ANTONIO ROBERTO GONCALVES

O QUE É ESTATÍTICA

1. INTRODUÇÃO:

A palavra estatística foi usada pela primeira vez nos século XVIII, pelo Alemão Gottfued Achennall. A história nos indica, que nos anos 3 000 A. C., já se usava fazer os censos na Babilônia, China e Egito.

Pode-se ver na Bíblia, no velho testamento, no livro dos números, uma instrução para fazer o levantamento dos homens de Israel apto para guerrear.

Tem-se um outro fato, que consta da Bíblia, que por ocasião do nascimento de Jesus, o censo feito em todo o Império Romano, a mando do imperador César Augusto, que fez com que José e Maria viajassem para Belém.

Assim a Estatística, era usada bem antes de ser batizada com esse nome. A palavra Estatística vem de Estatus, que significa Estado em Latim. Com essa palavra fazia-se as descrições e dados relativos aos Estados, logo a Estatística era para os governantes um meio de administração.

Com o passar do tempo, seu uso era cada vez mais freqüente no levantamento de dados referentes aos que existia em cada país, com a finalidade de se ter uma base para que se pudesse cobrar os impostos pelos governantes.

O uso da Estatística é cada vez maior, e aprimorada em sua técnica e métodos, até que se cunha o nome de Estatística para todo esse tipo de atividade.

Embora exista diversa definição de Estatística, uma que se está de acordo com o nosso objetivo é a de Dugé Bernonville:

> "Estatística é a parte da Matemática Aplicada que fornece métodos para a coleta, organização, descrição, análise e interpretação de dados e para a utilização dos mesmos na tomada de decisões".

2. FASES DO MÉTODO ESTATÍSTICO.

Podemos distinguir no método estatístico as seguintes fases:

1) - **Coleta de Dados**: Depois de cuidadoso planejamento e a devida determinação das características mensuráveis do fenômeno que se quer

pesquisar, damos inicio a coleta dos dados numéricos necessários a sua descrição.

A coleta de dados pode ser **direta** ou **indireta.**

A **coleta é direta** quando feito sobre elementos informativos de registros obrigatórios, (nascimentos, casamentos, importação e exportação de mercadorias) ou ainda, quando os dados são coletados pelo próprio pesquisador através de inquéritos e questionários, como é o caso de notas de verificação e de exames, do censo demográfico, etc..

A coleta direta de dados pode ser classificada em relação ao tempo:

- **Contínua**: quando feita continuamente. Ex: nascimentos e óbitos, freqüência de alunos, etc..
- **Periódica**: quando feita em intervalos constante de tempo. Ex: Censo, avaliações mensais, etc..
- **Ocasional**: quando feita extemporaneamente, a fim de atender a uma conjuntura ou a uma emergência. Ex: como é o caso das epidemias.

A **coleta se diz indireta** quando é do conhecimento de outros fenômenos relacionados com os fenômenos estudados. Como por exemplo, podemos citar a pesquisa sobre a mortalidade infantil, que é feita através de dados colhidos por uma coleta direta.

2) - **Critica dos Dados**: Obtidos os dados, eles devem ser cuidadosamente criticados, a procura de possíveis falhas e imperfeições, a fim de não incorrermos em erros grosseiros ou de certo vulto, que possam influir sensivelmente nos resultados.

3) - **Apuração de Dados**: Nada mais é do que a soma e o processamento dos dados obtidos e a mediante critérios e classificação.

4) - **Exposição ou apresentação dos Dados**: Os dados devem ser apresentados sob forma adequada (tabelas ou gráficos) tornando mais fácil o exame daquilo que esta sendo objeto de tratamento estatístico.

5) – Análise dos Resultados: O objetivo da Estatística é tirar conclusões sobre o todo (população) a partir de informações fornecidas por partes representativas do todo (amostra) e tiramos desses resultados conclusões e previsões.

Questionário:

1- O que é Estatística?

2- Cite as fases do método estatístico.

3- Para você o que é coletar dados?

4- Como podem ser apresentados ou expostos os dados?

5- O que é apurar os dados?

3. VARIÁVEIS

Variável é, convencionalmente, o conjunto de resultados possíveis de um fenômeno. As variáveis podem ser:

- Qualitativa – quando seus valores são expresso por atributos: sexo, cor da pele.
- Quantitativa – quando seus valores são expressos em números: salários dos operários, idade dos alunos de uma escola, etc..

As variáveis quantitativas podem ser:

- Variável continua: quando o valor pertence a um intervalo real, teoricamente, pode assumir qualquer valor. Ex: a altura dos empregados, peso das peças fabricadas, peso dos alunos de uma turma, etc.
- Variável discreta: quando o valor pertence a um conjunto finito, um conjunto enumerável. Ex: números de alunos de uma sala, quantidades de produtos fabricados, etc.

Resolva:

1 - Classifique as variáveis em **qualitativas** ou **quantitativas**

1) Universo: Alunos de uma escola. Variável: cor dos cabelos.

2) Universo: casais residentes em uma cidade. Variável: número de filhos.

3) Universo: as jogadas de um dado Variável: o ponto obtido em cada jogada

4) Universo: peças produzidas por certa máquina. Variável: número de peças produzidas por hora.

5) Universo: peças produzidas por certa máquina. Variável: diâmetro externo

2 – Escreva quais das variáveis abaixo são **quantitativas discreta** e quais são **quantitativas contínuas**.

a. População: Alunos de uma Faculdade. Variável: Peso dos alunos

b. População: Estação meteorológica de uma cidade. Variável: precipitação pluviométrica, durante o ano.

c. População: Bolsa de valores de São Paulo. Variável: número de ações negociadas

d. População: funcionários de uma empresa. Variável: salários

e. População: pregos produzidos por uma máquina. Variável: comprimento

f. População: Biblioteca da cidade de Ibaiti. Variável: número de volumes

3.1 População

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos população ou universo estatístico, sobre os quais se deseja obter informações para desenvolver alguma análise.

Ex: Os estudantes, por exemplo, constituem uma população, pois apresentam uma mesma característica: são os que estudam.

3.2 Amostra

É um conjunto reduzido da população, que mantém todas as características essenciais da mesma.

Uma boa amostra deve ser representativa e imparcial, ou seja, deve conter tudo que a população possui.

Há casos, como o de pesquisa sociais, econômicas e de opinião, em que os problemas de amostragem são de extrema complexidade. Mas existem também casos em que os problemas de amostragem são bem fáceis. Como exemplo, podemos citar a retirada de amostras para controle de qualidade dos produtos ou materiais de determinada industria.

3.3 Tipos de amostragem

Amostragem casual ou aleatória simples: é quando todos os elementos de uma população tem a mesma chance (probabilidade) de se selecionado.

Para manter a propriedade deve-se numerar todos os elementos da população, e obter os elementos que comporão a amostra, através de um sorteio ou do auxilio de uma tabela de números aleatórios.

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

- Numeramos os alunos de 01 a 90
- Escrevemos os números, de 01 a 90, em pedaços iguais de papel, e retiramos, um a um, nove números que formarão a amostra. Neste caso 10% da população.

Amostragem proporcional estratificada: Muitas vezes a população se divide em subpopulações (estratos).

É provável que a variável em estudo apresente, de estrato em estrato um comportamento heterogêneo e, dentro de cada estrato, um comportamento homogêneo, convém que o sorteio dos elementos da amostra leve em considerações tais estratos.. A amostragem proporcional estratificada, além de considerar a existência dos estratos, obtém os elementos da amostra proporcional ao número de elementos dos mesmos.

Ex: Supomos no exercício anterior que, dos noventas alunos, 54 sejam meninos e 36 sejam meninas, vamos obter a amostra proporcional estratificada.

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

10% de 54 = 5,4, ou seja, 5 amostras

10% de 36 = 3,6, ou seja, 4 amostras. Total 9 amostras.

Numeramos os alunos de 01 a 90, sendo que do 01 ao 54 correspondem aos meninos e do 55 ao 90 correspondem as meninas e realizamos o sorteio.

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

Ex: Supomos numa rua contendo 1000 prédios, dos quais desejamos obter uma amostra formada por 50 prédios: dividimos 1000 por 50 = 20, escolhemos por sorteio casual um numero de 01 a 20, o qual indicaria o primeiro elemento para a amostra, os demais elementos seriam periodicamente considerados de 20 em 20. Assim se o primeiro elemento sorteado for o cinco, o próximo seria o 25, 45, 65 e assim por diante.

Atividades:

1 – O que é população?

2 – Defina uma amostra

3 – Quais os tipos de amostragem?

Exercícios:

1. Na Feati, quer fazer-se um estudo sobre o peso dos alunos.Supondo-se que há 40 alunos, selecione uma amostra (20%) por:

a. Amostragem aleatória simples (descreva o procedimento)

R:...

b. Amostragem sistemática. (descreva o procedimento)

R:..

2. Numa outra escola, existem 125 crianças na faixa de 7 anos de idade e estejam distribuídas em cinco classes, com quantidades diferentes de alunos, conforme a tabela abaixo, calcule o número de elementos de cada extrato e o total da amostra para 10% da população. Complete na tabela.

Série	População	%	Amostra
A	20		
B	15		
C	35		
D	25		
E	30		
Total	125		

3. Na escola São Leopoldo, para estuda a preferência em relação a refrigerantes, sortearam-se 150 estudantes, entre os 1000 matriculados. Responda:

a) Qual população envolvida na pesquisa

b) Que tipo de amostragem foi utilizado e qual é a amostra considerada?

4. Uma população envolvida em uma pesquisa sobre a incidência de cárie dentária em escolares de W.B. é apresentada no quadro abaixo. Baseando-se nesses dados estratifique uma amostra com 200 elementos. Complete na tabela

Escola	População	%	amostra
A	500		
B	250		
C	440		
D	360		
Total	1550		

5. Na revendedora de doces Enairam, o setor de embalagens tem 20 funcionários, o setor de produção tem 40 funcionários, o setor de revenda tem 25 funcionários e o setor de entrega tem 15 funcionários. Determine uma amostra estratificada com 20 elementos.

6. Em uma cidade com 30 000 habitantes, deseja-se fazer uma pesquisa sobre a preferência por tipo de lazer entre pessoas de 20 anos de idade, levando em conta o sexo a que pertence.

a. Qual a população envolvida na pesquisa?

R:..

b. Supondo que na cidade haja 5 500 homens e 6 000 mulheres com 20 anos, determine uma amostra de 1200 pessoas. Complete na tabela

Sexo	**Número de pessoas**	**Amostra**
Total		

7. Quer fazer-se um estudo que estabeleça a relação entre a faixa salarial e o interesse por sexo, tomando-se um grupo de 1550 pessoas. A tabela abaixo indica o número de pessoas de determinadas faixas salariais. Determine uma amostra com 200 elementos.

Complete na tabela.

Faixa Salarial	**Numero de pessoas**	**Amostra**
Até 3 salários mínimos	776	
De 3 a 6 salários	387	
De 6 a 9 salários	232	
Acima de 9 salários	155	
Total	1550	

4. Distribuições de Freqüência.

Quando se resumem grandes massas de dados brutos, costuma-se freqüentemente distribui-los em classes ou categorias e determinar o número de indivíduos que pertencem a cada uma das classes, denominada freqüência de classe. Um arranjo tabular dos dados por classe, juntamente com as freqüências correspondentes, é denominado, distribuição de freqüência ou tabela de freqüência.

Tabela

Altura de 100 estudantes do sexo Masculino da Facibra

Altura (cm)	Nº. de Estudantes
151 - 158	5
159 - 166	18
167 - 174	42
175 - 182	27
183 - 190	8
Total	100

A primeira classe ou categoria, por exemplo, contém as alturas de 151 cm até 158 cm e é indicada pelo símbolo 151 – 158, como há 5 estudantes cujas alturas pertencem a essa classe, a freqüência que lhe corresponde é cinco.

Quando os elementos não foram numericamente organizados em tabelas, denominamos dados brutos ou tabela primitiva.

A tabela obtida após a ordenação dos dados, que em ordem crescente ou decrescente, recebe o nome de Rol.

4.1 Elementos de uma distribuição de freqüência.

1. **Classe**: Classe de freqüência ou, simplesmente, classes, são intervalos de variação de variável. As classes são representadas simbolicamente por i, sendo i = 1, 2, 3, ...k (onde k é o número total de classe da distribuição). Assim no nosso exemplo o intervalo 159 – 166, define a segunda classe (i = 2). Como a distribuição é formada de cinco classes, então k = 5. Para se determinar o número de classe de uma distribuição de freqüência podemos utilizar a fórmula k = $\sqrt{n}$, para n > 30, onde n é a quantidade de dados. É aconselhável que o número de classe fique entre 5 e 20 classes.
2. **Limites de Classes**: Denominamos limites de classe os extremos de cada classe. O maior número é o limite superior da classe (L_i) e o menor número é o limite inferior da classe (l_i). Na quarta classe temos que l_4 = 175 e L_4 = 182.
3. **Amplitude de um intervalo de classe**: é a medida do intervalo que define a classe. É obtida pela diferença entre o limite superior da classe e o limite inferior da classe, indicada por h_i. $h_i = L_i - l_i$

No exemplo da tabela acima temos que $h_1 = L_1 - l_1 \Leftrightarrow h_1 = 158 - 151 \Leftrightarrow h_1 = 7$

4. **Amplitude total da distribuição:** É a diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo). $AT = L_{(Max.)} - l_{(min.)}$

No nosso exemplo AT= 190 – 151 $\Leftrightarrow$ AT = 39

5. **Amplitude amostral**. É a diferença entre o valor máximo e o valor mínimo da amostra. $AA = x_{(max.)} - x_{(min)}$.

No exemplo temos que AA = 189 – 151 $\Leftrightarrow$ AA = 38

6. **Ponto médio de classe**. É o ponto que divide o intervalo de classe em duas partes. $x_i = \dfrac{l_i + L_i}{2}$, assim o ponto médio da quinta classe no nosso exemplo é:

   $x_i = \dfrac{183+190}{2} \Leftrightarrow x_5 = 124{,}3$

7. **Freqüência simples ou absoluta**: ou simplesmente freqüência de uma classe é o número de observações correspondentes a essa classe ou a esse valor. Assim no nosso exemplo temos que: f_1 = 5, f_2 = 18, f_3 = 42, f_4 = 27 e f_5 = 8. A soma de todas as freqüências é representada pelo símbolo de

somatório $\sum_{i=1}^{k} f_i$, como estamos estudando 100 alunos da Facibra, logo $\sum f_i = 100$

Exercícios.

1. As notas obtidas pelos 50 alunos da Feati na avaliação de matemática foram:

1	2	3	4	5	6	7	7	8	6
2	3	3	4	5	6	6	7	8	8
2	3	4	4	5	6	6	7	8	9
2	3	4	5	5	6	6	7	8	9
2	3	4	5	5	6	7	7	8	9

a. Agora complete a tabela de distribuição de freqüência abaixo.

i	notas	x_i (ponto médio)	f_i (freqüência absoluta)
1	$0 \leftarrow 2$		
2	$2 \leftarrow 4$		
3	$4 \leftarrow 6$		
4	$6 \leftarrow 8$		
5	$8 \leftarrow 10$		
			$\sum f_i = 50$

b. Agora Responda:

1) Qual a amplitude amostral?

2) Qual a amplitude da distribuição?

3) Qual o número de classe da distribuição?

4) Qual o limite inferior da quarta classe?

5) Qual o limite superior da classe 2?

6) Qual a amplitude do segundo intervalo de classe?

c. Complete:

$h_3 =$	n =	$l_1 =$	$L_3 =$	$x_5 =$	$f_4 =$

2. A tabela abaixo apresenta as vendas diárias de um determinado aparelho elétrico, durante um mês na firma Ventania Ltda.

14	12	11	13	14	13
12	14	13	14	11	12
12	14	10	13	15	11
15	13	16	17	14	14

Forme uma distribuição de freqüência na ordem crescente sem intervalos de classe.

3. Em uma fábrica foram testadas 400 lâmpadas, a duração das mesmas aparece na distribuição de freqüência abaixo.

Duração em horas	$300 \longleftarrow 400$	$400 \longleftarrow 500$	$500 \longleftarrow 600$	$600 \longleftarrow 700$	$700 \longleftarrow 800$	$800 \longleftarrow 900$	$900 \longleftarrow 1000$	$1000 \longleftarrow 1100$	Total
Nº. lamapdas	14	46	58	76	68	62	48	28	400

Observando a tabela responda:

a) Qual a amplitude de cada classe?

b) Qual a amplitude total da distribuição?

c) Qual o ponto médio da quinta classe?

d) Qual a porcentagem de lâmpadas com durabilidade máxima de 500 horas?

e) Qual a porcentagem de lâmpadas com durabilidade de 900 horas ou mais

4. As notas obtidas em matemática por 80 estudantes de uma escola X estão relacionadas abaixo:

68	84	75	82	68	90	62	88	76	93	73	79	88	66	83	62
59	85	75	61	65	75	87	74	62	95	78	63	72	79	78	71
94	77	69	74	68	60	96	78	89	61	75	95	60	88	93	75
62	67	97	78	85	76	65	71	75	65	80	73	57	60	82	79
53	74	86	67	73	81	72	63	76	75	85	77	73	78	71	76

a) Organize o rol colocando os dados em ordem crescente.

b) Qual é a menor nota? Qual é a maior nota?

c) Qual é a amplitude total?

d) Qual é a nota do estudante classificado em 10º lugar?

e) Organize os dados em classe considerando 5 como amplitude

f) Faça uma tabela com os pontos médios e a distribuição de freqüência

g) Quantos estudantes receberam nota superior ou igual a 85. Qual a porcentagem?

4.2 Tipos de Freqüência.

Freqüência Simples ou absoluta (f_i): são os valores que realmente representam o número de dados de cada classe. Como já vimos a soma das freqüências simples é igual ao número total dos dados. $\sum f_i = n$

Ex: No exercício 03, na quarta classe a freqüência absoluta é 76

Freqüências Relativas(f_r): são os valores das razões entre as freqüências simples e a freqüência total. $f_r = \frac{f_i}{\sum f_i}$. O propósito das freqüências relativas é o de permitir a analise ou facilitar as comparações.

Ex: No exercício 01, a freqüência relativa na terceira classe é $f_r = \frac{f_i}{\sum f_i}$ =

$f_r = \frac{13}{50}$ $f_r = 0{,}26$

Freqüências acumulada(F_a): é o total das freqüências de todos os valores inferiores ao limite superior do intervalo de uma dada classe. $F_a = f_1 + f_2 + f_3$ ...

Ex: No exercício 03, a freqüência acumulada até a sexta classe é:

Duração em horas	300 ← 400	400 ← 500	500 ← 600	600 ← 700	700 ← 800	800 ← 900	900 ← 1000	1000 ← 1100	Total
Nº lâmpadas	14	46	58	76	68	62	48	28	400

$F_a = 14 + 46 + 58 + 76 + 68 + 62$

$F_a = 324$

Freqüência acumulada relativa (Fra) de uma classe: é a freqüência acumulada da classe dividida pela freqüência total da distribuição. $Fra = \frac{F_i}{\sum f_i}$

Ex: No exercício 03 a freqüência acumulada relativa da classe três é:

$Fra = \frac{118}{400}$ Fra = 0,295

Exercícios

1. Complete a tabela abaixo:

i	classes	f_i	f_r	f_r (%)	F_a	F_{ra}	F_{ra} (%)
1	0 ← 8	4					
2	8 ← 16	10					
3	16 ← 24	14					
4	24 ← 32	9					
5	32 ← 40	3					
		$\sum$ = 40	$\sum$ = 1,0				

2. Dada a distribuição de freqüência ao lado determine:

a) $\sum f_i$

b) as freqüências relativas

x_i	3	4	5	6	7	8
f_i	2	5	12	10	8	3

c) as freqüências acumuladas

d) as freqüências relativas acumuladas

3. A tabela ao lado apresenta uma distribuição de freqüência das áreas de 400 lotes:

Áreas	300 ← 400	400 ← 500	500 ← 600	600 ← 700	700 ← 800	800 ← 900	900 ← 1000	1000 ← 1100	1000 ← 1200
Nº.	24	49	55	71	67	66	40	22	6

Com referencia a essa tabela determine:

a. a amplitude total

b. o limite superior da quinta classe;

c. o limite inferior da oitava classe;

d. o ponto médio da sétima classe

e. a amplitude do intervalo da segunda classe

f. a freqüência da quarta classe;

g. a freqüência relativa da sexta classe

h. o número de lotes cuja área não atinge 700 m^2;

i. o número de lotes cuja área seja igual ou maior que 900 m^2

j. a porcentagem dos lotes cuja área é de 500 m^2, no mínimo, mas inferior a 1 000 m^2;

k. a classe do 72º lote.

4.3 Representação gráfica de uma distribuição.

Uma distribuição de freqüência pode ser representada graficamente pelo histograma de freqüência, pelo polígono de freqüência e pelo polígono de freqüência acumulada.

Histograma

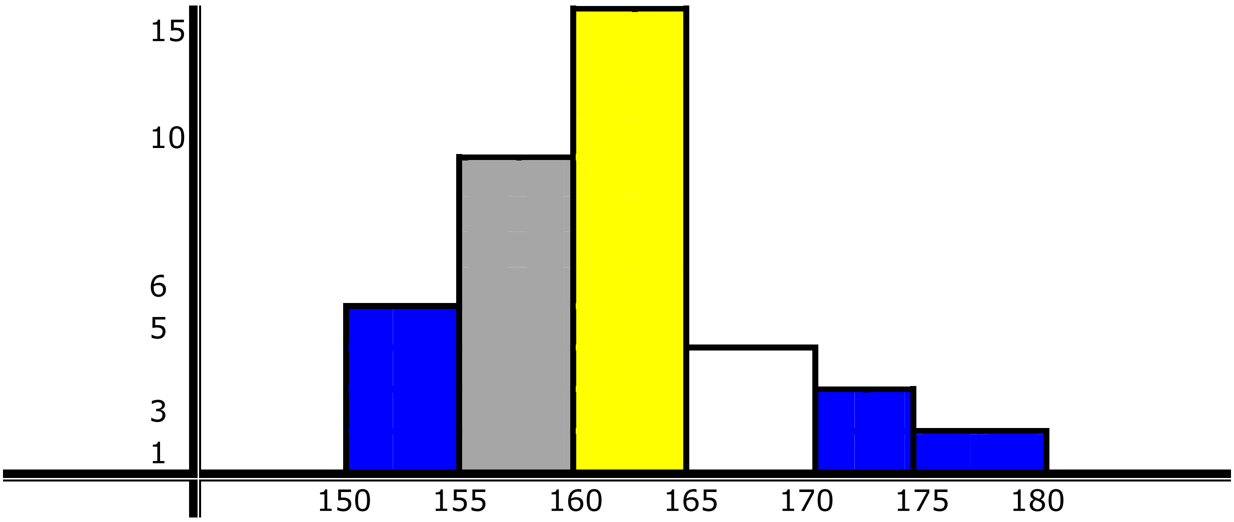
O histograma é formado por retângulos justapostos, sendo o número de retângulos igual ao número de intervalos de classes. A largura de cada retângulo é igual à amplitude do intervalo de classe, enquanto sua altura representa a freqüência do intervalo de classe.

Consideremos a tabela abaixo:

Altura em cm de estudantes da Escola do Sapé				
Classes	Pm	f	Fa	Fr (%)
150 ← 155	152,5	6	6	15
155 ← 160	157,5	10	16	25

160 ← 165	162,5	15	31	38
165 ← 170	167,5	5	36	12
170 ← 175	172,5	3	39	8
175 ← 180	177,5	1	40	2
		n = 40		100

O Histograma ficará assim representado



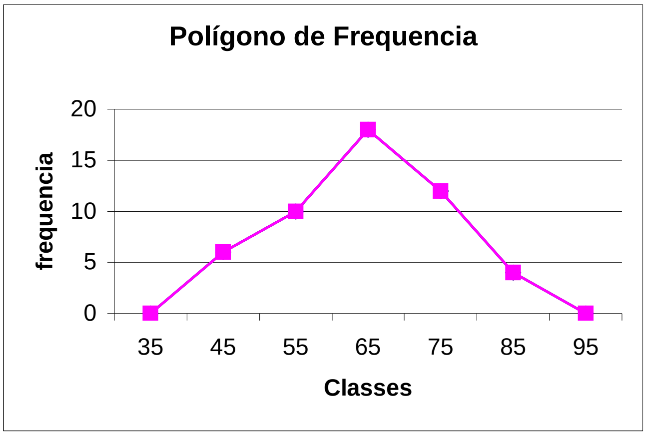
Polígono de Freqüência.

O polígono de freqüência é um gráfico de linha em que cada ponto é obtido considerando-se como valor de x o ponto médio do intervalo de classe e como valor de y a respectiva freqüência do intervalo. Consideramos também uma classe anterior à primeira e uma classe posterior à última, com freqüência nula, ligando todos os pontos temos o polígono de freqüência.

Ex: Observando a tabela de abaixo, construa um polígono de freqüência.

Tabela 02

Peso (Kg)	Ponto médio	Freqüência
40 ← 50	45	6
50 ← 60	55	10
60 ← 70	65	18
70 ← 80	75	12
80 ← 90	85	4



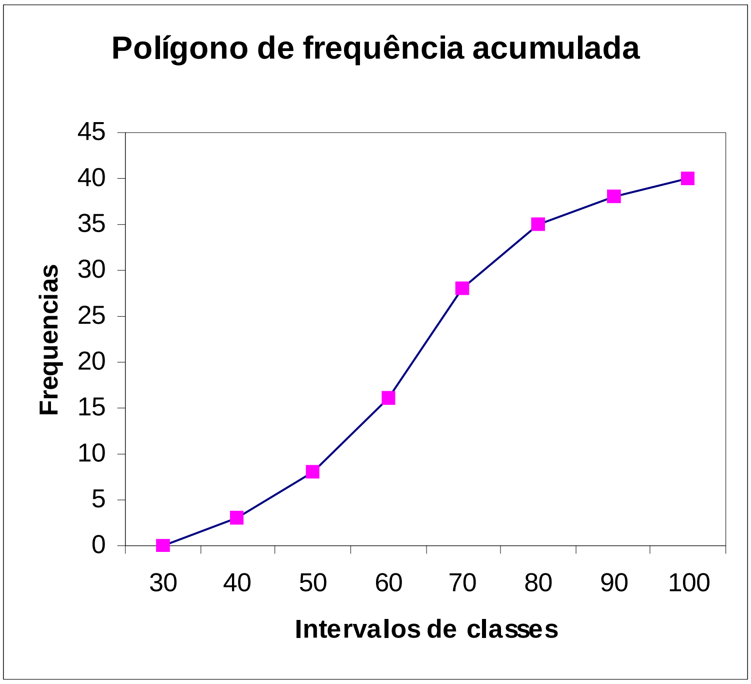
Polígono de freqüência acumulada ou Ogiva

É um gráfico de linha, em que são consideradas as freqüências acumuladas. Para indicarmos as freqüências acumuladas nas ogivas, anotamos a freqüência nula para o limite inferior da primeira da primeira classe e os limites superiores de todas as classes, em ordem crescente, da primeira à última. Devemos considerar como valor de x os limites superiores das classes e como y as respectivas freqüências.

Ex: Na tabela, para a distribuição dada construa um polígono de freqüência acumulada.

Tabela 03

Classe	f_i	F. acumulada
30 ← 40	3	3
40 ← 50	5	8
50 ← 60	8	16
60 ← 70	12	28
70 ← 80	7	35
80 ← 90	3	38
90 ← 100	2	40



Exercícios:

1) Construa o histograma para as distribuições de freqüência a seguir.

Classe	f_i
100 ← 120	4
120 ← 140	6
140 ← 160	8
160 ← 180	15
180 ← 200	20
200 ← 220	9

220 ← 240	2

2) Construa o polígono de freqüência para as distribuições abaixo:

Classe	0 ← 10	10 ← 20	20 ← 30	30 ← 40	40 ← 50	50 ← 60	60 ← 70
f_i	3	8	11	7	15	6	2

3) A tabela a seguir fornece os salários e o número de funcionários de uma empresa por categoria (ajudantes, oficiais, especialistas, mestres e gerentes).

Categoria	A	B	C	D	E
Salários (R$)	4 000,00	2 000,00	1 200,00	800,00	500,00
Nº trabalhadores	5	20	180	300	150

Calcule:

a) a freqüência relativa

b) a freqüência acumulada

c) a freqüência acumulada relativa da classe D

d) Construa um polígono de freqüência acumulada (ogiva)

4) O número de divórcios registrados em 1998 em determinada cidade brasileira esta discriminado na tabela seguinte, em que se identifica a duração dos casamentos.

Período em anos	0 ← 3	3 ← 6	6 ← 9	9 ← 12	12 ← 15	Mais de 15 anos
Número de divórcios	115	54	45	30	18	38

Pede-se para construir um histograma de freqüência, um polígono de freqüência e um polígono de freqüência acumulada.

5. GRÁFICO ESTATÍSTICO

O gráfico estatístico é uma forma de apresentação dos dados estatísticos, com o objetivo de produzir no investigador ou no público em geral, uma impressão mais rápida e viva do fenômeno em estudo.

Acontece uma correspondência entre os termos da série e de determinada figura geométrica, de modo que cada elemento da série seja representada por uma figura proporcional.

A representação gráfica de um fenômeno deve obedecer certos requisitos fundamentais para ser realmente útil:

a) **Simplicidade –** o gráfico deve ser destituído de detalhes de importância secundária, assim como traços desnecessários que possam levar o observador a uma análise morosa ou com erros.

b) **Clareza –** o gráfico deve possibilitar uma correta interpretação dos valores representativos do fenômeno em estudo.

c) **Veracidade** – o gráfico deve expressar a verdade sobre o fenômeno em estudo.

Os principais gráficos podem ser os diagramas, os cartogramas e os pictogramas.

Diagramas

São gráficos geométricos construídos no Sistema de Eixos Cartesianos. Podem ser:

Gráfico em linha ou em curva.

Constitui uma aplicação do processo de representação das funções num sistema de coordenadas cartesianas, utilizando a linha poligonal para representar a série.

Gráfico em Colunas ou em Barras.

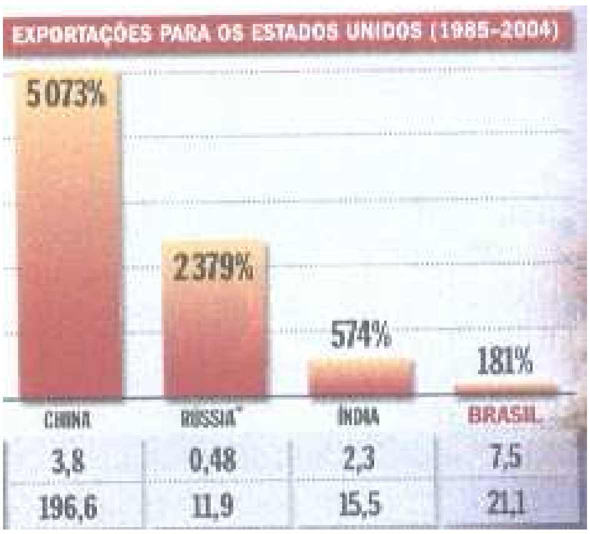
É a representação de uma série por meio de retângulos, dispostos verticalmente (em colunas) ou horizontalmente (em barras).

Quando em colunas, os retângulos têm a mesma base e as alturas são proporcionais aos dados das tabelas.

Quando em barras, os retângulos têm a mesma altura e os comprimentos são proporcionais aos respectivos dados.

Eles Ganham dos Americanos

Exportações em bilhões de dólares



Fontes: Ministério do Desenvolvimento e U.S. Census Bureau

Gráfico em Coluna ou em Barras Múltiplas.

Empregado quando se deseja representar, simultaneamente, dois ou mais fenômenos estudados com o objetivo de estabelecer comparações.

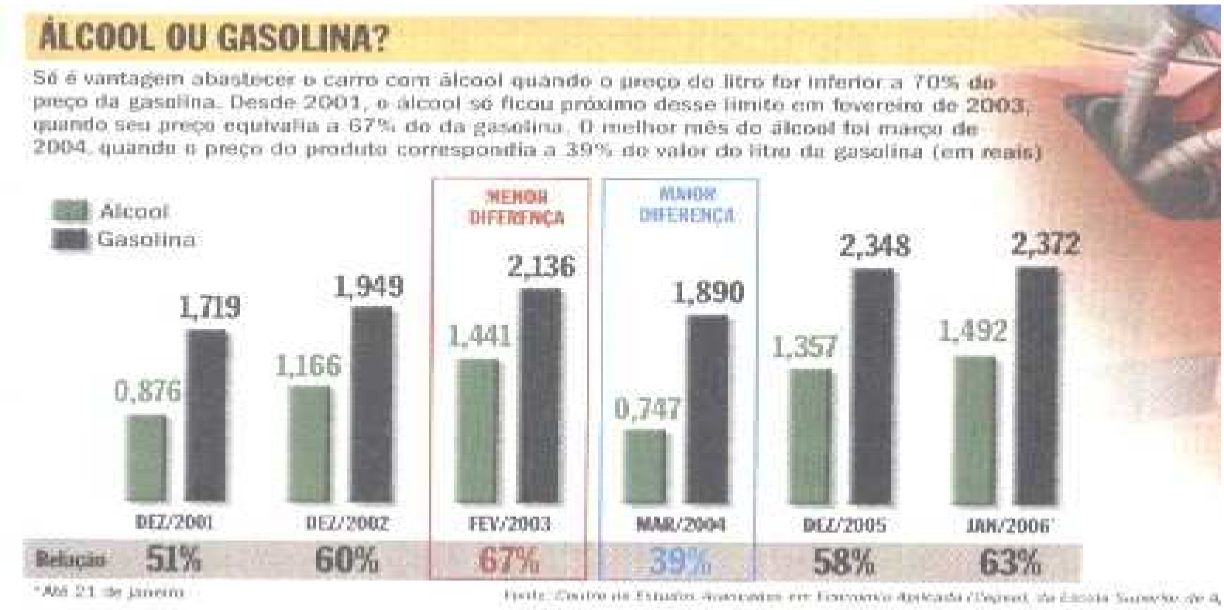
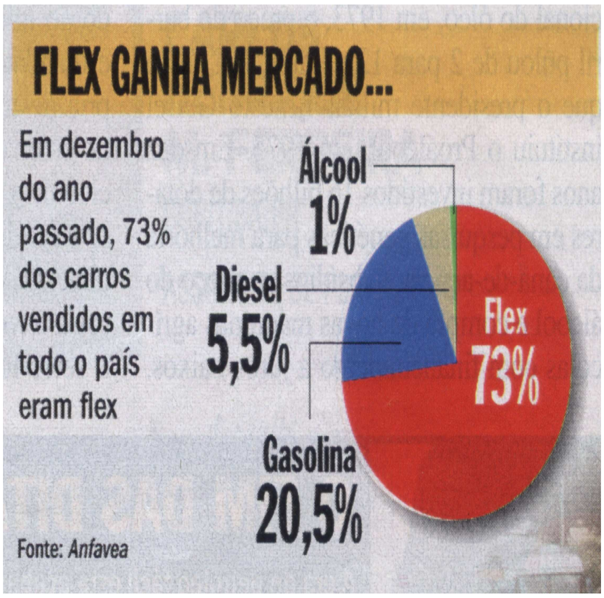
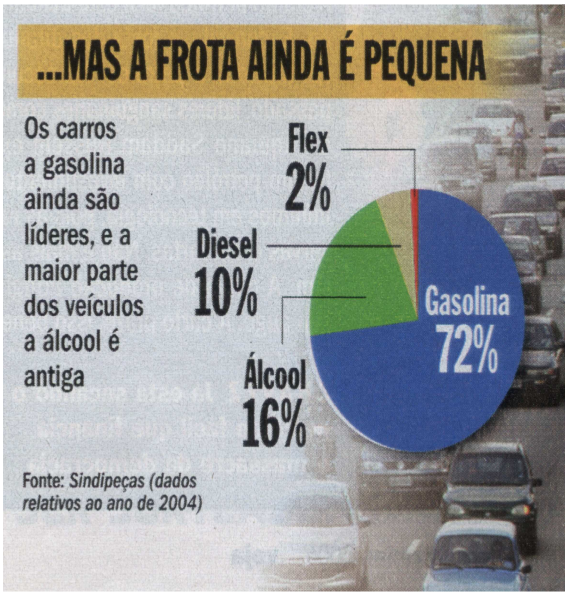


Gráfico de Setores

Este gráfico é construído tomando-se como base o círculo, e é empregado sempre que desejamos ressaltar a participação do dado no total.





Cartograma.

O Cartograma é a representação sobre uma carta geográfica.

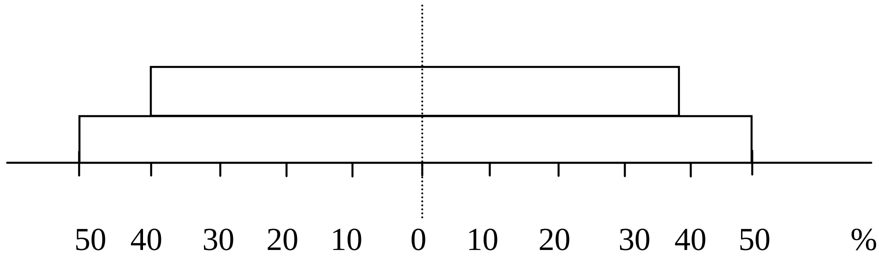
Pictograma.

É um dos processos gráficos que melhor fala ao público, pela sua forma ao mesmo tempo atraente e sugestiva. A representação gráfica consta de figuras.

Pictográfico:

Gráfico da Pirâmide.

sse gráfico é de fácil comunicação e freqüentemente utilizado em estudos demográficos tanto para séries simples quanto em séries conjugadas indicando, por exemplo, a produtividade escolar. Composto por barras sobrepostas que vão definindo a pirâmide, o eixo x orienta o meio exato da pirâmide. Para esse gráfico a série é sem total.



Na construção utiliza-se de modo semelhante a regra prática do gráfico em fita.

Exemplo

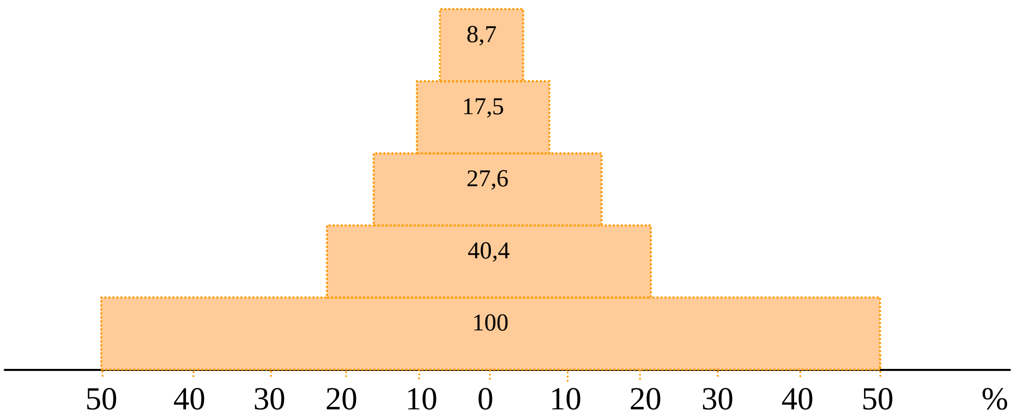
Um estudo de sobrevivência escolar no ensino fundamental realizado no Brasil em 1968 revelou que, para 1000 crianças matriculadas no 1º ano, 404 são promovidas ao 2º ano; 276 são levadas ao 3º; 175 são promovidas ao 4º e apenas 87 concluem a 4ª série.

Estudo de Sobrevivência Escolar, Ensino Fundamental no Brasil em 1968.

Ano	Número de crianças	%	cm
1º	1000	100	10
2º	404	40,4	4,0
3º	276	27,6	2,8
4º	175	17,5	1,8
Conclusão	87	8,7	0,9

Observação: No gráfico abaixo a base do gráfico ficou com 10 cm.

Estudo de Sobrevivência Escolar, Ensino Fundamental no Brasil em 1968.



Nos gráficos apresentados acima foi apresentado um esquema de construção que é útil quando se deseja construir um gráfico manualmente em uma folha de papel ou cartolina.

Os gráficos de colunas, barras, curvas e setores são especialmente úteis para representar dados categóricos.

Para as variáveis quantitativas contínuas representadas em séries de distribuição de freqüência em classe, geralmente são utilizados gráficos como o Histograma e o Polígono de Freqüência.

Outros Exemplos.
PORQUE O DOLAR NÃO PARA DE CAIR

Ele só pensa em namorar:

Exercícios de Gráficos Estatísticos

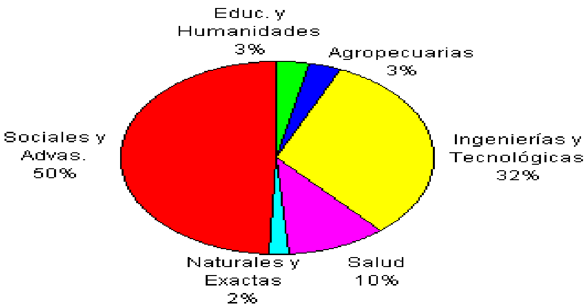
1. Em um grupo de pessoas foi pesquisada a variável “década de nascimento”. Os dados obtidos foram: 60, 70, 70, 70, 60, 80, 70, 70, 80, 60, 80, 90. Construa a tabela de freqüência, o gráfico de barras e de setores correspondentes á pesquisa.

2. As áreas das superfícies dos estados da região sudeste do Brasil, em valores aproximados são: SP, 250.000 Km^2 , Es, 46.000 Km^2 , RJ, 44.000 Km^2 , MG, 590.000 Km^2 . Construa um gráfico de barras registrando essa distribuição.

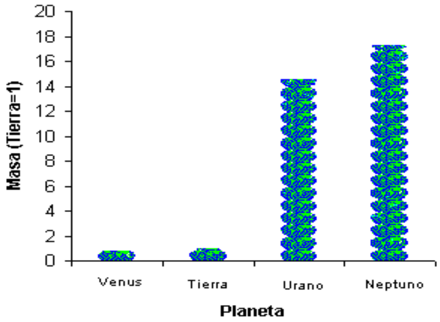
3. Os dados a seguir mostram a venda de revistas em quadrinhos no site www.sebodogibi.com.br no segundo semestre de 2005. Usando os eixos cartesianos construa um gráfico de linhas.

Meses	julho	agosto	setembro	outubro	novembro	Dezembro
Nº. de revistas	350	300	400	450	500	400

4. Na Espanha no ano de 2002 as matriculas nas faculdades locais, por áreas de conhecimento estão representadas no gráfico abaixo: Se na pesquisa foram ouvidos 5.600 estudantes encontre quantos estudantes responderam para cada área do conhecimento.



5. No gráfico abaixo está representado a massa de alguns planetas, tomados como unidade a massa da Terra. De acordo com o gráfico quantas vezes Netuno é maior que a Terra? E urano?



6. Em relação à sua turma, construa um gráfico de setores referente ao sexo masculino ou feminino de cada aluno.

6. MEDIDAS DE POSIÇÃO

O estudo que fizemos sobre distribuição de freqüência, até agora, permite-nos descrever, de modo geral, os grupos dos valores que uma variável pode assumir. Dessa forma, podemos localizar a maior concentração de valores de uma dada distribuição, isto é, se ela se localiza no início, no meio ou no final, ou ainda, se há uma distribuição por igual.

Porém, para ressaltar as tendências características de cada distribuição, isoladamente, ou em confronto com outras, necessitamos introduzir conceitos que se expressam através de números, que nos permitam traduzir essas tendências. Esses conceitos são denominados elementos típicos da distribuição e são as:

- Medidas de posição
- Medidas de variabilidade ou dispersão;

Dentre os elementos típicos, destacamos as medidas de posição – estatísticas que representam uma série de dados orientando-nos quanto à posição da distribuição em relação ao eixo horizontal (eixo das abscissas).

As medidas de posição mais importantes são as de medidas de tendência central, que recebem tal denominação pelo fato de os dados observados tenderem, em geral, a se agrupar em torno de valores centrais. Dentre as medidas de tendência central destacamos:

- A média aritmética;
- A mediana;
- A moda.

6.1 Média Aritmética

Média Aritmética é o quociente da divisão da soma dos valores da variável pelo numero deles.

$\bar{x} = \frac{\sum x_i}{n}$ sendo:

$\bar{x}$ é a média aritmética

x_i é os valores das variável

n é o número de valores.

A média é utilizada quando desejamos obter a medida de posição que possui a maior estabilidade.

Dados não agrupados:

Quando desejamos conhecer a média dos dados não agrupados, determinamos a média aritmética simples.

Exemplo: Sabendo que a produção leiteira da vaca Mimosa, durante uma semana foi de 10, 14, 13, 15, 16, 18 e 12 litros, temos para a produção média da semana:

Dados agrupados sem intervalo de classe.

Neste caso, como as freqüências são números indicadores da intensidade de cada valor na variável, elas funcionam como fatores de ponderação, o que nos leva a calcular a média aritmética ponderada, dada pela fórmula:

$\bar{x} = \frac{\sum xf_i}{\sum f_i}$, sendo:

$\bar{x}$ é a media aritmética

$\sum x_i f_i$ é a soma do produto dos valores da variável pelas respectivas freqüência

$\sum f_i$ é a soma ou o total das freqüências.

O modo mais prático de obtenção da média aritmética ponderada é abrir, na tabela, uma coluna correspondente aos produtos x .f_i.

Exemplo: Considere a distribuição relativa de 34 famílias de quatro filhos, tomando como variável o número de filhos do sexo masculino:

Nº de meninos	0	1	2	3	4
freqüência	2	6	10	12	4

Qual é a media aritmética?

Dados agrupados com intervalos de classe.

Neste caso convencionamos que todos os valores incluídos em um determinado intervalo de classe, coincidem com o seu pondo médio, e determinamos a media aritmética ponderada por meio da fórmula:

$\bar{x} = \dfrac{\sum x_i f_i}{\sum f_i}$, onde x_i é o ponto médio da classe.

Neste caso também, devemos abrir, na tabela uma coluna com os valores dos pontos médios e uma coluna com os produtos de x_i . f_i.

Exemplo:

6) O número de divórcios registrados em 1998 numa determinada cidade brasileira esta discriminado na tabela seguinte, em que se identifica a duração dos casamentos.

Período em anos	$0 \leftarrow 3$	$3 \leftarrow 6$	$6 \leftarrow 9$	$9 \leftarrow 12$	$12 \leftarrow 15$
Número de divórcios	115	54	45	30	18

Calcule o tempo médio de duração dos casamentos.

6.2 Moda

Denominamos moda o valor que ocorre com maior freqüência em uma série de valores.

Desse modo, o salário modal dos empregados de uma indústria é o salário mais comum, isto é, o salário recebido pelo maior número de empregados dessa indústria.

A moda pode não existir e se existir pode não ser única.

Dados não agrupados.

Quando lidamos com valores não agrupados, a moda é facilmente reconhecida, basta procurar o valor que mais se repete.

Ex:

Na série de dados: 7, 8, 9, 10, 10, 10, 11, 12, 13, 15 tem moda igual a 10,

Já na série 3, 5, 8, 10, 12, 13 não apresenta moda (amodal)

E na série 2, 3, 4, 4, 4, 5, 6, 7, 7, 7, 8, 9 tem duas moda: 4 e 7 (bimodal)

Dados agrupados sem intervalos de classe.

Uma vez agrupados os dados, é possível determinar imediatamente a moda: basta fixar o valor da variável de maior freqüência.

Ex:

x_i	3	4	5	6	7	8
f_i	2	5	12	10	8	3

A moda neste caso é o que possui maior freqüência, logo Mo = 5

Dados agrupados com intervalos de classe.

A classe que apresenta maior freqüência é denominada classe modal.

Pela definição, podemos afirmar que a moda, neste caso, é o valor dominante que está compreendido entre os limites da classe modal. O método mais simples para o cálculo da moda neste caso é tomar o ponto médio da classe modal. Damos a esse valor a denominação de moda bruta. Ex:

Duração em horas	300←400	400←500	500←600	600←700	700←800	800←900	900←1000	1000
Nº lâmpadas	14	46	58	76	68	62	48	28

A classe que apresenta maior freqüência é a classe 4, e seu ponto médio é 650, logo a moda é 650.

6.3 Mediana

Mediana é um conjunto de valores, ordenados segundo uma ordem de grandeza, é o valor situado de tal forma no conjunto que o separa em dois subconjuntos de mesmo número de elementos.

Dados não agrupados:

Dada uma série de dados, como, por exemplo: 2, 5, 6, 9, 10, 13, 15, 16, 18 a mediana é 10, já que, nessa série, há quatro elemento acima dele e quatro abaixo.

Se, porém a série tiver um número par de elementos, a mediana será por definição, o ponto médio dos números compreendidos entre os dois valores centrais da série. Assim a série de valores 2, 6, 7, 10, 12, 13, 18, 21 tem para a mediana o ponto médio entre 10 e 12, logo a mediana é 11.

Verificamos que, estando ordenado os valores de uma série e sendo n o número de elementos da série, o valor mediano será:

- O termo de ordem $\frac{n+1}{2}$ se n for impar;
- A média aritmética dos termos de ordem $\frac{n}{2}$ + 1 se n for par.

Dados agrupados sem intervalos de classe.

Neste caso, é o bastante identificar a freqüência acumulada imediatamente superior a metade da soma das freqüências. A mediana será aquele valor da variável que corresponde a tal freqüência acumulada.

Ex: Determine a mediana dessa distribuição, em que temos as diárias dos operários de uma fábrica.

Diária (R$)	Nº de operários	Freqüência acumulada
200,00	5	5
250,00	8	13
300,00	4	17
350,00	1	18

Como a soma das freqüências é 18 e sua metade é 9, a menor freqüência acumulada que supera esse valor é 13, que corresponde ao valor R$ 250,00, sendo esse o valor mediano, logo Md = 250,00

Dados agrupados com intervalos de classe.

Neste caso o problema consiste em determinar o ponto do intervalo em que está compreendida a mediana. Para tanto, temos inicialmente que determinar a classe na qual se acha a mediana – classe mediana – Tal classe será aquela correspondente à freqüência cumulada imediatamente superior a metade da soma das freqüências. Feito isso um problema de interpolação resolve a questão.

Ex:

Classe	f_i	Fa
0 ← 10	3	3
10 ← 20	8	11
20 ← 30	11	22
30 ← 40	7	29
40 ← 50	15	44
50 ← 60	6	50
60 ← 70	2	52

A soma das freqüências é 52 logo a metade é 26.

Portanto a classe mediana é 29. Como há 29 valores incluídos nas 4 primeiras classes da distribuição e como pretendemos determinar o valor que ocupa o 26º, vemos que esse valor esta na 4ª classe cuja freqüência é 7. Logo a mediana esta entre 30 e 40. Os 7 elementos estão na amplitude 10.

A diferença entre a metade da soma das freqüências e a freqüência acumulada da classe imediatamente anterior a quarta classe é 26 – 22 = 4.

A distancia entre 30 e a mediana chamaremos de x. na distancia x, temos 04 elementos. Na amplitude 10, temos 07 elementos. Podemos armar a proporção:

$\frac{x}{4} = \frac{10}{7}$ x = 5,71, portanto a mediana será Md = 30 + 5,71, Md = 35,71

Na prática seguimos os seguintes passos:

- Determinamos as freqüências acumuladas;
- Calculamos a metade da soma da freqüência; $\frac{\sum f_i}{2}$
- Marcamos a classe correspondente a freqüência acumulada imediatamente superior a metade da soma das freqüências e, em seguida empregamos a fórmula.

$$Md = l* + \frac{\left[\frac{\sum f_i}{2} - Fa(ant)\right]}{f*} \; . \; h*$$ onde:

l* é o limite inferior da classe mediana

Fa (ant) é a freqüência acumulada da classe anterior à classe mediana

f* é a freqüência simples da classe mediana

h* é a amplitude do intervalo da classe mediana.

Aplicando a formula para o exemplo acima fica:

EXERCÍCIOS

1) O Brasil extrai petróleo de sete bacias submarinas, num total de 91 plataformas. As plataformas estão assim distribuídas. Bacia de Santos (2), Bacia de Campos (27), Bacia do Espírito Santo (3), Bacia do Recôncavo (7), Bacia Sergipe/Alagoas (25), Bacia Potiguar (18) e Bacia do Ceará (9). Com esses dados calcule a média aritmética, a mediana e a moda.

2) Pesquisa realizada recentemente revela que nos últimos anos o consumo de cigarros vem crescendo entre as mulheres. Parte desse estudo permite a montagem de uma tabela de freqüências, que relaciona a quantidade de cigarros consumidos diariamente, entre 1000 mulheres fumantes. Calcular a média aritmética, a mediana e a moda.

Cigarros	f_i	Fa
15 ← 20	150	
20 ← 25	300	
25 ← 30	250	
30 ← 35	200	
35 ← 40	100	

3) Desejando lançar uma nova pasta dental, uma indústria pesquisou sobre os valores cobrados por nove marcas concorrentes e obteve os seguintes valores em reais: 1,12; 1,00; 1,07; 1,18; 1,60; 1,90; 0,90; 2,02; 1,70; 1,12. Calcule a média aritmética, a moda e a mediana desses valores.

4) Verificando a durabilidade de 48 pilhas elétricas, quando utilizadas sem interrupção, obtivemos os seguintes dados, em quantidades de dias:

notas	x_i (ponto médio)	f_i (freqüência absoluta)	Freqüência acumulada
0 ← 2		0	
2 ← 4		16	
4 ← 6		18	
6 ← 8		12	
8 ← 10		2	
total		48	

5) Os salários-hora de cinco funcionários de uma empresa são: R$ 75,00, R$ 90,00, R$ 83,00; R$ 142,00 e R$ 88,00, determine:

a) a média dos salários-hora.

b) o salário-hora mediano

6) Numa agencia bancária fez, no último dia do mês, uma pesquisa sobre os depósitos em cadernetas de poupança e seus respectivos valores em reais, desprezando-se os centavos e obtiveram os seguintes dados: Calcule a média aritmética, a mediana e a moda.

Valores	150	200	300	350	500	1000
Nº de depósitos	21	17	12	8	5	2

7) Determinada editora pesquisou o número de páginas das revistas mais vendidas em uma cidade e obteve a seguinte tabela: Calcule a média aritmética, a mediana e a moda.

Revistas	Nº páginas	F acum.
A	62	
B	90	
C	88	
D	92	
E	110	
F	86	
Total		

8) Em um supermercado, a reposição de pacotes de arroz, nesta segunda feira, permitiu a construção da seguinte tabela.

Marca do arroz	A	B	C	D	E
Quant. pacotes	120	60	280	200	140

Calcule a média aritmética, a mediana e a moda.

9) Uma distribuidora pesquisou o consumo de refrigerantes entre diferentes faixas etárias, para melhor direcionar sua campanha publicitária. Baseado nos dados, calcule a média aritmética, a mediana e a moda.

notas	f_i (freqüência absoluta)	Freqüência acumulada	x_i (ponto médio)	.
10 ← 14	60			
14 ← 18	100			
18 ← 22	130			
22← 26	90			
26 ← 30	20			
total	400			

7. MEDIDAS DE DISPERSÃO OU DE VARIABILIDADE.

Vimos anteriormente que um conjunto de valores pode ser convenientemente sintetizado, por meio de procedimentos matemáticos, em poucos valores representativos – média, mediana e moda. Tais valores podem servir de comparação para dar a posição de qualquer elemento do conjunto.

No entanto, quando se trata de interpretar dados estatísticos, mesmo aqueles já convenientemente simplificados, são necessários ter-se uma idéia retrospectiva de como se apresentam esses mesmos dados nas tabelas. Assim, não é o bastante dar uma das medidas de posição para caracterizar perfeitamente um conjunto de valores, pois, mesmo sabendo, por exemplo, que a temperatura média de duas cidades é a mesma, e igual a 24º, ainda assim somos levados a pensar a respeito do clima dessas cidades. Em uma delas poderá a temperatura variar entre limites de muito calor e de muito frio e haver, ainda, uma temperatura média de 24º. A outra poderá ter uma variação pequena de temperatura e possuir, portanto, no que se refere à temperatura, um clima mais favorável.

Vemos, então, que a média – ainda que considere como um número que tem a faculdade de representar uma série de valores – não pode por si mesmo, destacar o grau de homogeneidade ou heterogeneidade que existe entre os valores que compõem o conjunto.

Consideremos os seguintes conjuntos das variáveis x, y, e z.

X: 70, 70, 70, 70, 70.

Y: 68, 69, 70, 71, 72.

Z: 5, 15, 50, 120, 160.

Calculando a média aritmética de cada um desses conjuntos, obtemos 70.

Vemos, então, que os três conjuntos apresentam a mesma média aritmética, entretanto, é fácil notar que o conjunto X é mais homogêneo que os conjuntos Y e Z, já que todos os valores são iguais a média.

O conjunto Y é mais homogêneo que o conjunto Z, pois há menor diversificação entre cada um de seus valores e a média representativa.

Chamando de dispersão ou variabilidade a maior ou menor diversificação dos valores de uma variável em torno de um número de tendência central tomada como ponto de comparação, podemos dizer que o conjunto X apresenta dispersão ou variabilidade nula e que o conjunto Y apresenta uma dispersão ou variabilidade menor que o conjunto Z.

Portanto, para qualificar os valores de uma dada variável, ressaltando a maior ou menor dispersão ou variabilidade entre esses valores e a sua medida de posição, a Estatística recorre as medidas de dispersão ou de variabilidade.

Dessas medidas, estudaremos o Desvio absoluto Médio, a Variância, o Desvio Padrão.

7.1 Desvio Absoluto Médio

Dados não agrupados e dados agrupados sem intervalo de classe.

O Desvio Absoluto Médio mede o afastamento médio dos elementos da amostra em relação à média aritmética.

Para determinar quanto cada nota está afastada da média aritmética, basta efetuar a diferença entre a nota e a média. ($x - \bar{x}$). A essa diferença chamamos de **desvio da nota.**

Exemplo.

Em uma prova de Estatística Leonor obteve 8 de média aritmética e suas notas foram: 8,5; 9,5; 8,0; 7,0; e 7,0. Os desvio são:

8,5 – 8,0 = 0,5 → significa que a nota está 0,5 acima da média.

9,5 – 8,0 = 1,5 → significa que a nota está 1,5 acima da média.

8,0 – 8,0 = 0,0 → significa que a nota coincide com a média.

7,0 – 8,0 = -1,0 → significa que a nota está 1,0 abaixo da média.

O módulo de cada um desses desvios é chamado de desvio absoluto da nota correspondente. O desvio absoluto da nota 8,5 é | 0,5 | = 0,5; o desvio absoluto da nota 7,0 é | - 1,0 | = 1,0.

A média aritmética entre esses desvios absolutos é chamada de **Desvio Absoluto Médio**, que se indica por Dam.

Dam = $\frac{0,5+1,5+0+1,0+1,0}{5}=0,8$

Se for comparar com algum outro aluno que, por exemplo, tivesse Dam = 1,2 poderíamos concluir que Leonor teve um desempenho mais regular que o outro.

Dados agrupados com intervalo de classe.

Neste caso devemos substituir o valor da variável x pelo ponto médio da classe. Simbolicamente temos: Dam = $\frac{\sum |x-\bar{x}|}{n}$.

7.2 Variância.

A análise de variância é um teste estatístico amplamente difundido entre os analistas, e visa fundamentalmente verificar se existe uma diferença significativa entre as médias e se os fatores exercem influência em alguma variável dependente. A principal aplicação da variância é a comparação de médias oriundas de grupos diferentes, também chamados tratamentos, como por exemplo médias históricas de questões de satisfação, empresas que operam simultaneamente com diferentes rendimentos, entre muitas outras aplicações.

Como foi dito, a Variância (s^2) de um experimento é um número que tem relação com a média. Para obtê-la devemos:

- Calcular a média aritmética dos elementos da série;
- Obter a diferença entre cada termo e a média obtida: (x - $\bar{x}$), em que $\bar{x}$ é a média aritmética.
- Elevar essas diferenças ao quadrado e multiplicar por sua freqüência: $(x_i - \bar{x})^2 . f_i$;
- Adicionar todos os produtos dos quadrado das diferenças obtidas pelas freqüências: $\sum(x_i - \bar{x})^2 . f_i$
- Dividir essa soma pelo número de elementos da série:
- $s^2 = \dfrac{\sum(x_i - \bar{x})^2 . f_i}{n}$ para dados agrupados na tabela de distribuição de freqüência e $s^2 = \dfrac{\sum(x_i - \bar{x})^2}{n}$ para dados não agrupados .

A esse número s^2 damos o nome de variância da série.

Exemplo:

Calcule a variância para os seguintes exemplos:

1º) Dados não agrupados.

Revistas	Nº páginas	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
A	62		
B	90		
C	88		
D	92		
E	110		
F	86		
Total			

2º) Dados agrupados.

Duração em horas	$300 \leftarrow 400$	$400 \leftarrow 500$	$500 \leftarrow 600$	$600 \leftarrow 700$	$700 \leftarrow 800$	$800 \leftarrow 900$	$900 \leftarrow 1000$	Tota
Nº lâmpadas	14	46	58	76	68	62	48	

7.3 Desvio Padrão.

Desvio padrão é definido como a raiz quadrada positiva da variância. s = $\sqrt{s^2}$ assim s = $\sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$

Dados não agrupados.

Tomemos como exemplo op conjunto de valores da variável x: 40, 45, 48, 52, 54, 62, 70.

O modo mais prático para se obter o desvio padrão é formar uma tabela com duas colunas; uma para os valores de x_i e outra para os valores de x_i^2. Veja:

x_i	x_i^2

Dados agrupados sem intervalo de classe.

Como neste caso, temos a presença de freqüência, devemos leva-las em consideração, resultando a fórmula: s = $\sqrt{\frac{\sum f_i x_i^2}{n} - \left(\frac{\sum f_i x_i}{n}\right)^2}$.

Neste caso o modo mais prático para se obter o desvio padrão é abrir, na tabela dada, uma coluna para os produtos $f_i x_i$ e outra para $f_i x_i^2$. Lembrando que para se obter $f_i x_i^2$ basta multiplicar cada $f_i x_i$ pelo seu respectivo x_i.

Exemplo:

x_i	f_i	$f_i x_i$	$f_i x_i^2$
1	5		
2	4		
3	6		
4	3		
5	2		
Total			

Dados agrupados em intervalo de classe.

Neste caso devemos proceder como no exemplo anterior, somente substituindo o valor de x_i pelo ponto médio da classe.

Exemplo:

Classe	f_i	x_i	$f_i x_i$	$f_i x_i^2$

0 ← 10	3			
10 ← 20	8			
20 ← 30	11			
30 ← 40	7			
40 ← 50	15			
50 ← 60	6			
60 ← 70	2			
Total				

Exercícios

1) A distribuição de salários de uma empresa é dada na seguinte tabela:

Salários	Nº Funcionários
500,00	10
1 000,00	5
1 500,00	1
2 000,00	10
5 000,00	4
10 500,00	1
Total	31

a. Qual a amplitude total, a variância, o desvio padrão e o coeficiente de variação.

b. Supondo que sejam contratados dois novos funcionários de R$ 2 000,00 reais cada. A variância da nova distribuição de salários ficará menor, igual ou maior que a anterior?

2) Dois atiradores x e y obtiveram numa série de vinte tiros, num alvo com valores e obtiveram os resultados da tabela:

atiradores	Resultados				
	50	30	20	10	0
X	4	6	5	4	1
y	6	3	5	3	3

a) Qual a média de pontos de cada um dos atiradores?

b) Compare os desvios padrão dos dois atiradores de cada uma das séries e decida qual é o atirador de melhor desempenho.

c) Calcule o coeficiente de variação dos dois atiradores.

3) As séries A, B e C têm a mesma média, porém amplitude, variância e desvios padrão diferentes. Dada a tabela abaixo, calcule cada uma dessas medidas para essas séries.

A	10	10	10	10	10
B	10	8	12	8	12
C	12	10	6	10	12

4) Uma distribuidora pesquisou o consumo de refrigerantes entre diferentes faixas etárias, para melhor direcionar sua campanha publicitária. Baseado nos dados calcule a média dos consumidores, a amplitude, a variância e o desvio padrão.

notas	f_i (freqüência absoluta)
10 ← 14	60
14 ← 18	100
18 ← 22	130
22← 26	90
26 ← 30	20
total	400

5) Uma pesquisa realizada pela secretária de Saúde de uma cidade, visando conhecer os hábitos de higiene bucal da população, identificou num de seus itens o tipo de creme dental mais consumido e tabelaram os seguintes dados:

Tipo de creme	flúor	Bicarbonato de sódio	Menta e flúor	Flúor e Bicarbonato
Nº. de consumidores	80	20	60	40

Calcule:

a) a média aritmética desses valores;

b) o desvio padrão

c) a variância.

6) Calcule o desvio padrão para os dados da tabela abaixo:

Classe	f_i	X_i (ponto médio	f_ix_i	$F_ix_i^2$
150 ← 154	4			
154 ← 158	9			
158 ← 162	11			
162 ← 166	8			
166 ← 170	5			
170 ← 174	3			
Total	40			

7) Em uma turma de aluno, verificou-se através da análise das notas de 15 alunos, o seguinte desempenho: Calcule a média e o desvio padrão para os valores encontrados.

Alunos	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
Nota	43	45	90	60	80	67	75	10	75	63	80	55	97	93	75