

# ESTATÍSTICA

**Prof. Elizeu Martins  
de Oliveira Junior**

**2025/2**

**@elizeuoliveira.jr**

# Horário das aulas e presença

- 19:00 as 20:40
- 21:00 as 22:40

## Material e Anotações

**PROVAS**

**QUADRO**

## Participação nas aulas

**Quizzes e etc.**

**Celular**

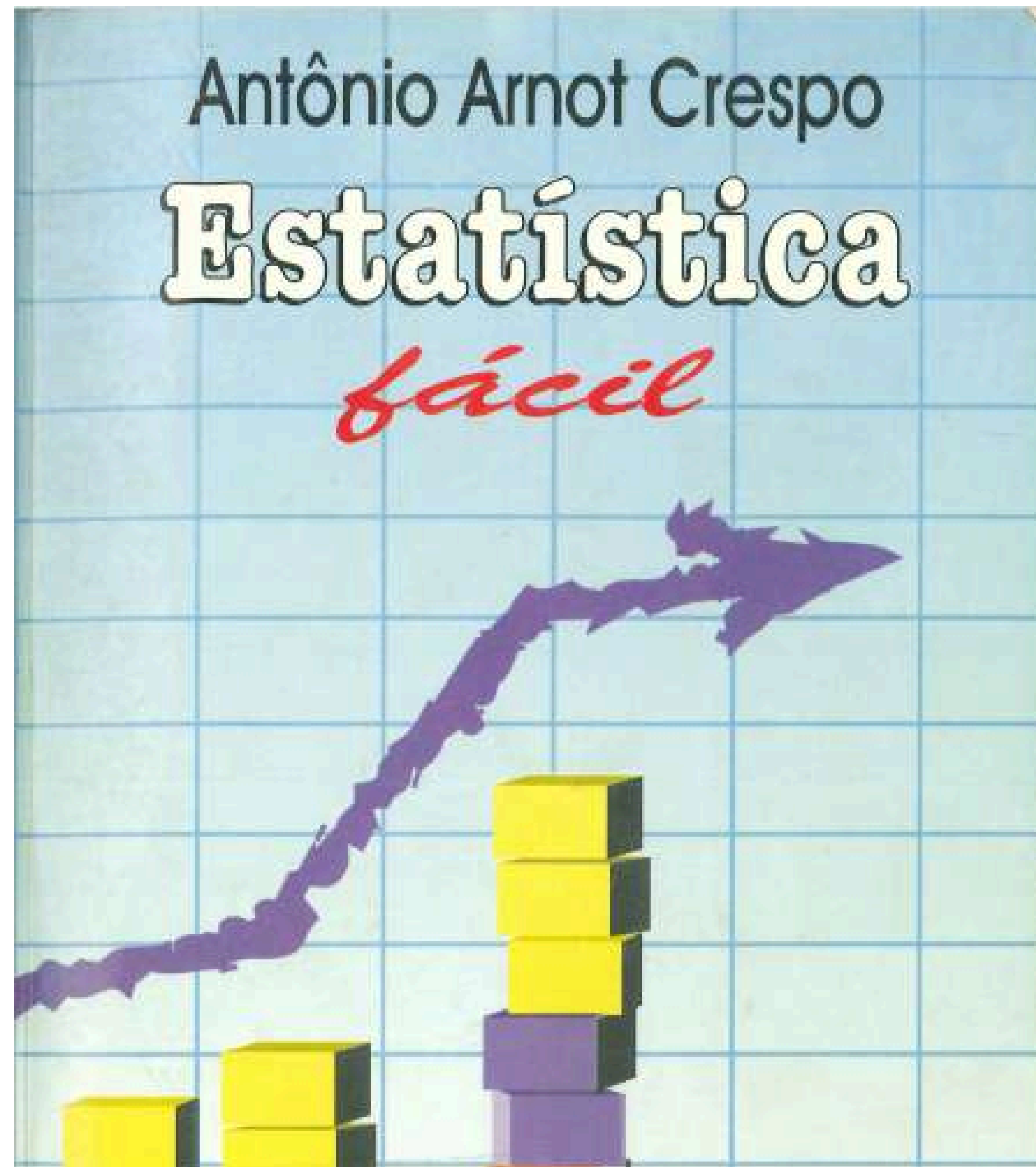
## SEMINÁRIO

**PROMPT/I.A.**

# Provas e Trabalhos

Avaliações (**A**) referente a UNIDADE 1, 2 e 3: Prova 1 representará 60% da nota, **Seminário 1, Quiz (Unidade 2) e Seminário 3** representará 30% da nota e Interação/Participação nas aulas e nas atividades representará 10% da nota;

# UNIDADE 1



# Por que estudar Estatística?

É porque uma grande parte das nossas *decisões* e *questionamentos* envolvem riscos é que precisamos estudar, e bem, a estatística.

Quantas pessoas em nossa cidade têm mais de 1,80m de altura?

Como está a saúde mental dos estudantes da graduação e Pós-graduação da Unemat?

Avaliar o potencial de mercado na área para determinar se abrir uma filial é viável.



# Eleições: Estatística nas urnas

## Pesquisas de Intenção

Entenda como as pesquisas medem a preferência dos eleitores.

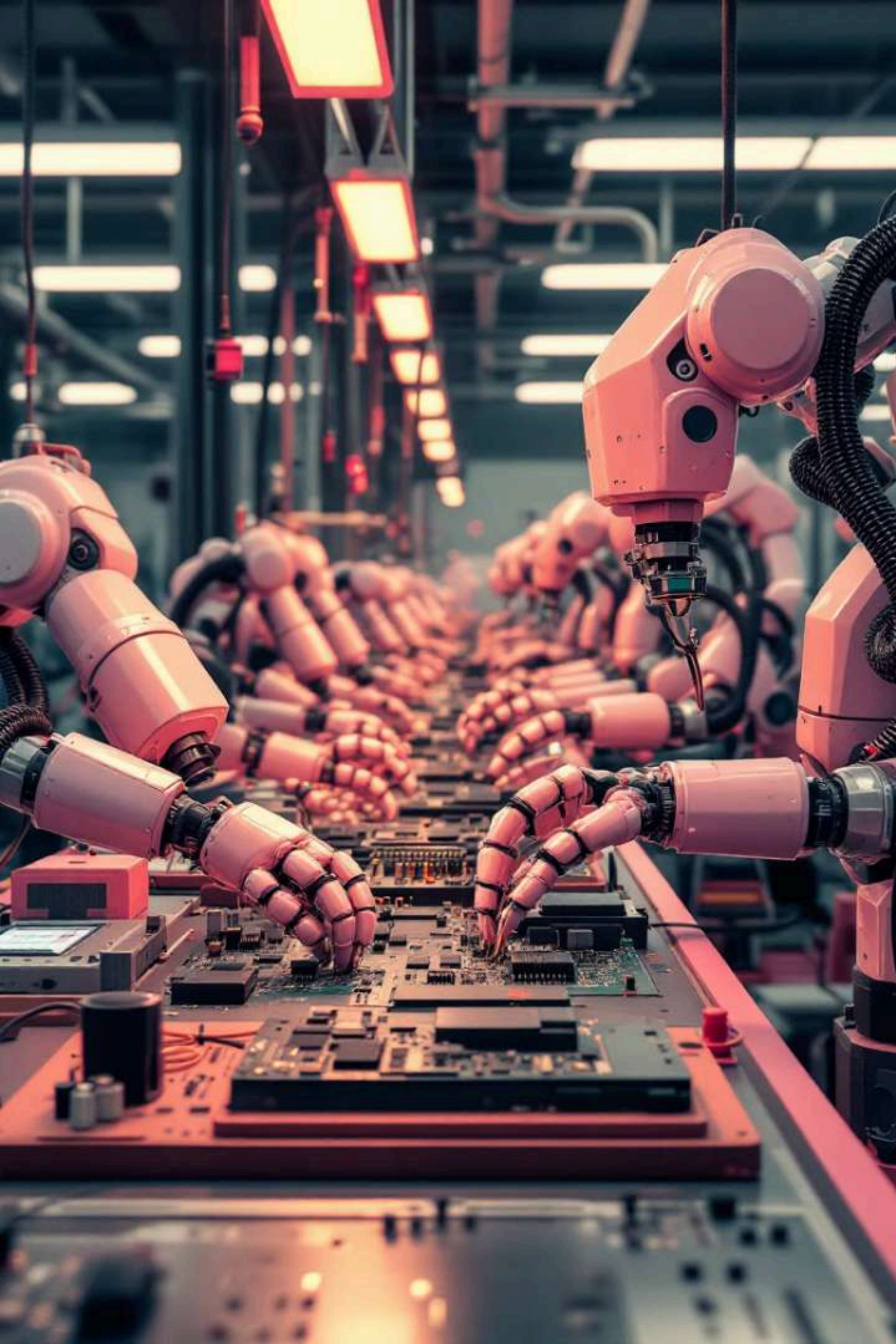
## Margem de Erro

Saiba como a margem de erro impacta a precisão das previsões.

## Intervalos de Confiança

Descubra como estimar a probabilidade de um candidato vencer.

O anúncio prévio, em termos percentuais, de que o candidato A é o provável vencedor pode ser feito após a contagem de 2% dos votos.



# Indústria: Controle estatístico da qualidade

1

## **Controle de Qualidade**

A estatística garante a conformidade dos produtos.

2

## **Amostragem**

A amostragem eficiente economiza recursos e tempo.

3

## **Gráficos de Controle**

Os gráficos monitoram a variação e detectam problemas.

O Recall é um exemplo, onde ocorre a retirada ou troca de produtos devido a uma falha em algum estágio de fabricação não detectado pelo controle de qualidade).



# Bancos: Estatística na análise de crédito

## **Análise de Risco**

A estatística avalia a probabilidade de perdas financeiras.

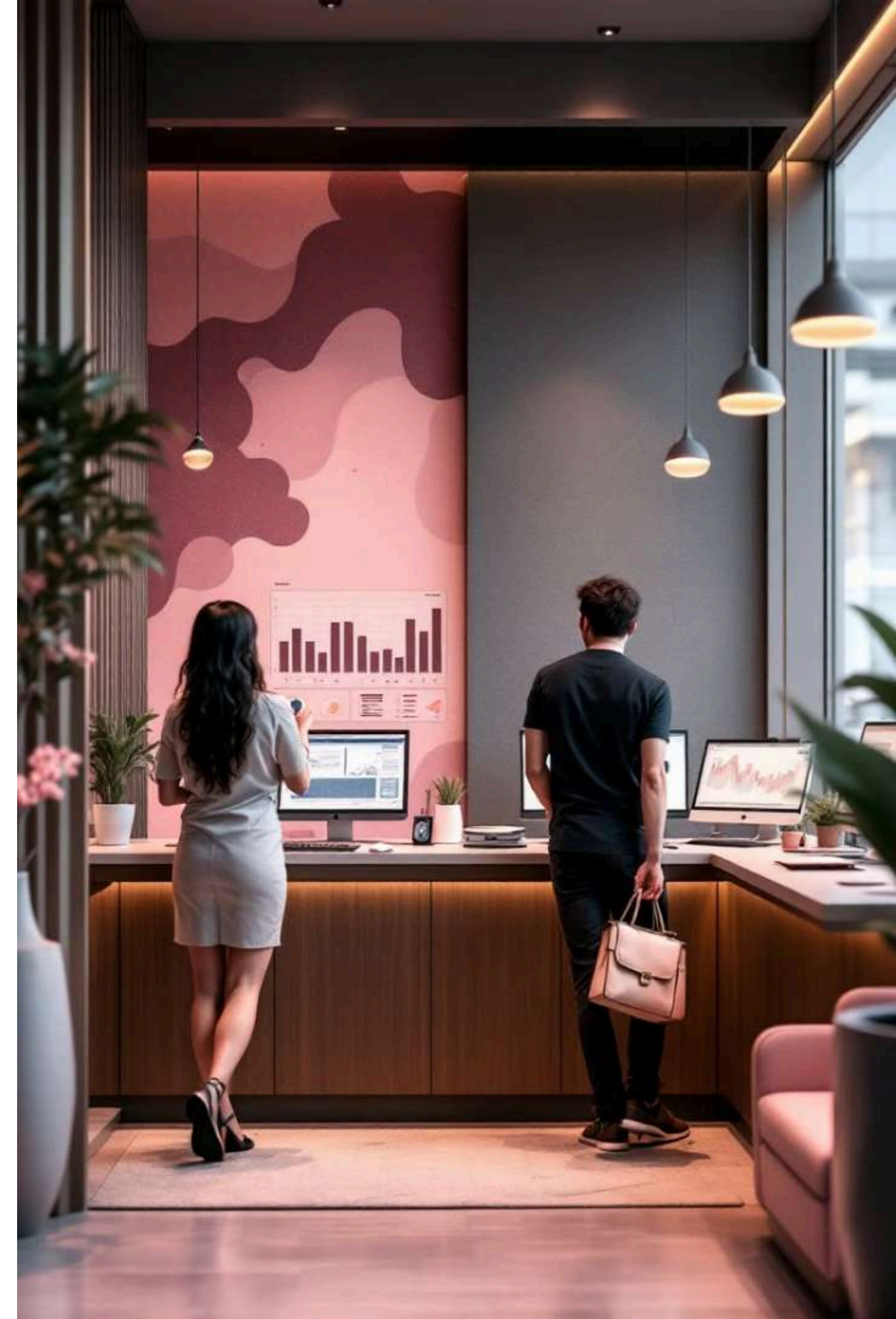
## **Credit Score**

O credit score quantifica a capacidade de pagamento.

## **Inadimplência**

A estatística prevê a probabilidade de não pagamento.

A estatística é fundamental na análise de crédito e na definição de seguros de vida e automóveis.





# IBGE: Estatística para entender o Brasil



## PIB

Produto Interno Bruto: a riqueza do país.



## IPCA

Índice de Preços ao Consumidor Amplo: a inflação.



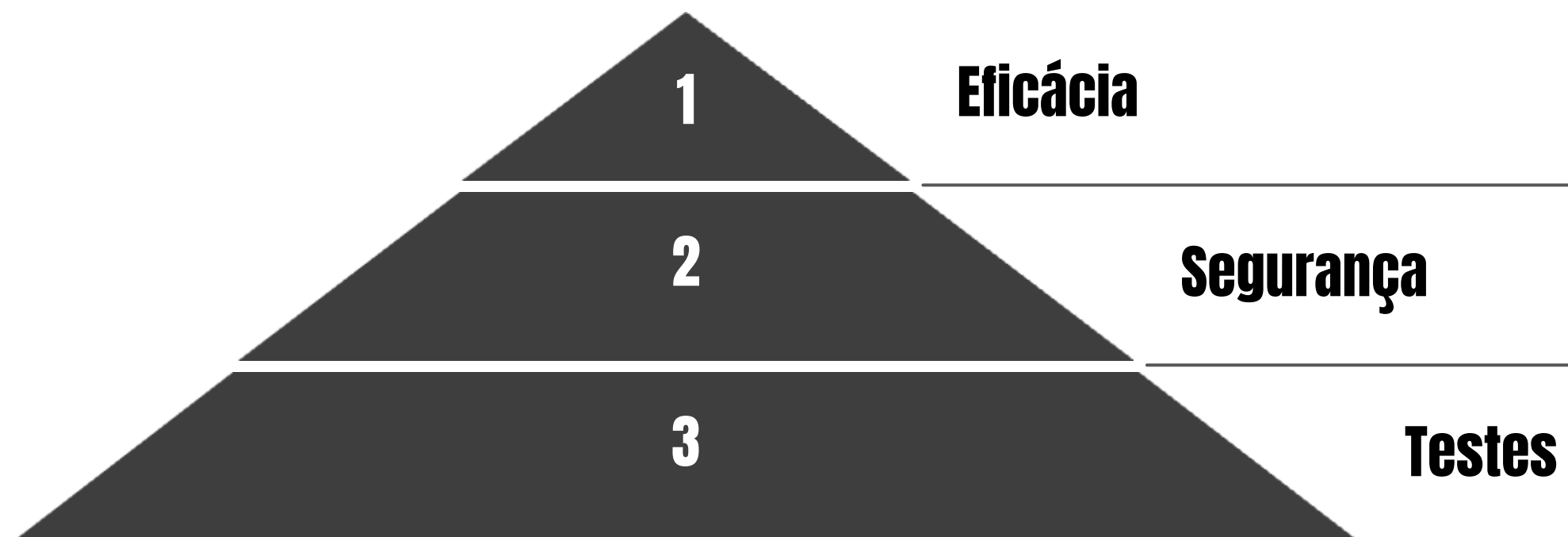
## Taxa de Desemprego

O percentual de pessoas sem trabalho.

O IBGE utiliza metodologias de pesquisas amostrais para calcular diversos índices, como taxa de mortalidade/natalidade e índices de analfabetismo.



# Farmácia e Medicina: Estatística para a saúde



A estatística é essencial nos testes clínicos para avaliar a eficácia de medicamentos e garantir a segurança dos pacientes. A tomada de decisão deve ser cautelosa.



# Meteorologia: Estatística e previsão do tempo

## Coleta de Dados

## Sensores e satélites coletam dados atmosféricos.

# Modelos Preditivos

# Modelos estatísticos preveem o clima futuro.

# Probabilidade

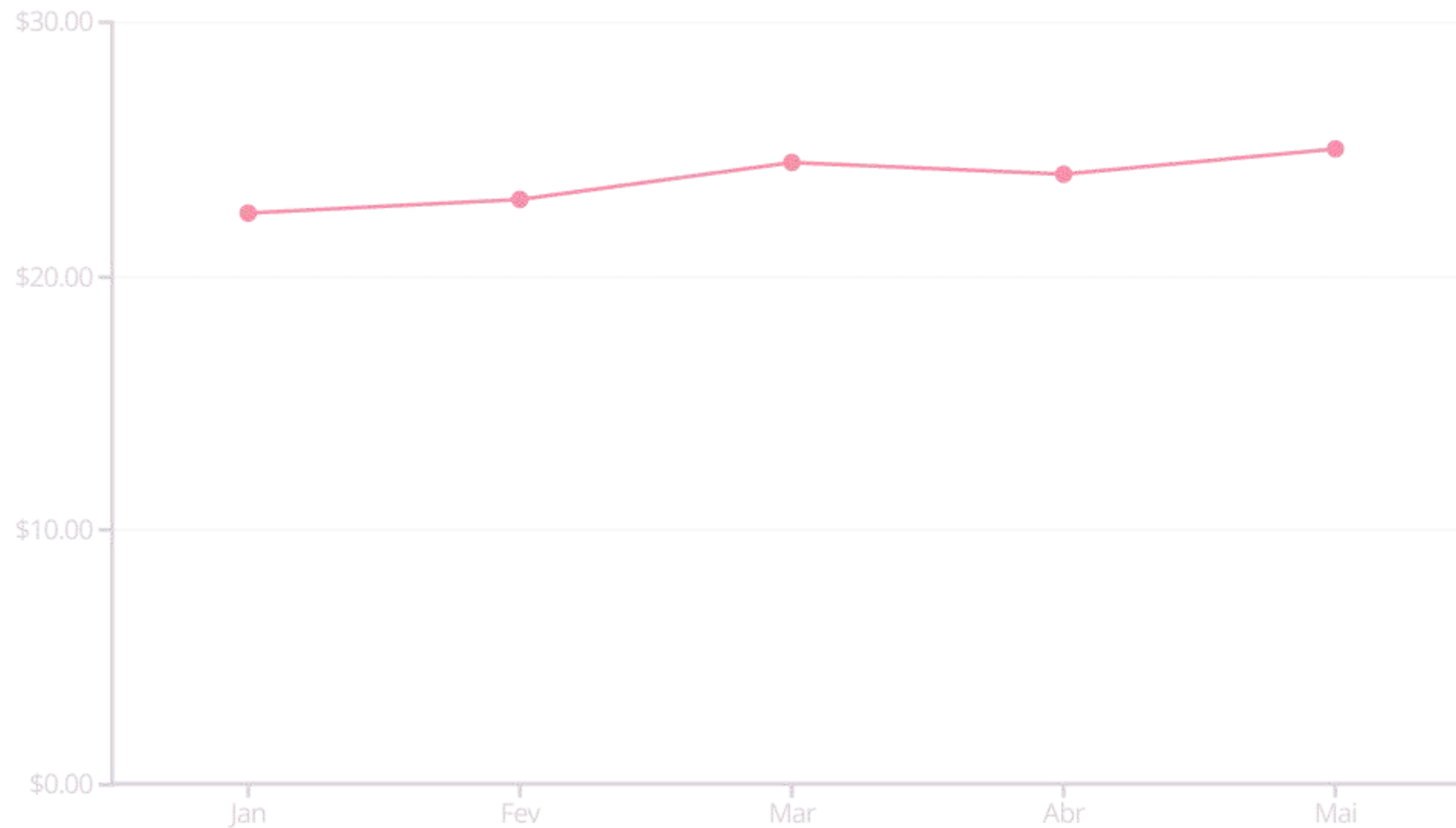
A estatística quantifica a chance de chuva.

A estatística, através de modelos preditivos e séries temporais, informa a "chance" de chuva, como 30% para hoje.





# Mercado de Valores: Estatística para investir



A estatística auxilia na análise de tendências, volatilidade e indicadores financeiros para a tomada de decisão na compra ou venda de ações, considerando o risco envolvido.



# Experimentação Agrícola: Estatística no campo

## Novas Sementes

A estatística avalia o rendimento de novas variedades.

## Fertilizantes

A estatística otimiza o uso de fertilizantes.

## Produtividade

A estatística aumenta a produtividade agrícola.

A estatística é fundamental em estudos de novas variedades de sementes (modificação genética) ou de fertilizantes para aumentar a produtividade agrícola.



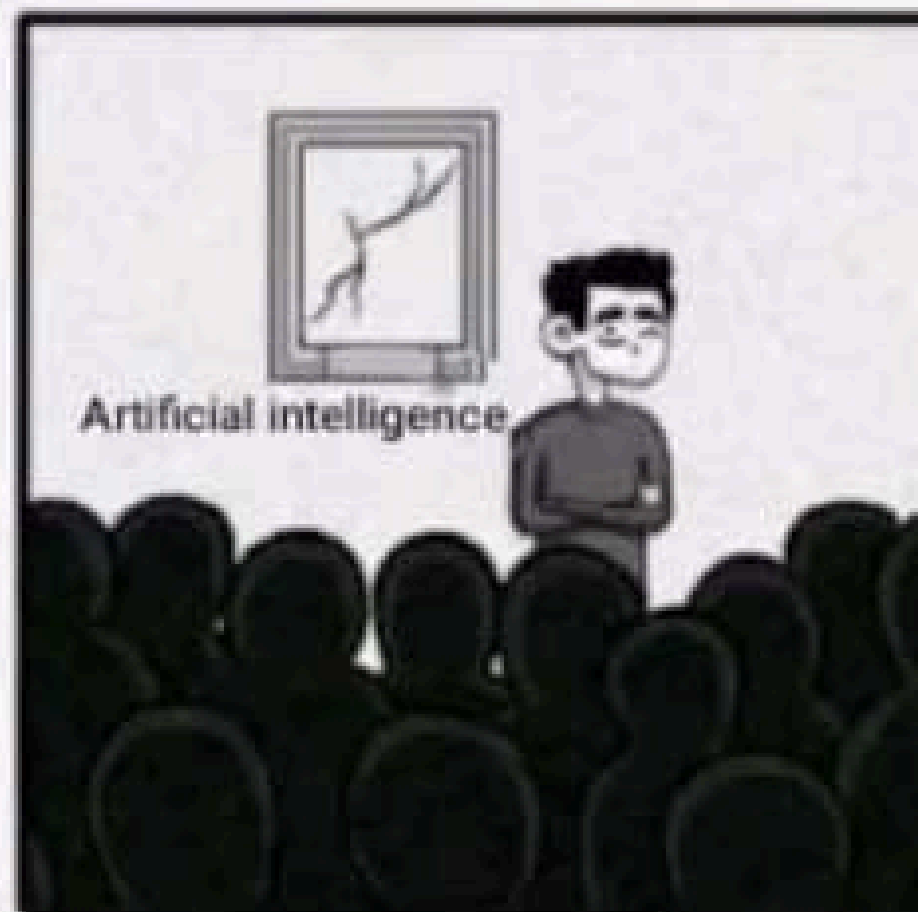
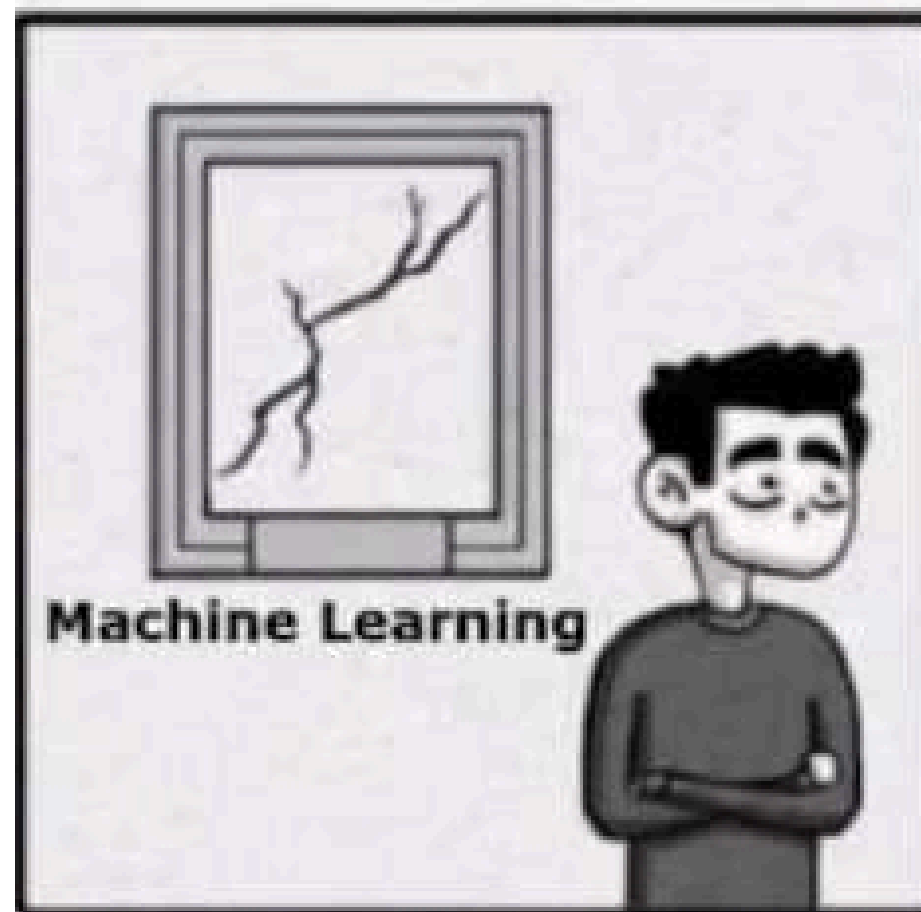
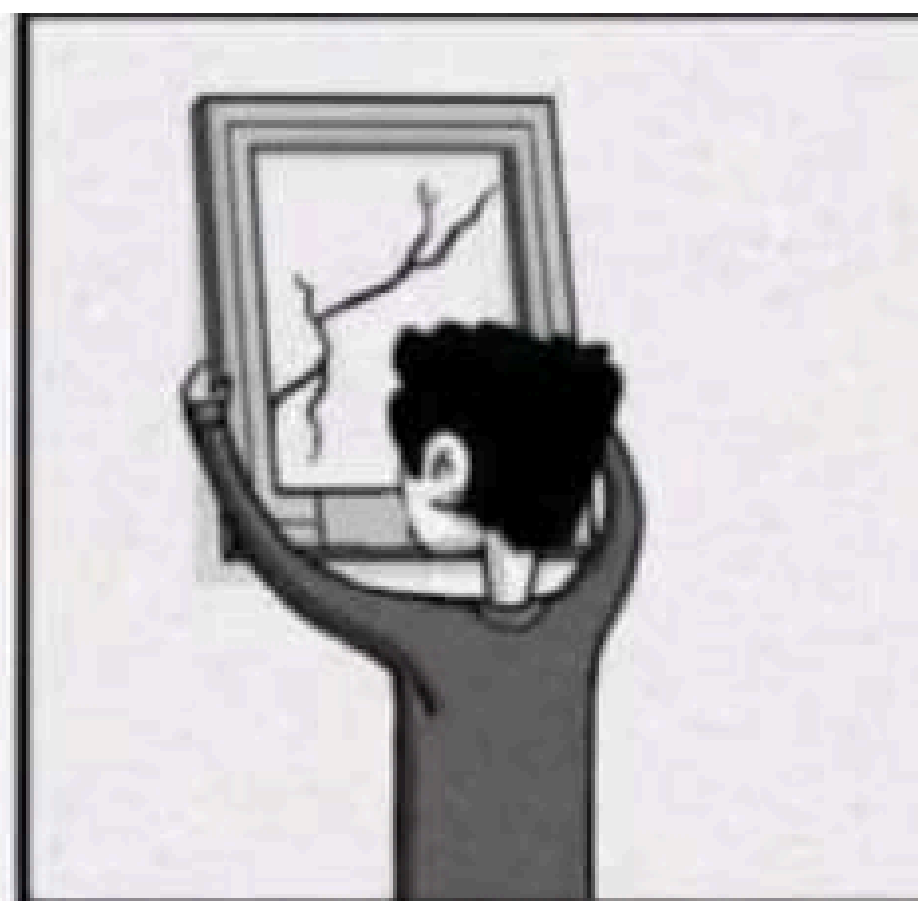
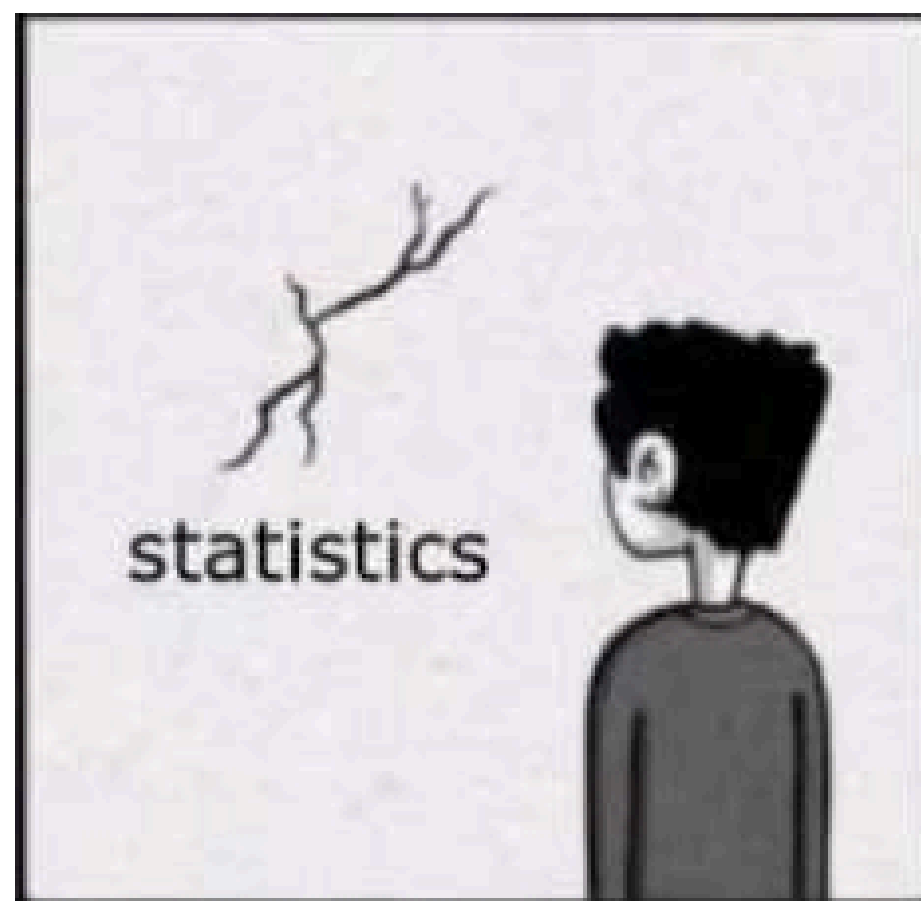


A Estatística é fundamental para a área de BigDATA e principalmente nos LLMs.



Ela ajuda a extrair informações úteis e insights dos dados, permitindo que os algoritmos determinem padrões e tomem decisões mais apropriadas.

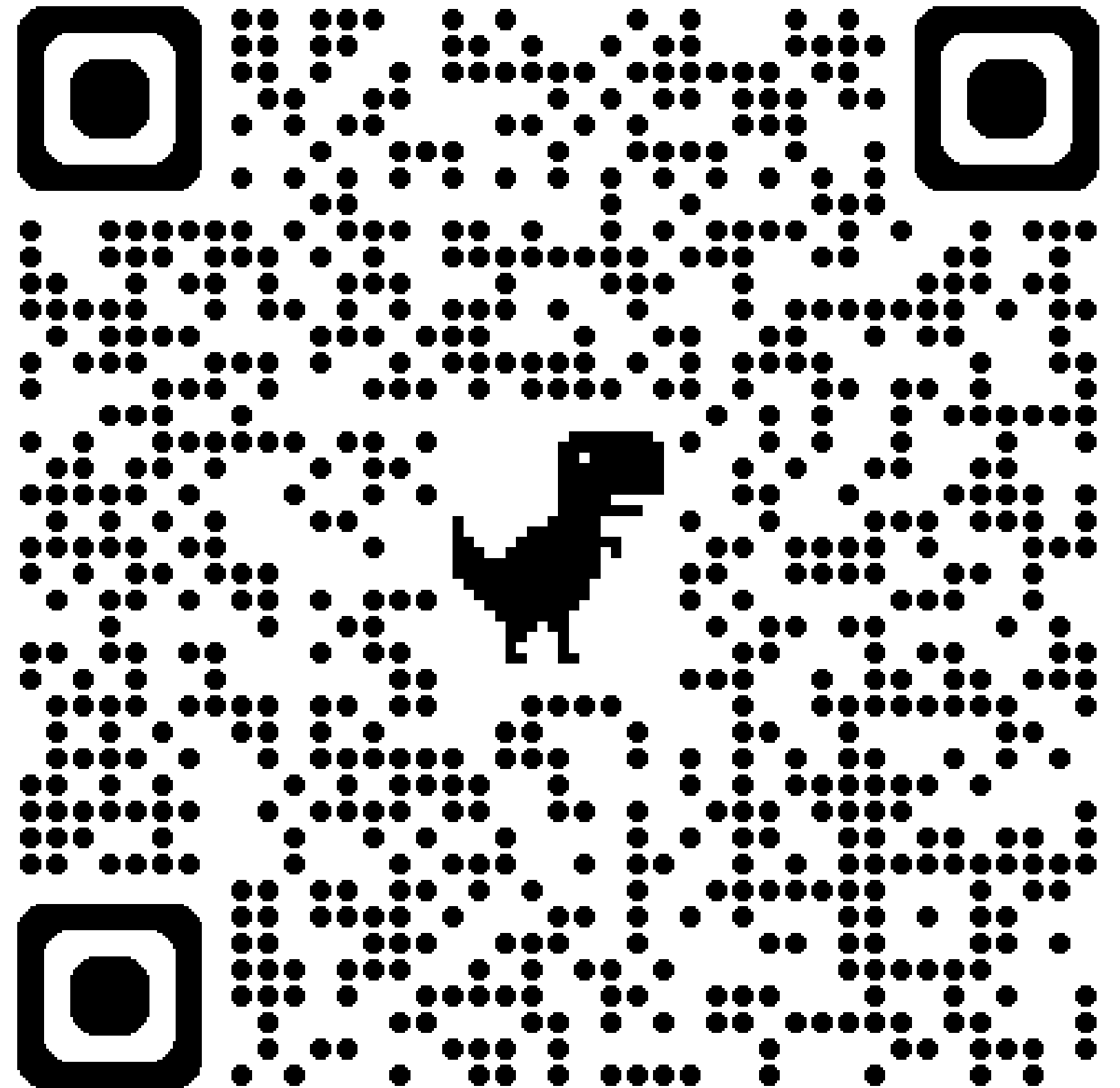




# PESQUISA DE INFORMAÇÃO E CONHECIMENTOS PRÉVIOS DE ESTATÍSTICA



<https://abre.ai/nfwm>



# DEFINIÇÃO

## Por que chama Estatística?

A palavra **estatística** vem do latim "status" que significa “estado”.

# DEFINIÇÃO

Tinha como função o registro de dados (nº de habitantes da população, nº de casamentos...) e a elaboração de tabelas e gráficos para descrever resumidamente um determinado país em números.



**A estatística pode ser dividida em duas:**

**ESTATÍSTICA DESCRITIVA:** Envolve a *coleta*, a *organização*, a *descrição*, dos dados e o cálculo e interpretação de coeficientes. Esta parte está associada a cálculos de médias, variâncias, estudo de gráficos, tabelas, etc.. É a parte mais conhecida.





# A estatística pode ser dividida em duas:

**ESTATÍSTICA INDUTIVA** ou **INFERENCIAL**: Envolve a *análise* e a *interpretação* dos dados (associados a uma margem de incerteza)





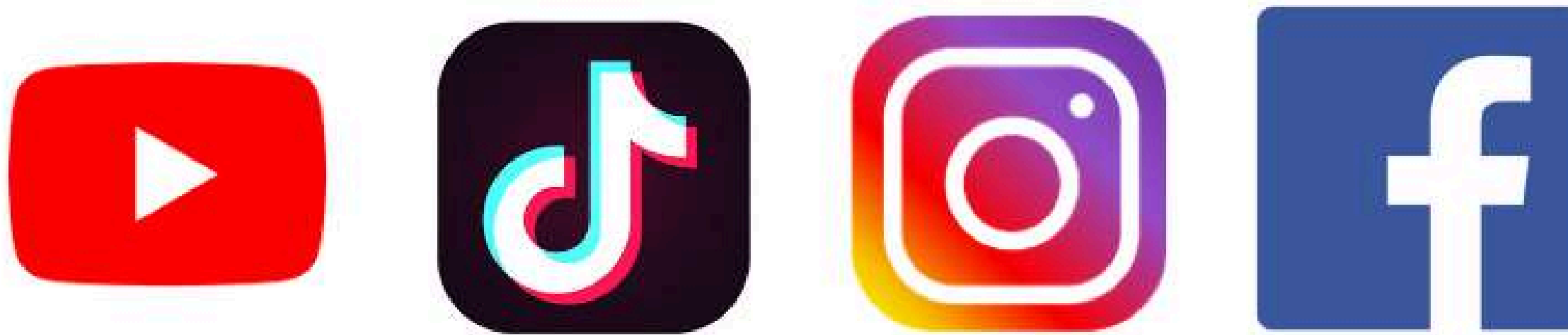
A Estatística Inferencial visa tirar conclusões sobre uma população a partir da análise de **amostras** retiradas dela, baseando-se nos resultados observados.

**Variável** é, convencionalmente, o conjunto de resultados possíveis de um fenômeno.

# VARIÁVEIS

Por exemplo:

- para o fenômeno "sexo" são dois os resultados possíveis: sexo masculino e sexo feminino;
- para o fenômeno "número de filhos" há um número de resultados possíveis expresso através dos números naturais: 0, 1, 2, 3, ...,  $n$ ;
- para o fenômeno "estatura" temos uma situação diferente, pois os resultados podem tomar um número infinito de valores numéricos.



- Plataforma de vídeo preferida: TikTok, YouTube, Instagram...  
As respostas são nomes, palavras. É uma **variável:**  
**qualitativa**
- Quantidade de vídeos que assiste por dia: 5, 10, 20... As  
respostas são números inteiros. É uma **variável:**  
**quantitativa discreta**
- Tempo que passa assistindo vídeos por dia: 30 minutos, 1  
hora e 15 minutos, 2 horas e 30 segundos... As respostas  
podem ter vírgulas ou frações de tempo. É uma variável:  
**quantitativa contínua**

**a. qualitativa** — quando seus valores são expressos por atributos: sexo (masculino — feminino), cor da pele (branca, preta, amarela, vermelha, parda) etc.;

**b. quantitativa** — quando seus valores são expressos em números (salários dos operários, idade dos alunos de uma escola etc.). Uma variável quantitativa que pode assumir, teoricamente, qualquer valor entre dois limites recebe o nome de **variável contínua**; uma variável que só pode assumir valores pertencentes a um conjunto enumerável recebe o nome de **variável discreta**.

# POPULAÇÃO E AMOSTRA

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos **população estatística** ou **universo estatístico**.

Na maioria das vezes, por impossibilidade ou inviabilidade econômica ou temporal, limitamos as observações referentes a uma determinada pesquisa a apenas uma parte da população. A essa parte proveniente da população em estudo denominamos **amostra**.

Uma amostra é um subconjunto finito de uma população.





# AMOSTRAGEM

Existe uma técnica especial — amostragem — para recolher amostras, que garante, tanto quanto possível, o acaso na escolha.

## **AMOSTRAGEM**

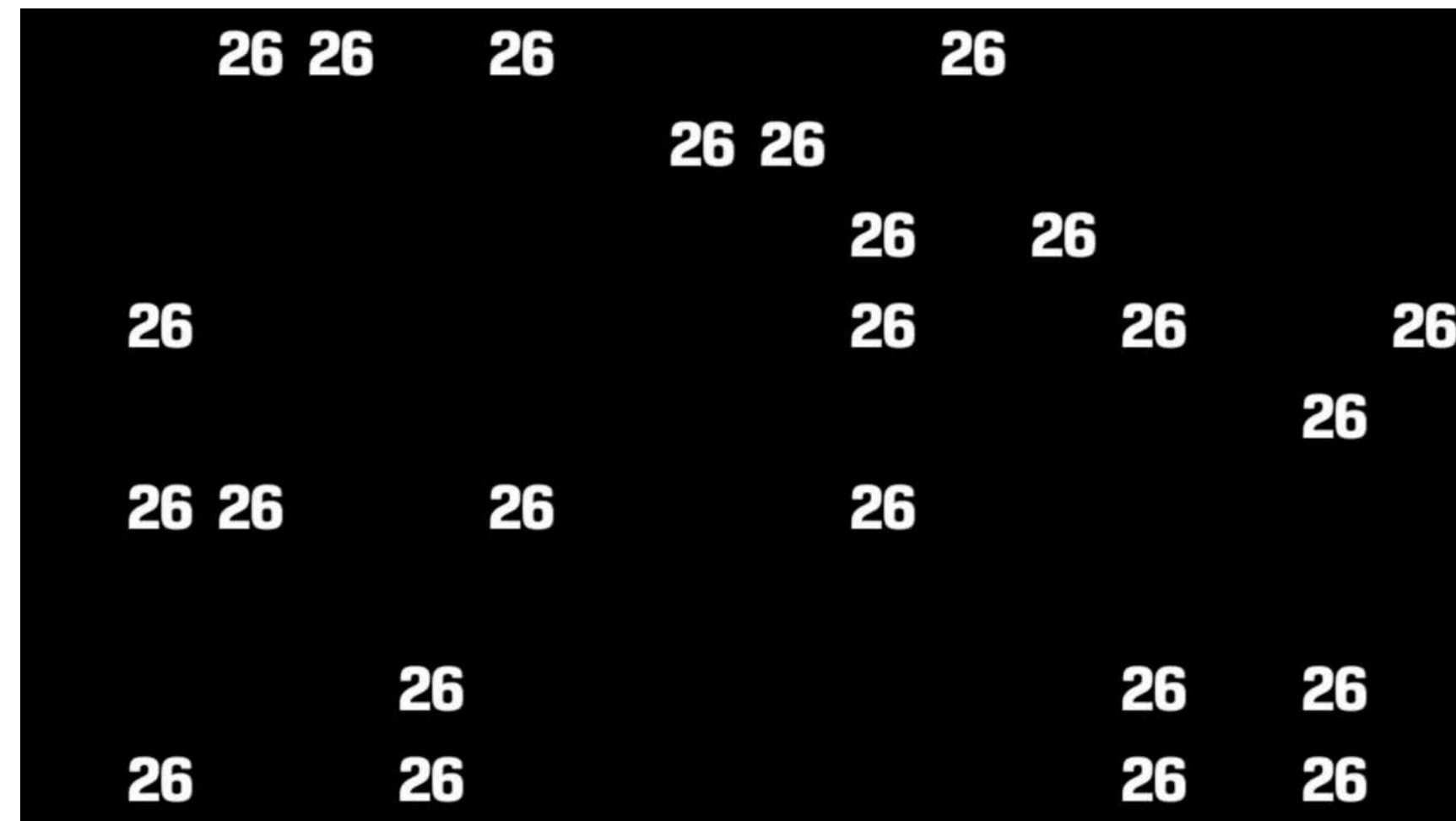
Existe uma técnica especial — **amostragem** — para recolher amostras, que garante, tanto quanto possível, o acaso na escolha.

Dessa forma, cada elemento da população passa a ter a mesma chance de ser escolhido, o que garante à amostra o caráter de representatividade, e isto é muito importante, pois, como vimos, nossas conclusões relativas à população vão estar baseadas nos resultados obtidos nas amostras dessa população.

# Amostragem casual ou aleatória simples

Este tipo de amostragem é equivalente a um sorteio lotérico.

Na prática, a amostragem casual ou aleatória simples pode ser realizada numerando-se a população de 1 a  $n$  e sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer,  $k$  números dessa seqüência, os quais corresponderão aos elementos pertencentes à amostra.



### **Exemplo:**

Vamos obter uma amostra representativa para a pesquisa da estatura noventa alunos de uma escola:

- a. Numeramos os alunos de 01 a 90.
- b. Escrevemos os números, de 01 a 90, em pedaços iguais de um mes papel, colocando-os dentro de uma caixa. Agitamos sempre a caixa p misturar bem os pedaços de papel e retiramos, um a um, nove números que formarão a amostra. Neste caso, 10% da população.

### **Exemplo:**

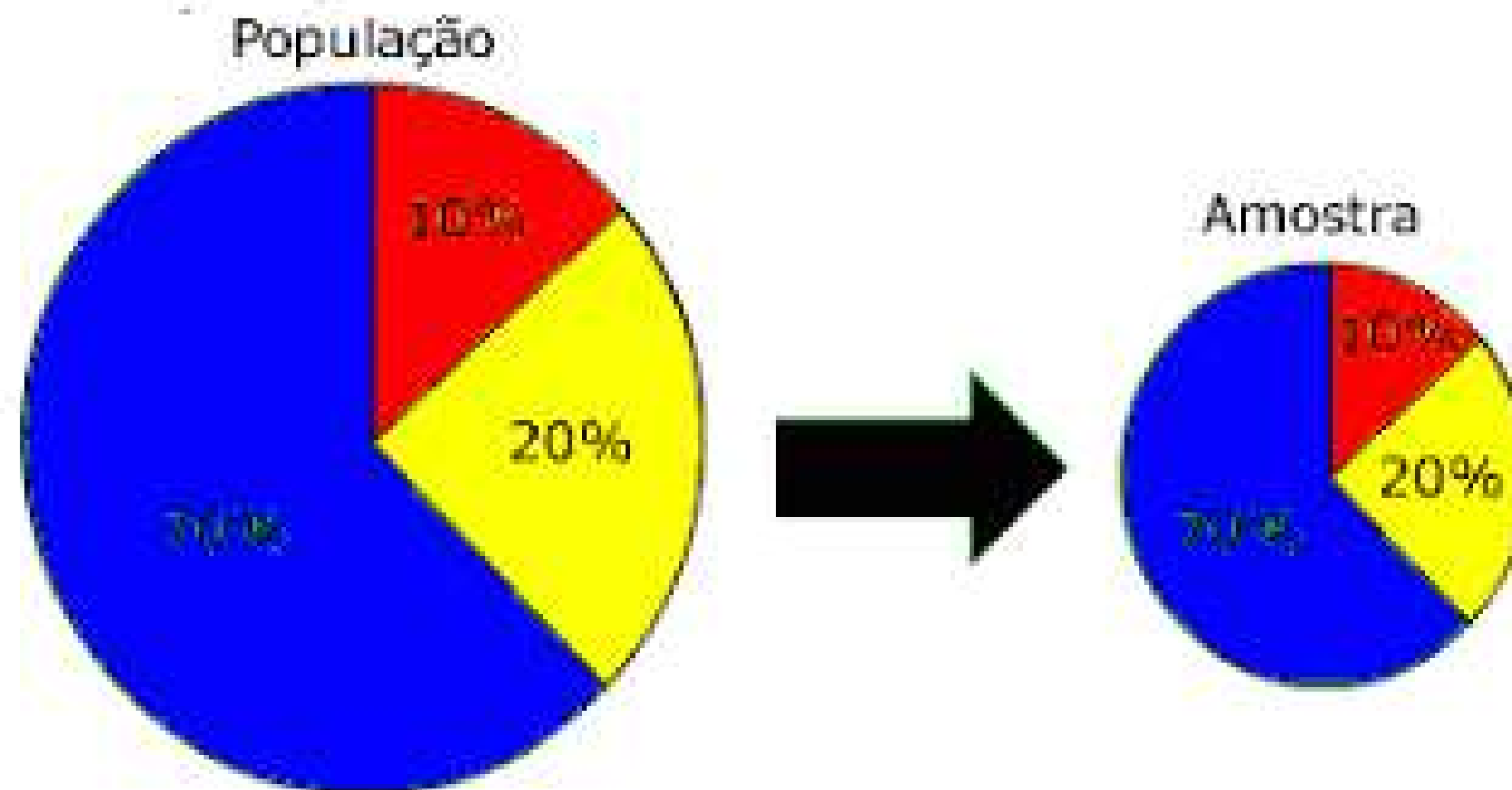
Vamos obter uma amostra representativa para a pesquisa da estatura noventa alunos de uma escola:

- a. Numeramos os alunos de 01 a 90.
- b. Escrevemos os números, de 01 a 90, em pedaços iguais de um mes papel, colocando-os dentro de uma caixa. Agitamos sempre a caixa p misturar bem os pedaços de papel e retiramos, um a um, nove números que formarão a amostra. Neste caso, 10% da população.

# Amostragem proporcional estratificada

Muitas vezes a população se divide em subpopulações — estratos.

Como é provável que a variável em estudo apresente, de estrato em estrato, um comportamento heterogêneo e, dentro de cada estrato, um comportamento homogêneo, convém que o sorteio dos elementos da amostra leve em consideração tais estratos.



## **Amostragem proporcional estratificada**

Muitas vezes a população se divide em subpopulações — estratos.

Como é provável que a variável em estudo apresente, de estrato em estrato, um comportamento heterogêneo e, dentro de cada estrato, um comportamento homogêneo, convém que o sorteio dos elementos da amostra leve em consideração tais estratos.

## **Amostragem proporcional estratificada**

Muitas vezes a população se divide em subpopulações — estratos.

Como é provável que a variável em estudo apresente, de estrato em estrato, um comportamento heterogêneo e, dentro de cada estrato, um comportamento homogêneo, convém que o sorteio dos elementos da amostra leve em consideração tais estratos.



É exatamente isso que fazemos quando empregamos a **amostragem proporcional estratificada**, que, além de considerar a existência dos estratos, obtém os elementos da amostra proporcional ao número de elementos dos mesmos.

### Exemplo:

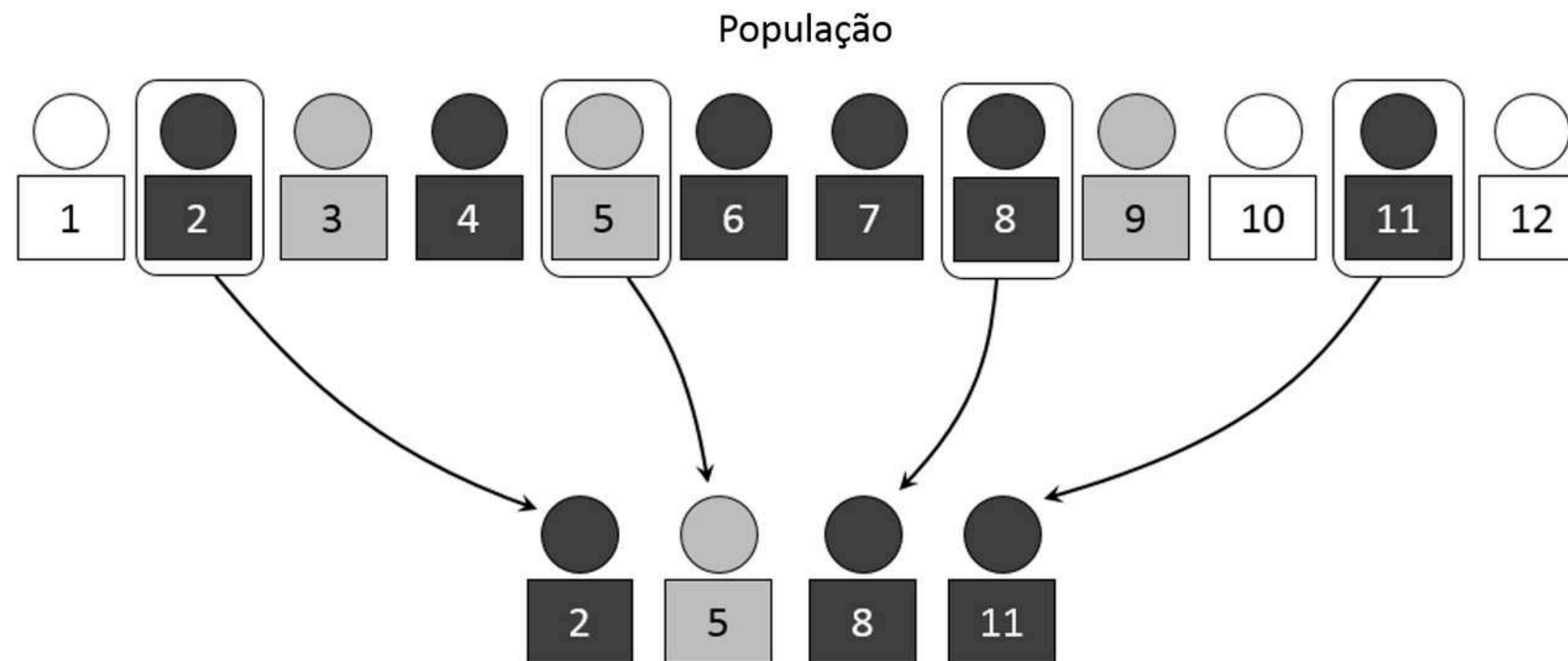
Supondo, no exemplo anterior, que, dos noventa alunos, 54 sejam meninos e 36 sejam meninas, vamos obter a amostra proporcional estratificada.

São, portanto, dois estratos (sexo masculino e sexo feminino) e queremos uma amostra de 10% da população. Logo, temos:

a.	SEXO	POPULAÇÃO	10%	AMOSTRA
	M	54	$\frac{10 \times 54}{100} = 5,4$	5
	F	36	$\frac{10 \times 36}{100} = 3,6$	4
	Total	90	$\frac{10 \times 90}{100} = 9,0$	9

# Amostragem sistemática

Quando os elementos da população já se acham ordenados, não há necessidade de construir o sistema de referência. São exemplos os prontuários médicos de um hospital, os prédios de uma rua, as linhas de produção etc. Nestes casos, a seleção dos elementos que constituirão a amostra pode ser feita por um sistema imposto pelo pesquisador. A esse tipo de amostragem denominamos **sistemática**.



Assim, no caso de uma linha de produção, podemos, a cada de produzidos, retirar um para pertencer a uma amostra da produção diária caso, estaríamos fixando o tamanho da amostra em 10% da população

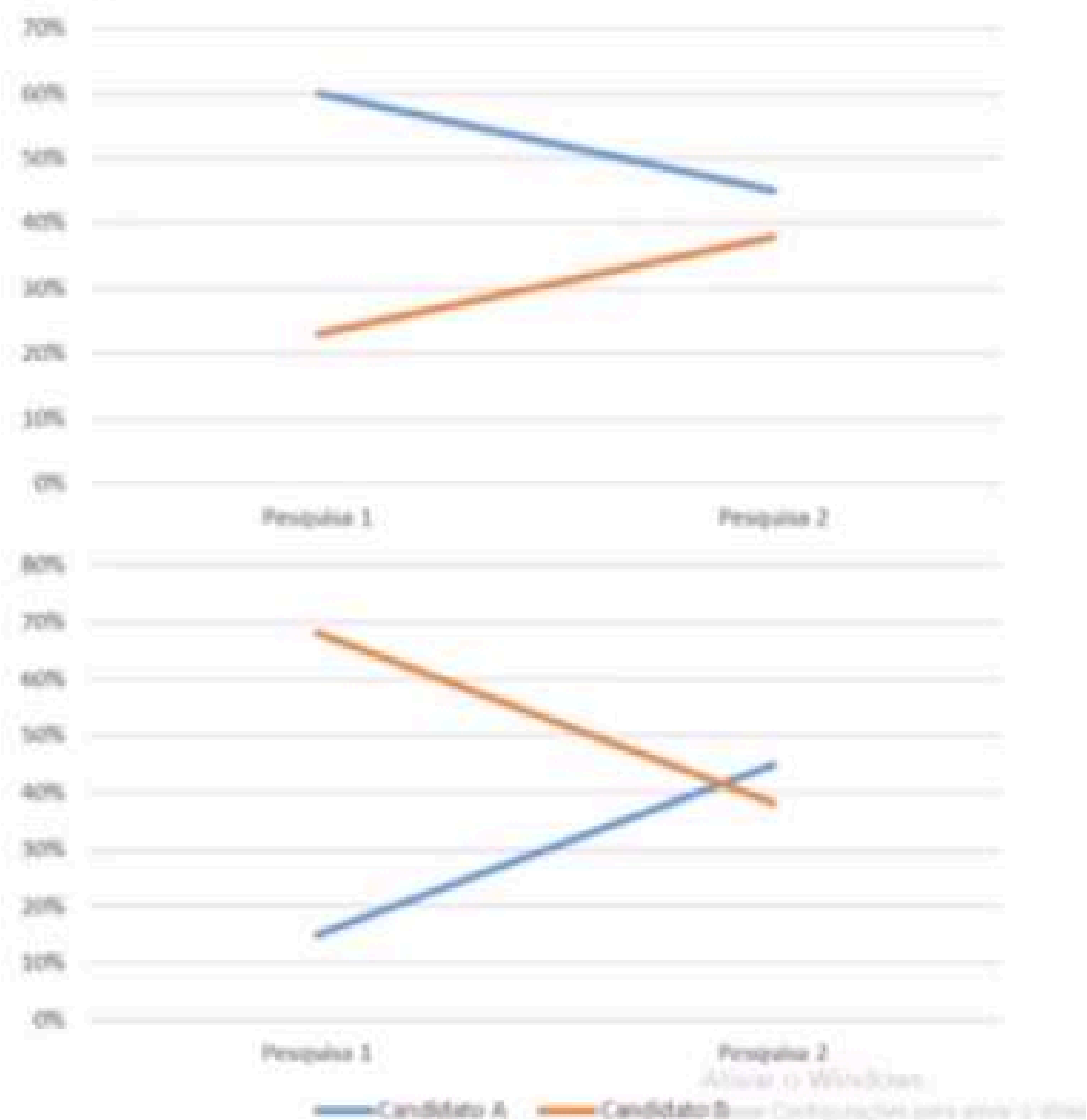
# Pesquisa de intenção de voto 2º turno

- Se as eleições fossem hoje...
  - Candidato A: 45%
  - Candidato B: 38%
- As eleições acontecerão em 15 dias.
- Quem vai ganhar com certeza?

# Pesquisa de intenção de voto 2º turno

- Candidato A: 45%
- Candidato B: 38%
- Brancos, nulos e indecisos: 17%
- Pesquisa de rejeição

- Tendência de crescimento
- Tendência de queda



# Margem de erro de 3%?

- Candidato A tem 45%, portanto pode ficar entre 42% e 48%
- Candidato B tem 38%, portanto pode ficar entre 35% e 41%
- Quem vai ganhar com certeza?

# Margem de erro de 3%?

- Candidato A tem 45%, portanto pode ficar entre 42% e 48%
- Candidato B tem 38%, portanto pode ficar entre 35% e 41%
- Quem vai ganhar com certeza?



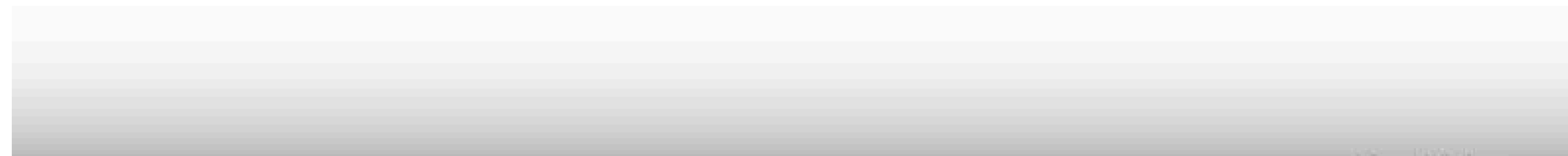
intervalo de  
confiança



Candidato A tem 45%, portanto pode ficar entre 42% e 48%, com 95% de chance

Candidato B tem 38%, portanto pode ficar entre 35% e 41%, com 95% de chance

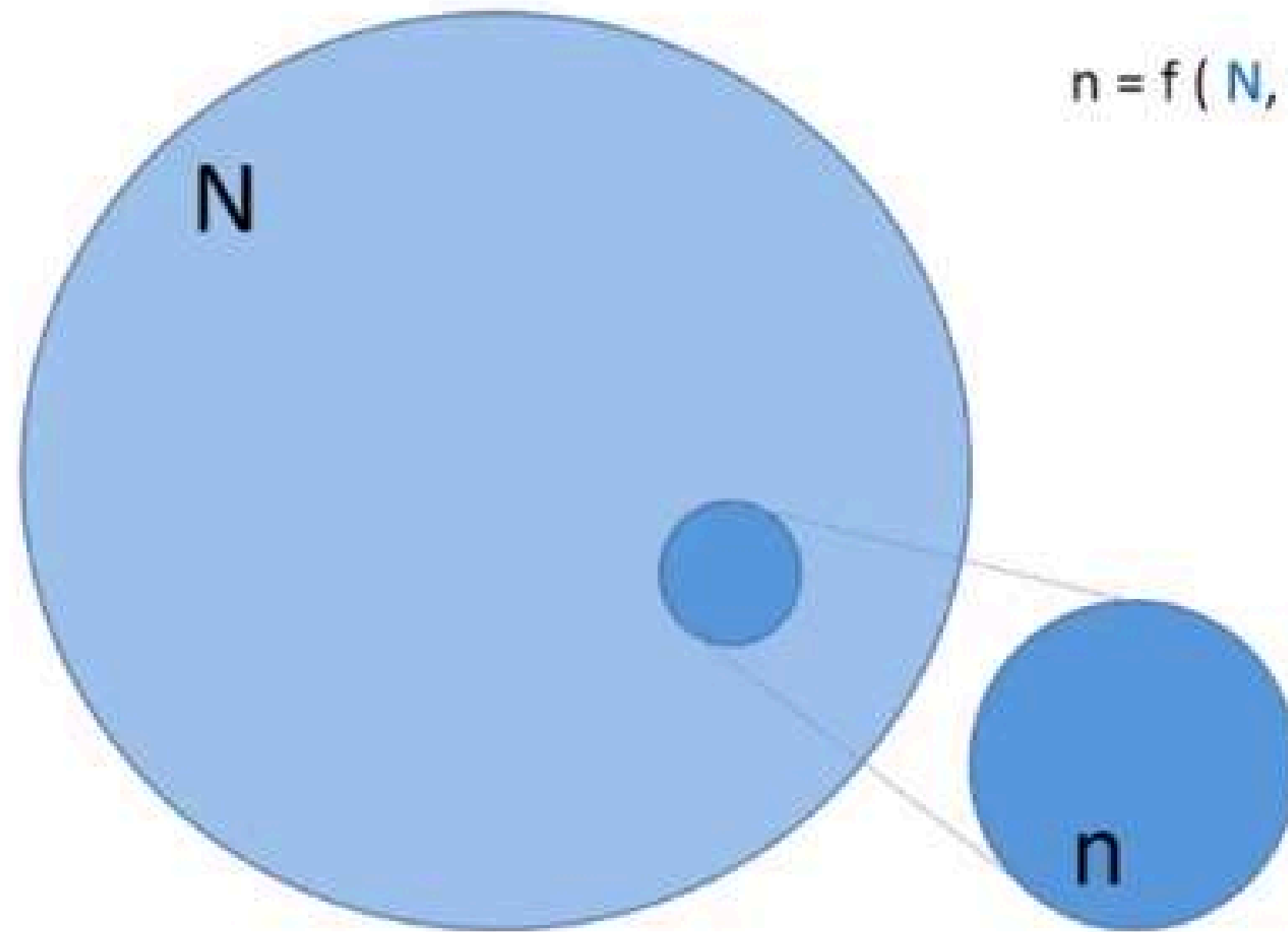
Ainda há uma chance de 5% do resultado ser diferente



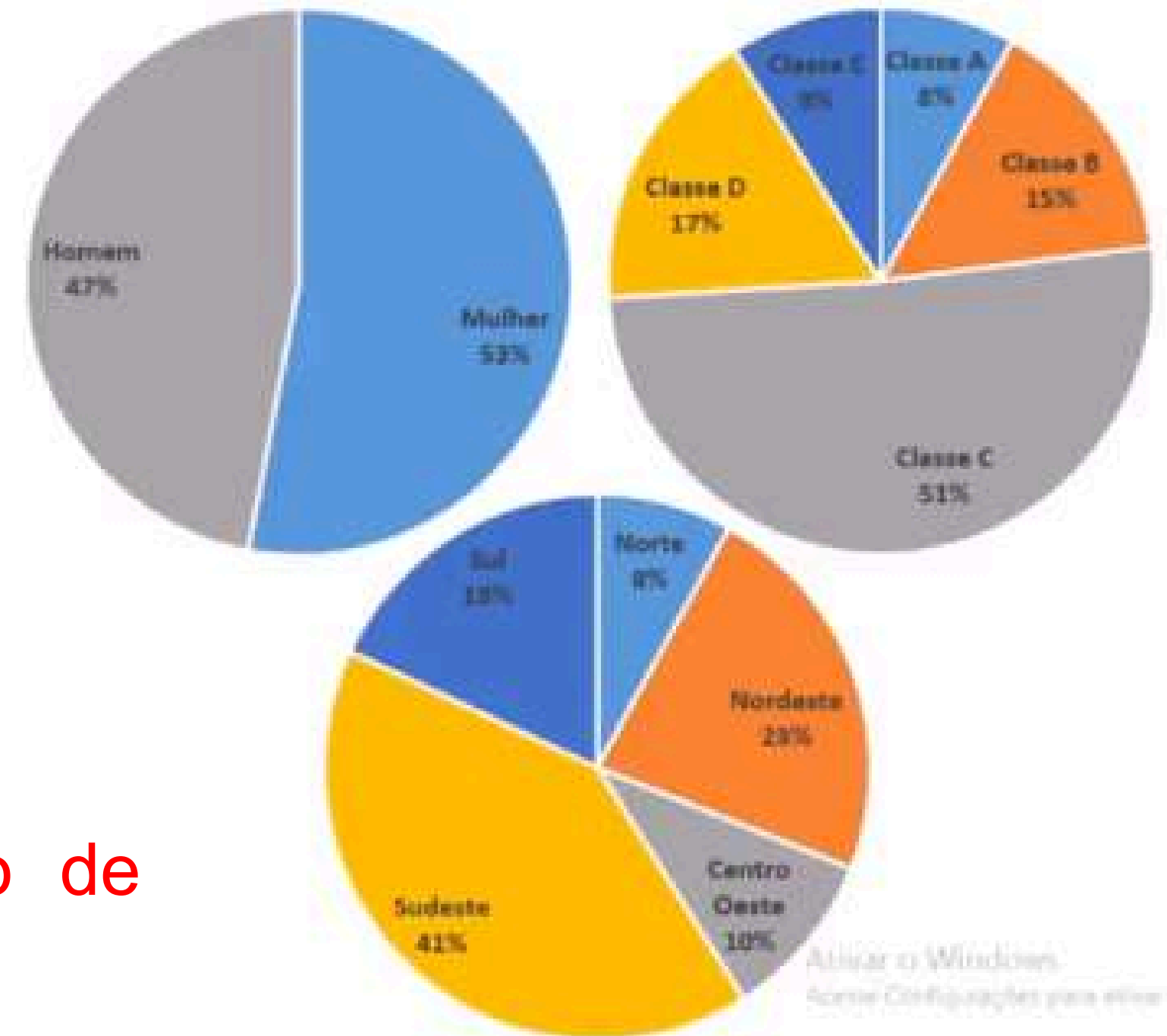
A margem de erro não esta sozinha, vem associada a um intervalo de confiança.

A chance de aquilo acontecer;

# De onde vem esse 95%?



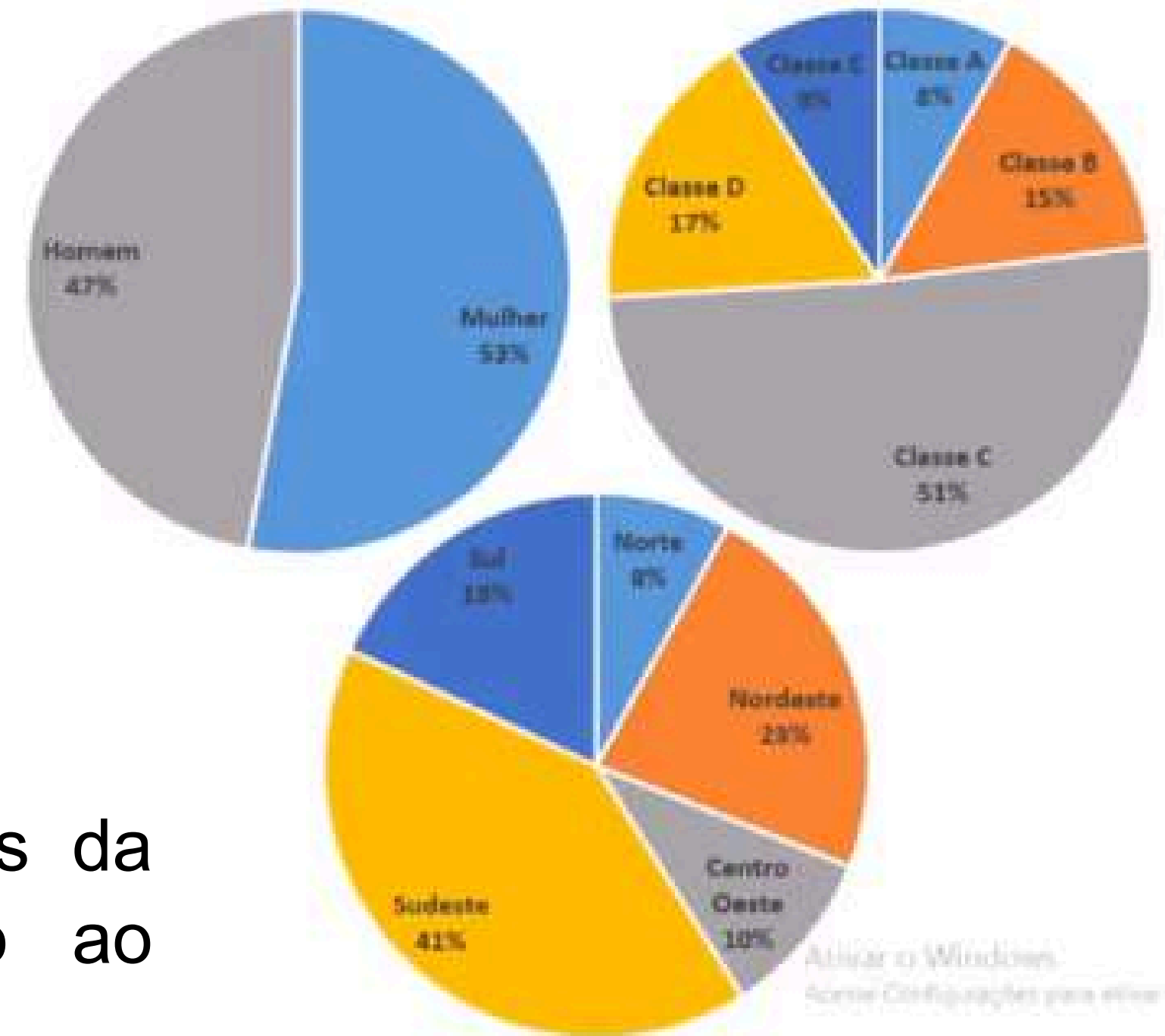
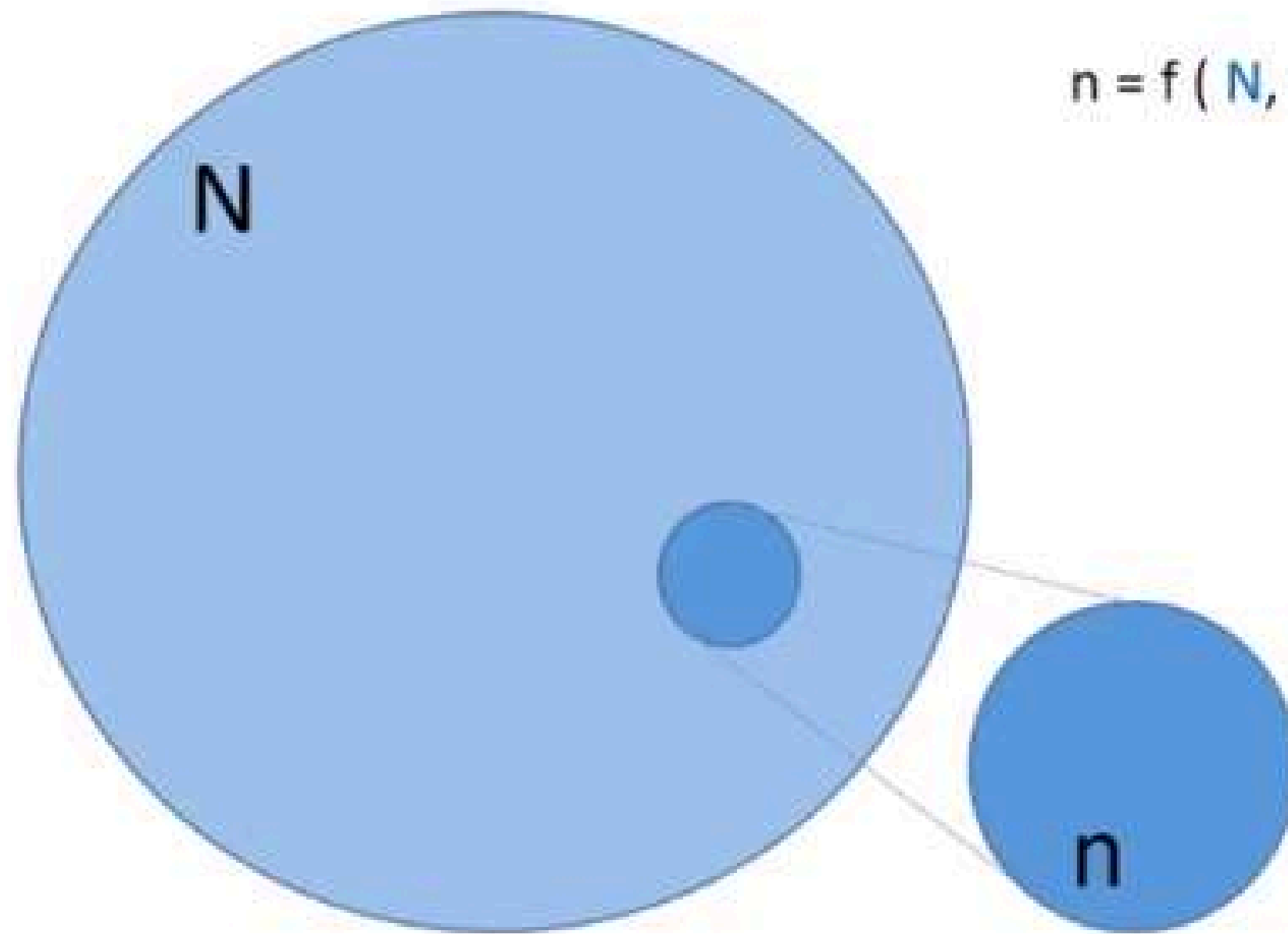
$n = f ( N, \text{variáveis significativas, margem de erro, confiança} )$



Tem a ver com a amostra, com o processo de amostragem;

# De onde vem esse 95%?

$n = f ( N, \text{variáveis significativas, margem de erro, confiança} )$



O intervalo de confiança é determinado antes da realização da pesquisa e está relacionado ao tamanho da amostra;

# Juntando margem de erro e confiança...

- Se as eleições fossem hoje...
  - Candidato A tem 45%, portanto pode ficar entre 42% e 48%, com 95% de chance
  - Candidato B tem 38%, portanto pode ficar entre 35% e 41%, com 95% de chance
- Em estatística inferencial, nunca há certeza!
- Quer dizer...