Question 1 – Write a Python program to replace all occurrences of a space, comma , or dot with a colon.

Sample Text- ‘Python Exercises, PHP exercises.’

Expected output: Python : Exercise ::PHP : Output exercises

Answer –To replace all occurrences of a space , commas , or dot with a colon in a given text, you can use the replace () method in python .Here’s a python program that demonstrates this .

Text = ‘ Python Exercises, PHP exercises.’

#Replace space, comma ,and dot with a colon

New\_text = text.replace (‘ ‘, ‘: ‘). Replace (‘ ‘, ‘: ‘).Replace (‘ ‘, ‘: ‘).

Print ( new \_text)

Output: Python : Exercise ::PHP : Output exercises

Question2- Create a dataframe using the dictionary below and remove everything (commas (,), ! , XXX, ,, etc.) from the columns except words.

Answer- 1. Import pandas as pd

Import pandas as pd

2. Create the dictionary :

Data ={‘ SUMMARY’ : [‘hello, world’, ‘xxxx test’, ‘123four , five ,six….]}

3.Create the dataframe :

Df= pd.dataframe(data)

4. Remove everything except words from the columns:

Df [‘SUMMARY’]= df[‘SUMMARY’].str.replace(‘[^a-zA—z/s]’, ‘’, regex= true)

5. Print the expected output:

Print (df)

The output will be:

SUMMARY

0 hello world

1. Test
2. Four five six

Question3- Create a function in Python to find all words that are the least 4 characters long in a string . The use of the re.compile () method is mandatory.

Answer- To find all words that are at least 4 charactres long in a string using the re.compile() method in python , you can follow these steps :

1. Import the re module :

Import re

1. Define a function , let’s call it find \_long\_words, that takes a string as input:

Def find \_long\_words(string):

1. Compile a regular expression pattern using re.compile () to match words that are at least 4 characters long:

Pattern = re.compile(r’\b\w{4,}\b’)

In this pattern , \b represents a word boundary ,\w matches any alphanumeric character, and {4,} specifies that the world should be at least 4 characters long.

1. Use the findall ()method of the compiled pattern to find all matches in the input string :

Long\_words = pattern.findall(string)

1. Return the list of long words :

Return long\_words

Here’s the complete code:

Import re

Def find\_long\_words(string):

Pattern = re.compile(r’\b\w{4,}\b’)

Long\_words = pattern .findall (string)

Return long\_words

You can then call the find \_long\_words() function with a string as an argument to find all words that are at least 4 characters long . For example

Text = “ this is a sample string with some lomg words like apple, banana, and orange.”

Result = find \_long\_words(text)

Print (result)

Output :

[‘ sample’, ‘string’, ‘with’, ‘some’, ‘long’, ‘words’, ‘like’, ‘apple’, ‘banana’, ‘orange’]

Queston4 - Create a function in Python to find all three , four , and five characters words in a string .The use of the re.compile() method mandatory .

Answer – To find all words that are the least three, four, or five characters long in a string using the re.compile method in Python , you can follow these steps :

1. Import the re module :

Import re

1. Define a function that takes a string as input:

Def find \_words (string)

1. Use the re.compile method to create a regular expression pattern that matches words of at least three, four , or five characters long.

Patten = re,compile (r’\b\w{3,5}\b’)

In this pattern ,\b represents a word boundary , \w matches any alphanumeric characters , and {3,5} specifies that the word should be at least three characters long and at the most five characters long .

1. Use the finda11 method of the pattern object to find all matches in the string .

Matches = pattern.find11(string)

1. Return the list of matches

Return matches

Putiing it all together , the complete function would look like this :

Import re

Def find\_words(string):

Pattern = re.compile(r’\b\w{3,5}\b’)

Matches = pattern .finda11(string)

Return matches

You can then call this function with a string to find all words that are at least three , four, or five characters long:

String = “this is a sample string with words of different lengths.”

Result = find\_words(string)

This will output a list of words that matches the specified criteria.

Question 5- Create a function in Python to remove the parenthesis in a list of string . The use of the re.compile() method is mandatory.

Answer – To remove th parentheses in a list of string using the re.compile() method in python , you can follow these steps:

1. Import the re module work with regular expressions.
2. Define a function , let’s call it remove \_parentheses, that takes a list of strings as input.
3. Inside the function , create a regular expression pattern using re. compile() to match the parentheses. The pattern should be”()”.
4. Letrate over each string in the input list using a for loop .
5. Use the re.sub () function to replace the parentheseses with an empty string in each string of list .Pass the regular expression pattern and an empty string as arguments to re. sub().
6. Return the modified list of strings.

Here’s the code implementation :

Import re

Def remove \_parentheses (string):

Pattern = re.compile(r”\(\”)

Modified \_strings =[]

For string in strings:

Modified \_string = re.sub(pattern ,””, string )

Modified \_strings.append (modified\_string)

Return modified \_strings

You can test the function withb the given the sample test :

Sample \_text =[“ example (.com)”, [hr@fliprobo (.com)](mailto:hr@fliprobo%20(.com)), “github(.com)”,

“ hello ( data science world)”, “data (scientist)” ]

Result = remove \_parentheses (sample \_text)

Print ( result)

Output:

[ ‘example.com’, ‘hr@fliprobo.com’, ‘github.com’, ‘Hello Data Science World ‘, ‘ Data scientist’]

Question 6- Create a function in Python to remove the parenthesis area from the text stored in the text file using Regular Expression .

Answer – To remove the parenthesis area from the text stored in a text file using regular expression Python , you can follow these steps :

1. Read the text file and store its content in a variable .
2. Import the re module for regular expressions.
3. Use the re.sub() function to replace the parenthesis area with an empty string.
4. Save the modified text back to the text file .

Here’s an example code sinppet that demonstrates this process:

Import re

# Read the text file and store its content in variable with open (‘filename.txt’, ‘r’ as file :

Text = file.read()

# Use regular expression to remove the parenthesis area

modified \_text = re.sub(r’\([^()]\*\),”, text)

# Save the modified text back to the text file

with open (‘filename :txt’, ‘w’) as file:

file .write ( modified\_text)

Question 7- Write a regular expression in Python to split a string into uppercase letters

Answer . To spilt a string into uppercase letters using regular expressions in Python , you can use the re module . Here’s an example of how you can achieve this:

Import re

Text = “ ImportanceofRegularExpressionsIn Python”

Result = re.fina11(‘[A-Z] [^A-Z]\*’, text)

Print ( result )

Output: [ ‘Importance ‘, ‘of’, ‘regular’ ,’Expression’, ‘In ‘ , Python ,]

Question 8- Create a function in Python to insert spaces between words starting with numbers .

Answer – To insert spaces between words starting with number in python , you can use regular expression and the re module. Here’s a function that achieves this:

Import re

Def insert \_ spaces (text):

# Use regular expression to find words starting with numbers

Pattern = r’(\d+) ([A-Za-z]+)’

Result = re.sub ( pattern , r’\1 \2’, text )

Return result

To Test the function , you can call it with the sample text provided:

Text = “ RegularExpression11sAn 21ImportantTopic 3InPython “

Output = insert spaces ( Text )

Print result

Question 9- Create a function in Python to insert spaces between words starting with capital letters or with numbers .

Answer – To insert spaces between words starting with capital letters or numbers In Python , youn can use regular expression and the re module . here’s a function that achieves this :

Import re

Def insert \_spaces (text):

# Use regular expression to find words starting with capital letters or numbers

Pattern = r’ ([A-Z] [a-z0-9] +|\d+)’

# Replace the matched words with a space followed by the word result = re.sub (pattern , r’ \1,text)

#Remove any leading or trailing spaces

Result = result.strip()

Return result

You can test the function with the given sample text :

Sample \_text = “RegularExpression11sAn2ImportantTopic3InPython”

Output = insert \_spaces (sample\_text)

Print ( output )

RegularExpression 1 Is An 2 Important Topic 3 InPython

Question10 – Use the github link below to read the data and create a dataframe. After creating the dataframe extract the first 6 letters of each country and store in the dataframe under a new column called first \_five\_letters.

Github Link- <https://raw.githubusercontent.com/dsrscientist/DSData/master/happiness_score_dataset.csv>

Question11 – Write a Python program to match a string that contains only upper and lowercase letters , numbers , And underscores.

Answer – Here’s a python program that users the re module to match the desired pattern :

Import re

Def match \_string (string):

Pattern = r’^[a-zA-Z0 -9\_]+$’

If re.match (pattern , string ):

Print (“ String matches the pattern “)

Else :

Print (“ string does not match the pattern “)

# Example usage

Match \_string(“Hello\_world123”) #Output :String matches the pattern

Match \_string(“Hello World”) #output :String does not match the pattern

Question12 – Write a Python program where a string will start with a specific number.

Answer- To write a Python program where a string starts with a specific number , you can use the start with () method . This method checks if a string starts with a specified substring and returns True Or false accordingly.

Here,s an example program that demonstrates this :

Def check \_starting \_number (string, number):

If string . startwith(str(number)):

Return True

Else:

Return false

# Example usage

String = “ 123abc”

Number =123

If check \_starting\_number (string, number):

Print (“The string starts with the specified number ,”)

Else:

Print (“the string does not starts with the specified number.”)

Question13 – Write a Python program to remove leading zeros from an IP address.

Answer- To Remove leading zeros from an ip address in python , you can use the split () method to split the ip address into its individual octets , and then use the int () function to convert each octect to an inteagar .finally , you can use the join () method to combine the octets back into a string without leading zeroes .

Here’s an example program that demonstrates this :

Def remove \_ leading \_zeros (ip \_ address ):

# split the IP address into octets

Octets = ip \_address.split(“,”)

# Remove leading zeros from each octet

Octets \_without \_zeros = [ str(int (octet)) for octect in octets ]

# Join the octets back into a string

Ip\_address \_without\_zeros= remove \_leading \_zeros (ip\_address)

Print (ip\_address\_without\_zeros) #Output :192.168.1.1

Question14 – Write a regular expression in python to match a date string in the form of Month Name followed by day number and year stored in a text file.

Answer- To match a date string in the from of the month Name followed by day number and year Python you can the following regular expression .

Import re ‘

Text = “ On August 15th 1947 that India was declared independent from british colonialism , and the reins of control were handed over to the leaderof the country .”

Pattern = r”\b [A-Z] [a-z]+) \d{1,2}(?:st|nd|rd|th)? \d{4}\b”

Matches = re.finda11(pattern ,text)

Question 15- Write a python program to search some literals string in a string .

Answer - import re

Patterns = [‘fox’, ‘dog’, ‘horse’,]

Text =’ The quick brown fox jumps over the lazy dog.’

For the pattern in patterns:

Print (‘searching for “%s” in “%S” -> % ( Pattern , Text),)

If re . search ( pattern , text ):

Print (‘matched !’)

Else:

Print ( ‘not Matched’)

Question16– Write a Python program to search a literals string in string and also find the location within the original string where the pattern occurs .

Answer – import re

Pattern = ‘fox’

Text= ‘ The quick brown fox jumps over the lazy dog .’

Match= re.serach(pattern, text)

S= match.start()

E= match .end()

Print (‘ found “%s” in “%S” from %d to %d,%\

( match.re.pattern, match.string,s,e))

Question 17– Write a Python program to find the substrings within string.

Answer- Import re

Text = ‘ Python excercies , PHP exercises, C# exercises’

Pattern = ‘ exercises’

For match in re.finda11(pattern, text):

Print (‘Found “%S” ‘ %match )

Question18 – Write a Python program to find the occurrence and position of the substrings within a string.

Answer- Import re

Text = ‘ Python excercies , PHP exercises, C# exercises’

Pattern = ‘ exercises’

For match in re.finditer(pattern, text):

S= match.start()

E=match.end ()

Print (‘found “%s” at %d:d’ % (text [s:e],s,e))

Question19 – Write a Python program to convert a date of yyyy-mm-dd format to dd-mm-yyyy format.

Answer – import re

Def change\_date\_format (dt):

Return re.sub(r’(\d{4})-(\d{1,2})-(\d{1,2})’, ‘\\3-\\2-\\1’,dt)

Dt1 = “2026-01-02”

Print (“original date in YYY-MM-DD Format :”,dt1)

Print (“New date in DD-MM-YYY Format : “, change\_date\_format(dt1)

Question20 – Create a function in Python to find all decimal numbers with a precision of 1 or 2 in a string . The use of the re.compile() method is mandatory .

Answer – To find all decimal numbers with a precision of 1 or 2 in a string using the re.compile() method in Python , You can follow these steps.

1. Import the re module .
2. Define a function ,lets call it find \_decimal\_numbers , that takes a string as input.
3. Use the re,compile () method to create a regular expression pattern that matches decimal numbers with a precision of 1 or 2. The pattern can be r’\d+\.d{1,2}’, where:

* \d+ matches one or more digits before the decimal point .
* \. Matches the decimal point.
* \d{1,2 } matches one or two digits after the decimal point

1. Use the re.finda11() method to find all matches of the pattern in the input string
2. Return the list of matched decimal numbers .

Here’s the implementation of the function:

Import re

Def find \_decimal \_number (string ):

Pattern = re. compile (r’\d+\.\d{1,2}.)

Decimal \_numbers = re.finda11(pattern, string )

Return decimal \_numbers

To test the function with the given sample text , you can call it like this:

Output

The output will be [’01.12’,’145.8’,’3.01’, ’27.25’,’0.25’], which matches the expected output.

Question21 – Write a Python program to separate and print the numbers and theirs position of agiven string .

Answer- Import re

# input.

Text = “ The following example creates an arrayList with a capacity of 50 elements. Four elements are for m in re.finditer(“\d+”,text):

Print (m.group (0))

Print (“ index position :”, m.start())

Question22 – Write a regular expression in python program to extract maximum/largest numeric value from string.

Answer – To extract the maximum or largest numeric vaule from a string using the regular expression in Python , you can follow these steps:

1. Import the re module , which provides support for regular expression.
2. Define the input string that contains the numeric values.
3. Use the re.finda11() function to find all the numeric values in the string.
4. Covert the extracted numeric values from strings to integers using int () function.
5. Use the max () function to find the maximum value from the list of integers.

Here’s the Python code that accomplishes this.

Import re

Input \_string = ‘ My marks in each semester are:947,896,926,524,734,950,642’

Numeric \_values = re.finda11(r’\d+’,input\_string)

Numeric\_values = [int (vaule) for value in numeric\_values]

Max\_vaule = max ( numeric\_values)

Print ( max\_vaule)

Question23 – Create a function in Python to insert spaces between words starting with capital letters. Answer - To insert spaces between words starting with capital letters in python , you can use regular expression and the re module . here’s function that achieves this.

Import re

Def insert \_spaces(text):

# Use regular expression to find words starting with capital letters pattern = r’ ([A-Z] [a-z]+)’

# Replace the found words with the same word followed by a space result = re.sub ( pattern ,r’\1’, text)

# Remove any leading or trailing spaces

Result = result.strip()

Return result

The output will be:

Regular expression is an important topic in Python

Question24 – Python regex to find sequences of one upper case letter followed by lower case letters .

Answer - To find sequences of one uppercase letter followed by lowercase letter using Python regex , you can use the following pattern :

Import re

Pattern = r’[A-Z] [a-z] +’

Text =” This is a simple Text with multiple matches “

Matches = re.finda11(pattern , text )

Print (matches)

This will output:

[ ‘This’, ‘Sample ‘,’ Text’, ‘Multiple’, ‘Matches’]

Question25 – Write a Python program to remove continuous duplicate words from sentence using

regular Expression .

Answer – To remove continuous duplicate words from a sentences using regular expression in Python , you can use the re module. Here’s a Python program that achieves this.

Import re ‘def remove \_duplicate (sentence):

Pattern = r’\b(\w+)(\s+\1\b)+’

Result = re.sub(pattern , r’\1, sentence)

Return result

# Example usage

Sentence = “ Hello hello world world”

Output = remove \_duplicates (sentence)

Question26 – Write a Python program using RegEx to accept string ending with alphanumeric character.

Answer – To write Python program using regular expression ( RegEx) to accept a string ending with an alphanumeric character , you can use the re module , Here’s example program.

Import re

Def check\_string(string):

Pattern = r”\w$”

Match = re. search ( pattern , string )

If match :

Return true

Else:

Return false

# Example usage

Input \_ string = input (“Enter a string:”)

If check\_ string (input \_string):

Print ( “string ends with an alphanumeric character”)

Else:  
print (“string does not end with an alphanumeric charcters”)

Question27 – Write a Python program using RegEx to extract the hashtags.

Ans. To extract the hashtags from a given text using regular expression in Python , you can use the re module. Here’s a Python program that accomplishes this:

Import re

Def extract \_hashtags ( text ):

Hashtags = re.finda11(r’#\w+’,text)

Return hashtags

#Sample text

Text = ‘RT @Kapil\_kausik: #doltiwal I mean #xyzabc is “hurt” by # Demonetization as the same has rendered USELESS <ED> <U+00BD<>ED><U+00B1><U+0089>”acquired funds “NO wo’

# Extract hashtags

Hashtags = extract \_hashtags(text)

# Print the extracted hashtags

Print (hashtags )

Question28 – Write a Python program using RegEx to remove <U+..> like symbols .

check the below sample text , there are strange symbols something of the sort <U+..> all over the place. You need to come up with general Regex expression that will cover all such symbols.

Answer – To remove the <U+..> Sumbols from the given text using regular expression in Python , you can follow these steps :

1. Import the re module to work with regular expressions.
2. Define a regular expression pattern that a matches the <U+..>symbols. The pattern r “ <u\+\w{4}>” can be used to match the symbol.

* R before the pattern denotes a raw string to avoid escaping special characters .
* <u\+matches the literal characters <u+.
* \w{4}matches any four alphanumeric characters .
* >matches the literal character>.

1. Use the re.sub () function to replace all occurrences of the pattern with an empty string.

* Pass the pattern , an empty string , and the input text as arguments to re.sub().
* Assign the result to variable , let’s say output\_text.

1. Print the output\_text to see the modified text without the <U+..> symbols .

Here’s the Python code that implements tha above steps.

Import re

Input \_text = “@Jags123456 Bharat band on 28??<ed><U+00A0><U+00BD><ed>< U+00B8><U+0082> Those who are protesting # demonetization are all different party leaders”

Pattern = r”<u\+\w{4}”

Print ( output \_text)

The excepted output will be:

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

The regular expression pattern r”<U\+\w{4}>”Matches all <U+…> symbols in the input text and replaces them with an empty string , effectively removing them from the text.

Question29 – Write a Python program to extract dates from the text stored in the text file .

Answer – To extract dates from the text file in Python , you can use regular expression . Here’s a Pytho program that demonstrates how to extract dates from a text file:

Import re

# Open the text file

With open (‘Filename.txt; ‘r’) as file:

Text = file.read()

# Define the regular expression [pattern for dates pattern = re.finda11(pattern , text)

# print the extracted dates

For date in dates :

Print (date)

Question30 – Create a function in python to remove all words from string of length between 2 And 4.

The use of the re.compile() method is mandatory.

Answer – To remove all words from a string of length between 2 and 4 in Python using re.compile() method , you can follow these steps :

1. Import the re module .
2. Define a function , lets call it remove\_words, that takes a string as input .
3. Use the re.compile() method to create a regular expression pattern that matches words of length between 2 and 4. The pattern can be defined as r’\b\w{2,4}\b’.
4. Use the re.Sub() method to replace all occurrences of the pattern with an empty string .
5. Return the modified string .

Here’s the code that implements the above steps:

Import re

Def remove \_words (string):

Pattern = re.compile (r’\b\w{2,4}\b’)

Modified \_string = re.sub( pattern , ‘’, string)

Return modified \_string

To test the function , you can use the provided sample text as input:

Sample\_text =” The following example creates an Arraylist with a capacity of 50 elements . 4 elements are then added to the arraylist and the arraylist is trimed accordingly.”