PMC2221A-Big Data Analytics (Core Course - Elective)-Web Scraping in Python

On

TITLE: Worldometer - COVID-19 Coronavirus Pandemic

Date: 09 Aug 2023

Submitted by:

Nandhu Jayan

22pmc139

Submitted to:

Sr. Italia Joseph Maria

(Assistant proffessor)

PG Department of computer applications

INTRODUCTION:

The rapid global spread of the COVID-19 pandemic has underscored the critical importance of accurate and up-to-date data in understanding and managing its impact. In response to this need, we have undertaken a web scraping project with the objective of extracting real-time COVID-19 statistics from the Worldometer website

PURPOSE OF THE PROJECT

The purpose of this project is to develop a web scraping solution for extracting real-time COVID-19 statistics and data from the Worldometer website. The scraped data is intended to be used for monitoring and analyzing the global and country-specific trends of the COVID-19 pandemic.

USE CASE: COVID-19 Data Monitoring and Analysis

The main use case of this web scraping project is to monitor and analyze the global and country-specific trends of the COVID-19 pandemic. By extracting data from the Worldometer website, we aim to achieve the following:

- **Data Monitoring:** The scraped data will provide a real-time snapshot of the current state of the pandemic, including the number of active cases, recoveries, and deaths for each country.
- **Trend Analysis:** By collecting historical data over time, we can analyze how the pandemic has evolved in terms of total cases, deaths, recoveries, and other relevant metrics. This can help in understanding the progression of the pandemic and identifying patterns.
- **Geographical Comparison:** The extracted data allows for comparing the impact of COVID-19 across different countries and regions. This information can aid in assessing the effectiveness of containment measures and healthcare responses.
- **Decision Support:** The data can be utilized by healthcare organizations, researchers, policymakers, and the general public to make informed decisions and recommendations related to the pandemic.

URL OF THE WEBPAGE:

https://www.worldometers.info/coronavirus/

TITLE OF THE PAGE:

Worldometer - COVID-19 Coronavirus Pandemic

EXTRACTED DATA:

From the Worldometer COVID-19 page, we are extracting the following attributes for each country:

Country

Other

Total Cases

New Cases

Total Deaths

New Deaths

Total Recovered

New Recovered

Active Cases

Serious Critical

Cases per 1M Population

Deaths per 1M Population

Total Tests

Tests per 1M Population

Population

GITHUB REPOSITORY:

https://github.com/nandhujayan/covid-19-web-scrapping-python.git

CONCLUSION

n conclusion, our web scraping project focused on monitoring and analyzing real-time COVID-19 data from the Worldometer website has successfully achieved its objectives and provided valuable insights into the ongoing pandemic. By extracting key metrics such as total cases, deaths, recoveries, and more, we have contributed to informed decision-making, public awareness, and research efforts.

Through the utilization of web scraping techniques, we have established a robust data collection mechanism that ensures the availability of up-to-date and accurate information. The data collected has been stored in a structured format, enabling seamless analysis and visualization for various stakeholders.