CIS 211 Lab 8 Name _____

Reference

Here are two helpful links: an introductory tutorial (`https://docs.python.org/3/howto/regex.html`), and a cheat sheet (`https://www.debuggex.com/cheatsheet/regex/python`).

Exercise 1

1.0 Simple Python syntax for using regular expression:

```
1  import re
2
3  result = re.search(r'pattern', 'string')
4  print(result.group(0))
5  print(result[0])
```

Don't forget the little `r` before the pattern string.

1.1 Replace line 3 of 1.0 with the following lines. What's printed for each of them? What does the pattern match?

(a) `result = re.search(r'[A-Z]*211', 'CIS210 CIS211 CIS212').`

(b) `result = re.search(r'(xy){2}', 'xy xyxy xyxyxy')`

(c) `result = re.match(r'[A-Z|a-z]+211\.', 'Cis211.  CIS210 CIS211')`

(d) What do `re.match` and `re.search` functions return? What is the main difference bewteen them?

Exercise 2

What does the code below print?

```
1  import re
2
3  result = re.match(r'(?P<course1>[A-Z|a-z]*210) (?P<course2>.*211)',
4                    'Cis210 CIS211 CIS212')
5  print(result.groupdict())
```

Challenge

Consider the set $S = \{$`''`, `'ab'`, `'aabb'`, `'aaabbb'`, `'aaaabbbb'`, ...$\}$.
Is it possible to give a regular expression that matches all the strings in set S, and matches no strings outside of S?