

# **DOCUMENTAÇÃO DO PROJETO**

## **PYTHON FOR DATA SCIENCE - NOTEBOOKLM**

Autor: **Diogo Zoboli**

*Técnico em Ciência de Dados | Cursando Eng. de Software |  
Medalhista Nacional | Python • SQL • Power BI*

Linkedin: [linkedin.com/in/zobolidiogo/](https://linkedin.com/in/zobolidiogo/)

GitHub: [github.com/zobolidiogo](https://github.com/zobolidiogo)

## **GLOSSÁRIO**

### **1. INTRODUÇÃO**

- 1.1. CONTEXTO DO DESAFIO
- 1.2. JUSTIFICATIVA DO TEMA

### **2. OBJETIVOS**

- 2.1. OBJETIVO GERAL
- 2.2. OBJETIVOS ESPECÍFICOS

### **3. METODOLOGIA**

- 3.1. FERRAMENTAS UTILIZADAS
- 3.2. PROCESSO DE TRABALHO

### **4. CONCLUSÃO**

### **5. REFERÊNCIAS**

## **1. INTRODUÇÃO**

Este projeto foi desenvolvido como parte do Bootcamp "GenAI & Dados", oferecido pela plataforma DIO em parceria com o Bradesco. O desafio proposto consiste na criação de um miniguia de estudo utilizando o Google NotebookLM como ferramenta de aprendizagem ativa com Inteligência Artificial.

### **1.1. CONTEXTO DO DESAFIO**

O projeto seguiu as seguintes etapas:

- Definição de um tema relevante para o desenvolvimento profissional;
- Curadoria de 3-5 fontes abertas e confiáveis;
- Criação de notebook temático no NotebookLM;
- Elaboração de perguntas estratégicas e prompts exploratórios;
- Geração de resumos estruturados e glossário de conceitos;
- Documentação completa do processo;
- Publicação em repositório GitHub.

## **1.2. JUSTIFICATIVA DO TEMA**

A escolha do tema "Python para Análise de Dados" se baseia na formação acadêmica do autor como Técnico em Ciência de Dados pelo CEDUP Timbó, concluída em dezembro de 2025.

Durante o curso técnico, o estudante desenvolveu competências sólidas em modelagem de dados, SQL, Power BI e análise de dados, identificando o Python como uma ferramenta essencial para expandir suas capacidades técnicas. As bibliotecas Pandas, NumPy e Matplotlib são fundamentais para qualquer profissional da área de dados, tornando este tema altamente relevante para o desenvolvimento profissional e acadêmico do autor.

Atualmente cursando Bacharelado em Engenharia de Software (Uninter), o autor busca consolidar conhecimentos que integrem análise de dados com desenvolvimento de software, preparando-se para atuar como Analista ou Cientista de Dados.

## **2. OBJETIVOS**

### **2.1. OBJETIVO GERAL**

Consolidar conhecimentos sobre as principais bibliotecas Python para análise de dados (Pandas, NumPy e Matplotlib), utilizando IA Generativa como ferramenta de aprendizagem ativa.

### **2.2. OBJETIVOS ESPECÍFICOS**

- Compreender as diferenças e semelhanças entre Pandas e NumPy;
- Conhecer técnicas de manipulação e transformação de dados com Pandas;
- Aplicar computação numérica com NumPy;
- Criar visualizações com Matplotlib;
- Desenvolver conjunto de prompts reutilizáveis para futuras consultas;
- Construir glossário de conceitos essenciais para referência rápida.

### **3. METODOLOGIA**

#### **3.1. FERRAMENTAS UTILIZADAS**

##### **3.1.1. Google NotebookLM**

Plataforma de IA Generativa para organização de conhecimento.

##### **3.1.2. Google Docs**

Elaboração da documentação completa.

##### **3.1.3. GitHub**

Versionamento e compartilhamento do projeto.

#### **3.2. PROCESSO DE TRABALHO**

##### **3.2.1. Procura de Fontes**

Seleção criteriosa de 4 fontes abertas e confiáveis:

- Documentações oficiais (Pandas, NumPy);
- Tutorial técnico especializado (Real Python - Matplotlib);
- Livro open-source de referência (Python Data Science Handbook).

##### **3.2.2. Criação do Notebook**

- Adição das fontes ao notebook;
- Exploração inicial do material;
- Identificação de conceitos-chave.

##### **3.2.3. Elaboração de Perguntas Estratégicas**

- Formulação de perguntas exploratórias;
- Teste de variações de prompts;
- Registro de respostas e suas referências.

### **3.2.4. Geração de Conteúdo**

- Solicitação de resumos estruturados;
- Criação de glossário de conceitos;
- Documentação e prompts reutilizáveis.

### **3.2.5. Organização e Documentação**

- Estruturação do mini guia no Google Docs;
- Criação de repositório no GitHub;
- README e documentação complementar.

## **4. CONCLUSÃO**

Este projeto demonstrou a eficácia do uso de Inteligência Artificial Generativa como ferramenta de aprendizagem ativa e organização de conhecimento. O Google NotebookLM se mostrou uma plataforma robusta para curadoria de fontes, elaboração de perguntas estratégicas e geração de conteúdo estruturado.

A metodologia aplicada permitiu consolidar conhecimentos sobre as principais bibliotecas Python para análise de dados (Pandas, NumPy e Matplotlib) de forma sistemática e crítica. A curadoria de fontes confiáveis, aliada à exploração guiada por prompts, resultou em um material de estudo completo e reutilizável.

O miniguia gerado representa não apenas uma entrega do desafio proposto pelo Bootcamp "GenAI & Dados" (DIO + Bradesco), mas também uma ferramenta prática para futuras consultas e aprofundamento em análise de dados com Python. Os resumos estruturados, glossário de conceitos e prompts reutilizáveis constituem uma base sólida para o desenvolvimento profissional do autor como Analista ou Cientista de Dados.

Este projeto reforça a importância de aliar pensamento crítico, curadoria de fontes e ferramentas de IA para maximizar o aprendizado e a organização do conhecimento técnico.

## 5. REFERÊNCIAS

MCKINNEY, Wes. pandas: powerful Python data analysis toolkit. Pandas Development Team, 2025.

Disponível em: <https://pandas.pydata.org/docs>.

Acesso em: 28 jan. 2026.

NUMPY COMMUNITY. NumPy Documentation. NumPy Developers, 2025.

Disponível em: <https://numpy.org/doc>.

Acesso em: 28 jan. 2026.

REAL PYTHON. Python Matplotlib Guide. Real Python, [s.d.].

Disponível em: <https://realpython.com/python-matplotlib-guide/>.

Acesso em: 28 jan. 2026.

VANDERPLAS, Jake. Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media, 2016.

Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook>.

Acesso em: 28 jan. 2026.