Python Regex Tutorial: re.match(),re.search(), re.findall(), re.split()

# Regular expressions (Regex)

A regular expressions is special kind of sequence of character that is used to match a string in different applications. One big example of regex is email validation.

When you sign in any website the normal pattern is that the browser sends a request to the server and ask if the email and password are correct. But if millions of users are sending requests then it is better to filter out emails which have a correct format. To check the format of email programmers use regex on the client-side application that makes sure that the email has the correct format. For example, the regex will help us to check if the user has entered “@” in the email at the correct place and “.com” at the end of the email. This type of checking is done using regex.

Let’s get to know about that special kind of sequence of character.

* **[ ]** *used to specify a special class. [1234]*
* **.** *(dot) it matches any character except a new line.*
* **\*** *it means zero or more*
* **+** *means one or more*
* **–** *it is used to express range:- [a-z]*

Now let tell you about some predefined sets

* **\d** *matches any decimal digits of class [0-9]*
* **\D** *matches any non-decimal character. [^0-9]*
* **\w** *matches any alphanumeric. [a-zA-z0-9 ]*
* **\W** *matches any non alphanumeric [^a-zA-Z0-9 ]*

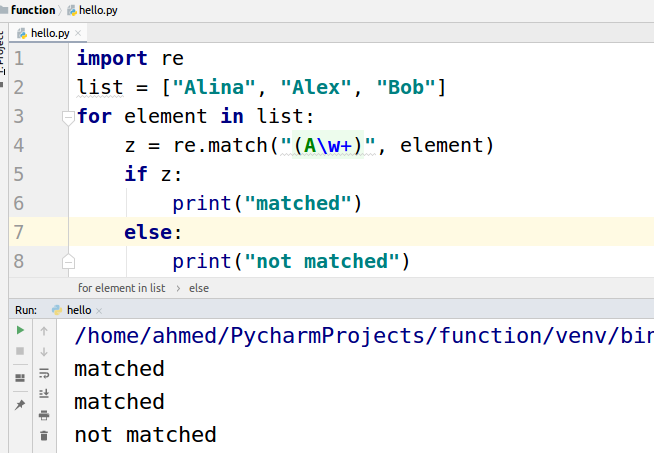
## re.match()

re.match() function checks whether a string matches a specific format or not. It returns true or false.

In the example below, we created a regex “ (A\w+) “ that means find a string that starts with “A” and \w means any alphanumeric and + stands for one or more. The whole string will be described as any string that starts with ‘A’ and contains any length.

|  |
| --- |
| import re list = ["Alina", "Alex", "Bob"] for element in list:  z = re.match("(A\w+)", element)  if z:  print("matched")  else:  print("not matched") |

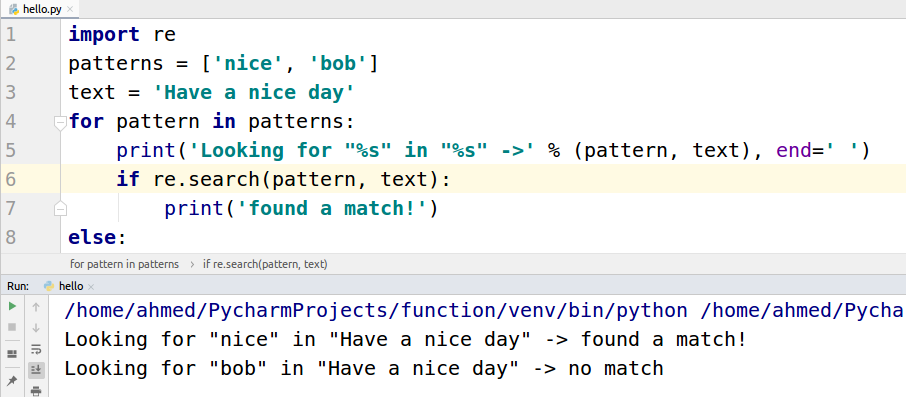
The following will be the output.



## re.search()

re.search() function is used to find the first occurrence in the required string. If we have a string “Have a nice day” and we want to find out whether “Bob” is present in this string or not we will use re.search()

|  |
| --- |
| import re patterns = ['nice', 'bob'] text = 'Have a nice day' for pattern in patterns:  print('Looking for "%s" in "%s" ->' % (pattern, text), end=' ')  if re.search(pattern, text):  print('found a match!') else:  print('no match') |

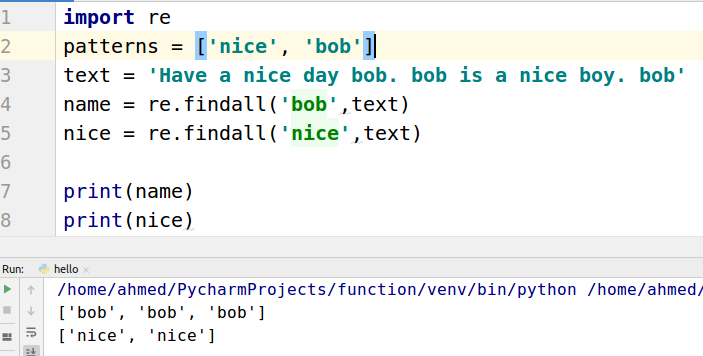


## re.findall()

The function re.search() returns when it finds the first occurrence but on the other hand re.findall() search for all the occurrence of the words in the string.

|  |
| --- |
| import re patterns = ['nice', 'bob'] text = 'Have a nice day bob. bob is a nice boy. bob' name = re.findall('bob',text) nice = re.findall('nice',text)  print(name) print(nice) |

The following will be the output.



## re.split()

Suppose you have a string of names space-separated and you want to separate them. For this purpose re.split() function is used.

|  |
| --- |
| import re  string = 'Ana B0B Ali' pattern = '\W'  result = re.split(pattern, string) print(result) |

The following will be the output.

