

Introduction to Programming for Scientists

Prof. Steven Ludtke

N420, sludtke@bcm.edu

Team Learning 2

Simple Math and Text Handling

Groups:

1:

Abreo Timothy	Timothy.Abreo@bcm.edu
AvalosMatthew	matthew.avalos@bcm.edu
OzcanZuhal	zuhal.ozcan@uth.tmc.edu

2:

Chander	Varuna	varuna.chander@bcm.edu
Benton	Susan	susan.benton@bcm.edu
Morales	Fatima	fatimam@bcm.edu

3:

Dupuis	Colleen	cdupuis@bcm.edu
Hu	Guo	guoh@bcm.edu
Coffin Stephanie	coffin@bcm.edu	
Smith Cameron	cameron.smith@bcm.edu	

4:

Kotu	Varshika	Varshika.kotu@gmail.com
LaPlante	Emily	elaplant@bcm.edu
Wang	Junkai	junkaiw@bcm.edu

5:

Liu	Yuwei	yuwei.liu@bcm.edu
Mangleburg Grant	manglebu@bcm.edu	
Shi	Huanan	huanan.shi@bcm.edu

6:

Mays	Jacquanae	jacquanae.mays@bcm.edu
Abeynaik	Shawn	abeynaik@bcm.edu
Nair	Vinay	Vinay.Nair@uth.tmc.edu
Park	Woojun	wpark@bcm.edu

Preliminaries:

Text in *italics* should be typed.

Type each block in a new “In []:” section. At the end of the group learning exercise, you will be asked to save your Jupyter notebook session and turn it in for a grade, so don’t overwrite your successful input/output blocks.

Nobody should leave class until your entire group has completed the exercise (or the class period ends).

Note, each member of the group must turn in their own independent notebook by midnight on the day of the group learning exercise. You are free to discuss and help one another (that is the point), but each person should do the actual work themselves, and submit their own .ipynb file by email to:

James.Bell@bcm.edu, muyuan.Chen@bcm.edu, sludtke@bcm.edu

- 1) Launch *jupyter notebook*
- 2) Create a new Notebook. You should get an “In []:” prompt
- 3) *1+1* (press shift-enter)
- 4) make sure you see 2 and get a new input line. If you don’t, get some help.

Simple loops (20 min):

- 1) Type this block:

```
for i in range(1,10,2):  
    print(i,i*i)
```

- 2) As with all blocks of code, press shift-enter when you’re done. This won’t be stated explicitly in the instructions any more.

- 3) Make sure you understand exactly what you got, and why you got it before proceeding. Discuss among your group if it isn’t clear.

- 4) Now try this:

```
last=input("Max: ")  
for i in range(1,last,2):  
    print(i,i*i)
```

- 5) When you run this, it will prompt you for a value. Type 10 and press enter (no shift on the enter). You will get an error. Figure out why (read the message), and make the necessary correction.

- 6) Now try this block:

```
x=1  
for i in range(2,30):  
    x=x*i  
    print(i,x)
```

7) again, make sure you understand the output before proceeding. Compare the final output line to 30! computed on a calculator.

8) Finally, try this block:

```
x=1
y=1.0
for i in range(2,30):
    x=x*i
    y=y*i
    print(i,x,y,x/y)
```

9) select the next 'In []' block and use the dropdown menu saying "Code" at the top of the page to change its type to "Markdown". You will now be able to enter text (rather than Python code) in that particular block. You can use this to provide explanations and answers to questions.

10) Use this block to briefly explain what you see in this final example and why, and press shift-enter to finalize it and open a new code block.

If statements (10 min):

1) Run this block:

```
from math import *
```

```
tot=0
for x in range(20):
    print(x,sin(x))
    tot+=sin(x)
print("total=",tot)
```

2) Now run this block:

```
from math import *
```

```
tot=0
for x in range(20):
    if (sin(x)>0) :
        print(x,sin(x))
        tot+=sin(x)
print("total=",tot)
```

3) Make sure you understand the result before moving on

String Manipulations (30 min)

1) Execute the following block. The "a=" line should be entered as a single line of text:

```
import urllib
seq=urllib.request.urlopen("https://www.ebi.ac.uk/ena/data/view/
M11717&display=text&download=tst&filename=M11717.txt").read()
```

- 2) The preceding block will download a particular entry from the ENA (European Nucleotide Archive) and put it in the variable "seq". We will now manipulate this block of text.
- 3) Figure out what protein is represented by the gene in this record (by looking at the string). Add a block to your notebook with your answer.
- 4) Extract just the sequence block from the string as a new string (there are several ways you could do this)
- 5) The sequence will have numbers, spaces and newlines in it. This is a typical DNA sequence representation. Come up with a strategy for removing all of this annotation, and produce another string containing just the sequence itself (acgt only).