

Week 1: An Overview of Python

Kevin Lee

Department of Statistics
Western Michigan University

Outline

1 Overview of Python

2 Python Basics

Python as a Programming Language

- Python is created by Guido van Rossum and first released in 1991.
- Python is an interpreted, high-level, general-purpose programming language.
 - Web development
 - Scientific computing
 - Data analytics
 - Etc.

Python as a Data Analytics Tool

- The Python environment makes it a perfect-fit for data analytics.
 - Simple and easy to learn
 - Has a lot of extensions and active community support
 - Open access to an extensive set of libraries (or packages)

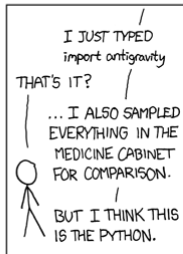
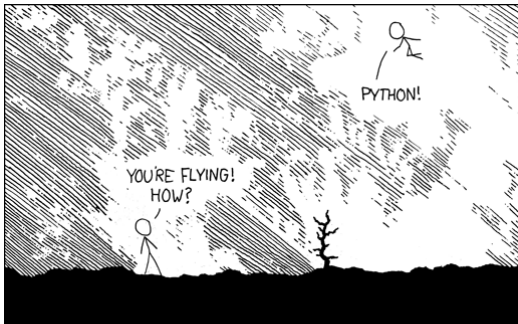
R vs. Python

- R came from the statistics community as a programming language for statistical computing and visualization.
- Python came from the computer science community as a general-purpose programming language.
- Both released in the early 1990s.

- When it comes to statistical analysis (modeling) and solving inference problem, R may be the better option.
- When it comes to machine learning and solving prediction problem, Python may be the better option.
- Choose the one that best fits your needs!

Popular Python Data Analytics Libraries

- Numerical & scientific computing: NumPy, SciPy
- Data manipulation & aggregation: Pandas
- Visualization: Matplotlib, Seaborn
- Machine learning: SciKit-Learn
- Deep learning: Keras, TensorFlow, Theano
- Text mining: NLTK, Gensim



How to Install Python?

- In this course, we will use Spyder which is a free integrated development environment (IDE) that is included in Anaconda.
- Anaconda can be downloaded from <https://www.anaconda.com/download>.
- Download “Python Demo.py” file from Elearning and open the file after launching Spyder.

Outline

1 Overview of Python

2 Python Basics