Python in 10 minutes

Part 9

Dr. Mark Williamson, PhD
Biostatistics, Epidemiology, and Research Design Core (BERDC)
Dakota Cancer Collaborative on Translational Activity (DaCCoTA)
University of North Dakota (UND)

Purpose:

- Quick, bite-size guides to basic usage and tasks in Python
- I’m no expert, I’ve just used it for various tasks, and it has made my life easier and allowed me to do things I couldn’t manually
- I’d like to share that working knowledge with you

Lesson 9: Library work

Last time, we tackled data transformation. Today, we'll step out into the wider world of Python add-ons. We'll examine Python modules, packages, and libraries. These include:

1) Built-in Add-ons (random, csv)

2) Built-in Add-ons 2 (collections, tkinter)

3) Third-party Add-ons (numpy, scipy, matplotlib)

4) Third-party Add-ons 2 (bokeh, pandas, pillow, requests)

Lesson 9: Add-on Definitions

Module:

- simple Python file that contains collections of functions and global variables
- code grouped together according to a particular defined function

Package:

- simple directory that contains collections of modules
- compiled pieces of code created by a third party to automate common tasks

Library:

- collection of related, reusable chucks of code that are imported and called with methods as needed
- the standard Python library includes functions, classes, objects, data types, etc.

Take Home Point: a module is a collection of code, a package is a collection of modules, and a library is a collection of packages

- https://www.geeksforgeeks.org/what-is-the-difference-between-pythons-module-package-and-library/
- https://thecleverprogrammer.com/2021/01/23/libraries-packages-and-modules-in-python/

Lesson 9: Add-on Installation

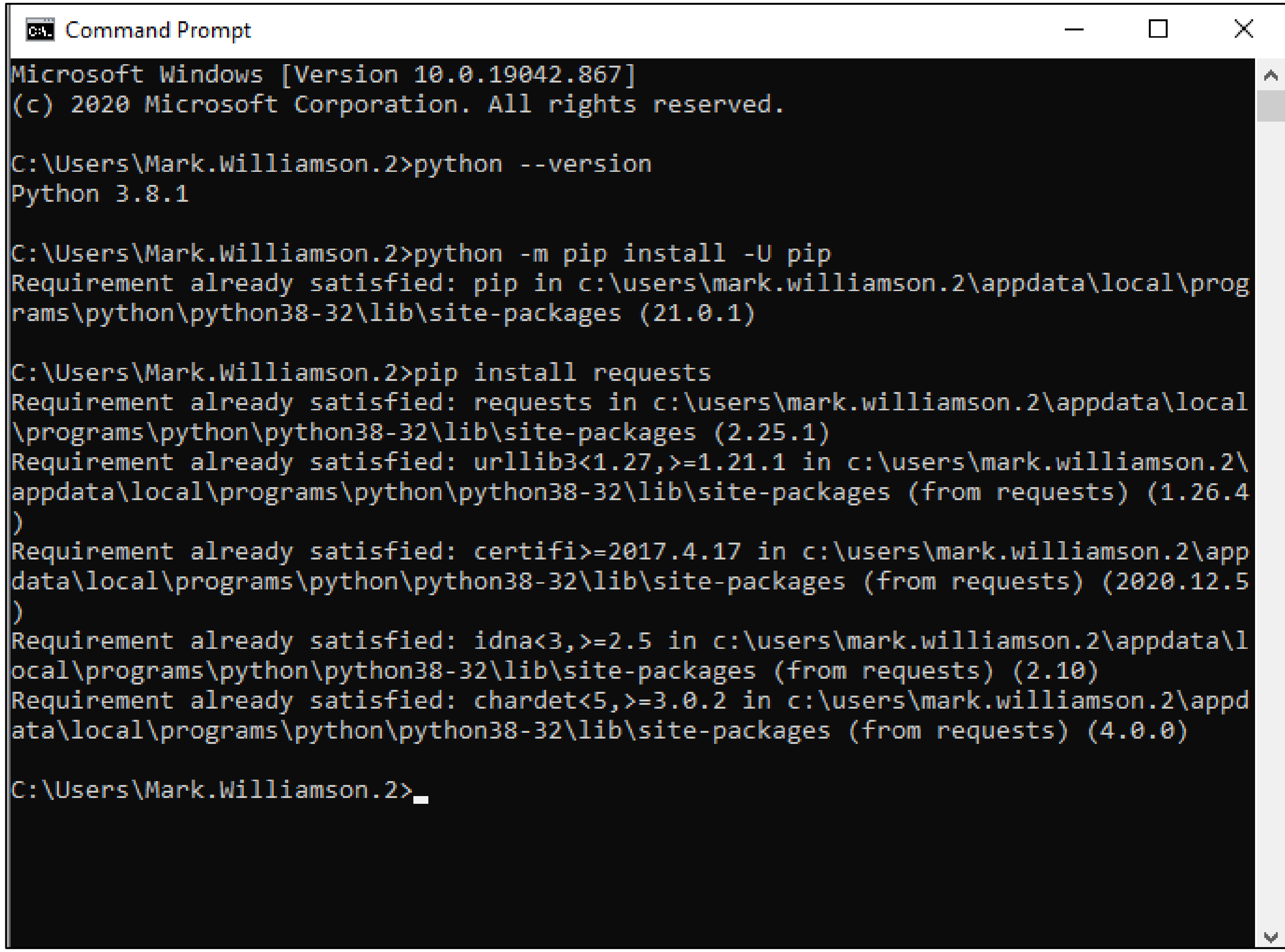
Command line:

- Text-based interface to install add-ons
- Accessing the command line:
  - Windows: search ‘command prompt’
  - Apple: search ‘terminal’
  - Linux: search ‘terminal’
- Check python version (Windows):

  >>>python --version
- Install pip if not already installed:
  - PIP is short for ‘PIP Installs Packages’
  - https://www.makeuseof.com/tag/install-pip-for-python/
- Upgrade pip (Windows, optional):

  >>>python -m pip install -U pip
- Install add-ons (Windows):

  >>>pip install *package_name*



```
Microsoft Windows [Version 10.0.19042.867]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\Users\Mark.Williamson.2>python --version
Python 3.8.1

C:\Users\Mark.Williamson.2>python -m pip install -U pip
Requirement already satisfied: pip in c:\users\mark.williamson.2\appdata\local\prog
rams\python\python38-32\lib\site-packages (21.0.1)

C:\Users\Mark.Williamson.2>pip install requests
Requirement already satisfied: requests in c:\users\mark.williamson.2\appdata\local
\programs\python\python38-32\lib\site-packages (2.25.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\mark.williamson.2\
appdata\local\programs\python\python38-32\lib\site-packages (from requests) (1.26.4
)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\mark.williamson.2\app
data\local\programs\python\python38-32\lib\site-packages (from requests) (2020.12.5
)
Requirement already satisfied: idna<3,>=2.5 in c:\users\mark.williamson.2\appdata\l
ocal\programs\python\python38-32\lib\site-packages (from requests) (2.10)
Requirement already satisfied: chardet<5,>=3.0.2 in c:\users\mark.williamson.2\appd
ata\local\programs\python\python38-32\lib\site-packages (from requests) (4.0.0)

C:\Users\Mark.Williamson.2>_
```

Lesson 9: Built-in Add-ons

Exploration

Write your own csv file using as many rows and items in each row as you like. You could even change the delimiter from a comma to something else (like a tab).

Description:

- random: generate data randomly
- csv: handle csv files

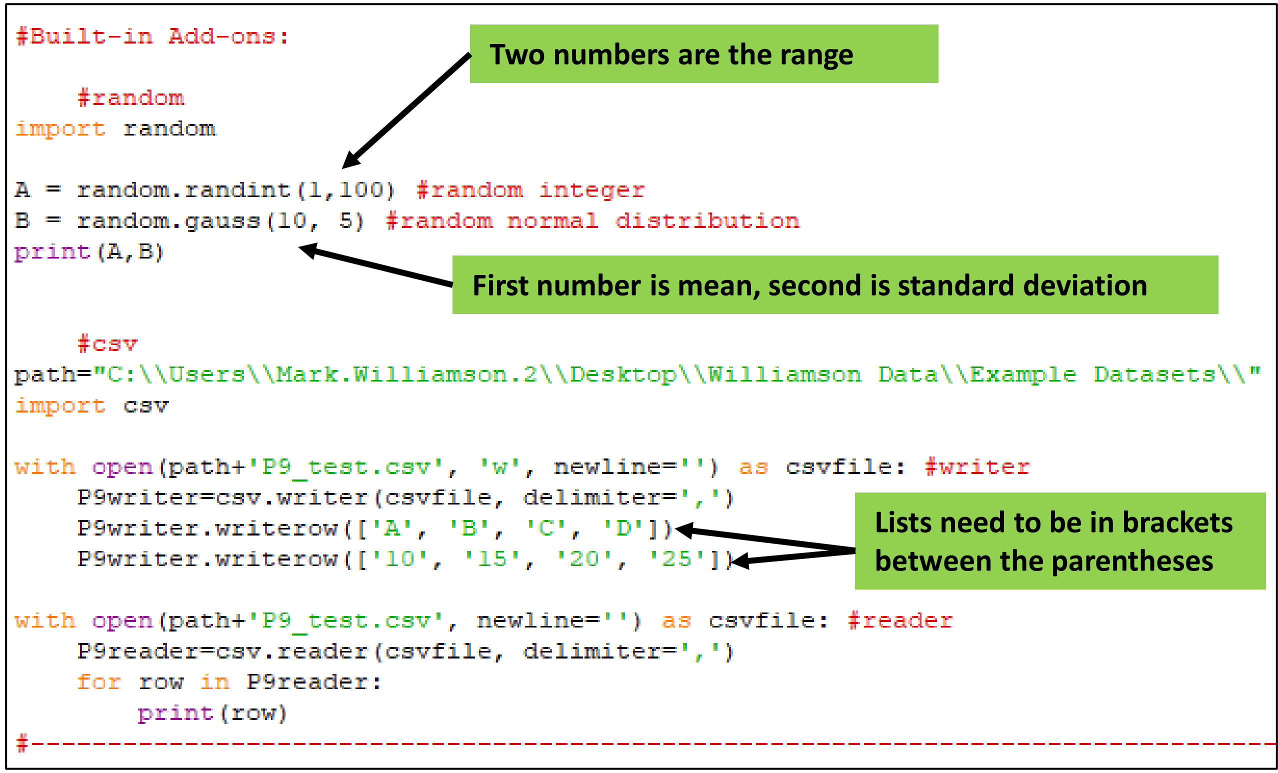
Procedure

#random

- Import **random**
- Create a variable called **A** and set it to a random integer using **random.randint()** function
- Create a variable called **B** and set it to a normally distributed number using **random.guass()** function
- Print both numbers

#csv

- Set a path to a folder of your choice
- Import csv
- Write a csv file called **P9_test.csv** and write two rows to it using **csv.writer()** function and **.writerow()** method
- Read the created csv file and print each row



```
#Built-in Add-ons:

    #random
import random

A = random.randint(1,100) #random integer
B = random.gauss(10, 5) #random normal distribution
print(A,B)


    #csv
path="C:\\Users\\Mark.Williamson.2\\Desktop\\Williamson Data\\Example Datasets\\"
import csv

with open(path+'P9_test.csv', 'w', newline='') as csvfile: #writer
    P9writer=csv.writer(csvfile, delimiter=',')
    P9writer.writerow(['A', 'B', 'C', 'D'])
    P9writer.writerow(['10', '15', '20', '25'])

with open(path+'P9_test.csv', newline='') as csvfile: #reader
    P9reader=csv.reader(csvfile, delimiter=',')
    for row in P9reader:
        print(row)
#--------------------------------------------------------------------------------
```

Two numbers are the range

First number is mean, second is standard deviation

Lists need to be in brackets between the parentheses

Illumination

Remember, functions take the form of *functionName(arguments)*, while methods take the form of *variable.methodName(aguments)*

Lesson 9: Built-in Add-ons 2

Exploration

Tkinter can make GUIs (graphical user interfaces) with all sorts of features. Try including buttons, canvases, entry forms, labels, menus, messages, etc. to your popup window.

Description:

- collections: storing collections of data
- tkinter: GUI development

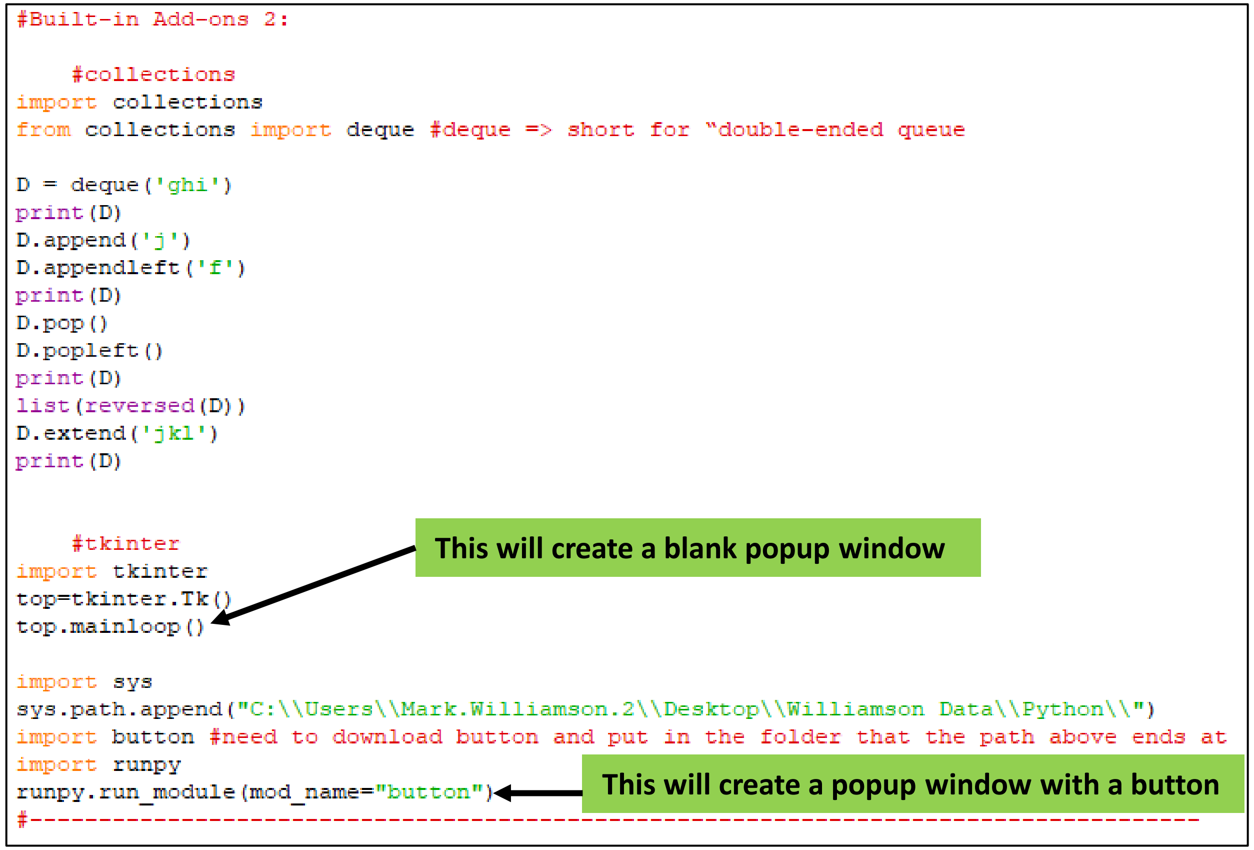
Procedure

#collections

- Import **collections** and **deque**
- Create a variable **D** and set it to a deque of ‘ghi’, then print
- add and remove items from the left and right of **D**, printing each time
- List **D** in reverse order, extend **D** by ‘jkl’, then print

#tkinter

- Import **tkinter**
- Create a variable **top** and set it to the function **tkinter.Tk()**
- Run the method **.mainloop()** on **top**
- Download the file **button.txt** and save as a python (.py) file in a folder
    - https://med.und.edu/daccota/_files/docs/berdc_docs/button.txt
- Import sys, then use sys.path.append with the path to the folder of **button.py**
- Import **button** and **runpy**
- Use the function **runpy.run_module** to run **button**



```
#Built-in Add-ons 2:

    #collections
import collections
from collections import deque #deque => short for “double-ended queue

D = deque('ghi')
print(D)
D.append('j')
D.appendleft('f')
print(D)
D.pop()
D.popleft()
print(D)
list(reversed(D))
D.extend('jkl')
print(D)


    #tkinter
import tkinter
top=tkinter.Tk()
top.mainloop()

import sys
sys.path.append("C:\\Users\\Mark.Williamson.2\\Desktop\\Williamson Data\\Python\\")
import button #need to download button and put in the folder that the path above ends at
import runpy
runpy.run_module(mod_name="button")
#------------------------------------------------------------------------------------
```

This will create a blank popup window

This will create a popup window with a button

Illumination

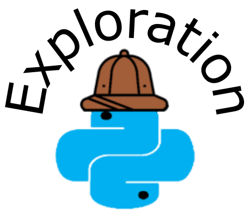
You will need to close the popup windows for the Python code to continue.

Lesson 9: Thirty-party Add-ons

Scipy.stats can run a variety of statistical functions, including ANOVA, Pearson correlation, T-tests, and chi-square tests,

Description:

- numpy: scientific computation
- scipy: scientific analysis
- matplotlib: 2D graph plotting



Try creating a scatter plot in matplotlib

Procedure

- *each add-on needs to be installed in command prompt first*

#numpy

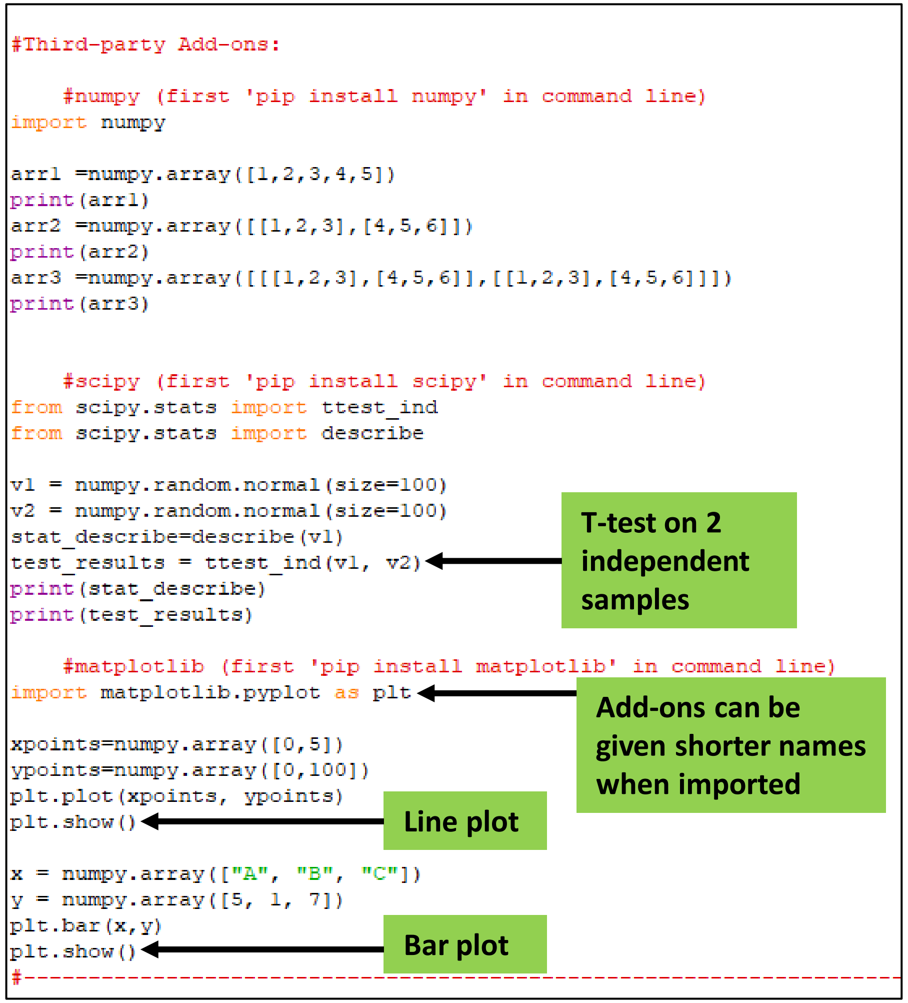
- Import **numpy**
- Create a 1D (**arr1**), 2D (**arr2**), and 3D (**arr3**) array and fill them with lists of numbers

#scipy

- Import **ttest_ind** and **describe** from **scipy.stats**
- Create two variables called **v1** and **v2**, each set to a list of 100 randomly distributed numbers
- Create a variable called **stat_describe** and set it to the description of **v1**
- Create a variable called **test_results** and set it to the t-test result between **v1** and **v2**

#matplotlib

- Import **matplotlib.pyplot** and set it to **plt**
- Create two numerical arrays (**xpoints**, **ypoints**), plot them (**plt.plot**) and display the plot (**plt.show**)
- Create a numerical and categorical array (**x**, **y**), plot them (**plt.bar**) and display the plot (**plt.show**)



```
#Third-party Add-ons:

    #numpy (first 'pip install numpy' in command line)
import numpy

arr1 =numpy.array([1,2,3,4,5])
print(arr1)
arr2 =numpy.array([[1,2,3],[4,5,6]])
print(arr2)
arr3 =numpy.array([[[1,2,3],[4,5,6]],[[1,2,3],[4,5,6]]])
print(arr3)


    #scipy (first 'pip install scipy' in command line)
from scipy.stats import ttest_ind
from scipy.stats import describe

v1 = numpy.random.normal(size=100)
v2 = numpy.random.normal(size=100)
stat_describe=describe(v1)
test_results = ttest_ind(v1, v2)
print(stat_describe)
print(test_results)

    #matplotlib (first 'pip install matplotlib' in command line)
import matplotlib.pyplot as plt

xpoints=numpy.array([0,5])
ypoints=numpy.array([0,100])
plt.plot(xpoints, ypoints)
plt.show()

x = numpy.array(["A", "B", "C"])
y = numpy.array([5, 1, 7])
plt.bar(x,y)
plt.show()
#----------------------------------------------------------------------
```

T-test on 2 independent samples

Add-ons can be given shorter names when imported

Line plot

Bar plot

Lesson 9: Thirty-party Add-ons

Illumination

The examples here just scratch the surface of the abilities of these add-ons

Description:

- bokeh: data visualization
- pillow (PIL): image manipulation
- pandas: data analysis
- requests: sending HTTP requests

Exploration

Try another image of your choice. Perhaps '*Oath of the Horatti*' by the same artist?

Procedure

#bokeh

- Import **output_file** and **show** from **bokeh.io** and **figure** from **bokeh.plotting**
- Set **plot** to the function **figure()** and show the **plot** with the label 'figure.html'
- Set **plot.line** to an x array and y array and the updated **plot** with the label 'line_glyph.html'

#pandas

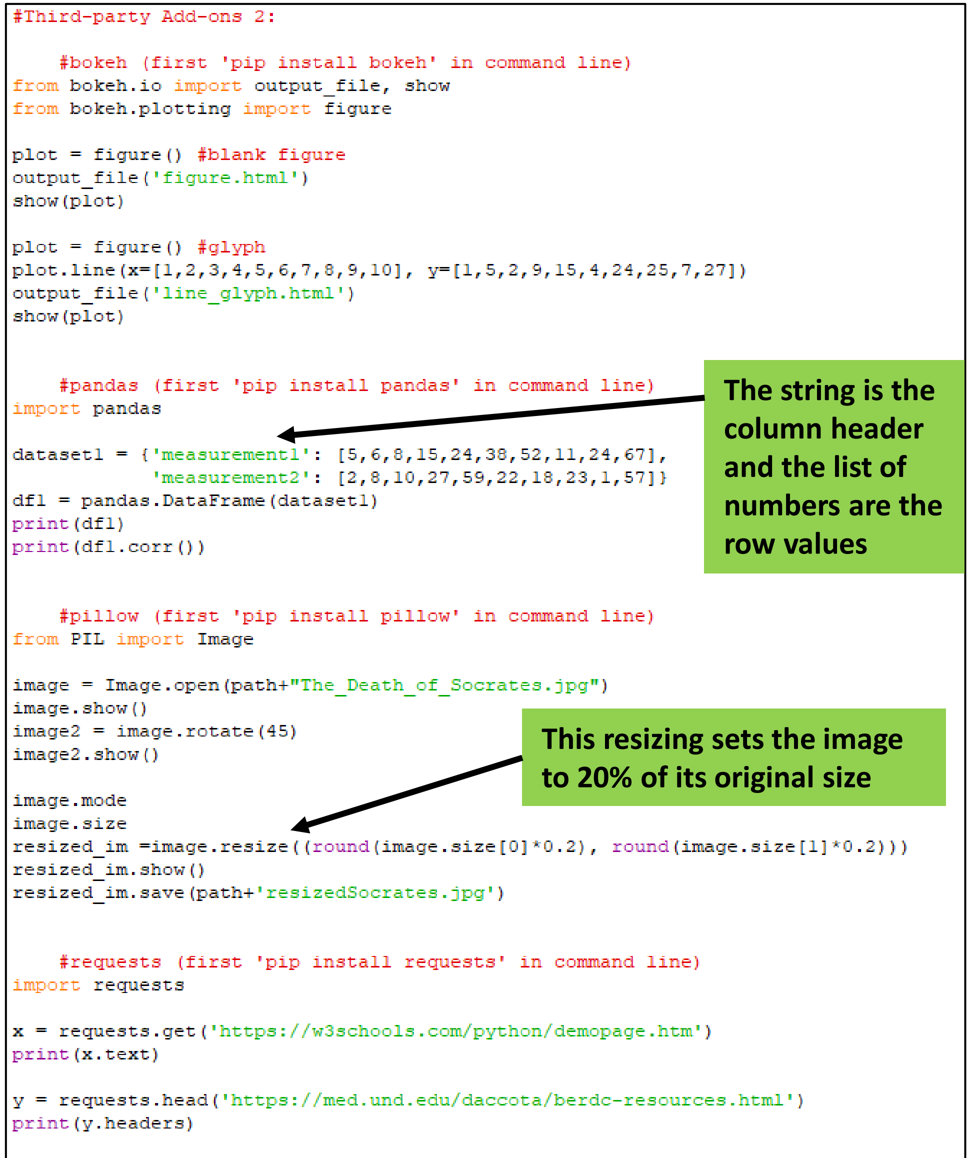
- Import **pandas**
- Create a dataset called **dataset1**, set it as a data frame called **df1**, then print both **df1** and its correlation value

#pillow

- Import **Image** from **PIL**, download the jpeg image below, rename it, and put in the path folder
  - https://upload.wikimedia.org/wikipedia/commons/8/8c/David_-_The_Death_of_Socrates.jpg
- Use the methods **.show, .rotate, .mode, .size, .resize,** and **.save** to manipulate the image

#requests

- Import **requests**, set the variables **x** and **y** to HTPP requests, and print the text or header of those request variables



```
#Third-party Add-ons 2:

    #bokeh (first 'pip install bokeh' in command line)
from bokeh.io import output_file, show
from bokeh.plotting import figure

plot = figure() #blank figure
output_file('figure.html')
show(plot)

plot = figure() #glyph
plot.line(x=[1,2,3,4,5,6,7,8,9,10], y=[1,5,2,9,15,4,24,25,7,27])
output_file('line_glyph.html')
show(plot)


    #pandas (first 'pip install pandas' in command line)
import pandas

dataset1 = {'measurement1': [5,6,8,15,24,38,52,11,24,67],
            'measurement2': [2,8,10,27,59,22,18,23,1,57]}
df1 = pandas.DataFrame(dataset1)
print(df1)
print(df1.corr())


    #pillow (first 'pip install pillow' in command line)
from PIL import Image

image = Image.open(path+"The_Death_of_Socrates.jpg")
image.show()
image2 = image.rotate(45)
image2.show()

image.mode
image.size
resized_im =image.resize((round(image.size[0]*0.2), round(image.size[1]*0.2)))
resized_im.show()
resized_im.save(path+'resizedSocrates.jpg')


    #requests (first 'pip install requests' in command line)
import requests

x = requests.get('https://w3schools.com/python/demopage.htm')
print(x.text)

y = requests.head('https://med.und.edu/daccota/berdc-resources.html')
print(y.headers)
```

The string is the column header and the list of numbers are the row values

This resizing sets the image to 20% of its original size

Lesson 9: Summary

- There is a bewildering variety of both built-in and third-party add-ons for Python
- Add-ons need to be imported before use and third-party ones need to be installed
- Please complete a brief assessment: https://und.qualtrics.com/jfe/form/SV_5os01AOLAuW4zhs

References:

- https://docs.python.org/3/library/collections.html
- https://docs.python.org/3/library/csv.html
- https://docs.python.org/3/library/random.html
- https://www.w3schools.com/python/module_requests.asp
- https://towardsdatascience.com/getting-started-with-bokeh-effortlessly-elegant-interactive-data-visualisations-in-python-703249565bb3
- https://www.w3schools.com/python/numpy_intro.asp
- https://www.w3schools.com/python/scipy_intro.asp
- https://www.w3schools.com/python/matplotlib_intro.asp
- https://www.w3schools.com/python/pandas/default.asp
- https://www.tutorialspoint.com/python_pillow/index.htm
- https://www.tutorialspoint.com/python/python_gui_programming.htm
- https://www.tutorialspoint.com/How-to-import-a-Python-module-given-the-full-path
- https://pythonbasics.org/tkinter/