

EST028 – Introdução à Estatística

Unidade 1

Lista de Exercícios

1. Exercícios sobre Variáveis, Tabelas e Gráficos

- 1- Para a notícia a seguir, identifique e classifique os tipos de variáveis utilizadas (Os dados constam da publicação: Dados Estatísticos de Turismo de 1998): “Este é o perfil, traçado pela Fundação Instituto de Pesquisas Econômicas, da USP, do turista brasileiro em 1998: o turista brasileiro tem entre 30 e 40 anos, renda média de R\$ 1800 e segundo grau completo. Viaja duas vezes por ano e de ônibus. Seu objetivo principal é visitar parentes e amigos. A maioria - 70% - viaja durante a alta estação. As viagens duram, em média, 72 dias.”
- 2- Identifique as variáveis abaixo como Qualitativas Ordinais (QO), Qualitativas Nominais (QN), Quantitativas Discretas (QD) ou Quantitativas Contínuas (QC):

[] A resposta de consumidores: preferem a marca A à marca B, se não têm preferência, ou se preferem a marca B à marca A

[] O número de consultas médicas feitas, por ano, por um associado de certo plano de saúde

[] A religião de pessoas que tentam o suicídio é codificada como 1, 2, 3, 4 ou 5, representando Protestante, Católico, Judeu, Outras ou Nenhuma

[] O teor de gordura, medido em gramas por 24 horas, nas fezes de crianças de 1 a 3 anos de idade.

[] O tipo de droga que os participantes de certo estudo tomaram, registrados como: Droga A, Droga B e placebo.

[] Opinião sobre determinado atendimento: Muito satisfeito, satisfeito, razoável, insatisfeito, muito insatisfeito.

- 3- Os transdutores de temperatura de um determinado tipo são enviados em lotes de 50. Uma amostra de 60 lotes foi selecionada e o número de transdutores fora das especificações em cada lote foi determinado, resultando nos dados a seguir:

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

- a. Determine a tabela de frequências dos valores observados de x (número de transdutores fora das especificações em um lote);
- b. Que proporção de lotes na amostra possui no máximo cinco transdutores fora das especificações? Que proporção tem menos de cinco? Que proporção possui no mínimo cinco unidades fora das especificações?
- c. Desenhe um gráfico de barras dos dados, usando a frequência relativa na escala vertical e comente suas características.
- 4- O artigo “Study on the Life Distribution of Microdrills” (*J. of Engr. Manufacture*, 2002:301-305) relatou as seguintes observações relativas à vida útil das brocas (número de furos que uma broca faz antes de quebrar), quando os furos são feitos em uma determinada liga de bronze.

11	14	20	23	31	36	39	44	47	50
59	61	65	67	68	71	74	76	78	79
81	84	85	89	91	93	96	99	101	104
105	105	112	118	123	136	139	141	148	158
161	168	184	206	248	263	289	322	388	513

- Por que uma distribuição de frequência não poder ter por base os intervalos de classe 0 → 50, 50 → 100, 100 → 150 e assim por diante?
 - Construa uma tabela de frequência e um histograma dos dados usando limites de classes 0, 50, 100, ... e então faça comentários sobre as características interessantes;
 - Que proporção das observações de vida útil dessa amostra é inferior a 100? Que proporção das observações é igual ou maior que 200?
- 5- Considere os dados a seguir sobre os tipos de queixas de saúde (J = inflamação de articulações, F = fadiga, B = dor nas costas, M = fadiga muscular, C = tosse, N = irritação nasal/coriza, O = outros) feitas por agricultores. Obtenha as frequências e as frequências relativas das diversas categorias, desenhe um gráfico de barras e um de pizza. (Os dados são consistentes com as porcentagens fornecidas no artigo “Physiological Effects of Work Stress and Pesticide Exposure in Tree Planting by British Columbia Silviculture Workers”, *Ergonomics*, 1993, p. 951-961.)

O	O	N	J	C	F	B	B	F	O	J	O	O	M
O	F	F	O	O	N	O	N	J	F	J	B	O	C
J	O	J	J	F	N	O	B	M	O	J	M	O	B
O	F	J	O	O	B	N	C	O	O	O	M	B	F
J	O	F	N										

- Quais são alguns dos benefícios obtidos quando se representa o conjunto de dados usando a distribuição de frequência?
- Cite algumas maneiras de representar graficamente dados quantitativos e qualitativos.
- Use um gráfico de dispersão para representar os dados da tabela, que são as horas trabalhadas e o salário por hora (em dólares) para uma amostra de 12 trabalhadores da produção. Descreva quaisquer tendências presentes.

Horas	Salário por hora	Horas	Salário por hora
33	13,16	40	13,65
37	9,98	33	12,05
34	10,79	28	10,54
40	11,71	45	10,33
35	11,80	37	11,57
33	11,51	28	10,17

- 9- Um nutricionista deseja comparar os efeitos de duas dietas, que ele identifica por dietas A e B. Um total de 100 pessoas é submetido à dieta A e outras 100 pessoas são submetidas à dieta B. O nutricionista percebeu que 46% das pessoas submetidas à dieta A tiveram redução de peso, e o mesmo aconteceu com 41% das pessoas submetidas à dieta B. Para uma apresentação em um evento científico, ele esboçou dois gráficos para esses dados, exibidos a seguir:

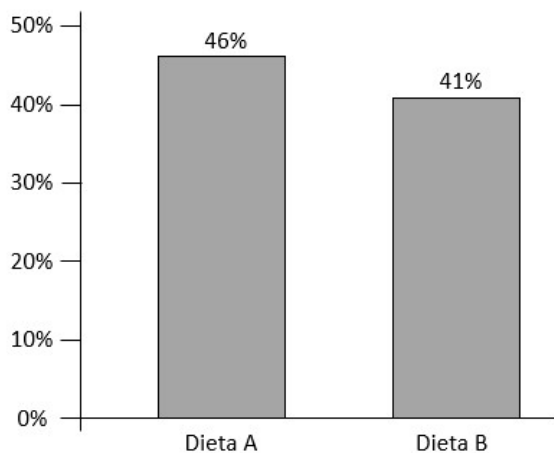


Gráfico A

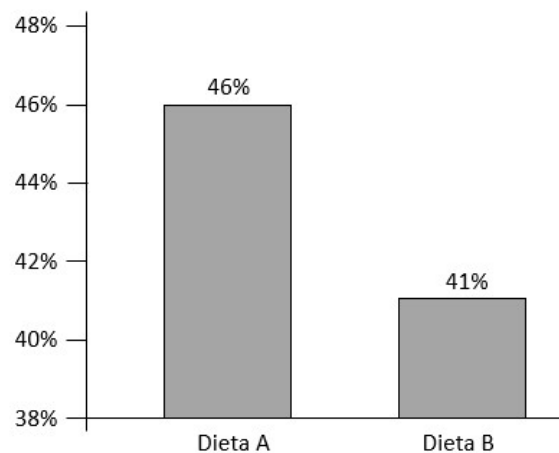
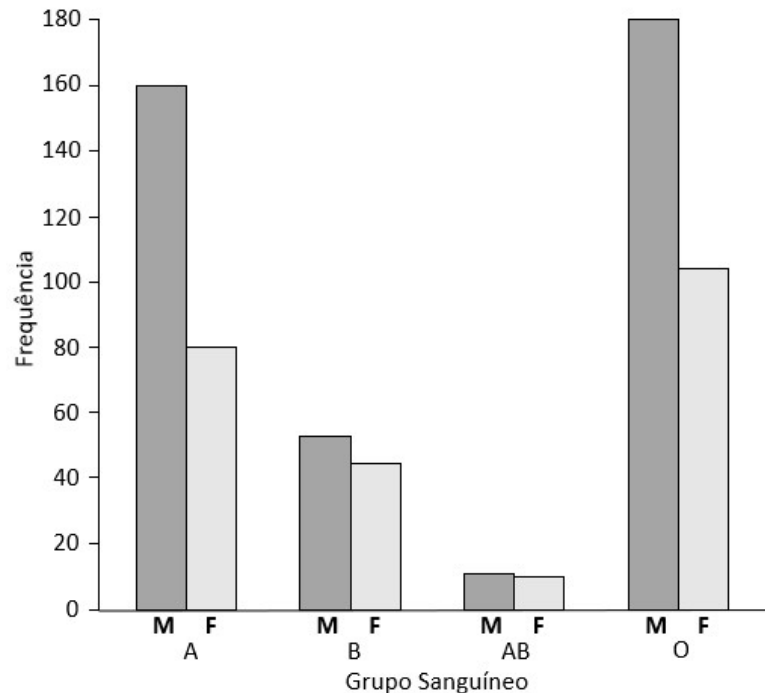


Gráfico B

Porém, somente um desses gráficos é adequado. Identifique o gráfico adequado (A ou B), e explique o erro cometido no gráfico que é inadequado.

10- Um estudante está conduzindo um estudo que objetiva descrever os doadores de sangue de um serviço de hemoterapia. Ele utiliza uma amostra de 651 doadores, sendo 410 de sexo masculino e 241 de sexo feminino. Dentre os doadores de sexo masculino, ele encontrou 161 portadores de sangue tipo A, 56 portadores de sangue tipo B, 14 portadores de sangue tipo AB e 179 portadores de sangue tipo O. Dentre os doadores de sexo feminino, ele encontrou 81 portadores de sangue tipo A, 44 portadores de sangue tipo B, 13 portadores de sangue tipo AB e 103 portadores de sangue tipo O. Para descrever esses dados, o estudante construiu o gráfico de barras abaixo. Neste gráfico, os homens são identificados por M, de masculino, e as mulheres por F, de feminino.

Seu professor, no entanto, lhe disse que esse gráfico contém um erro bastante sério.



- Comente o que está errado nesse gráfico;
- Esboce novamente um gráfico de barras para esses dados, de modo que esse erro não seja novamente cometido.

11- Construa um gráfico de dispersão utilizando os dados abaixo, sendo que os dados de alcatrão em cigarros tamanho *king* devem estar na escala horizontal e os dados de monóxido de carbono (CO) nos mesmos cigarros devem estar na escala vertical. Determine se parece haver uma relação entre o alcatrão e o CO em cigarros tamanho *king*. Em caso afirmativo, descreva a relação.

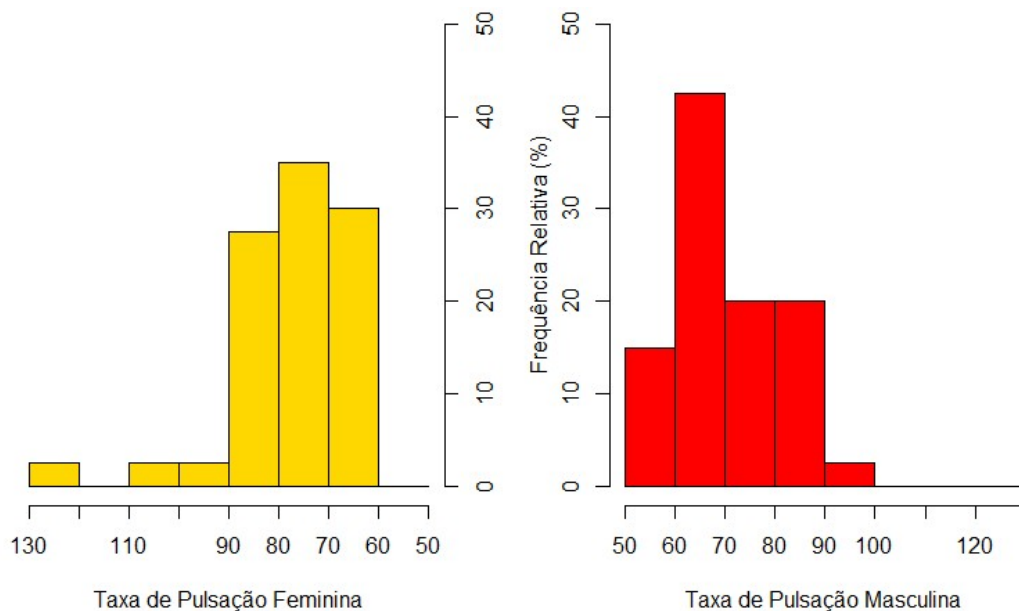
Marca	Tamanho King		
	Alcatrão	Nicotina	CO
Austin	20	1,1	16
Basic	27	1,7	16
Bristol	27	1,7	16
Cardinal	20	1,1	16
Cavalier	20	1,1	16
Chesterfield	24	1,4	17
Cimarron	20	1,1	16
Class /a	23	1,4	15
Doral	20	1,0	16
GPC	22	1,2	14
Higway	20	1,1	16

Jacks	20	1,1	16
Marker	20	1,1	16
Monaco	20	1,1	16
Monarck	20	1,1	16
Old Gold	10	1,8	14
Pall Mall	24	1,6	16
Pilot	20	1,1	16
Prime	21	1,2	14
Pyramid	25	1,5	18
Raleigh	23	1,3	15
Extra			
Sebring	20	1,1	16
Summit	22	1,3	14
Sundance	20	1,1	16
Worth	20	1,1	16

12- Declare se cada uma das seguintes observações é um exemplo de dados discretos ou contínuos:

- O número de suicídios nos Estados Unidos em um ano específico;
- A concentração de chumbo em uma amostra de água;
- A duração de tempo que um paciente de câncer sobrevive depois do diagnóstico;
- O número de abortos prévios que uma mãe grávida teve.

13- Ao usarmos histogramas para a comparação de dois conjuntos de dados, às vezes isso se torna difícil por termos que olhar um e outro histograma. Um *histograma de frequência relativa lado a lado* usa um formato que torna essa comparação muito mais fácil. Em lugar de frequências absolutas, devemos usar frequências relativas (porcentagens e proporções) para que as comparações não sejam distorcidas por tamanhos amostrais diferentes. Observe o histograma de frequências relativas lado a lado, mostrado a seguir e compare os dois conjuntos de dados.



2. Exercícios sobre Medidas de Tendência Central

- A propagação de trincas por fadiga em diversas peças de aeronaves tem sido objeto de muitos estudos nos últimos anos. Os dados a seguir consistem dos tempos de propagação (horas de voo/ 10^4) para atingir um determinado tamanho de trinca em furos de fixadores propostos para uso em aeronaves militares (“Statistical Crack Propagation in Fastener Holes under Spectrum Loading”, *J. Aircraft*, 1983, p. 1028-1032):

0,736	0,863	0,865	0,913	0,915	0,937	0,983	1,007
1,011	1,064	1,109	1,132	1,140	1,153	1,256	1,394

- Calcule e compare os valores da média e da mediana amostrais;
- Em quanto a maior observação da amostra pode ser diminuída sem afetar o valor da mediana?

Nos exercícios 2 e 3, determine se a afirmativa é verdadeira ou falsa. Se for falsa, reescreva-a em sua forma verdadeira

- A mediana é a medida de tendência central mais provável de ser afetada por um valor atípico.
- Alguns conjuntos de dados quantitativos não têm uma mediana.

Nos exercícios 4 e 5 faça o seguinte:

- Determine a média, a mediana e a moda dos dados, se possível. Se não for possível, explique o porquê a medida de tendência central não pôde ser determinada;
 - Qual é a medida de tendência central que melhor representa dos dados? Explique seu raciocínio.
- O tempo (em segundos) que uma amostra de sete carros esportivos leva para ir de zero a 60 milhas por hora: 3,7; 4,0; 4,8; 4,8; 4,8; 4,8; 5,1. (*Fonte: Motor Trend*)
 - Os mais velozes (em milhas por hora) em uma amostra de sete carros esportivos: 187,3; 181,8; 180,0; 169,3; 162,2; 158,1; 155,7. (*Fonte: Motor Trend*)
 - Em que sentido a média, a mediana e a moda são medidas de tendência central?
 - Em um editorial, o *Poughkeepsie Journal* colocou esta afirmativa: “O preço mediano – o preço exatamente no meio entre o mais alto e o mais baixo - ...” A afirmativa descreve corretamente a mediana? Por que sim, ou por que não?
 - Obteve-se uma amostra aleatória simples de páginas do *Merriam-Webster’s Collegiate Dictionary*, 11ª Edição, com um total de 1460 páginas. A seguir, estão listados os números de palavras definidas em cada uma das 10 páginas amostradas, estime o número total de palavras definidas no dicionário, utilizando a média e a mediana. É provável que essa seja uma estimativa precisa do número de palavras na língua inglesa?

51	63	36	43	34	62	73	39	53	79
----	----	----	----	----	----	----	----	----	----

- A média do peso dos 100 alunos de uma determinada turma de alunos é 68,4 kg. Essa turma possui 25 homens e 75 mulheres. O peso médio das mulheres dessa turma é 62,6 kg. Qual é o peso médio dos homens dessa turma?

3. Exercícios sobre Medidas de Dispersão

- O artigo “A Thin-Film Oxygen Uptake Test for the Evaluation of Automotive Crankcase Lubricants” (*Lubric. Engr.*, 1984, p. 75-83) informou os seguintes dados sobre o tempo de oxidação-indução (min) para diversos óleos comerciais:

87	103	130	160	180	195	132	145	211	105
145	153	152	138	87	99	93	119	129	

- Calcule a variância e o desvio padrão amostrais;
 - Se as observações fossem especificadas em horas, quais seriam os valores resultantes para a variância e para o desvio padrão amostrais? Responda sem reescrever os valores.
- Considere a e b como constantes e $y_i = ax_i + b$ para $i = 1, 2, \dots, n$. Quais são as relações entre \bar{x} e \bar{y} e entre s_x^2 e s_y^2 ?
 - Uma amostra de temperaturas iniciais de uma determinada reação química resultou em uma média amostral ($^{\circ}\text{C}$) de 87,3 e um desvio padrão amostral de 1,04. Quais são a média e o desvio padrão em $^{\circ}\text{F}$? (DICA: $F = \frac{9}{5}C + 32$),

- 4- Qual você considera ter maior variação: as rendas de uma amostra aleatória simples de 1.000 adultos selecionados da população geral, ou as rendas de uma amostra aleatória simples de 1.000 professores de estatística? Por que?
- 5- Calcule a variância e o desvio padrão dos dados apresentados no exercício 8 da seção anterior. Como esses valores afetam a confiança na precisão da estimativa da média encontrada anteriormente?
- 6- A seguir, estão listados os custos (em dólares) de viagens de ida e volta do aeroporto JFK na cidade de Nova York a São Francisco. Todos os voos envolvem uma escala e uma estada de duas semanas. As companhias aéreas são US Air, Continental, delta, United, American, Alaska e Northwest:

30 Dias Antes	244	260	264	264	278	318	280
1 Dia Antes	456	614	567	943	628	1088	536

Calcule os coeficientes de variação para cada um dos conjuntos de dados e os compare quanto à sua homogeneidade.

4. Exercícios sobre Quantis e Box Plots

- 1- Uma amostra de 20 garrafas de certo tipo de vidro foi selecionada e a resistência à pressão interna de cada garrafa foi determinada. Considere as seguintes informações parciais da amostra:

$$md = 202,2 \quad Q_1 = 196,0 \quad Q_3 = 216,8$$

Três observações menores: 125,8; 188,1; 193,7

Três observações maiores: 221,3; 230,5; 250,2

- a. Há algum outlier na amostra?
- b. Construa um box plot e comente suas características.
- 2- A concentração de cocaína no sangue (mg/L) foi determinada para uma amostra de indivíduos que morreram de delírio induzido por cocaína (ED) e para uma amostra de indivíduos que morreram de overdose de cocaína sem delírio. O tempo de sobrevivência das pessoas em ambos os grupos foi de, no máximo 6 horas. Os dados a seguir foram obtidos de um box plot comparativo do artigo “Fatal Excited Delirium Following Cocaine Use” (*J. of Forensic Sciences*, 1997, p.25-31):

ED

0	0	0	0	0,1	0,1	0,1	0,1	0,2
0,2	0,3	0,3	0,3	0,4	0,5	0,7	0,8	1,0
1,5	2,7	2,8	3,5	4,0	8,9	9,2	11,7	21,0

Não-ED

0	0	0	0	0	0,1	0,1	0,1	0,1	0,2
0,2	0,2	0,3	0,3	0,3	0,4	0,5	0,5	0,6	0,8
0,9	1,0	1,2	1,4	1,5	1,7	2,0	3,2	3,5	4,1
4,3	4,8	5,0	5,6	5,9	6,0	6,4	7,9	8,3	8,7
9,1	9,6	9,9	11,0	11,5	12,2	12,7	14,0	16,6	17,8

- a. Determine o resumo de cinco números das duas amostras;
- b. Há algum outlier nas amostras, se sim, qual/quais?
- c. Construa um box plot comparativo e use-o como base para comparar e destacar as diferenças das amostras ED e não-ED.
- 3- Obteve-se uma amostra aleatória simples de escores de classificação de crédito FICO, e esses escores ordenados estão listados a seguir. Encontre o resumo de cinco números, verifique a existência, ou não, de valores atípicos e faça o box plot.

664	693	698	714	751	753
779	789	802	818	834	836

- 4- Use os 40 comprimentos de coxas (cm) de mulheres listados abaixo. Encontre o resumo de cinco números, verifique a existência, ou não, de valores atípicos e faça o box plot.

27	31,1	32,1	33,2	33,8	36	36,2	36,6	37,3	37,5
38	38,1	38,2	38,2	38,2	38,5	38,7	39	39	39
39,1	39,2	39,4	39,7	39,9	40	40,2	40,2	40,3	41
41	41,5	41,6	41,6	42,3	42,8	43,2	43,4	43,8	48,6

- 5- Um estudante do autor pesou uma amostra aleatória simples de bifes Porter-house, e os resultados (em onças) estão listados a seguir. Os bifes devem ter 21 oz, pois estão listados no menu com peso de 20 onças, e eles perdem uma onça ao serem cozidos. Use a lista de pesos para encontrar: (a) a média; (b) a mediana; (c) a moda; (d) a amplitude; (e) o desvio padrão; (f) a variância; (g) o primeiro quartil; (h) o terceiro quartil; (i) faça o box plot.

17	19	21	18	20	18	19	20	20	21
----	----	----	----	----	----	----	----	----	----

5. Exercícios de Avaliações Anteriores

- 1- Foi feita uma pesquisa com os 160 empregados de uma empresa, querendo saber quantos filhos eles tinham. Complete a tabela de frequência dada abaixo:

n° de filhos	Freq. Abs.	Freq. Rel. (%)	Freq. Rel. Ac. (%)
0	32		
1			45
2	56		
3		15	
4	0		
5		5	
Total			

- 2- Abaixo são dados os tempos de espera (em minutos) de clientes no Banco New York (onde todos os clientes formam uma fila única) e no Banco Montreal (onde os clientes entram em três filas de guichês diferentes):

New York		Montreal	
Tempo de espera (min)	Freq. Abs.	Tempo de espera (min)	Freq. Abs.
0	0	0	1
1	2	1	0
2	0	2	2
3	2	3	1
4	1	4	3
5	1	5	3
6	6	6	3
7	9	7	5
8	3	8	3
9	0	9	2
10	0	10	1

Para cada um dos dois bancos acima:

- Calcule a média e o coeficiente de variação sabendo que a variância dos tempos de espera para os Bancos de New York e Montreal são 4,06 e 6,17, respectivamente;
- Faça os box plots, lado a lado;

- c. Comente sobre as diferenças ou semelhanças entre as distribuições, utilizando as informações obtidas nos itens a) e b).
- 3- A tabela de frequências abaixo mostra os gastos dos clientes de um supermercado registrados em um caixa expresso durante uma manhã:

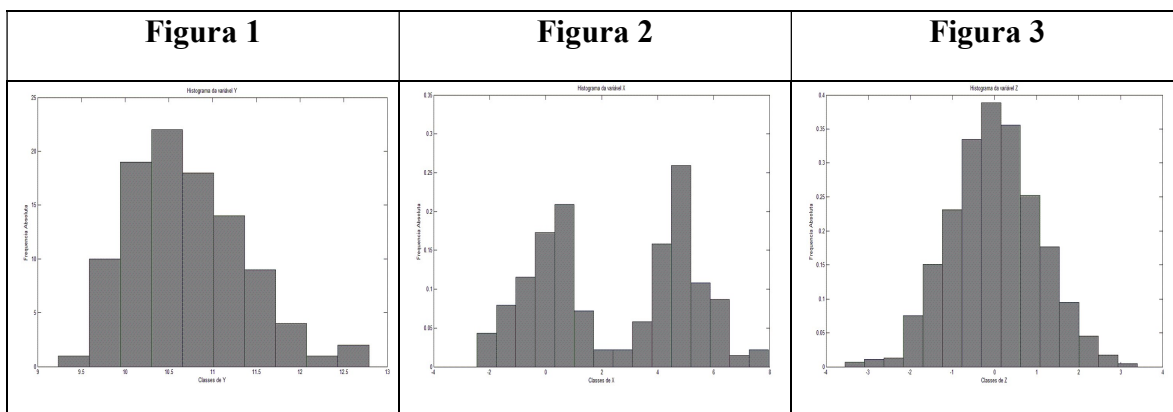
Gastos em reais	Número de Clientes
5 – 10	15
10 – 15	31
15 – 20	20
20 – 25	25
25 – 30	10
30 – 35	19

- a. Que porcentagem do total de clientes gastou pelo menos 20 reais?
- b. Que porcentagem do total de clientes gastou menos de 15 reais?
- 4- Para comparar os gastos familiares anuais com saúde, uma amostra aleatória de 100 pessoas, em cada uma de três cidades, foi selecionada e pesquisada. Os resultados são resumidos na tabela a seguir.

Estatísticas do gasto familiar anual com saúde para amostras de três cidades, em \$:

Estatística	Cidade A	Cidade B	Cidade C
Média	1800	1810	1820
Mediana	1800	1800	1800
Mínimo	1600	1600	1350
Máximo	2000	2200	2300
1° quartil	1700	1700	1700
3° quartil	1900	1900	1955
Desvio padrão	120	150	200

- a. Desenhe boxplos paralelos da distribuição de gastos familiares com saúde nas três cidades.
- b. Com base no gráfico obtido em (a), compare a variabilidade da distribuição neste contexto.
- c. Calcule o coeficiente de variação do gasto familiar com saúde para cada cidade.
- d. Com base em (c) e no desvio padrão, compare a variabilidade do gasto familiar com saúde nas três cidades.
- 5- Para cada figura abaixo, informe se a média amostral está próxima, à esquerda ou à direita da mediana amostral:



- 6- A dona de um restaurante registrou durante 6 meses quantos clientes ela recebia a cada semana. Os dados são apresentados abaixo já organizados do menor para o maior:

501 512 516 525 528 536 546 556 564 567 589 597 601
 603 605 612 615 624 629 635 642 645 648 651

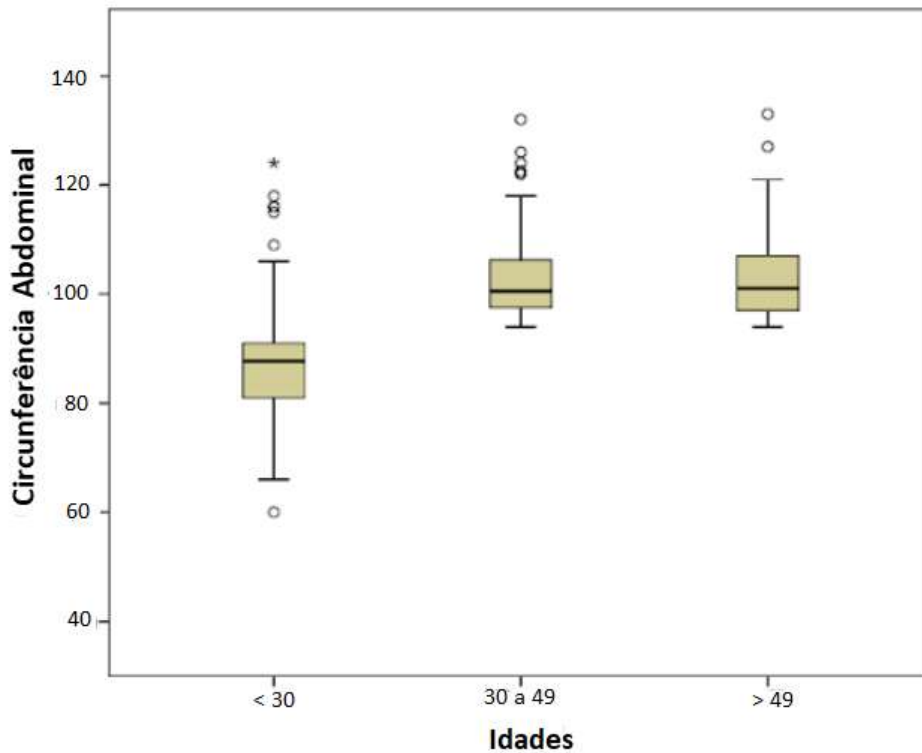
- Construa uma tabela de frequências contendo exatamente 6 classes e incluindo a frequência acumulada;
 - Construa o histograma da variável observada;
 - Sabendo que a média observada foi de 585,29 clientes, a distribuição observada pode ser considerada simétrica? Justifique sua resposta.
- 7- Abaixo são dadas as notas obtidas em um exame de estatística e em um exame de matemática para cada um dos 12 alunos de uma turma do ensino médio.

Aluno	Notas	
	Estatística	Matemática
1	50	45
2	50	54
3	53	53
4	58	61
5	59	55
6	60	66
7	62	59
8	64	70
9	65	57
10	66	65
11	68	72
12	80	90

- Calcule a média, mediana e moda para as notas das avaliações de matemática e de Estatística;
 - Você diria que a distribuição das notas (para cada uma das matérias) é simétrica? Justifique;
 - Sabendo que as variâncias observadas foram de 70,93 para as notas em estatística e 135,48 para as notas em matemática, calcule o coeficiente de variação para ambas as variáveis e responda: Quais notas apresentaram maior variabilidade? Justifique.
 - Encontre o resumo de cinco números para ambas as variáveis (notas de matemática e notas de estatística).
 - Calcule as distâncias interquartis e as regiões de observação típicas para ambas as variáveis. Existem valores atípicos para as variáveis notas em matemática e estatística, se sim, quais?
 - Desenhe um gráfico de dispersão bidimensional relacionando as variáveis notas de estatística, com notas de matemática, levando em consideração que a estatística depende da matemática.
- 8- Um personal trainer coletou dados sobre a quantidade de água que ele tomava por dia a partir de uma amostra de dias selecionada aleatoriamente. A partir disso, ele descobriu que o número médio de copos de água que ele bebia em um dia era igual a 4,4. Considerando a tabela abaixo, responda:

Quantidade de copos de água por dia	1	2	3	4	5	6	7
Frequência Absoluta	2	4	6	9	x	9	4

- Encontre o número de dias representado por x ;
 - Calcule o número mediano de copos de água que ele bebe por dia;
 - Calcule os quartis da quantidade de copos de água que ele bebe por dia;
 - Desenhe o Box Plot do número de copos de água bebido por dia.
- 9- Em 2018 foi realizada uma pesquisa para avaliar o estado nutricional dos funcionários técnico-administrativos da UFJF. Uma das variáveis coletadas foi a medida da circunferência abdominal em centímetros. Na figura a seguir temos a representação desta variável para 3 faixas etárias dos funcionários do sexo masculino. Assinale a afirmativa INCORRETA.



- (a) O grupo que apresenta menor variabilidade dos valores é o de idade inferior a 30 anos.
- (b) As distribuições dos grupos com idade entre 30 e 49 anos e idade superior a 49 anos são assimétricas.
- (c) A mediana do grupo com idade entre 30 e 49 anos é 100 centímetros
- (d) Todos os grupos apresentam dados discrepantes
- (e) O gráfico acima não apresenta a média da circunferência abdominal de cada faixa de idades

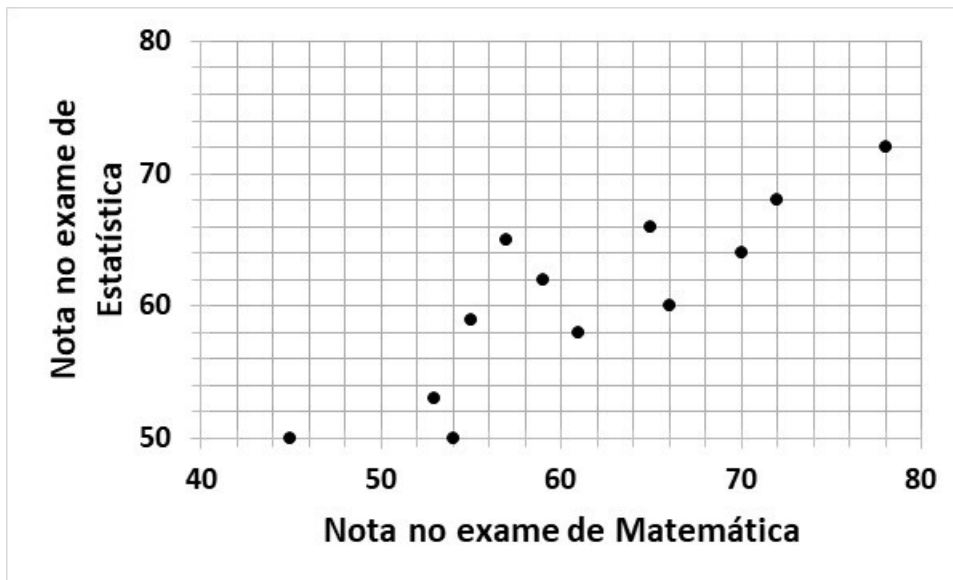
10- Leia atentamente as afirmações abaixo:

- I. A média amostral pode ser igual a zero.
- II. A variância amostral pode ser negativa.
- III. A mediana de uma amostra pode ser igual ao valor mínimo dessa amostra.
- IV. A média amostral pode ser superior à mediana.
- V. O desvio padrão pode ser nulo.

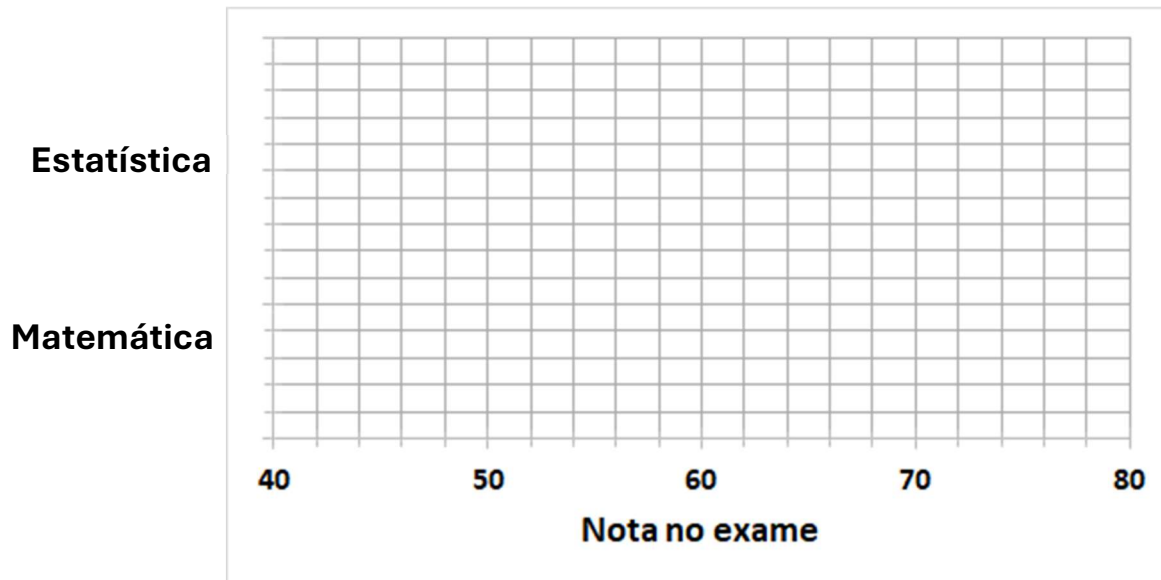
Marque a alternativa correta:

- (a) Todas as afirmações são verdadeiras.
- (b) Apenas as afirmações I, III, IV e V são verdadeiras.
- (c) Apenas as afirmações IV e V são verdadeiras.
- (d) Todas as afirmações são falsas.
- (e) Nenhuma das alternativas acima está correta.

11- O gráfico de dispersão abaixo fornece informações sobre as notas obtidas em um exame de estatística e em um exame de matemática para cada um dos 13 alunos de uma turma do ensino médio



a. No quadro abaixo, desenhe um box plot para as notas de cada exame.



- b. Calcule a média das notas de cada exame e, usando o símbolo “ \diamond ”, adicione esta informação ao gráfico solicitado em (a).
- c. Compare as duas distribuições. Comente as semelhanças e as suas diferenças.

12- Complete as duas afirmativas abaixo:

- I. Para que uma aluna seja aprovada numa disciplina, é preciso que o número resultante da soma das suas notas em três provas parciais, dividido por 3, seja maior que 70. Este número pode ser considerado como uma medida de _____.
- II. A diferença entre a nota máxima e a nota mínima obtidas numa prova pelos alunos de uma turma pode ser considerada como uma medida de _____.
- III. A nota na avaliação que mais se repete dentre as notas obtidas pelos alunos de uma turma pode ser considerado como uma medida de _____.

As respostas são, respectivamente:

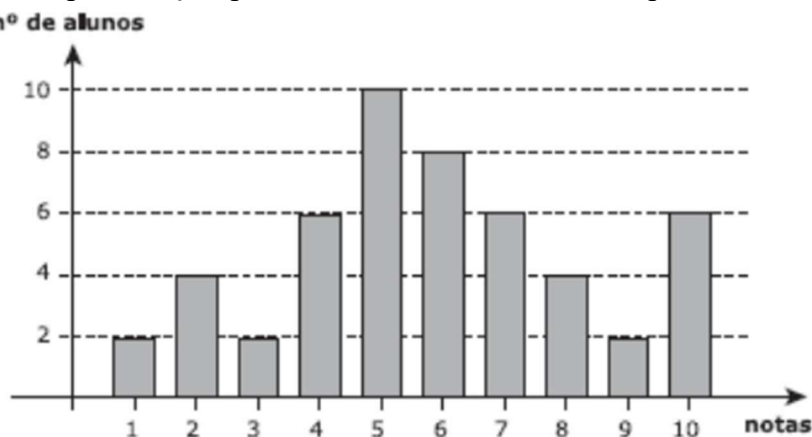
- (a) Dispersão, tendência central e dispersão.
- (b) Tendência central, tendência central e tendência central.
- (c) Tendência central, dispersão e tendência central.
- (d) Tendência central, dispersão e dispersão.

(e) Dispersão, dispersão e tendência central.

13- Observe no gráfico a seguir uma representação para as notas de 50 alunos do primeiro semestre de Ciências Econômicas numa determinada prova:

A mediana das notas dos 50 alunos de ciências Econômicas prova é igual a:

- (a) 3
- (b) 4
- (c) 5
- (d) 6
- (e) 7



nesta

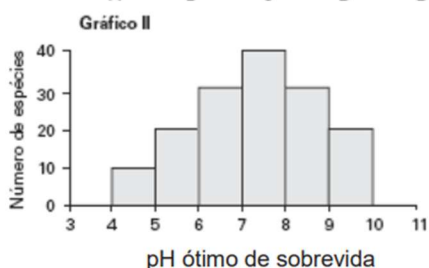
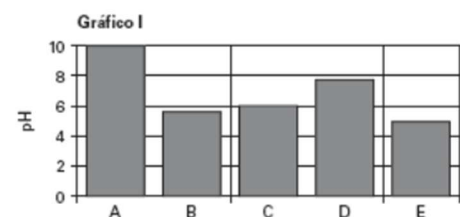
14- Marcos está de férias na casa do lago de seu amigo. Ele está curioso sobre a atividade aquática favorita das outras pessoas no lago. Então, ele entrevista 150 pessoas no lago e pergunta qual é sua atividade aquática favorita. A tabela abaixo resume os dados:

Atividade	Freq. Rel.
Passeio de barco	0,24
Pescaria	0,20
Esqui aquático/Wakeboard	0,20
Caiaque	0,16
Stand up paddle	0,12
Natação	0,08

Qual das seguintes afirmações é sustentada pela tabela?

- (a) 20 pessoas responderam pescaria como a sua atividade favorita.
- (b) 36 pessoas responderam passear de barco como a sua atividade favorita.
- (c) Esqui aquático/Wakeboard foi escolhida por mais pessoas do que Stand Up Paddle e natação juntas.
- (d) Passear de barco e caiaque somam metade das respostas.
- (e) Um quarto dos respondentes escolheram passear de barco como atividade favorita.

15- Um estudo caracterizou 5 ambientes aquáticos, nomeados de A a E, em uma região, medindo parâmetros físico-químicos de cada um deles, incluindo o pH nos ambientes:



O Gráfico I representa os valores de pH dos 5 ambientes. Utilizando o gráfico II, que representa a distribuição estatística de espécies em diferentes faixas de pH, pode-se esperar um maior número de espécies no ambiente:

- (a) A
- (b) B
- (c) C
- (d) D
- (e) E

16- Quais das seguintes afirmações são verdadeiras?

- I. Se o lado direito e o lado esquerdo de um histograma são a imagem espelhada um do outro, a distribuição é simétrica.

- II. Uma distribuição que é bastante dispersa ao longo do seu lado direito é dita assimétrica à direita.
- III. Se uma distribuição é assimétrica à direita, sua média é, usualmente, maior do que a sua mediana.

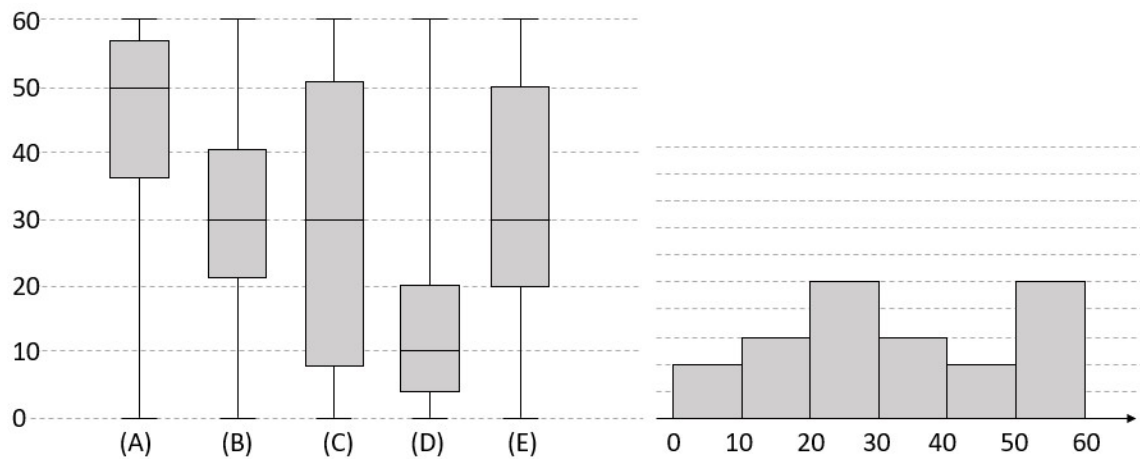
Escolha a alternativa com a resposta correta:

- (a) Apenas I
- (b) I e II
- (c) I e III
- (d) II e III
- (e) Nenhuma das alternativas acima.

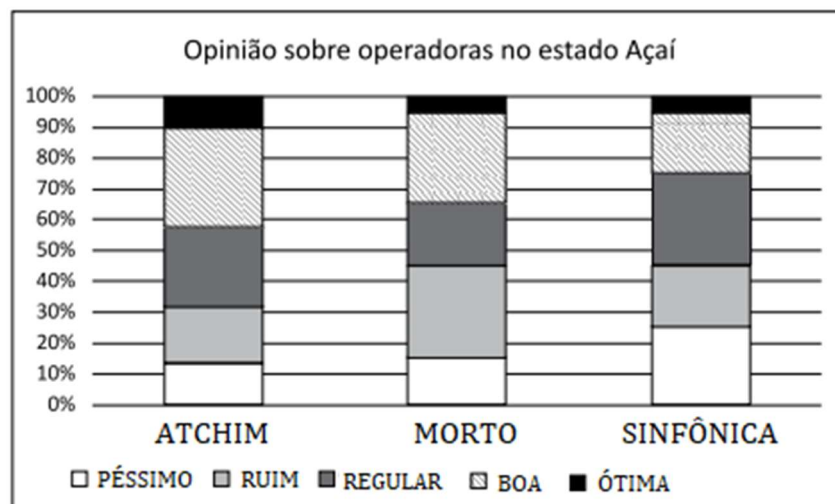
17- Em um conjunto de 100 observações numéricas, podemos afirmar que:

- (a) a média aritmética é maior que a mediana.
- (b) a mediana é maior que a moda.
- (c) 50% dos valores estão acima da média aritmética.
- (d) 50% dos valores estão abaixo da mediana.
- (e) 25% dos valores estão entre a moda e a mediana.

18- Relacione o histograma abaixo com o box plot correspondente.



19- O gráfico abaixo apresenta os resultados de uma pesquisa de opinião com os clientes de operadoras de telefonia móvel no estado Açaí.

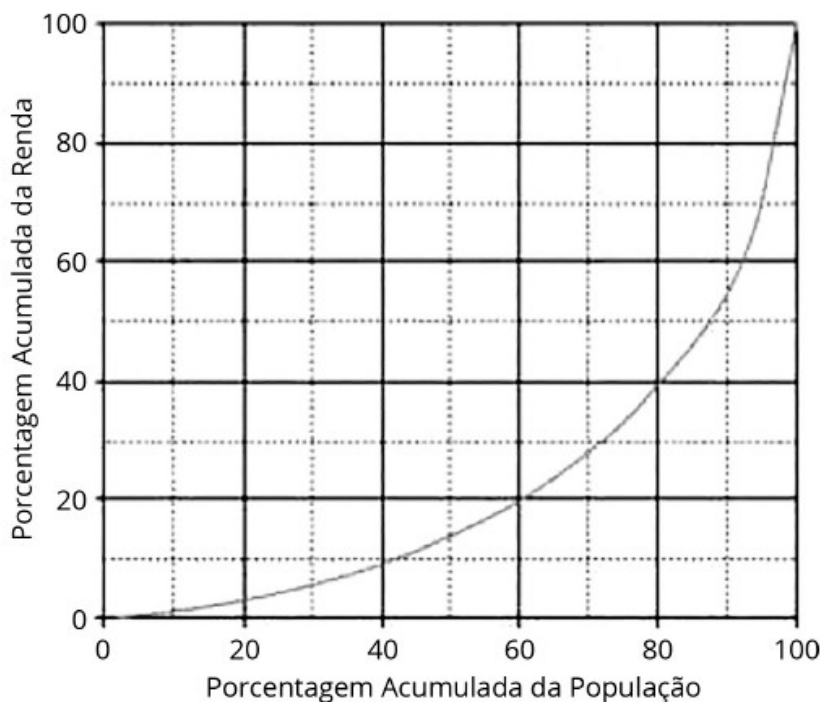


É correto afirmar sobre o gráfico que

- a) a operadora ATCHIM é a mais bem avaliada entre as três operadoras, pois a soma dos percentuais de “Ótima” e “Boa” é a maior entre as três operadoras.
- b) as opiniões dos clientes sobre as operadoras são semelhantes no estado Açaí, porque os percentuais são aproximadamente iguais.

- c) as opiniões dos clientes sobre as operadoras são positivas no estado Açaí, porque a soma dos percentuais de “Ótima” e “Boa” é superior a 50% para todas.
- d) a opinião dos clientes da operadora MORTO no estado Açaí é bastante negativa, porque a soma dos percentuais de “Ruim” e “Péssima” é cerca de 60%.
- e) o maior percentual de respostas dos clientes está concentrado na categoria “Boa” para todas as operadoras.

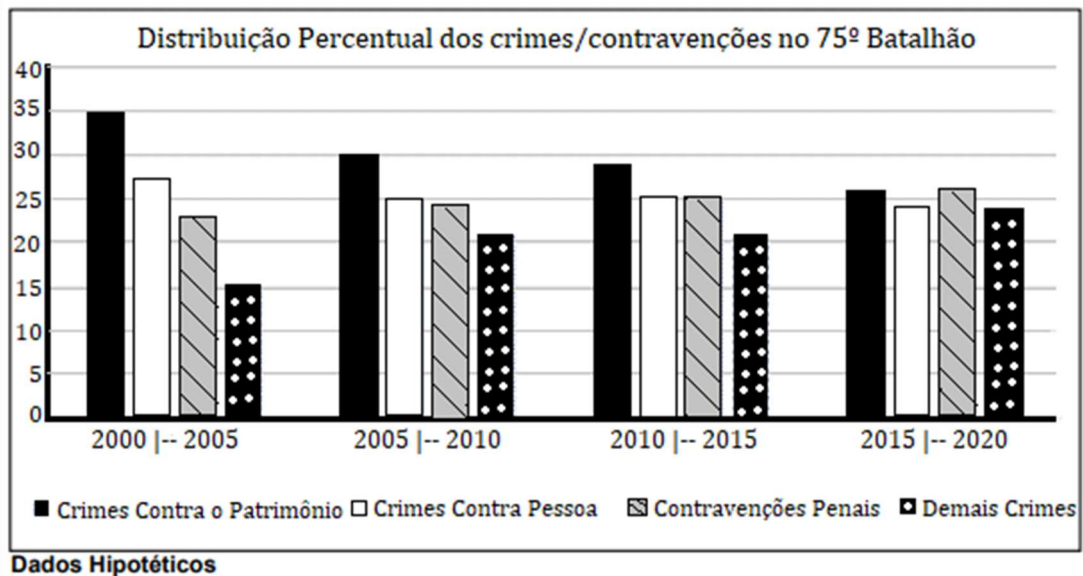
20- Uma forma de observar a desigualdade de um país é por meio da Curva de Lorenz, que fornece, para cada valor de x entre 0 e 100, o percentual da renda total do país auferido pelos $x\%$ de brasileiros de menor renda. Por exemplo, na Curva de Lorenz abaixo, constata-se que a renda total dos 60% de menor renda representou apenas 20% da renda total. De acordo com o mesmo gráfico, o percentual da renda total correspondente aos 20% de maior renda foi, aproximadamente, igual a



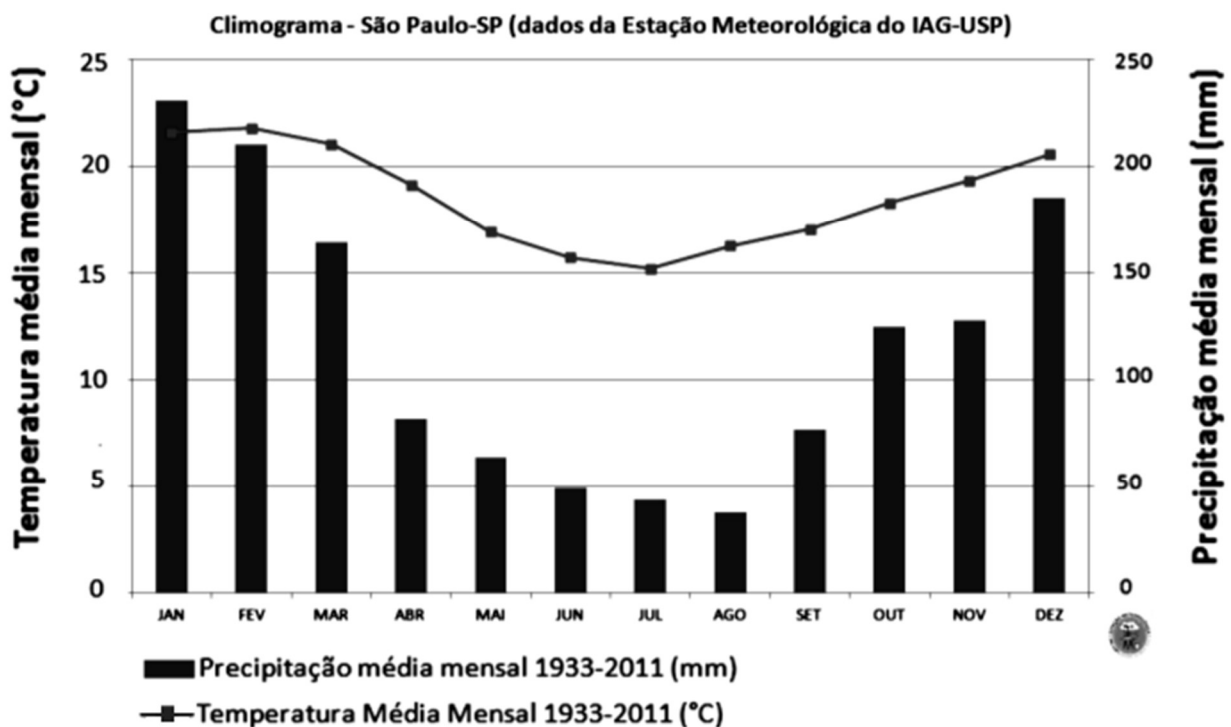
- a) 20%;
- b) 40%;
- c) 50%;
- d) 60%;
- e) 80%.

21- De acordo com os dados sobre a distribuição percentual dos crimes e contravenções no 75º Batalhão de Polícia Militar, apresentados no gráfico a seguir, marque a alternativa CORRETA:

- a) De 2000 a 2019 os crimes contra o patrimônio representaram em média 25% das infrações totais.
- b) Nos primeiros 10 anos do estudo as contravenções penais apresentaram decréscimo.
- c) Ao passo que os crimes contra o patrimônio e os demais crimes estão reduzindo, as contravenções penais estão aumentando.
- d) No período do estudo os crimes contra a pessoa estão reduzindo e as contravenções penais estão aumentando.
- e) A quantidade de crimes/contravenções está diminuindo com o passar dos anos.



22- A figura abaixo foi obtida a partir da série histórica da Estação Meteorológica do IAG-USP com os dados de temperatura e pluviosidade (volume de chuvas) obtidos para São Paulo entre 1933 e 2011.



A amplitude térmica ao longo dos dias nesta cidade pode ser muito grande, acima de 10 graus Celsius, o que não impede de haver um comportamento médio mensal relativamente estável e representativo de um padrão. Nesse contexto, analise as afirmativas abaixo.

- I. Podemos dizer que os meses mais secos são os meses de junho a agosto, que também é o período mais frio.
- II. É possível descrever o clima da cidade de São Paulo como sendo caracterizado por duas estações características: uma seca e outra úmida.
- III. Na cidade de São Paulo a amplitude anual (média) de temperatura é menor do que a amplitude diária de temperatura.
- IV. O gráfico mostra que mesmo no frio não existe registro de temperaturas abaixo de 15°C.

Quanto às conclusões lógicas verdadeiras inferidas a partir das informações acima, assinale a alternativa correta.

- (a) As afirmativas I, II, III e IV estão corretas.
- (b) Apenas afirmativas I, II e III estão corretas.
- (c) Apenas as afirmativas I e II estão corretas.
- (d) Apenas as afirmativas IV e III estão corretas.
- (e) Apenas as afirmativas I e III estão corretas

6. Avaliação 2025

1. Uma universidade decidiu estudar se características diferentes de seus alunos poderiam afetar suas notas finais; para tal coletou uma amostra de 1000 alunos. As características observadas foram: idade (em anos completos), gênero (masculino, feminino, outro), tempo de estudo diário, tempo gasto em redes sociais, tempo gasto na Netflix, tempo de sono (todos esses tempos medidos em horas, podendo ser completas ou não), trabalha em meio período (sim, não), frequência (percentual), qualidade da dieta (boa, razoável, pobre), frequência de exercícios na semana (de 0 a 7 dias), escolaridade dos pais (nenhuma, ensino médio, superior, mestre), qualidade da internet (ruim, média, boa), índice de saúde mental (de 1 a 10), participação em atividades extracurriculares (sim, não), nota final (pontuação de 0 a 100).

- a. Classifique as variáveis:

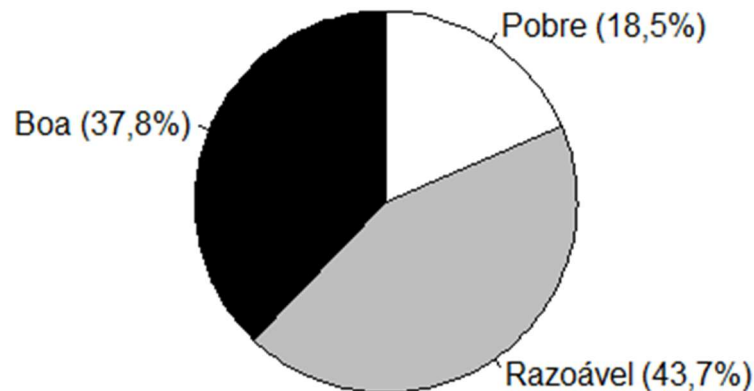
Idade: _____

Gênero: _____

Tempo de estudo diário: _____

Qualidade da Internet: _____

- b. Construa a tabela de frequências para a variável qualidade da dieta. As informações necessárias podem ser obtidas no gráfico abaixo:



Classes	Freq. absoluta	Freq. Rel. (%)	Freq. Rel. acumulada (%)
---------	----------------	----------------	--------------------------

Total

- c. Calcule a média e o terceiro quartil da variável idade (dada abaixo); e sabendo que a variância desta variável idade é igual a 5,33, calcule o seu coeficiente de variação:

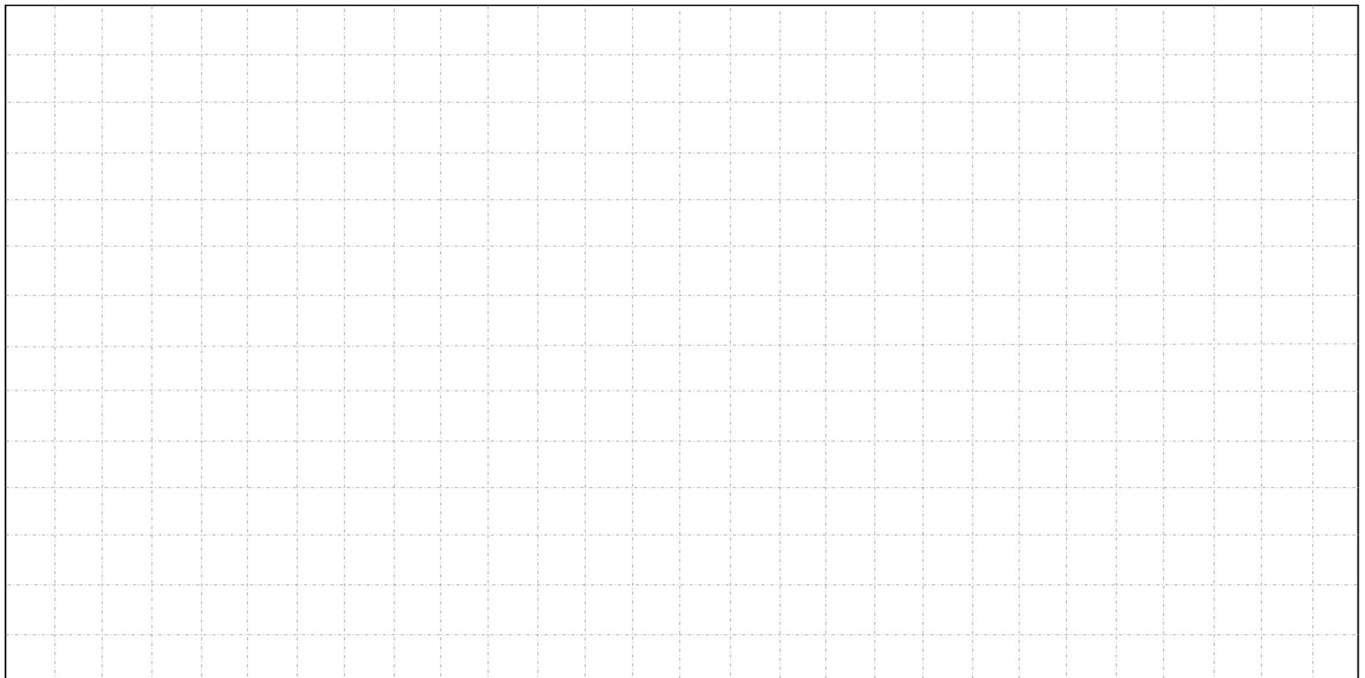
Idade (anos)	17	18	19	20	21	22	23	24
Frequência Absoluta	133	117	113	146	125	113	119	134

- d. Ao estudar o tempo gasto na Netflix de acordo com os diferentes gêneros observados, verificam-se os seguintes resultados:

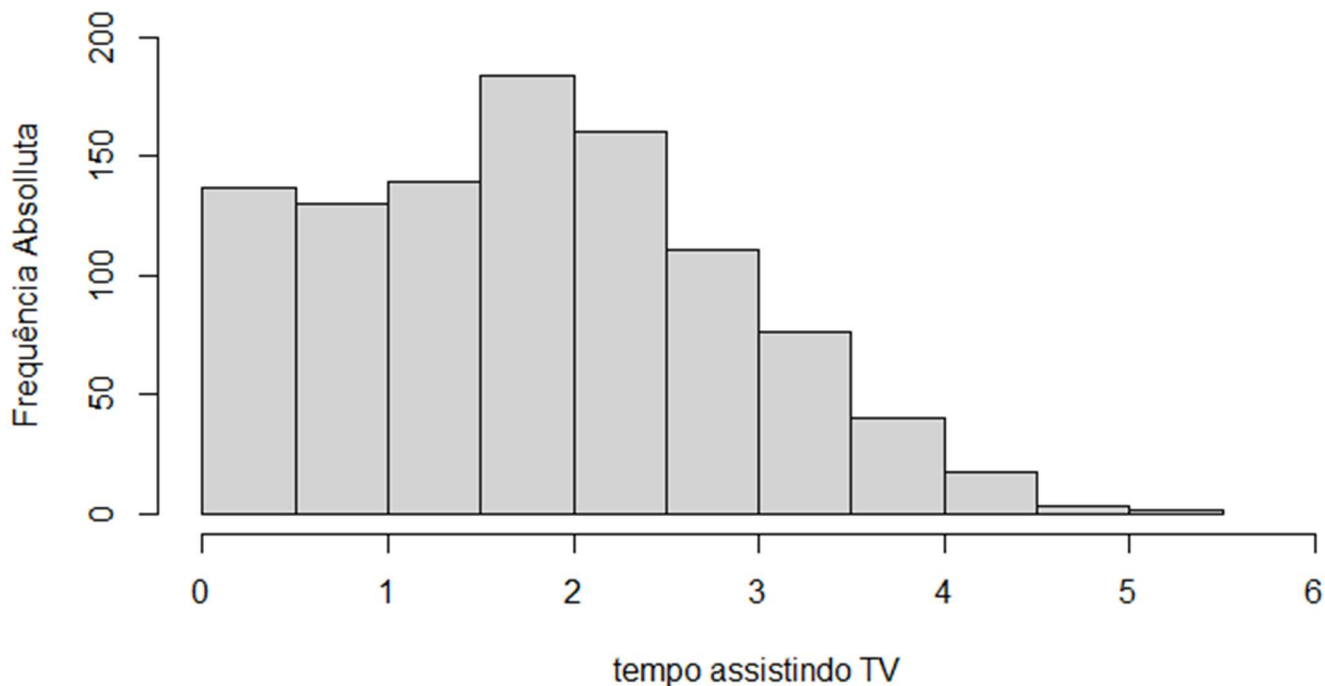
Gênero	mínimo	Q1	Q2	Q3	máximo	três menores valores	três maiores valores
Feminino	0	1	1,8	2,5	5,4	0 0 0	4,5 5,3 5,4
Masculino	0	1	1,8	2,6	4,9	0 0 0	4,3 4,6 4,9
Outro	0	1,125	2,05	2,875	5	0 0 0,1	4,3 4,3 5

Alguns dos três conjuntos de dados acima possuem valores atípicos? Se sim, qual(is)? Justifique sua resposta.

- e. Construa os três box plots, lado a lado, dos dados de tempo gasto na Netflix de acordo com os diferentes gêneros:

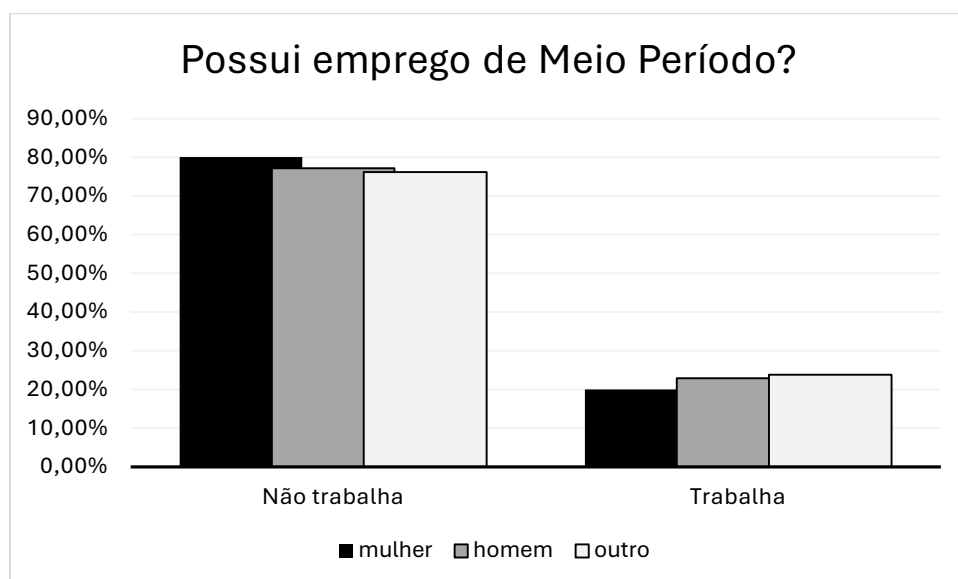


- f. Observe o histograma abaixo, sobre o tempo gasto na Netflix, independentemente do gênero, e discuta sobre a distribuição dos dados. Ela é simétrica? Justifique.

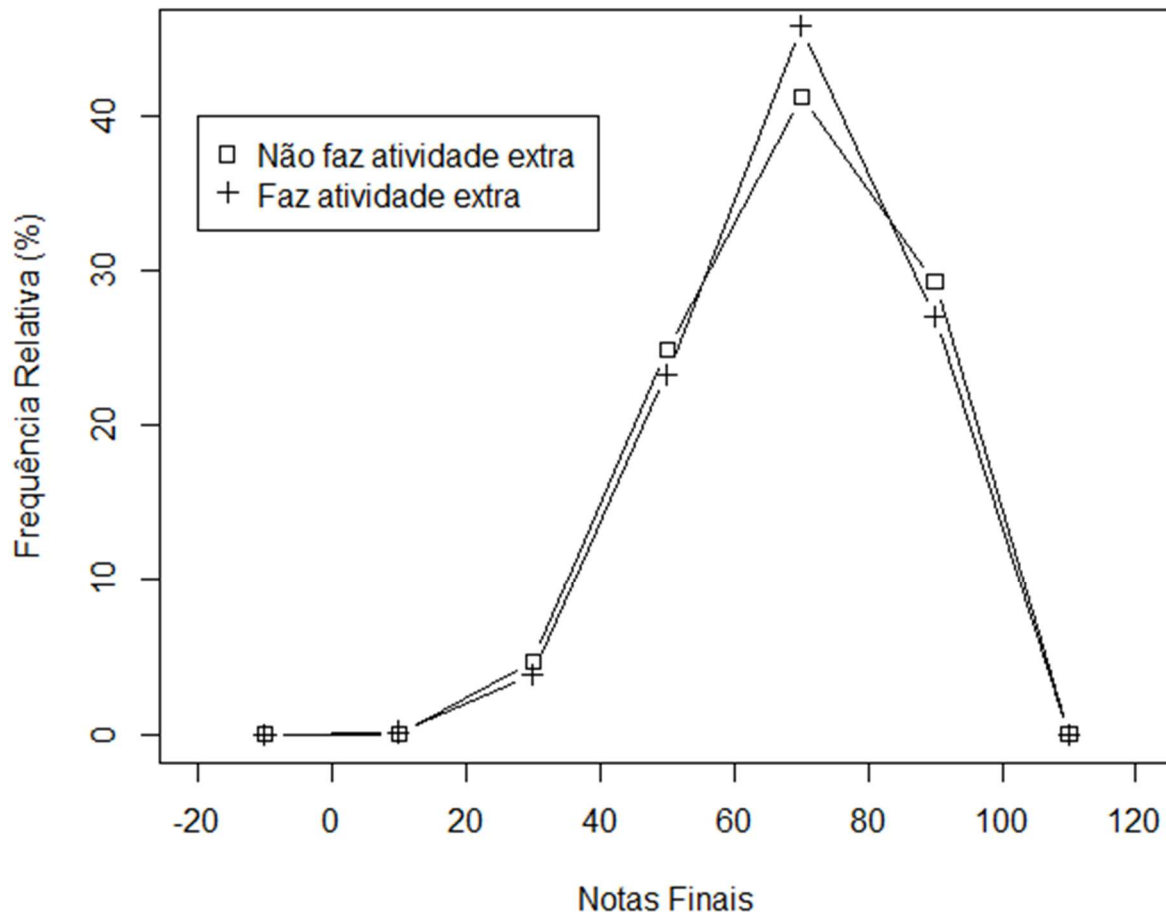


- g. Considerando que a média e a variância do tempo médio gasto na Netflix foram de 1,82 horas e 1,16 horas², calcule o quanto seria a média e a variância deste tempo se ele fosse calculado em segundos:
- h. Abaixo são dados dois gráficos para a visualização conjunta das variáveis: possuir ou não emprego de meio período relacionado com o gênero do aluno; e sobre as notas finais com relação a participar ou não de atividades extracurriculares. Utilize as informações de ambos os gráficos para classificar as afirmações abaixo em verdadeiras (V) e falsas (F):

- () Independentemente do gênero, mais de 70% dos alunos não possuem vínculo empregatício de meio período.
- () É possível afirmar que entre os alunos que não trabalham, a maioria se identifica como mulher.
- () Não é possível afirmar que, independentemente de trabalhar ou não, a distribuição de alunos de acordo com a identidade de gêneros é bem parecida.
- () O gênero dos alunos parece ter um impacto grande no fato deles estarem empregados ou não.



Polígono de Freq. Rel. das Notas vs Atividade Extracurricular



- () As notas entre 80 e 100 são um pouco mais frequentes para o grupo de alunos que não faz atividades extracurriculares.
- () Dentre os alunos que não fazem atividades extracurriculares, a frequência de notas abaixo de 60 (reprovados) é maior do que a frequência dessas notas no outro grupo.
- () O gráfico de polígonos aponta para distribuições de notas bem diferentes entre alunos que fazem, ou não atividades extracurriculares.
- () O gráfico de polígonos mostra que é possível tirar notas negativas.

Formulário

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad ; \quad q_{(p)} = \begin{cases} x_{(j+1)}, & \text{se } j \text{ for não inteiro} \\ \frac{x_{(j)} + x_{(j+1)}}{2}, & \text{se } j \text{ for inteiro} \end{cases} \quad ; \quad j = \frac{np}{100}$$

$$DIQ = Q_3 - Q_1 \quad ; \quad ROT = (Q_1 - 1,5DIQ ; Q_3 + 1,5DIQ)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad ; \quad s = \sqrt{s^2} \quad ; \quad CV = \frac{s}{\bar{x}} \times 100\%$$

Gabarito

Item 1:

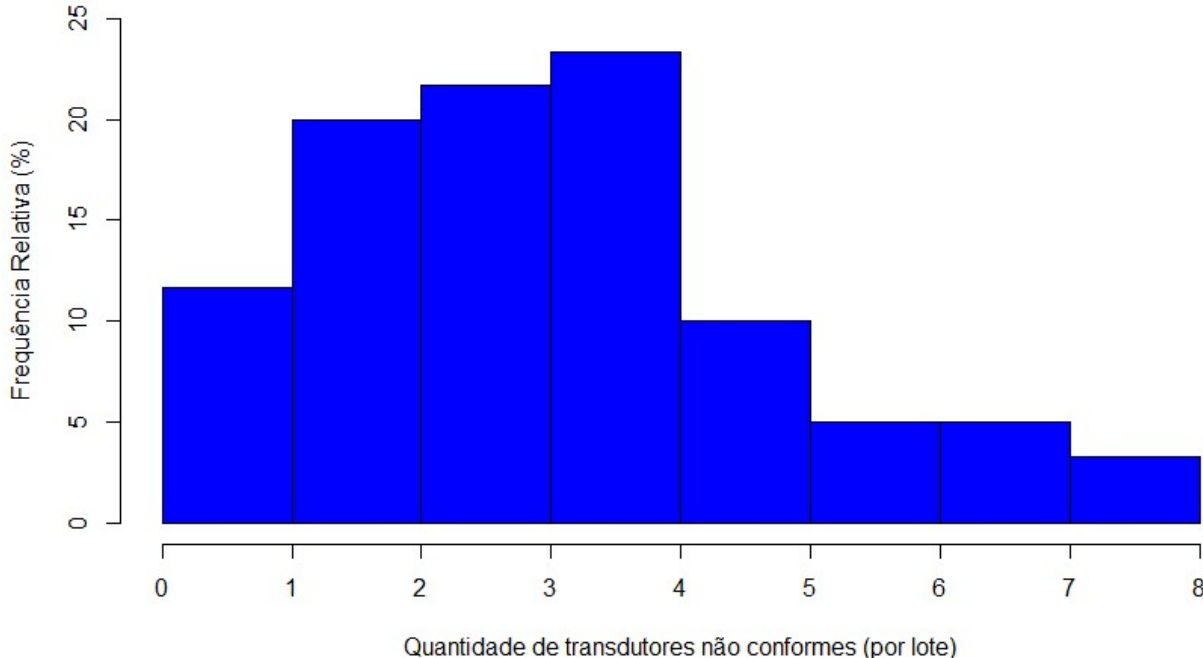
- 1- Idade – Quantitativa Discreta; Renda – Quantitativa Contínua; Escolaridade – Qualitativa Ordinal; Número de viagens no ano – Qualitativa Discreta; Meio de transporte – Qualitativa Nominal; Objetivo - Qualitativa Nominal; Estação (Alta ou Baixa) – Qualitativa Nominal; Duração da viagem – Quantitativa Discreta.
- 2- [QN]; [QD]; [QN]; [QC]; [QN]; [QO]
- 3- a.

Não Conformidade	Freq. Abs.	Freq. Rel. (%)	Freq. Rel. Ac. (%)
0	7	11,67	11,67
1	12	20	31,67
2	13	21,67	53,33
3	14	23,33	76,67
4	6	10	86,67
5	3	5	91,67
6	3	5	96,67
7	1	1,67	98,33
8	1	1,67	100
Total	60	100 *	

* o valor total da frequência Relativa difere da soma das frequências por erro de arredondamento.

b. $\frac{9167}{10000}$; $\frac{8667}{10000}$; $\frac{1333}{10000}$

c. O histograma está centrado em algum lugar entre 2 e 3 a distribuição se concentra principalmente entre os valores 0 e 4, sendo bastante esparsa entre 4 e 8, o que indica que um número de não conformidades superior à 4 ocorre, porém com menor frequência.



- 4- a. Os limites inferiores e superiores das classes são fechados, sendo assim, os valores limites pertencem a mais de uma classe.
- b.

Classe	Freq. Abs.	Freq. Rel. (%)	Freq. Rel. Ac. (%)
0 † 50	9	18	18
50 † 100	19	38	56
100 † 150	11	22	78
150 † 200	4	8	86
200 † 250	2	4	90

250 † 300	2	4	94
300 † 350	1	2	96
350 † 400	1	2	98
400 † 500	0	0	98
500 † 550	1	2	100
Total	50	100	

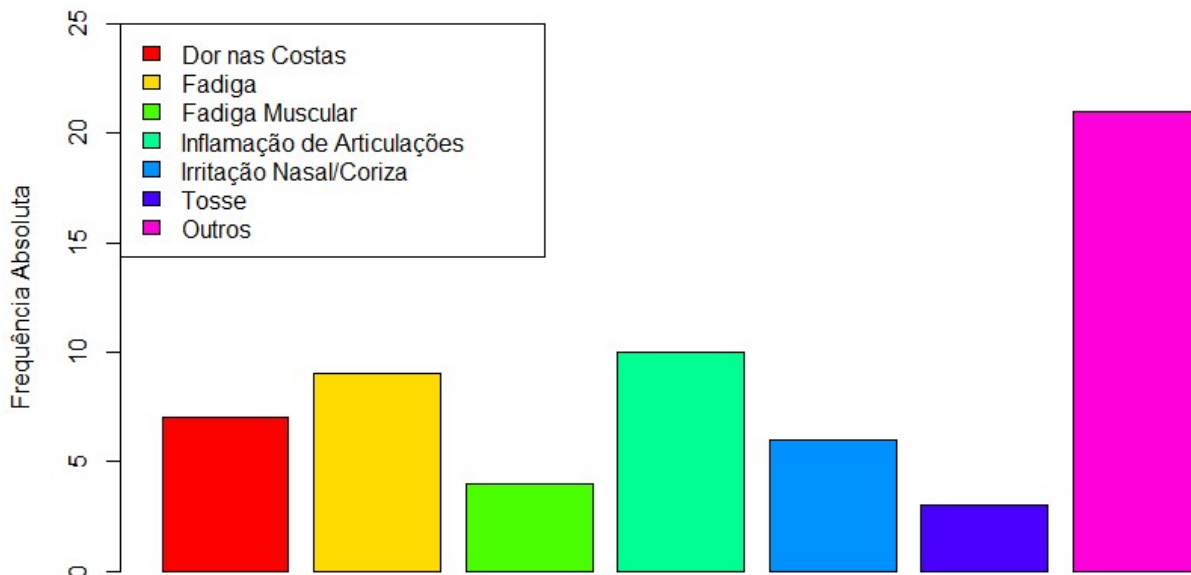
O valor 100 parece estar próximo do valor central. Há grande variação na vida útil, sendo que os dados estão mais concentrados nos valores inferiores à 200 e bastante dispersos até o valor máximo.

c. $\frac{56}{100}$; $\frac{14}{100}$

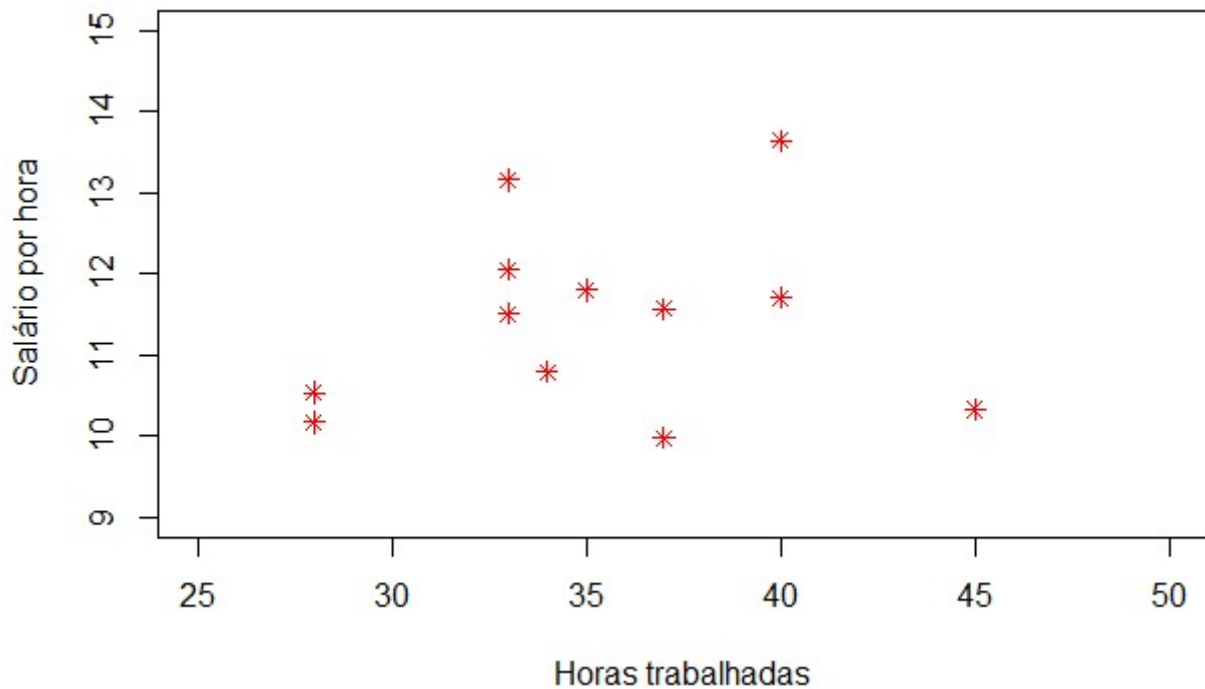
5-

Reclamação	Freq. Abs.	Freq. Rel. (%)
Dor nas Costas (B)	7	11,67
Fadiga (F)	9	15
Fadiga Muscular (M)	4	6,67
Inflamação de Articulações (J)	10	16,67
Irritação Nasal/Coriza (N)	6	10
Tosse (C)	3	5
Outros (O)	21	35
Total	60	100*

* o valor total da frequência Relativa difere da soma das frequências por erro de arredondamento.

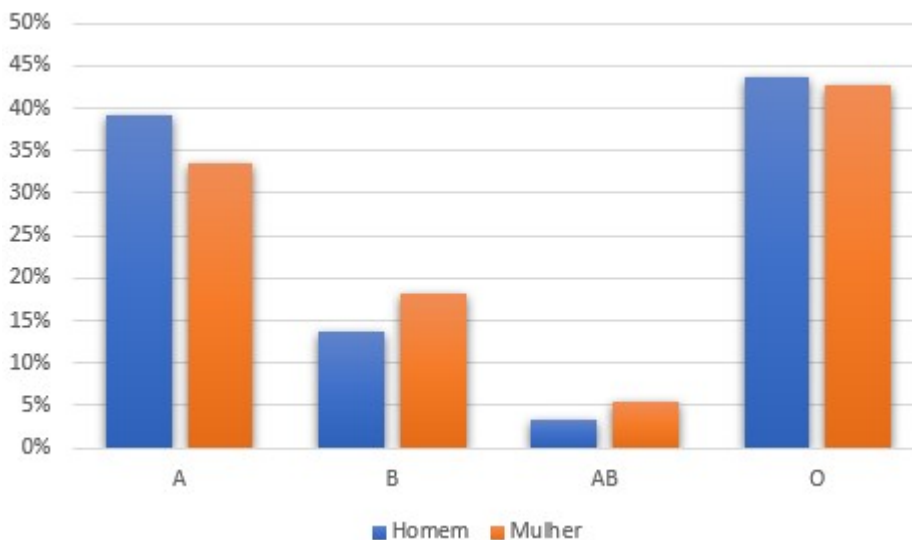


- 6- Ao organizar os dados em uma distribuição de frequências, pode-se ter uma visão geral dos dados de maneira resumida e simples e pode-se observar padrões de comportamento dos dados.
- 7- Quantitativo: histogramas e gráficos de dispersão / Qualitativo: diagrama de pizza e gráficos de barras.
- 8-



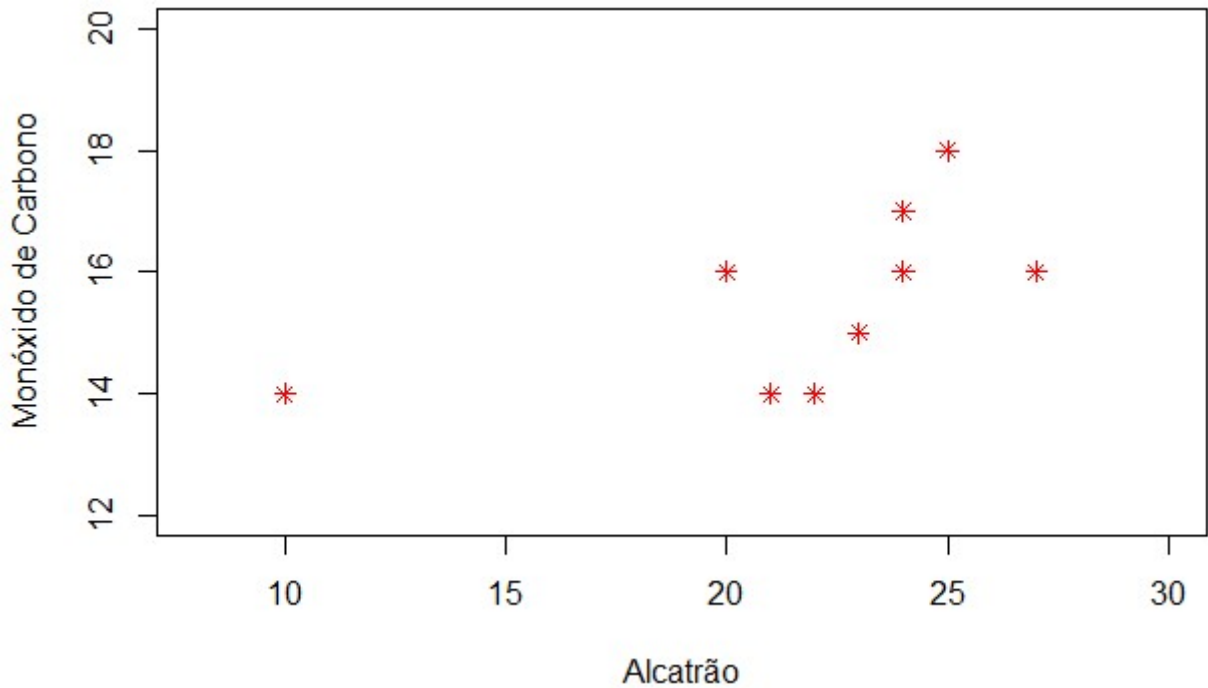
Não há uma tendência clara, mas parece que pessoas que trabalham poucas horas por semana (menos de 30) e pessoas que trabalham muitas horas (mais de 40) recebem pouco por hora trabalhada; já pessoas que trabalham entre 30 e 40 horas tem salários variados, provavelmente diferenciados pelo cargo exercido.

- 9- Apenas o gráfico A é adequado, já que o gráfico B tem o eixo vertical iniciando em 38% o que faz a diferença entre os resultados das duas dietas parecer maior do que ela é na realidade.
- 10- a. O gráfico foi feito utilizando a frequência absoluta, o que não permite a comparação entre as distribuições das frequências de indivíduos dentro de cada tipo sanguíneo de acordo com o sexo.



b.

- 11- Aparentemente, existe uma tendência crescente entre as duas variáveis, aumentando a quantidade de monóxido de carbono ao aumentar a quantidade de alcatrão, essa relação, no entanto não parece ser muito forte.



- 12- a. dados discretos
 b. dados contínuos
 c. dados discretos (se for considerado dias, meses ou anos completos) ou contínuos (se for considerado tempo corrido em horas, por exemplo)
 d. dados discretos
- 13- A taxa de pulsação masculina parece ser, num geral, mais baixa do que a das mulheres, sendo que mais de 50% dos homens tem taxa de pulsação inferior a 70, e apenas 30% das mulheres tem taxa abaixo de 70. Não foram observados homens com taxa superior a 100, isso já foi observado para as mulheres, embora em pouca quantidade.

Item 2:

- 1- a. $\bar{x} = 1,03$; $md = 1,01$. A mediana se encontra à esquerda da média, o que pode indicar uma leve assimetria à direita na distribuição dos dados, isso supondo que uma diferença de 0,02 seja importante neste tipo de variável.
 b. 0,383
- 2- Falso. A média é a medida de tendência central com a maior probabilidade de ser afetada por um valor atípico.
- 3- Falso. Todo conjunto de dados quantitativos tem uma mediana.
- 4- a. $\bar{x} = 4,57$, $md = 4,8$, $mo = 4,8$
 Mediana, uma vez que há uma possível assimetria
- 5- a. $\bar{x} = 170,63$, $md = 169,3$, $mo = amodal$
 b. Média, uma vez que a média e a mediana são semelhantes (possível simetria).
- 6- Elas usam abordagens diferentes para fornecer um valor (ou valores) central de conjunto de dados.
- 7- Não. O preço exatamente entre o maior e o menor é o ponto médio, não a mediana.
- 8- $\bar{x} = 53,3$; $md = 52$; $mo = amodal$. Usando a média de 53,3 palavras por página, a estimativa do número total de palavras é 77.765; já utilizando a mediana esse valor é de 75.920. Como as medidas estimadas se baseiam em uma pequena amostra e os números de palavras definidas nas páginas parecem variar bastante, as estimativas não podem ser consideradas precisas, basta ver o quanto elas se diferem entre si.
- 9- 85,5 kg

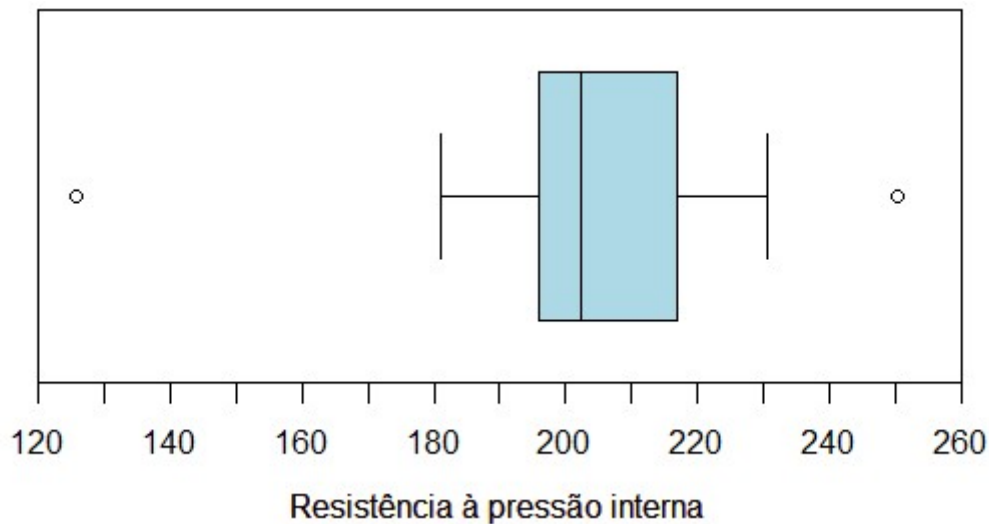
Item 3:

- 1- a. $s^2 = 1264,77$; $s = 35,56$ (minutos)
 b. $s^2 = 0,35$; $s = 0,59$ (horas)

- 2- a. $\bar{y} = a\bar{x} + b$; $s_y^2 = a^2s_x^2$
- 3- $s^2 = 189,14$; $s = 33,872$
- 4- As rendas de uma amostra aleatória simples de 1.000 adultos selecionados da população geral, já que dentro dessa população teremos a possibilidade de amostrar indivíduos sem renda, ou com renda muito baixa e indivíduos com rendas exorbitantes; já em uma amostra de 1000 professores de estatística os salários devem ser bem mais compatíveis entre si.
- 5- $s^2 = 245,12$; $s = 15,66$. Considerando que o desvio padrão é igual a 15,66 e a média é de 53,3, podemos dizer que existe uma variação grande dos dados, o que afeta consideravelmente a confiança nas estimativas calculadas anteriormente.
- 6- 30 dias antes: 8,56%; 1 dia antes: 33,75%. Os custos das passagens compradas com 30 dias de antecedência têm menos variação do que os custos de passagens compradas com 1 dia de antecedência. Há uma economia substancial na compra de passagens com 30 dias de antecedência.

Item 4:

- 1- a. Sim, já que $ROT = (164,8 ; 248)$, os outliers são os valores 125,8 e 250,2.
 b. Além da presença de outliers, a distribuição em torno da mediana é assimétrica, havendo uma maior concentração de observações entre o primeiro quartil e a mediana, também, os dados estão mais concentrados entre o primeiro e o terceiro quartis do que nas extremidades.



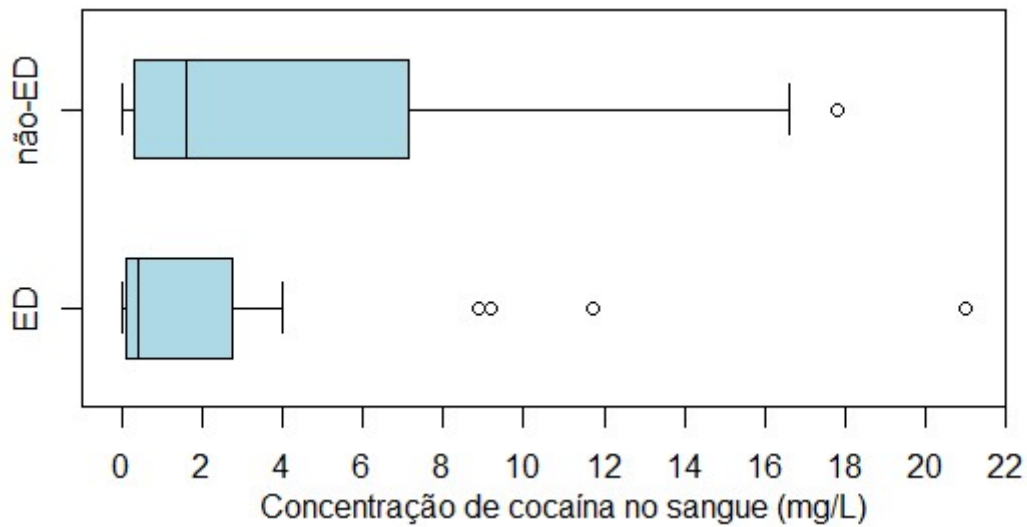
- 2- a.

Resumo dos 5 números	ED	Não-ED
Mínimo	0	0
Primeiro Quartil	0,1	0,3
Mediana	0,4	1,6
Terceiro Quartil	2,8	7,9
Máximo	21	17,8

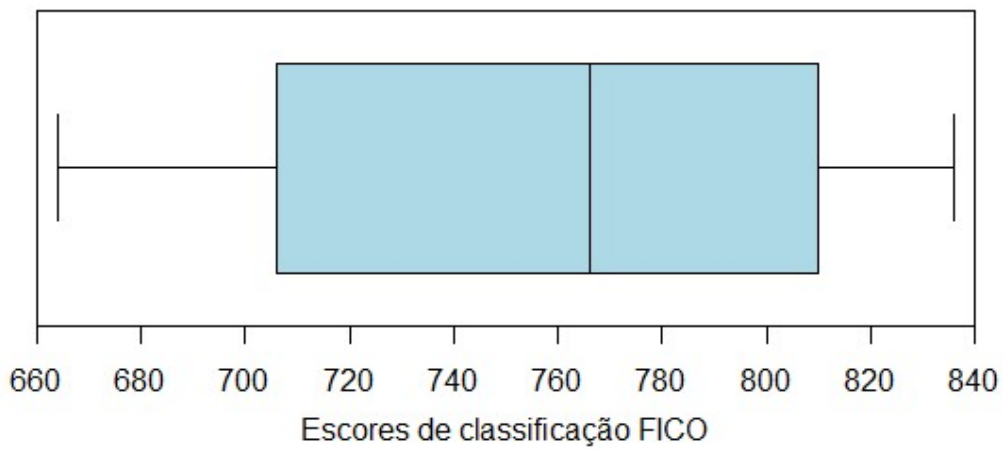
- b. Sim

	ED	Não-ED
DIQ	2,7	7,6
ROT	(-3,95 ; 6,85)	(-11,1 ; 19,3)
Outliers	8,9 ; 9,2 ; 11,7 ; 21	

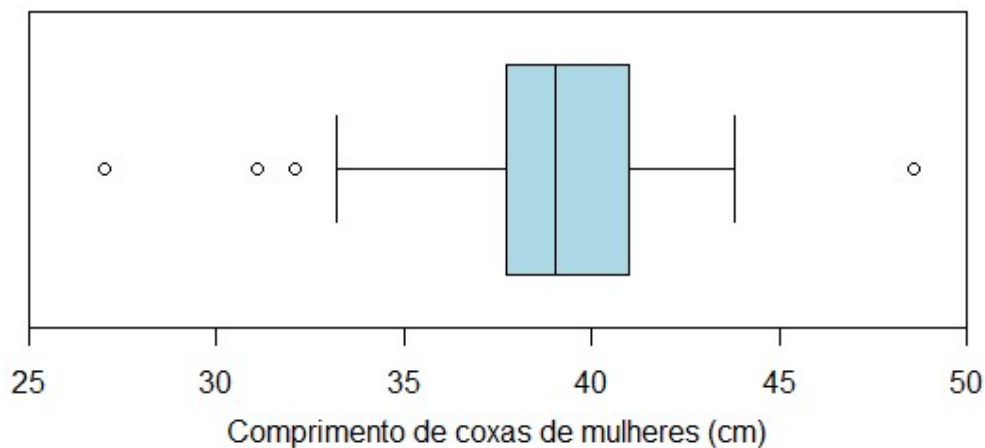
- c. Existem outliers apenas no conjunto de dados *ED*. Ambas as amostras são assimétricas à direita, tendo uma maior concentração nos valores mais baixos, os valores no conjunto *Não-ED* tendem a ser um pouco maiores do que a do conjunto *ED*. A dispersão no conjunto *Não-ED* é maior.



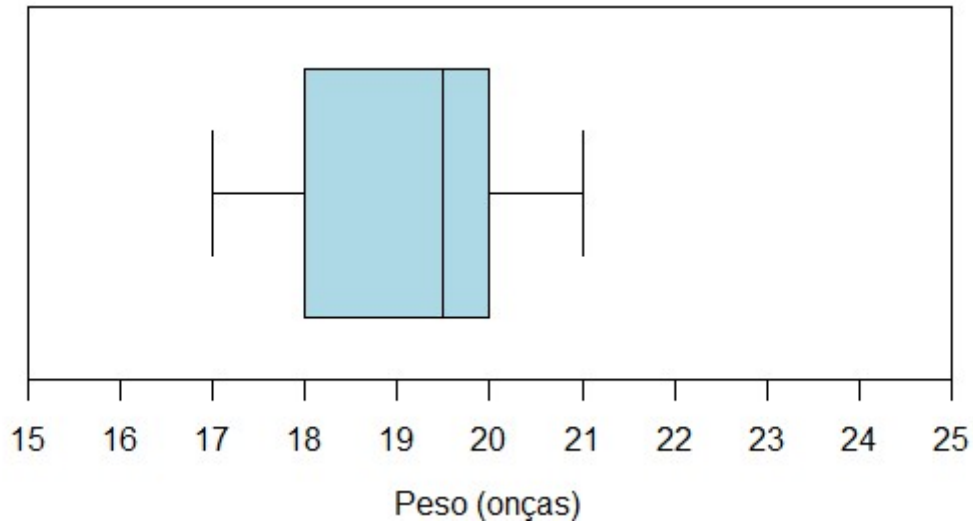
3- $\text{mín} = 664$; $Q_1 = 706$; $Q_2 = 766$; $Q_3 = 810$; $\text{máx} = 836$; $DIQ = 104$; $ROT = (550 ; 966)$. Esse conjunto de dados não possui valores atípicos.



4- $\text{mín} = 27 \text{ cm}$; $Q_1 = 37,75 \text{ cm}$; $Q_2 = 39,05 \text{ cm}$; $Q_3 = 41 \text{ cm}$; $\text{máx} = 48,6 \text{ cm}$; $DIQ = 3,25 \text{ cm}$; $ROT = (32,875 ; 45,875) \text{ cm}$. Esse conjunto de dados possui 4 valores atípicos.



5- $\bar{x} = 19,3 \text{ onças}$; $md = 19,5 \text{ onças}$; $mo = 20 \text{ onças}$; Amplitude = 4 onças; $s = 1,3 \text{ onça}$; $s^2 = 1,8 \text{ onça}^2$; $Q_1 = 18 \text{ onças}$; $Q_3 = 20 \text{ onças}$.



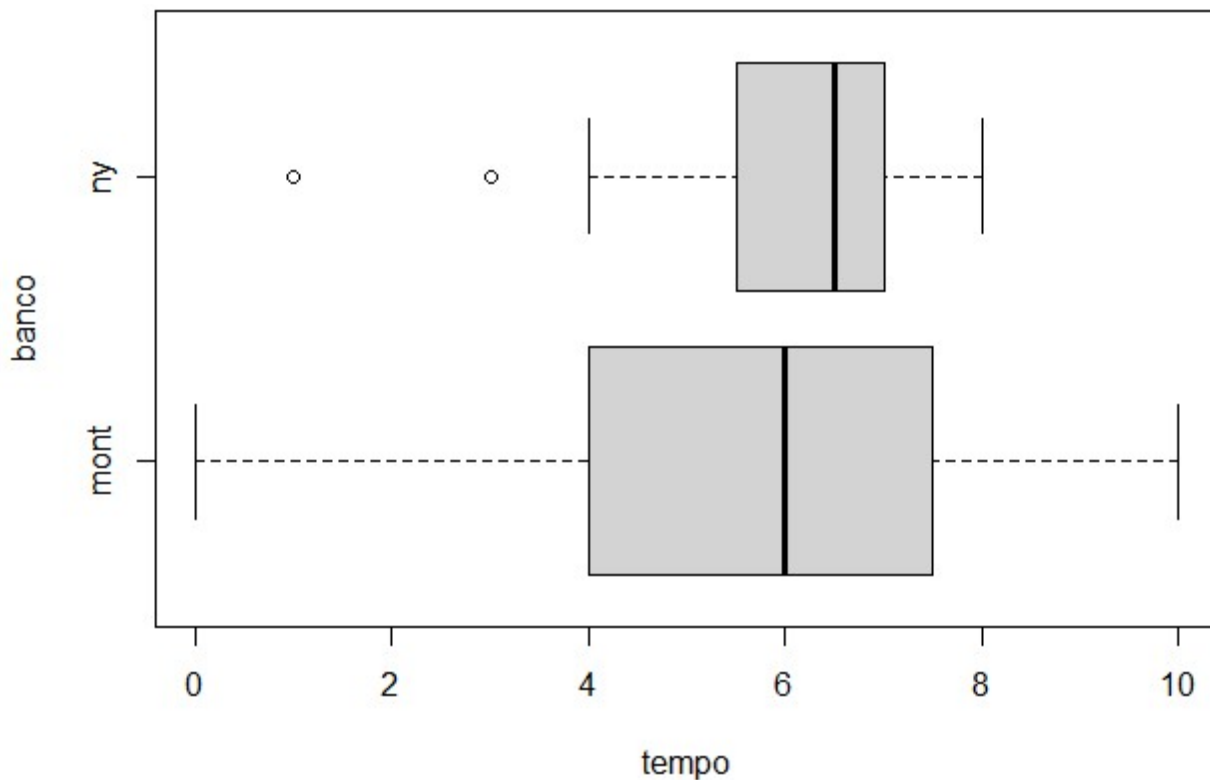
Item 5:

1-

n° de filhos	Freq. Abs.	Freq. Rel. (%)	Freq. Rel. Ac. (%)
0	32	20	20
1	40	25	45
2	56	35	80
3	24	15	95
4	0	0	95
5	8	5	100
Total	160	100	

2- a. $\bar{x}_{NY} = 5,83$ e $CV_{NY} = 34,53\%$; $\bar{x}_M = 5,79$ e $CV_{NY} = 42,90\%$

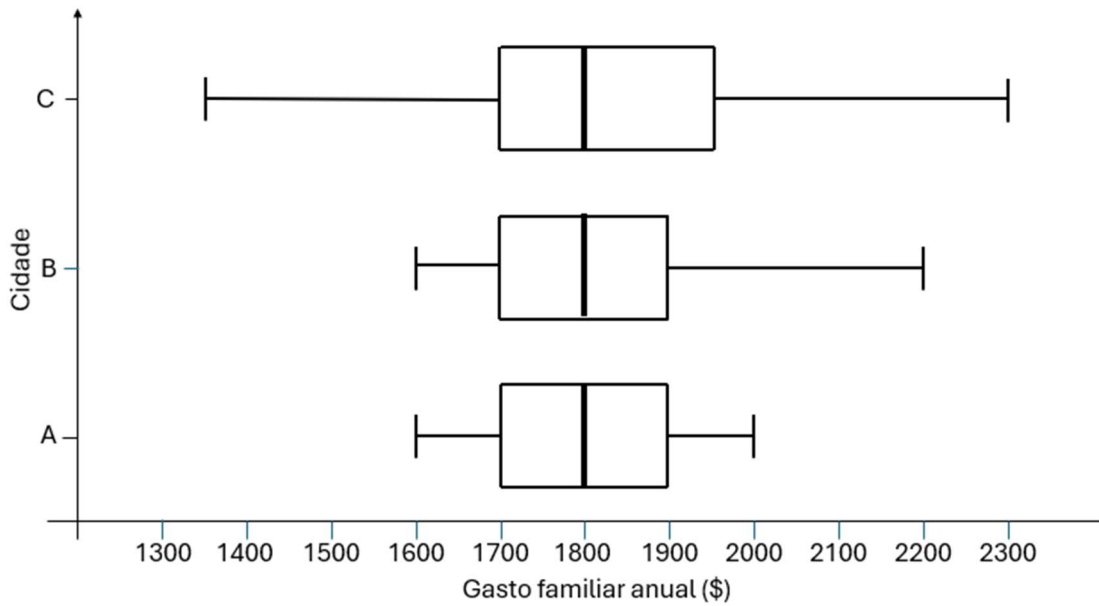
b.



c. Ambas as distribuições são assimétricas à esquerda, como pode ser observado no gráfico, e suportado pelo fato da ambas as médias calculadas no item (a) serem inferiores às medianas apresentadas no gráfico. Além disso, fica claro, tanto pelos coeficientes de variação como pelos box plots que os dados de montreal são mais heterogêneos que os dados de Nova York.

3- a. 45% ; b. 38,33%

4- a.



b. Os gastos das três cidades estão centrados no mesmo valor \$1.800, mas fica claro que os dados da cidade A estão menos dispersos, os dados da cidade B estão agrupados, como A, abaixo de \$1.800, porém dispersos acima de \$1.900. Já os dados da cidade C estão dispersos tanto acima de \$1.800 quanto abaixo de \$1.700.

c. $CV_A = 6,67\%$; $CV_B = 8,29\%$; $CV_C = 10,99\%$

d. tanto os box plots quanto os CV's indicam uma maior homogeneidade da cidade A, a cidade B apresenta homogeneidade intermediária e a cidade C apresenta a maior heterogeneidade.

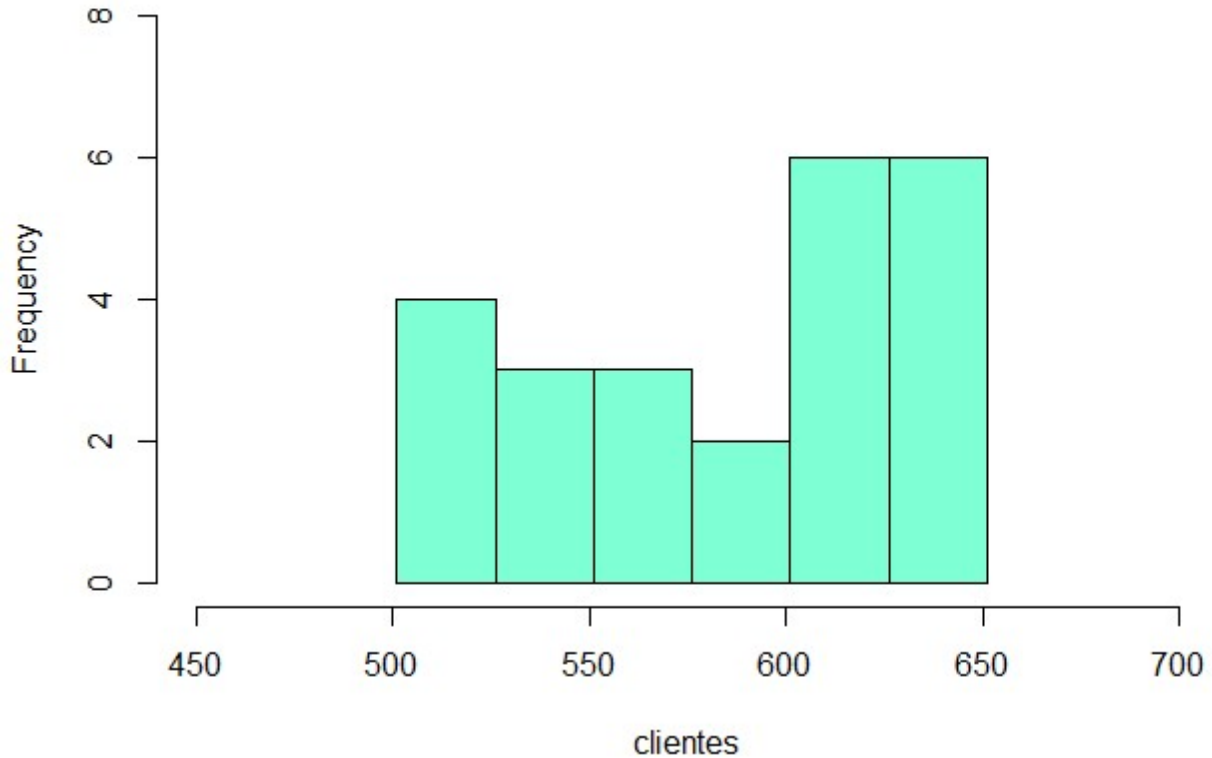
5- Figura 1 – média mais à direita ; Figuras 2 e 3 – médias próximas das medianas.

6- a.

Classe	Freq. Abs.	Freq. Rel. (%)	Freq. Rel. Ac. (%)
501 – 526	4	16,67	16,67
526 – 551	3	12,50	29,17
551 – 576	3	12,50	41,67
576 – 601	2	8,33	50,00
601 – 626	6	25,00	75,00
626 – 651	6	25,00	100,00
Total	24	100	

b.

Histogram of clientes



c. Não, já que o histograma não pode ser considerado como espelhado em torno de 585,29, além disso a média se encontra à esquerda da mediana (599), o que indica que essa variável apresenta uma assimetria à esquerda.

7- a. Estatística: média = 61,25 ; mediana = 61 ; moda = 50 / matemática: média = 62,25 ; mediana = 60 ; amodal

b. Para estatística, a média e a mediana são semelhantes, sendo assim, pode-se dizer que sua distribuição deve ser aproximadamente simétrica, já para matemática, a média e a mediana são mais afastadas, o que pode indicar uma assimetria na distribuição das notas.

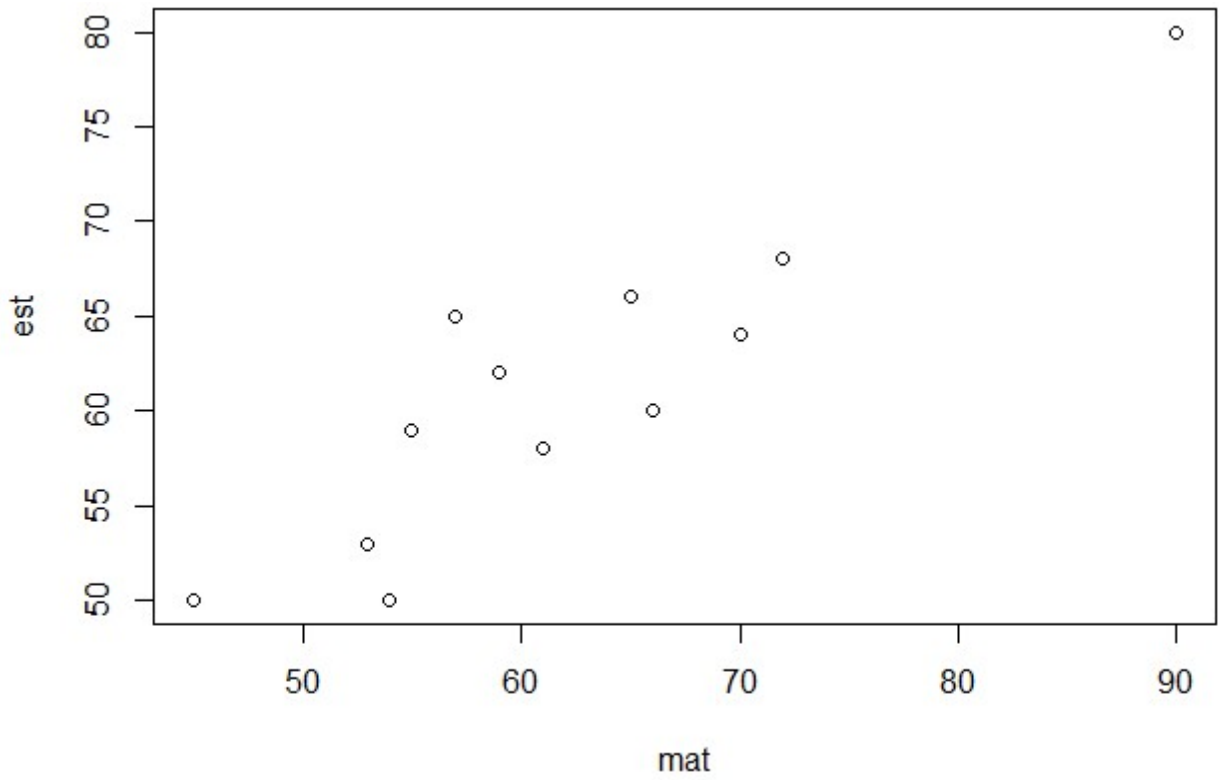
c. Estatística: CV = 13,75% / matemática: CV = 18,70%. Olhando pelo coeficiente de variação, pode-se dizer que as notas de estatística são mais homogêneas que as notas de matemática.

d.

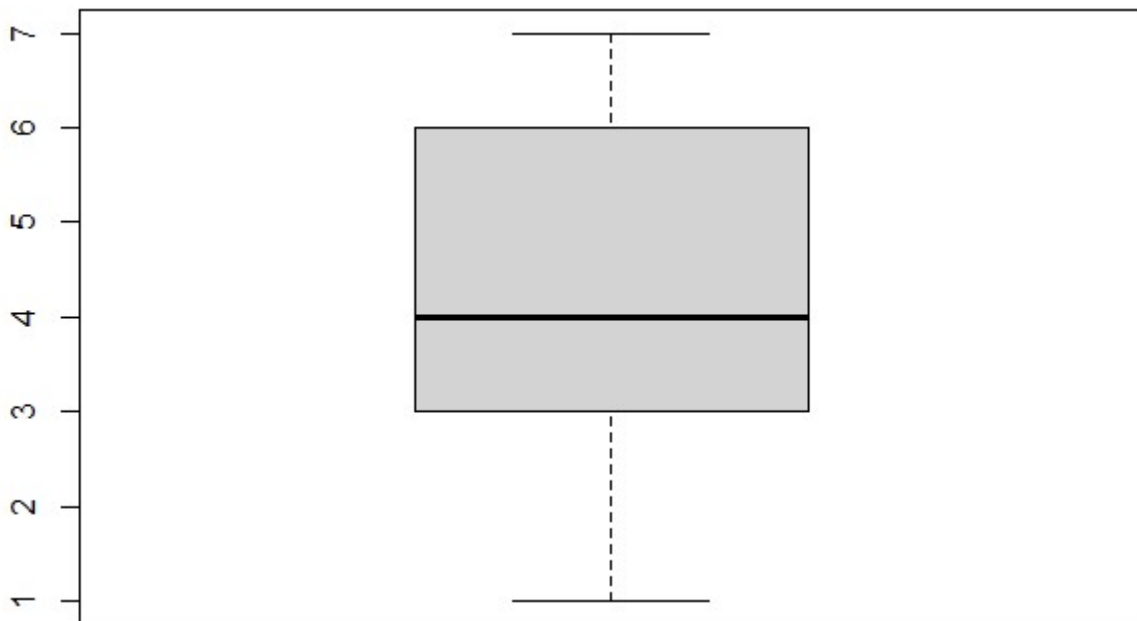
Resumo dos Cinco Números	Estatística	Matemática
Mínimo	50	45
Primeiro Quartil	55,5	54,5
Mediana	61	60
Terceiro Quartil	65,5	68
Máximo	80	90

e. Estatística: DIQ = 10 e ROT = (40,5 ; 80,5) não existem valores atípicos / Matemática: DIQ = 13,5 e ROT = (34,25 ; 88,25) existe um valor atípico, o 90.

f.



- 8- a. 6
 b. 4
 c. $Q_1 = 3$; $Q_2 = 4$; $Q_3 = 6$
 d. $DIQ = 3$ e $ROT = (-1,5 ; 10,5)$ não existem valores atípicos



- 9- item (a)
 10- item (b)
 11- item (a)
 12- item (c)
 13- item (d)
 14- item (b)
 15- item (d)
 16- item (e)
 17- item (d)
 18- item (d)

- 19- item (a)
- 20- item(d)
- 21- item (d)
- 22- item (b)

Item 6:

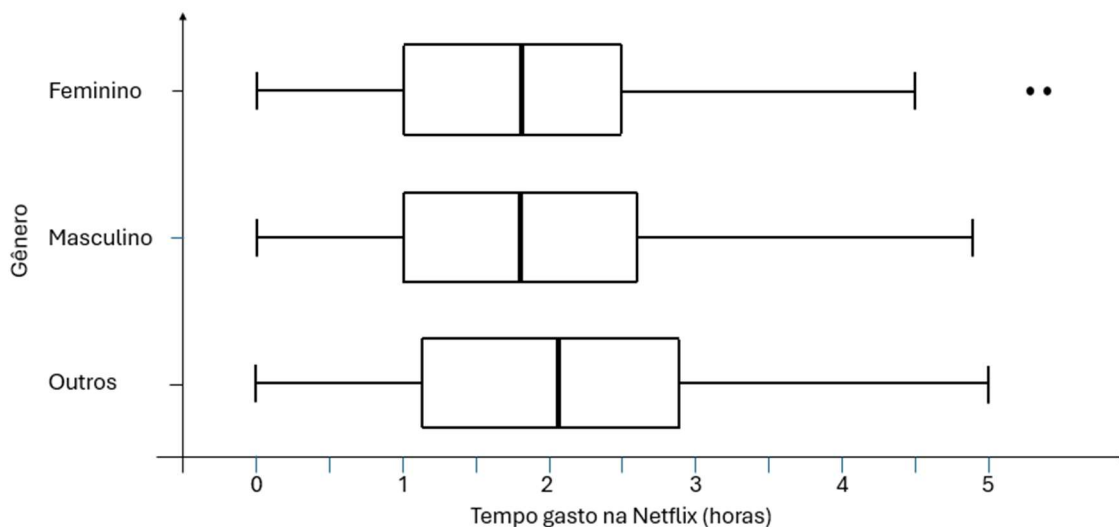
23- a. Idade: Quantitativa Discreta ; Gênero: Qualitativa Nominal; Tempo de estudo diário: Quantitativa Contínua; Qualidade da Internet: Qualitativa Ordinal

b.

Classes	Freq. absoluta	Freq. Rel. (%)	Freq. Rel. acumulada (%)
Pobre	185	18,5	18,5
Razoável	437	43,7	62,2
Boa	378	37,8	100
Total	1000	100	

c. $\bar{x} = 20,498$ anos; $Q_3 = 18,5$ anos; $CV = 11,26\%$

d. $ROT_{fem} = (-1,25 ; 4,75)$; $ROT_{masc} = (-1,4 ; 5)$; $ROT_{outro} = (-1,46 ; 5,46)$. Logo o único conjunto que possui valores atípicos é o do gênero feminino, e os valores atípicos são: 5,3 e 5,4 horas.



e.

f. A distribuição é assimétrica, já que a frequência varia pouco entre 0 e 1,5 horas (iniciando já próximo a 14%), aumenta entre 1,5 e 2,5 horas e depois cai rapidamente.

g. $\bar{x} = 6.552$ segundos ; $s^2 = 15.033.600$ segundos²

h. V ; F ; V ; F ; V ; V ; F ; F