Guía de Ejercitación 1. Estadística Descriptiva. Respuestas

Ej. 1. Cada gráfica presentada subraya mejor un determinado mensaje. No hay respuestas que estén bien o mal. Son subjetivas, pero lo importante es detenerse a pensar en ello.

Ej. 2

	Media	Mediana	Moda	Dispersión
CI {1; 2; 3; 4; 5; 6}	3,50	3,50	---	1,87
CII {1; 1; 1;6; 6; 6}	3,50	3,50	1 y 6	2,74
CIII {–13; 2; 3;4;5;20}	3,50	3,50	---	10,48

Ej. 4 Una mirada a los datos:
$n = 36; x_{min} = 31; x_{MAX} = 84; \bar{x} = \mathbf{65{,}86}$; media recortada al 10% (2 datos de cada extremo recortados) 66,66; $s = 12{,}16$; $q_1 = 59$;$q_2 = 67{,}5$; $q_3 = 75$; $q_4 = 84$ **v.** ≅ 28%; **vi.** ≅ 69%; ≅ 94%.

31 40 45 49 52 53 57 58 58 60 61 61 63 66 67 67 67 67 68 69 70 70 70 70 72 73 75 75 76 76 78 79 80 81 83 84,

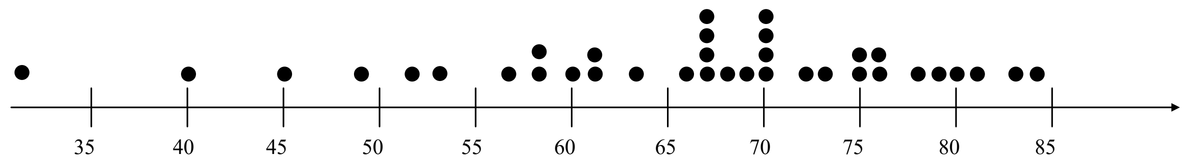


Tabla de Estadística Descriptiva de la colección de datos realizada con Excel y lectura de los parámetros que allí se detallan.

Desvío estándar $s = +\sqrt{s^2}$.

Error típico $\frac{s}{\sqrt{n}}$

Coeficiente de asimetría

$$\gamma = \frac{n}{(n-1)(n-2)s^3} \sum_{i=1}^{n} (x_i - \bar{x})^3$$

Medida de la asimetría de una distribución con respecto a su dispersión. Mayor, menor o igual a cero indican, respectivamente, simetría positivo (sesgo positivo), negativo o simétrica. (Si tiene un sólo pico)

Media	65,8611111
Error típico	2,02647273
Mediana	67,5
Moda	67
Desviación estándar	12,1588364
Varianza de la muestra	147,837302
Curtosis	0,84318992
Coeficiente de asimetría	-0,91814993
Rango	53
Mínimo	31
Máximo	84
Suma	2371
Cuenta	36

Curtosis estandarizada

$$\kappa = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)s^4} \sum_{i=1}^{n} (x_i - \bar{x})^4 \right] - \frac{n}{(n-1)(n-2)s^3}$$

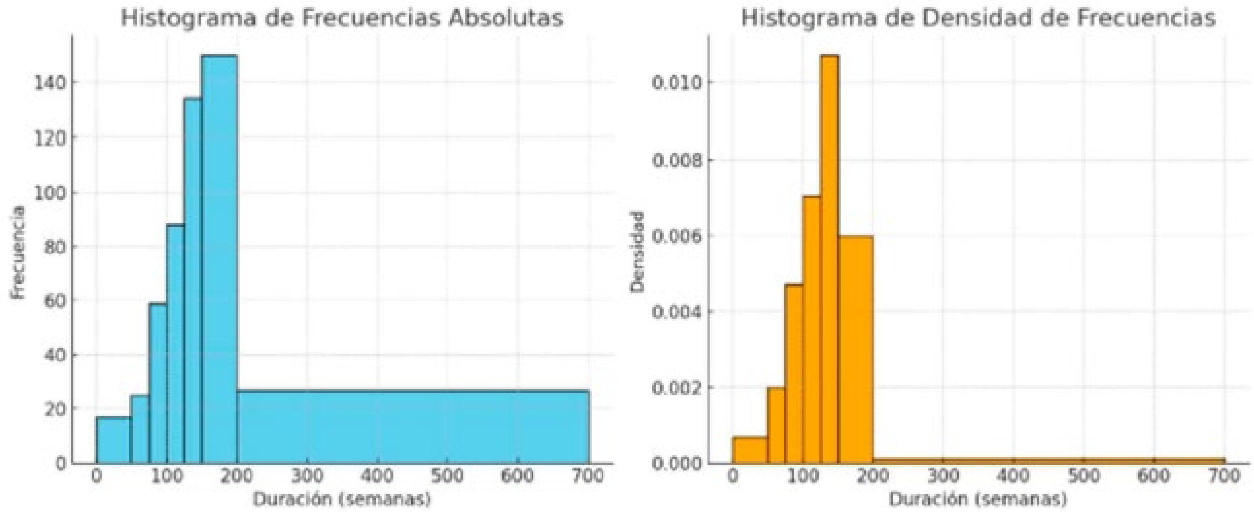
Medida de qué tan puntiaguda es una distribución. Si es mayor a 3 presenta un pico relativamente alto y recibe el nombre de leptocúrtica; si es menor a 3, la distribución es relativamente plana y recibe el nombre de platocúrtica; si es igual a 3, el pico de la distribución no es muy alto ni muy bajo y recibe el nombre de mesocúrtica. (Si tiene un sólo pico).

Ej. 5 **a)** Todo medido en semanas: $\bar{x} = 147{,}75$; $\tilde{x} = 136{,}3806$; $mo = 155{,}7554$; $s = 81{,}9771$.

b) 78,3898 semanas; **c)** 79,8%; **d)** 167,2306 semanas

e) En el histograma de frecuencias absolutas la altura de cada barra representa la cantidad de observaciones en ese intervalo específico. En cambio, el histograma de densidad de frecuencias ajusta la altura de las barras para corregir la influencia de intervalos de distinta amplitud, evitando así que las barras más anchas dominen visualmente la representación. La altura de cada barra se obtiene dividiendo la frecuencia absoluta por el ancho del intervalo y se destaca que la suma de las áreas de todas las barras equivale a 1.

Observación: Si los intervalos de clase tienen la misma amplitud, ambos histogramas serán similares. Sin embargo, si los intervalos de clase presentan distinta amplitud (como en este caso) se debe utilizar el histograma de densidad de frecuencias.



f) Dado que la media (147,75) es mayor que la mediana (136,3806), la distribución presenta una cola larga a la derecha (sesgo positivo), lo cual se confirma observando el histograma.

Ej. 11 Media: 24,02. Desvío estándar: 2,929. Mínimo: 17,5. Percentil 25: 22. Percentil 50: 23,75. Rango intercuartílico: 4. Cuartil 2: 23,75. Cuartil 3: 26.

El bigote inferior teórico abarcaría los valores de 16 a 22. Como el valor mínimo está contenido en él, a izquierda no hay datos inusuales. En forma semejante, el bigote superior teórico abarcaría los valores de 26 a 32, y el valor máximo está contenido en él.

Ejercicio I. Conjunto de datos $x_1, x_2, \dots, x_n$ donde $x_i \in \mathbf{R}$ con $i = 1, 2, \dots, n$.

a) $c = \bar{x}$

c) $\bar{y} = a\bar{x} + b$; $\tilde{y} = a\tilde{x} + b$; $s_y^2 = a^2 s_x^2$; $s_y = |a| s_x$

d) $\bar{x}_C = \frac{5}{9}(\bar{x}_F - 32)$ →media en °C: 18,81; $s_C = \frac{5}{9} s_F$ →desvío estándar en °C: 6,76