STAT679 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 is a dynamically typed, interpreted programming language

Created by Guido van Rossum 12 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

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.

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





Several options for running Python on your computer

Python interpreter

Jupyter: https://jupyter.org/

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

Suggestions from Allen Downey:

http://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 3.6.3 (default, Oct 4 2017, 06:09:05)
[GCC 4.2.1 Compatible Apple LLVM 8.0.0 (clang-800.0.42.1)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
keith@Steinhaus:~/demo$ python
Python 2.7.13 | Anaconda 4.4.0 (x86 64) | (default, Dec 20 2016, 23:05:08)
[GCC 4.2.1 Compatible Apple LLVM 6.0 (clang-600.0.57)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
Anaconda is brought to you by Continuum Analytics.
Please check out: http://continuum.io/thanks and https://anaconda.org
>>>
```

Python Interpreter on the Command Line

Python 3 vs Python 2 keith@Steinhaus:~/demo\$ python3 Python 3.6.3 (default, Oct 4 2017, 06:09:05) [GCC 4.2.1 Compatible Apple LLVM 8.0.0 (clarg-800.0.42.1)] on darwin Type "help", "copyright", "credits" or "license" for more information. keith@Steinhaus:~/demo\$ python Python 2.7.13 | Anaconda 4.4.0 (x86 64) | (default, Dec 20 2016, 23:05:08) [GCC 4.2.1 Compatible Aprile LLVM 6.0 (clang-600.0.57)] on darwin Type "help", "copyright", "credits" or "license" for more information. Anaconda is brought to you by Continuum Analytics. Please check out: http://continuum.lo/thanks and https://anaconda.org The **prompt** indicates that the I have Python 2 running inside

system is waiting for your input.

Anacaonda, by default.

Python Interpreter on the Command Line

```
keith@Steinhaus:~/demo$ python3
Python 3.6.3 (default, Oct 4 2017, 06:09:05)
[GCC 4.2.1 Compatible Apple LLVM 8.0.0 (clang-800.0.42.1)] on darwin
Type "help", "copyright", "credits" or "license" for more information
                                   Write Python commands (code) at the prompt
keith@Steinhaus:~/demo$ pythop
Python 2.7.13 | Anaconda 4 1.0 (x86 64) | (default, Dec 20 2016, 23:05:08)
[GCC 4.2.1 Compatible Apple LLVM 6.0 (clang-600.0.57)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
Anaconda is brought to you by Continuum Analytics.
Please check out: http://continuum.io/thanks and https://anaconda.org
```





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: https://jupyter.readthedocs.io/en/latest/running.html

Good tutorials:

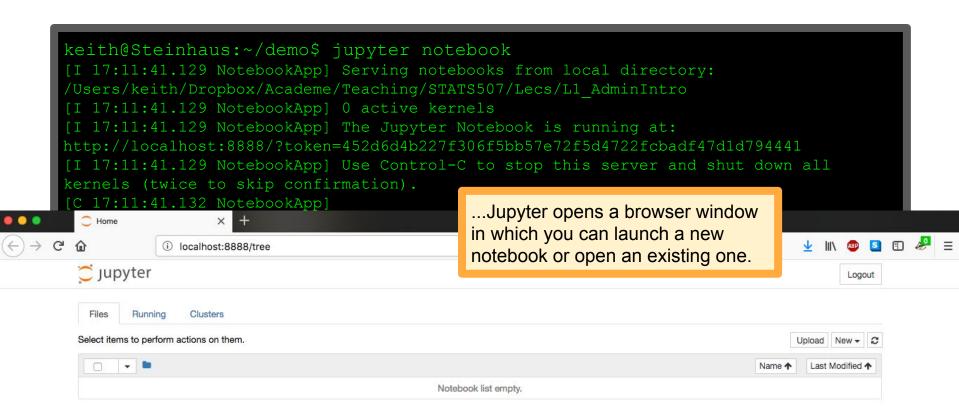
https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/execute.html

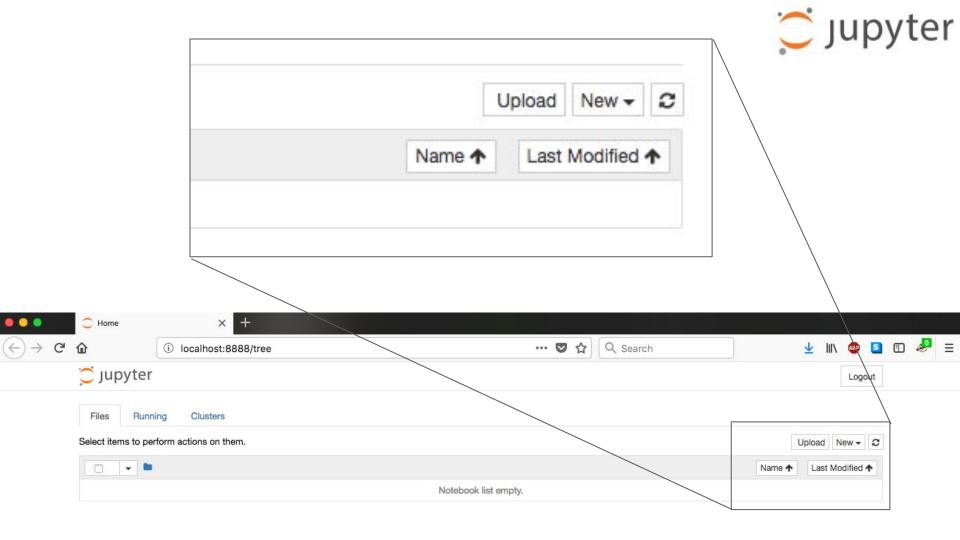
Running Jupyter

```
keith@Steinhaus:~/demo$ jupyter notebook
[I 17:11:41.129 NotebookApp] Serving notebooks from local directory:
/Users/keith/Dropbox/Academe/Teaching/STATS507/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
```

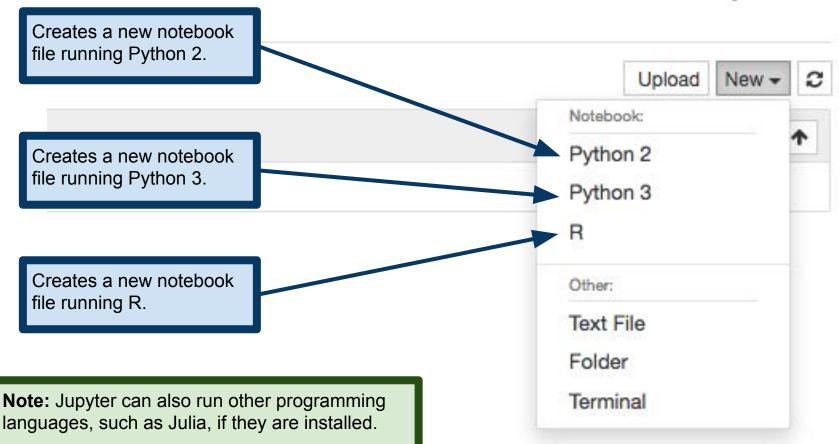
Jupyter provides some information about its startup process, and then...

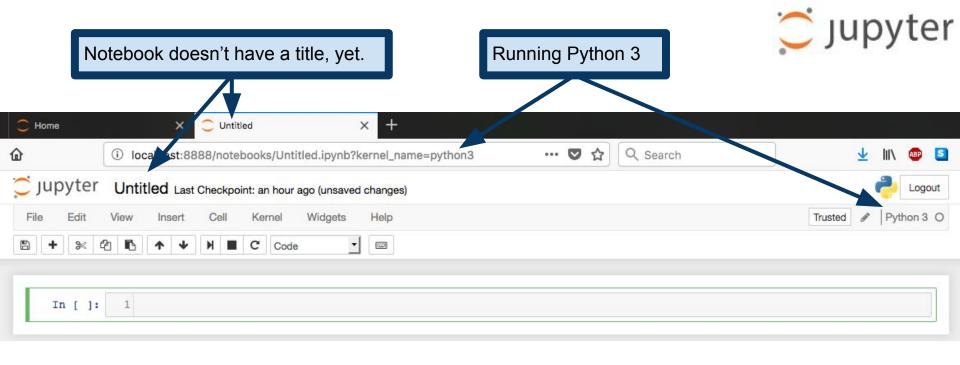
Running Jupyter

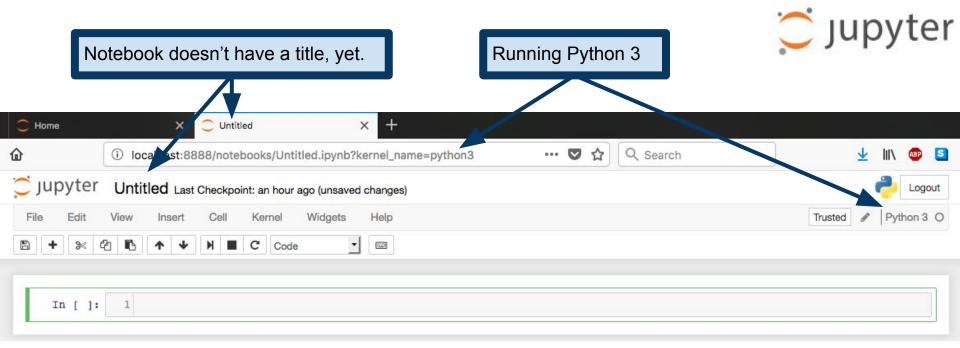




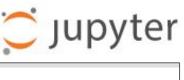


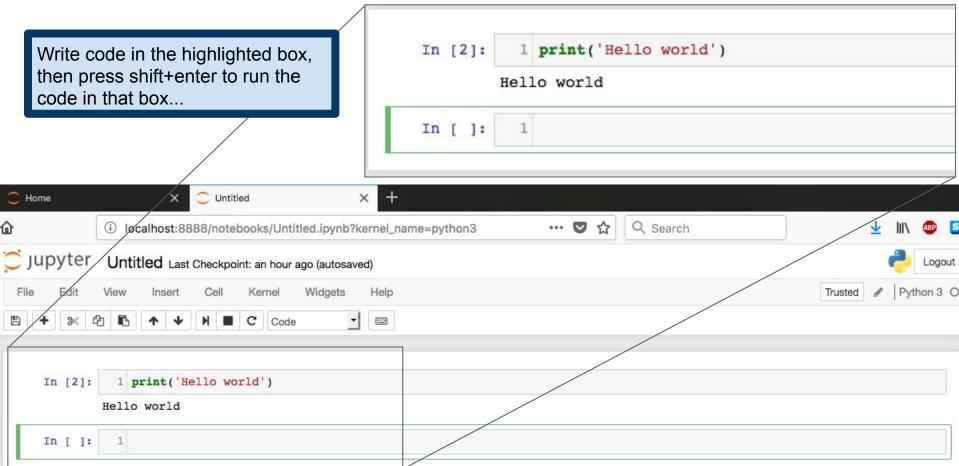




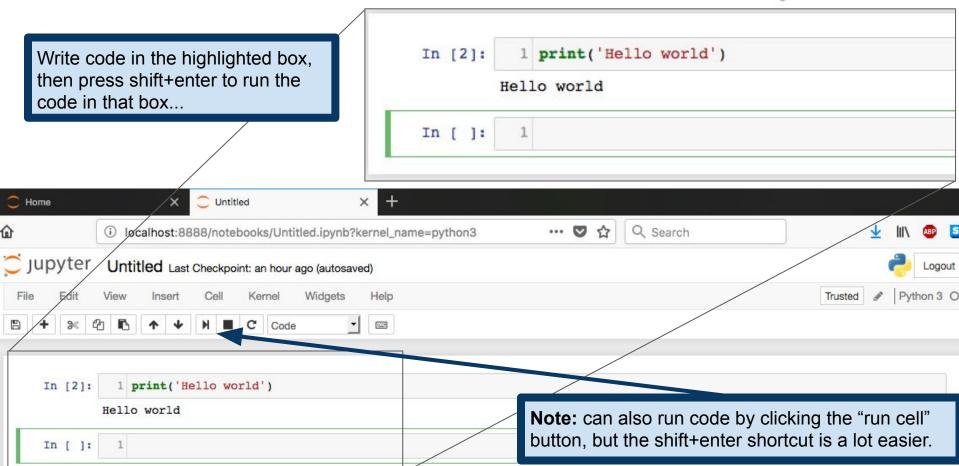


I'll leave it to you to learn about the other features by reading the documentation. For now, the green-highlighted box is most important. That's where we write Python code.

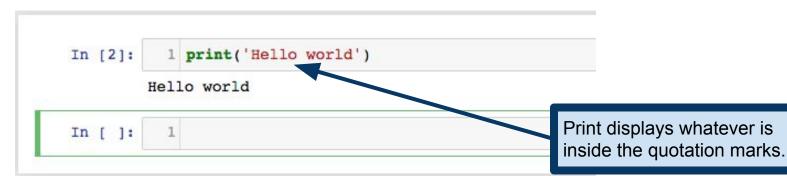








Our first function: print



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

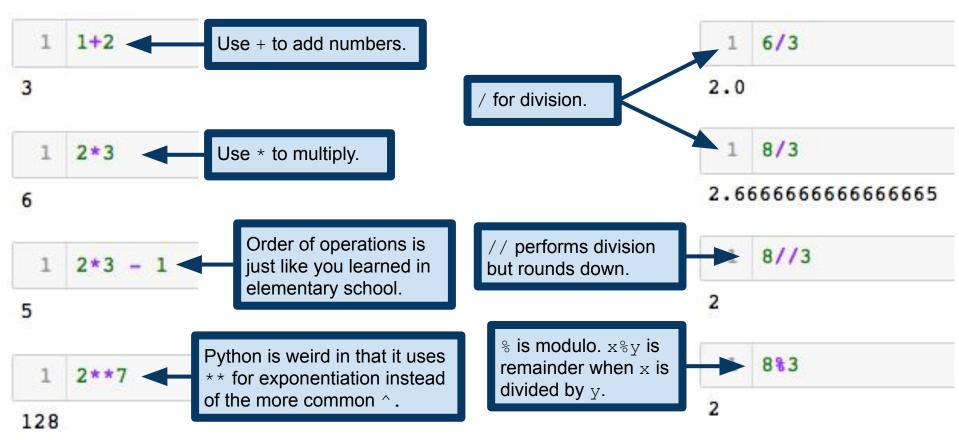
```
1 print('Hello world')
```

Hello world

```
1 print("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

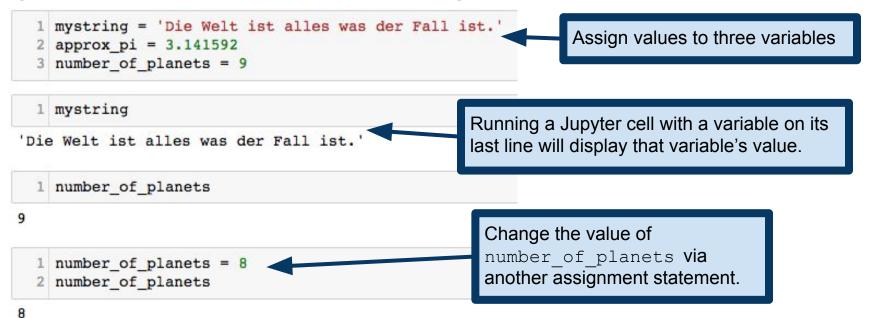
Assign a value to a variable via variable assignment

```
1 mystring = 'Die Welt ist alles was der Fall ist.'
 2 approx pi = 3.141592
                                                              Assign values to three variables
  3 number of planets = 9
  1 mystring
'Die Welt ist alles was der Fall ist.'
  1 number of planets
                                                              Change the value of
   number of planets = 8
                                                              number of planets via
  2 number of planets
                                                              another assignment statement.
```

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. Python uses what is called **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 \text{ approx pi} = 3.141592
  3 number of planets = 9
 1 mystring
                                         Python variable names can be arbitrarily long, and may
'Die Welt ist alles was der Fall ist.'
                                         contain any letters, numbers and underscore ( ), but
                                         may not start with a number. Variables can have any
  1 number of planets
                                         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
 1 number of planets = 8
 2 number of planets
                                          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 \text{ approx pi} = 3.141592
  3 number of planets = 9
  4 type(mystring)
str
                                                      Recall that type is one of the Python
  1 type(approx pi)
                                                      reserved words. Syntax highlighting
                                                      shows it as green, indicating that it is
float
                                                      a special word in Python.
  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
```

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

```
1 pi_int
```

Convert a string to an int:

```
1 int_from_str = int('8675309')
2 type(int_from_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:
     1 pi int
                    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
```

```
float_from_int = float(42)
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:
     1 pi int
                    what should be the value of
                    float from int?
```

Convert a string to an int:

float

```
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)
```

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

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

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

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

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!

```
# This is a comment.
# Python doesn't try to run code that is
# "commented out".

deuler = 2.71828 # Euler's number
'''Triple quotes let you write a multi-line comment
like this one. Everything between the first
triple-quote and the second one will be ignored
by Python when you run your program'''
print(euler)
```

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

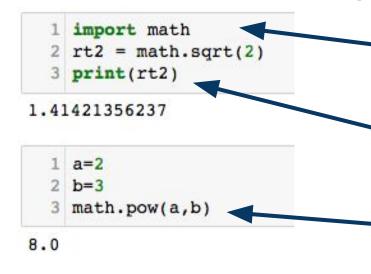
Function calls take the form function_name(function arguments)

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



Python math **module** provides a number of math functions. We have to **import** (i.e., load) the module before we can use it.

math.sqrt() takes one argument, returns its square root.

math.pow() takes two arguments. Returns the value obtained by raising the first to the power of the second.

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

```
import math
  rt2 = math.sqrt(2)
  print(rt2)

1.41421356237

1 a=2
  b=3
  math.pow(a,b)
```

8.0

Python math **module** provides a number of math functions. We have to **import** (i.e., load) the module before we can use it.

math.sqrt() takes one argument,
returns its square root.

math.pow() takes two arguments. Returns the value obtained by raising the first to the power of the second.

Note: in the examples below, we write math.sqrt() to call the sqrt() function from the math 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

```
import math
rt2 = math.sqrt(2)
print(rt2)
```

1.41421356237

```
1 a=2
2 b=3
3 math.pow(a,b)
```

Documentation for the Python math module: https://docs.python.org/3/library/math.html

Functions can be composed

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

```
1  a = 60
2  math.sin( (a/360)*2*math.pi )

0.8660254037844386

1  x = 1.71828
2  y = math.exp( -math.log(x+1))
3  y # approx'ly e^{-1}

math.sin() has as its argument an expression, which has to be evaluated before we can compute the answer.

Functions can even have the outputs of other functions as their arguments.
```

0.36787968862663156

Defining Functions

We can make new functions using function definition

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

```
1 def print_wittgenstein():
2    print("Die Welt ist alles")
3    print("was der Fall ist")
```

Let's walk through this line by line.

```
1 print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

We can make new functions using function definition

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

```
def print_wittgenstein():
    print( ble well ist alles")
    print("was der Fall ist")

1 print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

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.

We can make new functions using function definition

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

```
def print_wittgenstein():
    print("Die Welt ist alles")
    print("was der Fall ist")

The def keyword tells Python that we are defining a function.

Die Welt ist alles
was der Fall ist
```

We can make new functions using function definition

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

```
def print_wittgenstein()
    print("Die Welt ist alles")
    print("was der Fall ist")

1  print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

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

```
def print_wittgenstein(:
    print("Die Welt ist aries)
    print("was der Fall ist")

print_wittgenstein()

Die Welt ist alles
was der Fall ist

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).

Die Welt ist alles
was der Fall ist
```

We can make new functions using function definition

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

```
def print wittgenstein():

print("Die Welt ist alles")
print("was der Fall ist")

This is called the body of the function. This code is executed whenever the function is called.

print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

We can make new functions using function definition

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

```
1 def print_wittgenstein():
2    print("Die Welt ist alles")
3    print("was der Fall ist")
```

```
1 print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

Note: in languages like R, C/C++ and Java, code is organized into **blocks** using curly braces ({ and }). Python is **whitespace delimited**. So we tell Python which lines of code are part of the function definition using indentation.

We can make new functions using function definition

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

```
1 def print_wittgenstein():
2    print("Die Welt ist alles")
3    print("was der Fall ist")
```

This whitespace can be tabs, or spaces, so long as it's consistent. It is taken care of automatically by most IDEs.

```
1 print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

Note: in languages like R, C/C++ and Java, code is organized into **blocks** using curly braces ({ and }). Python is **whitespace delimited**. So we tell Python which lines of code are part of the function definition using indentation.

We can make new functions using function definition

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

```
def print_wittgenstein():
    print("Die Welt ist alles")
    print("was der Fall ist")

1    print_wittgenstein()
Die Welt ist alles
was der Fall ist

We have defined our function. Now, any
time we call it, Python executes the code in
the definition, in order.
```

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

```
def wittgenstein_sandwich(bread):
    print(bread)
    print_wittgenstein()
    print(bread)
    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.

here is a string

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

```
def wittgenstein_sandwicl (bread)
print(bread)
print_wittgenstein()
print(bread)

wittgenstein_sandwich('here is a string')

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

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.

was der Fall ist. here is a string

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

```
def wittgenstein sandwich (bread):
print(bread)
print(bread)
print(bread)
wittgenstein_sandwich(bread):
brint(bread)
wittgenstein_sandwich(bread):
brint(bread)
wittgenstein_sandwich(bread):
brint(bread)
wittgenstein_sandwich(bread):
brint(bread)
wittgenstein_sandwich(bread):
brint(bread)
wittgenstein quote, and then print the argument again.

here is a string
Die Welt ist Alles
```

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

```
def wittgenstein_sandwich(bread):
    print(bread)
    print_wittgenstein()
    print(bread)
wittgenstein_sandwich('here is a string')
```

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

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

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

Note: this last line is **not** part of the function body. We communicate this fact to Python by the indentation. Python knows that the function body is finished once it sees a line without indentation.

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

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.

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

```
def multiply_by_two(x):
    return 2*x
    multiply_by_two takes one
argument and returns that
argument, multiplied by two.

1    y = multiply_by_two(-1.5)
    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

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.

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)
```

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)
```

```
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)
```

```
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.

```
def wittgenstein_sandwich(bread):
    local_var = 1 # define a useless variable, just as example.
    print(bread)
    print_wittgenstein()
    print(bread)
    print(bread)
    print(bread)
```

Variables are **local**. Variables defined inside a function body can't be referenced outside.

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

