Lecture 02 - Intro to Python - Notebook (version after third lecture)

January 25, 2023

# 1 Lecture 2: Quick(ish) Introduction to Python

The goal of this lecture is to give you some of the basics. It's not possible for us to cover **everything** you'll need to know ahead of time. As graduate students, you are expected to be able to do some research and self-teaching on your own to build your coding skills, and of course you can *always* come ask me for help!

A nice reference for a lot of the computational skills we'll be covering (coding, unix command line, git) is the "Software Carpentry" set of lessons: https://software-carpentry.org/lessons/. You should seriously considering checking their tutorials for extra practice and more in-depth lessons!

This document is a "Jupyter Notebook". It's kind of like interactive mode, but also lets you intersperse text, html, etc, among it. (It's free and easy to set up.)

Python is a **whitespace-based language**. In C, C++, Java, and many other languages, you use braces to group code:

```
if (x == 1) {
    do_something();
}
```

and the spacing is just for readability. For example, the following code does the same thing:

```
if (x == 1) { do_something();
}
```

In Python you use indenting, and colons (:) to have the same effect. You also do not use semicolons (;) to end commands.

```
if x == 1:
    do something()
```

You have to be really careful to be consistent by either - always using tabs, or - always using spaces (and the same number)

```
[1]: x = 1

[2]: x

[2]: 1
```

```
[3]: if x == 1:
          print("hello")
     hello
 [4]: if x == 1:
                  print("hello")
     hello
 [7]: | if x == 1:
          print("hello")
           print("world")
         Cell In[7], line 3
           print("world")
       IndentationError: unexpected indent
 [8]: if x == 1:
          print("hello")
          print("world") # if you use a tab, Jupyter will fix it for you⊔
       →automatically! Your code editor might not.
     hello
     world
     Python uses if, for, and while statements like many other languages.
 [9]: x = 1
      while x < 10:
          x = x + 2
          print(x)
     3
     5
     7
     9
     11
[10]: for y in range(3, 6): # 3, 4, 5
          print(y)
     3
     4
     5
```

In a for loop, you can iterate over many different types of objects (lists, sets, tuples, dictionaries, strings, etc.)

range(a,b) is a way of looping over all of the integers between a (inclusive) and b (exclusive).

```
[12]: for z in "hello":
    print(z)

h
e
1
1
0

[13]: for z in [19, -100, "banana"]:
    print(z+1)
```

You may have noticed that Python is not a **statically-typed** language, which means you do not need to tell it whether a variable you are defining is an integer or a string or a list, etc. You just define it, and it figures it out.

But, there are still different types! You can always use the type function to check what type an object has.

```
[14]: L = [1, 2, 3] type(L)
```

[14]: list

```
[15]: L = (1,2,3)
     type(L)
[15]: tuple
[16]: L = \{1,2,3\}
     type(L)
[16]: set
[17]: sum([1,2,3])
[17]: 6
[18]: type(sum)
[18]: builtin_function_or_method
    Now we're going to discuss a bunch of the fundamental types in Python.
    1.1 Integers, Floating Point Numbers, and Complex Numbers
[19]: x = 7
     type(x)
[19]: int
[22]: import math
     math.pi
[22]: 3.141592653589793
[24]: x = 7.0
     type(x)
     X
[24]: 7.0
print(x)
    3.0
[26]: True
[27]: y = 0.1
     print(y)
```

```
0.1
```

[39]: bool

```
[28]: 0.1 + 0.1 + 0.1 - 0.3
[28]: 5.551115123125783e-17
[ ]: 5.55111 * 10^{-17}
[29]: 0.1 + 0.1 + 0.1 == 0.3
[29]: False
[30]: x = 0.1 + 0.1 + 0.1
      y = 0.3
      print(x==y)
      print(abs(x-y) < 0.00000000001)
     False
     True
[31]: 15 / 7
[31]: 2.142857142857143
[32]: 15 // 7
[32]: 2
[33]: 14 / 7
[33]: 2.0
[34]: z = complex(3, 5)
      t = complex(1,-1)
      print(z)
      type(z)
      z*t
     (3+5j)
[34]: (8+2j)
     1.2 Boolean
     A boolean is just a True or False value. That's it!
[39]: b = True
      type(b)
```

```
[40]: if b:
          print("hello")
     hello
[41]: if not b:
          print("hello")
[42]: t = (10 + 10 == 20) # Use "==" to test equality, and "=" to actually set_\( \sigma \)
       ⇔something equal
      print(t)
      type(t)
     True
[42]: bool
     1.3 None
     There is a weird object in Python called None. It's just a useful thing to have around, often as a
     default value until you assign something.
[44]: x = None
      type(x)
[44]: NoneType
[45]: y = None
      print(y)
      if y is None:
          print("y has the value None")
      y = 3
      if y is None:
          print("y has the value None")
     None
     y has the value None
[50]: x = 1
      y = x
      print(x is y)
     True
[59]: x = [1, 2, 3]
      y = [1, 2, 3]
      print(id(x))
      print(id(y))
      print(x == y)
```

print(x is y)

```
y = x
      x = [4, 5, 6]
      print(id(x))
      print(id(y))
     4437131456
     4437132032
     True
     False
     4437126656
     4437131456
[60]: x = [1, 2, 3]
      y = x
      x = [4, 5, 6]
      print(x)
      print(y)
      [4, 5, 6]
      [1, 2, 3]
     1.4 Strings
     A string is just a sequence of characters.
[63]: type('banana')
[63]: str
     You can do a million things with strings.
[64]: s = "banana"
      s.split("n")
[64]: ['ba', 'a', 'a']
     Use len to find the length of a string and + to concatenate two strings together.
[65]: one = "hello"
      two = "world"
      three = one + " " + two
      print(len(three))
      print(three)
     11
     hello world
[66]: three - two
```

```
Traceback (most recent call last)
       TypeError
       Cell In[66], line 1
       ---> 1 three - two
       TypeError: unsupported operand type(s) for -: 'str' and 'str'
[67]: three.len()
       AttributeError
                                                   Traceback (most recent call last)
       Cell In[67], line 1
       ---> 1 three.len()
       AttributeError: 'str' object has no attribute 'len'
     1.5 Lists
     A list is an ordered sequence of things.
[68]: L = [15, "banana", 7, False, [1, 2, 3]]
[69]: print(L)
     [15, 'banana', 7, False, [1, 2, 3]]
     Elements of lists are accessed with bracket notation, starting from 0.
[70]: L[0]
[70]: 15
[71]: L[1]
[71]: 'banana'
[72]: L[2]
[72]: 7
[73]: L[3]
[73]: False
[74]: L[4]
[74]: [1, 2, 3]
```

```
[77]: (L[4])[1]
[77]: 2
[78]: L[5]
       IndexError
                                                   Traceback (most recent call last)
       Cell In[78], line 1
       ----> 1 L[5]
       IndexError: list index out of range
     Use len(L) to get the length of a list.
[79]: len(L)
[79]: 5
[81]: print(L)
      len(L[4])
     [15, 'banana', 7, False, [1, 2, 3]]
[81]: 3
     You can set elements of the list manually as well.
[82]: L
[82]: [15, 'banana', 7, False, [1, 2, 3]]
[83]: L[1] = "apple"
[84]: L
[84]: [15, 'apple', 7, False, [1, 2, 3]]
[85]: L[8] = "can't set this"
       IndexError
                                                   Traceback (most recent call last)
       Cell In[85], line 1
       ----> 1 L[8] = "can't set this"
       IndexError: list assignment index out of range
```

You can sort lists:

```
[86]: R = [15, -20, 0]
R.sort()
print(R)
```

[-20, 0, 15]

Notice the . in the notation above. We'll talk about this more when we cover object-oriented programming, but what we're basically doing here is telling the list R to perform its <code>sort()</code> operation on itself.

You may wonder why we did len(R) instead of R.len()... it's kind of just a quirk. You get used to it.

A few more quick list functions:

```
[87]: R
```

[87]: [-20, 0, 15]

```
[88]: R.append(17) print(R)
```

[-20, 0, 15, 17]

[-20, 0, 15, 17, 7, 8, 9]

Lastly (for now) you can concatenate two lists together with the + sign.

```
[90]: [1,2,3] + [4,5,6]
```

[90]: [1, 2, 3, 4, 5, 6]

[-20, 0, 15, 17, 7, 8, 9] [100, -20, 0, 15, 17, 7, 8, 9]

[92]: 100 + R

```
TypeError Traceback (most recent call last)
Cell In[92], line 1
----> 1 100 + R

TypeError: unsupported operand type(s) for +: 'int' and 'list'
```

```
[93]: L
[93]: [15, 'apple', 7, False, [1, 2, 3]]
[101]: L[-6]
        IndexError
                                                     Traceback (most recent call last)
        Cell In[101], line 1
        ----> 1 L[-6]
        IndexError: list index out of range
[95]: L[-2]
[95]: False
[97]: L[-3]
[97]: 7
[98]: L[len(L)-1]
[98]: [1, 2, 3]
      1.6 Sets
      A list was an ordered sequence of things. A set is an unordered sequence of things with no
      repeats (just like in math).
  [4]: S = \{1, 2, 3, 4\}
       print(S)
      {1, 2, 3, 4}
  [5]: type(S)
  [5]: set
  [6]: T = \{3, 1, 4, 2\}
       print(T)
      {1, 2, 3, 4}
[109]: S = \{3, 7, 7, 7, 10\}
       print(S)
       T = \{3, 7, 10\}
       print(T)
```

```
print(S == {3, 7, 10})
      {10, 3, 7}
      {10, 3, 7}
      True
[105]: S == T
[105]: True
      You can't access elements using the bracket notation because there is no first element, second
      element, etc. You should never assume that you know the order Python will internally store your
      list in!
  [7]: S[2]
                                                      Traceback (most recent call last)
        TypeError
        Cell In[7], line 1
        ----> 1 S[2]
        TypeError: 'set' object is not subscriptable
  [8]: for element in S:
           print(element)
      1
      2
      3
      4
[112]: # we stopped here at the end of Friday, Jan 20
  []:
      Here are some functions you can do with sets:
  [9]: first = \{1,5,6\}
       second = \{2,4,5\}
 [10]: first.union(second)
 [10]: {1, 2, 4, 5, 6}
 [11]: print(first)
       print(second)
      {1, 5, 6}
      \{2, 4, 5\}
```

```
[12]: second.union(first)
[12]: {1, 2, 4, 5, 6}
[13]: first.intersection(second)
[13]: {5}
[14]: print(first)
      print(second)
     {1, 5, 6}
     \{2, 4, 5\}
[15]: first.difference(second) # all of the things IN first, and NOT IN second
[15]: {1, 6}
[16]: # By the way, you write comments in python by just starting the line with the
       ⇔pound key.
[17]: first
[17]: {1, 5, 6}
[18]: first.add(9)
      print(first)
     {1, 5, 6, 9}
[19]: first.add(5)
      print(first) # No duplicates!
     {1, 5, 6, 9}
[20]: first.remove(5)
[21]: print(first)
     {1, 6, 9}
[22]: first.remove(5)
      KeyError
                                                 Traceback (most recent call last)
      Cell In[22], line 1
      ---> 1 first.remove(5)
      KeyError: 5
```

```
[24]: first.discard(5) print(first)
```

{1, 6, 9}

# 1.7 Tuples

It starts to get a little tricky here! A tuple is an **ordered sequence** of things.

Wait... isn't that what a list was?

```
[25]: T = (1,2,3,4)
print(T)
type(T)
```

(1, 2, 3, 4)

[25]: tuple

```
[26]: L = [1,2,3,4]

L == T # They are different types of objects, so they can't be equal.
```

[26]: False

The key is that a tuple is what we call **immutable**. Once it's defined, it *cannot* be changed, ever, at all.

```
[27]: print(T) print(T[2])
```

(1, 2, 3, 4)

```
[28]: T[2] = 17
```

```
TypeError Traceback (most recent call last)
Cell In[28], line 1
----> 1 T[2] = 17

TypeError: 'tuple' object does not support item assignment
```

It is still possible to do things like concatenate two tuples to make a new bigger tuple, but it's a **new** bigger tuple, and the original one is still unchanged.

```
[29]: T
```

[29]: (1, 2, 3, 4)

```
[30]: T + (5,6)
```

```
[30]: (1, 2, 3, 4, 5, 6)
[31]: T
[31]: (1, 2, 3, 4)
[32]: T.append(5)
       AttributeError
                                                     Traceback (most recent call last)
       Cell In[32], line 1
       ---> 1 T.append(5)
       AttributeError: 'tuple' object has no attribute 'append'
[33]: T
[33]: (1, 2, 3, 4)
[34]: T = T + (5,6)
[35]: T
[35]: (1, 2, 3, 4, 5, 6)
     So, we define a new tuple with parentheses, but there's one catch: if your tuple has a single item,
     it needs a special bit of syntax.
[36]: x = (1)
      print(x)
      type(x)
     1
[36]: int
[37]: x = (1,)
      print(x)
      type(x)
     (1,)
[37]: tuple
     So, lists are mutable, tuples are immutable. Why do we need two different versions?
[39]: L = [1,2,3,4,5,6]
      print(5 in L)
      print(7 in L)
```

## True False

Under-the-hood, when you store things in a set, Python is being super smart about how it stores it. When you add an element to a set you really do not want python to have to scan one-by-one through all the things in the set to make sure it's not already there. So, it uses a clever technique called *hashing*.

You don't need to know the details right now, but the broad idea is that Python takes each thing in the set and assigns a number to it called its hash, and then uses the hashes to make sure there are no duplicates.

```
[40]: hash(17)
[40]: 17
[41]: hash("banana")
[41]: 1701994160590617831
[42]: hash((1,2,3,4))
[42]: 590899387183067792
[43]: hash([1,2,3])
                                                  Traceback (most recent call last)
       TypeError
       Cell In[43], line 1
       ---> 1 hash([1,2,3])
       TypeError: unhashable type: 'list'
 []:
[44]: {[1, 2, 3, 4], [1, 2], [7,8]}
       TypeError
                                                  Traceback (most recent call last)
       Cell In[44], line 1
       ----> 1 {[1, 2, 3, 4], [1, 2], [7,8]}
       TypeError: unhashable type: 'list'
[45]: \{(1, 2, 3, 4), (1, 2), (7, 8)\}
[45]: {(1, 2), (1, 2, 3, 4), (7, 8)}
```

The problem is that you **can't hash mutable things**. Once you get an object's hash, that needs to stay its hash forever. You could hash a list, then appending an element to the list would mean a new hash would have to be generated, and this would mess everything up.

Bottom line: Sometimes you need an immutable version of something, like to put it in a set.

```
[46]: {5, 17, [1,2,3]}
```

```
TypeError Traceback (most recent call last)
Cell In[46], line 1
----> 1 {5, 17, [1,2,3]}
TypeError: unhashable type: 'list'
```

```
[47]: {5, 17, (1,2,3)}
```

```
[47]: {(1, 2, 3), 17, 5}
```

```
[48]: {5, 17, {1,2,3}}
```

```
TypeError Traceback (most recent call last)
Cell In[48], line 1
----> 1 {5, 17, {1,2,3}}

TypeError: unhashable type: 'set'
```

Sets are mutable too! Sets must contain immutable things, but they themselves are mutable.

Of course we knew this, because we can do S.add(). So what if you want sets in your sets? There is an immutable version of a set called a frozenset.

```
[49]: { 5, 17, frozenset({1, 2, 3}) }
```

```
[49]: {17, 5, frozenset({1, 2, 3})}
```

When should you use a tuple versus a list? - If it's going to go in a set (or, as we'll see in a second, in a dictionary), it has to be immutable. Thus, use a tuple. - If you need to be able to add and remove things, use a list. - If the size will always stay the same, you probably want a tuple. For example, if you're representing xy-coordiates, use tuples.

## 1.8 Dictionaries

You can think of a list as kind of like a mathematical function whose inputs are the indices 0, 1, ... and whose outputs are the elements of the list.

```
[50]: L = ["apple", "banana", "pear"]
```

```
[]: | # 0 -> apple, 1 -> banana, 2 -> pear
     In a dictionary, the inputs don't have to be integers, they can be any (immutable) object.
[51]: # To define 17 -> apple, banana -> pear, (1, 2, 3) -> True
      d = {17:"apple", "banana":"pear", (1,2,3):True}
      print(d)
     {17: 'apple', 'banana': 'pear', (1, 2, 3): True}
     The inputs are called keys and the outputs are called values.
[52]: d[17]
[52]: 'apple'
[53]: d["banana"]
[53]: 'pear'
[54]: d[(1,2,3)]
[54]: True
[56]: d["pear"] = "hello"
      d["pear"]
      print(d)
     {17: 'apple', 'banana': 'pear', (1, 2, 3): True, 'pear': 'hello'}
     You can assign new values too
[57]: d[1] = "one"
      print(d)
     {17: 'apple', 'banana': 'pear', (1, 2, 3): True, 'pear': 'hello', 1: 'one'}
[58]: d[(2, 3, 5, 7)] = False
[59]: d
[59]: {17: 'apple',
       'banana': 'pear',
       (1, 2, 3): True,
       'pear': 'hello',
       1: 'one',
       (2, 3, 5, 7): False}
```

Dictionaries are *super* useful, but take some getting used to. The keys are hashed in the background, which makes looking up the value for a given key very fast.

```
[60]: for k in d.keys():
          print(k)
     17
     banana
     (1, 2, 3)
     pear
     1
     (2, 3, 5, 7)
[61]: for v in d.values():
          print(v)
     apple
     pear
     True
     hello
     one
     False
[62]: for pair in d.items():
          print(pair)
      # (key, value)
     (17, 'apple')
     ('banana', 'pear')
     ((1, 2, 3), True)
     ('pear', 'hello')
     (1, 'one')
     ((2, 3, 5, 7), False)
     1.9 Casting
     You can tell Python to turn an object of one type into an object of another type. This is called
     casting.
[63]: L = [3, 7, 7, 12]
      print(L)
     [3, 7, 7, 12]
[64]: T = tuple(L)
      print(T)
      print(L)
     (3, 7, 7, 12)
     [3, 7, 7, 12]
[65]: S = set(L)
      print(S)
```

```
{3, 12, 7}
[66]: print(L)
      list(set(L))
     [3, 7, 7, 12]
[66]: [3, 12, 7]
[67]: dict(L)
                                                  Traceback (most recent call last)
       TypeError
       Cell In[67], line 1
       ----> 1 dict(L)
       TypeError: cannot convert dictionary update sequence element #0 to a sequence
[68]: str(L)
[68]: '[3, 7, 7, 12]'
[69]: int(L)
                                                  Traceback (most recent call last)
       TypeError
       Cell In[69], line 1
       ----> 1 int(L)
       TypeError: int() argument must be a string, a bytes-like object or a real

⊔
        →number, not 'list'
[71]: d = {1:"one", 2:"two", 3:"three"}
[71]: {1: 'one', 2: 'two', 3: 'three'}
[72]: list(d)
[72]: [1, 2, 3]
[75]: for key in d:
          print(key, d[key])
     1 one
     2 two
     3 three
```

Time for some practice!

https://projecteuler.net/

Problem 1: mod, looping, and comprehensions

Problem 2: negative indexing

Problem 5: all / any, and thinking mathematically

If we list all the natural numbers below 10 that are multiples of 3 or 5, we get 3, 5, 6 and 9. The sum of these multiples is 23.

Find the sum of all the multiples of 3 or 5 below 1000.

```
[]: # mod - modulus
# a % b -- the remainder you get when you divide a by b
# list comprehensions
# +=
```

```
[77]: 19 % 4 # 3, because 4 goes into 19 four whole times, with a remainder of 3
```

[77]: 3

```
[78]: 16 % 4
```

[78]: 0

```
[83]: s = 0
# loop over the numbers from 1 to 1000
for number in range(1, 1000):
    if number % 3 == 0 or number % 5 == 0:
        s += number
print(s)
```

233168

```
[82]:  # "+="
# a shortcut to write s = s + number is "s += number"
```

```
[1]: # common mistake
s = 0
# loop over the numbers from 1 to 1000
for number in range(1, 1000):
    if number % 3 == 0 or number % 5 == 0:
        s += number
print(s)
# this is bad because it overwrites the built-in "sum" method
# len, list, set, tuple
```

233168

```
[2]: sum([1,2,3,4])
[2]: 10
[]: # list comprehension
     # a way of defining a list, maybe with a condition, in one line
[5]: multiples = [number**2 for number in range(1,100) if (number % 3 == 0 or number_

√% 5 == 0)]
     multiples
     # set-builder notation in math
     # \{n^2 : 1 \le n \le 100 \text{ and } n \text{ is a multiple of } 3 \text{ or } 5\}
[5]: [9,
      25,
      36,
      81,
      100,
      144,
      225,
      324,
      400,
      441,
      576,
      625,
      729,
      900,
      1089,
      1225,
      1296,
      1521,
      1600,
      1764,
      2025,
      2304,
      2500,
      2601,
      2916,
      3025,
      3249,
      3600,
      3969,
      4225,
      4356,
      4761,
      4900,
      5184,
```

```
5625,
6084,
6400,
6561,
7056,
7225,
7569,
8100,
8649,
9025,
9216,
9801]
```

```
[6]: sum([number for number in range(1,1000) if (number % 3 == 0 or number % 5 ==<math>_{\sqcup} _{\hookrightarrow}0)])
```

#### [6]: 233168

Each new term in the Fibonacci sequence is generated by adding the previous two terms. By starting with 1 and 2, the first 10 terms will be:

```
1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...
```

By considering the terms in the Fibonacci sequence whose values do not exceed four million, find the sum of the even-valued terms.

```
[11]: # negative indexing
# to get the last element in a list, you can use L[-1]
# to get the second to last, L[-2]
s = 2
L = [1, 2]
while L[-1] < 4_000_000: # "_" do nothing, but help you see grouping
    new_number = L[-1] + L[-2]
    L.append(new_number)
    if new_number % 2 == 0 and new_number < 4_000_000:
        s = s + new_number
print(s)</pre>
```

## 4613732

```
[13]: # another version
    # negative indexing
    # to get the last element in a list, you can use L[-1]
    # to get the second to last, L[-2]
    s = 0
    L = [1, 2]
    while L[-1] < 4_000_000: # "_" do nothing, but help you see grouping
        if L[-1] % 2 == 0:
            s = s + L[-1]</pre>
```

```
new_number = L[-1] + L[-2]
L.append(new_number)

print(s)
print(L)
```

#### 4613732

[1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040, 1346269, 2178309, 3524578, 5702887]

[3]:

#### [3]: 10

2520 is the smallest number that can be divided by each of the numbers from 1 to 10 without any remainder.

What is the smallest positive number that is evenly divisible by all of the numbers from 1 to 20?

```
[3]: | # all / any
```

```
[30]: # brute force but slow
num = 1
while True:
    if all([num % i == 0 for i in range(1, 21)]):
        print(num)
        break
num += 1
```

```
KeyboardInterrupt
                                            Traceback (most recent call last)
Cell In[30], line 4
      2 \text{ num} = 1
      3 while True:
            if all([num % i == 0 for i in range(1, 21)]):
---> 4
                 print(num)
      5
                 break
Cell In[30], line 4, in stcomp>(.0)
      2 \text{ num} = 1
      3 while True:
            if all([num \% i == 0 for i in range(1, 21)]):
                 print(num)
      6
                 break
KeyboardInterrupt:
```

```
[]: | # using mathematical thinking
[41]: num = 20*19
      primes = [10, 11, 12, 13, 14, 15, 16, 17, 18]
      while True:
          if all([num % i == 0 for i in primes]):
              print(num)
              break
          num += 20*19
     232792560
 []:
 []:
 []:
[32]: # least common multiple
      import math
[37]: math.lcm(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
[37]: 232792560
[]:
[]:
[14]: num = 1500
[16]: [(\text{num } \% \text{ i } == 0) \text{ for i in } \text{range}(1,11)]
[16]: [True, True, True, True, True, False, False, False, True]
[17]: # "all" and "any"
      # give them a list of Trues/Falses
      # all(L) gives you True if and only if every entry in L is "True"
      # any(L) gives you True if and only if "any" entry is L is "True"
[18]: all([(num % i == 0) for i in range(1,11)])
[18]: False
[19]: any([(num % i == 0) for i in range(1,11)])
[19]: True
 []:
```