

## ATIVIDADE\_02

### QUESTÃO 01

A Secretaria Municipal de Saúde realizou um levantamento com 20 participantes de um programa de prevenção de doenças crônicas. Os valores abaixo correspondem à idade (em anos) de cada participante:

48, 52, 41, 60, 55, 33, 70, 49, 46, 58, 65, 44, 39, 53, 61, 47, 42, 68, 56, 50

Com base nesses dados, faça o que se pede:

#### A. CALCULE O COEFICIENTE DE VARIAÇÃO

```
# Dados
idades <- c(48,52,41,60,55,33,70,49,46,58,65,44,39,53,61,47,42,68,56,50)

# Cálculo da média e do desvio padrão
media_id <- mean(idades)
dp_id <- sd(idades) #amostral por padrão

# Cálculo do coeficiente de variação (em %)
cv_id <- (dp_id/media_id)*100

# Imprimindo os resultados
list(Media = round(media_id,2), Desvio_Padrao = round(dp_id,4),Coeficiente_Variacao = round(cv_id,2))

$Media
[1] 51.85

$Desvio_Padrao
[1] 9.8956

$Coeficiente_Variacao
[1] 19.09
```

#### B. DETERMINE A ASSIMETRIA E A CURTOSE DA DISTRIBUIÇÃO ETÁRIA

```
g1 <- e1071::skewness(idades, type=2) #Assimetria
k_base <- e1071::kurtosis(idades, type=2) #curtose de Pearson

#Imprimindo os resultados
list(Assimetria = round(g1, 4), Curtose_base3 = round(k_base + 3, 4))

$Assimetria
[1] 0.1311

$Curtose_base3
[1] 2.4916
```

## **Especialização em Análise de Dados como Método de Apoio às Políticas Públicas**

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

---

### **C. INTERPRETE OS RESULTADOS**

#### **Resultados**

- \* Média  $\approx 51,85$  anos
- \* Desvio Padrão  $\approx 9,90$  anos
- \* Coeficiente de variação  $\approx 19,08\%$
- \* Assimetria =  $0,1311 \rightarrow$  Pequena assimetria à direita
- \* Curtose =  $2,4916 \rightarrow$  platicúrtica, mais achatada

#### **Interpretação**

A média mostra que a população analisada não é composta por pessoas jovens, mas sim por adultos na faixa dos 50 anos. O Coeficiente de variação em torno de 19%, indica que a dispersão das idades é moderada, ou seja, há variações entre os participantes, mas a concentração está em torno da faixa dos 50 anos. A assimetria levemente positiva, sugere que existem alguns idosos entre os entrevistados. Por fim, a curtose platicúrtica revela que as idades são espalhadas em torno da média, não tendo concentrações em uma única idade.

### **D. SUGESTÃO DE POLÍTICA PÚBLICA EM SAÚDE**

Os resultados demonstram que o governo deve concentrar os cuidados com doenças crônicas em adultos mais velhos, que estão entrando na faixa idosa. Os órgãos de saúde podem também fazer campanhas de prevenção com pessoas mais jovens para a longo prazo diminuir o número de adultos e idosos que possuem doenças crônicas.

### **QUESTÃO 02**

O Departamento de Saúde Municipal coletou dados sobre o número de consultas médicas realizadas por mês em três unidades de saúde (A, B e C) durante um ano. Os dados (número de consultas) são apresentados abaixo:

- **Unidade A:** 120, 130, 125, 140, 150, 135, 128, 142, 138, 145, 132, 136
- **Unidade B:** 90, 95, 100, 110, 105, 98, 102, 108, 100, 97, 93, 105
- **Unidade C:** 160, 155, 165, 170, 168, 162, 158, 175, 172, 160, 167, 169

## Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

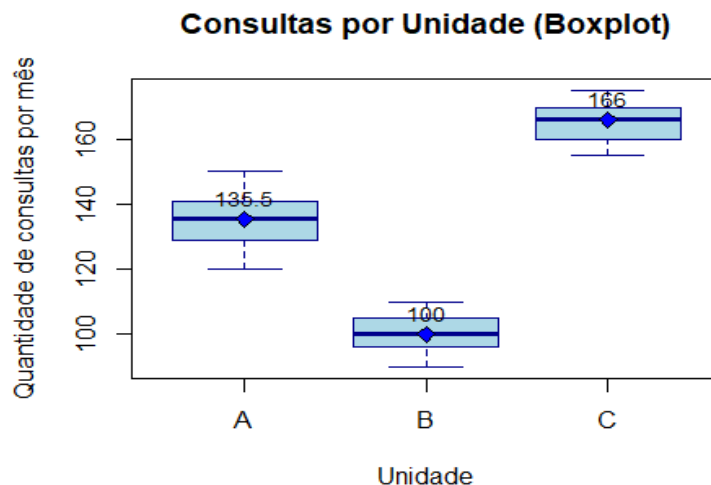
### A. CONSTRUA UM BOXPLOT COMPARATIVO DAS TRÊS UNIDADES DE SAÚDE

```
# Dados
Unidade_A <- c(120,130,125,140,150,135,128,142,138,145,132,136)
Unidade_B <- c( 90, 95,100,110,105, 98,102,108,100, 97, 93,105)
Unidade_C <- c(160,155,165,170,168,162,158,175,172,160,167,169)

df <- data.frame(
  unidade = factor(rep(c("A","B","C"), each = 12), levels = c("A","B","C"))
,
  consultas = c(Unidade_A,Unidade_B,Unidade_C)
)

#Gerando o gráfico
boxplot(consultas ~ unidade, data = df,
        col = "lightblue", border = "blue4",
        main = "Consultas por Unidade (Boxplot)",
        xlab = "Unidade", ylab = "Quantidade de consultas por mês")

# Visualizando as Medianas no gráfico
meds <- tapply(df$consultas, df$unidade, median)
points(x = seq_along(meds), y = meds, pch = 23, bg = "blue", cex = 1.2)
text(x = seq_along(meds), y = meds, labels = round(meds, 1),
     pos = 3, cex = 0.9)
```



**Especialização em Análise de Dados como Método de Apoio às Políticas Públicas**

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

**B. COMPARE A DISTRIBUIÇÃO DE CONSULTAS ENTRE AS UNIDADES E IDENTIFIQUE QUAL APRESENTA MAIOR VARIAÇÃO**

**Resultados**

*\* Unidade A*

- Mediana  $\approx 135,5$  consultas
- Variação moderada dos valores, indo de aproximadamente 125 a aproximadamente 140.

*\* Unidade B*

- Mediana = 100 consultas
- Variação baixa, indo de aproximadamente 100 a aproximadamente 110.
- É a unidade mais regular em termos de consultas.

*\* Unidade C*

- Mediana  $\approx 166$  consultas
- Valores maiores que 160

**Comparação**

A unidade C foi a que mais realizou consultas (mediana igual a 166). Já a unidade B teve o menor número (mediana igual a 100). A unidade A apresentou maior variação (caixa mais larga), enquanto a B foi a mais regular (caixa mais estreita).

**QUESTÃO 03**

Em um estudo para verificar a relação entre asma e incidência de gripe no outono, 150 crianças foram escolhidas ao acaso, dentre aquelas acompanhadas pelo Posto de Saúde de um bairro, nessa época do ano. Os resultados estão no quadro a seguir.

Asma \ Gripe	Sim	Não	Total
Sim	27	34	61
Não	42	47	89
Total	69	81	150

## Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

---

### A. DETERMINE O VALOR DO QUI-QUADRADO DE INDEPENDÊNCIA

```
#Dados - criando a tabela
tab <- base::matrix(c(27,34,
                      42,47),
                    nrow = 2, byrow = TRUE)

dimnames (tab) = list(Asma = c("Sim", "Não"),
                      Gripe = c("Sim", "Não"))

tab

      Gripe
Asma Sim Não
Sim  27  34
Não  42  47

addmargins(tab) #total de linhas/coluna

      Gripe
Asma Sim Não Sum
Sim  27  34  61
Não  42  47  89
Sum  69  81 150

#Calculando o qui-quadrado de independência
quiQuadrante <- chisq.test(tab, correct = FALSE)

#Imprimindo o resultado
list(Qui_quadrante = quiQuadrante)

$Qui_quadrante
Pearson's Chi-squared test

data:  tab
X-squared = 0.12498, df = 1, p-value = 0.7237

#list(Assimetria = round(g1, 4), Curtose_base3 = round(k_base + 3, 4))
```

### B. CALCULE O COEFICIENTE DE CONTINGÊNCIA E INTERPRETE SEU VALOR À LUZ DA FORÇA DA ASSOCIAÇÃO ENTRE AS VARIÁVEIS

```
# Usando o valor do teste anterior
quiQuadrante2 <- as.numeric(quiQuadrante$statistic)

#Calcula o Coeficiente de Contingência
n <- sum(tab)
C <- sqrt(quiQuadrante2/(quiQuadrante2+n))

#Imprime o resultado
list(Coeficiente_Contigencia = C)

$Coeficiente_Contigencia
[1] 0.02885289
```

---

Aluna: Silvana da Rocha Rodrigues - Matrícula: 20252014243

e-mail: [rodrigues.silvanar@ufrj.br](mailto:rodrigues.silvanar@ufrj.br) – GitHub - [https://github.com/RodriguesSR/especializacao-analise-dados/tree/main/AEVD/Atividade\\_02](https://github.com/RodriguesSR/especializacao-analise-dados/tree/main/AEVD/Atividade_02)

## Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

### *Interpretação do valor à luz da força da associação entre as variáveis*

#### Resultados

\* Estatística do Qui-quadrado  $\approx 0,125$

\*  $p$ -valor = 0,7237

\* Coeficiente de Contigência  $\approx 0,0289$

#### Interpretação

Como o valor do coeficiente foi de aproximadamente 0,029, o que indica associação muito fraca. Como o  $p$ -valor ( $\approx 0,72$ ) é alto, não há evidências de associação entre ter asma e gripe, logo as variáveis podem ser tratadas como independentes.

## QUESTÃO 04

Os dados a seguir correspondem à variável renda familiar e gasto com alimentação (em unidades monetárias) para uma amostra de 25 famílias.

X (Renda)	Y (Gasto)
3	1,5
5	2
10	6
10	7
20	10
20	12
20	15
30	8
40	10
50	20
60	20
70	25
70	30
80	25
100	40
100	35
100	40
120	30
120	40
140	40
150	50
180	40
180	50
200	60
200	50

## Especialização em Análise de Dados como Método de Apoio às Políticas Públicas

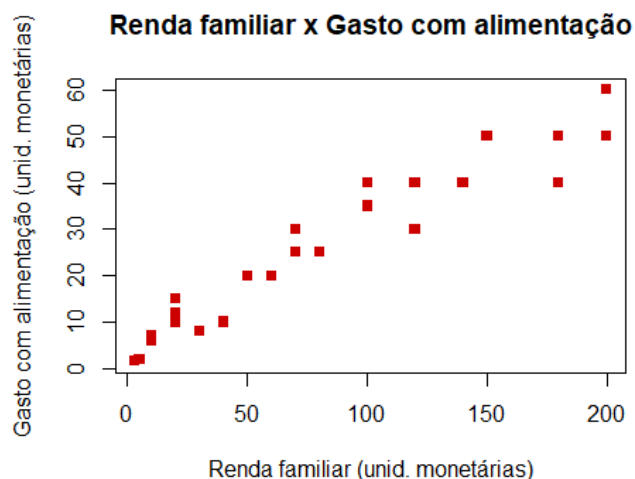
Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

- A. CONSTRUA UM GRÁFICO DE DISPERSÃO COM OS DADOS APRESENTADOS. QUAL O TIPO DE ASSOCIAÇÃO (POSITIVA, NEGATIVA OU INEXISTENTE) PARECE ESTAR PRESENTE ENTRE AS VARIÁVEIS?

```
#Dados
X <- c(3,5,10,10,20,20,20,30,40,50,60,70,70,80,100,100,100,120,120,140,150,180,180,200,200)
Y <- c(1.5,2,6,7,10,12,15,8,10,20,20,25,30,25,40,35,40,30,40,40,50,40,50,60,50)
i <- seq_along(X) #a mais

#Dispersão
plot(X, Y,
     pch = 15, col = "red3",
     main = "Renda familiar x Gasto com alimentação",
     xlab = "Renda familiar (unid. monetárias)",
     ylab = "Gasto com alimentação (unid. monetárias)")
```



- B. CALCULE O COEFICIENTE DE CORRELAÇÃO LINEAR DE PEARSON E INTERPRETE O VALOR OBTIDO

```
# Calculando o coeficiente de correlação linear de Person
r <- cor(X,Y,method = "pearson")

list(Coeficiente_Pearson = r)

$Coeficiente_Pearson
[1] 0.9540509
```

### **Interpretação**

Como o valor do coeficiente de correlação linear de Pearson foi de aproximadamente 0,95, indicando associação positiva muito forte. Ou seja, quanto maior a renda, maior o gasto com alimentação.

**Especialização em Análise de Dados como Método de Apoio às Políticas Públicas**

Disciplina: Análise Exploratória e Visualização de Dados

Professor: Marcelo Montillo Provenza

---

**C. DISCUTA BREVEMENTE DE QUE FORMA OS RESULTADOS ENCONTRADOS  
PODEM ORIENTAR A FORMULAÇÃO DE POLÍTICAS PÚBLICAS PARA MELHORAR  
RENDA FAMILIAR E GASTO COM ALIMENTAÇÃO**

*O resultado da análise do coeficiente de Pearson mostra que famílias com rendas maiores, tendem a ter um gasto mais elevado com alimentação do que famílias com a renda menor. Esse resultado mostra o impacto na segurança alimentar das famílias de baixa renda, mostrando que políticas de transferência de renda para estas famílias é fundamental para que não haja problemas relacionadas a má alimentação.*

*Vale ressaltar, que alimentos com pouca qualidade nutricional são mais baratos, e no caso de famílias de baixa renda, esta passa a ser uma opção de alimentos a serem consumidos. Isso pode gerar problemas graves nutricionais, além de outras comorbidades relacionadas a alimentação, impactando no Sistema Único de Saúde que em algum momento terá que atender essas famílias.*