**FACULDADE IMPACTA DE TECNOLOGIA**

**ANÁLISE EXPLORATÓRIA DE DADOS  
JOÃO MARCELO BOROVINA JOSKO**

**EXERCÍCIOS - LISTA 1**

**CST EM BANCO DE DADOS  
BRUNO SANCHES - RA:   
CÉSAR DE PAULA - RA:   
RONALDO DE MEDEIROS RIBEIRO - RA: 1700245**

**SÃO PAULO 2017**

Análise Exploratória de Dados - Prof. João Marcelo Borovina Josko - [Joao.Josko@impacta.edu.br](mailto:Joao.Josko@impacta.edu.br) Exercícios de Medidas de Resumo baseados nos exercícios Prof. Fernando de Pol Mayer

**Orientações:**

* Lista deve ser desenvolvida em grupos de **até 3 pessoas**, sem exceções
* Data da Entrega: **26*/*Set**
* Entrega via e-mail, com o seguinte cabeçalho no assunto “AED - Lista1"
* Coloque os nomes dos integrantes no trabalho e não no corpo do e-mail.

1. Estime as medidas de centro (média, mediana, moda) para amostras de altura de uma espécie de árvore (metros), coletadas em quatro áreas diferentes:

•Área A: 9,2 10,8 10,6 11,1 12,1 9,6 11,2 8,4 12,9 12,1 14,4 11,1 11,1 9,7 8,4 12,3 10,7 12,9 9,1 12,8;

•Área B: 12,5 18,5 21,3 14,3 18,5 19,0 10,8 23,1 17,4 10,7 14,3 16,3 18,0 7,1 12,8 14,7 11,3 8,2 13,8;

•Área C: 21,3 28,7 15,8 24,0 13,7 18,1 12,6 14,6 6,1 19,8 22,3 15,7 16,3 18,2 15,7 6,6 9,3 1,3 19,0;

•Área D: 13,7 8,6 14,9 10,2 14,0 10,5 15,0 5,2 10,0 11,7 18,7 9,3 7,9 6,5 11,5 12,0 8,3 8,3 9,8 4,7.

* 1. Calcule as medidas de centro: média, mediana e moda para cada área (b)Calcule a amplitude, a variância e o desvio-padrão para cada área

(c) Descreva comparativamente as quatro áreas quanto à altura das árvores, utilizando as estatísticas que você calculou.

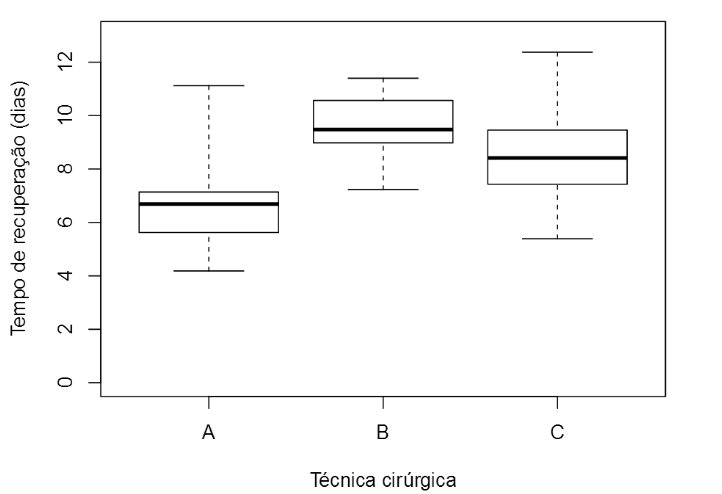
1. Um exame vestibular para uma faculdade tem 80 questões, sendo 40 de português e 40 de matemática. Para os 20 melhores classificados, apresentamos o número de acertos em cada disciplina.

•Português: 35, 35, 34, 32, 31, 30, 26, 26, 24, 23, 23, 12, 11, 20, 17, 12, 14, 20, 8, 10

•Matemática: 31, 29, 27, 28, 28, 26, 30, 28, 25, 23, 21, 32, 31, 20, 21, 25, 20, 13, 23, 20

* 1. Calcule as medidas de centro: média, mediana e moda para cada grupo
  2. Calcule as medidas de variabilidade: variância, desvio-padrão, e coeficiente de variação para cada grupo (c)Construa um gráfico de caixa (box plot) para cada grupo (em um mesmo gráfico para comparação)

1. Com todos os resultados obtidos, descreve comparativamente estes dois grupos em termos de medidas de tendência central, variabilidade, amplitude e distribuição (simetria) dos dados.
2. Você acha que os aprovados são melhores em português ou matemática?
3. Deseja-se comparar três técnicas cirúrgicas para a extração do dente siso. Cada uma das técnicas foi aplicada a 30 pacientes, e os resultados são apresentados em diagramas de caixa abaixo.

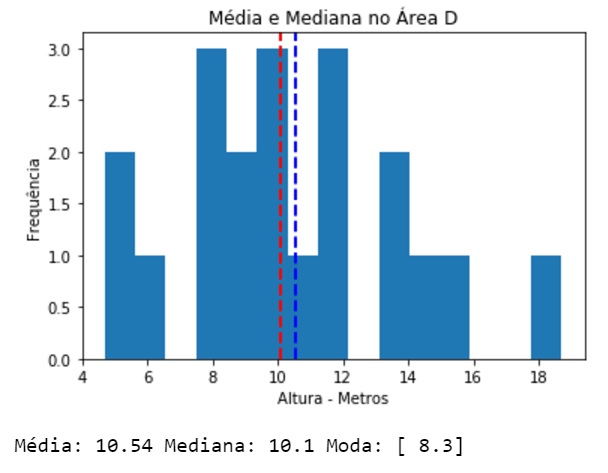
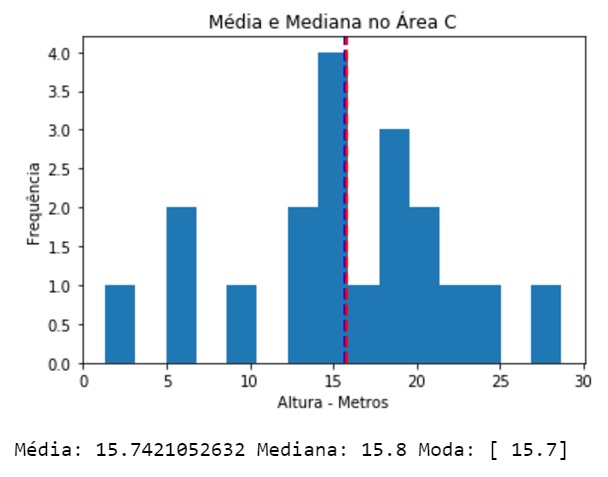
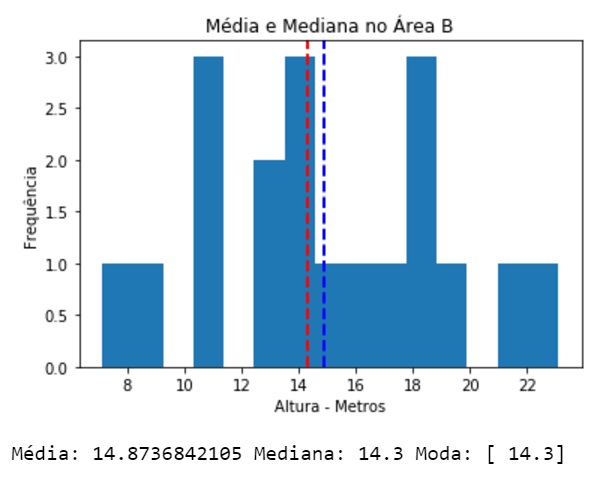
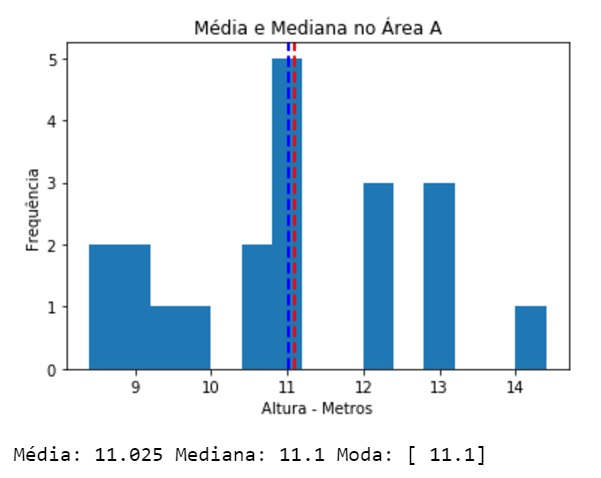


* 1. Encontre os valores (aproximados) para a mediana, os quartis, máximo e mínimo. (b)Discuta a variabilidade do tempo de recuperação em cada técnica.

(c)Se você é otimista, qual técnica escolheria?

1. Utilize as técnicas de análise descritiva e exploratória para discutir a possível relação existente entre a*resistência do material* e o *tipo de cola*. Tais atributos estão presentes no arquivo “resistencia.txt".

1 – A)



1 – B)

1 – C)

2 – A)

2 – B)

2 – C)

2 – D)

2 – E)

3 – A)

3 – B)

3 – C)

4)