

Automated Crypto Web Scraper

Problem Statement:

The increasing popularity of cryptocurrencies such as Bitcoin has led to a demand for real-time data tracking and analysis. Many cryptocurrency enthusiasts, traders, and analysts require automated systems to collect and store price data continuously for analysis, forecasting, and decision-making.

However, manually tracking cryptocurrency prices can be time-consuming and inefficient. Therefore, there is a need for an automated solution that can scrape real-time cryptocurrency price data (such as Bitcoin) from the CoinMarketCap website, save the data in a structured format (CSV), and run the scraper continuously on a schedule (every hour) without the need for manual intervention.

Input: URL of the cryptocurrency page to scrape (<https://coinmarketcap.com/currencies/bitcoin/>)

Output: A CSV file containing the cryptocurrency name, its price, and the timestamp for each data pull.

Challenges:

1. **Real-Time Data Retrieval:** The challenge is to retrieve up-to-date cryptocurrency prices, especially for popular currencies like Bitcoin, and store the information efficiently for future analysis.
2. **Automation:** The data collection needs to happen automatically at regular intervals (every hour in this case) to ensure that the data is up-to-date and useful for ongoing analysis.
3. **Data Storage:** The data must be stored in a structured format that is easy to read and analyze. The solution should ensure that the data is saved in a way that prevents data loss and enables the addition of new data without overwriting existing information.
4. **Handling Website Changes:** The CoinMarketCap website might change its layout, HTML structure, or class names, which could break the scraping logic. The system should be able to handle such changes or at least flag when an issue occurs.
5. **Error Handling:** Web scraping can encounter errors due to network issues, website unavailability, or unexpected HTML structure changes. The scraper should be robust and handle such errors gracefully without crashing.

Functional Requirements:

- **Web Scraping:** Extract real-time cryptocurrency price data (Bitcoin in this case) from a public website, CoinMarketCap, by parsing the HTML of the page.
- **Data Storage and Formatting:** Store the scraped data (cryptocurrency name, price, timestamp) in a CSV file for future analysis, ensuring that the file format supports easy access and organization of data.
- **Continuous Execution:** Set up an automation system that runs indefinitely and performs the data collection at periodic intervals (e.g., every hour), ensuring that the system collects data without manual intervention.
- **File Management:** Ensure that the scraper checks for the existence of the data file. If the file exists, it should append the new data; if not, it should create a new file to store the data.
- **Logging and Monitoring:** Print the collected data to the console for real-time monitoring and ensure that any issues are logged and handled (e.g., network failure, website changes)