

CONSTRUCCIÓN Y CARACTERÍSTICAS DE LAS TABLAS TIPO B

Este tipo de tablas, llamados también tabla para datos agrupados, suelen ser utilizadas cuando el número de resultados posibles que puede obtener una variable son tan amplios, que una Tabla Tipo A haría muy poco en resumirlos (estos datos representan un rango mayor a 15 aunque en este cursado lo tomaremos mayor a 10).

Debido a esta cantidad de valores, será necesario agruparlos mediante intervalos (la estadística los llama “Intervalos de Clases”).

Agrupar los valores de la variable en intervalos podría simplificar estas fuentes de datos. Por ejemplo, podríamos hablar de las frecuencias para los valores comprendidos entre 0-20, 20-40, 40-60, 60-80 y 80-100. En el intervalo 0-20 (que de ahora en adelante le llamaremos intervalo de clase), se sumarán las frecuencias de los datos cuyos resultados estén entre 0 y 20.

Intervalo de clase: Son cada uno de los grupos de datos empleados en las Tablas de Frecuencias Estadísticas, capaz de contener diversas medidas de una variable cuantitativa discreta o continua. Consta de un límite inferior (Li) y un límite superior (Ls).

Otro punto importante que en estos casos se debe definir, es la cantidad de intervalos de clase que se empleará en la tabla de frecuencia. Esta cantidad de intervalos no deberían ser muchos, debido a que no se cumpliría el objetivo de resumir la información, y no tan pocos intervalos, ya que se perdería mucha información.

No existe una fórmula, ni unos principios únicos para establecer el número de intervalos. En este cursado, optaremos por manejar un número de intervalos convenientes entre 5 y 10.

Numero de intervalos (Ni): Cantidad de intervalos con los cuales se compone una tabla de frecuencia.

Algunos autores han propuestos fórmulas que permiten ayudar en la tarea de conseguir el número ideal de intervalos.

1) La más conocida, establece el número de intervalos al obtener la raíz cuadrada del total de elementos considerados en el estudio.

$$N_i = \sqrt{n}$$

2) Cuando se trabajan con muestras mayores o iguales a 100, la fórmula obtiene un Ni superior o igual a 10, por lo tanto, recomendaremos para estos casos la siguiente formula:

$$N_i = 1 + 3,22 \log n$$

Cada intervalo posee un número máximo de resultados que puede agrupar. A este valor lo conoceremos como el “Ancho del Intervalo de Clase (A)”.

Ancho del intervalo de Clase (A): Su cálculo resulta de la división del Rango R entre el Número de Intervalos (N_i)

$$A = \frac{R}{N_i}$$

Equivale a la diferencia entre el Límite superior (Ls) y el Límite inferior (Lm) de cada intervalo. Matemáticamente se expresa:

$$A = Ls - Lm$$

Hay que aclarar, que el ancho puede variar entre los intervalos, pero por razones estéticas, comprensión y para facilitar el análisis, se recomienda manejar un ancho equivalente en todos los intervalos de clase.

Ejemplo tabla de frecuencia tipo B

Un sondeo realizado en el barrio de la ciudad de Resistencia sobre 30 mujeres con hijos, pretende hacer un estudio de las edades de las mismas cuando fueron madres por primera vez, y hallar la edad más representativa.

Las edades de las mujeres fueron:

17	17	19	19	31	21
18	27	21	22	24	19
25	24	24	23	20	29
21	19	22	21	20	20
19	19	23	20	21	21

Cuando recolectamos los datos por primera vez, estos se presentarán en forma confusa, lo que amerita ser ordenada.

A continuación, se detallan los siguientes pasos para la construcción de la tabla de frecuencia tipo B.

✚ **Paso 1: Siempre debemos calcular el rango R como primera medida. Este nos indica si nuestra tabla será de tipo A o B**

$$R = X_{max} - X_{min} \rightarrow R = 31 - 17 = 14$$

✚ **Paso 2: Determinar el número de intervalos (N_i)**

$$N_i = \sqrt{n} \rightarrow N_i = \sqrt{30} = 5,477 \cong 6 \text{ Intervalos}$$

Se debe siempre aproximar el número de intervalos al entero superior más próximo, recordando que este valor no será menor a 5, ni un valor mayor a 10. Nuestra tabla estará constituida entonces por 6 (seis) intervalos.

✚ **Paso 3: Determinar el ancho de cada intervalo.**

Con el Rango y el número de intervalos, podremos hallar el ancho:

$$A = \frac{R}{N_i} = \frac{14}{6} = 2,333$$

Aproximamos al entero superior $A = 3$ El ajuste del Ancho no podrá ser menor al valor obtenido inicialmente.

 Paso 4: Determinar el nuevo Rango (R')

En el momento de realizar el ajuste del ancho del intervalo, el rango se incrementa automáticamente. Este "Nuevo Rango" lo denotaremos como R'

El rango para datos agrupados es $R = Ls - Li$, para determinar el limite inferior del primer intervalo y el limite superior del ultimo intervalo debemos utilizar un nuevo rango.

Nuevo Rango (R'): Rango que es convenido por el Ancho de los intervalos a los decimales que son manejados en los datos objeto del estudio. Su cálculo se realiza multiplicando el Ancho ajustado por el Número de Intervalos:

$$R' = A \cdot N_i$$

Entonces el cálculo del nuevo rango es:

$$R' = 3 \cdot 6 = 18$$

El rango se incrementó en cuatro años. El incremento entre el nuevo rango (R') y el rango inicial R , se reparte entre el valor mínimo y el valor máximo. Es decir, al valor mínimo le quitamos y al valor máximo le agregamos.

$$R' - R = 18 - 14 = 4$$

$$X_{min}' = 17 - 2 = 15$$

$$X_{max}' = 31 + 2 = 33$$

Si $R' - R$ llegase a ser IMPAR, ejemplo: 5, por convención quitaremos 3 a X_{min} y agregaremos 2 a X_{max}

Este procedimiento permite encontrar los valores mínimos y máximos cuya resta sea igual al nuevo Rango (R')

$$R' = X_{max}' - X_{min}' = 33 - 15 = 18$$

Si el resultado de la diferencia entre el nuevo rango y el rango:

- Si es 0 esta bien el rango porque incluye a todos los datos
- Si es mayor que cero, se debe utilizar el rango nuevo porque hay datos que no se contemplan. Como en el ejemplo anterior-

 Paso 5: Determinar los intervalos de clases iniciales.

Con el ajuste de los valores mínimos y máximos, y el ancho, podremos armar cada intervalo de clase. El primer intervalo parte del valor mínimo (X_{min}), al cual le agregamos el ancho (A). El segundo intervalo parte del límite superior del intervalo anterior y se le agrega el ancho, y así sucesivamente hasta alcanzar el valor máximo (X_{max}).

N_i	Lm	Ls
1	15	18
2	18	21
3	21	24
4	24	27
5	27	30
6	30	33

Presenta algunos inconvenientes al momento de repartir las frecuencias a cada intervalo de clase, por ejemplo, existen 6 personas del total de encuestadas que tienen una edad de 21 años, los cuales podrían ser clasificados en el intervalo dos o en el tres.

Este caso se le conoce como el “Problema de la Ambigüedad”, y el cual debe ser solucionado antes de terminar la tabla de frecuencia tipo B.

Vamos a utilizar nosotros estará referido a los límites abiertos y cerrados. Los límites abiertos se identifican con paréntesis () y significa que el valor que se encuentre anexo al paréntesis, no pertenece al intervalo. Los límites cerrados se identifican con corchetes [] y significa que el valor que se encuentre anexo al corchete pertenece al intervalo.

Según nuestro ejemplo:

N_i	L_m	L_s
2	[18	21)
3	[21	24)

En el 2do intervalo van a estar comprendidos los datos recolectados de las variables desde 18 años hasta casi 21 años, pero no los que corresponden a 21 años. En el 3er intervalo sucede algo similar, este va a estar comprendido por los datos desde 21 años hasta casi 24

Paso 6: Determinar las frecuencias absolutas, frecuencias relativas y sus respectivas acumuladas.

Y para las tablas tipo B se agrega una columna más que estará referida a las "Marcas de clase"

Marcas de Clase (x_i): Se define como el punto medio (promedio) de un intervalo de clase. Es el valor representativo de cada intervalo.

$$x_i = \frac{L_s + L_m}{2}$$

Armemos nuestra tabla

Intervalos y marca de clase

N_i	L_m	L_s	Calculo auxiliar	x_i
1	[15	18)	(15+18):2	16,5
2	[18	21)	(18+21):2	19,5
3	[21	24)	(21+24):2	22,5
4	[24	27)	(24+27):2	25,5
5	[27	30)	(27+30):2	28,5
6	[30	33)	(30+33):2	31,5

Frecuencia absoluta acumulada

N_i	L_m	L_s	x_i	f_i	Calculo auxiliar	F_i
1	[15	18)	16,5	2	-----	2
2	[18	21)	19,5	11	(11+2)	13
3	[21	24)	22,5	10	(13+10)	23
4	[24	27)	25,5	4	(23+4)	27
5	[27	30)	28,5	2	(27+4)	29
6	[30	33)	31,5	1	(29+1)	30
	TOTAL			30		

Frecuencia relativa $h_i = \frac{f_i}{n} \cdot 100\%$

N_i	L_m	L_s	x_i	f_i	F_i	Calculo auxiliar	$H_i \%$
1	[15	18)	16,5	2	2	(2:30).100	6,7%
2	[18	21)	19,5	11	13	(11:30).100	36,7%
3	[21	24)	22,5	10	23	(10:30).100	33,3%
4	[24	27)	25,5	4	27	(4:30).100	13,3%
5	[27	30)	28,5	2	29	(2:30).100	6,7%
6	[30	33)	31,5	1	30	(1:30).100	3,3%
TOTAL				30			100%

Frecuencia relativa acumulada

N_i	L_m	L_s	x_i	f_i	F_i	$h_i \%$	Calculo auxiliar	$H_i \%$
1	[15	18)	16,5	2	2	6,7%	----	6,7%
2	[18	21)	19,5	11	13	36,7%	6,7%+3,7%	43,4%
3	[21	24)	22,5	10	23	33,3%	43,4%+33,3%	76,7%
4	[24	27)	25,5	4	27	13,3%	76,7%+13,3%	90%
5	[27	30)	28,5	2	29	6,7%	90%+6,7%	96,7%
6	[30	33)	31,5	1	30	3,3%	96,7%+3,3%	100%
TOTAL				30		100%		

El máximo valor del último intervalo también se pone el corchete, para de esta manera en los casos que se pudieran dar, este valor también este comprendido en el intervalo.

Características de las tablas tipo B

- El número de posibles valores que toma la variable es elevado. (Rango grande)
- Se utiliza para el tratamiento de variables cuantitativas (discretas y continuas).
- Su construcción es más compleja que en las tablas tipo A.
- La interpretación equivale a especificar la frecuencia de cada intervalo de clase.
- Presenta un componente adicional: las marcas de clase.

GRAFICOS ESTADISTICOS

Los gráficos nos permiten una fácil interpretación y análisis de los datos, al mostrar las frecuencias mediante símbolos, barras, polígonos y sectores.

Gráficos Estadísticos: Son representaciones visuales que emplean símbolos, barras, polígonos y sectores, de los datos contenidos en tablas de frecuencias.

Trataremos seis tipos de gráficos estadísticos:

- Gráficos de sectores
- Gráficos de columnas
- Histogramas
- Polígonos de frecuencia

- Ojivas
- Gráficos de líneas o de tendencia

GRÁFICOS DE COLUMNAS

Los gráficos de columnas representan las frecuencias a través de la altura de las mismas en un plano de ejes cartesianos.

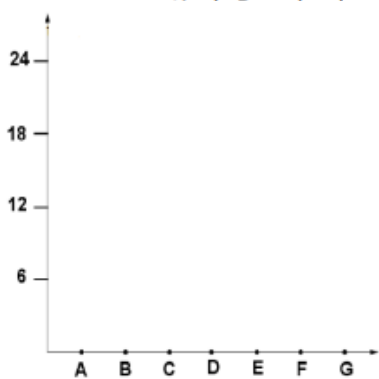
Ejemplo de gráficos de columnas Realizar un gráfico de barras a partir de la siguiente tabla de frecuencia:

Clase	Frecuencia
A	5
B	11
C	11
D	4
E	15
F	18
G	24
TOTAL	88

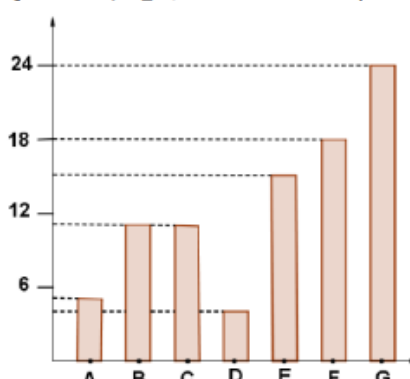
Solución

Para crear un gráfico de columnas, seguiremos dos pasos sencillos:

- **Paso 1:** Representar las escalas en los ejes horizontal y vertical del primer cuadrante de un plano de ejes cartesianos. En el eje vertical colocaremos las frecuencias y en el eje horizontal las clases. Estas clases deben tener un ancho igual para cada una de las columnas que graficaremos.
- **Paso 2:** A cada clase se representa con una columna (o barra) cuya altura concuerda con su frecuencia expuesta en el eje

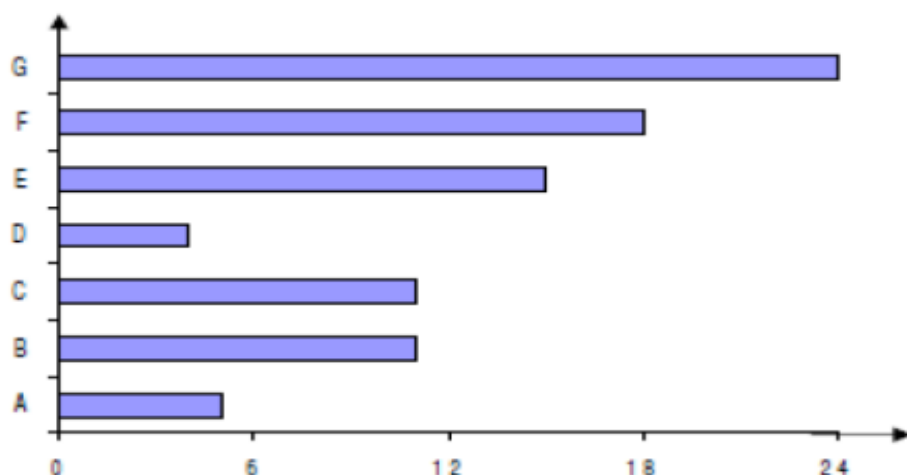


PASO 1



PASO 2

Podríamos realizar el gráfico de forma horizontal (conocido como gráfico de barras):



Características de los gráficos de columnas

- No muestran frecuencias acumuladas.
- Se prefiere para el tratamiento de datos cualitativos o cuantitativos discretos.
- La columna (o barra) con mayor longitud representa la mayor frecuencia.
- Son fáciles de elaborar. Suelen utilizarse para representar tablas tipo A.
- La sumatoria de las alturas de las columnas o las longitudes de las barras equivalen al 100% de los datos.

HISTOGRAMA Y POLÍGONOS DE FRECUENCIA

Se realiza sobre el primer cuadrante del plano cartesiano. La diferencia radica en que el histograma se utiliza más a menudo para representar tablas tipo B, donde el ancho de la columna equivale al ancho del intervalo de clase.

Las frecuencias absolutas se colocan en el eje vertical. Otra diferencia importante es que no existe espacio entre las barras.

Para construir el polígono de frecuencia se trazan líneas rectas entre los puntos medios de los techos de las columnas contiguas, partiendo desde el punto de origen del eje cartesiano (0,0) hasta el punto final definido en el eje horizontal.

Ejemplo de histogramas

Realizar un histograma a partir de la siguiente tabla de frecuencia tipo B:

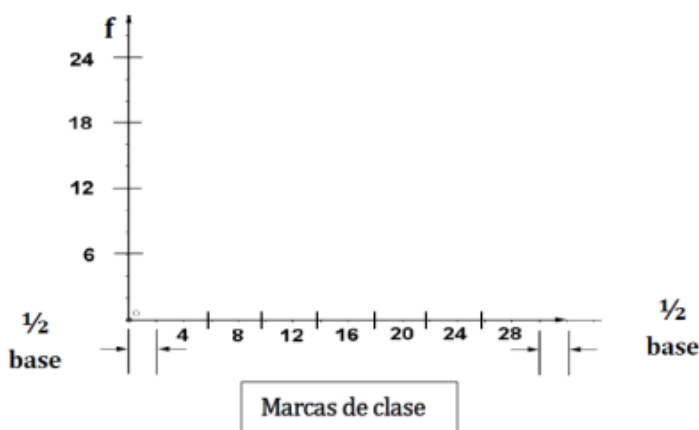
Lm	Ls	X_i	f_i
[2	6)	4	12
[6	10)	8	15
[10	14)	12	21
[14	18)	16	24
[18	22)	20	21
[22	26)	24	12
[26	30]	28	8
TOTAL			113

Solución

Al igual que en el gráfico de sectores y el gráfico de columnas, seguiremos 2 pasos

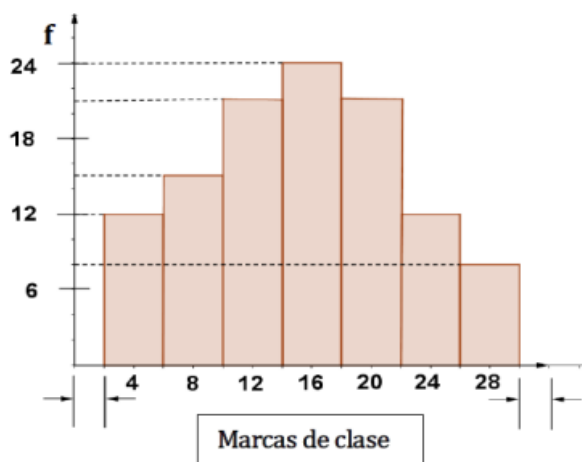
- **Paso 1:** Representar las escalas en los ejes horizontal y vertical del primer cuadrante de un plano de cartesiano. En el eje vertical colocaremos las frecuencias y en el eje horizontal las marcas de clases o los límites de los intervalos.

Para diferenciar este paso del anterior, trabajaremos con un ancho de columna único, y dejaremos la mitad de ese espacio entre el vertical y la primera columna, y el final del eje horizontal y la última columna. Para la construcción del histograma y 1 paso más para el polígono de frecuencia.

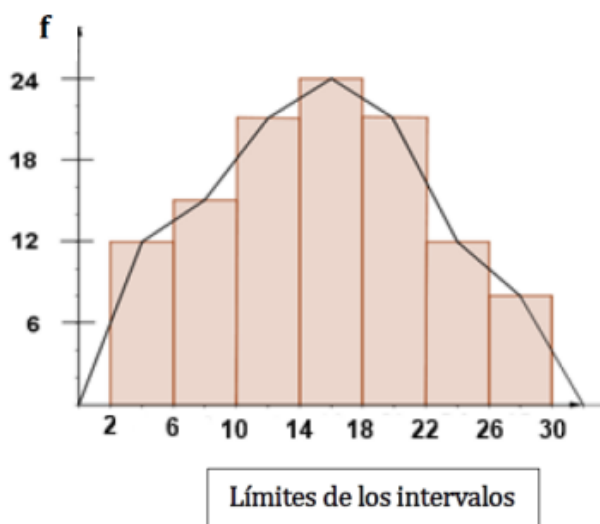


- **Paso 2:** Cada clase se representa con una columna cuya altura concuerda con su frecuencia expuesta en el eje

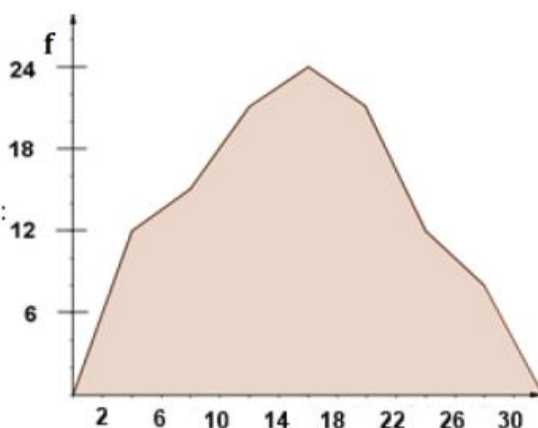
vertical. Podemos decir que la marca de clase 4 es representada por una frecuencia 12. Las barras estarán pegadas una junto a la otra, es decir, ocupando el total del ancho de cada intervalo de clase dispuesto en el gráfico.



▲ **Paso 3:** Trazar líneas rectas entre los puntos medios de los techos de columnas contiguas, partiendo desde el punto de origen (0,0) hasta el punto final definido en el eje horizontal.



Nuestro polígono de frecuencias sin el histograma quedaría de la siguiente forma:



Características de los histogramas

- No muestran frecuencias acumuladas.
- Se prefiere para el tratamiento de datos cuantitativos.
- La columna (o barra) con mayor altura representa la mayor frecuencia.
- Suelen utilizarse para representar tablas tipo B.
- La sumatoria de las alturas de las columnas equivalen al 100% de los datos.

Características de los polígonos de frecuencias

- No muestran frecuencias acumuladas.
- Se prefiere para el tratamiento de datos cuantitativos.
- El punto con mayor altura representa la mayor frecuencia.
- Suelen utilizarse para representar tablas tipo B.
- El área bajo la curva representa el 100% de los datos. El polígono de frecuencia está diseñado para mantener la misma área que de las columnas.

OJIVAS

En este gráfico se emplea un polígono de frecuencia con una característica muy particular: muestra las frecuencias absolutas o relativas acumuladas.

Ejemplo de ojivas

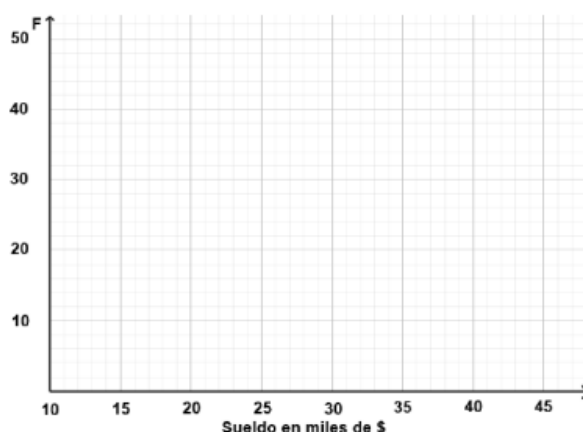
Los ingresos mensuales (en miles) de 50 trabajadores de una empresa se resumen en la siguiente tabla de frecuencia:

N_i	Lm	Ls	x_i	f_i	F_i	$h_i\%$	$H_i\%$
1	[10	15)	12,5	3	3	6	6
2	[15	20)	17,5	2	5	4	10
3	[20	25)	22,5	5	10	10	20
4	[25	30)	27,5	10	20	20	40
5	[30	35)	32,5	4	24	8	48
6	[35	40)	37,5	15	39	30	78
7	[40	45)	42,5	11	50	22	100
TOTAL				50		100	

Solución

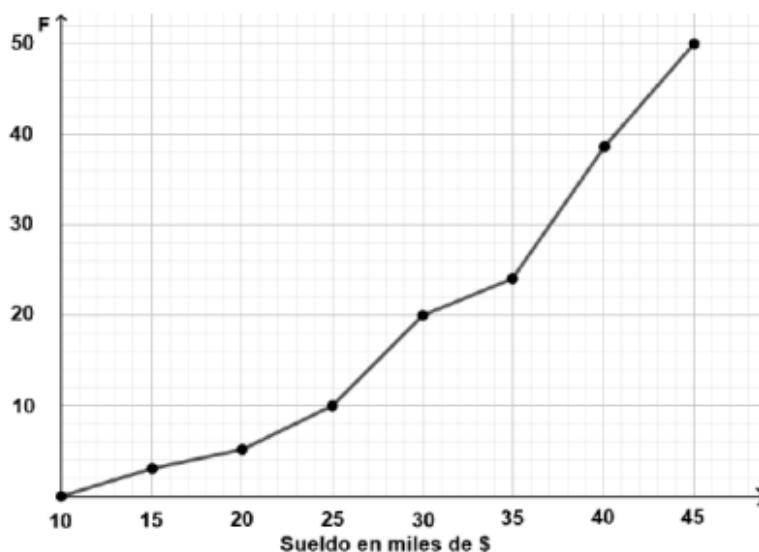
Paso 1: En un plano cartesiano (primer cuadrante), procedemos a establecer la escala de ambos ejes:

- En el eje vertical se colocarán las frecuencias, partiendo desde 0 hasta la última frecuencia acumulada (F o H).
- En el eje horizontal se ubicarán los límites superiores, partiendo desde el primer límite menor. En el caso que estos datos contengan demasiados dígitos, se realiza un recorte en sus dígitos y se especifica la igualdad.



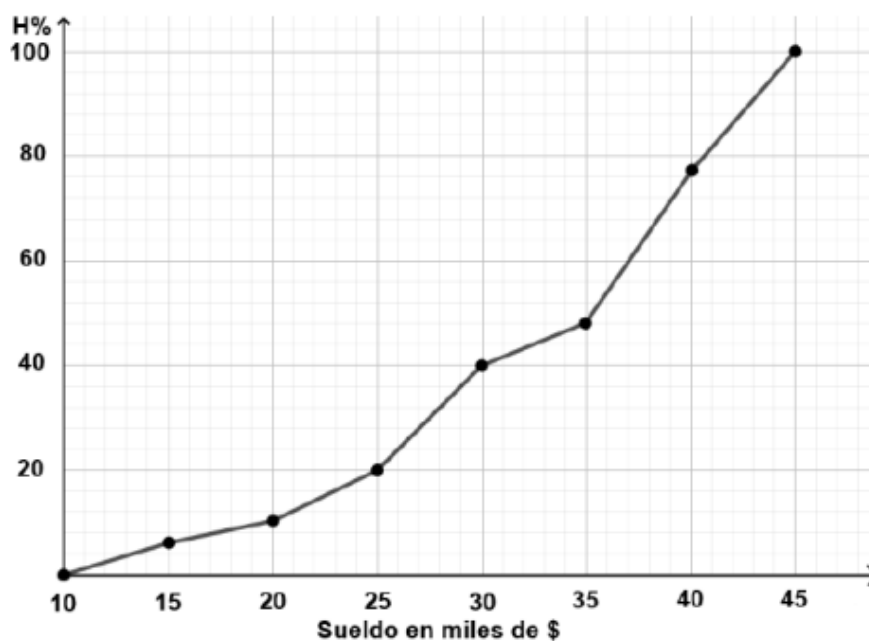
Paso 2: A cada límite superior le corresponde su frecuencia acumulada. El punto inicial será 0 (no existen datos por debajo de 10). Por ejemplo:

- Hasta un valor cercano a 15, existen acumulado 3 datos.
 - Hasta un valor cercano a 20 existen acumulado 5 datos.
- Gráficamente tenemos:



Podremos cambiar la escala del eje vertical para que trabaje con las frecuencias relativas acumuladas sin afectar el gráfico:

frecuencias relativas acumuladas sin afectar el gráfico:



Características de las ojivas

- Muestran frecuencias acumuladas.
 - Se prefiere para el tratamiento de datos cuantitativos.
 - El punto de inicio equivale a una frecuencia de 0.
 - Suelen utilizarse para representar tablas tipo B.
 - El punto final equivale al 100% de los datos
-