

PATRÍCIA MARIA LIVRARI

**INFLUÊNCIA DE VARIÁVEIS SOCIOECONÔMICAS NA NOTA DA
PROVA DE CIÊNCIAS DA NATUREZA E SUAS TECNOLOGIAS DO
ENEM – 2022 VIA *MACHINE LEARNING***

PATRÍCIA MARIA LIVRARI

**INFLUÊNCIA DE VARIÁVEIS SOCIOECONÔMICAS NA NOTA DA
PROVA DE CIÊNCIAS DA NATUREZA E SUAS TECNOLOGIAS DO
ENEM – 2022 VIA *MACHINE LEARNING***

Dissertação apresentada ao Programa de Pós-graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá como requisito parcial para obtenção do título de mestre em Bioestatística.

Orientador: Prof. Dr. Diogo Francisco Rossoni, Universidade Estadual de Maringá.

Membro do PBE: Prof. Dr. Brian Alvarez Ribeiro de Melo, Universidade Estadual de Maringá.

Membro Externo: Prof^a. Dr^a. Walkiria Maria de Oliveira Macerau, Universidade Estadual de Maringá.

Universidade Estadual de Maringá

Departamento de Estatística – DES

Programa de Pós-graduação em Bioestatística - PBE

Maringá – PR
2024

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá - PR, Brasil)

L788i Livrari, Patrícia Maria
Influência de variáveis socioeconômicas na nota da prova de ciências da natureza e suas tecnologias do ENEM - 2022 via *Machine Learning* / Patrícia Maria Livrari. -- Maringá, PR, 2024.
34 f. : il. color., tabs.

Orientador: Prof. Dr. Diogo Francisco Rossoni.
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Departamento de Estatística, Programa de Pós-Graduação em Bioestatística, 2024.

1. Exame Nacional do Ensino Médio (Brasil) . 2. Estatística. 3. Aprendizagem de máquina. I. Rossoni, Diogo Francisco, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Estatística. Programa de Pós-Graduação em Bioestatística. III. Título.

CDD 23.ed. 519.5

PATRÍCIA MARIA LIVRARI

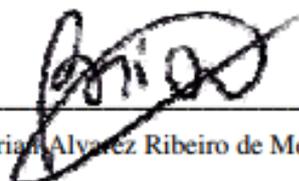
**Influência de variáveis socioeconômicas na nota da prova de
Ciências da Natureza e suas Tecnologias do ENEM – 2022 via
*Machine Learning***

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

BANCA EXAMINADORA



Prof. Dr. Diogo Francisco Rossoni
Universidade Estadual de Maringá – PBE/UEM



Prof. Dr. Brian Alvarez Ribeiro de Melo
Universidade Estadual de Maringá – PBE/UEM



Prof. Dra. Walciria Maria de Oliveira Macerau
Universidade Estadual de Maringá – DES/UEM

Maringá, 13 de março de 2024

À MEMORIA DE MEU PAI RUBENS LIVRARI
HOMEM SIMPLES... DE POUCO ESTUDO... MEU GUERREIRO... O MEU HERÓI!
QUE NOS DEIXOU A POUCO MAIS DE 1 ANO
QUE DESDE MUITO CEDO ME ENSINOU A REALIZAR MEUS SONHOS
QUE MESMO DO SEU JEITO... NUNCA ME DEIXOU SEM SEU AMPARO E PROTEÇÃO
PAI: ESTE MESTRADO É PARA O SENHOR!

AGRADECIMENTO

Ao meu bondoso Deus que em sua infinita graça e misericórdia me deu fé e forças para não desistir deste curso, mesmo diante da perda e dos desestimulos externos.

À minha mãezinha querida Terezinha Elizabeth por seu ânimo, alegria e constante vontade de viver e de vencer.

Ao meu esposo Edrei Daniel Vieira, nossos filhos Enzo Livrari Vieira e Giulia Livrari Vieira e minha prima Dr^a. Angela Francine Fusa, que estiveram ao meu lado todos os dias e me incentivaram a fazer e concluir este mestrado.

Ao meu orientador Prof. Dr. Diogo Francisco Rossoni que aceitou o desafio de me orientar na retomada integral deste trabalho.

Aos colegas do DES (Departamento de Estatística) e do PBE (Programa de Pós-Graduação em Bioestatística) da UEM, pelas importantes contribuições em minha carreira docente e acadêmica.

Por fim, e não menos importante, minha gratidão ao Professor Dr. Brian Alvarez Ribeiro de Melo e à Prof^a. Walkiria Maria de Oliveira Macerau por aceitaram o convite de avaliar este trabalho e pelas contribuições que com certeza enriquecerão ainda mais este trabalho.

RESUMO

O Exame Nacional do Ensino Médio (ENEM) é reconhecido como o segundo maior sistema de avaliação de ensino do mundo, ficando atrás apenas do exame chinês GAOKAO. Aplicado pela primeira vez em 1998 e com 64 questões de conhecimentos e 1 (um) questionário socioeconômico, o Exame passou por diversas melhorias e atualmente é composto por 180 questões objetivas, sendo 45 para cada uma das 4 áreas do conhecimento, além de 1 (uma) redação e 1 (um) questionário socioeconômico. Tornou-se também, um dos maiores mecanismos de inclusão dos jovens brasileiros no Ensino Superior público ou privado por meio do Sistema de Seleção Unificada (SISU), do Programa Universidade para Todos (ProUni) ou do Fundo de Financiamento Estudantil (FIES). É sabido que ao longo dos anos os milhares de dados do ENEM tem subsidiado inúmeros trabalhos científicos. Deste modo, este trabalho não pretende esgotar o assunto, mas com base nos dados do questionário sócioeconômico do ENEM de 2022 objetiva identificar as variáveis contidas no questionário, relacionando-a com a nota alcançada pelos estudantes na prova de *Ciências da Natureza e suas tecnologias*, utilizando a técnica de *Machine Learning*. Os resultados mostraram que 68% dos estudantes com a renda mensal familiar inferior ou igual a R\$3.030,00 alcançaram notas abaixo da média geral de 497,4 pontos e 32% dos estudantes com renda mensal familiar superior a R\$ 3.030,00 tiveram suas notas acima da média.

Palavras-chave: ENEM; Estatística, Machine Learning; Decision Tree.

ABSTRACT

The National High School Examination (*Exame Nacional do Ensino Médio - ENEM*) is recognized as the second largest education assessment system in the world, second only to the Chinese exam GAOKAO. Applied for the first time in 1998 and with 64 knowledge questions and 1 (one) socioeconomic questionnaire, the ENEM exam has undergone several improvements and currently consists of 180 objective questions, 45 for each of the 4 areas of knowledge, in addition to 1 (one) essay and 1 (one) socioeconomic questionnaire. It has also become one of the largest mechanisms for the inclusion of young Brazilians in public or private Higher Education through the Unified Selection System (*Sistema de Seleção Unificada - SISU*), the University for All Program (*Programa Universidade para Todos - ProUni*) or the Student Financing Fund (*Fundo de Financiamento Estudantil - FIES*). It is known that over the years, thousands of ENEM data have supported countless scientific works. Therefore, this work does not intend to exhaust the subject, but based on data from the 2022 *ENEM* socioeconomic questionnaire, aims to identify the variables contained in the questionnaire, relating it to the grade achieved by students in the Natural Sciences and its technologies test, using the "Machine Learning" technique. The results showed that 68% of students with a monthly family income of less than or equal to R\$3,030.00 achieved grades below the general average of 497.4 points and 32% of students with a monthly family income of more than R\$3,030.00 had their above average grades.

Keywords: ENEM; Statistic, Machine Learning; Decision Tree.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de estrutura de uma Árvore de Decisão.....	15
Figura 2 – Gráfico de Barras referente a questão socioeconômica Q005 do ENEM 2022 gerada pela linguagem de programação R.....	20
Figura 3 – Gráfico de Barras referente a questão socioeconômica Q006 do ENEM 2022 gerada pela linguagem de programação R.....	21
Figura 4 – Boxplot referentes às notas de Ciências da Natureza e suas tecnologias do ENEM 2022 gerada pela linguagem de programação R.....	22
Figura 5 – Gráfico de Pirâmide referentes às notas de Ciências da Natureza e suas tecnologias do ENEM 2022 e a questão socioeconômica Q001 gerada pela linguagem de programação R.....	22
Figura 6 – Importância do Classificador para as covariáveis socioeconômicas gerada pela linguagem de programação R.....	23
Figura 7 – Modelo: Árvore de Classificação das notas dos estudantes do ENEM 2022 na prova de Ciências da Natureza de acordo com as variáveis socioeconômicas gerada pela linguagem de programação R.....	24

LISTA DE TABELAS

Tabela 1 – Medidas utilizadas em um estudo de validação de Modelos de Classificação que produzem respostas Dicotomizadas..	17
Tabela 2 – Referente as 25 questões socioeconômicas do ENEM 2022 gerada pela linguagem de programação pelo R	19
Tabela 3 – Classificação dos Ramos do Modelo	23
Tabela 4 - Métricas para validar o Modelo	24

SUMÁRIO

1. Introdução	12
2. Métodos	14
2.1 Dados.....	14
2.2 Metodologia.....	14
3. Resultados e Discussões	18
3.1 Estatística Descritiva.....	18
3.1.1 Variáveis Socioeconômicas - ENEM 2022.....	20
3.1.2 Variável Nota de Ciências da Natureza e suas tecnologias – ENEM 2022....	21
3.2 Modelo de <i>Machine Learning</i>	23
3.2.1 Modelo e Validação.....	23
4. Conclusão	26
Referências	28
Anexos	30

Capítulo 1

INTRODUÇÃO

O Exame Nacional do Ensino Médio (ENEM) foi criado pelo Ministério da Educação, em 1998, no governo do então presidente, Fernando Henrique Cardoso, pela Portaria MEC nº 438 de 25/05/1998. O exame tinha como objetivo avaliar o desempenho dos estudantes da educação básica, após terem concluído o Ensino Médio. Conforme o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira INEP (2020) em sua primeira edição, em 1998, o ENEM teve a participação de 115.575 estudantes, em 184 municípios brasileiros, a prova era composta por 64 questões e um questionário socioeconômico para ser respondido.

No ano de 2004, o ENEM recebeu uma nova e importante atribuição, pois com sua nota, os estudantes passariam a ter acesso a bolsas de estudos parciais e integrais em faculdades e universidades privadas, por meio do Programa Universidade para Todos (ProUni). Este programa foi criado pelo governo do presidente Luíz Inácio Lula da Silva, pela Lei nº 11.096 de 13/01/2005 e abriu caminho à inclusão de milhares de jovens de baixa renda no ensino superior do país. Para o INEP (2005), a implantação do ProUni fez com que a quantidade de inscritos no ENEM passasse de 1.552.316, em 2004, para 3.004.491, em 2005. Em 2014, o ENEM obteve um recorde histórico com 9,5 milhões de inscrições. Este crescimento exponencial do número de inscritos e participantes no ENEM foi reconhecido mundialmente como o segundo maior sistema de avaliação em larga escala do mundo, ficando atrás apenas do exame chinês: GAOKAO, segundo GARCIA E FRANZOLIN (2021).

Em 2009, implantou-se o Novo ENEM sendo realizado em dois dias, composto por 180 questões objetivas, 45 para cada área do conhecimento, a saber: 1) *Linguagem, Códigos e suas tecnologias e Redação*; 2) *Ciências Humanas e suas Tecnologias*; 3) *Ciências da Natureza e suas tecnologias*; 4) *Matemática e suas tecnologias*. Vale ressaltar que neste mesmo ano, foi criado o Sistema de Seleção Unificado (SISU) que passou a utilizar a nota do ENEM como forma de ingresso do estudante no ensino superior público e, também, a partir de 2009, o INEP passou a utilizar a Teoria de Resposta ao Item (TRI) para a correção das provas.

Dentre as 4 áreas do conhecimento do ENEM, o interesse deste trabalho foi pela área do conhecimento *Ciências da Natureza e suas tecnologias*, que é composta por três relevantes disciplinas: Química, Física e Biologia. SANTOS E OLIOSI (2013) destacam a necessidade do ensino de *Ciências da Natureza e suas tecnologias*, pois o cidadão adquire pensamento crítico

das ciências e de suas implicações na sociedade contemporânea, além de ser fundamental para entender a relação com fatores históricos e sociais no processo de constituição das ciências da natureza de uma época específica.

Outro aspecto que chamou atenção foi a diferença significativa no desempenho dos estudantes de *Ciências da Natureza e suas tecnologias*, em relação às demais áreas do conhecimento. Segundo NAVARRO, IANELLO, MUNERATTO E WATANABE (2021, p. 14), em um estudo sobre o ENEM de 2017 destacou que “os dados apontaram que o desempenho desses jovens candidatos em ciências da natureza, comparativamente, com as outras áreas do conhecimento do ENEM, é marcante no que tange à quantidade de alunos que não chegam aos 600 pontos nesse exame”. E ainda, outro aspecto foi o baixo desempenho dos estudantes em *Ciências da Natureza e suas tecnologias*. ARAÚJO, GUIMARÃES E WARTHA (2021), destacam em sua pesquisa sobre o ENEM de 2012, 2013 e 2014, que “Os resultados encontrados apontam para uma dificuldade crescente no processo de aprendizagem, do sistema de ensino brasileiro para o ensino de *Ciências da Natureza e suas tecnologias*, visto o baixo desempenho dos participantes do exame.”

Também, alguns autores apontaram para a interação entre nota dos alunos e as questões socioeconômicas. Por exemplo, PIRES (2015), ao estudar os dados do ENEM de 2012, descobriu que “possuir renda familiar acima de 20 salários mínimos ou ter um dos pais com curso superior aumentam substancialmente as chances de alunos terem melhor desempenho nas provas objetivas e na redação”. Novamente, ARAÚJO, GUIMARÃES E WARTHA (2021), numa abordagem quantitativa, concluíram que “os resultados também apontam que o perfil socioeconômico dos participantes do ENEM de 2012 a 2014 do Estado de Sergipe são de pessoas que possuem renda familiar mais baixa normalmente associada a uma menor escolaridade dos pais, influenciando significativamente no desempenho”.

Diante do exposto, o objetivo deste artigo é identificar as variáveis contidas no questionário socioeconômico do ENEM 2022, relacionando-a com a nota alcançada pelos estudantes na prova de *Ciências da Natureza e suas tecnologias*, utilizando a técnica de *Machine Learning*.

Capítulo 2

Métodos

2.1. Dados

Os dados deste trabalho são referentes ao ENEM 2022, realizado nos dias 13/11/2021 e 20/11/2021. Eles foram fornecidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). São dados de acesso livre e estão expostos na plataforma digital <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/enem>. Neste ano foram mais de 3,4 milhões de inscritos, dos quais, 2.490.880 fizeram a prova no 1º dia, ou seja, uma abstenção de 26,7% e 2.355.395 fizeram a prova no 2º dia, uma abstenção de 31%.

Para a realização deste estudo, optou-se pela análise somente das notas da prova de *Ciências da Natureza e suas tecnologias* que envolvem questões de Biologia, Física e Química e pelo questionário socioeconômico, preenchido no ato da inscrição.

As análises propostas foram efetuadas através da linguagem de programação estatística **R**. Para se obter a população desejada manipulou-se os dados por meio de um filtro composto por estudantes com presença na prova de Ciências da Natureza, destes retiraram-se os treineiros e aqueles tiraram nota 0 (zero), reduzindo assim o banco de dados para 1.933.149 observações.

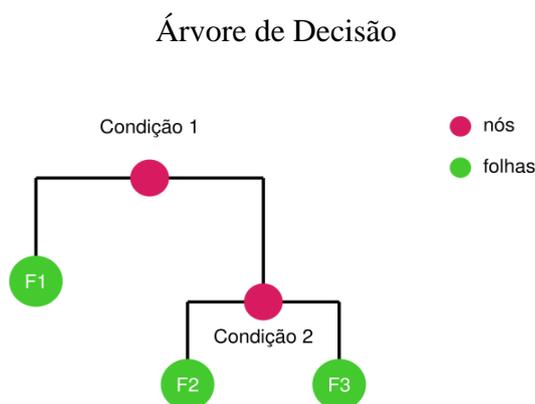
¹ Disponível em: www.gov.br/inep/pt-br/microdados/enem

2.2. Metodologia

Para o desenvolvimento deste trabalho utilizou-se a técnica de *Machine Learning*, também chamada de Aprendizado de Máquina que surgiu nos anos 1960 como um sub ramo da inteligência artificial. Ela tem a finalidade de extrair informações de dados brutos e representá-los por meio de um modelo matemático. De acordo com IZBICKI E SANTOS (2020) esta técnica era restrita à área computacional, mas nos anos 1990 aliou-se à estatística e passou a expandir seus horizontes, construindo ideias e métodos.

Na técnica de *Machine Learning* há um algoritmo chamado de *Árvore de Decisão*

que consiste em classificar os dados por meio de regras arbitrárias. É um modelo simples, com apresentação visual de fácil entendimento, lembrando um fluxograma. Segundo TACONELI (2008) uma *Árvore de Decisão* é chamada de *Árvore de Classificação* se a variável resposta for categórica, ou *Árvore de Regressão*, se numérica e ainda, uma *árvore de Classificação* ou *Regressão* é composta por nós, ramos e folhas. Cada nó representa um ponto de decisão que divide os dados de acordo com um atributo e os ramos que partem de cada nó conduzem a novos nós ou folhas, representando os caminhos de decisão. Observe a Figura 1:



Fonte: Izbicki e Santos 2020

Figura 1- Exemplo de estrutura de uma *Árvore de Decisão*

As *Árvores de Decisão* são construídas usando algoritmos que dividem repetidamente os dados em subconjuntos menores, com base em características específicas. Este processo na linguagem de programação é chamado de *particionamento recursivo*. O objetivo é criar uma estrutura que maximize a separação entre as diferentes classes ou resultados possíveis. Para construir as árvores basicamente precisa-se de 4 etapas: 1ª) a definição e execução de um critério de partição das amostras; 2ª) a aplicação do processo da poda; 3ª) a seleção do melhor modelo e 4ª) a classificação dos nós finais.

É muito importante lembrar que o valor da resposta Y na *Árvore de Decisão* é diferente da previsão na *Árvore de Regressão*. Nas árvores de *Classificação*, segundo IZBICKI E SANTOS (2020, p.156) a previsão para a resposta Y de uma observação com covariáveis x que estão em uma região R_k é dada por:

$$g(x) = \text{moda} \{y_i: x_i \in R_k\}$$

Ou seja, para prever o valor da resposta Y, observa-se a região a qual a observação X pertence e, então, calcula-se a moda dos valores da variável resposta das amostras do conjunto de treinamento pertencentes aquela mesma região.

Já a predição para a resposta Y das Árvores de Regressão é calculada pela média dos valores das variáveis resposta da amostra do conjunto de treinamento. Ou seja:

$$g(x) = \frac{1}{|\{i: x_i \in R_k\}|} \sum_{i: x_i \in R_k} y_i$$

Vale destacar que para a construção de uma *árvore* é necessário selecionar os atributos que irão separar os dados em cada *nó*, até a classificação total dos dados.

Para validar a precisão do modelo de classificação utilizou-se a análise das medidas da Sensibilidade, da Especificidade e da Acurácia que têm por base elementos estatísticos da teoria da decisão. A comparação entre as previsões do modelo e a real classificação será feita por um método que consiste em separar 80% das amostras para treinar o modelo e 20% para amostras teste. Com base na matriz de confusão (Tabela 1) LOUZADA (2009) definiu a medida Sensibilidade (S_e) como a razão:

$$S_e = \frac{VP}{n}$$

A Especificidade (E_s) por:

$$E_s = \frac{VN}{m}$$

Onde n é o total dos positivos e m é o total dos negativos.

Os estimadores da Sensibilidade e Especificidade seguem uma Distribuição Binomial com erros padrões de $S_e(1-S_e)/n$ e $E_s(1-E_s)/m$ respectivamente.

E a Acurácia como:

$$Acurácia = \frac{VP + VN}{VP + FN + VN + FP}$$

Medidas utilizadas em um estudo de validação de Modelos de Classificação que produzem respostas Dicotomizadas

Resultado do modelo de classificação	Real	
	positivo	negativo
positivo	<i>verdadeiro positivo (VP)</i>	<i>falso positivo (FP)</i>
negativo	<i>falso negativo (FN)</i>	<i>verdadeiro negativo (VN)</i>

Fonte: Louzada (2009)

Tabela1 - Medidas utilizadas em um estudo de validação de Modelos de Classificação que produzem respostas Dicotomizadas

É importante ressaltar que quanto maior a Acurácia mais preciso é o modelo. Em relação à sensibilidade e especificidade o teste foi em relação ao maior valor.

Capítulo 3

Resultados e Discussões

3.1 Estatística Descritiva

3.1.1 Variáveis Socioeconômicas – ENEM 2022

A análise dos dados desta pesquisa iniciou-se pela estatística descritiva das variáveis socioeconômicas contidas no questionário dos estudantes e respondidas no momento da inscrição do ENEM. Segundo BRUNI (2007) a estatística descritiva:

Tem como principal função resumir dados e informações investigadas, expondo-os da maneira mais prática e simples possível. Em muitos casos, há um grande número de dados e informações coletadas que poderiam atrapalhar o desenvolvimento das pesquisas e suas conclusões. Para simplificar, os dados são resumidos sob a forma de estatísticas. Para os pesquisadores descreverem os resultados de suas pesquisas de campo, eles criam tabelas e gráficos, dispondo seus dados de forma que possam ser interpretados mais facilmente. Bruni (2007, p.3).

A tabela abaixo contém a porcentagem das questões socioeconômicas e suas respectivas alternativas assinaladas:

Variáveis Socioeconômicas - ENEM 2022

	A	B	C	D	E	F	G	H
Q001	3.63%	16.41%	12.94%	11.24%	30.18%	9.63%	6.76%	9.21%
Q002	2.27%	11.49%	10.67%	11.74%	36.29%	12.90%	11.76%	2.88%
Q003	17.19%	19.73%	23.50%	20.85%	7.06%	11.67%		
Q004	13.90%	44.35%	6.35%	25.58%	5.09%	8.57%		
Q007	91.89%	4.55%	1.00%	2.56%				
Q008	0.63%	63.67%	24.71%	7.28%	3.71%			
Q009	0.63%	10.75%	40.89%	33.27%	5.66%			
Q010	50.19%	39.39%	9.10%	1.08%	0.23%			
Q011	75.28%	21.93%	2.50%	0.24%	0.04%			
Q012	1.28%	93.23%	5.07%	0.36%	0.07%			
Q013	51.63%	44.58%	3.29%	0.38%	0.09%			
Q014	35.13%	63.68%	1.14%	0.04%	0.01%			
Q015	86.45%	13.38%	0.15%	0.01%				
Q016	40.02%	50.87%	0.68%	0.03%	0.01%			
Q017	96.71%	3.22%	0.05%	0.01%				
Q018	74.54%	25.46%						
Q019	5.66%	64.29%	19.66%	7.13%	3.26%			
Q020	82.79%	17.21%						
Q021	77.29%	22.71%						
Q022	2.11%	15.49%	26.66%	31.53%	24.22%			
Q023	87.01%	12.99%						
Q024	41.11%	39.69%	10.41%	3.79%	1.81%			
Q025	8.72%	91.28%						

Tabela 2 - Referente as 25 questões socioeconômicas do ENEM 2022 gerada pela linguagem de programação R

Ao somar as alternativas A, B, C e D da Q001, observou-se que 44,22% dos pais ou os homens responsáveis pelos estudantes não completaram o Ensino Fundamental, 30,18% completaram o Ensino Médio e somente 16,39% completaram a faculdade. Já na Q002 é possível verificar que 36,17% das mães ou as mulheres responsáveis pelos estudantes não completaram o Ensino Fundamental, 36,29% completaram o Ensino Médio e somente 24,66% completaram a faculdade.

As questões Q003 e Q004 referem-se as ocupações dos pais ou responsáveis pelos estudantes. Em se tratando dos pais observou-se que 44,35% pertencem aos grupos C e D. Já as ocupações das mães com maior porcentagem estão nos grupos A e B com 54,42%.

Notou-se que as residências dos estudantes não possuem empregado(a) doméstico(a) (Q007 - 91%), têm apenas um banheiro (Q008 – 63,6%) e de dois a três quartos (Q009 – 82,9%).

Pelo questionário na Q010 observou-se que 50,19% das residências não possuem carro e 39,3% possuem apenas um. Já em relação à motocicleta na Q011, 75,28% não possuem e

21,93% possuem uma.

As questões Q012 a Q020 referem-se aos aparelhos domésticos encontrados nas residências. Assim verificou-se que em 93,2% das residências há uma geladeira, 51,6% não possuem freezer, 63,6% há uma máquina de lavar roupa (o tanquinho não serve), 86,4% não possuem máquina de secar roupa, 50,8% têm um forno micro-ondas, 96,7% não têm máquina de lavar louça, 74,5% não possuem aspirador de pó, 64,2% têm uma televisão em cores e 82,7% não têm aparelho de DVD.

A questão Q021 mostrou que 77,2% dos estudantes não possuem TV por assinatura. Na Q022 somente 2,1% das residências não têm telefone celular e 87% não têm telefone fixo (Q023). Em relação aos computadores, 44,2% das residências não possuem e 39,6% possuem apenas um computador (Q024). E por fim na questão Q025, 91,2% têm acesso à internet.

Após a análise da tabela apresenta-se agora as questões Q005 e Q006.

A Q005 está interessada em saber a quantidade de pessoas que moram na residência (incluindo você). Para responder esta questão basta analisar o gráfico abaixo e concluir que 27,3% moram três pessoas e 33,8% moram quatro pessoas.

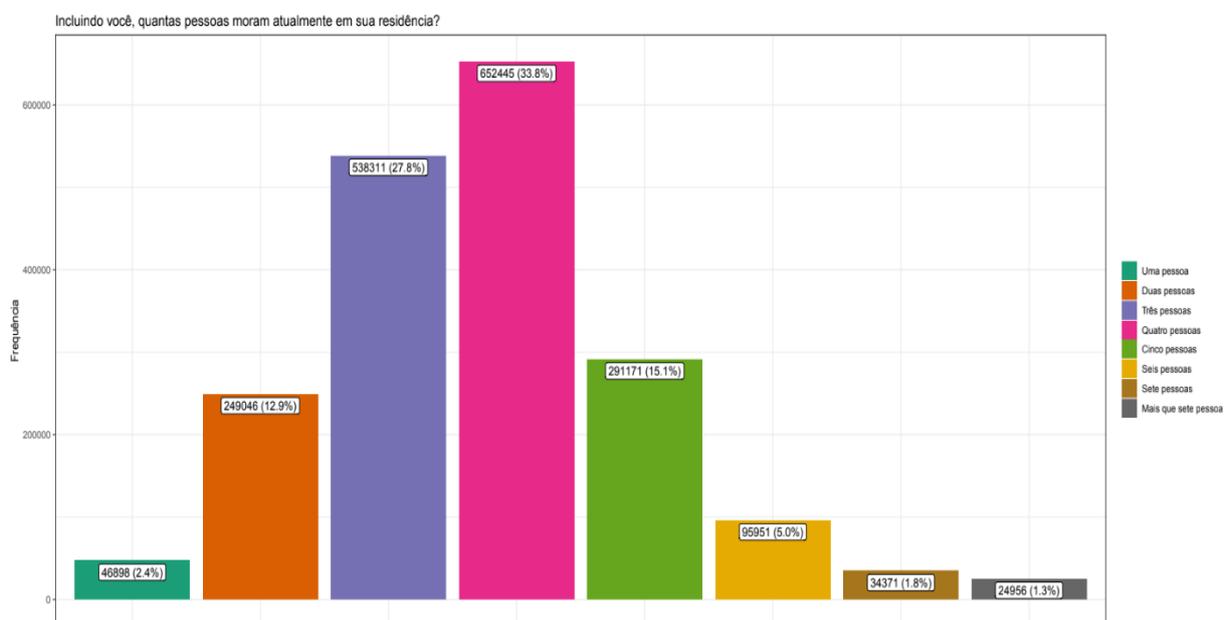


Figura 2 – Gráfico de Barras referente a questão socioeconômica Q005 do ENEM 2022 gerada pela linguagem de programação R

A Q006 trata da renda mensal familiar (somando a renda do estudante com a dos seus

familiares). Verificou-se que 28,7% dos estudantes têm renda de até R\$ 1.212,00; 16,8% têm renda de R\$ 1.212,01 até R\$ 1.818,00 e 12,2% têm renda de R\$ 1.818,01 até R\$ 2.424,00.

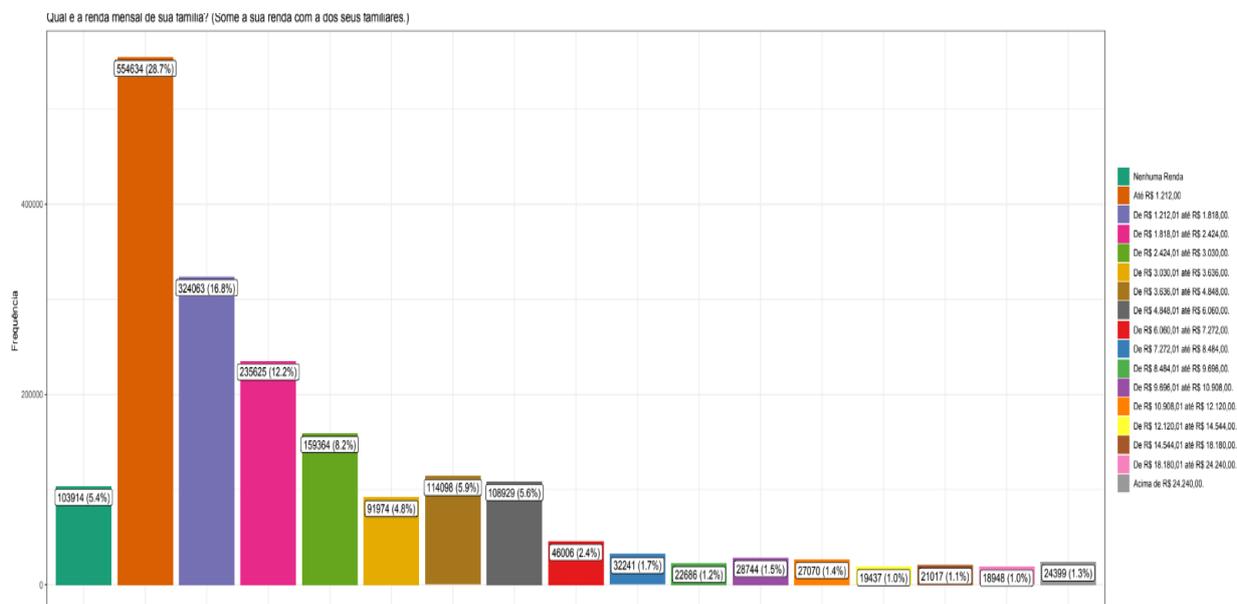


Figura 3 – Gráfico de Barras referente a questão socioeconômica Q006 do ENEM 2022 gerada pela linguagem de programação R

3.1.2 Variável Nota de Ciências da Natureza e suas tecnologias - ENEM 2022

Ao observar o gráfico *BoxPlot* e as medidas descritivas da média (Figura 4) percebeu-se que as notas dos estudantes do ENEM 2022, na prova de *Ciências da Natureza e suas tecnologias*, teve o mínimo de 313,6 pontos e o máximo de 875,3, média de 497,4 e mediana de 486,6. Os dados são assimétricos e a dispersão ficou em torno de 104,9 pontos. Constatou-se, ainda, que 25% dos dados são menores ou iguais a 440,75, 50% são menores ou iguais a 486,6 e 75% são menores ou iguais a 545,6 pontos.

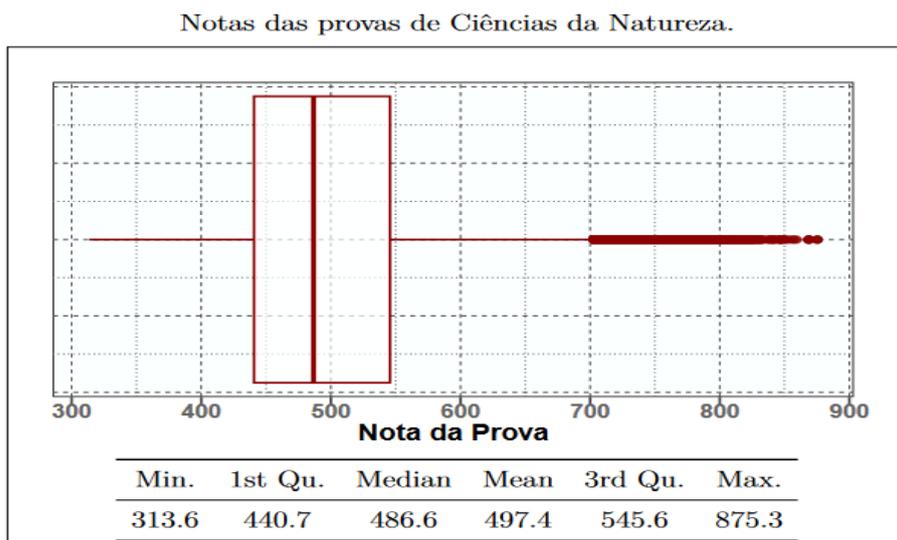


Figura 4 – Boxplot referente às notas de Ciências da Natureza e suas tecnologias do ENEM 2022 gerada pela linguagem de programação R

E por fim, ao cruzar as informações nota da prova de *Ciências da Natureza e suas tecnologias* com o grau de escolaridade do pai ou responsável, percebeu-se no gráfico de Pirâmide abaixo que o estudante cujo pai ou responsável completou a faculdade ou completou a pós-graduação tem as notas superiores à média de 497,42.

Nível de estudo do pai ou responsável versus nota categorizada

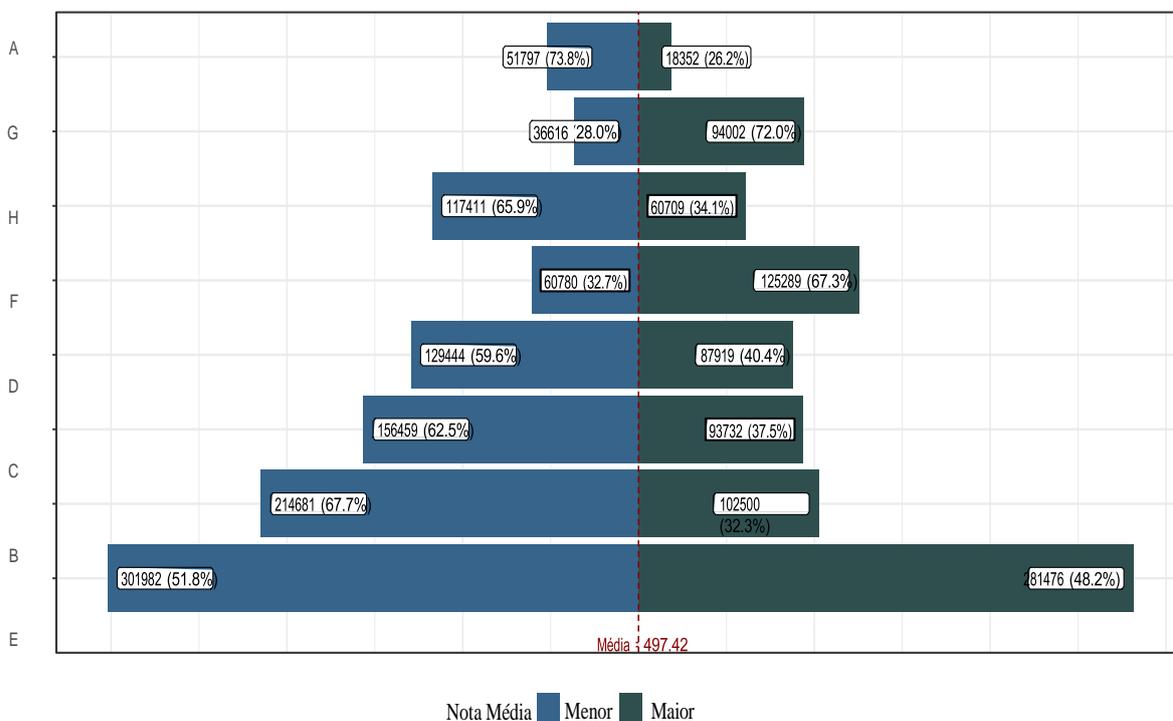


Figura 5– Gráfico de Pirâmide referentes às notas de Ciências da Natureza e suas tecnologias do ENEM 2022 e a questão socioeconômica Q001 gerada pela linguagem de programação R

3.2 Modelo de *Machine Learning*

3.2.1 Modelo e Validação

Na elaboração do Modelo considerou-se a pontuação dos estudantes na disciplina de Ciências da Natureza do ENEM 2022. As notas foram categorizadas como abaixo ou acima da pontuação média geral, servindo como variável resposta. Além disso, 25 questões socioeconômicas foram utilizadas como covariáveis. Dentre as 25 covariáveis, a questão Q006 foi a mais significativa, ocupando a posição de nó raiz na Árvore de Classificação, seguida das questões Q003 e Q004. A importância das variáveis pode ser observada na Figura 6.

Importância do Classificador para as covariáveis socioeconômicas

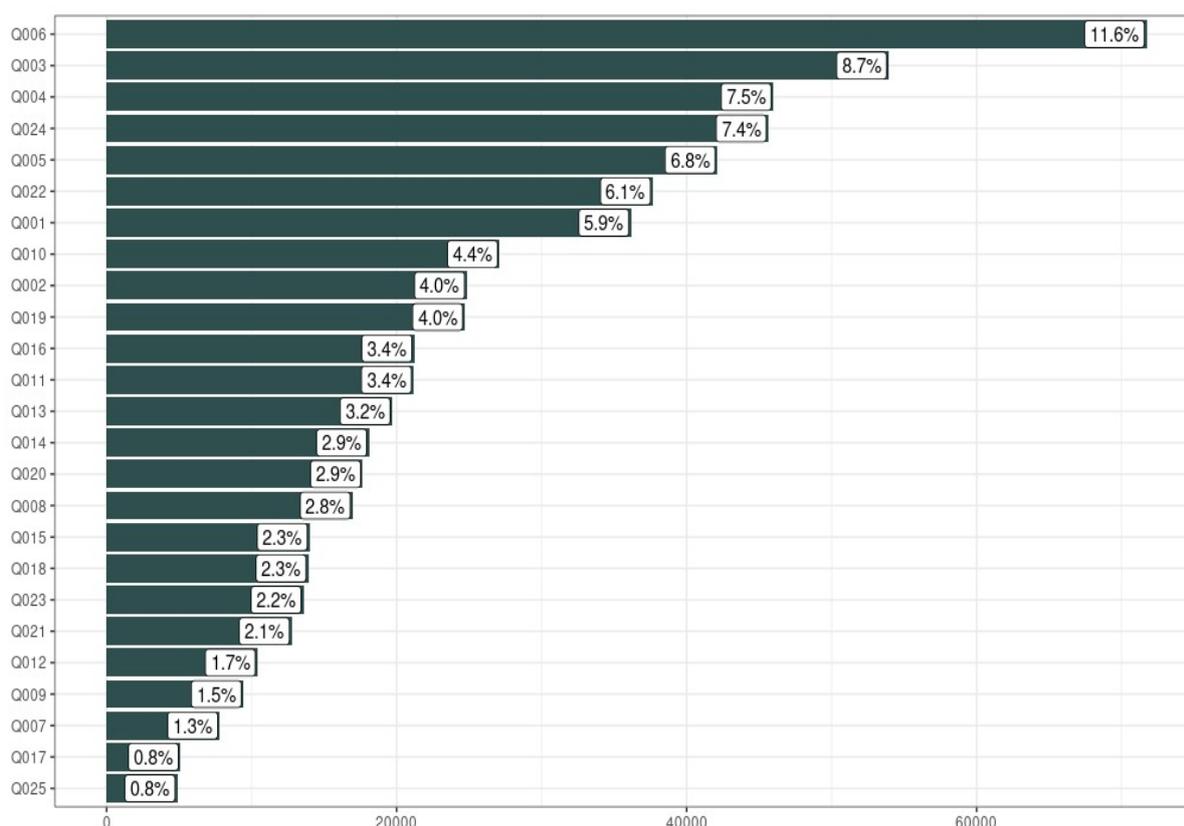


Figura 6 – Importância do Classificador para as covariáveis socioeconômicas gerada pela linguagem de programação R

Na construção do modelo levou-se em consideração as seguintes regras:

Classificação dos Ramos do Modelo

Regras	Cover
1. Quando Q006 for A ou B ou C ou D ou E e Q024 Não	39%
2. Quando Q006 for A ou B ou C ou D ou E e Q024 Sim e Q003 = G1, G2, G3 e não sei	22%
3. Quando Q006 for A ou B ou C ou D ou E e Q024 Sim e Q003 = outros grupos	7%
4. Quando Q006 for F ou G ou H ou I +	32%

Fonte: O autor

Tabela 3 – Classificação dos Ramos do Modelo

A Figura 7 contém o modelo – Árvore de Classificação

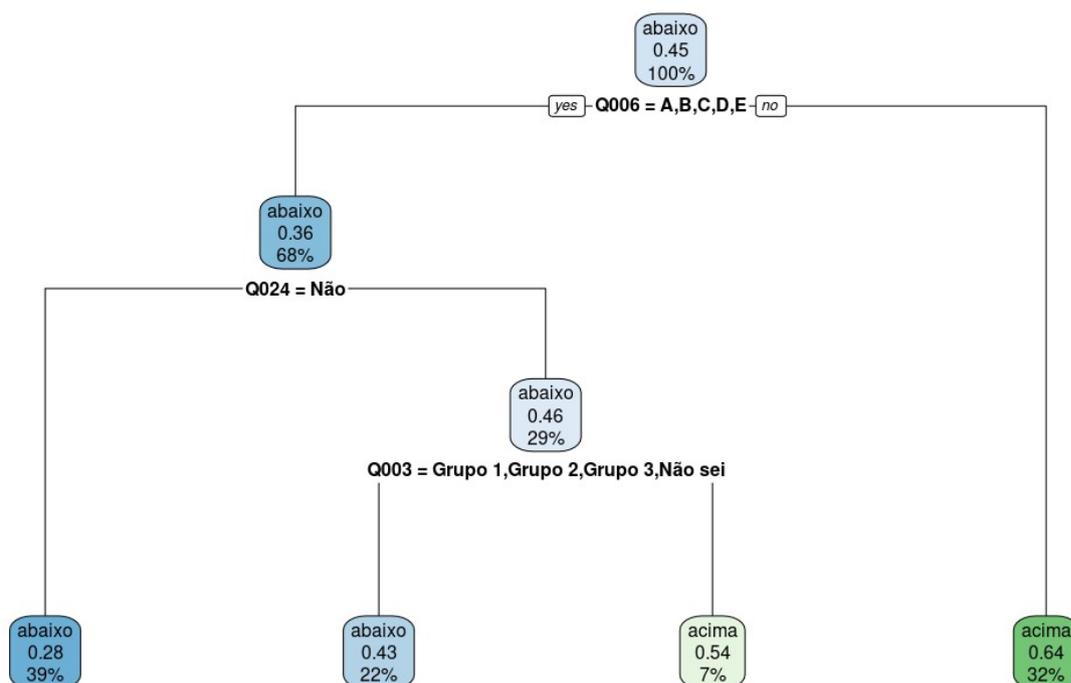


Figura 7 – Modelo: Árvore de Classificação das notas dos estudantes do ENEM 2022 na prova de Ciências da Natureza de acordo com as variáveis socioeconômicas gerada pela linguagem de programação R

Métricas para validar o modelo

Acurácia	Sensibilidade	Especificidade
0,6497	0,6655	0,6251

Fonte: O autor

Tabela 4 – Métricas para validar o Modelo

Concluiu-se que 68% os estudantes que realizaram a prova de Ciências da Natureza e suas tecnologias ENEM – 2022 com renda mensal familiar inferior ou igual a R\$3.030,00, têm média abaixo de 497,42 pontos. Entre esses estudantes:

- a) 39% não possuem computador e a profissão do pai ou responsável não está nos grupos 1, 2, 3 ou não sei;
- b) 22% possuem computador e a profissão do pai ou responsável está em outros grupos 1, 2, 3 ou não sei;
- c) 7% possuem computador e a profissão do pai ou responsável está em outros grupos;.

Além disso, 32% dos estudantes que possuem renda mensal familiar superior a R\$ 3.030,00 têm média acima de 497,42 pontos.

Dadas as condições do estudo e os valores mencionados Tabela 4 , o modelo foi considerado razoável.

Conclusão

Esta pesquisa mostrou por meio da Estatística Descritiva o perfil socioeconômico dos 1.933.149 estudantes que fizeram a prova de *Ciências da Natureza e suas tecnologias* do ENEM de 2022. Resumidamente este perfil apontou que:

- A porcentagem das mães ou responsáveis que concluíram Ensino Médio ou a Faculdade é superior aos pais ou responsáveis dos alunos;
- Em se tratando das ocupações dos pais observou-se que 44,35% pertencem aos grupos C e D, já as ocupações das mães com maior porcentagem estão nos grupos A e B com 54,42%;
- A maioria dos estudantes 63,1% têm renda mensal familiar inferior ou igual a R\$ 2.424,00 e somente 12,6% têm renda superior ou igual a R\$ 6.060,00;
- Na residência da maioria dos estudantes moram de 3 a 4 pessoas, tem apenas um banheiro, tem de dois a três quartos, não tem carro, não tem motocicleta, não tem empregado (a) doméstico (a), não possuem TV por assinatura, não tem telefone fixo, mas tem telefone celular e acesso à internet e 44% não tem computador;
- Em relação aos aparelhos domésticos a maioria possui: uma geladeira, uma máquina de lavar, um forno micro-ondas e uma televisão em cores. Não possuem: freezer, máquina de secar roupa, máquina de lavar roupa, aspirador de pó e DVD;
- Os estudantes cujos pais ou responsáveis completaram a faculdade ou completaram a pós-graduação têm as notas superiores à média de 497,42 pontos;
- Notou-se também que as notas dos estudantes na prova *Ciências da Natureza e suas tecnologias* teve mínimo de 313,6 pontos e máximo de 875,3, média de 497,4 e mediana de 486,6. Os dados são assimétricos e a dispersão ficou em torno de 104,9 pontos. Observou-se também que 25% dos dados são menores ou iguais a 440,75, 50% são menores ou iguais a 486,6 e 75% são menores ou iguais a 545,6 pontos.

A pesquisa também conseguiu fazer os seguintes apontamentos através do modelo Árvore de Decisão via *Machine Learning*:

- 68% dos estudantes com renda mensal familiar inferior ou igual a R\$ 3.030,00 têm médias abaixo de 497,42 pontos na prova de *Ciências da Natureza e suas tecnologias* do ENEM 2022;
- 32% dos estudantes com renda mensal familiar superior a R\$ 3.030,00 têm médias acima de 497,42 pontos na prova de *Ciências da Natureza e suas tecnologias* do ENEM 2022;

Concluiu-se assim, que a renda mensal familiar e a posse de computador, junto com a profissão do pai ou responsável, influenciaram significativamente a média alcançada pelos estudantes na prova de *Ciências da Natureza e suas tecnologias* do ENEM 2022.

Por fim, espera-se que esta pesquisa tenha contribuído para o desenvolvimento de um modelo estatístico que possa ser aplicado nas demais áreas do conhecimento da prova do ENEM e ainda, contribuir para entender melhor os milhares de dados produzidos pelo ENEM.

Referências

ARAÚJO, Robson F. R. de; GUIMARÃES, Ulisses V.; WARTHA, Edson J. Aspectos representacionais em questões do enem na área de ciências da natureza. ENCITEC. Santo Ângelo - Vol. 11, n. 1., p. 32-49, jan./abr. 2021.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. Portaria MEC nº 438, de 28 de maio de 1998. Institui o Exame Nacional do Ensino Médio - ENEM. Disponível em: http://www.crmariocovas.sp.gov.br/pdf/diretrizes_p0178-0181_c.pdf . Acesso em: 05 mar. 2022.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. Portaria MEC nº 109, de 27 de maio de 2009. Estabelece a sistemática para a realização do Exame Nacional do Ensino Médio no exercício de 2009. Disponível em: http://download.inep.gov.br/educacao_basica/enem/legislacao/2009/portaria_enem_2009_1.pdf . Acesso em: 05 mar. 2022.

BRUNI, Adriano L. Estatística Aplicada à Gestão Empresarial. São Paulo: Atlas, 2007.

ENEM 2014 tem 8,7 milhões de participantes e supera recorde do ano passado. Disponível em: <https://vestibular.brasescola.uol.com.br/enem/enem-2014-tem-87-milhoes-participantes-supera-recorde-ano-passado/327506.html> . Acessado em: 09 out. 2023.

GARCIA, Paulo; FRANZOLIN, Fernanda. A organização das questões com o foco na Biodiversidade no Exame Nacional do Ensino Médio: 10 anos de análises. Research, Society and Development. 10. e537101220925. 10.33448/rsd-v10i12.20925. 2021.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Microdados do Enem 2021. Brasília: Inep, 2022. Disponível em: < <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem> > . Acesso em: 25 mai. 2022.

INEP divulga resultados do enem 2005 por escola e município. Brasília: Inep, 2010. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/enem/inep-divulga-resultados-do-enem-2005-por-escola-e-municipio> . Acessado em: 09 out. 2023.

INEP: Nasce o maior exame brasileiro. Brasília: Inep, 2020. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enem/historico> . Acessado em: 18 out. 2023.

IZBICKI, Rafael; SANTOS, Tiago Mendonça dos. Aprendizado de máquina: uma abordagem estatística. São Carlos, SP: 2020.

LIMA, Danilo G.; LAGES, André M. G; SILVA, Rômulo P.; SANTOS, Jonathan de F.; Revisão introdutória dos indicadores de desigualdade e uma aplicação simplificada ao Brasil entre os anos de 1995 a 2014. RRCF, Fortaleza, v.10, n. 2, Jul./Dez. 2019.

LOUZADA, F.; AMARAL, G. J. A.; GUIRADO, L.; SILVA, P. H. F.; ABREU, H. J.; FERREIRA, M. R. P. Medidas estatísticas da capacidade preditiva de modelos de classificação em credit scoring. P@rtes (São Paulo), v. 68, p.7-28, 2009.

MARIANO, Diego. Métricas de avaliação em Machine Learning. 2021. Disponível: <https://bioinfo.com.br/metricas-de-avaliacao-em-machine-learning-acuracia-sensibilidade-precisao-especificidade-e-f-score/>. Acesso em: 28/02/2024.

MARIO FILHO. O que é Acurácia em Machine Learning? 2023. Disponível: <https://mariofilho.com/o-que-e-acuracia-em-machine-learning/>. Acesso em 27/02/2024.

MONICO, João F. G.; DAL PÓZ, Aluir P.; GALO, Maurício; SANTOS, Marcelo C. dos; OLIVEIRA, Leonardo, C.; Acurácia e Precisão: revendo os conceitos de forma Acurada. Bol. Ciênc. Geod., sec. Comunicações, Curitiba, v. 15, n o 3, p.469-483, jul-set, 2009.

NAVARRO, D.; IANELLO, M.; MUNERATTO, F.; WATANABE, G. Impactos do Conhecimento das Ciências Naturais para o Desempenho no ENEM: Considerações sobre a Desigualdade Científico-tecnológica para a Justiça Social. Revista Brasileira de Pesquisa em Educação em Ciências, [S. l.], p. e26002, 1-, 2021. DOI: 10.28976/1984-2686rbpec2021u12171246. Disponível em: <https://periodicos.ufmg.br/index.php/rbpec/article/view/26002>. Acesso em: 4 jun. 2022.

PIRES, A. Renda familiar e escolaridade dos pais: reflexões a partir dos microdados do ENEM 2012 do Estado de São Paulo. *Educação Temática Digital*, 17(3), 2015. P.535–541. <https://doi.org/10.20396/etd.v17i3.8638262>

SANTOS, Adailton Ferreira dos; OLIOSI, Elisa Cristina. A importância do ensino de ciências da natureza integrado à história da ciência e à filosofia da ciência: uma abordagem contextual. *Revista da FAAEBA: Educação e Contemporaneidade* [online]. 2013, vol.22, n.39, pp.195-204. ISSN 0104-7043.

TACONELI, C. A. Árvores de classificação multivariadas fundamentadas em coeficientes de dissimilaridade e entropia. São Paulo: Universidade de São Paulo, 2008.

Anexos

Variáveis Socioeconômicas

Sigla da variável	Descrição	Categoria-Descrição	
Q001	Até que série seu pai, ou homem responsável por você, estudou?	A	Nunca estudou.
		B	Não completou a 4ª série/5º ano do Ensino Fundamental.
		C	Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano do Ensino Fundamental.
		D	Completou a 8ª série/9º ano do Ensino Fundamental, mas não completou o Ensino Médio.
		E	Completou o Ensino Médio, mas não completou a Faculdade.
		F	Completou a Faculdade, mas não completou a Pós-graduação.
		G	Completou a Pós-graduação.
		H	Não sei.
Q002	Até que série sua mãe, ou a mulher responsável por você, estudou?	A	Nunca estudou.
		B	Não completou a 4ª série/5º ano do Ensino Fundamental.
		C	Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano do Ensino Fundamental.
		D	Completou a 8ª série/9º ano do Ensino Fundamental, mas não completou o Ensino Médio.
		E	Completou o Ensino Médio, mas não completou a Faculdade.
		F	Completou a Faculdade, mas não completou a Pós-graduação.
		G	Completou a Pós-graduação.
		H	Não sei.
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).	A	Grupo 1: Lavrador, agricultor sem empregados, bóia fria, criador de animais (gado, porcos, galinhas, ovelhas, cavalos etc.), apicultor, pescador, lenhador, seringueiro, extrativista.
		B	Grupo 2: Diarista, empregado doméstico, cuidador de idosos, babá, cozinheiro (em casas particulares), motorista particular, jardineiro, faxineiro de empresas e prédios, vigilante, porteiro, carteiro, office-boy, vendedor, caixa, atendente de loja, auxiliar administrativo, recepcionista, servente de pedreiro, repositor de mercadoria.
		C	Grupo 3: Padeiro, cozinheiro industrial ou em restaurantes, sapateiro, costureiro, joalheiro, torneiro mecânico, operador de máquinas, soldador, operário de fábrica, trabalhador da mineração, pedreiro, pintor, eletricista, encanador, motorista, caminhoneiro, taxista.
		D	Grupo 4: Professor (de ensino fundamental ou médio, idioma, música, artes etc.), técnico (de enfermagem, contabilidade, eletrônica etc.), policial, militar de baixa patente (soldado, cabo, sargento), corretor de imóveis, supervisor, gerente, mestre de obras, pastor, microempresário (proprietário de empresa com menos de 10 empregados), pequeno comerciante, pequeno proprietário de terras, trabalhador autônomo ou por conta própria.
		E	Grupo 5: Médico, engenheiro, dentista, psicólogo, economista, advogado, juiz, promotor, defensor, delegado, tenente, capitão, coronel, professor universitário, diretor em empresas públicas ou privadas, político, proprietário de empresas com mais de 10 empregados.
		F	Não sei.

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).	A	Grupo 1: Lavradora, agricultora sem empregados, bóia fria, criadora de animais (gado, porcos, galinhas, ovelhas, cavalos etc.), apicultora, pescadora, lenhadora, seringueira, extrativista.																																									
		B	Grupo 2: Diarista, empregada doméstica, cuidadora de idosos, babá, cozinheira (em casas particulares), motorista particular, jardineira, faxineira de empresas e prédios, vigilante, porteira, carteira, office-boy, vendedora, caixa, atendente de loja, auxiliar administrativa, recepcionista, servente de pedreiro, repositora de mercadoria.																																									
		C	Grupo 3: Padeira, cozinheira industrial ou em restaurantes, sapateira, costureira, joalheira, torneira mecânica, operadora de máquinas, soldadora, operária de fábrica, trabalhadora da mineração, pedreira, pintora, eletrícista, encanadora, motorista, caminhoneira, taxista.																																									
		D	Grupo 4: Professora (de ensino fundamental ou médio, idioma, música, artes etc.), técnica (de enfermagem, contabilidade, eletrônica etc.), policial, militar de baixa patente (soldado, cabo, sargento), corretora de imóveis, supervisora, gerente, mestre de obras, pastora, microempresária (proprietária de empresa com menos de 10 empregados), pequena comerciante, pequena proprietária de terras, trabalhadora autônoma ou por conta própria.																																									
		E	Grupo 5: Médica, engenheira, dentista, psicóloga, economista, advogada, juíza, promotora, defensora, delegada, tenente, capitã, coronel, professora universitária, diretora em empresas públicas ou privadas, política, proprietária de empresas com mais de 10 empregados.																																									
		F	Não sei.																																									
Q005	Incluindo você, quantas pessoas moram atualmente em sua residência?	1 a 20	<table border="1"> <tr><td>1</td><td>1, pois moro sozinho(a)</td></tr> <tr><td>2</td><td>2</td></tr> <tr><td>3</td><td>3</td></tr> <tr><td>4</td><td>4</td></tr> <tr><td>5</td><td>5</td></tr> <tr><td>6</td><td>6</td></tr> <tr><td>7</td><td>7</td></tr> <tr><td>8</td><td>8</td></tr> <tr><td>9</td><td>9</td></tr> <tr><td>10</td><td>10</td></tr> <tr><td>11</td><td>11</td></tr> <tr><td>12</td><td>12</td></tr> <tr><td>13</td><td>13</td></tr> <tr><td>14</td><td>14</td></tr> <tr><td>15</td><td>15</td></tr> <tr><td>16</td><td>16</td></tr> <tr><td>17</td><td>17</td></tr> <tr><td>18</td><td>18</td></tr> <tr><td>19</td><td>19</td></tr> <tr><td>20</td><td>20</td></tr> </table>		1	1, pois moro sozinho(a)	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11	12	12	13	13	14	14	15	15	16	16	17	17	18	18	19	19	20	20
1	1, pois moro sozinho(a)																																											
2	2																																											
3	3																																											
4	4																																											
5	5																																											
6	6																																											
7	7																																											
8	8																																											
9	9																																											
10	10																																											
11	11																																											
12	12																																											
13	13																																											
14	14																																											
15	15																																											
16	16																																											
17	17																																											
18	18																																											
19	19																																											
20	20																																											
Q006	Qual é a renda mensal de sua família? (Some a sua renda com a dos seus familiares).	<table border="1"> <tr> <td>A</td> <td>Nenhuma Renda</td> <td>K</td> <td>De R\$ 8.484,01 até R\$ 9.696,00.</td> </tr> <tr> <td>B</td> <td>Até R\$ 1.212,00</td> <td>L</td> <td>De R\$ 9.696,01 até R\$ 10.908,00.</td> </tr> <tr> <td>C</td> <td>De R\$ 1.212,01 até R\$ 1.818,00.</td> <td>M</td> <td>De R\$ 10.908,01 até R\$ 12.120,00.</td> </tr> <tr> <td>D</td> <td>De R\$ 1.818,01 até R\$ 2.424,00.</td> <td>N</td> <td>De R\$ 12.120,01 até R\$ 14.544,00.</td> </tr> <tr> <td>E</td> <td>De R\$ 2.424,01 até R\$ 3.030,00.</td> <td>O</td> <td>De R\$ 14.544,01 até R\$ 18.180,00.</td> </tr> <tr> <td>F</td> <td>De R\$ 3.030,01 até R\$ 3.636,00.</td> <td>P</td> <td>De R\$ 18.180,01 até R\$ 24.240,00.</td> </tr> <tr> <td>G</td> <td>De R\$ 3.636,01 até R\$ 4.848,00.</td> <td>Q</td> <td>Acima de R\$ 24.240,00.</td> </tr> <tr> <td>H</td> <td>De R\$ 4.848,01 até R\$ 6.060,00.</td> <td></td> <td></td> </tr> <tr> <td>I</td> <td>De R\$ 6.060,01 até R\$ 7.272,00.</td> <td></td> <td></td> </tr> <tr> <td>J</td> <td>De R\$ 7.272,01 até R\$ 8.484,00.</td> <td></td> <td></td> </tr> </table>			A	Nenhuma Renda	K	De R\$ 8.484,01 até R\$ 9.696,00.	B	Até R\$ 1.212,00	L	De R\$ 9.696,01 até R\$ 10.908,00.	C	De R\$ 1.212,01 até R\$ 1.818,00.	M	De R\$ 10.908,01 até R\$ 12.120,00.	D	De R\$ 1.818,01 até R\$ 2.424,00.	N	De R\$ 12.120,01 até R\$ 14.544,00.	E	De R\$ 2.424,01 até R\$ 3.030,00.	O	De R\$ 14.544,01 até R\$ 18.180,00.	F	De R\$ 3.030,01 até R\$ 3.636,00.	P	De R\$ 18.180,01 até R\$ 24.240,00.	G	De R\$ 3.636,01 até R\$ 4.848,00.	Q	Acima de R\$ 24.240,00.	H	De R\$ 4.848,01 até R\$ 6.060,00.			I	De R\$ 6.060,01 até R\$ 7.272,00.			J	De R\$ 7.272,01 até R\$ 8.484,00.		
A	Nenhuma Renda	K	De R\$ 8.484,01 até R\$ 9.696,00.																																									
B	Até R\$ 1.212,00	L	De R\$ 9.696,01 até R\$ 10.908,00.																																									
C	De R\$ 1.212,01 até R\$ 1.818,00.	M	De R\$ 10.908,01 até R\$ 12.120,00.																																									
D	De R\$ 1.818,01 até R\$ 2.424,00.	N	De R\$ 12.120,01 até R\$ 14.544,00.																																									
E	De R\$ 2.424,01 até R\$ 3.030,00.	O	De R\$ 14.544,01 até R\$ 18.180,00.																																									
F	De R\$ 3.030,01 até R\$ 3.636,00.	P	De R\$ 18.180,01 até R\$ 24.240,00.																																									
G	De R\$ 3.636,01 até R\$ 4.848,00.	Q	Acima de R\$ 24.240,00.																																									
H	De R\$ 4.848,01 até R\$ 6.060,00.																																											
I	De R\$ 6.060,01 até R\$ 7.272,00.																																											
J	De R\$ 7.272,01 até R\$ 8.484,00.																																											
Q007	Em sua residência trabalha empregado(a) doméstico(a)?	A	Não.																																									
		B	Sim, um ou dois dias por semana.																																									

		C	Sim, três ou quatro dias por semana.
		D	Sim, pelo menos cinco dias por semana.
Q008	Na sua residência tem banheiro?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q009	Na sua residência tem quartos para dormir?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q010	Na sua residência tem carro?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q011	Na sua residência tem motocicleta?	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q012	Na sua residência tem geladeira?	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q013	Na sua residência tem freezer (independente ou segunda porta da geladeira)?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.

Q014	Na sua residência tem máquina de lavar roupa? (o tanquinho NÃO deve ser considerado)	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q015	Na sua residência tem máquina de secar roupa (independente ou em conjunto com a máquina de lavar roupa)?	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q016	Na sua residência tem forno micro-ondas?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q017	Na sua residência tem máquina de lavar louça?	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q018	Na sua residência tem aspirador de pó?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q019	Na sua residência tem televisão em cores?	A	Não.
		B	Sim, uma.
		C	Sim, duas.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q020	Na sua residência tem aparelho de DVD?	A	Não.
		B	Sim.

Q021	Na sua residência tem TV por assinatura?	A	Não.
		B	Sim.
Q022	Na sua residência tem telefone celular?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q023	Na sua residência tem telefone fixo?	A	Não.
		B	Sim.
Q024	Na sua residência tem computador?	A	Não.
		B	Sim, um.
		C	Sim, dois.
		D	Sim, três.
		E	Sim, quatro ou mais.
Q025	Na sua residência tem acesso à Internet?	A	Não.
		B	Sim.