

## Relatório Técnico: Comparação de Grupos Etários

**Autor:** Felipe Adriel - Turma: TADS035

**Data:** 19 de abril de 2024

---

### 1. Hipóteses

Este estudo buscou comparar as distribuições etárias de dois grupos para avaliar se há diferenças estatisticamente significativas. As hipóteses testadas foram:

- **H (Hipótese Nula):** Não há diferença significativa entre as distribuições etárias dos grupos.
  - **H (Hipótese Alternativa):** O Grupo 2 apresenta valores etários significativamente maiores que o Grupo 1.
- 

### 2. Materiais e Métodos

**2.1 Dados** Foram analisados dois grupos com 10 observações cada:

- **Grupo 1:** Idades entre 12 e 40 anos.
- **Grupo 2:** Idades entre 15 e 43 anos.

#### 2.2 Fluxo de Análise

1. **Estatísticas Descritivas:** Cálculo de média, variância e intervalo interquartil (IQR) para caracterizar os dados.
2. **Teste de Normalidade (Shapiro-Wilk):** Verificação da normalidade dos dados com nível de significância  $= 0,05$ .
3. **Teste Comparativo:**
  - **Teste t de Student** para dados normais.
  - **Teste U de Mann-Whitney** para dados não normais.
4. **Visualização:** Histograma comparativo e diagrama de caixa (boxplot) para análise gráfica.

#### 2.3 Ferramentas

- **Python 3.10** com as bibliotecas:
    - **SciPy 1.10.0** para testes estatísticos.
    - **pandas 1.5.3** para manipulação de dados.
    - **Matplotlib 3.7.1** para visualização.
-

### 3. Resultados

#### 3.1 Estatísticas Descritivas

Grupo	Média	Variância	IQR
Grupo 1	24,5	76,06	13
Grupo 2	26,4	63,38	10

**Interpretação:**

- O Grupo 2 apresentou média ligeiramente superior ao Grupo 1.
- A variância do Grupo 1 foi maior, indicando maior dispersão nos dados.

#### 3.2 Teste de Normalidade (Shapiro-Wilk)

Grupo	Estatística W	p-valor	Conclusão
Grupo 1	0,9626	0,8258	Distribuição normal
Grupo 2	0,9604	0,7032	Distribuição normal

**Interpretação:**

Ambos os grupos seguiram distribuição normal ( $p > 0,05$ ), validando o uso do **Teste t de Student**.

#### 3.3 Teste t de Student

Estatística t	p-valor	Conclusão
-1,02	0,317	Não há diferença significativa ( $p > 0,05$ )

**Interpretação:**

Não houve evidências para rejeitar  $H_0$ . A diferença entre as médias dos grupos **não é estatisticamente significativa**.

#### 3.4 Visualização

- **Histograma:** Mostrou distribuição similar entre os grupos, com o Grupo 2 ligeiramente deslocado para idades mais altas.
- **Boxplot:** Confirmou a maior dispersão no Grupo 1 e a ausência de outliers significativos.

#### 4. Referências

1. **VIRTANEN, P. et al.** SciPy 1.0: Fundamental Algorithms for Scientific Computing. *Nature Methods*, v. 17, p. 261-272, 2020.
  2. **McKINNEY, W.** Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 2010.
  3. **HUNTER, J. D.** Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, v. 9, n. 3, p. 90-95, 2007.
- 

#### 5. Anexo

- **Repositório do Projeto:** GitHub.
  - **Script Completo:** Disponível no arquivo `eletiva-analise.py` anexo.
- 

#### Notas Adicionais:

- O relatório foi gerado em Markdown para facilitar a conversão para PDF/DOCX.
- As análises gráficas estão salvas em alta resolução no arquivo `comparacao_grupos.tiff`.
- O desafio obrigatório (análise de vendas) está implementado em `vendas-analise.py`, com resultados no arquivo `vendas.xlsx`.