

TOOL

ML Python Packages

Python Standard Library

Where	https://docs.python.org/3/library/index.html
What	A diverse set of modules included in all Python distributions providing a wide range of functions and the ability to interact with the operating system.
When	When you need to manipulate files on your computer, perform some mathematical operations, read files in some standard formats, process text strings that you have read in from a file, figure out where your program is spending most of its time when it runs... the list goes on.

IPython

Where	ipython.org
What	An extension of the default Python interpreter that provides powerful interactive capabilities to enhance your productivity.
When	When you need to interactively explore the data with which you are working or develop new data science pipelines by figuring out the tools you need and how they can best work with each other.

Jupyter

Where	jupyter.org
What	A system that enables you to create and share documents integrating code, documentation, commentary, results, and data visualizations.
When	When you want to put together a coherent and reproducible analysis documenting a problem on which you are working, its solution, and the results of your analyses.



Matplotlib

Where	matplotlib.org
What	A 2D plotting package that produces publication-quality figures in a variety of formats, with a high degree of customization.
When	When you need to make plots of your data, including line plots, scatter plots, bar charts, heatmaps, histograms, etc.

NumPy

Where	numpy.org
What	The fundamental package for scientific computing in Python, providing multidimensional arrays and a variety of functions for acting on the data stored in arrays.
When	When you want powerful number-crunching capabilities to work with array data, or as part of the larger Python data science ecosystem, which builds new operations on top of the core functionality provided by NumPy.

Pandas

Where	pandas.pydata.org
What	A fast, powerful, flexible, and easy-to-use data analysis and manipulation tool, especially suited to working with tabular data such as is contained in spreadsheets.
When	When you need to read in data from an Excel spreadsheet or CSV file and then process it further by writing your own analysis code.



Scikit-learn

Where	scikit-learn.org
What	A powerful and comprehensive package for doing machine learning in Python, with support for many algorithms as well as testing and cross-validation procedures.
When	When you want to build predictive models from data, such as classification or regression pipelines (supervised learning) or clustering and dimensionality reduction analyses (unsupervised learning)

SciPy

Where	scipy.org
What	An integrated set of packages for computational mathematics, science, and engineering, providing convenient Python interfaces to many well-established algorithms for scientific computing
When	When you want to find the roots of an equation, identify the parameters at the minimum value of a function, integrate a differential equation, fit your data to a model, do some signal or image processing, interpolate between two data points... the list goes on.

Seaborn

Where	seaborn.pydata.org
What	A data visualization package based on Matplotlib that provides a high-level interface for characterizing statistical properties of data.
When	When you need to visualize structure in your data set, overlaying different components of data into informative visual summaries, with the ability for customization through Matplotlib.

