import requests

from bs4 import BeautifulSoup

url = 'http://books.toscrape.com/catalogue/unqualified-how-god-uses-broken-people-to-do-big-things\_873/index.html'

response = requests.get(url)

infos = []

header = []

star = 0

links = []

if response.ok:

soup = BeautifulSoup(response.text,'lxml')

trs = soup.findAll('tr')

title = soup.find('div', {'class':'col-sm-6 product\_main'}).find('h1')

for tr in trs:

td = tr.find('td')

infos.append(td.text)

universal\_product\_code = infos[0]

price\_excluding\_tax = infos[2]

price\_including\_tax = infos[3]

number\_available = infos[5]

descripts = soup.find('article', {'class':'product\_page'}).findAll('p')

product\_description = descripts[3]

lis=soup.find('ul', {'class':'breadcrumb'}).findAll('li')

for li in lis:

a = li.find('a')

header.append(a)

category = header[2]

#product\_page\_url = soup.find('tr',{'class':'product\_page'})

stars = soup.find('div',{'col-sm-6 product\_main'}).findAll('p') #review\_rating

p = stars[2]

star = p['class']

image\_url = soup.find('div', {'class':'item active'}).find('img')

print('product\_page\_url: ' + url + '\n')

print('universal\_product\_code: ' + universal\_product\_code + '\n')

print('Title: ' + title.text + '\n')

print('price\_including\_tax: ' + price\_including\_tax + '\n')

print('price\_excluding\_tax: ' + price\_excluding\_tax + '\n')

print('number\_available: ' + number\_available + '\n')

print('product\_description: ' + product\_description.text + '\n')

print('category: ' + category.text + '\n')

print('review\_rating: ' + star[1] + '\n')

print('image\_url: ' + 'http://books.toscrape.com/' + image\_url['src'] + '\n')