

# Scraping Quotes from a Website and Saving to CSV and DOCX

## Introduction

This Python script scrapes quotes from the website "Quotes to Scrape" and saves them in both CSV and DOCX formats. The script ensures that up to 500 quotes are extracted, including the quote text, author name, and associated tags.

## Functionality

### 1. Automated File Naming:

- The script checks for existing files and assigns the next available filename to prevent overwriting.

### 2. Web Scraping Process:

- Uses the requests library to fetch HTML pages.
- Parses the content with BeautifulSoup to extract quotes, authors, and tags.
- Iterates through multiple pages until 500 quotes are collected.

### 3. Saving Data:

- Quotes are stored in a CSV file with columns for quote text, author, and tags.
- A DOCX file is also generated, formatting quotes in a structured manner.

## Implementation Details

- **Base URL:** The script constructs URLs dynamically to navigate pages.
- **Looping Condition:** The loop continues scraping until 500 quotes are collected or no more pages are available.
- **CSV Writing:** Data is stored using csv.writer, ensuring proper encoding and formatting.
- **DOCX Formatting:** Quotes are numbered and formatted with appropriate styles.

## Conclusion

This script efficiently scrapes and organizes quotes from a web source, making them easily accessible in structured formats. It demonstrates automation using Python's web scraping and file handling capabilities.