

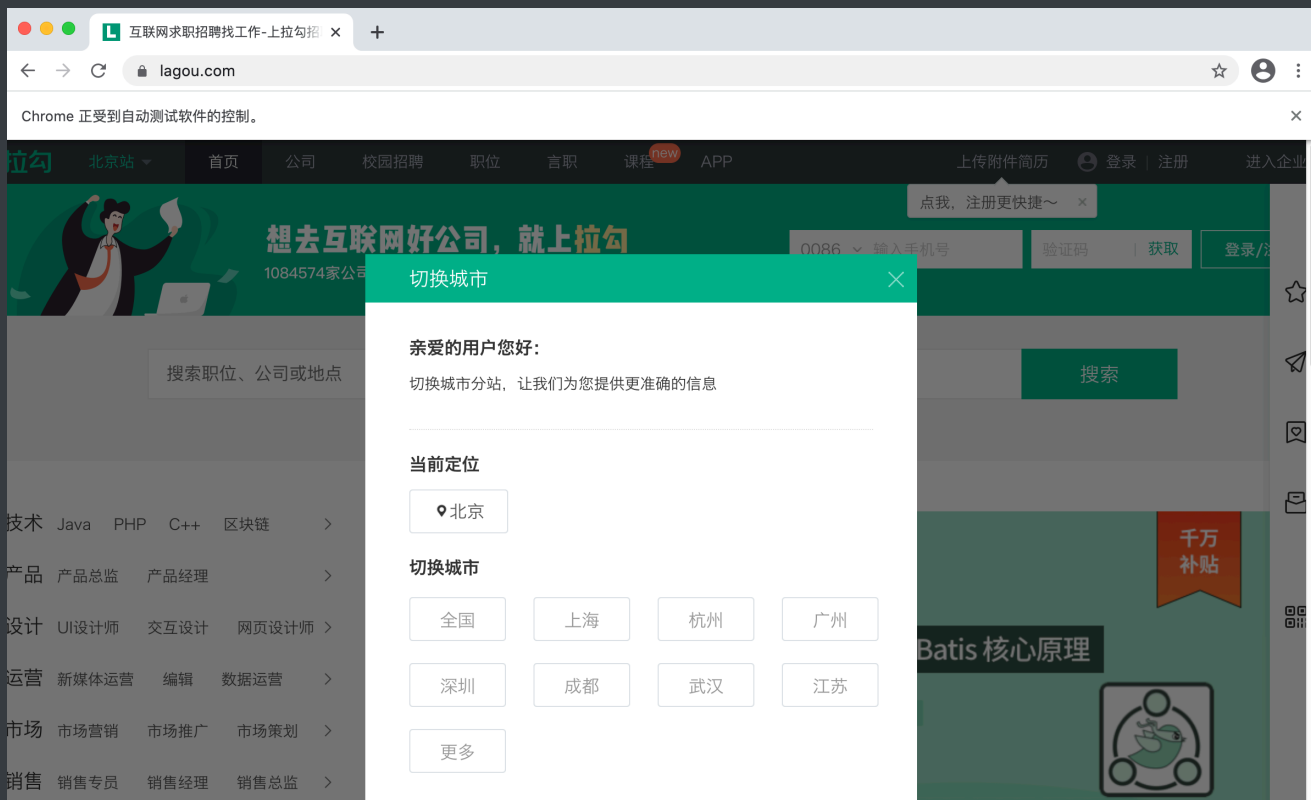
selenium各种神奇操作

selenium不但可以打开浏览器. 还可以对浏览器各种操作. 比如, 点击, 查找. 都可以.

我们直接上案例. 抓取拉钩网招聘python工程师的招聘信息

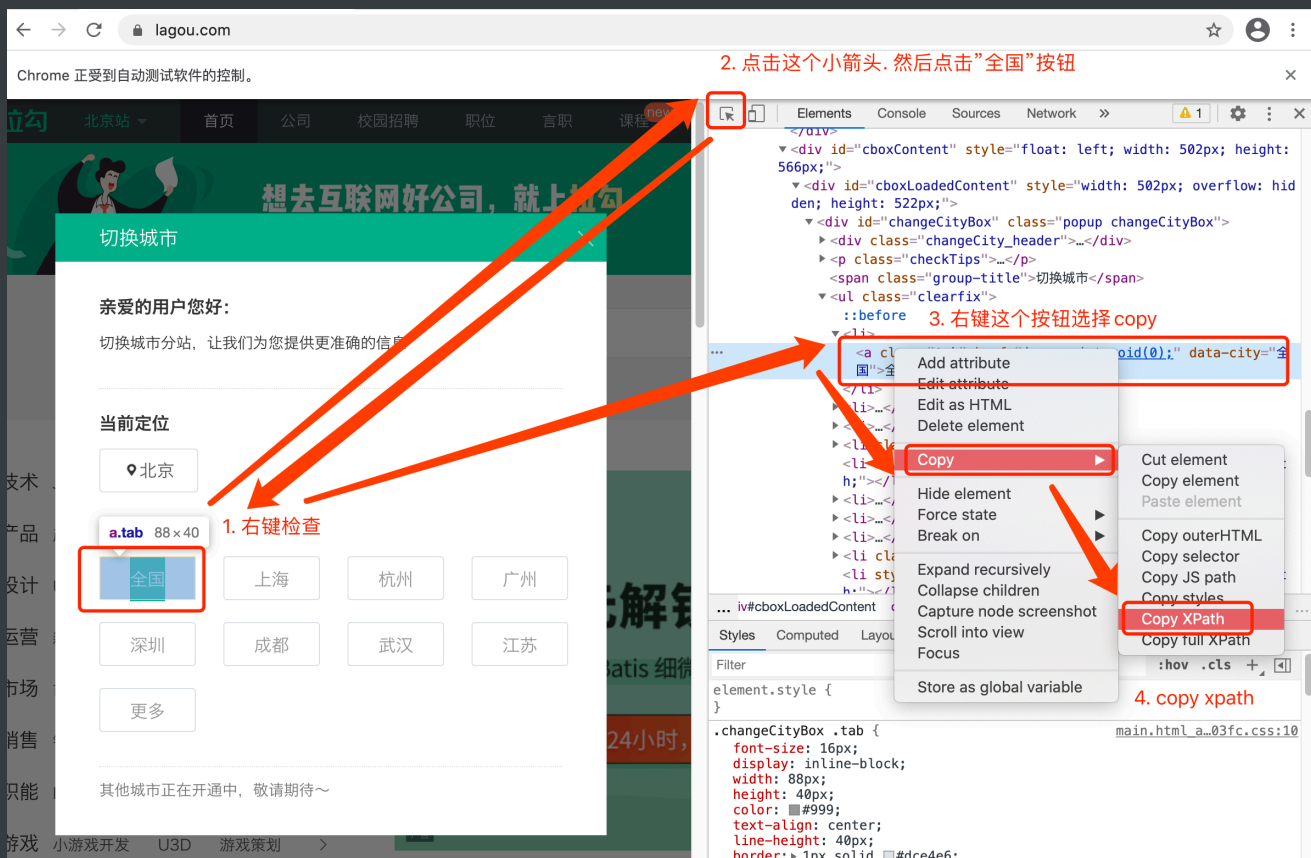
1. 准备工作

```
1 from selenium.webdriver import Chrome
2
3 web = Chrome()
4
5 web.get("http://lagou.com")
6
```



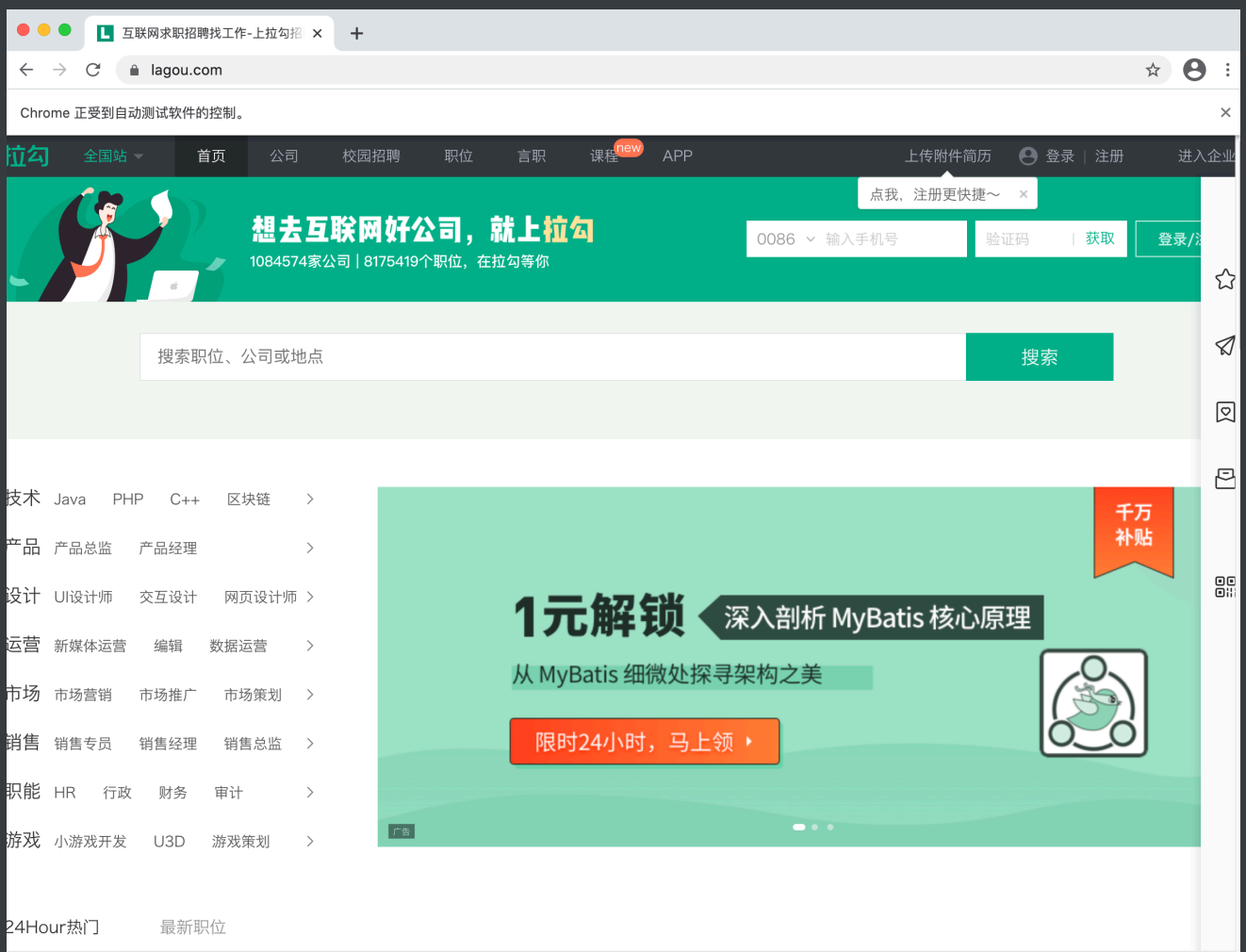
2. 点击-全国按钮

想要点击这个按钮. 我们需要先定位到这个按钮. 然后再点击
selenium想要定位某个元素. 太简单了.



有了xpath是不是明白了些什么.

```
1 from selenium.webdriver import Chrome
2
3 web = Chrome()
4
5 web.get("http://lagou.com")
6 btn = web.find_element_by_xpath('//*[@id="changeCityBox"]/ul/li[1]/a') # 找到全国按钮
7 btn.click() # 点击这个按钮
8
```



舒服!

3. 搜索python

人的过程: 找到文本框输入"python", 点击"搜索"按钮.

机器的过程: 找到文本框输入"python", 点击"搜索"按钮.

发现没, 用selenium最爽的地方就是这里. 人是怎么操作的. 机器就怎么操作. 爽到极点

```
1 # 找到文本框输入python, 点击搜索按钮
2 web.find_element_by_xpath('//*[@
  @id="search_input"]').send_keys("python")
3 web.find_element_by_xpath('//*[@
  @id="search_button"]').click()
```

send_keys() 这里要说一下. 如果我们给出的是一个字符串. 就是输入文本. 但是, 如果给出的是一个键盘指令, 那就按下键盘. 比如, 我想要按回车按钮. 就是这样的

```
1 from selenium.webdriver import Chrome
2 from selenium.webdriver.common.keys import
  Keys
3 import time
4 web = Chrome()
5
```

```
6 web.get("http://lagou.com")
7 btn = web.find_element_by_xpath('//*
    [@id="changeCityBox"]/ul/li[1]/a') # 找到全国
    按钮
8 btn.click() # 点击这个按钮
9
10 time.sleep(2) # 让浏览器反应一会儿
11
12 # 找到文本框输入python, 点击搜索按钮
13 web.find_element_by_xpath('//*
    [@id="search_input"]').send_keys("python",
    Keys.ENTER)
14 # web.find_element_by_xpath('//*
    [@id="search_button"]').click()
15
```

```
24
25 class Keys(object):
26     """
27     Set of special keys codes.
28     """
29
30     NULL = '\ue000'
31     CANCEL = '\ue001' # ^break
32     HELP = '\ue002'
33     BACKSPACE = '\ue003'
34     BACK_SPACE = BACKSPACE
35     TAB = '\ue004'
36     CLEAR = '\ue005'
37     RETURN = '\ue006'
38     ENTER = '\ue007'
39     SHIFT = '\ue008'
40     LEFT_SHIFT = SHIFT
41     CONTROL = '\ue009'
42     LEFT_CONTROL = CONTROL
43     ALT = '\ue00a'
44     LEFT_ALT = ALT
45     PAUSE = '\ue00b'
46     ESCAPE = '\ue00c'
47     SPACE = '\ue00d'
48     PAGE_UP = '\ue00e'
49     PAGE_DOWN = '\ue00f'
50     END = '\ue010'
51     HOME = '\ue011'
52     LEFT = '\ue012'
53     ARROW_LEFT = LEFT
```

keys里几乎包含了我们需要的所有特殊按键

4. 提取招聘信息

```
1 from selenium.webdriver import Chrome
2 from selenium.webdriver.common.keys import
  Keys
3 import time
4 web = Chrome()
5
6 web.get("http://lagou.com")
7 btn = web.find_element_by_xpath('//*[@
  id="changeCityBox"]/ul/li[1]/a') # 找到全国
  按钮
8 btn.click() # 点击这个按钮
9
10 time.sleep(2) # 让浏览器反应一会儿
11
12 # 找到文本框输入python, 点击搜索按钮
13 web.find_element_by_xpath('//*[@
  id="search_input"]').send_keys("python",
  Keys.ENTER)
14 time.sleep(1)
15 # web.find_element_by_xpath('//*[@
  id="search_button"]').click()
16
17
```

```
18 ls = web.find_elements_by_xpath('//*  
    [@id="s_position_list"]/ul/li') # 一次性提取多  
    个元素用elements  
19 for item in ls:  
20     name =  
        item.find_element_by_xpath('./div[1]/div[1]/d  
        iv[1]/a/h3').text  
21     addr =  
        item.find_element_by_xpath('./div[1]/div[2]/d  
        iv[1]/a').text  
22     print(name, addr)  
23     # 其他内容你自己琢磨吧  
24
```