Python has a built-in package called re, which can be used to work with Regular Expressions.

Regular expressions (regex) in Python are a powerful tool for pattern matching and searching within strings. Python's re module provides functions and methods to work with regular expressions.

|  |  |
| --- | --- |
| **Function** | **Description** |
| [findall](https://www.w3schools.com/python/python_regex.asp#findall) | Returns a list containing all matches |
| [search](https://www.w3schools.com/python/python_regex.asp#search) | [Returns a 1ST Match object if there is a match anywhere in the string](https://www.w3schools.com/python/python_regex.asp#matchobject) |
| [split](https://www.w3schools.com/python/python_regex.asp#split) | Returns a list where the string has been split at each match |
| [sub](https://www.w3schools.com/python/python_regex.asp#sub) | Replaces one or many matches with a string |

|  |  |  |
| --- | --- | --- |
| **Character** | **Example** | **Description** |
| **[]** | "[a-m]" | A set of characters |
| **\** | "\bhello" | **Signals a special sequence (can also be used to escape special characters)** |
| **.** | "he..o" | Any character (except newline character) |
| **^** | "^hello" | Starts with |
| **$** | "planet$" | Ends with |
| **|** | "falls**|**stays" | Either or |
| **()** | (apples|oranges) | Capture and group |

|  |  |
| --- | --- |
| **Set** | **Description** |
| [a-n] | lower case character, alphabetically between a and n |
| [A-N] | UPPER case character, alphabetically between A and N |
| [0-9] | digit between 0 and 9 |
| [a-zA-Z] | character alphabetically between a and z, lower case OR upper case |
| [a-zA-Z0-9] | character alphabetically between a and z, lower case OR upper case with all numeric Numbers |
| [**^**a-zA-Z0-9] | Non word(numeric and alpha char) |
| [arn] | Returns a match where one of the specified characters (a, r, or n) is present |
| [**^**arn] | character EXCEPT a, r, and n |
| [0123] | Returns a match where any of the specified digits (0, 1, 2, or 3) are present |

|  |  |  |
| --- | --- | --- |
| Character | Example(PATTERN) | Description |
| \b | r"\bRain" | Word boundry |
|  | r"Rain\b" |  |
|  | r"\bRain\b" |  |
| \A | "\AThe" | start with The |
| \Z | "Spain\Z" | **end with Spain** |
| \d | "\d" | digits (numbers from 0-9) |
| \D | "\D" | string **or non-digits** |
| \s | "\s" | **white space** |
| \S | "\S" | Non White Space |
| \w | "\w" | string contains any word characters  characters from a to Z,  digits from 0-9,  and the underscore \_ character) |
| \W | "\W" | string **DOES NOT contain any word characters** |

|  |  |  |
| --- | --- | --- |
| Character | Example | Description |
| . | "he..o" | Any character (except newline character) |
| .? | "he.?o" | Zero or one occurrences |
| .\* | "he.\*o" | Zero or more occurrences |
| .+ | "he.+o" | One or more occurrences |
| .{n} | "he.{2}o" | Exactly the specified number of occurrences |