

# VISUALIZAÇÃO DE DADOS

Tópicos Especiais em Automação e Controle I  
Leila Rodrigues  
UFC *Campus* Sobral - 2020.1



- Python é um ingresso na área de ciência de dados, inteligência artificial e aprendizado de máquina por causa da produção direta de pacotes, módulos poderosos e de qualidade.

- Também é uma linguagem multiparadigma. É tipado dinamicamente, de script, procedural e interpretada. Também pode suportar um tipo de programação orientada a objetos.
- É considerado um software estatístico, é especializado e adequado para estatísticas, análise e visualização de dados.

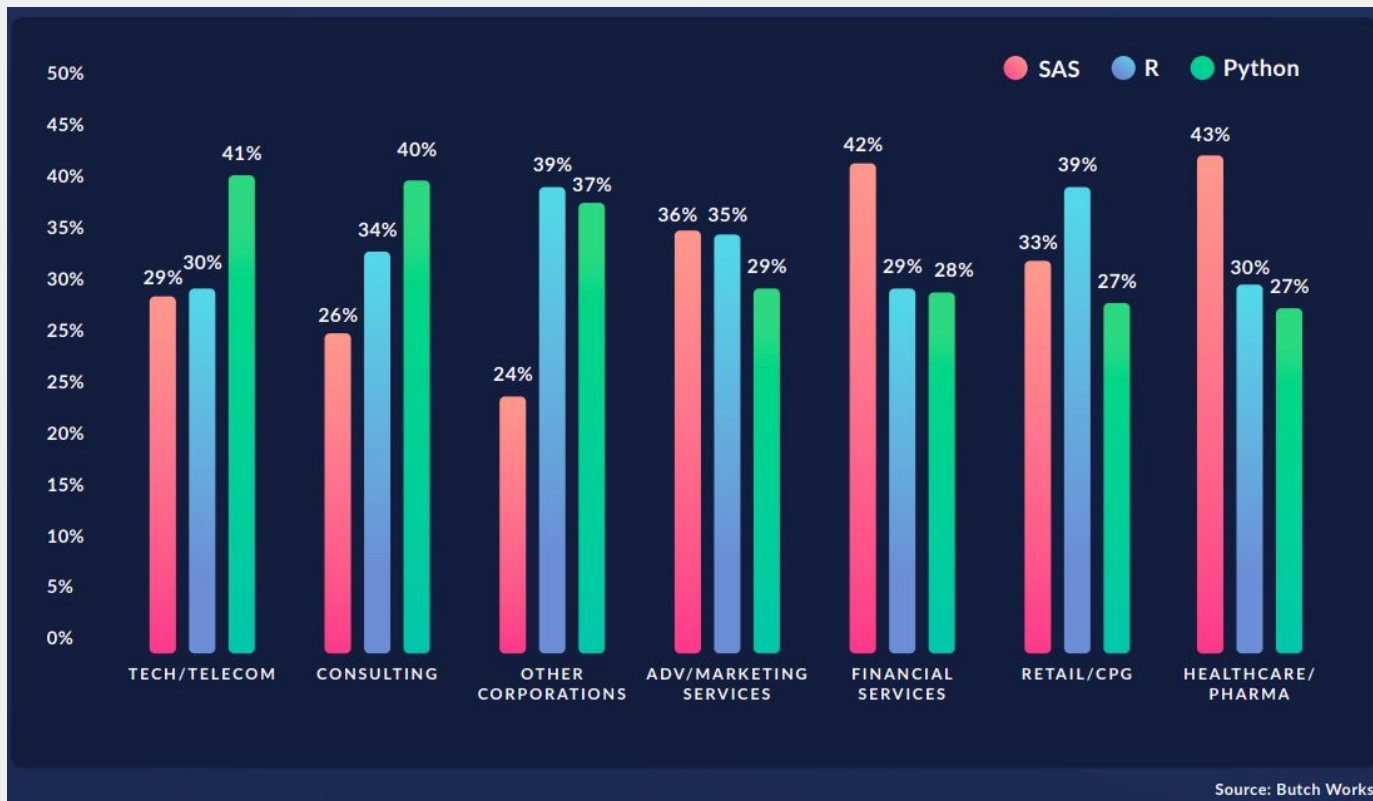


# Popularidade no Stack Overflow: Python vs. R



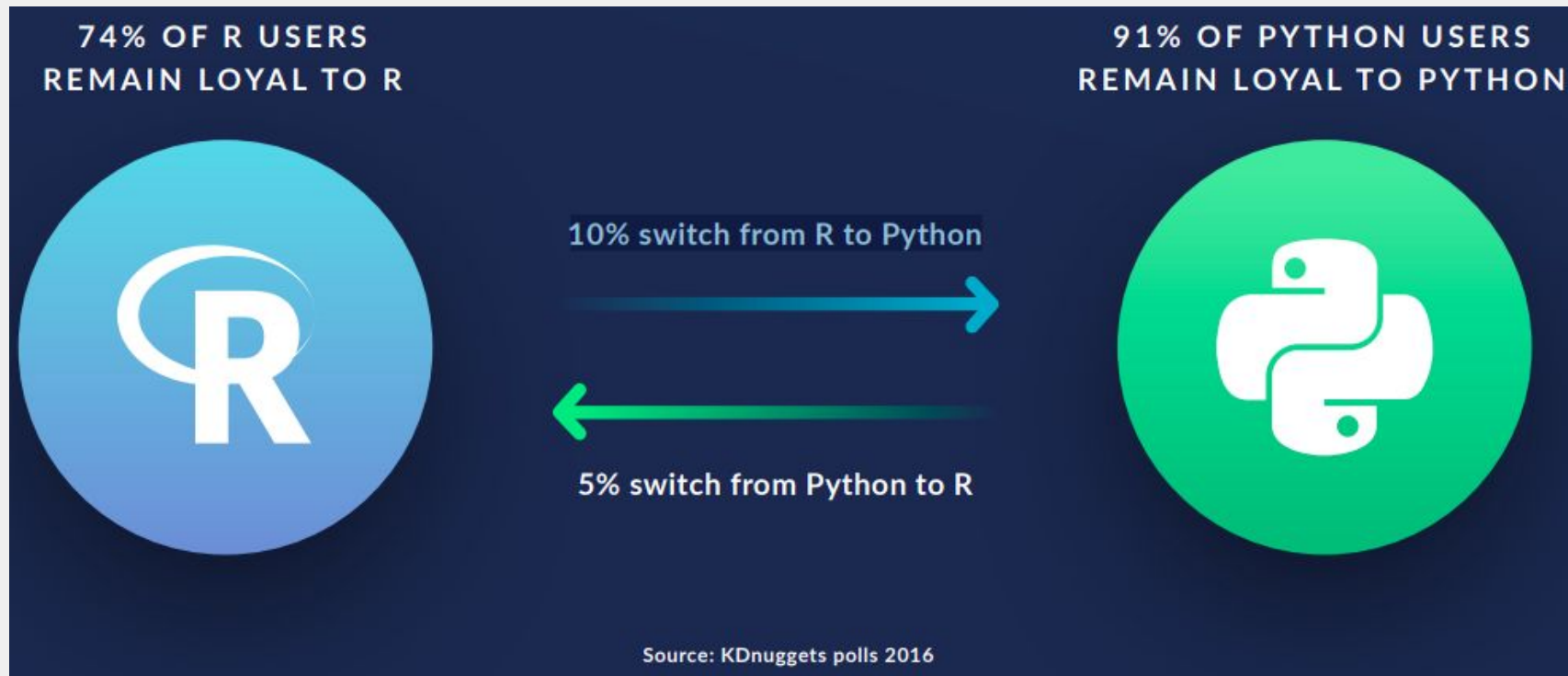
Fonte: <https://s3.amazonaws.com/assets.datacamp.com/email/other/Python+vs+R.pdf>

# Preferência de linguagem por setores da indústria





Fonte: <https://s3.amazonaws.com/assets.datacamp.com/email/other/Python+vs+R.pdf>

# Lealdade do usuário: Python vs. R



Fonte: <https://s3.amazonaws.com/assets.datacamp.com/email/other/Python+vs+R.pdf>

# Ranking de popularidade

	 Python	 R
2016	#3	#5
2017	#1	#6
2018	#1	#7
2019	#1	#5

Source: IEEE Spectrum, Sep 2019

Fonte: <https://s3.amazonaws.com/assets.datacamp.com/email/other/Python+vs+R.pdf>



- Python vs R: Melhor Linguagem para Data Science:  
<https://www.youtube.com/watch?v=W3J7EiSXA00>



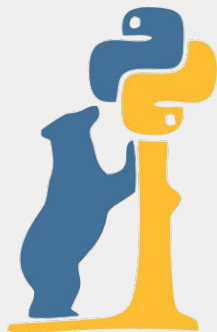
## PYTHON: PRINCIPAIS PACOTES



NumPy

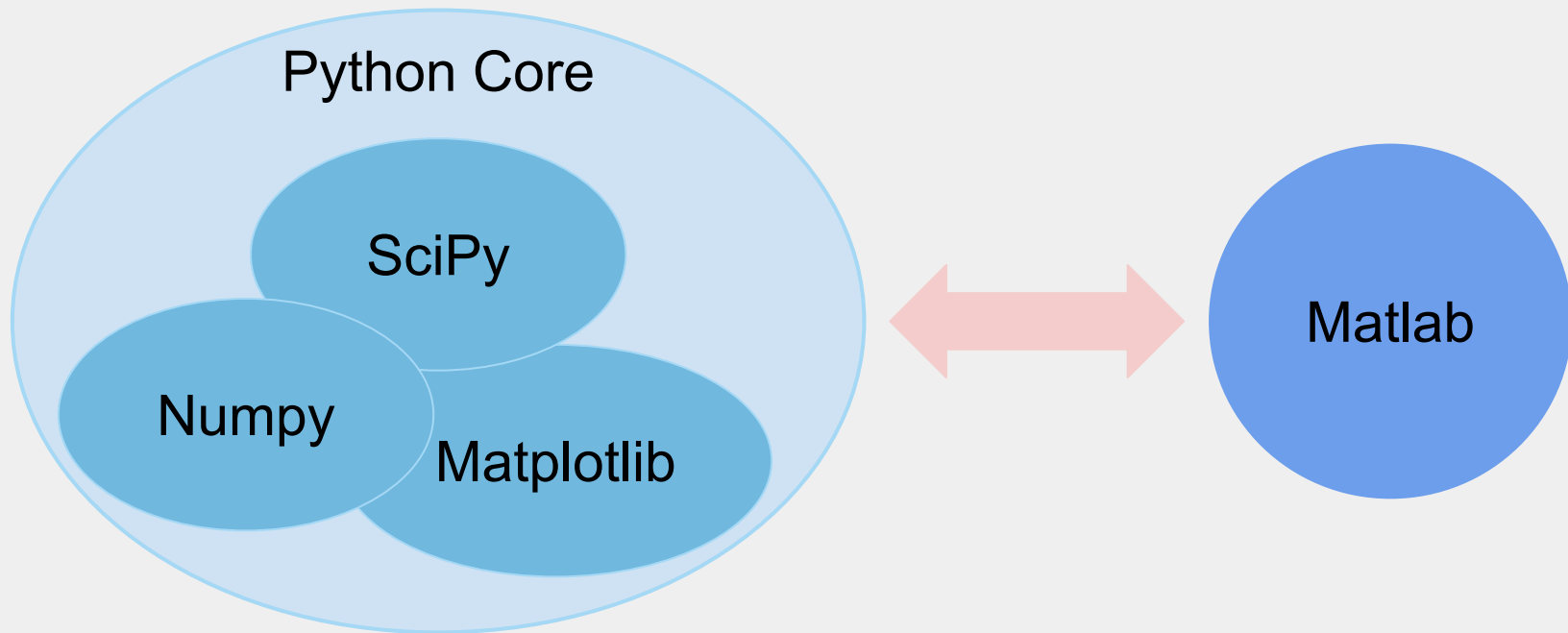
*matplotlib*

Pandas



Seaborn

- É um pacote para computação científica.
- Permite executar qualquer tipo de operações matemáticas e estatísticas.
- Usada para cálculos em Arrays Multidimensionais.
- Modelos de Machine Learning, Processamento de Imagem e Computação Gráfica.



- Ferramenta mais importante para data science atualmente em Python.
- Usada principalmente para manipulação e análise de dados.
- Possui funções para leitura de planilhas do Excel, documentos HTML ou bancos de dados SQL.

- É construído sobre o pacote NumPy.
- Pandas é como o “Excel” para Python.

28 comandos úteis de Pandas que talvez você não conheça

Séries

	Quantidade
0	100
1	91
2	128
3	87



Séries

	Classificação
0	1
1	4
2	5
3	3



Dataframe

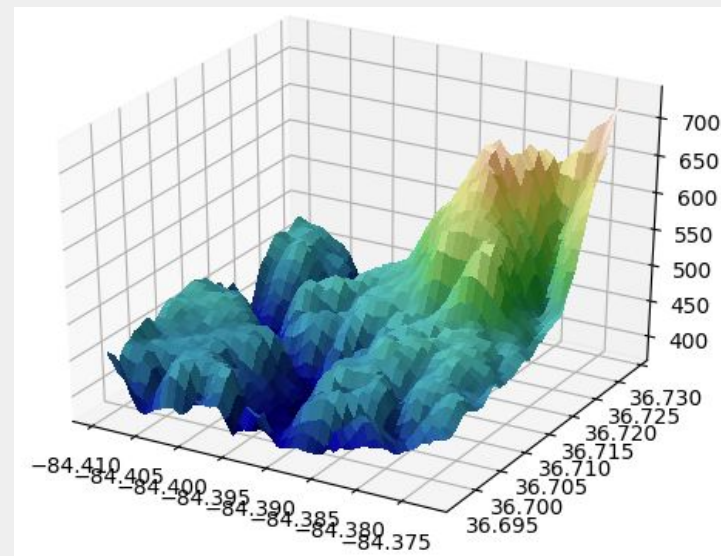
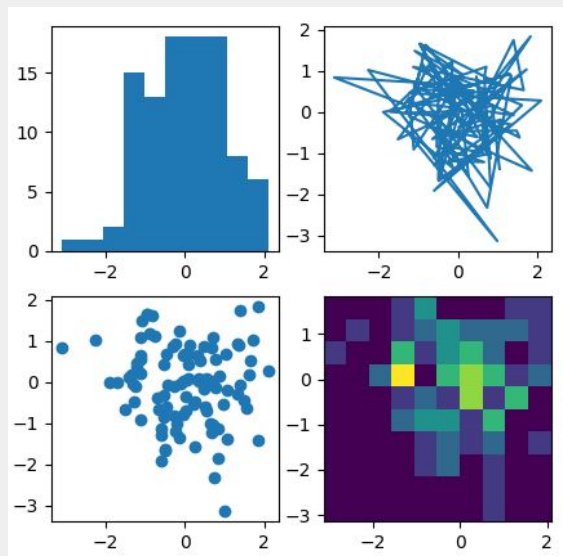
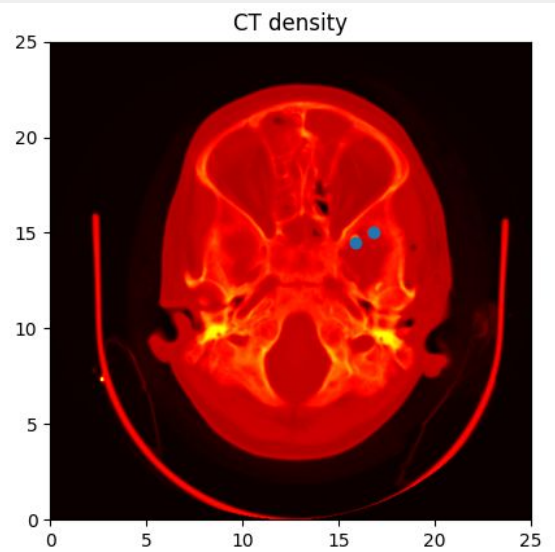
	Quantidade	Classificação
0	100	1
1	91	4
2	128	5
3	87	3

- Uma das ferramentas de plotagem em Python mais utilizadas.
- O módulo mais usado do Matplotlib é o Pyplot, que fornece uma interface como MATLAB, mas usa Python e é de código aberto.

Os 15 melhores comandos de Matplotlib que você vai ver hoje

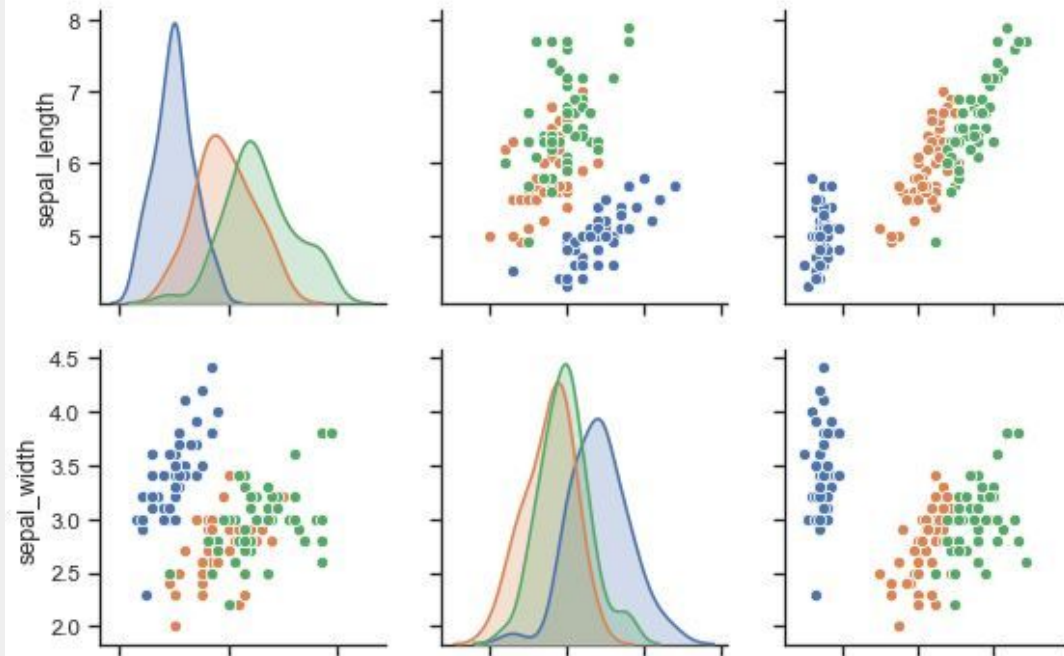
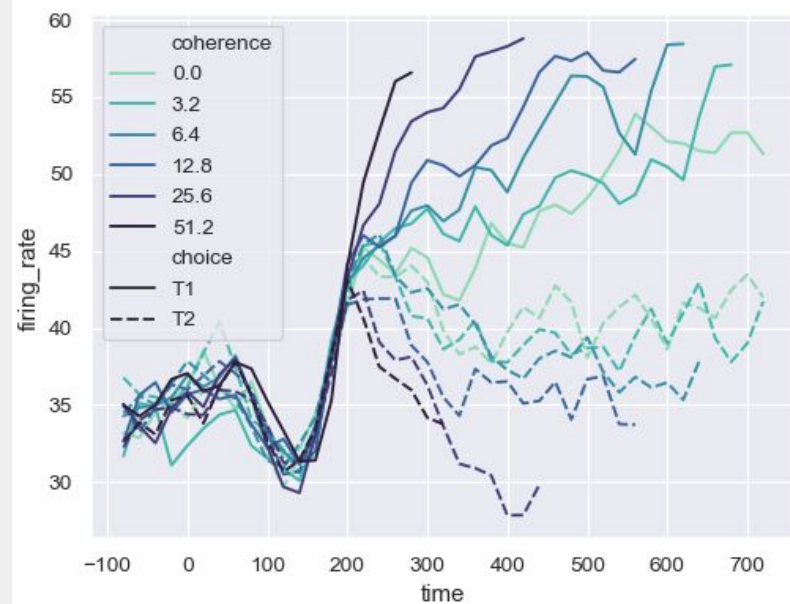


# MATPLOTLIB



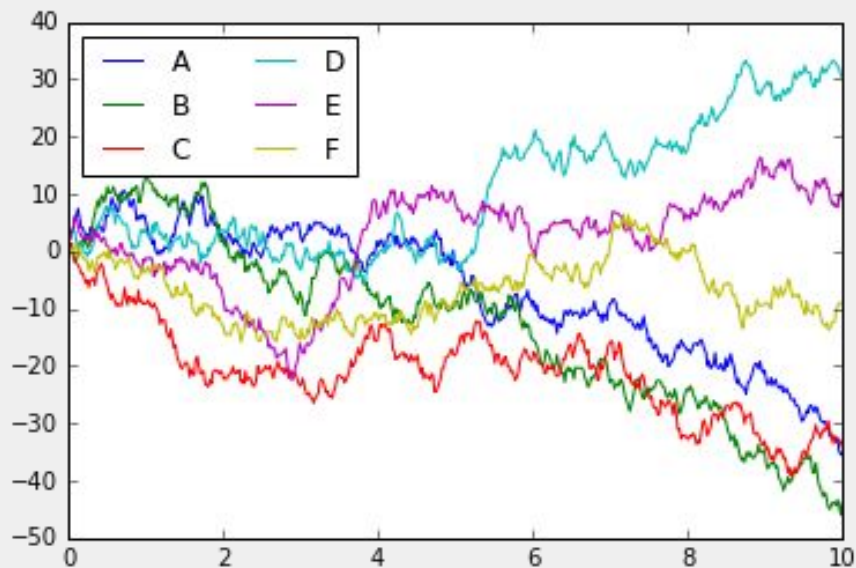
- Baseada no Matplotlib e estreitamente integrado às estruturas de dados do pandas.
- Fornece uma interface de alto nível para desenhar gráficos estatísticos atraentes e informativos.

Python Seaborn Tutorial For Beginners



- O Matplotlib pode ser personalizado, mas é difícil descobrir quais configurações são necessárias para tornar os gráficos mais atraentes.
- Ao trabalhar com o Pandas, o Matplotlib não funciona bem quando se trata de lidar com DataFrames.

# MATPLOTLIB X SEABORN



```
import matplotlib.pyplot as plt
plt.style.use('classic')
%matplotlib inline
plt.plot(x, y)
plt.legend('ABCDEF', ncol=2, loc='upper left');
```



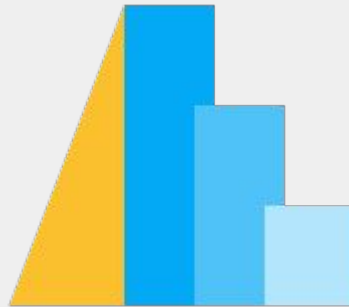
```
import seaborn as sns
sns.set()
plt.plot(x, y)
plt.legend('ABCDEF', ncol=2, loc='upper left');
```

# VISUALIZAÇÃO INTERATIVA DE DADOS



plotly

b<sup>okeh</sup>



- Biblioteca de código aberto, disponível para Python, Javascript e R.
- Pode ser usado como uma ferramenta offline e online.

Como criar gráficos interativos utilizando Plotly e Python

Plotly is a leading provider of interactive data visualization solutions for the enterprise. Plotly's cloud-based platform enables users to create, share, and collaborate on data visualizations across their organization. Plotly's solutions are used by a wide range of industries, including financial services, healthcare, and manufacturing.



- Como ingressar no mercado de Data Science
  - <https://youtu.be/gxyjGmOV540>
- PodCast: Data Science Academy
- PodCast: Hipster Ponto Tech
  - #73, #134
- PodCast: NerdCast (NerdTech)
  - #38 - Vingadores do Data Science

