Corrigés de la semaine 6

🕳 regexp_pythonid - Semaine 6 Séquence 6 🕳

```
# un identificateur commence par une lettre ou un underscore
    # et peut être suivi par n'importe quel nombre de
    # lettre, chiffre ou underscore, ce qui se trouve être \w
    # si on ne se met pas en mode unicode
    regexp_pythonid = "[a-zA-Z_]\w*"
                      🗕 regexp_pythonid - Semaine 6 Séquence 6 🗕
    # on peut aussi bien sûr l'écrire en clair
    regexp_pythonid2 = "[a-zA-Z_][a-zA-Z0-9_]*"
                      🗕 regexp_specials - Semaine 6 Séquence 6 🕳
    # il faut commencer par exactement 2 underscores
    # donc le caractère suivant doit être une lettre
2
    # ensuite on peut mettre ce qu'on veut comme alphanumérique,
    # mais avant les deux derniers underscores on ne peut pas avoir
    # un underscore
    # enfin pour traiter le cas où la partie centrale est réduite
    # à un seul caractère, on met une option - avec ()?
    regexp_specials = "__[a-zA-Z](\w*[a-zA-Z0-9])?__"
```

```
regexp_url - Semaine 6 Séquence 6
     # en ignorant la casse on pourra ne mentionner les noms de protocoles
1
     # qu'en minuscules
2
     i_flag = "(?i)"
3
4
     # pour élaborer la chaine (proto1|proto2|...)
     protos_list = ['http', 'https', 'ftp', 'ssh', ]
6
                 = "(?P<proto>" + "|".join(protos_list) + ")"
     protos
7
     # à l'intérieur de la zone 'user/password', la partie
9
     # password est optionnelle - mais on ne veut pas le ':' dans
10
     # le groupe 'password' - il nous faut deux groupes
11
     password
               = r"(:(?P<password>[^:]+))?"
12
13
     # la partie user-password elle-même est optionnelle
14
                 = r"((?P<user>\w+){password}@)?".format(**locals())
     user
15
16
     # pour le hostname on accepte des lettres, chiffres, underscore et '.'
17
     # attention à backslaher . car sinon ceci va matcher tout y compris /
18
                 = r''(?P<hostname>[\w\.]+)''
     hostname
19
     # le port est optionnel
21
                 = r"(:(?P<port>\d+))?"
22
     port
23
     # après le premier slash
24
                 = r"(?P<path>.*)"
     path
25
26
     # on assemble le tout
27
     regexp_url = i_flag + protos + "://" + user + hostname + port + '/' + path
```