410921208 楊右宇

from selenium import webdriver

from time import sleep

from selenium.webdriver.common.by import By

from selenium.webdriver.chrome.service import Service

import matplotlib.pyplot as plt

import matplotlib.image as img

import requests

import pandas as pd

import numpy as np

import requests

from bs4 import BeautifulSoup

import re

url = 'https://www.cdc.gov.tw/Bulletin/List/MmgtpeidAR5Ooai4-fgHzQ'

html = requests.get(url)

sp = BeautifulSoup(html.text, 'html.parser')

news\_all = sp.find\_all(class\_="content-boxes-v3")

news = []

for i in news\_all:

if "COVID-19確定病例" in i.text:

news += [i]

# print(news)

Date = []

Case\_Increase = []

Local = []

Overseas = []

herfs = []

for i in range(3):

data = re.findall(r'[\d+,]+', news[i].text)

herf = news[i].find('a')

h = 'https://www.cdc.gov.tw' + herf.attrs['href']

# print(herf)

herfs += [h]

# print(data)

Date += ['/'.join(data[0:3])]

Case\_Increase += [data[3].replace(',', '')]

Local += [data[5].replace(',', '')]

Overseas += [data[6].replace(',', '')]

# print(Date)

# print(Local)

# print(Case\_Increase)

# print(Overseas)

# print(herfs)

df = pd.DataFrame({'Date':Date,

'Case Increase':Case\_Increase,

'Local':Local,

'Overseas':Overseas,

'Press Link':herfs})

# df.index = df['Date']

from datetime import datetime

from datetime import date

now = datetime.now()

current\_time = now.strftime("%H:%M")

today = date.today()

# dd/mm/YY

d1 = today.strftime("%Y/%m/%d")

df.loc[3] = ['資料取得時間：', d1, current\_time, "", ""]

# df.loc[3] = '資料取得時間：'

df