



A INFLUÊNCIA DO NÍVEL SOCIOECONÔMICO (INSE) NO DESEMPENHO ESCOLAR (IDEB): UM ESTUDO ESTATÍSTICO DE ESCOLAS BRASILEIRAS EM 2021

Jamile Borges Boffo¹, Robson Erivelton Pereira², Marcos Aurélio Brambilla³, Thaise Moser Teixeira⁴

1 Acadêmica do Curso de Gestão do Conhecimento nas Organizações, Campus Maringá-PR, Universidade Cesumar - UNICESUMAR. jamileboffo@gmail.com;

2 Acadêmico do Curso de Gestão do Conhecimento nas Organizações, Campus Maringá-PR, Universidade Cesumar - UNICESUMAR. robson.pereira@unicesumar.edu.br;

3 Doutor em Economia e Docente do Curso de Gestão do Conhecimento nas Organizações, Campus Maringá-PR, Universidade Cesumar - UNICESUMAR. marcos.brambilla@unicesumar.edu.br;

4 Doutora em Ciência e Tecnologia Ambiental e Docente do Curso de Gestão do Conhecimento nas Organizações, Campus Maringá-PR, Universidade Cesumar - UNICESUMAR. thaise.teixeira@unicesumar.edu.br.

RESUMO

Este estudo investigou a influência do nível socioeconômico (INSE) no desempenho escolar, mensurado pelo Índice de Desenvolvimento da Educação Básica (IDEB), em escolas de ensino básico no Brasil. Utilizando dados de 2021, o objetivo foi verificar e quantificar a correlação e o impacto do INSE nos resultados educacionais. A análise estatística, realizada com dados de 7.194 escolas, confirmou uma correlação positiva e moderada ($r \approx 0,502$) entre as variáveis, indicando que, à medida que o INSE aumenta, o IDEB tende a ser maior. Por meio da regressão linear, constatou-se que o nível socioeconômico é um preditor estatisticamente significativo, explicando aproximadamente 25,2% da variação no desempenho escolar. O modelo revelou que, para cada aumento de um ponto no INSE, espera-se um acréscimo de 0,650 pontos no IDEB. A análise por Unidade Federativa corroborou essa tendência. Conclui-se que, embora outros fatores sejam relevantes, o desempenho escolar está intrinsecamente associado ao capital social e econômico da comunidade.

PALAVRAS-CHAVE: Desempenho Escolar, Nível Socioeconômico, IDEB.

1 INTRODUÇÃO

A busca por uma educação de qualidade no Brasil representa uma inspeção contínua e uma preocupação crescente para a gestão pública, tendo como objetivo garantir que cada estudante tenha acesso a métodos de ensino mais equitativos e com excelência. No entanto, para verificar a real eficácia das políticas públicas, garantir uma gestão do conhecimento funcional para amparar as tomadas de decisões, identificar as lacunas existentes, a aplicação de ferramentas de avaliação robustas torna-se necessária.

Nesse cenário, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) desempenha um papel central ao desenvolver os principais instrumentos de avaliação educacional governamental. Exemplos notáveis incluem o Sistema de Avaliação da Educação Básica (SAEB) e a Prova Brasil, os quais, por sua vez, fornecem dados para o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB). Essas métricas são, sem dúvida, significativas para a compreensão das variáveis que impulsionam os resultados positivos nas escolas (Capucho, 2024).

Contudo, observa-se uma lacuna na consideração de indicadores que podem impactar consideravelmente o desempenho acadêmico. Como apontado novamente pelo autor, o Nível Socioeconômico (NSE) dos estudantes não é plenamente integrado a essas avaliações. O Indicador de Nível Socioeconômico (INSE), por exemplo, ainda não é formalmente reconhecido como uma ferramenta principal de avaliação educacional, o que, lamentavelmente, pode levar a uma desconsideração da realidade social, cultural e econômica do aluno (Matias, 2018).



A análise de fatores escolares que impactam o desempenho dos alunos em avaliações de larga escala tem sido um foco importante da pesquisa educacional no Brasil. Estudos como o de Garcia, Rios Neto e Ribeiro (2021), por exemplo, utilizaram dados do Censo Escolar e do Enem para identificar os efeitos da infraestrutura e da qualificação docente na qualidade do ensino, destacando a relevância de se investigar essas variáveis.

Diante dessa perspectiva, o presente estudo tem como objetivo verificar e quantificar a correlação e o impacto do INSE nos resultados educacionais. Para tanto, foram utilizados os dados disponibilizados pelo IDEB e INSE, referentes ao ano de 2021, abrangendo o panorama das escolas de ensino básico em todo o Brasil.

2 MATERIAIS E MÉTODOS

Este estudo caracteriza-se como uma pesquisa quantitativa, de caráter correlacional e explicativo, que tem por objetivo principal verificar e quantificar a correlação e o impacto do INSE nos resultados educacionais. Os dados utilizados são provenientes de bases de dados públicas e oficiais da educação brasileira, sendo esses, do Índice de Nível Socioeconômico (INSE) e Índice de Desenvolvimento da Educação Básica (IDEB) realizados no ano de 2021, com abrangência de todas as escolas brasileiras.

Entre os bancos de dados utilizados, foi realizada uma filtragem entre as colunas de maior relevância para a construção das relações, então as variáveis utilizadas são: UF da escola, identificação da escola, média da INSE da escola, média do IDEB da escola. Também é possível entender melhor as variáveis e tipos de dados que cada uma delas comporta através da tabela 1:

Tabela 1: Descrição das variáveis e os tipos de dados suportados

VARIÁVEL	DESCRIÇÃO	TIPO DE DADO
UF	Unidade da Federação onde a escola está localizada	char
ID_ESCOLA	Código identificador da escola (para fins de diferenciação)	num
INSE	Valor da média do INSE da escola	num
IDEB	Valor da média do IDEB da escola	num

Devido a critérios rigorosos do INEP para que uma escola tenha seu IDEB divulgado, algumas escolas que possuíam valor nulo foram excluídas para a validação final.

Para atingir o objetivo deste estudo, a análise de dados foi construída em três etapas complementares. Inicialmente, foi realizada uma estatística descritiva para apresentar um panorama geral dos dados, destacando suas tendências e variabilidades. Em seguida, foi aplicada uma análise de correlação, que verificou a existência de uma associação entre as variáveis INSE e IDEB. E por fim, foi elaborada a análise de regressão, para modelar o impacto do nível socioeconômico no desempenho escolar, permitindo quantificar o grau dessa influência.

A estatística descritiva foi utilizada para agrupar e apresentar as principais características da amostra de escolas brasileiras. Com esse procedimento, foram calculadas as medidas de tendência central (média e mediana) e medidas de dispersão (desvio padrão) para as variáveis INSE e IDEB. Além disso, foram elaborados gráficos (como gráficos de barras) para visualizar a distribuição dos dados de cada variável. A



análise descritiva também inclui o cálculo das médias do IDEB e INSE desagregadas por Unidade da Federação (UF), permitindo identificar padrões e diferenças entre os estados. O objetivo é oferecer uma visão inicial e abrangente dos dados antes da análise das relações.

A análise de correlação teve como função investigar a existência, a força e a direção da associação linear entre o nível socioeconômico e o desempenho escolar. Para isso, foi calculado o coeficiente de correlação de Pearson(r) entre a variável INSE e a variável IDEB. A interpretação desse coeficiente r foi baseada na magnitude (próxima de 1 para correlação forte, próxima de 0 para correlação fraca) e seu sinal (positivo para relação direta, negativo para relação inversa). Adicionalmente, foi gerado um gráfico de dispersão para visualizar a nuvem de pontos e a tendência da relação entre as duas variáveis.

A análise de regressão foi utilizada a fim de modelar a influência do nível socioeconômico sobre o desempenho escolar. Esse método permite estabelecer uma equação linear na qual a IDEB (variável dependente) será explicada pela INSE (variável independente). Foram avaliados os coeficientes de regressão (incluindo coeficiente angular, que indicará a variação esperada no IDEB para cada unidade de variação no INSE), o coeficiente de determinação (R^2), que indicará a proporção da variância do IDEB explicada pelo INSE, e os valores- p associados aos coeficientes e ao modelo geral, para verificar a significância estatística da relação. O modelo resultante foi apresentado e discutido em termos de sua capacidade preditiva e relevância teórica.

3 RESULTADOS E DISCUSSÕES

A análise dos dados, como mencionada anteriormente, com o objetivo de investigar a influência do nível socioeconômico (INSE) no desempenho escolar (IDEB), foi conduzida por meio de estatística descritiva, análise de correlação e regressão linear simples. Os resultados obtidos oferecem apontamentos importantes sobre a relação entre essas duas variáveis no contexto da educação básica.

Primeiramente, a análise descritiva (Tabela 2) das 7194 observações revelou que a média do INSE foi de 4,88 (DP = 0,54), enquanto a do IDEB foi de 4,39 (DP = 0,70). A maior dispersão nos valores do IDEB em comparação ao INSE sugere que, embora o contexto socioeconômico seja um fator de base, os resultados educacionais entre as escolas são mais heterogêneos.

Tabela 2: Estatística Descritiva para INSE e IDEB

	INSE	IDEB
Média	4,888719766	4,395621351
Erro padrão	0,006378032	0,008255901
Mediana	4,97	4,4
Modo	5,29	4,3
Desvio padrão	0,540968452	0,700244448
Variância da amostra	0,292646866	0,490342287
Curtose	-0,623286143	0,768236306



Assimetria	-0,291986652	0,185446875
Intervalo	3,43	6,2
Mínimo	2,97	1,4
Máximo	6,4	7,6
Soma	35169,45	31622,1
Contagem	7194	7194
Nível de confiança(95,0%)	0,012502818	0,016183991

A investigação da associação entre as variáveis, por meio do coeficiente de Correlação de Pearson, revelou um valor de $r \approx 0,502$ (Tabela 3), confirmando uma relação positiva e de força moderada. Este valor indica uma tendência clara: escolas com maior média de INSE tendem a apresentar um IDEB correspondentemente maior. Tal achado corrobora a tese de que o desempenho escolar está intrinsecamente ligado ao capital social e econômico da comunidade, como aponta Matias (2018) ao afirmar que um nível socioeconômico mais elevado garante acesso a recursos e estímulos que impactam positivamente a escolarização.

Tabela 3: Coeficiente de Correlação de Pearson entre INSE e IDEB

	INSE	IDEB
INSE	1	
IDEB	0,502721	1

A visualização desta tendência no gráfico de dispersão (Figura 1) reforça essa relação linear ascendente.

Dispersão IDEB x INSE

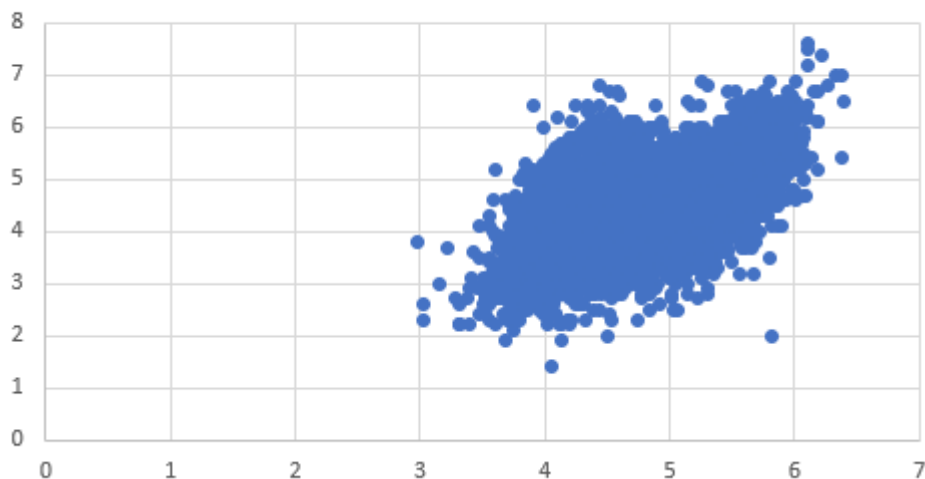


Figura 1: Gráfico de Dispersão do IDEB em função do INSE



Para quantificar a magnitude dessa influência, o modelo de regressão linear simples (Tabela 4) demonstrou que o INSE é um preditor estatisticamente significativo do IDEB. O coeficiente de determinação (R^2) de 0,252 indica que aproximadamente 25,2% da variação das notas do IDEB pode ser explicada pelas variações no INSE. Este é um poder explicativo considerável para uma única variável, reforçando o argumento de Capucho (2024), que também identificou forte associação entre o INSE e o desempenho escolar. No entanto, o modelo também evidencia que 74,8% da variabilidade se deve a outros fatores, o que não diminui a importância do INSE, mas o situa como uma peça-chave dentro de um cenário complexo que inclui gestão, corpo docente e infraestrutura. Adicionalmente, o coeficiente de regressão de 0,650 demonstra que, para cada aumento de um ponto na escala do INSE, espera-se um acréscimo médio de 0,650 pontos no IDEB, quantificando o impacto direto do contexto social no resultado educacional.

Tabela 4: Resultados da Análise de Regressão Linear Simples (IDEB x INSE)

Estatística de regressão	
R múltiplo	0,5027213
R-Quadrado	0,252728706
R-quadrado ajustado	0,252624803
Erro padrão	0,605367379
Observações	7194

ANOVA					
	gl	SQ	MQ	F	F de significação
Regressão	1	891,3822515	891,3822515	2432,349359	0
Resíduo	7192	2635,649821	0,366469664		
Total	7193	3527,032073			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	1,214354023	0,064897746	18,71180589	2,52071E-76	1,087135368	1,341572677	1,087135368	1,341572677
INSE	0,650736283	0,013194473	49,31885399	0	0,624871238	0,676601327	0,624871238	0,676601327

Por fim, a análise agregada por Unidade Federativa (Tabela 5) corrobora os achados em nível macro. Observa-se que estados com maiores médias de INSE, como Santa



Catarina (5,49) e Rio Grande do Sul (5,48), figuram entre aqueles com maiores médias de IDEB. Em contrapartida, estados com os menores indicadores de INSE, como Maranhão (4,18) e Pará (4,40), apresentam os maiores desafios no desempenho escolar. Essa consistência reforça a conclusão de que as desigualdades socioeconômicas regionais no Brasil se espelham diretamente nos resultados da educação básica.

Tabela 5: Médias de INSE e IDEB por Unidade Federativa (UF)

UF	Média de IDEB	Média de INSE
AC	4,572727273	4,784545455
AL	3,823880597	4,275970149
AM	3,768627451	4,42620915
BA	4,030136986	4,563013699
CE	4,578693624	4,284774495
DF	4,260869565	5,399130435
ES	4,808433735	4,984457831
GO	4,510707804	5,029872958
MA	3,591943128	4,186706161
MG	4,32476573	4,90605087
MS	4,28	5,1905
MT	4,707692308	5,323076923
PA	3,491891892	4,408918919
PB	4,296511628	4,369651163
PE	4,456076135	4,462986823
PI	4,119858156	4,197801418
PR	4,585791367	5,29506295
RJ	4,08952381	4,922857143
RN	4,1	4,762666667
RO	4,255813953	4,918604651
RR	4,65	4,61
RS	4,747297297	5,480945946
SC	4,448780488	5,491219512



SE	4,0109375	4,37328125
SP	4,619892833	5,329249833
TO	4,023809524	4,657142857

Em suma, os resultados confirmam que o nível socioeconômico é um preditor relevante do desempenho escolar. A correlação moderada e o poder explicativo de 25,5% demonstram que, embora o desempenho seja um fenômeno multifatorial, as condições sociais e econômicas da comunidade escolar exercem uma influência direta e mensurável sobre o IDEB.

O fato de o INSE explicar 25,2% da variância do IDEB posiciona o nível socioeconômico como um preditor relevante, mas também aponta que quase 75% da variação depende de outros fatores. Nesse sentido, os achados de Garcia, Rios Neto e Ribeiro (2021) são esclarecedores. Em um estudo com metodologia análoga, os autores concluíram que, dentre os diversos fatores escolares analisados, a qualificação docente foi o mais impactante para o desempenho dos alunos no Enem. Isso sugere que, embora o INSE estabeleça uma condição de base, a qualidade da prática docente pode ser um dos principais fatores que atuam na variação restante do desempenho escolar.

4 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo investigar a influência do nível socioeconômico (INSE) no desempenho escolar (IDEB) em escolas brasileiras no ano de 2021. Os resultados confirmaram uma correlação positiva e moderada entre as duas variáveis, indicando que, à medida que o INSE cresce, o desempenho escolar o acompanha. A análise de regressão linear reforça essa conclusão, demonstrando que a relação é estatisticamente significativa e que o INSE atua como um preditor relevante, explicado através da variação de aproximadamente 25,2% no IDEB.

Contudo, é relevante reconhecer algumas limitações. O uso de dados agregados por escola impede inferências diretas no nível do aluno, podendo incorrer na chamada falácia ecológica. Além disso, o delineamento transversal do estudo, com dados de um único ano, oferece um panorama estático da relação, não permitindo análises de causalidade ao longo do tempo. Outra limitação a ser considerada reside na própria natureza dos dados quantitativos utilizados. Embora o IDEB seja o indicador oficial para medir a qualidade da educação básica, sua construção, que envolve taxas de fluxo escolar (aprovação), não está isenta de críticas. Conforme aponta Carvalho (2001), a produção dessas estatísticas no cotidiano escolar pode sofrer pressões institucionais para a redução dos índices de repetência, de modo que os números nem sempre refletem a qualidade real do processo de ensino-aprendizagem. Assim, os resultados aqui encontrados devem ser entendidos dentro do escopo do indicador oficial, que possui suas próprias complexidades. Futuras pesquisas poderiam explorar essa dinâmica por meio de estudos longitudinais ou incorporar análises multinível que considerem tanto o contexto do aluno quanto o da escola.

Por fim, este trabalho reforça a forte correlação entre o capital social e o desempenho escolar medido por indicadores oficiais. No entanto, é preciso cautela para que a busca por melhores índices não se sobreponha à busca por uma aprendizagem efetiva. A 'maquiagem estatística' discutida por Carvalho (2001) serve como um alerta: políticas públicas devem ir além dos indicadores e focar nas condições reais de ensino, como a infraestrutura, a valorização docente e o apoio a alunos com dificuldades, para que a melhoria do IDEB signifique, de fato, uma melhoria na educação.



REFERÊNCIAS

CARVALHO, M. P. de. Estatísticas de desempenho escolar: o lado avesso. **Educação & sociedade**, v. 22, p. 231-252, 2001.

GARCIA, R. A.; RIOS-NETO, E. L. G.; MIRANDA-RIBEIRO, A. Efeitos rendimento escolar, infraestrutura e prática docente na qualidade do ensino médio no Brasil. **Revista brasileira de Estudos de População**, v. 38, p. e0152, 2021.

MATIAS, N. C. F. Relações entre nível socioeconômico, atividades extracurriculares e alfabetização. **Psico-USF**, v. 23, n. 3, p. 567-578, 2018.

CAPUCHO, P. H. P. A influência de custos, infraestrutura escolar, corpo pedagógico e nível socioeconômico no desempenho escolar: uma análise nas escolas municipais de Arapongas/PR. **Dissertação de Mestrado em Administração - Universidade Estadual de Londrina**, Londrina, 2024.