Spokeo Web Scraper

August 10, 2024

```python
[56]: import pandas as pd
      from selenium.webdriver.common.by import By
      from selenium.webdriver.support.ui import WebDriverWait
      from selenium.webdriver.support import expected_conditions as EC
      from selenium.webdriver.common.keys import Keys
      from selenium.common.exceptions import NoSuchElementException
```

```python
[72]: df = pd.read_excel('C:/Users/training.lhr/Desktop/VB Data-2023/VB Data Modified/
      ↪Zain Bhai Files/Phase 2 Files/Select States Files Part 1/Files To Upload/
      ↪CA_OH File 2.xlsx', sheet_name=0)

      phone_numbers = df["Phone Number"].tolist()
      area_codes = df["Area Code"].tolist()
      states = df["State"].tolist()

      data = []
      for i in range(len(phone_numbers)):
          row = {
              "phone_number": phone_numbers[i],
              "area_code": area_codes[i],
              "state": states[i]
          }
          data.append(row)
```

```python
[74]: spokeo_url = 'https://www.spokeo.com/'
      driver.get(spokeo_url)
```

```python
[81]: df = pd.read_excel('C:/Users/training.lhr/Desktop/VB Data-2023/VB Data Modified/
      ↪Zain Bhai Files/Phase 2 Files/Select States Files Part 1/Files To Upload/
      ↪CA_OH File 2.xlsx', sheet_name=0)

      phone_numbers = df["Phone Number"].tolist()
      area_codes = df["Area Code"].tolist()
      states = df["State"].tolist()

      data = []
      for i in range(len(phone_numbers)):
```

```python
    row = {
        "phone_number": phone_numbers[i],
        "area_code": area_codes[i],
        "state": states[i]
    }
    data.append(row)

login_button = WebDriverWait(driver, 10).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="header__"]/header/div/div/
 ↪div/div/div[3]/div[1]/div/a[2]')))
login_button.click()

email_input = WebDriverWait(driver, 10).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="email_address"]')))
email_input.send_keys("sybridthatsthem55@gmail.com")

password_input = WebDriverWait(driver, 10).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="password"]')))
password_input.send_keys("Welcome@123")

for phone_number in phone_numbers:

    search_bar_prev = WebDriverWait(driver, 10).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="homepage_hero_form"]/div/
 ↪div[1]/input')))
    search_bar_prev.send_keys(Keys.CONTROL + "a")
    search_bar_prev.send_keys(Keys.DELETE)
    search_bar_prev.send_keys(phone_number)

    search_button_prev = WebDriverWait(driver, 20).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="search"]')))
    driver.execute_script("arguments[0].click();", search_button_prev)

    wait = WebDriverWait(driver, 20)
    wait.until(EC.presence_of_element_located((By.XPATH,'//*[@id="header__"]/
 ↪header/div/div/div/div/div[2]/div/div/form/div/div/div/div/input')))

    try:

        see_results_button = WebDriverWait(driver, 10).until(EC.
 ↪presence_of_element_located((By.XPATH, '//*[@id="owner-0"]/div[4]/a')))
        driver.execute_script("arguments[0].click();", see_results_button)

        try:
            customer_name = driver.find_element(By.XPATH, '//
 ↪*[@id="possible-owner-0"]/a/div/span/div[1]/div[1]/ul/li/div').text
```

```python
            customer_address = driver.find_element(By.XPATH, '//
 ↪*[@id="possible-owner-0"]/a/div/span/div[1]/div[2]/span').text

        except NoSuchElementException:

            customer_name = ""
            customer_address = ""

    except NoSuchElementException:

        customer_name = ""
        customer_address = ""


    df.loc[df["Phone Number"] == phone_number, "Customer Name"] = customer_name
    df.loc[df["Phone Number"] == phone_number, "Customer Address"] =␣
 ↪customer_address

    df.to_excel("CA_OH File 2.xlsx", index=False)
    driver.get(spokeo_url)


df.to_excel("CA_OH File 2.xlsx", index=False)
```