REGULAR EXPRESSIONS
In Python

Arsalan Ali
@arslanchaos

Special Characters in Regex

```
.          - Period Means Any Character Except New Line
\d         - Digit (0-9)
\D         - Not a Digit (0-9)
\w         - Word Character (a-z, A-Z, 0-9, _)
\W         - Not a Word Character
\s         - Whitespace (space, tab, newline)
\S         - Not Whitespace (space, tab, newline)

\b         - Word Boundary (No characters before it)
\B         - Not a Word Boundary (Has characters before it)
^          - Begining of a string
$          - End of a string


[ ]        - Matches characters inside the bracket
[^ ]       - Not Matches characters inside the bracket
|          - Either OR
( )        - Group


--- Quantifiers ---
\*         - 0 ore more
\+         - 1 or more
?          - 0 or 1
{3}        - Exact number of times
{3,4}      - Range of numbers
```
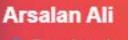
Special Characters

Arsalan Ali
@arslanchaos

Fetch Phone Numbers

```python
# Sample TEXT
text = '''
abcdefghijklmnopqurtuvwxyz
coreyms.com
0336-7564851
0301.5672890
0333*1548769
ABCDEFGHIJKLMNOPQRSTUVWXYZ. ^ $ * + ? { } [ ] \ | ( )
'''
# Import regex
import re
# Enter the pattern of the phone numbers
pattern = re.compile(r"\d{3}[-.*]\d{7}")
# Search for pattern in the TEXT
matches = pattern.findall(text)
# Loop to get the strings that matched the pattern
for numbers in matches:
    print(numbers)


# OUTPUT
0336-7564851
0301.5672890
0333*1548769
```
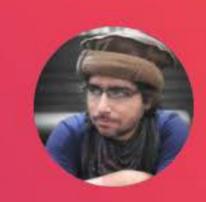
Arsalan Ali

@arslanchaos

Fetch Emails

```python
# Sample TEXT of emails
emails = '''
CoreyMSchafer@gmail.com
corey.schafer@university.edu
corey-321-schafer@my-work.net
'''

# Import regex
import re
# Create a pattern to match all emails
pattern = re.compile(r".*@.*(\w|\.)*")
# search for pattern in the TEXT
matches = pattern.findall(emails)
# Loop to get the emails
for email in matches:
    print(email)

# OUTPUT
CoreyMSchafer@gmail.com
corey.schafer@university.edu
corey-321-schafer@my-work.net
```

Fetch Email Addresses

Arsalan Ali
🐦 @arslanchaos

Fetch Domain Name

```python
# Sample TEXT with full address
urls = '''
https://www.google.com
http://coreyms.com
https://youtube.com
https://www.nasa.gov
'''

# Import Regex
import re
# Create a pattern to match entire address
pattern = re.compile(r"https?://(www\.)?(\w+\.\w+)")
# Find pattern in the TEXT
matches = pattern.findall(urls)
# Loop over the entire matches
for domain in matches:
# We've used groups above in patter so print only group 2
    print(domain.group(2))

# OUTPUT
google.com
coreyms.com
youtube.com
nasa.gov
```

Arsalan Ali

🐦 @arslanchaos

```python
# Sample TEXT with full address
text = '''
cat
mat
rat
bat
'''


# Import Regex
import re
# Create a pattern that ignores the name BAT
pattern = re.compile(r"[^b]at")
# Find pattern in the TEXT
matches = pattern.findall(text)
# Loop over the entire matches
for name in matches:
    print(name)


# OUTPUT
cat
mat
rat
```

Ignore Specific Pattern

Arsalan Ali
@arslanchaos