

ANÁLISE DE DADOS EDUCACIONAIS: DESEMPENHO DE ESTUDANTES EM MATEMÁTICA E REDAÇÃO

*ANALYSIS OF EDUCATIONAL DATA: PERFORMANCE OF STUDENTS
IN MATHEMATICS AND ESSAYS*

*ANÁLISE DE DADOS EDUCACIONAIS: DESEMPENHO DE
ESTUDANTES EM MATEMÁTICA E REDAÇÃO*

Artur Guilherme Acácio Cipriano

Andrey Pereira Tavares

Francisco Cleiton Vieira Gomes

Geovanna Miranda Rodrigues

Joana Victória Bezerra Venâncio

Lucas Alves dos Santos

Juvenaldo Florentino Canja

Resumo: O presente projeto tem como objetivo realizar uma análise detalhada de dados de escolas e estudantes, observando o desempenho nas disciplinas de Matemática e Redação. A proposta surgiu a partir da necessidade de compreender melhor os fatores que influenciam o aprendizado e de como as ferramentas tecnológicas podem auxiliar nesse processo. Por meio da captação e tratamento dos dados, foram aplicadas técnicas de análise e visualização utilizando a linguagem Python e as bibliotecas pandas, numpy, matplotlib. Essas ferramentas permitiram transformar dados brutos em informações organizadas, claras e visuais. O objetivo central da pesquisa é identificar padrões de aprovação e reprovação, comparando o desempenho dos estudantes e evidenciando possíveis diferenças entre escolas e disciplinas. Através dessas análises, foi possível compreender de forma mais ampla o panorama educacional, contribuindo para o desenvolvimento de estratégias pedagógicas mais eficazes e baseadas em evidências. Esse trabalho também reforça a importância da união entre educação e tecnologia, mostrando como a programação pode ser uma aliada poderosa para compreender o comportamento estudantil e promover melhorias no sistema educacional.

Palavras-chaves: Análise de dados, Desempenho escolar, Python, Educação.

Abstract: The objective of this project is to carry out a detailed analysis of data from schools and students, observing their performance in the disciplines of Mathematics and Writing. The purpose arose from the need to better understand the factors that influence learning and how technological tools can assist this process. By capturing and processing data, analysis and visualization techniques are applied using the Python language and the pandas, numpy, and matplotlib libraries. These tools will allow you to transform raw data into organized, clear and visual information. The central objective of the research is to identify patterns of passing and failing, comparing the performance of two students and evidencing possible differences between schools and disciplines. Through these analyses, it was possible to understand the educational panorama in a broader way, contributing to the development of more effective and evidence-based pedagogical strategies. This work also reinforces the importance of the

Artur Guilherme Acácio Cipriano, Andrey Pereira Tavares, Francisco Cleiton Vieira Gomes, Geovanna Miranda Rodrigues, Joana Victória Bezerra Venâncio, Lucas Alves dos Santos, Juvenaldo Florentino Canja

union between education and technology, showing how programming can be a powerful ally in understanding student behavior and promoting improvements in the educational system.

Keywords: Data analysis, School performance, Python, Education.

Resúmen: El presente proyecto tem como objetivo realizar un análisis detallado de datos de escuelas y estudiantes, observando o desempenho nas disciplinas de Matemática y Redação. A proposta surgiu a partir de la necesidad de comprender mejor los factores que influyen en el aprendiz y de como herramientas tecnológicas que pueden auxiliar ese proceso. Por medio de la captación y el tratamiento de datos, se aplicaron técnicas de análisis y visualización utilizando el lenguaje Python y las bibliotecas pandas, numpy, matplotlib. Estas herramientas permitirán transformar datos brutos en información organizada, clara y visual. El objetivo central de la investigación es identificar padrões de aprovação y reprovação, comparando el desempeño de dos estudiantes y evidenciando possíveis diferenças entre escuelas y disciplinas. A través de estos análisis, podremos comprender de forma más amplia el panorama educativo, contribuyendo para el desarrollo de estrategias pedagógicas más eficaces y basadas en evidencias. Este trabajo también refuerza la importancia de la unión entre la educación y la tecnología, mostrando como un programa puede ser una aliada poderosa para comprender el comportamiento estudiantil y promover mejoras en el sistema educativo.

Palabras clave: Análisis de datos, Desarrollo escolar, Python, Educación.

1 Introdução

A análise de dados educacionais tem se mostrado uma ferramenta essencial para compreender o desempenho dos estudantes e a eficiência das instituições de ensino. No contexto atual, em que a tecnologia faz parte de praticamente todos os setores da sociedade, a educação também vem se beneficiando dessas inovações. Por meio de dados sobre notas de Matemática e Redação, é possível identificar padrões de aprovação e reprovação, entender onde estão as principais dificuldades dos alunos e comparar o desempenho entre diferentes escolas. Essas informações podem servir de base para o desenvolvimento de políticas educacionais mais assertivas. Neste projeto, o uso da linguagem Python foi fundamental.

Ela permitiu automatizar análises, realizar cálculos de forma prática e criar gráficos explicativos, que facilitaram muito a interpretação dos resultados. A visualização dos dados se tornou um ponto-chave para que o grupo conseguisse enxergar o panorama completo do desempenho estudiantil. Dessa forma, a aplicação da tecnologia se mostra uma aliada importante da educação moderna, contribuindo para uma tomada de decisão mais elaborada e para o aprimoramento contínuo dos processos de ensino e aprendizagem.

2 Revisão Bibliográfica

Com o avanço da tecnologia, a educação passou a ter acesso a ferramentas computacionais que ajudam não apenas na gestão de instituições, mas também na análise do processo de aprendizado dos alunos. Segundo Silva (2022), a ciência de dados aplicada à educação é capaz de identificar lacunas de aprendizado e direcionar estratégias pedagógicas personalizadas.

Já Oliveira (2023) destaca que bibliotecas como pandas e numpy possibilitam uma manipulação eficiente de grandes volumes de dados, enquanto matplotlib e seaborn facilitam a criação de gráficos visuais e informativos. Essas tecnologias tornam o processo de análise mais acessível, interativo e produtivo.

Assim, a integração entre educação e análise de dados representa uma grande oportunidade de inovação, permitindo compreender o desempenho estudantil de forma mais profunda e contribuir para o aprimoramento da gestão escolar.

3 Materiais e Métodos

A pesquisa foi desenvolvida utilizando um conjunto de dados contendo as notas de estudantes em Matemática e Redação, além das informações sobre as escolas de origem. Todo o processo foi conduzido em Python, empregando bibliotecas específicas para cada função: pandas (leitura e manipulação dos dados); numpy (cálculos de médias e classificações); e matplotlib (criação de gráficos comparativos e de visualização).

O processo foi dividido em quatro etapas principais:

Importação e limpeza dos dados

Primeiramente, os dados foram importados para o ambiente Python. Utilizamos o pandas para ler o arquivo CSV e organizar as informações em formato de tabela.

Em seguida, realizamos a limpeza dos dados, removendo valores nulos, corrigindo erros de digitação e padronizando os nomes. Essa etapa foi essencial para garantir que os resultados finais fossem precisos e confiáveis.

1- Códigos e gráficos

```
import pandas as pd

dados = {
    'ID_Escola': [0, 1, 2, 3, 4],
    'Nome_Escola': ['Escola A', 'Escola B', 'Escola C', 'Escola D', 'Escola E'],
    'Tipo_Escola': ['Publica', 'Publica', 'Particular', 'Publica', 'Particular'],
    'Numero_Alunos': [2917, 2949, 1761, 4635, 1468],
    'Orcamento_Anual': [1910635, 1884411, 1056600, 3022020, 917500]
}

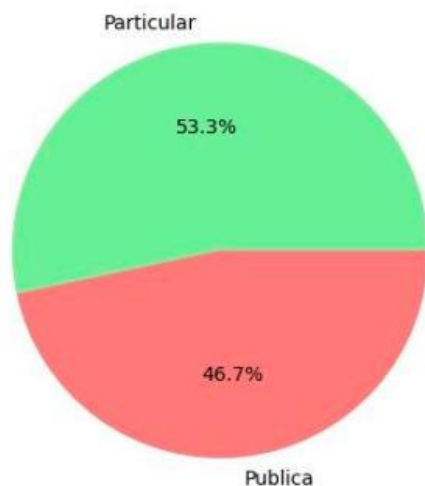
df = pd.DataFrame(dados)

df['Tipo_Escola'].value_counts().plot(kind='pie', autopct='%1.1f%%',
colors=['lightgreen', 'lightcoral'])
plt.title('Relação entre Escolas Públicas e Particulares')
plt.ylabel('')
plt.show()
```

Cálculo das médias e situação de aprovação

Com os dados limpos, partimos para os cálculos. Usando o numpy, criamos fórmulas para calcular as médias de cada aluno nas disciplinas de Matemática e Redação. Depois, aplicamos um critério para definir quem estava aprovado ou reprovado, adotando como média mínima o valor de 6,0. Essa parte foi responsável por transformar os dados brutos em informações úteis e interpretáveis.

Relação entre Escolas Públicas e Particulares



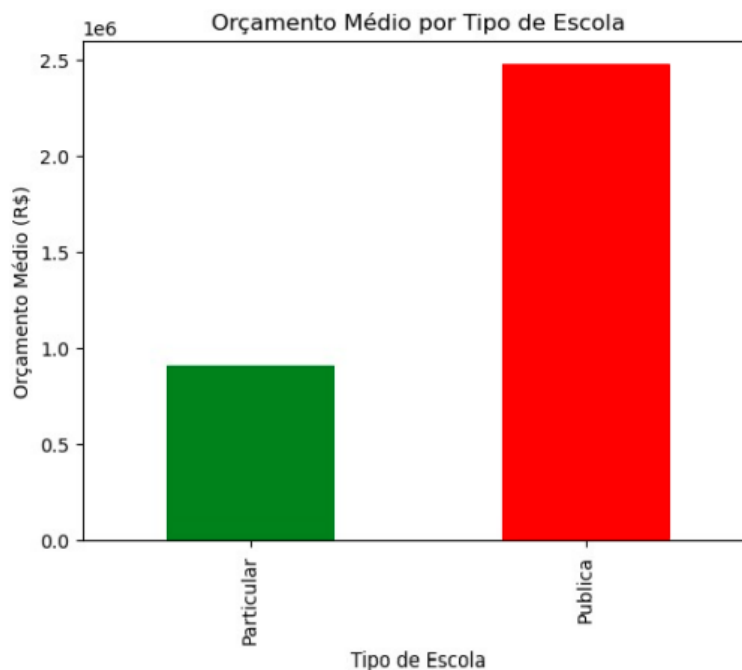
```
media_orcamento = df.groupby('Tipo_Escola')['Orcamento_Anual'].mean()

media_orcamento.plot(kind='bar', color=['green', 'red'])
plt.title('Orçamento Médio por Tipo de Escola')
plt.xlabel('Tipo de Escola')
plt.ylabel('Orçamento Médio (R$)')
plt.show()
```

Geração de gráficos comparativos

A próxima etapa foi criar gráficos para visualizar os resultados. Utilizamos o matplotlib para montar gráficos de barras e colunas que mostravam comparações entre escolas e disciplinas.

Essa parte foi muito interessante, pois permitiu observar, de maneira clara e intuitiva, quais instituições apresentaram melhor desempenho e onde estavam os maiores índices de reprovação.



```
import pandas as pd
import matplotlib.pyplot as plt

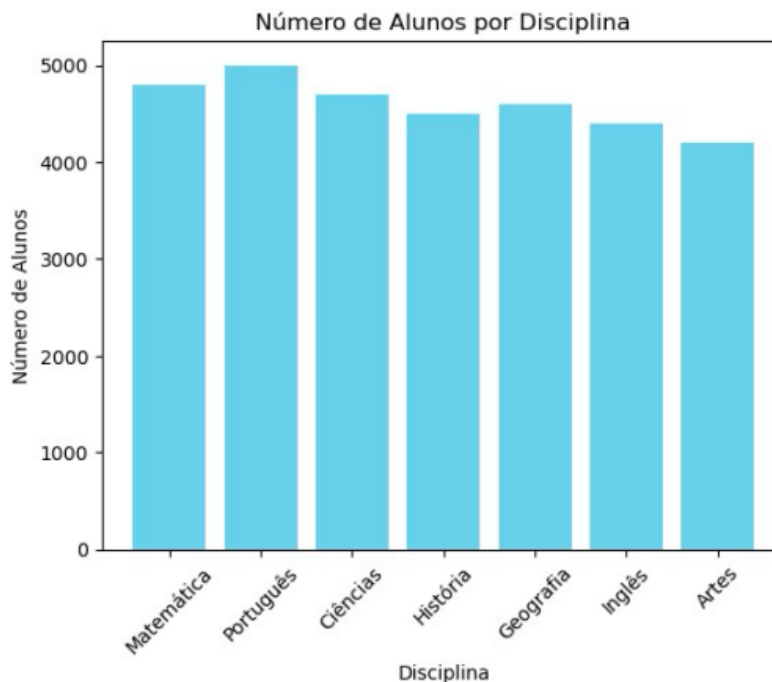
# Exemplo de base com disciplinas e número de alunos
df_disciplinas = pd.DataFrame({
    'Disciplina': ['Matemática', 'Português', 'Ciências', 'História', 'Geografia',
                  'Inglês', 'Artes'],
    'Numero_Alunos': [4800, 5000, 4700, 4500, 4600, 4400, 4200]
})

# Gráfico de barras
plt.bar(df_disciplinas['Disciplina'], df_disciplinas['Numero_Alunos'],
        color='skyblue')
plt.title('Número de Alunos por Disciplina')
plt.xlabel('Disciplina')
plt.ylabel('Número de Alunos')
plt.xticks(rotation=45)
plt.show()
```

Respostas às perguntas de análise

Por fim, o grupo respondeu a sete perguntas de análise propostas na atividade. Cada integrante ficou responsável por uma parte, o que garantiu que todos participassem do processo.

Essa divisão também proporcionou uma troca de aprendizados muito produtiva, pois cada aluno pôde contribuir com sua visão e interpretação dos resultados.



Os resultados obtidos foram bastante reveladores. A análise dos dados mostrou que o desempenho médio variou significativamente entre as escolas analisadas. Em geral, os alunos apresentaram um rendimento melhor em Redação do que em Matemática, o que pode estar relacionado a diferentes métodos de ensino ou ao perfil de aprendizado dos estudantes.

A taxa de aprovação em cada disciplina foi calculada e comparada, revelando padrões que ajudam a compreender o cenário educacional de forma mais ampla. Além disso, observamos uma correlação positiva entre as notas de Matemática e Redação, ou seja, alunos que se saíram bem em uma disciplina também tendiam a ter bom desempenho na outra.

Os gráficos criados facilitaram a visualização dos resultados e tornaram a análise mais dinâmica. Eles mostraram, de forma prática, as diferenças de desempenho entre escolas e ajudaram a identificar quais instituições precisavam de mais atenção pedagógica.

Considerações finais

A realização deste projeto demonstrou como o uso da programação e da análise de dados pode ser uma poderosa ferramenta para compreender o desempenho escolar. Através da linguagem Python, foi possível transformar informações numéricas em análises concretas, com gráficos e indicadores que ajudam a visualizar a realidade educacional de forma clara e objetiva.

O projeto reforçou a importância da tecnologia aplicada à educação, mostrando que dados bem interpretados podem auxiliar escolas e professores a tomar decisões

mais conscientes e assertivas. Além disso, o trabalho despertou em nós, estudantes, o interesse por unir diferentes áreas do conhecimento nesse caso, a lógica da programação com o raciocínio analítico da matemática e a interpretação crítica da educação.

Essa experiência foi extremamente enriquecedora, pois mostrou na prática que a tecnologia, quando usada com propósito, tem o poder de transformar realidades e gerar impacto positivo na sociedade.

Referências Bibliográficas

OLIVEIRA, M. A. **Introdução à Análise de Dados com Python**. Rio de Janeiro: LTC, 2023.

SILVA, J. R. **Ciência de Dados na Educação: Aplicações Práticas**. São Paulo: Atlas, 2022.

UNIFANOR WYDEN. **Modelo de Projeto de Extensão**. Fortaleza: 2025.

Editorial

Editor-chefe:

Vicente de Paulo Augusto de Oliveira Júnior
Centro Universitário Fanor Wyden
vicente.augusto@wyden.edu.br

Editor responsável:

Raimundo Rigoberto B. Xavier Filho
Centro Universitário Fanor Wyden
raimundo.bfilho@wyden.edu.br

Autor(es):

Artur Guilherme Acácio Cipriano
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*


Andrey Pereira Tavares
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Francisco Cleiton Vieira Gomes
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Geovanna Miranda Rodrigues Centro
Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Joana Victória Bezerra Venâncio
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Lucas Alves dos Santos
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Juvenaldo Florentino Canja 
Centro Universitário Fanor Wyden
virinoalyne@gmail.com
Contribuição: *Contextualização, escrita e desenvolvimento (orientação)*

Submetido em:

Aprovado em:

Publicado em:

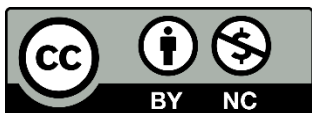
DOI:

Financiamento:

Como citar este trabalho:

(ABNT)

(APA)



© 2025 Revista de Educação à Distância. Centro Universitário Fanor Wyden – UniFanor Wyden. Este trabalho está licenciado sob uma licença *Creative Commons* Atribuição - Não comercial - Compartilhar 4.0 Internacional CC-BY NC 4.0 Internacional).