



Universidade Federal de Santa Catarina
Centro Tecnológico
Depto. de Informática e Estatística
INE 5405 - Probabilidade e Estatística

Índice de Desenvolvimento da Educação Básica (Ideb)

Trabalho 3

Autor(es): Jessica Regina dos Santos, Francisco Bortolanza e Myllena da Conceição Correa

Florianópolis, 2024

Índice

Introdução.....	3
Materiais e Métodos.....	4
Resultados e Discussões.....	5
Considerações Finais.....	17
Referências Bibliográficas.....	18

Introdução

O Índice de Desenvolvimento da Educação Básica (Ideb), criado em 2007 pelo Ministério da Educação do Brasil, é uma ferramenta estratégica que sintetiza a qualidade da educação básica em um único indicador. Ele combina dois conceitos fundamentais: a taxa de aprovação escolar, obtida pelo Censo Escolar, e a média de desempenho em avaliações bianuais de larga escala realizadas pelo Sistema de Avaliação da Educação Básica (Saeb).

Com uma escala de 0 a 10, o Ideb busca atingir a meta de média 6, correspondente a um sistema educacional de qualidade comparável ao de países desenvolvidos. O índice reflete tanto aspectos pedagógicos quanto socioeconômicos, propondo um equilíbrio entre a aprovação escolar e a qualidade do aprendizado. Ele também permite identificar desequilíbrios: sistemas que privilegiam exclusivamente a aprovação ou a retenção acabam demonstrando necessidade de ajustes, seja na gestão do fluxo escolar ou na qualidade da aprendizagem.

Além de ser um termômetro para reportar desigualdades educacionais e sociais, o Ideb é uma ferramenta que orienta políticas públicas voltadas à melhoria da qualidade da educação no Brasil. Este trabalho busca explorar os dados do Ideb entre os anos de 2005 e 2023, analisando estatísticas descritivas, formulando hipóteses e executando testes estatísticos para interpretar as informações de maneira estruturada e rigorosa.

Objetivos

Os objetivos deste estudo são:

- Compreender o comportamento histórico do Ideb, calculando estatísticas descritivas que incluam estimativas pontuais e intervalares para médias e proporções relevantes.
- Levantar hipóteses estatísticas, considerando fatores que possam estar associados ao desempenho do Ideb.
- Construir e executar testes de hipóteses, avaliando relações significativas nos dados e discutindo seus resultados.
- Apresentar conclusões baseadas em evidências, integrando análises descritivas e inferenciais para identificar tendências e propor insights sobre a qualidade da educação básica no Brasil.

Materiais e Métodos

A linguagem de programação escolhida pelo grupo foi Python. Utilizamos as bibliotecas de software *pandas*, *scipy.stats*, *matplotlib.pyplot*, *statsmodels.api*, *seaborn* e *numpy* para a análise e visualização dos dados. Essas bibliotecas oferecem ferramentas robustas para manipulação de dados tabulares, execução de análises estatísticas, construção de gráficos e visualizações avançadas, além de operações matemáticas e modelagem estatística. A combinação dessas bibliotecas possibilitou uma abordagem abrangente e eficiente para explorar e apresentar os resultados da base de dados.

A seguir, encontra-se a descrição e classificação das variáveis, assim como suas finalidades.

Variáveis Qualitativas

- Rede Escolar (rede): tipo de dependência administrativa, como: pública, privada, estadual e municipal.
 - Finalidade: diferenciar o desempenho entre escolas com diferentes gestões e níveis de investimento.
 - Tipo: qualitativa nominal (string).
- Tipo de Ensino (ensino): classificação do nível de ensino (fundamental I, fundamental II e médio).
 - Finalidade: identificar o desempenho por nível de ensino.
 - Tipo: qualitativa nominal (string).
- Anos Escolares (anos escolares): ciclo escolar dos alunos avaliados.
 - Finalidade: identificar o desempenho por ciclo escolar.
 - Tipo: qualitativa ordinal (string).

Variáveis Quantitativas

- Ano (ano): representa o período de coleta dos dados ao longo do tempo (2005-2021).
 - Finalidade: permite o acompanhamento do desempenho educacional; análise temporal das variações e tendências.
 - Tipo: quantitativa discreta (int64).
- Taxa de Aprovação (taxa aprovação): percentual de alunos aprovados ao final de cada ciclo escolar.
 - Finalidade: avaliar a eficácia do ensino em garantir a qualificação e aprovação dos alunos.
 - Tipo: quantitativa contínua (float64).
- Índice de Rendimento (indicador rendimento): proporção de alunos aprovados em cada ciclo escolar.
 - Finalidade: relacionar as taxas de aprovação, reprovação e abandono dos alunos.

- Tipo: quantitativa contínua (float64).
- Nota SAEB - Matemática (nota_saeb_matematica): nota obtida pelos alunos na disciplina de Matemática no Sistema de Avaliação da Educação Básica (SAEB).
 - Finalidade: medir conhecimentos específicos.
 - Tipo: qualitativa contínua (float64).
- Nota SAEB - Língua Portuguesa (nota_saeb_lingua_portuguesa): nota obtida pelos alunos na disciplina de Língua Portuguesa no Sistema de Avaliação da Educação Básica (SAEB).
 - Finalidade: medir conhecimentos específicos.
 - Tipo: qualitativa contínua (float64).
- Nota SAEB - Média Padronizada (nota_saeb_media_padronizada): média das notas de Matemática e Língua Portuguesa no SAEB.
 - Finalidade: fornecer uma métrica conjunta de desempenho nas duas disciplinas-chave.
 - Tipo: Quantitativa contínua (float64).
- Ideb (ideb): combina o rendimento dos alunos (aprovado ou não aprovado) com o desempenho acadêmico (nota do SAEB), formando uma medida de qualidade educacional.
 - Finalidade: avaliar o desempenho das escolas e seu progresso ao longo dos anos.
 - Tipo: quantitativa contínua (float64).

Resultados e Discussões

Nosso trabalho apresenta uma análise estatística descritiva e inferencial sobre um conjunto de dados educacionais, contendo 99 observações e 11 variáveis.

As variáveis analisadas incluem taxa de aprovação, ano, rede escolar, notas do SAEB (para Matemática e Língua Portuguesa), Ideb, entre outras.

Nosso objetivo principal é extrair informações relevantes sobre o desempenho educacional no Brasil e identificar possíveis desigualdades entre escolas públicas e privadas, além de tendências gerais.

Variáveis como “taxa_aprovacao”, “nota_saeb_matematica”, “nota_saeb_lingua_portuguesa”, “ideb”, etc, foram escolhidas para análise porque são quantitativas e contínuas, permitindo o cálculo de medidas estatísticas, como médias, desvios padrão e intervalos de confiança. Essas métricas são fundamentais para entender o desempenho educacional em aspectos como aprovação e resultados de aprendizagem em disciplinas-chave, oferecendo resultados importantes sobre a eficácia do sistema educacional. Por outro lado, variáveis categóricas (como “rede”) não são adequadas para esse tipo de análise direta, pois não possuem uma escala contínua. Essas variáveis são mais adequadas para as análises de proporções, distribuições ou classificações, que exploram diferenças entre grupos ou categorias.

Nossa análise foi fundamentada nos conceitos teóricos abordados ao longo do semestre na disciplina de Probabilidade e Estatística.

1. Análise Descritiva

Nossa análise inicial incluiu a inspeção dos dados e o cálculo de medidas descritivas como média, mediana, variância e desvio padrão para variáveis numéricas. Além disso, intervalos de confiança foram calculados para as médias das variáveis selecionadas.

As principais fórmulas utilizadas foram:

- Erro padrão da média (SE):

$$SE = s / \sqrt{n}$$

Onde:

- s é o desvio padrão da amostra.

$$s = \sqrt{[\Sigma(x - \bar{x})^2] / n - 1}$$

- n é o tamanho da amostra.
- Intervalo de Confiança:
Para um nível de confiança $1 - \alpha$, o IC é dado por:

$$IC(\mu; 100\gamma\%) = [\bar{x} \pm t(n - 1) \cdot SE]$$

Onde \bar{x} é a média amostral, $t(n - 1)$ é o valor crítico da distribuição t-Student e SE o erro padrão da média.

Por que usamos a distribuição t-Student (em vez da normal)?

Na maioria das situações práticas, o desvio padrão da população (σ) não é conhecido. Por isso, utilizamos o desvio padrão amostral (s) como uma estimativa. Essa substituição altera a distribuição de probabilidade associada ao cálculo do intervalo de confiança. Em vez de usar a distribuição normal (z), passamos a usar a t-Student, que leva em consideração a incerteza adicional introduzida pela estimativa do desvio padrão.

A distribuição t-Student é mais adequada especialmente quando temos amostras pequenas ($n \leq 30$), pois ela é mais "larga", refletindo o maior grau de incerteza. Mesmo para amostras grandes ($n > 30$), onde a t se aproxima da distribuição normal, sua utilização ainda garante maior precisão quando o desvio padrão populacional não é conhecido.

Portanto, o uso da t-Student é essencial para garantir que o nível de confiança desejado seja realmente alcançado, compensando a estimativa do desvio padrão populacional.

2. Resultados

Estatísticas Descritivas e Intervalos de Confiança

- **Taxa de Aprovação:**
 - Média de 87,97%, com desvio padrão de 8,19%. Varia de 70,5% a 99,1%.
 - Intervalo de confiança ($\approx 95\%$): [86,34%, 89,60%].
Em resumo: O intervalo de confiança de $\approx 95\%$ significa que, se repetíssemos o processo de amostragem e cálculo do intervalo inúmeras vezes, cerca de 95% desses intervalos calculados conteriam a verdadeira média populacional da taxa de aprovação. Em outras palavras, o intervalo [86,34%, 89,60%] é uma estimativa da faixa onde está localizada a média verdadeira da taxa de aprovação para todas as escolas (não apenas para as observadas na amostra), com 95% de confiança. Isso implica que temos uma margem de incerteza controlada baseada nos dados e no tamanho da amostra, mas não podemos afirmar com absoluta certeza que esse intervalo

contém a média populacional — existe uma chance de 5% de que ela esteja fora dele.

Com 95% de confiança, a verdadeira taxa média de aprovação para todas as escolas está entre 86,34% e 89,60%. Isso demonstra que, em geral, a aprovação dos estudantes é alta. Temos baixa variabilidade nos dados, o que é bem bom. A largura do intervalo é relativamente estreita ($\approx 3,26$ pontos percentuais), sugerindo que a estimativa da taxa média é precisa e que há pouca dispersão nos dados relacionados à aprovação. Essa consistência indica que poucas escolas se desviam significativamente da média.

Também temos um desempenho geral positivo. O intervalo demonstra que a maioria das escolas têm taxas de aprovação satisfatórias, reforçando a ideia de que os estudantes estão avançando nos níveis de ensino com frequência regular.

A alta média de 87,97%, aliada a uma dispersão relativamente baixa (desvio padrão de 8,19%), sugere que o desempenho das escolas é consistente em termos de aprovação. A maioria das escolas apresenta taxas acima de 70%, indicando que poucas instituições enfrentam desafios significativos. Contudo, seria interessante investigar as razões das taxas mais baixas em algumas escolas para identificar fatores que possam estar limitando o desempenho.

- **Nota SAEB - Matemática:**

- Média de 250,08, desvio padrão de 37,13, com valores de 174,86 a 334,72.
- Intervalo de confiança ($\approx 95\%$): [242,67, 257,48]. Com 95% de confiança, a média verdadeira da Nota SAEB - Matemática está nesse intervalo. Isso significa que, ao considerar uma amostra representativa, é altamente provável que a média real de todas as escolas esteja entre 242,67 e 257,48. Embora a média não seja considerada baixa, ela também não indica um desempenho excepcional, especialmente se houver metas de aprendizado estabelecidas que esperem notas significativamente mais altas. O intervalo relativamente estreito (diferença de aproximadamente 15 pontos) sugere uma boa precisão na estimativa da média com base nos dados coletados. A média de 250,08, combinada com uma maior variabilidade (desvio padrão de 37,13), aponta para desigualdades significativas no desempenho em matemática entre as escolas. Essa dispersão indica que algumas instituições estão alcançando excelentes resultados, enquanto outras enfrentam dificuldades consideráveis.

- **Nota SAEB - Língua Portuguesa:**

- Média: Média de 241,01, desvio padrão de 38,20, variando de 165,07 a 322,10.
- Intervalo de confiança ($\approx 95\%$): [233,39, 248,63]. Com 95% de confiança, a média verdadeira das notas de Língua Portuguesa para todas as escolas está entre 233,39 e 248,63. Isso indica um desempenho moderado, próximo ao que foi observado para Matemática. Houve variação menor no intervalo: a largura do intervalo ($\approx 15,24$ pontos) sugere que os dados amostrais fornecem uma estimativa relativamente precisa da média. O intervalo é ligeiramente inferior ao da Nota SAEB - Matemática ([242,67, 257,48]). Isso pode indicar que, em geral, o desempenho em Língua Portuguesa é um

pouco mais baixo, reforçando a necessidade de atenção específica a essa área.

Em geral, a situação em Língua Portuguesa reflete padrões semelhantes aos de Matemática, com uma média de 241,01 e uma dispersão significativa (desvio padrão de 38,20). Isso reforça a necessidade de atenção às disparidades educacionais. Identificar e atender as escolas com desempenho mais baixo é essencial para promover uma maior equidade no aprendizado.

- **Ideb:**

- Média: 4,75, desvio padrão: 1,13.
- Valores variam de 3,00 a 7,10, com projeções médias de 4,83.
Com uma média de 4,75 e valores variando de 3,00 a 7,10, o IDEB reflete um desempenho geral mediano das escolas. Embora algumas instituições estejam próximas ou acima das projeções médias (4,83), outras estão significativamente abaixo, indicando disparidades no alcance das metas educacionais. Esse cenário sugere a necessidade de intervenções direcionadas para garantir que todas as escolas tenham acesso a recursos e suporte suficientes para atingir suas projeções e melhorar o desempenho.
- Intervalo de confiança ($\approx 95\%$): [4.53, 4.98] indica que a média real, estimada em 4.75, está ligeiramente abaixo da projeção de 4.83, refletindo um desempenho geral mediano com disparidades significativas entre as escolas (valores variam de 3.00 a 7.10). Essa heterogeneidade sugere desigualdades no acesso a recursos e qualidade educacional. Enquanto algumas instituições estão próximas ou acima das metas, outras ficam bem abaixo, reforçando a necessidade de intervenções específicas para apoiar escolas com menor desempenho. Com ações focadas em recursos, monitoramento e compartilhamento de boas práticas, é possível alinhar o desempenho geral às metas projetadas.

3. Testes de Hipótese

Hipótese para a Média:

O objetivo desta análise é comparar as médias do IDEB entre escolas públicas e privadas.

- H_0 : Não há diferença significativa na média do Ideb entre as escolas públicas e privadas ($\mu_{pública} = \mu_{privada}$).
- H_1 : Há diferença significativa na média do Ideb ($\mu_{pública} \neq \mu_{privada}$).

Estatísticas Descritivas:

Para iniciar a análise, foram calculadas as seguintes estatísticas descritivas para as redes pública e privada:

- Médias: (valores médios do IDEB).
 - Rede pública: $\bar{x}_{pública} = 4,12$
 - Rede privada: $\bar{x}_{privada} = 6,10$
- Desvios padrão: (medida de dispersão).
 - Rede pública: $s_{pública} = 0,78$
 - Rede privada: $s_{privada} = 0,51$
- Tamanho das amostras: (os dados abrangem o ensino médio, fundamental I e fundamental II, coletados entre 2005 e 2021).
 - Rede pública: $n_{pública} = 27$
 - Rede privada: $n_{privada} = 27$
- Intervalos de confiança para as médias: (calculado para 95% de confiança).
 - Rede pública: $3,83 \leq \mu \leq 4,42$
 - Rede privada: $5,91 \leq \mu \leq 6,29$

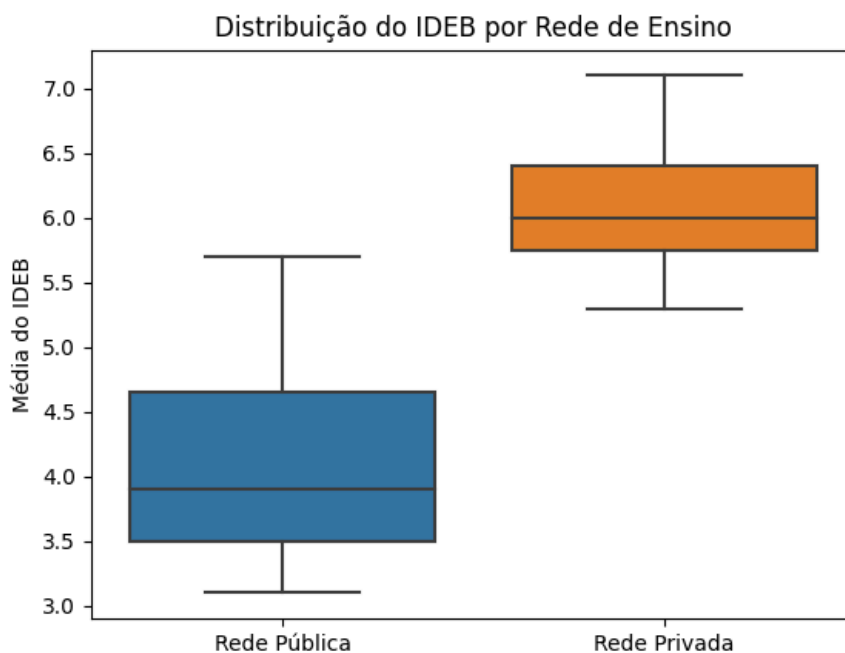


Figura 1. Gráfico de distribuição das médias do Ideb entre as redes pública e privada.

Análises Preliminares:

Antes de aplicar o teste t de Student, foram realizados testes para verificar a normalidade dos dados e a homogeneidade das variâncias, condições fundamentais para o uso do teste t clássico.

1. Teste de Normalidade (Shapiro-Wilk):
 - Hipóteses:

- H_0 : Os dados seguem uma distribuição normal.
- H_1 : Os dados não seguem uma distribuição normal.
- Resultados:
 - Rede pública: $W = 0,9180$, $p = 0,0354$. Como $p < 0,05$, rejeitamos H_0 e portanto os dados não seguem uma distribuição normal.
 - Rede privada: $W = 0,9353$, $p = 0,0932$. Como $p > 0,05$, aceitamos H_0 e portanto os dados seguem uma distribuição normal.

2. Teste de Homogeneidade de Variâncias (Levene):

- Hipóteses:
 - H_0 : As variâncias das duas amostras (pública e privada) são iguais.
 - H_1 : As variâncias das duas amostras são diferentes.
- Resultados:
 - $F = 3,6058$ e $p = 0,0631$. Como $p > 0,05$, aceitamos H_0 e portanto as variâncias das duas amostras podem ser consideradas homogêneas.

Embora os dados da rede pública não apresentem normalidade, o teste t é robusto em relação a desvios moderados dessa condição, especialmente quando as amostras têm tamanhos iguais e razoavelmente grandes, como no caso analisado ($n = 27$).

Aplicação do Teste t de Student:

- Hipóteses:
 - $H_0: \mu_{pública} = \mu_{privada}$ (não há diferença significativa entre as médias).
 - $H_1: \mu_{pública} \neq \mu_{privada}$ (há diferença significativa entre as médias).
- Resultados:
 - $t = -11,0358$ e $p = 0,0000$. Como $p < 0,05$, rejeitamos H_0 e portanto concluímos que há uma diferença significativa entre as médias do IDEB das escolas públicas e privadas.

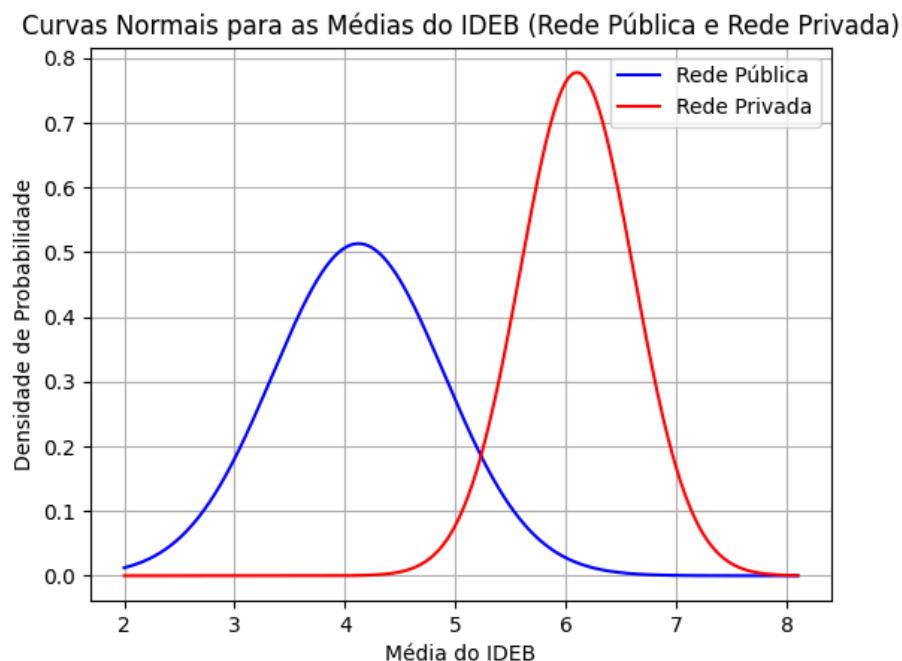


Figura 2. Gráfico de distribuição normal das médias do Ideb das redes pública e privada.

Conclusão:

A aplicação do teste t de Student nos permitiu testar a hipótese de que não há diferença significativa entre as médias do IDEB das redes pública e privada. O valor p obtido ($p = 0,0000$), que é inferior ao nível de significância de 0,05 (5%), nos leva a **rejeitar a hipótese nula (H_0)**. Assim, podemos concluir que há uma **diferença significativa** entre as médias do IDEB das duas redes de ensino, o que nos leva a **aceitar a hipótese alternativa (H_1)**.

Esse resultado evidencia uma disparidade no desempenho educacional entre as redes pública e privada no período analisado, indicando possíveis diferenças na qualidade do ensino e no acesso aos recursos educacionais.

Hipótese para Proporção:

O objetivo desta análise é verificar se a proporção de escolas com notas do SAEB em Matemática acima de 300 é diferente entre as redes públicas e privadas.

- H_0 : A proporção de escolas com nota acima de 250 difere significativamente entre as redes pública e privada ($P_{pública} = P_{privada}$).
- H_1 : A proporção de escolas com nota acima de 250 difere significativamente entre as redes pública e privada ($P_{pública} \neq P_{privada}$).

Proporções amostrais:

Para iniciar a análise, foram calculadas as proporções amostrais para as escolas da rede pública e privada, considerando os dados coletados entre 2005 e 2021, abrangendo os níveis de ensino fundamental I, fundamental II e ensino médio.

- Tamanho das amostras: (os dados abrangem o ensino médio, fundamental I e fundamental II, coletados entre 2005 e 2021).
 - Rede pública ($n_{pública}$): 27
 - Rede privada ($n_{privada}$): 27
- Número de amostras com nota acima de 250:
 - Rede pública ($n_{pública \text{ com nota } > 250}$): 13
 - Rede privada ($n_{privada \text{ com nota } > 250}$): 21
- Proporções amostrais:
 - $\hat{p}_{pública} = (n_{pública \text{ com nota } > 250}) / n_{pública} = 13/27 \approx 0,4815$
 - $\hat{p}_{privada} = (n_{privada \text{ com nota } > 250}) / n_{privada} = 21/27 \approx 0,7778$

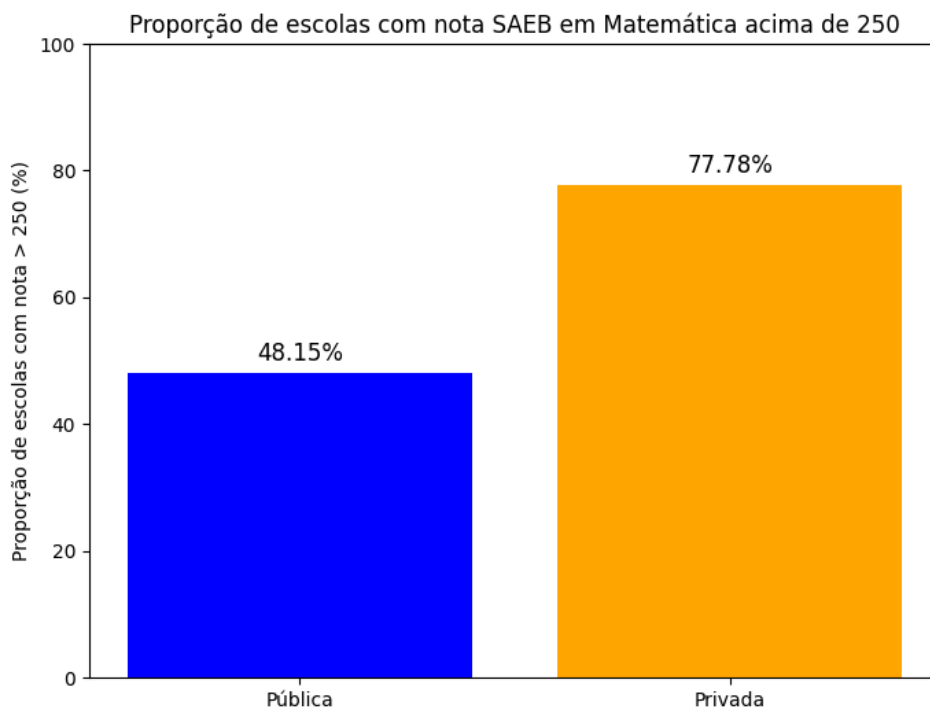


Figura 3. Gráfico de proporções de escolas com nota superior a 250 nas redes pública e privada.

Cálculo do erro padrão:

O erro padrão da diferença entre as proporções é calculado pela fórmula:

$$SE = \sqrt{[(\hat{p}_{pública}(1 - \hat{p}_{pública}) / n_{pública}) + (\hat{p}_{privada}(1 - \hat{p}_{privada}) / n_{privada})]}$$

Substituindo os valores:

$$SE = \sqrt{[(0,4815(1 - 0,4815) / 27) + (0,7778(1 - 0,7778) / 27)]} \approx 0,1973$$

Cálculo da Estatística de Teste z:

A estatística de teste z para a diferença entre as proporções é dada por:

$$z = (\hat{p}_{pública} - \hat{p}_{privada}) / SE = (0,4815 - 0,7778) / 0,1973 \approx - 2,3686$$

Determinação do valor-p:

Com base na estatística $z = - 2,3686$, o valor-p é aproximadamente 0,0179. Esse valor pode ser obtido a partir da tabela de distribuição normal padrão ou utilizando uma ferramenta estatística.

Conclusão:

Como o valor-p (0,0179) é menor que o nível de significância de 0,05 (5%), rejeitamos a hipótese nula (H_0). Ou seja, as proporções de escolas com nota acima de 250 no SAEB em matemática são **significativamente diferentes** entre as redes pública e privada. Em particular, as escolas privadas apresentam uma maior proporção de escolas com desempenho superior a 250.

Além disso, vale destacar que a nota do SAEB inferior a 250 indica uma deficiência na capacidade dos alunos de assimilar conceitos básicos de Matemática. Segundo o documento **"Escalas de Proficiência do SAEB"** publicado em 2020, notas abaixo de 250 refletem uma dificuldade em compreender os conceitos fundamentais exigidos para o avanço no ensino fundamental I, fundamental II e ensino médio. Isso sugere que, na rede pública, um número considerável de alunos não está atingindo os níveis de proficiência esperados para seu ciclo de aprendizado, o que é um indicativo de que há defasagens significativas no aprendizado.

Hipótese para Correlação ou Modelo de Regressão:

O objetivo desta análise é verificar se há uma tendência significativa de crescimento na média geral das notas do Ideb ao longo dos anos.

- H_0 : A média geral da nota do Ideb não apresenta tendência significativa de crescimento ao longo dos anos de 2005 a 2021, ou seja, o coeficiente da reta de regressão (β) é igual a zero ($\beta = 0$).
- H_1 : A média geral da nota do Ideb apresenta uma tendência significativa de crescimento ao longo dos anos de 2005 a 2021, ou seja, o coeficiente da reta de regressão (β) é diferente de zero ($\beta \neq 0$).

Cálculo e análise de Regressão Linear:

Para testar a hipótese, foi realizada uma regressão linear simples, onde a variável "ano" (variável independente) foi analisada em relação ao Ideb (variável dependente). O coeficiente da variável "ano" (β) da regressão linear representa a taxa de variação da média das notas do Ideb ao longo dos anos.

- Coeficiente da variável 'ano' (β): 0,0922, indicando que a cada ano, o Ideb tende a aumentar em 0,0857 pontos.
- Valor-p do coeficiente do 'ano': 0,000, que é muito inferior ao nível de significância de 0,05 (5%), indicando que o coeficiente é estatisticamente significativo.

Coeficiente	Valor	Erro Padrão	t-Valor	Valor-p
Ano	0,0922	0,023	4,046	0,0000

Figura 4. Tabela de resumo dos coeficientes de regressão calculados.

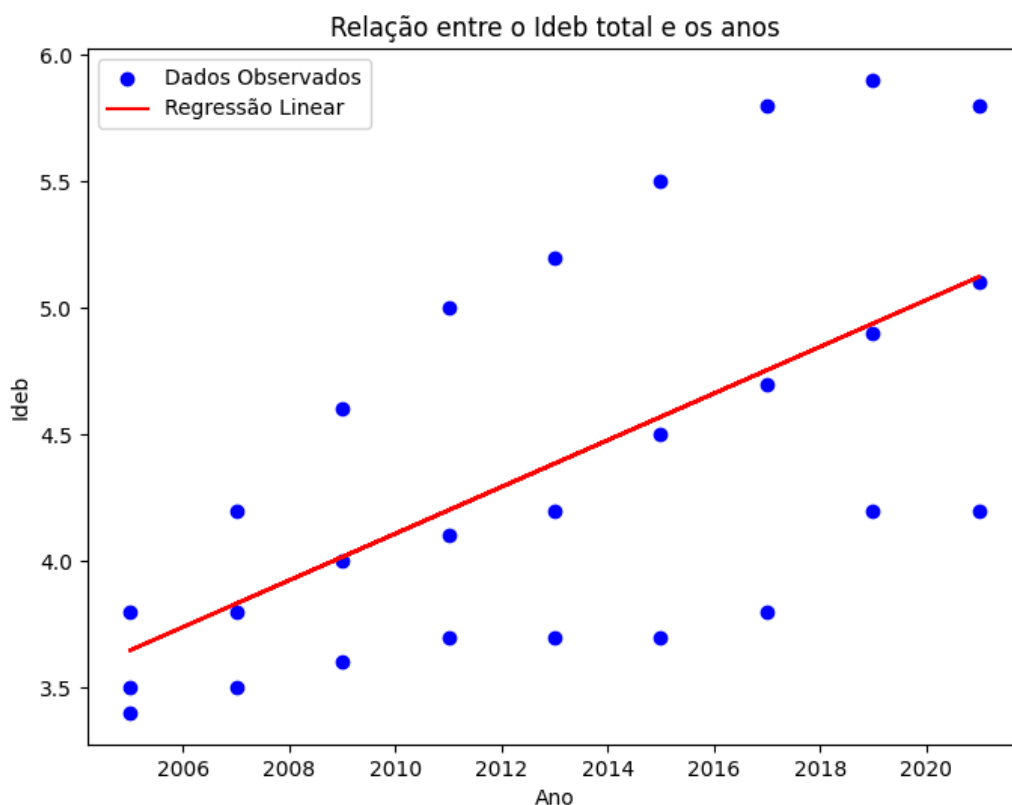


Figura 5. Gráfico de relação entre a média geral do Ideb e os anos.

Conclusão:

Com base no valor-p extremamente baixo (0,000) para o coeficiente da variável "ano", rejeitamos a hipótese nula (H_0). Assim, podemos concluir que a média geral das notas do Ideb apresenta uma tendência **significativa de crescimento ao longo dos anos** de 2005 a 2021. Em outras palavras, há uma tendência positiva no Ideb durante esse período, sugerindo que, em média, o desempenho dos estudantes no Brasil tem melhorado ao longo dos anos.

Considerações Finais

Com base nas análises realizadas, que englobaram estatísticas descritivas e testes estatísticos, foi possível identificar padrões significativos e disparidades no desempenho educacional brasileiro, conforme medido pelo Índice de Desenvolvimento da Educação Básica (Ideb) e outras métricas relacionadas.

As estatísticas descritivas demonstraram que, embora a taxa média de aprovação seja elevada, há uma variação considerável no desempenho acadêmico, especialmente entre escolas públicas e privadas. Essa discrepância foi confirmada por meio de testes de hipótese, como o teste t de Student, que evidenciou diferenças estatisticamente significativas entre as redes de ensino no Ideb e no desempenho em Matemática e Língua Portuguesa.

Além disso, a análise de proporções revelou que as escolas privadas apresentam um maior percentual de estudantes com desempenho acima de 250 pontos no SAEB, destacando desigualdades importantes no acesso à qualidade educacional. A regressão linear indicou uma tendência de crescimento no desempenho médio do Ideb ao longo dos anos, sugerindo progressos graduais no sistema educacional brasileiro.

Contudo, os resultados apontam para a necessidade de intervenções direcionadas que reduzam as disparidades entre as redes pública e privada, especialmente em áreas críticas como Matemática e Língua Portuguesa. A identificação e o suporte às escolas com menor desempenho são fundamentais para promover maior equidade no sistema educacional, garantindo que todas as instituições tenham recursos adequados para atingir as metas propostas.

Este estudo, ao analisar criticamente as tendências e desigualdades no Ideb, fornece subsídios valiosos para o desenvolvimento de políticas públicas mais eficazes, com foco na melhoria da qualidade e equidade da educação básica no Brasil.

Referências Bibliográficas

Livro referência da disciplina. Barbetta, Pedro Alberto. Estatística para cursos de engenharia e informática / Pedro Alberto Barbetta, Marcelo Menezes Reis, Antonio Cezar Bornia.– 2. ed.– São Paulo: Atlas, 2009.

Base de dados. Índice de Desenvolvimento da Educação Básica (Ideb)- Censo Escolar 2005- 2023- INEP. Base dos Dados. Disponível em: [Índice de Desenvolvimento da Educação Básica \(Ideb\) – Base dos Dados](#). Acesso em: set. 2024.

Página do Ideb. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). IDEB. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/ideb>. Acesso em: set. 2024.

Página de referência para o Teste de Levene. Disponível em: https://pt.wikipedia.org/wiki/Teste_de_Levene
Acesso em: nov. 2024.

Página de referência para o Teste de Shapiro Wilk. Disponível em: https://pt.wikipedia.org/wiki/Teste_de_Shapiro%E2%80%93Wilk
Acesso em: nov. 2024.

Página de referência para o Teste t de Student. Disponível em: https://pt.wikipedia.org/wiki/Teste_t_de_Student
Acesso em: nov. 2024.

Documento Escalas de Proficiência do SAEB. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/escalas_de_proficiencia_do_saeb.pdf
Acesso em: nov. 2024.