

Data Science do ZERO

Capítulo 03 - Dominando o Python

minerandodados.com.br

Porque Python para Data Science?

- Vasto suporte da comunidade.
- Bibliotecas diversas para **Análise de dados, Computação Científica e Visualização**.
- Uma linguagem mais disseminada na indústria (se comparado com R ou Scala).
- Mais fácil de aprender que Java.



Anaconda

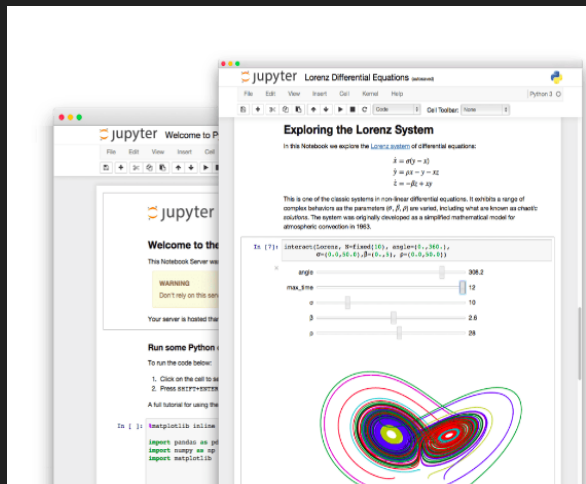
- Plataforma Open Source
- Contém centenas de pacotes para trabalhar com **Data Science**.
- Bibliotecas para Machine Learning e outras tarefas



ANACONDA®

Jupyter Notebook

- Software Open-Source
- Ferramenta interativa e permite diversas linguagens.
- Fácil usabilidade e bastante produtiva.
- Permite plotar gráficos e rodar scripts, tudo em um browser.
- Permite a documentação de scripts usando a linguagem **Markdown**.



The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

[Try it in your browser](#)

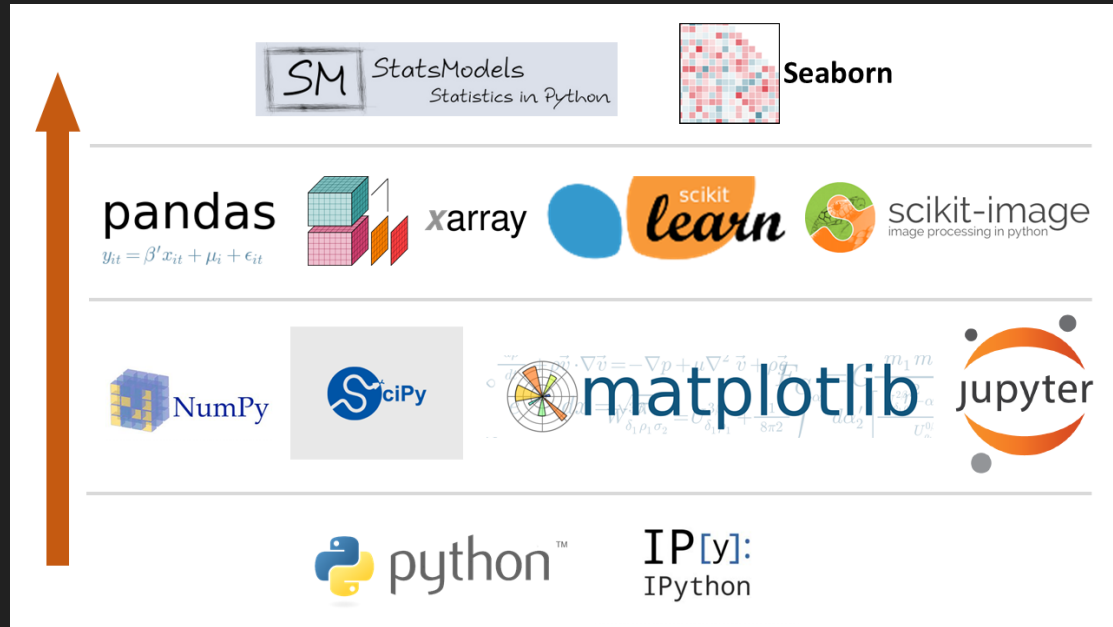
[Install the Notebook](#)



Bibliotecas Python

Bibliotecas mais utilizadas e suas aplicações.

NumPy - Voltada para **computação científica**.
Implementa **arrays multidimensionais**, facilita execução de operações matemáticas e lógicas como ordenação, seleção, transformações, operações estatísticas básicas etc.



Conclusões

- Vimos que existem diversas possibilidades de utilização de Data Science em diferentes segmentos.
- Essas tecnologias já estão mudando mundo em que vivemos.
- Nos próximos capítulos iremos iniciar de forma prática os estudos, te vejo lá..
- Estamos apenas iniciando a nossa jornada!

