project of regular expression

In [2]:
```
#question no 1

import re
text="india is. a   biggest,.economy"
x=re.sub(r"\s+|\.|,",":",text)
print(x)
```

```
india:is::a:biggest::economy
```

In [16]:
```
#question no 2
import pandas as pd
import re
data={'Teacher':['student,name!','xxxx Attendence','456marks,score:;in..?subject']}
df=pd.DataFrame(data)
df["Teacher"]=df["Teacher"].apply(lambda x:re.sub(r'[^\w\s]','',x))

print(df)
```

```
                Teacher
0           studentname
1        xxxx Attendence
2  456marksscoreinsubject
```

In [17]:
```
#question no 3
import re
pattern=re.compile(r"\w{4}")
text= "The example of a story be like there was a fox in the garden she was wonderi

matches=pattern.findall(text)

print(matches)
```

```
['exam', 'stor', 'like', 'ther', 'gard', 'wond', 'erin', 'here', 'ther']
```

In [19]:
```
#question no 4
import re
pattern=re.compile(r'\b\w{3}\b|\b\w{4}\b|\b\w{5}\b')
text= "The example of a story be like there was a fox in the garden she was wonderi

matches=pattern.findall(text)

print(matches)
```

```
['The', 'story', 'like', 'there', 'was', 'fox', 'the', 'she', 'was', 'here', 'an
d', 'there']
```

In [25]:
```
#question no 5
import re
pattern=re.compile(r'\w+')
text=['New(phone)','Work(shop)','Chat(corner)','hritik(sharma)']

matches=[pattern.findall(s) for s in text]

print(matches)
```

```
[['New', 'phone'], ['Work', 'shop'], ['Chat', 'corner'], ['hritik', 'sharma']]
```

In [ ]:
```
#question no 6
import re
pattern= re.compile
```

In [28]:
```
#question no 7
import re
text = "NothingIsImposibleTheWorldItselfSay'sIAmPossible"

pattern = re.compile(r'(?=[A-Z])')
result = pattern.split(text)

print(result)
```

```
['', 'Nothing', 'Is', 'Imposible', 'The', 'World', 'Itself', "Say's", 'I', 'Am',
'Possible']
```

In [34]:
```
#question no 8
import re
text='NothingIsImposible23TheWorldItself4Say'
pattern=r'(?<=\d)(?=[A-Za-z])'

x=re.sub(pattern,"  ",text)

print(x)
```

```
NothingIsImposible23  TheWorldItself4  Say
```

In [35]:
```
#question no 9
import re
text='NothingIsImposible23TheWorldItself4Say'
pattern=r'(?<=[a-z])([A-Z0-9])'

x=re.sub(pattern,r'\1',text)

print(x)
```

```
NothingIsImposible23TheWorldItself4Say
```

In [40]:
```
#question no 10
import pandas as pd

url = "https://raw.githubusercontent.com/dsrscientist/DSData/master/happiness_score
df = pd.read_csv(url)

df['first_five_letters'] = df['Country'].str[:6]
print(df.head())
```

```
       Country          Region  Happiness Rank  Happiness Score  \
0  Switzerland  Western Europe               1            7.587
1      Iceland  Western Europe               2            7.561
2      Denmark  Western Europe               3            7.527
3       Norway  Western Europe               4            7.522
4       Canada   North America               5            7.427

   Standard Error  Economy (GDP per Capita)   Family  \
0         0.03411                   1.39651  1.34951
1         0.04884                   1.30232  1.40223
2         0.03328                   1.32548  1.36058
3         0.03880                   1.45900  1.33095
4         0.03553                   1.32629  1.32261

   Health (Life Expectancy)  Freedom  Trust (Government Corruption)  \
0                   0.94143  0.66557                        0.41978
1                   0.94784  0.62877                        0.14145
2                   0.87464  0.64938                        0.48357
3                   0.88521  0.66973                        0.36503
4                   0.90563  0.63297                        0.32957

   Generosity  Dystopia Residual first_five_letters
0     0.29678            2.51738             Switze
1     0.43630            2.70201             Icelan
2     0.34139            2.49204             Denmar
3     0.34699            2.46531             Norway
4     0.45811            2.45176             Canada
```

In [42]:
```
#question no 11
import re
pattern= r'^[a-zA-Z0-9_]'
text='Hello folks i am student of.Datatrained my& batch no, is 1123 Ihave_done my:

formula=re.match(pattern,text)

print(formula)
```

```
<re.Match object; span=(0, 1), match='H'>
```

In [53]:
```
#question no 12
import re
text= "Cow was suffering from high fever and the deaseas was spreading quickly Cow
search=re.findall('Cow',text)
print(search)
```

```
['Cow', 'Cow']
```

In [ ]:
```
#question no 13
```

In [ ]:
```
#question no 14
```

In [2]:
```
#question no 15
import re

pattern= r'(fox|dog|horse)'
text= 'The quick brown fox jumps over the lazy dog.'

matches=re.findall(pattern,text)
print(matches)
```

```
['fox', 'dog']
```

In [12]:
```
#question no 16
import re
```

```
pattern='The quick brown fox jumps over the lazy dog.'

for yy in pattern:
    search=re.search('fox',pattern)
    print(search)
```

```
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
<re.Match object; span=(16, 19), match='fox'>
```

In [9]:

```
#question no 17
import re

text = 'Python exercises, PHP exercises, C# exercises'
pattern = r'exercises'

matches = re.findall(pattern, text)
print(matches)
```

```
['exercises', 'exercises', 'exercises']
```

In [ ]:

```
#question no 18
```

In [ ]:
```
#question no 19
```

In [16]:
```
#question no 20
import re
text="01.12 0132.123 2.31875 145.8 3.01 27.25 0.25"
prog=re.compile(r'\b\d+\.\d{1,2}\b')
result=prog.findall(text)
print(result)
```

```
['01.12', '145.8', '3.01', '27.25', '0.25']
```

In [ ]:
```
#question no 21
```

In [20]:
```
#question no 22
import re

sample_text = 'My marks in each semester are: 947, 896, 926, 524, 734, 950, 642'
numeric_values = [int(x) for x in re.findall(r'\d+', sample_text)]
max_value = max(numeric_values)
print("Maximum numeric value:", max_value)
```

```
Maximum numeric value: 950
```

In [21]:
```
#quetion no 23
import re
text= "RegularExpressionIsAnImportantTopicInPython"
match=re.findall('[A-Z][a-z]*', text)
print(match)
```

```
['Regular', 'Expression', 'Is', 'An', 'Important', 'Topic', 'In', 'Python']
```

In [22]:
```
#question no 24
import re

text = "This is a Test, and Another test. This Is Another Test."
pattern = r'[A-Z][a-z]+'

matches = re.findall(pattern, text)
print(matches)
```

```
['This', 'Test', 'Another', 'This', 'Is', 'Another', 'Test']
```

In [26]:
```
#question no 25
import re
text="Hello hello world world"
pattern = r'\b(\w+)(?:\s+\1\b)+'

result = re.sub(pattern,r'\1', text, flags=re.IGNORECASE)
print(result)
```

```
Hello world
```

In [32]:
```
#question no 26
```

In [35]:
```
#question no 27
import re
text="""RT @kapil_kausik: #Doltiwal I mean #xyzabc is "hurt" by #Demonetization as
pattern = r'#\w+'
hashtags = re.findall(pattern, text)
print(hashtags)
```

```
['#Doltiwal', '#xyzabc', '#Demonetization']
```

In [36]:
```
#question no 28
import re
text="@Jags123456 Bharat band on 28??<ed><U+00A0><U+00BD><ed><U+00B8><U+0082>Those
pattern=re.sub(r'<U\+[0-9A-Fa-f]+>', '', text)
print(pattern)
```

```
@Jags123456 Bharat band on 28??<ed><ed>Those who  are protesting #demonetization
are all different party leaders
```

In [37]:
```
#question no 29
import re
text='Ron was born on 12-09-1992 and he was admitted to school 15-12-1999.'
date_pattern = r'\b\d{2}-\d{2}-\d{4}\b'
dates = re.findall(date_pattern, text)
print(dates)
```

```
['12-09-1992', '15-12-1999']
```

In [38]:
```
#question no 30
import re
text="The following example creates an ArrayList with a capacity of 50 elements. 4
pattern = re.compile(r'\b\w{2,4}\b')
x=pattern.sub('',text)
print(x)
```

```
 following example creates  ArrayList  a capacity   elements. 4 elements   added
ArrayList   ArrayList  trimmed accordingly.
```

In [ ]:

In [ ]: