DATA 200 Spring 2020 Homework 3

Due February 21, 2019 at 11:59pm

Note: Please only use standard python covered in class lectures. Python libraries such as pandas, numpy are *not* allowed in your work.

1. Write python code to generate a csv file that contains student grades from user input. The program should prompt user to enter grade record (student ID, grade). It will continue to prompt for new record until the user specifies an empty record. It will then prompt the user for to enter a filename to save all the records:

```
Enter grade record (press Enter to stop): 003519342,22.5
Enter grade record (press Enter to stop): 007053869,20
Enter grade record (press Enter to stop):
Enter filename to save: hwl.csv
```

Once the filename is entered, a new file with the specified filename will be created with all the records specified by the user. You can use this program to generate input files to the next problem below.

(10 points)

2. Write Python code that prompts the user to specify a csv filename, read grades from the file, compute and output simple statistics of the grades. The input csv file should have the following format (i.e. student ID, points):

```
003519342,22.5
007053869,20
005255827,18
006034733,23.5
005467659,23
```

Process all the records in the file and print out the following information. Also, save the same information to a file (results.txt).

```
# of students in file: 5
Average of grades: 21.4
Highest grade: 23.5
Lowest grade: 18
```

Your code should be able to handle any input csv file with any number of student records. It should handle exceptions like incorrect filename (or file not found), empty grade file and non-numeric student ID and/or grade values.

(15 points)

3. Write Python code to read the provided file sonnet.txt and print out the top 15 most used words and the counts in the file by length 2, 3, 4, ..., 10 respectively. To achieve this, you can write several functions to perform the different tasks such as read from the file and print the information out. When processing the file content, you should remove these special characters: ?!.,;: (i.e. question mark, exclamation mark, period, comma, semi-colon and colon). The output from your code should look like the following:

(15 points) Top 15, length of word: 2 to 408 my 390 of 369 in 323 is 167 me 164 so 145 be 142 as 121 it 111 by 93 do 84 on 80 or 78 no 78 Top 15, length of word: 10 themselves 6 remembered 4 conscience 3 complexion 3 unkindness 2 unfathered 2 triumphant 2 temptation 2 speechless 2 resembling 2 possession 2 possessing 2 pilgrimage 2 perfection 2

obsequious 2