

17/September RE specifies a set of strings that matches it.
it will let you check if string matches Regular Expression

Regex — Corey Schafer

Raw String

→ string prefixed with `r` that tells Python not to handle ~~the~~ backslashes in special way. (✓)

» backslashes normally used to specify tabs, newlines `\t`

import re module i.e. `re`

```
print('\tTab')
```

`print(r'\tTab')` → backslash not handled in a special way.

re.compile method will allow us to separate out our patterns into a variable and reuse this variable

```
pattern = re.compile(r'abc')
```

```
matches = pattern.finditer(text to search)
```

```
for match in matches:  
    print(match)
```


Metacharacters (need to be escaped):

• \wedge \$ * + ? \{ } [] \ / ()

dot is special character.

for example

~~re.compile~~ print('r'.') → weird behaviour

~~re.compile~~ print('r\'.') →

• → Any character except any character except new line

\wedge → matches beginning of line (start with)

\$ → Matches end of the string (star end with)

\backslash → Matches previous character. (20 occurrence)

+ → Matches previous character (1 or more times (≥ 1))

? → 1 or more 0 times (0 or 1)

\wedge → matches only end of string

\wedge → matches - non-word characters = $[\wedge A-Za-z0-9]$

\wedge → matches alpha numeric in ASCII character set.
= $[a-zA-z0-9]$

\wedge → Matches any character which is not whitespace character.

\wedge → matches whitespace character - (tab, newline, vertical tab)

\wedge → not a decimal digit

\wedge → matches - digit = $[0-9]$

\wedge → matches empty string if it's not at end or start of word.

\wedge → matches empty string, but only at end and beginning of word

\wedge → Matches only start of string → "The"

Some Methods

- 1) `re.match (pattern, to-match)`
- 2) `re.sub (pattern, with-pattern, actual-word)` → replace matches
- 3) `re.findall (pattern, from-string)` → return list containing all matches
- 4) `re.finditer (pattern, text)`
- 5) `re.search (pattern, text)`
- 6) `re.compile ()`
- 7) `re.split ()` → return list containing all the matches