



HÁBITOS DE ESTUDO DOS ALUNOS DO ENSINO MÉDIO DURANTE O CONTEXTO DE ENSINO REMOTO EMERGENCIAL E O IMPACTO NAS NOTAS DO ENEM

Palavras-Chave: ENEM 2022, hábitos de estudo, modelos de regressão

Autores:

Leonardo dos Santos, IMECC – UNICAMP

Prof. Dr. Rafael Pimentel Maia (orientador), IMECC - UNICAMP

Introdução

Com o fechamento das escolas em março de 2020 devido à pandemia de COVID-19, tornou-se necessário adotar novas metodologias para manter o processo de aprendizagem dos alunos ¹. A principal alternativa foi o ensino remoto digital, cuja transição repentina representou um desafio para estudantes e docentes (Gularte et al., 2021). Apesar dos avanços durante o período, a aprendizagem seguiu comprometida Macedo (2021).

No Brasil, diversos métodos são utilizados para avaliar o desempenho escolar, com destaque para o Exame Nacional do Ensino Médio (ENEM)², criado em 1998 pelo Ministério da Educação (MEC). Desde sua reformulação em 2009, o ENEM passou a ser utilizado também como forma de acesso ao Ensino Superior, por meio do SiSU, ProUni e FIES.

Estudos indicam que o desempenho dos estudantes no ENEM está relacionado a fatores socioeconômicos (Melo et al., 2021). Durante a pandemia de COVID-19, o isolamento social e a crise econômica resultante afetaram diretamente essas condições (Silber, 2020), o que pode ter impactado negativamente o rendimento de estudantes no Ensino Médio e, consequentemente, nas provas do ENEM. Apesar disso, foram implementadas mudanças no exame, como o ENEM Digital³ e a possibilidade de reaplicação em casos de doenças infectocontagiosas, incluindo a COVID-19⁴.

Excepcionalmente em 2022, além do questionário socioeconômico, foi aplicado um questionário adicional sobre hábitos de estudo durante a pandemia, abordando temas como gestão do tempo, práticas de estudo, acesso à tecnologia e dificuldades enfrentadas pelos candidatos⁵. A análise desses dados permite compreender as condições de estudo no período, relacionar hábitos e condições socioeconômicas e avaliar possíveis impactos no desempenho no ENEM e na qualidade da aprendizagem.

¹https://www.todospelaeducacao.org.br/_uploads/_posts/425.pdf

²<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enem>

³<https://www.in.gov.br/en/web/dou/-/editais-n-34-de-28-de-abril-de-2022-396179710f>

⁴<https://www.in.gov.br/en/web/dou/-/editais-n-33-de-28-de-abril-de-2022-396385788>

⁵<https://www.gov.br/inep/pt-br/assuntos/noticias/enem/painel-apresenta-pesquisa-sobre-estudo-na-pandemia>

Materiais e Metodologia

Os dados utilizados neste estudo referem-se aos microdados anonimizados do ENEM 2022, disponibilizados publicamente pelo INEP⁶, em conformidade com a Lei Geral de Proteção de Dados (Lei Nº 13.709/2018)⁷. A base original contém cerca de 3,5 milhões de registros. Para este estudo, foram selecionados apenas os estudantes não-treineiros, presentes nos dois dias de prova, que responderam integralmente ao questionário de hábitos de estudo e não zeraram nenhuma prova, resultando em uma amostra final de 90.518 registros para a análise da nota de Matemática e suas Tecnologias.

As análises iniciais utilizaram técnicas descritivas univariadas e bivariadas, como tabelas, gráficos e correlações. Para investigar a relação entre hábitos de estudo e desempenho no ENEM 2022, foram aplicados modelos de regressão linear múltipla (Kutner, 2005).

Considere uma amostra de n indivíduos e uma variável resposta contínua Y_i , representando a nota obtida no ENEM. Seja $X_i = (X_{i1}, X_{i2}, \dots, X_{i,p-1})^t$ o vetor de variáveis explicativas do indivíduo i . O modelo de regressão linear múltipla é dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i,$$

em que $(\beta_0, \beta_1, \dots, \beta_{p-1})^t$ é o vetor de parâmetros e $\varepsilon_i \sim N(0, \sigma^2)$ são os erros aleatórios, assumidos independentes, com média zero e variância constante.

Os parâmetros do modelo foram estimados por máxima verossimilhança. A modelagem estatística inclui, além da estimação, a avaliação da qualidade do ajuste por meio de gráficos de resíduos e medidas como o *Mean Absolute Error (MAE)*; a seleção de variáveis com base no *Bayesian Information Criteria (BIC)*; e a interpretação dos resultados. Todas as análises serão realizadas na linguagem R (R Core Team, 2023).

Resultados

As análises iniciais revelaram que, entre os 90,5 mil estudantes analisados: 63% são mulheres; 90% têm 17 ou 18 anos; quase 40% vivem em famílias com renda de até 2 salários mínimos (R\$2.424,00 em 2022); e 96% não tiveram o ano letivo interrompido pela pandemia.

A Figura 1 ilustra a relação entre renda familiar e a nota na prova de Matemática. Observa-se que as notas tendem a se aproximar de 1.000 conforme a renda aumenta, embora haja muitos *outliers*. Além disso, destaca-se a forte desigualdade nos níveis de renda.

A Figura 2 apresenta a relação entre a nota na prova de Matemática e o estado de residência dos estudantes, com divisão visual por estado. Observa-se que estudantes das regiões Sul e Sudeste apresentam melhor desempenho. Relações semelhantes foram encontradas ao comparar renda e notas nas demais provas. Se desejar.

⁶<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>

⁷https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm

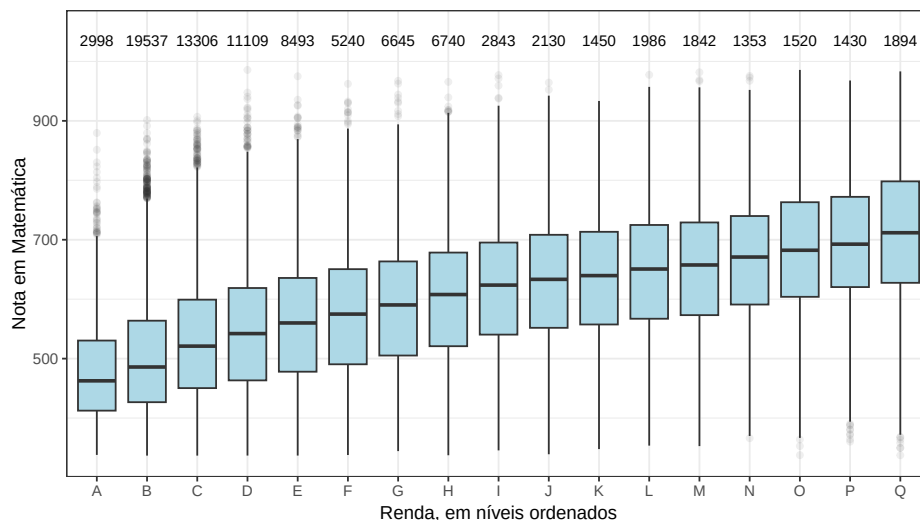


Figura 1: Diagrama de caixa para a distribuição da nota no exame de Matemática (ENEM 2022) segundo a faixa de renda familiar. No eixo X, "A" representa renda mensal inferior a 1 salário mínimo (R\$1.212,00) e "Q" renda superior a 20 salários mínimos (R\$24.240,00). Acima de cada gráfico, está indicado o tamanho do grupo correspondente.

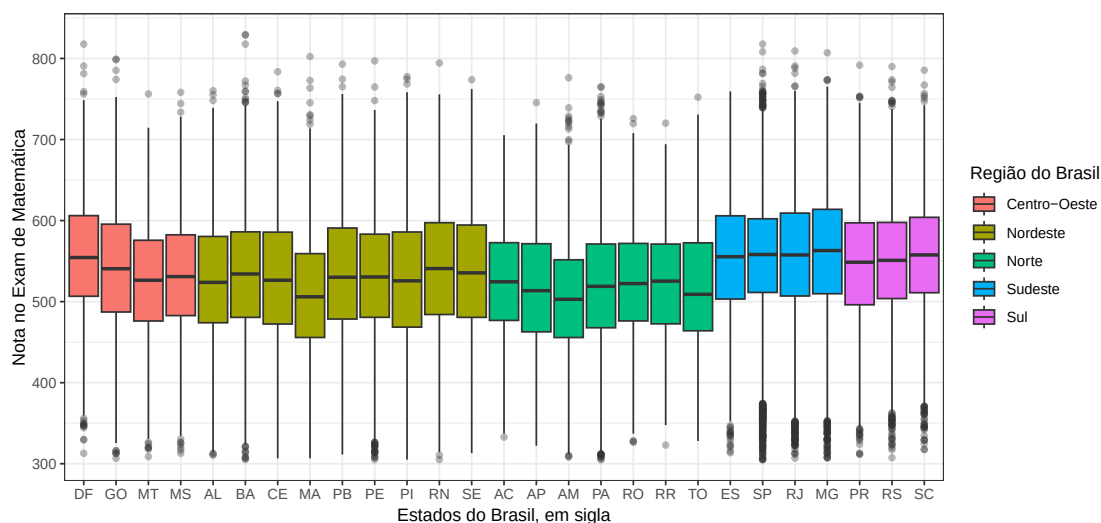


Figura 2: Diagrama de caixa para a distribuição da nota no exame de Matemática (ENEM 2022) segundo o estado da federação.

Modelos Lineares Ajustados

Os modelos para a nota do exame de Matemática foram ajustados separadamente para cada região brasileira devido às diferenças socioeconômicas, culturais e educacionais presentes entre elas. Ao ajustar os modelos por região, é possível capturar melhor as especificidades locais e melhorar a precisão das estimativas. A base de dados foi estratificada por região: Sul (11.882 registros), Sudeste (35.260 registros), Centro-Oeste (8.405 registros), Norte (8.582 registros) e Nordeste (26.292 registros).

Inicialmente, ao todo, haviam 118 variáveis preditoras, entre dados pessoais (faixa de idade, estado civil, etc.), dados socioeconômicos (renda mensal, quantidade de eletrodomésticos, escolaridade do pai e da mãe, etc.) e dados de hábitos de estudos (vínculo com a escola, motivos de desmotivação, se precisou de apoio financeiro ou para estudar, etc.). Para seleção do melhor modelo foi utilizado o método de seleção *stepwise*.

Após a seleção de variáveis, para a nota de Matemática, os modelos ajustados consideraram: 59 variáveis para a região Sudeste, 52 variáveis para a região Nordeste, 36 variáveis para a região Sul, 29 variáveis para a região Centro Oeste e 22 variáveis para a região Norte. Em sua maioria, as variáveis finais selecionadas pertencem ao questionário de hábitos de estudos, especificamente, se o estudante se sentiu desmotivado durante seus estudos; se durante o ensino remoto: participou de fóruns de discussão, treinou redação ou não se atrasou para as aulas; se precisou de ajuda para se manter informado no ano de 2021; se entrou nas aulas remotas sem atraso; e se fez resumo das videoaulas, podcasts assistidos. Quanto às variáveis demográficas e socioeconômicas, cita-se principalmente a renda, faixa etária, sexo e tipo de escola: privada ou pública.

A Figura 3 mostra as estimativas dos efeitos das 12 variáveis de maior impacto na nota em Matemática, em comparação à seus respectivos grupos de referência. Note que, assim como visto na Figura 1, a renda é uma variável preditora importante, dado que é a que possui as maiores estimativas de efeito, em comparação ao seu grupo de referência: até 1,5 salários mínimos. A participação em fóruns de discussão de matéria pelo estudante, muitas ou todas as vezes que este tinha alguma dúvida, de acordo com as estimativas, teve um impacto negativo em comparação ao grupo de referência: aqueles que não participaram nenhuma vez de fóruns de discussão. A idade também mostrou um efeito negativo: aqueles com idade maior que 19 anos obtiveram, em média, entre 20 e 30 pontos a menos que aqueles no grupo de referência: até 17 anos.

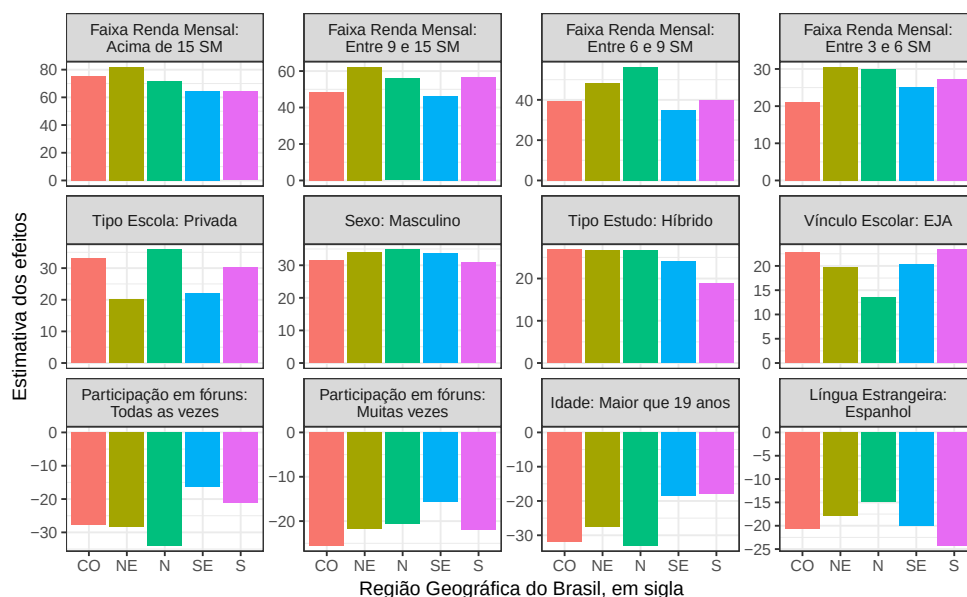


Figura 3: Estimativa dos efeitos, por região brasileira, das 12 variáveis de maior impacto na nota em Matemática em comparação à seus respectivos grupos de referência (Faixa Renda Mensal: Até 1,5 Salários Mínimos (SM); Tipo Escola: Pública; Sexo: Feminino; Participação em fóruns: Nenhuma vez; Idade: Até 17 anos; Tipo Estudo: Apenas presencial; Língua Estrangeira: Inglês; Vínculo Escolar: Ensino Médio Regular).

Conclusão

Pelos resultados observados, conclui-se que, de fato, os hábitos de estudos durante o contexto de ensino remoto emergencial de estudante impactaram em seu desempenho na prova de Matemática do ENEM 2022, dado que muitas das variáveis desse conjunto mostraram efeitos significativos. Vê-se que: o desempenho de alunos que estudaram de forma híbrida em 2021 obtiveram entre 15 e 25 pontos em comparação aqueles que estudaram somente no modo presencial; estudantes do formato EJA (Educação de Jovens e Adultos) obtiveram desempenho de até 25 pontos a mais aqueles que estudaram no formato de Ensino Médio Regular; e o desempenho de estudantes que participaram de fóruns de discussão online em 2021 foi entre 10 e 30 pontos menor do que aqueles que nunca participaram. Cita-se, também, algumas das variáveis demográficas e socioeconômicas que mostraram efeito significativo, como a renda, que mostrou um efeito de até 80 pontos, em comparação com aqueles que recebem até 1,5 salários mínimos; o sexo do estudante, em que homens obtiveram acima de 30 pontos em comparação à mulheres; e a idade, em que estudantes com 19 anos ou mais tiveram um desempenho menor, em média 20 pontos, em comparação aos estudantes com 17 anos ou menos.

Quanto aos modelos de regressão lineares múltiplos, conclui-se que os ajustes, por região, considerando tanto variáveis socioeconômicas quanto variáveis de hábitos de estudos resultou em modelos adequados, dado a análise dos resíduos graficamente e também pelo cálculo de estatísticas como o *Mean Absolute Error (MAE)*. Por fim, tendo em vista os resultados obtidos para a nota em Matemática e suas Tecnologias, se torna instigante a continuação das análises para as outras 3 provas aplicadas no ENEM 2022, e questiona-se se essas análises, mantendo a estratificação por região, concluirão em impactos semelhantes ou muito distintos. Ainda, questiona-se sobre a possibilidade da produção de um modelo de regressão multivariado, que simultaneamente ajuste, para cada região geográfica do Brasil, uma nota fixa, como a de Matemática, o que resultaria em um único modelo, em comparação a 5 (cinco) modelos separados.

Referências

- Gularte, F. N., Nascimento, F. K. V. d., and Carvalho, A. V. d. (2021). Enem in times of pandemic: challenges faced by teachers in the face of remote education of the state network of president kennedy to. *Research, Society and Development*, 10(14).
- Kutner, M. (2005). *Applied Linear Statistical Models*. McGraw-Hill international edition. McGraw-Hill Irwin.
- Macedo, R. M. (2021). Direito ou privilégio? desigualdades digitais, pandemia e os desafios de uma escola pública. *Estudos Históricos (Rio de Janeiro)*, 34(73):262–280.
- Melo, R. O., Freitas, A. C. d., Francisco, E. d. R., and Motokane, M. T. (2021). Impacto das variáveis socioeconômicas no desempenho do enem: uma análise espacial e sociológica. *Revista de Administração Pública*, 55(6):1271–1294.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Silber, S. D. (2020). A fragilidade econômica e financeira na pandemia do sars-covid-19. *Estudos Avançados*, 34(100):107–115.