

Relatório de Análise de Dados: Desigualdades no Desempenho das Redes de Ensino na Educação

João Pedro Soares Cavalcanti

15 de agosto de 2023

1 Introdução

Este relatório tem como objetivo apresentar a análise de dados feita para o processo seletivo da equipe UFRJ Analytica. Foi escolhida a base de dados *Índice de Desenvolvimento da Educação Básica (Ideb)* e a partir desta buscou-se responder a seguinte pergunta: Existe desigualdade de desempenho entre as redes de ensino na educação fundamental e do ensino médio? para melhor responder a pergunta foram estabelecidas 4 subperguntas mais específicas.

Para isso foram analisados diferentes aspectos relacionados ao desempenho das redes de ensino, como médias de notas, taxas de aprovação e reprovação, desempenho acadêmico e o IDEB.

A análise foi realizada utilizando a linguagem Python e suas bibliotecas de manipulação e visualização de dados, pandas, matplotlib e seaborn.

2 Metodologia

Para responder à pergunta de pesquisa, foram selecionadas uma das bases de dados pertinentes permitidas para a atividade: o *Índice de Desenvolvimento da Educação Básica (Ideb)* e nesta foi utilizada somente a tabela "Brasil"

Primeiramente, foram filtrados os dados da base do Ideb para obter apenas informações relacionadas à educação fundamental e do ensino médio. Em

seguida, os dados foram divididos em dois grupos: escolas municipais e estaduais, e escolas privadas e públicas.

A análise foi realizada utilizando o teste t de Student para comparar as médias de diferentes variáveis entre os grupos. As variáveis analisadas foram: nota média do Ideb, notas médias nas disciplinas de matemática e língua portuguesa, taxa de aprovação e indicador de rendimento.

Para cada comparação, foi calculado o valor de `t_stat` e `p_value`. O `p_value` é utilizado para determinar se a diferença entre as médias é estatisticamente significativa. Se o `p_value` for menor que 0,05, é considerado que há diferença estatisticamente significativa entre os grupos. Caso contrário, não há diferença estatisticamente significativa.

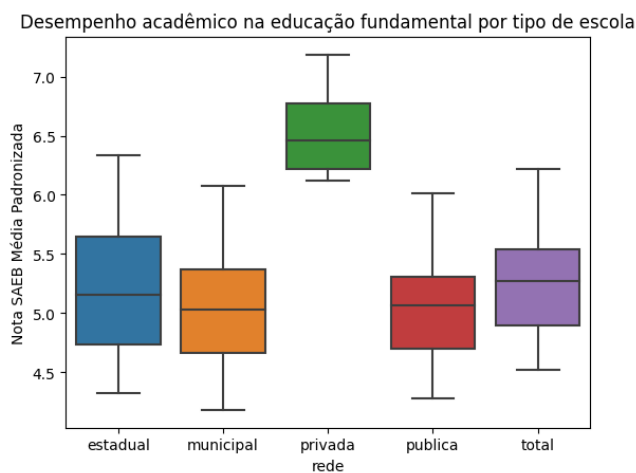
Por fim, foram construídos gráficos para visualizar as diferenças encontradas. Todos os procedimentos realizados foram implementados utilizando a linguagem Python e suas bibliotecas `pandas`, `matplotlib` e `seaborn`.

3 Análise dos Dados

Nesta seção, serão apresentados os resultados obtidos a partir da análise dos dados. Para cada subpergunta, será descrita as informações encontradas nos conjuntos de dados selecionados e discutida as principais descobertas.

Sub-pergunta 1: As escolas municipais apresentam um desempenho acadêmico inferior em relação às escolas estaduais na educação fundamental?

Para esta pergunta, foi avaliado o desempenho acadêmico das escolas municipais em relação às escolas estaduais na educação fundamental. Foi utilizado o teste t de Student para comparar as médias das notas SAEB padronizadas entre essas duas redes de ensino. (Valores positivos indicam uma média maior para escolas municipais e valores negativos indicam uma média maior para escolas estaduais)



	valor
t_stat	-0.867
p_value	0.393

A resposta para a questão é sim, porém não é estatisticamente significativa segundo o critério estabelecido.

Sub-pergunta 2: As escolas privadas têm um desempenho acadêmico superior em comparação com as escolas públicas na educação fundamental?

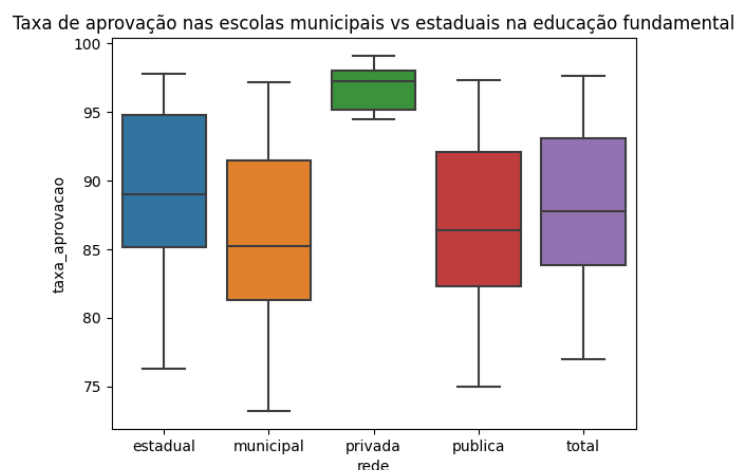
Aqui foi comparado o desempenho acadêmico das escolas privadas com as escolas públicas na educação fundamental. Novamente foi utilizado o teste t de Student. (Valores positivos indicam uma média maior para escolas privadas e valores negativos indicam uma média maior para escolas públicas)

	valor
t_stat	9.073
p_value	2.679e-12

A resposta para a questão é sim, com um resultado estatisticamente significativo segundo o nosso critério.

Subpergunta 3: A taxa de aprovação nas escolas municipais é menor do que nas escolas estaduais no ensino médio?

Foi investigado se a taxa de aprovação nas escolas municipais é menor do que nas escolas estaduais no ensino médio. Utilizamos o teste t de Student. (Valores positivos indicam uma média maior para escolas municipais e valores negativos indicam uma média maior para escolas estaduais)



	valor
t_stat	-1.311
p_value	0.1987

A resposta para a questão é sim, porém não é estatisticamente significativa segundo o critério estabelecido.

Subpergunta 4: A taxa de aprovação nas escolas privadas é maior do que nas escolas publicas no ensino médio?

Foi investigado se a taxa de aprovação nas escolas privadas é maior do que nas escolas publicas no ensino médio. Utilizamos o teste t de Student. (Valores

	valor
t_stat	3.215
p_value	0.002

positivos indicam uma média maior para escolas privadas e valores negativos indicam uma média maior para escolas públicas)

A resposta para a questão é sim, com um resultado estatisticamente significativo segundo o nosso critério.

4 Conclusões

É possível concluir que existem grandes desigualdades no ensino fundamental e médio, entretanto apenas em alguns critérios específicos. As escolas públicas municipais e estaduais não possuem grandes disparidades entre si com relação ao desempenho, sendo esta estatisticamente não significativa. Já para a comparação entre as escolas públicas e privadas, pode-se observar um grande diferencial de desempenho, encontrado como estatisticamente significativo.