

# Controle de Qualidade I

Ferramenta da Qualidade - Análise Multivariada

## Análise Multivariada

### ***Definição (O que?)***

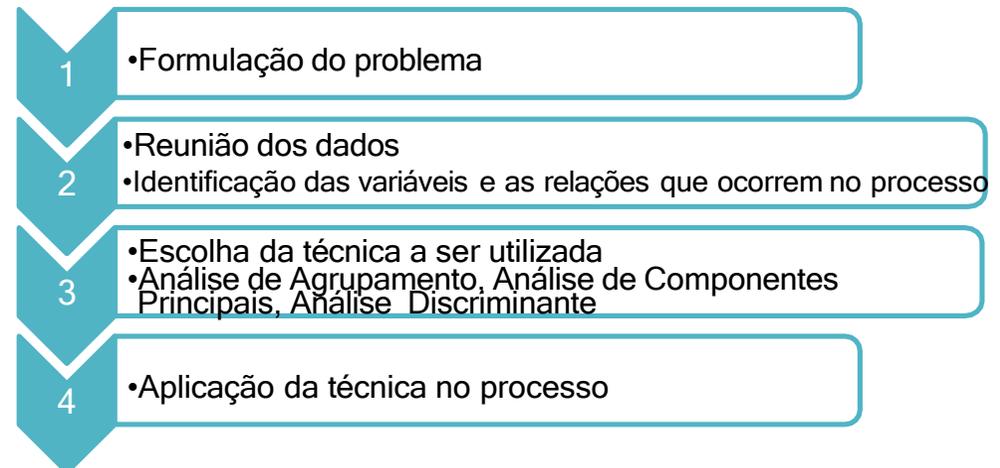
Análise Multivariada é uma ferramenta estatística no qual conseguimos avaliar diversas variáveis do conjunto de dados obtidos de uma vez só com o auxílio de técnicas e métodos que possuem finalidades diversas entre si.

É importante ressaltar que trata-se de uma análise voltada para gerar hipóteses, nem sempre gerando confirmações.

### ***Utilidade (Para que?)***

- Análise de mercado
- Garantia do controle de qualidade
- Otimização e controle do processo
- Pesquisa e desenvolvimento

### ***Procedimento (Como?)***



## Análise Multivariada

### *Procedimento(Como?)*

#### *Principais tipos de mapeamento*



#### **OBSERVAÇÃO**

Na Análise de Agrupamentos podendo subdividir em hierárquico e não hierárquico

## Análise Multivariada

### *Procedimento(Como?)*

#### **Análise de Agrupamentos / Clustering**

- Separa um conjunto heterogêneo em grupos homogêneos, sendo essa seleção definida por um coeficiente de semelhança ou dissimilaridade e um processo de aglomeração.
- Esse coeficiente tende a se basear em uma função de dissimilaridade, garantindo que eles sejam muito similares entre si assim como muito diferentes dos outros.
- No próximo slide temos alguns tipos de métodos de dissimilaridade

***OBS: O grande problema a ser enfrentado nessa análise é a quantidade absurda de possíveis agrupamentos***



# Análise Multivariada

## Procedimento (Como?)

### Análise de Agrupamentos / Clustering Métrica de Minkowski

$$d(i,j) = \sqrt[q]{(|x_{i1} - x_{j1}|^q + |x_{i2} - x_{j2}|^q + \dots + |x_{ip} - x_{jp}|^q)}$$

### Distância Camberra

$$d(x,y) = \sum_{i=1}^p \frac{|x_i - y_i|}{|x_i| + |y_i|}$$

### Distância de Mahalanobis

$$d(x,y) = \sqrt{(x-y)^T \Sigma^{-1} (x-y)}$$

Quanto maior o q, maior sensibilidade a métrica de distâncias maiores

Distância Manhattan: q = 1

Distância Euclidiana: q = 2

Distância Chebyshev: q = ∞

Distância é muito sensível a pequenas variações quando ambas as coordenadas estiverem muito próximas de zero.

Coordenadas cuja distância for a mesma, mas que tiverem módulo diferente apresentarão contribuições distintas

Σ sendo a matriz identidade = Distância Euclidiana

Σ sendo uma matriz diagonal = Distância Euclidiana normalizada

## Análise Multivariada

### Procedimento (Como?)

#### Análise de Agrupamentos / Clustering

#### Métodos para medir as distâncias entre os objetos:

Encadeamento único: Prioriza elementos mais próximos e deixando os mais distantes em segundo plano (menor valor da vizinhança permanece)

Encadeamento médio: Recalculo dos centróides e distâncias entre objetos cada vez que um objeto muda

Encadeamento completo: A distância entre objetos é a distância entre seus pontos mais distantes

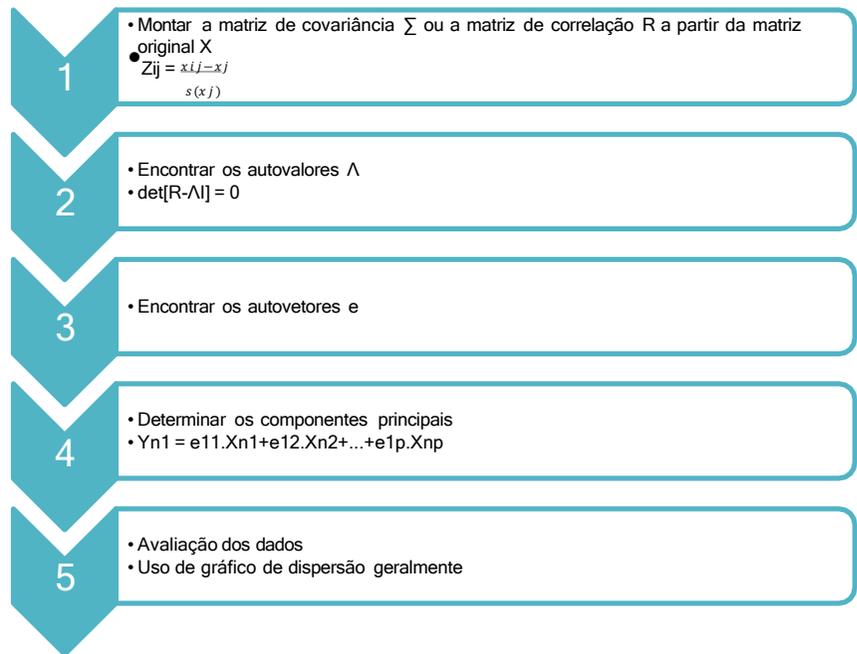
# Análise Multivariada

## Procedimento(Como?)

### Análise de Componentes Principais

Transforma linearmente um conjunto original de variáveis, inicialmente correlacionadas entre si, num conjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original.

Portanto, consiste em transformar um conjunto de variáveis originais em outro conjunto de variáveis de mesma dimensão denominadas de componentes principais



## Análise Multivariada

### Procedimento (Como?)

#### **Análise Fatorial**

Investigar os padrões ou relações latentes para um número grande de variáveis e determina se a informação pode ser resumida a um conjunto menor de fatores, no qual o fator pode ser definido como uma combinação linear das variáveis originais

**OBS: Pode ser exploratória ou confirmatória**

1

- Entrada de dados
- Verificar se é possível a aplicação da análise

2

- Definição do número de fatores a serem extraídos
- Critério do autovalor, do diagrama de inclinação, da porcentagem de variância acumulada,...

3

- Extração dos fatores
- Pode ser pelo método dos componentes principais

4

- Rotação dos fatores
- Atinge uma melhor distinção entre os fatores

## Análise Multivariada

### Exemplo de Aplicação

*Base de dados*

Cientes	Bolo 1	Bolo 2	Bolo 3	Bolo 4
1	20	5	7	49
2	18	9	2	45
3	11	35	15	7
4	10	3	4	26



*Matriz das distâncias*

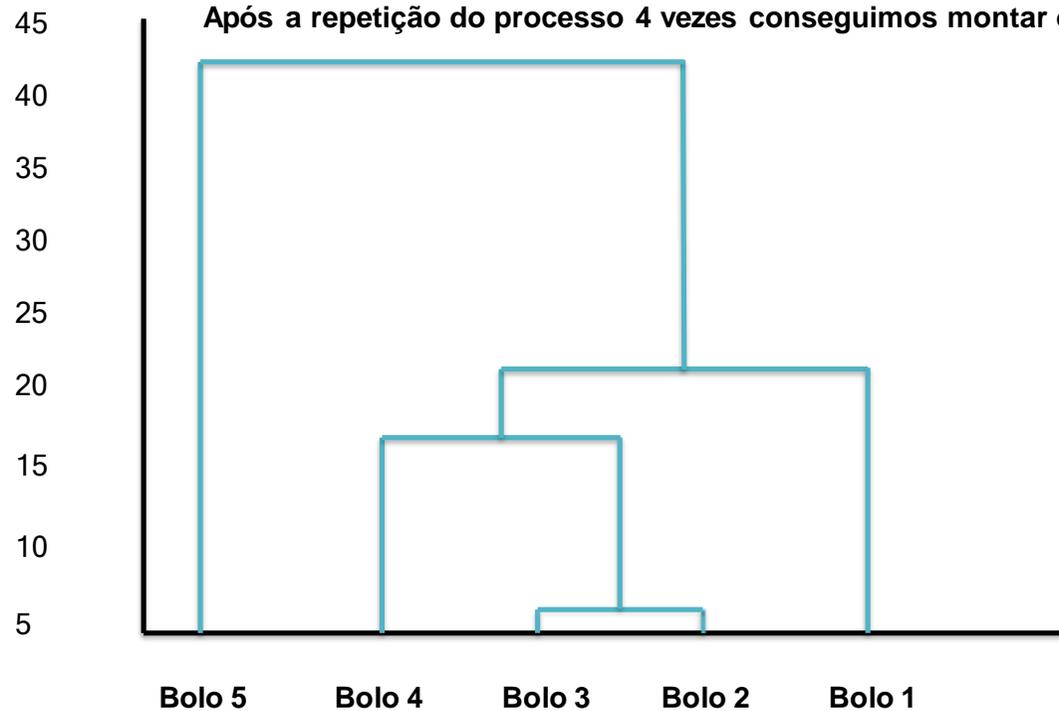
0,0	30,5	22,7	21,8	42,9
-	0,0	8,8	21,3	67,4
-	-	0,0	17,7	59,7
-	-	-	0,0	64,5
-	-	-	-	0,0

8,8 é a menor distância entre 2 variáveis distintas

## Análise Multivariada

### Exemplo de Aplicação

Após a repetição do processo 4 vezes conseguimos montar o seguinte dendograma



Concluimos então que o bolo 5 é heterogêneo em relação às outras  
**,ou seja, é o bolo menos comprado e por isso irá ser retirado do catálogo**

## Análise Multivariada

### ***Situação nominal e áreas de aplicação***

- Pode ser usada em qualquer área, podendo ser ressaltados as ciências exatas, as engenharias, as ciências da terra, a medicina, a psicologia e a administração
- Existem diversos métodos, e a escolha do ideal dependerá da sua necessidade
- A coleta de dados é geralmente colocada em tabelas, para então as informações serem aplicadas no mais adequado tipo de análise
- No exemplo usado vemos que a simplicidade da aplicação da análise de agrupamento hierárquico aglomerativo pode ser desvantajoso, já que os objetos intermediários são facilmente aglomerados em grupos.

### ***Fontes / Referências***

PAYE, Henrique de Sá; MELLO, Jaime Wilson Vargas de; MELO, Stefeson Bezerra de. Métodos de análise multivariada no estabelecimento de valores

de referência de qualidade para elementos-traço em solos. **Revista Brasileira de Ciência do Solo**, v. 36, n. 3, p. 1031-1042, 2012.

LINDEN, Ricardo. Técnicas de agrupamento. **Revista de Sistemas de Informação da**

**FSMA**, v. 4, n. 4, p. 18-36, 2009.] FACHEL, Jandyra Maria Guimarães. Análise fatorial.

1976.

MATOS, Daniel Abud Seabra; RODRIGUES, Erica Castilho. Análise fatorial. 2019.