

Data Science Course Tools

Tools used in this course

Python

Python is an interpreted, object-oriented, high-level programming language.

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.

Numpy

NumPy is the fundamental package for scientific computing with Python. It contains among other things:

- a powerful N-dimensional array object
- sophisticated (broadcasting) functions
- tools for integrating C/C++ and Fortran code
- useful linear algebra, Fourier transform, and random number capabilities

Pandas

Pandas is an open source python library that is built on top of NumPy.

It allows you do fast analysis as well as data cleaning and preparation.

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures.

Seaborn

Seaborn is a Python data visualization library based on matplotlib.

It provides a high-level interface for drawing attractive and informative statistical graphics.

SciKit Learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python.