### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

eleitorais

Medidas de posição

parla-

Medidas de dispersão

- 1 Revendo conteúdos da última aula RMarkdwon
- 2 Dados, Indicadores e escalas
- 3 Indicadores parlamentares e eleitorais
- 4 Medidas de posição
- 5 Medidas de dispersão
- 6 Surveys e planejamento

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

## Revendo conteúdos da última aula - RMarkdwon

## .rmd do zero e gh-pages

Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

#### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

# Dados, Indicadores e escalas

### Matriz de dados

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

```
# Atribuindo o dataframe de exercicio
dfe <- read_rds("dados/dfe.rds")
#</pre>
```

### dataMaid

```
Frederico
Bertholini
```

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

```
# Codebook dataMaid -- explicar
attr(dfe$id, "label") <- "Identificação"</pre>
attr(dfe$id, "shortDescription") <- "Variável de identificação do aluno"
attr(dfe$media, "label") <- "Média final"
attr(dfe$media, "shortDescription") <- "Nota do aluno ao final da disciplina"
attr(dfe$faltas. "shortDescription") <- "Total de faltas ao longo do semestre"
attr(dfe$turma, "shortDescription") <- "Turma do aluno"</pre>
attr(dfe$idade, "shortDescription") <- "Idade do aluno"</pre>
attr(dfe$interess. "label") <- "Interesse"</pre>
attr(dfe$interess, "shortDescription") <- "Prioridade de interesse do aluno"
attr(dfe$tempocup, "shortDescription") <- "Tempo de dedicação semanal do aluno"
attr(dfe$escola, "shortDescription") <- "Tipo de escola onde o aluno cursou ensino médi
attr(dfesetcivil, "shortDescription") <- "Estado civil do aluno"
# makeCodebook(dfe,file = "cb_dfe.rmd")
```

R Ladies - dataMaid

## Índices e escalas

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkd-

#### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

### Ordinais

- Escala
- Atribuições ao padrão de resposta
- Pode conter vários itens
  - Índice
- Compostos por itens (multidimensional)
- cada item é uma variável

### Critérios

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

- Validade (interna e externa)
- Variância
- Correlação (bivariada e multivariada)
- Outras propriedades consagradas (Januzzi)

Frederico Bertholini Revendo	Propriedades	Indicator	Indicatori	indicator 1
da última aula -	Relevância para a agenda política			
RMarkd- won  Dados, Indicadores e escalas  Indicadores parla- mentares e eleitorais  Medidas de posição  Medidas de dispersão  Surveys e planeja- mento	Validade de representação do conceito			
	Confiabilidade da medida			
	Cobertura populacional			
	Sensibilidade às ações previstas			
	Especificidade ao programa			
	Transparência metodológica na sua construção			
	Comunicabilidade ao público			
	Factibilidade operacional para sua obtenção			
	Periodicidade na sua atualização			
	Desagregabilidade populacional e territorial			
	Comparabilidade da série histórica			
	Total de propriedades (+):			

Obs.: Atribuir + quando a propriedade for verificada para o indicador

## tribuição de pontos

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

## Escalas de respostas

- amplitude
- Não resposta
- pesos

# Tipos

# ■ Likert

Feeling thermometer

Menos conhecidas (e utilizadas)

Wellos Collifecidas (e utilizadas

■ Bogartus (distância social)

Indicadores parlamentares e eleitorais

Bertholini

Revendo

RMarkdwon

Dados,
Indicadores
e escalas

conteúdos da última aula -

mentares e eleitorais Medidas de

Medidas de posição

Medidas de dispersão

■ Thurstone					
■ Guttman					
	Problems				
Children	Counting	Addition	Subtraction	Multiplication	Division
$S_1$	1	0	0	0	0
$S_2$	1	1	0	0	0
$S_3^-$	1	1	1	0	0
$S_A$	1	1	1	1	0

## Testes de confiabilidade - Cronbach

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkd-

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

### Exemplo Cervi

- B.11. O papel do Estado é: Garantir educação universitária universal e gratuita
- B.12. O papel do Estado é: Garantir acesso à saúde pública universal e gratuita
- B.13. O papel do Estado é: Garantia de renda mínima para cidadãos em situação de pobreza
- B.14. O papel do Estado é: Assegurar a oferta de cotas para minorias em universidades públicas
- B.15. O papel do Estado é: Promover Reforma agrária com desapropriação de grandes propriedades
- B.16. O papel do Estado é: Promoção de fontes alternativas de informação.

As alternativas de respostas cada uma das perguntas são:

- 1. Concordo fortemente
- 2. Concordo
- 3. Discordo
- Discordo fortemente

### Frederico Bertholini

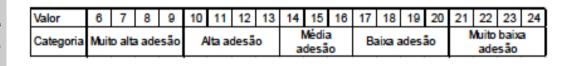
Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão



	SOMA	Nº CASOS	% CASOS	CATEGORIA	% CASOS
Frederico	6	77	17,3		
Bertholini	7	33	7,42	Adesão muito	47%
Revendo	8	55	12,36	alta	41%
conteúdos da última	9	44	9,89	†	
aula - RMarkd-	10	48	10,79		
won	11	52	11,69		44 549/
Dados, Indicadores	12	49	11,01	- Adesão alta	41,54%
e escalas	13	36	8,09	†	
Indicadores	14	18	4,04		
parla- mentares e	15	12	2,7	Adesão média	9,21%
eleitorais	16	11	2,47	†	
Medidas de posição	17	5	1,12		
Medidas de	18	2	0,45	Adesão baixa	2.259/
dispersão	19	2	0,45	Adesão baixa	2,25%
Surveys e planeia-	20	1	0,22	1	
mento	21	0	0		
	22	0	0	Adesão muito	0.008
	23	0	0	baixa	0,00%
	24	0	0	1	

## **Princípios**

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkd-

#### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

planeja

Surveys e

- Criterion-related
- Content
- Construct

Lógica:

$$x = t + e$$

### Onde:

x é a medida empírica a ser estudada

t é a parte explicada da variação

e é o erro randômico (aleatório)

Vamos olhar isso melhor em inferência.

## **Fórmula**

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento A fórmula para o cálculo do α de Cronbach é a seguinte:

$$\alpha = \frac{k}{k-1} x \frac{St^2 - \sum St^2}{St^2}$$

Onde:

K = número de variáveis que compõem o índice;

St<sup>2</sup> = Variância Total (a partir da soma das variâncias dos respondentes);

 $\sum Si^2$  = Soma das Variâncias das variáveis ou itens que compõem o índice.

### Cronbach

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

```
Usando o pacote psych
```

```
r4 <- psych::sim.congeneric()
crb <- psych::alpha(r4)
summary(crb)</pre>
```

```
Reliability analysis raw_alpha std.alpha G6(smc) average_r S/N median_r 0.74 \quad 0.74 \quad 0.69 \quad 0.42 \ 2.9 \quad 0.41
```

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e

eleitorais Medidas de posição

Medidas de

dispersão Surveys e

Composição do índice	α de Cronbach
B11 + B12 + B13 + B14 + B15 + B16	0,762
B11 + B12 + B13 + B14 + B15	0,756
B11 + B12 + B13 + B14 + B16	0,676
B11 + B12 + B13 + B15 + B16	0,694
B11 + B12 + B14 + B15 + B16	0,702
B11 + B13 + B14 + B15 + B16	0,754
B12 + B13 + B14 + B15 + B16	0,736

## Mais info

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

### Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento Pacote psych

Consistência Interna

### Indicadores estatísticos

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

#### Dados, Indicadores e escalas

Indicadores parlamentares e

eleitorais Medidas de

posição Medidas de dispersão

Surveys e planejamento Razão: Z = X/Y

Proporção: Z = X/(Y + X)

Percentagem: *Propor*ção \* 100

Taxa: eventos/exposiçãoaorisco

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

# Indicadores parlamentares e eleitorais

# Parlamentares - Índice de fracionalização

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

#### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

# Parlamentares - Índice de fracionalização máxima

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

#### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

# Parlamentares - Índice de fragmentação

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

#### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

## Parlamentares - Número efetivo de partidos

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

$$\frac{1}{\sum pe^2}$$

## Parlamentares - Renovação

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

$$Y_{ij} = [\beta_0 + \beta_1(\mathsf{Dose} - 300)] + [\varepsilon_{ij}]$$

## Eleitorais

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

### Mais recursos

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados,

Indicadores e escalas Indicadores

### parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento psData

rOpenGov

QoG

electoral

## Representações descritivas de dados - Tabela

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicador e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

```
# Com tabyl
dfe %>%
drop_na(interess,estcivil) %>% # retirando NAs
janitor::tabyl(interess,estcivil) %>% # tabela cruzada
janitor::adorn_percentages("col") %>%
janitor::adorn_pct_formatting()
```

```
interess Casado Solteiro
Secundário 29.4% 45.0%
Principal 70.6% 55.0%
```

## Representações descritivas de dados - Histograma

### Frederico Bertholini

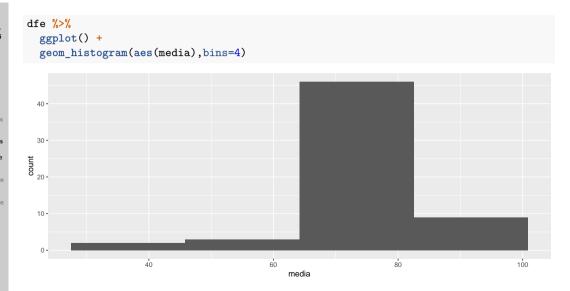
Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

#### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão



## Representações descritivas de dados - Gráfico de dispersão

### Frederico Bertholini

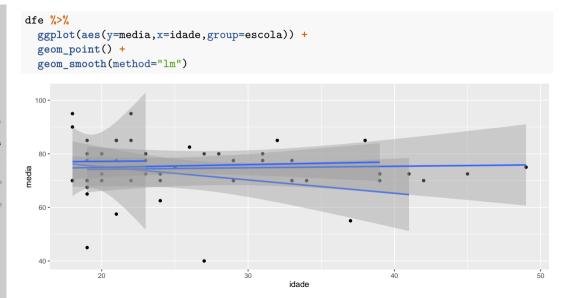
Revendo conteúdos da última aula -RMarkd-

Dados, Indicadores e escalas

### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão



## Representações descritivas de dados - Binário

### Frederico Bertholini

dfe %>%

drop\_na() %>%

ggplot() +

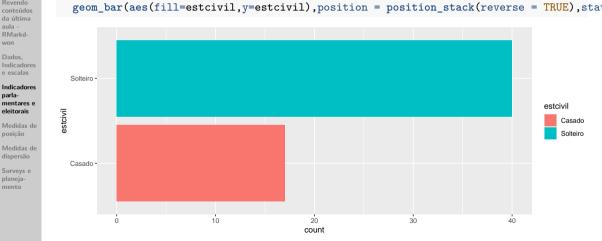
Revendo conteúdos da última aula -RMarkd-

Dados.

### Indicadores parlamentares e eleitorais

posição

dispersão



## Representações descritivas de dados - Box plot

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

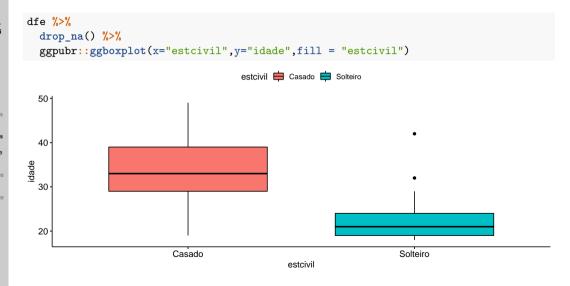
Dados, Indicadores e escalas

#### Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planeja-



### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

## Medidas de posição

Medidas de dispersão

Surveys e planejamento

# Medidas de posição

## Medidas de posição

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

### Medidas de posição

Medidas de dispersão

Surveys e planejamento As medidas de posição trazem informação sobre a localização dos dados no seu conjunto de possíveis valores.

### Média aritmética

### Frederico Bertholini

Revendo conteúdos da última aula -RMarkd-

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

### Medidas de posição

Medidas de dispersão

Surveys planejamento A média aritmética é definida pela soma das observações dividida pelo número total de observações.

Exemplo: sejam  $x_1,...,x_n$  n observações de uma variável X qualquer. A média é dada pela expressão

$$\frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

No R:

```
dfe %>%
  pull(media) %>%
  mean(na.rm=T)
```

[1] 74.375

## Média ponderada

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

#### Medidas de posição

Medidas de dispersão

Surveys e planejamento

### Mediana

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas Indicadores

Indicadores parlamentares e eleitorais

#### Medidas de posição

Medidas de dispersão

Surveys e planejamento A mediana é a observação que ocupa a posição central dos dados.

Exemplo 1 (número ímpar de observações): vamos considerar o seguinte conjunto de observações: 3, 7, 10, 5, 2, 1, 1.

Para calcular a mediana, primeiramente ordenamos os dados: 1, 1, 2, 3, 5, 7, 10.

É fácil verificar que a observação que ocupa a posição central é o valor 3. Portanto, 3 é a mediana desse conjunto de valores.

Exemplo 2 (número par de observações): vamos considerar agora o seguinte conjunto de observações: 15, 3, 2, 0, 9, 17.

Ordenanando os dados, temos: 0, 2, 3, 9, 15, 17.

Neste caso, a mediana será dada pela média entre as duas observações centrais, isto é, (3+9)/2=6.

### Moda

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Indicadores e escalas Indicadores

Indicadores parlamentares e eleitorais

#### Medidas de posição

Medidas de dispersão

Surveys e planejamento A moda é a observação mais frequente do conjunto de valores observados.

Exemplo: no conjunto de observações {3, 5, 10, 11, 11, 20}, a observação 11 aparece 2 vezes enquanto as demais apenas 1 vez. Portanto, 11 é a moda desse conjunto.

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

# Medidas de dispersão

## Medidas de dispersão

### Frederico Bertholini Revendo

conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys of planeja-

O resumo de um conjunto de dados por uma única medida de posição central não traz informação sobre a variabilidade das observações.

Um critério frequentemente usado para avaliar a dispersão de um conjunto de observações é aquele que mede a dispersão dos dados em torno da sua média.

### **Amplitude e Quantis**

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

planejamento Amplitude = valor máximo - valor mínimo

A mediana é uma medida que deixa metade dos dados abaixo dela e metade acima. De modo geral, podemos definir a medida  $q_p$ , 0 , tal que <math>p das observações sejam menores que  $q_p$ . Esta medida é chamada de quantil de ordem p.

Alguns quantis são bastante utilizados na prática. São eles

 $Q1=q_{25\%}$  (primeiro quartil)

 $Q2=q_{50\%}$  (mediana)

 $Q3 = q_{75\%}$  (terceiro quartil)

Observe que os  $q_{0\%}$  e  $q_{100\%}$  denotam, respectivamente, o mínimo e o máximo de conjunto de dados.

### Variância

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

# Medidas de dispersão

Surveys e

Sejam  $x_1, x_2, ..., x_n$  n observações de uma variável quantitativa X. Seja  $\bar{x}$  a média dessas observações.

Vamos observar, então, os desvios das observações em relação à média  $\bar{x}$ , isto é,  $x_i - \bar{x}$ , para i = 1, ..., n.

Ideia 1: usar como medida de dispersão a soma desses desvios, isto é,

$$\sum_{i=1}^{n}(x_{i}-\bar{x})$$

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parla-

parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento Problema 1: para qualquer conjunto de observações

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

Ideia 2: considerar a soma do quadrado dos desvios, dada por

$$\sum_{i=1}^{m}(x_{i}-\bar{x})^{2}$$

Revendo conteúdos da última aula -RMarkdwon

Indicadores e escalas Indicadores parla-

Dados.

parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento  $\textbf{Problema 2}: \ n\~{a}o \ podemos \ comparar \ conjuntos \ de \ observaç\~{o}es \ de \ tamanhos \ diferentes.$ 

Ideia 3: utilizar a média desses desvios quadráticos.

Sendo assim, definimos a medida de dispersão conhecida como variância da seguinte forma:

$$var(X) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

## Desvio padrão

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento A utilização da variância como medida de dispersão pode ainda causar um problema de interpretação, pois a sua dimensão é igual ao quadrado da dimensão dos dados. Sendo assim, costuma-se usar o desvio padrão, definido por

$$dp(X) = \sqrt{var(X)} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

## Coeficiente de variação

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

## Medidas de dispersão

Surveys e planejamento Medida relativa -> que proporção da média é o desvio-padrão

### vendo média, mediana e distribuições

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicador e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

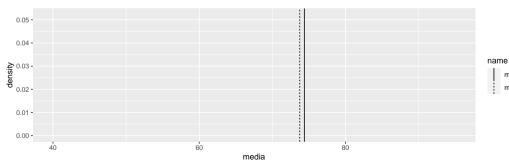
#### Medidas de dispersão

Surveys planejamento



media

mediana



### vendo média, mediana e distribuições

#### Frederico Bertholini

Revendo conteúdo da última aula -RMarkdwon

Dados, Indicador e escalas

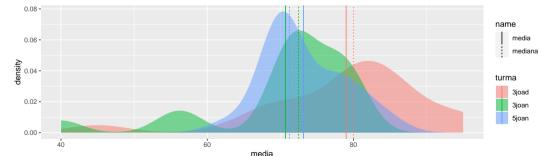
parlamentares e eleitorais

Medidas de posição

#### Medidas de dispersão

Surveys e planejamento





### Medidas de assimetria e curtose

#### Frederico Bertholini

Revendo conteúdos da última aula -RMarkdwon

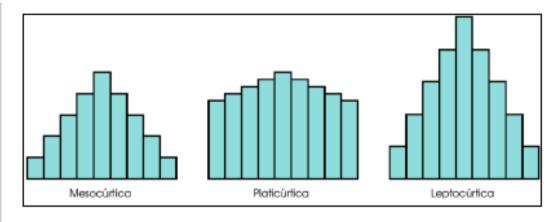
Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

#### Medidas de dispersão

Surveys e planejamento



Coeficiente de Assimteria de Pearson = média – moda

Assimetria adimensional =  $\frac{m\acute{e}dia - moda}{dp}$ 

Coeficiente quartílico de assimetria =  $3 * \frac{m\acute{e}dia - mediana}{dp}$ 

Revendo conteúdos da última aula -RMarkdwon

Dados, Indicadores e escalas

Indicadores parlamentares e eleitorais

Medidas de posição

Medidas de dispersão

Surveys e planejamento

# Surveys e planejamento