

▼ 2. 서울시 범죄 현황 분석

▼ 2.1 데이터 로드

```
import pandas as pd
```

```
df = pd.read_csv("/content/drive/MyDrive/스터디/데이터 주무르기/data/02_경찰청 서울특별시경찰청_경;
df
```

	구분	죄종	발생검거	건수
0	중부	살인	발생	1
1	중부	살인	검거	1
2	중부	강도	발생	3
3	중부	강도	검거	4
4	중부	강간,추행	발생	113
...
305	수서	강간,추행	검거	194
306	수서	절도	발생	1010
307	수서	절도	검거	475
308	수서	폭력	발생	1394
309	수서	폭력	검거	1200

310 rows × 4 columns

```
df["col"] = df["죄종"] + " " + df["발생검거"]
df
```

	구분	죄종	발생	검거	건수	col
0	중부	살인	발생	1	살인 발생	
1	중부	살인	검거	1	살인 검거	
2	중부	강도	발생	3	강도 발생	
3	중부	강도	검거	4	강도 검거	

데이터 프레임 재설정

```
df_1 = pd.pivot_table(data = df, values = "건수", index = "구분", columns="col")
df_1
```

coi	강간 검거	강간 발생	강간, 추행 검거	강간, 추행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생
구분												
강남	NaN	NaN	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0
강동	NaN	NaN	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0
강북	NaN	NaN	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0
강서	NaN	NaN	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0
관악	NaN	NaN	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0
광진	NaN	NaN	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0
구로	NaN	NaN	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0
금천	NaN	NaN	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0
남대문	NaN	NaN	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0
노원	NaN	NaN	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0
도봉	87.0	82.0	NaN	NaN	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0
동대문	NaN	NaN	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0
동작	NaN	NaN	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0
마포	NaN	NaN	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0
방배	50.0	63.0	NaN	NaN	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0
서대	NaN	NaN	149.0	159.0	1.0	2.0	3.0	3.0	524.0	1077.0	1029.0	1292.0

```
# 결측값 0으로 처리
df_1 = df_1.fillna(0)
df_1
```

coi	강 간 검 거	강 간 발 생	강 간, 추 행 검 거	강 간, 추 행 발 생	강 도 검 거	강 도 발 생	살 인 검 거	살 인 발 생	절 도 검 거	절 도 발 생	폭 력 검 거	폭 력 발 생
구 분												
강 남	0.0	0.0	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0
강 동	0.0	0.0	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0
강 북	0.0	0.0	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0
강 서	0.0	0.0	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0
관 악	0.0	0.0	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0
광 진	0.0	0.0	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0
구 로	0.0	0.0	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0
금 천	0.0	0.0	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0
남 대 문	0.0	0.0	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0
노 원	0.0	0.0	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0
도 봉	87.0	82.0	0.0	0.0	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0
동 대 문	0.0	0.0	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0
동 작	0.0	0.0	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0
마 포	0.0	0.0	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0
방 배	50.0	63.0	0.0	0.0	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0
서 대 문	0.0	0.0	149.0	159.0	1.0	2.0	3.0	3.0	524.0	1077.0	1029.0	1292.0
서 부	0.0	0.0	71.0	76.0	2.0	3.0	3.0	3.0	322.0	532.0	667.0	748.0
서	0.0	0.0	252.0	222.0	2.0	5.0	1.0	1.0	221.0	1672.0	1445.0	1752.0

초	0.0	0.0	253.0	333.0	3.0	5.0	4.0	4.0	801.0	1673.0	1415.0	1750.0
성동	0.0	0.0	98.0	115.0	3.0	3.0	5.0	5.0	565.0	1016.0	1103.0	1223.0
성북	77.0	95.0	0.0	0.0	2.0	2.0	0.0	0.0	233.0	439.0	610.0	672.0
송파	0.0	0.0	235.0	287.0	17.0	13.0	6.0	6.0	1006.0	2429.0	2280.0	2675.0
수서	0.0	0.0	194.0	195.0	1.0	2.0	2.0	1.0	475.0	1010.0	1200.0	1394.0
양천	90.0	109.0	0.0	0.0	3.0	3.0	5.0	5.0	718.0	1517.0	1270.0	1582.0
영등포	0.0	0.0	252.0	315.0	5.0	6.0	4.0	7.0	953.0	2188.0	2157.0	2701.0
용산	0.0	0.0	186.0	230.0	6.0	6.0	3.0	5.0	534.0	1111.0	1327.0	1617.0
이평	0.0	0.0	74.0	95.0	1.0	1.0	1.0	1.0	434.0	839.0	956.0	1092.0
종로	0.0	0.0	93.0	115.0	4.0	4.0	6.0	9.0	284.0	602.0	792.0	964.0
종암	44.0	47.0	0.0	0.0	0.0	1.0	4.0	4.0	347.0	549.0	630.0	758.0
중앙	0.0	0.0	153.0	183.0	6.0	5.0	9.0	10.0	889.0	1506.0	1724.0	2022.0
중부	0.0	0.0	76.0	113.0	4.0	3.0	1.0	1.0	495.0	943.0	830.0	997.0
혜	0.0	0.0	53.0	77.0	0.0	0.0	0.0	0.0	310.0	504.0	600.0	717.0

강간추행 통합

df_1["강간, 추행 검거"] = df_1["강간, 추행 검거"] + df_1["강간 검거"]

df_1["강간, 추행 발생"] = df_1["강간, 추행 발생"] + df_1["강간 발생"]

df_1 = df_1.drop(["강간 검거", "강간 발생"], axis = 1)

df_1

coi	강간, 추 행 검거	강간, 추 행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생
구 분										
강남	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0
강동	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0
강북	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0
강서	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0
관악	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0
광진	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0
구로	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0
금천	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0
남대문	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0
노원	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0
도봉	87.0	82.0	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0
동대문	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0
동작	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0
마포	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0
방배	50.0	63.0	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0
서										

```
df_r = df_1.reset_index()
df_r
```

col	구분	강간, 추행 검거	강간, 추행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생
0	강남	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0
1	강동	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0
2	강북	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0
3	강서	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0
4	관악	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0
5	광진	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0
6	구로	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0
7	금천	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0
8	남대문	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0
9	노원	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0
10	도봉	87.0	82.0	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0
11	동대문	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0
12	동작	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0
13	마포	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0
14	방배	50.0	63.0	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0
15	서대문	149.0	159.0	1.0	2.0	3.0	3.0	524.0	1077.0	1029.0	1292.0
16	서부	71.0	76.0	2.0	3.0	3.0	3.0	322.0	532.0	667.0	748.0
17	서초	253.0	333.0	3.0	5.0	4.0	4.0	801.0	1673.0	1415.0	1750.0
18	성동	98.0	115.0	3.0	3.0	5.0	5.0	565.0	1016.0	1103.0	1223.0

19	성북	77.0	95.0	2.0	2.0	0.0	0.0	233.0	439.0	610.0	672.0
20	송파	235.0	287.0	17.0	13.0	6.0	6.0	1006.0	2429.0	2280.0	2675.0
21	수서	194.0	195.0	1.0	2.0	2.0	1.0	475.0	1010.0	1200.0	1394.0
22	양천	90.0	109.0	3.0	3.0	5.0	5.0	718.0	1517.0	1270.0	1582.0
23	영등포	252.0	315.0	5.0	6.0	4.0	7.0	953.0	2188.0	2157.0	2701.0
24	용산	186.0	230.0	6.0	6.0	3.0	5.0	534.0	1111.0	1327.0	1617.0
25	은평	74.0	95.0	1.0	1.0	1.0	1.0	434.0	839.0	956.0	1092.0
26	종로	93.0	115.0	4.0	4.0	6.0	9.0	284.0	602.0	792.0	964.0
27	종암	44.0	47.0	0.0	1.0	4.0	4.0	347.0	549.0	630.0	758.0
28	중앙	153.0	183.0	6.0	5.0	9.0	10.0	889.0	1506.0	1724.0	2022.0
29	중남	76.0	113.0	4.0	3.0	1.0	1.0	495.0	943.0	830.0	997.0

```
!pip install googlemaps
```

```
Requirement already satisfied: googlemaps in /usr/local/lib/python3.7/dist-packages (4.6.0)
Requirement already satisfied: requests<3.0,>=2.20.0 in /usr/local/lib/python3.7/dist-packages (from googlemaps)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from googlemaps)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from googlemaps)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from googlemaps)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from googlemaps)
```

```
import googlemaps
```

```
gkeys = "*****"
```

```
gmaps = googlemaps.Client(key=gkeys)
```

```
# 서울중부경찰서 정보 불러오기
```

```
# 구글맵 활용
```

```
gmaps.geocode('서울중부경찰서', language='ko')
```

```
[{'address_components': [{'long_name': '27',
    'short_name': '27',
    'types': ['premise']},
    {'long_name': '수표로',
    'short_name': '수표로',
    'types': ['political', 'sublocality', 'sublocality_level_4']},
    {'long_name': '을지로동',
    'short_name': '을지로동',
```



```

    'types': ['political', 'sublocality', 'sublocality_level_2']},
    {'long_name': '중구',
     'short_name': '중구',
     'types': ['political', 'sublocality', 'sublocality_level_1']},
    {'long_name': '서울특별시',
     'short_name': '서울특별시',
     'types': ['administrative_area_level_1', 'political']},
    {'long_name': '대한민국',
     'short_name': 'KR',
     'types': ['country', 'political']},
    {'long_name': '100-032',
     'short_name': '100-032',
     'types': ['postal_code']},
    'formatted_address': '대한민국 서울특별시 중구 을지로동 수표로 27',
    'geometry': {'location': {'lat': 37.5636465, 'lng': 126.9895796},
                  'location_type': 'R00FTOP',
                  'viewport': {'northeast': {'lat': 37.56499548029149,
                                              'lng': 126.9909285802915},
                               'southwest': {'lat': 37.56229751970849, 'lng': 126.9882306197085}}},
    'place_id': 'ChIJc-9q5uSifDURLhQmr5wkXmc',
    'plus_code': {'compound_code': 'HX7Q+FR 대한민국 서울특별시',
                  'global_code': '8Q98HX7Q+FR'},
    'types': ['establishment', 'point_of_interest', 'police']}

```

서울 경찰서 불러오기

```
station_name = []
```

```

for name in df_r['구분']:
    station_name.append('서울' + str(name) + '경찰서')

```

```
station_name
```

```

['서울강남경찰서',
 '서울강동경찰서',
 '서울강북경찰서',
 '서울강서경찰서',
 '서울관악경찰서',
 '서울광진경찰서',
 '서울구로경찰서',
 '서울금천경찰서',
 '서울남대문경찰서',
 '서울노원경찰서',
 '서울도봉경찰서',
 '서울동대문경찰서',
 '서울 동작경찰서',
 '서울마포경찰서',
 '서울방배경찰서',
 '서울서대문경찰서',
 '서울서부경찰서',
 '서울서초경찰서',
 '서울성동경찰서',
 '서울성북경찰서',
 '서울송파경찰서',
 '서울수서경찰서',
 '서울양천경찰서',
 '서울영등포경찰서',
 '서울용산경찰서',
 '서울은평경찰서',

```

```
'서울종로경찰서',
'서울중암경찰서',
'서울중랑경찰서',
'서울중부경찰서',
'서울혜화경찰서']
```

```
station_address = []
station_lat = []
station_lng = []
# 경찰서의 이름으로 주소 가져오기
for name in station_name:
    tmp = gmaps.geocode(name, language='ko')
    station_address.append(tmp[0].get("formatted_address"))

    tmp_loc = tmp[0].get("geometry")

    station_lat.append(tmp_loc['location']['lat'])
    station_lng.append(tmp_loc['location']['lng'])

print(name + '-->' + tmp[0].get("formatted_address"))

서울강남경찰서-->대한민국 서울특별시 강남구 테헤란로114길 11
서울강동경찰서-->대한민국 서울특별시 강동구 성내로 57
서울강북경찰서-->대한민국 서울특별시 강북구 오패산로 406
서울강서경찰서-->대한민국 서울특별시 양천구 신월동 화곡로 73
서울관악경찰서-->대한민국 서울특별시 관악구 관악로5길 33
서울광진경찰서-->대한민국 서울특별시 광진구 구의동 자양로 167
서울구로경찰서-->대한민국 서울특별시 구로구 가마산로 235
서울금천경찰서-->대한민국 서울특별시 금천구 시흥대로73길 50
서울남대문경찰서-->대한민국 서울특별시 중구 한강대로 410
서울노원경찰서-->대한민국 서울특별시 노원구 하계동 노원로 283
서울도봉경찰서-->대한민국 서울특별시 도봉구 노해로 403
서울동대문경찰서-->대한민국 서울특별시 동대문구 약령시로21길 29
서울 동작경찰서-->대한민국 서울특별시 동작구 노량진로 148
서울마포경찰서-->대한민국 서울특별시 마포구 마포대로 183
서울방배경찰서-->대한민국 서울특별시 서초구 동작대로 204
서울서대문경찰서-->대한민국 서울특별시 서대문구 통일로 113
서울서부경찰서-->대한민국 서울특별시 은평구 진흥로 58
서울서초경찰서-->대한민국 서울특별시 서초구 서초3동 반포대로 179
서울성동경찰서-->대한민국 서울특별시 성동구 행당동 왕십리광장로 9
서울성북경찰서-->대한민국 서울특별시 성북구 삼선동 보문로 170
서울송파경찰서-->대한민국 서울특별시 송파구 가락동 중대로 221
서울수서경찰서-->대한민국 서울특별시 강남구 개포로 617
서울양천경찰서-->대한민국 서울특별시 양천구 목동동로 99
서울영등포경찰서-->대한민국 서울특별시 영등포구 영등포동1가 618-7
서울용산경찰서-->대한민국 서울특별시 용산구 백범로 329
서울은평경찰서-->대한민국 서울특별시 은평구 불광동 연서로 365
서울종로경찰서-->대한민국 서울특별시 종로구 종로1.2.3.4가동 율곡로 46
서울중암경찰서-->대한민국 서울특별시 성북구 중암로 135
서울중랑경찰서-->대한민국 서울특별시 중랑구 신내역로3길 40-10
서울중부경찰서-->대한민국 서울특별시 중구 을지로동 수표로 27
서울혜화경찰서-->대한민국 서울특별시 종로구 인의동 창경궁로 112-16
```

```
# 주소 목록 확인
station_address
```

```
['대한민국 서울특별시 강남구 테헤란로114길 11',
```

'대한민국 서울특별시 강동구 성내로 57',
 '대한민국 서울특별시 강북구 오패산로 406',
 '대한민국 서울특별시 양천구 신월동 화곡로 73',
 '대한민국 서울특별시 관악구 관악로5길 33',
 '대한민국 서울특별시 광진구 구의동 자양로 167',
 '대한민국 서울특별시 구로구 가마산로 235',
 '대한민국 서울특별시 금천구 시흥대로73길 50',
 '대한민국 서울특별시 중구 한강대로 410',
 '대한민국 서울특별시 노원구 하계동 노원로 283',
 '대한민국 서울특별시 도봉구 노해로 403',
 '대한민국 서울특별시 동대문구 약령시로21길 29',
 '대한민국 서울특별시 동작구 노량진로 148',
 '대한민국 서울특별시 마포구 마포대로 183',
 '대한민국 서울특별시 서초구 동작대로 204',
 '대한민국 서울특별시 서대문구 통일로 113',
 '대한민국 서울특별시 은평구 진흥로 58',
 '대한민국 서울특별시 서초구 서초3동 반포대로 179',
 '대한민국 서울특별시 성동구 행당동 왕십리광장로 9',
 '대한민국 서울특별시 성북구 삼선동 보문로 170',
 '대한민국 서울특별시 송파구 가락동 중대로 221',
 '대한민국 서울특별시 강남구 개포로 617',
 '대한민국 서울특별시 양천구 목동동로 99',
 '대한민국 서울특별시 영등포구 영등포동1가 618-7',
 '대한민국 서울특별시 용산구 백범로 329',
 '대한민국 서울특별시 은평구 불광동 연서로 365',
 '대한민국 서울특별시 종로구 종로1.2.3.4가동 율곡로 46',
 '대한민국 서울특별시 성북구 종암로 135',
 '대한민국 서울특별시 중랑구 신내역로3길 40-10',
 '대한민국 서울특별시 중구 을지로동 수표로 27',
 '대한민국 서울특별시 종로구 인의동 창경궁로 112-16']

구별 열 새로 생성하기

gu_name = []

for name in station_addreess:

tmp = name.split()

tmp_gu = [gu for gu in tmp if gu[-1] == '구'][0]

gu_name.append(tmp_gu)

df_r['구별'] = gu_name

df_r

col	구 분	강간, 추행 검거	강간, 추행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생	구 별
0	강남	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0	강남구
1	강동	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0	강동구
2	강북	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0	강북구
3	강서	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0	양천구
4	관악	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0	관악구
5	광진	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0	광진구
6	구로	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0	구로구
7	금천	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0	금천구
8	남대문	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0	종구
9	노원	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0	노원구
10	도봉	87.0	82.0	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0	도봉구
11	동대문	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0	동대문구
12	동작	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0	동작구
13	마포	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0	마포구
14	방	50.0	63.0	1.0	1.0	1.0	1.0	150.0	340.0	365.0	422.0	서초구

14	배	50.0	63.0	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0	소구
15	서대문	149.0	159.0	1.0	2.0	3.0	3.0	524.0	1077.0	1029.0	1292.0	서대문구
16	서부	71.0	76.0	2.0	3.0	3.0	3.0	322.0	532.0	667.0	748.0	은평구
17	서초	253.0	333.0	3.0	5.0	4.0	4.0	801.0	1673.0	1415.0	1750.0	서초구
18	성동	98.0	115.0	3.0	3.0	5.0	5.0	565.0	1016.0	1103.0	1223.0	성동구
19	성북	77.0	95.0	2.0	2.0	0.0	0.0	233.0	439.0	610.0	672.0	성북구
20	송파	235.0	287.0	17.0	13.0	6.0	6.0	1006.0	2429.0	2280.0	2675.0	송파구
21	수서	194.0	195.0	1.0	2.0	2.0	1.0	475.0	1010.0	1200.0	1394.0	강남구
22	양천	90.0	109.0	3.0	3.0	5.0	5.0	718.0	1517.0	1270.0	1582.0	양천구
23	영등포	252.0	315.0	5.0	6.0	4.0	7.0	953.0	2188.0	2157.0	2701.0	영등포구
24	용산	186.0	230.0	6.0	6.0	3.0	5.0	534.0	1111.0	1327.0	1617.0	용산구
25	은평	74.0	95.0	1.0	1.0	1.0	1.0	434.0	839.0	956.0	1092.0	은평구
26	종로	93.0	115.0	4.0	4.0	6.0	9.0	284.0	602.0	792.0	964.0	종로구
27	종암	44.0	47.0	0.0	1.0	4.0	4.0	347.0	549.0	630.0	758.0	성북구

금천 예외 처리

df_r.loc[df_r["구분"] == "금천",["구별"]] = "관악구"

df_r

col	구 분	강간, 추행 검거	강간, 추행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생	구 별
0	강남	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0	강남구
1	강동	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0	강동구
2	강북	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0	강북구
3	강서	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0	양천구
4	관악	351.0	403.0	3.0	3.0	12.0	12.0	1137.0	2229.0	2355.0	2614.0	관악구
5	광진	224.0	255.0	4.0	4.0	2.0	2.0	874.0	1664.0	1407.0	1676.0	광진구
6	구로	142.0	197.0	5.0	5.0	6.0	6.0	1008.0	1763.0	1823.0	2204.0	구로구
7	금천	93.0	117.0	6.0	7.0	4.0	4.0	553.0	999.0	1310.0	1471.0	관악구
8	남대문	52.0	73.0	2.0	2.0	1.0	1.0	368.0	699.0	494.0	579.0	종구
9	노원	148.0	169.0	3.0	3.0	3.0	4.0	646.0	1404.0	1781.0	2163.0	노원구
10	도봉	87.0	82.0	2.0	2.0	0.0	0.0	379.0	861.0	1022.0	1234.0	도봉구
11	동대문	142.0	161.0	5.0	6.0	8.0	10.0	713.0	1440.0	1597.0	1784.0	동대문구
12	동작	121.0	226.0	2.0	1.0	1.0	1.0	592.0	1341.0	1370.0	1631.0	동작구
13	마포	287.0	348.0	3.0	4.0	8.0	9.0	679.0	1483.0	1557.0	1844.0	마포구
14	방	50.0	63.0	1.0	1.0	1.0	1.0	150.0	340.0	365.0	422.0	서초

14	배	50.0	63.0	1.0	1.0	1.0	1.0	159.0	348.0	365.0	423.0	소구
15	서대문	149.0	159.0	1.0	2.0	3.0	3.0	524.0	1077.0	1029.0	1292.0	서대문구
16	서부	71.0	76.0	2.0	3.0	3.0	3.0	322.0	532.0	667.0	748.0	은평구
17	서초	253.0	333.0	3.0	5.0	4.0	4.0	801.0	1673.0	1415.0	1750.0	서초구
18	성동	98.0	115.0	3.0	3.0	5.0	5.0	565.0	1016.0	1103.0	1223.0	성동구
19	성북	77.0	95.0	2.0	2.0	0.0	0.0	233.0	439.0	610.0	672.0	성북구
20	송파	235.0	287.0	17.0	13.0	6.0	6.0	1006.0	2429.0	2280.0	2675.0	송파구
21	수서	194.0	195.0	1.0	2.0	2.0	1.0	475.0	1010.0	1200.0	1394.0	강남구
22	양천	90.0	109.0	3.0	3.0	5.0	5.0	718.0	1517.0	1270.0	1582.0	양천구
23	영등포	252.0	315.0	5.0	6.0	4.0	7.0	953.0	2188.0	2157.0	2701.0	영등포구
24	용산	186.0	230.0	6.0	6.0	3.0	5.0	534.0	1111.0	1327.0	1617.0	용산구
25	은평	74.0	95.0	1.0	1.0	1.0	1.0	434.0	839.0	956.0	1092.0	은평구
26	종로	93.0	115.0	4.0	4.0	6.0	9.0	284.0	602.0	792.0	964.0	종로구
27	종암	44.0	47.0	0.0	1.0	4.0	4.0	347.0	549.0	630.0	758.0	성북구
28	중앙	153.0	183.0	6.0	5.0	9.0	10.0	889.0	1506.0	1724.0	2022.0	중앙구
29	중부	76.0	113.0	4.0	3.0	1.0	1.0	495.0	943.0	830.0	997.0	중구

```
# 전처리 데이터 저장
```

```
df_r.to_csv("/content/drive/MyDrive/스터디/데이터 주무르기/data/02_preprocess.csv", index = False, €
```

▼ 2.2 피벗 테이블 활용하기

```
import numpy as np
```

```
# 피벗 테이블로 구별로 인덱스
```

```
crime_anal = pd.pivot_table(df_r, index='구별', aggfunc=np.sum)
```

```
crime_anal.head()
```

col	강간, 추 행 검거	강간, 추 행 발생	강도 검거	강도 발생	살인 검거	살인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생
구 별										
강 남 구	627.0	708.0	11.0	14.0	5.0	6.0	1491.0	2951.0	3111.0	3677.0
강 동 구	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0
강 북 구	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0
관 악 구	444.0	520.0	9.0	10.0	16.0	16.0	1690.0	3228.0	3665.0	4085.0

```
# 검거율 열 새로 생성
```

```
crime_anal['강간,추행 검거율'] = crime_anal['강간,추행 검거']/crime_anal['강간,추행 발생']
```

```
crime_anal['강도검거율'] = crime_anal['강도 검거']/crime_anal['강도 발생']
```

```
crime_anal['살인검거율'] = crime_anal['살인 검거']/crime_anal['살인 발생']
```

```
crime_anal['절도검거율'] = crime_anal['절도 검거']/crime_anal['절도 발생']
```

```
crime_anal['폭력검거율'] = crime_anal['폭력 검거']/crime_anal['폭력 발생']
```

```
del crime_anal['강간,추행 검거']
```

```
del crime_anal['강도 검거']
```

```
del crime_anal['살인 검거']
```

```
del crime_anal['절도 검거']
```

```
del crime_anal['폭력 검거']
```

```
crime_anal
```


구	강간, 추행 발생	강도 발생	살인 발생	절도 발생	폭력 발생	강간, 추행 검거율	강도 검거율	살인 검거율	절도 검거율	폭력 검거율
강남구	708.0	14.0	6.0	2951.0	3677.0	0.885593	0.785714	0.833333	0.505252	0.846070
강동구	188.0	15.0	3.0	1640.0	1942.0	0.888298	0.933333	1.000000	0.476220	0.873841
강북구	159.0	5.0	9.0	780.0	1817.0	0.830189	0.800000	0.888889	0.628205	0.894331
관악구	520.0	10.0	16.0	3228.0	4085.0	0.853846	0.900000	1.000000	0.523544	0.897185
광진구	255.0	4.0	2.0	1664.0	1676.0	0.878431	1.000000	1.000000	0.525240	0.839499
구로구	197.0	5.0	6.0	1763.0	2204.0	0.720812	1.000000	1.000000	0.571753	0.827132
노원구	169.0	3.0	4.0	1404.0	2163.0	0.875740	1.000000	0.750000	0.460114	0.823393
도봉구	82.0	2.0	0.0	861.0	1234.0	1.060976	1.000000	NaN	0.440186	0.828201
동대문구	161.0	6.0	10.0	1440.0	1784.0	0.881988	0.833333	0.800000	0.495139	0.895179
동작구	226.0	1.0	1.0	1341.0	1631.0	0.535398	2.000000	1.000000	0.441462	0.839975
마포구	348.0	4.0	9.0	1483.0	1844.0	0.824713	0.750000	0.888889	0.457856	0.844360
서대문구	159.0	2.0	3.0	1077.0	1292.0	0.937107	0.500000	1.000000	0.486537	0.796440
서초구	396.0	6.0	5.0	2021.0	2173.0	0.765152	0.666667	1.000000	0.475012	0.819144
성북구	100.0	1.0	1.0	1000.0	1000.0	0.800000	0.800000	0.800000	0.400000	0.800000

강동구	115.0	3.0	5.0	1016.0	1223.0	0.852174	1.000000	1.000000	0.556102	0.901881
성북구	142.0	3.0	4.0	988.0	1430.0	0.852113	0.666667	1.000000	0.587045	0.867133
송파구	287.0	13.0	6.0	2429.0	2675.0	0.818815	1.307692	1.000000	0.414162	0.852336
양천구	374.0	9.0	15.0	3206.0	4027.0	0.834225	0.888889	0.933333	0.557704	0.833623
영등포구	315.0	6.0	7.0	2188.0	2701.0	0.800000	0.833333	0.571429	0.435558	0.798593
용산구	230.0	6.0	5.0	1111.0	1617.0	0.808696	1.000000	0.600000	0.480648	0.820656
은평구	171.0	4.0	4.0	1371.0	1840.0	0.847953	0.750000	1.000000	0.551422	0.882065
종로구	192.0	7.0	9.0	1183.0	1711.0	0.760417	1.000000	0.666667	0.502113	0.829924
중구	186.0	5.0	2.0	1642.0	1576.0	0.688172	1.200000	1.000000	0.525579	0.840102
중										

```
con_list = ['강간,추행 검거율', '강도검거율', '살인검거율', '절도검거율', '폭력검거율']
```

```
# 검거율이 10이상인 데이터 통일
```

```
for column in con_list:
    crime_anal.loc[crime_anal[column] > 1, column] = 1
```

```
crime_anal
```

구	강간, 추행 발생	강도 발생	살인 발생	절도 발생	폭력 발생	강간, 추행 검거율	강도 검거율	살인 검거율	절도 검거율	폭력 검거율
강남구	708.0	14.0	6.0	2951.0	3677.0	0.885593	0.785714	0.833333	0.505252	0.846070
강동구	188.0	15.0	3.0	1640.0	1942.0	0.888298	0.933333	1.000000	0.476220	0.873841
강북구	159.0	5.0	9.0	780.0	1817.0	0.830189	0.800000	0.888889	0.628205	0.894331
관악구	520.0	10.0	16.0	3228.0	4085.0	0.853846	0.900000	1.000000	0.523544	0.897185
광진구	255.0	4.0	2.0	1664.0	1676.0	0.878431	1.000000	1.000000	0.525240	0.839499
구로구	197.0	5.0	6.0	1763.0	2204.0	0.720812	1.000000	1.000000	0.571753	0.827132
노원구	169.0	3.0	4.0	1404.0	2163.0	0.875740	1.000000	0.750000	0.460114	0.823393
도봉구	82.0	2.0	0.0	861.0	1234.0	1.000000	1.000000	NaN	0.440186	0.828201
동대문구	161.0	6.0	10.0	1440.0	1784.0	0.881988	0.833333	0.800000	0.495139	0.895179
동작구	226.0	1.0	1.0	1341.0	1631.0	0.535398	1.000000	1.000000	0.441462	0.839975
마포구	348.0	4.0	9.0	1483.0	1844.0	0.824713	0.750000	0.888889	0.457856	0.844360
서대문구	159.0	2.0	3.0	1077.0	1292.0	0.937107	0.500000	1.000000	0.486537	0.796440
서초구	396.0	6.0	5.0	2021.0	2173.0	0.765152	0.666667	1.000000	0.475012	0.819144
성북구	100.0	1.0	1.0	700.0	800.0	0.800000	1.000000	1.000000	0.400000	0.800000

강동구	115.0	3.0	5.0	1016.0	1223.0	0.852174	1.000000	1.000000	0.556102	0.901881
성북구	142.0	3.0	4.0	988.0	1430.0	0.852113	0.666667	1.000000	0.587045	0.867133
송파구	287.0	13.0	6.0	2429.0	2675.0	0.818815	1.000000	1.000000	0.414162	0.852336
양천구	374.0	9.0	15.0	3206.0	4027.0	0.834225	0.888889	0.933333	0.557704	0.833623
영등포구	315.0	6.0	7.0	2188.0	2701.0	0.800000	0.833333	0.571429	0.435558	0.798593
용산구	230.0	6.0	5.0	1111.0	1617.0	0.808696	1.000000	0.600000	0.480648	0.820656

열 이름 변경

```
crime_anal.rename(columns = {'강간,추행 발생':'강간,추행',
                              '강도 발생':'강도',
                              '살인 발생':'살인',
                              '절도 발생':'절도',
                              '폭력 발생':'폭력'}, inplace=True)
```

crime_anal

coi	강 간, 추 행	강 도	살 인	절 도	폭 력	강 간, 추 행 거 울	강 도 검 거 울	살 인 검 거 울	절 도 검 거 울	폭 력 검 거 울
구 별										
강 남 구	708.0	14.0	6.0	2951.0	3677.0	0.885593	0.785714	0.833333	0.505252	0.846070
강 동 구	188.0	15.0	3.0	1640.0	1942.0	0.888298	0.933333	1.000000	0.476220	0.873841
강 북 구	159.0	5.0	9.0	780.0	1817.0	0.830189	0.800000	0.888889	0.628205	0.894331
관 악 구	520.0	10.0	16.0	3228.0	4085.0	0.853846	0.900000	1.000000	0.523544	0.897185
광 진 구	255.0	4.0	2.0	1664.0	1676.0	0.878431	1.000000	1.000000	0.525240	0.839499
구 로 구	197.0	5.0	6.0	1763.0	2204.0	0.720812	1.000000	1.000000	0.571753	0.827132
노 원 구	169.0	3.0	4.0	1404.0	2163.0	0.875740	1.000000	0.750000	0.460114	0.823393
도 봉 구	82.0	2.0	0.0	861.0	1234.0	1.000000	1.000000	NaN	0.440186	0.828201
동 대 문 구	161.0	6.0	10.0	1440.0	1784.0	0.881988	0.833333	0.800000	0.495139	0.895179
동 작 구	226.0	1.0	1.0	1341.0	1631.0	0.535398	1.000000	1.000000	0.441462	0.839975
마 포 구	348.0	4.0	9.0	1483.0	1844.0	0.824713	0.750000	0.888889	0.457856	0.844360
서 대 문 구	159.0	2.0	3.0	1077.0	1292.0	0.937107	0.500000	1.000000	0.486537	0.796440
서 초 구	396.0	6.0	5.0	2021.0	2173.0	0.765152	0.666667	1.000000	0.475012	0.819144
성 동	115.0	3.0	5.0	1016.0	1222.0	0.852174	1.000000	1.000000	0.556102	0.801881

형구	115.0	3.0	5.0	1016.0	1223.0	0.852174	1.000000	1.000000	0.556102	0.901881
성북구	142.0	3.0	4.0	988.0	1430.0	0.852113	0.666667	1.000000	0.587045	0.867133
송파구	287.0	13.0	6.0	2429.0	2675.0	0.818815	1.000000	1.000000	0.414162	0.852336
..										

▼ 2.3 데이터 표현을 위해 다듬기

예

```
from sklearn import preprocessing

col = ['강간,추행', '강도', '살인', '절도', '폭력']

x = crime_anal[col].values
min_max_scaler = preprocessing.MinMaxScaler()

x_scaled = min_max_scaler.fit_transform(x.astype(float))
crime_anal_norm = pd.DataFrame(x_scaled, columns = col, index = crime_anal.index)

col2 = ['강간,추행 검거율', '강도검거율', '살인검거율', '절도검거율', '폭력검거율']
crime_anal_norm[col2] = crime_anal[col2]
crime_anal_norm
```

구 별	강간, 추 행	강도	살인	절도	폭력	강간, 추 행 검거율	강도검 거율	살인검 거율	절도검 거율
강남구	1.000000	0.928571	0.3750	0.886846	0.857442	0.885593	0.785714	0.833333	0.50525
강동구	0.169329	1.000000	0.1875	0.351307	0.251223	0.888298	0.933333	1.000000	0.47622
강북구	0.123003	0.285714	0.5625	0.000000	0.207547	0.830189	0.800000	0.888889	0.62820
관악구	0.699681	0.642857	1.0000	1.000000	1.000000	0.853846	0.900000	1.000000	0.52354
광진구	0.276358	0.214286	0.1250	0.361111	0.158281	0.878431	1.000000	1.000000	0.52524
구로구	0.183706	0.285714	0.3750	0.401552	0.342767	0.720812	1.000000	1.000000	0.57175
노원구	0.138978	0.142857	0.2500	0.254902	0.328442	0.875740	1.000000	0.750000	0.46011
도봉구	0.000000	0.071429	0.0000	0.033088	0.003843	1.000000	1.000000	NaN	0.44018
동대문구	0.126198	0.357143	0.6250	0.269608	0.196017	0.881988	0.833333	0.800000	0.49513
동작구	0.230032	0.000000	0.0625	0.229167	0.142558	0.535398	1.000000	1.000000	0.44146
마포구	0.424920	0.214286	0.5625	0.287173	0.216981	0.824713	0.750000	0.888889	0.45785
서대문구	0.123003	0.071429	0.1875	0.121324	0.024109	0.937107	0.500000	1.000000	0.48653
서초구	0.501597	0.357143	0.3125	0.506944	0.331936	0.765152	0.666667	1.000000	0.47501
성동구	0.052716	0.142857	0.2125	0.006405	0.000000	0.852174	1.000000	1.000000	0.55610

강구	0.052716	0.142857	0.3125	0.096405	0.000000	0.852174	1.000000	1.000000	0.55610
성북구	0.095847	0.142857	0.2500	0.084967	0.072327	0.852113	0.666667	1.000000	0.58704
송파	0.327476	0.857143	0.3750	0.673611	0.507338	0.818815	1.000000	1.000000	0.41416

CCTV 데이터와 결합

```
result_CCTV = pd.read_csv('/content/drive/MyDrive/스터디/데이터 주무르기/data/01_preprocessi
index_col='구별')
```

```
crime_anal_norm[['인구수', 'CCTV']] = result_CCTV[['인구수', '총계']]
```

```
crime_anal_norm.head()
```

	강간, 추 행	강도	살인	절도	폭력	강간, 추 행 검거 율	강도검 거율	살인검 거율	절도검 거율
구 별									
강 남 구	1.000000	0.928571	0.3750	0.886846	0.857442	0.885593	0.785714	0.833333	0.50525
강 동 구	0.169329	1.000000	0.1875	0.351307	0.251223	0.888298	0.933333	1.000000	0.47622
강 북 구	0.123003	0.285714	0.5625	0.000000	0.207547	0.830189	0.800000	0.888889	0.62820
관 악 구	0.699681	0.642857	1.0000	1.000000	1.000000	0.853846	0.900000	1.000000	0.52354
광 진 구	0.276358	0.214286	0.1250	0.361111	0.158281	0.878431	1.000000	1.000000	0.52524

범죄 열을 생성 np.sum 활용!!

```
col = ['강간, 추행', '강도', '살인', '절도', '폭력']
```

```
crime_anal_norm['범죄'] = np.sum(crime_anal_norm[col], axis=1)
```

```
crime_anal_norm.head()
```


구 별	강간, 추 행	강도	살인	절도	폭력	강간, 추 행 검거율	강도검 거율	살인검 거율	절도검 거율
강 남 구	1.000000	0.928571	0.3750	0.886846	0.857442	0.885593	0.785714	0.833333	0.50525
강 동 구	0.169329	1.000000	0.1875	0.351307	0.251223	0.888298	0.933333	1.000000	0.47622
강 북 구	0.123003	0.285714	0.5625	0.000000	0.207547	0.830189	0.800000	0.888889	0.62820

검거 열을 새로이 생성

```
col = ['강간,추행 검거율', '강도검거율', '살인검거율', '절도검거율', '폭력검거율']
```

```
crime_anal_norm['검거'] = np.sum(crime_anal_norm[col], axis=1)
```

```
crime_anal_norm.head()
```

구 별	강간, 추 행	강도	살인	절도	폭력	강간, 추 행 검거율	강도검 거율	살인검 거율	절도검 거율
강 남 구	1.000000	0.928571	0.3750	0.886846	0.857442	0.885593	0.785714	0.833333	0.50525
강 동 구	0.169329	1.000000	0.1875	0.351307	0.251223	0.888298	0.933333	1.000000	0.47622
강 북 구	0.123003	0.285714	0.5625	0.000000	0.207547	0.830189	0.800000	0.888889	0.62820
관 악 구	0.699681	0.642857	1.0000	1.000000	1.000000	0.853846	0.900000	1.000000	0.52354
광 진 구	0.276358	0.214286	0.1250	0.361111	0.158281	0.878431	1.000000	1.000000	0.52524

▼ 2.4 seaborn

```
# # 나눔 글꼴 설치
```

```
# !sudo apt-get install -y fonts-nanum
```

```
# !sudo fc-cache -fv
# !rm ~/.cache/matplotlib -rf

import matplotlib.pyplot as plt

import seaborn as sns
#seaborn 스타일 설정
sns.set_style("whitegrid")

# 플롯 스타일 설정
plt.style.use('fivethirtyeight')

# matplotlib을 사용하는 모든 plot에 나눔 글꼴 적용
plt.rc('font', family='NanumBarunGothic')
%matplotlib inline

sns.pairplot(crime_anal_norm, vars=["강도", "살인", "폭력"], kind='reg', size=3)
plt.show()
```

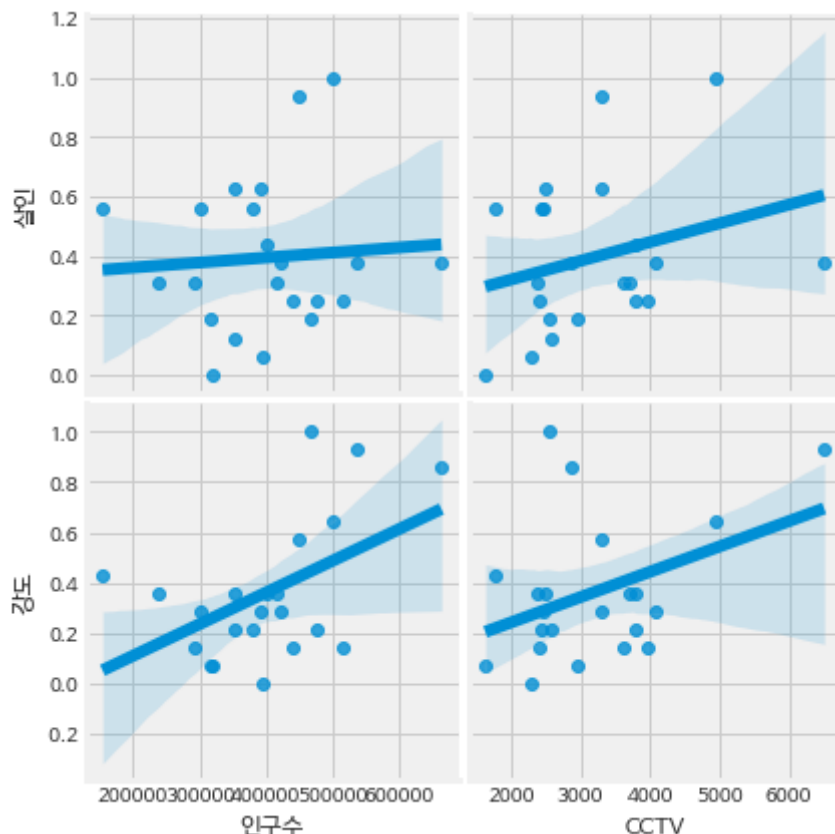
/usr/local/lib/python3.7/dist-packages/seaborn/axisgrid.py:2076: UserWarning: The `size` par

강도와 폭력, 살인과 폭력, 강도와 살인 모두 양의 상관관계가 있다.



```
sns.pairplot(crime_anal_norm, x_vars=["인구수", "CCTV"],
             y_vars=["살인", "강도"], kind='reg', size=3)
plt.show()
```

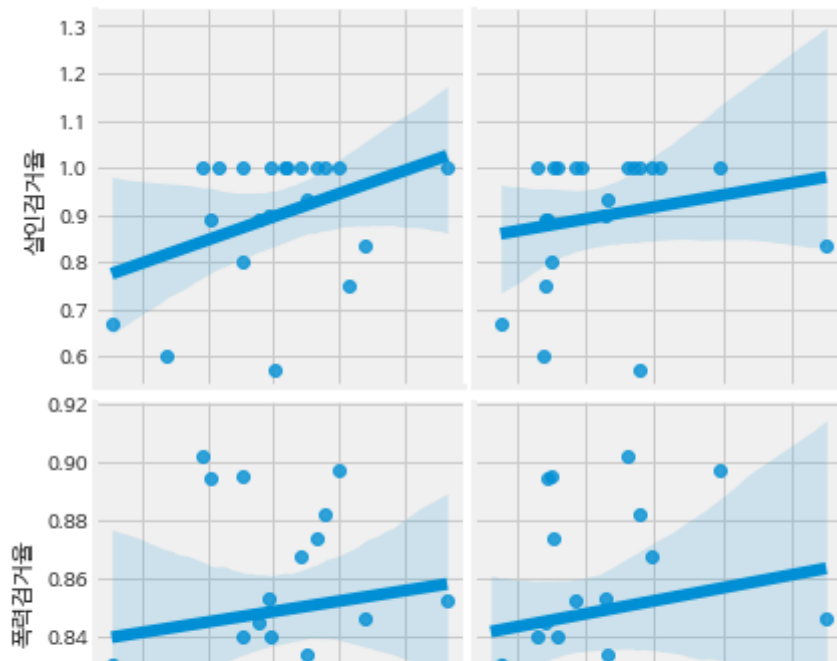
/usr/local/lib/python3.7/dist-packages/seaborn/axisgrid.py:2076: UserWarning: The `size` par
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/matplotlib/backends/backend_agg.py:214: RuntimeWarning:
font.set_text(s, 0.0, flags=flags)
/usr/local/lib/python3.7/dist-packages/matplotlib/backends/backend_agg.py:183: RuntimeWarning:
font.set_text(s, 0, flags=flags)



눈에 띄는 인사이트를 발견하지 못하였다. 굳이 찾아보자면 인구수가 증가할 수록 강도가 많이 일어나고 그런 지역일수록 CCTV가 많이 위치해 있는것으로 보인다.

```
sns.pairplot(crime_anal_norm, x_vars=["인구수", "CCTV"],
             y_vars=["살인검거율", "폭력검거율"], kind='reg', size=3)
plt.show()
```

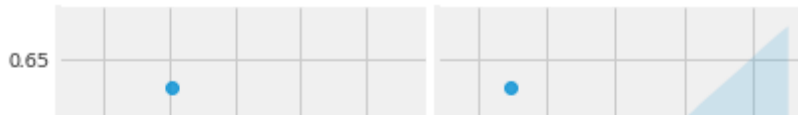
/usr/local/lib/python3.7/dist-packages/seaborn/axisgrid.py:2076: UserWarning: The `size` parameter is deprecated. Use `figure` or `facet_kwargs` instead.
warnings.warn(msg, UserWarning)



- 전반적으로 CCTV가 많은 지역에서 살인, 폭력 검거율이 좀 더 높은 모습을 보인다.
- 인구수가 많은 지역일수록도 검거율이 높아진다고 할 수 있다.
- 하지만 유의미한 양의 상관성은 띄지 않는것으로 판단한다.
- 살인 검거율은 1에 위치해 있는것이 많은것으로 보아 심각한 범죄인 살인범죄의 검거율은 매우 높다고 보여진다.

```
sns.pairplot(crime_anal_norm, x_vars=["인구수", "CCTV"],  
             y_vars=["절도검거율", "강도검거율"], kind='reg', size=3)  
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/axisgrid.py:2076: UserWarning: The `size` parameter
warnings.warn(msg, UserWarning)
```



- 강도 검거율은 오히려 CCTV와는 음의 상관관계를 보인다. 절도검거율은 CCTV와는 양의 상관관계를 보인다.
- 절도, 강도 검거율은 인구수와는 음의 상관관계를 보인다.



검거열 스케일링

```
tmp_max = crime_anal_norm['검거'].max()
crime_anal_norm['검거'] = crime_anal_norm['검거'] / tmp_max
crime_anal_norm_sort = crime_anal_norm.sort_values(by='검거', ascending=False)
crime_anal_norm_sort.head()
```

	강간, 추 행	강도	살인	절도	폭력	강간, 추 행 검거 율	강도검 거율	살 인 검거 율	절도검 거율	...
구 별										
성 동 구	0.052716	0.142857	0.3125	0.096405	0.000000	0.852174	1.000000	1.0	0.556102	0.9
광 진 구	0.276358	0.214286	0.1250	0.361111	0.158281	0.878431	1.000000	1.0	0.525240	0.8
중 랑 구	0.161342	0.285714	0.6250	0.296569	0.279175	0.836066	1.000000	0.9	0.590305	0.8
관 악 구	0.699681	0.642857	1.0000	1.000000	1.000000	0.853846	0.900000	1.0	0.523544	0.8
강 동 구	0.169329	1.000000	0.1875	0.351307	0.251223	0.888298	0.933333	1.0	0.476220	0.8

히트맵 출력

```
target_col = ['강간,추행 검거율', '강도검거율', '살인검거율', '절도검거율', '폭력검거율']
```

```
crime_anal_norm_sort = crime_anal_norm.sort_values(by='검거', ascending=False)
```

```
plt.figure(figsize = (10,10))
```

```
sns.heatmap(crime_anal_norm_sort[target_col], annot=True, fmt='f',
            linewidths=.5, cmap='RdPu')
```

```
plt.title('범죄 검거 비율 (정규화된 검거의 합으로 정렬)')
plt.show()
```

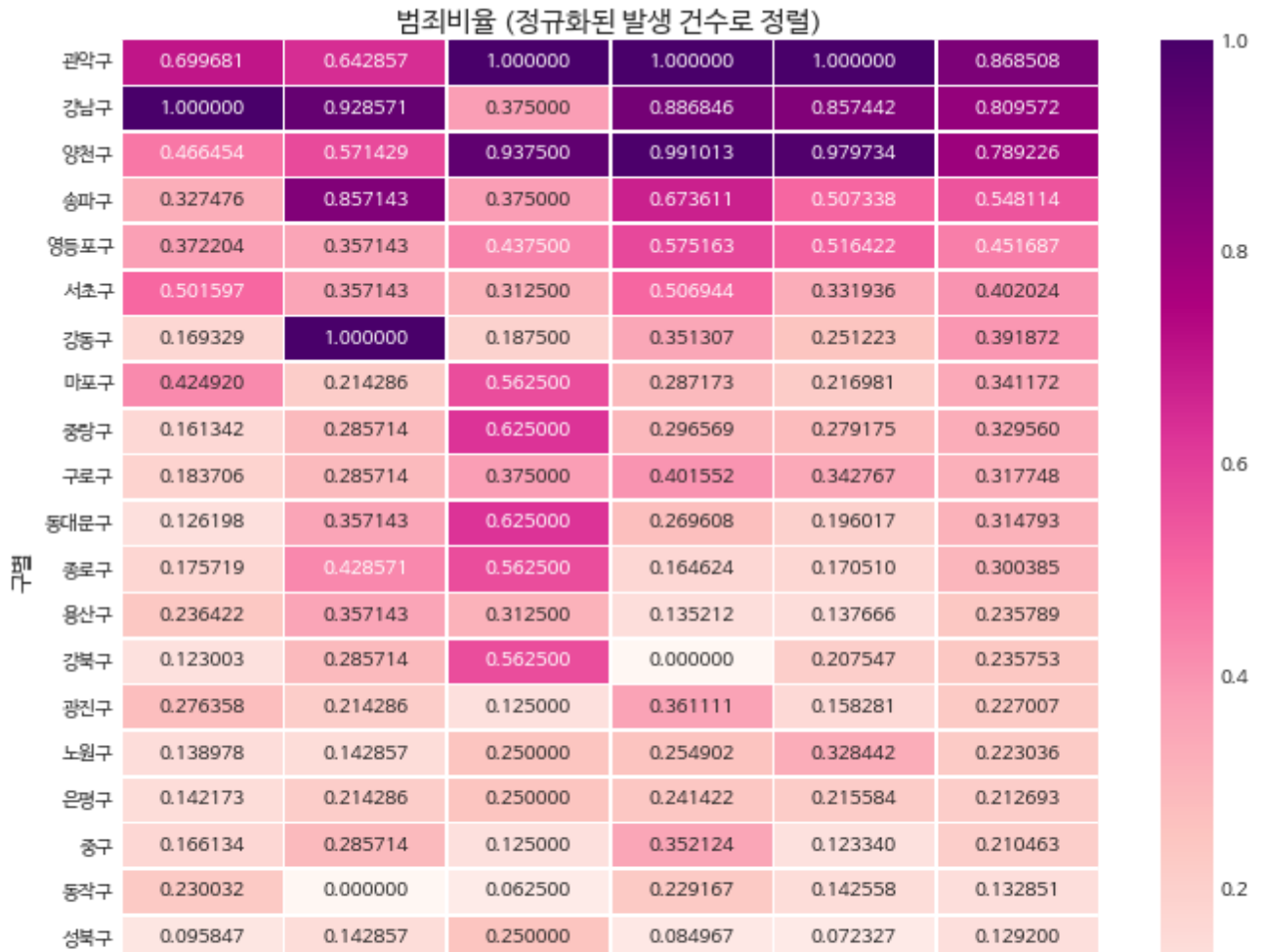


- 절도의 검거율이 다른 범죄의 검거율에 비해 낮은 것을 확인

```
# 히트맵 출력
target_col = ['강간,추행', '강도', '살인', '절도', '폭력', '범죄']

crime_anal_norm['범죄'] = crime_anal_norm['범죄'] / 5
crime_anal_norm_sort = crime_anal_norm.sort_values(by='범죄', ascending=False)

plt.figure(figsize = (10,10))
sns.heatmap(crime_anal_norm_sort[target_col], annot=True, fmt='f', linewidths=.5,
            cmap='RdPu')
plt.title('범죄비율 (정규화된 발생 건수로 정렬)')
plt.show()
```



- 범죄가 가장 많이 일어나는 지역은 관악구, 강남구, 양천구가 높게 관찰된다. 뒤로는 송파, 영등포, 서초가 뒤따르고 있다.

2.5 지도 시각화 도구 - Folium

```
#!pip install folium
```

```
import folium
```

```
# json으로 되어있는 서울시 구별 데이터 정보 로딩
```

```
import json
```

```
geo_path = '/content/drive/MyDrive/스터디/데이터 주무르기/data/02. skorea_municipalities_geo_simple'
```

```
geo_str = json.load(open(geo_path, encoding='utf-8'))
```

```
#서울시 중심의 위도와 경도 정보를 먼저 입력하고 경계선 설정
```

```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11,
                  tiles='Stamen Toner')
```

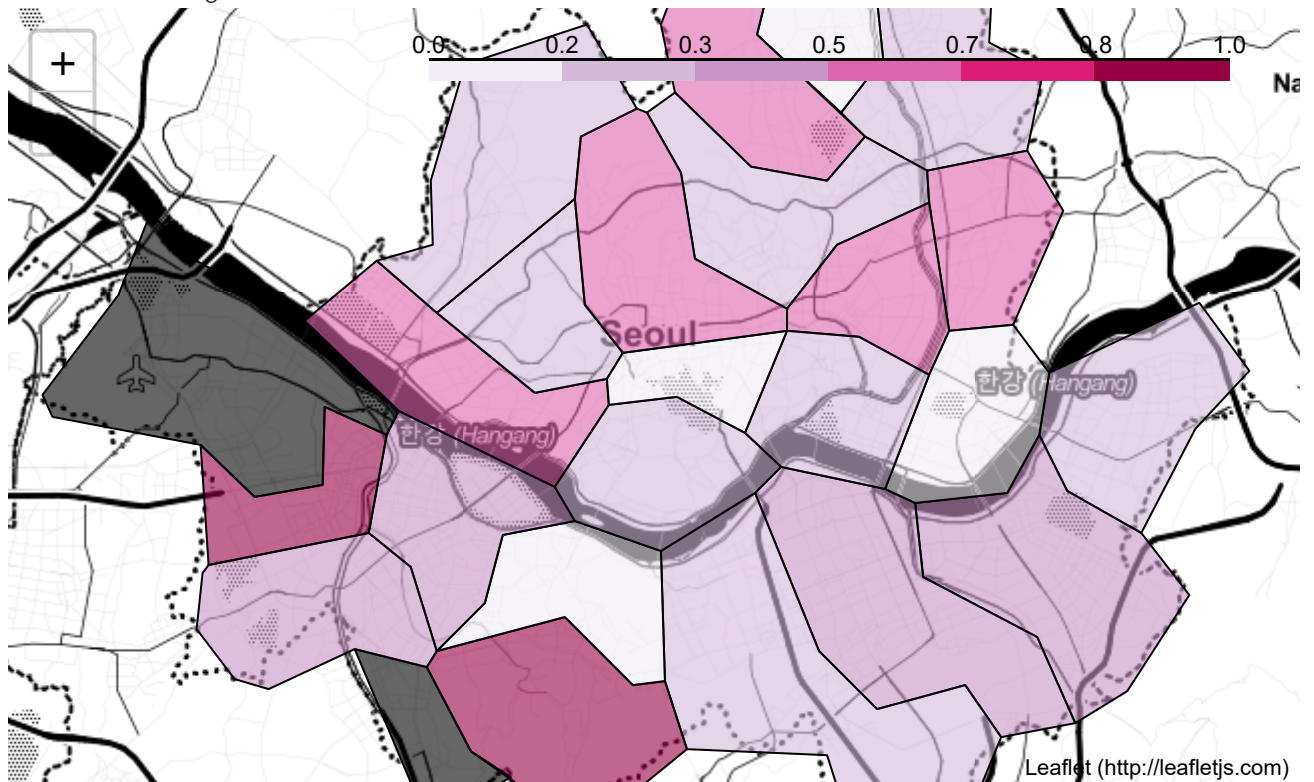
```
# 살인 발생 빈도 추출
```

```
map.choropleth(geo_data = geo_str,
               data = crime_anal_norm['살인'],
               columns = [crime_anal_norm.index, crime_anal_norm['살인']],
               fill_color = 'PuRd', #PuRd, YlGnBu)
```

```
key_on = 'feature.id')
```

map

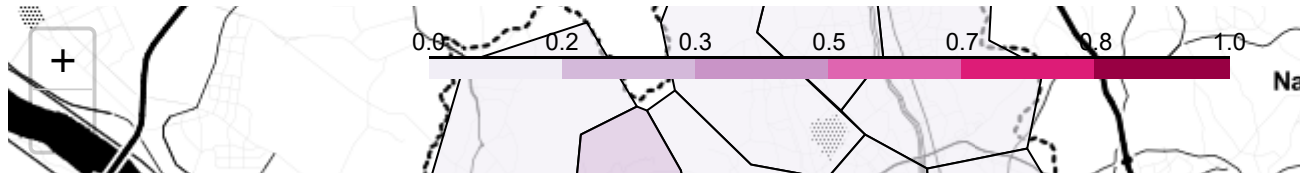
```
/usr/local/lib/python3.7/dist-packages/folium/folium.py:426: FutureWarning: The choropleth  
FutureWarning
```



```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11,  
                  tiles='Stamen Toner')  
# 강간, 추행 발생 빈도 추출  
map.choropleth(geo_data = geo_str,  
                data = crime_anal_norm['강간, 추행'],  
                columns = [crime_anal_norm.index, crime_anal_norm['강간, 추행']],  
                fill_color = 'PuRd', #PuRd, YlGnBu  
                key_on = 'feature.id')
```

map

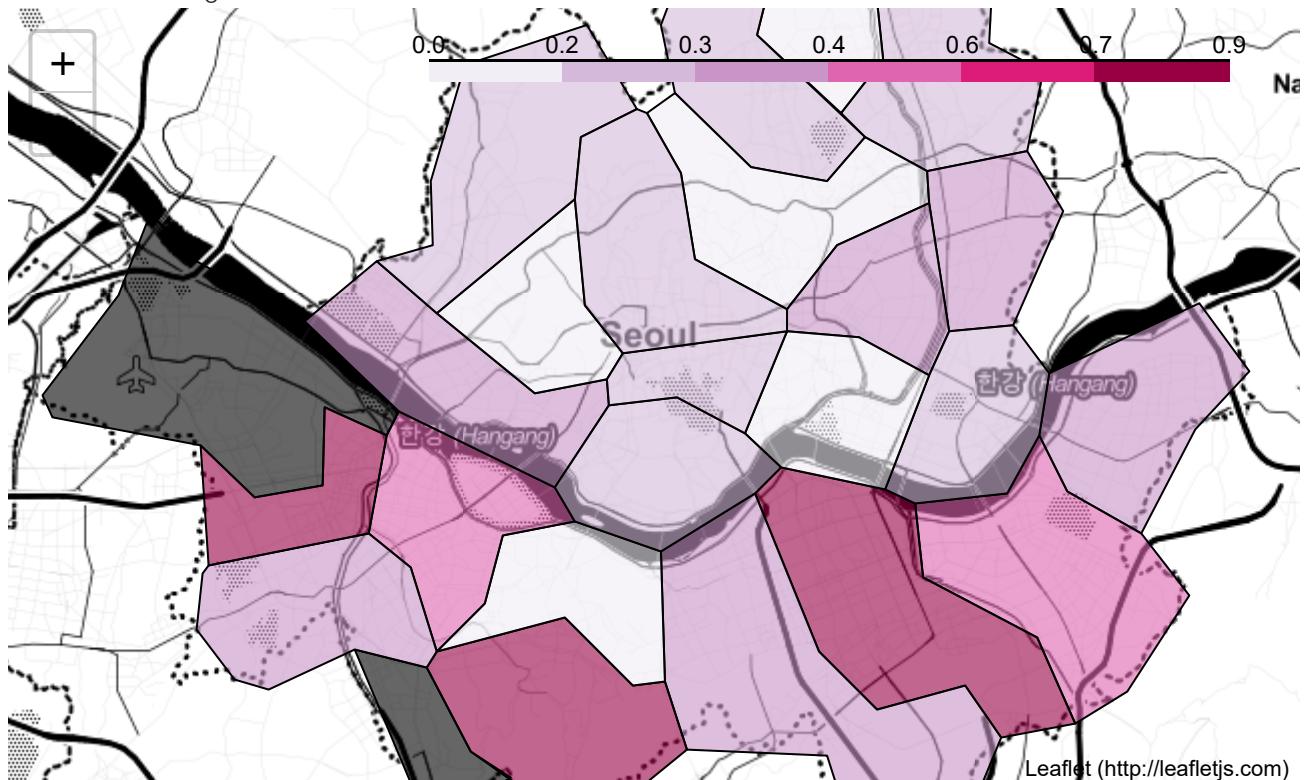

```
/usr/local/lib/python3.7/dist-packages/folium/folium.py:426: FutureWarning: The choropleth
FutureWarning
```



```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11,
                  tiles='Stamen Toner')
# 범죄 발생 빈도 추출
map.choropleth(geo_data = geo_str,
               data = crime_anal_norm['범죄'],
               columns = [crime_anal_norm.index, crime_anal_norm['범죄']],
               fill_color = 'PuRd', #PuRd, YlGnBu
               key_on = 'feature.id')
```

map

```
/usr/local/lib/python3.7/dist-packages/folium/folium.py:426: FutureWarning: The choropleth
FutureWarning
```



- 강남 지역
- 강서 지역의 범죄 빈도가 높다

하지만 이는 인구수에 비례해야한다.

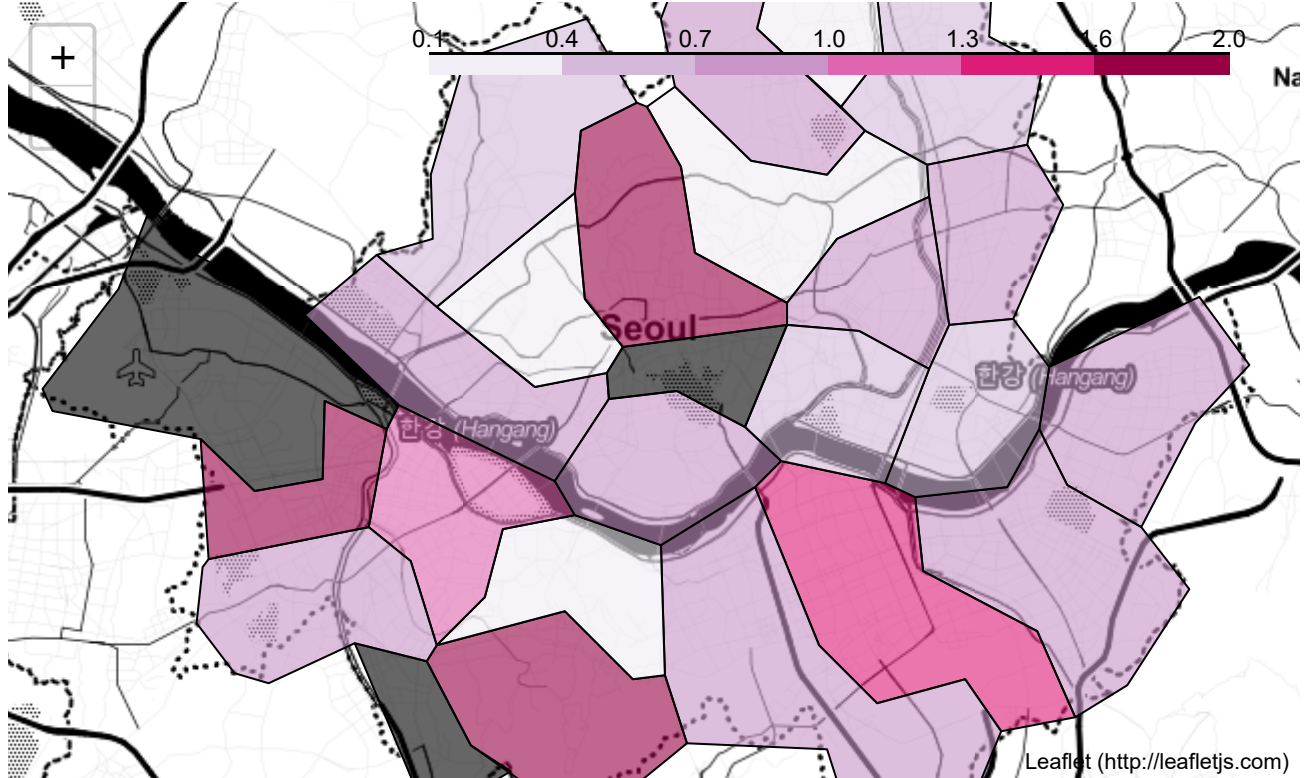
```
tmp_criminal = crime_anal_norm['범죄'] / crime_anal_norm['인구수'] * 1000000
```

```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11,
                  tiles='Stamen Toner')
```

```
map.choropleth(geo_data = geo_str,
               data = tmp_criminal,
               columns = [crime_anal.index, tmp_criminal],
               fill_color = 'PuRd', #PuRd, YlGnBu
               key_on = 'feature.id')
```

map

/usr/local/lib/python3.7/dist-packages/folium/folium.py:426: FutureWarning: The choropleth
FutureWarning



- 여전히 강서 지역의 범죄율이 높다
- 중구 및 종로구의 범죄율이 높아진게 보인다.

▼ 2.6 경찰서별 검거현황과 구별 범죄발생 현황을 표현하기

```
df_r['lat'] = station_lat
df_r['lng'] = station_lng
```

```
col = ['살인 검거', '강도 검거', '강간,추행 검거', '절도 검거', '폭력 검거']
tmp = df_r[col] / df_r[col].max()
```

```
df_r['검거'] = np.sum(tmp, axis=1)
```

```
df_r.head()
```

col	구분	강 간, 추행 검거	강 간, 추행 발생	강 도 검거	강 도 발생	살 인 검거	살 인 발생	절도 검거	절도 발생	폭력 검거	폭력 발생	구 별	lat
0	강남	433.0	513.0	10.0	12.0	3.0	5.0	1016.0	1941.0	1911.0	2283.0	강남구	37.50943
1	강동	167.0	188.0	14.0	15.0	3.0	3.0	781.0	1640.0	1697.0	1942.0	강동구	37.52851
2	강북	132.0	159.0	4.0	5.0	8.0	9.0	490.0	780.0	1625.0	1817.0	강북구	37.63730
3	강서	222.0	265.0	5.0	6.0	9.0	10.0	1070.0	1689.0	2087.0	2445.0	양천구	37.53978
	관												

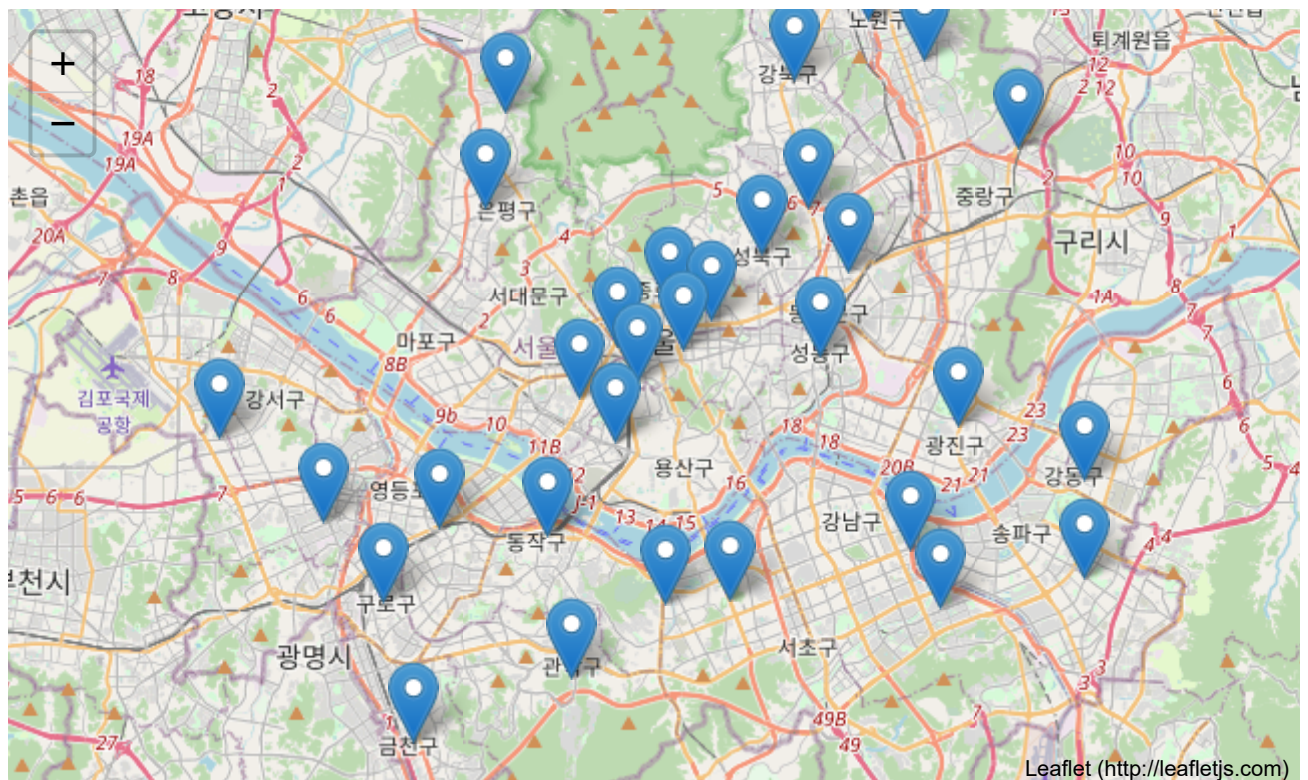
경찰서의 위치 확인

```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11)
```

```
for n in df_r.index:
```

```
    folium.Marker([df_r['lat'][n],
                    df_r['lng'][n]]).add_to(map)
```

map



경찰서의 검거율을 원의 넓이로 표현

```
map = folium.Map(location=[37.5502, 126.982], zoom_start=11)
```

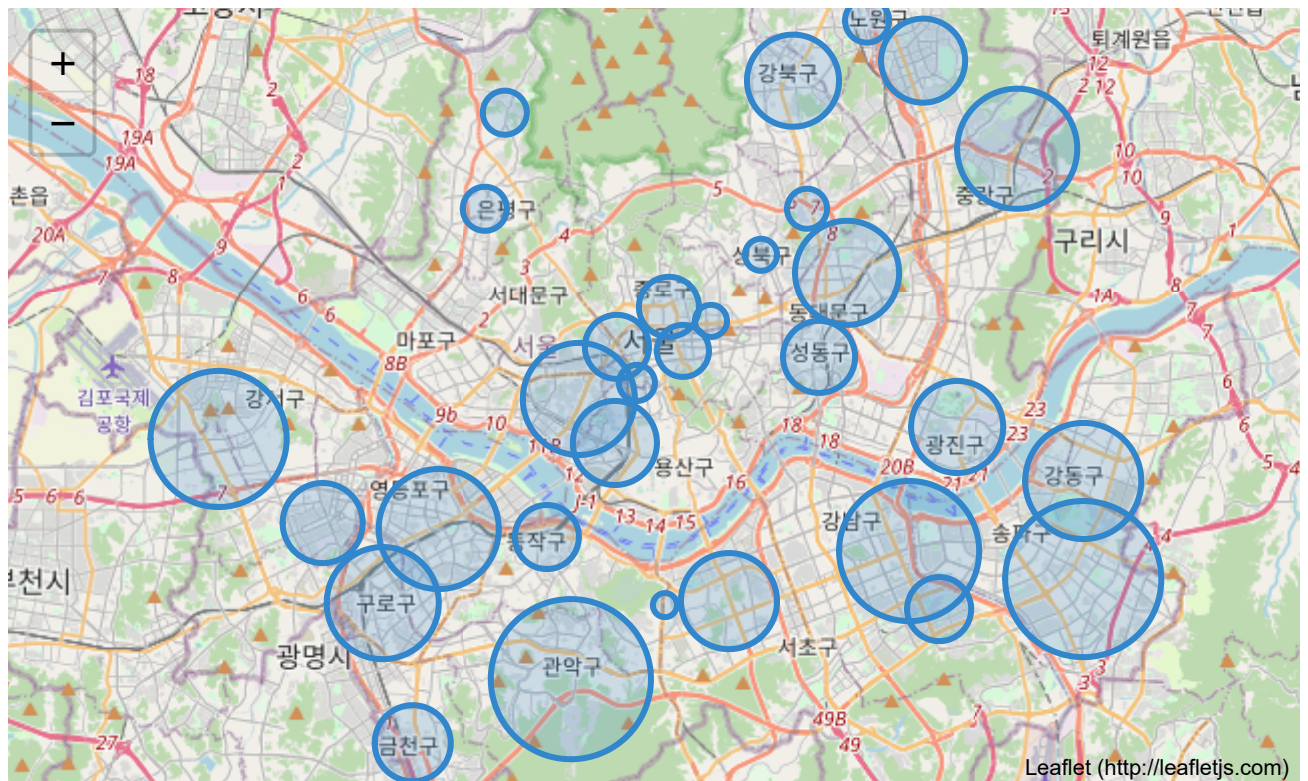


```

for n in df_r.index:
    folium.CircleMarker([df_r['lat'][n], df_r['lng'][n]],
                        radius = df_r['검거'][n]*10,
                        color='#3186cc', fill_color='#3186cc', fill=True).add_to(map)

```

map



```

map = folium.Map(location=[37.5502, 126.982], zoom_start=11)

```

```

map.choropleth(geo_data = geo_str,
               data = crime_anal_norm['범죄'],
               columns = [crime_anal_norm.index, crime_anal_norm['범죄']],
               fill_color = 'PuRd', #PuRd, YlGnBu
               key_on = 'feature.id')

```

```

for n in df_r.index:
    folium.CircleMarker([df_r['lat'][n], df_r['lng'][n]],
                        radius = df_r['검거'][n]*10,
                        color='#3186cc', fill_color='#3186cc', fill=True).add_to(map)

```

map

/usr/local/lib/python3.7/dist-packages/folium/folium.py:426: FutureWarning: The choropleth
FutureWarning

