

DESEMPENHO EDUCACIONAL E DESIGUALDADE SOCIAL: UM ESTUDO DE SUA RELAÇÃO ATRAVÉS DA ANÁLISE DE COMPONENTES PRINCIPAIS (ACP)

Lara Carla Freitas Rodrigues ¹
Marília Magalhães Maia ²
Sílvio Roberto Fernandes ³
Luiza Helena Félix de Andrade ⁴

RESUMO

Não é novidade que a problemática da manutenção das desigualdades sociais no Brasil, indicada pelas pesquisas publicadas em 2023 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), afeta de formas múltiplas os indivíduos que compõe a sociedade brasileira e isso engloba todo o contexto educacional em que estão submersos os estudantes brasileiros. Por isso, o presente artigo tem como objetivo analisar quais são as principais características socioeconômicas que influenciaram a proficiência dos estudantes que realizaram o Exame Nacional do Ensino Médio (ENEM) em 2022, já que o ENEM se tornou a principal porta de entrada para as instituições de nível superior no país e o desempenho de cada estudante nessa prova é um divisor de águas em sua trajetória acadêmica, profissional e social. Com esse propósito, primeiro foi necessário calcular a média geral das notas de 2022. O próximo passo foi a formação de dois grandes grupos: o primeiro com alunos com a média igual ou superior à média geral e o segundo com alunos com a média inferior à média geral. Finalmente, para cada um dos grupos, o método da Análise da Componente Principal (ACP) foi aplicado às informações do questionário socioeconômico com a finalidade de identificar a correlação entre cada pergunta deste questionário com a proficiência dos estudantes de cada grupo. Os resultados do estudo revelaram que a renda foi o principal fator determinante para os dois grupos, seguido do tipo de escola para o grupo de menor média e da quantidade de banheiros na casa para o grupo de maior média, demonstrando de forma explícita que o desempenho dos estudantes está, em sua maior porção, ligado ao contexto social e econômico em que estão inseridos e não só em sua dedicação individual.

Palavras-chave: ENEM, Proficiência, ACP, Desigualdade, Educação.

INTRODUÇÃO

Desde a antiguidade a humanidade busca entender quais os principais fatores que originam e contribuem com as desigualdades sociais. De fato, quando se fala sobre esse assunto, não há como não admitir que esse tema a muito tempo é debatido, porém nunca

¹Graduada pelo Curso de Ciência e Tecnologia da Universidade Federal Rural do Semi-Árido- Ufersa, lara.rodrigues@alunos.ufersa.edu.br;

²Mestranda do Curso de Engenharia de Teleinformática da Universidade Federal do Ceará- UFC, mariliamagmaia@alu.ufc.br.

³Doutor em Sistemas e Computação, Universidade Federal do Rio Grande do Norte - UFRN, silvio@ufersa.edu.br;

⁴Doutora em Engenharia de Teleinformática, Universidade Federal do Ceará- UFC, luizafelix@ufersa.edu.br;

solucionado. Prova disso são os recentes dados publicados pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em agosto de 2023, que investigaram o caráter multidimensional da pobreza e da vulnerabilidade no Brasil entre os biênios 2008-2009 e 2017-2018. O estudo indicou uma queda nos níveis de pobreza e vulnerabilidade no país na comparação entre esses anos. Por outro lado, as porcentagens indicaram a manutenção da desigualdade social no país, demonstrando que os mais afetados com a pobreza e a vulnerabilidade continuam sendo os grupos menos favorecidos historicamente, como os pretos, pardos, as mulheres e os trabalhadores rurais, reforçando assim o caráter estrutural da desigualdade no Brasil.

Sem dúvidas, é notório que a desigualdade é um cenário cotidiano na sociedade brasileira e, no contexto da educação, isso não é diferente. Como destaca o patrono da educação Paulo Freire (2005), a educação ao longo dos anos foi utilizada como prática de dominação das massas, sendo por muito tempo um privilégio da elite, onde apenas famílias ricas e brancas tinham condições de oferecer acesso a educação aos seus descendentes.

Porém, é possível observar que, com o passar do tempo, várias mudanças foram implementadas na educação do país a fim de torná-la mais acessível para todos. Um bom exemplo disso foi a criação, em 1996, da Lei de Diretrizes e Bases da Educação Nacional (LDBEN) com a finalidade de assegurar a formação básica para o povo. Além disso, em 1998, houve a reforma do ensino médio, de acordo com as Diretrizes Curriculares Nacionais do Ensino Médio, e a criação de métodos de avaliação da educação, como o Exame Nacional do Ensino Médio (ENEM). Esse último, no entanto, passou a ser a principal porta de entrada dos estudantes nas instituições de ensino superior e, diante do cenário de desigualdades, passou a contar com diversos mecanismos de compensação, como a Lei das Cotas (Lei nº 12.711), sancionada em 2012, que disponibiliza a reserva de vagas para diversas pessoas presentes em grupos de vulnerabilidade social, como aqueles estudantes de baixa renda, autodeclarados pretos, pardos e indígenas, pessoas portadoras de deficiências, além dos alunos advindos de escolas públicas.

Mesmo assim, o cenário de desigualdade continua. Por isso, o presente artigo tem como objetivo investigar, a partir dos dados do ENEM 2022, as principais características socioeconômicas que afetaram o desempenho dos estudantes que realizaram o exame em 2022. Essa análise foi feita com o auxílio do método da Análise das Componentes Principais (ACP ou PCA), após a aplicação de diversos filtros e a constatação da média geral dos estudantes para aquele ano.

Em vista disso, o presente trabalho apresentará primeiro uma breve explicação sobre a metodologia aplicada na pesquisa, posteriormente uma explicação direta sobre a importância do banco de dados do ENEM e o funcionamento do método ACP. Em seguida, os resultados serão apresentados, discutidos e seguidos pelas conclusões alcançadas.

METODOLOGIA

A presente pesquisa, é um estudo de natureza quantitativa, visto que, como define Babbie (2022), busca analisar os dados de maneira objetiva e estatística, além de utilizar técnicas matemáticas e estatísticas para examinar relações entre as variáveis e encontrar padrões através das amostras de uma população. Os dados utilizados nesse estudo foram os microdados referentes ao ENEM 2022, disponibilizados pelo Instituto Nacional de Ensino e Pesquisa Anísio Teixeira (INEP). Os microdados são amostragens balanceadas e que representam bem o conjunto completo de dados. Além disso, o ano de 2022 foi escolhido devido a obrigatoriedade da resposta do questionário durante a inscrição, o que colabora para uma maior robustez da pesquisa, fato esse que não aconteceu nos anos anteriores.

Neste trabalho, não foram utilizados todos os dados disponibilizados no banco de dados do INEP, visto que o foco da análise é a relação entre os fatores socioeconômicos e o desempenho dos estudantes, alguns filtros foram implementados para eliminar informações desnecessárias ou incompletas para esse estudo. A princípio, foram retirados os participantes que não fizeram a redação, faltaram algum dia de prova e os treineiros, ou seja, aqueles que ainda não tinham concluído o ensino médio e estavam realizando a prova por experiência. Além disso, já que alguns dados de respostas eram expressos em números e outros em letras, foi necessário padronizar as informações em um só formato, o de números, para que o algoritmo desenvolvido na linguagem Python conseguisse ser implantado. Sendo assim, o que antes eram as letras A,B,C,...,Q, se tornaram, na sequência, 1, 2 ,3,...,17.

Posteriormente, a fim de dividir os dados filtrados em dois grandes grupos, um de maior desempenho e outro de menor desempenho, foi calculada a média geral, levando em conta as notas obtidas nas 5 áreas do exame: Ciências Humanas, Ciências da Natureza, Linguagens, Códigos, Matemática e a Redação. Após o resultado, aqueles estudantes com a média igual ou maior ao valor obtido foram colocados no grupo de maior desempenho

e os que obtiveram a nota menor que a média geral foram colocados no grupo de menor desempenho. Esses dados foram organizados em 2 matrizes, chamadas aqui de A, a de maior desempenho, e B, a de menor desempenho.

Após a separação, como não seria mais necessário a informação das notas e médias, já que os grupos já estavam divididos, e com o objetivo de diminuir o esforço computacional, um segundo filtro foi aplicado aos dados retirando as colunas das notas de cada prova e a média de cada estudante. Sobrando assim 37 colunas referentes a faixa etária, sexo, estado civil, cor/raça, nacionalidade, situação de conclusão, ano que concluiu o ensino médio, escola, ensino, dependência administrativa, localização, situação de funcionamento e as 25 questões do questionário socioeconômico que podem ser obtidos do site do INEP (Brasil, 2023).

Com os dados preparados, foi possível utilizá-los no algoritmo ACP para obter a indicação dos fatores socioeconômicos que mais impactaram no desempenho dos estudantes. Finalmente, em posse dos resultados do ACP, foi feito um estudo mais profundo sobre as questões mais relevantes do questionário, de acordo com a ACP, e como as suas respostas estão distribuídas em cada grupo.

REFERENCIAL TEÓRICO

O ENEM foi um exame construído com o objetivo de avaliar os estudantes logo após a finalização do período da educação básica. Em seu primeiro ano, somente duas universidades o utilizavam, 91% dos participantes vinham de escolas privadas e somente 9% de escolas públicas. Já em 1999, o exame ganhou mais credibilidade e passou a ser aceito como porta de entrada de 93 instituições de ensino superior. Em 2009, com a criação do Sistema de Seleção Unificada (SISU), o ENEM mudou de formato, a prova passou a ter 180 questões objetivas e a redação, além de começar a certificar a conclusão do ensino médio, porém, o que permaneceu foi a presença do questionário socioeconômico. Já em 2013, o ENEM passou a ser o critério de seleção de todas as instituições federais de educação superior. Outra mudança importante ocorreu em 2022 quando a resposta do questionário socioeconômico passou a ser obrigatória para que a inscrição fosse aceita.

Com certeza, todas essas alterações, atreladas a política de inclusão com a insenção da taxa de inscrição para aqueles que se enquadravam como baixa renda, fizeram com que o ENEM se tornasse a principal ferramenta para o ingresso na educação superior,

fazendo com que em 2022 mais de 3 milhões de inscrições fossem contabilizadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

Concomitante a isso, os dados extraídos a partir do desempenho dos participantes e sua situação socioeconômica revelam um panorama bem amplo da realidade brasileira quanto as desigualdades sociais e educacionais, fato que contribui para que o ENEM seja um objeto de uso não só daqueles que estão iniciando seu nível superior, mas também de grandes pesquisadores os quais utilizam diversos métodos para retirar informações importantes do banco de dados disponibilizado pelo INEP.

Nesse cenário, devido ao grande número de dados e sua complexidade, é comum serem utilizadas algumas ferramentas de análise e otimização de dados para identificar as variáveis mais relevantes. Dentre elas, está o método intitulado como “Análise das componentes Principais” (ACP ou PCA) que foi o principal instrumento utilizado neste trabalho.

A ACP, inicialmente descrita por Pearson (1901) e posteriormente detalhada por Hotelling (1933, 1936), é uma técnica estatística de análise multivariada. Seu principal objetivo é transformar o conjunto original de variáveis, por vezes grandes e complexos, em um novo conjunto de variáveis substancialmente menor de variáveis linearmente independentes (Boldrini, 1980), que preservam a maior parte da informação original, se desfazendo as informações repetidas e identificando padrões nos dados. Essas novas variáveis são chamadas de componentes principais.

Cada componente principal apresenta propriedades importantes: são independentes entre si e são estimados de forma a maximizar, em ordem, a quantidade de informação retida. O que indica que o primeiro componente principal é o que carrega a maior parte da variação dos dados, o segundo carrega a segunda maior variação, e assim sucessivamente. A seguir, serão apresentadas as etapas de construção da ACP.

Etapa 1: Organização da Matriz de dados X. Considere a situação em que é possível observar ‘c’ características de ‘i’ indivíduos de uma população J. As características são representadas pelas variáveis $X_1, X_2, X_3, \dots, X_c$. A matriz de dados X é de ordem ‘c x i’ .

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1c} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ic} \end{bmatrix} \quad (1)$$

Etapa 2: Calcular a matriz de covariância. O primeiro passo é centralizar o dados. Isso significa subtrair a média de cada coluna de todos os valores dessa coluna, garantindo uma média zero para cada variável. Posteriormente, calcula-se a covariância, ela mede o grau de interdependência entre duas variáveis e é calculada pela Equação (2).

$$Cov(X_1, X_2) = \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_{11})(x_{i2} - \bar{x}_{12})}{n-1} \quad (2)$$

Sendo 'n' a quantidade de itens observados ou a população, caso não se trabalhe com amostra, 'x_{i1}' e 'x_{i2}' o valor dos dados, 'x̄₁₁' e 'x̄₁₂' a média dos valores da primeira e segunda variável, respectivamente. A matriz de covariância se apresenta da seguinte forma:

$$S = \begin{bmatrix} var(X_1, X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_c) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_i, X_1) & cov(X_i, X_2) & \cdots & var(X_i, X_c) \end{bmatrix} \quad (3)$$

A diagonal da matriz de covariância apresenta o resultado da covariância de uma variável por ela mesma, que resulta na variância. Essa etapa é muito importante, pois permite que ocorra a transformação dos dados originais na forma de um novo conjunto.

Etapa 3: Determinação dos componentes principais. Os componentes principais são o resultado da aplicação da equação característica da matriz S, ou seja, são os autovetores da matriz de covariância, isto é:

$$det[S - \lambda I] = 0 \quad (4)$$

Sendo a estrutura dos componentes descrita por:

$$V_i = v_{i1}X_1 + v_{i2}X_2 + \cdots + v_{in}X_n \quad (5)$$

Os componentes principais apresentam algumas propriedades. A variância do componente principal V_i é igual ao altovalor λ_i (Varella, 2008). Como na matriz de covariância a os autovetores tem norma igual a 1, é o seu autovalor que indicará o real tamanho do vetor (Varella, 2008). Assim, a ordenação dos CP ocorre de forma decrescente obedecendo os valores dos autovalores, sendo o CP1 o de maior autovalor, o

CP2 o segundo de maior autovalor e assim por diante (Varella, 2008). O total de variância das variáveis originais é igual ao somatório dos autovalores que representa o mesmo resultado da soma da variância das componentes principais e os componentes não são relacionados entre si (Varella, 2008). Além disso, os CP demonstram a relevância de cada variável para o conjunto como um todo, o que possibilita identificar os fatores que mais impactam, no caso desse estudo, os estudantes no grupo de maior e menor desempenho.

Etapa 4: Contribuição de cada componente principal. Por fim, calcularemos a seguir a medida que representa a contribuição de cada CP na variância dos dados do conjunto original (1). A Equação (6) demonstra como é feito seu cálculo, destacando que seu resultado é em porcentagem.

$$C_i = \frac{Var(V_i)}{\sum_{i=1}^n Var(V_i)} \times 100 = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \times 100 \quad (6)$$

RESULTADOS E DISCUSSÃO

Apesar da grande quantidade de dados referentes ao ano de 2022 do ENEM, como a análise foi realizada somente para esse ano, não foi necessário fazer nenhum processo de amostragem. O estudo avaliou todos os dados obtidos depois da série de filtragens descritas na metodologia.

Após a limpeza dos dados, o ponto de referência para a separação dos dois grupos a serem avaliados, o de maior desempenho e o de menor desempenho, foi a média geral calculada a partir da média das notas totais de cada estudante. O resultado foi de 551,35 pontos, com desvio padrão de 101,36 pontos, o que demonstra que as notas variam significativamente, com coeficiente de variação de 18,38%, fato que sugere que essa variação é moderada em relação a média das notas. A Figura 01 exemplifica como as matrizes ficaram organizadas.

As tabelas possuem 37 colunas, como estão descritas na seção sobre a metodologia deste trabalho e as linhas representam as respostas de cada estudante. O Quadro 01 traz a identificação de cada questão de acordo com o dicionário disponibilizado pelo INEP. Já nesse primeiro momento, é possível analisar que a maior parte dos estudantes, cerca de 53,42% está no grupo de baixo desempenho e, aproximadamente, 46,58% está no de maior desempenho.

Figura 01 – Modelo final das matrizes

TP_FAIXA_ETARIA	TP_SEXO	TP_ESTADO_CIVIL	TP_COR_RACA	TP_NACIONALIDADE	TP_ANO_CONCLUIU	TP_ESCOLA	TP_ENSINO	TP_DEPENDENCIA_ADM_ESC	...
5	6	1	2	1	2	1	0.0	0.0	...
2	6	1	1	1	0	3	1.0	0.0	...
8	6	1	2	1	5	1	0.0	0.0	...
4	13	1	3	1	1	1	0.0	0.0	...
3	6	1	3	1	1	1	0.0	0.0	...
...
5	6	1	1	1	4	1	0.0	0.0	...
3	13	1	1	1	1	1	0.0	0.0	...
12	13	1	3	1	16	1	0.0	0.0	...
11	6	1	2	1	11	1	0.0	0.0	...
2	13	1	1	1	0	2	1.0	2.0	...

Fonte: Autoria Própria.

Quadro 01- Descrição das variáveis

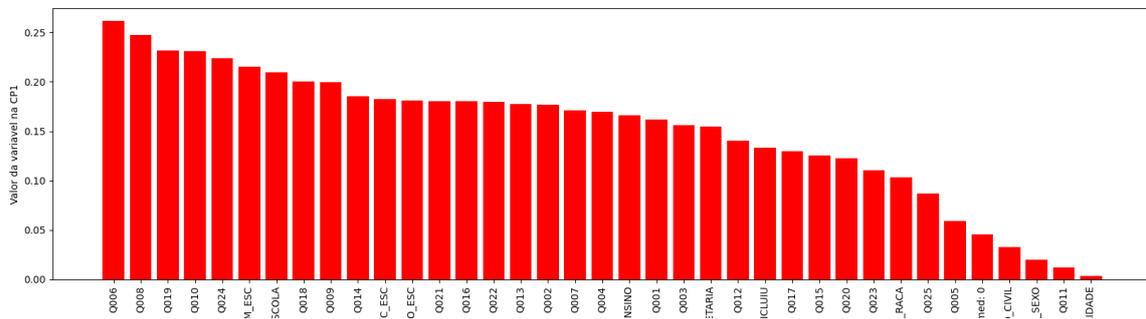
Nome da Variável	Descrição	Nome da Variável	Descrição
TP_FAIXA_ETARIA	Faixa etária ²	Q008	Na sua residência tem banheiro?
TP_SEXO	Sexo	Q009	Na sua residência tem quartos para
TP_ESTADO_CIVIL	Estado Civil	Q010	Na sua residência tem carro?
TP_COR_RACA	Cor/raça	Q011	Na sua residência tem motocicleta?
TP_NACIONALIDADE	Nacionalidade	Q012	Na sua residência tem geladeira?
TP_ST_CONCLUSAO	Situação de conclusão do Ensino Médio	Q013	Na sua residência tem freezer (independente ou segunda porta
TP_ANO_CONCLUIU	Ano de Conclusão do Ensino Médio	Q014	Na sua residência tem máquina de lavar roupa? (o tanquinho
TP_ESCOLA	Tipo de escola do Ensino Médio	Q015	Na sua residência tem máquina de secar roupa (independente ou em conjunto
TP_ENSINO	Tipo de instituição que concluiu ou	Q016	Na sua residência tem forno micro-
TP_DEPENDENCIA_ADM_ESC	Dependência administrativa (Escola) C	Q017	Na sua residência tem máquina de lavar louça?
TP_LOCALIZACAO_ESC	Localização (Escola)	Q018	Na sua residência tem aspirador de pó?
TP_SIT_FUNC_ESC	Situação de funcionamento (Escola)	Q019	Na sua residência tem televisão em
Q001	Até que série seu pai, ou o homem responsável por você,	Q020	Na sua residência tem aparelho de DVD?
Q002	Até que série sua mãe, ou a mulher responsável por você,	Q021	Na sua residência tem TV por assinatura?
Q003	A partir da apresentação de algumas ocupações divididas em grupos ordenados, indique o grupo que contempla a ocupação mais próxima	Q022	Na sua residência tem telefone celular?
Q004	A partir da apresentação de algumas ocupações divididas em grupos ordenados, indique o grupo que contempla a ocupação mais próxima	Q023	Na sua residência tem telefone fixo?
Q005	Incluindo você, quantas pessoas moram atualmente em sua	Q024	Na sua residência tem computador?
Q006	Qual é a renda mensal de sua família? (Some a sua renda com a	Q025	Na sua residência tem acesso à Internet?
Q007	Em sua residência trabalha		

Fonte: Adaptado do Dicionário de Microdados ENEM 2022 (INEP).

Posteriormente, com o banco de dados já pronto para o estudo, foi implementado o método do PCA e os resultados são expressos pelas Figuras 02 e 03. A primeira para o de menor desempenho e a segunda para a de maior desempenho. É digno de nota que o objetivo do presente trabalho é analisar a primeira componente principal apresentada pelo

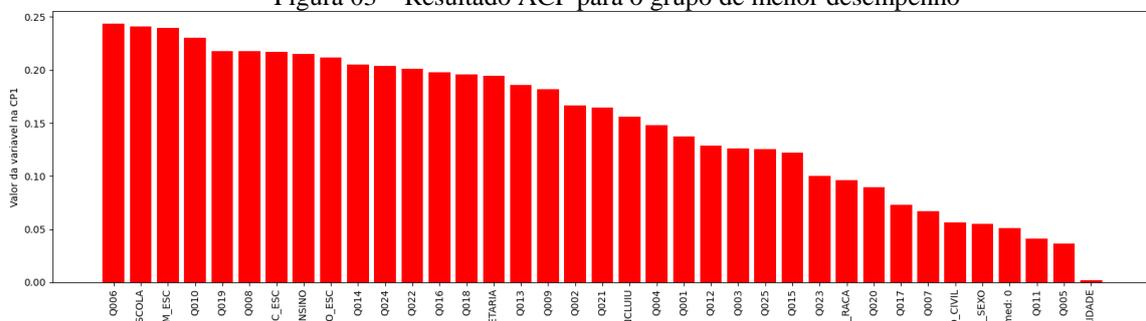
método, porém, como forma de complementar a análise, discutiremos também a segunda componente de cada grupo.

Figura 02 – Resultado ACP para o grupo de maior desempenho



Fonte: Autoria Própria.

Figura 03 – Resultado ACP para o grupo de menor desempenho



Fonte: Autoria Própria.

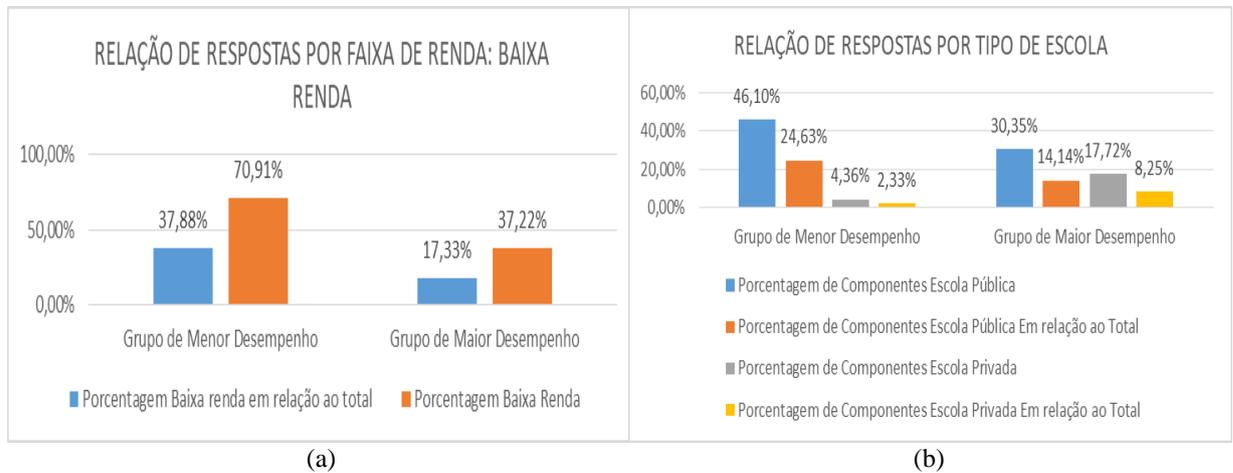
Em ambos os grupos, a questão Q006, referente a renda da família, se mostrou como a de maior relação com as demais questões. Por isso, a fim de investigar melhor o porquê a renda foi o principal fator nos dois grupos, uma análise mais específica foi feita. Nesse momento, foi avaliado quantos por cento em cada grupo, como também em relação aos dados gerais, representavam os estudantes de baixa renda, aqui definidos pelo critério adotado em 2022 pelo governo federal de até meio salário mínimo per capita.

Como resultado, a pesquisa demonstrou que 70,91% daqueles que compunham o grupo de menor desempenho era baixa renda, o que representa 37,88% dos participantes totais. Além disso, no grupo de alto desempenho, os alunos de baixa renda representaram 37,22% do grupo, sendo equivalente a 17,33% do total. A Figura 04 (a) apresenta o resumo dos resultados.

Também, o tipo de escola é indicado como a segunda componente que tem maior relação com o desempenho dos estudantes no grupo de menor desempenho. A Figura 04 (b) traz a distribuição dos estudantes de escola pública e privada, nesse caso foram retiradas aquelas pessoas que preferiram não responder as perguntas, e fica visível que, assim como acontece na renda, a proporção no grupo de baixo desempenho dos que

frequentam escola privada, 4,36%, em comparação ao grupo de maior desempenho, 17,72%, é bem discrepante.

Figura 04 – Duas principais variáveis do grupo de menor desempenho: (a) Renda; (b) Tipo de escola.



Fonte: Autoria Própria.

Essa disparidade tão elevada dos percentuais demonstra que os estudantes de baixa renda têm mais dificuldade de alcançar um alto desempenho, não por causa de sua capacidade cognitiva, mas pela estrutura limitada com que eles têm que desenvolver seu estudo, seja no núcleo familiar ao lidar com as dificuldades financeiras, como também nas escolas públicas que a maioria frequenta, onde não é oferecido um ensino que os permita está no mesmo patamar de conhecimento daqueles advindos de escolas privadas.

Sobre esse último ponto, os resultados do Índice de Desenvolvimento da Educação Básica (ideb) 2023, disponibilizados no site do governo federal, demonstram que há uma grande diferença ao longo dos anos entre a eficiência do ensino público e privado no Brasil. O resultado de 2023, especialmente para o ensino médio, mostra que os estudantes da rede pública, principalmente estadual, saem com uma desvantagem na ordem de 1,5 pontos em comparação com os do ensino privado.

Além disso, para o grupo de maior desempenho, a Q008, referente a quantidade de banheiros da casa, foi indicada como mais relevante. Como mostra Kamakura e Mazzon (2016), seguindo modelo de estratificação novo da Associação Brasileira de Empresas de Pesquisas (ABEP) estabelecido em 2015, quanto maior o número de banheiros, maior é a renda média da família e o contrário também é verdade, quanto menor a renda, menor a quantidade de banheiros. Deixando ainda mais evidente a associação direta do desempenho dos estudantes do ENEM 2022 com o nível socioeconômico o qual estavam inseridos.

CONSIDERAÇÕES FINAIS

Desde os tempos antigos a compreensão das desigualdades sociais tem sido um tema central e recorrente na reflexão humana. A análise feita a partir do presente artigo demonstrou de forma bem explícita que esse assunto, embora muito falado, ainda está longe de ser solucionado por completo. A aplicação da ACP nos dados do ENEM 2022 demonstraram uma relação direta entre a renda, tipo de escola, outros fatores socioeconômicos e o desempenho acadêmico dos estudantes.

O fato de que uma parcela significativa dos alunos de baixo desempenho no ENEM vem de famílias de baixa renda, em contraste com a maior proporção de estudantes não baixa renda no grupo de maior desempenho, escancara a profunda influência das relações socioeconômicas na construção educacional de cada estudante.

Desse modo, esse estudo reforça a necessidade de abordagens mais eficazes e integradas para combater as desigualdades sociais e educacionais. É de extrema importância a aplicação de políticas públicas que não apenas promovam o acesso à educação, mas também garantam a equidade na qualidade do ensino, permitindo que todos os estudantes, independente de sua origem socioeconômica, tenham a oportunidade de alcançar seu pleno potencial.

REFERÊNCIAS

Babbie, E. R. (2020). *The Practice of Social Research*. 15th edition. Cengage Learning. BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Microdados ENEM 2022. 2023. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/enem/divulgados-microdados-do-enem-2022>. Acesso em: 08 jan. 2024.

BRASIL. Assessoria de Comunicação Social do Inep. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Divulgados microdados do Enem 2022**. 2023. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/enem/divulgados-microdados-do-enem-2022>. Acesso em: 07 jun. 2024.

BRASIL. Lei nº 12711, de 29 de agosto de 2012. **Lei Nº 12.711, de 29 de Agosto de 2012**. Brasília

BOLDRINI, J. L. “Algebra Linea” Boldrini, J. L.; Costa, S.I.R.; Ribeiro, V. L., Wetzler, H.G., Álgebra Linear, Harper-Row, São Paulo, 1980.

Clem, E. L. do V., Vinhal, J. M., & Conceição, M. I. G. (2024). Desafios de estudantes de baixa renda na educação superior pública. *Caderno Pedagógico*, 21(5), e4178. <https://doi.org/10.54033/cadpedv21n5-045>

HIPPOLYTO, Luzia de Queiroz. Fatores institucionais associados ao desempenho em matemática por meio da análise de componentes principais (PCA) e da decomposição em fatores paralelos (PARAFAC). 2018.143f. – Tese (Doutorado) – Universidade Federal do Ceará, Programa de Pós-graduação em Educação, Fortaleza (CE), 2018.

INEP; EDUCAÇÃO, Ministério da; FEDERAL, Governo. **Resultados 2023 ideb**. Brasil: Inep, 2024. 54 p. Disponível em: https://download.inep.gov.br/ideb/apresentacao_ideb_2023.pdf. Acesso em: 11 set. 2024.

KAMAKURA, W.; MAZZON, J. A.. Critérios de Estratificação e Comparação de Classificadores Socioeconômicos no Brasil. **Revista de Administração de Empresas**, v. 56, n. 1, p. 55-70, jan. 2016.

LÉO RODRIGUES (Brasil). Agência Brasil. **IBGE: apesar da queda acentuada da pobreza, desigualdades se mantêm**: estudo foi divulgado nesta sexta-feira. Estudo foi divulgado nesta sexta-feira. 2023. Disponível em: <https://agenciabrasil.ebc.com.br/economia/noticia/2023-08/ibge-apesar-da-queda-acentuada-da-pobreza-desigualdades-se-mantem>. Acesso em: 30 ago. 2024.

SANTOS FILHO, F.F.B. Uma simulação comparativa do PCA e do PARAFAC em um estudo de caso da avaliação da efetividade do ensino. 2012.71 f. Trabalho de Conclusão de Curso (Licenciatura em Física) – Centro de Ciências, Universidade Federal do Ceará, Fortaleza, 2012.

SILVA, Thomaz Edson Veloso da. *Educometrics: from theory to application*. 2017. 82 f. Tese (Doutorado em Engenharia de Teleinformática) – Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2017.

Varella, C. A. A. *Análise de Componentes Principais*. Dissertação, Universidade Federal Rural do Rio de Janeiro, Seropédica – RJ, 2008.