

Exercise of Programming Language, Homework E4

Write 6 Python programs to solve the following questions. Please name your program files as *Q1.py*, *Q2.py*, and so on, *i.e.*, according to the serial number of questions. All data files you need can be obtained from the e3 system. Compress your programs along with the data files into a ZIP file and submit the ZIP to e3.

1. Sort the lines in file `animals.txt` in descend order (from greatest to least) and print the results on the screen. Your output should look like this:

```
Woody Cowboy Human
Tigger Tiger
Sulley Monster
Simba Lion
Pluto Dog
... etc ...
```

2. Given a list,

```
numbers = [164.9, 150.3, 148.7, 189.2, 160.0, 174.3, 154.2, 201.6, 184.9,
182.1, 179.4, 193.8]
```

Please write a program to calculate these values,

(1) The minimum. (2) The maximum. (3) The average. (4) The median.

Hint: Your program shall be applicable even if the number of elements in the list changed in the future.

For instance, if `numbers = [1, 3, 9]`, the median is 3; if `numbers = [1, 3, 4, 9]`, the median is 3.5.

3. There are two numbers x and y saved in the first and second line respectively in the file `xy.txt`. Write a program to read this file obtaining the value of x and y , and then calculate and print out the results of the following formula,

(1) C_y^x

(2) P_y^x

4. A secret bee decided to live in a garden full of flowers. When this bee collected pollen from a flower, the flower would die immediately. In the first day, this bee collected pollen from half of the flowers in the garden. On the way home, it decided to collect pollen from two more flowers. Next day, the bee collected pollen from half of the remaining flowers again, as well as from two additional ones on the way home. After repeating this behavior for 11 days, in the morning of the 12th day, the bee found that there was only one flower left. How many flowers were there in the garden before the bee started collecting pollen.

5. Given an integer n in your Python program, use “loop” to make a graph of n lines in the style shown below,

Value	$n = 1$	$n = 2$	$n = 3$
Figure	<pre> ** </pre>	<pre> * _ _ * * * * * </pre>	<pre> * _ _ _ _ * * * _ _ * * * * * * * </pre>

Please print the graph to the screen and also save it as a text file in the `output` folder.

6. Here is a fractional sequence: $\frac{5}{3}, \frac{13}{7}, \frac{33}{15}, \frac{81}{35}, \frac{197}{83}, \frac{477}{199}, \frac{1153}{479}, \dots$

What is the sum of the first 20 fractions of this sequence? (Hint: $\frac{a(t)}{b(t)} = \frac{2a_{(t-1)} + b_{(t-1)}}{a_{(t-1)} + 2}$)

7. Using the file `protein.txt`, write a program that will extract the “odd” number residues and concatenate them into a new sequence. Do the same to the “even” number residues. Note that here, in human languages, we have to consider the first residue of the protein sequence to be numbered as 1, rather than 0. Your output may look like this:

```

Odd residues: ACAQT...
Even residues: GNGLV...

```

8. Write a program to extract the exon fragments from the DNA sequence available in `genomic_dna.txt` and output the concatenated (merged) sequence of these exons into `genomic_dna.coding.txt`. The range of exons are listed below,
 Exon 1: form the 1st nucleotide to “GTA” (GTA included)
 Exon 2: from the 90th nucleotide to the end of the original sequence

9. FASTA file format is a commonly-used DNA and protein sequence file format. Sequence(s) in FASTA format look like this:

```

>sequence_one
ATCGATCGATCGATCGAT
>sequence_two
ACTAGCTAGCTAGCATCG
>sequence_three
ACTGCATCGATCGTACCT

```

where sequence_XXX is a header that describes the sequence (the “>” symbol indicates the start of the header line). Write a program that will create a FASTA file for the following sequences – make sure that all sequences are in upper case and only contain the standard bases A, T, G and C (and “-”). For any position with a base other than the standard ones, replace it with an “x”.

Sequence header	DNA sequence
ABC123	ATCGYACGATCGATCGATCGCYAGACGTATCG
DEF456	actgatogacgatogatcgaycacgact
HIJ789	ACTGAC-ACTGT--AOTGTA----CATGTG

An unfinished Python program `Q9.tmp.py` contains variables storing the above sequence information. You may finish this program and rename it as `Q9.py` before you submit it.

- There are several DNA sequence fragment files saved in the folder `dna_fragments`. Write a program to obtain the sequence fragments and concatenate (merge) the fragments in the order of filenames. For instance, the fragment from `dna_fragment_1.txt` should be concatenated in front of that from `dna_fragment_2.txt`.