

# Estatística descritiva

Introdução à metodologia de pesquisa

**Prof. Dr. Tetsu Sakamoto**

Instituto Metrópole Digital - UFRN

Sala A224, ramal 182

Email: [tetsu@imd.ufrn.br](mailto:tetsu@imd.ufrn.br)



**Você é uma pessoa alta?**



# Metodologia científica

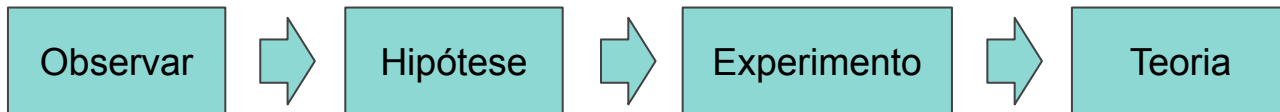
Você é uma pessoa alta?





# Metodologia de pesquisa

Você é uma pessoa alta?



**Métodos Descritivos** → levantar dados que descrevam uma população;

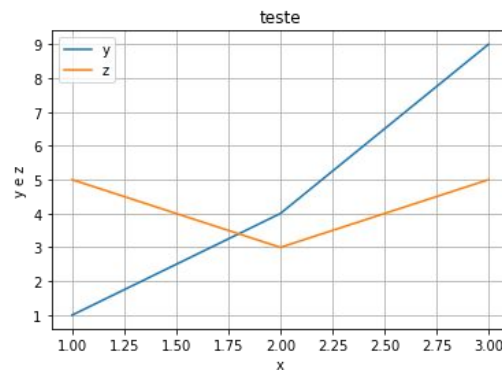
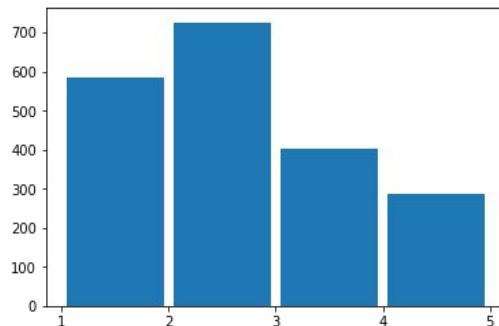
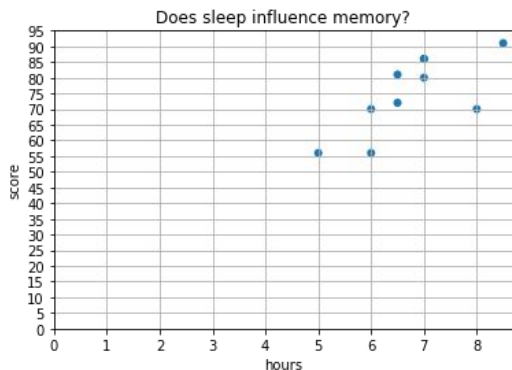
**Métodos Analíticos** → buscam esclarecer questões ou hipóteses específicas. Estudo de caso-controle, coorte, ...



# Estatística descritiva

Ramo da estatística que aplica várias técnicas que tentam descrever e sumarizar um conjunto de dados.

- Medidas de tendência central (Média, mediana, moda, ...)
- Medidas de dispersão (Variância, desvio padrão, ...)
- Medidas de associação (correlação)



**Como eu devo conduzir  
um experimento?**



**Você é uma pessoa alta  
entre os britânicos?**



Princípios importantes em um desenho experimental:

- Tamanho da amostra;
- Aleatoriedade;
- Blocagem → eliminar fontes de variação;
- Replicação → o experimento deve ser replicável;

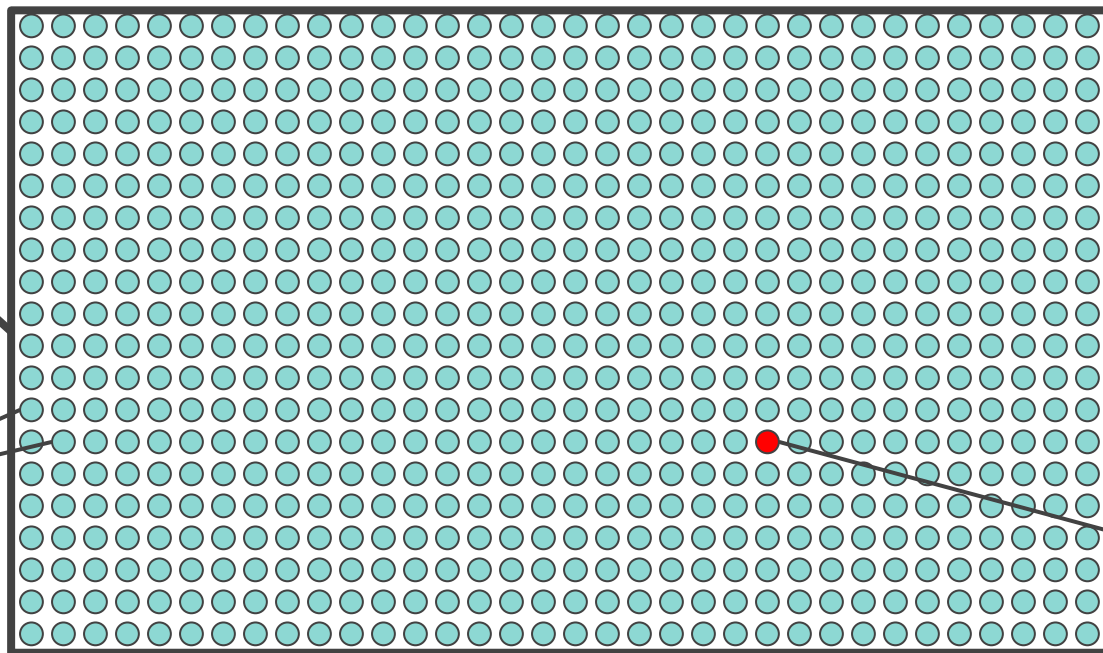
<https://www.dailymail.co.uk/news/article-2176733/At-home-worlds-tallest-family-Hes-7ft-shes-6ft-5in-new-baby-nearly-2ft-long.html>





População

Indivíduos

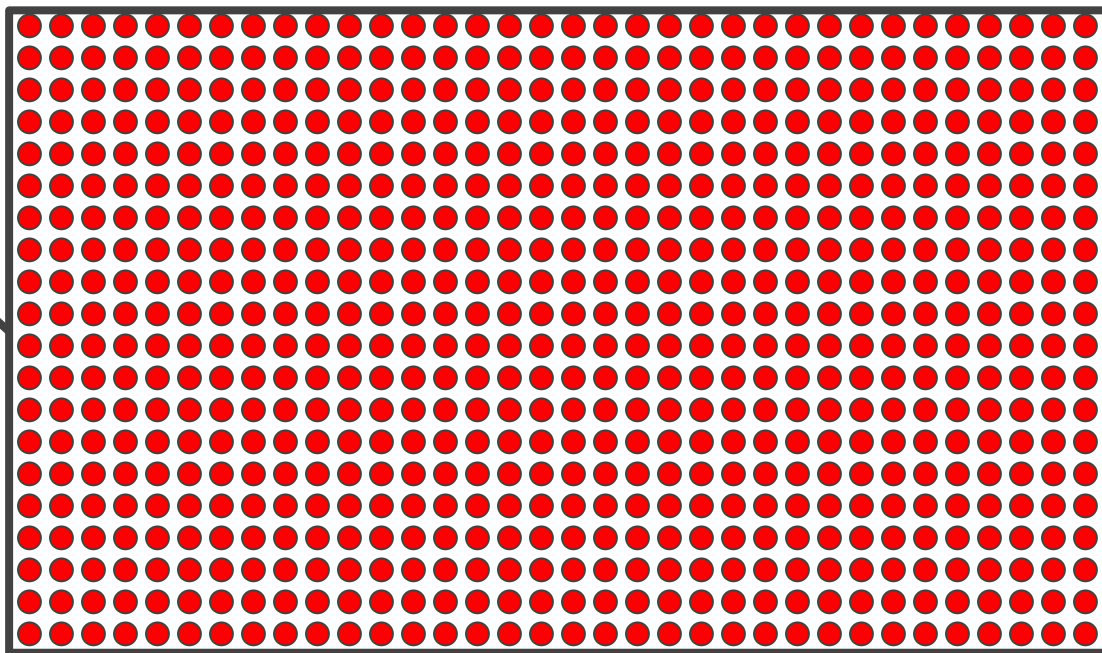


Indivíduo  
amostrado



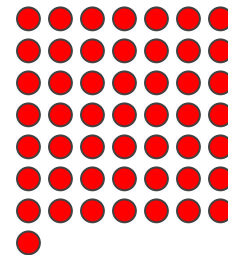
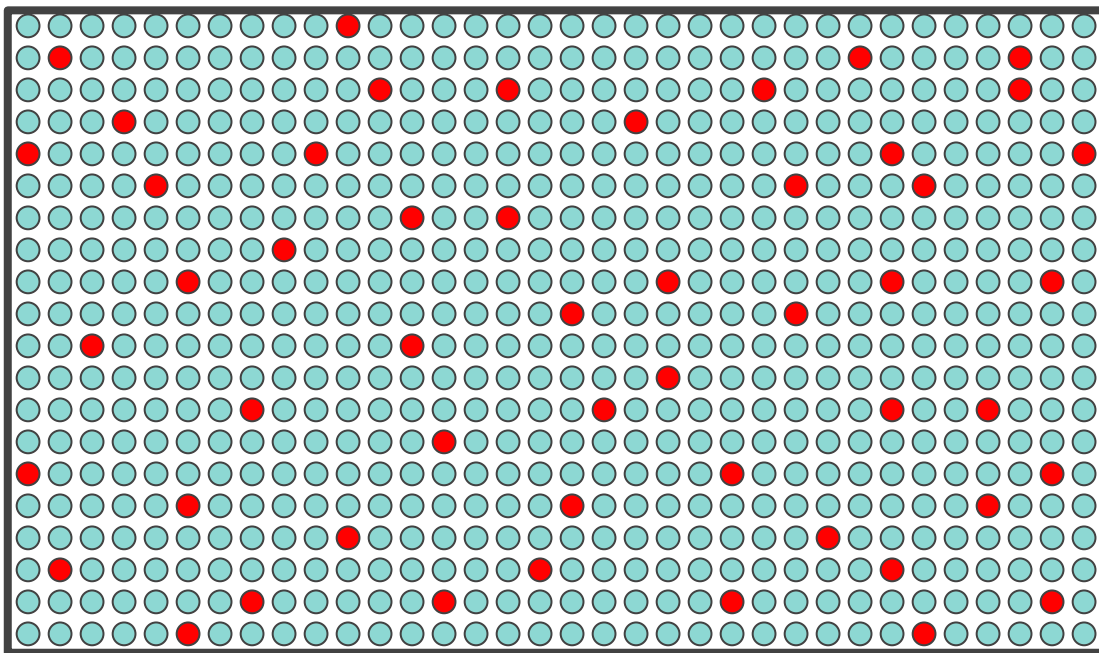
## Em um mundo ideal...

População





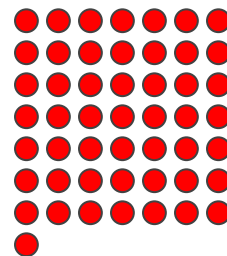
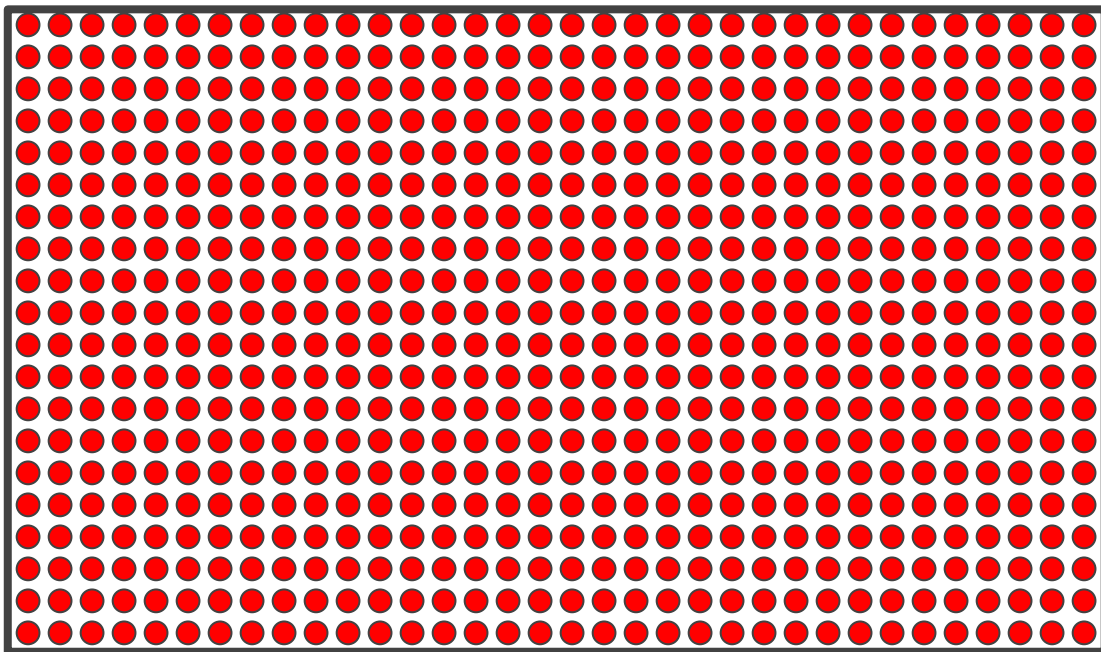
# Nem sempre é possível amostrar todos os indivíduos...



amostra



Média da  
população  
( $\mu$ )



Média da  
amostra  
( $\bar{x}$ )

$$\mu \approx \bar{x}$$



**Para que a média amostral se aproxime da média populacional...**

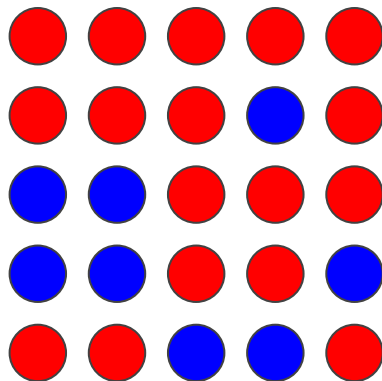
$$\bar{X} \approx \mu \text{ ou } (\bar{X} - \mu) \approx 0$$

erro amostral

- Tamanho da amostra;
- Aleatoriedade;
- Blocagem;
- Replicação;



# Influência do tamanho da amostra nos resultados

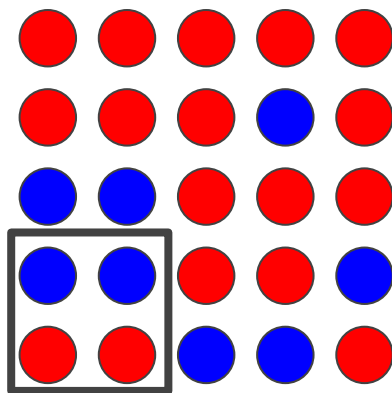


0.32 ●

0.68 ●



# Influência do tamanho da amostra nos resultados

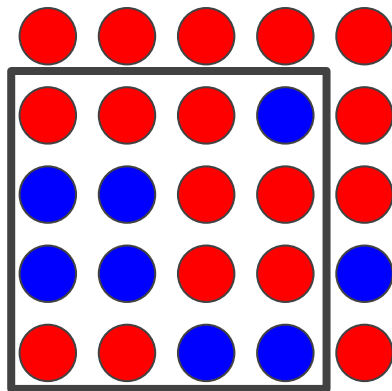


0.5 ●

0.5 ●



## Influência do tamanho da amostra nos resultados



0.4375 ●

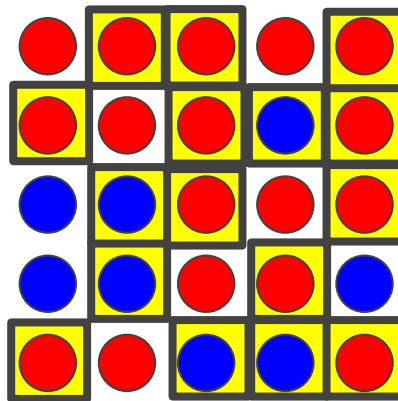
0.5625 ●





# Influência da aleatoriedade nos resultados

Todos os indivíduos possuem a mesma chance de serem amostrada



0.3125 ● (blue)

0.6875 ● (red)

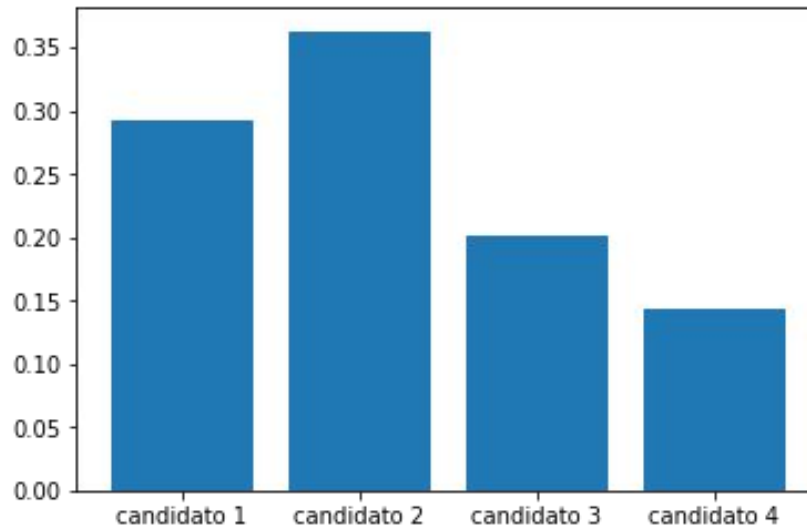
**Tamanho amostral +  
aleatoriedade + blocagem  
=  
amostra representativa  
da população**



# Intenção de votos

Você confia nesses dados?

O que você leva em conta para  
confiar nos dados?



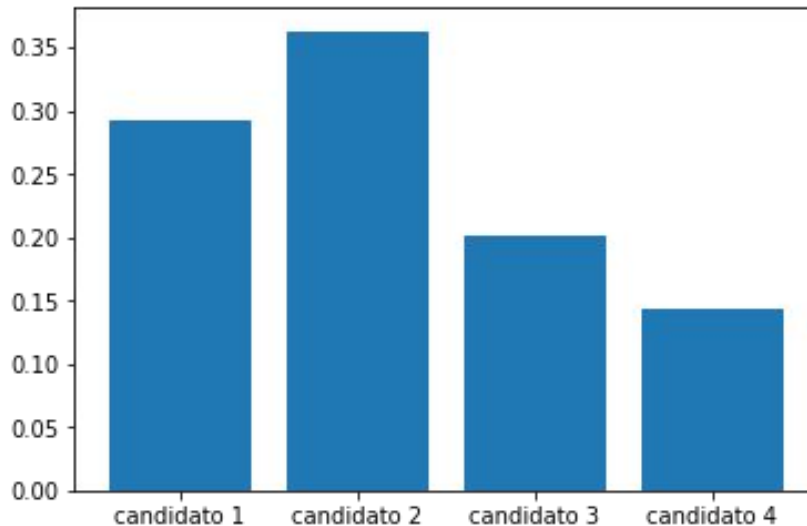


# Intenção de votos

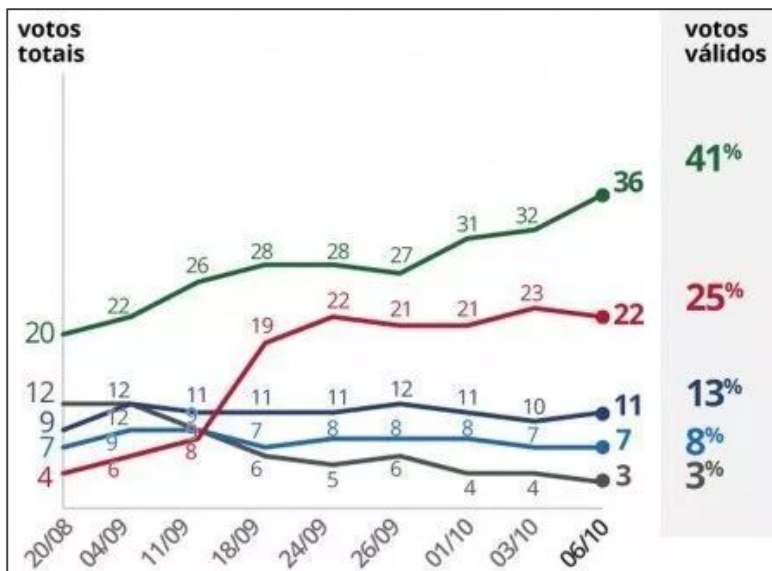
Você confia nesses dados?

O que você leva em conta para confiar nos dados?

- Quantas pessoas foram entrevistadas?
- As pessoas foram selecionadas de forma aleatória?
- Como a pesquisa foi conduzida?



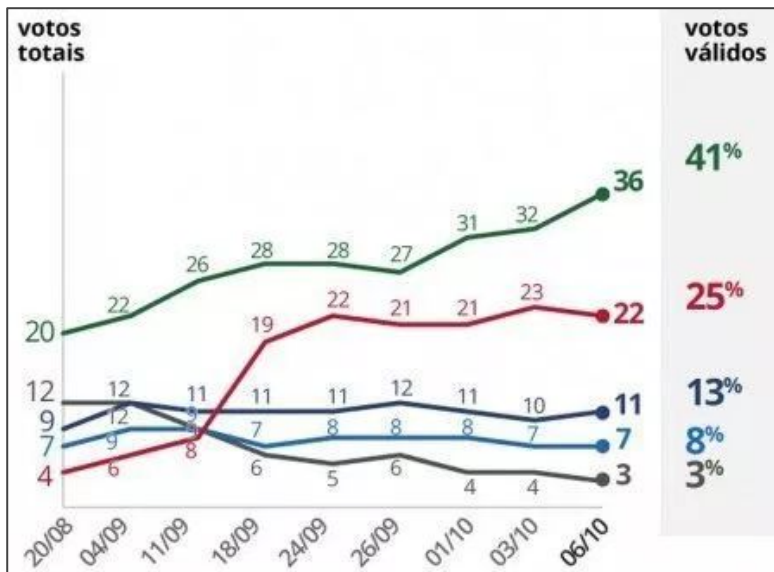
# Eleições 2018



19.552 eleitores em 382 municípios

5 e 6 de outubro

# Eleições 2018



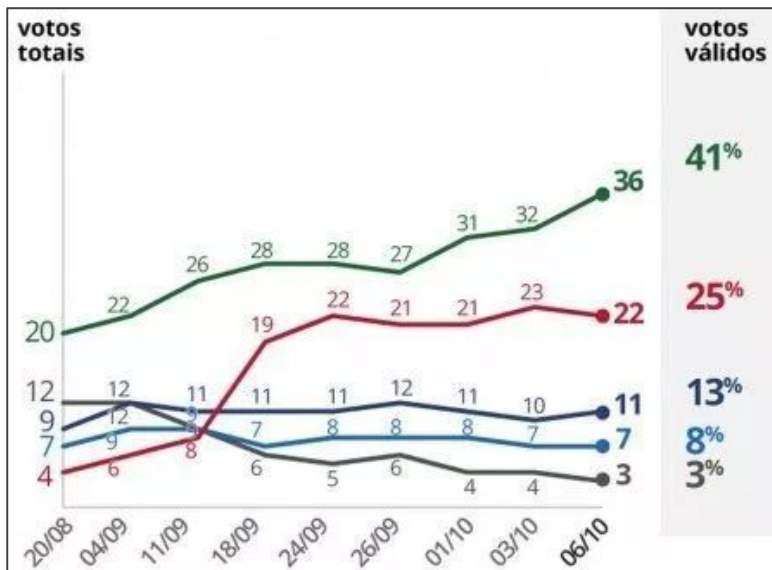
19.552 eleitores em 382 municípios

5 e 6 de outubro

(107.050.673 eleitores no total)

(5.570 municípios no Brasil)

# Eleições 2018



## Apuração

AGREGADOR DE PESQUISAS

DADOS ABERTOS DE PESQUISAS

FILTRO AVANÇADO DE PESQUISAS

ACERVO ANTIGO DE PESQUISAS

PÚBLICOS DO BRASIL

PRESIDENTE – BRASIL

1º turno

2º turno

Presidente

Governador

Senador

Dep. Federal

Dep. Estadual

Dep. Distrital

BRASIL

AC

AL

AM

AP

BA

CE

DF

ES

GO

MA

MG

MS

MT

PA

PB

PE

PI

PR

RJ

RN

RO

RR


RS

SC

SE

SP

TO




JAIR BOLSONARO

PSL

46,03%

Votos: 49.277.010




FERNANDO HADDAD

PT

29,28%

Votos: 31.342.051



CIRO GOMES











PDT

12,47%

Votos: 13.344.371

100,00% dos votos apurados

última atualização 15:44:58 - 24/10/2018

Candidato	Partido	Válidos	Votos
<div></div> <div>GERALDO ALCKMIN</div>	PSDB	4,76%	5.096.350
<div></div> <div>JOÃO AMOÊDO</div>	NOVO	2,50%	2.679.745
<div></div> <div>CABO DACIOLO</div>	PATRI	1,26%	1.348.323
<div></div> <div>HENRIQUE MEIRELLES</div>	MDB	1,20%	1.288.950
<div></div> <div>MARINA SILVA</div>	REDE	1,00%	1.069.578
<div></div> <div>ALVARO DIAS</div>	PODE	0,80%	859.601
<div></div> <div>GUILHERME BOULOS</div>	PSOL	0,58%	617.122
<div></div> <div>VERA</div>	PSTU	0,05%	55.762
<div></div> <div>EYMAEL</div>	DC	0,04%	41.710
<div></div> <div>JOÃO GOULART FILHO</div>	PPL	0,03%	30.176

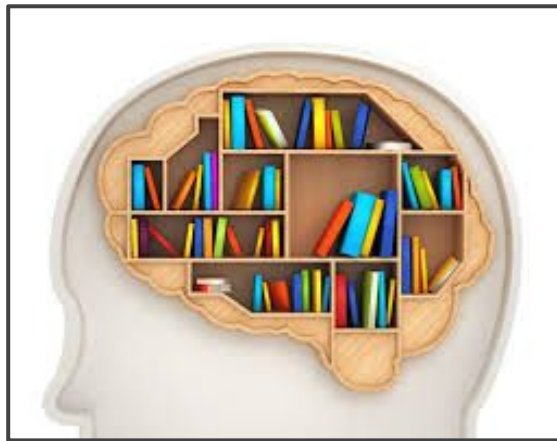
# Memória

Para tirar uma nota boa → Boa memória;

Como melhorar a memória?

- Comer bem a noite?
- Uma boa noite de sono?

Vários fatores influenciam a memória, mas sem haver uma forma de medir, não é possível estabelecer uma relação científica sobre a memória e os fatores em estudo.







**Como medir a memória?**



# Como medir a memória?

Existem vários testes que tentam medir a sua memória:

- <https://www.mindcrowd.org/> - Alzheimer's research
- <https://memtrax.com/test/>
- <https://testmybrain.org/>
- <https://www.humanbenchmark.com/dashboard>
- <https://brainlabs.me/>



# Conceito e definição operacional

**Conceitos** - objeto “ideal”, que não é diretamente observável.

- Memória;
- Inteligência;
- Fome;
- Saúde;
- Estresse;
- Felicidade;
- ...

**Definição operacional:** forma que torna um conceito observável e quantificável.

- Pontuação no teste de memória;
- Pontuação no teste de QI;
- Gramas de alimento ingerido;
- Taxa de batimentos cardíacos;
- Nível de cortisol (hormônio do estresse);
- Proporção de minutos sorrindo com os minutos não sorrindo;
- ...

**Eu devo dormir bem  
antes de uma prova?**





# Eu devo dormir bem antes de uma prova?



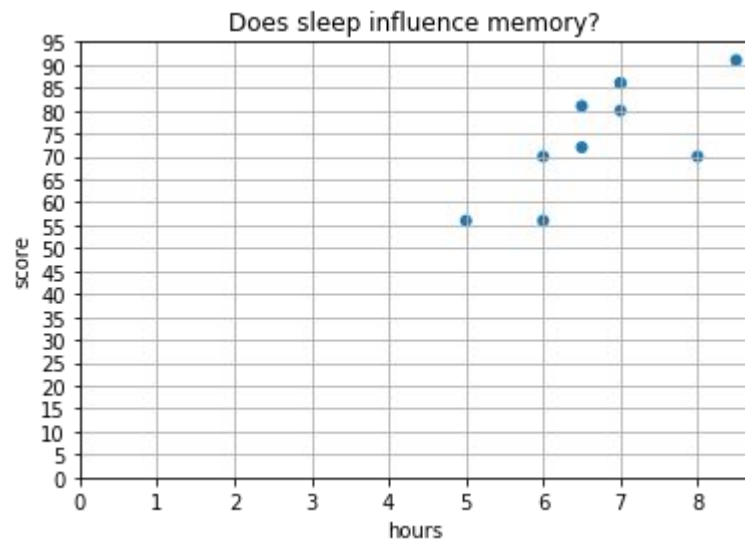
Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86



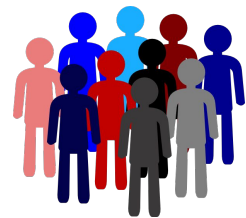
# Eu devo dormir bem antes de uma prova?



Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86



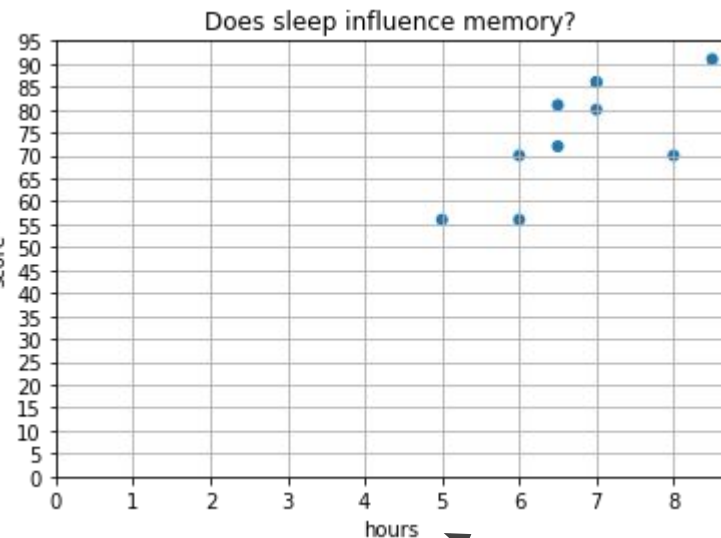
# Eu devo dormir bem antes de uma prova?



Variáveis

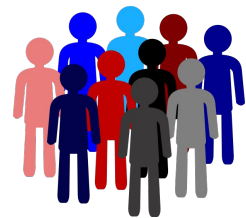
Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86

Variável dependente, resposta  
(Variável cujo valor depende  
da variável independente)



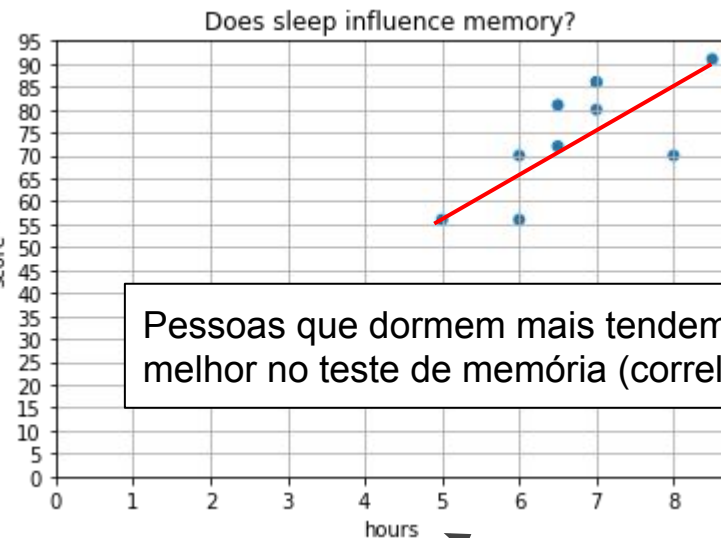
Variável independente, Predictor  
(variável que está sendo manipulado)

# Eu devo dormir bem antes de uma prova?



Variáveis	
Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86

Variável dependente, resposta  
(Variável cujo valor depende  
da variável independente)



Pessoas que dormem mais tendem a ir melhor no teste de memória (correlação).

Variável independente, Preditor  
(variável que está sendo manipulado)



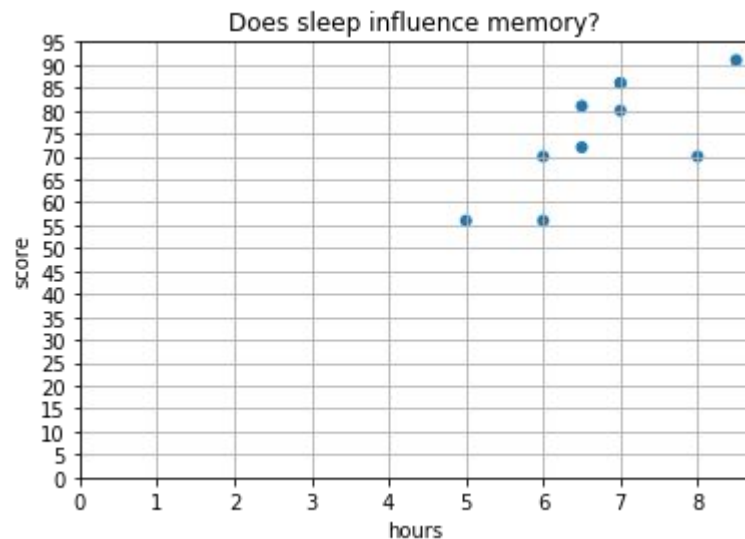
**Se eu deitar cedo, a  
minha memória com  
certeza será melhor  
amanhã?**





# Se eu deitar cedo...

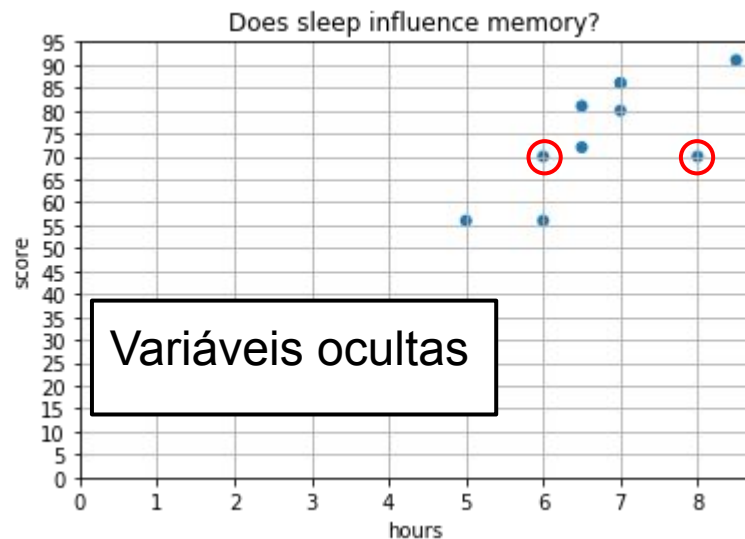
Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86





# Se eu deitar cedo...

Horas de sono	Pontuação
7	86
8	70
6	56
5	56
6	70
7	80
6.5	72
8.5	91
6.5	81
7	86





# Variáveis ocultas

Variáveis que não estão incluídas como variáveis independentes ou dependentes na análise, mas que podem afetar na interpretação da relação entre as variáveis em análise;

Em outras palavras: características de cada indivíduo que não estão registradas e não são consideradas na análise, mas que podem influenciar no resultado do experimento;

Em estudos de observação podemos verificar uma **correlação** (ou **associação**) entre duas variáveis (horas de sono e memória);

“Correlação **não** prova causa”.



# Correlação não prova a causa



Exemplo extremo:

- **Teoria dos arcos dourados de prevenção de conflitos**

“Dois países que possuem McDonald’s nunca entraram em guerra desde a sua entrada”  
(Thomas Friedman)

Se todos os países construíssem um McDonald’s, alcançaríamos a paz mundial.

# Correlação não prova a causa





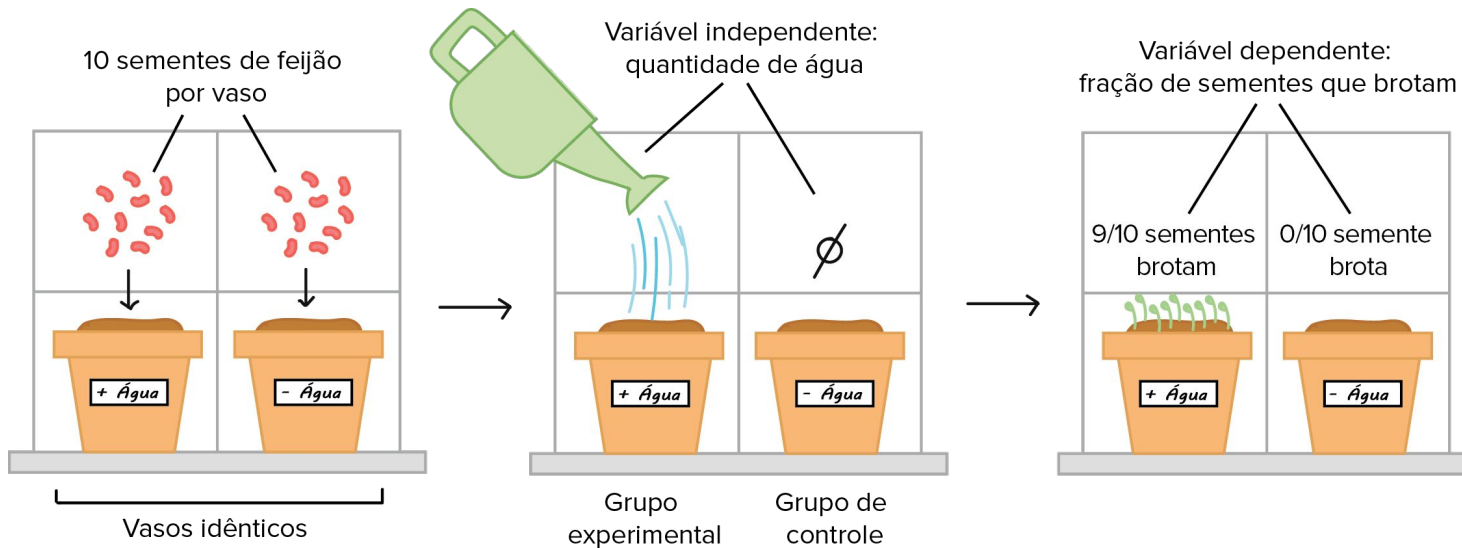
# Para provar que uma variável causa um efeito...

## Métodos analíticos - Caso-controle

- Grupo experimental
- Grupo controle

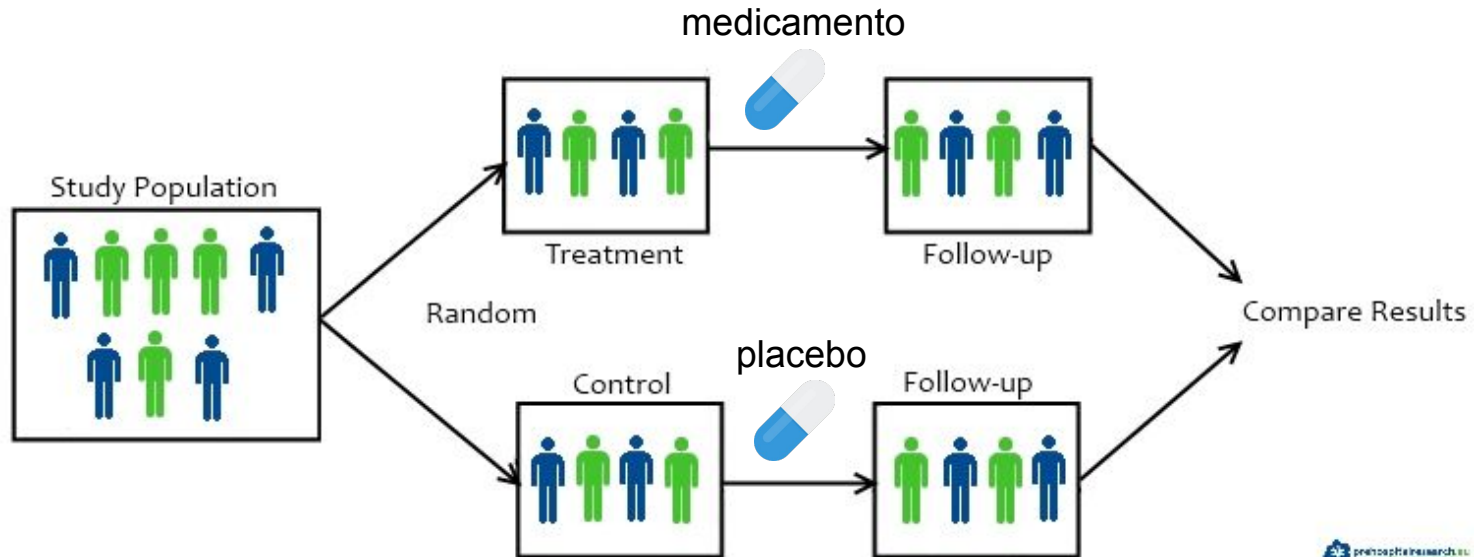
Exemplo: Feijão

# Experimento controlado





# Experimento controlado - teste de medicamento

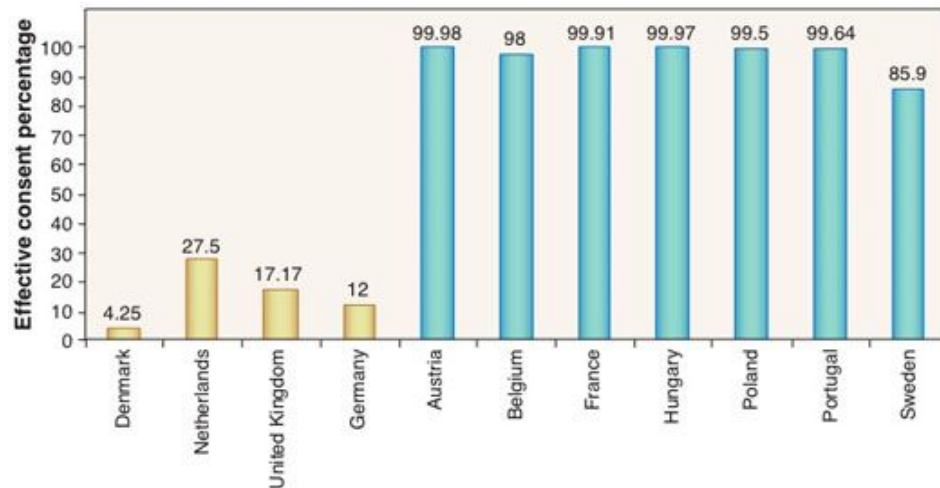


# Um pouco além do desenho experimental

Pessoas que consentem em doar os órgãos em caso de acidente no trânsito.

<http://danariely.com/2008/05/05/3-main-lessons-of-psychology/>

[https://www.ted.com/talks/dan\\_ariely\\_asks\\_are\\_we\\_in\\_control\\_of\\_our\\_own\\_decisions?language=pt-br](https://www.ted.com/talks/dan_ariely_asks_are_we_in_control_of_our_own_decisions?language=pt-br)



**Effective consent rates, by country.** Explicit consent (opt-in, gold) and presumed consent (opt-out, blue).

# Um pouco além do desenho experimental

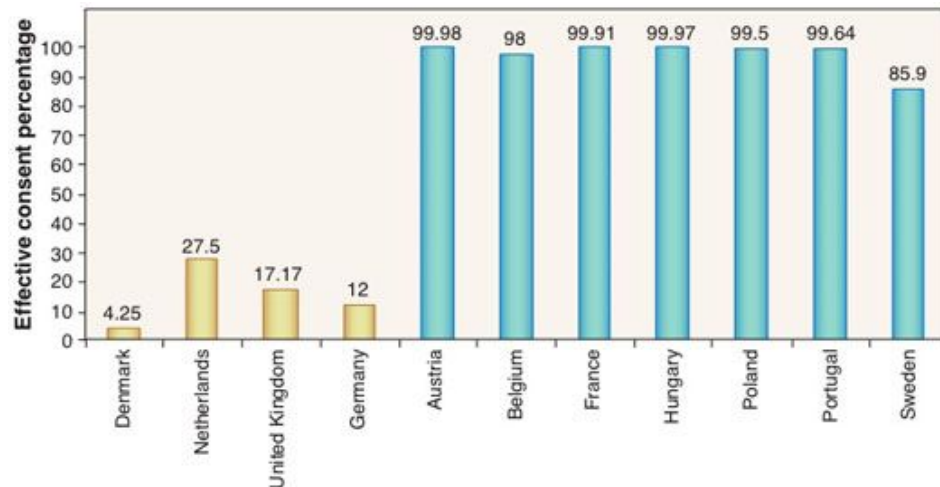
Diferença no formulário:

Amarelo → Você quer ser doador...

Azul → Você não quer ser doador...

<http://danariely.com/2008/05/05/3-main-lessons-of-psychology/>

[https://www.ted.com/talks/dan\\_ariely\\_asks\\_are\\_we\\_in\\_control\\_of\\_our\\_own\\_decisions?language=pt-br](https://www.ted.com/talks/dan_ariely_asks_are_we_in_control_of_our_own_decisions?language=pt-br)



**Effective consent rates, by country.** Explicit consent (opt-in, gold) and presumed consent (opt-out, blue).