

ESTATÍSTICA Computacional

Prof. Dr. José Luiz Cabral da Silva Júnior

E-mail: joseluizcabralturmas@gmail.com

ESTATÍSTICA COMPUTACIONAL

AULA 2

ALGUNS CONCEITOS IMPORTANTES

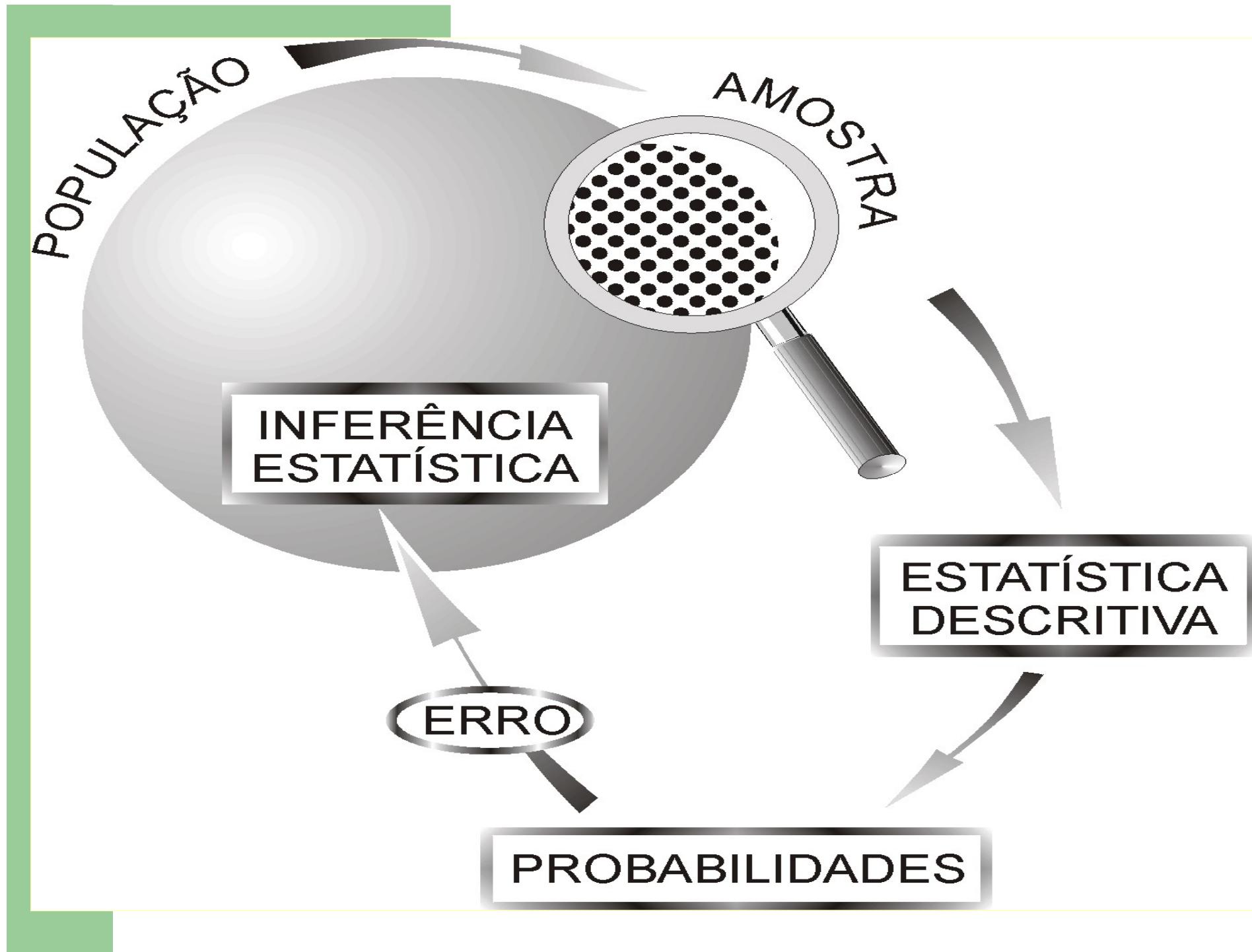
REVISÃO

- **Método:** é um conjunto de meios dispostos convenientemente para se chegar a um fim que se deseja.
- **Método experimental:** consiste em manter constantes todas as causas (fatores), menos uma, e variar esta causa de modo que o pesquisador possa descobrir seus efeitos, caso existam.
- **Método estatístico:** diante da impossibilidade de manter as causas constantes, admite todas essas causas presentes variando-as, registrando essas variações e procurando determinar, no resultado final, que influências cabem a cada uma delas.

O que é Estatística?

A Estatística basicamente se divide em 3 partes, a saber:

- *Estatística Descritiva*: Essa parte da Estatística utiliza números para descrever fatos. Compreende a coleta, a organização, o resumo e, em geral, a simplificação de informações que podem ser muito complexas.
- *Probabilidade*: É utilizada para analisar situações que envolvem o acaso.
- *Inferência*: Diz respeito a coleta, redução, análise e interpretação de dados amostrais, a partir do que, tira-se conclusões sobre a população na qual os dados (amostra) foram obtidos.

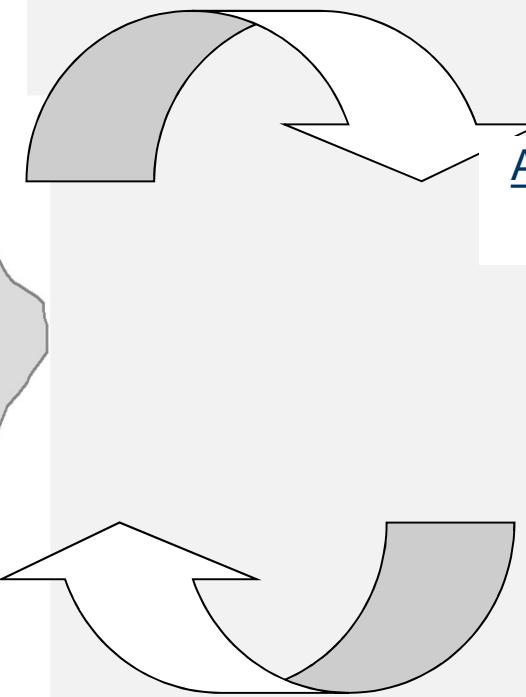


A Estatística

POPULAÇÃO: todos os possíveis consumidores



amostragem

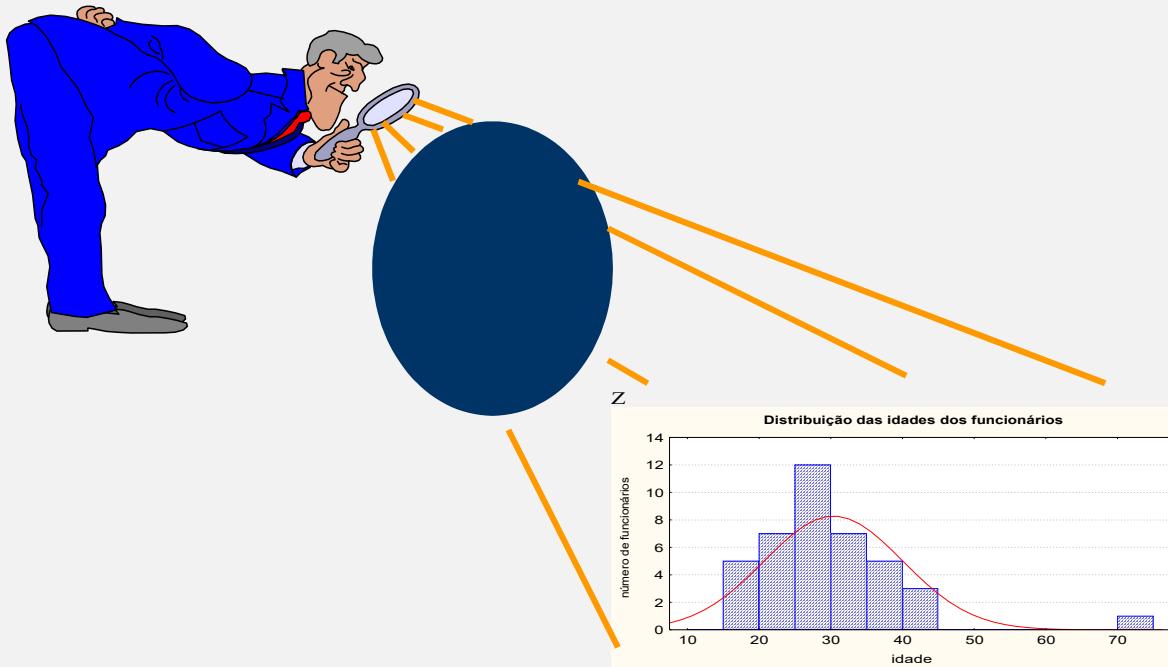


Amostra: um subconjunto dos consumidores

inferência

A essência de uma análise estatística é tirar conclusões sobre uma população, ou universo, com base em uma amostra de observações.

A Estatística



A estatística envolve técnicas para coletar, organizar, descrever, analisar e interpretar dados, ou provenientes de experimentos, ou vindos de estudos observacionais.

VARIÁVEIS

QUALITATIVAS (são atributos)

Exemplos:

- sexo
- religião
- naturalidade
- cor dos olhos
- faixa etária

QUANTITATIVAS (são numéricas)

DISCRETAS

Exemplos:

- quantidade de estudantes em uma disciplina
- quantidade de cômodos em uma residência

CONTÍNUAS

Exemplos:

- tempo de vôo entre duas cidades
- duração da bateria de telefone celular

TEORIA DA AMOSTRAGEM

É UMA TÉCNICA PARA ESCOLHER AMOSTRAS QUE PERMITEM GARANTIR A CASUALIDADE NA ESCOLHA. AS CATEGORIAS MAIS UTILIZADAS SÃO: AMOSTRAGEM CASUAL OU SIMPLES; AMOSTRAGEM PROPORCIONAL ESTRATIFICADA E AMOSTRAGEM SISTEMÁTICA.

Amostragem Simples ao Acaso

Também conhecida por amostragem ocasional, acidental, casual, randômica, etc., a amostragem simples ao acaso destaca-se por ser um processo de seleção bastante fácil e muito usado. Este tipo de amostragem é equivalente **a um sorteio lotérico**. Neste processo, todos os elementos da população têm igual probabilidade de serem escolhidos, não só antes de ser iniciado, como também até completar-se o processo de coleta.

Procedimento:

- 1 - Numerar todos os elementos da população.
- 2 - Efetuar sucessivos sorteios com reposição até completar o tamanho da amostra.

Amostragem Sistemática

Trata-se de uma variação da amostragem simples ao acaso, muito conveniente quando a população está naturalmente ordenada, pois quando os elementos da população já se acham ordenados, não há necessidade de construir o sistema de referência.

Procedimento:

1 - Seja N o tamanho da população e n o tamanho da amostra. Então, calcula-se o intervalo de amostragem N/n ou o inteiro mais próximo "a".

2 - Sortear um número x entre 1 e "a".

3 - Formar a amostra dos elementos correspondentes ao números:
 $x; x + a; x + 2a; \dots; x + (n-1)a$.

Exemplo: Seja $N = 95$ $n = 10$. Então, $95/10 = 9,5$, ou $a = 10$.

Sorteia-se um número de 1 a 10. Seja 4 ($x = 4$) o número sorteado. Logo, os elementos numerados por 4; 14; 24; ... serão os componentes da amostra.

Amostragem Estratificada

A amostragem estratificada pressupõe a divisão da população em subgrupos (estratos) de itens similares, procedendo-se então à amostragem em cada subgrupo.

Estratificar uma população é dividi-la em L subpopulações denominadas estratos, tais que: , onde os extratos são mutuamente exclusivos. Após a determinação dos extratos, seleciona-se uma amostra aleatória de cada subpopulação.

Exemplo: Obter uma amostra para a pesquisa da estatura de 90 alunos de uma escola, onde 54 sejam meninos e 36 sejam meninas. Teremos dois estratos (sexo masculino e sexo feminino) e queremos uma amostra de 20% da população, assim:

Masculino = logo amostra: 11

Feminino = logo amostra: 07
total: 18

Numeramos os alunos de 01 a 90, sendo que de 01 a 54 correspondem meninos e de 55 a 90, meninas. Efetuamos os sorteios até atingirmos 11 meninos e 07 meninas.

Outros tipos de Amostragem

Amostragem por conglomerados : Quando é possível identificar dentro da população subgrupos que representam uma miniatura da população, podemos escolher um desses subgrupos como a população e efetuarmos sorteios entre seus elementos. Na verdade, cada conglomerado pode ser encarado como uma minipopulação.

Amostragem múltipla: É um tipo de amostragem bastante utilizado em processos de Controle de Qualidade, nele retira-se uma primeira amostra e, conforme o resultado obtido, tiramos uma segunda amostra, uma terceira, etc.

TABELA

É UM QUADRO QUE RESUME UM CONJUNTO DE INFORMAÇÕES. A MANEIRA DE APRESENTAR ESSAS INFORMAÇÕES PODEM SER ESTABELECIDAS PELOS SEUS ELEMENTOS, CONFORME DISTRIBUIÇÃO A SEGUIR:

ELEMENTOS DE UMA TABELA

- CORPO;
- CABEÇALHO;
- COLUNA INDICADORA
- LINHAS;
- CASA OU CÉLULA;
- TÍTULO

TÍTULO

CABEÇALHO

Levantamento Sistemático da Produção Agrícola

1 - Área

Confronto das Estimativas Maio/Junho- Brasil

Produtos Agrícolas	Área (ha)		
	Maio	Junho	Variacão %
Total	61 249 539	61 304 436	0,1
Milho (em grão) - Total	12 871 064	12 861 250	-0,1
Milho (em grão) 1ª safra	7 640 048	7 767 071	-0,9
Milho (em grão) 2ª safra	5 031 016	5 094 179	1,3
Sorgo (em grão)	23 317 258	23 330 979	0,1
Sorgo (em grão)	647 475	664 230	2,6
Tribo (em grão)	2 166 749	2 201 261	1,6
Triticale (em grão)	63 836	62 045	-2,8

FONTE - Grupo de Coordenação de Estatísticas Agropecuárias - GCEA/IBGE, DPE, COAGRO - Levantamento Sistemático da Produção Agrícola - Junho 2010.

FONTE

LINHA

CÉLULA

COLUNA

SÉRIES ESTATÍSTICAS

SÃO AS TABELAS QUE APRESENTAM A DISTRIBUIÇÃO DE UM CONJUNTO DE DADOS ESTATÍSTICOS EM FUNÇÃO DA ÉPOCA, DO LOCAL OU DA ESPÉCIE.

HISTÓRICA – CRONOLÓGICA OU TEMPORAL;

GEOGRÁFICAS – TERRITORIAIS OU DE LOCALIZAÇÃO;

ESPECÍFICAS OU CATEGÓRICAS

SÉRIES HISTÓRICAS

ESTAS SÉRIES TAMBÉM CHAMADAS DE CRONOLÓGICAS OU TEMPORAIS APRESENTAM OS VALORES DE UMA VARIÁVEL DE UM CERTO LOCAL EM INTERVALOS DE TEMPO.

PRODUÇÃO DE CANA DE AÇUCAR

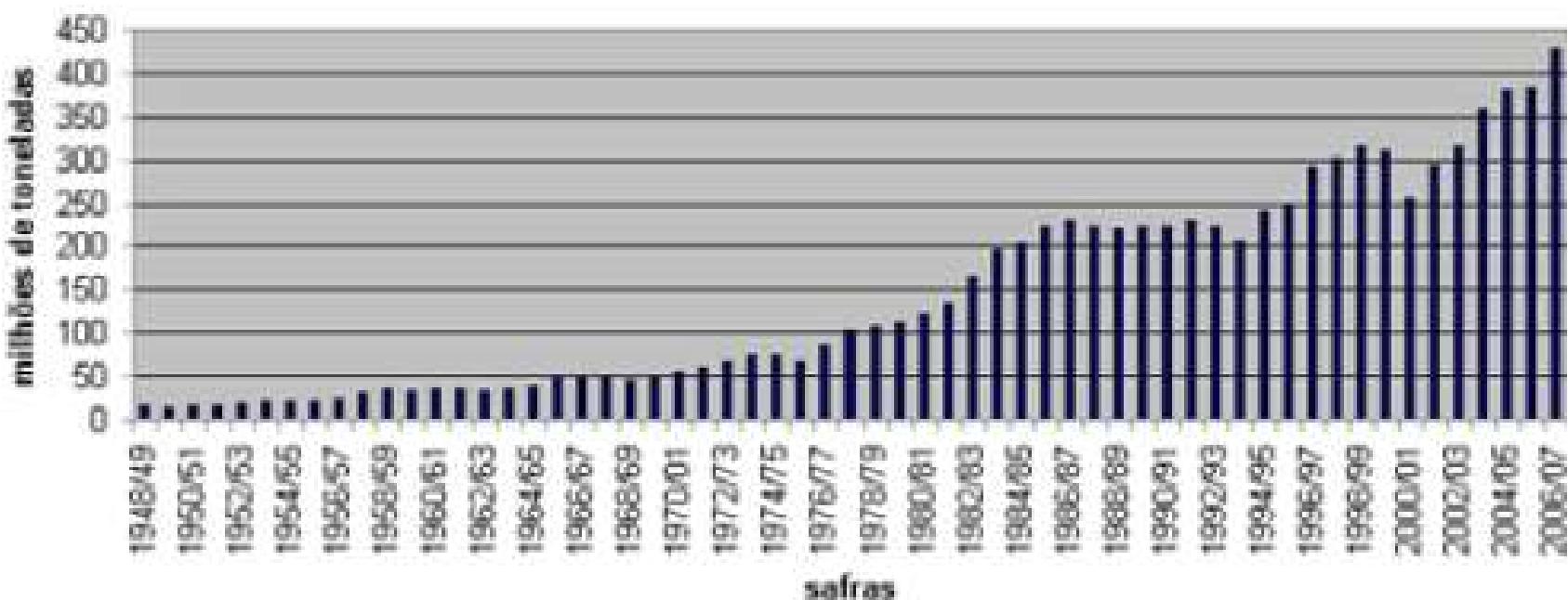


Fig. 1. Evolução da produção brasileira de cana-de-açúcar nas safras: 1948/1949 a 2006/2007.
Fonte: Brasil. Ministério da Agricultura, Pecuária e Abastecimento (2007).

SÉRIES GEOGRÁFICAS

DURAÇÃO MÉDIA DOS ESTUDOS SUPERIORES 1994

PAÍSES	NÚMERO DE ANOS
Itália	7,5
Alemanha	7,0
França	7,0
Holanda	5,9
Inglaterra	Menos de 4

FONTE: APA.

SÉRIES ESPECÍFICAS

ESTAS SÉRIES TAMBÉM CHAMADAS DE CATEGÓRICAS, APRESENTAM OS VALORES DE UMA VARIÁVEL, DE UM CERTO TEMPO E LOCAL, EM CATEGORIAS OU ESPECIFICAÇÕES.

PIB BRASILEIRO ENTRE 2000 E 2005

ANO DE REFERÊNCIA	NOVA SÉRIE	ANTIGA SÉRIE
2000	4,30%	4,40%
2001	1,30%	1,30%
2002	2,70%	1,9%
2003	1,10%	0,50%
2004	5,70%	4,90%
2005	2,90%	2,30%

Fonte: IBGE.

GRÁFICOS

O GRÁFICO ESTATÍSTICO É UMA FORMA DE APRESENTAÇÃO DOS DADOS, CUJO OBJETIVO É O DE PRODUZIR, NO INVESTIGADOR OU NO PÚBLICO EM GERAL, UMA IMPRESSÃO MAIS RÁPIDA E VIVA DO FENÔMENO EM ESTUDO, JÁ QUE OS GRÁFICOS FALAM MAIS RÁPIDO À COMPREENSÃO QUE AS SÉRIES (TABELAS).

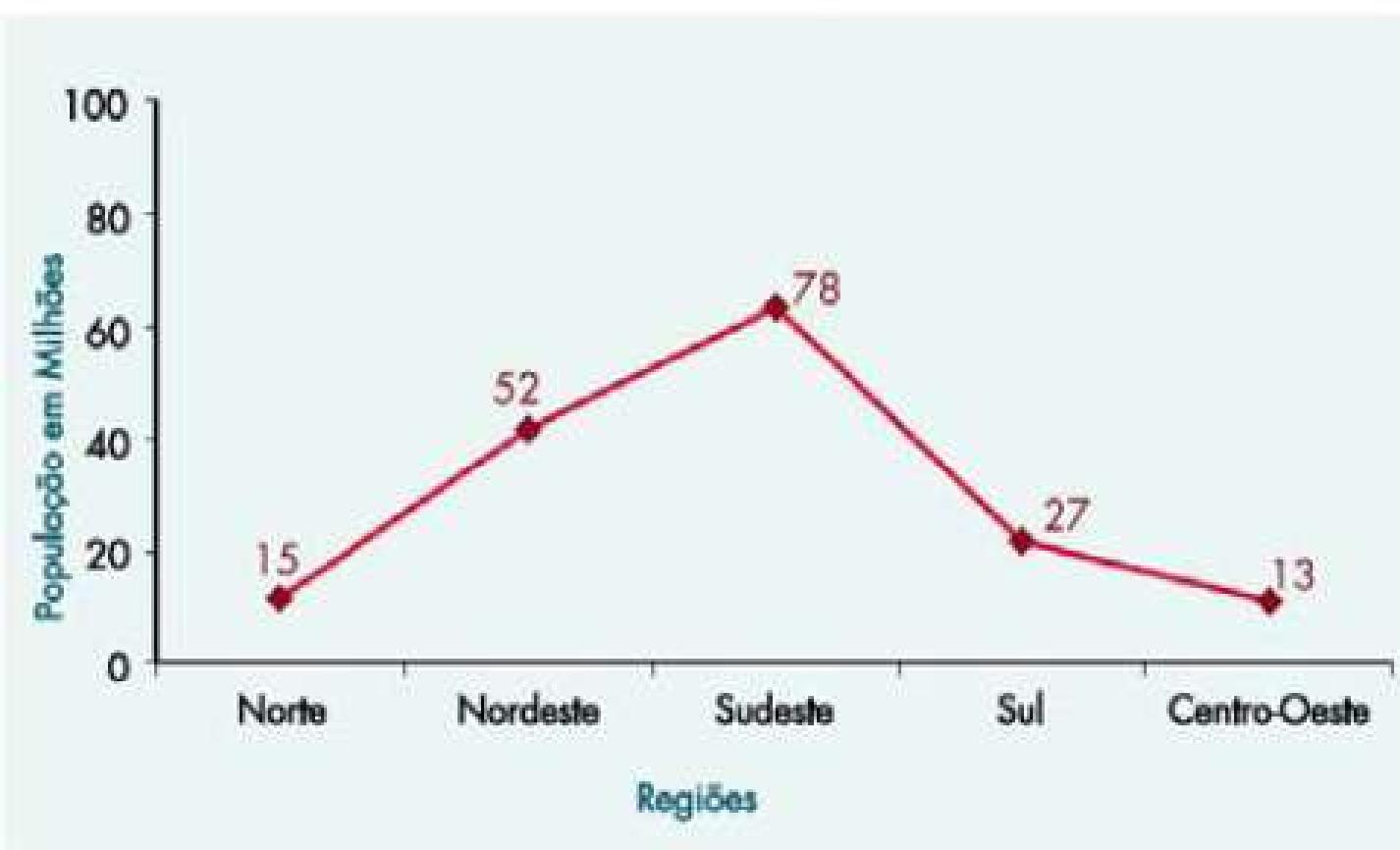
REQUISITOS FUNDAMENTAIS PARA A REPRESENTAÇÃO GRÁFICA:

- **SIMPLICIDADE;**
- **CLAREZA;**
- **VERACIDADE.**

FORMAS DE GRÁFICOS

GRÁFICO EM LINHA – UTILIZA-SE DA LINHA POLIGONAL PARA REPRESENTAR A SÉRIE ESTATÍSTICA; É RECOMENDADA A SUA UTILIZAÇÃO QUANDO SE DISPÕE DE UM CONJUNTO CONSIDERÁVEL DE DADOS.

**População recenseada e estimada, segundo as grandes regiões e
as unidades da Federação – 2007**



Fonte: IBGE, adaptado.

GRÁFICO EM COLUNAS OU EM BARRAS

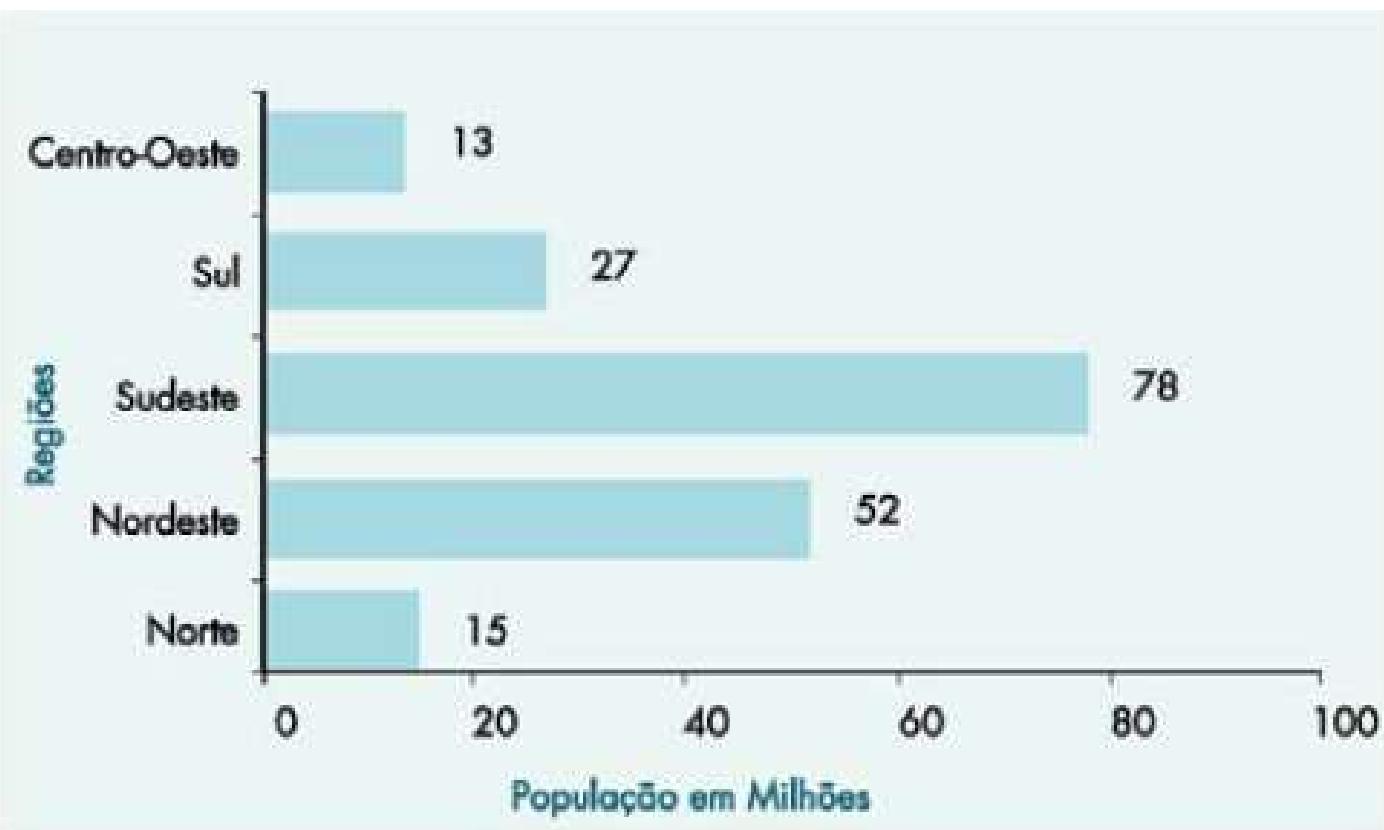
**REPRESENTA UMA SÉRIE POR MEIO
DE RETÂNGULOS, DISPOSTOS
VERTICALMENTE (COLUNAS) OU
HORIZONTALMENTE (BARRAS).**

População recenseada e estimada, segundo as grandes regiões e as unidades da Federação – 2007



Fonte: IBGE, adaptado.

População recenseada e estimada, segundo as grandes regiões e as unidades da Federação – 2007



Fonte: IBGE, adaptado.

GRÁFICO EM COLUNAS OU BARRAS MÚLTIPLAS

**EMPREGA-SE QUANDO SE QUER
REPRESENTAR, SIMULTANEAMENTE, DOIS
OU MAIS FENÔMENOS ESTUDADOS COM O
PROpósito DE COMPARAÇÃO.**

Distribuição percentual dos domicílios particulares permanentes, por forma de esgoto sanitário – 2004

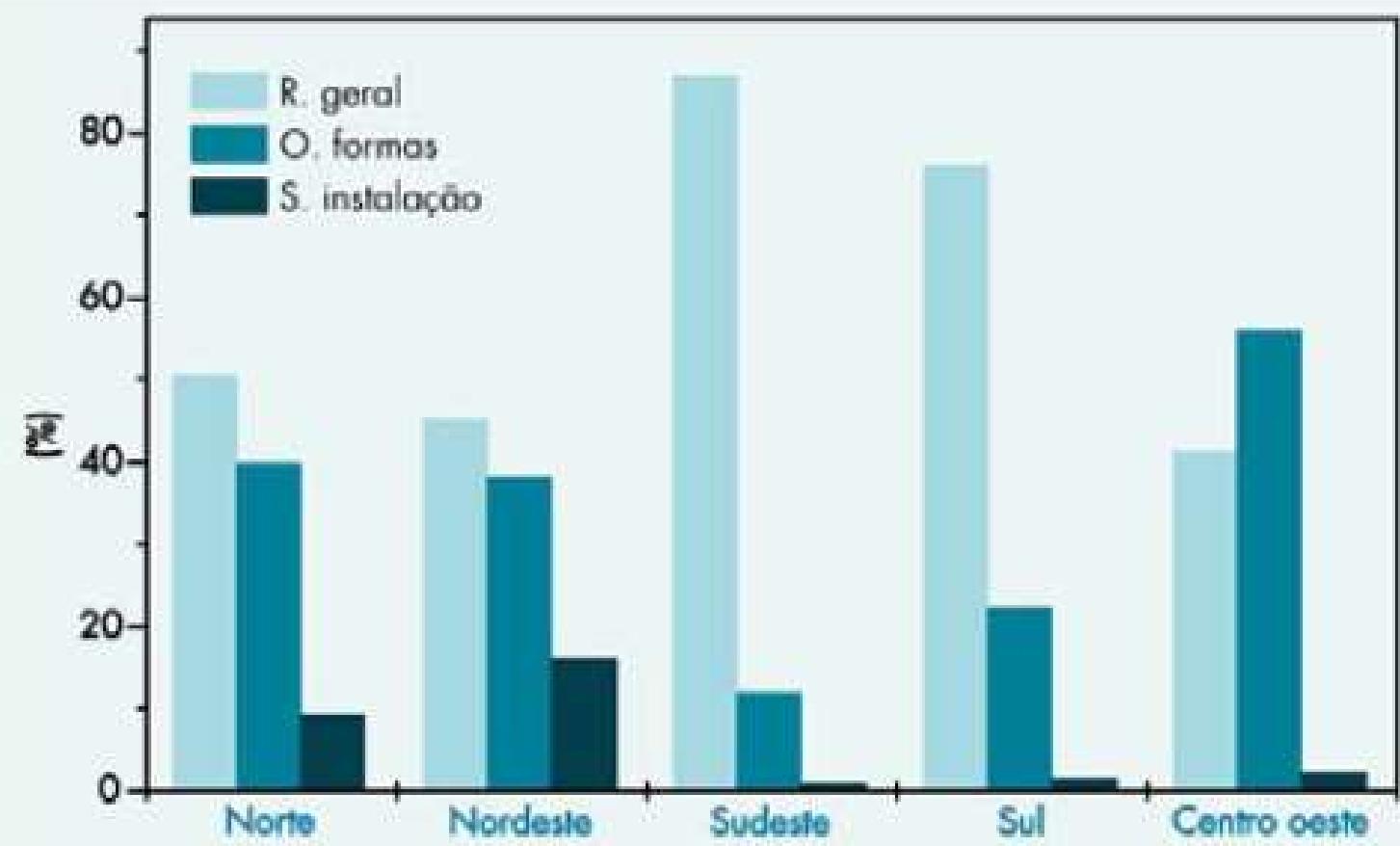
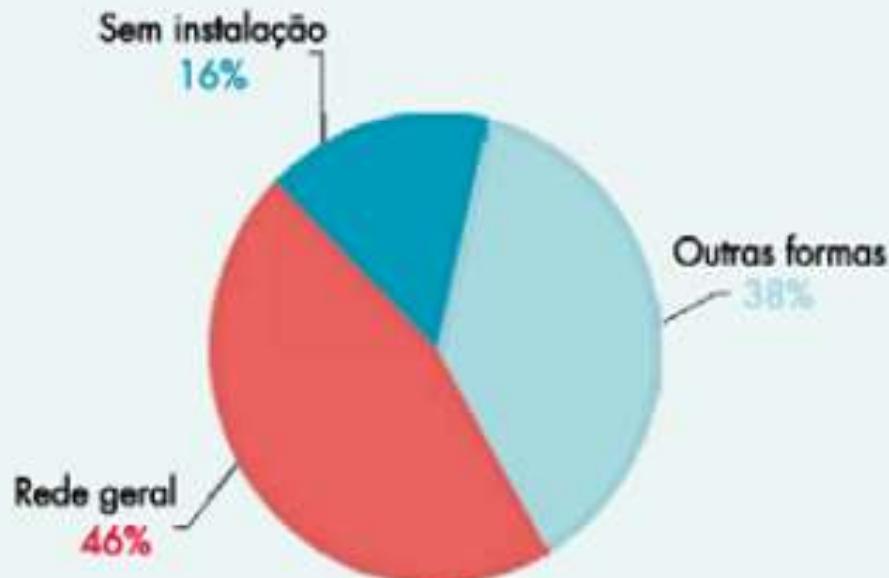


GRÁFICO EM SETORES

**CONSTITUI-SE COM BASE EM UM CÍRCULO,
E É EMPREGADO SEMPRE QUE
DESEJAMOS RESSALTAR A PARTICIPAÇÃO
DO DADO NO TOTAL.**

Distribuição percentual dos domicílios particulares permanentes, por forma de esgotamento sanitário – 2004

Região Nordeste



Fonte: IBGE.

Pictogramas

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

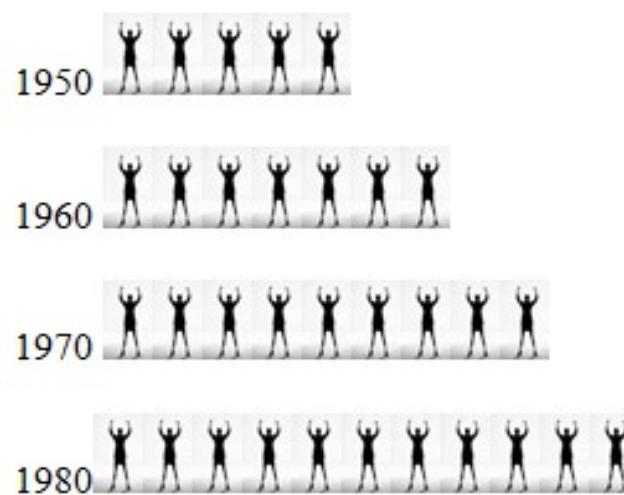
Exemplo:

**POPULAÇÃO DO BRASIL
1950 - 1980**

<i>ANOS</i>	<i>HABITANTES (milhares)</i>
1950	51.944
1960	70.191
1970	93.139
1980	119.071

FONTE: IBGE

**POPULAÇÃO DO BRASIL
1950 – 1980**



Cada símbolo representa 10.000.000 de habitantes
FONTE: IBGE

Cartogramas:

Distinguimos duas aplicações:

Representar dados absolutos (população) – neste caso lançamos mão, em geral, dos pontos, em número proporcional aos dados.

Representar dados relativos (densidade) – neste caso lançamos mão, em geral de hachuras.

Cartogramas:

exemplo:

**POPULAÇÃO PROJETADA DA REGIÃO SUL DO BRASIL
1990**

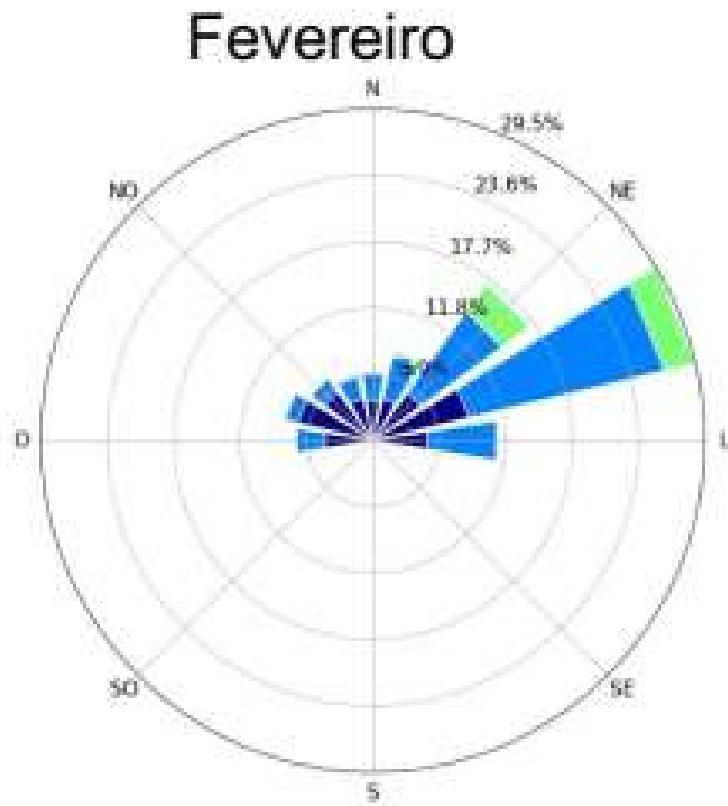
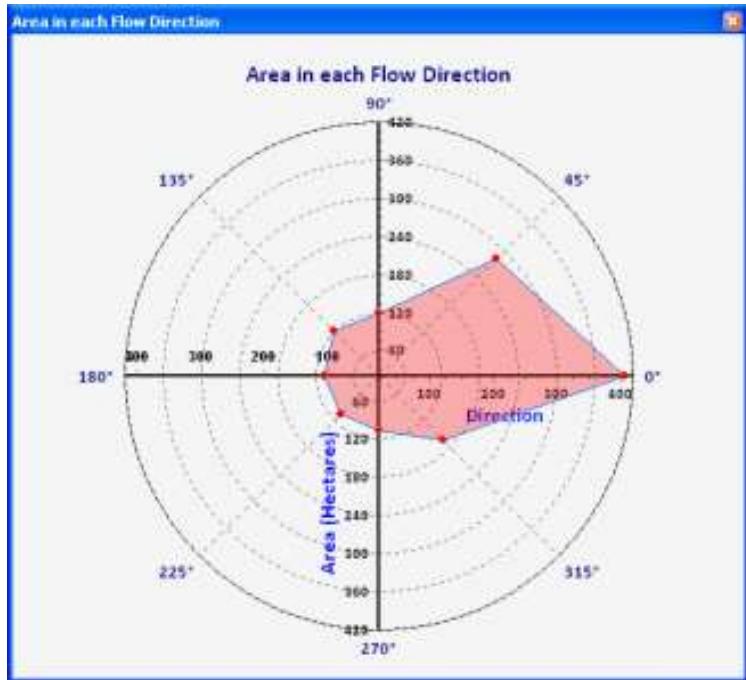
ESTADO	POPULAÇÃO (hab)	ÁREA(Km ²)	DENSIDADE
Paraná	9.137.700	199.324	45,8
Santa Catarina	4.461.400	95.318	46,8
Rio Grande do Sul	9.163.200	280.674	32,6

FONTE: IBGE



- Menos de 46,0 hab/km²
- Menos de 50,0 hab/km²
- Menos de 35,0 hab/km²

Gráfico POLAR:



Gráficos no EXCEL

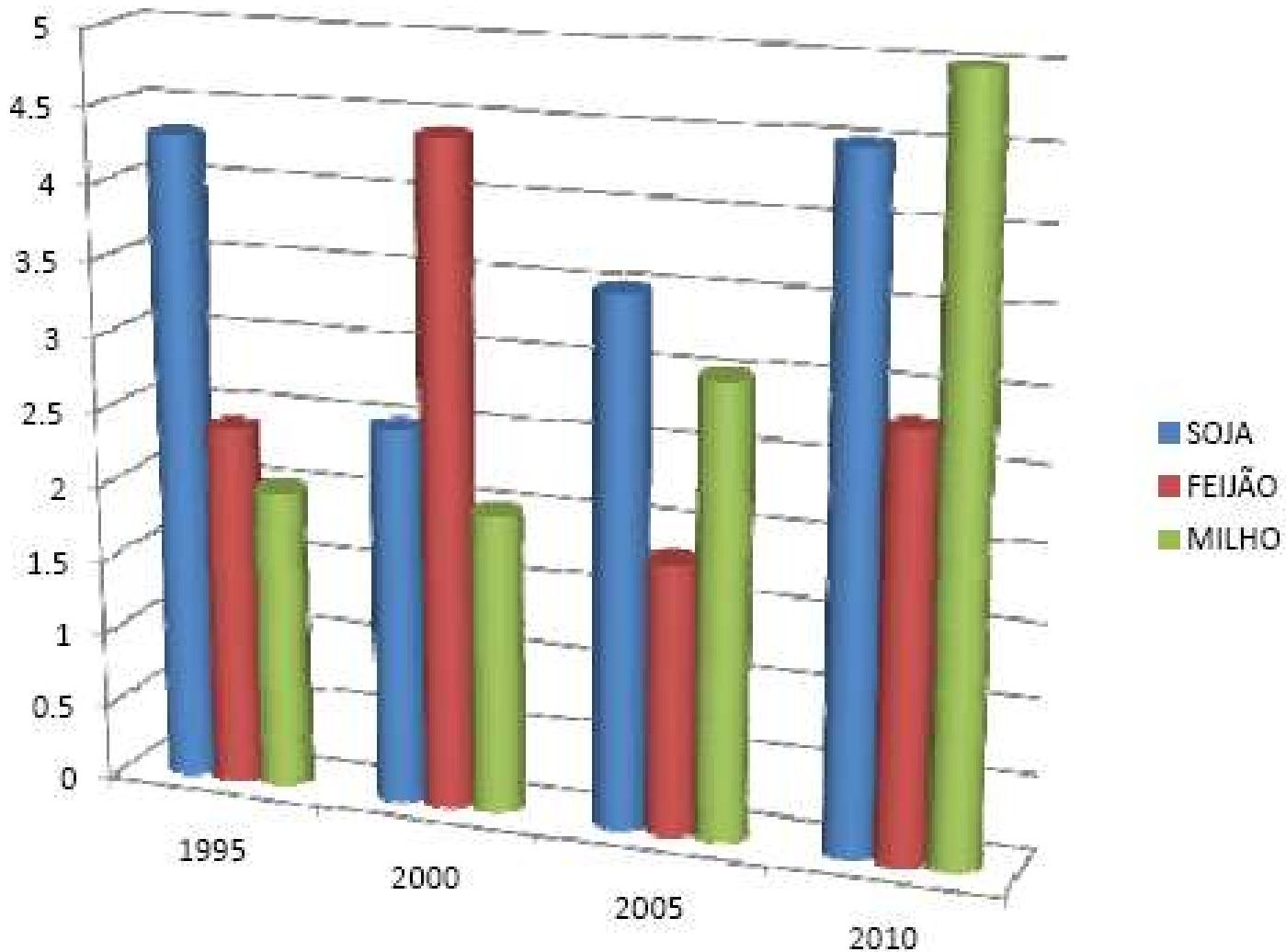
Para montar qualquer gráfico no Excel, é necessário que se tenha uma tabela pronta, na mesma planilha em que se construirá o gráfico.

Essa tabela é que fornecerá os dados para o gráfico.



Exemplo :

	SOJA	FEIJÃO	MILHO
1995	4,3	2,4	2
2000	2,5	4,4	2
2005	3,5	1,8	3
2010	4,5	2,8	5



SOMATÓRIO

A SOMA DOS NÚMEROS DO CONJUNTO
 $X = \{3, 6, 9, 12, 15\}$ CORRESPONDE A:

EQUAÇÃO:

$$\sum_{i=1}^n x_i = x_1 + x_2 + x_3 + \dots + x_n$$

$$\sum_{i=1}^5 x_i = 3 + 6 + 9 + 12 + 15 = 45$$

Exemplo :

EM UM TESTE DE MÁTEMÁTICA, APLICADA EM UMA CLASSE COM 20 ALUNOS, CHEGOU-SE AOS SEGUINTE RESULTADOS:

6	10	9	2	8
3	7	4	6	5
1	9	7	10	8
0	9	6	6	9

DETERMINAR:

$$\sum_{i=14}^{18} x_i = 10 + 8 + 0 + 9 + 6 = 33$$

$$\sum_{i=1}^5 x_i^2 = 36 + 100 + 81 + 4 + 64 = 285$$

PROPRIEDADES

$$1) \sum_{i=1}^n a x_i = a \sum_{i=1}^n x_i$$

$$2) \sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$$

EXEMPLO

ORDEM DO VALOR “i”	x_i	y_i	$x_i + y_i$
1	3	2	5
2	6	4	10
3	9	6	15
4	12	8	20
5	15	10	25

Exemplo :

$$\sum_{i=1}^5 x_i \cdot y_i =$$

$$\left(\sum_{i=1}^5 x_i \right) \bullet \left(\sum_{i=1}^5 y_i \right) = 4$$

DISTRIBUIÇÃO DE FREQUÊNCIA COM INTERVALO DE CLASSES

DISTRIBUIÇÃO DE FREQUÊNCIA

Distribuição de frequência:

é um tipo de tabela que condensa uma coleção de dados conforme as frequências (repetições de seus valores).

Frequência:

É o número de vezes em que se repete um determinado valor.
Como por exemplo, o número de alunos que apresentam
a mesma estatura.

ESTATURA (cm) DE 40 PESSOAS

DADOS BRUTOS

166	160	161	150	162	160	165	167	164	160
162	161	168	163	156	173	160	155	164	168
155	152	163	160	155	155	169	151	170	164
154	161	156	172	153	157	156	158	158	161



ESTATURA DE 40 PESSOAS

ESTATURA	FREQUÊNCIA
150 → 154	5
154 → 158	10
158 → 162	11
162 → 166	7
166 → 170	5
170 → 174	2
TOTAL	40

DADOS AGRUPADOS

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQUÊNCIA

CLASSE (i)

**SÃO INTERVALOS EM QUE OCORREM UM
DETERMINADO VALOR PARA A VARIÁVEL.**

150 → 154

CLASSE i_2

K = 6 / SEIS CLASSES

LIMITES DE CLASSE (I)

SÃO OS EXTREMOS DE CADA

CLASSE (LIMITE INFERIOR I_i E

LIMITE SUPERIOR I_s)

150 → 154

$$\begin{aligned}I_i &= 150 \\I_s &= 154\end{aligned}$$

AMPLITUDE DE UM INTERVALO DE CLASSE (h_i)

É A DIFERENÇA ENTRE O LIMITE DE CLASSE SUPERIOR E O LIMITE DE CLASSE INFERIOR DE UM CONJUNTO DE DADOS.

$$h_i = l_s - l_i$$

$$150 \mapsto 154$$

$$h = 154 \text{ cm} - 150 \text{ cm} = 4 \text{ cm}$$

AMPLITUDE TOTAL DA DISTRIBUIÇÃO (A_t)

É A DIFERENÇA ENTRE O LIMITE SUPERIOR (MÁXIMO) DA ÚLTIMA CLASSE E O LIMITE INFERIOR (MÍNIMO) DA PRIMEIRA CLASSE.

$$A_t = L_s(\text{max.}) - L_i(\text{min.})$$

$$A_t = 174 \text{ cm} - 150 \text{ cm} = 24 \text{ cm}$$

QUANDO AS CLASSES APRESENTAM O MESMO
INTERVALO, É VÁLIDA A RELAÇÃO:

A_t – AMPLITUDE TOTAL DA
DISTRIBUIÇÃO

$$\frac{A_t}{h} = k$$

h – AMPLITUDE DE UM INTERVALO DE
CLASSE

$$\frac{24}{4} \Leftrightarrow k = 6$$

AMPLITUDE AMOSTRAL

**AMPLITUDE AMOSTRAL É A DIFERENÇA
ENTRE O VALOR MÁXIMO E O VALOR
MÍNIMO DA AMOSTRA**

$$A_a = 173 \text{ cm} - 150 \text{ cm}$$

$$A_a = 23 \text{ cm}$$

FREQÜÊNCIA SIMPLES OU ABSOLUTA (f_i)

É O NÚMERO DE VEZES EM QUE SE REPETE UMA CLASSE OU UM DETERMINADO VALOR.

$$\sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k = n$$

ESTATURA DE 40 PESSOAS

ESTATURA	FREQUÊNCIA
150 → 154	4
154 → 158	9
158 → 162	11
162 → 166	8
166 → 170	5
170 → 174	3
TOTAL	$\sum f_i = 40$

NÚMERO DE CLASSES OU INTERVALOS DE CLASSE

REGRA DE STURGES (nº de CLASSES)

$$i = 1 + 3,3 \cdot \log n$$

i É O NÚMERO DE CLASSE;

n É O NÚMERO TOTAL DE DADOS.



FREQÜÊNCIA RELATIVA (f_r)

**É O QUOCIENTE ENTRE AS
FREQÜÊNCIA SIMPLES E A TOTAL.**

$$f_r = \frac{f_i}{\sum f_i}$$

FREQÜÊNCIA ACUMULADA(f_{ac})

É A SOMA DEFINIDA ENTRE A FREQÜÊNCIA ABSOLUTA OU RELATIVA DA SUA CLASSE COM A CLASSE ANTERIOR.

$$f_{ac} = \sum f_i \text{ ou } \sum f_r$$

NÚMERO DE CLASSES

O NÚMERO DE CLASSES DEPENDE
MUITO DAS HABILIDADES DO
PESQUISADOR E DAS NECESSIDADES
DAS RESPOSTAS PARA UMA VARIÁVEL
CONTÍNUA.

CRITÉRIO DA RAIZ

PARA UMA SÉRIE COM “ n ” ELEMENTOS,
INDICAREMOS POR Q O NÚMERO DE
CLASSES A SER UTILIZADO. PELO
CRITÉRIO DA RAIZ:

$$Q = \sqrt{n}$$

$$Q = \sqrt{40}$$

$$Q = 6,3 \approx 6$$

AMPLITUDE DA CLASSE - h

A AMPLITUDE DO INTERVALO DE CLASSE “**h**”, CALCULAMOS ATRAVÉS DA RELAÇÃO:

$$h = \frac{A_t}{Q}$$

EXEMPLO 1

DADOS BRUTOS

15	17	16	15	17	16	15	17	16	15
17	15	14	17	15	14	17	15	14	15
16	16	15	16	16	15	16	16	15	16
15	15	16	15	15	16	15	15	16	15
14	14	15	14	14	15	14	14	15	15

ROL

14 14 14 14 14 14 14 14 14 14 15

15 15 15 15 15 15 15 15 15 15 15

15 15 15 15 15 15 15 15 15 15 15

15 16 16 16 16 16 16 16 16 16 16

16 16 16 16 17 17 17 17 17 17

IDADES	f_i	$f_{i\text{ ac}}$	$f_r (\%)$	$f_{r\text{ ac}} (\%)$
14	09	09	18	18
15	22	31	44	62
16	13	44	26	88
17	06	50	12	100

DADOS BRUTOS

7	9,9	9,9	10	10	7	10	10	8	10
6	6	6	6	6,6	6,8	6,8	6,8	6,9	6,9
7	7	7	8	7	10	7	7	7	7
8	8	8	8	7	8	8	8	8	8
8	8	8	8	8	8	10	8	8	8
5,9	6	6	6	6	6	6	6	6	6
8	8	8	8,4	8	8,7	8,9	9	9	9
5	5	5,4	5,5	5,6	5,8	5,8	5,8	5,9	5,9
7	9,9	7	7	7,7	7,8	7,8	7,9	8	8

ROL

5	5	5,4	5,5	5,6	5,8	5,8	5,8	5,9	5,9
5,9	6	6	6	6	6	6	6	6	6
6	6	6	6	6,6	6,8	6,8	6,8	6,9	6,9
7	7	7	7	7	7	7	7	7	7
7	7	7	7	7,7	7,8	7,8	7,9	8	8
8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8
8	8	8	8,4	8	8,7	8,9	9	9	9
9	9	9	9	9	9	9,5	9,8	9,8	9,9
9,9	9,9	9,9	10	10	10	10	10	10	10

NÚMERO DE CLASSES

$$Q = \sqrt{n}$$

$$Q = \sqrt{100}$$

$$Q = 10$$

AMPLITUDE TOTAL

$$A_t = L_s(\max.) - L_i(\min.)$$

$$A_t = 10 - 5 = 5$$

AMPLITUDE DA CLASSE

$$h = \frac{A_t}{Q}$$

$$h = \frac{5}{10}$$

$$h = \frac{1}{2}$$

Classes	Notas	f_i	f_{i ac}	f_r (%)	f_{r ac} (%)
1	5,0 ← 5,5	3	3	3	3
2	5,5 ← 6,0	8	11	8	11
3	6,0 ← 6,5	13	24	13	24
4	6,5 ← 7,0	6	30	6	30
5	7,0 ← 7,5	14	44	14	44
6	7,5 ← 8,0	4	48	4	48
7	8,0 ← 8,5	26	74	26	74
8	8,5 ← 9,0	3	77	3	77
9	9,0 ← 9,5	9	86	9	86
10	9,5 ← 10,0	14	100	14	100

EXEMPLO 02

NO QUADRO A SEGUIR, CONSTA AS TEMPERATURAS MÁXIMAS DE 56 CIDADES DO MUNDO EM CERTO DIA DE JANEIRO. AGRUPÁ-LAS EM UMA DISTRIBUIÇÃO DE FREQUÊNCIA E INTERPRETAR OS RESULTADOS

DADOS BRUTOS

10 17 14 9 18 10 23 8 9 26 19

30 30 -13 4 8 6 7 7 28 -3 24

13 10 25 20 28 9 9 9 35 26 26

27 2 - 6 21 10 3 10 5 13 28 13

5 9 29 10 4 5 8 5 12 6 19

20

DETERMINAR:

- 1 – AMPLITUDE TOTAL;**
- 2 – NÚMERO DE CLASSES;**
- 3 – AMPLITUDE DAS CLASSES;**
- 4 – INTERVALOS;**
- 5 – TABELA DE FREQUÊNCIAS SIMPLES E ACUMULADAS**

CRITÉRIOS PARA DETERMINAR AS CLASSES

1. Raiz quadrada: $k = \sqrt{n}$
2. Log (Sturges): $k = 1 + 3,3 \log n$
3. ln (Milone): $k = -1 + 2 \cdot \ln n$

Milone (2004) acrescenta ainda que, adotando o princípio de que os agrupamentos devem ter no mínimo cinco e no máximo 20 classes, o critério da raiz é valido para $25 \leq n \leq 400$, o do log para $16 \leq n \leq 572.237$ e o do ln para $20 \leq n \leq 36.315$.

Por outro lado, o pesquisador pode definir o número de classes baseando-se em sua experiência.

ROL

- 13	- 6	- 3	2	3	4	4	5	5	5	5	6
6	7	7	8	8	8	9	9	9	9	9	10
10	10	10	10	10	12	13	13	13	14	17	18
19	19	20	20	21	23	24	25	26	26	26	
27	28	28	28	29	30	30	32	35			

AMPLITUDE TOTAL

$$A_t = 35 - (-13) = 48^{\circ}\text{C}$$

NÚMERO DE CLASSES

$$k = \sqrt{56}$$

$$k = 7,5 \text{ ou}$$

$$\textcolor{red}{K = 8}$$

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

$$k = 6,8 \approx 7$$

AMPLITUDE DAS CLASSES

$$A_c = \frac{A_t}{k} = \frac{48}{8} = 6$$

TABELA DE FREQUÊNCIA

Classes	Notas	f _i	f _{i ac}	f _r (%)	f _{r ac} (%)
1	- 13 ← - 7	1	1	1,79	1,79
2	- 7 ← - 1	2	3	3,57	5,36
3	- 1 ← 5	4	7	7,14	12,50
4	5 ← 11	22	29	39,28	51,78
5	11 ← 17	5	34	8,93	60,71
6	17 ← 23	7	41	12,50	73,21
7	23 ← 29	10	51	17,86	91,07
8	29 ← 35	5	56	8,93	100,00
TOTAL		$\sum = 56$		$\sum = 100$	

EXERCÍCIO

1) Os dados abaixo são relacionados com as idades em que morreram 40 pessoas doentes que se expuseram a um determinado tratamento. Fazer a tabela de distribuição de freqüência.

72 63 60 60 58 58 50 56 63 66 70 50 52 53 55 57 57 54 67 71 68 64
65 61 61 58 59 56 54 57 59 62 60 60 62 68 69 55 62 73

2) Os dados são referentes à Taxa de Mortalidade Infantil dos Municípios da Micro-Região Norte Mineira, ano de 2002. Faça a distribuição de freqüência dos dados.

32,3	62,2	10,3	22,0	13,1	9,9	11,9	20,0	36,4	23,5	18,0	22,6
20,3	38,3	19,6	27,2	28,9	18,4	27,3	21,7	23,7	13,9	36,3	32,9
29,7	25,4	23,8	15,7	17,0	39,2	22,7	29,9	18,3	33,0		

3) Com os dados do exercício nº 1 e do nº 2,
preencha :

- A) L1 e L1:
- B) L4 e L4:
- C) Amplitude do intervalo de classe:
- D) Amplitude total da distribuição:
- E) Amplitude total da amostra:
- F) Ponto médio da classe 1:
- G) Ponto médio da classe 4: