



By @kakashi_copiador

Aula 00 - Equipe Exatas

*CNU (Bloco 1 - Infraestrutura, Exatas e Engenharia) Conhecimentos Específicos
- Eixo Temático 5 - Geoprocessamento e
Análise de Dados - 2024 (Pós-Edital)*

Autor:

**Equipe Exatas Estratégia
Concursos, Monik Begname de
Castro, Alexandre Vastella**

15 de Janeiro de 2024

Índice

1) Introdução à estatística	3
2) Noções Iniciais sobre Estatística	4
3) Método Experimental x Método Estatístico	5
4) Dados Estatísticos	6
5) Variáveis Estatísticas	7
6) Séries Estatísticas	9
7) Distribuição de Frequências	12
8) Representação Gráfica das Distribuições de Frequências	17
9) Outros Gráficos e Representações	22
10) Questões Comentadas - Conceitos Iniciais - CESGRANRIO	28
11) Questões Comentadas - Variáveis Estatísticas - CESGRANRIO	29
12) Questões Comentadas - Representação Gráfica das Distribuições de Frequências - CESGRANRIO	30
13) Lista de Questões - Conceitos Iniciais - CESGRANRIO	31
14) Lista de Questões - Variáveis Estatísticas - CESGRANRIO	33
15) Lista de Questões - Representação Gráfica das Distribuições de Frequências - CESGRANRIO	35



INTRODUÇÃO À ESTATÍSTICA

A **Estatística** pode ser definida como a ciência que estuda os processos de **coleta, organização, análise e interpretação** de dados numéricos variáveis referentes a qualquer fenômeno. Podemos conceituá-la como **um conjunto de técnicas de coleta, organização, análise e interpretação de dados, aplicáveis a várias áreas do conhecimento, que auxiliam no processo de tomada de decisão.**

A Estatística se divide em três grandes ramos:

- **Estatística Descritiva (ou Dedutiva): responsável pela coleta, organização, descrição e resumo dos dados observados.** A partir de um determinado conjunto de dados, a Estatística Descritiva busca organizá-los em tabelas (ou gráficos) e estabelecer um sumário por meio de medidas descritivas como a média, os valores mínimo e máximo, o desvio padrão, entre outras.
- **Estatística Probabilística: responsável por estabelecer o modelo matemático probabilístico adotado para explicar os fenômenos aleatórios investigados pela Estatística.** Os resultados desses fenômenos aleatórios podem variar de uma observação para outra, o que dificulta muito a previsão de um resultado futuro. Por isso, a Teoria da Probabilidade é usada para medir a chance de ocorrência de determinados eventos.
- **Estatística Inferencial (ou Indutiva): responsável pela análise e interpretação dos dados.** A partir da análise de dados de uma amostra, a Estatística Indutiva estabelece inferências e previsões sobre a população, auxiliando na tomada de decisões. Além disso, busca generalizar conclusões a respeito da população a partir de uma amostra, analisando a representatividade, a significância e a confiabilidade dos resultados obtidos.



ESTATÍSTICA DESCRITIVA

É responsável pela coleta, organização, descrição e resumo dos dados observados.

ESTATÍSTICA PROBABILÍSTICA

É responsável por estabelecer o modelo matemático adotado para explicar fenômenos aleatórios.

ESTATÍSTICA INFERENCIAL

É responsável pela análise e interpretação dos dados.



CONCEITOS INICIAIS

Alguns conceitos iniciais da estatística costumam ser abordados em provas de concursos públicos, dentre os quais podemos citar:

- **População:** conjunto que contém **todos os indivíduos, objetos ou elementos** a serem estudados, que apresentam uma ou mais características em comum. A população pode ser **finita**, quando apresenta um número pequeno ou limitado de observações; ou **infinita**, quando apresenta um número muito grande ou ilimitado de observações.
- **Amostra:** subconjunto extraído da população para análise, devendo ser **representativo** daquele grupo. A partir das informações colhidas da amostra, os resultados obtidos podem ser utilizados para generalizar, inferir ou tirar conclusões acerca da população.
- **Censo (ou recenseamento):** estudo dos dados relativos a todos os elementos de uma população. O censo pode custar muito caro e demandar um tempo considerável, de forma que um estudo considerando apenas uma parcela da população pode ser uma alternativa mais simples, rápida e menos onerosa.
- **Amostragem:** processo que consiste na **seleção criteriosa dos elementos** a serem submetidos à investigação. Se forem cometidos erros no processo de seleção da amostra, muito provavelmente, o estudo ficará comprometido e os resultados serão tendenciosos. Portanto, devemos garantir que **a amostra seja representativa da população**.
- **Parâmetros:** descrições numéricas de características populacionais que raramente são conhecidas. Em geral, é muito caro ou demorado obter os dados da população inteira. Assim, **algumas medidas precisam ser estimadas a partir de critérios ou métodos definidos pelo pesquisador**, para representar características desconhecidas de uma população. Normalmente, os parâmetros populacionais são constantes para uma população.
- **Estatística (ou estimador):** medidas numéricas obtidas de amostras representativas extraídas da população. A partir das informações colhidas da amostra, **as estatísticas amostrais podem ser utilizadas para inferir ou tirar conclusões acerca dos parâmetros populacionais**, como a proporção de homens e mulheres na população brasileira.



MÉTODO EXPERIMENTAL X MÉTODO ESTATÍSTICO

Para a investigação de um fenômeno, temos a nossa disposição dois métodos:

- **Método Experimental:** consiste em **manter constantes as causas (fatores)**, com **exceção de uma, que é variada para que seus efeitos sejam descobertos**. Contudo, nem sempre poderemos aplicar o método experimental, pois os fatores que afetam um fenômeno podem não permanecer constantes enquanto variamos a causa que nos interessa. Por exemplo, para analisarmos a queda nas vendas de uma empresa que produz chocolates, teríamos que considerar vários fatores que não necessariamente permanecerão constantes durante toda a investigação do fenômeno, tais como o fluxo de turistas na localidade; a temperatura média; o preço do concorrente; etc.
- **Método Estatístico:** **admite e registra todas as possíveis variações das causas presentes, procurando determinar a influência de cada fator no resultado final**. Dessa forma, o método estatístico busca descobrir relações entre os fatores, como, por exemplo, a influência da temperatura média e do fluxo de turistas na venda de chocolates.



ESQUEMATIZANDO

MÉTODO EXPERIMENTAL

As **CAUSAS** são mantidas **CONSTANTES**, **COM EXCEÇÃO DE UMA**, que é **VARIADA** para que seus efeitos sejam descobertos.

MÉTODO ESTATÍSTICO

Admite e **REGISTRA TODAS AS POSSÍVEIS VARIAÇÕES DAS CAUSAS PRESENTES**, procurando determinar a influência de cada fator no resultado.



DADOS ESTATÍSTICOS

Os **dados estatísticos** constituem os valores resultantes do processo de **coleta de dados**. Os dados referem-se a um **conjunto de valores**, observações de um **fenômeno de interesse**, que denominamos de **variável**. Eles são organizados por meio de **variáveis** (a característica está sendo medida) e **observações** (elementos da amostra/população).

Uma **variável** é um **atributo** cujo valor pode variar de uma unidade de investigação para outra. Por exemplo, a unidade de investigação pode ser um morador de uma determinada cidade e a variável a sua altura. As **observações** são os **valores assumidos por uma variável** em uma das unidades investigadas.

Com relação ao número de observações coletadas, os dados são classificados em **univariados**, **bivariados** ou **multivariados**:

- a) **dados univariados**: quando uma única observação de cada unidade de investigação é registrada. Por exemplo: peso;
- b) **dados bivariados**: quando duas observações de cada unidade de investigação são registradas. Por exemplo: peso e altura;
- c) **dados multivariados**: quando mais de duas observações acerca de cada unidade de investigação são registradas. Por exemplo: peso, altura, sexo e idade.

Quanto à forma de apresentação, os dados podem ser classificados em **dados brutos** ou **rol**.

Dados Brutos

Os **dados brutos** são aqueles que **não foram numericamente organizados em ordem crescente ou decrescente**, ou seja, **estão na forma como foram coletados**. A tabela na qual os elementos não aparecem numericamente ordenados é denominada de **tabela primitiva**. Em geral, ela **oferece pouca ou nenhuma informação ao leitor**, sendo necessária uma organização dos dados, a fim de torná-los mais expressivos.

Rol

O **rol** é a organização dos dados brutos em **ordem de grandeza crescente ou decrescente**. Com os dados organizados em rol, podemos saber, com facilidade, qual o **menor** e o **maior** elemento de um conjunto de dados.



VARIÁVEIS ESTATÍSTICAS

A **variável estatística** consiste no **conjunto de características** que desejamos averiguar **estatisticamente**, podendo ser definida como o **objeto da pesquisa estatística**. As variáveis estatísticas podem ser classificadas em duas categorias: **qualitativas** e **quantitativas**.

Variáveis Qualitativas

As **variáveis qualitativas (ou categóricas)** são as características que **não podem ser descritas de forma numérica**, mas que podem ser definidas por meio de **qualidades (atributos ou categorias)** do indivíduo pesquisado. Elas podem ser classificadas em **nominais** ou **ordinais**:

- a) **variável qualitativa nominal**, as possíveis categorias **não podem** ser ordenadas. Por exemplo, a cor dos olhos dos moradores de uma determinada cidade (pretos, castanhos, azuis e verdes);
- b) **variável qualitativa ordinal**, as possíveis categorias **podem** ser ordenadas de alguma forma. Por exemplo, o grau de instrução dos funcionários de um determinado órgão (fundamental, médio, superior).

Variáveis Quantitativas

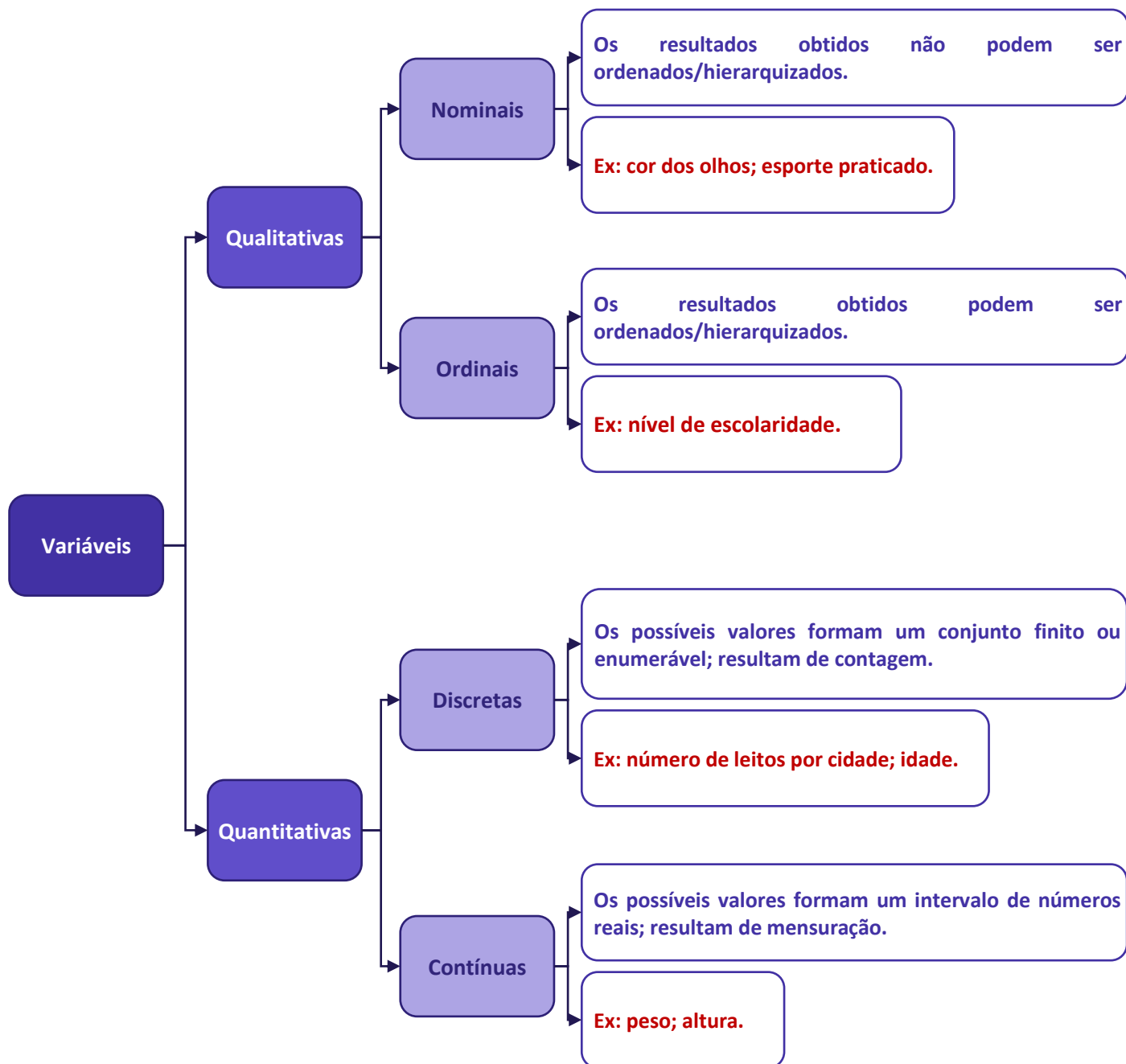
As **variáveis quantitativas** são características que podem ser descritas em termos de **quantidades (valores numéricos)**, obtidas por meio de contagem ou mensuração. Elas podem ser classificadas em **discretas** e **contínuas**:

- a) **variáveis quantitativas discretas**, os possíveis valores formam um **conjunto finito ou enumerável de números** e, geralmente, **resultam de um processo de contagem**. O número de ocorrências da característica em análise pode ser **contado**. Por exemplo, número de leitos disponíveis em um município.
- b) **variáveis quantitativas contínuas**, os possíveis valores formam um **intervalo de números reais** e, normalmente, **resultam de um processo de mensuração**. A característica pode ser medida em uma **escala contínua**, a qual podem ser associados um número infinito de possíveis valores. Por exemplo, altura dos habitantes de uma cidade.





RESUMINDO



SÉRIES ESTATÍSTICAS

Uma **série estatística** consiste em um conjunto de dados **organizado** com base em uma **característica comum**, ou seja, **uma mesma variável**. Ela normalmente é representada por meio de **tabelas** ou **gráficos**, conforme ficar melhor representada, a fim de sintetizar os dados estatísticos observados e torná-los mais compreensíveis.

Uma **tabela** é um **quadro** que **resume** um **conjunto de observações**, sendo composta de:

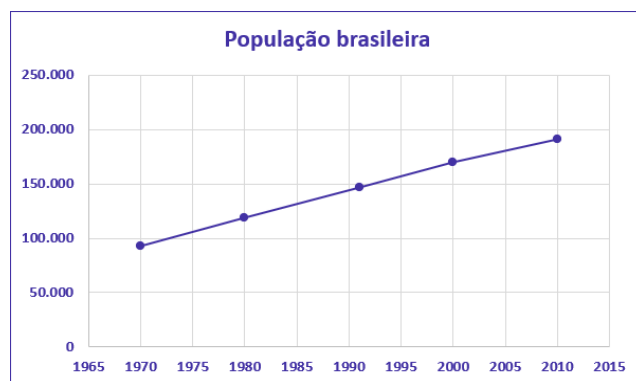
- a) corpo – conjunto de linhas e colunas com as informações sobre a variável em estudo;
- b) cabeçalho – parte superior que especifica o conteúdo das colunas;
- c) coluna indicadora – parte que indica o conteúdo das linhas;
- d) linhas – traços que facilitam a leitura dos dados;
- e) célula – espaço onde os dados são armazenados;
- f) título – identificação da tabela, contendo as informações sobre seu conteúdo;
- g) fonte – referência de onde os dados foram obtidos, localizada no rodapé.

Diagrama de uma tabela com legendas explicativas:

- Título:** População brasileira no período de 1970 a 2010 (x1000)
- Cabeçalho:** Anos, População
- Coluna indicadora:** 1970, 1980, 1991, 2000, 2010
- Corpo:** 93.134, 119.011, 146.825, 169.799, 190.755
- Célula:** 2000
- Fonte:** Fonte: Censo Demográfico (2010)
- Coluna numérica:** 93.134, 119.011, 146.825, 169.799, 190.755
- Linhas:** Linhas de dados

Anos	População
1970	93.134
1980	119.011
1991	146.825
2000	169.799
2010	190.755

Um **gráfico** é uma **forma clara e objetiva** de apresentar uma **série estatística**. Seu objetivo é **proporcionar uma compreensão mais rápida do fenômeno em estudo**. O gráfico deve ser destituído de detalhes sem importância (**ser simples**); permitir a correta interpretação dos valores representativos do fenômeno (**ser claro**); e transmitir a verdade sobre o fenômeno (**ser verossímil**). A tabela anterior pode ser representada por meio do gráfico ao lado.



Finalmente, podemos verificar a presença de **três elementos** nas séries estatística: o **tempo**, o **espaço** e a **espécie**. Conforme os elementos variem, a série pode ser classificada em três categorias: **temporal** (ou cronológica), **geográfica** (ou territorial) e **específica**.



Séries Temporais (ou Cronológicas)

É a série cujos dados são **dispostos segundo a época de ocorrência**. Enquanto o **tempo varia**, o **fato e o local permanecem constantes**. Também são chamadas de séries **históricas** ou **evolutivas**. A principal característica é o **fator cronológico variável**. Ao lado temos a série histórica da população residente no Brasil no período de 1970 a 2010, com frequência decenal.

População brasileira no período de 1970 a 2010 (x1000)

Anos	População
1970	93.134
1980	119.011
1991	146.825
2000	169.799
2010	190.755

Fonte: Censo Demográfico (2010)

Séries Geográficas (ou Territoriais)

É a série cujos dados são **dispostos segundo a localidade de ocorrência**. Enquanto o **local varia**, o **fato e o tempo permanecem constantes**. Também são chamadas de séries **espaciais** ou de **localização**. A principal característica é o **fator geográfico variável**. Ao lado temos a série geográfica da população urbana residente em cada uma das regiões brasileiras no ano de 2010.

População Urbana em 2010 (x1000)

Região	População
Norte	11.664
Nordeste	38.821
Sudeste	74.696
Sul	23.260
Centro-Oeste	12.482

Fonte: Censo Demográfico (2010)

Séries Específicas

É a série cujos dados são **dispostos segundo a modalidade de ocorrência**. Enquanto o **fato varia**, a **época e o local permanecem constantes**. Também são chamadas de séries **categóricas**. A principal característica é o **fator especificativo variável**. Ao lado temos uma série específica das populações urbana e rural residentes no Brasil no ano de 2010.

População Urbana e Rural em 2010 (x1000)

Zona	População
Urbana	93.134
Rural	119.011
Total	190.755

Fonte: Censo Demográfico (2010)



Séries Mistas (ou Compostas)

Muitas vezes, podemos ter a necessidade de apresentar, em uma única tabela, a variação de valores de mais de uma variável, isto é, combinar duas ou mais séries. As séries resultantes desse processo de combinação são chamadas de **séries mistas (ou compostas)** e apresentadas por meio de **tabelas de dupla entrada**.

O nome da nova série deve levar em consideração pelo menos dois elementos. Assim, **se for uma série mista de fato e tempo, denominaremos de série específico-temporal**. Ao lado temos uma série **específico-temporal** representando as populações de homens e mulheres residentes no Brasil, no período de 1970 a 2010, com variação decenal.

População do Brasil por Sexo de 1970 a 2010 (x1000)

Anos	Sexo	
	Homens	Mulheres
1970	46.327	46.807
1980	59.142	59.868
1991	72.485	74.340
2000	83.602	86.270
2010	93.406	97.348

Fonte: Censo Demográfico (2010)

Por sua vez, **se tivermos uma série mista de local e tempo, denominaremos de série geográfica-temporal**. Ao lado temos uma **série geográfico-temporal** representando as populações residentes em cada região brasileira, no período de 1970 a 2010, com variação decenal.

População do Brasil por Região de 1970 a 2010 (x1000)

Anos	Regiões				
	N	NE	SE	S	CO
1970	3.603	28.111	39.850	16.496	5.072
1980	5.880	34.815	51.737	19.031	7.545
1991	10.030	42.497	62.740	22.129	9.427
2000	12.900	47.741	72.412	25.107	11.636
2010	15.864	53.081	80.364	27.386	14.058

Fonte: Censo Demográfico (2010)

Por fim, devemos notar que podem existir séries compostas de três ou mais entradas, embora isso raramente aconteça, por conta da dificuldade de representação.



DISTRIBUIÇÃO DE FREQUÊNCIAS

Logo após a coleta de dados, temos o que chamamos de **dados brutos**. Os **dados brutos** fornecem pouca informação ao leitor, sendo necessário organizá-los. A simples organização dos dados em um **rol crescente** já ajuda bastante nesse sentido. Por exemplo, com os dados organizados em rol, facilmente identificamos os valores **mínimo** e **máximo** do conjunto de dados.

Rol Crescente				
85	115	129	143	161
89	115	129	143	165
96	123	134	148	168
98	123	135	153	170
99	124	135	154	171
103	126	135	155	171
104	126	137	157	171
105	126	137	158	173
113	127	137	159	175
114	128	142	161	175

Outra informação que conseguimos extrair dos dados organizados em rol crescente é que alguns tempos, como 126 min, 135 min, 137 min e 171 min, foram mais frequentes, ou seja, apareceram mais vezes durante a pesquisa.

Uma maneira mais concisa de mostrar os dados do rol é apresentar cada valor juntamente com o número de ocorrências (frequência), em vez de repeti-los. **A tabela que contém todos os valores com suas respectivas frequências é denominada de distribuição de frequências.**

Uma **distribuição de frequências** também **pode ser definida como uma série estatística na qual permanecem constantes o fato, o local e a época**. Ela pode ser classificada em dois tipos: **pontual (ou discreta)** e **intervalar (ou contínua)**.

Na **distribuição de frequências pontual**, são apresentados todos os dados coletados juntamente com suas respectivas frequências, **não havendo perda de valores**. Esse processo pode exigir muito espaço, especialmente quando o número de valores da variável tende a aumentar.

Tempo (min)	Freq.	Tempo (min)	Freq.	Tempo (min)	Freq.	Tempo (min)	Freq.
85	1	114	1	135	3	158	1
89	1	115	2	137	3	159	1
96	1	123	2	142	1	161	2
98	1	124	1	143	2	165	1
99	1	126	3	148	1	168	1



103	1	127	1	153	1	170	1
104	1	128	1	154	1	171	3
105	1	129	2	155	1	173	1
113	1	134	1	157	1	175	2

Quando a variável é contínua, o mais recomendável é agrupar os valores por **intervalos de classe**. Em vez de listar cada um dos valores, utilizamos uma **distribuição de frequências intervalar**, apresentando os **intervalos de classe** e as **frequências correspondentes**. Dessa forma, **perdemos a informação detalhada dos tempos médios**, mas ganhamos em termos de **praticidade**, simplificando o processo de análise de dados:

Tempo médio (X_i)	Frequência (f_i)
$85 \leq x < 100$	5
$100 \leq x < 115$	5
$115 \leq x < 130$	12
$130 \leq x < 145$	10
$145 \leq x < 160$	7
$160 \leq x < 175$	9
$175 \leq x < 190$	2

Para identificar uma classe, temos que conhecer os valores dos **limites inferior e superior da classe**, que delimitam um **intervalo de classe**. Desse modo, precisamos definir a natureza do intervalo de classe, se aberto ou fechado. Portanto, temos as seguintes notações para os diferentes tipos de intervalos:



Tipo de Intervalo	Notação matemática	Notação estatística	Significado
Intervalo aberto	$a < x < b$	$a - b$	Engloba todos os elementos entre a e b , mas não engloba a nem b .
Intervalo fechado à esquerda e aberto à direita	$a \leq x < b$	$a \vdash b$	Engloba todos os elementos entre a e b , inclusive a mas não b .
Intervalo aberto à esquerda e fechado à direita	$a < x \leq b$	$a \dashv b$	Engloba todos os elementos entre a e b , inclusive b mas não a .
Intervalo fechado	$a \leq x \leq b$	$a \dashv b$	Engloba todos os elementos entre a e b , inclusive a e b .

Em análises estatísticas, é comum encontramos **distribuições de frequências intervalares**, pois o **objetivo da estatística é justamente fazer um apanhado geral das características de um conjunto de dados, sem adentrar em detalhes de casos particulares**.



Elementos de uma Distribuição de Frequências

Agora, analisaremos cada elemento de uma distribuição de frequências, tomando como referência a tabela apresentada anteriormente:

Tempo médio (X_i)	Frequência (f_i)
$85 \leq x < 100$	5
$100 \leq x < 115$	5
$115 \leq x < 130$	12
$130 \leq x < 145$	10
$145 \leq x < 160$	7
$160 \leq x < 175$	9
$175 \leq x < 190$	2

Classe

As **classes** são os **intervalos** nos quais o fenômeno é subdividido. Podemos dizer que as classes são os intervalos ou subdivisões dos elementos que compõem um conjunto de dados. Na tabela anterior, a primeira classe é representada pelo intervalo $85 \leq x < 100$; a segunda, pelo intervalo $100 \leq x < 115$, e assim sucessivamente.

Existem duas maneiras de determinar o número "ideal" de classes, k , em função do número de dados da tabela, n . A primeira consiste em utilizar a fórmula de Sturges: $k = 1 + 3,3 \times \log n$. A outra, utilizada quando o número de dados é menor ou igual a 50, é por meio da fórmula: $k = \sqrt{n}$.

Limite de Classe

Cada classe tem um **limite inferior de classe** (l_{inf}), que é o menor número que pode pertencer à classe, e um **limite superior de classe** (l_{sup}), que é o maior número que pode pertencer à classe. Os **limites de uma classe são seus valores extremos**. Por exemplo, o limite inferior da primeira classe é 85, enquanto o limite superior é 100.

Amplitude de um Intervalo de Classe

A **amplitude de um intervalo de classe**, ou simplesmente **intervalo de classe**, é a **diferença entre os limites inferiores (ou superiores) de classes consecutivas**: $h = l_{sup} - l_{inf}$, em que l_{inf} é o limite inferior do intervalo de classe e l_{sup} é o limite superior do intervalo de classe. Embora desejável, a **amplitude do intervalo de classe nem sempre será constante ao longo de toda a distribuição de frequências intervalar**. Para exemplificar, a amplitude da primeira classe é $100 - 85 = 15$.



Amplitude Total

A **amplitude total** é a diferença entre o **limite superior da última classe** (limite superior máximo) e o **limite inferior da primeira classe** (limite inferior mínimo): $AT = l_{\max} - l_{\min}$. Quando todas as classes possuem a mesma amplitude, também podemos determinar o valor da amplitude total multiplicando o valor do intervalo de classe (h) pela quantidade de classes da distribuição (k): $AT = h \times k$. Em nosso exemplo, a amplitude total é calculada da seguinte maneira: $AT = l_{\max} - l_{\min} = 190 - 85 = 105$.

Ponto médio de classe

O **ponto médio** é a **média aritmética** simples dos **valores extremos de uma classe**, ou seja, a soma dos limites inferior e superior dividida por dois: $PM = (l_{\inf} + l_{\sup})/2$. Esse ponto também costuma ser chamado de **marca** ou **representante da classe**. Para praticar, vamos calcular os pontos médios de nossa distribuição de frequências: $PM = (l_{\inf} + l_{\sup})/2 = (85 + 100)/2 = 92,5$.

Frequência

Ao longo dessa aula, em várias oportunidades abordamos conceitos relacionados à **frequência**, isto é, ao **número de ocorrências de um determinado valor ou de uma certa classe**. Esse conceito é de grande relevância para a estatística descritiva e deve ser estudado de forma mais aprofundada. Existem quatro tipos de frequência, os quais serão analisados nas subseções seguintes:

- a) frequência absoluta simples (f_i);
- b) frequência absoluta acumulada (f_{ac});
- c) frequência relativa simples (F_i);
- d) frequência relativa acumulada (F_{ac}).

Frequência Absoluta Simples

A **frequência absoluta simples** corresponde ao número de observações correspondentes a uma determinada classe ou a um determinado valor. A frequência simples é simbolizada por f_i . Para a tabela em análise, temos que a frequência absoluta da primeira classe é $f_1 = 5$; da segunda classe é $f_2 = 5$.

A soma de todas as frequências corresponde ao número total de dados analisados: $\sum_{i=1}^k f_i = n$, sendo que a notação $\sum_{i=1}^k f_i$ representa o somatório das frequências de cada uma das k classes. Para o exemplo, temos:

$$\sum_{i=1}^7 f_i = f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 = 5 + 5 + 12 + 10 + 7 + 9 + 2 = 50$$



Frequência Absoluta Acumulada

A **frequência absoluta acumulada crescente** (f_{ac}) é a soma das frequências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe: $f_{ac_i} = f_1 + f_2 + f_3 + \dots + f_i$. No exemplo anterior, a frequência acumulada correspondente à quarta classe é: $f_{ac_4} = f_1 + f_2 + f_3 + f_4 = 5 + 5 + 12 + 10 = 32$, significando que 32 alunos estudam por um período igual ou superior a 85 minutos e inferior a 145 minutos (limite superior da quarta classe).

Frequência Relativa Simples

A **frequência relativa simples** corresponde à proporção de dados existentes em uma determinada classe. Para calcular a frequência relativa de uma classe, dividimos a frequência absoluta simples f_i pela frequência total (isto é, dividimos a parte pelo todo): $F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$. Em nosso exemplo, a frequência relativa da primeira classe é $F_1 = \frac{5}{50} = 0,10$; e da segunda classe, $F_2 = \frac{5}{50} = 0,10$. A soma de todas as frequências relativas deve ser igual a 100%:

Frequência Relativa Acumulada

A **frequência relativa acumulada crescente** (F_{ac}) é a proporção de valores inferiores ao limite superior do intervalo de uma dada classe: $F_{ac_i} = F_1 + F_2 + F_3 + \dots + F_i$. No exemplo anterior, a frequência acumulada correspondente à quarta classe é: $F_{ac_4} = F_1 + F_2 + F_3 + F_4 = 10\% + 10\% + 24\% + 20\% = 64\%$, significando que 64% dos alunos estudam por um período igual ou superior a 85 minutos e inferior a 145 minutos (limite superior da quarta classe).

Densidade de Frequência

A densidade de frequência de uma classe consiste no quociente entre a frequência da classe (absoluta ou relativa) e sua amplitude: $d_i = \frac{f_i}{h_i}$. Para a primeira classe do nosso exemplo, a densidade de frequência é $d_1 = \frac{f_1}{h_1} = \frac{5}{15} = 0,33$

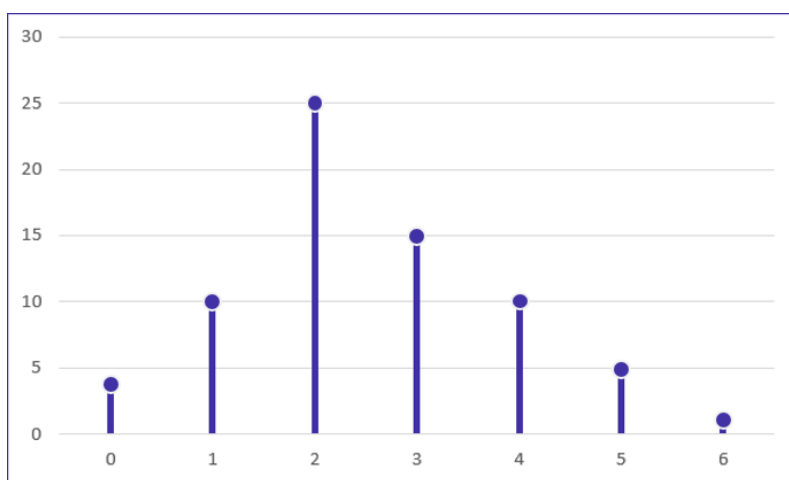


REPRESENTAÇÕES GRÁFICAS DAS DISTRIBUIÇÕES DE FREQUÊNCIA

Gráfico de Hastes ou Bastões

O **gráfico de hastes ou bastões** é muito utilizado para representar **dados não agrupados em classes**, o que normalmente ocorre com dados discretos. Nesse caso, não há perda de informação pois os valores da variável aparecem individualmente, conforme constam da amostra. Com relação a sua construção, basta representarmos as frequências simples absolutas ou relativas de cada elemento do conjunto de dados.

X_i	Frequência (f_i)
0	4
1	10
2	25
3	15
4	10
5	5
6	1



Repare que podemos reconstruir facilmente a tabela de frequências a partir do gráfico de hastes. De igual modo, conhecendo a tabela de frequências, podemos construir rapidamente o gráfico de hastes.

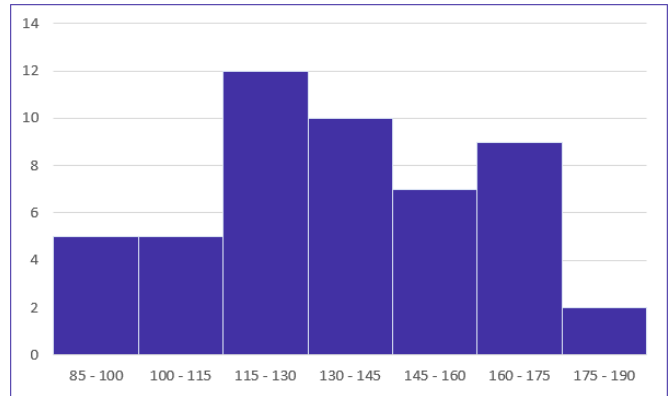
Histogramas

O **histograma** é um gráfico destinado a representar **dados agrupados em classes**, sendo composto por um conjunto de **retângulos contíguos (justapostos)** cujas bases estão situadas sobre o eixo horizontal (eixo x), de forma que os seus pontos médios devem coincidir com os pontos médios dos intervalos de classe e seus limites devem coincidir com os limites da classe.

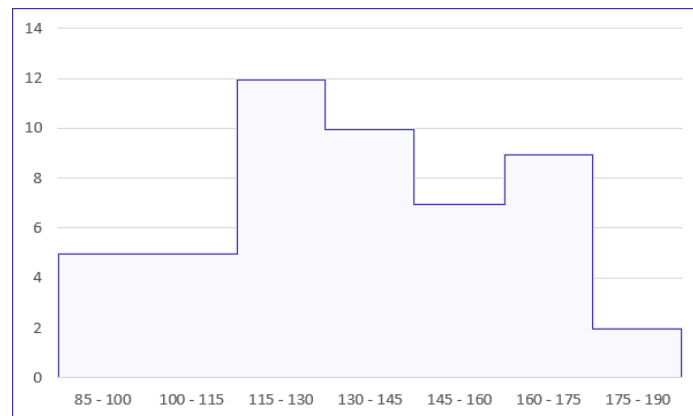
A quantidade de retângulos em um histograma é equivalente ao número de intervalos de classe. A largura de cada retângulo deve ser igual à amplitude do intervalo de classe, enquanto a altura precisa ser proporcional à frequência do intervalo de classe. Além disso, **a área do histograma é proporcional ao somatório das frequências**.



Tempo médio (X_i)	Frequência (f_i)
$85 \leq x < 100$	5
$100 \leq x < 115$	5
$115 \leq x < 130$	12
$130 \leq x < 145$	10
$145 \leq x < 160$	7
$160 \leq x < 175$	9
$175 \leq x < 190$	2



A diferença visual entre um **histograma** e um **gráfico de colunas** (estudaremos na próxima seção) é a **inexistência de separação** entre os retângulos adjacentes. Dito isso, é importante mencionarmos a existência do gráfico denominado de **polígono característica**, que construímos utilizando apenas os **contornos** do histograma.



Por fim, **o histograma pode ocasionar um certo nível de perda de informações**, pois os elementos da distribuição de frequência não são representados de forma individualizada, mas sim por meio de suas classes.

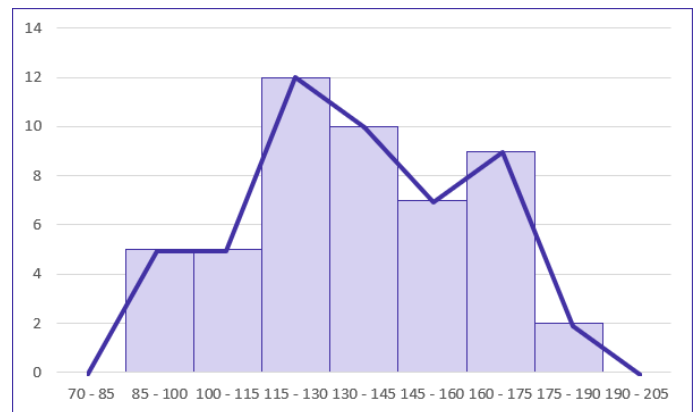
Polígono de Frequências

O **polígono de frequências** é um gráfico em linha obtido por meio da ligação, por segmentos de reta, dos pontos médios das bases superiores dos retângulos de um histograma. Também é necessário considerar a existência de uma classe anterior à primeira e outra posterior à última, ambas com a frequência nula.

Assim como o histograma, **o polígono de frequências apresenta área proporcional ao somatório das frequências**.



Tempo médio (X_i)	Ponto médio (PM_i)	Frequência (f_i)
$70 \leq x < 85$	77,5	0
$85 \leq x < 100$	92,5	5
$100 \leq x < 115$	107,5	5
$115 \leq x < 130$	122,5	12
$130 \leq x < 145$	137,5	10
$145 \leq x < 160$	152,5	7
$160 \leq x < 175$	167,5	9
$175 \leq x < 190$	182,5	2
$190 \leq x < 205$	197,5	0



Curva de Frequências

A curva de frequências é obtida a partir do polimento de um polígono de frequências. Em sentido geométrico, o polimento corresponde à eliminação dos vértices (cantos) da linha poligonal. **Esse processo suaviza os contornos do polígono de frequências, o que evidencia a verdadeira natureza dos dados em análise.**

O polígono de frequências fornece a imagem real do fenômeno investigado, enquanto a curva de frequência mostra sua tendência. Naturalmente, quando o conjunto de dados é grande, a linha poligonal se torna curva. Por isso, podemos afirmar que **a curva de frequência** antecipa o comportamento da distribuição para um número maior de dados.

O processo de polimento é realizado por meio da seguinte fórmula:

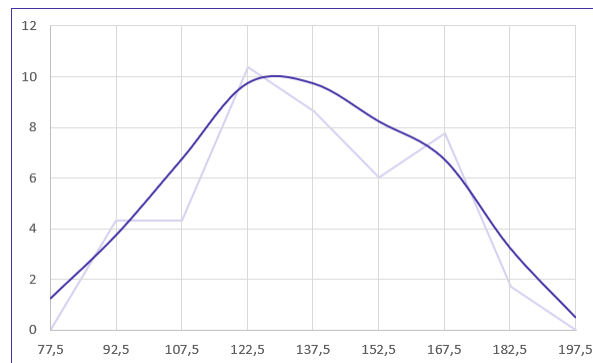
$$f_{c_i} = \frac{f_{ant} + 2 \times f_i + f_{post}}{4}$$

em que f_{c_i} é a frequência calculada da classe considerada (freq. polida); f_i é a frequência simples da classe considerada; f_{ant} é a frequência simples da classe anterior à da classe considerada; e f_{post} é a frequência simples da classe posterior à da classe considerada.

Vejamos como utilizá-la:

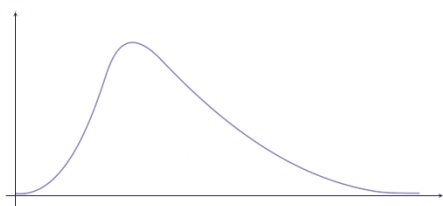


Tempo médio (X_i)	Ponto médio (PM_i)	Frequência (f_i)	Frequência Calculada (f_{c_i})
$70 \leq x < 85$	77,5	0	$f_{c_0} = \frac{0 + 2 \times 0 + 5}{4} = 1,25$
$85 \leq x < 100$	92,5	5	$f_{c_1} = \frac{0 + 2 \times 5 + 5}{4} = 3,75$
$100 \leq x < 115$	107,5	5	$f_{c_2} = \frac{5 + 2 \times 5 + 12}{4} = 6,75$
$115 \leq x < 130$	122,5	12	$f_{c_3} = \frac{5 + 2 \times 12 + 10}{4} = 9,75$
$130 \leq x < 145$	137,5	10	$f_{c_4} = \frac{12 + 2 \times 10 + 7}{4} = 9,75$
$145 \leq x < 160$	152,5	7	$f_{c_5} = \frac{10 + 2 \times 7 + 9}{4} = 8,75$
$160 \leq x < 175$	167,5	9	$f_{c_6} = \frac{7 + 2 \times 9 + 2}{4} = 6,75$
$175 \leq x < 190$	182,5	2	$f_{c_7} = \frac{9 + 2 \times 2 + 0}{4} = 3,25$
$190 \leq x < 205$	197,5	0	$f_{c_8} = \frac{2 + 2 \times 0 + 0}{4} = 0,50$



As curvas de frequências podem assumir as seguintes formas características:

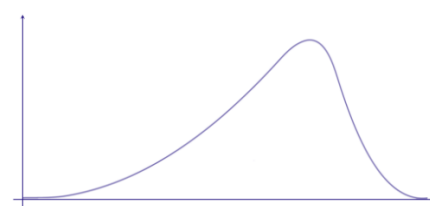
a) **curvas em forma de sino**: são curvas que apresentam concentração de valores em torno da região central da distribuição, podendo ser **simétricas** ou **assimétricas**. Quando assimétricas, as curvas ainda podem apresentar uma cauda mais alongada à esquerda (assimetria à esquerda) ou mais alongada à direita (assimetria à direita). Vejamos as possíveis configurações:



Curva assimétrica à direita.

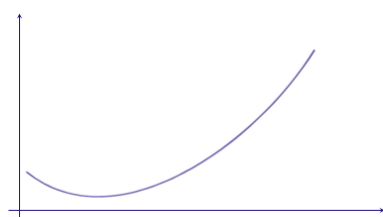


Curva Simétrica.

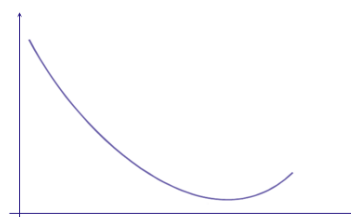


Curva assimétrica à esquerda.

b) **curvas em forma de jota**: são curvas que apresentam o ponto de **ordenada máxima** em uma das **extremidades**, representando distribuições extremamente assimétricas. As possíveis configurações são:



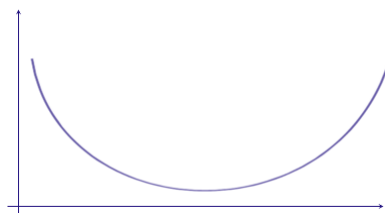
Curva em formato de J.



Curva em formato de J invertido.



c) **curvas em forma de U**: são curvas que apresentam as **ordenadas máximas** em **ambas as extremidades**.



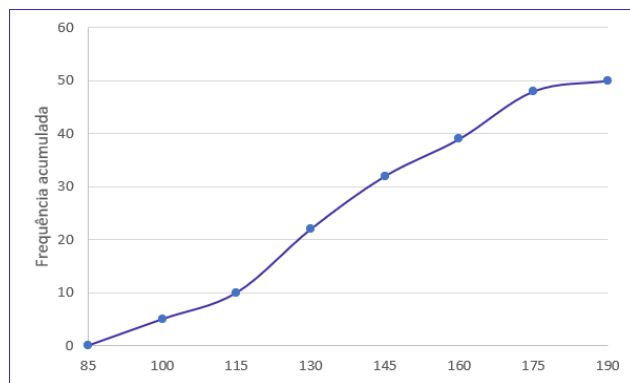
Curva em formato de U.

Ogiva (Polígono de Frequências Acumuladas)

O **gráfico de ogiva** corresponde a um **polígono de frequências acumuladas**. Esse gráfico é empregado na representação de distribuições de frequências acumuladas, sejam elas crescentes ou decrescentes. No eixo horizontal, colocamos as extremidades de cada classe e no eixo vertical as frequências acumuladas.

Ao contrário do polígono de frequências, a **ogiva** utiliza os pontos extremos das classes, e não os pontos médios. Na construção do polígono de frequências acumuladas, devemos considerar a existência de uma classe anterior à primeira, com frequência nula.

Tempo médio (X_i)	Frequência (f_i)	Frequência Acumulada (f_{ac})
$70 \leq x < 85$	0	0
$85 \leq x < 100$	5	5
$100 \leq x < 115$	5	10
$115 \leq x < 130$	12	22
$130 \leq x < 145$	10	32
$145 \leq x < 160$	7	39
$160 \leq x < 175$	9	48
$175 \leq x < 190$	2	50



OUTROS GRÁFICOS E REPRESENTAÇÕES

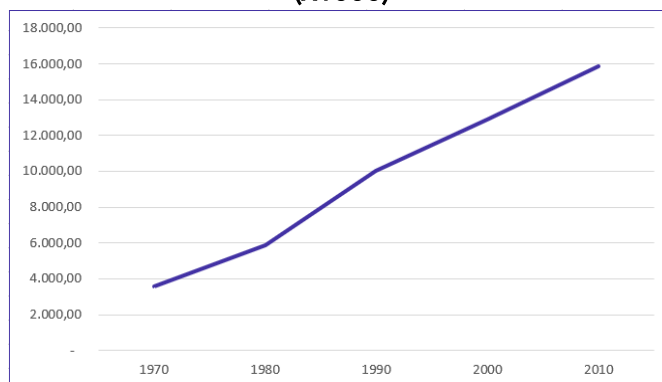
O principal objetivo dos gráficos estatísticos é proporcionar uma visualização mais rápida dos dados estatísticos ou do fenômeno sob investigação. A seguir vamos ver as principais formas de representação de dados estatísticos.

Gráficos em Linhas

Os **gráficos em linha** normalmente **são usados para representar dados de séries temporais**, com a finalidade de mostrar a variação dos valores de uma variável ao longo do tempo. Esse tipo de gráfico permite-nos comparar duas variáveis: uma é traçada no eixo x (horizontal) e a outra no eixo y (vertical). O eixo y geralmente indica uma quantidade, enquanto o eixo x representa uma unidade de tempo.

O gráfico ao lado mostra a evolução da população residente na Região Norte do Brasil no período de 1970 a 2010. O eixo horizontal indica a variação do tempo, enquanto o eixo vertical apresenta a população residente na região.

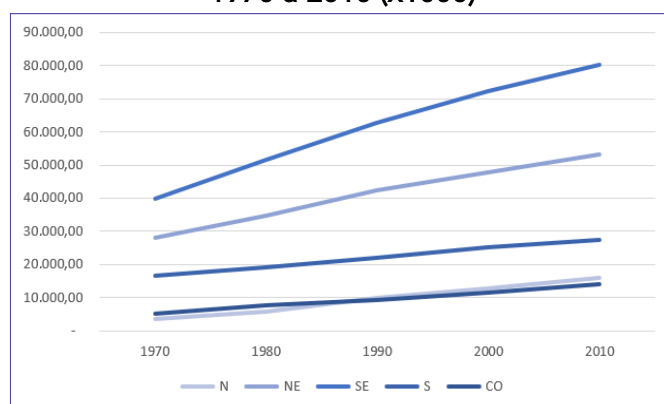
População da região Norte, de 1970 a 2010
(x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Também podemos elaborar um **gráfico de linhas múltiplas** para comparar a evolução da população residente nas Grandes Regiões do Brasil em diferentes períodos:

População brasileira, por Grandes Regiões, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

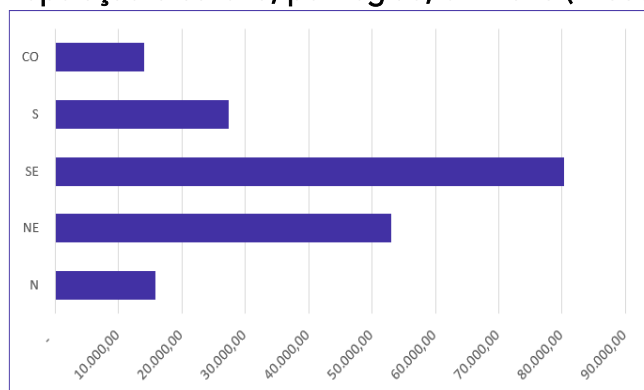


Gráficos em Barras

Os **gráficos em barra** são usados para representar distribuições de **dados categóricos ou qualitativos**. Uma série estatística é representada por um **conjunto de retângulos dispostos horizontalmente**, cada um indicando uma categoria particular, **os quais possuem a mesma altura e comprimentos proporcionais aos respectivos dados**.

A distribuição da população residente em cada região brasileira, no ano de 2010, pode ser representada por meio do gráfico de barras mostrado ao lado.

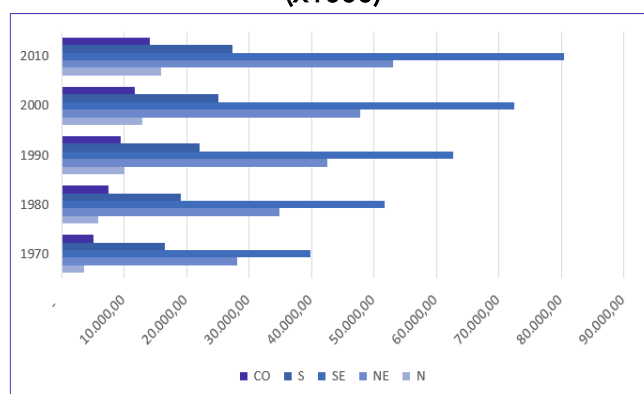
População brasileira, por região, em 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Para representar a evolução da população residente em cada região brasileira, podemos utilizar um **gráfico de barras justapostas**, como o apresentado ao lado.

População brasileira, por região, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

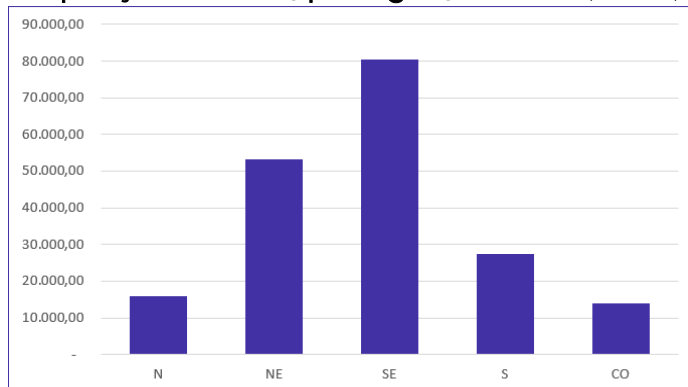
Gráficos em Colunas

Os **gráficos em coluna** também são usados para distribuições de **dados categóricos ou qualitativos**. A diferença básica é que, agora, uma série estatística é representada por um **conjunto de retângulos dispostos verticalmente**, cada um indicando uma categoria particular, todos com a mesma largura e alturas proporcionais aos respectivos dados.



Por exemplo, a distribuição da população residente em cada região brasileira, no ano de 2010, pode ser representada por meio do gráfico ao lado:

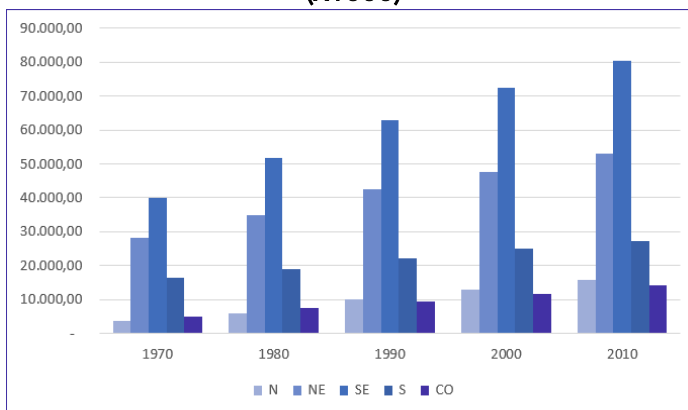
População brasileira, por região, em 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Também podemos utilizar um **gráfico de colunas justapostas** para representar a evolução da população residente em cada região brasileira, no período de 1970 a 2010. Dessa maneira, conseguimos apresentar mais informações em um espaço consideravelmente menor.

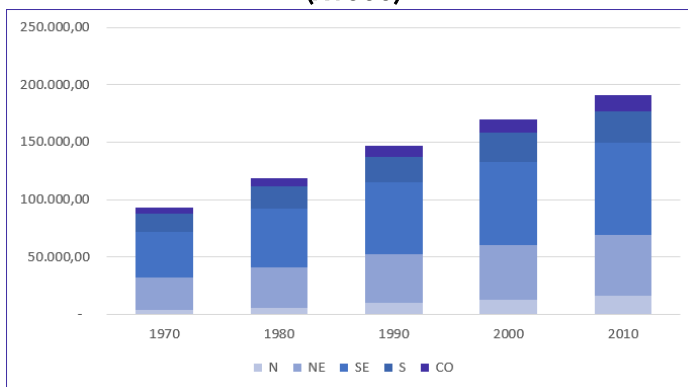
População brasileira, por região, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Adicionalmente, essas informações também podem ser representadas por meio de um **gráfico de colunas sobrepostas** (ou **gráfico de colunas empilhadas**). Esse tipo de gráfico é considerado uma extensão do formato tradicional, pois **permite analisarmos duas dimensões de uma variável categórica**, em vez de apenas uma. Cada coluna é dividida em várias partes que ficam empilhadas umas sobre as outras, cada uma correspondendo a um nível da segunda variável categórica.

População brasileira, por região, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

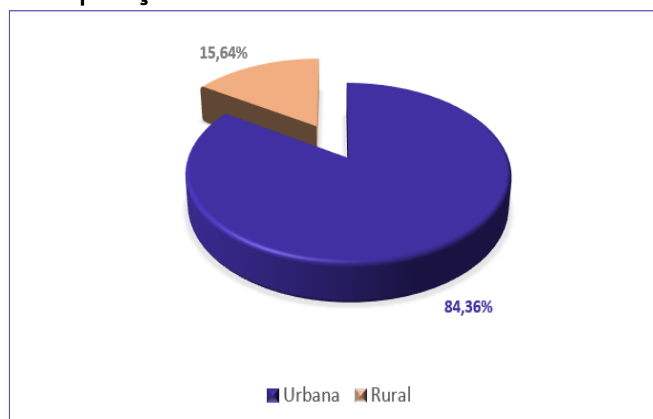


Gráfico em Setores

O **gráfico em setores** (também conhecido como **gráfico de pizza**) é usado para representar a **frequência relativa (porcentagem)** de uma variável categórica. Ele é formado por um círculo dividido em setores circulares, cada um representando uma categoria, cujos ângulos centrais são proporcionais às frequências relativas da categoria.

Para a construção do gráfico de setores, utilizaremos uma regra de três simples, em que as frequências relativas de cada categoria correspondem ao ângulo central que desejamos representar em relação à frequência total, que corresponde ao ângulo de 360°.

População Urbana e Rural do Brasil em 2010

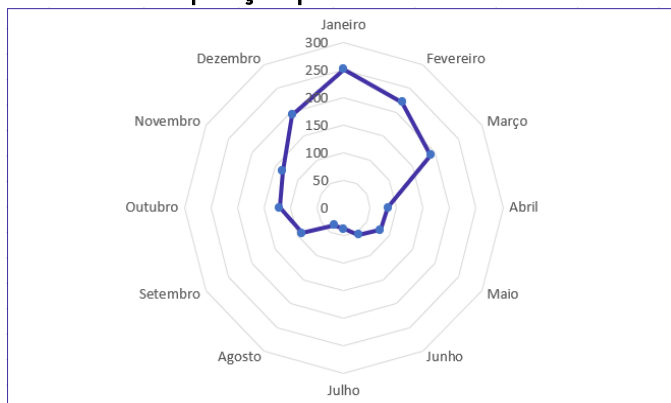


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

Gráfico Polar

O **gráfico polar** consiste em uma sequência de eixos igualmente espaçados (ângulos iguais), cada um representando uma das variáveis. Uma linha é desenhada ligando os valores de cada eixo. Esse tipo de gráfico é usado para representar **séries temporais cíclicas**, que apresentam uma determinada periodicidade, como é o caso da precipitação pluviométrica mensal de uma determinada cidade.

Precipitação pluviométrica média



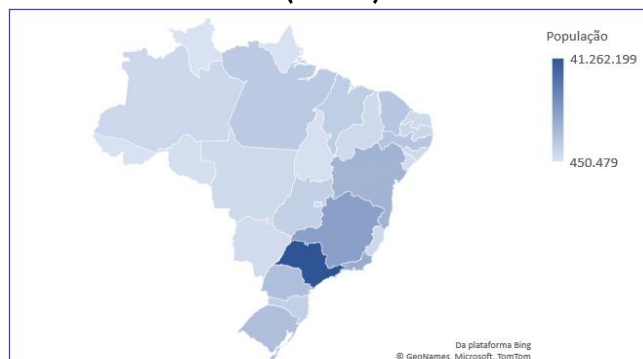
Fonte: Atlas Pluviométrico do Brasil



Cartograma

O **cartograma** é empregado com a finalidade de apresentar **dados estatísticos diretamente relacionados com áreas geográficas**. As áreas do cartograma podem ser preenchidas por pontos, hachuras ou cores. O significado do preenchimento será indicado em uma legenda. Vejamos a população residente em cada Estado brasileiro:

População residente no Brasil por Estado em 2010
(x1000)

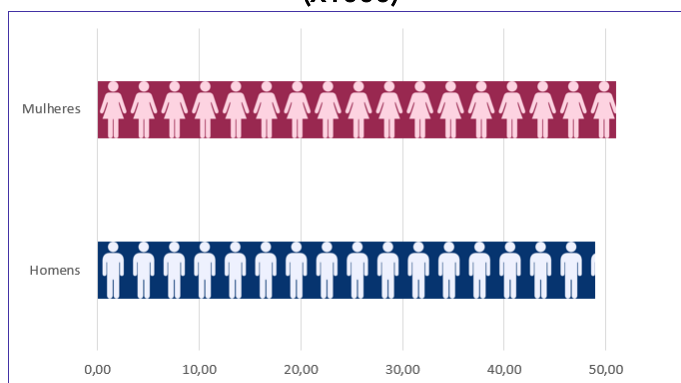


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

Pictograma

O **pictograma** **substitui valores por ícones**, tornando os dados mais atraentes e facilitando o entendimento acerca de um determinado fenômeno. Normalmente, uma legenda é utilizada para indicar o que cada ícone representa. Os ícones devem possuir o mesmo tamanho, mas podem aparecer fracionados para mostrar a respectiva fração de uma determinada quantidade. A proporção de homens e mulheres na população brasileira é apresentada no pictograma ao lado.

População residente no Brasil por Estado em 2010
(x1000)

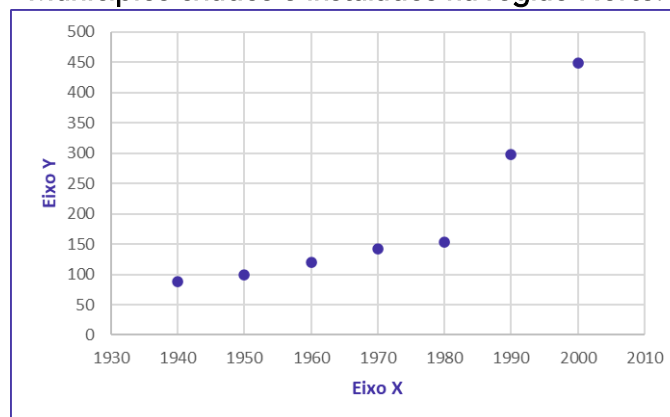


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

Gráfico de Dispersão

O **gráfico de dispersão** é uma **representação de pares ordenados** em um **plano cartesiano**, composto por um eixo vertical (ordenada) e um eixo horizontal (abscissa). Os dados são representados como uma coleção de pontos, cada um com o valor de uma variável determinando a posição no eixo horizontal e o valor da outra variável determinando a posição no eixo vertical. É uma ferramenta poderosa para estudar a relação entre duas variáveis.

Municípios criados e instalados na região Norte.



Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)



Diagrama de Ramos e Folhas

O **diagrama de ramos e folhas** fornece uma maneira rápida de representar graficamente a distribuição dos dados. Nesse diagrama, cada número é separado em duas partes. Em geral, de um lado ficam as unidades do número e do outro lado fica o restante desse número. Consideremos o seguinte rol crescente:

85, 89, 96, 98, 99, 103, 104, 105, 113, 114, 115, 115, 123, 123, 124, 126, 126, 126, 127, 128, 129, 129, 134, 135, 135, 135, 137, 137, 137, 142, 143, 143, 148, 153, 154, 155, 157, 158, 159, 161, 161, 165, 168, 170, 171, 171, 171, 173, 175, 175

Ao utilizarmos um diagrama de ramos e folhas para representar essa distribuição, no lado esquerdo, teremos as centenas e as dezenas indicando os ramos. Enquanto, no lado direito, teremos as unidades representando as folhas. As folhas, portanto, estão vinculadas aos ramos. Dessa maneira, a chave "9 | 6 8 9" significa que, no rol original, havia os números 96, 98 e 99.

8	5 9	Chave: 8 5 = 85
9	6 8 9	
10	3 4 5	
11	3 4 5 5	
12	3 3 4 6 6 6 7 8 9 9	
13	4 5 5 5 7 7 7	
14	2 3 3 8	
15	3 4 5 7 8 9	
16	1 1 5 8	
17	0 1 1 1 3 5 5	

Também é comum encontrarmos diagramas de ramos e folhas em que as unidades são separadas em dois grupos: de 0 a 4 e de 5 a 9. Nesse caso, teríamos o seguinte diagrama:

8	5 9	Chave: 8 5 = 85
9	6 8 9	
10	3 4	
10	5	
11	3 4	
11	5 5	
12	3 3 4	
12	6 6 6 7 8 9 9	
13	4	
13	5 5 5 7 7 7	
14	2 3 3	
14	8	
15	3 4	
15	5 7 8 9	
16	1 1	
16	5 8	
17	0 1 1 1 3	
17	5 5	

Não existe uma regra única para a construção do diagrama de ramos e folhas. O formato mais comumente encontrado é o que separa o número em duas partes, porém, **a depender da chave escolhida**, o número pode ser separado em mais partes.



QUESTÕES COMENTADAS – CESGRANRIO

Conceitos Iniciais

1. (CESGRANRIO/BB/2021) Responsável por entender o comportamento dos produtos oferecidos por determinado banco onde trabalhava, e preocupado com a quantidade enorme de dados disponíveis para a análise, um funcionário decidiu extrair um subconjunto desses dados.

Esse subconjunto é conhecido como

- a) amostra
- b) censo
- c) parâmetro
- d) população
- e) variável

Comentários:

A questão traz o conceito de amostra. Uma amostra é um subconjunto extraído da população para análise, que deve ser representativo. A partir das informações coletadas da amostra representativa, os resultados obtidos podem ser utilizados para generalizar, inferir ou tirar conclusões acerca da população. Portanto, o subconjunto coletado do banco de dados, que representa a população, é a amostra.

Vejamos as definições das demais alternativas:

Alternativa B: o censo, ou recenseamento, é um estudo dos dados referentes a todos os elementos de uma população.

Alternativa C: os parâmetros são descrições numéricas de características populacionais que raramente são conhecidas.

Alternativa D: uma população é um conjunto que contém todos os indivíduos, objetos ou elementos a serem estudados, que apresentam uma ou mais características em comum.

Alternativa E: uma variável é uma característica de relevância que é medida em cada componente da amostra ou população. As variáveis podem ter valores numéricos (variáveis quantitativas) ou não numéricos (variáveis qualitativas).

Gabarito: A.



QUESTÕES COMENTADAS – CESGRANRIO

Variáveis Estatísticas

1. (CESGRANRIO/BB/2021) Foi solicitado a um funcionário de um determinado banco que realizasse uma pesquisa, exclusivamente com variáveis do tipo qualitativa, sobre a satisfação dos clientes com os serviços oferecidos pela instituição.

Para atender a essa demanda utilizando os meios adequados, sua escolha de escalas de mensuração deve estar limitada às escalas

- a) intervalares e razão
- b) nominais e intervalares
- c) nominais e ordinais
- d) ordinais e intervalares
- e) ordinais e razão

Comentários:

As variáveis qualitativas representam as características que não podem ser descritas de forma numérica, mas que podem ser definidas por meio de qualidades (atributos ou categorias) do indivíduo pesquisado. Elas podem ser classificadas em nominais ou ordinais:

- i) variável qualitativa nominal (ou categórica), quando as possíveis categorias não podem ser ordenadas. Por exemplo, as cores dos olhos dos moradores de uma determinada cidade (pretos, castanhos, azuis e verdes);
- ii) variável qualitativa ordinal, as possíveis categorias podem ser ordenadas de alguma forma. Por exemplo, o grau de instrução dos funcionários de um determinado órgão (fundamental, médio e superior).

Gabarito: C.



QUESTÕES COMENTADAS – CESGRANRIO

Representação Gráfica das Distribuições de Frequências

1. (CESGRANRIO/BB/2021) Após a coleta de dados em um determinado contexto (variáveis A, B, C, ... X), uma das formas mais simples e iniciais de análise é a geração e a avaliação de um histograma para uma variável selecionada (ex: X), como por exemplo, em um estudo climático, em que os dados coletados poderiam incluir a temperatura máxima observada em toda a Terra ao longo de dez anos.

Nesse caso, o histograma adequado é um gráfico em que são apresentadas as

- a) últimas dez médias móveis da variável X
- b) somas das médias dos quadrados de cada valor de uma variável X
- c) variações de uma variável X ao longo do tempo
- d) médias históricas da variável X nos últimos sete dias
- e) frequências de uma variável X em intervalos de valores.

Comentários:

O histograma é um gráfico destinado a representar dados agrupados em classe, sendo composto por um conjunto de retângulos contíguos (justapostos) cujas bases ficam localizadas sobre o eixo horizontal (eixo x), de forma que os seus pontos médios devem coincidir com os pontos médios dos intervalos de classe e seus limites devem coincidir com os limites de cada classe. A quantidade de retângulos em um histograma é equivalente ao número de classes. A largura de cada retângulo é equivalente à amplitude do intervalo de classe, enquanto a altura deve ser proporcional à frequência da classe.

Logo, podemos estabelecer que o histograma é um gráfico em que são apresentadas as frequências de uma variável X em intervalos de valores.

Gabarito: E.



LISTA DE QUESTÕES – CESGRANRIO

Conceitos Iniciais

1. (CESGRANRIO/BB/2021) Responsável por entender o comportamento dos produtos oferecidos por determinado banco onde trabalhava, e preocupado com a quantidade enorme de dados disponíveis para a análise, um funcionário decidiu extrair um subconjunto desses dados.

Esse subconjunto é conhecido como

- a) amostra
- b) censo
- c) parâmetro
- d) população
- e) variável



GABARITO – CESGRANRIO

Conceitos Iniciais

1. LETRA A



LISTA DE QUESTÕES – CESGRANRIO

Variáveis Estatísticas

1. (CESGRANRIO/BB/2021) Foi solicitado a um funcionário de um determinado banco que realizasse uma pesquisa, exclusivamente com variáveis do tipo qualitativa, sobre a satisfação dos clientes com os serviços oferecidos pela instituição.

Para atender a essa demanda utilizando os meios adequados, sua escolha de escalas de mensuração deve estar limitada às escalas

- a) intervalares e razão
- b) nominais e intervalares
- c) nominais e ordinais
- d) ordinais e intervalares
- e) ordinais e razão



GABARITO – CESGRANRIO

Variáveis Estatísticas

1. LETRA C



LISTA DE QUESTÕES – CESGRANRIO

Representação Gráfica das Distribuições de Frequências

1. (CESGRANRIO/BB/2021) Após a coleta de dados em um determinado contexto (variáveis A, B, C, ... X), uma das formas mais simples e iniciais de análise é a geração e a avaliação de um histograma para uma variável selecionada (ex: X), como por exemplo, em um estudo climático, em que os dados coletados poderiam incluir a temperatura máxima observada em toda a Terra ao longo de dez anos.

Nesse caso, o histograma adequado é um gráfico em que são apresentadas as

- a) últimas dez médias móveis da variável X
- b) somas das médias dos quadrados de cada valor de uma variável X
- c) variações de uma variável X ao longo do tempo
- d) médias históricas da variável X nos últimos sete dias
- e) frequências de uma variável X em intervalos de valores.



GABARITO – CESGRANRIO

Representação Gráfica das Distribuições de Frequências

1. LETRA E



ESSA LEI TODO MUNDO CONHECE: PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.



1 Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.



2 Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).



3 Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).



4 Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).



5 Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).



6 Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).



7 Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.



8 O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.



Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.