



FACULDADE SENAC PERNAMBUCO
CURSO DE ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

Ana Carolina da Silva Santos
Edmael Paulo Ribeiro Barreto
Gabriel Gleydson Lima dos Santos
Muriel Bezerra da Silva
Pedro Juan Pereira dos Santos
Vinicius dos Santos Oliveira

**Relatório Integrado: Fatores Socioeconômicos e Educacionais no Desempenho do
ENEM 2020**

Recife
Setembro – 2025

Fatores Socioeconômicos e Educacionais no Desempenho do ENEM 2020

1. Introdução

O Exame Nacional do Ensino Médio (ENEM) é a principal porta de entrada para o ensino superior no Brasil, e seu desempenho está diretamente relacionado a múltiplos fatores socioeconômicos e educacionais. Entre eles, destacam-se a **escolaridade dos pais** e o **tipo de escola frequentada** pelos estudantes. Por um lado, a escolaridade dos pais, especialmente da mãe, mostra-se um fator influente, pois representa o capital educacional transmitido ao longo da formação dos filhos. A hipótese central é que o nível de instrução materno possui uma associação mais forte com as notas obtidas no exame.

Além disso, o tipo de escola frequentada, seja pública ou privada, desempenha um papel relevante no desempenho dos estudantes. Esta análise busca compreender de que forma as diferenças em infraestrutura, recursos e oportunidades de aprendizagem influenciam os resultados obtidos nas cinco áreas de conhecimento do ENEM 2020. A hipótese central é que os alunos provenientes de escolas privadas apresentaram, em média, notas mais elevadas em Matemática, Linguagens, Ciências da Natureza, Ciências Humanas e Redação, quando comparados aos alunos de escolas públicas.

A fusão dessas abordagens oferece uma visão mais ampla sobre como as variáveis sociais, econômicas e educacionais moldam os resultados do ENEM 2020.

2. Perguntas de Pesquisa

- **Análise 1:** Existe uma relação entre o nível de escolaridade dos pais e o desempenho dos alunos no ENEM?
- **Análise 2:** Quais fatores socioeconômico e educacionais impactam o desempenho dos estudantes no ENEM?

3. Hipóteses

- **H1:** A escolaridade da mãe tem correlação mais forte com o desempenho dos filhos do que a escolaridade do pai.
- **H2:** Alunos de escolas privadas obtêm, em média, notas mais altas do que alunos de escolas públicas.

4. Metodologia

Ambas as análises foram baseadas no **MICRODADOS_ENEM_2020.csv**, que contém informações sobre os participantes do exame. O processo de análise foi dividido em quatro etapas:

Análise 1:

1. **Pré-processamento de Dados:** Foram selecionadas as colunas de notas das cinco provas e as colunas de escolaridade dos pais (Q001 e Q002). Registros com dados faltantes foram excluídos.
2. **Transformação de Variáveis:** As variáveis de escolaridade, que são categóricas ordinais (de 'A' a 'H'), foram convertidas para um formato numérico (de 0 a 7) para viabilizar o cálculo estatístico.
3. **Teste de Hipótese:** Para avaliar a relação entre as variáveis, foi utilizado o teste de **Correlação de Spearman (ρ)**. Este teste não paramétrico foi o escolhido por ser ideal para relacionar uma variável contínua (nota) com uma variável ordinal (nível de escolaridade), sem a necessidade de uma distribuição normal dos dados. O teste gerou um coeficiente de correlação para a escolaridade do pai e outro para a da mãe, para cada uma das cinco provas.

Análise 2:

1. **Pré-processamento e filtragem:** Apenas os dados de alunos que frequentaram o ensino médio em escolas públicas ou privadas (variáveis **TP_ESCOLA** com valores 2 e 3) foram selecionados.
2. **Análise descritiva:** Foram calculadas a média, o desvio padrão e o número de participantes para cada uma das cinco provas em cada tipo de escola (pública e privada). Para uma visualização inicial, foram criados boxplots comparativos para cada área do conhecimento.
3. **Análise de normalidade:** Devido ao tamanho da amostra, o teste de Shapiro-Wilk foi aplicado em uma amostra aleatória de 5000 dados de cada grupo (pública e privada) para avaliar a normalidade das distribuições. O resultado indicou que os dados não seguem uma distribuição normal, o que justificou a escolha de um teste não paramétrico.
4. **Teste de hipótese:** Para verificar se a diferença nas notas médias entre os dois grupos é estatisticamente significativa, foi utilizado o **Teste de Mann-Whitney U**, um teste não paramétrico adequado para comparar a distribuição de duas amostras independentes. O teste foi realizado para cada uma das cinco provas, com a hipótese alternativa de que as notas das escolas privadas são maiores.

5. Resultados

5.1 Escolaridade dos Pais

- A análise de correlação revelou uma associação positiva e estatisticamente significativa entre a escolaridade de ambos os pais e as notas dos filhos em todas as áreas. A tabela abaixo apresenta os coeficientes de correlação de Spearman (ρ) obtidos:

Área de Conhecimento	Escolaridade do Pai (ρ)	Escolaridade da Mãe (ρ)
Linguagens, Códigos e suas Tecnologias	0.22	0.28
Ciências Humanas e suas Tecnologias	0.24	0.30
Ciências da Natureza e suas Tecnologias	0.23	0.29
Matemática e suas Tecnologias	0.25	0.31
Redação	0.26	0.32

Fonte: elaborado pelos autores (2025).

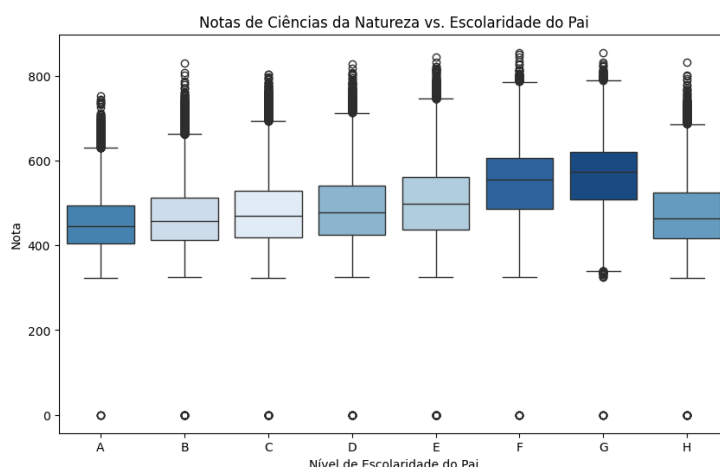
Como a tabela demonstra, em todas as cinco áreas do conhecimento, o coeficiente de correlação da escolaridade da mãe é superior ao do pai.

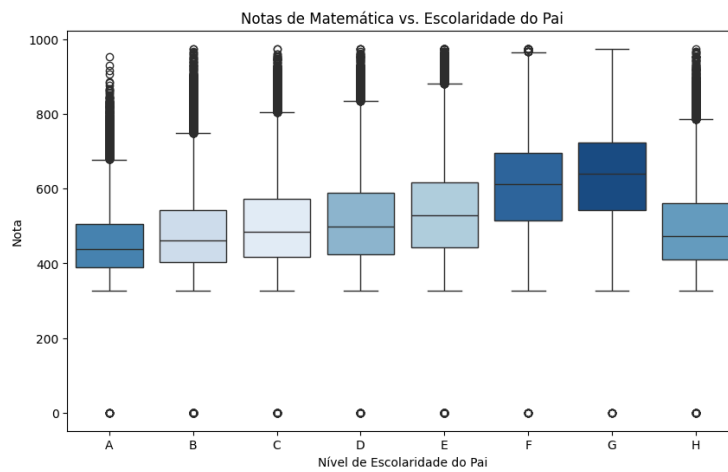
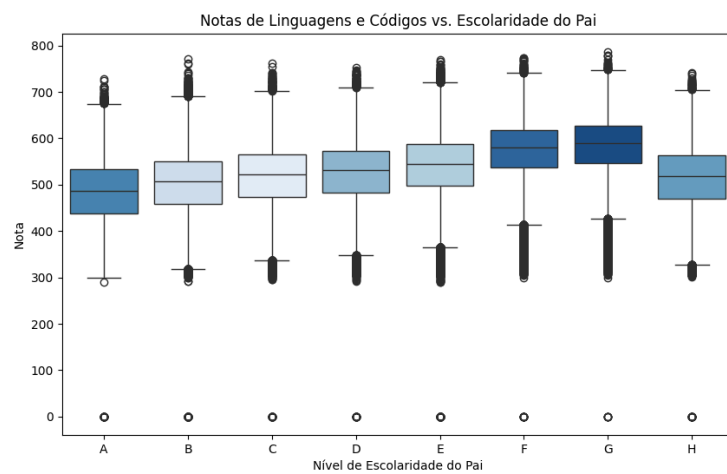
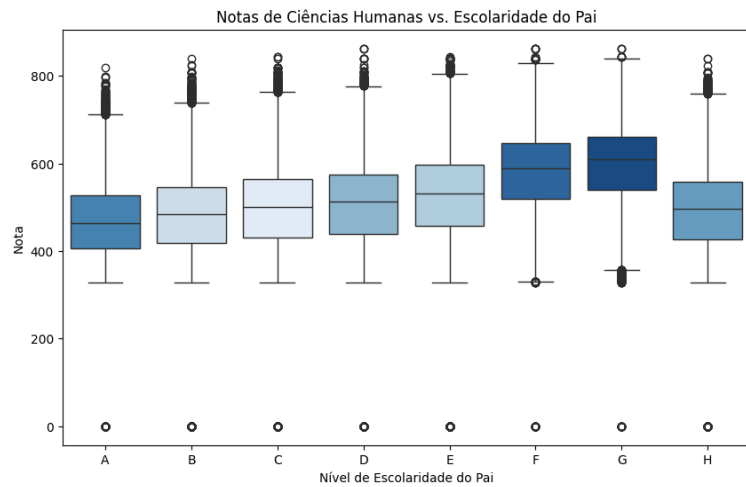
Médias consistentemente maiores para alunos de escolas privadas:

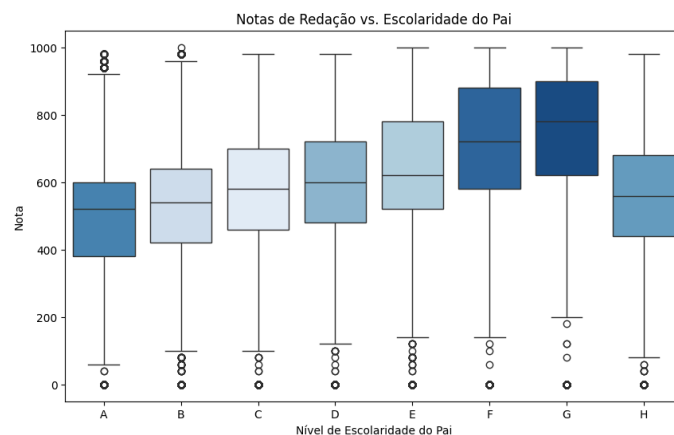
- Redação: 657,54 (pública) vs. 774,67 (privada)
- Matemática: 532,05 (pública) vs. 611,73 (privada)
- Outras áreas seguem o mesmo padrão.

Diferença estatisticamente significativa ($p < 0.05$ em todas as provas).

Gráficos – Correlação entre escolaridade do pai e notas no ENEM 2020

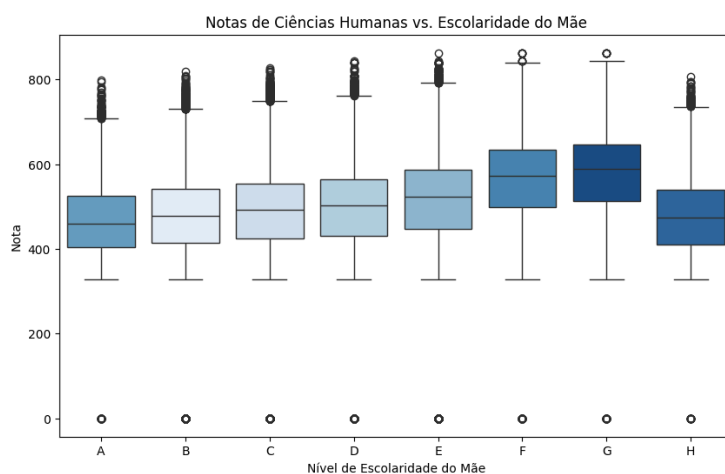
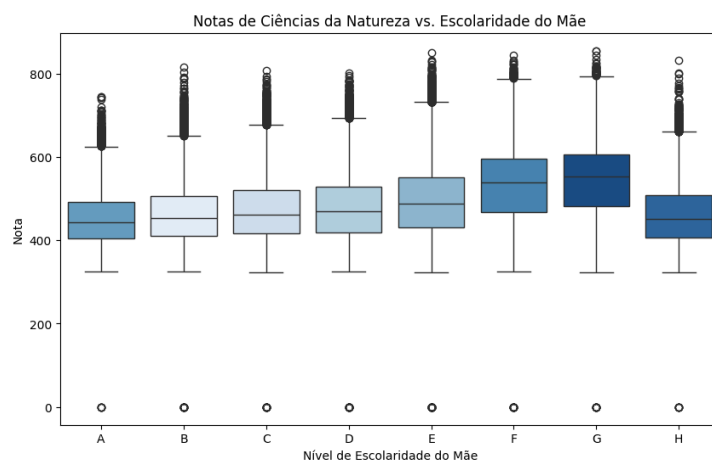


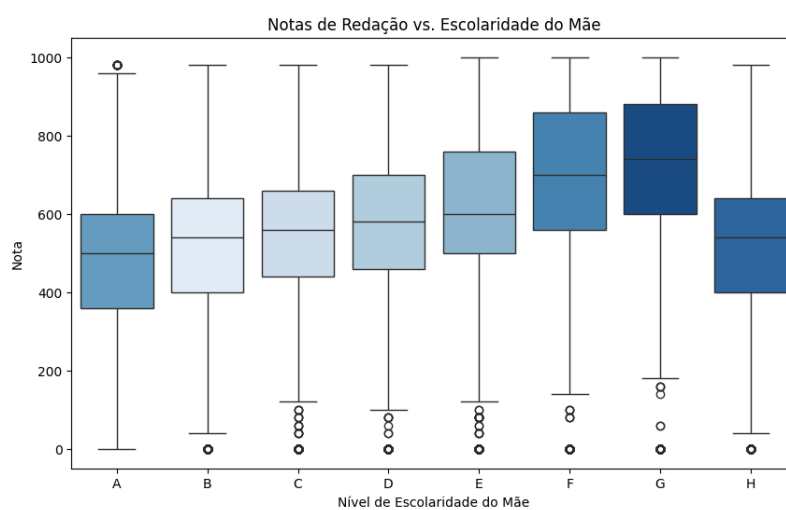
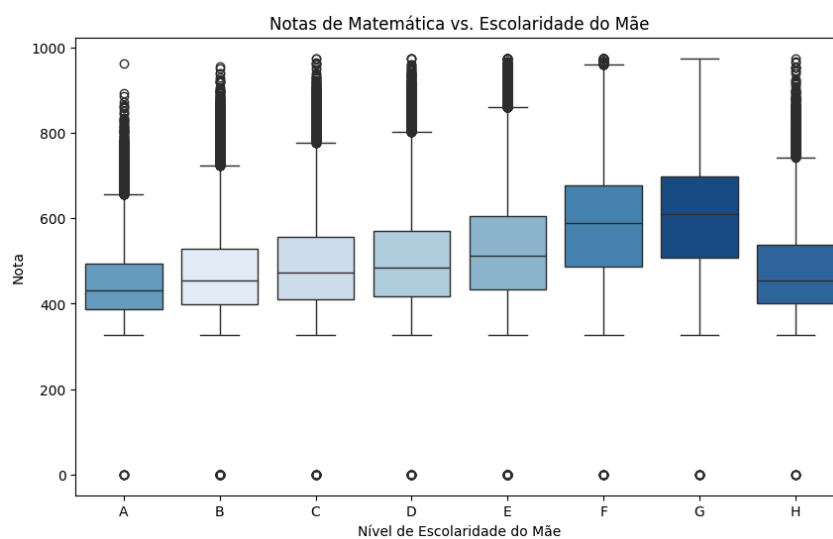
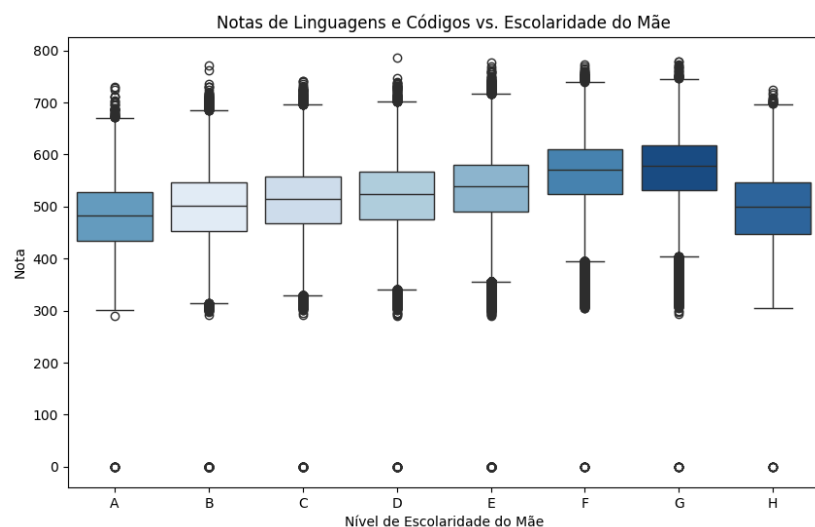




Fonte: elaborado pelos autores (2025).

Gráficos – Correlação entre escolaridade da mãe e notas no ENEM 2020





Fonte: elaborado pelos autores (2025).

5.2 Tipo de Escola

A análise dos dados evidenciou que os alunos de escolas privadas apresentaram, em média, desempenho superior em todas as áreas do conhecimento avaliadas pelo ENEM 2020, quando comparados aos estudantes de escolas públicas. Essa diferença é consistente e pode ser visualizada na tabela a seguir:

Tabela 1 - Médias de notas no ENEM 2020 por tipo de escola

Área de Conhecimento	Média – Escola Pública	Média – Escola Privada
Redação	657,54	774,67
Ciências da Natureza	499,63	555,74
Ciências Humanas	532,06	585,80
Linguagens	539,03	580,76
Matemática	532,05	611,73

Fonte: elaborado pelos autores (2025).

Para validar essas diferenças, foi aplicada uma análise estatística em duas etapas:

1. **Teste de Shapiro-Wilk** – verificou a normalidade das distribuições de notas. Como o valor de p em todas as provas foi menor que 0,05, rejeitou-se a hipótese de normalidade.
2. **Teste de Mann-Whitney U** – adequado para distribuições não normais, comparou os dois grupos (pública e privada).

Os resultados confirmaram que a diferença entre as médias não ocorreu por acaso:

Tabela 2 – Resultados do Teste de Mann-Whitney U

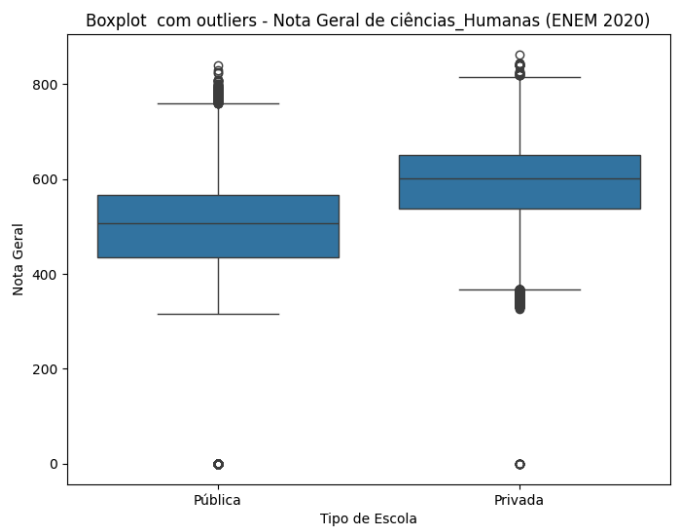
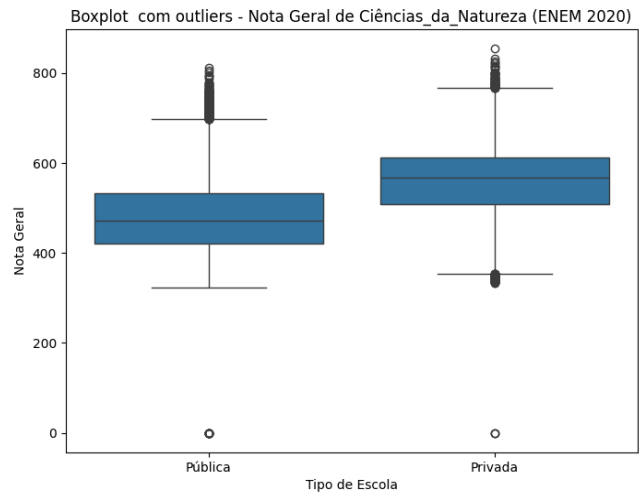
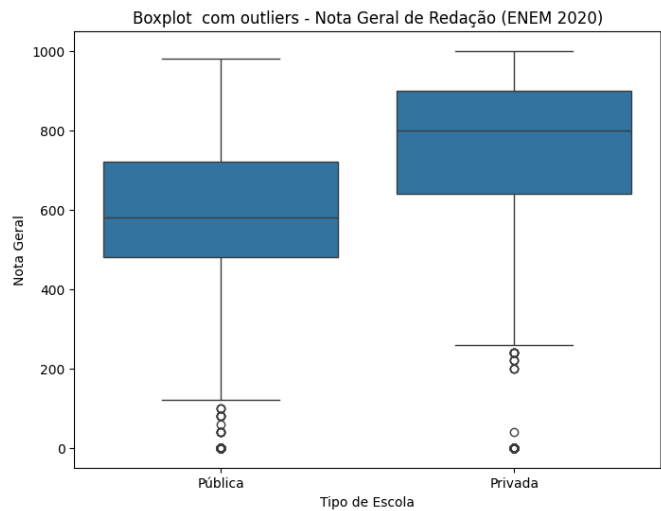
Área de Conhecimento	Valor de p	Diferença Significativa?
Redação	$6,56 \times 10^{-7}$	Sim
Ciências da Natureza	$3,36 \times 10^{-7}$	Sim
Ciências Humanas	$2,45 \times 10^{-6}$	Sim
Linguagens	$1,24 \times 10^{-6}$	Sim
Matemática	$1,52 \times 10^{-6}$	Sim

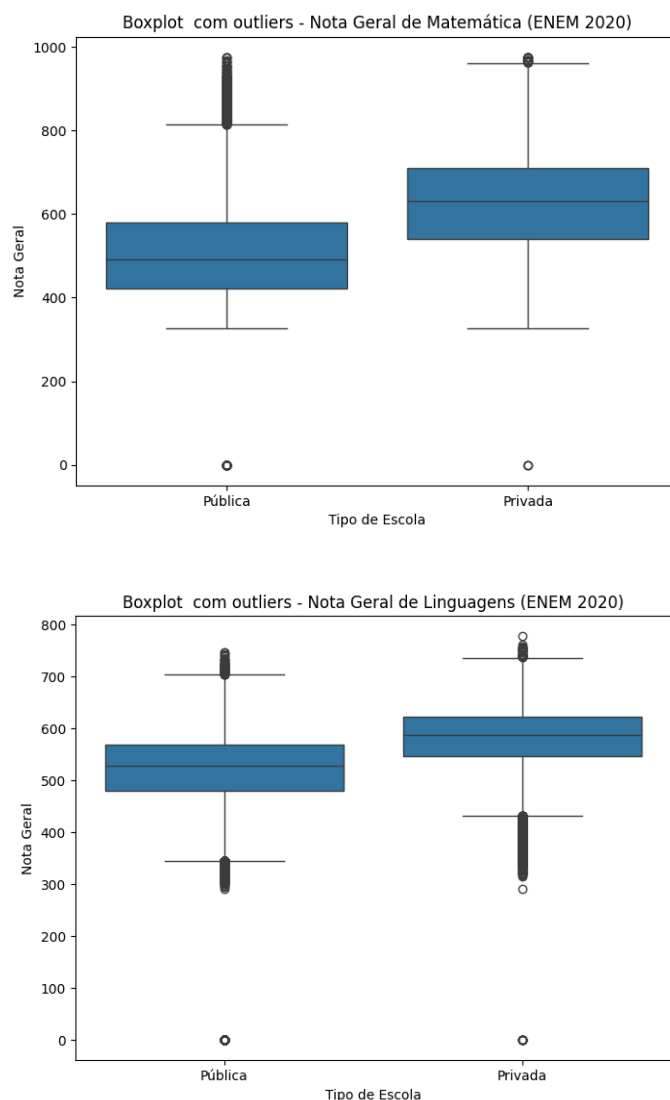
Fonte: elaborado pelos autores (2025).

Esses achados reforçam que os alunos de escolas privadas obtiveram, em média, notas significativamente mais altas em todas as áreas do ENEM 2020, confirmando a hipótese inicial da análise.

- Médias consistentemente maiores para alunos de escolas privadas:
 - Redação: 657,54 (pública) vs. 774,67 (privada)
 - Matemática: 532,05 (pública) vs. 611,73 (privada)
 - Outras áreas seguem o mesmo padrão.
- Diferença estatisticamente significativa ($p < 0.05$ em todas as provas).

Gráficos – Médias de notas por tipo de escola (pública x privada)





Fonte: elaborado pelos autores (2025).

6. Discussão

Os resultados apresentados revelam que tanto a escolaridade dos pais quanto o tipo de escola frequentada pelos estudantes constituem variáveis determinantes para o desempenho no ENEM 2020. Observou-se que a escolaridade materna apresentou correlações mais fortes com as notas obtidas em todas as áreas do exame, em comparação com a escolaridade paterna. Tal achado corrobora evidências sociológicas e educacionais que apontam a figura materna como principal mediadora no acompanhamento escolar e no incentivo ao estudo. Além disso, a elevada incidência de lares chefiados por mulheres no Brasil reforça o papel central do capital educacional materno no processo de formação acadêmica dos estudantes.

Paralelamente, a análise sobre o tipo de escola demonstrou que os alunos provenientes de instituições privadas obtiveram, em média, notas significativamente

mais altas do que os oriundos de escolas públicas em todas as áreas avaliadas. A aplicação do Teste de Mann-Whitney U confirmou que tais diferenças não são aleatórias, mas estatisticamente significativas. Este resultado reflete desigualdades estruturais do sistema educacional brasileiro, nas quais fatores como infraestrutura, qualidade docente, recursos pedagógicos e condições socioeconômicas do aluno exercem influência direta sobre o desempenho escolar.

Assim, a integração dessas análises evidencia que o desempenho acadêmico no ENEM é resultado de um conjunto complexo de fatores, em que tanto as condições familiares quanto institucionais se inter-relacionam, produzindo impactos cumulativos e significativos nos resultados.

7. Conclusão

Com base nas análises estatísticas realizadas, conclui-se que:

1. Existe correlação positiva entre a escolaridade dos pais e o desempenho dos estudantes, sendo a escolaridade materna o fator mais fortemente associado às notas obtidas no ENEM 2020.
2. Alunos de escolas privadas apresentaram, em média, notas significativamente mais altas em todas as áreas avaliadas, quando comparados aos de escolas públicas, confirmando a hipótese de disparidade entre os grupos.
3. Os resultados refletem desigualdades socioeconômicas e estruturais persistentes no sistema educacional brasileiro, nas quais a origem familiar e o tipo de escola exercem papel decisivo no desempenho acadêmico.

Dessa forma, o estudo reforça a necessidade de políticas públicas que busquem reduzir a desigualdade de oportunidades educacionais, por meio do fortalecimento da escola pública e da promoção de ações que envolvam a família no processo de aprendizagem. No campo da ciência de dados, a investigação evidencia a relevância da análise estatística como ferramenta para identificar padrões e subsidiar tomadas de decisão em contextos educacionais.

8. Referencial

MURIELBS. *Analise_enem_2020*. GitHub, 2020. Repositório disponível em: https://github.com/Murielbs/Analise_enem_2020/tree/main. Acesso em: 27 set. 2025.