

PROGRAM 5: REGULAR EXPRESSIONS IN PYTHON

Requirement:

Mr. Sherlock Holmes is investigating a suicide case of Mr. Kalman a noted celebrity. He strongly believes it is murder. But he needs to come up with strong evidence(s) to prove it. He receives an anonymous letter which contains some random words pasted from the newspaper letters. As you are the budding data analysts and RE experts he expects you to help him in finding out the real culprit. Please do help him. ALL THE BEST!

Random words received are: f, ope, ife, Mon, 50, +2, @, Kalman.org, toronto, Ind, 963819, SIA, league, dam, xer, ugs, ape, Cor, bit and end.

Note: You are free to create your own crime scene and characters.

BBC: Breaking news

Mr. Kalman, a popular Indian sportsman! found dead at Hotel in China, East Asia.

Chinese Crime Branch has specially appointed Mr. Sherlock Holmes to handle the case. Kalman found death by Hanging himself. Holmes investigates Kalman's suicide case in his direction, suspicious loss suspected of being a murder. However, he requires to come up with vital pieces of evidence to prove it.

Case details

Mr. Holmes Started the investigation from the crime spot. He has recognized there was no vital clue that he was expecting. Still, Holmes continued the hunt on collecting hints. Shockingly he spotted Kalman's dairy, which was left open with dropped ink marks. Mr. Holmes firmly believed that the dairy would be his crucial evidence for this case.

Soon he folded all his evidence and left the space quietly. Mr. Holmes started reading Kalman's dairy passionately and wondered how Mr. Kalman struggled and lifted success in his whole life. Match-fixing and game betting was one addictive activity that pushed Kalman to end life into the struggle. At the time, when Holmes built his initial inquiry. He was surprised by an anonymous letter that contains some random words pasted from the newspaper letters.

Anonymous letter

```
In [2]: import re
import numpy as np
import random

def letter_list():
    print('At the moment, Holmes receives an unknown anonymous letter that contains some random words pasted from the ne
    print("\n\n")
    teams_list = ["", "", "", "", ""]
    data = np.array([[ 'xer', 'ope', 'ife', 'mon', 'ind'], ['963819', 'sia', 'f', 'dam', 'league'], ['ugs', 'ape', 'cor',

    a="Random words received on anonymous letter"
    print("\t\t"+"*"+'-'*75+"*")
    print("\t\t"+"|{: ^75s}|".format(a.upper()))
    print("\t\t"+"*"+'-'*75+"*")

    row_format = "{:>15}" * (len(teams_list) + 1)
    print(row_format.format("", *teams_list))

    for team, row in zip(teams_list, data):
        print(row_format.format(team, *row))
```

In [3]: letter_list()

At the moment, Holmes receives an unknown anonymous letter that contains some random words pasted from the newspaper letters.

```
*-----*
|                RANDOM WORDS RECEIVED ON ANONYMOUS LETTER                |
*-----*

          xer          ope          ife          mon          ind
        963819        sia          f          dam        league
          ugs          ape          cor          bit          end
```

Mr. Holmes way of investigation : He finds Kalman's Diary

```
In [4]: def Last_page():
        a_file = open("WOLFOFWALLSTREET.txt", "r")

        list_of_lists = []
        for line in a_file:
            stripped_line = line.strip()
            line_list = stripped_line.split()

        a_file.close()

        return line_list
```

```
In [5]: print("\n\t*****")
print("\t\t\t\t\t KALMAN'S DIARY ")
a_file = open("WOLFOFWALLSTREET.txt", "r")
for line in a_file:
    stripped_line = line.strip()
    print("\n\n",stripped_line)
print("\n\t*****")
```

KALMAN'S DIARY

26th of october (monday),2020:As usual my day started with a hot bed coffee. took bath and I stared with my routine. I used to go for walking but due to the pandemic I am unable to go out. Every corner of the world is talking about this coronavirus. I am worried about the uncertainty of the life. I am blessed as I will be able to travel to toronto by the next week with my wife in a private jet. Usually I will keep my diary in the book shelf but today by mistake I kept it in my cupboard. damn !! I am growing older and I do not remember certain things. As we were planning to go to toronto next week, I suddenly remembered my last trip with my wife to Coorg. It was such a beautiful memory. We had a nice time there. Whenever I tease her, she used to scold me like. You are such a monster. I loved her very much and she was everything to me. We are almost 60 and we tried everything in Coorg. On the 2nd day of our trip, all of our plans were spoiled due to heavy rain. There was stagnant water outside our room's window and I do not actually believe that we did such a childish act. We were making paper boats to play in that stagnant water. We really acted like babies. The next day a friend of mine came to visit us and he accompanied with us to wildlife sanctuary My wife was endoscopic surgery patient and she was very afraid of wild animals. I made fun of her by comparing her with the monkey Anyways we had a great time there. Later that evening we went for a site seeing where we enjoyed the lake view using the telescope. Our trip was unforgettable. I hope that my next visit to toronto will be very exciting as we I and my "colleagues" have planned to go for a league match. I am much waiting for the next week to go my favorite hotel and my favorite room number 750. I am a bit emotional and sentimental so I always wish to stay in the same room number. More than 5,00,00 people have died due to coronavirus, but god has blessed me with good health and I am able to lead a very highlife with my loved one. My life had many hardships and I managed to overcome them. I hope my trip to toronto will be memorable one as I will spend time with my beloved wife. Note: 50 50000 750 +2 +21 @ @kalman.org sia asian asia euroleague damaged xerox fireplugs bugs paper grape escape correct bit coin end coffee flower envelope rabbit girlfriend trend weekend lender wallpaper habit corncakes corpsmen escaped mugs tape kidnaper fan rope friend life knife money jugs lemon monsoon monster drugs window blind boxer exercise.

Mr. Holmes suggests hints related to case and anonymous letter

```
In [6]: def first_word():
s= " "
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with ope :")
        x = re.findall("(ope)", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [7]: def second_word():
s=" "
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the 5 letter words ending with xer :")
        x = re.findall("..xer", txt)         # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [8]: def third_word():
        s=" "
        ch="yes"
        while(ch=="yes"):
            t=True
            while (t):
                txt = input("Enter the words ending with ife :")
                x = re.findall("ife$", txt)                # Findall
                if x == []:
                    t=True
                else:
                    s=s+txt+" "
                    t=False
            ch=input("Do you want to continue:")
        return s
```

```
In [9]: def fourth_word():
        s=" "
        ch="yes"
        while(ch=="yes"):
            t=True
            while (t):
                txt = input("Enter the words staring with mon :")
                x = re.findall(r"(mon|tues|wednes|thurs|fri|sat|sun)(day)", txt)                # Findall
                if x == []:
                    t=True
                else:
                    s=s+txt+" "
                    t=False
            ch=input("Do you want to continue:")
        return s
```

```
In [10]: def fifth_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words ending with ind :")
        x = re.findall("\w+ind", txt)          # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [11]: def six_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the number with 963819**** :")
        x = re.findall("\d{10}", txt)          # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [12]: def seven_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with sia :")
        x = re.findall("\w+[a-z]+", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [13]: def eight_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with f :")
        x = re.findall("(f)", txt)               # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```



```
In [14]: def nine_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with dam :")
        x = re.findall(r"\bdam", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [15]: def ten_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with league :")
        x = re.match(r"league\b", txt)         # match
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [16]: def eleven_word():
s=" "
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with ugs :")
        x = re.findall("\D{5}", txt)
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

Findall

```
In [17]: def tewlve_word():
s=" "
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with cor :")
        x = re.findall("\w", txt)
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

Findall

```
In [18]: def thirteen_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with ape :")
        x = re.findall("\w+ape*", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [19]: def fourteen_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with end :")
        x = re.findall("end\\Z", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [20]: def fifteen_word():
s=""
ch="yes"
while(ch=="yes"):
    t=True
    while (t):
        txt = input("Enter the words with bit:")
        x = re.findall("bit+", txt)           # Findall
        if x == []:
            t=True
        else:
            s=s+txt+" "
            t=False
    ch=input("Do you want to continue:")
return s
```

```
In [21]: def user_words():
words = []
ch = 'yes'
i = 1
print(" ENTER THE WORDS: \n\n")
while ch == 'yes':
    if i== 1:
        print("\033[1m"+"Words with ope"+" \033[0m")
        print("=====")
        word=first_word()
    elif i==2:
        print("\033[1m"+"Words with xer"+" \033[0m")
        print("=====")
        word=second_word()
    elif i==3:
        print("\033[1m"+"Words with ife"+" \033[0m")
        print("=====")
        word=third_word()
    elif i==4:
        print("\033[1m"+"Words with mon"+" \033[0m")
        print("=====")
        word=fourth_word()
    elif i==5:
        print("\033[1m"+"Words with ind"+" \033[0m")
        print("=====")
        word=fifth_word()
    elif i==6:
        print("\033[1m"+"Words with 963819"+" \033[0m")
        print("=====")
        word=six_word()
    elif i==7:
        print("\033[1m"+"Words with sia"+" \033[0m")
        print("=====")
        word=seven_word()
    elif i==8:
        print("\033[1m"+"Words with f"+" \033[0m")
        print("=====")
        word=eight_word()
    elif i==9:
        print("\033[1m"+"Words with dam"+" \033[0m")
        print("=====")
```

```

        word=nine_word()
    elif i==10:
        print("\033[1m"+"Words with league"+" \033[0m")
        print("=====")
        word=ten_word()
    elif i==11:
        print("\033[1m"+"Words with ugs"+" \033[0m")
        print("=====")
        word=eleven_word()
    elif i==12:
        print("\033[1m"+"Words with cor"+" \033[0m")
        print("=====")
        word=tewlve_word()
    elif i==13:
        print("\033[1m"+"Words with ape"+" \033[0m")
        print("=====")
        word=thirteen_word()
    elif i==14:
        print("\033[1m"+"Words with end"+" \033[0m")
        print("=====")
        word=fourteen_word()
    elif i==15:
        print("\033[1m"+"Words with bit"+" \033[0m")
        print("=====")
        word=fifteen_word()
    else:
        break

x = re.split("\s", word)                # split
for j in x:
    if j!=' ':
        words.append(j)

ch = input("\n"+" \033[1m"+"Do you want to continue with next match:"+" \033[0m")
print("*****")
if ch=='yes':
    i=i+1
else:
    break
while('' in words) :
    words.remove('')

```

```
return words
```

CLUES FOR INVESTIGATION

In [22]:

```
def last_page_text(Last_Page):  
    txt = ''  
    for i in Last_Page:  
        txt+=i  
        txt+= ' '  
    return txt
```

In [23]:

```
def word_find(i,txt):  
    l = ''  
    try:  
        assert re.search(i, txt)           # Search  
        l+=i  
        return l  
    except (AttributeError,AssertionError):  
        pass
```

In [24]:

```
def evidence_list(words,txt,evidence):  
    for i in words:  
        w=word_find(i,txt)  
        if w!=None:  
            evidence.append(w)  
  
    return evidence
```

```
In [25]: def investigation():
    evidence=[]
    Last_Page = Last_page()
    words=user_words()
    txt=last_page_text(Last_Page)
    evidence_list(words,txt,evidence)

    t=True
    while(t):
        random.shuffle(evidence)                                # shuffle
        print("\n\t"+"\\033[1m"+"Matched words are:"+"\\033[0m")
        print("\t"+"\\033[0m"+"=====")
        for i in evidence:
            print("\t",i)
        ch=input("\n Do you want to shuffle list:")
        if ch == "yes":
            t=True
        else:
            t=False
            pass
    return evidence
```



```
In [26]: final=[]  
final=investigation()
```

Words with ope

=====

Enter the words with ope :rope

Do you want to continue:no

Do you want to continue with next match:yes

Words with xer

=====

Enter the 5 letter words ending with xer :fixer

Do you want to continue:no

Do you want to continue with next match:yes

Words with ife

=====

Enter the words ending with ife :wife

Do you want to continue:no

~ ~ ~ ~ ~

UPDATING MORE DETAILS ON HINTS

```
In [27]: final
```

```
Out[27]: ['wife',  
          'drugs',  
          'corps',  
          'league',  
          'asia',  
          'monday',  
          'kidnaper',  
          'blind',  
          'fan',  
          'friend',  
          'rope',  
          'damage']
```

```
In [28]: txt = last_page_text(final)
```

```
In [48]: Case = re.sub('rope', 'nylon_rope', txt)           # sub
Case = re.sub('wife', 'kalman_wife', Case)
Case = re.sub('drugs', 'poison_drugs', Case)
Case = re.sub('corps', 'corps_men', Case)
Case = re.sub('league', 'match_league', Case)
Case = re.sub('blind', 'blindly', Case)
Case = re.sub('fan', 'hung_in_fan', Case)
```

```
In [49]: print("\nFinal Clues aboutt the case:")
print("=====\n")
w=[]
x = re.split("\s", Case)           # split
for j in x:
    if j!='':
        w.append(j)
for i in w:
    print(i)
```

```
Final Clues aboutt the case:
=====
```

```
kalman_wife
poison_drugs
corps_men
match_league
asia
monday
kidnaper
blindly
hung_in_fan
friend
nylon_rope
damage
```

Mr. Holmes Further Takes Actions on Clues

KALMAN'S WIFE AND HIS FRIENDS WILL BE CALLED FOR FURTHER INVESTIGATION.