

第7课 存储我们爬取到的数据答案

题目要求：

- 根据已给网址爬取果壳前10页标题和对应的网址信息，并将爬取到的数据写入到excel表格或则是CSV中。

使用技术点引导

- requests
- BeautifulSoup
- for循环
- openpyxl

地址：<https://www.guokr.com/ask/highlight/?page=1>

结果图片片段截图：

标题	网址
印度人把男人的生殖器叫林伽，把女人的生殖器叫喻尼，林伽和喻尼的交合，便是瑜伽。这是真还是假的	http://www.guokr.com/question/669761/
子弹能击穿多少本书，与那一本书的页数有何关系？	http://www.guokr.com/question/668948/
狗是大科的，猫是猫科的，猪是猪科的，然而狼是大科的，虎是猫科的等等，这种科到底根据什么规则来定？谁先被发现就谁老大？相信一定有大神可以为我解答科怎么出现的。	http://www.guokr.com/question/668829/
为什么往酒店房间里的消防喷头挂东西，会造成紧急消防系统启动？	http://www.guokr.com/question/667552/
为什么猪牛羊的瘦肉是红色的，鸡肉是粉色的，而鱼肉是白色的？鱼的肌肉里面不充血吗？为什么呢？	http://www.guokr.com/question/667536/
“三季人”是古文吗？	http://www.guokr.com/question/667020/
九江市传统文化学会说“克夫”的“克”其实是“影响”的意思，这是真的吗？	http://www.guokr.com/question/666893/
为何充电电池不是1.5V是1.2V，而干电池是1.5V？两者不能调到一致的水平么？有说法是因为充电电池需要充电所以，电压要比干电池下降20%，否则充不进去，我觉得似乎不太合理吧。。。	http://www.guokr.com/question/666252/
受精卵发育过程中，被吞掉的兄弟，变成了嵌合体的宝宝，这有可能吗？	http://www.guokr.com/question/665737/

```
1 import requests
2 from bs4 import BeautifulSoup
3 import openpyxl
4
5 wb = openpyxl.Workbook()
6 sheet = wb.active
7 sheet.title = '开课吧'
8 sheet['A1'] = '标题'
9 sheet['B1'] = '网址'
10 for page in range(1,11):
11     print('已下是第{}页数据: {}'.format(page))
12     res = requests.get('https://www.guokr.com/ask/highlight/?page={}'.format(page))
13     html = BeautifulSoup(res.text,'html')
14     data_ = html.find_all('ul',class_='ask-list-cp')
15     a=data_[0].find_all('div',class_='ask-list-detials')
16     for i in a:
17         res_data= i.find('a')
18         tit = res_data.text
19         print('标题: ',tit)
20         url = res_data['href']
21         print('网址: ',url)
22         sheet.append([tit,url])
23     print()
24 wb.save('kaikeba.xlsx')
25 wb.close()
```

