

Universidade Regional de Blumenau – FURB

Especialização em *Data Science*

Análise Multivariada de Dados

Profa. Dra. Adriana Kroenke

APRESENTAÇÃO

- ✓ Quem sou;
- ✓ Formação;
- ✓ O que faço;



ANÁLISE MULTIVARIADA DE DADOS

- ✓ Importância na tomada de decisão;
- ✓ Quantidade de dados;
- ✓ Qual modelo mais apropriado;
- ✓ Desenvolvimento computacional;

ANÁLISE MULTIVARIADA DE DADOS

✓ Vamos supor uma pesquisa relacionada ao comportamento/fidelidade do consumidor de uma rede de lojas...

A análise multivariada pode gerar informações relevantes para desenvolver estratégias na organização.

MODELOS DE ANÁLISE MULTIVARIADA DE DADOS

A análise multivariada trata de modelos com variáveis aleatórias e inter-relacionadas com objetivo de reduzir dados, selecionar ou agrupar, investigar dependência entre variáveis, fazer previsões, além de construir e verificar hipóteses.

(FÁVERO et. al., 2009)

MODELOS DE ANÁLISE MULTIVARIADA DE DADOS

- Inadimplência do consumidor;
- Fidelidade do cliente;
- Comportamento de emprego;
- Desempenho de uma organização;

MODELOS DE ANÁLISE MULTIVARIADA DE DADOS

Técnicas de Interdependência:

- ACP
- AF

Técnicas de Dependência:

- AD
- RL

ANÁLISE EXPLORATÓRIA DE DADOS

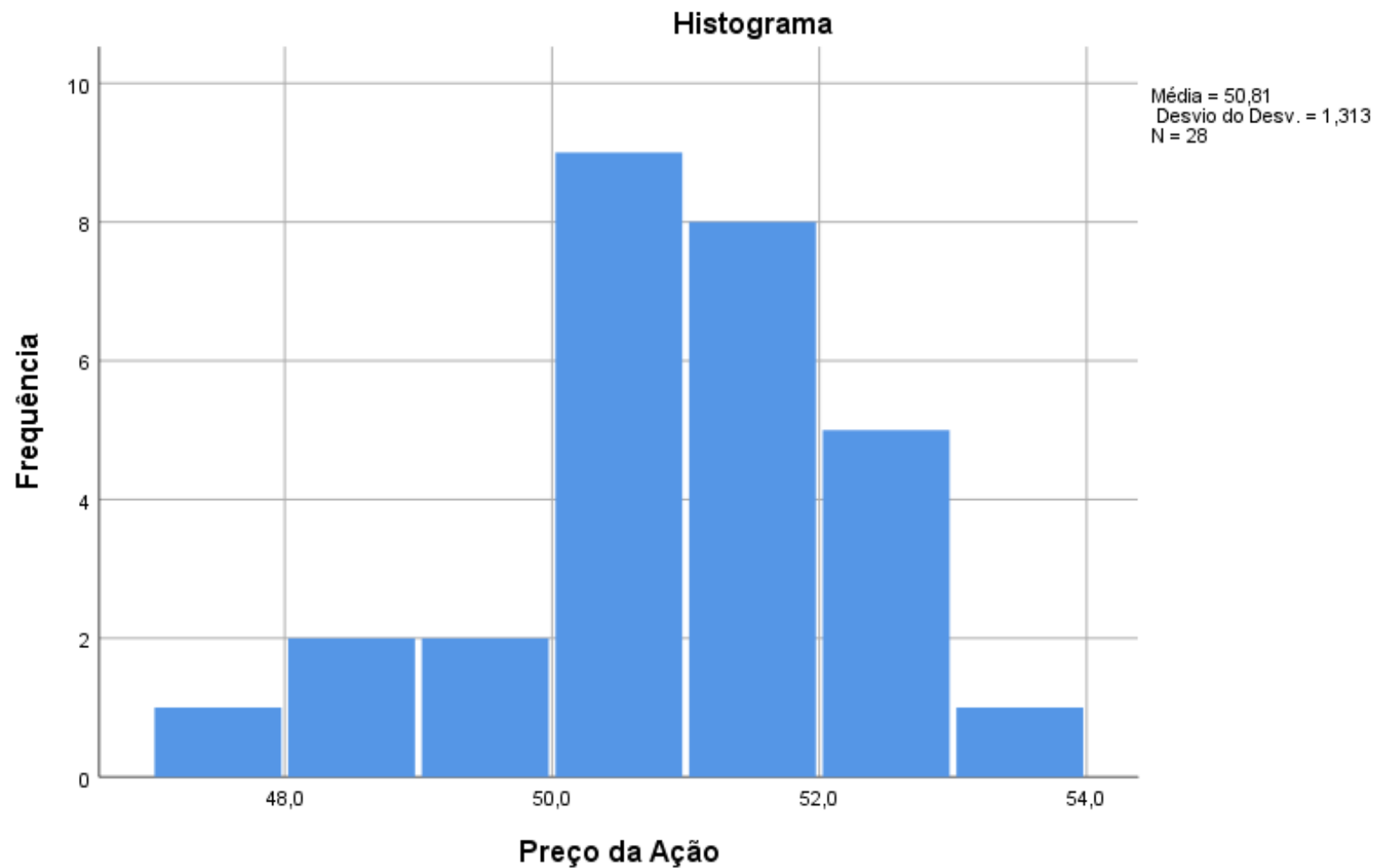
- Análise prévia;
- Preparo dos dados;
- Estatística descritiva;
- Exame gráfico;
- *Outliers*;
- Dados ausentes;

ANÁLISE EXPLORATÓRIA DE DADOS

- HISTOGRAMA

Variável quantitativa e valores representados no eixo horizontal. O eixo vertical representa a frequência da distribuição.

ANÁLISE EXPLORATÓRIA DE DADOS



ANÁLISE EXPLORATÓRIA DE DADOS

- Gráfico Ramo-e-Folhas

Apresenta as mesmas informações do histograma mantendo os dados originais.

ANÁLISE EXPLORATÓRIA DE DADOS

Preço da Ação Gráfico de Ramos e Folhas

Frequência Raiz & Folha

3,00 Extremes ($= < 48,5$)

,00 49 .

2,00 49 . 57

3,00 50 . 034

6,00 50 . 555667

→ 7,00 51 . 0011344

1,00 51 . 6

3,00 52 . 124

2,00 52 . 58

1,00 53 . 2

Largura do ramo: 1,0

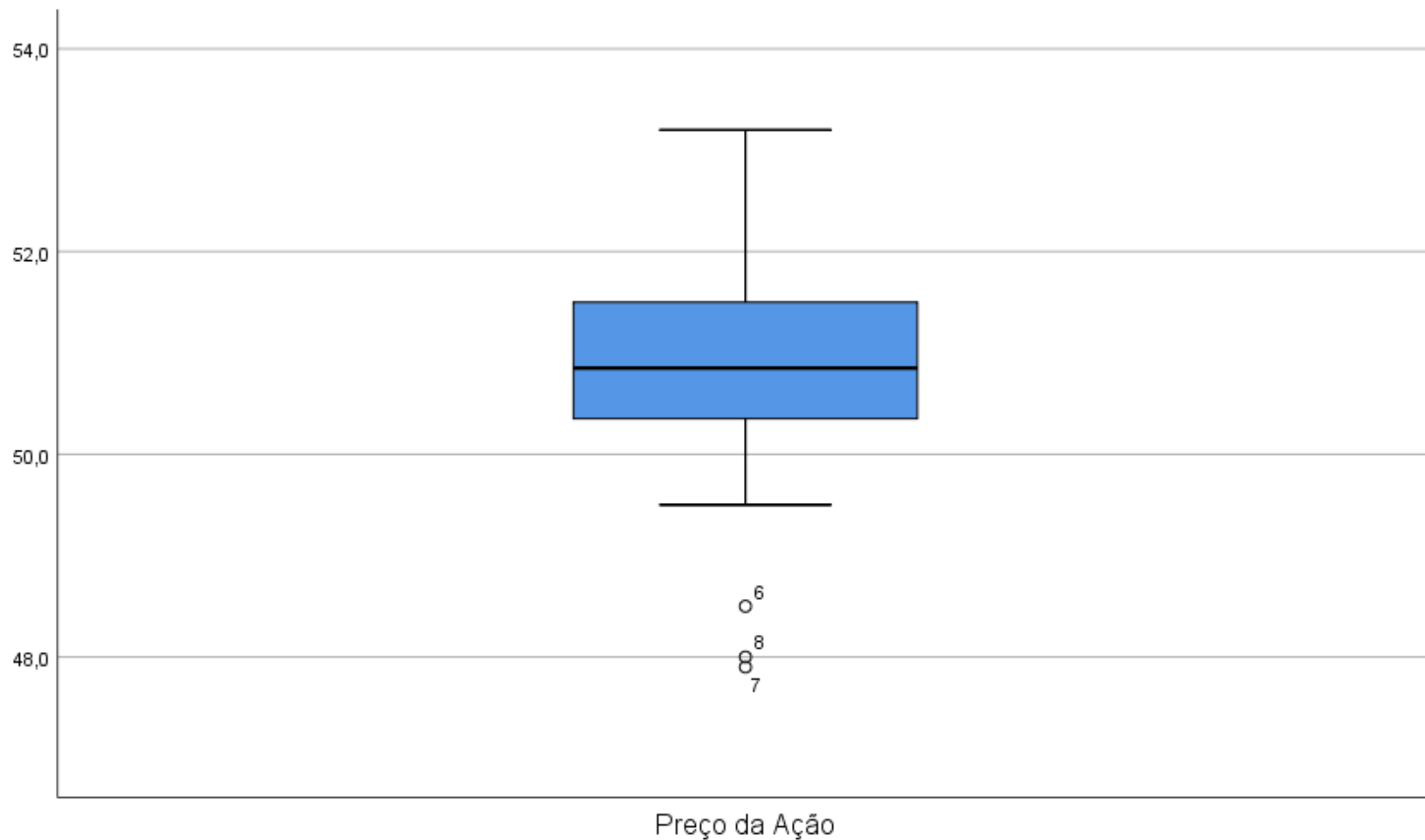
Cada folha: 1 caso(s)

ANÁLISE EXPLORATÓRIA DE DADOS

- *BOX-PLOT* ou Diagrama de extremos e quartis

Valores máximo e mínimo, 1º quartil, 2º quartil ou mediana, 3º quartil, *outliers* e extremos.

ANÁLISE EXPLORATÓRIA DE DADOS

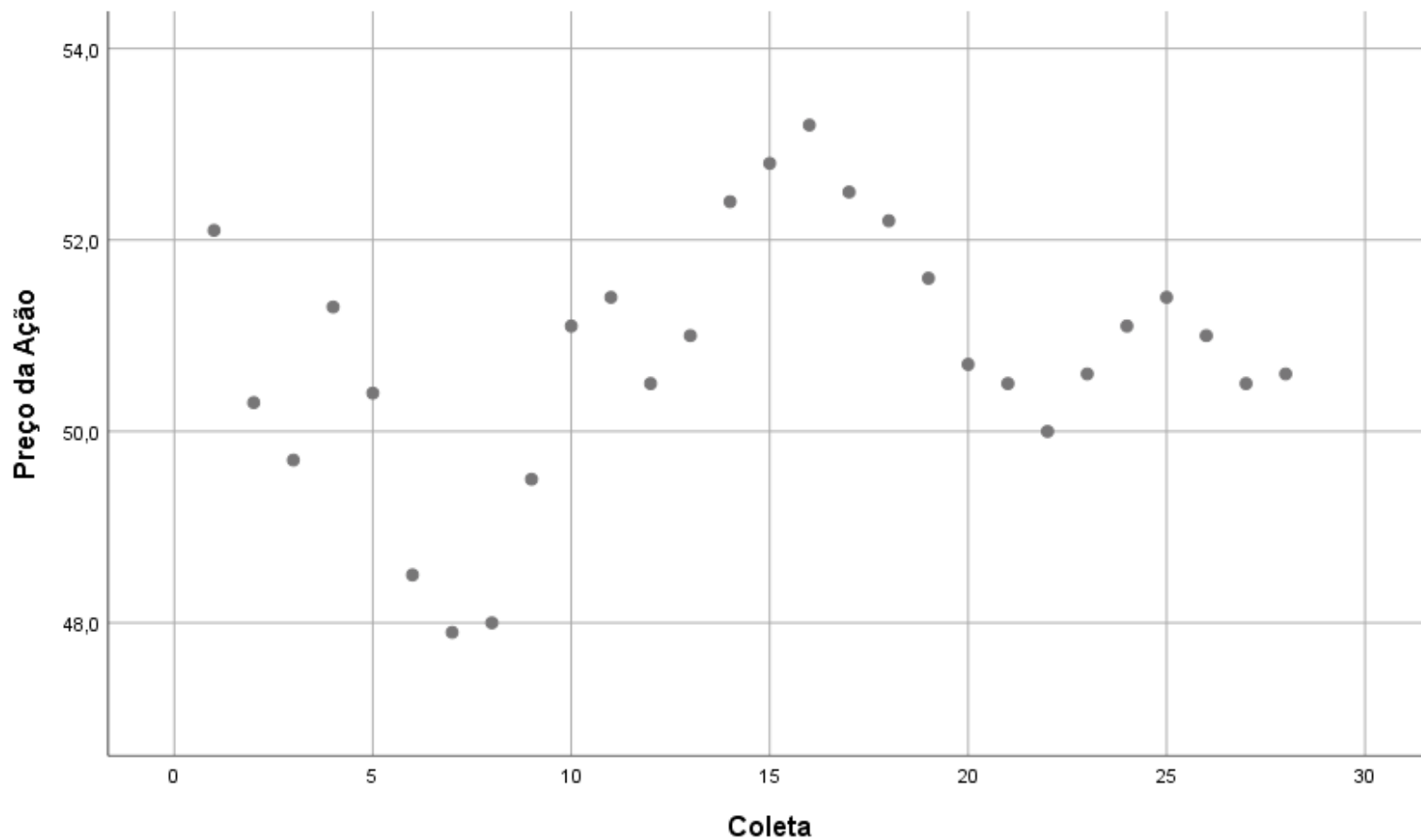


ANÁLISE EXPLORATÓRIA DE DADOS

- Gráfico de Dispersão

Busca verificar a relação entre variáveis

ANÁLISE EXPLORATÓRIA DE DADOS



ANÁLISE EXPLORATÓRIA DE DADOS

- Vamos supor o preço de uma determinada ação em um determinado dia. Coleta entre 10h e 17h obtendo 28 informações (a cada 15 mim).

52,1	50,3	49,7	51,3	50,4	48,5	47,9	48,0	49,5	51,1
51,4	50,5	51,0	52,4	52,8	53,2	52,5	52,2	51,6	50,7
50,5	50,0	50,6	51,1	51,4	51,0	50,5	50,6		

ANÁLISE EXPLORATÓRIA DE DADOS

52,1	50,3	49,7	51,3	50,4	48,5	47,9	48,0	49,5	51,1
51,4	50,5	51,0	52,4	52,8	53,2	52,5	52,2	51,6	50,7
50,5	50,0	50,6	51,1	51,4	51,0	50,5	50,6		

No SPSS...

ANÁLISE EXPLORATÓRIA DE DADOS

Exercício 1:

Considere os dados referente ao tempo médio de processamento da cenoura na fase de acabamento (em segundos) para 20 observações.

Determine:

1. Histograma

2. Gráfico de ramo-e-folhas

3. *Box-Plot*

ANÁLISE EXPLORATÓRIA DE DADOS

Dados Ausentes:

Valores “em branco” por vários motivos.

- Eliminação completa;
- Eliminação parcial;
- Substituição pela média;
- Interpolação

ANÁLISE EXPLORATÓRIA DE DADOS

Exemplo 2

-Dados ausentes

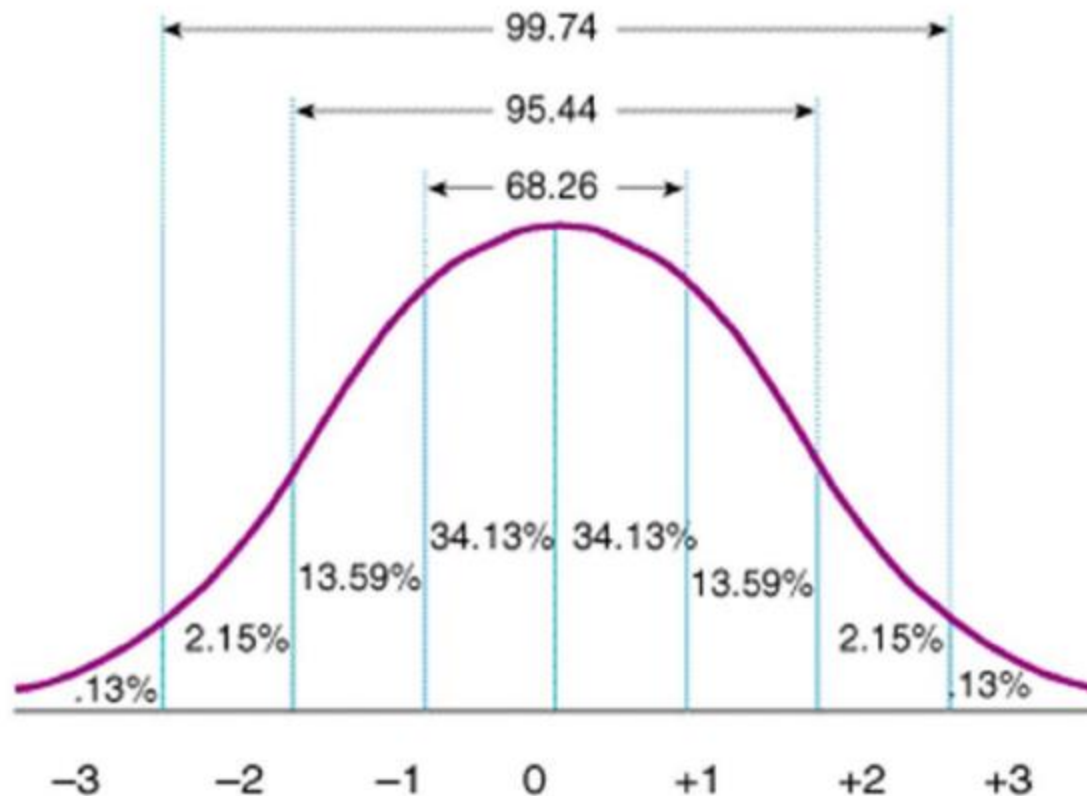
-Box Plot

DISTRIBUIÇÃO NORMAL MULTIVARIADA

Termo “Multivariado” apresenta divergências.

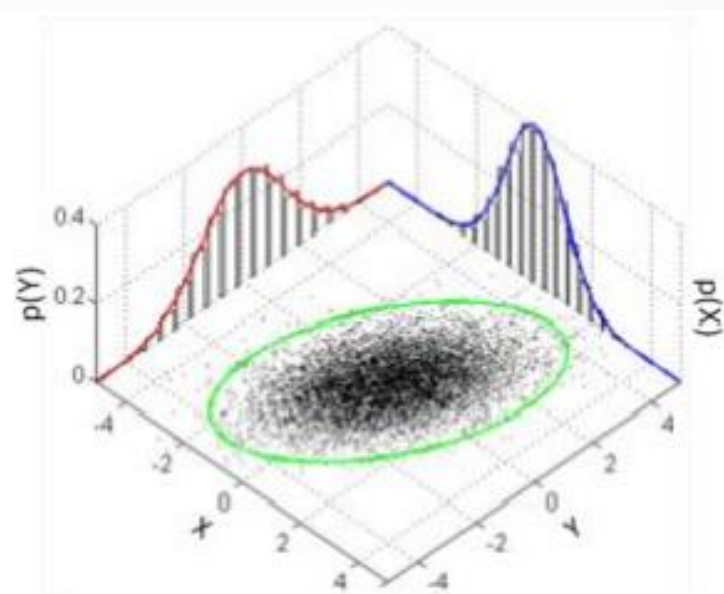
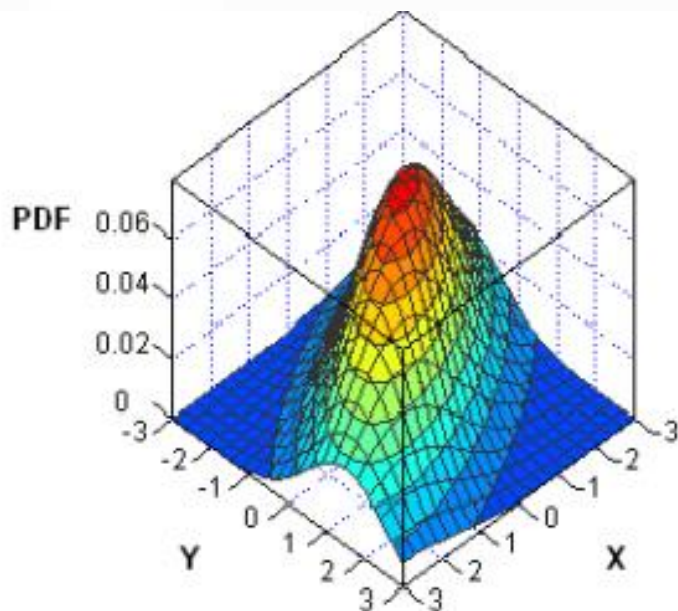
- ✓ Multivariado: analisar o relacionamento entre duas ou mais variáveis.
- ✓ Multivariado: situações nas quais todas as variáveis assumem, por hipótese, distribuição normal multivariada.

DISTRIBUIÇÃO NORMAL UNIVARIADA



Depende de dois parâmetros: média e variância

DISTRIBUIÇÃO NORMAL MULTIVARIADA



Depende de dois parâmetros:

- Vetor de médias de cada variável
- Matriz de covariâncias

DISTRIBUIÇÃO NORMAL MULTIVARIADA

Exercício 1

Verificar Normalidade

Gerar *Box Plot*

Verificar se há *outliers*

REFERÊNCIAS

FÁVERO, et al. **Análise de dados: modelagem multivariada para tomada de decisões.** Rio de Janeiro: Elsevier, 2009.

HAIR, Joseph F. et al. **Análise multivariada de dados.** Bookman, 2009, 688 p.

MAROCO, João. **Análise estatística com utilização do SPSS.** Lisboa: Sílabo, 2007.

MINGOTI, Sueli Aparecida. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada.** Belo Horizonte: UFMG, 2007.

PEREIRA, Alexandre. **Guia prático de utilização: análise de dados para ciências sociais e psicologia.** 5. ed. Lisboa: Sílabo, 2004.