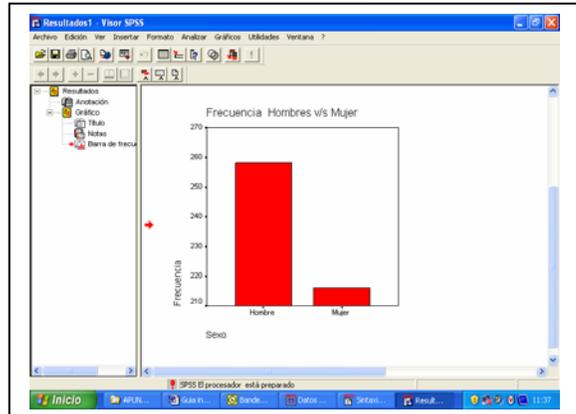
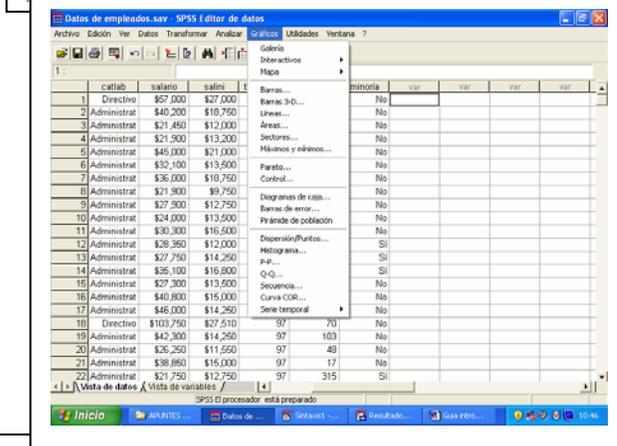
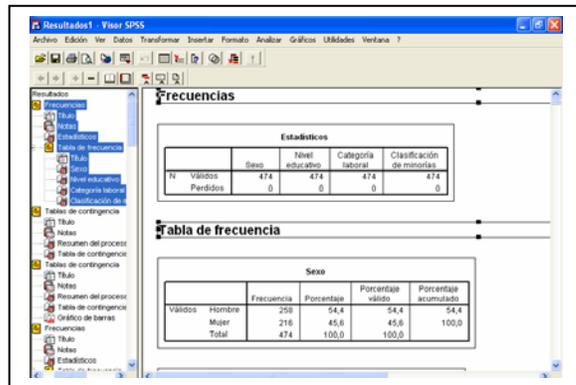
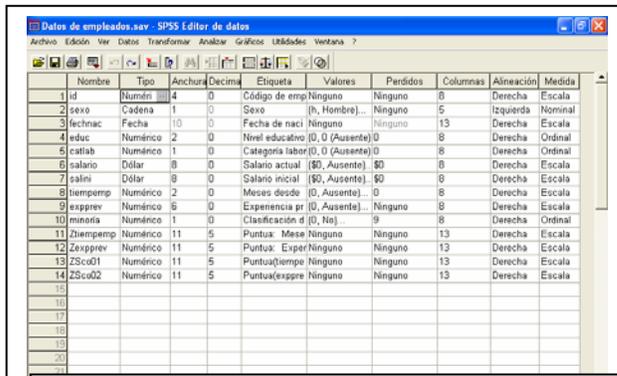
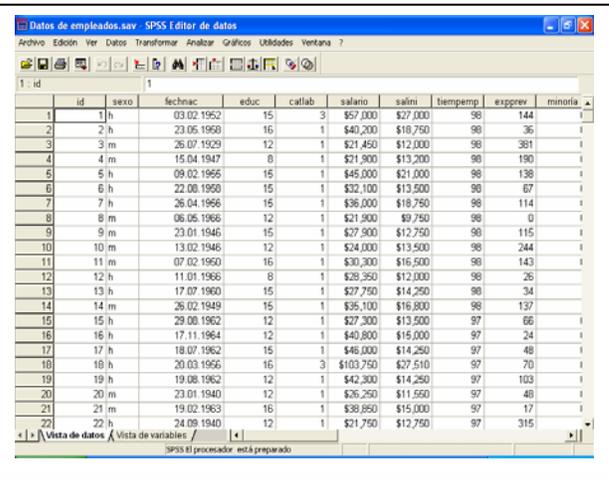
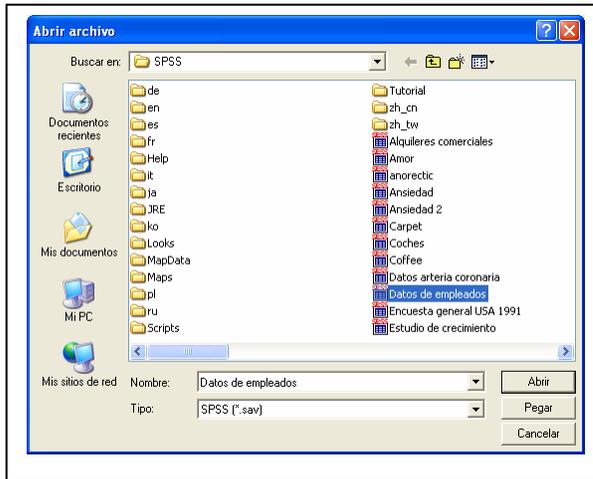


GUIA Introducción al SPSS¹

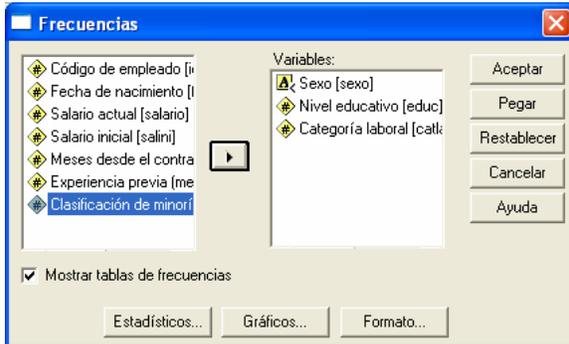
Abrir desde la carpeta SPSS el archivo "Datos de empleados.sav"
Archivo / Abrir / Datos



¹ SPSS (Statistical Package for the Social Sciences) es un programa estadístico muy usado en las Ciencias Sociales y Empresas de Investigación de Mercado

Tablas de frecuencias

Analizar/Estadísticos Descriptivos/Frecuencias



	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
Hombre	258	54,4	54,4	54,4
Mujer	216	45,6	45,6	100,0
Total	474	100,0	100,0	

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	8	53	11,2	11,2
12	190	11,2	40,1	51,3
14	6	1,3	1,3	52,5
15	116	24,5	24,5	77,0
16	59	12,4	12,4	89,5
17	11	2,3	2,3	91,8
18	9	1,9	1,9	93,7
19	27	5,7	5,7	99,4
20	2	,4	,4	99,8
21	1	,2	,2	100,0
Total	474	100,0	100,0	

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
Administrativo	363	76,6	76,6	76,6
Seguridad	27	5,7	5,7	82,3
Directivo	84	17,7	17,7	100,0
Total	474	100,0	100,0	

El botón **Restablecer** permite limpiar la pantalla del cuadro de diálogo

El botón **pegar** permite guardar la sintaxis del procedimiento para ser ejecutada en cualquier momento.

FRECUENCIES

VARIABLES=sexo educ catlab

/ORDER= ANALYSIS .

Tablas de contingencia

Analizar/Estadísticos descriptivos/Tablas de contingencia

(Recomendable para variables nominales y ordinales)



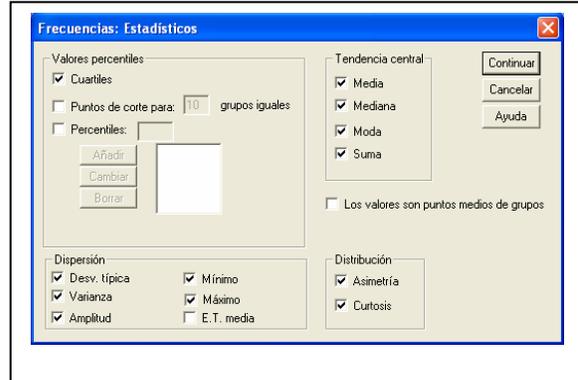
Sexo	Hombre	Recuento	Categoría laboral			Total
			Administrativo	Seguridad	Directivo	
Hombre	Recuento	157	27	74	258	
	% de Sexo	60,9%	10,5%	28,7%	100,0%	
	% de Categoría laboral	43,3%	100,0%	88,1%	54,4%	
Mujer	Recuento	206	0	10	216	
	% de Sexo	95,4%	,0%	4,6%	100,0%	
	% de Categoría laboral	56,7%	,0%	11,9%	45,6%	
Total	Recuento	363	27	84	474	
	% de Sexo	76,6%	5,7%	17,7%	100,0%	
	% de Categoría laboral	100,0%	100,0%	100,0%	100,0%	



Clasificación de minorías	Sexo	Hombre	Recuento	Categoría laboral			Total
				Administrativo	Seguridad	Directivo	
No	Hombre	Recuento	110	14	70	194	
		% de Sexo	56,7%	7,2%	36,1%	100,0%	
		% de Categoría laboral	38,9%	100,0%	87,5%	52,4%	
	Mujer	Recuento	166	0	10	176	
		% de Sexo	94,3%	,0%	5,7%	100,0%	
		% de Categoría laboral	60,1%	,0%	12,5%	47,6%	
	Total	Recuento	276	14	80	370	
		% de Sexo	74,6%	3,8%	21,6%	100,0%	
		% de Categoría laboral	100,0%	100,0%	100,0%	100,0%	
Sí	Hombre	Recuento	47	13	4	64	
		% de Sexo	73,4%	20,3%	6,3%	100,0%	
		% de Categoría laboral	54,0%	100,0%	3,8%	61,5%	
	Mujer	Recuento	40	0	0	40	
		% de Sexo	100,0%	,0%	,0%	100,0%	
		% de Categoría laboral	46,0%	,0%	,0%	38,5%	
	Total	Recuento	87	13	4	104	
		% de Sexo	83,7%	12,5%	3,8%	100,0%	
		% de Categoría laboral	100,0%	100,0%	100,0%	100,0%	

Descriptivos (Medidas de tendencia central y medidas de dispersión) Analizar/Estadísticos descriptivos/Frecuencias

Variables: Salario, Salini
Estadísticos

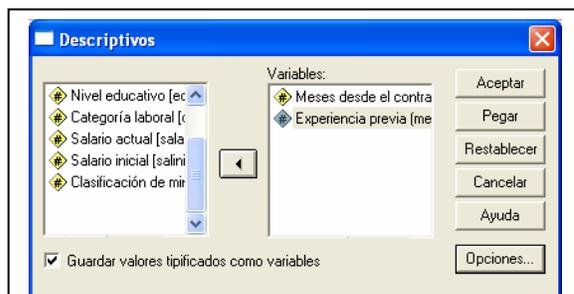


Estadísticos			
		Salario actual	Salario inicial
N	Válidos	474	474
	Perdidos	0	0
Media		\$34,419.57	\$17,016.09
Mediana		\$28,875.00	\$15,000.00
Moda		\$30,750	\$15,000
Desv. típ.		\$17,075.661	\$7,870.638
Varianza		291578214.5	61946944.96
Asimetría		2,125	2,853
Error típ. de asimetría		,112	,112
Curtosis		5,378	12,390
Error típ. de curtosis		,224	,224
Rango		\$119,250	\$70,980
Mínimo		\$15,750	\$9,000
Máximo		\$135,000	\$79,980
Suma		\$16,314,875	\$8,065,625
Percentiles	25	\$24,000.00	\$12,450.00
	50	\$28,875.00	\$15,000.00
	75	\$37,162.50	\$17,617.50

Descriptivos (Medidas de tendencia central y medidas de dispersión) Analizar/Estadísticos descriptivos/descriptivos

Variables: tiempemp, exprev
Opciones

En este menú si se selecciona guardar valores tipificados como variables se obtiene las puntuaciones z



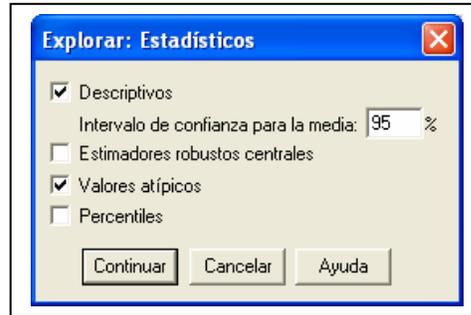
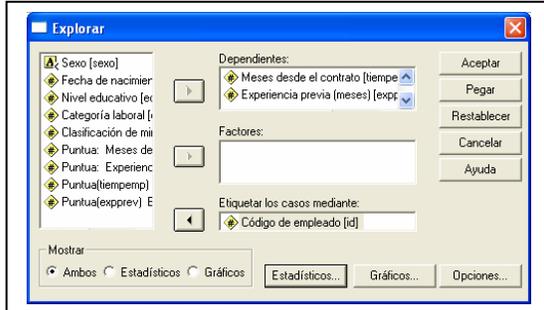
Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desv. tip.
Meses desde el contrato	474	63	98	81,11	10,061
Experiencia previa (meses)	474	0	476	95,86	104,586
N válido (según lista)	474				



Descriptivos (Descriptivos, Valores extremos, intervalos de confianza)

Analizar/Estadísticos descriptivos/explorar

Variables: salario, salini



Para mostrar la tabla de resultados se puede editar la tabla y seleccionar Pivotar/Paneles de pivotado.

Salario actual		Estadístico	Error tip.
Media		\$34,419.57	\$784.311
Intervalo de confianza para la media al 95%	Limite inferior	\$32,878.40	
	Limite superior	\$35,960.73	
Media recortada al 5%		\$32,455.19	
Mediana		\$28,875.00	
Varianza		291578214.453	
Desv. tip.		\$17,075.661	
Mínimo		\$15,750	
Máximo		\$135,000	
Rango		\$119,250	
Amplitud intercuartil		\$13,163	
Asimetría		2,125	,112
Curiosis		5,378	,224

Salario actual		Número del caso	Código de empleado	Valor
Mayores	1	29	29	\$135,000
	2	32	32	\$110,625
	3	18	18	\$103,750
	4	343	343	\$103,500
	5	446	446	\$100,000
Menores	1	378	378	\$15,750
	2	338	338	\$15,900
	3	411	411	\$16,200
	4	224	224	\$16,200
	5	90	90	\$16,200

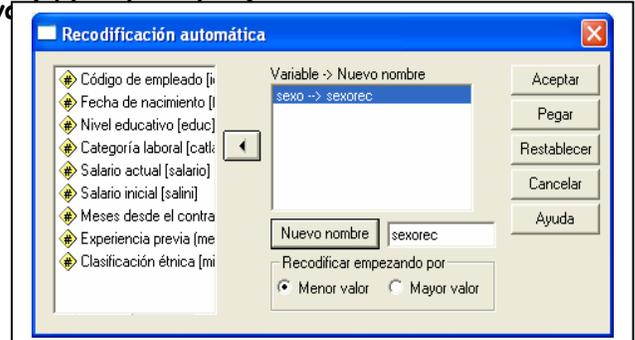
Si se agrega factores, los resultados se detallan por grupos



Salario actual		Estadístico	Error tip.	
Hombre	Media	\$41,441.78		
	Intervalo de confianza para la media al 95%	Limite inferior	\$39,051.19	
		Limite superior	\$43,832.37	
	Media recortada al 5%	\$39,445.87		
	Mediana	\$32,850.00		
	Varianza	380219336		
	Desv. tip.	\$19,499.214		
	Mínimo	\$19,650		
	Máximo	\$135,000		
	Rango	\$115,350		
	Amplitud intercuartil	\$22,675		
	Asimetría	1,639	,152	
	Curiosis	2,780	,302	
Mujer	Media	\$26,031.92		
	Intervalo de confianza para la media al 95%	Limite inferior	\$25,018.29	
		Limite superior	\$27,045.55	
	Media recortada al 5%	\$25,248.30		
	Mediana	\$24,300.00		
	Varianza	57123685.3		
	Desv. tip.	\$7,558.021		
	Mínimo	\$15,750		
	Máximo	\$58,125		
	Rango	\$42,375		
	Amplitud intercuartil	\$7,013		
	Asimetría	1,863	,166	
	Curiosis	4,641	,330	

Recodificar automáticamente (permite codificar variables)

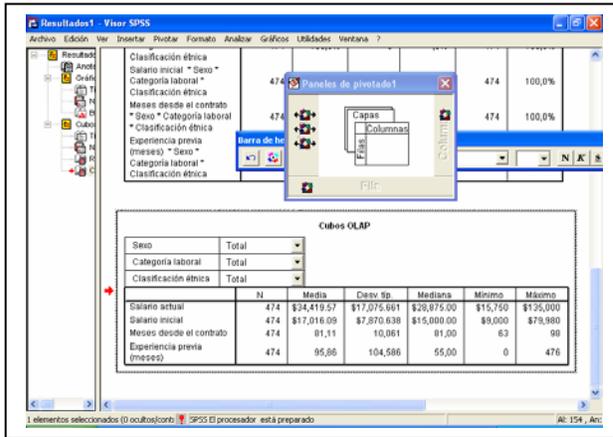
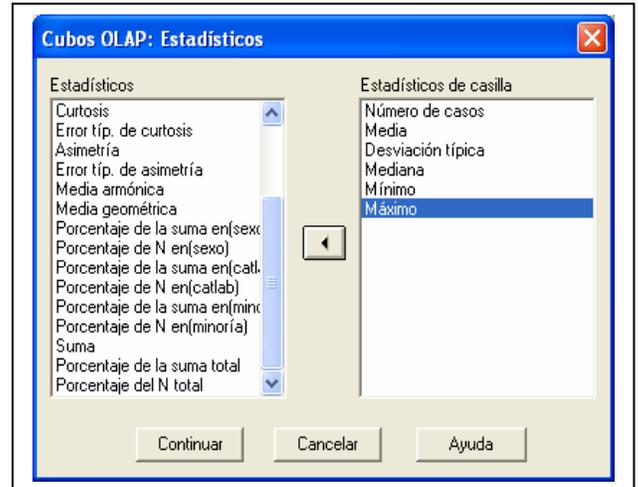
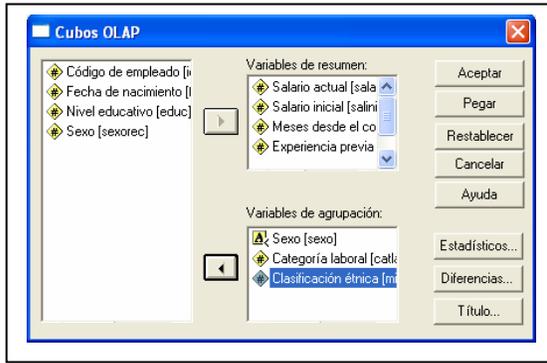
SEXO	SEXOREC	Sexo	Value Label
Old Value	New Value		
h	1	Hombre	
m	2	Mujer	



Informes Cubos OLAP (Descriptivos por grupos)

Analizar/Informes/Cubos OLAP

Se necesitan variables de escala y variables de Agrupación (ordinales y/o nominales)



Descriptivos para hombres administrativos

Sexo: Hombre
Categoría laboral: Administrativo
Clasificación étnica: Total

	N	Media	Desv. tip.	Mediana	Mínimo	Máximo
Salario actual	157	\$31,558.15	\$7,997.978	\$29,850.00	\$19,650	\$80,000
Salario inicial	157	\$15,861.21	\$2,564.694	\$15,750.00	\$9,750	\$31,980
Meses desde el contrato	157	81,73	10,670	82,00	63	98
Experiencia previa (meses)	157	94,22	94,429	58,00	5	476

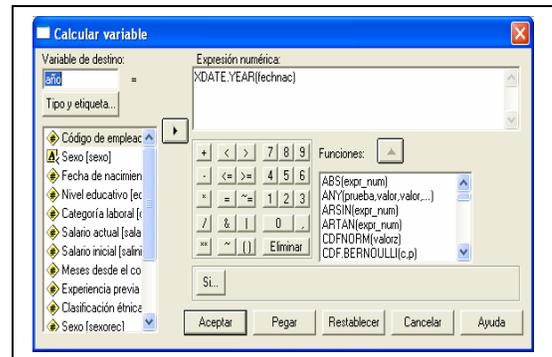
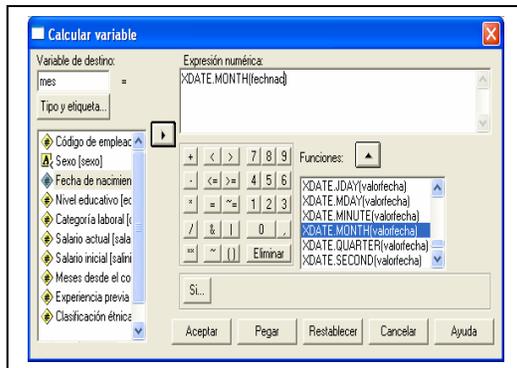
Descriptivos para mujeres administrativas

Sexo: Mujer
Categoría laboral: Administrativo
Clasificación étnica: Total

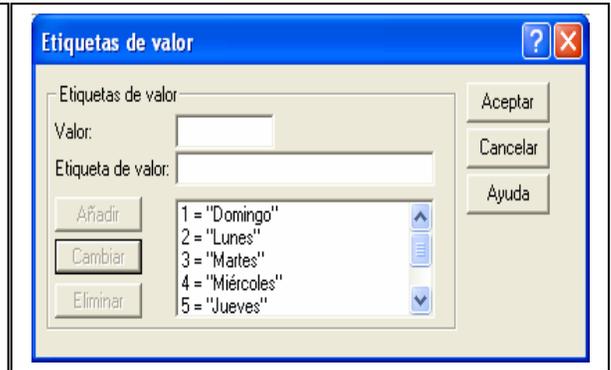
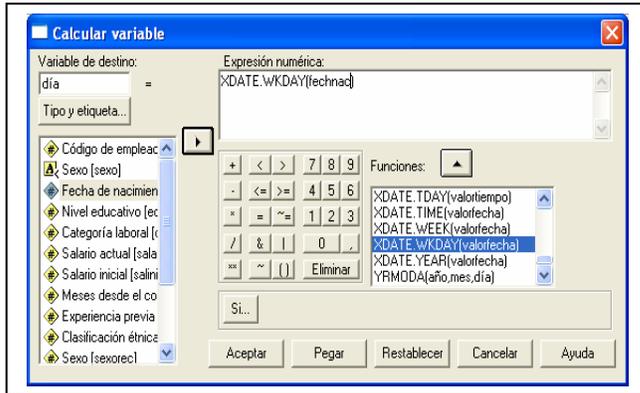
	N	Media	Desv. tip.	Mediana	Mínimo	Máximo
Salario actual	206	\$25,003.69	\$5,812.838	\$24,000.00	\$15,750	\$54,000
Salario inicial	206	\$12,750.75	\$2,391.056	\$12,000.00	\$9,000	\$19,980
Meses desde el contrato	206	80,56	9,658	81,00	63	98
Experiencia previa (meses)	206	78,04	95,550	33,00	0	412

Calcular (Crear nuevas variables)

Transformar/Calcular



XDATE.WKDAY(valorfecha). Numérico. Devuelve el número del día de la semana (un entero entre 1, Domingo, y 7, Sábado) a partir de un valor numérico en el formato de fecha de SPSS, tal como el creado por las funciones DATE.xxx o el leído por los formatos de entrada DATE.



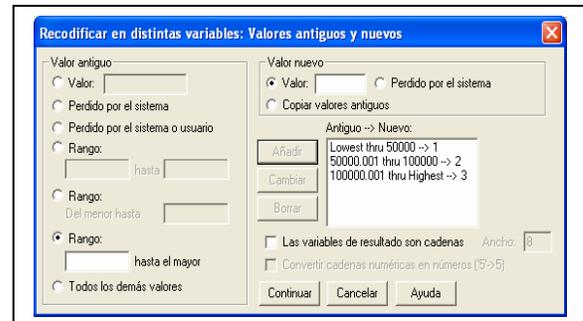
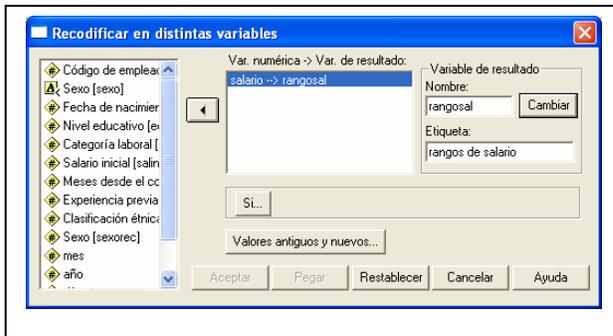
Transformar Recodificar (A partir de variables de escala reemplazar valores y crear rangos de valores)

Recodificar /en distinta variable

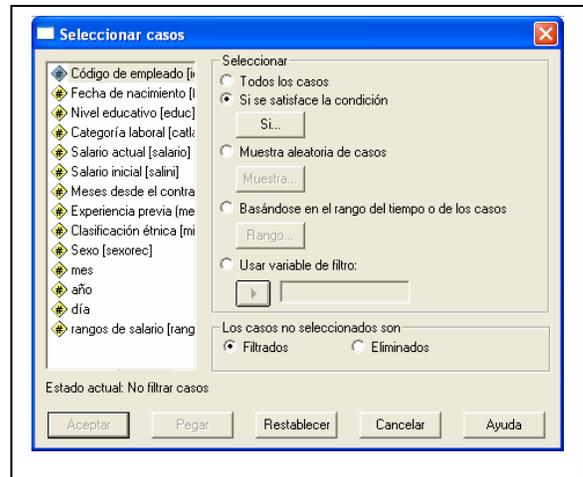
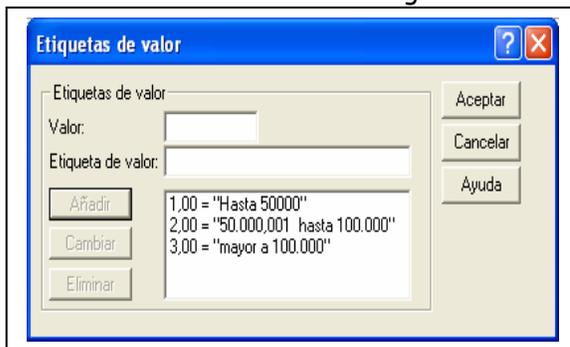
Variable de entrada: salario

Var de resultado: rangosal / Cambiar

Luego entrar a valores antiguos y nuevos



En vista de variables/ variable rangosal/ valores



Estudio de Caso: Caracterización del Mundo²

Considere el archivo Mundo 95, que contiene las siguientes variables de los países del Mundo en el año 1995:

Variable	Etiqueta	Etiqueta de Valor
país	País	
poblac	Población x 1000	
densidad	Habitantes x Km2	
urbana	Habitantes en ciudades (%)	
relig	Religión mayoritaria	
espvidaf	Esperanza de vida Femenina	
espvidam	Esperanza de vida Masculina	
alfabet	Alfabetización (%)	
inc_pob	Aumento de población (% anual)	
mortinf	Mortalidad infantil (Muertes por 1000 nacimientos vivos)	
pib_cap	Producto interno bruto per cápita	
región	Región Económica	1 = OCDE 2 = Europa Oriental 3 = Asia / Pacífico 4 = Africa 5 = Oriente Medio 6 = América Latina
calorías	Ingesta diaria de calorías	
sida	Casos de SIDA	
tasa_nata	Tasa de natalidad (por 1.000 habitantes)	
tasa_mort	Tasa de mortalidad (por 1.000 habitantes)	
tasasida	Casos de SIDA por 100.000 habitantes	
log_pib	Log(10) de PIB_CAP	

² Caso desarrollado por Sara Arancibia

logtsida	Log(10) de TASASIDA	
nac_def	Tasa nacimientos/defunciones	
fertilid	Número promedio de hijos	
log_pob	Log(10) de POBLAC	
cregran o	--	
alfabma s	Hombres alfabetizados (%)	
alfabfe m	Mujeres alfabetizadas (%)	
clima	Clima predominante	1 = Desierto 2 = Arido / Desierto 3 = Arido 5 = Tropical 6 = Mediterráneo 7 = Marítimo 8 = Templado 9 = Artico / Templado 10 = Artico

Usted debe realizar un informe donde compare los países en al menos los siguientes aspectos: Población, densidad, % de habitantes en ciudades, esperanza de vida, alfabetización (%), tasas de natalidad y mortalidad, número promedio de hijos por familia, tasa sida, considerando las variables nominales Región, Religión mayoritaria y clima predominante.

Para su informe debe considerar al menos los siguientes puntos:

- (i) Tres gráficos distintos con su interpretación.
- (ii) Tablas de frecuencia
- (iii) Tablas de contingencia
- (iv) Outliers (Valores extremos)
- (v) Medidas de tendencia central
- (vi) Medidas de dispersión
- (vii) Cubos OLAP
- (viii) Puntuaciones z

Solución:

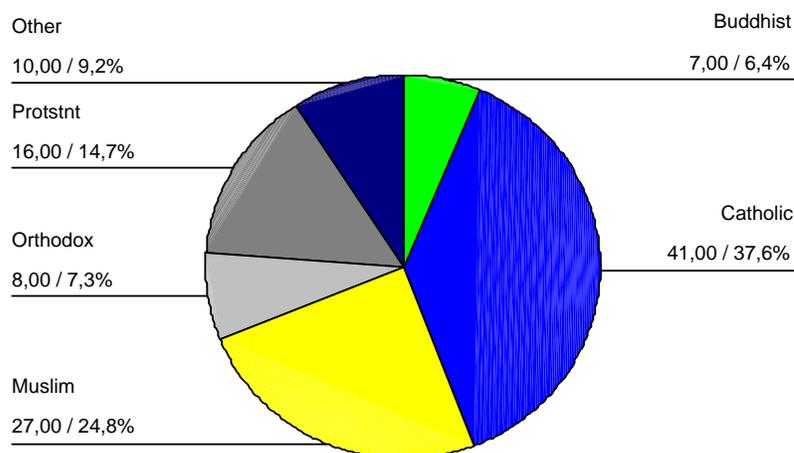
Comenzaremos el estudio determinando la frecuencia de las variables nominales; Región Económica, Religión Predominante y Clima Predominante de los países

Region or economic group				
		Frequency	Percent	Valid Percent
Valid	OECD	21	19,3	19,3
	East Europe	14	12,8	12,8
	Pacific/Asia	17	15,6	15,6
	Africa	19	17,4	17,4
	Middle East	17	15,6	15,6
	Latin America	21	19,3	19,3
	Total	109	100,0	100,0

La tabla de frecuencia muestra el número de países por Región económica. Se observan dos regiones con la mayor frecuencia, (21 países) las que corresponden a la Región OECD (Organización para la Cooperación y el Desarrollo Económico) y a la Región de Latino América, correspondiendo al 19,3% del total de países. La menor frecuencia se observa en Europa del Este con 14 países de un total de 109 países.

El gráfico siguiente muestra la frecuencia y porcentaje de países por Religión predominante.

Frecuencia y porcentaje de países
por Religión Predominante



Se observa que 41 países que representan el 37,6% del total de países considerados tienen como religión predominante a la religión Católica y 27 países a la religión Musulmana representando el 24,8% del total de países considerados.

Para generar el gráfico: Graficar/Sectores/Resumen para grupos de casos/Nº de casos/Religión Predominante. En el editor de gráficos se pide texto, valor y porcentaje y se colapsa los sectores a mayores del 5%.

La tabla de frecuencia para religión predominante muestra complementariamente al gráfico anterior que las religiones con menor frecuencia son las religiones Hindú, Judía, Taoísta y Tribal

Predominant religion

	Frequency	Percent	Valid Percent
Valid			
Animist	4	3,7	3,7
Buddhist	7	6,4	6,5
Catholic	41	37,6	38,0
Hindu	1	,9	,9
Jewish	1	,9	,9
Muslim	27	24,8	25,0
Orthodox	8	7,3	7,4
Protstnt	16	14,7	14,8
Taoist	2	1,8	1,9
Tribal	1	,9	,9
Total	108	99,1	100,0
Missing	1	,9	
Total	109	100,0	

Al cruzar las variables región y religión podemos observar en la tabla de contingencia que la Religión Predominante Animista pertenece a países de África. La religión predominante Católica se encuentra en todas las regiones excepto en la Región de Oriente donde la religión predominante es la Musulmana con 15 países de un total de 17 países de la región

Predominant religion * Region or economic group Crosstabulation

Count		Region or economic group						Total
		OECD	East Europe	Pacific/Asia	Africa	Middle East	Latn America	
Predominant religion	Animist				4			4
	Buddhist			7				7
	Catholic	10	5	1	5		20	41
	Hindu			1				1
	Jewish					1		1
	Muslim		1	5	6	15		27
	Orthodox	1	6			1		8
	Protstnt	10	2	1	2		1	16
	Taoist			2				2
	Tribal				1			1
Total		21	14	17	18	17	21	108

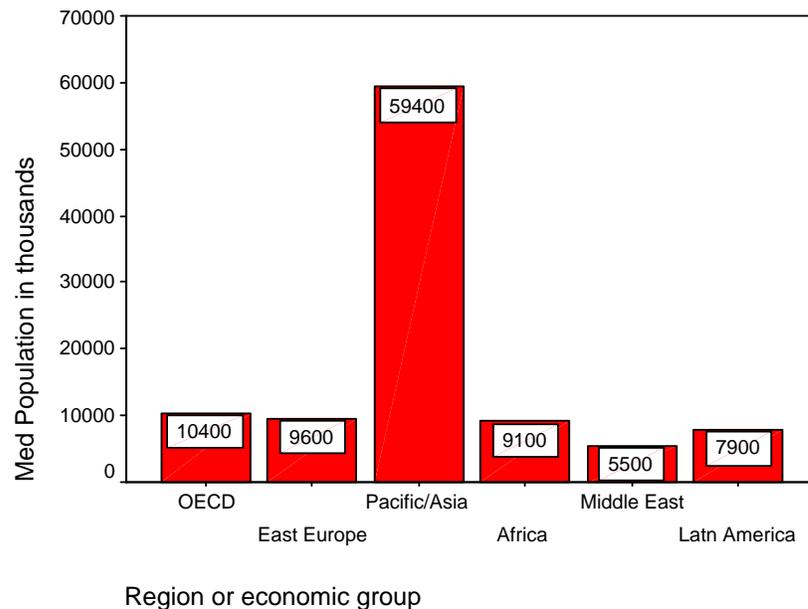
En relación al Clima Predominante se observa que las mayores frecuencias corresponden a los climas Templado y Tropical los que representan un 31,8% y 29,9% respectivamente, respecto al total de datos válidos.

Predominant climate

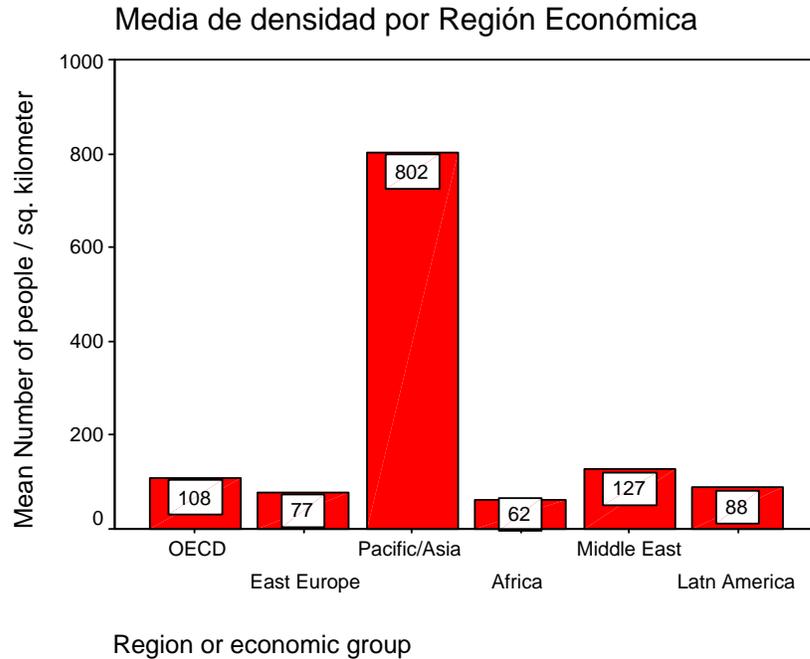
		Frequency	Percent	Valid Percent
Valid	desert	7	6,4	6,5
	arid / desert	5	4,6	4,7
	arid	6	5,5	5,6
	otro	5	4,6	4,7
	tropical	32	29,4	29,9
	mediterranean	10	9,2	9,3
	maritime	4	3,7	3,7
	temperate	34	31,2	31,8
	arctic / temp	4	3,7	3,7
	Total	107	98,2	100,0
Missing	System	2	1,8	
Total		109	100,0	

Ahora consideremos la población, densidad y habitantes que viven en ciudades. Podemos observar del gráfico correspondiente a la mediana de población por región económica que el 50% de los países del Asia/Pacífico tienen una población mayor o igual a 59.400.000 habitantes, valor notablemente alto en relación a las medianas del resto de las regiones las que oscilan entre 10.400.000 y 5.500.000 habitantes.

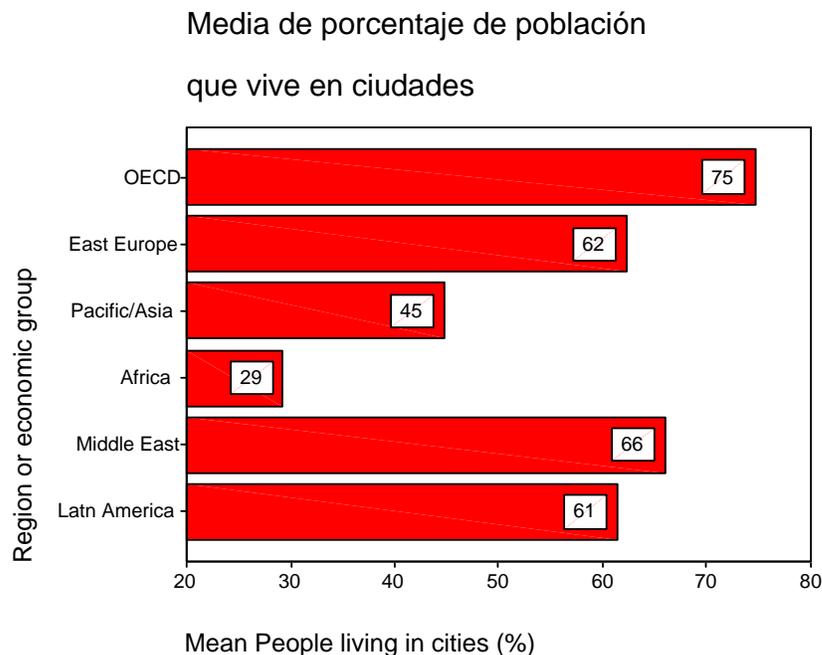
Mediana de Población por Región Económica



Coherente con lo anterior se observa que la mayor densidad por región económica corresponde a la región Asia/Pacífico con un valor promedio de 802 habitantes por km^2 , la que es considerablemente superior a la densidad promedio del resto de regiones, las que oscilan entre 127 y 62 habitantes por km^2 correspondiendo esta última a la región de África



Para generar el gráfico: *Graficar/Barras/Simples/Resumen para grupos de casos/N de casos/ Región Económica.*



En relación al porcentaje de personas que viven en ciudades, se observa del gráfico que el mayor porcentaje promedio corresponde a la Región OECD, con un 75% en promedio. Es considerable la diferencia con la región de África donde el promedio de población urbana es del 29%, seguido de Asia/Pacífico con un promedio del 45%.

Podemos complementar la información anterior con Cubos OLAP, los que muestran por grupos, los estadísticos que se necesiten conocer. Específicamente los Cubos siguientes muestran para las regiones OECD y África, el número de países el que corresponde a 21 y 18 países respectivamente. Se observa la media para cada una de las variables consideradas y la desviación estándar que muestra cuánto se desvían los datos, en promedio respecto a la media.

Al considerar el cubo correspondiente a la región OECD se observan los valores mínimo y máximo, es sorprendente observar que existen países con una densidad de 2,3 habitantes por km² y de 366 personas por km². Al considerar la población, dentro de los países del OECD se puede apreciar un valor mínimo de 263.000 habitantes en oposición al valor máximo de 260.800.000 habitantes. El mayor porcentaje de población urbana corresponde al 96% y el menor corresponde al 34%.

OLAP Cubes

Region or economic group: OECD
 Predominant climate: Total
 Predominant religion: Total

	Population in thousands	Number of people / sq. kilometer	People living in cities (%)
N	21	21	21
Mean	33085,10	107,981	74,71
Std. Deviation	57148,25	107,936	14,89
Minimum	263	2,3	34
Maximum	260800	366,0	96
Median	10400,00	80,000	77,00

OLAP Cubes

Region or economic group: Africa
 Predominant climate: Total
 Predominant religion: Total

	Population in thousands	Number of people / sq. kilometer	People living in cities (%)
N	18	18	18
Mean	18415,83	63,700	28,17
Std. Deviation	24331,33	79,823	14,70
Minimum	959	2,4	5
Maximum	98100	311,0	47
Median	8900,00	39,500	24,50

Al considerar el cubo correspondiente a la región de África se observa una media de población considerablemente más baja que la media de la Región OECD y que la variabilidad en la variable población del 132% es más baja que si se compara con la región del OECD cuyo coeficiente de variabilidad es del 172,7%. Por otra parte se observa para la población urbana un mínimo de 5% siendo el porcentaje máximo del 47%, valores muy bajos si se compara con la región del OECD. Al igual que la región OECD se observa un valor mínimo de densidad de

2,4 habitantes por km², en oposición al máximo cuya densidad es de 311 habitantes por km².

Para identificar a qué países corresponden estos valores máximos y mínimos se puede solicitar los valores extremos (outliers) que muestra los cinco valores mayores y menores.

Extreme Values

Region or economic group: OECD
Number of people / sq. kilometer

	Case Number	COUNTRY	Value	
Highest	1	70	Netherlands	366,0
	2	11	Belgium	329,0
	3	101	UK	237,0
	4	42	Germany	227,0
	5	56	Italy	188,0
Lowest	1	4	Australia	2,3
	2	49	Iceland	2,5
	3	21	Canada	2,8
	4	74	Norway	11,0
	5	71	New Zealand	13,0

Extreme Values

Region or economic group: Africa
Number of people / sq. kilometer

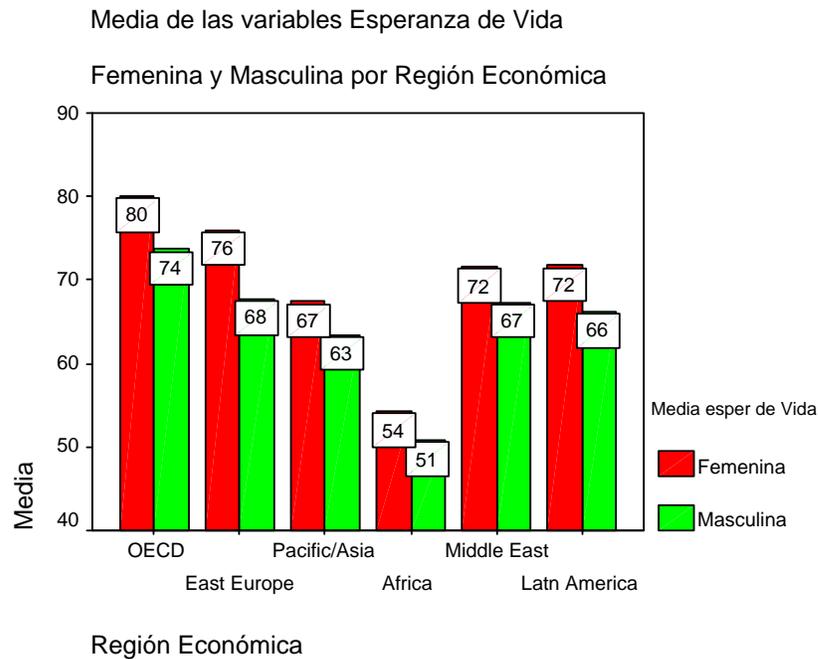
	Case Number	COUNTRY	Value	
Highest	1	85	Rwanda	311,0
	2	18	Burundi	216,0
	3	73	Nigeria	102,0
	4	40	Gambia	86,0
	5	103	Uganda	76,0
Lowest	1	14	Botswana	2,4
	2	39	Gabon	4,2
	3	22	Cent. Afri.R	5,0
	4	90	Somalia	10,0
	5	109	Zambia	11,0

Para generar las tablas: Analizar/Estadísticos Descriptivos/Explorar. En Factor colocar Región económica y etiquetar por país. En Estadísticos seleccionar Valores Atípicos.

Ahora consideraremos las variables: Esperanza de vida femenina, esperanza de vida masculina, tasa de natalidad, tasa de mortalidad, tasa sida, fertilidad y alfabetización

El gráfico siguiente muestra la media de esperanza de vida femenina y masculina por Región Económica. Se observa que en todas las regiones es mayor la media de esperanza de vida femenina que masculina siendo la región del OECD, la de mayor esperanza de vida, con un promedio de 80 y 74 años para mujeres y hombres respectivamente. Es notable la diferencia con África donde se observa que el

promedio de esperanza de vida es muy baja siendo la media de 54 y 51 años para mujeres y hombres respectivamente.



Para generar el gráfico: Graficar/Barras/Agrupados. Resumen para variables individuales/Media de las variables Esperanza de vida fem y masculina/eje de categorías Región Económica.

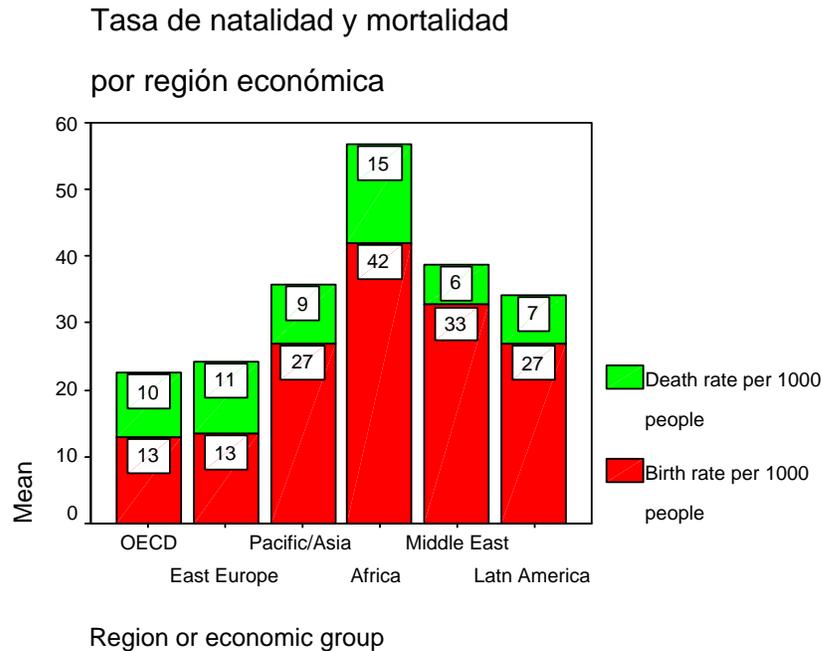
La tabla siguiente identifica los países con mayor y menor esperanza de vida

Valores Extremos (Outliers considerando todos los países)					
		Case Number	COUNTRY	Value	
Average female life expectancy	Highest	1	94	Switzerland	82
		2	57	Japan	82
		3	38	France	82
		4	21	Canada	81
		5	56	Italy	, ^a
	Lowest	1	103	Uganda	43
		2	1	Afghanistan	44
		3	22	Cent. Afri.R	44
		4	109	Zambia	45
		5	97	Tanzania	45
Average male life expectancy	Highest	1	55	Israel	76
		2	57	Japan	76
		3	26	Costa Rica	76
		4	49	Iceland	76
		5	47	Hong Kong	, ^b
	Lowest	1	103	Uganda	41
		2	97	Tanzania	41
		3	22	Cent. Afri.R	41
		4	85	Rwanda	43
		5	45	Haiti	43

a. Only a partial list of cases with the value 81 are shown in the table of upper extremes.

b. Only a partial list of cases with the value 75 are shown in the table of upper

El siguiente gráfico apilado compara la tasa de natalidad y mortalidad por región económica, mostrando que las mayores tasas corresponden a la región de África, las que indican que en promedio nacen 42 por cada 1.000 habitantes y mueren en promedio 15 por cada 1.000 habitantes. La menor tasa de natalidad en promedio corresponde a la región del OECD

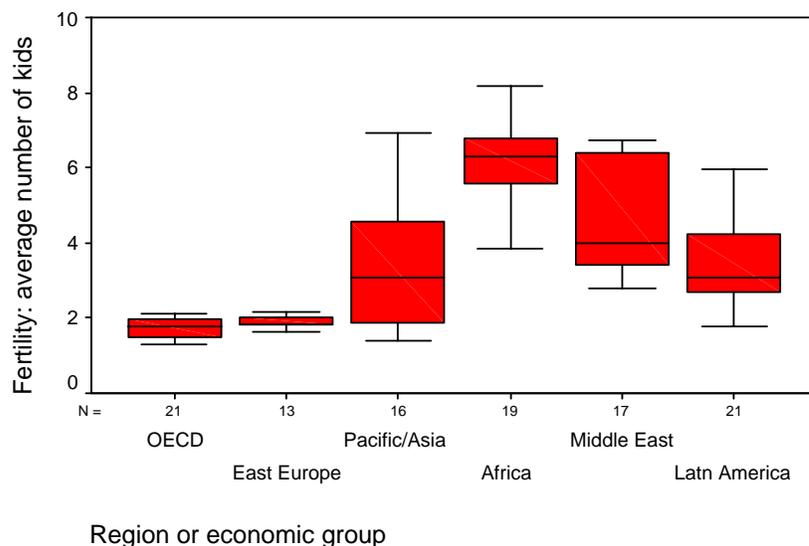


Esta información está muy de acuerdo con la variable fertilidad, la que indica el promedio de hijos por familia.

El diagrama de caja muestra por región económica que las mayores tasas de fertilidad se concentran en la Región de África mostrando que la mediana representada por la línea horizontal en las cajas se aproxima al valor 6 hijos por familia en promedio. El 50% de los datos de fertilidad de los países se encuentra en la caja la que va desde el primer cuartil al tercer cuartil. La tabla de descriptivos para fertilidad por región confirma la información entregada por el diagrama de caja.

Diagrama de caja

Fertilidad por región económica



Descriptives

Fertility: average number of kids

	Region or economic group					
	OECD	East	Pacific/As	Africa	Middle	Latn
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Mean	1,746	1,889	3,383	6,081	4,724	3,336
5% Trimmed Mean	1,751	1,886	3,298	6,088	4,721	3,280
Median	1,800	1,840	3,065	6,290	4,000	3,080
Variance	6,150E-02	1,772E-02	3,226	1,285	2,356	1,115
Std. Deviation	,248	,133	1,796	1,134	1,535	1,056
Minimum	1,3	1,7	1,4	3,8	2,8	1,8
Maximum	2,1	2,2	6,9	8,2	6,7	5,9
Range	,8	,5	5,5	4,4	3,9	4,2
Interquartile Range	,495	,190	2,795	1,380	3,165	1,655
Skewness	-,081	,530	,791	-,586	,231	,827
Kurtosis	-1,192	,932	-,563	,119	-1,812	,332

Como complemento de la información vemos que la tabla siguiente muestra las medidas de tendencia central y de dispersión para todas las variables consideradas en este apartado.

Descriptives

	Average female life	Average male life	Birth rate per 1000	Death rate per 1000	Number of aids cases /	Fertility: average number	People who read
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Mean	69,89	64,71	26,154	9,64	24,8271	3,558	77,95
5% Trimmed Mean	70,67	65,37	25,754	9,31	16,8072	3,475	79,74
Median	74,00	67,00	25,500	9,00	5,5512	3,065	87,50
Variance	115,241	88,926	154,112	18,400	2482,6	3,605	532,862
Std. Deviation	10,74	9,43	12,414	4,29	49,8252	1,899	23,08
Minimum	43	41	10,0	2	,00	1,3	18
Maximum	82	76	53,0	24	326,75	8,2	100
Range	39	35	43,0	22	326,75	6,9	82
Interquartile Range	12,75	12,75	21,000	4,00	23,2434	3,170	36,75
Skewness	-1,048	-1,020	,416	1,283	3,498	,665	-,955
Kurtosis	,054	,171	-1,163	1,754	15,008	-,933	-,250

Si consideramos sólo los países de las regiones OECD y África, observamos cómo cambian las medidas de tendencia central y dispersión ya que en todas las variables, los países de la región de África están con índice muy por debajo de los de la región OECD. Si queremos reconocer qué países en esas regiones tienen los cinco valores máximos y mínimos los podemos apreciar de la tabla de valores extremos.

Descriptives

Region or economic group: OECD

	Average female	Average male life	Birth rate per 1000	Death rate per	Number of aids	Fertility: average	People who read
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Mean	80,10	73,71	12,952	9,63	29,1052	1,746	97,67
5% Trimmed Mean	80,11	73,74	12,944	9,65	23,6322	1,751	98,22
Median	80,00	74,00	13,000	10,00	15,8713	1,800	99,00
Variance	1,390	1,314	2,748	1,633	1131,049	6,150E-02	11,333
Std. Deviation	1,18	1,15	1,658	1,28	33,6311	,248	3,37
Minimum	78	71	10,0	7	3,10	1,3	85
Maximum	82	76	16,0	12	157,94	2,1	100
Range	4	5	6,0	5	154,84	,8	15
Interquartile Range	2,00	1,50	2,000	2,00	24,2397	,495	2,00
Skewness	-,201	-,256	,302	-,169	3,090	-,081	-3,027
Kurtosis	-,827	,519	-,512	-,492	11,201	-1,192	10,370

Descriptives

Region or economic group: Africa

	Average female	Average male life	Birth rate per 1000	Death rate per	Number of aids	Fertility: average	People who read
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Mean	54,26	50,79	42,000	14,74	75,7491	6,081	47,26
5% Trimmed Mean	54,01	50,49	42,389	14,71	66,0056	6,088	47,29
Median	55,00	51,00	44,000	14,00	36,3077	6,290	50,00
Variance	63,649	52,731	41,111	25,538	7641,570	1,285	319,094
Std. Deviation	7,98	7,26	6,412	5,05	87,4161	1,134	17,86
Minimum	43	41	28,0	6	,13	3,8	18
Maximum	70	66	49,0	24	326,75	8,2	76
Range	27	25	21,0	18	326,61	4,4	58
Interquartile Range	12,00	11,00	5,000	7,00	112,6254	1,380	34,00
Skewness	,425	,352	-1,256	,126	1,562	-,586	,012
Kurtosis	-,434	-,458	,452	-,847	2,587	,119	-,964

Para generar la tabla: Analizar/Estadísticos Descriptivos/Explorar: esperanza de vida femenina y masculina, tasa de natalidad y mortalidad, promedio de hijos por familia etc. Factor: Región. Etiquetar por: país. Estadísticos: Valores Atípicos. Al editar la gráfica se borra lo que no se quiere mostrar.

Si queremos comparar Chile en esperanza de vida femenina y masculina, tasa de natalidad, tasa de mortalidad, fertilidad, tasa sida y alfabetización respecto al resto de países de la base de datos consideramos las puntuaciones z de cada una de ellas, las que nos muestran que:

- esperanza de vida femenina en Chile está sobre la media en 0,74 desviaciones estándares.
- esperanza de vida masculina en Chile está sobre la media en 0,65 desviaciones estándares.
- tasa de natalidad en Chile está bajo la media en 0,23 desviaciones estándares.
- tasa de mortalidad en Chile está bajo la media en 0,83 desviaciones estándares.
- fertilidad (promedio de hijos por familia) en Chile está bajo la media en 0,55 desviaciones estándares.
- tasa sida en Chile está bajo la media en 0,37 desviaciones estándares
- alfabetización (% de personas que saben leer) en Chile está sobre la media en 0,64 desviaciones estándares.

Caso: Encuesta Comunal

Debido a la cercanía de las elecciones municipales un conocido partido político del país lo (la) ha contratado a Usted con el fin de identificar el grado de satisfacción de los habitantes de la ciudad de Santiago con la comuna en donde residen, de esta manera poder realizar una campaña municipal dirigidas a atacar las características de las comunas donde se ha encontrado un grado de insatisfacción.

El grado de satisfacción de las personas encuestadas se cuantificaron porcentualmente con respecto a cada una de las características de la comuna puestas bajo tela de juicio, es decir, frente a la consulta de "¿Qué tan satisfecho está usted con X característica de su comuna?", el encuestado respondía en un rango de 0 a 100, donde 100 equivalía a "Muy Satisfecho", mientras que el valor 0 correspondía a "Completamente insatisfecho"

Las características de las comunas que fueron consultadas en esta encuesta fueron las siguientes:

- Nivel de Seguridad de la comuna.
- Cercanía a Servicios Básicos (Clínicas, Hospitales, Colegios, etc.).
- Cercanía a locales comerciales.
- Disponibilidad de lugares de esparcimiento.
- Desarrollo Vial (Estado de calles, iluminación, señalización, etc.).

La encuesta fue realizada 474 personas escogidas de manera aleatoria y los resultados de ésta se le entregaron tabulados en el archivo "Encuesta Comunal.sav" y en base a esta información se le solicita lo siguiente.

- a. Una de las creencias del partido político corresponde a que el grado de satisfacción con respecto a la comuna varía sustancialmente entre las personas que trabajan en la misma comuna en la que residen frente a los que trabajan en una comuna diferente a la que viven. Para salir de esta duda se le solicita que muestre gráficamente (en un solo gráfico) el porcentaje promedio de satisfacción de cada una de las características mencionadas para las personas que trabajan en la misma comuna en la que viven y para las personas que no trabajan en la misma comuna en la que viven. Comente los resultados más relevantes y muestre las sintaxis necesarias para llegar a este resultado.

Para obtener los resultados solicitados, en primer lugar se deben identificar aquellos encuestados que trabajan en la misma comuna en la que residen. Para estos efectos una de las soluciones posibles corresponde a la creación de una nueva variable por medio de la siguiente sintaxis.

```
if (comuna=comutrab) mismacom=1.
if (comuna<>comutrab) mismacom=0.
EXECUTE.
```

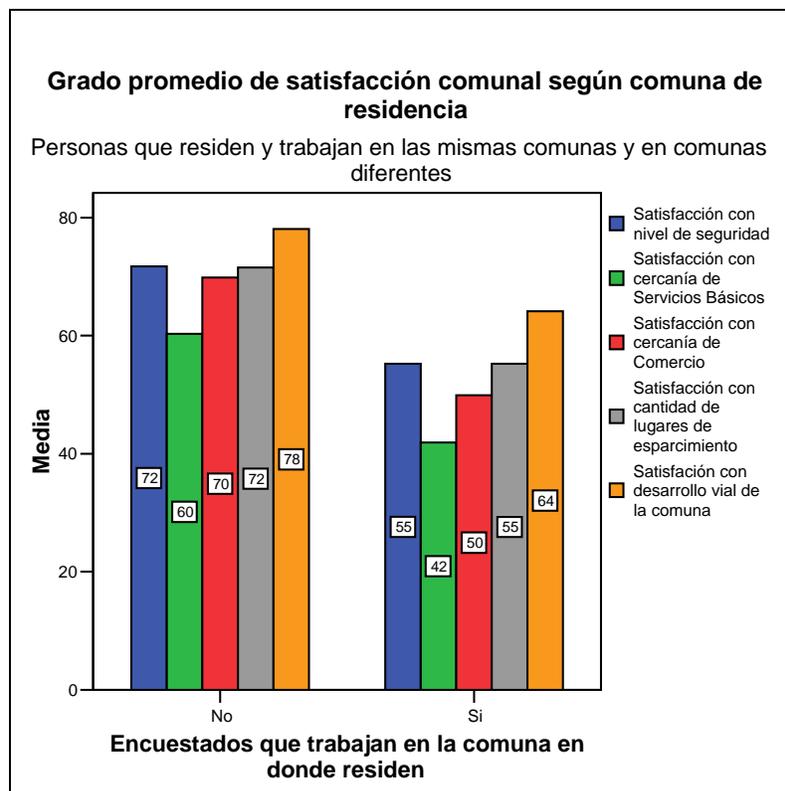
Luego, a la nueva variable se le deben asignar tanto una etiqueta como valores representativos de lo que se acaba de realizar. Para el caso particular corresponderían a las siguientes:

Etiqueta: "Encuestados que trabajan en la comuna en donde residen"

Valores: 0 → No.

1 → Si.

Mediante la generación de la variable detallada con anterioridad se cuenta con toda la información necesaria para obtener la información requerida. Para el caso particular se debe realizar un gráfico de barras agrupadas con resúmenes para diferentes variables, el cual se presenta a continuación.



El gráfico anterior muestra claramente que existe un mayor grado de satisfacción con las diferentes características comunales en las personas que trabajan en comunas diferentes a la que residen frente a las personas que trabajan en la misma comuna que residen, independiente de la característica observada, con diferencias de al menos 10 puntos porcentuales en cada una de las variables.

Lo anterior puede generarse por diversos motivos. Uno de ellos puede corresponder a que las personas que trabajan y residen en la misma comuna no tienen un punto de comparación objetivo con respecto a estas mismas características en otras comunas, tendiendo a ser más críticos con la comuna que conocen bien. Por otra parte, y debido al importante porcentaje de tiempo que pasan en la comuna donde residen/trabajan, conocen una mayor cantidad de carencias cotidianas en la propia comuna, punto de vista que no poseen personas que se encuentran la mayor parte del día en comunas ajenas a su residencia.

b. Luego de una ardua discusión, se determinó una fórmula para crear un porcentaje de "Satisfacción Total" de las personas con respecto a la comuna donde residen, dándole un peso específico a cada una de las variables particulares de satisfacción, este peso específico fue el siguiente:

- Nivel de Seguridad de la comuna. (35 %)
- Cercanía a Servicios Básicos (Clínicas, Hospitales, Colegios, etc.). (15 %)
- Cercanía a locales comerciales. (10 %)
- Disponibilidad de lugares de esparcimiento. (20 %)
- Desarrollo Vial (Estado de calles, iluminación, señalización, etc.). (20 %)

Se le solicita que genere la variable de Satisfacción Total y que además muestre una tabla con las cinco comunas (de residencia) con el menor porcentaje promedio de satisfacción total, con el fin de considerarlas como prioridad para comenzar con la campaña municipal. (Muestre todas las sintaxis necesarias).

Por medio de la definición presentada en el enunciado de la pregunta, lo primero que se debe realizar es el cálculo de la variable de satisfacción total. Este cálculo se detalla en la siguiente sintaxis.

```
COMPUTE stotal = ssegur * 0.35 + sservbas * 0.15 + scomerc * 0.10 + sesparc *  
0.20 + svial * 0.20 .  
VARIABLE LABELS stotal 'Porcentaje de satisfacción total con la comuna' .  
EXECUTE .
```

Luego, la pregunta es con respecto a las cinco comunas con menor porcentaje de satisfacción total, pero debido a que el archivo de trabajo corresponde a un archivo de **PERSONAS** encuestadas con su correspondiente comuna de residencia, no es el idóneo para contestar generar la información solicitada. Es por el motivo anterior que se debe generar un nuevo archivo con el promedio de satisfacción total de las personas por comuna. La sintaxis del procedimiento detallado es la que se presenta a continuación.

```
AGGREGATE /OUTFILE='C:\SPSS\SPSS 2008 51\Varios\Agregado EC.sav'
/BREAK=comuna
/stotal = MEAN(stotal)
/nencuest=N.
```

Por medio del nuevo archivo generado se pueden obtener, por medio de una tabla de valores atípicos, el detalle de las cinco comunas en donde los encuestados que residen en estas poseen los menores grados de satisfacción total. La tabla es la que se presenta a continuación.

Valores extremos				
<i>Menores</i>				
		Número del caso	Comuna de Residencia del Encuestado	Valor
stotal	1	33	VITACURA	58,73
	2	8	LA CISTERNA	58,79
	3	10	LA GRANJA	59,30
	4	6	HUECHURABA	59,33
	5	11	LA PINTANA	59,82

Entonces, las comunas a las cuales debiese orientar sus esfuerzos el partido político, son las comunas de Vitacura, La Cisterna, La Granja, Huechuraba y La Pintana.

c. A pesar de que cada una de las características que se han medido en la encuesta son de relevancia para la población, el partido político tiene la inquietud de que una de las variables que mas incidencia tiene en las personas es la de la seguridad, esto debido al aumento de secuestros Express, femicidios y asaltos en general. Es por esto que se le solicita lo siguiente:

- 1) Realice un análisis comparativo de los principales estadígrafos de tendencia central y dispersión (Media, Mediana, Mínimo, Máximo, desviación Estándar,

CV, etc.) de la satisfacción con la seguridad de la comuna según sexo. Comente lo más relevante.

Para obtener la información requerida se debe trabajar nuevamente con el archivo inicial. Desde este último se debe generar un cubo OLAP para la variable de satisfacción con la seguridad y como variable de segmentación el sexo. Los resultados obtenidos son los que se expresan en la tabla siguiente.

	Sexo del Encuestado	N	Media	Mediana	Desv. típ.	Mínimo	Máximo
Satisfacción con nivel de seguridad	Mujer	231	58,27	52,00	19,050	40	100
	Hombre	243	71,86	74,00	25,073	40	100
	Total	474	65,24	62,00	23,331	40	100

El cubo OLAP muestra que no existe una diferencia significativa en el número de encuestado de ambos sexos (231 mujeres y 243 hombres), sin embargo si existe una diferencia marcada en lo que corresponde a estadígrafos de tendencia central, ya que la satisfacción en seguridad es, en promedio, más alta para los hombres que para las mujeres. Lo anterior se ve reflejada tanto en la media como en la mediana, indicando que independiente de la comuna de residencia existe una clara muestra de insatisfacción de las mujeres con los niveles de seguridad presentes en su comuna. Lo anterior, con fines electorales, da un nicho claro sobre el cual proponer soluciones concretas y donde se podría tener una tasa de respuesta considerable.

Con respecto a la variabilidad de los datos, en las mujeres los puntajes de nivel de seguridad se desvían en promedio respecto a la media (58, 27) en aproximadamente 19 puntos, sin embargo para los hombres los valores del nivel de seguridad se desvían en promedio respecto a la media (71,86) en casi 25 puntos. Para ambos grupos el rango es el mismo. Por otra parte, al realizar una comparación de la variabilidad se aprecia que en el aspecto de seguridad existe una mayor variedad de respuestas en el caso de los hombres con respecto a las mujeres. El coeficiente de variabilidad (32,7% para las mujeres vs. Un 34,9% para los hombres).

- 2) Considerando los siguientes rangos de tiempo residiendo en la comuna: Hasta dos años; Más de 2 años hasta 5 años y Más de 5 años. Realice un análisis comparativo de los principales estadígrafos de tendencia central y dispersión (Media, Mediana, Mínimo, Máximo, desviación Estándar, CV, etc.) de la satisfacción con la seguridad de la comuna según los rangos de tiempo de residencia en la comuna. Comente lo más relevante.

Para este caso en particular, en primer lugar se deben generar los rangos de permanencia en la comuna de residencia en base a las especificaciones del enunciado. La sintaxis correspondiente a esta recodificación de datos es la siguiente.

RECODE

meses

(Lowest thru 24=1) (25 thru 60=2) (61 thru Highest=3) INTO ranmeses.

VARIABLE LABELS ranmeses 'Rango de Meses de Residencia'.

EXECUTE.

Ejecutada la sintaxis, se debe colocar valores a los resultados obtenidos:

1→ Hasta 2 Años.

2→ 2 - 5 Años.

3→ Más de 5 Años.

Con los rangos creados el análisis que se debe realizar es análogo al de la pregunta anterior. Los resultados del Cubo OLAP son los que se presentan a continuación.

	Rango de Meses	N	Media	Mediana	Desv. típ.	Mínimo	Máximo
Satisfacción con nivel de seguridad	Hasta 2 Años	142	68,37	68,00	24,864	40	100
	2 - 5 Años	110	61,95	58,50	21,121	40	100
	Más de 5 Años	222	64,86	62,00	23,209	40	100
	Total	474	65,24	62,00	23,331	40	100

Luego, se puede apreciar que en general las personas tienden a quedarse un tiempo considerable en la comuna de residencia, ya que de los encuestados cerca de un 47% de ellos reside en la misma comuna a lo menos 5 años. Por otra parte, los niveles de satisfacción con la seguridad de los encuestados tienen disímiles resultados en términos de promedio: Las personas con un mayor grado de conformidad con la seguridad de la comuna corresponden a los que tienen una menor permanencia en las mismas, esto puede darse debido a que las familias que tienen poco tiempo en las comunas aún no tienen toda la información de la historia de seguridad de los vecindarios que habitan y además por que en sus respuestas puede existir un sesgo dado por la conformidad del nuevo hogar que puede poseer un número importante de estos encuestados (La gran mayoría de las personas está feliz cuando se ha cambiado

de casa hace poco tiempo). Los resultados de los otros rangos de tiempo en la residencia son coherentes a lo expresado anteriormente, las personas con más bajo nivel de satisfacción son las que llevan entre 2 y 5 años en la comuna, por lo que ya tienen mayor información de seguridad disponible como para tener una opinión informada. Por último las personas con más de 5 años en la comuna el índice de satisfacción vuelve a subir y esto es que ya después de cinco años las personas están acostumbradas y/o agradas con su lugar de residencia. Los valores de la mediana son similares a los de la media, corroborando el sentido del análisis anterior.

Con respecto a la variabilidad de las respuestas, se da que no hay grandes diferencias entre los tres grupos, comparando los valores del coeficiente de variabilidad (36,4% para los de menos de dos años, 34,1% para los que tienen entre 2 y 5 años en la respectiva comuna y 35,8% para los que tienen más de 5 años en la comuna).

- 3) ¿Qué porcentaje del total de personas que viven en su comuna entre dos y cinco años son mujeres?. ¿Qué porcentaje del total de encuestados son hombres y viven hace más de cinco años en la misma comuna?. ¿Qué porcentaje del total de hombres vive en la misma comuna hace a lo más dos años?

Utilizando los rangos de tiempo de permanencia en la comuna creados para responder la pregunta anterior, se debe construir una tabla de contingencia entre las variables sexo y rangos de tiempo de permanencia en la comuna, con porcentajes para filas y columnas. La tabla de contingencia es la siguiente.

		Rango de Meses de Residencia			Total	
		Hasta 2 Años	2 - 5 Años	Más de 5 Años		
Sexo del Encuestado	Mujer	Recuento	40	80	111	231
		% de Sexo del Encuestado	17,3%	34,6%	48,1%	100,0%
		% de Rango de Meses de Residencia	28,2%	72,7%	50,0%	48,7%
		% del total	8,4%	16,9%	23,4%	48,7%
	Hombre	Recuento	102	30	111	243
		% de Sexo del Encuestado	42,0%	12,3%	45,7%	100,0%
		% de Rango de Meses de Residencia	71,8%	27,3%	50,0%	51,3%
		% del total	21,5%	6,3%	23,4%	51,3%
Total	Recuento	142	110	222	474	
	% de Sexo del Encuestado	30,0%	23,2%	46,8%	100,0%	
	% de Rango de Meses de Residencia	100,0%	100,0%	100,0%	100,0%	
	% del total	30,0%	23,2%	46,8%	100,0%	

Por medio de la tabla anterior se pueden contestar las tres preguntas planteadas:

- Del total de personas que viven en su comuna entre 2 y 5 años un 72,7% son mujeres.
 - Un 23,4% de los encuestados son hombres y viven hace más de 5 años en la misma comuna.
 - Del total de hombres, un 42% vive en la comuna de residencia respectiva a lo más hace dos años.
- d. El partido político, en base al universo de personas que contestaron esta encuesta, desea contratar como asesores a aquellos ciudadanos que tienen como porcentaje de satisfacción un valor mayor o igual al percentil 70 en cada una de las variables de satisfacción, con el fin de preparar un paquete de medidas concretas y óptimas para la campaña. Se le solicita encontrar el número de personas que cumplen con este requisito e identificar cuantas de ellas son hombres y cuantas de ellas son mujeres.

Para identificar a los posibles asesores del partido político, en primer lugar se debe calcular el percentil 70 de cada una de las variables de satisfacción de la encuesta. Estos porcentajes se encuentran expresados en la siguiente tabla.

Estadísticos			
Satisfacción con nivel de seguridad	N	Válidos	474
		Perdidos	0
	Percentiles	70	83,00
Satisfacción con cercanía de Servicios Básicos	N	Válidos	474
		Perdidos	0
	Percentiles	70	70,50
Satisfacción con cercanía de Comercio	N	Válidos	474
		Perdidos	0
	Percentiles	70	83,00
Satisfacción con cantidad de lugares de esparcimiento	N	Válidos	474
		Perdidos	0
	Percentiles	70	83,00
Satisfacción con desarrollo vial de la comuna	N	Válidos	474
		Perdidos	0
	Percentiles	70	98,00

Luego, existe un número diverso de formas de identificar el número de candidatos idóneos para ofrecerles el cargo de asesor. Una de estas formas es realizando una selección de casos seguido por una tabla de frecuencia de la variable sexo. La sintaxis de la selección de casos es la que se muestra a continuación.

USE ALL.

```

COMPUTE filter_$=((ssegur >= 83) & (sservbas >= 70.5) & (scomerc >= 83) &
(sesparc >= 83) & (svial >= 98)).
VARIABLE LABEL filter_$ '(ssegur >= 83) & (sservbas >= 70.5) & (scomerc >= 83) &
(sesparc >= 83) & (svial >= 98)'+
'(FILTER)'.
VALUE LABELS filter_$ 0 'No seleccionado' 1 'Seleccionado'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

```

Con los datos seleccionados se genera la tabla de frecuencias para obtener el número de hombres y mujeres que cumple con la condición señalada.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Mujer	1	50,0	50,0	50,0
	Hombre	1	50,0	50,0	100,0
	Total	2	100,0	100,0	

Luego, se aprecia que son dos personas las que cumplen con las condiciones de estar por sobre el percentil 70 en todos los niveles de satisfacción comunal y se tornan candidatos a ser asesores del partido político en caso de aceptar serlo. Una de estas personas es de género masculino mientras que la otra es de género femenino.

Caso: Municipalidades

La Asociación Chilena de Municipalidades debe mostrar la gestión del año 1999 de los distintos municipios del país. Para ello, han decidido contratarlo como asesor experto, proporcionándole un archivo de datos "Información comunas 1999", con los ingresos y gastos del año 1999 desglosado por los distintos ítems presupuestarios.

Usted sabe que las Municipalidades gozan de autonomía para la administración de sus finanzas. La Ley de Presupuestos de la Nación asigna recursos para atender sus gastos, además de los otros ingresos que pueden ser entregados por otras leyes o por los gobiernos regionales respectivos. Adicionalmente, la Ley Orgánica Constitucional de Municipalidades contempla un mecanismo de distribución solidaria de los ingresos propios entre las municipalidades del país, denominado Fondo Común Municipal (FCM).

Los ingresos propios están agrupados por los siguientes conceptos: rentas e inversiones, impuesto territorial, permisos de circulación, patentes, derechos de aseo, derechos varios y multas e intereses. Por su parte, los gastos están compuestos de Gastos en personal, bienes y servicios de consumo, transferencias, inversión real y otros gastos.

Su informe debe presentar antecedentes que respondan al menos- las siguientes inquietudes de los directivos de la Asociación Chilena de Municipalidades.

1.- Defina la variable Tipo (Rural y Urbana) y Zona (Norte, Centro, Sur) según los siguientes criterios

Si $tasarur \leq 50$ entonces Tipo es Urbana

Si $tasarur > 50$ entonces Tipo es Rural

Desde la primera a la cuarta región Zona=Norte

Desde la quinta a séptima y región metropolitana Zona= Centro

Desde la VIII a XII Zona = Sur.

- a) Muestre las sintaxis de las variables creadas e identifique cuáles regiones tienen más del 50% de sus comunas rurales. (10 puntos)
- b) ¿Cuántas comunas urbanas tiene la zona sur?. ¿Qué porcentaje representan respecto al total de comunas de Chile ?. ¿Cuántas comunas de la zona Centro son rurales y qué porcentaje representan respecto al total de comunas de Chile?. ¿Qué

porcentaje respecto al total de comunas rurales son de la zona centro? ¿Qué porcentaje respecto al total de comunas de la zona centro son rurales? (8 puntos)

a) Muestre las sintaxis de las variables creadas e identifique cuáles regiones tienen más del 50% de sus comunas rurales.

IF (tasarur <= 50) tipo = 1 .

EXECUTE .

IF (tasarur > 50) tipo = 2 .

EXECUTE .

En definición de las variables

1=Urbana

2= Rural

RECODE

regnum

(13=2) (1 thru 4=1) (8 thru 12=3)

(5 thru 7=2) INTO zona .

VARIABLE LABELS zona 'zona' .

EXECUTE .

En def de variables

1= Norte

2=Centro

3=Sur

Las regiones que tienen más del 50% de sus comunas rurales son, 4,6,7,9,10 ,11 y 12

Tabla de contingencia Código de la Región * tipo (Rural o Urbana)

		tipo (Rural o Urbana)		Total	
		Urbana	Rural		
Código de la Región	1	Recuento	5	5	10
		% de Código de la Región	50,0%	50,0%	100,0%
	2	Recuento	6	3	9
		% de Código de la Región	66,7%	33,3%	100,0%
	3	Recuento	8	1	9
		% de Código de la Región	88,9%	11,1%	100,0%
	4	Recuento	6	9	15
		% de Código de la Región	40,0%	60,0%	100,0%
	5	Recuento	35	3	38
		% de Código de la Región	92,1%	7,9%	100,0%
	6	Recuento	14	19	33
		% de Código de la Región	42,4%	57,6%	100,0%
	7	Recuento	10	20	30
		% de Código de la Región	33,3%	66,7%	100,0%
	8	Recuento	30	22	52
		% de Código de la Región	57,7%	42,3%	100,0%
	9	Recuento	14	17	31
		% de Código de la Región	45,2%	54,8%	100,0%
	10	Recuento	14	28	42
		% de Código de la Región	33,3%	66,7%	100,0%
	11	Recuento	4	6	10
		% de Código de la Región	40,0%	60,0%	100,0%
	12	Recuento	4	7	11
		% de Código de la Región	36,4%	63,6%	100,0%
	13	Recuento	46	6	52
		% de Código de la Región	88,5%	11,5%	100,0%
Total		Recuento	196	146	342
		% de Código de la Región	57,3%	42,7%	100,0%

b) ¿Cuántas comunas urbanas tiene la zona sur?. ¿Qué porcentaje representan respecto al total de comunas de Chile?. ¿Cuántas comunas de la zona Centro son rurales y qué porcentaje representan respecto al total de comunas de Chile?. ¿Qué porcentaje respecto al total de comunas rurales son de la zona centro? ¿Qué porcentaje respecto al total de comunas de la zona centro son rurales?

Tabla de contingencia zona * tipo (Rural o Urbana)

			tipo (Rural o Urbana)		Total
			Urbana	Rural	
zona	Norte	Recuento	25	18	43
		% de zona	58,1%	41,9%	100,0%
		% de tipo (Rural o Urbana)	12,8%	12,3%	12,6%
		% del total	7,3%	5,3%	12,6%
	Centro	Recuento	105	48	153
		% de zona	68,6%	31,4%	100,0%
		% de tipo (Rural o Urbana)	53,6%	32,9%	44,7%
		% del total	30,7%	14,0%	44,7%
	Sur	Recuento	66	80	146
		% de zona	45,2%	54,8%	100,0%
		% de tipo (Rural o Urbana)	33,7%	54,8%	42,7%
		% del total	19,3%	23,4%	42,7%
Total	Recuento	196	146	342	
	% de zona	57,3%	42,7%	100,0%	
	% de tipo (Rural o Urbana)	100,0%	100,0%	100,0%	
	% del total	57,3%	42,7%	100,0%	

La zona Sur tiene 66 comunas urbanas y representan el 19,3% del total de comunas. Del total de las comunas de la zona centro (153), 48 comunas son rurales y representan el 14% respecto al total de las comunas.

Respecto al total de comunas rurales el 32,9% son de la zona Centro.

Respecto al total de comunas de la zona centro el 31,4% son rurales

2. Calcule la población total proyectada 2000 para cada comuna y para la variable calculada de población total, estime por región el número de casos, mínimo, máximo, promedio, mediana, desv estándar y población total proyectada 2000. Muestre una sintaxis de todo el procedimiento y una tabla con los resultados para la Quinta Región. ¿Cuáles son las cinco comunas con menor y mayor población proyectada 2000?. Comente. (14 puntos)

```

COMPUTE poblac = v2000 + m2000 .
VARIABLE LABELS poblac 'Población total proyectada 2000' .
EXECUTE .

```

OLAP CUBES

```

poblac BY regnum
/CELLS=COUNT MIN MAX MEAN MEDIAN STDDEV SUM
/TITLE='Cubos OLAP'.

```

Cubos OLAP

Código de la Región: 5

	N	Mínimo	Máximo	Media	Mediana	Desv. típ.	Suma
Población total proyectada 2000	38	483	319141	41093,61	13767,00	69542,953	1561557

USE ALL.

```

COMPUTE filter_$=(regnum = 5).
VARIABLE LABEL filter_$ 'regnum = 5 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .

```

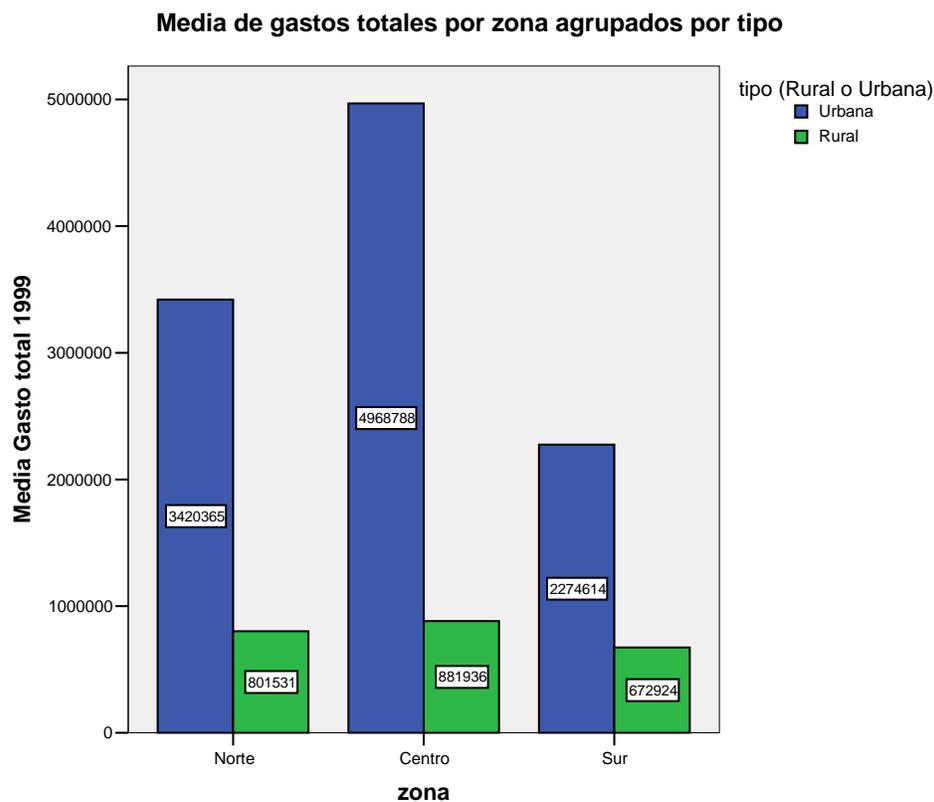
Valores extremos

			Número del caso	Nombre comuna	Valor
Población total proyectada 2000	Mayores	1	52	VIÑA DEL MAR	319141
		2	44	VALPARAISO	295203
		3	49	QUILPUE	122883
		4	70	SAN ANTONIO	89362
		5	51	VILLA ALEMANA	87589
	Menores	1	47	JUAN FERNANDEZ	483
		2	53	ISLA DE PASCUA	3618
		3	60	PAPUDO	4321
		4	62	ZAPALLAR	5191
		5	74	EL TABO	5870

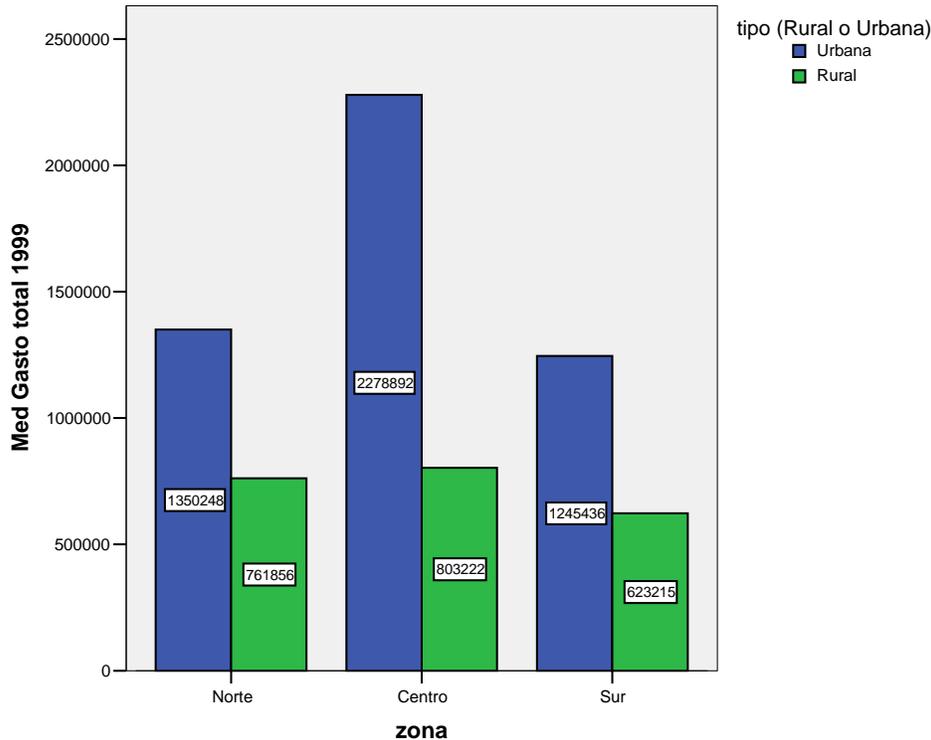
La Quinta Región cuenta con 38 comunas. La población total proyectada 2000 de las comunas de la Quinta Región oscila entre 483 y 319.141 habitantes, con un total regional de 1.561.557 habitantes. En promedio las comunas de la Región tienen 41093 habitantes aproximadamente con una dispersión de 13767 habitantes.. El 50% de las comunas de la región tiene hasta 13.767 habitantes.

En la tabla de valores extremos se observan las cinco comunas con mayor y menor población proyectada 2000 siendo Juan Fernández la comuna con menor población y Viña del Mar la comuna con mayor población.

3.- Realice dos gráficos, uno que permita visualizar la media y otro la mediana de los gastos totales de 1999 de las comunas por zona, agrupados por tipo (Rural o urbano). Comente.



Mediana de los gastos totales de las comunas por zona agrupados por tipo



El primer gráfico muestra la media de gastos totales de 1999 por zona, agrupados por tipo de comuna. Se observa que en todas las zonas la media de los gastos para las comunas urbanas supera notablemente la media de gastos para las comunas rurales, siendo notoria la diferencia para las comunas urbanas de la zona centro. En la zona centro se observan los mayores gastos promedios tanto para las comunas urbanas como para las rurales.

En el segundo gráfico se observa un comportamiento muy similar para la mediana de los gastos totales de las comunas por zona agrupados por tipo. Se destaca el valor de la mediana de gastos para las comunas de la zona centro y tipo urbanas donde el 50% de las comunas tienen un gasto mayor o igual a \$2.278.892.

4.- Considere la zona norte.

Determine por tipo (Rural y urbano) de comuna los siguientes estadísticos: Media, Mediana, Desviación estándar, mínimo, máximo, cuartiles para las variables: Ingreso propio permanente 1999, Inversión real 1999 y el Gasto total 1999 de las comunas. Muestre la sintaxis de todo el procedimiento y comente los estadísticos solicitados para la variable ingreso propio permanente. (10 puntos)

```

USE ALL.
COMPUTE filter_$=(zona = 1).
VARIABLE LABEL filter_$ 'zona = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .

```

```

SORT CASES BY tipo .
SPLIT FILE
  LAYERED BY tipo .

```

```

FRECUENCIAS
  VARIABLES=ipp99 inreal99 gtot99 /FORMAT=NOTABLE
  /NTILES= 4
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
  /ORDER= ANALYSIS .

```

Estadísticos

tipo (Rural o Urbana)			Ingreso propio permanente 1999	Inversión real 1999	Gasto total 1999	
Urbana	N	Válidos	25	25	25	
		Perdidos	0	0	0	
	Media		1561812,16	774042,40	3420365	
	Mediana		500620,00	365063,00	1350248	
	Desv. típ.		2193071,933	898764,56	3846333	
	Mínimo		63190	51779	512847	
	Máximo		7960986	3459527	14626857	
	Percentiles	25		234854,50	268306,50	1054640
		50		500620,00	365063,00	1350248
75			2542501,50	1100316,0	5258584	
Rural	N	Válidos	18	18	18	
		Perdidos	0	0	0	
	Media		117854,39	229733,94	801530,78	
	Mediana		73732,50	185260,50	761856,50	
	Desv. típ.		95372,830	150849,46	326001,0	
	Mínimo		6205	34761	353338	
	Máximo		354907	596981	1555382	
	Percentiles	25		42334,50	101249,50	559314,25
		50		73732,50	185260,50	761856,50
75			172067,50	339510,50	1012230	

Al considerar la zona norte, por tipo se observa 25 comunas de la zona norte de tipo urbanas las cuales presentan una media de ingresos propios muy sesgada por los valores extremos superiores alcanzando un monto de \$1.561.812 con una gran variabilidad entre las comunas alcanzando una desviación estándar que indica que los datos de ingresos se desvían en promedio respecto a la media en \$ 2.193.071 valor muy superior a la media. La mediana está muy por debajo de la media, indicando que el 50% de las comunas tienen un ingreso propio permanente de \$500.620, casi un tercio de la media. Es notable observar la gran diferencia entre los cuartiles de ingresos de las comunas, mostrando que un cuarto de las comunas tienen ingresos propios menores a \$234.854 y que el cuarto superior de ingresos tienen ingresos superiores a \$2542501.

Las comunas rurales muestran una realidad muy distinta, con valores muy inferiores a los de las comunas urbanas. Son en total 18 comunas rurales con un promedio de ingresos propios de \$117.854 bastante más alto que la mediana que indica que la mitad tiene ingresos propios superior a \$73.732 con un máximo de \$354.907. También se presenta una alta dispersión de los datos de ingresos entre las comunas pero con un valor algo menor a la media pero superior a la mediana. El primer cuartil indica que un cuarto de las comunas tienen un ingreso inferior o igual a \$42.334 mientras que el tercer cuartil indica que un cuarto de las comunas tiene un ingreso superior o igual a \$172067.