How to extract email from a text using re module in Python

You will learn

How to use the findall function of re module to extract email from a text

Slide this image to see a step by step practical explanation →

Author

Ashbab khan

Before jumping into the main syntax let's learn some basic concepts

What is the findall function in re module?

findall function is generally used to find something from a text whether it is a number, string, or anything

1. \w

This syntax returns the character from a text which contains the value 0-9 and a-Z and _

2. \.

This syntax returns . (dot) from a text

3. [ ]

This is used for different sets of chachaters for example, we use [\w\.] + for finding the \w syntax as well as \. syntax

1. [ ]+

This syntax returns multiple characters for example, if we use [\w] it means that it makes a list of one character but if use [\w]+ it makes a list of words

Importing re module and using our data

In [1]:
```
import re
```

In [67]:
```
data = '''Hii my name is ashbab khan and my
 you can find me on linkedin using @ashbab and my
 mail is demo@gmail.com and my second mail ww
 is abc-demo@gmail.com  ab@g.g'''

data
```

Out[67]:
```
'Hii my name is ashbab khan and my \n you can find me on linkedin using @ashbab
and my\n mail is demo@gmail.com and my second mail ww\n is abc-demo@gmail.com  a
b@g.g'
```

Using the findall function to extract email

4]:
```
extracted_email = re.findall('[\w\.-]+@[\w]+\.[\w]+',data)
extracted_email
```

4]:
```
['demo@gmail.com', 'abc-demo@gmail.com', 'ab@g.g']
```

Using this list of emails with pandas data frame

In [81]:
```
import pandas as pd
```

In [82]:
```
email = pd.DataFrame({'Email':extracted_email})
```

In [83]:
```
email
```

Out[83]:

	Email
0	demo@gmail.com
1	abc-demo@gmail.com
2	ab@g.g