

Regex

Andrew Rosen

September 9, 2018

Abstract

This is extra credit and an incentive for you to learn Regex. Each page is worth 5 points.

1 Regex Dictionary

For this section, I want you to answer questions about the dictionary file `words.txt`. If you're using a Unix system, feel free to use the dictionary file included on your computer. Some words begin with capital letters! Please include each regex with your answer.

1. I like animals. How many words contain the word 'cat' or 'dog' in them?
2. Four letters words are supposedly naughty. How many four letter words are there?
3. Do more words end in "ing" or "ion?"
4. How many words contain a "q" not immediately followed by a "u."
5. How many words have no vowels?
6. How many words with two vowels **in a row** are there?
7. How many words with at least two vowels are there? The vowels need not be adjacent, like in the previous problem.

If you are having trouble, please see this link.

2 More Regex

1. What is the difference between `.*` and `.+?` ?
2. How would you write a regex that matches the full name of someone whose last name is Nakamoto? You can assume that the first name that comes before it will always be one word that begins with a capital letter. The regex must match the following:

- 'Satoshi Nakamoto'
- 'Alice Nakamoto'
- 'RoboCop Nakamoto'

but not the following:

- 'satoshi Nakamoto' (where the first name is not capitalized)
- 'Mr. Nakamoto' (where the preceding word has a nonletter character)
- 'Nakamoto' (which has no first name)
- 'Satoshi nakamoto' (where Nakamoto is not capitalized)

3. Create a regex which matches the strings twenty, twenty-one, twenty-three, ..., ninety-nine