

**UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO - UFTM**  
**Programa de Mestrado em Matemática - PROFMAT**

**Jorge Leonardo Garcia**

**Busca de novos indicadores de aprendizado em Matemática através da  
comparação entre as bases de dados do Enem 2017 e Saeb 2017**

**Uberaba**  
**2019**

Jorge Leonardo Garcia

**Busca de novos indicadores de aprendizado em Matemática através da  
comparação entre as bases de dados do Enem 2017 e Saeb 2017**

Dissertação apresentada ao Programa de Mestrado  
em Matemática - PROFMAT da Universidade Federal  
do Triângulo Mineiro, como requisito parcial  
para obtenção do título de Mestre em Matemática.

Orientador: Prof. Leandro Cruvinel Lemes

Coorientador: Prof. Nome do Avaliador

Área de concentração: Mineração de dados - Educação

Uberaba

2019

**Catálogo na fonte: Biblioteca da Universidade Federal do  
Triângulo Mineiro**

G199b Garcia, Jorge Leonardo  
Busca de novos indicadores de aprendizado em Matemática  
através da comparação entre as bases de dados do Enem 2017 e  
Saeb 2017/ Jorge Leonardo Garcia. -- 2019.  
38 f. : il., graf., tab.

Dissertação (Mestrado Profissional em Matemática em Rede  
Nacional) -- Universidade Federal do Triângulo Mineiro, Uberaba,  
MG, 2019

Orientador: Prof. Dr. Leandro Cruvinel Lemes

1. Matemática (Ensino médio). 2. Estudantes - Avaliação. 3.  
Mineração de dados (Computação). I. Lemes, Leandro Cruvinel. II.  
Universidade Federal do Triângulo Mineiro. III. Título.

CDU 51:37.091.279.7


JORGE LEONARDO GARCIA

**Busca de novos indicadores de aprendizado em Matemática  
através da comparação entre as bases de dados do Enem 2017 e  
Saeb 2017**

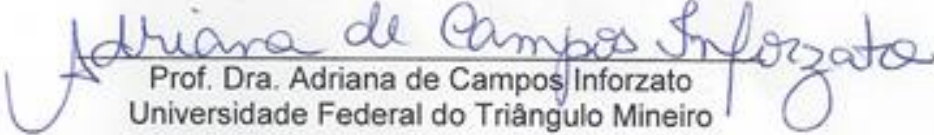
Dissertação apresentada ao Programa de Mestrado Profissional em Matemática em Rede Nacional, da Universidade Federal do Triângulo Mineiro, como requisito parcial para obtenção do título de Mestre em Matemática.

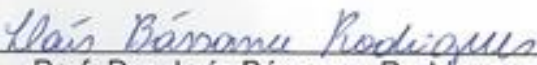
Uberaba, 04 de Outubro de 2019.

BANCA EXAMINADORA

  
Prof. Dr. Leandro Cruvinel Lemes  
Orientador

Universidade Federal do Triângulo Mineiro

  
Prof. Dra. Adriana de Campos Inforzata  
Universidade Federal do Triângulo Mineiro

  
Prof. Dra. Laís Bássame Rodrigues  
Universidade Federal de Uberlândia

## **AGRADECIMENTOS**

Agradeço ao professor Leandro Cruvinel pelos ensinamentos passado durante o curso, teve paciência, me incentivando e o mais importante, acreditou em mim a todo momento.

Ao Dream Team que não dei atenção o suficiente durante esse árduo período e mesmo assim estavam do meu lado, me apoiando, puxando a orelha e dando conselhos.

Aos amigos e familiares que me aguentaram reclamando, principalmente aqueles que me deram forças nos momentos difíceis, que disseram “você consegue”, que me ajudaram com correções, que me tirou várias vezes de momentos tristes, meu sincero obrigado.

## RESUMO

O presente trabalho apresenta uma análise descritiva dos conjuntos de dados do Enem 2017 e do Saeb 2017, obtidos no portal de dados abertos do Inep. Essa análise é limitada às notas de Matemática de candidatos oriundos de Escola Pública e visa determinar distinções entre grupos de candidatos por geolocalização, por desempenho nas provas e por assertividade de tópicos. Neste estudo, são propostos novos indicadores para diagnóstico de possíveis fragilidades dos alunos de Matemática de escolas públicas Brasileiras.

Palavras-chave: Enem; Saeb; Mineração dos dados; Educação Matemática.

## **ABSTRACT**

This document presents a descriptive analysis of the datasheets from Enem 2017 and Saeb 2017, obtained from Inep's open data portal. This analysis is limited to the math scores of public school candidates and aims to determine distinctions between candidate groups by geolocation, exam performance, and topic assertiveness. In this article, new indicators are proposed to diagnose possible weaknesses of mathematics students from Brazilian public schools.

Keywords: Enem; Saeb; Data mining; Mathematical education.

## LISTA DE FIGURAS

Figura 1 - Questão 167 (Caderno Amarelo Enem 2017).....	26
Figura 2 - Alunos com menos de 7 acertos (Caderno azul) - Enem.....	32
Figura 3 - Alunos com 40 acertos ou mais (Caderno Azul) - Enem.....	32



## **LISTA DE TABELAS**

Tabela 1 - Piores notas estaduais .....	24
Tabela 2 - Maiores notas estaduais .....	25
Tabela 3 - Matriz de referência 3ºano E.M - Saeb .....	35
Tabela 4 - Matriz de referência E.M - Enem.....	37

## LISTA DE GRÁFICOS

Gráfico 1 - Notas Saeb 2017 Particular x Estadual .....	20
Gráfico 2 - Enem 2017 x Saeb 2017 .....	24
Gráfico 3 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) .....	25
Gráfico 4 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) .....	26
Gráfico 5 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Centro-Oeste .....	27
Gráfico 6 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Centro-Oeste .....	27
Gráfico 7 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Nordeste .....	28
Gráfico 8 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Nordeste .....	28
Gráfico 9 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Norte .....	28
Gráfico 10 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Norte .....	29
Gráfico 11 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Sudeste.....	29
Gráfico 12 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Sudeste.....	29
Gráfico 13 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Sul .....	30
Gráfico 14 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Sul .....	30
Gráfico 15 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Amapá x Rio Grande do Sul .....	31
Gráfico 16 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Amapá x Rio Grande do Sul .....	31
Gráfico 17 - índice de acertos escola pública em Matemática (Enem 2017).....	32

## SUMÁRIO

1. INTRODUÇÃO.....	11
2. BIG DATA.....	13
2.1. KDD.....	13
2.2. Data-Driven Decision Making.....	15
3. SED .....	16
3.1. Saesp e A.A.P.....	17
4. AVALIAÇÕES NACIONAIS .....	19
4.1. Saeb.....	19
4.2. Enem.....	20
5. METODOLOGIA .....	22
6. RESULTADOS .....	24
6.1. Resultados Enem.....	27
7. CONCLUSÃO .....	33
8. REFERÊNCIAS .....	34
9. Anexos.....	35

## 1. INTRODUÇÃO

Com a aprovação da Lei n. 9394/96 (Brasil, 1996), foram instauradas no Brasil avaliações nacionais, dentre elas o Saeb (Sistema de Avaliação da Educação Básica) em 1.997 e o Enem (Exame Nacional do Ensino Médio) em 1.998, que possuem matrizes de referência com competências e habilidades muito similares.

Ambas as avaliações são realizadas pelo Inep (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) que disponibiliza os dados em formato CSV para download no próprio site (Inep, acesso 28/11/2018), contendo informações dos candidatos como: número de inscrição, ano da avaliação, sexo, cor, localização (municípios, estado e região), notas, gabaritos dos participantes, gabarito da avaliação, códigos das escolas, entre outros. Considerando os dados apenas de 2.017, ano escolhido para realização desta pesquisa, somam-se mais de 5 Gigabytes de informações. A necessidade de técnicas não convencionais de computação para realizar a análise, devido ao volume de dados, é caracterizada como um problema de Big Data, termo popularizado por John Mashey em meados dos anos 1.990.

A prova do Enem avalia o desempenho do estudante no ensino médio e, atualmente, é utilizada para acesso à educação superior por programas como o SiSU (Sistema de Seleção Unificada) e ProUni (Programa Universidade para Todos). Os resultados dessa avaliação são, ainda, critérios para financiamentos estudantis como o FIES (Fundo de Financiamento Estudantil). O exame é composto por 4 provas objetivas com 45 questões em cada uma das seguintes áreas: Linguagens, Códigos e suas Tecnologias, Ciências Humanas e suas Tecnologias, Ciências da Natureza e suas Tecnologias e Matemática e suas Tecnologias. Além de uma redação de texto dissertativo-argumentativo que completa a avaliação.

O Saeb (Sistema de Avaliação da Educação Básica) tem como objetivo obter um diagnóstico em larga escala. Um dos seus instrumentos avaliativos é a Prova Brasil que, no ano de 2.017, foi aplicada em 73.027 escolas, sendo que 59.388 tiveram os dados divulgados pelo portal Inep. A Prova Brasil avalia a qualidade oferecida pelo sistema educacional brasileiro. A prova é dividida em dois blocos de Língua Portuguesa e dois blocos de Matemática, sendo geralmente aplicada para alunos do 5º e 9º ano do ensino fundamental e para 3º ano do Ensino Médio (contendo 22 questões de Língua Portuguesa e 22 questões de Matemática).

O presente trabalho objetiva apresentar novos indicadores que ofereçam suporte para o planejamento didático do governo brasileiro. Para isso, as informações do principal instrumento de avaliação da educação brasileira, o Saeb, são comparadas com as informações do Enem, de forma a destacar diferenças que ainda não são consideradas como indicadores atuais. A comparação é feita através de técnicas de mineração de dados.

## 2. BIG DATA

O termo Big Data está diretamente relacionado ao conceito de volume de dados, que, popularizado por John Mashey (Mashey, 1998) em meados dos anos 1990, excede a capacidade de processamento e análise por meios convencionais. Anos depois, o conceito ganhou força por Doug Laney (Laney, 2001) que articulou o conceito inicialmente com 3vs : Volume, Velocidade e Variedade, e posteriormente adicionando outros 2 vs: Veracidade e Valor.

1. Volume: está relacionado com a quantidade de dados gerados, analisados e armazenados;
2. Velocidade: de coleta, análise e criação a qualquer momento;
3. Variedade (de dados): blogs, mídias sociais, atualizações, fotos, vídeos, arquivos de áudio, etc;
4. Veracidade: A confiabilidade dos dados em que estão sendo coletados;
5. Valor: da extração dos dados;

As informações nestes arquivos podem ser divididas em:

- Informações estruturadas: Possuem padrões ou formatos utilizados na leitura dos dados, como por exemplo arquivos em formato .CSV.
- Informações não estruturadas: Não possuem formatos padronizados para leitura, fotos e vídeos são exemplos.

Alguns exemplos de técnicas exploratórias de Big Data são: CRISP-DM (Cross Industry Standard Process for Data Mining), SEMMA (Sample, Explore, Modify, Model e Assess) e KDD (Knowledge Discovery in Databases) que será utilizado neste trabalho.

### 2.1. KDD

Para Fayyad (1996) o KDD (Knowledge Discovery in Database ou descoberta de conhecimento em banco de dados) é o processo geral da descoberta de conhecimento útil de dados durante a mineração dos dados. As etapas que constituem o KDD são: seleção, pré-processamento, transformação, mineração dos dados, interpretação e validação.

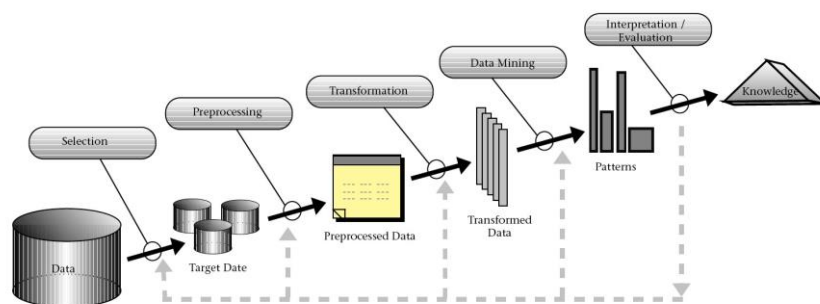
Segundo Fayyad (1996) as etapas para a realização do kdd são descritas

da seguinte forma:

1. A identificação da meta do KDD para o cliente;
2. A seleção um conjunto de dados ou um conjunto de variáveis;
3. Limpeza e pré-processamento: remoção de ruídos e a utilização de estratégias para lidar com falta de parte dos dados;
4. Redução e projeção de dados, ou seja, utilização de métodos de transformação com intuito de redução o número de blocos de variáveis;
5. Juntamente ao processo de KDD, são aplicados outros métodos particulares de mineração dos dados, por exemplo: a sumarização, classificação, regressão e agrupamento;
6. Escolha do algoritmo de mineração de dados: seleção de métodos, modelos e parâmetros a serem utilizados em busca de padrões nos dados;
7. Mineração de dados: para encontrar padrões de interesse;
8. Interpretação dos padrões minerados: pode ser necessário o retorno a qualquer uma das etapas anteriores para complemento;
9. Consolidação do conhecimento descoberto: incorporando esse conhecimento em outro sistema, outras ações, ou simplesmente apenas documentá-la.

A Figura 1 ilustra todas as as etapas do processo do KDD .

Figura 1 – Etapas do processo KDD



Fonte: Fayyad, Pietetsky-Shapiro, & Smyth (1996)

Com os resultados obtidos a partir do KDD, pode-se tomar decisões a partir destes dados, essa prática é conhecida como Data-Driven Decision Making.

## 2.2. Data-Driven Decision Making

Para Erik Brynjolfsson (2011) o Data-Driven Decision Making refere-se à prática de basear as decisões em torno de análise dos dados e não apenas à intuição.

Em 2013, o economista Erik Brynjolfsson juntamente com seus colegas do MIT e da escola Wharton da Penn realizaram um estudo sobre Data-decision making, classificando empresas de acordo com o quanto uma empresa utiliza os dados para a tomada de decisões (Provost & Fawcett, 2013) o resultado foi: quanto mais a empresa é orientada por dados, maior seu rendimento, chegando a um aumento de 4 a 6% na produtividade.

Outro exemplo de Data-Driven Decision Making é a prática utilizada pela equipe IBM Watson Education (Watson), chamada de Data-Driven Education, focada em melhorar resultados de aprendizagem, dando a professores ferramentas digitais para que eles sejam mais eficazes e ajudem os alunos a desempenharem o máximo de habilidades.

Alguns projetos que a equipe IBM Watson Education oferece como exemplo de Data-Driven Education são:

- Aplicar Conteúdo personalizado para estudantes com base no domínio com lições, atividades e tarefas baseadas no nível que estão hoje;
- Melhorar o desenvolvimento do vocabulário na infância com um App de aprendizagem de vocabulário que possa identificar áreas que exigem um foco adicional;
- Tutorial baseado em IA (Inteligência Artificial) para estudantes, de modo com que os professores/instrutores acompanhem o progresso dos alunos;

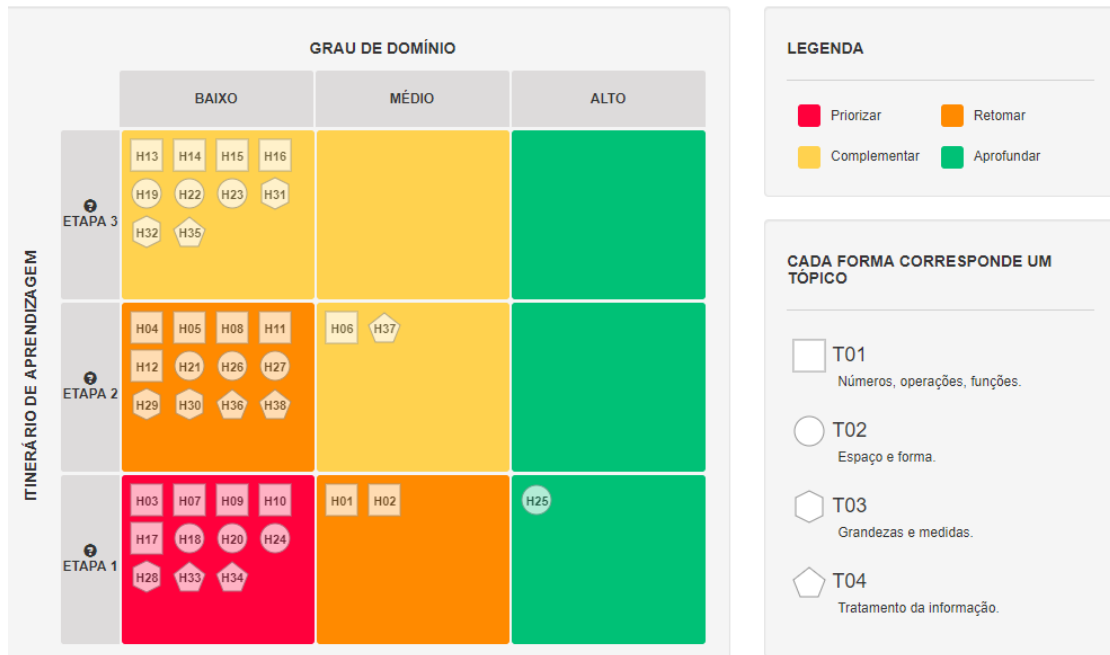
No estado de São Paulo, a SED (Secretaria Escolar Digital) é um exemplo de Data-Driven Education oferecido gratuitamente para o professor da rede estadual. A SED possui várias funções pedagógicas para a escola e o professor, uma delas é o SARA (Sistema de Acompanhamento dos Resultados de Avaliações), com ele é possível analisar quais as habilidades mais defasadas em Língua portuguesa e Matemática de uma determinada série/escola, com base nessa análise, são tomadas importantes decisões, por exemplo diagnosticar dificuldades de uma habilidade em um grupo de alunos, e a partir disso criar estratégias para auxiliar esse grupo.



### 3. SED

O Itinerário de Aprendizagem, assim chamado no sistema, favorece ao professor entender o comportamento que uma turma apresentou durante os anos anteriores, separado por habilidades e em qual nível cada habilidade está, de acordo com a turma ou aluno avaliado.

Figura 2 – Itinerário de aprendizagem.



Fonte: Secretaria da Educação do Estado de São Paulo (2017)

As figuras 2 e 3 ilustram o itinerário de uma sala de terceiro ano do Ensino médio de uma escola pública de Franca-SP. A legenda (no canto superior direito), mostra o nível em que cada habilidade a turma se encontra.

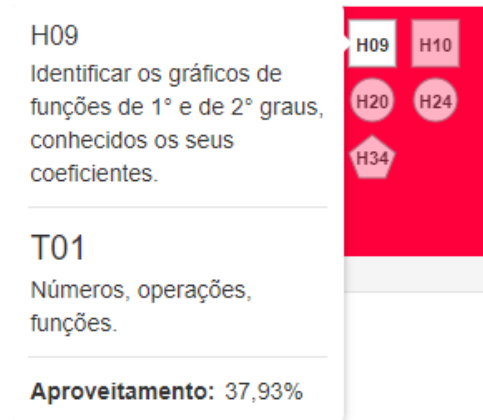
- Em verde, é necessário aprofundar o conteúdo;
- Em amarelo, é necessário complementar;
- Em laranja, é necessário retornar o conteúdo;
- Em vermelho, priorizar revisões para essa habilidade (estado crítico);

Já no canto inferior direito, formas geométricas demonstram tópicos específicos identificando qual a habilidade necessária.

- T01 - Números, operações e funções;
- T02 - Espaço e forma;
- T03 - Grandezas e medidas;
- T04 - Tratamento de informação;

Esse sistema apresenta dados concernentes às séries do 6º ano do Ensino Fundamental até o 3º ano do Ensino Médio e tem como objetivo auxiliar o professor, a cada início de ano letivo, a identificar as possíveis dificuldades da turma em geral.

Figura 3 – Itinerário de aprendizagem - Habilidades



Fonte:Secretaria da Educação do Estado de São Paulo, 2019

O sistema é alimentado pelo resultado de duas avaliações estaduais, A.A.P (Avaliação da Aprendizagem em Processo), aplicada semestralmente e o Saesp (Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo), realizada anualmente.

### 3.1. Saesp e A.A.P

A Avaliação de Aprendizagem em Processo (AAP) é aplicada duas vezes durante o ano, cujo foco é diagnosticar o nível de aprendizagem dos estudantes a partir do 2º Ano do Ensino Fundamental. O caderno de perguntas e respostas é formado por uma redação, questões dissertativas e de múltipla escolha de Língua Portuguesa e Matemática, tendo como base o Currículo Oficial do Estado de São Paulo (Secretaria da Educação do Estado de São Paulo, 2019)

O Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo, mais conhecido como Saesp, é uma avaliação aplicada anualmente para alunos do 3º, 5º, 7º e 9º anos do Ensino Fundamental e da 3ª ano do Ensino Médio tendo o propósito de produzir diagnósticos, visando orientar ações de gestores do ensino no monitoramento das políticas voltadas para a melhoria da qualidade educacional

(Secretaria da Educação do Estado de São Paulo, 2019). A prova de múltipla escolha contém questões de Língua Portuguesa, Matemática, Ciências Humanas, Ciências da Natureza e redação. Além de alimentar a SED, o Saresp tem o importante papel de fazer parte do índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp).

Criado em 2007, o Idesp é um dos principais indicadores da qualidade do ensino na rede estadual paulista (Secretaria da Educação do Estado de São Paulo). Os objetivos de cada escola são projetados levando em consideração esse índice, estabelecendo metas a serem alcançadas a cada ano.

O Saresp e a A.A.P tem como intuito gerar valor com os dados obtidos através de suas avaliações abrangendo o Estado de São Paulo.

## **4. AVALIAÇÕES NACIONAIS**

No final da década de 1990, algumas avaliações nacionais começaram a fazer parte da carreira acadêmica dos alunos de escolas públicas como forma de avaliação do ensino público, Saeb (Sistema de Avaliação da Educação Básica) em 1.997 e Enem (Exame Nacional do Ensino Médio) em 1.998.

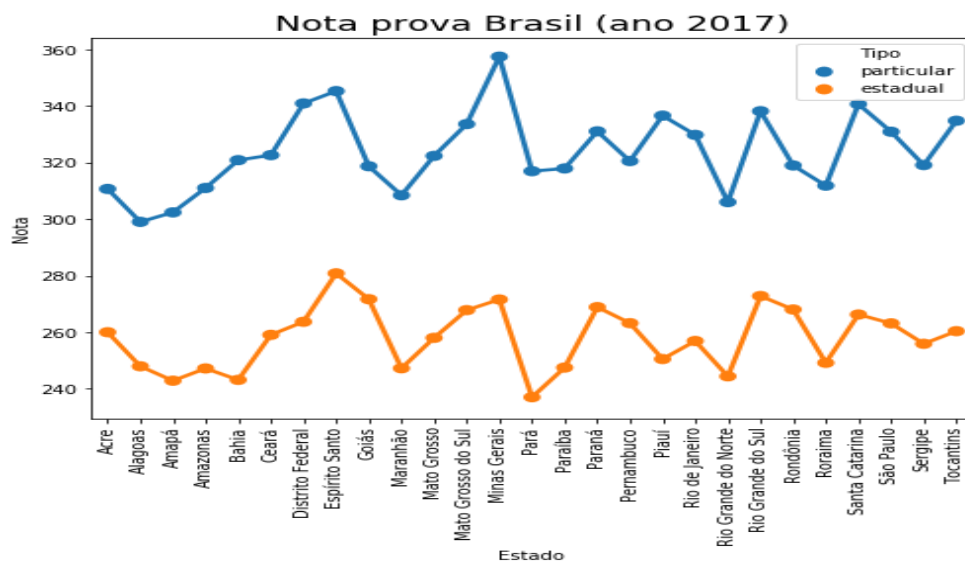
### **4.1. Saeb**

De acordo com o MEC/Inep, o SAEB (Sistema de Avaliação da Educação Básica) ou Prova Brasil, é uma avaliação nacional que tem como foco produzir diagnósticos em larga escala na educação básica.

Produzida pelo MEC/Inep, a avaliação segue as competências e habilidades das matrizes de referência (ANEXO 1) dividida em: dois blocos de Língua Portuguesa e dois blocos de Matemática, geralmente aplicada para alunos do 5º e 9º ano do Ensino Fundamental e para 3º ano do Ensino Médio (contendo 22 questões de Língua Portuguesa e 22 questões de Matemática).

A Prova Brasil é aplicada bianualmente, sendo em um ano a aplicação e no seguinte a divulgação dos resultados, avaliando o conhecimento dos alunos do Ensino Fundamental e Ensino Médio em Língua Portuguesa e Matemática das escolas estaduais, municipais e particulares (opcional para a escola). O caderno de questões da prova não é divulgado, seu conteúdo contém muitas habilidades e competências de Matemática similares às do ENEM, porém não há acesso às provas anteriores, apenas a questões “parecidas” para estudo.

Gráfico 1 - Notas Saeb 2017 Particular x Estadual



Fonte: Inep/Mec 2017

No ano de 2017 foi permitido a adesão à Prova Brasil para escolas particulares. O gráfico 1 apresenta a diferença das notas entre os ensinos particular e estadual.

## 4.2. Enem

O Enem (Exame Nacional do Ensino Médio) tem como proposta, apresentada pelo Ministério da Educação, ser utilizado como método de seleção unificada do processo seletivo das universidades públicas federais, abrindo portas de universidades.

Cada universidade possui quatro possibilidades para um exame seletivo:

- Como fase única, com o Sistema de Seleção Unificada (SiSU), informatizado e on-line;
- Como primeira fase;
- Combinado com o vestibular da instituição;
- Como fase única para as vagas remanescentes do vestibular.

O Enem atualmente é composto por uma redação e 180 questões (contendo cinco alternativas, sendo apenas uma a alternativa correta), divididas entre quatro áreas do conhecimento:

- Ciências Humanas e suas Tecnologias – 45 questões.
- Ciências da Natureza e suas Tecnologias – 45 questões.

- Linguagens, Códigos e suas Tecnologias – 45 questões.
- Matemática e suas Tecnologias – 45 questões
- Redação argumentativa

O exame é realizado anualmente, dividido em dois dias e tem papel fundamental para o acesso a universidades federais e particulares, pelo SiSU e ProUni (Programa Universidade para todos).

Destacando a importância dessas avaliações para aqueles que almejam o Ensino Superior, será analisado o conjunto dos dados do Saeb 2017 e Enem 2017.

## 5. METODOLOGIA

Os conjuntos de dados considerados (Enem 2017 e Saeb 2017) foram obtidos no portal Inep, sendo que os processos de extração de informação foram implementados com o uso da linguagem de programação *Python*. Em seguida, os códigos foram interpretados com uso do software *Jupyter notebook*. É importante destacar que tanto a linguagem, quanto o *software*, são ferramentas *open-source*. Foi instalado e executado em um computador com processador *Intel Core i5-4670*, CPU 3,40GHz, memória RAM 8 GB, com sistema operacional Windows 10 PRO.

A estratégia de extração de informação dos conjuntos de dados do Enem e Saeb foi o KDD (Knowledge Discovery in Database ou descoberta de conhecimento em banco de dados), processo que se realiza nas seguintes etapas: seleção, pré-processamento, transformação, mineração dos dados, interpretação e validação (Fayyad, Pietetsky-Shapiro, Smyth, 1996).

Na etapa de seleção, foram considerados do conjunto de dados do Enem 2017, os seguintes filtros: prova de Matemática (questões 136 até 180), alunos oriundos de escola pública que não são treineiros e que estavam presentes na prova de Matemática. Além disso, apenas as variáveis: código do município, respostas do aluno e nota do aluno foram consideradas neste trabalho. Nessa etapa, para que a memória RAM do computador pudesse armazenar o volume de dados sendo processado, o arquivo foi dividido em 10 partes.

Ainda na etapa de seleção, considerando o conjunto do Saeb 2017, foram aplicados os seguintes filtros: exclusão de escolas privadas, nota Saeb 2017 de Matemática, Sigla UF, município, código do município. Diferentemente do Enem, os dados encontrados no arquivo são pouco detalhados, ou seja, não contém a nota individual do aluno e nem seu gabarito. Aproximadamente, 89 municípios não apresentavam informações da avaliação do Saeb 2017. Essas cidades foram removidas do conjunto de dados na etapa de pré-processamento.

As respostas dos alunos que fizeram o Enem foram fornecidas através de uma sequência de 45 caracteres representando, respectivamente, as respostas desses participantes nas 45 questões da prova de Matemática. Na etapa de transformação, uma variável foi criada para cada resposta e a respectiva letra foi substituída pelo inteiro 1, caso o aluno tenha acertado, e pelo inteiro 0, caso o aluno tenha errado. A substituição de letras do alfabeto por inteiros faz com que o

computador use menos memória nos processos envolvidos. Ao fim desta etapa, as partes do arquivo principal, agora pré-processadas, filtradas e transformadas, foram novamente unidas em um único arquivo, uma vez que os processos envolvendo o volume total de dados já são comportados pela memória do computador utilizado.

Na etapa de mineração de dados, um dos objetivos foi a busca por visualizações mais resumidas dos dados. Para isso, fez-se o uso de gráficos de barras e estatísticas. Por fim, as informações mineradas foram interpretadas.



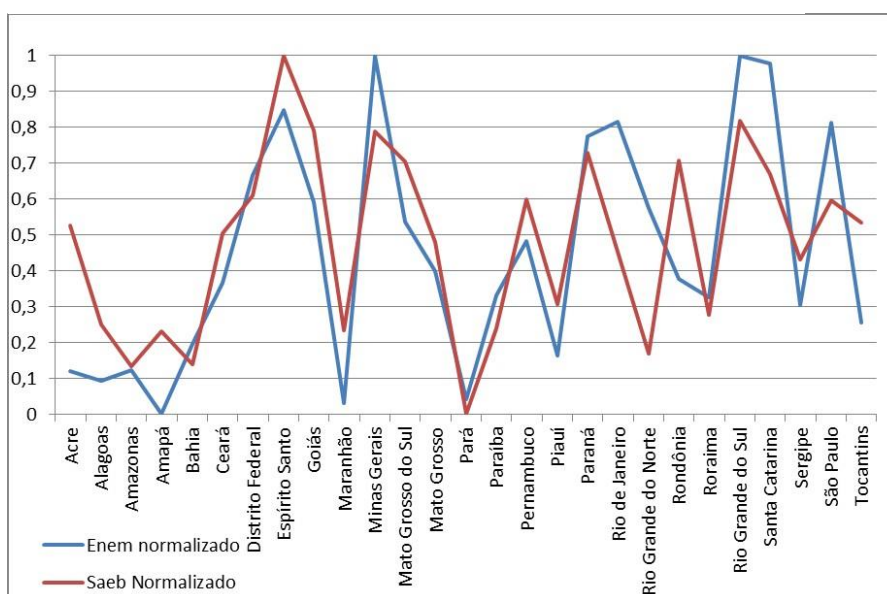
## 6. RESULTADOS

Considerando os registros, enquanto o Enem apresenta um total de 5.558 municípios, o Saeb possui 5.060. Com relação ao número de escolas, o Enem possui 23.072 e o Saeb possui 8.750. Enem abrangeu 99,61% destes e o Saeb 90,84%.

Observando os registros, nota-se que o Enem apresenta 5.558 municípios representados por estudantes de 23.072 escolas em comparação com o Saeb que contém 5060 municípios representados por estudantes de 8.750 escolas. Segundo o IBGE, o Brasil possui 5.570 municípios (IBGE, acesso em 01/09/2019), sendo assim o Enem abrangeu 99,61% destes e o Saeb 90,84%.

Comparando as notas normalizadas entre as médias dos estados brasileiros, percebe-se similaridade em vários pontos, pode-se observar como exemplo o caso do Pará, que contém baixo rendimento em ambas as avaliações. O gráfico 1 mostra a comparação das notas estaduais (Saeb e Enem) normalizadas.

Gráfico 2 - Enem 2017 x Saeb 2017



Fonte: Inep/Mec, 2017

As tabelas 1 e 2 indicam as três menores e as três maiores médias das notas.

Tabela 1 - Piores notas estaduais

Estados (Saeb)	Notas (Saeb)	Estados (Enem)	Notas (Enem)
Pará	237,30	Amapá	467,62
Amapá	242,86	Maranhão	469,23
Bahia	243,15	Pará	469,78

Fonte: Inep/Mec, 2017

Tabela 2 - Maiores notas estaduais

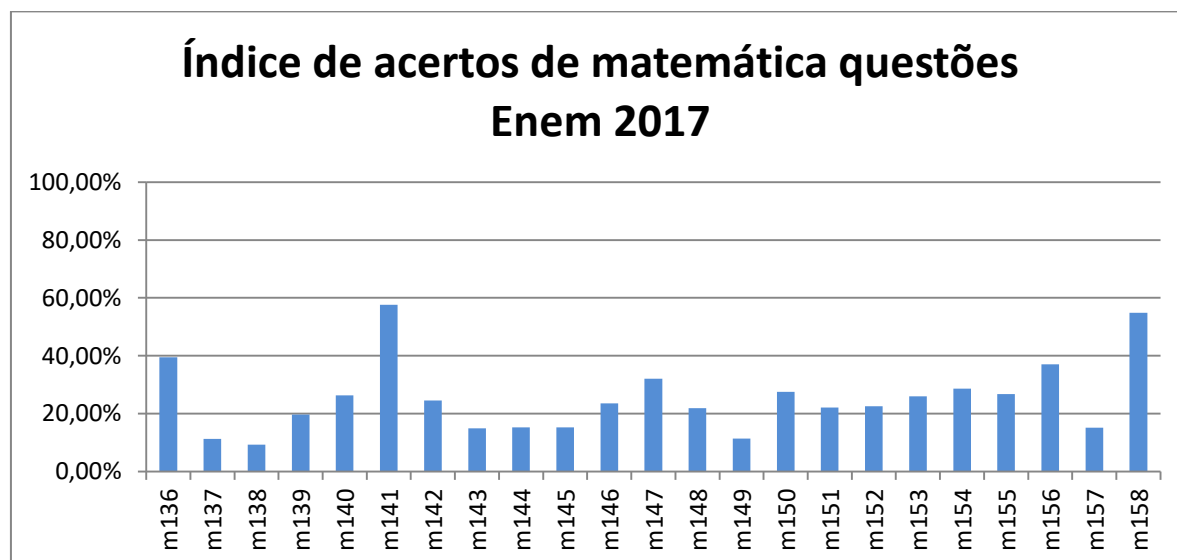
Estados (Saeb)	Notas (Saeb)	Estados (Enem)	Notas (Enem)
Espírito Santo	280,84	Rio Grande do Sul	518,21
Rio Grande do Sul	272,82	Minas Gerais	518,11
Goiás	271,70	Santa Catarina	517,01

Fonte: Inep/Mec, 2017

Os resultados obtidos do Saeb são limitados por conter apenas as notas referentes às escolas, pois tanto suas avaliações quanto o gabarito dos alunos não são publicados, impossibilitando análises mais detalhadas, porém as matrizes de referência das provas de Matemática de ambas as avaliações são feitas pelo Inep e possuem quase todas habilidades em comum (Inep, acesso 25/08/2019). Sendo assim, pela menor abrangência de dados fornecidos pelo Saeb, os registros do Enem podem ser uma alternativa viável para encontrar dificuldades apresentadas pelos alunos e a partir dessas, podem ser criadas estratégias e adoção de materiais específicos, personalizando o estudo de acordo com a necessidade dos alunos.

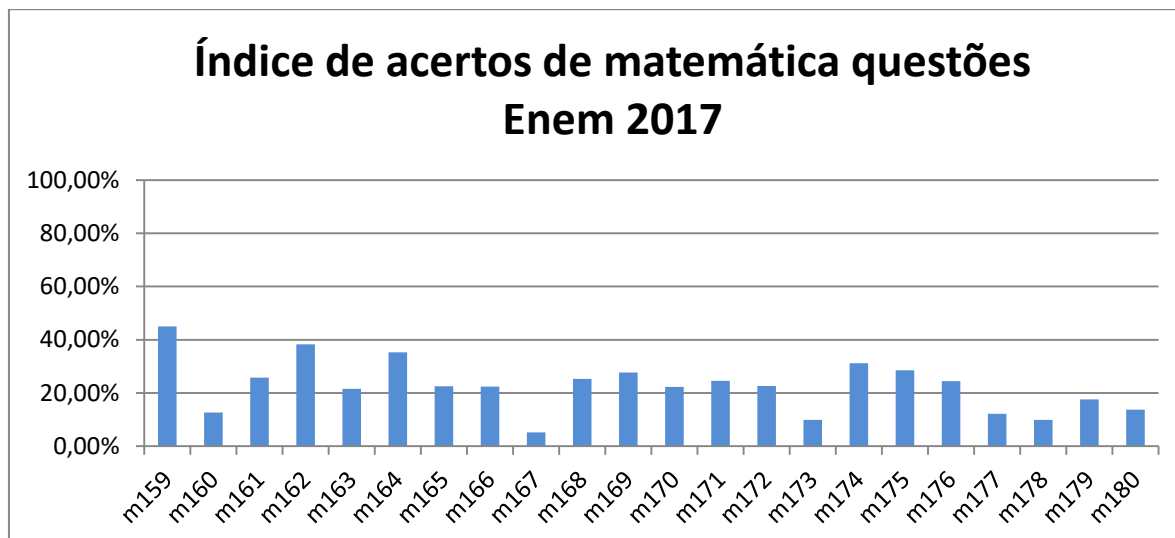
Um exemplo de análise é mostrado nos gráficos 2 e 3 que apontam o índice de acertos dos participantes. A prova amarela foi utilizada como legenda, ou seja, a coluna indicada por m136 corresponde a questão 136 da prova amarela e assim sucessivamente. Optou-se por esse registro porque as questões das provas de cores diferentes estão dispostas em locais diferentes.

Gráfico 3 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158)



Fonte: Inep/Mec, 2017

Gráfico 4 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180)



Fonte: Inep/Mec, 2017

Associando a dificuldade dos alunos com o índice de acerto, tem-se que a m167, ilustrada na Figura 1, foi considerada a questão mais difícil da prova, com quase 58000 (5,20%) acertos e com mais de um milhão de erros.

É possível, a partir destas informações, identificar focos alarmantes de dificuldade, por exemplo, as questões m138, m167, m173 e m178 (questões que envolvem razão, combinação e interpretação de gráficos) tiveram menos de 10% de acerto.

Figura 1 - Questão 167 (Caderno Amarelo Enem 2017)

**QUESTÃO 167**

O comitê organizador da Copa do Mundo 2014 criou a logomarca da Copa, composta de uma figura plana e o slogan "Juntos num só ritmo", com mãos que se unem formando a taça Fifa. Considere que o comitê organizador resolvesse utilizar todas as cores da bandeira nacional (verde, amarelo, azul e branco) para colorir a logomarca, de forma que regiões vizinhas tenham cores diferentes.



JUNTOS NUM SÓ RITMO

Disponível em: [www.pt.fifa.com](http://www.pt.fifa.com). Acesso em: 19 nov. 2013 (adaptado).

De quantas maneiras diferentes o comitê organizador da Copa poderia pintar a logomarca com as cores citadas?

- A 15
- B 30
- C 108
- D 360
- E 972

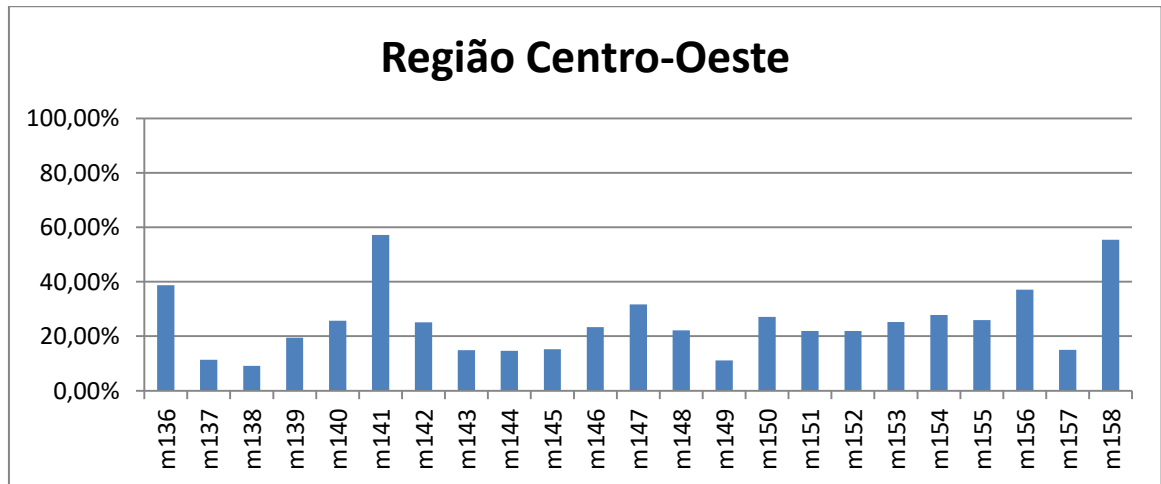
Fonte: Inep/Mec, 2017

Essas análises são relacionadas ao Brasil todo, com a aplicação dos filtros corretos podem-se obter resultados de uma região, estado, cidade ou até mesmo de uma escola específica, podendo assim ser uma alternativa para melhora de resultados de avaliações como o Saeb e o Enem.

**6.1. Resultados ENEM**

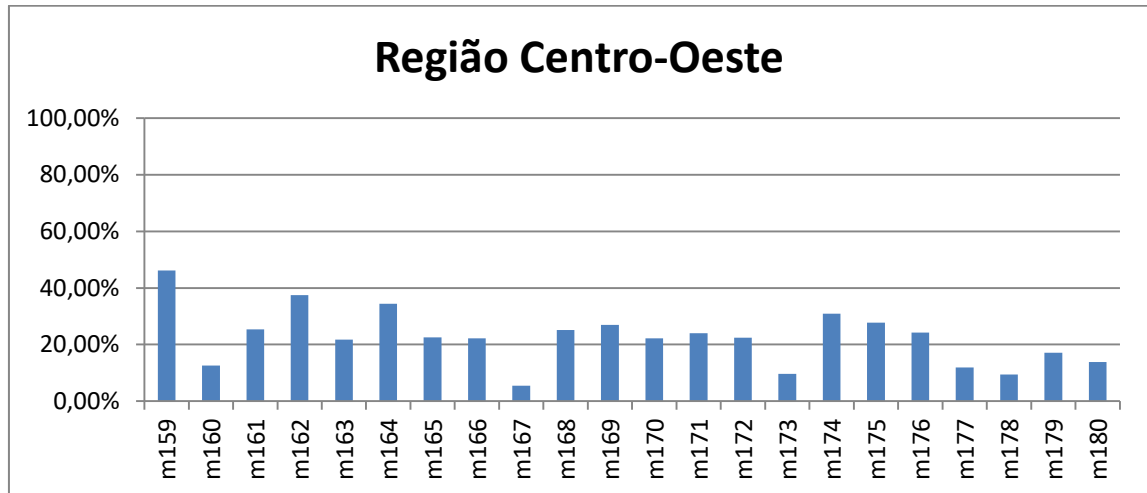
Utilizando os filtros dos gráficos 3 e 4, separando por região brasileira, temos os Gráficos 5 e 6 (Região Centro-Oeste), 7 e 8 (Região Nordeste), 9 e 10 (Região Norte), 11 e 12 (Região Sudeste), 13 e 14 (Região Sul).

Gráfico 5 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Centro-Oeste



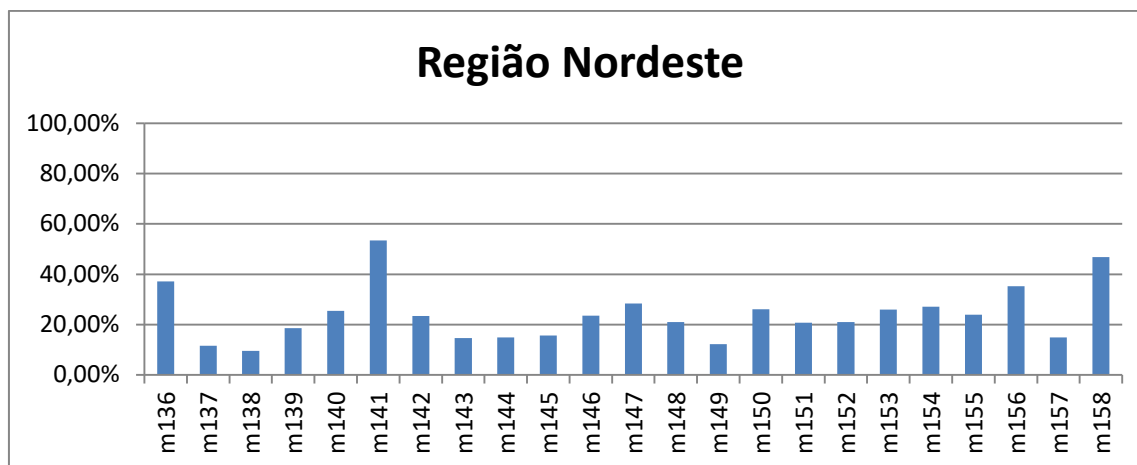
Fonte: Inep/Mec, 2017

Gráfico 6 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Centro-Oeste



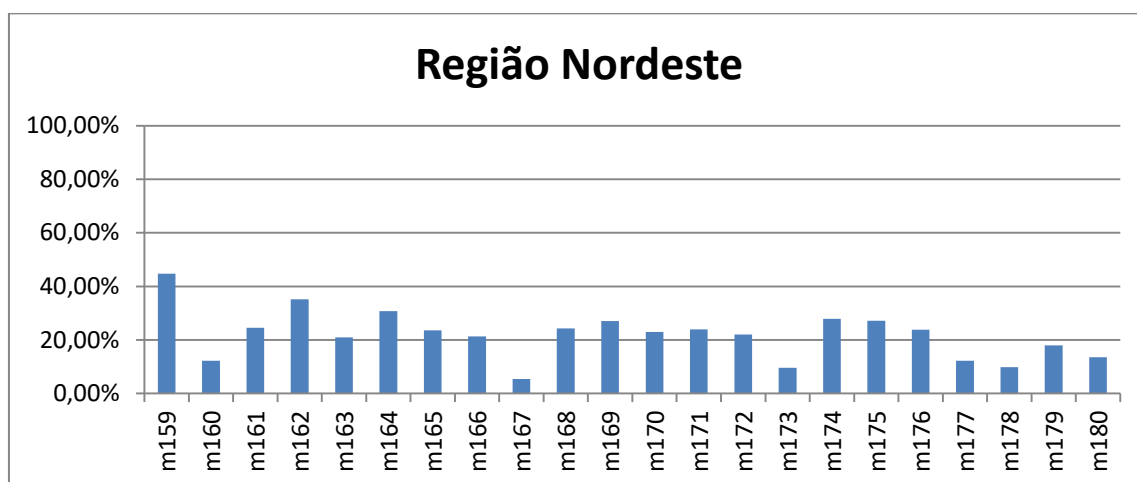
Fonte: Inep/Mec, 2017

Gráfico 7 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Nordeste



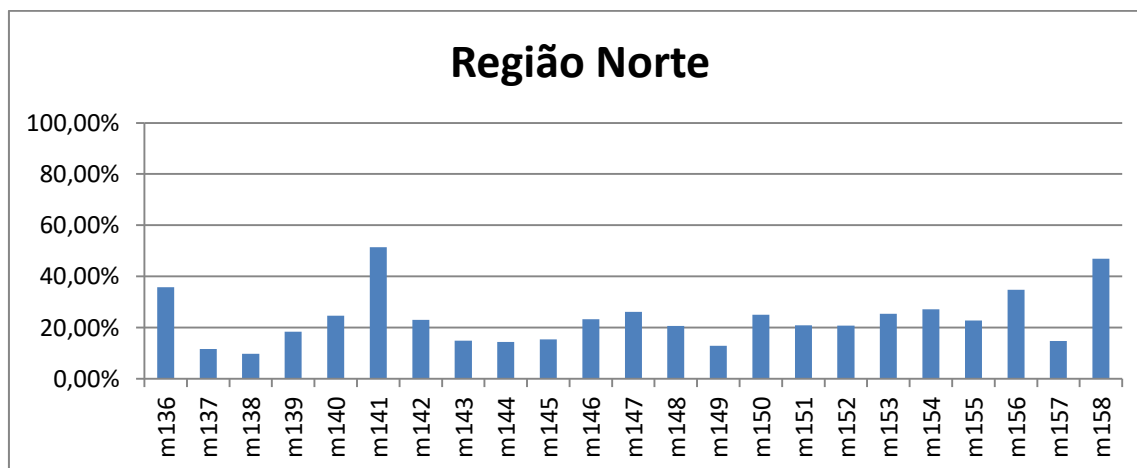
Fonte: Inep/Mec, 2017

Gráfico 8 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Nordeste



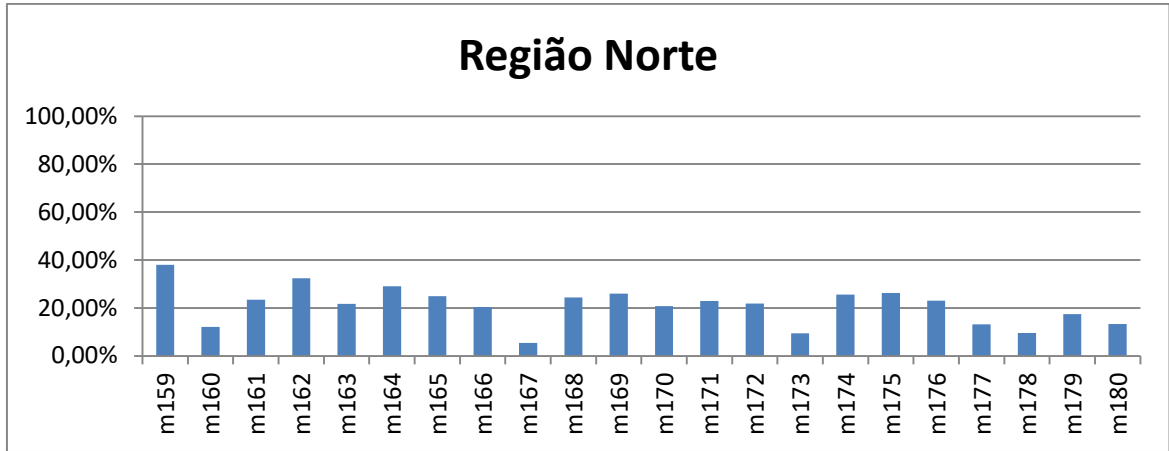
Fonte: Inep/Mec, 2017

Gráfico 9 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Norte



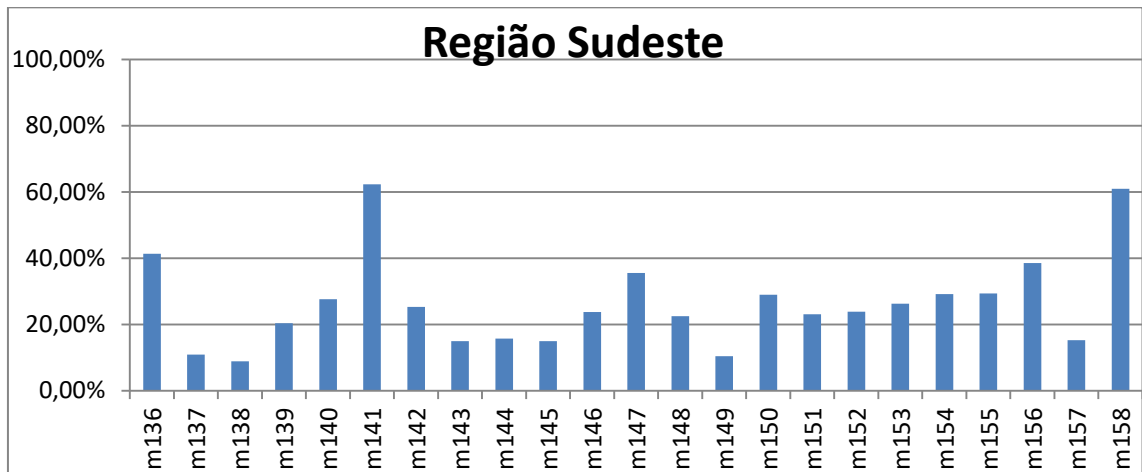
Fonte: Inep/Mec, 2017

Gráfico 10 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Norte



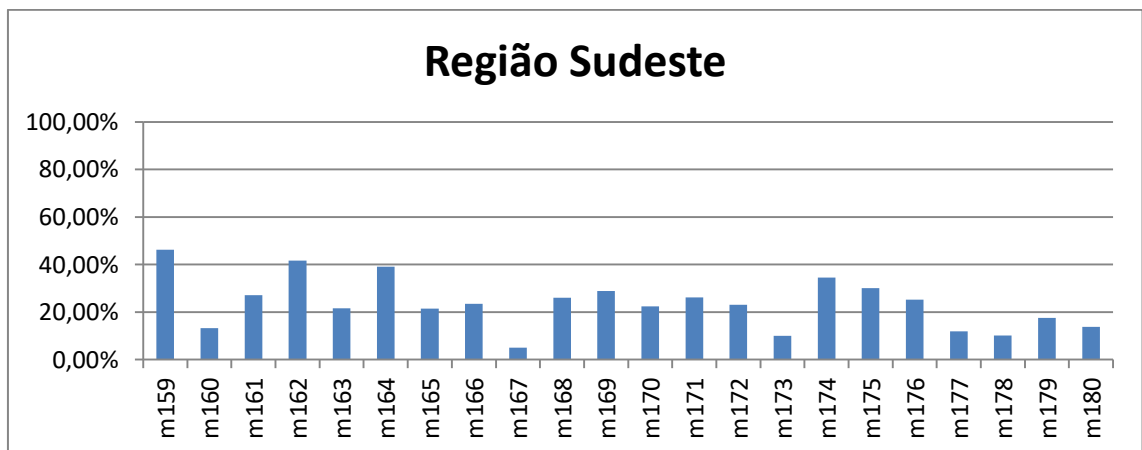
Fonte:Inep/Mec, 2017

Gráfico 11 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Sudeste



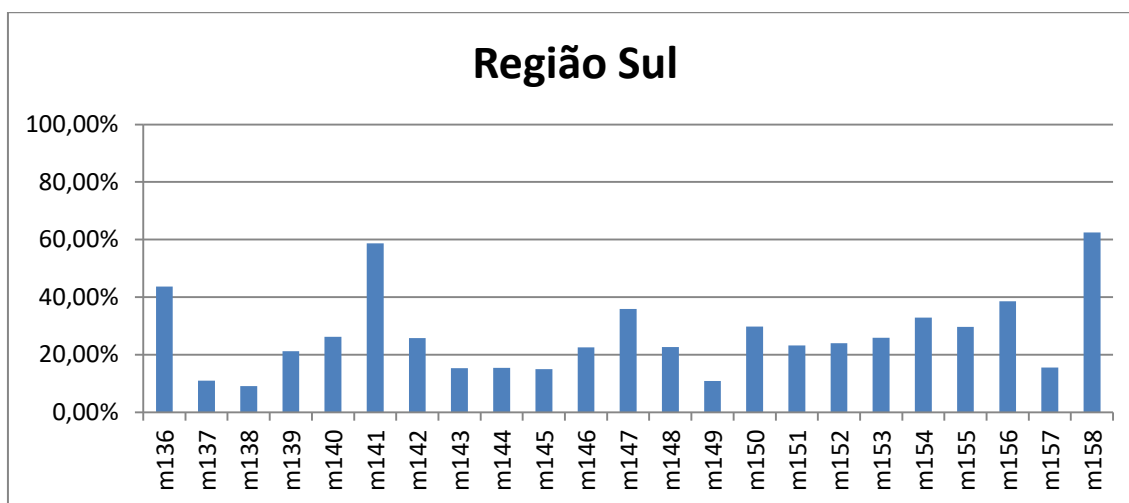
Fonte:Inep/Mec, 2017

Gráfico 12 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Sudeste



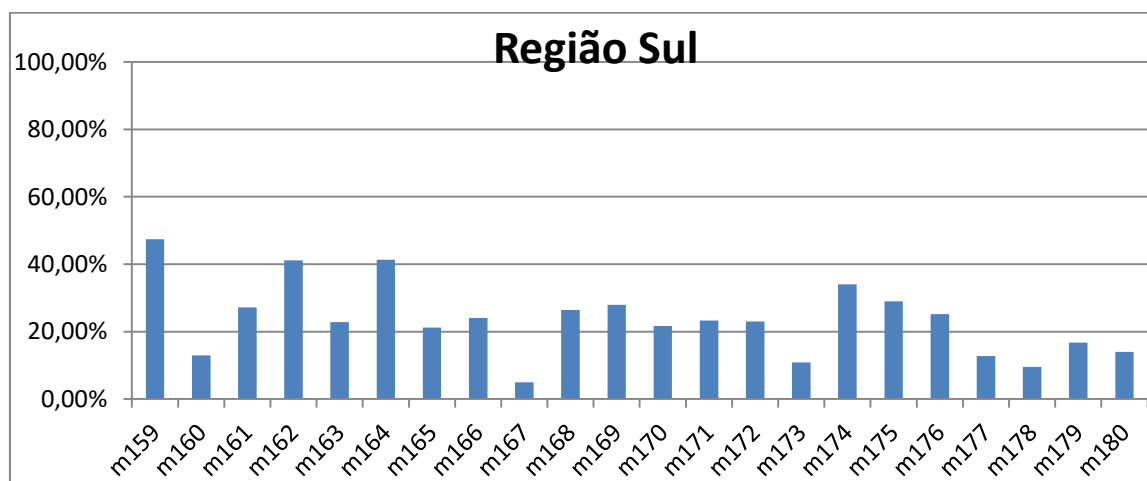
Fonte:Inep/Mec, 2017

Gráfico 13 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Região Sul



Fonte: Inep/Mec, 2017

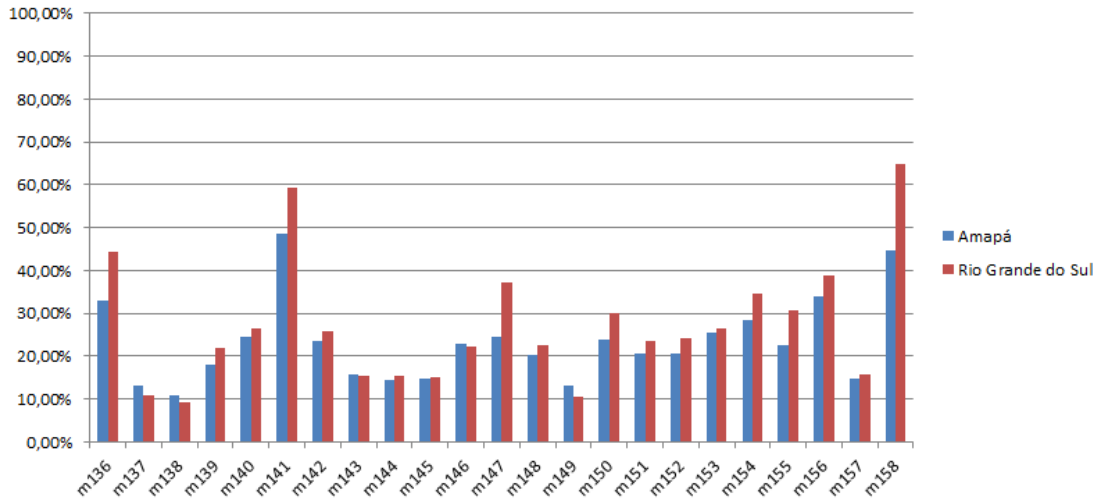
Gráfico 14 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Região Sul



Fonte: Inep/Mec, 2017

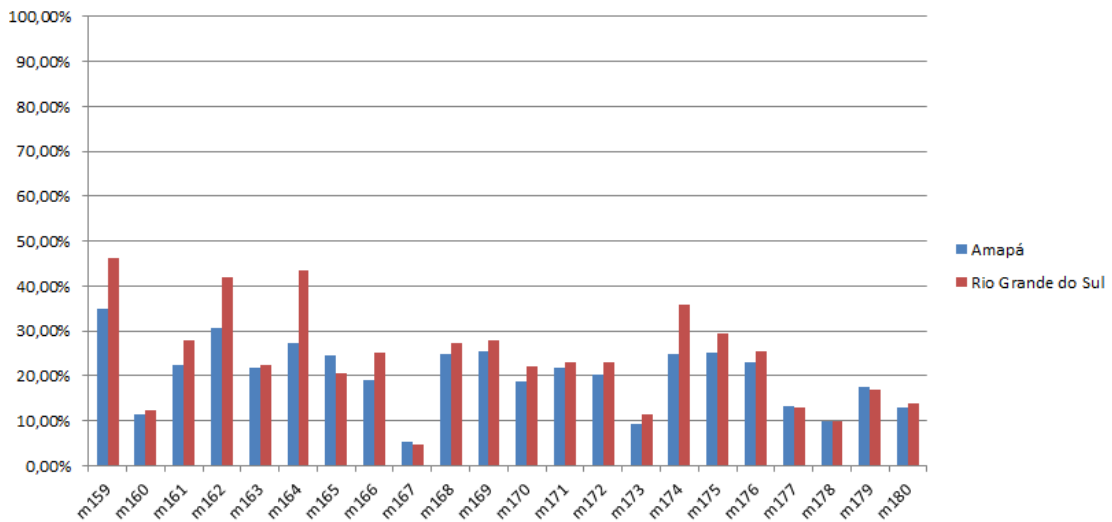
Foi feita a comparação do índice de acerto entre o estado com a maior nota (Rio Grande do Sul) com a menor nota (Amapá), ilustrada nos gráficos 15 e 16, apontando questões em que a diferença ultrapassava 20% de acertos (Questão m158).

Gráfico 15 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 136 até 158) - Amapá x Rio Grande do Sul



Fonte:Inep/Mec

Gráfico 16 - Índice de acertos escola pública Enem 2017 (Caderno amarelo questões 159 até 180) - Amapá x Rio Grande do Sul

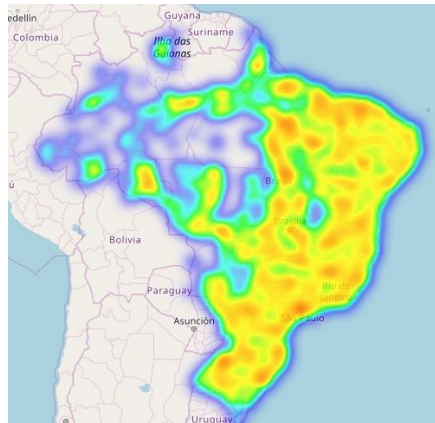


Fonte:Inep/Mec

Em relação ao número de acertos, observa-se que 51,54% dos candidatos acertaram apenas 10 questões ou menos das 45 questões presente na prova e 90,2% acertaram até 15 questões. Os mapas de calor (figuras 1 e 2) e o gráfico 17 ilustram a defasagem em Matemática no ENEM 2017 em relação ao número de acertos.



Figura 2 - Alunos com menos de 7 acertos (Caderno azul) - Enem



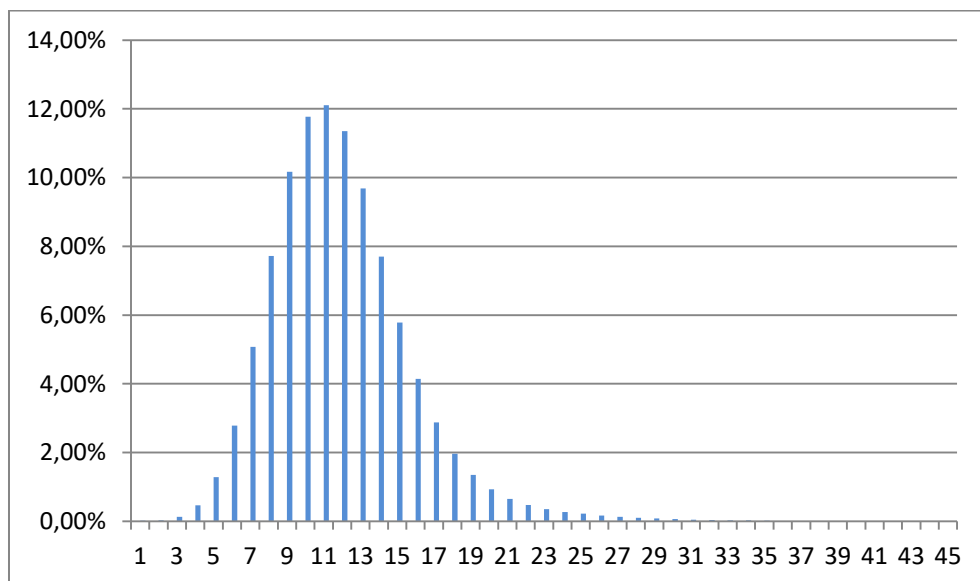
Fonte: Inep/Mec, 2017

Figura 3 - Alunos com 40 acertos ou mais (Caderno Azul) - Enem



Fonte: Inep/Mec, 2017

Gráfico 17 - índice de acertos escola pública em Matemática (Enem 2017)



Fonte: Inep/Mec, 2017

## 7. CONCLUSÃO

A análise dos dados do Enem 2017 expõem a defasagem do aprendizado de Matemática no ensino público em praticamente todo país, expondo que o índice médio de acerto por questão é de 24,01%, o que é relativamente baixo. Entender, de forma personalizada, as dificuldades que os alunos encontram nos conteúdos abordados no Ensino Médio é essencial para criar estratégias de resolução do problema. A produção e distribuição de material didático direcionado às especificidades de cada região é uma proposta viável, uma vez que a análise do conjunto de dados do Enem indica quais conteúdos os alunos apresentaram mais ou menos dificuldade por geolocalização. Logo, é cada vez mais importante buscar informações de valor para diagnósticos e intervenções efetivas a partir dos dados do Enem, abrangência desse exame e por sua representatividade em relação ao ensino e ao aprendizado em Matemática nas escolas públicas.

A mesma análise pode ser feita para qualquer ano ou para qualquer uma das outras avaliações (Ciências Humanas, Ciências Biológicas ou Linguagens), alterando apenas alguns filtros, abrindo possíveis caminhos para estudos futuros, com intuito de melhora na educação brasileira.

## 8. REFERÊNCIAS

Brasil. (20 de dezembro de 1996). Lei n.9394. *Diretrizes e bases da educação nacional*.

Erik Brynjolfsson, L. M. (22 de Abril de 2011). *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?* Acesso em 23 de Junho de 2019, disponível em Ssrn: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1819486](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486)

Fayyad, U., Pietetsky-Shapiro, G., & Smyth, P. (02 de Agosto de 1996). Knowledge Discovery and Data Mining:. *KDD-96 Proceedings*, pp. 82-88.

IBGE. (s.d.). *Instituto Brasileiro de Geografia e Estatística*. Acesso em 01 de Setembro de 2019, disponível em IBGE: <https://ibge.gov.br/>

Inep. (s.d.). *Inep*. Acesso em 28 de Novembro de 2018, disponível em Portal Inep - Mec: <http://inep.gov.br/microdados>

Laney, D. (06 de Fevereiro de 2001). Application Delivery Strategies. *Application Delivery Strategies*, p. file:949.

Mashey, J. R. (Abril de 1998). Big Data ... and the Next Wave of InfraStress.

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*.

Secretária da Educação do Estado de São Paulo. (s.d.). *Educação sp*. Acesso em 23 de junho de 2019, disponível em Educação sp: <https://www.educacao.sp.gov.br/saresp/>

Secretaria da Educação do Estado de São Paulo. (s.d.). *Secretaria Escolar Digital*. Acesso em 20 de Junho de 2019, disponível em Sed: <https://sed.educacao.sp.gov.br/Inicio>

Watson, I. (s.d.). *IBM Watson*. Acesso em 15 de Maio de 2019, disponível em IMB: <https://www.ibm.com/watson/education>

## 9. Anexos

### Anexo 1

Tabela 3 - Matriz de referência 3ºano E.M - SAEB

Temas	Descritores
<b>Espaço e Forma</b>	<ul style="list-style-type: none"> <li>• Identificar figuras semelhantes mediante o reconhecimento de relações de proporcionalidade.</li> <li>• Reconhecer aplicações das relações métricas do triângulo retângulo em um problema que envolva figuras planas ou espaciais.</li> <li>• Relacionar diferentes poliedros ou corpos redondos com suas planificações ou vistas.</li> <li>• Identificar a relação entre o número de vértices, faces e/ou arestas de poliedros expressa em um problema.</li> <li>• Resolver problema que envolva razões trigonométricas no triângulo retângulo (seno, cosseno, tangente).</li> <li>• Identificar a localização de pontos no plano cartesiano.</li> <li>• Interpretar geometricamente os coeficientes da equação de uma reta.</li> <li>• Identificar a equação de uma reta apresentada a partir de dois pontos dados ou de um ponto e sua inclinação.</li> <li>• Relacionar a determinação do ponto de interseção de duas ou mais retas com a resolução de um sistema de equações com duas incógnitas.</li> <li>• Reconhecer, dentre as equações do 2º grau com duas incógnitas, as que representam circunferências.</li> </ul>
<b>Grandezas e Medidas</b>	<ul style="list-style-type: none"> <li>• Resolver problema envolvendo o cálculo de perímetro de figuras planas.</li> <li>• Resolver problema envolvendo o cálculo de área de figuras planas.</li> <li>• Resolver problema envolvendo a área total e/ou volume de um sólido (prisma, pirâmide, cilindro, cone, esfera).</li> </ul>
<b>Números e Operações/Álgebra e Funções</b>	<ul style="list-style-type: none"> <li>• Identificar a localização de números reais na reta numérica.</li> <li>• Resolver problema que envolva variação proporcional, direta ou inversa, entre grandezas.</li> </ul>

- 
- Resolver problema que envolva porcentagem.
  - Resolver problema envolvendo equação do 2º grau.
  - Reconhecer expressão algébrica que representa uma função a partir de uma tabela.
  - Resolver problema envolvendo uma função do 1º grau.
  - Analisar crescimento/decrescimento, zeros de funções reais apresentadas em gráficos.
  - Identificar o gráfico que representa uma situação descrita em um texto.
  - Resolver problema envolvendo P.A./P.G. dada a fórmula do termo geral
  - Reconhecer o gráfico de uma função polinomial de 1º grau por meio de seus coeficientes.
  - Reconhecer a representação algébrica de uma função do 1º grau dado o seu gráfico.
  - Resolver problemas que envolvam os pontos de máximo ou de mínimo no gráfico de uma função polinomial do 2º grau.
  - Relacionar as raízes de um polinômio com sua decomposição em fatores do 1º grau.
  - Identificar a representação algébrica e/ou gráfica de uma função exponencial.
  - Identificar a representação algébrica e/ou gráfica de uma função logarítmica, reconhecendo-a como inversa da função exponencial.
  - Resolver problema que envolva função exponencial.
  - Identificar gráficos de funções trigonométricas (seno, cosseno, tangente) reconhecendo suas propriedades.
  - Determinar a solução de um sistema linear associando-o à uma matriz.
  - Resolver problema de contagem utilizando o princípio multiplicativo ou noções de permutação simples, arranjo simples e/ou combinação simples.
  - Calcular a probabilidade de um evento.
  - Resolver problema envolvendo informações apresentadas em tabelas e/ou gráficos.
  - Associar informações apresentadas em listas e/ou tabelas simples aos gráficos que as representam e vice-versa.

### **Tratamento da Informação**

## Anexo 2

Tabela 4 - Matriz de referência E.M - ENEM

Competência	Habilidades
<p><b>Competência de área 1 - Construir significados para os números naturais, inteiros, racionais e reais.</b></p>	<ul style="list-style-type: none"> <li>• Reconhecer, no contexto social, diferentes significados e representações dos números e operações - naturais, inteiros, racionais ou reais.</li> <li>• Identificar padrões numéricos ou princípios de contagem.</li> <li>• Resolver situação-problema envolvendo conhecimentos numéricos.</li> <li>• Avaliar a razoabilidade de um resultado numérico na construção de argumentos sobre afirmações quantitativas.</li> <li>• Avaliar propostas de intervenção na realidade utilizando conhecimentos numéricos.</li> </ul>
<p><b>Competência de área 2 - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.</b></p>	<ul style="list-style-type: none"> <li>• Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.</li> <li>• Identificar características de figuras planas ou espaciais.</li> <li>• Resolver situação-problema que envolva conhecimentos geométricos de espaço e forma.</li> <li>• Utilizar conhecimentos geométricos de espaço e forma na seleção de argumentos propostos como solução de problemas do cotidiano.</li> </ul>
<p><b>Competência de área 3 - Construir noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.</b></p>	<ul style="list-style-type: none"> <li>• Identificar relações entre grandezas e unidades de medida.</li> <li>• Utilizar a noção de escalas na leitura de representação de situação do cotidiano.</li> <li>• Resolver situação-problema que envolva medidas de grandezas.</li> <li>• Avaliar o resultado de uma medição na construção de um argumento consistente.</li> </ul>

---

**Competência de área 4 - Construir noções de variação de grandezas para a compreensão da realidade e a solução de problemas do cotidiano.**

- Avaliar proposta de intervenção na realidade utilizando conhecimentos geométricos relacionados a grandezas e medidas.
- Identificar a relação de dependência entre grandezas.

**Competência de área 5 - Modelar e resolver problemas que envolvem variáveis socioeconômicas ou técnico-científicas, usando representações algébricas.**

- Resolver situação-problema envolvendo a variação de grandezas, direta ou inversamente proporcionais.
- Analisar informações envolvendo a variação de grandezas como recurso para a construção de argumentação.
- Avaliar propostas de intervenção na realidade envolvendo variação de grandezas.
- Identificar representações algébricas que expressem a relação entre grandezas.

**Competência de área 6 - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.**

- Interpretar gráfico cartesiano que represente relações entre grandezas.
- Resolver situação-problema cuja modelagem envolva conhecimentos algébricos.
- Utilizar conhecimentos algébricos/geométricos como recurso para a construção de argumentação.
- Avaliar propostas de intervenção na realidade utilizando conhecimentos algébricos.
- Utilizar informações expressas em gráficos ou tabelas para fazer inferências.
- Resolver problema com dados apresentados em tabelas ou gráficos.
- Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.