

빅데이터 수집시스템 개발

빅데이터 수집/변환 모듈 개발

```
from datetime import datetime
import os.path
from selenium import webdriver
from selenium.webdriver.common.by import By

# 가장 브라우저 실행
browser = webdriver.Chrome('./chromedriver.exe')

# 기상청 날씨누리 이동
browser.get('https://www.weather.go.kr/m/obs-climate/land/city-obs.do')

# 파일 디렉터리 생성
dir = "./weather/{:%Y-%m-%d}".format(datetime.now())

if not os.path.exists(dir):
    os.makedirs(dir)

# 파일 생성 및 데이터 파싱
fname = "{Weather_%Y-%m-%d%H%M.csv}".format(datetime.now())
file = open(dir+'/' + fname, 'w', encoding='utf-8')

trs = browser.find_elements(By.CSS_SELECTOR, '#weather_table > tbody > tr')
for tr in trs:
    t1 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(1) > a').text
    t2 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(2)').text
    t3 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(3)').text
    t4 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(4)').text
    t5 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(5)').text
    t6 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(6)').text
    t7 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(7)').text
    t8 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(8)').text
    t9 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(9)').text
    t10 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(10)').text
    t11 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(11)').text
    t12 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(12)').text
    t13 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(13)').text
    t14 = tr.find_element(By.CSS_SELECTOR, 'td:nth-child(14)').text

    file.write('{},{},{},{},{},{},{},{},{},{},{},{},{},{},{},\n'.format(t1, t2, t3, t4, t5, t6, t7, t8, t9, t10, t11, t12, t13, t14))

file.close()
print('데이터 수집 완료...')
```

프로그램 실행 완료 확인

기상 데이터 수집

```
C:\Users\bigdata\Desktop\Workspace\Bigdata\ver  
C:\Users\bigdata\Desktop\Workspace\Bigdata\Cra  
    browser = webdriver.Chrome('./chromedriver.e  
데이터 수집 완료...  
  
Process finished with exit code 0
```

프로그램 실행 확인

```
Bigdata C:\Users\bigdata\Desktop\Worksp  
  ▾ Crawling  
    ▾ weather  
      ▾ 2022-03-04  
        Weather_2022-03-04-13-00.csv
```

파일 생성 확인

기상 데이터 수집(생성된 파일 notepad 확인)

1번 강릉

```
1 강릉, 20 이상, , 16.3,-8.1,15.7, , 18,남남서,4.2,1010.1,  
2 강진군, 20 이상, , 12.8,5.9,12.2, , 63,남서,2.4,1014.5,  
3 강화, 10.9, , 10.5,5.8,8.9, , 73,남남서,3.4,1011.6,  
4 거제, 10.4, , 11.7,2.4,11.1, , 53,남남서,2.1,1016.6,  
5 거창, 14.8, , 9.2,-6.2,9.2, , 33,북,0.7,1014.3,  
6 경주시, 14.9, , 12.9,-6.6,12.9, , 25,동남동,0.8,1014.6,  
7 고산, 18.7, , 13.6,7.3,10.8, , 66,남남동,10.9,1015.3,  
8 고창, 20 이상, , 14.2,-1.8,12.9, , 33,남남동,4.9,1014.2,  
9 고창군, 15.9, , 14.0,-0.8,12.8, , 36,남남동,4.4,1014.2,  
10 고흥, 10.7, , 12.5,5.1,12.5, , 61,서,1.4,1016.0,  
11 광양시, 12.3, , 12.1,2.7,11.3, , 53,남서,2.5,1016.4,  
12 광주,맑음,11.0,2,0,14.2,0.7,13.8, , 40,남서,2.4,1015.0,  
13 구미, 15.6, , 11.3,-4.4,11.3, , 33,북동,0.5,1014.6,  
14 군산, 11.0, , 14.2,0.0,12.9, , 38,남남동,4.8,1014.1,  
15 금산, 8.1, , 14.2,-2.7,13.7, , 31,서남서,2.6,1012.4,  
16 김해시, 8.2, , 11.5,0.8,11.5, , 48,남남서,0.7,1015.7,  
17 남원, 7.1, , 10.2,0.0,10.0, , 49,서,1.4,1015.4,  
18 남해, 7.2, , 12.3,4.7,12.3, , 60,북서,1.3,1016.0,  
19 대관령, 20 이상, , 8.5,-8.5,5.2, , 29,서북서,6.5,1011.1,  
20 대구,구름맑음,13.7,7,0,11.4,-4.3,11.4, , 33,북서,0.8,1014.9,  
21 대전,연무,10.0,3,0,14.4,-3.9,13.4, , 28,남서,3.9,1012.9,  
22 동두천, 18.6, , 11.7,2.1,10.2, , 52,남,3.8,1011.3,  
23 동해, 20 이상, , 17.3,-13.5,17.3, , 11,서남서,2.6,1010.4,  
24 목포,구름조금,14.9,4,0,13.2,4.0,11.5, , 54,남,5.1,1014.9,
```

```
78 창원,연무,5.4,9,0,11.4,0.7,10.5, , 48,남남동,2.5,1015.3,  
79 천안, 10.4, , 11.9,-0.9,10.8, , 41,남남동,2.9,1012.6,  
80 철원, 20 이상, , 10.5,1.0,9.2, , 52,남남서,2.9,1011.1,  
81 청송군, 20 이상, , 11.4,-6.5,11.2, , 28,북서,1.5,1013.0,  
82 청주,연무,9.0,2,0,13.5,-2.9,13.1, , 32,남남서,2.3,1012.4,  
83 추풍령, 20 이상, , 10.5,-7.7,8.7, , 27,남남동,3.7,1013.6,  
84 춘천, 14.3, , 8.8,-5.1,8.8, , 37,서,1,1012.2,  
85 충주, 9.2, , 11.5,-2.3,11.5, , 38,서,0.6,1012.3,  
86 태백, 19.7, , 9.8,-14.1,7.3, , 17,남서,5.2,1010.8,  
87 통영, 4.6, , 10.8,5.3,10.3, , 69,남남동,1.8,1016.4,  
88 파주, 20 이상, , 11.4,4.5,9.9, , 63,남남서,3.6,1011.6,  
89 포항,흐림,15.4,10,0,13.6,-5.5,12.9, , 26,남서,2.9,1014.5,  
90 함양군, 16.3, , 10.8,-9.5,10.8, , 23,북서,1.1,1014.6,  
91 함천, 12.2, , 12.4,-3.4,12.4, , 33,북북서,0.9,1014.4,  
92 해남, 10.0, , 14.0,2.8,13.7, , 47,남남서,2.2,1016.1,  
93 홍성,구름맑음,20 이상,6,0,14.3,-3.5,13.3, , 29,남서,3.8,1012.2,  
94 홍천, 20 이상, , 11.1,-1.6,11.1, , 41,서남서,1.2,1012.3,  
95 흑산도,구름조금,17.2,3,0,12.9,6.9,10.0, , 67,남,10.2,1014.2,
```

마지막 95번 흑산도