Practical Neural Networks From Scratch

Concepts, Techniques and Tools to Build Intelligent Systems

Module 2
Beyond Normal Python:
Python For Data Science and Machine Learning

Ali Samanipour

May. 2023

Ali Samanipour
linkedin.com/in/Samanipour

1. Setup Your Development Environment
2. NUMPY & Pandas
3. MATPLOTLIB
4. SEABOARN
5. SCIPY
6. SCIKIT-LEARN

Development Environment Setup- Step 1

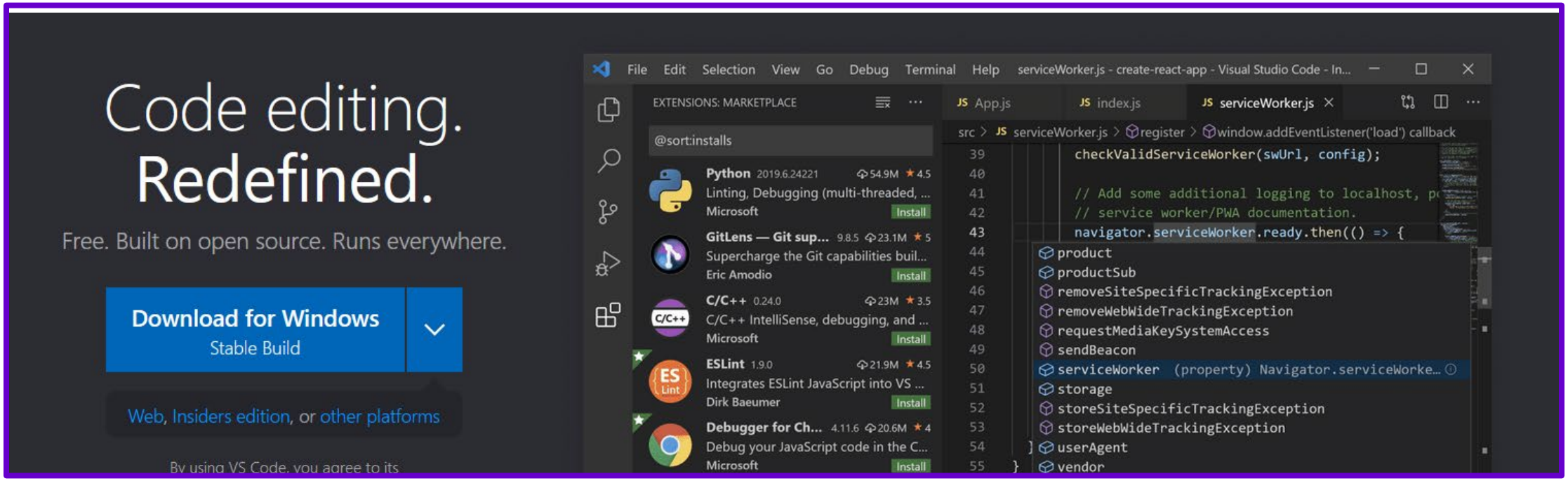
Download and install Python based on your OS.

https://www.python.org/downloads/

Development Environment Setup- Step 2

Download and install VSCode based on your OS.

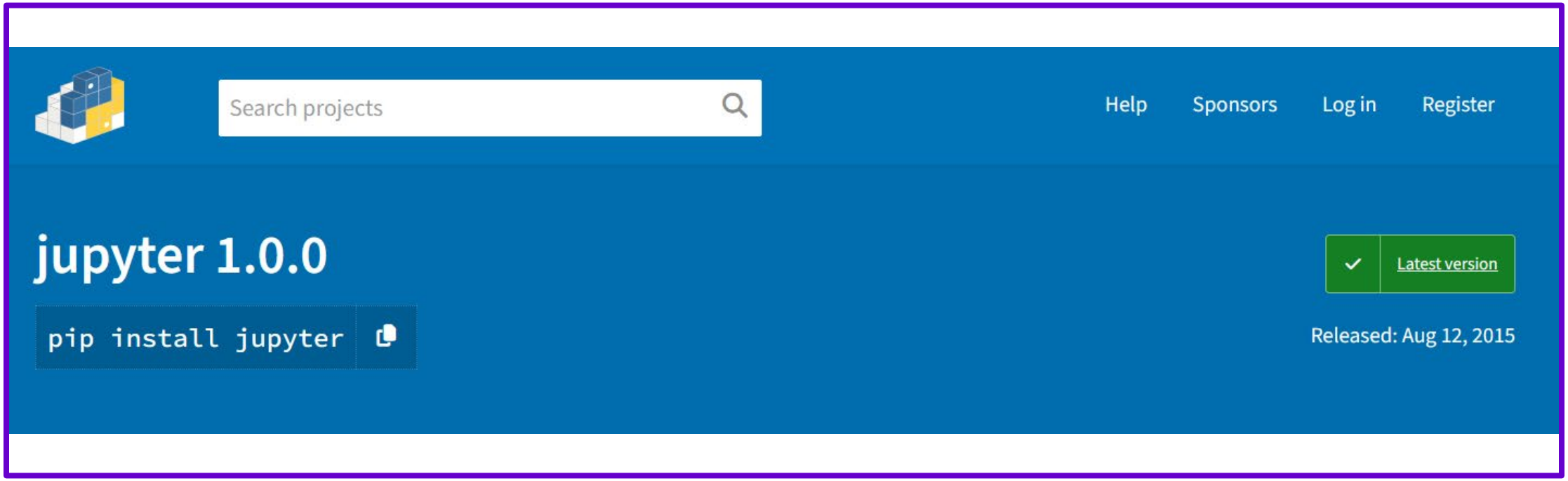
https://code.visualstudio.com/



Development Environment Setup- Step 3

Download and install Jupyter package.

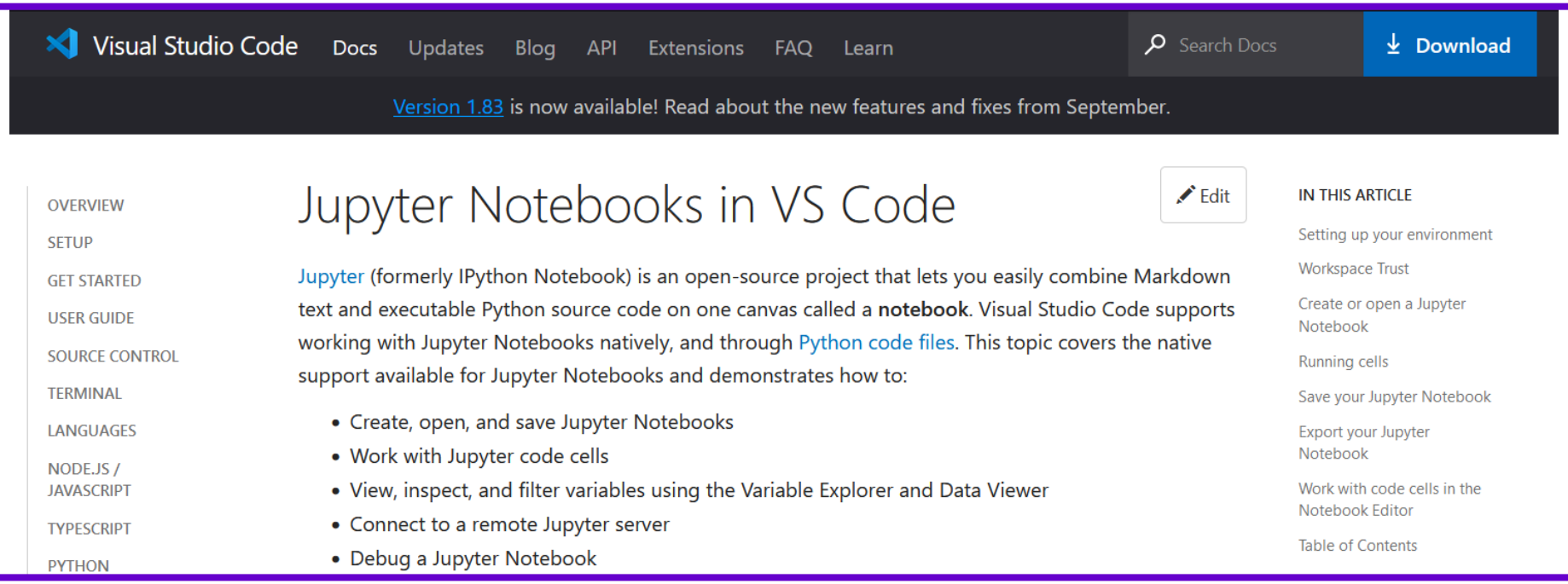
https://pypi.org/project/jupyter



Development Environment Setup- Step 4

Download and install Jupyter Extension on VSCode.

https://code.visualstudio.com/docs/datascience/jupyter-notebooks



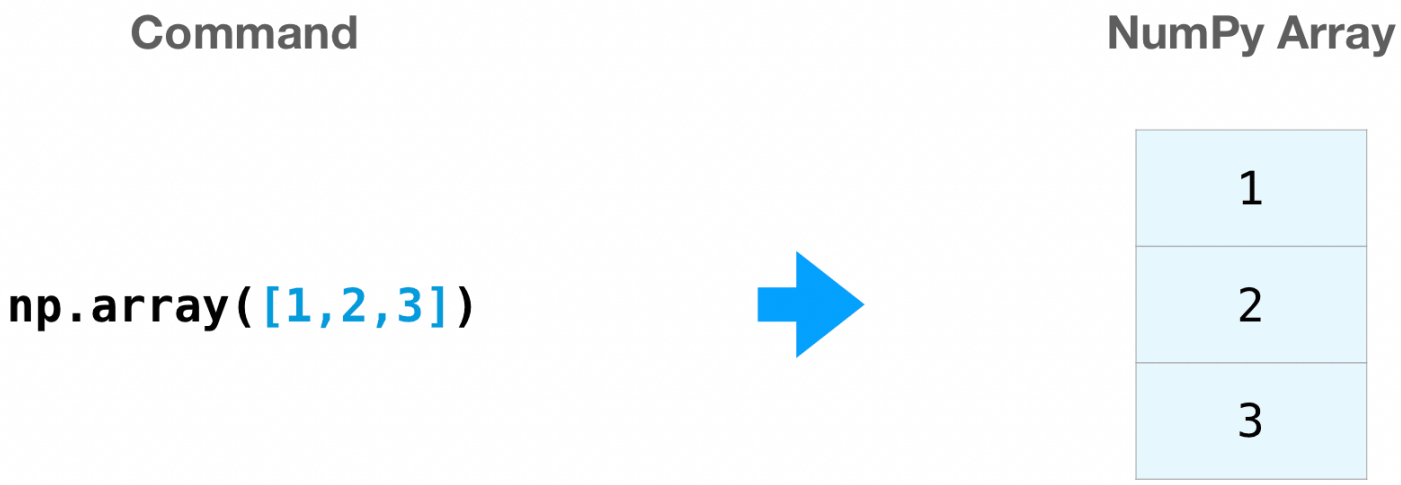
Alternative Development Environments

You could use anaconda, Jupyter, Google Colab, etc.

1. Setup Your Development Environment
2. NUMPY & Pandas
3. MATPLOTLIB
4. SEABOARN
5. SCIPY
6. SCIKIT-LEARN

NUMPY

NumPy is a Python library for **numerical computing**. It **provides powerful data structures**, such as n-dimensional arrays or "ndarrays", and a wide range of mathematical **functions for working with these arrays efficiently**.

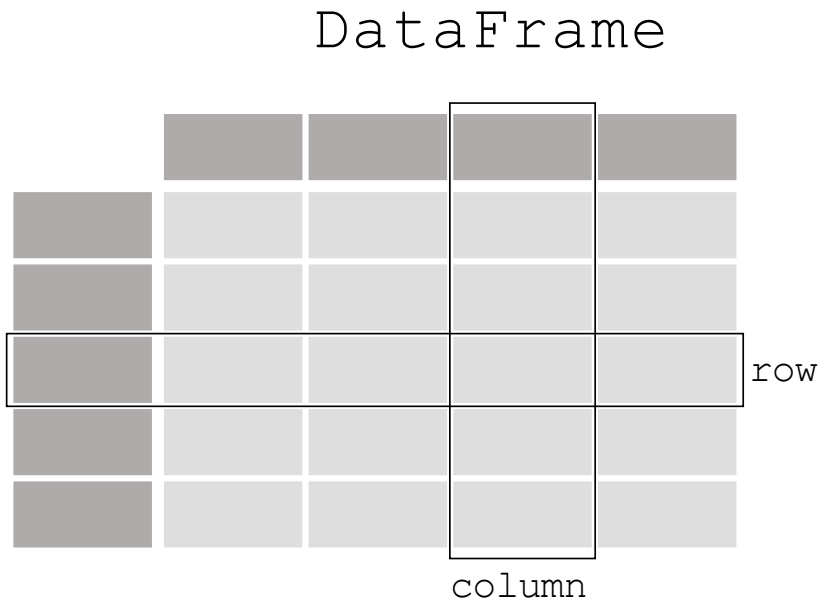


NUMPY

Pros	Cons
• Efficient and fast • Powerful data structures • Comprehensive mathematical functions • Integration with other Python libraries	• Steep learning curve • Memory usage • Lack of flexibility

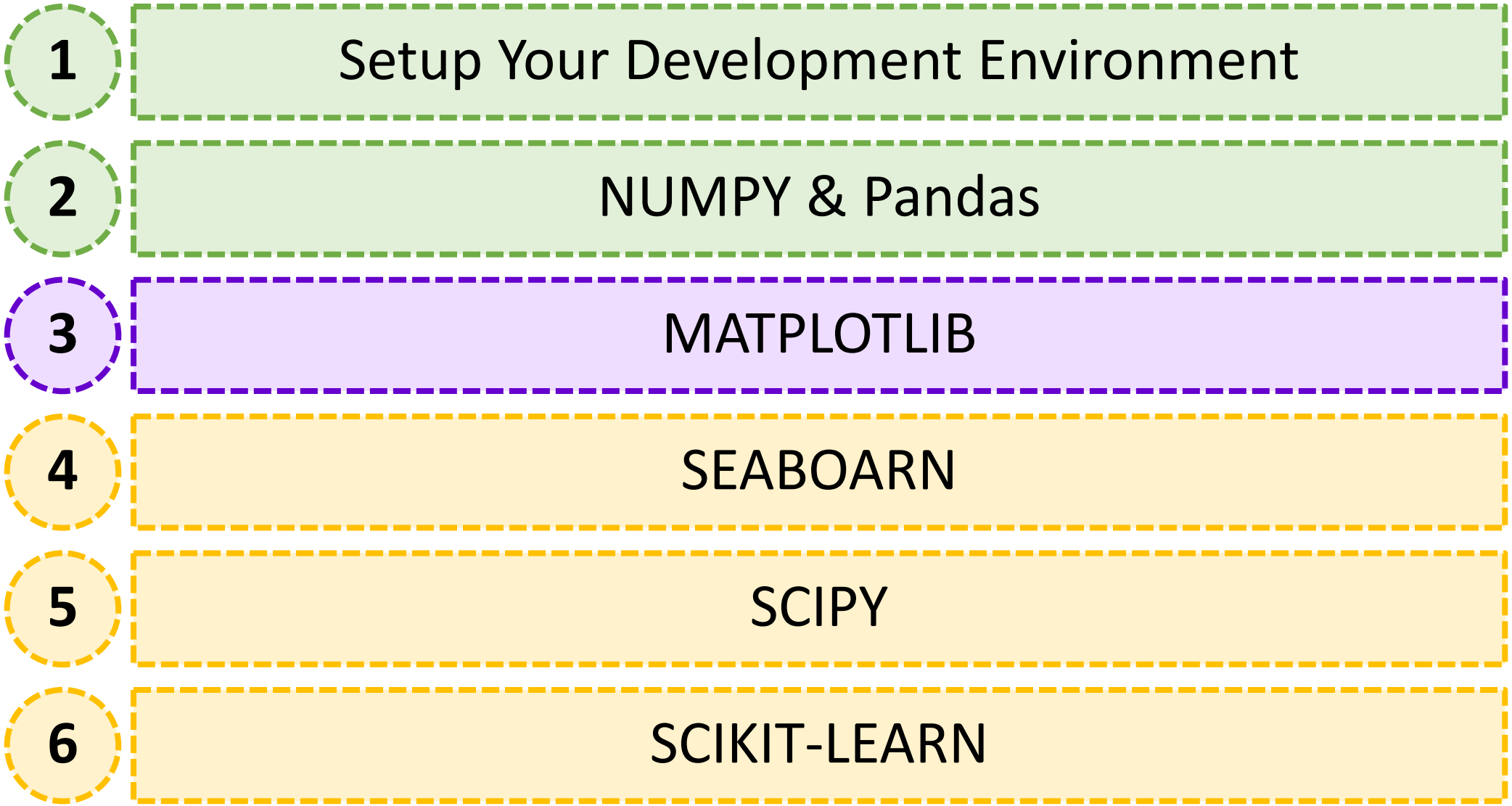
Pandas

Python Pandas is an open-source data manipulation and analysis library for the Python programming language. It provides a set of data structures for efficiently storing and manipulating large data sets, as well as a variety of tools for data analysis, cleaning, and preprocessing.



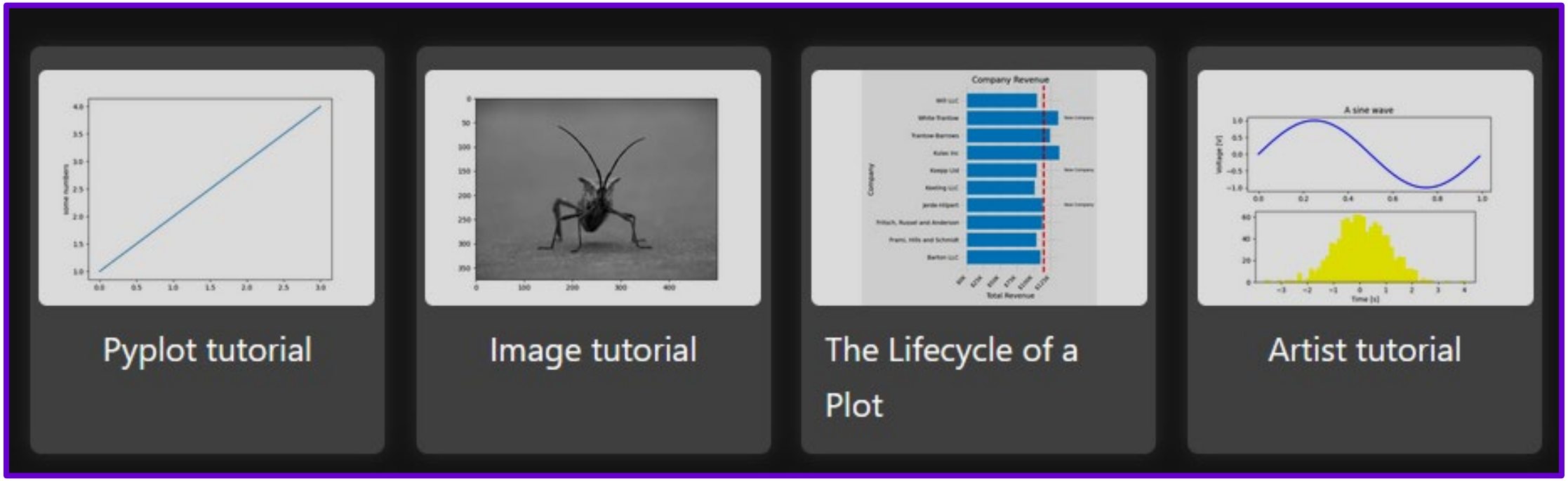
Pandas

Pros	Cons
• Easy-to-use and highly versatile • Provides powerful tools for handling large datasets • Supports a wide range of input and output formats • Offers a rich set of data visualization tools • Has a large and active community • Can be used in conjunction with other popular data science libraries	• Pandas can be memory-intensive • Some of the functions and methods can be complex • Can be slow when performing certain operations • May not always produce the desired results • Some users have reported issues with compatibility and portability between different versions of Pandas


1
Setup Your Development Environment
2
NUMPY & Pandas
3
MATPLOTLIB
4
SEABOARN
5
SCIPY
6
SCIKIT-LEARN

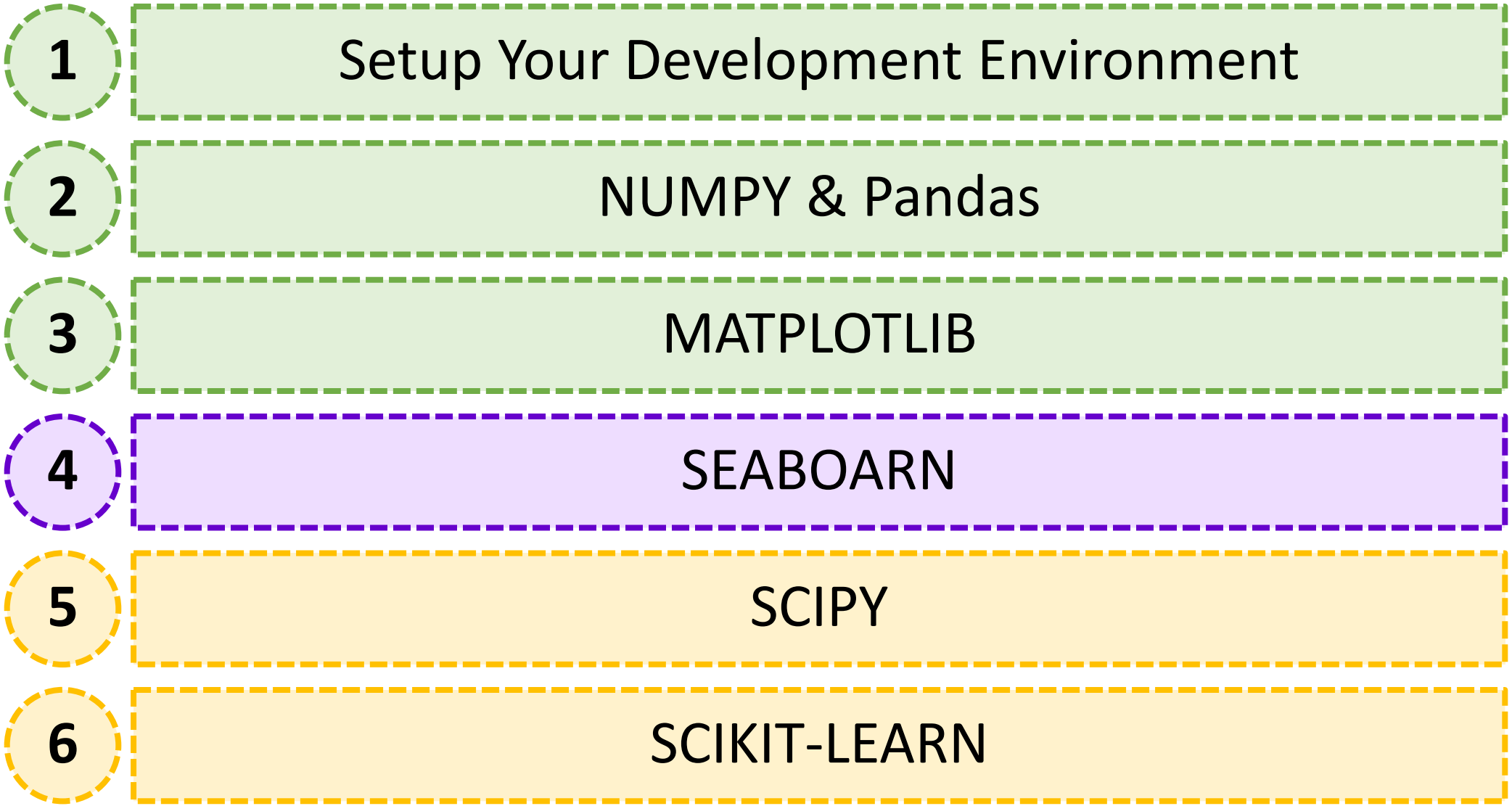
MATPLOTLIB

Matplotlib is a popular **data visualization** library for the Python programming language. It provides a way to create a wide range of static, animated, and interactive visualizations in Python.



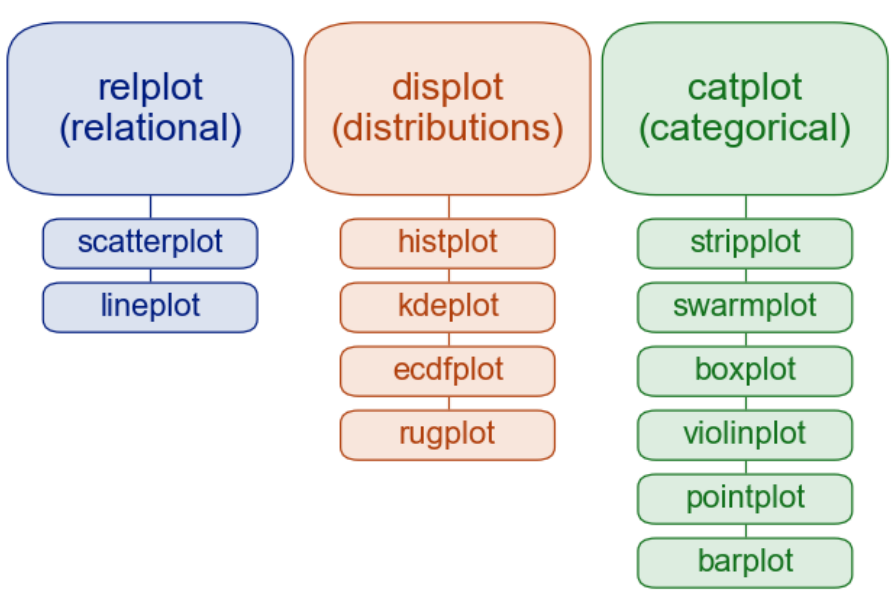
MATPLOTLIB

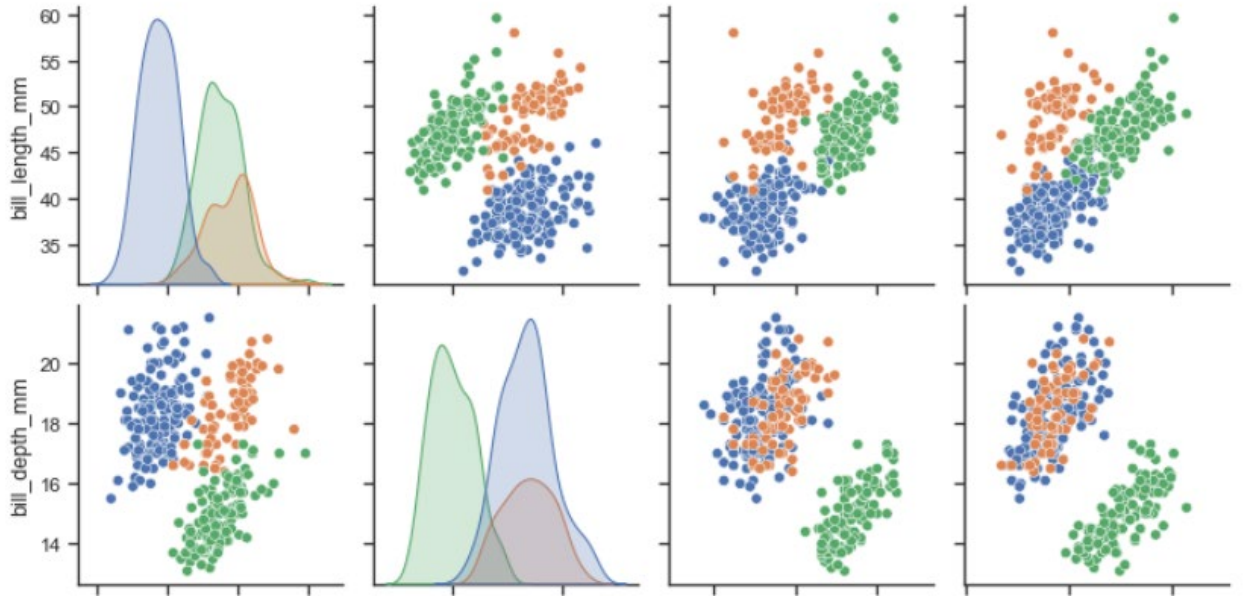
Pros	Cons
• Matplotlib is a widely used • It is highly customizable • It provides a wide range of plotting functionality • Can produce high-quality plots • It is well integrated with other Python libraries • It is easy to use and learn	• The syntax can be verbose and complex • The default settings for plots may not always be aesthetically pleasing • Does not provide as much interactivity • It can be slower for generating complex • Creating complex or advanced plots may require more coding


1
Setup Your Development Environment
2
NUMPY & Pandas
3
MATPLOTLIB
4
SEABOARN
5
SCIPY
6
SCIKIT-LEARN

SEABOARN

Seaborn is a Python **data visualization** library built **on top of Matplotlib**. It provides a high-level interface **for creating informative** and attractive **statistical graphics** in Python





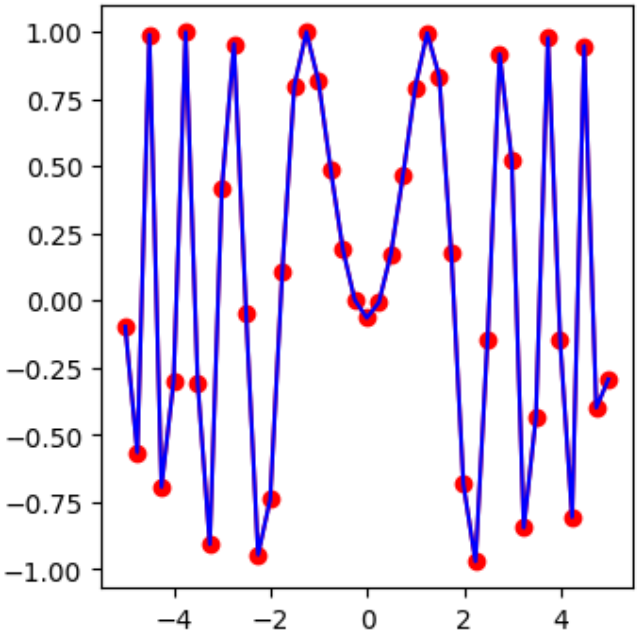
SEABOARN

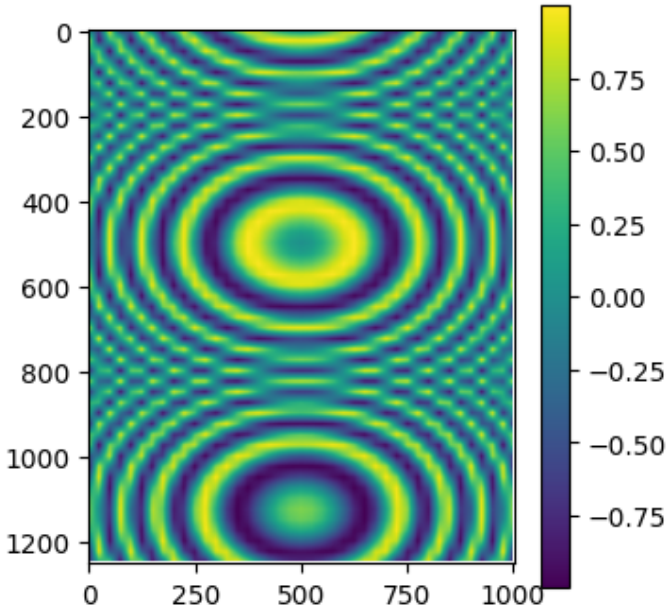
Pros	Cons
• Attractive and informative visualizations • User-friendly interface • Integration with Pandas • Versatility: Seaborn offers a wide range of visualization techniques,	• Limited scope: Seaborn is focused on statistical data visualization • Steep learning curve • Limited customization options

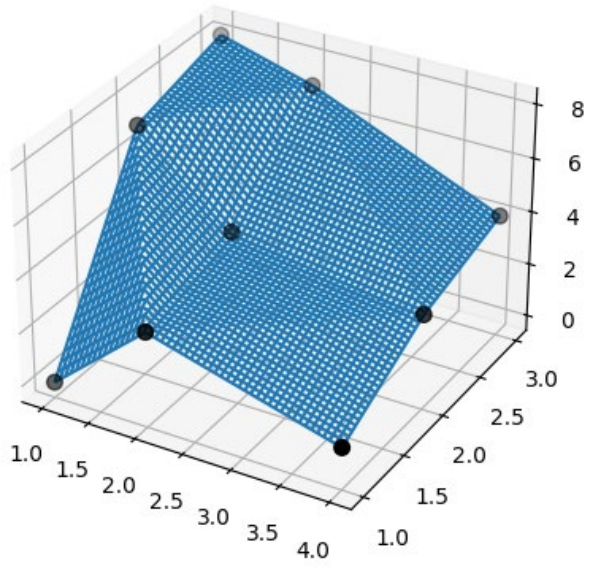
1. Setup Your Development Environment
2. NUMPY & Pandas
3. MATPLOTLIB
4. SEABOARN
5. SCIPY
6. SCIKIT-LEARN

SCIPY

Scipy is an open-source **scientific computing** library for Python that provides a collection of **functions for mathematics, science, and engineering**







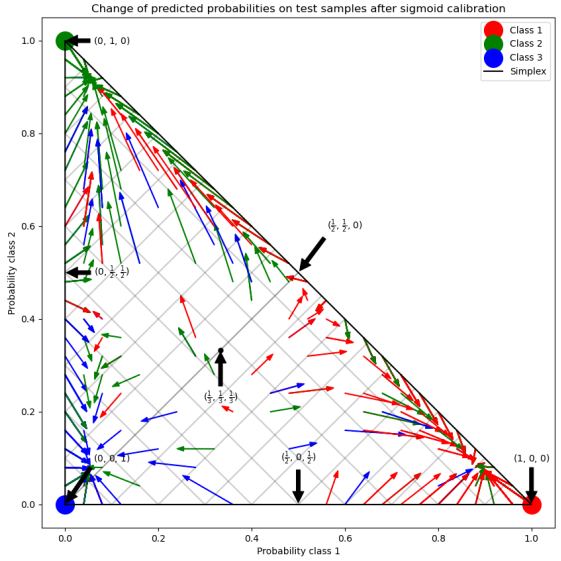
SCIPY

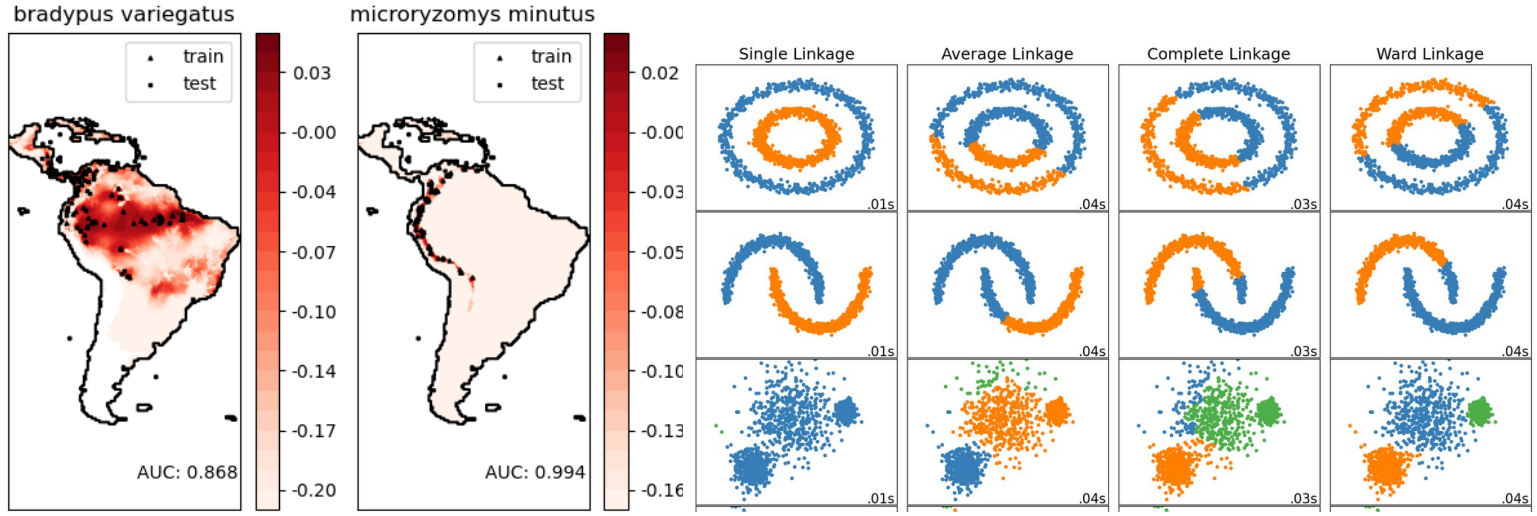
Pros	Cons
• Provides a comprehensive set of tools for scientific computing and numerical analysis • Built on top of NumPy • efficient numerical operations • Large and active community of users • Well-documented with many examples • Portable and cross-platform	• Can be complex and difficult to learn for beginners • Some functions may be computationally intensive • Some functions may have limitations • Requires careful consideration of precision • Some functions may not be as fast as optimized

1. Setup Your Development Environment
2. NUMPY & Pandas
3. MATPLOTLIB
4. SEABOARN
5. SCIPY
6. SCIKIT-LEARN

SCIKIT-LEARN

Python scikit-learn (also known as sklearn) is a **popular machine learning library** for the Python programming language. It provides a range of **supervised** and **unsupervised** learning algorithms





SCIKIT-LEARN

Pros	Cons
• It's a powerful and comprehensive machine learning library • Scikit-learn is easy to use • It is built on top of other popular scientific computing libraries for Python such as NumPy, SciPy, and matplotlib • It provides a range of tools for data preprocessing • Scikit-learn is well-documented • It is open-source and free to use	• it may not be the best choice for some specific tasks or datasets that require more specialized algorithms or models • It may not be the most efficient library for large-scale or complex machine learning tasks • It does not include some newer or more advanced machine learning techniques • Scikit-learn does not include built-in support for some popular machine learning frameworks such as TensorFlow or PyTorch

Course References

[1] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2021.

[2] T. Ghosh and S. K. B. Math, *Practical Mathematics for AI and Deep Learning: A Concise yet In-Depth Guide on Fundamentals of Computer Vision, NLP, Complex Deep Neural Networks and Machine Learning (English Edition)*. BPB Publications, 2022.

[3] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for Machine Learning*. Cambridge University Press, 2020.

[4] T. V. Geetha and S. Sendhilkumar, *Machine Learning: Concepts, Techniques and Applications*. CRC Press LLC, 2023.

[5] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2023.

[6] O. Theobald, *Machine Learning for Absolute Beginners: A Plain English Introduction (Third Edition)*. Scatterplot Press, 2021.

Accessing Course Resource

linkedin.com/in/Samanipour

t.me/SamaniGroup

github.com/Samanipour