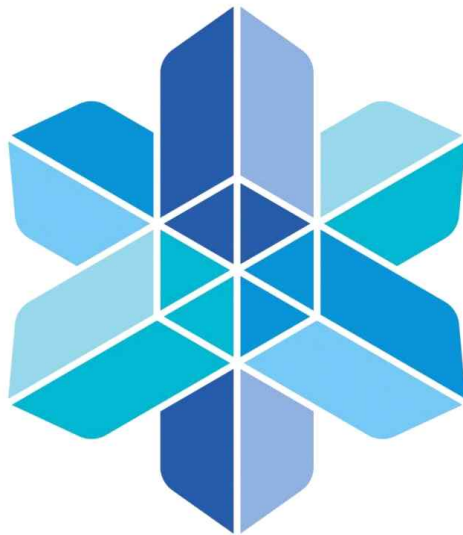


# 서울특별시 인구수 보고서

2019150030 - 염예찬

**한국공학대학교**  
TECH UNIVERSITY OF KOREA



# 목차

- 1. 서론
- 2. 주제
- 3. 공공데이터

## 4. 본문

- (I) 지역구별 평균 나이, 신생아수, 초등취학 연령 인구수 비교
- (II) 서울시 5대 범죄 발생현황과 평균연령과의 관계
- (III) 서울시 5대 범죄 발생현황의 발생 건수/검거 건수와 신생아수의 비교
- (IV) 자치구 전체 인구수 대비 범죄 건수와 신생아 수의 상관관계
- (V) 자치구의 초등학교 수가 많으면 신생아 수의 상관관계
- (VI) 자치구의 지역별 특징에 따른 신생아 수의 지역적 차이
- (VII) 자치구별 1인당 지역 내 총생산에 따른 신생아 수의 변화
- (VIII) 상급학교 진학률에 따른, 신생아수 차이

## 5. 결론

## 1. 서론

평소에 대한민국의 인구수별 분포와 점점 고령화 되가는 사회에 대해서 걱정을 하며, 대체 어떤 유의미한 해결책이 있는지에 대해서 관심이 있었다. 그러다 최근에 부모님이 청약에 당첨되셔서 서울에서 10년 넘게 살던 동네에서 다른 자치구의 신축 아파트로 이사를 가게 되었다. 이사를 가고 나서 이상한 점을 느꼈는데, 같은 5층에 사는 다른 3가구가 전부 유치원생에서 아기 정도의 나이 대를 가지고 있다는 점이였다. 이사를 간 초창기에는 별 생각이 안 들었지만 지나고 보니 아파트 단지가 항상 어린애들과 아기들로 시끌벅적하고, 30대의 젊은 부부가 많이 보이며 전체적으로 동네의 평균 연령이 낮으며 활력이 넘친다는 것을 느끼게 되었다.

어째서 이러한 체감차이가 나는지에 대해서 고심해보다가 이전 지역에 비해 달라진 여러 가지 점에 대해 생각해 봤다. 첫째는, 주택 청약이 이루어진 새 아파트다보니 이전 아파트에 비해 청약 당첨에 유리한 신혼부부나 젊은 부부의 접근성이 좋다는 점, 둘째, 이전 지역에 비해 초등학교가 많은 지역이라는 점, 셋째, 이전 지역에 비해 불량 학생들이나, 양아치, 길에서 흡연하는 사람이 유의미하게 적게 발견된다는 것 정도가 있었다. 과연 이러한 요소들이 그 지역 구내의 평균 연령, 신생아 수와 초등 취학 연령 인구와 얼마나 관련이 있는지에 대해 궁금해지기 시작했다.

따라서 신생아수와 평균 연령대와 초등 취학 연령 인구가 자치구별로 어느 정도의 차이가 있는지, 그러한 차이에 대한 유의미한 상관관계가 있는 요소가 있는지에 대해 조사한 보고서에 대해 작성할 것이고, 마지막으로 무너져버린 세대 비율을 되돌리기 위해 신생아수를 늘리기 위해 좋은 요소들이 어떠한 것인지에 대해 정리할 것이다.

## 2. 주제

서울 자치구별 신생아 수, 평균 연령대와 초등 취학 연령자 수에 대한 주변 환경의 영향  
- 자치구별 신생아수, 평균연령대, 초등취학 연령자수를 각각 지역별 지도상의 위치, 보육시설 이용 아동 비율, 사교육비 비율, 상급학교 진학률, 5대 범죄, 초등학교 수, 지역 내 총생산 등과의 비교를 진행할 것이다.

## 3. 공공 데이터

보육시설+이용+아동+비율(2019년)  
사교육비+현황(2019년)  
상급학교\_진학률\_(2019년)  
서울시 5대 범죄 발생현황 (2019년)  
서울시+자치구별+초등학교(2019년)  
자치구별 1인당 지역내총생산 및 수준지수(2019)  
주민등록인구(연령별)\_(2019년)

## 4. 본문

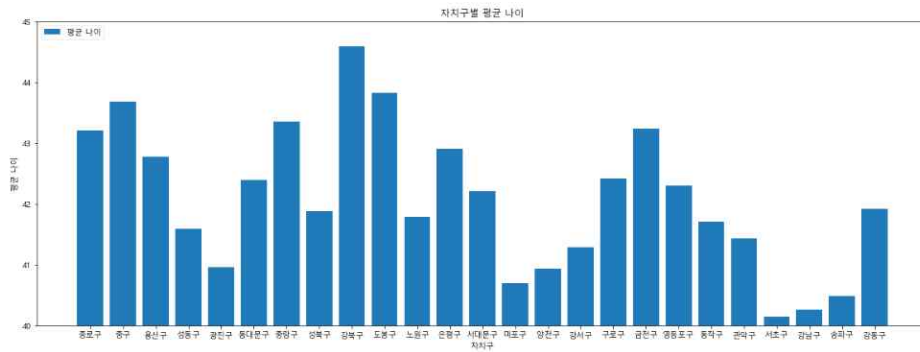
### (1) 지역구별 평균 나이, 신생아수, 초등취학 연령 인구수 비교

지역구별 평균 나이, 신생아수, 초등취학 연령을 비교하기 위하여 기본 쿼리와 이를 시각적으로 비교할 수 있는 그래프가 필요하다.

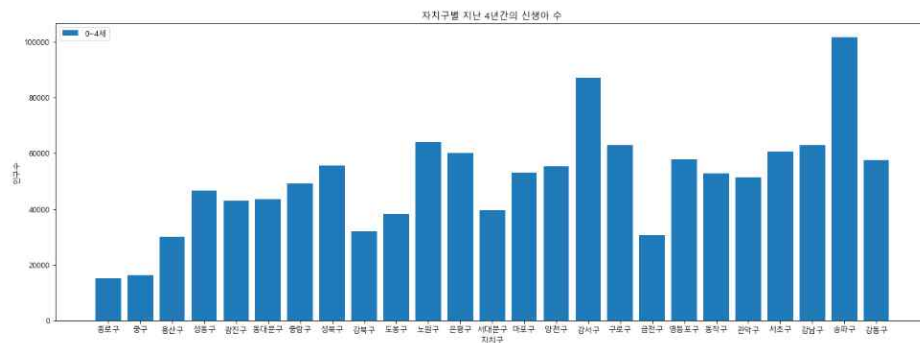
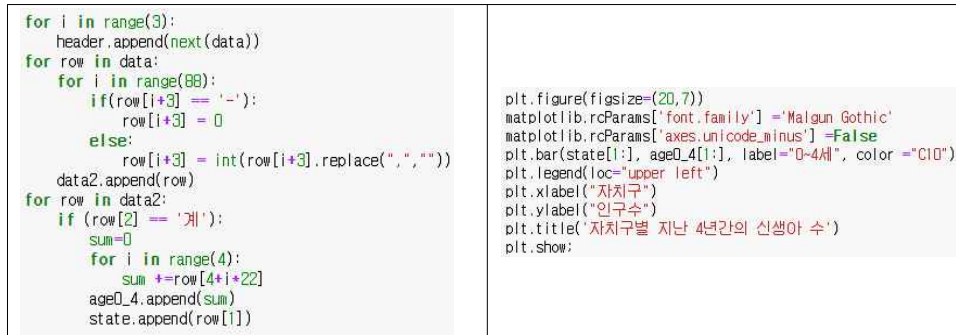
예상> 일반적으로 신생아수, 초등취학 연령 인구수는 비례하고, 평균나이는 반비례할 것이다.

#### ▶ 2019년도 평균나이 그래프

<pre>import csv import matplotlib.pyplot as plt import matplotlib import numpy as np  f = open("주민등록인구(연령별).csv") data = csv.reader(f) data2=[]  for row in range(3):     next(data)  for row in data:     for i in range(88):         if(row[i+3] == '-'):             row[i+3] = 0         else:             row[i+3] = int(row[i+3].replace(",",""))         data2.append(row)  age_avg = [] state = []</pre>	<pre>for row in data2:     if (row[2] == '계'):         sum=0         for i in range(21):             sum += row[4 + i] + (5 + i + 2) / row[3]             sum += row[26 + i] + (5 + i + 2) / row[25]             sum += row[48 + i] + (5 + i + 2) / row[47]             sum += row[70 + i] + (5 + i + 2) / row[69]         age_avg.append(sum/4)         state.append(row[1])  plt.figure(figsize=(20,7)) matplotlib.rcParams['font.family'] = 'Malgun Gothic' matplotlib.rcParams['axes.unicode_minus'] = False plt.bar(state[1:], age_avg[1:], label="평균 나이", color = "C10") plt.legend(loc="upper left") plt.xlabel("자치구") plt.ylabel("평균 나이") plt.ylim(40,45) plt.title("자치구별 평균 나이") plt.show;</pre>
--	---

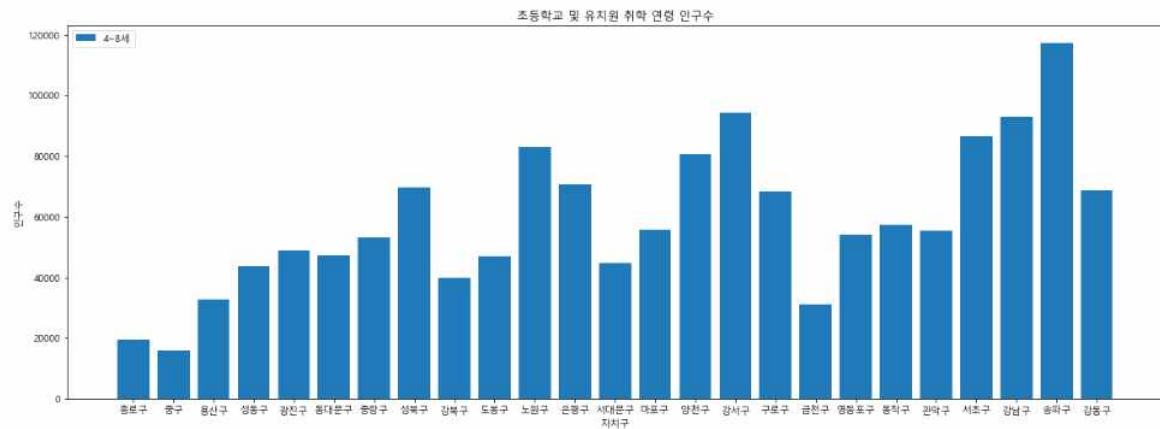


### ▶ 2019년기준 지난 4년간의 신생아수

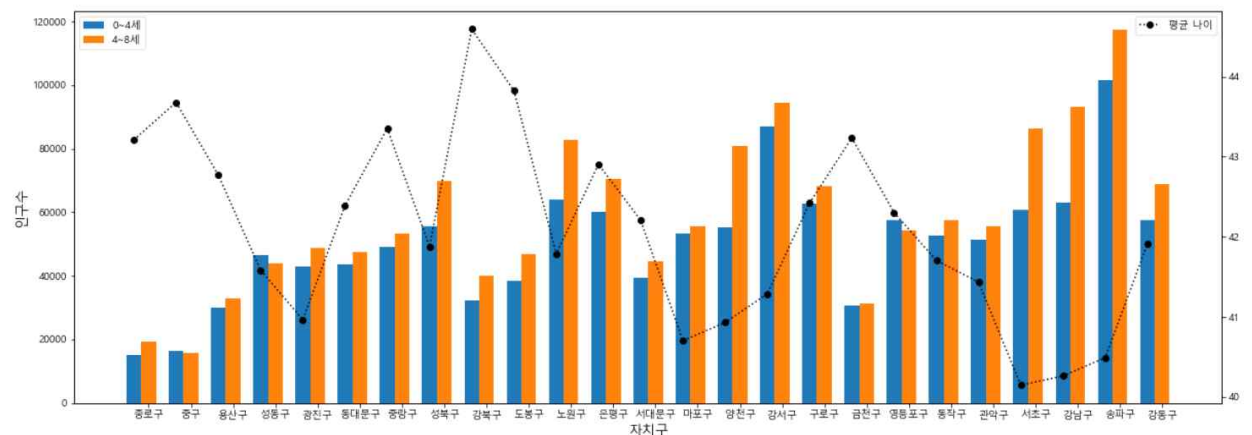
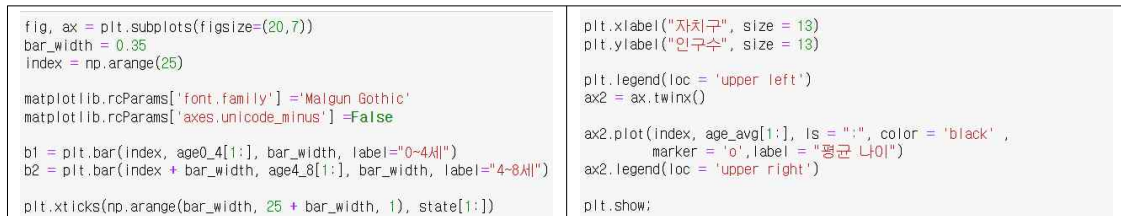


### ▶ 2019년도 초등학교 및 유치원 취학 연령 인구수 그래프





## ▶ 해석



- 위의 데이터들을 합쳐서 그래프로 나타낸 결과 예상한 의견과 같이 0~4세 즉 지난 4년간 신생아수가 높은 지역은 대체적으로 4~8세 즉 유치원~초등학생 취학연령 인원수도 높은 경향을 보였고, 그에 비해 평균나이는 반비례하게 나타났다.

이에 신생아 데이터들과 다른 데이터를 비교하면 동시에 유치원, 초등생 취학 연령 인원수, 평균나이라도 비교할 수 있으므로, 다음 데이터들과의 비교는 신생아 데이터를 대표로 비교한다.

결론 => 신생아 수와 유치원, 초등생 취학연령인원수는 비례하고, 평균 연령은 반비례한다.

## (II) 서울시 5대 범죄 발생현황과 평균연령과의 관계

서울시 5대 범죄 발생현황과 평균연령과의 상관관계에 대해서 비교해 볼 것이다. 이를 위하여 이전 데이터인 신생아 쿼리와 “서울시 범죄 발생현황\_2019.csv”에 대한 쿼리가 필요하다.

예상> 범죄가 많이 발생하는 지역이면 아이를 키우기 좋지 않은 환경이므로 신생아수가 적을 것이다.

### ▶ 2019년도 서울시 5대 범죄 발생현황 쿼리

```
f = open("서울시 범죄 발생현황_2019.csv")
data = csv.reader(f)
pd.set_option('display.max_row', 8)
for i in range(4):
    header.append(next(data))
df = pd.DataFrame(data)
df.columns = header[2]
df
```

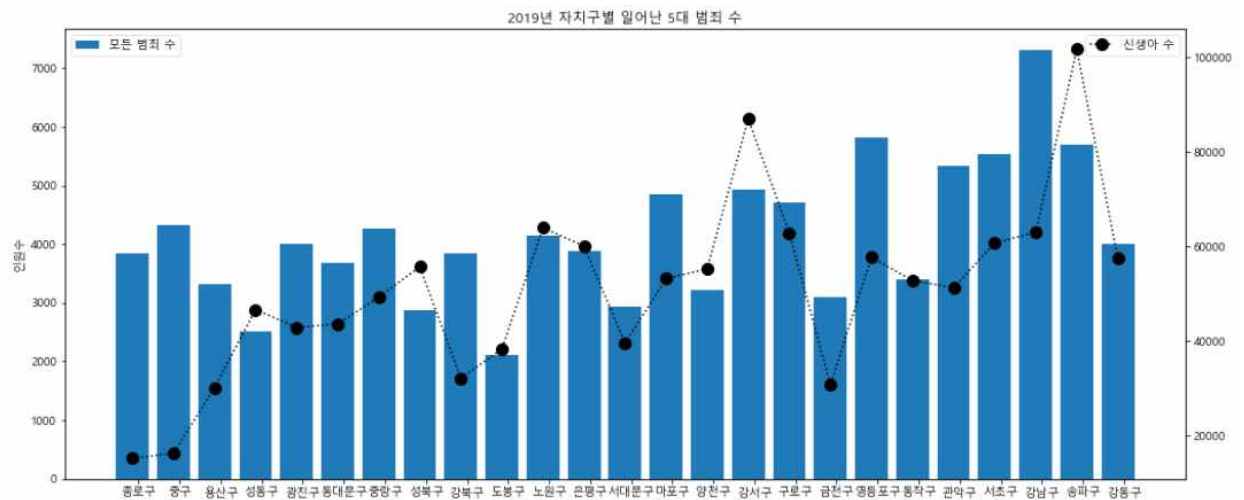
자치구별(1)	자치구별(2)	소계	소계	살인	살인	강도	강도	강간·강제추행	강간·강제추행	절도	절도	폭력	폭력	
0	합계	소계	103,668	74,805	136	128	136	133	6,469	6,007	42,204	21,284	54,723	47,253
1	종로구	3,846	4,117	4	5	7	8	238	1,143	1,515	1,069	2,082	1,892	
2	중구	4,327	2,804	2	1	6	5	195	115	2,202	1,050	1,922	1,633	
3	용산구	3,313	2,611	3	3	3	4	272	237	999	544	2,036	1,823	
4	성동구	2,512	1,838	6	5	9	10	133	96	970	511	1,394	1,216	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
21	관악구	5,328	3,810	13	10	10	11	408	335	2,223	1,085	2,674	2,369	
22	서초구	5,542	3,750	5	5	7	5	616	412	2,270	1,118	2,644	2,210	
23	강남구	7,304	5,069	5	3	5	6	666	562	2,970	1,339	3,658	3,159	
24	송파구	5,698	3,799	7	8	10	10	273	230	2,416	1,016	2,992	2,535	
25	강동구	4,014	2,955	6	5	15	14	180	155	1,650	907	2,163	1,874	

### ▶ 해석

```
fig, ax = plt.subplots(figsize=(17,7))
bar_width = 0.15
index = np.arange(25)

index = np.arange(25)
plt.bar(index, prime[0], label = "모든 범죄 수")
plt.legend(loc = 'upper left')
plt.ylabel("인원수")
ax2 = ax.twinx()
```

```
ax2.plot(index , age0_4[1:], ls = ":", color = 'black' ,
        marker = 'o', markersize = 10, label = "신생아 수")
ax2.legend(loc = 'upper right')
plt.xticks(np.arange(bar_width, 25 + bar_width, 1), state)
plt.title("2019년 자치구별 일어난 5대 범죄 수")
plt.show;
```



- 예상 외로 크게 관련이 없는 것처럼 보였다. 그 이유에 대해서 생각해보니 당연히 지역별로 사는 인구수가 다르므로 지역구별로 인원수 대비 범죄율이 아닌 단순 범죄 수라 오류가 난 것이라고 생각되었다. 따라서 다음 주제로 단순 범죄발생수가 아닌 (범죄 발생 건수/ 범죄 해결 건수) 로 치안에 대한 일정한 치안율을 만들어서 비교할 것이다.

### (Ⅲ) 서울시 5대 범죄 발생현황의 발생 건수/검거 건수와 신생아수의 비교

바로 전에 조사한 것의 데이터의 처리과정을 조금 수정하여 한 지역에서 발생한 발생 건수/검거 건수로 비교를 하여, 범죄가 많이 발생하여도 치안이 안정되면 신생아수가 높은지에 대해서 비교해 볼 것이다.

예상> 발생 건수가 많아도 검거 건수 비율이 높으면 아이를 키우기 좋은 환경이므로 신생아수가 많을 것이다.

#### ▶ 해석

```
matplotlib.rcParams['font.family'] = 'Malgun Gothic'
matplotlib.rcParams['axes.unicode_minus'] = False

b1 = plt.bar(index + 2 * bar_width, prime_rate[0], 5 * bar_width, color = 'red', label="소계")
b2 = plt.bar(index, prime_rate[1], bar_width, color = 'orange', label="살인")
b3 = plt.bar(index + bar_width, prime_rate[2], bar_width, color='yellow', label="강도")
b4 = plt.bar(index + 2 * bar_width, prime_rate[3], bar_width, color = 'pink', label="강간+강제추행")
b5 = plt.bar(index + 3 * bar_width, prime_rate[4], bar_width, color = 'wheat', label="절도")
b5 = plt.bar(index + 4 * bar_width, prime_rate[5], bar_width, color = 'lightblue', label="폭력")

plt.xticks(np.arange(bar_width, 25 * bar_width, 1), state)
```

```

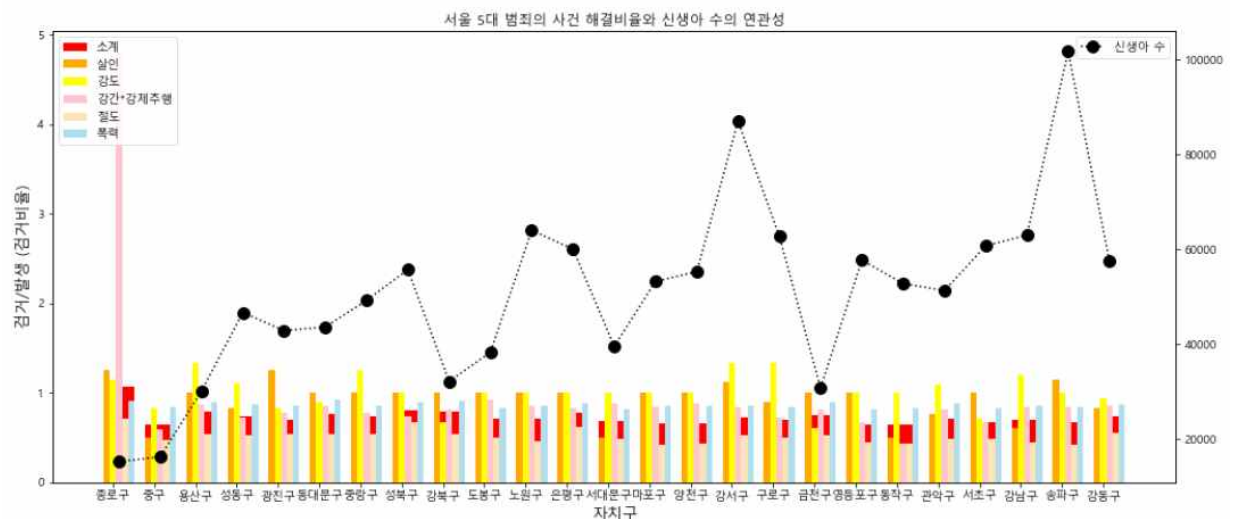
plt.xlabel("자치구", size = 13)
plt.ylabel("검거/발생 (검거비율)", size = 13)

plt.legend(loc = 'upper left')
ax2 = ax.twinx()

ax2.plot(index + 2 * bar_width, age0_4[1:], ls = ":", color = 'black',
        marker = 'o', markersize = 10, label = "신생아 수")
ax2.legend(loc = 'upper right')

plt.title("서울 5대 범죄의 사건 해결비율과 신생아 수의 연관성")
plt.savefig("서울 5대 범죄와 신생아 수의 연관성.jpg")
plt.show;

```



- 데이터를 분석하고 해석하는 과정에서 이 데이터 셋에 대해서 한 가지를 깨닫게 되었다. 발생 건수는 단순히 그 지역에서 범죄가 발생한 것이지만, 검거 건수는 그 지역에서 발생한 사건에 대한 검거가 아닌 다른 지역에서 도망쳐온 범죄자를 검거한 횟수도 포함되니까 종로구처럼 발생대비 검거율이 무려 500%에 해당하는 결과가 나온 것이다. 따라서 이런 식으로의 데이터 처리는 의미가 없으므로 다음 주제로 전체 인구수/발생 건수로 비교할 것이다.

#### (IV) 자치구 전체 인구수 대비 범죄 건수와 신생아 수의 상관관계

지금까지 조사한 데이터 셋과 다르게 이것은 정확히 그 지역의 인구수를 반영하고, 그에 대한 범죄건수를 보여주기 때문에 범죄건수가 낮으면 신생아수가 높을 것이다.

예상> 인구수대비 범죄건수가 낮으면 신생아수가 높을 것이다.



## ▶ 자치구별 총 인구수 쿼리

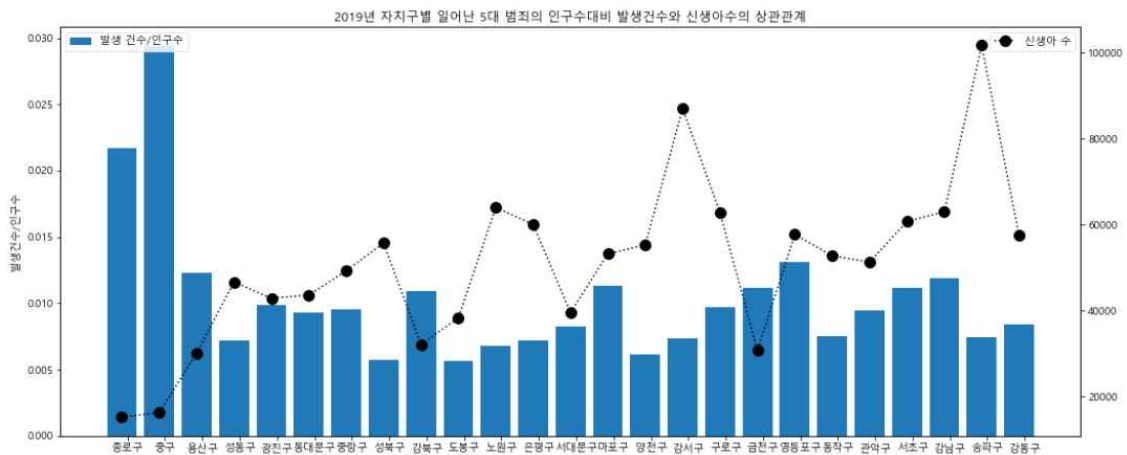
<pre>for row in data2:     sum = 0     if (row[2] == '계'):         for i in range(4):             sum+=row[i+3]         sivil.append(sum)         state.append(row[1])</pre>	<pre>df = pd.DataFrame(sivil) df = df.transpose() df.columns = state df</pre>
--	---

	종로구	중구	용산구	성동구	광진구	동대문구	중랑구	성북구	강북구	도봉구	...	강서구	구로구	금천구	영등포구	동작구	관악구	서초구	강남구
0	177097	147302	269073	347538	407013	397637	446337	499480	350248	374292	...	671287	486965	277396	444770	453654	561795	497018	614026

1 rows × 25 columns

## ▶ 해석

<pre>fig, ax = plt.subplots(figsize=(17,7)) bar_width = 0.15 index = np.arange(25)  index = np.arange(25) plt.bar(index, prime2, label = "발생 건수/인구수") plt.legend(loc = 'upper left') plt.ylabel("발생건수/인구수")</pre>	<pre>ax2 = ax.twinx() ax2.plot(index, age0_4[1:], ls = ":", color = 'black',          marker = 'o', markersize = 10, label = "신생아 수") ax2.legend(loc = 'upper right') plt.xticks(np.arange(bar_width, 25 + bar_width, 1), state) plt.title("2019년 자치구별 일어난 5대 범죄의 인구수대비 "          + "발생건수와 신생아수의 상관관계") plt.show;</pre>
---	--



- 결국 여러 번의 시행착오 끝에 예상한 결과가 나왔다. 인구수가 타 지역 대비 적고, 발생건수가 높은 종로구와 중구는 신생아수가 매우 낮게 나왔다
- 반대로 인구수가 타 지역대비 많고, 발생건수가 적은 성동구, 성북구, 노원구, 양천구, 강서구, 송파구 같은 지역은 신생아수가 높게 나왔다.

결론 => 인구수대비 5대 범죄 발생건수가 적으면 신생아수가 높게 나온다.

(V) 자치구의 초등학교 수가 많을 때 신생아 수의 상관관계.

초등학교 수가 많다는 것은 일반적으로 아이가 학교에 다닐 선택지가 많고, 교육열이 좋은 지역이므로 신생아 수도 많을 것이다.

예상> 자치구의 초등학교 수가 많으면 신생아 수가 많을 것이다.

▶ 자치구별 초등학교 수의 쿼리

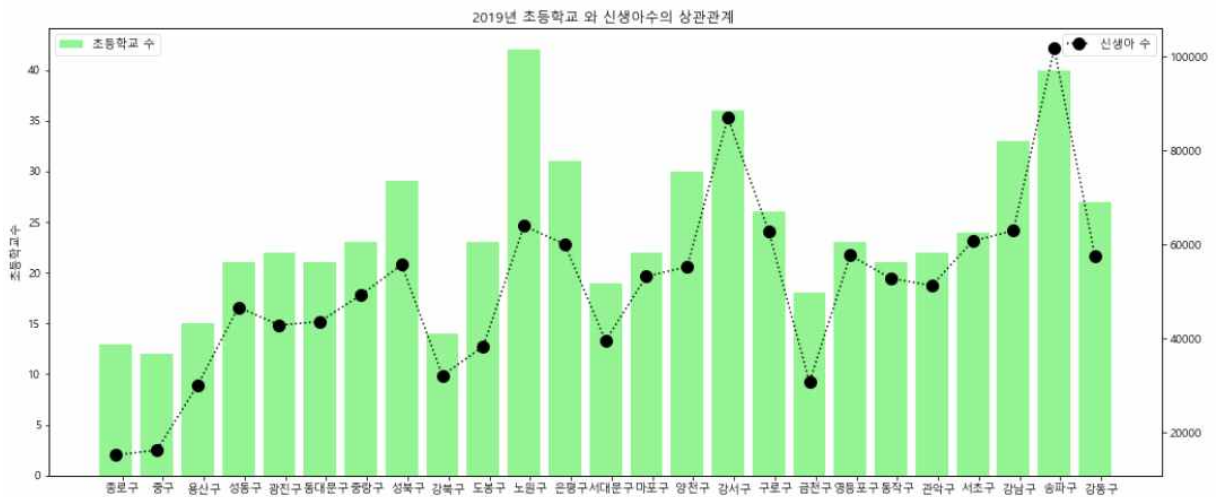
<pre>f = open("서울시+자치구별+초등학교.csv") data = csv.reader(f) elementary = [] header=[]  for row in range(5):     header.append(next(data))</pre>	<pre>for row in data:     elementary.append(row[44]) df = pd.DataFrame(elementary) df = df.transpose() df.columns = state df</pre>
---	--

	종로 구	중 구	용산 구	성동 구	광진 구	동대문 구	중랑 구	성북 구	강북 구	도봉 구	...	강서 구	구로 구	금천 구	영등포 구	동작 구	관악 구	서초 구	강남 구	송파 구	강동 구
0	13	12	15	21	22	21	23	29	14	23	...	36	26	18	23	21	22	24	33	40	27

1 rows × 25 columns

▶ 해석

```
index = np.arange(25)
plt.bar(index, elementary,color = 'lightgreen', label = "초등학교 수")
plt.legend(loc = 'upper left')
plt.ylabel("초등학교 수")
ax2 = ax.twinx()
ax2.plot(index , age0_4[1:], ls = ":", color = 'black' ,
         marker = 'o',markersize = 10,label = "신생아 수")
ax2.legend(loc = 'upper right')
plt.xticks(np.arange(bar_width, 25 + bar_width, 1), state)
plt.title("2019년 초등학교 와 신생아수의 상관관계 ")
plt.show;
```



- 예상대로 자치구의 초등학교 수가 많은 곳은 신생아수도 많았다. 초등학교의 숫자가 많으므로 어떤 초등학교에 보낼까에 대한 선택지가 많고, 초등학교가 많으므로 학원 등 교육시설도 더 많을 것이므로, 새로 이사를 오거나, 임신을 하는데 여러 유리한 점이 있어서 신생아 수가 많은 것 같다.

결론 => 초등학교 수가 많으면 선택의 폭이 넓어지기 때문에 신생아 수가 많아진다.

#### (VI) 자치구의 지역별 특징에 따른 신생아 수의 지역적 차이

도시가 성장함에 따라 도시공동화 현상이 일어나면 서울특별시의 중심부인 중구 쪽의 상주하는 인구가 적어지고, 유동 인구가 많아지기 때문에 서울특별시의 중심부의 신생아 수가 적을 것이다. 또한 요즘 강남 3구가 교육에도 유리하고, 살기 좋다는 주변인들의 인식 때문에 강남 3구도 신생아 수가 많을 것이다.

예상> 서울 중심부는 신생아 수가 적고, 교외로 갈수록 많아질 것이다. 또한 강남 3구의 신생아수는 많을 것으로 예상된다.

#### ▶ 서울특별시 지도

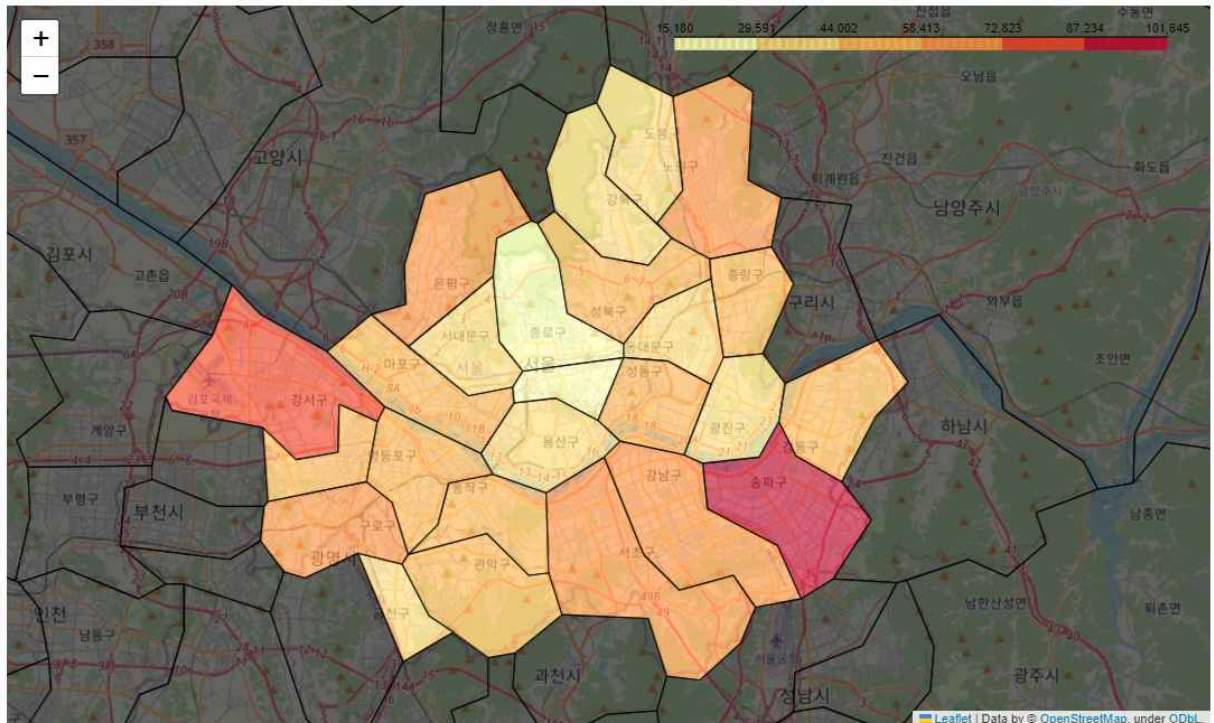
```
import pandas as pd
import numpy as np
import folium
import json

geo_str = json.load(open("skorea_municipalities_geo_simple.json",
                        encoding='utf-8'))
seoul_map = folium.Map(location=[37.5502, 126.982], zoom_start = 10.5)
seoul_map
```



► 해석

<pre>df = pd.DataFrame(state) df['인원수'] = age0_4 df = df.drop(0) df.columns = ['자치구', '인원수'] map_group = df.groupby('자치구')['인원수'].sum()</pre>	<pre>map_group.fillna(     method = 'pad',     inplace = True )</pre>
<pre>seoul_map = folium.Map(location=[37.5502, 126.982], zoom_start = 10.5) folium.Choropleth(     geo_data = geo_str,     data = map_group,     columns = ['인원수', '자치구'],     fill_color = 'YlOrRd',     key_on = 'feature.properties.name' ).add_to(seoul_map)  seoul_map</pre>	



- 예상대로 도심공동화 현상으로 인하여 서울의 중심부분인 종로구와 중구의 신생아 수가 적고, 강남 3구인 서초구, 강남구, 송파구의 신생아가 많았다, 특히하게 강서구의 신생아 수가 매우 많았다. 대체적으로 한강 이남의 지역이 한강 이북의 지역들보다 일반적으로 신생아가 많았다.

결론 => 도심 공동화 현상이 진행되면 도심의 신생아가 줄어든다. 또한 한강 이남지역은 한강 이북에 비하여 신생아 수가 많다.

## (Ⅶ) 자치구별 1인당 지역 내 총생산에 따른 신생아 수의 변화

대체로 결혼을 하기 전에 안정적으로 살 수 있는 일정 소득과, 거주지, 먹을거리, 입는 것 같은 의식주를 해결하고 결혼을 하고 싶어 하기 때문에, 결혼을 한 부부들이 많은 지역 즉 신생아 수가 많은 지역은 일반적으로 1인당 지역 내 총생산이 많을 것이다.

예상> 강남 3구는 일반적으로 잘 사는 동네라고 알려져 있기 때문에, 총생산량도 높을 것이고, 신생아 수도 많을 것이다.



## ▶ 자치구별 1인당 지역 내 총생산 쿼리

```
f = open("자치구별+1인당+지역내총생산+및+수준지수(2019).csv")
data = csv.reader(f)

header = []
GDP = []
state = []

for i in range(3):
    header.append(next(data))

for row in data:
    GDP.append(int(row[5].replace(",","")))
    state.append(row[1])
df = pd.DataFrame(GDP)
df = df.transpose()
df.columns = state
```

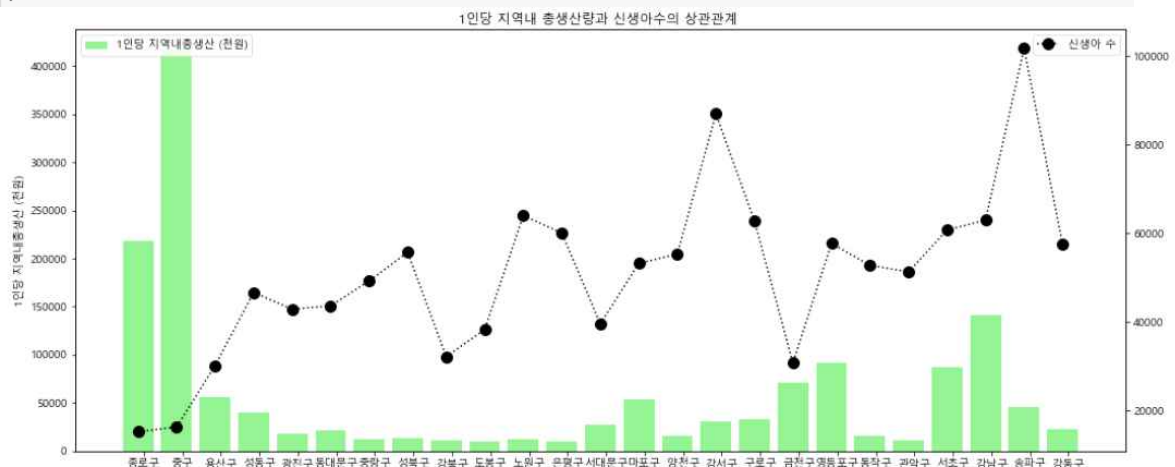
	종로구	중구	용산구	성동구	광진구	동대문구	중랑구	성북구	강북구	도봉구	...	강서구	구로구	금천구	영등포구	동작구	관악구	서초구	강남구	송파구	강동구
0	218704	417483	56571	40100	18211	21033	12174	13894	10743	10438	...	30176	32692	70750	91575	16197	11663	87695	141362	46033	22807

1 rows × 25 columns

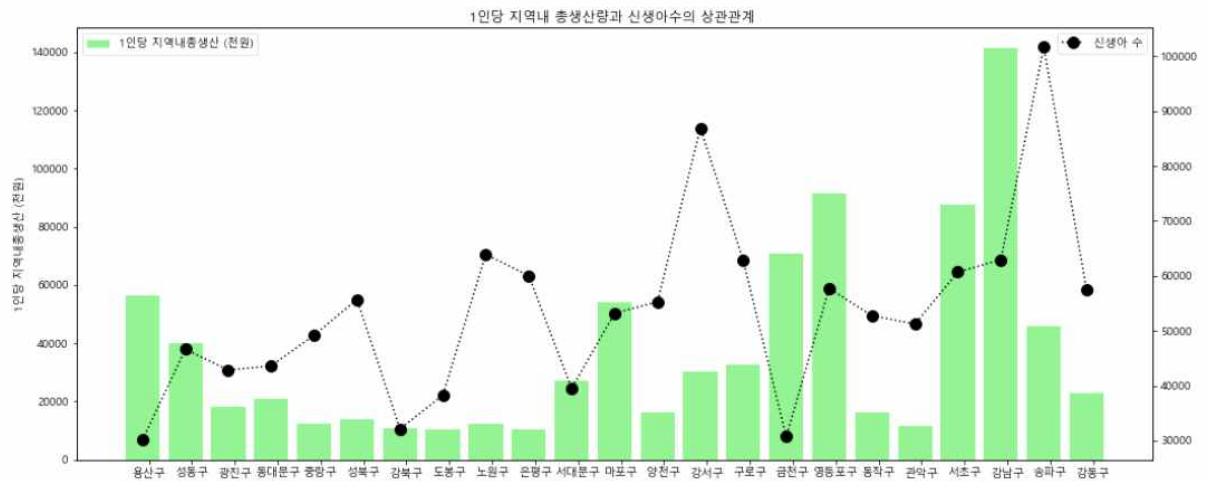
## ▶ 해석

```
fig, ax = plt.subplots(figsize=(17,7))
bar_width = 0.15
index = np.arange(25)

index = np.arange(25)
plt.bar(index, GDP,color = 'lightgreen', label = "1인당 지역내총생산 (천원)")
plt.legend(loc = 'upper left')
plt.ylabel("1인당 지역내총생산 (천원)")
ax2 = ax.twinx()
ax2.plot(index , age0_4[1:], ls = ":", color = 'black' ,
        marker = 'o',markerSize = 10,label = "신생아 수")
ax2.legend(loc = 'upper right')
plt.xticks(np.arange(bar_width, 25 + bar_width, 1), state)
plt.title("1인당 지역내 총생산량과 신생아수의 상관관계 ")
plt.show;
```



- 이때 종로구와 중구는 서울의 중심부라 너무 비싼 지역이고, 일반적으로 주거하지 않는 곳이며 회사원들이 많이 출근하는 상주하는 인원이 적고 이동 인원이 많은 지역이기 때문에, 너무 극단적이므로 중구와 종로구를 제외하고 다시 그래프를 그려본다.



- 예상 외로 1인당 지역 내 총생산량과 신생아 수와 크게 연관이 있어 보이지 않는다. 특히, 중구와 종로구와 용산구, 금천구는 각각 상업지대와 디지털단지가 있어서, GDP와 신생아 수가 정비례하지 않는 것이 당연하지만, 강남 3구를 제외한 다른 지역은 GDP 와 신생아수가 큰 관련이 있어보이지는 않는다.

결론 => 1인당 지역 내 총생산량과 신생아수는 크게 관련이 없다.

## (VIII) 상급학교 진학률에 따른, 신생아수 차이

상급학교 진학률이 높으면 공부에 관심이 있는 학부모들이 이사를 올 것이다. 또한 공부를 포기하고 불량한 행동을 하는 양아치들이 상대적으로 적다는 의미이므로 동네의 치안이 좋아질 것이고, 따라서 당연히 삶의 질이 좋은 동네이므로 신생아 수도 많아질 것이다.

예상> 상급학교 진학률이 높은 지역은, 신생아수도 많을 것이다.

### ▶ 자치구별 상급학교 진학률 쿼리

```
header = []
pre_edu = []
state = []

for i in range(3):
    header.append(next(data))
```

```
for row in data:
    if row[3] == '소계':
        pre_edu.append(float(row[6].replace(",","")))
        state.append(row[1])

df = pd.DataFrame(pre_edu)
df = df.transpose()
df.columns = state
df
```

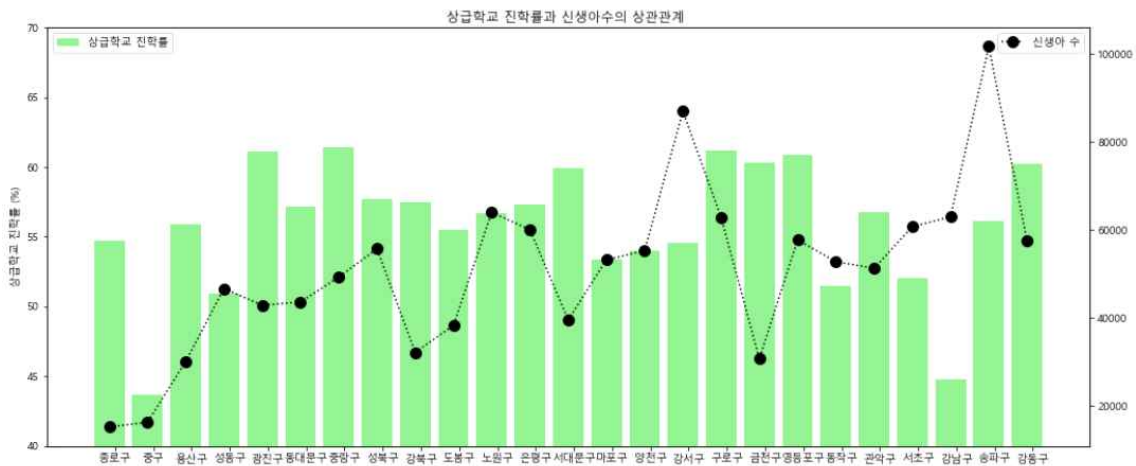
	종로구	중구	용산구	성동구	광진구	동대문구	중랑구	성북구	강북구	도봉구	...	강서구	구로구	금천구	영등포구	동작구	관악구	서초구	강남구	송파구	강동구
0	54.7	43.7	55.9	50.9	61.1	57.2	61.4	57.7	57.5	55.5	...	54.6	61.2	60.3	60.9	51.5	56.8	52.0	44.8	56.1	60.2

1 rows × 25 columns

## ► 해석

```
fig, ax = plt.subplots(figsize=(17,7))
bar_width = 0.15
index = np.arange(25)

plt.bar(index, pre_edu,color = 'lightgreen', label = "상급학교 진학률")
plt.ylim(40,70)
plt.legend(loc = 'upper left')
plt.ylabel("상급학교 진학률 (%)")
ax2 = ax.twinx()
ax2.plot(index , age0_4[1:], ls = ":", color = 'black' ,
         marker = 'o',markersize = 10,label = "신생아 수")
ax2.legend(loc = 'upper right')
plt.xticks(np.arange(bar_width, 25 + bar_width, 1), state)
plt.title("상급학교 진학률과 신생아수의 상관관계 ")
plt.show;
```



- 의외로 전혀 상관이 없었다. 오히려 종로구, 중구, 강남구 같은 경우는 되게 잘사는 동네임에도 불구하고 상급학교 진학률이 매우 저조하였고, 신생아 수와 상급학교 진학률은 전혀 상관이 없는 것처럼 보인다.

결론 => 상급학교 진학률과 신생아수는 전혀 관련이 없다.



## 5. 결론

지금까지 서울 자치구별 신생아 수, 평균 연령대와 초등 취학 연령 인원수에 대하여 5대 범죄, 자치구의 초등학교 개수, 자치구의 지역적인 특징, 지역 내 총생산량을 비교해 보았다. 이에 대한 결론으로 다음과 같은 결론들이 나왔다.

### 결론

- => 신생아 수와 유치원, 초등생 취학연령인원수는 비례하고, 평균 연령은 반비례한다.
- => 인구수대비 5대 범죄 발생건수가 적으면 신생아수가 높게 나온다.
- => 초등학교 수가 많으면 선택의 폭이 넓어지기 때문에 신생아 수도 많아진다.
- => 도심 공동화 현상이 진행되면 도시의 신생아가 줄어든다. 또한 한강 이남지역은 한강 이북에 비하여 신생아 수가 많다.
- => 1인당 지역 내 총생산량과 신생아수는 크게 관련이 없다.
- => 상급학교 진학률과 신생아수는 전혀 관련이 없다.

정리하자면, 신생아수와 유치원 초등생 취학 연령 인원수를 늘리고, 평균 연령을 낮추기 위해서는 인구수 대비 5대 범죄 발생 건수가 적으면 안정적이기 때문에 유리하고, 초등학교 수가 많아지면 교육적인 선택의 폭이 넓어지기 때문에 유리하고, 한강 이남지역은 한강 이북에 비하여 신생아 수가 많다.

하지만, 도심 공동화 현상이 진행되면 도시의 유동 인구가 증가하고, 상주인구가 줄어들기 때문에 전체적인 인구수가 줄어들어서 불리하고, 1인당 지역 내 총생산량은 신생아 수와 크게 관련이 없으며, 상급학교 진학률은 신생아 수와 전혀 관련이 없다.

따라서, 날이 가면 갈수록 증가하고 있는 평균연령을 낮추고, 신생아를 수를 늘리기 위해서 정부가 해야 할 정책은 초등학교수를 늘리고, 사법제도와 경찰 제도를 공정하게 하여 신혼부부나 연인들에게 하여금 아이를 낳아도 되는 안정적인 환경을 만들어주면 그나마 완화가 될 것이다.

## 6. 소감

지금까지 한국공학대의 컴퓨터공학과라는 과에 와서 단순히 조건문, 반복문, 함수 같은 것들만 배우면서 대체 이런 것들을 어디다 써먹을까 고민하면서, 흥미가 없었습니다.

하지만 이런 데이터처리 관련된 코딩과 패키지, 보고서를 배우고 작성하면서 생각 외로 코딩이 실생활에 도움이 많이 된다는 것을 느꼈습니다. 또한 이 주제에 대하여 통계를 내고 데이터 분석을 하면서 막연하게 제가 생각하고 예상했던 것과 실제 세상은 꽤 다르다는 것을 느끼게 되었습니다.