

Análise da Desigualdade na Educação Brasileira  
no Período da Pandemia  
A2 ANALYTICA - PS 2023.2

Wemerson Silva Caxias da Costa

wemersonscc@dcc.ufrj.br

Agosto, 2023



UFRJ

## Contents

<b>1</b>	<b>Introdução</b>	<b>3</b>
<b>2</b>	<b>Dados Utilizados</b>	<b>4</b>
2.1	Indicador Nível Socioeconômico (INSE) . . . . .	4
2.2	Indicadores Educacionais . . . . .	4
2.3	Índice de Desenvolvimento da Educação Básica (IDEB) . . . . .	5
2.4	Censo Escolar . . . . .	5
<b>3</b>	<b>Análise</b>	<b>7</b>
3.1	Hipóteses e Impressões Iniciais . . . . .	7
3.2	Em busca de contradições . . . . .	9
<b>4</b>	<b>Conclusões</b>	<b>12</b>

# 1 Introdução

Uma das coisas mais importantes para a prosperidade de um país é uma educação de qualidade, desde o nível de base para as crianças, até o nível superior para os adultos. Uma nação que investe em um sistema educacional sólido colhe benefícios que se estendem por gerações. Além de formar profissionais competentes, um ensino de qualidade também promove a cidadania ativa, o pensamento crítico e a capacidade de resolver problemas, preparando pessoas para contribuir de maneira significativa para o crescimento econômico, social e cultural de uma nação.

Essa questão ganha uma dimensão ainda mais urgente ao se abordar a desigualdade educacional no contexto brasileiro. A disparidade no acesso à educação de qualidade é um desafio persistente no Brasil, refletindo em consequências significativas para o desenvolvimento nacional. Sem um acesso equitativo à educação, resultado da falta de investimento adequado e de políticas públicas eficazes, temos em disparidades regionais, socioeconômicas e étnicas, perpetuando um ciclo de desigualdade que limita o potencial de progresso do país, além de perpetuar a marginalização de grupos historicamente excluídos, ampliando as divisões sociais e limitando a mobilidade social.

Todos os problemas acima foram intensificados significativamente com a questão da pandemia da COVID-19, pois com o fechamento das escolas e a transição abrupta para o ensino remoto, as disparidades socioeconômicas se aprofundaram. Muitos estudantes de famílias mais vulneráveis não tiveram acesso aos recursos necessários para o aprendizado online, como dispositivos eletrônicos e uma conexão estável à internet. Isso resultou em uma lacuna ainda maior no acesso à educação de qualidade entre aqueles que já estavam em desvantagem.

Neste relatório será feita uma análise sobre dados obtidos pelo governo relacionados a situação econômica dos estudantes e desempenhos escolares, mais especificamente no período da pandemia, isto é, do começo de 2020 até o final de 2021, onde buscaremos ter uma perspectiva mais clara afim de compreender o impacto dessa crise sanitária mundial na já complexa e desigual situação educacional no Brasil.

## 2 Dados Utilizados

Para este trabalho, temos bases já tratadas pela [basedosdados.org](https://basedosdados.org):

1. [Indicador Nível Socioeconômico \(INSE\) - Cobertura temporal 2013-2021](#)
2. [Indicadores Educacionais - Cobertura temporal 2006-2022](#)
3. [Índice de Desenvolvimento da Educação Básica \(IDEB\) - Cobertura temporal 2013-2021](#)
4. [Censo Escolar - Cobertura temporal 2007-2022](#)

### 2.1 Indicador Nível Socioeconômico (INSE)

A primeira dessa base de dados, que é sobre o INSE, é feita pelo INEP, e com base nos resultados do questionário do aluno do Saeb (Sistema de Avaliação da Educação Básica), tem como objetivo contextualizar resultados obtidos em avaliações e exames aplicados no âmbito da educação básica. Ela conta com as tabelas:

- brasil - indica o INSE ao nível do Brasil
- escola - indica o INSE ao nível de uma escola
- municipio - indica o INSE ao nível do município
- uf - indica o INSE ao nível do estado

Em cada uma das tabelas, o INSE é medido a partir das informações de posse de bens domésticos, renda e contratação de serviços pela família dos alunos e pelo nível de escolaridade de seus pais nas respostas de questionários socioeconômicos de exames com SAEB e ENEM.

### 2.2 Indicadores Educacionais

Os indicadores educacionais atribuem valor estatístico à qualidade do ensino, atendo-se não somente ao desempenho dos alunos mas também ao contexto econômico e social em que as escolas estão inseridas. Os dados foram obtidos pelo Inep. Ela conta com as tabelas:

- brasil - indicadores educacionais do Inep a nível Brasil
- brasil\_remuneracao\_docentes - remuneração média dos docentes em exercício na educação básica por escolaridade
- escola - indicadores educacionais do Inep a nível escola
- escola\_nivel\_socioeconomico - indicador de nível socioeconômico por escola.
- fluxo\_educacao\_superior - indicadores da trajetória acadêmica dos alunos em cursos de graduação

- `municipio` - indicadores educacionais do Inep a nível município
- `municipio_taxas_transicao` - têm por objetivo informar sobre a trajetória do estudante na educação básica
- `regiao` - indicadores educacionais do Inep a nível de região
- `uf` - indicadores educacionais do Inep a nível do estado
- `uf_remuneracao_docentes` - remuneração média dos docentes em exercício na educação básica por escolaridade

### 2.3 Índice de Desenvolvimento da Educação Básica (IDEB)

O IDEB foi criado em 2007 e reúne os resultados de dois conceitos igualmente importantes para a qualidade da educação: o fluxo escolar e as médias de desempenho nas avaliações. Ele é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho nas avaliações do Inep, o Sistema de Avaliação da Educação Básica (Saeb) – para as unidades da federação e para o país, e a Prova Brasil – para os municípios. O índice varia de zero a 10. Essa base de dados conta com as tabelas:

- `brasil` - indica o IDEB ao nível do Brasil
- `escola` - indica o IDEB ao nível de uma escola
- `municipio` - indica o IDEB ao nível do município
- `regiao` - indica o IDEB ao nível de região
- `uf` - indica o IDEB ao nível do estado

### 2.4 Censo Escolar

Censo Escolar é o principal instrumento de coleta de informações da educação básica e a mais importante pesquisa estatística educacional brasileira. É coordenado pelo Inep e realizado em regime de colaboração entre as secretarias estaduais e municipais de educação e com a participação de todas as escolas públicas e privadas do país. Essa base de dados conta com as tabelas.:

- `docente` - contém dados de cada professor de cada ano das escolas nacionais. A tabela identifica desde características pessoais dos profissionais, as disciplinas que eles ensinam, as suas formações e a qual escola e qual turma pertencem.
- `escola` - contém dados do censo escolar ao nível de cada escola. Os dados são organizados em mais de 300 variáveis que traçam um perfil completo da escola, caracterizando desde a infraestrutura física disponível, a equipe pedagógica, a infraestrutura pedagógica, marcas do período letivo como seu começo e fim, além de várias informações sobre os estudantes da escola.

- matrícula - contém dados de matrícula, como data de nascimento, sexo, raça, nacionalidade, e também outros aspectos mais específicos, como se o aluno há deficiências ou se precisa de um intérprete de libras para realizar avaliações do INEP
- turma - contém dados que caracterizam as turmas de cada escola brasileira. Entre os dados podemos encontrar as disciplinas feitas pela turma, o horário de começo das aulas, o número de alunos da turma e o nível de ensino no qual se encontra a turma. Há também vários outros aspectos mais específicos, como se a turma passa por atividades complementares, em quais dias da semana ela é ativa e se a turma conta com medidas de inclusão como aulas em libras e em braille

Como todas essas tabelas são muito grandes, especialmente essa última, escolheremos pequenos recortes para progredir no nosso estudo, tentando responder as perguntas que nós mesmos faremos.

## 3 Análise

### 3.1 Hipóteses e Impressões Iniciais

Todas as tabelas das bases de dados tem um número muito grande de colunas e linhas, então a ideia de restringirmos nossa busca é crucial para tentarmos realizar esse estudo, pois, olhar para as tabelas sem propósito definido não nos ajuda em nada, visto que uma tabela com centenas de milhões de linhas e centenas de colunas, para nós seres humanos, é ilegível. As perguntas que surgiram à minha cabeça imediatamente foram:

1. A pandemia fez o desempenho escolar piorar?
  - 1.1 A diferença foi muito grande entre as modalidades de ensino pública e privada?
    - 1.1.1 E entre as áreas rurais e urbanas?
  - 1.2 Quão impactante isso foi para os alunos de níveis (INSE) 1,2 e 3, que são os níveis onde a maior parte dos alunos não tem computador?
    - 1.2.1 E para o restante dos níveis (4,5,6,7,8)?
  - 1.3 O impacto da pandemia foi mais sentida em que nível de escolaridade (básica, média e superior)?
    - 1.3.1 E em quais tipos de rede (municipal, estadual e federal)?

Agora que já temos um ponto de partida, precisamos pensar: quais dessas perguntas conseguimos responder com os dados que temos? Começemos para a questão mais importante do tema: a questão 1.

Um bom começo é olharmos para a coluna que contém a taxa de aprovação dos estudantes de nível médio e fundamental, na base dos Indicadores Educacionais a nível Brasil, pois, soa intuitivo a ideia que a quantidade de aprovados na escola é diretamente proporcional a qualidade do estudo. Realizando nessa tabela a consulta abaixo

```

6 SELECT ano,
7 avg(taxa_aprovacao_ef) as media_aprovacao_ef,
8 avg(taxa_aprovacao_em) as media_aprovacao_em
9 FROM `basedosdados.br_inep_indicadores_educacionais.brasil` as bd
10 group by ano
11 having ano >= 2015 and ano <= 2021
12 order by ano desc
13

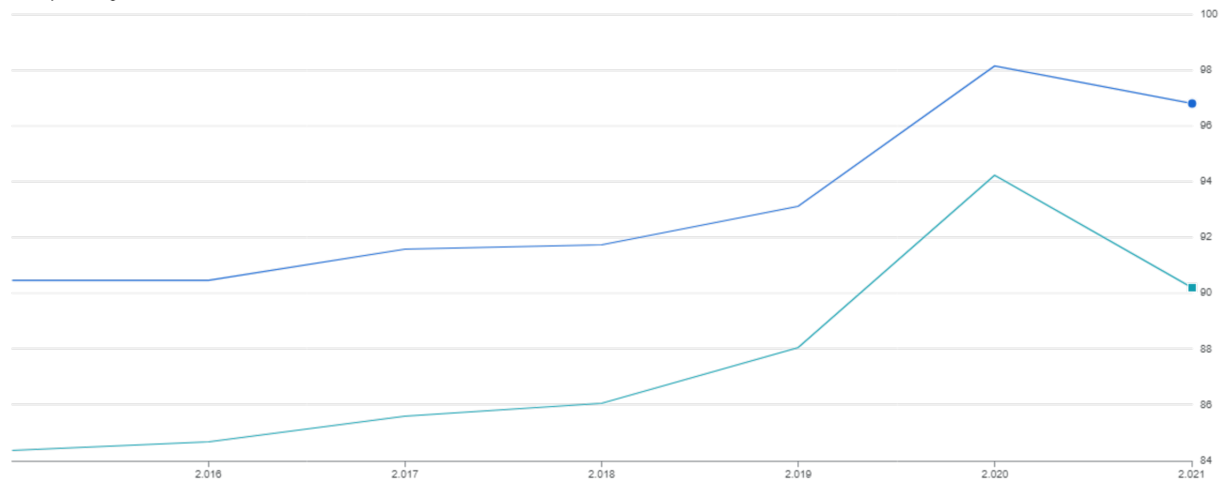
```

## Resultados da consulta

INFORMAÇÕES DO JOB	RESULTADOS	JSON	DETALHES DA EXECUÇÃO	GRÁFICO	PRÉ-VISUALIZAÇÃO	GRÁFICO D
Linha	ano	media_aprovacao_ef	media_aprovacao_em			
1	2021	96.77777777777771	90.205555555555534			
2	2020	98.155555555555566	94.22222222222229			
3	2019	93.100000000000009	88.049999999999983			
4	2018	91.73888888888889	86.055555555555557			
5	2017	91.572222222222209	85.611111111111128			
6	2016	90.483333333333332	84.661111111111111			
7	2015	90.466666666666669	84.38333333333334			

Obtemos um resultado que é mais fácil de analisar olhando para o gráfico:

Exibição em gráfico



Nos deparamos então com um grande problema, pois, como podemos perceber pela linhas, em 2020, o começo da pandemia, foi onde teve o pico (desde 2015, isto é, os 5 últimos anos antes da pandemia) de alunos aprovados, tanto no ensino médio quanto no ensino fundamental, a nível Brasil. O que contraria todo nosso pressuposto inicial que a pandemia prejudicou ainda mais a educação no nosso país (ou a ideia que uma educação de qualidade significaria um número maior de aprovados necessariamente). O que isso quer dizer?

Isso se dá por um motivo: como medida para controlar a questão da educação



durante a crise do Covid-19, o Conselho Nacional de Educação (CNE) definiu diretrizes que mudaram os critérios de aprovação da educação básica durante a pandemia, havendo a adoção do “continuum curricular”, o que, na prática, resultou na não reprovação dos estudantes, como forma de evitar punição para aqueles que não tiveram condições de acompanhar as aulas remotas.

Portanto, apenas os resultados dessa edição do Censo, e das outras bases de dados, não bastam para termos uma boa visão do impacto da pandemia na educação, já que não podem ser comparados aos resultados dos anos anteriores, devido a essa e possivelmente outras distorções, que acabam influenciando nos indicadores que utilizamos para atribuir valor estatístico à qualidade do ensino. Para termos de forma mais assertiva indicadores do impacto de dois anos de pandemia na educação brasileira, precisaremos esperar mais alguns anos.

Porém, essa conclusão foi feita a partir de um conhecimento externo (público). Haveria alguma forma, de apenas olhando apenas para os dados que já temos, conseguirmos chegar à conclusão de que talvez nossos dados não sejam “confiáveis”? Esse será o objetivo da próxima parte desse documento.

### 3.2 Em busca de contradições

Como dito no fim do parágrafo anterior, essa subseção tem o objetivo de encontrar alguns indicadores que coloquem em dúvida a validade dos dados que temos para as hipóteses que propusemos.

Nesse caso já sabemos por conhecimento comum que a nossa hipótese é verdadeira, em contradição com os nossos dados, então fica mais fácil, porém, esse tipo de verificação é importantíssima mesmo num caso onde nossa hipótese estivesse sendo provada correta, pois um erro comum é a tendência de darmos mais peso a informações que confirmam nossas crenças e acabarmos ignorando informações que contradizem essas ideias.

Ao verificar a validade dos dados, evitamos cair nesse viés de confirmação e garantimos que nossa análise seja objetiva e imparcial.

Um indicador que despertaria a dúvida do analista de dados que olhasse para essas bases às cegas seria o seguinte: a relação entre o IDEB e as médias de desempenho nas avaliações do Inep, (o Saeb); realizando a consulta

```
5 SELECT ano, rede, ensino, anos_escolares, nota_saeb_matematica, nota_saeb_lingua_portuguesa, ideb
6 FROM 'basedosdados.br_inep_ideb_brasil'
7 where ano >= 2019 and ano <= 2021 and rede = 'municipal' and anos_escolares = 'finais (6-9)'
```

Obtemos como resultado a seguinte tabela:

Linha	ano	rede	ensino	anos_escolares	nota_saeb_matematica	nota_saeb_lingua_pg	ideb
1	2019	municipal	fundamental	finais (6-9)	255.57	254.1	4.5
2	2021	municipal	fundamental	finais (6-9)	249.15	252.38	4.8

Como podemos ver pelo esquema acima, que nesse caso é o bastante para interpretarmos esses dados, sem a necessidade do auxílio de gráficos, por exemplo, pois tem poucas linhas e colunas, no ano de 2019, mesmo a nota do Saeb dos alunos do ensino fundamental da rede municipal sendo maior tanto

em português quanto em matemática, o Ideb foi maior em 2021 do que em 2019.

Isso se dá pois o Ideb combina taxas de aprovação com a prova do Saeb, o que faz com que o índice de desenvolvimento da educação básica suba mesmo com médias menores devido a porcentagem de aprovados que beirava os 100% durante a pandemia (por motivos já citados).

Outra informação que também poderia nos dar um norte é a quantidade de pessoas que participam dessas pesquisas. Porém, nas nossas bases de dados, não há muitas colunas que quantifiquem o número de pessoas que realizaram, por exemplo, as provas do Saeb, seja por escola ou por município, apenas constam as taxas de aprovação, o que não nos ajuda muito no sentido de contagem. Além disso, conforme os anos foram passando e chegamos na cobertura temporal da pandemia, as tabelas vão ficando mais esparsas, provavelmente pela dificuldade de se entrevistar/realizar provas durante a crise. Porém, numa tentativa de estimarmos mesmo assim, realizamos a consulta:

```
3 SELECT ano, sigla_uf, sum(quantidade_alunos_inse) as total_alunos
4 FROM basedosdados.br_inep_indicador_nivel_socioeconomico.uf`
5 group by ano, sigla_uf |
6 having ano = 2021 or ano = 2019
7
```

Que após tratarmos os dados separando-os por região, conseguimos gerar o gráfico:



Que como podemos perceber, demonstra uma queda em todas as regiões no número de participantes, sendo especialmente significativa no sudeste, o que indicaria naturalmente que alguma coisa está errada; se a pandemia não estivesse afetando a educação, o número de estudantes realizando as provas não cairia tão abruptamente num período de 2 anos.

## 4 Conclusões

Tendo em vista as análises e os dados apresentados, chegamos a conclusão que está muito cedo para medirmos assertivamente o impacto da pandemia na educação brasileira. Porém, mesmo assim, devemos nos lembrar dos problemas educacionais que já existiam antes, sendo responsabilidade de toda a sociedade cobrar o governo para promover políticas inclusivas, investir em infraestrutura digital e criar programas que reduzam as disparidades educacionais. Ações direcionadas, como a distribuição de dispositivos e recursos de aprendizado para famílias de baixa renda, a capacitação de educadores para o ensino remoto e a implementação de estratégias de recuperação do aprendizado perdido durante a pandemia, são passos cruciais para mitigar o déficit que embora não seja possível medir ainda, sabemos que existe. Ao reconhecer que um país só progredir quando há acesso igualitário à educação, o Brasil tem a oportunidade de lidar com o desafio da desigualdade educacional de forma sistemática, transparente e sustentável.

[Link para o Github](#)