

Disciplina: Estatística na Fisioterapia

Estatística Descritiva com Aplicações no Excel

Fisioterapia

Bertolo, L.A.



Capítulo 1 - Introdução

Por que estudar estatística?¹

Abordagem crescentemente quantitativa utilizada em todas as ciências. Uso de técnicas matemáticas na avaliação de controles de poluição, no planejamento de inventários, na análise de problemas de tráfego, no estudo dos efeitos de vários remédios, na avaliação de técnicas de ensino, na análise do comportamento competitivo de administradores e governos, no estudo de dietas e longevidade.

Aumento da quantidade de dados coletados, processados e apresentados ao público.

O avanço tecnológico permitiu dar maior ênfase à compreensão e à interpretação dos resultados ao tornar os cálculos relativamente fáceis.

A Estatística está presente em diversas áreas e em diversas abordagens. Podemos citar alguns exemplos de suas aplicações:

1. *Eleições*: Anúncio prévio, em termos percentuais, de que o candidato A é o provável vencedor. (A previsão pode ser feita após a contagem de 2% dos votos);
2. *Indústria*: Teste de lâmpadas flash" para máquinas fotográficas. O fabricante deve saber com antecedência se o lote fabricado pode ser colocado no mercado ou não (Ele pode estar sujeito a riscos !
- O *Recall* é um exemplo, onde ocorre a retirada ou troca de produtos devido a uma falha em algum estágio de fabricação não detectado pelo controle de qualidade);
3. *Meteorologia*: Informa que a "chance" de chuva para hoje é de 30% (utilização de métodos estatísticos de previsão);
4. *Mercado de Valores*: Decisão na compra ou venda de ações – "Risco";
5. *Instituições Bancárias/Crédito*: Análise de crédito; Seguro de Vida e/ou de automóveis;
6. *IBGE/Índices*: Censo; Taxa de desemprego; Inflação; Custo de vida; Valor do salário mínimo, cesta básica, índices de preços-INPC; Taxa de mortalidade/Natalidade; Índices de analfabetismo, etc..
7. *Farmácia/Medicina*: Analisar a entrada de um novo medicamento, verificando a sua eficácia. "Deve-se tomar muito cuidado" com a tomada de decisão;
8. *Pesquisas*: Estudo sobre hábitos migratórios de certo animal; Estimção do tamanho populacional de certas espécies;
9. *Experimentação Agrícola*: Estudos de uma nova variedade de semente (modificação genética) ou de fertilizantes;
10. *Telecomunicações*: A utilização da Internet gera informações preciosas de clientes ; Controle de tráfego telefônico, chamadas e consumo, forçam a utilização de modelos estatísticos complexos;
11. *Experimentação Zootécnica*: Estudo do desenvolvimento de um animal e análise de fatores que influenciam neste desenvolvimento tais como, alimentação, cativeiro e outros; Melhoramento genético; Pesquisas com vacinas animais entre outros.

O que é Estatística?

ESTATÍSTICA: é uma coleção de métodos² para planejar experimentos, obter e organizar dados, resumi-los, analisá-los, interpretá-los e deles extrair conclusões (*Triola, 1999*).

EX:- Com o objetivo de levantar conhecimento sobre algumas características de homens com doenças cardíacas da clínica especializada de Honolulu, no Havaí, decidiu-se pesquisar os pacientes com idade entre 45 e 67 anos. As características de interesse eram:

- 1 - Nível de instrução
- 2 - Peso
- 3 - Altura
- 4 - Idade
- 5 - Fumante
- 6 - Atividade física
- 7 - Taxa de glicose
- 8 - Taxa de colesterol sérico
- 9 - Pressão sanguínea sistólica.

¹ Ver Apêndice 1.43 – Estatística Médica, p. 61

² é um meio mais eficaz para atingir determinada meta.

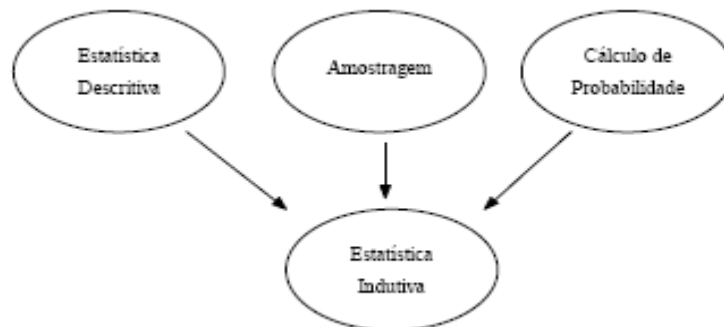
A origem da palavra Estatística está associada à palavra *status* (Estado em latim).

A Estatística se divide em:

ESTATÍSTICA DESCRITIVA: A coleta, a organização, a descrição dos dados, o cálculo e a interpretação de coeficientes. Esta parte está associada a cálculos de médias, variâncias, estudo de gráficos, tabelas, etc.. É a parte mais conhecida.

ESTATÍSTICA INDUTIVA ou INFERENCIAL: a análise e a interpretação dos dados, associado a uma margem de incerteza, cujos métodos que se fundamentam na teoria da probabilidade. Nela encontramos a *Estimação de Parâmetros*, *Testes de Hipóteses*, *Modelagens*, etc.

O diagrama seguinte mostra o contexto em que se situa o estudo da Estatística subdividido em Estatística *Descritiva* e Estatística *Indutiva* (ou *Inferencial*).



AMOSTRAGEM: é o ponto de partida (na prática) para todo um Estudo Estatístico.

Um dos objetivos da análise e interpretação de dados é buscar um modelo para as observações. Estes modelos podem ser essencialmente *determinísticos* ou *não-determinísticos* (*probabilísticos* ou *estocásticos*).

Nos determinísticos as condições sob as quais um experimento é executado determina o resultado do experimento.

Ex.: Um veículo viajando sempre a 100 km/h irá levar 2 horas para percorrer 200 km.

Aqui cabe apenas a otimização das variáveis de decisão para atingirmos o máximo ou o mínimo da função objetivo.

Nos modelos **não** determinísticos usa-se uma *Distribuição de Probabilidade*.

Ex.: Peças são fabricadas até que x peças, perfeitas, sejam produzidas; o número total de peças fabricadas é contado. Usa-se uma distribuição, no caso a *Geométrica*, para a tomada de decisões

1.1 - População e Amostra

O estudo de qualquer fenômeno, seja ele natural, social, econômico, exige a coleta e a análise de dados estatísticos. A coleta de dados é, pois, a fase inicial de qualquer pesquisa.

A *População* é a coleção de todas as observações *potenciais* sobre determinado fenômeno ou sobre um conjunto de indivíduos (possuindo estes pelo menos uma característica comum). A população é o conjunto Universo, podendo ser finita ou infinita.

- Finita – apresenta um número limitado de observações, que é passível de contagem.
- Infinita – apresenta um número ilimitado de observações que é impossível de contar e geralmente está associada a processos.

Senso – é uma coleção de dados relativos a todos os elementos de uma população.

Amostra da população é o conjunto de dados efetivamente observados, ou extraídos³. Sobre os dados da amostra é que se desenvolvem os estudos, com o objetivo de se fazerem inferências sobre a população. Devem ser escolhidas através de processos adequados que garantam o acaso na escolha.

Chama-se **amostragem** o processo de colher amostras. Nesse processo, cada elemento da população passa a ter a mesma chance de ser escolhido. Dentre os processos de amostragem, podem-se destacar três: *amostragem casual ou aleatória simples*, *amostragem proporcional estratificada* e *amostragem sistemática*.

³ Trata-se, portanto, de um subconjunto finito de uma população.

1.1.1 – Processos de Amostragem

a) Amostragem casual ou aleatória simples:

É um sorteio, por exemplo, para retirar uma amostra de 9 alunos de uma sala de 90 alunos, utiliza-se um sorteio com todos os números dos alunos escritos em papéis dentro de um saco. Para amostras grandes utiliza-se a Tabela de Números Aleatórios⁴. Assim para o exemplo da sala de aula, utilizando dois algarismos, através da leitura de uma linha (escolhida através de sorteio), obtém-se:

Como a população vai de 1 a 90 escolhe-se os 9 primeiros números dentro dessa faixa:

TABELA DE NÚMEROS ALEATÓRIOS:

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

69	129	95	123	81	93	105	95	96	80	87	110	139	75
123	60	72	86	108	120	57	113	65	108	90	137	74	106
109	84	121	60	128	100	72	119	103	128	80	99	149	85
77	91	51	100	63	107	76	82	110	63	131	65	114	103
104	107	63	117	116	86	115	62	122	92	102	113	74	78
69	116	82	95	72	121	52	80	100	85	117	85	102	106
94	84	123	42	90	91	81	116	73	79	98	82	69	102
100	79	101	98	110	95	67	77	91	95	74	90	134	94
79	92	73	83	74	125	101	82	71	75	101	102	78	108
125	56	86	98	106	72	117	89	99	86	82	57	106	90

Obtenha uma amostra formada de 26 elementos, tomando, inicialmente, a 1ª linha da esquerda para a direita

A.5 – O diretor de uma escola, na qual estão matriculados 280 meninos e 320 meninas, desejoso de conhecer as condições de vida extra-escolar de seus alunos e não dispondo de tempo para entrevistar todas as famílias, resolveu fazer um levantamento por amostragem, em 10% dessa clientela. Obtenha, para esse diretor, os elementos componentes da amostra.

b) Amostragem proporcional estratificada:

É comum termos populações que se dividam em sub-populações (estratos) e, como cada estrato pode ter um comportamento diferente do outro, a amostra deve considerar a existência desses estratos e a sua proporção em relação à população.

Exercícios de Aplicação

Supondo que uma sala de aula seja composta de 54 meninos e 36 meninas. Determine uma amostra de 9 pessoas:

Sexo	População	Cálculo Proporcional Regra de três simples	Amostra
Masculino	54	$54 \times 9 / 90 = 5,4$	5
Feminino	36	$36 \times 9 / 90 = 3,6$	4
Total	90	9	9

Posteriormente, utiliza-se a tabela de números aleatórios para escolher 5 meninos e 4 meninas da seguinte maneira:

Numeramos os alunos de 01 a 90, sendo de 01 a 54 correspondendo às meninas e de 55 a 90, os meninos. Na Tabela tomamos a primeira e a segunda colunas da esquerda, de cima para baixo, obtemos os seguintes números

40 ~~94~~ ~~91~~ 18 54 89 33 45 09 ~~00~~ ~~40~~ 48 83 ~~94~~ 72 75 05 77

Temos então:

meninos: 40 18 54 33 45

meninas: 89 83 72 75

PESQUISA

População: peso dos colegas de sua sala (incluindo você)

Amostra: correspondente a 40% da população

Sugestão: Tabela dos Números Aleatórios (5ª e 6ª colunas, de baixo para cima)

Exercícios Propostos

B.1 – Em uma escola existem 250 alunos, distribuídos em séries conforme a tabela. Obtenha uma amostra de 40 alunos e preencha a tabela.

Séries	População	Cálculo Proporcional	Amostra
1ª	35		
2ª	32		
3ª	30		
4ª	28		
5ª	35		
6ª	32		
7ª	31		
8ª	27		
Total	250		40

B.2 – Uma universidade apresenta o seguinte quadro relativo aos seus alunos do curso de Matemática. Obtenha uma amostra proporcional estratificada de 100 alunos.

Série	Qtde	Amostra
1ª	85	
2ª	70	
3ª	80	
4ª	75	
Total		100

B.3 – Uma cidade X apresenta o seguinte quadro relativo às suas escolas de 1º grau:

Escola	Homens	Mulheres	Total	Amostra		
				Homens	Mulheres	Total
A	80	95				
B	102	120				
C	110	92				
D	134	228				
E	150	130				
F	300	290				
Total						120

Obtenha uma amostra estratificada de 120 estudantes.

B.4 – Uma população encontra-se dividida em três estratos, com tamanhos, respectivamente, $n_1 = 40$, $n_2 = 100$ e $n_3 = 60$. Sabendo-se que, ao ser realizada uma amostragem estratificada proporcional, nove elementos da amostra foram retirados do 3º estrato, determine o número total de elementos da amostra.

c) Amostragem sistemática:

É quando a amostragem é feita através de um sistema possível de ser aplicado pois a população já se encontra ordenada.

Exercícios de Aplicação

- Em uma linha de produção, a cada 10 itens fabricados, retira-se 1 para inspeção, tem-se uma amostra de 10 % da população.
- em uma rua com 900 prédios, deseja-se uma amostra de 50. $900/50 = 18$ (50 grupos de 18 prédios cada). Faz-se um sorteio entre 1 e 18, por exemplo 4, então pesquisaremos o 4º prédio da rua, o 22º, o 40º, 58º, assim por diante.

Exercícios Propostos

C.1 – Mostre como seria possível retirar uma amostra de 32 elementos de uma população ordenada formada por 2.432 elementos.

Na ordenação geral, qual dos elementos abaixo seria escolhido para pertencer à amostra, sabendo-se que o elemento de ordem 1.420 a ela pertence?

1.648º 290º 725º 2.025º 1.120º.

1.2 - Análise exploratória dos dados

A inferência estatística permite chegar a conclusões amplas (gerais, sobre o universo estudado) a partir de uma série mais restrita (amostra) de informações (dados estatísticos). Daí a importância do ensino da estatística na formação de quase todas as categorias profissionais de nível superior.

Por meio da estatística fazemos, por exemplo, melhor análise dos conflitos e contradições que estamos pesquisando no momento e, assim, nós tomamos decisões que irão influenciar a resolução dos conflitos. Para tanto a estatística usa o método (pesquisa) estatístico composto basicamente de 5 fases:

1º) **coleta de dados** - Esta pode ser *direta* ou *indireta*. A coleta direta é feita sobre registros diversos: nascimento, casamento, óbitos, importação, registros escolares; ou ainda quando os dados são coletados diretamente pelo pesquisador através de questionários (ex: censo).

A *coleta direta* pode ser: contínua; periódica (censos); ocasional

A *coleta indireta* é uma coleta feita sobre dados colhidos de uma coleta direta (ex: mortalidade infantil)

2º) **apreciação ou crítica dos dados** - Os dados coletados devem ser observados, à procura de falhas e imperfeições, a fim de não causarem erro nos resultados.

Exemplo 1 : *Perguntas tendenciosas*. Foi realizada a seguinte pesquisa:

O tráfego contribui em maior ou menor grau do que a indústria para a poluição atmosférica?

Resposta: 45 % para o tráfego e 32 % para a indústria.

A indústria contribui em maior ou menor grau do que o tráfego para a poluição atmosférica?

Resposta: 24 % para o tráfego e 57 % para a indústria.

Exemplo 2: *Preservação da auto-imagem*. Em uma pesquisa telefônica 94 % dos entrevistados disseram que lavam as suas mãos após usar o banheiro, mas a observação em banheiros públicos esse percentual cai para 68 %.

Exemplo 3: *Más Amostras*. As pessoas devem ser escolhidas aleatoriamente para a pesquisa, como por exemplo, numa pesquisa de opinião na rua, deve-se entrevistar somente quem pisou em uma determinada marca pré-determinada na calçada.

Exemplo 4. *Más perguntas*. A pergunta deve conter o linguajar próprio do entrevistado. Geralmente, se o entrevistado não entender a pergunta, ele responderá qualquer coisa, pois tem vergonha de perguntar.

3º) **apuração dos dados** – É o processamento dos dados obtidos.

4º) **exposição dos dados** – Através de tabelas ou gráficos, tornando mais fácil seu exame e aplicação de um cálculo estatístico.

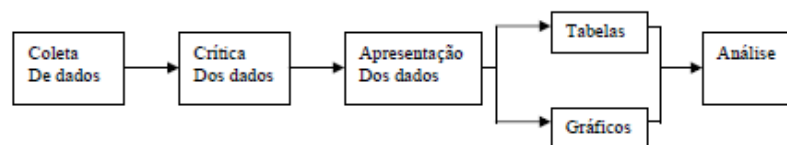
5º) **Análise dos Resultados** - Através de métodos de estatística indutiva ou inferencial obtêm-se conclusões e previsões de um todo através do exame de apenas uma parte desse todo.

Os erros e inconsistências ocorridos na coleta de dados devem ser corrigidos. As amostras de dados devem ser agrupadas de forma que seu manuseio, visualização e compreensão sejam simplificados.

Exemplos de onde empregar o método estatístico:

- Indústrias realizam pesquisa entre os consumidores para o lançamento de um novo produto;
- As pesquisas eleitorais fornecem elementos para que os candidatos direcionem a campanha;
- Emissoras de tevê utilizam pesquisas que mostram a preferência dos espectadores para organizar sua programação;
- A pesquisa do desempenho dos atletas ou das equipes em uma partida ou em um campeonato interfere no planejamento dos treinamentos.

Resumindo:



TRABALHO DE PESQUISA BIBLIOGRÁFICA

Leia o Capítulo 01 do livro **Estatística** do autor *Antonio Arnot Crespo* que se encontra na biblioteca e faça as seguintes atividades:

1. Complete:
O método experimental é o mais usado por ciências como:...
2. As ciências humanas e sociais, para obterem os dados que buscam, lançam mão de que método?
3. O que é Estatística?
4. Cite as fases do método estatístico.
5. Para você, o que é coletar dados?
6. Para que serve a crítica dos dados?
7. O que é apurar dados?
8. Como podem ser apresentados ou expostos os dados?
9. As conclusões, as inferências pertencem a que parte da Estatística?
10. Cite três ou mais atividades do planejamento empresarial em que a Estatística se faz necessária>
11. O método estatístico tem como um de seus fins:
 - a. estudar fenômenos estatísticos.
 - b. estudar qualidades concretas dos indivíduos que formam grupos.
 - c. determinar qualidades abstratas de grupos de indivíduos.

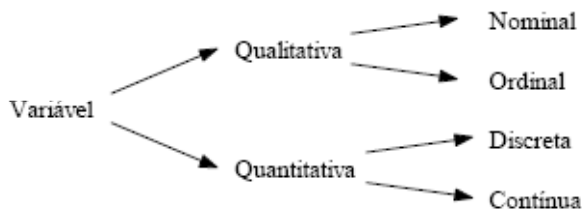
- d. estudar fenômenos numéricos.

1.2.1 - Tipos de variáveis

Variável é o conjunto de resultados possíveis de um fenômeno. Por exemplo:

- fenômeno “sexo” são dois os resultados possíveis: sexo masculino e sexo feminino;
- fenômeno “nº de alunos de uma sala de aula” há um número de resultados possíveis expresso através dos números naturais: 0, 1, 2, 3, ... , n
- para o fenômeno “peso” temos uma situação diferente, pois os resultados podem tomar um número infinito de valores numéricos dentro de um determinado intervalo.

Os dados coletados em uma primeira fase podem ser definidos como variáveis **qualitativas** (quando seus valores são expressos por atributos, p. ex, sexo, cor) ou **quantitativas** (quando os seus valores são expressos em números), de acordo com a seguinte figura:



Ex.: Para uma população de peças produzidas em um determinado processo, poderíamos ter:

Variável	Tipo
Estado: Perfeita ou defeituosa	Qualitativa Nominal
Qualidade: 1 ^a , 2 ^a ou 3 ^a categoria	Qualitativa Ordinal
N ^o de peças defeituosas	Quantitativa Discreta
Diâmetro das peças	Quantitativa Contínua

A variável quantitativa pode ser **contínua**, quando assume qualquer valor entre dois limites (ex: peso, altura, medições), ou pode ser **discreta**, quando só pode assumir valores pertencentes a um conjunto enumerável (ex: nº de filhos, contagem em geral)⁵.

Designamos as variáveis pelas letras latinas, x, y, z, etc. Por exemplo seja uma população (ou amostra) {2,3,4,5,9}, indicando por x a variável relativa ao fenômeno que deu origem à população de resultados acima, temos:

$$x \in \{2,3,4,5,9\}$$

Exercícios Propostos

1. Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas)
 - a. População (ou Universo): alunos de uma escola.
Variável: cor dos cabelos -
 - b. População: casais residentes em uma cidade
Variável: nº de filhos -
 - c. População: as jogadas de um dado
Variável: o ponto obtido em cada jogada -
 - d. População: peças produzidas por certa máquina.
Variável: nº de peças produzidas por hora -
 - e. População: peças produzidas por certa máquina.
Variável: diâmetro externo -
2. Diga quais das variáveis abaixo são discretas e quais são contínuas:
 - a. População (ou Universo): crianças de uma cidade.
Variável: cor dos olhos -
 - b. População: dados de uma estação meteorológica de uma cidade.
Variável: precipitação pluviométrica, durante o ano -
 - c. População: dados da Bolsa de Valores de São Paulo - IBOVESPA.
Variável: número de ações negociadas -
 - d. População: funcionários de uma empresa.
Variável: salários -
 - e. População: pregos produzidos por uma máquina.
Variável: comprimento -

⁵ De modo geral, as medições dão origem a variáveis contínuas e as contagens a variáveis discretas.

- f. População: casais residentes numa cidade.
Variável: sexo dos filhos -
- g. População: propriedades agrícolas do Brasil.
Variável: produção de soja -
- h. População: segmentos de reta.
Variável: comprimento -
- i. População: bibliotecas de Catanduva.
Variável: nº de volumes -
- j. População: aparelhos produzidos em uma linha de montagem.
Variável: nº de defeitos por unidade -
- k. População: indústrias de uma cidade.
Variável: retorno sobre o capital próprio empregado -

1.2.2 – Séries Estatísticas

Um dos objetivos da Estatística é sintetizar os valores que uma ou mais variáveis podem assumir, para que tenhamos uma visão global da variação dessa ou dessas variáveis. E isto ela consegue, inicialmente, apresentando esses valores em tabelas e gráficos.

1.2.2.1 -Tabela

É uma pequena tábua, ou quadro, composta de linhas e colunas, que se resume um conjunto de observações. A construção de uma tabela depende dos dados coletados que serão resumidos e dispostos em forma tabular, o que significa dizer que **são colocados em série e apresentados em quadros ou tabelas**.

Tabela é a disposição gráfica das séries de acordo com uma determinada ordem de classificação.

Título

Cabeçalho

Corpo

Rodapé

As partes principais de uma tabela são:

- Título – conjunto de informações, respondendo às perguntas: O quê? Quando? Onde?
- Cabeçalho – onde é designada a natureza do conteúdo de cada coluna
- Corpo (com as informações organizadas em linhas e colunas);
- Rodapé (com a fonte e eventuais legendas). Fonte se refere à entidade que organizou a Tabela ou forneceu os dados expostos.

Exemplo

Título		
Cabeçalho		
Nível de retinol sérico	n	%
Aceitável ⁽²⁾	89	55,3
Baixo	65	40,4
Deficiente	7	4,3
Total	161	100

Corpo

(1) 24 – 72 meses
(2) Aceitável: 20,0 – 49,9 µg/dl; baixo: 10,0 – 19,9 µg/dl; deficiente: <10,0 µg/dl
Fonte: Prado MS et al., 1995

Séries Estatísticas

Denominamos série estatística toda tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, do local ou da espécie.

Podemos, então, inferir que numa série estatística observamos a existência de três elementos ou fatores: o tempo, o espaço e a espécie.

Conforme varie um dos elementos da série, podemos classificá-la em histórica, geográfica e específica.

Tabela 1 – Série geográfica, espaciais, territoriais ou de localização - cujos dados são dispostos em correspondência com a região geográfica, isto é, variam com o local.

População brasileira de alguns estados – 1970	
Estados	Número de Habitantes
Bahia	7.583.140
Rio de Janeiro	4.794.578
São Paulo	17.958.693
Paraná	6.997.682

Título: População brasileira de alguns estados – 1970
Cabeçalho: Estados, Número de Habitantes
Corpo: Bahia, Rio de Janeiro, São Paulo, Paraná

Tabela 2 – Série específica ou categórica, cujos dados são dispostos de acordo com itens específicos ou categorias, em determinado tempo e local.

Índice de Preços ao Consumidor Semanal (IPC-S), itens com maiores influências positivas, variações percentuais ao mês		
Item	Variação percentual ao mês até	
	08-05-2004	17-05-2004
Mamão da Amazônia (papaia)	23,64	33,34
Tarifa elétrica residencial	1,58	1,61
Batata - inglesa	20,12	15,42
Manga	15,63	23,14
Leite tipo longa vida	3,25	3,25
Tomate	-3,97	9,76
Plano e seguro saúde	0,72	0,72
Cebola	1,69	8,38
Vasodilatador para pressão arterial	2,47	2,63
Açúcar cristal	8,68	8,26

Tabela 3 – Série Temporal ou Cronológica – descrevem os valores da variável, em determinado local, discriminados segundo intervalos de tempo variáveis.

Índice de Preços ao Consumidor Semanal (IPC-S), itens com maiores influências negativas, variações percentuais ao mês		
Item	Variação percentual ao mês até	
	08-05-2004	17-05-2004
Passagem aérea	-6,78	-6,81
Melão	-1,19	-23,01
Laranja-pêra	-1,90	-3,64
Frango em pedaços	2,03	-1,83
Pimentão	3,70	-4,70
Maçã argentina	-6,95	-8,22
Contra-filé	-3,69	-2,93
Corvina	-2,83	-5,12
Óleo de soja	-0,16	-1,45
Maçã nacional	-7,98	-5,41

Fonte: Divisão de Gestão de Dados

Tabela 4 – Série geográfica

Índice de Preços ao Consumidor Semanal (IPC-S), segundo municípios de 12 capitais, variações percentuais ao mês		
Município	Variação percentual ao mês até	
	08-05-2004	17-05-2004
IPC-S	0,52	0,59
Belém	0,59	0,50
Belo Horizonte	1,05	0,99
Curitiba	0,96	1,37
Distrito Federal	0,35	0,54
Florianópolis	0,86	0,85
Fortaleza	0,89	0,98
Goiânia	0,56	1,06
Porto Alegre	0,37	0,58
Recife	0,14	0,36
Rio de Janeiro	0,48	0,59
Salvador	0,59	0,57
São Paulo	0,30	0,27

PROVIBR - Divisão de Censo de Cidades

Como vimos, a tabela nos fornece uma visão e noção geral dos dados pesquisados.

As tabelas 2, 3 e 4 são, também, exemplos de séries temporais.

Contrária a esta tendência, uma corrente mais moderna, liderada por Tukey, utiliza principalmente técnicas visuais, representações pictóricas dos dados, em oposição aos dados numéricos.

1.2.2.2 – Fazendo uma tabela no Excel

A grosso modo, o Excel é uma **matriz** (67 mil linhas por 256 colunas na versão 2003 e de 1 milhão de linhas por 16 mil colunas na versão 2007) e mais um “monte” de ferramentas para manusear os elementos desta matriz⁶

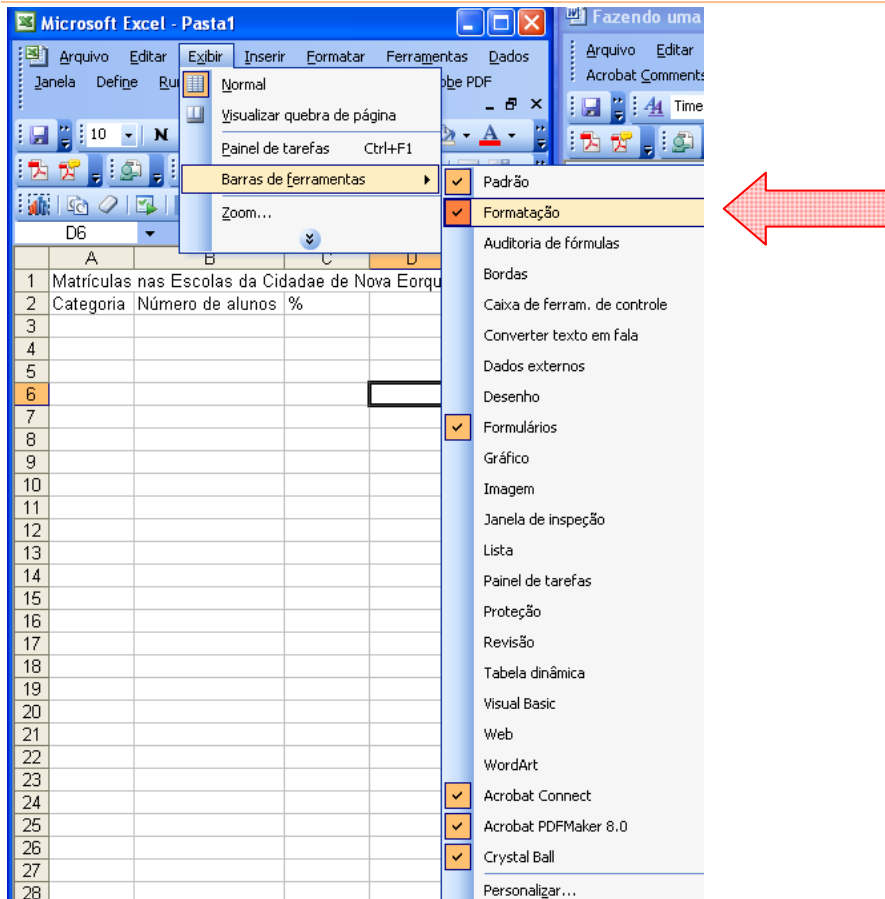
Mão na massa:

1. Clicar na célula A1 digitar o texto: Matrículas nas Escolas da Cidade de Nova Eorque - SP - 2010. Pressione ENTER
2. Clicar na célula A2 e digitar o texto: Categoria. Pressione ENTER
Clicar em B2 e digitar: Número de alunos. Pressione ENTER
Clicar em C2 e digitar: %. Pressione ENTER
3. Fazer um ajuste dando um duplo clique entre as colunas B e C. O cursor muda de símbolo (uma cruz com setas nos dois sentidos na linha horizontal).

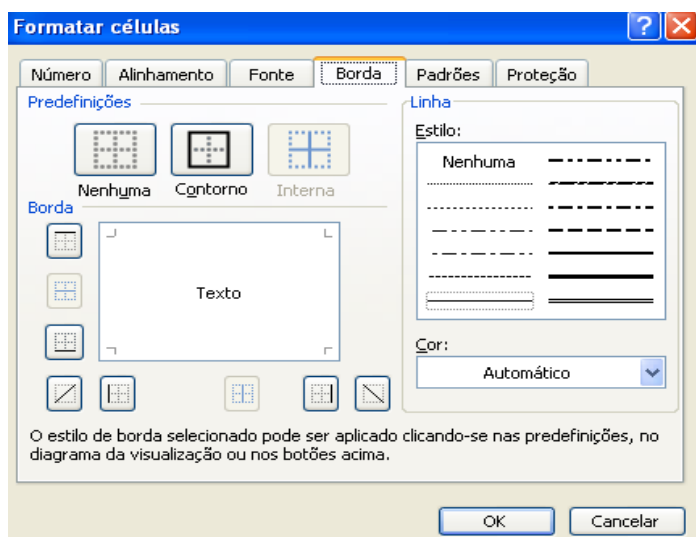
	A	B	C	D	E	F
1	Matrículas nas Escolas da Cidade de Nova Eorque - SP - 20					
2	Categoria	Número de				
3			%			

4. No menu principal em Exibir, selecione Barra de Ferramentas e a seguir Formatação (no menu suspenso). Aparecerá uma série de ícones de ferramentas de formatação agrupados numa barra de ferramentas. Coloque-a num lado da planilha.

⁶ Ver os textos [Básico de Estatística](#), Introdução ao Excel, etc. disponíveis no site do Prof. Bertolo



5. Clicar na célula A1. A célula fica destacada e apresenta um “quadrado” no canto inferior direito. O ponteiro do mouse também mudou de símbolo. Agora aparece uma cruz com linhas largas. Clique em qualquer lugar na célula destacada e arraste até a célula C1. Agora fica destacado um **intervalo de células** chamado A1:C1.
6. Naquela nova barra de ferramentas, clique no botão **Mesclar e Centralizar**. A coisa ainda não ficou boa.
7. Clique outra vez na célula A1 e depois no menu principal selecione **Formatar** e a seguir escolha no menu suspenso a opção **Células**. Aparecerá uma janela:



Você poderia usar um atalho para fazer isso usando as **teclas de atalho CTRL + 1**. Na caixa **Formatar Células** escolher a aba **Alinhamento**. Aparece uma nova janela:



Na seção **Controle de Texto** marque a caixa de seleção **Quebrar texto automaticamente** e depois **OK**.

A seguir clique entre as linhas 1 e 2 e o cursor muda de símbolo (agora aparece uma cruz com a linha vertical apontando nos dois sentidos). Mantendo o botão esquerdo do mouse pressionado, arraste - o até a linha de baixo. O espaçamento entre as linhas fica maior.

8. Clique na célula **A3** e comece a introduzir os valores constantes nas colunas 1 e 2 (apenas) da tabela até a linha 5. A coisa fica *anssim*:

	A	B	C	D	E	F	G
1	Matrículas nas Escolas da Cidade de Nova Eorque - SP - 2010						
2	Categoria	Número de alunos	%				
3	1º grau	19286					
4	2º grau	1681					
5	3º grau	234					
6							
7							
8							
9							
10							
11							
12							
13							
14							

Beleza?

9. Clique em **A6** e digite **Total**.
 10. Clique em **B6** e introduza a fórmula: **= SOMA(B3:B5) ENTER**. Aparece 21201
 11. Clique na célula **C3** e introduza a fórmula: **= B3/B6** (Você poderia clicar nas células desejadas ao invés de digitá-las). Não pressione **ENTER** ainda.

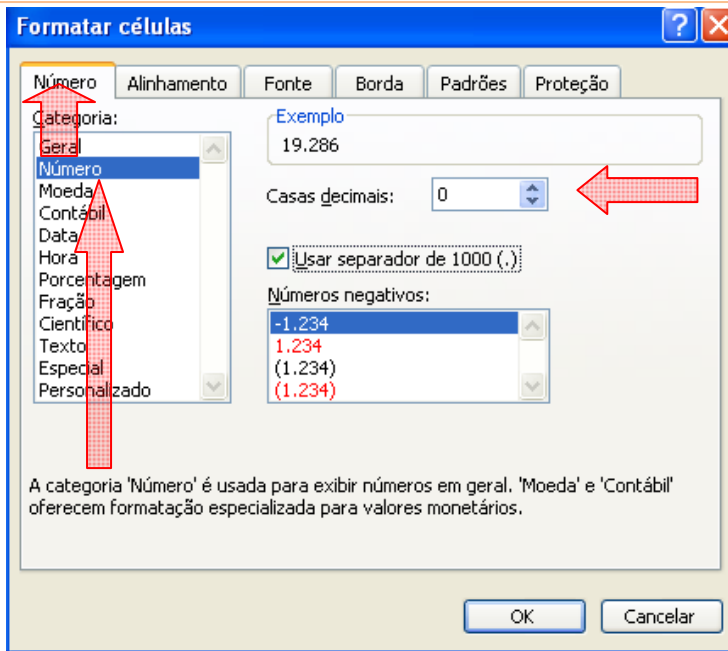
A seguir vá com o cursor à barra de fórmulas e clique em qualquer lugar sobre a célula **B6**.

	A	B	C
1	Matrículas nas Escolas da Cidade de Nova Eorque - SP - 2010		
2	Categoria	Número de alunos	%
3	1º grau	19286	=B3/B6
4	2º grau	1681	
5	3º grau	234	
6	Total	21201	
7			

A seguir pressione a tecla **F4**. Você com isso, “dolarizou” a célula, isto é, tornou-a **Absoluta**.

Agora pressione **ENTER**.

12. Clique no canto direito inferior da célula **C3** (no quadradinho preto) e o mouse muda de símbolo, passando agora a mostrar uma cruzinha de linhas finas. Mantendo o botão esquerdo do mouse pressionado, arraste até a célula **C6** (nesta deverá aparecer 1).
 13. Clique em qualquer lugar na célula **B3** e arraste até a célula **B6** (Não arraste pelo quadradinho agora). Daí vá ao menu **Formatar** e selecione **Células** (ou use a tecla de atalho **CTRL + 1**).



Na janela **Formatar Células**, selecione a aba **Número** e escolha a **Categoria Número**. Configure para zero casas decimais no botão de rotações.

Marque a caixa de seleção **Usar separador de 1000 (.)**.

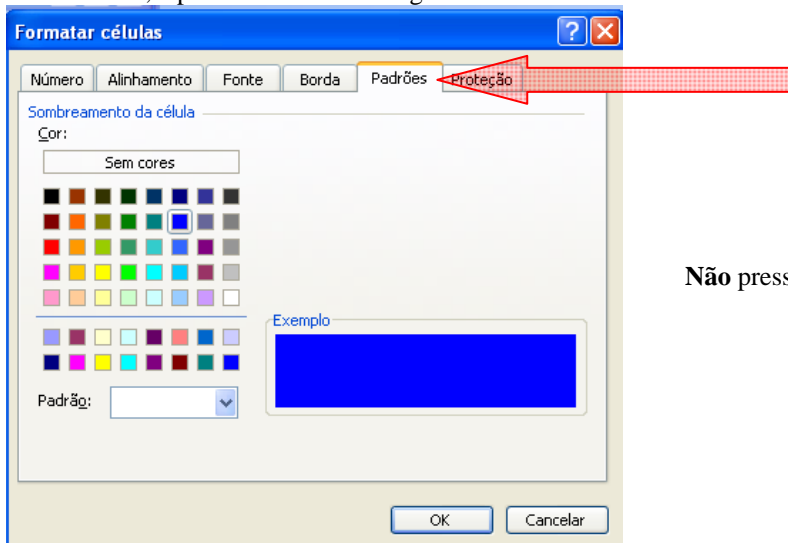
Pressione o botão **OK**.

14. Faça a mesma coisa para as células da coluna de % (porcentagem), escolhendo a **Categoria Porcentagem** na aba **Número**. Configure para 1 casa decimal com o botão de rotações. Pressione **OK** e a coisa ficou *assim*:

	A	B	C
1	Matriculas nas Escolas da Cidade de Nova Eorque - SP - 2010		
2	Categoria	Número de alunos	%
3	1º grau	19.286	91,0%
4	2º grau	1.681	7,9%
5	3º grau	234	1,1%
6	Total	21.201	100,0%
7			

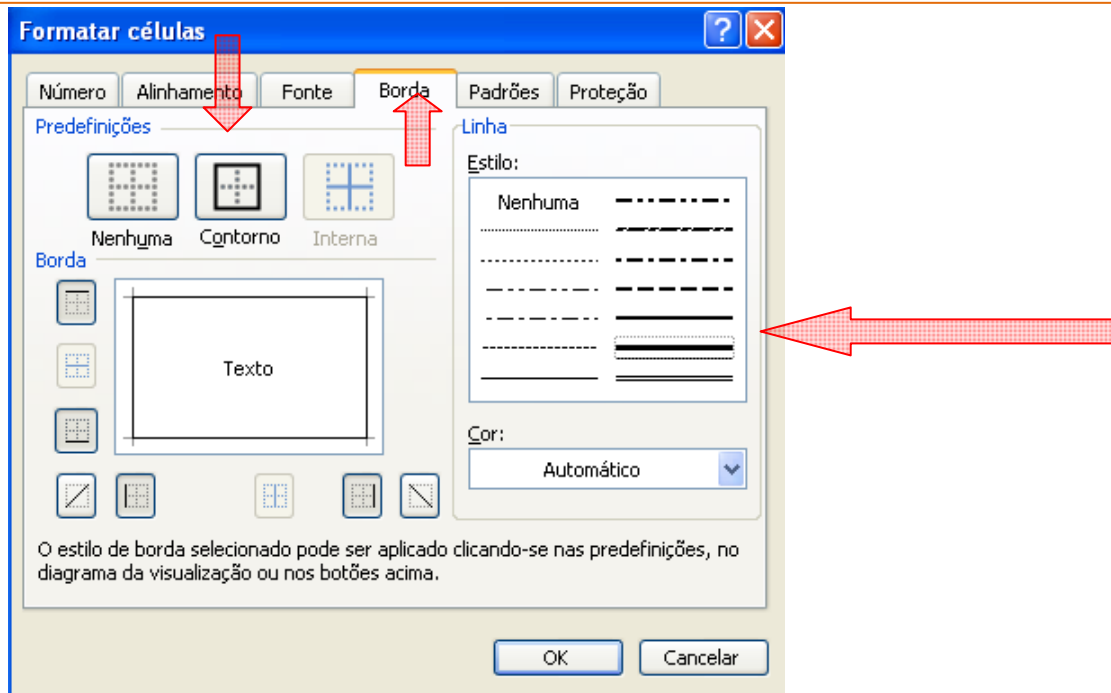
15. Vamos enfeitar o jogue!!!!

Selecione A1, depois **CTRL + 1** e a seguir escolha a aba **Padrões**. Para sombreamento da célula escolha a cor azul.



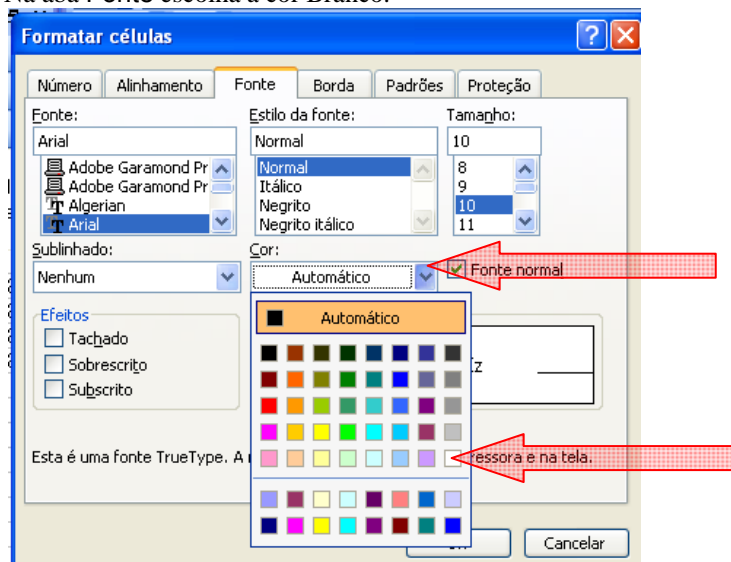
Não pressione **OK**.

Na aba **Borda** selecione **Contorno** e em **Linha** selecione a penúltima da coluna à direita.



Não pressione OK.

Na aba Fonte escolha a cor Branco.



Pressione OK.

16. Clique em A2 e arraste até C2. (Não arraste pelo quadradinho agora)

A seguir CTRL+1.

Na aba Alinhamento, na seção Alinhamento de texto, clique na caixa de listagem Horizontal e selecione Centralizar seleção. Não pressione OK.



Na aba Fonte na seção Estilo da Fonte, selecione Negrito. Não pressione OK.

Na aba Borda selecione Interna e a linha “última da coluna da esquerda”.

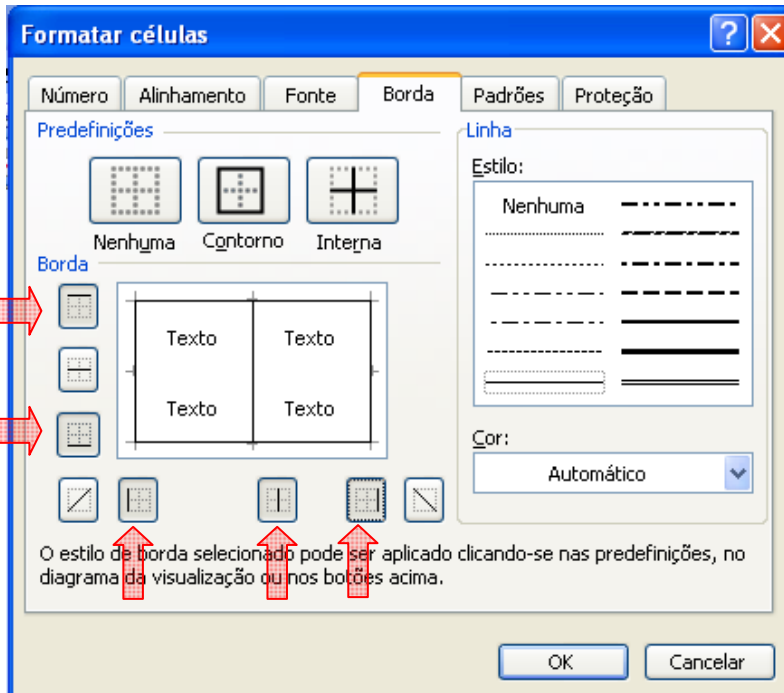
Na aba Padrões selecione a cor azul claro (a última da 6ª coluna a partir da esquerda).

Pressione OK.

17. Clicando entre as colunas B e C reajuste o tamanho da coluna B.

18. Selecione o intervalo A3:C6 e a seguir CTRL + 1.

Na aba Borda selecione os botões:



Pressione

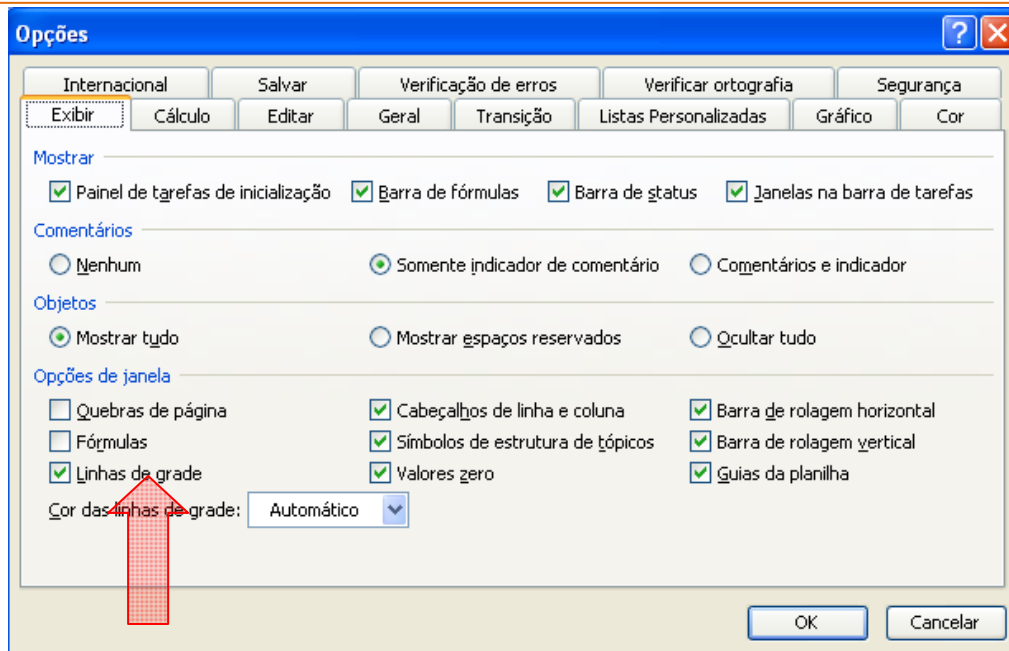
OK.

19. Clique na Célula A7 e digite: Dados fictícios. escolha o tamanho da fonte como 8.

20. Selecione o intervalo A7:C7 e a seguir CTRL + 1.

Na aba Borda, selecione o primeiro botão (borda superior). Escolha a linha estilo (penúltima da coluna da direita). Pressione OK

21. No menu principal selecione Ferramentas, depois selecione Opções... no menu suspenso. Aparece a janela



Escolha a aba Exibir e em Opções de janela desmarque a caixa de seleção Linha de grade. Ficou mais bonitinho.

22. Agora CTRL + C e cole especial (imagem de metarquivo avançado) no Word.

1.2.2.3 – Dados Absolutos e Dados Relativos

Os dados estatísticos resultantes da coleta direta da fonte, sem outra manipulação senão a contagem ou medida, são chamados de **dados absolutos**. Sua leitura é sempre enfadonha e inexpressiva; embora esses dados traduzam um resultado exato e fiel, não têm a virtude de ressaltar de imediato as suas conclusões numéricas. Daí o uso dos dados relativos.

Os **dados relativos** são o resultado de comparações por quocientes (razões ou índices) que se estabelecem entre os dados absolutos, e têm por finalidade realçar ou facilitar as comparações entre quantidades. São traduzidos, em geral, por meio de porcentagens, índices, coeficientes e taxas.

PORCENTAGENS

Considere a série

Matrículas nas Escolas da Cidade de Nova Eorque – SP - 2010

Categoria	Número de Alunos
1º grau	19.286
2º grau	1.681
3º grau	234
Total	21.201

Dados fictícios

Calculemos as porcentagens dos alunos de cada grau:

$$1^\circ \text{ grau} \rightarrow (19.286 \times 100)/21.201 = 90,96 = 91,0$$

$$2^\circ \text{ grau} \rightarrow (1.681 \times 100)/21.201 = 7,92 = 7,9$$

$$3^\circ \text{ grau} \rightarrow (234 \times 100)/21.201 = 1,10 = 1,1$$

Com esses dados, podemos formar uma nova coluna na série em estudo:

Matrículas nas Escolas da Cidade de Nova Eorque – SP - 2010

Categoria	Número de Alunos	%
1º grau	19.286	91,0
2º grau	1.681	7,9
3º grau	234	1,1
Total	21.201	100,0

Dados fictícios

Os valores dessa nova coluna nos dizem que, de cada 100 alunos da cidade A, 91 estão matriculados no 1º grau, 8, aproximadamente, no 2º grau e 1 no 3º grau.

O emprego da porcentagem é de grande valia quando é nosso intuito destacar a participação da parte no todo.

ÍNDICES

São razões entre duas grandezas tais que uma não inclui a outra. Exemplos:

$$\text{Índice cefálico} = \frac{\text{diâmetro transversal do crânio}}{\text{diâmetro longitudinal do crânio}} \times 100$$

$$\text{Quociente Intelectual} = \frac{\text{idade mental}}{\text{idade cronológica}} \times 100$$

$$\text{Densidade demográfica} = \frac{\text{população}}{\text{superfície}}$$

Índices econômicos

$$\text{Produção per capita} = \frac{\text{valor total da produção}}{\text{população}}$$

$$\text{Consumo per capita} = \frac{\text{consumo do bem}}{\text{população}}$$

$$\text{Renda per capita} = \frac{\text{renda}}{\text{população}}$$

$$\text{Receita per capita} = \frac{\text{receita}}{\text{população}}$$

COEFICIENTES

São razões entre o número de ocorrências e o número total (número de ocorrências e número de não-ocorrências). Exemplos:

$$\text{Coeficiente de natalidade} = \frac{\text{número de nascimentos}}{\text{população total}}$$

$$\text{Coeficiente de mortalidade} = \frac{\text{número de óbitos}}{\text{população total}}$$

Coeficientes educacionais:

$$\text{Coeficiente de evasão escolar} = \frac{\text{número de alunos evadidos}}{\text{número inicial de matrículas}}$$

$$\text{Coeficiente de aproveitamento escolar} = \frac{\text{número de alunos aprovados}}{\text{número final de matrículas}}$$

$$\text{Coeficiente de recuperação escolar} = \frac{\text{número de alunos recuperados}}{\text{número de alunos em recuperação}}$$

TAXAS

São os coeficientes multiplicados por uma potência de 10 (10, 100, 1.000, etc.) para tornar o resultado mais inteligível. Exemplo:

Taxas de mortalidade = coeficiente de mortalidade x 1.000

Taxa de natalidade = coeficiente de natalidade x 1.000

Taxa de evasão escolar = coeficiente de evasão escolar x 100

EXEMPLOS

1. A cidade de Catanduva apresentou 73.398 matrículas na 1ª série, no início do ano de 2009, e 68.381 no final do ano. A cidade de Nova Caputira apresentou, respectivamente, 43.612 e 41.247 matrículas. Qual cidade que apresentou maior evasão escolar?

Catanduva → taxa de evasão escolar = coeficiente de evasão escolar x 100 =

$$\text{Coeficiente de evasão escolar} = \frac{\text{número de alunos evadidos}}{\text{número inicial de matrículas}} = \frac{73.398 - 68.381}{73.398} \times 100 = 6,84\%$$

Nova Caputira → taxa de evasão escolar = coeficiente de evasão escolar x 100 =

$$\text{Coeficiente de evasão escolar} = \frac{\text{número de alunos evadidos}}{\text{número inicial de matrículas}} = \frac{43.612 - 41.247}{43.612} \times 100 = 5,42\%$$

A cidade que apresentou maior evasão escolar foi Catanduva (motivo? Presídio?)

Exercícios Propostos

1. Uma escola registrou em março, na 1ª série, a matrícula de 40 alunos e a matrícula efetiva, em dezembro, de 35 alunos. Qual foi a taxa de evasão? Resp: 12,5%
2. Calcule a taxa de aprovação de um professor de uma classe de 45 alunos, sabendo que obtiveram aprovação 36 alunos. Resp: 80%.
3. Considere a série estatística:

Séries	Número de Matriculados	%
1ª	546	
2ª	328	
3ª	280	
4ª	120	
Total	1.274	100,0

Complete-a, determinando as porcentagens com uma casa decimal e fazendo a compensação, se necessário.

4. Uma escola apresentava, no final do ano, o seguinte quadro:

Séries	Número de Matriculados	
	MARÇO	NOVEMBRO
1ª	480	475
2ª	458	456
3ª	436	430
4ª	420	420
Total	1.794	1.781

- a. Calcule a taxa de evasão, por série.
 - b. Calcule a taxa de evasão da escola.
5. Considere a tabela abaixo:

**EVOLUÇÃO DAS RECEITAS
DO CAFÉ INDUSTRIALIZADO
JAN./ABR. - 2009**

Meses	Valor (US\$ milhões)
Janeiro	33,3
Fevereiro	54,1
Março	44,5
Abril	52,9
Total	184,8

Dados fictícios

- Complete-a com uma coluna de taxas percentuais.
 - Como se distribuem as receitas em relação ao total?
 - Qual o desenvolvimento das receitas de um mês para o outro?
 - Qual o desenvolvimento das receitas em relação ao mês de janeiro?
- São Paulo tinha, em 1989, uma população projetada de 32.361.700 habitantes. Sabendo que sua área terrestre é de 248.256 km², calcule a sua densidade demográfica.
 - Considerando que Minas Gerais, em 1988, apresentou (dados fornecidos pelo IBGE):
 - População projetada: 15.345.800 habitantes;
 - Superfície: 586.624 km²;
 - Nascimentos: 337.859;
 - Casamentos: 110.473.
 Calcule:
 - O índice de densidade demográfica;
 - A taxa de natalidade;
 - A taxa de nupcialidade.
 - Uma frota de 40 caminhões, transportando, cada um, 8 toneladas, dirige-se a duas cidades A e B. Na cidade A são descarregados 65% desses caminhões, por 7 homens, trabalhando 7 horas. Os caminhões restantes seguem para a cidade B, onde 4 homens gastam 5 horas para o seu descarregamento. Em que cidade se obteve melhor produtividade?
 - Um professor preencheu um quadro, enviado pela Delegacia de Ensino, com os seguintes dados:

Série e Turma	Nº de Alunos 30.03	Nº de Alunos 30.11	Promovidos sem recuperação	Retidos sem recuperação	Em Recuperação	Recuperados	Não-Recuperados	Total Geral	
								Promovidos	Retidos
1º B	49	44	35	03	06	05	01	40	04
1º C	49	42	42	00	00	00	00	42	00
1º E	47	35	27	00	08	03	05	30	05
1º F	47	40	33	06	01	00	01	33	07
Total	192	161	137	09	15	08	07	145	16

Calcule:

- A taxa de evasão, por classe;
- A taxa de evasão total;
- A taxa de aprovação, por classe;
- A taxa de aprovação geral;
- A taxa de recuperação, por classe;
- A taxa de recuperação geral;
- a taxa de reprovação na recuperação geral;
- a taxa de aprovação, sem recuperação;
- a taxa de retidos, sem recuperação.

1.2.2.3 - Gráficos estatísticos

Os gráficos também são uma forma de apresentação dos dados coletados na pesquisa. O gráfico precisa ter clareza, simplicidade e veracidade.

Existem vários tipos de gráficos: gráfico de segmento, de setor, de barras e de colunas.

A palavra, gráfico, refere-se à grafia, ou às artes gráficas, ou ao que delas se ocupa. Um gráfico pode ser representado por desenho ou figuras geométricas.

Os principais gráficos serão sucintamente descritos a seguir. É importante lembrar que os modernos programas computacionais de Edição de Texto, Planilha Eletrônica e Banco de Dados facilitam em muito a manipulação com gráficos. Dentre esses programas o destaque é o Excel.

A representação gráfica de um fenômeno de vê obedecer a certos requisitos fundamentais, para ser realmente útil:

- Simplicidade – o gráfico deve ser destituído de detalhes, bem como traços desnecessários.
- Clareza – deve possibilitar uma correta interpretação dos valores representativos do fenômeno em estudo.
- Veracidade – deve expressar a verdade sobre o fenômeno em estudo.

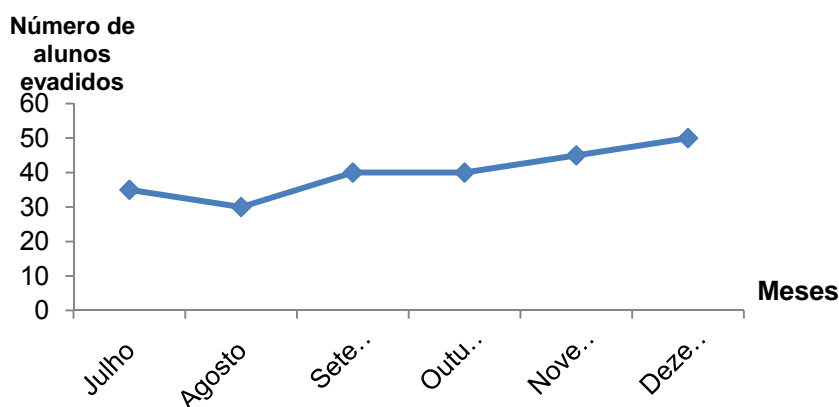
Os gráficos dividem-se nos seguintes tipos: *diagramas*, *cartogramas* e *pictogramas*.

Gráfico de Linha

A tabela seguinte mostra o número de alunos evadidos de uma determinada escola de Ensino Médio no segundo semestre de 2002 (uma série temporal, cujos dados são dispostos de acordo com o tempo).

Meses do 2º semestre	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
Nº de alunos evadidos	35	30	40	40	45	50

De acordo com a tabela, construiremos um gráfico de segmentos cujos pares ordenados serão (julho, 35), (agosto, 30), (setembro, 40), (outubro, 40), (novembro, 45) e (dezembro, 50).



Os gráficos de linhas são muito utilizados para mostrar a evolução durante um certo período (*séries temporais*). O gráfico permite visualizar muito bem o crescimento, o decréscimo ou a estabilidade do objeto a ser analisado.

Exemplo

O gráfico nos mostra a mortalidade por tuberculose na cidade e no estado de São Paulo, de 1986 a 1999, em número de mortes por cem mil habitantes (gráfico publicado na *Folha de S. Paulo*, de 27 de fevereiro de 2000):



Gráfico de Linhas Comparativas

Estimativa baseada nos números do 1º semestre/99

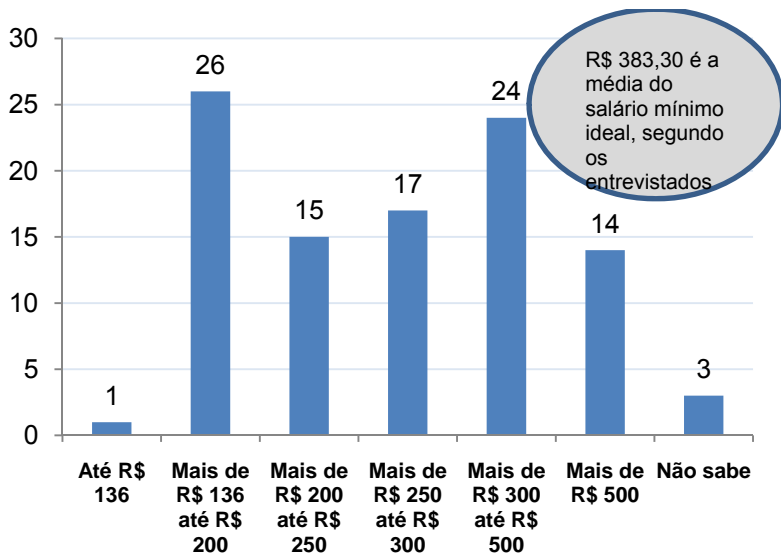
Gráfico de barras

No geral, é usado para comparar as freqüências de mesma variável em um determinado momento.

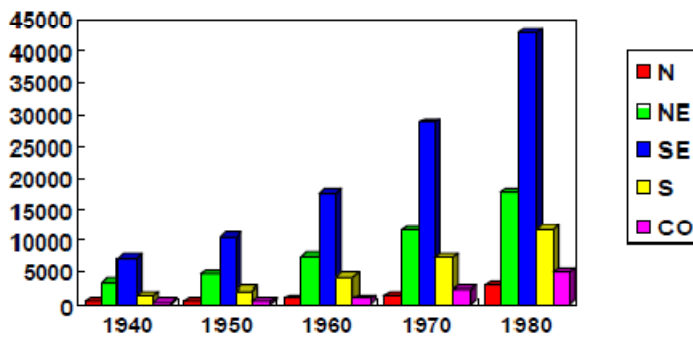
Freqüência é o número de vezes que cada resposta ou cada intervalo de resposta aparece na tabela do pesquisador, a qual pode ter sido feita por amostragem.

Exemplo

A opinião dos paulistanos sobre o salário mínimo (em porcentagem)



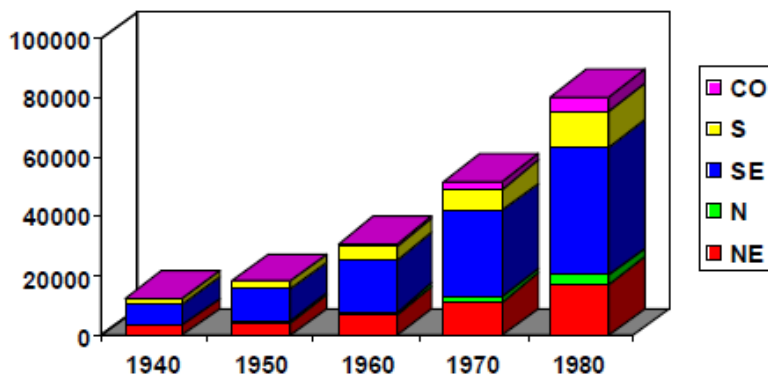
População Urbana do Brasil por Região de 1940 a 1980 (x 1000)



Colunas Justapostas

Fonte: Anuário Estatístico (1984)

População Urbana do Brasil por Região de 1940 a 1980 (x 1000)

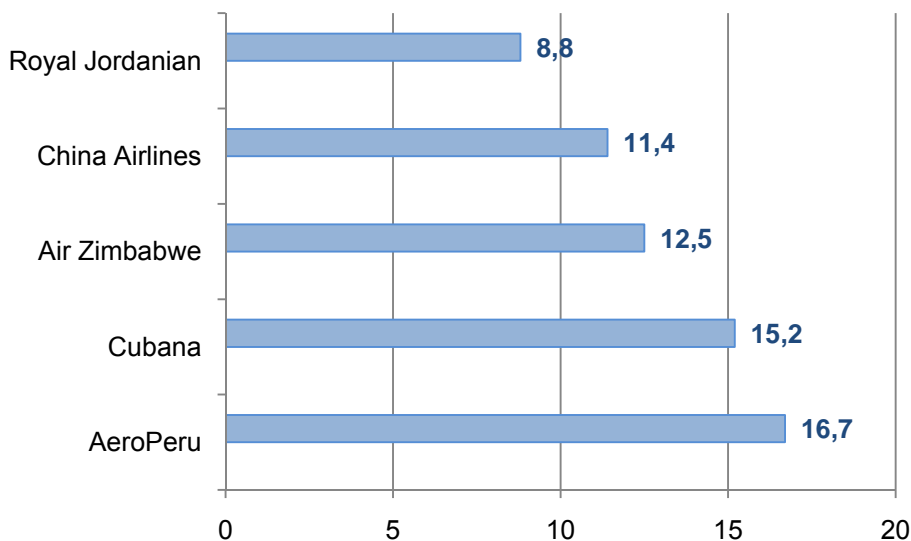


Colunas Sobrepostas

Fonte: Anuário Estatístico (1984)

Gráfico de barras horizontais

Número de acidentes por milhão de decolagens



Revista Veja, São Paulo: Abril, ano 31, n.º 36, 9 set. 1

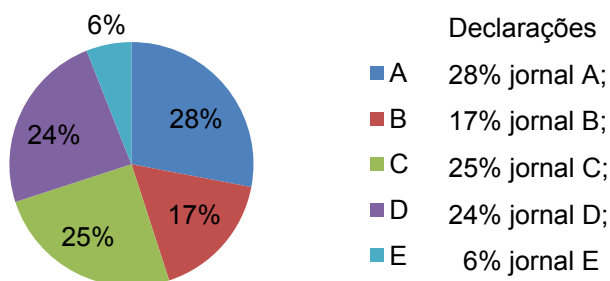
Gráfico de setores (Pie Charts)

O gráfico de setores também é muito usado e faz enxergar melhor as partes do todo (representadas em porcentagens). Para ser traçado, é necessário saber que o círculo todo tem 360°, o que corresponde a 100%. De acordo com a pesquisa, o círculo será repartido, efetuando-se uma regra de três simples.

Para a construção do gráfico de setores, são necessários a pesquisa, a tabela, a fonte, o compasso, o transferidor e as cores.

Exemplo

O gráfico de setores a seguir foi construído com os resultados de uma pesquisa que perguntou a 1.500 pessoas qual jornal diário elas mais gostam de ler:



Sabemos que, para calcularmos o número de pessoas, e para determinar o ângulo do setor circular, segundo a pesquisa, basta resolver uma regra de três simples:

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ A \text{ ----- } 28\% \end{array}$$

$$A = \frac{1500 \times 28}{100} = 420 \text{ pessoas}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ B \text{ ----- } 17\% \end{array}$$

$$B = \frac{1500 \times 17}{100} = 255 \text{ pessoas}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ E \text{ ----- } 6\% \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ A^\circ \text{ ----- } 28\% \\ A^\circ = 100,8^\circ \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ B^\circ \text{ ----- } 17\% \\ B^\circ = 61,2^\circ \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ E^\circ \text{ ----- } 6\% \end{array}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ C \text{ ----- } 25\% \end{array}$$

$$C = \frac{1500 \times 25}{100} = 375 \text{ pessoas}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ D \text{ ----- } 24\% \end{array}$$

$$D = \frac{1500 \times 24}{100} = 360 \text{ pessoas}$$

$$E^\circ = 21,6^\circ$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ C^\circ \text{ ----- } 25\% \\ C^\circ = 90^\circ \end{array}$$

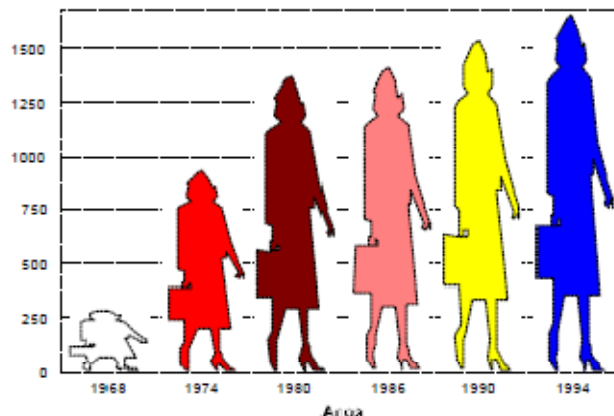
$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ D^\circ \text{ ----- } 24\% \\ D^\circ = 86,4^\circ \end{array}$$

$$E = \frac{1500 \times 6}{100} = 90 \text{ pessoas}$$

Gráfico Pictorial - Pictograma

O gráfico de pictorial tem por objetivo despertar a atenção do público em geral, muito desses gráficos apresentam grande dose de originalidade e de habilidade na arte de apresentação dos dados.

Evolução da matrícula no Ensino Superior no Brasil de 1968 a 1994 (x 1000)



Fonte: Grandes números da educação brasileira março de 1996

1.2.2.4 – Como Fazer Gráficos no Excel?

O Microsoft Office Excel oferece suporte para vários tipos de gráficos com a finalidade de ajudar a exibir dados de maneiras que sejam significativas para a audiência. Quando desejar criar um gráfico ou modificar um gráfico existente, você poderá escolher uma ampla gama de subtipos de gráficos disponíveis para cada um dos seguintes tipos de gráficos:

[Gráficos de colunas](#) [Gráficos de linhas](#) [Gráficos de pizza](#) [Gráficos de barras](#)
[Gráficos de área](#) [Gráficos de dispersão \(XY\)](#) [Gráficos de ações](#)
[Gráficos de superfície](#) [Gráficos de rosca](#) [Gráficos de bolhas](#) [Gráficos de radar](#)

Para se construir um gráfico no Excel devemos primeiro construir uma tabela como fizemos nas seções 1.2.2.2 e 1.2.2.3. Para este exemplo consideremos a tabela da seção 1.2.2.3 da página 20.

A tabela seguinte mostra o **número de alunos evadidos** de uma determinada escola de Ensino Médio (Escola Capitão Bertolo) no **segundo semestre de 2009** (uma *série temporal*, cujos dados são dispostos de acordo com o tempo).

Meses do 2º semestre	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
Nº de alunos evadidos	35	30	40	40	45	50

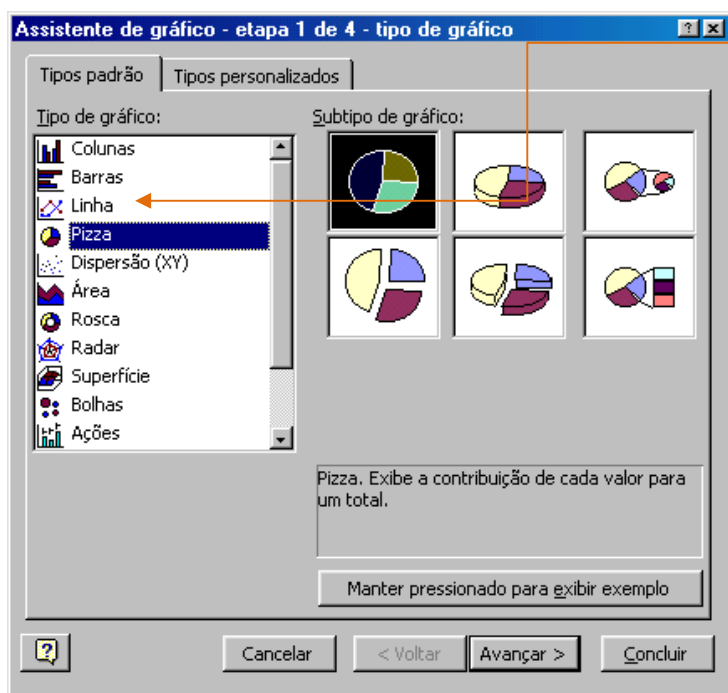
- Faça isto no Excel.
- Abra uma nova planilha e na célula A1 e digite o cabeçalho: Escola Capitão Bertolo.
- Na célula A2 digite: Meses do 2º semestre
- Na célula B2 digite: Julho.
- Clique na célula B2 e a seguir com o ponteiro do mouse sobre o quadradinho do canto inferior direito, arraste adiante pela linha até completar o mês de Dezembro.
- Na célula A3 digite: Nº de alunos evadidos.
- Digite os valores da tabela a partir da célula B3.
- Formate a tabela (lembre-se das teclas de atalho CTRL + 1) para ficar com o aspecto:

Escola Capitão Bertolo						
Meses do 2º semestre	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
Nº de alunos evadidos	35	30	40	40	45	50

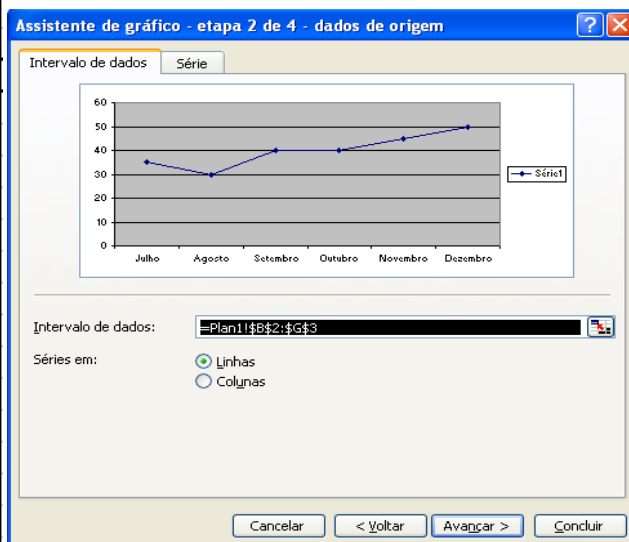
De acordo com a tabela, construiremos um gráfico de segmentos cujos pares ordenados serão (julho, 35), (agosto, 30), (setembro, 40), (outubro, 40), (novembro, 45) e (dezembro, 50).

Para isso, selecione o intervalo B2:G3. No menu principal clique em Inserir e, no menu suspenso selecione Gráfico.... Na barra

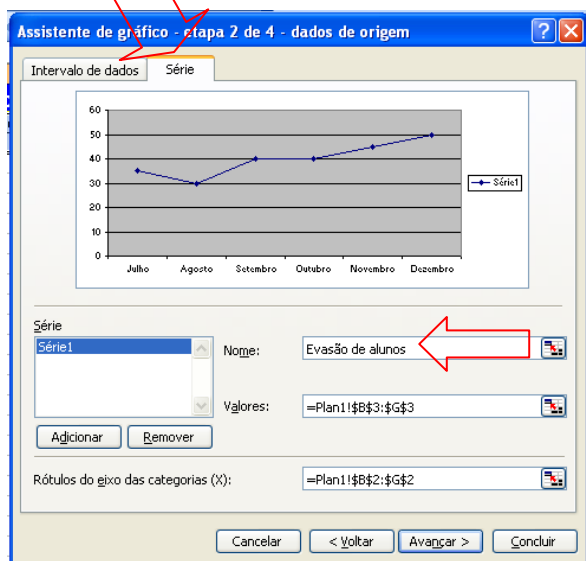
de ferramentas, você poderia ter escolhido o botão com ícone . Aparecerá uma janela [Assistente de gráfico – etapa 1 de 4 – tipo de gráfico](#).



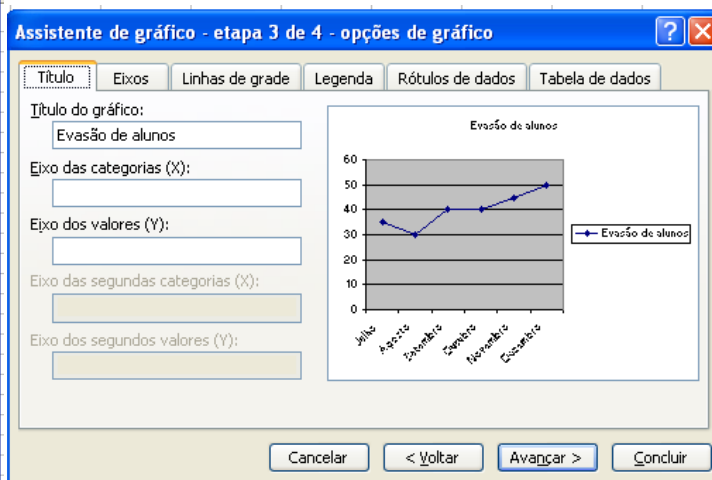
Na seção **Tipo de gráfico**, selecione **Linha** e, na seção **Subtipo de gráfico** escolha o primeiro da 2ª linha. Clique em **Avançar**. Aparecerá a Janela etapa 2:



Na aba **série**, clique na caixa **Nome**: e digite: **Evasão de alunos**.



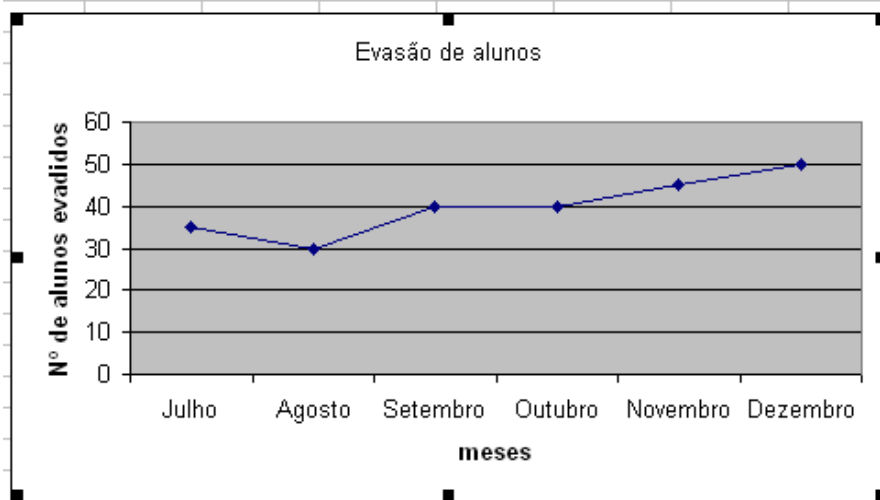
A seguir clique em **Avançar**. Aparecerá a janela da etapa 3:



Selecione a aba **Título** e na caixa **Eixo das categorias (X)**, digite: **meses**.

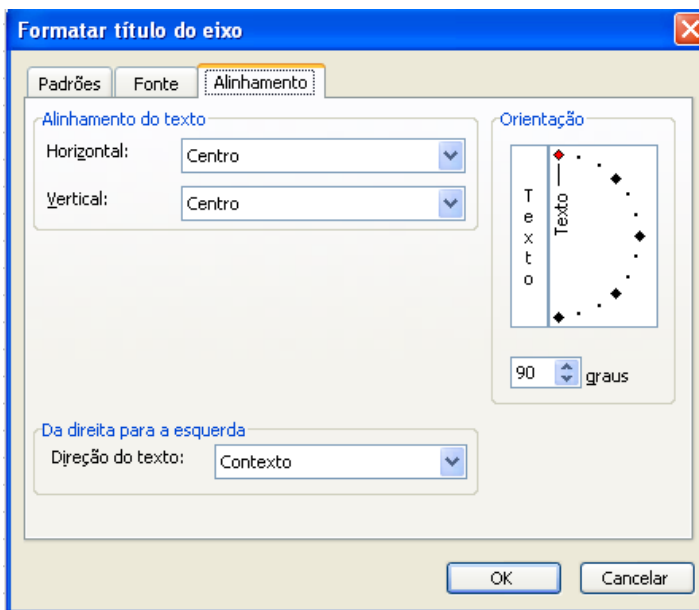
Na caixa **Eixo dos valores (Y)**, digite: **Nº de alunos evadidos**.

Vejam que podemos formatar os **Eixos**, **Linha de grade**, **Legenda**, **Rótulos de dados** selecionando a aba correspondente. Clique em **Avançar**, na etapa 4, selecione o botão **Como objeto em:** (Plan1) e depois **Concluir**.

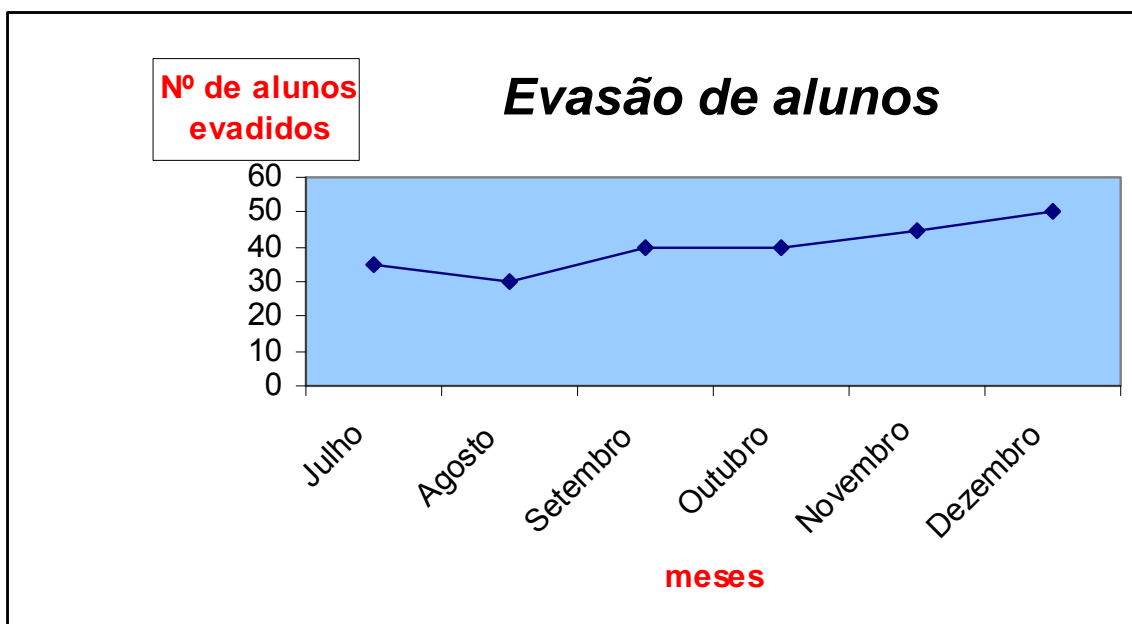


A maravilha está pronta. Se você não quiser a **Legenda** (Evasão de alunos) que está à direita, basta clicar sobre ela e deletá-la.

Você pode Formatar o **Padrão**, **Fonte** e **Alinhamento** do **Título dos eixos**. Clique no Título do eixo com o botão direito e selecione: Formatar título de eixo... E aparecerá a janela:



Usando Formatação procure deixar o gráfico como segue:

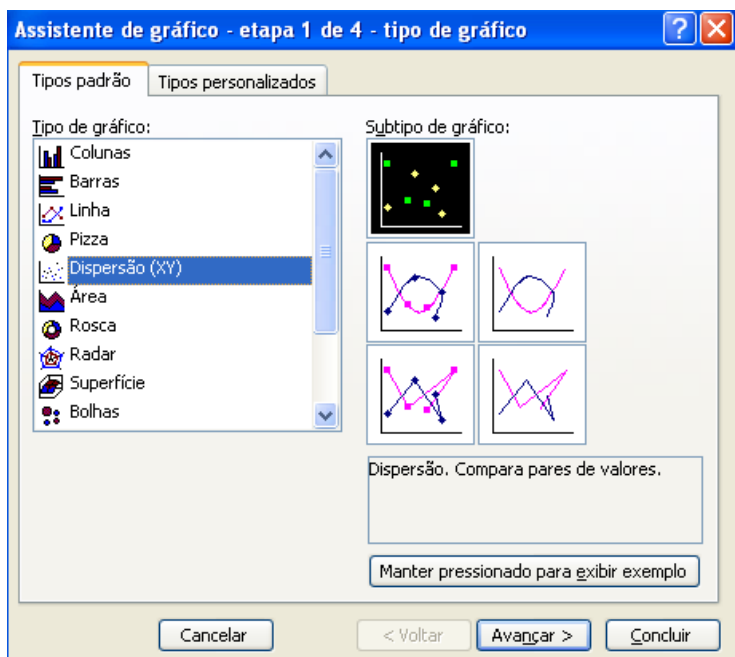


Os gráficos de linhas são muito utilizados para mostrar a evolução durante certo período (**séries temporais**).

O gráfico permite visualizar muito bem o crescimento, o decréscimo ou a estabilidade do objeto a ser analisado.

Gráfico de Linhas Comparativas

Um gráfico de linhas comparativas pode ser feito escolhendo-se como Tipo padrão **Dispersão (XY)**:

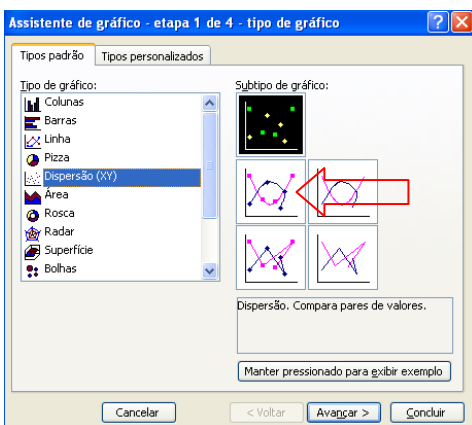


Para exemplificar vamos construir um gráfico mostrando a **mortalidade por tuberculose** na cidade e no estado de São Paulo, de 1986 a 1999, em número de mortes por 100 mil habitantes (gráfico publicado na Folha de S. Paulo, de 27 de Fevereiro de 2.000).

Construa a Tabela:

Mortalidade por tuberculose na cidade e estado de São Paulo - 1986 a 1999 - Nº de emortes por 100 mil habitantes		
Ano	Estado	Cidade
1986	3,92	4,53
1987	4,03	4,57
1988	4,33	5,84
1989	4,12	6,03
1990	4,54	5,83
1991	4,55	6,53
1992	4,21	5,86
1993	4,04	6,23
1994	4,21	6,43
1995	4,66	6,23
1996	4,32	7,04
1997	4,44	8,02
1998	4,05	6,06
1999	4,51	6,49

Depois escolhendo o Subtipo (1ª coluna e 2ª linha)



Construa um gráfico como este:

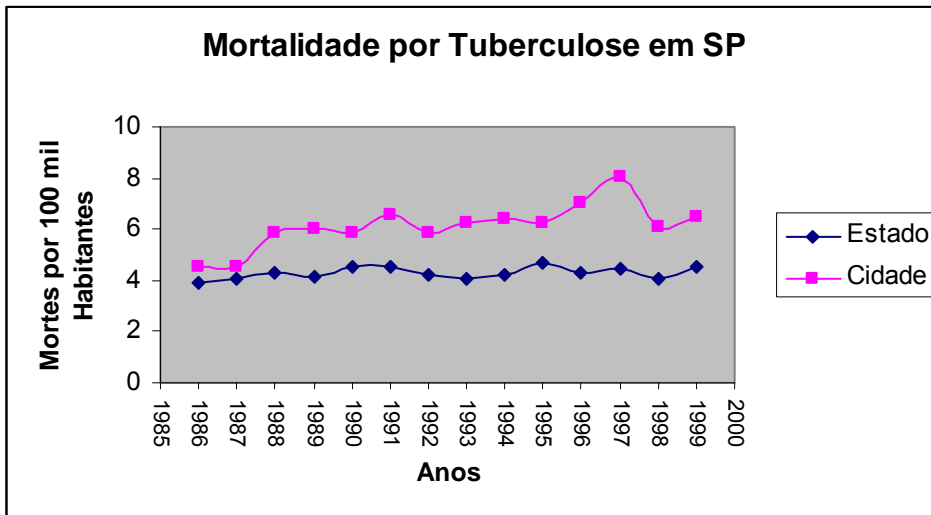
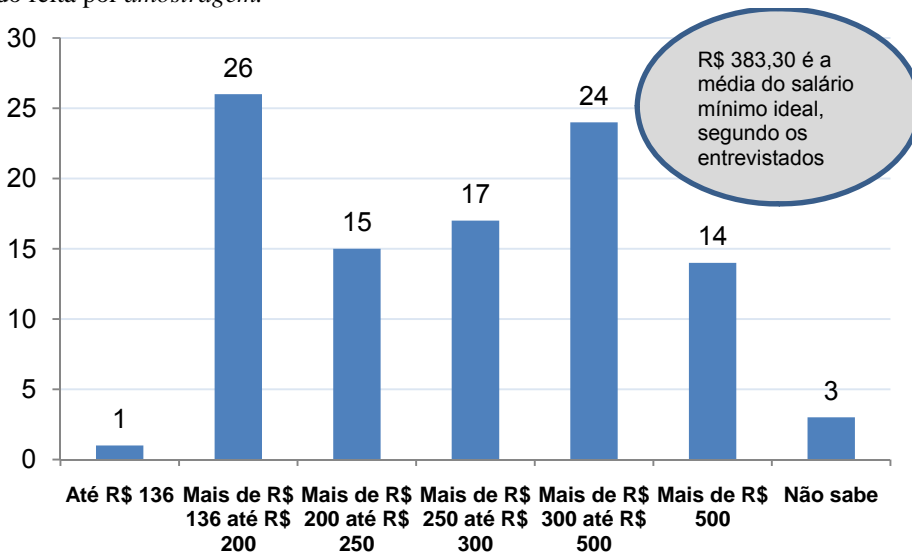


Gráfico de Barras

No geral, é usado para comparar as frequências de mesma variável em um determinado momento.

Frequência é o número de vezes que cada resposta ou cada intervalo de resposta aparece na tabela do pesquisador, a qual pode ter sido feita por amostragem.



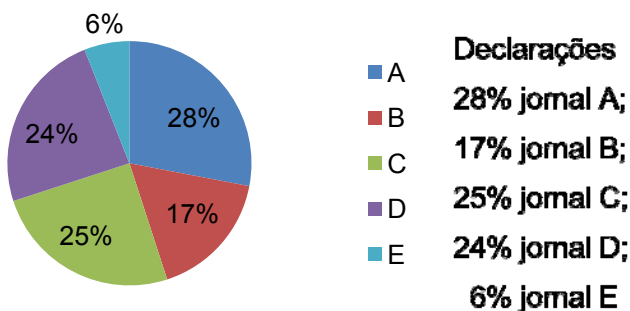
A opinião dos paulistanos sobre o salário mínimo (em porcentagem)

Gráfico de Setores (Pie Charts)

O gráfico de setores também é muito usado e faz enxergar melhor as partes do todo (representadas em porcentagens). Para ser traçado, é necessário saber que o círculo todo tem 360°, o que corresponde a 100%. De acordo com a pesquisa, o círculo será repartido, efetuando-se uma regra de três simples.

Para a construção do gráfico de setores, são necessários a **pesquisa**, a **tabela**, a **fonte**, o **compasso**, o **transferidor** e as **cores**.

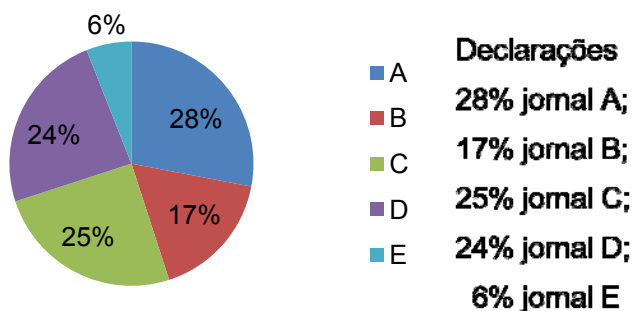
Exemplo - O gráfico de setores a seguir foi construído com os resultados de uma pesquisa que perguntou a 1.500 pessoas qual jornal diário elas mais gostam de ler:



Um EXEMPLO para esquentar

Exercícios de Aplicação

Exemplo - O gráfico de setores a seguir foi construído com os resultados de uma pesquisa que perguntou a 1.500 pessoas qual jornal diário elas mais gostam de ler:



Sabemos que, para calcularmos o número de pessoas, e para determinar o ângulo do setor circular, segundo a pesquisa, basta resolver uma regra de três simples:

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ A \text{ ----- } 28\% \\ 420 \text{ pessoas} \\ A^\circ = 100,8^\circ \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ A^\circ \text{ ----- } 28\% \end{array}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ B \text{ ----- } 17\% \\ B^\circ = 61,2^\circ \\ 255 \text{ pessoas} \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ B^\circ \text{ ----- } 17\% \end{array}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ C \text{ ----- } 25\% \\ C^\circ = 90^\circ \\ 375 \text{ pessoas} \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ C^\circ \text{ ----- } 25\% \end{array}$$

$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ D \text{ ----- } 24\% \\ D^\circ = 86,4^\circ \\ 360 \text{ pessoas} \end{array}$$


$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ D^\circ \text{ ----- } 24\% \end{array}$$

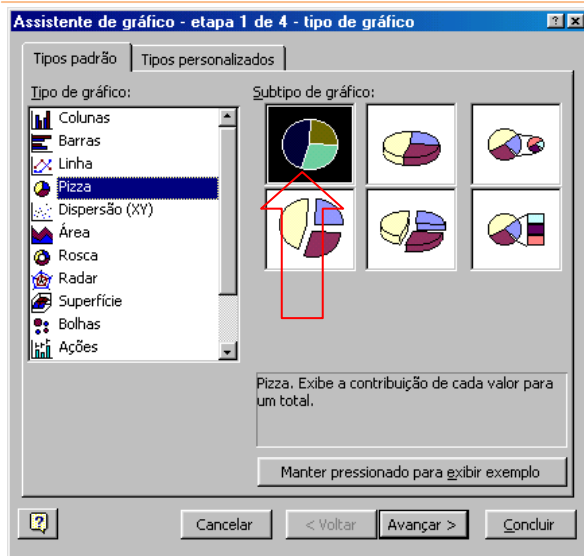
$$\begin{array}{l} 1.500 \text{ ----- } 100\% \\ E \text{ ----- } 6\% \\ E^\circ = 21,6^\circ \quad 90 \text{ pessoas} \end{array}$$

$$\begin{array}{l} 360^\circ \text{ ----- } 100\% \\ E^\circ \text{ ----- } 6\% \end{array}$$

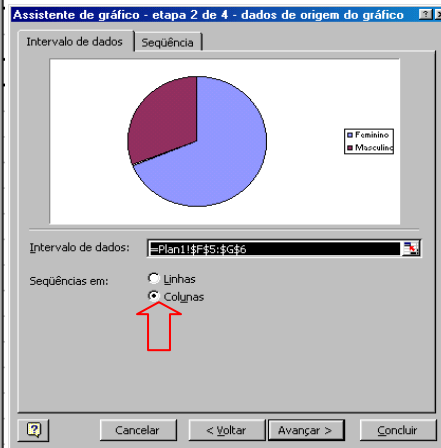
Primeiramente construa a Tabela:

Pesquisa de 1.500 leitores diário de jornal	
Jornal	Declarações
A	28%
B	17%
C	25%
D	24%
E	6%

Selecione na tabela o intervalo que se quer representar, neste caso, o intervalo de células B3:B7. Escolha a opção gráfico . Na janela Tipo de gráfico, escolha a alternativa Pizza e selecionar o Subtipo de gráfico, como mostrado.

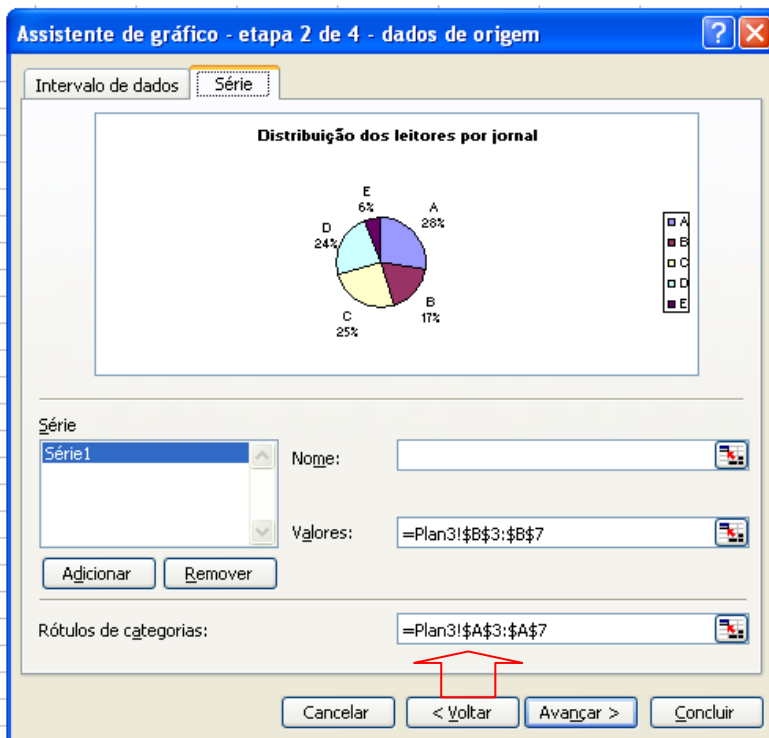


Marque a opção **Seqüência em:** Colunas e depois **Avançar**



Não pressione **Avançar** ainda. Tá com pressa?

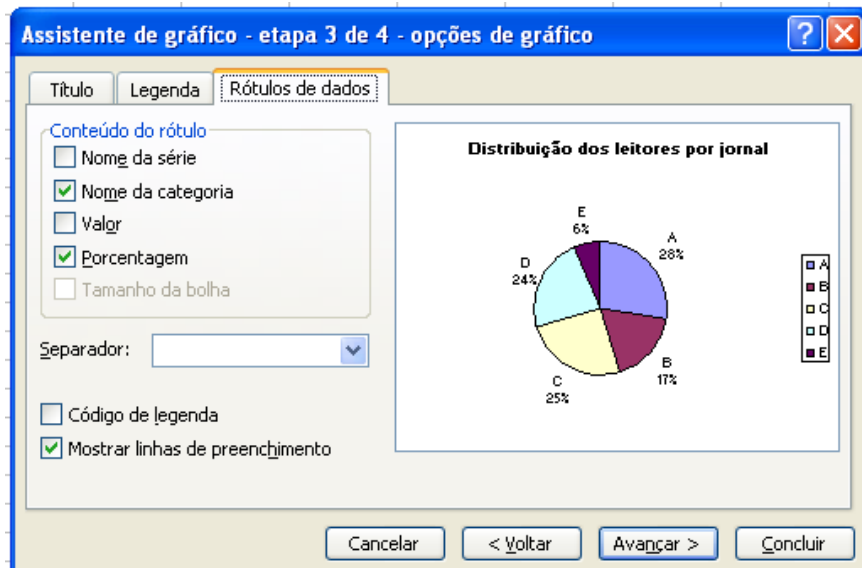
Clique na aba **Série** (ou **Seqüência**) e em **Rótulos de categorias** selecione o intervalo de células A3:A7. A Janela ficará *assim*:



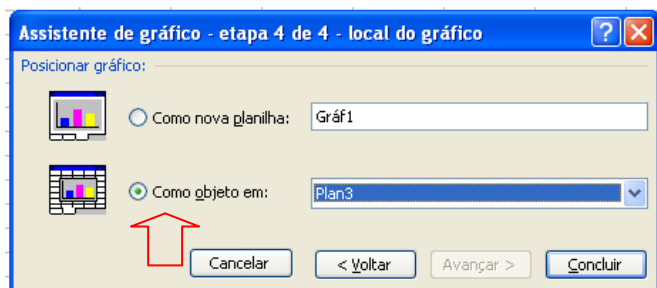
Agora sim, clique em **Avançar** e na aba **Título** e digite em **Título do gráfico:** Distribuição dos leitores por jornal

Se desejamos excluir a legenda que é feita automaticamente devemos abrir a aba **Legenda** e clicar em **Mostrar legenda**, apagando a marca de inclusão.

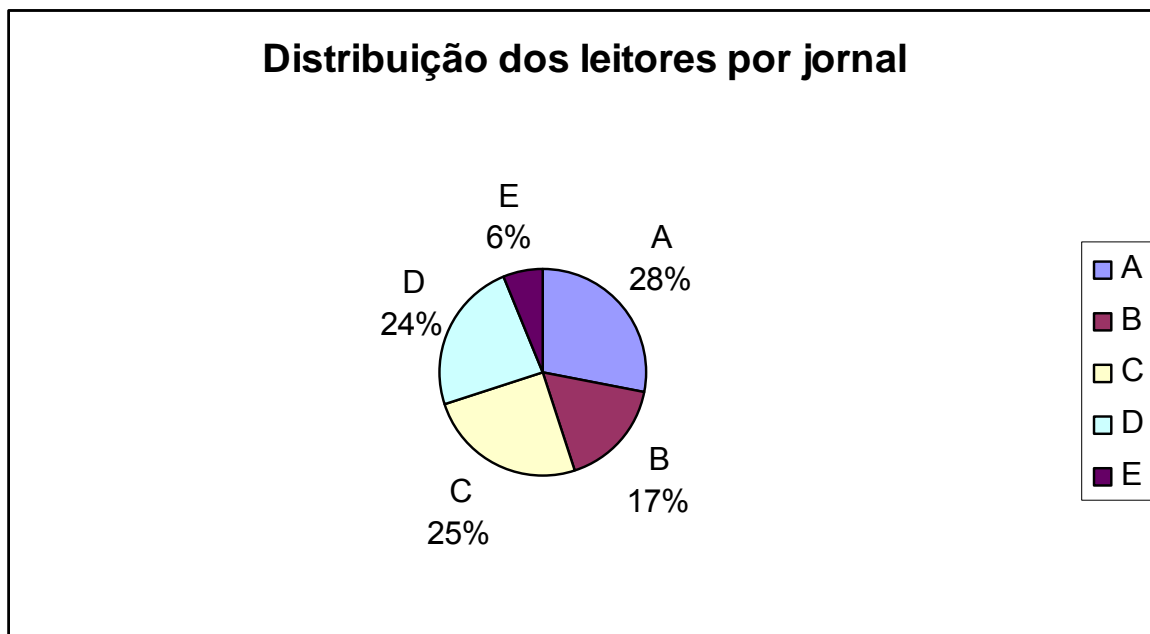
Clique na aba **Rótulo de dados** e na secção **Conteúdo do rótulo** marque as caixas como a figura



Avançar e



O BONITINHO fica assim:



Exercícios Propostos

1. Represente a série abaixo usando o gráfico em linha e demarcando a área de excesso.

COMÉRCIO EXTERIOR BRASIL – 1979-88

Anos	Quantidade (1.000 t)	
	Exportação	Importação
1979	98.010	75.328
1980	109.100	71.855
1981	123.994	64.066
1982	119.990	60.718
1983	178.790	55.056
1984	141.737	53.988
1985	146.351	48.870
1986	133.832	60.605
1987	142.382	61.975
1988	169.396	58.085

2. Represente as tabelas usando o gráfico em colunas

CHEGADA DE VISITANTES BRASIL - 1983 - 86

ANOS	Número (Milhares)
1983	1.420
1984	1.596
1985	1.736
1986	1.934

ENTREGA DE GASOLINAS PARA CONSUMO BRASIL - 1985 - 88

ANOS	Quantidade (1.000 m ²)
1985	9.793
1986	11.095
1987	9.727
1988	9.347

3. Usando o gráfico de barras, represente as tabelas:

PRODUÇÃO DE OVOS DE GALINHA BRASIL - 1988

REGIÃO	Quantidade (1.000 dúzias)
Norte	66.092
Nordeste	356.810
Sudeste	937.463
Sul	485.098
Centro-Oeste	118.468

CONSUMO DE BORRACHA NA INDÚSTRIA BRASIL - 1988

ESPECIFICAÇÃO	Quantidade (t)
Pneumáticos	238.775
Câmaras de Ar	14.086
Correias	4.472
Material de conserto	19.134
Outros artefatos	4.647

4. Represente as tabelas por meio de gráficos de setores:

ÁREA TERRESTRE BRASIL

REGIÕES	Relativa (%)
Norte	45,25
Nordeste	18,28
Sudeste	10,85
Sul	6,76
Centro-Oeste	18,86

PRODUÇÃO DE LAMINADOS NÃO- PLANOS BRASIL - 1988

UNIDADES DA FEDERAÇÃO	Produção (1.000 t)
Norte	2.773
Nordeste	1.326
Sudeste	1.059
Sul	476
Centro-Oeste	797

5. Represente a tabela por meio de gráfico de colunas justapostas e sobrepostas

ANOS	DEPENDÊNCIAS ADMINISTRATIVAS		
	FEDERAIS	ESTADUAIS	PARTICULARES
1986	452.577	198.156	358.002
1987	510.638	275.797	477.208
1988	451.701	300.947	356.006

Fonte: IBGE

1.2.3 - Agrupamentos de Dados e Distribuição de Freqüências

1.2.3.1 – Sem Intervalos de Classe

Quando se vai fazer um levantamento de uma população, um dos passos é retirar uma amostra desta população e obter dados relativos à variável desejada nesta amostra.

Cabe à Estatística sintetizar estes dados na forma de tabelas e gráficos que contenham, além dos valores das variáveis, o número de elementos correspondentes a cada variável. A este procedimento está associado o conceito de:

- **Dados brutos:** É o conjunto de dados numéricos obtidos e que ainda não foram organizados.
- **Rol:** É o arranjo dos dados brutos em ordem crescente (ou decrescente).
- **Amplitude (H):** É a diferença entre o maior e o menor dos valores observados.

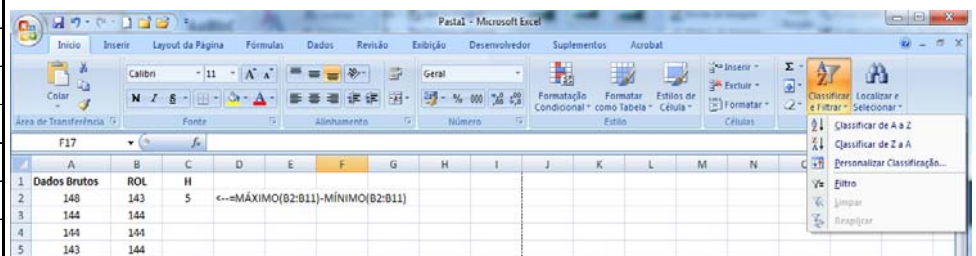
Ex.: População = pressão arterial no pico da sístole em várias medidas num paciente (em mmHg).
 Dados brutos: {148, 144, 144, 143, 145, 148, 145, 144, 148, 148 }
 Rol : {143, 144, 144, 144, 145, 145, 148, 148, 148, 148 }
 $H = 148 - 143 = 5$


Como Fazer isso no Excel?

- Abra uma pasta e na Plan1 insira o intervalo dos dados:

	A
1	Dados Brutos
2	148
3	144
4	144
5	143
6	145
7	148
8	145
9	144
10	148
11	148

Agora copie o intervalo para A2:A11 para B2:B11, re-nomeando o título para ROL.
 A seguir procure classificar os dados em ordem crescente, usando a guia **Início** e o grupo **Edição**. Assim:



Clicando no botão  **Classificar de A a Z** teremos um intervalo de dados ordenados do menor para o maior.

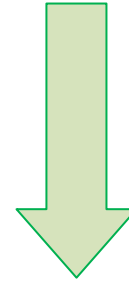
- Como encontrar a Amplitude no Excel?

	A	B
1	Dados Brutos	ROL
2	148	143
3	144	144
4	144	144
5	143	144
6	145	145
7	148	145
8	145	148
9	144	148
10	148	148
11	148	148

Dê o título **H** para a célula C1.

Na célula C2 introduza a função composta
`=MÁXIMO(B2:B11)-MÍNIMO(B2:B11)`

O resultado final de sua planilha será:



	A	B	C	D	E	F	G
1	Dados Brutos	ROL	H				
2	148	143	5	<code><--=MÁXIMO(B2:B11)-MÍNIMO(B2:B11)</code>			
3	144	144					
4	144	144					
5	143	144					
6	145	145					
7	148	145					
8	145	148					
9	144	148					
10	148	148					
11	148	148					

• **Frequência absoluta (f_i):** É o número de vezes que um elemento aparece na amostra. Temos,

$\sum_i^k f_i = n$, onde n é o número total de dados da amostra e k é o número de valores diferentes na amostra.

• **Frequência Relativa (fr_i):** $fr_i = \frac{f_i}{n}$ e $\sum_i^k fr_i = 1$

• **Frequência Absoluta Acumulada (F_i):** É a soma da frequência absoluta do valor da variável i com todas as frequências absolutas anteriores.

• **Frequência Relativa Acumulada (FR_i):** $FR_i = \frac{F_i}{N}$

Ex.: População = pressão arterial no pico da sístole em várias medidas num paciente (em mmHg).

Dados brutos: { 148, 144, 144, 143, 145, 148, 145, 144, 148, 148 }

Rol : { 143, 144, 144, 144, 145, 145, 148, 148, 148, 148 }

$H = 148 - 143 = 5$

Obs.: Esse tipo de tabela é denominado "**Distribuição de Frequências**".

x_i	f_i	fr_i	F_i	FR_i
143	1	0,1	1	0,1
144	3	0,3	4	0,4
145	2	0,2	6	0,6
148	4	0,4	10	1,0
Σ	10	1		

Questão: Como implementar um programa computacional que forneça o rol de 2000 valores de uma determinada variável, de maneira ótima? Assista a aula e verás.....

Exercícios Propostos

1. Embora os gastroenterologistas infantis reconhecessem a utilidade diagnóstica do teor de gordura fecal, até 1984 não existia um padrão de referência desta medida para crianças brasileiras. Para preencher esta lacuna, o *Prof. Francisco Penna*, titular de Pediatria da UFMG, examinou 43 crianças saudáveis que produziram os valores da tabela abaixo, expressos em g/24 horas) (Penna et al. 1987).

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

- Construa um ROL em ordem crescente.
 - Qual é o valor máximo? E o mínimo? E a amplitude?
 - Construa uma tabela de distribuição de frequências.
2. Dados relativos a uma amostra de 36 funcionários de uma população de 2.000 funcionários da Empresa *Aparelhos Ortopédicos S.A*, obtido junto ao Departamento Pessoal da empresa.

Nº	ESTADO CIVIL	GRAU DE INSTRUÇÃO	Nº DE FILHOS	SALARIO (X SAL. MIN)	IDADE ANOS MESES		REGIAO DE PROCEDÊNCIA
1	solteiro	1º grau	—	4	26	03	interior
2	casado	1º grau	1	4,56	32	10	capital
3	casado	1º grau	2	5,25	36	05	capital
4	solteiro	2º grau	—	5,73	20	10	outro
5	solteiro	1º grau	—	6,26	40	07	outro
6	casado	1º grau	0	6,66	28	00	interior
7	solteiro	1º grau	—	6,86	41	00	interior
8	solteiro	1º grau	—	7,39	43	04	capital
9	casado	2º grau	1	7,59	34	10	capital
10	solteiro	2º grau	—	7,44	23	06	outro
11	casado	2º grau	2	8,12	33	06	interior
12	solteiro	1º grau	—	8,46	27	11	capital
13	solteiro	2º grau	—	8,74	37	05	outro
14	casado	1º grau	3	8,95	44	02	outro
15	casado	2º grau	0	9,13	30	05	interior
16	solteiro	2º grau	—	9,35	38	08	outro
17	casado	2º grau	1	9,77	31	07	capital
18	casado	1º grau	2	9,8	39	07	outro
19	solteiro	superior	—	10,53	25	08	interior
20	solteiro	2º grau	—	10,76	37	04	interior
21	casado	2º grau	1	11,06	30	09	outro
22	solteiro	2º grau	—	11,59	34	02	capital
23	solteiro	1º grau	—	12,00	41	00	outro
24	casado	superior	0	12,79	26	01	outro
25	casado	2º grau	2	13,23	32	05	interior
26	casado	2º grau	2	13,6	35	00	outro
27	solteiro	1º grau	—	13,85	46	07	outro
28	casado	2º grau	0	14,69	29	08	interior
29	casado	2º grau	5	14,71	40	06	interior
30	casado	2º grau	2	15,99	35	10	capital
31	solteiro	superior	—	16,22	31	05	outro
32	casado	2º grau	1	16,61	36	04	interior
33	casado	superior	3	17,26	43	07	capital
34	solteiro	superior	—	18,75	33	07	capital
35	casado	2º grau	2	19,40	48	11	capital
36	casado	superior	3	23,30	42	02	interior

- Preencha o quadro abaixo considerando a variável grau de instrução:

Grau de Instrução x_i	Qtde de Funcionários	Frequência Absoluta n_i	Frequência Relativa f_i	Frequência Acumulada Absoluta N_i	Frequência Acumulada Relativa F_i
1º Grau					
2º Grau					
Superior					
Total					

b. Fazer a mesma coisa para a variável *nº de filhos*:

Número de Filhos x_i	Frequência Absoluta n_i	Frequência Relativa f_i	Frequência Acumulada Absoluta N_i	Frequência Acumulada Relativa F_i
0				
1				
2				
3				
5				
Total				

c. Fazer a mesma coisa para a variável *Região de procedência*.

Como calcular essas frequências sem intervalo de classe no Excel?

- Dê os títulos x_i , x_i agrupado e frequência (f_i) para as células D1, E1 e F1, respectivamente.
- Com o cursor na célula D2 insira =B2.
- Pule para a célula D3 e introduza a fórmula: =SE(B3:\$B\$11<>B2;B3;"")
- Arraste pela alça, copiando esta fórmula até a D11 e a sua planilha ficará assim:

	A	B	C	D	E	F
1	Dados Brutos	Rol	H	x_i	x_i agrupado	frequência(f_i)
2	148	143	5	143		
3	144	144		144		
4	144	144				
5	143	144				
6	145	145		145		
7	148	145				
8	145	148		148		
9	144	148				
10	148	148				
11	148	148				

Como juntar os valores em células adjacentes? Ajude o Bertolo a fazer isso e será recompensado!!! Pesquise kra.... Você não quer ser fisioterapeuta top?

Mas enquanto isso. Vamos tocar o “bonde” fazendo a coisa manualmente:

	A	B	C	D	E	F
1	Dados Brutos	Rol	H	X_i	X_i agrupado	freqüência(f_i)
2	148	143	5	143	143	
3	144	144		144	144	
4	144	144			145	
5	143	144			148	
6	145	145		145		
7	148	145				
8	145	148		148		
9	144	148				
10	148	148				
11	148	148				

Calculando as freqüências no Excel:

- Agora selecione o intervalo F2:F6 e introduza a função:
- **=FREQUÊNCIA(B2:B11;E2:E6)** e não dê **ENTER** não!. Calma!!!!
- Agora pressione **F2** e após, ao mesmo tempo, **CTRL +SHIFT + ENTER**.
- O que aconteceu?
- Apareceu **1** em E2, **3** em F3, **2** em F4, **4** em F5 e **0** em F6. Olhando para a barra de fórmulas vemos:
- **{=FREQUÊNCIA(B2:B11;E2:E6)}**, a fórmula ficou entre chaves. O que significa isto?
- O Excel reconheceu a fórmula como uma matriz (*array*).
- Você viu que o intervalo foi de E2 até E6. É isso mesmo. A explicação encontra-se no Apêndice da apostila.
- A coisa agora ficou assim:

The screenshot shows the Excel interface with the formula bar displaying `{=FREQUÊNCIA(B2:B11;E2:E6)}`. The spreadsheet shows the following data:

	A	B	C	D	E	F	G	H
1	Dados Brutos	Rol	H	X_i	X_i agrupado	freqüência(f_i)		
2	148	143	5	143	143	1		
3	144	144		144	144	3		
4	144	144			145	2		
5	143	144			148	4		
6	145	145		145		0		
7	148	145						
8	145	148		148				
9	144	148						
10	148	148						
11	148	148						
12								
13								

Na célula E7 digite: Total

Para encontrar o total introduzimos na célula F7 a função SOMA:

=SOMA(F2:F5)

Resultado Final

	A	B	C	D	E	F	G	H	I
1	Dados Brutos	Rol	H	X_i	X_i agrupado	freqüência(f_i)	fr_i	F_i	Fr_i
2	148	143	5	143	143	1	0,1	1	0,1
3	144	144		144	144	3	0,3	4	0,4
4	144	144			145	2	0,2	6	0,6
5	143	144			148	4	0,4	10	1
6	145	145		145		0			
7	148	145			Total	10	1		
8	145	148		148					
9	144	148							
10	148	148							
11	148	148							

E aí? Quebre a cabeça e preencha (calculando no Excel) as colunas G, H e I.

1.2.3.2 – Histograma no Excel? Para uma distribuição de freqüências SEM intervalo de classe

O histograma é uma forma gráfica de apresentar a distribuição de freqüências de uma variável. O histograma é um gráfico de barras verticais construído com os resultados da tabela de freqüências.


Histograma é o gráfico de barras das freqüências de uma variável

1. CONSTRUÇÃO DE UM HISTOGRAMA DE FREQÜÊNCIAS ABSOLUTAS

Abra uma pasta dê a ela o nome HISTOGRAMA.

Re-nomeie PLAN1 para HISTOGRAMA também.

Aqui vamos mostrar a construção do histograma das freqüências absolutas utilizando os recursos gráficos do Excel 2007.

Entre a seguir: em **A1** com o rótulo **HISTOGRAMA**. Em **B3: Dados**. Selecione as células **E2:G2**, clique no botão  **Mesclar e Centralizar** e insira o rótulo **Freqüências**. Introduza agora os rótulos **Seleção** em **D3**, **Absolutas** em **E3**, **Relativas** em **F3** e **Acumuladas** em **G3**.

Procure colocar bordas externas na cor e sombreamentos nestas células. Escolha

Introduza agora no intervalo **B4:B29**, os valores:

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

No intervalo **D4:D10** introduza os valores:

11	12	13	14	15	16	17
----	----	----	----	----	----	----

No intervalo E4:E10 introduzir a função FREQÜÊNCIA como antes pressionando **Ctrl + Shift + Enter** ao sair para inserir uma matriz. Assim: {=FREQÜÊNCIA(B4:B29;D4:D11)}

Depois na célula **F4** introduza a fórmula =E4/26 e arraste o canto inferior direito até à célula **F10**.

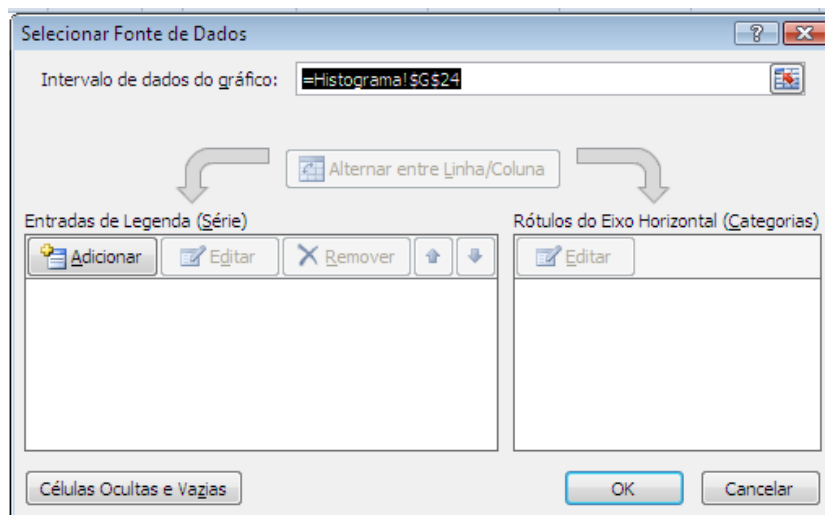
Agora vá à célula **G4** e introduza =F4 e depois na **G5** introduza = G4+F5 e arraste a fórmula até **G10**.


No final a sua planilha ficará como:

	A	B	C	D	E	F	G	H
1	HISTOGRAMA							
2								
3					Frequências			
4		Dados		Seleção	Absolutas	Relativas	Acumuladas	
5		14		11	2	7,7%	7,7%	
6		12		12	5	19,2%	26,9%	
7		13		13	6	23,1%	50,0%	
8		11		14	7	26,9%	76,9%	
9		12		15	3	11,5%	88,5%	
10		13		16	2	7,7%	96,2%	
11		16		17	1	3,8%	100,0%	
12		14			0			
13		14						
14		15						
15		17						
16		14						
17		11						
18		13						
19		14						
20		15						
21		13						
22		14						
23		13						
24		14						
25		13						
26		15						
27		16						
28		12						
29		12						
30								
31								
32								
33								
34								
35								

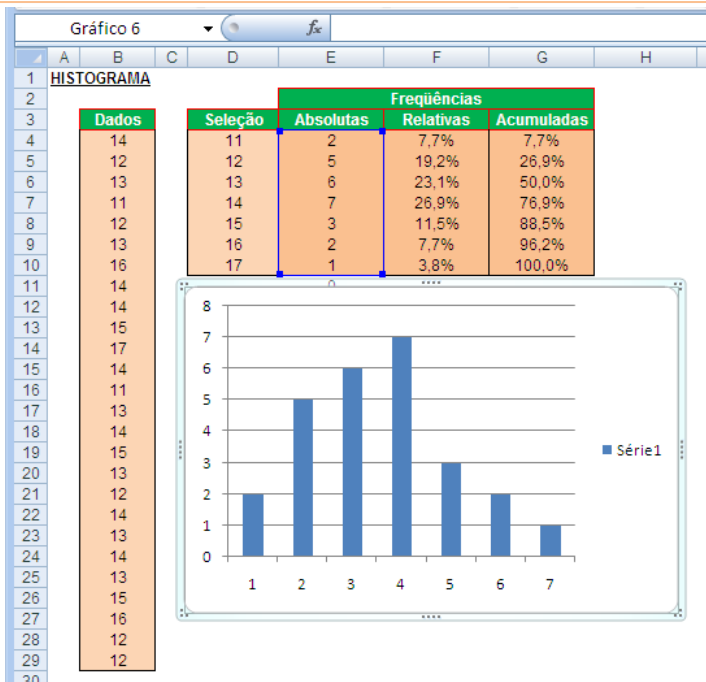
Agora com o cursor do Excel posicionado em qualquer célula da planilha **Histograma** vá até a guia **Inserir**, no grupo **Gráficos**, selecione o tipo de gráfico **Colunas** e o subtipo **colunas agrupadas** da categoria **Coluna 2D**. Este gráfico compara valores entre categorias usando retângulos verticais.

Aparecerá uma janela de *plotagem* e serão acrescentadas três guias de grupos de ferramentas de gráfico: **Design**, **Layout** e **Formatar**. Na guia **Design** e no grupo **Dados** dê um clique no botão **Selecionar Dados** e irá aparecer a janela *Selecionar Fonte de Dados*:

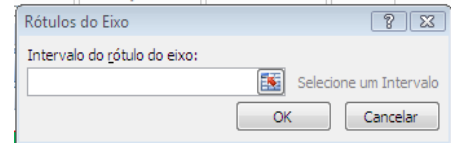


Na caixa Intervalo de dados do gráfico, clique no botão à direita  e a seguir selecione o intervalo **E4:E10**. Aparecerá na caixa **=Histograma!\$E\$4:\$E\$10**. Na categoria *Entradas de Legenda (Série)* aparecerá **Série 1** e na categoria *Rótulos do Eixo Horizontal (Categorias)* os números 1, 2, 3, 4, 5. O gráfico será automaticamente construído.

O gráfico assume o aspecto:



A seguir na Categoria *Rótulos do Eixo Horizontal (Categorias)* clique no botão **Editar** e aparecerá a janela:

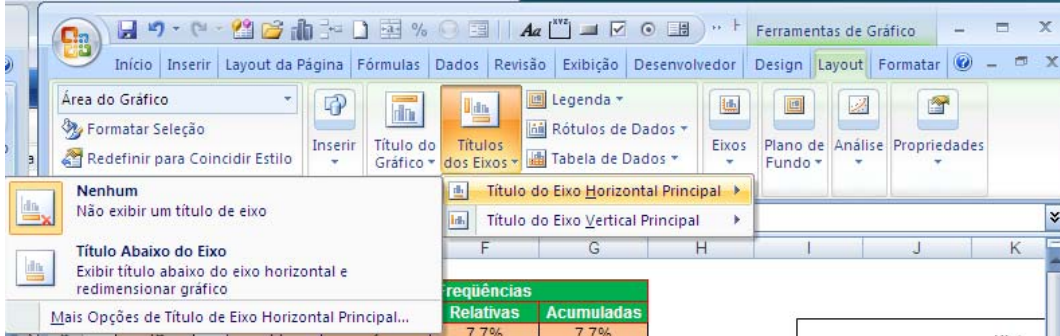


Clicando no botão à direita introduzimos o intervalo D4:D10 clicando e arrastando sobre as células do intervalo na planilha.

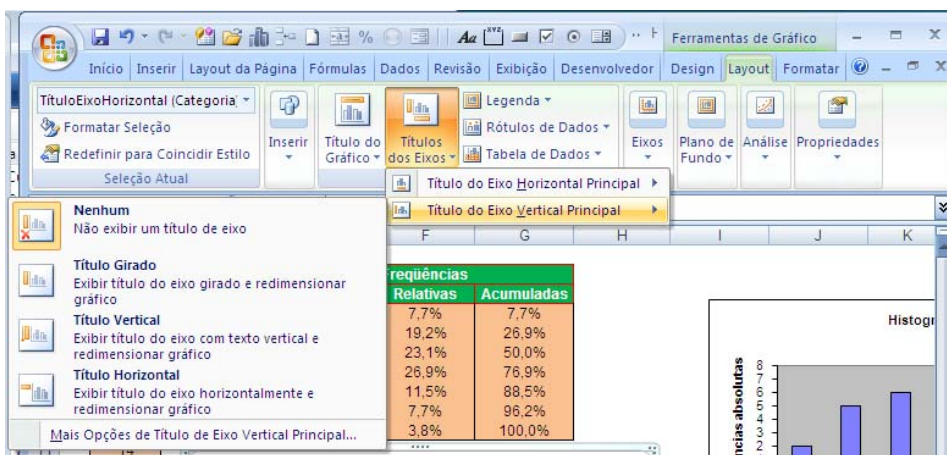
Agora na Categoria *Entradas de Legenda (Série)* clique no botão **Editar** e na caixa de edição **Nome da série** digite **Frequências Absolutas** e a seguir OK.

Agora vamos colocar o título do gráfico e dos eixos indo à guia **Layout**, grupo **Rótulos** e no ícone **Título do Gráfico** e selecionando no menu suspenso **Acima do Gráfico**. O título colocado automaticamente (Frequências Absolutas) ficará marcado. Substitua o título para **Histograma**.

Clicando no ícone **Título dos Eixos** e selecione **Título do Eixo Horizontal Principal** e a seguir **Título Abaixo do Eixo**. Aparecerá uma caixa marcada **Título do Eixo**. Nesta caixa edite o texto substituindo-o para **Seleção**.

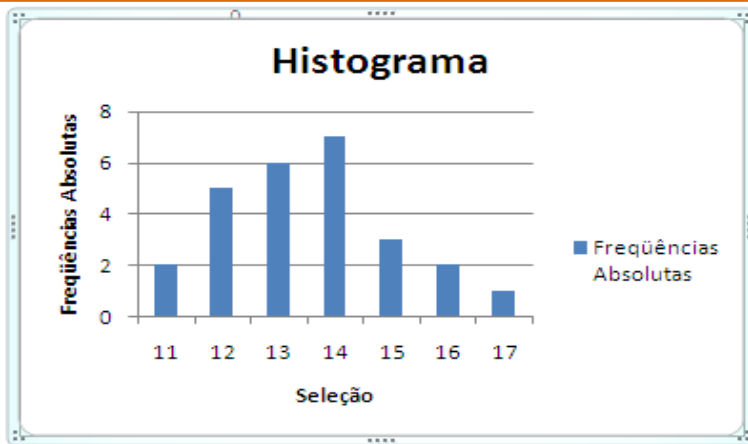


Repetindo para o eixo vertical, selecione na guia **Layout**, no grupo **Rótulos** o ícone **Títulos dos Eixos**, **Título do Eixo Vertical Principal** e **Título Girado**.



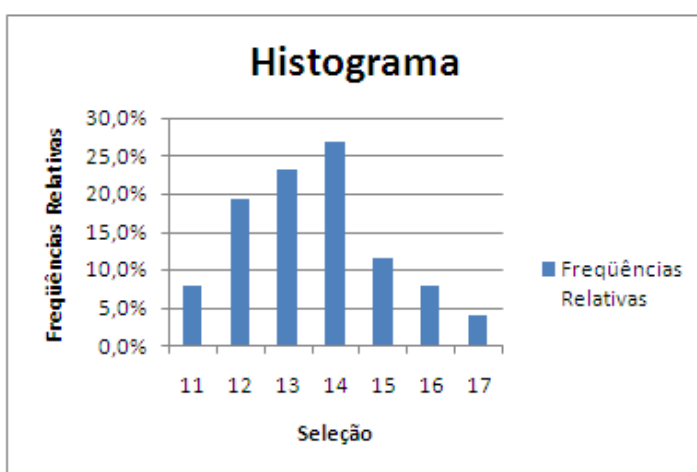
Aparecerá uma caixa marcada com o texto **Título do Eixo**. Mude-o para **Frequências Absolutas**.

Finalmente o gráfico ficará assim:



2. CONSTRUÇÃO DE UM HISTOGRAMA DE FREQUÊNCIAS RELATIVAS

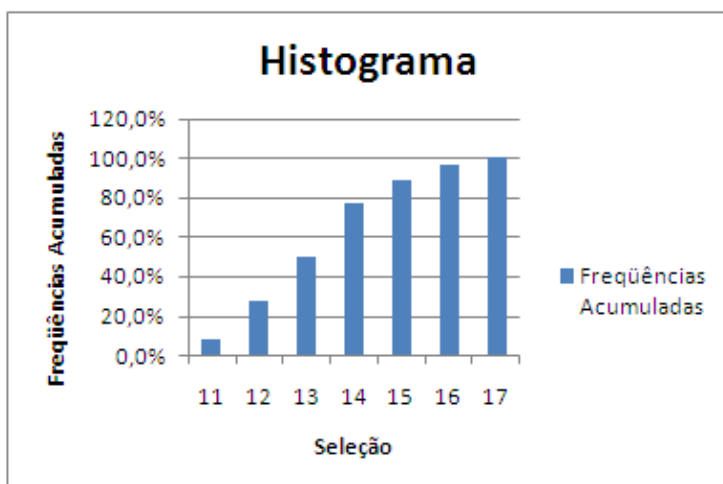
Repetindo os passos anteriores, substituindo o intervalo o intervalo **E4:E10** por **F4:F10** ficaremos com **=Histograma!\$E\$4:\$E\$10** na caixa de edição e modificando o necessário, teremos:



Poderíamos também, através de um **Ctrl + C** e um **Ctrl + V** obter um novo gráfico e nele fazer todas as modificações necessárias para o gráfico ficar como este ao lado.

3. CONSTRUÇÃO DE UM HISTOGRAMA DE FREQUÊNCIAS ACUMULADAS

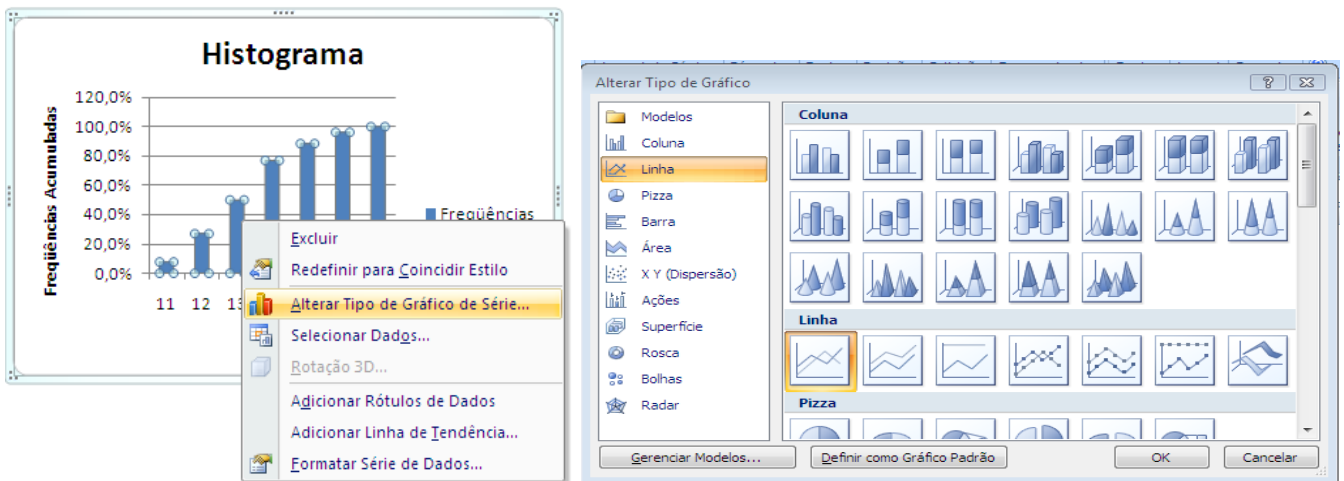
Da mesma forma que acima poderemos obter o gráfico das frequências acumuladas.



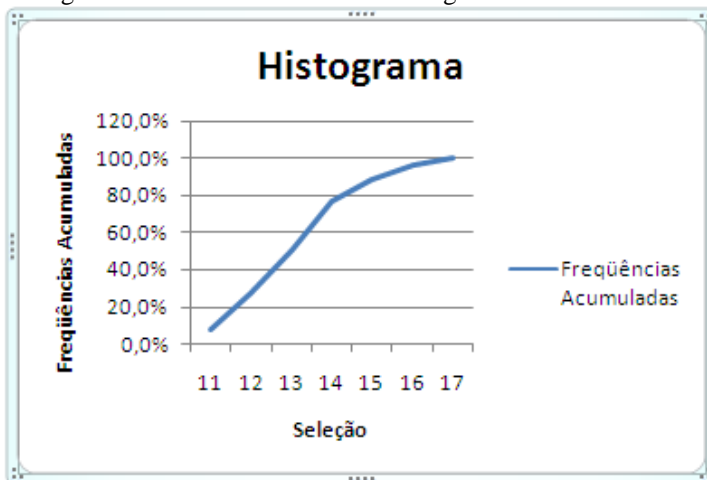
4. REPRESENTAÇÃO GRÁFICA DIFERENTE

As colunas dos exemplos anteriores podem ser substituídas por linhas, as quais damos o nome de *poligonal*.

Para mudarmos o tipo de gráfico das frequências acumuladas para linha, basta clicar com o botão direito do mouse sobre uma coluna e no menu suspenso selecionar **Alterar Tipo de Gráfico de Série...**. Na janela *Alterar Tipo de Gráfico*, selecionar a categoria **Linha** e o tipo **Linhas**



O histograma de barras se transformará no gráfico abaixo:



Este gráfico mostra então a *poligonal* da distribuição de frequências acumuladas que é denominada OGIVA.

5. HISTOGRAMA COMBINADO

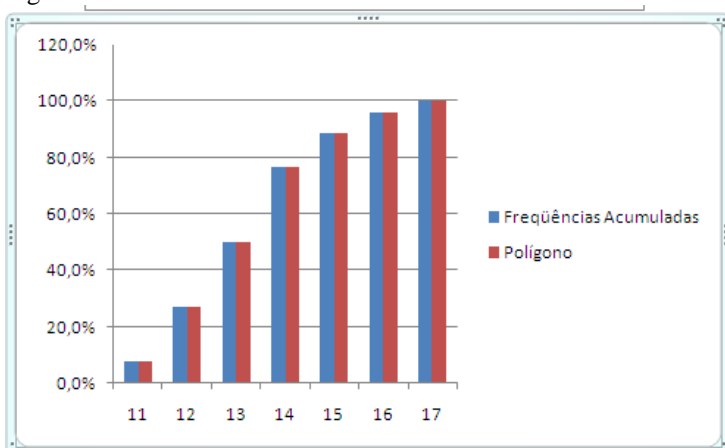
Podemos construir um gráfico combinado (colunas e linhas) da seguinte forma:

a. Marcar o intervalo de células de frequências relativas acumuladas **G4:G10** na planilha histograma e daí ir para a guia **Inserir**, grupo **Gráfico** e escolher o tipo **Colunas Agrupadas**. Automaticamente é gerado o gráfico como antes. A seguir, depois de clicar sobre a área do gráfico vá para a guia **Design**, grupo **Dados** e selecione o botão **Selecionar Dados**

b. Na janela *Selecionar Fonte de Dados*, na categoria *Entradas de Legenda (Série)* clique no botão **Editar** e dê a série o nome de **Frequências Acumuladas**. Agora na categoria *Rótulos do Eixo Horizontal (Categorias)* clique no botão **Editar** e entre com as células da coluna Seleção, assim **=Histograma!\$D\$4:\$D\$10**.

c. Novamente na categoria *Entradas de Legenda* clique no botão **Adicionar**. Dê à série o nome de **Polígono** e entre com os valores da coluna **Acumuladas**.

O gráfico ficará assim:

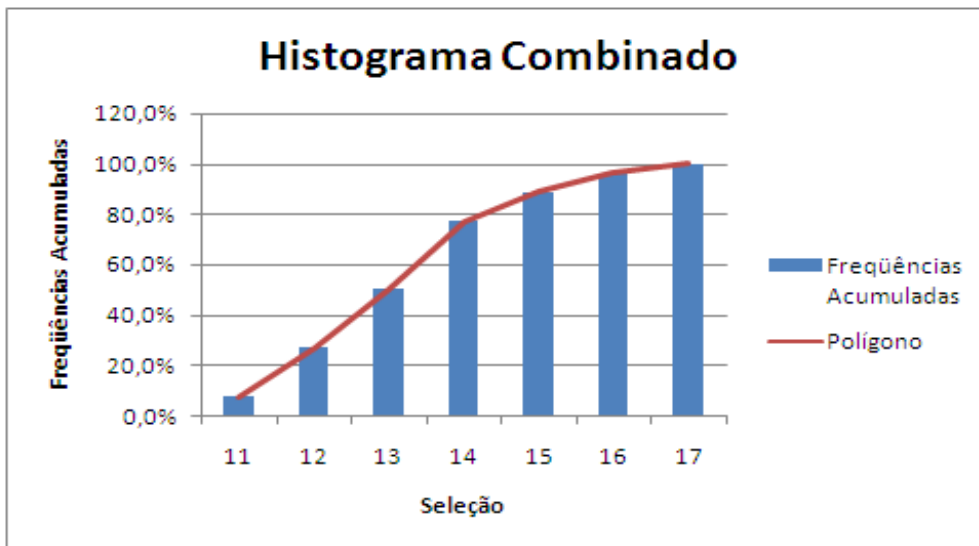


d. Clicando com o botão direito do mouse sobre a série **Polígono**, marcamos a série toda.

5. Clique outra vez com botão direito e aparece o menu suspenso



Selecionando o item **Alterar Tipo de Gráfico de Série...**, escolhendo o tipo **Linhas**, ficamos depois de colocados os títulos nos eixos e no gráfico como antes:



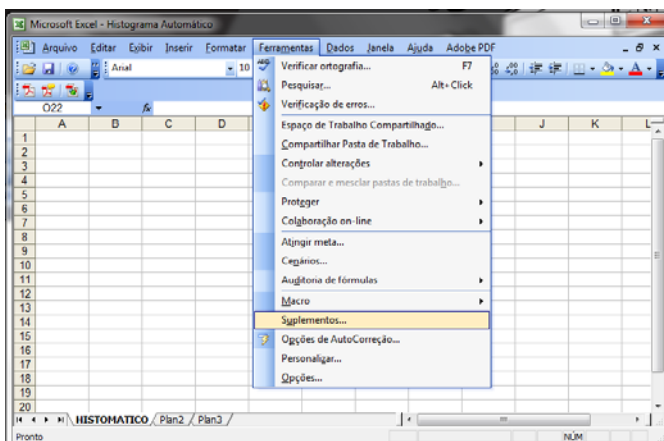
1.2.3.3 – Histograma no Excel. Fazendo uso da Ferramenta Análise de Dados

O Bill Gates sensibilizado com você por achar muito trabalhoso a construção de histograma no Excel, desenvolveu uma ferramenta que constrói o tal histograma automaticamente (obviamente, após alimentada com os dados).

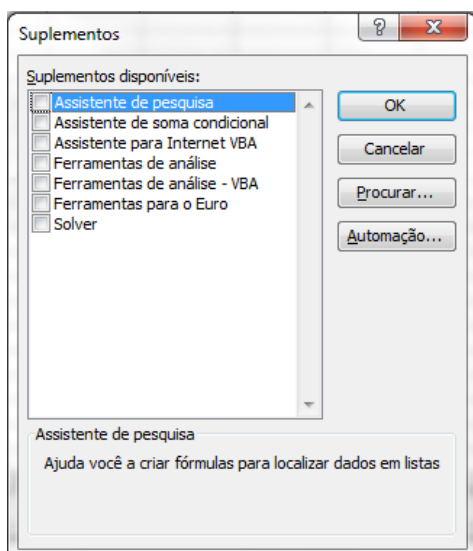
Primeiramente precisamos instalar esta Ferramenta de Análise. Para isso, Abra uma pasta e dê a ela o nome HISTOGRAMA AUTOMÁTICO. A seguir re-nomeie a PLAN1 para HISTOMATICO.

Vamos configurar a nossa ferramenta de Análise de Dados. Isso é feito uma vez só. O Excel fica configurado até uma nova modificação.

Clique no menu **Ferramentas**. No menu suspenso clique sobre **Suplementos...**



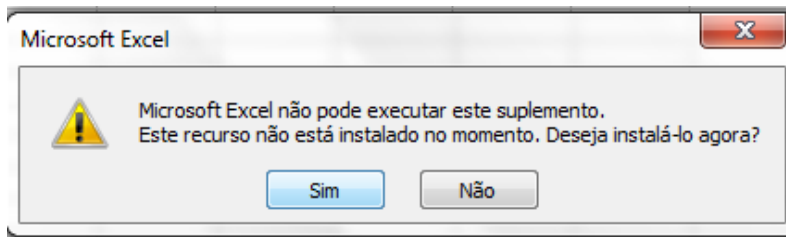
Isto fará aparecer a janela **Suplementos** com os *Suplementos disponíveis*:



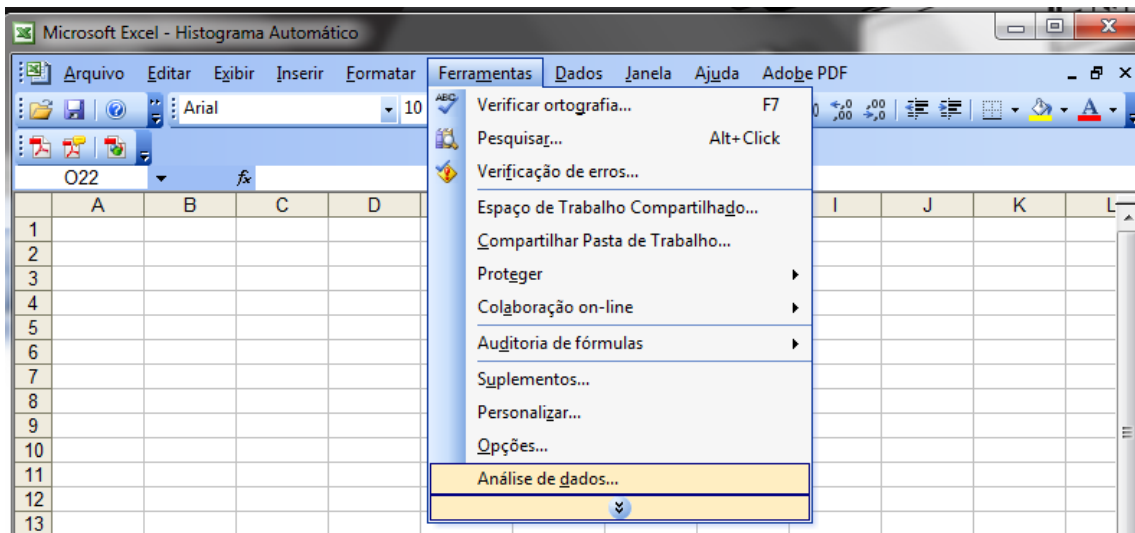
Nesta janela marque (se ainda não estiver marcada) a caixa de verificação **Ferramentas de Análise**¹. A seguir pressione o botão **OK**.

Nesta janela marque (se ainda não estiver marcada) a caixa de verificação **Ferramentas de Análise**⁷. A seguir pressione o botão **OK**.

Vai aparecer agora se a caixa de verificação estava inicialmente desmarcada, a janela:



Isto fará aparecer esta ferramenta **Análise de dados...** no menu **Ferramentas**.



Isto não precisará mais ser feito quando o Excel for aberto novamente. Uma vez basta.

Agora vamos mostrar a construção do histograma das frequências absolutas utilizando a Ferramenta de Análise Histograma do Excel 2007.

Entre a seguir: em **A1** com o rótulo **HISTOGRAMA AUTOMÁTICO**. Em **B3: Dados**. Introduza agora os rótulos **Seleção** em **D3**.

Procure colocar bordas externas e sombreamentos nestas células (a gosto do cliente!).

Introduza agora no intervalo **B4:B29**, os valores:

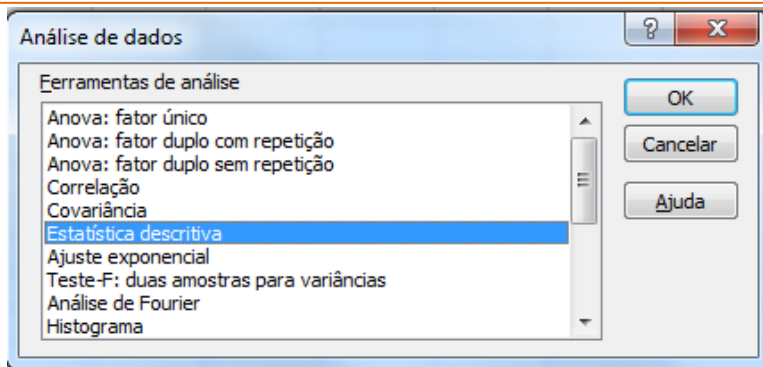
14	12	13	11	12	13	16	14	14	15
17	14	11	13	14	15	13	12	14	13
14	13	15	16	12	12				

No intervalo **D4:D10** introduza os valores:

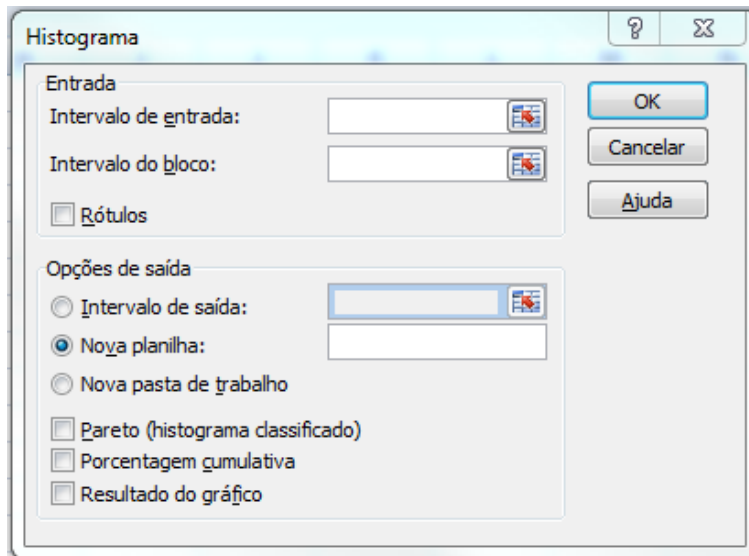
11	12	13	14	15	16	17
----	----	----	----	----	----	----

No menu Ferramentas clique em **Análise de Dados**, aparecerá a janela **Análise de dados**.

⁷ Fornece ferramentas de análise de dados para análises estatísticas e de engenharia.



Selecione Histograma e a seguir pressione o botão OK que fará aparecer a janela Histograma.



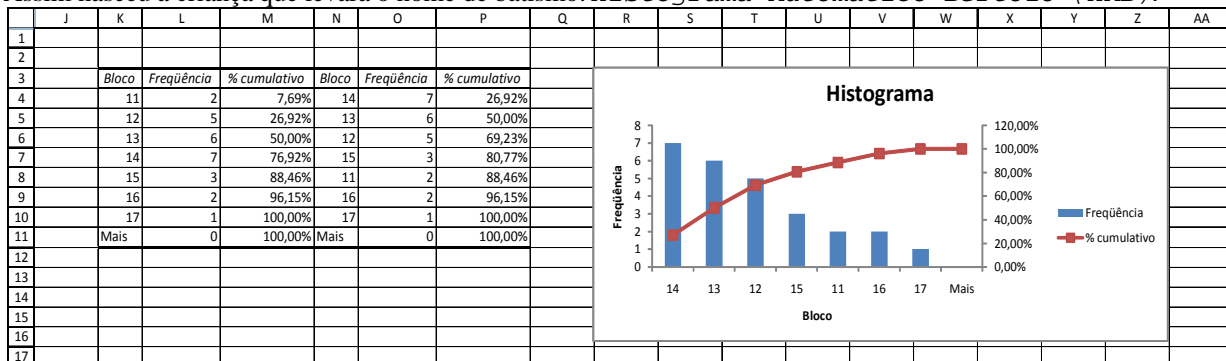
Na caixa Intervalo de entrada: selecione o intervalo B4:B29.

Na caixa Intervalo do bloco: selecione o intervalo D4:D10.

Em Opções de saída clique no botão de rádio Intervalo de saída: e a seguir selecione a célula K3.

Marque as caixas Pareto (histograma classificado), Porcentagem cumulativa, Resultado do gráfico. A seguir pressione o botão OK.

Assim nasceu a criança que levará o nome de batismo: *Histograma Automático Bertolo (HAB)*.



Que facilidade!!! Hein. Thanx Bill Gates.

1.2.4 – Classes e Distribuição de Frequências COM intervalo de classes

As classes são um artifício para condensar o número de elementos diferentes de uma amostra. Imagine construir uma tabela para 200 valores diferentes? São úteis quando a variável for **contínua**.

Os principais pré-requisitos para uma boa definição de classes em um conjunto de dados são:

Pré-requisitos para definição de classes

- As classes devem abranger todas as observações;
- O extremo superior de uma classe é o extremo inferior da classe subsequente (simbologia: $\left[\text{---} \right)$, intervalo fechado à esquerda e aberto à direita);
- Cada valor observado deve enquadrar-se em apenas uma classe;
- $k \leq 25$, de um modo geral, sendo k o número de classes;
- As unidades das classes devem ser as mesmas dos dados.

Cálculo de k (opções não rígidas):

- Fórmula de Sturges:

$$k = 1 + \log_2 N = 1 + 3,32 \log N = 1 + 1,44 \ln N$$

- $k \cong \sqrt{N}$

Obs.: N é o número de elementos diferentes da amostra e em muitas vezes pode ser considerado $N = n$

Geralmente, temos ainda:

- Intervalo da classe (h): $h \cong H / k$
- Ponto médio da classe (x_i): Ponto médio entre o limite inferior e o limite superior de cada classe .

Exercícios de Aplicação

Framingham é uma pequena cidade americana localizada perto de Boston, *Massachussets*. Em 1948 foi selecionada como local adequado para desenvolvimento de um estudo prospectivo cujo objetivo era verificar como os hábitos de vida das pessoas influenciam o risco de desenvolvimento de doenças cardíacas. Muitos resultados, hoje completamente integrados à prática cardiológica, como a necessidade de controle de nível de colesterol, foram primeiramente obtidos neste estudo.

Os valores da tabela seguinte são as medidas em *mg/dl* do colesterol, referentes ao segundo exame realizado em 1952 em 80 pacientes.

Construir a *tabela de distribuição de freqüências*.

278	118	171	179	212	175	242	248	182	219	213	167	233
194	180	184	247	255	233	192	250	221	255	291	227	201
226	277	243	233	170	185	277	121	209	317	150	121	209
242	194	209	200	146	209	184	161	276	196	219	200	217
184	217	196	243	276	228	363	292	199	150	165	229	244
209	209	217	250	167	234	242	192	209	200	255	479	265
179	250											

SOLUÇÃO

Faça isto manualmente assim:

- Encontre o ROL
- Determine o número de classes: $\sqrt{80} \cong 9$.
- Determine os limites inferior e superior de cada classe:

$l_1 = 118$	$L_1 = 158$
$l_2 = 158$	$L_2 = 198$
$l_3 = 198$	$L_3 = 238$
$l_4 = 238$	$L_4 = 278$
$l_5 = 278$	$L_5 = 318$
$l_6 = 318$	$L_6 = 358$
$l_7 = 358$	$L_7 = 398$
$l_8 = 398$	$L_8 = 438$
$l_9 = 438$	$L_9 = 479$

d.

Classe	Frequência Absoluta f_i	Frequência Relativa fr_i	Frequência Acumulada Absoluta F_i	Frequência Acumulada Relativa FR_i
118 --- 158	6	0,08	6	0,08
158 --- 198	21	0,26	27	0,34
198 --- 238	28	0,35	55	0,69
238 --- 278	19	0,24	74	0,93
278 --- 318	4	0,05	78	0,98
318 --- 358	0	0,00	78	0,98
358 --- 398	1	0,01	79	0,99
398 --- 438	0	0,00	79	0,99
400 ---	1	0,01	80	1,00
Total	80	1,00		

Fazendo Distribuição com intervalo de classes no Excel

Abrindo uma pasta para este exercício, escolhemos Plan1 e introduza os dados brutos, classifique em ordem crescente e determine o valor máximo, mínimo, o tamanho da amostra, o número de classes e a amplitude total. No final a sua planilha deverá ficar parecida com esta:

	A	B	C	D	E	F
1	Funções Estatísticas					
2						
3	Dados Brutos	ROL		Funções		
4	278	118	Maior Valor	479	=MÁXIMO(B4:B83)	
5	118	121	Menor Valor	118	=MÍNIMO(B4:B83)	
6	171	121	Tamanho da Amostra	80	=CONT.VALORES(A4:A83)	
7	179	146	Número de classes	9	=ARRED(RAIZ(D6);0)	
8	212	150	Amplitude Total	361	=MÁXIMO(B4:B83)-MÍNIMO(B4:B83)	
9	175	150	Amplitude das classes	40	=INT(D8/D7)	
10	242	161				
11	248	165				
12	182	167				
13	219	167				
14	213	170				
15	167	171				
16	233	175				
17	194	179				
18	180	179				
19	184	180				
20	247	182				
21	255	184				
22	233	184				
23	192	184				
24	250	185				
25	221	192				
26	255	192				
27	291	194				
28	227	194				
29	201	196				
30	226	196				
31	277	199				
32	243	200				
33	233	200				
34	170	200				
35	185	201				
36	277	209				
37	121	209				
38	209	209				
39	317	209				
40	150	209				
41	121	209				
42	209	209				
43	242	212				
44	194	213				
45	209	217				
46	200	217				
47	146	217				
48	209	219				
49	184	219				
50	161	221				
51	276	226				
52	196	227				
53	219	228				
54	200	229				
55	217	233				
56	184	233				
57	217	233				
58	196	234				
59	243	242				
60	276	242				
61	228	242				
62	363	243				
63	292	243				
64	199	244				
65	150	247				
66	165	248				
67	229	250				
68	244	250				
69	209	250				
70	209	255				
71	217	255				
72	250	255				
73	167	265				
74	234	276				
75	242	276				
76	192	277				
77	209	277				
78	200	278				
79	255	291				
80	479	292				
81	265	317				
82	179	363				
83	250	479				

Agora vamos agrupar os dados em classes e calcular as frequências de valores em cada classe. Para isso, coloque os títulos Classe, Limite Inferior e Limite Superior, respectivamente, nas células C11, D11, E11 e F11. A seguir, introduza os rótulos dos intervalos de classe de C12 até C20, como mostrado na planilha abaixo. Use a sua criatividade e espírito de pesquisa e encontre uma maneira para o Excel fazer este preenchimento automaticamente. Depois coloque o cursor na célula D12 e introduza a fórmula =D5. Na célula D13, introduza a seguinte fórmula: =D12+\$D\$9 e arraste a alça até D20. Na célula E12, coloque a fórmula: =D12+\$D\$9-0,01. Por que esse tal 0,01? Para o intervalo ficar aberto à direita. Arraste o resultado até E20. Agora arredonde o resultado 477,99 para 479, para tomar esse valor e os outros superiores a ele.

Agora selecione o intervalo F12 a F21 e introduza a fórmula: =FREQUÊNCIA(B4:B83;E12:E\$20). Não aperte o ENTER. Calma, tá com pressa? Vá pescar hehehe...Pressione a tecla F2 e depois, ao mesmo tempo, as teclas **CTRL + SHIFT + ENTER**, para transformar numa fórmula de matriz. Aparecerá os resultados da planilha abaixo (parte dela, por falta de espaço).

	A	B	C	D	E	F
1	Funções Estatísticas					
2						
3	Dados Brutos	ROL		Funções		
4	278	118	Maior Valor	479	=MÁXIMO(B4:B83)	
5	118	121	Menor Valor	118	=MÍNIMO(B4:B83)	
6	171	121	Tamanho da Amostra	80	=CONT.VALORES(A4:A83)	
7	179	146	Número de classes	9	=ARRED(RAIZ(D6);0)	
8	212	150	Amplitude Total	361	=MÁXIMO(B4:B83)-MÍNIMO(B4:B83)	
9	175	150	Amplitude das classes	40	=INT(D8/D7)	
10	242	161				
11	248	165	Classe	Limite Inferior	Limite Superior	Frequências Absolutas
12	182	167	118 - 158	118	157,99	6
13	219	167	158 - 198	158	197,99	21
14	213	170	198 - 238	198	237,99	28
15	167	171	238 - 278	238	277,99	19
16	233	175	278 - 318	278	317,99	4
17	194	179	318 - 358	318	357,99	0
18	180	179	358 - 398	358	397,99	1
19	184	180	398 - 438	398	437,99	0
20	247	182	438 -	438	479	1
21	255	184				0
22	233	184			Total	80
23	192	184				
24	250	185				
25	221	192				

Rapadura é doce, mas não é mole!!!!

Exercícios Propostos

- Complete a distribuição de frequências abaixo, determinando as frequências simples:

i	X_i	f_i	F_i
1	2	...	2
2	3	...	9
3	4	...	21
4	5	...	29
5	6	...	34
Total		34	

- Conhecidas as idades de 50 pacientes potencialmente suicidas:

84	68	33	52	47	73	68	61	73	77
74	71	81	91	65	55	57	35	85	88
59	80	41	50	53	65	76	85	73	60

67	41	78	56	94	35	45	55	64	74
65	94	66	48	39	69	89	98	42	54

Obtenha a distribuição de frequência, tendo 30 para limite inferior da primeira classe e 10 para o intervalo de classe.

3. Melanoma é uma neoplasia de grande potencial maligno responsável por grande parte das mortes por câncer de pele. O reconhecimento clínico do tumor em estágio inicial é difícil e de fundamental importância, já que o único procedimento terapêutico realmente eficaz é a excisão cirúrgica precoce. Na análise de 35.695 biópsias realizadas entre 2001 e 2011 no Hospital do Câncer de Catanduva, a tabela mostra a distribuição dos casos da localização anatômica dos melanomas primitivos de pele, sendo a cabeça/pescoço mais comuns:

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

Forme uma distribuição de frequências sem intervalo de classe.

4. Considerando as notas de um teste de inteligência aplicado a 100 alunos:

64	78	66	82	74	103	78	86	103	87
73	95	82	89	73	92	85	80	81	90
78	86	78	101	85	98	75	73	90	86
86	84	86	76	76	83	103	86	84	85
76	80	92	102	73	87	70	85	79	93
82	90	83	81	85	72	81	96	81	85
68	96	86	70	72	74	84	99	81	89
71	73	63	105	74	98	78	78	83	96
95	94	88	62	91	83	98	93	83	76
94	75	67	95	108	98	71	92	72	73

Forme uma tabela de distribuição de frequências.

5. Complete a tabela abaixo:

i	classes	f_i	fr_i	F_i	FR_i
1	0 ---- 8	4
2	8 ---- 16	10
3	16 ---- 24	14
4	24 ---- 32	9
5	32 ---- 40	3
Total		40	1,00		

6. Dada a distribuição de frequências:

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

Determine:

- $\sum f_i$
- As frequências relativas;
- As frequências acumuladas;
- As frequências relativas acumuladas.

7. A tabela abaixo apresenta uma distribuição de freqüências das áreas de 400 lotes:

ÁREAS (m ²)	300 --- 400	400 --- 500	500 --- 600	600 --- 700	700 --- 800	800 --- 900	900 --- 10.000	10.000 --- 11.000	11.000 --- 12.000
Nº de lotes	14	46	58	76	68	62	48	22	6

Com referência a essa tabela, determine:

- A amplitude total;
 - O limite superior da quinta classe;
 - O limite inferior da oitava classe;
 - O ponto médio da sétima classe;
 - A amplitude do intervalo da segunda classe;
 - A freqüência da quarta classe;
 - A freqüência relativa da sexta classe;
 - A freqüência acumulada da quinta classe;
 - O número de lotes cuja área não atinge 700 m²;
 - O número de lotes cuja área atinge e ultrapassa 800 m²;
 - A porcentagem dos lotes cuja área não atinge 600 m²;
 - A porcentagem dos lotes cuja área seja maior ou igual a 900 m²;
 - A porcentagem dos lotes cuja área é de 500 m², no mínimo, mas inferior a 1.000 m²;
 - A classe do 72º lote;
 - Até que classe, estão incluídos 60% dos lotes.
8. A distribuição abaixo indica o número de acidentes ocorridos com 70 motoristas de uma empresa de ônibus:

Nº de Acidentes	0	1	2	3	4	5	6	7
Nº de Motoristas	20	10	16	9	6	5	3	1

Determine:

- O número de motoristas que não sofreram nenhum acidente;
 - O número de motoristas que sofreram pelo menos 4 acidentes;
 - O número de motoristas que sofreram menos de 3 acidentes;
 - O número de motoristas que sofreram no mínimo 3 e no máximo 5 acidentes;
 - A porcentagem dos motoristas que sofreram no máximo 2 acidentes.
9. Complete os dados que faltam na distribuição de freqüência:

i	x _i	f _i	fr _i	F _i
1	0	1	0,05	...
2	1	...	0,15	4
3	2	4
4	3	...	0,25	13
5	4	3	0,15	...
6	5	2	...	18
7	6	19
8	7
Total		20	1,00	

e

i	classes	x _i	f _i	fr _i	F _i
1	0 ---- 2	1	4	0,04	...
2	2 ---- 4	...	8
3	4 ---- 6	5	...	0,18	30
4	...	7	27	0,27	...
5	8 ---- 10	...	15	...	72
6	10 ---- 12	83
7	...	13	10	0,10	93
8	14 ---- 16	0,07	...
Total					

1.2.5 – Representação Gráfica de uma Distribuição de Frequência com intervalos de classe

A distribuição de frequência com intervalo de classe pode ser representada graficamente:

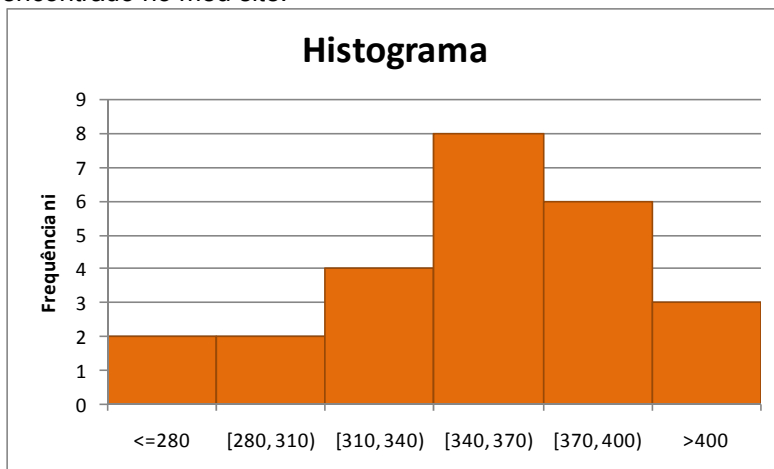
- pelo **histograma**,
- pelo **polígono de frequência** (que é um gráfico em linha, sendo as frequências marcadas sobre perpendiculares ao eixo horizontal, levantadas pelos pontos médios dos intervalos de classe),
- pelo **polígono de frequência acumulada** (marcando-se as frequências acumuladas sobre perpendiculares ao eixo horizontal, levantadas nos pontos correspondentes aos limites superiores dos intervalos de classe) e
- pela **ogiva de Galton**.

1.2.5.1 – Pelo Histograma

Seja a tabela de Distribuição de Frequências com intervalo de classe de uma amostra aleatória de tamanho, vinte e cinco, das Vendas diárias em milhares de uma empresa, analisada anteriormente num exemplo:

i	Classe	Ponto médio da classe - x_i	Frequência Absoluta n_i	Frequência Acumulada N_i
1	280 --- 310	295	3	3
2	310 --- 340	325	4	7
3	340 --- 370	355	6	13
4	370 --- 400	385	7	20
5	400 ---	415	5	25
	Total		25	

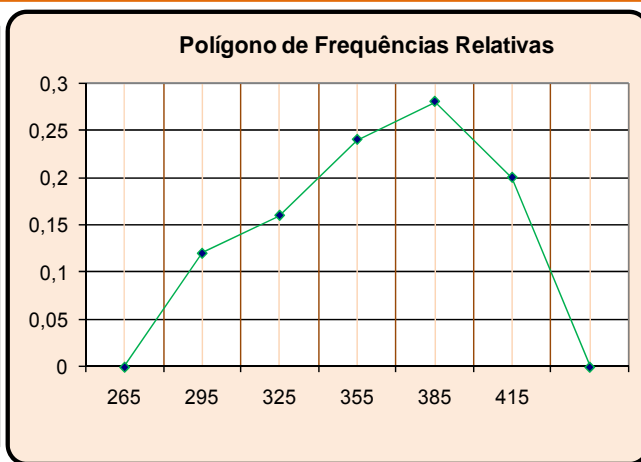
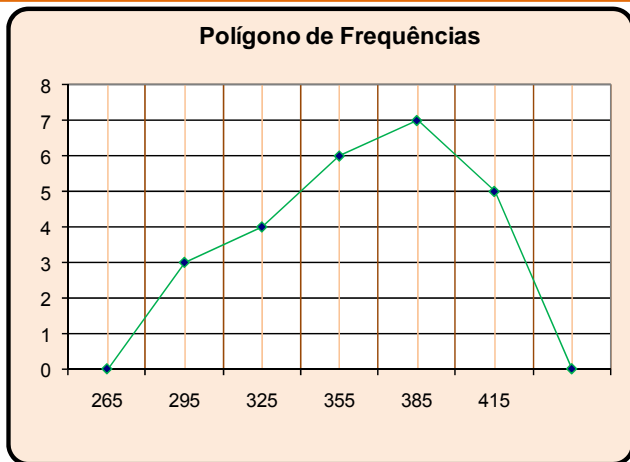
A maneira de construir um histograma com intervalo de classe feito, de uma maneira profissional⁸, no Excel pode ser encontrado no meu site.



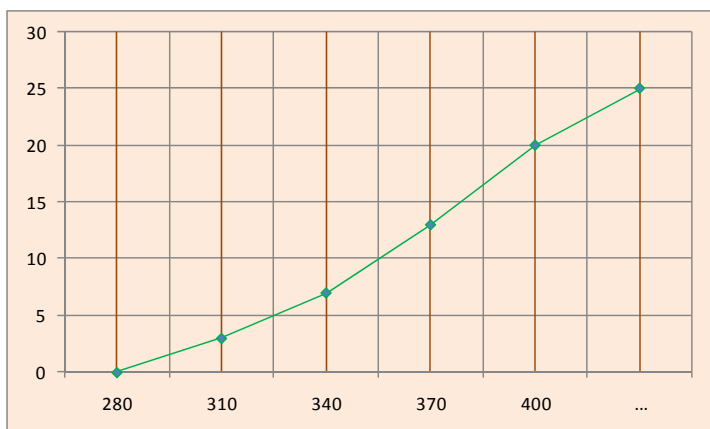
1.2.5.2 – Pelo Polígono de Frequências

Consideremos a mesma Tabela de Distribuição de Frequências dada acima e marquemos os pontos médios dos intervalos de classe no eixo horizontal.

⁸ Se você curte o Excel procure estudar esta planilha que usa recursos avançados (para os “craques”).

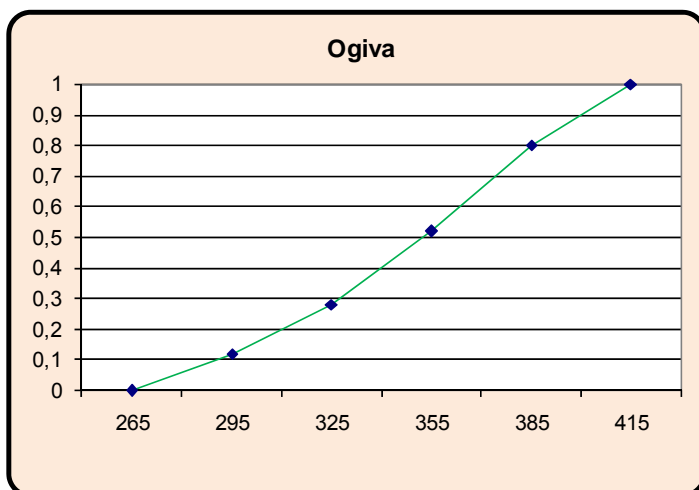


1.2.5.3 – Pelo Polígono de Frequências Acumuladas



1.2.5.4 – Pela ogiva de Galton

É constituído pelo polígono das frequências RELATIVAS acumuladas



Exercícios de Revisão

01 Numa pequena empresa com 20 funcionários a distribuição dos salários, em reais, é a seguinte:

800,00, 800,00, 12.000,00, 12.000,00, 800,00, 600,00, 1.300,00, 600,00, 800,00, 1.000,00, 500,00, 600,00, 700,00, 800,00, 700,00, 12.000,00, 500,00, 600,00, 800,00 e 500,00.

Coloque os valores em ordem crescente e, em seguida, monte uma tabela de distribuição de frequências.

Resposta

500,00; 500,00; 500,00; 600,00; 600,00; 600,00; 600,00; 600,00; 700,00; 700,00; 800,00; 800,00; 800,00; 800,00; 800,00; 800,00; 1.000,00; 1.300,00; 12.000,00; 12.000,00; 12.000,00.

Salários dos funcionários de uma empresa X

Salário dos funcionários	Frequência
500,00	3
600,00	4
700,00	2
800,00	6
1.000,00	1
1.300,00	1
12.000,00	3

Empresa X

02 De acordo com a tabela construída no exercício anterior, encontrar a frequência percentual de cada salário e repetir a sua tabela acrescentando mais uma coluna, de valores percentuais.

Salário dos funcionários da empresa X

Salário dos funcionários	Frequência	Percentual
500,00	3	15%
600,00	4	20%
700,00	2	10%
800,00	6	30%
1.000,00	1	5%
1.300,00	1	5%
12.000,00	3	15%

Empresa X

03 Em certa eleição do grêmio de um colégio, havia 3 equipes de candidatos A, B e C. Apesar dos conflitos – que a própria estatística aponta com os votos nulos e brancos – vamos completar a tabela e mostrar a equipe vencedora?

Equipes	Porcentagem do total de votos	Número de votos
A	26%	182
B	24%	168
C	22%	154
Nulos ou brancos	28%	196

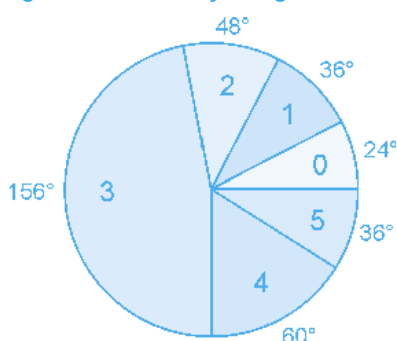
Como o número de votos brancos e nulos corresponde a um número maior que os votos das equipes, isso representa um conflito. No entanto, a equipe vencedora foi a A, com 182 votos.

04 Foi aplicada, em duas séries, A e B, de um colégio X, uma prova com 5 questões. A representação dos resultados da série A foi feita por meio de uma tabela, e os resultados da série B foram apresentados em um gráfico.

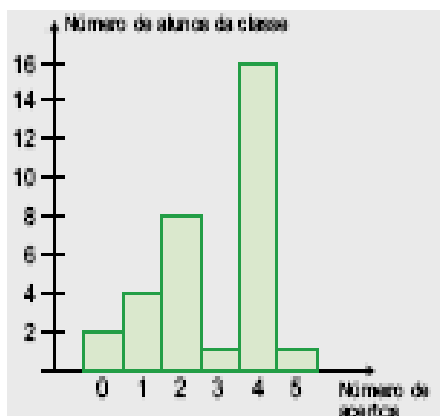
Série A – Acertos das questões da prova

Complete a tabela acima e represente, por meio de um gráfico de setores, os resultados percentuais.

Agora, vamos traçar o gráfico de setores.



Série B



Passes os dados para uma tabela como a utilizada para a série A.

Série B – Acertos das questões da prova

05 Determine as medidas aproximadas dos ângulos a serem utilizados num gráfico de setores, correspondentes às porcentagens:

- a) 2,5%
100% ----- 360°
2,5% ----- x → x = 9°
- b) 80%
100% ----- 360°
80% ----- x → x = 288°
- c) 52%
100% ----- 360°

Número de Acertos	Frequência	Porcentagem
0	2	6,25%
1	4	12,5%
2	8	25%
3	1	3,125%
4	16	50%
5	1	3,125%
Total	32	100%

$$52\% \text{ ----- } x \rightarrow x = 187,2^\circ$$

d) 37%

$$100\% \text{ ----- } 360^\circ$$

$$37\% \text{ ----- } x \rightarrow x = 133,2^\circ$$

06 Em uma prova de Português de 40 questões objetivas, um aluno errou 12 questões. Qual será seu percentual de acerto?

Total de questões = 40

Questões erradas = 12 questões

Questões certas = 28 questões

$$p = \frac{28}{40} 100 \Rightarrow p = 70\%$$

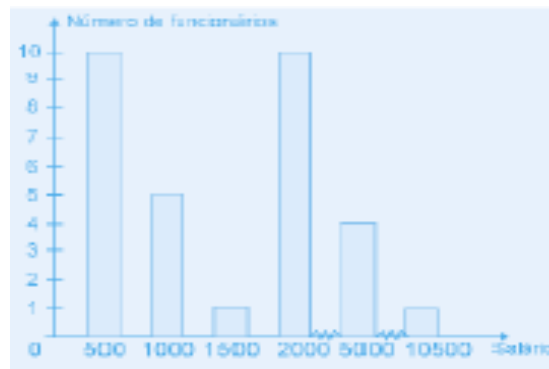
Seu percentual de acerto é 70%.

07 É dada a tabela abaixo, que corresponde à distribuição de salários de uma empresa A. Complete a tabela com valores percentuais e, em seguida, faça um gráfico de barras **número de funcionários x salários**.

Distribuição de salários

Salário (em R\$)	Número de funcionários	Percentual
500,00	10	32,3
1.000,00	5	16,1
1.500,00	1	3,2
2.000,00	10	32,3
5.000,00	4	12,9
10.500,00	1	3,2
Total	31	= 100%

Empresa A



1.3 – Mais sobre Gráficos

Cartograma - É a representação de uma carta geográfica. Este tipo de gráfico é empregado quando o objetivo é o de figurar os dados estatísticos diretamente relacionados com as áreas geográficas ou políticas.

Dados absolutos (população) – usa-se pontos proporcionais aos dados.

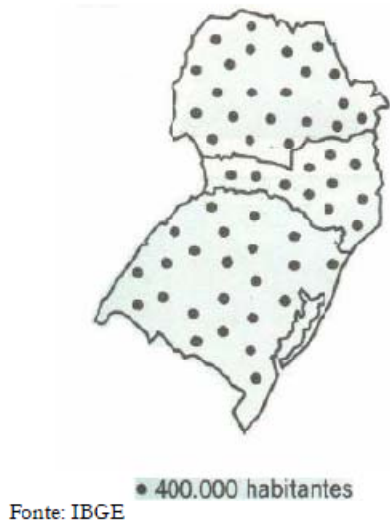
Dados relativos (densidade) – usa-se hachaduras.

População da Região Sul do Brasil - 1990

Estado	População (hab.)	Área (m 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

População da Região Sul do Brasil – 1990



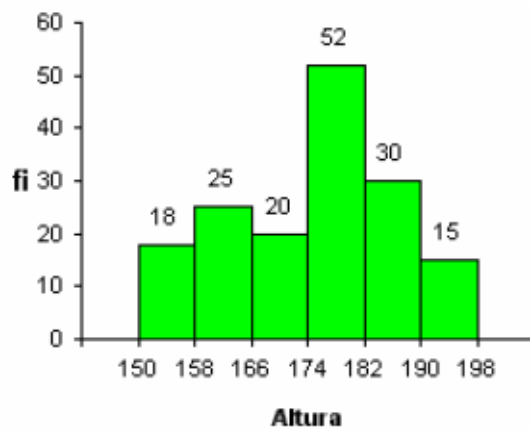
Densidade populacional da Região Sul do Brasil – 1990



3.7 Gráficos utilizados para a análise de uma distribuição de freqüência

3.7.1 Histograma

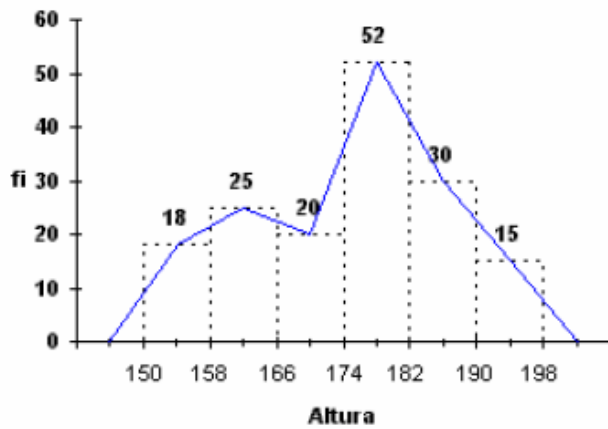
Alturas em centímetros de 160 alunos do IMES – Catanduva - 2010



Fonte: Departamento de Estatística (1990)

Polígono de Freqüências

Alturas em centímetros de 160 alunos do IMES – Catanduva - 2010



Fonte: Departamento de Estatística (1990)

Ogivas

Alturas em centímetros de 160 alunos do IMES – Catanduva - 2010

Ogiva Crescente

Ogiva Decrescente

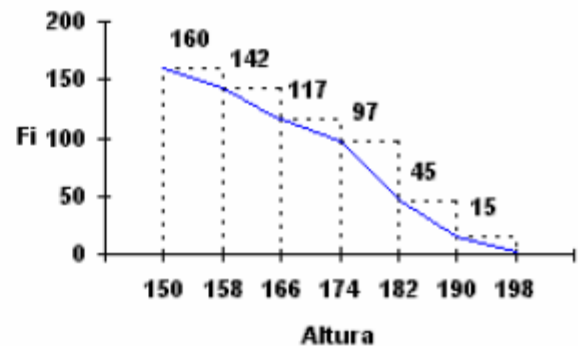
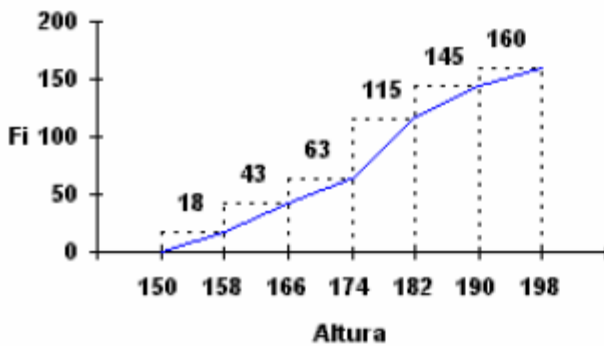
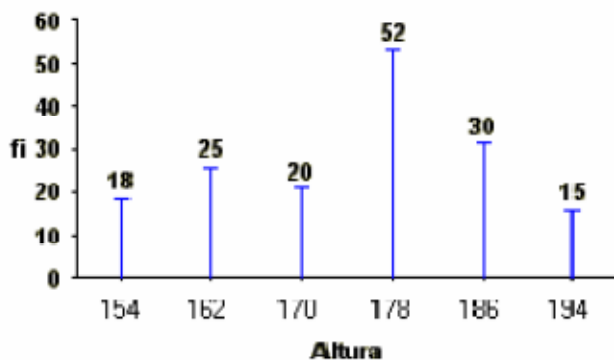


Gráfico em segmentos de reta vertical

É utilizado para representar uma distribuição de frequência pontual, onde os segmentos de reta são proporcionais às respectivas frequências absolutas.

Alturas em centímetros de 160 alunos do IMES – Catanduva - 2010



Como se interpreta um histograma?

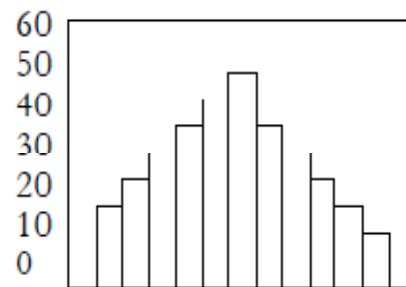
A representação gráfica da distribuição da variável, por **histogramas**. Este gráfico é utilizado para variáveis contínuas.

Características:

- Cada barra representa a frequência do intervalo respectivo;
- Os intervalos devem ter a mesma amplitude;
- As barras devem estar todas juntas.

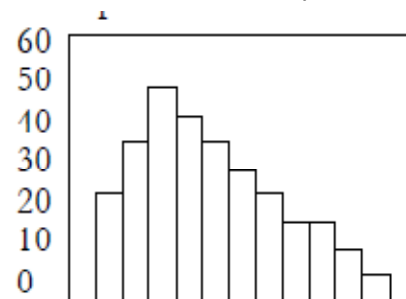
1. Histograma

A simples observação da forma do histograma permite algumas conclusões. Veja a figura 1 ao lado. A medida dos dados está no centro do desenho. As freqüências mais altas também estão no centro da figura. Nos processos industriais, esta é a forma desejável.



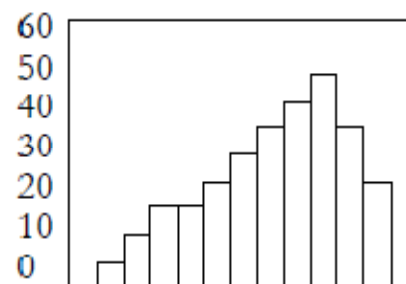
A figura 2 apresenta um histograma com assimetria positiva. A média dos dados está localizada à esquerda do centro da figura e a cauda à direita é alongada. Esta ocorre quando o limite inferior é controlado ou quando não podem ocorrer valores abaixo de determinado limite.

2. Histograma com assimetria positiva



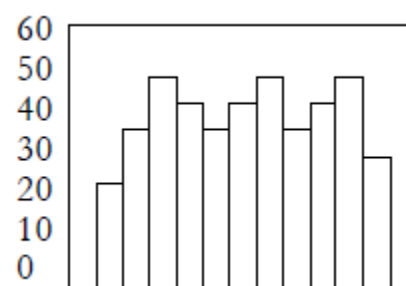
A figura 3 apresenta um histograma com assimetria negativa. A média dos dados está localizada à direita do centro da figura e a cauda à esquerda é alongada. Esta forma ocorre quando o limite superior é controlado ou quando não podem ocorrer valores acima de certo limite

3. Histograma com assimetria negativa



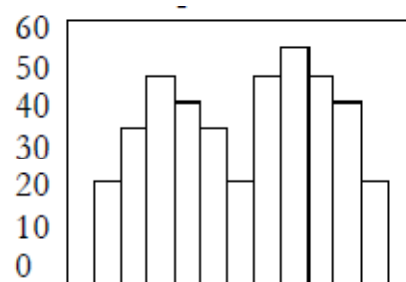
4. Histograma em plateau

A figura 4 mostra um histograma em *plateau*, Isto é, com exceção das primeiras e das últimas classes, todas as outras têm freqüências quase iguais. Essa forma ocorre quando se misturam várias distribuições com diferentes médias.



5. Histograma com dois picos

A figura 5 mostra um histograma com *dois picos*, ou duas modas. As freqüências são baixas no centro da figura, mas existem dois picos fora do centro. Esta forma ocorre quando duas distribuições com médias bem diferentes se misturam. Podem estar misturados, por exemplo, os produtos de dois turnos de trabalho.



1.4 – APÊNDICES

Curva de frequência – curva polida

Como, em geral, os dados coletados pertencem a uma amostra extraída de uma população, pode-se imaginar as amostras tornando-se cada vez mais amplas e a amplitude das classes ficando cada vez menor, o que nos permite concluir que o contorno do polígono de frequências tende a se transformar numa curva (curva de frequência), mostrando, de modo mais evidente, a verdadeira natureza da distribuição da população.

Pode-se dizer, então, que, enquanto que o polígono de frequência nos dá a imagem real do fenômeno estudado, a curva de frequência nos dá a imagem tendenciosa.

Assim, após o traçado de um polígono de frequência, é desejável, muitas vezes, que se faça um polimento, de modo a mostrar o que seria tal polígono com um número maior de dados.

Esse procedimento é claro, não nos dará certeza absoluta que a curva polida seja tal qual a curva resultante de um grande número de dados. Porém, pode-se afirmar que ela assemelha-se mais a curva de frequência que o polígono de frequência obtido de uma amostra limitada.

O polimento, geometricamente, corresponde à eliminação dos vértices da linha poligonal. Consegue-se isso com a seguinte fórmula:

$$fc_i = \frac{f_{\text{ant.}} + 2f_i + f_{\text{post.}}}{4}$$

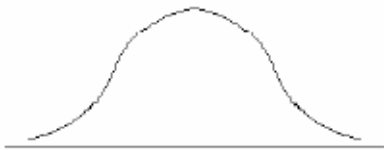
Curvas em forma de sino

As curvas em forma de sino caracterizam-se pelo fato de apresentarem um valor máximo na região central.

Distinguem-se as curvas em forma de sino em: **simétrica** e **assimétrica**

a) Curva simétrica

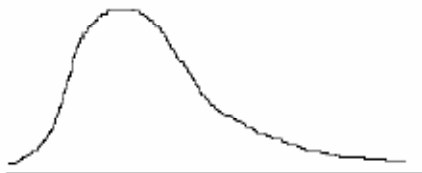
Esta curva caracteriza-se por apresentar o valor máximo no ponto central e os pontos equidistantes desse ponto terem a mesma frequência.



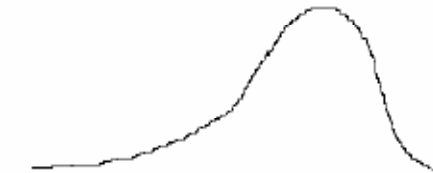
b) Curvas assimétricas

Na prática, não se encontram distribuições perfeitamente simétricas. As distribuições obtidas de medidas reais são mais ou menos assimétricas, em relação à frequência máxima. Assim, as curvas correspondentes a tais distribuições apresentam a **cauda** de um lado da ordenada máxima mais longa do que o outro. Se a cauda mais longa fica a direita é chamada assimétrica positiva ou enviesada à direita, se a cauda se alonga a esquerda, a curva é chamada assimétrica negativa ou enviesada à esquerda.

Assimétrica Positiva



Assimétrica Negativa

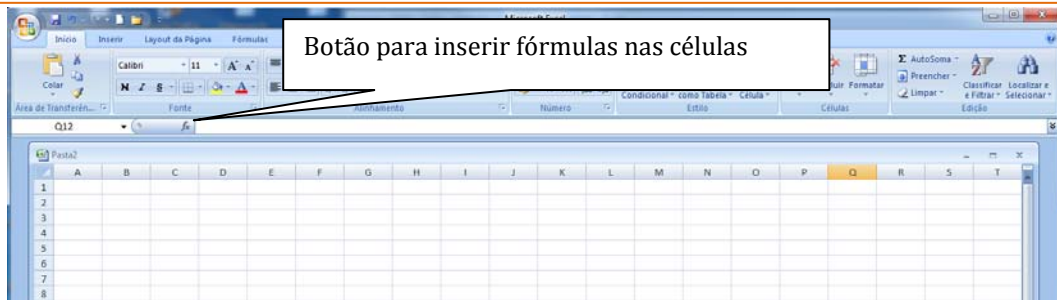


1.4.1 – Geração de Números Aleatórios no Excel

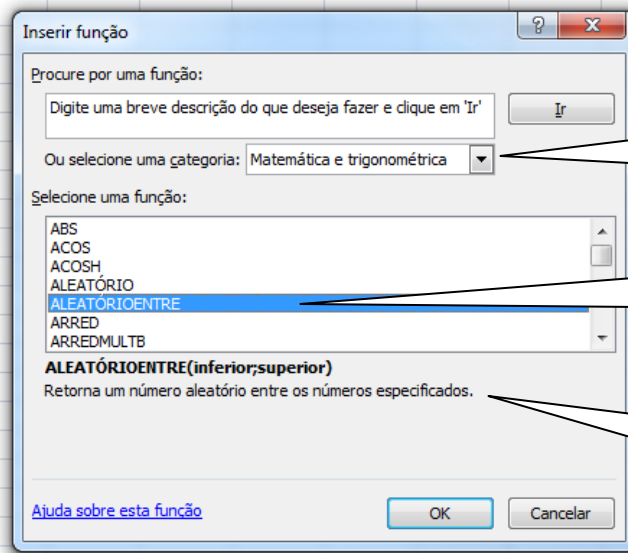
A planilha Excel possui uma função embutida chamada ALEATÓRIOENTRE que gera números aleatórios num intervalo estabelecido previamente. Esta função se encontra na categoria Matemática e trigonométrica.

Para gerar uma série de números aleatórios entre 1 e 90 usados nos nossos exemplos de amostragens, fazemos o seguinte:

- Abrimos o aplicativo Excel com um nova pasta.
- Posicionamos o cursor na célula A1 e depois clicamos no botão para inserir fórmulas que se encontra na barra de fórmulas:



- Uma vez clicado o botão aparecerá a seguinte caixa de diálogo:

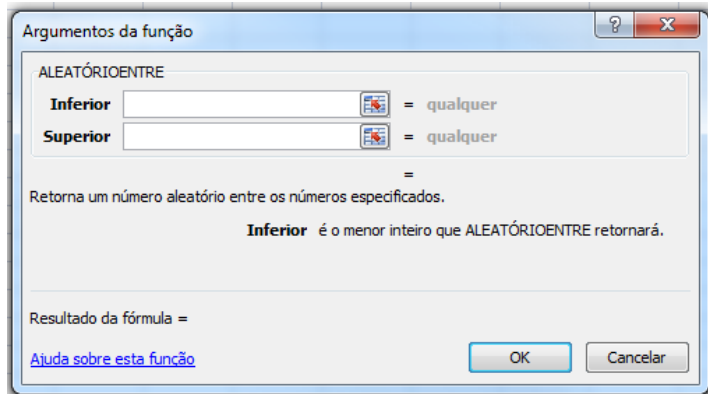


Nesta caixa de combinação escolha a categoria Matemática e trigonométrica

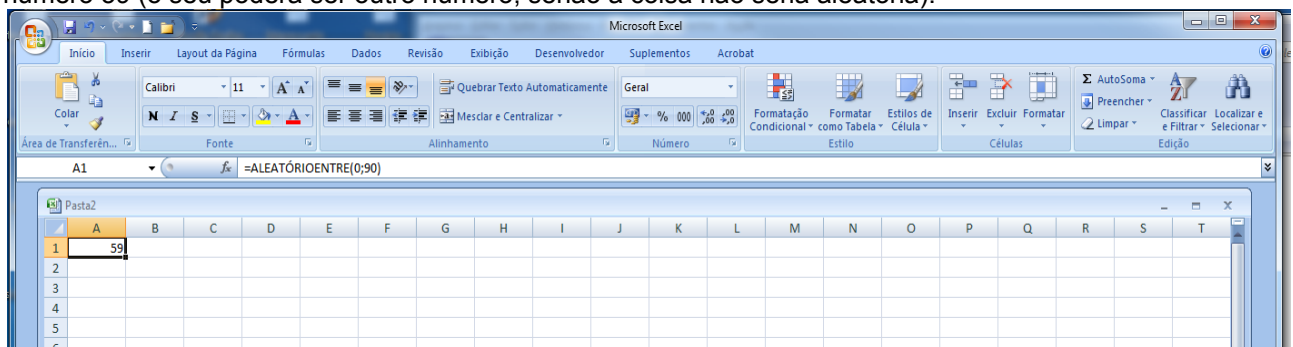
Nesta caixa de listagem selecione a função ALEATÓRIOENTRE.

Vejam que aparece a sintaxe e a ação da função selecionada. E também uma AJUDA no canto inferior esquerdo.

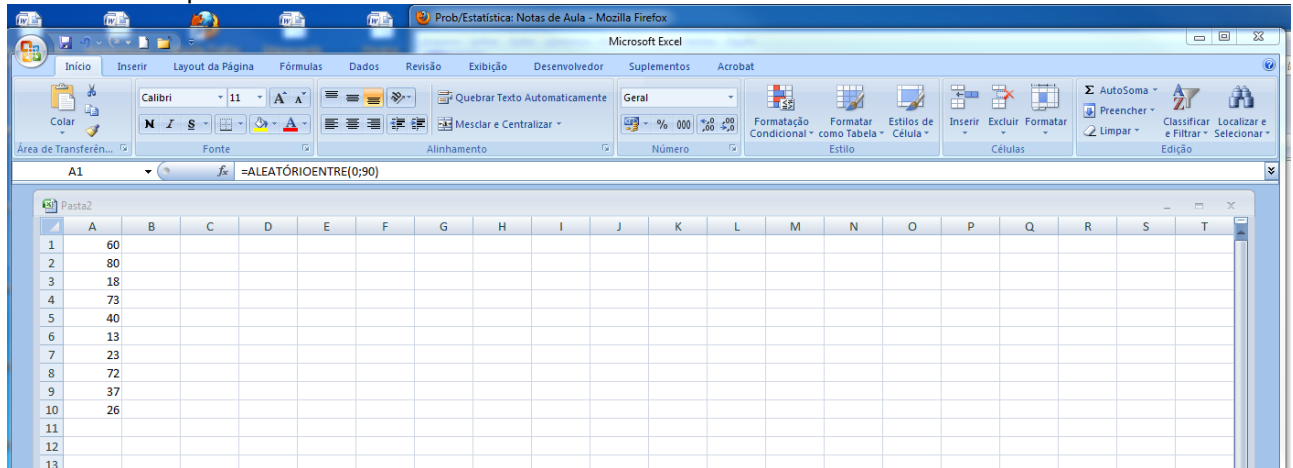
- Clicando em OK aparecerá uma nova caixa de diálogo para inserção dos argumentos da função (a isso dá-se o nome de assistente ou wizard do Office):



- Preencha as caixa Inferior (com o número 0) e a caixa Superior (com o número 90). Observe que o resultado da fórmula é Volátil. Isto porque a cada vez que você der um ENTER será gerado um n° diferente. Clicando em OK temos a inserção de um n° aleatório entre zero e 90 na célula A1. No nosso exemplo foi o número 59 (o seu poderá ser outro número, senão a coisa não seria aleatória).



Clicando com o mouse no canto inferior da célula A1 e arrastando até A10 encontramos 9 números gerados aleatoriamente pelo Excel:



É assim que se faz. Gostou?

1.4.2 – Como eliminar células em branco de um intervalo de células? A sua pesquisa, lembra-se?

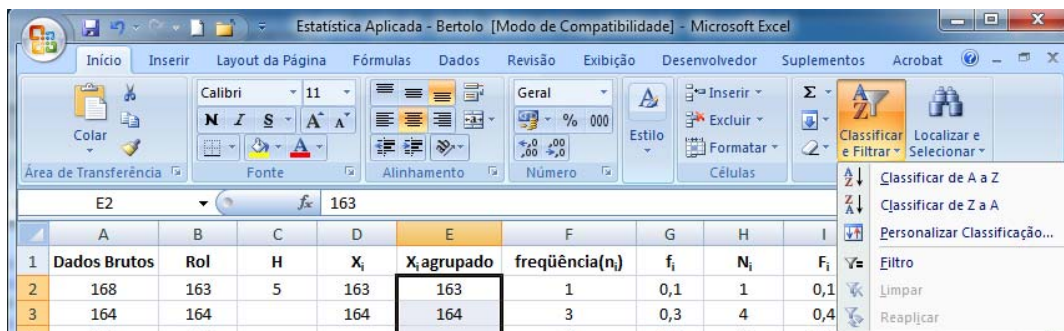
Bem, você pediu as bênçãos do Santo Google e ele o iluminou neste momento difícil que o professor proporcionou gratuitamente. Como num milagre você executou os passos abaixo e chegou lá. Cuidado para não ser canonizado.

Passo #01- Destacar o intervalo **D2:D11** e pressionar **Ctrl + C**.

Passo #02- Clique na célula **E2** e depois então, na guia **Início**, no grupo de ferramentas **Área de transferência**, selecione **Colar Valores**. Os valores da coluna **D** serão repetidos na coluna **E**, mas as células em branco ficam vazias, sem as formulas (diferentemente daquelas correspondentes na coluna **D**).

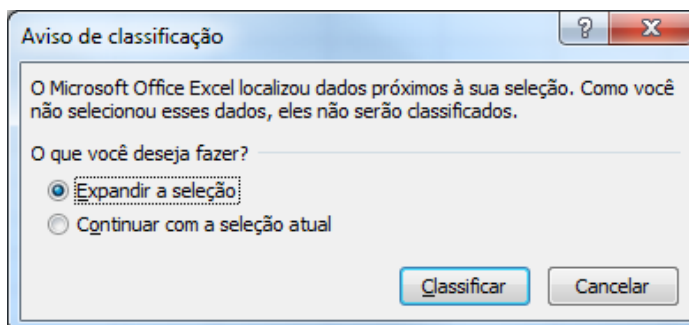
Passo #03- Pressione **ESC** para apagar aquele chato do luminoso.

Passo #04- Ainda na guia **Início**, vá ao grupo de ferramentas **Edição** e clique no botão **Classificar e Filtrar**.



Passo #05- No menu suspenso selecione **Classificar de A a Z**.

Passo #06- Na caixa de diálogo **Aviso de classificação**, clique no botão de opção **Continuar com a seleção atual**.



Passo #07- Pressione o botão **Classificar**.

O milagre aconteceu!!!! Você é um gênio abençoado! Procure agora derramar amor sobre o Planeta Terra, dê esmolas, cesta básica aos pobres, enfim você venceu!!!!

1.4.3 – Estatística Médica

A má utilização de dados médicos ou de outras ciências, com interpretações errôneas ou mal intencionadas, tem produzido um grande ceticismo em relação à estatística. Podemos encontrar muitos professores, clínicos ou mesmo profissionais de ciências básicas, que acham que "bom senso" é suficiente para tratar desses dados, que qualquer coisa que exija prova estatística não pode ter valor prático ou que procurarão um matemático se tiverem algum problema estatístico em seu trabalho.

É importante compreender, no entanto, que cada vez que se toma a média de duas leituras de pipeta, por exemplo, ou o número médio de dias de internação no hospital de pacientes com determinada doença, se está usando um método estatístico. Do mesmo modo, usamos métodos estatísticos para concluir que a pressão arterial do paciente está normal ou que um tratamento cirúrgico é melhor que outro. Deste modo, a *Estatística Médica* se constitui realmente nos princípios da *Medicina Quantitativa*. Muito da medicina depende, direta ou indiretamente, de contagem ou mensuração. Isto é óbvio em pesquisa e verdadeiro também no diagnóstico. Quando um médico, examinando um nódulo no seio de uma paciente, leva em consideração sua idade e história familiar, logo está usando dados quantitativos, ou seja, a incidência de tumores por idade e por família.

Convencionou-ser chamar de *Bioestatística* o conjunto de conceitos e métodos científicos usados no tratamento da variabilidade nas ciências médicas e biológicas. A Bioestatística fornece métodos para se tomar decisões ótimas na presença de incerteza, estabelecendo faixas de confiança para a eficácia dos tratamentos e verificando a influência de fatores de risco no aparecimento de doenças. Dentro da área biológica, compreende-se por *Biometria* a ciência que estuda as medidas de seres vivos.

A ênfase crescente do papel dos métodos quantitativos na prática da medicina torna imperativo que o estudante de medicina, assim como o médico, tenha algum conhecimento de estatística.

O estudante aprende na escola o melhor método de diagnose e terapêutica; depois de formado dependerá necessariamente de trabalhos apresentados em reuniões, jornais e revistas médicas, para aprender novos métodos de terapia, assim como os progressos em diagnose e técnica cirúrgica. Portanto, deverá estar apto a avaliar por si próprio os resultados de outros pesquisadores, deverá decidir quando uma nova técnica ou método pode substituir os antigos. Deverá estar apto a responder à pergunta da mãe que o consulta sarampo. Deverá estar apto a dar à família do paciente segurança quanto ao diagnóstico, o que pode depender de sua capacidade de avaliar adequadamente os resultados de exames laboratoriais, como também de seu conhecimento sobre a relação da idade, sexo e outras condições do paciente a uma determinada doença. Os novos conhecimentos virão através do trabalho de pesquisa realizado por ele próprio ou por outros. Deve, portanto, ser capaz de selecionar, da massa de informações, àquelas que forem válidas e que resistirem aos testes científicos rígidos. Deve desenvolver um ceticismo sadio em relação a tudo que lê.

Uma noção básica em planejamentos experimentais referentes à pesquisa médica, é o reconhecimento de que os indivíduos diferem não apenas uns dos outros, mas também em relação a si próprios, de dia para dia ou mesmo de hora para hora. Certa quantidade de variação é normal, mas a questão que desafia o médico é determinar quando uma variação específica se torna patológica (referente à doença). Para isso, o estudante deve aprender como medir a variação em indivíduos normais e definir qual é o limite de variação normal. Deve aprender que há algum erro aleatório presente em cada medida ou contagem feita. É altamente improvável que duas contagens sucessivas de glóbulos, feitas na mesma amostra de sangue, sejam idênticas.

Quando, porém, uma diferença se torna maior que o erro de mensuração? Para tratar seu paciente do melhor modo possível, o médico deve saber responder às questões como essa. Para cada medida ou determinação fornecida pelo laboratório, o médico deve conhecer a variação que é parte do próprio método, para saber quando uma dada variação representa uma mudança real no paciente. Sempre que novos métodos de terapia são introduzidos, é necessário saber se são realmente superiores, isto é, mais eficientes aos velhos métodos. Será necessário fazer-se a avaliação crítica do estudo experimental, verificando principalmente se as medidas foram realizadas de modo a produzir resultados fidedignos, se o fator em prova foi o único fator de diferença entre o grupo experimental e o grupo controle, se a diferença entre os resultados obtidos nos dois grupos foi maior que aquela que poderia ser atribuída ao acaso. Apenas depois da avaliação, através de técnicas estatísticas adequadas, da fidedignidade dos resultados, comparabilidade dos grupos experiência e controle e, significância da diferença encontrada é que podemos tirar as conclusões relativas ao novo método.

O número de falhas encontradas em publicações médicas enfatiza a necessidade de uma avaliação crítica da literatura. Para ajudar o estudante a desenvolver tal atitude são necessários certos conceitos estatísticos básicos e certa familiaridade com a terminologia mais usada.

Métodos estatísticos são essenciais no estudo de situações em que as variáveis de interesse estão sujeitas, inerentemente, a flutuações aleatórias. Este é o caso da área da saúde principalmente na Medicina. Mesmo tomando-se um grupo de pacientes homogêneos, observa-se grande variabilidade, por exemplo, no tempo de sobrevivência após um tratamento adequado. Dosagens de características hematológicas flutuam não só entre indivíduos, como também no mesmo indivíduo em ocasiões diferentes. Na realidade, há variações entre diferentes pacientes para qualquer

variável de interesse clínico. Portanto, para se estudar problemas clínicos, é necessária uma metodologia capaz de tratar a variabilidade de forma adequada.

Deve-se notar, entretanto, que ao tratar um paciente, o médico se vale da experiência de eventos anteriores, vivenciada pessoalmente ou transmitida por outros através de livros e artigos. Assim, a Estatística pode ser vista como ferramenta de organização e validação do conhecimento médico.

1.4.4 – Dados Biométricos

Nas áreas médica e biológica coletam-se dados de pessoas, de animais experimentais e de fenômenos físicos e químicos. Interessam aos pesquisadores dessas áreas dados sobre mortalidade infantil, eficiência de medicamentos, incidência de doenças, causas de morte etc.. Os dados referem-se a *variáveis*, classificadas como qualitativas, ordinais e quantitativas.

O dado qualitativo está fortemente presente nas ciências da saúde. Características de pessoas, raça, doenças, etc., são frequentemente medidas como variáveis categóricas. No entanto, mercê de sua tradição escolástica, as ciências da saúde ainda mostram preferência pelas medidas contínuas de alta precisão: medir em miligramas, milímetros etc.. Engajado no conhecimento do biólogo, o cientista da saúde tende a isentar-se de reflexão sobre a importância das estratégias de mensuração e análise providas por outras ciências, como a estatística.

Exemplo 1.1. *Com o objetivo de levantar conhecimento sobre algumas características de homens com doenças cardíacas da clínica especializada de Honolulu, no Havaí, decidiu-se pesquisar os pacientes com idade entre 45 e 67 anos. As características de interesse eram:*

- 1 - Nível de instrução
- 2 - Peso
- 3 - Altura
- 4 - Idade
- 5 - Fumante
- 6 - Atividade física
- 7 - Taxa de glicose
- 8 - Taxa de colesterol sérico
- 9 - Pressão sanguínea sistólica.

1.5 – BIBLIOGRAFIA COMPLEMENTAR

Chambers JM, Cleveland WS, Kleiner B, Tukey PA. *Graphical methods for data analysis*. Wadsworth & Brooks/Cole Publishing Company, Advanced Books & Software. Pacific Grove, California, 1983.

Daly F, Hand DJ, Jones MC, Lunn AD, McConway KJ. *Elements of Statistics*, 1999.

Dixon WJ e Massey FJ. *Introduction to Statistical Analysis*. 2nd edit. The Maple Press Company, York, 1957.

Hand DJ et al. *A handbook of small data sets*. Chapman&Hall, 1994.

Triola MF. *Introdução à Estatística*. LTC Livros Técnicos e Científicos Editora, 1999.