Grande guerreiro! Na aula de hoje veremos Estatística que é um assunto fundamental nos concursos da EEAR e da AFA.

Então, vem comigo para mais uma batalha!

“***Nenhuma deficiência resiste ao trabalho*.” (Bernardinho)**

SUMÁRIO

ESTATÍSTICA

1. CONCEITOS BÁSICOS

1.1. ESTATÍSTICA

É a ciência que utiliza a coleta de dados, sua classificação, sua apresentação, sua análise e sua interpretação para se tomar algum tipo de decisão.

1.2. ESTATÍSTICA DESCRITIVA

É o ramo da Estatística que se ocupa em organizar e descrever os dados que podem ser expressos por tabelas e gráficos.

1.3. ESTATÍSTICA INFERENCIAL

É o ramo da Estatística que utiliza técnicas de análise e interpretação de dados, a partir de uma amostra de uma população, e fornece conclusões sobre este conjunto.

1.4. POPULAÇÃO

Na coleta de dados sobre determinado assunto, chama-se população estatística, o conjunto formado por todos os elementos que possam oferecer dados relativos ao assunto em questão.

Podemos dizer que população é qualquer conjunto que reúna **todos** os elementos que tenham pelo menos uma característica comum, objeto de estudo.

1.5. AMOSTRA

É um subconjunto de uma população. A seleção da amostra pode ser feita de várias maneiras, dependendo, entre outros fatores, do grau de conhecimento que temos da população, da quantidade de recursos disponíveis e outros fatores. A seleção da amostra deve fornecer um subconjunto de valores mais parecido possível com a população original.

EXEMPLO: Uma pesquisa típica de audiência na televisão utiliza uma amostra de 5000 lares e, com base nestes dados, formula conclusões acerca de uma população de todos os milhões de lares no país.

1.6. CENSO

É o método utilizado para um trabalho com uma população.

1.7. DADOS ESTATÍSTICOS

Os dados são denominados quantitativos quando são representados por números ou medidas, como por exemplo as alturas de uma população, o número de filhos e o salário bruto. Quando os dados representam contagens são discretos e quando representam mensurações são contínuos.

Os dados são chamados de qualitativos ou nominais quando são definidos por categorias tais como: cor dos olhos, sexo, nível de escolaridade, naturalidade.

1.8. AMPLITUDE DE UMA AMOSTRA

A amplitude total dos dados é a diferença entre o valor máximo e o valor mínimo da amostra.

EXEMPLO:

Um pesquisador, contratado pela empresa de cervejas, deseja estudar quantas cervejas por semana seus clientes bebem. A amostra com 10 clientes resultou nos seguintes números: 2,3,7,1, 10, 11, 5, 2, 8, 9.

A amplitude desta amostra é igual a 11-1=10.

1.9. ROL

Os dados coletados em uma amostra podem ser organizados em tabelas ou gráficos. Para isso, antes devemos organizá-los em sequências crescentes ou decrescentes denominadas **Rol.**

EXEMPLO:

No exemplo anterior organizando em ordem crescente temos: 1, 2, 2, 3, 5, 7, 8, 9, 10, 11.

1.10. DADOS BRUTOS

Podemos considerar como dados brutos aqueles que não estão numericamente organizados.

EXEMPLO: 2,3,7,1, 10, 11, 5, 2, 8, 9.

1.11. VARIÁVEIS

Uma outra definição que aparece na análise de dados estatísticos é o conceito de variável.

Uma variável é quantitativa quando seus valores podem ser representados por contagem (variáveis quantitativas discretas) ou mensuração (variáveis quantitativas contínuas).

Uma variável é qualitativa quando apresentam como resultado um atributo, qualidade ou preferência de um entrevistado.

1.12. SÉRIES ESTATÍSTICAS

Uma série estatística é um conjunto de dados estatísticos que fazem referência aos seguintes fatores: tempo, local e fenômeno.

Os exemplos mais comuns de séries estatísticas são:

- A Série Temporal, Cronológica ou Histórica;
- Varia o Tempo
- A Série Geográfica, Territorial ou Espacial;
- Varia o Local
- A Série Específica ou Especificativa.

- Varia o Fenômeno

Elementos **essenciais** em uma tabela:

1) **Título** – é a indicação contida na parte superior da tabela, onde deve estar definido o **fato** observado.

2) **Corpo** – é constituído por **linhas** e **colunas**, que fornecem o conteúdo das informações prestadas.

3) **Cabeçalho** – é a parte da tabela que apresenta a natureza do que contém cada coluna

Há ainda os elementos que complementam a tabela como por exemplo:

1) **Fonte** – designação da instituição que forneceu os dados estatísticos.

EXEMPLO: Datafolha, IBOPE, IBGE, INPE e etc.

2) **Notas** – esclarecimentos de natureza geral.

2. TABELAS DE FREQUÊNCIAS

Um processo que possibilita uma leitura mais sucinta dos dados é a construção de uma tabela de frequências.

EXEMPLO 1

Uma entrevista com 20 pessoas é realizada no estado do Rio de Janeiro. O objetivo da pesquisa era saber qual o time do entrevistado.

Dos 20 entrevistados foram encontrados os seguintes resultados para a **frequência absoluta** dos entrevistados:

Flamengo(f_1=10)

Vasco(f_2-6)

Fluminense($f_3=2$)

Botafogo($f_4=1$)

Note que $f_1+f_2+f_3+f_4=20$.

Definimos frequência relativa absoluta assumido por uma variável como a razão entre a frequência absoluta e o número total de dados.

$$f_{ri} = \frac{f_i}{n}$$

TIME	FREQUÊNCIA ABSOLUTA (f_i)	FREQUÊNCIA RELATIVA(f_{ri})	PORCENTAGEM
FLAMENGO	10	$\frac{10}{20}$	50%
VASCO	6	$\frac{6}{20}$	30%
FLUMINENSE	3	$\frac{3}{20}$	15%
BOTAFOGO	1	$\frac{1}{20}$	5%
TOTAL	20	1	100%

EXEMPLO 2:

Para avaliar o tempo de permanência em um supermercado, o gerente mensurou em minutos o tempo de permanência de 20 clientes na loja. Os tempos estão representados na tabela abaixo.

49	52	56	52	50
54	57	60	48	59
48	49	57	53	55

51	53	52	55	57

Para representar esses dados em uma tabela de frequência devemos:

a) calcular a amplitude da amostra.

60 – 48 = 12 cm

b) Dividir o intervalo em subintervalos de mesmo comprimento. Como exemplo, tomemos 4 subintervalos de comprimento igual a 3: [48, 51[; [51,54[; [54,57[; [57,60]. Esses subintervalos são chamados de classes e o comprimento de cada um é chamado de amplitude da classe.

c) Conte quantas observações se situam em cada classe, respeitando os intervalos fechados à esquerda e abertos à direita, e coloque as observações numa tabela do tipo abaixo.

Comprimento	fi	fri	fac	frac
[48, 51[	5	$\frac{5}{20}=0,25=25\%$	5	25%
[51, 54[	6	$\frac{6}{20}=0,3=30\%$	11	55%
[54, 57[	4	$\frac{4}{20}=0,20=20\%$	15	75%
[57, 60]	5	$\frac{5}{20}=0,25=25\%$	20	100%
Total	20			

Onde,

fi: frequência que o intervalo aparece na distribuição;

fri (Frequência relativa): É a percentagem do valor dos dados em relação ao total da amostra;

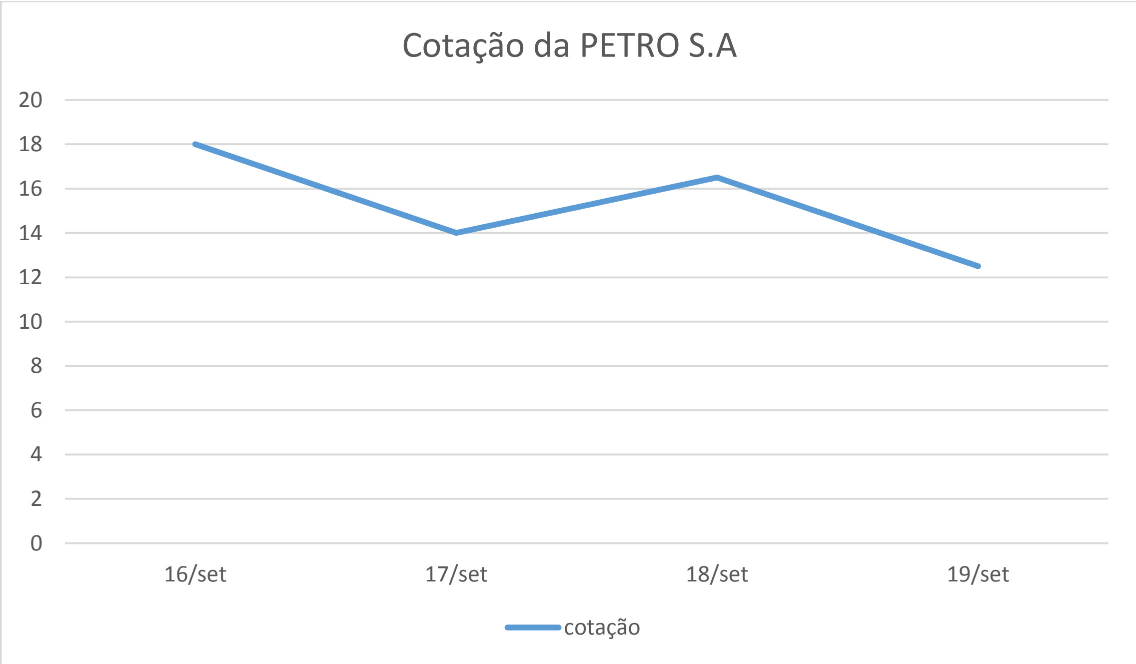
fac (frequência acumulada): É a soma das frequências absolutas começando pelo menor valor;

frac (frequência relativa acumulada): É a porcentagem do valor das frequências acumuladas em relação ao total da amostra.

3. GRÁFICOS

3.1. GRÁFICO EM LINHA

Esse tipo de gráfico é usado sobretudo quando temos observações temporais de uma variável em estudo e desejamos representá-la no tempo (abscissa) afim de reconhecer possíveis tendências e/ou sazonalidade (comportamento periódicos repetidos). O exemplo a seguir ilustra bem a utilidade do gráfico em linha para a evolução do preço da ação da empresa PETRO S.A.



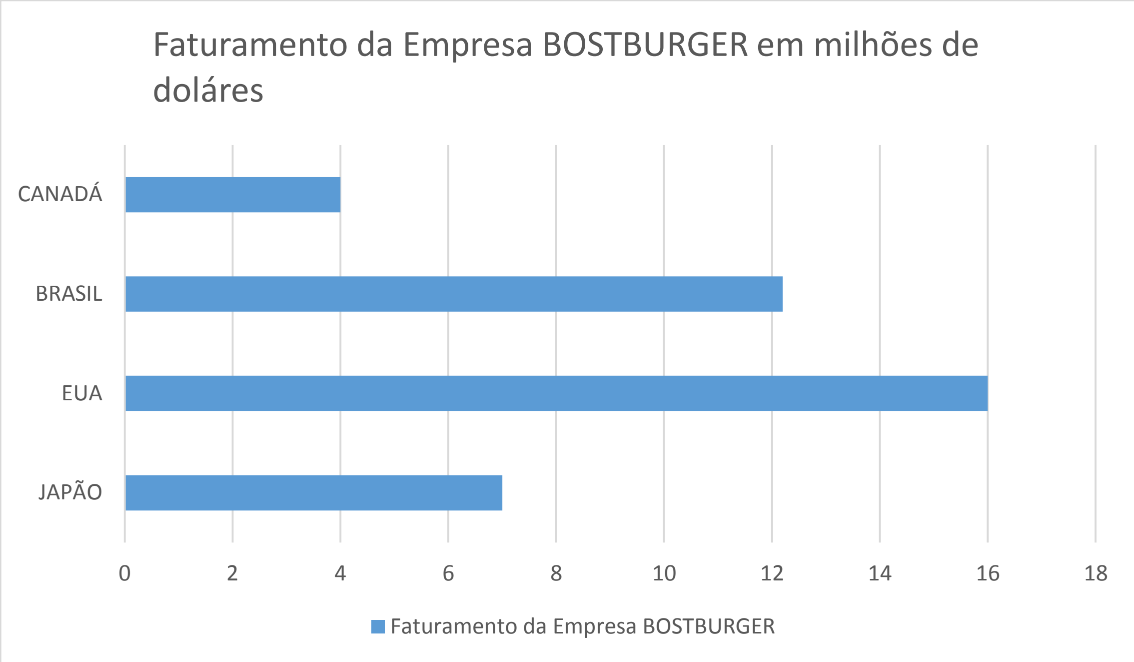
3.2. GRÁFICO EM BARRAS HORIZONTAIS

Os dados que estejam organizados em colunas ou linhas em uma tabela podem ser representados em um gráfico de barras horizontais. Gráficos de barras ilustram comparações entre itens individuais.

Considere a utilização de um gráfico de barras horizontais quando:

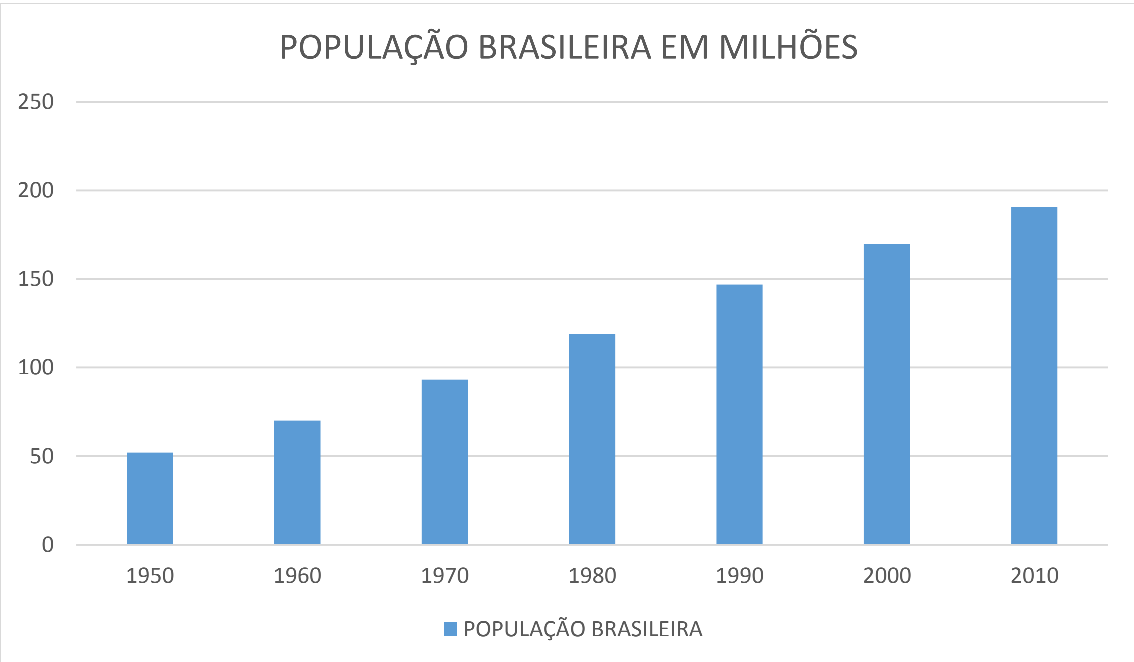
- Os rótulos dos eixos forem longos.

- Os valores mostrados forem durações.



3.3. GRÁFICO DE COLUNAS

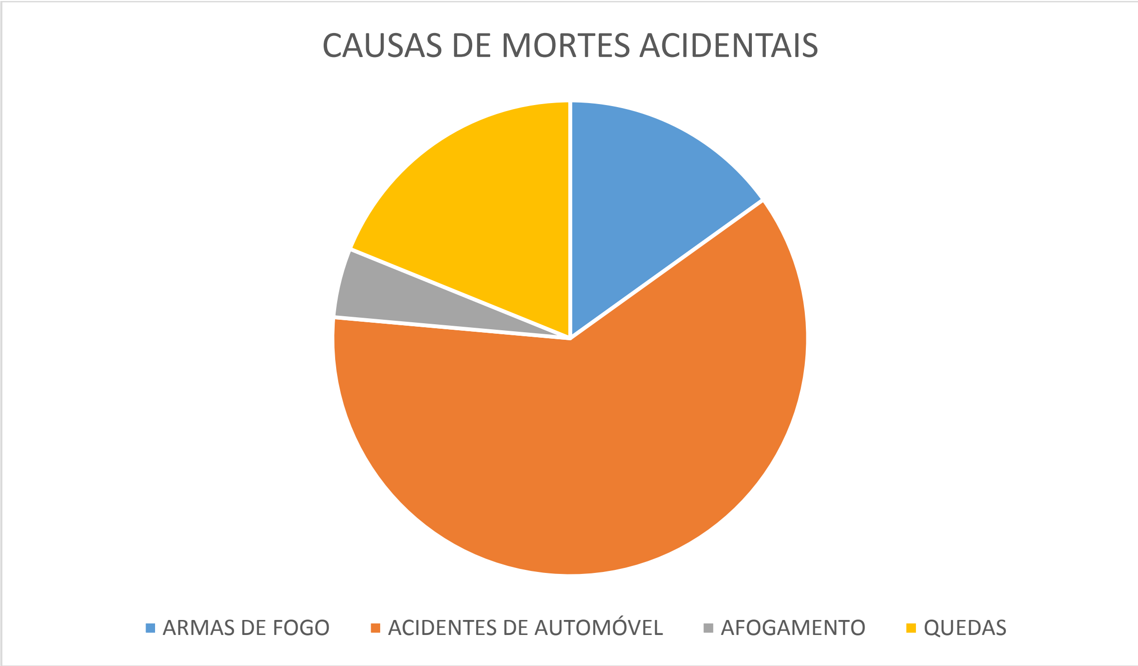
A ideia é expressar informações individualizadas, e representadas por barras cuja altura representa a frequência nas categorias. Vejamos o exemplo a seguir, representando em barras o número de cópias de jornais (em milhares de exemplares) em alguns países.



3.4. GRÁFICO EM SETORES

É utilizado quando se deseja mostrar partes do total, conforme ocorre em produções, vendas e orçamentos de países e etc.

EXEMPLO:



3.5. HISTOGRAMA

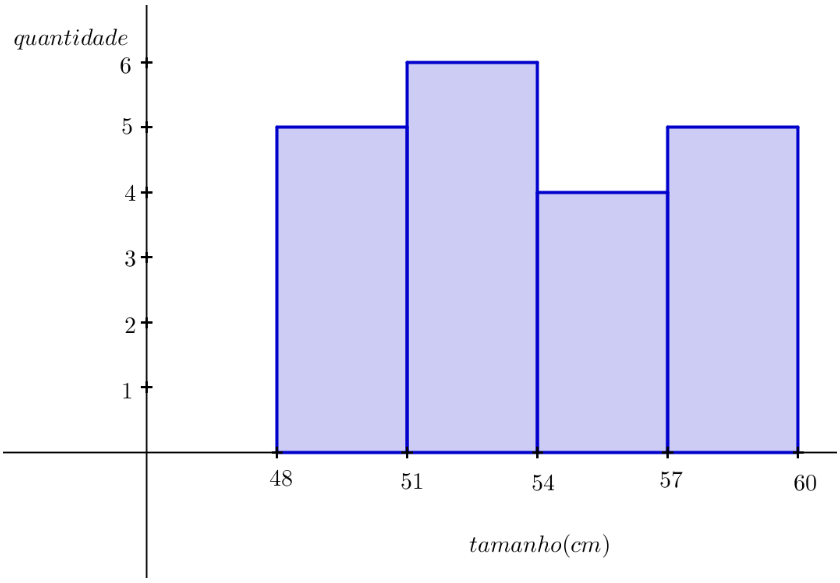
Quando as classes são intervalos reais, a interpretação da distribuição de frequências em um sistema de eixos é feita por um tipo de gráfico chamado Histograma.

EXEMPLO:

Voltando ao exemplo 2

Tempo	fi
[48, 51[	5
[51, 54[	6
[54, 57[	4
[57, 60]	5
Total	20

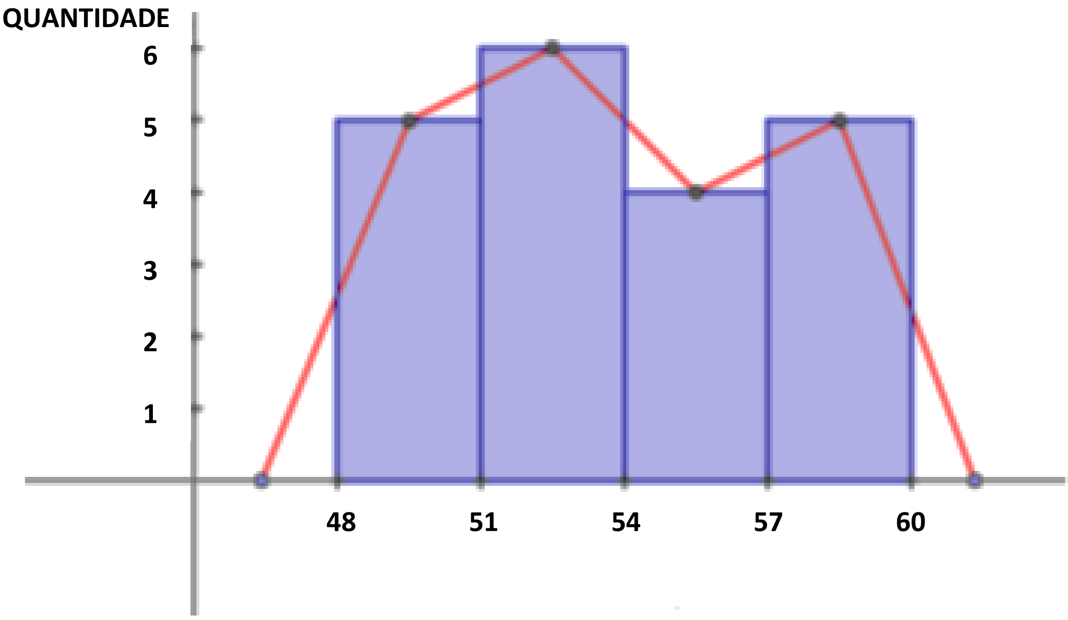
O Histograma referente à tabela é



3.6. POLÍGONO DE FREQUÊNCIA

O polígono de frequência é obtido unindo-se os pontos médios da parte superior de cada retângulo do histograma com segmentos de reta. É importante notar que tanto o histograma quanto o polígono de frequência indicam a frequência absoluta de cada classe.

Voltando ao exemplo 2 temos:



O gráfico em linhas representativo de uma distribuição de frequências acumuladas é chamado **polígono de frequências acumuladas ou ogiva de Galton**. No caso de dados agrupados em intervalos de classe, os pontos do gráfico são os pontos correspondentes aos limites superiores das classes das bases superiores dos retângulos.

3.7. CARTOGRAMA

Um cartograma é um mapa que mostra informação quantitativa mantendo um certo grau de precisão geográfica das unidades espaciais mapeadas.

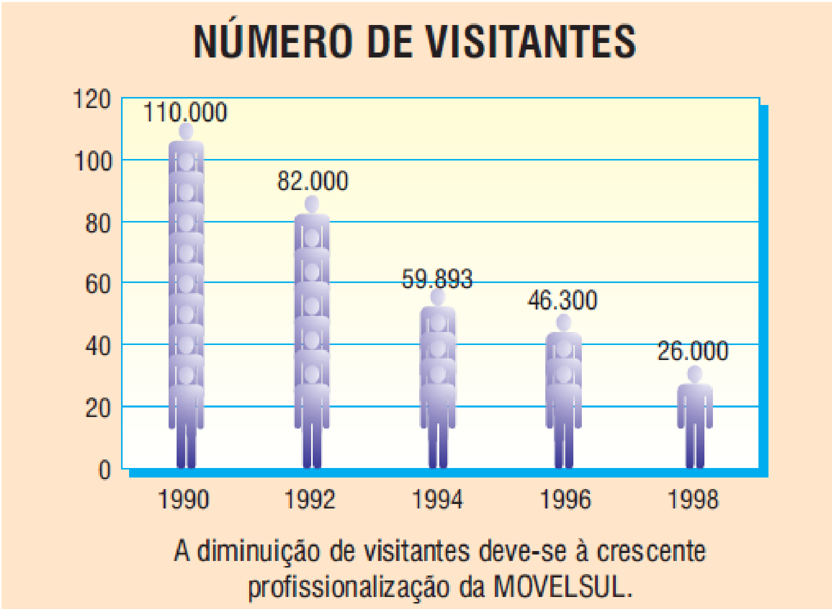
EXEMPLO:

% de Prefeitos Eleitos em 2012: PT

3.8. PICTOGRAMA

É comum em jornais e revistas ilustrar os vários tipos de gráficos com figuras relacionadas ao assunto, tornando-os mais atraentes. Esses são os pictogramas.

EXEMPLO:



Fonte: Google

4. MEDIDAS DE CENTRALIDADE: MÉDIA ARITMÉTICA, MÉDIA ARITMÉTICA PONDERADA, MEDIANA, MODA.

Para apresentar os conceitos a seguir, vamos considerar um grupo de 10 estudantes com as seguintes idades: 12, 13, 14, 14, 14, 15, 15, 16, 16, 17.

4.1. MÉDIA ARITMÉTICA E MÉDIA ARITMÉTICA PONDERADA

Seja x uma variável quantitativa e $x_1, x_2, x_3, \ldots, x_n$ os valores assumidos por essa variável. A média aritmética $(\overline{x})$ de x é igual a soma de todos os valores assumidos pela variável dividida pelo número de valores, ou seja,

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

EXEMPLO: $\overline{x} = \dfrac{12+13+14+14+14+15+15+16+16+17}{10} = 14,6$

PROPRIEDADES DA MÉDIA

1) Ao adicionarmos um mesmo valor a cada um dos valores assumidos pela variável, a média aritmética fica adicionada desse valor.

2) Ao multiplicarmos cada um dos valores assumidos pela variável por um mesmo valor, a média aritmética fica multiplicada por esse valor.

Seja x uma variável quantitativa e $x_1, x_2, x_3, \ldots, x_k$ os valores assumidos por essa variável, com frequências absolutas respectivamente iguais a $n_1, n_2, n_3, \ldots, n_k$. A média aritmética $(\overline{x})$ desses valores é igual a soma de cada um os valores assumidos pela variável multiplicado pela sua frequência dividida pela soma das frequências, ou seja, $\overline{x} = \dfrac{\sum_{i=1}^{k} x_i \cdot n_i}{\sum_{i=1}^{k} n_i} = \dfrac{x_1 \cdot n_1 + x_2 \cdot n_2 + \cdots + x_k \cdot n_k}{n_1 + n_2 + \cdots + n_k}$.

A média aritmética é muito influenciada por valores discrepantes (“outliers”).

4.2. MEDIANA

Seja $x_1 \le x_2 \le x_3 \le \cdots \le x_n$ os valores ordenados assumidos pela variável quantitativa x. A mediana (Me) é dada por:

$$\mathrm{Me} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ é ímpar} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ é par} \end{cases}.$$

Dessa forma, a mediana é tal que a quantidade de valores menores ou iguais à mediana é igual à quantidade de valores maiores ou iguais à mediana.

A mediana é uma medida de centralidade menos sensível a valores discrepantes.

EXEMPLO: No nosso exemplo a mediana é $\frac{x_5 + x_6}{2} = \frac{14 + 15}{2} = 14,5$.

4.3. MODA

A moda de um conjunto é o valor que ocorre mais vezes ou de maior frequência simples (absoluta ou relativa) numa distribuição de frequências.

No nosso exemplo a moda é igual a 14.

A moda pode também não existir ou não ser única.

EXEMPLO: 0,1,1,2,2, 3 tem modas 1 e 2 (bimodal).

5. MEDIDAS DE DISPERSÃO: DESVIO MÉDIO, VARIÂNCIA, DESVIO PADRÃO

Considere uma turma de 5 alunos em todos tiraram nota 5 e outra com a mesma quantidade de alunos com as seguintes notas: 1, 3, 5, 7 e 9. Essas duas turmas têm a mesma média aritmética e a mesma mediana que é 5. Mas a dispersão dos valores é completamente diferente e pode ser calculado.

5.1. DESVIO MÉDIO ABSOLUTO (DMA)

O desvio em relação à média aritmética é a diferença entre cada valor e a média aritmética.

$$\delta_i = x_i - \overline{x}$$

A soma de todos os desvios em relação à média aritmética é sempre nula.

O desvio médio absoluto (DMA) é a média aritmética dos desvios em módulo.

$$\mathrm{DMA} = \frac{\sum_{i=1}^{n} |\delta_i|}{n} = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n}$$

O desvio médio absoluto da segunda turma do nosso exemplo é $\mathrm{DMA} = \frac{4+1+0+1+4}{5} = 2$.

5.2. VARIÂNCIA POPULACIONAL

A **variância populacional** (σ^2) é a média aritmética da soma dos quadrados dos desvios em relação à média de um conjunto de números.

$$\sigma^2 = \frac{\sum_{i=1}^{n} \delta_i^2}{n} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n}$$

Outra fórmula para o cálculo da variância populacional é $\sigma^2 = \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]$.

5.3. PROPRIEDADES DA VARIÂNCIA

Se adicionarmos um mesmo valor a cada um dos valores assumidos pela variável, a variância não se altera.

Se multiplicarmos cada um dos valores assumidos pela variável por um mesmo valor, a variância fica multiplicada pelo quadrado desse valor.

Se a variância for calculada sobre uma amostra em vez de sobre toda a população, teremos então a chamada **variância amostral** que é dada por $S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$

5.4. DESVIO PADRÃO POPULACIONAL

O **desvio padrão populacional** (σ) é a média quadrática dos desvios em relação à média de um conjunto de números ou a raiz quadrada da variância.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \delta_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

Outra fórmula para o cálculo do desvio padrão populacional é $\sigma = \sqrt{\frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]}$.

Se os dados forem agrupados em intervalos de classe, então o quadrado de cada desvio deve ser multiplicado pela sua frequência absoluta (n_i) ou relativa (f_i).

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 \cdot f_i}{n}} = \sqrt{\sum_{i=1}^{k} (x_i - \bar{x})^2 \cdot fr_i}$$

5.5. PROPRIEDADES DO DESVIO PADRÃO

Se adicionarmos um mesmo valor a cada um dos valores assumidos pela variável, o desvio padrão não se altera.

Se multiplicarmos cada um dos valores assumidos pela variável por um mesmo valor, o desvio padrão fica multiplicado por esse valor.

OBSERVAÇÃO

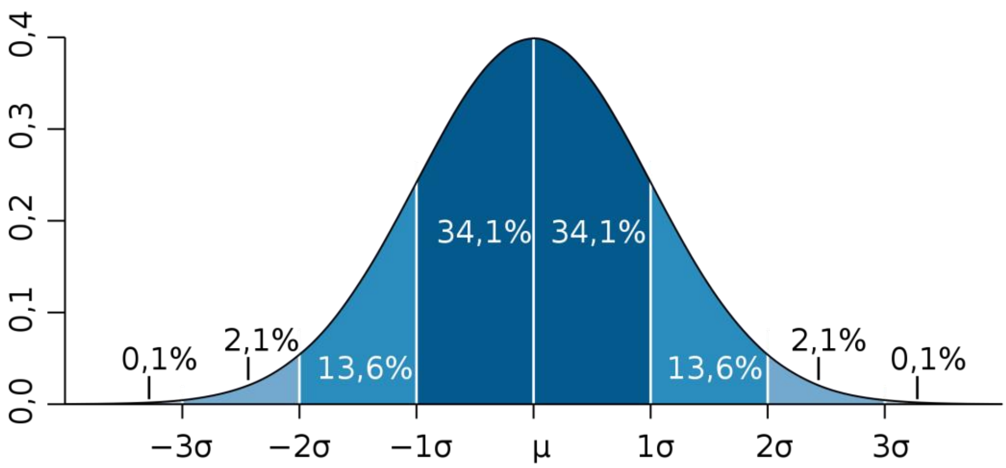
Se o desvio padrão for calculado sobre uma amostra em vez de sobre toda a população, teremos então o chamado **desvio padrão amostral** que é dado por $S = \sqrt{\dfrac{\sum_{i=1}^{n} \delta_i^2}{n-1}} = \sqrt{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$ ou para dados agrupados por $S = \sqrt{\dfrac{\sum_{i=1}^{k} \delta_i^2 \cdot f_i}{n-1}} = \sqrt{\dfrac{\sum_{i=1}^{k} (x_i - \bar{x})^2 \cdot f_i}{n-1}}$.

A razão entre o desvio padrão amostral e o populacional é $\sqrt{\dfrac{n}{n-1}}$ chamado fator de correção de Bessel.

Se os valores se distribuem de acordo com uma distribuição normal (unimodal, gaussiana, simétrica, de afunilamento médio) podemos dizer que:

- 68% dos valores encontram-se a uma distância da média inferior a um desvio padrão.
- 95% dos valores encontram-se a uma distância da média inferior a duas vezes o desvio padrão.
- 99,7% dos valores encontram-se a uma distância da média inferior a três vezes o desvio padrão.

Esta informação é conhecida como a regra dos "68-95-99,7".



Fonte: Wikipédia

Exemplo: Sejam 10 estudantes com as seguintes idades: 13, 13, 14, 14, 14, 14, 15, 15, 15, 16.

Idade	n_i	δ_i	δ_i^2
13	2	$-1,3$	$1,69$
14	4	$-0,3$	$0,09$
15	3	$0,7$	$0,49$
16	1	$1,7$	$2,89$

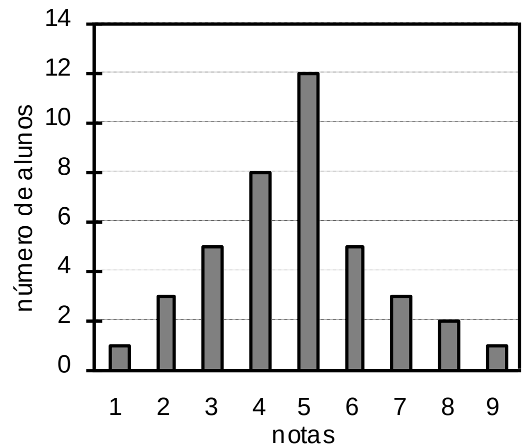
O desvio padrão populacional é $\sigma = \sqrt{\dfrac{2 \cdot 1,69 + 4 \cdot 0,09 + 3 \cdot 0,49 + 1 \cdot 2,89}{2 + 4 + 3 + 1}} = \sqrt{0,81} = 0,9$.

1. (EEAr 2000) Numa prova de matemática, três classes obtiveram as seguintes médias e desvios:
classe A : $\overline{x} = 4,5$ e $\delta = 2,5$
classe B : $\overline{x} = 4,5$ e $\delta = 3,1$
classe C : $\overline{x} = 4,5$ e $\delta = 2,8$

Se for sorteado um aluno em cada classe, em qual delas é mais provável que a nota desse aluno esteja entre 3,0 e 6,0 ?

a) Classe A
b) Classe B
c) Classe C
d) Classes B e C

2. (EEAr 2000) Os resultados da prova de Ciências aplicada a uma turma de um certo colégio estão apresentados no gráfico. Baseado neste gráfico, podemos afirmar que a porcentagem de alunos dessa turma com nota inferior a $5,0$, nessa prova de Ciências, foi de



a) $37,5\%$
b) $42,5\%$
c) $47,5\%$
d) $52,5\%$

3. (EEAr 2002) A altura de 80 homens de uma comunidade está distribuída de acordo com a tabela. A porcentagem de homens com altura maior ou igual a $1,80\,m$ é

altura (m)	número de homens
1,60 |— 1,65	04
1,65 |— 1,70	12
1,70 |— 1,75	18
1,75 |— 1,80	26
1,80 |— 1,85	10
1,85 |— 1,90	08
1,90 |— 1,95	02
Total	80

a) 25%
b) 30%
c) 60%
d) 75%

4. (EEAr 2003) Um teste de inteligência, aplicado aos alunos das 4^{as} séries do Ensino Fundamental da Escola A, apresentou os seguintes resultados:

Pontos	n.º de alunos	Pontos	n.º de alunos
90 |— 95	40	115 |— 120	140
95 |— 100	60	120 |— 125	120
100 |— 105	140	125 |— 130	30
105 |— 110	160	130 |— 135	20
110 |— 115	180	135 |— 140	10

A frequência relativa da classe modal é
a) 0,2
b) 0,22
c) 0,25
d) 0,5

5. (EEAr 2004) Na distribuição a seguir, as 6 classes foram justapostas, isto é, o limite superior de uma classe é o inferior da classe seguinte. Se f_i é frequência absoluta, então a frequência acumulada absoluta da 4.ª classe é

Classes	5 |-- 10	10 |-- 15	15 |-- 20	20|-- 30	25|-- 30	30 |--
f_i	8	12	21	16	6	3

a) 57
b) 20
c) 41
d) 16

6. (EEAr 2004) A média de um conjunto de quatro valores é $4{,}25$. Se aumentarmos de 5 unidades o menor desses valores, e diminuirmos de 3 unidades o maior deles, a nova média será
a) 4,75
b) 5,25
c) 5
d) 5,5

7. (EEAr 2005) A tabela traz as idades, em anos, dos filhos de 5 mães.

Nome da Mãe	Ana	Márcia	Cláudia	Lúcia	Eloisa
Idade dos filhos	7, 10, 12	11, 15	8, 10, 12	12, 14	9, 12, 15, 16, 18

A idade modal desses 15 filhos é inferior à idade média dos filhos de Eloísa em _______ anos(s).
a) 4
b) 3
c) 2
d) 1

8. (EEAr 2005) Na distribuição dos salários de 800 empregados de uma empresa, o ponto médio da 4.ª classe é R$ 1400,00. Se as 8 classes dessa distribuição têm a mesma amplitude de R$ 200,00 e são do tipo [a, b[, então a 6.ª classe não inclui, com certeza, o salário de R$
a) 1900,00
b) 1850,00
c) 1800,00
d) 1750,00

9. (EEAr 2005) Sejam $x_1, x_2, x_3, \ldots, x_{81}$ os valores ordenados de uma variável X. A mediana desse conjunto de valores é igual a

a) x_{41}

b) x_{40}

c) $\dfrac{x_{40} + x_{41}}{2}$

d) $\dfrac{x_{41} + x_{42}}{2}$

10. (EEAr 2006) Os resultados de uma pesquisa realizada com 20 alunos de uma escola, a respeito da área da carreira pretendida, estão apresentados na tabela:

Área	Frequência Absoluta	Frequência Relativa
Humanas	8	M
Biológicas	P	0,35
Exatas	R	S
Total	20	1,00

Os valores de M, P, R e S são, respectivamente:

a) 0,35; 5; 7 e 0,35.

b) 0,4; 7; 5 e 0,4.

c) 0,4; 7; 5 e 0,25.

d) 0,25; 5; 7 e 0,25.

11. (EEAr 2006) Sendo f_i as frequências absolutas, a classe mediana da distribuição é a:

classe	[10,20[	[20,30[	[30,40[	[40,50[	[50,60[	[60,70[	[70,80[
f_i	25	18	10	05	09	12	15

a) 2ª

b) 3ª

c) 4ª

d) 5ª

12. (EEAr 2006) A tabela mostra as idades dos alunos matriculados no Centro de Educação Infantil "X", em 2005.

Idade (anos)	Número de alunos
2	3
3	3
4	5
5	14
6	25
Total	50

A média das idades dos alunos dessa escola, em anos, é, aproximadamente:

a) 4,1
b) 4,5
c) 5,1
d) 5,6

13. (EEAr 2007) A tabela a seguir traz o resultado de uma prova de Ciências. Nela, x_i são as notas e f_i são as frequências absolutas. Agrupando os dados em 5 classes do tipo $[a,b[$, de amplitude $1,5$, sendo o limite inferior da 1.ª classe a nota $1,5$, a frequência absoluta da 3.ª classe da nova tabela será igual a

x_i	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5
f_i	1	2	2	3	5	6	7	8	9	7	6	5	4	3	2

a) 14
b) 19
c) 24
d) 29

14. (EEAr 2007) Os dados de uma pesquisa, cujo objetivo era saber o número de filhos, por família, realizada em uma certa comunidade, estão na tabela:

Nº de filhos	0	1	2	3	4	5
Nº de famílias	2	8	10	14	18	15

É correto afirmar que o número

a) modal de filhos é maior que o número médio.
b) médio de filhos coincide com o número modal.
c) mediano e o número modal de filhos são iguais.
d) modal, o mediano e o número médio de filhos são iguais.

15. (EEAr 2007) Quando o objetivo de uma pesquisa é comparar o comportamento de uma mesma variável em populações com números diferentes de elementos, a frequência mais conveniente é a

a) total.
b) relativa.
c) absoluta.
d) acumulada.

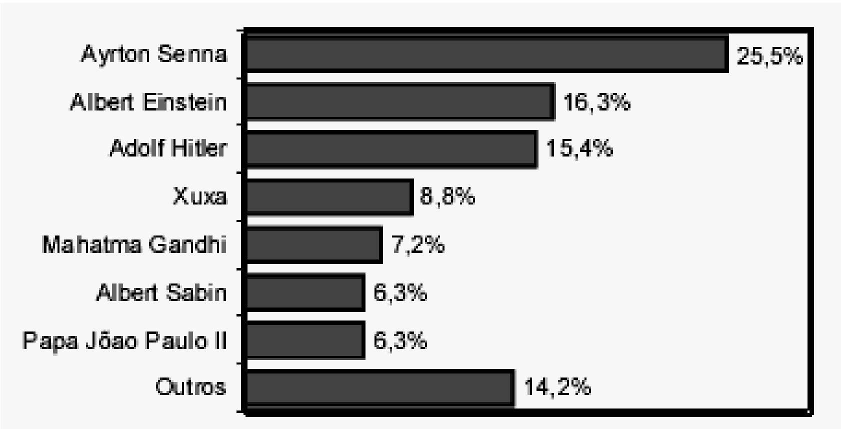
16. (EEAr 2007) Feito um levantamento sobre a altura dos 50 alunos da 5ª série A de um colégio, chegou-se aos seguintes resultados:

Altura (cm)	nº de alunos	Altura	nº de alunos
150 $\vdash\!\!\rightarrow$ 154	6	162 $\vdash\!\!\rightarrow$ 166	8
154 $\vdash\!\!\rightarrow$ 158	12	166 $\vdash\!\!\rightarrow$ 170	6
158 $\vdash\!\!\rightarrow$ 162	14	170 $\vdash\!\!\rightarrow$ 174	4

Nessas condições, o número de alunos da 5ª A que não atingem 1,58 m de altura, e a porcentagem de alunos cuja altura é maior ou igual a 1,62 m são, respectivamente,

a) 12 e 12%
b) 12 e 20%
c) 18 e 36 %
d) 18 e 20%

17. (EEAr 2008) A revista Época publicou, em janeiro de 2000, os resultados de uma pesquisa por ela realizada em setembro de 1999. Cada participante indicava o nome de uma personalidade mundialmente conhecida, do século XX, da qual ele mais se lembrava. O gráfico a seguir traz o percentual de pessoas que indicaram cada uma dessas personalidades.



Sabendo que participaram dessa pesquisa 60 mil pessoas, Ayrton Senna foi indicado por ______ pessoas.

a) 12 800
b) 15 300

c) 16 900
d) 18 600

18. (EEAr 2008) Seja a distribuição de frequência, onde f_i é a frequência simples absoluta:

x_i	4	8	10	12	20
f_i	9	10	16	30	35

A média dessa distribuição é
a) 10,28
b) 11,17
c) 13,36
d) 14,15

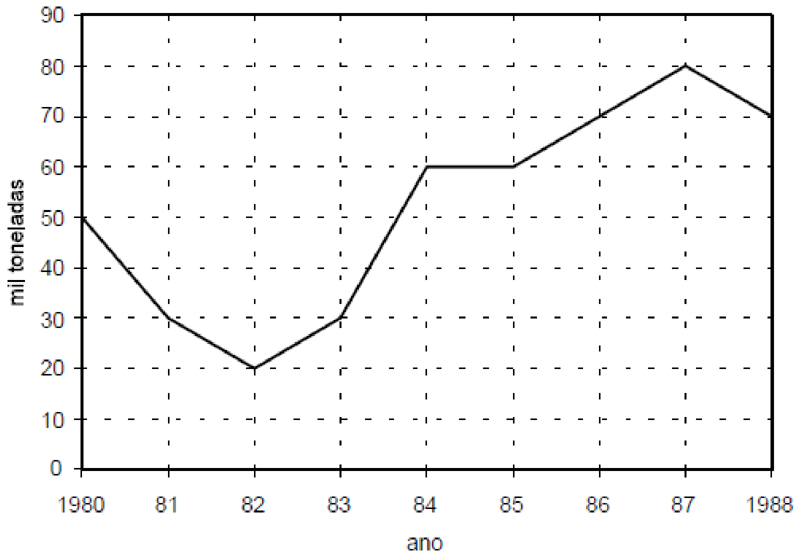
19. (EEAr 2008) Segundo a distribuição de frequências, o número de funcionários que ganham a partir de 4 salários mínimos e menos de 10 é

Número de salários mínimos	Número de funcionários
0 |-- 2	95
2 |-- 4	75
4 |-- 6	45
6 |-- 8	35
8 |-- 10	30
10 |-- 12	20

a) 110.
b) 130.
c) 185.
d) 205.

20. (EEAr 2009) Na 5ª série A do Colégio X, numa prova de Ciências, 8 alunos obtiveram notas menores que 4; 15 alunos, notas de 4 a 6; 20 alunos, notas entre 6 e 8; e apenas 2, notas a partir de 8.
A nota modal da 5ª série A, nessa prova de Ciências, foi
a) 8.
b) 7.
c) 6.
d) 5.

21. (EEAr 2010) O gráfico representa a produção de arroz, em milhares de toneladas, em certo país, no período 1980-1988.



Pelo gráfico, pode-se concluir que, no período 1980-1988, nesse país, a produção média anual de arroz, em mil toneladas, é, aproximadamente,

a) 64.
b) 60.
c) 58.
d) 52.

22. (EEAr 2010) Os salários mensais, em reais, dos 24 funcionários de uma empresa são

800	840	880	880	1000	1050	1060	1060
1100	1150	1200	1210	1230	1250	1280	1300
1340	1380	1450	1480	1500	1500	1520	1550

O salário mensal mediano dessa empresa, em reais, é:

a) 1200
b) 1210
c) 1220
d) 1230

23. (EEAr 2011) Um teste de Matemática foi aplicado em duas turmas distintas de uma escola, a primeira com 40 alunos e a segunda com 20. As médias aritméticas das notas da primeira e da segunda turma forma, respectivamente, $6,0$ e $7,0$. Assim, a média aritmética das notas dos 60 alunos foi aproximadamente

a) 6,1

b) 6,3
c) 7,2
d) 7,5

24. (EEAr 2012) Em um supermercado, Ana pesquisou o preço de cinco marcas de molho de tomate e obteve os seguintes valores, em reais: 2,05; 1,92; 2,16; 1,98 e 2,11. O valor mediano, em reais, é

a) 2,05.
b) 1,92.
c) 1,74.
d) 1,85.

GABARITO

1. Observe que, comparando-se duas amostras que possuam a mesma média aritmética, aquela de menor desvio padrão apresentará resultados mais próximos da média do que a outra.

Como o intervalo $[3,0;6,0]$ é simétrico em relação à média aritmética $\overline{x}=4,5$, a probabilidade do aluno ter uma nota nesse intervalo é maior na classe que tem mais notas próximas à média, ou seja, naquela que tem o menor desvio médio, que é a classe A.

RESPOSTA: A

2. O total de alunos é $n=1+3+5+8+12+5+3+2+1=40$.
A porcentagem de alunos com nota inferior a $5,0$ é

$$F_4=\frac{N_4}{n}=\frac{1+3+5+8}{40}=\frac{17}{40}=42,5\%.$$

RESPOSTA: B

3.

$$\frac{10+8+2}{80}=\frac{1}{4}=25\%$$

RESPOSTA: A

4. A classe modal é a 5^a classe (110|— 115). A frequência relativa dessa classe é:

$$f_5=\frac{180}{40+60+140+160+180+140+120+30+20+10}=\frac{180}{900}=\frac{20}{100}=0,2.$$

RESPOSTA: A

5. $F_4=f_1+f_2+f_3+f_4=8+12+21+16=57$

RESPOSTA: A

6. Sejam $x_1 \le x_2 \le x_3 \le x_4$ cuja média é $\frac{x_1+x_2+x_3+x_4}{4}=4,25 \Leftrightarrow x_1+x_2+x_3+x_4=17$.

A nova média é dada por $\frac{(x_1+5)+x_2+x_3+(x_4-3)}{4}=\frac{(x_1+x_2+x_3+x_4)+2}{4}=\frac{17+2}{4}=4,75$.

RESPOSTA: A

7.

A idade modal é 12 anos.

A idade média dos filhos de Eloísa é $\frac{9+12+15+16+18}{5}=14$.

A idade modal desses 15 filhos é inferior à idade média dos filhos de Eloísa em $14-12=2$ anos.

RESPOSTA: C

8. Se R$ 1400,00 é o ponto médio da 4.ª classe e ela tem amplitude R$ 200,00, então a 4ª classe é 1300|−−1500.

Portanto, a 5ª classe é 1500|−−1700 e a 6ª classe 1700|−−1900, ou seja, $[1700,1900[$ que não inclui R$ 1900,00.

RESPOSTA: A

9. Como a quantidade de termos é ímpar a ordem da mediana é $\frac{81+1}{2}=41$. Logo, a mediana é x_{41}.

RESPOSTA: A

10.

A frequência relativa de humanas é $M=\frac{8}{20}=0,4$.

A frequência relativa da área biológica é $\frac{P}{20}=0,35 \Leftrightarrow P=7$.

Como o total de alunos é 20, temos: $8+7+R=20 \Leftrightarrow R=5$.

A frequência relativa de exatas é $S=\frac{5}{20}=0,25$.

RESPOSTA: A

11.

O total de observações é $n=25+18+10+5+9+12+15=94$.

Como $n=94$ é par, então a mediana é dada por $\dfrac{x_{\frac{94}{2}}+x_{\left(\frac{94}{2}+1\right)}}{2}=\dfrac{x_{47}+x_{48}}{2}$, como x_{47} e x_{48} estão na 4ª classe, então é essa a classe mediana.

12.

$$\bar{x}=\frac{2\cdot 3+3\cdot 3+4\cdot 5+5\cdot 14+6\cdot 25}{50}=5,1$$

RESPOSTA: C

13.

CLASSES		f_i
1ª classe	1,5 |-- 3,0	5
2ª classe	3,0 |-- 4,5	14
3ª classe	4,5 |--6,0	24
4ª classe	6,0 |-- 7,5	18
5ª classe	7,5 |-- 9,0	9

Logo, a frequência absoluta da 3ª classe é 24.

RESPOSTA: C

14.

A moda é 4.

A média é $\bar{x}=\dfrac{0\cdot 2+1\cdot 8+2\cdot 10+3\cdot 14+4\cdot 18+5\cdot 15}{2+8+10+14+18+15}=\dfrac{217}{67}\approx 3,24$.

A ordem da mediana é $\dfrac{67+1}{2}=34$ e o seu valor é 3.

Logo, o número modal de filhos é maior que o número médio

RESPOSTA: A

15.

A frequência relativa é obtida pela razão entre a frequência absoluta e o número total de elementos. Portanto, ela permite comparar amostras com quantidades de elementos diferentes.

RESPOSTA: A

16. O números de alunos que não atingem $1,58\,\text{m}$ é $6+12=18$.

A porcentagem de alunos cuja altura é maior ou igual a $1,62\,\text{m}$ é $\dfrac{8+6+4}{6+12+14+8+6+4}=\dfrac{18}{50}=36\%$.

RESPOSTA: C

17.

$$\frac{x}{60000}=25,5\% \Leftrightarrow x=0,255\cdot 60000=15300$$

RESPOSTA: B

18.

$$\overline{x}=\frac{4\cdot 9+8\cdot 10+10\cdot 16+12\cdot 30+20\cdot 35}{9+10+16+30+35}=\frac{1336}{100}=13,36$$

RESPOSTA: C

19. O número de funcionários que ganham a partir de 4 salários mínimos e menos de 10 é $45+35+30=110$.

RESPOSTA: A

20. A classe modal é $[6,8[$ e, portanto, a moda é $\dfrac{6+8}{2}=7$.

RESPOSTA: B

21.

$$\overline{x}=\frac{50+30+20+30+60+60+70+80+70}{9}=\frac{470}{9}\approx 52$$

RESPOSTA: D

22.

Como $n=24$ é par, o salário mediano é $\dfrac{x_{\left(\frac{24}{2}\right)}+x_{\left(\frac{24}{2}+1\right)}}{2}=\dfrac{x_{12}+x_{13}}{2}=\dfrac{1210+1230}{2}=1220$.

RESPOSTA: C

23.

$$\overline{x} = \frac{40 \cdot 6,0 + 20 \cdot 7,0}{40 + 20} = \frac{380}{60} \approx 6,3$$

RESPOSTA: B

24. Ordenando os valores, temos: $1,92; 1,98; 2,05; 2,11; 2,16$.

Como $n = 5$ é ímpar, a ordem da mediana é $\frac{5+1}{2} = 3$ e seu valor é $x_3 = 2,05$.

RESPOSTA: A