

An Introduction to Python: The SAS Programmers Guide

TT05 – Bradley Harris

An Introduction to Python: The SAS Programmers Guide



Contents

- Review motivation for the project
- Introduce Python programming language and how to get started.
- Introduce Pandas module for data manipulation and analysis.
- Demonstrate some common tasks of a SAS programmer in Python.
- Highlight areas for further learning.
- Conclusions.

An Introduction to Python: The SAS Programmers Guide



Motivation for the Project

- Lots of talk in the industry currently about Machine Learning, AI etc.
- R and Python gaining further popularity as a result, given their ability to carry out these techniques.
- Learn to walk before you can run!
- Objectives:
 - Explore the Python programming language, and it's viability for performing common tasks of the pharmaceutical industry SAS programmer.
 - Create a 'user guide' for SAS programmers wishing to pick up Python.
 - Suggest routes for further learning.
 - Review the plausibility of Python's use within the pharmaceutical industry.

An Introduction to Python: The SAS Programmers Guide



What is Python?

- Freely available general purpose programming language.
- High-level.
- Interpreted.
- Object-oriented.
- Extensive standard library, with ability to extend using third party libraries.



An Introduction to Python: The SAS Programmers Guide

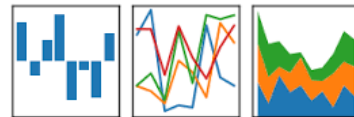


What is Pandas?

- Python module providing high-performance, easy-to-use data structures and data analysis tools.
- Pandas makes it possible for use to effectively work with dataset (dataframe) objects.
- The tools provided allow us to program like a SAS programmer, within Python.
- Pandas is the bread and butter of a programmer/analyst working in Python.

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



An Introduction to Python: The SAS Programmers Guide



Getting Started

- First needed an environment to program in.
- Anaconda – launcher for a number of programming environments as well as simplified management of third-party python packages.
Download Link: <https://www.anaconda.com/distribution/>.
- Programming environments:
 - Spyder.
 - Jupyter Notebook.





Demonstration

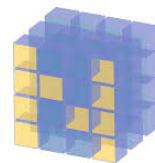


An Introduction to Python: The SAS Programmers Guide



Further Learning

- NumPy – ‘Numerical Python’ a core library for scientific computing.
- Matplotlib - 2D plotting library, which can produce publication quality figures in a variety of hardcopy formats and interactive environments across platforms.
- Scikit-Learn - Module for machine learning in Python. Contains simple and efficient tools for data mining and data analysis
- See paper for suggested reading!



NumPy

matplotlib



An Introduction to Python: The SAS Programmers Guide



Conclusions

- Python is a user friendly and easy to pick up language, great for programmers looking to add another language to their skillset.
- Python is an extremely useful tool for all aspects of data analysis, the ability to add 3rd party modules makes it useable for a wide range of tasks.
- Python could be used in the Pharmaceutical industry, but maybe not in the exact way we use SAS.
- Python opens up the door for greater automation in our daily work, which previously we would have thought to impossible.
- Adding Python programming to our skillset would make us more effective programmers, and more employable. Python has an endless number of uses in almost any industry.



Any Questions?

