

Code Alpha Internship

Task 1: Web Scraping

Name: Ramya Manjunath Angadi

Internship Domain: Data Analyst Task Name: Web Scraping

1. Introduction

Web scraping is the process of automatically extracting data from websites using programming tools. Instead of manually copying information from web pages, web scraping allows data to be collected efficiently and accurately.

2. Objective

The objective of this task is to collect relevant data from a public website using Python and create a structured dataset for further analysis.

3. Tools and Technologies Used

- Programming Language: Python •
Code Editor: Visual Studio Code
 - Libraries Used: ○ Requests ○
BeautifulSoup
 - Data Storage Format: CSV file
-

4. Methodology

1. A public website was selected for data extraction.
 2. The Requests library was used to send an HTTP request to the website.
 3. The BeautifulSoup library was used to parse the HTML content.
 4. Relevant data (quotes and author names) was extracted using HTML tags and class names.
 5. The extracted data was saved into a CSV file to create a custom dataset.
-

5. Output

The program successfully extracted quotes and their respective authors from the website and stored them in a CSV file named quotes.csv.

6. Conclusion

This task helped in understanding the basics of web scraping, HTML structure, and automated data collection. The use of Python and BeautifulSoup made it easy to extract and organize web data efficiently.