# 14.1\_Scripts\_Modules

April 17, 2023

## 1 Introduction to Python for Open Source Geocomputation



• Instructor: Dr. Wei Kang

• Class Location and Time: ENV 336, Mon & Wed 12:30 pm - 1:50 pm

#### Content:

- Scripts
- Modules
- Packages

#### 1.0.1 Two working modes modes: interactively and running scripts

- Interactive working mode (interpreter (code cell in a Jupyter Notebook))
  - You can use the interpreter to build up and test pieces of code until you get them working to your liking, at which point you can save them to the text file/script.
- Running scripts:
  - Scripts are useful when you want to permanently save some code with an eye for reuse later.

Complementary: using both an interpreter and scripts together is a common use pattern for scientific programming in Python.

# 2 What is a script?

- write the code in text files
  - using a text editor
  - Integrated Development Environment (IDE)

- \* more features such as build automation, code highlighting, testing and debugging.
- \* examples: PyCharm, Spyder, visual studio
- The file has the .py extension

#### 2.0.1 Writing and Running Python Scripts

Create a new file from the Jupyter directory called hello.py and enter the following into this
file

```
print('Hello World!')
```

- Running the script hello.py
  - Shell (Windows: Git BASH, Mac: terminal):

```
python hello.py
```

- Python interpreter:

```
>>> exec(open("hello.py").read())
```

```
[1]: exec(open("hello.py").read())
```

Hello World

## 3 What is a Module?

- More organized "script"
  - A module is a set of functions and data structures.
- A module is a file containing Python code that can be **reused** in other Python code files.
- The file of a Python module has the .py extension.
- A module is often reused by import statements.

### 3.1 Standard Python Modules

Python comes with a rich set of standard modules (learn more):

- Numeric and Mathematical Modules
  - numbers Numeric abstract base classes
  - math Mathematical functions
  - cmath Mathematical functions for complex numbers
  - decimal Decimal fixed point and floating point arithmetic
  - fractions Rational numbers
  - random Generate pseudo-random numbers
  - statistics Mathematical statistics functions
- Functional Programming Modules
  - itertools Functions creating iterators for efficient looping
  - functools Higher-order functions and operations on callable objects
  - operator Standard operators as functions

## 3.2 Using Modules

- getting the whole module: import module\_name
- only getting a specific function/class/attribute: from module\_name import something

```
[2]: import math
```

Use math. tab to inspect all the objects (functions, variables, etc) in the current namespace

[3]: math.

```
Cell In[3], line 1
math.

SyntaxError: invalid syntax
```

[4]: math.isinf?

```
[5]: help(math.isinf)
```

Help on built-in function isinf in module math:

```
isinf(x, /)
```

Return True if x is a positive or negative infinity, and False otherwise.

```
[6]: dir(math) # used to find out which names a module defines
```

```
[6]: ['__doc__',
      '__file__',
      '__loader__',
      '__name__',
      '__package__',
      '__spec__',
      'acos',
      'acosh',
      'asin',
      'asinh',
      'atan',
      'atan2',
      'atanh',
      'ceil',
      'comb',
      'copysign',
      'cos',
      'cosh',
      'degrees',
```

```
'dist',
'e',
'erf',
'erfc',
'exp',
'expm1',
'fabs',
'factorial',
'floor',
'fmod',
'frexp',
'fsum',
'gamma',
'gcd',
'hypot',
'inf',
'isclose',
'isfinite',
'isinf',
'isnan',
'isqrt',
'lcm',
'ldexp',
'lgamma',
'log',
'log10',
'log1p',
'log2',
'modf',
'nan',
'nextafter',
'perm',
'pi',
'pow',
'prod',
'radians',
'remainder',
'sin',
'sinh',
'sqrt',
'tan',
'tanh',
'tau',
'trunc',
'ulp']
```

Use the form module.function\_name to call a function that lives inside the module.

```
[7]: math.sqrt(4)
 [7]: 2.0
     Use the form module.attribute to call an attribute that lives inside the module.
 [8]: math.pi
 [8]: 3.141592653589793
     Use an alias to rename the module upon import, which is handy when you have a long module
     name and want to save some typing.
 [9]: import math as m
[10]: m.pi
[10]: 3.141592653589793
[11]: pi = 10
[12]: m.pi
[12]: 3.141592653589793
     from module_name import something
[13]: dir() # get a list of names comprising the attributes of the current namespac
[13]: ['In',
       'Out',
       '_',
       '_10',
       '_12',
       '_6',
       '_7',
       '_8',
       '__builtin__',
       '__builtins__',
```

'\_\_doc\_\_',
'\_\_loader\_\_',
'\_\_name\_\_',
'\_\_package\_\_',
'\_\_spec\_\_',
'\_dh',
'\_i',

```
'_i1',
       '_i10',
       '_i11',
       -
'_i12',
       '_i13',
       '_i2',
       '_i3',
       '_i4',
       '_i5',
       '_i6',
       '_i7',
       '_i8',
       '_i9',
       '_ih',
       '_ii',
       '_iii',
       '_oh',
       'exit',
       'get_ipython',
       'm',
       'math',
       'open',
       'pi',
       'quit']
[14]: from math import floor
[15]: help(floor)
     Help on built-in function floor in module math:
     floor(x, /)
          Return the floor of x as an Integral.
          This is the largest integer \leftarrow x.
[16]: math.floor(4.1)
[16]: 4
[17]: floor(4.1)
[17]: 4
[18]: dir()
```

```
[18]: ['In',
         'Out',
         '_',
         '_10',
         '_12',
         '_13',
         '_16',
         '_17',
         '_6',
         '_7',
         '_8',
         '__',
'___',
         '__builtin__',
'__builtins__',
        '__doc__',
'__loader__',
'__name__',
'__package__',
'__spec__',
         '_dh',
         '_i',
         '_i1',
         '_i10',
         '_i11',
         '_i12',
         '_i13',
         '_i14',
         _
'_i15',
         '_i16',
         '_i17',
         '_i18',
         '_i2',
         '_i3',
         '_i4',
         '_i5',
         '_i6',
         '_i7',
         _
'_i8',
         '_i9',
         '_ih',
         '_ii',
         '_iii',
         '_oh',
         'exit',
         'floor',
         'get_ipython',
```

```
'm',
        'math',
        'open',
        'pi',
        'quit']
[19]: from math import * # import everything in the math module, not recommended
[20]: dir()
[20]: ['In',
        'Out',
        '_',
        '_10',
        '_12',
        '_13',
        '_16',
        '_17',
        '_18',
        '_6',
        '_7',
        '_8',
        '__',
        '__builtin__',
        '__builtins__',
'__doc__',
'__loader__',
'__name__',
'__package__',
        '__spec__',
'_dh',
        '_i',
        '_i1',
        '_i10',
        '_i11',
        '_i12',
        '_i13',
        '_i14',
        '_i15',
        '_i16',
        '_i17',
        '_i18',
        '_i19',
        '_i2',
        '_i20',
        '_i3',
```

```
'_i4',
'_i5',
'_i6',
'_i7',
'_i8',
'_i9',
'_ih',
'_ii',
_
'_iii',
'_oh',
'acos',
'acosh',
'asin',
'asinh',
'atan',
'atan2',
'atanh',
'ceil',
'comb',
'copysign',
'cos',
'cosh',
'degrees',
'dist',
'e',
'erf',
'erfc',
'exit',
'exp',
'expm1',
'fabs',
'factorial',
'floor',
'fmod',
'frexp',
'fsum',
'gamma',
'gcd',
'get_ipython',
'hypot',
'inf',
'isclose',
'isfinite',
'isinf',
'isnan',
'isqrt',
'lcm',
```

```
'ldexp',
       'lgamma',
       'log',
       'log10',
       'log1p',
       'log2',
       'm',
       'math',
       'modf',
       'nan',
       'nextafter',
       'open',
       'perm',
       'pi',
       'pow',
       'prod',
       'quit',
       'radians',
       'remainder',
       'sin',
       'sinh',
       'sqrt',
       'tan',
       'tanh',
       'tau',
       'trunc',
       'ulp']
[21]: import math
[22]: math.pi
[22]: 3.141592653589793
[23]: dir()
[23]: ['In',
       'Out',
       '_',
       '_10',
       '_12',
       '_13',
       '_16',
       '_17',
       '_18',
       '_20',
       '_22',
```

```
'_6',
'_7',
'_8',
'__',
'__',
'__builtin__',
'__builtins__',
'__doc__',
'__loader__',
__name__',
'__package__',
'__spec__',
'_dh',
'_i',
'_i1',
'_i10',
'_i11',
'_i12',
'_i13',
'_i14',
_
'_i15',
'_i16',
'_i17',
'_i18',
-
'_i19',
'_i2',
'_i20',
-
'_i21',
'_i22',
'_i23',
'_i3',
'_i4',
'_i5',
'_i6',
-
'_i7',
'_i8',
'_i9',
'_ih',
'_ii',
'_iii',
'_oh',
'acos',
'acosh',
'asin',
'asinh',
'atan',
'atan2',
```

```
'atanh',
'ceil',
'comb',
'copysign',
'cos',
'cosh',
'degrees',
'dist',
'e',
'erf',
'erfc',
'exit',
'exp',
'expm1',
'fabs',
'factorial',
'floor',
'fmod',
'frexp',
'fsum',
'gamma',
'gcd',
'get_ipython',
'hypot',
'inf',
'isclose',
'isfinite',
'isinf',
'isnan',
'isqrt',
'lcm',
'ldexp',
'lgamma',
'log',
'log10',
'log1p',
'log2',
'm',
'math',
'modf',
'nan',
'nextafter',
'open',
'perm',
'pi',
'pow',
'prod',
```

```
'quit',
'radians',
'remainder',
'sin',
'sinh',
'sqrt',
'tan',
'tau',
'tau',
'trunc',
'ulp']
```

#### 3.2.1 Issues with from math import \*

While this approach might seem like a good idea, it can create namespace clashes which arise if the module has objects with names that are identitical to those in the current namespace. In other words if we already had a floor object in our namespace, it is now overwritten and points to the object in the math module and not the original floor object.

So using the explicit module name or alias to protect the namespace is generally **good practice**. It also lets us reuse the same name across different modules.

#### 3.2.2 Further learning about python standard modules

- There are many more python standard modules that are very useful.
  - Follow the following tutorials/documentations to learn more about them:
    - \* Brief Tour of the Standard Library
    - \* Python 3 Module of the Week
    - \* Python Built-in Modules from tutorialsteacher
- These standard python modules are building blocks of many external modules and projects

```
'08.1_Strings_Lists.ipynb',
       '15.2_pandas.ipynb',
       '08.2_Lists.ipynb',
       'Untitled.ipynb',
       '10.2_Sets_Dictionaries.ipynb',
       '02.2_Program-Variables-Operators.ipynb',
       '14.1_Mapping.ipynb',
       '02.1_Installation-Notebook-GitHub.ipynb',
       '04.2_ScalarDataTypes.ipynb',
       'pics',
       '14.2_Numpy.ipynb',
       '13.1_00P(2).ipynb',
       'untitled.txt',
       '04_Git.ipynb',
       '10.1_Lists_Tuples.ipynb',
       '11.1_Functions(2).ipynb',
       '.ipynb_checkpoints',
       '04.1_Functions.ipynb',
       'data',
       '15.1_Matplotlib.ipynb',
       '07.2_Strings.ipynb',
       '14.1_Scripts_Modules.ipynb',
       '06.2_Conditionals_Strings.ipynb',
       '14.1 OOP(3).ipynb']
[27]: os.listdir('/Users/wk0110/My Drive (weikang9009@gmail.com)/teaching')
[27]: ['Advanced GIS',
       'grading',
       '.DS_Store',
       'Geospatial Technology and Urban Environments',
       'resources',
       'Canvas.docx',
       'Course design.docx',
       'Urban and geography materials.docx',
       'Geographic_data_analysis',
       'Intro to Python',
       'rubrics',
       'Teaching and Learning with Jupyter.docx',
       'course development',
       'Data science',
       'Email to students',
       'UNT',
       'Scaffolding group project.docx',
       'Spatial data science',
       'Intro to GIS']
```

#### 3.3 Writing our own module

We will create a module and include the function that was used to Check whether two words (strings) are the reverse of each other (Part B of HW4) in the module. Our module will be named hw4.py and the function will be called is\_reverse and take two arguments:

- string1
- string2

It will return True or False.

From the Jupyter Notebook Dashboard, click "New"-"Text File" to create the file hw4.py so that its contents are as follows (You may copy your solution to Part B of HW4 to this file instead)::

```
[28]: def is reverse(string1, string2):
        ⇔reverse of the other
        Parameters
        _____
        string1: str
                 A string containing one word
        string2: str
                 A string containing one word
        Return
               : bool
                 True if string1 is the reverse of string2; Otherwise, False
        11 11 11
        length1 = len(string1)
        length2 = len(string2)
        if length1 != length2:
            return False
        for i in range(length1):
            if string1[i] != string2[-1-i]:
               return False
        return True
```

Importing the module gives access to its objects, using the module.object syntax

Now we can use our module by importing the function from the module:

```
[29]: from hw4 import is_reverse
```

ModuleNotFoundError
Cell In[29], line 1

Traceback (most recent call last)

```
----> 1 from hw4 import is_reverse
       ModuleNotFoundError: No module named 'hw4'
 []: dir()
[30]: is_reverse("Happy", "Happy")
[30]: False
[31]: is_reverse("Happy", "yppaH")
[31]: True
     Or we import the module:
[36]: import hw4
[37]: hw4.is_reverse("ha","ds")
[37]: False
[38]: hw4?
[39]: dir(hw4)
[39]: ['__builtins__',
       '__cached__',
       '__doc__',
       '__file__',
       '__loader__',
       '__name__',
'__package__',
       '__spec__',
       'is_reverse']
[40]: hw4.is_reverse("Happy", "Happy")
[40]: False
[41]: hw4.is_reverse("Happy", "yppaH")
[41]: True
```

#### 3.3.1 Documenting your module with Docstrings

- It is good practice to document your modules so that users know what and how to use the objects in the module.
- A Python docstring is a string used to document a Python module, class, function or method, so programmers can understand what it does without having to read the details of the implementation.
- It is a common practice to generate online (html) documentation automatically from docstrings.
- Everything between the pair of *triple quotes* is our function's **docstring**.
- There are some standard conventions for writing docstrings that provide consistency.

```
[42]: hw4.is_reverse?

[43]: help(hw4.is_reverse)
```

Help on function is\_reverse in module hw4:

is\_reverse(string1, string2)
 compare two words (strings) and return True if one of the words is the
reverse of the other

Parameters
----string1 : str

A string containing one word

string2 : str

A string containing one word

Return

: bool

True if string1 is the reverse of string2; Otherwise, False

#### 3.3.2 Locating your module

When issuing an import statement Python searches a list of directories:

- The directory containing the input script (or the current directory when no file is specified).
- PYTHONPATH (a list of directory names, with the same syntax as the shell variable PATH).
- The installation-dependent default (by convention including a site-packages directory, handled by the site module).

Over time you will likely be accumutating more of your own modules, and rather than continuing to copy them to new working directories (and have multiple copies of the same file on your system) you can create your own central directory to contain your Python modules. How you do this depends on what operating system you are on. The Python Documentation wesbite provide

• instructions for windows

• instructions for Unix type systems (including Mac)

## 4 Python Packages

- Suitable for a large application that includes many modules
- Allow for a hierarchical structuring of the *module* namespace using dot notation .. In the same way that modules help avoid collisions between global variable names, packages help avoid collisions between module names.
- special file called \_\_init\_\_.py (which may be empty) tells Python that the directory is a Python package, from which modules can be imported.

## 4.1 Importing a package

Given this structure, if the pkg directory resides in a location where it can be found (in one of the directories contained in sys.path), you can refer to the two modules with dot notation (pkg.mod1, pkg.mod2) and import them with the syntax you are already familiar with:

```
import pkg.mod1, pkg.mod2
```

If mod1 has a function bar() and mod2 has a function foo(), we can import and call these functions as shown below:

```
import pkg.mod1, pkg.mod2
     pkg.mod1.bar()
     pkg.mod2.foo()
     Or
     from pkg.mod1 import bar
     bar()
     from pkg.mod2 import foo
     foo()
[44]: import scipy
[45]: scipy.__file__ #location of the package on your computer
[45]: '/Users/wk0110/opt/anaconda3/lib/python3.9/site-packages/scipy/__init__.py'
[46]: scipy.__version__ #version of the package
[46]: '1.9.3'
[47]: dir(scipy)
[47]: ['LowLevelCallable',
       '__numpy_version__',
       '__version__',
       'cluster',
       'fft',
```

```
'fftpack',
       'integrate',
       'interpolate',
       'io',
       'linalg',
       'misc',
       'ndimage',
       'odr',
       'optimize',
       'show_config',
       'signal',
       'sparse',
       'spatial',
       'special',
       'stats',
       'test']
[48]: from scipy import stats
[49]: dir(stats)
[49]: ['ConstantInputWarning',
       'DegenerateDataWarning',
       'FitError',
       'NearConstantInputWarning',
       'NumericalInverseHermite',
       '__all__',
       '__builtins__',
       '__cached__',
       '__doc__',
       '__file__',
       '__loader__',
       '__name__',
'__package__',
       '__path__',
       '__spec__',
       '_axis_nan_policy',
       '_biasedurn',
       '_binned_statistic',
       '_binomtest',
       '_boost',
       '_common',
       '_constants',
       '_continuous_distns',
       '_crosstab',
       '_discrete_distns',
       '_distn_infrastructure',
```

```
'_distr_params',
'_entropy',
'_fit',
'_hypotests',
'_hypotests_pythran',
'_kde',
'_ksstats',
'_levy_stable',
'_mannwhitneyu',
'_morestats',
'_mstats_basic',
'_mstats_extras',
'_multivariate',
'_mvn',
'_page_trend_test',
'_qmc',
'_qmc_cy',
'_relative_risk',
'_resampling',
'_rvs_sampling',
'_sobol',
'_statlib',
'_stats',
'_stats_mstats_common',
'_stats_py',
'_tukeylambda_stats',
'_unuran',
'_variation',
'_warnings_errors',
'alexandergovern',
'alpha',
'anderson',
'anderson_ksamp',
'anglit',
'ansari',
'arcsine',
'argus',
'barnard_exact',
'bartlett',
'bayes_mvs',
'bernoulli',
'beta',
'betabinom',
'betaprime',
'biasedurn',
'binned_statistic',
'binned_statistic_2d',
```

```
'binned_statistic_dd',
'binom',
'binom_test',
'binomtest',
'boltzmann',
'bootstrap',
'boschloo_exact',
'boxcox',
'boxcox_llf',
'boxcox_normmax',
'boxcox_normplot',
'bradford',
'brunnermunzel',
'burr',
'burr12',
'cauchy',
'chi',
'chi2',
'chi2_contingency',
'chisquare',
'circmean',
'circstd',
'circvar',
'combine_pvalues',
'contingency',
'cosine',
'cramervonmises',
'cramervonmises_2samp',
'crystalball',
'cumfreq',
'describe',
'dgamma',
'differential_entropy',
'dirichlet',
'distributions',
'dlaplace',
'dweibull',
'energy_distance',
'entropy',
'epps_singleton_2samp',
'erlang',
'expon',
'exponnorm',
'exponpow',
'exponweib',
'f',
'f_oneway',
```

```
'fatiguelife',
'find_repeats',
'fisher_exact',
'fisk',
'fit',
'fligner',
'foldcauchy',
'foldnorm',
'friedmanchisquare',
'gamma',
'gausshyper',
'gaussian_kde',
'genexpon',
'genextreme',
'gengamma',
'genhalflogistic',
'genhyperbolic',
'geninvgauss',
'genlogistic',
'gennorm',
'genpareto',
'geom',
'gibrat',
'gilbrat',
'gmean',
'gompertz',
'gstd',
'gumbel_1',
'gumbel_r',
'gzscore',
'halfcauchy',
'halfgennorm',
'halflogistic',
'halfnorm',
'hmean',
'hypergeom',
'hypsecant',
'invgamma',
'invgauss',
'invweibull',
'invwishart',
'iqr',
'jarque_bera',
'johnsonsb',
'johnsonsu',
'kappa3',
'kappa4',
```

```
'kde',
'kendalltau',
'kruskal',
'ks_1samp',
'ks_2samp',
'ksone',
'kstat',
'kstatvar',
'kstest',
'kstwo',
'kstwobign',
'kurtosis',
'kurtosistest',
'laplace',
'laplace_asymmetric',
'levene',
'levy',
'levy_l',
'levy_stable',
'linregress',
'loggamma',
'logistic',
'loglaplace',
'lognorm',
'logser',
'loguniform',
'lomax',
'mannwhitneyu',
'matrix_normal',
'maxwell',
'median_abs_deviation',
'median_test',
'mielke',
'mode',
'moment',
'monte_carlo_test',
'mood',
'morestats',
'moyal',
'mstats',
'mstats_basic',
'mstats_extras',
'multinomial',
'multiscale_graphcorr',
'multivariate_hypergeom',
'multivariate_normal',
'multivariate_t',
```

```
'mvn',
'mvsdist',
'nakagami',
'nbinom',
'ncf',
'nchypergeom_fisher',
'nchypergeom_wallenius',
'nct',
'ncx2',
'nhypergeom',
'norm',
'normaltest',
'norminvgauss',
'obrientransform',
'ortho_group',
'page_trend_test',
'pareto',
'pearson3',
'pearsonr',
'percentileofscore',
'permutation_test',
'planck',
'pmean',
'pointbiserialr',
'poisson',
'power_divergence',
'powerlaw',
'powerlognorm',
'powernorm',
'ppcc_max',
'ppcc_plot',
'probplot',
'qmc',
'randint',
'random_correlation',
'rankdata',
'ranksums',
'rayleigh',
'rdist',
'recipinvgauss',
'reciprocal',
'relfreq',
'rice',
'rv_continuous',
'rv_discrete',
'rv_histogram',
'rvs_ratio_uniforms',
```

```
'scoreatpercentile',
'sem',
'semicircular',
'shapiro',
'siegelslopes',
'sigmaclip',
'skellam',
'skew',
'skewcauchy',
'skewnorm',
'skewtest',
'somersd',
'spearmanr',
'special_ortho_group',
'statlib',
'stats',
'studentized_range',
't',
'test',
'theilslopes',
'tiecorrect',
'tmax',
'tmean',
'tmin',
'trapezoid',
'trapz',
'triang',
'trim1',
'trim_mean',
'trimboth',
'truncexpon',
'truncnorm',
'truncweibull_min',
'tsem',
'tstd',
'ttest_1samp',
'ttest_ind',
'ttest_ind_from_stats',
'ttest_rel',
'tukey_hsd',
'tukeylambda',
'tvar',
'uniform',
'unitary_group',
'variation',
'vonmises',
'vonmises_line',
```

```
'wald',
 'wasserstein_distance',
 'weibull_max',
 'weibull_min',
 'weightedtau',
 'wilcoxon',
 'wishart',
 'wrapcauchy',
 'yeojohnson',
 'yeojohnson_llf',
 'yeojohnson_normmax',
 'yeojohnson_normplot',
 'yulesimon',
 'zipf',
 'zipfian',
 'zmap',
 'zscore']
stats.arcsine?
```

### 4.1.1 Installing python packages

- pip: pip install scipy
  - The Python Package Index (PyPI) is a repository of software for the Python programming language.
- conda: conda install scipy
  - Anaconda Packages

## 4.2 Further readings on creating a python package

Read the Real Python tutorial on Python packages to learn more about creating modules and packages.

Read the open source book Python Packages to learn more about the modern and efficient workflows for creating Python packages.