



Intro to Python

Last Updated: 3/1/2019

Python is great for working with data, and is rapidly gaining popularity for lots of other things too!

Python 2x vs 3x

[Install Python](#) in the default location, either "C:\Python27" or "C:\Python36" or if you are using the latest version of Python then "C:\Python37". Install and update Python site packages (Numpy, Scipy, and such) with [Pip](#) which is the default Python site package installer using either "python -m pip install site_package_name" or "python -m pip install --upgrade site_package_name". If you have both Python 2.7 and Python 3 installed, then you will have 2 versions of Pip installed too.

To make sure that the right version of Python is running the right Python script: In the first line of the Python script have either "#! python3" or "#! python2". This will guarantee that Python 2 will not complain about running a Python 3 script with features that are not supported in Python 2 and also guarantee that Python 3 will not complain about running a Python 2 script with features that have been discontinued in Python 3. Note that Python 2.7 will not be maintained past 2020.

If you have Python 3 installed, Python 2.7 scripts can be run at the command prompt with "py -2 my_python_2_script.py" and Python 3 scripts can be run with "py -3 my_python_3_script.py". We recommend using the above method though with the "#!". By setting your PATH to Python

3 first that will have Python 3 run before Python 2 when you type python in the command prompt. You can reverse this for Python 2.

Jupyter Notebooks

Many people use Jupyter Notebooks as a place to work on and store code. Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

Geospatial Data Python

Always let Python install Python in the default location, either "C:\Python27" or "C:\Python36" or if you are using the latest version of Python then "C:\Python37". Remember that there are differences between Python 2.7 and Python 3. Never have ArcGIS or QGIS install Python on your computer if you intend to use Python outside of ArcGIS or QGIS which includes having programs like PyScripter, [GDAL](#), etc.

ArcPy Tips

Another good package to install is Pylint, which is a linter for Python which will improve the quality of your Python code. Pylint got its name from the old Unix command lint which was used to find problems in C programs before these problems were found by the compiler or at run time. For kicks, you can run some of ESRI's ArcPy code through Pylint and the results that Pylint generates will be a real eye-opener as to why ESRI code is so buggy.

Python for Data Cleaning

The [pandas](#) library is very popular for data transformation and analysis work. Pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language. A great way to get started with pandas is the [10 Minutes to pandas tutorial](#). In combination with the [Numpy](#) package, Python is an excellent language to work with data.

Python for Other Stuff

Python is used in an abundance of other scenarios, such as machine learning, web development, software development, and task automation. For examples of what python is used for, check out the [Applications for Python](#) page. For a list of python packages, explore the [Python Package Index](#).