



Python do zero à classificação de textos 🚀

Aula 1

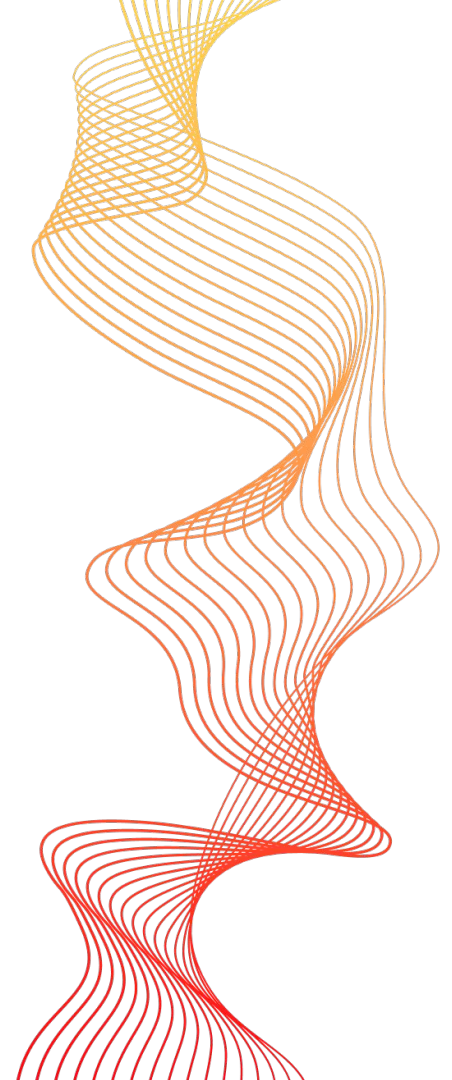
André Gerardi
LabDados
Escola de Direito FGV SP





Ciência de Dados

Visão geral



O que é Ciência de Dados?

A Ciência de Dados é um campo multidisciplinar que abarca a coleta, processamento, análise e interpretação de conjuntos de dados com a finalidade de extrair informações.

Para os acadêmicos, a exploração de dados permite a identificação de padrões, tendências e relações, podendo ser aplicada em diversas áreas de pesquisa.

O propósito fundamental consiste em converter informações de natureza quantitativa em informações, insights e conhecimento.

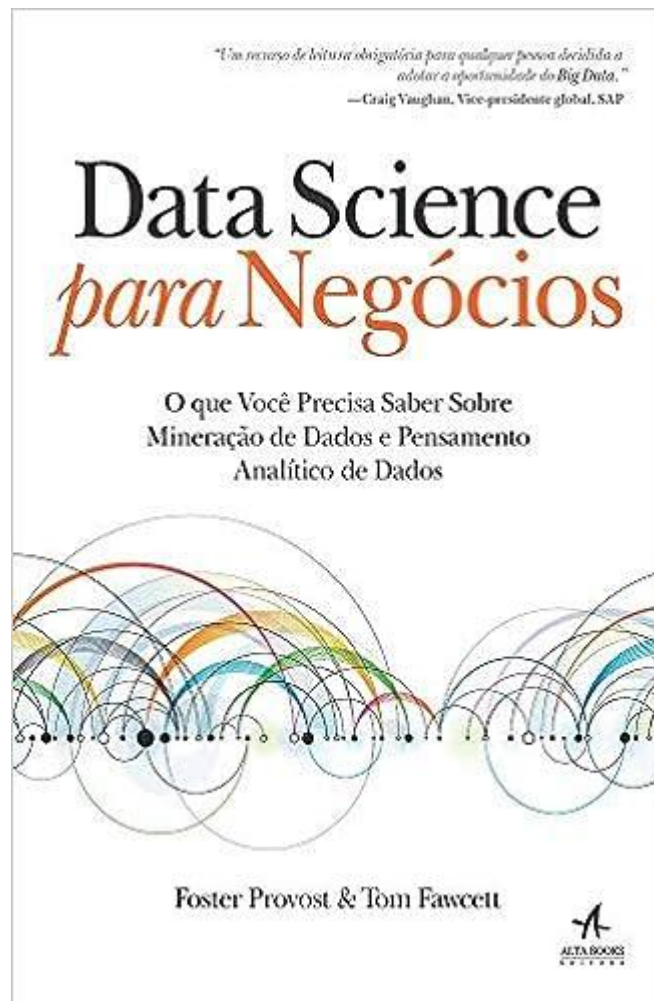
Aplicações:

- Dispositivos;
- Big Data;
- Mercado Financeiro;
- Entretenimento (Netflix)
- Varejo (Amazon)
- Gestão Marketing (Coleta dados)
- Saúde (Predição)
- Direito?



Livro:

- Livro para iniciantes
- Focado em exemplos de negócios;
- Mostra como pensar com dados;
- Como estruturar problemas data-driven;
- Como aplicar Ciência de Dados em diversas áreas;
- Insights para quem quer começar a trabalhar com dados;



Aplicações da Ciência de Dados

As aplicações da Ciência de Dados são vastas e abrangem diversas áreas do conhecimento e setores da sociedade.

No contexto acadêmico, auxilia na análise de grandes conjuntos de dados de pesquisa, na geração de insights e na criação de modelos preditivos que podem orientar investigações, por exemplo.





Descrição:

- Linguagem de alto nível;

Vantagens:

- Diversas opções para análise estatística;
- Grande ferramental para visualização de dados;
 - Analista de dados
 - Shine

Desvantagens:

- Pouca integração com Frameworks Web;
- Menos opções para Deep Learning;
- Atualização menos frequente;
- Linguagem usada exclusivamente no meio acadêmico;



Descrição:

- Linguagem de alto nível;
- Fácil aprendizado;

Vantagens:

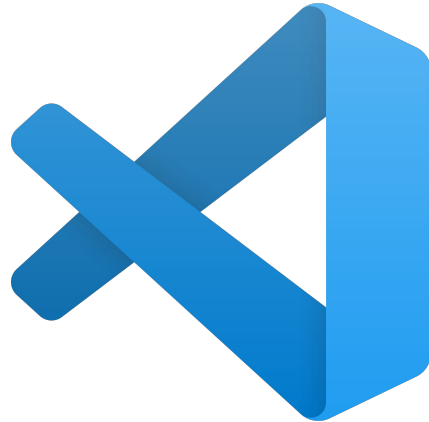
- Melhor qualidade dos pacotes de visualização de dados;
- Deep Learning e IA (Keras, Pytorch, TensorFlow);
- Forte integração com frameworks da indústria (Flask, Django, [Streamlit](#));
- Integração Cloud fácil (bigtechs disponibilizam bibliotecas para integração - AWS, Azure, GCP)

Desvantagens:

- Menos opções para análise estatística e visualização de dados se comparado ao R.



Programação com Jupyter Notebook e Visual Studio Code



Intro

A produção de novos profissionais na área de tecnologia é uma questão urgente e existem muitos obstáculos que impedem as pessoas de ingressarem nesse setor.

As barreiras mais comuns são: exigência de conhecimento prévio, custos, curva de aprendizado, acesso à internet, estereótipos.

É por isso que vamos falar sobre duas ferramentas essenciais para qualquer programador: o Jupyter Notebook e o Visual Studio Code.



O que são interpretadores?

Computadores só podem executar programas escritos em linguagem de máquina, ou seja, binária: Assembly ([010010001...](#))



Os interpretadores, chamados de IDE, traduzem as instruções escritas em uma linguagem de programação para linguagem de máquina.



Python >

Jupyter Notebook

O Jupyter Notebook é uma ferramenta de programação interativa que permite a criação e compartilhamento de documentos que contêm código, visualizações e texto explicativo.

Suporta várias linguagens de programação, incluindo Python, R, Julia.

Uma das principais características do Jupyter Notebook é a sua capacidade de executar o código em blocos separados, chamados de células.

linha de comando CMD:

```
conda install -c anaconda python
```

Portal Anaconda:

```
https://www.anaconda.com/download
```



VS Code

O Jupyter Notebook é uma ferramenta de programação interativa que permite a criação e compartilhamento de documentos que contêm código, visualizações e texto explicativo.

Suporta várias linguagens de programação, incluindo Python, R, Julia.

Uma das principais características do Jupyter Notebook é a sua capacidade de executar o código em blocos separados, chamados de células.



Acesse sua conta do Google Drive!