

# 데이터 크롤링과 정제

---

과제2: 커피 매장 찾기

# 커피 매장 검색

- 할리스 커피 매장 찾기
  - <http://www.hollys.co.kr>
  - 홈페이지 상단 메뉴에서 Store를 선택

창업설명회

기업특판

HOLLYS

멤버십 로그인

멤버십 가입

고객센터

English

Menu

HOLLYS Mail

Membership

HOLLYS News

Store

HOLLYS is

가맹문의

매장검색

Membership

할리스 멤버십 'CROWN'으로 다양한 혜택을 즐겨보세요

HOLLYS

Membership

Home > Store > 매장검색

매장검색

지도검색

· 찾고자 하시는 지역을 클릭하세요.

매장검색

지역검색

시/도

구/군

매장검색

검색

서비스별 매장검색

DT 매장

테라스

24시간

흡연시설

주차

\* 할리스 전 매장에서 무선인터넷 서비스 이용이 가능합니다 (단, 휴게소 및 특수매장 제외)

지역	매장명	현황	주소	매장 서비스	전화번호
서울 구로구	오류동점	영업중	서울특별시 구로구 경인로 224 (오류동, 칸타빌레5차)		02-2625-4854
경기 고양시 일산동구	부천역점	영업중	경기도 부천시 부일로 448 (삼곡동, 프리온빌딩) 110호, 204호	<div>DT</div> <div>P</div>	032-665-8255
경기 화성시	화성병점점	영업중	경기도 화성시 병점동로 105 1,2 층(신안동)		031-233-4686
서울 송파구	잠실학원사거리점	영업중	서울특별시 송파구 석촌호수로 130 (상천동, 신원빌딩) 상건물 1-1와 3월지		02-3431-4662

2

# HTML 코드 확인

## ■ 매장 정보 확인

### • <tbody>태그 내부

- <td> 태그 첫번째: 매장이 있는 지역
- <td> 태그 두번째: 매장 이름
- <td> 태그 네번째: 매장 주소
- <td> 태그 다섯번째: 매장 전화번호

```
522 <tbody>
523
524 <!--
525 <td class="noline center_t">
526
527 </td>
528
529 <td class="noline center_t">부산 사하구</td>
530 <td class="center_t"><a href="#" onclick="javascript:storeView(997); return false;">부산아트몰링영풍문고점</a></td>
531 <td class="center_t tdp0">영업중</td>
532 <td class="center_t"><a href="#" onclick="javascript:storeView(997); return false;">부산광역시 사하구 낙동남로 1413 (하단동, 아트몰링) 13층</a></td>
533 <td class="center_t">
534
535 </td>
536 <td class="center_t">051-201-0621</td>
537 </tr>
538
539 <!--
540 <td class="noline center_t">
541
542 </td>
543
544 <td class="noline center_t">경기 광명시</td>
545 <td class="center_t"><a href="#" onclick="javascript:storeView(996); return false;">광명GIDC점</a></td>
546 <td class="center_t tdp0">영업중</td>
547 <td class="center_t"><a href="#" onclick="javascript:storeView(996); return false;">경기도 광명시 일직로 43 일직동 514 지하2층 G249, G250</a></td>
548 <td class="center_t">
549
550 </td>
551 <td class="center_t">02-6914-6550</td>
552 </tr>
553
554 <!--
555 <td class="noline center_t">
556
557 </td>
558
```

## ■ 매장 정보 확인

### • <tbody>태그 내부

- <td> 태그 첫번째: 매장이 있는 지역
- <td> 태그 두번째: 매장 이름
- <td> 태그 네번째: 매장 주소
- <td> 태그 다섯번째: 매장 전화번호

지도검색

매장검색

지역검색

시/도

구군

매장검색

검색

서비스별 매장검색

☐ DT 매장 ☐ 테라스 ☐ 24시간

☐ 휴먼사일 ☐ 주차

\* 홀리스 전 매장에서 무선인터넷 서비스가 이용 가능합니다 (단, 휴게소 및 특수매장 제외)

매장명	현황	주소	매장 서비스	전화번호
서울 구로구	영업중	서울특별시 구로구 경인로 224 (오류동, 인터빌레5차)		02-2625-4854
경기 고양시 일신동구	부천역점	경기도 부천시 부일로 448 (삼국동, 프리온빌딩) 110호, 204호		032-665-8255
경기 화성시	화성병점	경기도 화성시 병점동로 105 1, 2층 (인안동)		031-233-4686
서울 송파구	잠실학원사거리점	서울특별시 송파구 석촌호수로 130 (삼전동, 신원빌딩) 상전동 1-1의 3필지		02-3431-4662
경북 구미시	구미시청점	경상북도 구미시 송정대로 54 (송정동)		054-452-8297
전남 순천시	순천조점	전라남도 순천시 송호로 109 (조례동, 디지털프라자전지랜드)		061-902-7288

```
<div class="lnb_wr"></div>
<script type="text/javascript"></script>
<div class="content">
  <div class="location"> Home > Store > 매장검색 </div>
  <h2></h2>
  <!-- 지도 및 매장검색 -->
  <fieldset>
    <legend></legend>
    <div class="store_view01" style="margin-bottom:5px;"></div>
    <fieldset>
      <div class="tableType01">
        <table class="tb_store" summary>
          <caption></caption>
          <colgroup></colgroup>
          <thead></thead>
          <tbody>
            <tr>
              <td class="noline center_t">
                <!--
                <div class="store_view01">
                  <img alt="Hollys Logo" />
                </div>
                </td>
              <td class="noline center_t">서울 구로구</td>
              <td class="center_t">
                <a href="#" onclick="javascript:storeView(1083); return false;">점</a>
              </td>
              <td class="center_t">영업중</td>
              <td class="center_t">
                <a href="#" onclick="javascript:storeView(1083); return false;">점</a>
              </td>
              <td class="center_t"></td>
            </tr>
          </tbody>
        </table>
      </div>
    </fieldset>
  </div>
```

# 나머지 매장 정보 확인

## ■ URL 주소

- <https://www.hollys.co.kr/store/korea/korStore2.do?pageNo=1&sid=&gugun=&store=>
  - ‘pageNo=’ 다음에 페이지 번호가 붙음
  - pageNo=1 ~ pageNo=51
- 마지막 페이지
  - <https://www.hollys.co.kr/store/korea/korStore2.do?pageNo=51&sid=&gugun=&store=>

\* 할리스 전 매장에서 무선인터넷 서비스 이용이 가능합니다 (단, 휴게소 및 특수매장 제외)

지역	매장명	현황	주소	매장 서비스	전화번호
경기 수원시 팔달구	수원역점	영업중	경기도 수원시 팔달구 향교로 6 (매산로1가)		031-257-3313
서울 종로구	종로코리안리빌딩점	영업중	서울특별시 종로구 종로5길 68 코리안리빌딩 1층 할리스	📞	070-4733-4828
경기 성남시 분당구	성남터미널점	영업중	경기도 성남시 분당구 성남대로 925번길 16 (아탑동, 분당종합터미널) 1층 할리스	📞	031-725-1004
서울 마포구	합정역점	영업중	서울특별시 마포구 양화로 36 (합정동 374-1) 할리스	📞	02-6204-1234
부산 해운대구	부산달맞이점	영업중	부산광역시 해운대구 달맞이길 199, 2~3층 (중동 1488-14) 할리스	📞 📶	051-731-3410
경기 성남시 수정구	신흥역점	영업중	경기도 성남시 수정구 산성대로 265 (신흥동) 2층 할리스	📞	031-626-6140
서울 관악구	신림점	영업중	서울특별시 관악구 신림로 353-1	📞 📶 📶	02-877-0019
서울 중구	태평로점	영업중	서울특별시 중구 세종대로 64, 해남빌딩 1층 (태평로2가 70-5) 할리스	📞	02-755-7795
서울 서대문구	신촌점	영업중	서울특별시 서대문구 연세로 34 (창천동 31-12) 할리스	📞 📶	02-393-2004

매장 마지막 페이지

# 과제 내용

- 1. 매장 찾기에서 1~51페이지까지 모든 매장의 정보를 스크레이핑
  - 지역, 매장명, 매장 주소, 전화번호
  - 수집된 정보는 csv 파일로 저장함
  - 결과물
    - csv파일: `hollys_branches.csv` (utf-8로 인코딩)
    - 제출 파일: `hw02_coffeeshop_to_csv.py`
- 2. 저장된 csv파일을 읽어서 사용자가 입력한 지역에 있는 매장 정보 출력
  - 제출 파일: `hw02_find_coffeeshop.py`

지역	매장명	현황	주소	매장 서비스	전화번호
서울 구로구	오류동점	영업중	서울특별시 구로구 경인로 224 (오류동, 칸타빌레5차) .		02-2625-4854
경기 고양시 일산동구	부천역점	영업중	경기도 부천시 부일로 448 (심곡동, 프리존빌딩) 110호, 204호		032-665-8255
경기 화성시	화성병점점	영업중	경기도 화성시 병점동로 105 1,2층(진안동)		031-233-4686
서울 송파구	잠실학원사거리점	영업중	서울특별시 송파구 석촌호수로 130 (삼전동, 신원빌딩) 삼전동 1-1외 3필지		02-3431-4662

# 과제2의 #1번 실행 결과

## ■ 1번 실행 결과

### • 화면 출력

[ 1]: 매장이름: 부산사상광장점, 지역: 부산 사상구, 주소: 부산광역시 사상구 광장로 22 (괘법동) 2층 사상구 괘법동 565-2, 전화번호: 051-322-4117  
[ 2]: 매장이름: 분당파크타운점, 지역: 경기 성남시 분당구, 주소: 경기도 성남시 분당구 내정로 170 (수내동, 동신파크빌딩) ., 전화번호: 031-713-8570  
[ 3]: 매장이름: 하남덕풍점, 지역: 경기 하남시, 주소: 경기도 하남시 덕풍북로 2 (덕풍동) ., 전화번호: 070-4791-4731  
... (중간 생략)  
[503]: 매장이름: 태평로점, 지역: 서울 중구, 주소: 서울특별시 중구 세종대로 64, 해남빌딩 1층 (태평로2가 70-5) 할리스, 전화번호: 02-755-7795  
[504]: 매장이름: 신촌점, 지역: 서울 서대문구, 주소: 서울특별시 서대문구 연세로 34 (창천동 31-12) 할리스, 전화번호: 02-393-2004  
전체 매장 수: 504  
hollys\_branches.csv 파일 저장 완료

### • csv 파일 내용

매장이름	위치(시,구)	주소	전화번호
부산사상광장점	부산 사상구	부산광역시 사상구 광장로 22 (괘법동) 2층 사상구 괘법동 565-2	051-322-4117
분당파크타운점	경기 성남시 분당구	경기도 성남시 분당구 내정로 170 (수내동, 동신파크빌딩) .	031-713-8570
하남덕풍점	경기 하남시	경기도 하남시 덕풍북로 2 (덕풍동) .	070-4791-4731
대전오류점	대전 중구	대전광역시 중구 계룡로 876 메종드메디컬빌딩 104호, 202호 .	042-522-3141
계룡대점	충남 계룡시	충청남도 계룡시 신도안면 신도안1길 43 (계룡스포츠펀터) 계룡스포츠펀터 1층	042-551-5508
구미오태DT점	경북 구미시	경상북도 구미시 오태동 609 .	054-461-0348
화정DI점	경기 고양시 덕양구	경기도 고양시 덕양구 화신로 254 (화정동)	031-971-4097
부산광복점	부산 중구	부산광역시 중구 광복로 70 (광복동2가) .	051-253-8033
양평북한강DI점	경기 양평군	경기도 양평군 양서면 양수로118번길 29 할리스	031-8079-7090
완주삼봉점	전북 완주군	전라북도 완주군 삼례읍 신기길 5-3 수계리 1212-4	063-261-2003
잠실새내역점	서울 송파구	서울특별시 송파구 올림픽로12길 8 (잠실동, 제일빌딩) .	02-416-2085
화성남양점	경기 화성시	경기도 화성시 남양읍2279-2 .	031-8067-7760

# 과제2의 #2번 실행 결과

2. 저장된 csv 파일을 읽어서 사용자가 입력한 도시 정보를 검색하고 결과를 출력 (무한 반복, 'quit'을 입력하면 종료)

검색할 매장의 도시를 입력하세요: 대구 북구

검색된 매장 수: 6

```
[ 1]: ['대구광역시 북구 동화천로 229 (연경동) 1층~2층', '053-982-6877']
[ 2]: ['대구광역시 북구 대학로 17길 3 (산격동 1393-11) 할리스 B1~3층', '053-951-9990']
[ 3]: ['대구광역시 북구 구암로 59 비전스퀘어 1,2층 .', '053-324-6122']
[ 4]: ['대구광역시 북구 침산남로 157 (침산동) .', '053-353-5544']
[ 5]: ['대구 북구 성북로 36,2층 601,602,603,604 473-3, 침산초등학교 맞은편', '053-353-5412']
[ 6]: ['대구광역시 북구 동천로 125 (동천동) .', '053-321-0099']
```

검색할 매장의 도시를 입력하세요: 서울 종로

검색된 매장 수: 11

```
[ 1]: ['서울특별시 종로구 종로 60-1 종로2가 102-3', '02-736-3399']
[ 2]: ['서울특별시 종로구 통일로 230 (무악동, 경희궁롯데캐슬) 상가동 1층', '010-2594-7308']
[ 3]: ['서울특별시 종로구 성균관로 6 (명륜2가 204) 1~4층', '02-743-8526']
[ 4]: ['서울특별시 종로구 청계천로 81 1,2,3,4층', '02-735-5462']
[ 5]: ['서울특별시 종로구 동호로 401(종로5가) .', '02-2266-1989']
[ 6]: ['서울특별시 종로구 대학로 101 (연건동, 서울대학교의과대학)', '02-741-5815']
[ 7]: ['서울특별시 종로구 새문안로5길 31, 센터포인트광화문빌딩 1층 (도림동 65, 센터포인트 광화문) 할리스', '02-6031-6213']
[ 8]: ['서울특별시 종로구 새문안로 58 (신문로2가, LG광화문빌딩) 1층, 지하1층', '02-720-6994']
[ 9]: ['서울특별시 종로구 대명길 6 2~4층 할리스', '02-747-0348']
[10]: ['서울특별시 종로구 세종대로 161, 2~3층 (세종로 200) 할리스', '02-720-5582']
[11]: ['서울특별시 종로구 종로5길 68 코리안리빌딩 1층 할리스', '070-4733-4828']
```

검색할 매장의 도시를 입력하세요: quit  
종료 합니다.





**Questions?**