

2022~2024 여름철 시도별

평균 기온에 관한 분석

대구가 정말 가장 더울까?

주제

기상청 폭염 데이터를 활용하여 2022~2024년까지의
시도별 평균 기온을 분석하여 대비책을 마련하는 것

선정 배경 및 분석 범위

대구는 '대프리카'로 불릴 만큼 여름철 더위가 심각한 지역입니다.

대구가 정말로 가장 더운 지역인지, 그렇다면 그 이유는 무엇인지
분석하고자 하며 원본 데이터인 2019~2024년 전국 각 지역의 폭염
데이터에서 일부를 추출하여 최근 3개년 간인 2022~2024년까지의
특별시와 광역시 폭염 관련 데이터를 분석하고자 합니다.

데이터 관련 설명

	A	B	C	D	E	F	G	H	I	J	K	L	
1	일시	지점	폭염여부(O/X)	최고체감온도	최고기온(°C)	평균기온(°C)	최저기온(°C)	평균상대습도	폭염특보(O/X)	폭염영향예보	열대야(O/X)	자외선지수(단계)	
2	2023.8.1	북강릉(104)	X	31.8	31.4	27	22.8	78.6	O	경고	O	낮음	
3	2023.8.1	동해(106)	X	30.6	30	26.2	22.8	86.5	O	주의	O	낮음	
4	2023.8.1	서울(108)	O	34.6	34.2	29.9	25.5	71.6	O	경고	O	매우높음	
5	2023.8.1	인천(112)	O	33.8	32.8	29.3	26.5	75.4	O	경고	O	높음	
6	2023.8.1	원주(114)	O	33.4	33.7	28.1	23	73.6	O	경고	X	낮음	
7	2023.8.1	울릉도(115)	O	33	31	27.6	24.7	85.5	O	주의	O	매우높음	
8	2023.8.1	수원(119)	O	34.5	34.4	29.2	24.8	73.3	O	경고	X	매우높음	
9	2023.8.1	영월(121)	O	33.2	33.2	27.4	23.5	77.3	O	주의	X	낮음	
10	2023.8.1	충주(127)	O	34.4	33.6	27.2	22.9	82.8	O	경고	X	매우높음	
11	2023.8.1	서산(129)	O	34.8	34.8	29.1	24.6	77.4	O	경고	X	매우높음	
12	2023.8.1	울진(130)	X	30.6	29.6	25.9	22.2	82.3	O	경고	O	매우높음	
13	2023.8.1	청주(131)	O	35	35.7	30.4	25.3	66.8	O	경고	O	매우높음	
14	2023.8.1	대전(133)	O	34.6	34.6	29.5	24.8	70.4	O	경고	O	매우높음	
15	2023.8.1	추풍령(135)	X	32.9	32.3	26.7	22.6	80.5	O	경고	X	매우높음	
16	2023.8.1	안동(136)	O	34.5	33.9	28.7	23.9	78.3	O	경고	X	매우높음	
17	2023.8.1	상주(137)	O	35	34.2	28.8	24.1	80.9	O	경고	X	매우높음	
18	2023.8.1	포항(138)	O	33	32.7	28.5	24.2	71	O	경고	O	매우높음	
19	2023.8.1	군산(140)	O	35	35	29.3	24.9	77.5	O	경고	O	매우높음	
20	2023.8.1	대구(143)	O	33.9	34.8	29.7	24.9	64.8	O	경고	X	매우높음	
21	2023.8.1	전주(146)	O	35	34.9	29.6	24.9	74	O	경고	O	매우높음	
22	2023.8.1	울산(152)	X	32.8	32.9	28.2	23.6	68.5	O	경고	X	매우높음	
23	2023.8.1	창원(155)	O	33.8	33.8	29	24.7	66.5	O	경고	X	매우높음	
24	2023.8.1	광주(156)	O	35.3	35.1	29.9	25.2	82.1	O	경고	O	매우높음	
25	2023.8.1	부산(159)	O	33.3	33.2	29.6	26.8	62.6	O	경고	O	매우높음	
26	2023.8.1	통영(162)	O	33.4	33	28.5	25.4	74.6	O	경고	O	매우높음	
27	2023.8.1	목포(165)	O	34.4	34	29	25.5	77.8	O	경고	O	매우높음	
28	2023.8.1	여수(168)	X	32.8	31.6	28.5	25.5	74.5	O	경고	O	매우높음	

출처 : 기상청 기상자료개방포털

데이터 파일명 : HW2205.csv ~ HW2407.csv

컬럼 종류

수치형 데이터: 최고체감온도, 최고기온, 평균기온, 최저기온, 평균상대습도

범주형 데이터: 폭염여부(O/X), 폭염특보(O/X), 폭염영향예보(없음/관심/주의/경고), 자외선지수(낮음/보통/높음/매우높음)

데이터 확인 및 전처리 과정

- 1. 첫 번째 줄이 컬럼명인지 확인
- 2. 불필요한 칼럼 및 행 확인
- 3. 결측치 확인
- 4. 이상치 처리
- 5. 중복 데이터 처리

(1) 첫 번째 줄이 컬럼명 맞는지 확인

HW2206

첫 번째 줄이 컬럼명이 맞다.

✓ 0.0s

Python

	일시	지점	폭염여부 (O/X)	최고체감온도 (°C)	최고기온 (°C)	평균기온 (°C)	최저기온 (°C)	평균상대습도 (%)	폭염특보 (O/X)	폭염영향예보 (단계)	열대야 (O/X)	자외선지수 (단계)
0	2022.6.1	북강릉 (104)	X	25.4	28.7	21.5	15.7	37.3	X		X	높음
1	2022.6.1	동해(106)	X	25.1	28.0	20.8	16.2	46.6	X		X	높음
2	2022.6.1	서울(108)	X	26.1	28.6	22.1	16.2	41.6	X		X	매우높음
3	2022.6.1	인천(112)	X	23.7	25.4	20.4	16.5	48.5	X		X	매우높음
4	2022.6.1	원주(114)	X	25.5	29.2	21.4	12.8	35.0	X		X	매우높음
...
4855	2022.6.30	북춘천 (93)	X	27.5	24.2	21.8	20.2	99.9	X		X	보통
4856	2022.6.30	하동(932)	X	32.4	31.2	26.9	24.0	85.0	X	관심	X	높음
4857	2022.6.30	철원(95)	X	27.1	24.0	21.5	19.9	98.5	X		X	보통
4858	2022.6.30	동두천 (98)	X	27.0	23.7	21.6	20.1	100.0	X		X	보통
4859	2022.6.30	파주(99)	X	26.6	23.3	21.6	20.7	99.8	X		X	보통

4860 rows x 12 columns

데이터 확인 및 전처리 과정

1. 첫 번째 줄이 컬럼명인지 확인

2. 불필요한 칼럼 및 행 확인

3. 결측치 확인

4. 이상치 처리

5. 중복 데이터 처리

- (2) 불필요한 컬럼이나 행이 있는지 확인

지역이 너무 세분화되어있어 특별시와 광역시만 추출

```
# 특정 지역의 데이터만 뽑아주는 함수
def make_city_df(df, city):
    # 지점 컬럼에서 지역명만 빼기
    list_data = []

    for i in df['지점']:
        list_data.append(i.split('(')[0])

    df['지점'] = list_data
    return df[df['지점'] == city]

def concat_data(city):
    data = pd.concat([make_city_df(HW2206, city), make_city_df(HW2207, city), make_city_df(HW2208, city),
                        make_city_df(HW2306, city), make_city_df(HW2307, city), make_city_df(HW2308, city),
                        make_city_df(HW2406, city), make_city_df(HW2407, city)], ignore_index = True)
    return data
```

✓ 0.0s

Python

```
seoul = concat_data('서울')
daegu = concat_data('대구')
busan = concat_data('부산')
incheon = concat_data('인천')
daejeon = concat_data('대전')
gwangju = concat_data('광주')
ulsan = concat_data('울산')
```

✓ 0.0s

Python

데이터 확인 및 전처리 과정

- 1. 첫 번째 줄이 컬럼명인지 확인
- 2. 불필요한 칼럼 및 행 확인
- 3. 결측치 확인
- 4. 이상치 처리
- 5. 중복 데이터 처리

- (1) 결측치 확인

```
# 결측치 확인하는 함수 생성
def city_na(city):
    return city.isna().sum()[city.isna().sum() >= 1]

city_list = [seoul, daegu, busan, incheon, daejeon, gwangju, ulsan]
city_name = ['서울', '대구', '부산', '인천', '대전', '광주', '울산']

for i in range(len(city_list)):
    print(f'[{city_name[i]} 결측치]\n{city_na(city_list[i])}\n')

✓ 0.0s Python
```

[서울 결측치]
최고기온(°C) 1
평균기온(°C) 1
최저기온(°C) 1
dtype: int64

[대구 결측치]
최고기온(°C) 1
평균기온(°C) 1
dtype: int64

[부산 결측치]
최고기온(°C) 1
평균기온(°C) 1
dtype: int64

[인천 결측치]
최고기온(°C) 1
평균기온(°C) 1
dtype: int64

[대전 결측치]
평균기온(°C) 1
최저기온(°C) 1
dtype: int64

[광주 결측치]
최고체감온도(°C) 1
최고기온(°C) 1
평균기온(°C) 1
dtype: int64

[울산 결측치]
최고기온(°C) 1
평균기온(°C) 1
dtype: int64

데이터 확인 및 전처리 과정

1. 첫 번째 줄이 컬럼명인지 확인
2. 불필요한 칼럼 및 행 확인
3. 결측치 확인
4. 이상치 처리
5. 중복 데이터 처리

```
# 결측치 평균으로 대체하는 함수 생성
def na_fill(city):
    city['최고체감온도(°C)'] = city['최고체감온도(°C)'].fillna(city['최고체감온도(°C)'].mean())
    city['최고기온(°C)'] = city['최고기온(°C)'].fillna(city['최고기온(°C)'].mean())
    city['평균기온(°C)'] = city['평균기온(°C)'].fillna(city['평균기온(°C)'].mean())
    city['최저기온(°C)'] = city['최저기온(°C)'].fillna(city['최저기온(°C)'].mean())
```

✓ 0.0s

Python

```
# 결측치 평균으로 대체
for i in city_list:
    na_fill(i)

for i in range(len(city_list)):
    print(f'[{city_name[i]} 결측치]\n{city_na(city_list[i])}\n')
# 결측치가 모두 처리되어 빈 시리즈 출력
```

✓ 0.0s

Python

```
[서울 결측치]
Series([], dtype: int64)

[대구 결측치]
Series([], dtype: int64)

[부산 결측치]
Series([], dtype: int64)

[인천 결측치]
Series([], dtype: int64)

[대전 결측치]
Series([], dtype: int64)

[광주 결측치]
Series([], dtype: int64)

[울산 결측치]
Series([], dtype: int64)
```


데이터 확인 및 전처리 과정

1. 첫 번째 줄이 컬럼명인지 확인
2. 불필요한 칼럼 및 행 확인
3. 결측치 확인
4. 이상치 처리
5. 중복 데이터 처리

- (2) 이상치 처리

```
fig, axes = plt.subplots(5, 1, figsize = (20, 20))

for j in range(len(temp_name)):
    best_temp_list = []

    for i in city_list:
        best_temp_list.append(i[temp_name[j]])

    axes[j].boxplot(best_temp_list, tick_labels = city_name)
    axes[j].set_title(f'지역별 {temp_name[j]} 분포')
    axes[j].set_xlabel('지역')
    if j == 4:
        axes[j].set_ylabel('습도(%)')
    else:
        axes[j].set_ylabel('기온(°C)')

plt.rcParams['font.family'] = 'AppleGothic'
plt.tight_layout()
plt.show()
```

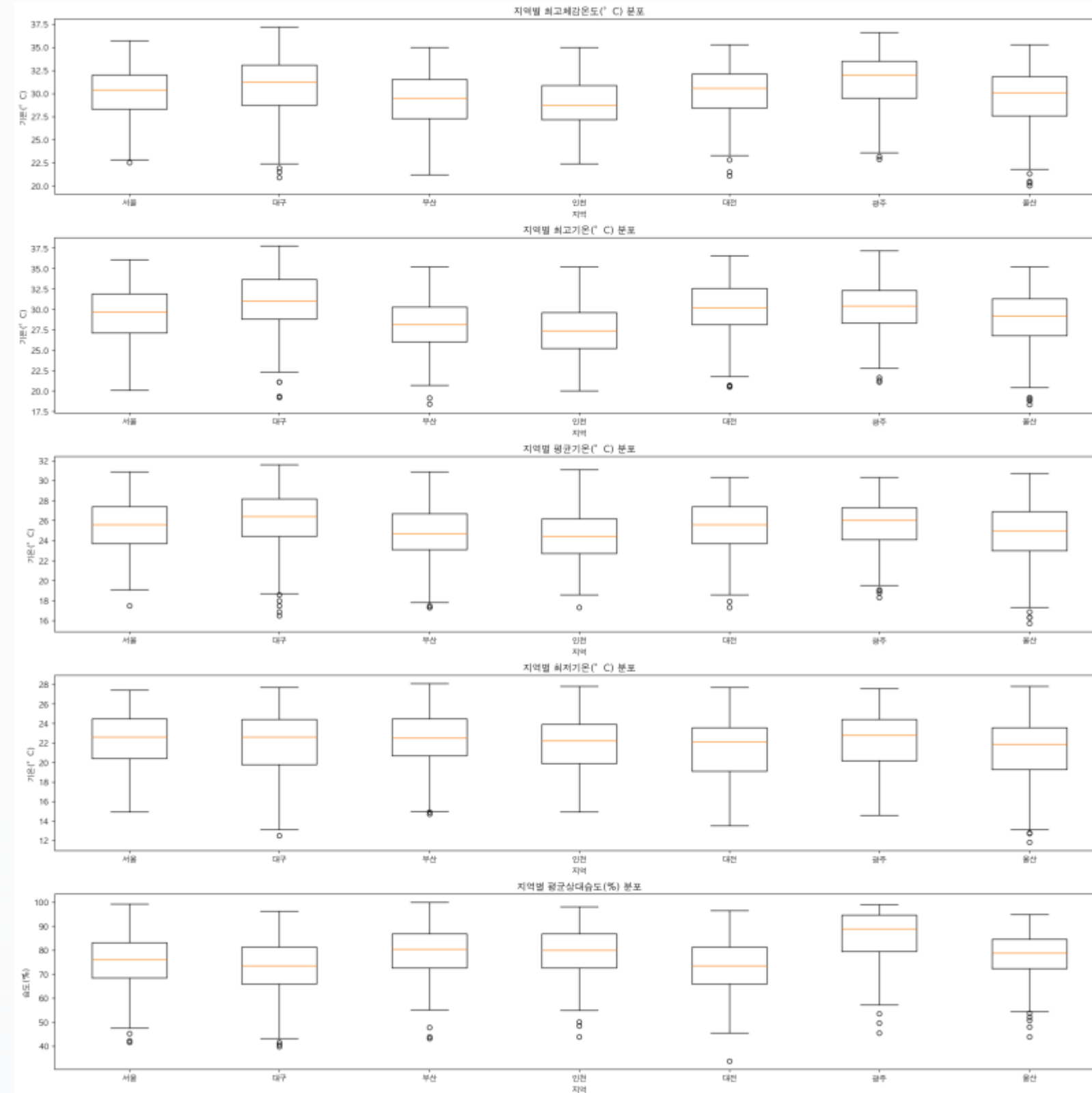
✓ 0.5s

Python

도시별 수치형 데이터의 이상치를 Box Plot으로 확인하기

데이터 확인 및 전처리 과정

1. 첫 번째 줄이 컬럼명인지 확인
2. 불필요한 칼럼 및 행 확인
3. 결측치 확인
4. 이상치 처리
5. 중복 데이터 처리



각 지역별로 이상치가 존재하긴 하지만 크게 벗어나는 온도는 아니며 실제 온도 데이터라고 판단되어 제거하지 않고 남기기로 결정

데이터 확인 및 전처리 과정

1. 첫 번째 줄이 컬럼명인지 확인
2. 불필요한 칼럼 및 행 확인
3. 결측치 확인
4. 이상치 처리
5. 중복 데이터 처리

- (3) 중복 데이터 처리

```
city_list = [seoul, daegu, busan, incheon, daejeon, gwangju, ulsan]

for i in range(len(city_list)):
    print(f'{city_name[i]} 중복 데이터 수 : {city_list[i].duplicated().sum()}')

# 중복 데이터 없음
```

```
✓ 0.0s
서울 중복 데이터 수 : 0
대구 중복 데이터 수 : 0
부산 중복 데이터 수 : 0
인천 중복 데이터 수 : 0
대전 중복 데이터 수 : 0
광주 중복 데이터 수 : 0
울산 중복 데이터 수 : 0
```

Python

모든 지역에서 중복 데이터는 발견되지 않음

데이터 가공

1. 폭염영향예보(단계) 컬럼 가공

- (1) 폭염영향예보(단계) 공백 값 없음으로 채워넣기

+ 코드

+ Markdown

```
seoul['폭염영향예보(단계)'].value_counts()
```

✓ 0.0s

python

폭염영향예보(단계)

101

관심 70

주의 46

경고 20

Name: count, dtype: int64

```
city_list = [seoul, daegu, busan, incheon, daejeon, gwangju, ulsan]
```

```
for i in city_list:
```

```
i['폭염영향예보(단계)'] = i['폭염영향예보(단계)'].replace(' ', '없음')
```

✓ 0.0s

python

```
seoul['폭염영향예보(단계)'].value_counts()
```

✓ 0.0s

python

폭염영향예보(단계)

없음 101

관심 70

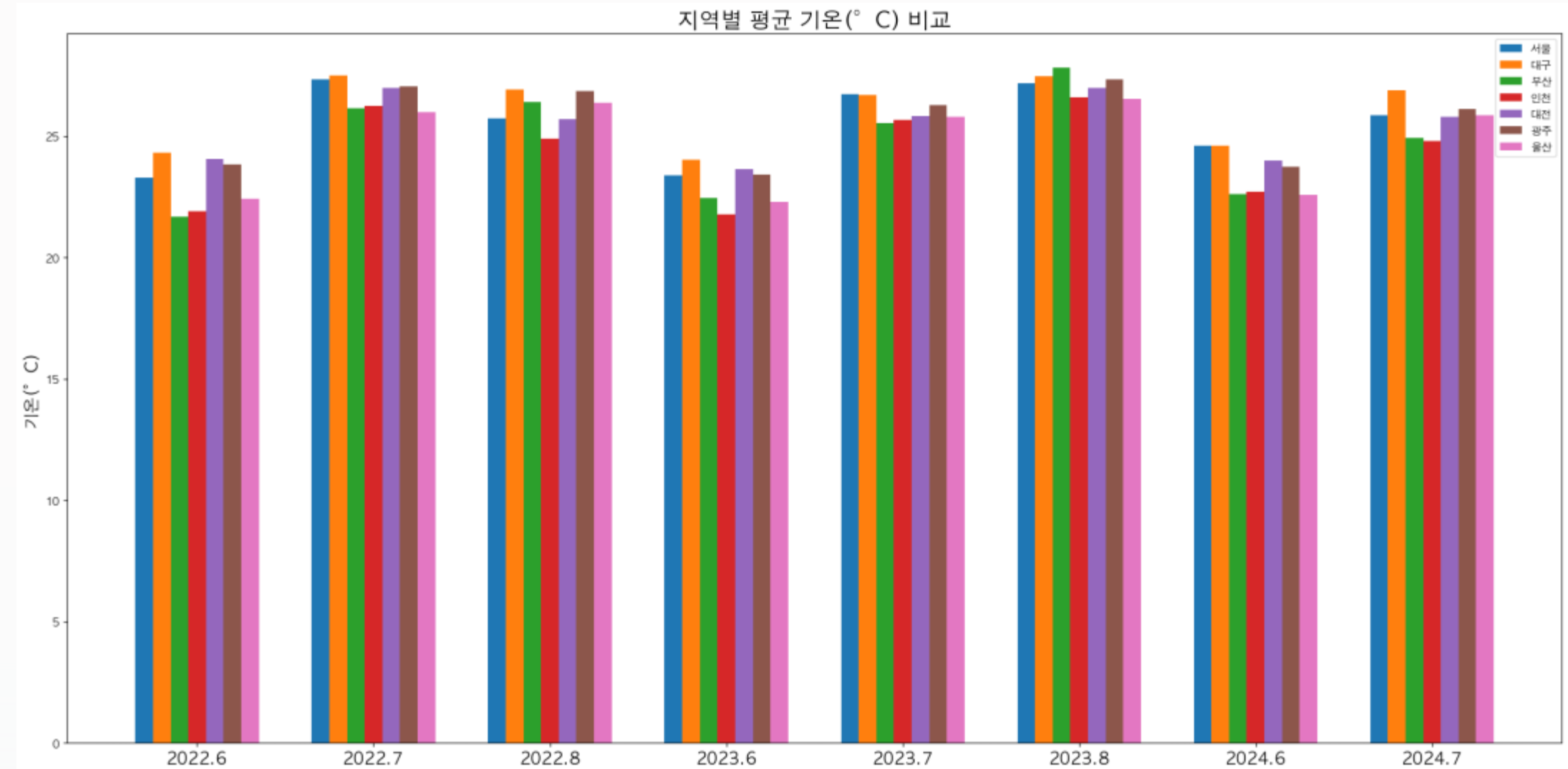
주의 46

경고 20

Name: count, dtype: int64

데이터 분석

1. 연도별 지역에 따른 평균 기온 비교
2. 열대야 발생 여부에 따른 평균 기온 비교
3. 습도와 체감 온도 간의 관계 분석
4. 지역별 폭염영향예보(단계) 비율 비교



- > 전체적으로 대구가 높은 기온을 보였다.
- > 부산의 경우 전체적으로 낮은 기온을 보였지만 유독 2023.8에 높은 기온을 보였다.

데이터 분석

1. 연도별 지역에 따른 평균 기온 비교
2. 열대야 발생 여부에 따른 평균 기온 비교
3. 습도와 체감 온도 간의 관계 분석
4. 지역별 폭염영향예보(단계) 비율 비교

2023년 부산 역대 최고로 더웠다...연평균기온 16.0도

입력:2024-01-16 11:16:35 | 수정:2024-01-16 16:25:55



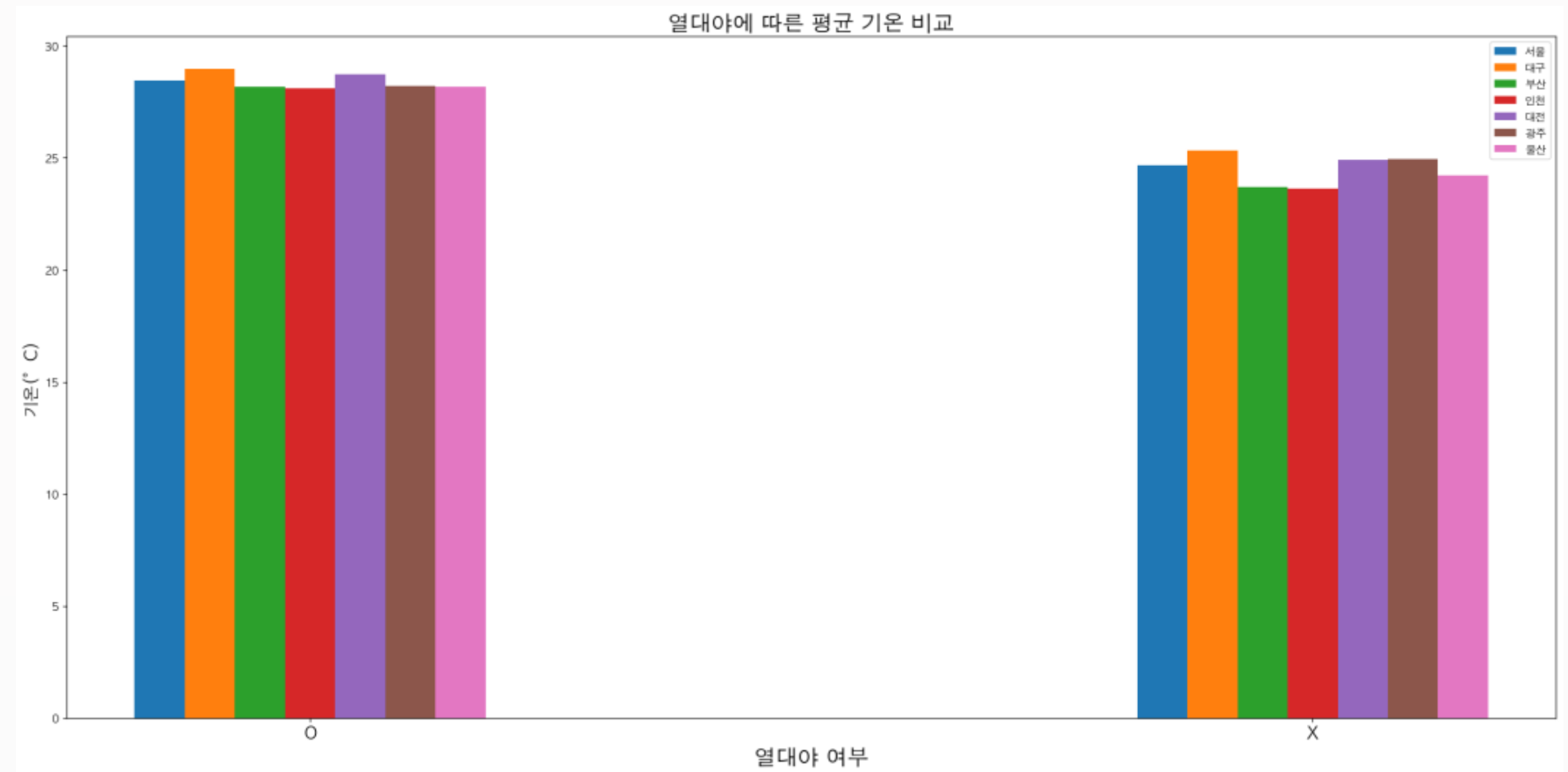
부산기상청은 “북태평양을 비롯해 전반적으로 우리나라 동쪽에서 고기압성 흐름이 발달한 만큼 남풍계열의 따뜻한 바람이 자주 불어 기온이 높은 날이 많았다”고 설명했다.

지난해는 이례적으로 기온 변동이 컸던 해이기도 했다. 겨울철인 1월과 11월, 12월의 기온 변동이 두드러졌다. 따뜻한 이동성고기압 영향을 받은 후 시베리아에서 기압능이 발달하면서 동시에 북동아시아에 남북 흐름이 강화되며 북극 주변 찬 공기가 우리나라로 유입되는 패턴이 반복됐다.

<https://www.busan.com/view/busan/view.php?code=2024011610474318709>

데이터 분석

1. 연도별 지역에 따른 평균 기온 비교
2. 열대야 발생 여부에 따른 평균 기온 비교
3. 습도와 체감 온도 간의 관계 분석
4. 지역별 폭염영향예보(단계) 비율 비교

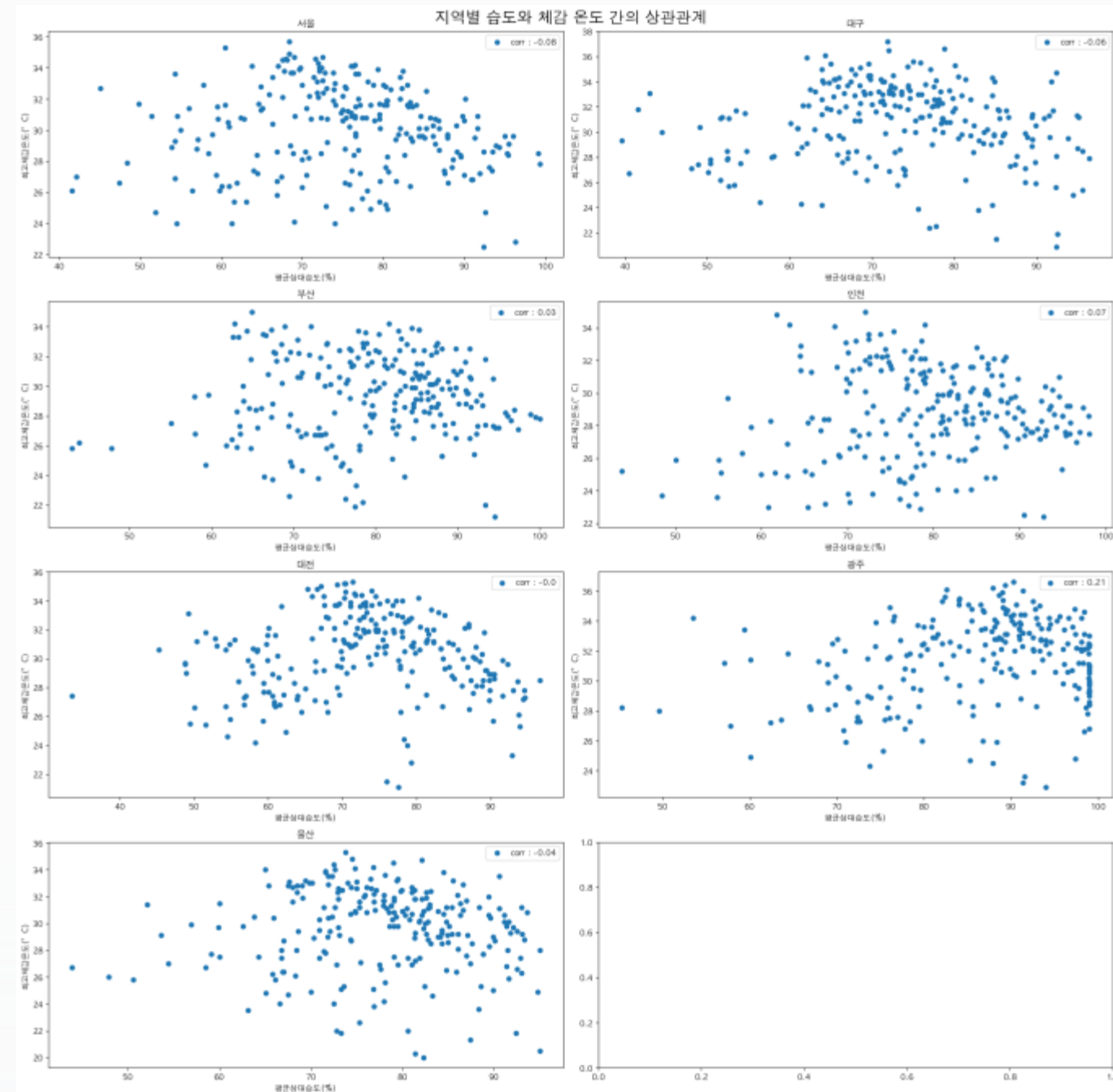


- > 모든 지역에서 열대야가 있을때가 없을때보다 평균 기온이 높았다.
- > 대구가 열대야 여부에 관계없이 가장 기온이 높다.
- > 바다가 있는 부산, 인천, 울산은 비교적 온도가 낮다.

데이터 분석

1. 연도별 지역에 따른 평균 기온 비교
2. 열대야 발생 여부에 따른 평균 기온 비교
3. 습도와 체감 온도 간의 관계 분석
4. 지역별 폭염영향예보(단계) 비율 비교

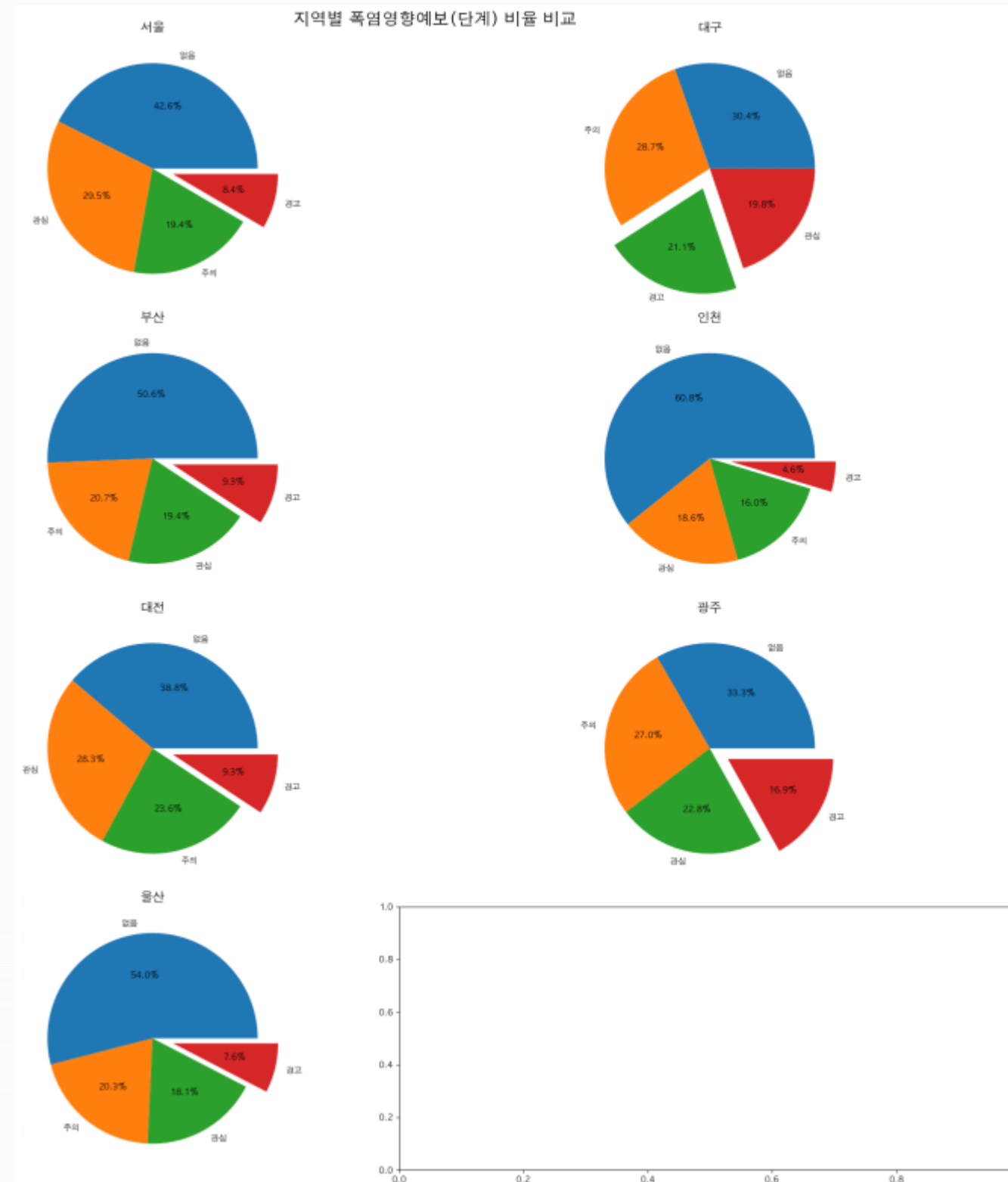
습도가 높을수록 불쾌 지수가 높아져 심리적 요소인 체감 온도도 높아질 것으로 생각하여 두 변수 간에 관계가 있을 것으로 생각하고 분석 진행



> 습도와 체감 온도 간의 관계가 있을 것으로 생각했지만 큰 상관관계는 발견하지 못하였다.
 > 아마도 상관분석으로는 관계를 파악하기 힘들고 평균온도와 상대습도를 비롯한 여러 변수를 설명변수로 두고 체감온도를 반응변수로 둔 회귀분석을 진행해서 인과관계를 파악해야 할 것으로 생각된다.

데이터 분석

1. 연도별 지역에 따른 평균 기온 비교
2. 열대야 발생 여부에 따른 평균 기온 비교
3. 습도와 체감 온도 간의 관계 분석
4. 지역별 폭염영향예보(단계) 비율 비교



> 대구랑 광주가 폭염영향예보 경고 비율이 각각 21.1%, 16.9%로 높게 나타났다.

데이터 분석

- 1. 연도별 지역에 따른 평균 기온 비교
- 2. 열대야 발생 여부에 따른 평균 기온 비교
- 3. 습도와 체감 온도 간의 관계 분석
- 4. 지역별 폭염영향예보(단계) 비율 비교



<https://news.kbs.co.kr/news/pc/view/view.do?ncd=7991718>

뉴스7(광주)

[이슈대담] 폭염에 손 잡은 ‘광프리카×대프리카’

입력 2024.06.19 (19:25) | 수정 2024.06.19 (20:41)



앵커

'대프리카'라는 말 혹시 들어보셨습니까?

우리나라에서 가장 더운 지역으로 꼽히는 대구의 여름을 뜨거운 아프리카 대륙에 빗댄 말인데요.

몇 년 전부터는 광주의 더위도 만만치 않다고 해서 '광프리카'라는 신조어까지 나왔습니다.

답변

광주하고 대구가 똑같이 내륙에 위치한 분지형 도시입니다.

분지형 지형이 열기에 갇히기 때문에 외부로 못 나갑니다.

그러다 보니까 더 덥게 느껴질 수가 있는데요.

우리가 대구와 광주의 가장 큰 차이는 습도 문제입니다.

광주가 대구보다 한 10% 정도 습도가 높습니다.

그러다 보니까 대구가 온도는 높은 반면 체감 온도는 광주가 훨씬 더 높게 발생한다고 보시면 될 것 같습니다.

기사를 보고 광주의 체감 온도가 높은지 분석해보고 싶었지만
시간이 부족해 진행하지 못하여 아쉬움을 남김

분석 결과

> 다양한 비교 분석에서 대구가 더운 곳으로 판명되었다.

[아프리카만큼 더운 대구, 대프리카]

대구광역시의 엠블럼은 산이 세 개 겹쳐 있는 모양이다. 시가지의 남과 북에 높은 산이 둘러싸고 있는 분지 지형을 형상화한 것이다. 이처럼 북쪽에는 팔공산, 남쪽에는 앞산과 비슬산이 가로막고 있다 보니 더운 공기가 빠져나가기 힘든 구조이다. 사방에서 불어오는 바람은 대구[대구광역시]를 둘러싼 산을 좀처럼 넘지 못하는데 겨우 넘어온 고온건조한 공기는 대구 분지에 계속해서 쌓여 찜통더위의 원인으로 작용한다. 게다가 인구밀도와 도시화 비율이 높은 광역시여서 지형적·인구학적 원인으로 열섬현상까지 나타나는 실정이다. 최근 들어 여름철 대구에는 순간적으로 천둥과 번개가 치며 국지성 호우가 쏟아졌다가 금세 개는 현상이 나타나 아열대 기후의 조짐이 아니냐는 반응도 나오고 있다.

종합 정리



<싸이 흠뻑쇼>

> 지형적 특징으로 인해 대구의 기온을 낮추기엔 한계가 있을 것으로 생각된다. 실제로도 피할 수 없다면 즐기자는 역발상으로 대구지역의 폭염을 활용한 물총 축제, 호러연극 축제 등과 같은 사업을 추진 중이라고 한다. 앞으로도 흠뻑쇼 등을 비롯한 더위를 식힐 수 있는 행사가 많아져 대구 시민들이 여름철 더위를 잘 이겼으면 좋겠다.



<대구워터페스티벌>

감사합니다.

A simple presentation template suitable for
both long and short articles