

INTRODUÇÃO À VISUALIZAÇÃO DE DADOS COM PYTHON

SILVA, A. G.¹; LINS, A. A.¹; MELO, L. F. S.¹; SILVA, S. G. L.; PUCA, C. V. L.¹; SANTOS, L. B.²

¹PET PAESPE Ciência e Tecnologia, UFAL, Campus A. C. Simões, ²Tutor do Grupo PET PAESPE Ciência e Tecnologia, UFAL, Campus A. C. Simões

E-mail: aristides.silva@ctec.ufal.br, petct.ufal@gmail.com

TIPO DA PROPOSTA: Oficina

DURAÇÃO: 120 minutos

QUANTIDADE DE OFERTA: Será ministrada uma única vez.

NÚMERO DE VAGAS: A oficina depende da disponibilidade de computadores de um laboratório de informática, então estimamos 20 participantes.

1. MINISTRANTES

- Aristides Guilherme da Silva: discente não-bolsista do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: aristides.silva@ctec.ufal.br.
- Arestides Alves Lins: discente colaborador do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: arestides.lins@ctec.ufal.br.
- Luis Fernando Silva Melo: discente bolsista do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: luis.melo@ctec.ufal.br.
- Lucas Gabriel Silva Santos: discente bolsista do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: lucas.santos@ctec.ufal.br.
- Cleydson Victor de Lima Puça: discente bolsista do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: cleydson.puca@ctec.ufal.br.
- Luciano Barbosa dos Santos: Docente tutor do PET PAESPE Ciência e Tecnologia - Universidade Federal de Alagoas. E-mail: lbsantos@ctec.ufal.br

2. PROPOSTA

Eixo Temático: Engenharias.

Em um mundo repleto de informações, transformar dados em conhecimento visual é importante para a formação e para o exercício profissional. De acordo com Casagrande e Silva (2024), a análise de dados tem um grande impacto em diversas áreas do nosso cotidiano. O crescimento dos dados digitais e o avanço das tecnologias tem permitido análises mais detalhadas e sofisticadas de grandes volumes de informações, superando as ferramentas estatísticas tradicionais. Segundo Zhang, Wolfram e Ma (2023), essas novas tecnologias possibilitam o processamento de dados em escalas e velocidades antes inimagináveis, fornecendo uma base mais sólida para modelos complexos que apoiam decisões estratégicas. Para isso, propõe-se desenvolver atividades práticas com dados, a fim de explorar padrões, comunicar ideias e compreender a realidade por meio de gráficos e análises simples utilizando os fundamentos da visualização de dados em Python, com auxílio das bibliotecas Pandas, Matplotlib e NumPy.

3. OBJETIVO GERAL

Capacitar os participantes a compreender e aplicar os conceitos básicos de visualização de dados utilizando Python e as bibliotecas Pandas, NumPy e Matplotlib, promovendo habilidades práticas para manipular, analisar e representar dados de forma clara e acessível. Busca-se fomentar o pensamento crítico e a inclusão digital, mostrando como a ciência de dados pode ser uma ferramenta transformadora na formação cidadã, capaz de apoiar decisões e promover a comunicação efetiva de informações complexas.

4. METODOLOGIA

A oficina será centrada em atividades práticas e interativas, conduzidas por meio do ambiente virtual do Google Colaboratory (GOOGLE, 2025), em Python, com foco na construção do conhecimento por meio da experimentação. O primeiro momento será dedicado à adaptação dos participantes ao ambiente de desenvolvimento e à importação das bibliotecas necessárias. Em seguida, será realizada a leitura de um conjunto de dados em formato CSV, abordando a importância da estrutura dos dados e a necessidade de conversão de formatos, quando necessário. Após a leitura dos dados, os participantes serão orientados a renomear colunas para facilitar a manipulação, aplicar funções descritivas para sumarizar as informações e compreender as principais características do conjunto. Serão apresentados os conceitos de séries

e DataFrames, destacando suas diferenças e aplicações. Em seguida, será realizada a limpeza dos dados, incluindo a identificação e correção de erros de tipografia e o tratamento de valores ausentes. Logo após, serão explorados diferentes tipos de gráficos, como dispersão, histogramas, boxplots e barras, discutindo a escolha mais adequada para cada tipo de dado e objetivo de análise. Os gráficos serão construídos com a biblioteca matplotlib, com foco na personalização: ajuste de escalas, organização dos dados no espaço gráfico, modificação de cores e elementos visuais para aprimorar a comunicação das informações. Por fim, será realizada uma análise estatística descritiva, com medidas de tendência central e dispersão, além da visualização de séries temporais.

5. CONTEÚDO PROGRAMÁTICO

Tabela 1. Conteúdo Programático

Conteúdos Propostos	Tempo Médio (min)
Apresentação da proposta da oficina e da plataforma Google Colab	10
Importação e explicação sobre as bibliotecas (Pandas, Numpy, Matplotlib)	10
Leitura e tratamento de dados em .CSV	30
Criação de gráficos para descrever os dados	40
Geração e interpretação de medidas de tendência central	20
Finalização da oficina com resumo do que foi ministrado	10

Fonte: Autores, 2025.

6. MATERIAIS

Necessita de um laboratório de informática que possua computadores com acesso a internet.

Tabela 2. Materiais utilizados para oficina/minicurso

Tipo de Material	Quantidade
Computadores	20
Projektor	1

Fonte: Autores, 2025.

5. REFERÊNCIAS

- CASAGRANDE, Vitor Moreira; SILVA, Thiago Pereira da. Uma introdução à análise de dados usando pandas, matplotlib e seaborn. In: **ESCOLA REGIONAL DE INFORMÁTICA DE MATO GROSSO (ERI-MT)**, 13. , 2024, Alto Araguaia/MT. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 196-203. ISSN 2447-5386. DOI: <https://doi.org/10.5753/eri-mt.2024.245927>.
- GOOGLE. **Google Colaboratory**. 2025. Disponível em: <https://colab.research.google.com/>. Acesso em: 12 jul. 2025.
- HARRIS, C. R.; MILLMAN, K. J.; van der WALT, S. J. Array programming with NumPy. **Nature**, v. 585, n. 7825, p. 357–362, 2020.
- HUNTER, J. D. Matplotlib: A 2D graphics environment. **Computing in Science & Engineering**, v. 9, n. 3, p. 90-95, 2007. DOI: 10.1109/MCSE.2007.55.
- MCKINNEY, Wes. **Pandas**. Disponível em: <https://pandas.pydata.org/>. Acesso em: 12 jul. 2025.
- PYTHON SOFTWARE FOUNDATION. **Python Language Site: Documentation**, 2020. Disponível em: . Acesso em: 12 jul. 2025.
- ZHANG, Jin; WOLFRAM, Dietmar; MA, Feicheng. The impact of big data on research methods in information science. **Data and Information Management**, v. 7, n. 2, 2023.