

Crawling scroll webpage tutorial:

Step 1: Inspect elements and select **Network** tab.

The screenshot shows a web browser with the URL <https://cafef.vn/vi-mo-dau-tu.chn>. The Network tab is open, displaying a list of requests. The first request is 'tssp?domain=cafef.vn&path=...' with a status of 200 and a type of text/plain. The second request is 'tssp?domain=cafef.vn&path=...' with a status of 200 and a type of text/plain. The third request is 'tssp?domain=cafef.vn&path=...' with a status of 200 and a type of text/plain. The fourth request is 'tssp?domain=cafef.vn&path=...' with a status of 200 and a type of text/plain. The fifth request is 'tracking.gif?...' with a status of 200 and a type of image/gif. The 'tracking.gif?' request is highlighted.

Step 2: Scroll until the scroll action is called.

Find a link that contains a list of news.

In this case is: <https://cafef.vn/timeline/33/trang-3.chn>, format: <https://cafef.vn/timeline/33/trang-{0}.chn>

The screenshot shows a web browser with the URL <https://cafef.vn/vi-mo-dau-tu.chn>. The Network tab is open, displaying a list of requests. The first request is 'trang-2.chn' with a status of 200 and a type of text/html. The second request is 'get-ist-expert-stock.chn' with a status of 200 and a type of text/html. The third request is 'ajax-brandcontent.chn' with a status of 200 and a type of text/html. The fourth request is 'photo1603629735672-1603629735812228992148.jpg' with a status of 200 and a type of image/jpeg. The fifth request is 'photo1603629611841-16036296120361471926332.jpg' with a status of 200 and a type of image/jpeg. The sixth request is 'photo1603629309566-16036293098071300023525.png' with a status of 200 and a type of image/png. The seventh request is 'photo1603628391964-16036283921621034768304.jpg' with a status of 200 and a type of image/jpeg. The eighth request is 'photo1603596833070-1603596833501171056738.jpg' with a status of 200 and a type of image/jpeg. The ninth request is 'photo1603553386545-1603553386782903979748.jpeg' with a status of 200 and a type of image/jpeg. The 'trang-2.chn' request is highlighted.

Thông tin kinh tế vĩ mô, hoạt động

THỜI SỰ CHUNG KHOẢN BẤT ĐỘNG SẢN DOANH NGHIỆP NGÂN HÀNG TÀI CHÍNH QUỐC TẾ VĨ MÔ SỐNG THỊ TRƯỜNG ...

Dữ liệu Top 200

1 ngày trước

Nhiều địa phương liên tục chấp thuận các dự án điện gió điện mặt trời sử dụng nhiều diện tích đất rừng và cho phép nộp tiền trồng rừng thay thế. Tuy nhiên mức độ cấp thiết để...

Tập đoàn, tổng công ty nhà nước nợ nước ngoài 341.591 tỷ đồng

1 ngày trước

Nợ vay nước ngoài của các tập đoàn, tổng công ty nhà nước đến cuối năm 2019 đạt 341.591 tỷ

Elements Console Sources Network Performance Memory Application Security Lighthouse

Filter Hide data URLs All XHR JS CSS Img Media Font Doc WS Manifest Other Has blocked cookies Blocked Requests

Name

- trang-2.chn
- get-list-expert-stock.chn
- ajax-brandcontent.chn
- photo1603629735672-1603629735812228992148.jpg
- photo1603629611841-16036296120361471926332.jpg
- photo1603629309566-16036293098071300023525.png
- photo1603628391064-16036283921621034768304.jpg
- photo1603596833070-1603596833501171056738.jpg
- photo1603553586545-1603553586782903979748.jpeg

33 requests 22.5 kB transferred 1.3 MB resources

Headers Preview Response Initiator Timing Cookies

Tập đoàn, tổng công ty nhà nước nợ nước ngoài 341.591 tỷ đồng

Type here to search

11:07 PM 10/26/2020

Step 3: Check other links and get the scroll action link format

<https://cafef.vn/timeline/33/trang-3.chn>

Thông tin kinh tế vĩ mô, hoạt động

https://cafef.vn/timeline/33/trang-3.chn

Nhiều tập đoàn Hàn Quốc muốn đến Khánh Hòa

Kinh tế vĩ mô - Đầu tư

Các tập đoàn lớn của Hàn Quốc đang xúc tiến kế hoạch đầu tư ở Khánh Hòa trong nhiều lĩnh vực điện máy, công nghiệp

World Bank: Mỗi năm thiên tai gây thiệt hại hơn 10 tỷ đô cho Việt Nam, 60% do bão lũ

Kinh tế vĩ mô - Đầu tư

Trên toàn quốc, trung bình thiệt hại tài sản công và tư hàng năm theo sức mua tương đương lên đến 8,1 tỷ đô la Mỹ, tương đương khoảng 2,7 tỷ đô la Mỹ giá trị thực. Khoảng 60%...

Gỡ thẻ vàng IUU: Ba năm vẫn 'mong manh'

Type here to search

11:07 PM 10/26/2020

Code flow:

Get start url data (in this case is <https://cafef.vn/vi-mo-dau-tu.chn>) → get the highlight news in this page → call the start function with scroll action links → get the news in current page.