

Instituto Federal de Brasília *Campus* Estrutural
Licenciatura em Matemática

EVANDRO DA SILVA MAGALHÃES JÚNIOR

**FATORES SOCIOECONÔMICOS ASSOCIADOS AO
DESEMPENHO DOS ESTUDANTES NA PROVA DE
MATEMÁTICA DO EXAME NACIONAL DO ENSINO
MÉDIO**

BRASÍLIA
JANEIRO DE 2022

Evandro da Silva Magalhães Júnior

FATORES SOCIOECONÔMICOS ASSOCIADOS AO DESEMPENHO DOS ESTUDANTES NA PROVA DE MATEMÁTICA DO EXAME NACIONAL DO ENSINO MÉDIO

Trabalho de conclusão de curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia de Brasília Campus Estrutural como parte dos requisitos necessários à obtenção do título de Licenciado em Matemática.

Orientador: Adriana Barbosa de Souza

Coorientador: Pedro Carvalho Brom

Brasília

Janeiro de 2022



MINISTÉRIO DA EDUCAÇÃO
Instituto Federal de Educação, Ciência e Tecnologia de Brasília

FICHA DE APROVAÇÃO EM BANCA EXAMINADORA

Trabalho de Conclusão de Curso

Discente: EVANDRO DA SILVA MAGALHÃES JÚNIOR

Título: FATORES SOCIOECONÔMICOS ASSOCIADOS AO DESEMPENHO DOS ESTUDANTES NA PROVA DE MATEMÁTICA DO EXAME NACIONAL DO ENSINO MÉDIO

Trabalho aprovado em: 04 / 02 / 2022.

Brasília - DF, 04 de Fevereiro de 2022.

Banca Examinadora

Orientadora (Presidente): Ma. Adriana Barbosa de Souza

Coorientador: Esp. Pedro Carvalho Brom

Examinador (membro): Dr. Wembesom Mendes Soares

Examinador (membro): Me. Bruno Marx de Aquino Braga

Documento assinado eletronicamente por:

- Bruno Marx de Aquino Braga, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 11/02/2022 14:54:11.
- Wembesom Mendes Soares, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 07/02/2022 22:03:12.
- Pedro Carvalho Brom, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 07/02/2022 15:37:31.
- Adriana Barbosa de Souza, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 05/02/2022 11:07:36.

Este documento foi emitido pelo SUAP em 04/02/2022. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifb.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 349161

Código de Autenticação: 5b705a2533



Resumo

O presente trabalho tem como objetivo analisar os fatores socioeconômicos associados ao desempenho dos estudantes na prova de matemática do Exame Nacional do Ensino Médio, com objetivo de entender os fatores que influenciam no desempenho final do discente na prova de matemática. A partir do uso da nota do ENEM como forma de ingresso à universidade, a forma de se ensinar e avaliar tem sofrido grandes mudanças no ensino médio. Essa mudança gerou um grande impacto na forma de se lecionar e avaliar, com muitas escolas adotando metodologias de ensino voltada para a aprovação em exames como ENEM ou similares, foi criada uma cultura de aprovação em vestibulares que resume a qualidade de ensino ao desempenho nessas avaliações, fazendo com que a matriz curricular do ENEM se torne um parâmetro para os currículos das escolas de ensino médio do país. Com o resultado das análises, espera-se contribuir com a predição de problemas de rendimento escolar que podem vir a acontecer, tornando mais claro o perfil dos estudantes que pode apresentar alguma dificuldade, tomando com referência, o perfil socioeconômico que ele possui. E a partir dos dados socioeconômicos, espera-se que os gestores escolares possam ter uma ideia maior dos fatores socioeconômicos que levam os estudantes a apresentarem melhor ou pior desempenho no ENEM, buscando melhorar as metodologias utilizadas pelo colégio, não deixando o resultado final como o principal resultado do processo educacional.

Palavras-chaves: Avaliação de Larga Escala; Análise de Dados; Educação.

Abstract

The present work aims to analyze the socioeconomic factors associated with the performance of students in the mathematics test of the National High School Exam, in order to understand the factors that influence the student's final performance in the mathematics test. From the use of the ENEM grade as a way of entering the university, the way of teaching and evaluating has undergone major changes in high school. This change had a great impact on the way of teaching and evaluating, with many schools adopting teaching methodologies aimed at passing exams such as ENEM or similar, a culture of approval in vestibular was created that summarizes the quality of teaching to performance in these evaluations, making the ENEM curriculum a parameter for the curricula of high schools in the country. With the result of the analyses, it is expected to contribute to the prediction of school performance problems that may happen, making the profile of students who may have some difficulty clearer, taking as a reference, the socioeconomic profile they have. With the result of the analyses, it is expected to contribute to the prediction of school performance problems that may happen, making the profile of students who may have some difficulty clearer, taking as a reference, the socioeconomic profile they have.

Keywords: Large Scale Rating; Data analysis; Education.

Lista de ilustrações

Figura 1 – Bloco 1 - Informação dos Participantes	25
Figura 2 – Bloco 2 - Nota dos Participantes	26
Figura 3 – Bloco 3 - Questionário Socioeconômico	27

Lista de tabelas

Tabela 1 – Escala Temporal Avaliações de Larga Escala no Brasil	15
Tabela 2 – Nota Separada em Quintil	23
Tabela 3 – Divisão Blocos dos Dados	24

Lista de abreviaturas e siglas

ANEB	Avaliação Nacional da Educação Básica
ANRESC	Avaliação Nacional do Rendimento Escolar
ANA	Avaliação Nacional da Alfabetização
ACM	Análise de Correspondência Múltipla
CNE	Conselho Nacional de Educação
CNPE	Centro Brasileiro de Pesquisas Educacionais
ENEM	Exame Nacional do Ensino Médio
ENADE	Exame Nacional de Desempenho dos Estudantes
IFB	Instituto Federal de Brasília
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LDB	Lei de Diretrizes e Bases da Educação Nacional
PCN	Parâmetros Curriculares Nacionais
SAEB	Sistema Nacional da Educação Básica

Sumário

1	INTRODUÇÃO	8
2	AVALIAÇÕES NO BRASIL E O INEP	11
3	CRIAÇÃO DAS AVALIAÇÕES DE LARGA ESCALA	14
4	ESTRUTURA DO ENEM	17
5	MATERIAIS E MÉTODOS	19
5.1	Material	19
5.2	Métodos	20
5.2.1	Análise de Correspondência Múltipla	20
5.2.2	Desenvolvimento Algébrico da Análise de Correspondência	21
6	RESULTADOS E DISCUSSÃO	23
6.1	Detalhamento dos Dados	23
6.1.1	Análise de Múltiplas Correspondências - MCA	25
	Considerações finais	29
	REFERÊNCIAS	31
	APÊNDICES	34
	APÊNDICE A – TABELA DE DESCRIÇÃO DAS VARIÁVEIS . . .	35

1 Introdução

Esse trabalho é focado no Exame Nacional do Ensino Médio, que é uma das maiores avaliações de larga escala do país. O ENEM foi criado em 1998, pela Portaria MEC nº 438 da lei Nº 9.131 com objetivo de avaliar os conhecimentos adquiridos pelos estudantes ao final da educação básica. Com o passar dos anos, o exame acabou sofrendo diversas mudanças de estrutura e metodologia, até que em 2009 começou a ser utilizado como forma de acesso à universidade, o que muda completamente o funcionamento e estrutura da avaliação no país, o que modifica drasticamente a forma de avaliar em larga escala.

Desta forma, o ENEM ocupa uma posição de destaque entre as avaliações de larga escala no Brasil, e muitas vezes moldando os conteúdos que são ensinados nas escolas brasileiras. No ano de 2019 o ENEM contou com 5.095.270 participantes, demonstrando ser uma prova com grande alcance e que atinge estudantes de escolas públicas e particulares de todo o Brasil.

A partir do uso da nota do ENEM como forma de ingresso à universidade, a forma de se ensinar e avaliar tem sofrido grandes mudanças no ensino médio. Essa mudança gerou um grande impacto na forma de se lecionar e avaliar, com muitas escolas adotando metodologias de ensino voltada para a aprovação em exames como ENEM ou similares, foi criada uma cultura de aprovação em vestibulares que resume a qualidade de ensino ao desempenho nessas avaliações, fazendo com que a matriz curricular do ENEM se torne um parâmetro para os currículos das escolas de ensino médio do país.

Vale ressaltar que ENEM não tinha como objetivo ditar os conteúdos que são ensinados nas escolas, já que a prova é dividida em habilidades e competências, buscando purificar e abranger diversos conhecimentos. Porém a cultura de acesso a universidade a partir de vestibulares acabou se moldando a esse modelo de avaliação, influenciando a estrutura curricular nas escolas do ensino médio e criando metodologias de ensino baseadas nessa estrutura.

Dentro dessa política de acesso às universidades, estudantes de famílias com maior renda possuem maior incentivo e oportunidades para se prepararem para a avaliação. O Enem, sendo um exame com estrutura definida, possibilita que muitas metodologias de preparação para a prova sejam construídas estrategicamente para que o estudante treine com foco na melhora de desempenho da prova.

Escolas e cursos preparatórios com maior investimento e estrutura possibilitam essa oportunidade de treinamento e preparo focado na estrutura da prova, assim, seus estudantes alcançam os melhores desempenhos. Em geral, essas escolas e cursos com grande investimento são instituições da rede privada de ensino, às quais só podem ser

acessadas pelos alunos de baixa renda por meio de incentivos como bolsas de estudos ou cursos solidários.

É necessário oportunizar o acesso a esse tipo de conhecimento também na rede pública de ensino, é importante que os alunos aprendam como estudar: qual a melhor metodologia de estudo, quais os objetivos desta avaliação, como treinar uma situação de prova, além de receberem a orientação adequada quanto ao seu desempenho durante os estudos, não só após realizada a prova final.

O Exame tem a proposta de avaliar o conhecimento dos estudantes através de situações - problema, nas quais os itens são criados baseados nas competências. A resolução de problemas, segundo Rabelo (2013, p.51), é destacada como referência à autonomia intelectual do estudante no final da educação básica, mediada pelos princípios da cidadania e do trabalho. E, além disso, as competências para a participação social incluem a criatividade, capacidade de solucionar problemas, senso crítico e domínio dos conhecimentos.

De acordo com Rabelo (2013, p.5), o Enem “[...] tem uma perspectiva de avaliação de competências e com provas contextualizadas e interdisciplinares, que abrangiam as diversas áreas de conhecimento objeto de estudo do ensino médio”. Este estudo aborda a área de conhecimento de matemática e suas tecnologias como objeto de estudo, em que, busca-se entender o desempenho dos estudantes a partir dos fatores socioeconômicos que são informados ao fazer a inscrição no programa. Para essa análise serão utilizados os dados disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) referentes ao ano de 2019.

Segundo Carvalho (2017, p.17) os primeiros estudos sobre a influência dos fatores socioeconômicos no desempenho escolar foram realizados entre os anos de 1950 e 1960. Nas últimas décadas muitos autores estudaram o tema, formando um grande acervo sobre os fatores associados ao desempenho escolar. Neste momento foi observado que grande parte dos estudantes tinham seu desempenho explicado pelas condições socioeconômicas, ambiente letrado, importância que os pais ou familiares atribuem à educação, qualidade do ensino nas escolas, formação e prática docente, estrutura física, gestão escolar e outros aspectos. Dentre os autores que pesquisam na área temos Carvalho, Rabelo, Lima, Soares que fizeram contribuições muito importantes para o segmento.

Os dados disponibilizados pelo INEP permitem que sejam feitas diversas análises, principalmente relacionadas ao desempenho dos estudantes. Segundo (Lima et al., p.100), com o vasto volume de dados que são disponibilizados, a análise de perfil do estudante e predição de resultados, que correlacionam com os dados socioeconômicos tem sido pouco explorados, existindo uma demanda de trabalhos com objetivo de entender como os fatores socioeconômicos podem afetar o desempenho dos estudantes na avaliação, gerando um campo de pesquisa muito amplo e ainda pouco explorado e com grande potencial para entender melhor o perfil dos estudantes e do sistema educacional brasileiro.

O presente trabalho tem como objetivo analisar os fatores socioeconômicos associados ao desempenho dos estudantes na prova de matemática do Exame Nacional do Ensino Médio, com objetivo de entender os fatores que influenciam no desempenho final do discente na prova de matemática. Além de fazer uma análise exploratória dos dados socioeconômicos do ENEM, propondo mostrar a associação, se houver, entre o desempenho na prova de matemática e o perfil socioeconômico.

Com o resultado das análises, espera-se contribuir com a predição de problemas de rendimento escolar que podem vir a acontecer, tornando mais claro o perfil dos estudantes que podem apresentar alguma dificuldade, tomando com referência, o perfil socioeconômico que ele possui. E a partir dos dados socioeconômicos, espera-se que os gestores escolares possam ter uma ideia maior dos fatores socioeconômicos que levam os estudantes a apresentarem melhor ou pior desempenho no ENEM, buscando melhorar as metodologias utilizadas pelo colégio, não deixando o resultado final como o principal resultado do processo educacional.

2 Avaliações no Brasil e o INEP

O INEP, que é uma instituição vinculada ao Ministério da Educação que tem como objetivo promover estudos, pesquisas e avaliações, buscando entender e acompanhar o sistema educacional brasileiro, possibilitando a implementação de políticas públicas mais eficientes para a educação. Atualmente o INEP realiza pesquisas e avaliações em todos os níveis e modalidades de ensino, a maior delas e foco principal do estudo é o ENEM, que desde 2013 vem ocupando um papel de extrema importância no processo educacional do país.

Segundo Saviani(2012, p.292) O INEP começou a ser cogitado em 1937, com o nome de Instituto Nacional de Pedagogia, aparecendo na lei Nº 378, partindo pela iniciativa do então ministro da Educação e Saúde Pública, Gustavo Capanema, porém o Instituto Nacional De Estudos Pedagógicos (INEP) só foi criado no ano seguinte pelo Decreto-Lei Nº 580, de 30 de julho de 1938, já na vigência do Estado Novo.

De acordo com o Decreto-Lei Nº 580, o INEP precisa seguir as Competências:

- a) organizar documentação relativa à história e ao estudo atual das doutrinas e das técnicas pedagógicas, bem como das diferentes espécies de instituições educativas;
- b) manter intercâmbio, em matéria de pedagogia, com as instituições educacionais do País e do estrangeiro;
- c) promover inquéritos e pesquisas sobre todos os problemas atinentes à organização do ensino, bem como sobre os vários métodos e processos pedagógicos;
- d) promover investigações no terreno da psicologia aplicada à educação, bem como relativamente ao problema da orientação e seleção profissional;
- e) prestar assistência técnica aos serviços estaduais, municipais e particulares de educação, ministrando-lhes, mediante consulta ou independentemente desta, esclarecimentos e soluções sobre os problemas pedagógicos;
- f) divulgar, pelos diferentes processos de difusão, os conhecimentos relativos à teoria e à prática pedagógicas (Brasil, 1938).

Essas competências norteiam os caminhos das pesquisas e avaliações desenvolvidas pelo INEP até 1952. Vele destacar que a relação com a psicologia e diferentes métodos e processos pedagógicos, como mostra SAVIANI(2012, p. 293):

Ora, a alínea “c” já determinava a promoção de “inquéritos e pesquisas sobre todos os problemas atinentes à organização do ensino, bem como sobre os vários métodos e processos pedagógicos”. Portanto, aí já estavam contempladas as investigações no terreno da psicologia. É compreensível, contudo, esse destaque, pois efetivamente, desde sua criação até o início da década de 1950, a psicologia conformou o nascente pensamento pedagógico exercendo uma hegemonia incontestável, como foi confirmado no estudo de Aparecida Joly Gouveia (1971), “A pesquisa educacional no Brasil”.

A primeira fase do INEP vai de sua criação até o ano de 1955, Anísio Teixeira, que assumiu a diretoria no INEP em 1952, criou o Centro Brasileiro de Pesquisas Educacionais (CBPE) e a partir criação dessa iniciativa, a psicologia passou a ter uma papel mais ativo na educação. Após isso, na década de 60, por conta do contexto político, começaram a debater temas como, a importância econômica da educação e o financiamento do ensino, assim o pensamento pedagógico tendeu a incorporar a área da economia da educação.

A década de 70 foi um momento de muito desenvolvimento educacional, além da expansão da educação tecnicista a psicologia da educação continuou se desenvolvendo, abrindo o debate entre o behaviorismo e o cognitivismo. E ainda nessa década, em 1972 por meio do decreto Nº 71.407, de 20 de novembro de 1972 o Instituto Nacional de Estudos Pedagógicos passou a ser chamado de Instituto Nacional de Estudos e Pesquisas Educacionais, mantendo a sigla INEP. E no final da década de 70 o INEP teve a transferência de sua sede do Rio de Janeiro para Brasília.

Entre as décadas de 80 e 90 o INEP passou por momentos muito difíceis, que acarretaram em diversas mudanças na estrutura do instituto. Como mostra SAVIANI(2012, p. 296), que após as mudanças no período da "Nova República":

o Inep sofreu uma mudança mais radical no início da primeira gestão de Fernando Henrique Cardoso na Presidência da República, iniciada em 1º de janeiro de 1995. Em 8 de março desse ano, assumiu a Diretoria-Geral do Inep Maria Helena Guimarães de Castro, que se afastou em 23 de junho do mesmo ano para assumir a Secretaria de Avaliação e Informação Educacional (Sediae) do MEC, sendo substituída na direção do Inep pelo diretor-executivo Og Roberto Dória. SAVIANI(2012, p. 296)

E efetivamente em 24 de fevereiro de 1997 Maria Helena Castrou voltou a assumir a presidência e no mês seguinte, em 14 de março de 1997, a Lei Nº 9.448 transformou o INEP em uma autarquia federal. Então, depois de permanecer esquecido e quase instinto, o INEP deixou de ter suas funções de fomento as pesquisas educacionais, organização da documentação e disseminação das informações educacionais, passando a ter a responsabilidade de avaliar a educação brasileira em todos os seus níveis e modalidades. SAVIANI(2012, p. 297).

Com toda a mudança estrutural e definição do novo objetivo da instituição, a partir desse momento, o INEP passou a desenvolver novas formas de avaliar a educação brasileira, buscando atingir a maior parcela possível da população. Então tem-se o surgimento as avaliações de larga escala, onde cada nível e modalidade tem sua avaliação.

3 Criação das Avaliações de Larga Escala

Nas linhas de Rabelo as avaliações de larga escala vêm ocupando um papel de extrema importância na educação brasileira, trazendo diferentes possibilidades para sua aplicação. Segundo Rabelo (2013, p.2), as experiências com avaliações educacionais no Brasil são motivadas pelas necessidades políticas de produzir indicadores com objetivo de aprimorar a qualidade do ensino oferecido pelas instituições dos mais de 5,5 mil municípios, que englobam mais de 190 milhões de habitantes.

O início da utilização das avaliações de larga escala se deu com a necessidade de criar parâmetros para medir o nível da educação no país. Com a introdução do conceito de avaliação de larga escala foram criados os Parâmetros Curriculares Nacionais (PCN) por uma necessidade da constituição de 1988 no artigo 206, que garante o padrão de qualidade como um dos princípios que o ensino deve ser ministrado.

Além disso, a Lei de Diretrizes e Bases da Educação Nacional (LDB) lei 9.394 de 1994 título 4 artigo 9 diz: "Assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino", assim o Estado precisa oferecer um padrão de qualidade para a educação. A partir dessa obrigação, surgem outras preocupações como garantir a qualidade de ensino para estudantes de todos os níveis sociais e estabelecer como parâmetro de gestão a observação de fatores socioeconômicos.

As políticas educacionais têm passado por um processo de muitas mudanças, com o objetivo de melhorar e desenvolver a educação no país. No início da década de 90, começaram a ser criadas políticas públicas com a finalidade de analisar os níveis da educação no Brasil. Foram criados diversos sistemas de avaliação educacional como: Sistema Nacional de Avaliação da Educação Básica (SAEB), Avaliação Nacional da Educação Básica (ANEB), Avaliação Nacional do Rendimento Escolar (ANRESC) e Avaliação Nacional da Alfabetização (ANA); além do o Exame Nacional do Ensino Médio (ENEM) e do Exame Nacional de Desempenho dos Estudantes (ENADE). Cada um desses sistemas tem o objetivo de avaliar a educação de diferentes grupos.

Rabelo (2013, p.6) ordena todas essas avaliações em uma escala temporal, temos:

Tabela 1 – Escala Temporal Avaliações de Larga Escala no Brasil

Avaliação	Ano
ENC	1996
ENEM	1998
ENCCEJA	2000
ENADE	2004
Prova Brasil	2005
Provinha Brasil	2008
Novo ENEM	2009
Prova Docente	2010
ANA	2013

Fonte: RABELO(2013, p.6)

Nesse sentido, a avaliação tem sido foco de análise de pesquisadores, em uma gama diversificada de estudos, envolvendo níveis de ensino, diferentes atores e organizações que interagem com tais políticas. (SUDBRACK e COCCO, 2014, p.348)

Segundo Andrade e Soares(2006), a qualidade do ensino não pode ser vista como a média do desempenho cognitivo:

A média do desempenho cognitivo dos alunos de uma dada escola não pode ser tomada como uma medida de sua qualidade, já que escolas diferentes têm alunos com perfis socioeconômicos muito diferentes e é amplamente conhecida a influência do nível socioeconômico no desempenho dos alunos da educação básica. Diante disso, tomar a medida de desempenho como medida de qualidade é favorecer de forma inaceitável as escolas particulares que atendem a alunos socioeconomicamente mais favorecidos. (ANDRADE e SOARES, 2006, p.118).

Segundo Soares (2005, p. 94 e 95), a desigualdade socioeconômica, especialmente no Brasil, é um dos principais fatores que geram dificuldade e afetam o desempenho escolar dos estudantes, tornando necessário levar em conta o contexto socioeconômico do estudante para avaliar seu desempenho escolar.

A busca pela qualidade nas escolas deve ser feita de forma organizada e planejada por uma equipe especializada. Segundo Carvalho (2017):

Salienta-se que a busca por essa qualidade somente pode ser alcançada por meio de uma análise acurada da escola, da situação vigente, das ações realizadas, dos programas implementados, dos materiais, dos recursos humanos, do Projeto Político Pedagógico (PPP), do entorno socioeconômico e político no qual está inserida a escola. (Carvalho, 2017, p.16).

Assim, as avaliações de larga escala têm um papel de muita importância no processo educacional brasileiro, não só como uma porta de entrada as universidades, mas como uma forma de representar e monitorar a qualidade de ensino no Brasil. O INEP realiza avaliações cobrindo toda educação básica em vários âmbitos até o ensino superior, onde são disponibilizados dados que possibilitam que essas análises sejam feitas.

4 Estrutura do ENEM

O exame conta com uma estrutura bem definida em que os participantes fazem a avaliação em 2 dias. No primeiro dia os estudantes fazem a prova de Linguagens e Ciências Humanas e no segundo dia a prova de Ciências da Natureza, Matemática e a Redação. Com um total de 180 questões sendo distribuídas igualmente em 45 itens para cada prova, o estudante tem 5 horas no primeiro dia e 5 horas e meia no segundo dia para a realização da prova. Vale ressaltar que a avaliação vem sofrendo diversas mudanças com a reformulação do ensino médio.

A elaboração de itens é orientada por uma matriz de referência, que serve como base para elaboração de itens de múltipla escolha que alimentam um banco nacional de questões. Esses itens são utilizados para analisar o desempenho do estudante, possibilitando a interpretação quantitativa das escalas de proficiência construídas após a aplicação da prova.

As matrizes de referências contemplam as habilidades essenciais para a determinada etapa do ensino. Além de ser levado em consideração aquilo que é possível ser avaliado por meio de itens de múltipla escolha, o foco da matemática é a resolução de problemas.

O exame contempla 7 competências e serão levados em consideração os fatores socioeconômicos informados pelos participantes na pesquisa socioeconômica realizada na fase de inscrição.

- **Competência de área 1** - Construir significados para os números naturais, inteiros, racionais e reais.
- **Competência de área 2** - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.
- **Competência de área 3** - Construir noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.
- **Competência de área 4** - Construir noções de variação de grandezas para a compreensão da realidade e a solução de problemas do cotidiano.
- **Competência de área 5** - Modelar e resolver problemas que envolvem variáveis socioeconômicas ou técnico-científicas, usando representações algébricas.

- **Competência de área 6** - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.
- **Competência de área 7** - Compreender o caráter aleatório e não determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.

No ano de 2017, foi aprovada a lei nº13.415/2017, por meio da qual foram estabelecidas novas mudanças no ensino médio. O novo modelo traz uma nova organização curricular e a ampliação da carga horária de 800 para 1000 horas anuais e a oferta de diferentes possibilidades de escolha pelos estudantes, a partir da criação de itinerários formativos, serão incluídas formações técnicas e profissionais de forma a introduzir os estudantes no mercado de trabalho. Essas mudanças vão impactar diretamente a estrutura das avaliações de larga escala, principalmente do ENEM, que é a prova mais realizada pelos estudantes do ensino médio.

Contudo, mesmo com essas mudanças, o fator socioeconômico continua sendo um ponto importante a ser estudado e de forte impacto na sociedade, principalmente nas regiões territoriais nas quais cada itinerário será distribuído, tornando possível que determinados estudantes não tenham acesso a uma determinada área de estudo devido à região que reside.

5 Materiais e Métodos

5.1 Material

Com o objetivo de apresentar a influência dos fatores socioeconômicos no desempenho dos estudantes na prova de matemática do ENEM, a metodologia adotada consiste em analisar os dados da edição de 2019, a fim de observar fatores que afetam o desempenho na avaliação entre estudantes de diferentes níveis socioeconômicos.

Os microdados do ENEM¹ contêm as informações respondidas pelos estudantes durante a inscrição no exame, eles atendem por informações específicas sobre o estudante, informações sobre os itens, gabaritos e notas. Neste estudo, serão consideradas as variáveis que contêm relação com os fatores socioeconômicos e a nota dos estudantes na prova de matemática e suas tecnologias dos estudantes que realizaram a prova de matemática e preencheram o questionário socioeconômico, no intuito de que os resultados sejam os mais precisos possíveis.

A base de dados utilizada no estudo contém 5.095.270 linhas referentes aos estudantes e um total de 36 colunas referentes às variáveis. As variáveis que serão consideradas podem ser encontradas na tabela do Anexo A. Além disso, serão consideradas as variáveis do questionário respondido pelos estudantes especificamente sobre o perfil socioeconômico. Essas variáveis são representadas por Q001 até Q025 e, com essas variáveis, é possível ter informações como a escolaridade dos responsáveis, ocupação profissional dos responsáveis, informações sobre a família dos participantes e informações sobre a residência.

A análise dos dados será feita a partir da Análise de correspondência Múltipla (ACM), mesmo sendo feita uma análise geométrica e visual, o detalhamento dos cálculos pode exigir conhecimento de alguns pontos de estatística multivariada.

A partir do ano de 2018, a lei N^o 13.709, também conhecida como lei de proteção de dados, começou a garantir que alguns cuidados fossem tomados ao tratar com dados abertos, onde as empresas que coletam esse material são responsáveis por garantir o sigilo total das informações pessoais dos colaboradores. O INEP garante que os dados sejam disponibilizados de forma anônima, garantindo sigilo total e respeitando a privacidade dos participantes da avaliação.

Todas as análises foram desenvolvidas utilizando a linguagem R e na interface do Rstudio na versão 4.1.1 (2021-08-10) e todo o material utilizado para fazer as análises está disponível no repositório: <https://github.com/evandro-jr/TCC>.

¹ Disponível em: <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/enem>

5.2 Métodos

5.2.1 Análise de Correspondência Múltipla

A análise de correspondência múltipla (ACM) é utilizada para explorar as relações entre duas ou mais variáveis qualitativas. Segundo (PRADO, 2012, p. 18), “Na análise de correspondência múltipla, os dados são apresentados por meio de duas ou mais variáveis categóricas, onde a dimensionalidade está ligada às categorias de cada variável.”

E em sua aplicação a ACM, os dados devem estar distribuídos em uma tabela de contingência multidimensional que são ajustados na forma de uma matriz retangular, onde as linhas apresentam os objetos de estudo e as colunas as categorias de todas as variáveis. Ainda segundo (PRADO, 2012, p. 19), a ACM tem como objetivo:

Representar geometricamente as linhas e colunas da matriz resultante da tabela multidimensional em um subespaço ótimo de dimensão reduzida, de modo que, seja possível, observar as relações entre objetos, variáveis e categorias.(PRADO, 2012, p. 19).

Além disso a ACM é utilizada para analisar um conjunto de observações descritas por um conjunto de variáveis nominais, onde cada uma dessas variáveis possui vários níveis (Abdi Valentin, 2007 p. 2). O que vai de acordo com as variáveis que são trabalhadas no projeto, já que cada estudante tem em média 3 opções de resposta no questionário socioeconômico.

Uma das maiores dificuldades de se trabalhar com uma grande quantidade é a falta de poder computacional para processar esses dados em tempo hábil. E ao fazer o uso da ACM, segundo (Abdi Valentin, 2007 p. 2) “A ACM também pode acomodar variáveis quantitativas, recodificando - as como “caixas”. Por exemplo, uma pontuação com um intervalo de -5 a +5 poderia ser transformada em uma variável nominal com três níveis: menos que 0, igual a 0, ou maior que 0.”. Desta forma, fazer esse agrupamento de alguns estudantes é uma boa alternativa para lidar com um vasto números de dados.

A interpretação da ACM é baseada na proximidade entre os pontos em um gráfico de duas ou três dimensões. Quando dois pontos estão próximos um do outro, tendem a selecionar os mesmos níveis para as variáveis nominais. E na observação geométrica do gráfico, a proximidade entre duas variáveis, elas são definidas em dois casos:

Em primeiro lugar, a proximidade entre níveis de diferentes variáveis nominais faz com que esses níveis tendem a aparecer juntos nas observações. Segundo, porque os níveis da mesma variável nominal não podem ocorrer juntos, precisamos de um tipo diferente de interpretação para este caso. Aqui a proximidade entre os níveis significa que o grupos de observações associados a esses dois níveis são semelhantes.(Abdi Valentin, 2007 p. 8)

Assim, a partir da análise gráfica da base de dados por meio da ACM, serão observadas as variáveis que se mostram mais relevantes quando colocadas ao lado das notas de matemática dos estudantes.

5.2.2 Desenvolvimento Algébrico da Análise de Correspondência

Seja X , com os elementos x_{ij} , uma tabela de duas vias $I \times J$ de frequência ou sem contagem de escala. Tomando $I > J$ e assumindo que X possui a coluna J completa, as linhas e as colunas da tabela de contingência X correspondem a diferentes categorias de diferentes características. Se $\mathbf{n}$ é o total das frequências na matriz de dados X , primeiro construímos uma matriz de proporção $P = P_{ij}$ dividindo cada elemento de X por $\mathbf{n}$. Portanto,

$$P_{ij} = \frac{X_{ij}}{n}, i = 1, 2, \dots, I \quad ; \quad j = 1, 2, \dots, J$$

A matriz $\mathbf{P}$ é chamada de matriz de correspondência. Definimos os vetores de soma de linhas e colunas r e c respectivamente, e as diagonais das matrizes D_r e D_c com os elementos de r e c nas diagonais. Portanto:

$$r_i = \sum_{j=1}^J P_{ij} = \sum_{j=1}^J \frac{X_{ij}}{n} \quad , \quad i = 1, 2, \dots, I \quad \text{ou} \quad \begin{matrix} r \\ (I \times 1) \end{matrix} = \begin{matrix} P & 1_J \\ (1 \times J) & (J \times 1) \end{matrix}$$

$$c_j = \sum_{i=1}^I P_{ij} = \sum_{i=1}^I \frac{X_{ij}}{n} \quad , \quad j = 1, 2, \dots, J \quad \text{ou} \quad \begin{matrix} c \\ (J \times 1) \end{matrix} = \begin{matrix} P' & 1_I \\ (I \times 1) & (1 \times I) \end{matrix}$$

onde 1_J é $j \times 1$ e 1_I é $I \times 1$ são vetores de de 1's.

Temos as diagonais

$$D_r = \text{diag}(r_1, r_2, \dots, r_i)$$

$$D_c = \text{diag}(c_1, c_2, \dots, c_j)$$

e definimos a raiz quadrada das matrizes

$$D_r^{\frac{1}{2}} = \text{diag}(\sqrt{r_1}, \dots, \sqrt{r_i}) \quad D_r^{-\frac{1}{2}} = \text{diag}\left(\frac{1}{\sqrt{r_1}}, \dots, \frac{1}{\sqrt{r_i}}\right)$$

$$D_c^{\frac{1}{2}} = \text{diag}(\sqrt{c_1}, \dots, \sqrt{c_j}) \quad D_c^{-\frac{1}{2}} = \text{diag}\left(\frac{1}{\sqrt{c_1}}, \dots, \frac{1}{\sqrt{c_i}}\right)$$

para fins de dimensionamento. A análise de correspondência pode ser formulada a partir da criação de uma medida de dispersão, como um problema de mínimos quadrados ponderados para selecionar $\hat{P} = \{\hat{P}_{ij}\}$, uma matriz de classificação reduzida para minimizar

$$\sum_{i=1}^I \sum_{j=1}^J \frac{(P_{ij} - \hat{P}_{ij})^2}{x_i c_j} = \text{Tr}[(D_r^{-\frac{1}{2}} \cdot (P - \hat{P}) \cdot D_c^{-\frac{1}{2}}) \cdot (D_r^{-\frac{1}{2}} \cdot (P - \hat{P}) \cdot D_c^{-\frac{1}{2}})']$$

$$\text{Assim } \frac{(P_{ij} - \hat{P}_{ij})}{\sqrt{x_i c_j}} \text{ é o } (i, j) \text{ elemento de } D_r^{-\frac{1}{2}} \cdot (P - \hat{P}) \cdot D_c^{-\frac{1}{2}}$$

A análise de correspondência é primariamente uma técnica gráfica projetada para representar a associação de um espaço de baixa dimensão, assim ele pode ser utilizado como um método de dimensionamento e pode ser aproveitado como um complemento para outros métodos de análise. Uma referência para continuar os estudos em análise de correspondência é o livro *Multiple Correspondence Analysis* (ABDI e VALENTIN, 2007) que detalha o método e apresenta alguns exemplos, e para a parte computacional o livro *Practical Guide To Principal Component Methods in R* (KASSAMBARA, 2017) que apresenta o conteúdo de forma mais prática com aplicações na linguagem R.

6 Resultados e Discussão

6.1 Detalhamento dos Dados

Com o grande volume de dados fornecidos pelo INEP a respeito do ENEM, algumas estratégias possibilitam uma maior facilidade para trabalhar com essas informações. Durante o processo de tratamento de dados, foram selecionadas 44 variáveis, que estão representadas na tabela do anexo A, onde foram vistos somente os estudantes que possuem todas as informações respondidas no questionário socioeconômico e que estavam presentes no dia da prova de matemática. Com esse fracionamento, a base de dados ficou com uma dimensão de 2.120.104 linhas e 44 colunas, o que não prejudica os resultados pois a população continua com dados suficientemente grandes para serem trabalhados, e durante os testes feitos não foram apresentadas diferenças significativas em relação a base de dados completa. Além de possibilitar focar no grupo de interesse da pesquisa e reduzir a dimensão da base, facilitando o trabalho de processamento desse material.

Porém, como ainda existe uma limitação computacional, foram criadas novas variáveis para representar as notas dos estudantes nos 5 componentes curriculares, separando em 5 grupos diferentes. Nesse processo estão sendo transformadas variáveis quantitativas em variáveis qualitativas. As notas dos estudantes são representadas no intervalo (0, 1000) e para fazer a aplicação da ACM, essas variáveis foram transformadas e separadas em 5 níveis, conforme a tabela abaixo:

Tabela 2 – Nota Separada em Quintil

Notas	Representação
(0,421]	1
(421,467]	2
(467,524]	3
(524,606]	4
(606,986]	5

Essas variáveis foram criadas e agrupadas a base de dados, sendo nomeadas como: "nota-matematica", "nota-natureza", "nota-humanas", "nota-linguagem" e "nota-redacao". Após ser feita toda a parte de mineração de dados e manipulação das variáveis, é possível analisar as notas dos estudantes em relação as respostas do questionário socioeconômico, que são o objeto alvo do estudo.

Com o grande volume de dados utilizados, foram tomadas estratégias para tornar mais fácil o tratamento dos dados. As primeiras análises dos dados foram feitas em blocos, diminuindo o volume e possibilitando uma visão mais específica sobre a relação entre as variáveis de cada segmento. Os blocos foram separados seguindo a divisão que é feita pelo INEP na criação da base de dados, sendo definidos como bloco 1: Dados do Participante, bloco 2: Dados da Prova Objetiva e Bloco 3: Dados do questionário Socioeconômico

Tabela 3 – Divisão Blocos dos Dados

Bloco 1	Bloco 2	Bloco 3
NU-IDADE	NOTA-MATEMATICA	Q001
TP-SEXO	NOTA-NATUREZA	Q002
TP-COR-RACA	NOTA-HUMANAS	Q003
SG-UF-NACIMENTO	NOTA-LINGUAGENS	Q004
TP-ST-CONCLUSAO	NOTA-REDACAO	Q005
TP-ESCOLA		Q006
TP-ENSINO		Q007
NOTA-MATEMATICA		Q008
		Q009
		Q010
		Q011
		Q012
		Q013
		Q014
		Q015
		Q016
		Q017
		Q018
		Q019
		Q020
		Q021
		Q022
		Q023
		Q024
		Q025
	NOTA-MATEMATICA	
	NOTA-NATUREZA	
	NOTA-HUMANAS	
	NOTA-LINGUAGENS	
	NOTA-REDACAO	

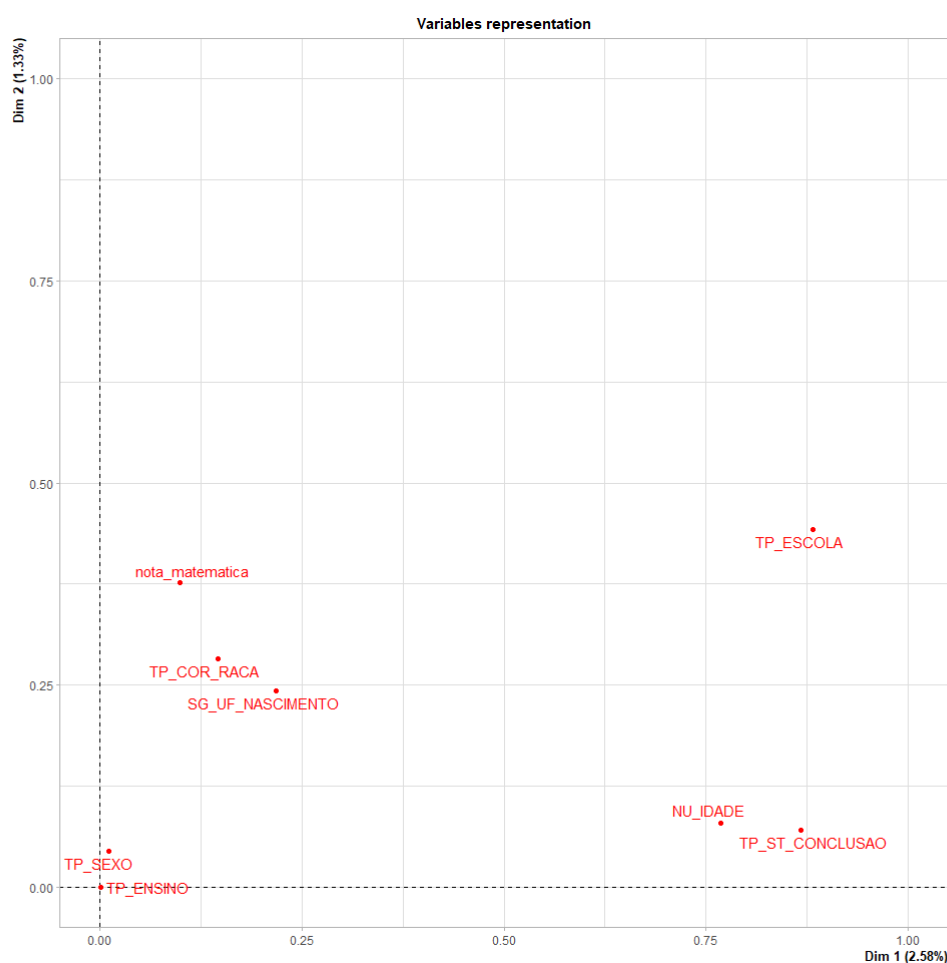
Os blocos foram definidos fundamentados no questionário que os estudantes respondem ao efetuarem a inscrição na prova, seguindo o dicionário de dados disponibilizado junto com a base, as variáveis são divididas em três partes, a primeira corresponde aos dados do estudante, a segunda às notas de cada participante, que nesse caso estão sendo consideradas as variáveis criadas referente as notas, e o terceiro bloco o questionário socioeconômico.

Com essa divisão dos dados em blocos é possível ter uma noção maior do desempenho dos estudantes em relação a cada um dos blocos montados, possibilitando ter uma noção de como as variáveis mais relevantes para o estudo se comportavam, guiando o caminho da pesquisa nas futuras análises.

6.1.1 Análise de Múltiplas Correspondências - MCA

Como já visto, a Análise de Correspondência tem como objetivo visualizar geometricamente as informações da base de dados. Assim, após a manipulação dos dados e construção dos blocos foram construídos gráficos para melhor visualizar os resultados, buscando encontrar a proximidade entre as diferentes variáveis qualitativas em relação a nota de matemática. Devida a dificuldade para trabalhar com o grande volume de dados, foram criados três gráficos, um para representar cada bloco. Assim é possível ter uma análise mais específica sobre as informações, sem perder nenhum nível de interpretação e reduzindo a quantidade de processamento que seria necessário.

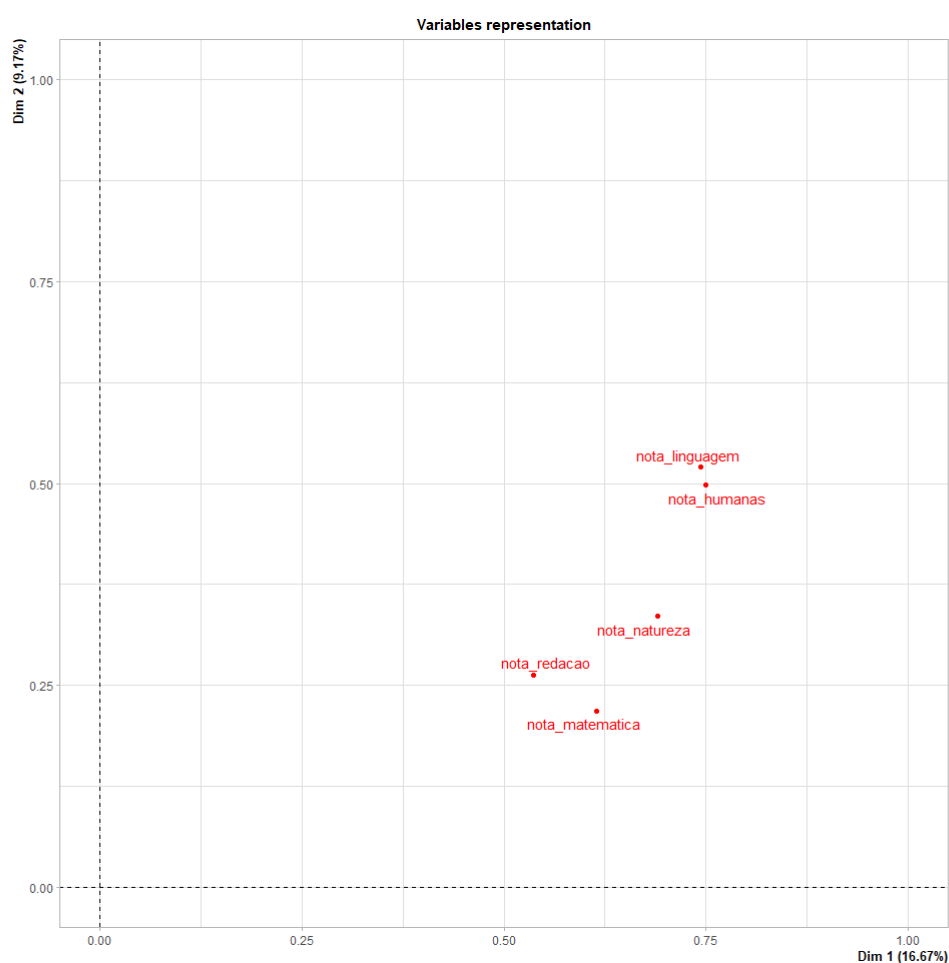
Figura 1 – Bloco 1 - Informação dos Participantes



O primeiro gráfico é referente ao bloco 01, que trata das informações sobre o estudante. Nesse primeiro bloco consegue-se ver a proximidade entre as variáveis referentes a nota de matemática, a cor do participante e a região onde ele nasceu, mostrando que existe uma influência desses indicadores em relação a nota do estudante no componente de matemática. Também é possível observar mais três grupos que não apresentam grande relação com a nota de matemática, então o sexo, o tipo de ensino, o tipo de escola, a idade e a situação de conclusão do ensino médio não apresentam influência na nota de matemática.

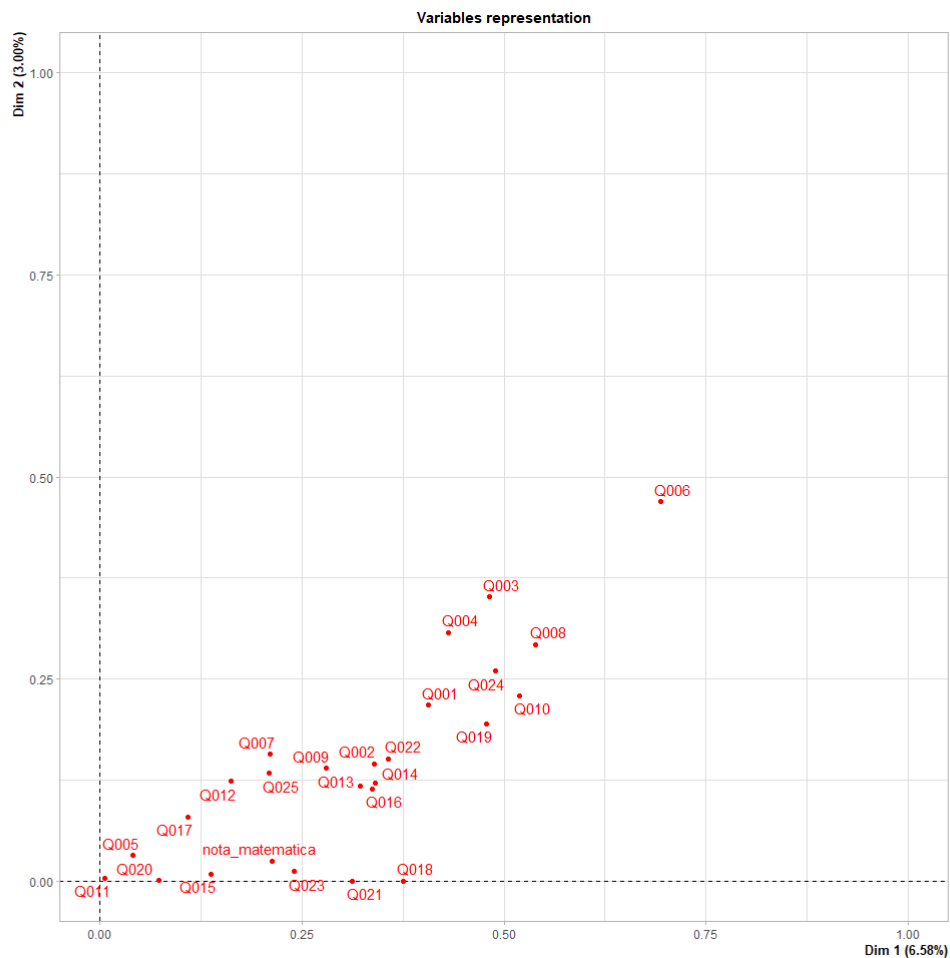
Trazendo para a realidade, os dados mostram que a região onde o estudante mora tem uma relação com sua nota, o que vai de acordo com o que foi mostrado anteriormente, onde pessoas que moram em regiões melhor estruturadas tendem a ter acesso a um ensino melhor qualidade. Além da cor dos estudantes, mostrando que a política de cotas aprovada pela LEI N^o 12.711, de 29 de agosto de 2012, tem um papel importante para tornar o acesso a educação de qualidade mais igualitário.

Figura 2 – Bloco 2 - Nota dos Participantes



O segundo gráfico, referente ao bloco 2 que trata da nota dos participantes, mostra basicamente dois blocos, um contendo as notas de linguagens e ciências humanas, que não apresentam grande influência na nota de matemática e o segundo bloco contendo a nota da redação, nota de matemática e a nota de ciências da natureza. Mostrando uma tendência de desempenho do participante em disciplinas consideradas de "exatas" e de "humanas", além da nota de redação que está associada as disciplinas exatas, o que pode ser explicado pelo modelo de avaliação da redação que é dividida em 5 competências, possibilitando que seja feita seguindo uma metodologia bem estruturada para obter um ótimo resultado.

Figura 3 – Bloco 3 - Questionário Socioeconômico



O terceiro gráfico, referente ao bloco 3, trata das informações respondidas no questionário socioeconômico dos participantes. Pode ser visto três grandes grupos, onde tem-se a variável Q006 mais distante da nota de matemática, outro grupo com as variáveis Q003, Q004, Q005, Q008, Q024, Q001, Q010 e Q019 que se aproximam da nota de matemática mas ainda se encontram distantes, e o terceiro, grupo mais próximo, contendo a maioria das variáveis. Dentro desse terceiro grupo ainda pode ser feita uma divisão, mostrando que as variáveis mais próximas da nota de matemática são Q015, Q021, Q018 e Q023.

Entre essas variáveis que apresentam maior influência na nota de matemática, a maioria são variáveis referentes a bens de consumo, o que torna visível a influência que a estrutura da residência está diretamente relacionada com a nota final em matemática. E como foi visto no primeiro gráfico, a região onde o participante nasceu tem relação com sua nota, mostrando que moradores de regiões mais vulneráveis e com menos estrutura residencial, podem vir a apresentar pior desempenho.

Considerações Finais

Com esse estudo, foi possível confirmar algumas informações que foram mostradas por alguns autores que atuam na área, mostrando que existem fatores socioeconômicos que influenciam no desempenho do participante na nota de matemática. E os dados mostraram que fatores como região onde o participante nasceu, cor e variáveis referentes a estrutura de sua residência tem influência direta com seu rendimento. Como também foi possível observar, fatores como o sexo, idade, situação da conclusão, tipo de escola e o tipo de ensino não apresentaram tanta influência sobre a desempenho em matemática.

Por meio do real impacto que a situação socioeconômica pode ter no desempenho do participantes, não pode - se deixar de intervir para evitar que o processo educacional seja prejudicado, como destaca CARVALHO(2017, p.140):

"[...]a escola precisaria ter o cuidado para não legitimar as desigualdades sociais por meio da educação, como nos alerta Bourdieu (1998), para quem os problemas econômicos e sociais não poderiam ser superados apenas pela educação, mas, sim, considerando também a origem social dos estudantes."

Como foi visto a influência da região onde o participante nasceu, pode ser feita uma visão mais detalhada sobre essas variáveis. Comparando as regiões onde cada um nasceu, reside e estuda. Já que o Brasil é um país heterogêneo, com diferentes contextos e culturas, regiões podem vir a ter diferentes condições de qualidade de ensino, como foi possível ver no estudo de CARVALHO(2017, p.141), onde as regiões norte e nordeste apresentavam maior vulnerabilidade, sendo necessário adotar políticas mais eficazes para possibilitar uma melhor qualidade de ensino.

Além disso, pesquisas futuras podem ser realizadas no sentido de aprofundar a análise dos fatores socioeconômicos, incluindo variáveis que não foram utilizadas nesse estudo como a influência de outros fatores socioeconômicos, buscando diferentes perspectivas, enriquecendo a pesquisa. Outra sugestão é tratar das demais áreas do conhecimento, buscando uma análise mais profunda do desempenho do estudantes em todos os componentes avaliados na prova.

A partir das análises, espera - se que as informações extraídas sejam aproveitadas para a construção de modelos preditivos ou explicativos, buscando poder observar o perfil desses estudantes, impulsionando e intervindo nos discentes que apresentam pior desempenho e acompanhando os estudantes com perfil de bom desempenho para que se mantenham nessa situação.

Por fim, essas informações são importantes para que gestores escolares possam intervir de forma mais eficiente, tornando mais claro o perfil socioeconômico dos estudantes, podendo criar, implementar ou adaptar metodologias que se enquadram melhor em cada realidade, buscando não deixar o resultado final da avaliação como o principal resultado do processo educacional, visando o desenvolvimento do estudante ao final do ciclo educacional básico.

Referências

ABDI, VALENTIN. **Multiple Correspondence Analysis**. Neil Salkind (Ed.) (2007). Encyclopedia of Measurement and Statistics. Program in Cognition and Neurosciences, MS: Gr.4.1, The University of Texas at Dallas.

ALBOUKADEL KASSAMBARA, FABIAN MUNDT (2020). **factoextra: Extract and Visualize the Results of Multivariate Data Analyses**. R package version 1.0.7. <<https://CRAN.R-project.org/package=factoextra>> .

AGRESTI, Alan. **An Introduction to Categorical Data Analysis**. 3^a ed.
Florida, Estados Unidos: Wielely, 2019, p. 65-88

BRASIL. Ministério da Educação. **Lei nº13.415**. 16 de fevereiro de 2017. Brasília - DF, 2017. Disponível em: <www.planalto.gov.br/ccivil03/ato2015-2018/2017/lei/l13415.htm> . Acesso em julho de 2021.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, [2016].
Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>
. Acesso em: Novembro de 2021.

BRASIL. **Matriz de Referência para o Enem (2015)**.
Disponível em: <http://portal.inep.gov.br/matriz-de-referencia>. Acesso em Julho 2021.

BRASIL. **Lei de Diretrizes e Bases da Educação Nacional Nº 9.394** Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/19394.htm> . Acesso em novembro de 2021

BRASIL. Ministério da Educação. **PORTARIA MEC Nº 438**. 28 de maio de 1998. Brasília - DF. Disponível em: <http://www.crmariocovas.sp.gov.br/pdf/diretrizes_p0178-0181_c.pdf> . Acesso em janeiro de 2022.

BRASIL. **LEI Nº 9.131**. 24 de novembro de 1995. Brasília - DF. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/19131.htm>. Acesso em: Janeiro de 2022.

BRASIL. **LEI Nº 13.702**. 14 de agosto de 2018 .Brasília - DF. Disponível em:
<http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13702.htm>. Acesso em:
Janeiro de 2022

BRASIL. **LEI Nº 378**. 13 de janeiro de 1937. Rio de Janeiro. Disponível em:
<https://www2.camara.leg.br/legin/fed/lei/1930-1939/lei-378-13-janeiro-1937-398059-publicacaooriginal-6-pe.html>>. Acesso em: Janeiro de 2022

BRASIL. **LEI Nº 580**. 30 de julho de 1938 .Rio de Janeiro. Disponível em: <<https://www2.camara.leg.br/legin/fed/declei/1930-1939/decreto-lei-580-30-julho-1938-350924-publicacao.html>>. Acesso em: Janeiro de 2022

BRASIL. **LEI Nº 9.448** . 14 de março de 1997 .Brasília - DF. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9448.htm>. Acesso em: Janeiro de 2022

BRASIL. **LEI Nº 12.711**. 29 de agosto de 2012. Brasília - DF? disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12711.htm>. Acesso em: Janeiro de 2022.

BRASIL. Secretaria de Educação Fundamental.**Parâmetros Curriculares Nacionais**. Brasília - DF : MEC/SEF, 1998.

CARVALHO, Francisca Eliane Dias de. **Fatores socioeconômicos associados ao desempenho dos estudantes na prova de redação do Exame Nacional do Ensino Médio (ENEM)**. 2017. 179f. –Tese (Doutorado) – Universidade Federal do Ceará, Programa de Pós-graduação em Educação Brasileira, Fortaleza (CE), 2017.

KASSAMBARA. **Practical Guide To Principal Component Methods in R: PCA, M(CA), FAMD, MFA, HCPC, factoextra**. Volume 2 de Multivariate Analysis. Editora STHDA, 2017.

LIMA, PRICILA DA SILVA NEVES et al. **Análise de dados do Enade e Enem: uma revisão sistemática da literatura**. Avaliação: Revista da Avaliação da Educação Superior (Campinas) [online]. 2019, v. 24, n. 1 [Acessado 20 Agosto 2021] , pp. 89-107. Disponível em: <<<https://doi.org/10.1590/S1414-40772019000100006>>.Epub25\protect\unhbox\voidb@x\protect\penalty\@M\Abr\protect\unhbox\voidb@x\protect\penalty\@M\2019.ISSN1982-5765.<https://doi.org/10.1590/S1414-40772019000100006>.>

MATT DOWLE, ARUN SRINIVASAN (2021). **data.table: Extension of ‘data.frame’**. R package version 1.14.2. <https://CRAN.R-project.org/package=data.table>.

MAX KUHN (2021).**caret: Classification and Regression Training** . R package version 6.0-90. <https://CRAN.R-project.org/package=caret>.

PRADO, M. V. B. **Métodos de análise de correspondência múltipla: estudo de caso aplicado à avaliação da qualidade do café**. 2012. 79 p. Dissertação (Mestrado em Estatística e Experimentação)-Universidade Federal de Lavras, Lavras, 2012.

RABELO, Mauro. **Avaliação Educacional: fundamentos, metodologias e aplicações no contexto brasileiro**. Coleção PROFMAT. Sociedade Brasileira de Matemática. 1ª edição. Rio de Janeiro. 2013.

SAVIANI. **O Inep, o diagnóstico da educação brasileira e a Rbep**. R. bras. Est. pedag., Brasília, v. 93, n. 234, [número especial], p. 291-322, maio/ago. 2012. Disponível em: <http://rbep.inep.gov.br/ojs3/index.php/rbep/article/view/3565>. Acesso

em: Janeiro 2022.

SEBASTIAN LE, JULIE JOSSE, FRANCOIS HUSSON (2008). **FactoMineR: An R Package for Multivariate Analysis** . Journal of Statistical Software, 25(1), 1-18. 10.18637/jss.v025.i01

SUDBRACK, E. M.; COCCO, E. M. **AValiação em Larga Escala NO BRASIL: POTENCIAL INDUTOR DE QUALIDADE? / Large scale evaluation in Brazil: potential quality inducer?**. Roteiro, [S. l.], v. 39, n. 2, p. 347–370, 2014. Disponível em: <<https://portalperiodicos.unoesc.edu.br/roteiro/article/view/4231>> . Acesso em: 3 ago. 2021.

SOARES, Jose Francisco. ANDRADE, Renato Júdice de Andrade. **Nível socioeconômico, qualidade e eqüidade das escolas de Belo Horizonte**. Ensaio: aval. pol. públ. Educ., Rio de Janeiro, v.14, n.50, p. 107-126, jan./mar. 2006

SOARES, Jose Francisco. **Qualidade e eqüidade na educação básica brasileira: fatos e possibilidades** Soares, José Francisco. "Qualidade e equidade na educação básica brasileira: fatos e possibilidades."Os desafios da educação no Brasil. Rio de Janeiro: Nova Fronteira 1 (2005): 87-114.

WICKMAM et al., (2019). **Welcome to the tidyverse**. Journal of Open Source Software, 4(43), 1686, <https://doi.org/10.21105/joss.01686>

Apêndices

APÊNDICE A – Tabela de Descrição das Variáveis

Tabela referente as informações das variáveis utilizadas no estudo. Contendo as variáveis, sua descrição e o tipo de cada indicador.

Descrição Variáveis		
Variável	Descrição	Tipo
NU-ANO	Ano do exame	Qualitativa Ordinal
NU-IDADE	Idade do estudante	Qualitativa Ordinal
TP-SEXO	Sexo do estudante	Qualitativa
TP-COR-RACA	Cor/Raça do estudante	Qualitativa
SG-UF-NASCIMENTO	Sigla da Unidade da Federação de nascimento	Qualitativa
TP-ST-CONCLUSAO	Situação de conclusão do Ensino Médio	Qualitativa
TP-ESCOLA	Tipo de escola do Ensino Médio	Qualitativa
TP-ENSINO	Tipo de instituição que concluiu ou concluirá o Ensino Médio	Qualitativa
TP-PRESENCA-MT	Presença na prova objetiva de Matemática	Qualitativa
NU-NOTA-CN	Nota da prova de Ciências da Natureza	Quantitativa
NU-NOTA-CH	Nota da prova de Ciências Humanas	Quantitativa
NU-NOTA-LC	Nota da prova de Linguagens e Códigos	Quantitativa
NU-NOTA-MT	Nota da prova de Matemática	Quantitativa
NU-NOTA-REDACAO	Nota da prova de redação	Quantitativa
Q001	Até que série seu pai, ou o homem responsável por você, estudou?	Qualitativa
Q002	Até que série sua mãe, ou a mulher responsável por você, estudou?	Qualitativa
Q003	A partir da apresentação de algumas ocupações divididas em grupos ordenados, indique o grupo que contempla a ocupação mais próxima da ocupação do seu pai ou do homem responsável por você. (Se ele não estiver trabalhando, escolha uma ocupação pensando no último trabalho dele).	Qualitativa
Q004	A partir da apresentação de algumas ocupações divididas em grupos ordenados, indique o grupo que contempla a ocupação mais próxima da ocupação da sua mãe ou da mulher responsável por você. (Se ela não estiver trabalhando, escolha uma ocupação pensando no último trabalho dela).	Qualitativa

Descrição Variáveis		
Variável	Descrição	Tipo
Q005	Incluindo você, quantas pessoas moram atualmente em sua residência?	Qualitativa
Q006	Qual é a renda mensal de sua família? (Some a sua renda com a dos seus familiares.)	Qualitativa
Q007	Em sua residência trabalha empregado(a) doméstico(a)?	Qualitativa
Q008	Na sua residência tem banheiro?	Qualitativa
Q009	Na sua residência tem quartos para dormir?	Qualitativa
Q010	Na sua residência tem carro?	Qualitativa
Q011	Na sua residência tem motocicleta?	Qualitativa
Q012	Na sua residência tem geladeira?	Qualitativa
Q013	Na sua residência tem freezer (independente ou segunda porta da geladeira)?	Qualitativa
Q014	Na sua residência tem máquina de lavar roupa? (o tanquinho NÃO deve ser considerado)	Qualitativa
Q015	Na sua residência tem máquina de secar roupa (independente ou em conjunto com a máquina de lavar roupa)?	Qualitativa
Q016	Na sua residência tem forno micro-ondas?	Qualitativa
Q017	Na sua residência tem máquina de lavar louça?	Qualitativa
Q018	Na sua residência tem aspirador de pó?	Qualitativa
Q019	Na sua residência tem televisão em cores?	Qualitativa
Q020	Na sua residência tem aparelho de DVD?	Qualitativa
Q021	Na sua residência tem TV por assinatura?	Qualitativa
Q022	Na sua residência tem telefone celular?	Qualitativa
Q023	Na sua residência tem telefone fixo?	Qualitativa
Q024	Na sua residência tem computador?	Qualitativa
Q025	Na sua residência tem acesso à Internet?	Qualitativa

Documento Digitalizado Público

TCC EVANDRO

Assunto: TCC EVANDRO
Assinado por: Antonio Neto
Tipo do Documento: Trabalho de Conclusão de Curso - TCC
Situação: Finalizado
Nível de Acesso: Público
Tipo do Conferência: Documento Original

Documento assinado eletronicamente por:

■ **Antonio Dantas Costa Neto**, COORDENADOR DE CURSO - FUC1 - ES-GRAD-LM, em 18/08/2022 11:48:30.

Este documento foi armazenado no SUAP em 18/08/2022. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifb.edu.br/verificar-documento-externo/> e forneça os dados abaixo:

Código Verificador: 377194

Código de Autenticação: 60b551a789

