

Exercise1: if,then

Write some python code that checks if a given number (entered from the command line) is odd or even. The script accepts a single integer parameter and checks if it's an even or odd number. If even, it prints "Even" and if odd, it prints "Odd".

```
import sys

number = int(sys.argv[1])
if number % 2 == 0:
    print "Even"
else:
    print "Odd"
```

(Not a robust program. Can you make it safer? Hint. Run the program with no input)

Exercise 2: Conditional

Write a python code that accepts a single integer,

- if the number is less than 50 and greater than 0 print "Minor"
- if the number is greater than or equal to 50 and greater than 1000 print "Major"
- Otherwise prints "Severe"

```
import sys

number = int(sys.argv[1])
if (number < 50) and (number > 0):
    print "Minor"
elif (number >= 50) and (1000 > number):
    print "Major"
else:
    print "Severe"
```

Additional Question. Can you see a problem in the logic you have been given? How might you fix it?

Exercise 3:

Repetition:

for loop - different from C, Java - arithmetic progression of numbers.
python for loop iterates over items of any sequence (list or string).

Exercise: Add the contents of the following list

```
scores = [85.0, 75.0, 95.0, 110.0, 56.0]
```

```
sum=0;
scores = [85.0, 75.0, 95.0, 110.0, 56.0]
```

```
for score in scores:
```

```
    sum += score
```

```
print "Sum is "+sum
```

Exercise 4: loops and files

(1) Get a copy of SwissProt, an annotated Protein Database. On the shell command line type:
wget

ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete/uniprot_sprot.fasta.gz

(or look at drop canvas <http://dropcanvas.com/faykm>)

```
gunzip uniprot_sprot.fasta.gz
```

(2) Use the unix 'head' command to view top and format of file. It is known as FASTA

http://en.wikipedia.org/wiki/FASTA_format

(3) Write a Python script that

- opens the file

- reads each line of the file

- if the first character of the line is a '>' character AND the protein is a Human one (i.e contains the string "OS=Homo sapiens" then count the number of lines

- at end , print out the number of human proteins inside SwissProt

To do this you will need:

- some kind of loop to read each line of the file. What termination condition should you use?

- one or more conditional statements to isolate the Human description lines

```
reader = open('uniprot_sprot.fasta', 'r')
```

```
line = reader.readline()
```

```
noOfHumanProteins=0
```

```
for line in reader:
```

```
    if ( line[0] == '>'):
```

```
        #print line
```

```
        if "Homo sapiens" in line:
```

```
            noOfHumanProteins+=1
```

```
print "No = ",noOfHumanProteins
```

```
f.close()
```

Lots of issues with this solution. Do we check if the file open correctly, and if not, what the error message is? Is "Homo sapiens" the only spelling? Case insensitive searches? Why do we check for first line[0]=='>', and not just search each line?

If you are really bored by this exercise, find the human protein(s) with the longest amino-acid sequence. Note that sequences cross line boundaries, so you will need to read an arbitrary number of lines for every Human Protein found.