

Capítulo 1 – Descrição de Dados: Tabelas e Gráficos

1. No quadro a seguir é dada uma relação de nomes de variáveis de pesquisa (qualitativas ou quantitativas) e categorias de variáveis qualitativas. Identifique, primeiramente, as 15 variáveis. Para as variáveis qualitativas, identifique as respectivas categorias e para as variáveis quantitativas, classifique-as como discreta ou contínua.

Pousada	Pequena	Norte
Rural	Sexo do animal	Resort
Objetivo da viagem	Hotel histórico	Área do território
Cor ou raça	Turismo de pesca	Lazer
Sobrepeso	Amarela	Nordeste
Tipo de turismo	Turismo social	Hotel fazenda
Situação do domicílio	Turismo de esporte	Peso normal
Média	Turismo de saúde	Sul
Rendimento familiar	Sem declaração	Fêmea
Sudeste	Centro-Oeste	Cama & Café
Economicamente ativa	Não economicamente ativa	Densidade demográfica
Classificação do paciente	Turismo rural	Condição de atividade
Macho	Turismo de negócios e eventos	Hotel
Indígena	Visita familiar	Região geográfica
Preta	Meio de hospedagem	Abaixo do peso
Urbana	Ecoturismo	Branca
Turismo de estudos e intercâmbio	População do território	Flat/apart hotel
Classificação da empresa	Serviço	Distância entre o município e a capital
Turismo de aventura	Obeso	Turismo cultural
Grande	Parda	Turismo de sol e praia

2. Faça cada um dos cálculos indicados a seguir, apresentando os resultados com

- (i) 2 casas decimais
(ii) 3 casas decimais
(iii) 4 casas decimais

(a) $100 \times \frac{20}{30}$

(b) $2 + \frac{5}{7}$

(c) $\frac{17}{4}$

(d) $\frac{17}{31} + \frac{3}{31}$

(e) $\frac{12}{19} \times \frac{2}{13}$

3. Para o Departamento Financeiro da empresa analisada na apostila, obteve-se a seguinte informação sobre o sexo dos 23 funcionários:

M F F M M M F F M M M M
M F M M F F M M M F F

onde M = Masculino e F = Feminino. Construa uma tabela de frequências para esses dados, bem como os gráficos de colunas e de setores.

4. Na Tabela 1, temos informações sobre o sexo, a matéria predileta (**P**ortuguês, **M**atemática, **H**istória, **G**eografia ou **C**iências) no ensino médio e a nota (número de questões certas) em um teste de múltipla escolha com 10 questões de matemática, ministrado no primeiro dia de aula de calouros de uma universidade (dados fictícios).

- (a) Classifique as variáveis envolvidas.
- (b) Construa a tabela de frequências apropriada para cada uma das variáveis.
- (c) Construa gráficos apropriados para ilustrar as distribuições de frequência.

Tabela 1: Dados sobre sexo, matéria predileta e nota de alunos

Sexo	Predileta	Nota	Sexo	Predileta	Nota	Sexo	Predileta	Nota
F	H	5	M	M	2	M	H	3
M	M	8	M	G	4	M	M	5
F	P	8	M	G	9	F	P	5
F	H	6	M	M	7	F	G	5
M	C	5	M	M	1	M	C	7
M	H	6	F	P	8	M	H	4
F	M	8	F	G	5	F	M	7
F	P	4	M	G	9	F	P	7
F	H	2	M	P	5	F	M	6
M	C	6	F	M	8	M	G	6
F	P	8	F	G	6	M	H	9
F	M	8	M	P	5	M	G	6
F	M	7	M	P	5	F	M	5
F	M	5	F	P	9	M	M	8

5. Na Tabela 2, temos as frequências acumuladas do número de sinistros por apólice de seguro do ramo Automóveis.

- (a) Para quantas apólices não houve qualquer sinistro?
- (b) Para quantas apólices houve mais de um sinistro?
- (c) Para quantas apólices houve no máximo 2 sinistros?
- (d) Complete a tabela, calculando as frequências simples absolutas e relativas e também as frequências acumuladas relativas.

Tabela 2: Número de sinistros por apólice

Número de sinistros	Número de apólices
0	2913
≤ 1	4500
≤ 2	4826
≤ 3	4928
≤ 4	5000

6. Em um conjunto de dados referentes ao gasto médio por pessoa em uma amostra de 50 turistas visitando Niterói, o valor mínimo foi de R\$22,35 e o valor máximo foi de R\$127,50. Quais seriam os limites de classes para uma distribuição de frequências com 5 classes de mesmo tamanho, se esses limites devem ser números *inteiros*, escolhidos de forma a se fazer o menor ajuste necessário na amplitude geral?

7. Escolha, na Figura 1, o gráfico que melhor descreve cada uma das afirmações a seguir.

- (a) Mais da metade dos elementos pertence à categoria A.
- (b) Metade dos elementos está na categoria A e a outra metade está igualmente distribuída entre as categorias B, C, D.
- (c) Os elementos estão divididos uniformemente entre as 4 categorias.
- (d) Metade dos elementos está na categoria A. Em relação à outra metade, as categorias C e D têm metade dos elementos da categoria B.

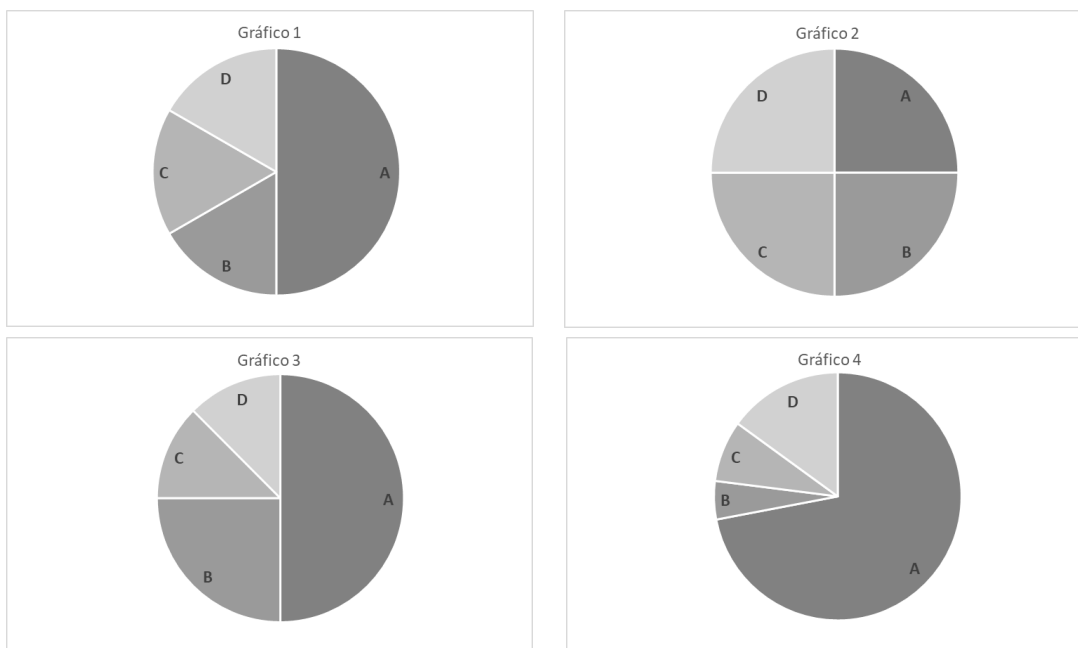


Figura 1: Gráficos para o Exercício 7

8. A seguir, são dadas quatro afirmativas, duas das quais descrevem corretamente alguma informação contida na Tabela 3. Identifique essas afirmativas e explique porque as outras duas estão erradas.

- (a) Menos de 10 domicílios têm renda inferior a 82 unidades monetárias.
- (b) 95% dos domicílios têm renda inferior a 50 unidades monetárias.
- (c) Mais de 80% dos domicílios têm renda inferior a 10 unidades monetárias.
- (d) A renda máxima entre esses 500 domicílios é de 50 unidades monetárias.

Tabela 3: Distribuição da renda de 500 domicílios (dados fictícios)

Renda (u.m.)	Número de Domicílios		Número Acumulado de Domicílios	
	Absoluto	Relativo (%)	Absoluto	Relativo (%)
Menos de 3	285	57,0	285	57,0
3 † 5	80	16,0	365	73,0
5 † 10	45	9,0	410	82,0
10 † 20	35	7,0	445	89,0
20 † 50	30	6,0	475	95,0
50 ou mais	25	5,0	500	100,00
Total	500	100,00		

9. Complete a tabela abaixo:

Classe	Freq. Simples		Freq.Acumulada
	Absoluta	Relativa (%)	Absoluta
0	8	4,0	
1	16		
2			60
3	54		
4			144
5			166
6	20		
7			
TOTAL			

10. Construa uma tabela de frequências com 4 classes de mesmo tamanho para os dados a seguir. Os limites das classes devem ser números inteiros. Faça o ajuste mínimo necessário na amplitude.

Obs: Os valores mínimo e máximo dos dados são 12,8 e 59,4.

33,0	43,5	22,7	24,8	35,1	55,8	27,7	58,9	37,0	46,0
54,8	59,0	55,1	59,4	52,9	14,8	12,8	34,0	13,8	27,2
19,3	25,8	27,4	53,0	57,6	13,5	43,1	38,5	40,9	17,9
43,6	37,1	34,3	31,4	28,9	13,5	25,2	42,6	46,5	29,3

11. Complete a tabela abaixo:

Classe	Freq. Simples		Freq.Acumulada	
	Absoluta	Relativa (%)	Absoluta	Relativa (%)
20 † 30			8	
30 † 40	12	10,00		
40 †			46	
50 † 60			105	
† 70				
TOTAL				

12. A seguir temos a relação de notas de 150 candidatos a um emprego. Construa uma tabela de frequências com 5 classes de mesmo tamanho. Os limites das classes devem ser números inteiros e o limite inferior da primeira classe tem que ser 0.

Obs: Os valores mínimo e máximo das notas são 0,9 e 99,0.

73,1	78,3	41,1	0,7	4,1	72,9	59,6	81,9	43,6	47,3	60,2	25,9	58,3	37,3	65,0
17,7	57,9	52,9	35,1	23,7	77,9	90,1	25,7	60,0	56,7	43,6	37,1	96,2	96,5	91,5
64,1	5,3	49,8	46,0	2,1	58,6	37,2	76,9	61,4	38,3	32,0	52,0	13,0	81,6	88,0
51,7	68,9	93,5	39,8	13,3	61,8	55,6	4,2	26,5	49,5	42,7	58,1	53,4	42,5	8,0
67,9	95,0	61,9	94,6	96,8	47,9	69,5	45,8	33,7	94,6	48,5	9,3	57,8	38,4	47,9
63,1	25,8	69,4	58,9	39,5	89,5	26,1	22,2	8,6	51,4	60,9	29,3	44,1	22,2	68,0
64,8	48,7	57,9	58,9	50,1	21,5	48,1	71,6	43,4	73,1	21,6	27,1	79,1	74,3	16,2
52,2	78,4	41,9	49,2	36,7	48,9	31,1	20,3	74,1	98,9	84,8	55,7	95,2	44,9	43,8
72,7	19,5	81,4	42,6	89,3	34,2	79,8	37,8	92,4	20,6	73,4	71,0	46,4	43,7	30,3
28,9	49,4	72,5	50,4	89,0	53,7	60,5	25,7	99,0	67,8	0,9	28,0	93,6	10,6	2 9,1

Capítulo 2 – Descrição de Dados: Resumos Numéricos

13. Para os dados do Exercício 10, faça o seguinte:

- (a) Arredonde os números para números inteiros.
- (b) Construa o diagrama de ramo e folhas, tendo o cuidado de especificar a escala.
- (c) Calcule a média dos dados arredondados, sabendo que $\sum_{i=1}^{40} = 1441$.
- (d) Calcule os 3 quartis.
- (e) Construa o boxplot.

14. Para os dados do Exercício 12, faça o seguinte:

- (a) Arredonde os números para números inteiros.
- (b) Construa o diagrama de ramo e folhas, tendo o cuidado de especificar a escala.
- (c) Calcule a média dos dados arredondados, sabendo que $\sum_{i=1}^{150} = 7792$.
- (d) Calcule os 3 quartis.
- (e) Construa o boxplot.

15. Quatro amigos trabalham em um supermercado em tempo parcial com os seguintes salários horários:

Pedro:	R\$ 3,50	João:	R\$ 2,60
Marcos:	R\$ 3,80	Luiz:	R\$ 2,20

Se Pedro trabalha 10 horas por semana, João 12 horas, Marcos 15 horas e Luiz 8 horas, qual é o salário horário médio desses quatro amigos?

16. Na UFF, o coeficiente de rendimento (CR) semestral dos alunos é calculado como uma média das notas finais nas disciplinas cursadas, levando em conta a carga horária (ou crédito) das disciplinas, de modo que disciplinas com maior carga horária têm maior peso no CR. Suponha que um aluno tenha cursado cinco disciplinas em um semestre, obtendo médias finais de 7, 5; 6, 1; 8, 3; 6, 5; 7, 5. As três primeiras disciplinas tinham carga horária de 4 horas semanais, a quarta, carga horária de 6 horas e a última, duas horas semanais. Calcule o CR do aluno nesse semestre.

17. Em uma pesquisa sobre atividades de lazer realizada com uma amostra de 20 alunos de um campus universitário, perguntou-se o número de horas que os alunos gastaram “navegando” na internet na semana anterior. Os resultados obtidos foram os seguintes:

15	30	18	8	10	12	15	14	12	10
18	12	6	20	18	16	10	12	15	9

- (a) Construa o boxplot. Há valores discrepantes segundo a regra $1,5 \times AIQ$?
 - (b) Calcule os escores padronizados. Há valores discrepantes segundo o teorema de Chebyshev?
18. No final do ano 2005, o dono de um pequeno escritório de administração deu a seus oito funcionários uma gratificação de 250 reais, paga junto com o salário de dezembro. Se em novembro o salário médio desses funcionários era de 920 reais, qual o salário médio em dezembro? Que propriedades você utilizou para chegar a esse resultado?
19. No mês de dissídio de determinada categoria trabalhista, os funcionários de uma empresa tiveram reajuste salarial de 8,9%. Se no mês anterior ao dissídio o salário médio desses funcionários era de 580 reais, qual o valor do salário médio depois do reajuste? Que propriedades você utilizou para chegar a esse resultado?
20. O número médio de empregados das empresas industriais do setor de fabricação de bebidas em determinado momento era de 117 empregados, enquanto o número mediano era de 27. Dê uma explicação para a diferença entre essas medidas de tendência central.

Capítulo 3 – Correlação

21. Considere os dados sobre consumo de cigarros e mortes por câncer de pulmão, apresentados na Tabela 4. Calcule o coeficiente de correlação entre as variáveis, interpretando o resultado.

Tabela 4: Consumo de cigarros (X) e morte por câncer de pulmão (Y)

País	X	Y	Países	X	Y
Islândia	240	63	Holanda	490	250
Noruega	255	100	Suíça	180	180
Suécia	340	140	Finlândia	1125	360
Dinamarca	375	175	Grã-Bretanha	1150	470
Canadá	510	160	Estados Unidos	1275	200
Austrália	490	180			

22. Calcule o coeficiente de correlação entre preço de venda e área da casa, para os dados da Tabela 5, interpretando o resultado. Para facilitar o seu trabalho, você tem os seguintes resultados:

$$\sum_{i=1}^{59} X_i = 14433 \quad \sum_{i=1}^{59} Y_i = 10470 \quad \sum_{i=1}^{59} X_i Y_i = 2667063$$

$$\sum_{i=1}^{59} X_i^2 = 3736397 \quad \sum_{i=1}^{59} Y_i^2 = 1976794$$

Tabela 5: Vendas de casas em Boulder, Colorado (1995)

Preço (Y) (1000 US\$)	Área (X) (m ²)	Preço (Y) (1000 US\$)	Área (X) (m ²)	Preço (Y) (1000 US\$)	Área (X) (m ²)
113	126	163	227	186	228
114	158	168	228	187	219
120	126	168	249	187	222
120	126	169	244	188	279
122	158	169	263	188	249
123	126	170	234	190	317
129	229	171	283	192	304
137	196	172	286	193	195
140	262	173	268	195	217
142	272	175	223	195	232
143	189	175	270	200	234
146	158	175	231	200	322
146	218	176	249	200	304
148	276	177	285	207	300
149	218	178	243	270	252
152	302	178	251	290	322
153	168	180	279	300	353
157	302	180	189	320	349
157	289	181	153	328	388
160	277	185	316		

Capítulo 1 – Descrição de Dados: Tabelas e Gráficos

1. (a) Objetivo da viagem: Qualitativa
 - Serviço
 - Lazer
 - Visita familiar
- (b) Cor ou raça: Qualitativa
 - Preta
 - Branca
 - Indígena
 - Amarela
 - Parda
 - Sem declaração
- (c) Tipo de turismo: Qualitativa
 - Turismo de estudos e intercâmbio
 - Turismo de aventura
 - Turismo de pesca
 - Turismo social
 - Turismo de esporte
 - Turismo de saúde
 - Turismo rural
 - Turismo de negócios e eventos
 - Ecoturismo
 - Turismo de sol e praia
 - Turismo cultural
- (d) Situação do domicílio: Qualitativa
 - Urbana
 - Rural
- (e) Rendimento familiar: Quantitativa Contínua
- (f) Classificação do paciente: Qualitativa
 - Sobrepeso
 - Obeso
 - Peso normal
 - Abaixo do peso
- (g) Classificação da empresa: Qualitativa
 - Pequena
 - Média
 - Grande
- (h) Densidade demográfica: Quantitativa Contínua
- (i) Sexo do animal: Qualitativa
 - Macho
 - Fêmea
- (j) Meio de hospedagem: Qualitativa
 - Flat/apart hotel
 - Hotel
 - Cama & Café

- Pousada
 - Hotel histórico
 - Hotel fazenda
- (k) População do território: Quantitativa Discreta
- (l) Área do território: Quantitativa Contínua
- (m) Condição de atividade: Qualitativa
- Economicamente ativa
 - Não economicamente ativa
- (n) Região geográfica: Qualitativa
- Norte
 - Nordeste
 - Sudeste
 - Sul
 - Centro-Oeste
- (o) Distância entre o município e a capital: Quantitativa Contínua

2. (a) $100 \times \frac{20}{30} = 66,66666667$

(i) 2 casas decimais: 66,67

(ii) 3 casas decimais: 66,667

(iii) 4 casas decimais: 66,6667

(b) $2 + \frac{5}{7} = 2,714285714285714$

(i) 2 casas decimais: 2,71

(ii) 3 casas decimais: 2,714

(iii) 4 casas decimais: 2,7143

(c) $\frac{17}{4}$

(i) 2 casas decimais: 4,25

(ii) 3 casas decimais: 4,250

(iii) 4 casas decimais: 4,2500

(d) $\frac{17}{31} + \frac{3}{31}$

Primeira solução: fazendo cada parcela individualmente

$$\frac{17}{31} = 0,5483870967741935 \quad \frac{3}{31} = 0,0967741935483871$$

- (i) 2 casas decimais: $0,55 + 0,10 = 0,65$
- (ii) 3 casas decimais: $0,548 + 0,097 = 0,645$
- (iii) 4 casas decimais: $0,5484 + 0,0968 = 0,6452$

Segunda solução: fazendo a soma primeiro

$$\frac{17}{31} + \frac{3}{31} = \frac{20}{31} = 0,6451612903225806$$

- (i) 2 casas decimais: 0,65
- (ii) 3 casas decimais: 0,645
- (iii) 4 casas decimais: 0,6452

Embora nesse exemplo não tenha dado diferença, é sempre recomendável fazer o menor número de divisões e multiplicações possível antes de arredondar.

(e) $\frac{12}{19} \times \frac{2}{13}$

Primeira solução: fazendo cada fator individualmente

$$\frac{12}{19} = 0,6315789473684211 \quad \frac{2}{13} = 0,1538461538461538$$

- (i) 2 casas decimais: $0,63 \times 0,15 = 0,09$
- (ii) 3 casas decimais: $0,632 \times 0,154 = 0,097$

(iii) 4 casas decimais: $0,6316 \times 0,1538 = 0,0971$

Segunda solução: fazendo a multiplicação primeiro

$$\frac{12}{19} \times \frac{2}{13} = \frac{24}{247} = 0,097165991902834$$

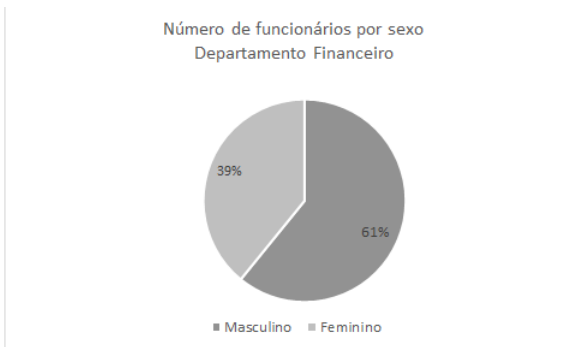
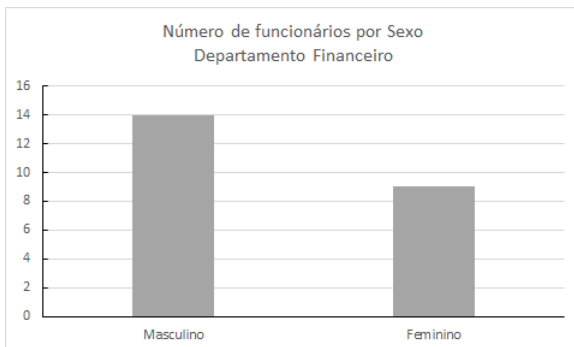
(i) 2 casas decimais: 0,10

(ii) 3 casas decimais: 0,097

(iii) 4 casas decimais: 0,0972

3. .

Funcionários do Depto Financeiro		
Sexo	Número	%
Masculino	14	60,87
Feminino	9	39,13
Total	23	100,00



4. (a) Sexo: Qualitativa

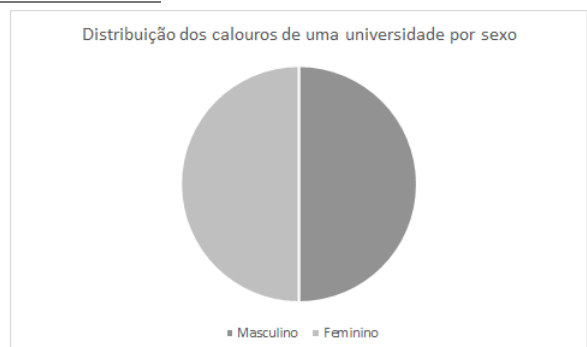
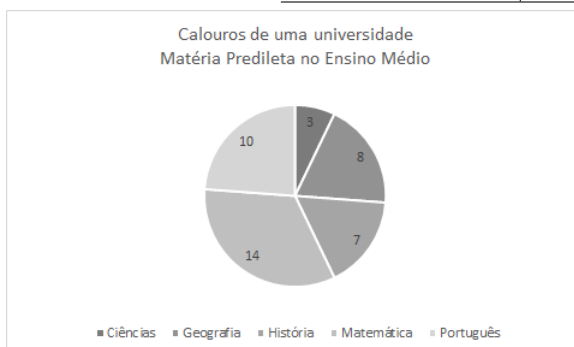
Matéria Predileta: Qualitativa

Nota: Quantitativa discreta

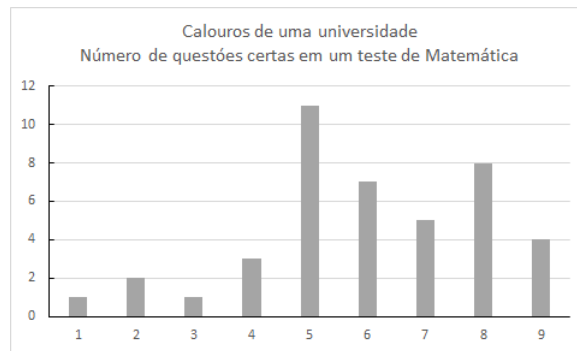
(b) .

Calouros de uma universidade		
Sexo	Número	%
Masculino	21	50,0
Feminino	21	50,0
Total	23	100,00

Calouros de uma universidade		
Matéria predileta	Número	%
Ciências	3	7,14
Geografia	8	19,05
História	7	16,67
Matemática	14	33,33
Português	10	23,81
Total	42	100,00



Calouros de uma universidade				
Nota	Número	%	Número acumulado	%
1	1	2,38	1	2,38
2	2	4,76	3	7,14
3	1	2,38	4	9,52
4	3	7,14	7	16,67
5	11	26,19	18	42,86
6	7	16,67	25	59,52
7	5	11,90	30	71,43
8	8	19,05	38	90,48
9	4	9,52	42	100,00
Total	42	100,00		



5. Note que a tabela nos dá as frequências acumuladas. Da última classe, vemos que há 5000 apólices.

- (a) 2913 apólices com 0 sinistro
 (b) Mais de um sinistro - temos que tirar as apólices com 1 ou menos.
 Mais de um sinistro: $5000 - 4500 = 500$ apólices
 (c) No máximo 2 é o mesmo que ≤ 2 . Logo, são 4826 apólices.
 (d) .

Número de sinistros	Apólices		Apólices	
	Número	%	Número acumulado	%
0	2913	58,26	2913	58,26
1	1587	31,74	4500	90,00
2	326	6,52	4826	96,52
3	102	2,04	4928	98,56
4	72	1,44	5000	100,00

Note como cada frequência simples foi obtida:

$$1587 = 4500 - 2913$$

$$326 = 4826 - 4500$$

$$102 = 4928 - 4826$$

$$72 = 5000 - 4928$$

6. Amplitude exata: $127,50 - 22,35 = 105,15$

Próximo múltiplo de 5: 110

Comprimento de classe: $110/5 = 22$

Distribuindo a diferença nos dois extremos, as classes seriam:

$$20 \text{ † } 42 \quad 42 \text{ † } 64 \quad 64 \text{ † } 86 \quad 86 \text{ † } 108 \quad 108 \text{ † } 130$$

7. (a) Gráfico 4
 (b) Gráfico 1
 (c) Gráfico 2
 (d) Gráfico 3.

8. • Afirmativa (a) - errada
 Note que essa afirmativa envolve números que constam da terceira linha da tabela. Mas o que temos aí é o seguinte: 82% dos domicílios têm renda inferior a 10 u.m.
- Afirmativa (b) - Correta
 Note que essa afirmativa envolve a quinta linha da tabela: na classe 20 † 50 temos uma frequência acumulada de 95%.
- Afirmativa (c) - Correta
 82% dos domicílios (portanto, mais de 80%) têm renda inferior a 10 u.m.
- Afirmativa (d) - Errada
 Na última classe não temos informação sobre a renda máxima; só podemos dizer que a renda máxima é maior que 50 u.m.

9. .

Classe	Freq. Simples		Freq. Acumulada	
	Absoluta	Relativa (%)	Absoluta	Relativa (%)
0	8	4,0	8	4,0
1	16	8,0	24	12,0
2	36	18,0	60	30,0
3	54	27,0	114	57,0
4	30	15,0	144	72,0
5	22	11,0	166	83,0
6	20	10,0	186	93,0
7	14	7,0	200	100,0
TOTAL	200	100,0		

Passo 1: Número total n de observações é obtido a partir das informações da primeira classe e a com isso todas as frequências relativas podem ser obtidas:

$$4,0\% = 100 \times \frac{8}{n} \Rightarrow 4n = 800 \Rightarrow n = 200$$

Passo 2: Frequências acumuladas da classe 0

Passo 3: Frequências acumuladas da classe 1: $16 + 8 = 24$

Passo 4: Frequência simples da classe 2: $60 - 24 = 36$

Passo 5: Frequência acumulada da classe 3: $60 + 54 = 114$

Passo 6: Frequência simples da classe 4: $144 - 114 = 30$

Passo 7: Frequência simples da classe 5: $166 - 144 = 22$

Passo 8: Frequência acumulada da classe 6: $166 + 20 = 186$

Passo 9: Frequência acumulada da classe 7 é o número total de observações

Passo 8: Frequência simples da classe 7: $200 - 186 = 14$

10. Amplitude exata: $59,4 - 12,8 = 46,6$

Próximo múltiplo de 4: $48 = 46,6 + 1,4$

Comprimento de classe: $48/4 = 12$

Limite inferior da primeira classe: $12 = 12,8 - 0,8$.

Começando de 12 e somando a nova amplitude 48, o novo limite superior será 60, que inclui o valor máximo dos dados. Logo, não teremos problema ao usar 12 como valor mínimo.

Valor dos dados	Frequência simples		Frequência acumulada	
	Absoluta	Relativa (%)	Absoluta	Relativa (%)
12 † 24	8	20,0	8	20,0
24 † 36	13	32,5	21	52,5
36 † 48	10	25,0	31	77,5
48 † 60	9	22,5	40	100,00
Total	40	100,00		

11. A tabela completa é

Classe	Freq. Simples		Freq. Acumulada	
	Absoluta	Relativa (%)	Absoluta	Relativa (%)
20 - 30	8	6,67	8	6,67
30 - 40	12	10,00	20	16,67
40 - 50	26	21,67	46	38,33
50 - 60	59	49,17	105	87,50
60 - 70	15	12,50	120	100,00
TOTAL				

12. Amplitude exata: $99,0 - 0,9 = 98,1$

Próximo múltiplo de 5: $100 = 98,1 + 1,9$

Comprimento de classe: $100/5 = 20$

Limite inferior da primeira classe: $0 = 0,9 - 0,9$.

Começando de 0 e somando a nova amplitude 100, o novo limite superior será 100, que inclui o valor máximo dos dados. Logo, não teremos problema ao usar 0 como valor mínimo.

Valor dos dados	Frequência simples		Frequência acumulada	
	Absoluta	Relativa (%)	Absoluta	Relativa (%)
0 - 20	15	10,00	15	10,00
20 - 40	33	22,00	48	32,00
40 - 60	47	31,33	95	63,33
60 - 80	34	22,67	129	86,00
80 - 100	21	14,00	150	100,00
Total	150	100,00		

Capítulo 2 – Descrição de Dados: Resumos Numéricos

13. (a) Os números arredondados são

33	44	23	25	35	56	28	59	37	46
55	59	55	59	53	15	13	34	14	27
19	26	27	53	58	14	43	39	41	18
44	37	34	31	29	14	25	43	47	29

(b) .

Escala

$$1 \mid 3 = 13$$

1	3	4	4	4	5	8	9		
2	3	5	<u>5</u>	<u>6</u>	7	7	8	9	9
3	1	3	4	<u>4</u>	<u>5</u>	7	7	9	
4	1	3	3	4	4	<u>6</u>	<u>7</u>		
5	3	3	5	5	6	8	9	9	9

(c) $\bar{x} = \frac{1441}{40} = 36,025$

(d) $n = 40$ – par

Mediana: deixa 20 observações abaixo e 20 acima; logo, a mediana é a média da vigésima e da vigésima primeira observações:

$$Q_2 = \frac{34 + 35}{2} = 34,5$$

O primeiro quartil é a mediana das 20 menores observações; como 20 é par, o primeiro quartil vai ser a média da décima e da décima primeira observações.

$$Q_1 = \frac{25 + 26}{2} = 25,5$$

O terceiro quartil é a mediana das 20 maiores observações; como 20 é par, o terceiro quartil vai ser a média da décima e da décima primeira observações da metade superior.

(c) $\bar{x} = \frac{7792}{150} = 51,9467$

(d) $n = 150$ – par

Mediana: deixa 75 observações abaixo e 75 acima; logo, a mediana é a média da observações de ordem 75 e 76.

$$Q_2 = \frac{50 + 50}{2} = 50$$

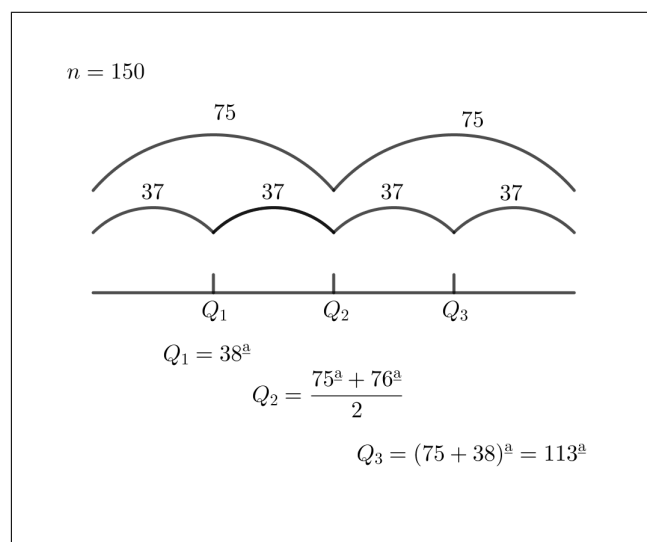
O primeiro quartil é a mediana das 75 menores observações; como 75 é ímpar, o primeiro quartil vai deixar 37 observações para baixo e 37 para cima, ou seja, o primeiro quartil vai ser a trigésima oitava observação.

$$Q_1 = 34$$

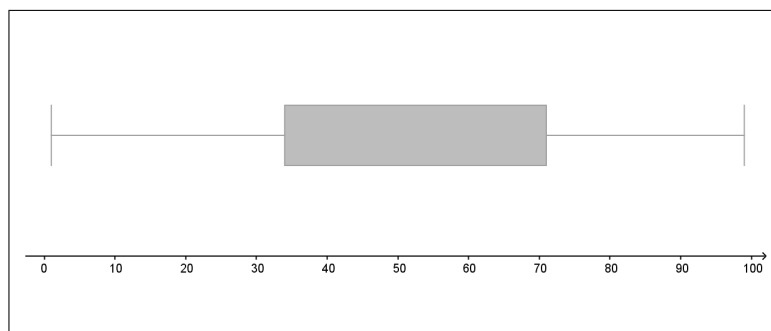
O terceiro quartil é a mediana das 75 maiores observações; como 75 é ímpar, o terceiro quartil vai ser a trigésima oitava observação da metade superior. Como na metade inferior já tem 75 observações, o terceiro quartil vai ser a observação de ordem $75+38=113$.

$$Q_3 = 71$$

Veja o esquema a seguir.



(e) $AIQ = 71 - 34 = 37$ $Q_1 - 1,5AIQ < 0$ $Q_3 + 1,5AIQ > 99$ \Rightarrow não há outliers



15. Para calcular o salário horário médio, temos que dividir o total dos vencimentos pelo total de horas trabalhadas pelos quatro amigos.

$$\begin{aligned}
 \bar{x} &= \frac{10 \times 3,50 + 12 \times 2,6 + 15 \times 3,80 + 8 \times 2,20}{10 + 12 + 15 + 8} \\
 &= \frac{10 \times 3,50 + 12 \times 2,6 + 15 \times 3,80 + 8 \times 2,20}{45} \\
 &= \frac{10}{45} \times 3,50 + \frac{12}{45} \times 2,6 + \frac{15}{45} \times 3,80 + \frac{8}{45} \times 2,20 \\
 &= \frac{140,8}{45} = 3,1289
 \end{aligned}$$

Note que o salário médio é uma média ponderada dos salários individuais, com o peso sendo definido pelo número de horas de trabalho.

16. A carga horária semanal total é $4 + 4 + 4 + 6 + 2 = 20$. Logo, o CR do aluno é

$$CR = \frac{4}{20} \times 7,5 + \frac{4}{20} \times 6,1 + \frac{4}{20} \times 8,3 + \frac{6}{20} \times 6,5 + \frac{2}{20} \times 7,5$$

$$= \frac{141,6}{20} = 7,08$$

17. O diagrama de ramos-e-folhas é o seguinte:

```

0 | 6 8 9
1 | 0 0 0 2 2 2 2 4 5 5 5 6 8 8 8
2 | 0
3 | 0
    
```

Os quartis são

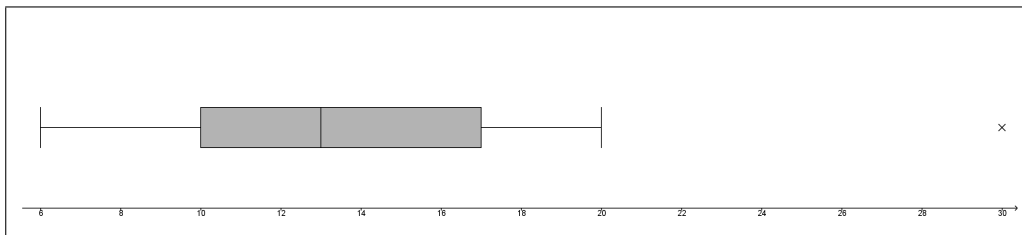
$$Q_2 = \frac{x_{(10)} + x_{(11)}}{2} = \frac{12 + 14}{2} = 13$$

$$Q_1 = \frac{x_{(5)} + x_{(6)}}{2} = \frac{10 + 10}{2} = 10$$

$$Q_3 = \frac{x_{(15)} + x_{(16)}}{2} = \frac{16 + 18}{2} = 17$$

Regra dos outliers:

$$AIQ = 17 - 10 = 7 \quad Q_1 - 1,5AIQ = -0,5 \quad Q_3 + 1,5AIQ = 27,5 \quad 30 \text{ é outlier}$$



A média, a variância e o desvio-padrão são

$$\bar{x} = \frac{6 + 8 + 9 + \dots + 20 + 30}{20} = \frac{280}{20} = 14$$

$$\sigma^2 = \frac{6^2 + 8^2 + 9^2 + \dots + 20^2 + 30^2}{20} - 14^2 = \frac{4456}{20} - 196 = 26,8$$

$$\sigma = \sqrt{(26,8)} = 5,1769$$

O escore padronizado de uma observação x é $z = \frac{x - \bar{x}}{\sigma}$. Na tabela a seguir, temos as observações e seus respectivos escores padronizados. Podemos ver que há apenas um escore padronizado maior que 3, correspondente à observação de valor 30, que é, assim, valor discrepante pelo Teorema de Chebyshev.

Observação	Escore padronizado	Observação	Escore padronizado
6	-1,5454	14	0
8	-1,1590	15	0,1932
9	-0,9659	15	0,1932
10	-0,7727	15	0,1932
10	-0,7727	16	0,3864
10	-0,7727	18	0,7727
12	-0,3864	18	0,7727
12	-0,3864	18	0,7727
12	-0,3864	20	1,1590
12	-0,3864	30	3,0907

18. Todos os salários ficaram aumentados em 250 reais. Se chamamos de x_i o salário do funcionário i no mês de novembro e de y_i o salário desse mesmo funcionário em dezembro, então

$$y_i = x_i + 250.$$

De acordo com a Propriedade 2, o salário médio em dezembro é

$$\bar{y} = \bar{x} + 250 = 920 + 250 = 1170 \text{ reais}$$

19. Seja x_i o salário do funcionário i no mês anterior ao dissídio. Depois do aumento, seu salário passa a ser

$$y_i = x_i + 0,089x_i = 1,089x_i$$

Logo, todos os salários ficam multiplicados por 1,089 e, pela Propriedade 3, a média também fica multiplicada por este valor, ou seja, depois do dissídio, o salário médio passa a ser

$$\bar{y} = 1,089\bar{x} = 1,089 \times 580 = 631,62 \text{ reais}$$

20. A diferença se deve à existência de grandes empresas no setor de bebidas, com muitos empregados. Como vimos, a média é bastante influenciada pelos valores discrepantes.

Capítulo 3 – Correlação

21. Na tabela a seguir temos as informações importantes para o cálculo do coeficiente de correlação.

País	X	Y	X ²	Y ²	XY
Islândia	240	63	57600	3969	15120
Noruega	255	100	65025	10000	25500
Suécia	340	140	115600	19600	47600
Dinamarca	375	140	140625	30625	65625
Canadá	510	160	260100	25600	81600
Austrália	490	180	240100	32400	88200
Holanda	490	250	240100	62500	122500
Suiça	180	180	32400	32400	32400
Finlândia	1125	360	1265625	129600	405000
Grã-Bretanha	1150	470	1322500	220900	540500
Estados Unidos	1275	200	1625625	40000	255000
Soma	6430	2278	5365300	607594	1679045

- Médias

$$\bar{X} = \frac{6430}{11} = 584,5455 \quad \bar{Y} = \frac{2278}{11} = 207,0909$$

- Variâncias

$$\sigma_X^2 = \frac{5365300}{11} - 584,5455^2 = 146061,157 \quad \sigma_Y^2 = \frac{607594}{11} - 207,0909^2 = 12349,1736$$

- Covariância

$$\text{Cov}(X, Y) = \frac{1679045}{11} - 584,5455 \times 207,0909 = 31586,40496$$

- Coeficiente de correlação

$$\rho(X, Y) = \frac{31586,40496}{\sqrt{146061,157 \times 12349,1736}} = 0,743727574$$

Há uma forte correlação linear positiva, ou seja, aumentando o consumo de cigarros, aumenta o número de mortes por câncer de pulmão.

22. Eis as informações relevantes:

- Médias

$$\bar{X} = \frac{14433}{59} = 244,627119 \quad \bar{Y} = \frac{10470}{59} = 177,457627$$

- Variâncias

$$\sigma_X^2 = \frac{3736397}{59} - 244,627119^2 = 3486,335536 \quad \sigma_Y^2 = \frac{1976794}{59} - 177,457627^2 = 2013,773628$$

- Covariância

$$\text{Cov}(X, Y) = \frac{2667063}{59} - 244,627119 \times 177,457627 = 1793,509624$$

- Coeficiente de correlação

$$\rho(X, Y) = \frac{1793,509624}{\sqrt{3486,335536 \times 2013,773628}} = 0,676883$$

Há uma forte correlação linear positiva, ou seja, aumentando a área da casa, aumenta o preço de venda.