**Estatística com foco em Ciência de Dados e Business Analytics**

**Lista de exercício do Módulo 1 com respostas**

1. Determine se essas variáveis de uma coleta de dados de alunos são qualitativas ou quantitativas. Se quantitativas, especifique se são contínuas ou discretas. Também identifique a tipo de escala de medida: nominal, ordinal, intervalo ou razão.

(a) Gênero (M, F)

Qualitativa, nominal

(b) Número de irmãos

Quantitativa, discreta, razão

(c) Ano na universidade

Qualitativa, ordinal

(d) Carreira

Qualitativa, nominal

(e) Avaliação da comida no campus (ruim, ok, boa, excelente)

Qualitativa, ordinal

(f) Você mora no campus? (Sim/Não)

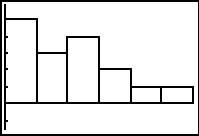
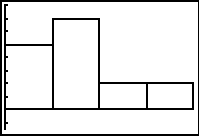
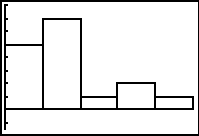
Qualitativa, nominal

(g) Número médio de horas de estudo por semana

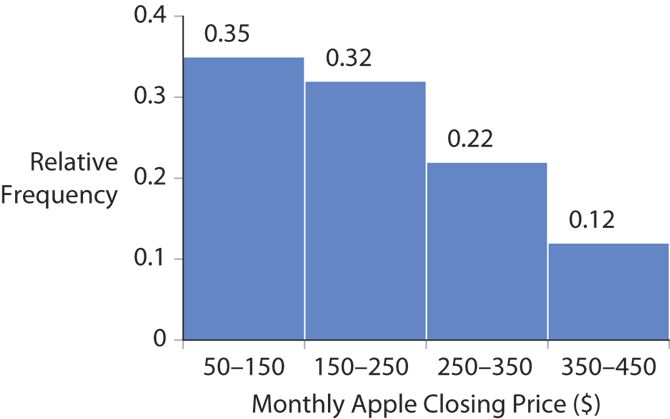
Quantitativa, contínua, razão

2. Uma professora pergunta aos seus alunos quantas horas, em média, eles trabalham por semana. As respostas são: 21, 30, 0, 0, 15, 0, 15, 12, 10, 32.5, 0, 15, 10, 15, 5, 25.

Um histograma começando no ponto 0 horas e usando largura da barra de 10 seria:

A.  B.  C. 

3. O histograma abaixo mostra a distribuição do preço mensal da ação da Apple entre os anos 2007 e 2011.



(a) A distribuição é simétrica ou assimétrica (positiva ou negativa)?

Assimétrica positiva

(b) Nesse período de 5 anos, qual é o valor da ação mensal aproximado mínimo e o valor máximo?

Mínimo $50; máximo $450

(c) Nesse período de 5 anos, qual classe teve a maior frequência relativa?

A classe $50-$150

4. Uma empresa de contabilidade paga para seu assistente administrativo $30,000, cinco escriturários $35,000 cada, dois contadores $70,000 cada, e o dono da firma $270,000.

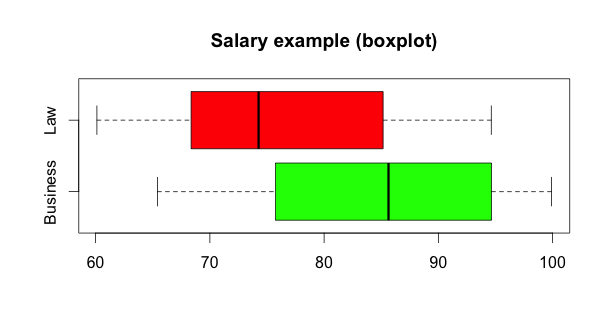
(a) Qual a mediana dos salários dessa firma?

30, 35, 35, 35, 35, 35, 70, 70, 270

(b) Suponha que o dono reduza seu salário para $200,000. Como a média dos salários seria afetada (aumentaria, diminuiria, manteria a mesma)? E a mediana?

A média também reduz, mas a mediana não alteraria.

5. A figura mostra os boxplots de dados de salário para pessoas na área de negócios (business) e na área de direito (law).



(a) Qual a mediana dos salários na área de negócios?

Aproximadamente $87

(b) Aproximadamente qual o percentual de pessoas na área de direito tem salário abaixo da mediana dos salários em negócios?

3 quartis, ou seja, ¾ = 75%

6. As notas de uma prova de estatística são:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| 75 | 25 | 75 | 62 | 80 | 85 | 80 | 99 | 90 | 60 |
| 86 | 92 | 40 | 74 | 72 | 65 | 87 | 70 | 85 | 70 |
|  | | | | | | | | | |

O primeiro quartil é 66,25, o segundo é 75, e o terceiro é 85,75. Encontre:

(a) A amplitude interquartil (IQR)

IQR = Q3 – Q1 = 85,75 – 66,25 = 19,5

(b) O limite inferior

Q1 – 1,5xIQR = 66,25 – 1,5x19,5 = 37

(c) O limite superior

Q3 – 1,5xIQR = 85,75 + 1,5x19,5 = 115

(d) Tem algum outlier nos dados?

Sim, tem uma nota menor que o limite inferior 37. A nota 25 é um outlier.

7. A tabela abaixo apresenta os retornos anuais (em percentual) para os fundos de eletrônicos e energia da Fidelity.

|  |  |  |
| --- | --- | --- |
| **Ano** | **Eletrônicos** | **Energia** |
| 2005 | 13,23 | 9,36 |
| 2006 | 1,97 | 32,33 |
| 2007 | 2,77 | 21,03 |
| 2008 | -50 | -35,21 |
| 2009 | 81,65 | 14,71 |

(a) Calcule a média e o desvio padrão para cada fundo.

|  |  |  |
| --- | --- | --- |
|  | Média | Desvio padrão |
| *Eletrônicos* | 9,92 | 47,07 |
| *Energia* | 8,44 | 25,86 |

(b) Qual fundo teve o maior retorno médio?

O fundo Eletrônicos teve o maior retorno médio.

(c) Qual fundo foi o mais arriscado nesse periodo?

O mais arriscado é o que tem maior variância (mais variabilidade = mais risco). Logo, o fundo Eletrônicos é o mais arriscado.

(d) Encontre a correlação e descreva a relação entre os dois fundos.

A correlação é 57%, os fundos são moderadamente relacionados.