

1. 분석 목적

- 공공자전거 신규 가입자 정보를 일별로 분석하여 신규가입과 날씨와 연관성이 있는지 확인

2. 분석 개요

- 분석 과제 : 공공자전거의 신규 가입자 수와 날씨와의 연관성
- 분석 데이터
 - 서울 열린데이터 광장(<http://data.seoul.go.kr/>)의 서울특별시 공공자전거 신규가입자 정보(일별) (open API, excel 파일)
 - 기상청 기상자료개방포털(<https://data.kma.go.kr/cmmn/main.do>)의 기상청 기상관측자료(종관_일자료) (open API 파일)
- 분석 도구 : Jupyter Notebook(python)
- 분석 내용 :
 - 공공자전거 신규 가입자 수 분석
 - 날씨 데이터 최저/최고 기온, 평균 전운량/풍속, 강수량 분석

3. 분석 과정

사람들이 자전거를 타고 싶을 때는 날씨가 맑고 너무 덥지도 않고 춥지도 않은 날에 타고 싶다고 생각하여 날씨가 좋은 날에 자전거를 타고 싶은 사람들이 많이 가입한다고 생각했다.

공공자전거 날짜별 신규 가입 데이터 불러오기

```
In [1]: from selenium import webdriver
        from bs4 import BeautifulSoup
        import time
        import codecs
        import pandas as pd
```

- rent_dt : 대여 일자
- member_cd : 사용자 코드
- sex_cd : 성별
- year_cd : 연령대 코드
- mem_cnt : 신규 가입자 수

```
In [2]: data_list = {}
driver = webdriver.Chrome("chromedriver.exe")
startcnt = 1
endcnt = 1000

while endcnt <= 9000:
    time.sleep(1)

    driver.get(
        'http://openapi.seoul.go.kr:8088/57454f474b67677837337173794a4f/xml/cycleNewMemberRentInfoDay/'+str(startcnt)+'/'+str(endcnt)+'/')

    source = driver.page_source

    s1 = BeautifulSoup(source, "html.parser")

    find_list = ['rent_dt', 'member_cd', 'sex_cd', 'year_cd', 'mem_cnt']

    for each in find_list:
        find = s1.find_all(each)
        #data_list.append(find)
        if data_list.get(each) is None:
            data_list[each]=[x.text for x in find]
        else:
            data_list[each]=data_list[each]+[x.text for x in find]

    startcnt += 1000
    endcnt += 1000
```

```
In [3]: data = pd.DataFrame(data_list)
```

```
In [4]: data
```

Out[4]:

| | rent_dt | member_cd | sex_cd | year_cd | mem_cnt |
|------|------------|-----------|--------|---------|---------|
| 0 | 2018-07-19 | 회원-내국인 | | 70대~ | 2194 |
| 1 | 2018-07-18 | 회원-내국인 | | 70대~ | 2123 |
| 2 | 2018-07-17 | 회원-내국인 | | 70대~ | 2133 |
| 3 | 2018-07-16 | 회원-내국인 | | 70대~ | 2161 |
| 4 | 2018-07-15 | 회원-내국인 | | 70대~ | 2937 |
| ... | ... | ... | ... | ... | ... |
| 8090 | 2017-01-01 | 회원-내국인 | F | ~10대 | 3 |
| 8091 | 2017-01-01 | 회원-내국인 | F | 50대 | 5 |
| 8092 | 2017-01-01 | 회원-내국인 | F | 40대 | 11 |
| 8093 | 2017-01-01 | 회원-내국인 | F | 30대 | 16 |
| 8094 | 2017-01-01 | 회원-내국인 | F | 20대 | 53 |

8095 rows × 5 columns

```
In [5]: data.rename(columns = {"rent_dt": "가입일자"}, inplace = True)
```

```
In [6]: data.rename(columns = {"member_cd": "사용자코드"}, inplace = True)
data.rename(columns = {"sex_cd": "성별"}, inplace = True)
data.rename(columns = {"year_cd": "연령대코드"}, inplace = True)
data.rename(columns = {"mem_cnt": "가입 수"}, inplace = True)
```

```
In [7]: data
```

```
Out[7]:
```

| | 가입일자 | 사용자코드 | 성별 | 연령대코드 | 가입 수 |
|------|------------|--------|-----|-------|------|
| 0 | 2018-07-19 | 회원-내국인 | | 70대~ | 2194 |
| 1 | 2018-07-18 | 회원-내국인 | | 70대~ | 2123 |
| 2 | 2018-07-17 | 회원-내국인 | | 70대~ | 2133 |
| 3 | 2018-07-16 | 회원-내국인 | | 70대~ | 2161 |
| 4 | 2018-07-15 | 회원-내국인 | | 70대~ | 2937 |
| ... | ... | ... | ... | ... | ... |
| 8090 | 2017-01-01 | 회원-내국인 | F | ~10대 | 3 |
| 8091 | 2017-01-01 | 회원-내국인 | F | 50대 | 5 |
| 8092 | 2017-01-01 | 회원-내국인 | F | 40대 | 11 |
| 8093 | 2017-01-01 | 회원-내국인 | F | 30대 | 16 |
| 8094 | 2017-01-01 | 회원-내국인 | F | 20대 | 53 |

8095 rows × 5 columns

```
In [8]: data.dtypes
```

```
Out[8]: 가입일자      object
사용자코드    object
성별          object
연령대코드    object
가입 수       object
dtype: object
```

open api 파일은 2018-07-19까지 밖에 없어서 추가로 xlsx 파일을 불러올 것이다.

```
In [9]: data1 = pd.read_excel("서울특별시 공공자전거 신규가입자 정보(일별)_20180720_20181231.xlsx")
```

```
In [10]: data1
```

```
Out[10]:
```

| | 가입일자 | 사용자코드 | 연령대코드 | 성별 | 가입 수 |
|------|------------|--------|---------|-----|------|
| 0 | 2018-07-20 | 회원-내국인 | AGE_001 | NaN | 221 |
| 1 | 2018-07-20 | 회원-내국인 | AGE_002 | NaN | 1189 |
| 2 | 2018-07-20 | 회원-내국인 | AGE_003 | NaN | 348 |
| 3 | 2018-07-20 | 회원-내국인 | AGE_004 | NaN | 201 |
| 4 | 2018-07-20 | 회원-내국인 | AGE_005 | NaN | 91 |
| ... | ... | ... | ... | ... | ... |
| 1126 | 2018-12-31 | 회원-내국인 | AGE_003 | NaN | 49 |
| 1127 | 2018-12-31 | 회원-내국인 | AGE_004 | NaN | 22 |
| 1128 | 2018-12-31 | 회원-내국인 | AGE_005 | NaN | 10 |
| 1129 | 2018-12-31 | 회원-내국인 | AGE_006 | NaN | 2 |
| 1130 | 2018-12-31 | 회원-내국인 | AGE_007 | NaN | 1 |

1131 rows × 5 columns

```
In [11]: data1.dtypes
```

```
Out[11]: 가입일자      datetime64[ns]  
사용자코드      object  
연령대코드      object  
성별            float64  
가입 수         int64  
dtype: object
```

```
In [12]: import datetime
```

```
data1['가입일자'] = data1['가입일자'].dt.strftime("%Y-%m-%d")
```

In [13]: data1

Out[13]:

| | 가입일자 | 사용자코드 | 연령대코드 | 성별 | 가입 수 |
|------|------------|--------|---------|-----|------|
| 0 | 2018-07-20 | 회원-내국인 | AGE_001 | NaN | 221 |
| 1 | 2018-07-20 | 회원-내국인 | AGE_002 | NaN | 1189 |
| 2 | 2018-07-20 | 회원-내국인 | AGE_003 | NaN | 348 |
| 3 | 2018-07-20 | 회원-내국인 | AGE_004 | NaN | 201 |
| 4 | 2018-07-20 | 회원-내국인 | AGE_005 | NaN | 91 |
| ... | ... | ... | ... | ... | ... |
| 1126 | 2018-12-31 | 회원-내국인 | AGE_003 | NaN | 49 |
| 1127 | 2018-12-31 | 회원-내국인 | AGE_004 | NaN | 22 |
| 1128 | 2018-12-31 | 회원-내국인 | AGE_005 | NaN | 10 |
| 1129 | 2018-12-31 | 회원-내국인 | AGE_006 | NaN | 2 |
| 1130 | 2018-12-31 | 회원-내국인 | AGE_007 | NaN | 1 |

1131 rows × 5 columns

In [14]: data2 = pd.read_excel("서울특별시 공공자전거 신규가입자 정보(일별)_20190101_20191130.xlsx")

In [15]: data2

Out[15]:

| | 가입일자 | 사용자코드 | 연령대코드 | 성별 | 가입 수 |
|------|------------|--------|---------|-----|------|
| 0 | 2019-01-01 | 회원-내국인 | AGE_001 | NaN | 64 |
| 1 | 2019-01-01 | 회원-내국인 | AGE_002 | NaN | 196 |
| 2 | 2019-01-01 | 회원-내국인 | AGE_003 | NaN | 51 |
| 3 | 2019-01-01 | 회원-내국인 | AGE_004 | NaN | 37 |
| 4 | 2019-01-01 | 회원-내국인 | AGE_005 | NaN | 17 |
| ... | ... | ... | ... | ... | ... |
| 2503 | 2019-11-30 | 회원-내국인 | AGE_005 | F | 23 |
| 2504 | 2019-11-30 | 회원-내국인 | AGE_005 | M | 41 |
| 2505 | 2019-11-30 | 회원-내국인 | AGE_006 | F | 7 |
| 2506 | 2019-11-30 | 회원-내국인 | AGE_006 | M | 8 |
| 2507 | 2019-11-30 | 회원-내국인 | AGE_007 | M | 1 |

2508 rows × 5 columns

```
In [16]: data2.dtypes
```

```
Out[16]: 가입일자      datetime64[ns]  
사용자코드      object  
연령대코드      object  
성별            object  
가입 수        int64  
dtype: object
```

```
In [17]: data2['가입일자'] = data2['가입일자'].dt.strftime("%Y-%m-%d")
```

```
In [18]: data2.dtypes
```

```
Out[18]: 가입일자      object  
사용자코드      object  
연령대코드      object  
성별            object  
가입 수        int64  
dtype: object
```

```
In [19]: data2
```

```
Out[19]:
```

| | 가입일자 | 사용자코드 | 연령대코드 | 성별 | 가입 수 |
|------|------------|--------|---------|-----|------|
| 0 | 2019-01-01 | 회원-내국인 | AGE_001 | NaN | 64 |
| 1 | 2019-01-01 | 회원-내국인 | AGE_002 | NaN | 196 |
| 2 | 2019-01-01 | 회원-내국인 | AGE_003 | NaN | 51 |
| 3 | 2019-01-01 | 회원-내국인 | AGE_004 | NaN | 37 |
| 4 | 2019-01-01 | 회원-내국인 | AGE_005 | NaN | 17 |
| ... | ... | ... | ... | ... | ... |
| 2503 | 2019-11-30 | 회원-내국인 | AGE_005 | F | 23 |
| 2504 | 2019-11-30 | 회원-내국인 | AGE_005 | M | 41 |
| 2505 | 2019-11-30 | 회원-내국인 | AGE_006 | F | 7 |
| 2506 | 2019-11-30 | 회원-내국인 | AGE_006 | M | 8 |
| 2507 | 2019-11-30 | 회원-내국인 | AGE_007 | M | 1 |

2508 rows × 5 columns

```
In [20]: data.shape
```

```
Out[20]: (8095, 5)
```

```
In [21]: type(data1)
```

```
Out[21]: pandas.core.frame.DataFrame
```

```
In [22]: merge = pd.concat([data, data1, data2])
```

C:\anaconda\lib\site-packages\ipykernel_launcher.py:1: FutureWarning: Sorting because non-concatenation axis is not aligned. A future version of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=False'.

To retain the current behavior and silence the warning, pass 'sort=True'.

"""Entry point for launching an IPython kernel.

```
In [23]: merge
```

Out[23]:

| | 가입 수 | 가입일자 | 사용자코드 | 성별 | 연령대코드 |
|------|------|------------|--------|-----|---------|
| 0 | 2194 | 2018-07-19 | 회원-내국인 | | 70대~ |
| 1 | 2123 | 2018-07-18 | 회원-내국인 | | 70대~ |
| 2 | 2133 | 2018-07-17 | 회원-내국인 | | 70대~ |
| 3 | 2161 | 2018-07-16 | 회원-내국인 | | 70대~ |
| 4 | 2937 | 2018-07-15 | 회원-내국인 | | 70대~ |
| ... | ... | ... | ... | ... | ... |
| 2503 | 23 | 2019-11-30 | 회원-내국인 | F | AGE_005 |
| 2504 | 41 | 2019-11-30 | 회원-내국인 | M | AGE_005 |
| 2505 | 7 | 2019-11-30 | 회원-내국인 | F | AGE_006 |
| 2506 | 8 | 2019-11-30 | 회원-내국인 | M | AGE_006 |
| 2507 | 1 | 2019-11-30 | 회원-내국인 | M | AGE_007 |

11734 rows × 5 columns

```
In [24]: merge.shape
```

Out[24]: (11734, 5)

```
In [25]: bikejoin = merge
```

In [26]: bikejoin

Out[26]:

| | 가입 수 | 가입일자 | 사용자코드 | 성별 | 연령대코드 |
|------|------|------------|--------|-----|---------|
| 0 | 2194 | 2018-07-19 | 회원-내국인 | | 70대~ |
| 1 | 2123 | 2018-07-18 | 회원-내국인 | | 70대~ |
| 2 | 2133 | 2018-07-17 | 회원-내국인 | | 70대~ |
| 3 | 2161 | 2018-07-16 | 회원-내국인 | | 70대~ |
| 4 | 2937 | 2018-07-15 | 회원-내국인 | | 70대~ |
| ... | ... | ... | ... | ... | ... |
| 2503 | 23 | 2019-11-30 | 회원-내국인 | F | AGE_005 |
| 2504 | 41 | 2019-11-30 | 회원-내국인 | M | AGE_005 |
| 2505 | 7 | 2019-11-30 | 회원-내국인 | F | AGE_006 |
| 2506 | 8 | 2019-11-30 | 회원-내국인 | M | AGE_006 |
| 2507 | 1 | 2019-11-30 | 회원-내국인 | M | AGE_007 |

11734 rows × 5 columns

In [27]: bikejoin.dtypes

Out[27]: 가입 수 object
가입일자 object
사용자코드 object
성별 object
연령대코드 object
dtype: object

공공자전거 결측치 확인

신규 가입자 수만 추출할 것 이므로 성별 결측치는 그대로 둘 것이다.

In [28]: bikejoin.isnull().sum()

Out[28]: 가입 수 0
가입일자 0
사용자코드 0
성별 3231
연령대코드 0
dtype: int64

공공자전거 데이터 전처리

신규 가입자 수만 추출


```
In [29]: bikejoin.dtypes
```

```
Out[29]: 가입 수      object  
가입일자      object  
사용자코드    object  
성별          object  
연령대코드    object  
dtype: object
```

```
In [30]: bikejoin['가입 수'] = pd.to_numeric(bikejoin['가입 수'])
```

```
In [31]: bikejoin.dtypes
```

```
Out[31]: 가입 수      int64  
가입일자      object  
사용자코드    object  
성별          object  
연령대코드    object  
dtype: object
```

```
In [32]: bikejoin = bikejoin.groupby('가입일자')['가입 수'].sum()
```

```
In [33]: bikejoin
```

```
Out[33]: 가입일자  
2017-01-01      221  
2017-01-02      182  
2017-01-03      198  
2017-01-04      189  
2017-01-05      190  
...  
2019-11-26      920  
2019-11-27      688  
2019-11-28      712  
2019-11-29      671  
2019-11-30     1100  
Name: 가입 수, Length: 1062, dtype: int64
```

```
In [34]: type(bikejoin)
```

```
Out[34]: pandas.core.series.Series
```

```
In [35]: bikejoin = pd.DataFrame(bikejoin)
```

In [36]: bikejoin

Out[36]:

가입 수

가입일 자

| | |
|------------|------|
| 2017-01-01 | 221 |
| 2017-01-02 | 182 |
| 2017-01-03 | 198 |
| 2017-01-04 | 189 |
| 2017-01-05 | 190 |
| ... | ... |
| 2019-11-26 | 920 |
| 2019-11-27 | 688 |
| 2019-11-28 | 712 |
| 2019-11-29 | 671 |
| 2019-11-30 | 1100 |

1062 rows × 1 columns

In [37]: bikejoin.reset_index(inplace=True)

In [38]: bikejoin

Out[38]:

가입일 자 가입 수

| | | |
|------|------------|------|
| 0 | 2017-01-01 | 221 |
| 1 | 2017-01-02 | 182 |
| 2 | 2017-01-03 | 198 |
| 3 | 2017-01-04 | 189 |
| 4 | 2017-01-05 | 190 |
| ... | ... | ... |
| 1057 | 2019-11-26 | 920 |
| 1058 | 2019-11-27 | 688 |
| 1059 | 2019-11-28 | 712 |
| 1060 | 2019-11-29 | 671 |
| 1061 | 2019-11-30 | 1100 |

1062 rows × 2 columns

In [39]: bikejoin = bikejoin.sort_values(by=['가입일 자'])

In [40]: bikejoin

Out[40]:

| | 가입일자 | 가입 수 |
|------|------------|------|
| 0 | 2017-01-01 | 221 |
| 1 | 2017-01-02 | 182 |
| 2 | 2017-01-03 | 198 |
| 3 | 2017-01-04 | 189 |
| 4 | 2017-01-05 | 190 |
| ... | ... | ... |
| 1057 | 2019-11-26 | 920 |
| 1058 | 2019-11-27 | 688 |
| 1059 | 2019-11-28 | 712 |
| 1060 | 2019-11-29 | 671 |
| 1061 | 2019-11-30 | 1100 |

1062 rows × 2 columns

In [41]: bikejoin.dtypes

Out[41]: 가입일자 object
가입 수 int64
dtype: object

In [42]: bikejoin.rename(columns = {"가입일자": "날짜"}, inplace = True)

In [43]: bikejoin

Out[43]:

| | 날짜 | 가입 수 |
|------|------------|------|
| 0 | 2017-01-01 | 221 |
| 1 | 2017-01-02 | 182 |
| 2 | 2017-01-03 | 198 |
| 3 | 2017-01-04 | 189 |
| 4 | 2017-01-05 | 190 |
| ... | ... | ... |
| 1057 | 2019-11-26 | 920 |
| 1058 | 2019-11-27 | 688 |
| 1059 | 2019-11-28 | 712 |
| 1060 | 2019-11-29 | 671 |
| 1061 | 2019-11-30 | 1100 |

1062 rows × 2 columns

날씨 데이터 불러오기

- tm : 날짜
- min_ta : 최저 기온
- max_ta : 최고 기온
- avg_tca : 평균 전운량 -> 맑음(0-2), 구름조금(3-5), 구름많음(6-8), 흐림(9-10이상)
- sum_rn : 일 강수량
- iscs : 일기 현상
- avg_ws : 평균 풍속
- stn_nm : 지역
- stn_id : 지점 번호

```
In [44]: from urllib.request import urlopen
        from urllib.parse import urlencode, unquote, quote_plus
        import urllib
        import requests
        import json
        import pandas as pd
```

```
In [45]: url = 'http://data.kma.go.kr/apiData/getData'
```

```
In [46]: params = '?' + urlencode({
        quote_plus("type"): "json",
        quote_plus("dataCd"): "ASOS",
        quote_plus("dateCd"): "DAY",
        quote_plus("startDt"): "20170101",
        quote_plus("endDt"): "20191130",
        quote_plus("stnIds"): "108",
        quote_plus("schListCnt"): "800",
        quote_plus("pageIndex"): "1",
        quote_plus("apiKey"): "XfV3VGyAZTRP/zPdbyN%2BR15NpGdAjhBndaCbr9QuWgK5%2BF
        kvhTuv7XTDZPvaCg3f"
    })

    req = urllib.request.Request(url + unquote(params))
```

```
In [47]: params1 = '?' + urlencode({
        quote_plus("type"): "json",
        quote_plus("dataCd"): "ASOS",
        quote_plus("dateCd"): "DAY",
        quote_plus("startDt"): "20170101",
        quote_plus("endDt"): "20191130",
        quote_plus("stnIds"): "108",
        quote_plus("schListCnt"): "800",
        quote_plus("pageIndex"): "2",
        quote_plus("apiKey"): "XfV3VGyAZTRP/zPdbyN%2BR15NpGdAjhBndaCbr9QuWgK5%2BF
        kvhTuv7XTDZPvaCg3f"
    })

    req1 = urllib.request.Request(url + unquote(params1))
```

```
In [48]: response_body = urlopen(req, timeout=60).read() # get bytes data
weather_data1 = json.loads(response_body) # convert bytes data to json data
```

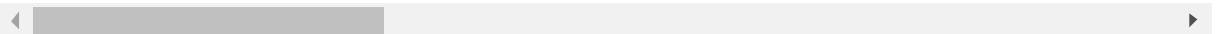
```
In [49]: res1 = pd.DataFrame(weather_data1[3]['info'])
```

```
In [50]: res1
```

```
Out[50]:
```

| | AVG_M1_5_TE | N9_9_RN | SUM_LRG_EV | MIN_RHM_HRMT | STN_ID | MAX_INS_WS_WD | MA |
|-----|-------------|---------|------------|--------------|--------|---------------|-----|
| 0 | 10.2 | 0.3 | 0.7 | 1342.0 | 108 | 20.0 | |
| 1 | 10.1 | NaN | 0.9 | 1355.0 | 108 | 230.0 | |
| 2 | 9.9 | NaN | 1.3 | 1254.0 | 108 | 180.0 | |
| 3 | 9.8 | NaN | 1.5 | 1503.0 | 108 | 290.0 | |
| 4 | 9.7 | NaN | 1.6 | 943.0 | 108 | 50.0 | |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 795 | 6.7 | NaN | 2.9 | 1603.0 | 108 | 320.0 | |
| 796 | 6.9 | NaN | 3.4 | 1708.0 | 108 | 200.0 | |
| 797 | 7.0 | NaN | 2.9 | 1650.0 | 108 | 290.0 | |
| 798 | 7.2 | NaN | 2.2 | 1142.0 | 108 | 70.0 | |
| 799 | 7.3 | 0.0 | 1.5 | 1234.0 | 108 | 200.0 | |

800 rows × 63 columns



```
In [51]: response_body1 = urlopen(req1, timeout=60).read() # get bytes data
weather_data2 = json.loads(response_body1) # convert bytes data to json data
```

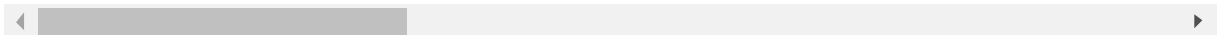
```
In [52]: res2 = pd.DataFrame(weather_data2[3]['info'])
```

In [53]: res2

Out[53]:

| | AVG_M1_5_TE | N9_9_RN | SUM_LRG_EV | MIN_RHM_HRMT | STN_ID | MAX_INS_WS_WD | MA |
|-----|-------------|---------|------------|--------------|--------|---------------|-----|
| 0 | 7.4 | 0.0 | 1.7 | 1705 | 108 | 270 | |
| 1 | 7.5 | 0.0 | 3.2 | 1345 | 108 | 340 | |
| 2 | 7.5 | 0.0 | 2.2 | 1335 | 108 | 200 | |
| 3 | 7.6 | 3.5 | 1.5 | 1505 | 108 | 270 | |
| 4 | 7.7 | NaN | 1.6 | 1323 | 108 | 200 | |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 259 | 15.0 | NaN | 1.7 | 1439 | 108 | 270 | |
| 260 | 14.8 | NaN | 1.8 | 1427 | 108 | 290 | |
| 261 | 14.7 | NaN | 1.9 | 1152 | 108 | 320 | |
| 262 | 14.5 | NaN | 2.0 | 1409 | 108 | 320 | |
| 263 | 14.5 | 0.1 | 1.5 | 1409 | 108 | 320 | |

264 rows × 58 columns



In [54]: weather1 = res1[['TM', 'MIN_TA', 'MAX_TA', 'AVG_TCA', 'SUM_RN', 'AVG_WS', 'STN_NM', 'STN_ID']]

In [55]: weather2 = res2[['TM', 'MIN_TA', 'MAX_TA', 'AVG_TCA', 'SUM_RN', 'AVG_WS', 'STN_NM', 'STN_ID']]

In [56]: weather1

Out[56]:

| | TM | MIN_TA | MAX_TA | AVG_TCA | SUM_RN | AVG_WS | STN_NM | STN_ID |
|-----|------------|--------|--------|---------|--------|--------|--------|--------|
| 0 | 2017-01-01 | -1.6 | 6.9 | 7.0 | NaN | 1.5 | 서울 | 108 |
| 1 | 2017-01-02 | 1.8 | 9.2 | 7.6 | 0.3 | 2.1 | 서울 | 108 |
| 2 | 2017-01-03 | -2.3 | 7.7 | 1.1 | NaN | 1.8 | 서울 | 108 |
| 3 | 2017-01-04 | 1.0 | 8.9 | 2.6 | NaN | 1.7 | 서울 | 108 |
| 4 | 2017-01-05 | -0.1 | 7.3 | 8.4 | NaN | 3.1 | 서울 | 108 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 795 | 2019-03-07 | 3.1 | 12.7 | 1.5 | NaN | 2.2 | 서울 | 108 |
| 796 | 2019-03-08 | 0.6 | 13.4 | 0.1 | NaN | 1.9 | 서울 | 108 |
| 797 | 2019-03-09 | -0.4 | 14.9 | 5.4 | NaN | 1.3 | 서울 | 108 |
| 798 | 2019-03-10 | 6.7 | 15.3 | 7.1 | NaN | 1.2 | 서울 | 108 |
| 799 | 2019-03-11 | 3.5 | 10.0 | 4.9 | 0.0 | 1.8 | 서울 | 108 |

800 rows × 8 columns

In [57]: weather2

Out[57]:

| | TM | MIN_TA | MAX_TA | AVG_TCA | SUM_RN | AVG_WS | STN_NM | STN_ID |
|-----|------------|--------|--------|---------|--------|--------|--------|--------|
| 0 | 2019-03-12 | -0.3 | 7.9 | 4.9 | 0.0 | 2.7 | 서울 | 108 |
| 1 | 2019-03-13 | -0.4 | 8.5 | 1.8 | 0.0 | 3.3 | 서울 | 108 |
| 2 | 2019-03-14 | -1.7 | 10.0 | 5.5 | 0.0 | 1.9 | 서울 | 108 |
| 3 | 2019-03-15 | 1.6 | 11.3 | 7.4 | 3.5 | 1.6 | 서울 | 108 |
| 4 | 2019-03-16 | 0.2 | 9.2 | 5.5 | NaN | 1.8 | 서울 | 108 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 259 | 2019-11-26 | 1.2 | 13.7 | 3.9 | NaN | 2.0 | 서울 | 108 |
| 260 | 2019-11-27 | 0.3 | 6.8 | 4.0 | NaN | 2.7 | 서울 | 108 |
| 261 | 2019-11-28 | -0.6 | 10.4 | 1.8 | NaN | 2.0 | 서울 | 108 |
| 262 | 2019-11-29 | -3.4 | 6.9 | 0.9 | NaN | 2.0 | 서울 | 108 |
| 263 | 2019-11-30 | -1.4 | 9.3 | 6.0 | NaN | 1.7 | 서울 | 108 |

264 rows × 8 columns

In [58]: weather = pd.concat([weather1, weather2])

In [59]: weather

Out[59]:

| | TM | MIN_TA | MAX_TA | AVG_TCA | SUM_RN | AVG_WS | STN_NM | STN_ID |
|-----|------------|--------|--------|---------|--------|--------|--------|--------|
| 0 | 2017-01-01 | -1.6 | 6.9 | 7.0 | NaN | 1.5 | 서울 | 108 |
| 1 | 2017-01-02 | 1.8 | 9.2 | 7.6 | 0.3 | 2.1 | 서울 | 108 |
| 2 | 2017-01-03 | -2.3 | 7.7 | 1.1 | NaN | 1.8 | 서울 | 108 |
| 3 | 2017-01-04 | 1.0 | 8.9 | 2.6 | NaN | 1.7 | 서울 | 108 |
| 4 | 2017-01-05 | -0.1 | 7.3 | 8.4 | NaN | 3.1 | 서울 | 108 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 259 | 2019-11-26 | 1.2 | 13.7 | 3.9 | NaN | 2.0 | 서울 | 108 |
| 260 | 2019-11-27 | 0.3 | 6.8 | 4.0 | NaN | 2.7 | 서울 | 108 |
| 261 | 2019-11-28 | -0.6 | 10.4 | 1.8 | NaN | 2.0 | 서울 | 108 |
| 262 | 2019-11-29 | -3.4 | 6.9 | 0.9 | NaN | 2.0 | 서울 | 108 |
| 263 | 2019-11-30 | -1.4 | 9.3 | 6.0 | NaN | 1.7 | 서울 | 108 |

1064 rows × 8 columns

```
In [60]: weather.rename(columns = {'TM' : '날짜'}, inplace = True)
weather.rename(columns = {'MIN_TA' : '최저기온'}, inplace = True)
weather.rename(columns = {'MAX_TA' : '최고기온'}, inplace = True)
weather.rename(columns = {'AVG_TCA' : '평균전운량'}, inplace = True)
weather.rename(columns = {'SUM_RN' : '일강수량'}, inplace = True)
weather.rename(columns = {'AVG_WS' : '평균풍속'}, inplace = True)
weather.rename(columns = {'STN_NM' : '지역'}, inplace = True)
weather.rename(columns = {'STN_ID' : '지점번호'}, inplace = True)
```

```
In [61]: weather
```

```
Out[61]:
```

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|-----|------|
| 0 | 2017-01-01 | -1.6 | 6.9 | 7.0 | NaN | 1.5 | 서울 | 108 |
| 1 | 2017-01-02 | 1.8 | 9.2 | 7.6 | 0.3 | 2.1 | 서울 | 108 |
| 2 | 2017-01-03 | -2.3 | 7.7 | 1.1 | NaN | 1.8 | 서울 | 108 |
| 3 | 2017-01-04 | 1.0 | 8.9 | 2.6 | NaN | 1.7 | 서울 | 108 |
| 4 | 2017-01-05 | -0.1 | 7.3 | 8.4 | NaN | 3.1 | 서울 | 108 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 259 | 2019-11-26 | 1.2 | 13.7 | 3.9 | NaN | 2.0 | 서울 | 108 |
| 260 | 2019-11-27 | 0.3 | 6.8 | 4.0 | NaN | 2.7 | 서울 | 108 |
| 261 | 2019-11-28 | -0.6 | 10.4 | 1.8 | NaN | 2.0 | 서울 | 108 |
| 262 | 2019-11-29 | -3.4 | 6.9 | 0.9 | NaN | 2.0 | 서울 | 108 |
| 263 | 2019-11-30 | -1.4 | 9.3 | 6.0 | NaN | 1.7 | 서울 | 108 |

1064 rows × 8 columns

```
In [62]: weather.isnull().sum()
```

```
Out[62]: 날짜          0
최저기온          0
최고기온          1
평균전운량         0
일강수량        672
평균풍속           3
지역              0
지점번호          0
dtype: int64
```

```
In [63]: weather[weather['평균풍속'].isnull()]
```

```
Out[63]:
```

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 286 | 2017-10-14 | 9.0 | 20.5 | 4.9 | NaN | NaN | 서울 | 108 |
| 338 | 2017-12-05 | -8.2 | -0.4 | 1.6 | 0.1 | NaN | 서울 | 108 |
| 339 | 2017-12-06 | -4.5 | 6.0 | 7.0 | 1.2 | NaN | 서울 | 108 |


```
In [64]: weather[weather['최고기온'].isnull()]
```

Out[64]:

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 284 | 2017-10-12 | 8.8 | NaN | 7.1 | NaN | 2.0 | 서울 | 108 |

In [65]: weather[255:310]

Out[65]:

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 255 | 2017-09-13 | 15.9 | 27.5 | 0.0 | NaN | 1.9 | 서울 | 108 |
| 256 | 2017-09-14 | 14.8 | 29.5 | 1.5 | NaN | 1.7 | 서울 | 108 |
| 257 | 2017-09-15 | 17.8 | 27.9 | 3.9 | NaN | 3.6 | 서울 | 108 |
| 258 | 2017-09-16 | 16.7 | 27.6 | 3.1 | NaN | 3.6 | 서울 | 108 |
| 259 | 2017-09-17 | 19.0 | 29.2 | 3.5 | NaN | 2.7 | 서울 | 108 |
| 260 | 2017-09-18 | 16.4 | 27.6 | 1.8 | NaN | 2.1 | 서울 | 108 |
| 261 | 2017-09-19 | 20.1 | 25.8 | 7.9 | 8.0 | 3.0 | 서울 | 108 |
| 262 | 2017-09-20 | 14.1 | 24.9 | 0.8 | NaN | 2.4 | 서울 | 108 |
| 263 | 2017-09-21 | 14.2 | 27.0 | 0.5 | NaN | 1.6 | 서울 | 108 |
| 264 | 2017-09-22 | 15.4 | 26.8 | 3.5 | NaN | 2.1 | 서울 | 108 |
| 265 | 2017-09-23 | 17.6 | 27.3 | 6.6 | 0.0 | 1.6 | 서울 | 108 |
| 266 | 2017-09-24 | 18.8 | 27.9 | 6.9 | 0.0 | 2.0 | 서울 | 108 |
| 267 | 2017-09-25 | 17.4 | 30.1 | 2.5 | NaN | 1.4 | 서울 | 108 |
| 268 | 2017-09-26 | 19.4 | 31.4 | 4.9 | 0.0 | 1.9 | 서울 | 108 |
| 269 | 2017-09-27 | 19.9 | 27.3 | 8.8 | 0.5 | 1.7 | 서울 | 108 |
| 270 | 2017-09-28 | 14.2 | 25.3 | 2.5 | NaN | 2.8 | 서울 | 108 |
| 271 | 2017-09-29 | 11.2 | 23.5 | 3.1 | NaN | 2.2 | 서울 | 108 |
| 272 | 2017-09-30 | 14.7 | 25.9 | 3.5 | NaN | 1.5 | 서울 | 108 |
| 273 | 2017-10-01 | 15.5 | 21.2 | 8.9 | 4.0 | 2.3 | 서울 | 108 |
| 274 | 2017-10-02 | 15.6 | 29.4 | 2.6 | 1.5 | 2.3 | 서울 | 108 |
| 275 | 2017-10-03 | 13.4 | 23.6 | 5.8 | NaN | 2.3 | 서울 | 108 |
| 276 | 2017-10-04 | 10.7 | 24.3 | 6.6 | NaN | 1.8 | 서울 | 108 |
| 277 | 2017-10-05 | 13.9 | 23.4 | 8.0 | NaN | 2.1 | 서울 | 108 |
| 278 | 2017-10-06 | 16.2 | 23.3 | 8.5 | 0.5 | 2.3 | 서울 | 108 |
| 279 | 2017-10-07 | 16.9 | 28.8 | 7.5 | 0.0 | 1.2 | 서울 | 108 |
| 280 | 2017-10-08 | 19.3 | 28.7 | 4.8 | NaN | 1.4 | 서울 | 108 |
| 281 | 2017-10-09 | 19.8 | 27.6 | 6.5 | 0.0 | 1.5 | 서울 | 108 |
| 282 | 2017-10-10 | 18.6 | 24.8 | 9.3 | 0.4 | 1.3 | 서울 | 108 |
| 283 | 2017-10-11 | 12.2 | 21.7 | 9.3 | 20.0 | 2.9 | 서울 | 108 |
| 284 | 2017-10-12 | 8.8 | NaN | 7.1 | NaN | 2.0 | 서울 | 108 |
| 285 | 2017-10-13 | 6.1 | 18.9 | 3.4 | NaN | 3.2 | 서울 | 108 |
| 286 | 2017-10-14 | 9.0 | 20.5 | 4.9 | NaN | NaN | 서울 | 108 |
| 287 | 2017-10-15 | 9.0 | 23.0 | 5.4 | NaN | 1.7 | 서울 | 108 |
| 288 | 2017-10-16 | 13.6 | 22.0 | 7.9 | NaN | 1.7 | 서울 | 108 |
| 289 | 2017-10-17 | 9.2 | 23.9 | 0.4 | NaN | 1.3 | 서울 | 108 |
| 290 | 2017-10-18 | 14.2 | 19.1 | 9.3 | 0.1 | 2.4 | 서울 | 108 |
| 291 | 2017-10-19 | 11.9 | 23.2 | 3.8 | NaN | 1.8 | 서울 | 108 |

~~~~~

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 292 | 2017-10-20 | 11.1 | 24.2 | 0.1   | NaN  | 1.5  | 서울 | 108  |
| 293 | 2017-10-21 | 11.0 | 25.2 | 0.0   | NaN  | 1.3  | 서울 | 108  |
| 294 | 2017-10-22 | 11.0 | 23.5 | 1.8   | NaN  | 3.1  | 서울 | 108  |
| 295 | 2017-10-23 | 8.6  | 20.6 | 0.0   | NaN  | 1.4  | 서울 | 108  |
| 296 | 2017-10-24 | 9.9  | 20.9 | 3.0   | NaN  | 1.9  | 서울 | 108  |
| 297 | 2017-10-25 | 9.5  | 20.6 | 2.9   | NaN  | 1.5  | 서울 | 108  |
| 298 | 2017-10-26 | 10.7 | 18.1 | 8.6   | NaN  | 1.2  | 서울 | 108  |
| 299 | 2017-10-27 | 11.9 | 24.2 | 1.9   | NaN  | 1.2  | 서울 | 108  |
| 300 | 2017-10-28 | 10.6 | 24.8 | 0.1   | NaN  | 2.0  | 서울 | 108  |
| 301 | 2017-10-29 | 6.3  | 17.2 | 0.0   | NaN  | 3.3  | 서울 | 108  |
| 302 | 2017-10-30 | 2.5  | 13.2 | 0.0   | NaN  | 2.4  | 서울 | 108  |
| 303 | 2017-10-31 | 3.6  | 16.0 | 1.8   | NaN  | 1.8  | 서울 | 108  |
| 304 | 2017-11-01 | 11.5 | 18.1 | 8.5   | NaN  | 2.1  | 서울 | 108  |
| 305 | 2017-11-02 | 14.0 | 18.0 | 8.8   | 1.0  | 1.6  | 서울 | 108  |
| 306 | 2017-11-03 | 6.1  | 16.6 | 5.6   | 1.5  | 3.2  | 서울 | 108  |
| 307 | 2017-11-04 | 4.5  | 13.7 | 2.4   | NaN  | 2.5  | 서울 | 108  |
| 308 | 2017-11-05 | 2.6  | 15.4 | 3.3   | NaN  | 1.5  | 서울 | 108  |
| 309 | 2017-11-06 | 8.0  | 16.5 | 5.5   | NaN  | 1.5  | 서울 | 108  |

2017년 10월 12일과 최저 기온이 비슷하고 가까운 날짜의 최고 기온의 평균을 결측치로 대체

In [66]: weather[285:288]

Out[66]:

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 285 | 2017-10-13 | 6.1  | 18.9 | 3.4   | NaN  | 3.2  | 서울 | 108  |
| 286 | 2017-10-14 | 9.0  | 20.5 | 4.9   | NaN  | NaN  | 서울 | 108  |
| 287 | 2017-10-15 | 9.0  | 23.0 | 5.4   | NaN  | 1.7  | 서울 | 108  |

In [67]: weather['최고기온'][285:288].mean()

Out[67]: 20.8

In [68]: weather['최고기온'].fillna(20.8, inplace=True)

```
In [69]: weather.isnull().sum()
```

```
Out[69]: 날짜                0
최저기온                0
최고기온                0
평균전운량              0
일강수량              672
평균풍속                3
지역                  0
지점번호              0
dtype: int64
```

```
In [70]: weather[weather['최고기온'].isnull()]
```

```
Out[70]: 날짜  최저기온  최고기온  평균전운량  일강수량  평균풍속  지역  지점번호
_____
```

```
In [71]: weather.loc[284]
```

```
Out[71]: 날짜                2017-10-12
최저기온                8.8
최고기온              20.8
평균전운량              7.1
일강수량              NaN
평균풍속                2
지역                  서울
지점번호              108
Name: 284, dtype: object
```

'평균풍속' 결측치의 가까운 날짜 10개의 평균을 결측치로 대체하였다.

```
In [72]: weather['평균풍속'][281:292].mean()
```

```
Out[72]: 1.9799999999999998
```

```
In [73]: weather.loc[286]
```

```
Out[73]: 날짜                2017-10-14
최저기온                9
최고기온              20.5
평균전운량              4.9
일강수량              NaN
평균풍속              NaN
지역                  서울
지점번호              108
Name: 286, dtype: object
```

```
In [74]: weather['평균풍속'].loc[286] = 1.98
```

C:\anaconda\lib\site-packages\pandas\core\indexing.py:205: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame  
  
See the caveats in the documentation: [http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
self.\_setitem\_with\_indexer(indexer, value)

```
In [75]: weather.loc[286]
```

```
Out[75]: 날짜          2017-10-14  
최저기온              9  
최고기온            20.5  
평균전운량           4.9  
일강수량            NaN  
평균풍속             1.98  
지역                서울  
지점번호             108  
Name: 286, dtype: object
```

```
In [76]: weather['평균풍속'][333:345]
```

```
Out[76]: 333    3.3  
334    1.6  
335    1.7  
336    1.5  
337    3.4  
338    NaN  
339    NaN  
340    1.7  
341    1.8  
342    1.0  
343    2.0  
344    3.5  
Name: 평균풍속, dtype: float64
```

```
In [77]: weather['평균풍속'][333:345].mean()
```

```
Out[77]: 2.15
```

```
In [78]: weather.loc[338:339]
```

```
Out[78]:
```

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 338 | 2017-12-05 | -8.2 | -0.4 | 1.6   | 0.1  | NaN  | 서울 | 108  |
| 339 | 2017-12-06 | -4.5 | 6.0  | 7.0   | 1.2  | NaN  | 서울 | 108  |

```
In [79]: weather['평균풍속'].fillna(2.15, inplace=True)
```

```
In [80]: weather.loc[338:339]
```

```
Out[80]:
```

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 지역 | 지점번호 |
|-----|------------|------|------|-------|------|------|----|------|
| 338 | 2017-12-05 | -8.2 | -0.4 | 1.6   | 0.1  | 2.15 | 서울 | 108  |
| 339 | 2017-12-06 | -4.5 | 6.0  | 7.0   | 1.2  | 2.15 | 서울 | 108  |

```
In [81]: weather.isnull().sum()
```

```
Out[81]: 날짜            0
최저기온            0
최고기온            0
평균전운량          0
일강수량          672
평균풍속            0
지역              0
지점번호           0
dtype: int64
```

나머지 결측치인 '일강수량'은 결측치가 자료의 반 이상을 차지하여 정확하게 판단할 수 없으므로 0으로 결측치를 대체할 것이다.

```
In [82]: weather['일강수량'].fillna(0, inplace=True)
```

```
In [83]: weather.isnull().sum()
```

```
Out[83]: 날짜            0
최저기온            0
최고기온            0
평균전운량          0
일강수량            0
평균풍속            0
지역              0
지점번호           0
dtype: int64
```

## 날씨 데이터 전처리

지역과 지점번호는 편의상 추출한 것이므로 삭제할 것이다.

```
In [84]: weather = weather.drop(['지역', '지점번호'], axis = 1)
```

In [85]: weather

Out[85]:

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 |
|-----|------------|------|------|-------|------|------|
| 0   | 2017-01-01 | -1.6 | 6.9  | 7.0   | 0.0  | 1.5  |
| 1   | 2017-01-02 | 1.8  | 9.2  | 7.6   | 0.3  | 2.1  |
| 2   | 2017-01-03 | -2.3 | 7.7  | 1.1   | 0.0  | 1.8  |
| 3   | 2017-01-04 | 1.0  | 8.9  | 2.6   | 0.0  | 1.7  |
| 4   | 2017-01-05 | -0.1 | 7.3  | 8.4   | 0.0  | 3.1  |
| ... | ...        | ...  | ...  | ...   | ...  | ...  |
| 259 | 2019-11-26 | 1.2  | 13.7 | 3.9   | 0.0  | 2.0  |
| 260 | 2019-11-27 | 0.3  | 6.8  | 4.0   | 0.0  | 2.7  |
| 261 | 2019-11-28 | -0.6 | 10.4 | 1.8   | 0.0  | 2.0  |
| 262 | 2019-11-29 | -3.4 | 6.9  | 0.9   | 0.0  | 2.0  |
| 263 | 2019-11-30 | -1.4 | 9.3  | 6.0   | 0.0  | 1.7  |

1064 rows × 6 columns

## 데이터 합치기

In [86]: weather.dtypes

Out[86]: 날짜 object  
최저기온 float64  
최고기온 float64  
평균전운량 float64  
일강수량 float64  
평균풍속 float64  
dtype: object



In [87]: bikejoin

Out[87]:

|      | 날 짜        | 가 입 수 |
|------|------------|-------|
| 0    | 2017-01-01 | 221   |
| 1    | 2017-01-02 | 182   |
| 2    | 2017-01-03 | 198   |
| 3    | 2017-01-04 | 189   |
| 4    | 2017-01-05 | 190   |
| ...  | ...        | ...   |
| 1057 | 2019-11-26 | 920   |
| 1058 | 2019-11-27 | 688   |
| 1059 | 2019-11-28 | 712   |
| 1060 | 2019-11-29 | 671   |
| 1061 | 2019-11-30 | 1100  |

1062 rows × 2 columns

In [88]: df = pd.merge(weather, bikejoin, on="날 짜")

In [89]: df

Out[89]:

|      | 날 짜        | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|------|------------|---------|---------|-----------|---------|---------|-------|
| 0    | 2017-01-01 | -1.6    | 6.9     | 7.0       | 0.0     | 1.5     | 221   |
| 1    | 2017-01-02 | 1.8     | 9.2     | 7.6       | 0.3     | 2.1     | 182   |
| 2    | 2017-01-03 | -2.3    | 7.7     | 1.1       | 0.0     | 1.8     | 198   |
| 3    | 2017-01-04 | 1.0     | 8.9     | 2.6       | 0.0     | 1.7     | 189   |
| 4    | 2017-01-05 | -0.1    | 7.3     | 8.4       | 0.0     | 3.1     | 190   |
| ...  | ...        | ...     | ...     | ...       | ...     | ...     | ...   |
| 1057 | 2019-11-26 | 1.2     | 13.7    | 3.9       | 0.0     | 2.0     | 920   |
| 1058 | 2019-11-27 | 0.3     | 6.8     | 4.0       | 0.0     | 2.7     | 688   |
| 1059 | 2019-11-28 | -0.6    | 10.4    | 1.8       | 0.0     | 2.0     | 712   |
| 1060 | 2019-11-29 | -3.4    | 6.9     | 0.9       | 0.0     | 2.0     | 671   |
| 1061 | 2019-11-30 | -1.4    | 9.3     | 6.0       | 0.0     | 1.7     | 1100  |

1062 rows × 7 columns

```
In [90]: df.isnull().sum()
```

```
Out[90]: 날짜            0
최저기온            0
최고기온            0
평균전운량          0
일강수량            0
평균풍속            0
가입 수            0
dtype: int64
```

```
In [91]: df.dtypes
```

```
Out[91]: 날짜            object
최저기온            float64
최고기온            float64
평균전운량          float64
일강수량            float64
평균풍속            float64
가입 수            int64
dtype: object
```

## 4. 분석 결과

신규가입 수가 5,000명 이상이면 매우 많고, 100명 이하면 매우 적다고 판단하였다.

```
In [92]: cloud_high = df.sort_values(by='평균전운량', ascending=False).head(50)
```

In [93]: cloud\_high

Out[93]:

|      | 날 짜        | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|------|------------|---------|---------|-----------|---------|---------|-------|
| 498  | 2018-05-16 | 21.8    | 24.4    | 10.0      | 45.0    | 1.9     | 634   |
| 674  | 2018-11-08 | 10.5    | 14.7    | 10.0      | 64.0    | 1.8     | 281   |
| 866  | 2019-05-19 | 17.9    | 21.6    | 10.0      | 22.0    | 1.2     | 1331  |
| 603  | 2018-08-29 | 22.7    | 27.4    | 10.0      | 42.0    | 2.4     | 557   |
| 190  | 2017-07-10 | 22.6    | 25.4    | 10.0      | 144.5   | 2.3     | 255   |
| 195  | 2017-07-15 | 23.4    | 27.3    | 10.0      | 42.5    | 2.2     | 423   |
| 621  | 2018-09-16 | 21.0    | 23.9    | 10.0      | 2.0     | 1.7     | 2127  |
| 976  | 2019-09-06 | 21.6    | 29.1    | 10.0      | 2.4     | 2.1     | 675   |
| 226  | 2017-08-15 | 20.8    | 24.0    | 10.0      | 93.5    | 3.6     | 865   |
| 495  | 2018-05-12 | 11.8    | 16.1    | 10.0      | 32.0    | 0.9     | 536   |
| 625  | 2018-09-20 | 17.8    | 20.1    | 10.0      | 5.0     | 1.8     | 951   |
| 977  | 2019-09-07 | 22.6    | 29.5    | 10.0      | 2.8     | 6.0     | 424   |
| 499  | 2018-05-17 | 18.8    | 23.5    | 10.0      | 83.0    | 1.4     | 312   |
| 426  | 2018-03-04 | 7.6     | 11.9    | 10.0      | 11.5    | 1.2     | 818   |
| 422  | 2018-02-28 | 1.8     | 7.8     | 10.0      | 25.0    | 2.0     | 213   |
| 231  | 2017-08-20 | 22.0    | 25.8    | 10.0      | 124.5   | 2.9     | 309   |
| 539  | 2018-06-26 | 20.2    | 26.9    | 10.0      | 71.5    | 2.4     | 635   |
| 182  | 2017-07-02 | 21.9    | 26.8    | 10.0      | 92.0    | 2.5     | 513   |
| 476  | 2018-04-23 | 8.4     | 10.7    | 10.0      | 59.0    | 2.2     | 304   |
| 52   | 2017-02-22 | -0.9    | 2.9     | 9.9       | 5.1     | 2.7     | 77    |
| 673  | 2018-11-07 | 10.3    | 15.7    | 9.9       | 2.0     | 1.1     | 748   |
| 485  | 2018-05-02 | 8.7     | 19.7    | 9.9       | 12.0    | 1.9     | 651   |
| 934  | 2019-07-26 | 23.7    | 25.9    | 9.9       | 60.0    | 1.7     | 634   |
| 933  | 2019-07-25 | 25.6    | 27.4    | 9.9       | 9.7     | 1.2     | 901   |
| 221  | 2017-08-10 | 23.9    | 26.3    | 9.9       | 5.5     | 2.0     | 456   |
| 699  | 2018-12-03 | 9.3     | 13.5    | 9.9       | 9.5     | 1.4     | 339   |
| 919  | 2019-07-11 | 19.5    | 25.9    | 9.9       | 2.9     | 1.3     | 1951  |
| 94   | 2017-04-05 | 10.8    | 14.2    | 9.9       | 37.5    | 2.4     | 214   |
| 458  | 2018-04-05 | 6.7     | 9.5     | 9.9       | 10.5    | 2.0     | 464   |
| 183  | 2017-07-03 | 21.9    | 27.4    | 9.9       | 67.5    | 3.5     | 641   |
| 544  | 2018-07-01 | 21.0    | 22.8    | 9.9       | 83.5    | 0.8     | 689   |
| 1002 | 2019-10-02 | 20.0    | 24.2    | 9.9       | 23.1    | 3.0     | 795   |
| 443  | 2018-03-21 | 1.1     | 5.3     | 9.9       | 1.0     | 1.7     | 546   |
| 552  | 2018-07-09 | 20.1    | 24.2    | 9.9       | 14.0    | 1.1     | 895   |
| 188  | 2017-07-08 | 24.6    | 27.6    | 9.9       | 4.5     | 2.6     | 745   |
| 379  | 2018-01-16 | 0.1     | 6.8     | 9.8       | 0.0     | 0.7     | 202   |
| 59   | 2017-03-01 | 2.0     | 11.4    | 9.8       | 5.0     | 1.8     | 515   |

~~~~~

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|-------|------|------|
| 440 | 2018-03-18 | 5.4 | 12.5 | 9.8 | 0.5 | 1.2 | 1233 |
| 437 | 2018-03-15 | 8.4 | 16.6 | 9.8 | 27.0 | 2.1 | 538 |
| 203 | 2017-07-23 | 26.1 | 28.1 | 9.8 | 133.5 | 2.1 | 627 |
| 939 | 2019-07-31 | 24.0 | 27.5 | 9.8 | 62.3 | 2.2 | 762 |
| 936 | 2019-07-28 | 25.0 | 28.2 | 9.8 | 25.6 | 2.2 | 990 |
| 620 | 2018-09-15 | 21.0 | 26.0 | 9.8 | 0.2 | 1.0 | 3810 |
| 980 | 2019-09-10 | 22.1 | 26.5 | 9.8 | 63.2 | 2.4 | 792 |
| 640 | 2018-10-05 | 15.7 | 19.0 | 9.8 | 36.5 | 1.7 | 606 |
| 982 | 2019-09-12 | 18.7 | 22.9 | 9.8 | 3.9 | 2.6 | 1838 |
| 156 | 2017-06-06 | 14.5 | 23.9 | 9.8 | 15.5 | 2.5 | 1247 |
| 608 | 2018-09-03 | 21.4 | 25.8 | 9.8 | 34.5 | 1.0 | 913 |
| 553 | 2018-07-10 | 20.4 | 26.5 | 9.8 | 16.5 | 0.7 | 1347 |
| 225 | 2017-08-14 | 23.1 | 27.1 | 9.6 | 1.5 | 3.6 | 946 |

In [94]: cloud_high['가입 수'].mean()

Out[94]: 785.5

In [95]: cloud_low = df.sort_values(by='평균전운량', ascending=False).tail(50)

In [96]: cloud_low

Out[96]:

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|------|-------|------|------|------|
| 647 | 2018-10-12 | 5.2 | 17.4 | 0.0 | 0.0 | 1.3 | 2032 |
| 638 | 2018-10-03 | 11.2 | 24.2 | 0.0 | 0.0 | 1.3 | 6966 |
| 481 | 2018-04-28 | 10.5 | 21.7 | 0.0 | 0.0 | 2.0 | 3379 |
| 292 | 2017-10-21 | 11.0 | 25.2 | 0.0 | 0.0 | 1.3 | 2566 |
| 294 | 2017-10-23 | 8.6 | 20.6 | 0.0 | 0.0 | 1.4 | 1358 |
| 300 | 2017-10-29 | 6.3 | 17.2 | 0.0 | 0.0 | 3.3 | 1522 |
| 591 | 2018-08-17 | 21.7 | 33.8 | 0.0 | 0.0 | 1.8 | 2799 |
| 301 | 2017-10-30 | 2.5 | 13.2 | 0.0 | 0.0 | 2.4 | 625 |
| 369 | 2018-01-06 | -6.8 | 2.9 | 0.0 | 0.0 | 1.5 | 192 |
| 311 | 2017-11-09 | 3.7 | 16.8 | 0.0 | 0.0 | 1.2 | 695 |
| 313 | 2017-11-11 | 2.7 | 10.7 | 0.0 | 0.0 | 2.2 | 733 |
| 318 | 2017-11-16 | -3.4 | 6.4 | 0.0 | 0.0 | 1.5 | 264 |
| 320 | 2017-11-18 | -4.8 | 2.6 | 0.0 | 0.0 | 3.4 | 217 |
| 344 | 2017-12-12 | -12.3 | -5.2 | 0.0 | 0.0 | 2.1 | 83 |
| 345 | 2017-12-13 | -12.0 | -4.4 | 0.0 | 0.0 | 1.7 | 81 |
| 348 | 2017-12-16 | -8.1 | -2.6 | 0.0 | 0.0 | 3.0 | 124 |
| 514 | 2018-06-01 | 17.5 | 30.2 | 0.0 | 0.0 | 1.9 | 2109 |
| 501 | 2018-05-19 | 10.3 | 24.7 | 0.0 | 0.0 | 1.5 | 4342 |
| 708 | 2018-12-12 | -5.0 | 2.4 | 0.0 | 0.0 | 2.2 | 263 |
| 725 | 2018-12-29 | -10.5 | -3.0 | 0.0 | 0.0 | 2.9 | 168 |
| 24 | 2017-01-25 | -10.2 | 0.9 | 0.0 | 0.0 | 1.5 | 72 |
| 729 | 2019-01-02 | -8.8 | 0.2 | 0.0 | 0.0 | 1.7 | 295 |
| 994 | 2019-09-24 | 13.9 | 26.5 | 0.0 | 0.0 | 1.8 | 3194 |
| 36 | 2017-02-06 | -4.4 | 2.6 | 0.0 | 0.0 | 3.0 | 125 |
| 42 | 2017-02-12 | -7.4 | 4.2 | 0.0 | 0.0 | 2.1 | 179 |
| 44 | 2017-02-14 | -4.0 | 6.5 | 0.0 | 0.0 | 2.5 | 166 |
| 45 | 2017-02-15 | -3.9 | 8.0 | 0.0 | 0.0 | 1.4 | 209 |
| 57 | 2017-02-27 | -1.8 | 10.7 | 0.0 | 0.0 | 1.8 | 389 |
| 92 | 2017-04-03 | 4.3 | 19.2 | 0.0 | 0.0 | 2.1 | 787 |
| 398 | 2018-02-04 | -12.8 | -5.2 | 0.0 | 0.0 | 3.1 | 105 |
| 118 | 2017-04-29 | 9.6 | 23.7 | 0.0 | 0.0 | 2.7 | 2319 |
| 119 | 2017-04-30 | 13.7 | 27.8 | 0.0 | 0.0 | 2.6 | 2555 |
| 120 | 2017-05-01 | 12.8 | 28.3 | 0.0 | 0.0 | 1.9 | 2201 |
| 122 | 2017-05-03 | 14.5 | 30.2 | 0.0 | 0.0 | 2.0 | 2956 |
| 870 | 2019-05-23 | 15.0 | 29.2 | 0.0 | 0.0 | 1.7 | 3290 |
| 868 | 2019-05-21 | 11.7 | 23.8 | 0.0 | 0.0 | 2.9 | 3595 |
| 395 | 2018-02-01 | -8.8 | 1.2 | 0.0 | 0.0 | 1.6 | 104 |

Out[96]:

| | 날짜 | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|-------|-------|------|------|------|
| 854 | 2019-05-07 | 8.4 | 21.9 | 0.0 | 0.0 | 2.3 | 3559 |
| 850 | 2019-05-03 | 10.1 | 26.1 | 0.0 | 0.0 | 1.6 | 4612 |
| 849 | 2019-05-02 | 10.6 | 25.3 | 0.0 | 0.0 | 2.3 | 3957 |
| 138 | 2017-05-19 | 15.1 | 27.9 | 0.0 | 0.0 | 2.6 | 1514 |
| 832 | 2019-04-15 | 3.9 | 18.1 | 0.0 | 0.0 | 1.8 | 2833 |
| 821 | 2019-04-04 | 3.0 | 15.7 | 0.0 | 0.0 | 2.5 | 1921 |
| 389 | 2018-01-26 | -17.8 | -10.7 | 0.0 | 0.0 | 2.6 | 42 |
| 779 | 2019-02-21 | -2.6 | 9.0 | 0.0 | 0.0 | 1.5 | 666 |
| 167 | 2017-06-17 | 20.6 | 29.9 | 0.0 | 0.0 | 2.4 | 2534 |
| 753 | 2019-01-26 | -6.5 | 2.4 | 0.0 | 0.0 | 2.4 | 484 |
| 740 | 2019-01-13 | -3.0 | 7.6 | 0.0 | 0.0 | 1.4 | 534 |
| 735 | 2019-01-08 | -7.2 | 0.5 | 0.0 | 0.0 | 2.6 | 329 |
| 387 | 2018-01-24 | -16.3 | -10.7 | 0.0 | 0.0 | 3.4 | 64 |

In [97]: cloud_low['가입 수'].mean()

Out[97]: 1522.16

평균 전운량의 수치가 높으면 비가 올 확률도 있고, 흐리기 때문에 맑은 날씨보다 습한 느낌이 있어서 가입 수가 적을 것이라고 예상했는데 평균 전운량의 수치가 높은 쪽과 낮은 쪽의 가입 수는 눈에 띄는 큰 차이가 없었다. 하지만 평균 전운량의 수치가 낮은 쪽이 가입 수의 평균이 더 높았다.

In [98]: wind_high = df.sort_values(by='평균풍속', ascending=False).head(50)

In [99]:

```
wind_high
```

Out[99]:

| | 날 짜 | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|------|------------|---------|---------|-----------|---------|---------|-------|
| 977 | 2019-09-07 | 22.6 | 29.5 | 10.0 | 2.8 | 6.0 | 424 |
| 50 | 2017-02-20 | -5.2 | 5.4 | 2.5 | 0.5 | 4.2 | 125 |
| 40 | 2017-02-10 | -9.3 | -2.5 | 0.9 | 0.0 | 4.1 | 64 |
| 460 | 2018-04-07 | 1.4 | 8.6 | 5.1 | 0.2 | 4.1 | 949 |
| 64 | 2017-03-06 | -2.8 | 4.0 | 0.5 | 0.0 | 4.0 | 219 |
| 235 | 2017-08-24 | 23.9 | 28.1 | 7.9 | 14.0 | 4.0 | 794 |
| 65 | 2017-03-07 | -4.0 | 2.9 | 1.0 | 0.0 | 3.9 | 205 |
| 1049 | 2019-11-18 | 0.4 | 12.7 | 7.1 | 14.7 | 3.9 | 538 |
| 47 | 2017-02-17 | -5.2 | 9.0 | 3.1 | 4.0 | 3.9 | 171 |
| 397 | 2018-02-03 | -10.4 | -2.1 | 1.3 | 0.5 | 3.8 | 118 |
| 53 | 2017-02-23 | -4.2 | 2.1 | 1.1 | 0.0 | 3.8 | 113 |
| 13 | 2017-01-14 | -10.5 | -5.4 | 0.0 | 0.0 | 3.8 | 44 |
| 39 | 2017-02-09 | -7.1 | -2.1 | 1.5 | 0.0 | 3.8 | 100 |
| 703 | 2018-12-07 | -9.6 | -1.1 | 1.1 | 0.0 | 3.8 | 175 |
| 423 | 2018-03-01 | -4.3 | 3.0 | 4.4 | 0.5 | 3.8 | 420 |
| 125 | 2017-05-06 | 10.0 | 19.3 | 0.9 | 0.0 | 3.8 | 967 |
| 992 | 2019-09-22 | 17.3 | 19.6 | 9.4 | 0.4 | 3.8 | 1806 |
| 21 | 2017-01-22 | -10.9 | -3.3 | 0.0 | 0.0 | 3.8 | 43 |
| 386 | 2018-01-23 | -14.6 | -5.3 | 0.3 | 0.0 | 3.7 | 69 |
| 227 | 2017-08-16 | 20.5 | 26.8 | 9.6 | 2.0 | 3.7 | 1011 |
| 133 | 2017-05-14 | 11.0 | 20.7 | 0.6 | 0.0 | 3.7 | 3483 |
| 101 | 2017-04-12 | 8.8 | 16.4 | 1.9 | 0.0 | 3.7 | 831 |
| 225 | 2017-08-14 | 23.1 | 27.1 | 9.6 | 1.5 | 3.6 | 946 |
| 226 | 2017-08-15 | 20.8 | 24.0 | 10.0 | 93.5 | 3.6 | 865 |
| 257 | 2017-09-16 | 16.7 | 27.6 | 3.1 | 0.0 | 3.6 | 3929 |
| 256 | 2017-09-15 | 17.8 | 27.9 | 3.9 | 0.0 | 3.6 | 2541 |
| 765 | 2019-02-07 | -8.2 | 6.3 | 4.3 | 0.5 | 3.5 | 437 |
| 1012 | 2019-10-12 | 13.3 | 23.3 | 0.6 | 0.0 | 3.5 | 4367 |
| 103 | 2017-04-14 | 10.1 | 17.7 | 6.9 | 6.0 | 3.5 | 363 |
| 183 | 2017-07-03 | 21.9 | 27.4 | 9.9 | 67.5 | 3.5 | 641 |
| 463 | 2018-04-10 | 6.9 | 21.9 | 6.3 | 5.0 | 3.5 | 1149 |
| 343 | 2017-12-11 | -11.0 | -2.5 | 0.3 | 0.0 | 3.5 | 88 |
| 723 | 2018-12-27 | -12.5 | -6.5 | 0.5 | 0.0 | 3.5 | 159 |
| 387 | 2018-01-24 | -16.3 | -10.7 | 0.0 | 0.0 | 3.4 | 64 |
| 938 | 2019-07-30 | 26.3 | 30.8 | 8.8 | 1.2 | 3.4 | 1905 |
| 320 | 2017-11-18 | -4.8 | 2.6 | 0.0 | 0.0 | 3.4 | 217 |
| 442 | 2018-03-20 | 1.6 | 9.1 | 6.9 | 0.0 | 3.4 | 563 |

~~~~~

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 336 | 2017-12-04 | -6.1 | 4.5  | 1.4   | 0.1  | 3.4  | 205  |
| 867 | 2019-05-20 | 13.0 | 20.2 | 4.5   | 8.7  | 3.4  | 3188 |
| 144 | 2017-05-25 | 14.4 | 24.4 | 0.9   | 0.0  | 3.4  | 1239 |
| 26  | 2017-01-27 | -5.7 | 3.1  | 1.6   | 3.0  | 3.4  | 83   |
| 332 | 2017-11-30 | -5.2 | 2.7  | 2.0   | 0.0  | 3.3  | 208  |
| 799 | 2019-03-13 | -0.4 | 8.5  | 1.8   | 0.0  | 3.3  | 830  |
| 115 | 2017-04-26 | 10.5 | 20.2 | 2.0   | 0.0  | 3.3  | 1061 |
| 159 | 2017-06-09 | 17.9 | 25.7 | 6.9   | 0.1  | 3.3  | 1443 |
| 300 | 2017-10-29 | 6.3  | 17.2 | 0.0   | 0.0  | 3.3  | 1522 |
| 116 | 2017-04-27 | 7.9  | 19.8 | 0.4   | 0.0  | 3.3  | 976  |
| 29  | 2017-01-30 | -7.9 | -0.9 | 2.9   | 0.1  | 3.3  | 67   |
| 8   | 2017-01-09 | -3.1 | 4.3  | 4.3   | 0.0  | 3.3  | 149  |
| 305 | 2017-11-03 | 6.1  | 16.6 | 5.6   | 1.5  | 3.2  | 829  |

```
In [100]: wind_high['가입 수'].mean()
```

```
Out[100]: 854.06
```

```
In [101]: wind_low = df.sort_values(by='평균풍속', ascending=False).tail(50)
```

In [102]:

```
wind_low
```

Out[102]:

|      | 날 짜        | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|------|------------|---------|---------|-----------|---------|---------|-------|
| 504  | 2018-05-22 | 15.4    | 23.2    | 9.0       | 12.5    | 1.1     | 1966  |
| 596  | 2018-08-22 | 23.4    | 37.6    | 4.5       | 0.0     | 1.1     | 1817  |
| 527  | 2018-06-14 | 18.8    | 25.1    | 9.5       | 29.0    | 1.1     | 1147  |
| 923  | 2019-07-15 | 22.1    | 27.4    | 7.5       | 7.7     | 1.1     | 1803  |
| 541  | 2018-06-28 | 20.6    | 27.3    | 9.5       | 26.5    | 1.1     | 1238  |
| 403  | 2018-02-09 | -5.3    | 6.1     | 9.1       | 0.0     | 1.1     | 147   |
| 462  | 2018-04-09 | 0.1     | 15.4    | 2.6       | 0.0     | 1.1     | 1411  |
| 656  | 2018-10-21 | 7.2     | 21.1    | 2.4       | 0.0     | 1.1     | 3960  |
| 658  | 2018-10-23 | 9.8     | 16.9    | 5.4       | 5.0     | 1.1     | 1304  |
| 347  | 2017-12-15 | -7.5    | 0.9     | 7.3       | 0.0     | 1.0     | 113   |
| 620  | 2018-09-15 | 21.0    | 26.0    | 9.8       | 0.2     | 1.0     | 3810  |
| 189  | 2017-07-09 | 23.8    | 30.3    | 9.6       | 32.5    | 1.0     | 655   |
| 681  | 2018-11-15 | 4.7     | 16.2    | 4.8       | 0.0     | 1.0     | 995   |
| 624  | 2018-09-19 | 18.9    | 26.6    | 9.5       | 0.2     | 1.0     | 2112  |
| 657  | 2018-10-22 | 8.8     | 20.6    | 5.3       | 0.0     | 1.0     | 2068  |
| 341  | 2017-12-09 | -5.6    | 5.4     | 2.8       | 0.0     | 1.0     | 223   |
| 738  | 2019-01-11 | -0.4    | 7.2     | 6.1       | 0.0     | 1.0     | 462   |
| 711  | 2018-12-15 | -7.8    | 3.2     | 3.6       | 0.0     | 1.0     | 313   |
| 608  | 2018-09-03 | 21.4    | 25.8    | 9.8       | 34.5    | 1.0     | 913   |
| 966  | 2019-08-27 | 22.2    | 29.3    | 8.1       | 20.0    | 1.0     | 3214  |
| 697  | 2018-12-01 | 0.8     | 12.4    | 3.6       | 0.0     | 1.0     | 1002  |
| 634  | 2018-09-29 | 14.6    | 27.0    | 0.4       | 0.0     | 1.0     | 7170  |
| 692  | 2018-11-26 | 1.6     | 13.6    | 0.6       | 0.0     | 1.0     | 661   |
| 691  | 2018-11-25 | 1.3     | 9.8     | 3.8       | 0.0     | 1.0     | 922   |
| 384  | 2018-01-21 | -3.7    | 5.5     | 2.0       | 0.0     | 1.0     | 266   |
| 760  | 2019-02-02 | -3.2    | 6.1     | 7.5       | 0.0     | 1.0     | 486   |
| 763  | 2019-02-05 | -3.6    | 10.1    | 3.9       | 0.0     | 1.0     | 725   |
| 648  | 2018-10-13 | 6.1     | 19.8    | 0.0       | 0.0     | 1.0     | 3690  |
| 1021 | 2019-10-21 | 10.9    | 24.6    | 2.1       | 0.0     | 1.0     | 2142  |
| 362  | 2017-12-30 | -1.2    | 4.5     | 9.3       | 0.7     | 0.9     | 166   |
| 602  | 2018-08-28 | 22.5    | 26.1    | 9.0       | 96.5    | 0.9     | 611   |
| 346  | 2017-12-14 | -11.2   | -1.9    | 5.9       | 0.0     | 0.9     | 102   |
| 660  | 2018-10-25 | 7.0     | 21.2    | 4.0       | 0.0     | 0.9     | 1932  |
| 741  | 2019-01-14 | -2.4    | 5.3     | 4.5       | 0.0     | 0.9     | 310   |
| 543  | 2018-06-30 | 21.6    | 27.9    | 9.1       | 14.5    | 0.9     | 2094  |
| 495  | 2018-05-12 | 11.8    | 16.1    | 10.0      | 32.0    | 0.9     | 536   |
| 716  | 2018-12-20 | -1.1    | 10.4    | 3.1       | 0.0     | 0.9     | 539   |

-----

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 672 | 2018-11-06 | 8.0  | 17.5 | 6.9   | 0.0  | 0.9  | 1091 |
| 712 | 2018-12-16 | -2.6 | 2.1  | 9.3   | 0.3  | 0.9  | 315  |
| 706 | 2018-12-10 | -8.1 | 4.6  | 3.0   | 0.0  | 0.9  | 231  |
| 846 | 2019-04-29 | 8.3  | 17.3 | 5.8   | 0.0  | 0.9  | 2861 |
| 421 | 2018-02-27 | 0.6  | 6.9  | 8.5   | 0.0  | 0.9  | 369  |
| 370 | 2018-01-07 | -5.7 | 2.8  | 7.5   | 0.0  | 0.8  | 253  |
| 601 | 2018-08-27 | 20.2 | 24.9 | 8.9   | 19.5 | 0.8  | 1295 |
| 544 | 2018-07-01 | 21.0 | 22.8 | 9.9   | 83.5 | 0.8  | 689  |
| 546 | 2018-07-03 | 22.5 | 32.6 | 6.0   | 0.1  | 0.8  | 2254 |
| 376 | 2018-01-13 | -6.6 | -1.2 | 8.3   | 0.4  | 0.8  | 89   |
| 379 | 2018-01-16 | 0.1  | 6.8  | 9.8   | 0.0  | 0.7  | 202  |
| 553 | 2018-07-10 | 20.4 | 26.5 | 9.8   | 16.5 | 0.7  | 1347 |
| 545 | 2018-07-02 | 20.6 | 25.3 | 9.6   | 59.5 | 0.7  | 530  |

```
In [103]: wind_low['가입 수'].mean()
```

```
Out[103]: 1291.02
```

평균 풍속이 높으면 바람이 강해 걷기가 힘들 때도 있고 몸을 가누지 못할 때도 있어서 자전거를 타기 힘든 환경이기 때문에 가입 수가 적을 것이라고 예상하였다. 역시 눈에 띄는 큰 차이는 없지만, 평균값을 비교해보면 평균 풍속이 낮은 쪽이 평균 가입 수가 더 높다.

```
In [104]: rain_high = df.sort_values(by='일강수량', ascending=False).head(50)
```

```
In [105]: rain_high
```

Out[105]:

|     | 날 짜        | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|-----|------------|---------|---------|-----------|---------|---------|-------|
| 190 | 2017-07-10 | 22.6    | 25.4    | 10.0      | 144.5   | 2.3     | 255   |
| 203 | 2017-07-23 | 26.1    | 28.1    | 9.8       | 133.5   | 2.1     | 627   |
| 231 | 2017-08-20 | 22.0    | 25.8    | 10.0      | 124.5   | 2.9     | 309   |
| 602 | 2018-08-28 | 22.5    | 26.1    | 9.0       | 96.5    | 0.9     | 611   |
| 226 | 2017-08-15 | 20.8    | 24.0    | 10.0      | 93.5    | 3.6     | 865   |
| 182 | 2017-07-02 | 21.9    | 26.8    | 10.0      | 92.0    | 2.5     | 513   |
| 544 | 2018-07-01 | 21.0    | 22.8    | 9.9       | 83.5    | 0.8     | 689   |
| 499 | 2018-05-17 | 18.8    | 23.5    | 10.0      | 83.0    | 1.4     | 312   |
| 539 | 2018-06-26 | 20.2    | 26.9    | 10.0      | 71.5    | 2.4     | 635   |
| 183 | 2017-07-03 | 21.9    | 27.4    | 9.9       | 67.5    | 3.5     | 641   |
| 674 | 2018-11-08 | 10.5    | 14.7    | 10.0      | 64.0    | 1.8     | 281   |
| 980 | 2019-09-10 | 22.1    | 26.5    | 9.8       | 63.2    | 2.4     | 792   |
| 939 | 2019-07-31 | 24.0    | 27.5    | 9.8       | 62.3    | 2.2     | 762   |
| 934 | 2019-07-26 | 23.7    | 25.9    | 9.9       | 60.0    | 1.7     | 634   |
| 545 | 2018-07-02 | 20.6    | 25.3    | 9.6       | 59.5    | 0.7     | 530   |
| 476 | 2018-04-23 | 8.4     | 10.7    | 10.0      | 59.0    | 2.2     | 304   |
| 951 | 2019-08-12 | 24.8    | 28.5    | 8.9       | 56.2    | 1.3     | 725   |
| 641 | 2018-10-06 | 16.3    | 21.7    | 7.5       | 55.5    | 2.2     | 3206  |
| 498 | 2018-05-16 | 21.8    | 24.4    | 10.0      | 45.0    | 1.9     | 634   |
| 195 | 2017-07-15 | 23.4    | 27.3    | 10.0      | 42.5    | 2.2     | 423   |
| 603 | 2018-08-29 | 22.7    | 27.4    | 10.0      | 42.0    | 2.4     | 557   |
| 94  | 2017-04-05 | 10.8    | 14.2    | 9.9       | 37.5    | 2.4     | 214   |
| 968 | 2019-08-29 | 20.1    | 26.4    | 5.6       | 36.9    | 2.2     | 2148  |
| 640 | 2018-10-05 | 15.7    | 19.0    | 9.8       | 36.5    | 1.7     | 606   |
| 975 | 2019-09-05 | 21.5    | 27.6    | 9.0       | 36.4    | 2.0     | 772   |
| 885 | 2019-06-07 | 17.0    | 24.7    | 7.4       | 35.3    | 1.9     | 2923  |
| 608 | 2018-09-03 | 21.4    | 25.8    | 9.8       | 34.5    | 1.0     | 913   |
| 186 | 2017-07-06 | 23.9    | 34.6    | 7.1       | 34.0    | 1.7     | 38    |
| 189 | 2017-07-09 | 23.8    | 30.3    | 9.6       | 32.5    | 1.0     | 655   |
| 495 | 2018-05-12 | 11.8    | 16.1    | 10.0      | 32.0    | 0.9     | 536   |
| 234 | 2017-08-23 | 24.6    | 30.2    | 9.4       | 31.5    | 2.6     | 665   |
| 527 | 2018-06-14 | 18.8    | 25.1    | 9.5       | 29.0    | 1.1     | 1147  |
| 176 | 2017-06-26 | 20.1    | 29.9    | 8.8       | 29.0    | 2.2     | 843   |
| 437 | 2018-03-15 | 8.4     | 16.6    | 9.8       | 27.0    | 2.1     | 538   |
| 940 | 2019-08-01 | 23.5    | 30.1    | 7.8       | 26.9    | 1.7     | 1711  |
| 541 | 2018-06-28 | 20.6    | 27.3    | 9.5       | 26.5    | 1.1     | 1238  |
| 936 | 2019-07-28 | 25.0    | 28.2    | 9.8       | 25.6    | 2.2     | 990   |



-----

|      | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|------|------------|------|------|-------|------|------|------|
| 1007 | 2019-10-07 | 13.3 | 17.9 | 9.5   | 25.4 | 1.8  | 842  |
| 422  | 2018-02-28 | 1.8  | 7.8  | 10.0  | 25.0 | 2.0  | 213  |
| 327  | 2017-11-25 | -3.3 | 7.3  | 7.5   | 24.0 | 1.8  | 162  |
| 252  | 2017-09-11 | 18.6 | 24.6 | 7.6   | 24.0 | 1.7  | 1619 |
| 1002 | 2019-10-02 | 20.0 | 24.2 | 9.9   | 23.1 | 3.0  | 795  |
| 974  | 2019-09-04 | 21.2 | 25.5 | 9.3   | 22.7 | 2.5  | 770  |
| 196  | 2017-07-16 | 23.4 | 28.8 | 9.4   | 22.5 | 1.4  | 1184 |
| 866  | 2019-05-19 | 17.9 | 21.6 | 10.0  | 22.0 | 1.2  | 1331 |
| 489  | 2018-05-06 | 15.0 | 20.7 | 9.3   | 22.0 | 1.1  | 2288 |
| 1048 | 2019-11-17 | 5.6  | 10.5 | 9.3   | 21.1 | 3.1  | 623  |
| 356  | 2017-12-24 | -0.6 | 5.7  | 6.6   | 20.0 | 1.9  | 207  |
| 282  | 2017-10-11 | 12.2 | 21.7 | 9.3   | 20.0 | 2.9  | 1049 |
| 966  | 2019-08-27 | 22.2 | 29.3 | 8.1   | 20.0 | 1.0  | 3214 |

```
In [106]: rain_high['가입 수'].mean()
```

```
Out[106]: 890.78
```

일 강수량은 결측치가 너무 많아서 그 결측치를 0으로 처리했기 때문에 제대로 분석을 할 수 없었다. 일 강수량이 적은 쪽과 비교는 할 수 없지만, 가입 수의 평균이 1,000명 미만이기 때문에 적은 편이라는 것을 알 수 있다.

```
In [107]: hightemp_high = df.sort_values(by='최고기온', ascending=False).head(50)
```

In [108]: `hightemp_high`

Out[108]:

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 575 | 2018-08-01 | 27.8 | 39.6 | 0.8   | 0.0  | 1.7  | 1153 |
| 574 | 2018-07-31 | 27.8 | 38.3 | 2.8   | 0.0  | 1.5  | 1463 |
| 589 | 2018-08-15 | 28.3 | 38.0 | 6.4   | 1.0  | 1.2  | 1792 |
| 565 | 2018-07-22 | 25.3 | 38.0 | 3.5   | 0.0  | 1.4  | 1949 |
| 577 | 2018-08-03 | 30.0 | 37.9 | 1.0   | 0.0  | 1.9  | 1192 |
| 576 | 2018-08-02 | 30.3 | 37.9 | 4.0   | 0.0  | 1.6  | 1173 |
| 596 | 2018-08-22 | 23.4 | 37.6 | 4.5   | 0.0  | 1.1  | 1817 |
| 588 | 2018-08-14 | 27.7 | 37.2 | 6.9   | 0.0  | 1.3  | 1581 |
| 573 | 2018-07-30 | 26.2 | 36.9 | 3.1   | 0.0  | 1.2  | 1696 |
| 564 | 2018-07-21 | 24.9 | 36.9 | 0.3   | 0.0  | 1.5  | 2270 |
| 567 | 2018-07-24 | 27.2 | 36.8 | 4.4   | 0.0  | 1.6  | 1524 |
| 945 | 2019-08-06 | 27.9 | 36.8 | 6.8   | 0.0  | 1.5  | 1672 |
| 584 | 2018-08-10 | 25.7 | 36.8 | 3.5   | 0.0  | 1.4  | 1298 |
| 572 | 2018-07-29 | 26.4 | 36.7 | 2.4   | 0.0  | 1.2  | 1918 |
| 949 | 2019-08-10 | 26.6 | 36.5 | 5.1   | 0.0  | 1.5  | 2107 |
| 914 | 2019-07-06 | 24.6 | 36.1 | 3.4   | 0.0  | 1.7  | 3584 |
| 942 | 2019-08-03 | 25.6 | 36.0 | 3.6   | 3.0  | 1.3  | 1893 |
| 585 | 2018-08-11 | 26.4 | 36.0 | 2.9   | 0.0  | 1.5  | 1991 |
| 581 | 2018-08-07 | 26.8 | 35.9 | 3.6   | 0.0  | 1.5  | 1352 |
| 944 | 2019-08-05 | 25.6 | 35.9 | 2.5   | 0.0  | 1.1  | 1706 |
| 953 | 2019-08-14 | 25.9 | 35.7 | 4.0   | 0.0  | 1.5  | 1536 |
| 566 | 2018-07-23 | 29.2 | 35.7 | 7.9   | 0.0  | 1.8  | 1643 |
| 950 | 2019-08-11 | 26.3 | 35.7 | 6.8   | 6.7  | 2.5  | 1782 |
| 590 | 2018-08-16 | 24.1 | 35.6 | 3.3   | 0.0  | 1.7  | 1905 |
| 205 | 2017-07-25 | 24.6 | 35.4 | 4.3   | 0.0  | 1.9  | 949  |
| 570 | 2018-07-27 | 27.2 | 35.4 | 2.3   | 0.0  | 1.9  | 1352 |
| 582 | 2018-08-08 | 27.6 | 35.3 | 3.0   | 0.0  | 2.1  | 1394 |
| 952 | 2019-08-13 | 25.1 | 35.3 | 7.4   | 6.4  | 1.5  | 1528 |
| 586 | 2018-08-12 | 27.2 | 35.3 | 6.5   | 5.0  | 1.5  | 1892 |
| 215 | 2017-08-04 | 25.9 | 35.3 | 6.3   | 0.0  | 2.1  | 952  |
| 580 | 2018-08-06 | 28.3 | 35.3 | 7.6   | 6.5  | 1.6  | 1161 |
| 571 | 2018-07-28 | 28.0 | 35.2 | 5.5   | 7.5  | 1.2  | 1417 |
| 913 | 2019-07-05 | 20.2 | 35.0 | 3.8   | 0.0  | 1.5  | 2912 |
| 927 | 2019-07-19 | 22.8 | 35.0 | 4.6   | 0.0  | 1.5  | 2173 |
| 578 | 2018-08-04 | 28.0 | 34.9 | 3.4   | 0.0  | 2.4  | 1613 |
| 200 | 2017-07-20 | 26.0 | 34.9 | 2.4   | 0.0  | 2.2  | 1024 |
| 216 | 2017-08-05 | 27.6 | 34.8 | 4.1   | 0.0  | 1.9  | 902  |

-----

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 563 | 2018-07-20 | 24.6 | 34.7 | 0.4   | 0.0  | 1.5  | 2069 |
| 948 | 2019-08-09 | 26.0 | 34.6 | 4.9   | 0.0  | 1.3  | 1871 |
| 186 | 2017-07-06 | 23.9 | 34.6 | 7.1   | 34.0 | 1.7  | 38   |
| 218 | 2017-08-07 | 26.4 | 34.4 | 3.3   | 0.0  | 2.4  | 998  |
| 587 | 2018-08-13 | 26.7 | 34.4 | 7.3   | 1.0  | 1.4  | 1510 |
| 943 | 2019-08-04 | 25.0 | 34.4 | 5.8   | 8.6  | 1.2  | 2174 |
| 593 | 2018-08-19 | 23.0 | 34.4 | 7.9   | 0.0  | 1.2  | 4211 |
| 199 | 2017-07-19 | 24.5 | 34.1 | 5.9   | 0.0  | 1.8  | 1056 |
| 562 | 2018-07-19 | 23.2 | 34.1 | 0.1   | 0.0  | 1.6  | 2194 |
| 173 | 2017-06-23 | 21.3 | 34.1 | 4.1   | 0.0  | 2.0  | 1760 |
| 568 | 2018-07-25 | 27.7 | 34.1 | 6.3   | 0.0  | 2.4  | 1606 |
| 217 | 2017-08-06 | 28.2 | 34.0 | 6.9   | 5.0  | 1.9  | 987  |
| 559 | 2018-07-16 | 24.0 | 34.0 | 0.4   | 0.0  | 1.7  | 2161 |

In [109]: hightemp\_high['가입 수'].mean()

Out[109]: 1678.02

In [110]: lowtemp\_high = df.sort\_values(by='최저기온', ascending=False).head(50)

In [111]: lowtemp\_high

Out[111]:

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량  | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|-------|------|------|
| 576 | 2018-08-02 | 30.3 | 37.9 | 4.0   | 0.0   | 1.6  | 1173 |
| 577 | 2018-08-03 | 30.0 | 37.9 | 1.0   | 0.0   | 1.9  | 1192 |
| 566 | 2018-07-23 | 29.2 | 35.7 | 7.9   | 0.0   | 1.8  | 1643 |
| 589 | 2018-08-15 | 28.3 | 38.0 | 6.4   | 1.0   | 1.2  | 1792 |
| 580 | 2018-08-06 | 28.3 | 35.3 | 7.6   | 6.5   | 1.6  | 1161 |
| 217 | 2017-08-06 | 28.2 | 34.0 | 6.9   | 5.0   | 1.9  | 987  |
| 571 | 2018-07-28 | 28.0 | 35.2 | 5.5   | 7.5   | 1.2  | 1417 |
| 578 | 2018-08-04 | 28.0 | 34.9 | 3.4   | 0.0   | 2.4  | 1613 |
| 945 | 2019-08-06 | 27.9 | 36.8 | 6.8   | 0.0   | 1.5  | 1672 |
| 575 | 2018-08-01 | 27.8 | 39.6 | 0.8   | 0.0   | 1.7  | 1153 |
| 574 | 2018-07-31 | 27.8 | 38.3 | 2.8   | 0.0   | 1.5  | 1463 |
| 588 | 2018-08-14 | 27.7 | 37.2 | 6.9   | 0.0   | 1.3  | 1581 |
| 568 | 2018-07-25 | 27.7 | 34.1 | 6.3   | 0.0   | 2.4  | 1606 |
| 582 | 2018-08-08 | 27.6 | 35.3 | 3.0   | 0.0   | 2.1  | 1394 |
| 216 | 2017-08-05 | 27.6 | 34.8 | 4.1   | 0.0   | 1.9  | 902  |
| 202 | 2017-07-22 | 27.5 | 32.5 | 9.5   | 0.5   | 2.3  | 753  |
| 569 | 2018-07-26 | 27.3 | 33.7 | 5.8   | 0.0   | 1.5  | 1592 |
| 201 | 2017-07-21 | 27.3 | 32.4 | 8.6   | 0.0   | 2.5  | 826  |
| 579 | 2018-08-05 | 27.2 | 33.9 | 6.5   | 2.0   | 2.6  | 1636 |
| 214 | 2017-08-03 | 27.2 | 33.9 | 5.6   | 0.0   | 2.2  | 989  |
| 586 | 2018-08-12 | 27.2 | 35.3 | 6.5   | 5.0   | 1.5  | 1892 |
| 570 | 2018-07-27 | 27.2 | 35.4 | 2.3   | 0.0   | 1.9  | 1352 |
| 567 | 2018-07-24 | 27.2 | 36.8 | 4.4   | 0.0   | 1.6  | 1524 |
| 583 | 2018-08-09 | 27.1 | 32.5 | 7.3   | 2.0   | 1.5  | 1024 |
| 581 | 2018-08-07 | 26.8 | 35.9 | 3.6   | 0.0   | 1.5  | 1352 |
| 587 | 2018-08-13 | 26.7 | 34.4 | 7.3   | 1.0   | 1.4  | 1510 |
| 949 | 2019-08-10 | 26.6 | 36.5 | 5.1   | 0.0   | 1.5  | 2107 |
| 213 | 2017-08-02 | 26.6 | 33.0 | 5.1   | 0.0   | 1.3  | 995  |
| 585 | 2018-08-11 | 26.4 | 36.0 | 2.9   | 0.0   | 1.5  | 1991 |
| 572 | 2018-07-29 | 26.4 | 36.7 | 2.4   | 0.0   | 1.2  | 1918 |
| 218 | 2017-08-07 | 26.4 | 34.4 | 3.3   | 0.0   | 2.4  | 998  |
| 938 | 2019-07-30 | 26.3 | 30.8 | 8.8   | 1.2   | 3.4  | 1905 |
| 950 | 2019-08-11 | 26.3 | 35.7 | 6.8   | 6.7   | 2.5  | 1782 |
| 573 | 2018-07-30 | 26.2 | 36.9 | 3.1   | 0.0   | 1.2  | 1696 |
| 203 | 2017-07-23 | 26.1 | 28.1 | 9.8   | 133.5 | 2.1  | 627  |
| 194 | 2017-07-14 | 26.1 | 33.0 | 7.0   | 0.0   | 2.1  | 912  |
| 948 | 2019-08-09 | 26.0 | 34.6 | 4.9   | 0.0   | 1.3  | 1871 |

-----

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 200 | 2017-07-20 | 26.0 | 34.9 | 2.4   | 0.0  | 2.2  | 1024 |
| 215 | 2017-08-04 | 25.9 | 35.3 | 6.3   | 0.0  | 2.1  | 952  |
| 947 | 2019-08-08 | 25.9 | 32.7 | 6.6   | 0.0  | 1.6  | 1871 |
| 953 | 2019-08-14 | 25.9 | 35.7 | 4.0   | 0.0  | 1.5  | 1536 |
| 584 | 2018-08-10 | 25.7 | 36.8 | 3.5   | 0.0  | 1.4  | 1298 |
| 597 | 2018-08-23 | 25.7 | 32.2 | 9.5   | 3.0  | 1.7  | 805  |
| 204 | 2017-07-24 | 25.6 | 27.3 | 9.4   | 3.0  | 1.2  | 810  |
| 937 | 2019-07-29 | 25.6 | 30.0 | 8.0   | 0.5  | 2.6  | 2147 |
| 942 | 2019-08-03 | 25.6 | 36.0 | 3.6   | 3.0  | 1.3  | 1893 |
| 944 | 2019-08-05 | 25.6 | 35.9 | 2.5   | 0.0  | 1.1  | 1706 |
| 933 | 2019-07-25 | 25.6 | 27.4 | 9.9   | 9.7  | 1.2  | 901  |
| 219 | 2017-08-08 | 25.5 | 31.0 | 8.0   | 0.0  | 1.9  | 1332 |
| 565 | 2018-07-22 | 25.3 | 38.0 | 3.5   | 0.0  | 1.4  | 1949 |

```
In [112]: lowtemp_high['가입 수'].mean()
```

```
Out[112]: 1404.5
```

```
In [113]: hightemp_low = df.sort_values(by='최고기온', ascending=False).tail(50)
```

In [114]: hightemp\_low



Out[114]:

|     | 날짜         | 최저기온  | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|------|-------|------|------|------|
| 726 | 2018-12-30 | -11.3 | -0.9 | 0.8   | 0.0  | 1.2  | 223  |
| 29  | 2017-01-30 | -7.9  | -0.9 | 2.9   | 0.1  | 3.3  | 67   |
| 20  | 2017-01-21 | -10.0 | -0.9 | 5.3   | 2.1  | 1.6  | 32   |
| 14  | 2017-01-15 | -11.5 | -0.9 | 0.0   | 0.0  | 2.3  | 89   |
| 351 | 2017-12-19 | -6.5  | -1.0 | 1.0   | 0.0  | 2.1  | 73   |
| 358 | 2017-12-26 | -7.9  | -1.1 | 0.3   | 0.0  | 2.9  | 134  |
| 703 | 2018-12-07 | -9.6  | -1.1 | 1.1   | 0.0  | 3.8  | 175  |
| 743 | 2019-01-16 | -10.1 | -1.1 | 4.1   | 0.0  | 2.3  | 317  |
| 391 | 2018-01-28 | -9.3  | -1.2 | 3.3   | 0.0  | 2.6  | 120  |
| 372 | 2018-01-09 | -6.1  | -1.2 | 3.5   | 0.5  | 2.9  | 147  |
| 376 | 2018-01-13 | -6.6  | -1.2 | 8.3   | 0.4  | 0.8  | 89   |
| 23  | 2017-01-24 | -11.6 | -1.3 | 0.0   | 0.0  | 2.0  | 32   |
| 766 | 2019-02-08 | -10.2 | -1.3 | 6.1   | 0.0  | 2.3  | 328  |
| 19  | 2017-01-20 | -7.8  | -1.4 | 4.8   | 4.0  | 3.0  | 53   |
| 12  | 2017-01-13 | -8.6  | -1.4 | 1.8   | 0.6  | 2.9  | 56   |
| 401 | 2018-02-07 | -13.4 | -1.6 | 1.0   | 0.0  | 1.5  | 92   |
| 406 | 2018-02-12 | -9.6  | -1.8 | 4.1   | 0.0  | 2.7  | 109  |
| 346 | 2017-12-14 | -11.2 | -1.9 | 5.9   | 0.0  | 0.9  | 102  |
| 767 | 2019-02-09 | -8.6  | -1.9 | 4.4   | 0.0  | 2.7  | 376  |
| 768 | 2019-02-10 | -9.1  | -1.9 | 7.5   | 0.0  | 1.7  | 415  |
| 705 | 2018-12-09 | -11.8 | -2.1 | 2.0   | 0.0  | 1.5  | 197  |
| 397 | 2018-02-03 | -10.4 | -2.1 | 1.3   | 0.5  | 3.8  | 118  |
| 39  | 2017-02-09 | -7.1  | -2.1 | 1.5   | 0.0  | 3.8  | 100  |
| 405 | 2018-02-11 | -9.5  | -2.2 | 3.5   | 0.0  | 2.7  | 111  |
| 349 | 2017-12-17 | -10.8 | -2.3 | 1.1   | 0.0  | 1.9  | 109  |
| 40  | 2017-02-10 | -9.3  | -2.5 | 0.9   | 0.0  | 4.1  | 64   |
| 343 | 2017-12-11 | -11.0 | -2.5 | 0.3   | 0.0  | 3.5  | 88   |
| 359 | 2017-12-27 | -10.3 | -2.6 | 0.0   | 0.0  | 1.9  | 111  |
| 348 | 2017-12-16 | -8.1  | -2.6 | 0.0   | 0.0  | 3.0  | 124  |
| 725 | 2018-12-29 | -10.5 | -3.0 | 0.0   | 0.0  | 2.9  | 168  |
| 21  | 2017-01-22 | -10.9 | -3.3 | 0.0   | 0.0  | 3.8  | 43   |
| 390 | 2018-01-27 | -15.9 | -3.5 | 1.8   | 0.0  | 1.4  | 99   |
| 22  | 2017-01-23 | -12.6 | -4.1 | 0.0   | 0.0  | 2.2  | 34   |
| 345 | 2017-12-13 | -12.0 | -4.4 | 0.0   | 0.0  | 1.7  | 81   |
| 704 | 2018-12-08 | -11.4 | -4.4 | 1.4   | 0.0  | 2.6  | 187  |
| 392 | 2018-01-29 | -11.6 | -4.7 | 0.1   | 0.0  | 2.7  | 86   |
| 373 | 2018-01-10 | -10.3 | -4.8 | 0.9   | 0.3  | 3.0  | 93   |

-----

|     | 날짜         | 최저기온  | 최고기온  | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|-------|-------|------|------|------|
| 399 | 2018-02-05 | -11.8 | -5.1  | 0.1   | 0.0  | 3.0  | 89   |
| 344 | 2017-12-12 | -12.3 | -5.2  | 0.0   | 0.0  | 2.1  | 83   |
| 398 | 2018-02-04 | -12.8 | -5.2  | 0.0   | 0.0  | 3.1  | 105  |
| 386 | 2018-01-23 | -14.6 | -5.3  | 0.3   | 0.0  | 3.7  | 69   |
| 375 | 2018-01-12 | -15.3 | -5.4  | 2.3   | 0.0  | 1.7  | 79   |
| 13  | 2017-01-14 | -10.5 | -5.4  | 0.0   | 0.0  | 3.8  | 44   |
| 400 | 2018-02-06 | -12.8 | -5.8  | 1.1   | 0.0  | 3.1  | 67   |
| 723 | 2018-12-27 | -12.5 | -6.5  | 0.5   | 0.0  | 3.5  | 159  |
| 724 | 2018-12-28 | -14.4 | -6.6  | 0.6   | 0.0  | 2.5  | 150  |
| 374 | 2018-01-11 | -13.6 | -7.4  | 0.4   | 0.0  | 2.8  | 76   |
| 388 | 2018-01-25 | -16.4 | -9.5  | 3.6   | 0.0  | 2.8  | 56   |
| 387 | 2018-01-24 | -16.3 | -10.7 | 0.0   | 0.0  | 3.4  | 64   |
| 389 | 2018-01-26 | -17.8 | -10.7 | 0.0   | 0.0  | 2.6  | 42   |

```
In [115]: hightemp_low['가입 수'].mean()
```

```
Out[115]: 118.5
```

```
In [116]: lowtemp_low = df.sort_values(by='최저기온', ascending=False).tail(50)
```

In [117]:

```
lowtemp_low
```

Out[117]:

|     | 날짜         | 최저기온  | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|------|-------|------|------|------|
| 15  | 2017-01-16 | -8.9  | 4.0  | 0.1   | 0.0  | 1.4  | 105  |
| 768 | 2019-02-10 | -9.1  | -1.9 | 7.5   | 0.0  | 1.7  | 415  |
| 41  | 2017-02-11 | -9.2  | 0.8  | 0.3   | 0.0  | 2.8  | 90   |
| 32  | 2017-02-02 | -9.2  | 3.3  | 1.3   | 0.0  | 2.7  | 119  |
| 391 | 2018-01-28 | -9.3  | -1.2 | 3.3   | 0.0  | 2.6  | 120  |
| 40  | 2017-02-10 | -9.3  | -2.5 | 0.9   | 0.0  | 4.1  | 64   |
| 10  | 2017-01-11 | -9.4  | 1.5  | 1.9   | 0.0  | 2.1  | 88   |
| 736 | 2019-01-09 | -9.4  | 1.3  | 4.9   | 0.0  | 1.3  | 305  |
| 405 | 2018-02-11 | -9.5  | -2.2 | 3.5   | 0.0  | 2.7  | 111  |
| 406 | 2018-02-12 | -9.6  | -1.8 | 4.1   | 0.0  | 2.7  | 109  |
| 703 | 2018-12-07 | -9.6  | -1.1 | 1.1   | 0.0  | 3.8  | 175  |
| 402 | 2018-02-08 | -9.7  | 2.0  | 0.3   | 0.0  | 1.2  | 150  |
| 30  | 2017-01-31 | -9.9  | 0.2  | 3.1   | 0.0  | 1.7  | 75   |
| 20  | 2017-01-21 | -10.0 | -0.9 | 5.3   | 2.1  | 1.6  | 32   |
| 743 | 2019-01-16 | -10.1 | -1.1 | 4.1   | 0.0  | 2.3  | 317  |
| 24  | 2017-01-25 | -10.2 | 0.9  | 0.0   | 0.0  | 1.5  | 72   |
| 766 | 2019-02-08 | -10.2 | -1.3 | 6.1   | 0.0  | 2.3  | 328  |
| 373 | 2018-01-10 | -10.3 | -4.8 | 0.9   | 0.3  | 3.0  | 93   |
| 359 | 2017-12-27 | -10.3 | -2.6 | 0.0   | 0.0  | 1.9  | 111  |
| 397 | 2018-02-03 | -10.4 | -2.1 | 1.3   | 0.5  | 3.8  | 118  |
| 725 | 2018-12-29 | -10.5 | -3.0 | 0.0   | 0.0  | 2.9  | 168  |
| 13  | 2017-01-14 | -10.5 | -5.4 | 0.0   | 0.0  | 3.8  | 44   |
| 727 | 2018-12-31 | -10.5 | -0.5 | 1.6   | 0.0  | 1.2  | 245  |
| 349 | 2017-12-17 | -10.8 | -2.3 | 1.1   | 0.0  | 1.9  | 109  |
| 21  | 2017-01-22 | -10.9 | -3.3 | 0.0   | 0.0  | 3.8  | 43   |
| 343 | 2017-12-11 | -11.0 | -2.5 | 0.3   | 0.0  | 3.5  | 88   |
| 346 | 2017-12-14 | -11.2 | -1.9 | 5.9   | 0.0  | 0.9  | 102  |
| 726 | 2018-12-30 | -11.3 | -0.9 | 0.8   | 0.0  | 1.2  | 223  |
| 393 | 2018-01-30 | -11.4 | -0.8 | 5.8   | 2.9  | 1.3  | 79   |
| 704 | 2018-12-08 | -11.4 | -4.4 | 1.4   | 0.0  | 2.6  | 187  |
| 14  | 2017-01-15 | -11.5 | -0.9 | 0.0   | 0.0  | 2.3  | 89   |
| 23  | 2017-01-24 | -11.6 | -1.3 | 0.0   | 0.0  | 2.0  | 32   |
| 392 | 2018-01-29 | -11.6 | -4.7 | 0.1   | 0.0  | 2.7  | 86   |
| 399 | 2018-02-05 | -11.8 | -5.1 | 0.1   | 0.0  | 3.0  | 89   |
| 705 | 2018-12-09 | -11.8 | -2.1 | 2.0   | 0.0  | 1.5  | 197  |
| 345 | 2017-12-13 | -12.0 | -4.4 | 0.0   | 0.0  | 1.7  | 81   |
| 344 | 2017-12-12 | -12.3 | -5.2 | 0.0   | 0.0  | 2.1  | 83   |

-----

|     | 날짜         | 최저기온  | 최고기온  | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|-------|-------|------|------|------|
| 723 | 2018-12-27 | -12.5 | -6.5  | 0.5   | 0.0  | 3.5  | 159  |
| 22  | 2017-01-23 | -12.6 | -4.1  | 0.0   | 0.0  | 2.2  | 34   |
| 400 | 2018-02-06 | -12.8 | -5.8  | 1.1   | 0.0  | 3.1  | 67   |
| 398 | 2018-02-04 | -12.8 | -5.2  | 0.0   | 0.0  | 3.1  | 105  |
| 401 | 2018-02-07 | -13.4 | -1.6  | 1.0   | 0.0  | 1.5  | 92   |
| 374 | 2018-01-11 | -13.6 | -7.4  | 0.4   | 0.0  | 2.8  | 76   |
| 724 | 2018-12-28 | -14.4 | -6.6  | 0.6   | 0.0  | 2.5  | 150  |
| 386 | 2018-01-23 | -14.6 | -5.3  | 0.3   | 0.0  | 3.7  | 69   |
| 375 | 2018-01-12 | -15.3 | -5.4  | 2.3   | 0.0  | 1.7  | 79   |
| 390 | 2018-01-27 | -15.9 | -3.5  | 1.8   | 0.0  | 1.4  | 99   |
| 387 | 2018-01-24 | -16.3 | -10.7 | 0.0   | 0.0  | 3.4  | 64   |
| 388 | 2018-01-25 | -16.4 | -9.5  | 3.6   | 0.0  | 2.8  | 56   |
| 389 | 2018-01-26 | -17.8 | -10.7 | 0.0   | 0.0  | 2.6  | 42   |

```
In [118]: lowtemp_low['가입 수'].mean()
```

```
Out[118]: 120.68
```

최고 기온과 최저 기온이 너무 높으면 가입 수가 적을 줄 알았지만, 가입 수가 꽤 있는 편이고, 최고 기온과 최저 기온이 너무 낮으면 가입 수가 적다.

```
In [119]: head = df.sort_values(by='가입 수', ascending=False).head(43)
```

In [120]:

```
head
```

Out[120]:

|     | 날 짜        | 최 저 기 온 | 최 고 기 온 | 평 균 전 운 량 | 일 강 수 량 | 평 균 풍 속 | 가 입 수 |
|-----|------------|---------|---------|-----------|---------|---------|-------|
| 853 | 2019-05-06 | 9.8     | 19.8    | 0.3       | 0.0     | 3.1     | 8525  |
| 852 | 2019-05-05 | 13.6    | 26.5    | 5.1       | 0.0     | 2.2     | 8006  |
| 859 | 2019-05-12 | 14.7    | 28.9    | 4.3       | 0.0     | 1.5     | 7396  |
| 858 | 2019-05-11 | 13.2    | 28.3    | 2.0       | 0.0     | 1.8     | 7220  |
| 634 | 2018-09-29 | 14.6    | 27.0    | 0.4       | 0.0     | 1.0     | 7170  |
| 631 | 2018-09-26 | 16.2    | 24.8    | 5.0       | 0.0     | 1.7     | 7149  |
| 851 | 2019-05-04 | 12.0    | 27.4    | 4.9       | 0.0     | 1.7     | 7015  |
| 630 | 2018-09-25 | 10.8    | 24.3    | 2.1       | 0.0     | 1.1     | 6973  |
| 638 | 2018-10-03 | 11.2    | 24.2    | 0.0       | 0.0     | 1.3     | 6966  |
| 848 | 2019-05-01 | 11.2    | 22.2    | 1.9       | 0.0     | 2.3     | 6760  |
| 824 | 2019-04-07 | 5.6     | 20.3    | 4.4       | 0.0     | 1.7     | 6689  |
| 865 | 2019-05-18 | 19.1    | 25.9    | 8.5       | 0.0     | 2.1     | 6520  |
| 894 | 2019-06-16 | 16.1    | 27.2    | 3.9       | 0.0     | 1.6     | 6425  |
| 613 | 2018-09-08 | 15.9    | 27.7    | 0.3       | 0.0     | 1.7     | 6355  |
| 614 | 2018-09-09 | 15.6    | 28.9    | 1.4       | 0.0     | 1.2     | 6298  |
| 844 | 2019-04-27 | 4.1     | 18.2    | 4.8       | 0.0     | 1.7     | 6294  |
| 971 | 2019-09-01 | 19.2    | 27.8    | 6.1       | 0.0     | 1.4     | 6256  |
| 873 | 2019-05-26 | 18.6    | 31.1    | 6.0       | 0.0     | 1.8     | 6237  |
| 970 | 2019-08-31 | 18.5    | 28.9    | 2.3       | 0.0     | 1.9     | 6127  |
| 642 | 2018-10-07 | 12.3    | 21.1    | 2.9       | 0.0     | 1.7     | 6099  |
| 886 | 2019-06-08 | 16.0    | 27.6    | 5.9       | 0.0     | 1.4     | 5916  |
| 880 | 2019-06-02 | 14.1    | 27.8    | 0.6       | 0.0     | 1.7     | 5877  |
| 879 | 2019-06-01 | 14.8    | 23.7    | 5.5       | 0.0     | 1.9     | 5874  |
| 635 | 2018-09-30 | 14.0    | 21.6    | 0.6       | 0.0     | 2.3     | 5862  |
| 900 | 2019-06-22 | 17.7    | 30.6    | 2.1       | 0.0     | 1.5     | 5830  |
| 901 | 2019-06-23 | 19.1    | 31.0    | 3.0       | 0.0     | 1.8     | 5721  |
| 998 | 2019-09-28 | 19.4    | 29.2    | 3.8       | 0.0     | 1.7     | 5721  |
| 872 | 2019-05-25 | 18.4    | 28.6    | 7.9       | 0.0     | 1.6     | 5715  |
| 551 | 2018-07-08 | 17.7    | 28.1    | 7.0       | 0.0     | 1.5     | 5706  |
| 999 | 2019-09-29 | 17.9    | 28.6    | 0.3       | 0.0     | 1.7     | 5664  |
| 830 | 2019-04-13 | 6.2     | 17.9    | 5.6       | 0.0     | 1.8     | 5648  |
| 838 | 2019-04-21 | 12.2    | 20.2    | 7.9       | 0.3     | 1.9     | 5605  |
| 985 | 2019-09-15 | 19.5    | 30.0    | 5.5       | 0.0     | 1.7     | 5576  |
| 964 | 2019-08-25 | 20.2    | 29.9    | 3.4       | 0.0     | 1.5     | 5437  |
| 550 | 2018-07-07 | 19.5    | 29.2    | 3.6       | 0.0     | 1.6     | 5428  |
| 629 | 2018-09-24 | 14.2    | 22.1    | 1.1       | 0.0     | 2.1     | 5406  |
| 983 | 2019-09-13 | 18.3    | 29.3    | 4.3       | 0.1     | 1.7     | 5360  |

-----

|      | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|------|------------|------|------|-------|------|------|------|
| 963  | 2019-08-24 | 21.5 | 27.3 | 5.5   | 0.0  | 1.5  | 5286 |
| 644  | 2018-10-09 | 10.7 | 19.8 | 9.0   | 0.0  | 1.1  | 5278 |
| 837  | 2019-04-20 | 9.7  | 19.1 | 7.1   | 1.2  | 1.9  | 5178 |
| 1003 | 2019-10-03 | 19.8 | 27.3 | 6.3   | 4.6  | 2.8  | 5045 |
| 1009 | 2019-10-09 | 7.7  | 20.4 | 1.9   | 0.0  | 1.7  | 5011 |
| 845  | 2019-04-28 | 9.9  | 16.6 | 8.1   | 0.0  | 1.6  | 5006 |

가입 수가 많은 쪽을 보면 주로 주말이나 공휴일이다. 주로 너무 덥지도 않고 춥지도 않은 4, 5, 9, 10월이 많다.

```
In [121]: head[head['평균전운량'] > 6]
```

Out[121]:

|      | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|------|------------|------|------|-------|------|------|------|
| 865  | 2019-05-18 | 19.1 | 25.9 | 8.5   | 0.0  | 2.1  | 6520 |
| 971  | 2019-09-01 | 19.2 | 27.8 | 6.1   | 0.0  | 1.4  | 6256 |
| 872  | 2019-05-25 | 18.4 | 28.6 | 7.9   | 0.0  | 1.6  | 5715 |
| 551  | 2018-07-08 | 17.7 | 28.1 | 7.0   | 0.0  | 1.5  | 5706 |
| 838  | 2019-04-21 | 12.2 | 20.2 | 7.9   | 0.3  | 1.9  | 5605 |
| 644  | 2018-10-09 | 10.7 | 19.8 | 9.0   | 0.0  | 1.1  | 5278 |
| 837  | 2019-04-20 | 9.7  | 19.1 | 7.1   | 1.2  | 1.9  | 5178 |
| 1003 | 2019-10-03 | 19.8 | 27.3 | 6.3   | 4.6  | 2.8  | 5045 |
| 845  | 2019-04-28 | 9.9  | 16.6 | 8.1   | 0.0  | 1.6  | 5006 |

```
In [122]: head['평균전운량'].mean()
```

Out[122]: 4.013953488372093

```
In [123]: head['평균풍속'].mean()
```

Out[123]: 1.7325581395348837

한편, 평균 전운량 즉, 구름이 많아서 흐린 것과 가입 수는 밀접한 관계가 없는 것처럼 보인다. 필자는 평균 전운량이 6 이상이면 구름이 많다고 판단했는데, 평균 전운량이 6 이상인 요소가 엄청 많지는 않지만 43개의 요소 중에 9개가 있었다. 일 강수량은 대부분 0이고 평균 풍속의 평균은 1.7 정도로 수치가 낮은 편이다.

```
In [124]: tail = df.sort_values(by='가입 수', ascending=False).tail(43)
```



In [125]:

```
tail
```

Out[125]:

|     | 날짜         | 최저기온  | 최고기온  | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|-------|-------|------|------|------|
| 390 | 2018-01-27 | -15.9 | -3.5  | 1.8   | 0.0  | 1.4  | 99   |
| 35  | 2017-02-05 | -1.5  | 2.3   | 7.3   | 0.5  | 2.1  | 93   |
| 373 | 2018-01-10 | -10.3 | -4.8  | 0.9   | 0.3  | 3.0  | 93   |
| 401 | 2018-02-07 | -13.4 | -1.6  | 1.0   | 0.0  | 1.5  | 92   |
| 9   | 2017-01-10 | -7.4  | 1.1   | 0.8   | 0.0  | 3.2  | 92   |
| 41  | 2017-02-11 | -9.2  | 0.8   | 0.3   | 0.0  | 2.8  | 90   |
| 399 | 2018-02-05 | -11.8 | -5.1  | 0.1   | 0.0  | 3.0  | 89   |
| 14  | 2017-01-15 | -11.5 | -0.9  | 0.0   | 0.0  | 2.3  | 89   |
| 376 | 2018-01-13 | -6.6  | -1.2  | 8.3   | 0.4  | 0.8  | 89   |
| 343 | 2017-12-11 | -11.0 | -2.5  | 0.3   | 0.0  | 3.5  | 88   |
| 10  | 2017-01-11 | -9.4  | 1.5   | 1.9   | 0.0  | 2.1  | 88   |
| 352 | 2017-12-20 | -8.5  | -0.1  | 7.1   | 0.2  | 1.1  | 86   |
| 392 | 2018-01-29 | -11.6 | -4.7  | 0.1   | 0.0  | 2.7  | 86   |
| 26  | 2017-01-27 | -5.7  | 3.1   | 1.6   | 3.0  | 3.4  | 83   |
| 344 | 2017-12-12 | -12.3 | -5.2  | 0.0   | 0.0  | 2.1  | 83   |
| 345 | 2017-12-13 | -12.0 | -4.4  | 0.0   | 0.0  | 1.7  | 81   |
| 16  | 2017-01-17 | -7.1  | 3.4   | 4.4   | 0.0  | 1.5  | 80   |
| 11  | 2017-01-12 | -6.1  | 1.0   | 2.1   | 0.0  | 3.1  | 80   |
| 375 | 2018-01-12 | -15.3 | -5.4  | 2.3   | 0.0  | 1.7  | 79   |
| 393 | 2018-01-30 | -11.4 | -0.8  | 5.8   | 2.9  | 1.3  | 79   |
| 52  | 2017-02-22 | -0.9  | 2.9   | 9.9   | 5.1  | 2.7  | 77   |
| 350 | 2017-12-18 | -5.9  | 3.3   | 6.8   | 3.4  | 1.2  | 76   |
| 374 | 2018-01-11 | -13.6 | -7.4  | 0.4   | 0.0  | 2.8  | 76   |
| 30  | 2017-01-31 | -9.9  | 0.2   | 3.1   | 0.0  | 1.7  | 75   |
| 351 | 2017-12-19 | -6.5  | -1.0  | 1.0   | 0.0  | 2.1  | 73   |
| 24  | 2017-01-25 | -10.2 | 0.9   | 0.0   | 0.0  | 1.5  | 72   |
| 25  | 2017-01-26 | -7.9  | 3.4   | 6.9   | 0.3  | 2.2  | 70   |
| 386 | 2018-01-23 | -14.6 | -5.3  | 0.3   | 0.0  | 3.7  | 69   |
| 29  | 2017-01-30 | -7.9  | -0.9  | 2.9   | 0.1  | 3.3  | 67   |
| 400 | 2018-02-06 | -12.8 | -5.8  | 1.1   | 0.0  | 3.1  | 67   |
| 387 | 2018-01-24 | -16.3 | -10.7 | 0.0   | 0.0  | 3.4  | 64   |
| 40  | 2017-02-10 | -9.3  | -2.5  | 0.9   | 0.0  | 4.1  | 64   |
| 12  | 2017-01-13 | -8.6  | -1.4  | 1.8   | 0.6  | 2.9  | 56   |
| 388 | 2018-01-25 | -16.4 | -9.5  | 3.6   | 0.0  | 2.8  | 56   |
| 19  | 2017-01-20 | -7.8  | -1.4  | 4.8   | 4.0  | 3.0  | 53   |
| 28  | 2017-01-29 | -0.9  | 2.5   | 9.6   | 4.5  | 1.8  | 51   |
| 13  | 2017-01-14 | -10.5 | -5.4  | 0.0   | 0.0  | 3.8  | 44   |

-----

|     | 날짜         | 최저기온  | 최고기온  | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|-------|-------|-------|------|------|------|
| 21  | 2017-01-22 | -10.9 | -3.3  | 0.0   | 0.0  | 3.8  | 43   |
| 389 | 2018-01-26 | -17.8 | -10.7 | 0.0   | 0.0  | 2.6  | 42   |
| 186 | 2017-07-06 | 23.9  | 34.6  | 7.1   | 34.0 | 1.7  | 38   |
| 22  | 2017-01-23 | -12.6 | -4.1  | 0.0   | 0.0  | 2.2  | 34   |
| 23  | 2017-01-24 | -11.6 | -1.3  | 0.0   | 0.0  | 2.0  | 32   |
| 20  | 2017-01-21 | -10.0 | -0.9  | 5.3   | 2.1  | 1.6  | 32   |

가입 수가 적은 쪽을 보면, 주로 겨울인 12, 1, 2월이며, 강수량의 수치가 높은 7월도 한 개 있다.

In [126]: `tail[tail['평균전운량'] > 6]`

Out[126]:

|     | 날짜         | 최저기온 | 최고기온 | 평균전운량 | 일강수량 | 평균풍속 | 가입 수 |
|-----|------------|------|------|-------|------|------|------|
| 35  | 2017-02-05 | -1.5 | 2.3  | 7.3   | 0.5  | 2.1  | 93   |
| 376 | 2018-01-13 | -6.6 | -1.2 | 8.3   | 0.4  | 0.8  | 89   |
| 352 | 2017-12-20 | -8.5 | -0.1 | 7.1   | 0.2  | 1.1  | 86   |
| 52  | 2017-02-22 | -0.9 | 2.9  | 9.9   | 5.1  | 2.7  | 77   |
| 350 | 2017-12-18 | -5.9 | 3.3  | 6.8   | 3.4  | 1.2  | 76   |
| 25  | 2017-01-26 | -7.9 | 3.4  | 6.9   | 0.3  | 2.2  | 70   |
| 28  | 2017-01-29 | -0.9 | 2.5  | 9.6   | 4.5  | 1.8  | 51   |
| 186 | 2017-07-06 | 23.9 | 34.6 | 7.1   | 34.0 | 1.7  | 38   |

In [127]: `tail['평균전운량'].mean()`

Out[127]: 2.5953488372093014

In [128]: `tail['평균풍속'].mean()`

Out[128]: 2.4255813953488374

한편, 평균 전운량과 가입 수는 밀접한 관계가 없는 것처럼 보이고, '평균전운량'의 평균이 약 2.6으로 낮은 편이다. 오히려 가입 수가 많은 데이터의 평균이 약 4로 더 높다. 그리고 평균 풍속도 큰 차이가 없는데 가입 수가 많은 데이터의 평균은 약 1.7, 가입 수가 적은 쪽은 2.4이다. 가입 수가 적은 데이터가 조금 높다.

## 5. 분석 결론

평균 전운량과 평균 풍속은 공공자전거의 신규 가입자 수에 큰 영향을 주지는 않지만, 어느 정도는 영향을 준다. 또한, 일 강수량이 높아도 가입자 수에 어느 정도 영향을 준다. 날씨 중에서 가장 큰 영향을 주는 요소는 기온이다. 기온이 많이 떨어지는 날씨는 가입자 수에 영향을 많이 주고 기온이 너무 높으면 봄, 가을 날씨에 비교해서 가입자 수가 떨어진다. 따라서 평균 전운량과 평균 풍속, 일 강수량은 많은 영향을 주지는 않고 기온은 많은 영향을 준다. 즉, 봄, 가을 기온에 가입자 수가 많다.