



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS CORRENTE
CURSO: ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

CELSO BADU GOMES

**UM PANORAMA DOS RESULTADOS DO ENEM 2022 FRENTE A PANDEMIA DE
COVID-19**

CORRENTE - PI
2023

CELSO BADU GOMES

UM PANORAMA DOS RESULTADOS DO ENEM 2022 FRENTE A PANDEMIA DE
COVID-19

Trabalho de Conclusão de Curso (artigo científico)
apresentado como exigência parcial para obtenção
do diploma do Curso de Análise e Desenvolvimento
de Sistemas do Instituto Federal de Educação,
Ciência e Tecnologia do Piauí, Campus Corrente.

Orientador: Prof. Me. Felipe Gonçalves dos Santos.
Coorientador: Prof. Me. Rafael Emanuel
Costa

Aprovada em: ____/____/2023.

BANCA EXAMINADORA

Prof. Me. Felipe Gonçalves dos Santos (Orientador)
Instituto Federal do Piauí (IFPI)

Prof. Me. Rafael Emanuel Costa (Coorientador)
Instituto Federal do Piauí (IFPI)

Prof. Me. Juan Morysson Viana Marciano
Instituto Federal do Piauí (IFPI)

Prof. Me. Jurandir Cavalcante Lacerda Júnior
Instituto Federal do Piauí (IFPI)

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Gomes, Celso Badu

G633p Um panorama dos resultados do Enem 2022 frente a pandemia de
Covid - 19 / Celso Badu Gomes. - 2024.
39 p.: il. color.

Trabalho de Conclusão de Curso (Análise e Desenvolvimento de
Sistemas) - Instituto Federal de Educação, Ciência e Tecnologia do
Piauí, Campus Corrente, 2024.

Orientador : Prof Me. Felipe Gonçalves dos Santos.

Coorientador : Prof Me. Rafael Emanuel Costa.

1. Análise e Desenvolvimento de Sistemas - estudo e ensino. 2.
Enem. 3. Pandemia - Covid 19. 4. Inep. I.Título.

CDD - 004

Elaborado por Tefischer Huanderson Soares e Sousa CRB 3/1251

UM PANORAMA DOS RESULTADOS DO ENEM 2022 FRENTE A PANDEMIA DE COVID-19

Celso Badu Gomes¹

Felipe Gonçalves dos Santos²

Rafael Emanuel Costa³

RESUMO

O entendimento sobre os aspectos educacionais de um país é fundamental para estabelecer metas de melhorias no ensino. No Brasil, a entidade responsável por promover estudos, pesquisas e avaliações no âmbito da educação é o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Inep, com o objetivo, dentre outros, de analisar a educação básica do país. Com o iminente impacto causado pela pandemia de Covid-19, se fez necessária, no âmbito educacional, uma pesquisa que buscasse analisar e compreender os impactos da pandemia na educação brasileira. Diante disso, através da análise exploratória de dados do Exame Nacional do Ensino Médio, Enem, do ano de 2019 e 2022, foi possível diagnosticar, de modo geral, que as médias do Enem de 2022 não sofreram tanta alteração com relação a 2019, ao contrário, até melhoraram um pouco, o que mudou realmente foi a quantidade de inscritos, que caíram drasticamente, a quantidade de ausentes no exame e os que zeraram a redação. O fato novo foi o questionário de hábitos de estudos durante a pandemia, onde apenas 37% dos candidatos responderam. No entanto, apesar do momento difícil em questão, a maioria dos estudantes não abandonaram a escola, estudando de forma híbrida, mas que necessitam muito de ajuda de terceiros, e sentiram falta das aulas no modo presencial. Além disso, os relatos sobre as faltas de suporte tecnológicos, como equipamentos e acesso à internet foram numerosos. No entanto, isso não afetou as médias das notas desses candidatos, apesar da maioria relatar sentir-se pouco preparado no processo educacional.

Palavras chaves: Tecnologias; Enem; Pandemia; Educação.

¹ Concluinte do curso de análise e desenvolvimento de sistemas do Instituto Federal do Piauí, campus Corrente; Licenciado em Pedagogia pela Universidade Estadual do Piauí-UESPI, 2016; licenciado em História pela Faculdade Internacional do Delta do Parnaíba, 2015 e Pós-graduado pelo Núcleo de Educação a Distância da Universidade estadual do Piauí- NEAD-UESPI em gestão educação ambiental, 2017.

² Professor mestre do Instituto Federal do Piauí, campus Corrente

³ Professor mestre do Instituto Federal do Piauí, campus Corrente

ABSTRACT

The understanding of a country's educational aspects is crucial for establishing goals for improving education. In Brazil, the entity responsible for promoting studies, research, and evaluations in the field of education is the National Institute for Educational Studies and Research Anísio Teixeira, INEP, with the goal, among others, of analyzing the country's basic education. With the imminent impact caused by the Covid-19 pandemic, a necessary educational survey was conducted to analyze and understand the impacts of the pandemic on Brazilian education. In light of this, through exploratory data analysis of the National High School Exam, ENEM, for the years 2019 and 2022, it was possible to diagnose, in general, that the ENEM averages for 2022 did not undergo much change compared to 2019; on the contrary, they even improved slightly. What really changed was the number of registrants, which dropped drastically, the number of absentees in the exam, and those who scored zero in the essay. The new factor was the study habits questionnaire during the pandemic, where only 37% of candidates responded. However, despite the difficult situation, most students did not drop out of school, studying in a hybrid manner but requiring significant assistance from others. They also missed in-person classes. Additionally, reports about the lack of technological support, such as equipment and internet access, were numerous. Nevertheless, this did not affect the average scores of these candidates, although the majority reported feeling somewhat unprepared in the educational process.

Keywords: Technologies; Enem; Pandemic; Education.

SUMÁRIO

1 INTRODUÇÃO.....	6
2 REFERENCIAL TEÓRICO.....	7
2.1 EXAME NACIONAL DO ENSINO MÉDIO.....	7
2.2 CIÊNCIAS DE DADOS.....	8
2.3 NOÇÕES DE ESTATÍSTICA.....	9
2.3.1 Amostragem estatística.....	9
2.3.2 Estatística descritiva.....	10
2.3.3 Distribuição.....	10
2.4 TRABALHOS RELACIONADOS.....	12
4 METODOLOGIA.....	15
4.1 FONTES E TIPOS DE AMOSTRAS.....	15
4.2 PROCEDIMENTO DE COLETA DE DADOS.....	15
4.3 CARACTERIZAÇÃO DAS AMOSTRAS E ESTUDAS DAS ANÁLISES DE DADOS.....	16
4.4 CRITERIOS DE SELEÇÃO DE DADOS.....	17
4.5 ARQUITETURA DAS FERRAMENTAS.....	17
4.6 TECNOLOGIAS UTILIZADOS.....	18
4.6.1 Tratamentos dos dados.....	19
5 RESULTADOS E DISCUSSÃO.....	20
5.1 QUESTIONÁRIO DE HÁBITOS DE ESTUDOS DURANTE A PANDEMIA....	24
REFERÊNCIAS.....	37

1 INTRODUÇÃO

Na lei de Diretrizes e Bases (LDB – Lei nº 9394 de 20 de dezembro de 1996) em seu artigo 2º retrata a educação como um dever do estado e direito de todo cidadão inspirados no princípio da liberdade e solidariedade humana, no qual tem por finalidade o pleno desenvolvimento do educando e o preparo para o exercício da cidadania. Na mesma lei, em seu artigo 35º, na seção que trata do ensino médio, define essa etapa como a parte final do ciclo da educação básica, com duração mínima de três anos, que tem como finalidades “[...] O aprimoramento do educando como pessoa humana, incluindo a formação ética e o desenvolvimento da autonomia intelectual e do pensamento crítico. ”

No mesmo sentido, a Constituição Federal de 1988 em seu artigo 206º, dentre outros, possui a garantia de um padrão de qualidade que deve ser oferecido pelo sistema público de ensino no Brasil. Necessitando assim de medidas avaliativas que irão medir esse padrão de qualidade. Nesse sentido, em 1988, foi instituído pelo Ministério da Educação – MEC, o Exame Nacional do Ensino Médio - Enem, que buscava comparar o desempenho dos estudantes ao término do ensino médio. Em 2009 foi aperfeiçoado e passou a ser utilizado como mecanismo de acesso à educação superior no Brasil, (INEP, 2009).

O Enem é uma avaliação que mede o desempenho dos estudantes ao término da educação básica, e serve como porta de acesso à educação superior no Brasil e em Portugal. São mais de 85 milhões de candidatos disputando uma vaga em universidades públicas e privadas de todo o país até 2022. A forma como as universidades brasileiras oferecem as vagas podem variar de acordo com a instituição, no entanto, por todo o Brasil, são mais de 60 universidades de todos os estados que aceitam as notas do Enem como ingresso em algumas das centenas dos cursos ofertados pelas mesmas, através do SisU, que é um Sistema de Seleção Unificada que utiliza as notas do Enem como parâmetro para aprovação dos candidatos à vagas universitárias brasileiras, (BRASIL, 2023).

As vagas oferecidas através do SisU são exclusivamente dedicadas às universidades públicas federais e estaduais, para as universidades particulares existem outros meios disponíveis que, também, podem se utilizar das notas do Enem como acesso para os estudantes. No final de 2019 esses estudantes que buscavam por uma dessas vagas através das notas do Enem tiveram que enfrentar outro problema, a pandemia de Covid-19.

Segundo um levantamento feito por um site de notícias, em março de 2020 foram detectados os primeiros casos de Covid-19, e pouco tempo depois começaram os decretos de isolamentos sociais no Brasil e no mundo, deixando apenas serviços essenciais em funcionamento. E tendo em vistas tais mudanças, principalmente na educação, muitas soluções foram propostas para que as escolas continuassem a manter o seu processo educacional, sejam por aulas remotas ou por outros meios que pudesse garantir o ensino e aprendizagem dos alunos durante a pandemia.

O método que pode ter sido mais utilizado foi a aula remota, onde tecnologias e ambientes digitais foram utilizados. Esse método, segundo Alqahtani e Rajkhan (2020), é conhecido no mundo todo e é bastante explorado com o avanço dos ambientes de ensino online. Embora esse tipo de educação já tenha sido disponível antes da pandemia, esse é um tipo de ensino que ainda era incomum nas instituições de ensino brasileiras nos quais as aulas eram predominantemente de modo presencial, Gularte, Nascimento, Carvalho (2021).

Diante disso, o Inep, órgão responsável pela organização do exame, disponibiliza uma base em seu site onde são reunidas todas as informações detalhadas relacionados ao exame, sendo uma avaliação realizada todos os anos, desde a sua criação, em vários municípios brasileiros. Porém, com a chegada da pandemia em 2020, houve uma mudança na aplicação do exame, pois antes eram aplicadas sempre nos meses de outubro e novembro, e por questão de segurança, o exame referente a 2020, foi aplicado em janeiro de 2021, sendo importante aqui salientar a inclusão do Enem digital, modelo opcional para o aluno de fazer as mesmas avaliações de modo digital.

Diante de todo o disposto e com o impacto da pandemia, fez-se necessário um estudo no qual possamos compreender os impactos na educação básica e como o principal exame de desempenho do ensino médio brasileiro pode ter sido afetado, principalmente nos hábitos educacionais dos estudantes e suas respectivas notas no exame. E para identificarmos esses aspectos, utilizamos técnicas de mineração de dados, ou análise de dados, que segundo Neto, Soares, Coutinho e Teles (2022) é um processo computacional aplicado em diversas áreas visando obter informações e gerar conhecimento.

Os objetos desse estudo são utilizar os microdados do Enem de 2019 e 2022 disponibilizados pelo Inep para comparar os resultados obtidos através da análise de dados. Além de buscar responder ao questionamento: Houve uma discrepância nas notas gerais dos estudantes no período de pandemia? Também busca responder, através do questionário feito pelo Inep na hora da inscrição, e disponibilizado no site do Inep, qual o hábito dos candidatos, com relação aos estudos no período pandêmico. Isso fez diferença nas notas obtidas, de acordo com as respostas? E para responder a todos esses questionamentos foi feito o levantamento dos dados, um processamento a priori, e uma análise exploratória comparativa.

2 REFERENCIAL TEÓRICO

2.1 EXAME NACIONAL DO ENSINO MÉDIO

O Enem recebe milhões de inscrições pela internet todos os anos, por se tratar da porta de entrada para a educação superior a muitos estudantes e produto de avaliação de desempenho da educação básica, é no ato da inscrição do candidato que os participantes devem preencher um formulário socioeconômico, onde contém informações sobre faixa etária de idade, renda familiar, sexo, cor, raça, tipo de escolar concluinte do ensino médio e outras.

Em 2022, como praticamente todas as escolas funcionaram de maneira remota ou híbrida, devido ao surto pandêmico, o Inep, além dos questionários habituais, disponibilizou também um questionário de hábitos educacionais, e é com base nessas bases de dados disponibilizados pelo Inep que foi realizado o estudo deste trabalho.

2.2 CIÊNCIAS DE DADOS

As pesquisas que trabalham no campo tecnológico ou em outras áreas relacionadas à ciência de dados estão sujeitas a métodos científicos, matemáticos, estatísticos, que podem ser usadas para analisar e manipular dados. Se é uma ferramenta ou processo feito para análise de dados, para obter algum tipo de

informação, provavelmente a pesquisa se enquadra na ciência de dados. (CETAX, 2022). Para Aggarwal (2015) a mineração de dados é definida como:

A mineração de dados é o estudo de coleta, limpeza, processamento, análise e obtenção de informações e ideias úteis de dados. Existe uma grande variação em termos de domínios problemáticos, aplicativos, formulações e representações de dados encontradas em aplicativos reais. Portanto, “Mineração de dados” é um termo abrangente usado para descrever esses diferentes aspectos de processamento de dados (AGGARWAL, 2015).

É importante ressaltar que, como as demais áreas da tecnologia, a ciência de dados possui um ciclo de vida nos quais envolve seus projetos, pois ela não segue o mesmo padrão para todos os projetos, assim, cada projeto possui suas necessidades específicas e requer uma adaptação adequada a cada modelo. O rápido avanço computacional tem permitido análises de quantidade de dados cada vez maiores, possibilitando a descoberta de padrões e *insights* até mesmo em tempo real. E nesse sentido, o ciclo de vida da ciência de dados pode ser dividido em cinco partes, Gonçalves (2018).

- **Entendimento do problema**

Essa é uma das etapas mais importantes de todo o ciclo, pois é a partir do conhecimento da problemática que estratégias para alcançar os resultados esperados serão traçados.

- **Coleta de dados**

Uma vez definida a problemática, é preciso extrair e coletar os dados, essa é uma etapa fundamental para entender quais tipos de dados serão coletados para a pauta do projeto. Costa (2011) classifica essa etapa, quanto ao fator e tempo, em três níveis: A continuar, é quando uma análise é feita de forma contínua; A periódica, é quando uma análise é feita de tempos em tempos, como o censo demográfico feito pelo IBGE de 10 em 10 anos e; A ocasional, que é feita apenas em uma determinada ocasião para atender a um objetivo.

- **Processamento dos dados**

Depois da coleta, e hora de começar a analisá-los. Nessa etapa é preciso ficar atento aos dados faltantes, duplicados ou formatados de forma não-convencional, como campos inválidos, inconsistências nos cadastros etc.

- **Exploração dos dados**

Essa é a etapa onde o projeto começa a tomar vida, nessa fase é necessário ter habilidades analíticas e criativas para pensar em ideias e hipóteses a serem validadas. Também é nessa fase onde existem várias ferramentas e bibliotecas em várias linguagens de programação que auxiliam nessa etapa, dos quais serão vistas mais adiante neste trabalho.

- **Comunicação de resultados e Feedback**

É nessa fase que podemos ter a disseminação efetiva dos resultados, podendo assim concluir o estudo efetuado.

Portanto, este trabalho trata do estudo e aplicação de técnicas de ciências de dados nos microdados do Enem de 2019 e 2022, passando pelas etapas de ciências de dados já mencionadas, no qual a fase da coleta de dados será feita no nível ocasional, com ênfase na exploração dos dados.

2.3 NOÇÕES DE ESTATÍSTICA

As aplicações da estatística se desenvolveram de tal forma que, hoje, praticamente todo o campo de estudo se beneficia da utilização de métodos estatísticos. Os fabricantes fornecem melhores produtos a menores valores mediante técnicas de controle de qualidade. Espécies ameaçadas são protegidas por regulamentos e leis que regem as estimativas estatísticas de modificação de tamanho da população. Assim como a educação brasileira, a estatística faz parte do seu cotidiano, para a distribuição de recursos, avaliação do desempenho escolar, indicadores educacionais e outros.

Costa (2011) define a estatística como um ramo da matemática ao estudo dos processos de obtenção, coleta, organização, apresentação, descrição, análises e interpretação de dados numéricos para sua utilização em tomadas de decisões. Já Vickery (2011) afirma que o campo da estatística é muito amplo, porém nem todos os seus conceitos são obrigatórios para desenvolver um estudo com o auxílio da ciência de dados. São apresentados nesta seção, alguns conceitos fundamentais sobre estatísticas que auxiliam na análise de dados deste trabalho.

2.3.1 Amostragem estatística

Uma amostragem, de modo geral, é a coleta das informações de parte da população chamada amostra mediante métodos adequados de seleção dessas unidades, Costa (2011). Já a amostragem é considerada uma técnica especial de escolher amostras, onde cada elemento da população possui a mesma probabilidade de ser escolhido, pois assim, a amostra garante a representatividade da população, no entanto, temos que observar os dados para não ocorrer “abusos estatísticos” com amostras insuficientes, números imprecisos, gráficos enganosos e outros fatores que podem distorcer os resultados obtidos.

Para exemplificar de forma simples esse tipo de amostragem, podemos considerar que queremos analisar a qualidade do ensino médio no Brasil, a população desejada seria de todos os concludentes do ensino médio do país, no entanto, por razões lógicas, não dispomos de todos esses dados. Pode ser utilizada apenas uma amostra que representa toda a população, desde que a amostra possa ter a representatividade dessa população, poderemos realizar inferências sobre a população em sua totalidade e assim chegar a uma conclusão.

2.3.2 Estatística descritiva

Estatística descritiva consiste em um conjunto de técnicas, para planejar, organizar, coletar, resumir, classificar, examinar, explicar, comunicar e analisar dados em tabelas, gráficos ou outros recursos visuais que não sejam de cálculos de estimativa de parâmetros representativos, tendo assim uma interpretação de coeficientes e exposições que permitem descrever um fenômeno. (COSTA, 2011).Na

área da análise estatística descritiva, onde coletamos, organizamos e descrevemos os dados, podemos apenas descrever e analisar o conjunto de dados, sem tirar conclusões, pois esse tipo de estatística apenas nos auxilia a descrever os dados e compreender suas características.

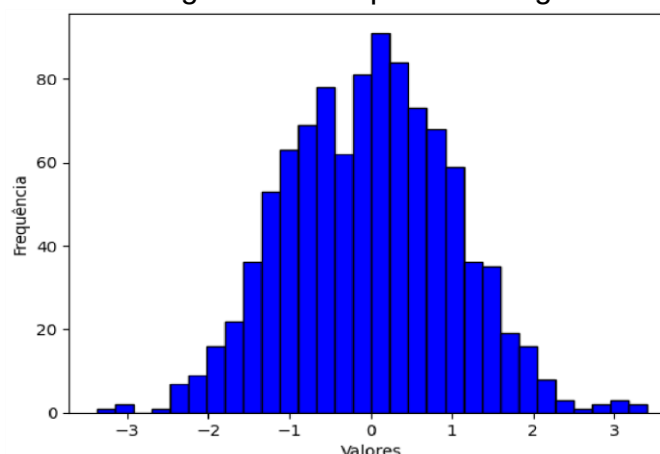
O objetivo dessa etapa é apresentar as descrições da aparência da amostra que possuímos. Geralmente esse modelo de estatística é obtida através de dados com medidas de tendência central, como **média**, que se trata do valor médio dos dados analisados; **mediana**, que se os dados forem ordenados de maneira crescente, esse valor seria a metade do conjunto de dados, caso seja dividido exatamente pela metade, e a **moda**, que é o valor com maior ocorrência no conjunto da amostra.

2.3.3 Distribuição

As estatísticas descritivas, apesar de úteis, podem mascarar informações importantes sobre a amostra estudada. Caso um conjunto de dados contenham valores que sejam muito maiores que outros, a média será distorcida e não poderá ser considerada uma representação fidedigna da análise, (SOUZA, 2021).

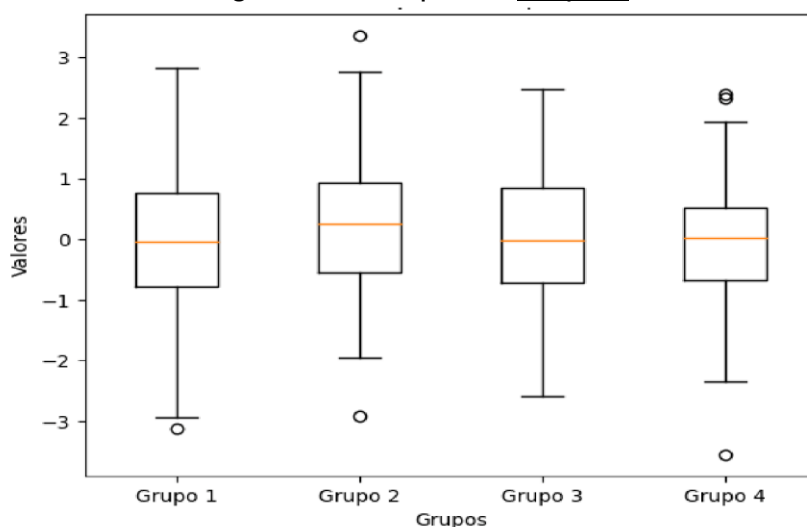
Para representar uma distribuição, pode ser feito o uso de um histograma, gráficos e tabelas. Um histograma é uma espécie de gráfico de barras que demonstra uma distribuição de frequências. No histograma a base de cada uma das barras representa uma classe e a altura representa a quantidade ou frequência absoluta com que o valor de cada classe ocorre, (SIQUEIRA, 2023). A Figura 1 mostra um exemplo de histograma.

Figura 1. Exemplo de histograma.



Fonte: Elaboração própria (2023).

Outro exemplo de gráfico que foi utilizado neste trabalho foi o do tipo *boxplot*, ele é capaz de mostrar a distribuição quantitativa dos dados de um jeito que facilita a comparação entre as variáveis, ou através dos níveis categóricos das variáveis. Para a utilização desse modelo de gráfico, é necessário que entendamos algumas especificidades dele. Essa caixa (“*box*”) mostra os quartis da base de dados enquanto os “*whiskers*” mostram o resto da distribuição, exceto os pontos que são chamados de *outliers*. O modelo de *boxplot* pode ser visualizado na Figura 2.

Figura 2. Exemplo de *boxplot*.

Fonte: Elaboração própria (2023).

O *Boxplot* nos fornece uma análise visual da posição, dispersão, simetria, caudas e valores discrepantes, chamados de *Outliers* do conjunto de dados, onde a:

Posição – Em relação à posição dos dados, observa-se a linha central do retângulo (a mediana ou segundo quartil).

Dispersão – A dispersão dos dados pode ser representada pelo intervalo interquartílico que é a diferença entre o terceiro quartil e o primeiro quartil (tamanho da caixa), ou ainda pela amplitude que é calculada da seguinte maneira: valor máximo – valor mínimo. Embora a amplitude seja de fácil entendimento, o intervalo interquartílico é uma estatística mais robusta para medir variabilidade uma vez que não sofre influência de *outliers*.

Simetria – Um conjunto de dados que tem uma distribuição simétrica, terá a linha da mediana no centro do retângulo. Quando a linha da mediana está próxima ao primeiro quartil, os dados são assimétricos positivos e quando a posição da linha da mediana é próxima ao terceiro quartil, os dados são assimétricos negativos. Vale ressaltar que a mediana é a medida de tendência central mais indicada quando os dados possuem distribuição assimétrica, uma vez que a média aritmética é influenciada pelos valores extremos.

Caudas – As linhas que vão do retângulo até aos outliers podem fornecer o comprimento das caudas da distribuição.

Outliers – Já os *outliers* indicam possíveis valores discrepantes. No *boxplot*, as observações são consideradas *outliers* quando estão abaixo ou acima do limite de detecção de *outliers*.

2.4 TRABALHOS RELACIONADOS

A área de pesquisa em ciência de dados já há algum tempo vem ganhando seu espaço no campo educacional. Segundo Rodrigues (2014) dentre os anos de 2006 a 2010, os trabalhos relacionados a mineração de dados eram apenas quinze no Brasil, no entanto, de 2011 a 2014 esse número foi elevado a 50 artigos, mostrando uma evolução muito importante sobre o tema. Diante do crescente e com a grande quantidade de dados educacionais disponibilizados no Brasil, sendo o Inep seu

principal articulador. Nesse quesito Soares (2021), mediante uma revisão sistemática focada em trabalhos de mineração de dados educacionais, mostrou que de 19 trabalhos revisados, 11 são focados no desempenho acadêmico dos estudantes, inclusive exames como o Enem.

No trabalho de Souza (2021), foi realizada uma análise do perfil dos participantes do Enem do ano de 2019, onde foram utilizadas técnicas de estatísticas e de ciência de dados nos quais tinha como finalidade identificar características relevantes dos participantes do exame no ano de 2019 e possíveis relações dessas características com o desempenho deles.

Ele realizou neste trabalho um panorama bem detalhado dos participantes do Enem desta edição, mostrando através de gráficos e tabelas, os resultados obtidos no estudo. Diante disso, ele pode concluir que as desigualdades sociais de quaisquer países têm grande influência na qualidade da educação, uma vez que foram analisados dados de renda familiar desses candidatos.

Essas desigualdades, que podem ser descritas não apenas pela renda familiar, mas também pelo tipo de escola oriunda do estudante, pelo grau de escolaridade dos pais etc. Podem ser fatores inerentes aos resultados do exame dos estudantes ao término do ensino médio, fatores esses que também serão analisados nesta pesquisa.

Foi mostrado na pesquisa de Sousa (2021) que mais de 89% dos candidatos deste ano estão entre 15 e 19 anos, mostrou ainda que as maiores notas são dos candidatos com idade entre 15 e 19 anos, além disso, candidatos do gênero feminino teve maior participação neste ano, cerca de 59%, mostrando também um equilíbrio nas médias das notas, e que o sexo não é um fator determinante para obter maior ou menor média.

No trabalho de Leria (2022) *et al.* trata de um assunto muito discutido hoje no meio acadêmico e da pesquisa educacional que é o acesso do estudante com deficiência visual ao ensino superior, por meio de análises dos microdados do Enem de 2017 e 2018. Onde é relatado que, o Enem, além de ser o principal instrumento de acesso ao ensino superior, também é instrumento de políticas públicas para ofertas de vagas subsidiadas pelo estado. O foco principal é a investigação dos fatores relacionados ao rendimento acadêmico em uma avaliação de larga escala do Enem.

Os autores investigaram nessa pesquisa exploratória com abordagem quantitativa a participação, características socioeconômicas e demográficas e o desempenho dos estudantes com deficiência visual no exame em comparação com os estudantes sem nenhum tipo de deficiência, tendo como objetivo identificar se há algum tipo de vulnerabilidade no acesso dessas pessoas à educação superior.

Diante da crescente dos estudantes com deficiência, cerca de 257% de 2009 a 2018, segundo o Inep, apesar dessa crescente, autores relatam que a quantidade de pessoas com deficiência na educação superior no Brasil pode chegar a menos de 1% em relação às vagas oferecidas, contudo, para compensar essas desigualdades nos processos seletivos em universidades, é utilizado uma política de cotas tanto para universidades públicas quanto para privadas, à alunos regressos de escolas públicas e negros, sendo obrigatória em universidades federais a partir de 2018.

Os resultados obtidos na pesquisa de Leria (2022) *et al* demonstraram que há um baixa muito grande de participação de candidatos de pessoas com deficiência visual no exame, além de possuírem menores rendas, maiores idades, menores participações de pessoas do sexo feminino e menor concentração na região sul e sudeste em relação à participantes sem nenhum tipo de deficiência. No entanto, um fato importante é que o desempenho desses candidatos é similar aos dos candidatos

sem deficiência, desmistificando, assim, o conceito de incapacidade dessas pessoas, os pesquisadores relatam ainda, que de acordo com a severidade da deficiência os resultados podem variar, um sinal de falta de acessibilidade no exame que pode prejudicar esses candidatos no acesso à Educação Superior.

No trabalho realizado por Rocha (2022) *et al.* foi feito uma análise dos dados dos candidatos do Enem na região nordeste nos anos de 2017 a 2019, mostrando o quão é importante esses tipos de estudo para a sociedade educacional de nosso país, pois é através desses dados disponibilizados de forma simples e gratuitas pelo Inep que podemos lapidar e tirar conclusões sobre a educação do Brasil, e não só dados do Enem, mas muitas outras bases de dados também são disponibilizadas, como a base da SAEB, que é um sistema de avaliação da educação básica brasileira, Enade, que é um exame nacional de desempenho dos estudantes do ensino superior do sistema federal de ensino e outras, todas de suma importância para sabermos quais caminhos traçar para termos uma educação de mais qualidade.

Ainda sobre essas pesquisas de exploração de dados, não podemos deixar de fora um fato muito importante que assolou o mundo na última década, a pandemia de coronavírus, que, de um jeito ou de outro, afetou a todos de maneira emocional e cognitiva, principalmente a área educacional, onde tiveram que lidar, meio a um caos, com a novas formas de se ensinar e aprender.

No trabalho de Neto (2022) *et al.* foi feito um estudo do impacto da pandemia de Covid-19 no exame do Enem, nos anos de 2019 e 2020, onde foi detectado uma melhoria nas notas médias gerais e um aumento na disparidade das notas entre os candidatos com maior e menor renda, mostrando que muitos estudantes onde a renda média familiar é baixa, podem não possuir acesso à internet em casa, ou mesmo não ter um eletrônico que possibilita esse acesso, e com o isolamento social e as dificuldades educacionais enfrentados por muitos alunos no período, houve um aumento na evasão escolar.

Diante disso, e buscando conhecer melhor o tema, foi analisado um artigo sobre os desafios enfrentados no ensino remoto de uma escola da rede estadual do estado do Tocantins no período da pandemia no ano de 2021. Gularte, Nascimento e Carvalho (2021), nesse trabalho destacam a situação pelo qual o mundo se passa em um momento de pandemia, focado nas novas perspectivas e mudanças educacionais atribuídos ao momento em questão, destacando os desafios enfrentados pelos profissionais da área de educação com as novas formas de ensino e a preparação desses estudantes para o Enem.

O trabalho se deu em forma de questionário de pesquisa exploratória, onde foram atribuídas 10 questões, objetivas e subjetivas, à professores atuantes no ensino aprendizagem dos alunos, relacionadas às metodologias utilizadas e mudanças que foram sofridas durante a pandemia. Os autores relatam que os resultados mostraram com a pandemia a preparação dos alunos para o Enem se tornara bastante desafiadora para os docentes que vinham tentando se reinventar de várias formas para que os alunos pudessem atingir as competências e habilidades adequadas para o exame.

Notou-se que durante a pesquisa, as mudanças vivenciadas pelos docentes foram significativas, nos quais eles precisaram se desdobrar e se reinventar para alcançar o aprendizado e rendimentos adequados para os alunos, pois esse foi um período onde muitos professores e alunos passaram por dificuldades emocionais e alta probabilidade de desenvolver transtornos, seja pela situação causada pela pandemia seja pelas condições de trabalho na escola.

Foi destacado também que na turma de 3º ano do ensino médio, último ano da educação básica e início do ensino superior, onde os professores pesquisados buscaram metodologias que facilitasse o entendimento dos alunos, no entanto, mesmo que os professores pesquisados instigassem os alunos a aprender, eles não tinham muito interesse. Diante disso, foi destacado que houve um alto índice de evasão escolar, principalmente no ensino médio, podendo ter prejudicado o futuro de vários alunos.

Conforme apresentado, existem vários trabalhos relacionados a análises de dados do Enem e alguns poucos sobre a pandemia de Covid-19, por ser algo recente ainda em nosso meio. No entanto, não foi identificado nenhum trabalho ainda que retrate as médias dos candidatos no Enem em épocas antes e depois da pandemia no que tange aos seus hábitos estudantis, no qual o próprio Inep disponibilizou um questionário sobre tais hábitos. Por tanto, a distribuição deste estudo é realizar uma análise exploratória dos dados do Enem pós pandemia (2022), para identificar quais eram os hábitos dos estudantes no período de pandemia, com relação à educação, e se isso teve impacto nas notas deles.

4 METODOLOGIA

As pesquisas exploratórias têm como objetivo proporcionar maiores familiaridades com o problema, com vista a torná-los mais explícitos ou a constituir hipóteses. (Gil, 2022). Minayo e Sanches (1993, p 247) afirmam que a abordagem quantitativa objetiva trazer à luz dados, indicadores e tendências observáveis [...] grandes aglomerados de dados, classificando-os e tornando-os inteligíveis através de variáveis. Dessa forma, foi realizada uma pesquisa de cunho exploratório com abordagem quantitativa e com análise de dados dos microdados do Enem, tendo como base as características socioeconômicas e hábitos de estudo durante a pandemia de Covid-19 dos candidatos que realizaram o exame.

4.1 FONTES E TIPOS DE AMOSTRAS

No sítio do Inep⁴, consta um repositório com os dados de todas as edições anteriores do Enem, podendo ser acessado por meio do link “Enem” e, então, por ano de edição. Para este estudo, utilizamos os dados do exame de 2019 e 2022. Aqui vale ressaltar que a partir do ano de 2021 o Inep mudou o padrão na forma de como os dados foram divulgados para atender a Lei Geral de Proteção de Dados (LGPD), uma vez que 44% das informações foram retiradas da base de dados, pois com as mudanças foram retiradas 60 colunas da base de dados, passando de 136 para 76 apenas (Inep, 2021), no entanto, nada que prejudique que os resultados desejados fossem alcançados neste trabalho.

Do questionário de hábitos de estudo durante a pandemia, para alcançar alguns dos objetivos deste estudo, foram analisadas os seguintes grupos temáticos do dicionário: a) inscrição; b) participação; c) situação de matrícula escolar e percepção da própria aprendizagem; d) gestão do tempo e planejamento de estudos; e) práticas de estudo e pesquisas; f) tecnologias e tipos de acesso; g) problemas na rotina de

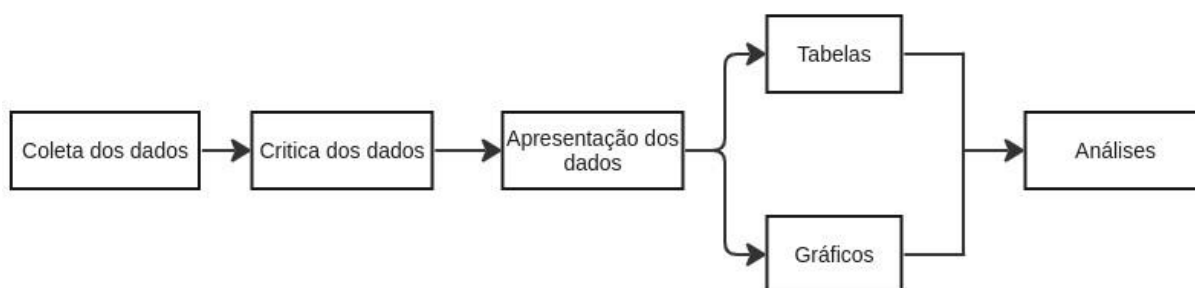
⁴ <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/enem>

estudos; h) dificuldade de infraestrutura; i) ajuda de terceiros e j) avaliação sobre a própria experiência.

4.2 PROCEDIMENTO DE COLETA DE DADOS

A estatística descritiva envolve a organização e sumarização dos dados através de métodos simples onde é analisado um conjunto de indivíduos. A Figura 3 demonstra através de um diagrama as etapas da estatística descritiva dos dados.

Figura 3. Diagrama da estatística descritiva dos dados.



Fonte: Elaboração própria (2023).

Após a definição do problema a ser estudado e do planejamento da pesquisa, o passo seguinte é a coleta de dados, que consiste na busca ou compilação dos dados das variáveis, componentes do fenômeno a ser estudados. O próximo passo é uma revisão crítica que tem a finalidade de suprimir os valores estranhos ao levantamento no intuito de evitar futuros enganos. Sendo assim, o próximo passo consiste na organização do conjunto de dados de modo prático e racional através de tabelas e/ou gráficos, dando resultados a análise estatística.

4.3 CARACTERIZAÇÃO DAS AMOSTRAS E ESTUDAS DAS ANÁLISES DE DADOS

Para caracterizar as amostras, foram utilizados critérios de exclusão e de seleção, como mostrado na Tabela 1.

O Inep disponibiliza, juntamente com os microdados, um documento denominado “Dicionário de Variáveis” que inclui a descrição das informações contidas nos microdados, definindo o tipo de informação (numérica ou alfanumérica) e possíveis conteúdo ou campos válidos, como, por exemplo, para a variável “TP_SEXO”, que é numérica com dois valores possíveis: “M” – masculino e “F” – feminino.

A Tabela 1, do dicionário de variáveis, mostra as variáveis utilizadas para as análises após as exclusões das variáveis desnecessárias para o cumprimento dos objetivos.

Tabela 1. Dicionário de variáveis socioeconômicas.

Variável	Descrição
NU_INSCRICAO	Número de inscrição
IN_TREINEIRO	Indica se o candidato fez a prova apenas para treinar seus conhecimentos
TP_PRESENÇA_CH	Presença na prova objetiva de Ciências da Humanas
TP_PRESENÇA_CN	Presença na prova objetiva de Ciências da Natureza
TP_PRESENÇA_LC	Presença na prova objetiva de Linguagens Códigos e suas tecnologias
TP_PRESENÇA_MT	Presença na prova objetiva de Matemática
NU_NOTA_CN	Nota na prova de Ciências da Natureza
NU_NOTA_CH	Nota na prova de Ciências Humanas
NU_NOTA_LC	Nota na prova de Linguagens e Códigos
NU_NOTA_MT	Nota na prova de Matemática
NU_NOTA_REDACAO	Nota da prova de Redação
Q006	Renda mensal da família
Q025	Posse de internet na residência

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Em 2022, por se tratar de um período pós pandêmico, o Inep considerou realizar um questionário de hábitos de estudos durante a pandemia. Foram questões elaboradas exclusivamente sobre os hábitos que os candidatos tinham durante o período do ano de 2021, questões essas referentes a situações de matrículas, infraestrutura e outras, nesse quesito, utilizamos para nossa pesquisa apenas as questões que respondessem a nossos objetivos, como descrito na tabela abaixo.

Tabela 2. Dicionário de variáveis utilizadas (Questionário sobre hábitos de estudo durante a pandemia).

Variável	Descrição
NU_INSCRICAO	Número de inscrição
PARTICIPAÇÃO	Tipo de opção de resposta ao questionário
Q001	Situação de acordo ao vínculo escolar durante a pandemia
Q002	Situação de acordo a realidade do estudante
Q003	Situação de acordo a experiência do estudante
Q004	Processo de aprendizagem durante a pandemia
Q029	Dificuldade de infraestrutura para estudar em 2021
Q031	Precisou de ajuda de alguém para estudar em 2021

Q034	O quanto o candidato se sente preparado para conduzir o processo de aprendizagem em 2021
------	--

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

4.4 CRITÉRIOS DE SELEÇÃO DE DADOS

Para o presente estudo, foram eliminados dados da base que não iam ser utilizadas, para reduzir o processamento da máquina e retornar resultados válidos e corretos relacionado ao público em questão, bem como dados ausentes, ou considerado nulos nas tabelas, assim como informações daqueles candidatos que não esteve presente em alguns dos dias de aplicação do Enem, os candidatos treineiros, bem como os que continham nota zero na redação. Assim, a base de dados foi reduzida em cerca de 52%.

4.5 ARQUITETURA DAS FERRAMENTAS

As ferramentas de arquitetura de software têm um papel crucial no desenvolvimento de projetos de *software*. Elas conferem características e funcionalidades que auxiliam a assegurar a qualidade, a eficiência e a escalabilidade das soluções criadas. Diante disso, destaca-se as principais características das ferramentas utilizadas nesse projeto.

A linguagem utilizada foi *Python*⁵ Em sua versão 3, ela é uma linguagem poderosa e de fácil aprendizado. Possui uma estrutura de dados eficiente de alto nível de abordagem simples, essencial para programação orientada a objetos. É uma linguagem de natureza interpretada com código fonte *Python*, gerenciador de memória, sistemas de tipos dinâmicos e várias bibliotecas padrões. Possui uma sintaxe e a tipagem dinâmica, aliadas com sua natureza interpretativa, fazem da linguagem ideal para programas e desenvolvimento de aplicações rápidas em diversas áreas e na maioria das plataformas. (*Python Software Foundation*, 2023).

A biblioteca *Pandas*⁶ É uma biblioteca de código aberto do *Python* para manipulação de análises de dados, licenciada pela BSD, que possui estruturas de dados e ferramentas de análises de dados de alto desempenho e fácil de usar para a linguagem *Python*.

*Matplotlib*⁷ É uma biblioteca para fazer gráficos 2D de *arrays* em *Python*. Ela é utilizada na forma *pythonic*, orientada a objetos, embora seja utilizada principalmente em *Python* puro, ele faz uso intenso de outra biblioteca, a *NumPy*. Ela é uma biblioteca projetada com a filosofia de que você deve ser capaz de criar gráficos simples com apenas alguns comandos, ou apenas UM.

A biblioteca *Seaborn*⁸ Serve para realizar gráficos estatísticos em *python*, baseada em *matplotlib*, ela é integrada intimamente às estruturas de dados do *pandas*, ajudando na compreensão e exploração dos dados. Suas funções de plotagem operam em *dataframes* e *arrays* contendo conjuntos de dados inteiros e realizam internamente o mapeamento semântico e agregação estatística necessários para

⁵ <https://docs.python.org/3/tutorial/>

⁶ <https://pandas.pydata.org/pandas-docs/stable/#module-pandas>

⁷ <https://matplotlib.org/3.8.2/users/project/history.html>

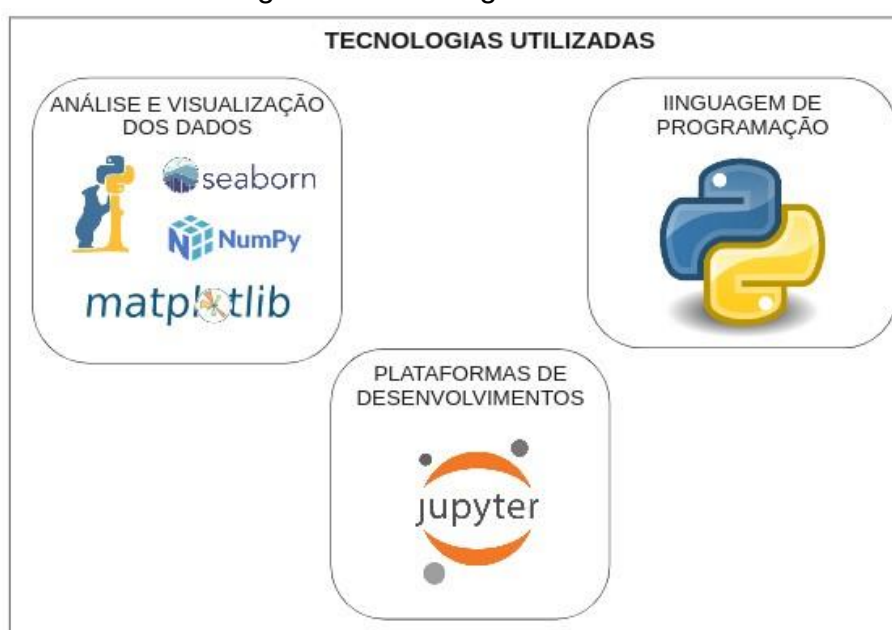
⁸ <https://seaborn.pydata.org/tutorial/introduction.html>

produzir gráficos informativos. Sua API declarativa orientada a conjuntos de dados permite que se concentre no significado dos diferentes elementos de seus gráficos, em vez dos detalhes de como desenhá-los.

4.6 TECNOLOGIAS UTILIZADOS

A Figura 4 mostra o desenho sobre todas as tecnologias utilizadas no processo de análise de dados deste trabalho. Nesse tópico, as tecnologias referentes à análise e visualização dos dados e linguagem de programação, já foram muito bem descritas no tópico anterior, por tanto, será feito apenas uma análise visual e sucinta na Figura abaixo.

Figura 4. Tecnologias utilizadas.



Fonte. Elaboração própria (2023).

Como vimos anteriormente, a linguagem de programação utilizada foi a *Python*, em sua versão 3, ela possui uma gama de bibliotecas que nos auxiliam nas análises, como visto no tópico anterior, no qual foi usada a biblioteca *Pandas*, onde foi feita toda a leitura, tratamento e processamento dos dados. A biblioteca *Matplotlib* é utilizada na criação de diversos tipos de gráficos para os mais variados tipos de dados, assim como a *Seaborn*, e *Numpy*, que auxiliam na formatação dos *layouts*.

Para o ambiente de execução, foi utilizado o *Jupyter Notebook*. Que é um *software* gratuito, com padrões abertos e ambiente de desenvolvimento iterativo baseado na *web* para *notebooks*, códigos e dados. Ele possibilita que máquinas que não tenham um *hardware* tão potente possam lidar com bases de dados massivas, uma vez que ele utiliza o próprio navegador como compilador dos códigos.

4.6.1 Tratamentos dos dados

Os arquivos de formato .csv que contém os dados do Enem de 2019 e 2022 são descarregados do sítio do Inep de forma compacta, e possuem tamanho, descompactado, de cerca de 2.56 GB. Felizmente, a linguagem *Python* nos proporciona várias bibliotecas que facilitam na hora de analisar os dados, e umas delas é a biblioteca *zipfile* que permite que os dados sejam analisados sem a necessidade de descompactá-los, economizando memória e processamento, ideal para computadores de baixo poder de processamento.

Para auxiliar no processo de entendimento das colunas a serem exploradas, foi utilizado o dicionário dos microdados, que também estão disponíveis no sítio do Inep, pois aqui vale ressaltar a importância do processo de preparação dos dados como uma etapa essencial na análise de dados, por tanto, os conjuntos de dados dos anos de 2019 e 2022 passaram pelos mesmos processos de preparação.

A análise de dados refere-se ao processo crítico de realizar investigações iniciais em dados para descobrir padrões, detectar anomalias, testar hipóteses e verificar suposições com a ajuda de estatísticas descritivas e representações gráficas (Komorowski et al., 2016). O objetivo dessa etapa neste estudo foi extrair dados que pudessem ser efetivamente transformados em informações importantes para entender os impactos no Enem causados pela pandemia. Além da análise exploratória dos dados, o principal foco dessa etapa foi gerar informações comparativas e analisar se houve impacto (i.e., se houve diferenças).

Do conjunto total de dados, foram selecionadas as colunas referentes às questões socioeconômicas, informações de realização da prova, e dados sobre o participante, além do questionário sobre hábitos de estudo na pandemia realizado no Enem de 2022. Para diversas análises de informações, consideramos apenas os participantes presentes e que não tenham tirado nota zero na redação, o qual é caracterizado por estar presente, ao menos no primeiro dia de prova. Após a seleção das características abordadas, geramos análises exploratórias, com o uso de gráficos e tabelas, utilizando-se de apenas um *Jupyter notebook* carregando os microdados dos dois anos para propósitos de comparação.

5 RESULTADOS E DISCUSSÃO

Nesta seção serão apresentados os estudos realizados nas análises descritivas dos dados obtidos nas bases a partir dos microdados dos participantes do Enem do ano de 2019 e 2022. Serão analisados, em alguns casos, apenas os candidatos que não obtiveram nota zero na redação e que esteve presente em pelo menos um dos dias do exame. As análises, descartando esses casos, serão apenas para o dicionário de variáveis, onde é preciso essa crítica dos dados para que os resultados sejam mais precisos.

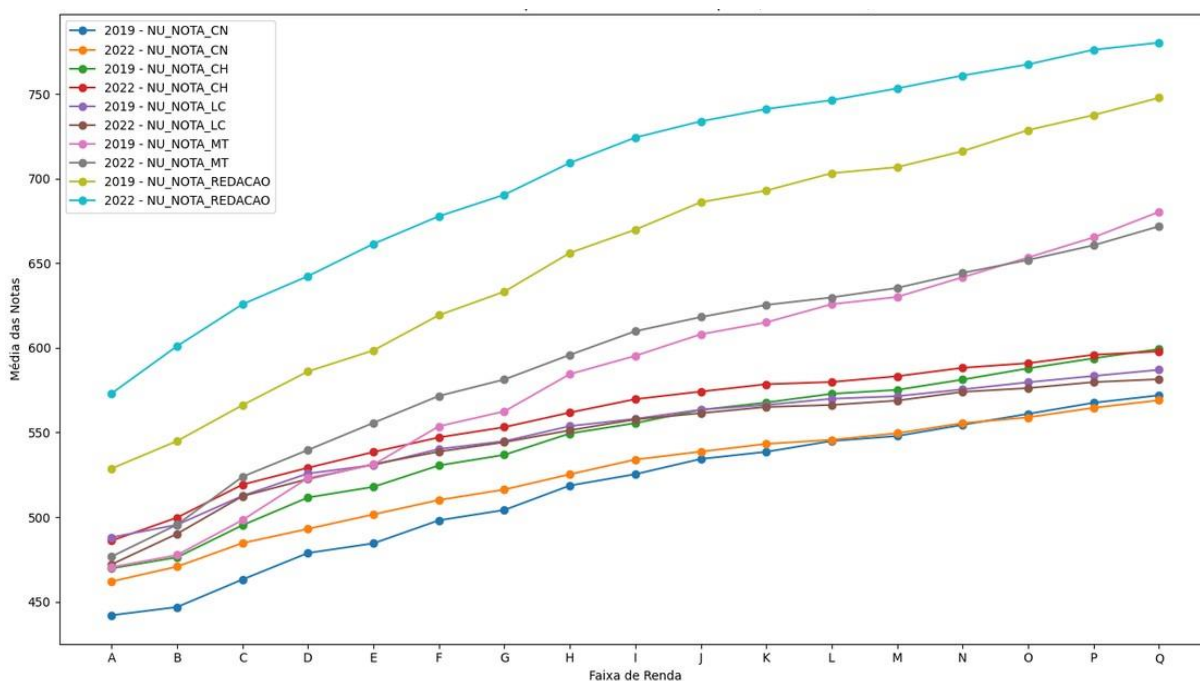
Para Lima e Brighenti (2023) em bases de dados reais, é comum que alguns dados possam estar incompletos ou inconsistentes, pois dados de determinados atributos com informações ausentes ou pouco detalhadas, são considerados incompletos (*outliers*).

Por tanto, e levando em consideração essa afirmativa, os dados oficiais divulgados pelo Inep e os dados obtidos pela base de dados do Enem estão descritos na Tabela 3.

A redação do Enem é realizada no primeiro dia de aplicação do exame, e em 2019 cerca de 143.000 estudantes zeraram a redação. Já em 2022 foram cerca de 129.000 estudantes. Para chegar a esses resultados temos que considerar os candidatos que estiveram presentes, ao menos no primeiro dia de aplicação, e levando em consideração que em 2019 foram cerca de 3,9 milhões de candidatos presentes no primeiro dia e em 2022 foram 2,4 milhões de candidatos, podemos concluir que 3,6% e 5,2%, respectivamente, dos candidatos zeraram a redação.

Em 2019 foram 60% inscritos do sexo feminino e 40% do sexo masculino, assim como em 2022, quando foram 61% do sexo feminino e 39% do sexo masculino, estatística essa que desde 2015, os candidatos do Enem são, em sua ampla maioria do sexo feminino. Nesse quesito, quando se trata da cor e raça declarados pelos candidatos, a maioria são declarados como pardos, cerca de 46% em 2019 e 43% em 2022. Assim como a maioria vem de escolas públicas e outras várias gamas de variáveis que poderiam ser descritas aqui, no entanto, nosso foco está em desvendar os hábitos desses estudantes no período da pandemia, sem deixar de observar algumas peculiaridades em algumas variáveis do Enem, como visto no gráfico abaixo, onde podemos observar as médias de acordo a faixa etária da renda do candidato dos dois anos. Para entendermos melhor a legenda do Gráfico 2, onde está escrito NU_NOTA_CN trata-se do eixo temático da avaliação do Enem de Ciências da Natureza, NU_NOTA_CH, Ciências Humanas, NU_NOTA_LC, Linguagens e Códigos, NU_NOTA_MT, Matemática e NU_NOTA_REDAÇÃO, Redação. Assim como no eixo x, as letras representam a faixa etária da renda familiar do candidato, indo de “A” R\$: 998.00 a “Q” R\$: 24.240.00.

Gráfico 2. Média das notas de acordo com a faixa etária da renda familiar dos candidatos de 2019 e 2022.



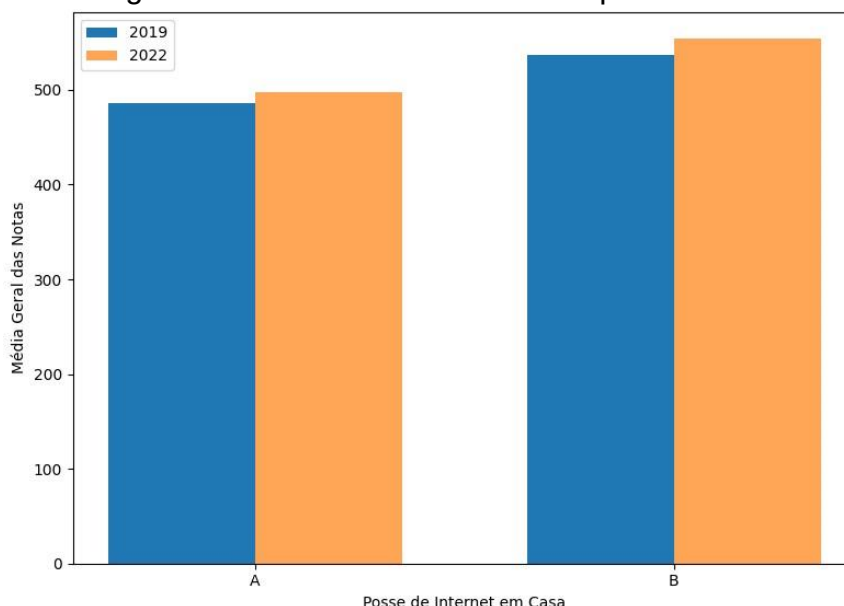
Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Mesmo com a pandemia, o IDEB do ensino médio no Brasil, em 2019 e 2021 não foi alterado, ou seja, segundo Inep, em ambos os anos a média foi de 4,2, igualando a média nacional e sendo as maiores médias desde 2005. No entanto, a média das escolas da rede estadual foi apenas 3,9 em um cenário em que 84,5% dos estudantes matriculados no ano de 2021 fazem parte dessa rede de ensino.

A renda das famílias brasileiras diminuiu bastante no período da pandemia, pois muitos tiveram que parar suas atividades em decorrência dos isolamentos sociais por causa da Covid-19, o Senado Federal (2020) estima que cerca de 68% das famílias diminuam suas rendas em 2020. Em geral, as médias dos candidatos, tanto em redação quanto nas demais disciplinas, vão aumentando de acordo com a renda familiar. Um fator interessante de se observar é que em 2019 as médias são menores que em 2022, pois em 2019 não tinham pandemia ainda, e as aulas aconteciam normalmente nas escolas, em 2022, apesar do retorno normal das atividades escolares, o país tinha acabado de passar por uma pandemia onde a maioria das atividades escolares eram dadas de maneira remota ou semi-presencial, as maiores mudanças podem ser observadas em redação, quando em 2022 as notas foram maiores.

Em tempos modernos onde a tecnologia faz parte do dia a dia do cidadão, as ferramentas de acesso a internet se tornaram parte crucial no ensino aprendizagem dos estudantes brasileiros, alavancando ainda mais no período da pandemia com a suspensão das aulas presenciais. Diante disso, no questionário socioeconômico, no ano de 2019, 22,39% do total de candidatos inscritos no Enem responderam que possuíam acesso a internet em casa, e 77,61% disseram que não possuíam. Em 2022 esse número era de 9,4% para os que não possuíam e de 90,6% para os que possuíam. No Gráfico 3, no eixo x, “A” está representado os candidatos que não possuem internet em casa e “B” os que possuem.

Gráfico 3. Média geral das notas de acordo com a posse de internet em casa.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

É notório que quando o estudante possui o auxílio de uma rede de internet para realizar suas pesquisas e auxiliar nos estudos do dia a dia, seu rendimento escolar se torna maior, como é mostrado no Gráfico 3, onde as médias gerais dos candidatos são maiores em 2022, tanto para os que possuem internet em casa quanto para os que não possuem.

5.1 QUESTIONÁRIO DE HÁBITOS DE ESTUDOS DURANTE A PANDEMIA

Na seção a seguir, serão mostrados os hábitos de estudos dos candidatos do Enem no período da pandemia de Covid-19 em 2022. Foram analisadas questões relevantes dos grupos temáticos da situação de matrícula escolar e percepção própria de aprendizagem; gestão do tempo e planejamento de estudos; práticas de estudo e pesquisa; tecnologia e tipos de acesso; problemas na rotina de estudos; dificuldades de infraestrutura; ajuda de terceiros e avaliação sobre a própria experiência.

No questionário de hábitos de estudo durante o ano de 2021, os candidatos não eram obrigados a responder a todas as questões, bem como, se optasse por responder ao questionário, não necessariamente deveria responder as questões na sequência, ou seja, a partir do momento que o candidato optasse por responder, ele poderia alternar suas repostas.

O Inep divulgou um quantitativo de 928.564 estudantes que responderam ao questionário, havendo assim, uma perda de apenas 17 estudantes em comparação com a base dos microdados destinada pelo Inep no site. Na tabela 4, mostra um panorama geral, através de números, dos candidatos do Enem de 2022 sobre o questionário de hábitos de estudo dos candidatos durante a pandemia.

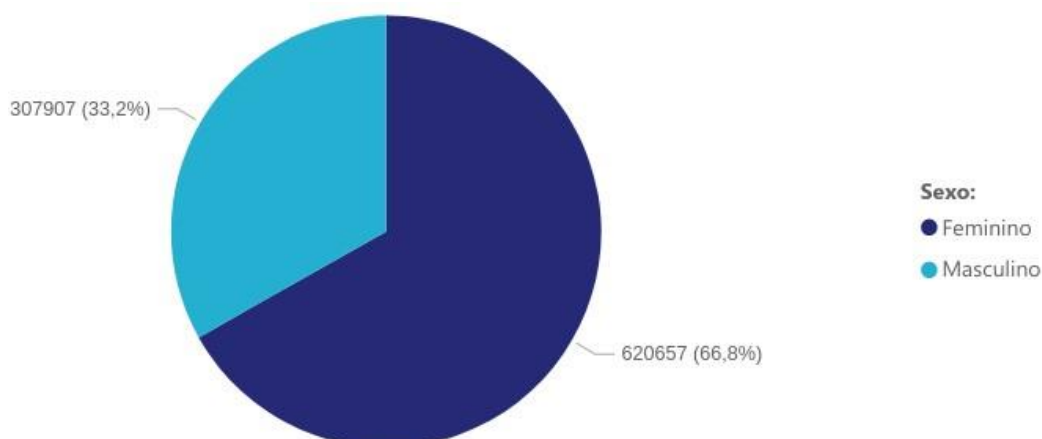
Tabela 4. Distribuição dos candidatos referente ao questionário de hábitos de estudo durante a pandemia.

Descrição	Quantidade
Total de candidatos inscritos	2.467.086
candidatos que responderam ao questionário	928.547
candidatos que não responderam ao questionário	346.348
candidatos que optaram para responder depois	1.192.191

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

O dicionário de variáveis do questionário de hábitos de estudos é vasto e contém os questionamentos sobre quase tudo que o estudante vivenciou durante o período de pandemia com relação à vida escolar, no entanto, os candidatos possuíam três opções, a de responder o questionário imediatamente, a de não responder e a de responder depois, diante do fato, cada pergunta possui uma quantidade específica de respostas, variando de questão para questão. De modo geral, o questionário de hábitos de estudos durante a pandemia nos traz um panorama onde 66,8% dos candidatos eram só sexo feminino, e 33,2% masculino, Gráfico 4.

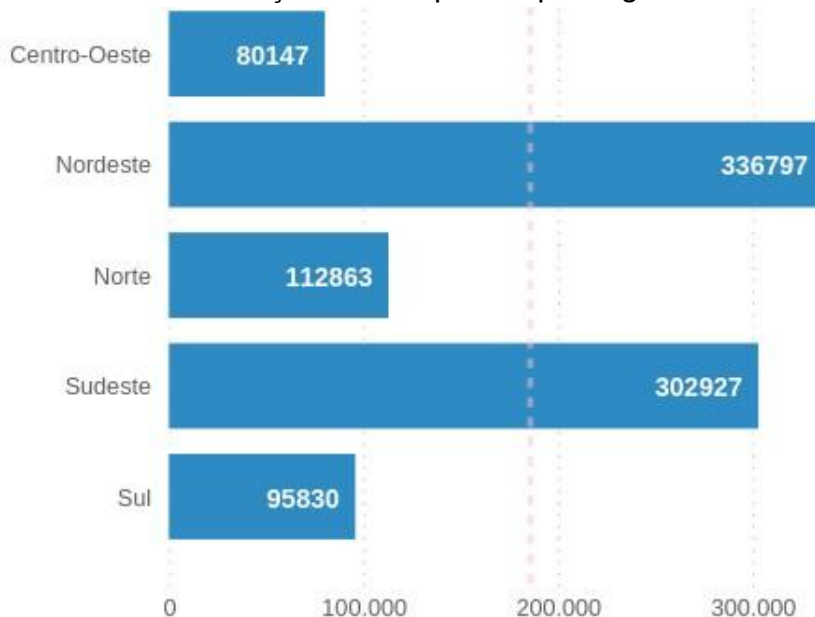
Gráfico 4. Distribuição dos participantes por sexo.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Assim como as mulheres são maioria na prova do Enem, tanto em 2019 quanto em 2022, e os mesmos candidatos que foram inscritos para o exame também foram submetidos ao questionário, é justificável sua maioria nas respostas do questionário de hábitos de estudos durante a pandemia também sejam dos candidatos do sexo feminino.

Gráfico 5. Distribuição das respostas por regiões do Brasil.



Fonte: Inep 2022.

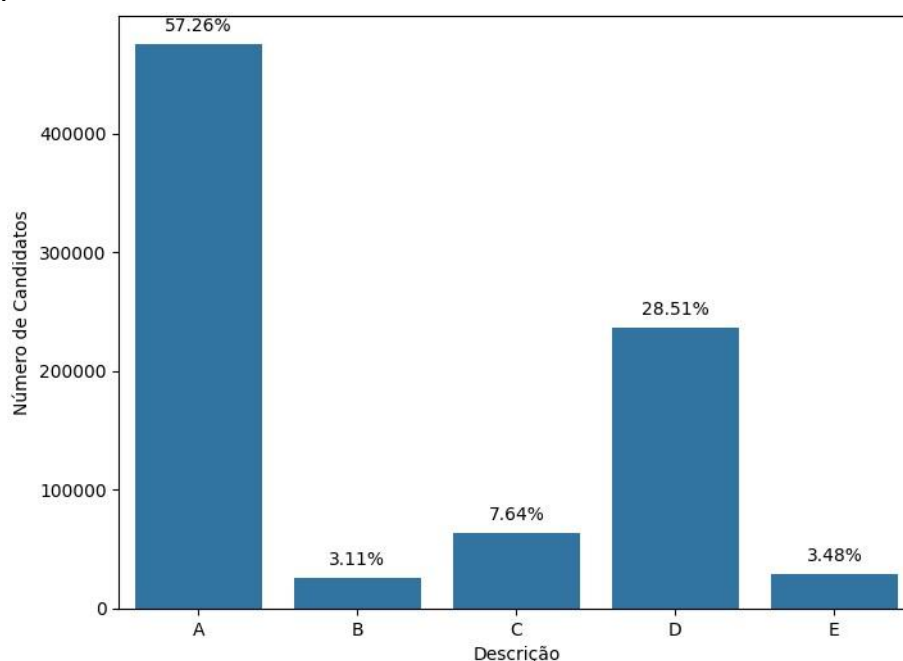
Mesmo sendo apenas a segunda região mais populosa do país, a região nordeste foi onde houve mais candidatos dispostos a responder o questionário, fato esse que está em contraste com a quantidade de inscritos no Enem do mesmo ano, assim como os demais quantitativos também estão em paralelos à quantidade de inscritos no Enem de 2022. A descrição do Gráfico 6 está distribuída da seguinte forma, no seu eixo X, na linha horizontal do gráfico, como observado na Tabela 5.

Tabela 5. Distribuição das variáveis do eixo x do Gráfico 6.

Variável	Descrição
A	Os alunos estiveram matriculados no ensino médio regular
B	O aluno esteve matriculado em uma turma de EJA
C	O aluno esteve matriculado no ensino médio integrada à educação profissional
D	O aluno concluiu o ensino médio antes de 2021
E	O aluno não esteve matriculado nesse período

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 6. Situação que está de acordo com vínculo escolar do estudante durante a pandemia.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

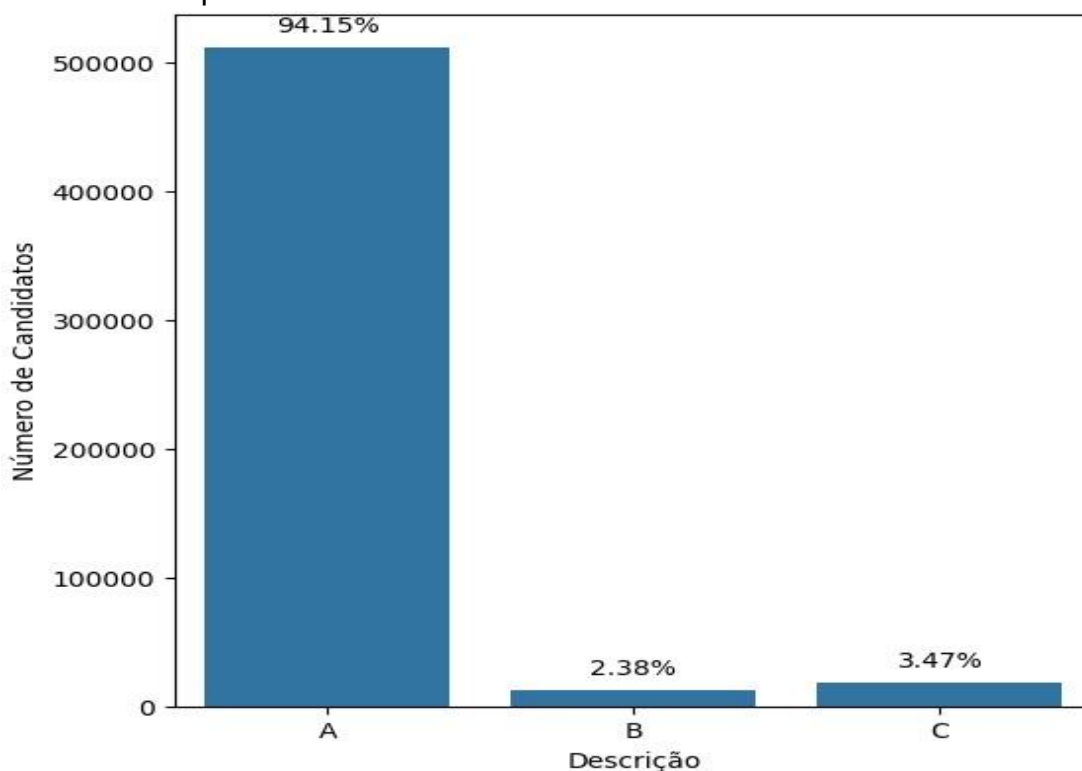
Podemos observar que a maioria dos alunos pesquisados e que responderam ao questionário esteve matriculado durante a pandemia, seguido de já terem concluído o ensino médio antes de 2021, mostrando que apesar do período difícil de pandemia a maioria dos estudantes preferiram continuar os estudos sem interromper.

Tabela 6. Distribuição das variáveis do eixo x do Gráfico 7.

Variável	Descrição
A	O estudante continuou os estudos sem interromper o ano letivo
B	O estudante decidiu interromper os estudos no primeiro ano da pandemia, mas se matriculou em 2021
C	Ele foi levado a interromper os estudos no primeiro ano da pandemia por não ter recebido suporte da escola, mas matriculou-se novamente em 2021

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 7. Considerando a continuidade do vínculo escolar na pandemia, situações que estão de acordo com a realidade do estudante.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

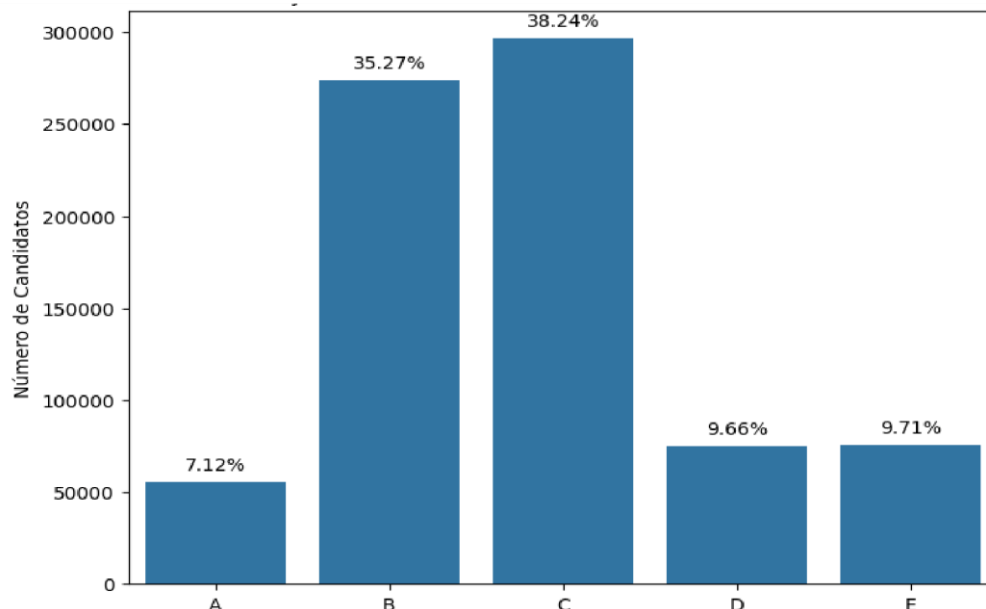
Podemos observar no Gráfico 7 que a continuidade do vínculo escolar do estudante durante o período de pandemia, de acordo com a realidade dele, não foi alterada, ou seja, a maioria continuou seus estudos normalmente e os poucos que saíram matricularam-se novamente. Num período em que a taxa de abandono escolar no ensino superior foi de 16,3% em 2020, no questionário levantado no gráfico acima os alunos que permaneceram não desistiram de continuar seus estudos e buscar uma vaga nas universidades brasileiras. Instituto alicerce (2022).

Tabela 7. Distribuição das variáveis do eixo x do Gráfico 8.

Variável	Descrição
A	Eu apenas tive aulas presenciais nesse período
B	Eu apenas estudei em casa de forma remota nesse período
C	Eu estudei de forma híbrida, ou seja, em casa remotamente e em aulas presenciais
D	Eu realizei atividades de estudo diversas, mas sem estar matriculado(a) no ensino médio
E	Eu não estive matriculado(a) regularmente, nem realizei atividades de estudo

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 8. Distribuição das experiências educacionais dos candidatos do Enem 2022.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

No Gráfico 8 estão distribuídas as respostas dos candidatos com relação ao ano de 2021 (o segundo ano da pandemia), no que diz respeito à situação do estudante de acordo com a experiência dele. Essa foi a terceira pergunta realizada pelo Inep aos candidatos do Enem de 2022.

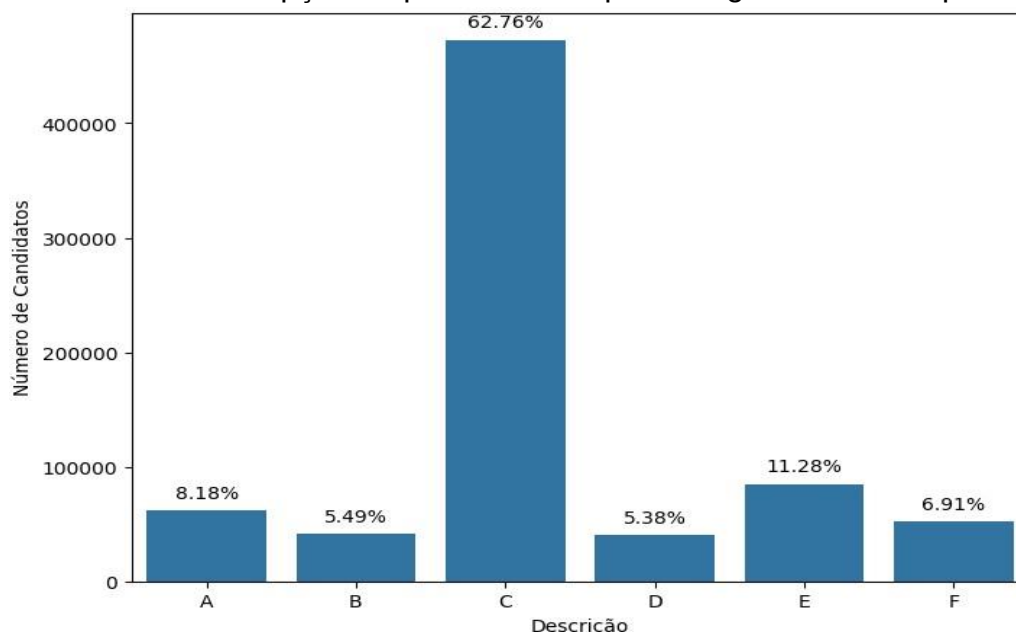
Podemos ver que a maioria dos estudantes, neste período, estudaram de forma remota e híbrida, mas que também houve muitos alunos que tiveram aulas presenciais. A quantidade de quem não esteve matriculado no período também é significativa, pois podemos supor que muitos ficaram fora da escola por medo ou por ter que trabalhar para complementar o sustento da família, uma vez que já observamos anteriormente uma queda nas rendas das famílias no período, ou até mesmo por ter contraído o vírus de Covid-19.

Tabela 8. Distribuição das variáveis do eixo x do Gráfico 9.

Variável	Descrição
A	Como eu aprendi mais no ensino remoto (estudando somente em casa)
B	Eu aprendi mais estudando de forma híbrida
C	Aprende mais estudando de modo presencial do que nas formas híbridas ou remotas
D	Eu aprendi no ensino remoto ou híbrido na mesma medida de quando eu estava no ensino presencial
E	Eu não estive matriculado(a) em 2021, mas aprendi estudando por conta própria
F	Eu não estive matriculado(a) em 2021 e nem estudei

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 9. Percepção do processo de aprendizagem durante a pandemia.

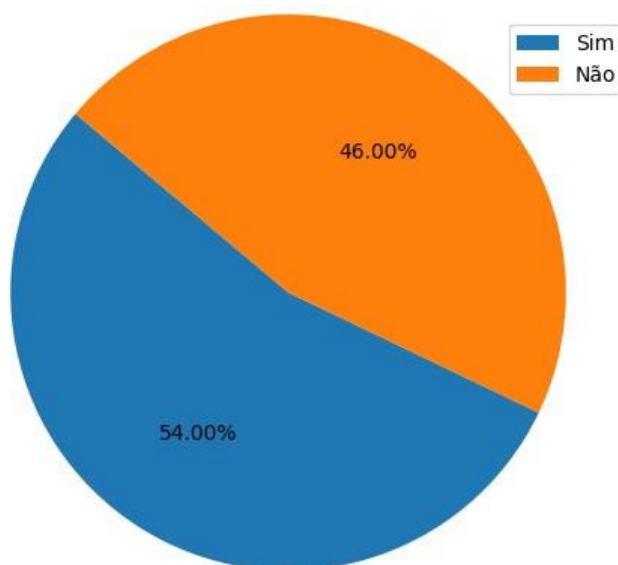


Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Quando perguntado sobre como o candidato percebeu seu processo de aprendizagem durante a pandemia, a maioria respondeu que aprendeu mais estudando de modo presencial do que nas formas híbridas ou remotas, representado pela variável C.

Uma das maiores causas de evasão escolar nesse período foi justamente a falta de acesso a internet, uma vez que 47 milhões de pessoas não tinham acesso à internet, Instituto alicerce, (2022). Em contraste com o Gráfico 10, podemos perceber que quando perguntado sobre as dificuldades de infraestrutura para estudar durante a pandemia em 2021, a maioria dos candidatos responderam que não tiveram tantas dificuldades em infraestruturas para estudar, uma diferença de apenas 5% dos que responderam que tiveram algumas dificuldades, mostrando assim que a falta de infraestrutura pode ter sido uma das maiores dificuldades para os estudantes nesse período.

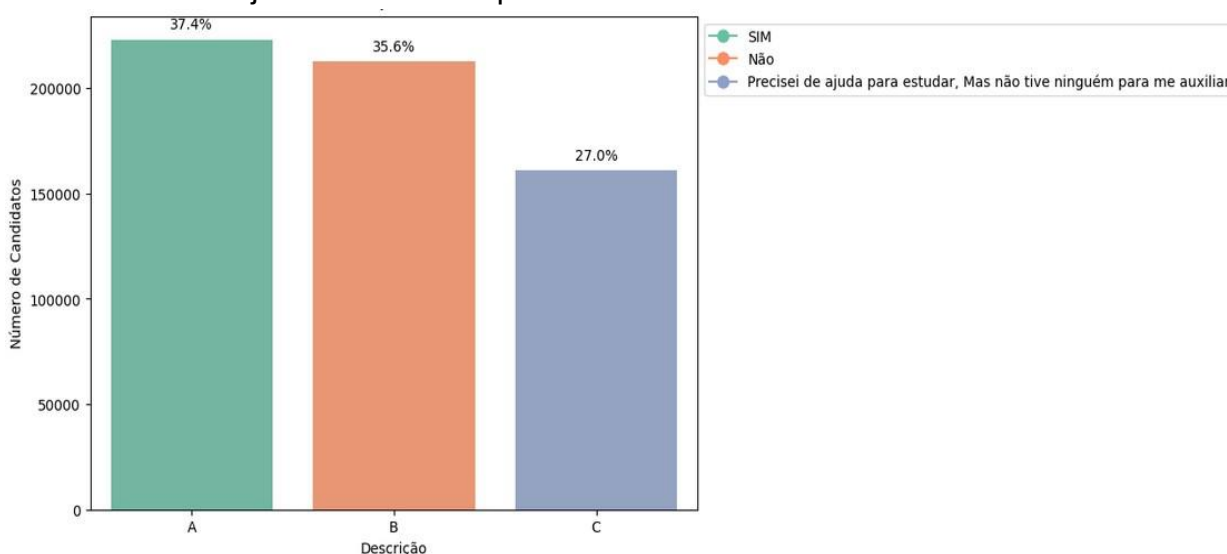
Gráfico 10. Dificuldades de infraestrutura para estudar durante a pandemia.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Sabemos que a falta de estrutura pode ter sido um dos principais fatores de dificuldades no processo de ensino aprendizagem durante a pandemia de Covid-19 no país, uma vez que quase todas as escolas operaram na forma remota ou híbrida. A falta de equipamentos adequados em casa, como computadores e internet, foram um dos principais pontos relatados pelos estudantes, onde 58,55% dos que responderam reclamaram da falta de conexão ou conexão ruim com internet, e 20,55% relataram não ter computador ou notebook disponível por ter que dividir o equipamento com familiares, talvez, por causa dessas dificuldades foi que muitos disseram ter necessidade de ajuda de terceiros.

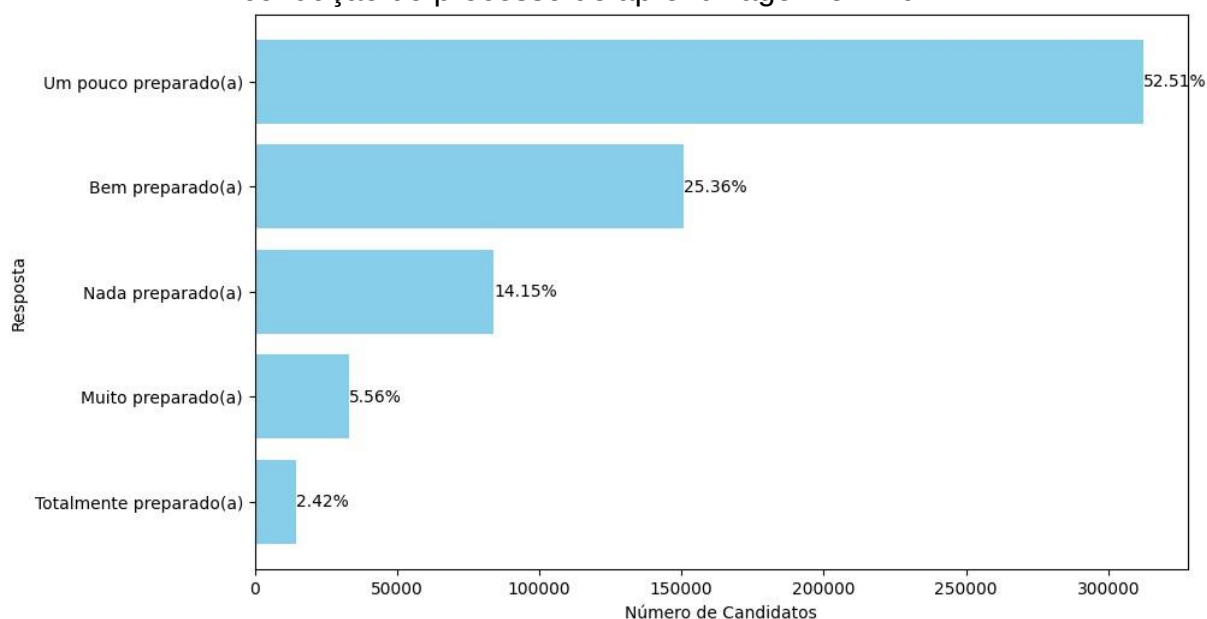
Gráfico 11. Ajuda de terceiros para estudar e se manter informado em 2021.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Quando perguntado sobre ajuda de terceiros para estudar ou manter-se informado durante a pandemia, muitos dos candidatos responderam que sim, seja ajuda de colegas, parentes, pais, ou até mesmo de professores, mostrando as dificuldades que por si só, os estudantes tinham na hora de estudar, precisão essa que podem ter sido geradas por necessidades já vistas anteriormente, falta de acesso a internet e dificuldades de infraestrutura. Uma pequena minoria também respondeu que precisou, mas que, porém, não encontrou essa ajuda, pois talvez essa falta de ajuda esteja relacionada com a ausência dessas pessoas mais próximas, pois foi uma época em que a zona de risco de contágio se estendeu por todo lugar.

Gráfico 12. Distribuição sobre o quanto o candidato será preparado na condução do processo de aprendizagem em 2022.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Uma minoria dos candidatos que responderam ao questionário afirma que estão totalmente preparados no processo de condução do ensino aprendizagem em 2021. Mas apesar de todas as dificuldades já relatadas, mais da metade sente-se pouco preparado nesse processo em meio a um ano atípico para a educação, e outros meios sociais, econômicos etc.

Para se ter uma ideia dessas distribuições de notas do Enem, no Brasil a cada dois anos é feito um levantamento do IDEB escolar de estudantes do Ensino Fundamental I, II e 3ª série do ensino médio, para diagnosticar a evolução, ou não, de indicadores de melhoria na qualidade da educação, uma nota que vai de 0 a 10. Em 2019 e 2021, como já mencionado, essa taxa para o ensino médio não se alterou, ficando em 4,2, apesar da pandemia, Inep¹¹.

A distribuição do eixo X, do Gráfico 13, onde estão localizadas as variáveis das matérias estão de acordo a Tabela 9 abaixo:

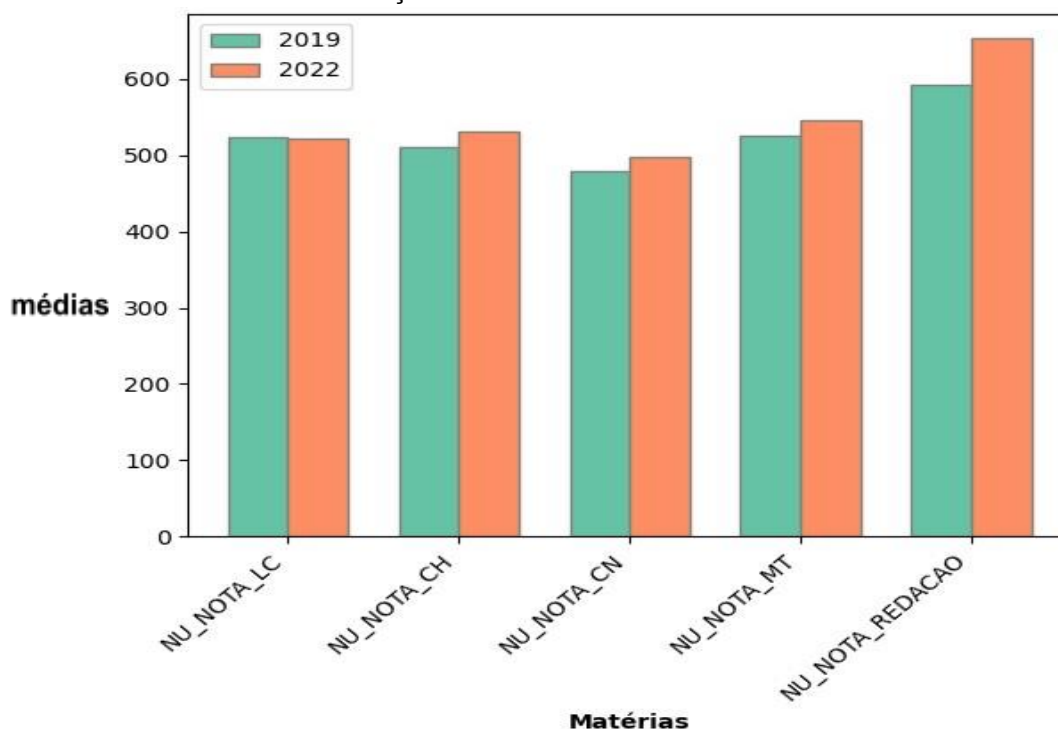
¹¹ https://download.inep.gov.br/institucional/apresentacao_saeb_ideb_2021.pdf

Tabela 9. Distribuição das variáveis do eixo x do Gráfico 13.

Variável	Descrição
NU_NOTA_LC	Linguagens e códigos
NU_NOTA_CH	Ciências humanas
NU_NOTA_CN	Ciências da natureza
NU_NOTA_MT	Matemática
NU_NOTA_REDAÇÃO	Redação

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 13. Distribuição das médias dos candidatos em 2019 e 2022.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Por tanto, como pode ser observado no Gráfico 13, a não alteração dessas pontuações em 2022 pode ter levado a uma melhora nas médias dos estudantes, pois apesar da maioria dos estudantes ter respondido que em 2021 estudou de forma híbrida, Gráfico 6, a maioria dos estudantes continuou seus estudos, Gráfico 7, e se sentiam um pouco preparado, Gráfico 12, na condução do processo e aprendizagem.

Apesar de todos os indicadores do SAEB, que é um Sistema de Avaliação da Educação Básica brasileira, e do IDEB, que é o Índice de Desenvolvimento da Educação Básica brasileira feita a cada dois anos, terem caído em 2021, retratando assim as dificuldades de aprendizado obtidas pelos estudantes no período de pandemia, taxas essas atreladas a evolução nas respostas ao teste do SAEB e do IDEB, relacionado ao insucesso do estudante no período, onde as médias são mais acentuadas no terceiro do ensino médio, a pesquisa nos mostra avanços discretos nos eixos temáticos da avaliação do Enem de 2022 em relação a 2019, razões essas que podem estar atreladas a pontos específicos de preparação para o exame, uma vez que se trata da porta de entrada para o ensino superior.

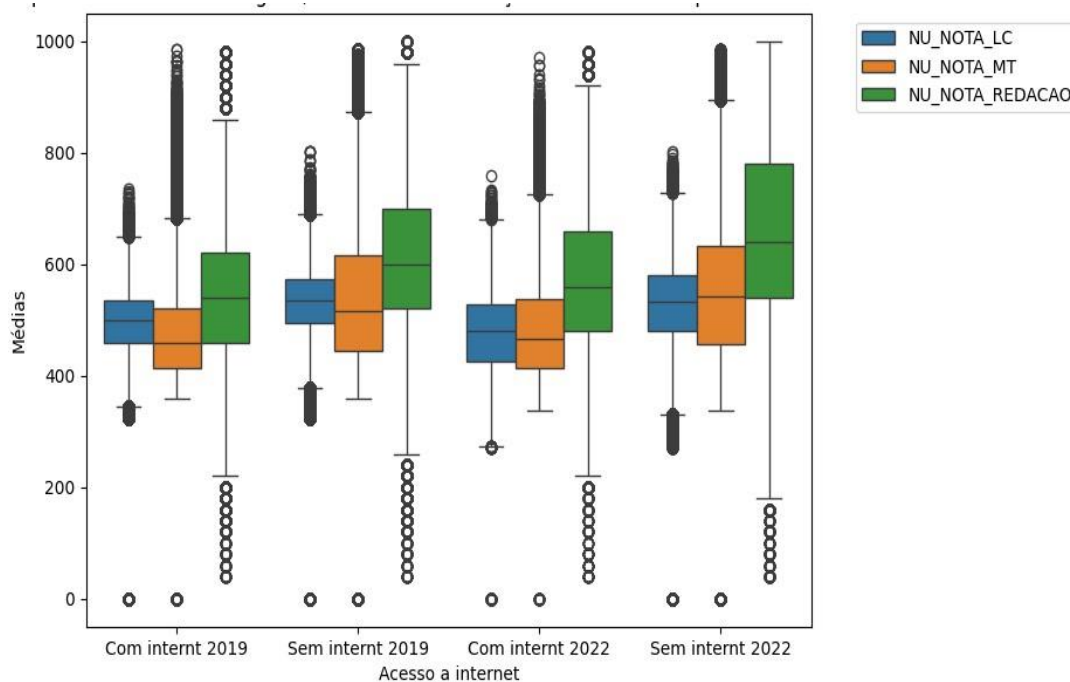
A distribuição das variáveis a direita, do Gráfico 14, onde estão localizadas as variáveis das matérias estão de acordo a Tabela 10 abaixo:

Tabela 10. Distribuição das variáveis das disciplinas do Gráfico 14.

Variável	Descrição
NU_NOTA_LC	Linguagens e códigos
NU_NOTA_MT	Matemática
NU_NOTA_REDAÇÃO	Redação

Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Gráfico 14. Distribuição das médias das notas e português, matemática e redação de acordo com o acesso à internet.



Fonte: Elaboração própria baseado nos microdados do Enem (2023).

Segundo o Inep¹², o meio mais utilizado para estudar ou manter-se informado durante a pandemia de Covid em 2021, foi o aparelho celular, com cerca de 64% dos estudantes fazendo uso dessa ferramenta, seguida de computador/notebook, com 22%. Já no que diz respeito às dificuldades de infraestrutura, mais da metade dos participantes, cerca de 54%, relataram dificuldades com equipamentos (computador/notebook) pouco disponíveis, por ter que dividir com algum familiar etc., no entanto, atrelado a isso, mais de 26% relataram dificuldades de acesso à internet, como já observado em gráficos e números anteriores.

Diante disso, no Gráfico 14, é notório que alunos que não possuem acesso as essas tecnologias possuem maiores dificuldades em obter médias mais altas nas áreas de conhecimentos do Enem, o intrigante é que apesar dos relatos, em 2022, ano pós pandemia, as notas dos candidatos no Enem foram melhores para os candidatos que não possuíam acesso as essas ferramentas e tecnologias.

Ainda sobre o gráfico acima, os resultados trazem referência a duas das principais disciplinas estudadas e pesquisadas nos relatórios de pesquisa de desempenho estudantil do país, além da redação, que possui nota específica e separada das demais notas no exame do Enem. Diante disso, e com relação ao ensino médio, em 2021 a evolução da proficiência média na disciplina de matemática do SAEB, houve uma queda de 7 pontos com relação a 2019, é de 3 pontos em português com relação ao mesmo ano, havendo uma queda, por tanto, em todos os indicadores da educação básica do SAEB e IDEB do país nesse período.

6 CONSIDERAÇÕES FINAIS

Com a pandemia de Covid-19 ocorrida entre os anos de 2020 e 2023, a educação brasileira passou por um grande processo de mudanças e adaptações para não parar o processo de ensino aprendizagem à população educacional do país, que em 2021 eram mais de 7 milhões de estudantes só no ensino médio, distribuídas em mais de 29 mil escolas, onde 92% das escolas de educação básica do Brasil adotaram estratégias de mediação, escola aluno, remota ou híbrida, Inep (2021).

O Enem, exame aplicado pelo Inep muito importante para os estudantes do último ano do ensino médio, e para indicadores da educação básica como um todo, teve que passar por mudanças durante o período de pandemia, com a implementação da possibilidade de realização das provas por meios eletrônicos, o Enem digital, e a mudança nas datas de aplicação no ápice da pandemia, sendo o Enem de 2020 aplicado em 2021, mantendo assim, o sonho vivo de muitos estudantes de se ingressarem em uma universidade.

Diante do disposto, este trabalho encontra-se enquadrado na aplicação de ciências de dados, dados educacionais brasileiros, e teve como objetivo a aplicação de técnicas estatísticas e da ciência de dados para descoberta de conhecimentos sobre os candidatos do Enem do ano de 2019 e 2020, bem como a reflexão sobre o questionário de hábitos de estudo de educação praticados pelos estudantes durante o período de pandemia de Covid.

Inicialmente, a análise referente aos candidatos do Enem de 2019, mostrou que a quantidade de inscritos no exame tinha sido umas das menores até então, assim

12

<https://app.powerbi.com/view?r=eyJrIjoiaWVhYTY4MWUtOWUzOS00MWEyLTNmOTctMzdhdjU0GU3ZiA0liwldiCl6lI2Zic0dk3LWM4YWMTNgIxZS05NzhmLWVhNGMwNzc0MzRiZiJ9>

como poucos candidatos foram considerados ausentes neste ano, no entanto, com a chegada da pandemia, o número de inscritos despencou e os candidatos considerados ausentes aumentaram proporcionalmente. Um fato interessante de observar nas análises entre os dois períodos é que as médias gerais dos candidatos aumentaram, mesmo sobre relatos de dificuldades de aprendizado, infraestrutura e relacionamento.

A base de dados do Enem de 2019 e 2022 possuem 76 colunas cada, um decréscimo em relação a outros anos, antes da Lei Geral de Proteção de Dados (LGPD) entrar em vigor, havendo várias variáveis de grupos temáticos que questionam a vida socioeconômica, educacional familiar de vários candidatos, em 2022 esse tipo de questionário evolui para tratar de assuntos educacionais referente a pandemia, havendo mais de 80 questionamentos sobre tal. Por tanto, esse trabalho se limitou a realizar as análises focando no panorama dos resultados do Enem em meio a pandemia de Covid-19, deixando assim muitas variáveis, já estudadas em bibliografias recentes, sobre vários outros conteúdos que poderiam ser explorados.

REFERÊNCIAS

AGGARWAL, Charu C. **Data Mining: The Textbook. 1. ed. New York, USA: Springer, 2015. v. 1E-book.** Disponível em: <https://doi.org/10.1007/978-3-319-14142-8>. disponível em: <https://downloads.editoracientifica.com.br/articles/210303626.pdf>. acesso em 09 de nov. de 2023.

ALQAHTANI, Ammar Y; RAJKHAN, Albraa A. **E-learning critical success factors during the covid-19 pandemic: a comprehensive analysis of e-learning managerial perspectives.** Education Science, MDPI, v.10,n.9, p.2016, ago,2020. Disponível em: <https://www.mdpi.com/2227-7102/10/9/216>. Acesso em 03 de novembro de 2023.

BRASIL, **Sistema de seleção unificada - SISU**, vagas. Disponível em: <https://sisu.mec.gov.br/#!/vagas?categoria=curso&id=37#target>. acesso em 13 de nov. de 2023.

CETAX. **A diferença entre ciência de dados e análise de dados. 2022.** Disponível em: [A diferença entre ciência de dados e análise de dados \(cetax.com.br\)](https://cetax.com.br). Acesso em: 08 nov. de 2023.

CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL, promulgada em 5 de outubro de 1988. Disponível em: https://www2.senado.leg.br/bdsf/bitstream/handle/id/518231/CF88_Livro_EC91_2016.pdf. Acesso em: 01 de novembro de 2023.

COSTA, Paulo Roberto da. **Rede e-tec Brasil, Estatística**, Santa Maria, RS, 2011. Disponível em: https://www.ufsm.br/app/uploads/sites/413/2018/11/04_estatistica.pdf. acesso em 09 de nov. de 2023.

em: <https://towardsdatascience.com/8-fundamental-statistical-concepts-for-data-science-9b4e8a0c6f1c>. Acesso em: 08 nov. 2023.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002. disponível em: https://files.cercomp.ufg.br/weby/up/150/o/Anexo_C1_como_elaborar_projeto_de_pesquisa_-_antonio_carlos_gil.pdf. acesso em 05 de dez. de 2023

GONÇALVES, P. **Afinal, como se desenvolve um projeto de Data Science?** 2018. Disponível em: [Afinal, como se desenvolve um projeto de Data Science? | by Pollyanna Gonçalves | TechBlogHotmart | Medium](#). Acesso em: 22 nov. 2023.

Gularte, F. N.; Nascimento, F. K. V. do; Carvalho, A. V. de. **ENEM em tempos de pandemia: desafios enfrentados pelos docentes frente o ensino remoto da rede estadual de Presidente Kennedy/TO**. Research, Society and Development, v. 10, n. 14, p.e571101422339, nov. 2021.

INEP, **apresentação dos microdados foi reformulada visando a plena aderência à Lei Geral de Proteção de Dados Pessoais**. Arquivos podem ser acessados no portal do Inep. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/institucional/inep-publica-microdadosdo-enem-2020-e-do-censo-escolar-2021>. Acesso em 14 de nov. de 2023.

INEP, **Brasil, Resultado Enem 2019**, Disponível em: https://download.inep.gov.br/educacao_basica/enem/downloads/2019/apresentacao_resultados_enem_2019.pdf. Acesso em 21 de nov. de 2023.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP), **Exame Nacional do Ensino Médio (ENEM)**. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enem>. Acesso em 01 de novembro de 2023.

Instituto Alicerce (2022). **As principais consequências da pandemia na educação**. Disponível em: https://blog.institutoalicerceedu.org.br/as-principais-consequencias-da-pandemia-na-educacao/?gad_source=1&gclid=Cj0KCQiAj_CrBhD-ARIsAliMxT8mDKaZLR1lt0ejSCn0yNNAHYtt0m8O-CpTA3N-doSDiVVleGjhLjkaArDHEALw_wcB. Acesso em 15 de dez. 2023.

Komorowski, M.; Marshall, D. C.; Saliccioli, J. D.; Crutain, Y. **Exploratory data analysis. Secondary Analysis of Electronic Health Records**, Springer International Publishing, p. 185–203, 2016. Disponível em: https://www.ncbi.nlm.nih.gov/books/NBK543630/pdf/Bookshelf_NBK543630. acesso em 15 de nov. de 2023

LEI Nº 9.394 DE 20 DE DEZEMBRO DE 1996. Estabelece as diretrizes e bases da educação nacional. Disponível em: http://portal.mec.gov.br/seesp/arquivos/pdf/lei9394_ldbn1.pdf. Acesso em 01 de novembro de 2023.

LERIA Lucinda A. *et al.* **O acesso do estudante com deficiência visual à educação superior: análise dos microdados do exame nacional do ensino médio (ENEM).** Disponível em: <https://www.scielo.br/ij/edur/a/zMjBLWmLQHRBnYxY53XqC4x/>. Acesso em 06 de agosto de 2023.

LIMA, Caio César Viana de; BRIGHENTI, Carla Regina Guimarães. **Desempenho de estudantes de Minas Gerais no Exame Nacional do Ensino Médio considerando variáveis socioeconômicas.** Universidade Federal de São João Del Rei - MG. Disponível em: <https://www.scielo.br/ij/ep/a/bKqgdWptzcZP4573NKgTZtt/?format=pdf&lang=pt>. Acesso em 31 de ago. de 2023

LINHA DO TEMPO DO CORONA VIRUS NO BRASIL. Disponível em: <https://www.sanarmed.com/linha-do-tempo-do-coronavirus-no-brasil>. Acesso em 01 de novembro de 2023.

MINAYO, Maria Cecília de S.; SANCHES, Odécio. **Quantitativo-qualitativo: oposição ou complementaridade?** *Cadernos de Saúde Pública*, Rio de Janeiro, v. 9, n. 3, p. 237-248, jul./set. 1993. <<https://doi.org/10.1590/S0102-311X1993000300002>>. disponível em: <https://www.scielo.br/ij/csp/a/Bgpmz7T7cNv8K9Hg4J9fJDb/?format=pdf&lang=pt>. acesso em 05 de dez. de 2023

NETO Nelson Weber, SOARES Raimundo C., COUTINHO Luciano Reis, TELES, Ariel Soares.: **A Pandemia da COVID-19 impactou o ENEM? Uma Análise Comparativa de Dados dos Anos de 2019 e 2020.** Disponível em: <https://seer.ufrgs.br/index.php/renote/article/view/126655/86731>. Acesso em 03 de novembro de 2023.

ORGANIZAÇÃO PAN-AMERICANO DA SAÚDE. **Histórico da pandemia da covid 19.** Disponível em: <https://www.paho.org/pt/covid19/historico-da-pandemia-covid-19>. Acesso em 01 de novembro de 2023.

PYTHON SOFTWARE FOUNDATION. Python Language Site: Documentation, 2020. Disponível em: <https://docs.python.org/pt-br/3/tutorial/>. Acesso em 13 de nov de 2023

Reis, E.A., Reis I.A. (2002) **Análise Descritiva de Dados. Relatório Técnico do Departamento de Estatística da UFMG.** Disponível em: <https://www.est.ufmg.br/portal/wp-content/uploads/2023/01/RTE-02-2002.pdf>. Acesso em 13 de nov. de 2023

ROCHA, Francisco Bruno Nascimento da. *et al.* **Análise dos microdados de matemática do ENEM de 2017 a 2019 do Nordeste.** Research, Society and Development, v. 11, n. 10, e207111032716, 2022.

Rodrigues, R. L.; Ramos, J. L. C.; Silva, J. C. S.; Gomes, A. S. **A literatura brasileira sobre mineração de dados educacionais.** Workshops do Congresso Brasileiro de Informática na Educação, Sociedade Brasileira de Computação - SBC, nov. 2014. Disponível em: <http://milanesa.ime.usp.br/rbie/index.php/wcbie/article/view/3286>. Acesso em 06 de novembro de 2023.

Senado Federal (2020). **Epidemia diminui a renda dos brasileiros.** Disponível em: <https://www12.senado.leg.br/institucional/datasenado/publicacaodatasenado?id=epidemia-diminui-renda-de-brasileiros>. Acesso em 15, dez. de 2023.

SIQUEIRA, Daniel, (2023). **Histograma: o que é, exemplos, gráficos e tipos.** Disponível em: <https://www.alura.com.br/artigos/o-que-e-um-histograma#:~:text=O%20objetivo%20de%20um%20histograma,em%20torno%20deste%20valor%20central>. Acesso em 10 de nov. de 2023.

Soares, R. D. C. et al. **Mineração de dados da educação básica brasileira usando as bases do INEP: Uma revisão sistemática da literatura.** RENOTE, UFRGS, v. 19, n. 1, p. 361–370, jul. 2021. Disponível em: <https://seer.ufrgs.br/index.php/renote/article/view/118526>. Acesso em 06 de novembro de 2023

Souza, Thiago Oliveira de. **Análise de dados: um estudo do perfil dos participantes do enem 2019 / Thiago Oliveira de Souza. - 2021. 44 f. : il.** Disponível em: https://repositorio.ufersa.edu.br/bitstream/prefix/6916/1/ThiagoOS_MONO.pdf. Acesso em 08 de junho de 2023.

UOL, **Universidades públicas que aceitam o enem**, por Miguel Souza, 2023. Disponível em: <https://vestibular.brasilecola.uol.com.br/enem/universidades-publicas-que-aceitam-o-enem.htm>. Acesso em 07 de novembro de 2023.

VICKERY, R. **8 Fundamental Statistical Concepts for Data Science.** 2021. Disponível em: <https://towardsdatascience.com/8-fundamental-statistical-concepts-for-data-science9b4e8a0c6f1c>. acesso em 06 de dez. de 2023