

USP – ICMC – Análise Exploratória de Dados

3ª lista de exercícios – 1º/2010

Parte 1

1. Compare os dados de distribuição de renda dos anos de 1970 e 1980 usando medidas de posição e dispersão.

Distribuição de renda – população com rendimento

Rendimento mensal em SM	% da população – 1970	% da população – 1980
Até 1/2 SM	28,4	12,5
Mais de 1/2 a 1 SM	32,2	20,8
Mais de 1 a 2 SM	21,7	31,1
Mais de 2 a 5 SM	12,7	23,6
Mais de 5 a 10 SM	3,3	7,2
Mais de 10 a 20 SM	1,3	3,2
Mais de 20 SM	0,4	1,6

Fonte: IBGE – Folha de São Paulo (modificado)

4. A figura ao lado procura transmitir o quão barato é anunciar na televisão em São Paulo no horário nobre. Faça um comentário a respeito.

VOCÊ SABIA?

**Atingir mil pessoas pela
televisão no horário
nobre em São Paulo
custa o mesmo
que 5 xícaras de café**

2. Na tabela abaixo são apresentadas variáveis que procuram quantificar a utilização de instrumentos de *marketing* em determinados setores industriais. Proponha um critério que lhe permita concluir se um dado instrumento é bem utilizado. Com base nas respostas da tabela aponte os instrumentos bem utilizados de acordo com o critério proposto.

Variáveis	Frequências das respostas			
	1 e 2	3	4 e 5	n
Questionário sobre as necessidades do consumidor	93	1	4	98
Informações com vendedores sobre necessidades do consumidor	49	8	41	98
Informações com fornecedores sobre necessidades do consumidor	50	6	28	94
Conversa com os consumidores sobre ações dos concorrentes	34	17	47	98
Controle de qualidade das matérias-primas antes da fabricação	5	7	85	97
Utiliza embalagem como proteção do produto	48	1	49	98
Utiliza propaganda em jornais	89	4	5	98
Faz orçamento sem compromisso	37	4	57	98
Oferece descontos para promover as vendas	34	10	54	98
Faz treinamento dos vendedores através de cursos	94	1	2	97

1:nenhum, 2:pouco, 3:moderado, 4:bastante e 5:total

3. Em uma competição um atirador dispara 200 vezes. Tiros à esquerda do alvo são considerados distâncias negativas e à direita são considerados positivas. A distribuição das distâncias é dada abaixo :

Distâncias (cm)	-30 + -20	-20 + -10	-10 + 0	0	0 — 10	10 + 20	20 + 30
Número de tiros	5	15	30	90	43	9	8

Calcule a distância média (a) dos tiros à esquerda do alvo, (b) dos tiros à direita do alvo, (c) dos tiros que erram o alvo e (d) de todos os tiros.

Parte 2

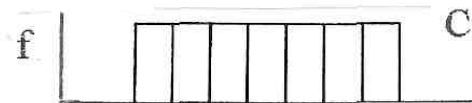
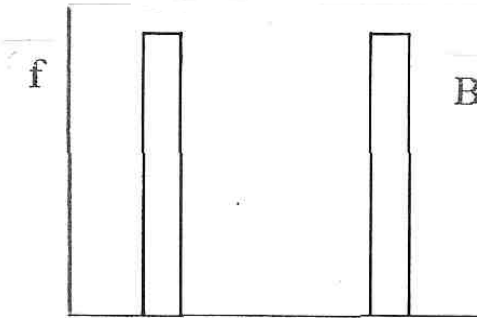
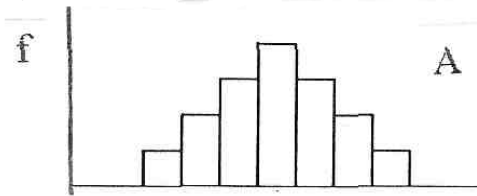
	Empresa A	Empresa B	
1.	Média	75,00	80,00
	Mediana	72,00	70,00
	Moda	67,00	62,00
	Q_1 e Q_3	62,00 e 78,00	65,00 e 85,00
	Desvio padrão	13,00	17,00

A tabela ao lado se refere a salários (em \$) de duas empresas. Compare-as usando medidas de posição, dispersão e assimetria.

2. (a) “Uma medida de dispersão deve ser baseada em todas os valores da amostra”. Apresente medidas com e medidas sem esta propriedade.
(b) “Uma medida de dispersão deve mostrar a variabilidade em relação a uma medida de posição”. Apresente medidas com e medidas sem esta propriedade.
(c) Em que situação o coeficiente de variação é maior que 100% ? O que você pode dizer sobre os itens da amostra neste caso ?
4. Na distribuição de freqüências dos gastos semanais de 105 famílias os números de famílias nas classes 100 † 200 e 300 † 400 não foram fornecidos. Entretanto, sabemos que a mediana e a moda valem 250 e 240, respectivamente.
 - (a) Complete a tabela e calcule a média aritmética.
 - (b) Apresente em um gráfico as 3 medidas de posição.

Parte 3

1. Os gráficos abaixo mostram histogramas obtidos de amostras de mesmo tamanho de três variáveis diferentes (A , B e C).



- (a) Indique nos gráficos a média, mediana, moda e amplitude para cada variável.
- (b) Apresente a ogiva da variável B .
- (c) Qual histograma apresenta maior desvio padrão ?
4. (a) Considere uma variável qualitativa que possui dois possíveis valores codificados como 0 e 1. Qual o significado, se existir, da média ?
- (b) Se o "sexo médio" (masculino=0, feminino=1) de uma amostra vale 0,75, o que isto significa ?

2. (a) Na estação de medição A temos precipitação anual média de 712,3 mm e desvio padrão 319,7 mm. Na estação B estes valores são 1446,3 mm e 319,3 mm, respectivamente. Com base nestes dados pode-se afirmar que a variabilidade é quase idêntica? Justifique. Pode-se afirmar que na estação A a quantidade de precipitação é maior? Justifique.

(b) As duas amostras abaixo se referem a estatura e superfície corporal de um grupo de 10 estudantes. Qual das duas variáveis apresenta maior dispersão?

Estaturas (cm) : 162 175 169 180 173 169 158 176 154 162

Superfícies (m^2) : 1,56 1,76 1,80 1,93 1,86 1,66 1,46 1,77 1,52 1,51

3. Assinale verdadeiro (V) ou (F) e justifique aquela(s) onde você respondeu falso.

(a) () Numa prova aplicada a 26 estudantes todos obtiveram a mesma nota. O desvio padrão vale 0.

(b) () 50% dos valores de uma amostra situam-se acima da moda.

(c) () O coeficiente de variação é uma medida que expressa o quociente entre o desvio padrão e a mediana.

(d) () Quando queremos verificar em uma relação de produtos aquele que apresenta maior aceitação utilizamos a moda.

(e) () A mediana é mais sensível a valores extremos que a média.

(f) () Numa pesquisa foi perguntado qual o bairro de preferência dos entrevistados na cidade de Vitória. Dos resultados obtidos foram calculadas média e moda.

(g) () Quanto menor a amplitude interquartil, maior o agrupamento dos dados em torno da mediana.