

**Análise de Dados Sociais com Python: Taxa de analfabetismo da População de
15 anos ou mais por Município do Maranhão**

CÍCERO GABRIEL DE SOUSA CARDOSO

Universidade Estadual do Maranhão

cicerocardoso@aluno.uema.br

ALICE FERNANDA BORGES CHAVES

Universidade Estadual do Maranhão

alicechaves034@gmail.com

PAULO FERNANDES DA SILVA JUNIOR

Universidade Estadual do Maranhão

paulofernandes@professor.uema.br

ANA ELISA OLIVEIRA SILVA

Universidade Estadual do Maranhão

anna.eosilva01@gmail.com

BRUNO RAPHAEL ANDRADE SANTOS

Universidade Estadual do Maranhão

bruno.20240060243@aluno.uema.br

SAMYRA DE ARAUJO LOBO SILVA

Universidade Estadual do Maranhão

samyra.20240002025@aluno.uema.br

JÚLIO CÉSAR COSTA DA SILVA

Universidade Estadual do Maranhão

julio.20240045597@aluno.uema.br

RESUMO

Neste é apresentada uma análise de dados sociais, para a identificação de indicadores da educação do Estado do Maranhão do período de 1991 a 2022. Os dados foram obtidos da plataforma Ipeadata, processados na linguagem *Python*, permitindo a organização dos dados para aquisição de informação. A análise foi realizada com o uso das bibliotecas *pandas*, *matplotlib* e *seaborn*, possibilitando a construção de figuras que apresentam dos indicadores por município do Maranhão. Como produto foi desenvolvido um *dashboard* interativo com a biblioteca *Streamlit*, em uma aplicação web, com a

geração de *widgets* para a visualização dinâmica dos resultados. O estudo busca identificar indicadores educacionais e os padrões existentes nos municípios maranhenses, contribuindo para uma melhor compreensão desse cenário socioeducacional.

Palavras-chave: Indicadores sociais; *Python*; Maranhão; *dashboard* interativo.

1 INTRODUÇÃO

O analfabetismo permanece como um dos principais desafios sociais do Brasil, sobretudo em regiões historicamente marcadas por desigualdades educacionais e socioeconômicas. Embora o país tenha avançado significativamente nas últimas décadas, os dados ainda evidenciam disparidades relevantes entre estados e municípios, o que reforça a necessidade de estudos que analisem os indicadores de forma detalhada e contextualizada (IBGE, 2023). Nesse sentido, compreender a evolução do analfabetismo em um período histórico amplo contribui para fundamentar políticas públicas voltadas à inclusão social e à melhoria da qualidade de vida da população (UNESCO, 2019).

O Estado do Maranhão apresenta índices de analfabetismo superiores à média nacional, especialmente entre cidadãos com mais de 15 anos. Essa realidade evidencia a importância de se desenvolver instrumentos capazes de organizar e interpretar dados de maneira eficiente, a fim de apoiar a tomada de decisão por parte de gestores públicos e pesquisadores (IPEA, 2022). O uso de tecnologias de análise de dados e visualização interativa surge, assim, como um recurso estratégico para transformar grandes volumes de informações em conhecimento aplicável (CHEN; CHIANG; STOREY, 2012)

O analfabetismo é um indicador importante do desenvolvimento humano e da inclusão social, bem como remete à necessidade de soluções acessíveis, transparentes e de baixo custo para análise e disseminação de dados públicos. Ao aliar metodologias de ciência de dados a ferramentas de código aberto, busca-se oferecer um produto acadêmico e prático que contribua para o debate e para o desenho de estratégias de enfrentamento do analfabetismo no Maranhão (CAMPANTE; GONÇALVES; GONÇALVES, 2024).

Neste trabalho são apresentados indicadores sociais vinculados ao analfabetismo

no Estado do Maranhão no período de 1991 a 2022. A implementação foi realizada na linguagem *Python* para processar e analisar séries históricas disponibilizadas pelo Ipeadata. Com a exploração dos dados, foi desenvolvido um *dashboard* interativo com a biblioteca *Streamlit*, que permite a visualização dinâmica dos resultados, por meio de figuras e widgets, ampliando a compreensão dos padrões e tendências. Este artigo está organizado em mais quatro partes além desta Introdução. Na seção 2 é apresentada a fundamentação teórica do trabalho. Na seção 3 são indicados os Materiais e Métodos utilizados. Na seção 4 os Resultados obtidos são apresentados e discutidos e na seção 5 são indicadas as Considerações Finais.

2 DESENVOLVIMENTO

2.1.1 FUNDAMENTAÇÃO TEÓRICA

O *Business Intelligence* (BI) é um conceito abrangente que envolve arquiteturas, ferramentas, bancos de dados, aplicações e metodologias voltadas à análise de dados. Dessa forma, seu principal objetivo consiste em possibilitar o acesso interativo às informações, permitindo que gestores, analistas e demais profissionais possam manipulá-las de forma adequada e realizar análises aprofundadas, identificando tendências e padrões relevantes, de modo a apoiar o processo de tomada de decisões. (SHARDA; DELEN; TURBAN, 2010).

O processo do *Business Intelligence* envolve diversas etapas que se iniciam com a extração de dados brutos de diferentes fontes, que podem ter variados formatos e conexões. Em seguida, ocorre o processo de ETL (*Extract, Transform and Load*), responsável por tratar, padronizar e aplicar regras de negócio aos dados antes de carregá-los no *Data Warehouse* (DW). No DW, os dados ficam unificados e modelados de forma multidimensional, permitindo a criação de cubos e indicadores, o que configura a camada semântica para análise. Por fim, os painéis, relatórios e dashboards são construídos por meio de ferramentas de BI, que tornam as informações intuitivas e de fácil compreensão (DUARTE, 2020).

A análise de dados se consolida como um dos componentes centrais do BI, pois é por meio dela que as informações armazenadas podem ser examinadas,

interpretadas e transformadas em insights relevantes para a organização (PAIVA, 2023). Para isso, diversas ferramentas podem ser utilizadas, destacando-se o Python, amplamente aplicado em áreas como desenvolvimento web, ciência de dados e aprendizado de máquina, beneficiando-se de uma ampla variedade de bibliotecas que agilizam o desenvolvimento, permitem a implementação de soluções complexas e ampliam significativamente as possibilidades de aplicação da linguagem.

No contexto do presente projeto, três bibliotecas do *Python* terão papel central no processo de análise de dados: *pandas*, *matplotlib* e *seaborn*. A biblioteca *pandas* destaca-se pela manipulação de dados estruturados por meio de estruturas como *DataFrame*, permitindo organização e transformação eficiente das informações (MCKINNEY, 2021). Já a *matplotlib* é amplamente utilizada para a construção de visualizações gráficas, oferecendo recursos para a criação de figuras personalizadas e adequadas a diferentes contextos analíticos (HUNTER, 2007). Por fim, o *seaborn* complementa esse processo ao fornecer representações estatísticas mais sofisticadas e visualmente atraentes, integrando-se de forma direta com o *pandas* e o *matplotlib* (WASKOM, 2021). A junção dessas três bibliotecas é fundamental para este projeto, pois garante um fluxo completo que envolve a organização, a exploração e a visualização clara dos dados analisados.

2.1.2 MATERIAIS E MÉTODOS

O desenvolvimento do trabalho seguiu uma abordagem combinada de pesquisa bibliográfica e implementação prática do sistema, estruturada nas seguintes etapas:

- Pesquisa, coleta e análise de livros, artigos e sites especializados, a fim de entender e aplicar os conhecimentos retidos na análise de dados sociais, especificamente aqueles relacionados aos indicadores sociais do Ipeadata - sistema mantido pelo Instituto de Pesquisa Econômica Aplicada (IPEA), que disponibiliza, de forma gratuita, uma ampla base de dados econômicos, sociais e demográficos do Brasil e do mundo.

- Processo de preparação dos dados coletados para serem manuseados

corretamente por códigos em *Python*;

- Construção do código em *Python* para organizar, tratar, explorar e visualizar as informações de forma funcional e eficaz, utilizando bibliotecas específicas;

- Uso da biblioteca *streamlit* no código em *Python* para a apresentação dos resultados, consistindo em uma forma de visualizar e interagir com os dados, transformando pequenas linhas de código em uma aplicação *web*.

Inicialmente, por meio de pesquisas direcionadas, foram identificados livros, artigos científicos e sites especializados que abordavam “análise de dados” como tema principal. A partir desse referencial, tornou-se possível a exploração e tratamento dos indicadores sociais, permitindo a construção das figuras desenvolvidas neste trabalho, utilizando dados do estado do Maranhão no período de 1991 a 2022.

Considerando o site Ipeadata como fonte de dados, foram baixadas manualmente as planilhas associadas às taxas de analfabetismo da população de 15 anos ou mais por município no Brasil, envolvendo o período de 1991 a 2022. A partir dos dados contidos nessas planilhas, foi possível a seleção específica dos registros referentes aos municípios do estado do Maranhão, considerando a segmentação por alfabetizados e não alfabetizados com 15 anos ou mais. A partir do levantamento realizado foi executada a avaliação dos dados coletados, para evitar informações imprecisas, parciais ou duplicadas, reduzindo o risco de conclusões equivocadas (ENGINEERING BRASIL, 2023). Dessa forma, a análise feita por segmentação permitiu identificar tendências e semelhanças nas taxas de alfabetização entre diferentes regiões do Maranhão e visualizar sua evolução ao longo do tempo.

Na execução do trabalho foi desenvolvido um script em Python que leu planilhas .csv, limpou linhas vazias e organizou os dados por siglas, municípios, datas e percentuais médios de alfabetizados e não alfabetizados, utilizando a biblioteca pandas para análise e manipulação de dados (McKINNEY, 2010). Para visualização, Matplotlib e Seaborn foram empregados na criação e armazenamento de figuras (HUNTER, 2007). A camada de apresentação foi

desenvolvida com Streamlit, que converte scripts Python em interface web interativa, dispensando HTML, CSS ou JavaScript e permitindo a criação de um aplicativo intuitivo e eficaz com poucas linhas de código (MHADHBI, 2024).

2.2 RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os resultados da análise exploratória e descritiva da taxa de analfabetismo da população com 15 anos ou mais nos municípios do Maranhão.

.1 VISUALIZAÇÃO E ANÁLISE DESCRIPTIVA DOS DADOS BRUTOS

O estudo inicial do conjunto de dados foi realizado com a biblioteca Pandas, contudo, os dados não tratados apresentavam problemas de organização, como colunas desestruturadas, valores sem formatação e tipos de variáveis inconsistentes. Diante disso, foi necessário substituir algumas colunas, remover linhas com valores NaN e criar novas variáveis denominadas 'Percentual Médio de Não Alfabetizados' e 'Percentual Médio de Alfabetizados', obtidas a partir da média dos percentuais de cada ano. O código referente a essa etapa de limpeza encontra-se no Quadro 1. Como resultado, é possível compreender cada variável e visualizar as taxas percentuais anuais, percentual de alfabetizados e não alfabetizados como demonstrado na Tabela 2.

Quadro 1 - Código para limpeza e pré-processamento do conjunto de dados

```
import pandas as pd

def dataset_limpo():

    # Carrega e remove a primeira linha irrelevante
    df = pd.read_csv("bancodados.csv", sep=";", decimal=",").drop(0)

    # Renomeia colunas principais
    df = df.rename(columns={

        "Unnamed: 0": "Sigla",
        "Unnamed: 1": "Código",
        "Unnamed: 2": "Município",
        "Unnamed: 3": "1991 (%)",
        "Unnamed: 4": "2000 (%)",
```

```

    "Unnamed: 5": "2010 (%)",
    "Unnamed: 6": "2022 (%)"
})

# Ajusta tipo e calcula médias
df["Código"] = df["Código"].astype(int)
anos = ["1991 (%)", "2000 (%)", "2010 (%)", "2022 (%)"]
df["% Médio Não Alfabetizados"] = df[anos].mean(axis=1)
df["% Médio Alfabetizados"] = 100 - df["% Médio Não Alfabetizados"]

# Retorna o DataFrame limpo e com as novas colunas
return df

```

Fonte: Elaborado pelo autor.

Com o conjunto de dados organizado pode-se definir o que seria apresentado graficamente, como, quais municípios que apresentam as maiores e menores taxas de analfabetismo, qual a progressão dessa situação, qual a distribuição desses dados e como a taxa de analfabetismo se compara ao longo dos anos entre municípios.

Tabela 2 - Dados organizados e tratados

Sigla	Código	Município	1991 (%)	2000 (%)	2010 (%)	2022 (%)	Média de não Alfabetizados (%)	Média de Alfabetizados (%)
MA	2100105	Afonso Cunha	61.02	42.22	34.64	23.22	40.2750	59.7250
MA	2100204	Alcântara	48.65	30.26	22.44	18.48	29.9575	70.0425

Fonte: Elaborado pelo autor.

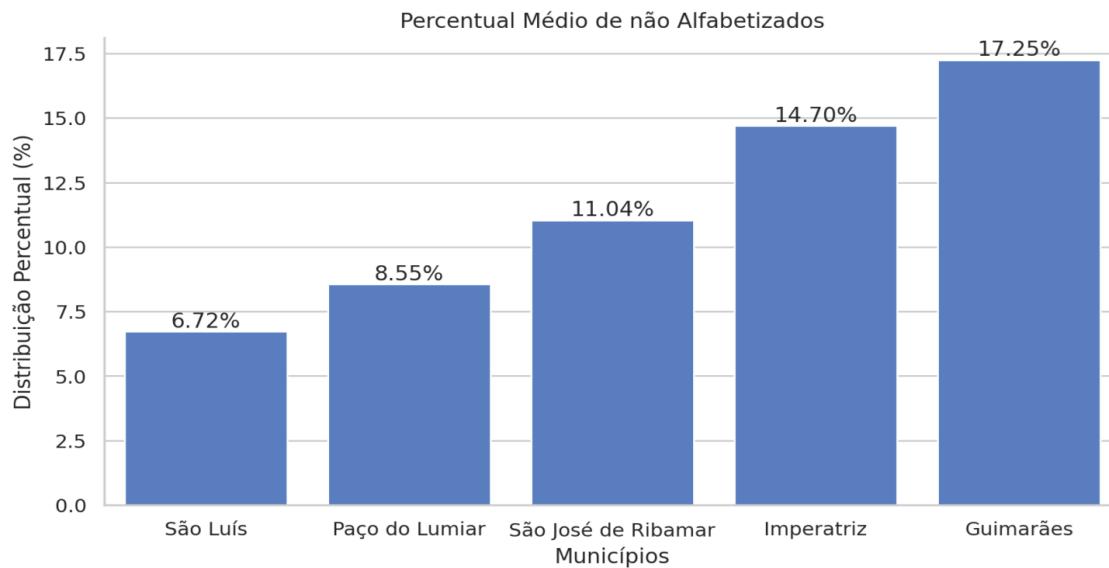
2.2.2 ANÁLISE DA DISTRIBUIÇÃO PERCENTUAL DO ANALFABETISMO POR MUNICÍPIO

A partir dos dados organizados e tratados, desenvolveu-se em Python, com o apoio da biblioteca *Streamlit*, uma aplicação capaz de estruturar opções de filtragem e geração gráfica. O *Dashboard* interativo permite escolher o tipo de

dado que deseja filtrar: percentual médio de alfabetizados e não alfabetizados, o tipo de ordenação, crescente, decrescente, e a quantidade de municípios que deseja listar. Após a escolha das opções, é gerado automaticamente uma figura de barras com o valor percentual correspondente ao município na extremidade superior da barra, como observado na Figura 1.

A análise da taxa de analfabetismo nos municípios do Maranhão evidencia desigualdades significativas. A capital, São Luís, apresenta o menor índice, com 6,72% da população não alfabetizada, resultado influenciado pela concentração econômica e pela centralização dos serviços educacionais. Já municípios próximos, como Paço do Lumiar e São José de Ribamar, registram percentuais superiores, mas que, quando comparados ao restante do estado, tornam-se relativamente baixos. Assim, ressalta-se a importância de uma análise mais ampla, articulando as taxas de analfabetismo a outros indicadores socioeconômicos e educacionais, a fim de compreender suas causas e efeitos.

Figura 1 - Figura crescente do percentual médio de não alfabetizados nos municípios de São Luís, Paço do Lumiar, São José de Ribamar, Imperatriz e Guimarães.

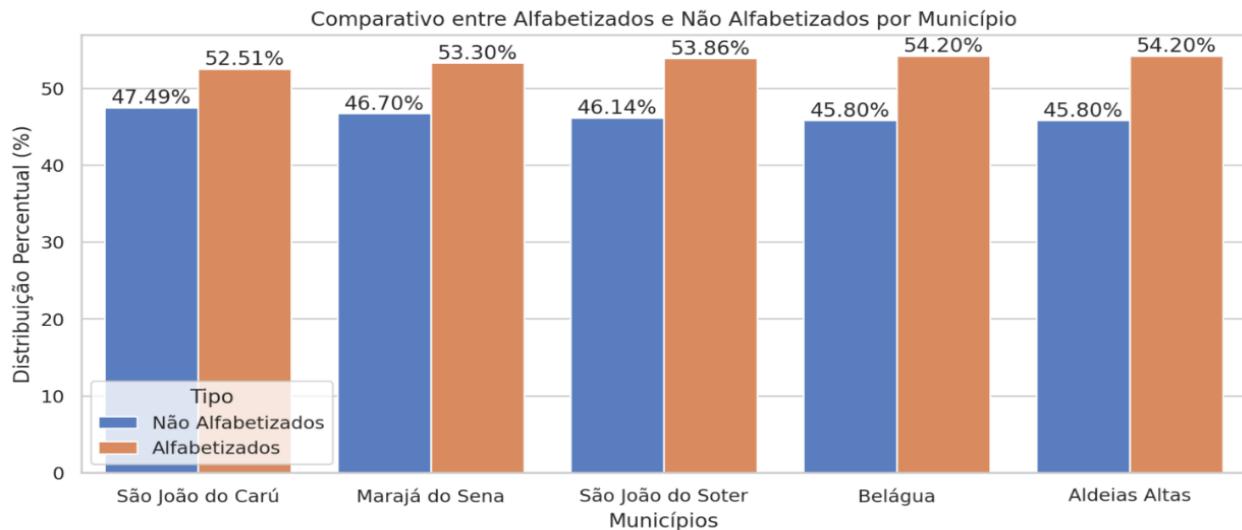


Fonte: Elaborado pelo autor.

2.2.3 ESTUDO COMPARATIVO ENTRE POPULAÇÃO ALFABETIZADA E NÃO-ALFABETIZADA

Na etapa de pré-processamento dos dados, foi usado o processo da engenharia de atributos, processo de transformação de dados originais em novas informações, formulando novas colunas e características. Dessa forma, uma nova sessão do painel foi desenvolvida para facilitar a visualização comparativa das taxas da população alfabetizada e não alfabetizada, possibilitando ao usuário acessar um conjunto de opções semelhantes à sessão de visualização gráfica por município, porém sem a escolha por filtro. Na Figura 2 são indicadas a taxa de analfabetismo da população em comparação ao percentual da população alfabetizada.

Figura 2 - Comparativo percentual da taxa da população não alfabetizada em relação à população alfabetizada por município do Maranhão.



Fonte: Elaborado pelo autor.

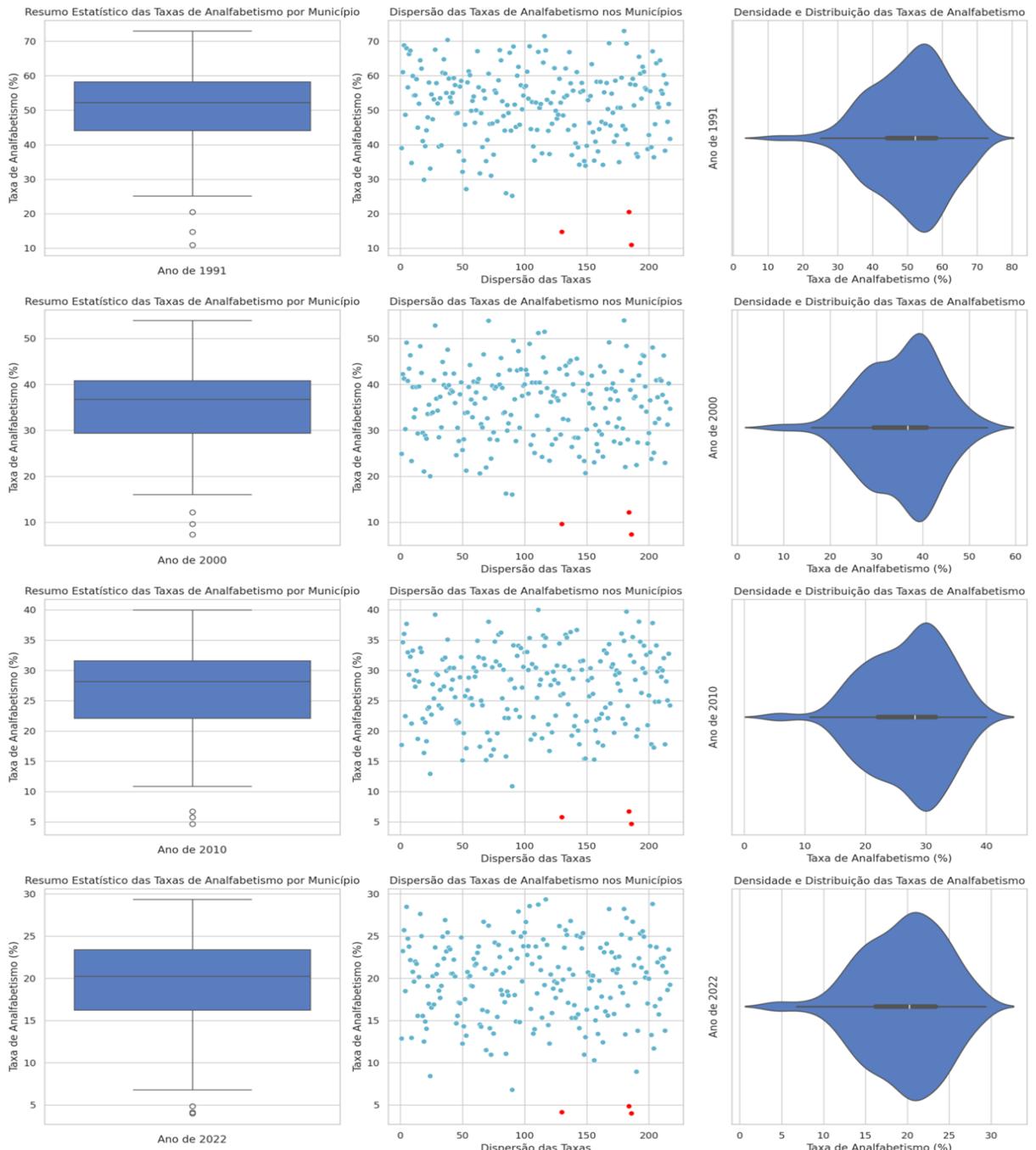
Ao analisar os percentuais de alfabetização nos municípios do Maranhão, percebe-se uma distribuição desigual. Um exemplo relevante é o município de São João do Carú, que apresenta a maior taxa de analfabetismo do estado. Nesse município, a proporção de não alfabetizados chega a se aproximar da de alfabetizados, configurando um quadro crítico. Essa realidade está associada ao baixo desempenho educacional refletido no IDEB dos anos finais do ensino

fundamental, que registrou nota 4, inferior à meta nacional de 6, segundo dados do IBGE. Tal resultado compromete a continuidade da trajetória escolar e contribui para a reprodução de dificuldades de aprendizagem, o que perpetua altos índices de analfabetismo. Diante desse cenário, evidencia-se a necessidade de políticas públicas voltadas à melhoria da qualidade educacional e à redução das desigualdades regionais.

2.2.4 DISTRIBUIÇÃO NUMÉRICA DA TAXA DE ANALFABETISMO AO LONGO DOS ANOS

No estudo dos dados numéricos, é fundamental avaliar a distribuição de valores, identificação de mínimos e máximos, dados acima da média, chamado de *outliers*, e como estão correlacionadas essas variáveis, sendo gerado o gráfico de dispersão, densidade e distribuição com uso das bibliotecas *ScatterPlot*, *ViolinPlot* e *BoxPlot*, como observado na Figura 3. No resultado podem ser destacados os *outliers* e a concentração de valores ao longo dos anos de 1991 a 2022.

Figura 3 - Análise descritiva da taxa de analfabetismo em municípios do Maranhão,
1991-2022.



Fonte: Elaborado pelo autor.

Como observado na Figura 3, pelo *ScatterPlot* é possível avaliar a dispersão dos valores por ano, destacando os pontos cianos, que representam os valores mais agrupados e perto da mediana, e os pontos vermelhos identificados como valores muito acima da mediana.

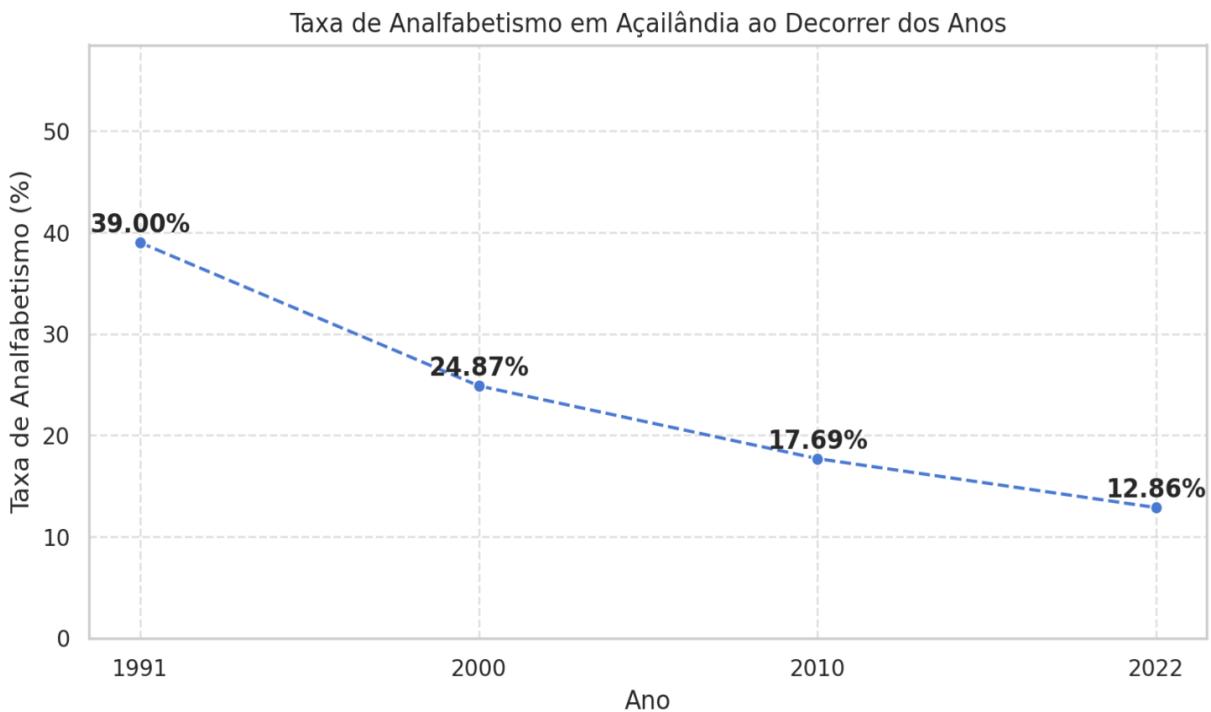
O *BoxPlot* fornece um resumo estatístico da distribuição de cada ano, dividindo os dados em quatro quartis, onde o primeiro e o quarto representam 25% da amostra, o segundo 50% (a mediana) e o terceiro 75%, além disso, o Intervalo Interquartil (IIQ), representado pela linha preta no centro do retângulo azul, indica a dispersão dos 50% centrais dos dados. A análise do Box Plot revela uma clara tendência de queda na mediana da taxa de analfabetismo ao longo dos anos. O *ViolinPlot* evidencia a densidade e distribuição dos dados: as partes mais estreitas representam valores muito acima ou abaixo da mediana, enquanto a largura da “onda” indica a concentração de valores próximos, que neste caso demonstra redução da dispersão.

A partir dos resultados pode-se avaliar que os valores estão bem distribuídos, não sendo necessária a aplicação de padronização ou normalização, recursos que só seriam exigidos em algoritmos de aprendizado de máquina sensíveis a *outliers*. Observa-se ainda que a mediana e a variação diminuem ao longo dos anos, refletindo a tendência de crescimento da população alfabetizada.

2.2.5 PROGRESSÃO PERCENTUAL DA TAXA DE ANALFABETISMO AO LONGO DOS ANOS

Entre 1991 e 2022, os dados sobre alfabetização no Maranhão foram organizados por município e ano, o que possibilita analisar a evolução das taxas de analfabetismo, identificar tendências e examinar possíveis causas, além de subsidiar estratégias de mitigação desse problema educacional. Para isso, a aplicação Web oferece uma seção onde o usuário pode filtrar municípios e visualizar, em figuras de linha, a progressão das taxas ao longo dos anos, como observado na Figura 4. A linha permite analisar como está a tendência da taxa de analfabetismo no município específico, neste caso, o município de Açailândia tende a ter uma baixa na população não alfabetizada, além disso, é possível identificar que no ano de 1991 foi o período em que mais teve casos e o ano de 2022 foi o ano em que apresentou a menor taxa.

Figura 4 – Progressão da taxa de analfabetismo por município através de Figura de linha.



Fonte: Elaborado pelo autor.

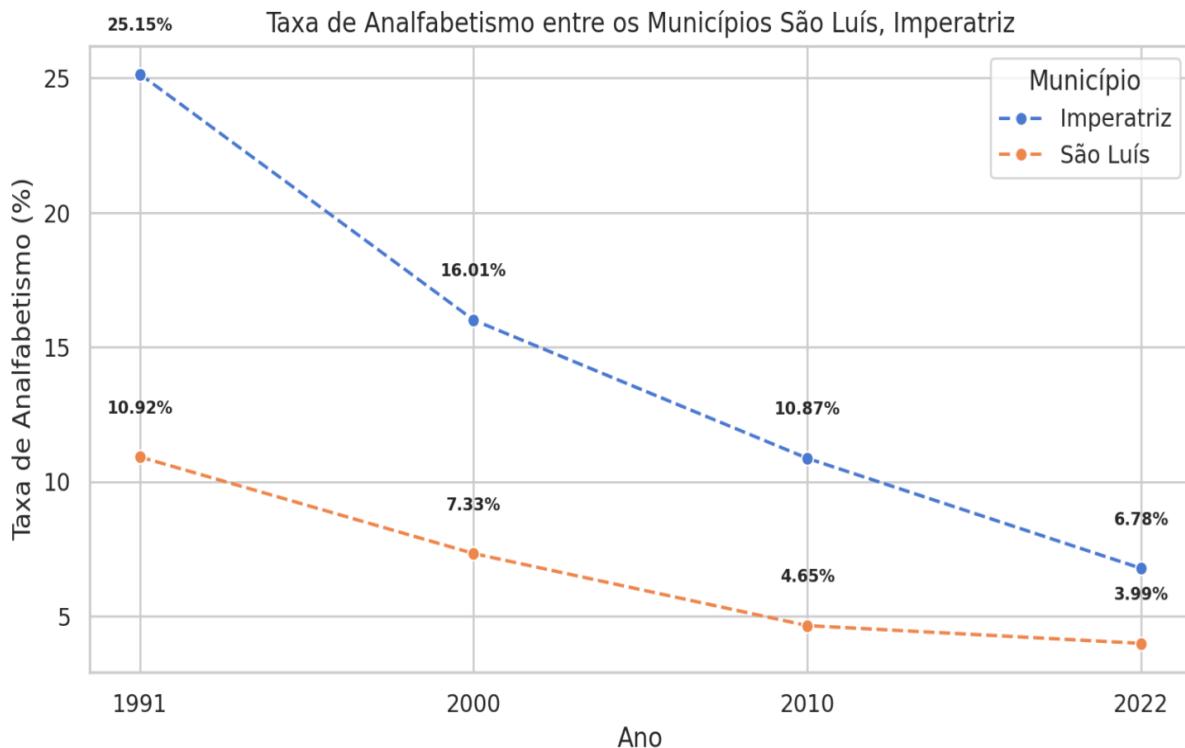
2.2.6 PROGRESSÃO DA TAXA DE ANALFABETISMO EM MUNICÍPIOS SELECIONADOS

A progressão da taxa de analfabetismo, foi elaborada uma visualização que permita a comparação entre dois ou mais municípios, possibilitando o estudo das variações ao longo dos anos e se tem alguma correlação desses dados municipais.

O *Dashboard* interativo agora inclui uma sessão específica que possibilita ao usuário selecionar os municípios desejados, e a partir disso uma figura de linhas coloridas é gerada automaticamente, facilitando a compreensão da progressão, como é demonstrado na Figura 5. A análise da progressão da taxa de analfabetismo revela tendência de queda em São Luís e Imperatriz. Em 1991, Imperatriz apresentava 25,15%, mais que o dobro da capital, que registrava 10,92%. Já em 2022, os índices caíram para 6,78% e 3,99%, respectivamente. Essa redução expressiva evidencia avanços na educação básica ao longo dos

anos e reforça a necessidade de analisar a tendência de queda das taxas de analfabetismo no Maranhão como um todo.

Figura 5 - Taxa de Analfabetismo entre os Municípios São Luís e Imperatriz, 1991-2022.



Fonte: Elaborado pelo autor.

3 CONCLUSÃO

Este estudo desenvolveu um modo de avaliação interativa para monitorar e analisar indicadores sociais do Ipeadata, com foco nas taxas de analfabetismo no Maranhão entre 1991 e 2022. A metodologia adotada utilizou a linguagem Python e suas bibliotecas, possibilitando análise estatística e visualização dos resultados. Durante o pré-processamento, foram removidos registros inconsistentes ou ausentes. Os indicadores foram organizados em percentual médio de alfabetizados e não alfabetizados, a partir da média dos percentuais anuais. Essa abordagem evidenciou a redução do analfabetismo no período analisado; contudo, alguns municípios ainda registram índices elevados, chegando a 47,49% de não alfabetizados. Os resultados

demonstram que a combinação entre pré-processamento adequado e o uso eficiente do Python constitui uma ferramenta eficaz para o monitoramento de indicadores sociais, oferecendo uma base consistente para análises futuras e aplicável a diferentes contextos e regiões.

4 REFERÊNCIAS

HUNTER, John D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, v. 9, n. 3, p. 90-95, 2007.

MCKINNEY, Wes. *pandas: powerful Python data analysis toolkit*. 2021. Disponível em: <https://pandas.pydata.org/pandas-docs/version/1.3/pandas.pdf>. Acesso em: 19 set. 2025.

WASKOM, Michael L. Seaborn: statistical data visualization. *Journal of Open-Source Software*, v. 6, n. 60, p. 3021, 2021.

PAIVA, Felipe Andrade de. *Business Intelligence e gestão de negócio: estudo de caso sobre a aplicação de Business Intelligence na obtenção de informações*. 2023. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Computação) – Escola Politécnica e de Artes, Pontifícia Universidade Católica de Goiás, Goiânia, 2023.

DUARTE, Alan. *Tudo o que você precisa saber para ingressar na área de BI* [e-book]. Entendendo BI, 2020. Disponível em: <https://entendendobi.com/wp-content/uploads/2020/12/E-Book-Business-Intelligence.pdf>

SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. *Business Intelligence: uma abordagem gerencial*. 3. ed. Upper Saddle River: Pearson, 2010

IBGE – Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional por Amostra de Domicílios Contínua: Educação 2023. Rio de Janeiro: IBGE, 2023.

UNESCO. Global report on adult learning and education (GRALE 4). Paris: UNESCO, 2019. I

PEA – Instituto de Pesquisa Econômica Aplicada. Ipeadata: base de indicadores sociais e econômicos. Brasília: IPEA, 2022. Disponível em: .

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, v. 36, n. 4, p. 1165-1188, 2012.

OLIPHANT, T. E. Python for scientific computing. *Computing in Science & Engineering*, v. 9, n. 3, p. 10-20, 2007.

CAMPANTE, F.; GONÇALVES, L.; GONÇALVES, R. Ciência de dados aplicada às

políticas públicas. São Paulo: Atlas, 2024.

ROMERO, C.; VENTURA, S. Educational data mining and learning analytics: an updated review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, v. 14, n. 2, p. e1502, 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. SÃO JOÃO DO CARU (MA) — *Panorama*. Disponível em: <https://cidades.ibge.gov.br/brasil/ma/sao-joao-do-caru/panorama>. Acesso em: 20 set. 2025.

BORGES, L. E. Python para Desenvolvedores. 3. ed. Rio de Janeiro: Novatec, 2014.

McKINNEY, Wes. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 1. ed. O'Reilly Media, 2010.

HUNTER, J. D. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, v. 9, n. 3, p. 90–95, 2007.

IPEA. *Ipeadata: base de dados macroeconômico, regional e social*. Disponível em: <https://www.ipeadata.gov.br/Default.aspx>. Acesso em: 20 set. 2025.

ASIMOV ACADEMY. *Streamlit: guia completo*. Disponível em: <https://hub.asimov.academy/blog/streamlit-guia-completo/>. Acesso em: 18 set. 2025.

ENGBD. *Preparação de dados: o que é*. Disponível em: <https://blog.engdb.com.br/preparacao-de-dados-o-que-e/>. Acesso em: 18 set. 2025.