Introduction to Python

Peter M. Dahlgren, JMG

2016-11-24

The Python programming language

The Python programming language was developed 1991 and is an easy to read programming language that many journalists and data scientists use. It's easy to get started compared to many other languages.

Python a interpretative language, which means that Python reads and executes text files. Many other programming languages compile to binary, which means that a text file is converted to machine code. But we don't have to bother with that, fortunately.

Install Python

Python is free and open source. It works on Windows, Mac and Linux.

We will use **Anaconda**, which also installs many add-ons that enhance Python (like notebooks).

> **Download and install Anaconda with Python 3.5**
>
> Choose the 64-bit version (or 32-bit version if you don't know the difference). The download is about 400 MB so it'll take some time on a slow connection. Just click OK to accept the defaults during installation.

You should install Sublime Text 3 as well. It's a text editor. We will probably need them during the course. (Just install the beta, it's stable.)

Install Git too while you're at it. It's a popular version control software that helps keep track of file changes. Maybe we will have some time to use them.

Create your first Python program

You can execute Python code in two ways (don't worry, we will cover everything on the lecture if you have no idea what is going on):

Run code via a notebook

The first way is to run code via an interactive notebook, called Jupyter Notebook. If you have installed Anaconda, a shortcut to Jupyter Notebook has been installed. Simply click on the shortcut to start it, otherwise:

1. Open a command prompt (Windows) or terminal (Mac).
2. Type `jupyter notebook` and press Enter.
3. A web page should open.
4. Create a notebook, write the following in a cell and press Ctrl+Enter to execute:
5. If everything works, `Hello World!` will appear on your screen.

Run code via a file

The second and old-fashioned way is to run code via your command prompt or terminal.

1. Create a text file named `helloworld.py` and save it somewhere on your computer. Note that Python files end with `.py`.
2. Put this text inside the text file and save it:
3. Open a command prompt (Windows) or terminal (Mac) and execute your program by writing `python helloworld.py` and press Enter.
4. If everything works, `Hello World!` will appear on your screen.

Lecture objectives

After the lectures, you will have some knowledge in how to:

- Getting data by scraping web pages
- Getting data by API
- Write a simple news robot

The examination for this part of the course is to write a news robot that follows simple rules and produce a journalistic text.

You will be given code that you can modify, so you don't have to write anything from scratch.

We want you to be able to *think* like a programmer, not necessarily become a programmer.

Code for workshops

All Python code for workshops and examinations are located at GitHub: https://github.com/peterdalle/mij

Resources to learn Python

Web sites

Do you have to buy a book? No, internet is often more than enough. Google is your friend.

- Learn Python basics: http://www.pythonforbeginners.com/
- Pythons own site is helpful: https://www.python.org/about/gettingstarted/
- Ask questions, or find answers, about Python on StackOverflow: http://stackoverflow.com/questions/tagged/python
- Find code projects on GitHub: https://github.com/
- Python documentation at is mandatory reading if you're advanced: https://docs.python.org/3/

Books

- McKinney, W. (2013). Python for Data Analysis. Beijing: O'Reilly.

Python too easy for you?

Follow through the lectures by doing everything in R instead.