Aula 00 (Prof. Djefferson Maranhão)

TCDF (Auditor de Controle Externo)

Estatística - 2021 (Pré-Edital)

Autor:

Equipe Exatas Estratégia Concursos

14 de Agosto de 2021

Sumário

1 – Introdução à Estatística 5

2 – Conceitos Iniciais 7

2.1 – População 7
2.2 – Amostra 7
2.3 – Censo 7
2.4 – Amostragem 7
2.5 – Parâmetros 8
2.6 – Estatística (ou estimador) 8

3 – Método Experimental x Método Estatístico 11

4 – Dados estatísticos 12

4.1 – Dados Brutos 12
4.2 – Rol 13

5 – Variáveis Estatísticas 13

5.1 – Variáveis Qualitativas 13
5.2 – Variáveis Quantitativas 14

6 – Séries Estatísticas 18

6.1 Séries Temporais (ou Cronológicas) 18
6.2 Séries Geográficas (ou Territoriais) 19
6.3 Séries Específicas 19
6.4 Séries Mistas (ou Compostas) 20

7 Distribuição de Frequências 22

7.1 Elementos de uma Distribuição de Frequências 24
7.1.1 Classe 25
7.1.2 Limite de Classe 26
7.1.3 Amplitude de um Intervalo de Classe 26
7.1.4 Amplitude Total 27
7.1.5 Ponto médio de classe 27
7.2 Frequência 29
7.2.1 Frequência Absoluta Simples 29
7.2.2 Frequência Absoluta Acumulada 30
7.2.3 Frequência Relativa Simples 32

7.2.4 Frequência Relativa Acumulada 34
7.3 Densidade de Frequência 35

8. Representações Gráficas das Distribuições de Frequência 40

8.1 – Gráfico de Hastes ou Bastões 40
8.2 – Histogramas 40
8.3 – Polígono de Frequências 41
8.4 – Curva de Frequências 42
8.5 – Ogiva (Polígono de Frequências Acumuladas) 45

9. Outros Gráficos e Representações 52

9.1 – Gráficos em Linhas 52
9.2 – Gráficos em Barras 53
9.3 – Gráficos em Colunas 54
9.4 – Gráfico em Setores 55
9.5 – Gráfico Polar 56
9.6 – Cartograma 57
9.7 – Pictogramas 57
9.8 – Gráfico de Dispersão 58
9.9 – Diagrama de Ramos e Folhas 58

Questões Comentadas 66

Lista de Questões 96

Questões Complementares Comentadas 114

Questões da FCC 114
Questões da FGV 130
Questões da FUNDATEC 141
Questões da VUNESP 145

Lista de Questões Complementares 158

Questões da FCC 158
Questões da FGV 164
Questões da FUNDATEC 168
Questões da VUNESP 170

Gabarito 177



Olá, amigos do Estratégia Concursos, tudo bem?

O meu nome é Djefferson Maranhão. Sou Graduado em Ciência da Computação pela Universidade Federal do Maranhão (UFMA) e em Engenharia Civil pela Universidade Estadual do Maranhão (UEMA). Desde 2015, sou Auditor da Controladoria Geral do Estado do Maranhão (2015 - 5º lugar). Antes, porém, exerci os cargos de Analista de Sistemas na UFMA (2010 - 1º lugar) e no TJ-MA (2011 - 1º lugar).

Estive na posição de vocês por muito tempo e sei o quanto a vida de um concurseiro é atribulada! São vários assuntos para se dominar em um curto espaço de tempo. Nesse sentido, quero compartilhar com vocês o que aprendi ao longo dos anos, para auxiliá-los nessa jornada rumo à aprovação.

Um grande abraço,

Djefferson Maranhão

1 – 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 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	ESTATÍSTICA PROBABILÍSTICA	ESTATÍSTICA INFERENCIAL
É responsável pela coleta, organização, descrição e resumo dos dados observados.	É responsável por estabelecer o modelo matemático adotado para explicar fenômenos aleatórios.	É responsável pela análise e interpretação dos dados.



(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.

2 – Conceitos Iniciais

Nessa seção, 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.

2.1 – 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.

2.2 – 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.

2.3 – 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.

2.4 – 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.

2.5 – 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, são **medidas que normalmente 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.

2.6 – Estatística (ou estimador)

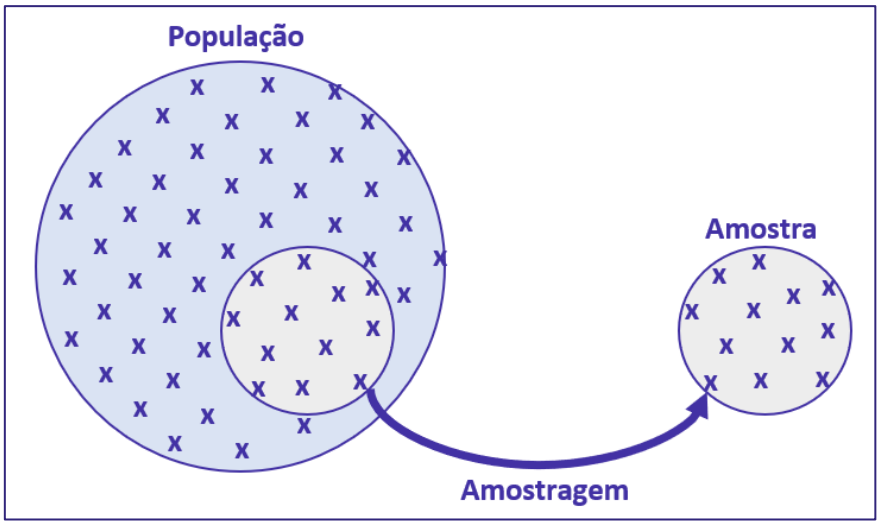
As **estatísticas são medidas numéricas obtidas de amostras representativas extraídas da população**. A partir das informações colhidas da amostra, as **estatísticas amostrais obtidas podem ser utilizadas para inferir ou tirar conclusões acerca dos parâmetros populacionais**, como a proporção de homens 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

POPULAÇÃO CONJUNTO de TODOS os elementos a serem estudados, que apresentam uma ou mais características em comum.	**CENSO** ESTUDO dos dados relativos a TODOS os ELEMENTOS de uma população.
AMOSTRA SUBCONJUNTO extraído DA POPULAÇÃO para análise, devendo ser representativo daquele grupo	**AMOSTRAGEM** PROCESSO que consiste na seleção criteriosa dos elementos a serem submetidos à investigação.
PARÂMETROS Descrições numéricas de características da POPULAÇÃO, que normalmente precisam ser estimadas.	**ESTATÍSTICAS** Medidas numéricas extraídas de AMOSTRAS representativas extraídas da população.







(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. Portanto, o censo requer a realização de um número maior de entrevistas.

Gabarito: Errado.

3 – MÉTODO EXPERIMENTAL X MÉTODO ESTATÍSTICO

Para a investigação de um fenômeno, temos a nossa disposição dois métodos: **experimental** e **estatístico**. De forma resumida, o **método experimental consiste em manter constantes as causas (fatores), com exceção de uma, que é variada para que seus efeitos sejam descobertos**. Contudo, nem sempre poderemos aplicar o método experimental, pois os fatores que afetam um fenômeno podem não permanecer constantes enquanto variamos a causa que nos interessa.

Por exemplo, para analisarmos uma queda nas vendas de uma empresa nacional que produz chocolates finos, teríamos que considerar vários fatores que não necessariamente permanecerão constantes durante toda a investigação do fenômeno, tais como a região, o fluxo de turistas na localidade; a temperatura média; o preço do concorrente; o mês de férias; etc.

Assim, diante da impossibilidade de manter as causas ou fatores constantes, **o método estatístico admite e registra todas as possíveis variações das causas presentes, procurando determinar a influência de cada fator no resultado final**. Dessa forma, o método estatístico busca descobrir relações entre os fatores, como, por exemplo, a influência da temperatura média e do fluxo de turistas na venda de chocolates finos.



ESQUEMATIZANDO

MÉTODO EXPERIMENTAL	MÉTODO ESTATÍSTICO
As CAUSAS são mantidas CONSTANTES, COM EXCEÇÃO DE UMA, que é VARIADA para que seus efeitos sejam descobertos.	Admite e REGISTRA TODAS AS POSSÍVEIS VARIAÇÕES DAS CAUSAS PRESENTES, procurando determinar a influência de cada fator no resultado.

4 – DADOS ESTATÍSTICOS

Os **dados estatísticos** constituem os valores resultantes da etapa de coleta de dados. Os dados referem-se a um conjunto de valores, que geralmente são organizados por variáveis (o que está sendo medido) 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**:

a) **dados univariados**: quando uma única observação de cada indivíduo é registrada. Por exemplo: peso;
b) **dados bivariados**: quando duas observações de cada indivíduo são registradas. Por exemplo: peso e altura;
c) **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**.

4.1 – 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 Concursos, em que a escolha dos alunos ocorreu de forma aleatória, não havendo qualquer ordenação de valores.

Aluno	Tempo	Aluno	Tempo	Aluno	Tempo	Aluno	Tempo	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.

4.2 – 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**

5 – 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 Concursos estudam por dia, então nossa variável é o número de horas estudadas por dia. As variáveis estatísticas podem ser classificadas, inicialmente, em duas categorias: **qualitativas** e **quantitativas**.

5.1 – Variáveis Qualitativas

As **variáveis qualitativas** 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 (ou categórica)**, 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).

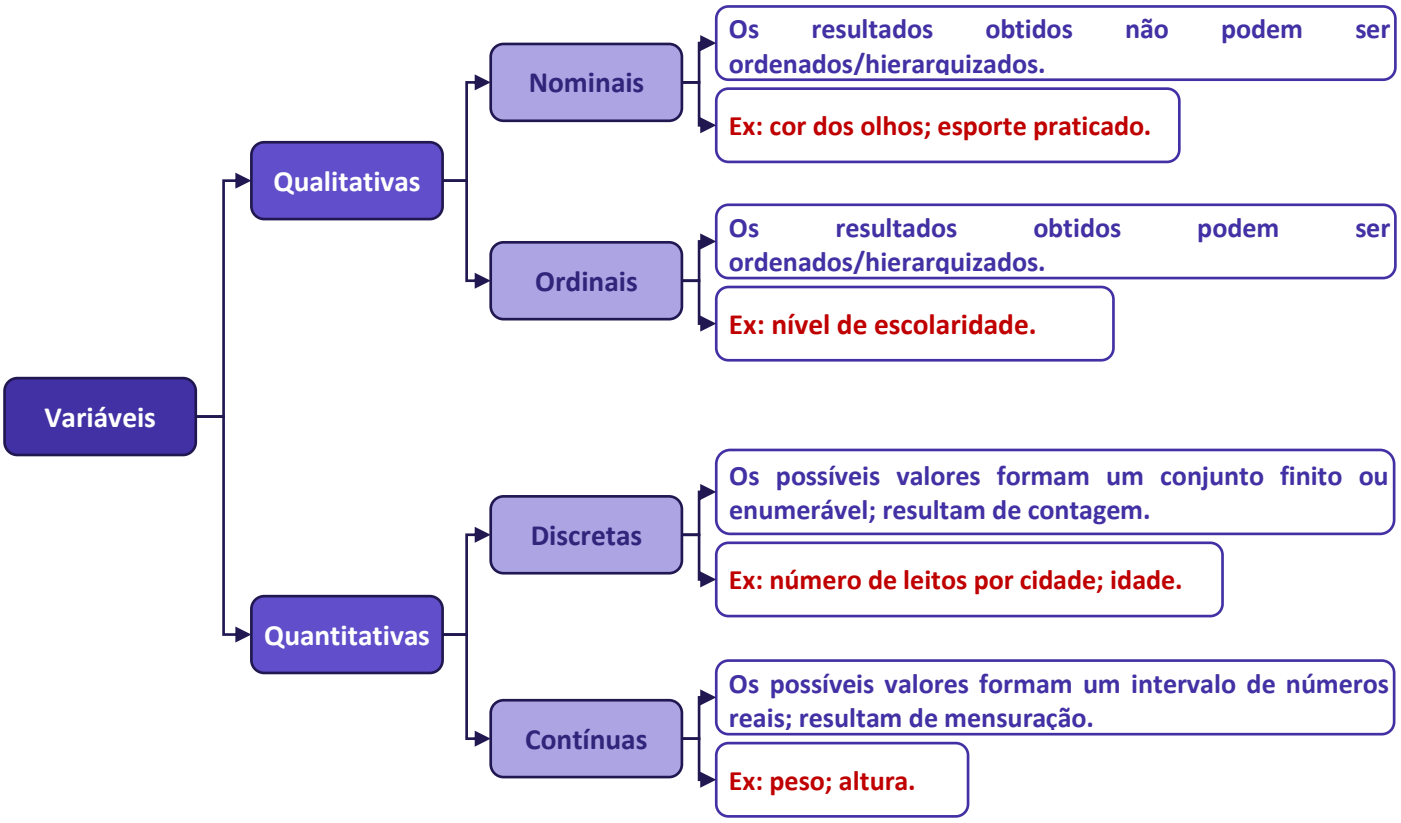
5.2 – Variáveis Quantitativas

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**. 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**. Por exemplo, a altura dos moradores de uma determinada cidade.





HORA DE PRATICAR!

(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.

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

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

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

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

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

Comentários:

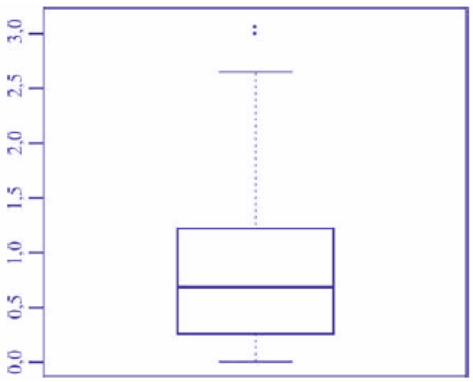
As variáveis podem ser classificadas em:

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

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**.

Gabarito: D.

(CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

X representa uma variável qualitativa ordinal.

Comentários:

As variáveis podem ser classificadas em:

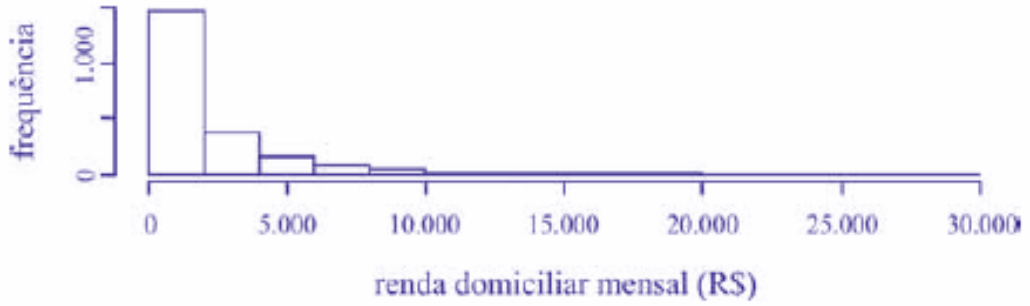
a) variáveis quantitativas: são as características que podem ser medidas em uma escala quantitativa, isto é, numérica, podendo ser contínuas ou discretas.

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. Podem ser nominais ou ordinais.

Analisando as informações dadas na questão, percebemos que X pode adotar valores que vão de 0 (mínimo) a 3,10 (máximo). Logo, como X assume valores numéricos, dizemos que ela é uma variável quantitativa.

Gabarito: Errado.

(CESPE/TELEBRAS/2015)



Uma empresa coletou e armazenou em um banco de dados diversas informações sobre seus clientes, entre as quais estavam o valor da última fatura vencida e o pagamento ou não dessa fatura. Analisando essas informações, a empresa concluiu que 15% de seus clientes estavam inadimplentes. A empresa recolheu ainda dados como a unidade da Federação (UF) e o CEP da localidade em que estão os clientes. Do conjunto de todos os clientes, uma amostra aleatória simples constituída por 2.175 indivíduos prestou também informações sobre sua renda domiciliar mensal, o que gerou o histograma apresentado. Com base nessas informações e no histograma, julgue o item a seguir.

O CEP da localidade dos clientes e o valor da última fatura vencida são variáveis quantitativas.

Comentários:

O CEP, embora possua uma representação numérica, não exprime quantidades e, portanto, não é uma variável quantitativa. Esse código número apenas qualifica as ruas de um determinado endereço, podendo ser classificada como uma variável qualitativa.

O valor da última fatura vencida, por sua vez, é sim uma variável quantitativa.

Gabarito: Errado.

(CESPE/TELEBRAS/2015) Roberto comprou, por R$ 2.800,00, rodas de liga leve para seu carro, e, ao estacionar no shopping, ficou indeciso sobre onde deixar o carro, pois, caso o coloque no estacionamento público, correrá o risco de lhe roubarem as rodas, ao passo que, caso o coloque no estacionamento privado, terá de pagar R$ 70,00, com a garantia de que eventuais prejuízos serão ressarcidos pela empresa administradora.

Considerando que p seja a probabilidade de as rodas serem roubadas no estacionamento público, que X seja a variável aleatória que representa o prejuízo, em reais, ao deixar o carro no estacionamento público, e que Y seja a variável aleatória que representa o valor, em reais, desembolsado por Roberto ao deixar o carro no estacionamento pago, julgue o item subsequente.

A variável aleatória Y é contínua.

Comentários:

As variáveis quantitativas podem ser classificadas em:

a) variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua, nesse caso, valores fracionários fazem sentido;

b) variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros.

No problema em questão, a variável aleatória Y pode assumir o valor R$ 70,00, se Roberto deixar o carro no estacionamento pago; ou o valor R$ 0,00, se Roberto não deixar o carro no estacionamento pago. Portanto, Y é uma variável aleatória discreta.

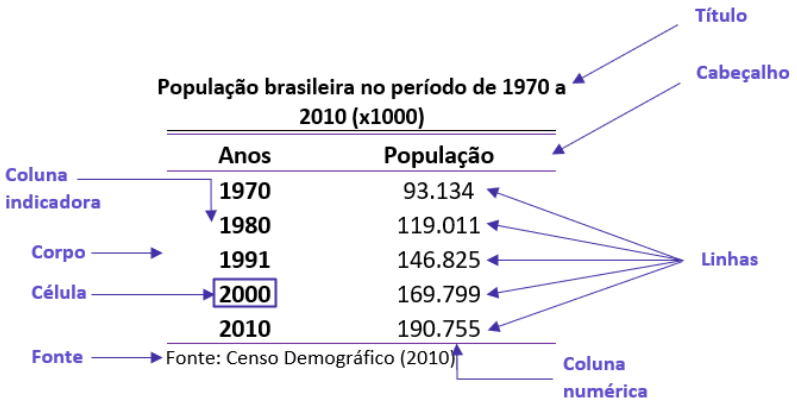
Gabarito: Errado.

6 – 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:

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



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 intepretação dos valores representativos do fenômeno (ser claro); e transmitir a verdade sobre o fenômeno (ser 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.

6.1 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)

6.2 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.

População Urbana em 2010 (x1000)

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

Fonte: Censo Demográfico (2010)

6.3 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.

População Urbana e Rural em 2010 (x1000)

Zona	População
Urbana	93.134
Rural	119.011
Total	**190.755**

Fonte: Censo Demográfico (2010)

6.4 Séries Mistas (ou Compostas)

Muitas vezes, temos a necessidade de apresentar, em uma única tabela, a variação de valores de mais de uma variável, isto é, fazer uma combinação de 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)

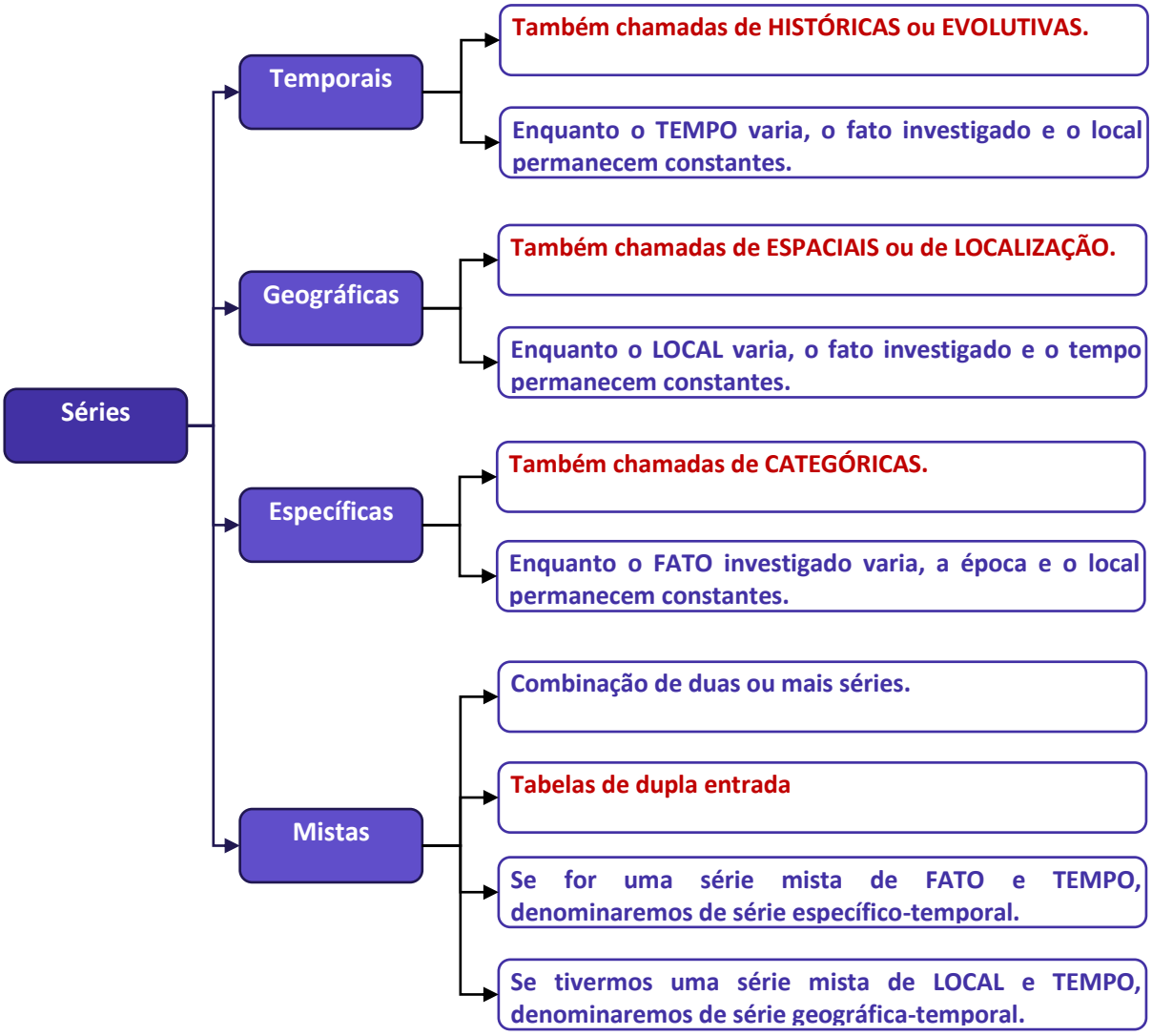
Anos	Regiões				
	N	NE	SE	S	CO
1970	3.603	28.111	39.850	16.496	5.072
1980	5.880	34.815	51.737	19.031	7.545
1991	10.030	42.497	62.740	22.129	9.427

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.





7 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 Concursos, 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	Aluno	Tempo	Aluno	Tempo	Aluno	Tempo	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, esse dado fornece pouca informação ao leitor, sendo necessário organizar os dados, com o propósito de aumentar sua capacidade de informação. A simples organização dos dados em um **rol crescente** já aumenta bastante a capacidade de informação. Com os dados organizados em rol, conseguimos verificar que o menor tempo observado foi 85 minutos, e o maior, 175 minutos, o que nos fornece uma **amplitude total** (AT = 175 - 85 = 70) de variação da ordem de 70 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, ao invés de repeti-los. **O número de ocorrências de um determinado valor recebe o nome de frequência (é o número de vezes que cada tempo de estudo aparece na tabela). 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.	Tempo (min)	Freq.	Tempo (min)	Freq.	Tempo (min)	Freq.
85	**1**	114	**1**	135	**3**	158	**1**
89	**1**	115	**2**	137	**3**	159	**1**
96	**1**	123	**2**	142	**1**	161	**2**
98	**1**	124	**1**	143	**2**	165	**1**
99	**1**	126	**3**	148	**1**	168	**1**
103	**1**	127	**1**	153	**1**	170	**1**
104	**1**	128	**1**	154	**1**	171	**3**
105	**1**	129	**2**	155	**1**	173	**1**
113	**1**	134	**1**	157	**1**	175	**2**

Nesses casos, quando a variável é contínua, o mais recomendável é agrupar os valores por intervalos de classe. Desse modo, ao invés 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	$a \leq x \leq b$	$a \mapsto b$	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 o apanhado geral das características de um conjunto de dados, sem adentrar em detalhes de casos particulares.

7.1 Elementos de uma Distribuição de Frequências

Agora, analisaremos cada um dos elementos de uma distribuição de frequências de forma mais detalhada. Tomaremos como referência a tabela apresentada anteriormente:

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

7.1.1 Classe

As classes são os intervalos nos quais o fenômeno é subdividido. Isto é, 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
		$\boldsymbol{n = 50}$

Existem duas maneiras de determinar o número de classes, k, em função do número de dados da tabela, n. A primeira consiste em utilizar a fórmula de Sturges:

$$\boldsymbol{k = 1 + 3{,}3 \times log\, n}$$

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

$$\boldsymbol{k = \sqrt{n}}$$

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:

$$k = \sqrt{n}$$

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

7.1.2 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

7.1.3 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 ($h = l_{sup} - l_{inf}$)
$85 \leq x < 100$	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.

7.1.4 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:

$$AT = l_{máx} - l_{mín}$$

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:

$$AT = l_{máx} - l_{mín}$$

$$AT = 190 - 85 = 105$$

7.1.5 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**.

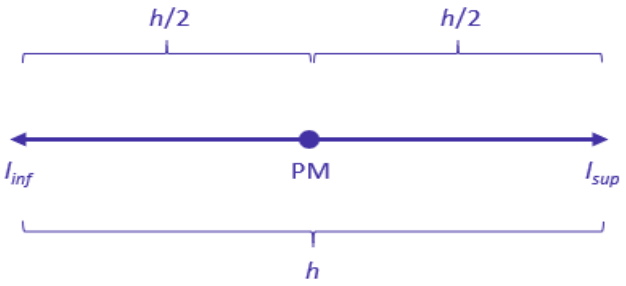
$$PM = \frac{(l_{inf} + l_{sup})}{2}$$

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

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

Veja que os pontos médios formaram uma progressão aritmética, vez que 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, **podemos afirmar que, 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:

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

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

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

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

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

7.2 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:

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

7.2.1 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	$85 \vdash 100$	5
2	$100 \vdash 115$	5
3	$115 \vdash 130$	12
4	$130 \vdash 145$	10
5	$145 \vdash 160$	7
6	$160 \vdash 175$	9
7	$175 \vdash 190$	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	$85 \vdash 100$	5
2	$100 \vdash 115$	5
3	$115 \vdash 130$	12
4	$130 \vdash 145$	10
5	$145 \vdash 160$	7
6	$160 \vdash 175$	9
7	$175 \vdash 190$	2
		$\sum_{i=1}^{7} f_i = 50$

7.2.2 Frequência Absoluta Acumulada

A **frequência absoluta acumulada crescente** (f_{ac}) é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: $f_{ac_4} = f_1 + f_2 + f_3 + f_4 = 5 + 5 + 12 + 10 = 32$, o que significa existem 32 alunos que estudam diariamente entre 85 e 145 minutos (limite superior da quarta classe).

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

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 o total das frequências de todos os valores superiores ao limite inferior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: $f_{ad_4} = f_4 + f_5 + f_6 + f_7 = 10 + 7 + 9 + 2 = 28$, o que significa existem 28 alunos que estudam diariamente entre 130 (limite inferior da quarta classe) e 190 minutos.

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

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	$85 \vdash 100$	5	50
2	$100 \vdash 115$	5	45
3	$115 \vdash 130$	12	40
4	$130 \vdash 145$	10	28
5	$145 \vdash 160$	7	18
6	$160 \vdash 175$	9	11
7	$175 \vdash 190$	2	2
		$\sum_{i=1}^{7} f_i = 50$	

7.2.3 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):

$$\boldsymbol{F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}}$$

Em nosso exemplo, as frequências relativas são:

i	Tempos (min)	Frequência Absoluta (f_i)	Frequência Relativa (F_i)
1	$85 \vdash 100$	5	$F_1 = \frac{5}{50} = 0{,}10$
2	$100 \vdash 115$	5	$F_2 = \frac{5}{50} = 0{,}10$
3	$115 \vdash 130$	12	$F_3 = \frac{12}{50} = 0{,}24$
4	$130 \vdash 145$	10	$F_4 = \frac{10}{50} = 0{,}20$
5	$145 \vdash 160$	7	$F_5 = \frac{7}{50} = 0{,}14$
6	$160 \vdash 175$	9	$F_6 = \frac{9}{50} = 0{,}18$
7	$175 \vdash 190$	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	$85 \vdash 100$	5	$F_1 = \frac{5}{50} \times 100\% = 10\%$
2	$100 \vdash 115$	5	$F_2 = \frac{5}{50} \times 100\% = 10\%$
3	$115 \vdash 130$	12	$F_3 = \frac{12}{50} \times 100\% = 24\%$
4	$130 \vdash 145$	10	$F_4 = \frac{10}{50} \times 100\% = 20\%$
5	$145 \vdash 160$	7	$F_5 = \frac{7}{50} \times 100\% = 14\%$
6	$160 \vdash 175$	9	$F_6 = \frac{9}{50} \times 100\% = 18\%$
7	$175 \vdash 190$	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%:

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

Podemos incluir essa informação na representação tabular:

i	Tempos (min)	Frequência Absoluta (f_i)	Frequência Relativa (F_i)
1	$85 \vdash 100$	5	10%
2	$100 \vdash 115$	5	10%
3	$115 \vdash 130$	12	24%
4	$130 \vdash 145$	10	20%
5	$145 \vdash 160$	7	14%
6	$160 \vdash 175$	9	18%
7	$175 \vdash 190$	2	4%
		$\sum_{i=1}^{7} f_i = 50$	$\sum_{i=1}^{7} F_i = 100\%$

7.2.4 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: $F_{ac_4} = F_1 + F_2 + F_3 + F_4 = 10\% + 10\% + 24\% + 20\% = 64\%$, o que significa que 64% dos alunos estudam diariamente entre 85 e 145 minutos (limite superior da quarta classe).

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

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: $F_{ad_4} = F_4 + F_5 + F_6 + F_7 = 20\% + 14\% + 18\% + 4\% = 56\%$, o que significa que 56% dos alunos estudam diariamente entre 130 (limite inferior da quarta classe) e 190 minutos.

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

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	$85 \vdash 100$	5	10%	100%
2	$100 \vdash 115$	5	10%	90%
3	$115 \vdash 130$	12	24%	80%
4	$130 \vdash 145$	10	20%	56%
5	$145 \vdash 160$	7	14%	36%
6	$160 \vdash 175$	9	18%	22%
7	$175 \vdash 190$	2	4%	4%
		$\sum_{i=1}^{7} f_i = 50$	$\sum_{i=1}^{7} F_i = 100\%$	

7.3 Densidade de Frequência

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

$$densidade = \frac{frequência\ absoluta\ ou\ relativa}{amplitude}$$

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

Para o nosso exemplo, as densidades de frequência são:

i	Tempos (min)	Frequência Absoluta (f_i)	Densidade de Frequência (d)
1	$85 \vdash 100$	5	$d_1 = \frac{f_1}{h_1} = \frac{5}{15} = 0{,}33$
2	$100 \vdash 115$	5	$d_2 = \frac{f_2}{h_2} = \frac{5}{15} = 0{,}33$
3	$115 \vdash 130$	12	$d_3 = \frac{f_3}{h_3} = \frac{12}{15} = 0{,}80$
4	$130 \vdash 145$	10	$d_4 = \frac{f_4}{h_4} = \frac{10}{15} = 0{,}66$

5	$145 \vdash 160$	7	$d_5 = \frac{f_5}{h_5} = \frac{7}{15} = 0{,}46$
6	$160 \vdash 175$	9	$d_6 = \frac{f_6}{h_6} = \frac{9}{15} = 0{,}60$
7	$175 \vdash 190$	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.	$k = 1 + 3{,}3 \times log\ n$ $k = \sqrt{n}$
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).	$AT = l_{máx} - l_{mín}$ $AT = h \times k$
Ponto Médio	Média aritmética simples dos valores extremos de uma classe.	$PM = \frac{(l_{inf} + l_{sup})}{2}$ $PM = l_{inf} + \frac{h}{2}$ $PM = l_{sup} - \frac{h}{2}$
Frequência Absoluta Simples	Número de observações correspondentes a uma determinada classe ou a um determinado valor.	f_i
Frequência Absoluta Acumulada	Total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe	$f_{ac_i} = f_1 + f_2 + f_3 + \cdots + f_i$
Frequência Relativa Simples	Proporção de dados existentes em uma determinada classe.	$F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$

Frequência Relativa Acumulada	Proporção de valores inferiores ao limite superior do intervalo de uma dada classe.	$F_{ac_i} = F_1 + F_2 + F_3 + \cdots + F_i$
Densidade de Frequência	Quociente entre a frequência da classe (absoluta ou relativa) e sua amplitude	$d = \frac{f}{h}$



(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 simples absoluta do primeiro elemento é igual a 20%.

Gabarito: B.

(CESPE/Polícia Federal/2018)

	DIA				
	1	2	3	4	5
X (quantidade diária de drogas apreendidas, em kg)	10	22	18	22	28

Tendo em vista que, diariamente, a Polícia Federal apreende uma quantidade X, em kg, de drogas em determinado aeroporto do Brasil, e considerando os dados hipotéticos da tabela precedente, que apresenta os valores observados da variável X em uma amostra aleatória de 5 dias de apreensões no citado aeroporto, julgue o item.

A tabela em questão descreve a distribuição de frequências da quantidade de drogas apreendidas nos cinco dias que constituem a amostra.

Comentários:

A tabela apenas registra as quantidades de drogas apreendidas por dia. Para que ela configurasse uma distribuição de frequências da quantidade de drogas apreendidas nos cinco dias, era necessário que cada valor estivesse associado a uma frequência, o que não ocorreu. Uma distribuição de frequência dos dados acima seria assim:

Valor	Frequência
10	1
18	1
22	2
28	1

Gabarito: Errado.

(CESPE/BACEN/2013)

2 4 8 4 8 1 2 32 12 1 5 7 5 5 3 4 24 19 4 14

Os dados mostrados acima representam uma amostra, em minutos, do tempo utilizado na armazenagem de formulários no almoxarifado central de certa instituição por diversos funcionários.

Com base nesses dados, julgue o próximo item.

A distribuição de frequência acumulada para tempo de armazenagem observado na amostra inferior a 8 minutos é igual a 13, o que corresponde a uma frequência relativa superior a 0,60.

Comentários:

Não precisamos construir a distribuição de frequências para responder à questão. Sabemos que o total de observações é 20, pois é esse o número de termos.

Vamos destacar as observações que são inferiores a 8:

2 4 8 **4** 8 **1 2** 32 12 **1 5 7 5 5 3 4** 24 19 **4** 14

São 13 números inferiores a 8. Portanto, a frequência acumulada para um tempo menor do que 8 é 13.

Para calcular a frequência relativa, basta dividir a frequência absoluta pelo total de observações:

$$\frac{8}{13} = 0{,}61$$

Logo, a frequência relativa acumulada é maior do que 0,60.

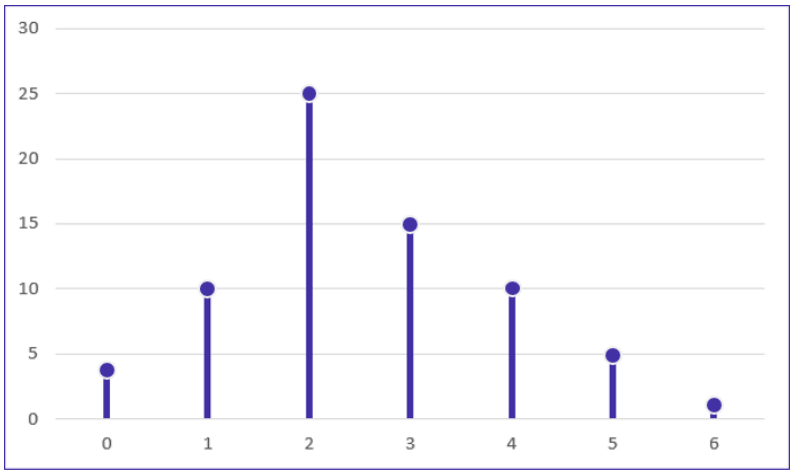
Gabarito: Certo.

8. Representações Gráficas das Distribuições de Frequência

8.1 – 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



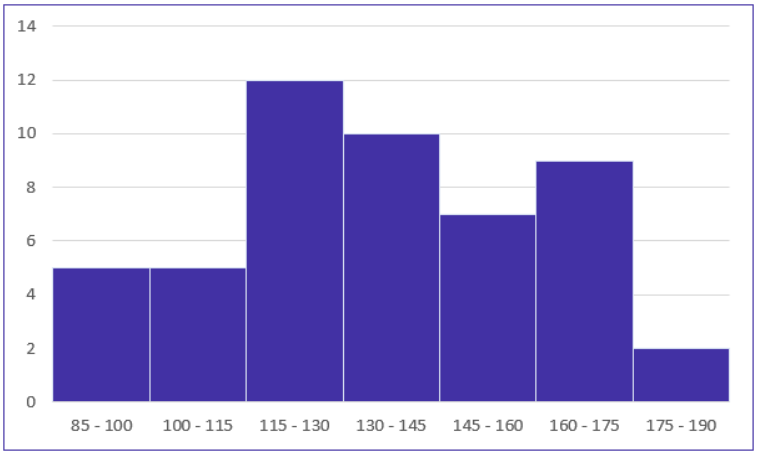
Repare que podemos reconstruir facilmente a tabela de frequências a partir do gráfico de hastes. De igual modo, conhecendo a tabela de frequências, podemos construir rapidamente o gráfico de hastes.

8.2 – 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.

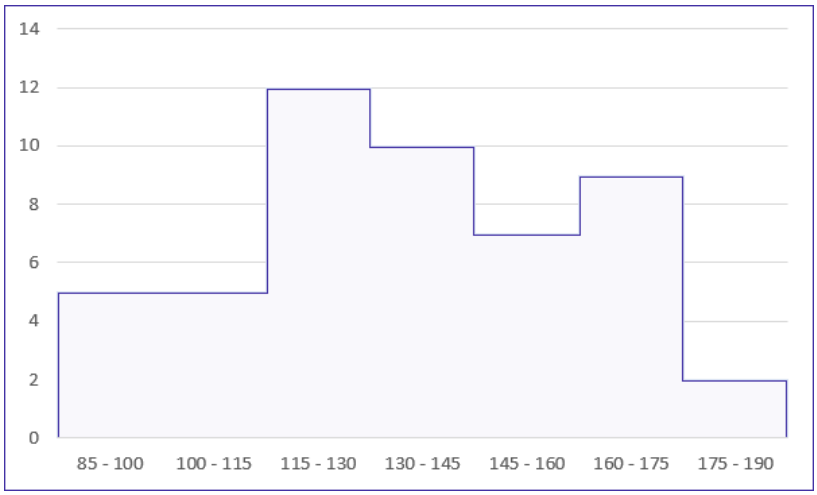
A quantidade de retângulos em um histograma é equivalente ao número de intervalos de classe. A largura de cada retângulo deve ser igual à amplitude do intervalo de classe, enquanto a altura precisa ser proporcional à frequência do intervalo de classe. Além disso, a área do histograma é proporcional ao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



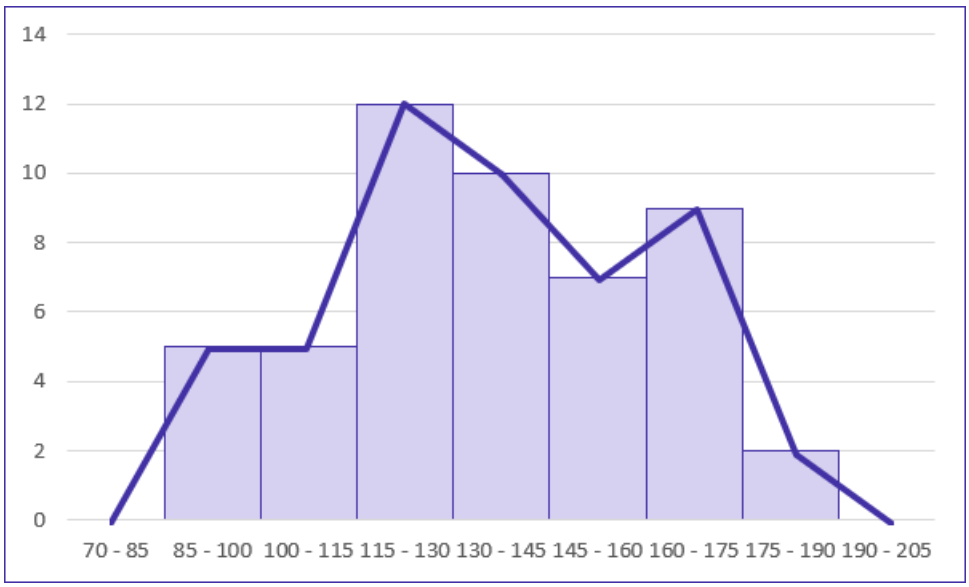
Por fim, **o histograma pode ocasionar um certo nível de perda de informações, pois os elementos da distribuição de frequência não são representados de forma individualizada, mas sim por meio de suas classes**.

8.3 – 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



8.4 – 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.

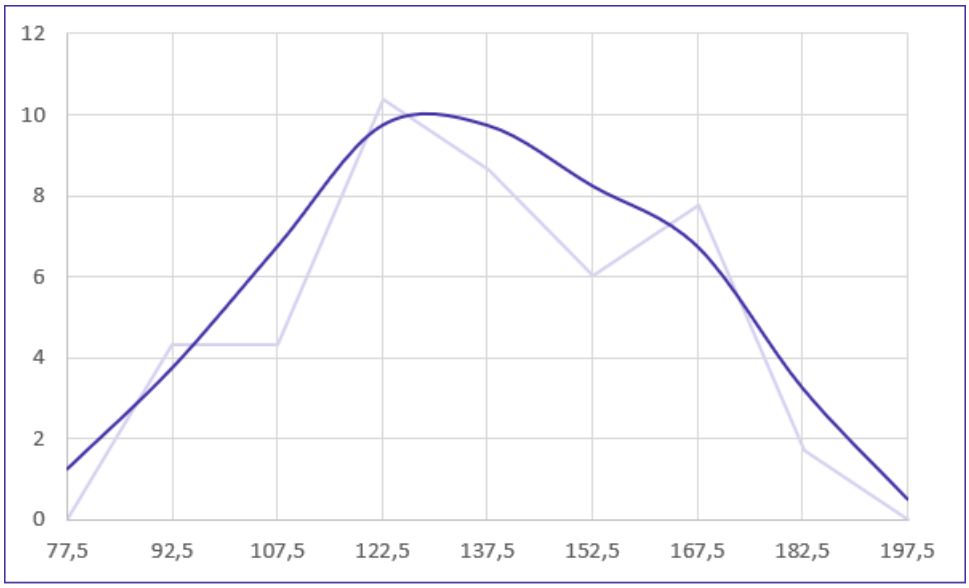
O processo de polimento é realizado por meio da seguinte fórmula:

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

Em que fc_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.

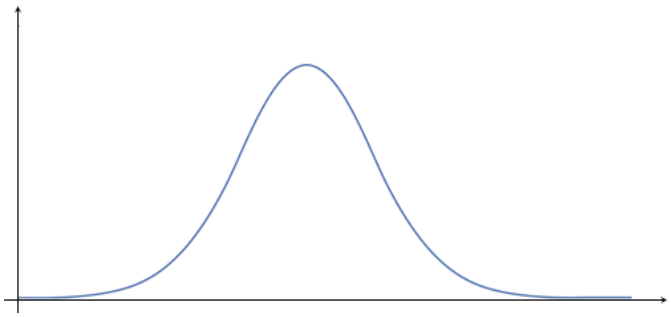
Vejamos como utilizá-la:

Tempo médio (X_i)	Ponto médio (PM_i)	Frequência (f_i)	Frequência Calculada (fc_i)
$70 \leq x < 85$	77,5	0	$fc_0 = \frac{0 + 2 \times 0 + 5}{4} = 1{,}25$
$85 \leq x < 100$	92,5	5	$fc_1 = \frac{0 + 2 \times 5 + 5}{4} = 3{,}75$
$100 \leq x < 115$	107,5	5	$fc_2 = \frac{5 + 2 \times 5 + 12}{4} = 6{,}75$
$115 \leq x < 130$	122,5	12	$fc_3 = \frac{5 + 2 \times 12 + 10}{4} = 9{,}75$
$130 \leq x < 145$	137,5	10	$fc_4 = \frac{12 + 2 \times 10 + 7}{4} = 9{,}75$
$145 \leq x < 160$	152,5	7	$fc_5 = \frac{10 + 2 \times 7 + 9}{4} = 8{,}75$
$160 \leq x < 175$	167,5	9	$fc_6 = \frac{7 + 2 \times 9 + 2}{4} = 6{,}75$
$175 \leq x < 190$	182,5	2	$fc_7 = \frac{9 + 2 \times 2 + 0}{4} = 3{,}25$
$190 \leq x < 205$	197,5	0	$fc_8 = \frac{2 + 2 \times 0 + 0}{4} = 0{,}50$

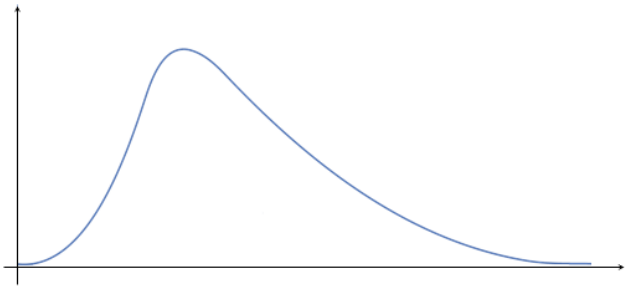


As curvas de frequências podem assumir as seguintes formas características:

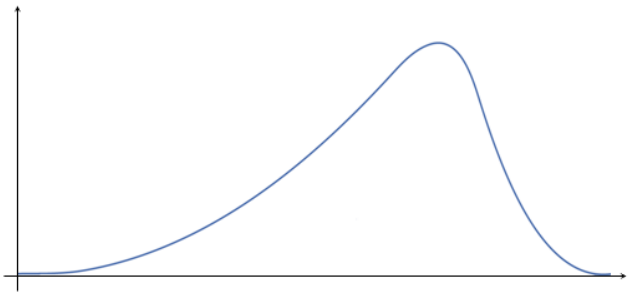
a) **curvas em forma de sino**: são curvas que apresentam concentração de valores em torno da região central da distribuição. Tais curvas podem ser simétricas ou assimétricas. Quando assimétricas, as curvas ainda podem apresentar uma cauda mais alongada à esquerda (assimetria à esquerda) ou mais alongada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



Curva Simétrica.

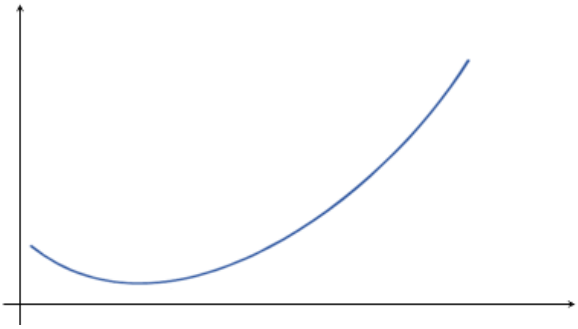


Curva assimétrica à direita.

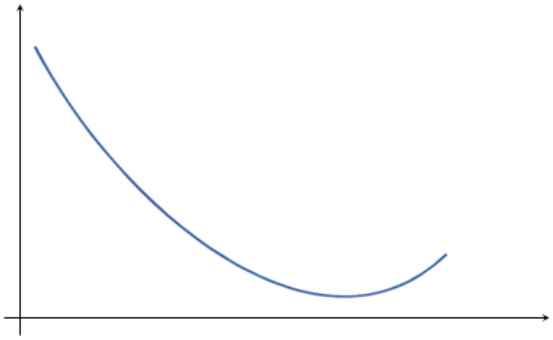


Curva assimétrica à esquerda.

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:

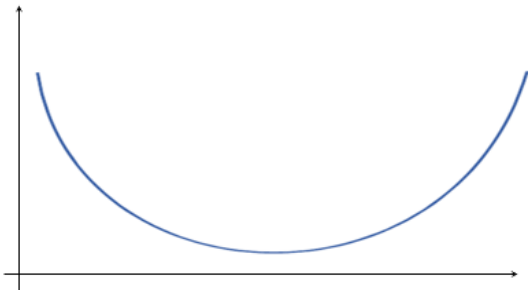


Curva em formato de J.



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.

8.5 – 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

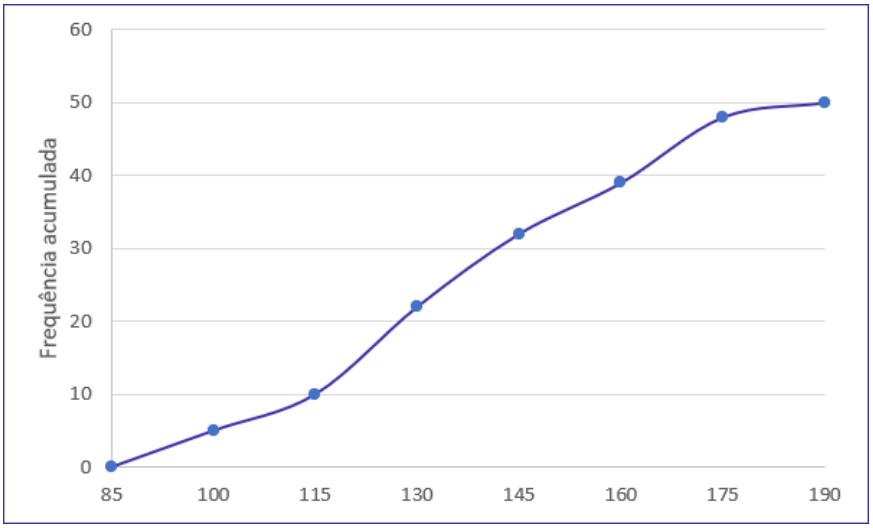
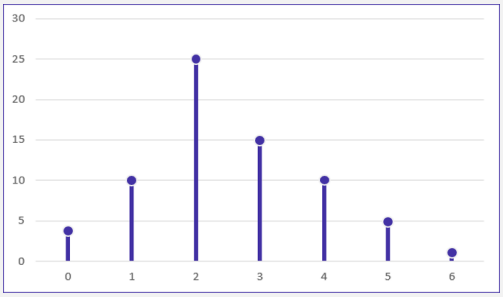
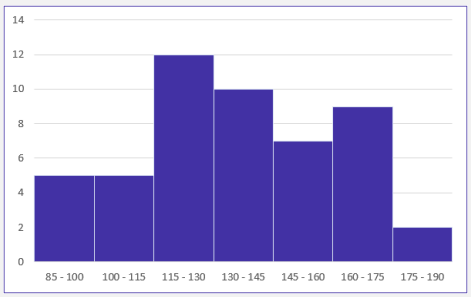
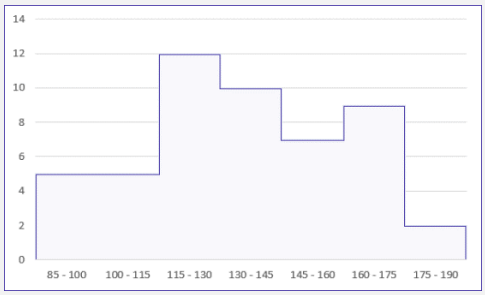




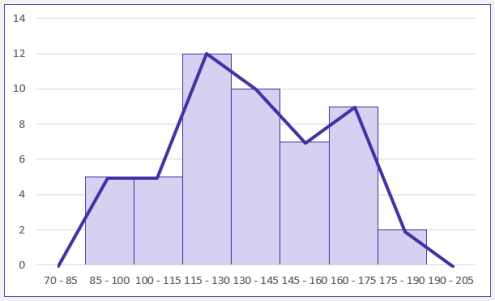
Gráfico	Definição
	O **gráfico de hastes ou bastões** é muito utilizado para representar dados não agrupados em classes, o que normalmente ocorre com dados discretos.



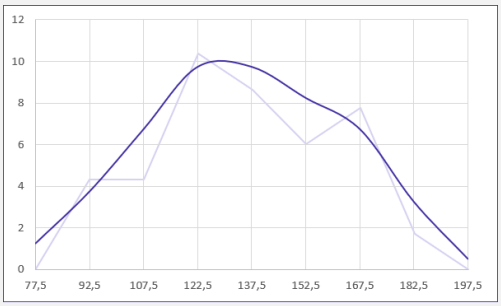
O **histograma** é um gráfico destinado a representar **dados agrupados em classe**, sendo composto por um conjunto de **retângulos contíguos (justapostos).**



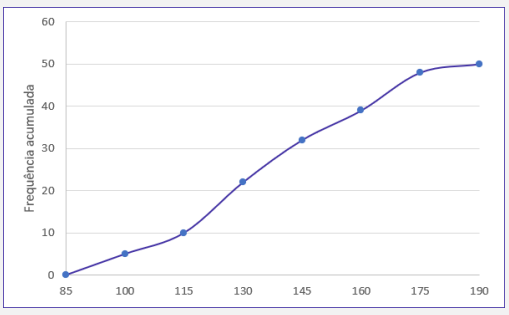
A **poligonal característica** é construída utilizando apenas os contornos do histograma.



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.



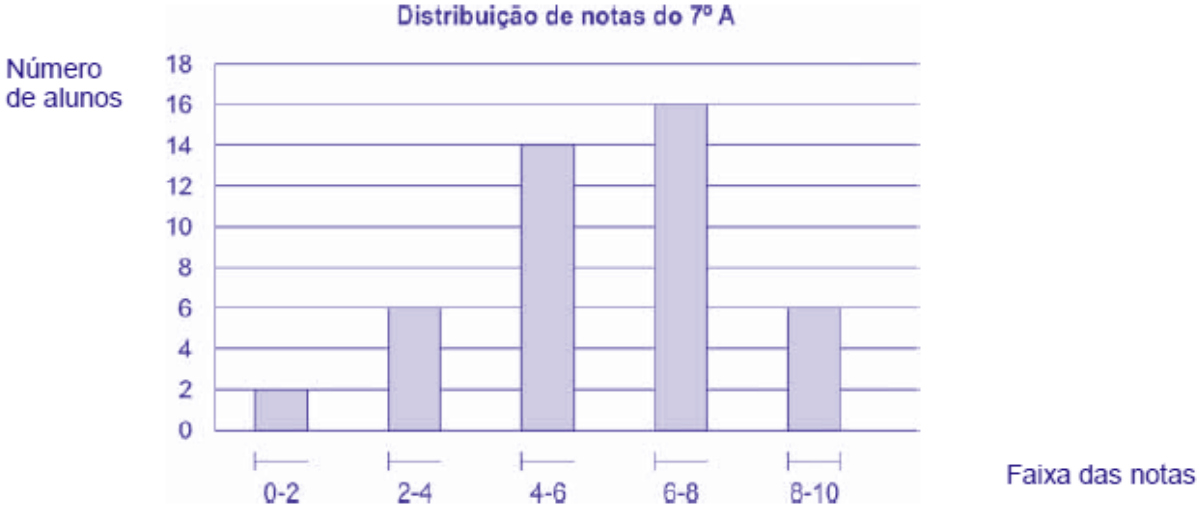
A **curva de frequências** é obtida a partir do polimento de um polígono de frequências



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

HORA DE PRATICAR!

(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, incialmente, 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	$a \leq x \leq b$	$a \dashv\vdash b$	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.

(TRT 14ª Região/2018) De um histograma e uma tabela de frequências absolutas, elaborados para analisar a distribuição dos salários dos empregados em uma empresa, obtém-se a informação que 24 empregados ganham salários com valores pertencentes ao intervalo (2.000; 4.000], em reais, que apresenta uma densidade de frequência de $0,75 \times 10^{-4}(R\$)^{-1}$

Densidade de frequência de um intervalo é o resultado da divisão da respectiva frequência relativa pela amplitude deste intervalo. Em um intervalo do histograma que está sendo analisado, com uma amplitude de R$ 3.000,00 e uma densidade de frequência de $1 \times 10^{-4}(R\$)^{-1}$, tem-se que o correspondente número de empregados é igual a

a) 40.

b) 36.

c) 30.

d) 48.

e) 42.

Comentários:

Vamos destacar os dados da questão:

- densidade de frequência: $0{,}75 \times 10^{-4}(R\$)^{-1}$;
- amplitude da classe: $4000 - 2000 = 2000$;
- número de funcionários na classe: 24; e
- total de funcionários na classe: T.

A frequência relativa é calculada pela fórmula a seguir:

$$Frequência\ relativa\ = \frac{Nº\ de\ funcionários\ na\ classe}{Total\ de\ funcionários} = \frac{24}{T}$$

Então, pela fórmula da densidade de frequência, teremos:

$$Densidade\ de\ frequência = \frac{Frequência\ relativa}{Amplitude}$$

$$0{,}75 \times 10^{-4} = \frac{\frac{24}{T}}{2000}$$

$$0{,}15 = \frac{24}{T}$$

$$T = 160$$

Agora, vamos aplicar a mesma metodologia para a classe do histograma com amplitude de 3000 e densidade de frequência de 1×10^{-4}. Considerando o número de funcionários na classe igual a N, teremos a seguinte frequência relativa:

$$Frequência\ relativa\ = \frac{Nº\ de\ funcionários\ na\ classe}{Total\ de\ funcionários} = \frac{N}{160}$$

Então, pela fórmula da densidade de frequência, teremos:

$$Densidade\ de\ frequência = \frac{Frequência\ relativa}{Amplitude}$$

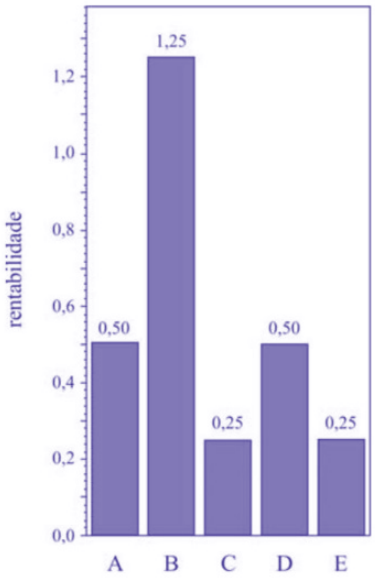
$$10^{-4} = \frac{\frac{N}{160}}{3000}$$

$$0{,}3 = \frac{N}{160}$$

$$N = 48$$

Gabarito: D.

(CESPE/FUNPRESP/2016)



O gráfico ilustra cinco possibilidades de fundos de investimento com suas respectivas rentabilidades. Considerando que as probabilidades de investimento para os fundos A, B, C e D sejam, respectivamente, P(A) = 0,182; P(B) = 0,454; P(C)= 0,091; e P(D)= 0,182, julgue o item subsequente.

O gráfico apresentado é um histograma.

Comentários:

O gráfico não é um histograma por dois motivos: primeiro porque há uma separação entre as colunas, o que não ocorre em um histograma, e sim em um gráfico de colunas; segundo porque um histograma representa dados que estão agrupados em intervalos de classe, e não em categorias, como é o caso.

Gabarito: Errado.

9. Outros Gráficos e Representações

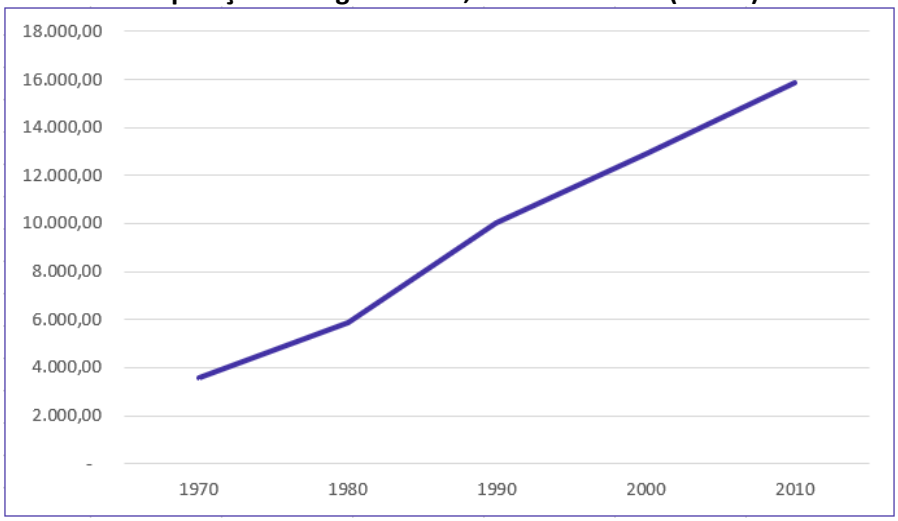
O principal objetivo dos gráficos estatísticos é proporcionar uma visualização mais rápida dos dados estatísticos ou do fenômeno sob investigação. A seguir vamos ver as principais formas de representação de dados estatísticos.

9.1 – Gráficos em Linhas

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

Por exemplo, o gráfico a seguir mostra a evolução da população residente na Região Norte do Brasil no período de 1970 a 2010. Veja que o eixo horizontal está indicando o decurso do tempo, enquanto o eixo vertical apresenta a população residente na região.

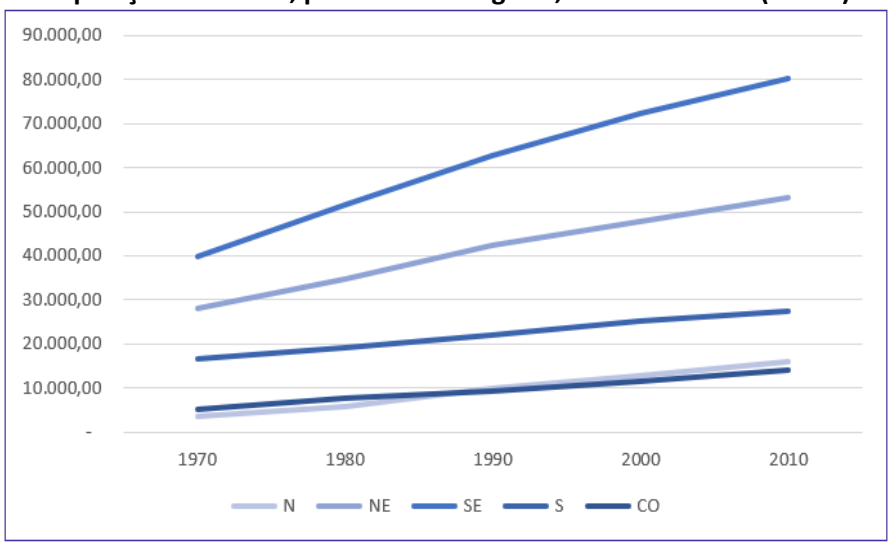
População da região Norte, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

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

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



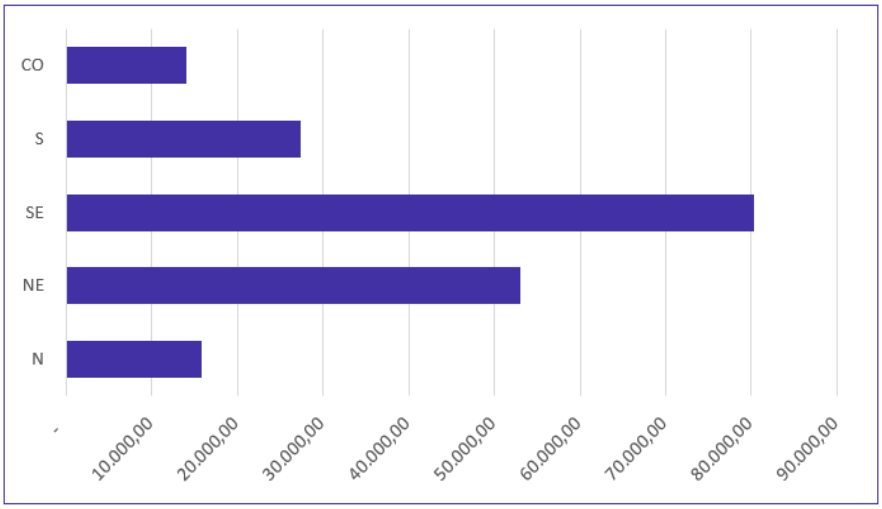
Fonte: Censo Demográfico 1970/2010 (IBGE)

9.2 – Gráficos em Barras

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

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

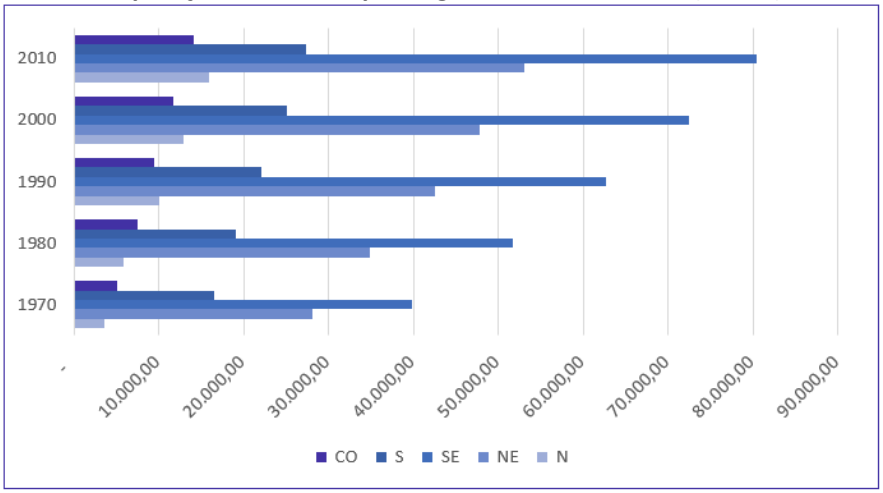
População brasileira, por região, em 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Adicionalmente, podemos utilizar um **gráfico de barras justapostas** para representar a evolução da população residente em cada região brasileira, como o mostrado a seguir:

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



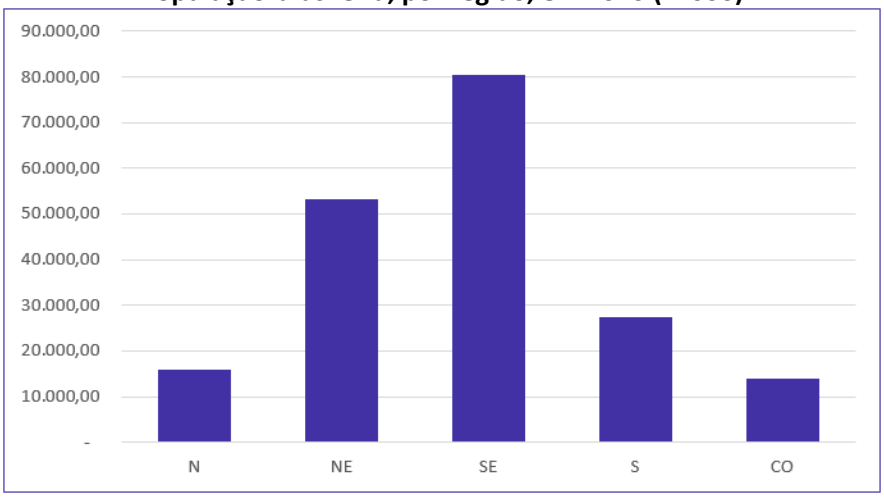
Fonte: Censo Demográfico 1970/2010 (IBGE)

9.3 – Gráficos em Colunas

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

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

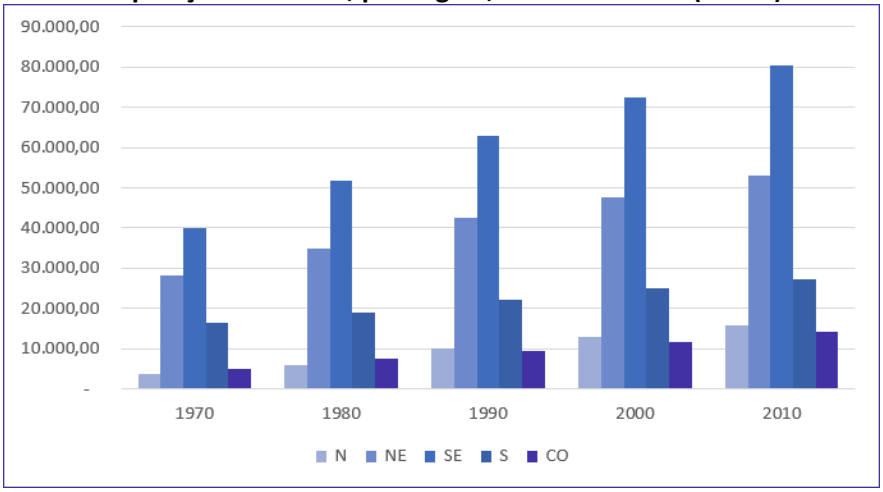
População brasileira, por região, em 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

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

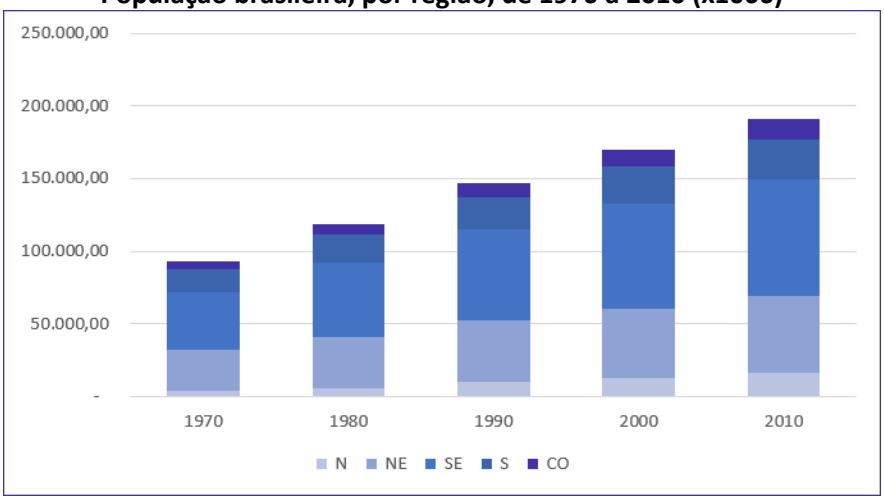
População brasileira, por região, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

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

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



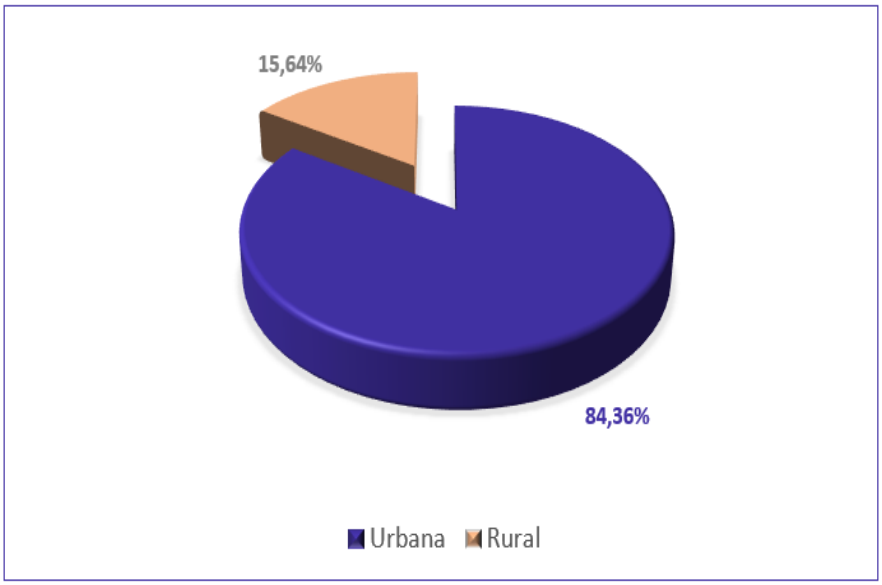
Fonte: Censo Demográfico 1970/2010 (IBGE)

9.4 – Gráfico em Setores

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

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

População Urbana e Rural do Brasil em 2010

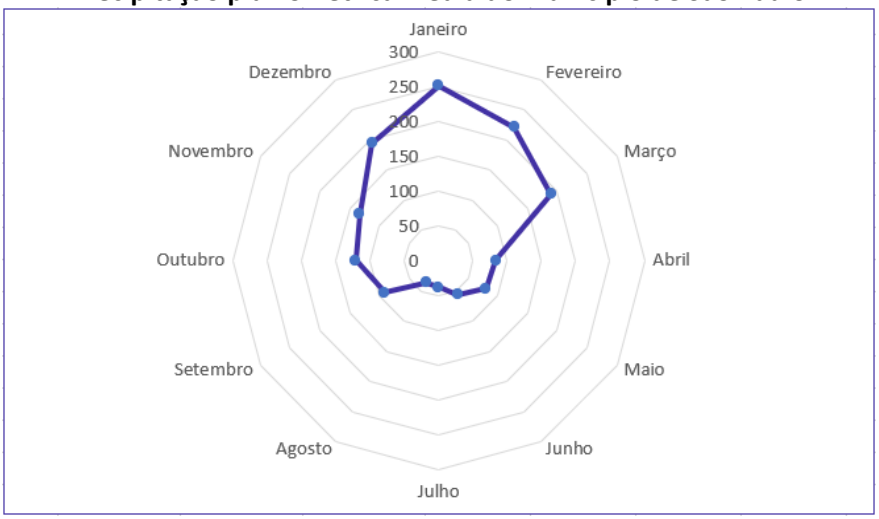


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

9.5 – Gráfico Polar

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

Precipitação pluviométrica média do município de São Paulo

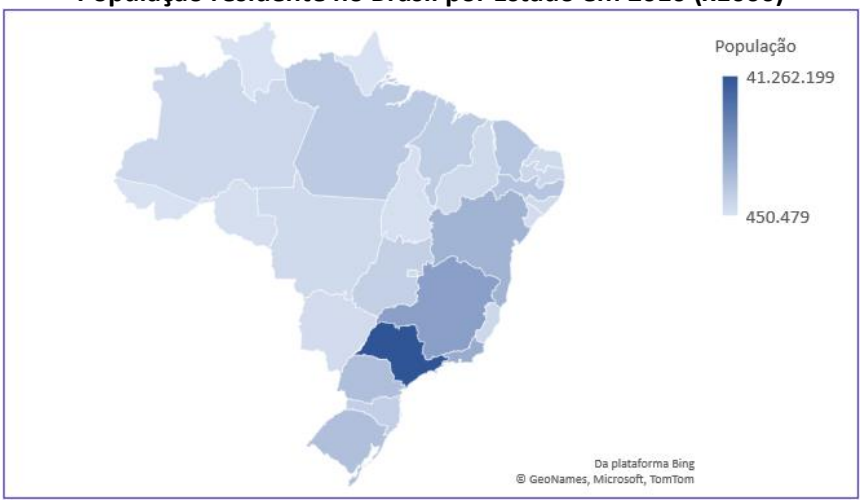


Fonte: Atlas Pluviométrico do Brasil (Serviço Geológico do Brasil)

9.6 – Cartograma

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

População residente no Brasil por Estado em 2010 (x1000)

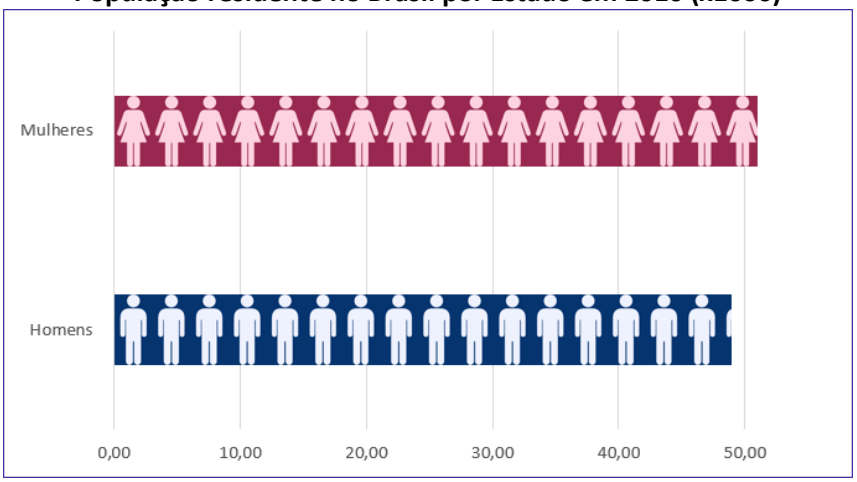


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

9.7 – Pictogramas

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

População residente no Brasil por Estado em 2010 (x1000)

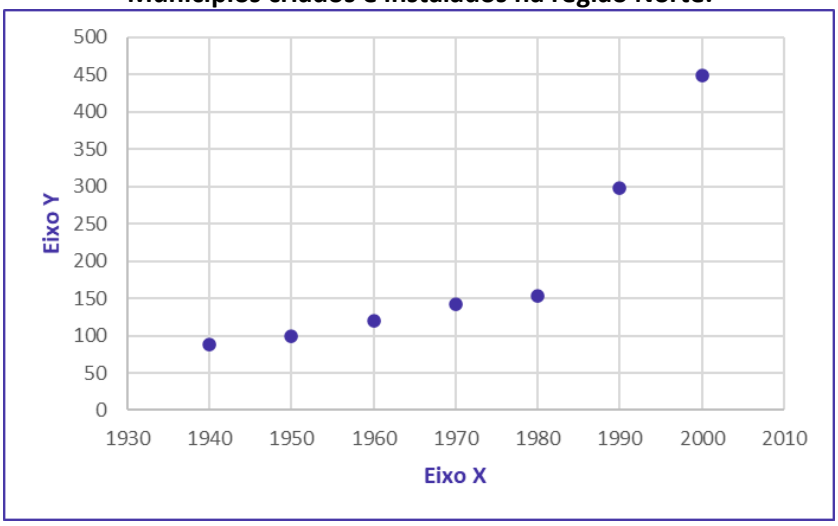


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

9.8 – Gráfico de Dispersão

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

Municípios criados e instalados na região Norte.



Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

9.9 – Diagrama de Ramos e Folhas

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

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

A representação utilizando um diagrama de ramos e folhas ficaria assim:

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

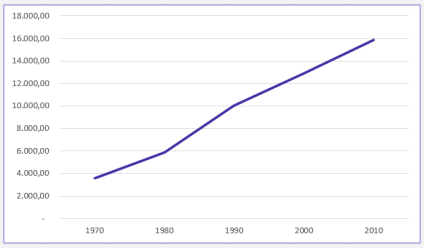
Repare que, no lado esquerdo, temos as centenas e as dezenas representando os ramos. Por sua vez, no lado direito, temos as unidades representando as folhas. As folhas, portanto, estão vinculadas aos ramos. Dessa maneira, a chave "**9 |** 6 8 9" significa que, no rol original, havia os números **9**6, **9**8 e **9**9.

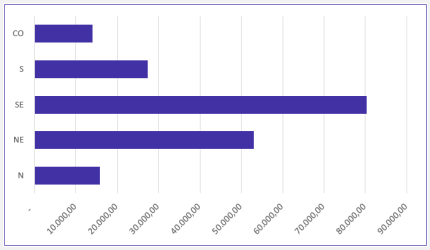
Também é comum encontrarmos diagramas de ramos e folhas em que as unidades são separadas em dois grupos: de 0 a 4 e de 5 a 9. Nesse caso, teríamos o seguinte diagrama:

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

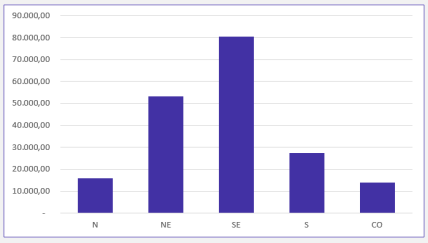
Por fim, é importante observarmos que não existe uma regra única para a construção do diagrama de ramos e folhas. O formato mais comumente encontrado é o que separa o número em duas partes, porém, a depender da chave escolhida, o número pode ser separado em mais partes. **Assim, ao resolver questões que envolvem esse tipo de diagrama, devemos observar a chave adotada**.



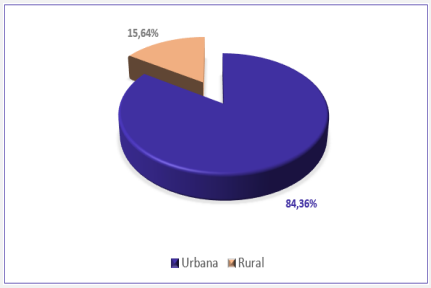
Gráfico	Definição
	Os **gráficos em linha** normalmente são usados para representar a variação dos valores de uma variável ao longo do tempo. Esse tipo de gráfico permite-nos comparar duas variáveis: uma é traçada no eixo x (horizontal) e a outra no eixo y (vertical).



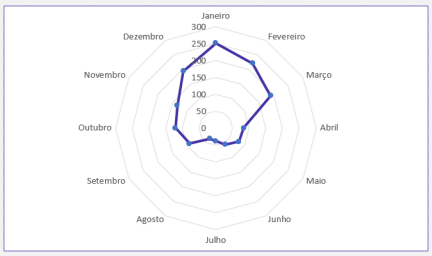
Os **gráficos em barra** normalmente são usados para representar distribuições de dados categóricos ou qualitativos. Uma série estatística é representada por um conjunto de retângulos dispostos **horizontalmente**, cada um indicando uma categoria particular.



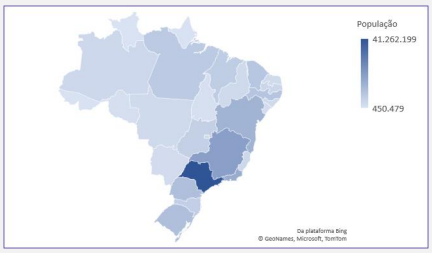
Os **gráficos em coluna** também são usados para distribuições de dados categóricos ou qualitativos. A diferença básica é que, agora, uma série estatística é representada por um conjunto de retângulos dispostos **verticalmente**, cada um indicando uma categoria particular.



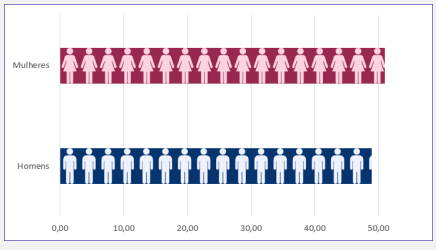
O **gráfico em setores** é usado para representar a frequência relativa (porcentagem) de uma variável categórica, sendo formado por um círculo dividido em setores circulares, cada um representando uma categoria.



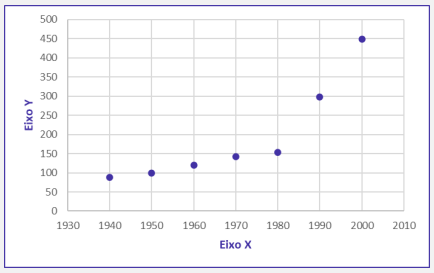
O **gráfico polar** consiste em uma sequência de eixos igualmente espaçados (ângulos iguais), cada um representando uma das variáveis. Uma linha é desenhada ligando os valores de cada eixo.



O **cartograma** é empregado com a finalidade de apresentar dados estatísticos diretamente relacionados com áreas geográficas.



O **pictograma** substitui valores por ícones, tornando os dados mais atraentes e facilitando o entendimento acerca de um determinado fenômeno.



O **gráfico de dispersão** é uma representação de pares ordenados em um plano cartesiano, composto por um eixo vertical (ordenada) e um eixo horizontal (abcissa). Os dados são representados como uma coleção de pontos.

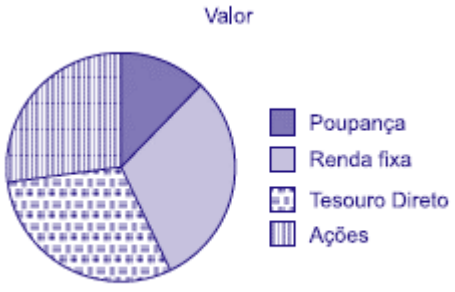
8	5 9	Chave: 8 | 5 = 85
9	6 8 9	
10	3 4 5	
11	3 4 5 5	
12	3 3 4 6 6 6 7 8 9 9	
13	4 5 5 5 7 7 7	
14	2 3 3 8	
15	3 4 5 7 8 9	

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

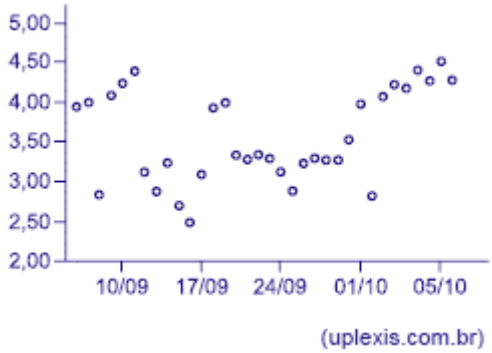
HORA DE PRATICAR!

(UNICAMP/2019) Assinale dentre os exemplos a seguir, o gráfico de dispersão.

a)



b)

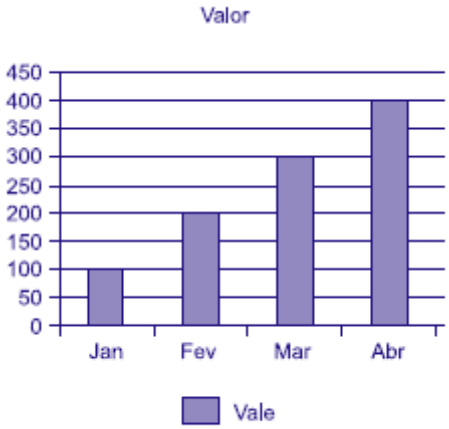


c)



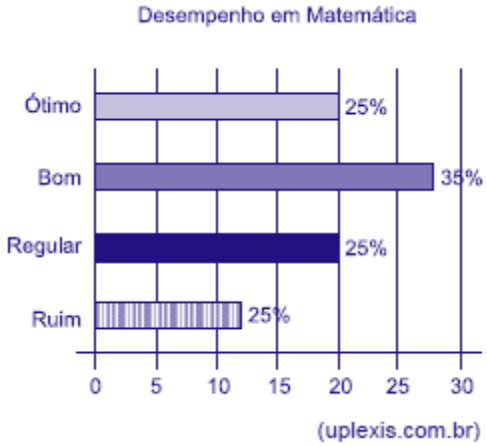
(uplexis.com.br)

d)



(uplexis.com.br)

e)



(uplexis.com.br)

Comentários:

Analisando cada gráfico, temos:

Letra A: o gráfico de setores é formado por um círculo dividido em setores circulares, cada um representando uma categoria, cujos ângulos centrais são proporcionais às frequências relativas da categoria. Logo, a alternativa está errada.

Letra B: o gráfico de dispersão é uma representação gráfica que analisa a relação entre duas variáveis quantitativas — uma de causa e outra de efeito, chamadas de variáveis independente e dependente, respectivamente. Esse tipo de diagrama traz números simultâneos das duas variáveis, deixando visível se o que acontece em uma variável interfere na outra.

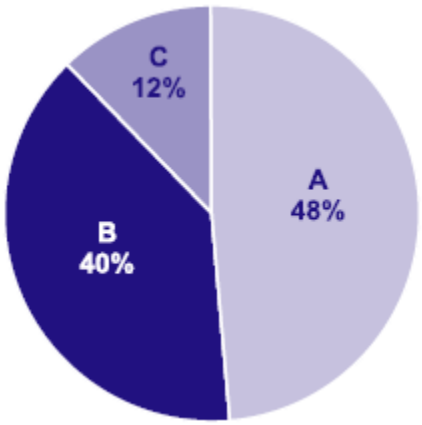
Letra C: o gráfico em linhas é o gráfico em que os pontos são geralmente usados para controlar alterações ao longo do tempo e para facilitar a identificação de tendências ou de anomalias. Alternativa Errada.

Letra D: os gráficos em colunas assim como os gráficos de barras são usados para distribuições de dados categóricos ou qualitativos. A diferença básica é que, no primeiro, uma série estatística é representada por um conjunto de retângulos dispostos verticalmente, cada um indicando uma categoria particular, todos com a mesma largura e alturas proporcionais aos respectivos. Alternativa Errada.

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

Gabarito: B.

(Pref. Campinas/2019) Uma empresa atua em três segmentos de mercado, A, B e C. O gráfico de setores mostra a distribuição percentual, por segmento, da receita total obtida por essa empresa em 2018.



Sabendo-se que a receita obtida no segmento A superou a receita obtida no segmento B em R$ 64 milhões, é correto afirmar que a receita obtida no segmento C foi igual a

a) R$ 98 milhões.

b) R$ 96 milhões.

c) R$ 94 milhões.

d) R$ 88 milhões.

e) R$ 86 milhões.

Comentários:

Consideremos T o total de receitas obtidas pela empresa em 2018. Pelas informações apresentadas no gráfico de pizza, temos:

$$A = 0{,}48 \times T$$

$$B = 0{,}4 \times T$$
$$C = 0{,}12 \times T$$

Sabendo que a receita obtida no segmento A superou a receita obtida no segmento B em R$ 64 milhões, então:

$$A = B + 64$$
$$0{,}48 \times T = 0{,}4 \times T + 64$$
$$0{,}8 \times T = 64$$
$$T = 800\ milhões$$

Logo,

$$C = 0{,}12 \times T$$
$$C = 12 \times 800$$
$$C = 96\ milhões$$

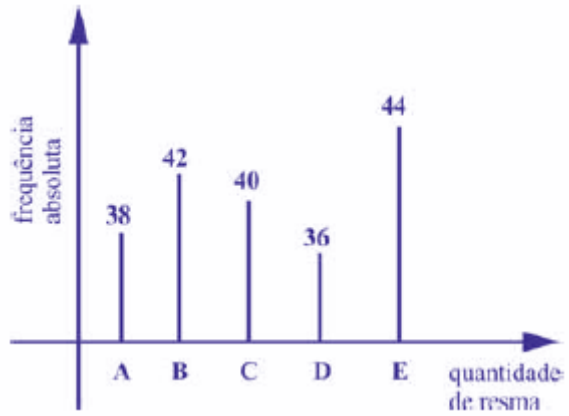
Gabarito: B.

(CESPE/CBM-AL/2017) Na tabela a seguir, A, B, C, D e E são as quantidades de resmas de papel A4 consumidas, em quatro meses, pelas seções administrativas I, II, III, IV e V, respectivamente. Apesar de não mostrar explicitamente essas quantidades, a tabela apresenta as frequências absolutas e (ou) relativas de algumas dessas quantidades.

Seção	Quantidades de Resmas	Frequência Absoluta	Frequência Relativa
I	A	38	19%
II	B		
III	C		20%
IV	D	36	
V	E	44	
	Total		**100%**

Considerando que cada uma dessas resmas, juntamente com a embalagem, tem forma de um paralelepípedo retângulo reto que mede 5 cm × 21 cm × 30 cm, julgue o item seguinte.

O gráfico de barras verticais a seguir apresenta as frequências absolutas de resmas consumidas pelas cinco seções.



Comentários:

Podemos usar uma regra de três simples para completar a tabela. Primeiro, vamos encontrar o valor que representa 100%:

$$38 - 0{,}19$$
$$Q - 1{,}00$$
$$Q = \frac{38 \times 1{,}00}{0{,}19} = 200$$

Encontrando o valor de 20%:

$$200 - 1{,}00$$
$$C - 0{,}20$$
$$C = \frac{200 \times 0{,}2}{1{,}00} = 40$$

Agora, basta sabermos a frequência absoluta na seção II. Como foram consumidas 200 resmas, então o número de resmas consumidas pela seção II foi:

$$200 - 38 - 40 - 36 - 44 = 42$$

Dessa forma, nas seções I, II, III, IV e V foram consumidas, respectivamente 38, 42, 40, 36 e 44 resmas de papel A4, conforme mostra o gráfico.

Gabarito: Certo.

QUESTÕES COMENTADAS

1. (CESPE/Polícia Federal/2018)

	DIA				
	1	2	3	4	5
X (quantidade diária de drogas apreendidas, em kg)	10	22	18	22	28

Tendo em vista que, diariamente, a Polícia Federal apreende uma quantidade X, em kg, de drogas em determinado aeroporto do Brasil, e considerando os dados hipotéticos da tabela precedente, que apresenta os valores observados da variável X em uma amostra aleatória de 5 dias de apreensões no citado aeroporto, julgue o item.

A tabela em questão descreve a distribuição de frequências da quantidade de drogas apreendidas nos cinco dias que constituem a amostra.

Comentários:

A tabela apenas registra as quantidades de drogas apreendidas por dia. Para que ela configurasse uma distribuição de frequências da quantidade de drogas apreendidas nos cinco dias, era necessário que cada valor estivesse associado a uma frequência, o que não ocorreu. Uma distribuição de frequência dos dados acima seria assim:

Valor	Frequência
10	1
18	1
22	2
28	1

Gabarito: Errado.

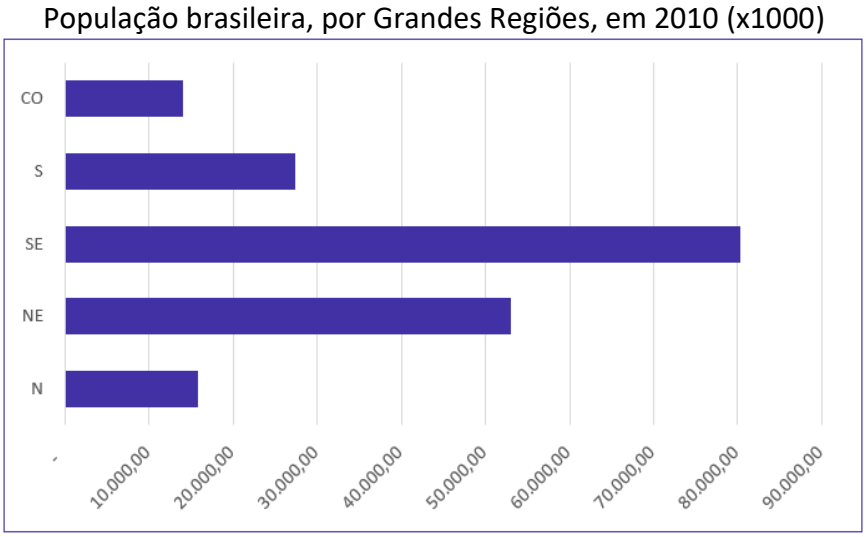
2. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

O gráfico de barras é adequado para a análise de variáveis qualitativas ordinais ou quantitativas discretas, pois permite investigar a presença de tendência nos dados.

Comentários:

O gráfico em barras normalmente é usado para representar distribuições de dados categóricos ou qualitativos. Por meio desse gráfico, uma série estatística é representada por um conjunto de retângulos dispostos horizontalmente, cada um indicando uma categoria em particular, os quais possuem a mesma altura e comprimentos proporcionais aos respectivos dados.

Para ficar claro, apresentamos um gráfico de barras como exemplo:



Fonte: Censo Demográfico 1970/2010 (IBGE)

Gabarito: Certo.

3. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

A representação de diagramas de barras, de linha e de pizza possui escala de medida nominal e tem a moda como medida de tendência central.

Comentários:

Os diagramas de barras, de linha e de pizza são geralmente usados para representar dados nominais ou categóricos. Como veremos na aula de medidas de posição, a única medida de tendência central que podemos utilizar no caso de dados categóricos é a moda, pois ela apenas registra a categoria que mais se repete em uma amostra, sem envolver outras operações matemáticas.

Gabarito: Certo.

4. (CESPE/IPHAN/2018). A tabela a seguir mostra as quantidades de bibliotecas públicas presentes em 20 microrregiões brasileiras.

90	66	78	82
77	60	64	90

87	85	67	91
82	70	81	80
69	78	90	67

A partir desses dados, pretende-se construir um gráfico de distribuição de frequências com quatro classes de igual amplitude. Os valores mínimo e máximo de cada classe devem ser números inteiros.

Considerando essas informações, julgue o item subsequente, relativo ao gráfico de distribuição a ser apresentado.

A amplitude de cada classe deverá ser superior a 6.

Comentários:

A questão forneceu um conjunto de dados e disse que se pretende construir um gráfico de distribuição de frequências com quatro classes de igual amplitude. Para encontrar a amplitude de cada classe, podemos utilizar a fórmula da amplitude total:

$$AT = h \times k \Longrightarrow h = \frac{AT}{k}$$

Analisando a tabela, identificamos os limites mínimo e máximo:

$$l_{máx} = 91$$
$$l_{mín} = 60$$

Calculando a amplitude total, temos:

$$AT = l_{máx} - l_{mín}$$

$$AT = 91 - 60$$

$$AT = 31$$

Aplicando a fórmula anterior, temos:

$$h = \frac{AT}{k} = \frac{31}{4} = 7{,}75$$

O número inteiro mais próximo deste resultado é 8 (superior a 6).

Logo, o item está correto.

Gabarito: Certo.

5. (CESPE/IPHAN/2018) A tabela a seguir mostra as quantidades de bibliotecas públicas presentes em 20 microrregiões brasileiras.

90	66	78	82
77	60	64	90
87	85	67	91
82	70	81	80
69	78	90	67

A partir desses dados, pretende-se construir um gráfico de distribuição de frequências com quatro classes de igual amplitude. Os valores mínimo e máximo de cada classe devem ser números inteiros. Considerando essas informações, julgue o item subsequente, relativo ao gráfico de distribuição a ser apresentado.

A última classe deverá variar de 84 a 91.

Comentários:

A questão não informou qual o limite inferior do primeiro intervalo, portanto, poderíamos ter adotado qualquer valor. Isso alteraria totalmente o resultado esperado pelo avaliador. Dessa forma, considero que a questão deveria ter sido anulada. De todo modo, vamos fazer a questão como a banca esperava que ela fosse resolvida.

Para encontrar a amplitude de cada classe, podemos utilizar a fórmula da amplitude total:

$$AT = h \times k \Longrightarrow h = \frac{AT}{k}$$

Analisando a tabela, identificamos os limites mínimo e máximo:

$$l_{máx} = 91$$
$$l_{mín} = 60$$

Calculando a amplitude total, temos:

$$AT = l_{máx} - l_{mín}$$

$$AT = 91 - 60$$

$$AT = 31$$

Aplicando a fórmula anterior, temos:

$$h = \frac{AT}{k} = \frac{31}{4} = 7{,}75$$

Para o número inteiro mais próximo deste resultado pode ser 8 ou 7. Vamos iniciar calculando para o menor valor, 7:

$$60 \vdash\dashv 67$$
$$68 \vdash\dashv 75$$
$$76 \vdash\dashv 83$$
$$84 \vdash\dashv 91$$

Assim, concluímos que o examinador optou por esse valor para o cálculo. Novamente, a questão não explicitou que o valor mínimo deveria ser 60, logo, poderíamos ter adotado qualquer valor, o que alteraria esse resultado.

Gabarito: Certo.

6. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

O histograma é um diagrama de retângulos contíguos com base na curtose das faixas de valores da variável e com área igual à diferença da frequência absoluta da respectiva faixa.

Comentários:

Um histograma é uma representação gráfica (um gráfico de barras verticais ou barras horizontais) da distribuição de frequências de um conjunto de dados quantitativos contínuos. Ele pode ser construído com base em valores absolutos ou frequência relativa ou densidade.

Em se tratando de densidade, a frequência relativa do intervalo *i*, (f_{ri}), é representada pela área de um retângulo, colocado acima do ponto médio da classe *i*. Consequentemente, a área total do histograma (igual a soma das áreas de todos os retângulos) será igual a 1.

Assim, ao construir o histograma, cada retângulo deverá ter área proporcional à frequência relativa (ou à frequência absoluta) correspondente. No caso em que os intervalos são de tamanhos iguais, as alturas dos retângulos serão iguais às frequências relativas (ou iguais às frequências absolutas) dos intervalos correspondentes.

Desse modo, a área não tem relação com a diferença de frequência absoluta de um certo intervalo. A área de cada retângulo é proporcional à frequência da classe, sem qualquer operação de diferença envolvida.

Outro ponto relevante é que o histograma não tem relação com a curtose dos intervalos. A curtose é uma medida que indica o grau de achatamento de uma distribuição de frequências.

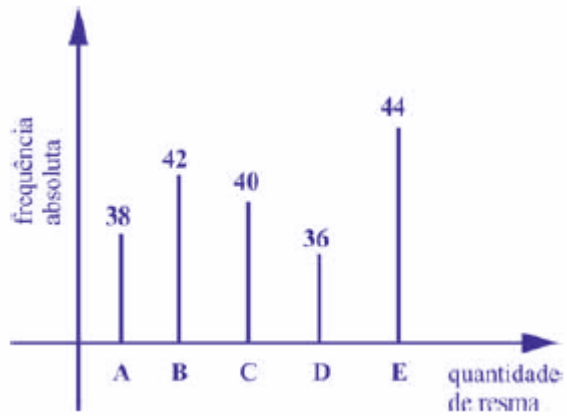
Gabarito: Errado.

7. (CESPE/CBM-AL/2017). Na tabela a seguir, A, B, C, D e E são as quantidades de resmas de papel A4 consumidas, em quatro meses, pelas seções administrativas I, II, III, IV e V, respectivamente. Apesar de não mostrar explicitamente essas quantidades, a tabela apresenta as frequências absolutas e (ou) relativas de algumas dessas quantidades.

Seção	Quantidades de Resmas	Frequência Absoluta	Frequência Relativa
I	A	38	19%
II	B		
III	C		20%
IV	D	36	
V	E	44	
	Total		**100%**

Considerando que cada uma dessas resmas, juntamente com a embalagem, tem forma de um paralelepípedo retângulo reto que mede 5 cm × 21 cm × 30 cm, julgue o item seguinte.

O gráfico de barras verticais a seguir apresenta as frequências absolutas de resmas consumidas pelas cinco seções.



Comentários:

Podemos usar uma regra de três simples para completar a tabela. Primeiro, vamos encontrar o valor que representa 100%:

$$38 - 0{,}19$$

$$Q - 1{,}00$$

$$Q = \frac{38 \times 1{,}00}{0{,}19} = 200$$

Encontrando o valor de 20%:

$$200 - 1{,}00$$

$$C - 0{,}20$$

$$C = \frac{200 \times 0{,}2}{1{,}00} = 40$$

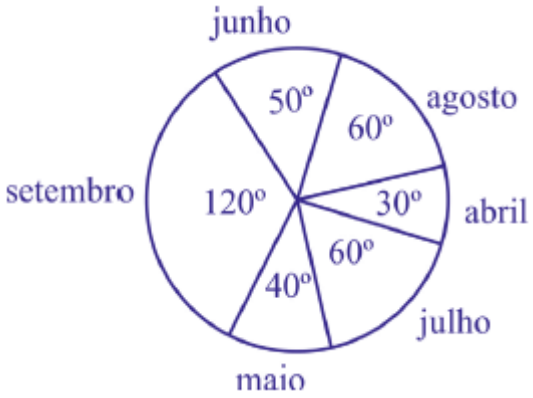
Agora, basta sabermos a frequência absoluta na seção II. Como foram consumidas 200 resmas, então o número de resmas consumidas pela seção II foi:

$$200 - 38 - 40 - 36 - 44 = 42$$

Dessa forma, nas seções I, II, III, IV e V foram consumidas, respectivamente 38, 42, 40, 36 e 44 resmas de papel A4, conforme mostra o gráfico.

Gabarito: Certo.

8. (CESPE/CBM-AL/2017) O gráfico de setores a seguir mostra a distribuição das quantidades de incêndios em determinada região, nos meses de abril a setembro de determinado ano.



Sabendo-se que nesses meses ocorreram 1.548 incêndios nessa região, julgue o item que se segue.

A frequência relativa à classe "incêndios no mês de setembro" é superior a 30%.

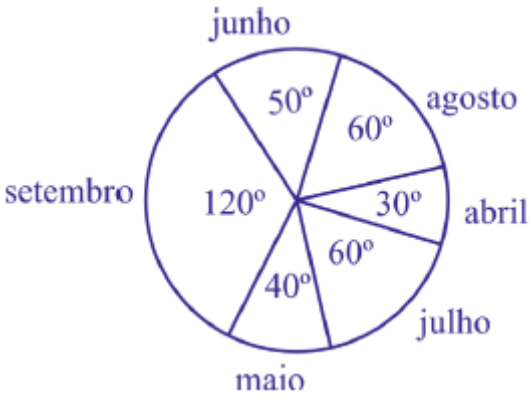
Comentários:

Para resolvermos essa questão, basta aplicarmos uma regra de três simples. Se uma circunferência total tem 360° e corresponde a 100%, então o ângulo de 120° corresponderá a:

$$Set. = \frac{120}{360} = \frac{1}{3} \approx 33{,}33\%$$

Gabarito: Certo.

9. (CESPE/CBM-AL/2017) O gráfico de setores a seguir mostra a distribuição das quantidades de incêndios em determinada região, nos meses de abril a setembro de determinado ano.



Sabendo-se que nesses meses ocorreram 1.548 incêndios nessa região, julgue o item que se segue.

Nos meses de maio e junho ocorreram mais de 400 incêndios nessa região.

Comentários:

Para resolver a questão, devemos aplicar uma regra de três simples. Isto é, se a circunferência total possui 360° e corresponde a 100%, então 360° também equivale ao número total de incêndios, 1548.

Podemos verificar também que os meses de maio e junho somam um total de 90°. Calculando 90°, temos:

$$360 - 1548$$

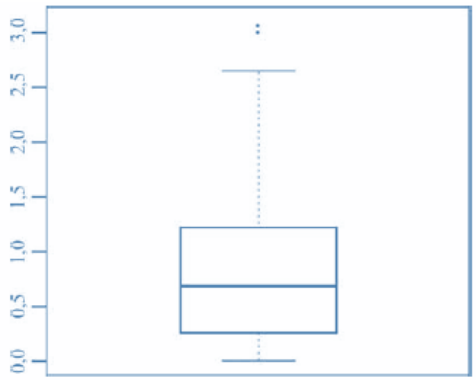
$$90 - X$$

$$X = \frac{139.320}{360} = 387$$

Logo, a quantidade de incêndios durante os meses de maio e junho não foi superior a 400, como afirma o item.

Gabarito: Errado.

10. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

X representa uma variável qualitativa ordinal.

Comentários:

As variáveis podem ser classificadas em:

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

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. Podem ser nominais ou ordinais.

Analisando as informações dadas na questão, percebemos que X pode adotar valores que vão de 0 (mínimo) a 3,10 (máximo). Logo, como X assume valores numéricos, dizemos que ela é uma variável quantitativa.

Gabarito: Errado.

11. (CESPE/TCE-PA/2016)

Número diário de denúncias registradas (X)	Frequência Relativa
0	0,3
1	0,1

2	0,2
3	0,1
4	0,3
Total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

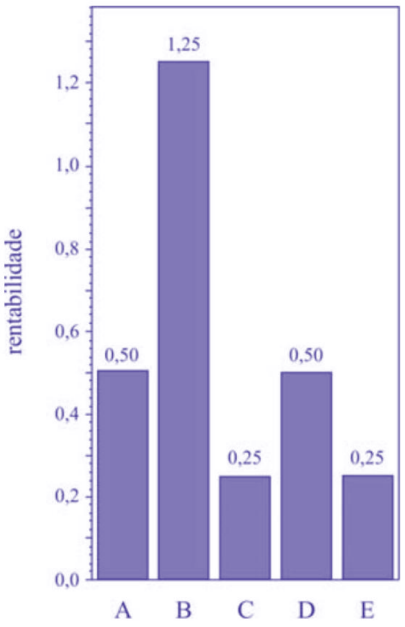
A variável X é do tipo qualitativo nominal.

Comentários:

De acordo com a tabela X corresponde à quantidade diária de denúncias. Logo, X é uma variável quantitativa discreta, pois assume valores inteiros.

Gabarito: Errado.

12. (CESPE/FUNPRESP/2016)



O gráfico ilustra cinco possibilidades de fundos de investimento com suas respectivas rentabilidades. Considerando que as probabilidades de investimento para os fundos A, B, C e D sejam, respectivamente, P(A) = 0,182; P(B) = 0,454; P(C)= 0,091; e P(D)= 0,182, julgue o item subsequente.

O gráfico apresentado é um histograma.

Comentários:

O gráfico não é um histograma por dois motivos: primeiro porque há uma separação entre as colunas, o que não ocorre em um histograma, e sim em um gráfico de colunas; segundo porque um histograma representa dados que estão agrupados em intervalos de classe, e não em categorias, como é o caso.

Gabarito: Errado.

13. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

Na análise exploratória, o histograma é um gráfico adequado para descrever a distribuição da quantidade de detentos por região em 2013.

Comentários:

O histograma é destinado a representar dados agrupados em classes, mas o enunciado apresenta dados com valores individualizados. Como a informação não está agrupada em classes, o histograma não representaria adequadamente os dados da tabela.

Gabarito: Errado.

14. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

Em 2013, mais de 55% da população carcerária no Brasil se encontrava na região Sudeste.

Comentários:

Vamos analisar os dados da região sudeste. Sabemos que o total da população carcerária é 555 mil. Se a região sudeste contém 306 mil, basta calcularmos o percentual desse valor com relação ao total. Assim:

$$X = \frac{306}{555} \times 100\% = 55{,}13\%$$

Gabarito: Certo.

15. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85

Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

A quantidade total de vagas existentes no sistema penitenciário brasileiro em 2013 era de 340 mil vagas.

Comentários:

Analisando a tabela, temos a informação do total de 555 mil detentos, e a informação do déficit de vagas no sistema penitenciário 215 mil.

Ora, para sabermos a quantidade real de vagas, basta subtrairmos o déficit do total.

$$555 - 215 = 340$$

Gabarito: Certo.

16. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

O déficit relativo de vagas observado na região Sudeste, em 2013, foi superior ao déficit relativo de vagas registrado na região Centro-oeste no mesmo período.

Comentários:

Precisamos calcular o percentual do déficit das regiões sudeste e centro-oeste. Para isso, basta dividirmos o déficit de vagas da região pela respectiva quantidade de detentos dessa região. Assim:

$$Sudeste = \frac{120}{306} = 39{,}21\%$$

$$Centro - Oeste = \frac{24}{51} = 47{,}05\%$$

Portanto, o déficit é maior na região Centro-Oeste.

Gabarito: Errado.

17. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema

penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

No ano considerado, a quantidade média de detentos por 100 mil habitantes na região Nordeste foi superior ao número médio de detentos por 100 mil habitantes na região Centro-oeste.

Comentários:

Precisamos calcular a razão entre a quantidade de detentos e a população das respectivas regiões. Assim, para as regiões Nordeste e Centro-Oeste, temos que:

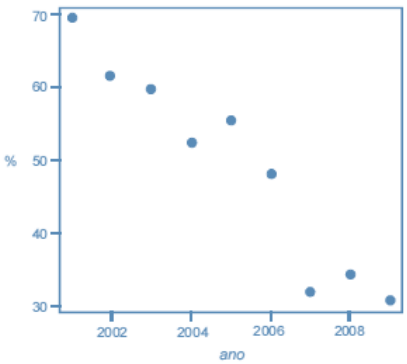
$$Nordeste = \frac{94 \times 10^3}{55 \times 10^6} = 1{,}7 \times 10^{-3}$$

$$Centro - Oeste = \frac{51 \times 10^3}{15 \times 10^6} = 3{,}4 \times 10^{-3}$$

Portanto, a quantidade é maior na região Centro-Oeste.

Gabarito: Errado.

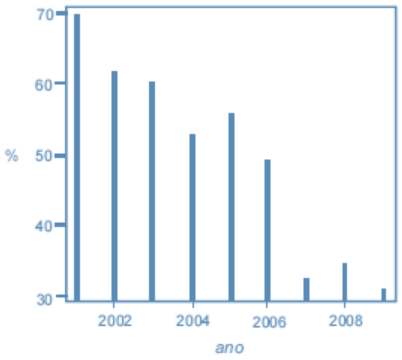
18. (CESPE/DEPEN/2015)



Dado que a participação dos presidiários em cursos de qualificação profissional é um aspecto importante para a reintegração do egresso do sistema prisional à sociedade, foram realizados levantamentos estatísticos, nos anos de 2001 a 2009, a respeito do valor da educação e do trabalho em ambientes prisionais. Cada um desses levantamentos, cujos resultados são apresentados no gráfico, produziu uma estimativa anual do percentual P de indivíduos que participaram de um curso de qualificação profissional de curta duração, mas que não receberam o diploma por motivos diversos. Em 2001, 69,4% dos presidiários que participaram de um curso de qualificação profissional não receberam o diploma. No ano seguinte, 2002, esse percentual foi reduzido para 61,5%, caindo, em 2009, para 30,9%.

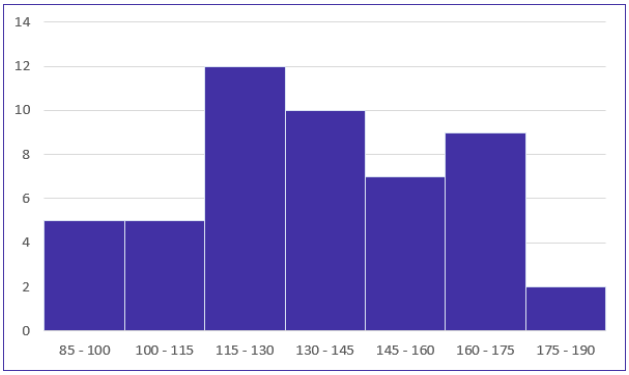
A partir das informações e do gráfico apresentados, julgue o item que se segue.

Se os percentuais forem representados por barras verticais, conforme o gráfico a seguir, então o resultado será denominado histograma.



Comentários:

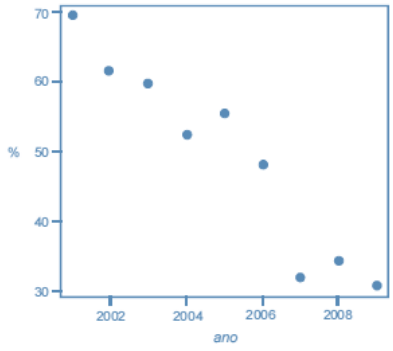
O histograma é destinado a representar dados contínuos agrupados em classes (em intervalos), como o mostrado na figura a seguir.



Em realidade, a assertiva apresentou um gráfico de colunas (barras verticais) que representa dados agrupados por valores.

Gabarito: Errado.

19. (CESPE/DEPEN/2015)



Dado que a participação dos presidiários em cursos de qualificação profissional é um aspecto importante para a reintegração do egresso do sistema prisional à sociedade, foram realizados levantamentos estatísticos, nos anos de 2001 a 2009, a respeito do valor da educação e do trabalho em ambientes prisionais. Cada um desses levantamentos, cujos resultados são apresentados no gráfico, produziu uma estimativa anual do percentual P de indivíduos que participaram de um curso de qualificação profissional de curta duração, mas que não receberam o diploma por motivos diversos. Em 2001, 69,4% dos presidiários que participaram de um curso de qualificação profissional não receberam o diploma. No ano seguinte, 2002, esse percentual foi reduzido para 61,5%, caindo, em 2009, para 30,9%.

A partir das informações e do gráfico apresentados, julgue o item que se segue.

Os dados apresentados são suficientes para que se possa afirmar que o total de presidiários que participaram de um curso de qualificação profissional de curta duração e que não receberam o diploma em 2008 foi superior ao total referente ao ano de 2007.

Comentários:

As informações presentes no gráfico foram dispostas em termos percentuais e não por meio de valores absolutos, portanto, não podemos afirmar que o número de presidiários que participaram do curso e não receberam diploma foi maior em um ano do que em outro. Apenas podemos afirmar que o percentual de presidiários que participaram dos cursos e não receberam o diploma foi maior em 2008 que o percentual em 2007.

Gabarito: Errado.

20. (CESPE/TELEBRAS/2015)



Uma empresa coletou e armazenou em um banco de dados diversas informações sobre seus clientes, entre as quais estavam o valor da última fatura vencida e o pagamento ou não dessa fatura. Analisando essas informações, a empresa concluiu que 15% de seus clientes estavam inadimplentes. A empresa recolheu ainda dados como a unidade da Federação (UF) e o CEP da localidade em que estão os clientes. Do conjunto de todos os clientes, uma amostra aleatória simples constituída por 2.175 indivíduos prestou também informações sobre sua renda domiciliar mensal, o que gerou o histograma apresentado. Com base nessas informações e no histograma, julgue o item a seguir.

O CEP da localidade dos clientes e o valor da última fatura vencida são variáveis quantitativas.

Comentários:

O CEP, embora possua uma representação numérica, não exprime quantidades e, portanto, não é uma variável quantitativa. Esse código número apenas qualifica as ruas de um determinado endereço, podendo ser classificada como uma variável qualitativa.

O valor da última fatura vencida, por sua vez, é sim uma variável quantitativa.

Gabarito: Errado.

21. (CESPE/TELEBRAS/2015) Roberto comprou, por R$ 2.800,00, rodas de liga leve para seu carro, e, ao estacionar no shopping, ficou indeciso sobre onde deixar o carro, pois, caso o coloque no estacionamento público, correrá o risco de lhe roubarem as rodas, ao passo que, caso o coloque no estacionamento privado, terá de pagar R$ 70,00, com a garantia de que eventuais prejuízos serão ressarcidos pela empresa administradora.

Considerando que p seja a probabilidade de as rodas serem roubadas no estacionamento público, que X seja a variável aleatória que representa o prejuízo, em reais, ao deixar o carro no estacionamento público, e que Y seja a variável aleatória que representa o valor, em reais, desembolsado por Roberto ao deixar o carro no estacionamento pago, julgue o item subsequente.

A variável aleatória Y é contínua.

Comentários:

As variáveis quantitativas podem ser classificadas em:

a) variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua, nesse caso, valores fracionários fazem sentido;

b) variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros.

No problema em questão, a variável aleatória Y pode assumir o valor R$ 70,00, se Roberto deixar o carro no estacionamento pago; ou o valor R$ 0,00, se Roberto não deixar o carro no estacionamento pago. Portanto, Y é uma variável aleatória discreta.

Gabarito: Errado.

22. (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. Portanto, o censo requer a realização de um número maior de entrevistas.

Gabarito: Errado.

23. (CESPE/TJ-SE/2014)

Quantidade	São Paulo (j =1)	Rio de Janeiro (j =2)	Minas Gerais (j =3)	Rio Grande do Sul (j =4)	Total
Casos novos (X, em milhões)	5	2	1	2	**18**
Casos pendentes (Y, em milhões)	16	8	3	2	**48**
Processos baixados (Z, em milhões)	5	2	2	2	**18**
Sentenças e Decisões (W, em milhões)	4	3	1	1	**16**

O quadro acima mostra uma síntese da movimentação processual dos tribunais de justiça dos estados de São Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul e do total da justiça estadual no Brasil em 2010. Considere que o estoque de processos em andamento no estado j (Ej), no final de 2010, seja um indicador que se define como $E_j = X_j + Y_j - Z_j - W_j$, em que j = 1, 2, ..., 27; X_j representa o número de casos novos registrados em 2010 no estado j; Yj seja a quantidade de casos pendentes no estado j (i.e., casos anteriores que não foram solucionados até o final de 2010); Z_j denota o total de processos baixados (arquivados) no estado j durante 2010 e W_j seja o número de sentenças e decisões proferidas no estado j até o final de 2010. Considere, por fim, que, para todos os efeitos, o Distrito Federal seja um estado. Com base nessas informações e no quadro acima, julgue o item que se segue.

O quadro apresentado é uma tabela de contingência que mostra o cruzamento entre uma variável qualitativa nominal com 4 níveis de resposta (estados) e outra variável qualitativa com quatro níveis de resposta (casos novos, pendentes, baixados e resolvidos).

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:

i) 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.

ii) 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:

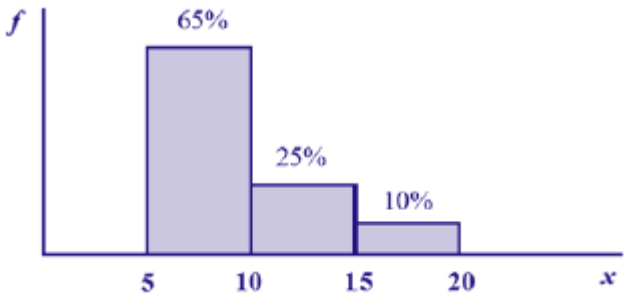
i) variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos.

ii) 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).

No problema em questão, a tabela não representa uma variável qualitativa, mas sim quantitativa, vez que podemos atribuir valores às variáveis X, Y, Z e W.

Gabarito: Errado.

24. (CESPE/STF/2013)



Com referência à figura acima, que mostra a distribuição da renda mensal — x, em quantidades de salários mínimos (sm) — das pessoas que residem em determinada região, julgue o item subsequente.

A variável x, por possuir quatro níveis de respostas, é do tipo qualitativa ordinal.

Comentários:

O gráfico representa um histograma, em que as pessoas foram classificadas de acordo com as quantidades de salários-mínimos recebidos. Reparem na existência de três intervalos de classe (5 a 10; 10 a 15; e 15 a

20). Portanto, a variável x exprime a quantidade de salários-mínimos, sendo classificada como uma variável quantitativa.

Gabarito: Errado.

25. (CESPE/BACEN/2013)

2 4 8 4 8 1 2 32 12 1 5 7 5 5 3 4 24 19 4 14

Os dados mostrados acima representam uma amostra, em minutos, do tempo utilizado na armazenagem de formulários no almoxarifado central de certa instituição por diversos funcionários.

Com base nesses dados, julgue o próximo item.

A distribuição de frequência acumulada para tempo de armazenagem observado na amostra inferior a 8 minutos é igual a 13, o que corresponde a uma frequência relativa superior a 0,60.

Comentários:

Não precisamos construir a distribuição de frequências para responder à questão. Sabemos que o total de observações é 20, pois é esse o número de termos.

Vamos destacar as observações que são inferiores a 8:

2 4 8 **4** 8 **1 2** 32 12 **1 5 7 5 5 3 4** 24 19 **4** 14

São 13 números inferiores a 8. Portanto, a frequência acumulada para um tempo menor do que 8 é 13.

Para calcular a frequência relativa, basta dividir a frequência absoluta pelo total de observações:

$$\frac{8}{13} = 0{,}61$$

Logo, a frequência relativa acumulada é maior do que 0,60.

Gabarito: Certo.

26. (CESPE/BACEN/2013)

2 4 8 4 8 1 2 32 12 1 5 7 5 5 3 4 24 19 4 14

Os dados mostrados acima representam uma amostra, em minutos, do tempo utilizado na armazenagem de formulários no almoxarifado central de certa instituição por diversos funcionários.

Com base nesses dados, julgue o próximo item.

É inviável a elaboração de um histograma em decorrência do fato de ser este um conjunto de dados quantitativos discretos; dessa forma, apenas por meio de um gráfico de barras pode ser realizada a representação gráfica.

Comentários:

O item está **errado**. É sim possível elaborar um histograma para os dados apresentados, bastaria, para tanto, organizá-los em intervalos de classe.

Gabarito: Errado.

27. (CESPE/TRE-ES/2011)

Quantidade de eleitores	Quantidade de municípios
$0 \vdash 2.000$	364
$2.000 \vdash 4.000$	1.000
$4.000 \vdash 6.000$	3.000
$6.000 \vdash 8.000$	1.000
$8.000 \vdash 10.000$	200
Total	**5.564**

A tabela acima apresenta uma distribuição hipotética das quantidades de eleitores que não votaram no segundo turno da eleição para presidente da República bem como os números de municípios em que essas quantidades ocorreram. Com base nessa tabela, julgue o item seguinte, relativo à análise exploratória de dados.

Na tabela de frequências, o uso de intervalos de classe permite concluir que a variável em questão é contínua.

Comentários:

As variáveis quantitativas podem ser classificadas em:

a) variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua, nesse caso, valores fracionários fazem sentido;

b) variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros.

A questão apenas agrupou a quantidade de eleitores em classes, mas continua sendo uma variável discreta.

Gabarito: Errado.

28. (CESPE/TRE-ES/2011)

Cargo	Candidatos	Candidatos aptos	Eleitos
Presidente da República	9	9	1
Governador de Estado	170	156	27
Senador	272	234	54
Deputado Federal	6.021	5.058	513
Deputado Estadual/Distrital	15.268	13.076	1.059
Total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.

A variável "cargo" classifica-se como uma variável qualitativa ordinal.

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:

i) 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;

ii) 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:

i) variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos;

ii) 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 nesses conceitos, podemos afirmar que o cargo é realmente uma variável qualitativa, pois classifica os indivíduos, e ordinal, pois estabelece uma ordem para os cargos (a tabela está ordenada pela coluna número de eleitos).

Gabarito: Certo.

29. (CESPE/TRE-ES/2011)

Cargo	Candidatos	Candidatos aptos	Eleitos
Presidente da República	9	9	1
Governador de Estado	170	156	27
Senador	272	234	54
Deputado Federal	6.021	5.058	513
Deputado Estadual/Distrital	15.268	13.076	1.059
Total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.

Considerando-se a representação das quantidades de eleitos para cada cargo em um gráfico de pizza, a fatia desse gráfico correspondente ao cargo de deputado federal terá ângulo superior a 120°.

Comentários:

Inicialmente, precisamos descobrir o quanto o ângulo de 120° corresponde do valor total. Sabemos que 360° é o ângulo total, logo, 120° corresponde a 1/3 de 360°. Portanto, para afirmarmos que o cargo de deputado

federal corresponde ao ângulo de 120°, ele deve necessariamente corresponder a 1/3 do total de candidatos eleitos. Vejamos:

$$\frac{1658}{3} \cong 552{,}66$$

Logo, para que o cargo de deputado federal tivesse um ângulo superior a 120°, deveríamos ter 553 candidatos eleitos. Não é o caso, pois, de acordo com a tabela, foram eleitos 513 candidatos.

Gabarito: Errado.

30. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

Os extremos mínimo e máximo da variável X foram, respectivamente, iguais a 1 e 80.

Comentários:

Na tabela temos que a variável X assume valores que vão de 0 a 5. As frequências absolutas correspondem a quantas vezes esses valores se repetem. Portanto, concluímos que os valores mínimo e máximo de X são, respectivamente, 0 e 5.

Gabarito: Errado

31. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

O número de municípios que têm obras sob suspeita de irregularidades é superior a 120.

Comentários:

Para respondermos a questão, basta entendermos que precisamos fazer um somatório de todas as obras que estão sob suspeita de irregularidades. Se $X = 0$ se refere ao fato de nenhuma obra estar sob suspeita de irregularidade, então partiremos da soma $X = 1$ a $X = 5$. Assim, somaremos todas as frequências absolutas:

$$47 + 30 + 20 + 6 + 1 = 104$$

Gabarito: Errado.

32. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

O total de municípios considerado no levantamento foi superior a 180.

Comentários:

Para sabermos o total de municípios, basta somarmos todas as frequências absolutas:

$$80 + 47 + 30 + 20 + 6 + 1 = 184$$

Gabarito: Certo.

33. (CESPE/ANTAQ/2009)

	Variável	2003	2004	2005	2006	2007
Exportação	X	40	46	50	52	54
Importação	Y	20	21	22	24	27
Total	X+Y	60	67	72	76	81

Internet: <www.portodesantos.com> (com adaptações).

Considerando a tabela acima, que apresenta a movimentação anual de cargas no porto de Santos de 2003 a 2007, em milhões de toneladas/ ano e associa as quantidades de carga movimentadas para exportação e importação às variáveis X e Y, respectivamente, julgue o item subsequente.

As séries estatísticas apresentadas na tabela formam três séries temporais.

Comentários:

Sabemos que uma série temporal é uma sequência de dados observados em um período de tempo. No caso da tabela, temos três séries cronológicas correspondentes às variáveis X (exportação); Y (importação); e X+Y (total).

Gabarito: Certo.

34. (CESPE/TCU/2008) Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (X) e vendidos (Y) em determinado município, nos anos de 2005 a 2007.

Ano	Número de imóveis	
	Ofertados (X)	Vendidos (Y)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correios Braziliense, 29/4/2008, p.17 (com adaptações)

Com respeito ao texto, considere que cada imóvel ofertado em determinado ano seja classificado como vendido ou não-vendido, e, a um imóvel e classificado como vendido seja atribuído um valor Z = 1, e, ao imóvel classificado como não-vendido, seja atribuído um valor Z = 0. Supondo-se que as classificações dos imóveis como vendido ou não-vendido em um dado ano possam ser consideradas como sendo realizações de uma amostragem aleatória simples, julgue os itens a seguir.

A variável Z é classificada como variável qualitativa nominal, pois representa o atributo do imóvel como vendido ou não-vendido.

Comentários:

Analisando os dados fornecidos no enunciado, observamos que a variável Z é uma variável qualitativa que assume os valores nominais “vendido” ou “não vendido”. No entanto, a questão atribuiu valores numéricos à variável, 1 para vendido e 0 para não vendido. Diante disso, a variável Z também passou a ser classificada como variável quantitativa.

Gabarito: Errado.

35. (CESPE/TCU/2008) Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (X) e vendidos (Y) em determinado município, nos anos de 2005 a 2007.

Ano	Número de imóveis	
	Ofertados (X)	Vendidos (Y)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correios Braziliense, 29/4/2008, p.17 (com adaptações)

Considerando as informações do texto, julgue os itens subsequentes.

A variável X forma uma série estatística denominada série temporal.

Comentários:

Vamos analisar os elementos da tabela. Temos uma série em que o fato (número de imóveis) e o local (um determinado município) permanecem constantes, mas há variação do tempo (2005, 2006, 2007). Portanto, temos todos os elementos para classificar a série em temporal ou cronológica.

Gabarito: Certo.

36. (CESPE/SEFAZ-MT/2004) Considere a seguinte situação hipotética.

Um órgão do governo recebeu pela Internet denúncias de sonegação de impostos estaduais contra 600 pequenas empresas. Denúncias contra outras 200 pequenas empresas foram encaminhadas pessoalmente para esse órgão. Para a apuração das denúncias, foram realizadas auditorias nas 800 empresas denunciadas. Como resultado dessas auditorias, foi elaborada a tabela abaixo, que apresenta um quadro das empresas denunciadas e os correspondentes débitos fiscais ao governo. Das empresas denunciadas, observou-se que apenas 430 tinham débitos fiscais.

	Valor do débito fiscal (VDF), em R$ mil, apurado após auditoria na empresa denunciada				
Forma de recebimento da denúncia	$0 < VDF < 1$	$1 \leq VDF < 2$	$2 \leq VDF < 3$	$3 \leq VDF \leq 4$	Total
Pela internet	60	100	50	30	240
Pessoalmente	20	120	40	10	190

Total	80	220	90	40	430*

Nota: *Para as demais empresas, VDF=0

Com base na situação hipotética acima e de acordo com as informações apresentadas, julgue o item que se segue.

O valor total dos débitos fiscais apurados após as auditorias feitas nas empresas denunciadas é inferior a R$ 500 mil.

Comentários:

Para responder essa questão, trabalharemos com os limites inferiores das classes, buscando minimizar os valores dos débitos, para ver se é possível que o total seja menor de R$ 500 mil.

Débito Mínimo	Frequência	Produto
0	80	0
1	220	220
2	90	180
3	40	120
Total		**520**

Portanto, o débito total é de, no mínimo, R$ 520 mil. Logo, é impossível que o valor total dos débitos fiscais seja inferior a R$ 500 mil.

Gabarito: Errado.

37. (CESPE/SEFAZ-AL/2002) Julgue o seguinte item.

Em uma distribuição de frequências para um conjunto de n indivíduos, pode-se calcular as frequências relativas, dividindo-se cada frequência absoluta pela amplitude da correspondente classe ou do intervalo.

Comentários:

A frequência relativa é calculada pela divisão entre a frequência absoluta simples e a frequência total. Na realidade, a questão fez referência ao cálculo da densidade de frequência.

Gabarito: Errado.

38. (CESPE/TC-DF/2002). Julgue o item seguinte, relativo a estatística e matemática financeira.

Por Estatística Descritiva entende-se um conjunto de ferramentas, tais como gráficos e tabelas, cujo objetivo é apresentar, de forma resumida, um conjunto de observações.

Comentários:

A definição está perfeita. A Estatística Descritiva estuda técnicas para coleta, organização e sumarização dos dados.

Gabarito: Certo.

39. (CESPE/SEFAZ-AL/2002). Julgue os seguintes itens.

Um censo consiste no estudo de todos os indivíduos da população considerada.

Comentários:

O conceito apresentado está correto. O censo consiste na obtenção dos dados de todos os elementos da população considerada.

Gabarito: Certo.

40. (CESPE/SEFAZ-AL/2002). Julgue os seguintes itens.

Como a realização de um censo tipicamente é muito onerosa e (ou) demorada, muitas vezes é conveniente estudar um subconjunto próprio da população, denominado amostra.

Comentários:

O item está correto. Como normalmente é muito caro e demorado obter os dados de toda a população, trabalhamos apenas com uma parte dela, que chamamos de amostra.

Gabarito: Certo.

LISTA DE QUESTÕES

1. (CESPE/Polícia Federal/2018)

	DIA				
	1	2	3	4	5
X (quantidade diária de drogas apreendidas, em kg)	10	22	18	22	28

Tendo em vista que, diariamente, a Polícia Federal apreende uma quantidade X, em kg, de drogas em determinado aeroporto do Brasil, e considerando os dados hipotéticos da tabela precedente, que apresenta os valores observados da variável X em uma amostra aleatória de 5 dias de apreensões no citado aeroporto, julgue o item.

A tabela em questão descreve a distribuição de frequências da quantidade de drogas apreendidas nos cinco dias que constituem a amostra.

2. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

O gráfico de barras é adequado para a análise de variáveis qualitativas ordinais ou quantitativas discretas, pois permite investigar a presença de tendência nos dados.

3. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

A representação de diagramas de barras, de linha e de pizza possui escala de medida nominal e tem a moda como medida de tendência central.

4. (CESPE/IPHAN/2018). A tabela a seguir mostra as quantidades de bibliotecas públicas presentes em 20 microrregiões brasileiras.

90	66	78	82
77	60	64	90
87	85	67	91
82	70	81	80
69	78	90	67

A partir desses dados, pretende-se construir um gráfico de distribuição de frequências com quatro classes de igual amplitude. Os valores mínimo e máximo de cada classe devem ser números inteiros.

Considerando essas informações, julgue o item subsequente, relativo ao gráfico de distribuição a ser apresentado.

A amplitude de cada classe deverá ser superior a 6.

5. (CESPE/IPHAN/2018) A tabela a seguir mostra as quantidades de bibliotecas públicas presentes em 20 microrregiões brasileiras.

90	66	78	82
77	60	64	90
87	85	67	91
82	70	81	80
69	78	90	67

A partir desses dados, pretende-se construir um gráfico de distribuição de frequências com quatro classes de igual amplitude. Os valores mínimo e máximo de cada classe devem ser números inteiros. Considerando essas informações, julgue o item subsequente, relativo ao gráfico de distribuição a ser apresentado.

A última classe deverá variar de 84 a 91.

6. (CESPE/IPHAN/2018). Julgue o item subsequente, referente à análise exploratória de dados.

O histograma é um diagrama de retângulos contíguos com base na curtose das faixas de valores da variável e com área igual à diferença da frequência absoluta da respectiva faixa.

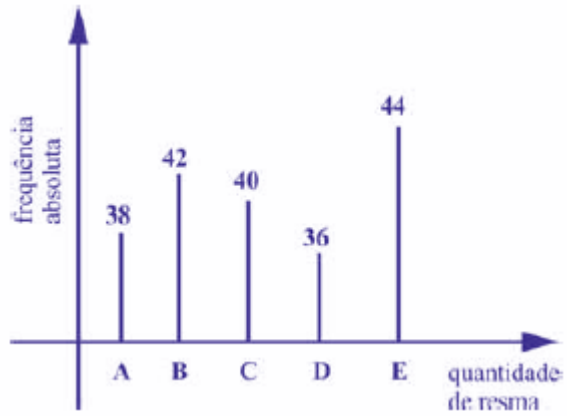
7. (CESPE/CBM-AL/2017). Na tabela a seguir, A, B, C, D e E são as quantidades de resmas de papel A4 consumidas, em quatro meses, pelas seções administrativas I, II, III, IV e V, respectivamente. Apesar de não mostrar explicitamente essas quantidades, a tabela apresenta as frequências absolutas e (ou) relativas de algumas dessas quantidades.

Seção	Quantidades de Resmas	Frequência Absoluta	Frequência Relativa
I	A	38	19%
II	B		
III	C		20%

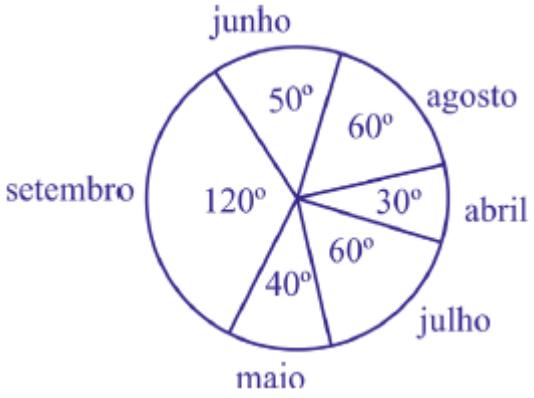
IV	D	36	
V	E	44	
	Total		**100%**

Considerando que cada uma dessas resmas, juntamente com a embalagem, tem forma de um paralelepípedo retângulo reto que mede 5 cm × 21 cm × 30 cm, julgue o item seguinte.

O gráfico de barras verticais a seguir apresenta as frequências absolutas de resmas consumidas pelas cinco seções.



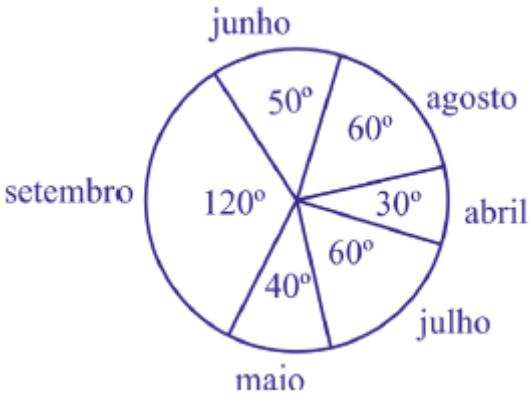
8. (CESPE/CBM-AL/2017) O gráfico de setores a seguir mostra a distribuição das quantidades de incêndios em determinada região, nos meses de abril a setembro de determinado ano.



Sabendo-se que nesses meses ocorreram 1.548 incêndios nessa região, julgue o item que se segue.

A frequência relativa à classe “incêndios no mês de setembro” é superior a 30%.

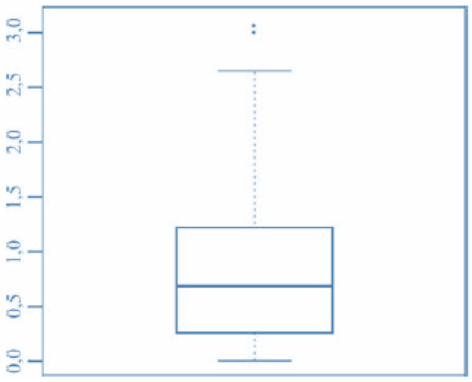
9. (CESPE/CBM-AL/2017) O gráfico de setores a seguir mostra a distribuição das quantidades de incêndios em determinada região, nos meses de abril a setembro de determinado ano.



Sabendo-se que nesses meses ocorreram 1.548 incêndios nessa região, julgue o item que se segue.

Nos meses de maio e junho ocorreram mais de 400 incêndios nessa região.

10. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

X representa uma variável qualitativa ordinal.

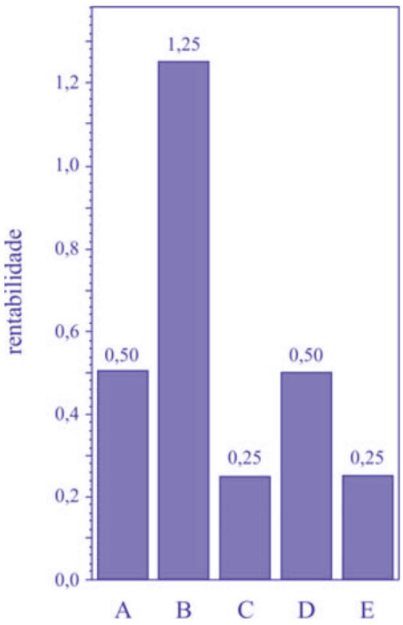
11. (CESPE/TCE-PA/2016)

Número diário de denúncias registradas (X)	Frequência Relativa
0	0,3
1	0,1
2	0,2
3	0,1
4	0,3
Total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

A variável X é do tipo qualitativo nominal.

12. (CESPE/FUNPRESP/2016)



O gráfico ilustra cinco possibilidades de fundos de investimento com suas respectivas rentabilidades. Considerando que as probabilidades de investimento para os fundos A, B, C e D sejam, respectivamente, P(A) = 0,182; P(B) = 0,454; P(C)= 0,091; e P(D)= 0,182, julgue o item subsequente.

O gráfico apresentado é um histograma.

13. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

Na análise exploratória, o histograma é um gráfico adequado para descrever a distribuição da quantidade de detentos por região em 2013.

14. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

Em 2013, mais de 55% da população carcerária no Brasil se encontrava na região Sudeste.

15. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

A quantidade total de vagas existentes no sistema penitenciário brasileiro em 2013 era de 340 mil vagas.

16. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17

Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

O déficit relativo de vagas observado na região Sudeste, em 2013, foi superior ao déficit relativo de vagas registrado na região Centro-oeste no mesmo período.

17. (CESPE/DEPEN/2015)

Região	Quantidade de detentos no sistema penitenciário brasileiro (mil pessoas)	Déficit de vagas no sistema penitenciário (mil vagas)	População brasileira (milhões de habitantes)
Norte	37	13	17
Centro-Oeste	51	24	15
Nordeste	94	42	55
Sudeste	306	120	85
Sul	67	16	28
Total	555	215	200

Ministério da Justiça — Departamento Penitenciário Nacional — Sistema Integrado de Informações Penitenciárias – InfoPen, Relatório Estatístico Sintético do Sistema Prisional Brasileiro, dez./2013

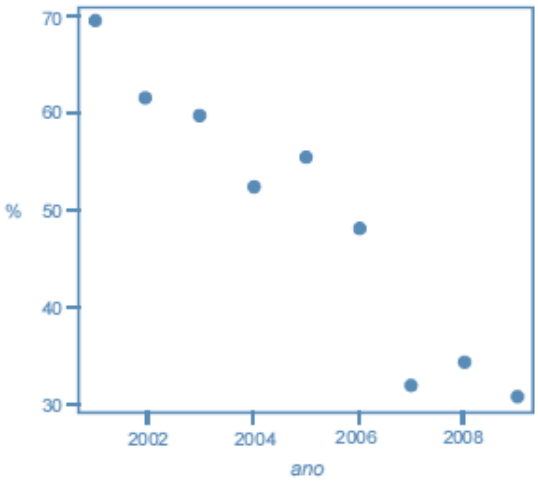
Internet: <www.justica.gov.br> (com adaptações).

A tabela mostrada apresenta a quantidade de detentos no sistema penitenciário brasileiro por região em 2013. Nesse ano, o déficit relativo de vagas — que se define pela razão entre o déficit de vagas no sistema penitenciário e a quantidade de detentos no sistema penitenciário — registrado em todo o Brasil foi superior a 38,7%, e, na média nacional, havia 277,5 detentos por 100 mil habitantes.

Com base nessas informações e na tabela apresentada, julgue o item a seguir.

No ano considerado, a quantidade média de detentos por 100 mil habitantes na região Nordeste foi superior ao número médio de detentos por 100 mil habitantes na região Centro-oeste.

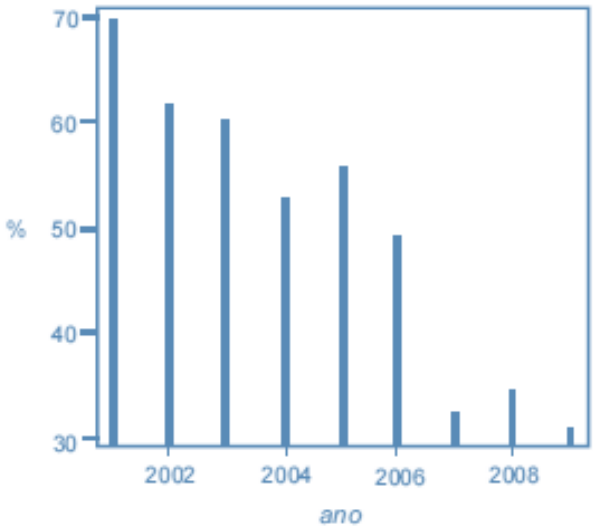
18. (CESPE/DEPEN/2015)



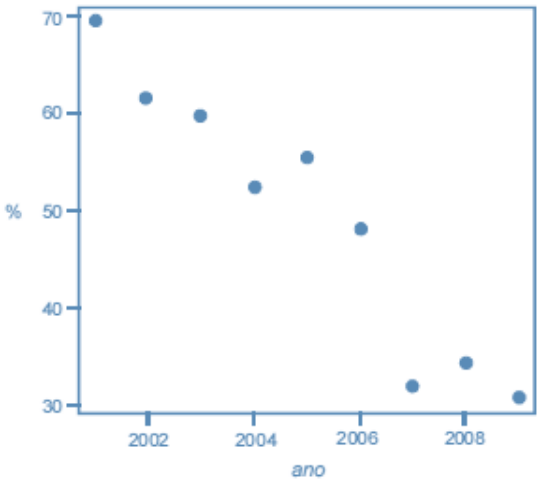
Dado que a participação dos presidiários em cursos de qualificação profissional é um aspecto importante para a reintegração do egresso do sistema prisional à sociedade, foram realizados levantamentos estatísticos, nos anos de 2001 a 2009, a respeito do valor da educação e do trabalho em ambientes prisionais. Cada um desses levantamentos, cujos resultados são apresentados no gráfico, produziu uma estimativa anual do percentual P de indivíduos que participaram de um curso de qualificação profissional de curta duração, mas que não receberam o diploma por motivos diversos. Em 2001, 69,4% dos presidiários que participaram de um curso de qualificação profissional não receberam o diploma. No ano seguinte, 2002, esse percentual foi reduzido para 61,5%, caindo, em 2009, para 30,9%.

A partir das informações e do gráfico apresentados, julgue o item que se segue.

Se os percentuais forem representados por barras verticais, conforme o gráfico a seguir, então o resultado será denominado histograma.



19. (CESPE/DEPEN/2015)

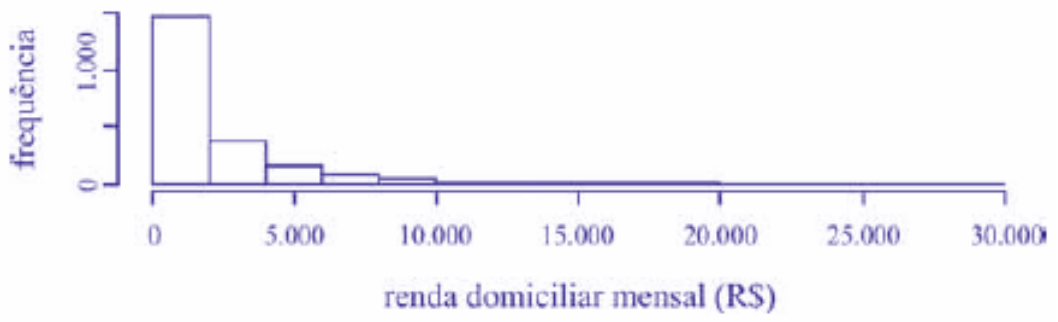


Dado que a participação dos presidiários em cursos de qualificação profissional é um aspecto importante para a reintegração do egresso do sistema prisional à sociedade, foram realizados levantamentos estatísticos, nos anos de 2001 a 2009, a respeito do valor da educação e do trabalho em ambientes prisionais. Cada um desses levantamentos, cujos resultados são apresentados no gráfico, produziu uma estimativa anual do percentual P de indivíduos que participaram de um curso de qualificação profissional de curta duração, mas que não receberam o diploma por motivos diversos. Em 2001, 69,4% dos presidiários que participaram de um curso de qualificação profissional não receberam o diploma. No ano seguinte, 2002, esse percentual foi reduzido para 61,5%, caindo, em 2009, para 30,9%.

A partir das informações e do gráfico apresentados, julgue o item que se segue.

Os dados apresentados são suficientes para que se possa afirmar que o total de presidiários que participaram de um curso de qualificação profissional de curta duração e que não receberam o diploma em 2008 foi superior ao total referente ao ano de 2007.

20. (CESPE/TELEBRAS/2015)



Uma empresa coletou e armazenou em um banco de dados diversas informações sobre seus clientes, entre as quais estavam o valor da última fatura vencida e o pagamento ou não dessa fatura. Analisando essas informações, a empresa concluiu que 15% de seus clientes estavam inadimplentes. A empresa recolheu ainda dados como a unidade da Federação (UF) e o CEP da localidade em que estão os clientes. Do conjunto de todos os clientes, uma amostra aleatória simples constituída por 2.175 indivíduos prestou também informações sobre sua renda domiciliar mensal, o que gerou o histograma apresentado. Com base nessas informações e no histograma, julgue o item a seguir.

O CEP da localidade dos clientes e o valor da última fatura vencida são variáveis quantitativas.

21. (CESPE/TELEBRAS/2015) Roberto comprou, por R$ 2.800,00, rodas de liga leve para seu carro, e, ao estacionar no shopping, ficou indeciso sobre onde deixar o carro, pois, caso o coloque no estacionamento público, correrá o risco de lhe roubarem as rodas, ao passo que, caso o coloque no estacionamento privado, terá de pagar R$ 70,00, com a garantia de que eventuais prejuízos serão ressarcidos pela empresa administradora.

Considerando que p seja a probabilidade de as rodas serem roubadas no estacionamento público, que X seja a variável aleatória que representa o prejuízo, em reais, ao deixar o carro no estacionamento público, e que Y seja a variável aleatória que representa o valor, em reais, desembolsado por Roberto ao deixar o carro no estacionamento pago, julgue o item subsequente.

A variável aleatória Y é contínua.

22. (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.

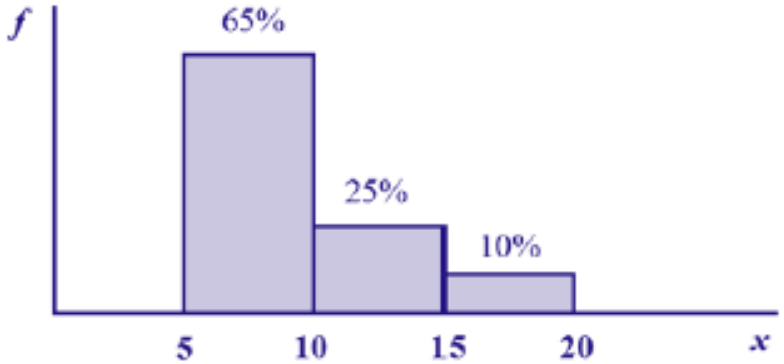
23. (CESPE/TJ-SE/2014)

Quantidade	São Paulo (j =1)	Rio de Janeiro (j =2)	Minas Gerais (j =3)	Rio Grande do Sul (j =4)	Total
Casos novos (X, em milhões)	5	2	1	2	**18**
Casos pendentes (Y, em milhões)	16	8	3	2	**48**
Processos baixados (Z, em milhões)	5	2	2	2	**18**
Sentenças e Decisões (W, em milhões)	4	3	1	1	**16**

O quadro acima mostra uma síntese da movimentação processual dos tribunais de justiça dos estados de São Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul e do total da justiça estadual no Brasil em 2010. Considere que o estoque de processos em andamento no estado j (Ej), no final de 2010, seja um indicador que se define como $E_j = X_j + Y_j - Z_j - W_j$, em que j = 1, 2, ..., 27; X_j representa o número de casos novos registrados em 2010 no estado j; Yj seja a quantidade de casos pendentes no estado j (i.e., casos anteriores que não foram solucionados até o final de 2010); Z_j denota o total de processos baixados (arquivados) no estado j durante 2010 e W_j seja o número de sentenças e decisões proferidas no estado j até o final de 2010. Considere, por fim, que, para todos os efeitos, o Distrito Federal seja um estado. Com base nessas informações e no quadro acima, julgue o item que se segue.

O quadro apresentado é uma tabela de contingência que mostra o cruzamento entre uma variável qualitativa nominal com 4 níveis de resposta (estados) e outra variável qualitativa com quatro níveis de resposta (casos novos, pendentes, baixados e resolvidos).

24. (CESPE/STF/2013)



Com referência à figura acima, que mostra a distribuição da renda mensal — x, em quantidades de salários mínimos (sm) — das pessoas que residem em determinada região, julgue o item subsequente.

A variável x, por possuir quatro níveis de respostas, é do tipo qualitativa ordinal.

25. (CESPE/BACEN/2013)

2 4 8 4 8 1 2 32 12 1 5 7 5 5 3 4 24 19 4 14

Os dados mostrados acima representam uma amostra, em minutos, do tempo utilizado na armazenagem de formulários no almoxarifado central de certa instituição por diversos funcionários.

Com base nesses dados, julgue o próximo item.

A distribuição de frequência acumulada para tempo de armazenagem observado na amostra inferior a 8 minutos é igual a 13, o que corresponde a uma frequência relativa superior a 0,60.

26. (CESPE/BACEN/2013)

2 4 8 4 8 1 2 32 12 1 5 7 5 5 3 4 24 19 4 14

Os dados mostrados acima representam uma amostra, em minutos, do tempo utilizado na armazenagem de formulários no almoxarifado central de certa instituição por diversos funcionários.

Com base nesses dados, julgue o próximo item.

É inviável a elaboração de um histograma em decorrência do fato de ser este um conjunto de dados quantitativos discretos; dessa forma, apenas por meio de um gráfico de barras pode ser realizada a representação gráfica.

27. (CESPE/TRE-ES/2011)

Quantidade de eleitores	Quantidade de municípios
$0 \vdash 2.000$	364
$2.000 \vdash 4.000$	1.000
$4.000 \vdash 6.000$	3.000
$6.000 \vdash 8.000$	1.000
$8.000 \vdash 10.000$	200
Total	**5.564**

A tabela acima apresenta uma distribuição hipotética das quantidades de eleitores que não votaram no segundo turno da eleição para presidente da República bem como os números de municípios em que essas

quantidades ocorreram. Com base nessa tabela, julgue o item seguinte, relativo à análise exploratória de dados.

Na tabela de frequências, o uso de intervalos de classe permite concluir que a variável em questão é contínua.

28. (CESPE/TRE-ES/2011)

Cargo	Candidatos	Candidatos aptos	Eleitos
Presidente da República	9	9	1
Governador de Estado	170	156	27
Senador	272	234	54
Deputado Federal	6.021	5.058	513
Deputado Estadual/Distrital	15.268	13.076	1.059
Total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.

A variável "cargo" classifica-se como uma variável qualitativa ordinal.

29. (CESPE/TRE-ES/2011)

Cargo	Candidatos	Candidatos aptos	Eleitos
Presidente da República	9	9	1
Governador de Estado	170	156	27
Senador	272	234	54
Deputado Federal	6.021	5.058	513
Deputado Estadual/Distrital	15.268	13.076	1.059
Total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.

Considerando-se a representação das quantidades de eleitos para cada cargo em um gráfico de pizza, a fatia desse gráfico correspondente ao cargo de deputado federal terá ângulo superior a 120°.

30. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

Os extremos mínimo e máximo da variável X foram, respectivamente, iguais a 1 e 80.

31. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

O número de municípios que têm obras sob suspeita de irregularidades é superior a 120.

32. (CESPE/ALECE/2011)

X	0	1	2	3	4	5
Frequência Absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue o item a seguir.

O total de municípios considerado no levantamento foi superior a 180.

33. (CESPE/ANTAQ/2009)

	Variável	2003	2004	2005	2006	2007
Exportação	X	40	46	50	52	54
Importação	Y	20	21	22	24	27
Total	X+Y	60	67	72	76	81

Internet: <www.portodesantos.com> (com adaptações).

Considerando a tabela acima, que apresenta a movimentação anual de cargas no porto de Santos de 2003 a 2007, em milhões de toneladas/ ano e associa as quantidades de carga movimentadas para exportação e importação às variáveis X e Y, respectivamente, julgue o item subsequente.

As séries estatísticas apresentadas na tabela formam três séries temporais.

34. (CESPE/TCU/2008) Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (X) e vendidos (Y) em determinado município, nos anos de 2005 a 2007.

Ano	Número de imóveis	
	Ofertados (X)	Vendidos (Y)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correios Braziliense, 29/4/2008, p.17 (com adaptações)

Com respeito ao texto, considere que cada imóvel ofertado em determinado ano seja classificado como vendido ou não-vendido, e, a um imóvel e classificado como vendido seja atribuído um valor Z = 1, e, ao imóvel classificado como não-vendido, seja atribuído um valor Z = 0. Supondo-se que as classificações dos imóveis como vendido ou não-vendido em um dado ano possam ser consideradas como sendo realizações de uma amostragem aleatória simples, julgue os itens a seguir.

A variável Z é classificada como variável qualitativa nominal, pois representa o atributo do imóvel como vendido ou não-vendido.

35. (CESPE/TCU/2008) Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (X) e vendidos (Y) em determinado município, nos anos de 2005 a 2007.

Ano	Número de imóveis	
	Ofertados (X)	Vendidos (Y)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correios Braziliense, 29/4/2008, p.17 (com adaptações)

Considerando as informações do texto, julgue os itens subsequentes.

A variável X forma uma série estatística denominada série temporal.

36. (CESPE/SEFAZ-MT/2004) Considere a seguinte situação hipotética.

Um órgão do governo recebeu pela Internet denúncias de sonegação de impostos estaduais contra 600 pequenas empresas. Denúncias contra outras 200 pequenas empresas foram encaminhadas pessoalmente para esse órgão. Para a apuração das denúncias, foram realizadas auditorias nas 800 empresas denunciadas. Como resultado dessas auditorias, foi elaborada a tabela abaixo, que apresenta um quadro das empresas denunciadas e os correspondentes débitos fiscais ao governo. Das empresas denunciadas, observou-se que apenas 430 tinham débitos fiscais.

	Valor do débito fiscal (VDF), em R$ mil, apurado após auditoria na empresa denunciada				
Forma de recebimento da denúncia	$0 < VDF < 1$	$1 \leq VDF < 2$	$2 \leq VDF < 3$	$3 \leq VDF \leq 4$	Total
Pela internet	60	100	50	30	240
Pessoalmente	20	120	40	10	190
Total	80	220	90	40	430*

Nota: *Para as demais empresas, VDF=0

Com base na situação hipotética acima e de acordo com as informações apresentadas, julgue o item que se segue.

O valor total dos débitos fiscais apurados após as auditorias feitas nas empresas denunciadas é inferior a R$ 500 mil.

37. (CESPE/SEFAZ-AL/2002) Julgue o seguinte item.

Em uma distribuição de frequências para um conjunto de n indivíduos, pode-se calcular as frequências relativas, dividindo-se cada frequência absoluta pela amplitude da correspondente classe ou do intervalo.

38. (CESPE/TC-DF/2002) Julgue o item seguinte, relativo a estatística e matemática financeira.

Por Estatística Descritiva entende-se um conjunto de ferramentas, tais como gráficos e tabelas, cujo objetivo é apresentar, de forma resumida, um conjunto de observações.

39. (CESPE/SEFAZ-AL/2002) Julgue os seguintes itens.

Um censo consiste no estudo de todos os indivíduos da população considerada.

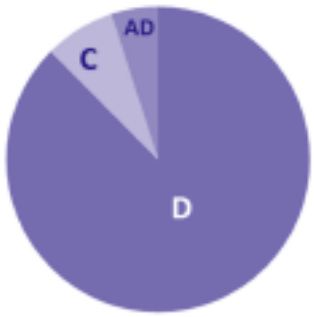
40. (CESPE/SEFAZ-AL/2002) Julgue os seguintes itens.

Como a realização de um censo tipicamente é muito onerosa e (ou) demorada, muitas vezes é conveniente estudar um subconjunto próprio da população, denominado amostra.

QUESTÕES COMPLEMENTARES COMENTADAS

Questões da FCC

41. (FCC/Pref. Macapá/2018) Para fazer um gráfico de setores que representasse o número de alunos canhotos, destros e ambidestros na sala de aula, Renato coletou os dados indicados na tabela abaixo.



	Nº DE ALUNOS DA SALA DE AULA
Destros (D)	35
Canhotos (C)	3
Ambidestros (AD)	2

Sabendo-se que um círculo pode ser dividido em 360°, quantos graus tem o setor circular correspondente aos canhotos no gráfico correto feito por Renato?

a) 18°.

b) 25°.

c) 7,5°.

d) 24°.

e) 27°.

Comentários:

O gráfico em setores é usado para representar a frequência relativa (porcentagem) de uma variável categórica. Para a construção do gráfico de setores, utilizaremos uma regra de três simples, em que as frequências relativas de cada categoria correspondem ao ângulo central que desejamos representar em relação à frequência total, que corresponde ao ângulo de 360°.

Podemos considerar que 360° corresponde a 100% dos alunos. Temos um total de 40 alunos.

Queremos saber o percentual de canhotos. Então, montamos uma regra de três:

$$360° \ - \ 40$$
$$X \ - \ 3$$

Calculando, temos que:

$$\frac{360}{x} = \frac{40}{3}$$

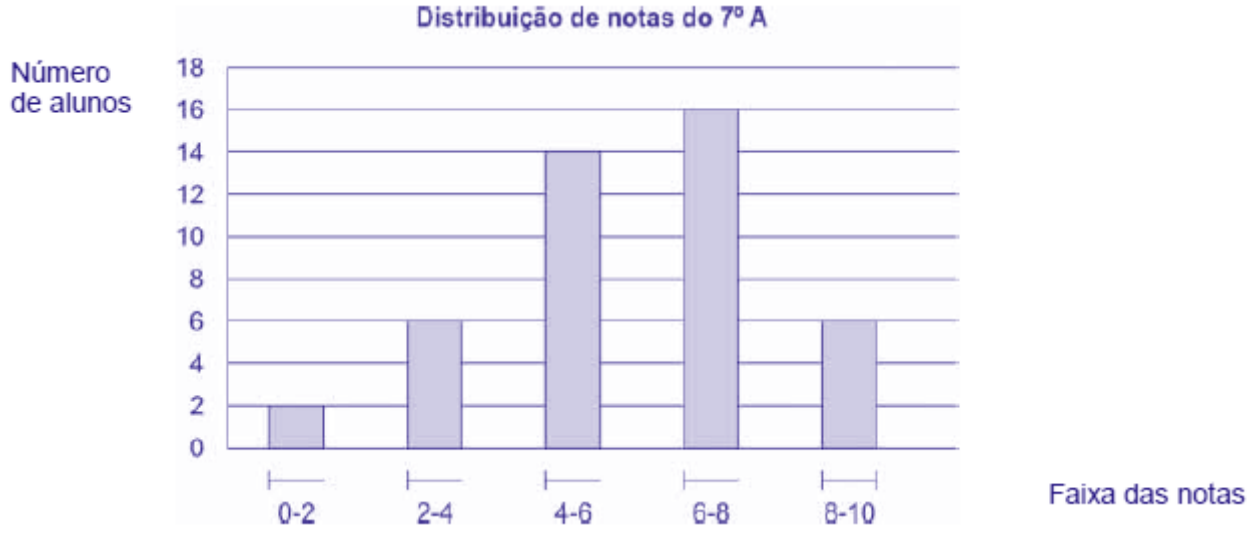
$$40x = 1080$$

$$x = \frac{1080}{40}$$

$$x = 27°$$

Gabarito: E.

42.(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, incialmente, calcular o total de alunos:

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

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

$$16 + 6 = 22$$

Logo, a porcentagem será de:

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

Gabarito: B.

43. (FCC/TRT 14ª Região/2018) De um histograma e uma tabela de frequências absolutas, elaborados para analisar a distribuição dos salários dos empregados em uma empresa, obtém-se a informação que 24 empregados ganham salários com valores pertencentes ao intervalo (2.000; 4.000], em reais, que apresenta uma densidade de frequência de $0,75 \times 10^{-4}(R\$)^{-1}$

Densidade de frequência de um intervalo é o resultado da divisão da respectiva frequência relativa pela amplitude deste intervalo. Em um intervalo do histograma que está sendo analisado, com uma amplitude de R$ 3.000,00 e uma densidade de frequência de $1 \times 10^{-4}(R\$)^{-1}$, tem-se que o correspondente número de empregados é igual a

a) 40.

b) 36.

c) 30.

d) 48.

e) 42.

Comentários:

Vamos destacar os dados da questão:

- densidade de frequência: $0{,}75 \times 10^{-4}(R\$)^{-1}$;

- amplitude da classe: $4000 - 2000 = 2000$;

- número de funcionários na classe: 24; e

- total de funcionários na classe: T.

A frequência relativa é calculada pela fórmula a seguir:

$$Frequência\ relativa\ = \frac{Nº\ de\ funcionários\ na\ classe}{Total\ de\ funcionários} = \frac{24}{T}$$

Então, pela fórmula da densidade de frequência, teremos:

$$Densidade\ de\ frequência = \frac{Frequência\ relativa}{Amplitude}$$

$$0{,}75 \times 10^{-4} = \frac{\frac{24}{T}}{2000}$$

$$0{,}15 = \frac{24}{T}$$

$$T = 160$$

Agora, vamos aplicar a mesma metodologia para a classe do histograma com amplitude de 3000 e densidade de frequência de 1×10^{-4}. Considerando o número de funcionários na classe igual a N, teremos a seguinte frequência relativa:

$$Frequência\ relativa\ = \frac{Nº\ de\ funcionários\ na\ classe}{Total\ de\ funcionários} = \frac{N}{160}$$

Então, pela fórmula da densidade de frequência, teremos:

$$Densidade\ de\ frequência = \frac{Frequência\ relativa}{Amplitude}$$

$$10^{-4} = \frac{\frac{N}{160}}{3000}$$

$$0{,}3 = \frac{N}{160}$$

$$N = 48$$

Gabarito: D.

44. (FCC/ARTESP/2017) Foi solicitado para uma empresa de transportes que fizesse um levantamento da idade da frota dos seus caminhões que operavam em um trecho de rodovia com tráfego intenso. O gerente da empresa entregou a seguinte tabela:

Caminhão	A	B	C	D	E	F	G	H	I	J	K	L	M
Idade (anos)	15	20	5	2	3	3	17	23	16	8	6	9	14
Caminhão	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	
Idade (anos)	10	10	6	8	2	4	18	5	9	11	26	14	

De posse deste levantamento, a analista de operações da autarquia solicitante organizou uma distribuição de frequência para organizar melhor os dados para serem analisados. Sendo assim, em um primeiro momento, fez-se necessário encontrar o número de intervalos (K) e a classe (C), que são expressos, respectivamente, por

a) 5 e 4,80.

b) 5 e 5,20.

c) 6 e 4,80.

d) 6 e 5,20.

e) 4 e 5.

Comentários:

Iniciaremos a resolução calculando a amplitude total. Sabemos que o menor valor é 2 e o maior valor é 26. Assim, a amplitude total fica:

$$AT = 26 - 2 = 24.$$

Temos 25 caminhões, logo, o valor correspondente de intervalos será:

$$K = \sqrt{25} = 5.$$

Portanto, com K intervalos de amplitude C, conseguiremos cobrir as 24 unidades:

$$AT = C \times K$$

Podemos montar a equação da seguinte forma:

$$C = \frac{AT}{K} = \frac{24}{5}$$

$$C = 4{,}8$$

Portanto, a alternativa correta contém as combinações $K = 5$ e $C = 4{,}8$.

Gabarito: A.

45. (FCC/TRT 20ª Região/2016) Um gráfico corresponde a um histograma apresentando a distribuição dos salários dos funcionários lotados em um determinado órgão público. No eixo das abscissas constam os intervalos de classe (fechados à esquerda e abertos à direita) dos salários em R$ e no eixo das ordenadas as respectivas densidades de frequências em (R$) $^{-1}$. Densidade de frequência de um intervalo é definida como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Se 135 funcionários ganham salários com valores pertencentes ao intervalo [3.000, 6.000) com uma densidade de frequência de 1 × 10 $^{-4}$ (R$) $^{-1}$, então o número de funcionários que ganham salários com valores pertencentes ao intervalo [6.000, 8.000) com uma densidade de frequência de 2 × 10 $^{-4}$ (R$) $^{-1}$ é igual a:

a) 300

b) 180

c) 270

d) 150

e) 90

Comentários:

O enunciado da questão traz informações importantes para a resolução do problema.

Sabemos que a densidade de frequência (d), é a relação entre a **frequência relativa** (f_r) e a amplitude de classe (h). Assim:

$$d = \frac{f_r}{h}$$

Temos a informação de que 135 funcionários pertencem à classe [3.000, 6.000). Logo, a amplitude dessa classe será:

$$h = 6.000 - 3.000 = 3.000$$

Se 135 funcionários ganham salários com valores entre o intervalo [3.000, 6.000) com uma densidade de frequência de 10^{-4}, então precisamos encontrar a frequência relativa (f_r). Para isso, podemos fazer:

$$d = \frac{f_r}{h}$$

$$10^{-4} = \frac{f_r}{3000}$$

$$f_r = 3.000 \times 10^{-4}$$
$$f_r = 0{,}3$$

Então, temos 30% das observações nesta classe.

Além disso, foi afirmado que a classe [6.000, 8.000) tem densidade de frequência 2×10^{-4}. Esta classe tem amplitude:

$$h = 8.000 - 6.000 = 2.000$$

Aplicando a fórmula da densidade, temos:

$$d = \frac{f_r}{h}$$

$$2 \times 10^{-4} = \frac{fr}{2000}$$

$$fr = 2.000 \times 2 \times 10^{-4} = 0{,}4$$

Portanto, já sabemos que esta classe apresenta 40% das observações. Então, montando uma regra de três, temos:

$$30\% \ - \ 135 \ funcionários$$

$$40\% \ - \ x$$

$$0{,}3x = 0{,}4 \times 135$$

$$x = \frac{54}{0{,}3} = 180$$

Gabarito: B.

46. (FCC/TRT 3ª Região/2015) Em um histograma representando os preços unitários de microcomputadores em estoque, observa-se que no eixo das abscissas constam os intervalos de classe em R$ e no eixo das ordenadas as respectivas densidades de frequências em $(R\$)^{-1}$. Densidade de frequência de um intervalo de classe é o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um determinado intervalo de classe com amplitude igual a R$ 2.500,00 apresenta uma densidade de frequência, em $(R\$)^{-1}$, igual a $12{,}8 \times 10^{-5}$. Se o número de microcomputadores deste intervalo é igual a 48, então o número total de microcomputadores em estoque é igual a

a) 150.

b) 120.

c) 240.

d) 160.

e) 96.

Comentários:

Sabemos que a densidade de frequência (d), é a relação entre a **frequência relativa** (f_r) e a amplitude de classe (h). Assim:

$$d = \frac{f_r}{h}$$

Temos que, para determinada classe, a amplitude vale 2.500 e a densidade de frequência vale $12{,}8 \times 10^{-5}$. Aplicando na fórmula, temos:

$$12{,}8 \times 10^{-5} = \frac{f_r}{2500}$$

$$f_r = 12{,}8 \times 10^{-5} \times 2.500 = 0{,}32$$

Sabemos também que a frequência relativa vale 0,32. Isso significa que 32% das observações estão nesta classe. O enunciado informa que 32% correspondem a 48 microcomputadores. Para determinar o número total de computadores, fazemos uma regra de três:

$$32\% \ - \ 48$$

$$100\% \ - \ x$$

$$0{,}32x = 48$$

$$x = \frac{48}{0{,}32}$$

$$x = 150$$

Gabarito: A.

47. (FCC/TRT 16ª Região/2014) A distribuição das medidas dos comprimentos, em cm, de uma determinada peça em estoque de uma fábrica está representada em um histograma com todos os intervalos de classe fechados à esquerda e abertos à direita. No eixo horizontal constam os intervalos de classe e no eixo vertical as respectivas densidades de frequências, em cm^{-1}. Define-se densidade de frequência de um intervalo de classe como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude deste intervalo. Verifica-se com relação ao histograma, que o intervalo de classe [2 , 6), em cm, apresenta uma densidade de frequência igual a 0,028 cm^{-1}. Dado que o número de peças em estoque com medidas iguais ou superiores a 2 cm e inferiores a 6 cm é igual a 84, obtém-se que o número total destas peças em estoque é

a) 750.

b) 400.

c) 500.

d) 250.

e) 1.000.

Comentários:

Vamos calcular a densidade de frequência:

$$d = \frac{f_r}{h}$$

Em que d é a densidade de frequências; f_r é a frequência relativa; e h é a amplitude.

Aplicando os valores do enunciado, temos:

$$0{,}28 = \frac{f_r}{6-2}$$

$$f_r = 0{,}28 \times 4$$

$$f_r = 0{,}112$$

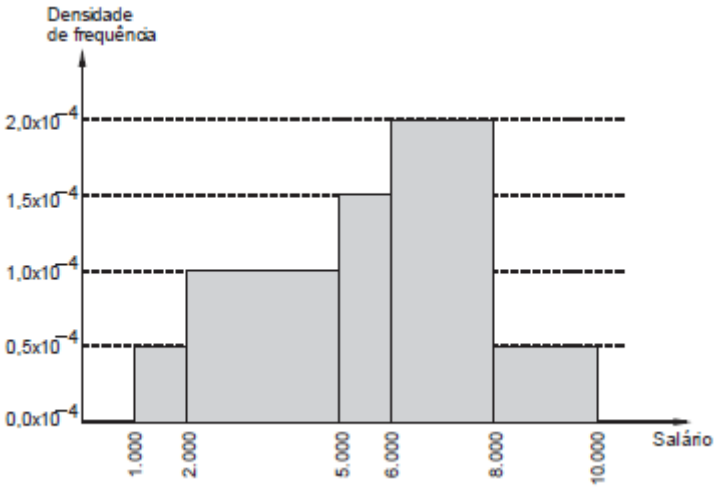
O número de peças em estoque com medidas iguais ou superiores a 2 cm e inferiores a 6 cm é igual a 84, logo:

$$f_r = \frac{84}{T}$$

$$T \times f_r = 84$$

$$T = \frac{84}{0{,}112} T = 750 \textbf{Gabarito:}$$
$$T = 750 \textbf{Gabarito:}$$

48. (FCC/TRT 13ª Região/2014) O histograma, abaixo, refere-se à distribuição dos salários dos funcionários lotados em um setor de um órgão público. No eixo das abscissas constam os intervalos de classe em R$ (todos fechados à esquerda e abertos à direita) e no eixo das ordenadas as respectivas densidades de frequências em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo.



Se o número de funcionários que tem um salário inferior a R$ 5.000,00 é igual a 56, então verifica-se que o número de funcionários que tem um salário igual ou superior a R$ 2.000,00 e inferior a R$ 8.000,00 é igual a

a) 104.

b) 152.

c) 136.

d) 144.

e) 120.

Comentários:

Vamos iniciar a resolução montando uma tabela com os dados fornecidos pelo histograma. Assim:

	Classes	f_r	f_{ac}
1	[0,1000)	0	0
2	[1000,2000)	0,05	0,05
3	[2000,5000)	0,30	0,35
4	[5000,6000)	0,15	0,50
5	[6000,8000)	0,40	0,90
6	[8000,10000)	0,10	1,00
		1,00	

Na tabela, f_r é a frequência relativa e f_{ac} é frequência acumulada.

Analisando a tabela percebemos que o número de empregados com salários inferiores a 5000 corresponde às três primeiras classes. O número de funcionários com salário igual ou superior a 2000 e inferior a 8000 corresponde às 3ª, 4ª e 5ª classes.

Vamos calcular a densidade de frequência:

$$d_f = \frac{f_r}{h}$$

A amplitude da terceira classe é 3000, temos:

$$1{,}0 \times 10^{-4} = \frac{f_r}{3000}$$

$$f_r = 0{,}3$$

Para encontrarmos o número de funcionários com salário igual ou superior a 2000 e inferior a 8000, deve-se calcular a f_{ac}. Já sabemos que a frequência acumulada é obtida pela soma das frequências relativas de todos os valores da variável (salário), menores ou iguais ao valor considerado, como apresentado na quarta coluna da nossa tabela.

Assim, os 56 empregados com salários inferiores a 5000 correspondem à f_{ac} de 0,35. Para achar a frequência acumulada correspondente ao número de funcionários com salário igual ou superior a 2000 e inferior a 8000, deve-se somar as f_r da 3ª, 4ª e 5ª $(0{,}35 + 0{,}15 + 0{,}40 = 0{,}85)$.

Utilizando a regra de três, temos:

$$0{,}35 \rightarrow 56$$

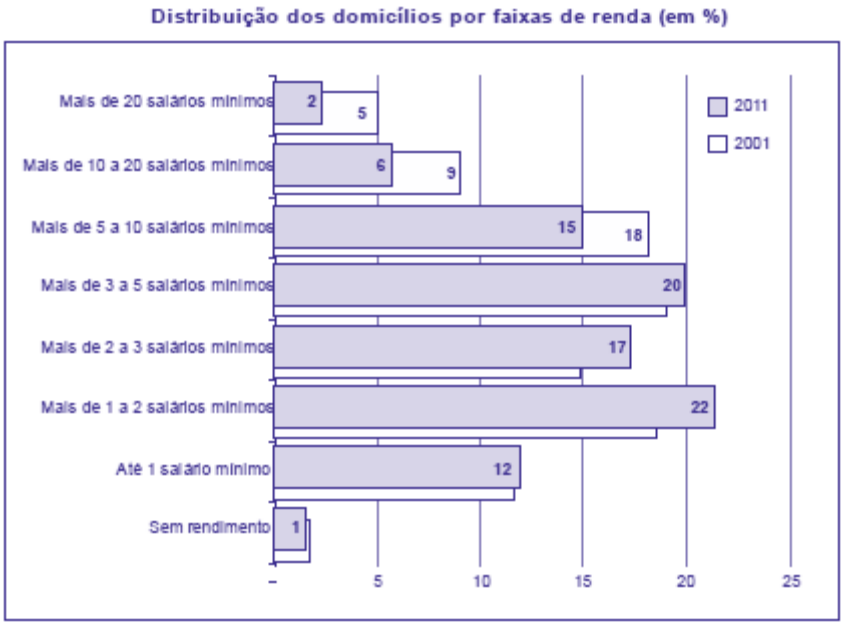
$$0{,}85 \rightarrow x$$

$$0{,}35x = 47{,}60$$

$$x = 136$$

Gabarito: C.

49. (FCC/ALE-PE/2014) Analise o gráfico e as afirmativas abaixo.



(http://blogs.estadao.com.br/vox-publica/tag/psdb/. Acessado em: 20/02/2013)

I. O crescimento da faixa de renda de "até 1 salário mínimo" pode ser reflexo dos programas de transferência de renda governamentais, os quais garantem um rendimento mínimo às pessoas carentes.

II. O crescimento da faixa de renda de "até 1 salário mínimo" é fruto do emprego formal gerado no país dentro do período da pesquisa, pois, os beneficiários dos programas de transferência de renda governamentais entram na categoria sem rendimento.

III. No gráfico, percebe-se uma diminuição da concentração de renda no país. Isso fica evidente quando analisamos, por exemplo, a faixa de renda "mais de 5 a 10 salários mínimos".

IV. No gráfico, a faixa "mais de 1 a 2 salários mínimos" é a que apresenta a maior diferença em pontos percentuais entre os dados de 2001 e 2011.

Está correto o que se afirma APENAS em

a) I e II.

b) II e III.

c) III e IV.

d) I e III.

e) II e IV.

Comentários:

Analisando as afirmações, temos o seguinte:

Item I: no gráfico percebemos que de fato houve um aumento na faixa de renda de “até 1 salário-mínimo”, porém, não temos nenhuma informação que justifique esse aumento. O que torna a alternativa correta é o emprego da possibilidade, não a afirmação categórica.

Item II: como dito no item anterior não temos nenhuma informação sobre o motivo do aumento da renda da respectiva faixa, portanto, o item está errado.

Item III: o gráfico mostra o aumento nas faixas abaixo de "mais de 5 a 10 salários-mínimos" e reduções nas faixas acima de "mais de 3 a 5 salários-mínimos". Assim, podemos sugerir que houve diminuição da concentração de renda no país. Logo, o item está correto.

Item IV: não temos informações suficientes no gráfico para esta afirmativa. Portanto, o item está incorreto.

Gabarito: D.

50. (FCC/TRT 6ª Região/2012) A distribuição dos 500 preços unitários de um equipamento é representada por um histograma em que no eixo das abscissas constam os intervalos de classe e no eixo das ordenadas estão assinaladas as respectivas densidades de frequências, em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo de classe como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um intervalo de classe no histograma apresenta uma amplitude de R$ 2,50 com uma densidade de frequência igual a 0,096. A quantidade de preços unitários referente a este intervalo é

a) 96.

b) 120.

c) 144.

d) 150.

e) 192.

Comentários:

Sabemos que a densidade de frequência (d), é a relação entre a **frequência relativa** (f_r) e a amplitude de classe (h). Assim:

$$d = \frac{f_r}{h}$$

Dados do enunciado:

- Amplitude de classe = 2,50 ($h = 2{,}50$);
- Densidade de frequência = 0,096 ($d_f = 0{,}096$).

Substituindo esses valore na equação, temos:

$$d_f = \frac{f_r}{h}$$

$$f_r = 0{,}096 \times 2{,}50 = 0{,}24 = 24\%$$

Então, 24% dos preços estão localizados neste intervalo.

Calculando a quantidade para 500 preços, temos:

$$24\% \times 500 = 120$$

Gabarito: B.

51. (FCC/TRT 6ª Região/2012) A função de distribuição empírica abaixo, $F_{200}(x)$, refere-se a uma pesquisa realizada em 200 residências, escolhidas aleatoriamente, em que x é o número verificado de pessoas que trabalham em cada residência.

$$F_{200}(x) = \begin{cases} 0,00\ se\ x < 0 \\ 0,10\ se\ 0 \leq x < 1 \\ 0,40\ se\ 1 \leq x < 2 \\ 0,65\ se\ 2 \leq x < 3 \\ 0,85\ se\ 3 \leq x < 4 \\ 0,95\ se\ 4 \leq x < 5 \\ 1,00\ se\ x \geq 5 \end{cases}$$

O número de residências desta pesquisa em que se verificou possuir pelo menos uma pessoa que trabalha e menos que 4 é

a) 150.

b) 160.

c) 170.

d) 180.

e) 190.

Comentários:

Inicialmente, vamos montar uma tabela com as frequências acumuladas para melhor visualizarmos a questão:

Classe	Frequência Acumulada
$0 \leq x < 1$	**0,1**
$1 \leq x < 2$	0,4
$2 \leq x < 3$	0,65
$3 \leq x < 4$	**0,85**
$4 \leq x < 5$	0,95
$x \geq 5$	1

A primeira linha nos diz que 10% das residências não apresentam ninguém que trabalhe.

A quarta linha nos diz que em 85% das residências há menos de 4 pessoas trabalhando.

Se o total de residências é 200 e temos a informação de que em 85% há menos de 4 pessoas trabalhando, logo, há 170 residências nesta classe. Queremos eliminar as residências em que não trabalha ninguém, os 10% = 20 residências. Então:

$$170 - 20 = 150$$

Gabarito: A.

52. (FCC/TRT 1ª Região/2011) Um histograma representa a distribuição dos preços unitários de venda de um determinado equipamento no mercado. No eixo das ordenadas estão assinaladas as respectivas densidades de frequência para cada intervalo em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo de classe como sendo o quociente da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um intervalo de classe do histograma corresponde aos preços unitários maiores ou iguais a R$ 32,00 e inferiores a R$ 44,50 com uma densidade de frequência igual a $1,6 \times 10^{-2}$ $(R\$)^{-1}$. Se todos os intervalos de classe do histograma têm a mesma frequência relativa, então um intervalo de classe com densidade de frequência igual a $5,0 \times 10^{-3}$ $(R\$)^{-1}$ apresenta uma amplitude de

a) R$ 64,00.

b) R$ 48,00.

c) R$ 40,00.

d) R$ 32,00.

e) R$ 24,00.

Comentários:

Calculando a amplitude do intervalo:

$$44{,}50 - 32 = 12{,}50$$

Sabemos que a densidade de frequência (d), é a relação entre a **frequência relativa** (f_r) e a amplitude de classe (h). Assim:

$$d = \frac{f_r}{h}$$

Em que d é a densidade de frequência; f_r é a frequência relativa e h é a amplitude.

Aplicando os valores do enunciado na fórmula, temos:

$$d = \frac{f_r}{h}$$

$$1{,}6 \times 10^{-2} = \frac{f_r}{12{,}50}$$

$$f_r = (1{,}6 \times 10^{-2}) \times 12{,}50 = 20 \times 10^{-2} = 0{,}2$$

A questão informou que todos os intervalos têm a mesma frequência. Assim, podemos concluir que todos eles têm frequência de 20%.

Considerando o intervalo com densidade de 5,0 × 10−3 (R$)$^{-1}$:

$$5 \times 10^{-3} = \frac{0{,}2}{h}$$

$$h = \frac{0{,}2}{5 \times 10^{-3}}$$

$$h = 40$$

Gabarito: C.

53. (FCC/DPE-SP/2010) Sobre estatística aplicada, é correto o que se afirma em:

a) Parâmetros são medidas características de grupos, determinadas por meio de uma amostra aleatória.

b) A estatística descritiva é a técnica pela são coletados dados de uma amostra, a partir do que são tomadas decisões sobre uma determinada população.

c) A caracterização de uma população se dá por meio da observação de todos os seus componentes que a integram.

d) A estatística inferencial compreende um conjunto de técnicas destinadas à síntese de dados numéricos.

e) Censo é o processo utilizado para se medir as características de todos os membros de uma dada população.

Comentários:

Vejamos cada uma das alternativas:

- Letra A: na aula de amostragem, veremos que os parâmetros são as medidas características da população. Os valores obtidos da amostra são as estimativas da população. Logo, a alternativa A está errada.
- Letra B: é a Estatística Inferencial que possui técnicas para tomada de decisões sobre a população a partir de uma amostra. Portanto, a alternativa B está errada.
- Letra C: utilizamos justamente a Estatística Inferencial para caracterizar toda a população a partir da observação de uma parte de seus elementos (amostra). Dessa forma, a alternativa C está errada.
- Letra D: é a Estatística Descritiva que possui técnicas para síntese de dados numéricos. Assim, a alternativa D está errada.
- Letra E: o censo consiste na coleta de dados de todos os elementos de uma população. Portanto, a alternativa E está correta.

Gabarito: E.

Questões da FGV

54. (FGV/COMPESA/2018) A COMPESA, em uma pesquisa de satisfação dos usuários, preparou um formulário para traçar os perfis de seus clientes e o grau de satisfação com os serviços da empresa,

Em um formulário, ela solicitou os dados a seguir.

I. Idade.

II. Grau de escolaridade.

III. Faixa de renda familiar.

IV. Nota dada ao serviço.

Assinale a opção que contempla apenas variáveis categóricas.

a) I e II.

b) II e III.

c) III e IV.

d) I, 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:

i) 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.

ii) 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:

i) variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos.

ii) 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. Grau de escolaridade. (Categórica)

III. Faixa de renda familiar. (Categórica)

IV. Nota dada ao serviço. (Quantitativa)

Gabarito: B.

55. (FGV/IBGE/2017) Uma vez concluída a etapa de críticas de dados, relativa a um conjunto de registros obtido através de uma pesquisa de campo, inicia-se o trabalho de tabulação e elaboração de gráficos. Durante as análises, algumas variáveis surgem com destaque.

I. Número de indivíduos por faixa etária;

II. Percentuais do nível de escolaridade; e

III. Pares de valores de consumo e renda.

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

a) Colunas, linha e setores;

b) Setores, cartograma e barras;

c) Barras, setores e linha;

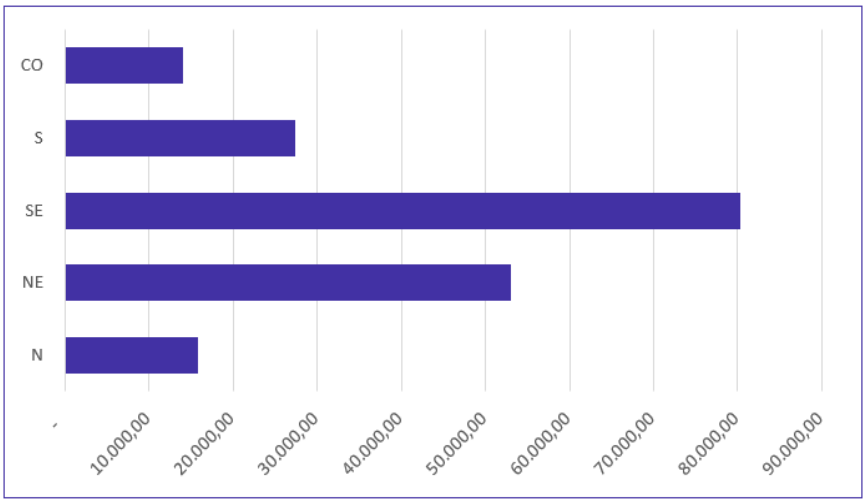
d) Linha, colunas e cartograma;

e) Colunas, setores e linha.

Comentários:

O gráfico de barras é um gráfico com barras retangulares que apresenta dados categorizados em que o comprimento das barras é proporcional aos valores representados. Este gráfico é construído no plano cartesiano. Há duas formas de representar as barras, verticalmente ou horizontalmente. O gráfico de barras verticais também é chamado de gráfico de colunas. Nesse gráfico, um eixo do gráfico mostra especificamente o que está sendo comparado enquanto o outro eixo representa valores associados. Vejamos um exemplo:

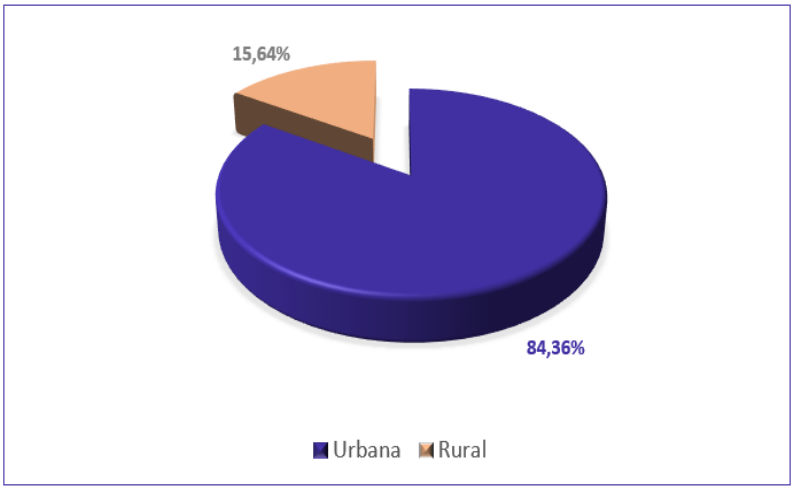
População brasileira, por Grandes Regiões, em 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

O gráfico de setores consiste em representar informações na forma de setor circular. Também chamado de gráfico de pizza, os valores podem ser expressos por números ou em percentuais. Para tanto, basta considerar que 100% dos dados equivalem a um setor circular de 360°. Cada ângulo do setor circular corresponde a uma porcentagem dos dados. Vejamos um exemplo:

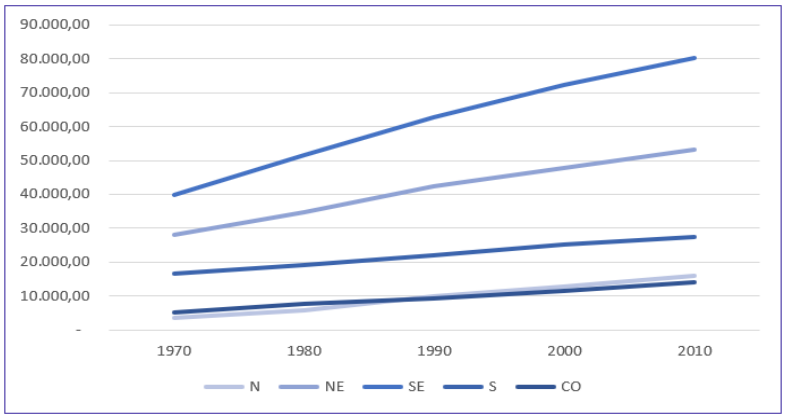
População Urbana e Rural do Brasil em 2010



Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

O gráfico de linha é um tipo de gráfico que nos permite comparar duas variáveis: uma é traçada no eixo x (horizontal) e a outra no eixo y (vertical). O eixo y geralmente indica uma quantidade, enquanto o eixo x representa uma unidade de tempo. Tipicamente, o eixo y representa a variável dependente e o eixo x representa a variável independente. Um gráfico de linhas é normalmente desenhado para representar dados de séries temporais. A finalidade é mostrar a variação dos valores de uma variável ao longo do tempo. Por exemplo,

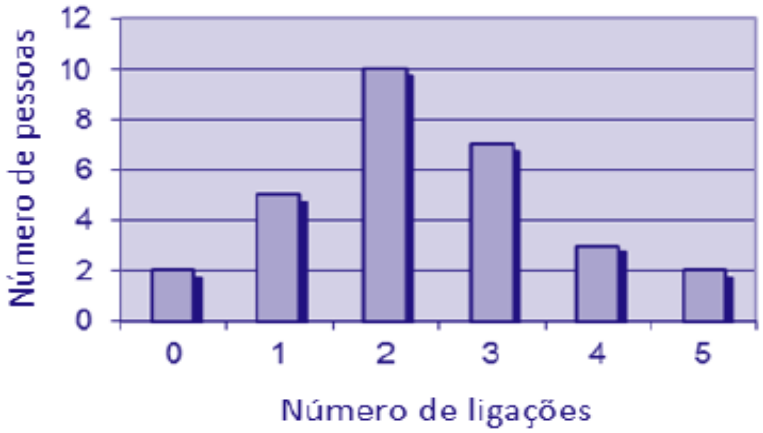
População brasileira, por Grandes Regiões, de 1970 a 2010 (x1000)



Fonte: Censo Demográfico 1970/2010 (IBGE)

Gabarito: E.

56. (FGV/CGE-MA/2014) No setor A de uma empresa foi feita uma auditoria para descobrir quantas vezes cada pessoa fazia ligações pessoais do seu celular no período de trabalho de 14 às 17 horas de um único dia. O resultado está no gráfico a seguir.



O número de pessoas que trabalham no setor A dessa empresa é

a) 15.

b) 22.

c) 27.

d) 29.

e) 42.

Comentários:

Ao analisarmos o gráfico, percebemos a existência de um eixo horizontal, que nos mostra o número de ligações feitas 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	2
1	5
2	10
3	7
4	3
5	2

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

$$2 + 5 + 10 + 7 + 3 + 2 = 29\ pessoas$$

Gabarito: D.

57. (FGV/AL-BA/2014) Observe a tabela de frequências a seguir, que se refere aos saldos em conta, num determinado dia, de duzentas contas-correntes:

Saldos em conta (R$)	Frequência
Até 100,00	8
De mais de 100,00 a 300,00	28
De mais de 300,00 a 500,00	46
De mais de 500,00 a 700,00	54
De mais de 700,00 a 900,00	44
De mais de 900,00 a 1.100,00	13
De mais de 1.100,00 a 1.300,00	6
Acima de 1.300,00	1

A frequência relativa acumulada de saldos em R$ 900,00 é igual a

a) 22%.

b) 36%.

c) 54%.

d) 90%.

e) 97%.

Comentários:

Para resolvermos a questão, precisamos saber a frequência acumulada, da seguinte forma:

Saldos em conta (R$)	Frequência	Frequência acumulada
Até 100,00	8	=8
De mais de 100,00 a 300,00	28	= 8 + 28 = 36
De mais de 300,00 a 500,00	46	= 36 + 46 = 82
De mais de 500,00 a 700,00	54	= 82 + 54 = 136
De mais de 700,00 a 900,00	44	= 136 + 44 =180
De mais de 900,00 a 1.100,00	13	
De mais de 1.100,00 a 1.300,00	6	
Acima de 1.300,00	1	
Total	**200**	

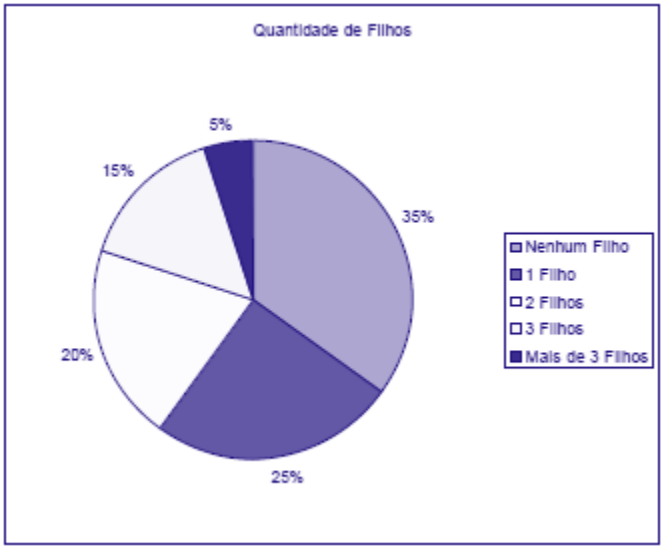
Não é necessário preencher toda a tabela, tendo em vista que a questão pede o saldo em 900,00.

Portanto, verificamos que a frequência absoluta acumulada até 900 é 180. Para determinar a frequência relativa acumulada, basta dividir este valor (180) por 200, pois são 200 observações (soma das frequências):

$$\frac{180}{200} = 0{,}9 = 90\%$$

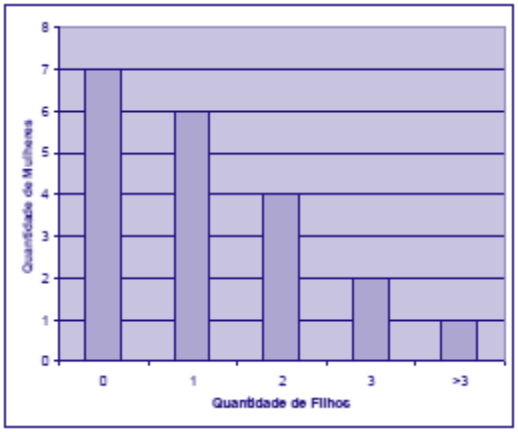
Gabarito: D.

58. (FGV/CODEBA/2010) O gráfico de pizza apresentado abaixo fornece os resultados de uma pesquisa feita com 20 mulheres acerca da quantidade de filhos que cada uma possui.

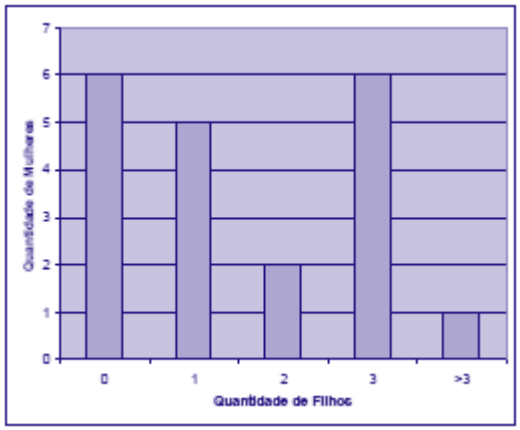


Assinale o gráfico de colunas que melhor representa os dados apresentados acima.

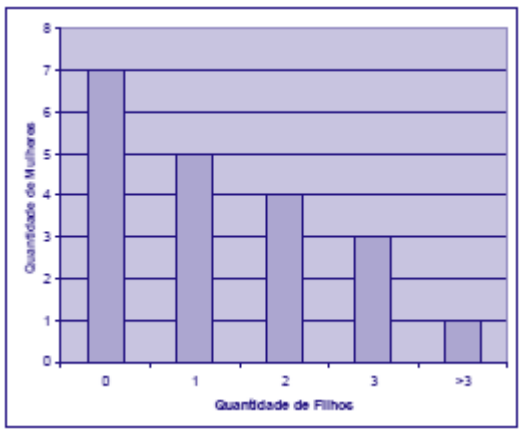
a)



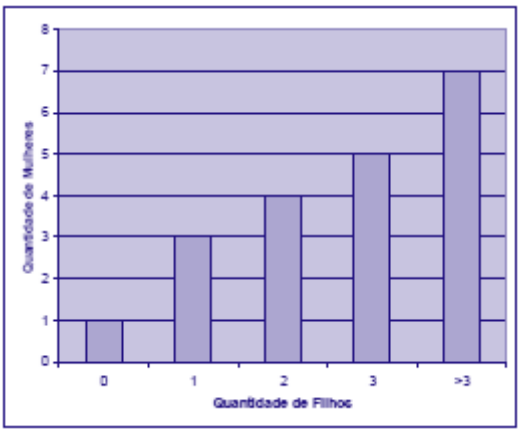
b)



c)



d)



e)



Comentários:

Segundo o enunciado, a pesquisa foi realizada com um total de 20 mulheres.

Agora, interpretando as informações do gráfico, temos:

- nenhum filho = 35%. Logo, 35% de 20 mulheres corresponde a: $0{,}35 \times 20 = 7\ mulheres$;
- um filho = 25%. Logo, 25% de 20 mulheres corresponde a: $0{,}25 \times 20 = 5\ mulheres$.

Olhando para as alternativas, percebemos que não se faz necessário calcular todas as informações do gráfico. Portanto, ficamos entre as alternativas a) ou c), cujas primeiras colunas demonstram que 7 mulheres não tiveram nenhum filho.

Passamos para a segunda informação calculada, em que 5 mulheres tiveram apenas 1 filho. Com isso eliminamos a letra a) e temos nosso gabarito, letra c).

Gabarito: C.

59. (FGV/Senado/2008) Entre os tipos de representações gráficas a seguir, o único que não é definido para uma variável quantitativa contínua é:

a) histograma.

b) diagrama ramo-e-folhas.

c) diagrama de setores.

d) diagrama "Box-Plot".

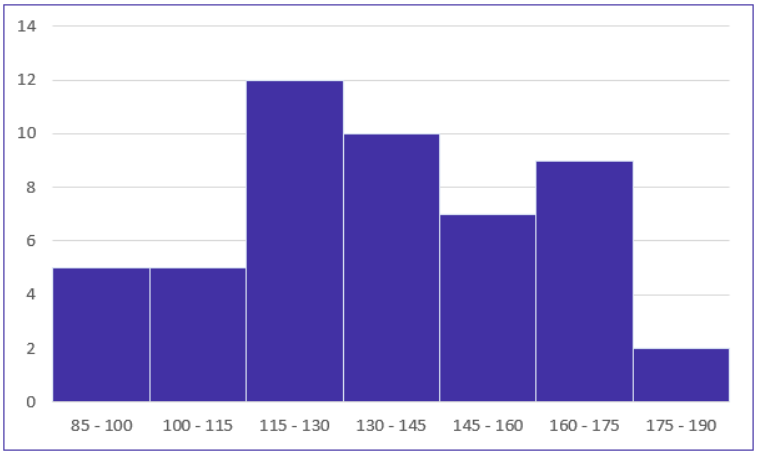
e) Polígono de frequências.

Comentários:

Uma variável quantitativa contínua é aquela que tem características mensuráveis, que assume valores em uma escala contínua (na reta real). Analisando as alternativas temos:

Letra A: em um histograma associamos a cada intervalo de classe uma determinada densidade de frequência. No caso de variáveis quantitativas, é sempre possível estabelecer intervalos de classe, portanto, é sempre factível traçar um histograma. Exemplo de histograma:

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



Na leitura do histograma tem-se que a classe 85 - 100 tem frequência 5. A classe 100 - 115 tem frequência 5. A classe 115 - 130 tem frequência 12. E assim por diante.

Observamos que é possível abranger **todos** os valores reais do intervalo entre 85 a 190. E, por esse motivo, o histograma pode representar uma variável contínua.

Letra B: o diagrama de ramos e folhas impõe que listemos todas as realizações da variável. Desta forma, como a variável assume infinitos valores reais, é impossível listar todos eles, portanto, não seria viável a construção do diagrama de ramos e folhas. No entanto, para representar uma amostra da população, tal diagrama poderá ser utilizado. Exemplo de diagrama de ramos e folhas:

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

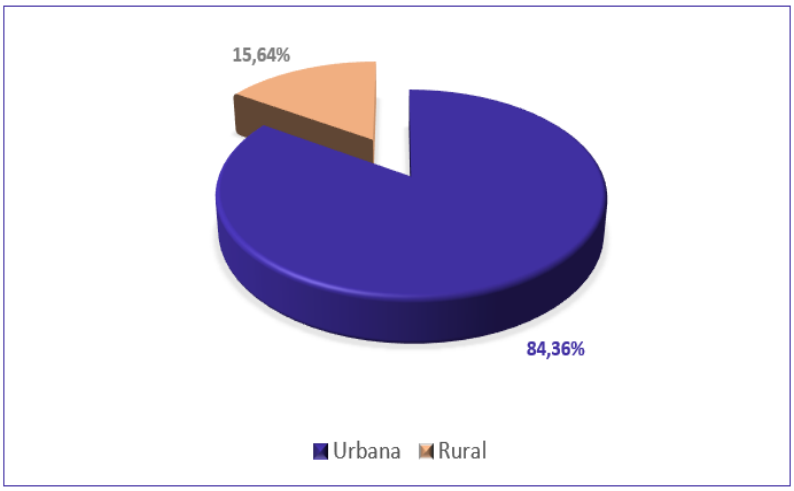
Dados utilizados:

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

Repare que, no lado esquerdo, temos as centenas e as dezenas representando os ramos. Por sua vez, no lado direito, temos as unidades representando as folhas. As folhas, portanto, estão vinculadas aos ramos. Dessa maneira, a chave "**9 |** 6 8 9" significa que, no rol original, havia os números 96, 98 e 99. Podemos perceber que a amostra da variável contínua pôde ser representada. Porém, não conseguiríamos representar todos os possíveis de dada população num intervalo real.

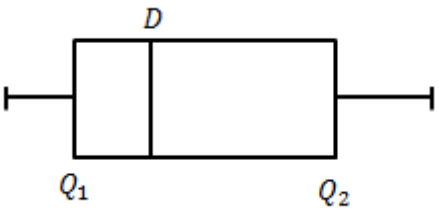
Letra C: o diagrama de setores é usualmente utilizado para representar variáveis qualitativas, motivo pelo qual a banca considerou esta alternativa como o gabarito da questão. Exemplo de diagrama de setores:

População Urbana e Rural do Brasil em 2010

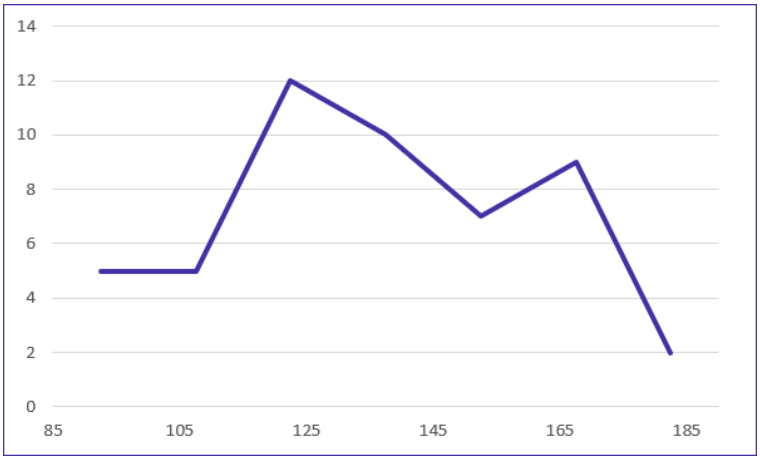


Fonte: Sinopse do Censo Demográfico de 2010 (IBGE)

Letra D: o box-plot impõe a determinação do primeiro quartil, da mediana, do terceiro quartil, e dos valores extremos. Tudo isso pode ser calculado para uma variável quantitativa contínua. Exemplo de box-plot:



Letra E: o polígono de frequência é o gráfico em linhas obtido através da união dos pontos médios das bases superiores dos retângulos de um histograma (que utiliza colunas verticais para mostrar as frequências). Portanto, se o histograma pode ser elaborado para variáveis quantitativas contínuas, então o polígono de frequência também pode. Exemplo de polígono de frequência:



Gabarito: C.

Questões da FUNDATEC

60. (FUNDATEC/BRDE/2017) Um grupo de pessoas foi consultada sobre a marca de veículos de sua preferência. Nessa pesquisa, 100 pessoas escolheram a marca Hyundai de um total de 500 participantes. As que escolheram a marca Ford correspondem a 30%. As pessoas restantes escolheram entre as marcas Chevrolet e Volkswagen. A frequência absoluta das pessoas que escolheram a marca Volkswagen é igual a diferença entre o número de pessoas que escolheram as marcas Ford e Hyundai. A frequência relativa entre as pessoas que escolheram a marca Chevrolet corresponde a:

a) 10%.

b) 20%.

c) 30%.

d) 40%.

e) 50%.

Comentários:

Segundo o enunciado, do total de 500 participantes, 100 escolheram Hyundai.

Além disso, 30% dos 500 participantes escolheram Ford. Logo,

$$0{,}3 \times 500 = 150.$$

Os 150 que escolheram Ford subtraídos dos 100 que escolheram Hyundai é igual ao número de pessoas que escolheram Volkswagen:

$$150 - 50 = 100$$

Portanto, o número de pessoas que escolheram a Chevrolet foi:

$$500 - 100 - 150 - 50 = 200$$

A frequência relativa é obtida dividindo a frequência absoluta (200 pessoas que escolheram Chevrolet) pelo número total de observações (total de 500 pessoas consultadas). Calculando a frequência relativa das pessoas que escolheram Chevrolet, temos:

$$\frac{200}{500} = \frac{40}{100} = 40\%$$

Logo, a frequência relativa de pessoas que escolheram Chevrolet é de 40%.

Gabarito: D.

61. (FUNDATEC/BRDE/2015) Assinale a alternativa que representa a nomenclatura dos três gráficos abaixo, respectivamente.

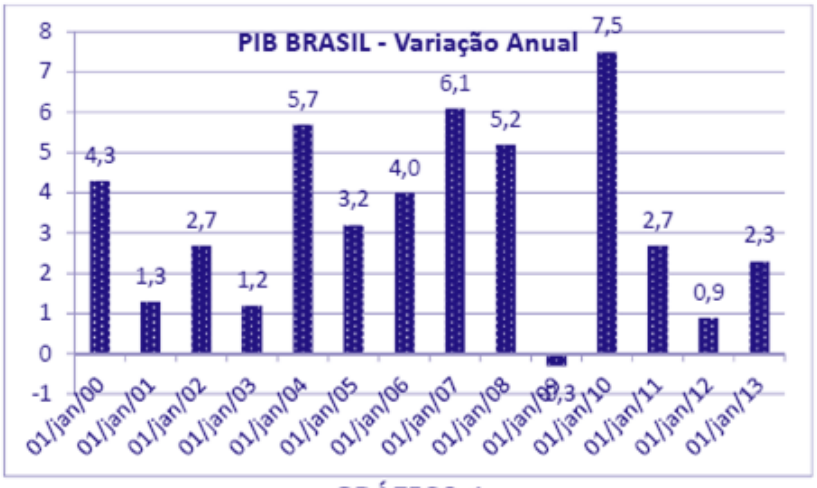


GRÁFICO 1

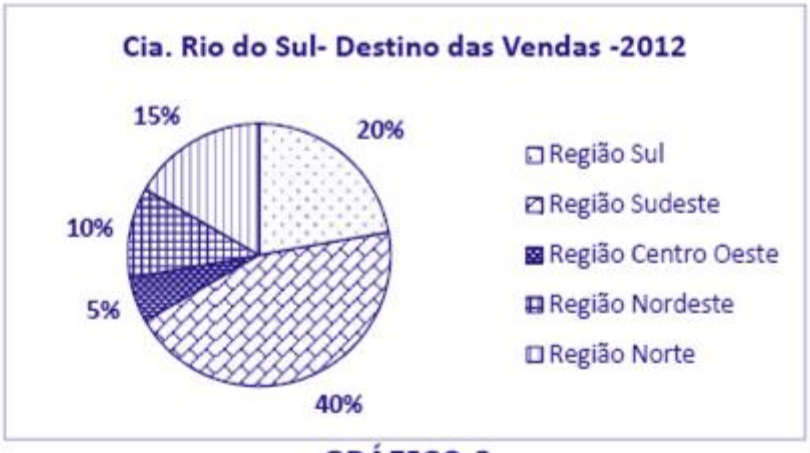


GRÁFICO 2



GRÁFICO 3

a) Gráfico de Setores – Gráfico de Barras – Gráfico de Linha.

b) Gráfico de Pareto – Gráfico de Pizza – Gráfico de Tendência.

c) Gráfico de Barras – Gráfico de Setores – Gráfico de Linha.

d) Gráfico de Linhas – Gráfico de Pizza – Gráfico de Barras.

e) Gráfico de Tendência – Gráfico de Setores – Gráfico de Linha.

Comentários:

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

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

Os **gráficos em linha**, normalmente, são usados para representar dados de séries temporais, com a finalidade de mostrar a variação dos valores de uma variável ao longo do tempo. Esse tipo de gráfico permite-nos

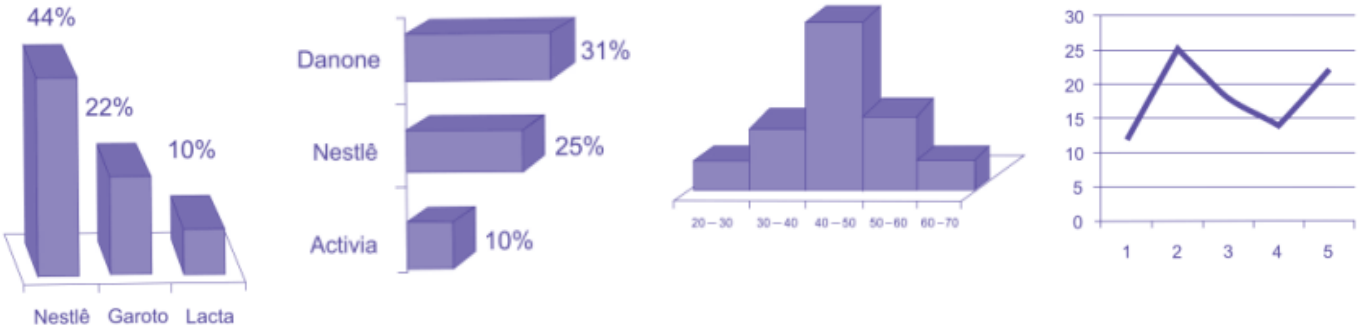
comparar duas variáveis: uma é traçada no eixo x (horizontal) e a outra no eixo y (vertical). O eixo y geralmente indica uma quantidade, enquanto o eixo x representa uma unidade de tempo.

O **gráfico de hastes ou bastões** é frequentemente usado para representar dados não agrupados em classes (normalmente dados discretos), em que não há perda de informação. Em relação a sua construção, basta utilizar as frequências absolutas ou relativas simples de cada elemento.

Apenas com estes conceitos, já somos capazes de julgar a alternativa correta. No caso, respectivamente, gráficos de barras, setores e linhas.

Gabarito: C.

62. (FUNDATEC/SEFAZ-RS/2009) Na revista Seleções de setembro de 2009 foram publicados os resultados da 8ª edição da Pesquisa Marcas de Confiança, também apresentados através de alguns tipos de gráficos, dentre eles, os dois primeiros que estão abaixo, sendo os outros dois fictícios. A sequência correta dos gráficos abaixo é, respectivamente,



a) Polígono, Colunas, histograma, barras.

b) Histograma, colunas, barras, polígono.

c) Colunas, barras, histograma, polígono

d) Histograma, polígono, barras, colunas.

e) Barras, colunas, histograma, polígono

Comentários:

Na primeira figura, temos um gráfico de colunas. Os **gráficos de colunas** são usados para distribuições de dados categóricos ou qualitativos, uma série estatística é representada por um conjunto de retângulos dispostos verticalmente, cada um indicando uma categoria particular, todos com a mesma largura e alturas proporcionais aos respectivos dados.

Na segunda figura, temos um **gráfico de barras**, que também são usados para distribuições de dados categóricos ou qualitativos. A diferença é que as barras são colocadas ao longo do eixo vertical.

A terceira figura corresponde a um **histograma**. Nesse caso, este gráfico é destinado a representar dados agrupados em classe, de modo que haverá uma perda de informação, pois o trabalho realizado nas distribuições de frequências é com classes de elementos, e não com elementos dispostos individualmente.

A quarta figura é o **polígono de frequência** que é o gráfico em linhas obtido através da união dos pontos médios das bases superiores dos retângulos de um Histograma (que utiliza colunas verticais para mostrar as frequências).

Gabarito: C.

Questões da VUNESP

63. (VUNESP/CM Serrana/2019) A tabela apresenta o número de acertos dos 200 candidatos a um cargo, em um concurso interno, composto por uma prova contendo 5 questões de múltipla escolha.

Número de acertos	Número de candidatos
0	8
1	22
2	26
3	32
4	44
5	68

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

a) 68 candidatos acertaram a questão de número 5.

b) 144 candidatos acertaram, no mínimo, 3 questões.

c) 26 candidatos acertaram, no máximo, duas questões.

d) Mais da metade do número total dos candidatos acertaram menos da metade do número total das questões.

e) Menos da metade do número total dos candidatos acertaram mais da metade do número total das questões.

Comentários:

Analisando as alternativas temos:

Letra A: Errado. 68 candidatos acertaram as 5 questões da prova, essa é a informação dada pela tabela, e não somente a questão número 5.

Letra B: Correto. Acertar no mínimo 3 questões significa acertar 3 ou 4 ou 5 questões. Somando o número de candidatos correspondentes a esses acertos, de acordo com a tabela, teremos $32 + 44 + 68 = 144$.

Letra C: Errado. A tabela informa que 26 candidatos acertaram exatamente duas questões. Para considerar quantas pessoas acertaram no máximo duas questões, precisamos incluir as pessoas que acertaram exatamente 2, exatamente 1 e exatamente nenhuma questão, ou seja, 56 candidatos acertaram no máximo 2 questões, $(8 + 22 + 26 = 56)$.

Letra D: Errado. Como visto no item acima, 56 candidatos acertaram 2 questões ou menos, o que corresponde a menos da metade da prova. Esse valor não corresponde a mais da metade do número total de candidatos (200).

Letra E: Errado. De acordo com a letra C, 56 pessoas acertaram menos da metade da prova. Se o total de candidatos corresponde a 200, então $200 - 56 = 144$ candidatos acertaram mais da metade da prova, número superior a metade do número total de candidatos, que é igual a 100. Também podemos recorrer à letra B para confirmarmos o resultado, $32 + 44 + 68 = 144$ candidatos que acertaram mais da metade da prova.

Gabarito: B.

64. (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:

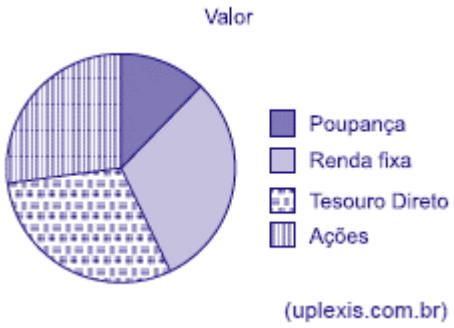
- Letra A: amostra é um subconjunto de uma população. Os dados coletados e/ou selecionados de uma população estatística por um procedimento definido, denotado de amostragem.
- Letra B: unidades são elementos da amostra (unidades amostrais) ou da população (unidade populacional).

- 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).

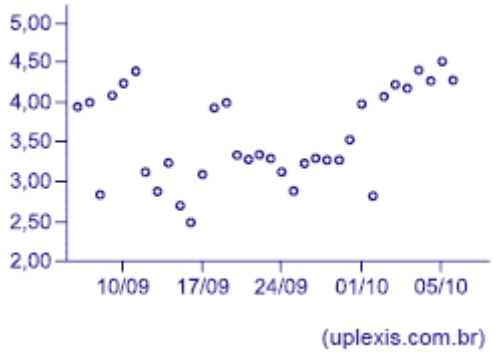
Gabarito: D.

65. (VUNESP/UNICAMP/2019) Assinale dentre os exemplos a seguir, o gráfico de dispersão.

a)



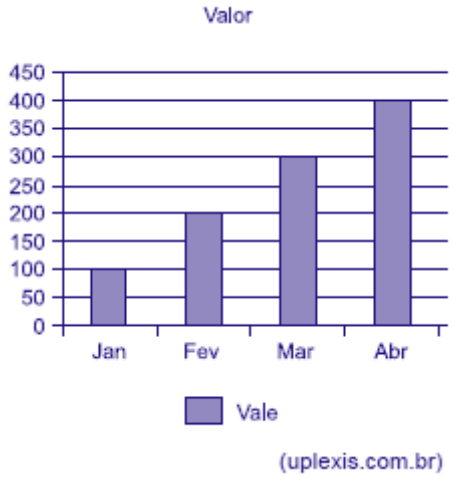
b)



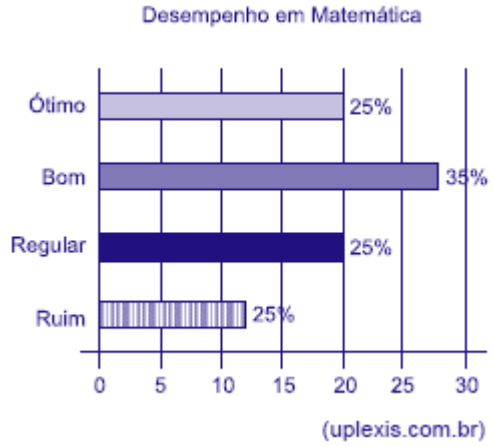
c)



d)



e)



Comentários:

Analisando cada gráfico, temos:

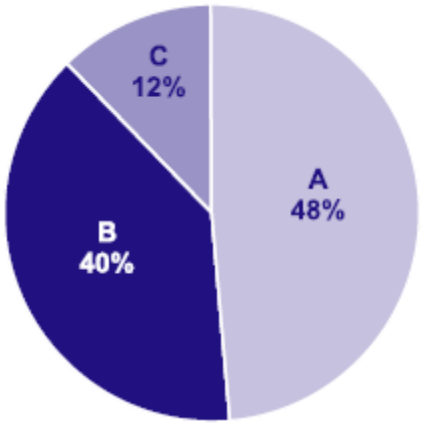
- Letra A: o gráfico de setores é formado por um círculo dividido em setores circulares, cada um representando uma categoria, cujos ângulos centrais são proporcionais às frequências relativas da categoria. Logo, a alternativa está errada.
- Letra B: o gráfico de dispersão é uma representação gráfica que analisa a relação entre duas variáveis quantitativas — uma de causa e outra de efeito, chamadas de variáveis independente e dependente, respectivamente. Esse tipo de diagrama traz números simultâneos das duas variáveis, deixando visível se o que acontece em uma variável interfere na outra.
- Letra C: o gráfico em linhas é o gráfico em que os pontos são geralmente usados para controlar alterações ao longo do tempo e para facilitar a identificação de tendências ou de anomalias. Alternativa Errada.
- Letra D: os gráficos em colunas assim como os gráficos de barras são usados para distribuições de dados categóricos ou qualitativos. A diferença básica é que, no primeiro, uma série estatística é

representada por um conjunto de retângulos dispostos verticalmente, cada um indicando uma categoria particular, todos com a mesma largura e alturas proporcionais aos respectivos. Alternativa Errada.

- Letra E: os gráficos em barras normalmente são usados para representar distribuições de dados categóricos ou qualitativos. Uma série estatística é representada por um conjunto de retângulos dispostos horizontalmente, cada um indicando uma categoria particular, os quais possuem a mesma altura e comprimentos proporcionais aos respectivos dados. Alternativa Errada.

Gabarito: B.

66. (VUNESP/Pref. Campinas/2019) Uma empresa atua em três segmentos de mercado, A, B e C. O gráfico de setores mostra a distribuição percentual, por segmento, da receita total obtida por essa empresa em 2018.



Sabendo-se que a receita obtida no segmento A superou a receita obtida no segmento B em R$ 64 milhões, é correto afirmar que a receita obtida no segmento C foi igual a

a) R$ 98 milhões.

b) R$ 96 milhões.

c) R$ 94 milhões.

d) R$ 88 milhões.

e) R$ 86 milhões.

Comentários:

Consideremos T o total de receitas obtidas pela empresa em 2018. Pelas informações apresentadas no gráfico de pizza, temos:

$$A = 0{,}48 \times T$$
$$B = 0{,}4 \times T$$
$$C = 0{,}12 \times T$$

Sabendo que a receita obtida no segmento A superou a receita obtida no segmento B em R$ 64 milhões, então:

$$A = B + 64$$
$$0{,}48 \times T = 0{,}4 \times T + 64$$
$$0{,}8 \times T = 64$$
$$T = 800\ milhões$$

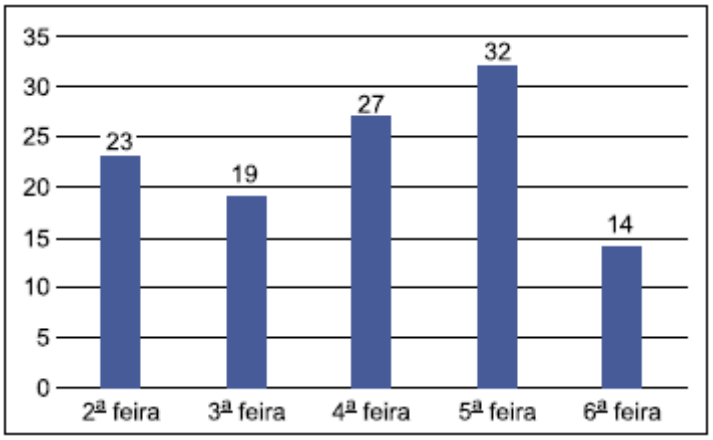
Logo,

$$C = 0{,}12 \times T$$
$$C = 12 \times 800$$
$$C = 96\ milhões$$

Gabarito: B.

67. (VUNESP/Pref. Campinas/2019) Considere o gráfico a seguir.

Atendimento ao público: número de pessoas por dia



Se, ao invés da situação exposta no gráfico, o número de pessoas atendidas por dia tivesse sido sempre igual e mantido o número total de atendimentos, o número de pessoas que teriam sido atendidas na 5ª feira teria sido inferior ao que consta no gráfico em

a) 6 unidades.

b) 8 unidades.

c) 7 unidades.

d) 10 unidades.

e) 9 unidades.

Comentários:

Para saber qual seria o número de pessoas atendidas por dia se fossem atendidas a mesma quantidade todos os dias, basta calcularmos a média dos atendimentos da semana. Assim, temos:

$$23 + 19 + 27 + 32 + 14 = 115$$
$$\frac{115}{5} = 23$$

Desse modo, como anteriormente haviam sido atendidas 32 pessoas, então o número de pessoas atendidas na 5ª feira foi inferior ao que consta no gráfico em:

$$32 - 23 = 9\ pessoas.$$

Gabarito: E.

68. (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.

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

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

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

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

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

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:

i) 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.

ii) 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:

i) variáveis nominais: não existe uma ordenação entre as categorias. Exemplos: sexo, estado civil, cor dos olhos.

ii) 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).

Gabarito: D.

69. (VUNESP/IPSM-SJC/2018) A tabela apresenta o número de acertos de 100 candidatos a uma vaga de emprego, em uma avaliação contendo 5 questões de múltipla escolha.

Número de acertos	Número de candidatos
0	4
1	11
2	13
3	16
4	22
5	34

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

a) 22 candidatos acertaram a questão de número 4.

b) Menos da metade dos candidatos acertaram mais da metade das questões.

c) 28 candidatos acertaram, no máximo, duas questões.

d) Mais da metade dos candidatos acertaram menos da metade das questões.

e) 15 candidatos acertaram, pelo menos, uma questão.

Comentários:

Analisando as alternativas, temos:

- **Letra A**: observando a última linha da tabela, concluímos que 34 candidatos que gabaritaram a prova. Logo, ao menos esses 34 acertaram a quarta questão da prova. Portanto, se há 34 que a acertaram, é possível formar um subgrupo contendo 22 candidatos. Porém, o examinador deu a entender que devemos considerar que a palavra "exatamente" está implícita no enunciado, tornando a alternativa incorreta.

- **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 $(16 + 22 + 34)$, concluímos que 72 candidatos acertaram mais de metade das questões. Isso corresponde a mais de 50% do total de alunos. Logo, alternativa incorreta.

- **Letra C**: De fato, 4 candidatos erraram todas; 11 candidatos acertaram (exatamente) uma questão e 13 candidatos acertaram (exatamente) duas questões, formando um total de $4 + 11 + 13 = 28$. 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 $(4 + 11 + 13)$, concluímos que 28 candidatos acertaram mais de metade das questões. Isso corresponde a menos de 50% do total de alunos. Assim, a alternativa está incorreta.

- **Letra E**: aqui temos o mesmo caso da letra A. Precisamos supor implícita a palavra "exatamente". **Exatamente** 15 candidatos acertaram, pelo menos, uma questão. Dos 100 candidatos, só 4 erraram todas as questões da prova. Os outros $100 - 4 = 96$ acertaram pelo menos uma questão. Dessa forma, a alternativa está incorreta.

Gabarito: C.

70. (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 absoluta simples do primeiro elemento é igual a 20%.

Gabarito: B.

71. (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 quinto elemento é

a) 12%.

b) 84%.

c) 5.

d) 6.

e) 42.

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 simples do quinto elemento é igual a 12%.

Gabarito: A.

72. (VUNESP/TJM-SP/2017) A tabela apresenta o número de acertos dos 600 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	204
4	132
3	96
2	78
1	66
0	24

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 168 candidatos acertaram, no mínimo, 2 questões.

d) 264 candidatos acertaram, no máximo, 3 questões.

e) 132 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 $78 + 66 + 24 = 168$. Portanto, temos que 168 pessoas acertaram menos da metade da prova. Logo, menos da metade dos candidatos (300) acertaram menos da metade da prova. Dessa forma, alternativa incorreta.

- **Letra B**: já temos a informação na letra A de que 168 pessoas acertaram menos da metade da prova. Logo, a quantidade de pessoas que acertaram mais da metade da prova é: $600 - 168 = 432$. Portanto, temos que 432 pessoas acertaram mais da metade da prova. Logo, mais da metade das pessoas (300) acertou mais de 50% da prova. Logo, alternativa incorreta.

- **Letra C**: já vimos na letra A que 168 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: $204 + 132 + 96 + 78 = 432 + 78 = 510$. Assim, podemos afirmar que a alternativa está incorreta.

- **Letra D:** 168 candidatos que acertaram até duas questões (vide letra A). Somando com os 96 que acertaram exatamente 3 questões, temos $168 + 96 = 264$. Portanto, alternativa correta.

- **Letra E:** sabemos que 132 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.

73. (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.

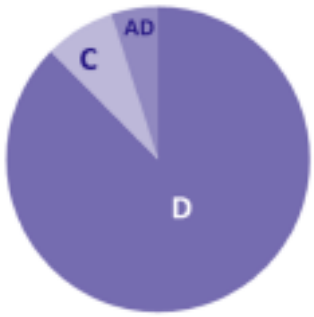
Portanto, a alternativa correta é a letra E.

Gabarito: E.

LISTA DE QUESTÕES COMPLEMENTARES

Questões da FCC

41. (FCC/Pref. Macapá/2018) Para fazer um gráfico de setores que representasse o número de alunos canhotos, destros e ambidestros na sala de aula, Renato coletou os dados indicados na tabela abaixo.

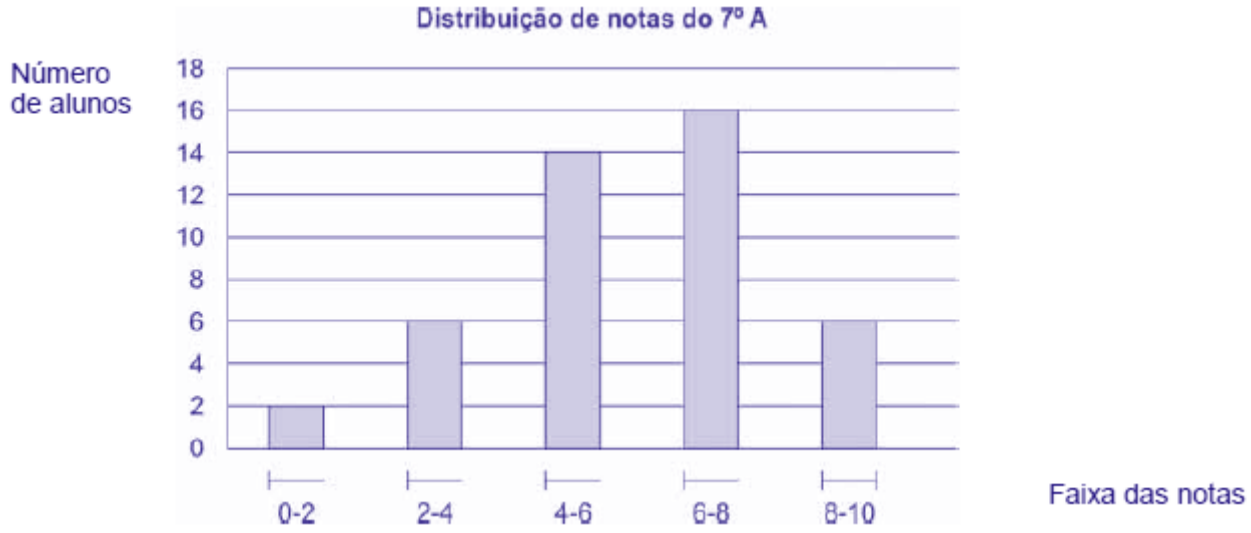


	Nº DE ALUNOS DA SALA DE AULA
Destros (D)	35
Canhotos (C)	3
Ambidestros (AD)	2

Sabendo-se que um círculo pode ser dividido em 360°, quantos graus tem o setor circular correspondente aos canhotos no gráfico correto feito por Renato?

a) 18°.
b) 25°.
c) 7,5°.
d) 24°.
e) 27°.

42.(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%.

43. (FCC/TRT 14ª Região/2018) De um histograma e uma tabela de frequências absolutas, elaborados para analisar a distribuição dos salários dos empregados em uma empresa, obtém-se a informação que 24 empregados ganham salários com valores pertencentes ao intervalo (2.000; 4.000], em reais, que apresenta uma densidade de frequência de $0,75 \times 10^{-4}(R\$)^{-1}$

Densidade de frequência de um intervalo é o resultado da divisão da respectiva frequência relativa pela amplitude deste intervalo. Em um intervalo do histograma que está sendo analisado, com uma amplitude de R$ 3.000,00 e uma densidade de frequência de $1 \times 10^{-4}(R\$)^{-1}$, tem-se que o correspondente número de empregados é igual a

a) 40.
b) 36.
c) 30.
d) 48.
e) 42.

44. (FCC/ARTESP/2017) Foi solicitado para uma empresa de transportes que fizesse um levantamento da idade da frota dos seus caminhões que operavam em um trecho de rodovia com tráfego intenso. O gerente da empresa entregou a seguinte tabela:

Caminhão	A	B	C	D	E	F	G	H	I	J	K	L	M

Idade (anos)	15	20	5	2	3	3	17	23	16	8	6	9	14
Caminhão	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	
Idade (anos)	10	10	6	8	2	4	18	5	9	11	26	14	

De posse deste levantamento, a analista de operações da autarquia solicitante organizou uma distribuição de frequência para organizar melhor os dados para serem analisados. Sendo assim, em um primeiro momento, fez-se necessário encontrar o número de intervalos (K) e a classe (C), que são expressos, respectivamente, por

a) 5 e 4,80.
b) 5 e 5,20.
c) 6 e 4,80.
d) 6 e 5,20.
e) 4 e 5.

45. (FCC/TRT 20ª Região/2016) Um gráfico corresponde a um histograma apresentando a distribuição dos salários dos funcionários lotados em um determinado órgão público. No eixo das abscissas constam os intervalos de classe (fechados à esquerda e abertos à direita) dos salários em R$ e no eixo das ordenadas as respectivas densidades de frequências em (R$) $^{-1}$. Densidade de frequência de um intervalo é definida como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Se 135 funcionários ganham salários com valores pertencentes ao intervalo [3.000, 6.000) com uma densidade de frequência de 1×10^{-4} (R$) $^{-1}$, então o número de funcionários que ganham salários com valores pertencentes ao intervalo [6.000, 8.000) com uma densidade de frequência de 2×10^{-4} (R$) $^{-1}$ é igual a:

a) 300
b) 180
c) 270
d) 150
e) 90

46. (FCC/TRT 3ª Região/2015) Em um histograma representando os preços unitários de microcomputadores em estoque, observa-se que no eixo das abscissas constam os intervalos de classe em R$ e no eixo das ordenadas as respectivas densidades de frequências em $(R\$)^{-1}$. Densidade de frequência de um intervalo de classe é o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um determinado intervalo de classe com amplitude igual a R$ 2.500,00 apresenta uma densidade de frequência, em $(R\$)^{-1}$, igual a $12{,}8 \times 10^{-5}$. Se o número de microcomputadores deste intervalo é igual a 48, então o número total de microcomputadores em estoque é igual a

a) 150.
b) 120.
c) 240.
d) 160.
e) 96.

47. (FCC/TRT 16ª Região/2014) A distribuição das medidas dos comprimentos, em cm, de uma determinada peça em estoque de uma fábrica está representada em um histograma com todos os intervalos de classe fechados à esquerda e abertos à direita. No eixo horizontal constam os intervalos de classe e no eixo vertical as respectivas densidades de frequências, em cm^{-1}. Define-se densidade de frequência de um intervalo de classe como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude deste intervalo. Verifica-se com relação ao histograma, que o intervalo de classe [2 , 6), em cm, apresenta uma densidade de frequência igual a 0,028 cm^{-1}. Dado que o número de peças em estoque com medidas iguais ou superiores a 2 cm e inferiores a 6 cm é igual a 84, obtém-se que o número total destas peças em estoque é

a) 750.

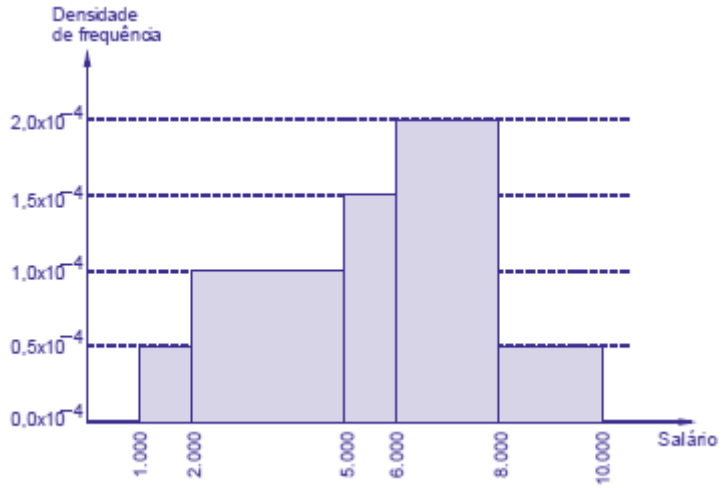
b) 400.

c) 500.

d) 250.

e) 1.000.

48. (FCC/TRT 13ª Região/2014) O histograma, abaixo, refere-se à distribuição dos salários dos funcionários lotados em um setor de um órgão público. No eixo das abscissas constam os intervalos de classe em R$ (todos fechados à esquerda e abertos à direita) e no eixo das ordenadas as respectivas densidades de frequências em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo.



Se o número de funcionários que tem um salário inferior a R$ 5.000,00 é igual a 56, então verifica-se que o número de funcionários que tem um salário igual ou superior a R$ 2.000,00 e inferior a R$ 8.000,00 é igual a

a) 104.

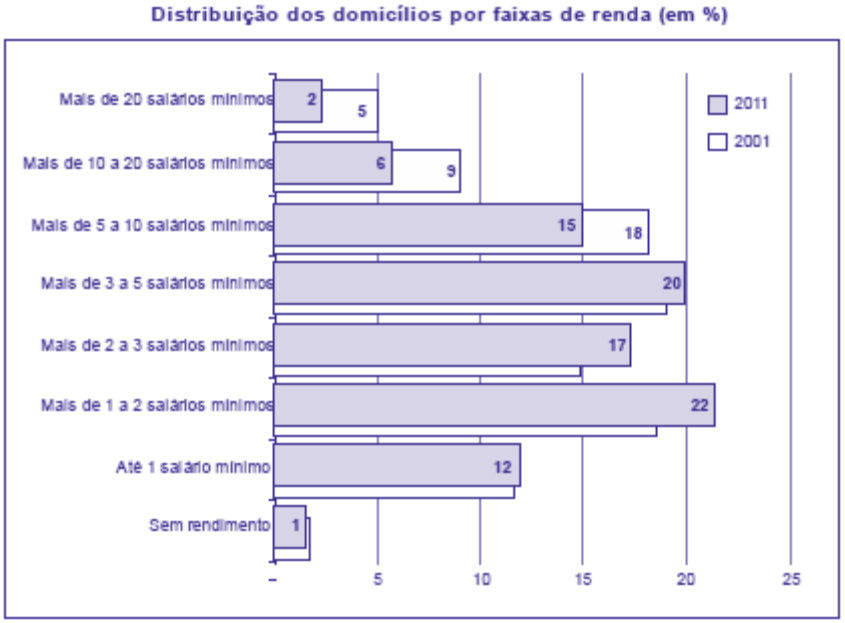
b) 152.

c) 136.

d) 144.

e) 120.

49. (FCC/ALE-PE/2014) Analise o gráfico e as afirmativas abaixo.



(http://blogs.estadao.com.br/vox-publica/tag/psdb/. Acessado em: 20/02/2013)

I. O crescimento da faixa de renda de "até 1 salário mínimo" pode ser reflexo dos programas de transferência de renda governamentais, os quais garantem um rendimento mínimo às pessoas carentes.

II. O crescimento da faixa de renda de "até 1 salário mínimo" é fruto do emprego formal gerado no país dentro do período da pesquisa, pois, os beneficiários dos programas de transferência de renda governamentais entram na categoria sem rendimento.

III. No gráfico, percebe-se uma diminuição da concentração de renda no país. Isso fica evidente quando analisamos, por exemplo, a faixa de renda "mais de 5 a 10 salários mínimos".

IV. No gráfico, a faixa "mais de 1 a 2 salários mínimos" é a que apresenta a maior diferença em pontos percentuais entre os dados de 2001 e 2011.

Está correto o que se afirma APENAS em

a) I e II.

b) II e III.

c) III e IV.

d) I e III.

e) II e IV.

50. (FCC/TRT 6ª Região/2012) A distribuição dos 500 preços unitários de um equipamento é representada por um histograma em que no eixo das abscissas constam os intervalos de classe e no eixo das ordenadas estão assinaladas as respectivas densidades de frequências, em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo de classe como sendo o resultado da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um intervalo de classe no histograma apresenta uma amplitude de R$ 2,50 com uma densidade de frequência igual a 0,096. A quantidade de preços unitários referente a este intervalo é

a) 96.
b) 120.
c) 144.
d) 150.
e) 192.

51. (FCC/TRT 6ª Região/2012) A função de distribuição empírica abaixo, $F_{200}(x)$, refere-se a uma pesquisa realizada em 200 residências, escolhidas aleatoriamente, em que x é o número verificado de pessoas que trabalham em cada residência.

$$F_{200}(x) = \begin{cases} 0,00\ se\ x < 0 \\ 0,10\ se\ 0 \leq x < 1 \\ 0,40\ se\ 1 \leq x < 2 \\ 0,65\ se\ 2 \leq x < 3 \\ 0,85\ se\ 3 \leq x < 4 \\ 0,95\ se\ 4 \leq x < 5 \\ 1,00\ se\ x \geq 5 \end{cases}$$

O número de residências desta pesquisa em que se verificou possuir pelo menos uma pessoa que trabalha e menos que 4 é

a) 150.
b) 160.
c) 170.
d) 180.
e) 190.

52. (FCC/TRT 1ª Região/2011) Um histograma representa a distribuição dos preços unitários de venda de um determinado equipamento no mercado. No eixo das ordenadas estão assinaladas as respectivas densidades de frequência para cada intervalo em $(R\$)^{-1}$. Define-se densidade de frequência de um intervalo de classe como sendo o quociente da divisão da respectiva frequência relativa pela correspondente amplitude do intervalo. Um intervalo de classe do histograma corresponde aos preços unitários maiores ou iguais a R$ 32,00 e inferiores a R$ 44,50 com uma densidade de frequência igual a $1,6 \times 10^{-2}$ $(R\$)^{-1}$. Se todos os intervalos de classe do histograma têm a mesma frequência relativa, então um intervalo de classe com densidade de frequência igual a $5,0 \times 10^{-3}$ $(R\$)^{-1}$ apresenta uma amplitude de

a) R$ 64,00.

b) R$ 48,00.

c) R$ 40,00.

d) R$ 32,00.

e) R$ 24,00.

53. (FCC/DPE-SP/2010) Sobre estatística aplicada, é correto o que se afirma em:

a) Parâmetros são medidas características de grupos, determinadas por meio de uma amostra aleatória.

b) A estatística descritiva é a técnica pela são coletados dados de uma amostra, a partir do que são tomadas decisões sobre uma determinada população.

c) A caracterização de uma população se dá por meio da observação de todos os seus componentes que a integram.

d) A estatística inferencial compreende um conjunto de técnicas destinadas à síntese de dados numéricos.

e) Censo é o processo utilizado para se medir as características de todos os membros de uma dada população.

Questões da FGV

54. (FGV/COMPESA/2018) A COMPESA, em uma pesquisa de satisfação dos usuários, preparou um formulário para traçar os perfis de seus clientes e o grau de satisfação com os serviços da empresa,

Em um formulário, ela solicitou os dados a seguir.

I. Idade.

II. Grau de escolaridade.

III. Faixa de renda familiar.

IV. Nota dada ao serviço.

Assinale a opção que contempla apenas variáveis categóricas.

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

55. (FGV/IBGE/2017) Uma vez concluída a etapa de críticas de dados, relativa a um conjunto de registros obtido através de uma pesquisa de campo, inicia-se o trabalho de tabulação e elaboração de gráficos. Durante as análises, algumas variáveis surgem com destaque.

IV. Número de indivíduos por faixa etária;

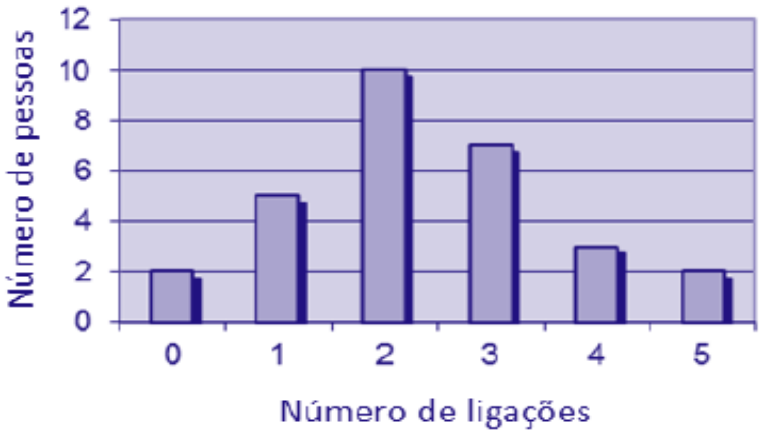
V. Percentuais do nível de escolaridade; e

VI. Pares de valores de consumo e renda.

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

a) Colunas, linha e setores;
b) Setores, cartograma e barras;
c) Barras, setores e linha;
d) Linha, colunas e cartograma;
e) Colunas, setores e linha.

56. (FGV/CGE-MA/2014) No setor A de uma empresa foi feita uma auditoria para descobrir quantas vezes cada pessoa fazia ligações pessoais do seu celular no período de trabalho de 14 às 17 horas de um único dia. O resultado está no gráfico a seguir.



O número de pessoas que trabalham no setor A dessa empresa é

a) 15.
b) 22.
c) 27.
d) 29.
e) 42.

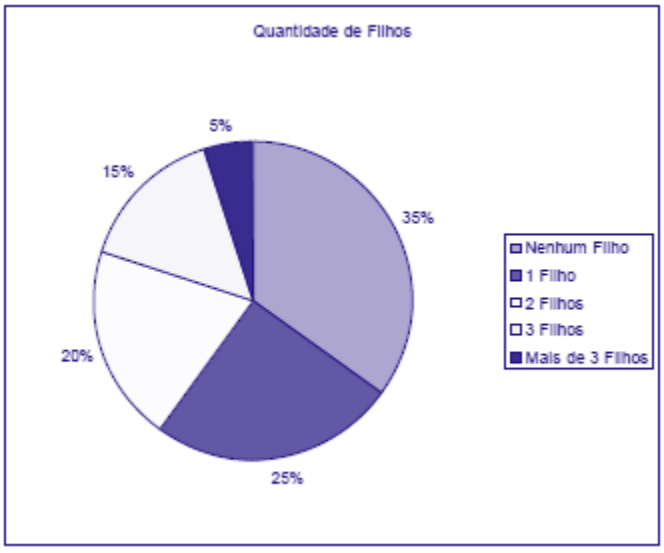
57. (FGV/AL-BA/2014) Observe a tabela de frequências a seguir, que se refere aos saldos em conta, num determinado dia, de duzentas contas-correntes:

Saldos em conta (R$)	Frequência
Até 100,00	8
De mais de 100,00 a 300,00	28
De mais de 300,00 a 500,00	46
De mais de 500,00 a 700,00	54
De mais de 700,00 a 900,00	44
De mais de 900,00 a 1.100,00	13
De mais de 1.100,00 a 1.300,00	6
Acima de 1.300,00	1

A frequência relativa acumulada de saldos em R$ 900,00 é igual a

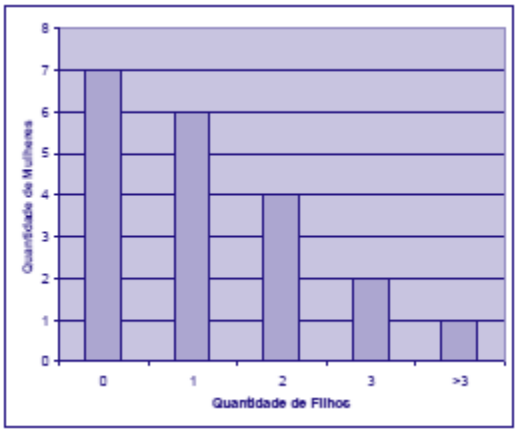
a) 22%.
b) 36%.
c) 54%.
d) 90%.
e) 97%.

58. (FGV/CODEBA/2010) O gráfico de pizza apresentado abaixo fornece os resultados de uma pesquisa feita com 20 mulheres acerca da quantidade de filhos que cada uma possui.

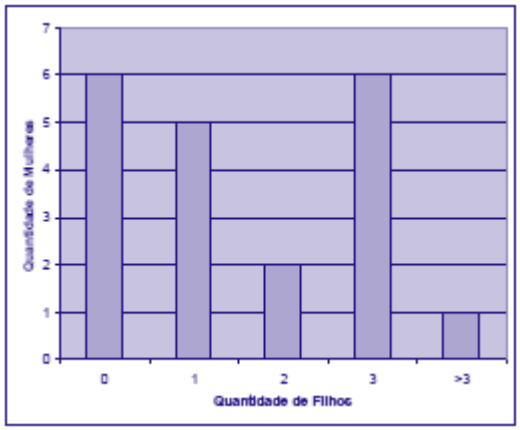


Assinale o gráfico de colunas que melhor representa os dados apresentados acima.

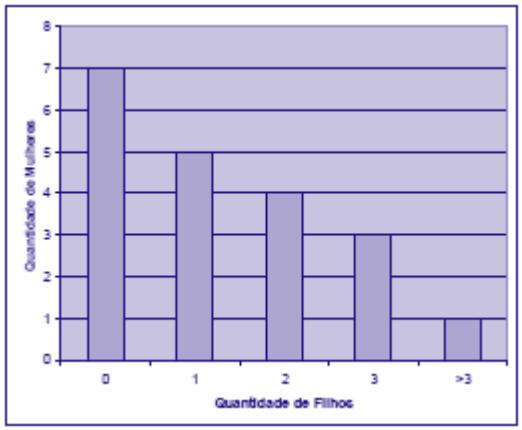
a)


Quantidade de Mulheres
0
1
2
3
4
5
6
7
8
0
1
2
3
>3
Quantidade de Filhos

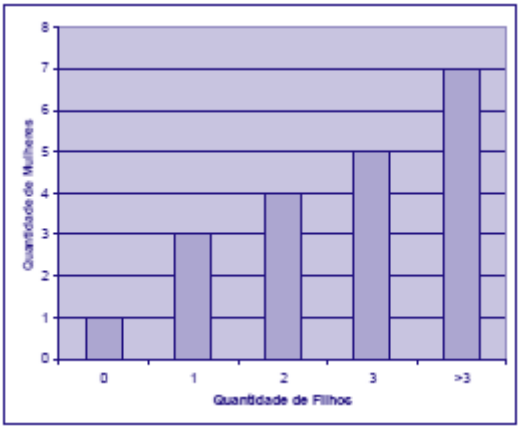
b)


Quantidade de Mulheres
0
1
2
3
4
5
6
7
0
1
2
3
>3
Quantidade de Filhos

c)


Quantidade de Mulheres
0
1
2
3
4
5
6
7
8
0
1
2
3
>3
Quantidade de Filhos

d)


Quantidade de Mulheres
0
1
2
3
4
5
6
7
8
0
1
2
3
>3
Quantidade de Filhos

e)



59. (FGV/Senado/2008) Entre os tipos de representações gráficas a seguir, o único que não é definido para uma variável quantitativa contínua é:

a) histograma.
b) diagrama ramo-e-folhas.
c) diagrama de setores.
d) diagrama "Box-Plot".
e) Polígono de frequências.

Questões da FUNDATEC

60. (FUNDATEC/BRDE/2017) Um grupo de pessoas foi consultada sobre a marca de veículos de sua preferência. Nessa pesquisa, 100 pessoas escolheram a marca Hyundai de um total de 500 participantes. As que escolheram a marca Ford correspondem a 30%. As pessoas restantes escolheram entre as marcas Chevrolet e Volkswagen. A frequência absoluta das pessoas que escolheram a marca Volkswagen é igual a diferença entre o número de pessoas que escolheram as marcas Ford e Hyundai. A frequência relativa entre as pessoas que escolheram a marca Chevrolet corresponde a:

a) 10%.

b) 20%.

c) 30%.

d) 40%.

e) 50%.

61. (FUNDATEC/BRDE/2015) Assinale a alternativa que representa a nomenclatura dos três gráficos abaixo, respectivamente.

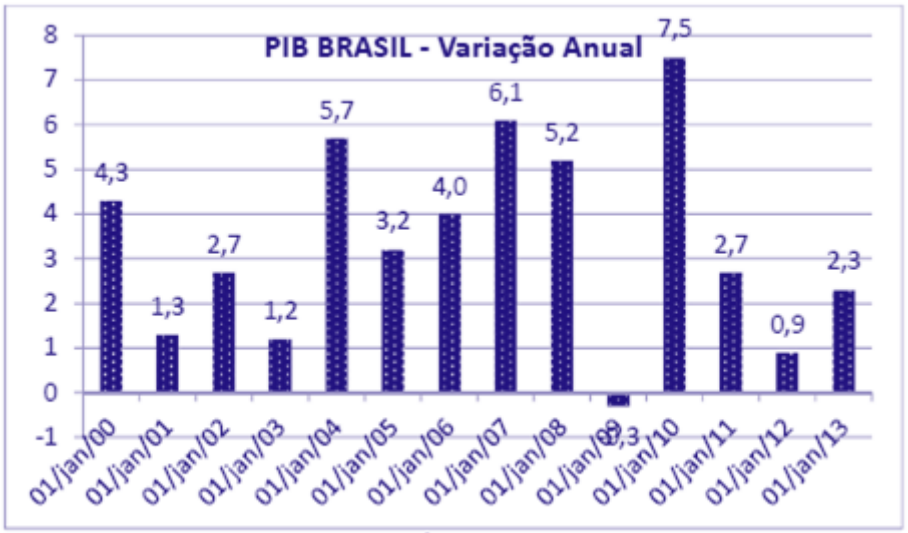


GRÁFICO 1

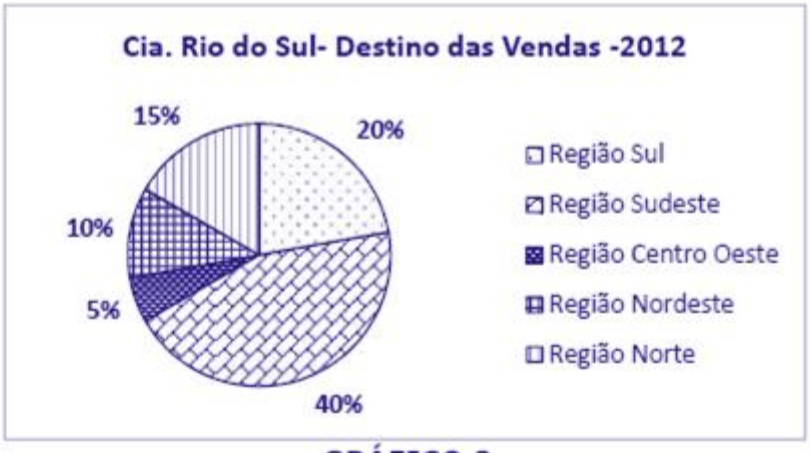


GRÁFICO 2



GRÁFICO 3

a) Gráfico de Setores – Gráfico de Barras – Gráfico de Linha.

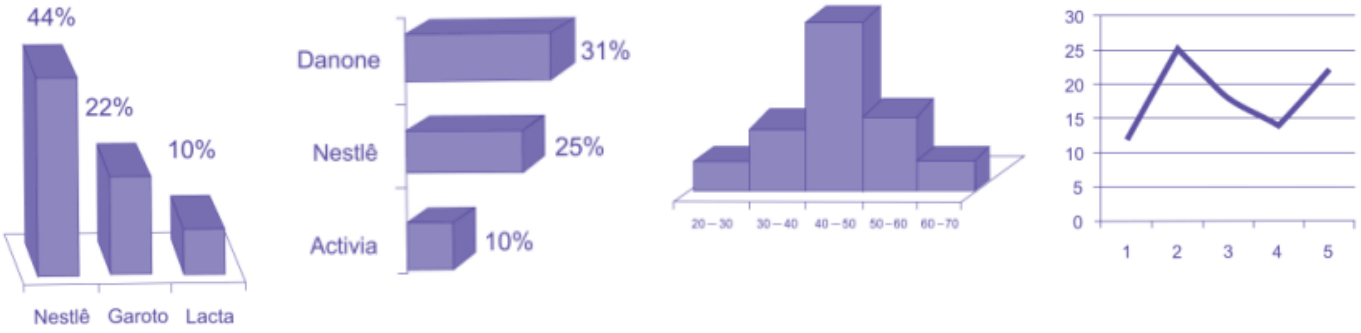
b) Gráfico de Pareto – Gráfico de Pizza – Gráfico de Tendência.

c) Gráfico de Barras – Gráfico de Setores – Gráfico de Linha.

d) Gráfico de Linhas – Gráfico de Pizza – Gráfico de Barras.

e) Gráfico de Tendência – Gráfico de Setores – Gráfico de Linha.

62. (FUNDATEC/SEFAZ-RS/2009) Na revista Seleções de setembro de 2009 foram publicados os resultados da 8ª edição da Pesquisa Marcas de Confiança, também apresentados através de alguns tipos de gráficos, dentre eles, os dois primeiros que estão abaixo, sendo os outros dois fictícios. A sequência correta dos gráficos abaixo é, respectivamente,



a) Polígono, Colunas, histograma, barras.

b) Histograma, colunas, barras, polígono.

c) Colunas, barras, histograma, polígono

d) Histograma, polígono, barras, colunas.

e) Barras, colunas, histograma, polígono

Questões da VUNESP

63. (VUNESP/CM Serrana/2019) A tabela apresenta o número de acertos dos 200 candidatos a um cargo, em um concurso interno, composto por uma prova contendo 5 questões de múltipla escolha.

Número de acertos	Número de candidatos
0	8
1	22
2	26
3	32
4	44
5	68

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

a) 68 candidatos acertaram a questão de número 5.

b) 144 candidatos acertaram, no mínimo, 3 questões.

c) 26 candidatos acertaram, no máximo, duas questões.

d) Mais da metade do número total dos candidatos acertaram menos da metade do número total das questões.

e) Menos da metade do número total dos candidatos acertaram mais da metade do número total das questões.

64. (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.

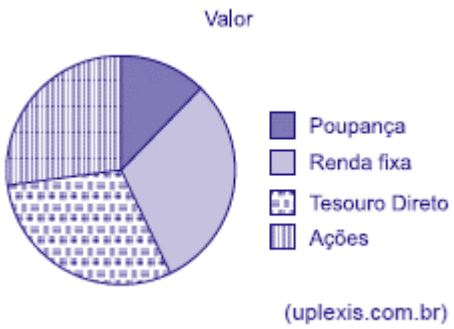
c) Censo.

d) População.

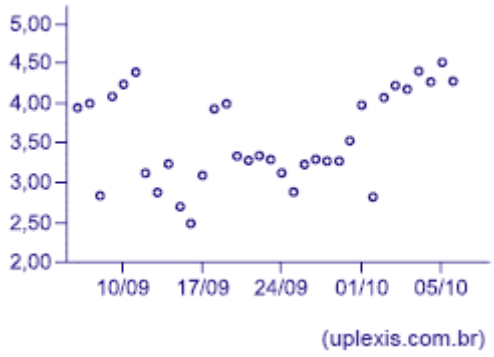
e) Variáveis.

65. (VUNESP/UNICAMP/2019) Assinale dentre os exemplos a seguir, o gráfico de dispersão.

a)



b)

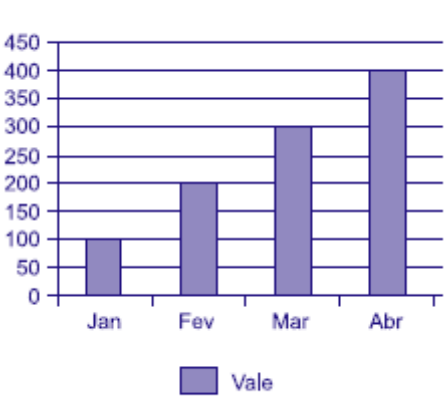


c)



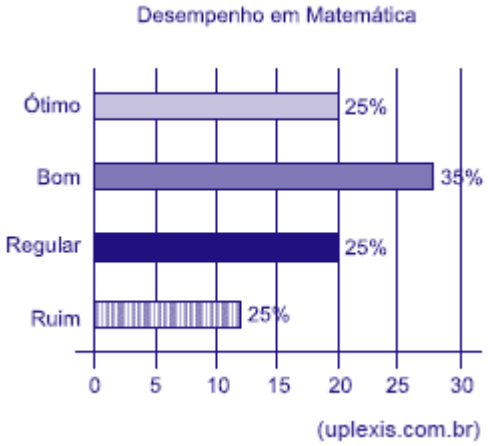
(uplexis.com.br)

d)



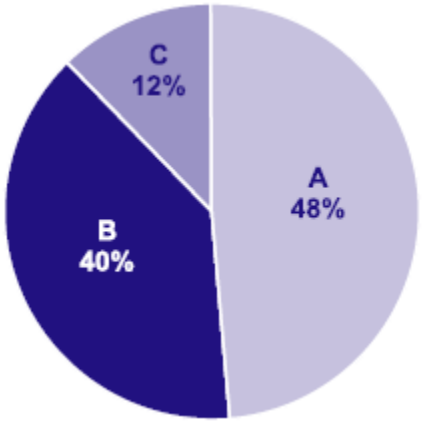
(uplexis.com.br)

e)



(uplexis.com.br)

66. (VUNESP/Pref. Campinas/2019) Uma empresa atua em três segmentos de mercado, A, B e C. O gráfico de setores mostra a distribuição percentual, por segmento, da receita total obtida por essa empresa em 2018.



Sabendo-se que a receita obtida no segmento A superou a receita obtida no segmento B em R$ 64 milhões, é correto afirmar que a receita obtida no segmento C foi igual a

a) R$ 98 milhões.

b) R$ 96 milhões.

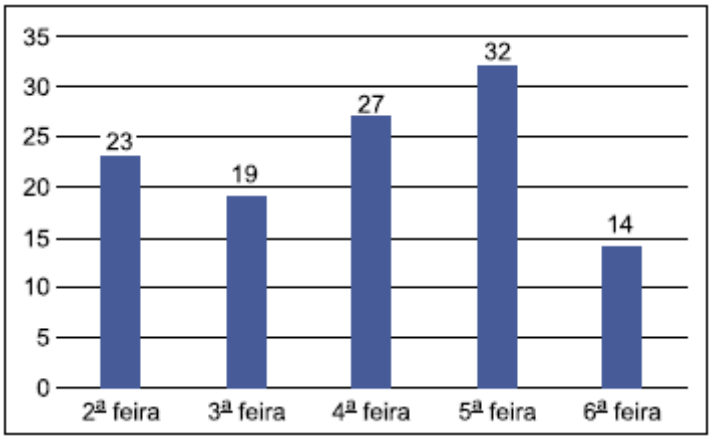
c) R$ 94 milhões.

d) R$ 88 milhões.

e) R$ 86 milhões.

67. (VUNESP/Pref. Campinas/2019) Considere o gráfico a seguir.

Atendimento ao público: número de pessoas por dia



Se, ao invés da situação exposta no gráfico, o número de pessoas atendidas por dia tivesse sido sempre igual e mantido o número total de atendimentos, o número de pessoas que teriam sido atendidas na 5ª feira teria sido inferior ao que consta no gráfico em

a) 6 unidades.
b) 8 unidades.
c) 7 unidades.

d) 10 unidades.
e) 9 unidades.

68. (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.

a) Quantitativas referem-se às variáveis ordinal ou discreta; as qualitativas referem-se às variáveis nominal ou contínua.
b) Quantitativas referem-se às variáveis ordinal ou contínua; as qualitativas referem-se às variáveis nominal ou discreta.
c) Quantitativas referem-se às variáveis nominal ou contínua; as qualitativas referem-se às variáveis discreta ou ordinal.
d) Quantitativas referem-se às variáveis contínua ou discreta; as qualitativas referem-se às variáveis nominal ou ordinal.
e) Quantitativas referem-se às variáveis nominal ou ordinal; as qualitativas referem-se às variáveis contínua ou discreta.

69. (VUNESP/IPSM-SJC/2018) A tabela apresenta o número de acertos de 100 candidatos a uma vaga de emprego, em uma avaliação contendo 5 questões de múltipla escolha.

Número de acertos	Número de candidatos
0	4
1	11
2	13
3	16
4	22
5	34

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

a) 22 candidatos acertaram a questão de número 4.
b) Menos da metade dos candidatos acertaram mais da metade das questões.
c) 28 candidatos acertaram, no máximo, duas questões.
d) Mais da metade dos candidatos acertaram menos da metade das questões.
e) 15 candidatos acertaram, pelo menos, uma questão.

70. (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.

71. (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 quinto elemento é

a) 12%.
b) 84%.
c) 5.
d) 6.
e) 42.

72. (VUNESP/TJM-SP/2017) A tabela apresenta o número de acertos dos 600 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	204
4	132
3	96
2	78
1	66
0	24

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 168 candidatos acertaram, no mínimo, 2 questões.

d) 264 candidatos acertaram, no máximo, 3 questões.

e) 132 candidatos acertaram a questão de número 4.

73. (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.

GABARITO

1. ERRADO
2. CERTO
3. CERTO
4. CERTO
5. CERTO
6. ERRADO
7. CERTO
8. CERTO
9. ERRADO
10. ERRADO
11. ERRADO
12. ERRADO
13. ERRADO
14. CERTO
15. CERTO
16. ERRADO
17. ERRADO
18. ERRADO
19. ERRADO
20. ERRADO
21. ERRADO
22. ERRADO
23. ERRADO
24. ERRADO
25. CERTO
26. ERRADO
27. ERRADO
28. CERTO
29. ERRADO
30. ERRADO
31. ERRADO
32. CERTO
33. CERTO
34. ERRADO
35. CERTO
36. ERRADO
37. ERRADO
38. CERTO
39. CERTO
40. CERTO
41. LETRA E
42. LETRA B
43. LETRA D
44. LETRA A
45. LETRA B
46. LETRA A
47. LETRA A
48. LETRA C
49. LETRA D
50. LETRA B
51. LETRA A
52. LETRA C
53. LETRA E
54. LETRA B
55. LETRA E
56. LETRA D
57. LETRA D
58. LETRA C
59. LETRA C
60. LETRA D
61. LETRA C
62. LETRA C
63. LETRA B
64. LETRA D
65. LETRA B
66. LETRA B
67. LETRA E
68. LETRA D
69. LETRA C
70. LETRA B
71. LETRA A
72. LETRA D
73. LETRA E

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

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

1. 

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

2. 

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

3. 

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

4. 

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

5. 

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

6. 

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

7. 

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

8. 

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



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