ISSN: 2358-8829

APRENDIZAGEM BASEADA EM DADOS NA EDUCAÇÃO BÁSICA: ANÁLISE DE INDICADORES EDUCACIONAIS REALIZADA POR ESTUDANTES DO ENSINO MÉDIO

Brena Maia Santos[1]
Maria Gabriela Pires Belfort Freire[2]

RESUMO

O presente trabalho apresenta um relato de experiência pedagógica desenvolvido com estudantes do ensino médio, no qual foi proposta a análise de indicadores educacionais a partir de bases de dados públicas e confiáveis. A atividade foi realizada por um grupo de seis alunos ao longo do ano letivo de 2025 e teve como objetivo promover o desenvolvimento do pensamento crítico e do letramento em dados por meio da investigação de informações relacionadas à realidade educacional brasileira. Durante a realização da proposta, os estudantes foram responsáveis pela coleta, organização e tratamento de bases de dados referentes ao ano de 2024. A análise concentrou-se em variáveis relevantes para a compreensão do cenário educacional do país, tais como distorção idade-série, evasão escolar e índices de reprovação. Os resultados obtidos nas análises realizadas pelos alunos evidenciaram diferentes padrões regionais nos indicadores investigados. Observou-se, por exemplo, maior incidência de distorção idade-série na região Norte, índices mais elevados de evasão escolar na região Nordeste e taxas de reprovação mais altas na região Sul. Tais resultados indicam a presença de desigualdades educacionais entre diferentes contextos regionais do país. Do ponto de vista pedagógico, a experiência possibilitou o desenvolvimento de habilidades relacionadas à interpretação e representação gráfica de dados, além de favorecer uma compreensão mais aprofundada dos estudantes acerca dos desafios evidenciados pelos indicadores educacionais analisados.

Palavras-chave: Indicadores Educacionais, Análise De Dados, Letramento Em Dados, Educação Básica.

INTRODUÇÃO

A crescente disponibilidade de dados sobre diferentes aspectos da sociedade tem ampliado a necessidade de desenvolver, desde a educação básica, conhecimentos e habilidades relacionados à interpretação de informações, uma vez que a vida social, acadêmica e profissional está cada vez mais organizada por dados (GEBRE, 2022; PINTO et. al., 2023). Nesse contexto, torna-se importante que os estudantes desenvolvam não apenas a capacidade de acessar informações quantitativas, mas

[1] Professora Efetiva do Instituto Federal do Sertão Pernambucano - Campus Salgueiro, brena.maia@ifsertao-pe.edu.br

[2]Estudante do Ensino Médio da Escola Professora Olindina Roriz Dantas, Belém do São Francisco-PE, gabrielabelfort@aluno.escolaolindina.com

ISSN: 2358-8829

também de analisá-las de maneira crítica e contextualizada, considerando os contextos de produção, circulação e uso social dos dados. Visto isso, o letramento de dados tem sido sugerido como uma competência importante que os indivíduos precisam para ter sucesso em uma sociedade intensiva em dados (GRAY; GERLITZ; BOUNEGRU, 2018; GEBRE, 2022).

A utilização pedagógica de bases de dados públicas pode favorecer o desenvolvimento do letramento em dados e do pensamento crítico. Ao participar de etapas como seleção, organização, tratamento, representação e interpretação de informações, os estudantes têm a oportunidade de compreender que os dados não constituem apenas conjuntos de números, mas representações de situações reais que precisam ser analisadas e contextualizadas. ATENAS, HAVEMANN e PRIEGO (2015) vêem o valor educacional dos dados abertos como um componente fundamental em atividades de aprendizagem baseada em pesquisa e cenários, onde sua implementação pode aprimorar o letramento digital e apoiar o desenvolvimento de habilidades críticas, analíticas, colaborativas e de cidadania.

A partir dessa perspectiva, este trabalho apresenta um relato de experiência pedagógica desenvolvido com estudantes do segundo ano do ensino médio. A experiência teve como objetivo promover o desenvolvimento do pensamento crítico e do letramento em dados por meio da investigação de indicadores educacionais brasileiros. Para isso, os estudantes participaram de diferentes etapas de análise de bases de dados públicas, envolvendo a coleta, organização, tratamento e representação das informações.

METODOLOGIA

Esta experiência pedagógica foi desenvolvida com seis estudantes do segundo ano do ensino médio da escola Professora Olindina Roriz Dantas, situada na zona rural de Belém do São Francisco-PE, durante o ano letivo de 2025. A proposta teve como objetivo demonstrar como os dados podem ser transformados em informações úteis, de fácil compreensão e visualização, além de estimular o desenvolvimento do pensamento crítico por meio da análise de indicadores educacionais obtidos em bases de dados públicas e confiáveis.

ISSN: 2358-8829

O primeiro passo consistiu na pesquisa de bases de dados públicas, confiáveis e atualizadas. Para o desenvolvimento do trabalho, os alunos utilizaram planilhas disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), órgão responsável pela produção e divulgação das principais informações estatísticas, geográficas, econômicas e sociais do Brasil (IBGE, 2025). Os dados selecionados foram referentes ao ano de 2024, últimos dados publicados na época da experiência.

As planilhas, que já estavam no formato CSV, passaram por processos de filtragem e limpeza dos dados, com o objetivo de eliminar inconsistências e garantir a qualidade das informações antes de sua utilização. Posteriormente, os dados foram importados para o Visual Studio Code (VS Code), ambiente de desenvolvimento utilizado para o tratamento dos dados e para a construção e visualização dos gráficos empregados nas análises. Para manipulação dos dados foi utilizada a linguagem Python, com importação das bibliotecas Numpy, Pandas e Matplotlib.

Os indicadores educacionais analisados foram distorção idade-série, taxas de aprovação, reprovação e evasão escolar. A partir dessas informações, buscou-se demonstrar uma comparação dos indicadores em cada região do país (Norte, Nordeste, Sudeste, Sul, Centro-Oeste), identificando padrões e desigualdades presentes no cenário educacional brasileiro.

Por fim, os resultados obtidos foram compartilhados e discutidos em sala de aula, o que permitiu que os estudantes correlacionassem os indicadores analisados com a realidade vivenciada por eles. Essa etapa favoreceu a reflexão crítica sobre as desigualdades educacionais presentes no país e contribuiu para uma compreensão mais aprofundada do contexto educacional brasileiro.

Vale ressaltar, também, o uso de ferramentas de Inteligência Artificial (IA) no suporte ao desenvolvimento das atividades pedagógicas e na elaboração deste trabalho. A IA foi empregada como recurso auxiliar durante a programação com a linguagem Python, auxiliando os estudantes no entendimento da sintaxe e na resolução de erros de código durante o tratamento e a geração dos gráficos de dados. Adicionalmente, utilizou-se a ferramenta na etapa de redação científica, atuando como apoio na estruturação das seções do artigo e na revisão e correção ortográfica do texto.

ISSN: 2358-8829

REFERENCIAL TEÓRICO

Aprendizagem baseada em dados

O trabalho pedagógico com dados envolve diferentes etapas, como a seleção de fontes, coleta, organização, tratamento, análise, representação e interpretação das informações. Fagerlund, Palsa e Mertala (2025), ao analisarem estudos relacionados ao letramento em dados na educação básica, destacam a importância de experiências que utilizem dados reais e situações relacionadas ao cotidiano dos estudantes. Essa abordagem possibilita que os alunos assumam uma posição mais ativa diante das informações, participando do processo investigativo e desenvolvendo conhecimentos que ultrapassam o domínio de procedimentos técnicos.

Essa perspectiva também está relacionada às orientações da Base Nacional Comum Curricular (BNCC), que enfatiza a necessidade de desenvolver uma relação crítica, ética e responsável com as tecnologias digitais e com as informações disponíveis na sociedade (BRASIL, 2018). Nesse sentido, a utilização pedagógica de bases de dados pode articular conhecimentos tecnológicos, científicos e sociais, favorecendo situações de aprendizagem nas quais os estudantes utilizem recursos digitais para investigar fenômenos e construir interpretações fundamentadas em evidências.

Na experiência apresentada neste trabalho, essa perspectiva manifesta-se na participação dos estudantes em diferentes etapas da investigação. Em vez de receberem dados previamente organizados para análise, os alunos realizaram a pesquisa das bases em fontes abertas, selecionaram as informações pertinentes, efetuaram procedimentos de filtragem e limpeza e, posteriormente, utilizaram recursos computacionais para tratamento e representação dos dados. A atividade, portanto, aproxima-se de uma perspectiva investigativa, na qual o contato com dados reais constitui parte do próprio processo de aprendizagem.

Letramento em dados e pensamento crítico

O letramento em dados está relacionado à capacidade de compreender, utilizar, analisar e comunicar informações produzidas a partir de dados. É uma competência que não se limita ao domínio de ferramentas ou procedimentos técnicos, uma vez que

ISSN: 2358-8829

envolve também a compreensão do contexto em que os dados são produzidos e a capacidade de avaliar as informações apresentadas. Para Gebre (2022), há quatro orientações do letramento de dados: desenvolvimento de competência, investigação com dados, conscientização sobre dados pessoais e engajamento cívico. A relação entre análise de dados e pensamento crítico também é destacada por Koga (2022), ao compreender o letramento estatístico como uma competência de ordem superior associada à interpretação e à avaliação crítica de informações.

Essa compreensão é particularmente relevante quando os dados utilizados representam fenômenos sociais. A interpretação de um indicador não pode ser reduzida à identificação do maior ou menor valor observado, pois exige compreender aquilo que está sendo medido, as condições em que a informação foi produzida e os possíveis significados dos padrões identificados. Dessa maneira, o trabalho com dados pode favorecer uma postura questionadora, na qual os estudantes são estimulados a analisar não apenas os resultados, mas também aquilo que pode ou não ser inferido a partir deles.

No presente trabalho, o letramento em dados é compreendido, portanto, como um conjunto de conhecimentos e habilidades relacionados à localização, organização, tratamento, representação, interpretação e questionamento de informações baseadas em dados. Essa concepção está diretamente relacionada ao objetivo da experiência, uma vez que os estudantes participaram do processo de transformação de bases de dados em representações que posteriormente foram analisadas e discutidas em sala de aula.

Indicadores educacionais e compreensão da realidade brasileira

Os indicadores educacionais são instrumentos importantes para a análise do sistema educacional. Por meio de medidas quantitativas, é possível acompanhar aspectos relacionados ao fluxo e ao rendimento escolar, bem como realizar comparações entre diferentes períodos, regiões e contextos. Sua utilização permite transformar informações produzidas por sistemas de coleta de dados em elementos que podem contribuir para a compreensão do cenário educacional brasileiro.

Entre os indicadores analisados na experiência aqui mencionada estão a distorção idade-série e as taxas de aprovação, reprovação e evasão escolar. Esses

ISSN: 2358-8829

indicadores representam diferentes dimensões da trajetória educacional dos estudantes. A distorção idade-série permite observar situações em que a idade do estudante não corresponde à etapa de ensino esperada, enquanto as taxas de aprovação, reprovação e evasão fornecem informações relacionadas ao fluxo e ao rendimento escolar.

A análise conjunta dessas informações possibilita uma compreensão mais ampla da realidade educacional, uma vez que diferentes indicadores podem revelar diferentes perspectivas de uma mesma situação. Nesse sentido, a comparação entre indicadores, períodos e regiões constitui uma estratégia importante para identificar padrões e diferenças no cenário educacional.

A dimensão territorial também assume importância nesse tipo de análise. A comparação entre as regiões brasileiras permite observar como determinadas situações educacionais se distribuem no território nacional, evidenciando diferenças que podem não ser percebidas quando os dados são apresentados apenas de forma agregada para o país. Dessa maneira, os indicadores podem ser utilizados não apenas para descrever uma realidade, mas também para problematizar desigualdades e estimular discussões sobre os diferentes contextos educacionais.

Ferramentas computacionais para o tratamento e a visualização de dados

O desenvolvimento de atividades envolvendo a análise de dados requer ferramentas que auxiliem na organização, no tratamento e na representação das informações. Nesse contexto, o uso de linguagens de programação se destaca pela possibilidade de automatizar operações e manipular conjuntos de dados por meio de bibliotecas específicas. Para o desenvolvimento do trabalho proposto, os estudantes usaram a linguagem Python, uma linguagem de programação interpretada de código aberto e uso geral, bem adequada para tarefas padrão de programação, como limpeza de dados, interação com recursos web e análise de texto (HARRIS, 2020).

Para o desenvolvimento das atividades, o ambiente de desenvolvimento escolhido foi Visual Studio Code (VS Code), que permite a escrita, a edição e a execução de códigos em Python. Além disso, bibliotecas como NumPy, Pandas e Matplotlib oferecem recursos que facilitam diferentes etapas do processo de análise. O NumPy é uma biblioteca de código aberto desenvolvida pela comunidade, que fornece

ISSN: 2358-8829

um objeto array multidimensional em Python junto com funções conscientes de arrays que operam nele, enquanto o Pandas é uma biblioteca com estruturas de dados e ferramentas robustas para trabalhar com conjuntos de dados estruturados (HARRIS, 2020; MCKINNEY, 2011). Já o Matplotlib é um pacote gráfico 2D usado para desenvolvimento de aplicações, scripts interativos e geração de imagens, que contribui para a identificação de padrões e a comparação entre os dados analisados (HUNTER et al., 2007).

RESULTADOS E DISCUSSÃO

Após a organização dos dados, os estudantes iniciaram o processo de representação dos mesmos através de gráficos, que serão mostrados a seguir. A Figura 1 apresenta as taxas de aprovação e reprovação dos alunos do ensino médio, em cada região do país. A maior taxa de aprovação foi a da região Centro-Oeste, enquanto a região Sul registrou o menor índice de aprovação. O gráfico mostra ainda que a região Sul apresenta a maior taxa de reprovação, enquanto o Norte possui o menor índice de reprovação. Durante a interpretação deste gráfico, os alunos levantaram o seguinte questionamento: *"Que medidas poderiam ser adotadas para reduzir a reprovação escolar no Brasil?"*.

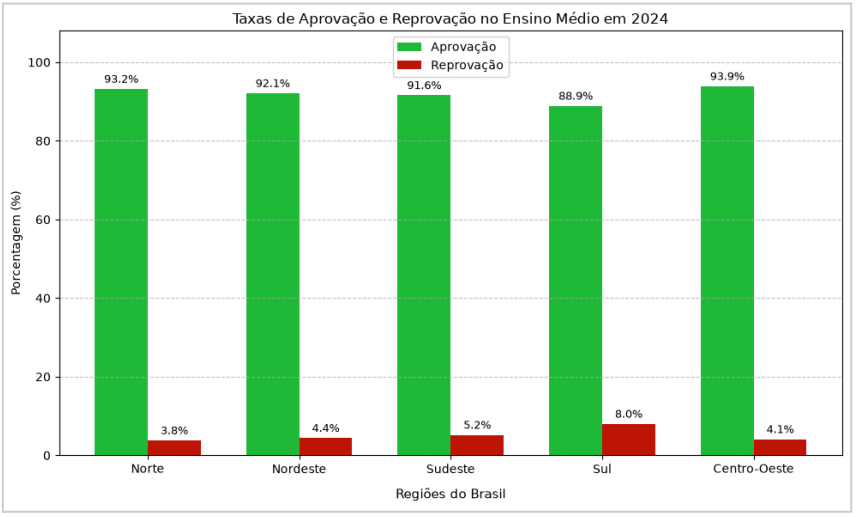

Figura 1. Taxas de aprovação e reprovação no ensino médio em 2024

Podemos notar na Figura 1, que a porcentagem de aprovação junto à de reprovação não atinge 100% dos casos. Isso acontece por conta da evasão de estudantes. A Figura 2 apresenta a taxa de evasão no ensino médio, de acordo com a região. O

ISSN: 2358-8829

gráfico mostra que o Nordeste apresenta a maior taxa de evasão escolar, com 3,5%, enquanto a região Centro-Oeste possui a menor taxa, com 2%. No processo de interpretação destes resultados, os alunos levantaram o questionamento: *"Quais fatores sociais e econômicos podem explicar a maior taxa de evasão escolar na região Nordeste em comparação às demais regiões brasileiras?"*

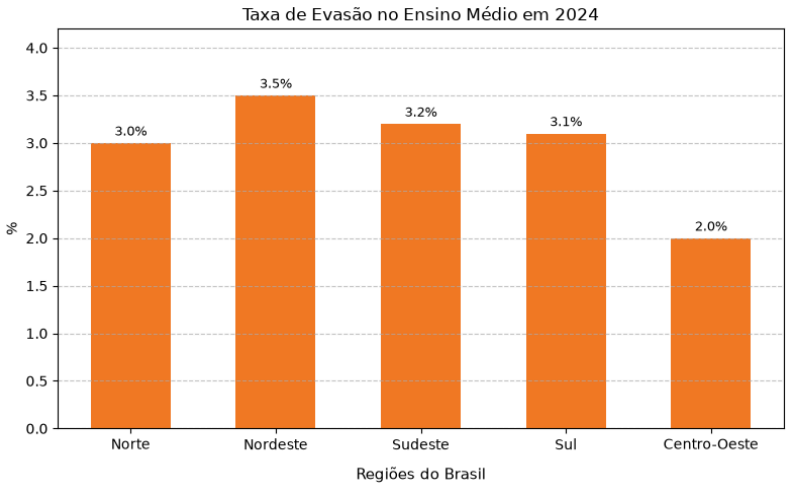

Figura 2. Taxa de evasão no ensino médio no ano de 2024

A Figura 3 apresenta os resultados obtidos sobre a distorção idade-série. Podemos observar que o Norte possui a maior taxa, com 26,8%. Já o Sudeste apresenta a menor taxa, com 13,9%. Estes resultados trouxeram aos alunos questionamentos como: *"Por que alguns estudantes chegam ao ensino médio com idade superior à esperada para essa etapa escolar?"*.

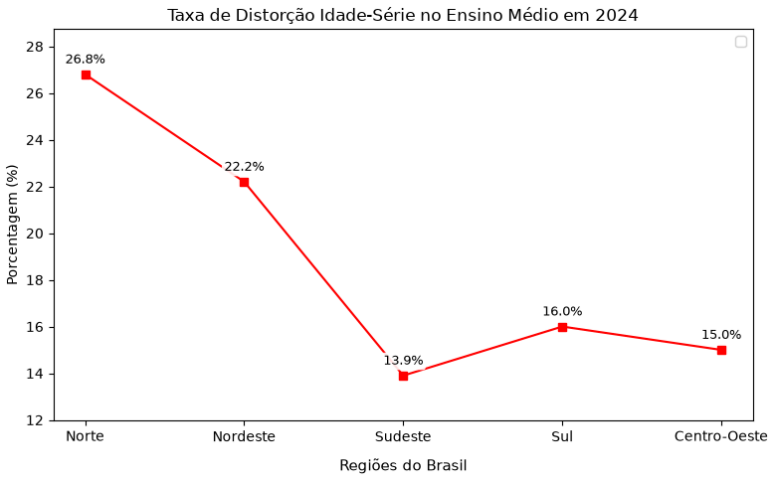

Figura 3. Taxa de Distorção Idade-Série no Ensino Médio em 2024

ISSN: 2358-8829

Nesse processo de análise dos resultados, os dados deixaram de ser compreendidos apenas como valores numéricos e passaram a constituir elementos para a formulação de perguntas e a discussão de possíveis explicações para a realidade educacional analisada.

CONSIDERAÇÕES FINAIS

O presente trabalho apresentou um relato de experiência pedagógica baseado na utilização de bases de dados públicas para a análise de indicadores educacionais por estudantes do ensino médio. A proposta permitiu integrar o uso prático da linguagem Python e de bibliotecas para manipulação e visualização de dados à interpretação crítica do contexto educacional brasileiro referentes ao ano de 2024.

Os resultados obtidos evidenciaram padrões regionais relevantes no cenário educacional do país, revelando desigualdade em indicadores chave como distorção idade-série, taxas de evasão e de reprovação. A atividade demonstrou que o contato direto com dados reais possibilita que os estudantes ultrapassem a simples manipulação técnica de ferramentas computacionais, transformando números agregados em questionamentos sobre determinantes sociais e econômicos das desigualdades educacionais. Do ponto de vista pedagógico, a experiência contribuiu significativamente para o fortalecimento do letramento em dados e do pensamento crítico dos alunos.

AGRADECIMENTOS

Agradecemos a Paulo Dantas, diretor presidente da Agrodan e fundador da Escola Professora Olindina Roriz Dantas, projeto social da empresa, pelas oportunidades proporcionadas aos estudantes por meio dessa iniciativa. A escola tem contribuído para a construção de um futuro com mais possibilidades para estudantes da zona rural de Belém do São Francisco, oferecendo experiências que ampliam seus horizontes e fazem a diferença em suas trajetórias. Agradecemos, também, aos estudantes Adrian, Geisilany, Marcos Gabriel, Maria Gabriela, Mariane e Yasmim, do segundo ano do ensino médio (turma de 2025), que participaram da experiência relatada neste trabalho, pelo empenho, dedicação e envolvimento nas atividades desenvolvidas, fundamentais para a realização do projeto.

ISSN: 2358-8829

REFERÊNCIAS

ATENAS, Javiera; HAVEMANN, Leo; PRIEGO, Ernesto. Open data as open educational resources: Towards transversal skills and global citizenship. **Open praxis**, v. 7, n. 4, p. 377-389, 2015.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, DF: MEC, 2018. Disponível em: https://www.gov.br/mec/pt-br/cne/base-nacional-comum-curricular-bncc. Acesso em: 13/10/2026.

HARRIS, Charles R. et al. Array programming with NumPy. **nature**, v. 585, n. 7825, p. 357-362, 2020.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Plano de Dados Abertos 2024-2025**. Rio de Janeiro: IBGE, 2024. Disponível em: https://www.ibge.gov.br/acesso-informacao/dados-abertos.html. Acesso em: 16 set. 2025.

GEBRE, Engida. Conceptions and perspectives of data literacy in secondary education. **British Journal of Educational Technology**, v. 53, n. 5, p. 1080-1095, 2022.

GRAY, Jonathan; GERLITZ, Carolin; BOUNEGRU, Liliana. Data infrastructure literacy. **Big Data & Society**, v. 5, n. 2, p. 2053951718786316, 2018.

HUNTER, John D. Matplotlib: A 2D graphics environment. **Computing in science & engineering**, v. 9, n. 03, p. 90-95, 2007.

MCKINNEY, Wes et al. pandas: a foundational Python library for data analysis and statistics. **Python for high performance and scientific computing**, v. 14, n. 9, p. 1-9, 2011.

PINTO, María et al. A strategic approach to information literacy: data literacy. A systematic review. 2023.

PYTHON SOFTWARE FOUNDATION. **Python Documentation**. Disponível em: https://www.python.org/doc/. Acesso em: 13/10/2026.