

## **Aula 00 - Equipe Exatas**

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

Autor:

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

11 de Novembro de 2024

# Índice

|   |     |
|---|-----|
| 1) Apresentação do Curso .....  | 4   |
| 2) Aviso .....  | 5   |
| 3) Introdução à Estatística - CNU .....   | 6   |
| 4) Conceitos Iniciais - CNU .....   | 10  |
| 5) Dados Estatísticos - CNU .....   | 15  |
| 6) Variáveis Estatísticas - CNU .....   | 17  |
| 7) Séries Estatísticas - CNU .....  | 21  |
| 8) Distribuição de Frequências - CNU .....  | 26  |
| 9) Representação Gráfica das Distribuições de Frequências - CNU .....                               | 44  |
| 10) Questões Comentadas - Conceitos Iniciais - Cesgranrio .....                                     | 55  |
| 11) Questões Comentadas - Variáveis Estatísticas - Cesgranrio .....                                 | 57  |
| 12) Questões Comentadas - Distribuições de Frequência - CESGRANRIO .....                            | 61  |
| 13) Questões Comentadas - Representação Gráfica das Distribuições de Frequências - Cesgranrio ..... | 63  |
| 14) Questões Comentadas - Outros Gráficos e Representações - Cesgranrio .....                       | 64  |
| 15) Aviso importante - Orientação de estudo .....   | 67  |
| 16) Questões Comentadas - Introdução à Estatística - Inéditas .....                                 | 68  |
| 17) Questões Comentadas - Conceitos Iniciais - Inéditas .....                                       | 69  |
| 18) Questões Comentadas - Variáveis Estatísticas - Inéditas .....                                   | 71  |
| 19) Questões Comentadas - Séries Estatísticas - Inéditas .....                                      | 74  |
| 20) Questões Comentadas - Distribuições de Frequência - Inéditas .....                              | 75  |
| 21) Questões Comentadas - Representação Gráfica das Distribuições de Frequências - Inéditas .....   | 85  |
| 22) Questões Comentadas - Outros Gráficos e Representações - Inéditas .....                         | 90  |
| 23) Lista de Questões - Conceitos Iniciais - Cesgranrio .....                                       | 99  |
| 24) Lista de Questões - Variáveis Estatísticas - Cesgranrio .....                                   | 101 |
| 25) Lista de Questões - Distribuições de Frequência - CESGRANRIO .....                              | 104 |
| 26) Lista de Questões - Representação Gráfica das Distribuições de Frequências - Cesgranrio .....   | 106 |
| 27) Lista de Questões - Outros Gráficos e Representações - Cesgranrio .....                         | 108 |
| 28) Lista de Questões - Introdução à Estatística - Inéditas .....                                   | 111 |



## Índice

|   |     |
|---|-----|
| 29) Lista de Questões - Conceitos Iniciais - Inéditas .....                                     | 113 |
| 30) Lista de Questões - Variáveis Estatísticas - Inéditas .....                                 | 115 |
| 31) Lista de Questões - Séries Estatísticas - Inéditas .....                                    | 118 |
| 32) Lista de Questões - Distribuições de Frequência - Inéditas .....                            | 120 |
| 33) Lista de Questões - Representação Gráfica das Distribuições de Frequências - Inéditas ..... | 126 |
| 34) Lista de Questões - Outros Gráficos e Representações - Inéditas .....                       | 129 |



## APRESENTAÇÃO DO CURSO

Olá, pessoal! Tudo bem?

É com grande satisfação damos início ao nosso curso de **Estatística!**

Os professores **Luana Brandão** e **Djefferson Maranhão** ficarão responsáveis pela elaboração do **Livro Digital**.

Antes de continuarmos, vamos apresentar os professores do material escrito:

**Luana Brandão:** Oi, pessoal! O meu nome é Luana Brandão e sou professora de Estatística do Estratégia Concursos. Sou Graduada, Mestre e Doutora em Engenharia de Produção, pela Universidade Federal Fluminense. Passei nos concursos de Auditor Fiscal (2009/2010) e Analista Tributário (2009) da Receita Federal e de Auditor Fiscal do Estado do Rio de Janeiro (2010). Sou Auditora Fiscal do Estado do RJ desde 2010. Vamos juntos nesse caminho até a aprovação?  **@professoraluanabrandao**

**Djefferson Maranhão:** Olá, caros alunos! Meu nome é Djefferson Maranhão, sou professor de Estatística do Estratégia Concursos. Graduado e Mestre em Ciência da Computação pela Universidade Federal do Maranhão (UFMA). Desde 2015, sou Auditor da Controladoria Geral do Estado do Maranhão (2015 - 5º lugar). Também exerci os cargos de Analista de Sistemas na UFMA (2010 - 1º lugar) e no TJ-MA (2011 - 1º lugar). Sinto-me honrado em fazer parte de sua jornada rumo à aprovação.

O material escrito em **PDF** está sendo construído para ser sua fonte **autossuficiente** de estudos. Isso significa que o livro digital será **completo e voltado para o seu edital**, justamente para que você não perca o seu precioso tempo procurando o conteúdo que será cobrado na sua prova. Ademais, sempre que necessário, você poderá fazer perguntas sobre as aulas no **fórum de dúvidas**. **Bons estudos!**



## AVISO IMPORTANTE!



Olá, Alunos (as)!

Passando para informá-los a respeito da **disposição das questões** dentro do nosso material didático. Informamos que a escolha das bancas, dentro dos nossos Livros Digitais, é feita de maneira estratégica e pedagógica pelos nossos professores a fim de proporcionar a melhor didática e o melhor direcionamento daquilo que mais se aproxima do formato de cobrança da banca do seu concurso.

Assim, o formato de questões divididas por tópico facilitará o seu processo de estudo, deixando mais alinhado às disposições constantes no edital.

No mais, continuaremos à disposição de todos no Fórum de dúvidas!

Atenciosamente,

Equipe Exatas

Estratégia Concursos



## INTRODUÇÃO À ESTATÍSTICA

A origem da estatística remonta às civilizações antigas, em que vários povos coletavam e registravam dados populacionais e econômicos de interesse do Estado. Nessa época, também eram realizadas estimativas das riquezas individuais e familiares, as quais eram utilizadas para determinar o montante de impostos a serem pagos pela população.

O termo estatística se originou da palavra *status*, que significa Estado em latim. O termo era utilizado para designar um conjunto de dados, relativos aos Estados, que os governantes utilizavam com a finalidade de controle fiscal e segurança nacional. O primeiro a utilizar a palavra foi Schmeitzel, ainda no século XVII, em latim. Depois, foi adotada pelo acadêmico alemão Godofredo Achenwall.

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**. Ou ainda, 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**.

Os avanços computacionais tornaram a Estatística mais acessível e permitiram aplicações mais sofisticadas em diferentes áreas do conhecimento. Nesse cenário, os softwares estatísticos passaram a disponibilizar ferramentas antes inimagináveis, voltadas para planejamento de experimentos, teste de hipóteses, cálculos de confiabilidade, criação de gráficos complexos e elaboração de modelos preditivos.

A Estatística pode ser dividida em três grandes ramos: Estatística Descritiva (ou dedutiva), Estatística Probabilística e Estatística Inferencial (ou indutiva). Alguns autores, porém, consideram a Estatística Probabilística como parte da Estatística Inferencial.

A **Estatística Descritiva (ou Dedutiva)** é **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.

A **Estatística Probabilística** é **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.



A **Estatística Inferencial (ou Indutiva)** é **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.



(AVANÇA-SP/Pref. Lorena/2024) A estatística é a ciência que estuda as técnicas necessárias para coletar, organizar, apresentar, analisar e interpretar os dados, a fim de extrair informações a respeito de uma população.

Dentro da estatística, qual é a parte que trabalha com organização e apresentação dos dados?

- a) Estatística dedutiva
- b) Estatística descritiva
- c) Estatística inferencial
- d) Probabilidade
- e) Estatística indutiva



## Comentários:

A parte da estatística que se ocupa da organização e apresentação dos dados é a Estatística Descritiva. Essa área é responsável por resumir e descrever as características de um conjunto de dados, utilizando técnicas como tabelas, gráficos e medidas de tendência central e dispersão. Portanto, a alternativa B está correta.

Vejam as demais alternativas:

Alternativa A: Incorreta. A Estatística Dedutiva não é um termo amplamente utilizado na área de estatística. A dedução é um processo lógico que parte de premissas gerais para chegar a conclusões específicas. No contexto estatístico, esse termo não é comum, e por isso essa alternativa não representa a parte da estatística que organiza e apresenta os dados.

Alternativa C: Incorreta. A Estatística Inferencial é a área da estatística que, a partir de uma amostra, faz generalizações ou inferências sobre uma população maior. Ela utiliza métodos para estimar parâmetros populacionais, testar hipóteses e prever comportamentos futuros. Embora seja fundamental para a análise estatística, seu foco não é a organização ou apresentação dos dados, mas sim a interpretação e inferência a partir deles.

Alternativa D: Incorreta. A Probabilidade é um ramo da matemática que estuda a chance ou a possibilidade de ocorrência de eventos. Ela serve de base para a estatística inferencial, mas não se dedica à organização ou apresentação de dados. A probabilidade fornece as ferramentas teóricas para modelar eventos aleatórios e incertos.

Alternativa E: Incorreta. A Estatística Indutiva é outro termo usado para se referir à Estatística Inferencial. A indução é o processo de partir de observações específicas para chegar a conclusões gerais. Portanto, assim como a estatística inferencial, a estatística indutiva está preocupada em fazer generalizações sobre a população com base em dados amostrais, e não na organização ou apresentação dos dados.

**Gabarito: B.**

**(VUNESP/Câmara Municipal de São Carlos/2013) Trata-se da estatística que tem as atribuições de obtenção, organização, redução e representação de dados estatísticos, bem como a obtenção de algumas informações que auxiliam na descrição de um fenômeno observado. Esta estatística é denominada**

- a) Coletora.
- b) Celetista.
- c) Populacional.



- d) Amostral.
- e) Descritiva.

**Comentários:**

A Estatística está dividida em três grandes ramos, a saber:

- a) **Estatística Descritiva (ou Dedutiva)** - responsável pela coleta (ou obtenção), organização, redução e representação de dados estatísticos;
- b) **Estatística Probabilística** - responsável por estabelecer o modelo matemático probabilístico adotado para explicar os fenômenos aleatórios investigados pela Estatística; e
- c) **Estatística Inferencial (ou Indutiva)** - responsável pela análise e interpretação dos dados.

Portanto, a alternativa correta é a letra E.

**Gabarito: E.**



## CONCEITOS INICIAIS

Neste tópico, apresentaremos alguns conceitos iniciais da estatística que costumam ser abordados em provas de concursos públicos, dentre os quais podemos citar: população, amostra, censo, amostragem, parâmetros e estatísticas.

### População

Uma **POPULAÇÃO** é um 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

Uma **AMOSTRA** é um **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. Como exemplo, podemos citar as pesquisas eleitorais, em que uma amostra de eleitores deve ser extraída de acordo com a proporcionalidade de gênero, idade, grau de instrução e classe social.

### Censo

O **CENSO**, ou recenseamento, é um 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. Como exemplos, podemos citar a pesquisa sobre o grau de escolaridade dos habitantes brasileiros, o estudo sobre a renda dos brasileiros e a pesquisa de emprego.

### Amostragem

A **AMOSTRAGEM** é um 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,



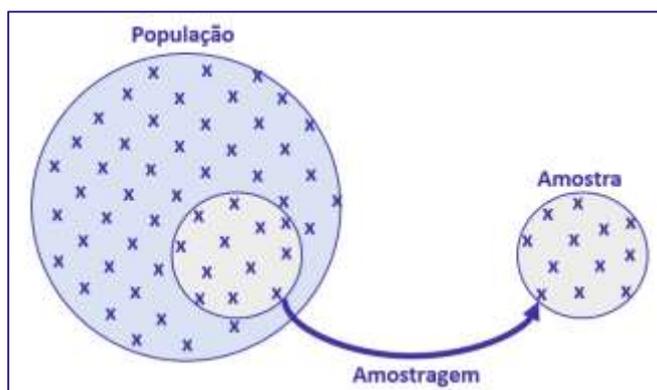
devemos garantir que a amostra seja representativa da população. Isso significa que, com exceção de pequenas discrepâncias inerentes à aleatoriedade existente no processo de amostragem, uma amostra deve possuir as mesmas características básicas da população, no que diz respeito às variáveis que desejamos pesquisar.

## Parâmetros

Os **PARÂMETROS** são **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 (por exemplo, a proporção de homens e mulheres na população brasileira). Normalmente, os parâmetros populacionais são constantes para uma população.

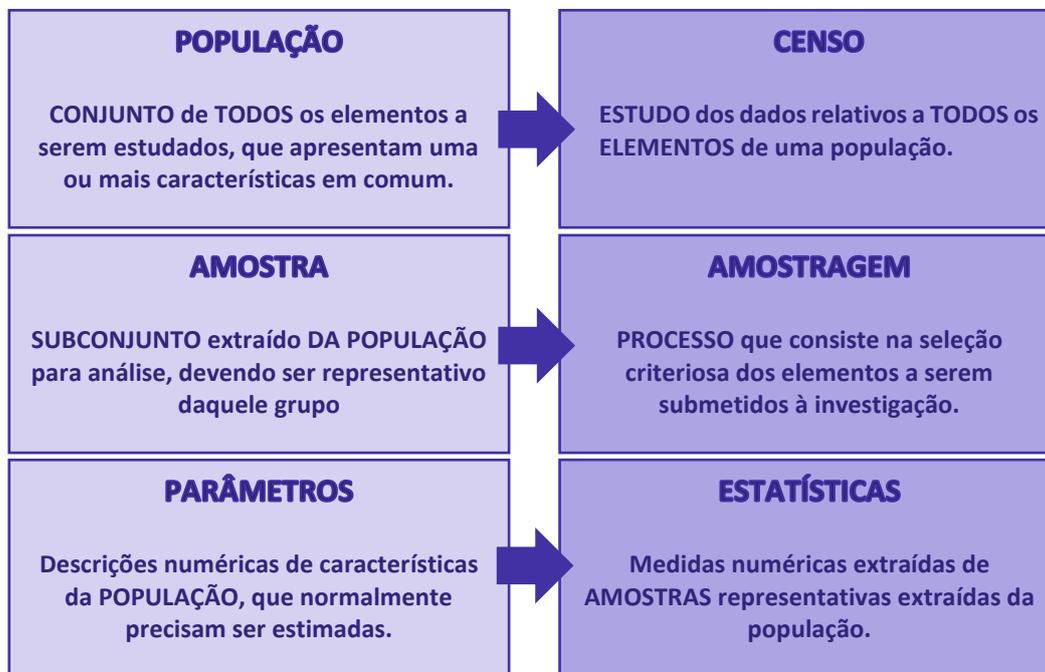
## Estatística (ou estimador)

As **ESTATÍSTICAS** são **MEDIDAS NUMÉRICAS OBTIDAS DE AMOSTRAS** representativas extraídas da população. A partir das informações colhidas da amostra, as **estatísticas amostrais obtidas podem ser utilizadas para inferir ou tirar conclusões acerca dos parâmetros populacionais**, como a proporção de homens e mulheres na população brasileira. De forma resumida, as estatísticas (ou estimadores) são **descrições numéricas de características amostrais**. Normalmente, as estatísticas amostrais diferem de uma amostra para outra.





## ESQUEMATIZANDO



## HORA DE PRATICAR!

(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.

Vejam as definições das demais alternativas:

- Letra B: o censo, ou recenseamento, é um estudo dos dados referentes a todos os elementos de uma população.
- Letra C: os parâmetros são descrições numéricas de características populacionais que raramente são conhecidas.
- Letra 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.
- Letra 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.**

**(VUNESP/Pref. Valinhos/2019) O grupo completo de unidades elementares de pessoas, objetos ou coisas é denominado, para a estatística, de**

- a) Amostra.
- b) Unidades.
- c) Censo.
- d) População.
- e) Variáveis.

### Comentários:

Para resolvermos a questão vamos conceituar cada uma das alternativas:

- Letra A: amostra é um subconjunto de uma população. Os dados coletados e/ou selecionados de uma população estatística por um procedimento definido, denotado de amostragem.
- Letra B: unidades são elementos da amostra (unidades amostrais) ou da população (unidade populacional).
- Letra C: censo consiste na análise de todos os elementos da população.



- Letra D: população é o grupo completo de unidades elementares de pessoas, objetos ou coisas, ou seja, constitui todo o universo de informações de que se necessita.
- Letra E: variável é a 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: D.**

(CESPE/DEPEN/2015) O diretor de um sistema penitenciário, com o propósito de estimar o percentual de detentos que possuem filhos, entregou a um analista um cadastro com os nomes de 500 detentos da instituição para que esse profissional realizasse entrevistas com os indivíduos selecionados.

A partir dessa situação hipotética e dos múltiplos aspectos a ela relacionados, julgue o item, referente a técnicas de amostragem.

A diferença entre um censo e uma amostra consiste no fato de esta última exigir a realização de um número maior de entrevistas.

**Comentários:**

Um censo é a análise de todos os elementos de determinada população. Trabalhamos com uma amostra quando obtemos informações de apenas uma parte dessa população. O censo, portanto, requer a realização de um número maior de entrevistas.

**Gabarito: Errado.**



## DADOS ESTATÍSTICOS

Os **dados estatísticos** constituem os valores resultantes da **coleta de dados**. Os dados referem-se a um **conjunto de valores**, os quais são organizados por meio de **variáveis** (a característica está sendo medida) e **observações** (elementos da amostra/população). É o caso, por exemplo, dos valores obtidos na pesquisa de peso, altura, idade e sexo de uma determinada amostra de indivíduos/população.

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

- dados univariados**: quando uma única observação de cada indivíduo é registrada. Por exemplo: peso;
- dados bivariados**: quando duas observações de cada indivíduo são registradas. Por exemplo: peso e altura;
- dados multivariados**: quando mais duas observações acerca de cada indivíduo 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**. Como exemplo de dados brutos, podemos citar uma relação dos tempos médios de estudo diário, em minutos, de 50 alunos do Estratégia, em que a seleção dos alunos ocorreu de forma aleatória, não havendo qualquer ordenação de valores.

| Aluno | Tempo |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1     | 143   | 11    | 113   | 21    | 170   | 31    | 124   | 41    | 105   |
| 2     | 142   | 12    | 143   | 22    | 158   | 32    | 137   | 42    | 154   |
| 3     | 161   | 13    | 159   | 23    | 123   | 33    | 153   | 43    | 99    |
| 4     | 126   | 14    | 168   | 24    | 96    | 34    | 129   | 44    | 114   |
| 5     | 134   | 15    | 123   | 25    | 98    | 35    | 148   | 45    | 161   |
| 6     | 137   | 16    | 135   | 26    | 135   | 36    | 173   | 46    | 128   |
| 7     | 171   | 17    | 135   | 27    | 129   | 37    | 126   | 47    | 175   |
| 8     | 85    | 18    | 175   | 28    | 126   | 38    | 104   | 48    | 137   |
| 9     | 155   | 19    | 115   | 29    | 103   | 39    | 157   | 49    | 165   |
| 10    | 171   | 20    | 89    | 30    | 171   | 40    | 127   | 50    | 115   |



Esse tipo de tabela, em que os elementos não aparecem numericamente ordenados, é denominada de **tabela primitiva**. A **tabela primitiva**, em geral, **oferece pouca ou nenhuma informação ao leitor**, sendo necessário haver 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. Os dados do nosso exemplo, isto é, os tempos médios de estudo diário, podem ser organizados em ordem crescente ou decrescente:

| Rol (em ordem 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 |

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



## VARIÁVEIS ESTATÍSTICAS

A **variável estatística** consiste no **conjunto de características** que desejamos **averiguar estatisticamente**. Ela também pode ser definida como o **objeto da pesquisa estatística**. Por exemplo, se nosso interesse é conhecer quantas horas os alunos do Estratégia estudam diariamente, então nossa variável é o número de horas estudadas por dia. As variáveis estatísticas podem ser classificadas, inicialmente, 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**:

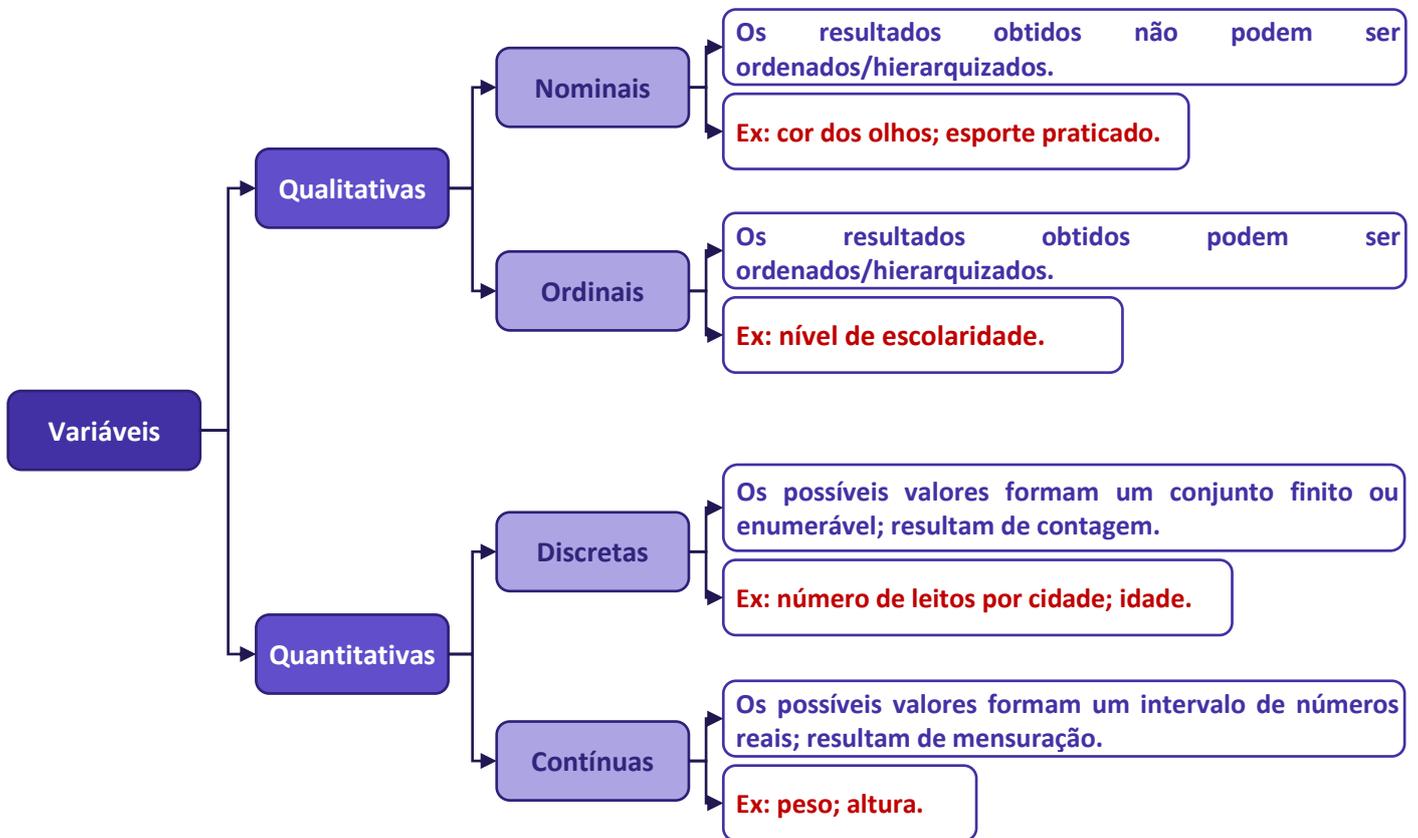
- a) **variável qualitativa nominal**, as possíveis categorias **não podem** ser ordenadas. Por exemplo, a cor dos olhos dos moradores de uma determinada cidade (pretos, castanhos, azuis e verdes);
- b) **variável qualitativa ordinal**, as possíveis categorias **podem** ser ordenadas de alguma forma. 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, o número de leitos abertos em hospitais de uma determinada cidade;
- 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, de modo a não haver lacunas ou interrupções. Por exemplo, a altura dos moradores de uma determinada cidade.





(FGV/TJ-MS/2024) A afirmação que descreve corretamente os conceitos e as relações entre variáveis quantitativas e qualitativas é:

- a) variáveis quantitativas podem ser obtidas a partir de variáveis qualitativas em situações particulares;
- b) variáveis quantitativas são características observadas em populações, enquanto variáveis qualitativas são características observadas em amostras;
- c) variáveis quantitativas são características subjetivas, enquanto variáveis qualitativas são características objetivas;



- d) variáveis quantitativas são contínuas, enquanto variáveis qualitativas são discretas;
- e) variáveis quantitativas são representadas por números inteiros.

### Comentários:

Vamos analisar cada alternativa:

Alternativa A: **Correta**. De fato, variáveis quantitativas podem ser obtidas a partir de variáveis qualitativas em situações particulares. Em algumas situações, variáveis qualitativas podem ser transformadas em variáveis quantitativas por meio de codificação. Por exemplo, atribuir valores numéricos a categorias como "baixo", "médio" e "alto" para representar níveis de satisfação.

Alternativa B: **Incorreta**. Tanto variáveis quantitativas quanto qualitativas podem ser observadas em populações e amostras. A distinção entre as duas não depende do fato de serem observadas em populações ou amostras, mas sim de suas naturezas. Variáveis quantitativas são numéricas e mensuráveis, enquanto variáveis qualitativas são categóricas e descritivas.

Alternativa C: **Incorreta**. Variáveis quantitativas são objetivas porque são mensuráveis e expressas numericamente (por exemplo, altura, peso, idade). Variáveis qualitativas são descritivas e categóricas, e podem ser tanto objetivas (por exemplo, cor dos olhos) quanto subjetivas (por exemplo, satisfação com um serviço).

Alternativa D: **Incorreta**. Variáveis quantitativas podem ser tanto contínuas (por exemplo, altura, peso) quanto discretas (por exemplo, número de filhos, quantidade de carros em um estacionamento). Variáveis qualitativas são categóricas e podem ser nominais (sem ordem) ou ordinais (com ordem).

Alternativa E: **Incorreta**. Variáveis quantitativas podem ser representadas por números inteiros (por exemplo, número de irmãos) ou por números fracionários (por exemplo, altura, peso). Elas não são restritas a números inteiros.

**Gabarito: A.**

(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:

- variável qualitativa nominal (ou categórica): as possíveis categorias não podem ser ordenadas. Por exemplo, as cores dos olhos dos moradores de uma determinada cidade (pretos, castanhos, azuis e verdes);
- 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.**

**(VUNESP/CM São Joaquim Barra/2018) A Estatística descritiva faz uso de variáveis, que são classificadas como quantitativas ou qualitativas. Assinale a alternativa correta em relação a essas variáveis.**

- Quantitativas referem-se às variáveis ordinal ou discreta; as qualitativas referem-se às variáveis nominal ou contínua.
- Quantitativas referem-se às variáveis ordinal ou contínua; as qualitativas referem-se às variáveis nominal ou discreta.
- Quantitativas referem-se às variáveis nominal ou contínua; as qualitativas referem-se às variáveis discreta ou ordinal.
- Quantitativas referem-se às variáveis contínua ou discreta; as qualitativas referem-se às variáveis nominal ou ordinal.
- Quantitativas referem-se às variáveis nominal ou ordinal; as qualitativas referem-se às variáveis contínua ou discreta.

### Comentários:

As variáveis podem ser classificadas em:

- variáveis quantitativas: são as características que podem ser medidas em uma escala quantitativa, isto é, numérica. Elas podem ser contínuas ou discretas.
- variáveis qualitativas (ou categóricas): são as características que não possuem valores quantitativos. Nesse caso, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Elas podem ser nominais ou ordinais.

**Gabarito: D.**



## SÉRIES ESTATÍSTICAS

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

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

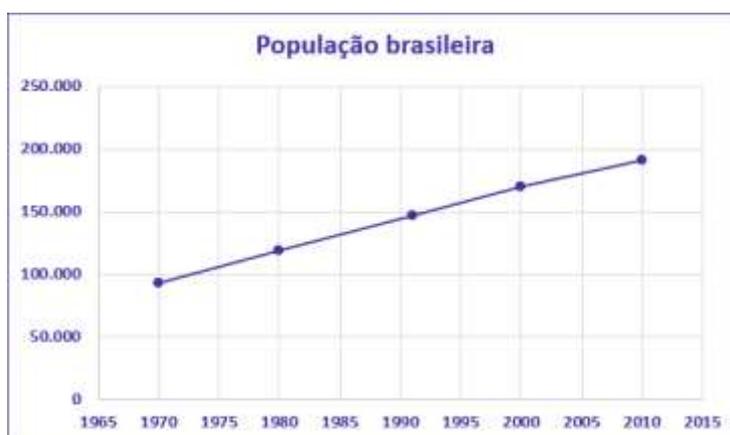
- corpo – conjunto de linhas e colunas com as informações sobre a variável em estudo;
- cabeçalho – parte superior que especifica o conteúdo das colunas;
- coluna indicadora – parte que indica o conteúdo das linhas;
- linhas – traços que facilitam a leitura dos dados;
- célula – espaço onde os dados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

| 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)

Diagrama de anotações: Título (População brasileira no período de 1970 a 2010 (x1000)), Cabeçalho (Anos, População), Coluna indicadora (Anos), Corpo (dados), Célula (2000, 169.799), Linhas (linhas de dados), Coluna numérica (População), Fonte (Fonte: Censo Demográfico (2010)).

Um **gráfico** é uma **forma clara e objetiva** de apresentar uma **série estatística**. O objetivo é **proporcionar uma compreensão mais rápida do fenômeno em estudo**. Para isso, 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 série estatística apresentada na tabela anterior pode ser representada graficamente da seguinte forma:





### Tabela

- Quadro que resume um conjunto de observações.
- Composta de cabeçalho, corpo, coluna indicadora, linhas, células, título e fonte.

### Gráfico

- Forma simples e clara de apresentar uma série estatística.
- Proporcionar uma compreensão mais rápida do fenômeno em estudo.
- Deve ser simples, claro e verossímil.

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 **temporal** (ou cronológica), **geográfica** (ou territorial) e **específica**.

## Séries Temporais (ou Cronológicas)

É a série cujos dados são **dispostos segundo a época de ocorrência**. Isto é, 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**.

A seguir temos a série histórica da população residente no Brasil no período de 1970 a 2010, com frequência decenal. Percebam que o **tempo varia**; contudo, o fato que está sendo analisado (quantidade populacional) e o local alvo da pesquisa (Brasil), continuam iguais.

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. Isto é, 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**.

A seguir temos a série geográfica da população urbana residente em cada uma das regiões brasileiras no ano de 2010. Percebam que o **local (região) varia**; contudo, o fato que está sendo analisado (quantidade populacional) e o tempo (ano de 2010) permanecem constantes.

| <b>Região</b> | <b>População</b> |
|---------------|------------------|
| 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. Isto é, 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**.

A seguir temos uma série específica das populações urbana e rural residentes no Brasil no ano de 2010. Percebam que os fatos analisados variam (população urbana x população rural); contudo, o tempo (2010) e o local de análise (Brasil) são constantes.

| <b>Zona</b>  | <b>População</b> |
|--------------|------------------|
| Urbana       | 93.134           |
| Rural        | 119.011          |
| <b>Total</b> | <b>190.755</b>   |

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**. A seguir 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**. A seguir temos uma série geográfico-temporal representando as populações residentes em cada região brasileira, no período de 1970 a 2010, com variação decenal.

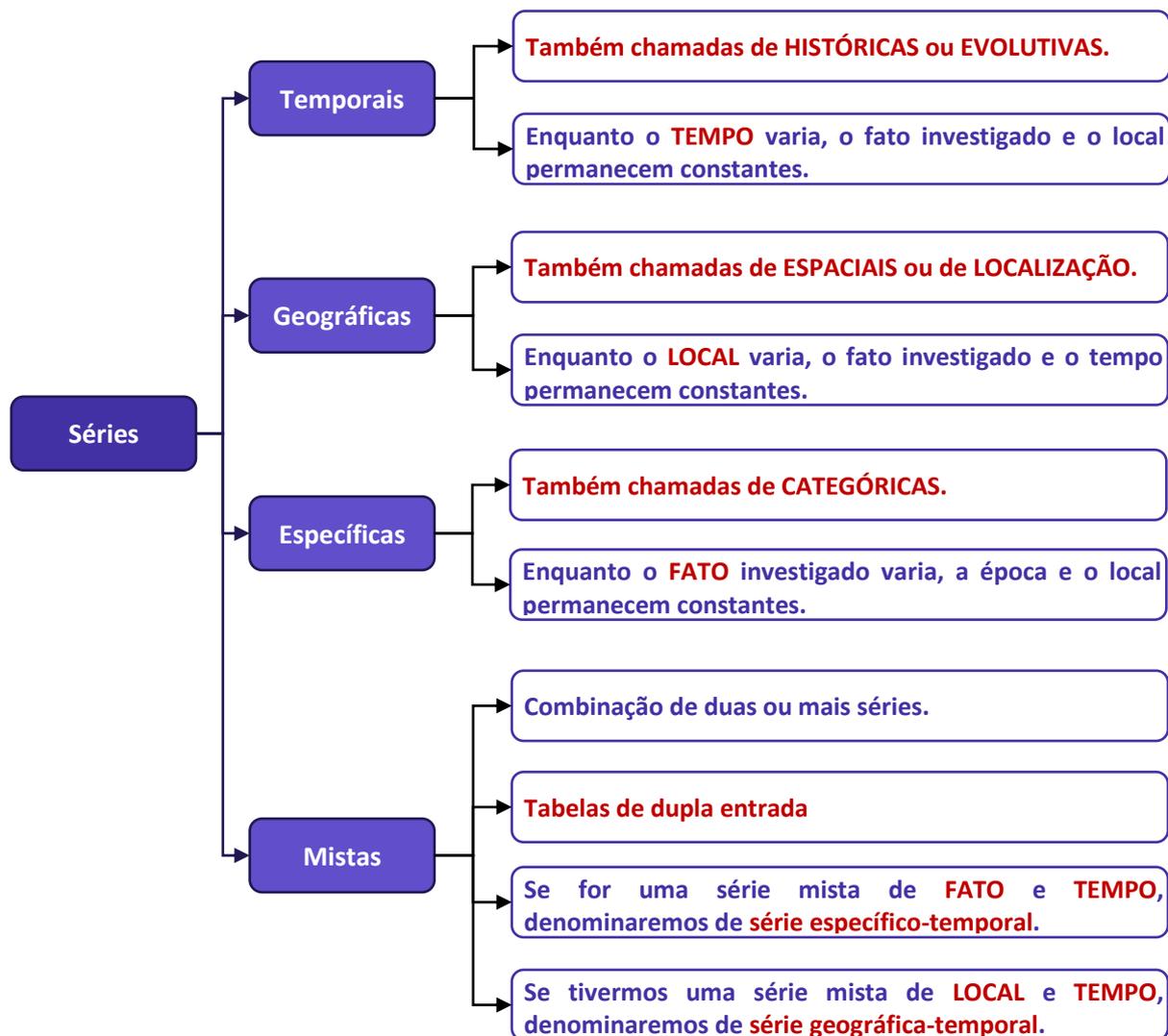
**População do Brasil por Região de 1970 a 2010 (x1000)**

| 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

Vimos anteriormente que, logo após a coleta de dados, temos o que denominamos de **dados brutos**. Como exemplo de dados brutos, citamos uma pesquisa de tempo médio de estudo diário, em minutos, envolvendo 50 alunos do Estratégia, em que os alunos foram escolhidos de maneira aleatória, não havendo qualquer organização dos valores observados. Por terem sido apresentados na forma em que foram coletados, são denominados de **dados brutos**.

| Aluno | Tempo |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1     | 143   | 11    | 113   | 21    | 170   | 31    | 124   | 41    | 105   |
| 2     | 142   | 12    | 143   | 22    | 158   | 32    | 137   | 42    | 154   |
| 3     | 161   | 13    | 159   | 23    | 123   | 33    | 153   | 43    | 99    |
| 4     | 126   | 14    | 168   | 24    | 96    | 34    | 129   | 44    | 114   |
| 5     | 134   | 15    | 123   | 25    | 98    | 35    | 148   | 45    | 161   |
| 6     | 137   | 16    | 135   | 26    | 135   | 36    | 173   | 46    | 128   |
| 7     | 171   | 17    | 135   | 27    | 129   | 37    | 126   | 47    | 175   |
| 8     | 85    | 18    | 175   | 28    | 126   | 38    | 104   | 48    | 137   |
| 9     | 155   | 19    | 115   | 29    | 103   | 39    | 157   | 49    | 165   |
| 10    | 171   | 20    | 89    | 30    | 171   | 40    | 127   | 50    | 115   |

Normalmente, esses dados fornecem pouca informação ao leitor, sendo necessário organizá-los, com o propósito de aumentar sua capacidade informativa. A simples organização dos dados em um **rol crescente** já ajuda bastante nesse sentido. Com os dados organizados em rol, conseguimos verificar que o menor tempo observado foi de 85 minutos, e o maior, de 175 minutos, o que nos fornece uma **amplitude total** ( $AT = 175 - 85 = 90$ ) de variação da ordem de 90 minutos.

| 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 vezes em que ocorre, em vez de repeti-los. O **número de ocorrências de um determinado valor recebe o nome de frequência**. 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: distribuição de frequências **pontual (ou discreta)** e distribuição de frequências **intervalar (ou contínua)**.

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

| 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     |

Nesse caso, **quando a variável é contínua**, o mais **recomendável é agrupar** os valores por **intervalos de classe**. Desse modo, em vez de listar cada um dos valores que ocorrem, utilizamos uma **distribuição de frequências intervalar**, listando os **intervalos de classe** e as **frequências correspondentes**.

| 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 |

Procedendo dessa forma, **perdemos a informação detalhada dos tempos médios**, mas ganhamos em termos de **praticidade**, o que simplifica o processo de análise de dados. Examinando a tabela acima, percebemos facilmente que a maioria dos alunos estuda diariamente entre 115 e 130 minutos, enquanto uma minoria alcança entre 175 e 190 minutos.

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 $b$ .         |

Por fim, é importante salientarmos que, em análises estatísticas, constantemente 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 de forma detalhada cada elemento de uma distribuição de frequências. Tomaremos 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, temos as seguintes classes:

| Classe    | Intervalo          | Frequência ( $f_i$ )       |
|-----------|--------------------|----------------------------|
| 1ª Classe | $85 \leq x < 100$  | 5                          |
| 2ª Classe | $100 \leq x < 115$ | 5                          |
| 3ª Classe | $115 \leq x < 130$ | 12                         |
| 4ª Classe | $130 \leq x < 145$ | 10                         |
| 5ª Classe | $145 \leq x < 160$ | 7                          |
| 6ª Classe | $160 \leq x < 175$ | 9                          |
| 7ª Classe | $175 \leq x < 190$ | 2                          |
|           |                    | <b><math>n = 50</math></b> |

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:

$$k = 1 + 3,3 \times \log n$$

Outra abordagem, utilizada quando o número de dados é menor ou igual a 50, é por meio da fórmula:

$$k = \sqrt{n}$$



Vamos aproveitar para calcular o número de classes do nosso exemplo:

a) pela fórmula de Sturges:

$$k = 1 + 3,3 \times \log n$$

$$k = 1 + 3,3 \times \log 50$$

$$k = 1 + 3,3 \times 1,7$$

$$k = 1 + 5,61 = 6,61$$

b) pela outra fórmula:

$$k = \sqrt{n}$$

$$k = \sqrt{50} = 7,07$$

## 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.**

Vamos identificar os limites inferiores e superiores do nosso exemplo:

| Classes            | Limite Inferior ( $l_{inf}$ ) | Limite Superior ( $l_{sup}$ ) |
|--------------------|-------------------------------|-------------------------------|
| $85 \leq x < 100$  | 85                            | 100                           |
| $100 \leq x < 115$ | 100                           | 115                           |
| $115 \leq x < 130$ | 115                           | 130                           |
| $130 \leq x < 145$ | 130                           | 145                           |
| $145 \leq x < 160$ | 145                           | 160                           |
| $160 \leq x < 175$ | 160                           | 175                           |
| $175 \leq x < 190$ | 175                           | 190                           |

## Amplitude de um Intervalo de Classe

A **amplitude de um intervalo de classe**, ou simplesmente **intervalo de classe**, é a **distância entre os limites inferiores (ou superiores) de classes consecutivas**. Ela é obtida pela diferença entre dois limites inferiores (ou superiores) consecutivos:

$$h = l_{sup} - l_{inf}$$

em que  $l_{inf}$  é o limite inferior do intervalo de classe e  $l_{sup}$  é o limite superior do intervalo de classe.



Dando continuidade ao nosso exemplo, vejamos como calcular as amplitudes dos intervalos:

| Classes            | Limite Inferior ( $l_{inf}$ ) | Limite Superior ( $l_{sup}$ ) | Amplitude de Classe ( $h = l_{sup} - l_{inf}$ ) |
|--------------------|-------------------------------|-------------------------------|---|
| $85 \leq x < 100$  | 85                            | 100                           | $100 - 85 = 15$                                 |
| $100 \leq x < 115$ | 100                           | 115                           | $115 - 100 = 15$                                |
| $115 \leq x < 130$ | 115                           | 130                           | $130 - 115 = 15$                                |
| $130 \leq x < 145$ | 130                           | 145                           | $145 - 130 = 15$                                |
| $145 \leq x < 160$ | 145                           | 160                           | $160 - 145 = 15$                                |
| $160 \leq x < 175$ | 160                           | 175                           | $175 - 160 = 15$                                |
| $175 \leq x < 190$ | 175                           | 190                           | $190 - 175 = 15$                                |

Embora seja desejável, a amplitude do intervalo de classe nem sempre será constante ao longo de toda a distribuição de frequências intervalar.

## 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)**. Portanto, corresponde à diferença entre o último e o primeiro elemento de um conjunto de dados ordenado de forma crescente:

$$AT = l_{m\acute{a}x} - l_{m\acute{i}n}$$

Note que, **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:

$$AT = l_{m\acute{a}x} - l_{m\acute{i}n}$$
$$AT = 190 - 85 = 105$$



## 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. Esse ponto divide a classe em duas partes iguais. Também costuma ser chamado de **marca** ou **representante da classe**.

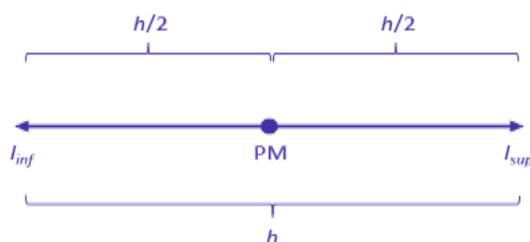
$$PM = \frac{(l_{inf} + l_{sup})}{2}$$

Para praticar, vamos calcular os pontos médios de nossa distribuição de frequências:

| Tempo médio ( $X_i$ ) | Ponto médio ( $PM_i$ )        | Frequência ( $f_i$ ) |
|-----------------------|-------------------------------|----------------------|
| $85 \leq x < 100$     | $\frac{85 + 100}{2} = 92,5$   | 5                    |
| $100 \leq x < 115$    | $\frac{100 + 115}{2} = 107,5$ | 5                    |
| $115 \leq x < 130$    | $\frac{115 + 130}{2} = 122,5$ | 12                   |
| $130 \leq x < 145$    | $\frac{130 + 145}{2} = 137,5$ | 10                   |
| $145 \leq x < 160$    | $\frac{145 + 160}{2} = 152,5$ | 7                    |
| $160 \leq x < 175$    | $\frac{160 + 175}{2} = 167,5$ | 9                    |
| $175 \leq x < 190$    | $\frac{175 + 190}{2} = 182,5$ | 2                    |

Veja que os pontos médios formaram uma progressão aritmética, pois a diferença entre dois pontos médios consecutivos foi constante e igual a 15. Isso ocorreu porque o intervalo de classe,  $h$ , também foi constante (e igual a 15) em toda a distribuição. Assim, **quando o intervalo de classes é constante, a diferença entre os pontos médios também será constante e igual ao intervalo de classe**.

Adicionalmente, sabendo que o ponto médio divide a classe em duas partes iguais, podemos derivar outras relações envolvendo o próprio ponto médio, a amplitude de classe e os limites inferior e superior.



Dada a figura anterior, podemos obter os limites de uma classe por meio das seguintes expressões:

$$l_{inf} = PM - \frac{h}{2}$$

$$l_{sup} = PM + \frac{h}{2}$$

Além disso, podemos encontrar o ponto médio de uma classe a partir das seguintes relações:

$$PM = l_{inf} + \frac{h}{2}$$

$$PM = l_{sup} - \frac{h}{2}$$

## 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. Nesse contexto, é importante sabermos que existem quatro tipos de frequência, os quais serão analisados nas subseções seguintes:

- frequência absoluta simples ( $f_i$ );
- frequência absoluta acumulada ( $f_{ac}$ );
- frequência relativa simples ( $F_i$ );
- 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.



| i | Tempos (min) | Frequência ( $f_i$ ) |
|---|--------------|----------------------|
| 1 | 85 - 100     | 5                    |
| 2 | 100 - 115    | 5                    |
| 3 | 115 - 130    | 12                   |
| 4 | 130 - 145    | 10                   |
| 5 | 145 - 160    | 7                    |
| 6 | 160 - 175    | 9                    |
| 7 | 175 - 190    | 2                    |

A frequência simples é simbolizada por  $f_i$ . No exemplo anterior, temos:  $f_1 = 5$ ,  $f_2 = 5$ ,  $f_3 = 12$ ,  $f_4 = 10$ ,  $f_5 = 7$ ,  $f_6 = 9$  e  $f_7 = 2$ .

A soma de todas as frequências é igual ao número total de dados analisados:

$$\sum_{i=1}^k f_i = n$$

em que a notação  $\sum_{i=1}^k f_i$  representa o somatório das frequências de cada uma das  $k$  classes.

Para a distribuição em análise, temos:

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

Agora, podemos incluir essa informação na representação tabular:

| i | Tempos (min) | Frequência ( $f_i$ ) |
|---|--------------|----------------------|
| 1 | 85 - 100     | 5                    |
| 2 | 100 - 115    | 5                    |
| 3 | 115 - 130    | 12                   |
| 4 | 130 - 145    | 10                   |
| 5 | 145 - 160    | 7                    |
| 6 | 160 - 175    | 9                    |
| 7 | 175 - 190    | 2                    |

$$\sum_{i=1}^7 f_i = 50$$


## 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. No exemplo apresentado anteriormente, 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).

$$f_{ac_i} = f_1 + f_2 + f_3 + \dots + f_i$$

A **frequência absoluta acumulada crescente** é calculada de cima para baixo, da seguinte forma:

- 1) repetimos a frequência absoluta da **primeira** classe;
- 2) os demais valores da frequência absoluta são obtidos a partir da soma da frequência acumulada anterior com a frequência absoluta da classe correspondente;
- 3) a frequência acumulada crescente sempre termina com o valor de  $n$ .

| i | Tempos (min) | Frequência ( $f_i$ ) | Frequência Acumulada |
|---|--------------|----------------------|----------------------|
| 1 | 85 † 100     | 5                    | 5                    |
| 2 | 100 † 115    | 5                    | 10                   |
| 3 | 115 † 130    | 12                   | 22                   |
| 4 | 130 † 145    | 10                   | 32                   |
| 5 | 145 † 160    | 7                    | 39                   |
| 6 | 160 † 175    | 9                    | 48                   |
| 7 | 175 † 190    | 2                    | 50                   |

$$\sum_{i=1}^7 f_i = 50$$

A **frequência absoluta acumulada decrescente** ( $f_{ad}$ ) é a soma das frequências de todos os valores superiores ao limite inferior do intervalo de uma determinada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é:  $f_{ad_4} = f_4 + f_5 + f_6 + f_7 = 10 + 7 + 9 + 2 = 28$ , significando que 28 alunos estudam por um período igual ou superior a 130 minutos (limite inferior da quarta classe) e inferior a 190 minutos.

$$f_{ad_i} = f_i + f_{i+1} + f_{i+2} + \dots + f_k$$



A **frequência absoluta acumulada decrescente** é calculada de baixo para cima, da seguinte forma:

- 1) repetimos a frequência absoluta da **última** classe;
- 2) os demais valores da frequência acumulada decrescente são obtidos a partir da soma da frequência acumulada anterior com a frequência absoluta da classe correspondente;
- 3) a frequência acumulada decrescente sempre termina com o valor de  $n$ .

| $i$ | Tempos (min) | Frequência ( $f_i$ ) | Frequência Acumulada |
|-----|--------------|----------------------|----------------------|
| 1   | 85 † 100     | 5                    | 50                   |
| 2   | 100 † 115    | 5                    | 45                   |
| 3   | 115 † 130    | 12                   | 40                   |
| 4   | 130 † 145    | 10                   | 28                   |
| 5   | 145 † 160    | 7                    | 18                   |
| 6   | 160 † 175    | 9                    | 11                   |
| 7   | 175 † 190    | 2                    | 2                    |

$$\sum_{i=1}^7 f_i = 50$$

## 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, as frequências relativas são:

| $i$ | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Frequência Relativa ( $F_i$ ) |
|-----|--------------|-------------------------------|-------------------------------|
| 1   | 85 † 100     | 5                             | $F_1 = \frac{5}{50} = 0,10$   |
| 2   | 100 † 115    | 5                             | $F_2 = \frac{5}{50} = 0,10$   |
| 3   | 115 † 130    | 12                            | $F_3 = \frac{12}{50} = 0,24$  |



|   |           |    |                              |
|---|-----------|----|------------------------------|
| 4 | 130 † 145 | 10 | $F_4 = \frac{10}{50} = 0,20$ |
| 5 | 145 † 160 | 7  | $F_5 = \frac{7}{50} = 0,14$  |
| 6 | 160 † 175 | 9  | $F_6 = \frac{9}{50} = 0,18$  |
| 7 | 175 † 190 | 2  | $F_7 = \frac{2}{50} = 0,04$  |

Para representar esses valores em termos de porcentagem, basta multiplicarmos por 100%. A tabela ficaria assim:

| i | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Frequência Relativa ( $F_i$ )             |
|---|--------------|-------------------------------|---|
| 1 | 85 † 100     | 5                             | $F_1 = \frac{5}{50} \times 100\% = 10\%$  |
| 2 | 100 † 115    | 5                             | $F_2 = \frac{5}{50} \times 100\% = 10\%$  |
| 3 | 115 † 130    | 12                            | $F_3 = \frac{12}{50} \times 100\% = 24\%$ |
| 4 | 130 † 145    | 10                            | $F_4 = \frac{10}{50} \times 100\% = 20\%$ |
| 5 | 145 † 160    | 7                             | $F_5 = \frac{7}{50} \times 100\% = 14\%$  |
| 6 | 160 † 175    | 9                             | $F_6 = \frac{9}{50} \times 100\% = 18\%$  |
| 7 | 175 † 190    | 2                             | $F_7 = \frac{2}{50} \times 100\% = 4\%$   |

O propósito das frequências relativas é facilitar a realização de comparações de classes individuais com o total das observações. Na tabela anterior, por exemplo, conseguimos verificar facilmente que 24% das observações pertencem à quarta classe e que 18% das observações pertencem à sexta classe.

Repare que a soma de todas as frequências relativas deve ser igual a 100%:

$$\sum_{i=1}^k F_i = 100\%$$

Podemos incluir essa informação na representação tabular:

| i | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Frequência Relativa ( $F_i$ ) |
|---|--------------|-------------------------------|-------------------------------|
| 1 | 85 † 100     | 5                             | 10%                           |
| 2 | 100 † 115    | 5                             | 10%                           |
| 3 | 115 † 130    | 12                            | 24%                           |



|   |           |                         |                            |
|---|-----------|-------------------------|----------------------------|
| 4 | 130 † 145 | 10                      | 20%                        |
| 5 | 145 † 160 | 7                       | 14%                        |
| 6 | 160 † 175 | 9                       | 18%                        |
| 7 | 175 † 190 | 2                       | 4%                         |
|   |           | $\sum_{i=1}^7 f_i = 50$ | $\sum_{i=1}^7 F_i = 100\%$ |

## 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. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é:  $F_{ac4} = F_1 + F_2 + F_3 + F_4 = 10\% + 10\% + 24\% + 20\% = 64\%$ , 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).

$$F_{ac_i} = F_1 + F_2 + F_3 + \dots + F_i$$

A **frequência relativa acumulada crescente** é calculada de cima para baixo, da seguinte forma:

- 1) repetimos a frequência relativa da **primeira** classe;
- 2) os demais valores são obtidos a partir da soma da frequência relativa acumulada anterior com a frequência relativa da classe correspondente;
- 3) a frequência relativa acumulada sempre termina com o valor de 100%.

| i | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Frequência Relativa ( $F_i$ ) | Frequência Acumulada ( $F_{ac}$ ) |
|---|--------------|-------------------------------|-------------------------------|-----------------------------------|
| 1 | 85 † 100     | 5                             | 10%                           | 10%                               |
| 2 | 100 † 115    | 5                             | 10%                           | 20%                               |
| 3 | 115 † 130    | 12                            | 24%                           | 44%                               |
| 4 | 130 † 145    | 10                            | 20%                           | 64%                               |
| 5 | 145 † 160    | 7                             | 14%                           | 78%                               |
| 6 | 160 † 175    | 9                             | 18%                           | 96%                               |
| 7 | 175 † 190    | 2                             | 4%                            | 100%                              |
|   |              | $\sum_{i=1}^7 f_i = 50$       | $\sum_{i=1}^7 F_i = 100\%$    |                                   |

A **frequência relativa acumulada decrescente** ( $F_{ad}$ ) a proporção de valores superiores ao limite inferior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência



acumulada correspondente à quarta classe é:  $F_{ad_4} = F_4 + F_5 + F_6 + F_7 = 20\% + 14\% + 18\% + 4\% = 56\%$ , significando que 56% dos alunos estudam por um período igual ou superior a 130 minutos (limite inferior da quarta classe) e inferior a 190 minutos.

$$F_{ad_i} = F_i + F_{i+1} + F_{i+2} + \dots + F_k$$

A **frequência relativa acumulada decrescente** é calculada de baixo para cima, da seguinte forma:

- 1) repetimos a frequência relativa da **última** classe;
- 2) os demais valores são obtidos a partir da soma da frequência acumulada anterior com a frequência relativa da classe correspondente;
- 3) a frequência acumulada sempre termina com o valor de 100%.

| i | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Frequência Relativa ( $F_i$ ) | Frequência Acumulada ( $F_{ad}$ ) |
|---|--------------|-------------------------------|-------------------------------|-----------------------------------|
| 1 | 85 - 100     | 5                             | 10%                           | 100%                              |
| 2 | 100 - 115    | 5                             | 10%                           | 90%                               |
| 3 | 115 - 130    | 12                            | 24%                           | 80%                               |
| 4 | 130 - 145    | 10                            | 20%                           | 56%                               |
| 5 | 145 - 160    | 7                             | 14%                           | 36%                               |
| 6 | 160 - 175    | 9                             | 18%                           | 22%                               |
| 7 | 175 - 190    | 2                             | 4%                            | 4%                                |

$\sum_{i=1}^7 f_i = 50$        $\sum_{i=1}^7 F_i = 100\%$

## Densidade de Frequência

A densidade de frequência de uma classe consiste no quociente entre a frequência da classe (absoluta ou relativa) e sua amplitude:

$$densidade = \frac{\text{freq. absoluta ou relativa}}{\text{amplitude}}$$

$$d_i = \frac{f_i}{h_i}$$



Para o nosso exemplo, as densidades de frequência são:

| i | Tempos (min) | Frequência Absoluta ( $f_i$ ) | Densidade de Frequência (d)                    |
|---|--------------|-------------------------------|--|
| 1 | 85 - 100     | 5                             | $d_1 = \frac{f_1}{h_1} = \frac{5}{15} = 0,33$  |
| 2 | 100 - 115    | 5                             | $d_2 = \frac{f_2}{h_2} = \frac{5}{15} = 0,33$  |
| 3 | 115 - 130    | 12                            | $d_3 = \frac{f_3}{h_3} = \frac{12}{15} = 0,80$ |
| 4 | 130 - 145    | 10                            | $d_4 = \frac{f_4}{h_4} = \frac{10}{15} = 0,66$ |
| 5 | 145 - 160    | 7                             | $d_5 = \frac{f_5}{h_5} = \frac{7}{15} = 0,46$  |
| 6 | 160 - 175    | 9                             | $d_6 = \frac{f_6}{h_6} = \frac{9}{15} = 0,60$  |
| 7 | 175 - 190    | 2                             | $d_7 = \frac{f_7}{h_7} = \frac{2}{15} = 0,13$  |



| Item                                | Definição  | Símbolos e Fórmulas                                 |
|-------------------------------------|--|---|
| Número de Classes                   | As classes são os intervalos nos quais o fenômeno é subdividido.   | $k = 1 + 3,3 \times \log n$<br>ou<br>$k = \sqrt{n}$ |
| Limites de Classe                   | Correspondem aos valores extremos.   | $l_{inf}$ e $l_{sup}$                               |
| Amplitude de um Intervalo de Classe | Distância entre os limites inferiores (ou superiores) de classes consecutivas.   | $h = l_{sup} - l_{inf}$                             |
| Amplitude total                     | Diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo). | $AT = l_{máx} - l_{mín}$<br>$AT = h \times k$       |



|                                      |   |  |
|--------------------------------------|---|--|
| <b>Ponto Médio</b>                   | Média aritmética simples dos valores extremos de uma classe.  | $PM = \frac{(l_{inf} + l_{sup})}{2}$ $PM = l_{inf} + \frac{h}{2}$ $PM = l_{sup} - \frac{h}{2}$ |
| <b>Frequência Absoluta Simples</b>   | Número de observações correspondentes a uma determinada classe ou a um determinado valor.               | $f_i$  |
| <b>Frequência Absoluta Acumulada</b> | Total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe | $f_{ac_i} = f_1 + f_2 + f_3 + \dots + f_i$   |
| <b>Frequência Relativa Simples</b>   | Proporção de dados existentes em uma determinada classe.  | $F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$   |
| <b>Frequência Relativa Acumulada</b> | Proporção de valores inferiores ao limite superior do intervalo de uma dada classe.                     | $F_{ac_i} = F_1 + F_2 + F_3 + \dots + F_i$   |
| <b>Densidade de Frequência</b>       | Quociente entre a frequência da classe (absoluta ou relativa) e sua amplitude                           | $d = \frac{f}{h}$  |



(CESGRANRIO/CAIXA/2024) Uma pesquisa foi realizada com todos os 200 funcionários de uma empresa. Uma das variáveis a ser investigada é o número de dependentes por funcionário. Os dados sobre essa variável foram organizados e estão resumidos na Tabela a seguir.

Número de dependentes por funcionários



| Número de dependentes | Frequência absoluta |
|-----------------------|---------------------|
| 0                     | 20                  |
| 1                     | 40                  |
| 2                     | 50                  |
| 3                     | 60                  |
| 4                     | 25                  |
| 5                     | 5                   |

Considerando-se essas informações, o percentual de funcionários dessa empresa que têm até 2 dependentes, inclusive, é igual a

- a) 60%
- b) 55%
- c) 50%
- d) 30%
- e) 25%

#### Comentários:

Para calcular o percentual de funcionários que têm até 2 dependentes, precisamos somar as frequências absolutas dos funcionários que têm 0, 1 ou 2 dependentes e, em seguida, dividir pelo total de funcionários. Vamos primeiro somar as frequências absolutas:

$$\text{frequência de 0 dependentes} = 20$$

$$\text{frequência de 1 dependente} = 40$$

$$\text{frequência de 2 dependentes} = 50$$

Total de funcionários com até 2 dependentes:

$$20 + 40 + 50 = 110$$

Como o total de funcionários da empresa é 200, podemos calcular o percentual:

$$\text{Percentual} = \left(\frac{110}{200}\right) \times 100\% = 55\%$$

Portanto, o percentual de funcionários dessa empresa que têm até 2 dependentes, inclusive, é igual a 55%.

**Gabarito: B.**



(VUNESP/IPSM-SJC/2018) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão.

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

5 4 6 1 2 5 3 1 3 3

4 4 1 5 5 6 1 2 5 1

3 4 5 1 1 6 6 2 1 1

4 4 4 3 4 3 2 2 2 3

6 6 3 2 4 2 6 6 2 1

A frequência simples relativa do primeiro elemento é

- a) 10%.
- b) 20%.
- c) 50.
- d) 7.
- e) 20.

#### Comentários:

A frequência absoluta é o número de vezes em que uma determinada variável assume um valor, enquanto a frequência relativa é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado).

Para os dados apresentados, temos que:

| Valor obtido no lançamento | Frequência absoluta | Frequência relativa  |
|----------------------------|---------------------|----------------------|
| 1                          | 10                  | $10/50 = 0,2 = 20\%$ |
| 2                          | 9                   | $9/50 = 0,18 = 18\%$ |
| 3                          | 8                   | $8/50 = 0,16 = 16\%$ |
| 4                          | 9                   | $9/50 = 0,18 = 18\%$ |
| 5                          | 6                   | $6/50 = 0,12 = 12\%$ |
| 6                          | 8                   | $8/50 = 0,16 = 16\%$ |
| Total                      | 50                  | 100%                 |

Portanto, a frequência relativa do primeiro elemento é igual a 20%.

**Gabarito: B.**

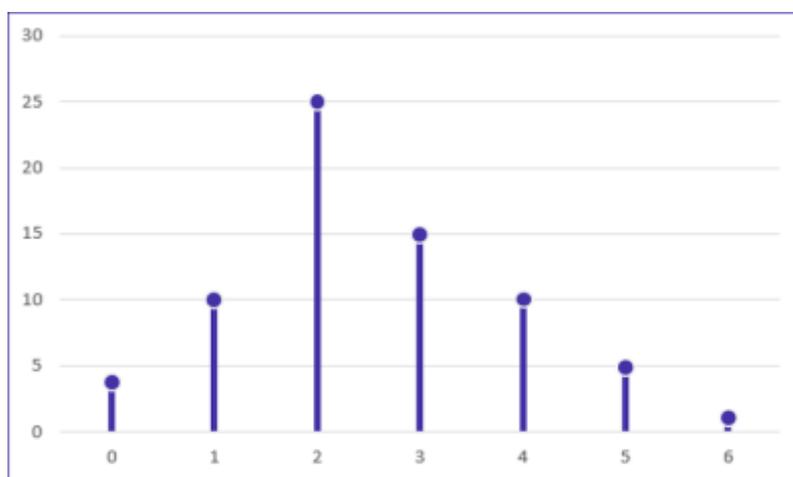


## REPRESENTAÇÕES GRÁFICAS DAS DISTRIBUIÇÕES DE FREQUÊNCIA

### 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

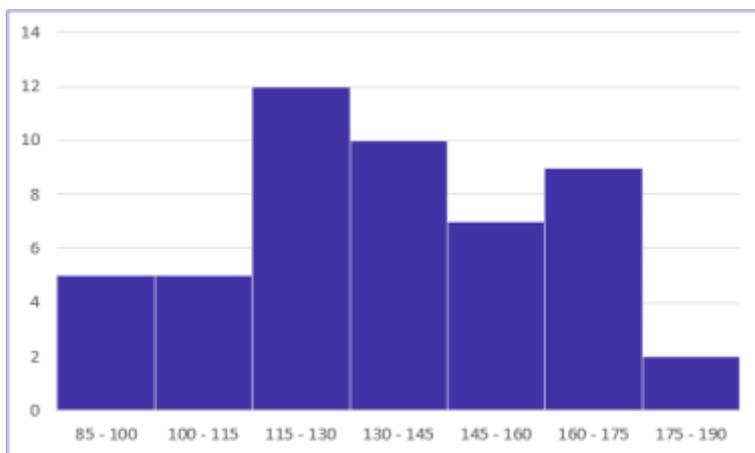
O **histograma** é um gráfico destinado a representar **dados agrupados em classe**, sendo composto por um conjunto de **retângulos contíguos (justapostos)** cujas bases ficam localizadas sobre o eixo horizontal (eixo x), de forma que os seus pontos médios devem coincidir com os pontos médios dos intervalos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



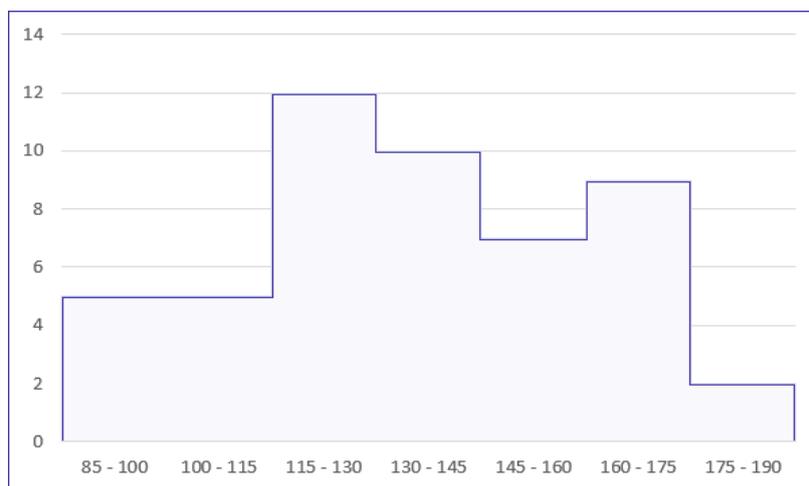
precisa ser proporcional à frequência do intervalo de classe. Além disso, a área do histograma é proporcional ao somatório das frequências.

| 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 básica entre um **histograma** e um **gráfico de colunas** (estudaremos na próxima seção) é a separação entre os retângulos adjacentes. **Veja que não existe separação entre os retângulos no caso do histograma.**

Dito isso, é importante mencionarmos a existência do gráfico denominado de **poligonal 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.

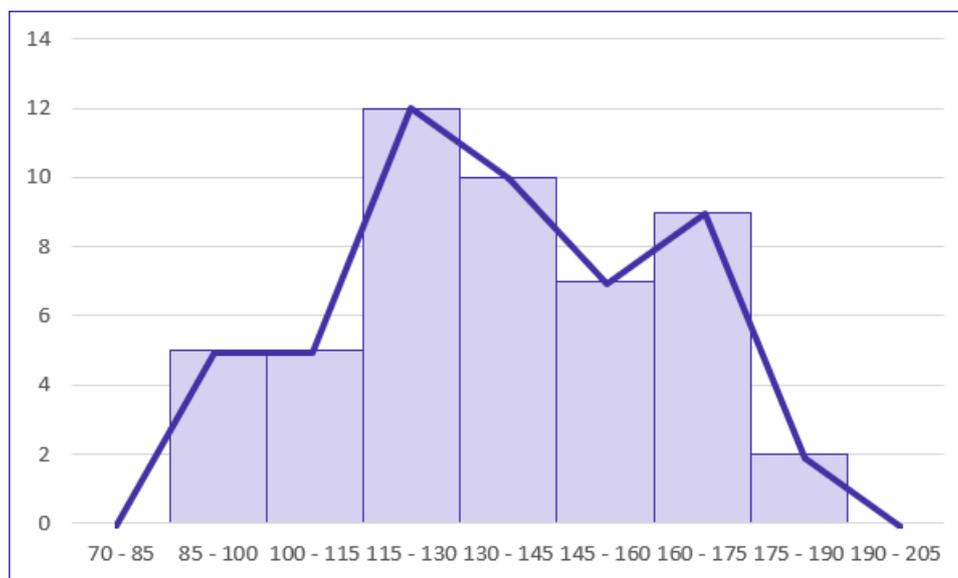


## Polígono de Frequências

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:

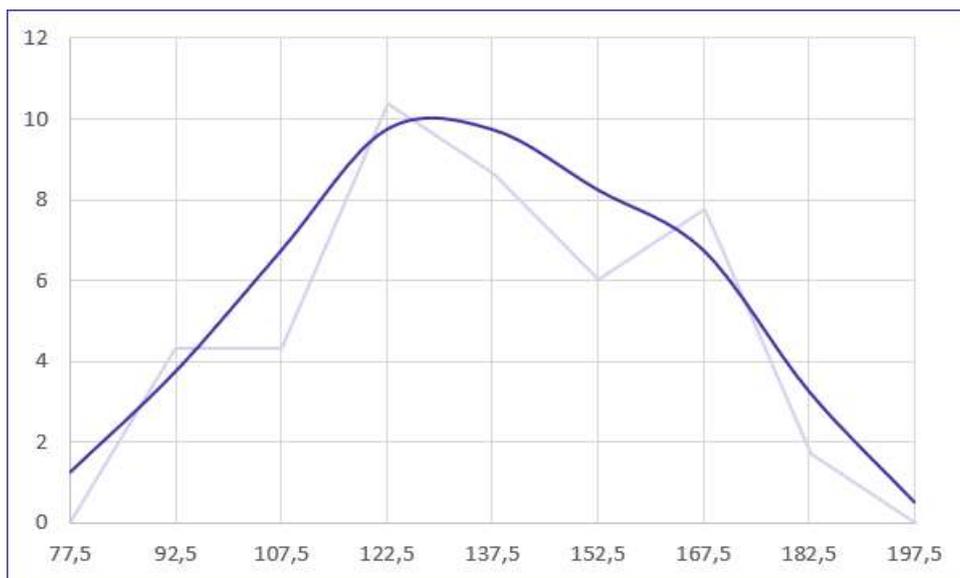
$$f_{c_i} = \frac{f_{ant} + 2 \times f_i + f_{post}}{4}$$

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:

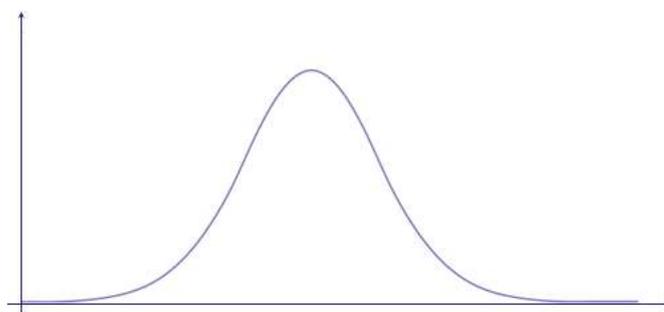
| Tempo médio ( $X_i$ ) | Ponto médio ( $PM_i$ ) | Frequência ( $f_i$ ) | Frequência Calculada ( $f_{c_i}$ )                |
|-----------------------|------------------------|----------------------|---|
| $70 \leq x < 85$      | 77,5                   | 0                    | $f_{c_0} = \frac{0 + 2 \times 0 + 5}{4} = 1,25$   |
| $85 \leq x < 100$     | 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                    | $f_{c_8} = \frac{2 + 2 \times 0 + 0}{4} = 0,50$   |



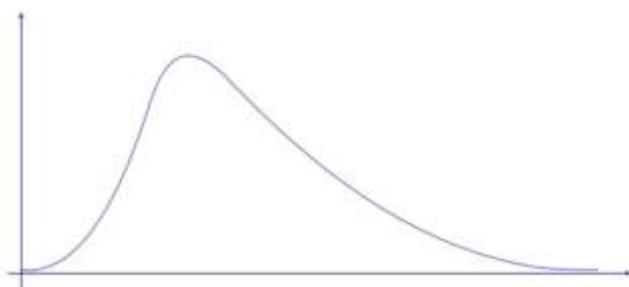


As curvas de frequências podem assumir as seguintes formas características:

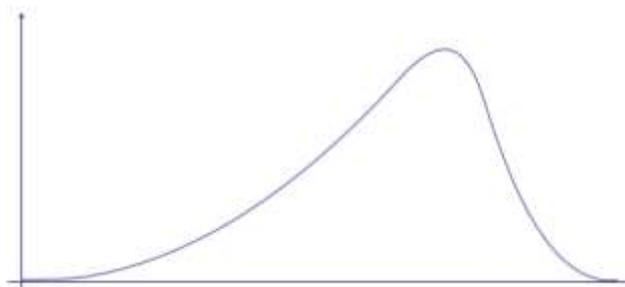
a) **curvas em forma de sino**: são curvas que apresentam concentração de valores em torno da região central da distribuição. Tais curvas podem ser simétricas ou assimétricas. Quando assimétricas, as curvas ainda podem apresentar uma cauda mais alongada à esquerda (assimetria à esquerda) ou mais alongada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



Curva Simétrica.



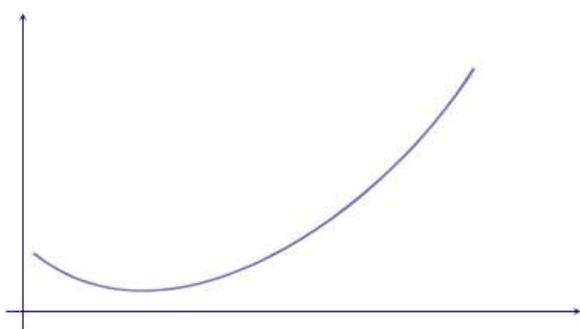
Curva assimétrica à direita.



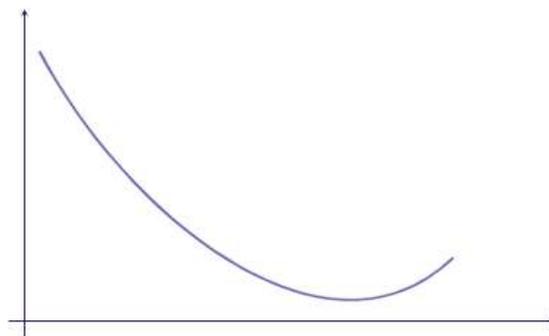
Curva assimétrica à esquerda.



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:

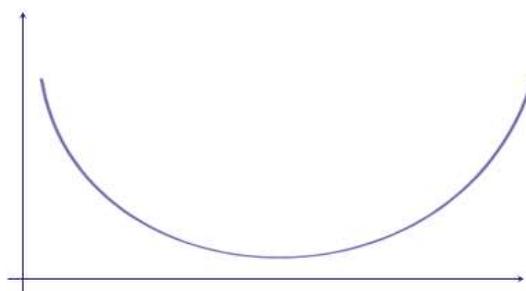


Curva em formato de J.



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.

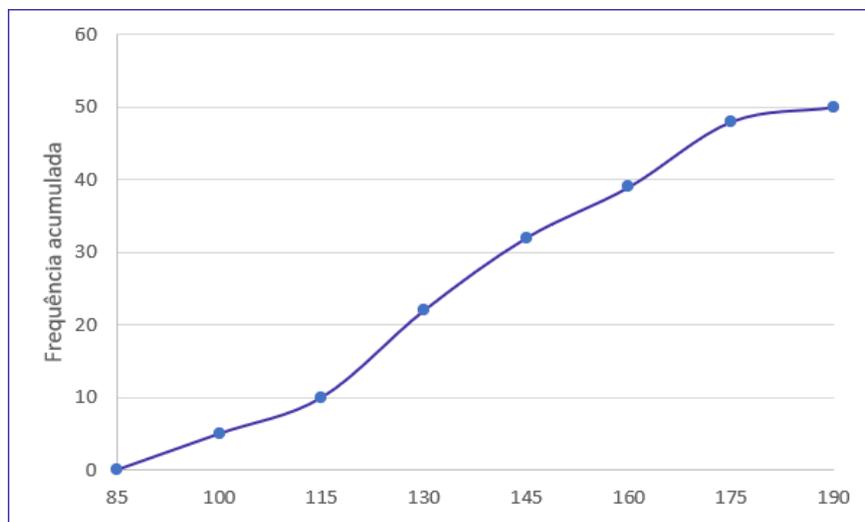
## Ogiva (Polígono de Frequências Acumuladas)

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.



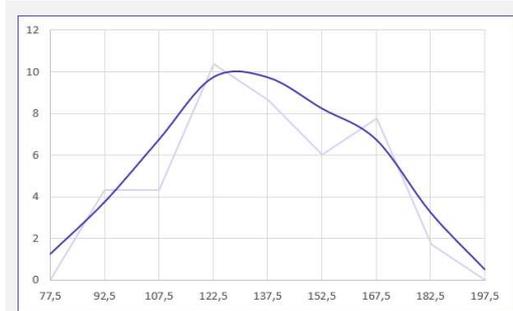
| Tempo médio ( $X_i$ ) | Frequência ( $f_i$ ) | Frequência Acumulada ( $f_{ac}$ ) |
|-----------------------|----------------------|-----------------------------------|
| $70 \leq x < 85$      | 0                    | 0                                 |
| $85 \leq x < 100$     | 5                    | 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                                |



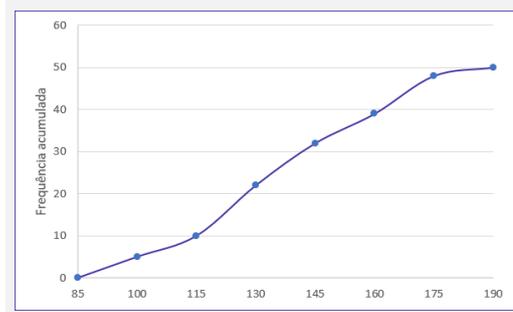


| Gráfico   | Definição  |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
|---|------------|------------|--------|---|---------|----|---------|----|---------|----|---------|----|---------|---|---------|---|--|---|---------|---|---|
| <table border="1"><thead><tr><th>Classe</th><th>Frequência</th></tr></thead><tbody><tr><td>0</td><td>4</td></tr><tr><td>1</td><td>10</td></tr><tr><td>2</td><td>25</td></tr><tr><td>3</td><td>15</td></tr><tr><td>4</td><td>10</td></tr><tr><td>5</td><td>5</td></tr><tr><td>6</td><td>1</td></tr></tbody></table>  | Classe     | Frequência | 0      | 4 | 1       | 10 | 2       | 25 | 3       | 15 | 4       | 10 | 5       | 5 | 6       | 1 | <p>O <b>gráfico de hastes ou bastões</b> é muito utilizado para representar dados não agrupados em classes, o que normalmente ocorre com <b>dados discretos</b>.</p>           |   |         |   |   |
| Classe  | Frequência |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 0   | 4          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 1   | 10         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 2   | 25         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 3   | 15         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 4   | 10         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 5   | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 6   | 1          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| <table border="1"><thead><tr><th>Classe</th><th>Frequência</th></tr></thead><tbody><tr><td>85-100</td><td>5</td></tr><tr><td>100-115</td><td>5</td></tr><tr><td>115-130</td><td>12</td></tr><tr><td>130-145</td><td>10</td></tr><tr><td>145-160</td><td>7</td></tr><tr><td>160-175</td><td>9</td></tr><tr><td>175-190</td><td>2</td></tr></tbody></table>   | Classe     | Frequência | 85-100 | 5 | 100-115 | 5  | 115-130 | 12 | 130-145 | 10 | 145-160 | 7  | 160-175 | 9 | 175-190 | 2 | <p>O <b>histograma</b> é um gráfico destinado a representar <b>dados agrupados em classe</b>, sendo composto por um conjunto de <b>retângulos contíguos (justapostos)</b>.</p> |   |         |   |   |
| Classe  | Frequência |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 85-100  | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 100-115   | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 115-130   | 12         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 130-145   | 10         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 145-160   | 7          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 160-175   | 9          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 175-190   | 2          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| <table border="1"><thead><tr><th>Classe</th><th>Frequência</th></tr></thead><tbody><tr><td>85-100</td><td>5</td></tr><tr><td>100-115</td><td>5</td></tr><tr><td>115-130</td><td>12</td></tr><tr><td>130-145</td><td>10</td></tr><tr><td>145-160</td><td>7</td></tr><tr><td>160-175</td><td>9</td></tr><tr><td>175-190</td><td>2</td></tr></tbody></table>   | Classe     | Frequência | 85-100 | 5 | 100-115 | 5  | 115-130 | 12 | 130-145 | 10 | 145-160 | 7  | 160-175 | 9 | 175-190 | 2 | <p>A <b>poligonal característica</b> é construída utilizando apenas os <b>contornos do histograma</b>.</p>   |   |         |   |   |
| Classe  | Frequência |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 85-100  | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 100-115   | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 115-130   | 12         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 130-145   | 10         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 145-160   | 7          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 160-175   | 9          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 175-190   | 2          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| <table border="1"><thead><tr><th>Classe</th><th>Frequência</th></tr></thead><tbody><tr><td>70-85</td><td>0</td></tr><tr><td>85-100</td><td>5</td></tr><tr><td>100-115</td><td>5</td></tr><tr><td>115-130</td><td>12</td></tr><tr><td>130-145</td><td>10</td></tr><tr><td>145-160</td><td>7</td></tr><tr><td>160-175</td><td>9</td></tr><tr><td>175-190</td><td>2</td></tr><tr><td>190-205</td><td>0</td></tr></tbody></table> | Classe     | Frequência | 70-85  | 0 | 85-100  | 5  | 100-115 | 5  | 115-130 | 12 | 130-145 | 10 | 145-160 | 7 | 160-175 | 9 | 175-190  | 2 | 190-205 | 0 | <p>O <b>polígono de frequências</b> é um gráfico em linha <b>obtido por meio da ligação, por segmentos de reta, dos pontos médios</b> das bases superiores dos retângulos de um histograma.</p> |
| Classe  | Frequência |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 70-85   | 0          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 85-100  | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 100-115   | 5          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 115-130   | 12         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 130-145   | 10         |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 145-160   | 7          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 160-175   | 9          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 175-190   | 2          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |
| 190-205   | 0          |            |        |   |         |    |         |    |         |    |         |    |         |   |         |   |  |   |         |   |   |





A curva de frequências é obtida a partir do **polimento de um polígono de frequências**.



O gráfico de ogiva é empregado na **representação de distribuições de frequências acumuladas**, sejam elas crescentes ou decrescentes



(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

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

Comentários:

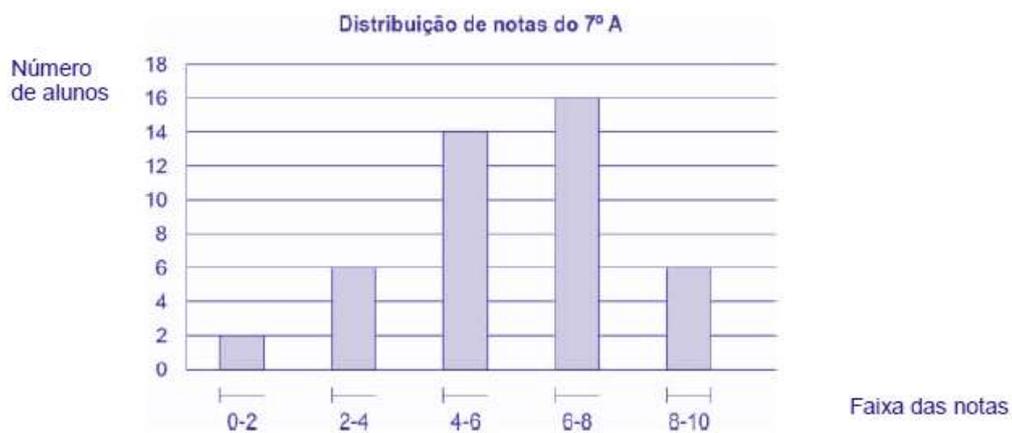


O histograma é um gráfico destinado a representar dados agrupados em classe, sendo composto por um conjunto de retângulos contíguos (justapostos) cujas bases ficam localizadas sobre o eixo horizontal (eixo x), de forma que os seus pontos médios devem coincidir com os pontos médios dos intervalos de classe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

(FCC/SEDU-ES/2018) Depois da aplicação de uma prova para todos os alunos do 7º A, a professora Marli fez um histograma com a distribuição das notas, conforme indicado abaixo.



Se a média de notas da sala foi igual a seis, a porcentagem dos alunos que tiraram nota maior ou igual a média da sala nessa prova foi de

- a) 54%.
- b) 50%.
- c) 56%.
- d) 60%.
- e) 52%.

**Comentários:**

Vamos, inicialmente, calcular o total de alunos:



$$2 + 6 + 14 + 16 + 6 = 44$$

Reparem na notação intervalar adotada para as colunas, vamos relembrar o significado dela:

| Tipo de Intervalo                               | Notação matemática | Notação estatística | Significado  |
|---|--------------------|---------------------|--|
| 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 fechado                               | $a \leq x \leq b$  | $a \dashv b$        | Engloba todos os elementos entre $a$ e $b$ , inclusive $a$ e $b$ .       |

Agora, com base nas informações do gráfico, percebemos que existem 22 pessoas com média maior ou igual a 6, representadas pelas duas últimas colunas do gráfico. Assim:

$$16 + 6 = 22$$

Logo, a porcentagem será de:

$$\frac{22}{44} = 0,5 = 50\%$$

**Gabarito: B.**



## 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.**



**2. (CESGRANRIO/UNIRIO/2019) As pesquisas da área da saúde se utilizam dos estudos da bioestatística. Nesses estudos, um dos conceitos estatísticos fundamentais é o de população, que se constitui em**

- a) conjunto de elementos a serem observados.
- b) valor obtido com maior frequência.
- c) resultado do sorteio aleatório para definir o que será estudado.
- d) dado aritmético resultante da raiz quadrada de cálculos logarítmicos.
- e) subconjunto de dados numéricos que se pretendem ordenar.

**Comentários:**

Uma população é um 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.

**Gabarito: A.**



## QUESTÕES COMENTADAS – CESGRANRIO

### Variáveis Estatísticas

1. (CESGRANRIO/BNDES/2024) O uso de dados quantitativos e qualitativos é fundamental para a compreensão e a análise da situação de um país, servindo de importante subsídio à tomada de decisões. Esses dados são mensurados a partir de diferentes variáveis. Sobre essas variáveis, verifica-se que

- a) estado civil, grau de instrução, região e salários são variáveis qualitativas.
- b) salários, idade, grau de instrução e estado civil são variáveis quantitativas.
- c) número de filhos é uma variável quantitativa discreta.
- d) grau de instrução é uma variável qualitativa nominal.
- e) região é uma variável quantitativa discreta.

#### Comentários:

A questão pede para determinar qual alternativa descreve corretamente as variáveis em termos de serem qualitativas ou quantitativas, e se são discretas ou contínuas. Vamos analisar cada opção:

Alternativa A: **Incorreta**. Estado civil, grau de instrução, região e salários são variáveis qualitativas.

- Estado civil: Variável qualitativa nominal.
- Grau de instrução: Variável qualitativa ordinal (pois pode ser ordenada, por exemplo, ensino fundamental, médio, superior).
- Região: Variável qualitativa nominal.
- Salários: Variável quantitativa contínua (salários são valores numéricos que podem assumir qualquer valor dentro de um intervalo).

Esta opção está incorreta porque salários são variáveis quantitativas, não qualitativas.

Alternativa B: **Incorreta**. Salários, idade, grau de instrução e estado civil são variáveis quantitativas.

- Salários: Variável quantitativa contínua.
- Idade: Variável quantitativa contínua.
- Grau de instrução: Variável qualitativa ordinal.
- Estado civil: Variável qualitativa nominal.

Esta opção está incorreta porque grau de instrução e estado civil são variáveis qualitativas, não quantitativas.



Alternativa C: **Correta**. Número de filhos é uma variável quantitativa discreta.

- Número de filhos: Variável quantitativa discreta (pois representa contagens inteiras que não podem ser fracionadas).

Alternativa D: **Incorreta**. Grau de instrução é uma variável qualitativa nominal.

- Grau de instrução: Variável qualitativa ordinal (possui uma ordem ou hierarquia natural).

Esta opção está incorreta porque grau de instrução é uma variável ordinal, não nominal.

Alternativa E: **Incorreta**. Região é uma variável quantitativa discreta.

- Região: Variável qualitativa nominal (representa categorias sem ordem ou valor numérico intrínseco).

Esta opção está incorreta porque região é uma variável qualitativa, não quantitativa.

**Gabarito: C.**

**2. (CESGRANRIO/CEF/2024) Instituições financeiras realizam empréstimos na expectativa de receberem de volta o valor do empréstimo acrescido de juros. Nem sempre, todavia, o cliente paga ao banco o valor que lhe é devido, e isso se denomina inadimplência. A probabilidade de um cliente inadimplir varia de acordo com as características do cliente, como renda, idade ou se o cliente possui casa própria. Quando um cliente solicita um empréstimo, um modelo matemático utiliza essas características para estimar a probabilidade de o cliente inadimplir. Essa probabilidade estimada será usada para classificar o cliente conforme o seu risco de crédito. Esse risco pode variar de AA (para clientes cuja probabilidade de inadimplir é quase nula), passando por B, C, D, E, F, G e finalmente H, à medida que a probabilidade de inadimplência vai aumentando. Segundo a apresentação de resultados do terceiro trimestre de 2023, a Caixa Econômica Federal tem 92,7% de sua carteira com risco de crédito entre AA e C e encerrou 2023 com uma inadimplência de 2,16%.**

**A variável risco de crédito descrita acima é uma variável**

- a) escalar
- b) discreta
- c) contínua
- d) nominal
- e) ordinal

**Comentários:**

A variável "risco de crédito" descrita é uma variável qualitativa ordinal. Isso ocorre porque os clientes são classificados em categorias (AA, B, C, D, E, F, G, H) que representam níveis de risco de crédito baseados na



probabilidade de inadimplência. Essas categorias têm uma ordem natural, indo de AA (risco quase nulo) até H (risco mais elevado), refletindo um aumento na probabilidade de inadimplência. Portanto, apesar de as categorias não serem numéricas, elas possuem uma hierarquia que permite comparações de magnitude entre elas.

**Gabarito: E.**

**3. (CESGRANRIO/IPEA/2024) Um instituto de pesquisa está planejando uma pesquisa em que uma das perguntas diz respeito à escolaridade do entrevistado. A pergunta admite os valores 1, 2, 3, 4 e 5, abrangendo desde ensino fundamental incompleto (1) até ensino superior completo (5).**

**Essa é uma variável**

- a) discreta
- b) contínua
- c) ordinal
- d) nominal
- e) natural

**Comentários:**

A escolaridade é uma variável qualitativa ordinal. Apesar de utilizar números (1 a 5), esses valores representam categorias com uma ordem ou hierarquia (do ensino fundamental incompleto ao ensino superior completo). Os números servem apenas para indicar a posição na escala de escolaridade e não para realizar operações aritméticas.

**Gabarito: C.**

**4. (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:

- 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);
- 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

### Distribuições de Frequência

1. (CESGRANRIO/CAIXA/2024) Uma pesquisa foi realizada com todos os 200 funcionários de uma empresa. Uma das variáveis a ser investigada é o número de dependentes por funcionário. Os dados sobre essa variável foram organizados e estão resumidos na Tabela a seguir.

Número de dependentes por funcionários

| Número de dependentes | Frequência absoluta |
|-----------------------|---------------------|
| 0                     | 20                  |
| 1                     | 40                  |
| 2                     | 50                  |
| 3                     | 60                  |
| 4                     | 25                  |
| 5                     | 5                   |

Considerando-se essas informações, o percentual de funcionários dessa empresa que têm até 2 dependentes, inclusive, é igual a

- a) 60%
- b) 55%
- c) 50%
- d) 30%
- e) 25%

#### Comentários:

Para calcular o percentual de funcionários que têm até 2 dependentes, precisamos somar as frequências absolutas dos funcionários que têm 0, 1 ou 2 dependentes e, em seguida, dividir pelo total de funcionários.

Vamos primeiro somar as frequências absolutas:

$$\text{frequência de 0 dependentes} = 20$$



$$\text{frequência de 1 dependente} = 40$$

$$\text{frequência de 2 dependentes} = 50$$

Total de funcionários com até 2 dependentes:

$$20 + 40 + 50 = 110$$

O total de funcionários da empresa é 200.

Agora, calculamos o percentual:

$$\text{Percentual} = \left(\frac{110}{200}\right) \times 100\% = 55\%$$

Portanto, o percentual de funcionários dessa empresa que têm até 2 dependentes, inclusive, é igual a 55%.

**Gabarito: B.**



## 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.

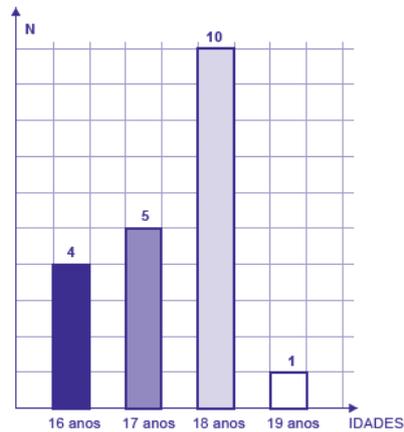
**Gabarito: E.**



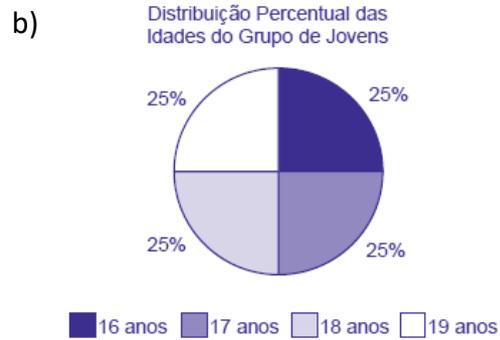
## QUESTÕES COMENTADAS – CESGRANRIO

### Outros Gráficos e Representações

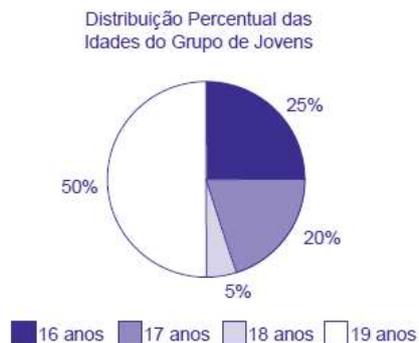
1. (CESGRANRIO/BASA/2024) Um grupo é formado por jovens cujas idades variam entre 16 e 19 anos. O gráfico a seguir representa a distribuição do número (N) de jovens de acordo com suas idades.



O gráfico de setores que reaperesenta, em forma percentual, as frequências das idades desse grupo de jovens é



e)



### Comentários:

Para representar as frequências das idades do grupo de jovens em um gráfico de setores (gráfico de pizza), precisamos calcular a porcentagem que cada idade representa em relação ao total de jovens.

O gráfico de barras define como estão distribuídas as idades dos jovens:

- 16 anos: 4 jovens
- 17 anos: 5 jovens
- 18 anos: 10 jovens
- 19 anos: 1 jovem

Calculando o total de jovens no grupo, temos:

$$\text{Total de jovens} = 4 + 5 + 10 + 1 = 20 \text{ jovens}$$

Agora, podemos calcular a porcentagem de cada idade

- 16 anos:

$$\text{Porcentagem} = \left(\frac{4}{20}\right) \times 100\% = 20\%$$

- 17 anos:

$$\text{Porcentagem} = \left(\frac{5}{20}\right) \times 100\% = 25\%$$

- 18 anos:

$$\text{Porcentagem} = \left(\frac{10}{20}\right) \times 100\% = 50\%$$

- 19 anos:

$$\text{Porcentagem} = \left(\frac{1}{20}\right) \times 100\% = 5\%$$



O gráfico de setores que representa, em forma percentual, as frequências das idades do grupo são aquelas que dividem o círculo em setores correspondentes a 20%, 25%, 50% e 5% para as idades de 16, 17, 18 e 19 anos, respectivamente. Portanto, a resposta correta é a Alternativa A.

**Gabarito: A.**



## AVISO IMPORTANTE!



Olá, alunos (as)!

Informamos que não temos mais questões da banca, referente ao assunto tratado na aula de hoje, em virtude de baixa cobrança deste tópico ao longo dos anos. No entanto, para complementar o estudo e deixar sua preparação em alto nível, preparamos um caderno de questões inéditas que servirá como treino e aprimoramento do conteúdo.

Em caso de dúvidas, não deixe de nos chamar no Fórum de dúvidas!

Bons estudos!

Estratégia Concursos



## QUESTÕES COMENTADAS – INÉDITAS

### Introdução à Estatística

1. (INÉDITA/2022) O ramo da estatística que tem as atribuições de obter, organizar, reduzir e representar os dados estatísticos, bem como de obter informações que auxiliem na descrição de um fenômeno observado é denominada de:

- a) Estatística Indutiva.
- b) Estatística Inferencial.
- c) Estatística Populacional.
- d) Estatística Probabilística.
- e) Estatística Descritiva.

#### Comentários:

A Estatística está dividida em três grandes ramos, a saber:

- a) **Estatística Descritiva (ou Dedutiva)** - responsável pela coleta (ou obtenção), organização, redução e representação de dados estatísticos;
- b) **Estatística Probabilística** - responsável por estabelecer o modelo matemático probabilístico adotado para explicar os fenômenos aleatórios investigados pela Estatística; e
- c) **Estatística Inferencial (ou Indutiva)** - responsável pela análise e interpretação dos dados.

Portanto, a alternativa correta é a letra E.

Gabarito: E.



## QUESTÕES COMENTADAS – INÉDITAS

### Conceitos Iniciais

1. (INÉDITA/2022) Um funcionário decidiu extrair um subconjunto dos dados da empresa em que trabalhava, para analisar o comportamento dos produtos oferecidos por ela. Esse subconjunto é conhecido como

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

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

Alternativa C: 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).

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

Gabarito: C.

2. (INÉDITA/2022) A coleção de unidades observacionais, que podem ser pessoas, objetos ou resultados experimentais, com uma ou mais características em comum, é denominada, para a estatística, de

- a) amostra.



- b) população.
- c) unidades.
- d) censo.
- e) variáveis.

### Comentários:

Para resolvermos a questão vamos conceituar cada uma das alternativas:

- Letra A: amostra é um subconjunto de uma população. Os dados coletados e/ou selecionados de uma população estatística por um procedimento definido, denotado de amostragem.
- Letra B: população é o grupo completo de unidades elementares de pessoas, objetos ou coisas, ou seja, constitui todo o universo de informações de que se necessita.
- Letra C: unidades são elementos da amostra (unidades amostrais) ou da população (unidade populacional).
- Letra D: censo consiste na análise de todos os elementos da população.
- Letra E: variável é a 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: B.



## QUESTÕES COMENTADAS – INÉDITAS

### Variáveis Estatísticas

1. (INÉDITA/2022) As variáveis podem ser classificadas como quantitativas ou qualitativas.

- I. As quantitativas referem-se às variáveis ordinal ou discreta;
- II. As qualitativas referem-se às variáveis nominal ou contínua;
- III. As quantitativas referem-se às variáveis contínua ou discreta;

Assinale a alternativa que contempla apenas as assertivas corretas.

- a) I.
- b) II.
- c) I e II.
- d) III.
- e) II e III.

#### Comentários:

As variáveis podem ser classificadas em:

a) variáveis quantitativas: são as características que podem ser medidas em uma escala quantitativa, isto é, numérica. Elas podem ser contínuas ou discretas:

- variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros. Geralmente são o resultado de contagens. Exemplos: número de membros de uma família, número de carros produzidos por dia.
- variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua (na reta real), nesse caso, valores fracionários fazem sentido. Geralmente são medidas por meio de algum instrumento. Exemplos: peso e altura.

b) variáveis qualitativas (ou categóricas): são as características que não possuem valores quantitativos. Nesse caso, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Elas podem ser nominais ou ordinais:

- variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos.
- variáveis ordinais: existe uma ordenação entre as categorias. Exemplos: grau de instrução (fundamental, médio, superior), nível de aprendizado (básico, intermediário, avançado), nível de queimadura (1º, 2º, 3º graus).



Gabarito: D.

2. (INÉDITA/2022) Um município realizou uma pesquisa exclusivamente com variáveis do tipo qualitativa, sobre a satisfação dos usuários quanto aos serviços públicos disponibilizados. As escalas de mensuração que podem ser empregadas na pesquisa são:

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

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 medidas em escalas nominais ou ordinais:

- Escala nominal: as variáveis são divididas em categorias segundo um ou mais de seus atributos (características) e não podem ser ordenadas. Por exemplo, as cores dos olhos dos moradores de uma determinada cidade (pretos, castanhos, azuis e verdes);
- Escala 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: A.

3. (INÉDITA/2022) Uma empresa realizou uma pesquisa de satisfação para traçar os perfis de seus clientes e o grau de satisfação com os serviços prestados por ela. Em um formulário, ela solicitou os dados a seguir.

I. Idade.

II. Sexo.

III. Escolaridade.

IV. Nota atribuída ao serviço.

Assinale a opção que contempla apenas variáveis categóricas.

- a) I e II.
- b) III e IV.



- c) I, II e III.
- d) II e III.
- e) I, II e IV.

### Comentários:

As variáveis podem ser classificadas em:

a) variáveis quantitativas: são as características que podem ser medidas em uma escala quantitativa, isto é, numérica. Elas podem ser contínuas ou discretas:

- variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros. Geralmente são o resultado de contagens. Exemplos: número de membros de uma família, número de carros produzidos por dia.
- variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua (na reta real), nesse caso, valores fracionários fazem sentido. Geralmente são medidas por meio de algum instrumento. Exemplos: peso e altura.

b) variáveis qualitativas (ou categóricas): são as características que não possuem valores quantitativos. Nesse caso, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Elas podem ser nominais ou ordinais:

- variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos.
- variáveis ordinais: existe uma ordenação entre as categorias. Exemplos: grau de instrução (fundamental, médio, superior), nível de aprendizado (básico, intermediário, avançado), nível de queimadura (1°, 2°, 3° graus).

Com base nos conceitos apresentados, concluímos que:

- I. Idade. (Quantitativa)
- II. Sexo. (Categórica)
- III. Escolaridade. (Categórica)
- IV. Nota atribuída ao serviço. (Quantitativa)

Gabarito: D.



## QUESTÕES COMENTADAS – INÉDITAS

### Séries Estatísticas

1. (INÉDITA/2022) Acerca das séries estatísticas, considere as definições apresentadas a seguir
- I. É a série cujos dados são dispostos segundo a época de ocorrência. Isto é, enquanto o tempo varia, o fato e o local permanecem constantes.
- II. É a série cujos dados são dispostos segundo a modalidade de ocorrência. Isto é, enquanto o fato varia, a época e o local permanecem constantes.
- III. É a série cujos dados são dispostos segundo a localidade de ocorrência. Isto é, enquanto o local varia, o fato e o tempo permanecem constantes.

Assinale a alternativa que relaciona corretamente a definição à série estatística respectiva:

- a) série temporal, série geográfica, série específica
- b) série específica, série temporal, série geográfica
- c) série temporal, série específica, série geográfica
- d) série geográfica, série específica, série temporal
- e) série específica, série geográfica, série temporal

#### Comentários:

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 temporal (ou cronológica), geográfica (ou territorial) e específica. Vejamos:

- série temporal: é a série cujos dados são dispostos segundo a época de ocorrência. Isto é, enquanto o tempo varia, o fato e o local permanecem constantes.
- série geográfica: é a série cujos dados são dispostos segundo a localidade de ocorrência. Isto é, enquanto o local varia, o fato e o tempo permanecem constantes.
- série específica: é a série cujos dados são dispostos segundo a modalidade de ocorrência. Isto é, enquanto o fato varia, a época e o local permanecem constantes.

Gabarito: C.



## QUESTÕES COMENTADAS – INÉDITAS

### Distribuições de Frequência

1. (INÉDITA/2022) A tabela apresenta o número de acertos de 50 candidatos a uma vaga de emprego, em uma avaliação contendo 5 questões de múltipla escolha.

| Número de Acertos | Número de Candidatos |
|-------------------|----------------------|
| 0                 | 6                    |
| 1                 | 12                   |
| 2                 | 14                   |
| 3                 | 8                    |
| 4                 | 7                    |
| 5                 | 3                    |

Com base nas informações apresentadas na tabela, é correto afirmar que

- a) Exatamente 7 candidatos acertaram a questão de número 4.
- b) Mais da metade dos candidatos acertaram mais da metade das questões.
- c) 32 candidatos acertaram, no máximo, duas questões.
- d) Menos da metade dos candidatos acertaram menos da metade das questões.
- e) Exatamente 45 candidatos acertaram, pelo menos, uma questão.

#### Comentários:

Analisando as alternativas, temos:

- Letra A: o enunciado não fornece informações suficientes para que possamos afirmar que os 7 candidatos que acertaram 4 questões também acertaram a quarta questão.
- Letra B: os candidatos das três últimas linhas da tabela acertaram mais da metade das questões. São aqueles que acertaram 3, 4 ou 5 questões. Somando as quantias correspondentes ( $8 + 7 + 3$ ), concluímos que 18 candidatos acertaram mais de metade das questões. Isso corresponde a menos de 50% do total de alunos. Logo, alternativa incorreta.



- Letra C: De fato, 6 candidatos erraram todas; 12 candidatos acertaram (exatamente) uma questão e 14 candidatos acertaram (exatamente) duas questões, formando um total de  $6 + 12 + 14 = 32$ . Portanto, alternativa correta.
- Letra D: os candidatos das três primeiras linhas da tabela acertaram menos da metade das questões. São aqueles que acertaram 0, 1 ou 2 questões. Somando as quantias correspondentes ( $6 + 12 + 14$ ), concluímos que 32 candidatos acertaram menos da metade das questões. Isso corresponde a mais de 50% do total de alunos. Assim, a alternativa está incorreta.
- Letra E: dos 50 candidatos, só 6 erraram todas as questões da prova. Os outros  $50 - 6 = 44$  acertaram pelo menos uma questão. Dessa forma, a alternativa está incorreta.

Gabarito: C.

2. (INÉDITA/2022) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão.

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

4 6 6 5 4 5 4 5 5 4

6 6 2 5 6 2 1 4 3 1

2 5 6 6 3 2 5 6 6 1

5 5 5 1 5 5 6 5 5 4

5 5 2 5 2 4 4 5 3 6

A frequência absoluta simples do primeiro elemento é

- a) 4.
- b) 8.
- c) 6.
- d) 11.
- e) 18.

Comentários:

A frequência absoluta é o número de vezes em que uma determinada variável assume um valor, enquanto a frequência relativa é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado).

Para os dados apresentados, temos que



| Valor obtido no lançamento | Frequência absoluta |
|----------------------------|---------------------|
| 1                          | 4                   |
| 2                          | 6                   |
| 3                          | 3                   |
| 4                          | 8                   |
| 5                          | 18                  |
| 6                          | 11                  |
| Total                      | 50                  |

Portanto, a frequência absoluta simples do primeiro elemento é igual a 4.

Gabarito: A.

3. (INÉDITA/2022) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão. Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

4 6 6 5 4 5 4 5 5 4

6 6 2 5 6 2 1 4 3 1

2 5 6 6 3 2 5 6 6 1

5 5 5 1 5 5 6 5 5 4

5 5 2 5 2 4 4 5 3 6

A frequência relativa simples do quinto elemento é

- a) 5%.
- b) 12%.
- c) 6%.
- d) 36%.
- e) 22%.

Comentários:



A frequência absoluta é o número de vezes em que uma determinada variável assume um valor, enquanto a frequência relativa é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado).

Para os dados apresentados, temos que:

| Valor obtido no lançamento | Frequência absoluta | Frequência relativa   |
|----------------------------|---------------------|-----------------------|
| 1                          | 4                   | $4/50 = 0,08 = 8\%$   |
| 2                          | 6                   | $6/50 = 0,12 = 12\%$  |
| 3                          | 3                   | $3/50 = 0,06 = 6\%$   |
| 4                          | 8                   | $8/50 = 0,16 = 16\%$  |
| 5                          | 18                  | $18/50 = 0,36 = 36\%$ |
| 6                          | 11                  | $11/50 = 0,22 = 22\%$ |
| Total                      | 50                  | 100%                  |

Portanto, a frequência relativa simples do quinto elemento é igual a 36%.

Gabarito: D.

4. (INÉDITA/2022) A tabela apresenta o número de acertos dos 100 candidatos que realizaram a prova da segunda fase de um concurso, que continha 5 questões de múltipla escolha.

| Número de acertos | Número de candidatos |
|-------------------|----------------------|
| 5                 | 65                   |
| 4                 | 15                   |
| 3                 | 8                    |
| 2                 | 7                    |
| 1                 | 3                    |
| 0                 | 2                    |



Analisando-se as informações apresentadas na tabela, é correto afirmar que

- a) Mais da metade dos candidatos acertou menos de 50% da prova.
- b) Menos da metade dos candidatos acertou mais de 50% da prova.
- c) Exatamente 90 candidatos acertaram, no mínimo, 2 questões.
- d) 20 candidatos acertaram, no máximo, 3 questões.
- e) 15 candidatos acertaram a questão de número 4.

#### Comentários:

Analisando as alternativas, temos:

- Letra A: A prova tinha 5 questões. Metade da prova corresponde a 2,5 questões. Portanto, acertaram "menos da metade da prova" quem teve 0, 1 ou 2 acertos, o que perfaz um total de  $2 + 3 + 7 = 12$ . Portanto, temos que 12 pessoas acertaram menos da metade da prova. Logo, menos da metade dos candidatos (50) acertaram menos da metade da prova. Dessa forma, alternativa incorreta.
- Letra B: já temos a informação na letra A de que 12 pessoas acertaram menos da metade da prova. Logo, a quantidade de pessoas que acertaram mais da metade da prova é:  $100 - 12 = 88$ . Portanto, temos que 88 pessoas acertaram mais da metade da prova. Logo, mais da metade das pessoas (50) acertou mais de 50% da prova. Logo, alternativa incorreta.
- Letra C: já vimos na letra A que 12 candidatos acertaram 0, 1 ou 2 questões; portanto, acertaram no máximo duas questões. Já a quantidade dos que acertaram no mínimo duas questões é dada por:  $100 - 2 - 3 = 95$ . Assim, podemos afirmar que a alternativa está incorreta.
- Letra D: 12 candidatos que acertaram até duas questões (vide letra A). Somando com os 8 que acertaram exatamente 3 questões, temos  $12 + 8 = 20$ . Portanto, alternativa correta.
- Letra E: sabemos que 15 candidatos acertaram exatamente quatro questões, mas não temos como saber quantos acertaram a quarta questão da prova. Além disso, nas demais linhas da tabela, ou seja, nas linhas referentes às pessoas que acertaram 1 questão, 2 questões, 3 questões, ou 5 questões, é perfeitamente possível termos mais gente acertando a questão 04. Assim, alternativa incorreta.

Gabarito: D.

5. (INÉDITA/2022) Os salários de 50 funcionários de uma empresa estão distribuídos conforme a tabela a seguir:



| Salários<br>(em salários<br>mínimos) | Funcionários |
|--------------------------------------|--------------|
| 1                                    | 18           |
| 2                                    | 12           |
| 3                                    | 10           |
| 5                                    | 6            |
| 8                                    | 3            |
| 10                                   | 1            |

Pode-se concluir que

- a) o total da folha de pagamentos é de 115 salários.
- b) 60% dos trabalhadores ganham mais ou igual a 3 salários.
- c) 10% dos trabalhadores ganham 10 salários.
- d) 30% dos trabalhadores detêm mais de 50% da renda total.
- e) 60% dos trabalhadores detêm menos de 20% da renda total.

Comentários:

Alternativa A: Incorreta. O total da folha de pagamentos equivale 136 salários, conforme os cálculos a seguir:

| Salários<br>(X) | Funcionários<br>(f) | X×f |
|-----------------|---------------------|-----|
| 1               | 18                  | 18  |
| 2               | 12                  | 24  |
| 3               | 10                  | 30  |



|       |   |     |
|-------|---|-----|
| 5     | 6 | 30  |
| 8     | 3 | 24  |
| 10    | 1 | 10  |
| Total |   | 136 |

Alternativa B: Incorreta. Os trabalhadores que ganham 3 salários ou mais são  $10+6+3+1=20$  funcionários de um total de 50. Isso equivale a  $20 \div 50 = 40\%$  do total.

| Salários (X) | Funcionários (f) | Frequência Relativa |
|--------------|------------------|---------------------|
| 1            | 18               | 36%                 |
| 2            | 12               | 24%                 |
| 3            | 10               | 20%                 |
| 5            | 6                | 12%                 |
| 8            | 3                | 6%                  |
| 10           | 1                | 2%                  |
| Total        | 50               | 100%                |

Alternativa C: Incorreta. Apenas 1 funcionário ganha dez salários, o que corresponde a  $1 \div 50 = 2\%$ .

Alternativa D: Correta. A renda total é de 136 salários, como já calculamos na letra A. Vamos agora determinar 50% disso:

$$0,5 \times 136 = 68 \text{ salários}$$

São 50 funcionários. Portanto, 30% deles correspondem a  $0,3 \times 50 = 15$  funcionários. Os 15 funcionários que mais ganham são os das três últimas linhas da tabela. Eles têm salários de:

$$10 + 8 + 8 + 8 + 5 + 5 + 5 + 5 + 5 + 5 + 3 + 3 + 3 + 3 + 3 = 79$$

Vejam que esses 15 trabalhadores (30% dos trabalhadores) detêm mais de 30% da renda total (mais de 68 salários)



Alternativa E: Incorreta. A renda total é de 136 salários, como já vimos na letra A. Então, 20% disso corresponde a 27,2 salários. Como são 50 funcionários ao todo, 60% disso resultam em 30 funcionários. Eles têm salários de:

| Salários (X) | Funcionários (f) |
|--------------|------------------|
| 1            | 18               |
| 2            | 12               |
| <b>Total</b> | <b>30</b>        |

Então, os funcionários ganham 30 salários, o que é maior que 20% da renda total, ou seja, maior que 27,2 salários.

Gabarito: D.

6. (INÉDITA/2022) A distribuição de frequências é um agrupamento de dados em classes, de tal forma que contabilizamos o número de ocorrências em cada classe. Com relação ao conceito de frequência, avalie as afirmações a seguir.

I - frequência absoluta é o número de observações correspondente a cada classe. A frequência absoluta é, geralmente, chamada apenas de frequência.

II - frequência relativa é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado).

III - frequência acumulada é o total acumulado (soma) de todas as classes anteriores até a classe atual, podendo ser: frequência acumulada absoluta; frequência acumulada relativa ou frequência acumulada percentual.

Assinale a alternativa que contempla todas as assertivas corretas.

- a) I.
- b) II.
- c) II e III.
- d) I, II e III.
- e) I e III.

Comentários:

Com relação à definição de frequência, podemos citar:



- frequência absoluta ( $f_i$ ): é o número de observações correspondente a cada classe. A frequência absoluta é, geralmente, chamada apenas de frequência.
- frequência relativa ( $f_{ri}$ ): é o resultado obtido da divisão entre a frequência absoluta - o valor que é observado na população - e a quantidade de elementos da amostra.
- frequência acumulada: é o total acumulado (soma) de todas as classes anteriores até a classe atual, podendo ser: frequência acumulada absoluta ( $f_i$ ), frequência acumulada relativa ( $f_{ri}$ ) ou frequência acumulada percentual ( $p_i$ ).

Portanto, todas as assertivas estão corretas.

Gabarito: D.

7. (INÉDITA/2022) A distribuição dos salários dos funcionários em uma empresa está representada pela tabela de frequências absolutas abaixo, considerando que todos os intervalos de classe são fechados à esquerda e abertos à direita.

| Ordem da Classe | Classe de Salários (R\$ 1.000) |                 | Frequências |
|-----------------|--------------------------------|-----------------|-------------|
|                 | Limite Inferior                | Limite Superior |             |
| 1ª              | 1                              | 2               | 15          |
| 2ª              | 2                              | 3               | 25          |
| 3ª              | 3                              | 4               | 45          |
| 4ª              | 4                              | 5               | 35          |
| 5ª              | 5                              | 6               | 30          |

Se inexistem funcionários que ganham salários inferiores a R\$ 1.000,00 ou salários iguais ou superiores a R\$ 6.000,00, então a porcentagem de funcionários que ganham, pelo menos, R\$ 2.000,00 e menos que R\$ 4.000,00 é igual a

- a) 43,33%.
- b) 45,00%.
- c) 46,67%.
- d) 50,00%.
- e) 53,33%.



Comentários:

Inicialmente, precisamos calcular as frequências relativas de cada classe:

| Ordem da Classe | Classe de Salários (R\$ 1.000) |                 | Frequências | Frequências Relativas |
|-----------------|--------------------------------|-----------------|-------------|-----------------------|
|                 | Limite Inferior                | Limite Superior |             |                       |
| 1ª              | 1                              | 2               | 15          | 10,00%                |
| 2ª              | 2                              | 3               | 25          | 16,67%                |
| 3ª              | 3                              | 4               | 45          | 30,00%                |
| 4ª              | 4                              | 5               | 35          | 23,33%                |
| 5ª              | 5                              | 6               | 30          | 20,00%                |
| <b>TOTAL</b>    |                                |                 | <b>150</b>  |                       |

O percentual de pessoas que ganham, pelo menos, R\$ 2.000,00 e menos que R\$ 4.000,00 é igual a soma dos percentuais da 2ª e da 3ª classe:

$$16,67 + 30,00 = 46,67\%$$

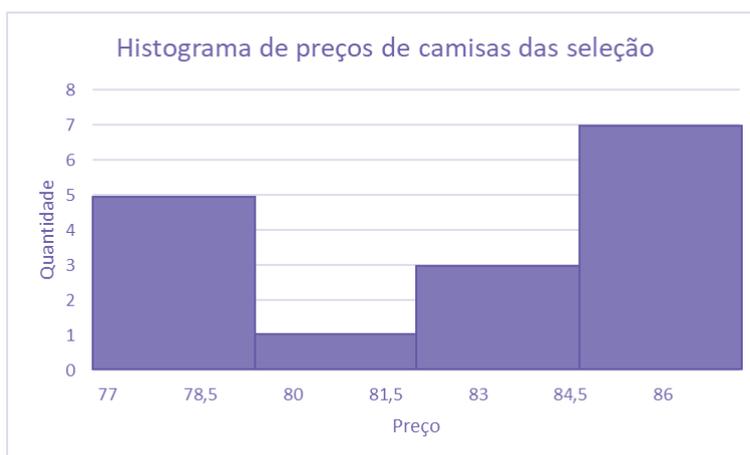
Gabarito: C.



## QUESTÕES COMENTADAS – INÉDITAS

### Representação Gráfica das Distribuições de Frequências

1. (INÉDITA) O histograma a seguir mostra a quantidade de camisas da seleção brasileira, por faixas de preço, vendidas em determinada região.



Sabendo-se que o menor valor encontrado foi de R\$ 77,00; que o maior valor foi de R\$ 87,00; e que os valores foram divididos em 4 classes de mesma amplitude, marque o conjunto de dados consistente com o histograma:

- a) 77, 77, 77, 78, 79, 84, 84, 84, 85, 85, 85, 85, 86, 86, 86, 87;
- b) 77, 77, 77, 78, 79, 84, 84, 84, 85, 85, 86, 86, 87;
- c) 77, 77, 77, 78, 79, 81, 83, 84, 84, 85, 85, 85, 86, 86, 86, 87;
- d) 77, 77, 78, 79, 80, 84, 84, 84, 85, 85, 85, 85, 86, 86, 86, 87;
- e) 77, 77, 77, 78, 79, 81, 81, 84, 84, 85, 85, 85, 86, 86, 86, 87.

#### Comentários:

Calculando a amplitude de cada classe:

$$\frac{87 - 77}{4} = 2,50$$

De posse dessa informação, podemos formar as classes do histograma:

$$77 \text{ } \vdash \text{ } 79,50$$



79,50 † 82,00

82,00 † 84,50

84,50 † 87,00

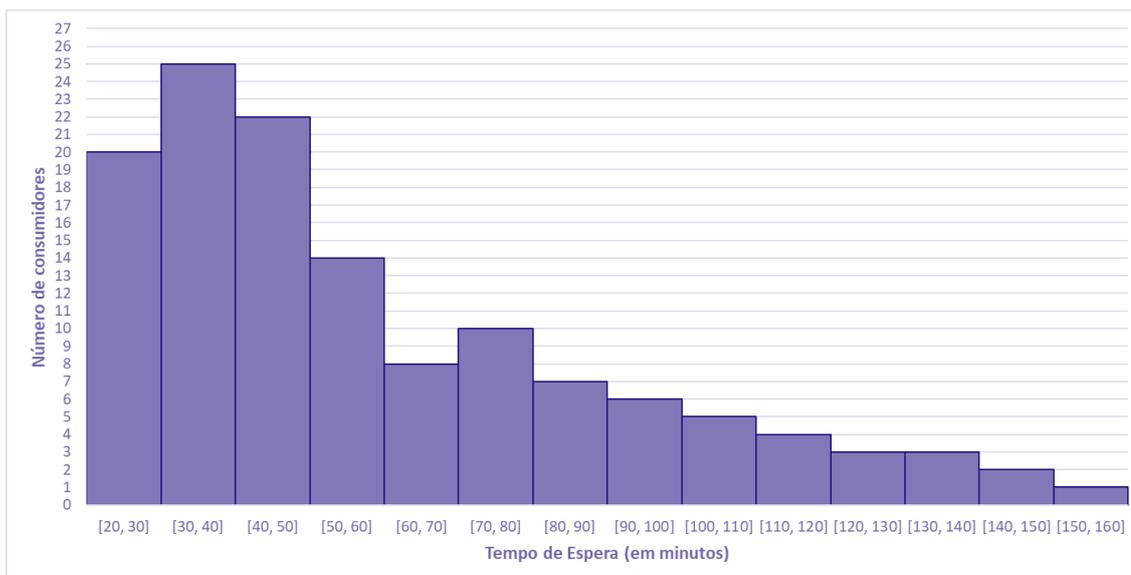
Agora, podemos acrescentar as frequências de cada classe:

| Classes       | Frequências |
|---------------|-------------|
| 77 † 79,50    | 5           |
| 79,50 † 82,00 | 1           |
| 82,00 † 84,50 | 3           |
| 84,50 † 87,00 | 7           |

Analisando as alternativas, concluímos que a letra C está correta.

Gabarito: C.

2. (INÉDITA) Os tempos de espera, em minutos, para o atendimento de 130 consumidores em uma central telefônica estão registrados no gráfico a seguir.



Com base no gráfico, afirmar que o tempo de espera de

a) mais da metade dos consumidores foi superior a 1 hora.



- b) 37 consumidores foi entre 50 min e 80 min.
- c) 70% dos consumidores foi inferior a 1 hora.
- d) 10,0% dos consumidores foi entre 1 h 50 min e 2 h e 40 min.
- e) no mínimo 3 pessoas, foi superior a 2 h e 30 min.

### Comentários:

Primeiro, podemos montar a tabela de frequências acumuladas para simplificar a análise das alternativas:

| Classes   | Frequência Absoluta | Frequência Absoluta Acumulada | Frequência Relativa | Frequência Relativa Acumulada |
|-----------|---------------------|-------------------------------|---------------------|-------------------------------|
| 20 † 30   | 20                  | 20                            | 15,38%              | 15,38%                        |
| 30 † 40   | 25                  | 45                            | 19,23%              | 34,62%                        |
| 40 † 50   | 22                  | 67                            | 16,92%              | 51,54%                        |
| 50 † 60   | 14                  | 81                            | 10,77%              | 62,31%                        |
| 60 † 70   | 8                   | 89                            | 6,15%               | 68,46%                        |
| 70 † 80   | 10                  | 99                            | 7,69%               | 76,15%                        |
| 80 † 90   | 7                   | 106                           | 5,38%               | 81,54%                        |
| 90 † 100  | 6                   | 112                           | 4,62%               | 86,15%                        |
| 100 † 110 | 5                   | 117                           | 3,85%               | 90,00%                        |
| 110 † 120 | 4                   | 121                           | 3,08%               | 93,08%                        |
| 120 † 130 | 3                   | 124                           | 2,31%               | 95,38%                        |
| 130 † 140 | 3                   | 127                           | 2,31%               | 97,69%                        |
| 140 † 150 | 2                   | 129                           | 1,54%               | 99,23%                        |



|              |            |            |             |             |
|--------------|------------|------------|-------------|-------------|
| 150 † 160    | 1          | 130        | 0,77%       | 100,00%     |
| <b>Total</b> | <b>130</b> | <b>130</b> | <b>100%</b> | <b>100%</b> |

Agora, analisando as alternativas, podemos afirmar que:

Alternativa A: Incorreta. O tempo de espera foi inferior a 50 minutos para mais da metade (51,54%) das pessoas.

Alternativa B: Incorreta. O tempo de espera foi entre 50 min e 80 min para  $14 + 8 + 10 = 32$  pessoas.

Alternativa C: Incorreta. O tempo de espera foi inferior a 1h para 62,31% das pessoas.

Alternativa D: Correta. Para 10,0% dos consumidores, o tempo de espera foi entre 1h e 50 min e 2h e 40min. Somando as frequências relativas referentes às classes 110 † 120, 120 † 130, 130 † 140, 140 † 150 e 150 † 160, temos:

$$3,08\% + 2,31\% + 2,31\% + 1,54\% + 0,77\% = 10,0\%$$

Alternativa E: Incorreta. O tempo de espera foi superior a 2h e 30 min para apenas uma pessoa.

**Gabarito: D.**

**3. (INÉDITA)** Após a coleta de dados, uma das formas de se iniciar a análise dos dados é por meio da geração e a avaliação de um histograma para uma variável selecionada. Em estudos climatológicos, por exemplo, o histograma pode ser utilizado para se estudar a precipitação mensal das chuvas de uma determinada região. Nesse contexto, o histograma é um gráfico em que são apresentadas as

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

**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.

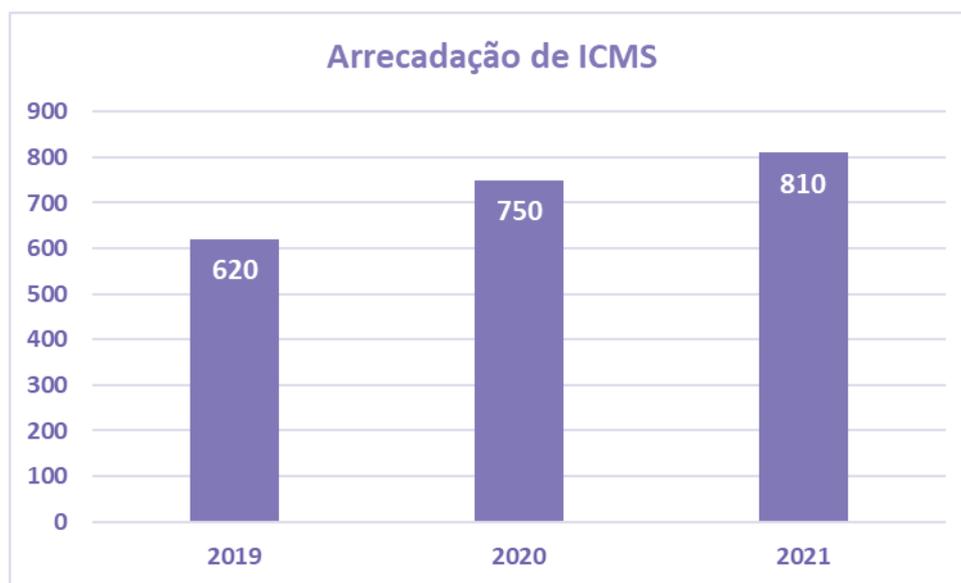
**Gabarito: A.**



## QUESTÕES COMENTADAS – INÉDITAS

### Outros Gráficos e Representações

1. (INÉDITA) O gráfico a seguir apresenta informações sobre a arrecadação de ICMS de determinado Estado brasileiro.



Com base nas informações do gráfico, avalie as afirmações a seguir.

- I. De 2019 a 2021, a arrecadação de ICMS desse Estado aumentou.
- II. De 2020 para 2021, houve um aumento de mais de 10% na arrecadação de ICMS desse Estado.
- III. É crescente a variação da arrecadação de ICMS desse Estado, de 2019 para 2020 e de 2020 para 2021.

Está correto o contido em

- a) apenas I.
- b) apenas I e II.
- c) apenas I e III.
- d) apenas II e III.
- e) I, II e III.

**Comentários:**

Analisando as assertivas:



**Item I. Correto.** De 2019 a 2021, a arrecadação de ICMS desse Estado aumentou. De acordo com o gráfico, a cada ano que passa, a arrecadação de ICMS é maior.

**Item II. Incorreto.** Em 2021, houve uma arrecadação de R\$ 810 milhões e, em 2020, a arrecadação foi de R\$ 750 milhões. Assim, o aumento da arrecadação foi de:  $810 - 750 = \text{R\$ } 60$  milhões. Como 10% de 750 é  $0,10 \times 750 = 75$  milhões, o aumento da arrecadação foi menor que 10%.

**Item III. Correto.** De 2019 para 2020, a variação foi de  $750 - 620 = 130$  milhões. E, de 2020 para 2021, a variação foi de  $810 - 750 = \text{R\$ } 60$  milhões. Logo, a variação da arrecadação é crescente.

**Gabarito: C.**

**2. (INÉDITA) Durante uma pesquisa de campo, algumas variáveis se destacam.**

(i) número de indivíduos por faixa de massa corporal;

(ii) percentuais do nível de escolaridade; e

(iii) número de indivíduos por região.

**Portanto, os tipos de gráficos considerados adequados a serem empregados em cada caso são, respectivamente:**

a) cartograma, linha e setores;

b) setores, cartograma e barras;

c) histograma, setores e linha;

d) histograma, setores e cartograma;

e) pictograma, setores e linha.

**Comentários:**

Analisando as assertivas:

**Item I.** O histograma é o gráfico mais indicado para representação do número de indivíduos por faixa de massa corporal.

**Item II.** O gráfico de setor é recomendado para representação de percentuais de variáveis categóricas, como no caso do nível de escolaridade.

**Item III.** Para representação do nível de renda por região, podemos utilizar um cartograma, que consiste em representar um território indicando de maneira proporcional os valores de uma determinada variável.



Gabarito: D.

3. (INÉDITA) A respeito de gráficos estatísticos, avalie as afirmações a seguir.

I. Em um gráfico de colunas, as alturas de cada retângulo são proporcionais aos dados.

II. Um gráfico de barras apresenta a frequência absoluta (ou relativa) cumulativa.

III. Em um gráfico de linhas, os dados de categorias são distribuídos uniformemente ao longo do eixo vertical, e todos os valores são distribuídos igualmente ao longo do eixo horizontal.

Assinale a alternativa que contempla todas as assertivas corretas.

- a) I
- b) I e II
- c) II e III
- d) I, II e III
- e) I e III

**Comentários:**

Analisando as assertivas:

**Item I. Correto.** Em um gráfico de colunas, as alturas de cada retângulo são proporcionais aos dados.

**Item II. Incorreto.** Não necessariamente precisa ser a frequência acumulada. O gráfico de barras também pode ser construído com base na frequência absoluta simples.

**Item III. Incorreto.** Normalmente, em um gráfico de linhas, as categorias estão distribuídas ao longo do eixo horizontal e os dados estão distribuídos proporcionalmente aos seus valores ao longo do eixo vertical.

Gabarito: A.

4. (INÉDITA) A respeito das representações gráficas, avalie as afirmações a seguir.

I. Os gráficos de setores apresentam os dados em um círculo, no qual cada setor indica a frequência acumulada de um item observado.

II. Os gráficos de barras verticais apresentam os dados por meio de colunas retangulares em que a altura de cada coluna corresponde à frequência (absoluta ou relativa) dos itens observados.



III. Os pictogramas exibem os dados por meio de símbolos que possuem o mesmo tamanho e valor.

Assinale a alternativa que contempla apenas as assertivas corretas.

- a) I.
- b) I e II.
- c) II e III.
- d) I, II e III.
- e) I e III.

#### Comentários:

Analisando-se cada item, temos que:

**Item I: Incorreto.** O gráfico de setores é um diagrama circular no qual os valores de cada categoria representada são proporcionais às respectivas **frequências simples**, podendo vir acompanhado de porcentagens. Normalmente é utilizado para dados qualitativos nominais.

**Item II: Correto.** Um gráfico de barras resume um conjunto de dados categóricos utilizando um número de barras de mesma largura, cada uma delas representando uma categoria particular. A altura de cada barra é proporcional a uma frequência específica, seja ela absoluta ou relativa.

**Item III: Correto.** Os pictogramas são gráficos que utilizam imagens para chamar atenção, podendo ser representado por gráficos de barras, colunas, linhas e circulares. As imagens devem possuir o mesmo tamanho, mas podem aparecer fracionadas para mostrar a respectiva fração de uma determinada quantidade.

**Gabarito: C.**

5. (INÉDITA) Acerca das representações gráficas, considere:

- I. O gráfico de setores não é adequado para representar variáveis categóricas.
- II. O gráfico de barras não é adequado para representar variáveis quantitativas contínuas.
- III. Histograma é um gráfico que apresenta a distribuição de frequências de uma variável por meio de retângulos justapostos, sendo a área de cada retângulo proporcional à frequência observada da classe correspondente.

Está correto o que consta apenas em:



- a) I.
- b) III.
- c) I e II.
- d) I e III.
- e) II e III.

### Comentários:

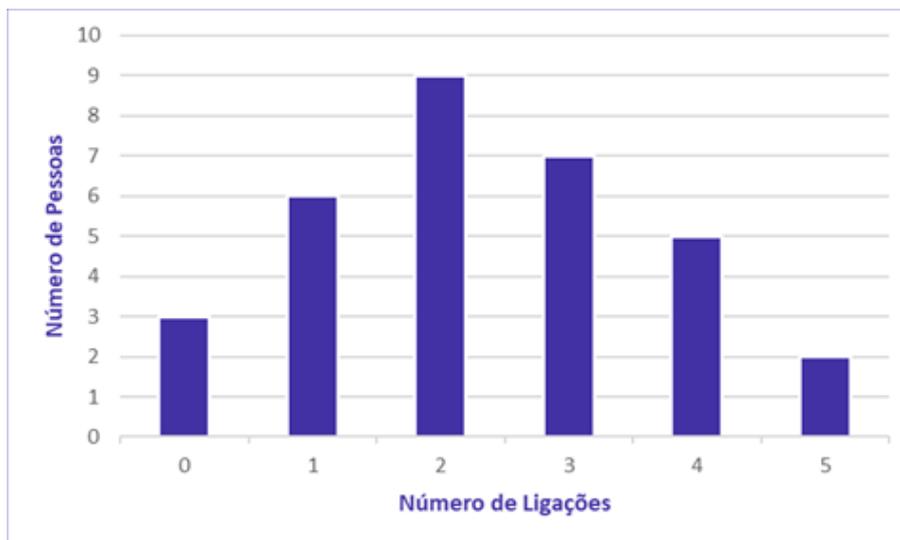
**Item I - Incorreto.** O gráfico de setores é indicado para representar variáveis categóricas. Nesse caso, cada fatia da "pizza", ou setor, é utilizado para representar uma categoria da variável.

**Item II - Correto.** Como cada barra tem que ser associada a um único valor, não temos como representar variáveis contínuas.

**Item III - Correto.** A assertiva apresentou a exata definição de um histograma.

**Gabarito: E.**

6. (INÉDITA) Em uma empresa foi realizada uma auditoria para descobrir com que frequência os funcionários atendiam ligações pessoais no horário de trabalho de um único dia. O resultado está representado no gráfico a seguir.



O número de pessoas que trabalham nessa empresa é

- a) 29.
- b) 30.



- c) 31.
- d) 32.
- e) 33.

**Comentários:**

Ao analisarmos o gráfico, percebemos a existência de um eixo horizontal, que nos mostra o número de ligações atendidas no período; e um eixo vertical, que nos apresenta o número de pessoas. Assim, podemos montar uma tabela atribuindo os valores do gráfico:

| Nº de ligações | Nº de pessoas |
|----------------|---------------|
| 0              | 3             |
| 1              | 6             |
| 2              | 9             |
| 3              | 7             |
| 4              | 5             |
| 5              | 2             |

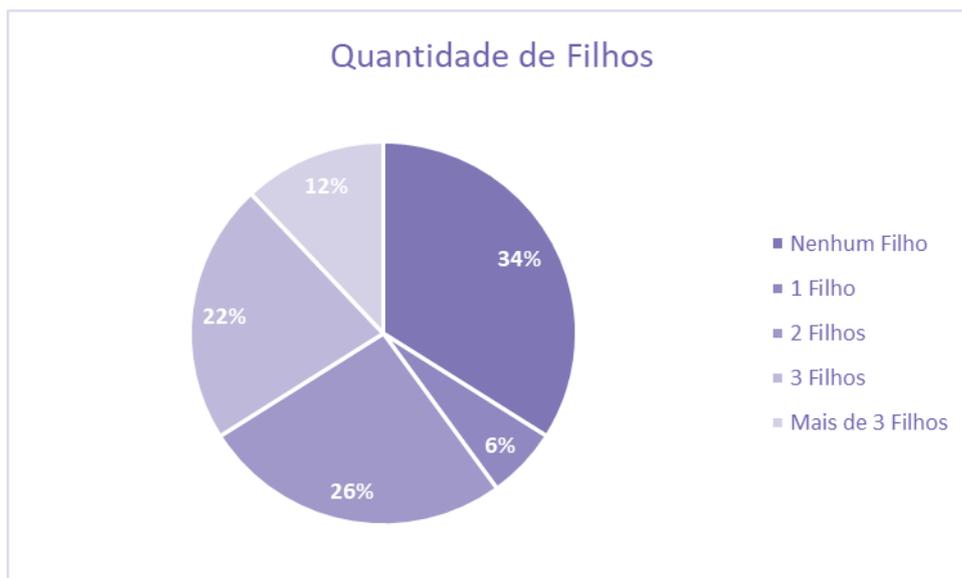
Logo, somando todos os valores que representam o número de pessoas, temos:

$$3 + 6 + 9 + 7 + 5 + 2 = 32 \text{ pessoas}$$

**Gabarito: D.**

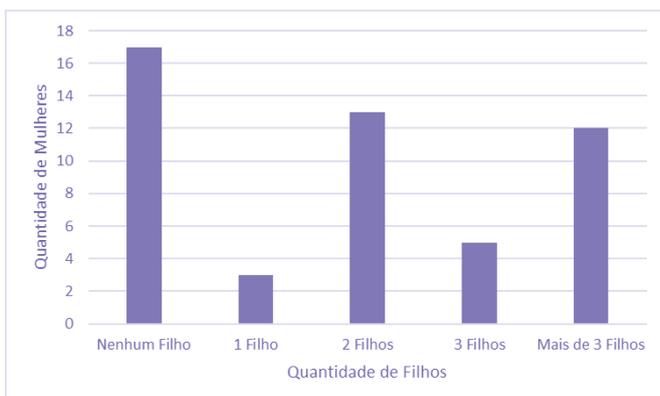
7. (INÉDITA) O gráfico de setor apresentado a seguir fornece os resultados de uma pesquisa realizada com 50 mulheres acerca da quantidade de filhos.



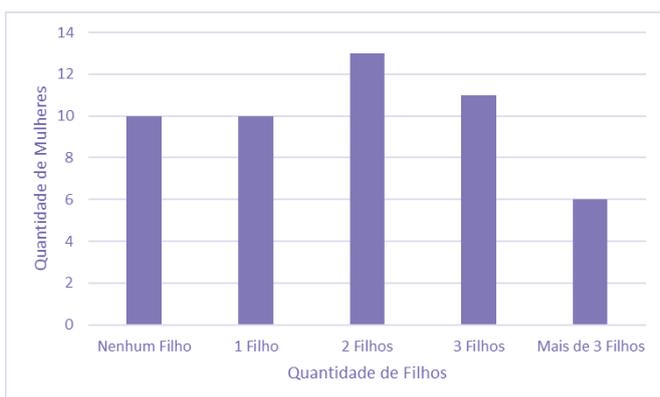


Assinale o gráfico de colunas que representa os dados apresentados acima.

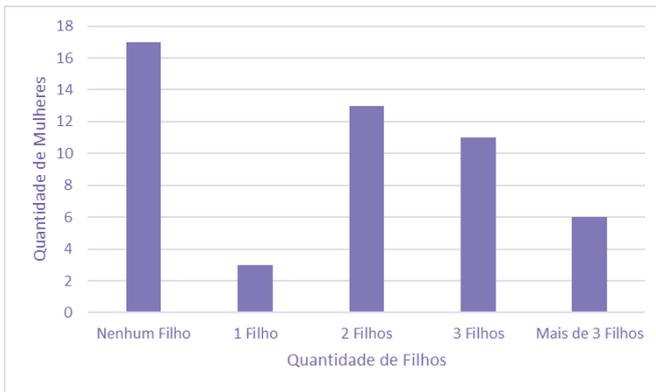
a)



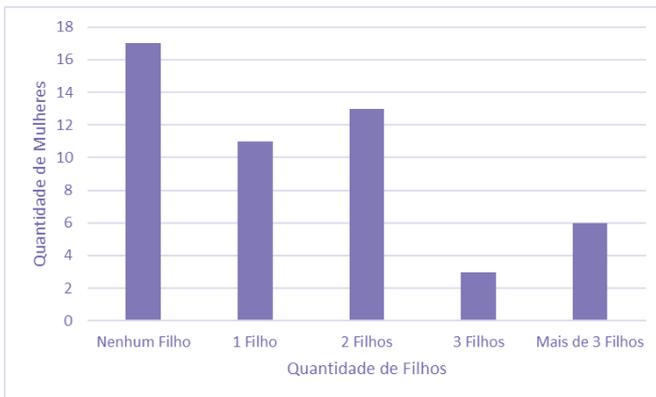
b)



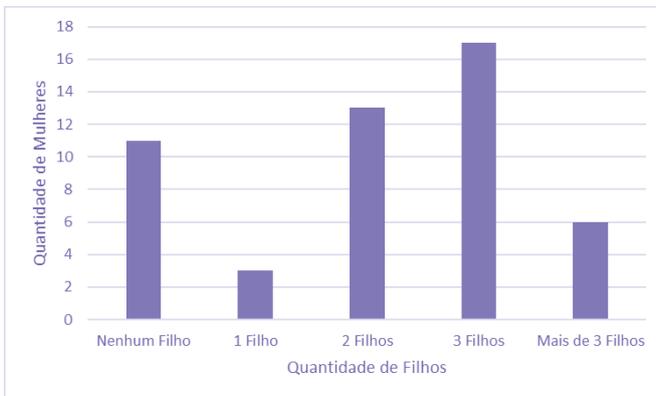
c)



d)



e)



### Comentários:

Segundo o enunciado, a pesquisa foi realizada com um total de 50 mulheres. Interpretando as informações do gráfico, temos:

- nenhum filho = 34%. Logo, 34% de 50 mulheres correspondem a:  $0,34 \times 50 = 17$  mulheres;
- um filho = 6%. Logo, 6% de 50 mulheres correspondem a:  $0,06 \times 50 = 3$  mulheres.



- três filhos = 22%. Logo, 22% de 50 mulheres correspondem a:  $0,22 \times 50 = 11$  *mulheres*.

Olhando para as alternativas, percebemos que não se faz necessário calcular todas as colunas do gráfico. Portanto, ficamos entre as alternativas a) ou c), cujas primeiras colunas demonstram que 17 mulheres não tiveram nenhum filho.

Passamos para a terceira informação calculada, em que 11 mulheres tiveram 3 filhos. Com isso eliminamos a letra a) e temos nosso gabarito, letra c).

**Gabarito: C.**

**8. (INÉDITA) Dentre os tipos de representações gráficas a seguir, o único que não pode ser utilizado para representar uma variável categórica é:**

- a) histograma.
- b) diagrama de barras justapostas.
- c) diagrama de colunas.
- d) diagrama de setores.
- e) diagrama de colunas empilhadas.

**Comentários:**

Variáveis qualitativas (ou categóricas) são as características que não possuem valores quantitativos. Nesse caso, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Os gráficos normalmente utilizados para representação desse tipo de variável são: gráfico de barras, gráfico de colunas e gráfico de setores. Por sua vez, o histograma é indicado para variáveis quantitativas, sendo a nossa resposta.

**Gabarito: A.**



## 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

2. (CESGRANRIO/UNIRIO/2019) As pesquisas da área da saúde se utilizam dos estudos da bioestatística. Nesses estudos, um dos conceitos estatísticos fundamentais é o de população, que se constitui em

- a) conjunto de elementos a serem observados.
- b) valor obtido com maior frequência.
- c) resultado do sorteio aleatório para definir o que será estudado.
- d) dado aritmético resultante da raiz quadrada de cálculos logarítmicos.
- e) subconjunto de dados numéricos que se pretendem ordenar.



## GABARITO – CESGRANRIO

### Conceitos Iniciais

1. LETRA A

2. LETRA A



## LISTA DE QUESTÕES – CESGRANRIO

### Variáveis Estatísticas

1. (CESGRANRIO/BNDES/2024) O uso de dados quantitativos e qualitativos é fundamental para a compreensão e a análise da situação de um país, servindo de importante subsídio à tomada de decisões. Esses dados são mensurados a partir de diferentes variáveis. Sobre essas variáveis, verifica-se que

- a) estado civil, grau de instrução, região e salários são variáveis qualitativas.
- b) salários, idade, grau de instrução e estado civil são variáveis quantitativas.
- c) número de filhos é uma variável quantitativa discreta.
- d) grau de instrução é uma variável qualitativa nominal.
- e) região é uma variável quantitativa discreta.

2. (CESGRANRIO/CEF/2024) Instituições financeiras realizam empréstimos na expectativa de receberem de volta o valor do empréstimo acrescido de juros. Nem sempre, todavia, o cliente paga ao banco o valor que lhe é devido, e isso se denomina inadimplência. A probabilidade de um cliente inadimplir varia de acordo com as características do cliente, como renda, idade ou se o cliente possui casa própria. Quando um cliente solicita um empréstimo, um modelo matemático utiliza essas características para estimar a probabilidade de o cliente inadimplir. Essa probabilidade estimada será usada para classificar o cliente conforme o seu risco de crédito. Esse risco pode variar de AA (para clientes cuja probabilidade de inadimplir é quase nula), passando por B, C, D, E, F, G e finalmente H, à medida que a probabilidade de inadimplência vai aumentando. Segundo a apresentação de resultados do terceiro trimestre de 2023, a Caixa Econômica Federal tem 92,7% de sua carteira com risco de crédito entre AA e C e encerrou 2023 com uma inadimplência de 2,16%.

A variável risco de crédito descrita acima é uma variável

- a) escalar
- b) discreta
- c) contínua
- d) nominal
- e) ordinal

3. (CESGRANRIO/IPEA/2024) Um instituto de pesquisa está planejando uma pesquisa em que uma das perguntas diz respeito à escolaridade do entrevistado. A pergunta admite os valores 1, 2, 3, 4 e 5, abrangendo desde ensino fundamental incompleto (1) até ensino superior completo (5).



**Essa é uma variável**

- a) discreta
- b) contínua
- c) ordinal
- d) nominal
- e) natural

**4. (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
2. LETRA E
3. LETRA C
4. LETRA C



## LISTA DE QUESTÕES – CESGRANRIO

### Distribuições de Frequência

1. (CESGRANRIO/CAIXA/2024) Uma pesquisa foi realizada com todos os 200 funcionários de uma empresa. Uma das variáveis a ser investigada é o número de dependentes por funcionário. Os dados sobre essa variável foram organizados e estão resumidos na Tabela a seguir.

Número de dependentes por funcionários

| Número de dependentes | Frequência absoluta |
|-----------------------|---------------------|
| 0                     | 20                  |
| 1                     | 40                  |
| 2                     | 50                  |
| 3                     | 60                  |
| 4                     | 25                  |
| 5                     | 5                   |

Considerando-se essas informações, o percentual de funcionários dessa empresa que têm até 2 dependentes, inclusive, é igual a

- a) 60%
- b) 55%
- c) 50%
- d) 30%
- e) 25%



## GABARITO – CESGRANRIO

### Distribuições de Frequência

#### 1. LETRA B



## 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

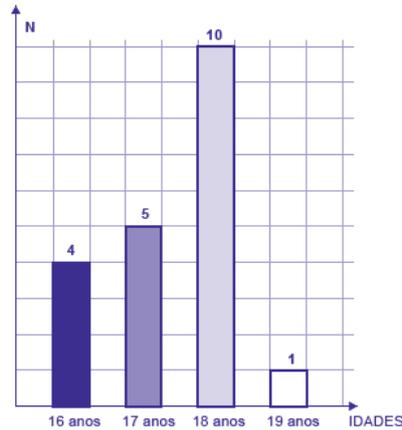
#### 1. LETRA E



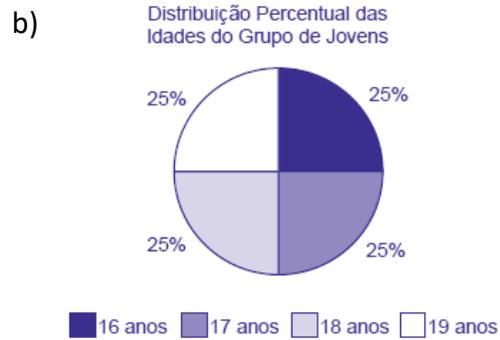
## LISTA DE QUESTÕES – CESGRANRIO

### Outros Gráficos e Representações

1. (CESGRANRIO/BASA/2024) Um grupo é formado por jovens cujas idades variam entre 16 e 19 anos. O gráfico a seguir representa a distribuição do número (N) de jovens de acordo com suas idades.



O gráfico de setores que reaperesenta, em forma percentual, as frequências das idades desse grupo de jovens é



e)

Distribuição Percentual das Idades do Grupo de Jovens



## GABARITO – CESGRANRIO

### Outros Gráficos e Representações

1. LETRA A



## LISTA DE QUESTÕES – INÉDITAS

### Introdução à Estatística

1. (INÉDITA/2022) O ramo da estatística que tem as atribuições de obter, organizar, reduzir e representar os dados estatísticos, bem como de obter informações que auxiliem na descrição de um fenômeno observado é denominada de:

- a) Estatística Indutiva.
- b) Estatística Inferencial.
- c) Estatística Populacional.
- d) Estatística Probabilística.
- e) Estatística Descritiva.



## GABARITO – INÉDITAS

### Introdução à Estatística

1. LETRA E



## LISTA DE QUESTÕES – INÉDITAS

### Conceitos Iniciais

1. (INÉDITA/2022) Um funcionário decidiu extrair um subconjunto dos dados da empresa em que trabalhava, para analisar o comportamento dos produtos oferecidos por ela. Esse subconjunto é conhecido como

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

2. (INÉDITA/2022) A coleção de unidades observacionais, que podem ser pessoas, objetos ou resultados experimentais, com uma ou mais características em comum, é denominada, para a estatística, de

- a) amostra.
- b) população.
- c) unidades.
- d) censo.
- e) variáveis.



## GABARITO – INÉDITAS

### Conceitos Iniciais

1. LETRA C

2. LETRA B



## LISTA DE QUESTÕES – INÉDITAS

### Variáveis Estatísticas

1. (INÉDITA/2022) As variáveis podem ser classificadas como quantitativas ou qualitativas.

- I. As quantitativas referem-se às variáveis ordinal ou discreta;
- II. As qualitativas referem-se às variáveis nominal ou contínua;
- III. As quantitativas referem-se às variáveis contínua ou discreta;

Assinale a alternativa que contempla apenas as assertivas corretas.

- a) I.
- b) II.
- c) I e II.
- d) III.
- e) II e III.

2. (INÉDITA/2022) Um município realizou uma pesquisa exclusivamente com variáveis do tipo qualitativa, sobre a satisfação dos usuários quanto aos serviços públicos disponibilizados. As escalas de mensuração que podem ser empregadas na pesquisa são:

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

3. (INÉDITA/2022) Uma empresa realizou uma pesquisa de satisfação para traçar os perfis de seus clientes e o grau de satisfação com os serviços prestados por ela. Em um formulário, ela solicitou os dados a seguir.

- I. Idade.
- II. Sexo.
- III. Escolaridade.
- IV. Nota atribuída ao serviço.

Assinale a opção que contempla apenas variáveis categóricas.



- a) I e II.
- b) III e IV.
- c) I, II e III.
- d) II e III.
- e) I, II e IV.



## GABARITO – INÉDITAS

### Variáveis Estatísticas

1. LETRA D

2. LETRA A

3. LETRA D



## LISTA DE QUESTÕES – INÉDITAS

### Séries Estatísticas

1. (INÉDITA/2022) Acerca das séries estatísticas, considere as definições apresentadas a seguir
- I. É a série cujos dados são dispostos segundo a época de ocorrência. Isto é, enquanto o tempo varia, o fato e o local permanecem constantes.
- II. É a série cujos dados são dispostos segundo a modalidade de ocorrência. Isto é, enquanto o fato varia, a época e o local permanecem constantes.
- III. É a série cujos dados são dispostos segundo a localidade de ocorrência. Isto é, enquanto o local varia, o fato e o tempo permanecem constantes.

Assinale a alternativa que relaciona corretamente a definição à série estatística respectiva:

- a) série temporal, série geográfica, série específica
- b) série específica, série temporal, série geográfica
- c) série temporal, série específica, série geográfica
- d) série geográfica, série específica, série temporal
- e) série específica, série geográfica, série temporal



## GABARITO – INÉDITAS

### Séries Estatísticas

1. LETRA C



## LISTA DE QUESTÕES – INÉDITAS

### Distribuições de Frequência

1. (INÉDITA/2022) A tabela apresenta o número de acertos de 50 candidatos a uma vaga de emprego, em uma avaliação contendo 5 questões de múltipla escolha.

| Número de Acertos | Número de Candidatos |
|-------------------|----------------------|
| 0                 | 6                    |
| 1                 | 12                   |
| 2                 | 14                   |
| 3                 | 8                    |
| 4                 | 7                    |
| 5                 | 3                    |

Com base nas informações apresentadas na tabela, é correto afirmar que

- a) Exatamente 7 candidatos acertaram a questão de número 4.
- b) Mais da metade dos candidatos acertaram mais da metade das questões.
- c) 32 candidatos acertaram, no máximo, duas questões.
- d) Menos da metade dos candidatos acertaram menos da metade das questões.
- e) Exatamente 45 candidatos acertaram, pelo menos, uma questão.

2. (INÉDITA/2022) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão.

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

4 6 6 5 4 5 4 5 5 4

6 6 2 5 6 2 1 4 3 1

2 5 6 6 3 2 5 6 6 1

5 5 5 1 5 5 6 5 5 4

5 5 2 5 2 4 4 5 3 6



A frequência absoluta simples do primeiro elemento é

- a) 4.
- b) 8.
- c) 6.
- d) 11.
- e) 18.

3. (INÉDITA/2022) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão. Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

4 6 6 5 4 5 4 5 5 4  
6 6 2 5 6 2 1 4 3 1  
2 5 6 6 3 2 5 6 6 1  
5 5 5 1 5 5 6 5 5 4  
5 5 2 5 2 4 4 5 3 6

A frequência relativa simples do quinto elemento é

- a) 5%.
- b) 12%.
- c) 6%.
- d) 36%.
- e) 22%.

4. (INÉDITA/2022) A tabela apresenta o número de acertos dos 100 candidatos que realizaram a prova da segunda fase de um concurso, que continha 5 questões de múltipla escolha.

| Número de acertos | Número de candidatos |
|-------------------|----------------------|
| 5                 | 65                   |
| 4                 | 15                   |
| 3                 | 8                    |



|   |   |
|---|---|
| 2 | 7 |
| 1 | 3 |
| 0 | 2 |

Analisando-se as informações apresentadas na tabela, é correto afirmar que

- a) Mais da metade dos candidatos acertou menos de 50% da prova.
- b) Menos da metade dos candidatos acertou mais de 50% da prova.
- c) Exatamente 90 candidatos acertaram, no mínimo, 2 questões.
- d) 20 candidatos acertaram, no máximo, 3 questões.
- e) 15 candidatos acertaram a questão de número 4.

5. (INÉDITA/2022) Os salários de 50 funcionários de uma empresa estão distribuídos conforme a tabela a seguir:

| Salários<br>(em salários<br>mínimos) | Funcionários |
|--------------------------------------|--------------|
| 1                                    | 18           |
| 2                                    | 12           |
| 3                                    | 10           |
| 5                                    | 6            |
| 8                                    | 3            |
| 10                                   | 1            |

Pode-se concluir que

- a) o total da folha de pagamentos é de 115 salários.
- b) 60% dos trabalhadores ganham mais ou igual a 3 salários.
- c) 10% dos trabalhadores ganham 10 salários.
- d) 30% dos trabalhadores detêm mais de 50% da renda total.



e) 60% dos trabalhadores detêm menos de 20% da renda total.

6. (INÉDITA/2022) A distribuição de frequências é um agrupamento de dados em classes, de tal forma que contabilizamos o número de ocorrências em cada classe. Com relação ao conceito de frequência, avalie as afirmações a seguir.

I - frequência absoluta é o número de observações correspondente a cada classe. A frequência absoluta é, geralmente, chamada apenas de frequência.

II - frequência relativa é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado).

III - frequência acumulada é o total acumulado (soma) de todas as classes anteriores até a classe atual, podendo ser: frequência acumulada absoluta; frequência acumulada relativa ou frequência acumulada percentual.

Assinale a alternativa que contempla todas as assertivas corretas.

- a) I.
- b) II.
- c) II e III.
- d) I, II e III.
- e) I e III.

7. (INÉDITA/2022) A distribuição dos salários dos funcionários em uma empresa está representada pela tabela de frequências absolutas abaixo, considerando que todos os intervalos de classe são fechados à esquerda e abertos à direita.

| Ordem da Classe | Classe de Salários (R\$ 1.000) |                 | Frequências |
|-----------------|--------------------------------|-----------------|-------------|
|                 | Limite Inferior                | Limite Superior |             |
| 1 <sup>a</sup>  | 1                              | 2               | 15          |
| 2 <sup>a</sup>  | 2                              | 3               | 25          |
| 3 <sup>a</sup>  | 3                              | 4               | 45          |
| 4 <sup>a</sup>  | 4                              | 5               | 35          |
| 5 <sup>a</sup>  | 5                              | 6               | 30          |



Se inexistem funcionários que ganham salários inferiores a R\$ 1.000,00 ou salários iguais ou superiores a R\$ 6.000,00, então a porcentagem de funcionários que ganham, pelo menos, R\$ 2.000,00 e menos que R\$ 4.000,00 é igual a

- a) 43,33%.
- b) 45,00%.
- c) 46,67%.
- d) 50,00%.
- e) 53,33%.



## GABARITO – INÉDITAS

### Distribuições de Frequência

1. LETRA C
2. LETRA A

3. LETRA D
4. LETRA D

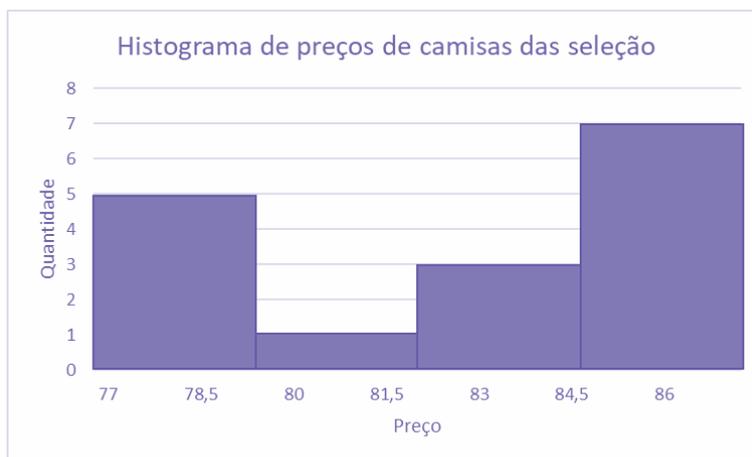
5. LETRA D
6. LETRA D
7. LETRA C



## LISTA DE QUESTÕES – INÉDITAS

### Representação Gráfica das Distribuições de Frequências

1. (INÉDITA/2022) O histograma a seguir mostra a quantidade de camisas da seleção brasileira, por faixas de preço, vendidas em determinada região.

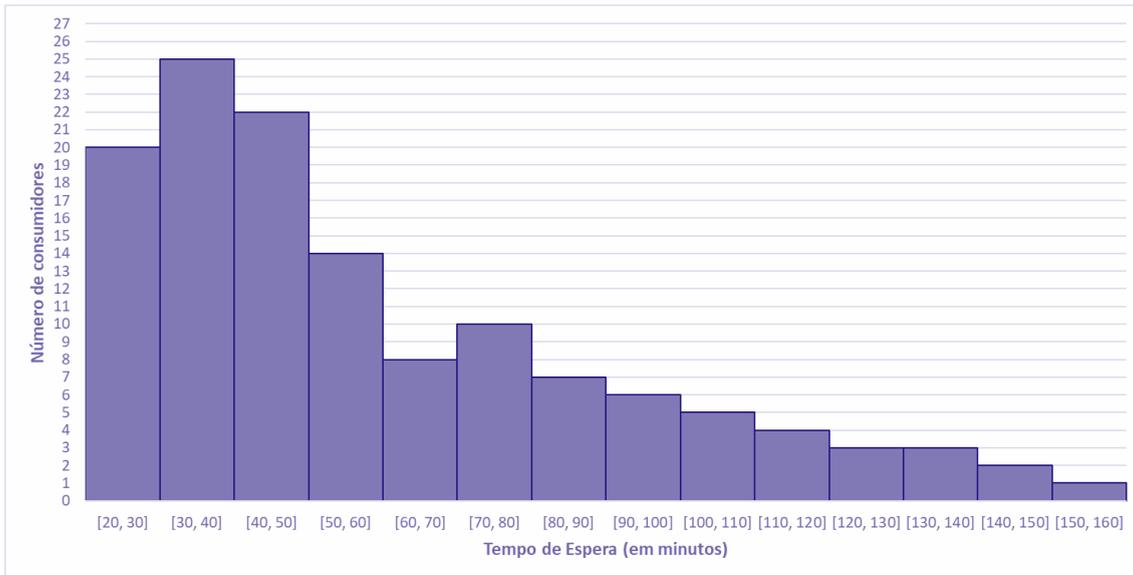


Sabendo-se que o menor valor encontrado foi de R\$ 77,00; que o maior valor foi de R\$ 87,00; e que os valores foram divididos em 4 classes de mesma amplitude, marque o conjunto de dados consistente com o histograma:

- a) 77, 77, 77, 78, 79, 84, 84, 84, 85, 85, 85, 85, 86, 86, 86, 87;
- b) 77, 77, 77, 78, 79, 84, 84, 84, 85, 85, 86, 86, 87;
- c) 77, 77, 77, 78, 79, 81, 83, 84, 84, 85, 85, 85, 86, 86, 86, 87;
- d) 77, 77, 78, 79, 80, 84, 84, 84, 85, 85, 85, 85, 86, 86, 86, 87;
- e) 77, 77, 77, 78, 79, 81, 81, 84, 84, 85, 85, 85, 86, 86, 86, 87.

2. (INÉDITA/2022) Os tempos de espera, em minutos, para o atendimento de 130 consumidores em uma central telefônica estão registrados no gráfico a seguir.





Com base no gráfico, afirmar que o tempo de espera de

- a) mais da metade dos consumidores foi superior a 1 hora.
- b) 37 consumidores foi entre 50 min e 80 min.
- c) 70% dos consumidores foi inferior a 1 hora.
- d) 10,0% dos consumidores foi entre 1 h 50 min e 2 h e 40 min.
- e) no mínimo 3 pessoas, foi superior a 2 h e 30 min.

3. (INÉDITA/2022) Após a coleta de dados, uma das formas de se iniciar a análise dos dados é por meio da geração e a avaliação de um histograma para uma variável selecionada. Em estudos climatológicos, por exemplo, o histograma pode ser utilizado para se estudar a precipitação mensal das chuvas de uma determinada região. Nesse contexto, o histograma é um gráfico em que são apresentadas as

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



## GABARITO – INÉDITAS

### Representação Gráfica das Distribuições de Frequências

1. LETRA C

2. LETRA B

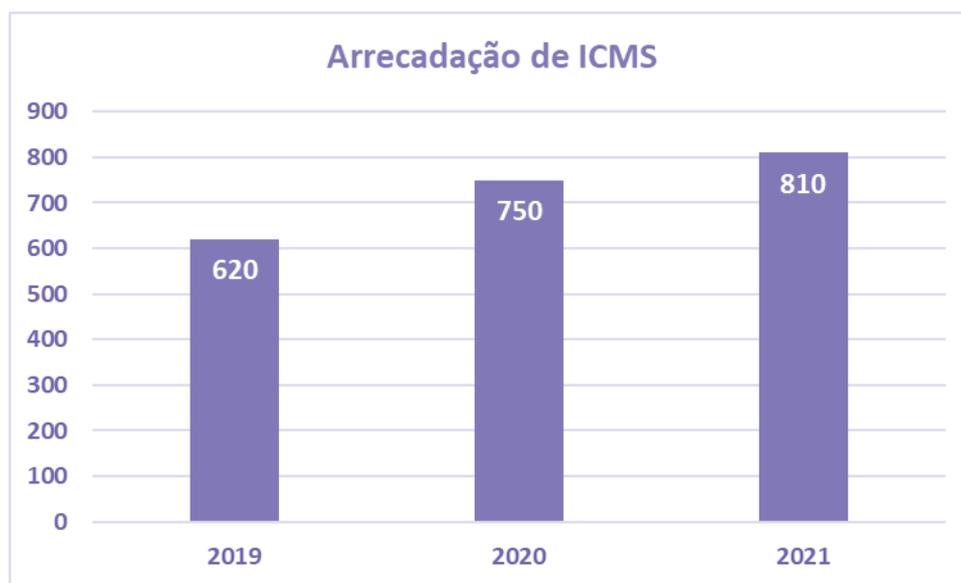
3. LETRA A



## LISTA DE QUESTÕES – INÉDITAS

### Outros Gráficos e Representações

1. (INÉDITA) O gráfico a seguir apresenta informações sobre a arrecadação de ICMS de determinado Estado brasileiro.



Com base nas informações do gráfico, avalie as afirmações a seguir.

- I. De 2019 a 2021, a arrecadação de ICMS desse Estado aumentou.
- II. De 2020 para 2021, houve um aumento de mais de 10% na arrecadação de ICMS desse Estado.
- III. É crescente a variação da arrecadação de ICMS desse Estado, de 2019 para 2020 e de 2020 para 2021.

Está correto o contido em

- a) apenas I.
- b) apenas I e II.
- c) apenas I e III.
- d) apenas II e III.
- e) I, II e III.

2. (INÉDITA) Durante uma pesquisa de campo, algumas variáveis se destacam.

- (i) número de indivíduos por faixa de massa corporal;



- (ii) percentuais do nível de escolaridade; e
- (iii) número de indivíduos por região.

Portanto, os tipos de gráficos considerados adequados a serem empregados em cada caso são, respectivamente:

- a) cartograma, linha e setores;
- b) setores, cartograma e barras;
- c) histograma, setores e linha;
- d) histograma, setores e cartograma;
- e) pictograma, setores e linha.

3. (INÉDITA) A respeito de gráficos estatísticos, avalie as afirmações a seguir.

I. Em um gráfico de colunas, as alturas de cada retângulo são proporcionais aos dados.

II. Um gráfico de barras apresenta a frequência absoluta (ou relativa) cumulativa.

III. Em um gráfico de linhas, os dados de categorias são distribuídos uniformemente ao longo do eixo vertical, e todos os valores são distribuídos igualmente ao longo do eixo horizontal.

Assinale a alternativa que contempla todas as assertivas corretas.

- a) I
- b) I e II
- c) II e III
- d) I, II e III
- e) I e III

4. (INÉDITA) A respeito das representações gráficas, avalie as afirmações a seguir.

I. Os gráficos de setores apresentam os dados em um círculo, no qual cada setor indica a frequência acumulada de um item observado.

II. Os gráficos de barras verticais apresentam os dados por meio de colunas retangulares em que a altura de cada coluna corresponde à frequência (absoluta ou relativa) dos itens observados.

III. Os pictogramas exibem os dados por meio de símbolos que possuem o mesmo tamanho e valor.



Assinale a alternativa que contempla apenas as assertivas corretas.

- a) I.
- b) I e II.
- c) II e III.
- d) I, II e III.
- e) I e III.

5. (INÉDITA) Acerca das representações gráficas, considere:

I. O gráfico de setores não é adequado para representar variáveis categóricas.

II. O gráfico de barras não é adequado para representar variáveis quantitativas contínuas.

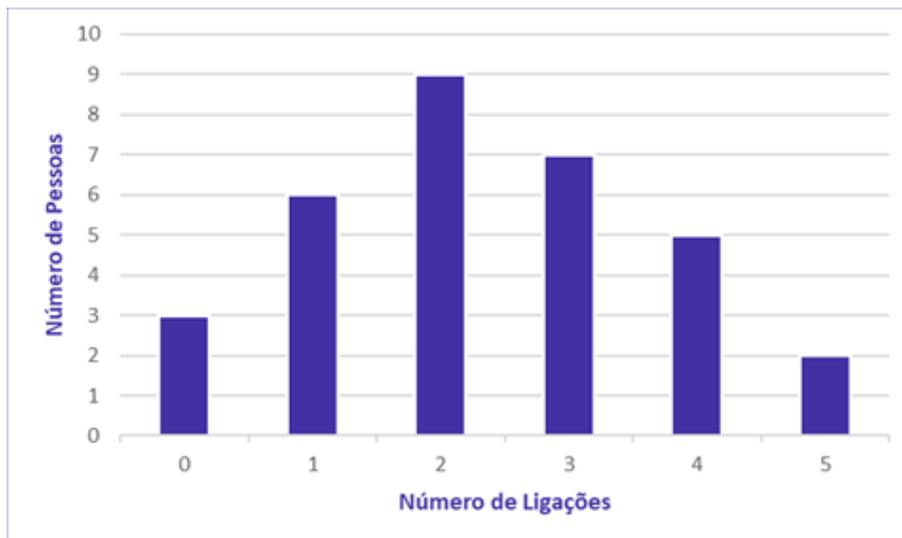
III. Histograma é um gráfico que apresenta a distribuição de frequências de uma variável por meio de retângulos justapostos, sendo a área de cada retângulo proporcional à frequência observada da classe correspondente.

Está correto o que consta apenas em:

- a) I.
- b) III.
- c) I e II.
- d) I e III.
- e) II e III.

6. (INÉDITA) Em uma empresa foi realizada uma auditoria para descobrir com que frequência os funcionários atendiam ligações pessoais no horário de trabalho de um único dia. O resultado está representado no gráfico a seguir.

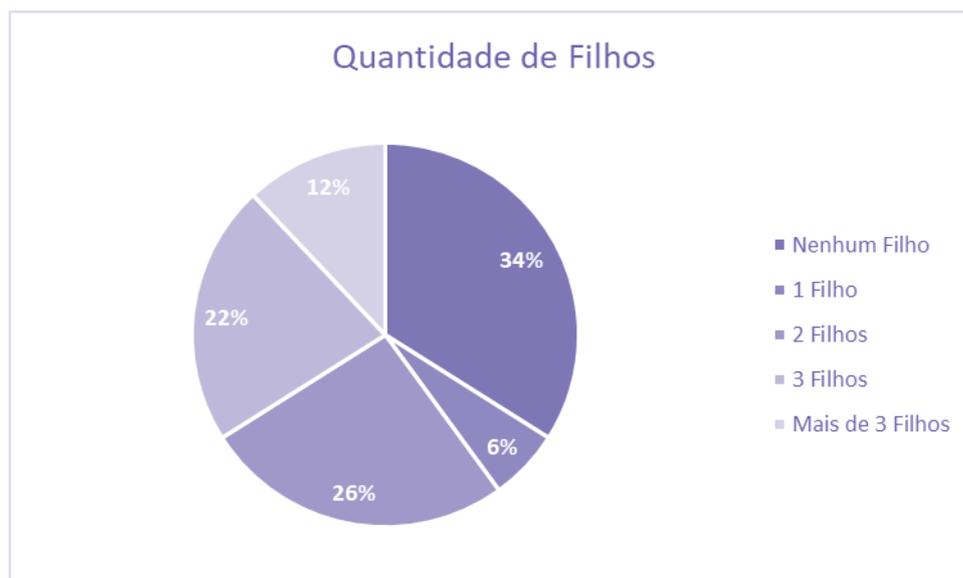




O número de pessoas que trabalham nessa empresa é

- a) 29.
- b) 30.
- c) 31.
- d) 32.
- e) 33.

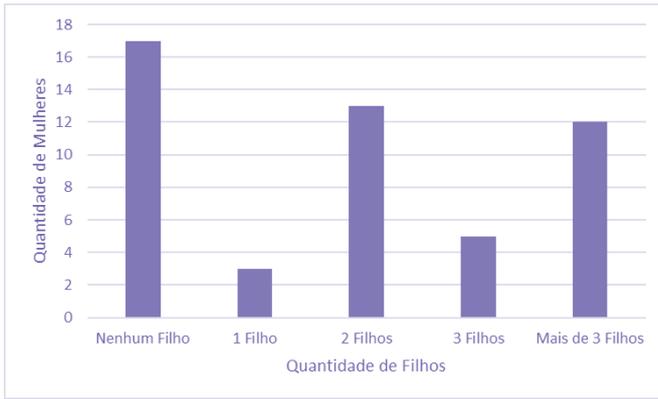
7. (INÉDITA) O gráfico de setor apresentado a seguir fornece os resultados de uma pesquisa realizada com 50 mulheres acerca da quantidade de filhos.



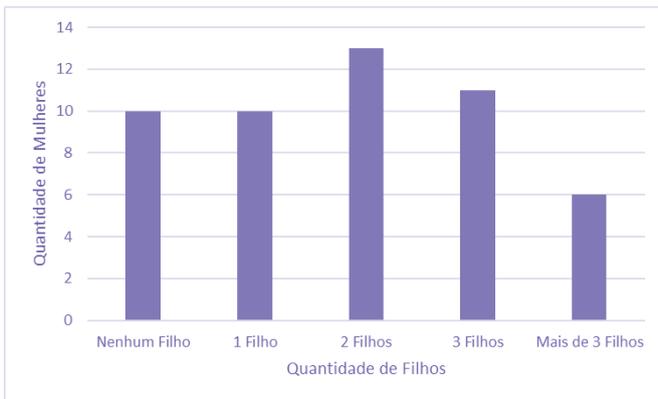
Assinale o gráfico de colunas que representa os dados apresentados acima.



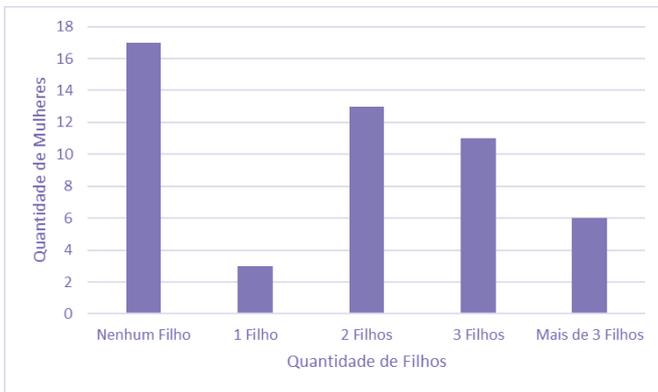
a)



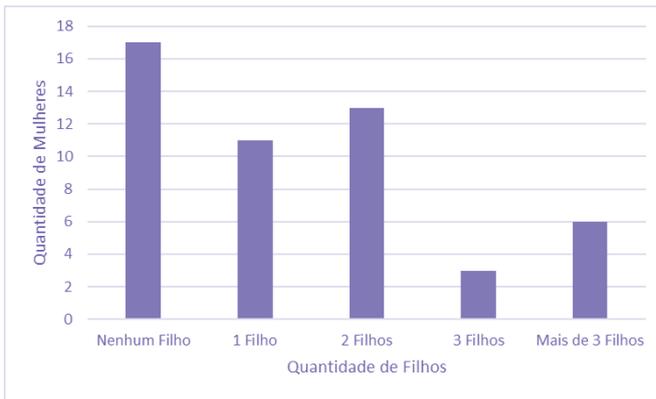
b)



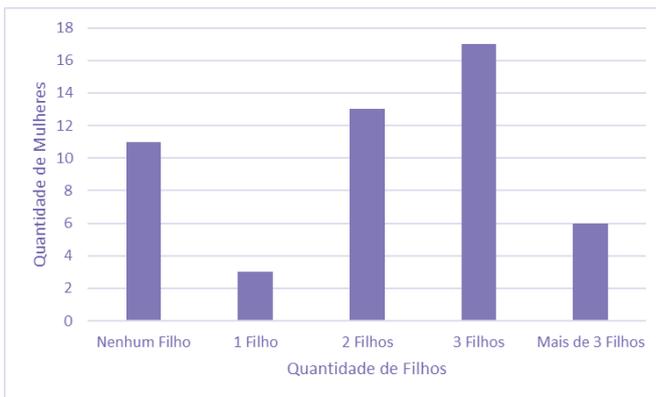
c)



d)



e)



8. (INÉDITA) Dentre os tipos de representações gráficas a seguir, o único que não pode ser utilizado para representar uma variável categórica é:

- a) histograma.
- b) diagrama de barras justapostas.
- c) diagrama de colunas.
- d) diagrama de setores.
- e) diagrama de colunas empilhadas.



## GABARITO – INÉDITAS

### Outros Gráficos e Representações

1. LETRA C
2. LETRA D
3. LETRA A

4. LETRA C
5. LETRA E
6. LETRA D

7. LETRA C
8. LETRA A



# 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.