

5. CONSTRUÇÃO E INTERPRETAÇÃO DE TABELAS E GRÁFICOS ESTATÍSTICOS

1. Organização dos Dados: Construção de Tabelas

- 1.1. Cada linha é um registro (entrevistado), cada coluna uma pergunta (variável)
- 1.2. Cada linha é uma pergunta (variável), cada coluna um registro (entrevistado)
- 1.3. Digitação de tabelas em Planilha Eletrônica (Excel):
 - 1.3.1. Pode-se digitar no formato 1.2 e copiar e colar transposto para o formato 1.1.

1.1. Tabela mais apropriada para análise estatística, particularmente em programas de estatística

Ficha	nome	idade	Altura	sexo	Cidade	Escolaridade	Pontuação
1	AAS	17	1,62	F	Rio Preto	1º grau	110
2	CMM	72	1,70	M	Barrinha	2º grau	81
3	CNS	26	1,55	F	Jardinópolis	3º grau	162

1.2. Tabela preferida por alguns para digitação das fichas (sequência com Enter)

Ficha	1	2	3	Tipo de Variável	Escala
nome	AAS	CMM	CNS	Qualitativa Nominal	
Idade (anos)	17	72	26	Quantitativa Razão	Discreta
Altura (m)	1,62	1,70	155	Quantitativa Razão	Contínua
sexo	F	M	F	Qualitativa Nominal	
Cidade	Rio Preto	Barrinha	Jardinópolis	Qualitativa nominal	
Escolaridade	1º grau	2º grau	3º grau	Qualitativa Ordinal	
Pontuação (0 a 200)	110	81	162	Quantitativa Intervalo	Contínua

- 1.4. Obs: Escala de Likert: 1 a 5 (discordo fortemente – concordo fortemente)
 - 1.4.1. Trata-se de uma escala Qualitativa Ordinal
 - 1.4.2. Prática comum: considerar como quantitativa - intervalar e até mesmo razão
 - 1.4.2.1. Avaliação docente: média para efeito de comparação
- 1.5. Importante:
 - 1.5.1. Numerar cada ficha e lançar os dados na linha/coluna correspondente
 - 1.5.1.1. Possibilita a revisão e verificação posterior da digitação
 - 1.5.2. Definir unidade e faixa de variação de cada medida
 - 1.5.3. Revisar tabela para localizar erros de digitação (dados fora da faixa de variação)
 - 1.5.4. Dados perdidos: deixar em branco ou inserir a média da variável (bem identificada)
- 1.6. Para análise em alguns softwares de Estatística:
 - 1.6.1. Alguns programas exigem que as variáveis categóricas sejam reduzidas a números
 - 1.6.1.1. Sexo M = 1; F=0 Cidade: Barrinha = 1; Rio Preto = 2; Jardinópolis = 3, etc.

2. Medidas Descritivas de Dados Observados (Estatística Descritiva)

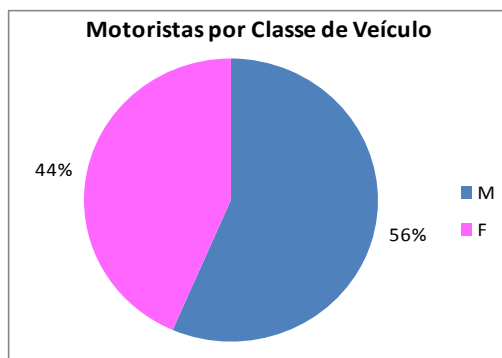
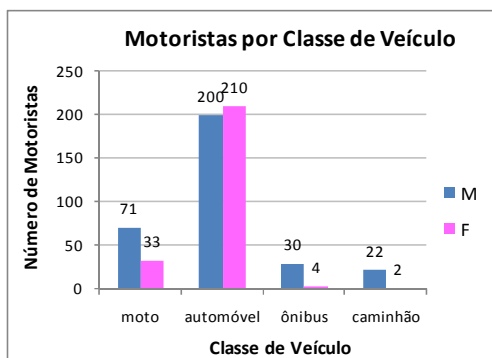
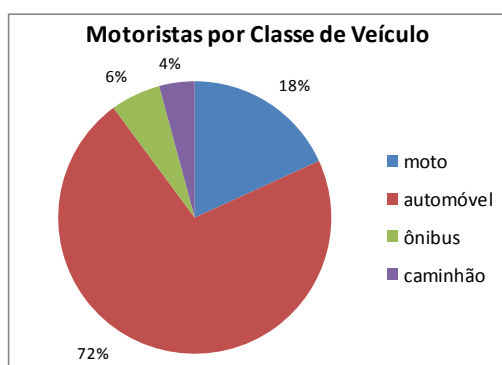
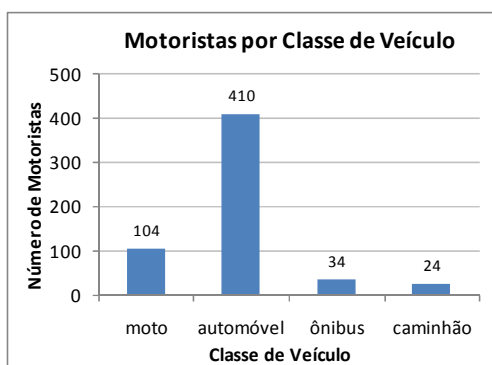
- 2.1. Estatística Descritiva: resume algumas características dos dados coletados
- 2.2. Dados Qualitativos: agrupa indivíduos em classes
 - 2.2.1. Frequência: nº de dados em cada classe (contagem simples)
 - 2.2.2. Frequência relativa: porcentagem de indivíduos em cada classe
 - 2.2.2.1. Total da classe dividido pelo total geral
 - 2.2.3. Tabelas cruzadas: duas variáveis correlacionadas, por exemplo, sexo e veículo
- 2.3. Quantitativos: frequência (grupos convencionados), frequência relativa (%) e também medidas centro (média, mediana, moda), dispersão (variância, desvio), correlação

Frequência Absoluta				
MOTORISTA	moto	automóvel	ônibus	caminhão
M	71	200	30	22
F	33	210	4	2
SubTotal	104	410	34	24
TOTAL	572			

Frequência Relativa					
MOTORISTA	moto	automóvel	ônibus	caminhão	SubTotal
M	12%	35%	5%	4%	56%
F	6%	37%	1%	0%	44%
SubTotal	18%	72%	6%	4%	100%

3. Gráficos

- 3.1. Obs: existem muitos tipos de gráfico, depende de o que se quer evidenciar
- 3.2. Frequência: Qualitativos: gráficos de barra
 - 3.2.1. Eixo x: classe - eixo Y: contagem dentro da classe
 - 3.2.1.1. No gráfico há espaço entre as colunas (barras)
- 3.3. Quantitativos: Histograma
 - 3.3.1. Eixo x: classe - eixo Y: contagem dentro da classe
 - 3.3.2. A classe é definida pelo pesquisador (intervalo de idade, de altura, etc.)
 - 3.3.2.1. No gráfico não há espaço entre as colunas (barras)
- 3.4. Frequência relativa: gráficos de pizza
 - 3.4.1. Porcentagem de elementos dentro da classe
 - 3.4.1.1. Total da classe dividido pelo total geral, expresso em %
- 3.5. Correlação: gráficos xy

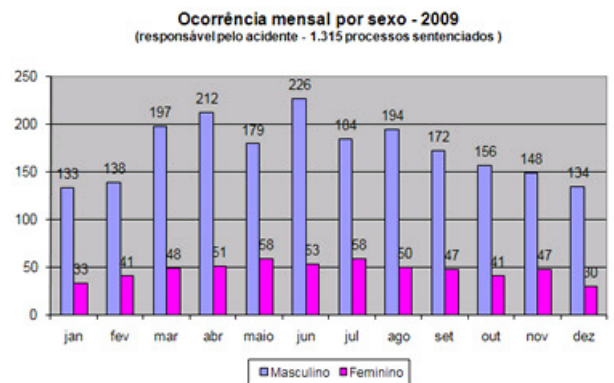
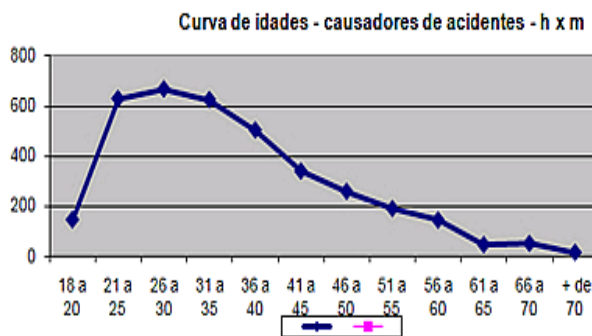


Obs: número absoluto, não informa número de motoristas em cada idade

- 3.6. Determinação do número de classes (i= número de classes; n = número de elementos)

3.6.1. Critério da Raiz: $i \cong \sqrt{n}$

3.6.2. Fórmula de Stuges: $i \cong 1 + 3,3 \log n$



Atenção: os dois gráficos acima indicam número absoluto de ocorrências, sem considerar o número de pessoas em cada classe (idosos que dirigem, homens e mulheres que dirigem). Seria importante incluir gráficos com a porcentagem dentro de cada classe.

Exemplo de gráfico mais complexo, que resume muitas informações

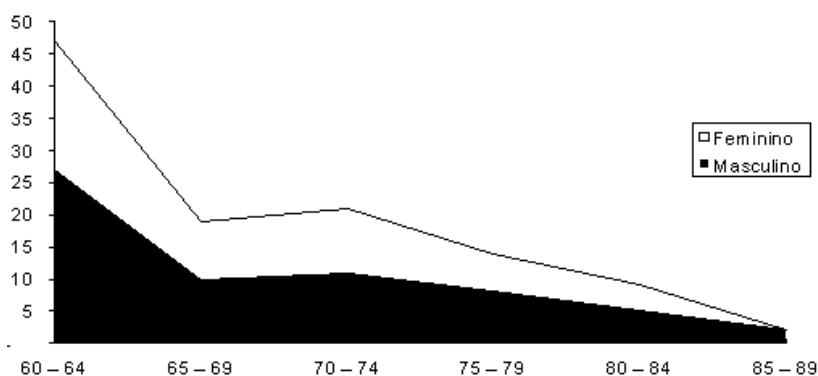
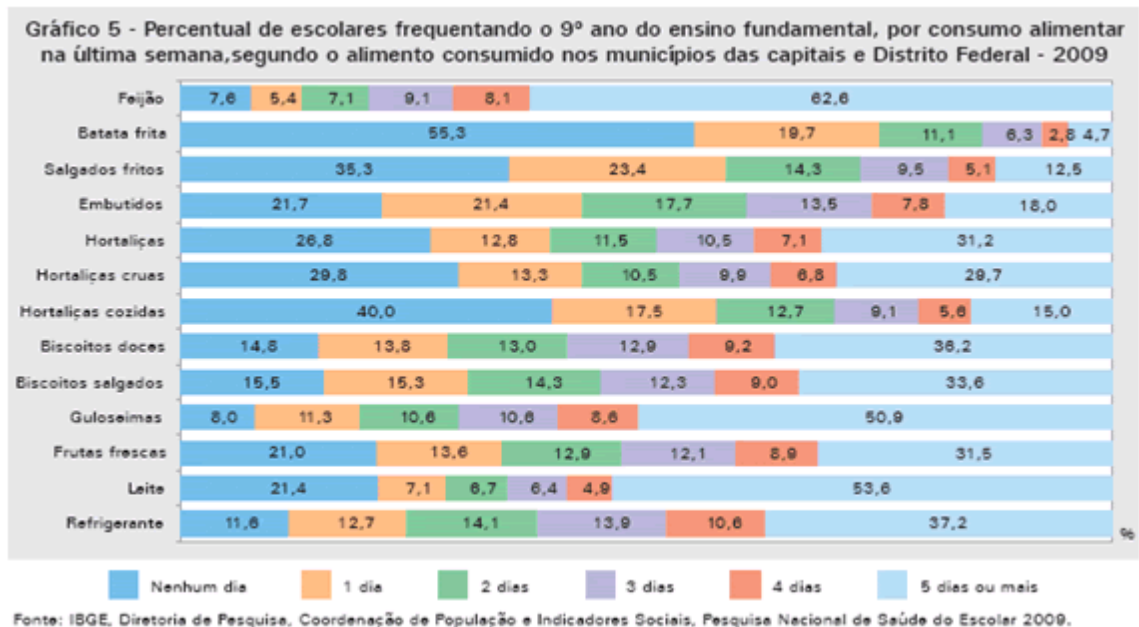


Figura 1 - Distribuição do número de idosos acidentados no trânsito, segundo sexo e idade, Ribeirão Preto-SP, 1998

O gráfico acima também não indica a frequência relativa, isto é, não considera o número absoluto de pessoas em cada classe. Isso não significa que o gráfico está incorreto e sim que o foco do gráfico está no número absoluto de pessoas acidentadas.

4. Histograma

Table 3-1 Colorado Live Births by Mother's Age									
Year	Total births	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
1975	40,148	88	6,627	14,533	12,565	4,885	1,211	222	16
1980	49,716	57	6,530	16,642	16,081	8,349	1,842	198	12
1985	55,115	90	5,634	16,242	18,065	11,231	3,464	370	13
1990	53,491	91	5,975	13,118	16,352	12,444	4,772	717	15
1995	54,310	134	6,462	12,935	14,286	13,186	6,184	1,071	38
2000	65,429	117	7,546	15,865	17,408	15,275	7,546	1,545	93

* Note: The sum of births may not add up to the total number of births due to unknown or unusually high age (50 and over) of the mother.

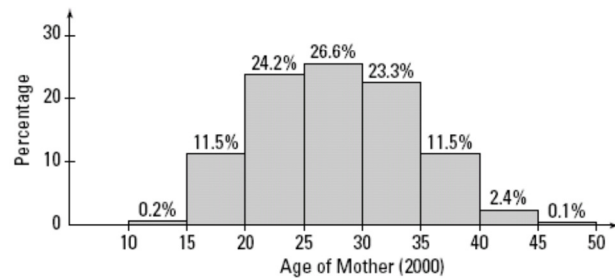
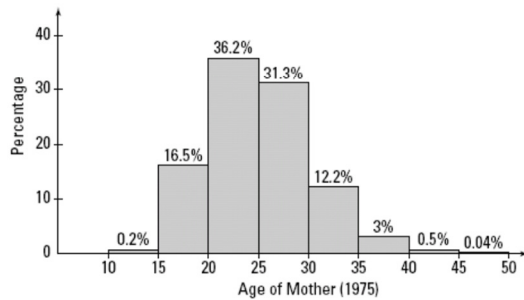
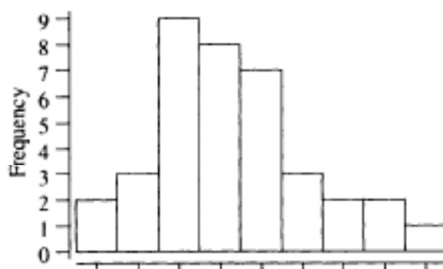
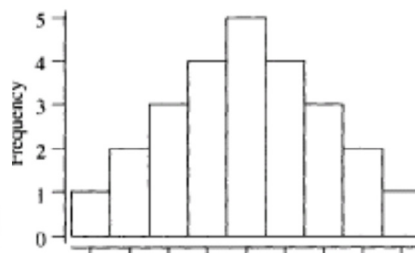


Figure 3-4: Colorado live births, by age of mother for 1975 and 2000.

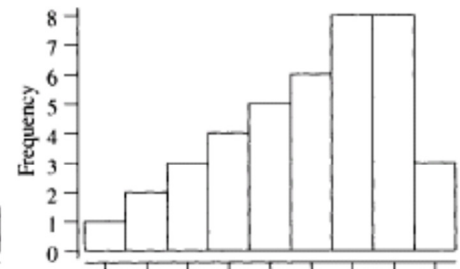
4.1. Inclinação do Histograma (skewness)



Inclinado à direita



Sem inclinação



Inclinado à esquerda

4.2. Frequência Acumulada

Tabela 1: Horas assistindo TV por dia (pesquisa com 20.200 crianças nas férias)

Hours	Frequency	Upper limit	Cumulative frequency
0	0	0	0
0-1	4,300	1	4,300
1-3	6,900	3	4,300+6,900 = 11,200
3-5	4,900	5	4,300+6,900+4,900 = 16,100
5-10	2,000	10	4,300+6,900+4,900+2,000 = 18,100
10-24	2,100	24	4,300+6,900+4,900+2,000+2,100 = 20,200

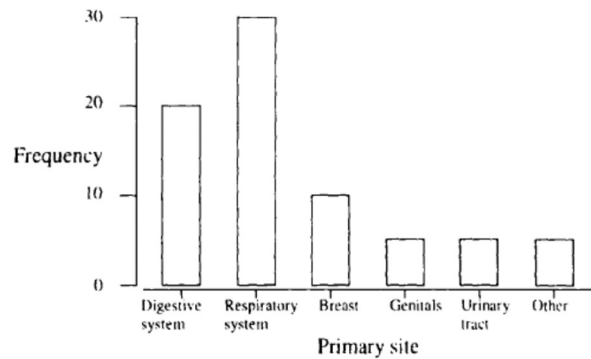


5. Exemplos:

5.1. Dados Qualitativos:

Table 2.3

Primary site	Frequency
Digestive system	20
Respiratory	30
Breast	10
Genitals	5
Urinary tract	5
Other	5



5.2. Dados Quantitativos

Table 2.5

IQ score	Frequency
80-94	8
95-109	14
110-124	24
125-139	16
140-154	13

Table 2.6

Class limits	Class boundaries	Class width	Class marks
80-94	79.5-94.5	15	87.0
95-109	94.5-109.5	15	102.0
110-124	109.5-124.5	15	117.0
125-139	124.5-139.5	15	132.0
140-154	139.5-154.5	15	147.0

EXAMPLE 2.7 Group the following weights into the classes 100 to under 125, 125 to under 150, and 150 to under 175.

111	120	127	129	130	145	145	150	153	155	160
161	165	167	170	171	174	175	177	179	180	180
185	185	190	195	195	201	210	220	224	225	230
245	248									

Table 2.7

Weight	Frequency
100 to under 125	2
125 to under 150	5
150 to under 175	10
175 to under 200	10
200 to under 225	4
225 to under 250	4

6. Obs: Sistema de Caule e Folhas

1.1. Forma visual de apreender uma distribuição (pouco utilizada)

1.2. Exemplo: Pontuação de 30 estudantes num jogo qualquer (boliche, por exemplo)

1.2.1. Neste caso, registra-se a dezena na coluna da esquerda e cada instância da unidade na linha correspondente à dezena

