Homework 8

Download the following files from Collab/Resources/Python/Data/HW8 and save them in the same directory on your computer:

```
original.py
input-1000.txt       (1000 molecule definitions)
input-10000.txt      (10,000 molecule definitions)
```

The code works with a fake equation that is meant to bear a vague similarity to a chemistry calculation. Specifically, it computes a function.

$$E = \sum_{j<i} \left(\frac{e^{(r_{ij}q_i)} e^{(r_{ij}q_j)}}{r_{ij}} - \frac{1}{a} \right)$$

Where `q` is the "charge", a is a constant, and `rij` is a distance given by

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

The data are fake "coordinates" of "molecules".

Task 1:

a) Run the `original.py` script using the `input-10000.txt` data file (command line argument 1) and a cut off of 0.9 (command line argument 2). If your computer chokes on processing 10,000 molecules, switch to using the `input-1000.txt` data file for this entire assignment.

b) Post the console output as first part of your Collab submission.

c) Run the script 4 more times and calculate (for all 5 script runs):
   average for "Time to read coord file"
   average for "Time to calculate E"
   average for "Total Execution Time"
   Post the averages as second part of your Collab submission.

Task 2:

a) Optimize the python code to improve its performance. Eliminate as many for loops as you can and apply other optimization tricks that were discussed in class. Change the data structure if necessary to achieve this. Add appropriate timing calls into the code.

b) Run your optimized script 5 times and post the output as described for Task 1 to Collab. Make sure that the calculated values for `Num Pairs` and `Total E` are the same as obtained for the `original.py` script. Upload the optimized python script file to Collab.

c) Document all optimization steps with a brief rationale. Based on averaged timing values, calculate the performance improvement of the optimized code relative to the `original.py` code. Include the calculated performance improvement in the documentation and upload the documentation as separate file to Collab.

Note: You will not be graded based on achieved performance improvement, but rather on your approach and documentation of the optimization.