import regex as re

import requests

from bs4 import BeautifulSoup

from selenium.webdriver.common.by import By

from selenium import webdriver

from selenium.webdriver.chrome.service import Service

from webdriver\_manager.chrome import ChromeDriverManager

from selenium.webdriver.common.keys import Keys as KeysBrowser

import pandas as pd

import time

import os

from lxml import etree

**PARTIE I**

driver = webdriver.Chrome(service=Service(ChromeDriverManager().install()))

driver.get("https://iccwbo.org/dispute-resolution-services/arbitration/icc-case-information/")

soup=BeautifulSoup(driver.page\_source)

ICC=soup.find("div", class\_="content\_\_pad arbitration\_\_pad")

nom\_arbitres =[]

for child in ICC.findAll("li"):

nom\_arbitres = nom\_arbitres + [child.text]

print(nom\_arbitres)

set\_tribunaux = soup.find\_all("a", text=re.compile("Case ID"))

Case\_ID = []

Case\_Statut = []

Date = []

Sector\_Industry = []

for compo in set\_tribunaux:

arbitration = re.search("case-id-\d{5}", compo.get("href")).group()

driver.get("https://iccwbo.org/dispute-resolution-services/arbitration/icc-case-information/" + arbitration)

soup2=BeautifulSoup(driver.page\_source)

test=soup2.find("div", class\_ ="page\_\_intro-content")

for child in test.findAll("h1"):

Case\_ID = Case\_ID + [child.next.text]

for child in test.findAll("h1"):

Case\_Statut = Case\_Statut + [child.next.next.next.next.next.next.next.next.next.text]

for child in test.findAll("h1"):

Date = Date + [child.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.text]

for child in test.findAll("h1"):

Sector\_Industry = Sector\_Industry + [child.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.text]

print(Case\_ID)

print(Case\_Statut)

print(Date)

print(Sector\_Industry)

**PARTIE II (PAS UTILE POUR LE MOMENT)**

brute = [nom\_arbitres, Date, Case\_ID, Case\_Statut, Sector\_Industry]

my\_df=pd.DataFrame(brute)

my\_df.T

**A REPRENDRE**

set\_tribunaux = soup.find\_all("a", text=re.compile("Case ID"))

Case\_ID = []

Case\_Statut = []

Date = []

Sector\_Industry = []

Fin = False

while Fin == False :

for compo in set\_tribunaux:

arbitration = re.search("case-id-\d{5}", compo.get("href")).group()

driver.get("https://iccwbo.org/dispute-resolution-services/arbitration/icc-case-information/" + arbitration)

soup2=BeautifulSoup(driver.page\_source)

test=soup2.find("div", class\_ ="page\_\_intro-content")

Page\_Suivante = "https://iccwbo.org/dispute-resolution-services/arbitration/icc-case-information/?fwp\_paged=2"

if Page\_Suivante != "":

driver.get(Page\_Suivante)

time.sleep(1)

else:

Fin = True

for child in test.findAll("h1"):

Case\_ID = Case\_ID + [child.next.text]

for child in test.findAll("h1"):

Case\_Statut = Case\_Statut + [child.next.next.next.next.next.next.next.next.next.text]

for child in test.findAll("h1"):

Date = Date + [child.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.text]

for child in test.findAll("h1"):

Sector\_Industry = Sector\_Industry + [child.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.next.text]

print(Case\_ID)

print(Case\_Statut)

print(Date)

print(Sector\_Industry)