SI 507 Lab #8

October 25

Today's Plan

Lab #8

- Regular Expressions (REGEX)
- Project Help

What is a regex?

- A text string
  - foo
  - [A-Z]+:\d+
  - (?<=[a-z])(?=[A-Z])

What is a regex?

- A text string
  - foo matches the string *foo*
  - [A-Z]+:\d+
  - (?<=[a-z])(?=[A-Z])

What is a regex?

- A text string
  - foo matches the string *foo*
  - [A-Z]+:\d+ matches string fragments like *F:1* and *GO:30*
  - (?<=[a-z])(?=[A-Z])

What is a regex?

- A text string
  - foo matches the string *foo*
  - [A-Z]+:\d+ matches string fragments like *F:1* and *GO:30*
  - (?<=[a-z])(?=[A-Z]) matches the position in the string *CamelCase* where we shift from a lower-case letter to an upper-case letter

What is a regex for?

- **Find** text within a larger body of text
- **Validate** that a string conforms to a desired format
- **Replace** or **insert** text at matched positions
- **Split** strings

Activity

Regex Practice

Goals: Spend time practicing regex and getting to know some common characters.

Activity 1: Recommended Approach

```
1  import re
2
3  str1 = 'From: stephen.marquard@uct.ac.za Sat Jan walk@12  5 09:14:16
4  res1 = re.findall('\S+@\S+', str1)
5  print(res1)
6
7  ### guess res1
8  # one or more non-whitespace
9  # @
10 # one or more non-whitespace
11 ### stephen.marquard@uct.ac.za
12 ### walk@12
13
```

---Break down the regex
---Use a cheat sheet to help

---Make a prediction about what will print

Activity 1: Instructions

- 30 minutes
- Cheat Sheet:
  - **https://bit.ly/3Eoo6PE**
- Activity Link:
  - **https://bit.ly/3e2YPQz**
- If you finish early:
  - Ask questions about projects

Homework

Current Assignments

iTunes (Project 1): Due 10/17

20 Questions (Tree) (Project 2): Due 11/18

OR

Kevin Bacon (Graphs) (Project 3):

Due: 11/18

Sources

https://runestone.academy/runestone/books/published/py4e-int/regex/regex-group.html (activity 1)

https://docs.google.com/document/d/12jSMPc_-xca93TXG58IyAAH0dbVDNnRbqCuiLmyC49M/edit# (the key I made for activity 1)

https://www.rexegg.com/ (for best results, open in Firefox)

https://www.debuggex.com/cheatsheet/regex/python (searchable regex cheat sheet)

https://pythex.org/ (allows you to check your regular expressions!)