



Materia: Estadística

Actividad 1: Datos No Agrupados

Nombre del alumno (a): Josefa Pérez Magaña

Licenciatura: Administración y Estrategia de Negocios

2do. Cuatrimestre

Nombre del Profesor: Jorge Sebastián Domínguez Torres

Villahermosa, Tab., 10 de marzo de 2024.

Desarrollo de la actividad:

Se realiza una encuesta sobre el tiempo que lleva a los trabajadores llegar a la empresa, el objetivo es mejorar el horario laboral de manera general. Los datos recaudados fueron los siguientes:

20, 25, 13, 12, 45, 75, 8, 6, 20, 25, 20, 32, 16, 25, 12, 6, 5, 25, 40, 35, 45, 15, 15, 12, 8, 16, 25, 20, 35, 5, 8, 20, 25, 13, 12, 45, 75, 8, 6, 20, 25, 20, 32, 16, 25, 12, 6, 5, 25, 40, 35, 45, 15, 15, 12

1) Determina la media, mediana y moda

Hallar la media

$$\bar{x} = \frac{\sum x}{n}$$

$$x = 20 + 25 + 13 + 12 + 45 + 75 + 8 + 6 + 20 + 25 + 20 + 32 + 16 + 25 + 12$$

$$+ 6 + 5 + 25 + 40 + 35 + 45 + 15 + 15 + 12 + 8 + 16 + 25 + 20 + 35 + 5$$

$$+ 8 + 20 + 25 + 13 + 12 + 45 + 75 + 8 + 6 + 20 + 25 + 20 + 32 + 16 + 25$$

$$+ 12 + 6 + 5 + 25 + 40 + 35 + 45 + 15 + 15 + 12 = \mathbf{1,221}$$

$$\bar{x} = \frac{\sum(1221)}{55} = \mathbf{22.2}$$

$$\bar{x} = \mathbf{22.2} \rightarrow \text{Valor de la media}$$

Hallar la mediana

Ordenar de menor a mayor los datos

5, 5, 5, 6, 6, 6, 6, 8, 8, 8, 8, 12, 12, 12, 12, 12, 12, 13, 13, 15, 15, 15, 15, 16, 16, 16, 20, **20**, 20, 20, 20, 20, 20, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 32, 32, 35, 35, 35, 40, 40, 45, 45, 45, 45, 75, 75 27 datos están situados a la derecha y a la izquierda de la mediana por lo tanto la mediana es 20
 $Me = 20 \rightarrow$ Valor de la mediana

Hallar la moda

5, 5, 5, 6, 6, 6, 6, 8, 8, 8, 8, 12, 12, 12, 12, 12, 12, 13, 13, 15, 15, 15, 15, 16, 16, 16, 20, 20, 20, 20, 20, 20, 20, 20, 20, **25, 25, 25, 25, 25, 25, 25, 25**, 32, 32, 35, 35, 35, 40, 40, 45, 45, 45, 45, 75, 75

Valor	Frecuencia
5	3
6	4
8	4
12	6
13	2
15	4
16	3
20	7
25	9
32	2
35	3
40	2
45	4
75	2

Por lo tanto, la moda es el 25 toda vez que es el valor que mas se repite en los datos

$M_o = 25 \rightarrow$ Valor de la Moda

2) Determina el conjunto de la población al 30%, 55% y 75% (Tema de cuantiles)

Ordenar de menor a mayor los datos, determinando sus posiciones (X)

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄
5	5	5	6	6	6	6	8	8	8	8	12	12	12
X ₁₅	X ₁₆	X ₁₇	X ₁₈	X ₁₉	X ₂₀	X ₂₁	X ₂₂	X ₂₃	X ₂₄	X ₂₅	X ₂₆	X ₂₇	X ₂₈
12	12	12	13	13	15	15	15	15	16	16	16	20	20
X ₂₉	X ₃₀	X ₃₁	X ₃₂	X ₃₃	X ₃₄	X ₃₅	X ₃₆	X ₃₇	X ₃₈	X ₃₉	X ₄₀	X ₄₁	X ₄₂
20	20	20	20	20	25	25	25	25	25	25	25	25	25
X ₄₃	X ₄₄	X ₄₅	X ₄₆	X ₄₇	X ₄₈	X ₄₉	X ₅₀	X ₅₁	X ₅₂	X ₅₃	X ₅₄	X ₅₅	
32	32	35	35	35	40	40	45	45	45	45	75	75	

Determinar cuartil al 30%

$$Q_k = \frac{\%}{100} (n + 1)$$

donde:

Q_k es la posición del cuartil

"%" es el porcentaje del cuartil que deseo encontrar.

"n" es la cantidad total de datos de la población.

$$Q_k = \frac{30}{100}(55 + 1)$$

$$Q_k = 0.30(56)$$

$$Q_k = 16.8$$

Como la posición no es un número entero, generalmente se interpola (se suma y se divide entre dos) entre los dos valores más cercanos en la lista ordenada de datos.

En este caso, entre el dato en la posición 16 y el dato en la posición 17

$$\frac{X_{16}}{12} \quad \frac{X_{17}}{12}$$

Por lo tanto, el cuartil al 30% de los datos proporcionados es 12.

Determinar cuartil al 55%

$$Q_k = \frac{\%}{100}(n + 1)$$

donde:

Q_k es la posición del cuartil

"%" es el porcentaje del cuartil que deseo encontrar.

"n" es la cantidad total de datos de la población.

$$Q_k = \frac{55}{100}(55 + 1)$$

$$Q_k = 0.55(56)$$

$$Q_k = 30.8$$

Como la posición no es un número entero, generalmente se interpola (se suma y se divide entre dos) entre los dos valores más cercanos en la lista ordenada de datos.

En este caso, entre el dato en la posición 30 y el dato en la posición 31

$$\frac{X_{30}}{20} \quad \frac{X_{31}}{20}$$

Por lo tanto, el cuartil al 55% de los datos proporcionados es 20.

Determinar cuartil al 75%

$$Q_k = \frac{75}{100}(n + 1)$$

donde:

Q_k es la posición del cuartil

"%" es el porcentaje del cuartil que deseo encontrar.

"n" es la cantidad total de datos de la población.

$$Q_k = \frac{75}{100}(55 + 1)$$

$$Q_k = 0.75(56)$$

$$Q_k = 42$$

Como la posición es un número entero.

En este caso, la posición 42

$$\frac{X_{41} \quad X_{42} \quad X_{43}}{25 \quad 25 \quad 32}$$

Por lo tanto, el cuartil al 75% de los datos proporcionados es 25.

3) Calcula la varianza y la desviación estándar

Tomando en cuenta que los datos recabados corresponden a la población de todos los trabajadores de la empresa.

Para calcular la varianza debemos calcular primero la media de los datos

Hallar la media

$$\bar{x} = \frac{\sum x}{n}$$

$$\begin{aligned} x &= 20 + 25 + 13 + 12 + 45 + 75 + 8 + 6 + 20 + 25 + 20 + 32 + 16 + 25 + 12 \\ &+ 6 + 5 + 25 + 40 + 35 + 45 + 15 + 15 + 12 + 8 + 16 + 25 + 20 + 35 + 5 \\ &+ 8 + 20 + 25 + 13 + 12 + 45 + 75 + 8 + 6 + 20 + 25 + 20 + 32 + 16 + 25 \\ &+ 12 + 6 + 5 + 25 + 40 + 35 + 45 + 15 + 15 + 12 = \mathbf{1,221} \end{aligned}$$

$$\bar{x} = \frac{\sum(1221)}{55} = \mathbf{22.2}$$

$$\bar{x} = \mathbf{22.2} \rightarrow \mathbf{Valor de la media}$$

Hallar la varianza para la población

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}$$

$$\sigma^2 = \frac{\sum \begin{aligned} &((20 - 22.2)^2 + (25 - 22.2)^2 + (13 - 22.2)^2 + (12 - 22.2)^2 + (45 - 22.2)^2 \\ &+(75 - 22.2)^2 + (8 - 22.2)^2 + (6 - 22.2)^2 + (20 - 22.2)^2 + (25 - 22.2)^2 \\ &+(20 - 22.2)^2 + (32 - 22.2)^2 + (16 - 22.2)^2 + (25 - 22.2)^2 + (12 - 22.2)^2 \\ &+(6 - 22.2)^2 + (5 - 22.2)^2 + (25 - 22.2)^2 + (40 - 22.2)^2 + (35 - 22.2)^2 \\ &+(45 - 22.2)^2 + (15 - 22.2)^2 + (15 - 22.2)^2 + (12 - 22.2)^2 + (8 - 22.2)^2 \\ &+(16 - 22.2)^2 + (25 - 22.2)^2 + (20 - 22.2)^2 + (35 - 22.2)^2 + (5 - 22.2)^2 \\ &+(8 - 22.2)^2 + (20 - 22.2)^2 + (25 - 22.2)^2 + (13 - 22.2)^2 + (12 - 22.2)^2 \\ &+(45 - 22.2)^2 + (75 - 22.2)^2 + (8 - 22.2)^2 + (6 - 22.2)^2 + (20 - 22.2)^2 \\ &+(25 - 22.2)^2 + (20 - 22.2)^2 + (32 - 22.2)^2 + (16 - 22.2)^2 + (25 - 22.2)^2 \\ &+(12 - 22.2)^2 + (6 - 22.2)^2 + (5 - 22.2)^2 + (25 - 22.2)^2 + (40 - 22.2)^2 \\ &+(35 - 22.2)^2 + (45 - 22.2)^2 + (15 - 22.2)^2 + (15 - 22.2)^2 + (12 - 22.2)^2 \end{aligned}}{55}$$

$$\sigma^2 = \frac{\sum \begin{aligned} &((-2.2)^2 + (2.8)^2 + (-9.2)^2 + (-10.2)^2 + (22.8)^2 \\ &+(52.8)^2 + (-14.2)^2 + (-16.2)^2 + (-2.2)^2 + (2.8)^2 \\ &+(-2.2)^2 + (9.8)^2 + (-6.2)^2 + (2.8)^2 + (-10.2)^2 \\ &+(-16.2)^2 + (-17.2)^2 + (2.8)^2 + (17.8)^2 + (12.8)^2 \\ &+(22.8)^2 + (-7.2)^2 + (-7.2)^2 + (-10.2)^2 + (-14.2)^2 \\ &+(-6.2)^2 + (2.8)^2 + (-2.2)^2 + (12.8)^2 + (-17.2)^2 \\ &+(-14.2)^2 + (-2.2)^2 + (2.8)^2 + (-9.2)^2 + (-10.2)^2 \\ &+(22.8)^2 + (52.8)^2 + (-14.2)^2 + (-16.2)^2 + (-2.2)^2 \\ &+(2.8)^2 + (-2.2)^2 + (9.8)^2 + (-6.5)^2 + (2.8)^2 \\ &+(-10.2)^2 + (-16.2)^2 + (-17.2)^2 + (2.8)^2 + (17.8)^2 \\ &+(12.8)^2 + (22.8)^2 + (-7.2)^2 + (-7.2)^2 + (-10.2)^2 \end{aligned}}{55}$$

$$\sigma^2 = \frac{\sum \begin{aligned} &(4.84 + 7.84 + 84.64 + 104.04 + 519.84 \\ &+2787.84 + 201.64 + 262.44 + 4.84 + 7.84 \\ &+4.84 + 96.04 + 38.44 + 7.84 + 104.04 \\ &+262.44 + 295.84 + 7.84 + 316.84 + 163.84 \\ &+519.84 + 51.84 + 51.84 + 104.04 + 201.64 \\ &+38.44 + 7.84 + 4.84 + 163.84 + 295.84 \\ &+201.64 + 4.84 + 7.84 + 84.64 + 104.04 \\ &+519.84 + 2787.84 + 201.64 + 262.44 + 4.84 \\ &+7.84 + 4.84 + 96.04 + 38.44 + 7.84 \\ &+104.04 + 262.44 + 295.84 + 7.84 + 316.84 \\ &+163.84 + 519.84 + 51.84 + 51.84 + 104.04) \end{aligned}}{55}$$

$$\sigma^2 = \frac{12,936.80}{55}$$

$$\sigma^2 = 235.21t^2 \rightarrow \text{Valor de la varianza (tiempo al cuadrado)}$$

Hallar la desviación estándar calculando la raíz cuadrada de la desviación estándar

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}$$

$$\sigma = \sqrt{235.21}$$

$$\sigma = 15.34t$$

- 4) Construye una tabla de frecuencias respetando, rango, intervalos, amplitud, así como las diferentes frecuencias (absoluta, relativa, acumulada y porcentajes)

Ordenar los datos en orden ascendente y luego contar cuántas veces aparece cada valor

Valor	Frecuencia
5	3
6	4
8	4
12	6
13	2
15	4
16	3
20	7
25	9
32	2
35	3
40	2
45	4
75	2

Determinar el **rango** de los datos

El rango de estos datos es la diferencia entre el valor máximo y el valor mínimo:

$$R = X_{\max} - X_{\min}$$

$$R = 75 - 5$$

R = 70 → números que se incluirán en la tabla

Determinar un número adecuado de **intervalos**:

Regla de Sturges

$$k = 1 + 3.322 \log(n)$$

Donde n = número de datos en la población

$$k = 1 + 3.322 \log(55)$$

$$k = 5.0073331852324709186627029111913$$

Por lo tanto: 5 intervalos.

Determinar la **amplitud** del intervalo (ancho de cada intervalo):

Dividir el rango por el número de intervalos.

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

$$A = \frac{70}{5}$$

$$A = 14$$

Por lo tanto, la amplitud es de 14 datos.

Esto divide los datos en 5 intervalos con amplitud 14.

Cada fila muestra cuántas veces los datos caen dentro de ese intervalo. Por ejemplo, hay 26 datos en el intervalo de 5 a 18, 18 datos en el intervalo de 19 a 32, y así sucesivamente tal como se muestra en la siguiente tabla.

Intervalo	Frecuencia
5 - 18	26
19 - 32	18
33 - 46	9
47 - 60	0
61 - 74	2

Cálculos para cada columna:

- **Intervalo:** Se establece el límite inferior y el límite superior de acuerdo con la amplitud (14)
- **Marca de clase:** Es la diferencia entre los límites dividido entre dos.
- **Frecuencia:** Muestra cuántas veces los datos de la población se encuentran dentro de cada intervalo.
- **Frecuencia Acumulada:** Es la suma acumulativa de las frecuencias.
- **Frecuencia Relativa:** Se calcula dividiendo la frecuencia por el total de observaciones (55).
- **Porcentaje:** Se calcula multiplicando la frecuencia relativa por 100.
-

Intervalo (Edad) Lim Inf – Lim Sup			Marca de clase x	Frecuencias	Frecuencia Acumulada	Frecuencia Relativa	Porcentaje
5	-	19	12	26	26	0.47273	47.27 %
19	-	33	26	18	44	0.32727	32.73 %
33	-	47	40	9	53	0.16364	16.36 %
47	-	61	54	0	53	0.00000	0.00 %
61	-	75	68	2	55	0.03636	3.64 %
Σ				55		1	100 %

Fuente:

Universidad del Sureste, *Antología*

<https://plataformaeducativauds.com.mx/assets/docs/libro/LAN/f7a932d4725d180a79ef009dbb5b5379-LC-LAN204%20%20ESTADISTICA.pdf>,

Videos tutoriales consultados:

Media mediana y moda

<https://youtu.be/0DA7Wtz1ddg?feature=shared>

Cuantiles

https://youtube.com/playlist?list=PLeySRPnY35dG_Wo7ngNsT60lvahhW8UQy&si=nxyRMg98hDRe7ovi

Varianza y desviación estándar

<https://youtu.be/oZRaDwnpXkY?si=vZDsPR8MMYSTNzZB>

Tabla de frecuencias

<https://www.youtube.com/watch?v=Nm7log51vFA>