

Tema 6 – Unidad 1:

Tipos de variables

Tipos de gráficos de 1, 2 y 3 variables

- 1.1 Tipos de variables**
- 1.2 Gráficos de 1, 2 y 3 variables**
- 1.3 Edición de los gráficos resultantes**

Interacción y Visualización de la Información

Pablo.Bermejo@uclm.es

Tipos de variables y gráficos

1



Introducción

- Solo conocemos un número limitado de gráficos.
- Herramientas como SPSS y Excel nos ofrecen una amplia gama de representaciones de nuestros datos.
- En este tema nos centraremos en los gráficos capaces de mostrar de 1 a 3 variables.

Tipos de variables y gráficos

2

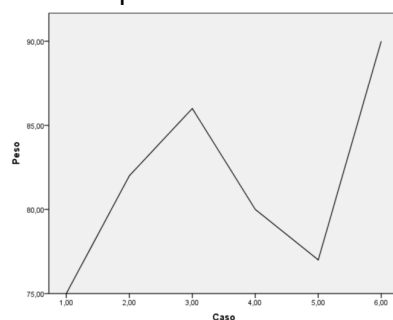
1.1 Tipos de Variables

1.1.1 Introducción

- Un conjunto de **datos** es una serie de mediciones objetivas que no aportan nada por sí mismas.

75
82
86
80
77
90

- Cuando los datos son representados, ordenados y/o contextualizados, entonces obtenemos **información**:
 - La media de peso en varones es 81.67 kg.

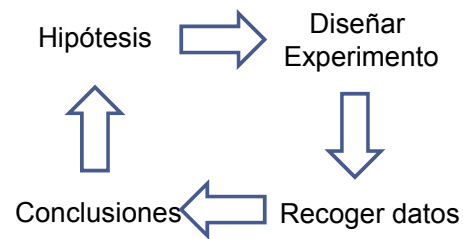


1.1.1 Introducción

- A partir de la información, podemos inferir **conocimiento** si:
 - Comparamos con otra información
 - La media de varones ha aumentado respecto a la década anterior
 - Predecimos mediciones futuras
 - Se calcula que en la próxima década el peso aumentará una media de 1.3 kg.
 - Realizamos conexiones entre otras informaciones similares
 - La media de peso en mujeres es de 10 kg menos que la media de peso en los varones



Davenport y Prusak (1999)



Método científico

1.1.2 Tipos de variables

- Antes de recoger una muestra, es necesario decidir qué tipos de datos son más adecuados para nuestras intenciones.
- El programa IBM SPSS es muy sensible al tipo de datos, y es necesario indicarle correctamente qué tipo representa cada columna.
- Los datos se recogen para una variable, que es la característica que interesa ser estudiada.
- Tipos de datos o variables:
 - **Cualitativa = Categórica = Discreta**: dividen la variable en un conjunto concreto de posibles valores.
 - **Cuantitativa o Numérica**: valores pertenecientes al dominio \mathbb{N} , \mathbb{Z} o \mathbb{R} .

1.1.2 Tipos de variables

- **Variables Cualitativas**
- **Nominales:** los posibles valores no tienen semántica de orden, y por lo tanto su orden es arbitrario:
 - Color = {Rojo, Verde, Amarillo}
 - Sexo = {Mujer, Hombre}
- **Ordinales:** los valores pueden ser ordenados según algún criterio dependiente de cada caso.
 - Fiebre = {No, Febrícula, Fiebre}
 - Dolor Abdominal = {No, Poco, Normal, Mucho}

1.1.2 Tipos de variables

- **Variables Numéricas**
 - Números naturales: 0,1,2,...
 - Números enteros: -5,3,8, ...
 - Números reales: -2.6, 2.9, 5.8...
 - NO se suelen utilizar números imaginarios para inferencias estadísticas.
- En ocasiones es cuestión nuestra decidir cómo tratar una variable:
 - Paquetes tabaco/día = 0, 0.5, 1, 1.5,... ¿Ordinal o Numérica?
- Toda variable numérica puede convertirse en categórica:
 - Dependiendo del caso y del algoritmo utilizado, en ocasiones se obtienen *mejores* resultados con la versión discreta que con la numérica.

1.1.2 Tipos de variables

- ...pero al discretizar perdemos granularidad en la información:
 - Edad numérica: valor natural entre 0 y 100.
 - Edad discreta = {<20, [20-39], [40-60], >60}
- Cuando la variable es numérica, **se recomienda crear otra variable** con su versión discreta, y experimentar con ambas versiones de la variable.
- Discretización manual vs. Automática:
 - Manual: el experto puede aconsejar los rangos de valores más recomendables
 - Nivel de PSA: los rangos de interés para el cáncer de próstata son: <3, [3-6], (6-11], >11
 - Automática: indicando otra variable de interés *C*, existen algoritmos que calculan los rangos más informativos en los que se puede dividir la variable *A* para predecir el valor de *C*.

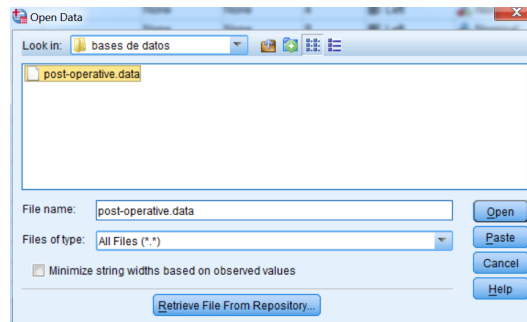
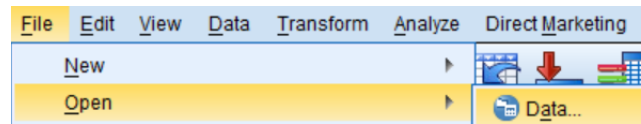
1.1.3 Inserción de datos en SPSS

- En <http://archive.ics.uci.edu/ml/datasets.html> disponéis de un amplísimo repositorio con distintos tipos de bases de datos con los que trabajar.
- Vamos a introducir en SPSS la base de datos: *Post-Operative Patient*:
 - 80 pacientes en sala de recuperación tras una operación.
 - 7 variables descriptivas, más una octava que dice adónde fue enviado el paciente tras pasar unos días: cuidados intensivos, a casa, habitación general.
 - Puede obtenerse de <http://archive.ics.uci.edu/ml/datasets/Post-Operative+Patient>

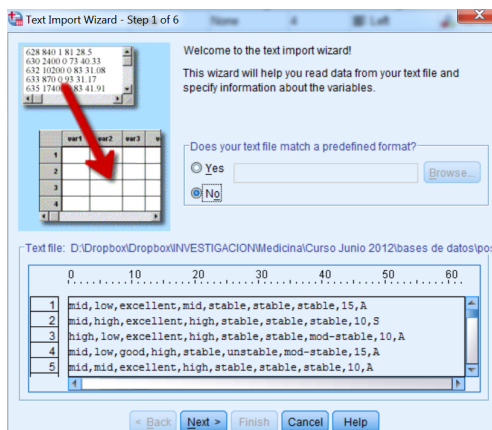
1.1.3 Inserción de datos en SPSS

• 1) Importar base de datos

- La base de datos tiene un formato (.data) no reconocible por SPSS, pero éste dispone una herramienta de importación con la que indicarle cómo importar los datos.



1.1.3 Inserción de datos en SPSS - Importar



How are your variables arranged?

☒ Delimited - Variables are delimited by a specific character (i.e., comma, tab).

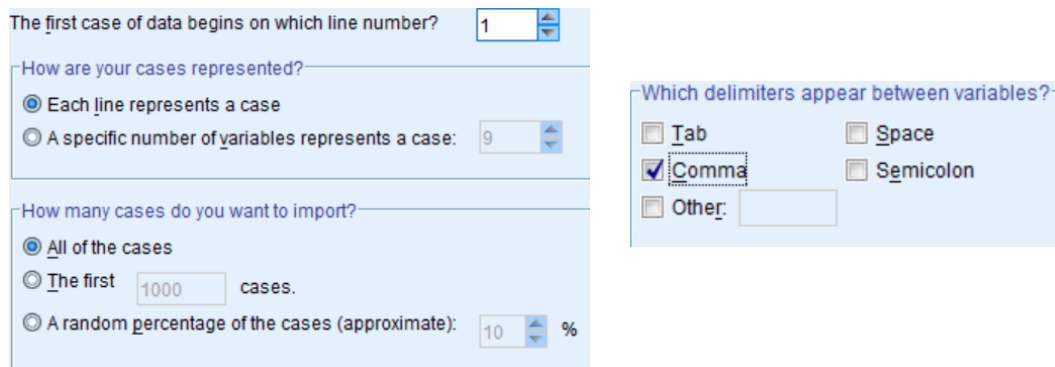
☐ Fixed width - Variables are aligned in fixed width columns.

Are variable names included at the top of your file?

☐ Yes

☒ No

1.1.3 Inserción de datos en SPSS - Importar



The first case of data begins on which line number? 1

How are your cases represented?

- ☒ Each line represents a case
- ☐ A specific number of variables represents a case: 9

How many cases do you want to import?

- ☒ All of the cases
- ☐ The first 1000 cases.
- ☐ A random percentage of the cases (approximate): 10 %

Which delimiters appear between variables?

- ☐ Tab
- ☐ Space
- ☒ Comma
- ☐ Semicolon
- ☐ Other:












En los siguientes 2 pasos hacer Next y Finish, ya que introduciremos los tipos de datos manualmente.

1.1.3 Inserción de datos en SPSS - Tipos

- **2) Especificar el tipo de cada variable**
- SPSS divide nuestros datos en 2 vistas:
 - **Data View**, o vista de datos: el valor de cada variable (columnas) en cada caso (filas)
 - **Variable View**, o vista de variables: vista que nos permite indicar el tipo de variable, nombre, y posibles valores para las categóricas.
 - **Name:** nombre que queremos darle a la variable
 - **Type:** cómo se trata la variable. Si se harán tests, proporciones, etc,... tiene que ser numérica.
 - **Numeric:** variable numérica
 - **Comma:** variable numérica que usa una coma cada 3 cifras, y el punto para indicar decimales.
 - **Dot:** al revés que el anterior
 - **Scientific notation:** variable numérica con el exponente E
 - **Date:** una fecha
 - **Custom currency:** variable numérica que indica una moneda definida
 - **String:** tipo string
 - **Width y Decimals:** anchura máxima y número de decimales permitidos.

1.1.3 Inserción de datos en SPSS - Tipos

- **Label:** descripción personal de la variable
- **Values:** valor numérico que codifique cada valor discreto
- **Missing:** valor o carácter que representa un valor perdido
- **Measure:** cómo se trata la variable
 - **Scale:** variable numérica continua
 - **Ordinal:** cuidado, el orden por defecto es alfabético
 - **Nominal:** variable nominal
- **Role:** marcar como Target si es una variable objetivo o clase a predecir.
- Según Type y Measure, SPSS representa la variable de distintas formas:

Measurement Level	Data Type			
	Numeric	String	Date	Time
Scale (Continuous)		n/a		
Ordinal				
Nominal				

1.1.3 Inserción de datos en SPSS - Tipos

- Configurar desde la Vista de Variables el siguiente formato para la base de datos importada:

Nombre	Valores
TempInterna	{high, mid, low}
TempSuperficial	{high, mid, low}
PorcentajeOxígeno	{excellent, good, fair, poor}
PresionArterial	{high, mid, low}
EstabilidadTempSuperficial	{stable, mod-stable, unstable}
EstabilidadTempInterna	{stable, mod-stable, unstable}
EstabilidadPresionArterial	{stable, mod-stable, unstable}
ComodidadDelPaciente	[0-100]
Decision	{I, S, A}
	I= UCI
	S= a casa
	A = habitación general

1.1.3 Inserción de datos en SPSS - Tipos

- ¡Ojo!: al insertar variables ordinales, es conveniente insertarlas primero con valores numéricos, y después codificar con su nombre textual. Si se hace al revés, SPSS supone que el orden de los posibles valores es por orden alfabético.
- Así que si sucede como en este caso, en el que la bbdd viene codificada textualmente, hay que convertirla previamente a números (opción buscar→reemplazar todos... en cualquier aplicación como el bloc de notas). Para más tarde recodificarla en SPSS.

1.1.3 Inserción de datos en SPSS - Tipos

- Al seleccionar **View→Value Label**, podemos cambiar de vista original a codificada.

	TempInter...	TempSuperficial	PorcentajeOxígeno	PresiónArterial	EstabilidadTempSuperficial	EstabilidadTempInterna	EstabilidadPresiónArterial	ComodidadDelPaciente	Decision
1	mid	low	excellent	mid	stable	stable	stable	15,0	A
2	1	high	3	high	2	stable	stable	10,0	S
3	high	low	excellent	high	stable	stable	mod-stable	10,0	A
4	mid	low	good	high	stable	unstable	mod-stable	15,0	A
5	mid	mid	excellent	high	stable	stable	stable	10,0	A
6	high	low	good	mid	stable	stable	unstable	15,0	S

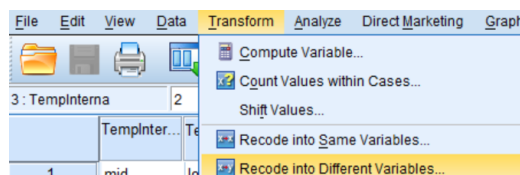
	TempInter...	TempSuperficial	PorcentajeOxígeno	PresiónArterial	EstabilidadTempSuperficial	EstabilidadTempInterna	EstabilidadPresiónArterial	ComodidadDelPaciente	Decision
1	1	0	3	1	2	2	2	15,0	2
2	1	2	3	2	2	2	2	10,0	1
3	2	0	3	2	2	2	1	10,0	2
4	1	0	2	2	2	0	1	15,0	2
5	1	1	3	2	2	2	2	10,0	2
6	2	0	2	1	2	2	0	15,0	1

1.1.3 Inserción de datos en SPSS - Discretizar

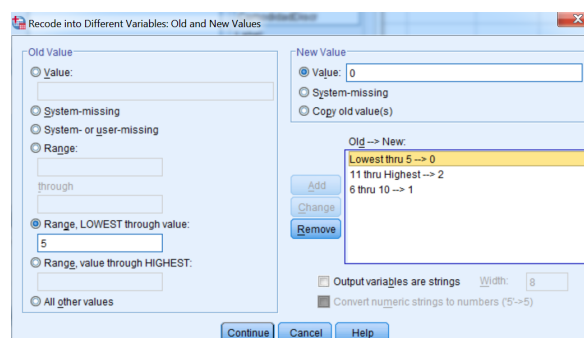
- **3) Discretizar valores numéricos**
- Es conveniente tener una versión categórica de nuestras variables numéricas:
 - Interpretación más sencilla
 - En ocasiones ayuda en la predicción o pruebas estadísticas.
- Se pueden discretizar de forma manual o automática

1.1.3 Inserción de datos en SPSS - Discretizar

- **Discretización manual:**

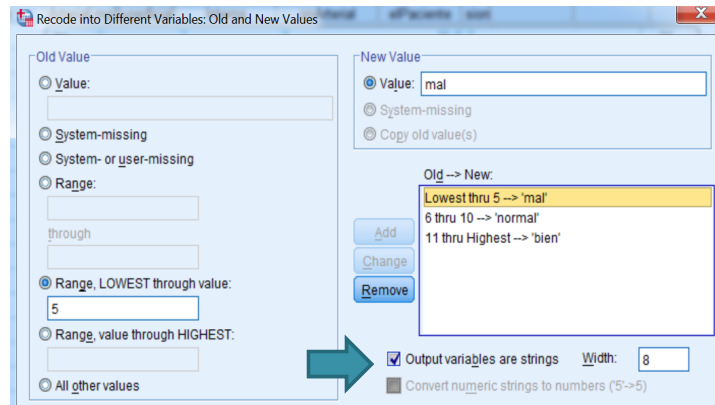


Seleccionamos la única variable numérica, y añadimos los siguientes intervalos:



1.1.3 Inserción de datos en SPSS - Discretizar

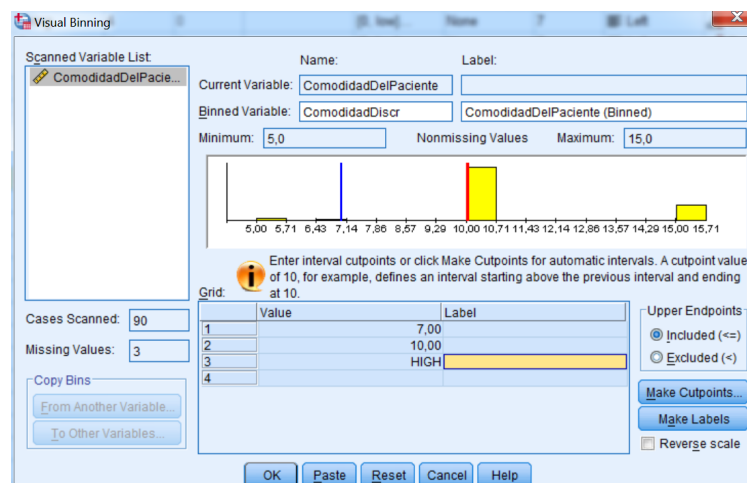
- Ahora volvemos a Vista de Variables, y codificamos 0 como mal, 1 como normal, y 2 como bien.
- También se podían haber indicado los valores textuales indicando que son Strings en la ventana anterior:



- Pero esto provocaría que el orden de los valores sea alfabético, lo cual no es deseable.

1.1.3 Inserción de datos en SPSS - Discretizar

- **Discretización semi-automática:** en **Transform→Visual Binning**, podemos ver un histograma de la variable seleccionada,

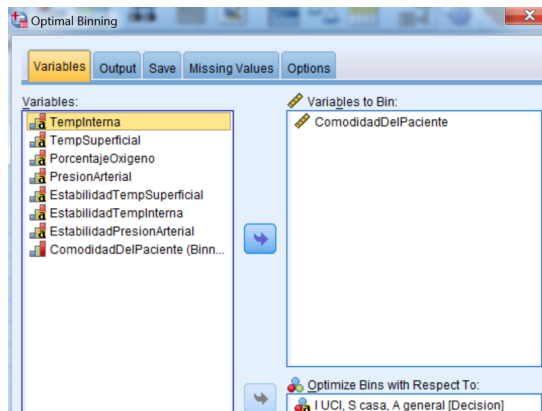


- HIGH representa el mayor valor presente, en este caso 15.
- Los puntos creados pueden moverse con el ratón

1.1.3 Inserción de datos en SPSS - Discretizar

- Pulsando el botón **Make Cutpoints**, podemos seleccionar criterios para que los intervalos se creen automáticamente:
 - Por igual anchura (seleccionando 2 parámetros, el tercero es automático)
 - Por igual frecuencia: di el número de bin. Es mejor que os creéis una variable con una distribución que incluya más valores que ComodidadDelPaciente.
- **Discretización automática**

En **Transform → Optimal Binning**, podemos seleccionar una variable objetivo respecto de la cual optimizar los intervalos a crear para la variable numérica.



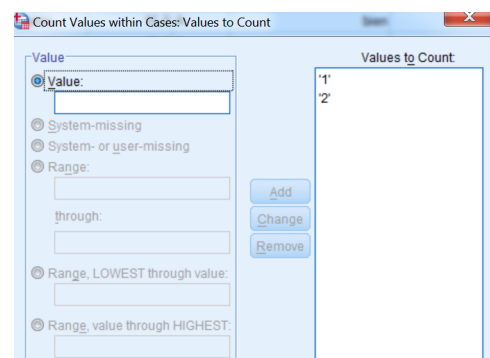
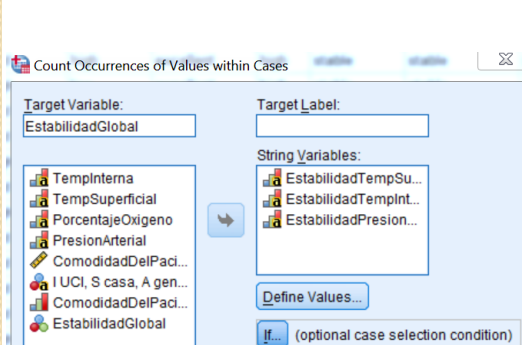
Si la variable objetivo no está correlacionada con la variable numérica, puede no encontrarse una división óptima, y entonces el proceso no se realiza, como ocurre en este caso.

1.1.3 Inserción de datos en SPSS - Crear

- **4) Crear variables nuevas:** es posible que necesitemos calcular variables producto de otras, como suma de factores de riesgo, conjunto de condiciones cumplidas,...

Transform → count values with cases, nos sirve para contar cuántas **variables categóricas** en cada caso tienen el valor requerido.

- Para las variables *EstabilidadTempInterna*, *EstabilidadTempSuperficial* y *EstabilidadPresionArterial*, calcular para cada caso cuántas de estas variables tienen al menos un valor de 1 ('normal').



1.1.3 Inserción de datos en SPSS - Crear

- **Transform→Compute Variable** nos permite realizar operaciones matemáticas sobre variables numéricas
- Ej: disponiendo del peso y altura, calcular el IMC

Target Variable:		Numeric Expression:
IMC	=	peso / (altura * altura)

- En nuestra bbdd solo disponemos de 1 variable numérica, podéis practicar con vuestras propias bbdds.

1.2 Gráficos de 1, 2 ó 3 dimensiones

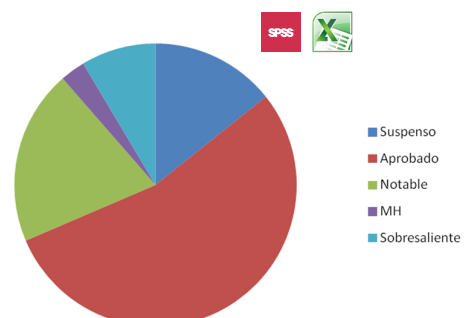
1.2.1 Introducción

- **Visualización de la Información** es el uso de representaciones visuales de un conjunto de datos en un formato de 2 dimensiones (papel, pantalla), con el objetivo de amplificar nuestro conocimiento acerca del mismo:
 - Descubrir conocimiento inesperado
 - Profundizar en lo esperado
- Normalmente un conjunto de datos es un grupo de registros descritos cada uno por ***n* variables**:
 - $n \leq 3$ representación directa
 - $n > 3$ técnicas de visualización multivariada; su comprensión no es directa y requieren entrenamiento.

1.2.2 Tipos de Gráficos - circular

- **Diagrama de sectores, pastel o circular:**
 - Representar porcentajes correspondientes a la **frecuencia** de cada posible valor de una **variable categórica**.
 - El tamaño de cada sector representa el porcentaje, siendo 100% el total
 - Efectivo para comparar proporciones dentro de un mismo diagrama de sectores
 - Inadecuado para comparar entre distintos diagramas de sectores

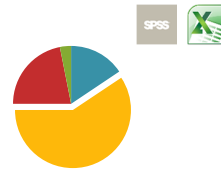
Nota	Cantidad
Suspense	5
Aprobado	19
Notable	7
Sobresaliente	3
MH	1



1.2.2 Tipos de Gráficos - circular

- Seccionado o explotado:

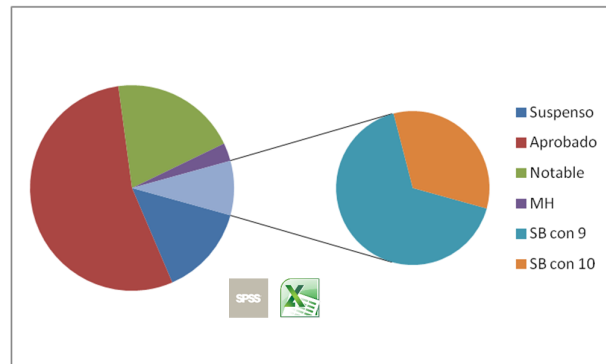
- Separa 1 ó más sectores
- Resaltar una proporción



- Con subdiagrama circular:

- Permite desglosar una variable en varias

Nota	Cantidad
Suspenso	5
Aprobado	19
Notable	7
MH	1
SB con 9	2
SB con 10	1



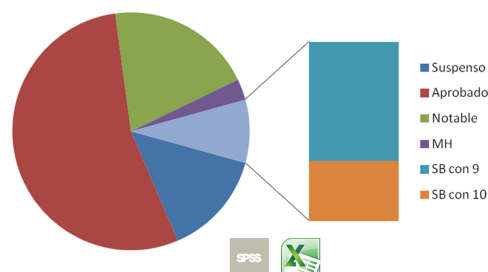
Tipos de variables y gráficos

29

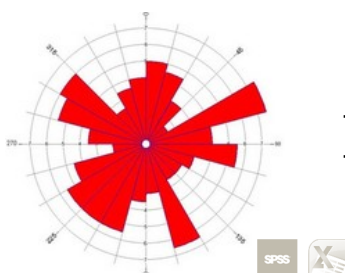
1.2.2 Tipos de Gráficos - circular

- Con subdiagrama de barras:

- Igual que el anterior, siendo el subgráfico una diagrama de barras que también representa porcentaje.



- De área polar:



- Mismo ángulo
- La cantidad se representa por la altura

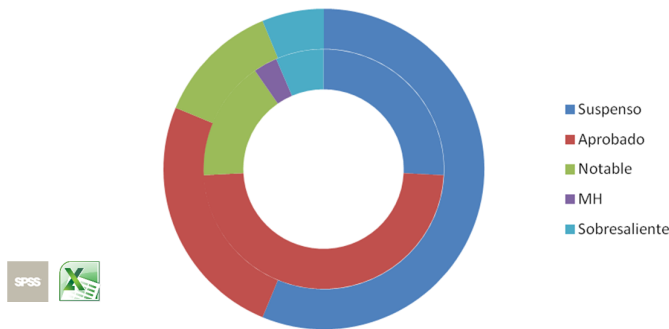
Tipos de variables y gráficos

30

1.2.2 Tipos de Gráficos - circular

- Diagrama de anillo
 - Como el diagrama de sectores, pero añade una segunda variable categórica como referencia: notas en Julio y Septiembre.

	Suspense	Aprobado	Notable	MH	Sobresaliente	
Julio		8	15	5	1	2
Septiembre		9	4	2	0	1



'Han suspendido en mucha mayor proporción los presentados en Septiembre que en Julio'

1.2.2 Tipos de Gráficos - circular

- **SPSS** tiene su gestor de gráficos en **Graphs→ chart builder...**

Estadístico cuyo valor indica el ángulo de los sectores

Indicar posición de los sectores

Indicar orden de los sectores y leyenda

Títulos del gráfico

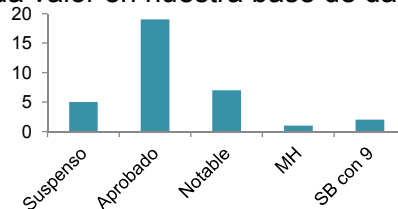
Variable categórica cuyos valores dividirán el pastel

1.2.2 Tipos de Gráficos - barras

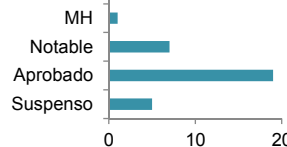
- **Diagrama de barras o columnas:**

- Simples:

- El eje horizontal representa los posibles valores de una **variable categórica**
- El eje vertical representa la **cantidad de veces o frecuencia** que aparece de cada valor en nuestra base de datos.



- Las barras pueden ser horizontales, en cuyo caso los papeles de los ejes horizontal y vertical **se intercambian**.



Tipos de variables y gráficos

33

1.2.2 Tipos de Gráficos - barras

Decidir orden de los valores

Decidir la escala del eje Y

Lineal: normal Log: útil cuando la diferencia de valores es demasiado grande como para apreciarse

Basic Elements → Transpose: para poner las barras horizontales

Tipos de variables y gráficos

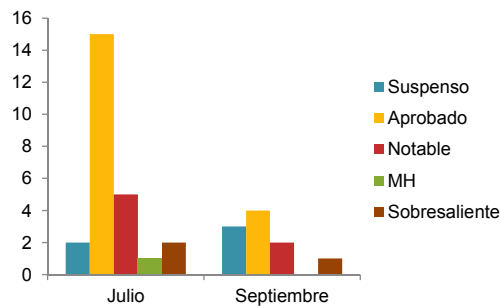
34

1.2.2 Tipos de Gráficos - barras

- Agrupadas o clusterizadas:

- Contrastar una variable categórica en función de otra variable categórica

	Suspense	Aprobado	Notable	MH	Sobresaliente
Julio	2	15	5	1	2
Septiembre	3	4	2	0	1



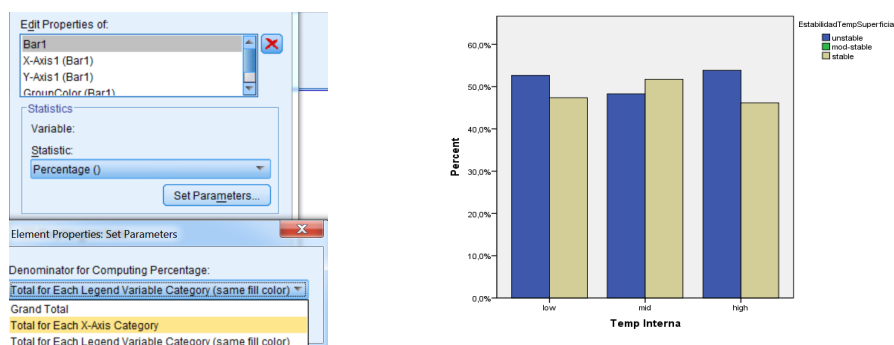
1.2.2 Tipos de Gráficos - barras

Variable que divide cada grupo

Variable que divide el eje x

1.2.2 Tipos de Gráficos - barras

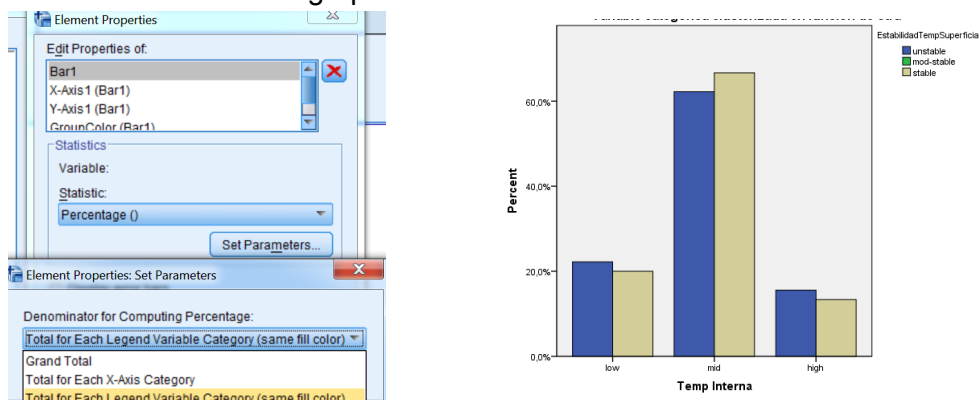
- Si se utilizan **frecuencias en el eje Y** (para elegir *percent* se elimina la variable para la que se indicaba su media), y se quieren comparar las categorías de la variable de agrupación (EstabilidadTempSuperficial) según cada valor de la variable de referencia del eje X (TempInterna), hay que seleccionar que en cada grupo divida entre el número de casos de la categoría correspondiente del eje X.



'Cuando la Temperatura Interna=low, los pacientes inestables son el 52%, y los estables el 48%'

1.2.2 Tipos de Gráficos - barras

- Si se utilizan **frecuencias en el eje Y**, y se quiere comparar cómo se reparten las categorías de la variable de agrupación (EstabilidadTempSuperficial) a lo largo de cada valor de la variable de referencia del eje X (TempInterna), hay que seleccionar que en cada grupo divida entre el número de casos de la categoría correspondiente de la variable de agrupación.

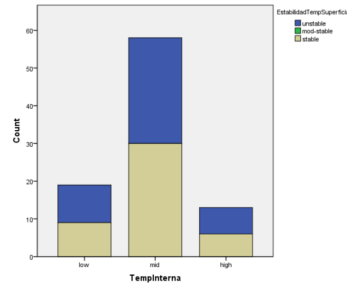


'Casi todos los casos estables se dan cuando la temperatura interna es media'

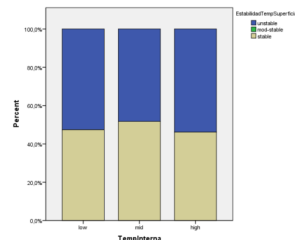
1.2.2 Tipos de Gráficos - barras

- Apiladas o compuestas

- El mismo objetivo que las agrupadas, pero en vez de hacer grupo de barras por cada valor de la variable de referencia, estas se apilan en cada valor de la variable secundaria.



- También se pueden apilar las frecuencias, y el total obviamente es 100:



Tipos de variables y gráficos

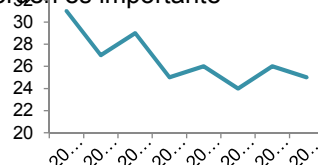
39

1.2.2 Tipos de Gráficos - líneas

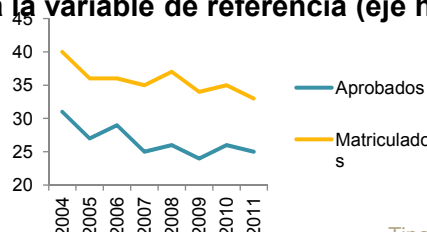
- Diagrama de líneas

- Simples

- Valor de una variable numérica según cambia el valor de otra variable numérica
- Muy útil:
 - Para observar tendencias a lo largo del tiempo en intervalos iguales
 - Cuando el orden es importante



- Se puede observar la evolución de varias **variables (eje vertical)** **respecto a la variable de referencia (eje horizontal)**, todas ellas siendo numéricas.

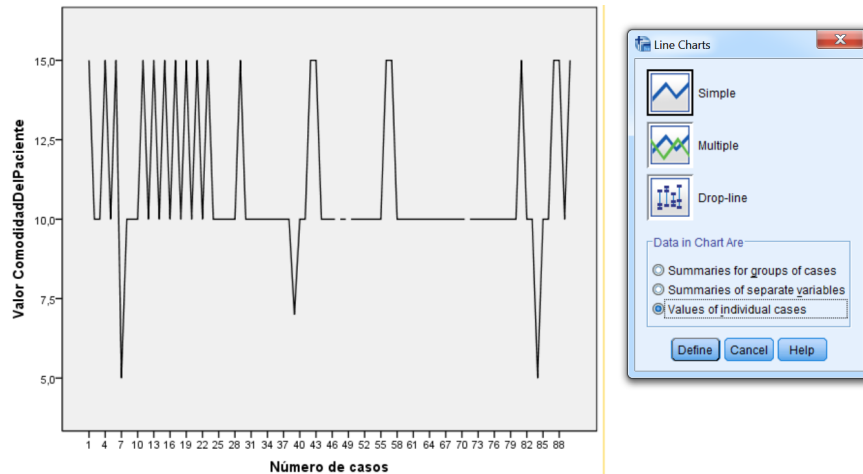


Tipos de variables y gráficos

40

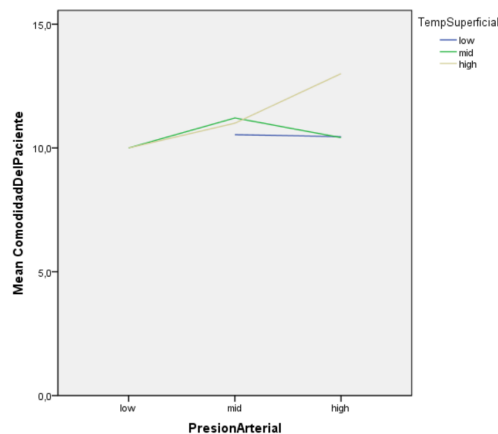
1.2.2 Tipos de Gráficos - líneas

- Para simplemente mostrar los valores de cada caso, hay que ir a **Graphs → Legacy Dialogs → line**, ya que desde el Chart builder no podemos seleccionar un eje X que muestre los casos



1.2.2 Tipos de Gráficos - líneas

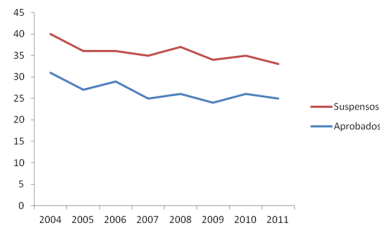
- Con múltiples líneas, se puede indicar un estadístico (Media en este caso), de la variable numérica para cada posible valor de otras 2 categorías, una indicada con leyenda y la otra en el eje x.



1.2.2 Tipos de Gráficos - líneas

- Apiladas

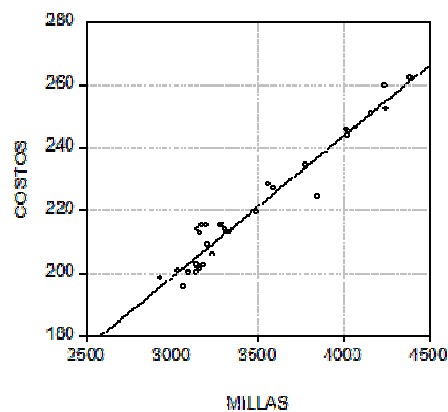
- Muestra la cantidad con que aporta cada variable secundaria en cada punto de la variable de referencia
- Estas no son de interés, ya que las barras apiladas son más sencillas de leer y representan lo mismo



1.2.2 Tipos de Gráficos - líneas

- **Diagrama de dispersión**

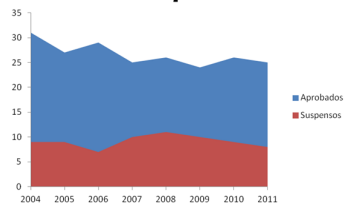
- Son más útiles que las líneas cuando hay muchos puntos próximos y queremos marcarlos en vez de unirlos.
- Además de tendencias, pueden mostrar **patrones o clusters**.
- Si se imprimen sobre una recta de regresión, nos indican el grado de error cometido por la ecuación.



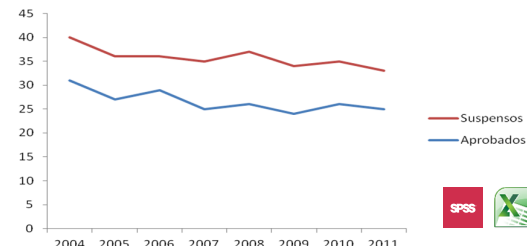
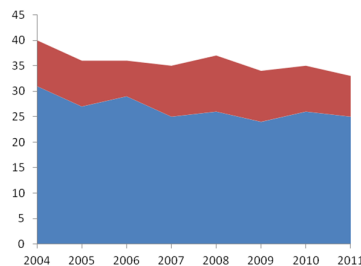
1.2.2 Tipos de Gráficos - áreas

- **Diagrama de áreas**

- Misma función que el de líneas, pero **puede dar la falsa impresión de que las variables están apiladas**. No son recomendables.



- Áreas apiladas



- También se pueden apilar frecuencias, como en las anteriores.

1.2.2 Tipos de Gráficos - histogramas

- **Histogramas**

- Representación con barras verticales de **una variable discreta**
- Si es continua, se puede discretizar.
- Para cada rango, una columna vertical con altura que indica **frecuencia o cantidad de veces** en las que ese rango de valores se da en la base de datos.
- Con SPSS mucho más sencillo que en Excel

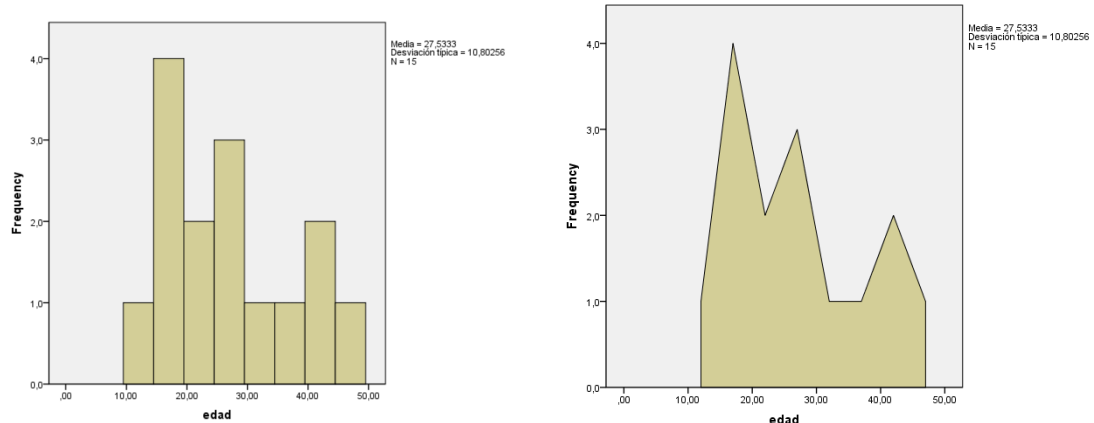


- Crear una variable 'edad' de tipo Scale con los siguientes valores

edad
18,00
35,00
29,00
40,00
23,00
19,00
43,00
26,00
12,00
18,00
15,00
29,00
33,00
24,00
49,00

1.2.2 Tipos de Gráficos - histogramas

- Para construir el histograma, podemos discretizar la variable previamente, como hemos visto antes.
- Ó dejamos a SPSS que decida automáticamente los intervalos.
- Si nos quedamos solo con la forma, esto se conoce como Polígono de Frecuencias.



Tipos de variables y gráficos

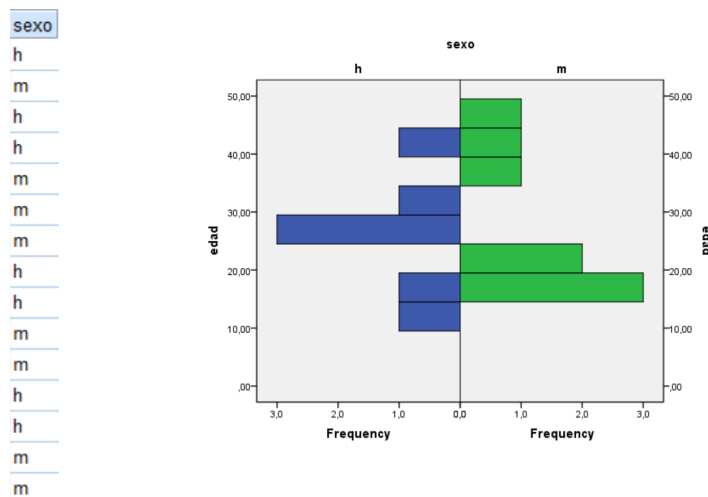
47

1.2.2 Tipos de Gráficos - histogramas

- [Histograma piramidal](#)  

Muestra el valor de la variable de interés, según las 2 categorías de otra variable de referencia.

Añadiendo una variable nominal con los siguientes valores:



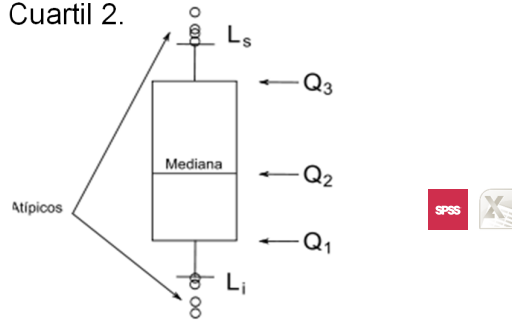
Tipos de variables y gráficos

48

1.2.2 Tipos de Gráficos – caja y bigotes

• Diagrama de Caja y Bigotes

- Resumen estadístico de la distribución y dispersión de una variable
- El tope de la caja representa el Cuartil 3 (los valores inferiores comprenden el 75% de los datos), y la base el Cuartil 1. El centro es la mediana o Cuartil 2.

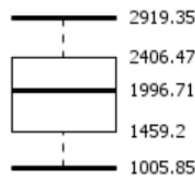


- Los bigotes (barras verticales) representan los límites superior e inferior de los datos, excepto aquellos casos cuyo valor supere:
 - $1.5 \times \text{IQR}$: atípicos
 - $3 \times \text{IQR}$: outliers

1.2.2 Tipos de Gráficos – caja y bigotes

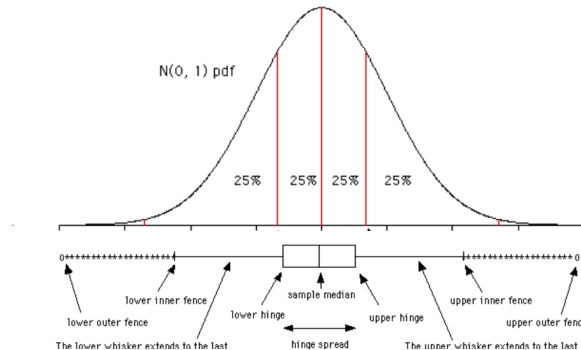
+10345.67

Citizen Nr.	Income	Key Value
24	10,345.67	Maximum
23	2,919.35	Upper Bound
22	2,787.02	
21	2,784.72	
20	2,696.83	
19	2,412.51	
18	2,400.43	Q3: $(18. + 19.) / 2 = 2,406.47$
17	2,367.84	
16	2,333.37	
15	2,285.53	
14	2,214.87	
13	2,069.79	
12	1,923.62	Median: $(12. + 13.) / 2 = 1,996.71$
11	1,819.22	
10	1,773.34	
9	1,597.54	
8	1,589.48	
7	1,494.65	
6	1,423.74	Q1: $(6. + 7.) / 2 = 1,459.2$
5	1,391.92	
4	1,334.88	
3	1,184.53	
2	1,125.78	
1	1,005.85	Minimum / Lower Bound



1.2.2 Tipos de Gráficos – caja y bigotes

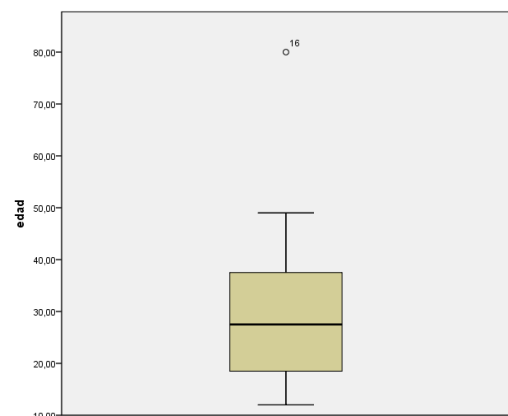
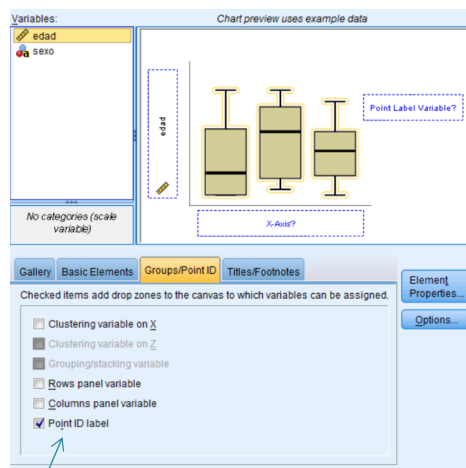
- Cuando la mediana está por el centro de la caja y los bigotes tienen una extensión similar, se dice que los datos son más o menos **simétricos (distribución normal)**.



- Entre la mediana y un cuartil, hay el mismo número de casos (25%) que entre la mediana y el otro cuartil; por eso, cuanto más pegada esté la mediana a un cuartil respecto del otro, significa que la dispersión de datos en ese rango de valores es menor.

1.2.2 Tipos de Gráficos – caja y bigotes

- Si en nuestra variable edad añadimos un caso con 80 años, éste será considerado como un outlier.



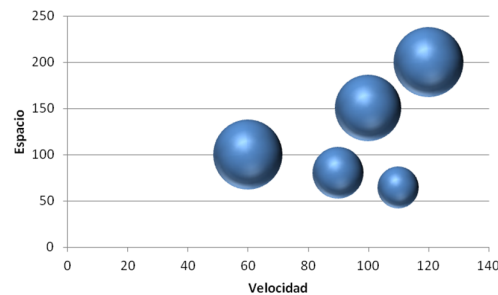
Para identificar el caso outlier

1.2.2 Tipos de Gráficos

• Diagrama de Burbujas

- 3 variables numéricas: ejes X-Y, y el tamaño de la burbuja es proporcional a la tercera variable.
- Una burbuja por registro: no sirve para bases de datos con muchos casos
- ó una burbuja por valor de una cuarta variable categórica

Velocidad	Espacio	Tiempo
80	100	1.25
90	80	0.88888889
100	150	1.5
110	64	0.58181818
120	200	1.66666667



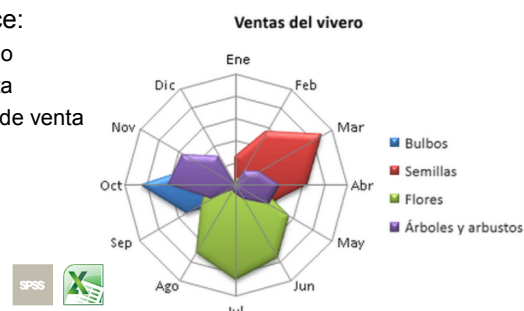
1.2.2 Tipos de Gráficos

• Diagrama radial o de estrella

- Círculo dividido en tantos tramos como valores tenga una variable categórica X
- Una estrella por valor de una variable categórica Y
- La estrella tiene una esquina por cada valor de X que aparece en la base de datos simultáneamente a Y
- El alcance de cada esquina, desde el centro, es proporcional a una variable numérica Z
- → 2 variables categóricas y 1 numérica

• Ejemplo de Office:

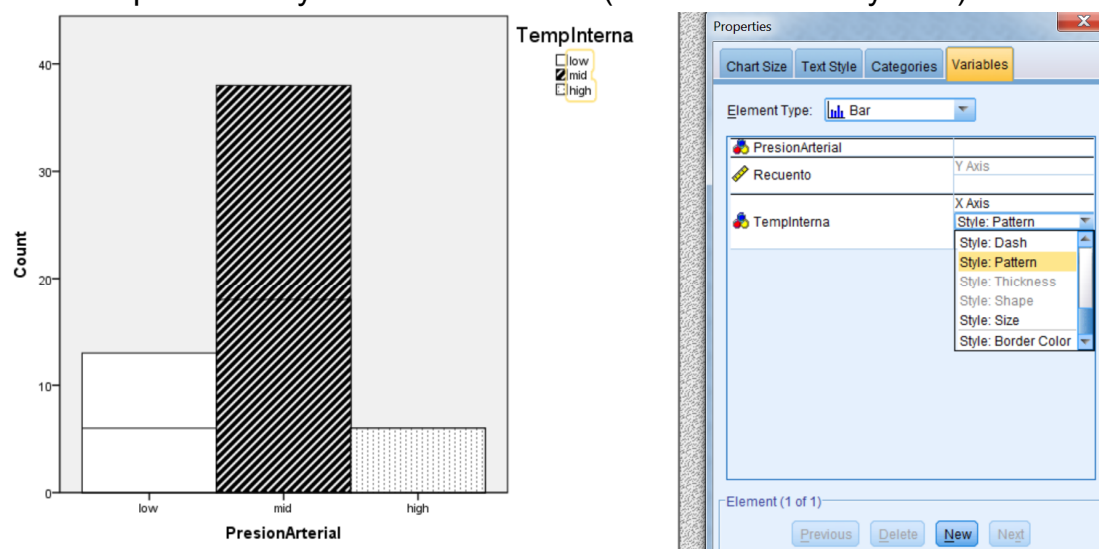
- X → mes del año
- Y → tipo de venta
- Z → frecuencia de venta



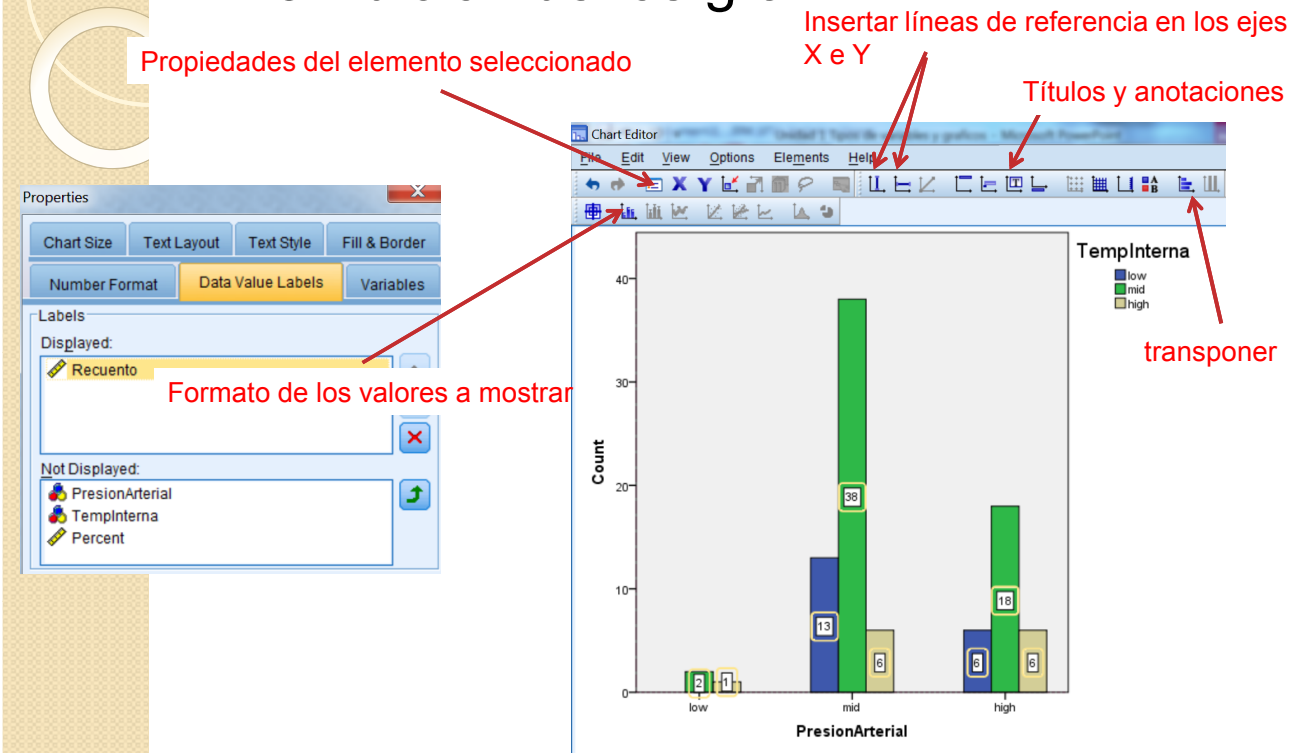
1.3 Edición de los gráficos

1.3 Edición de los gráficos

- Una vez los gráficos han sido creados, el resultado puede editarse haciendo doble clic sobre él.
- [Texturas](#): si vamos a imprimir el trabajo en blanco y negro, tenemos que crear leyendas con texturas (doble clic en la leyenda)



1.3 Edición de los gráficos



Tipos de variables y gráficos

57

Conclusiones

- Tanto Excel como SPSS ofrecen varios tipos de gráficos.
- SPSS gestiona de manera más formal el tipo de variables:
 - Debemos tener paciencia e introducir los datos de la forma más correcta posible
 - Obtendremos resultados erróneos si indicamos mal el tipo de variable
- Gráficos más avanzados como Caja y Bigotes, o Histogramas, son más sencillos de crear utilizando SPSS.
- La edición de los gráficos es más sencilla con Excel, aunque en las 2 herramientas todo es configurable.

Tipos de variables y gráficos

58



PRÁCTICAS

Selecciona la base de datos que desees de <http://archive.ics.uci.edu/ml/datasets.html> , o utiliza la tuya propia y haz lo siguiente:

- 1) Configura el tipo de cada variable
- 2) Crea gráficos de distinto tipo
- 3) Edítalos para impresión en blanco y negro, y escribe comentarios sobre ellos