EM624 Midterm Exam - Fall 2022

Results for Section 1 will be submitted in a separate .doc/.pdf file.

Please submit your results for Section 2 and 3 in one <u>single</u> .py file.

Section 1: General questions

1. How you define software engineering and how is related to computer science?
2. In Python, can you describe when pandas may be a better option than lists and dictionaries? EX04 could be an example
3. Describe the 3-4 most relevant things to make your code easier to read and maintain
4. Why is important to define a testing strategy when coding? How would you perform testing for your code?
5. Why/when you may want to use functions in Python?

Section 2: Code checking

The following scripts are modified versions of students' solutions to exercises proposed a previous semester.
You will find a brief descriptions and the student's solution.
For each of them:

a. Check if it's doing what it was supposed to do
b. Describe what is wrong, if any
c. Fix it

If an input is required from the user, be sure the input testing is performed in the proper way.

<u>Please note: a .py file with the 3 snippets is provided on Canvas to help you working with the code</u>

1. Write a script that takes a string of 3 characters as input from the user and prints a string where for every character in the original, there are two characters (example: 'The' → 'TTThhheee')

```
N = input("Enter your characters: ")
L = []
for letters in N:
    letters.split()
    L.append(letters)
    print (L*3)
```

2. Write a script that calculates the 2 longest words of a text stored in a file and print them from the longest to the smaller of the 2. Please note:
   - Assume that the file contains *n* records, each one composed by 1 word. Words can be present more than once, but only unique words need to be considered
   - A sample word_list.csv file is attached for testing.

```
handle = open('word_list.csv','r')
top2 = ["",""]
for line in handle:
```

```
   #For each line in the file, strip the input and put it into the word variable
   word = line.strip()
   #Compare the length of each incoming word to the length of each word in each
position
   for i in range(0,2):
      top2.sort(key = len)
      if (len(word) < len(top2[i])):
         top2[i] = word
#Print the words
print ("\nThe 2 longest words are:"), top2
```

3. Write a script that takes a character (i.e. a string of length 1) as input from the user and returns True if it is a consonant, False otherwise. A check on the length of the input string and its being alphabetical is required and if not, send a message to the user and ask again

```
while True:
    #prompts and receives user input
    char = input('Please enter an alphabetical character:')
    if len(char) > 1: #checks if input is more than one character
        print ('Invalid input')
    else:
        if char == 'a' or 'e' or 'i' or 'o' or 'u' or 'y': #checks if input is a
vowel
            print ('False')
    else:
        print ('True')
```

Section 3: Writing code

4. Write a script reads a file “ai_trends.txt”, into a list of words, eliminates from the list of words the words in the file “stopwords_en.txt” and then calculates
    a. The 5 most frequent words and the 5 least frequent words
    b. The average occurrence of the words. Occurrence is the number of times a word is appearing in the text
    c. The longest and the shortest words
    d. The average word length. This is based on the unique words: each word counts as one

5. Write a script reads a file “cars.csv”, into a pandas structure and then print
    a. the first 3 rows and the last 3 of the dataset
    b. the cars with the 5 lowest values of the ratio average-mileage/horsepower

   For the name of the cars in your prints in b., use "car ID" + "company" + "body-style", with a "-" in between.