

# 빅데이터 처리 및 응용 자유주제

전국 미세먼지 데이터 분석

2017204016 이지환

# 구현 기능 및 설명

## 1. 데이터 수집- 미세먼지 측정소 주소 저장하기

DUST FEEL	
전국 측정소 총 455개의 미세먼지 측정소가 있습니다.	동네지정 ▾
1 페이지 (1 - 10)	PM10높은순 PM10낮은순
충북 칠금동 측정소 12월 9일 오후 8시 갱신됨	미세먼지 - 매우나쁨 <b>169</b> $\mu\text{g}/\text{m}^3$ 초미세먼지 - 매우나쁨 <b>79</b> $\mu\text{g}/\text{m}^3$
충북 충주시 칠금동 874번지(칠금동 주민센터 옥상)(칠금 중앙로 37)칠금금릉동 주민센터 옥상 도시대기	
인천 백령도 측정소 12월 9일 오후 8시 갱신됨	미세먼지 - 매우나쁨 <b>152</b> $\mu\text{g}/\text{m}^3$ 초미세먼지 - 매우나쁨 <b>100</b> $\mu\text{g}/\text{m}^3$
인천 옹진군 백령면 연화리산241-2 국가배경농도	
충북 복대동 측정소 12월 9일 오후 8시 갱신됨	미세먼지 - 나쁨 <b>145</b> $\mu\text{g}/\text{m}^3$ 초미세먼지 - 매우나쁨 <b>93</b> $\mu\text{g}/\text{m}^3$
충북 청주시 흥덕구 복대동111 산단육거리 교차점광장 도로변대기	

- Beautifulsoup을 이용해 dustfeel 사이트에  
서 455개의 미세먼지 측정소의 주소를 크롤링  
한 뒤 리스트로 저장

# 구현 기능 및 설명

## 1. 데이터 수집-위도 경도 수집

- Chromedriver를 이용해 주소리스트에 저장된 값을 하나씩 Google map 검색창에 입력하여 url 주소에 있는 위도 경도 값을 위도 경도 리스트에 저장한다

otebooks/Desktop/빅처리%20자유%20이치환/csv1%20미세먼지%20측정소%20위도%20경도%20selenium으로%20불러와%20저장하기.ipynb

www.BANDICAM.com

jupyter csv1 미세먼지 측정소 위도 경도 selenium으로 불러와 저장하기 Last Checkpoint: 17분 전 (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

2. selenium에서 google map에 접속에 455개의 주소를 입력하고 경도 위도를 리스트에 저장한다

```
In [ ]: from selenium import webdriver
        from selenium.webdriver.common.keys import Keys #내뭉치우기 위해 필요한 모듈
        from time import sleep

        latitude=[]
        longitude=[]

        driver = webdriver.Chrome('../driver/chromedriver')
        driver.implicitly_wait(3)
        driver.get("https://www.google.co.kr/maps/")

        for i in range(len(addr_list)):
            select=driver.find_element_by_name('q')
            select.send_keys(addr_list[i])
            driver.find_element_by_xpath("//*[id='searchbox-searchbutton']").click()
            sleep(2)
            a=driver.current_url
            local=a.split('@')[1]
            latitude.append(local.split(',')[0]) #위도
            longitude.append(local.split(',')[1]) #경도
            sleep(3)
            for j in range(1, len(addr_list[i])+1):
                select.send_keys(Keys.BACKSPACE)
                #sleep(0.5)
            sleep(1)
        driver.close()
```

3. 첫 주소만 현재 나의 위치로 저장되어, 첫주소만 수정한다

```
In [ ]: driver = webdriver.Chrome('../driver/chromedriver')
        driver.implicitly_wait(3)
        driver.get("https://www.google.co.kr/maps/")
        select=driver.find_element_by_name('q')
        select.send_keys(addr_list[0])
        driver.find_element_by_xpath("//*[id='searchbox-searchbutton']").click()
        sleep(5)
```

하십시오.

오전 1:27 2019-12-09

# 구현 기능 및 설명

## 1. 데이터 수집- df1

### 4.dataframe에 저장하기

```
addr_all={'주소':addr_list, '위도':latitude, '경도':longitude}  
df=pd.DataFrame(addr_all)  
df.head()
```

	주소	위도	경도
0	경기 양주시 백석읍 꿈나무로 199꿈나무도서관	37.792424	126.9845236
1	대구 달성군 현풍면 부리 247 (상수도사업본부 달성사업소)(현풍중앙로 144-34...	35.6999534	128.4412756
2	경기 이천시 설성면 신필리산 88-5(전파연구소 입구)	37.1387735	127.5449728
3	경기 포천시 삼육사로 2186번길 11-15선단보건지소	37.8533363	127.1570327
4	충북 음성군 금왕읍 무극로370 (금왕공설운동장내)	36.9627761	127.5755361

주소에 따른 위도의 경도를 하나의 dataframe으로 저장

# 구현 기능 및 설명

## 2. 데이터 추가 및 변경

	주소	위도	경도
0	경기 양주시 백석읍 꿈나무로 199꿈나무도서관	37.792424	126.9845236
1	대구 달성군 현풍면 부리 247 (상수도사업본부 달성사업소)(현풍중앙로 144-34...	35.6999534	128.4412756

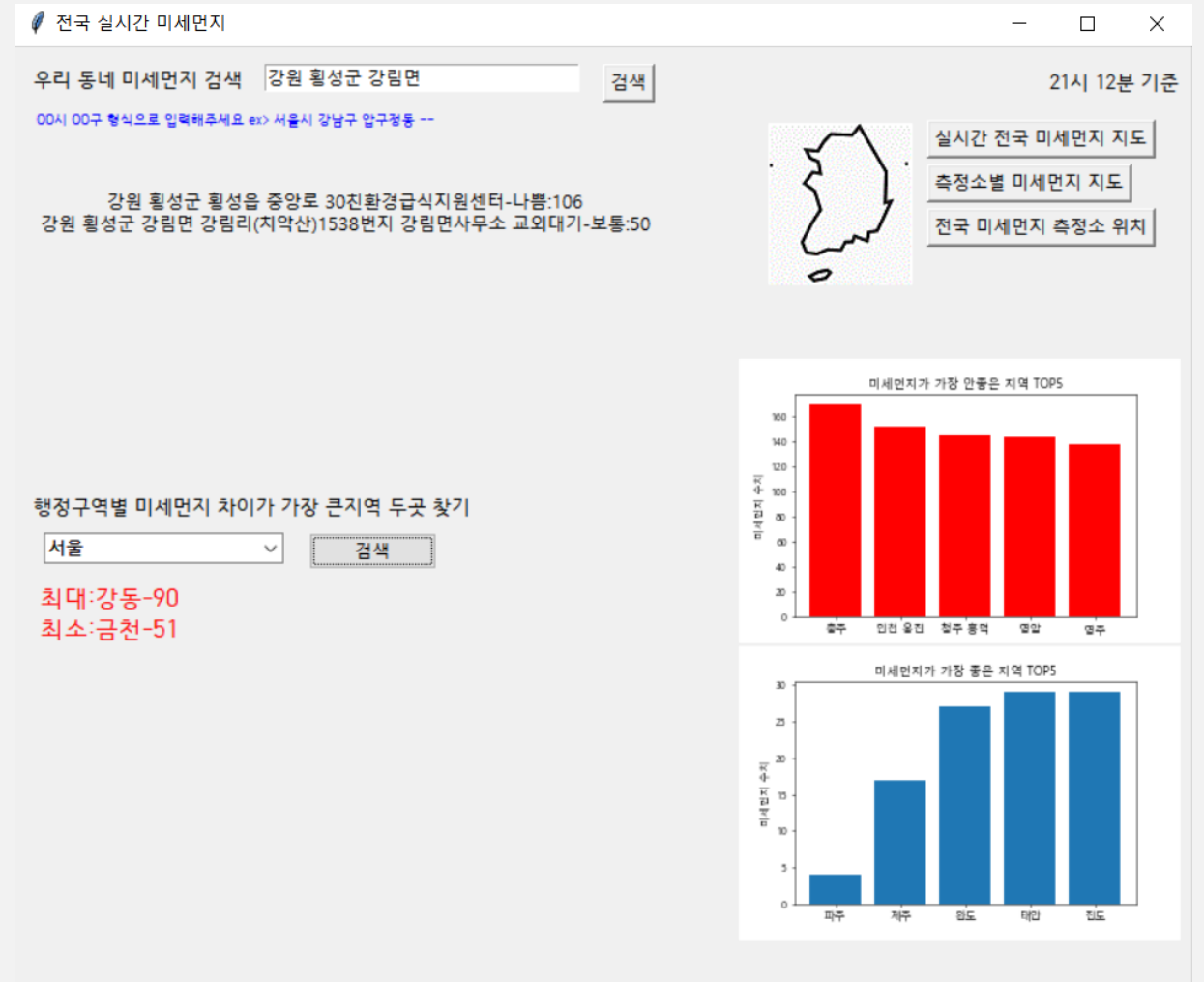
	주소	위도	경도	지역	시군	ID
0	경기 양주시 백석읍 꿈나무로 199꿈나무도서관	37.792424	126.984524	경기	양주	양주
1	대구 달성군 현풍면 부리 247 (상수도사업본부 달성사업소)(현풍중앙로 144-34...	35.699953	128.441276	대구	달성	대구 달성

지도로 시각화하기 위해서 dataframe에 지역별 고유 ID를 생성해서 column에 추가

# 구현 기능 및 설명

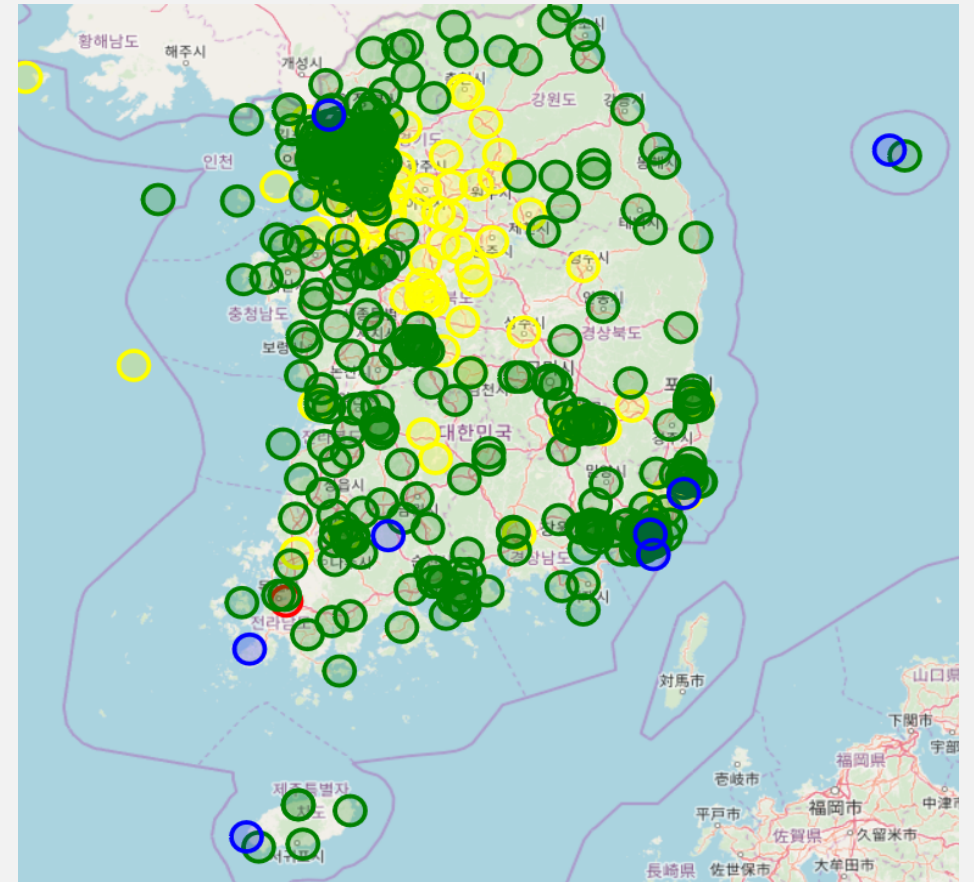
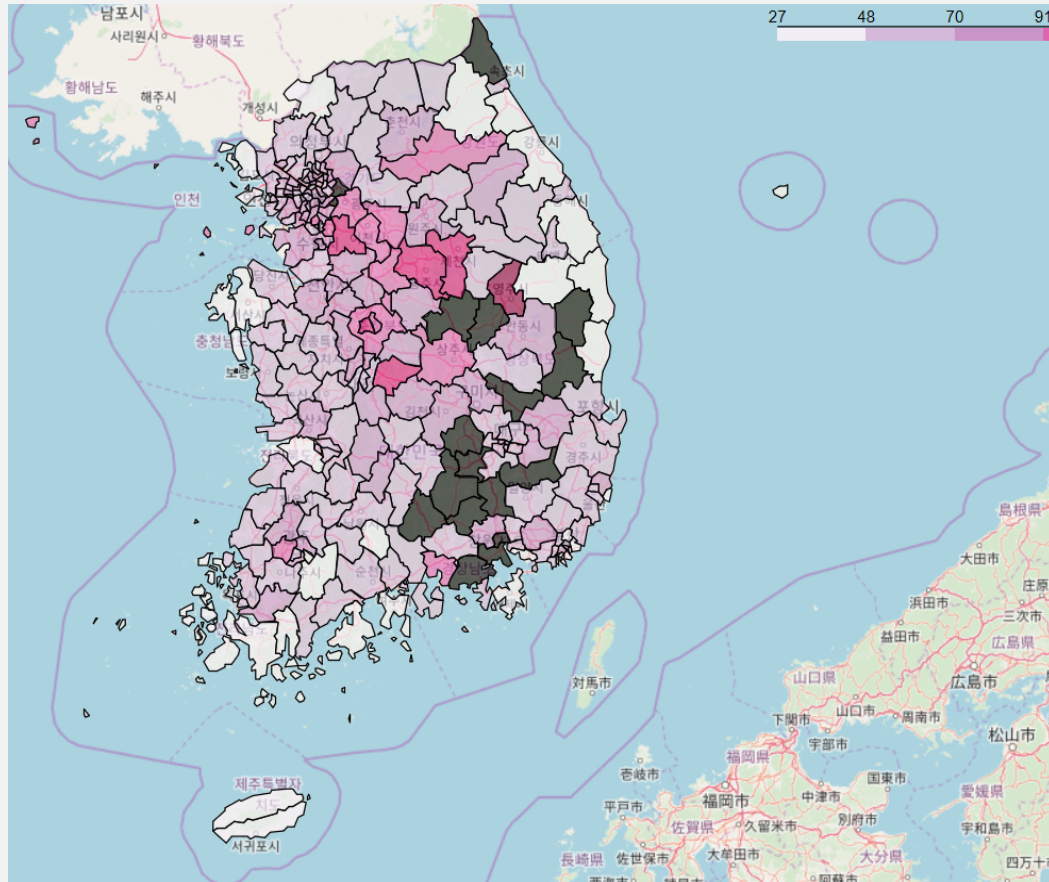
## 3. tkinter로 구현

- 실시간 미세먼지 정보를 한눈에 볼 수 있도록 tkinter로 구현
- 미세먼지 정보가 1시간마다 업데이트 되기 때문에, 스레드 주기를 1시간으로 설정함



# 구현 기능 및 설명

## 3. tkinter로 구현- 전국 미세먼지 수치 지도로 시각화



## 구현 기능 및 설명

### 3. tkinter로 구현- 행정 구역별 미세먼지 차이가 큰 지역 두 곳 찾기

행정구역별 미세먼지 차이가 가장 큰지역 두곳 찾기

충남 ▼

검색

최대:아산-89

최소:서산-32

행정구역별 미세먼지 차이가 가장 큰지역 두곳 찾기

서울 ▼

검색


최대:강동-90

최소:금천-51

동일한 행정구역이여도 미세먼지 수치가 차이가 난다는 점을 알 수 있다.



## 3. tkinter로 구현- 사용자가 검색한 위치 주변의 미세먼지 값 알려주기

 전국 실시간 미세먼지

우리 동네 미세먼지 검색

서울 영등포구 양산로23

검색

00시 00분 형식으로 입력해주세요 ex> 서울시 강남구 압구정동 --

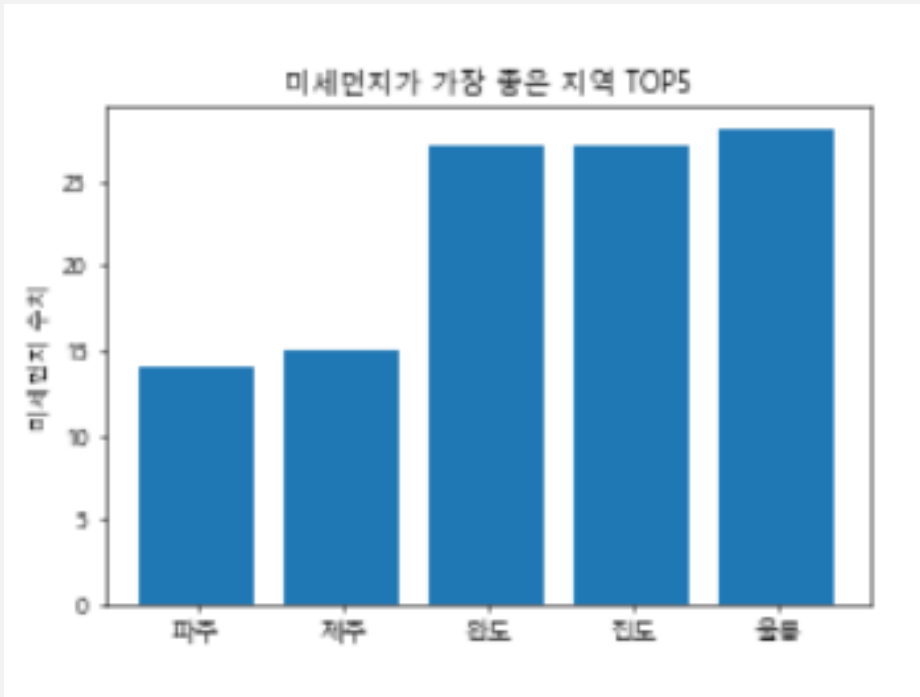
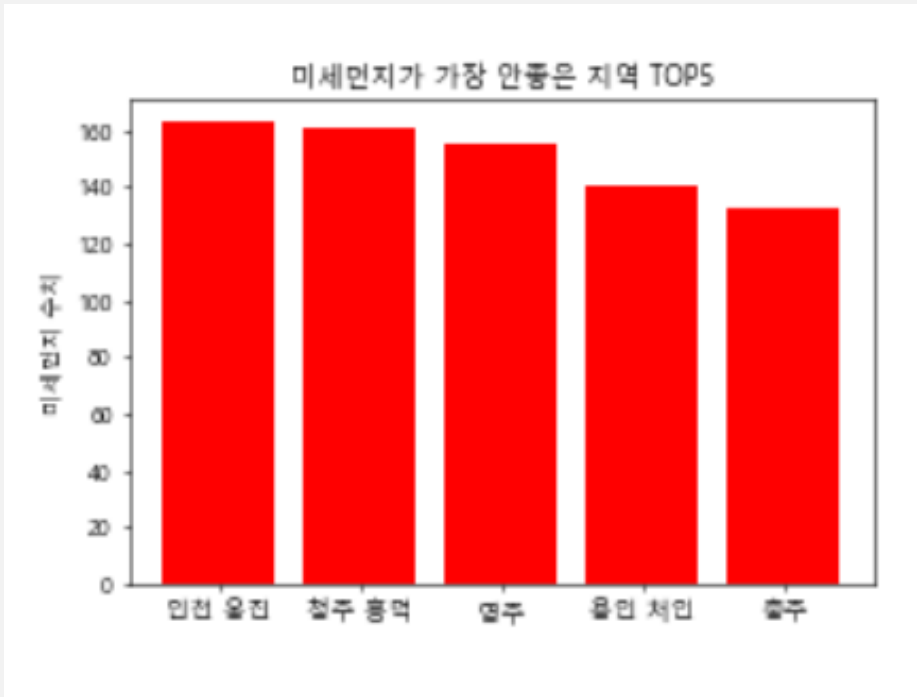
서울 영등포구 영중로 37(영등포시장사거리)-보통:80

서울 영등포구 양산로23길 11당산1동 주민센터-보통:56

해당 지역 주변 측정소들의 미세먼지 값들을 한눈에 알 수 있다

# 구현 기능 및 설명

## 3. tkinter로 구현- 미세먼지가 가장 좋은지역/ 안 좋은지역 상위 5개 지역 뽑아내기



jupyter main code 이지환 Last Checkpoint: 3시간 전 (unsaved changes)



Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted

Python 3

Run Code

```
warnings.filterwarnings('ignore')
```

```
In [2]: window=Tk()

class startDust:

    def __init__(self,window):
        self.window=window

        self.window.title('전국 실시간 미세먼지')
        self.window.geometry('800x640')

        self.display()

    def display(self):

        self.label1=Label(self.window,text='우리 동네 미세먼지 검색',font=('함초롬돋움',10))
        self.label1.place(x=10,y=10)

        self.entry1=Entry(self.window,width=30)
        self.entry1.place(x=170,y=12)

        self.button1=Button(self.window,text='검색',command=self.click1)
        self.button1.place(x=400,y=12)

        info=Label(self.window,text='00시 00분 형식으로 입력해주세요 ex> 서울시 강남구 압구정동 ~~',font=('함초롬돋움',7),fg='blue')
        info.place(x=12,y=40)

        self.var2=StringVar()
        self.txt=Label(self.window,textvariable=self.var2)
        self.txt.place(x=15,y=80)

        self.now=datetime.now()
        self.time='%s시 %s분 기준'%(self.now.hour,self.now.minute)
        self.now_time=Label(self.window,text=self.time,font=('함초롬돋움',10))
        self.now_time.place(x=700,y=12)
```

여기엔 여기에 입력하십시오.