Headers uit HTML 1

Neem de code van het volgende voorbeeld over.

```
import re

string = "De kat in de hoed wist dat de vleermuis daar was."
pattern = "de"

matches = re.findall(pattern, string, flags=re.I)  # De vlag re.I maakt de zoekopdracht case-insensitive

print(matches)
```

De output zijn alle delen van de tekst die de regex matchen. Probeer maar!

Stel je hebt een webpagina. Dus de string is een webpagina, bijvoorbeeld als volgt:

```
string="<!DOCTYPE html><html><head><title>Mijn webpagina</title></head><body><h1>Mijn eerste kop</h1><p>Dit is een paragraaf.</p><h1>Mijn tweede kop</h1><p>Dit is nog een paragraaf.</p></body></html>"
```

Zet deze string nu in de code en vervang daarmee regel 3 (de regel van "De kat....").

Pas nu de reguliere expressie zodat alle <h1> stukken uit de tekst worden gehaald.

Opdracht

Lees de tekst hierboven!

Pas de voorbeeldcode aan.

Aanpassing 1

Zet op regel 3 de html-code zoals hierboven is gegeven.

Aanpassing 2

Verander de reguliere expressie (regex) zodat je alle h1 elementen uit de HTML-code laat afdrukken.

Inleveren

1. Aangepaste code
2. Een screendump waarin je laat zien dat je code werkt.
   Let op, in de screendump moet je ook de **datum en tijd van Windows** zichtbaar maken!

Revision #7
Created 2023-07-26 18:08:11 UTC by Max
Updated 2023-07-31 13:44:49 UTC by Max