

Formação Cientista de Dados



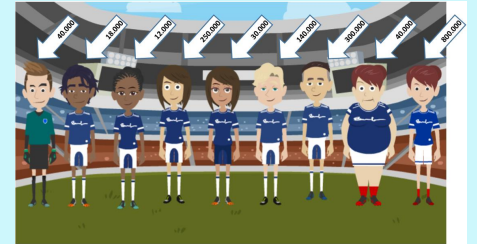
Estatística I: Medidas de Centralidade e Variabilidade



Variabilidade

- Variância
- Desvio Padrão
- Amplitude
- Não Centrais: Quartis

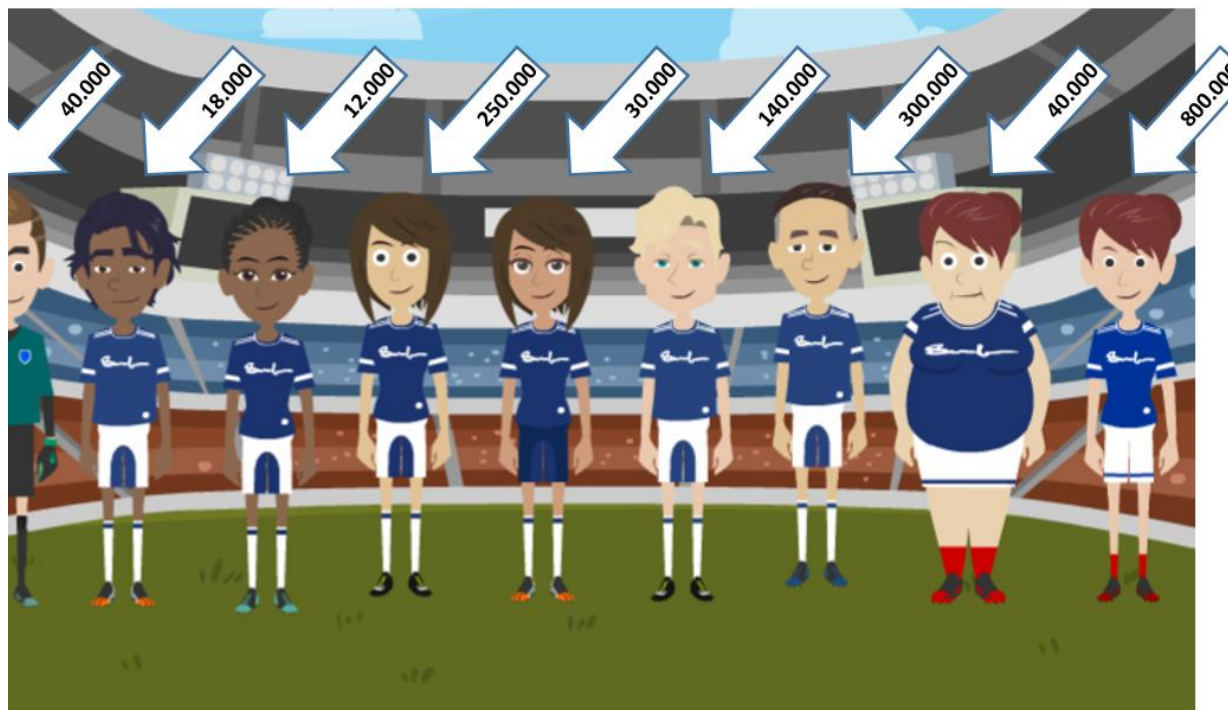
Variância



$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

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

57.939.654.321



240.706,57

Desvio Padrão



Amplitude

800.000

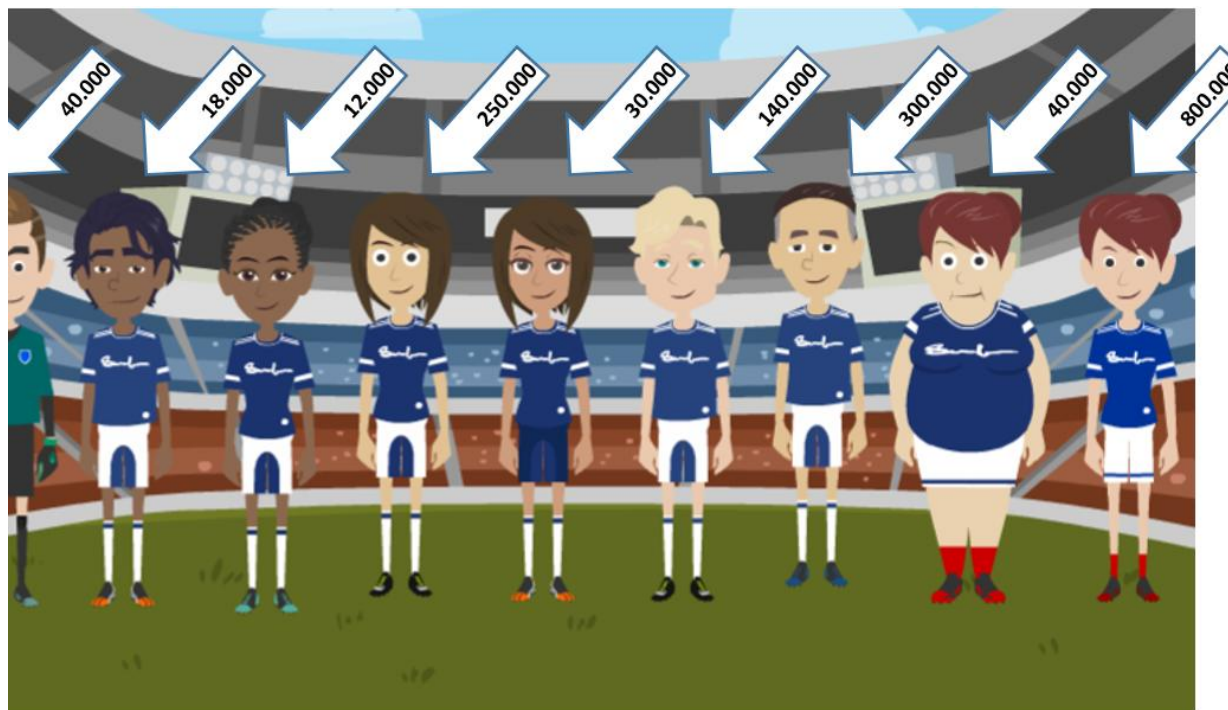
12.000

788.000

Medidas não Centrais

- Quartis

- Q1 : 25% dos menores valores 30.000
- Q2: 50% (equivale a mediana) 40.000
- Q3: 75% dos maiores valores 250.000



Resumo do Time de Futebol

Média: 181.111

Q1: 30.000

Q2: 40.000

Q3: 250.000

Desvio Padrão: 240.706,57

Variância: 57.939.654.321

Amostra <>
População

ESTATÍSTICA

AMOSTRA

POPULAÇÃO

Média

\bar{X}

μ

Desvio Padrão

S

σ



Qual Maior Desvio
Padrão?

$A = \{10, 20, 31, 12\}$

$B = \{100, 110, 113, 123\}$

Semelhante: 9,5



Média: 181.111

Q1: 30.000

Q2: 40.000

Q3: 250.000

Desvio Padrão: 240.706,57

Variância: 57.939.654.321

Desvio Padrão e Variância!

- Qual usar?
- O ideal é usar ambos!
- Variância é Influência por valores muito distantes
- Você pode analisar a variação destacando ou não discrepâncias (variância, desvio padrão)

Variabilidade



- Variância
- Desvio Padrão
- Amplitude
- Não Centrais: Quartis

