Name: Dhaval Gogri

SMU ID : 47444609

Course : Advanced Application Programming

Quest : Quest 2 – Regex

*#Quest: Regex, Files, Urls***import** re  
**import** requests  
  
  
\_\_STUDENT\_ID\_\_ = **"47444609"** *# replace with your 8 digit student id*\_\_QUEST\_NAME\_\_ = **"Code "** *# replace with your coding name - max 15 characters***def** count\_vowels(mystr):  
 matchObject = re.findall(**r'[aeiou]'**, mystr, re.IGNORECASE)  
 **return** matchObject.\_\_len\_\_()  
  
  
**def** is\_valid\_python\_hex(mystr):  
 matchObject = re.search(**r'^(0{1}x{1})[A-F0-9]+$'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return True  
 else**:  
 **return False  
  
def** has\_vowel(mystr):  
 matchObject = re.search(**r'[aeiou]'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return True  
 else**:  
 **return False  
  
  
def** is\_integer(mystr):  
 matchObject = re.search(**r'^(-?)[0-9]+$'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return True  
 else**:  
 **return False  
  
  
def** get\_extension(mystr):  
 matchObject = re.search(**r'^[a-z].\*[.]{1}[a-z]+$'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return** (**''** + re.findall(**r'[a-z]\*$'**, mystr, re.IGNORECASE)[0])  
 **else**:  
 **return 'NONE'  
  
def** is\_number(mystr):  
 matchObject = re.search(**r'^(-?)[0-9]+?[.]{0,1}[0-9]\*$'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return True  
 else**:  
 **return False  
  
 return True  
  
def** convert\_date\_format(mystr):  
 matchObject = re.search(**r'^(\d{4})-(\d{1,2})-(\d{1,2})$'**, mystr, re.IGNORECASE)  
 **if** matchObject:  
 **return** re.sub(**r'(\d{4})-(\d{1,2})-(\d{1,2})'**, **'\\2-\\3-\\1'**, mystr)  
 **else**:  
 **return 'NONE'***#File functions***def** readFileCountLines(filename):  
 **with** open(filename, **'r'**) **as** myfile:  
 data = myfile.read()  
 print(data)  
 matchObject = re.findall(**r'\n'**, data, re.IGNORECASE)  
 print(matchObject.\_\_len\_\_() + 1)  
 **return** (matchObject.\_\_len\_\_() + 1)  
  
  
  
**def** readFileCountStringOccurrences(filename, stringval):  
 **with** open(filename, **'r'**) **as** myfile:  
 data = myfile.read()  
 matchObject = re.findall(**r'('** + re.escape(stringval) + **')'**, data, re.IGNORECASE)  
 print(matchObject)  
 **return** matchObject.\_\_len\_\_()  
  
**def** readFileSumDigitsGreaterThanNumber(filename, number):  
 **with** open(filename, **'r'**) **as** myfile:  
 data = myfile.read()  
 matchObject = re.findall(**r'\d[0-9]+'**, data, re.IGNORECASE | re.MULTILINE)  
 print(matchObject)  
 sum = 0  
 **for** allNumbers **in** matchObject:  
 **if** int(allNumbers) > number:  
 sum = sum + int(allNumbers)  
 **return** sum  
  
  
**def** remove\_all\_but\_alpha(mystr):  
 *""" remove all characters that are not alpha a-z A-Z  
 >>> remove\_all\_but\_alpha('hey-99-where8isthe\*\*big\_table\*\*') -> 'heywhereisthebigtable'  
 """* matchObject = re.findall(**r'[a-z]+'**, mystr, re.IGNORECASE)  
 print(matchObject)  
 **if** matchObject:  
 allAplha = **""  
 for** alpha **in** matchObject:  
 allAplha = allAplha + alpha  
 **return** allAplha  
 **else**:  
 **return 'NONE'***#URL functions***def** readurlCountStringOccurrences(urlname, stringval):  
 response = requests.get(urlname)  
 matchObject = re.findall(**r'('** + re.escape(stringval) + **')'**, response.content.decode(**'utf-8'**), re.IGNORECASE)  
 **if** matchObject:  
 **return** matchObject.\_\_len\_\_()  
  
 **return** 0  
  
  
**def** readurlCountValidPhoneNumbers(urlname):  
 response = requests.get(urlname)  
 matchObject = re.findall(**r'([0-9]{3}[.][0-9]{3}[.][0-9]{4})|([0-9]{3}[-][0-9]{3}[-][0-9]{4})|([0-9]{10})'**, response.content.decode(**'utf-8'**), re.IGNORECASE)  
 **if** matchObject:  
 **return** matchObject.\_\_len\_\_()  
  
 **return** 0  
  
  
  
**if** \_\_name\_\_ == **'\_\_main\_\_'**:  
 print (**"To test your code execute: python test\_QuestFilesUrls.py or on command line execute: pytest "**)

OUTPUT :

