



웹 크롤링 + 웹 페이지 만들기

해외축구 리그별 순위표

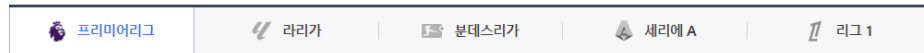
도성탁

차례

- 개요
- 개발환경
- 결과물
- 보완점
- 기타

개요

- 주제 : 해외축구 리그별 순위표
- 크롤링 사이트 : 네이버 스포츠 해외축구 순위(팀 순위)



2024/25 시즌 팀 순위

2024-25 현재

순위	팀	경기수	승점	승	무	패	득점	실점	득실차
1	맨체스터 시티 FC	3	9	3	0	0	9	2	7
2	리버풀 FC	3	9	3	0	0	7	0	7
3	브라이턴 앤 호브 알비온 FC	3	7	2	1	0	6	2	4
4	아스널 FC	3	7	2	1	0	5	1	4
5	뉴캐슬 유나이티드 FC	3	7	2	1	0	4	2	2
6	브렌트퍼드 FC	3	6	2	0	1	5	4	1
7	에버턴 FC	3	6	2	0	1	4	4	0
8	AFC 본머스	3	5	1	2	0	5	4	1
9	노팅엄 포리스트 FC	3	5	1	2	0	3	2	1
10	토트넘 홋스퍼 FC	3	4	1	1	1	6	3	3
11	첼시 FC	3	4	1	1	1	7	5	2
12	풀럼 FC	3	4	1	1	1	3	3	0
13	웨스트햄 유나이티드 FC	3	3	1	0	2	4	5	-1
14	맨체스터 유나이티드 FC	3	3	1	0	2	2	5	-3
15	레스터 시티 FC	3	1	0	1	2	3	5	-2
16	크리스탈 팰리스 FC	3	1	0	1	2	2	5	-3
17	입스위치 타운 FC	3	1	0	1	2	2	7	-5
18	울버햄튼 원더러스 FC	3	1	0	1	2	3	9	-6
19	사우샘프턴 FC	3	0	0	0	3	1	5	-4
20	에버턴 FC	3	0	0	0	3	2	10	-8

■ 챔피언스 리그 직행 ■ 챔피언스 리그 예선 ■ 유로파 리그 ■ 강등 직행



개발환경

- Visual Studio Code – 코드 편집기
- Python – 작성 언어
- Selenium – 웹 크롤링
- Flask – 파이썬 기반 웹 프레임워크
- HTML / CSS / JavaScript – 웹 페이지 구성

결과물

- 각 리그별 순위표를 크롤링하여 하나의 페이지에 기록하는 웹 페이지

리그별 팀 순위

프리미어리그 (잉글랜드)

순위	팀	경기수	승점
1 챔피언스 리그 직행	맨체스터 시티 FC	3	9
2 챔피언스 리그 직행	리버풀 FC	3	9
3 챔피언스 리그 직행	브라이턴 앤 호브 알비온 FC	3	7
4 챔피언스 리그 직행	아스널 FC	3	7
5 유로파 리그	뉴캐슬 유나이티드 FC	3	7
6	브렌트퍼드 FC	3	6
7	에버턴 FC	3	6
8	AFC 본머스	3	5
9	노팅엄 포리스트 FC	3	5
10	토트넘 홋스퍼 FC	3	4
11	첼시 FC	3	4
12	풀럼 FC	3	4
13	웨스트햄 유나이티드 FC	3	3
14	맨체스터 유나이티드 FC	3	3
15	레스터 시티 FC	3	1
16	크리스탈 팰리스 FC	3	1
17	임스워치 타운 FC	3	1
18 강등 직행	울버햄튼 원더러스 FC	3	1
19 강등 직행	사우스햄프턴 FC	3	0
20 강등 직행	에버턴 FC	3	0

라리가 (스페인)

순위	팀	경기수	승점
1 챔피언스 리그 직행	FC 바르셀로나	4	12
2 챔피언스 리그 직행	레알 마드리드	4	8
3 챔피언스 리그 직행	클루브 아틀레티코 데 마드리드	4	8
4 챔피언스 리그 직행	비야레알 CF	4	8
5 유로파 리그	지로나 FC	4	7
6 유로파 리그	데포르티보 알라베스	4	7
7	CA 오사수나	4	7
8	RC 셀타 데 비고	4	6
9	CD 레가네스	4	5
10	RCD 마요르카	4	5
11	라요 바예카노	4	4
12	레알 소시에다드	4	4
12	아틀레틱 클루브	4	4
14	RCD 에스파뇰	4	4
15	레알 바야돌리드	4	4
16	헤타페 CF	3	3
17	레알 베티스 발롱피에	3	2
18 강등 직행	UD 라스팔마스	4	2
19 강등 직행	세비야 FC	4	2
20 강등 직행	발렌시아 CF	4	1

분데스리가 (독일)

순위	팀	경기수	승점
1 챔피언스 리그 직행	1. FC 하이덴하임	2	6
2 챔피언스 리그 직행	FC 바이에른 뮌헨	2	6
3 챔피언스 리그 직행	RB 라이프치히	2	6
4 챔피언스 리그 직행	보루시아 도르트문트	2	4
5 챔피언스 리그 직행	FC 우니온 베를린	2	4
6 유로파 리그	보루시아 뮌헨글라트바흐	2	3
6	VfL 볼프스부르크	2	3
8	TSV 바이어 04 레버쿠젠	2	3
9	SC 프라이부르크	2	3
9	아인트라흐트 프랑크푸르트	2	3
11	TSG 1899 호펜하임	2	3
12	FSV 마인츠 05	2	2
13	SV 베르더 브라운	2	2
14	VfB 슈투트가르트	2	1
15	FC 아우크스부르크	2	1
16	홀슈타인 킬	2	0
17 강등 직행	FC 장크트파울리	2	0
17 강등 직행	VfL 보름	2	0

보완점

- 5개 리그의 순위 표를 하나의 페이지에 게시할 때의 가독성 하락 문제
- 완성된 웹 페이지의 단조로운 디자인

개선 방향

- 리그 선택 박스를 만들어서 리그 선택 시 해당 리그의 순위표만 게시
- 추가적인 디자인 요소를 적용하여 가독성 상승

기타

div#wfootballTeamRecordBody.record_tbl 1000 x 682

◀ 2024-25 ▶ 현재

순위	팀	경기수	승점	승	무	패	득점	실점	득실차
1	맨체스터 시티 FC	3	9	3	0	0	9	2	7
2	리버풀 FC	3	9	3	0	0	7	0	7
3	브라이턴 앤 호브 알비온 FC	3	7	2	1	0	6	2	4
4	아스널 FC	3	7	2	1	0	5	1	4
5	뉴캐슬 유나이티드 FC	3	7	2	1	0	4	2	2
6	브렌트퍼드 FC	3	6	2	0	1	5	4	1
7	애스턴빌라 FC	3	6	2	0	1	4	4	0
8	AFC 본머스	3	5	1	2	0	5	4	1
9	노팅엄 포리스트 FC	3	5	1	2	0	3	2	1
10	토트넘 홋스퍼 FC	3	4	1	1	1	6	3	3
11	첼시 FC	3	4	1	1	1	7	5	2
12	풀럼 FC	3	4	1	1	1	3	3	0
13	웨스트햄 유나이티드 FC	3	3	1	0	2	4	5	-1
14	맨체스터 유나이티드 FC	3	3	1	0	2	2	5	-3
15	레스터 시티 FC	3	1	0	1	2	3	5	-2
16	크리스탈 팰리스 FC	3	1	0	1	2	2	5	-3
17	입스위치 타운 FC	3	1	0	1	2	2	7	-5
18	올버햄튼 원더러스 FC	3	1	0	1	2	3	9	-6
19	사우스햄프턴 FC	3	0	0	0	3	1	5	-4
20	에버턴 FC	3	0	0	0	3	2	10	-8

```

<div class="record_tbl" id="wfootballTeamRecordBody">
  <table summary="팀 순위" cellspacing="0">
    <caption>팀 순위</caption>
    <colgroup>
      <col width="45">
      <col width="*">
      <col width="80">
      <col width="80">
      <col width="80">
      <col width="80">
      <col width="80">
      <col width="80">
      <col width="80">
    </colgroup>
    <thead>
    </thead>
    <tbody>
      <tr>
        <!-- [0] 상위링크는 best 클래스 추가 -->
        <td class="num best">
          <div class="inner_champsLeague">
            <strong>1</strong>
            <span class="state_bar1">
              <span class="blind">챔피언스 리그 직행</span>
            </span>
          </div>

```

기타

```
# 크롤링 함수 정의
def get_team_data(url):
    options = webdriver.ChromeOptions()
    options.add_argument("--headless") # 화면 없이 실행
    options.add_argument("--no-sandbox")
    options.add_argument("--disable-dev-shm-usage")

    driver = webdriver.Chrome(service=Service(ChromeDriverManager().install()), options=options)
    driver.get(url)
    time.sleep(3) # 페이지 로딩 대기

    teams = []
    try:
        table = driver.find_element(By.CSS_SELECTOR, 'table[summary="팀 순위"]')
        rows = table.find_elements(By.TAG_NAME, 'tr')[1:] # 헤더 제외

        for row in rows:
            columns = row.find_elements(By.TAG_NAME, 'td')
            if len(columns) > 0:
                rank = columns[0].text.strip()
                team_name = columns[1].text.strip()
                games_played = columns[2].text.strip()
                points = columns[3].text.strip()

                teams.append({
                    'rank': rank,
                    'team_name': team_name,
                    'games_played': games_played,
                    'points': points,
                })

    except Exception as e:
        print(f"Error: {e}")

    finally:
        driver.quit() # 드라이버 종료

    return teams
```


기타

```
# 개별 리그 크롤링 엔드포인트
@app.route('/premier')
def premier_league():
    url = 'https://sports.news.naver.com/wfootball/record/index'
    teams = get_team_data(url)
    return render_template('premier.html', teams=teams)

@app.route('/la')
def la_liga():
    url = 'https://sports.news.naver.com/wfootball/record/index?category=primera&tab=team'
    teams = get_team_data(url)
    return render_template('la.html', teams=teams)

@app.route('/bundes')
def bundesliga():
    url = 'https://sports.news.naver.com/wfootball/record/index?category=bundesliga&tab=team'
    teams = get_team_data(url)
    return render_template('bundes.html', teams=teams)

@app.route('/seria')
def seria_a():
    url = 'https://sports.news.naver.com/wfootball/record/index?category=seria&tab=team'
    teams = get_team_data(url)
    return render_template('seria.html', teams=teams)

@app.route('/league1')
def ligue_1():
    url = 'https://sports.news.naver.com/wfootball/record/index?category=ligue1&tab=team'
    teams = get_team_data(url)
    return render_template('league1.html', teams=teams)
```

```
# 모든 리그 데이터를 한 페이지에서 보여주는 엔드포인트
@app.route('/')
def index():
    premier_url = 'https://sports.news.naver.com/wfootball/record/index'
    la_url = 'https://sports.news.naver.com/wfootball/record/index?category=primera&tab=team'
    bundes_url = 'https://sports.news.naver.com/wfootball/record/index?category=bundesliga&tab=team'
    seria_url = 'https://sports.news.naver.com/wfootball/record/index?category=seria&tab=team'
    ligue_url = 'https://sports.news.naver.com/wfootball/record/index?category=ligue1&tab=team'

    teams1 = get_team_data(premier_url)
    teams2 = get_team_data(la_url)
    teams3 = get_team_data(bundes_url)
    teams4 = get_team_data(seria_url)
    teams5 = get_team_data(ligue_url)

    return render_template('index.html',
                           teams1=teams1,
                           teams2=teams2,
                           teams3=teams3,
                           teams4=teams4,
                           teams5=teams5)
```

A vertical bar on the left side of the slide with a color gradient from orange at the top to blue at the bottom.

감사합니다