**Charities Bureau Web Scraping**

This Python script automates web scraping of the New York State Charities Bureau website to extract information about registered charities. It utilizes the Selenium library for web automation and Pandas for data manipulation. Additionally, it demonstrates uploading the scraped data to an AWS S3 bucket for storage.

**Instructions-**

**Prerequisites**

1. Ensure you have Python installed on your system.
2. Install necessary Python packages:

bash

!pip install selenium pandas webdriver-manager boto3

**Usage**

1. Clone the repository or download the script file (**M10\_webscraper\_assignment.ipybn**).
2. Run the script using Jupyter Lab:
3. The script will scrape the Charities Bureau website, extract data from the table, save it as a CSV file locally (**my\_charities\_data.csv**), create an AWS S3 bucket (if not already created), and upload the CSV file to the bucket with a timestamped filename.

**Code Explanation**

* **Web Scraping**: The script uses Selenium to automate web browsing and extract data from the Charities Bureau website.
* **Data Processing**: It organizes the extracted data into a Pandas DataFrame for easy manipulation and analysis.
* **AWS S3 Integration**: The script interacts with AWS S3 using the **boto3** library to create a bucket and upload the CSV file for storage.

**File Descriptions**

* **M10\_webscraper\_assignment.ipybn**: The main Python script for web scraping and AWS S3 interaction.
* **charities\_bureau\_scrape\_20240415130036.csv**: The final output CSV file containing the scraped data is in the s3 bucket.
* **README.md**: This file, providing instructions and an overview of the script.

Feel free to modify the script as needed for your specific use case. If you encounter any issues or have suggestions for improvement, please open an issue or pull request on GitHub.