STAT606
Computing for Data Science and Statistics

Lecture 1: Introduction to Python
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

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

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.

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

Several options for running Python on your computer

- Python interpreter
  - Jupyter: [https://jupyter.org/](https://jupyter.org/)
  - PythonAnywhere: [https://www.pythonanywhere.com/](https://www.pythonanywhere.com/)
- Suggestions from Allen Downey:
  - [https://www.allendowney.com/wp/books/think-python-2e/](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


*Note:* Jupyter recommends Anaconda: [https://www.anaconda.com/](https://www.anaconda.com/)

I mildly recommend against Anaconda, but it’s your choice
Python Interpreter on the Command Line

keith@Steinhaus:~/demo$ python3
Python 3.11.2 (v3.11.2:878ead1acl1, 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

The prompt indicates that the system is waiting for your input.
Python Interpreter on the Command Line

Write Python commands (code) at the prompt
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

Good tutorials:
https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook
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] 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

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]
...Jupyter opens a browser window in which you can launch a new notebook or open an existing one.
Creates a new notebook file running Python 2.

Creates a new notebook file running Python 3.

Creates a new notebook file running R.

Note: Jupyter can also run other programming languages, such as Julia, if they are installed.
Notebook doesn’t have a title, yet.

Running Python 3

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 code in the highlighted box, then press shift+enter to run the code in that box...

```python
print('Hello world')
Hello world
```
Write code in the highlighted box, then press shift+enter to run the code in that box...

Note: can also run code by clicking the “run cell” button, but the shift+enter shortcut is a lot easier.
Our first function: \texttt{print}

If you haven’t already guessed, \texttt{print} takes a Python \texttt{string} and prints it. Of course, “print” here means to display a string, not literally print it on a printer!

\textbf{Note:} if you know Python 2, you’ll notice that \texttt{print} is a bit different in Python 3. That is because in Python 2, \texttt{print} was a \texttt{statement}, whereas in Python 3, \texttt{print} is a \texttt{function}. 

\begin{itemize}
  \item \texttt{In [2]: print('Hello world')} \quad \texttt{Hello world}
  \item \texttt{In [ ]: print('Hello world')} \quad \texttt{Hello world}
  \item \texttt{In [ ]: print("Hello world!")} \quad \texttt{Hello world!}
\end{itemize}

Print displays whatever is inside the quotation marks.

Can also use double quotes
Use `+` to add numbers.

Use `*` to multiply.

Order of operations is just like you learned in elementary school.

Python is weird in that it uses `**` for exponentiation instead of the more common `^`.

`/` for division.

`//` performs division but rounds down.

`%` is modulo. `x % y` is remainder when `x` is divided by `y`. 
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)
Variables in Python

**Variable** is a name that refers to a value

Assign a value to a variable via **variable assignment**

Assign values to three variables

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

Change the value of `number_of_planets` via another assignment statement.

```python
1 mystring
'Die Welt ist alles was der Fall ist.'
```

```python
1 number_of_planets
9
```

```python
1 number_of_planets = 8
2 number_of_planets
8
```
Variables in Python

**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
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
8
```

**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.

Running a Jupyter cell with a variable on its last line will display that variable’s value.

Change the value of `number_of_planets` via another assignment statement.
Variables in Python

**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
3 number_of_planets = 9
```

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

**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.
Variables in Python

Sometimes we do need to know the type of a variable

Python `type()` function does this for us

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

type(mystring)
```

```
str
```

```python
type(approx_pi)
```

```
float
```

```python
type(number_of_planets)
```

```
int
```

Recall that `type` is one of the Python reserved words. Syntax highlighting shows it as green, indicating that it is a special word in Python.
Variables in Python

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

Convert a float to an int:

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

float

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

int
```

Convert a string to an int:

```python
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.
Variables in Python

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`

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.

Test your understanding: what should be the value of `float_from_int`?
Variables in Python

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

**Convert a float to an int:**

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

```
float
```

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

```
int
```

**Convert a string to an int:**

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

```
int
```

```python
1 int_from_str
```

```
8675309
```

**Test your understanding:**

what should be the value of `float_from_int`?

```python
1 float_from_int = float(42)
2 type(float_from_int)
```

```
float
```

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

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

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

```python
1 goat_int = int('goat')
```

```bash
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](https://docs.python.org/3/library/exceptions.html#ValueError)
Variables in Python

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

```
answer = 2*does_not_exist
```

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

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

2.71828

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!
Functions in Python

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
Functions in Python

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.
Functions in Python

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

```plaintext
1.41421356237
```

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

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

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

```plaintext
8.0
```

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

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
1 import math
2 rt2 = math.sqrt(2)
3 print(rt2)
1.41421356237
```

```python
1 a=2
2 b=3
3 math.pow(a,b)
8.0
```

**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!

Documentation for the Python `math` module:
[https://docs.python.org/3/library/math.html](https://docs.python.org/3/library/math.html)
Functions in Python

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}
0.36787968862663156
```

*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.
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

Defining Functions

We can make new functions using **function definition**

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

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

print_wittgenstein()
```

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.
Defining Functions

We can make new functions using function definition

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

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

print_wittgenstein()
```

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

```
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.
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")
```

```
1  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).
Defining Functions

We can make new functions using **function definition**

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

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

print_wittgenstein()

Die Welt ist alles
was der Fall ist
```

This is called the **body** of the function. This code is executed whenever the function is called.
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")
```

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.
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")
```

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

**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.
Defining Functions

We can make new functions using **function definition**

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

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

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

```
print_wittgenstein()
```

Die Welt ist alles
was der Fall ist
Defining Functions

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

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

This function takes one argument, prints it, then prints our Wittgenstein quote, then prints the argument again.
Defining Functions

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

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.
Defining Functions

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

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

wittgenstein_sandwich('here is a string')
```

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.
Defining Functions

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

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

wittgenstein_sandwich('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.
Defining Functions

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

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

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.

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.
Defining Functions

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

```python
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
```
Defining Functions

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

```python
def multiply_by_two(x):
    return 2*x
```

`multiply_by_two` takes one argument and returns that argument, multiplied by two.

```python
multiply_by_two(5)
```

```
10
```

```python
y = multiply_by_two(-1.5)
print(y)
```

```
-3.0
```
Defining Functions

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

```python
def multiply_by_two(x):
    return 2*x
multiply_by_two(5)
```

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.

```python
y = multiply_by_two(-1.5)
print(y)
```

-3.0
Defining Functions

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

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

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

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

-3.0
Defining 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

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.
Defining 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

```
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.
Defining Functions

```python
1 def wittgenstein_sandwich(bread):
2     local_var = 1  # define a useless variable, just as example.
3     print(bread)
4     print_wittgenstein()
5     print(bread)
6     print(bread)
```

```
NameError: Traceback (most recent call last)
<ipython-input-96-8745f5bed0d2> in <module>()
   4     print_wittgenstein()
   5     print(bread)
----> 6 print(bread)
NameError: name 'bread' is not defined
```

```
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
```

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

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

```python
1 type(print_wittgenstein)
function

1 print_wittgenstein
<function __main__.print_wittgenstein>

1 print(print_wittgenstein)
<function print_wittgenstein at 0x10a0a0aa0>
```

This number is the address in memory where `print_wittgenstein` is stored. It may be different on your computer.