# Semi-project

2팀: 서태원, 김예린, 김의중, 장윤식, 조희정

### CONTENTS



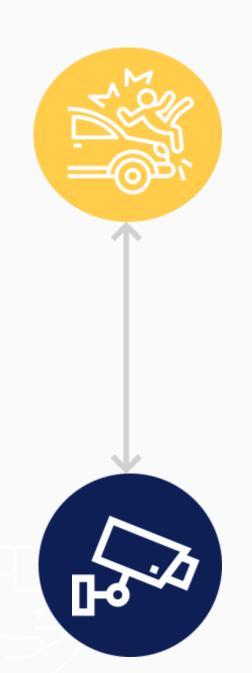








### 분석 주제 및 목표

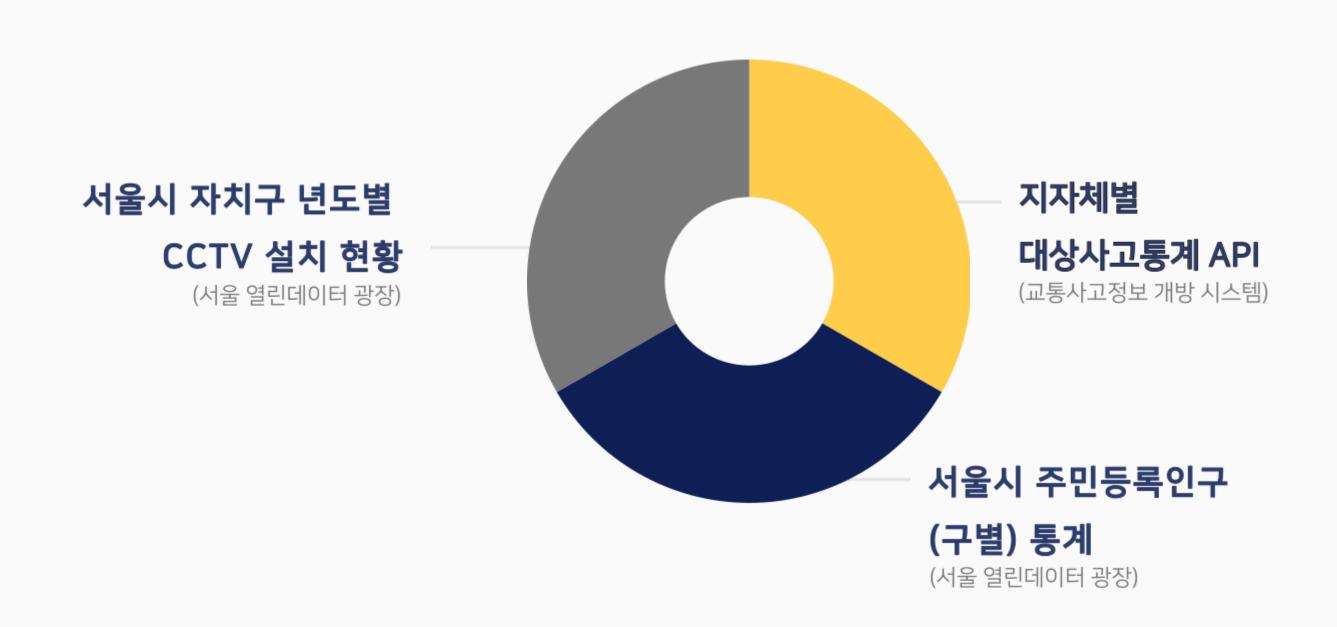


서울시 각 구별 뺑소니 건수와 CCTV 설치 개수와의 연관성



CCTV가 많을수록 뺑소니 건수가 적을것이다

## 데이터 수집 방법



## 데이터 수집 방법

#### 서울시 자치구 년도별 CCTV 설치 현황

| 2  | 구분   | 총계     | 2012년 이 | 2012년 | 2013년 | 2014년 | 2015년 | 2016년 |
|----|------|--------|---------|-------|-------|-------|-------|-------|
| 3  | 계    | 77,032 | 7,667   | 2,200 | 3,491 | 4,439 | 6,582 | 8,129 |
| 4  | 종로구  | 1,772  | 813     | 0     | 0     | 210   | 150   | 1     |
| 5  | 중 구  | 2,333  | 16      | 114   | 87    | 77    | 236   | 240   |
| 6  | 용산구  | 2,383  | 34      | 71    | 234   | 125   | 221   | 298   |
| 7  | 성동구  | 3,602  | 448     | 125   | 212   | 105   | 339   | 310   |
| 8  | 광진구  | 2,588  | 35      | 57    | 100   | 187   | 98    | 52    |
| 9  | 동대문구 | 2,497  | 1,090   | 146   | 60    | 29    | 111   | 233   |
| 10 | 중랑구  | 3,296  | 302     | 24    | 253   | 88    | 141   | 161   |
| 11 | 성북구  | 3,958  | 83      | 78    | 170   | 230   | 323   | 594   |
| 12 | 강북구  | 2,462  | 0       | 0     | 24    | 65    | 105   | 243   |
| 13 | 도봉구  | 1,629  | 39      | 22    | 96    | 181   | 79    | 159   |
| 14 | 노원구  | 2,415  | 0       | 97    | 193   | 77    | 516   | 331   |
| 15 | 은평구  | 3,791  | 14      | 3     | 44    | 332   | 329   | 555   |
| 16 | 서대문구 | 2,940  | 730     | 253   | 155   | 108   | 137   | 223   |
| 17 | 마포구  | 2,421  | 58      | 93    | 77    | 61    | 162   | 454   |
| 18 | 양천구  | 3,312  | 1,413   | 167   | 181   | 143   | 180   | 350   |
| 19 | 강서구  | 2,560  | 55      | 129   | 106   | 213   | 188   | 168   |
| 20 | 구로구  | 4,075  | 852     | 216   | 349   | 187   | 268   | 326   |
| 21 | 금천구  | 2,374  | 0       | 0     | 178   | 80    | 361   | 133   |
| 22 | 영등포구 | 3,778  | 572     | 136   | 238   | 123   | 209   | 248   |

### 서울시 주민등록인구 (구별) 통계

| 기간』   | 자치구   | 세대.        | 합계          |            |            |  |  |
|-------|-------|------------|-------------|------------|------------|--|--|
|       |       |            | Al⊿         | 남자』        | 여자』        |  |  |
|       | 합계』   | 4,189,839』 | 10,204,057』 | 5,007,379』 | 5,196,678』 |  |  |
|       | 종로구』  | 72,118』    | 161,922』    | 79,322』    | 82,600』    |  |  |
|       | 중구』   | 59,651』    | 134,409』    | 66,452』    | 67,957』    |  |  |
|       | 용산구♪  | 106,607』   | 245,102』    | 119,636』   | 125,466』   |  |  |
|       | 성동구♪  | 128,727』   | 307,161.    | 152,006』   | 155,155』   |  |  |
|       | 광진구♪  | 158,390』   | 372,104』    | 181,251.   | 190,853』   |  |  |
|       | 동대문구, | 159,426』   | 370,312,    | 183,444    | 186,868』   |  |  |
|       | 중랑구』  | 177,400』   | 415,677』    | 206,772』   | 208,905』   |  |  |
|       | 성북구♪  | 188,304』   | 461,617』    | 224,504』   | 237,113』   |  |  |
|       | 강북구』  | 141,229』   | 330,704.    | 162,211.   | 168,493』   |  |  |
|       | 도봉구』  | 136,898』   | 350,272』    | 172,006』   | 178,266』   |  |  |
|       | 노원구♪  | 219,736』   | 571,212,    | 277,866』   | 293,346』   |  |  |
| 2016. | 은평구』  | 201,891』   | 495,937』    | 241,048』   | 254,889』   |  |  |
| 2010  | 서대문구』 | 136,766』   | 325,871.    | 156,597』   | 169,274』   |  |  |
|       | 마포구』  | 169,432』   | 390,887』    | 186,668』   | 204,219』   |  |  |
|       | 양천구♪  | 177,434』   | 481,845』    | 238,163』   | 243,682』   |  |  |
|       | 강서구♪  | 245,045』   | 602,104』    | 293,734』   | 308,370』   |  |  |
|       | 구로구,  | 172,220』   | 449,600』    | 225,411.   | 224,189』   |  |  |
|       | 금천구♪  | 104,483』   | 254,654』    | 130,429』   | 124,225』   |  |  |
|       | 영등포구』 | 165,729』   | 406,779』    | 204,712.   | 202,067』   |  |  |

### 분석 프로세스



#### 데이터 추출

### 뺑소니 사건현황 데이터

▶ 반복문으로 연도별(2016~2021) 서울시의 데이터만 추출



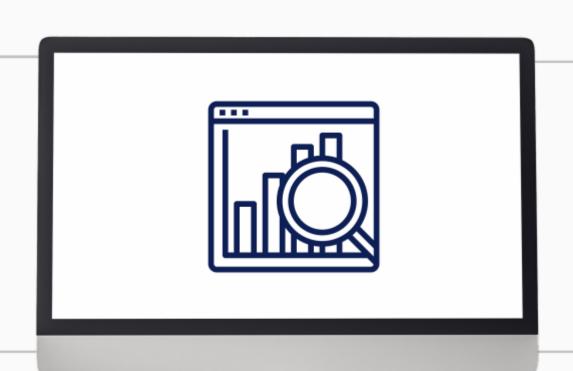
#### 데이터 가공 및 분석 변수 처리

뺑소니 사건 현황 데이터

▶ 필요없는 열 데이터삭제

서울시 CCTV 데이터

▶ 신규 CCTV 발생 건수를 누적 데이터로 변환



#### 데이터 결합

인구수 대비 뺑소니 사고 건수 연도별 데이터 열방향으로 결합◀



#### 데이터시각화

- heatmap(seaborn) <
- jointplot(seaborn)
  - plot(matplotlib)
- rcParams(matplotlib)
- choropleth map(folium)
  - pie(matplotlib)



### 분석 프로세스

```
arr_col = df_year.columns.values

df_year[arr_col] = df_year[arr_col].cumsum()

df_year[arr_col]
```

▶ cumsum함수 사용해서 cctv갯수 구한부분

```
cctv_columns = df_cctv.columns.values
for column in cctv_columns:
    df_cctv[column] = df_cctv[column].str.replace(',','')|
    df_cctv[column] = df_cctv[column].apply(pd.to_numeric)
```

모든 데이터 정수(int64)변환,

','콤마 제거(ex)1,000→ 1000)

▶ 기존 데이터프레임을 iterrows함수를 통해 새로운 데이터 프레임 만들기

### 시각화

- 100

- 80

- 60

- 40

- 20

|     |        |            | 년도        | 별 뺑소니 사    | 건 수       |           |
|-----|--------|------------|-----------|------------|-----------|-----------|
|     | 강남구 -  | 101.000000 | 69.000000 | 75.000000  | 65.000000 | 96.000000 |
|     | 서초구 -  | 61.000000  | 61.000000 | 73.000000  | 58.000000 | 66.000000 |
|     | 강서구 -  | 14.000000  | 25.000000 | 31.000000  | 56.000000 | 52.000000 |
|     | 영등포구 - | 44.000000  | 35.000000 | 39.000000  | 27.000000 | 36.000000 |
|     | 금천구 -  | 19.000000  | 25.000000 | 36.000000  | 30.000000 | 34.000000 |
|     | 구로구 -  | 32.000000  | 61.000000 | 44.000000  | 35.000000 | 32.000000 |
|     | 노원구 -  | 32.000000  | 19.000000 | 27.000000  | 49.000000 | 32.000000 |
|     | 송파구 -  | 47.000000  | 34.000000 | 36.000000  | 30.000000 | 32.000000 |
|     | 종로구 -  | 30.000000  | 23.000000 | 17.000000  | 20.000000 | 31.000000 |
|     | 광진구 -  | 31.000000  | 28.000000 | 32.000000  | 29.000000 | 31.000000 |
|     | 은평구 -  | 17.000000  | 22.000000 | 21.000000  | 20.000000 | 31.000000 |
|     | 강동구 -  | 36.000000  | 16.000000 | 37.000000  | 29.000000 | 29.000000 |
| 자치구 | 관악구 -  | 18.000000  | 19.000000 | 21.000000  | 23.000000 | 28.000000 |
| K   | 중랑구 -  | 13.000000  | 25.000000 | 29.000000  | 25.000000 | 26.000000 |
|     | 용산구 -  | 34.000000  | 22.000000 | 41.000000  | 33.000000 | 25.000000 |
|     | 동작구 -  | 21.000000  | 26.000000 | 18.000000  | 18.000000 | 22.000000 |
|     | 동대문구 - | 34.000000  | 40.000000 | 33.000000  | 29.000000 | 20.000000 |
|     | 서대문구 - | 29.000000  | 20.000000 | 15.000000  | 20.000000 | 19.000000 |
|     | 성북구 -  | 25.000000  | 29.000000 | 29.000000  | 15.000000 | 19.000000 |
|     | 강북구 -  | 25.000000  | 26.000000 | 26.000000  | 20.000000 | 18.000000 |
|     | 양천구 -  | 27.000000  | 11.000000 | 16.000000  | 20.000000 | 17.000000 |
|     | 마포구 -  | 56.000000  | 53.000000 | 42.000000  | 31.000000 | 16.000000 |
|     | 성동구 -  | 15.000000  | 7.000000  | 24.000000  | 11.000000 | 15.000000 |
|     | 중구 -   | 6.000000   | 14.000000 | 13.000000  | 16.000000 | 14.000000 |
|     | 도봉구 -  | 23.000000  | 23.000000 | 14.000000  | 12.000000 | 10.000000 |
|     |        | 2016       | 2017      | 2018<br>년도 | 2019      | 2020      |

| CCTV | 운영     | 수 |
|------|--------|---|
| 4770 | 000000 | ) |

- 6000

- 5000

- 4000

- 3000

- 2000

- 1000

|    | 강남구 -  | 3023.000000 | 4022.000000 | 4770.000000 | 5559.000000 | 6501.000000 |
|----|--------|-------------|-------------|-------------|-------------|-------------|
|    | 관악구 -  | 2587.000000 | 3274.000000 | 3937.000000 | 4577.000000 | 4908.000000 |
|    | 구로구 -  | 2198.000000 | 2738.000000 | 3226.000000 | 3660.000000 | 4075.000000 |
| 9  | 영등포구 - | 1526.000000 | 1837.000000 | 2495.000000 | 2560.000000 | 3773.000000 |
|    | 성북구 -  | 1478.000000 | 1938.000000 | 2805.000000 | 3519.000000 | 3770.000000 |
|    | 서초구 -  | 1435.000000 | 2495.000000 | 2923.000000 | 3281.000000 | 3701.000000 |
|    | 은평구 -  | 1277.000000 | 1680.000000 | 2315.000000 | 3372.000000 | 3660.000000 |
|    | 성동구 -  | 1539.000000 | 2413.000000 | 2803.000000 | 3065.000000 | 3526.000000 |
|    | 양천구 -  | 2434.000000 | 2573.000000 | 2713.000000 | 2987.000000 | 3312.000000 |
|    | 중랑구 -  | 969.000000  | 1131.000000 | 1304.000000 | 2353.000000 | 3292.000000 |
| ,  | 서대문구 - | 1606.000000 | 1962.000000 | 2199.000000 | 2542.000000 | 2939.000000 |
|    | 송파구 -  | 667.000000  | 908.000000  | 1450.000000 | 2518.000000 | 2753.000000 |
| 자자 | 광진구 -  | 529.000000  | 1204.000000 | 1669.000000 | 2381.000000 | 2556.000000 |
| K  | 강동구 -  | 927.000000  | 1200.000000 | 1577.000000 | 1933.000000 | 2547.000000 |
|    | 강서구 -  | 859.000000  | 1365.000000 | 1624.000000 | 2081.000000 | 2437.000000 |
| 4  | 동대문구 - | 1669.000000 | 1805.000000 | 2002.000000 | 2211.000000 | 2434.000000 |
|    | 강북구 -  | 437.000000  | 443.000000  | 835.000000  | 1835.000000 | 2423.000000 |
|    | 마포구 -  | 905.000000  | 1264.000000 | 1607.000000 | 2101.000000 | 2401.000000 |
|    | 용산구 -  | 983.000000  | 1334.000000 | 1459.000000 | 1766.000000 | 2383.000000 |
|    | 금천구 -  | 752,000000  | 948.000000  | 1488 000000 | 1857.000000 | 2365.000000 |
|    | 노원구 -  | 1214.000000 | 1389 000000 | 1605.000000 | 1925.000000 | 2311.000000 |
|    | 동작구 -  | 974.000000  | 1245.000000 | 1545 000000 | 1867.000000 | 2286.000000 |
|    | 중구 -   | 770.000000  | 1142.000000 | 1528.000000 | 1683.000000 | 2044.000000 |
|    | 종로구 -  | 1174.000000 | 1435.000000 | 1520.000000 | 1529.000000 | 1729.000000 |
|    | 도봉구 -  | 576.000000  | 710.000000  | 932.000000  | 1130.000000 | 1298.000000 |
|    |        | 2016        | 2017        | 2018<br>년도  | 2019        | 2020        |

#### CCTV 운영 장비 수 대비 뺑소니 사건 수

- 0.06

- 0.05

- 0.04

- 0.03

- 0.02

- 0.01

|     | 강서구 -  | 0.016298 | 0.018315 | 0.019089            | 0.026910 | 0.021338  |
|-----|--------|----------|----------|---------------------|----------|-----------|
|     | 종로구 -  | 0.025554 | 0.016028 | 0.011184            | 0.013080 | 0.017929  |
|     | 서초구 -  | 0.042509 | 0.024449 | 0.024974            | 0.017678 | 0.017833  |
|     | 강남구 -  | 0.033411 | 0.017156 | 0.015723            | 0.011693 | 0.014767  |
|     | 금천구 -  | 0.025266 | 0.026371 | 0.024194            | 0.016155 | 0.014376  |
|     | 노원구 -  | 0.026359 | 0.013679 | 0.016822            | 0.025455 | 0.013847  |
|     | 광진구 -  | 0.058601 | 0.023256 | 0.019173            | 0.012180 | 0.012128  |
|     | 송파구 -  | 0.070465 | 0.037445 | 0.024828            | 0.011914 | 0.011624  |
|     | 강동구 -  | 0.038835 | 0.013333 | 0.023462            | 0.015003 | 0.011386  |
|     | 용산구 -  | 0.034588 | 0.016492 | 0.028101            | 0.018686 | 0.010491  |
|     | 동작구 -  | 0.021561 | 0.020884 | 0.011650            | 0.009641 | 0.009624  |
|     | 영등포구 - | 0.028834 | 0.019053 | 0.015631            | 0.010547 | 0.009541. |
| 자치구 | 은평구 -  | 0.013312 | 0.013095 | 0.009071            | 0.005931 | 0.008470  |
| Κ,  | 동대문구 - | 0.020371 | 0.022161 | 0.016484            | 0.013116 | 0.008217  |
|     | 중랑구 -  | 0.013416 | 0.022104 | 0.022:239           | 0.010625 | 0.007898  |
|     | 구로구 -  | 0.014559 | 0.022279 | 0.013639            | 0.009563 | 0.007853  |
|     | 도봉구 -  | 0.039931 | 0.032394 | 0.015021            | 0.010619 | 0.007704  |
|     | 강북구 -  | 0.057208 | 0.058691 | 0.031138            | 0.010899 | 0.007429  |
|     | 중구 -   | 0.007792 | 0.012259 | 0.008508            | 0.009507 | 0.006849  |
|     | 마포구 -  | 0.061878 | 0.041930 | 0.026136            | 0.014755 | 0.006664  |
| ,   | 서대문구 - | 0.018057 | 0.010194 | 0.006821            | 0.007868 | 0.006465  |
|     | 관악구 -  | 0.006958 | 0.005803 | 0.005334            | 0.005025 | 0.005705  |
|     | 양천구 -  | 0.011093 | 0.004275 | 0.005898            | 0.006696 | 0.005133  |
|     | 성북구 -  | 0.016915 | 0.014964 | 0.010339            | 0.004263 | 0.005040  |
|     | 성동구 -  | 0.009747 | 0.002901 | 0.008562            | 0.003589 | 0.004254  |
|     |        | 2016     | 2017     | 20 <b>1</b> 8<br>년도 | 2019     | 2020      |

### 산점도 분석

```
#독립표본 2016년 뺑소니 & 2016년 cotv
stats.ttest_ind(df_tot_2016["횟수"], df_tot_2016["CCTV 갯수"])
```

Ttest\_indResult(statistic=-9.461060880318874, pvalue=1.5044704755337304e-12)

```
#독립표본 2017년 뺑소니 & 2017년 cotv
stats.ttest_ind(df_tot_2017["횟수"], df_tot_2017["CCTV 갯수"])
```

Ttest\_indResult(statistic=-10.001797294846918, pvalue=2.518048059917295e-13)

```
#독립표본 2018년 뺑소니 & 2018년 octv
stats.ttest_ind(df_tot_2018["횟수"], df_tot_2018["CCTV 갯수"])
```

Ttest\_indResult(statistic=-11.0217905739632, pvalue=9.526862941424921e-15)

```
#독립표본 2019년 뺑소니 & 2019년 cotv
stats.ttest_ind(df_tot_2019["횟수"], df_tot_2019["CCTV 갯수"])
```

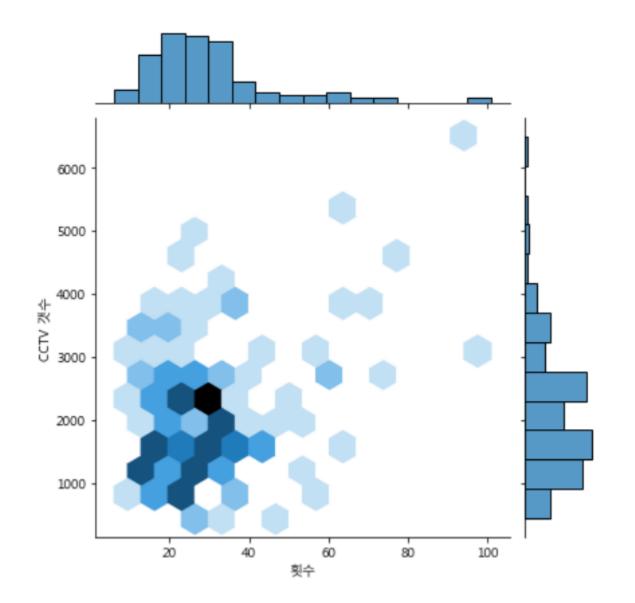
Ttest\_indResult(statistic=-12.647776812981668, pvalue=6.78313604865569e-17)

```
#독립표본 2020년 벨소니 & 2020년 cctv
stats.ttest_ind(df_tot_2020["횟수"], df_tot_2020["CCTV 갯수"])
```

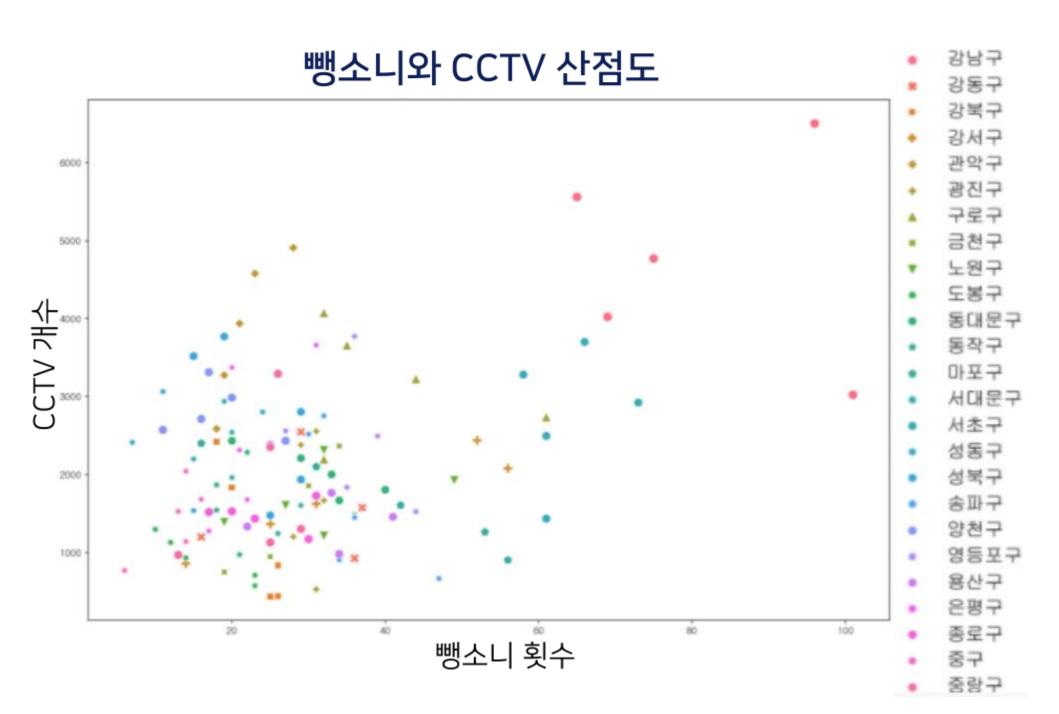
Ttest\_indResult(statistic=-13.675832722920964, pvalue=3.542932962562021e-18)

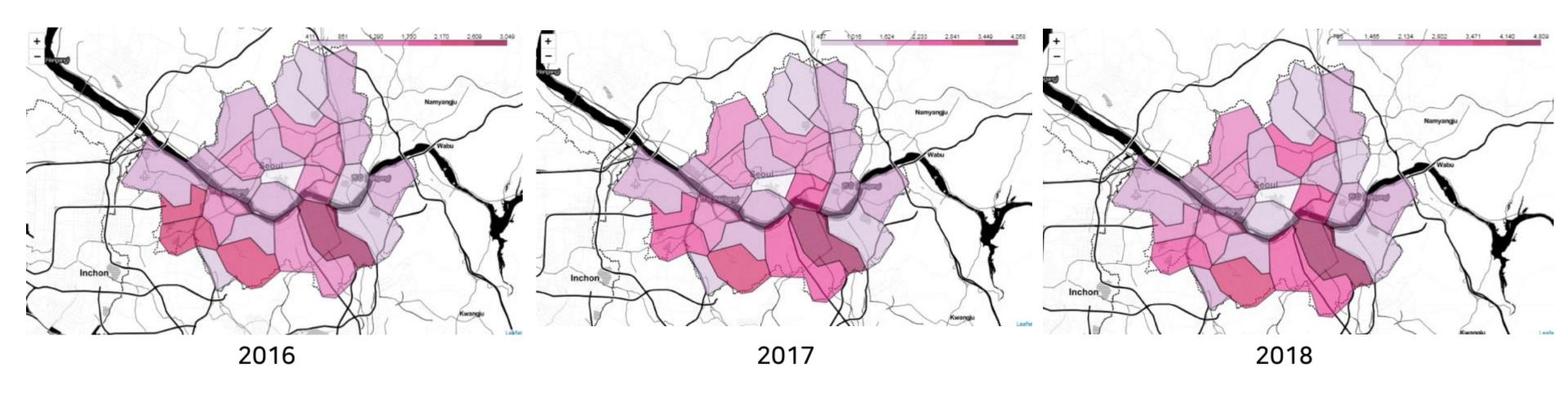
```
import scipy as sp #과학계산을 위하 사용# 통계학기법사용시 활용
from scipy import stats # statistics
import seaborn as sns
sns.jointplot(x=df_tot['횟수'], y=df_tot['CCTV 갯수'], kind='hex')
```

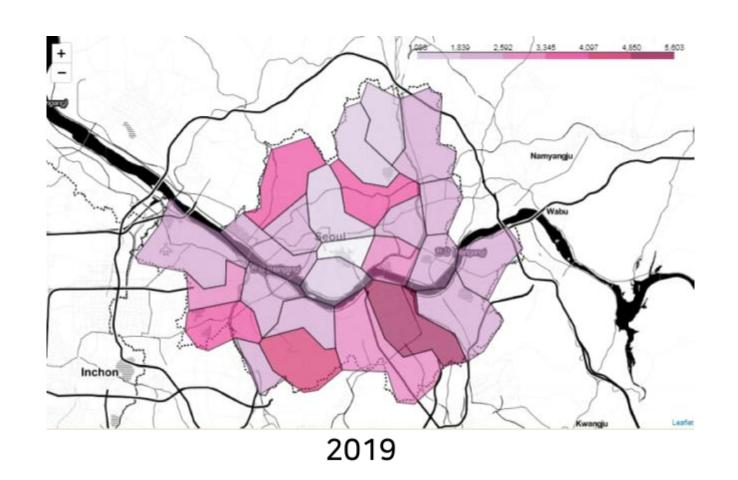
<seaborn.axisgrid.JointGrid at Oxbf8b908>

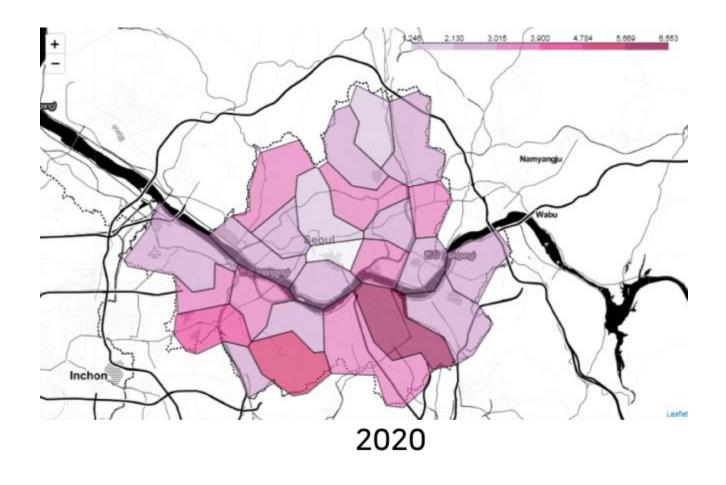


```
for name, group in groups :
   sns.scatterplot(x=group['횟수'],
              y=group['CCTV 갯수'],
              hue='자치구', # different colors by group
               style='자치구', # different shapes by group
               s=150, # marker size
               data=df_acc_cctv,
               legend=False)
plt.legend(fontsize=15, loc='upper left', bbox_to_anchor=(1.05, 1))
plt.title('뺑소니와 CCTV 산점도', fontsize=20)
plt.xlabel('뺑소니 횟수', fontsize=14)
plt.ylabel('CCTV 갯수', fontsize=14)
plt.show()
```

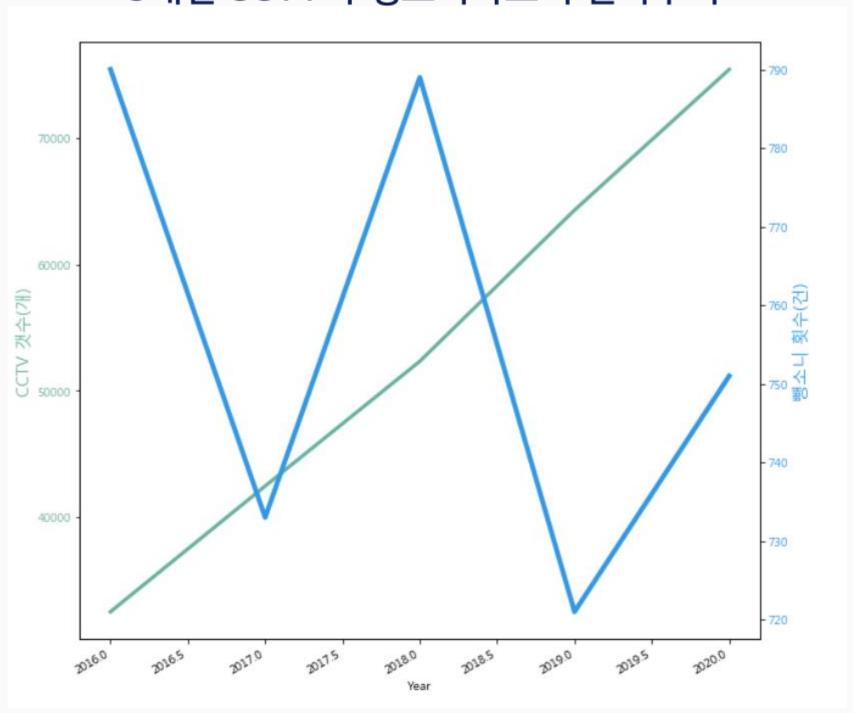


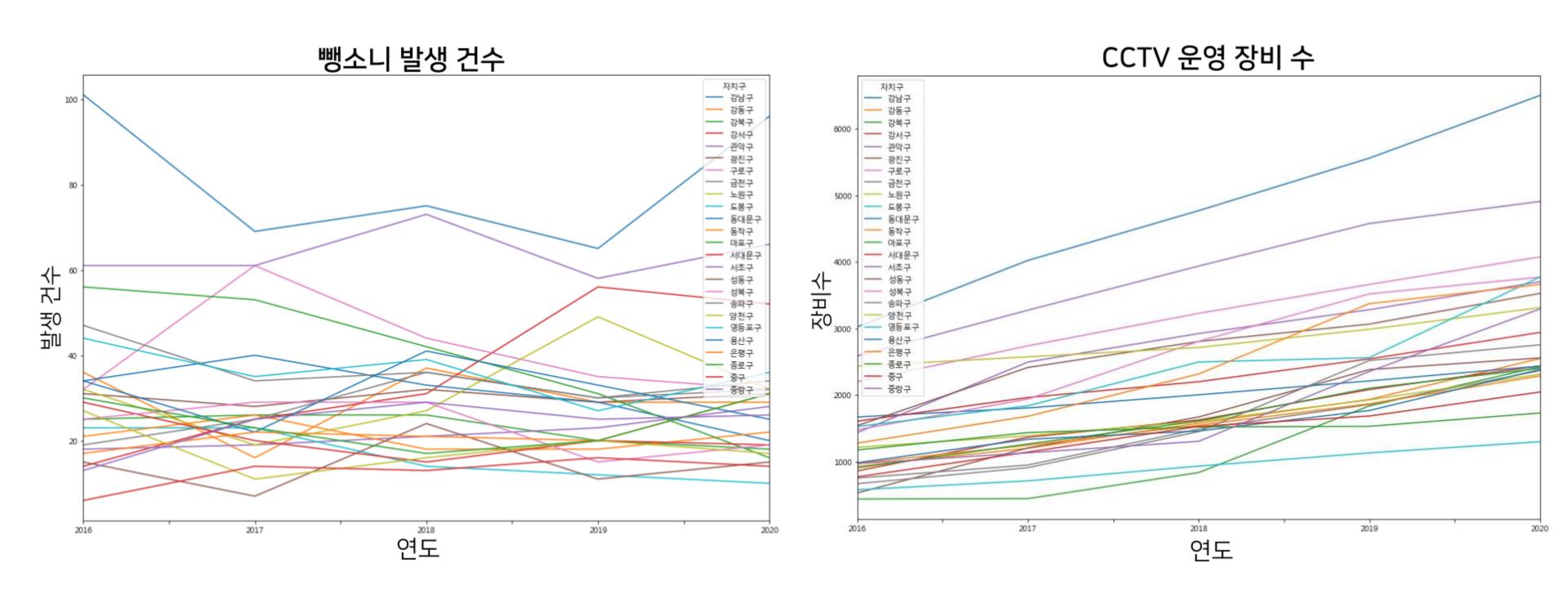






5개년 CCTV와 뺑소니자료의 변화추이

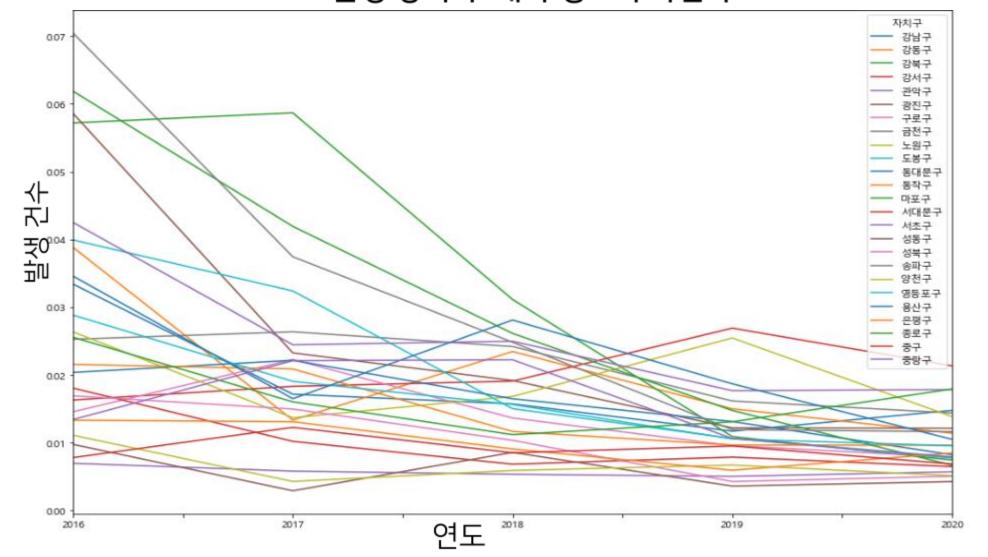




### 두 데이터간의 피어슨 상관관계분석

```
year_data={'년도':year_index,'CCTV':df_tot_CCTV,'뺑소니':df_tot_run}
year_data
df_year_data=pd.DataFrame(year_data)
pivot_year = pd.pivot_table(df_year_data, index='년도', aggfunc=np.sum)
pivot_year
```

#### CCTV 운영 장비 수 대비 뺑소니 사건 수



| corr_df.head() |      |     |     |         |  |
|----------------|------|-----|-----|---------|--|
|                | 년도   | 자치구 | 횟수  | CCTV 갯수 |  |
| 0              | 2016 | 강남구 | 101 | 3023    |  |
| 1              | 2016 | 강동구 | 36  | 927     |  |
| 2              | 2016 | 강북구 | 25  | 437     |  |
| 3              | 2016 | 강서구 | 14  | 859     |  |
| 4              | 2016 | 관악구 | 18  | 2587    |  |
|                |      |     |     |         |  |

|         | 횟수       | CCTV 갯수  |
|---------|----------|----------|
| 횟수      | 1.000000 | 0.356749 |
| CCTV 갯수 | 0.356749 | 1.000000 |

corr\_df.corr()

|      | CCTV  | 뺑소니 |
|------|-------|-----|
| 년도   |       |     |
| 2016 | 32508 | 790 |
| 2017 | 42455 | 733 |
| 2018 | 52331 | 789 |
| 2019 | 64292 | 721 |
| 2020 | 75424 | 751 |



CCTV와 뺑소니 발생건수는 약한 양의 상관관계

## 후속 과제



CCTV 개수와 다른 데이터와의 관계 인구수가 아닌 구별 차량대수를 반영한 값 미래 뺑소니 발생 건수 수치 예측 10개년 데이터 반영

주택가격 ○
 평당가
 혼인 및 이혼 수

## THANK YOU!