

ESTADÍSTICA

Práctica 2

Grupo 6 - Práctica con SPSS.

Alumno: Elvi Mihai Sabau Sabau.

DNI: 51254875L

2.6 Ejercicios.

1. Clasificar las siguientes variables razonando la respuesta:

Nombre	CN: Los nombres no se pueden ordenar, a no ser que sea alfabéticamente.
Edad (en años)	MD: La edad es una cifra exacta.
Longitud de 150 tornillos producidos en una fábrica	MC: La longitud del tornillo nunca será una cifra exacta.
Número de páginas de una serie de libros de estadística	MD: La cantidad de páginas se puede contar, y es una cifra exacta.
Tiempo requerido para responder las llamadas en un call center	MC: El tiempo nunca será una medida exacta.
Calificación en un examen (Susp., Ap., etc.)	CO: No son cifras, pero se pueden ordenar.
Diámetro de una esfera	MD: El diámetro (como medida) es una cifra exacta, otra cosa sería plasmarlo.
Color de pelo	CN: Los colores, por defecto, no se pueden ordenar, a no ser que se tenga en cuenta un valor representando dicho color, como por ejemplo, su valor en HEX (RGB).
Grado de satisfacción laboral en una empresa	CO: El grado se puede ordenar, aunque no esté representado con una cifra.
Resultado de tirar dos dados	MD: La suma de ambas cifras siempre será un valor exacto.

2. En la tabla adjunta tenemos la representación de un conjunto de datos obtenidos de una población; se trata de una muestra de 20 ordenadores de una tienda informática, de los cuales observamos varias características; para cada ordenador obtenemos datos correspondientes a las variables:

- X_1 : Tiempo en segundos de conexión a Internet en un determinado periodo de tiempo.
- X_2 : Sistema Operativo preinstalado. (2)
- X_3 : Marca. (1)
- X_4 : Número de periféricos.
- X_5 : Precio en euros.

(1)

1. LENOVO
2. ASUS
3. ACER
4. DELL
5. APPLE
6. HP
7. SAMSUNG
8. MSI

(2)

1. Windows 10
2. LINUX
3. Mac OS
4. Windows 8
5. UNIX
6. Windows 7

	X_1	X_2	X_3	X_4	X_5
Ordenador	Tiempo conexión (s)	S.O.	Marca	Periféricos	Precio
1	1500	6	2	4	810
2	1989	4	1	6	1340
3	2900	3	5	6	2000
4	2400	3	5	5	1920
5	1200	1	8	6	2325
6	1200	1	2	3	1020
7	1999	4	4	4	934
8	1050	3	5	4	1570
9	2300	4	7	3	1098
10	825	1	2	4	945
11	3995	5	4	3	1299
12	1999	2	6	4	1985
13	2085	2	8	7	2020
14	1020	4	3	3	955
15	994	1	1	5	800
16	1375	1	3	4	1195
17	3000	2	6	5	2145
18	945	1	1	3	999
19	1999	5	6	5	1585
20	2284	3	5	4	1720

Creamos las variables:

N...	Type	Label	Values	Miss...	...	Align	Measure	Role
X1	Numeric	4	0	Tiempo de conexion (s)	None	None	8	Right	Scale	Input
X2	Numeric	1	0	S.O.	{1, W...	None	8	Right	Nomi...	Input
X3	Numeric	1	0	Marca	{1, L...	None	8	Right	Nomi...	Input
X4	Numeric	1	0	Perifericos	None	None	8	Right	Scale	Input
X5	Numeric	4	0	Precio	None	None	8	Right	Scale	Input

Damos nombre a los valores:

Value Labels

Value: 6

Label: Windows 7

Add

Change

Remove

1 = "Windows 10"
2 = "Linux"
3 = "Mac OS"
4 = "Windows 8"
5 = "UNIX"
6 = "Windows 7"

OK

Cancel

Help

Value Labels

Value: 8

Label: MSI

Add

Change

Remove

3 = "ACER"
4 = "DELL"
5 = "APPLE"
6 = "HP"
7 = "SAMSUNG"
8 = "MSI"

OK

Cancel

Help

Y, introducimos los datos:

	X1	X2	X3	X4	X5
1	1500	6	2	4	810
2	1989	4	1	6	1340
3	2900	3	5	6	2000
4	2400	3	5	5	1920
5	1200	1	8	6	2325
6	1200	1	2	3	1020
7	1999	4	4	4	934
8	1050	3	5	4	1570
9	2300	4	7	3	1098
10	825	1	2	4	945
11	3995	5	4	3	1299
12	1999	2	6	4	1985
13	2085	2	8	7	2020
14	1020	4	3	3	955
15	994	1	1	5	800
16	1375	1	3	4	1195
17	300	2	6	5	2145
18	945	1	1	3	999
19	1999	5	6	5	1585
20	2284	3	5	4	1720

a) Clasificar las variables. Dar una explicación razonada de la clasificación.

X1: MC, ya que, aunque el valor esté en segundos, el tiempo de conexión no es exacto.

X2: CN, no se puede ordenar las marcas.

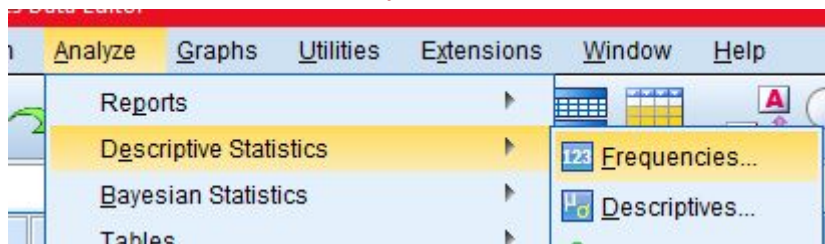
X3: CN, no se puede ordenar el nombre del S.O.

X4: MD, el número de periféricos es un valor exacto.

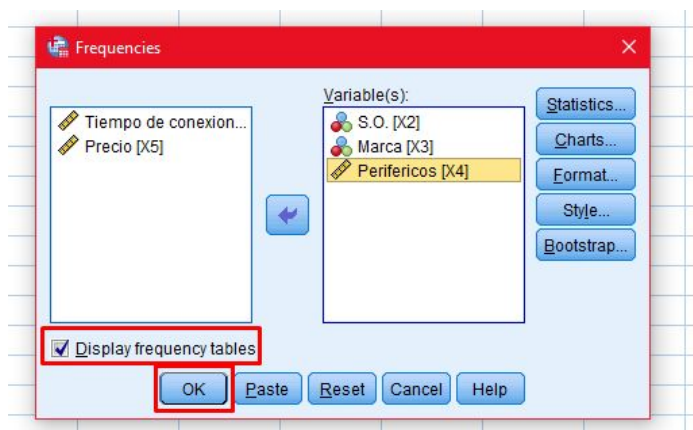
X5: MD, el precio es un valor exacto.

b) Construir las tablas de frecuencias de las variables X2, X3 y X4.

Primero, vamos al submenú, y seleccionamos, "Frecuencias".



También, chequeamos el checkbox de "Mostrar tablas de frecuencia" y le damos a OK.



¿Qué conclusiones puedes extraer de las tablas de frecuencias?

S.O.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Windows 10	6	30,0	30,0	30,0
	Linux	3	15,0	15,0	45,0
	Mac OS	4	20,0	20,0	65,0
	Windows 8	4	20,0	20,0	85,0
	UNIX	2	10,0	10,0	95,0
	Windows 7	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

De la tabla de SO, podemos concluir que, la mayoría de los equipos viene con W10, y que solo uno viene con W7, mientras tanto, el resto de equipos poseen una variedad de SO restantes.

Marca					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LENOVO	3	15,0	15,0	15,0
	ASUS	3	15,0	15,0	30,0
	ACER	2	10,0	10,0	40,0
	DELL	2	10,0	10,0	50,0
	APPLE	4	20,0	20,0	70,0
	HP	3	15,0	15,0	85,0
	SAMSUNG	1	5,0	5,0	90,0
	MSI	2	10,0	10,0	100,0
	Total	20	100,0	100,0	

De la tabla de Marcas, podemos decir que, la mayoría de dispositivos es de apple, mientras que hay un promedio de ordenadores de marcas Lenovo, Asus, Acer, Dell y MSI, siendo Samsung la marca con menor cantidad de ordenadores.

Perifericos					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	5	25,0	25,0	25,0
	4	7	35,0	35,0	60,0
	5	4	20,0	20,0	80,0
	6	3	15,0	15,0	95,0
	7	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

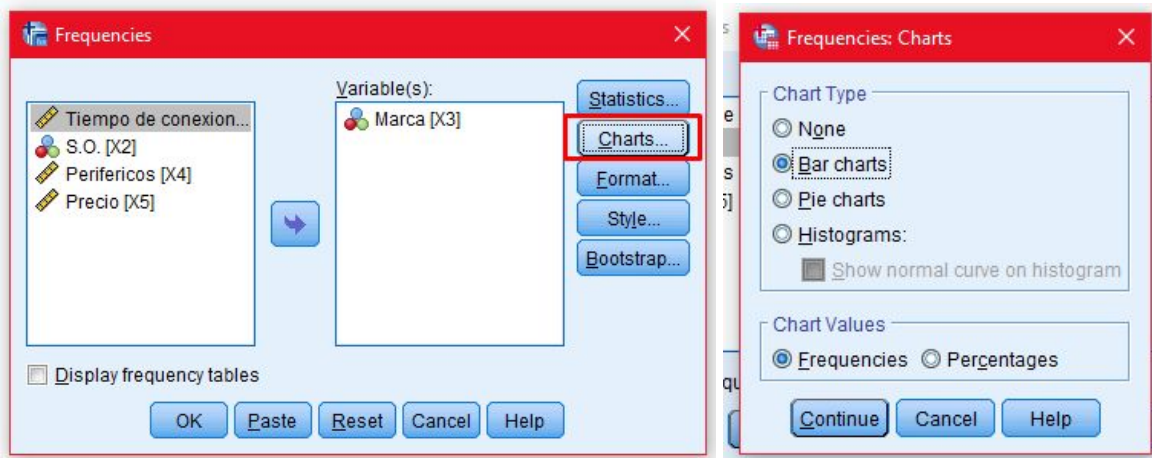
De la tabla de Periféricos, podemos deducir que, la mayoría de ordenadores vienen con 4 periféricos, y sólo un ordenador con un total de 7 periféricos, y el resto oscilan entre 3, 5 y 6 periféricos.

c) En un estudio estadístico, ¿para qué utilizarías las frecuencias relativas en vez de las frecuencias absolutas?

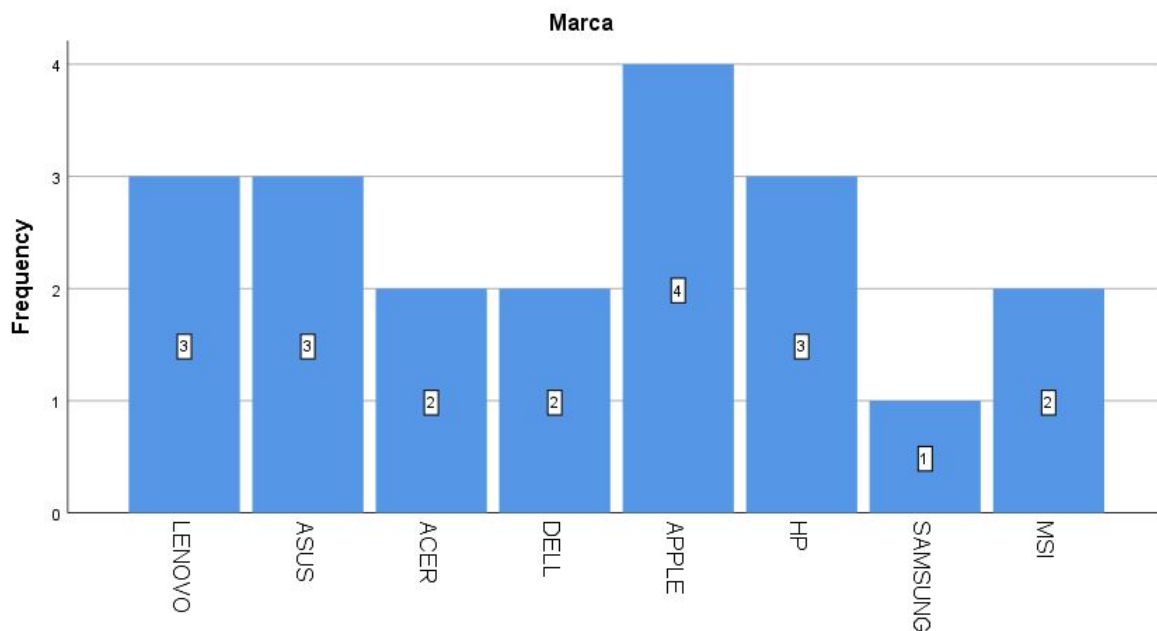
La tabla de frecuencias relativas, se puede usar tanto para obtener una información sobre el porcentaje de elementos analizados, como para comprobar que los resultados son todos válidos.

d) Representar la variable X3 de dos formas distintas: mediante un diagrama de barras y mediante un diagrama de sectores en el que se incluyan los porcentajes.

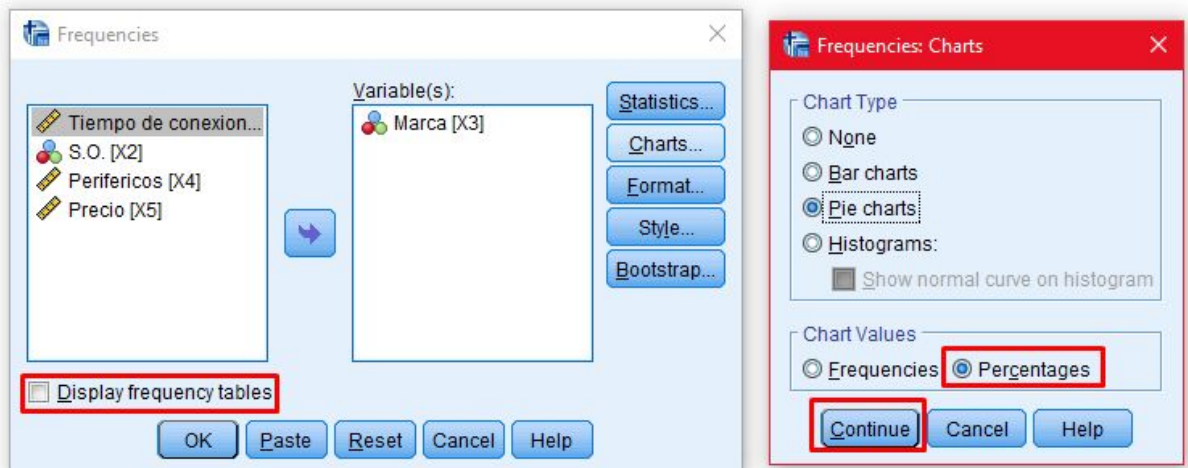
Para ello, volvemos al submenú de “Frecuencias” Pero esta vez, dejamos solo la variable X3, desmarcando la casilla de “Mostrar tablas de frecuencias”, y pinchamos en “Charts” (Gráficos), y seleccionamos, “Bar Charts”, dejando el tipo de valor por defecto (frecuencias).



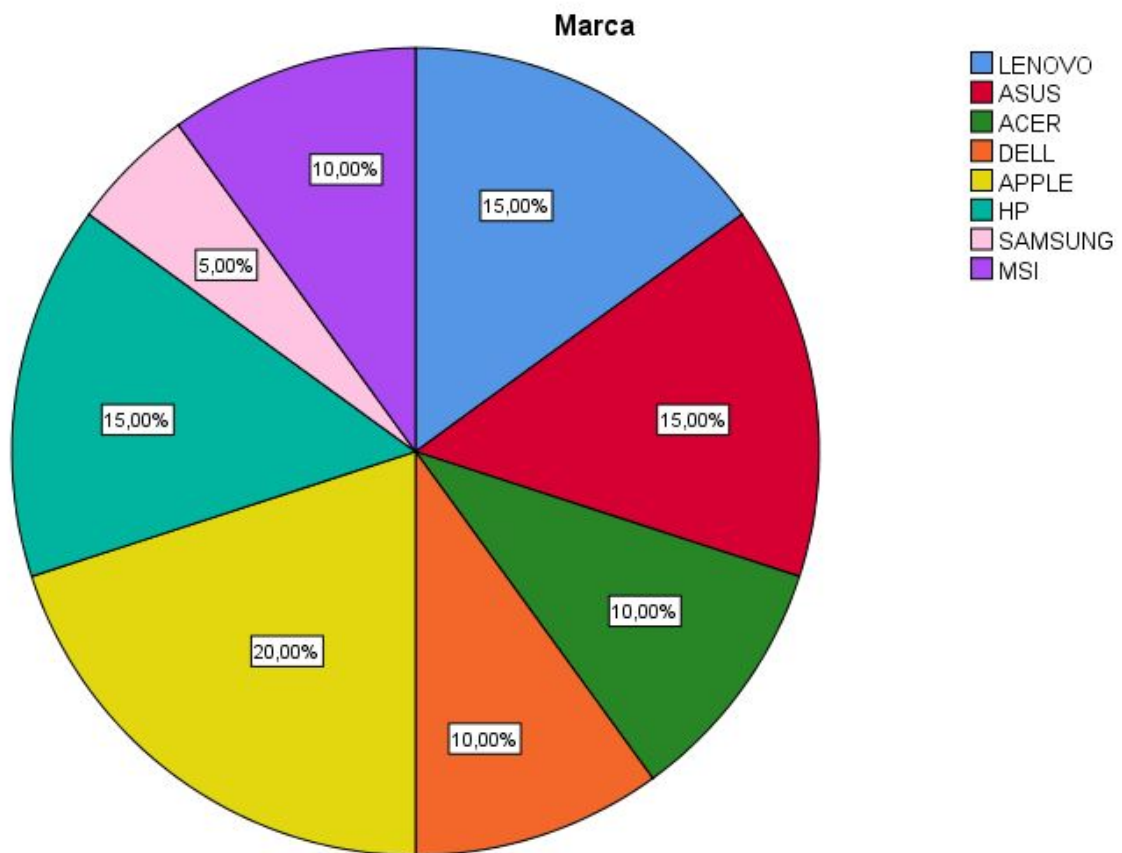
Le damos a “Continue” y se nos abrirá una ventana mostrando la tabla de frecuencias con el gráfico de barras.



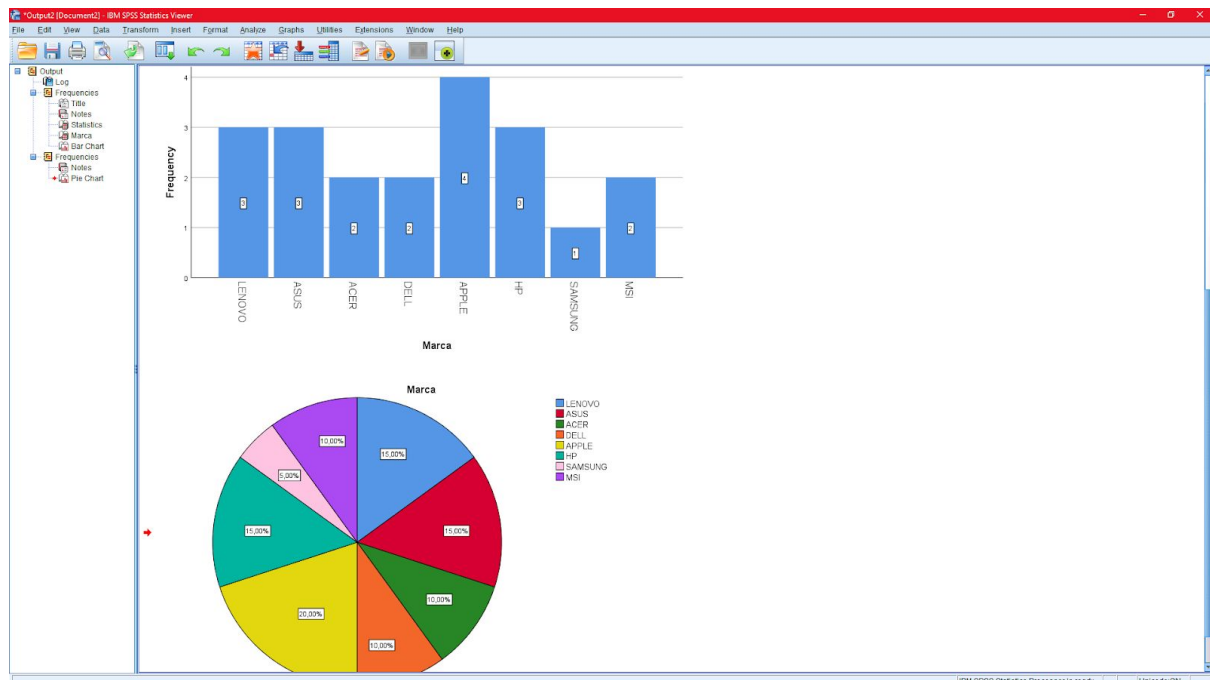
Y ahora, para añadir el gráfico de sectores, lo que haremos será volver al submenú de “Frecuencias” y realizar el mismo proceso, desmarcando la casilla de “Mostrar tablas de frecuencias”, seleccionando el gráfico de sectores, y especificando el tipo de valor, seleccionando el tipo “porcentajes”.



Al darle a “continue” se nos añadirá en la misma ventana, el gráfico de sectores.



Ejemplo del resultado:

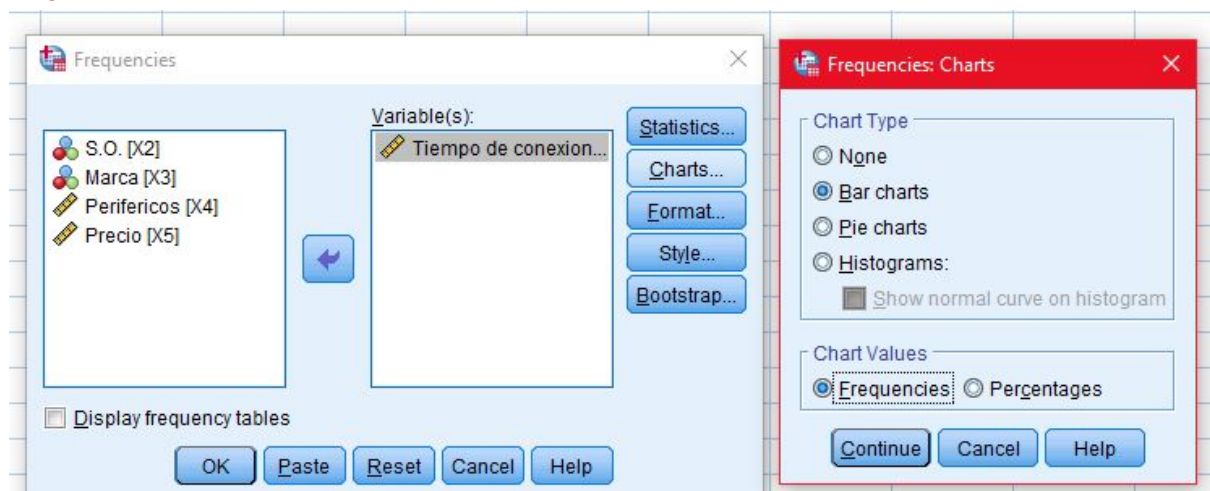


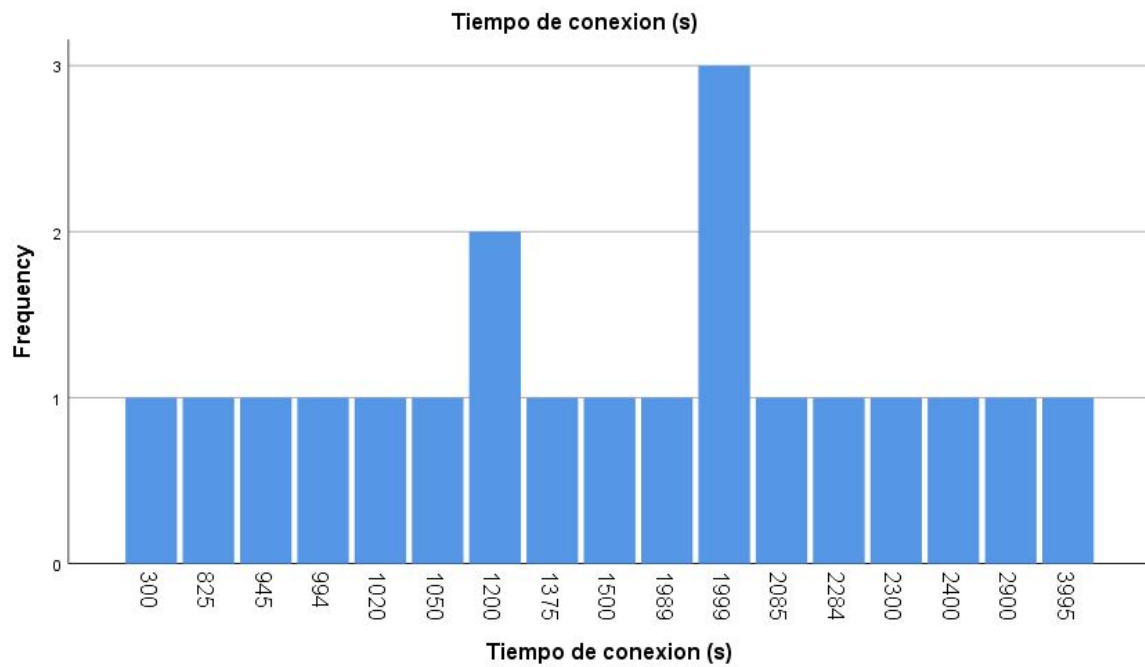
¿Qué conclusiones extraes? ¿Qué ventajas/desventajas ofrece el diagrama de sectores frente al diagrama de barras?

Debido al tipo de datos que estamos representando, el diagrama de sectores es mejor, ya que resulta más fácil de interpretar los resultados, de una manera visual.

e) Representar mediante un diagrama de barras la variable X1. Interpreta el resultado.

Realizamos el mismo proceso que hemos realizado en el apartado anterior para generar un diagrama de barras.





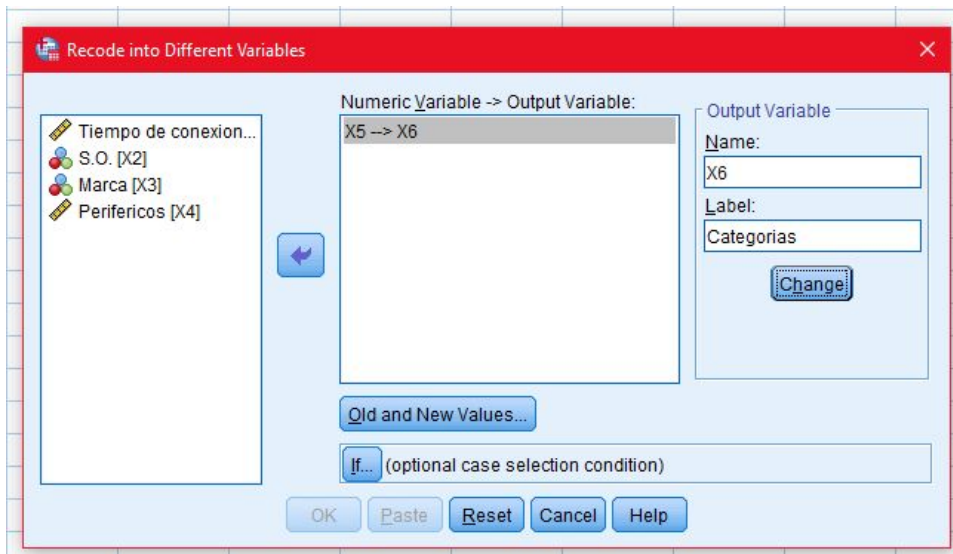
Este diagrama nos da a interpretar cuántos ordenadores poseen el mismo precio, en este caso sólo hay 5 ordenadores con un precio similar, 2 al precio de 1200, y 3 con el precio de 1999, mientras que el resto de ordenadores tienen un precio único en comparación con el resto.

f) Obtén el gráfico de sectores de la variable X5 agrupando los precios de los ordenadores en dos categorías:

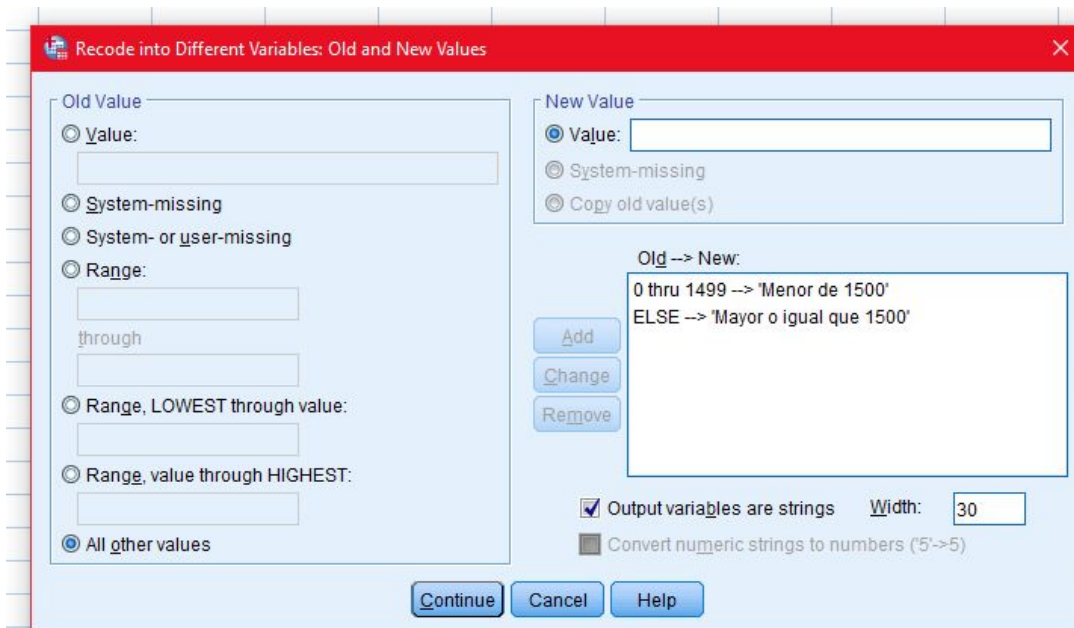
1 - menos de 1500 euros, 2 - más de 1500 euros.

Incluye los porcentajes en el gráfico y explica lo que observas.

Primero, vamos al submenú "Transform" > "Recodificar variables en diferentes variables" y creamos una variable que nos servirá para categorizar los precios.

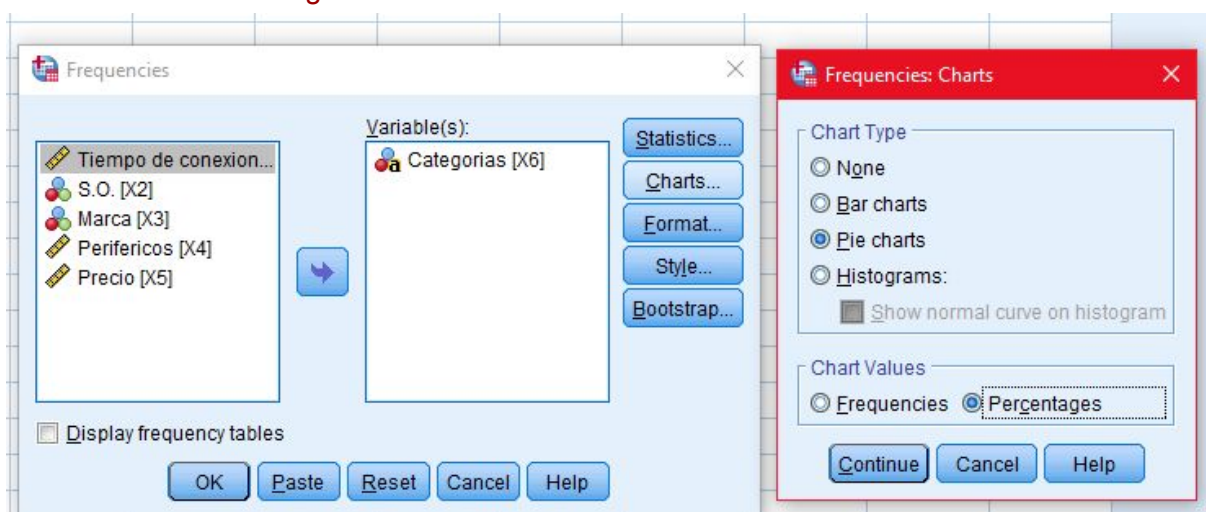


Asignamos el rango de valores, de 0 - 1499, y el resto (por defecto, mayor o igual que 1500).

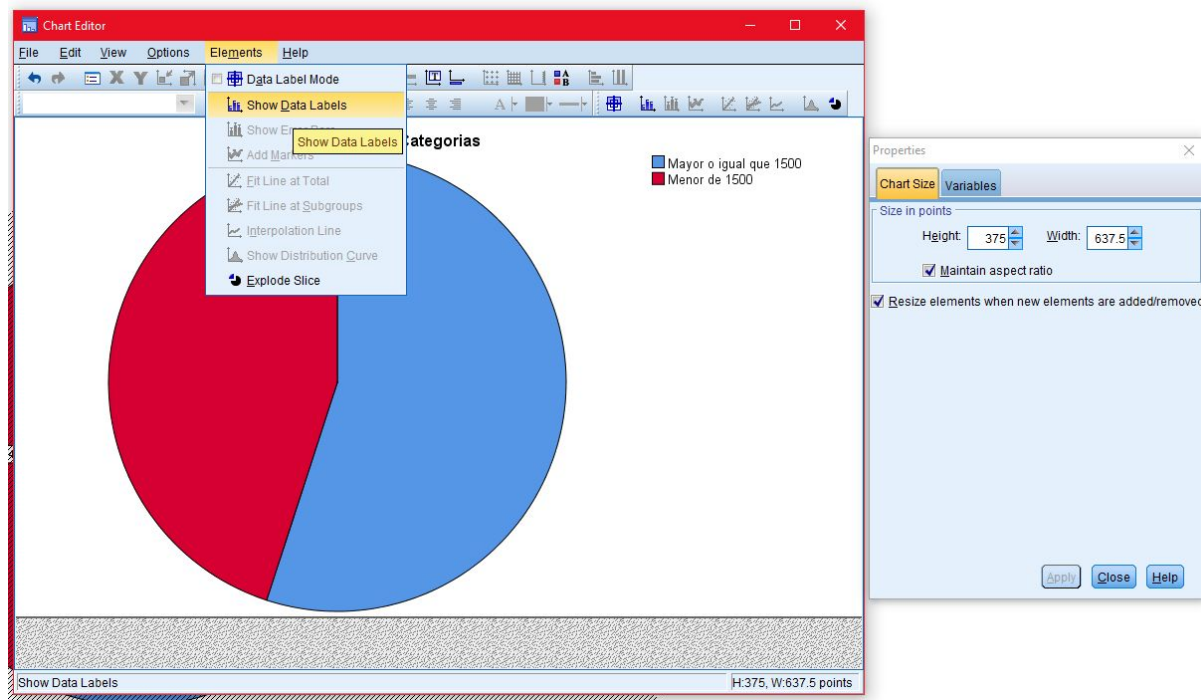


	X1	X2	X3	X4	X5	X6
1	1500	6	2	4	810	Mayor o igual que 1500
2	1989	4	1	6	1340	Mayor o igual que 1500
3	2900	3	5	6	2000	Mayor o igual que 1500
4	2400	3	5	5	1920	Mayor o igual que 1500
5	1200	1	8	6	2325	Menor de 1500
6	1200	1	2	3	1020	Menor de 1500
7	1999	4	4	4	934	Mayor o igual que 1500
8	1050	3	5	4	1570	Menor de 1500
9	2300	4	7	3	1098	Mayor o igual que 1500
10	825	1	2	4	945	Menor de 1500
11	3995	5	4	3	1299	Mayor o igual que 1500
12	1999	2	6	4	1985	Mayor o igual que 1500
13	2085	2	8	7	2020	Mayor o igual que 1500
14	1020	4	3	3	955	Menor de 1500
15	994	1	1	5	800	Menor de 1500
16	1375	1	3	4	1195	Menor de 1500
17	300	2	6	5	2145	Menor de 1500
18	945	1	1	3	999	Menor de 1500
19	1999	5	6	5	1585	Mayor o igual que 1500
20	2284	3	5	4	1720	Mayor o igual que 1500

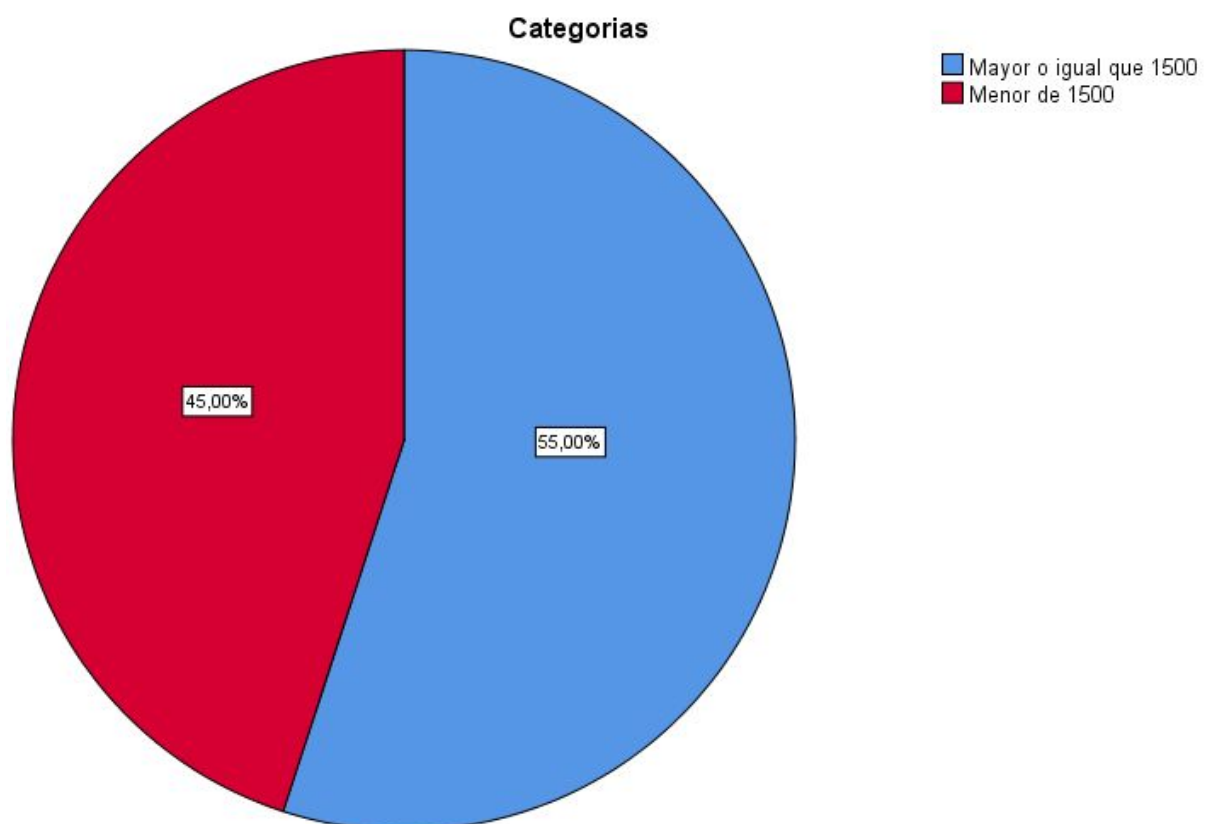
Procedemos a crear el gráfico:



Para incluir los valores (el %) en el gráfico, hacemos doble click encima de él, para abrir el editor de gráficos, vamos al submenú “Elementos” y pinchamos en “Mostrar datos”, los cambios se aplicarán automáticamente.



Al cerrar el editor, podremos observar el gráfico con los datos sobre el, teniendo un modelo más descriptivo que el que viene por defecto.



g) ¿Qué conclusión global puede extraerse de este estudio?

Que hay pocos ordenadores a la venta con muchos periféricos, que la mayoría de ordenadores a la venta son Macs, y que el precio de venta de los ordenadores es bastante caro