**Scraping and Data Analysis with Python**

**Code:**

import selenium

import pandas as pd

from selenium import webdriver

from selenium.webdriver.common.by import By

from selenium.webdriver.chrome.options import Options

import pandas as pd

# Creating Chrome Webdriver Instance

chrome\_options = Options()

chrome\_options.add\_argument("--headless")

# Initialising WebDriver with Chrome options

driver = webdriver.Chrome(options=chrome\_options)

# Navigating to the URL

driver.get("http://www.sunsirs.com/futures-price-2023-0927-daily.html")

table = driver.find\_element(By.XPATH, '//span[@id="date"]//parent::div//following-sibling::div/table/tbody')

# Getting all the rows in table\_rows

table\_rows = table.find\_elements(By.TAG\_NAME, "tr")

# Extract the data from the table

table\_data = []

for indexrow in range(len(table\_rows)): #Iterating for rows

    table\_row = []

    table\_col = driver.find\_elements(By.XPATH,f"//tbody/tr[{indexrow+1}]/td") #getting all column elements

    for index in range(len(table\_col)): #iterating for column

        element=driver.find\_element(By.XPATH,f"//tbody/tr[{indexrow+1}]/td[{index+1}]") #getting each data from current row

        textdata = element.get\_attribute("textContent") #invoking text from element

        textdata=textdata.strip()   #removing extra space and new line

        table\_row.append(textdata) #storing all column data in a single array of the current row

    table\_data.append(table\_row)

df = pd.DataFrame(table\_data[1:], columns=table\_data[0])

#storing data in excel file

df.to\_excel('RawData.xlsx',index=False)

# Counting total number of rows in the DataFrame

total\_rows = df.shape[0]

print(f"total number of data rows is::{total\_rows}")

# Converting the "09-28" column to a numeric data type

df["09-28"] = df["09-28"].str.replace(",", "").astype(float)

# Finding row with the highest closing price in the "09-28" column

max\_closing\_row = df[df["09-28"] == df["09-28"].max()]

# Extracting commodity name from the highest closed price row

commodity\_with\_highest\_price = max\_closing\_row.iloc[0]["Commodity"]

#Handling if name is not there

if(commodity\_with\_highest\_price):

    print(f"Commidity name is:{commodity\_with\_highest\_price}")

else:

    print('Commodity name not found')

# Extracting highest closing price

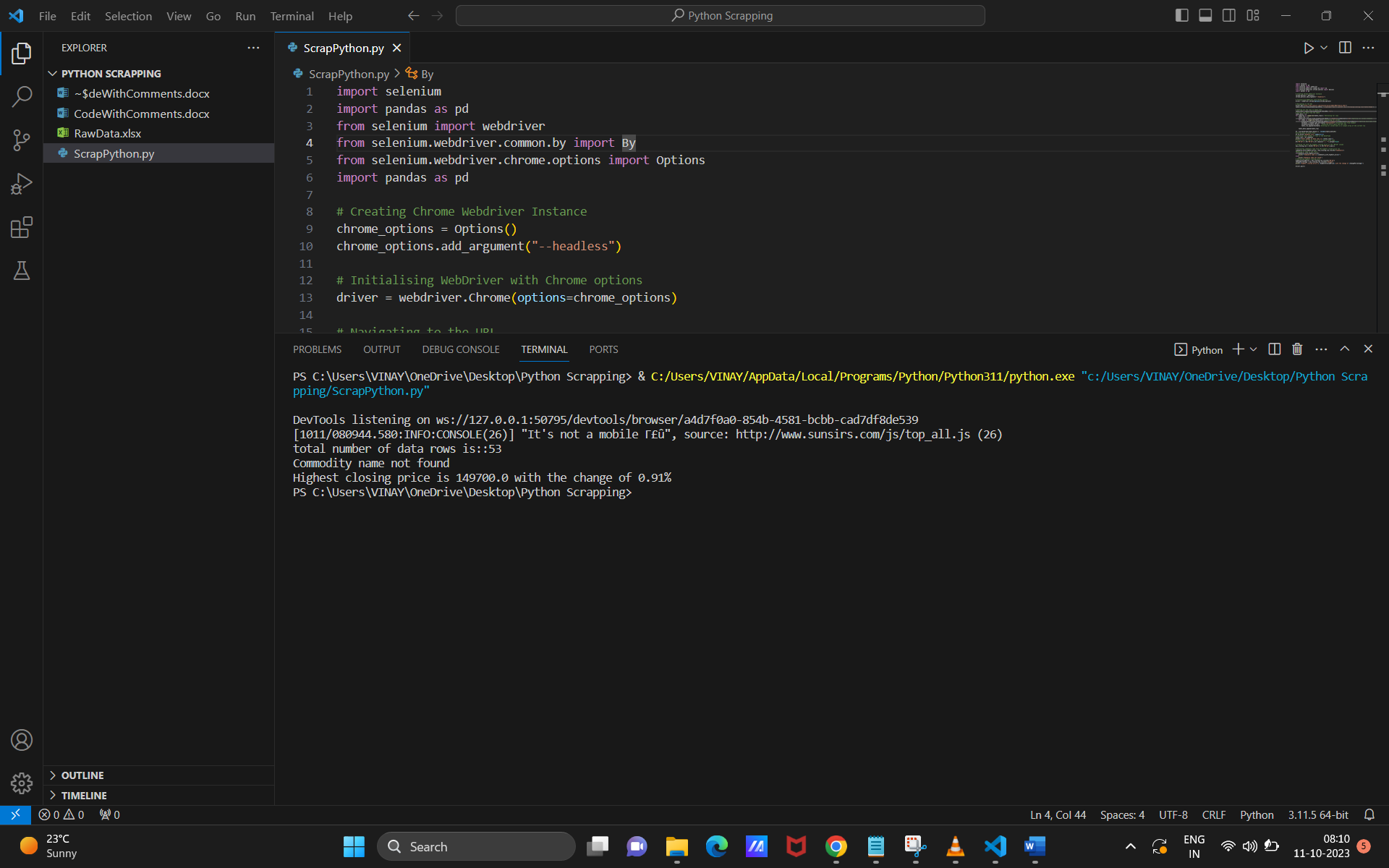
highestClosingPrice = max\_closing\_row.iloc[0]["09-28"]

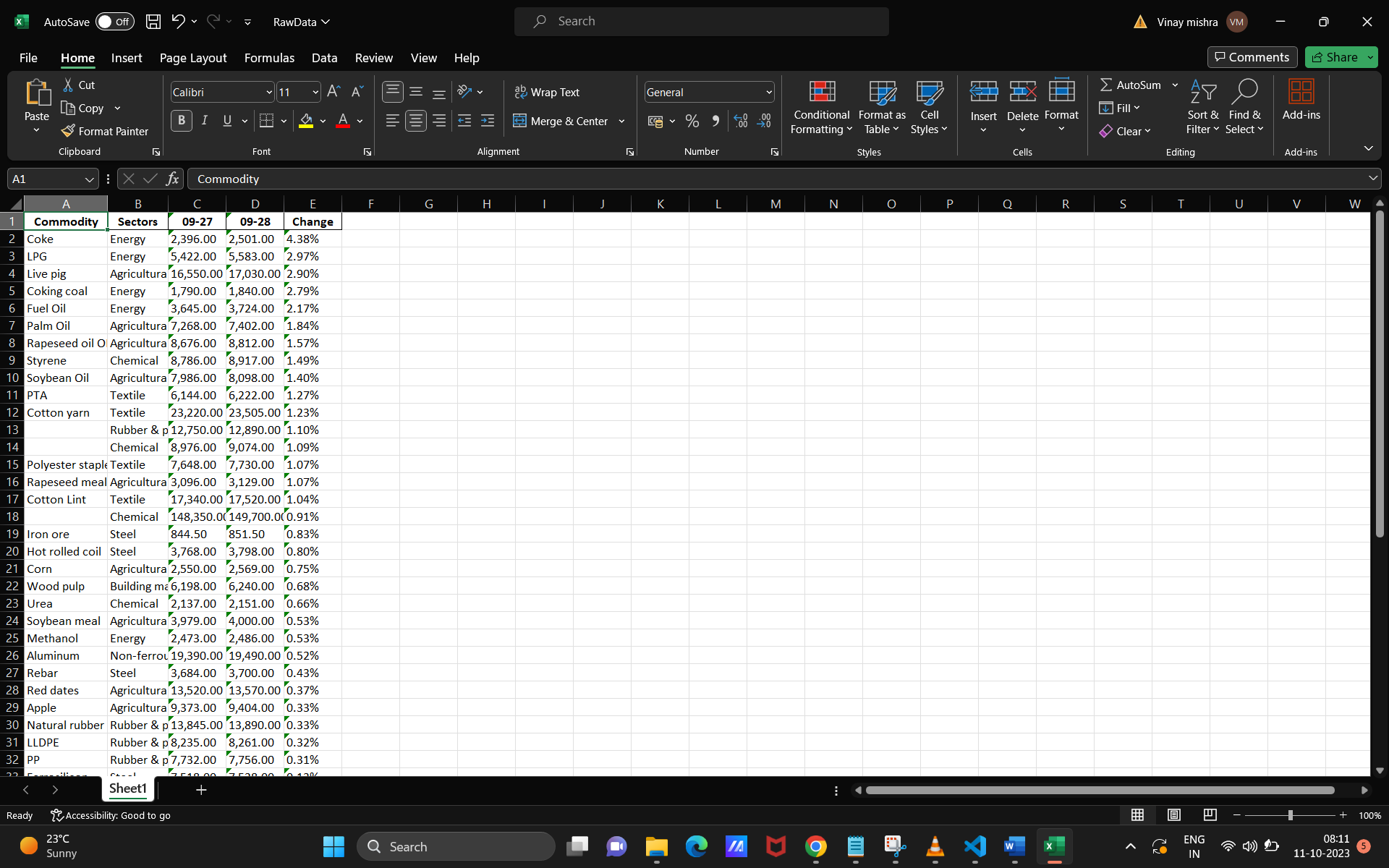
ChangePercentage = max\_closing\_row.iloc[0]["Change"]

print(f"Highest closing price is {highestClosingPrice} with the change of {ChangePercentage}")

driver.quit()

**Outputs:**

****

****