



UEPB

**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

MÁRCIO EDGLAITON ROSA FARIAS

USO DE DATA SCIENCE NA ANÁLISE DAS PROVAS DO ENEM

**CAMPINA GRANDE
2021**

MÁRCIO EDGLAITON ROSA FARIAS

USO DE DATA SCIENCE NA ANÁLISE DAS PROVAS DO ENEM

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação Bacharelado em Ciência da Computação na Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharelado em Computação.

Área de concentração: Ciência de dados e Big Data.

Orientador: Prof. Dr. Vladimir Costa de Alencar.

**CAMPINA GRANDE
2021**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

F224u Farias, Márcio Edglaíton Rosa.
Uso de Data Science na análise das provas do ENEM
[manuscrito] / Márcio Edglaíton Rosa Farias. - 2022.
52 p.

Digitado.

Trabalho de Conclusão de Curso (Graduação em
Computação) - Universidade Estadual da Paraíba, Centro de
Ciências e Tecnologia, 2022.

"Orientação : Prof. Dr. Vladimir Costa de Alencar ,
Coordenação do Curso de Computação - CCT."

1. Ciência de Dados. 2. Educação. 3. Big Data. I. Título

21. ed. CDD 005.7

MÁRCIO EDGLAITON ROSA FARIAS

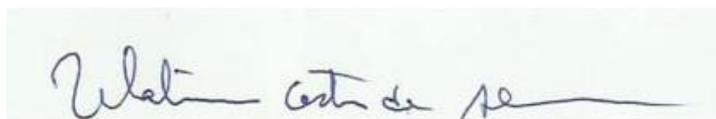
USO DE DATA SCIENCE NA ANÁLISE DAS PROVAS DO ENEM

Trabalho de Conclusão de Curso apresentado ao curso de Graduação em Bacharelado em Computação da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharelado em Computação.

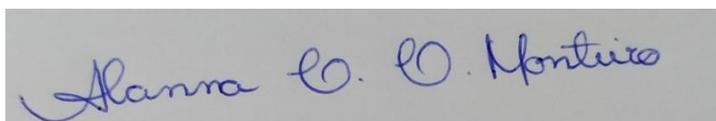
Área de concentração: Ciência de dados e Big Data.

Aprovada em: 29 de novembro de 2021.

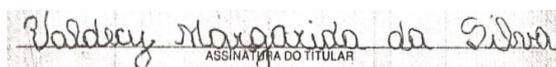
BANCA EXAMINADORA



Prof. Dr. Vladimir Costa de Alencar (Orientador)
Universidade Estadual da Paraíba (UEPB)



Prof. Me. Alanna Camylla Coelho Monteiro
Universidade Estadual da Paraíba (UEPB)



ASSINATURA DO TITULAR

Profa. Dra. Valdecy Margarida da Silva
Universidade Estadual da Paraíba (UEPB)

À minha avó Francisca Maria Rosa (*in
memoriam*) por todo apoio e incentivo,
DEDICO.

AGRADECIMENTOS

Agradeço primeiramente a Deus, nosso pai e criador, por me conceder essa oportunidade de crescimento intelectual através da graduação, e por me dar as forças necessárias para não desistir diante das dificuldades encontradas ao longo desses anos.

Agradeço a minha família, em especial a meu pai, minha mãe e meus irmãos, pelo apoio e compreensão nas minhas ausências, por sempre me incentivarem e não medirem esforços para me ajudar durante toda minha graduação.

Aos meus professores e toda equipe da universidade, que me acolheram no início da minha vida acadêmica, em especial a meu orientador Vladimir, por sempre acreditar no meu potencial e pela paciência nos momentos de dificuldades.

Aos amigos que construí durante minha trajetória acadêmica, o pessoal da coordenação do curso, das cantinas e de outros cursos pelas palavras de incentivo e encorajamento.

Aos meus colegas de curso, não irei citar os nomes para não esquecer alguém, por sempre estarem ao meu lado e por compartilharem sempre comigo os momentos bons e também nas horas difíceis.

A todos que de uma forma ou de outra contribuíram para que este sonho fosse realizado, que Deus possa abençoá-los grandemente.

“Conhecimento não tem valor a menos que você o coloque em prática.”

(Anton Chekhov)

RESUMO

O Exame Nacional do Ensino Médio - ENEM tornou-se ao longo dos anos, a principal ferramenta utilizada pelo MEC para realizar a avaliação dos estudantes brasileiros no último ano do Ensino Médio. Ele tem como principal objetivo possibilitar o ingresso dos estudantes no Ensino Superior e também para avaliar o desempenho dos concluintes do Ensino Médio. O presente trabalho tem como proposta realizar uma análise nos dados do Enem 2019, a fim de comprovar a sua importância e características gerais. A metodologia utilizada foi uma pesquisa experimental de forma quantitativa. Foi utilizado um dataset com microdados do ENEM 2019, disponibilizados pelo Ministério da Educação no site do INEP. Foram realizadas análises utilizando Python e técnicas de Big Data analytics para tratar os dados e logo em seguida foi utilizado o Tableau para construção dos gráficos e mapas. Observou-se algumas características do perfil dos inscritos e também o seu desempenho no exame. Os dados foram visualizados utilizando as ferramentas Excel e o Tableau Public para facilitar na compreensão dos mesmos. A partir das análises realizadas tem-se que no ENEM 2019 a maioria dos estudantes inscritos é proveniente de escolas públicas o que justifica a quantidade de pedidos de isenção e que a região sudeste se destaca nas médias de Matemática e Ciências da Natureza já a região Norte apresenta as menores notas. Dentre os autores que se utilizou no embasamento deste trabalho destaca-se Carneiro (2012), Travitzki (2013), (CURI; MENEZES FILHO, 2013) e (FIGUEIRÊDO; NOGUEIRA; SANTANA, 2014) que contribuíram para a compreensão da importância do ENEM. Conclui-se que a cada ano o Exame cresce e busca adaptar-se à realidade dos estudantes.

Palavras-Chave: Ciência de dados. Educação. Big Data.

ABSTRACT

The National High School Exam -ENEM- became, over the Years, the main tool used by the MEC to evaluate the Brazilian students in the last year of the high school. Its main objective is to make possible the admission of students in higher education and evaluate the performance of the students of high school. The present study aims to analyze data from ENEM 2019 to assure its importance and general characteristics. The methodology used was an experimental research of quantitative nature. A dataset with microdata about ENEM 2019 made available by the Ministry of Education in the website of INEP was used. Python and Big Data analytics were used in the analysis of the data and Tableau was used in the construction of the charts and maps. Some characteristics regarding the profile and performance of the students were noticed and were displayed through Microsoft Excel and public Tableau so they could be better understood. The analysis shows that most of students who take the ENEM come from public schools, which explains the quantity of requests of registration tax exemption and that the Southeast region excels at Mathematics and Science while the North region has the lowest grades. Among the authors whose works were used as basis of this paper, the emphasis falls on Carneiro (2012), Travitzki (2013), (Cury; Menezes; Filho 2013) e (Figueirêdo; Nogueira; Santana, 2014), whose works contribute to the comprehension of how important ENEM is. In conclusion, year after year, the ENEM matures and adapts to the social reality of the students.

Keywords: Data Science. Education. Big Data.

LISTA DE ILUSTRAÇÕES

| | |
|---|----|
| Figura 1- Print da tela do ambiente virtual do Google Colab..... | 29 |
| Figura 2- Código para definir uma função em Python utilizando o Google Colab..... | 29 |
| Figura 3- Código utilizado para explorar as variáveis..... | 30 |
| Figura 4- Exemplo de tela do software Tableau utilizado na visualização de dados das médias do ENEM em todo território brasileiro..... | 32 |
| Figura 5- Comparativo do número de inscritos no ENEM entre os anos 2011-2019..... | 34 |
| Figura 6- Porcentagem dos participantes inscritos que solicitaram isenção da taxa de inscrição por região geográfica no Brasil no ano de 2019..... | 35 |
| Figura 7- Percentual dos candidatos que realizaram o ENEM 2019 em braile..... | 36 |
| Figura 8- Número de inscritos portadores de autismo que participaram do ENEM..... | 37 |
| Figura 9- Número de inscritos no ENEM segundo o grau de escolaridade da mãe ou mulher responsável..... | 38 |
| Figura 10- Infográfico sobre as características dos participantes do ENEM 2019..... | 39 |
| Figura 11- Porcentagem de participantes do ENEM 2019 por cor/raça..... | 39 |
| Figura 12- Percentual de alunos por tipo de escola..... | 40 |
| Figura 13- Representação do tipo de escola do inscrito no ENEM 2019..... | 40 |
| Figura 14- Representação da opção de língua estrangeira pelos candidatos do ENEM..... | 41 |
| Figura 15- Situação dos estudantes de ensino médio inscritos no ENEM 2019..... | 42 |
| Figura 16- Mapa que demonstra a média das idades dos participantes..... | 42 |
| Figura 17- Mapa do Brasil representando a nota média da redação por estado..... | 43 |
| Figura 18- Infográfico sobre a prova da redação do ENEM 2019..... | 44 |
| Figura 19- Mapa do Brasil representando a média da nota da prova de matemática por estado..... | 44 |
| Figura 20- Mapa do Brasil representando as médias das notas dos participantes em Ciências da Natureza..... | 45 |

LISTA DE QUADROS

| | |
|--|----|
| Quadro 1- Pastas e Arquivos dos Microdados do ENEM 2019..... | 26 |
| Quadro 2- Pasta de Dados e Arquivos dos Microdados do ENEM 2019. | 26 |
| Quadro 3- Pasta Dicionário e Arquivos dos Microdados do ENEM 2019. | 27 |
| Quadro 4- Pasta Leia-Me e Documentos Técnicos e Arquivos dos Microdados do ENEM 2019. | 27 |
| Quadro 5- Faixa etária dos candidatos inscritos no ENEM 2019..... | 38 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|--|
| ASCII | American Standart Code for Information Interchange. |
| CN | Ciências da Natureza |
| CONADE | Conselho Nacional dos Direitos da Pessoa com Deficiência. |
| ENADE | Exame Nacional de Desempenho de Estudantes. |
| ENEM | Exame Nacional do Ensino Médio. |
| FIES | Fundo de Financiamento Estudantil |
| INEP | Instituto Nacional de Estudos e Pesquisas Educacionais Anízio Teixeira |
| MEC | Ministério da Educação |
| PROUNI | Programa Universidade para Todos |
| REUNI | (Plano) de Reestruturação e Expansão das Universidades Federais |
| SAEB | Sistema de Avaliação da Educação Básica |
| SAS | Statistical Analysis System |
| SISU | Sistema de Seleção Unificada |
| SPSS | Statistical Product and Service Solutions |
| TRI | Teoria de Resposta ao Item |

SUMÁRIO

| | |
|---|----|
| 1. INTRODUÇÃO | 12 |
| 2. OBJETIVOS | 14 |
| 2.1 OBJETIVO GERAL | 14 |
| 2.2 OBJETIVOS ESPECÍFICOS..... | 14 |
| 2.3 JUSTIFICATIVA | 14 |
| 3. FUNDAMENTAÇÃO TEÓRICA | 15 |
| 3.1 O ENEM..... | 15 |
| 3.1.1. Estrutura do Exame Nacional do Ensino Médio - ENEM | 16 |
| 3.1.2. Trajetória do ENEM..... | 18 |
| 3.1.3. Características Gerais do ENEM | 20 |
| 3.2 DATA SCIENCE..... | 21 |
| 3.3 BIG DATA..... | 22 |
| 3.4 VISUALIZAÇÃO DE DADOS..... | 23 |
| 4. METODOLOGIA | 25 |
| 4.1 OS MICRODADOS ENEM 2019 | 25 |
| 4.1.1 A utilização de Microdados | 25 |
| 4.1.2 Descrição dos Microdados do ENEM 2019..... | 26 |
| 4.2 LINGUAGEM DE PROGRAMAÇÃO | 28 |
| 4.2.1 Google Colab e Python | 28 |
| 4.3 ESTATÍSTICA DESCRITIVA (ANÁLISE DE DADOS)..... | 32 |
| 5. RESULTADOS E DISCUSSÕES | 34 |
| 5.1 CARACTERÍSTICAS DOS PARTICIPANTES | 35 |
| 5.2 DESEMPENHO DOS CANDIDATOS DO ENEM 2019..... | 43 |
| 6. CONCLUSÃO | 47 |
| REFERÊNCIAS | 49 |

1. INTRODUÇÃO

É fato notório a existência de exames de vestibulares desde o século passado. Porém, a necessidade de selecionar e classificar candidatos para ingressarem no ensino superior chegou ao ápice no início da década de 70, com os vestibulares unificados para suprir a demanda nacional pelo ensino superior (CHIQUETTO; KRAPAS, 2012; MATTOS, 2012).

No Brasil, o exame seletivo que possibilita o ingresso dos estudantes do ensino médio nas universidades é o Exame Nacional do Ensino Médio (ENEM), que vem substituir os antigos vestibulares adotados/utilizados pelas universidades brasileiras e o Ministério da Educação (MEC) a partir do ano de 2009, depois de sua reformulação, que antes (1998-2008) era utilizado para avaliar o desempenho dos estudantes do ensino médio (INEP, 2019; INEP, 2020).

Para se ingressar na universidade brasileira é necessário que o estudante, no ENEM, realize o exame de redação e das quatro áreas do conhecimento (ciências da natureza e suas tecnologias; ciências humanas e suas tecnologias; linguagens, códigos e suas tecnologias e matemática e suas tecnologias) (INEP, 2009).

Para participar do ENEM o estudante concluinte do ensino médio deve informar através de um questionário a sua condição socioeconômica de caráter obrigatório no ato da inscrição, possibilitando equiparar entre diferentes sistemas de ensino, regiões e estratos sociais, contextualizando o desempenho dos sujeitos, que participam deste certame, em situações pessoais, familiares, trajetória escolar, condição econômica, experiência de trabalho, como também nas percepções sobre a escola (GREMAUD et al., 2007; INEP, 2009).

Para Lopes e Péres (2010), o ENEM além de servir como porta para entrada no ensino superior, serve como controle dos currículos da educação básica, e do desempenho da educação. Já Travitzki (2013) o ENEM surgiu com o propósito de servir de instrumento de autoavaliação de conhecimentos para os alunos, com o passar dos anos o exame foi modificando seus objetivos, se transformando num instrumento central na política educacional brasileira, com características variadas e inúmeras funções.

De acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE) a primeira edição (1998) do Exame reuniu 157,2 mil inscritos sendo que apenas 115,6 mil participantes realizaram o certame. O Instituto aponta que um grande aumento de participantes no ENEM ocorreu em 2004, quando o Programa Universidade para Todos (ProUni) foi criado pelo Ministério da Educação, que também passou a utilizar o desempenho dos estudantes na prova para distribuir bolsas em Instituições de Ensino Superior privadas (IBGE, 2020).

Para Hey (2009), especialista em Políticas Públicas de Educação, o ENEM não muda a realidade educacional, que é marcada pela desigualdade de oportunidades que a muito vem sendo praticada com os vestibulares tradicionais que provocavam uma exclusão significativa para aqueles que não tinham acesso a uma educação de qualidade.

Autores como Carneiro (2012) e Travitzki (2013), também relatam alguns efeitos do ENEM sobre o Ensino Médio. De acordo com o primeiro Autor, o Exame não contribui para identificar e resolver os problemas relacionados com as diferentes condições de “partida” dos indivíduos que a ele se submetem. Já o Segundo, a forma de divulgação das notas do Exame traz um “ranqueamento” entre as instituições de ensino, estimulando a competição e, por conseguinte, o sentimento de superioridade nas escolas já bem conceituadas - que, num círculo vicioso, continuam a ser bem avaliadas - e de frustração dos alunos nas escolas menos conceituadas que, em situação adversa, não conseguem equilibrar o quadro, mantendo-se em condição inferior (CARNEIRO,2012; TRAVITZKI, 2013).

Somente onze anos após sua criação, o ENEM passou a ser utilizado também como processo seletivo para entrada em universidades públicas, por meio do sistema de seleção unificada (SISU) e é também empregado para concessão de bolsas em universidades privadas, por meio do programa Universidade para Todos - PROUNI (SILVA & MELETTI, 2014).

Valendo salientar que, de acordo com Santos (2011), o exame possuía uma proposta interdisciplinar e contextualizada, tendo sido sua criação influenciada pelas reformas no sistema educacional. Já Travitzki (2013), comenta que o ENEM tinha inicialmente o objetivo de avaliar a capacidade de raciocínio dos alunos, sua habilidade para resolver problemas, suas competências gerais. O ENEM teve na sua fase inicial, o caráter de ser voluntário e individual, no qual dependia da disponibilidade dos estudantes, tendo como principal objetivo possibilitar sua autoavaliação. (BRASIL, 2008b).

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral deste trabalho foi realizar uma análise dos Microdados do ENEM 2019, utilizando técnicas de Ciências de Dados.

2.2 OBJETIVOS ESPECÍFICOS

- Compreender os Microdados do ENEM 2019 utilizando Data Science.
- Analisar as características socioeconômicas e culturais dos participantes da prova do ENEM 2019.
- Apresentar as principais características do ENEM.
- Verificar características do ENEM 2019 com relação ao ano anterior.
- Compreender o contexto do ENEM desde sua criação até o modelo atual.
- Analisar os números de inscritos no ENEM nos anos de 2011 a 2019.

2.3 JUSTIFICATIVA

O tema desta pesquisa foi escolhido a partir do interesse de pesquisa da área da análise de dados voltados para o ENEM, por reconhecê-lo como uma ferramenta importante para o desenvolvimento da educação brasileira, podendo assim observar os problemas que muitas vezes interferem nesse processo.

Partimos da premissa de que a cada ano o ENEM apresenta características diferentes e que necessitamos observar as características dos participantes e do próprio exame para nos situarmos quanto à importância do ENEM para a educação brasileira.

Este trabalho será voltado para a visualização e análise de dados, por reconhecer que muitos aspectos do ENEM necessitam ser compreendidos, para termos ciência da realidade dos alunos que realizam o exame e observar seus aspectos fundamentais.

3. FUNDAMENTAÇÃO TEÓRICA

3.1 O ENEM

A partir da reforma do ensino médio proposta pela LDB 9.394/96, foram propostas mudanças significativas nos paradigmas de ensino nos quais prevaleciam a memorização de conteúdos passando-se a utilizar competências, habilidades e atitudes de acordo com os PCN (BRASIL, 1997).

Segundo o INEP (2019), os resultados do ENEM deverão possibilitar:

I - A constituição de parâmetros para a autoavaliação do participante, com vistas à continuidade de sua formação e a sua inserção no mercado de trabalho;

II - A criação de referência nacional para o aperfeiçoamento dos currículos do ensino médio;

III - A utilização do Exame como mecanismo único, alternativo ou complementar para acesso à educação superior, especialmente a ofertada pelas instituições federais de educação superior;

IV - O acesso a programas governamentais de financiamento ou apoio ao estudante da educação;

V - A sua utilização como instrumento de seleção para ingresso nos diferentes setores do mundo do trabalho;

VI - O desenvolvimento de estudos e indicadores sobre a educação brasileira.

O ENEM disponibiliza as médias e os percentuais dos alunos em cada um dos quatro níveis de proficiência e da redação dos estudantes que participaram da prova, por escola, para cada uma das áreas de conhecimentos consideradas.

De acordo com INEP (2021), os participantes do exame podem utilizar suas notas para ter acesso ao Sistema de Seleção Unificada (SISU) e ao Programa Universidade para Todos (PROUNI). As notas obtidas também são aceitas em mais de 50 instituições de educação superior portuguesas, podem também pleitear financiamento estudantil em programas do governo, como o Fundo de Financiamento Estudantil (FIES).

Quanto ao critério para participação do exame temos que qualquer pessoa que já concluiu o ensino médio ou está concluindo a etapa pode fazer o ENEM para ter acesso à educação superior. Os participantes que ainda não concluíram o ensino médio podem

participar como “treineiros” e seus resultados no exame servem somente para autoavaliação de conhecimentos (INEP, 2021).

Nas seções 3.1.1 irá ser exposto de forma sucinta como o ENEM é organizado atualmente no nosso país. Já na seção 3.1.2 será abordado um pouco sobre o histórico e trajetória do ENEM. Em seguida tem-se a seção 3.1.3 na qual é explicado as principais características do Exame.

3.1.1 Estrutura do Exame Nacional do Ensino Médio – ENEM

O ENEM é estruturado a partir de uma matriz que indica a associação entre conteúdos, competências e habilidades básicas próprias ao jovem e ao jovem adulto, na fase de desenvolvimento cognitivo e social correspondente ao término da escolaridade básica (INEP, 2005).

Com relação às habilidades que o Exame Nacional do Ensino Médio requer dos estudantes temos que elas decorrem das competências adquiridas e referem-se ao plano imediato do saber fazer, através das ações e operações, as habilidades aperfeiçoam-se e articulam-se, possibilitando nova organização das competências (BRASIL, 1998).

Há, entretanto, aqueles que não acreditam que o ENEM seja um método eficaz para realizar a seleção dos candidatos que pretendem concorrer a uma vaga no ensino superior. Para o professor Nilson Machado, da Faculdade de Educação da USP, e um dos idealizadores do exame no formato até 2009, o ENEM não foi projetado para ser um processo seletivo, por isso apresenta atualmente tantos desvios (FOLHA DE SÃO PAULO, 2012), São eles:

- 1) Com a transformação de uma prova de 63 questões em quatro provas, uma para cada área em que se organiza o Ensino Médio, com 45 questões cada uma, o teste ficou excessivamente longo para o conteúdo que examina;

- 2) A premissa de que a prova deve ser contextualizada, na prática, tem sido tratada como se a palavra “contexto” significasse “com muito texto”. Os enunciados são muito extensos e, na maioria das vezes, o estudante leva um bom tempo até que a resposta correta possa ser encontrada;

- 3) A utilização da sofisticada Teoria da Resposta ao Item (TRI), na correção das provas, é o mais grave de todos os desvios. As limitações na qualidade e na quantidade dos itens dos bancos de questões inviabilizam o sucesso de tal recurso. Sem contar os problemas logísticos: roubos de provas, quebras de sigilo, inadequações na pré-testagem têm se sucedido e afetam significativamente a integridade e a credibilidade da prova.

De acordo com Andrade, D. F. e Tavares e Valle (2000), a Teoria da Resposta ao Item (TRI) é uma metodologia que sugere formas de representar a relação entre a probabilidade de um indivíduo dar uma certa resposta a um item e seus traços latentes. Traços latentes são características do indivíduo que não podem ser observadas diretamente, isto é, não existe um aparelho capaz de medi-las diretamente, como, por exemplo, um termômetro que mede diretamente a temperatura. Portanto, essas características são mensuradas através de variáveis secundárias que sejam relacionadas com o traço latente em estudo.

Para Andrade, a TRI é um modelo matemático que determina a probabilidade de um indivíduo dar uma resposta ao item em função de parâmetros do item e da habilidade do respondente. Leva-se em conta para aplicação da TRI três fatores:

- i) A natureza do item - Dicotômico ou não dicotômico.
- ii) O número de populações envolvidas - Apenas uma ou mais de uma.
- iii) A quantidade de traços latentes - Apenas um ou mais de um.

No Brasil, a TRI é utilizada desde 1995 nas provas do Sistema Nacional de Avaliação da Educação Básica (SAEB), que mede o desempenho dos estudantes do ensino fundamental e médio. Em 2009, foi usada pelo ENEM com o objetivo de garantir a comparação das notas do exame daquele ano com as próximas edições do Exame.

Atualmente o ENEM é aplicado em dois domingos consecutivos, sendo que no primeiro domingo os alunos respondem 45 questões objetivas de **Ciência Humanas e suas Tecnologias**, 45 questões de **Linguagens, Códigos e suas Tecnologias** e elaboram uma **redação**. Já no segundo domingo a prova é composta por 45 questões de **Matemática e suas Tecnologias** e 45 questões de **Ciências da Natureza e suas Tecnologias**. O caderno de provas já vem com o nome do participante sendo quatro modelos (cores) em cada dia, o aluno recebe também um cartão resposta onde irá marcar as questões de múltipla escolha e redigir a redação para ser corrigida pela comissão avaliadora (BRASIL ESCOLA, 2018).

O ENEM é um exame que aponta unicamente o desempenho do estudante em termos de nota (CURI; MENEZES FILHO, 2013). Ele não considera nenhuma métrica que possa contextualizar o resultado obtido, de forma a reconhecer que mesmo com desempenho inferior, alguns alunos podem ter tido um efeito de esforço próprio maior no resultado, pois quando apenas a nota final é considerada, as utilizações do resultado obtido no ENEM ficam condicionados ao desempenho em termos de nota, que por sua vez depende de variáveis que impactam diretamente no resultado dos alunos nas provas (FIGUEIRÊDO; NOGUEIRA; SANTANA, 2014).

De acordo com INEP (2021) o ENEM apresenta os seguintes eixos cognitivos (comuns a todas as áreas de conhecimento)

I. Dominar linguagens (DL): dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica e das línguas espanhola e inglesa.

II. Compreender fenômenos (CF): construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos históricos geográficos, da produção tecnológica e das manifestações artísticas.

III. Enfrentar situações-problema (SP): selecionar, organizar, relacionar, interpretar dados e informações representadas de diferentes formas, para tomar decisões e enfrentar situações-problema.

IV. Construir argumentação (CA): relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.

V. Elaborar propostas (EP): recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.

3.1.2 Trajetória do ENEM

O ENEM foi criado em 1998 pelo Ministério da Educação com o propósito de avaliar o desempenho dos estudantes do ensino médio, visando examinar o desenvolvimento de competências e habilidades necessárias ao exercício pleno da cidadania (INEP, 2011).

A criação do ENEM se deu por meio da Portaria do Ministério da Educação (MEC) Nº 438/1998, onde constavam os seguintes objetivos, além do já citado: II- servir de referência para os egressos do Ensino Médio, III- fornecer dados para o acesso à Educação Superior, IV- servir de acesso aos cursos profissionalizantes.

Na sua fase de concepção, o ENEM não era um exame obrigatório, mas tinha como exigência o pagamento de uma taxa de inscrição que excluía aqueles que não tinham condições financeiras favoráveis. Segundo o Censo de educação superior (INEP, 2011), na sua primeira edição, o exame contou com cerca de 115.600 participantes.

Segundo o INEP (2019) em 1999, na segunda edição do exame, 93 instituições de Ensino Superior já aderiram aos resultados do ENEM em seus processos de seleção, ainda de forma independente, com a iniciativa vinda das próprias instituições.

Conforme o Ministério da Educação (2002), os três principais objetivos que conduziram a criação do exame nacional do ensino médio foram:

- Oferecer uma referência para que cada cidadão possa proceder a sua autoavaliação tendo em vista suas escolhas futuras, tanto em relação ao mercado de trabalho quanto em relação a continuidade dos estudos.
- Estruturar uma avaliação da educação básica que sirva como modalidade alternativa ou complementar aos processos de seleção nos diferentes setores do mundo do trabalho.
- Estruturar uma avaliação da educação básica que sirva como modalidade alternativa ou complementar aos exames de acesso aos cursos profissionalizantes pós-médios e ao ensino superior.

Para Travitzki (2013), entre os anos de 1998 e 2008 o objetivo com o ENEM era mais voltado para a preparação dos indivíduos para o exercício da cidadania, enquanto que a partir de 2009 passou a ter um objetivo direcionado a preparação para o Ensino Superior (Figura 1).

Em 2009, ocorreu a reformulação metodológica do ENEM e de sua utilização como forma de seleção nos processos seletivos de muitas universidades federais. Reformularam-se as Matrizes de Referência do Exame, que passaram a ser estruturadas por competências em quatro áreas do conhecimento – Linguagens, Códigos e suas tecnologias, Matemática e suas tecnologias, Ciências Humanas e suas tecnologias e Ciências da Natureza e suas tecnologias, apresentando, ainda, eixos cognitivos comuns a todas as áreas (INEP, 2019).

Nesse mesmo ano o MEC modificou novamente o ENEM, adicionando mais três objetivos aos quatro já especificados. O ENEM passou a ser usado também para a certificação no nível de conclusão do Ensino Médio, avaliação de desempenho dos ingressantes na Educação Superior e para avaliação de desempenho acadêmico dos concluintes do Ensino Médio, gerando o resultado por escola, conforme regulamentação das Portarias Nº 462/2009 (MEC) e Nº 109/2009 (INEP).

A partir das reformulações ocorridas no exame, verificou-se que o ENEM passou a ser muito requisitado pelos estudantes que estão concluindo o Ensino Médio, sendo que o mesmo já era muito procurado desde o início da utilização dos resultados obtidos pelos alunos no PROUNI que realiza concessão de bolsas para o sistema de ensino superior privado, já por volta de 2011 quando muitas universidades aderiram Sistema de Seleção Unificada (SISU) o número de inscritos superou 5 milhões (GONÇALVES JR e BARROSO, 2014).

No ano de 2004, o Programa Universidade Para Todos (PROUNI) foi criado pela Medida Provisória nº 213/2004, convertida posteriormente na lei nº 11.096/2005. Este

programa tem a função de fornecer bolsas de estudos integrais e parciais, para que alunos de baixa renda possam fazer cursos superiores nas universidades privadas. Em 2006, devido à implementação do PROUNI, foi acrescentado aos objetivos do ENEM, por meio da Portaria INEP nº 07/2006, o quarto objetivo, conforme segue “IV - possibilitar a participação e criar condições de acesso a programas governamentais” (BRASIL, 2006, p. 1).

Em 2005 esta medida virou Lei nº 11.096 e disponibilizou o artigo 3º que diz:

Art.3º O estudante a ser beneficiado pelo PROUNI será pré-selecionado pelos resultados e pelo perfil socioeconômico do Exame Nacional do Ensino Médio – ENEM ou outros critérios a serem definidos pelo Ministério da Educação, e na etapa final, selecionado pela instituição de ensino superior, segundo seus próprios critérios, à qual competirá, também, aferir as informações prestadas pelo candidato (BRASIL, 2005, p. 1).

Em 2007, o governo criou o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), que visa ampliar o acesso e permanência no ensino superior. O REUNI foi responsável por expandir o número de universidades no Brasil, criando diversos campi no interior dos estados (de Paula, 2017). O Programa contribuiu de forma significativa para o aumento de participantes das regiões Norte e Nordeste do Brasil a partir de 2010, tornando a proporção de participantes por região mais próxima à proporção populacional das regiões. O número de inscritos aumentou consideravelmente ano após ano até 2014, atingindo quase 9 milhões de estudantes, quando começou a sofrer oscilações e a diminuir (Santos et al., 2019).

3.1.3 Características Gerais do ENEM

Nesta edição do Exame Nacional do Ensino Médio analisamos os dados de um universo de 5.095.270 estudantes inscritos, os quais se encontravam espalhados pelas cinco regiões brasileiras da seguinte forma: 598.985 (Norte); 1.741.107 (Nordeste); 1.793.813(Sudeste); 538.358 (Sul); 423.005 (Centro-oeste).

Um fato que deve ser levado em consideração é que apesar da região sudeste apresentar o maior número de candidatos representando 35% do total, temos que a região nordeste é a que mais se aproxima desta quantidade, ficando 34%. As regiões norte, sul e centro-oeste apresentam respectivamente 12%,10% e 8% aproximadamente.

Com relação aos candidatos que não compareceram nos dois dias para realizar o exame temos um total de 1.160.010 o que representa 23%, neste cenário destaca-se a região sudeste

com 36% dos faltosos e a região na qual os participantes mais compareceram foi a centro-oeste que teve apenas 8% de faltosos.

Diante desses números é perceptível que existem fatores que influenciam para que os candidatos que realizaram suas inscrições desistam de realizar o certame, tendo em vista que essa desistência acarreta prejuízos não apenas para os candidatos, mas também para o Ministério de Educação pelos gastos na impressão de milhões de cadernos de prova que são desperdiçados. Acarretando como penalidade para aquele candidato que conseguiu a isenção da taxa de inscrição, a impossibilidade de conseguir novamente no ano seguinte.

Dentre as características dos participantes do ENEM 2019 temos aqueles que solicitaram atendimento nome social para se adequarem a sua identidade de gênero distribuída pelo Brasil da seguinte forma: região norte (20); região nordeste (119); região sudeste (193); região sul (36) e região centro-oeste (26).

O nome social é atualmente utilizado para identificar aquelas pessoas que são transsexuais ou travestis que apesar de serem registradas com um nome relacionado ao seu sexo de nascimento, possuem um nome pelo qual são identificados e que pode ser informado no ato da inscrição, afim de evitar certos constrangimentos com relação a forma de tratamento.

No cenário da sociedade atual, ainda existe muito preconceito com relação aqueles indivíduos que utilizam o nome social ao realizar cadastros ou inscritos em exames.

3.2 DATA SCIENCE

A Data Science é um termo muito utilizado na atualidade por possibilitar o estudo de determinadas informações desde sua coleta, transformação e análise dos dados. De acordo com PASSOS (2016), ela é descrita como a ciência responsável pela análise e utilização de dados que incorpora técnicas e teorias de diversas áreas, como lógica, matemática, estatística, computação, engenharia e economia. Pode ser definida como um conjunto de técnicas utilizadas no processamento e análise de dados, com intuito de fornecer informações para decisões inteligentes. Para tanto, mescla-se diversas áreas do conhecimento, desde conceitos simples de estatística até complexos algoritmos.

Segundo Boschetti e Massaron (2016), a Ciência de Dados é um domínio de conhecimento relativamente novo, embora seus componentes principais tenham sido estudados e pesquisados por muitos anos pela comunidade de Ciência da Computação. Seus componentes incluem álgebra linear, modelagem estatística, visualização, linguística corporal,

análise de gráficos, aprendizado de máquina, inteligência de negócios, armazenamento e recuperação de dados.

Segundo Vanderplas (2016), Data Science é o conjunto de habilidades interdisciplinares que estão se tornando cada vez mais importantes em muitas aplicações em toda a indústria acadêmica.

3.3 BIG DATA

Big Data pode ser definido de acordo com Silva (2016), como um conjunto de dados que para serem utilizados que necessitam de métodos e tecnologia avançadas, por apresentarem características como velocidade, variedade, complexidade e grande volume de dados.

Para Martins (2014), Big Data possui outras características com relação ao tipo de dados, podendo considerar estruturados aqueles que possuem estrutura rígida e foram projetadas a partir de um modelo; semiestruturado àquele que necessitam estar armazenados em Sistemas Gerenciadores de Banco de dados e não possuem um tipo de estrutura e os não-estruturados aqueles que não tem formato, não seguem regras e não são previsíveis.

Os dados utilizados nesta pesquisa são dados estruturados, que foram obtidos dos microdados do Enem do ano de 2019. Devido ao seu grande volume de dados obtidos, foi necessário o provisionamento de uma infraestrutura em nuvem para o armazenamento e processamento e análise dos dados. Foi utilizado o Google Drive para o armazenamento dos dados em nuvem, e para processamento e análise dos dados utilizamos o Google Colab (DA SILVA, 2020).

De acordo com Da Silva (2020), o Google Colab ou Colaboratório é um espaço digital disponível na nuvem, hospedado pelo Google e que pode ser utilizado de forma gratuita, visando auxiliar a utilização da linguagem de programação Python. Este ambiente é uma máquina virtual chamada de notebook e que permite programar e rodar códigos em Python bem como comentar, criar análises, gráficos, relatórios, modelos de aprendizado de máquina e Deep Learning, dentre outras funcionalidades.

Para utilizarmos o Google Colab não necessitamos realizar a instalação de nenhum software e nem necessitamos de um supercomputador pois a plataforma disponibiliza um cluster de computadores, necessitando apenas de ter acesso a internet. Este foi um dos motivos de optarmos por esta plataforma por não termos uma máquina com as configurações necessárias para rodar os códigos (DA SILVA, 2020).

Outra característica importante do Google Colab é a possibilidade de dividir blocos de texto e códigos no mesmo arquivo, facilitando os comentários dos códigos, a compilação do código separada, elaboração simultânea de relatório, clareza e organização dos estudos (DA SILVA, 2020).

3.4 VISUALIZAÇÃO DE DADOS

Frequentemente depara-se com banco de dados que apresentam uma quantidade considerável de informações, podendo conter milhões de elementos de dados, como é o caso do ENEM, pois ele é um exame que abrange todo o país. Para trabalhar com essa grande quantidade de dados se faz necessário utilizar novas técnicas que nos possibilitem uma melhor visualização dos dados, contribuindo de forma significativa na análise dos dados (CAIRO, 2012; MANOVICH, 2011; DUR, 2014; JL VALERO, 2014).

A visualização de dados é uma das principais habilidades necessárias para o profissional que trabalha na área de ciências de dados sendo um diferencial nessa área. Nos dias atuais com o crescimento da utilização da tecnologia fica mais evidente a necessidade de utilizar meios mais eficientes de exibir informações através de gráficos de grandes volumes de dados.

O emprego das técnicas de visualização de dados (Dataviz) nesse campo é uma estratégia emergente e de inovação que tem sido explorado em distintas áreas do conhecimento e recebido cada vez mais atenção em estudos e pesquisas na atualidade (CAIRO, 2012; MANOVICH, 2011; DUR, 2014; JL VALERO, 2014).

De acordo com Lemos (2002), a visualização de dados ampliou seu alcance com as tecnologias digitais da sociedade em rede e da cibercultura que possibilitou que os dados invisíveis, isto é, armazenados em redes digitais, somados aos demais elementos combinatórios (gráficos, mapas, dados científicos, números etc.), fossem transmutados para o ambiente Web e conferissem um maior enriquecimento às estruturas gráficas.

Meirelles (2011) conceitua a visualização de dados como “representações de dados que podem assumir diferentes formas, tais como sistemas de notação, mapas, diagramas, explorações de dados interativos, e outras invenções gráficas”. A visualização de dados é o processo de utilização de tecnologias mediadas por computador e digitais para exibir informações quantitativas e qualitativas. Neste sentido, o termo "visualização de dados" é usado aqui para especificar este processo de transformar dados numéricos ou informativos científicos para distinguir a visualização de dados de outras visualizações, tais como a arte não baseada em dados. As visualizações de dados estão cada vez mais se tornando complexas

para narrativas sofisticadas que se utilizam de mapas com dados que permitem interação (VIÉGAS, 2013).

Por ser uma ferramenta poderosa, a mesma pode ser usada para distorcer ou mascarar informações verdadeiras. Contudo, se utilizar eticamente, o processo de visualização pode auxiliar a enxergar o mundo de outra maneira, revelar padrões ocultos e tendências que não estão disponíveis de outra forma. Em outras palavras, a visualização de dados tem como especialidade contar histórias. Além disso, visualização é um processo de mapeamento de informações para imagens, para isso, criam-se regras e metodologias que interpretam dados numa dinâmica muitas vezes complexa (MURRAY, 2013).

4 METODOLOGIA

4.1 OS MICRODADOS ENEM 2019

Os Microdados do ENEM fazem parte dos dados abertos do MEC e podem ser obtidos de forma gratuita pelo site do INEP. Eles são o menor nível de desagregação de dados recolhidos por meio do Exame. Eles atendem à demanda por informações específicas ao disponibilizar as provas, os gabaritos, as informações sobre os itens, as notas e o questionário respondido pelos inscritos no ENEM (INEP, 2019).

Os dados disponibilizados vêm numa pasta compactada se fazendo necessário extrair os arquivos para poderem ser utilizados. Ele apresenta 5 pastas na qual tive acesso às provas, os gabaritos, as informações sobre os itens, as notas e os questionários resolvidos pelos participantes do Exame.

Ele apresenta uma pasta de dados contendo dois arquivos do tipo excel, `ITENS_PROVA_2019` uma base contendo informações sobre as provas e `MICRODADOS_ENEM_2019` que facilita a utilização dos resultados contendo os questionários respondidos pelos participantes (INEP, 2019).

A pasta dicionário de dados contém informações sobre as variáveis que são indispensáveis para a compreensão mais completa da organização do banco de dados e da própria estrutura dos instrumentos utilizados.

4.1.1 A utilização de Microdados

Os microdados se constituem no menor nível de desagregação de dados recolhidos por pesquisas, avaliações e exames realizados. As informações podem ser obtidas via download, em formato ASCII, e contêm inputs (canais de entrada) para leitura por meio dos softwares SAS e SPSS (INEP, 2021).

As bases de microdados estão organizadas de forma a serem compreendidas por softwares específicos, o que agiliza o processo de tratamento e cálculos estatísticos. No Brasil, o Instituto Nacional de Pesquisas Educacionais Anísio Teixeira (Inep), é o maior produtor de microdados relativos à educação: Censo Escolar, Censo da Educação Superior, Prova Brasil, Saeb, Enem são algumas das principais bases de microdados do ministério da educação (INEP,2021).

Eles fazem parte dos dados abertos do MEC e podem ser obtidos de forma gratuita no site do INEP. Os microdados vem num arquivo compactado contendo em sua composição 5 pastas: DADOS, DICIONÁRIO, INPUTS, LEIA-ME E DOCUMENTOS TÉCNICOS e PROVAS E GABARITOS. A pasta de dados é composta por dois arquivos do tipo Excel que são ITENS_PROVA_2019 e MICRODADOS_ENEM_2019. Possui também uma pasta chamada DICIONÁRIO que auxilia na compreensão das variáveis que serão analisadas. A pasta LEIA-ME E DOCUMENTOS TÉCNICOS é composta pelo Edital_enem_2019, Leia-Me_Enem_2019, manual_de_redação_do_enem_2019 e a matriz_referencia_enem (INEP, 2021).

4.1.2 Descrição dos Microdados do ENEM 2019

Quadro1- Pastas e Arquivos dos Microdados do ENEM 2019.

| | |
|---------------|-------------------------------|
| PASTAS | 5 Pastas: |
| | DADOS |
| | DICIONÁRIO |
| | LEIA-ME E DOCUMENTOS TÉCNICOS |
| | INPUTS |
| | PROVAS E GABARITOS |

Fonte: INEP-Microdados do ENEM 2019.

Quadro 2-Pasta de Dados e Arquivos dos Microdados do ENEM 2019.

| Dados | 2 Arquivos: | Descrição: |
|--------------|--------------------------|---|
| | MICRODADOS_ENEM_2019.csv | Informações gerais sobre a realização das provas, a caracterização do participante e da escola que ele declarou ter frequentado e as notas das provas objetivas e da redação. |

| | | |
|--|-----------------------|---|
| | ITENS_PROVA _2019.csv | Informações gerais sobre os itens das provas. |
|--|-----------------------|---|

Fonte: INEP-Microdados do ENEM 2019.

Quadro 3- Pasta Dicionário e Arquivos dos Microdados do ENEM 2019.

| Dicionário | 2 Arquivos: | Descrição |
|-------------------|---|--|
| | Dicionário_Microdados_ENEM_2019.xlsx Dicionário_Microdados_ENEM_2019.ods | Informações gerais sobre as variáveis contidas na base e sobre as perguntas e alternativas do questionário aplicado. |

Fonte: INEP-Microdados do ENEM 2019.

Quadro4- Pasta Leia-Me e Documentos Técnicos e Arquivos dos Microdados do ENEM 2019.

| Leia-Me e Documentos Técnicos | 4 Arquivos: | Descrição |
|--------------------------------------|------------------------------------|--|
| | Leia-me_ENEM_2019.pdf | Breve descrição do ENEM, bem como das informações sobre as bases e os arquivos disponibilizados nos Microdados. |
| | Matriz_referencia_ENEM.pdf | Apresentação da Matriz de Referência, que compreende os eixos cognitivos, as competências e as habilidades avaliadas em cada área de conhecimento do Ensino Médio. |
| | Edital_ENEM_2019.pdf | Edital de publicação do ENEM 2019 |
| | Manual_de_redacao_do_ENEM_2019.pdf | Expõem a metodologia de avaliação da redação, bem como o que se espera do participante em cada uma. das competências avaliadas |

Fonte: INEP-Microdados do ENEM 2019.

4.2 LINGUAGEM DE PROGRAMAÇÃO

A linguagem de programação utilizada neste trabalho foi o python por ser uma das mais utilizadas para análise de dados, apresentando como principal característica ser uma linguagem de propósito geral. De acordo com Manzano (2011), ela é classificada como uma linguagem que pode ser utilizada tanto para desenvolvimento de programas comerciais quanto de programas científicos.

Por apresentar uma sintaxe clara e concisa que favorece a legibilidade do código fonte, tornando a linguagem mais produtiva, Python é uma das linguagens mais bem aceitas no meio científico em geral, principalmente por cientistas de áreas que não estão necessariamente ligadas a tecnologia (BORGES, 2014).

4.2.1 Google Colab e Python

De acordo com Carneiro et al. (2018), o ambiente Google Colaboratory ou Colab é um serviço de nuvem gratuito que é hospedado pelo Google para incentivar a pesquisa de Aprendizado de Máquina e Inteligência Artificial. O ambiente digital é uma máquina virtual, ou notebook, no qual o estudante pode desenvolver suas práticas de programação em Python e aplicar os conceitos de Ciências de Dados.

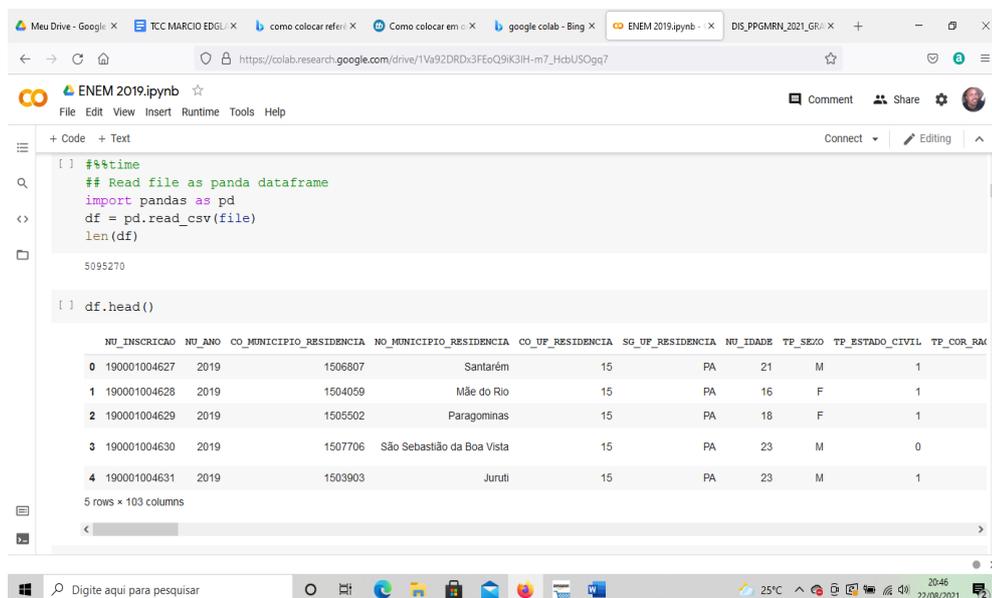
Segundo VISHAKHALALL (2018) os principais benefícios do Colab, são:

- Suporte para linguagem Python;
- Aceleração de GPU de forma gratuita;
- Possui bibliotecas pré-instaladas: todas as principais bibliotecas Python, como o TensorFlow, o ScikitLearn, o Matplotlib (para gerar gráficos), o Seaborn (para gerar gráficos, entre muitas que estão pré-instaladas e prontas para serem usadas;
- Construído com base no Jupyter Notebook;
- Recursos de colaboração: permite que os desenvolvedores usem e compartilhem o Jupyter notebook entre si sem precisar baixar, instalar ou executar qualquer coisa que não seja um navegador;
- Os notebooks do Google colab são armazenados no drive.

Têm-se como uma característica importante do Google Colab a possibilidade de distribuir separadamente blocos de texto e código dentro do mesmo arquivo, facilitando assim no suporte a comentários no código, compilação separada, clareza e organização nos estudos (SILVA, 2021).

Observa-se na Figura 1, um print da tela do Google Colab, onde verifica-se uma função que tem como objetivo realizar a leitura do arquivo .csv que contém as variáveis e os dados que serão analisados. E ao utilizar a função `df.head()` é exibido uma pequena parte dos dados contidos no arquivo que é composto por 103 colunas.

Figura 1-Print da tela do ambiente virtual do Google Colab.



```

[ ] %%time
## Read file as panda dataframe
import pandas as pd
df = pd.read_csv(file)
len(df)

5095270

[ ] df.head()

```

| | NU_INSCRICAO | NU_ANO | CO_MUNICIPIO_RESIDENCIA | NO_MUNICIPIO_RESIDENCIA | CO_UF_RESIDENCIA | SG_UF_RESIDENCIA | NU_IDADE | TP_SEZO | TP_ESTADO_CIVIL | TP_COR_RAC |
|---|--------------|--------|-------------------------|----------------------------|------------------|------------------|----------|---------|-----------------|------------|
| 0 | 190001004627 | 2019 | 1506807 | Santarém | 15 | PA | 21 | M | 1 | |
| 1 | 190001004628 | 2019 | 1504059 | Mãe do Rio | 15 | PA | 16 | F | 1 | |
| 2 | 190001004629 | 2019 | 1505502 | Paragominas | 15 | PA | 18 | F | 1 | |
| 3 | 190001004630 | 2019 | 1507706 | São Sebastião da Boa Vista | 15 | PA | 23 | M | 0 | |
| 4 | 190001004631 | 2019 | 1503903 | Juruti | 15 | PA | 23 | M | 1 | |

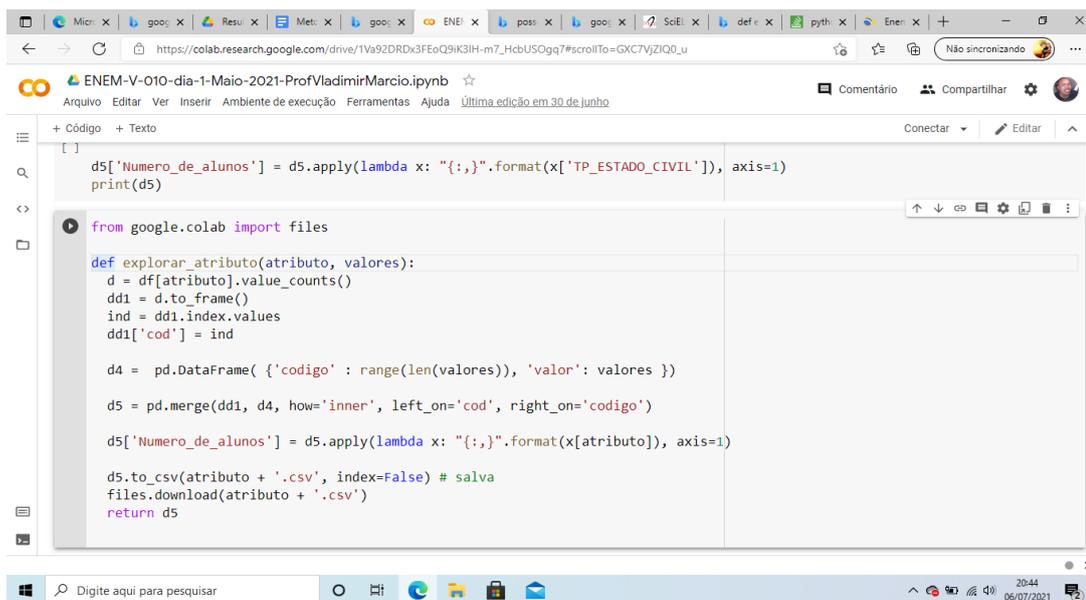
5 rows x 103 columns

Fonte: Google Colaboratory, 2021.

Para Guido van Rossum (1995), o python é uma linguagem de programação de alto nível, orientada a objetos que possui uma sintaxe clara, que facilita a leitura do código e o desenvolvimento rápido de programas. É dinamicamente tipificada e permite que as variáveis tomem diversos tipos, em diferentes locais do programa.

A Figuras 2 e 3, representam códigos que foram utilizados para dividir os microdados em partes menores para serem utilizados em outras ferramentas de visualização de dados como o Tableau e o Excel.

Figura 2-Código para definir uma função em Python utilizando o Google Colab.



```

d5['Numero_de_alunos'] = d5.apply(lambda x: "{:,".format(x['TP_ESTADO_CIVIL']), axis=1)
print(d5)

from google.colab import files

def explorar_tributo(atributo, valores):
    d = df[atributo].value_counts()
    dd1 = d.to_frame()
    ind = dd1.index.values
    dd1['cod'] = ind

    d4 = pd.DataFrame( {'codigo' : range(len(valores)), 'valor': valores })

    d5 = pd.merge(dd1, d4, how='inner', left_on='cod', right_on='codigo')

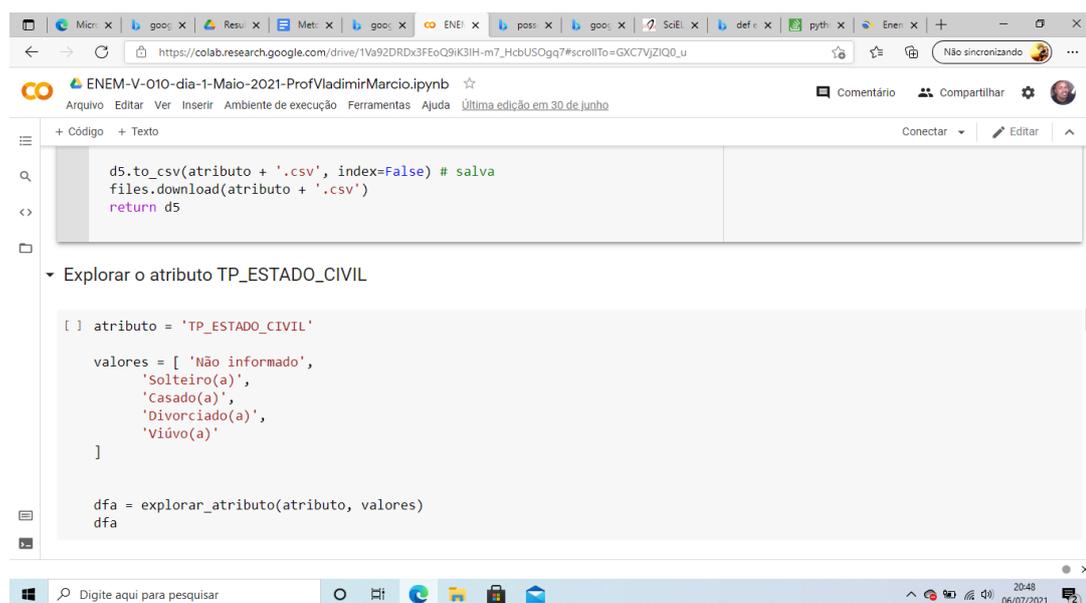
    d5['Numero_de_alunos'] = d5.apply(lambda x: "{:,".format(x[atributo]), axis=1)

    d5.to_csv(atributo + '.csv', index=False) # salva
    files.download(atributo + '.csv')
    return d5

```

Fonte: Google Colab.

Figura 3-Código utilizado para explorar as variáveis.



```

d5.to_csv(atributo + '.csv', index=False) # salva
files.download(atributo + '.csv')
return d5

▼ Explorar o atributo TP_ESTADO_CIVIL

[ ] atributo = 'TP_ESTADO_CIVIL'

valores = [ 'Não informado',
            'Solteiro(a)',
            'Casado(a)',
            'Divorciado(a)',
            'Viúvo(a)'
          ]

dfa = explorar_tributo(atributo, valores)
dfa

```

Fonte: Google Colab.

Para a construção dos gráficos e mapas, foi utilizada uma ferramenta gratuita chamada Tableau Public que é uma plataforma de análise visual (Tableau Public, 2021.1).

Neste trabalho decidiu-se utilizar as variáveis mais gerais a fim de poder compreender as principais características dos participantes do ENEM 2019 em cada região do nosso país.

As variáveis utilizadas na pesquisa foram:

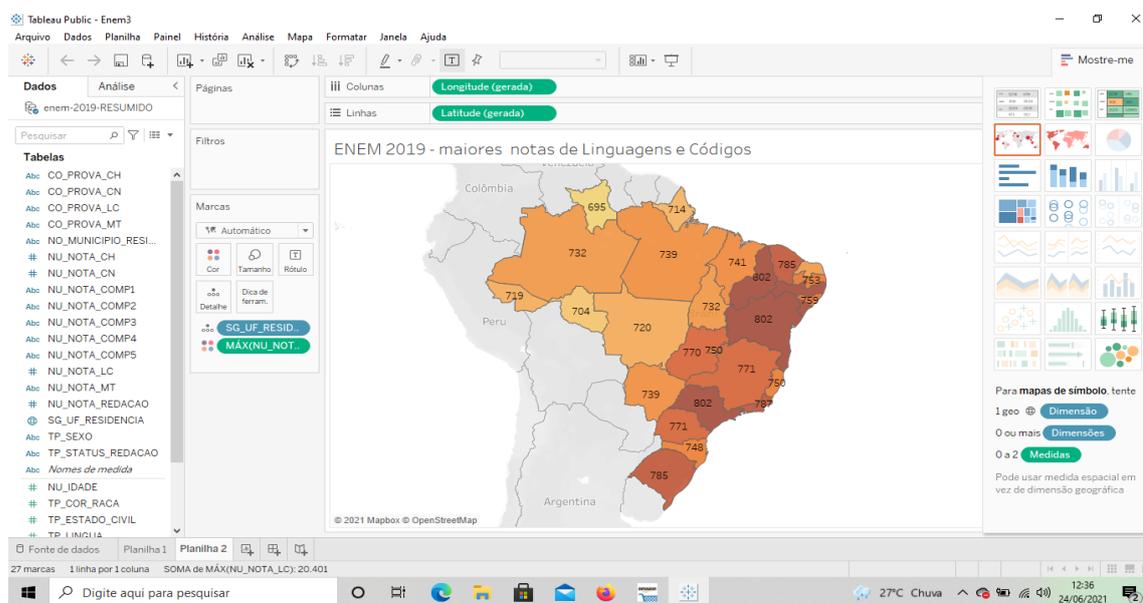
- **NU_IDADE** é uma variável numérica representa a idade dos participantes do exame.

- **TP_SEXO** é uma variável alfanumérica que representa o sexo do participante tendo as opções: M-masculino e F-feminino.
- **TP_ESTADO_CIVIL** é uma variável numérica que representa o estado civil informada pelos participantes tendo as seguintes opções: '0'-Não declarado; 1-Solteiro (a); 2-Casado (a)/Mora com companheiro(a); 3-Divorciado(a)/Desquitado(a)/Separado(a);4-Viúvo.

- **TP_COR_RACA** é uma variável numérica que representa a cor/raça informada pelos participantes tendo as seguintes opções: 0-Não declarado; 1-Branca; 2-Preta; 3-Parda; 4-Amarela e 5-Indígena.
- **TP_NACIONALIDADE** é uma variável numérica que representa a nacionalidade informada pelos participantes tendo como opções: 0-Não informado; 1-Brasileiro (a); 2-Brasileiro(a) Naturalizado(a); 3-Estrangeiro(a); 4-Brasileiro(a) Nato(a), nascido(a) no exterior.
- **SG_UF_NASCIMENTO:** é uma variável alfanumérica que representa a sigla da Unidade da Federação informada pelo aluno.
- **TP_ST_CONCLUSAO:** é uma variável que representa a situação de conclusão do Ensino Médio informada pelo aluno que pode ser uma das seguintes opções:1-Já conclui o Ensino Médio;2-Estou cursando e concluirei o Ensino Médio em 2019;3-Estou cursando e concluirei o Ensino Médio após 2019;4 Não conclui e não estou cursando o Ensino Médio.
- **TP_ESCOLA** é uma variável do tipo numérica que representa qual o tipo de escola que o aluno frequentou durante o Ensino Médio fornecendo as seguintes opções:1-Não respondeu; 2-Pública3-Privada; 4-Exterior.
- **TP_ENSINO** é uma variável que representa o tipo de instituição que o aluno concluiu ou irá concluir o Ensino Médio apresentando as seguintes informações:1-Ensino regular; 2-Educação Especial-Modalidade Substitutiva.
- **TP_DEPENDENCIA_ADM_ESC** é uma variável numérica que representa a dependência administrativa da escola tendo como opções:1-Federal; 2-Estadual; 3-Municipal e 4-Privada.
- **TP_LOCALIZACAO_ESC** é uma variável numérica que representa a situação de funcionamento da escola possuindo as seguintes informações: 1-Em atividade; 2-Paralisada; 3-Extinta.

A ferramenta utilizada para construção dos mapas utilizada foi o Tableau Public (2021.1), que é software que tem como objetivo contribuir para análise dos dados possibilitando trabalhar com várias informações de forma simples e intuitiva (Figura 4).

Figura 4-Exemplo de tela do software Tableau utilizado na visualização de dados das médias do ENEM em todo território brasileiro.



Fonte: Dados da pesquisa.

Utilizou-se neste trabalho a versão gratuita, o Tableau Public (2021.1) que pode ser baixado para contribuir com o aprendizado da ferramenta e suas potencialidades. É uma ferramenta que é fácil de utilizar e que possibilita gerar os mapas de forma simples e intuitiva, auxiliando na tarefa de análise de dados.

4.3 ESTATÍSTICA DESCRITIVA (ANÁLISE DE DADOS)

Após a coleta dos dados, tivemos que realizar o tratamento dos dados para que eles pudessem ser utilizados nas ferramentas que auxiliam na análise dos mesmos. Um dos problemas enfrentados nessa fase do trabalho foi o tamanho dos arquivos de dados que dificultava a execução do software. Sendo necessário dividir os dados em partes menores para serem trabalhadas de forma individual.

Foram analisados alguns aspectos relacionados às questões socioeconômicas dos participantes do ENEM, dentre eles a escolaridade dos pais, a renda e se o aluno necessitou de tratamento especializado.

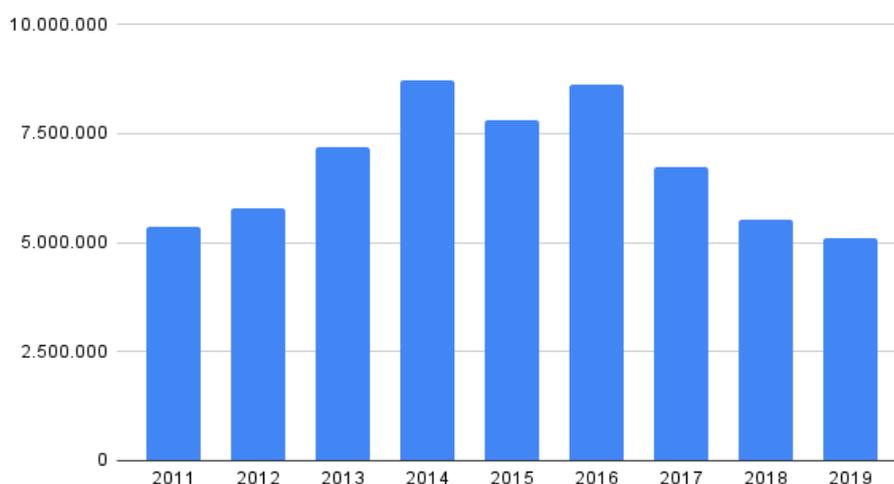
5 RESULTADOS E DISCUSSÕES

O gráfico da Figura 5 mostra o número de inscritos entre os anos de 2011 a 2019 no ENEM de acordo com INEP. Na Figura 5, as colunas representam o quantitativo anual de inscritos no ENEM para o período citado (2011-2019).

Observando essa Figura 5, percebe-se que houve uma redução significativa no número de inscritos para realizar o exame a partir do ano de 2016, sendo os anos de 2014 e 2016 os de maiores inscritos e, o de 2019 com o menor número de inscritos para o ENEM. Assim sendo, é verificado que o Enem de 2019 contou com um total de 5.095.308 candidatos inscritos (Figura 5).

Dentre as principais causas para essa expressiva redução da procura dos estudantes brasileiros em realizar o certame, deve-se a falta de motivação do corpo discente, por perceber a desvalorização que vem ocorrendo com relação à educação nesses últimos anos. E também pela falta de igualdade entre o ensino ofertado pelas escolas públicas em comparação com as escolas particulares, tendo em vista que as mesmas disponibilizam uma educação de qualidade o que favorece para um bom desempenho dos seus alunos.

Figura 5-Comparativo do número de inscritos no ENEM entre os anos de 2011-2019.



Fonte: Dados do INEP.

5.1 CARACTERÍSTICAS DOS PARTICIPANTES DO ENEM 2019

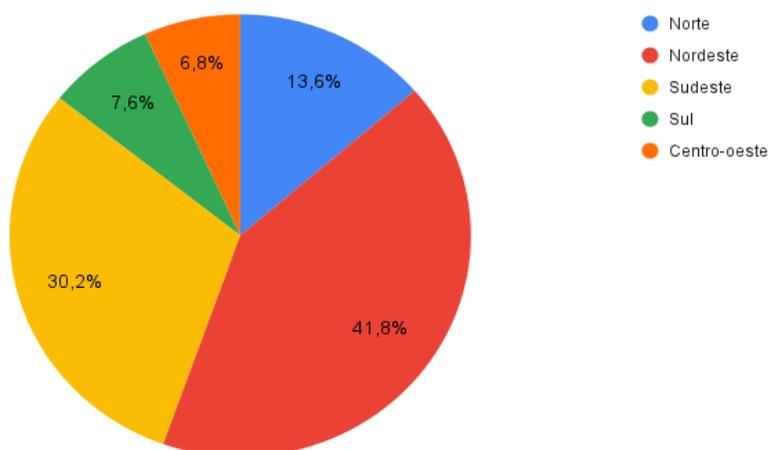
De acordo com Corti (2013), a gratuidade da taxa de inscrição do ENEM foi instituída em 2001 com objetivo de conseguir atrair maior quantidade de concluintes do ensino médio.

A partir da análise do gráfico da Figura 6, verifica-se que a maioria dos participantes que solicitaram isenção na taxa de inscrição foi da Região Nordeste correspondendo a 1.244.925 candidatos, aproximadamente 42%, do total do Brasil, já a Região Centro-oeste foi a que apresentou menor número de participantes (201.741) que solicitaram gratuidade da inscrição.

Um dos principais critérios utilizados para selecionar aqueles estudantes que não tem condições de realizar o pagamento da taxa de inscrição, é verificar se o mesmo possui um cadastro único para benefícios sociais, ou seja, Cadúnico.

De acordo com BRASIL (2007) o Cadúnico é um instrumento de identificação e caracterização socioeconômica das famílias brasileiras de baixa renda, a ser obrigatoriamente utilizado para seleção de beneficiários e integração de programas sociais do Governo Federal.

Figura 6-Porcentagem dos participantes inscritos que solicitaram isenção da taxa de inscrição por região geográfica no Brasil no ano de 2019.



Fonte: Dados da pesquisa

Dentre os inscritos no ano de 2019, a maioria conseguiu obter isenção da taxa de inscrição, o que representou cerca de 58% do total, demonstrando uma redução se comparada ao ano anterior que chegou a 63% de acordo com os microdados do ENEM 2018 (INEP, 2018).

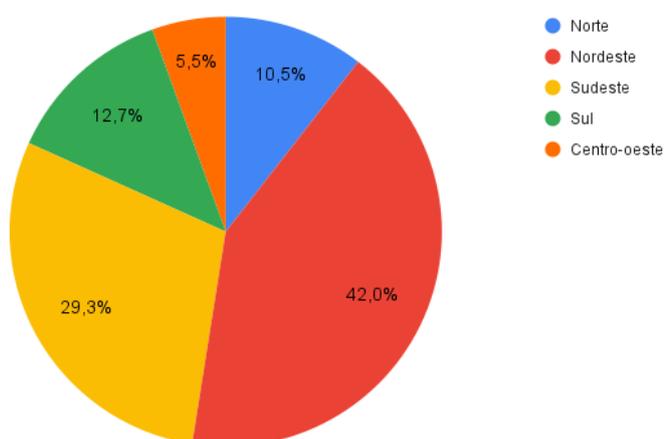
Na Figura 6 constata-se que apesar da Região Sudeste ser a que apresenta maior quantidade de inscritos contabilizando 1.793.813 de candidatos, tem-se que apenas 30,2% solicitaram a isenção da taxa de inscrição, tendo como principal critério para obtenção da referida isenção, o fato do candidato ser inscrito no Cadastro Único para benefícios sociais.

Já as Regiões Sul e Centro-Oeste foram as que apresentaram menor percentual de candidatos que solicitaram a isenção da taxa de inscrição.

Na Figura 7, observou-se o quantitativo dos candidatos que solicitaram atendimento específico e especializado por apresentarem limitações com relação a aplicação do exame. Segundo o INEP, existem três tipos de provas para os candidatos que no ato da inscrição comprovarem possuir algum tipo de necessidade especial por meio de laudo médico, que são: prova em braile para pessoas cegas; prova ampliada para pessoas com baixa visão e prova com leitor para pessoas com algum tipo de deficiência visual. Após o término do prazo de inscrições, os funcionários do INEP entram em contato com esses inscritos para confirmar a necessidade do auxílio e também de tempo extra para realizar o certame.

De acordo com Junqueira e Lacerda (2019), esses procedimentos diferenciados no Exame para pessoas com deficiência foram incluídos paulatinamente no Enem. Porém no ano de 2010, o Conselho Nacional dos Direitos da Pessoa com Deficiência (CONADE) realizou (Recomendação nº 1, de 15 de julho de 2010), que instruiu a inclusão do princípio da acessibilidade em editais e concursos públicos a fim de garantir a igualdade de condições à pessoa surda, como os demais candidatos, determinando expressamente medidas indispensáveis para remoção de barreiras que impeçam a plena e livre concorrência, sem prejuízo de outras que porventura venham a ser adotadas (JUNQUEIRA; LACERDA, 2019, p. 9).

Figura 7- Percentual dos candidatos que realizaram o ENEM 2019 em braile.



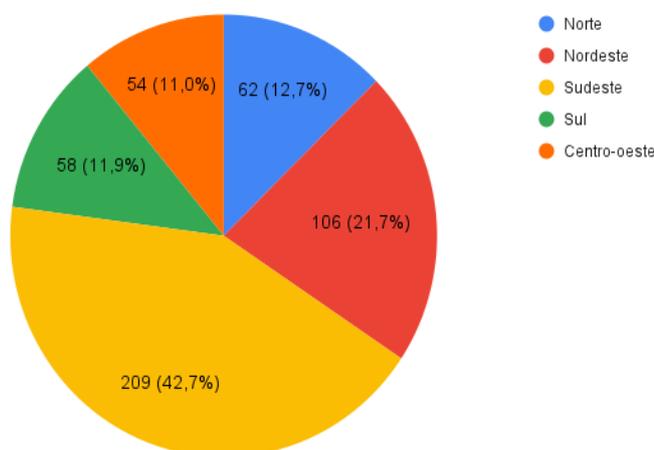
Fonte: Dados da pesquisa

Os participantes surdos ou deficientes auditivos tiveram a possibilidade de solicitar, no momento da inscrição, o recurso para atendimento especializado vídeo prova em Libras. Esses participantes receberam, além do caderno de questões impresso, vídeos contendo a tradução correspondente na Língua Brasileira de Sinais - Libras. Esse recurso para atendimento especializado introduzido em caráter experimental em 2017 permaneceu nas edições 2018 e 2019 (INEP, 2021).

Na figura 7 tem-se que a Região Nordeste apresentou 76 candidatos que tiveram a necessidade de atendimento especializado representando 42%, já a Região com menor quantidade de participantes deficientes foi a Centro-oeste com apenas 10 participantes que comprovaram tal limitação, totalizando 5,5%.

Salientamos ainda, que com relação aos inscritos portadores do espectro autista tem-se na Figura 8, que a Região Sudeste apresentou maior percentual de candidatos portadores da referida doença, já as Regiões Norte, Sul e Centro-oeste apresentaram aproximadamente o mesmo percentual de participantes que declararam possuir esse problema.

Figura 8-Número de inscritos portadores de autismo que participaram do ENEM 2019.



Fonte: Dados

A Figura 8 demonstra o percentual dos candidatos que se declararam possuir autismo, dislexia, discalculia e déficit de atenção recebem um auxílio de um leitor e transcritor para ajudá-lo a responder o exame.

No Quadro 2 temos a quantidade de participantes por faixas de idades menores que 17 anos, entre 26 e 30 anos e com mais de 70 anos. Destaca-se que a região sudeste é a que apresenta o maior número de idosos participantes do ENEM, já com relação aos menores de

idade temos que a região nordeste apresenta uma pequena maioria se comparado a região sudeste.

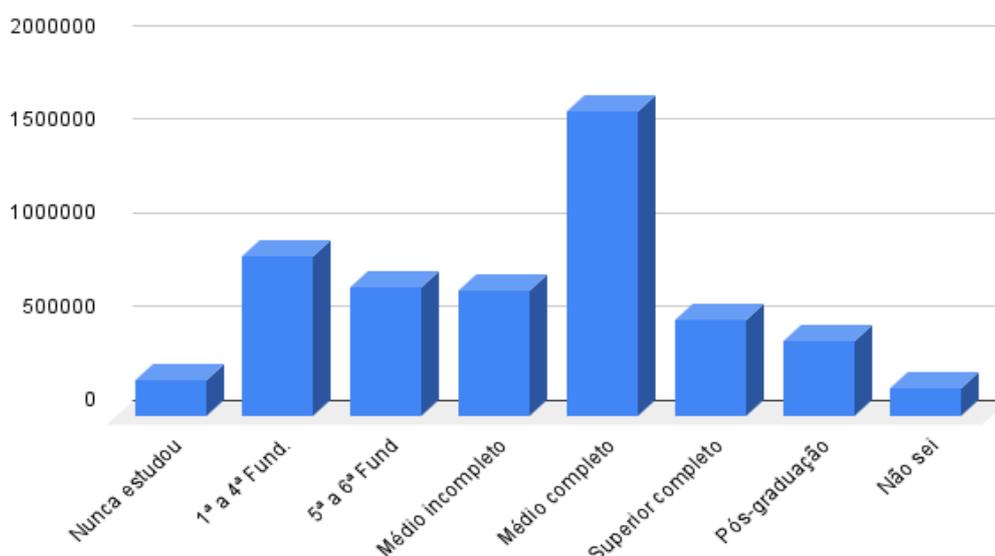
Quadro 5- Faixa etária dos candidatos inscritos no ENEM 2019.

| Região geográfica | <17 | 26-30 | >70 |
|--------------------------|---------------|--------------|---------------|
| Brasil | 400.663 | 403.780 | 717 |
| Norte | 37.885 | 62.974 | 54 |
| Nordeste | 142.585 | 145.999 | 188 |
| Sudeste | 139.341 | 122.063 | 345 |
| Sul | 41.328 | 39.183 | 73 |
| Centro-oeste | 39.524 | 33.561 | 57 |

Fonte: Sinopse ENEM 2019.

Observando a figura 9 tem-se que 32% dos participantes informaram que sua mãe possui ensino médio completo representando a maioria. O fator da escolaridade dos pais influi de forma significativa no desempenho dos candidatos, pois à medida que o nível de escolaridade dos pais aumenta, favorece para uma melhor condição socioeconômica, o que oportuniza uma melhor preparação do candidato ao exame.

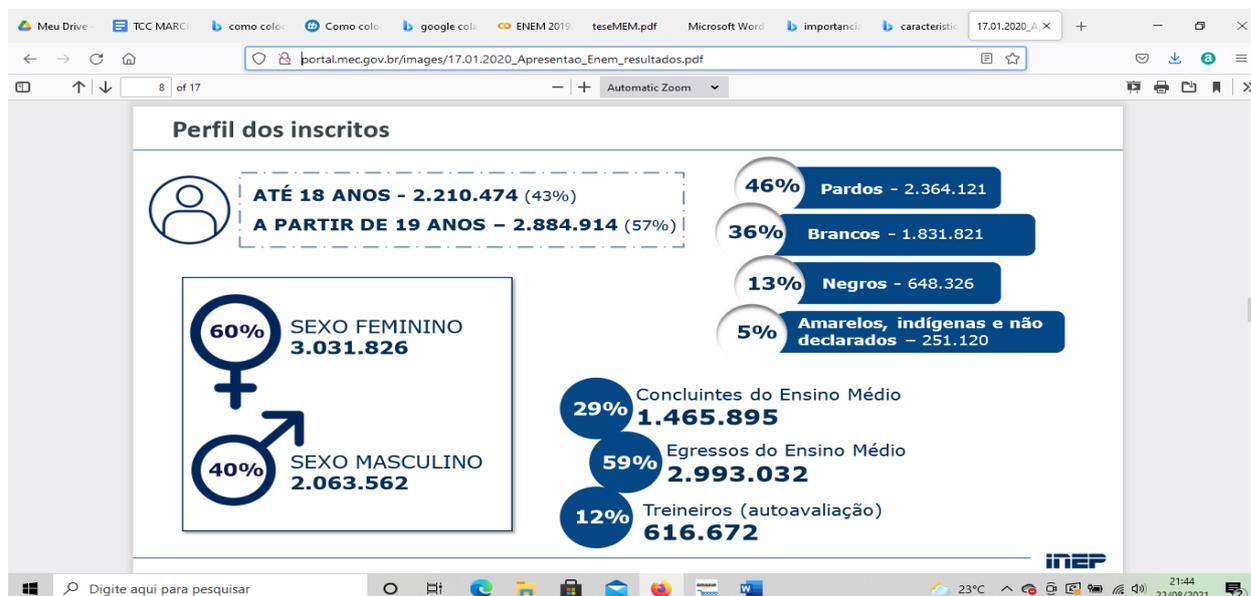
Figura 9- Número de inscritos no ENEM segundo o grau de escolaridade da mãe ou mulher responsável.



Fonte: Dados da pesquisa

Ao analisar a Figura 10 verifica-se características importantes dos participantes do ENEM 2019, como também se constata as semelhanças com as versões anteriores do exame.

Figura 10-Infográfico sobre as características dos participantes do ENEM 2019.

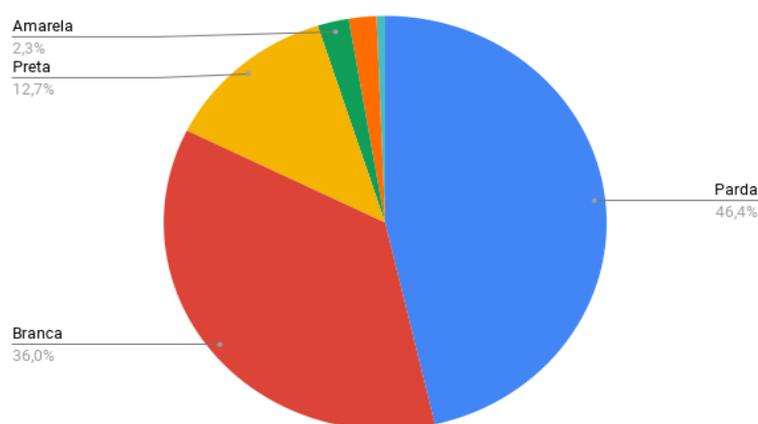


Fonte: Ministério da Educação, 2019,
http://portal.mec.gov.br/images/17.01.2020_Apresentao_Enem_resultados.pdf

Para caracterizar os participantes do exame foram observados os seguintes caracteres: a) cor e raça; b) tipo de escola que frequentou; c) opção de língua estrangeira; d) da situação de conclusão do ensino; e) idade dentre outras.

De acordo com a Figura 11, fica evidente que a maioria dos participantes 46,4% declarou possuir a cor parda. Em seguida, tem-se a cor branca com 36% e a cor negra representa 12,7% dos participantes.

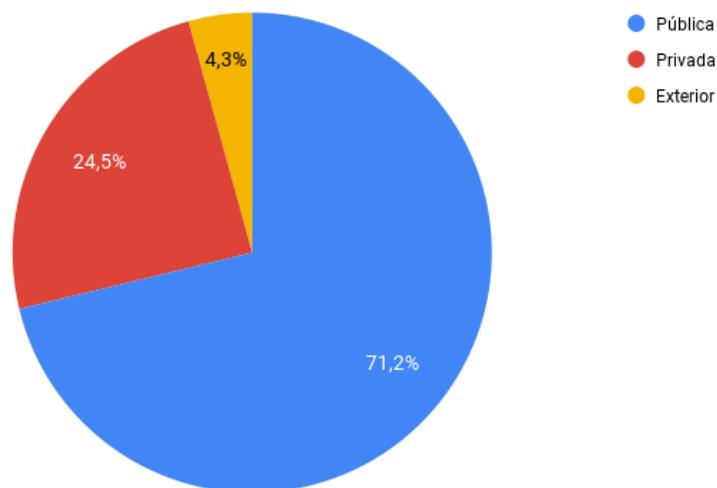
Figura 11-Porcentagem de participantes do ENEM 2019 por cor/raça.



Fonte: Dados da pesquisa.

Mais de 3,5 milhões de inscritos, o que representa 74% da quantidade de alunos, são alunos provenientes de escolas públicas e 26% alunos de escolas privadas (Figura 12).

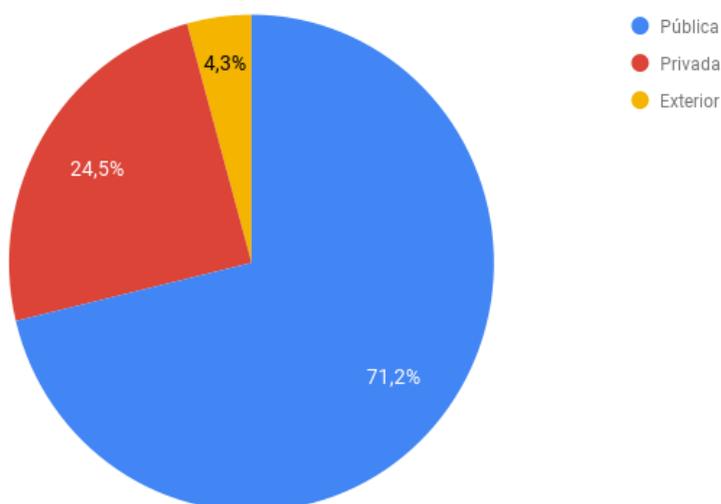
Figura 12- Percentual de alunos por tipo de escola.



Fonte: Dados da pesquisa.

O tipo de escola que o aluno frequentou no ensino médio também representa um fator relevante de ser analisado, pela influência que apresenta no resultado do desempenho do aluno, ao verificar-se na figura 12 tem-se que a maior parte dos inscritos é proveniente de escola pública num total de 3.629.376, mas que as escolas particulares tiveram um quantitativo de 1.247.262 e apenas 218.632 informaram que estudaram no exterior (Figura 13).

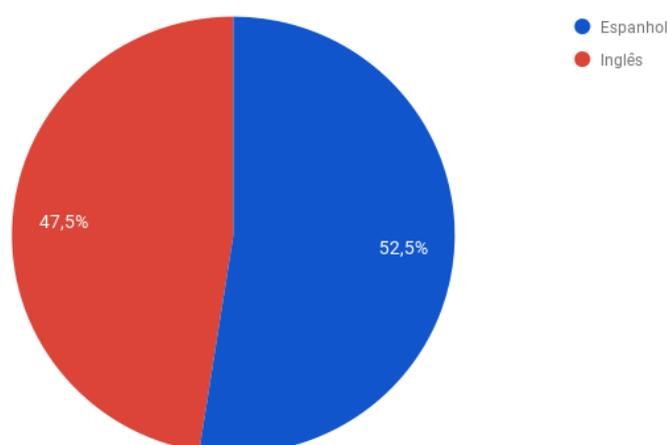
Figura 13- Representação do tipo de escola do inscrito no ENEM 2019.



Fonte: Dados da pesquisa.

Ao se inscrever no ENEM o aluno necessita escolher uma língua estrangeira que apresente maior domínio, mas de acordo com a Figura 14, comprova-se uma tendência de escolha do espanhol apesar do inglês ser bastante utilizado, apresentando uma pequena diferença entre as duas opções.

Figura 14- Representação da opção de língua estrangeira pelos candidatos do ENEM.

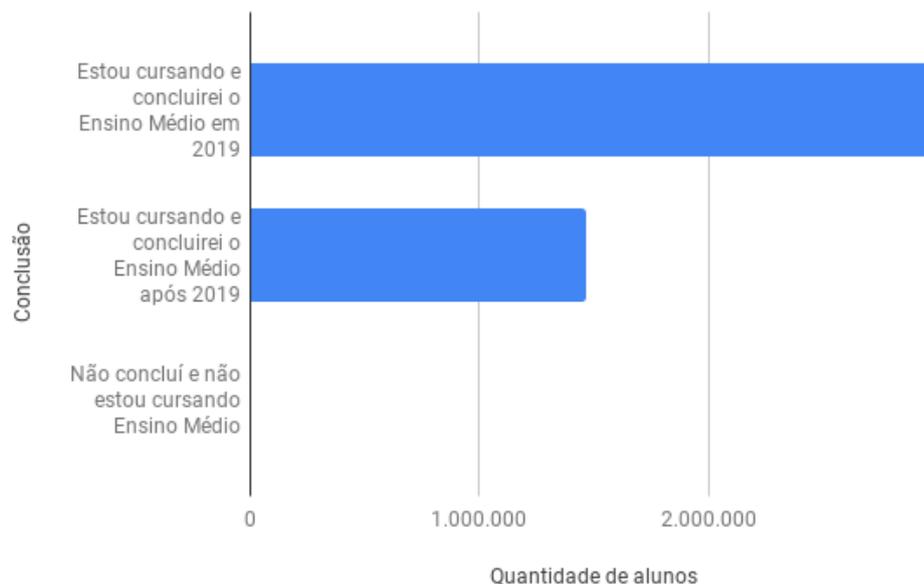


Fonte: Dados da pesquisa.

A Figura 15 mostra os dados dos estudantes de acordo com o questionário respondido no ato da inscrição do ENEM que estão cursando o ensino médio e concluirá no ano de 2019 ou após 2019, como também não concluir e não estar cursando o ensino médio no ano de 2019.

Com relação à situação dos candidatos do ENEM 2019 se verifica que a maioria, (67%), está cursando o ensino médio e concluirá no mesmo ano, os demais (33%) irão concluir após 2019 (Figura 15). Verifica-se, assim, que cerca de 3 milhões de estudantes estavam aptos a utilizar suas notas para ingressar no Ensino Superior brasileiro.

Figura 15- Situação dos estudantes de ensino médio inscritos no ENEM de 2019.

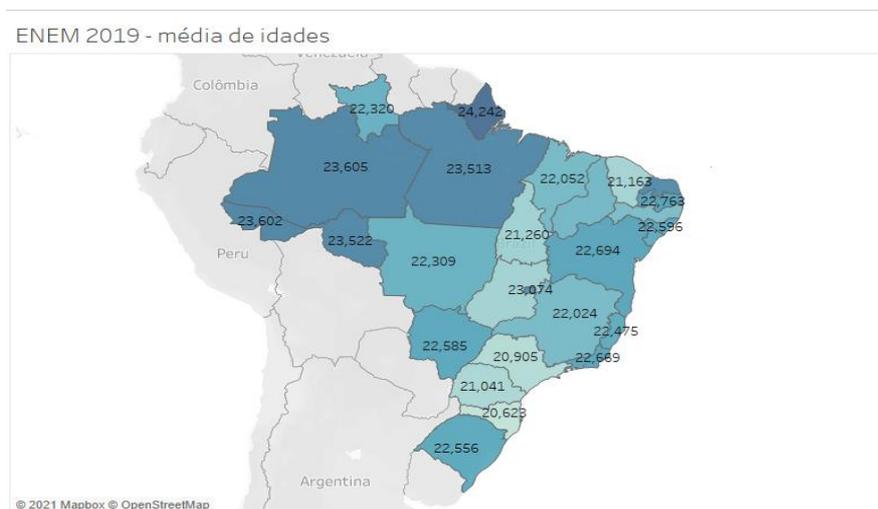


Fonte: Dados da pesquisa.

Com relação à média de idades dos participantes do ENEM 2019 (Figura16), é observado que, dentre as Regiões Geográficas Nacionais, a Sul foi a que apresentou a menor média (21,407 anos, aproximadamente) para os que se inscreveram no certame (ENEM 2019). O estado de Santa Catarina, dentre todos, foi o de menor idade média (20,6 anos), mostrando, assim, que teve os candidatos mais jovens, em média, do Brasil a concorrer às vagas nas Universidades brasileiras. Por outro lado, os candidatos mais velhos, também em média (24,2 anos), foram os do Amapá.

A Figura 16 mostra como está distribuído a média de idade dos participantes do ENEM sendo que a maioria dos estados obtiveram uma média de 22 anos dos participantes inscritos.

Figura 16- Mapa que demonstra a média das idades dos participantes.



Fonte: Dados da pesquisa.

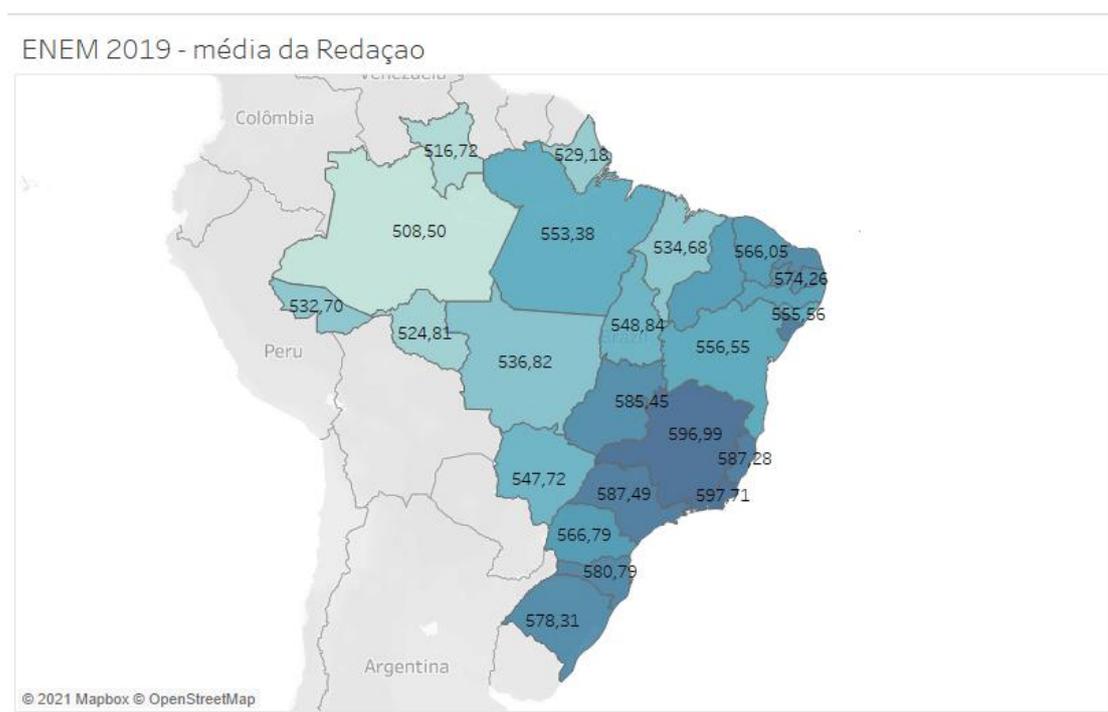
5.2 DESEMPENHO DOS CANDIDATOS DO ENEM 2019

Observando os aspectos do desempenho dos participantes do ENEM 2019, analisou-se como foi o rendimento dos participantes com relação à redação, matemática e ciências da natureza.

Na Figura 17, tem-se que os estados que apresentam uma cor mais escura são aqueles que obtiveram maiores médias com relação a nota da redação, ou seja, são aqueles nos quais uma grande quantidade de participantes obteve bons resultados. Na região sudeste foi o estado do Rio de Janeiro que se destacou obtendo uma média de 598 pontos aproximadamente. Já na região norte temos o estado do Amazonas que obteve a menor média obtendo 508 pontos.

De acordo com BERMÚDEZ (2020), o ENEM 2019 teve um total de 53 redações nota 1000 sendo que a maior parte dos candidatos são do estado de Minas Gerais e mulheres, no qual das 13 redações com nota máxima, 10 foram escritas por mulheres. Esse resultado é muito expressivo por demonstrar o quanto as mulheres estão se preparando e se esforçando para realizar o ENEM, em busca de conseguir uma vaga no Ensino Superior, e expressa a capacidade do público feminino de alcançar bons resultados num exame cujo objetivo é dar oportunidades àqueles que mais se destacaram.

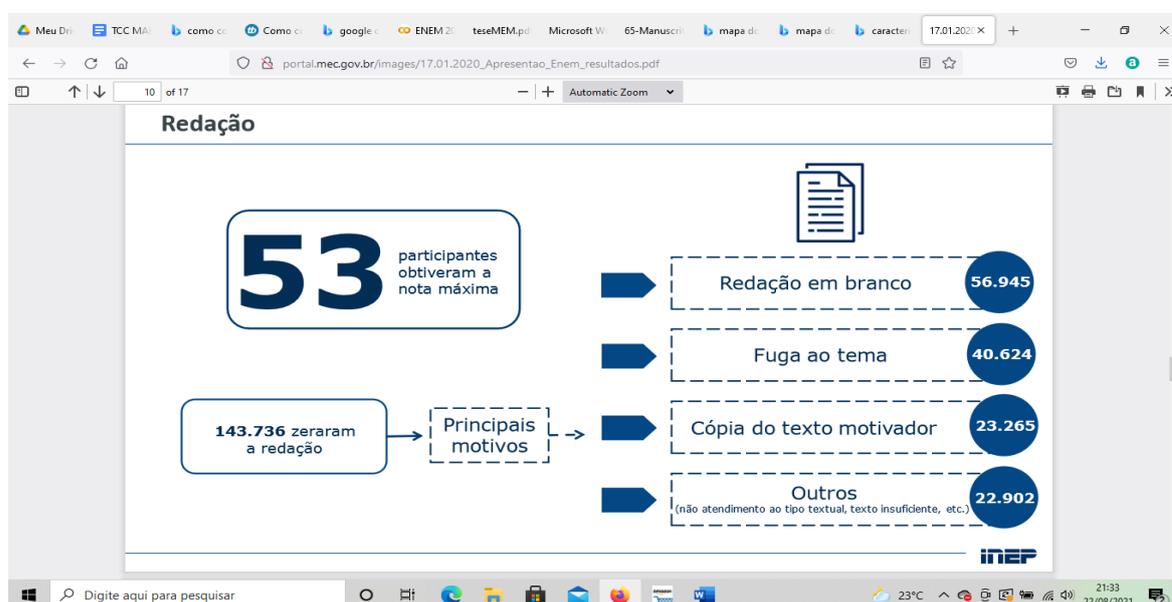
Figura 17- Mapa do Brasil representando a nota média da redação por estado.



Fonte: Dados da pesquisa.

Na figura abaixo observa-se informações referentes à prova de redação e a quantidade de participantes que não obtiveram êxito nesta parte tão importante do exame. É de conhecimento de todos que a nota da redação tem influenciado de forma significativa no desempenho do estudante, tendo como uma das principais causas desse problema a falta de acesso a uma educação de qualidade que os possibilite realizar o certame com a mesma proficiência daqueles que têm acesso a escolas particulares.

Figura 18- Infográfico sobre a prova da redação do ENEM 2019.



Fonte: Dados do MEC.

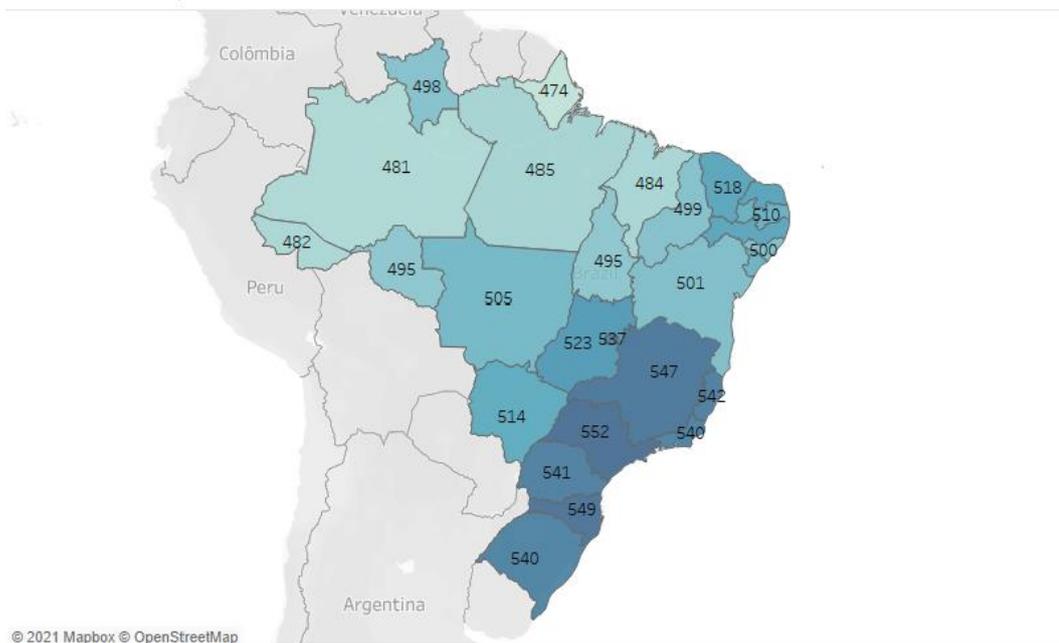
http://portal.mec.gov.br/images/17.01.2020_Apresentao_Enem_resultados.pdf

Ao analisar-se a Figura 19 que o estado de São Paulo foi o que obteve as maiores notas da prova de matemática com média de 552 pontos, já o estado com menor média foi o estado do Amapá. No mapa tem-se que os estados que apresentam uma cor azul mais escura foram os que obtiveram maiores médias e a maioria se concentra na região sudeste do país.

Vale salientar que a média é obtida através da soma dos resultados e dividida pelo número de participantes.

Figura 19- Mapa do Brasil representando a média da nota da prova de matemática por estado.

ENEM 2019 - prova de matemática

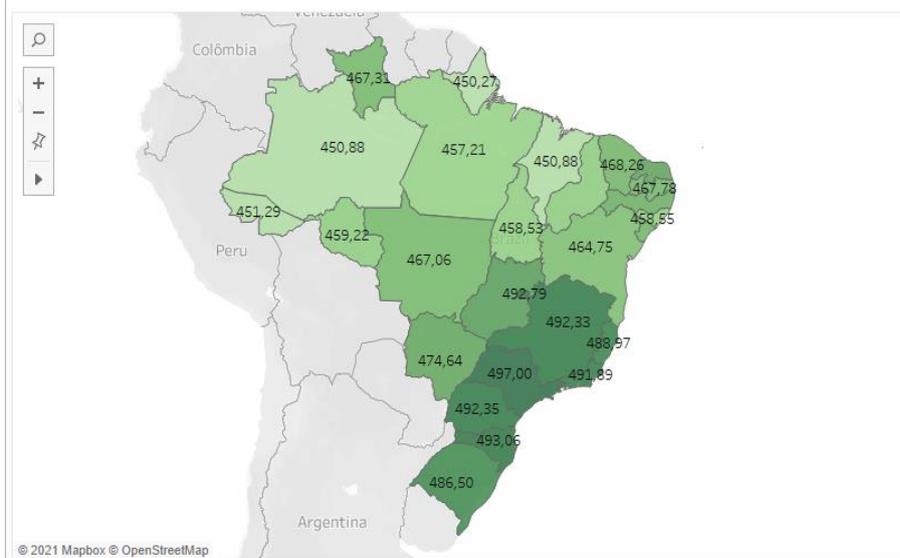


Fonte: Dados da pesquisa.

Já na Figura 20 deteve-se a analisar as médias das notas obtidas pelos participantes do ENEM 2019 em Ciências da natureza, novamente vemos que a região sudeste se destaca no mapa sendo que estado de São Paulo foi o que obteve maior média (497) já o estado de Amapá foi a que teve menor nota (450,27).

Figura 20- Mapa do Brasil representando as médias das notas dos participantes em Ciências da Natureza.

ENEM 2019 - média das notas Ciências da Natureza



Fonte: Dados da pesquisa.

É possível observar na figura acima que a parte com a cor mais escura exibe os estados que se destacaram nessa parte do exame demonstrando assim a necessidade dos estados que apresentaram médias menores auxiliar os estudantes nessa área de Ciência da Natureza, que é tão importante quanto as outras e que influencia no resultado final.

6 CONCLUSÃO

Conclui-se com esse trabalho a importância do ENEM para os estudantes do ensino médio e ingresso nas Universidades brasileiras, devido ao enfoque, a abrangência, o caráter e a inovação do Exame, bem como, para o mercado de trabalho.

Conclui-se, também, de acordo com o contexto socioeconômico dos participantes do certame, que a maioria dos estudantes inscritos no ENEM 2019, é proveniente de escolas públicas. Isso reflete na quantidade de solicitação da taxa de isenção de inscrição para as provas.

Conjectura-se, que de acordo com os que solicitaram atendimentos específico-especializados no ENEM 2019, que os nordestinos são os que mais se utilizam das provas específicas na modalidade braile para os candidatos com deficiência visual, por outro lado, a região sudeste é a que mais se utiliza dos leitores e transcritores para os que possuem autismo, dislexia, discalculia e déficit de atenção para realizar a Exame Nacional do Ensino Médio.

Após a análise das notas do ENEM 2019 por estado verificou-se que com relação às médias por estado obtidas na prova de Matemática e Ciências da Natureza tem-se que a região sudeste foi a que apresentou a maior quantidade de candidatos que obtiveram bons resultados (São Paulo- média de 552 pontos em Matemática e 497 pontos em CN) já na região norte o estado de Amapá foi o que apresentou as menores notas ficando com média 474 pontos na prova de Matemática e 450,27 pontos em CN.

Constata-se que no ENEM 2019 apesar do estado de Minas Gerais ter atingido a maior quantidade de notas máximas na redação foi o estado do Rio de Janeiro que obteve maior média (598 pontos), já as menores notas ocorreram na região Norte, com destaque para o estado do Amazonas que ficou com uma média de 508 pontos.

Considerando que o ENEM no contexto atual, como acreditam alguns autores representa o substituto dos antigos vestibulares, que na maioria das vezes, ainda apresentava a necessidade de o estudante decorar o conteúdo para poder realizar com êxito o exame. Verifica-se também que apesar do ENEM apresentar esse caráter mais multidisciplinar e que tem como objetivo avaliar o nível dos estudantes, que necessitam resolver questões utilizando o raciocínio e a interpretação que são necessários para resolver as questões.

É notório que as transformações que ocorreram com o ENEM ao longo dos últimos anos trazem para a comunidade estudantil uma nova forma de percepção e de necessidade de se dedicar mais para poder alcançar seus objetivos. Dentre elas está o fato do exame ser dividido em dois domingos, o aumento da quantidade de questões e também na forma de avaliar o desempenho dos estudantes.

Com relação a redução da quantidade de inscritos e participantes do exame pode-se compreender que ao longo dos anos vem ocorrendo uma desvalorização e falta de incentivo para aqueles que desejam adentrar ao Ensino Superior a fim de se formar e ter uma profissão que possa lhe dar condições de viver melhor.

Como sugestões para trabalhos futuros poder-se-ia investigar os principais fatores que mais influenciam na desmotivação dos jovens de se preparar logo após sair do ensino médio, e como a escola poderia intervir nesse processo buscando despertar nos estudantes esse interesse e necessidade de ampliar seus conhecimentos.

REFERÊNCIAS

B. Santos, C. G. Oliveira, L. O. H. Topin, O. M. Mendizabal, and R. Barwaldt. **Analysis of Candidates Profile for the National Entrance Exams for Admission to Brazilian Universities**. Proceedings - Frontiers in Education Conference, FIE, 2019-October: 1–8, 2019. ISSN 15394565. doi: 10.1109/FIE43999.2019.9028381.

BERMÚDEZ, Ana Carla. **Enem 2019:53 candidatos tiraram nota mil na redação; 143 mil tiraram nota zero**. Do UOL, em São Paulo, 17 de janeiro de 2020. Disponível em: <<https://educacao.uol.com.br/noticias/2020/01/17/enem-2019-53-candidatos-tiraram-nota-mil-na-redacao-143-mil-tiraram-zero.htm>>. Acesso em: 04 de março de 2021.

BRAGA, Andrei S.; SILVA, Geraldo M.; BARROS, Marcos C. Cloud computing. **Instituto de Computação-Universidade Estadual de Campinas**, p. 18, 2012.

BRASIL. Lei nº 11.096, de 13 de janeiro de 2005. **Institui o Programa universidade para Todos -PROUNI**.

BRASIL. Medida provisória Nº 213, DE 10 DE SETEMBRO DE 2004. **Institui o Programa universidade para Todos- PROUNI**.

BRASIL. **Constituição da República Federativa do Brasil**.1998.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. **Relatório Pedagógico do Enem 2008**. Brasília, DF, 2008b.

CAIRO, A. **El arte funcional: infografía y visualización de información**.Madrid: Alamu, 2012.

CAMPOS,Lorraine Vilela. **A prova do Enem na prática**.Brasil Escola, 06 de março de 2018. Disponível em:<<https://vestibular.brasilescola.uol.com.br/enem/a-prova-enem-na-pratica.htm>>Acesso em 23 de agosto de 2020.

CARNEIRO, T., Da Nobrega, R. V. M. Nepomuceno, T., Bian, G. B., DeAlbuquerque, V. H. C., and Filho, P. P. R. (2018). **Performance Analysis of GoogleColaboratory as a Tool for Accelerating Deep Learning Applications**.IEEE Access 12 de março de 2020.

CARNEIRO, Verônica Lima. **As avaliações estandardizadas e o papel do exame nacional do ensino médio (ENEM) na etapa final da educação básica**. Revista Exitus, Vol 02, nº01, Universidade do Oeste do Pará, Santarém/PA. Jan./Jun. 2012.

CURI, A. Z. MENEZES FILHO, N. A. **Mensalidade escolar, background familiar e os resultados do Exame Nacional do Ensino Médio (ENEM)**. Pesquisa e Planejamento Econômico, v. 43, n.2, p. 223–254, 2013. Disponível em: <<http://ppe.ipea.gov.br/index.php/ppe/article/view/1469/1131>>.

DA SILVA, MartonyDemes. **Aplicação da Ferramenta Google Colaboratory para o Ensino da Linguagem Python**. In: **ESCOLA REGIONAL DE ENGENHARIA DE SOFTWARE (ERES)**, 4., 2020, Evento Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2020. p. 67-76. DOI: <https://doi.org/10.5753/eres.2020.13717>.

em: 16/08/2019.

FIGUEIRÊDO, E.; NOGUEIRA, L.; SANTANA, F. L. **Igualdade de oportunidades: Analisando o papel das circunstâncias no desempenho do ENEM**. Revista Brasileira de Economia, v. 68, n. 3, p. 373–392, 2014.

FOLHA DE SÃO PAULO. **A loteria Enem**. São Paulo, 2012. Disponível em: <<http://www.andifes.org.br/index.php?option=comcontent&view=article&id=6155:a-loteria-enem-&catid=508Itemid=100017>>. Acesso em: 20 de outubro de 2020.

GONÇALVES JUNIOR, Wanderley; BARROSO, Marta Feijó. **As questões de física e o desempenho dos estudantes no ENEM**. Revista Brasileira de Ensino de Física, São Paulo, SP, v. 36, n. 1, p. 1-16. jan./mar. 2014. Disponível em: . Acesso em: 08 jul. 2015. ISSN 1806-1117.

Guido van Rossum. **Python reference manual**. Technical report, Amsterdam,

HEY, A. P. Notícias da UFMG: **Professora da USP questiona Enem como forma de democratização do acesso à universidade**. In: Semana de conhecimento e cultura da UFMG. 2009. Disponível em: < www.ufmg.br/online/arquivos/013510.shtml >. Acesso em: 28 de outubro de 2020.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais. Anísio Teixeira. Ministério da Educação. Disponível em: <<http://www.enem.inep.gov.br/conheca-o-inep>>.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatísticas do Exame Nacional de Ensino Médio 2018**. Brasília: Inep, 2019. Disponível em: <<http://portal.inep.gov.br/xxxxxx>>. Acesso em: 06 de julho de 2021.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2019**. Brasília: Inep, 2019. Disponível em: Acesso em: 08 de ago 2021.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatísticas do Exame Nacional de Ensino Médio 2018**. Brasília: Inep, 2019. Disponível em: <<http://portal.inep.gov.br/xxxxxx>>. Acesso em: 26 jul. 2020.

JUNQUEIRA, Rogério Diniz; LACERDA, Cristina Broglia Feitosa de. **Avaliação de estudantes surdos e deficientes auditivos sob um novo paradigma: Enem em Libras**. Revista Educação Especial, Santa Maria, p. 1-17, mar. 2019. Disponível em: <https://periodicos.ufsm.br/educacaoespecial/article/view/28732>. Acesso em: 26 dez. 2019.

LEMOS, A. **Cibercultura, tecnologia e vida social na cultura contemporânea**. Porto Alegre: Sulina, 2002.

M. F. C. de Paula. **Políticas de democratização da educação superior brasileira: limites e desafios para a próxima década**. Avaliação: Revista da Avaliação da Educação Superior (Campinas), 22(2):301–315, 2017. ISSN 1414-4077. doi: 10.1590/s1414-40772017000200002.

MANOVICH, L. **Trending: The Promises and the Challenges of Big Social Data**. Disponível em: <http://goo.gl/IqIlgGF> Acesso em: 3 jun. 2015.

Ministério da Educação. Secretaria de Educação Média e Tecnológica. **Parâmetros nas avaliações em larga escala: prova Brasil e ENEM**. Revista Brasileira de Educação Especial, Marília, v. 20, p. 53-68, Março 2014.

NASCIMENTO, Anarcisa de Freitas. **Aprimoramento do atendimento especializado para pessoas com Transtorno do Espectro Autista na redação do Enem**. 2020.

OLIVEIRA PEREIRA, Poliana Francibele de et al. Uma Análise Temporal de Dados Abertos do Ensino Superior utilizando o Software de Visualização Tableau. **Anais do Simpósio Ibero-Americano de Tecnologias Educacionais**, [S.l.], p. 192-198, June 2018. ISSN 2594-388X. Disponível em: <<https://publicacoes.rexlab.ufsc.br/old/index.php/sited/article/view/292>>. Acesso em: 24 June 2021.

PASSOS, Daniele Sandles. Big Data, **Data Science e seus contributos para o avanço no uso do open source intelligence**. **Sistema & Gestão**. Revista Eletrônica. Volume 11, Número 4, pp.392-396, 2016, Universidade Nova de Lisboa.

SANTOS, J. M. C. T. **Exame Nacional do Ensino Médio: entre a regulação da qualidade do ensino médio e o vestibular**. Educar em Revista. n. 40, p. 195-205, abr./jun. 2011.

Silva, M. C. V., & Meletti, S. M. F. (2014). **Estudantes com necessidades educacionais especiais nas avaliações em larga escala: Prova brasil e enem.** *Revista Brasileira de Educaçao Especial*, 20(1), 53–67. <https://doi.org/10.1590/S1413-65382014000100005>.

SILVA, Martony Demes da. **Aplicação da Ferramenta Google Colaboratory no Ensino de Ciências de Dados.***In: SIMPÓSIO BRASILEIRO DE SISTEMAS COLABORATIVOS (SBSC), 17., 2021, Evento Online.***Anais** [...] Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 13-22. ISSN 2326-2842.

TODOS – PROUNI. Diário Oficial da União, Brasília, 14 jan. 2005a. p.7.

TRAINA, Agma JM et al. **Visualização de Dados em Sistemas de Bases de Dados Relacionais.** In: **SBBD.** 2001. p. 95-109.

TRAVITZKI, Rodrigo. **Enem: Limites e possibilidades do exame nacional do ensino médio enquanto indicador de qualidade escolar.** Faculdade de Educação da Universidade de São Paulo. São Paulo. 2013.

Universidade para Todos - PROUNI, regula a atuação de entidades beneficentes de assistência social no ensino superior. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Mpv/213.htm> Acesso em: 14 ago. 2021.

VISHAKHALALL (2018). **Google colab — the beginner’s guide.** Acesso em fev. 2021.