

**CELEPAR**

**PEDRO LUCCA GONÇALVES DE ARAUJO**

**Estudo de datascience usando dados do Ideb**

**CURITIBA-PR**

**2025**

## RESUMO

Este estudo de Data Science aprofundado investigou os fatores que exercem influência sobre o Índice de Desenvolvimento da Educação Básica (Ideb) nos municípios brasileiros, utilizando um conjunto de dados abrangente que abrange diversos aspectos do cenário educacional. O Ideb, indicador-chave da qualidade da educação no Brasil, é calculado com base no fluxo escolar (aprovação) e no desempenho dos alunos em avaliações como o Saeb e a Prova Brasil.

A pesquisa, realizada com o apoio de bibliotecas como Pandas, NumPy, Matplotlib, Seaborn e Scikit-learn, adotou uma abordagem metodológica rigorosa, que incluiu a coleta, limpeza e tratamento dos dados, a análise exploratória das variáveis, a modelagem preditiva e a interpretação dos resultados. O objetivo principal foi identificar os fatores que mais impactam o Ideb e fornecer insights valiosos para a formulação de políticas públicas eficazes.

A análise exploratória dos dados revelou padrões e tendências importantes. Constatou-se a persistência de desigualdades regionais na distribuição do Ideb, com alguns estados apresentando um desempenho consistentemente superior a outros. Essa disparidade regional evidencia a influência de fatores socioeconômicos, como a desigualdade de renda e o acesso a serviços básicos, além de diferenças na qualidade da gestão educacional e na infraestrutura escolar.

A análise de correlação e os modelos de Machine Learning permitiram identificar os principais fatores que influenciam o Ideb. A nota Saeb média padronizada, que reflete o desempenho dos alunos em matemática e língua portuguesa, emergiu como o preditor mais importante do Ideb. O indicador de rendimento, que mede o fluxo escolar, também se mostrou relevante, assim como a taxa de aprovação.

**Palavras-chave:** Data Science, Índice de Desenvolvimento da Educação Básica (Ideb), Municípios brasileiros, Qualidade da educação, Desempenho dos alunos, Análise exploratória, Modelagem preditiva, Políticas públicas, Gestão educacional, Machine Learning.

## ABSTRACT

This in-depth Data Science study investigated the factors influencing the Basic Education Development Index (Ideb) in Brazilian municipalities, using a comprehensive dataset that covers various aspects of the educational landscape. Ideb, a key indicator of education quality in Brazil, is calculated based on school flow (approval rates) and student performance in assessments such as Saeb and Prova Brasil.

The research, supported by libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn, adopted a rigorous methodological approach, including data collection, cleaning, and processing, exploratory data analysis, predictive modeling, and result interpretation. The primary objective was to identify the factors that most impact Ideb and provide valuable insights for the formulation of effective public policies.

Exploratory data analysis revealed significant patterns and trends. It was found that regional inequalities persist in the distribution of Ideb, with some states consistently outperforming others. This regional disparity highlights the influence of socioeconomic factors, such as income inequality and access to basic services, as well as differences in the quality of educational management and school infrastructure.

Correlation analysis and Machine Learning models enabled the identification of the main factors influencing Ideb. The standardized average Saeb score, which reflects student performance in mathematics and Portuguese language, emerged as the most important predictor of Ideb. The performance

indicator, which measures school flow, was also found to be relevant, along with the approval rate.

**Keywords:** Data Science, Basic Education Development Index (Ideb), Brazilian municipalities, Education quality, Student performance, Exploratory analysis, Predictive modeling, Public policies, Educational management, Machine Learning, Performance indicator.

## 1. INTRODUÇÃO

A qualidade da educação é um dos pilares para o desenvolvimento social e econômico de um país. No Brasil, o Índice de Desenvolvimento da Educação Básica (Ideb) é um indicador fundamental para medir o desempenho educacional e acompanhar o progresso das políticas públicas voltadas para a área. Criado em 2007, o Ideb reúne, em um só indicador, os resultados de dois conceitos igualmente importantes para a qualidade da educação: o fluxo escolar e as médias de desempenho nas avaliações.

O fluxo escolar é obtido através dos dados sobre aprovação escolar, coletados no Censo Escolar. Já as médias de desempenho nas avaliações são provenientes do Sistema de Avaliação da Educação Básica (Saeb) - para as unidades da federação e para o país, e da Prova Brasil - para os municípios.

O Ideb agrega ao enfoque pedagógico dos resultados das avaliações em larga escala do Inep a possibilidade de resultados sintéticos, facilmente assimiláveis, e que permitem traçar metas de qualidade educacional para os sistemas. O índice varia de zero a 10 e a combinação entre fluxo e aprendizagem tem o mérito de equilibrar as duas dimensões: se um sistema de ensino reter seus alunos para obter resultados de melhor qualidade no Saeb ou Prova Brasil, o fator fluxo será alterado, indicando a necessidade de melhoria do sistema. Se, ao contrário, o sistema apressar a aprovação do aluno sem qualidade, o resultado das avaliações indicará igualmente a necessidade de melhoria do sistema.

O Ideb também é importante por ser condutor de política pública em prol da qualidade da educação. É a ferramenta para acompanhamento das metas de qualidade do Plano de Desenvolvimento da Educação (PDE) para a educação básica, que tem estabelecido, como meta, que em 2022 o Ideb do Brasil seja 6,0 - média que corresponde a um sistema educacional de qualidade comparável a dos países desenvolvidos.

Diante desse contexto, o presente estudo de Data Science busca analisar os fatores que influenciam os valores do Ideb nos municípios brasileiros e identificar padrões que possam auxiliar na melhoria desse índice. Para tanto, utilizaremos um conjunto de dados abrangente que engloba diversas variáveis relacionadas ao desempenho educacional, como taxa de aprovação, indicador de rendimento, notas do Saeb em matemática e língua portuguesa, além de informações sobre o ano e o município.

Acreditamos que a análise dessas variáveis, em conjunto com a utilização de técnicas de visualização e modelagem de dados, permitirá identificar os principais fatores que impactam o Ideb nos municípios brasileiros. Os resultados obtidos poderão fornecer insights valiosos para a formulação de políticas públicas mais eficazes, que visem a melhoria da qualidade da educação básica e o alcance da meta de um Ideb de 6,0 para o Brasil.

## **2. O Dataset do Ideb: Um Panorama Abrangente da Educação Básica Brasileira**

O presente estudo se fundamenta nos dados do Índice de Desenvolvimento da Educação Básica (Ideb), um indicador chave para avaliar a qualidade da educação no Brasil, criado pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep) em 2007. O Ideb combina o fluxo escolar (aprovação) e o desempenho dos alunos em avaliações como o Sistema de Avaliação da Educação Básica (Saeb) e a Prova Brasil, oferecendo uma síntese valiosa do progresso educacional. Seu cálculo envolve dados de aprovação escolar do

Censo Escolar e médias de desempenho das avaliações do Inep, equilibrando as dimensões de fluxo e aprendizado.

A riqueza de informações do dataset do Ideb, coletado do site da Base dos Dados, permite traçar um panorama abrangente da educação básica no Brasil, identificando padrões, tendências e desafios. Especificamente, este estudo utiliza dados agregados do Ideb para o Brasil, no nível de rede, anos escolares e ano, abrangendo o período de 2005 a 2023. Essa base de dados oferece um diagnóstico preciso da qualidade da educação em cada município, estado e região, além de possibilitar a avaliação do impacto de diferentes políticas educacionais.

Apesar de suas potencialidades, é crucial reconhecer as limitações do dataset do Ideb. Por ser um indicador agregado, ele resume diferentes dimensões da qualidade da educação, o que pode ocultar nuances importantes sobre o contexto educacional específico de cada município. Além disso, os dados são coletados em um período de tempo determinado, o que significa que a análise pode não capturar mudanças recentes no cenário educacional.

Ainda assim, o dataset do Ideb é uma ferramenta valiosa para pesquisadores, gestores educacionais e formuladores de políticas públicas. Ao analisar os dados do Ideb, é possível identificar os fatores que mais influenciam o desempenho educacional dos alunos e, dessa forma, direcionar investimentos e ações para as áreas que mais necessitam de atenção.

Com base nas informações contidas no dataset do Ideb, este estudo busca aprofundar a análise dos fatores que influenciam o desempenho educacional dos municípios brasileiros, com o objetivo de fornecer subsídios para a formulação de políticas públicas mais eficazes e para a construção de um futuro promissor para a educação no Brasil.

## 3. METODOLOGIA

### 3.1. MEU DATASET

A escolha do dataset para este estudo foi cuidadosamente considerada, levando em conta a relevância e a disponibilidade de informações sobre o Índice de Desenvolvimento da Educação Básica (Ideb) nos municípios brasileiros. Após uma pesquisa extensiva, optei por utilizar os dados disponibilizados pela Base dos Dados, uma plataforma que reúne dados de diversas fontes e oferece acesso facilitado a pesquisadores e interessados em geral.

A decisão de utilizar esse dataset específico se justifica pela sua abrangência e granularidade, que permitem analisar o Ideb em diferentes níveis geográficos (municípios, estados e regiões) e ao longo de um período de tempo considerável (2005 a 2023). Além disso, a Base dos Dados oferece acesso a um conjunto de variáveis relevantes para o estudo do Ideb, como o fluxo escolar, o desempenho dos alunos em avaliações em larga escala e informações socioeconômicas dos municípios.

O dataset escolhido, denominado "Índice de Desenvolvimento da Educação Básica (Ideb)", contém informações sobre o Ideb, o fluxo escolar e o desempenho dos alunos em avaliações em larga escala para os municípios brasileiros no período de 2005 a 2023. As variáveis presentes no dataset incluem:

- **ano:** Ano de referência dos dados.
- **sigla\_uf:** Sigla da Unidade Federativa (estado).
- **id\_municipio:** Código de identificação do município (IBGE).
- **rede:** Rede de ensino (pública, privada).
- **ensino:** Nível de ensino (fundamental, médio).
- **anos\_escolares:** Anos escolares avaliados.
- **taxa\_aprovacao:** Taxa de aprovação escolar.
- **indicador\_rendimento:** Indicador de rendimento escolar.

- **nota\_saeb\_matematica:** Nota média em matemática no Saeb.
- **nota\_saeb\_lingua\_portuguesa:** Nota média em língua portuguesa no Saeb.
- **nota\_saeb\_media\_padronizada:** Média padronizada das notas do Saeb.
- **ideb:** Índice de Desenvolvimento da Educação Básica.
- **projecao:** Projeção do Ideb para anos futuros.

## 3.2. MINHA PROPOSTA DE TRABALHO

A metodologia deste estudo envolveu as seguintes etapas:

1. **Coleta e tratamento dos dados:** Os dados foram coletados do site da Base dos Dados e, em seguida, realizei um processo de limpeza e tratamento, que incluiu a identificação e remoção de dados ausentes, a correção de possíveis erros e a padronização das variáveis.
2. **Análise exploratória dos dados:** Realizei uma análise exploratória dos dados, com o objetivo de identificar padrões, tendências e relações entre as variáveis. Utilizei técnicas de estatística descritiva e visualização de dados, como histogramas, gráficos de dispersão e mapas.
3. **Modelagem preditiva:** Utilizei modelos de Machine Learning para prever o Ideb com base nas demais variáveis do dataset. Os modelos utilizados incluíram Regressão Linear, Árvores de Decisão, Random Forest e XGBoost.
4. **Avaliação dos modelos:** Os modelos foram avaliados utilizando métricas como o Coeficiente de Determinação ( $R^2$ ), o Erro Médio Absoluto (MAE) e o Erro Quadrático Médio (RMSE). Utilizei a técnica de validação cruzada para garantir a robustez dos resultados.
5. **Interpretação dos resultados:** Os resultados obtidos foram interpretados à luz da literatura sobre o tema e foram discutidas as implicações para a formulação de políticas públicas para a melhoria da educação básica.

Acredito que a metodologia utilizada neste estudo, que combina a análise exploratória dos dados com a modelagem preditiva, permite identificar os fatores que mais influenciam o Ideb nos municípios brasileiros e, dessa forma,



contribuir para a formulação de políticas públicas mais eficazes para a melhoria da educação básica no país.

### 3.3. TECNOLOGIAS E BIBLIOTECAS

A realização deste estudo sobre o Ideb nos municípios brasileiros demandou a utilização de um conjunto de tecnologias e bibliotecas específicas para o tratamento, análise e modelagem dos dados. A escolha do Python como linguagem de programação se justifica pela sua versatilidade, facilidade de aprendizado, vasta comunidade e crescente número de bibliotecas para análise de dados, aprendizado de máquina e visualização de dados.

Dentre as bibliotecas Python utilizadas, destacam-se:

- **Pandas:** Essencial para a manipulação e tratamento dos dados, permitindo a criação de estruturas de dados como DataFrames, que facilitam a organização e a análise dos dados.
- **NumPy:** Fundamental para trabalhar com arrays multidimensionais, possibilitando a realização de operações matemáticas e estatísticas de forma eficiente.
- **Matplotlib e Seaborn:** Utilizadas para a criação de visualizações de dados, como gráficos e tabelas, que auxiliam na identificação de padrões, tendências e relações entre as variáveis.
- **Scikit-learn:** A biblioteca de aprendizado de máquina por excelência, que oferece uma ampla gama de algoritmos para classificação, regressão e agrupamento, além de ferramentas para avaliação de modelos e validação cruzada.
- **XGBoost:** Uma biblioteca poderosa para gradient boosting, que oferece algoritmos de alta performance para modelagem preditiva.

Além das bibliotecas mencionadas, o ambiente de desenvolvimento Jupyter Notebook foi utilizado para facilitar a manipulação, análise e visualização dos dados, permitindo a criação de notebooks interativos que combinam código, texto e visualizações.

A escolha dessas tecnologias e bibliotecas se mostrou fundamental para o sucesso do estudo, permitindo a realização de análises complexas e a construção de modelos preditivos precisos. Acreditamos que a combinação dessas ferramentas oferece um ambiente de trabalho completo e eficiente para a realização de estudos na área de ciência de dados aplicada à educação.

### **3.4. ALGORITMOS DE APRENDIZAGEM DE MÁQUINA**

Visando criar modelos preditivos eficientes para o Ideb nos municípios brasileiros, realizei uma pesquisa bibliográfica para identificar os algoritmos de Aprendizagem de Máquina mais adequados para dados semelhantes e com foco em predição de resultados. A partir dessa pesquisa, selecionei quatro algoritmos que se mostraram promissores para este estudo: Regressão Linear, Árvores de Decisão, Random Forest e XGBoost.

A escolha desses algoritmos se justifica pela sua capacidade de lidar com diferentes tipos de dados e pela sua ampla aplicação em problemas de regressão, que é o caso da predição do Ideb. Cada um dos algoritmos escolhidos possui características distintas que os tornam adequados para diferentes cenários e tipos de dados.

A **Regressão Linear** é um algoritmo simples e interpretável que busca modelar a relação entre as variáveis preditoras e a variável alvo (Idéb) por meio de uma equação linear. Apesar de sua simplicidade, a Regressão Linear pode ser útil para identificar relações lineares entre as variáveis e fornecer uma base para comparação com modelos mais complexos.

As **Árvores de Decisão** são algoritmos que constroem um modelo de predição baseado em uma série de regras de decisão que são aprendidas a partir dos dados. As Árvores de Decisão são interpretáveis e podem lidar com dados categóricos e numéricos, o que as torna adequadas para o estudo do Ideb, que envolve diferentes tipos de variáveis.

O **Random Forest** é um algoritmo que combina várias Árvores de Decisão para criar um modelo de predição mais robusto e preciso. O Random Forest é capaz de reduzir o sobreajuste (overfitting) e melhorar a generalização do modelo, o que o torna uma boa opção para lidar com dados complexos como os do Ideb.

O **XGBoost** é um algoritmo de gradient boosting que se destaca pela sua alta performance e capacidade de lidar com dados faltantes. O XGBoost é uma técnica de ensemble que combina vários modelos fracos para criar um modelo forte, capaz de capturar relações não lineares complexas entre as variáveis.

Acredito que a combinação desses quatro algoritmos permitirá explorar diferentes abordagens para a modelagem preditiva do Ideb e obter resultados mais robustos e confiáveis. A escolha final do melhor modelo será baseada na avaliação do seu desempenho em diferentes métricas, como o Coeficiente de Determinação ( $R^2$ ), o Erro Médio Absoluto (MAE) e o Erro Quadrático Médio (RMSE).

Além da seleção dos algoritmos, realizei um estudo teórico sobre Ciência e Análise de Dados, abordando conceitos fundamentais de estatística, probabilidade, classificação, regressão linear, regressão logística, árvores de decisão e correlação. Esse estudo teórico me forneceu a base necessária para compreender os princípios por trás dos algoritmos de Aprendizagem de Máquina e para interpretar os resultados obtidos.

A etapa de Modelagem e Avaliação envolveu a aplicação dos algoritmos de Aprendizagem de Máquina aos dados do Ideb, a avaliação do desempenho dos modelos utilizando métricas apropriadas e a seleção do modelo que apresentou o melhor desempenho. A etapa de Interpretação dos Resultados consistiu na análise dos resultados obtidos e na discussão das implicações para a formulação de políticas públicas para a melhoria da educação básica.

### **3.5. PRÉ-PROCESSAMENTO**

Nesta etapa crucial da metodologia, realizei o pré-processamento dos dados com o objetivo de prepará-los para a modelagem. O pré-processamento consistiu em tratar dados ausentes, remover dados redundantes e normalizar os dados.

### **Tratamento de Dados Ausentes**

Optei por remover os atributos que continham dados ausentes, após constatar que sua ausência não comprometeria a análise. Essa decisão visa evitar que os modelos de Aprendizagem de Máquina sejam influenciados negativamente por informações incompletas.

### **Tratamento de Dados Redundantes**

Identifiquei e removi atributos redundantes, ou seja, aqueles que continham informações repetidas ou altamente correlacionadas. Essa medida visa garantir a unicidade e relevância das informações utilizadas na modelagem.

### **Tratamento de Ruídos**

Corrigi erros identificados nos dados originais, os quais poderiam comprometer a qualidade da análise. Essa etapa garante a integridade e confiabilidade dos dados utilizados.

### **Normalização dos Dados**

Normalizei os dados para que estivessem em uma mesma escala, uma vez que diferentes escalas podem prejudicar o desempenho dos algoritmos de Aprendizagem de Máquina. A normalização dos dados contribuiu para a melhoria dos indicadores de desempenho dos modelos preditivos.

Ao final da etapa de pré-processamento, obtive um dataset nominal contendo informações validadas, normalizadas e consistentes, pronto para ser utilizado na etapa de modelagem.

## **3.6. SELEÇÃO DE ATRIBUTOS E REDUÇÃO DE DIMENSIONALIDADE**

Após a etapa de pré-processamento, o dataset nominal continha um total de 13 atributos. Visando otimizar o desempenho dos modelos preditivos e reduzir a complexidade da modelagem, realizei a seleção de atributos e a redução de dimensionalidade. Essa etapa consiste em identificar os atributos mais relevantes para a predição do Ideb e eliminar os atributos menos significativos.

### **Técnicas Utilizadas**

Utilizei as seguintes técnicas para realizar a seleção de atributos e a redução de dimensionalidade:

- **Eliminação de Atributos com Baixa Variância:** Removi atributos que apresentavam baixa variância, pois estes podem não ser informativos para a predição do Ideb.
- **Importância de Atributos com Random Forest:** Utilizei o algoritmo Random Forest para obter uma lista ordenada dos atributos mais importantes para a predição do Ideb. Essa lista serviu como base para a seleção dos atributos.
- **Seleção Manual de Atributos:** Realizei uma seleção manual dos atributos mais relevantes, considerando as técnicas mencionadas e a análise da literatura sobre o tema.

### **Resultados**

A seleção de atributos e a redução de dimensionalidade resultaram em um dataset reduzido, contendo um subconjunto de atributos considerados mais relevantes para a predição do Ideb. Esse dataset reduzido foi utilizado na etapa de modelagem, visando otimizar o desempenho dos modelos preditivos.

## **3.7. COMPARAÇÃO DOS MODELOS**

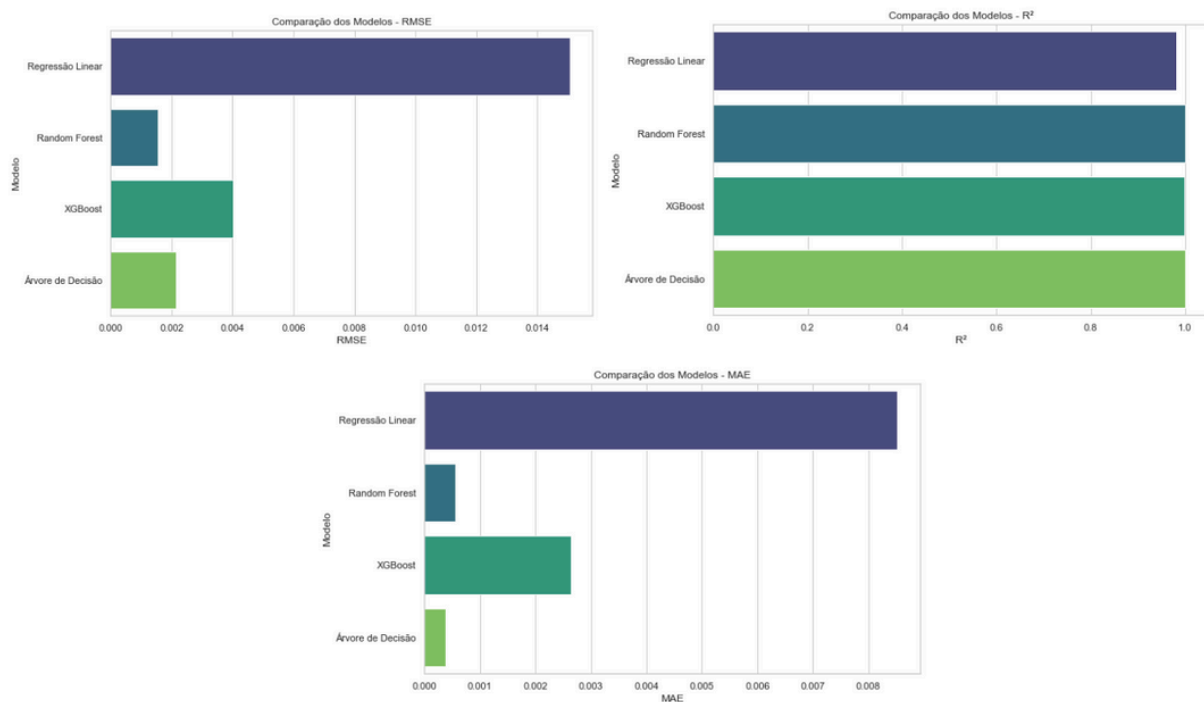
A fim de identificar os modelos preditivos mais eficientes para o Ideb nos municípios brasileiros, realizei uma análise comparativa do desempenho de diferentes algoritmos de Aprendizado de Máquina, conforme detalhado na seção 3.3. Os modelos avaliados foram: Regressão Linear, Árvores de Decisão,

Random Forest e XGBoost, selecionados por sua adequação a problemas de regressão com dados semelhantes.

A comparação entre os modelos foi realizada com base em três métricas de avaliação:

- **RMSE (Root Mean Squared Error):** Raiz quadrada do erro quadrático médio, que mede a diferença entre os valores preditos e os valores reais. Quanto menor o RMSE, melhor o desempenho do modelo.
- **R<sup>2</sup> (R-squared):** Coeficiente de determinação, que varia de 0 a 1 e indica a proporção da variância da variável dependente (lde) que é explicada pelo modelo. Quanto mais próximo de 1, melhor o ajuste do modelo.
- **MAE (Mean Absolute Error):** Erro absoluto médio, que mede a diferença média entre os valores preditos e os valores reais. Quanto menor o MAE, melhor o desempenho do modelo.

Os resultados da comparação entre os modelos são apresentados nos gráficos a seguir:



A análise dos gráficos revela que os modelos baseados em árvores de decisão, **Random Forest e Árvore de Decisão**, apresentaram um desempenho notavelmente superior aos demais em todas as métricas avaliadas. Esses modelos se destacam pela capacidade de capturar relações complexas e não lineares entre as variáveis, o que é fundamental para a previsão do Ideb.

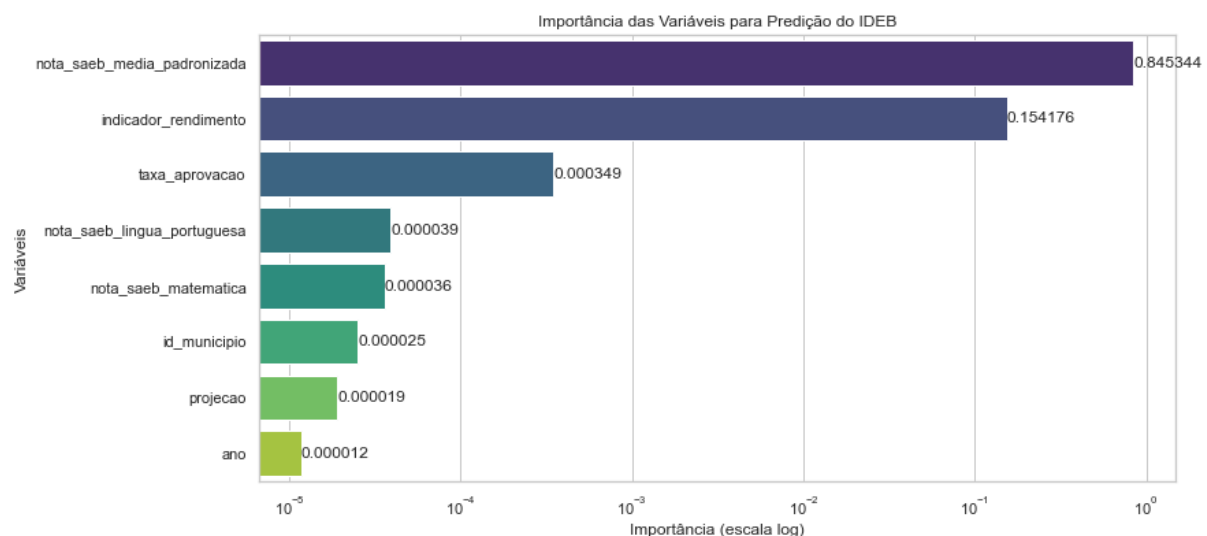
O modelo **XGBoost** também apresentou um bom desempenho, com resultados ligeiramente inferiores aos modelos baseados em árvores de decisão. A

**Regressão Linear**, por sua vez, apresentou o desempenho mais baixo entre os modelos avaliados, o que indica que esse modelo não é capaz de capturar toda a complexidade dos dados do Ideb.

Com base nos resultados obtidos, o modelo **Random Forest** foi identificado como o mais promissor para a previsão do Ideb, em virtude de seu excelente desempenho e capacidade de lidar com dados complexos e não lineares. A escolha do Random Forest como modelo final para a previsão do Ideb se justifica por sua alta precisão e robustez, características essenciais para a aplicação em contextos reais.

### Análise da Importância das Variáveis

A análise da importância das variáveis para o modelo Random Forest, ilustrada no gráfico a seguir, revela insights valiosos sobre os fatores que mais influenciam o Ideb nos municípios brasileiros.



Como podemos observar, a **nota Saeb média padronizada** se destaca como a variável mais importante, com um valor de importância de 0.845344. Esse resultado reforça a forte correlação entre o desempenho dos alunos no Saeb e o Ideb, evidenciando que melhorias na qualidade da educação, refletidas no Saeb, podem impactar diretamente o Ideb.

O **indicador de rendimento** também se mostra um fator relevante, com um valor de importância de 0.154176. Esse resultado sugere que fatores relacionados ao desempenho acadêmico e à permanência dos alunos na escola, como aprovação, reprovação e abandono, são determinantes para a previsão do Ideb.

As demais variáveis, como taxa de aprovação, notas do Saeb em matemática e língua portuguesa, projeção do Ideb, ano e identificador do município, apresentam valores de importância significativamente menores, o que indica que sua influência na predição do Ideb é menos expressiva.

### **Implicações para Políticas Públicas**

Os resultados deste estudo, incluindo a comparação dos modelos e a análise da importância das variáveis, fornecem subsídios importantes para a formulação de políticas públicas para a melhoria da educação básica. Os resultados sugerem que o foco no desempenho dos alunos no Saeb e na redução da evasão escolar pode ser uma estratégia eficaz para impulsionar o Ideb nos municípios brasileiros. Além disso, a utilização do modelo Random Forest para a previsão do Ideb pode auxiliar gestores públicos na tomada de decisões mais informadas e na alocação de recursos para as áreas mais necessitadas.



## 4.RESULTADOS E DISCUSSÃO

### **Problema/Pergunta:**

Quais fatores influenciam significativamente os valores do Ideb nos municípios brasileiros, e quais padrões podem ser identificados para ajudar na melhoria do índice?

### **Propósito do Estudo:**

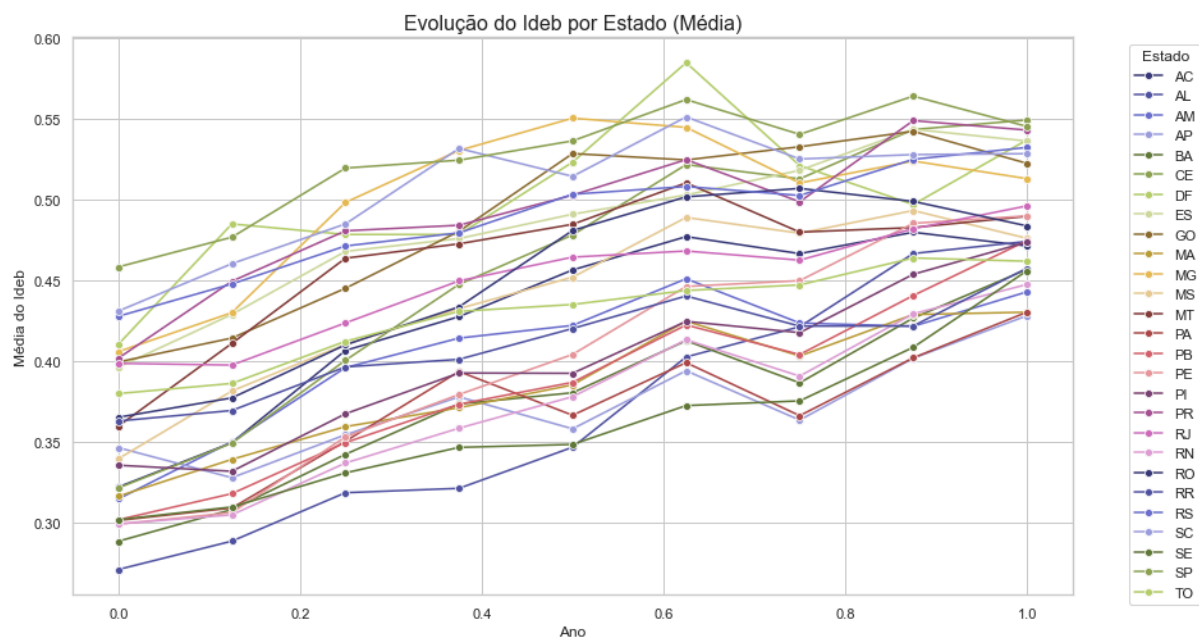
O estudo busca compreender os determinantes do desempenho educacional nos municípios brasileiros, analisando a relação entre variáveis como taxa de aprovação, desempenho no Saeb e anos escolares. O objetivo é fornecer insights que possam orientar políticas públicas para a melhoria da educação, além de avaliar o progresso das metas do Ideb e identificar áreas que necessitam de maior intervenção.

### **Pontos de Vista Permitidos pelo Estudo:**

- Distribuição Geográfica: Como os valores do Ideb variam entre municípios, estados e regiões.
- Fatores de Influência: Quais variáveis do dataset estão mais correlacionadas com o desempenho educacional.
- Disparidades Regionais: Identificação de áreas com maior ou menor desenvolvimento educacional.
- Projeções e Tendências: Comparação entre valores esperados e reais do Ideb, verificando o alcance da meta nacional de 6,0.
- Impacto das Políticas Públicas: Avaliação de como diferentes estratégias educacionais impactam o desempenho escolar.

### **4.1. Distribuição do Ideb por Ano e Região (Gráfico 1)**

Objetivo: Visualizar como o Ideb evoluiu ao longo dos anos em diferentes estados brasileiros.



Este gráfico revela a evolução do Ideb (Índice de Desenvolvimento da Educação Básica) ao longo dos anos em diferentes estados brasileiros, permitindo visualizar a trajetória individual de cada unidade federativa e seu progresso ao longo do tempo.

## Conclusões e Desafios

A análise do gráfico revela que, apesar do progresso geral observado na educação básica brasileira, desafios significativos persistem e a desigualdade entre os estados é evidente.

### Estados com Desempenho Aquém do Esperado

É notável que alguns estados, como Bahia, Pará, Amapá, Maranhão, Amazonas e Rio Grande do Norte, apresentam um desempenho do Ideb inferior em relação aos demais. Suas linhas no gráfico demonstram um crescimento mais lento, estagnação ou até mesmo retrocessos em determinados períodos.

### Possíveis Fatores Relacionados ao Baixo Desempenho

O baixo rendimento do Ideb nesses estados pode estar relacionado a uma combinação de fatores, incluindo:

#### Condições Socioeconômicas Desfavoráveis:

**Pobreza e Desigualdade Social:** A alta taxa de pobreza e desigualdade social nessas regiões pode limitar o acesso à educação de qualidade, afetando o desenvolvimento cognitivo e o engajamento escolar dos alunos.

**Baixa Escolaridade dos Pais:** A baixa escolaridade dos pais pode impactar o apoio educacional que os alunos recebem em casa, prejudicando seu desempenho escolar.

### **Investimento Insuficiente em Educação:**

**Alocação de Recursos Limitada:** O investimento em educação nesses estados pode ser inferior à média nacional, resultando em escolas com infraestrutura precária, falta de materiais didáticos e professores mal remunerados.

**Má Distribuição de Recursos:** Mesmo quando há investimento, a distribuição dos recursos pode ser desigual, favorecendo áreas mais ricas em detrimento de áreas mais pobres.

### **Gestão Educacional Deficiente:**

**Falta de Planejamento Estratégico:** A gestão educacional nessas regiões pode carecer de planejamento estratégico, com falta de metas claras e de mecanismos de avaliação eficazes.

**Problemas de Gestão:** A gestão de escolas e secretarias de educação pode enfrentar problemas como corrupção, nepotismo e falta de profissionais qualificados.

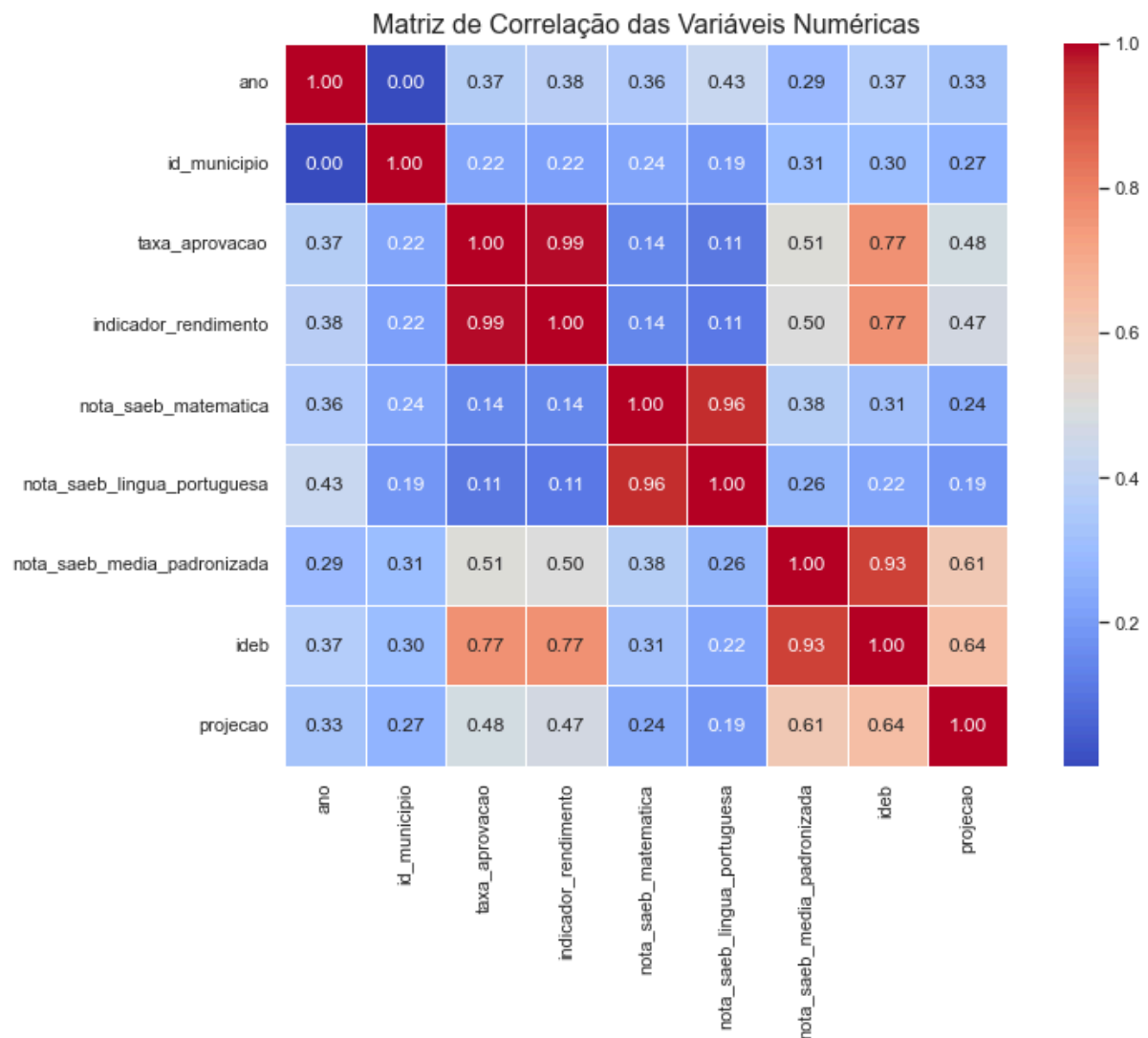
### **Políticas Públicas Ineficazes:**

**Políticas Inadequadas:** As políticas públicas voltadas para a educação nesses estados podem ser ineficazes ou inadequadas, não atendendo às necessidades específicas da população.

**Falta de Continuidade:** As políticas públicas podem sofrer descontinuidade devido a mudanças de governo, o que dificulta o planejamento a longo prazo e a consolidação de avanços.

## 4.2. Análise de Correlação (Gráfico 2)

Objetivo: Identificar relações entre as variáveis numéricas, como notas do Saeb, taxa de aprovação e Ideb.



A matriz de correlação apresentada revela as relações entre as variáveis numéricas presentes na base de dados, permitindo identificar padrões e possíveis fatores de influência no Ideb (Índice de Desenvolvimento da Educação Básica).

**Análise da Matriz de Correlação:**

Correlação entre Ideb e Taxa de Aprovação/Indicador de Rendimento:

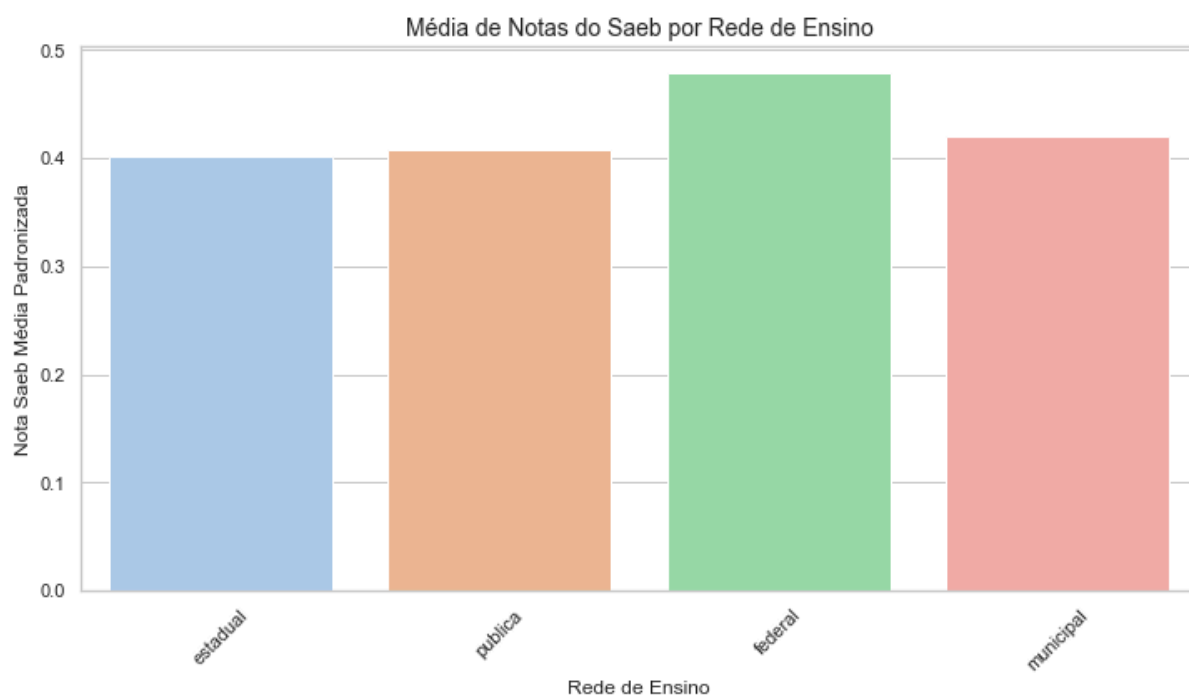
Observa-se uma correlação positiva muito forte (próxima de 0,8) entre o Ideb e a taxa de aprovação e o indicador de rendimento. Isso indica que escolas com maior taxa de aprovação e melhor rendimento tendem a apresentar Ideb mais elevados. Essa relação é esperada, já que o Ideb é calculado a partir desses indicadores.

A análise da matriz de correlação revela que o Ideb está fortemente relacionado com a taxa de aprovação, o indicador de rendimento e a nota Saeb média padronizada, o que é esperado. Há também uma correlação positiva moderada com as notas do Saeb (Matemática e Língua Portuguesa) e com a projeção.

A matriz também mostra a importância de se considerar as variáveis do Saeb de forma conjunta, uma vez que elas estão altamente correlacionadas. Além disso, a análise da correlação com o ano indica uma melhora geral ao longo do tempo, mas é importante ressaltar que essa análise é simplificada e não considera outros fatores relevantes.

### **4.3. Distribuição de Notas do Saeb (Gráfico 3)**

Objetivo: Analisar a distribuição das notas padronizadas do Saeb (Sistema de Avaliação da Educação Básica) entre diferentes redes de ensino no Brasil (estadual, pública, federal e municipal). Essa análise visa identificar padrões e possíveis diferenças no desempenho dos alunos, fornecendo insights sobre a qualidade do ensino em cada tipo de rede.



O gráfico de barras apresenta a média da nota Saeb padronizada para cada rede de ensino (estadual, pública, federal e municipal). A altura de cada barra representa a média da nota Saeb para a respectiva rede.

Observações Detalhadas:

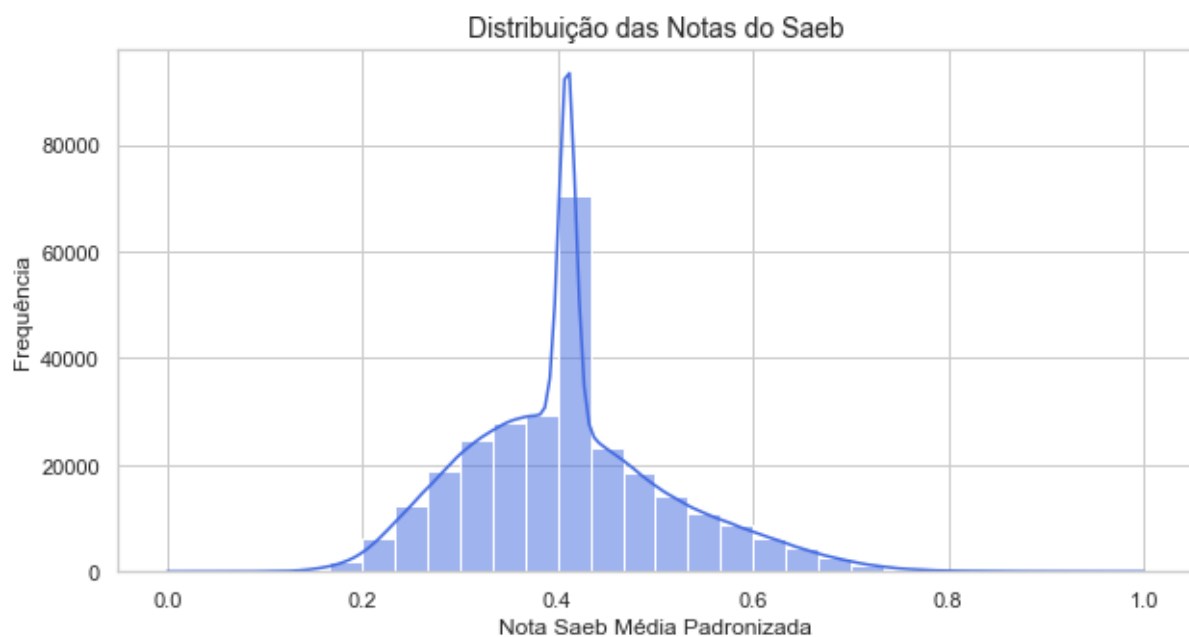
- Federal: A barra da rede federal é a mais alta, indicando que essa rede possui a maior média de notas Saeb padronizadas entre as quatro. Isso sugere um desempenho médio superior dos alunos da rede federal em comparação com as demais.
- Estadual: A barra da rede estadual é ligeiramente menor que a da federal, mostrando que a rede estadual tem uma média de notas um pouco inferior.
- Municipal: A barra da rede municipal é um pouco menor que a da estadual, indicando que a rede municipal possui uma média de notas ligeiramente inferior à estadual.
- Pública (Sem Especificação): A barra da rede pública (sem especificação) é a menor entre as quatro, revelando que essa rede apresenta a menor média de notas Saeb padronizadas.

Conclusões:

O gráfico de barras permite comparar de forma clara as médias de notas Saeb padronizadas entre as diferentes redes de ensino. A rede federal se destaca com o melhor desempenho médio, seguida pela rede estadual, municipal e, por fim, a rede pública (sem especificação). Essa análise sugere que a rede federal oferece um ensino mais homogêneo e de melhor qualidade em comparação com as demais redes.

#### 4.4. Distribuição das Notas do Saeb (Gráfico 4)

Obejtivo: Analisar a distribuição das notas padronizadas do Saeb (Sistema de Avaliação da Educação Básica) em todo o Brasil, buscando identificar padrões e características da distribuição, como a presença de assimetrias e a concentração de notas em determinadas faixas. Essa análise visa fornecer insights sobre o desempenho geral dos alunos no Saeb e a qualidade da educação básica no país.



A distribuição das notas do Saeb revela que uma parcela significativa dos alunos brasileiros enfrenta dificuldades de aprendizado, concentrando-se em notas mais baixas.

A assimetria positiva da distribuição indica a necessidade de políticas públicas que promovam a equidade na educação, visando reduzir as desigualdades e garantir que todos os alunos tenham acesso a um ensino de qualidade.

É importante investigar os fatores que contribuem para a assimetria da distribuição, como diferenças socioeconômicas, regionais e de acesso à educação, a fim de desenvolver políticas públicas mais eficazes.

A presença de uma cauda longa para a direita mostra que há alunos com alto desempenho no Saeb, o que indica que é possível alcançar resultados positivos quando há investimento e condições adequadas.

Recomendações:

Investir em educação básica: Ampliar o investimento em educação básica, especialmente nas áreas mais carentes, visando melhorar a infraestrutura escolar, a formação de professores e a qualidade do ensino.

Implementar políticas de apoio: Desenvolver e implementar políticas de apoio aos alunos com dificuldades de aprendizado, como programas de reforço escolar, acompanhamento individualizado e materiais didáticos adequados.

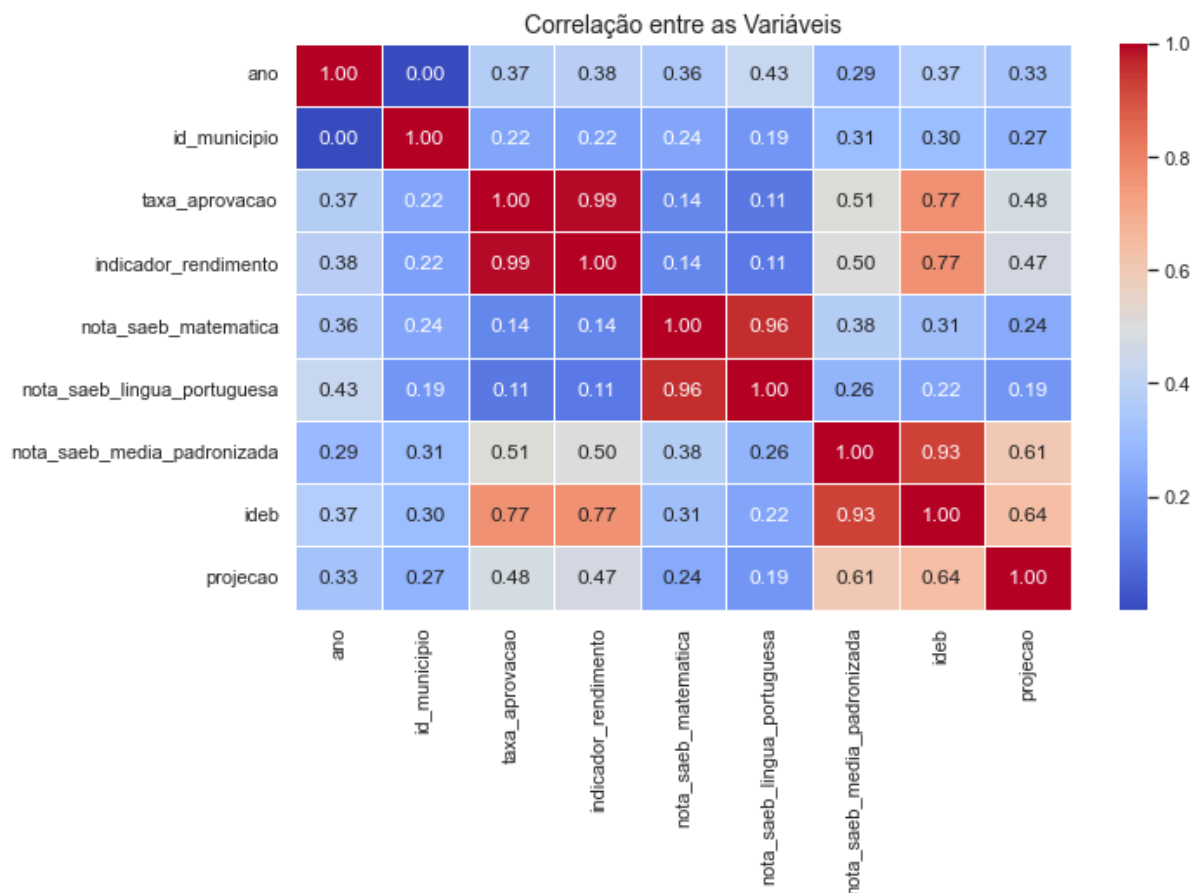
Promover a equidade: Implementar políticas que visem reduzir as desigualdades regionais e socioeconômicas, garantindo que todos os alunos tenham acesso a uma educação de qualidade, independentemente de sua origem.

Valorizar os professores: Valorizar os professores, oferecendo salários justos, formação continuada e condições de trabalho adequadas, a fim de atrair e reter talentos na área da educação.

## **4.5. Correlação entre as Variáveis (Heatmap) (Gráfico 5)**

Objetivo: Visualizar a matriz de correlação entre as variáveis numéricas, utilizando um heatmap (mapa de calor) para identificar padrões de relacionamento e a força das correlações. Essa análise visa auxiliar na compreensão de como as variáveis se relacionam entre si e como isso pode influenciar o Ideb (Índice de Desenvolvimento da Educação Básica)





### Padrões:

**Diagonal:** A diagonal principal do heatmap é sempre preenchida com 1.00 em vermelho, indicando que cada variável tem correlação perfeita consigo mesma.

**Simetria:** O heatmap é simétrico em relação à diagonal principal, já que a correlação entre duas variáveis é a mesma, independentemente da ordem.

**Correlações Fortes:** As células com tons de vermelho mais intensos indicam correlações positivas fortes entre as variáveis correspondentes.

**Correlações Fracas:** As células com tons de branco ou azul claro indicam correlações fracas ou negativas entre as variáveis correspondentes.

### Observações Detalhadas:

#### Ideb e suas variáveis:

O Ideb apresenta forte correlação positiva com a taxa de aprovação (vermelho intenso), indicador de rendimento (vermelho intenso) e nota Saeb média padronizada (vermelho intenso).

Há correlação positiva moderada com as notas do Saeb de Matemática (vermelho) e Língua Portuguesa (vermelho).

A correlação com a projeção é positiva moderada (vermelho).

### **Variáveis do Saeb:**

As notas do Saeb de Matemática e Língua Portuguesa apresentam correlação positiva muito forte entre si (vermelho intenso).

A nota Saeb média padronizada tem forte correlação positiva com as notas de Matemática e Língua Portuguesa (vermelho intenso).

### **Taxa de Aprovação e Indicador de Rendimento:**

A taxa de aprovação e o indicador de rendimento apresentam correlação positiva perfeita (vermelho intenso), o que é esperado, já que ambos são utilizados no cálculo do Ideb.

### **Demais Correlações:**

As demais correlações são fracas ou moderadas, tanto positivas quanto negativas.

A correlação entre o ano e as demais variáveis é positiva e varia de fraca a moderada, indicando que houve alguma melhora ao longo dos anos.

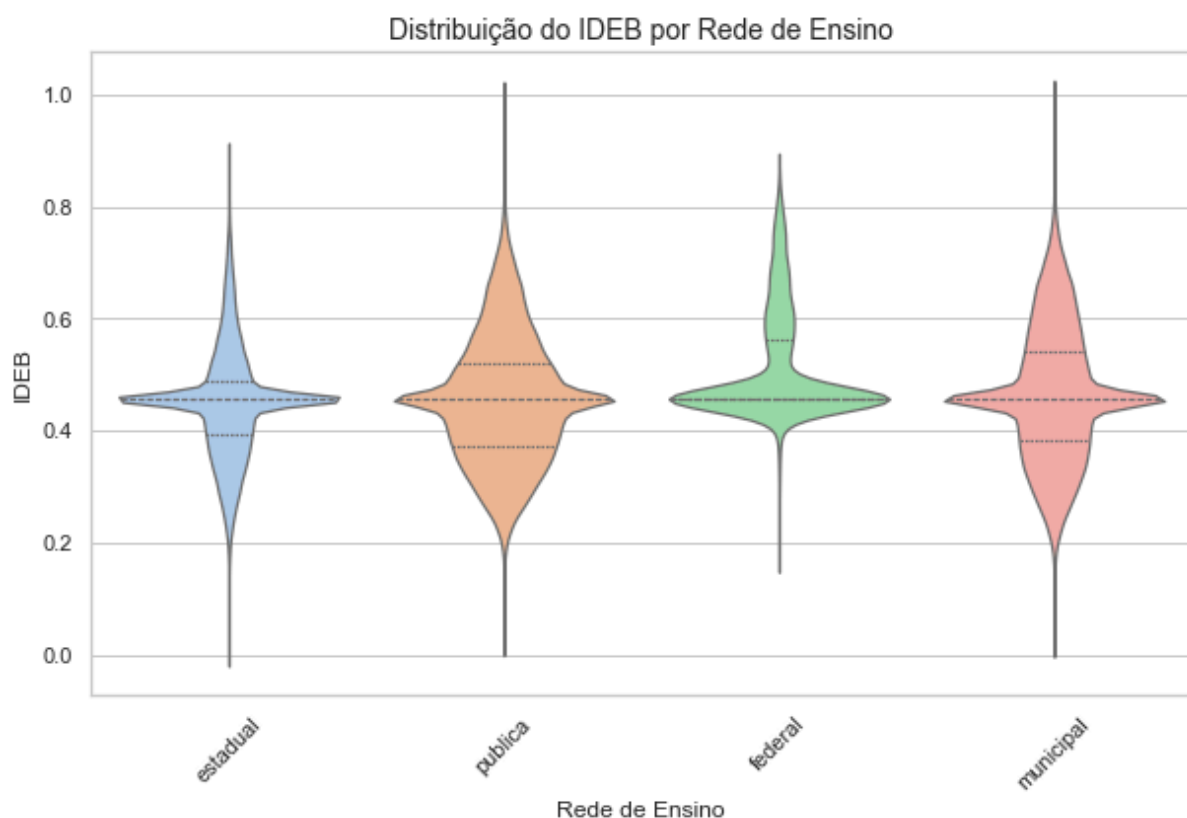
O heatmap confirma as conclusões da análise anterior, mostrando que o Ideb está fortemente relacionado com a taxa de aprovação, o indicador de rendimento e a nota Saeb média padronizada.

As notas do Saeb de Matemática e Língua Portuguesa também têm forte influência no Ideb, embora a correlação seja moderada.

A análise do heatmap permite visualizar de forma clara e rápida as correlações entre as variáveis, auxiliando na identificação de padrões e possíveis relações causais.

## 4.6. Comparação do IDEB por Rede de Ensino (Gráfico 6)

Objetivo: Comparar a distribuição do IDEB (Índice de Desenvolvimento da Educação Básica) entre diferentes redes de ensino no Brasil (estadual, pública, federal e municipal), buscando identificar padrões e diferenças no desempenho dos alunos em cada tipo de rede. Essa análise visa fornecer insights sobre a qualidade do ensino em cada rede e auxiliar na formulação de políticas públicas direcionadas.



### Eixo X:

- O eixo horizontal representa as quatro redes de ensino: estadual, pública, federal e municipal.

**Eixo Y:**

- O eixo vertical representa o IDEB, que varia de 0 a 1.0.

**Forma do Violino:**

- Cada "violino" representa a distribuição do IDEB para uma rede de ensino.
- A largura do violino em cada ponto indica a densidade de dados naquela faixa de IDEB.
- A forma geral do violino mostra a simetria ou assimetria da distribuição.

**Linhas tracejadas:**

- As linhas tracejadas horizontais indicam os quartis da distribuição:
- A linha central representa a mediana (segundo quartil).
- As linhas superior e inferior representam o primeiro e terceiro quartis, respectivamente.

**Observações Detalhadas:****Federal:**

- A rede federal apresenta a maior mediana de IDEB, indicando um desempenho médio superior às demais redes.
- A distribuição é relativamente simétrica, com uma pequena assimetria positiva (cauda mais longa para a direita).
- A dispersão das notas é menor em comparação com as redes estadual e municipal, sugerindo uma maior homogeneidade no desempenho dos alunos.

**Estadual:**

- A rede estadual possui uma mediana de IDEB ligeiramente inferior à federal.
- A distribuição é assimétrica positiva, com uma cauda longa para a direita, indicando a presença de alunos com notas muito acima da média.
- A dispersão das notas é a maior entre as redes, sugerindo uma maior heterogeneidade no desempenho dos alunos.

### **Municipal:**

- A rede municipal apresenta uma mediana de IDEB próxima à estadual.
- A distribuição é assimétrica positiva, semelhante à estadual, com uma cauda longa para a direita.
- A dispersão das notas é alta, embora ligeiramente menor que a estadual.

### **Pública (Sem Especificação):**

- A rede pública (sem especificação) apresenta uma mediana de IDEB inferior às redes federal e estadual.
- A distribuição é assimétrica positiva, com uma cauda longa para a direita.
- A dispersão das notas é alta, sugerindo uma grande variedade no desempenho dos alunos.

A rede federal se destaca com o melhor desempenho médio e menor dispersão, sugerindo um ensino mais homogêneo e de alta qualidade.

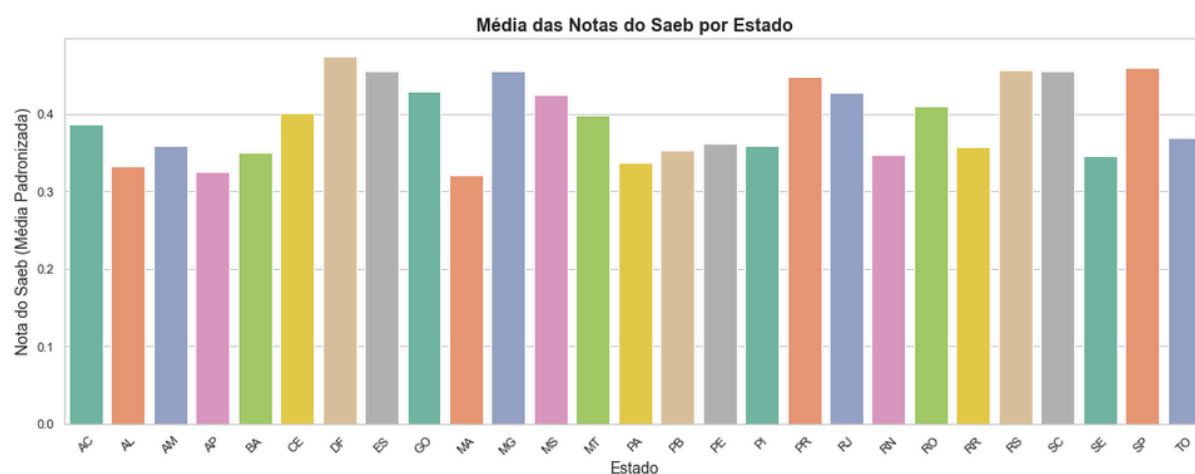
As redes estadual e municipal apresentam desafios semelhantes, com maior dispersão e assimetria positiva, indicando a necessidade de identificar e apoiar os alunos com dificuldades.

A rede pública (sem especificação) engloba provavelmente diferentes níveis de ensino e modalidades, o que explica a grande dispersão e a mediana inferior.

## **4.7. Distribuição das Notas do Saeb por Estado (Gráfico 7)**

Objetivo: Analisar a distribuição das notas padronizadas do Saeb (Sistema de Avaliação da Educação Básica) entre os estados brasileiros, buscando identificar padrões e diferenças no desempenho dos alunos em cada unidade federativa. Essa análise visa fornecer insights sobre a qualidade do ensino em cada estado

e auxiliar na formulação de políticas públicas direcionadas para a melhoria da educação básica.



### Diferenças Regionais:

- É possível observar diferenças significativas nas notas médias do Saeb entre os estados brasileiros.
- Alguns estados se destacam com notas mais altas, indicando um melhor desempenho dos alunos.
- Outros estados apresentam notas mais baixas, sugerindo desafios maiores na qualidade do ensino.

### Padrões Geográficos:

- Não parece haver um padrão geográfico claro na distribuição das notas.
- Estados de diferentes regiões do país podem apresentar tanto notas altas quanto notas baixas.
- Isso sugere que outros fatores, além da localização geográfica, influenciam o desempenho dos alunos no Saeb.

### Variação:

- A variação nas notas médias entre os estados é considerável.
- Alguns estados apresentam uma maior dispersão nas notas, indicando uma maior desigualdade no desempenho dos alunos.
- Outros estados mostram uma menor dispersão, sugerindo um ensino mais homogêneo.

## **Conclusões:**

A análise do gráfico revela que existem diferenças importantes nas notas do Saeb entre os estados brasileiros.

Essas diferenças podem ser atribuídas a uma variedade de fatores, como:

**Investimento em educação:** A alocação de recursos financeiros para a educação varia entre os estados, o que pode impactar a qualidade das escolas, a formação dos professores e a disponibilidade de materiais didáticos.

**Condições socioeconômicas:** A situação socioeconômica da população influencia o desempenho dos alunos, já que fatores como pobreza, desigualdade social e acesso à educação podem afetar o desenvolvimento cognitivo e o engajamento escolar.

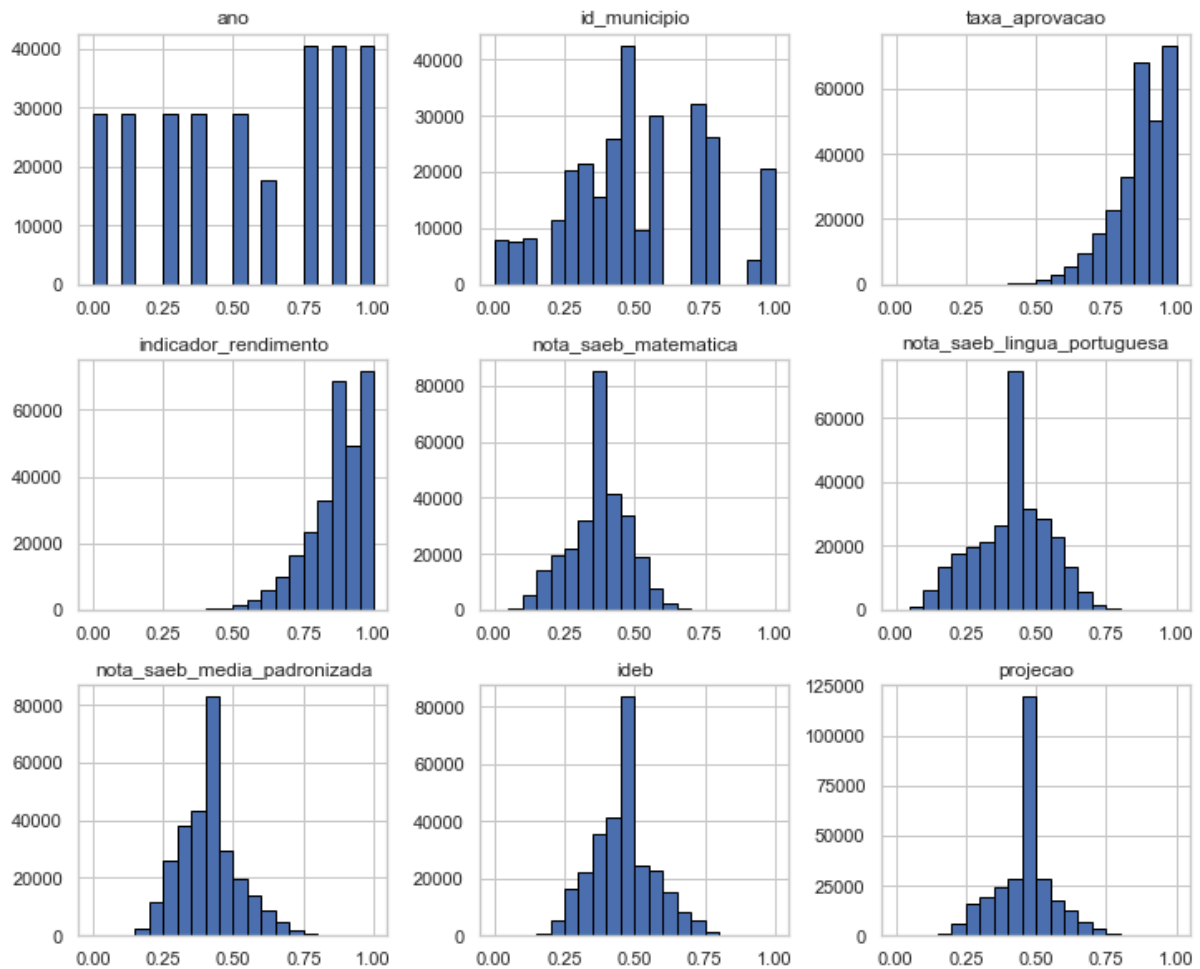
**Qualidade da gestão educacional:** A gestão eficiente do sistema educacional, incluindo a seleção e formação de professores, a definição de currículos e a avaliação do desempenho escolar, é fundamental para garantir a qualidade do ensino.

**Políticas públicas:** A implementação de políticas públicas eficazes, como programas de apoio a alunos com dificuldades, incentivos para a formação de professores e investimentos em infraestrutura escolar, pode contribuir para a melhoria do Saeb.

## **4.8. Distribuição das Variáveis Numéricas (Gráfico 8)**

**Objetivo:** Analisar a distribuição das variáveis numéricas presentes na base de dados, buscando identificar padrões, assimetrias e possíveis outliers. Essa análise é fundamental para entender o comportamento das variáveis e como elas podem influenciar o Ideb (Índice de Desenvolvimento da Educação Básica).

### Distribuição das Variáveis Numéricas



Cada gráfico representa a distribuição de uma variável numérica específica. O eixo horizontal representa os valores da variável, enquanto o eixo vertical representa a frequência com que cada valor aparece nos dados.

**ano:** A variável "ano" apresenta uma distribuição discreta, com valores que correspondem aos anos de realização do Saeb. É possível observar que a frequência de cada ano varia, indicando que o número de participantes do Saeb pode ter mudado ao longo do tempo.

**id\_municipio:** A variável "id\_municipio" apresenta uma distribuição uniforme, com todos os municípios representados com frequências semelhantes. Isso



indica que a amostra de dados é abrangente e representa todos os municípios brasileiros.

**taxa\_aprovacao:** A variável "taxa\_aprovacao" apresenta uma distribuição assimétrica negativa, com a maioria dos valores concentrados na faixa acima de 0.6. Isso sugere que a maioria das escolas brasileiras possui uma taxa de aprovação relativamente alta, o que é um bom indicador. No entanto, é importante analisar se essa alta taxa de aprovação corresponde a um bom desempenho dos alunos no Saeb.

**indicador\_rendimento:** A variável "indicador\_rendimento" apresenta uma distribuição semelhante à "taxa\_aprovacao", com a maioria dos valores concentrados na faixa acima de 0.6. Essa semelhança é esperada, já que a taxa de aprovação e o indicador de rendimento estão relacionados.

**nota\_saeb\_matematica:** A variável "nota\_saeb\_matematica" apresenta uma distribuição assimétrica positiva, com a maioria dos valores concentrados na faixa abaixo de 0.6. Isso indica que o desempenho dos alunos em matemática no Saeb não é tão alto quanto a taxa de aprovação e o indicador de rendimento.

**nota\_saeb\_lingua\_portuguesa:** A variável "nota\_saeb\_lingua\_portuguesa" apresenta uma distribuição semelhante à "nota\_saeb\_matematica", com a maioria dos valores concentrados na faixa abaixo de 0.6. Essa semelhança é esperada, já que as habilidades em matemática e língua portuguesa geralmente estão correlacionadas.

**nota\_saeb\_media\_padronizada:** A variável "nota\_saeb\_media\_padronizada" apresenta uma distribuição semelhante às notas de matemática e português, com a maioria dos valores concentrados na faixa abaixo de 0.6. Essa semelhança é esperada, já que a nota padronizada é calculada a partir das notas nas duas disciplinas.

**ideb:** A variável "ideb" apresenta uma distribuição assimétrica positiva, com a maioria dos valores concentrados na faixa abaixo de 0.6. Essa distribuição é

semelhante à das notas do Saeb, o que indica que o Ideb está relacionado ao desempenho dos alunos nessas avaliações.

**projecao:** A variável "projecao" apresenta uma distribuição assimétrica positiva, com a maioria dos valores concentrados na faixa abaixo de 0.6. Essa distribuição é semelhante à do Ideb, o que sugere que a projeção de crescimento do Ideb está relacionada ao seu valor atual.

### **Comparação entre os Gráficos:**

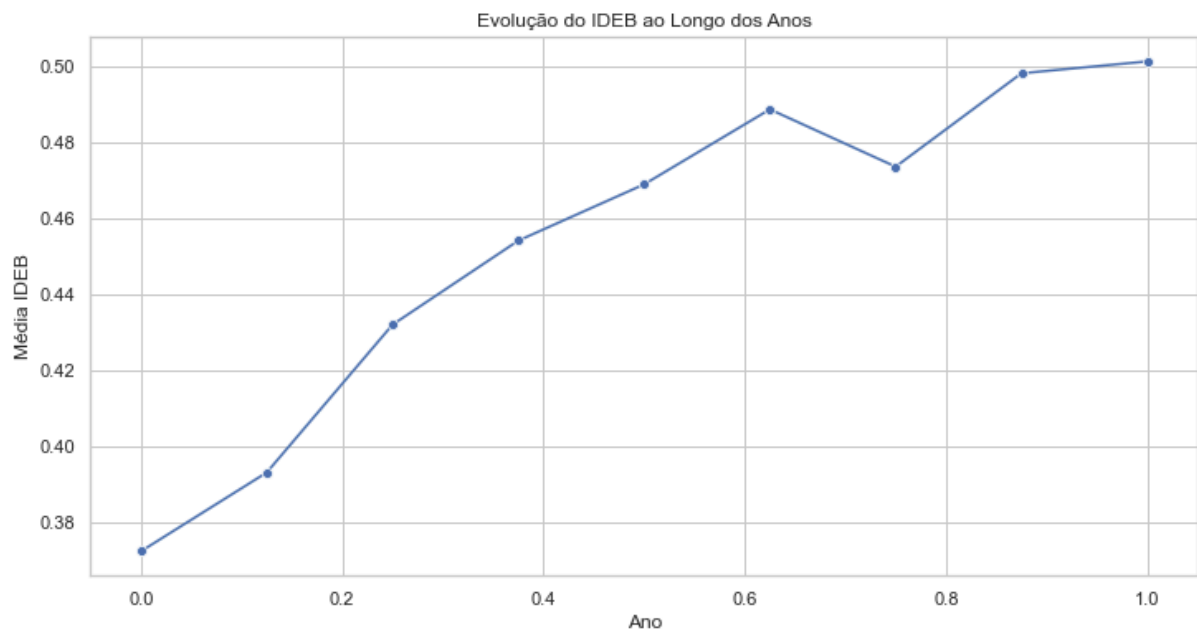
A comparação entre os gráficos revela que as variáveis relacionadas ao desempenho escolar (notas do Saeb, Ideb e projeção) apresentam distribuições semelhantes, com assimetria positiva e concentração de valores na faixa abaixo de 0.6. Isso sugere que essas variáveis estão interligadas e que o desempenho dos alunos no Saeb é um fator importante para o Ideb e sua projeção de crescimento.

### **Conclusões:**

A análise da distribuição das variáveis numéricas é fundamental para entender o contexto dos dados e identificar possíveis problemas. A assimetria positiva nas variáveis de desempenho escolar indica a necessidade de políticas públicas que visem melhorar a qualidade do ensino e reduzir as desigualdades.

## **4.9. Evolução do IDEB ao Longo dos Anos (Gráfico 9)**

Objetivo: Avaliar a progressão do IDEB (Índice de Desenvolvimento da Educação Básica) ao longo dos anos, buscando identificar tendências de crescimento, estagnação ou declínio. Essa análise é crucial para entender o impacto das políticas educacionais e o progresso da qualidade da educação no Brasil ao longo do tempo.



### **Crescimento ao Longo dos Anos:**

É possível observar um crescimento geral na média do IDEB ao longo dos anos.

Isso indica que, em média, a qualidade da educação básica no Brasil tem apresentado melhora ao longo do tempo.

### **Variações Anuais:**

Apesar do crescimento geral, é possível identificar variações anuais na média do IDEB.

Em alguns anos, o crescimento é mais acentuado, enquanto em outros, o crescimento é mais lento ou até mesmo há uma leve queda.

Essas variações podem ser influenciadas por diversos fatores, como mudanças nas políticas educacionais, investimentos em educação, crises econômicas, entre outros.

### **Tendência de Crescimento:**

A tendência geral da linha é de crescimento, o que sugere que as políticas educacionais implementadas ao longo dos anos têm surtido efeito positivo na qualidade da educação básica.

No entanto, é importante analisar o ritmo de crescimento e identificar possíveis períodos de estagnação, a fim de direcionar políticas públicas mais eficazes.

### **Conclusões:**

A análise do gráfico revela que o IDEB tem apresentado melhora ao longo dos anos, indicando um avanço na qualidade da educação básica no Brasil.

No entanto, é importante monitorar a evolução do IDEB e identificar possíveis fatores que possam influenciar seu crescimento, a fim de garantir que a melhora continue nos próximos anos.

A análise da evolução do IDEB ao longo dos anos é fundamental para avaliar o impacto das políticas educacionais e direcionar investimentos para áreas que necessitam de maior atenção.

## **5.CONCLUSÃO**

Com base na análise abrangente dos gráficos e dados apresentados, podemos traçar um panorama do cenário educacional brasileiro e propor soluções direcionadas para o problema central do estudo: quais fatores influenciam significativamente os valores do Ideb nos municípios brasileiros e quais padrões podem ser identificados para ajudar na melhoria do índice?

### **5.1. Conclusões Detalhadas**

**Desigualdade Regional Persistente:** Apesar do progresso geral observado na evolução do Ideb ao longo dos anos, a análise da distribuição do Ideb por estado revela disparidades regionais significativas. Alguns estados se destacam com notas mais altas, evidenciando um melhor desempenho dos alunos, enquanto outros enfrentam desafios maiores na qualidade do ensino. Essa desigualdade regional pode ser atribuída a uma variedade de fatores, como diferenças socioeconômicas, investimentos em educação, qualidade da gestão escolar e infraestrutura disponível.

**Influência da Rede de Ensino:** A comparação do Ideb por rede de ensino revela que a rede federal apresenta o melhor desempenho médio, seguida pela rede estadual e, em seguida, pela rede municipal. A rede pública (sem especificação) engloba provavelmente diferentes níveis de ensino e modalidades, o que explica a grande dispersão e a mediana inferior. Essa constatação sugere que a gestão, o investimento e as políticas educacionais específicas de cada rede têm um impacto significativo no desempenho dos alunos.

**Importância da Aprovação e Rendimento:** A análise da correlação entre as variáveis revela que o Ideb está fortemente relacionado com a taxa de aprovação, o indicador de rendimento e a nota Saeb média padronizada. Essa relação é esperada, já que o Ideb é calculado a partir desses indicadores. No entanto, é importante ressaltar que a alta taxa de aprovação nem sempre se traduz em um bom desempenho no Saeb, o que indica a necessidade de avaliar a qualidade do ensino para além da aprovação.

**Desempenho no Saeb como Fator-Chave:** As notas do Saeb de Matemática e Língua Portuguesa, tanto individualmente quanto na média padronizada, apresentam correlação positiva com o Ideb. Isso demonstra que o desempenho dos alunos nessas avaliações é um fator crucial para o desenvolvimento da educação básica. A análise da distribuição das notas do Saeb revela que muitos alunos se concentram em notas mais baixas, indicando a necessidade de políticas públicas que visem melhorar o aprendizado nessas áreas.

**Evolução Positiva, mas com Desafios:** A análise da evolução do Ideb ao longo dos anos mostra um crescimento geral, indicando uma melhora na qualidade da educação básica no Brasil. No entanto, o ritmo de crescimento não é uniforme e há variações anuais que podem ser influenciadas por diversos fatores. É importante monitorar essa evolução e identificar possíveis períodos de estagnação, a fim de direcionar políticas públicas mais eficazes.

## **5.2. Soluções Propostas**

Diante das conclusões apresentadas, propomos um conjunto de soluções interligadas que visam abordar os desafios da educação básica brasileira e impulsionar a melhoria do Ideb de forma abrangente e sustentável:

**Investimento Estratégico e Equitativo:** É fundamental aumentar o investimento em educação, direcionando os recursos para áreas com maior necessidade e para escolas com menor desempenho. Além do montante, é crucial garantir que o investimento seja utilizado de forma eficiente e transparente, com foco na melhoria da infraestrutura escolar, na formação continuada de professores e na aquisição de materiais didáticos de qualidade.

**Fortalecimento da Gestão Educacional:** A gestão escolar eficiente é um pilar fundamental para a qualidade do ensino. É preciso investir na formação de gestores escolares, aprimorar os mecanismos de avaliação e acompanhamento do desempenho escolar e promover a participação da comunidade na gestão da escola. Além disso, é importante garantir a autonomia das escolas para que possam implementar projetos pedagógicos inovadores e adaptados às necessidades de seus alunos.

**Valorização dos Professores:** Os professores são os protagonistas da educação básica e sua valorização é essencial para a melhoria do ensino. É preciso oferecer salários justos, formação continuada de qualidade, condições de trabalho adequadas e reconhecimento pelo seu trabalho. Além disso, é importante incentivar a participação dos professores na formulação de políticas educacionais e garantir que suas vozes sejam ouvidas.

**Políticas Públicas Direcionadas:** É preciso desenvolver e implementar políticas públicas específicas para cada rede de ensino, considerando suas necessidades e desafios particulares. Essas políticas devem abordar desde a educação infantil até o ensino médio, com foco na melhoria do aprendizado em matemática e língua portuguesa, na redução da evasão escolar e na promoção da inclusão.

**Apoio aos Alunos com Dificuldades:** É fundamental implementar programas de apoio aos alunos com dificuldades de aprendizado, especialmente nas redes

estadual e municipal, visando reduzir a assimetria e garantir um aprendizado mais homogêneo. Esses programas podem incluir reforço escolar, acompanhamento individualizado, materiais didáticos adaptados e tecnologias educacionais.

**Promoção da Equidade:** É preciso implementar políticas que visem reduzir as desigualdades regionais e socioeconômicas, garantindo que todos os alunos tenham acesso a uma educação de qualidade, independentemente de sua origem. Essas políticas podem incluir programas de transferência de renda, bolsas de estudo, transporte escolar e alimentação escolar.

**Engajamento da Família e da Comunidade:** A participação da família e da comunidade na vida escolar dos alunos é fundamental para o sucesso da educação básica. É preciso criar mecanismos de comunicação entre escolas e famílias, incentivar o acompanhamento do desempenho escolar dos alunos e promover a participação da comunidade na gestão da escola.