데이터 크롤링과 정제

과제 1-2번: 네이버 날씨

네이버 날씨 정보 가져오기 #1 (제출 파일: hw01-02.py)

- 네이버에서 "대구 날씨" 입력:
 - url: https://search.naver.com/search.naver?query=대구+날씨
 - 한글 깨짐으로 quote('대구+날씨')로 검색
- ■위치 정보
 - <div class="title_area _area_panel">
 - <h2 class="title">중구 동인동1가</h2>





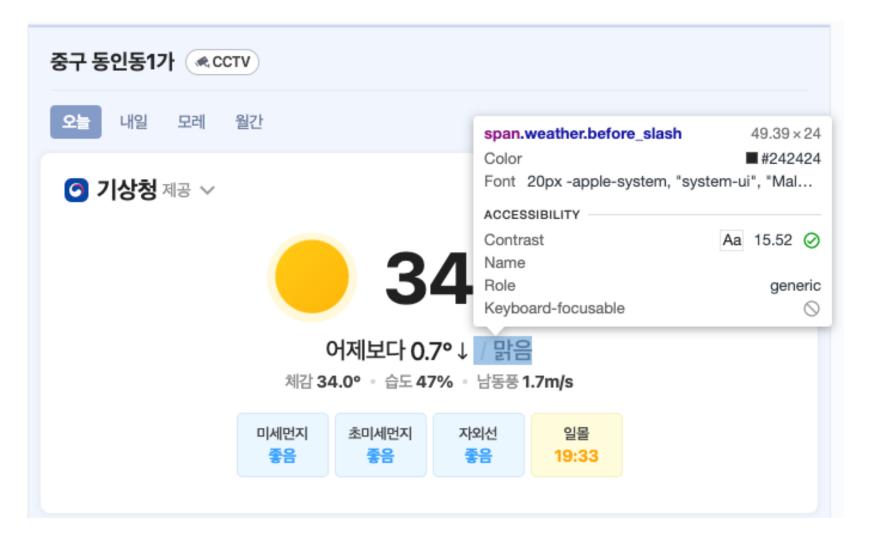
- ■날씨 정보 가져오기
 - <div class="weather_info">



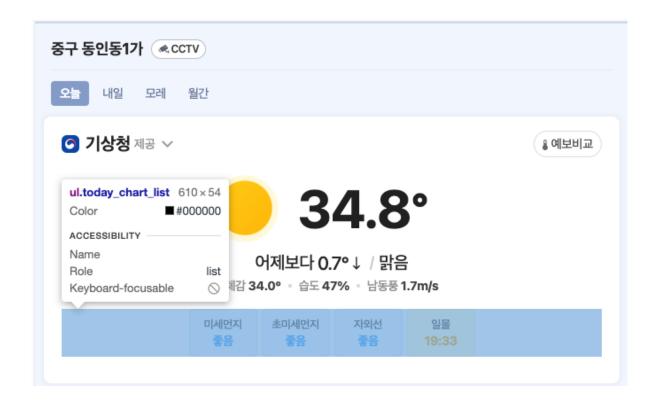
- ■현재 온도
 - <div class="temperature_text">



- ■날씨 상태
 - 맑음



- ■공기 상태
 - •
 - 가 4개 존재



```
▼ == $0
 ▼
  ▼<a nocr onclick="return go0therCR(this, 'a=nco x7t*4&r=1&i=8815
   7038_00000065035F&u=' + urlencode(this.href));" href="?where=ne
   xearch&sm=tab etc&pkid=227&gvt=0&guery=%EB%8C%80%EA%B5%AC%EA%B
   4%91...B%8F%99%EC%9D%B8%EB%8F%991%EA%B0%80%20%EB%AF%B8%EC%84%B8%E
    B%A8%BC%EC%A7%80">
     <strong class="title">미세먼지</strong>
     <span class="txt">좋음</span>
   </a>
  ▼
  ▶ <a nocr onclick="return goOtherCR(this, 'a=nco x7t*4&r=1&i=8815
   7038_00000065035F&u=' + urlencode(this.href));" href="?where=ne"
   xearch&sm=tab etc&pkid=227&qvt=0&query=%EB%8C%80%EA%B5%AC%EA%B
    4%91...C%9D%B8%EB%8F%991%EA%B0%80%20%EC%B4%88%EB%AF%B8%EC%84%B8%E
    B%A8%BC%EC%A7%80">...</a>
  ▶...
 ▶...
```

- ■시간대별 날씨
 - <div class="graph_inner _hourly_weather">
 - 하위에 많은 터그 존재



```
▼<div class="graph_inner _hourly_weather">
▼
 ▶...
 ▶...
 ▶...
 ▶...
 ▼
  ▼<dl class="graph_content">
   ▶ <dt class="time">...</dt>
   ▶ <dd class="weather box">...</dd>
   ▼<dd class="degree_point">
   ▶ <div class="inner">...</div> == $0
   </dd>
   </dl>
  ▶ ...
 ▶...
 ▶...
 <\li class="_li">...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
 ▶...
```

출력 결과

```
현재 위치: 중구 동인동1가
현재 온도: 현재 온도4.0°
날씨 상태:
        맑음
공기 상태:
미세먼지 보통
초미세먼지 나쁨
자외선 좋음
일몰 17:58
시간대별 날씨 및 온도
13시 맑음
          7°
    맑음
14시
          7°
    맑음
15시
          7°
    맑음
          7°
16시
    맑음
17시
          7°
18시 구름많음
             5°
    구름많음
             4°
19시
    구름많음
             30
20시
21시
    흐림
    흐림
22시
23시
    흐림
내일 구름많음
             1°
. . . (중간 생략)
```

8



Questions?