

[| 카카오톡 봇 강좌 | >](#)

## 정말 왕초보들을 위한 jsoup 파싱방법



이슬두병 챗봇 고수 1:1 채팅

2020.04.28. 12:20 조회 501

댓글 11 URL 복사

강좌할 주제

=> 파싱방법

강좌 내용

=>

안녕하세요.

정말 파싱에 대해서 감도 못잡고, jsoup에 대해서는 아무것도 모르는 저같은 초보들을 위해서 씁니다.

고수분들 보면 복잡한 함수로 후딱후딱 파싱을 한줄로 끝내버리는데, 저희같은 초보들은 강좌를 보고도 이해가 되지 않는것이 사실입니다.

제가 그랬거든요. ㅠㅠ

그래서 아주 쉽게 파싱하는 방법을 알아보도록 하겠습니다.

일단 저는 코인투자자라서 코인가격을 파싱해보려고 합니다.

저번글에서 썼듯이 만약 거래소에서 json형태로 정보를 던져준다면 아주 쉽게 파싱이 끝납니다.

[\(json을 통한 파싱방법 참고\)](#)

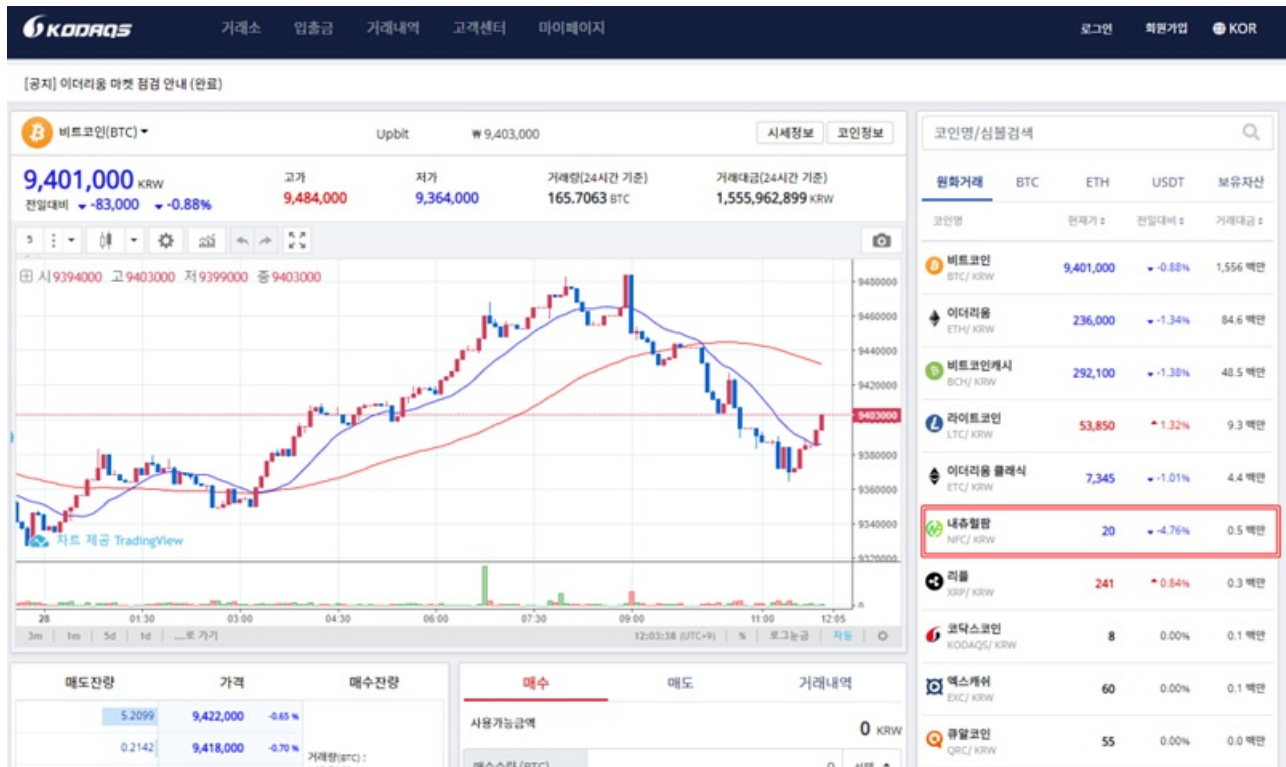
하지만 중소 거래소나, json을 제공하지 않는 거래소의 경우 직접 파싱을 할 수밖에 없습니다.

아래 주소의 빨간색 네모부분을 파싱해보도록 하겠습니다.

소스는 아래와 같으며 스크린샷과 함께 설명하도록 하겠습니다.

```
function response(room, msg, sender, isGroupChat, replier, imageDB, packageName) {  
  
    if(msg=="/nfc가격" || msg == "/NFC가격") {  
  
        var url_nfc = 'https://www.kodags.com/trd/order/dashboard.do'; // 파싱할 주소  
  
        var data_nfc = org.jsoup.Jsoup.connect(url_nfc).get();  
        data_nfc = data_nfc.select("span[class=float-left]").text();  
        var data_nfc_16 = data_nfc.split(" ")[16];  
  
        replier.reply(data_nfc_16);  
  
        var data_nfc1 = org.jsoup.Jsoup.connect(url_nfc).get();  
        data_nfc1 = data_nfc1.select("span[class=d-block]").text();  
        var data_nfc_2_20 = data_nfc1.split(" ")[20];  
        var data_nfc_2_21 = data_nfc1.split(" ")[21];  
        var data_nfc_2_22 = data_nfc1.split(" ")[22];  
        var data_nfc_2_23 = data_nfc1.split(" ")[23];  
  
        replier.reply("가격 : " + data_nfc_2_20 + "\n" + "변동율 : " + data_nfc_2_21 + "\n" + "거래대금 :  
" + data_nfc_2_22 + "( " + data_nfc_2_23 + " )");  
  
    }  
}
```

주소는 <https://www.kodas.com/trd/order/dashboard.do> 이며, 빨간 네모부분을 파싱할겁니다.



이페이지에서 파싱하고 싶은 부분을 누른 후 오른쪽 버튼을 누른 후 검사를 클릭합니다.

The screenshot shows the Chrome DevTools 'Elements' panel. The HTML structure of the trading dashboard is displayed. The 'price-list-tr-nfc-kw' element is selected, and its styles are visible in the 'Styles' panel. The styles include 'float: left !important;' and 'font-size: 12px;'. A diagram on the right shows the box model with a width of 52.516 and a height of 35.

```

class="pointer">...</tr>
<input type="hidden" id="ltc-kw" value="50">
<tr id="price-list-tr-etc-kw" name="이더리움 클래식, ETC-0" class="pointer">...</tr>
<input type="hidden" id="etc-kw" value="5">
<tr id="price-list-tr-nfc-kw" name="내츄럴팜, NFC-0" class="pointer">
  <td>
    <span class="float-left m-1">...</span>
    <span class="float-left">...</span> == $0
  </td>
  <td class="text-right mr-3 ratePrice text-primary">...</td>
  <td class="text-primary" style="padding: 19px 0.75em 0px 0px !important; text-align: right;">...</td>
  <td style="padding: 19px 0.75em 0px 0px !important;">...</td>
</tr>
<input type="hidden" id="nfc-kw" value="1">
<tr id="price-list-tr-xrp-kw" name="리플 XRP-0" class="pointer">...</tr>
  
```

... #div\_priceList #price-list-table tbody #price-list-tr-nfc-kw td span.float-left

Styles

Filter

element.style {

.float-left {

float: left !important;

\*, ::before, ::after {

box-sizing: border-box;

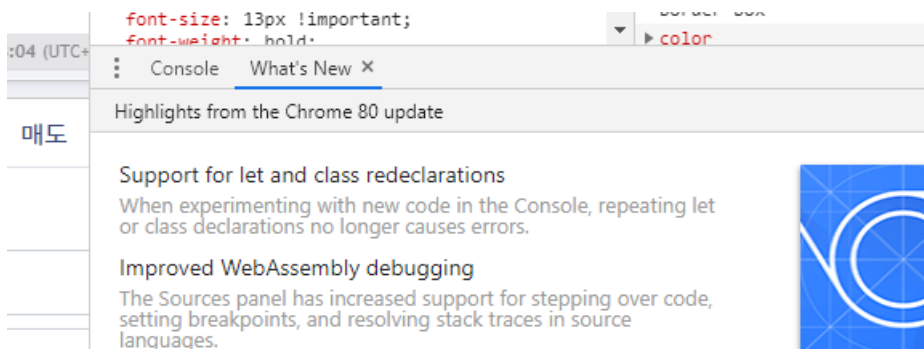
Inherited from td

.ctype-table td:first-child {

font-size: 12px;

padding: 0.7rem 0;

.ctype-table td:first-child {



화면 오른쪽에 뭔가 복잡한 것이 나오지만, 당황하실 것이 없습니다.

이제 파싱하고 싶은 곳에 마우스를 가져다 대면 오른쪽에서 뭔가 왔다갔다 하는게 보이실 겁니다.



이제 제가 파싱하고 싶은 부분이 나왔군요

이제 저 class 오른쪽에 따옴표부분을 잘 기억하시고, 위의 소스를 메모장에 붙여넣으신 후 보도록 하겠습니다.

```
var url_nfc = 'https://www.kodaqs.com/trd/order/dashboard.do'; // 파싱할 주소
```

이부분은 파싱하고 싶은 주소를 넣어주는 부분입니다.

```
var data_nfc = org.jsoup.Jsoup.connect(url_nfc).get();
```

이부분은 저 주소에 있는 부분을 몽땅 가져와서 data\_nfc라는 변수에 넣었다고 생각하시면 됩니다.

```
data_nfc = data_nfc.select("span[class=float-left]").text();
```

이부분이 중요한데, 위에서 따옴표부분을 기억해달라고 했죠?  
그부분을 select해서 data\_nfc변수에 text형태로 넣어주는 부분입니다.

다른 파싱강좌 보시면 복잡한 html 태그들 보셨죠?  
.text()함수 때문에 태그는 하나도 없이, 텍스트만 data\_nfc변수에 들어가게 됩니다.

```
replier.reply(data_nfc);
```

해보시면 아래와 같이 class에 float-left라는 속성을 가진 텍스트가 전부 출력이 되게 됩니다.

```
비트코인 BTC/ KRW 이더리움 ETH/
KRW 비트코인캐시 BCH/ KRW 라이
트코인 LTC/ KRW 이더리움 클래식
ETC/ KRW 내츄럴팜 NFC/ KRW 리플
XRP/ KRW 코닥스코인 KODAQS/
KRW 엑스캐쉬 EXC/ KRW 큐알코인
QRC/ KRW 테더 USDT/ KRW 헥스코
인 HEG/ KRW 올리브마운틴 OMC/
KRW 넥소 NEXO/ KRW 비트원코인
BWC/ KRW 지비코인 GBC/ KRW
```

보시는대로 텍스트는 띄어쓰기 형태로 출력이 되며, 우리가 찾고싶은 값은 17번째에 있군요

```
var data_nfc_16 = data_nfc.split(" ")[16];
```

이런식으로 하시면 띄어쓰기를 제외하고 17번째의 값을 가져와서 data\_nfc\_16이라는 변수를 골라낼 수 있습니다.

```
var data_nfc1 = org.jsoup.Jsoup.connect(url_nfc).get();
data_nfc1 = data_nfc1.select("span[class=d-block]").text();
var data_nfc_2_20 = data_nfc1.split(" ")[20];
var data_nfc_2_21 = data_nfc1.split(" ")[21];
var data_nfc_2_22 = data_nfc1.split(" ")[22];
var data_nfc_2_23 = data_nfc1.split(" ")[23];
```

이제 이부분도 이해가 되시겠죠?

가격이 있는 부분의 속성이 d-block이기 때문에 거기서 값을 가져온 것입니다.

저같은 왕초보들을 위한 파싱방법을 설명드렸으며, 조금이라도 도움이 되었으면 좋겠네요.

긴글 읽어주셔서 감사드립니다.

( 말머리 등록 해주세요! )



이슬두병님의 게시글 더보기 >

❤️ 좋아요 6 💬 댓글 11

🔗 공유 | 신고

댓글 등록순 최신순 ↻

댓글알림 ☐



지풍



2020.04.28. 12:52 답글 쓰기



lunarbox



2020.04.28. 15:53 답글 쓰기



JHK



2020.04.28. 16:32 답글 쓰기



도미 doami2005



2020.04.28. 16:47 답글 쓰기



보라댕이



2020.04.28. 20:09 답글 쓰기



delta

span[class=float-left] = span.float-left

2020.04.30. 10:35 답글 쓰기



토르륵

실제는 거래소 api 가 제공된다는거... ㅠ

2020.04.30. 11:52 답글 쓰기



이슬두병 작성자

맞지만 거래소 api가 제공되지 않는 듣보잡(?) 거래소의 경우 파싱을 해야해서 작성했습니다 ㅠㅠ

2020.05.06. 16:28 답글 쓰기



토르륵

이슬두병 우하! 듣보잡! 그럼 따라하기해볼게영 ㅎㅎ

2020.05.06. 19:03 답글 쓰기



Yunbot



2020.07.03. 23:05 답글 쓰기



Kiri



2020.07.06. 10:15 답글 쓰기

Hibot

댓글을 남겨보세요



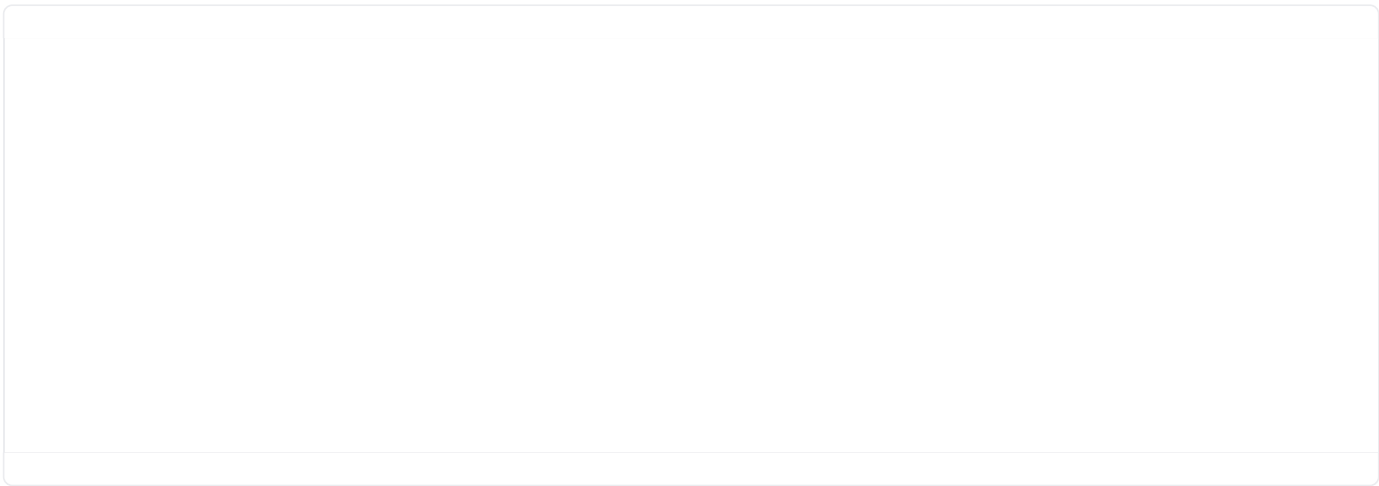
등록

글쓰기

답글

목록

▲ TOP



' 카카오톡 봇 강좌 |' 게시판 글

이 게시판 새글 구독하기 ☐

카카오톡 이용자 보호조치 푸는법 [11]	디돈	2020.05.01.
카카오톡 부계정 만드는법 🤖 [4]	디돈	2020.04.30.
정말 왕초보들을 위한 jsoup 파싱방법 🤖 [11]	이슬두병	2020.04.28.
파일스트림 강좌(기초 간략) [9]	Leesy006	2020.04.27.
[고급] [안드로이드] 기기 로컬 언어설정 가져오기 🤖 [5]	jomin398	2020.04.26.

1 2 3 전체보기

이 카페 인기글



"시초가 중국발" 발언 관련 해명글 및 부탁  
의 글  
타 메신저 사용시 무한반복  
트위치봇  
New5 🗨️3  
👍0 🗨️1

[ 치\*\*, 한\*\* ] 님 강제 탈퇴  
AlphaDo  
👍1 🗨️2

함수를 입력받아 그래프 그리기  
Steve28  
👍2 🗨️4

오픈이발 괴롭히기(메봇R)  
ERROR  
👍0 🗨️7



틱택토 (Player vs Bot)

1 2 3 4 5

예전에 어떤분이 카톡봇을 이용하여 카페에  
게시글을 올렸었는데  
리페  
👍0 🗨️1

안녕하세요!  
트위터봇  
👍0 🗨️6

카톡방 나가기  
hajuhee01  
👍0 🗨️16