

Análise Exploratória do Desempenho no ENEM e sua Relação com Fatores Socioeconômicos na Gestão Social

Luiz Henrique Rodrigues Farias (UFERSA)
Ana Cláudia Souza Vidal de Negreiros (UFERSA)
Miriam Karla Rocha (UFRN)

As desigualdades educacionais no Brasil evidenciam o impacto das diferenças socioeconômicas e do acesso à tecnologia no desempenho escolar dos estudantes. Este estudo investiga como fatores socioeconômicos, acesso à tecnologia e o tipo de escola influenciam o desempenho no Exame Nacional do Ensino Médio (ENEM). Para investigar essas relações, foram analisados os microdados do ENEM 2023, aplicando técnicas estatísticas, como ANOVA (Análise de Variância) e o Teste de Tukey, para identificar padrões e diferenças significativas no desempenho dos estudantes. Os resultados mostraram que a renda familiar e o acesso a tecnologias, como computadores e celulares, influenciam diretamente o desempenho acadêmico, com estudantes que possuem maior acesso a esses recursos apresentando melhores notas no ENEM, de forma geral.

Palavras-chave: Desigualdades educacional, fatores socioeconômicos, ANOVA.

1. Introdução

O desenvolvimento acadêmico desde a infância até o ensino médio, seguido pela busca por capacitação no ensino superior, é uma realidade comum para muitos cidadãos que almejam melhorar suas perspectivas de vida. Nesse contexto, o Exame Nacional do Ensino Médio (ENEM) destaca-se como um dos principais instrumentos de avaliação para o ingresso em cursos superiores (INEP, 2020). De acordo com Oliveira (2016), o ENEM vai muito além de um exame avaliativo, pois funciona como um importante subsídio para reestruturação do ensino médio no Brasil ao avaliar as lacunas no aprendizado para auxílio de ajustes às práticas pedagógicas e políticas públicas.

De acordo com o Ministério da Educação (INEP, 2020),

O ENEM tem, ainda, papel fundamental na implementação da Reforma do Ensino Médio, ao apresentar, nos itens da prova, os conceitos de situação-problema, interdisciplinaridade e contextualização, que são, ainda, mal compreendidos e pouco habituais na comunidade escolar. A prova do ENEM, ao entrar na escola, possibilita a discussão entre professores e alunos dessa nova concepção de ensino preconizada pela LDB — Lei de Diretrizes e Bases da Educação Nacional (Lei nº 9.394/1996) —, pelos Parâmetros Curriculares Nacionais e pela Reforma do Ensino Médio, norteadores da concepção do exame.

Portanto, além de servir como uma ferramenta de competição, o ENEM desempenha um papel social significativo ao avaliar o desempenho dos alunos no ensino médio e fornecer subsídios para a melhoria do sistema educacional público brasileiro. Essa avaliação pode ser feita a partir de dados gerados pelo ENEM.

Nesse contexto, os microdados gerados pelo ENEM, contém diversas informações fornecidas

pelos candidatos, assim, oferecem uma rica fonte de análise que pode fundamentar políticas públicas voltadas à educação (INEP, 2023). Com o avanço da tecnologia e a crescente disponibilidade de dados, a utilização de métodos estatísticos (Schleger, 2021), como a Análise de Variância (ANOVA) e o teste de Tukey, permite identificar se esses fatores como renda familiar, acesso à tecnologia e tipo de escola influenciam diretamente o desempenho acadêmico.

A visualização dos resultados por meio de gráficos, como boxplots, facilita a compreensão de padrões e discrepâncias, permitindo que gestores públicos tomem decisões embasadas para corrigir deficiências, como a falta de acesso à internet em comunidades carentes. A célebre frase de William Edwards Deming (apud DAYCHOUM, 2007), “O que não pode ser medido, não pode ser gerenciado”, reforça que, sem dados concretos e mensuráveis sobre o desempenho dos alunos e as condições educacionais, obtidos por meio dos microdados do ENEM, torna-se impossível gerenciar ou melhorar o sistema educacional de maneira eficiente.

Desta forma, o objetivo deste estudo é esclarecer como fatores socioeconômicos, o acesso à tecnologia e o tipo de escola afetam o desempenho dos estudantes no ENEM, utilizando métodos estatísticos para identificar diferenças e padrões que podem ajudar na criação de políticas públicas mais eficazes na educação. Assim, o tratamento e a análise adequada dos microdados do ENEM poderão possibilitar uma compreensão mais profunda das desigualdades educacionais e oferecem uma base sólida para a formulação de políticas públicas que promovam um sistema educacional mais inclusivo e eficiente no Brasil.

2. Referencial Teórico

Neste tópico, serão abordados temas relacionados ao âmbito socioeconômico da educação, explorando fatores como a rede de apoio familiar, as diferenças entre escolas públicas e privadas e o acesso à tecnologia influenciam o desempenho educacional. Além do mais, será discutido o papel do Exame Nacional do Ensino Médio (ENEM) na avaliação do ensino básico e sua importância para a formulação de políticas educacionais. Ao final, será apresentada uma análise da relevância da análise de dados na educação, destacando como essa prática pode contribuir para a melhoria da qualidade do ensino e para a equidade nas oportunidades educacionais.

2.1 Determinantes socioeconômicos da educação

A disponibilidade de recursos para a realização de uma atividade é um fator crucial em qualquer tarefa, e no contexto educacional isso não é diferente. Como levantado por Juliano e Yunes (2014), indivíduos que contam com uma rede de apoio, como pais e familiares, que fornecem

suporte para o desenvolvimento acadêmico, têm uma vantagem em comparação com aqueles cujas famílias dependem de sua contribuição para a renda familiar.

A interpretação de Luz (2006) sugere que o desempenho educacional e as desigualdades observadas são resultantes de interações entre diversas variáveis. Dentre essas, destacam-se fatores micro, como a escolaridade dos pais e a renda familiar, assim como fatores macro, incluindo o sistema educacional, que abrange a formação dos professores, os recursos disponíveis e as políticas públicas que visam promover o desenvolvimento educacional. Portanto, é fundamental identificar os principais fatores determinantes na educação e sua relevância na formação dos indivíduos, com o objetivo de garantir que a parte da sociedade mais dependente de instituições públicas, como creches, escolas de ensino fundamental e médio, e universidades, tenha acesso a um ensino de qualidade que promova maior acessibilidade às populações mais vulneráveis.

Moraes e Belluzzo (2014) ressaltam uma característica importante na comparação do desempenho entre escolas públicas e privadas: a tendência de que as instituições privadas frequentemente obtêm melhores resultados educacionais. Os autores sugerem que essa diferença pode ser atribuída a fatores adicionais que influenciam a qualidade do serviço educacional, como arranjos administrativos e práticas pedagógicas. Assim, mesmo com um controle de insumos educacionais mais restrito, as escolas privadas conseguem direcionar suas atividades de forma mais eficaz, resultando em um desempenho superior dos alunos em avaliações como o ENEM. Esse argumento enfatiza uma característica que tende a se repetir frequentemente: a superioridade no desempenho educacional das escolas do setor privado em comparação com as do setor público. Esse fator secundário reforça o acesso a recursos que possibilitam investimentos privados em educação pelas famílias, eliminando a necessidade de esperar por medidas públicas que garantam uma melhor qualidade de ensino. Assim, as famílias podem optar por um ensino de maior qualidade no setor privado, especialmente quando não têm acesso a um ensino público satisfatório.

Além do acesso a recursos para a educação, é essencial considerar o acesso à informação por meio das tecnologias. No Brasil, os relatórios da pesquisa TIC Educação 2016 (CGI.br, 2016) mostram que 52% dos alunos do 5º ao 9º ano utilizam celulares—dispositivos mais acessíveis em termos de custo em comparação com computadores e tablets—para atividades escolares, e esse percentual aumenta para 74% entre os estudantes do ensino médio. É importante destacar que esse estudo foi realizado antes da pandemia de COVID-19, um evento que evidenciou ainda mais a importância do uso de tecnologias na educação.

2.2 Exame Nacional do Ensino Médio

De acordo com o Ministério da Educação (INEP, 2024),

“o Exame Nacional do Ensino Médio foi instituído em 1998, com o objetivo de avaliar o desempenho escolar dos estudantes ao término da educação básica. Em 2009, o exame aperfeiçoou sua metodologia e passou a ser utilizado como mecanismo de acesso à educação superior”.

O exame, conforme definido pelo Ministério da Educação (INEP, 2024), consiste em 180 questões objetivas organizadas em quatro áreas de conhecimento: linguagens, códigos e suas tecnologias; ciências humanas e suas tecnologias; ciências da natureza e suas tecnologias; e matemática e suas tecnologias. Além disso, é exigida a elaboração de uma redação dissertativo-argumentativa, que deve ser fundamentada em uma situação ou problema central contextualizado.

Como se trata de uma avaliação do desempenho escolar, a inscrição no ENEM requer a disponibilização de algumas informações socioeconômicas. Esses dados ajudam a compreender o perfil dos participantes, sua situação econômica e outras características que podem contribuir para o planejamento de ações voltadas à reestruturação e à melhoria do ensino básico.

2.3 Análise de dados na educação

O surgimento da escrita representou um marco no desenvolvimento humano, possibilitando o registro e o aprimoramento contínuo do conhecimento (FISCHER, 2009). Com os avanços da Tecnologia da Informação, ocorreu uma nova transformação, facilitando o acesso e a disseminação de informações, além de promover a colaboração entre indivíduos (ROGERS, 2017). A produção massiva de dados educacionais, como os resultados do Exame Nacional do Ensino Médio (ENEM), oferece a oportunidade de visualizar informações que antes eram inacessíveis e, por isso, pode ser fundamental para a avaliação educacional.

A análise de dados, nesse contexto, torna-se essencial para organizar e interpretar essas informações em um ambiente saturado de dados (GARCIA e OLIVEIRA, 2012). A interpretação dos resultados do ENEM é crucial para formular ações e tomar decisões baseadas em evidências, permitindo a identificação de padrões e tendências no desempenho escolar dos estudantes. Assim, a análise de dados contribui para uma compreensão mais precisa da qualidade do ensino e das desigualdades educacionais, proporcionando *insights* que podem ser utilizados para melhorar políticas públicas e práticas pedagógicas.

Nas instituições educacionais, de acordo com o Ministério da Educação (INEP, 2023), a análise de dados desempenha um papel estratégico, permitindo que educadores e gestores avaliem o desempenho dos alunos e das ações implementadas nas escolas. Essa abordagem pode personalizar estratégias de ensino e aprendizagem, melhorando a eficácia das intervenções educacionais. Portanto, a análise de dados não só ajuda as organizações a se adaptarem e

inovarem, mas também influencia diretamente a eficiência e a qualidade da educação oferecida, contribuindo para um futuro mais equitativo e sustentável na formação dos indivíduos.

3. Procedimentos metodológicos

A análise do desempenho acadêmico no ENEM é fundamental para compreender as dinâmicas educacionais no Brasil. O acesso aos microdados do ENEM 2023, disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), possibilitou uma investigação detalhada das variáveis socioeconômicas e educacionais que influenciam os resultados dos estudantes.

A metodologia adotada neste estudo é caracterizada como exploratória e de natureza quantitativa, com o propósito de estabelecer um caminho claro dividido em etapas que incluem: (i) **acesso aos microdados do ENEM 2023**, (ii) **extração, transformação e carregamento dos dados** e (iii) **análise estatística**, com o auxílio de gráficos para visualizar a distribuição dos dados e fornecer *insights* relevantes sobre a qualidade da educação no país.

3.1. Acesso aos Microdados do ENEM 2023

De acordo com o Ministério da Educação, os microdados disponibilizados pelo Inep (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), através do gov.br, reúnem um conjunto informações detalhadas relacionadas a exames, avaliações e pesquisas realizadas pelo Instituto. Esses dados são organizados de forma que não permitam a identificação de pessoas, cumprindo as normas de proteção à privacidade. Essa reestruturação dos arquivos tem como objetivo garantir a segurança das informações pessoais, de acordo com as regulamentações vigentes.

Os microdados fornecem, além dos resultados individuais dos alunos (sem qualquer identificação pessoal), uma gama de outras informações que são preenchidas no momento da inscrição. Isso inclui aspectos sociodemográficos e educacionais que são valiosos para o nosso objetivo de análise, que é identificar como variáveis como renda familiar, acesso à tecnologia e o tipo de escola frequentada influenciam o desempenho acadêmico. Para isso, serão exploradas as informações contidas na Tabela 1.

As variáveis selecionadas neste estudo, obtidas dos microdados do ENEM 2023, foram escolhidas por sua importância para entender como fatores socioeconômicos e educacionais influenciam o desempenho acadêmico. As informações coletadas na inscrição permitem organizar os estudantes em grupos específicos, facilitando a identificação de padrões e a análise do impacto dessas condições no resultado das provas. Essa abordagem ajuda a evidenciar desigualdades no sistema educacional brasileiro de forma objetiva.

Tabela 1 – Questionário socioeconômico do ENEM 2023

Pergunta	Agrupamentos de Respostas
Qual é a renda mensal de sua família?	A: Nenhuma Renda
	B: Até R\$ 1.320,00
	C: De R\$ 1.320,01 até R\$ 1.980,00
	D: De R\$ 1.980,01 até R\$ 2.640,00
	E: De R\$ 2.640,01 até R\$ 3.300,00
	F: De R\$ 3.300,01 até R\$ 3.960,00
	G: De R\$ 3.960,01 até R\$ 5.280,00
	H: De R\$ 5.280,01 até R\$ 6.600,00
	I: De R\$ 6.600,01 até R\$ 7.920,00
	J: De R\$ 7.920,01 até R\$ 9.240,00
	K: De R\$ 9.240,01 até R\$ 10.560,00
	L: De R\$ 10.560,01 até R\$ 11.880,00
	M: De R\$ 11.880,01 até R\$ 13.200,00
	N: De R\$ 13.200,01 até R\$ 15.840,00
	O: De R\$ 15.840,01 até R\$ 19.800,00
	P: De R\$ 19.800,01 até R\$ 26.400,00
	Q: Acima de R\$ 26.400,00
Na sua residência tem computador?	A: Não
	B: Sim, um
	C: Sim, dois
	D: Sim, três
	E: Sim, quatro ou mais
Na sua residência tem telefone celular?	A: Não
	B: Sim, um
	C: Sim, dois
	D: Sim, três
	E: Sim, quatro ou mais
Qual o tipo de escola do Ensino Médio que você frequentou?	1: Não Respondeu
	2: Pública
	3: Privada
Na sua residência tem acesso à internet?	A: Não
	B: Sim

Fonte: adaptado (INEP, 2023).

3.2. Extração, transformação e carregamento dos dados

Diante do grande volume de dados, cerca de 1,74 GB contendo mais de 3,9 milhões de inscritos, foi necessário realizar um tratamento cuidadoso para garantir que os dados estivessem prontos para análise. Esse tratamento foi feito com o uso da linguagem de programação Python, aproveitando as bibliotecas Pandas e Dask. Ambas são amplamente utilizadas por sua flexibilidade e capacidade de integração com outras ferramentas de análise. A biblioteca Dask, em particular, oferece a funcionalidade de processamento paralelo, o que é essencial para lidar

com grandes volumes de dados sem sobrecarregar o hardware, permitindo o processamento eficiente dos microdados do ENEM.

Para efeito de comparação, um software comum como o Excel possui um limite de 1.048.576 linhas, o que equivale a apenas 26,65% das linhas presentes nos microdados do ENEM 2023. Portanto, o uso de uma ferramenta mais robusta como o Dask foi imprescindível para manipular essa quantidade de informações.

O processo de ETL (extração, transformação e carregamento) foi aplicado para extrair apenas os dados essenciais para a análise. Nesse caso, foram considerados apenas os candidatos que participaram de todas as provas e que não se identificaram como treineiros no momento da inscrição. Além disso, foram selecionadas as colunas referentes às variáveis de interesse, como renda familiar, presença de computadores e celulares na residência, tipo de escola de ensino médio (pública ou privada), acesso à internet na residência e a nota média dos candidatos.

Com esse tratamento inicial, o tamanho do arquivo foi drasticamente reduzido de 1,74 GB para aproximadamente 46,2 MB, o que representa uma redução de 97,34%. Esse processo permitiu que o arquivo se tornasse mais leve e viável para a realização das análises estatísticas, como a ANOVA e o Teste de Tukey, sem maiores limitações de hardware.

3.3. Análise estatística

A Análise de Variância (ANOVA) é um método estatístico utilizado para comparar médias entre grupos e identificar diferenças significativas. Na sua forma mais simples, a ANOVA one-way é aplicada para comparar três ou mais grupos, enquanto versões mais complexas, como a ANOVA de dois fatores, permitem análises com apenas dois grupos ou múltiplos fatores simultaneamente (Pagano e Gauvreau, 2004). Este método busca identificar se pelo menos um grupo apresenta um comportamento distinto em relação aos demais ou se as variações observadas são meramente resultado de flutuações internas dentro dos grupos. Quando a variação entre os grupos é maior do que a variação observada dentro de cada grupo, a ANOVA sugere a existência de uma diferença significativa entre os grupos analisados. Logo, o método analisa a variância em comparação a um nível de significância pré-estabelecido, normalmente definido em 5% (0,05), sendo utilizado também neste trabalho para validar duas hipóteses:

- H_0 : Não há diferença no desempenho médio entre os grupos.
- H_1 : Pelo menos um grupo apresenta um desempenho médio diferente dos outros.

Se o valor-p calculado pela ANOVA for inferior a 0,05, a hipótese nula (H_0) é rejeitada, indicando que pelo menos um grupo tem um desempenho médio diferente dos demais. Por outro lado, se o valor-p for igual ou superior a 0,05, conclui-se que não existem diferenças significativas no desempenho médio entre os grupos.

Caso a ANOVA indique a existência de diferenças significativas entre os grupos (rejeição da hipótese nula), torna-se necessário aplicar um método que permita identificar quais grupos apresentam médias significativamente diferentes. De acordo com Hair et al. (2009), o Teste de Tukey é ideal para essa finalidade devido possuir poucos grupos, uma vez que a ANOVA apenas indica se existe uma diferença significativa entre os grupos, mas não especifica qual grupo é superior ou inferior. Assim, o Teste de Tukey fornece uma análise mais detalhada, esclarecendo como os diferentes fatores impactam o desempenho acadêmico.

Os gráficos boxplot, frequentemente chamados de diagramas de caixa, são ferramentas gráficas que permitem visualizar a distribuição de dados. Eles ajudam a identificar a dispersão, a centralidade e possíveis outliers (valores extremos) em um conjunto de dados. No gráfico, diversos componentes são destacados: (i) a caixa, que representa o intervalo entre o primeiro quartil (Q1), que abrange 25% dos dados, e o terceiro quartil (Q3), que abrange 75%, refletindo assim 50% dos dados centrais, conhecido como intervalo interquartil (IQR); (ii) a mediana, indicada pela linha central da caixa, que divide o conjunto de dados em duas partes iguais, com cerca de 50% dos valores acima e 50% abaixo da mediana; (iii) as extensões, que são linhas que se estendem 1,5 vezes o IQR a partir de Q1 e Q3; e (iv) os outliers, que aparecem como pontos além das extensões e podem indicar variações nos dados ou erros de coleta.

As principais vantagens dos gráficos boxplot incluem a visualização clara da distribuição dos dados, a capacidade de identificar outliers que podem comprometer a confiabilidade das análises, e sua versatilidade para comparar a distribuição entre diferentes grupos.

3.4. Ferramentas e bibliotecas

A análise de grandes conjuntos de dados exige ferramentas especializadas que combinem eficiência computacional, robustez estatística e clareza na apresentação dos resultados. Neste projeto, o processamento inicial dos microdados do ENEM 2023 será feito com o Dask, devido à sua capacidade de lidar com grandes volumes de dados de forma paralela. Após a redução do volume durante o tratamento, a manipulação será transferida para o Pandas, que oferece maior flexibilidade para análises detalhadas e integração com bibliotecas de visualização, como o Matplotlib.

Para as análises estatísticas, o projeto utilizará o Statsmodels, que oferece recursos avançados para regressões lineares e testes de comparação entre grupos, como o Teste de Tukey. A função OLS será usada para modelar relações entre variáveis, enquanto o `pairwise_tukeyhsd` realizará o Teste de Tukey pós-ANOVA para identificar diferenças significativas entre grupos. Para a visualização dos resultados, a biblioteca Matplotlib será empregada, oferecendo grande flexibilidade na criação de gráficos, como linhas, barras, dispersão, histogramas e boxplots. Seu

módulo pyplot integra-se facilmente com DataFrames do Pandas, facilitando a criação de gráficos diretamente a partir das tabelas de dados, o que torna o processo analítico mais ágil e visualmente acessível. Essa abordagem garante que cada ferramenta seja utilizada em sua melhor capacidade, assegurando eficiência no tratamento e na interpretação dos dados.

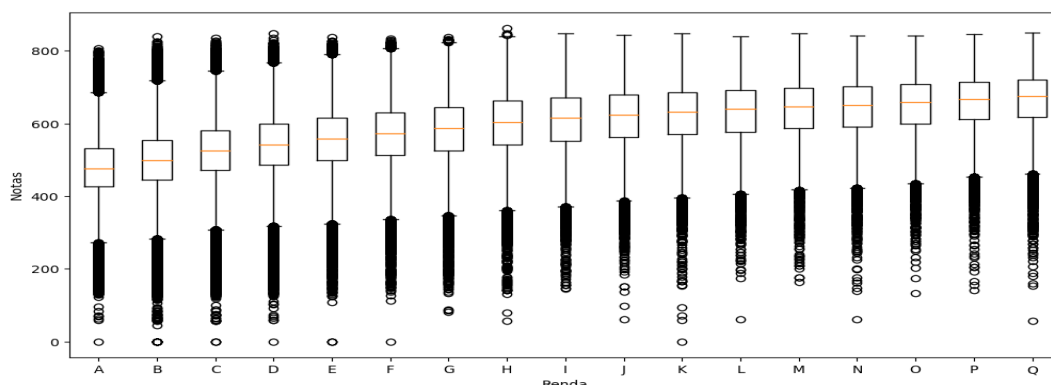
4. Resultados e Discussões

Nesta seção, são apresentados os principais resultados da pesquisa, organizados de forma clara para facilitar a compreensão. Para a análise, são utilizados métodos estatísticos mencionados anteriormente, garantindo maior embasamento e confiabilidade ao estudo. Os dados são ilustrados por meio de gráficos e tabelas, permitindo identificar padrões e tendências relevantes para a discussão.

4.1 Desempenho dos estudantes com base na renda familiar

A análise das notas médias ENEM 2023 entre as diferentes faixas de renda dos inscritos é apresentada a seguir, utilizando gráficos boxplots que ilustram a distribuição de cada grupo, organizados em ordem crescente de renda familiar, como demonstrado na Figura 1.

Figura 1 – Notas por renda familiar



Fonte: autoria própria, 2025.

Observa-se uma tendência de aumento nas notas do ENEM conforme a faixa de renda, conforme descrito na Tabela 2, o que pode ser confirmado posteriormente por meio da ANOVA.

Tabela 2 - Rótulos de renda familiar

Rótulo	Renda Familiar
A	Nenhuma Renda
B	Até R\$ 1.320,00
C	De R\$ 1.320,01 até R\$ 1.980,00.
D	De R\$ 1.980,01 até R\$ 2.640,00.
E	De R\$ 2.640,01 até R\$ 3.300,00.
F	De R\$ 3.300,01 até R\$ 3.960,00.
G	De R\$ 3.960,01 até R\$ 5.280,00.
H	De R\$ 5.280,01 até R\$ 6.600,00.

I	De R\$ 6.600,01 até R\$ 7.920,00.
J	De R\$ 7.920,01 até R\$ 9.240,00.
K	De R\$ 9.240,01 até R\$ 10.560,00.
L	De R\$ 10.560,01 até R\$ 11.880,00.
M	De R\$ 11.880,01 até R\$ 13.200,00.
N	De R\$ 13.200,01 até R\$ 15.840,00.
O	De R\$ 15.840,01 até R\$ 19.800,00.
P	De R\$ 19.800,01 até R\$ 26.400,00.
Q	Acima de R\$ 26.400,00.

Fonte: autoria própria, 2025.

Os boxplots por faixa de renda mostram uma possível relação entre a renda e as notas no ENEM. Para entender melhor essa questão a Figura 2 mostra a ANOVA realizada para verificar se há diferenças significativas entre as notas dos grupos.

Figura 2 - Análise de variância de notas por renda familiar

	sum_sq	df	F	PR(>F)
C(Q006)	4.531659e+09	16.0	39204.97129	0.0
Residual	1.565381e+10	2166826.0	NaN	NaN

Fonte: autoria própria, 2025.

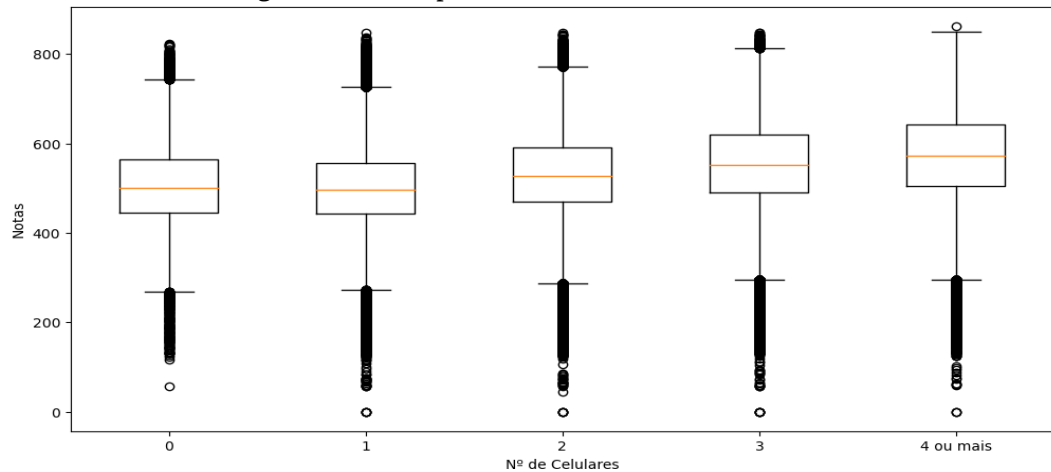
Como o valor-p é inferior a 0,05, a hipótese nula (H_0) é rejeitada, indicando que pelo menos um dos grupos apresenta uma diferença significativa entre as médias. Para identificar qual grupo possui a média maior ou menor o Teste de Tukey foi realizado.

De acordo com o Teste de Tukey, todas as análises comparativas rejeitaram a hipótese nula (H_0), indicando que existem diferenças significativas entre os grupos. Além disso, as diferenças são crescentes à medida que se comparam os grupos de A a P. Assim, as médias dos inscritos mostram um comportamento crescente, sugerindo que, quanto maior a renda familiar, maior tende a ser a nota no ENEM.

4.2 Impacto do acesso a celulares

Seguindo o mesmo raciocínio, pode-se analisar a distribuição dos inscritos em grupos através da Figura 3. Além disso, nota-se uma possível relação entre a nota no ENEM e o número de celulares, a qual será investigada mais detalhadamente por meio da ANOVA, Figura 4. Como o valor-p, $PR(>F)$, é inferior a 0,05, a hipótese nula é rejeitada, indicando que pelo menos um dos grupos apresenta uma diferença significativa entre as médias. Para identificar qual grupo possui a média maior ou menor, será aplicado o Teste de Tukey, Tabela 3, para comparações em pares dos grupos descritos na Quadro 1.

Figura 3 – Notas por número de celulares



Fonte: Autoria própria, 2025.

Figura 4 - Análise de variância de notas por número de celulares

	sum_sq	df	F	PR(>F)
C(Q022)	1.335194e+09	4.0	38370.123135	0.0
Residual	1.885028e+10	2166838.0	NaN	NaN

Fonte: autoria própria, 2025.

Tabela 3 – Teste de Tukey para número de celulares

Multiple Comparison of Means - Tukey HSD, FWER=0.01						
=====						
group1	group2	meandiff	p-adj	lower	upper	reject

A	B	-5.9209	0.0	-7.3165	-4.5254	True
A	C	23.6675	0.0	22.302	25.0329	True
A	D	47.4519	0.0	46.0917	48.8121	True
A	E	65.4403	0.0	64.0618	66.8187	True
B	C	29.5884	0.0	28.9605	30.2163	True
B	D	53.3728	0.0	52.7566	53.9891	True
B	E	71.3612	0.0	70.7055	72.0169	True
C	D	23.7844	0.0	23.2397	24.3292	True
C	E	41.7728	0.0	41.1838	42.3618	True
D	E	17.9884	0.0	17.4118	18.565	True

Fonte: autoria própria, 2025.

Portanto, a análise das diferenças das médias revela uma correlação positiva entre a quantidade de celulares na residência e a nota no ENEM. No entanto, observa-se uma exceção para as famílias que possuem apenas um celular, cuja nota no ENEM foi inferior à das famílias sem nenhum celular.

Quadro 1 - Rótulos de número de celulares

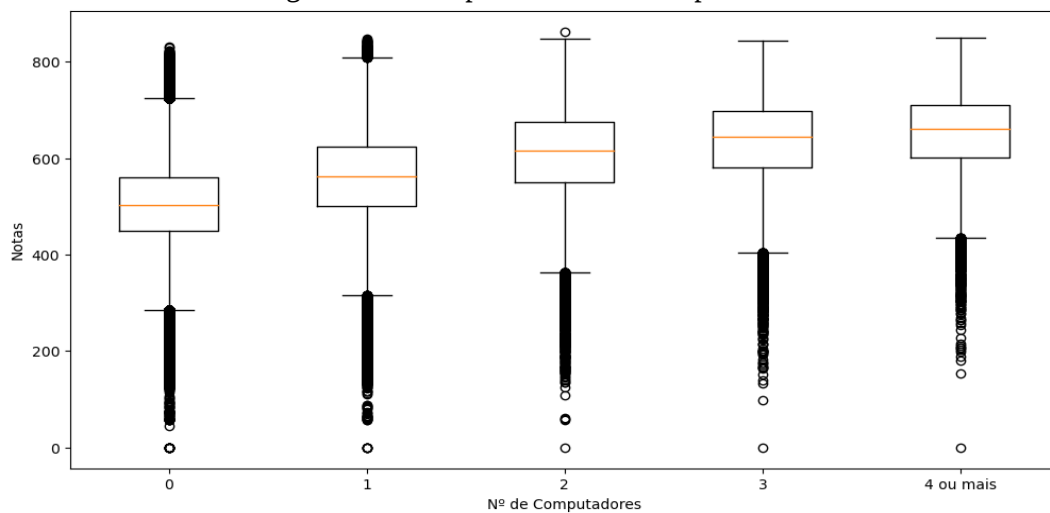
Rótulo	Quantidades de Celulares na Residência
A	Não.
B	Sim, um.
C	Sim, dois.
D	Sim, três.
E	Sim, quatro ou mais.

Fonte: autoria própria, 2025.

4.3 Impacto do acesso a computadores

Seguindo a mesma linha de raciocínio, pode-se analisar a distribuição dos inscritos em grupos através na Figura 5. Nota-se uma possível relação entre a nota no ENEM e o número de computadores, que será analisada mais detalhadamente por meio da ANOVA, Figura 6.

Figura 5 – Notas por número de computadores



Fonte: autoria própria, 2025.

Figura 6 - Análise de variância de notas por número de computadores

	sum_sq	df	F	PR(>F)
C(Q024)	3.701742e+09	4.0	121651.418413	0.0
Residual	1.648373e+10	2166838.0	NaN	NaN

Fonte: autoria própria, 2025.

Como o valor-p, $PR(>F)$, é inferior a 0,05, a hipótese nula (H_0) é rejeitada, indicando que pelo menos um dos grupos apresenta uma diferença significativa entre as médias. Para identificar qual grupo possui a média maior ou menor, será aplicado o Teste de Tukey na tabela 4 para comparações em pares descritos na quadro 2.

Tabela 4 – Teste de Tukey para número de computadores

Multiple Comparison of Means - Tukey HSD, FWER=0.01						
=====						
group1	group2	meandiff	p-adj	lower	upper	reject

A	B	56.8189	0.0	56.3977	57.2402	True
A	C	104.5946	0.0	103.899	105.2902	True
A	D	129.8895	0.0	128.7897	130.9893	True
A	E	145.979	0.0	144.4081	147.5499	True
B	C	47.7757	0.0	47.0597	48.4917	True
B	D	73.0706	0.0	71.9578	74.1834	True
B	E	89.1601	0.0	87.58	90.7401	True
C	D	25.2949	0.0	24.0521	26.5378	True
C	E	41.3844	0.0	39.7102	43.0586	True
D	E	16.0895	0.0	14.211	17.9679	True

Fonte: autoria própria, 2025.

Quadro 2 - Rótulos de número de computadores

Rótulo	Quantidades de Computadores na Residência
A	Não.
B	Sim, um.
C	Sim, dois.
D	Sim, três.
E	Sim, quatro ou mais.

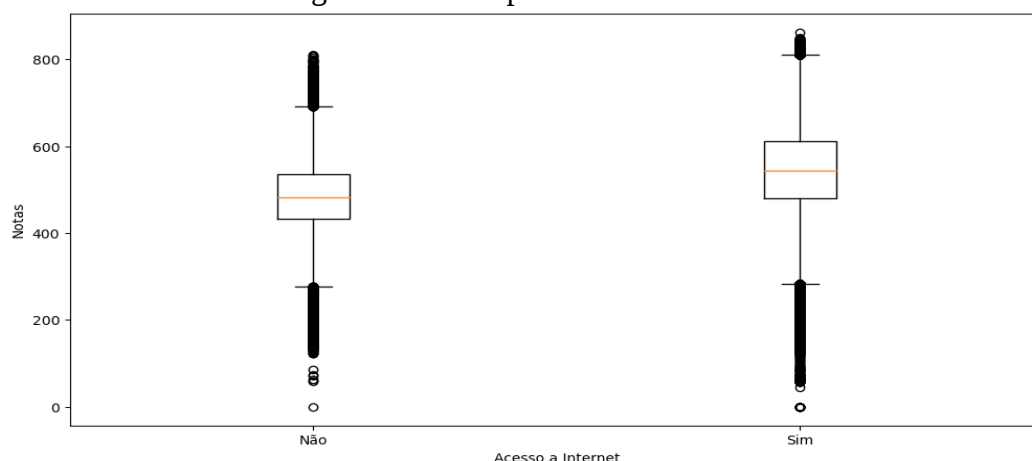
Fonte: autoria própria, 2025.

Portanto, a análise das diferenças das médias no Teste de Tukey revela uma correlação positiva entre a quantidade de computadores na residência e a nota no ENEM.

4.4 Impacto do acesso a internet

Seguindo a mesma linha de raciocínio, pode-se analisar a distribuição dos inscritos em grupos através da figura 7. Observa-se uma possível relação entre a nota no ENEM e o acesso à internet, que será investigada mais detalhadamente por meio da ANOVA, Figura 8. Como o valor-p é inferior a 0,05, a hipótese nula (H_0) é rejeitada, indicando que pelo menos um dos grupos apresenta uma diferença significativa entre as médias. Para identificar qual grupo possui a média maior ou menor, será aplicado o Teste de Tukey, Tabela 5, a seguir para comparações em pares descritos no Quadro 3. Assim, a análise das diferenças das médias no Teste de Tukey indica uma correlação positiva quando há acesso à internet na residência.

Figura 7 – Notas por acesso a internet



Fonte: autoria própria, 2025.

Figura 8 - Análise de variância de notas por acesso a internet

	sum_sq	df	F	PR(>F)
C(Q025)	6.782953e+08	1.0	75344.485286	0.0
Residual	1.950718e+10	2166841.0	NaN	NaN

Fonte: autoria própria, 2025.

Tabela 5 – Teste de Tukey para acesso a internet

Multiple Comparison of Means - Tukey HSD, FWER=0.01						
=====						
group1	group2	meandiff	p-adj	lower	upper	reject

A	B	62.3197	0.0	61.7349	62.9045	True

Fonte: autoria própria, 2025.

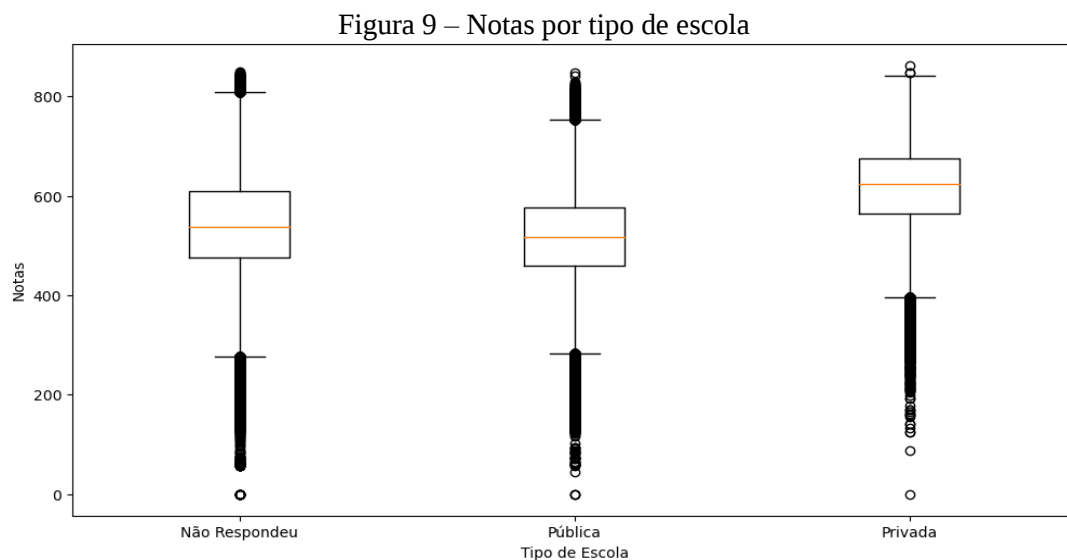
Quadro 3 - Rótulos de acesso a internet

Rótulo	Quantidades de Celulares na Residência
A	Não.
B	Sim.

Fonte: autoria própria, 2025.

4.5 Influência da modalidade de ensino médio

Ao se tratar da distribuição das notas do ENEM entre os grupos de alunos das modalidades de ensino público, privado e não declarado, tem-se a distribuição das notas descrita pela figura 9. Observa-se que os alunos provenientes de escolas particulares apresentaram resultados superiores no ENEM, o que pode ser comprovado por meio da análise ANOVA, Figura 10.



Fonte: autoria própria, 2025.

Figura 10 - Análise de variância de notas por tipo de escola

	sum_sq	df	F	PR(>F)
C(TP_ESCOLA)	1.783982e+09	2.0	105035.085157	0.0
Residual	1.840149e+10	2166840.0	NaN	NaN

Fonte: autoria própria, 2025.

Como o valor-p é inferior a 0,05, a hipótese nula (H_0) é rejeitada, indicando que pelo menos um dos grupos apresenta uma diferença significativa entre as médias. Para identificar qual grupo possui a média maior ou menor, será aplicado o Teste de Tukey, tabela 6, a seguir para comparações em pares descritos no quadro 4.

Tabela 6 – Teste de Tukey para tipo de escola

Multiple Comparison of Means - Tukey HSD, FWER=0.01						
group1	group2	meandiff	p-adj	lower	upper	reject
1	2	-28.1026	0.0	-28.4918	-27.7133	True
1	3	72.1976	0.0	71.5729	72.8223	True
2	3	100.3002	0.0	99.6578	100.9425	True

Fonte: autoria própria, 2025.

Quadro 4 - Rótulos de acesso a internet

Rótulo	Quantidades de Celulares na Residência
1	Não Respondeu
2	Pública
3	Privada

Fonte: autoria própria, 2025.

Assim, a análise das diferenças das médias no Teste de Tukey indica que alunos de escolas privadas alcançam melhores resultados no ENEM.

4.6 Apresentação dos resultados e algumas considerações

A análise dos dados revela algumas tendências interessantes sobre o desempenho dos estudantes no ENEM, com base em fatores como renda familiar e acesso à tecnologia. Na Quadro 5, é apresentado um resumo dos principais resultados que emergiram de cada subtópico.

Quadro 5 – Resumo de resultados e considerações

Determinantes do Desempenho	Resultado Obtido	Considerações
Desempenho com base na renda familiar	O aumento da renda familiar está associado a um crescimento nas notas do ENEM.	Estudantes de famílias com maior renda tendem a obter melhores notas, refletindo desigualdades no acesso a recursos.
Impacto do acesso a celulares	Observa-se uma relação entre a quantidade de celulares na residência e as notas no ENEM, com exceção das famílias que possuem apenas um celular, as quais apresentam notas mais baixas do que aquelas sem nenhum celular.	O acesso a celulares pode facilitar o acesso à informação, mas sua quantidade não é diretamente proporcional ao desempenho.
Impacto do acesso a computadores	Aumento no número de computadores na residência está positivamente relacionado com melhores notas no ENEM.	A presença de computadores em casa reforça o aprendizado digital, associado a um desempenho acadêmico superior.
Impacto do acesso à internet	Alunos de escolas privadas tendem a ter um desempenho superior ao de alunos de escolas públicas.	A infraestrutura digital nas escolas privadas favorece o aprendizado e o acesso a recursos online, impactando positivamente os resultados.

Fonte: autoria própria, 2025.

Esses resultados mostram que fatores como a renda familiar e o acesso a tecnologias impactam diretamente no desempenho dos estudantes. De forma geral, quem tem mais acesso a recursos, como celulares, computadores e internet, tende a se sair melhor nas provas. No entanto, também

notou-se que o simples fato de ter um celular não garante um desempenho alto, e que as diferenças entre escolas públicas e privadas ainda são um reflexo das desigualdades de acesso e infraestrutura.

5. Conclusões

Antes de tomar decisões para solucionar os desafios da educação brasileira, é fundamental compreender a natureza do problema, suas características, obstáculos, fatores de influência e outros parâmetros relevantes. Esse entendimento permite embasar a tomada de decisões de forma mais assertiva, resultando em soluções que contribuam significativamente para a formulação de políticas públicas voltadas à promoção de uma educação mais inclusiva e eficiente. Além disso, incentiva uma reflexão aprofundada sobre as dinâmicas que moldam o acesso e a qualidade educacional no Brasil.

Os resultados das análises mostradas neste estudo ratificam a necessidade de políticas públicas voltadas para a redução das desigualdades educacionais. A relação entre a renda familiar e o desempenho acadêmico demonstra a importância de programas de assistência estudantil e incentivos financeiros que garantam condições benéficas ao aprendizado para alunos de baixa renda. Além disso, os obstáculos no acesso à tecnologia se relacionam com o desenvolvimento dos estudantes, deixando nítida a urgência de investimentos em infraestrutura digital nas escolas públicas e na ampliação da oferta de dispositivos e internet para alunos em situação de vulnerabilidade social e financeira.

Outro ponto relevante é que, em média, estudantes de escolas públicas apresentam um desempenho inferior ao de alunos de escolas particulares, reforçando a necessidade de uma revisão curricular, maior capacitação dos professores e mais investimentos em materiais didáticos. Essas ações podem proporcionar maior equidade entre indivíduos de diferentes condições socioeconômicas e tornar o ensino mais competitivo. Programas como o Novo Ensino Médio e novas estratégias pedagógicas podem ser fundamentais para diminuir essas desigualdades.

Portanto, fica evidente que fatores socioeconômicos, como a renda familiar e o acesso à informação por meio de computadores, celulares e internet, desempenham um papel essencial no desenvolvimento educacional. As maiores discrepâncias, evidenciadas pelo Teste de Tukey, foram observadas justamente na análise da renda familiar, com diferenças de até 182,20 pontos no ENEM, e no acesso a computadores na residência, com variações de até 129,89 pontos. Além disso, esses fatores estão interligados, pois famílias com quatro ou mais computadores tendem a possuir maior capacidade aquisitiva para adquiri-los, o que está diretamente associado a uma renda familiar mais elevada.

Embora este estudo tenha focado na análise dos microdados do ENEM 2023, é importante considerar direções para futuros trabalhos, estes envolvendo a análise contínua dos dados e a perspectiva temporal. Comparar dados de diferentes anos possibilitaria a visualização de tendências no desempenho dos estudantes ao longo do tempo, além de proporcionar um entendimento mais aprofundado sobre o impacto de políticas educacionais e mudanças sociais nas desigualdades educacionais. O alinhamento constante em análises fundamentadas em dados permite não apenas um monitoramento mais eficiente, mas também uma adaptação mais rápida das estratégias educacionais, baseadas em informações atualizadas. Esse tipo de análise contínua pode ajudar na criação de intervenções mais eficazes, garantindo que todas as ações sejam direcionadas com base em informações reais, contribuindo para transformar a realidade educacional e oferecendo uma formação de qualidade para todos os estudantes, independentemente de sua condição socioeconômica.

REFERÊNCIAS

INEP. **ENEM – Exame Nacional do Ensino Médio: fundamentação teórico-metodológica.**

Brasília: Inep, 2020. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/ENEM_exame_nacional_do_ensino_medio_fundamentacao_teorico_metodologica.pdf. Acesso em: 10 out. 2024.

INEP. **Microdados e sinopse estatística do ENEM 2023 disponíveis.** Disponível em:

<https://www.gov.br/inep/pt-br/assuntos/noticias/ENEM/microdados-e-sinopse-estatistica-do-ENEM-2023-disponiveis>. Acesso em: 10 out. 2024.

SCHLEGER, Carolina Hilda; ZUCOLOTTI, Benjamim; ZANCAN FRANTZ, Rafael.

Métodos estatísticos no contexto de big data. Salão do Conhecimento, v. 7, n. 7, 2021. Acesso em: 10 out. 2024.

LUZ, Luciana Soares. **Os determinantes do desempenho escolar: a estratificação educacional e o efeito valor adicionado.** Associação Brasileira de Estudos Populacionais, 2006.

INEP. **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Exame Nacional do Ensino Médio.** Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/ENEM>. Acesso em: 05 out. 2024.

MORAES, André Guerra Esteves de; BELLUZZO, Walter. **O diferencial de desempenho escolar entre escolas públicas e privadas no Brasil**. *Revista de Economia da Universidade Federal de Santa Catarina*, v. 14, n. 2, p. 63-82, mai.-ago. 2014.

CGI.br – COMITÊ GESTOR DA INTERNET NO BRASIL. **TIC Educação 2016 – Indicadores**. Disponível em: <https://cetic.br/pt/noticia/cetic-br-pesquisa-o-uso-de-celular-por-alunos-para-a-realizacao-de-atividades-escolares/>. Acesso em: 05 out. 2024.

FISCHER, S. R. **A História da Escrita**. São Paulo: UNESP, 2009.

ROGERS, David L. **Transformação digital: repensando o seu negócio para a era digital**. 1. ed. São Paulo: DVS Editora, 2017.

GARCIA, Edivaldo; OLIVEIRA, Elizandra Roseli Macedo de. **Gestão da informação, inovação e inteligência competitiva: como transformar a informação em vantagem competitiva nas organizações**. Rio de Janeiro: Editora Ciência Moderna, 2012.

JULIANO, Maria Cristina Carvalho; YUNES, Maria Angela Mattar. Reflexões sobre rede de apoio social como mecanismo de proteção e promoção de resiliência. *Ambiente & Sociedade*, v. 17, n. 3, set. 2014.

PAGANO, M.; GAUVREAU, K. **Bioestatística: Princípios e Aplicações**. 2. ed. Porto Alegre: Artmed, 2004.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise Multivariada de Dados**. 6. ed. Porto Alegre: Bookman, 2009.

OLIVEIRA, Thiago Soares de. **O ENEM: breves considerações sobre importância avaliativa e reforma educacional**. 2016. PUCRS, RS, Brasil.

DAYCHOUM, Merhi. **40 ferramentas e técnicas de gerenciamento**. Rio de Janeiro: Brasport, 2007.