

Assignment 4

Web Crawling and Extracting Information-Part1

Running Instructions:

```
python3 task1.py
```

```
python3 task2.py
```

Task 1

Firstly, I read the main URL & saves the page along with all the pages of countries present in the given file in HTML format. I have saved the pages into the "COUNTRIES" folder in html format.

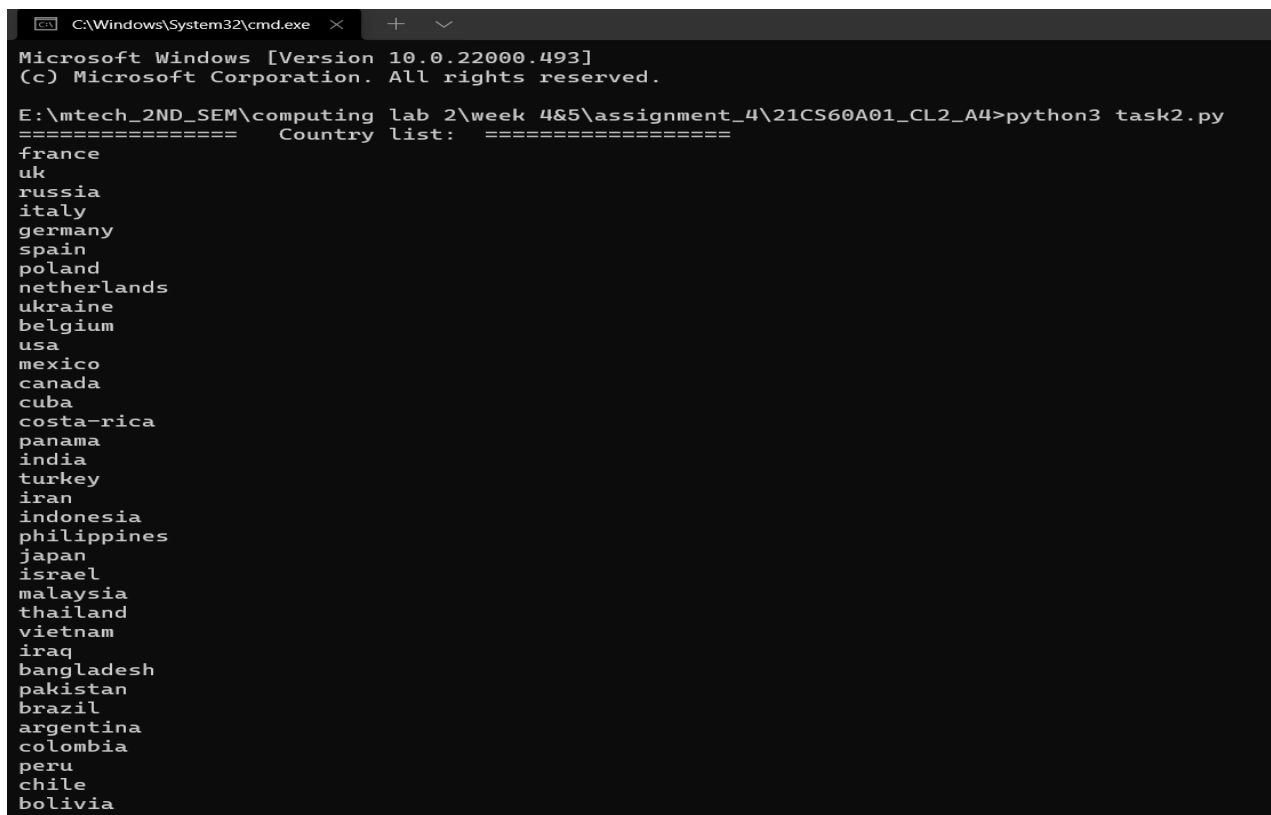
Required modules of python:

```
requests (pip install requests)
```

Task 2

In task 2, I have extracted only 4 fields which are available in the html file of countries. They are Total Cases, Active Cases, Total Recovered and Total deaths. I can't extract other fields as they are not present in the html files of the individual countries.

Screenshots of the program:



```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.22000.493]
(c) Microsoft Corporation. All rights reserved.

E:\mtech_2ND_SEM\computing lab 2\week 4&5\assignment_4\21CS60A01_CL2_A4>python3 task2.py
===== Country list: =====
france
uk
russia
italy
germany
spain
poland
netherlands
ukraine
belgium
usa
mexico
canada
cuba
costa-rica
panama
india
turkey
iran
indonesia
philippines
japan
israel
malaysia
thailand
vietnam
iraq
bangladesh
pakistan
brazil
argentina
colombia
peru
chile
bolivia
```

```
C:\Windows\System32\cmd.exe
fiji
papua-new-guinea
new-caledonia
new-zealand
Enter the country(country name in lower case): india
Generating LALR tables
Enter the field no.(1-8)
1.Total cases
2.Active cases
3.Total deaths
4.Total recovered
5.Total tests
6. New case
7.New death
8.New recovered
1
Total cases in india = 42822473

<india><Total_cases><42822473>

Want to explore other datas on same country (y/n)> y
Enter the field no.(1-8)
1.Total cases
2.Active cases
3.Total deaths
4.Total recovered
5.Total tests
6. New case
7.New death
8.New recovered
3
Total deaths in india = 511935

<india><Total_deaths><511935>

Want to explore other datas on same country (y/n)> y
Enter the field no.(1-8)
1.Total cases
2.Active cases
3.Total deaths
```

```
C:\Windows\System32\cmd.exe
Total deaths in india = 511935

<india><Total_deaths><511935>

Want to explore other datas on same country (y/n)> y
Enter the field no.(1-8)
1.Total cases
2.Active cases
3.Total deaths
4.Total recovered
5.Total tests
6. New case
7.New death
8.New recovered
4
Total recovered in india = 42086383

<india><Total_recovered><42086383>

Want to explore other datas on same country (y/n)> n
Do you want to explore other countries (y/n)? y
Enter the country(country name in lower case): bangladesh
Enter the field no.(1-8)
1.Total cases
2.Active cases
3.Total deaths
4.Total recovered
5.Total tests
6. New case
7.New death
8.New recovered
1
Total cases in bangladesh = 1931304

<bangladesh><Total_cases><1931304>

Want to explore other datas on same country (y/n)> n
Do you want to explore other countries (y/n)? n

E:\mtech_2ND_SEM\computing lab 2\week 4&5\assignment_4\21CS60A01_CL2_A4>
```