Regular Expressions in Python Tutorial

Michelle Fullwood

PyLadies Meetup Aug 2013

Motivating regular expressions

Scenario: You're evaluating Acme Company's products for a hospital. You're given a text file containing purchase records from Acme. Looking through the text file, however, you see that purchases from other companies are included as well, with no mention of which ones come from which!

```
1  PCOD   QTY    DEPT      COST
   A169   100    Micro     0.58
3  PDA1   1      Xray      600.00
   X280   5      ER        199.99
5  ...
```

Luckily, you know that Acme's product codes consist of one uppercase letter followed by three digits.

Motivating regular expressions

So you get to work with a script:

```
import string
for line in open('purchaserecords.txt','r'):
    if line[0] in string.uppercase and \
        line[1].isdigit() and \
        line[2].isdigit() and \
        line[3].isdigit():
            print line
    else:
        continue
```

It's a bit clunky, but it works.

Motivating regular expressions

The next step in your evaluation is to collate evaluations emailed to you by hospital staff. These were free text, so you need to extract the product codes from within them to know which evaluation refers to which product.

```
1  ...The gloves(P180) felt sticky...
   ...The X701 vacuum cleaner really sucked!...
3  ...
```

You might be able to think of ways to program this, but really...

It's time to bust out regular expressions.