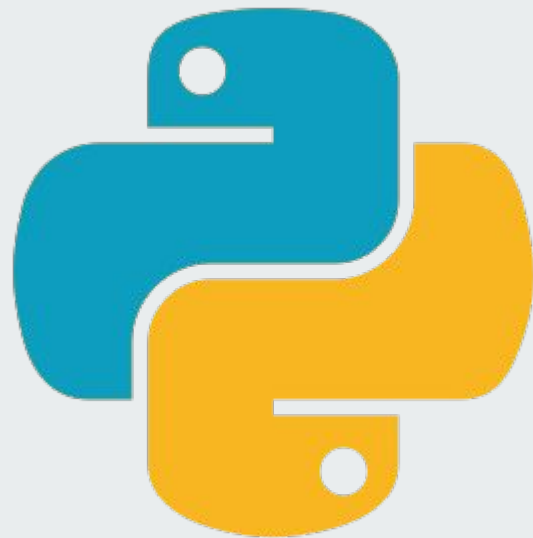




Live de Python #27

Raspando datas e fazendo downloads / Web Scraping #8





Roteiro

- Requests (stream)
- Calendar
- Dateutil - rrule
- CODE !!!



Stream dos requests

Por padrão, quando você faz um request, o body da resposta é baixado imediatamente. Você pode substituir esse comportamento e adiar o download do corpo de resposta até acessar o atributo `Response.content` com o parâmetro de fluxo.

```
tarball_url = 'https://github.com/requests/requests/tarball/master'  
r = requests.get(tarball_url, stream=True)
```



Stream dos requests

Neste ponto, apenas os cabeçalhos de resposta foram baixados e a conexão permanece aberta, permitindo-nos assim tornar a recuperação do conteúdo condicional

```
if int(r.headers['content-length']) < TOO_LONG:  
    content = r.content  
    ...
```

<http://docs.python-requests.org/en/master/user/advanced/#body-content-workflow>



Calendar

Este módulo permite que você publique calendários como o programa Unix `cal` e fornece funções úteis adicionais relacionadas ao calendário.

O módulo é basicamente reduzido em duas classes (existem mais) e algumas funções com utilidades.



Calendar - Calendar

```
In [1]: import calendar  
  
In [2]: cal = calendar.Calendar()  
  
In [3]: list(cal.iterweekdays())  
Out[3]: [0, 1, 2, 3, 4, 5, 6]
```



Calendar - Calendar

```
In [7]: cal.itermonthdays2(2017, 11) # Lista a relação do dia com o dia da semana
Out[7]: <generator object Calendar.itermonthdays2 at 0x7f2fcc1cc048>
```

```
In [8]: cal.monthdayscalendar(2017, 11)
```

```
Out[8]:
[[0, 0, 1, 2, 3, 4, 5],
 [6, 7, 8, 9, 10, 11, 12],
 [13, 14, 15, 16, 17, 18, 19],
 [20, 21, 22, 23, 24, 25, 26],
 [27, 28, 29, 30, 0, 0, 0]]
```



Calendar - TextCalendar

```
In [1]: import calendar

In [2]: tcal = calendar.TextCalendar()

In [3]: tcal.formatmonth(2017, 11).split('\n')
Out[3]:
['    November 2017',
 'Mo Tu We Th Fr Sa Su',
 '    1  2  3  4  5',
 ' 6  7  8  9 10 11 12',
 '13 14 15 16 17 18 19',
 '20 21 22 23 24 25 26',
 '27 28 29 30',
 '']
```




Calendar - TextCalendar

```
In [6]: import calendar
```

```
In [7]: tcal = calendar.LocaleTextCalendar(locale='pt_BR.utf8')
```

```
In [8]: tcal.formatmonth(2017, 11).split('\n')
```

```
Out[8]:
```

```
['  novembro 2017',  
 'se te qu qu se sá do',  
 '    1  2  3  4  5',  
 ' 6  7  8  9 10 11 12',  
 '13 14 15 16 17 18 19',  
 '20 21 22 23 24 25 26',  
 '27 28 29 30',  
 '']
```



Dateutil - rrule

```
In [1]: from datetime import datetime

In [2]: from dateutil.rrule import rrule, DAILY

In [3]: for x in rrule(DAILY, count=3, dtstart=datetime.now()):
...:     print(x)
...:
2017-11-06 20:20:51
2017-11-07 20:20:51
2017-11-08 20:20:51
```

—

CODE !!!!