose rows and Columns.

18.-Resultados sobre Características cualitativas basales según sexo y tratamiento recibido.

		Placebo		IECA		Ca Antagonista + Diurético	
		N	%	n	%	n	%
Sexo	Varón	19	57,6%	19	55,9%	16	48,5%
	Mujer	14	42,4%	15	44,1%	17	51,5%

- Analyze → Descriptive Statistics → Crosstabs
- Row(s): sexo (sexo)
- Column(s): tratamiento aplicado [fármaco]
- Cell(s): Percentages. Activar: Column.

PRACTICA N° 16: TRABAJO DE APLICACIÓN

SANEAMIENTO AMBIENTAL

1.- ¿Su vivienda es?

- Propia ()
 - Alquilada ()
 - Otros1 _____
- Especifique

2.- ¿El material de construcción de su vivienda es?

- Material Noble ()
- Adobe ()
- Caña ()
- Estera ()
- Triplay ()
- Otros2 _____

3.- ¿Cuántas habitaciones tiene:
Número: _____

- Dormitorios ()
 - Baños ()
 - Cocina ()
 - Sala ()
 - Comedor ()
 - Otros3 _____
- Especifique

4.- ¿Cuántas personas duermen por dormitorio?
Número: _____

5.- ¿Su vivienda cuenta con jardín?
SI () NO ()

III SERVICIOS BASICOS DE SALUBRIDAD

6.- ¿Cómo se abastece de agua?
Red Intradomiciliaria
Pileta ()
Pozo ()
Otros6 _____

Especifique

7.- ¿En qué almacena el agua?

- Cisterna ()
- Cilindro ()
- Balde ()
- Lata ()
- Otros7 _____

8.- ¿Hierve el agua que toma?

SI () NO ()

9.- ¿Cuenta con energía eléctrica?

SI () NO ()

Horario

10. ¿Dónde elimina sus excretas?

- Desague ()
- Letrina ()
- Otros10 _____

¿Cómo desecha su basura?

11. Camión de basura ()
Arrojan al campo ()
La queman ()
Contenedor ()
Otros11 _____

12. ¿Qué animales domésticos tiene en su vivienda?

- Aves ()
- Perros ()
- Gatos ()
- Cerdos ()
- Otros12 _____

13. ¿Qué animales transmisores de enfermedades existen en su vivienda?

- Moscas ()
- Roedores ()
- Cucarachas ()
- Zancudos ()
- Pulgas ()
- Otros13 _____

IV. ASPECTOS DE SALUD

14. ¿Cuáles son los síntomas y signos más frecuentes de enfermedades que ha sufrido su familia?
 SINTOMA TIEMPO DE DURACION
 Tos : _____
 Diarrea: _____
 Fatiga: _____
 Silbido pecho: _____
 Ardor al orinar: _____
 Dolor de cabeza : _____
 Granos: _____
 Picazón: _____
 Otros14 _____
 Especifique
15. ¿Cuando algún miembro de su familia enferma, a dónde acude?
 Puesto de Salud ()
 Centro de Salud ()
 Hospital ()
 Centro Particular ()
 Farmacia ()
 Botica ()
 Curandero ()
16. ¿Algún miembro de su familia fuma?
 SI () NO ()
 Cuántos: _____
17. ¿Algún miembro de su familia ingiere licor? SI () NO ()
 Cuántos: _____
18. ¿Cuántas veces al mes ingiere licor?
 (0) (1) (2) (3) (4) (+ de 4)
19. ¿Alguien de su familia ha fallecido en el presente año que residió en su hogar?
 SI () NO ()
 Edad: _____
 Causa:

20. ¿En el presente año ha nacido alguien en su familia?
 SI () NO ()
 Especifique el sexo: F () M ()
21. ¿Qué método de planificación familiar usa?
 Anticonceptivos:
 Orales () Locales ()
 Intramusculares ()
 Quirúrgicos
 Vasectomía ()
 Ligadura de trompas ()
 Método Natural:
 Billing () Regla ()
 Ninguno ()
22. ¿Hay gestantes en su hogar?
 SI () NO ()
 Si su respuesta es SI:
 a) Tiene control de embarazo
 SI () NO ()
 b) Estavacunada contra el tétano.
 SI () NO ()
 c) ¿Dónde se atendió el último parto?
 Hospital ()
 Domicilio ()
 Otros22 _____
 Especifique
23. ¿Están vacunados sus hijos?
 SI () NO ()
 Si están vacunados, contra qué?
 Sarampión ()
 Tosferina ()
 TBC ()
 Tétano ()
 Polio ()
 Otros23 () Verificar Tarjeta

NUTRICION

24. ¿Qué alimentos consumen durante la semana?

- Pescado () Frutas ()
 Carne () Mantequilla ()
 Pollo () Pan ()

25. ¿Cuántas veces al día se alimentan?

Nº: _____

- D A L C
 () () () ()

26. ¿Qué medio utiliza para la cocción de sus alimentos?

- Leña () Carbón ()
 Kerosene () Gas ()
 Electricidad ()
 Otros26 _____

V. ASÉCTOS SOCIOECONÓMICOS

27. ¿Pertenece a alguna asociación de base? SI () NO ()

- Club de Madres ()
 Vaso de Leche ()
 Comedor Popular ()
 Otros27 _____

 Especifique

28. ¿Qué tipo de información escrita prefiere?

- Libro ()
 Revista ()
 Periódicos ()
 Otros28 _____

29. ¿Qué tipo de música prefiere?

- Chicha () Salsa ()
 Rock ()
 Otros29: _____

30. ¿Cuántas personas trabajan en su familia?

Nº: _____

31. ¿Qué artefactos eléctricos posee?

- TV ()
 Video ()
 Radio ()
 Cocina a Gas ()
 Refrigerador ()
 Otros31 _____

32. ¿Cada qué tiempo le pagan?

Diario () Quincenal ()
Semanal () Mensual ()

¿Cuál es ingreso económico familiar?

33. S/.

< 250 ()
250 - 500 ()
500 - 1000 ()
1000 - 3000 ()
> 3000 ()

34. ¿Cuánto gasta en alimentos mensualmente?

S/. _____

35. ¿En qué tipo de colegio estudian sus hijos?

Nacional ()

Particular ()

36. ¿Cuántos de sus hijos terminaron de estudiar el presente año?

¿Cuántos?

Inicial () _____

Primaria () _____

Secundaria () _____

Superior () _____

37. ¿El medio de transporte que utiliza diariamente para movilizarse es?

PROPIO () NO PROPIO ()

A PARTIR DE LOS DATOS RECOLECTADOS:

- Definir las variables y grabar el archivo como: XXXXXXXX (Escriba un nombre máximo con 8 caracteres). Luego realice el análisis de los datos y los gráficos que le faciliten responder las siguientes preguntas:

1. Frecuencia de tipo de vivienda
2. Promedio de material de construcción de las viviendas
3. Número de habitaciones promedio
4. Número de dormitorios / casa
5. Número de baños / casa
6. Promedio del número de personas que duermen por dormitorio
7. Promedio de número de hijos por familia
8. Promedio de jardines por casa
9. Frecuencia de abastecimiento de agua

10. Frecuencia de almacenamiento de agua
11. Frecuencia de viviendas con energía eléctrica
12. Frecuencia de eliminación de excretas
13. Frecuencia de formas para desechar la basura
14. Frecuencia de animales domésticos en las viviendas
15. Frecuencia de animales transmisores de enfermedades en las viviendas
16. Frecuencia de enfermedades de síntomas y signos
17. Frecuencia de asistencia de las personas al enfermarse
18. Frecuencia de personas que fuman
19. Frecuencia de personas que ingieren licor
20. Frecuencia al mes de personas que consumen licor
21. Frecuencia de personas que han fallecido
22. Tabla de nacimiento según sexo
23. Frecuencia de método de planificación según anticonceptivos, quirúrgicos y método natural
24. Frecuencia de gestantes en su hogar
25. Frecuencia de personas vacunadas
26. Frecuencia de consumo de alimentos
27. El medio más común de cocción de los alimentos
28. Institución a la que pertenecen
29. El tipo de información escrita de mayor preferencia
30. La música de mayor preferencia
31. Número de personas que trabajan
32. El artefacto eléctrico más común
33. El tiempo de pago
34. El ingreso económico promedio
35. El promedio de gastos en alimentos
36. Tipo de colegio más común en la cual estudian los hijos
37. Número de Hijos que concluyeron estudios
38. El medio de transporte más común

UNIVERSIDAD PRIVADA SAN PEDRO
FACULTAD DE MEDICINA HUMANA – ESCUELA DE MEDICINA

**EVALUACIÓN DEL LOGRO DE LAS COMPETENCIAS BÁSICAS
 EN EL MANEJO DE SOFTWARE ESTADÍSTICO”**

Alumno(a): Código:

PC N°: Ciclo: I I Grupo: Semestre:

Instrucciones: Esta evaluación para el manejo de un paquete informático, sólo tiene dos dominios de evaluación: **Si** cuando se tenga el manejo o dominio de la habilidad (1 punto cada **Si**) y **No** cuando el manejo sea deficiente o nulo (0 puntos)

I. Crea variables adecuadas para un cuestionario ya aplicado	Sí	No
1. Define en el computador las variables del cuestionario según tipo, formato de columna, etiqueta, valores perdidos o datos en blanco.		
2. Describe la variable en los casos en que el nombre de la variable no sea autoexplicativo o se requiera una explicación más detallada.		
II. Introduce y reingresa datos de todo tipo como paso previo para realizar un análisis de los datos.		
3. Registra los datos por filas o columnas		
4. Toma decisiones para registrar o no los datos provenientes de cuestionarios que sean dudosos o ilógicos, utilizando el valor definido en la variable definida para estos casos.		
5. Evita errores al registrar datos		
6. Une archivos añadiendo nuevos casos o nuevas variables		
III. Realiza el análisis de los datos.		
Creando y transformando las variables, mediante la:		
a) Recodificación de variables		
b) Creación de nuevas o transformando las ya existentes utilizando operadores diversos		
c) Creación de nuevas variables o de nuevos valores a variables existentes si una determinada condición se cumple		
8. Segmenta archivos para analizar datos de grupos por separado		
9. Selecciona casos o una parte de los sujetos para realizar el análisis		

IV. Examina las variables de una manera bastante exhaustiva		
10. Describe variables de tipo: a) Nominales: Moda b) Ordinales: mediana, rango, máximo, mínimo c) De Intervalo: Histograma, media, desviación std, rango, coeficientes		
V. Relaciona variables con índices estadísticos mediante:		
11. Tablas de contingencia		
12. Coeficiente de correlación de Pearson entre dos variables cualitativas		
13. Una representación gráfica (nube de puntos) de la relación entre dos variables editando la gráfica para hallar la recta de regresión y sus parámetros (constante y pendiente)		
VI. Obtiene y edita Gráficas de diferentes tipos que representan adecuadamente los datos y facilitan su interpretación.		
14. De barras		
15. De líneas		
16. De área y diferencias de medias		
17. De Sectores		
18. De Pareto		
19. De Cajas		
20. De barras de error		