# STAT606 Computing for Data Science and Statistics

Lecture 1: Introduction to Python

# Python: Overview



Python is a **dynamically typed**, **interpreted** programming language Created by Guido van Rossum in 1991 Maintained by the Python Software Foundation

Design philosophy: simple, readable code

Python syntax differs from R, Java, C/C++, MATLAB whitespace delimited limited use of brackets, semicolons, etc

# Python: Overview

python<sup>™</sup>

Python is a **dynamically typed**, **nterpreted** programming language Created by Guido van Rossum (\* 1991 Maintained by the Python Software Foundation

Design philosophy: simple, readable code

Python syntax differs from R, Java, C/C++, MATLAB whitespace delimited limited use of brackets, semicolons, etc In many languages, when you declare a variable, you must specify the variable's **type** (e.g., int, double, Boolean, string). Python does not require this.

# Python: Overview

Python is a **dynamically typed interpreted** programming language Created by Guido van Rossum in 1991 Maintained by the Python Software Foundation

Design philosophy: simple, readable code

Python syntax differs from R, Java, C/C++, MATLAB whitespace delimited limited use of brackets, semicolons, etc

**Compiled vs interpreted languages:** compiled languages are (generally) faster than interpreted languages, typically at the cost of being more complicated.

Some languages (e.g., C/C++ and Java) are **compiled**: we write code, from which we get a runnable program via **compilation**. In contrast, Python is **interpreted**: A program, called the **interpreter**, runs our code directly, line by line.

python

Image credit: https://www.python.org/community/logos/

# **Running Python**



Several options for running Python on your computer

Python interpreter

Jupyter: <u>https://jupyter.org/</u>

PythonAnywhere: https://www.pythonanywhere.com/

Suggestions from Allen Downey:

https://www.allendowney.com/wp/books/think-python-2e/

Your homeworks must be handed in as Jupyter notebooks

But you should also be comfortable with the interpreter and running Python on the command line

Installing Jupyter: https://jupyter.readthedocs.io/en/latest/install.html

Note: Jupyter recommends Anaconda: https://www.anaconda.com/

I mildly recommend against Anaconda, but it's your choice

Image credit: https://www.python.org/community/logos/

#### Python Interpreter on the Command Line

```
keith@Steinhaus:~/demo$ python3
Python 3.11.2 (v3.11.2:878ead1ac1, Feb 7 2023, 10:02:41) [Clang 13.0.0
(clang-1300.0.29.30)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
keith@Steinhaus:~/demo$ python
Python 2.7.16 (default, Jun 5 2020, 22:59:21)
[GCC 4.2.1 Compatible Apple LLVM 11.0.3 (clang-1103.0.29.20)
(-macos10.15-objc- on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

#### Python Interpreter on the Command Line

Python 3 vs Python 2

keith@Steinhaus:~/demo\$ python3
Python 3.11.2 (v3.11.2:878eadlac1, Feb 7 2023, 10:02:41) [Clang 13.0.0
(clang-1300.0.29.30)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
keith@Steinhaus:~/demo\$ python
Python 2.7.16 (default, Jun 5 2020, 22:59:21)
[GCC 4.2.1 Compatible Apple LLVM 11.0.3 (clang-1103.0.29.20)
(-macos10.15-objc- on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>

The **prompt** indicates that the system is waiting for your input.

#### Python Interpreter on the Command Line



# Python in Jupyter



Creates "notebook files" for running Julia, Python and R

Example notebook:

https://nbviewer.jupyter.org/github/jrjohansson/

scientific-python-lectures/blob/master/Lecture-4-Matplotlib.ipynb

Clean, well-organized presentation of code, text and images, in one document

Installation: https://jupyter.readthedocs.io/en/latest/install.html

Documentation on running: <u>https://jupyter.readthedocs.io/en/latest/running.html</u> Good tutorials:

https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/execute.html

## **Running Jupyter**

keith@Steinhaus:~/demo\$ jupyter notebook

[I 17:11:41.129 NotebookApp] Serving notebooks from local directory: /Users/keith/Dropbox/Academe/Teaching/STAT606/Lecs/L1\_AdminIntro [I 17:11:41.129 NotebookApp] 0 active kernels [I 17:11:41.129 NotebookApp] The Jupyter Notebook is running at: http://localhost:8888/?token=452d6d4b227f306f5bb57e72f5d4722fcbadf47d1d794441 [I 17:11:41.129 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation). [C 17:11:41.132 NotebookApp]

Copy/paste this URL into your browser when you connect for the first time, to login with a token:

http://localhost:8888/?token=452d6d4b227f306f5bb57e72f5d4722fcbadf47d1d794441
[I 17:11:41.635 NotebookApp] Accepting one-time-token-authenticated connection from
::1

Jupyter provides some information about its startup process, and then...

# Running Jupyter

er opens a browser window n you can launch a new	
ok or open an existing one.	💵 🛐 🗊 🤌
Upload Net	N - 2
	Upload Nev Name <b>↑</b> Last Modi











I'll leave it to you to learn about the other features by reading the documentation. For now, the green-highlighted box is most important. That's where we write Python code.



Write c then pr code in	ode in the highlighted box, ess shift+enter to run the that box	In [2]: 1 print('Hello world') Hello world	
C Home	X C Untitled X	+	
ŵ	i løcalhost:8888/notebooks/Untitled.ipynb?kern	nel_name=python3 ···· 🛡 🏠 🔍 Search 👱 💷	ABP S
💢 jupyter	Untitled Last Checkpoint: an hour ago (autosaved)		Logout
File Edit	View Insert Cell Kernel Widgets He	Help Trusted I Pyth	ion 3 O
B + x 6	2 🖪 🛧 🔸 🕅 🔳 C Code 🗾		
In [2]:	<pre>1 print('Hello world') Hello world</pre>		
In [ ]:	1		



Write c then pr code in	ode in the highlighted box, ress shift+enter to run the n that box	In [2]: He In []:	1 print('Hello world' ello world	)
C Home	X 📿 Untitled X	+		
ŵ	i løcalhost:8888/notebooks/Untitled.ipynb?ke	rnel_name=python3	🚥 🔽 🏠 🔍 Search	🚽 III 🐵 🛽
Ċ jupyter	Untitled Last Checkpoint: an hour ago (autosaved	0		Cogout
File Edit	View Insert Cell Kernel Widgets	Help		Trusted 🖋 Python 3 C
B + × 0				
In [2]:	<pre>1 print('Hello world')</pre>			
	Hello world		Note: can also run cod	e by clicking the "run cell" ter shortcut is a lot easier
in []:	1		batton, but the shift en	

## Our first function: print

In [2]:	<pre>1 print('Hello world')</pre>	
Hello world		
In [ ]:	1	Print displays whatever is inside the quotation marks.

If you haven't already guessed, print takes a Python **string** and prints it. Of course, "print" here means to display a string, not literally print it on a printer!

**Note:** if you know Python 2, you'll notice that print is a bit different in Python 3. That is because in Python 2, print was a **statement**, whereas in Python 3, print is a **function**.

#### Can also use double quotes

Hello world



Hello world!

#### Arithmetic in Python



#### Data Types

Programs work with values, which come with different types

Examples:

The value 42 is an integer

The value 2.71828 is a floating point number (i.e., decimal number) The value "bird" is a string (i.e., a string of characters)

Variable's type determines what operations we can and can't perform e.g., 2\*3 makes sense, but what is `cat' \* `dog'? (We'll come back to this in more detail in a later lecture)

Variable is a name that refers to a value

Assign a value to a variable via variable assignment



Variable is a name that refers to a value

**Note:** unlike some languages (e.g., C/C++ and Java), you don't need to tell Python the type of a variable when you declare it. Instead, Python figures out the type of a variable automatically. This has the amusing name **duck typing**, which we will return to in a few lectures.

Assign a value to a variable via variable assignment



Variable is a name that refers to a value

**Note:** unlike some languages (e.g., C/C++ and Java), you don't need to tell Python the type of a variable when you declare it. Instead, Python figures out the type of a variable automatically. This has the amusing name **duck typing**, which we will return to in a few lectures.

Assign a value to a variable via variable assignment

```
1 mystring = 'Die Welt ist alles was der Fall ist.'
2 approx_pi = 3.141592
3 number_of_planets = 9
```

1 mystring

```
'Die Welt ist alles was der Fall ist.'
```

1 number\_of\_planets

```
9
```

```
1 number_of_planets = 8
2 number_of_planets
```

Python variable names can be arbitrarily long, and may contain any letters, numbers and underscore (\_), but may not start with a number. Variables can have any name, except for the Python 3 reserved keywords: None continue for lambda try True def from nonlocal while and del global not with as elif if or yield assert else import pass break except in raise

Sometimes we do need to know the type of a variable

Python type () function does this for us

```
1 mystring = 'Die Welt ist alles was der Fall ist.'
2 approx_pi = 3.141592
3 number_of_planets = 9
4 type(mystring)
```

str

1 type(approx_pi)	Recall that type is one of the Python reserved words. Syntax highlighting
float	shows it as green, indicating that it is a special word in Python.
1. And the second se	

1 type(number\_of\_planets)

int

**Note:** changing a variable to a different type is often called **casting** a variable to that type.

We can (sometimes) change the type of a Python variable

```
Convert a float to an int:
```

1	approx_pi =	3.141592
2	type(approx_	pi)

float

1 pi\_int = int(approx\_pi)
2 type(pi\_int)

int

1 pi\_int

Convert a string to an int:

1	int_from_s	str =	int('8675309'	)
2	type(int_	Erom_s	str)	

int

1 int\_from\_str

8675309

**Note:** changing a variable to a different type is often called **casting** a variable to that type.

We can (sometimes) change the type of a Python variable

```
Convert a float to an int:
```

```
1 approx_pi = 3.141592
2 type(approx_pi)
```

float

```
1 pi_int = int(approx_pi)
2 type(pi_int)
```

int



Test your understanding: what should be the value of float\_from\_int? Convert a string to an int:

```
1 int_from_str = int('8675309')
2 type(int_from_str)
```

int

1 int\_from\_str

8675309

1 float\_from\_int = float(42)
2 type(float from int)

**Note:** changing a variable to a different type is often called **casting** a variable to that type.

We can (sometimes) change the type of a Python variable

```
Convert a float to an int:
```

```
1 approx_pi = 3.141592
2 type(approx_pi)
```

float

```
1 pi_int = int(approx_pi)
2 type(pi_int)
```

int



Test your understanding: what should be the value of float\_from\_int? Convert a string to an int:

```
1 int_from_str = int('8675309')
2 type(int_from_str)
```

int

1 int\_from\_str

8675309

1 float\_from\_int = float(42)
2 type(float from int)

```
float
```

We can (sometimes) change the type of a Python variable

But if we try to cast to a type that doesn't make sense...

```
1 goat_int = int('goat')
ValueError Traceback (most recent call last)
<ipython-input-72-6ee721a55259> in <module>()
----> 1 goat_int = int('goat')
```

ValueError: invalid literal for int() with base 10: 'goat'

ValueError signifies that the type of a variable is okay, but its value doesn't make sense for the operation that we are asking for. https://docs.python.org/3/library/exceptions.html#ValueError

Variables must be declared (i.e., must have a value) before we evaluate them

```
1 answer = 2*does_not_exist
NameError Traceback (most recent call last)
<ipython-input-78-7576ff000ce0> in <module>()
----> 1 answer = 2*does_not_exist
```

NameError: name 'does\_not\_exist' is not defined

NameError signifies that Python can't find anything (variable, function, etc) matching a given name. <u>https://docs.python.org/3/library/exceptions.html#NameError</u>

# Comments in Python

Comments provide a way to document your code Good for when other people have to read your code But *also* good for you! Comments explain to a reader (whether you or someone else) what your code is *meant* to do, which is not always obvious from reading the code itself!

```
1 # This is a comment.
2 # Python doesn't try to run code that is
3 # "commented out".
4 euler = 2.71828 # Euler's number
5 '''Triple quotes let you write a multi-line comment
6 like this one. Everything between the first
7 triple-quote and the second one will be ignored
8 by Python when you run your program'''
9 print(euler)
```

We've already seen examples of functions: e.g., type() and print()

Function calls take the form function name(function arguments)

A function takes zero or more arguments and returns a value

We've already seen examples of functions: e.g., type() and print()

Function calls take the form function\_name(function arguments)

A function takes zero or more arguments and returns a value



Note: in the examples below, we write math.sqrt() to call the sqrt() function from the math module. This "dot" notation will show up a lot this semester, so get used to it!

We've already seen examples of functions: e.g., type() and print()

Function calls take the form function\_name(function arguments)

A function takes zero or more arguments and returns a value



Note: in the examples below, we write <code>math.sqrt()</code> to call the <code>sqrt()</code> function from the <code>math</code> module. This notation will show up a lot this semester, so get used to it!

We've already seen examples of functions: e.g., type() and print()

Function calls take the form function\_name(function arguments)

A function takes zero or more arguments and returns a value

```
1 import math
2 rt2 = math.sqrt(2)
3 print(rt2)
1.41421356237
1 a=2
2 b=3
3 math.pow(a,b)
8.0
```

Documentation for the Python math module: <a href="https://docs.python.org/3/library/math.html">https://docs.python.org/3/library/math.html</a>

Functions can be composed

Supply an expression as the argument of a function Output of one function becomes input to another



#### 0.36787968862663156

We can make new functions using function definition

Creates a new function, which we can then call whenever we need it



https://en.wikipedia.org/wiki/Ludwig\_Wittgenstein

We can make new functions using function definition

Creates a new function, which we can then call whenever we need it



This line (called the **header** in some documentation) says that we are defining a function called print\_wittgenstein, and that the function takes no argument.

```
print_wittgenstein()
```

Die Welt ist alles was der Fall ist

We can make new functions using function definition



We can make new functions using function definition

Creates a new function, which we can then call whenever we need it



Any arguments to the function are giving inside the parentheses. This function takes no arguments, so we just give empty parentheses. In a few slides, we'll see a function that takes arguments.

We can make new functions using function definition

Creates a new function, which we can then call whenever we need it



The colon (:) is required by Python's syntax. You'll see this symbol a lot, as it is commonly used in Python to signal the start of an indented block of code. (more on this in a few slides).

We can make new functions using function definition



We can make new functions using function definition



We can make new functions using function definition



We can make new functions using function definition



After defining a function, we can use it anywhere, including in other functions

```
1 def wittgenstein_sandwich(bread):
2     print(bread)
3     print_wittgenstein()
4     print(bread)
5 wittgenstein_sandwich('here is a string')
```

here is a string Die Welt ist Alles was der Fall ist. here is a string This function takes one argument, prints it, then prints our Wittgenstein quote, then prints the argument again.

After defining a function, we can use it anywhere, including in other functions

1	def wittgenstein sandwick (bread)
2	print(bread)
3	print wittgenstein()
4	print(bread)
5	wittgenstein_sandwich('here is a string')

This function takes one argument, which we call bread. All the arguments named here act like variables within the body of the function, but not outside the body. We'll return to this in a few slides.

```
here is a string
Die Welt ist Alles
was der Fall ist.
here is a string
```

After defining a function, we can use it anywhere, including in other functions



Body of the function specifies what to do with the argument(s). In this case, we print whatever the argument was, then print our Wittgenstein quote, and then print the argument again.

```
here is a string
Die Welt ist Alles
was der Fall ist.
here is a string
```

After defining a function, we can use it anywhere, including in other functions



Now that we've defined our function, we can call it. In this case, when we call our function, the variable bread in the definition gets the value 'here is a string', and then proceeds to run the code in the function body.

After defining a function, we can use it anywhere, including in other functions



Using the return keyword, we can define functions that produce results

```
1 def multiply_by_two(x):
2 return 2*x
3 multiply_by_two(5)
```

10

1 y = multiply\_by\_two(-1.5)
2 print(y)

-3.0

Using the return keyword, we can define functions that produce results



1 y = multiply\_by\_two(-1.5)
2 print(y)

-3.0

Using the return keyword, we can define functions that produce results

```
1 def multiply_by_two(x):
2 return 2*x
3 multiply by two(5)
```

10

1 y = multiply\_by\_two(-1.5)
2 print(y)

So when Python executes this line, it takes the integer 5, which becomes the parameter x in the function multiply\_by\_two, and this line evaluates to 10.

-3.0

Using the return keyword, we can define functions that produce results

```
1 def multiply_by_two(x):
2 return 2*x
3 multiply_by_two(5)
```

10

```
1 y = multiply_by_two(-1.5)
2 print(y)
-3.0
```

Alternatively, we can call the function and assign its result to a variable, just like we did with the functions in the math module.

Using the return keyword, we can define functions that produce results

```
1 def multiply_by_two(x):
2 return 2*x
3 multiply_by_two(5)
```

10

1 y = multiply\_by\_two(-1.5)
2 print(y)
-3.0

Notice that the argument is a float, now, instead of an int. This doesn't bother Python at all. We know how to multiply a float by an integer.

Using the return keyword, we can define functions that produce results

```
1 def multiply_by_two(x):
2 return 2*x
3 multiply_by_two(5)
```

10

```
1 y = multiply_by_two(-1.5)
2 print(y)
```

-3.0

```
1 multiply_by_two('goat')
```

'goatgoat'

2\*'goat' is 'goatgoat'?! It makes sense, but... where did that come from? We'll see what's going on here in a few lectures.

```
1 def wittgenstein sandwich(bread):
        local var = 1 # define a useless variable, just as example.
  2
  3
        print(bread)
        print_wittgenstein()
  4
  5
        print(bread)
   print(bread)
NameError
                                           Traceback (most recent call last)
<ipython-input-96-8745f5bed0d2> in <module>()
            print wittgenstein()
      4
           print(bread)
      5
----> 6 print(bread)
                                                      Variables are local. Variables defined inside a
NameError: name 'bread' is not defined
                                                      function body can't be referenced outside.
  1 print(local var)
NameError
                                           Traceback (most recent call last)
<ipython-input-97-38c61bb47a8e> in <module>()
----> 1 print(local var)
NameError: name 'local var' is not defined
```

When you define a function, you are actually creating a variable of type **function** 

Functions are objects that you can treat just like other variables

```
1 type(print_wittgenstein)
```

function

```
1 print_wittgenstein
```

<function \_\_main\_.print\_wittgenstein>

This number is the address in memory where print\_wittgenstein is stored. It may be different on your computer.

1 print(print\_wittgenstein)

<function print\_wittgenstein at 0x10aa0aaa0>