

Web Scrapping

Extracting contacts from LinkedIn

Import Libraries

```
In [3]: from bs4 import BeautifulSoup
import requests
import lxml
import selenium
from selenium import webdriver
from webdriver_manager.chrome import ChromeDriverManager
from time import sleep
from selenium.webdriver.common.by import By
import numpy as np
import pandas as pd
import time
```

crate a function to extract contacts from LinkedIn

```
In [9]: def LinkedIn_contacts(email,password,number_of_contacts):
    #creat 3 empty Lists:
    first_name=[]
    last_name=[]
    phone_number=[]
    # defining driver and get the LinkedIn homepage:
    driver = webdriver.Chrome(ChromeDriverManager().install())
    driver.get("https://www.linkedin.com/home")
    sleep(2)
    #sending Email:
    username=driver.find_element(by="id",value="session_key").send_keys(str(email))
    sleep(1)
```

```

#sending Password:
password=driver.find_element(by="id",value="session_password").send_keys(str(password))
sleep(5)
# Singing In:
sign=driver.find_element(by="xpath",
                        value="/html/body/main/section[1]/div/div/form[1]/div[2]/button").click()

sleep(30)
# Clicking on network button:
network=driver.find_element(by="xpath" ,
                           value='/html/body/div[5]/header/div/nav/ul/li[2]/a')

network.click()
sleep(5)
# Showing the list of contacts by clicking on contact button:
contact_list=driver.find_element(By.XPATH,
                                "/html/body/div[5]/div[3]/div/div/div/div/div[2]/div/div/div/div/div/div/div/section[1]/div/div[2]/a")
contact_list.click()
sleep(5)
for i in np.arange(1,number_of_contacts+1):
    try:

        start_time=time.time()
        print(i)
        sleep(3)
        # get contact infos
        contact=driver.find_element(By.XPATH,
                                    f'/html/body/div[5]/div[3]/div/div/div/div/div/div/div/div[2]/div/div/main/div/div/div/div[1]/ul/li[{i}]')
        sleep(2)
        contact.click()
        sleep(4)
        # get contact name
        name=driver.find_element(By.XPATH,
                                "/html/body/div[3]/div/div/div[2]/div/div[2]/div[1]/p")
        first_name.append(name.text)
        sleep(1)
        # get contact phone number and Last name
        phone_or_lastname=driver.find_element(By.XPATH,
                                              "/html/body/div[3]/div/div/div[2]/div/div[2]/div[2]/p").text
        int_phone=phone_or_lastname.replace("-", "").replace("+", "").replace(" ", "")
        if int_phone.isdigit():
            phone=phone_or_lastname

```

```

        phone_number.append(phone)
        last_name.append(np.nan)

    else :
        lastname=phone_or_lastname
        last_name.append(lastname)
        phone=driver.find_element(By.XPATH,
                                   "/html/body/div[3]/div/div/div[2]/div/div[2]/div[3]/p").text
        phone_number.append(phone)
    sleep(2)
    driver.find_element(By.XPATH, "/html/body/div[3]/div/div/button").click()
    sleep(2)
    end_time=time.time()
    elapsed_time = end_time - start_time

    print(f"Elapsed time: {elapsed_time:.2f} seconds")
except :
    print("can't find this element!")
    continue
# stored data in Dataframe than in a csv file:
contact_df=pd.DataFrame({"First_Name":first_name,
                          "Last_Name":last_name,
                          "Phone_Number":phone_number})
return contact_df.to_csv("linkdIn_contacts.csv")

```

In [8]: LinkedIn_contacts("your_Email", "Your_Password", number_of_your_contacts)