**EM624 Midterm Exam - Spring 2023-2**

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

**Please submit your results for Section 2 and 3 in one single .py file.**

**Section 1: General questions**

1. How you define agile methodology?
2. In Python, can you describe when pandas may be a better option than lists and dictionaries?
3. Why it is important to write a story out of the data to get insights on a given data-supported fact?
4. 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:

* 1. Check if it’s doing what it was supposed to do
  2. Describe what is wrong, if any
  3. Fix it

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

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

1. Write a script that takes a string as input from the user and prints a string where for every character in the original, there are 4 characters (example: 'The' → 'TTTThhhheeee')

N = input("Enter your characters: ")

L = []

for letters in N:

letters.split()

L.append(letters)

print (L\*4)

1. 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.

# --Original code

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

1. Write a script that takes a character (i.e. a string of length 1) as input from the user and returns if the character is a consonant or not. A check on the length of the input string 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**

1. From the file “keanumycins.txt” create a list of the words in it, eliminate from that list the words in the file “stopwords\_en.txt” and then from the resulting list calculate and print the following:
   1. The number of unique words
   2. The 5 most frequent words
   3. The relevance of the 5 most frequent words. The relevance is calculated by dividing the frequency of the word by the number of unique words.