

# Questionário acerca do Teste de Vaga Softplan

Sadi Júnior Domingos Jacinto

23 de Abril de 2018

## 1 Considerações Iniciais

Apresento aqui algumas considerações que considero relevantes acerca do teste:

1. O tempo médio de execução da minha solução foi de 23 segundos.
2. Arriscando-me a parecer idiota, mas acredito que esse tempo cairia bastante se meu poder computacional fosse maior.
3. No teste, era requerido que os resultados fossem exibidos de forma ordenada porém, no mesmo, não era estava explícito se o ordenamento deveria ser feito pelo título ou pelo número de visualizações.
4. Com relação ao item anterior, já irei adiantar que realizei o ordenamento baseado no número de visualizações.

## 2 Melhorias

A solução poderia ser melhorada com o uso da biblioteca *Selenium*, pois o *site Youtube* possui uma quantidade razoável de conteúdo gerado por *JavaScript*. Não posso afirmar nada a respeito do uso da biblioteca *lxml* pois não a conheço.

## 3 Teoria

### 3.1 *Site* com conteúdo gerado por *JavaScript*

Nesse tipo de situação, eu iria primeiramente fazer uma breve análise para saber qual a quantidade e complexidade do conteúdo do *site* gerado em *JavaScript* e, após chegar a uma conclusão, usaria uma das seguintes bibliotecas ou *frameworks*:

- *Selenium*:  
Para utilizar o *browser* e, assim, agilizar e melhorar o processo.
- *Scrapy*:  
Para escrever *spiders* que irão atrás do conteúdo requerido.
- *Splash*:  
Para utilizar a renderização que o mesmo realiza em páginas com *JavaScript*.

Também é possível integrar as soluções, porém, ainda não possuo conhecimento para tal feito.

### 3.2 *Site* sem conteúdo gerado por *JavaScript*

Nesse caso, eu usaria a biblioteca *BeautifulSoup* visando obtenção e análise do código HTML da página. Também poderia usar a biblioteca *lxml*, caso o *site* em questão possuía conteúdo em XML. E, provavelmente, uniria as duas soluções com a biblioteca *urllib*.