

Ejercicio 03

En una clase de secundaria se han medido los pesos de sus 36 alumnos, obteniendo los siguientes resultados:

78 60 74 56 56 69 79 64 66 52 79 68 70 79 60 65 54 62
64 83 66 75 70 71 52 74 64 51 49 71 63 70 72 51 74 53

- a) Agrúpalos convenientemente y dibuja el histograma y el polígono de frecuencias correspondientes.
- b) Calcula la media, moda y mediana de la distribución.
- c) Calcula la desviación típica y el porcentaje de datos que hay en los intervalos $(\bar{x} - \sigma, \bar{x} + \sigma)$ y $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$

En primer lugar debemos agrupar los datos convenientemente. El rango de esta distribución es 34 (la diferencia entre el mayor y el menor de los valores 83-49), con lo que una agrupación podría ser en intervalos de longitud 5 de esta forma:

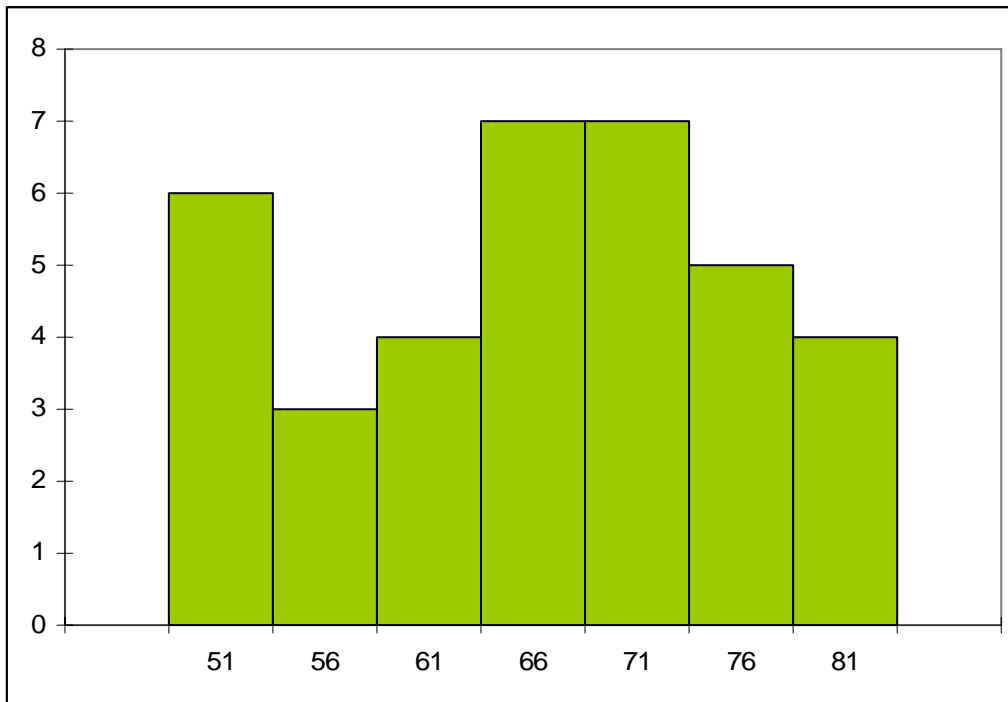
(48'5 , 53'5] (53'5 , 58'5] (58'5 , 63'5] (63'5 , 68'5] (68'5 , 73'5] (73'5 , 78'5] (78'5 , 83'5]
Empezamos por 48,5 porque el valor más pequeño de la distribución es 49. Las marcas de clase (punto medio de cada intervalo) son, respectivamente **51, 56, 61, 66, 71, 76 y 81.**

Hacemos el recuento de los datos y los ordenamos en la tabla:

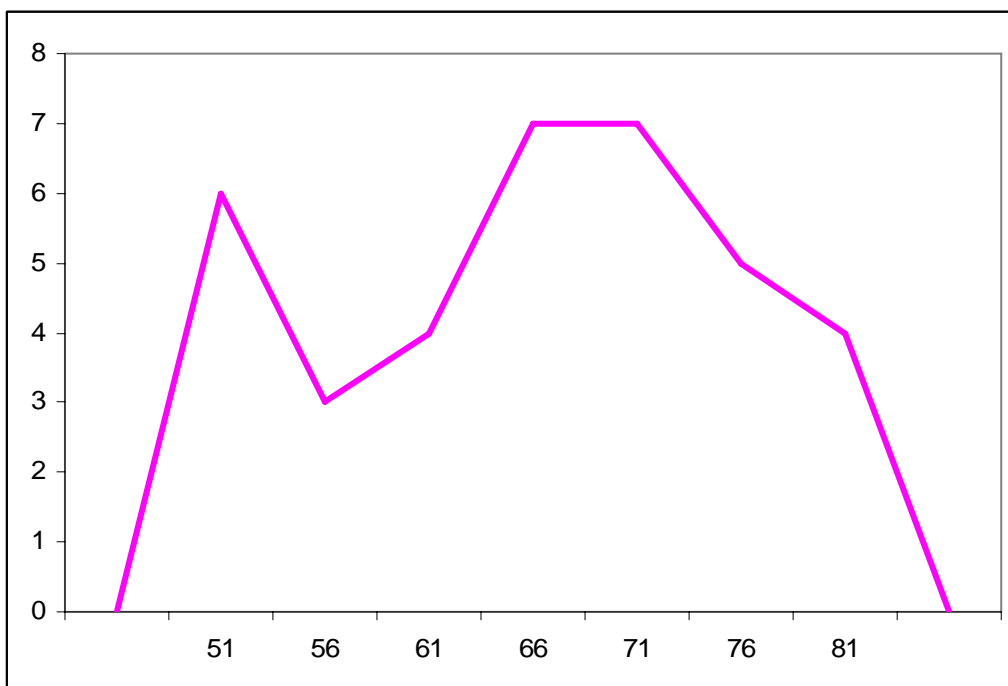
Intervalo	Marca de clase	frecuencia
(48'5 , 53'5]	51	6
(53'5 , 58'5]	56	3
(58'5 , 63'5]	61	4
(63'5 , 68'5]	66	7
(68'5 , 73'5]	71	7
(73'5 , 78'5]	76	5
(78'5 , 83'5]	81	4
		36

Como ya sabes, dibujar el histograma consiste en representar las frecuencias de cada valor mediante rectángulos. La base de cada rectángulo es el intervalo dado y su área es proporcional a la frecuencia. Si como es habitual (pero ten cuidado porque a veces no

sucede así) todos los rectángulos tienen la misma base, la altura del rectángulo correspondiente será proporcional a la frecuencia.



El polígono de frecuencias es la línea poligonal que se obtiene cuando unimos los puntos centrales de las bases superiores de los rectángulos del histograma. Observa que en la construcción de este polígono aparecen dos intervalos extremos que tienen frecuencia 0.



b) Cálculo de la moda, mediana y media de la distribución dada.

La moda es el valor que más se repite. En nuestro caso estamos ante una distribución “bimodal”, puesto que hay dos datos que tienen la misma frecuencia, que son el 66 y el 71.

La mediana de una distribución con un número par de datos es la media aritmética de los dos que ocupan las posiciones centrales. En nuestro caso, como hay 36 datos, los dos centrales son los que ocupan las posiciones 18 y 19. Vamos a la tabla y vemos que

esos valores son el 66 en ambos casos, luego la mediana sería $\frac{66 + 66}{2} = 66$.

Para calcular la media, aplicamos la fórmula cuando tenemos datos agrupados por frecuencias:

$$\bar{x} = \frac{\sum_1^7 x_i f_i}{\sum_1^7 f_i} = \frac{51 \cdot 6 + 56 \cdot 3 + 61 \cdot 4 + 66 \cdot 7 + 71 \cdot 7 + 76 \cdot 5 + 81 \cdot 4}{6 + 3 + 4 + 7 + 7 + 5 + 4} = \frac{2381}{36} = 66,14$$

c) Desviación típica e intervalos:

Calculamos la varianza y luego la desviación típica que es la raíz cuadrada

$$v = \frac{\sum_1^7 f_i x_i^2}{36} - \bar{x}^2 = \frac{15606 + 9408 + 14884 + 30492 + 35287 + 28880 + 26244}{36} - 4374,5 = \frac{160801}{36} - 4374,5 = 4466,69 - 4374,5 = 92,19$$

De manera que $\sigma = \sqrt{v} \rightarrow \sigma = 9,60$

El intervalo $(\bar{x} - \sigma, \bar{x} + \sigma)$ es el (56'54, 75'74) y dentro de él hay 18 datos, que suponen el 50% del total.

El intervalo $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$ es el (46'94, 85'34) y dentro de él están el 100% de los datos.