

# Relatório de Análise Estatística e Testes de Hipóteses

## Proficiência em Análise de Dados com Python

### 1. Introdução

Este relatório sumariza as principais análises estatísticas conduzidas utilizando a linguagem de programação **Python**. O objetivo é demonstrar a aplicação de testes de hipóteses estatísticas para extrair conclusões significativas a partir de diferentes conjuntos de dados, cobrindo cenários de comparação de proporções (Teste A/B) e comparação de médias (Teste t e ANOVA).

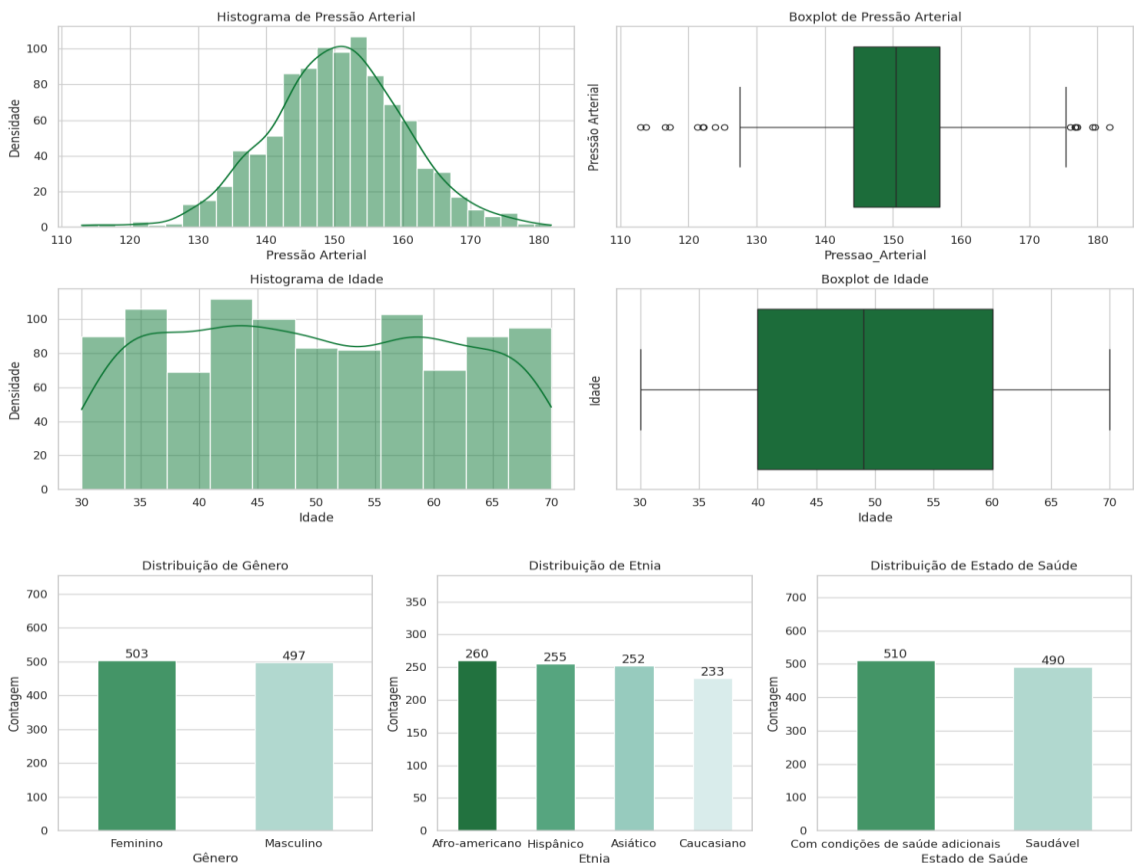
### 2. Metodologia e Ferramentas

A análise foi realizada em ambiente Python, utilizando bibliotecas especializadas que são padrão na ciência de dados e estatística, garantindo robustez e replicabilidade do processo:

- Pandas e NumPy:** Para manipulação, limpeza e organização dos dados.
- Matplotlib e Seaborn:** Para visualização e inspeção exploratória (não detalhado, mas importado).
- SciPy.stats:** A biblioteca fundamental para a execução de todos os testes de hipóteses estatísticas (Teste T, ANOVA, Qui-Quadrado, K-S e Levene).

### 3. Análises e Resultados

As análises foram segmentadas em quatro estudos de caso principais, todos utilizando um nível de significância ( $\alpha$ ) de **0.05**.



---

### Estudo de Caso I: Análise de Teste A/B (Proporções)

**Objetivo:** Determinar se a Versão B de uma página web apresenta uma Taxa de Conversão (TC) significativamente diferente ou superior à Versão A.

Versão	Visitantes (n)	Conversões	Taxa de Conversão (TC)
A	1000	115	11.5%
B	1000	132	13.2%

- **Teste Aplicado:** Teste Qui-Quadrado de Independência (chi2\_contingency).
- **Hipótese Nula (H0):** Não há diferença significativa entre as taxas de conversão (TC\_A = TC\_B).
- **Resultado:**
  - Estatística Qui-Quadrado: 1.68
  - **Valor-p: 0.193**
- **Conclusão:** Como o Valor-p (0.193) é maior que  $\alpha$  (0.05), a hipótese nula **não foi rejeitada**. Não há evidências estatísticas suficientes para afirmar que a Versão B é significativamente melhor que a Versão A, apesar da diferença observada na amostra (13.2% vs. 11.5%).

---

### Estudo de Caso II: Teste T de Uma Amostra (Média)

**Objetivo:** Verificar se a média da pressão arterial de uma amostra é significativamente diferente do valor populacional de referência de 120 mmHg.

- **Amostra:** 30 observações de Pressão Arterial com Média = 122.5.
- **Teste Aplicado:** Teste T de uma Amostra (ttest\_1samp).
- **Hipótese Nula (H0):** A média da amostra é igual a 120 mmHg.
- **Resultado:**
  - Estatística T: 2.131
  - **Valor-p: 0.041**
- **Conclusão:** Como o Valor-p (0.041) é menor que  $\alpha$  (0.05), a hipótese nula **foi rejeitada**. A amostra analisada apresenta uma média de pressão arterial significativamente diferente do valor de referência de 120 mmHg.

---

### Estudo de Caso III: Teste T de Duas Amostras Independentes (Medicamento vs. Placebo)

**Objetivo:** Comparar se o grupo que recebeu o medicamento novo apresenta uma pressão arterial média diferente do grupo que recebeu placebo.

- **Pré-requisitos Verificados:**
  - **Normalidade (Teste K-S):** A distribuição de ambos os grupos (Medicamento e Placebo) segue uma distribuição normal (p-valores de 0.536 e 0.999, respectivamente).
  - **Homogeneidade das Variâncias (Teste de Levene):** As variâncias dos grupos são consideradas iguais (p-valor de 0.505).
- **Teste Aplicado: Teste T de Duas Amostras Independentes (ttest\_ind)** com variâncias iguais.
- **Hipótese Nula (H0):** Não há diferença nas médias de pressão arterial entre os grupos.
- **Resultado:**
  - Estatística T: -2.396
  - **Valor-p: 0.02**
- **Conclusão:** Com o Valor-p (0.02) menor que  $\alpha$  (0.05), a hipótese nula **foi rejeitada**. Existe uma **diferença estatisticamente significativa** na pressão arterial média entre os pacientes que receberam o novo medicamento e os que receberam placebo.

---

### Estudo de Caso IV: Análise de Variância (ANOVA One-Way)

**Objetivo:** Comparar a altura média entre três grupos de atletas (Grupo A, B e C) para determinar se pelo menos um grupo é diferente.

- **Pré-requisito Verificado:**
    - **Homogeneidade das Variâncias (Teste de Levene):** As variâncias dos três grupos são consideradas iguais (p-valor de 0.702).
  - **Teste Aplicado: ANOVA One-Way (f\_oneway).**
  - **Hipótese Nula (H0):** Todas as médias de altura são iguais ( $\mu_A = \mu_B = \mu_C$ ).
  - **Resultado:**
    - Estatística F: 0.174
    - **Valor-p: 0.841**
  - **Conclusão:** Como o Valor-p (0.841) é muito maior que  $\alpha$  (0.05), a hipótese nula **não foi rejeitada**. Não há diferença estatisticamente significativa na altura média entre os três grupos de atletas.
-

#### 4. Resumo das Conclusões

O processo demonstrou a capacidade de:

1. **Comparar Proporções (Teste A/B):** A diferença de 1.7% na Taxa de Conversão observada entre as Versões A (11.5%) e B (13.2%) **não é estatisticamente significativa**, sugerindo que o custo de implementação da Versão B pode não se justificar apenas pelos dados de conversão.
2. **Validar Médias (Teste T):** Foi comprovado estatisticamente que a amostra de pressão arterial **não representa** o valor de 120 mmHg.
3. **Avaliar Intervenções (Teste T Duas Amostras):** O novo medicamento **demonstra ter um efeito estatisticamente significativo** na pressão arterial quando comparado ao placebo.
4. **Comparar Múltiplos Grupos (ANOVA):** Não há diferenças significativas entre as médias de altura dos três grupos de atletas.

Este conjunto de análises, executado integralmente em Python, atesta a proficiência na aplicação de métodos estatísticos rigorosos para a tomada de decisões baseada em dados.

Fontes e conteúdo relacionado