Estatística descritiva

Medidas de dispersão

Prof. Dr. Tetsu Sakamoto
Instituto Metrópole Digital - UFRN
Sala A224, ramal 182
Email: tetsu@imd.ufrn.br

Slides e notebook em:

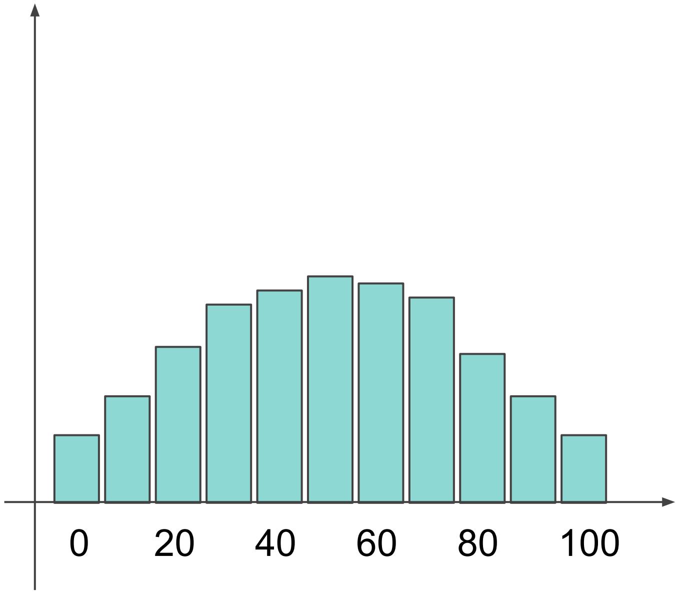
github.com/tetsufmbio/IMD0033/aula06

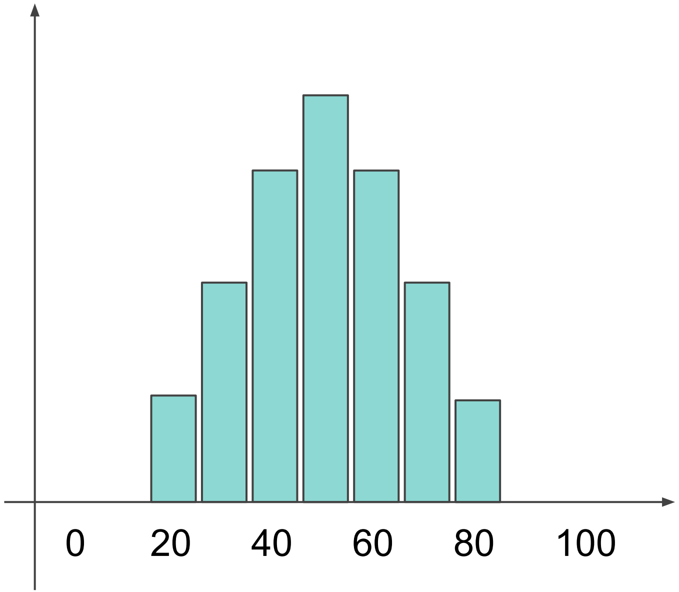
Objetivos da aula

Medidas de dispersão de dados

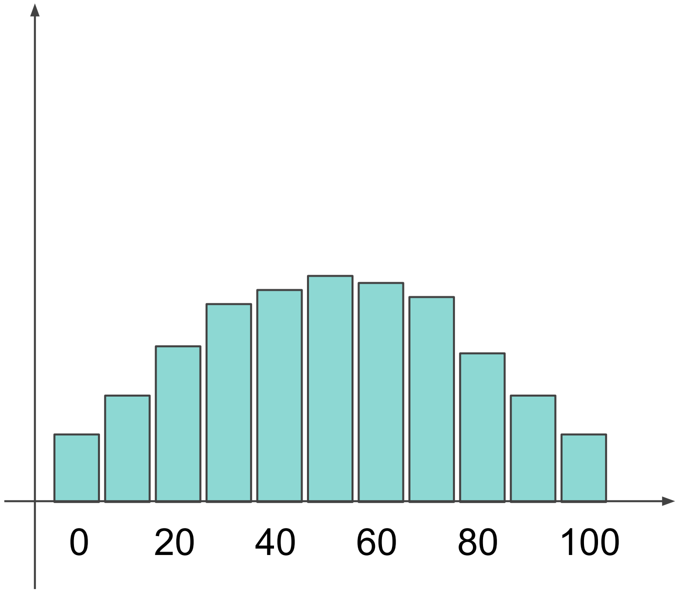
- Amplitude;
- Amplitude entre quartis;
- Variância e Desvio Padrão;

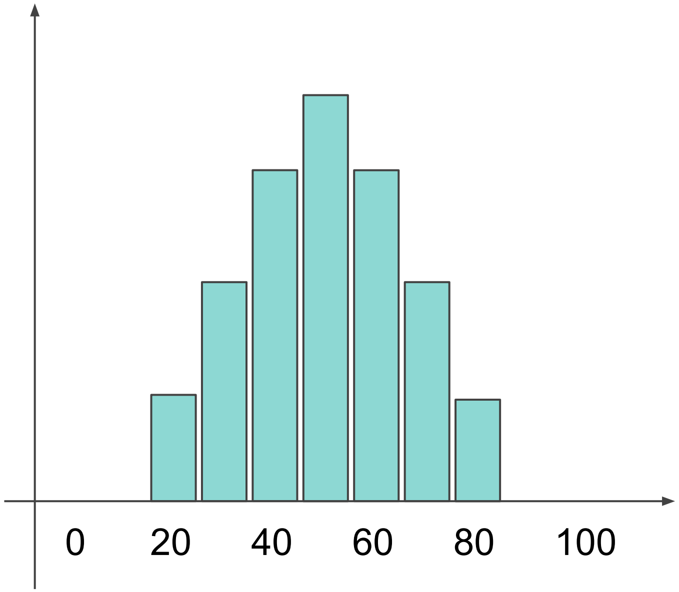
Observe os seguintes dados



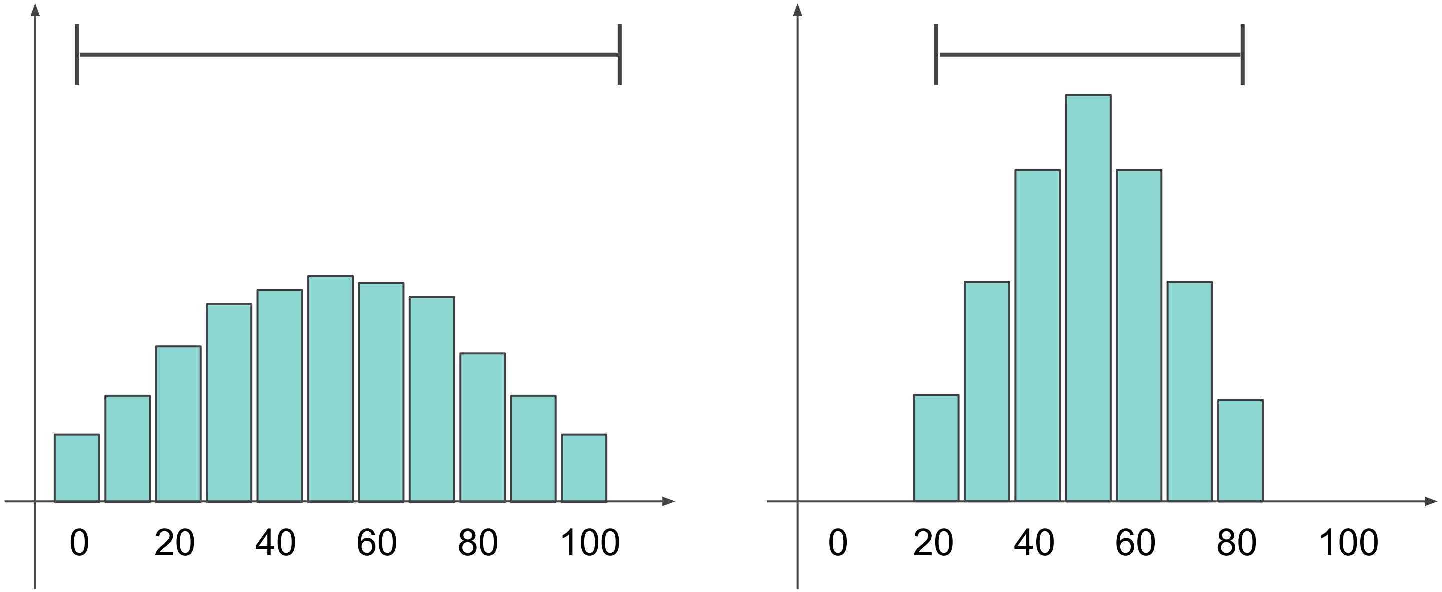


O que você diria sobre a média, mediana e moda?





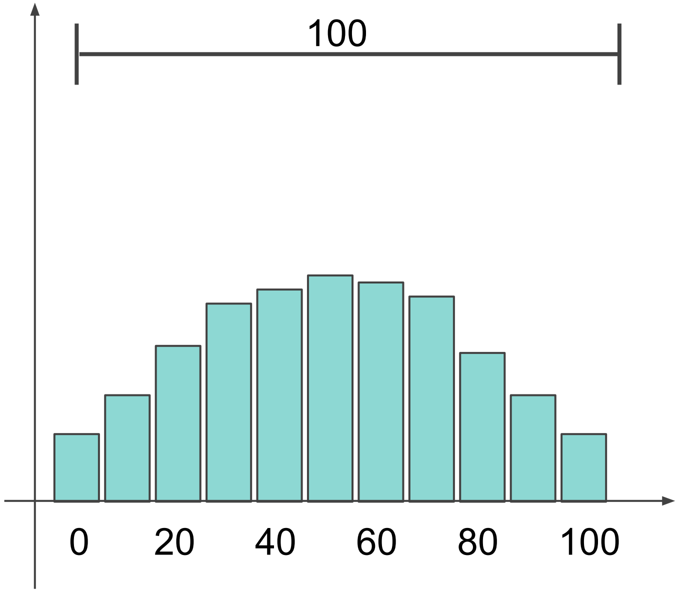
Medidas de dispersão (variabilidade)

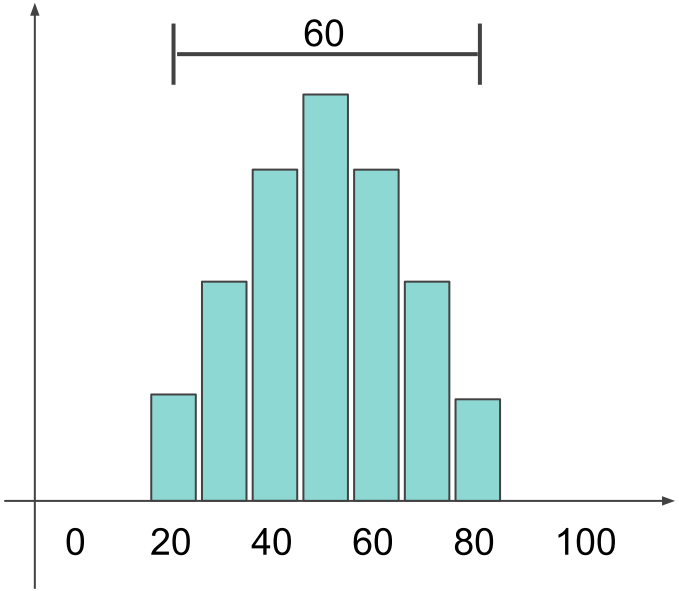


Medidas que tentam descrever o grau de dispersão dos dados.

Amplitude

Diferença entre o valor máximo e o valor mínimo dos dados.





Amplitude entre quartis (IQR)

Quartil - Divide os dados em 4 partes.

Em um quartil, definimos 3 posições:

Q1: Compreende até 25% dos dados;

Q2: Compreende até 50% dos dados (mediana);

Q3: Compreende até 75% dos dados.

Amplitude entre quartis (IQR)

A = { 1, 4, 5, 8, 8, 9, 10, 12, 14, 15, 16 }

Q1	Q2	Q3
5	9	14

IQR = 9

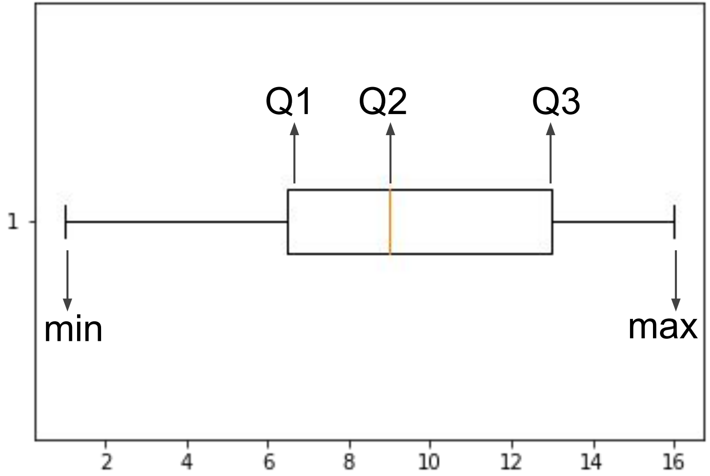
Amplitude entre quartis

A = { 1, 4, 5, 8, 8, 9, 10, 12, 14, 15, 16, 19 }

Q1	Q2	Q3
6.5	9.5	14.5

IQR = 8

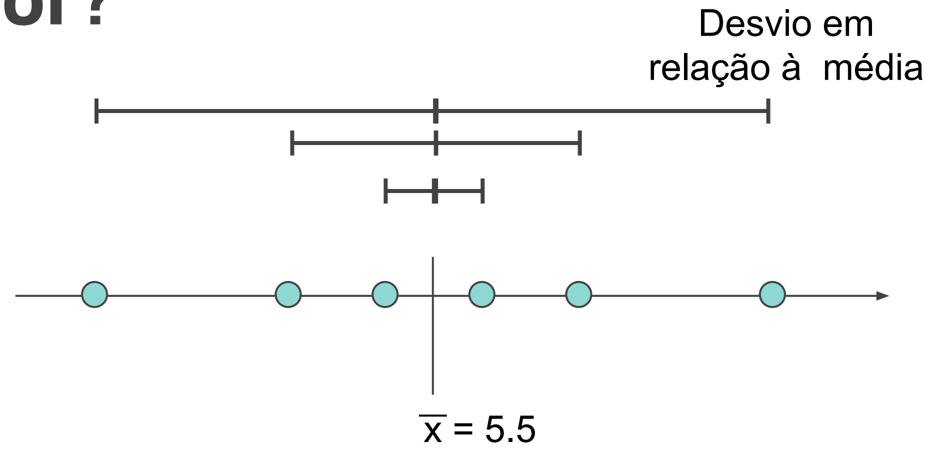
Representação gráfica em Boxplot



$A = \{ 1, 4, \boxed{5}, 8, 8, \boxed{9}, 10, 12, \boxed{14}, 15, 16 \}$

Como descrever a dispersão com apenas um valor?

x_i	$x_i - \bar{x}$
2	2 - 5.5 = -3.5
4	4 - 5.5 = -1.5
5	5 - 5.5 = -0.5
6	6 - 5.5 = 0.5
7	7 - 5.5 = 1.5
9	9 - 5.5 = 3.5



Média do desvio = somatória(xi -x)/n = 0

Como descrever a dispersão com apenas um valor?

Desvio absoluto

x_i	$x_i - \overline{x}$	$\lvert x_i - \overline{x} \rvert$
2	2 - 5.5 = -3.5	|2 - 5.5| = 3.5
4	4 - 5.5 = -1.5	|4 - 5.5| = 1.5
5	5 - 5.5 = -0.5	|5 - 5.5| = 0.5
6	6 - 5.5 = 0.5	|6 - 5.5| = 0.5
7	7 - 5.5 = 1.5	|7 - 5.5| = 1.5
9	9 - 5.5 = 3.5	|9 - 5.5| = 3.5

Média do desvio absoluto = somatória(|xi -x|)/n = 5.5

Como descrever a dispersão com apenas um valor?

Quadrado do Desvio

xi	xi - $\overline{x}$	| xi - $\overline{x}$ |	(xi - $\overline{x}$)**2
2	2 - 5.5 = -3.5	|2 - 5.5| = 3.5	(2 - 5.5)**2 = 12.25
4	4 - 5.5 = -1.5	|4 - 5.5| = 1.5	(4 - 5.5)**2 = 2.25
5	5 - 5.5 = -0.5	|5 - 5.5| = 0.5	(5 - 5.5)**2 = 0.25
6	6 - 5.5 = 0.5	|6 - 5.5| = 0.5	(6 - 5.5)**2 = 0.25
7	7 - 5.5 = 1.5	|7 - 5.5| = 1.5	(7 - 5.5)**2 = 2.25
9	9 - 5.5 = 3.5	|9 - 5.5| = 3.5	(9 - 5.5)**2 = 12.25

Média do quadrado do desvio

=

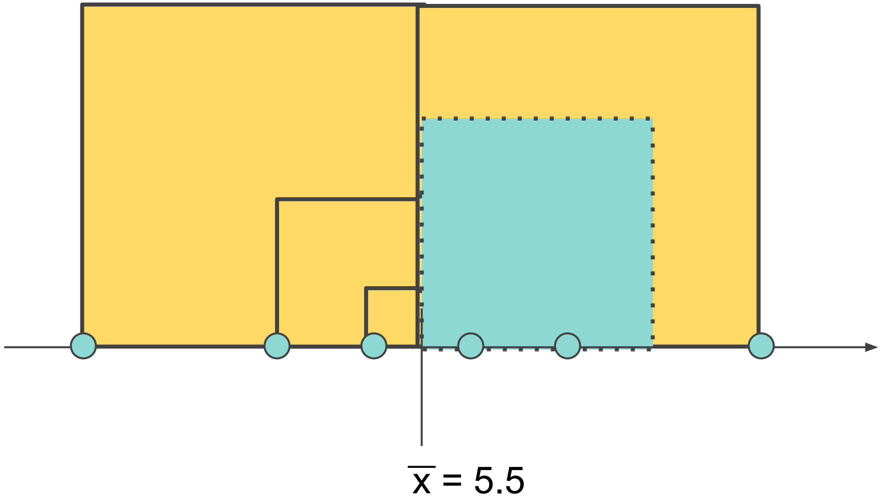
somatória((xi -x)**2)/n

=

4.916

$\boxed{\text{Variância}}$

Quando estamos calculando a variância...

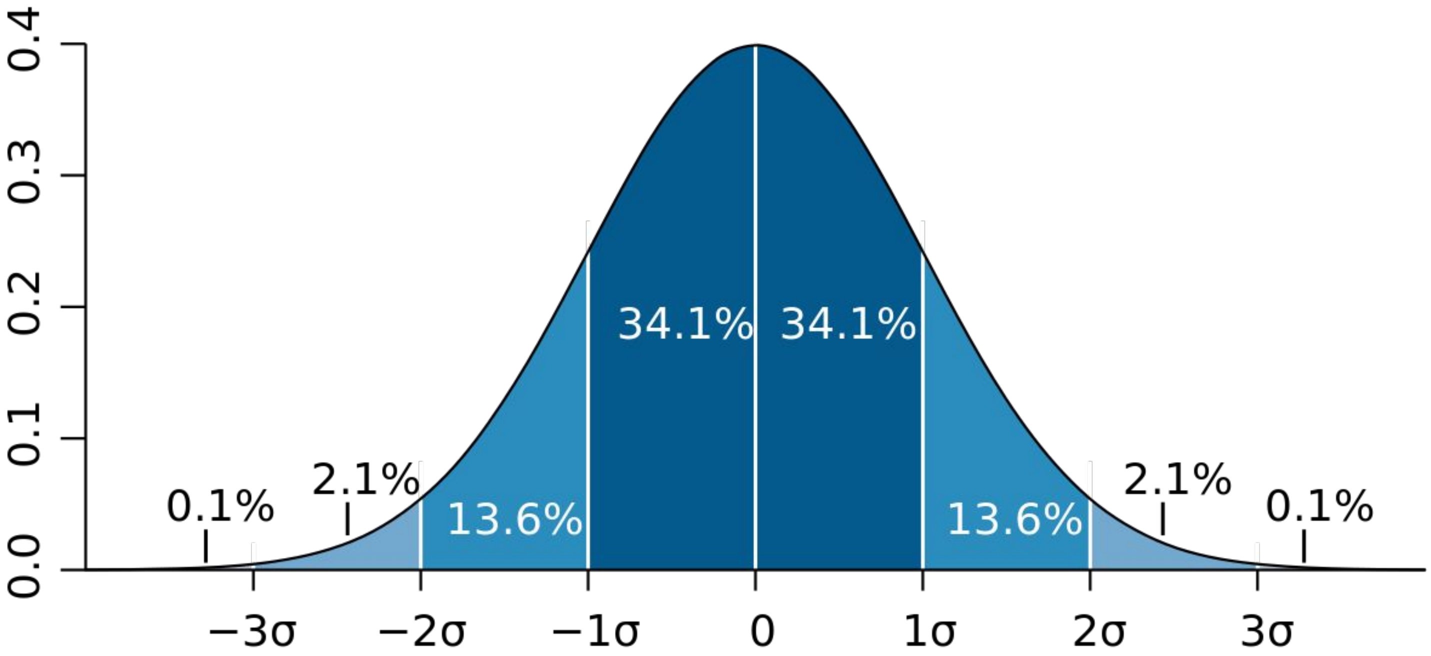


Desvio padrão

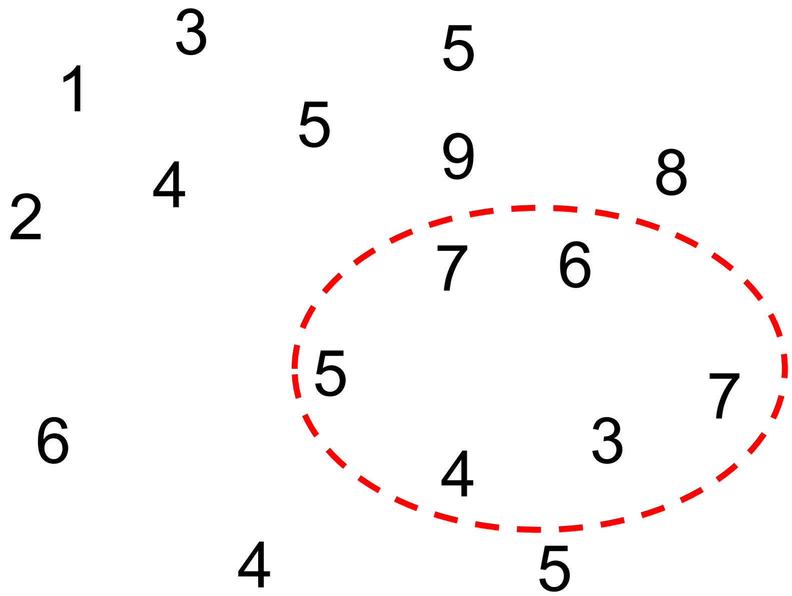
Variância = $\sigma^2 = (\Sigma(xi - x)^2)/n)$

Desvio padrão = $\sigma = ((\Sigma(xi - x)^2)/n)^{0.5}$

Desvio padrão



Desvio padrão da amostra



Desvio padrão da amostra

Na maioria das vezes, uma amostragem não consegue representar toda a variabilidade da população, por isso utilizamos a correção de Bessel, onde determina que o desvio padrão da amostra, caso utilizado para estimar o desvio padrão de uma população corresponde a:

$$\text{Desvio padrão amostral} = s = ((\Sigma(xi - x)^2)/(n-1))^{0.5}$$