

Relatório Técnico – Mini Projeto Web:

ODS 4 - Educação de Qualidade

Autores: Caio Barbosa, Davi Vaz e Lucas Almeida **Data:** Outubro de 2025 **Fonte dos Dados:** World Bank SDG Atlas 2023 – Goal 4: Quality Education

1. Objetivo Geral

Text

O objetivo deste projeto foi desenvolver uma aplicação web interativa para analisar e visualizar dados relacionados ao **ODS 4 – Educação de Qualidade**, com ênfase nos impactos da pandemia de COVID-19 na aprendizagem dos estudantes. O projeto visa demonstrar como o uso de **Pandas**, **Plotly** e **Streamlit** pode ser aplicado na análise exploratória de dados e na criação de dashboards informativos e acessíveis.

2. Tecnologias Utilizadas

A tabela a seguir detalha as tecnologias empregadas no desenvolvimento do projeto, categorizando-as por sua função principal:

Categoria	Ferramenta	Função Principal
Linguagem	Python	Manipulação e análise de dados
Biblioteca de Análise	Pandas	Leitura, limpeza e transformação de dados
Visualização	Plotly Express / Graph Objects	Criação de gráficos interativos
Framework Web	Streamlit	Desenvolvimento da aplicação web
Hospedagem	Google Colab + Pyngrok	Execução e compartilhamento via link público

3. Fonte e Estrutura dos Dados

Os dados utilizados foram extraídos do **World Bank SDG Atlas 2023**, especificamente do arquivo `Chart6_student_learning_impacts.csv`. Este dataset foca nos impactos da pandemia na aprendizagem dos estudantes.

As principais colunas do dataset são:

Coluna	Descrição
Country	País analisado
Closure length weeks	Duração do fechamento das escolas (em semanas)
Average learning losses (SD)	Perdas médias de aprendizagem (em desvio padrão)
share_closed	Proporção de escolas fechadas
years_lost	Anos de aprendizagem perdidos
Source	Fonte do estudo
Notes	Observações complementares

Não foram encontrados valores ausentes, e o dataset contém dados de 19 países.

4. Etapas de Desenvolvimento

4.1. Carregamento e Pré-processamento dos Dados

O arquivo CSV foi carregado diretamente no ambiente do Google Colab. A função `load_data()` com cache (`@st.cache_data`) foi utilizada para garantir um melhor desempenho ao recarregar a aplicação, otimizando o processo de leitura e preparação dos dados.

4.2. Análise Exploratória de Dados (EDA)

Foram realizadas análises estatísticas descritivas e gráficas para compreender o comportamento das variáveis e identificar padrões. Os principais achados incluem:

- **Duração média de fechamento das escolas:** 26,3 semanas.
- **Perda média de aprendizagem:** -0.18 SD (desvio padrão).
- **Correlação forte negativa (-0.76):** Observou-se uma correlação significativa entre o tempo de fechamento das escolas e as perdas de aprendizagem, indicando que períodos mais longos de fechamento estão associados a maiores perdas.

5. Estrutura da Aplicação Streamlit

A aplicação web desenvolvida foi organizada em quatro seções principais, acessíveis através de um menu lateral, proporcionando uma navegação intuitiva:

5.1. Visão Geral

Esta seção oferece uma introdução ao ODS 4, uma descrição detalhada do dataset e as métricas principais do estudo, tais como:

- Países analisados
- Média de semanas de fechamento das escolas
- País com a maior perda de aprendizagem (México)
- Média de anos de aprendizagem perdidos

Inclui também uma tabela completa dos dados para consulta.

5.2. Análise Exploratória

Nesta seção, são exibidos:

- Estatísticas descritivas (`df.describe()`) do dataset.
- Histogramas das variáveis “Semanas de Fechamento” e “Perdas de Aprendizagem” , para visualizar a distribuição dos dados.
- Matriz de correlação entre variáveis numéricas, gerada com `px.imshow()` , para identificar relações entre as variáveis.

5.3. Visualizações Interativas

Esta parte da aplicação contém visualizações dinâmicas que permitem uma exploração aprofundada dos dados:

- **Gráfico de dispersão interativo:** Apresenta a relação entre o tempo de fechamento das escolas e as perdas médias de aprendizagem, com uma linha de tendência via OLS (Ordinary Least Squares).
- **Gráfico de barras ordenado:** Exibe as perdas de aprendizagem por país, facilitando a comparação.
- **Slider interativo:** Permite filtrar países com base na duração do fechamento das escolas, oferecendo flexibilidade na análise.

5.4. Insights e Conclusões

Esta seção finaliza a análise, apresentando:

- **Principais insights analíticos:** Destacando a correlação negativa entre fechamento e perdas, e o México como um *outlier* com -0.55 SD e 48 semanas de fechamento.
- **Variação significativa entre países:** Enfatizando a diversidade dos impactos observados.
- **Recomendações para políticas públicas:** Sugestões baseadas nos dados para mitigar os efeitos negativos na educação.
- **Limitações do estudo e próximos passos:** Discussão sobre as restrições da pesquisa e direções futuras.

6. Execução no Google Colab

A aplicação foi adaptada para ser executada no Google Colab, utilizando o **Pyngrok** para permitir o acesso via um link público. O processo de configuração envolve:

```
from pyngrok import ngrok
ngrok.set_auth_token("seu_token") # Substitua "seu_token" pelo seu token de
autenticação do ngrok
public_url = ngrok.connect(8501, "http").public_url
```

O aplicativo é iniciado com o seguinte comando no ambiente do Colab:

```
!streamlit run streamlit_app_colab.py --server.port 8501 --server.headless
true
```

Após a inicialização, o link público para acessar a aplicação é exibido automaticamente.

7. Resultados Obtidos

O projeto resultou em:

- Uma visualização clara e interativa dos impactos da pandemia na educação.
- A identificação de padrões e correlações de forma intuitiva.
- Uma interface funcional e organizada, acessível a usuários sem conhecimento técnico prévio.

8. Conclusões

Este projeto demonstrou efetivamente como a combinação de **Pandas** para manipulação de dados e **Streamlit** para desenvolvimento de dashboards pode facilitar a criação de ferramentas interativas para a análise de dados relacionados aos Objetivos de Desenvolvimento Sustentável (ODS). A abordagem adotada não apenas torna a exploração educacional acessível e visualmente atrativa, mas também promove uma compreensão mais profunda dos desafios enfrentados pela educação global durante a pandemia de COVID-19.

9. Próximos Passos Sugeridos

Para aprimorar e expandir o projeto, sugere-se os seguintes próximos passos:

- **Expandir o dataset:** Incluir novos indicadores, como proficiência em leitura e matemática, para uma análise mais abrangente.
- **Implementar filtros avançados:** Adicionar filtros por região, gênero ou faixa etária, permitindo análises mais segmentadas.
- **Criar modelos preditivos:** Desenvolver modelos para prever a recuperação de aprendizagem ou o impacto de futuras crises.

- **Implantar a aplicação permanentemente:** Realizar a implantação em plataformas como Streamlit Cloud, Render ou Hugging Face Spaces para disponibilidade contínua.