

```

1  def getdatacsv(kwlist, geo='HK'):
2      from selenium.webdriver.chrome.options import Options
3      from selenium.common.exceptions import TimeoutException
4      from selenium.webdriver.support.ui import WebDriverWait
5      from selenium.webdriver.support import expected_conditions as EC
6      from selenium.webdriver.common.by import By
7      import time
8
9      #set download automatically
10     options = Options()
11     options.add_experimental_option("prefs", {
12         "download.default_directory": r'C:\Users\yeung\Downloads\g6p1',
13         "download.prompt_for_download": True,
14         "download.directory_upgrade": True,
15         "safebrowsing.enabled": True
16     })
17
18     from selenium import webdriver
19     driver = webdriver.Chrome(r'C:\Users\yeung\Google Drive\xecelerate\small
programme\chromedriver')
20     input('after change download directory press enter')
21
22     #check if kwlist have chinese char.
23     import re
24     ischi = lambda x: not bool(re.match(r'^[a-zA-Z\-\ ']+', x))
25
26     #translate kwlist into utf-8 encoding
27     trslated = []
28     for i in kwlist:
29         if ischi(i):
30             trslated.append('%'+%.join(re.findall(r'(?x)(\w\w)',
31                 str(i.encode('utf-8')))))
32         else:
33             trslated.append(i)
34
35     #generate query keywords
36     indexkeyword = '%e5%8f%a3%e7%bd%a9'
37     linkkeyword = []
38     if len(trslated) > 3:
39         while len(trslated) >= 3:
40             linkkeyword.append(indexkeyword+', '+'.join(trslated[:3]))
41             trslated = trslated[3:]
42             if len(trslated)<3 and len(trslated)>0:
43                 linkkeyword.append(indexkeyword+', '+'.join(trslated))
44     else:
45         linkkeyword.append(indexkeyword+', '+'.join(trslated))
46
47     if geo=='HK':
48         for i in linkkeyword:
49             urlhk =
50                 'https://trends.google.com/trends/explore?date=2020-01-01%202020-06-30&geo=HK
&q='
51             requesturl = urlhk+i
52             driver.get(requesturl)
53             time.sleep(8)
54             #click the download buttom
55
56             xpath =
57                 '/html/body/div[2]/div[2]/div/md-content/div/div/div[1]/trends-widget/ng-incl
58                 ude/widget/div/div/div/widget-actions/div/button[1]/i'
59             try:
60                 element_present = EC.presence_of_element_located((By.XPATH, xpath))
61                 WebDriverWait(driver, 600).until(element_present)
62             except TimeoutException:
63                 print("Timed out waiting for page to load")
64             download = driver.find_element_by_xpath(xpath)

```

```
62         download.click()
63         time.sleep(5)
64
65     kwlist=['a','b','c','出入境', '肺炎', '消毒酒精', '洗手液', 'mask', 'facemask', '封關',
66            '紙巾', '世界末日', 'end of the world', '肺炎徵狀', 'covid syndrome', '廁紙', 'toilet
        paper', 'air-purifier', '空氣淨化器', '出境', '入境', '去旅行', '健康', '行山', '旅行',
        'travelling', '移民', 'china', '中國', '香港']
66     getdatacsv(kwlist, geo='HK')
```