

Report of "Regular Expression" Assignment

Yuhuan Huang, Aug 2025

- Console Output:

```
Test Case 1:
Original: <p>Great product!</p> &#128516;
Cleaned: great product 128516
=====

Test Case 2:
Original: Check it out at https://example.com and www.test.com
Cleaned: check it out at and wwtestcom
=====

Test Case 3:
Original: Product ID: AB1234 - 5* rating!!!
Cleaned: product id ab1234 5 rating
=====

Test Case 4:
Original: EXCELLENT product... very HAPPY :)
Cleaned: excellent product very happy
=====

Test Case 5:
Original: Size: 10.5cm x 20cm - Perfect fit!
Cleaned: size 105cm x 20cm perfect fit
=====

...

Cleaned text:
['great product 128516', 'check it out at and wwtestcom', 'product id ab1234 5 rating', 'excellent product very happy', 'size 105cm x 20cm perfect fit', 'contact
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...

TestResults(failed=0, attempted=8)
```

- Reflections

1. What metacharacters represent any character except newlines, start of strings, end of strings, and escape special characters?

`.` is used to match a single character; `^` is used to match the start of a string; `$` is used to match the end of a string; `\` is used to escape special characters, making other metacharacters like `.` lose their special meaning in matching.

2. What characters are used to delimit matching letters or digits?

The square brackets `[]` is used to match the single characters inside the brackets. For example, `[a-z]` is used to match all lower-case letters while `[0-9]` is used to match all single digits. What's more, `\d` can also be used to match any digit while `\w` can be used to match all kinds of letters and digits.

3. What is the difference between "greedy" and "non-greedy" matching?

Greedy matching means to match as much as possible, while non-greedy matching means to stop when the first match happens.