

3. Python

Data Management
Spring & Summer 2018
OSIPP, Osaka U

Shuhei Kitamura

Goal

Get to know about Python!

Outline of the Python part

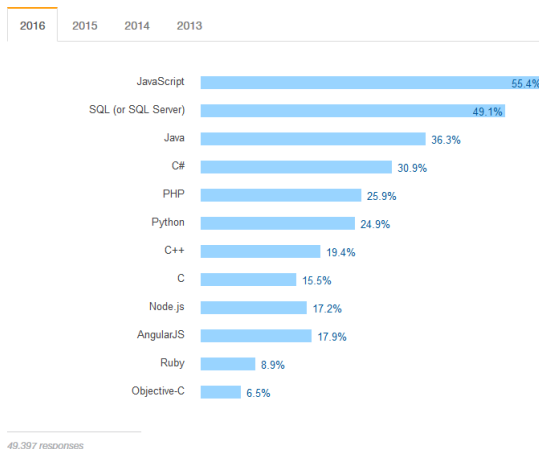
- What is Python?
- Install Python
- Install Jupyter Notebook
- Basics
- Clean data
- Analyze data
- Application - Web scraping

What is Python?

- Python is a general purpose programming language created by Guido van Rossum, a Dutch programmer.
- Free and open source

Popularity of Python

I. Most Popular Technologies



A quarter of developers in Stack Overflow used Python in 2016 (the response rate = 88%).

Versions

Two versions: Python 3.x and 2.x. Which one to use?

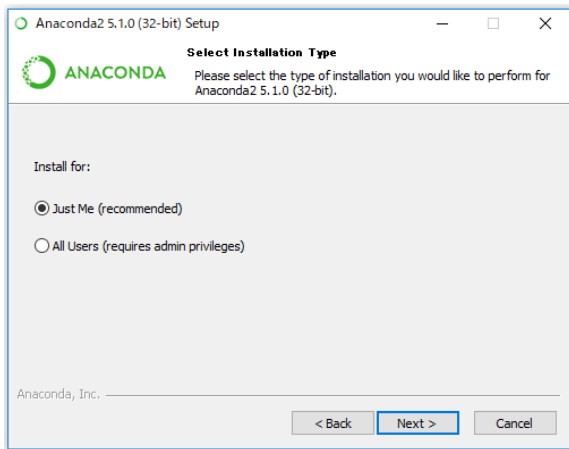
- “Python 2.x is legacy, Python 3.x is the present and future of the language.”
- Differences; e.g., $3/2 = 1$ in Python 2.x but $3/2 = 1.5$ in Python 3.x
- Unless Python 2.x is required (e.g. ArcGIS Desktop), a safer option would be Python 3.x.

Install Python

Install Python using the [Anaconda](#) Distribution

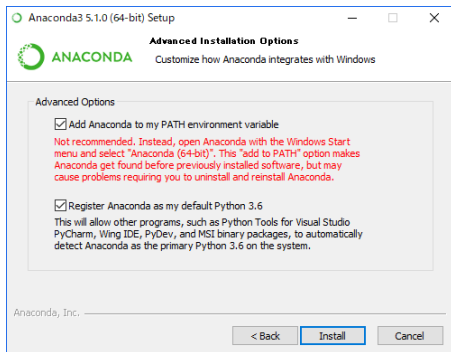
- Include Jupyter Notebook, a web application supporting over 40 programming languages.

Install Python (cont.)



Choose “Just Me (recommended).”

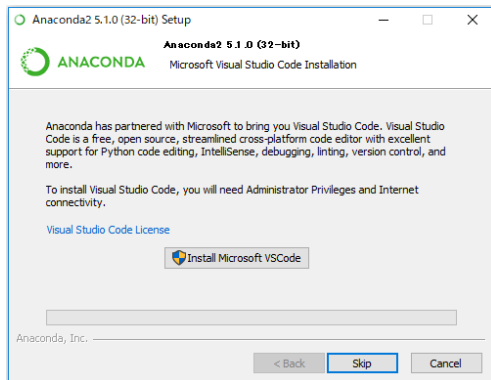
Install Python (cont.)



Check BOTH.

- **CAUTION:** Double-check whether any previously installed software uses Python (e.g. ArcGIS) before starting installation.
- For Arc, a [path file](#) should be added to Anaconda3/Lib/site-packages after installation. Otherwise, Python won't recognize ArcPy.

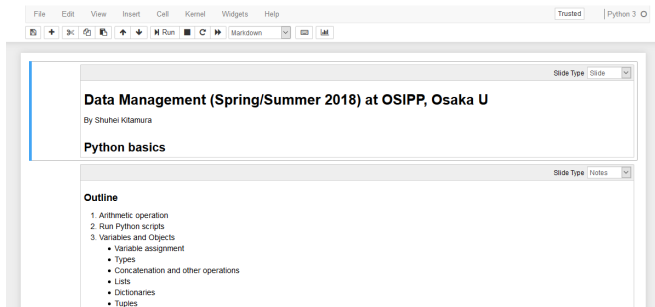
Install Python (cont.)



Install Visual Studio Code (a text editor for coding), if you like.

- Other options: Sublime Text, Vim

Start Jupyter Notebook in your browser



- Type “cmd” in the search box at the bottom of your screen. Command Prompt pops up.
 - For other OS, see [this guide](#).
- Type “jupyter notebook” in the command line. This starts Jupyter Notebook in your web browser.
- In the Notebook, go to your local repository (“dm_tutorial”).
- Click python_basics.ipynb. A screen like the above picture shows up.

Appendix: References

References

Python

- Lectures in Quantitative Economics: ([English](#))
- Matsuo Lab at U of Tokyo: ([Japanese](#))
- An Introduction to Python for Economists: ([English](#))
- DataCamp (not free): ([English](#))