Python for Data Science

Matthew Cooper Healy mhealy@mst.edu

Python

Python was conceived in the late 1980s, and its implementation began in December 1989 by Guido van Rossum at Centrum Wiskunde & Informatica (CWI) in the Netherlands as a successor to the ABC language (itself inspired by SETL) capable of exception handling and interfacing with the Amoeba operating system.

Learn Python in One Video

Basic Python

Why Python?

- Open Source
- Amazing Community
- Short Learning Curve
- General Purpose Language

Libraries for Data Science

- Numpy
- Pandas
- Matplotlib
- Seaborn
- Scikit-Learn

Examples

Jupyter Notebooks

Loading Data

Doing Analysis

Visualization

Complete Tutorial

Shortcomings(vs R)

- Less helpful visualization suites
- Smaller statistical community
- Less focused on data science

Infographic

Conclusion

Python is an excellent language for rapid testing and prototyping. However, your tools should suit what *you* want to do with them, so it may/may not be right for *you*.

TL;DR: Give it a try, see what you think!