

INFLUÊNCIA DA RENDA NO DESEMPENHO DOS ESTUDANTES NO ENEM: UMA ANÁLISE LONGITUDINAL

THE INFLUENCE OF INCOME ON STUDENT PERFORMANCE
IN THE ENEM: A LONGITUDINAL ANALYSIS

Antônio Cesar Lopes Garcez Filho

Especialista em Ciência de Dados na Administração Pública pela Universidade Federal de Santa Maria (Santa Maria/Brasil).
Analista de Tecnologia da Informação na Universidade Federal de Santa Maria (Santa Maria/Brasil).
E-mail: antonio.garcez@ufsm.br

Suellen Pereira dos Santos

Especialista em Ciência de Dados na Administração Pública pela Universidade Federal de Santa Maria (Santa Maria/Brasil).
Técnica em Enfermagem no Hospital Universitário de Santa Maria (Santa Maria/Brasil).
E-mail: spereiradosantos49@gmail.com

Marcelo Trevisan

Doutor em Administração pela Universidade Federal do Rio Grande do Sul (Porto Alegre/Brasil).
Professor na Universidade Federal de Santa Maria (Santa Maria/Brasil).
E-mail: marcelo.trevisan@ufsm.br

Lucas Almeida dos Santos

Doutor em Administração pela Universidade Federal de Santa Maria (Santa Maria/Brasil).
Professor na Universidade Franciscana (Santa Maria/ Brasil).
E-mail: luksanttos@gmail.com

Sheila Kocourek

Doutora em Serviço Social pela Pontifícia Universidade Católica do Rio Grande do Sul (Porto Alegre/Brasil).
Professora na Universidade Federal de Santa Maria (Santa Maria/Brasil).
E-mail: sheilakocourek@gmail.com

Recebido em: 16 de junho de 2025

Aprovado em: 15 de agosto de 2025

Sistema de Avaliação: Double Blind Review

RGD | v. 22 | n. 2 | p. 202-225 | jul./dez. 2025

DOI: <https://doi.org/10.25112/rgd.v22i2.4318>

RESUMO

Este estudo analisa a influência da renda familiar sobre o desempenho dos estudantes brasileiros no Exame Nacional do Ensino Médio (Enem), por meio de uma análise longitudinal abrangente das edições do exame desde sua criação em 1998. Dada a centralidade do Enem como instrumento de acesso ao ensino superior e de aferição da qualidade da educação básica, também buscou-se compreender em que medida as desigualdades socioeconômicas, com ênfase na renda, impactam os resultados dos candidatos ao longo do tempo. A pesquisa caracteriza-se como descritiva, de natureza aplicada, com abordagem quantitativa e fundamentação bibliográfica. Foram utilizadas técnicas de ciência de dados e visualização gráfica por meio da linguagem Python, permitindo a análise comparativa dos microdados disponibilizados pelo Inep. Os resultados apontam que há uma relação diretamente proporcional entre nível de renda e desempenho: estudantes de maior renda apresentam médias mais altas em todas as edições analisadas. Apesar de políticas públicas como Sisu, ProUni e Fies ampliarem o acesso ao ensino superior, a renda permanece como forte indicador de desigualdade educacional no Brasil. Conclui-se que o Enem, embora democratize parcialmente o acesso à universidade, ainda reflete disparidades estruturais que necessitam ser enfrentadas por meio de políticas educacionais mais equitativas e integradas. Por fim, o estudo destaca a importância do uso de dados abertos para a formulação de políticas públicas e propõe futuras pesquisas com variáveis complementares que ampliem a compreensão das desigualdades no desempenho educacional.

Palavras-chave: Enem. Desempenho escolar. Desigualdade social. Renda familiar. Políticas educacionais.

ABSTRACT

This study analyzes the influence of family income on the performance of Brazilian students in the National High School Exam (Enem) through a comprehensive longitudinal analysis of the exam editions since its creation in 1998. Given the centrality of Enem as both an instrument for access to higher education and a measure of the quality of basic education, the study also aims to understand the extent to which socioeconomic inequalities—particularly income—impact candidates' results over time. This is a descriptive, applied research with a quantitative approach and bibliographic foundation. Data science techniques and graphical visualization tools were employed using the Python programming language, allowing for comparative analysis of microdata made available by Inep. The results show a directly proportional relationship between income level and performance: higher-income students consistently achieve higher average scores across all analyzed editions. Although public policies such as Sisu, ProUni, and Fies have expanded access to higher education, income remains a strong indicator of educational inequality in Brazil. The study concludes that while Enem partially democratizes access to university, it still reflects structural disparities that must be addressed through more equitable and integrated educational policies. Finally, the study highlights the importance of open data in public policy formulation and suggests future research incorporating complementary variables to broaden the understanding of educational performance disparities.

Keywords: Enem. Academic performance. Social inequality. Family income. Educational policies.

1 INTRODUÇÃO

A equidade no acesso à educação e às oportunidades proporcionadas por esta são pilares fundamentais para o desenvolvimento social e econômico de um país. No contexto brasileiro, o Exame Nacional do Ensino Médio (Enem) destaca-se como um dos mais importantes instrumentos de avaliação do desempenho dos estudantes concluintes da educação básica. Os dados disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), autarquia federal responsável pela aplicação do exame, são uma fonte riquíssima para a análise do efeito de desigualdades na educação.

O exame foi criado no ano de 1998, inicialmente com o propósito de avaliar a qualidade do ensino médio no país. A partir de 2009, ampliou-se para servir como uma das principais vias de acesso ao ensino superior, oferecendo aos estudantes oportunidades de ingresso em cursos diversos de instituições públicas e privadas. O Sistema de Seleção Unificada (Sisu) que permite ao estudante o ingresso em universidades públicas, o Programa Universidade para Todos (ProUni) que concede bolsas parciais e integrais em instituições privadas e o Fundo de Financiamento Estudantil (Fies) que oferece financiamento para estudantes cursarem o ensino superior em instituições privadas, utilizam a nota do Enem como critério para a seleção de candidatos.

Dentre as variáveis socioeconômicas que compõem as bases de dados disponibilizadas, a renda se destaca como uma das mais relevantes, sendo apontada como fator determinante para o nível educacional das pessoas. Neste contexto, o nível educacional, por sua vez, determina o nível de renda que os estudantes alcançarão, sendo caracterizado, portanto, o círculo vicioso da desigualdade no Brasil (Helene, 2024).

Dado este contexto, este estudo busca responder os seguintes questionamentos: Em que medida a renda familiar dos candidatos influencia seu desempenho no Enem e como essa relação se comporta ao longo do tempo, revelando possíveis tendências ou padrões recorrentes nas edições do exame? Para a análise longitudinal das 26 edições do Enem é essencial evidenciar mudanças e continuidades no impacto da desigualdade de renda ao longo do tempo, contribuindo para um entendimento mais completo da dinâmica educacional no país. A compreensão desses aspectos é fundamental para a formulação de políticas públicas que promovam a equidade na educação, ampliando o acesso e melhorando a qualidade do ensino.

Diante do contexto e da problemática levantada, esta pesquisa tem por objetivo geral de analisar a relação entre a renda familiar e o desempenho dos candidatos no Enem ao longo do tempo, identificando e compreendendo os efeitos da desigualdade socioeconômica no acesso à educação superior. Para tanto, são aplicadas técnicas e ferramentas de ciência de dados, especialmente a visualização de dados,

que possibilitam o desdobramento do objetivo geral, nos específicos, conforme tem-se: averiguar estatisticamente a relação entre os níveis de renda familiar e o desempenho dos estudantes nas provas do Enem, considerando diferentes edições do exame; identificar variações e tendências temporais no impacto da renda sobre o desempenho, observando padrões de desigualdade educacional ao longo das edições; interpretar os achados empíricos à luz de estudos e teorias sobre equidade educacional, fornecendo subsídios para o aprimoramento de políticas públicas de acesso ao ensino superior.

Além desta introdução, o artigo está estruturado em outras cinco seções. Na seção 2, são apresentados os estudos relacionados, contextualizando a relevância do tema e discutindo trabalhos prévios na área. A seção 3 descreve a metodologia adotada, incluindo a coleta, o pré-processamento dos dados e as técnicas de análise utilizadas. Na seção 4, são discutidos os resultados da pesquisa, destacando os principais achados e as implicações da relação entre renda e desempenho no Enem. Por fim, a seção 5 apresenta as considerações finais, ressaltando as contribuições do estudo, limitações e sugestões para pesquisas futuras.

2 FUNDAMENTAÇÃO TEÓRICA

Diante da relevância do Exame Nacional do Ensino Médio (ENEM) como instrumento de acesso ao ensino superior e de mensuração de desigualdades educacionais no Brasil, torna-se fundamental compreender os aspectos teóricos que embasam sua implementação, seus efeitos e suas limitações no enfrentamento das desigualdades socioeconômicas. Para isso, este referencial teórico abordará a trajetória do ENEM, seu papel nas políticas públicas de inclusão, a correlação entre desempenho e variáveis socioeconômicas — como renda e escolaridade dos pais — e os principais estudos que evidenciam essas relações. A seguir, serão apresentados autores e pesquisas recentes que contribuem para o entendimento crítico da eficácia do exame como política de equidade e seus desdobramentos na estrutura do ensino superior brasileiro.

2.2 ENEM: HISTÓRICO, POLÍTICAS PÚBLICAS, ESTUDOS E INDICADORES

O Enem, instituído em 1998 e consolidado como principal mecanismo de ingresso ao ensino superior a partir de 2009, desempenha papel central na democratização do acesso à graduação no Brasil, integrando-se a programas como Sisu, ProUni e Fies. Como observado por Travitzki (2023), o Enem contribui para ampliar a equidade educacional, especialmente quando articulado às políticas de ação afirmativa, embora sua eficácia dependa dessa articulação institucional.

A reformulação do exame buscou não apenas avaliar competências do ensino médio, mas estimular a reorganização curricular e incentivar a mobilidade acadêmica de estudantes provenientes de contextos desfavorecidos (Malusá; Ordones; Ribeiro, 2013). Com a criação do Sistema de Seleção Unificada (Sisu) em 2010 e a instituição da Lei de Cotas em 2012 (Lei nº 12.711), o ENEM passou a servir de critério essencial para ingresso público, vinculando renda, cor/raça e escolaridade pública como requisitos para reserva de vagas e promoção da equidade.

A atuação conjunta do Enem com programas como ProUni e Fies é vista como essencial para viabilizar o sonho da graduação para jovens de baixa renda: esses programas utilizam a pontuação do ENEM como base para oferecimento de bolsas e financiamentos, ampliando a inclusão no ensino superior privado e público. Apesar das melhorias observadas, estudos apontam que o exame isoladamente não garante equidade plena. Segundo análise da UFPR, as políticas de acesso precisam de uma base compensatória sólida, pois as desigualdades de base na educação básica ainda limitam a eficácia do ENEM como instrumento único de seleção (Oliveira, 2015). Essa base teórica contextualiza o ENEM como instrumento estratégico de inclusão educacional, articulado às políticas públicas de acesso e complementado por estrutura normativa inclusiva, permitindo compreender seu impacto social e sua função formativa no sistema de educação superior brasileiro.

No cenário educacional brasileiro, a renda familiar é um determinante crítico do desempenho no Enem, com estudantes de estratos mais altos apresentando desempenho significativamente superior. Dados levantados pela USP e Unifesp, com base em mais de 757 mil registros do Enem2019, revelam que alunos de classes A e B, frequentadores de escolas privadas e com pais de escolaridade superior alcançam notas médias de 130 a 140 pontos acima de estudantes negros, de baixa renda e de escolas públicas. Essa diferença reforça a existência de barreiras estruturais que implicam menor acesso aos cursos mais concorridos.

Estudos recentes, como o de Peres, Campos e Moraes (2025), indicam que há uma tendência clara: quanto maior a renda familiar e o grau de escolaridade da mãe, maiores as notas em Matemática no ENEM 2022, com média superior a 600 pontos entre candidatas cujas mães têm ensino superior. Essa evidência reforça a persistência de desigualdades socioeconômicas refletidas no desempenho acadêmico. Outro estudo conduzido por pesquisadores da USP e Unifesp reforça essa correlação, destacando que estudantes negros, de baixa renda, que frequentaram escolas públicas, obtêm média na faixa de 130 a 140 pontos inferior à de estudantes brancos, de renda média a alta, e que frequentaram escolas privadas.

Na concepção de Mariano et al. (2023), que analisaram o desempenho de estudantes da modalidade EJA e ensino regular em 2018 (utilizando Coarsened Exact Matching e propensão de escore), apontam desigualdades persistentes. Por meio dos estudos, os autores apontaram que alunos da EJA apresentam

desempenho inferior nas provas do exame, mesmo controlando variáveis socioeconômicas, indicando que o tipo de ensino impacta o resultado de forma independente.

Sob a perspectiva da análise longitudinal, o estudo pioneiro do LaPOpE/UFRJ (Bartholo; Castro; Klitzke e Carvalhaes, 2022) revelou que, entre 2013 e 2021, a diferença de participação no Enem entre jovens de nível socioeconômico muito baixo e muito alto dobrou de 11 para 23 pontos percentuais. Tal desigualdade se intensificou durante a pandemia, sendo mais acentuada nas regiões Sudeste e Norte do país. Outrossim, estudos socioeducacionais complementares, como os da lede com base no PIRLS (2021/2023), demonstram que 49% dos alunos de renda mais baixa permanecem abaixo do nível básico de leitura, enquanto 83,9% dos de renda elevada atingem níveis adequados. Essa disparidade de 58 pontos percentuais é a maior entre os países avaliados.

Brighenti e Lima (2023) analisaram o desempenho dos estudantes concluintes do ensino médio provenientes de escolas de Minas Gerais no Enem 2019. Os autores utilizaram estatística descritiva e consideraram as variáveis dependência administrativa escolar, sexo, raça, renda, escolaridade dos pais e notas, no qual, o tipo de dependência administrativa da escola (federal, estadual, municipal e privada) foi identificada como a variável mais influente no desempenho dos estudantes.

Feijó, França e Pinho Neto (2022) investigaram a relação entre a escolaridade dos pais e o desempenho educacional dos filhos tendo como fonte os microdados do Enem 2017, identificando que filhos de pai e mãe com escolaridade de nível superior obtiveram melhor desempenho do que os filhos de pais sem esse nível de escolaridade, principalmente em Redação e Matemática. Quando os pesquisadores controlaram variáveis como renda, tamanho da família, infraestrutura domiciliar e escola (efeito líquido), esta disparidade no desempenho entre os dois grupos foi mantida, mas com uma magnitude consideravelmente menor.

Por sua vez, Mariano, Silva, Santos e Benevides (2023) compreenderam que as disparidades no desempenho acadêmico entre estudantes da Educação de Jovens e Adultos (EJA) e do ensino regular no Exame Nacional do Ensino Médio (Enem) no Brasil. Utilizando dados do Enem e do Censo Escolar de 2018, empregou-se o algoritmo *Coarsened Exact Matching* (CEM) e o método de Pareamento por Escore de Propensão para isolar os efeitos da EJA de outros fatores que influenciam os resultados educacionais. Os resultados destacam que o ensino regular demonstra superioridade na promoção de habilidades cognitivas essenciais para a formação profissional e o desenvolvimento do capital humano, em comparação com a EJA.

Outro estudo procurou testar a eficácia de algoritmos de regressão linear múltipla para prever o desempenho dos estudantes utilizando os microdados referentes à edição de 2019 do exame (Novaes; Piton-Gonçalves, 2021). A amostra utilizada foi composta por 18.908 registros referentes a candidatos residentes na região metropolitana de Ribeirão Preto, São Paulo, com pontuação nas cinco provas,

desconsiderando os ausentes, treineiros e desclassificados. Foram propostos sete modelos de regressão múltipla, cada um com uma combinação distinta de variáveis escolhidas, e seis deles foram considerados ótimos pelos autores para prever o desempenho dos estudantes em cada uma das provas. A nota de Matemática foi a variável dependente em quatro modelos, os demais modelos buscam prever as notas de Linguagens e Códigos, Ciências da Natureza e Ciências Humanas.

Carmo, Heckler e Carvalho (2020) analisaram o desempenho médio dos candidatos do estado do Rio Grande do Sul no Enem 2019 através de recursos de visualização de dados e observaram que alunos provenientes das classes sociais A e B; oriundos de escolas privadas ou federais; brancos, pardos ou amarelos; que possuem acesso à internet em casa e cujos pais possuem maior grau de ensino; obtiveram melhor desempenho.

Percebe-se nos estudos citados, uma diversidade quanto à metodologia e variáveis analisadas, bem como em relação à amostra utilizada que se restringe a uma edição específica ou a determinado local de aplicação da prova. No presente artigo, procura-se, conforme metodologia descrita a seguir, realizar uma análise abrangente, que considere dados de todas as edições do exame.

3 METODOLOGIA

O presente estudo que tem como objetivo analisar a relação entre a renda familiar e o desempenho dos candidatos no Enem ao longo do tempo, identificando e compreendendo os efeitos da desigualdade socioeconômica no acesso à educação superior, apresenta-se como quantitativo e de natureza descritiva, voltado para a descrição, previsão e explicação, além de abordar dados mensuráveis ou observáveis, conforme definido por Sampieri, Collado e Lucio (2013) e Marconi e Lakatos (2022). As pesquisas descritivas, por sua vez, objetivam descrever as características de uma população, ou identificar relações entre variáveis (Marconi e Lakatos, 2022). No presente estudo o foco está em identificar padrões, tendências e correlações nos dados de forma objetiva, a partir de medidas estatísticas e técnicas de visualização de dados.

O público-alvo da pesquisa consiste nos candidatos que realizaram o Enem entre os anos de 1998 e 2023, sendo uma amostra composta por indivíduos de todas as regiões do Brasil que obtiveram nota maior do que zero em todas as provas, englobando diferentes perfis socioeconômicos e educacionais. Esse recorte permite uma análise abrangente e diversificada, considerando variáveis que refletem a heterogeneidade da população brasileira.

A coleta de dados deu-se a partir dos microdados disponibilizados pelo Inep em sua página de Dados Abertos, incluindo informações detalhadas sobre as notas dos candidatos, além de dados socioeconômicos

de indivíduos que responderam aos questionários aplicados durante o exame. As tabelas incluem as notas dos candidatos, bem como informações socioeconômicas. Ainda, a coleta de dados foi desenvolvida por meio de etapas, sendo que primeira etapa do estudo consistiu na obtenção dos dados através do *download* dos 26 arquivos do tipo *comma-separated values* (CSV) que contemplam as edições do exame desde sua criação em 1998 até o ano de 2023. Como houve mudanças no questionário socioeconômico e na metodologia da avaliação ao longo deste período, algumas tabelas apresentam divergência em sua estrutura, conforme Quadro 1.

Quadro 1 - Estrutura das tabelas de microdados

Ano	Número de Linhas	Número de Colunas	Ano	Número de Linhas	Número de Colunas
1998	150.944	163	2011	5.366.931	128
1999	308.959	156	2012	5.791.065	115
2000	378.421	154	2013	7.173.563	130
2001	1.609.886	276	2014	8.722.248	130
2002	1.820.404	252	2015	7.746.427	105
2003	1.873.079	222	2016	8.627.179	105
2004	1.542.165	239	2017	6.731.278	78
2005	2.991.271	257	2018	5.513.733	78
2006	3.743.370	258	2019	5.095.171	76
2007	3.584.577	258	2020	5.783.109	76
2008	4.018.050	258	2021	3.389.832	76
2009	4.148.720	339	2022	3.476.105	76
2010	4.611.614	105	2023	3.933.955	76

Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

A análise dos dados foi conduzida utilizando a linguagem de programação *Python*, em um ambiente *Jupyter Notebook*, com o auxílio das bibliotecas *Pandas* para manipulação de dados e *Matplotlib* para a criação das visualizações gráficas. Os dados referentes a cada uma das edições contêm um dicionário de dados, que a partir destes, foi realizado o mapeamento da variável de interesse e verificou-se a presença dela em cada uma das tabelas. O Quadro 2 contém o resultado deste mapeamento, detalhando o nome da variável referente à renda em cada uma das tabelas de microdados obtidas.

Quadro 2 - Mapeamento das variáveis referentes à renda

Ano	Nome da Variável	Ano	Nome da Variável
1998	Q120	2011	Q004
1999	Q24	2012	Q003
2000	Q23	2013	Q003
2001	Q15	2014	Q003
2002	Q15	2015	Q006
2003	Q22	2016	Q006
2004	Q22	2017	Q006
2005	Q23	2018	Q006
2006	Q23	2019	Q006
2007	Q23	2020	Q006
2008	Q23	2021	Q006
2009	Q21	2022	Q006
2010	Q04	2023	Q006

Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Para facilitar a consulta aos dados de interesse, diminuindo o tempo de processamento, visto que as tabelas totalizam 53,5 *gigabytes* e mais de 180 milhões de registros, optou-se pela concatenação dos registros e campos válidos em um novo arquivo CSV. Os nomes das variáveis também foram padronizados nesta etapa. O mapeamento feito anteriormente (Quadro 2) foi fundamental para a compilação dos dados de maneira automatizada.

Ainda, realizou-se a limpeza dos dados com a remoção de registros desnecessários referentes a candidatos com alguma nota inválida em alguma das provas e remoção dos campos desnecessários. Por conseguinte, criou-se um campo contendo a média aritmética das cinco notas válidas (Ciências da Natureza, Ciências Humanas, Linguagens e Códigos, Matemática e Redação) das edições a partir do ano de 2009. Nas edições anteriores, foi calculada a média das duas notas válidas (Objetiva e Redação). Este campo foi chamado de NOTA.

Como a prova passou por mudanças em sua estrutura e metodologia de cálculo da nota, não foi utilizado o valor real das médias dos candidatos. Para comparar o desempenho dos candidatos em cada edição, um campo com a nota normalizada em valores de 0 (equivalente à menor média válida da edição) a 100 (equivalente à maior nota válida), também foi criado. É importante destacar que este campo calculado objetiva facilitar a visualização da diferença de desempenho entre as categorias analisadas dentro de

uma edição, mas não é apropriado para comparar, por exemplo, se o desempenho dos candidatos em determinado ano foi melhor ou pior que o ano anterior. A nota normalizada foi calculada por meio da fórmula:

$$Nota Normalizada = \frac{(Nota Candidato - Nota Mínima)}{(Nota Máxima - Nota Mínima)} \times 100$$

Assim, como forma de demonstrar os dados, a tabela resultante ficou composta pelos campos *ANO*, *RENTA* e *NOTA_NORMALIZADA*.

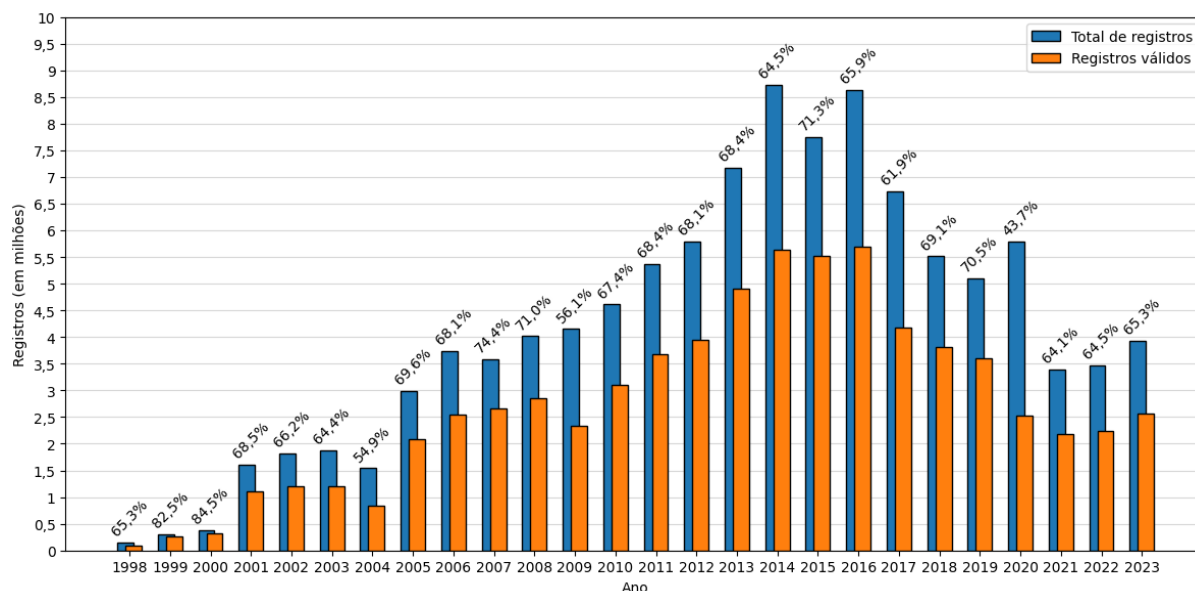
Ainda, cabe salientar, conforme exposto nos resultados deste estudo, a necessidade de exclusão dos microdados de 2010, devido à incompatibilidade das categorias de renda, que não permitiram uma padronização direta com os demais anos. Tal decisão foi necessária para preservar a consistência interna das análises comparativas, no qual reconhece-se, contudo, que essa exclusão pode comprometer parcialmente a análise longitudinal, uma vez que 2010 representa o primeiro ano pós-reformulação do Enem. Para minimizar essa limitação, avaliou-se a possibilidade de recategorização aproximada, mas a alta granularidade das faixas originais geraria perda de precisão e vieses interpretativos, motivo pelo qual se optou pela exclusão.

4 RESULTADOS E DISCUSSÕES

A partir das análises realizadas com base nos microdados do Enem, esta seção apresenta os principais resultados obtidos e discute suas implicações à luz da literatura especializada. Inicialmente, é apresentada uma análise geral sobre a amostra utilizada neste estudo e, logo em seguida, parte-se para a análise da influência da variável renda no desempenho dos estudantes.

Diante do contexto investigado, o número de registros que atendem ao critério utilizado neste estudo para a seleção da amostra (notas maiores do que zero em todas as provas), em cada ano, é detalhado na Figura 1. Percebe-se o aumento no número de candidatos a partir do ano de 2004, quando foi instituído o Prouni.

Figura 1 - Percentual de registros com notas válidas por ano

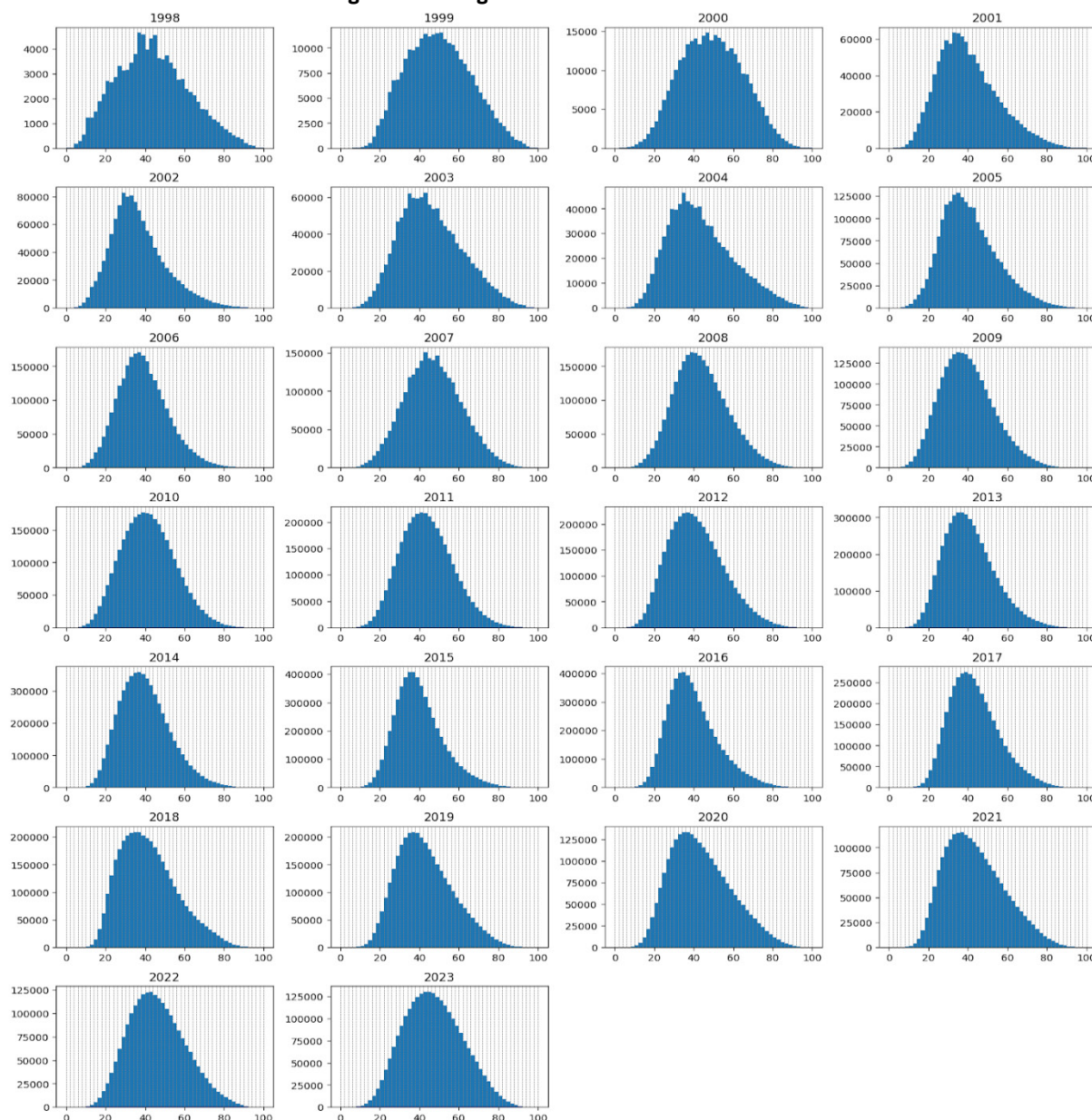


Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Observa-se também, por meio da Figura 1, uma redução drástica no percentual de registros válidos, que caiu de 70,5%, em 2019, para 43,7%, em 2020, em razão da pandemia de Covid-19. Neste ano, ocorreu a maior abstenção registrada no exame (Agência Brasil, 2021; Correio Braziliense, 2021).

Por conseguinte, os histogramas da Figura 2 representam as distribuições da nota normalizada em cada ano, no qual, percebe-se que a distribuição se manteve consistente ao longo das edições do exame, com pequenas variações na concentração dos valores. A maioria dos histogramas apresenta uma forma simétrica com um pico central, que se assemelha a uma distribuição normal (gaussiana), indicando que, em geral, as notas se concentram em torno de um valor médio, com poucos valores nas extremidades.

Figura 2 - Histogramas das notas normalizadas



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

No que tange a classificação da renda familiar dos candidatos, esta sofreu alterações ao longo das edições, como pode-se observar nos dicionários de dados. Assim, optou-se pela utilização da classificação utilizada no período de 1998 a 2009, por ser a mais resumida e, para compatibilidade com a classificação

adotada em 2011 e no período de 2012 a 2023, unificou-se no nível E todas as faixas superiores a 10 salários-mínimos. A classificação resultante é mostrada no quadro 3.

Quadro 3 - Classificação adotada para os níveis de renda (em salários-mínimos)

Nível	Renda familiar total
A	Até 1 salário
B	Mais de 1 a 2 salários
C	Mais de 2 a 5 salários
D	Mais de 5 a 10 salários
E	Mais do que 10 salários

Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Não foi possível adaptar a classificação de renda utilizada em 2010 (quadro 4), pois as faixas de renda utilizadas neste ano divergiram das classificações adotadas em outras edições. Sob este viés, salienta-se que a exclusão dos dados de 2010, necessária em razão da incompatibilidade das categorias de renda, também impacta a análise longitudinal. Embora a omissão não comprometa os padrões gerais observados, esta limita a avaliação das tendências no período imediatamente posterior à reformulação do Enem. Contudo, entende-se que, com a realização de pesquisas futuras poder-se-ia aplicar técnicas de recategorização proporcional ou agrupamentos por percentis de renda para incluir o ano de 2010, o que proporcionaria uma compreensão mais completa da evolução histórica. Logo, os dados deste ano não compuseram a análise.

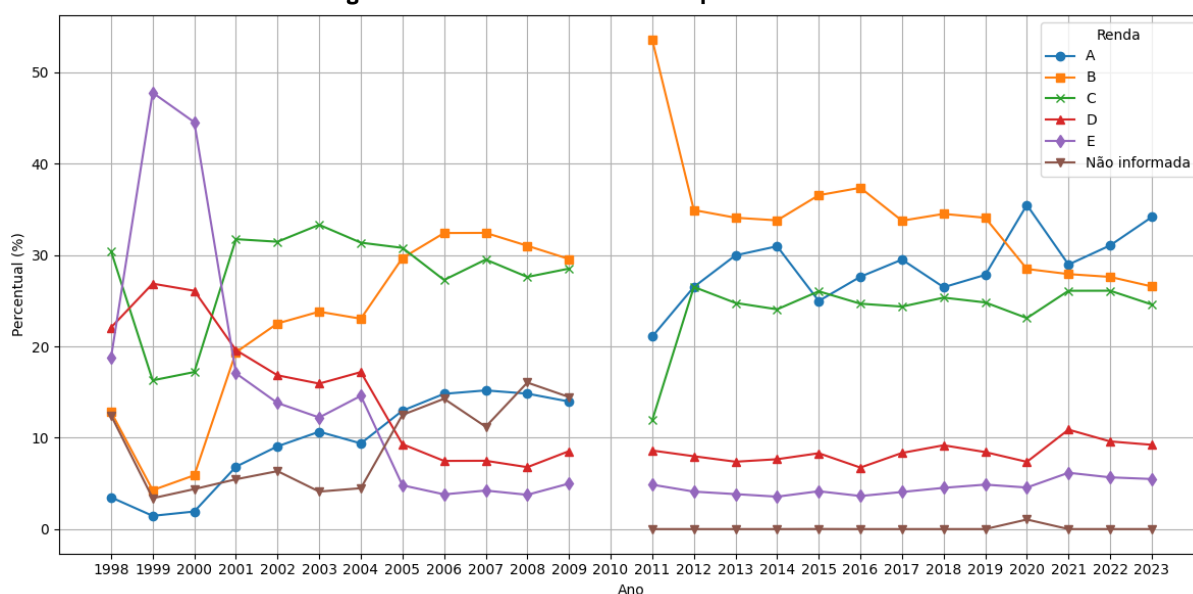
Quadro 4 - Classificação de renda utilizada no ano de 2010 (em salários-mínimos)

Nível	Renda familiar total
A	Até 1 salário
B	Mais de 1 a 3 salários
C	Mais de 3 a 6 salários
D	Mais de 6 a 9 salários
E	Mais de 9 a 12 salários
F	Mais de 12 a 15 salários
G	Mais do que 15 salários
H	Nenhuma renda

Fonte: Adaptado pelos autores a partir dos dicionários de dados do Enem (2024).

Ademais, observa-se na Figura 3 que houve, na amostra, bastante variação quanto à proporção de candidatos de cada nível de renda durante as edições, com uma tendência ao aumento de candidatos do nível A nos últimos anos, superando o nível B que representava a maioria desde 2006. Destarte, o número de candidatos com renda não informada em algumas edições, notadamente nos anos de 2006 e 2008, além da relativa baixa popularização do exame nos anos iniciais, podem ajudar a explicar a variação na distribuição.

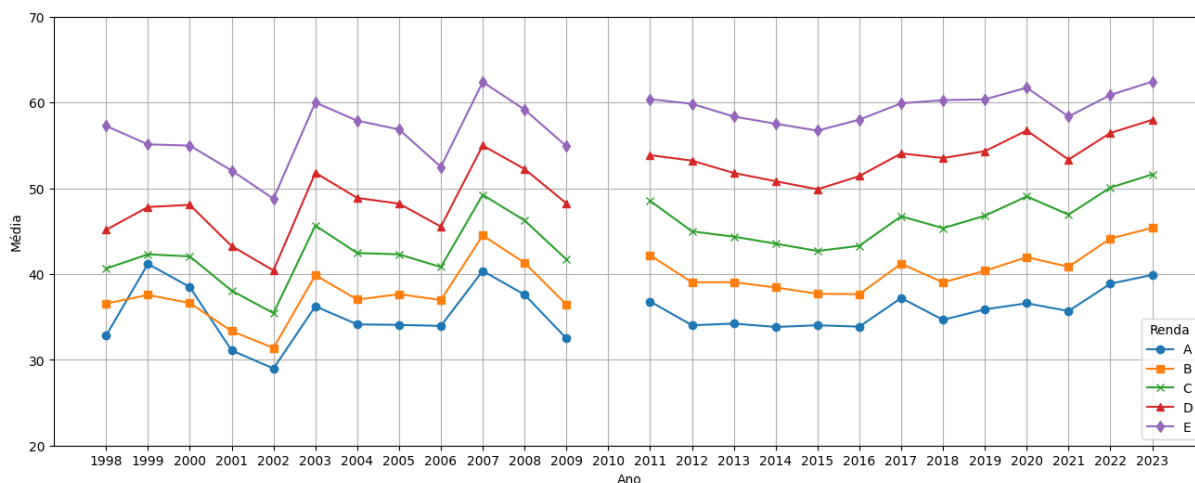
Figura 3 - Percentual de candidatos por faixa de renda



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Ainda, por meio da Figura 4, encontra-se ilustrado como foi o desempenho médio de indivíduos de cada um dos níveis de renda. Percebe-se clara estratificação, com o nível de desempenho relacionando-se diretamente ao nível de renda.

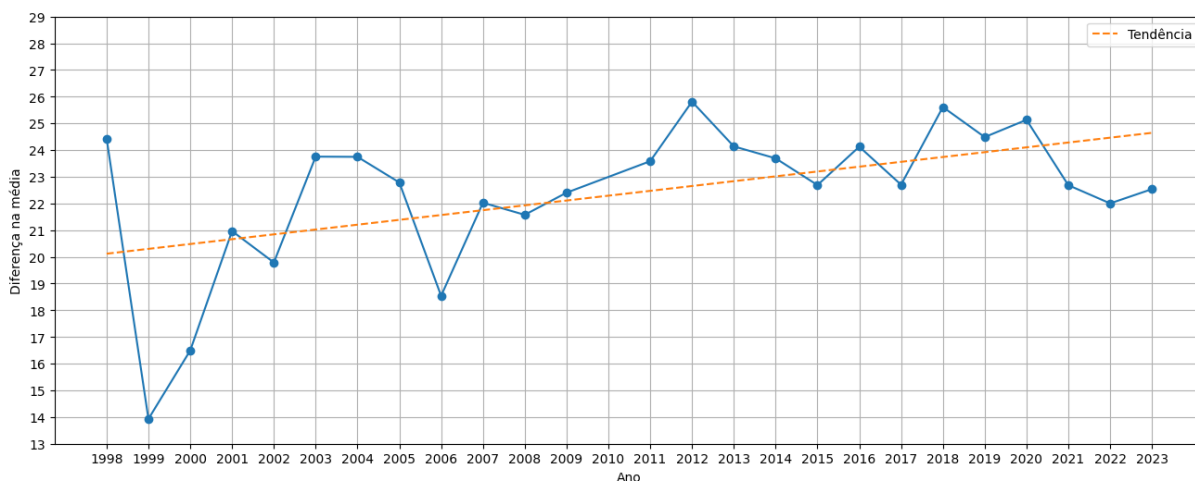
Figura 4 - Nota normalizada média por faixa de renda



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Em média, o desempenho dos candidatos do nível E, de maior renda, foi 22,39 pontos maior do que o dos candidatos do nível A. Na Figura 5 ilustra-se como comportou-se a diferença de desempenho entre estes níveis que ocupam extremidades opostas da escala.

Figura 5 - Diferença entre as médias dos níveis E e A em todo o período analisado

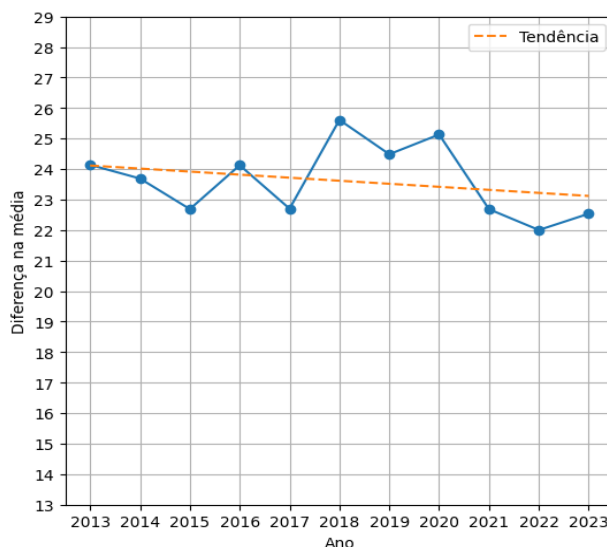


Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Através da técnica de regressão linear, foi estabelecida a linha de tendência para os valores da série. Quando todo o período é analisado (Figura 5), tem-se uma tendência ao aumento da superioridade de desempenho dos indivíduos do nível de renda E em relação aos do nível A, com coeficiente angular de

0,18. Quando os últimos 10 anos são analisados, observa-se uma ligeira tendência de diminuição, com coeficiente angular de $-0,1$; como é possível observar na Figura 6.

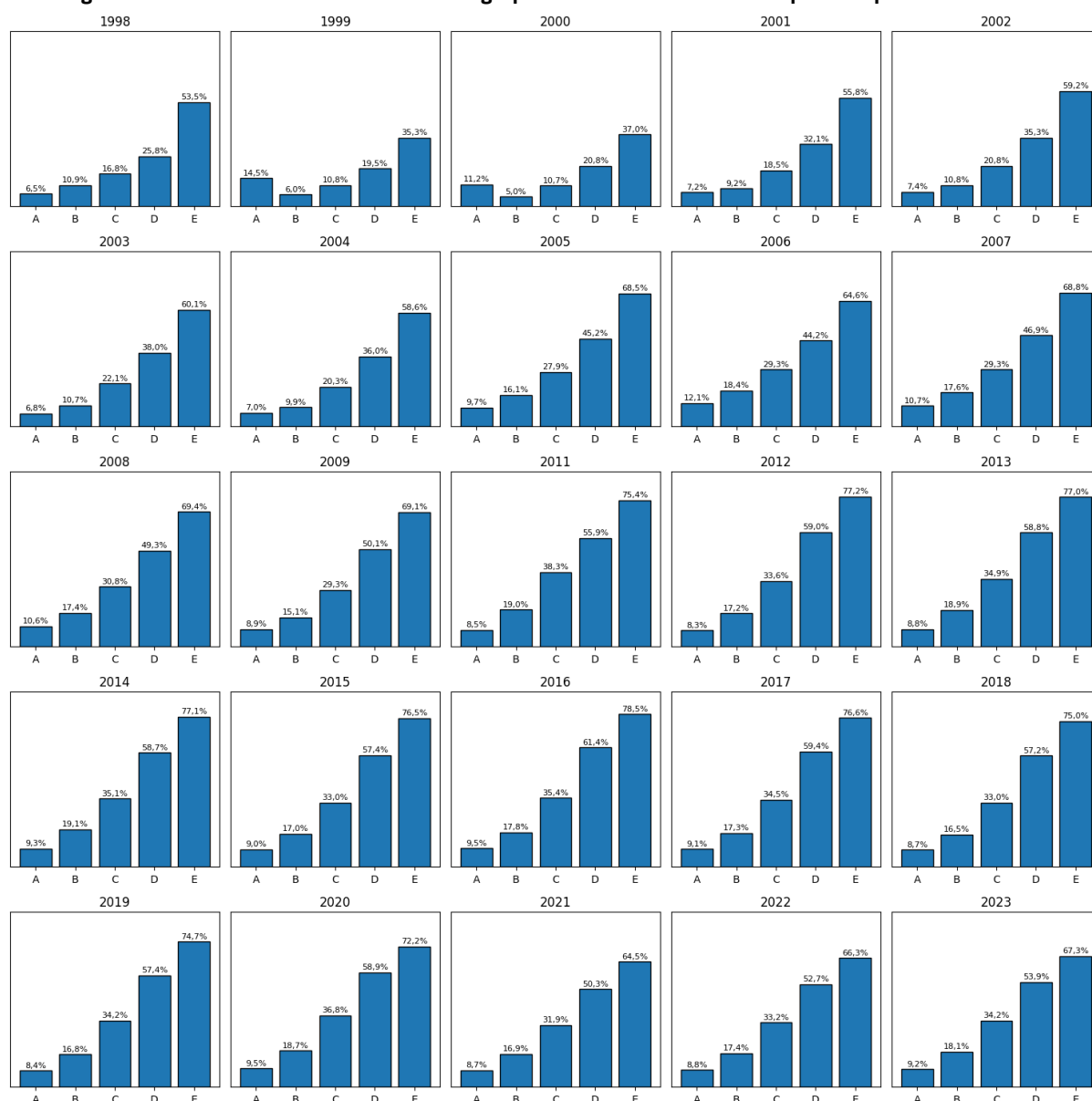
Figura 6 - Diferença entre as médias dos níveis E e A de 2013 a 2023



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Quando a análise é direcionada ao percentual de candidatos de cada grupo cuja nota normalizada estava incluída no quartil superior de cada ano, a desigualdade de rendimento entre os grupos é destacada, como é possível verificar nos gráficos da Figura 7.

Figura 7 - Percentual de candidatos de cada grupo com nota normalizada no quartil superior de cada ano



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

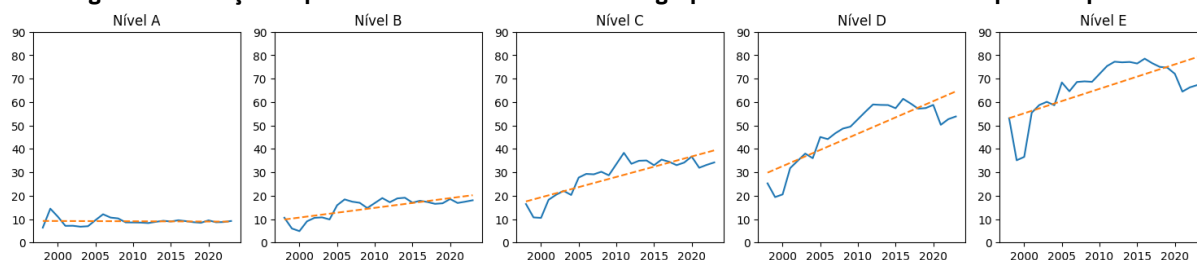
A análise da distribuição dos estudantes no quartil superior de desempenho reforça ainda mais essa desigualdade: os grupos com maior renda continuam a ocupar, majoritariamente, as faixas mais altas de desempenho, demonstrando o que Bartholo et al. (2022) identificaram como intensificação das desigualdades de acesso e permanência educacional, sobretudo após a pandemia. Embora a tendência de

desigualdade tenha diminuído levemente na última década, os coeficientes angulares ainda demonstram a permanência de uma lacuna relevante entre os extremos de renda.

Além disso, os dados revelam que, mesmo com a implementação de políticas de ação afirmativa e ampliação do acesso ao ensino superior, como o ProUni e o SisU, as disparidades de desempenho não foram significativamente reduzidas. Conforme Travitzki (2023), a eficácia dessas políticas depende de um conjunto mais amplo de medidas compensatórias desde a educação básica, o que não tem sido suficiente. Tal argumento é corroborado por Senkevics (2021), ao demonstrar que o aumento no acesso não implicou, necessariamente, redução na desigualdade educacional.

Para verificar as tendências nos dados apresentados na Figura 7, analisou-se a evolução de cada grupo aplicando-se regressão linear. Assim, conforme observa-se na figura 8, o nível de renda A teve leve tendência de diminuição na proporção de candidatos que conseguiram estar entre as 25% melhores notas de cada edição. O nível D foi o que apresentou maior tendência de aumento. Os coeficientes angulares obtidos foram -0,02; 0,41; 0,86; 1,38 e 1,03 para os níveis A, B, C, D e E, respectivamente.

Figura 8 - Evolução do percentual de candidatos de cada grupo com nota normalizada no quartil superior



Fonte: Adaptado pelos autores a partir dos dados do Enem (2024).

Do ponto de vista metodológico, os resultados obtidos pela aplicação de ciência de dados e regressão linear dialogam com estudos como os de Novaes e Piton-Gonçalves (2021), que apontam para a eficácia da modelagem estatística na predição de desempenho, mas que também destacam os limites do uso exclusivo de variáveis cognitivas para explicar a performance dos estudantes. Nesse sentido, a inclusão de variáveis socioeconômicas, como neste estudo, amplia a capacidade explicativa das análises.

Ainda, os resultados apresentados neste estudo confirmam uma tendência histórica já destacada pela literatura: a renda familiar é um fator determinante para o desempenho no Enem, expressando desigualdades estruturais e persistentes no sistema educacional brasileiro. A análise longitudinal revelou que, em todas as edições do exame entre 1998 e 2023 (exceto 2010), candidatos pertencentes ao grupo de maior renda (nível E) apresentaram desempenho médio superior aos de menor renda (nível A), com uma diferença média de 22,39 pontos. Essa constatação reforça o que Helene (2024) define como "círculo

vicioso da desigualdade”, no qual a baixa renda limita o acesso a oportunidades educacionais, e a baixa escolarização retroalimenta a exclusão econômica.

Embora as políticas de expansão das universidades federais e as ações afirmativas tenham contribuído para a redução da desigualdade de acesso ao ensino superior, aumentando, ainda que de maneira insuficiente, a participação dos mais pobres nas universidades (Senkevics, 2021), os dados demonstram que não houve redução na desigualdade de desempenho dos estudantes no Enem, conforme suas faixas de renda familiar. Isto indica a carência de políticas públicas eficazes para a melhoria da qualidade da educação básica dos mais pobres.

Contudo, os achados aqui apresentados reafirmam que a renda continua sendo uma variável central para a compreensão das desigualdades educacionais no Brasil. Conforme Carmo, Heckler e Carvalho (2020), os estudantes de classes sociais mais altas, oriundos de escolas privadas e com melhores condições socioeconômicas, apresentam melhor desempenho devido a um conjunto de fatores estruturais, como acesso à internet, suporte educacional e ambiente doméstico favorável ao estudo. Esses elementos não apenas aumentam as chances de sucesso no Enem, mas também evidenciam a urgência de políticas públicas intersetoriais voltadas à promoção da equidade educacional.

Por fim, ao considerar a teoria de Vigotsky (2007) sobre o papel da interação social e do contexto cultural no desenvolvimento cognitivo, pode-se inferir que o ambiente socioeconômico do estudante tem um papel significativo em sua trajetória educacional. O apoio social e cultural, tanto dentro quanto fora da escola, tem um impacto profundo no processo de aprendizagem. Para os estudantes de famílias de baixa renda, o acesso a um ambiente estimulante, com recursos educacionais adequados, é muitas vezes limitado, o que reflete diretamente em seu desempenho em exames como o Enem.

5 CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo analisar a influência da renda dos candidatos no desempenho obtido no Exame Nacional do Ensino Médio (Enem), a partir de uma perspectiva longitudinal que contemplou todas as edições do exame desde sua criação em 1998. Através da aplicação de técnicas de ciência de dados, com ênfase na visualização de dados e análise estatística, foi possível constatar que a renda familiar segue sendo um dos principais determinantes do desempenho dos estudantes no Enem, reproduzindo desigualdades históricas no acesso a uma educação de qualidade.

É importante ressaltar que o desempenho educacional é também influenciado por outros aspectos socioeconômicos e condições específicas de saúde e bem-estar dos estudantes que não foram abordados neste estudo. Além disso, a autodeclaração de renda pelos candidatos pode ser uma limitação

significativa, uma vez que depende da veracidade e precisão das informações fornecidas. Fatores como desconhecimento, erros de preenchimento do formulário socioeconômico ou até mesmo manipulação intencional podem introduzir viés nos dados.

Os dados demonstraram que estudantes pertencentes aos estratos de renda mais elevada obtêm, de forma recorrente, médias superiores em comparação àqueles de menor renda, configurando um padrão que se manteve praticamente inalterado ao longo de duas décadas. Essa constatação reforça a literatura especializada, que já indicava a forte correlação entre capital econômico, escolarização dos pais e oportunidades educacionais. A análise longitudinal também apontou que, embora políticas públicas de acesso ao ensino superior, como Sisu, ProUni e Fies tenham ampliado as possibilidades de ingresso, ainda são insuficientes para romper com os mecanismos estruturais de exclusão.

Com base nos resultados obtidos, entende-se a necessidade de proposições para a implementação de políticas públicas voltadas para a redução das desigualdades educacionais no Brasil, tais como ampliação de investimentos em infraestrutura para escolas públicas, especialmente em áreas periféricas e regiões mais carentes, para garantir o acesso a recursos educacionais essenciais, como computadores e internet; o fortalecimento de ações afirmativas, como as cotas para alunos de escolas públicas e de baixa renda, além de programas que garantam a permanência desses estudantes no ensino superior; e a capacitação contínua de professores, com foco em métodos de ensino inclusivos, que considerem a diversidade socioeconômica dos estudantes.

As proposições de políticas públicas aqui discutidas visam não apenas diminuir a disparidade no desempenho dos estudantes do Enem com base na renda, mas também estabelecer uma base sólida para a igualdade de oportunidades educacionais. É essencial que o governo continue avançando em sua agenda de inclusão educacional, investindo em educação de qualidade, suporte contínuo e ações afirmativas, para que o sistema educacional brasileiro não apenas amplie o acesso, mas também ofereça condições reais de sucesso a todos os estudantes, independentemente de sua classe social. O fortalecimento dessas políticas é crucial para garantir que as próximas gerações de estudantes tenham as ferramentas necessárias para alcançar seu pleno potencial e contribuir para o desenvolvimento do país.

Outrossim, conclui-se que o Enem, ainda que represente uma importante ferramenta de democratização do acesso à universidade, reflete e amplifica desigualdades já presentes na trajetória escolar dos estudantes brasileiros. Os resultados evidenciam a necessidade urgente de políticas públicas mais integradas e intersetoriais, que promovam a equidade desde a educação básica e assegurem condições mais igualitárias de preparação para os exames nacionais. Além disso, a análise reforça a importância do uso de dados educacionais abertos como instrumento de formulação de políticas públicas baseadas em evidências.

Contudo, o conjunto das evidências reforça que a desigualdade econômica e educacional está intrinsecamente ligada ao desempenho escolar e à reprodução social no Brasil. A análise longitudinal dessas variáveis, aliada ao uso de técnicas de ciência de dados (visualização, regressão, *matching*), permite compreender padrões sistemáticos e propor políticas públicas pautadas na equidade e inclusão.

Nesta esteira de pensamento, entende-se que este estudo apresentou alguns desafios importantes, visto que, os dados de renda são autodeclarados pelos candidatos, o que pode introduzir vies decorrente de erros, omissões ou respostas intencionais inverídicas. Além disso, não foram incluídas variáveis contextuais relevantes, como infraestrutura escolar, tempo de estudo, apoio familiar, capital cultural e acesso a recursos pedagógicos, que podem influenciar significativamente o desempenho, bem como a falta de variáveis que possibilitem análises frente a diversidade regional brasileira, que apresenta grandes disparidades educacionais e socioeconômicas. Essas limitações foram consideradas neste estudo, pois seriam, caso o questionário apresentasse, consideradas na interpretação dos resultados e abrem caminhos para pesquisas futuras mais abrangentes e integradas.

Como sugestões para estudos futuros, recomenda-se o aprofundamento da análise com a inclusão de variáveis qualitativas, como acesso a recursos pedagógicos, histórico de repetência e tempo de estudo diário, além da aplicação de modelos estatísticos multivariados que considerem interações entre fatores socioeconômicos e institucionais. Ademais, propõe-se que investigações futuras explorem os efeitos de políticas públicas específicas sobre os grupos mais vulneráveis, bem como a eficácia de programas de permanência no ensino superior no combate às desigualdades educacionais. Por fim, também se sugere, para trabalhos futuros, a adoção de técnicas de recategorização proporcional, como agrupamentos por percentis de renda, permitindo a integração mais precisa dos microdados de 2010 sem comprometer a comparabilidade.

REFERÊNCIAS

AGÊNCIA BRASIL. **Segundo dia de Enem tem abstenção de 55,3%**. Brasília, 18 jan. 2021. Disponível em: <https://agenciabrasil.ebc.com.br/educacao/noticia/2021-01/segundo-dia-de-Enem-tem-abstencao-de-55%2C3>. Acesso em: 23 jul. 2024.

BARTHOLO, T.; CASTRO, J. A. R. de; KLITZKE, J.; CARVALHAES, A. **Desigualdade de participação no Enem: uma análise longitudinal (2013–2021)**. Rio de Janeiro: Laboratório de Políticas Públicas Educacionais (LaPOpE/UFRJ), 2022.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. **Base de dados do ENEM**. Disponível em: <https://www.gov.br/inep/pt-br>. Acesso em: 20 jul. 2024.

BRIGHENTI, A. M.; LIMA, M. F. Análise do desempenho dos estudantes de escolas públicas e privadas no Enem 2019: um estudo em Minas Gerais. **Revista Brasileira de Educação**, v. 28, 2023.

CARMO, D. G.; HECKLER, M. V.; CARVALHO, M. D. M. Desempenho dos estudantes gaúchos no Enem 2019: uma análise visual dos dados. **Revista Educação e Realidade**, v. 45, n. 4, p. 1–22, 2020.

CARMO, R. V. do; HECKLER, W. F.; CARVALHO, J. V. de. Análise do desempenho dos estudantes do Rio Grande do Sul no Enem 2019. **Revista RENOTE**, v. 18, n. 2, p. 378–387, 2020. DOI: <https://doi.org/10.22456/1679-1916.110257>.

CORREIO BRAZILIENSE. **Enem 2020 tem a maior abstenção da história**; mais de 50% não fizeram a prova. Brasília, 17 jan. 2021. Disponível em: <https://www.correiobraziliense.com.br/cidades-df/2021/01/4902363-Enem-2020-tem-a-maior-abstencao-da-historia-mais-de-50--nao-fizeram-a-prova.html>. Acesso em: 23 jul. 2024.

FEIJÓ, J. R.; FRANÇA, J. M. S. D.; PINHO NETO, V. R. D. Desempenho dos estudantes ao final do ensino médio: mensurando a influência direta e indireta da educação dos pais. **Revista Brasileira de Economia**, Rio de Janeiro, v. 76, n. 1, p. 30–56, 2022. DOI: <https://doi.org/10.5935/0034-7140.20220003>.

FEIJÓ, V. M.; FRANÇA, S. P.; PINHO NETO, J. D. Desigualdade educacional e escolaridade dos pais: uma análise com dados do Enem 2017. **Revista Brasileira de Estudos Educacionais**, v. 27, n. 68, p. 56–78, 2022.

HELENE, A. Educação e desigualdade: o círculo vicioso da renda no Brasil. **Cadernos de Pesquisa**, São Paulo, v. 54, 2024.

HELENE, O. **Desigualdades econômicas e educacionais, um cruel círculo vicioso**. Jornal da USP, São Paulo, 9 maio 2024. Disponível em: <https://jornal.usp.br/?p=754279>. Acesso em: 23 jul. 2024.

IEDE – INTERDISCIPLINARIDADE E EVIDÊNCIAS NO DEBATE EDUCACIONAL. **Desigualdades na alfabetização em leitura entre crianças brasileiras**: evidências do PIRLS 2021. São Paulo, 2023. Disponível em: <https://iedebate.org.br>. Acesso em: 23 jul. 2025.

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Dados abertos**: microdados do Enem. Brasília, 2024a. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/Enem>. Acesso em: 23 jul. 2024.

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Exame Nacional do Ensino Médio (Enem)**. Brasília, 2024b. Disponível em: <https://www.gov.br/Inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/Enem>. Acesso em: 23 jul. 2024.

LIMA, C. C. V. de; BRIGHENTI, C. R. G. Desempenho de estudantes de Minas Gerais no Exame Nacional do Ensino Médio considerando variáveis socioeconômicas. **Educação e Pesquisa**, São Paulo, v. 49, e253303, 2023. DOI: <https://doi.org/10.1590/S1678-4634202349253303>.

MALUSÁ, D.; ORDONES, C.; RIBEIRO, L. O Exame Nacional do Ensino Médio (Enem) e as mudanças no ensino médio: uma proposta de reestruturação curricular. **Revista Educação Pública, Rio de Janeiro**, v. 13, n. 2, 2013.

MARCONI, M. A.; LAKATOS, E. M. **Metodologia científica**. 8. ed. São Paulo: Atlas, 2022.

MARIANO, D. L. et al. Diferenças no desempenho acadêmico entre estudantes da EJA e do ensino regular: uma análise do Enem. **Educar em Revista**, Curitiba, v. 39, 2023.

MARIANO, F. Z.; SILVA, M. C.; SANTOS, M. M.; BENEVIDES, A. A. Desigualdade de desempenho no ensino médio: evidências sobre a educação de jovens e adultos. **Nova Economia**, Belo Horizonte, v. 33, n. 3, p. 719–747, 2023. <https://doi.org/10.1590/0103-6351/7950>

NOVAES, A. A.; PITON-GONÇALVES, J. Predicting student performance by multiple regression. **Sigmae**, Alfenas, v. 10, n. 1, p. 82–98, 2021.

OLIVEIRA, D. A. Políticas públicas de acesso ao ensino superior: inclusão ou reforço da desigualdade? **Revista da Associação Nacional de Pós-Graduação e Pesquisa em Educação – ANPEd**, Niterói, v. 23, n. 67, p. 55–75, 2015.

PERES, M.; CAMPOS, T.; MORAES, R. Desigualdades educacionais no Enem 2022: a influência da renda e escolaridade materna no desempenho em Matemática. **Revista Brasileira de Políticas Educacionais**, Brasília, v. 10, n. 1, 2025 (no prelo).

PPUFU. Plataforma de Produção Universitária da Universidade Federal do Uruguai. **Base de dados do Enem 2022**. Disponível em: <https://www.ppufu.edu.uy>. Acesso em: 20 jul. 2024.

SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, P. B. **Metodologia de pesquisa**. 5. ed. Tradução de Daisy Vaz de Moraes. Porto Alegre: Penso, 2013.

SENKEVICS, A. S. **O acesso, ao inverso:** desigualdades à sombra da expansão do ensino superior brasileiro, 1991-2020. 2021. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2021. DOI: 10.11606/T.48.2021.tde-11012022-103758.

TRAVITZKI, R. O papel do Enem na democratização do ensino superior: avanços e limites das políticas afirmativas. **Revista Brasileira de Educação**, São Paulo, v. 28, e028069, 2023. <https://doi.org/10.1590/s1413-2478202328069>

UNIFESP; USP – Universidade Federal de São Paulo; Universidade de São Paulo. **Desempenho no Enem e desigualdade:** uma análise baseada nos microdados de 2019. São Paulo: Relatório Técnico, 2023.

VIGOTSKY, L. S. **A formação social da mente:** o desenvolvimento dos processos psicológicos superiores. 7. ed. São Paulo: Martins Fontes, 2007.