

Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

ATIVIDADE 01

QUESTÃO 01

O quadro abaixo apresenta o número de homicídios e a população de quatro regiões brasileiras em um determinado ano.

Região	Homicídio	População	Taxa
A	3.089	6.521.253	
B	1.472	2.156.904	
C	369	235.604	
D	2.561	4.562.358	

- a) Calcule a taxa de homicídios por 100 mil habitantes para cada região.

Região A	Região B	Região C	Região D
3.089 — 6.521.253	1.472 — 2.156.904	369 — 235.604	2.561 — 4.562.358
A — 100.000	B — 100.000	C — 100.000	D — 100.000
$A = \frac{3.089 \cdot 100.000}{6.521.253}$	$B = \frac{1.472 \cdot 100.000}{2.156.904}$	$C = \frac{369 \cdot 100.000}{235.604}$	$D = \frac{2.561 \cdot 100.000}{4.562.358}$
$A \approx 47,37$	$B \approx 68,56$	$C \approx 156,62$	$D \approx 56,13$

- b) Se a Região B reduzir seus homicídios em 15%, qual será a nova taxa?

$$\text{Redução de 15 \%} \Rightarrow 1.472 \cdot 0,85 = 1.251,2$$

$$\text{Nova taxa: } \frac{1.251,2 \cdot 100.000}{2.156.904} \approx 58$$

- c) Faça um comentário geral do quadro sobre as quatro regiões.

Uma leitura simplificada da tabela, considerando-se somente o valor absoluto de homicídios, poderia levar a conclusão equivocada de que a Região A é a mais perigosa. Porém, após calcular a taxa de homicídios por 100 mil habitantes, nota-se que a Região C é a que possui a maior taxa, mesmo sendo a que possui menor quantidade de homicídios. Isso ocorre devido a proporção entre homicídios e o tamanho da população, fazendo com que a Região A tenha uma taxa 3 vezes menor do que a Região C.

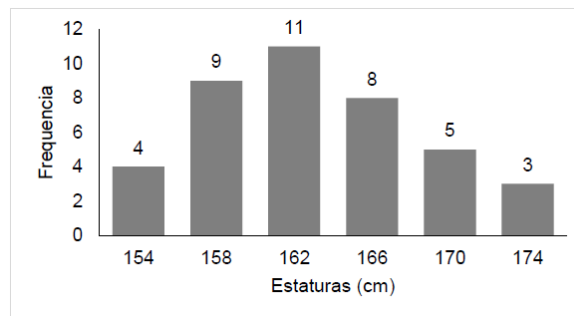
Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

QUESTÃO 02

Considere o gráfico abaixo que apresenta a distribuição das alturas (em centímetros) dos alunos em uma escola no ano letivo de 2015.



a) Identifique e classifique a variável de interesse.

*A variável de interesse é a **estatura dos alunos** e ela é classificada como sendo uma **variável quantitativa contínua**.*

b) Calcule a média, mediana e moda.

$$\begin{aligned} \text{média} &= \frac{154 \cdot 4 + 158 \cdot 9 + 162 \cdot 11 + 166 \cdot 8 + 170 \cdot 5 + 174 \cdot 3}{4 + 9 + 11 + 8 + 5 + 3} = \frac{6520}{40} = 163 \\ \text{mediana} &\Rightarrow 4 + 9 + 11 + 8 + 5 + 3 = 40 \Rightarrow \text{mediana entre 20 e 21} \\ \text{mediana} &= \frac{162 + 162}{2} = 162 \\ \text{moda} &= 162 \end{aligned}$$

QUESTÃO 03

O quadro a seguir apresenta a distribuição de frequências relativas a uma variável quantitativa contínua, organizada por classes. Algumas informações estão ausentes. Complete o quadro abaixo.

Classes	Fr Abs	Fac Abs	Fr (%)	Fac (%)
5 — 10	4	4	8	8
10 — 15	12	16	24	32
15 — 20	13	29	26	58
20 — 25	19	48	38	96
25 — 30	2	50	4	100
Total	50			

Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

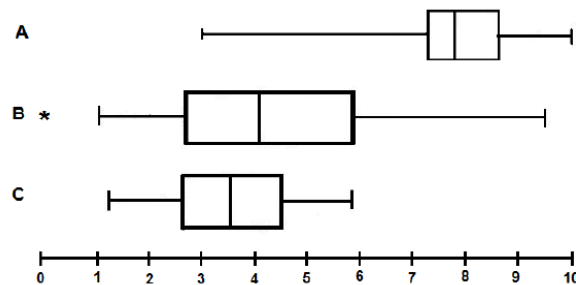
Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

Classe 5-10	Classe 10-15	Classe 15-20
$Fr_{abs} = 4 = Fac_{abs}$	$Fr_{abs} = 16 - 4 = 12$	$Fr = 58 - 32 = 26 \%$
$Fr(\%) = \frac{4}{50} \cdot 100 = 8\%$	$Fac = 24 + 8 = 32\%$	$Fr = \frac{Fr_{abs}}{total} \times 100$
		$26 = \frac{x}{50} \cdot 100 \Rightarrow 100x = 1.300 \Rightarrow x = 13$
		$Fac_{abs} = 16 + 13 = 29\%$
Classe 20-25	Classe 25-30	
$Fac = 58 + 38 = 96$	$Fr = 100 - 96 = 4\%$	
	$Fr = \frac{Fr_{abs}}{total} \times 100 \Rightarrow 4 = \frac{x}{50} \cdot 100 \Rightarrow 100x = 200 \Rightarrow x = 2$	
	$Fac_{abs} = 48 + 2 = 50$	

QUESTÃO 04

Os desenhos esquemáticos a seguir mostram as distribuições das notas de três classes A, B e C. Analise-os e descreva o comportamento dessas três distribuições.



- **Classe A**
 - A turma possui média aproximadamente igual a 8, ou seja alta.
 - A variedade de notas é pequena, uma vez que o retângulo da box é pequeno.
 - A assimetria vista entre os bigodes, com o esquerdo maior, demonstra que alguns alunos tiraram notas muito baixas (próximo a 3).
 - Não há outliers.
- **Classe B**
 - A turma possui média aproximadamente igual a 4,5, ou seja baixa.
 - A assimetria vista entre os bigodes, com o direito maior, demonstra que alguns alunos tiraram notas muito altas (próximo a 9,5). O que mostra um cenário contrário ao verificado na Classe A.
 - Há um outlier baixo, próximo a 1.
- **Classe C**
 - A turma é que possui menor média, aproximadamente igual a 3,5.
 - Diferente das outras duas classes, o gráfico é simétrico, mostrando homogeneidade nas notas.
 - Não há outliers.