

데이터 분석및 시각화

프로젝트

2조 박수인, 김종욱, 김성혜, 서경녹, 권진우

목차

저출산과 고령화, 코로나



1. 개요



2. 팀 구성원 역할



3. 프로젝트 일정



4. 프로젝트 수행 결과



5. 팀원 자체 평가



1. 개요

사회적 이슈 시각화-저출산과 고령화, 코로나

1. 노인 인구 분포를 지도 형태로 시각화 (CSV)
2. 초등학생 학년별 학생 수 비교 및 시각화 (JSON)
3. 저출산과 고령화 시각화 (XML)
4. 코로나 발생 시기와 현재의 마스크 가격 변동 시각화 (HTML)

2. 팀 구성원 구성 및 역할

이름

박수인(팀장)

김종욱

김성혜

서경녹

권진우

역할

요구사항 분석, CSV 데이터 크롤링 및 시각화, 프로그램 통합, PPT 제작

요구사항 분석, XML 및 CSV 데이터 크롤링 및 시각화, PPT 제작 및 발표

요구사항 분석, XML 및 CSV 데이터 크롤링 및 시각화, PPT 제작 및 통합

요구사항 분석, HTML 데이터 크롤링 및 시각화, PPT 제작 및 통합

요구사항 분석, 데이터 크롤링 및 시각화





3. 프로젝트 일정

(230206-230210)

230206

요구사항 분석, 계획 및 방향 수립, 역할 분담, 데이터
크롤링

230207-230209

데이터 크롤링, 전처리, 시각화, DB or EXCEL 저장

230209

프로그램 통합, PPT 제작, 발표 준비

230210

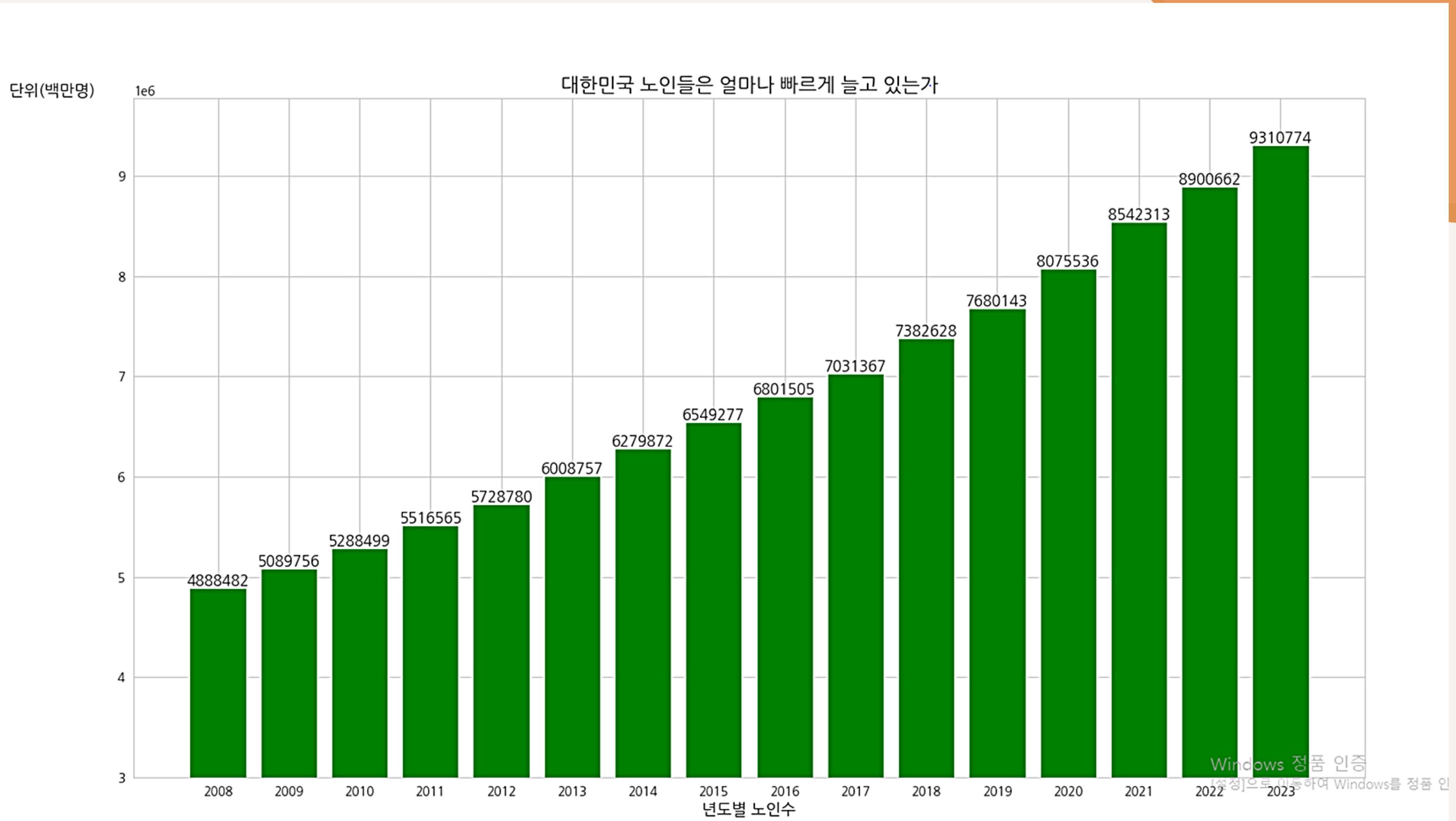
발표

크롤링과 시각화



4. 프로젝트 수행 결과

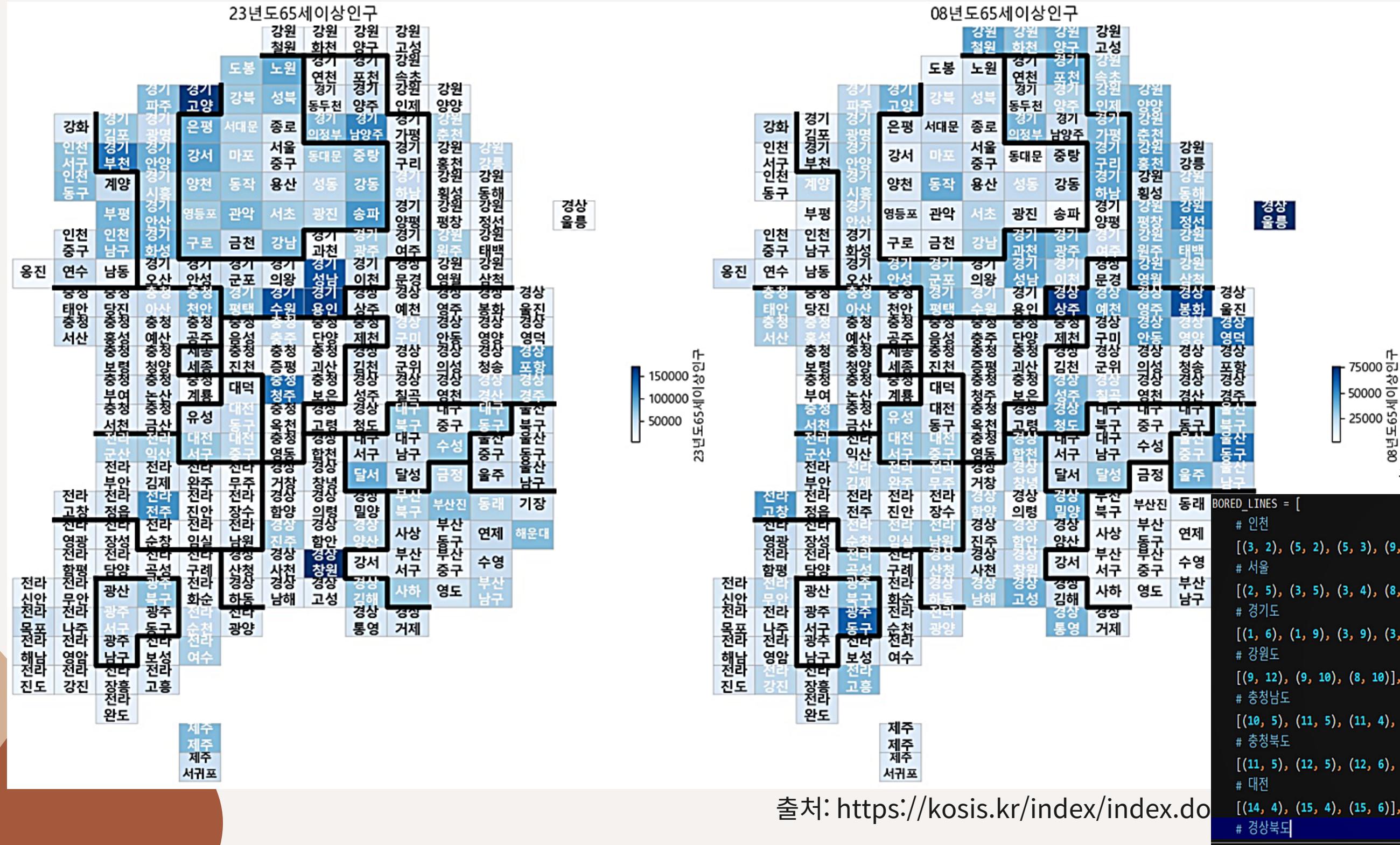
심화되는 고령화 괜찮을까?



출처: <https://kosis.kr/index/index.do>

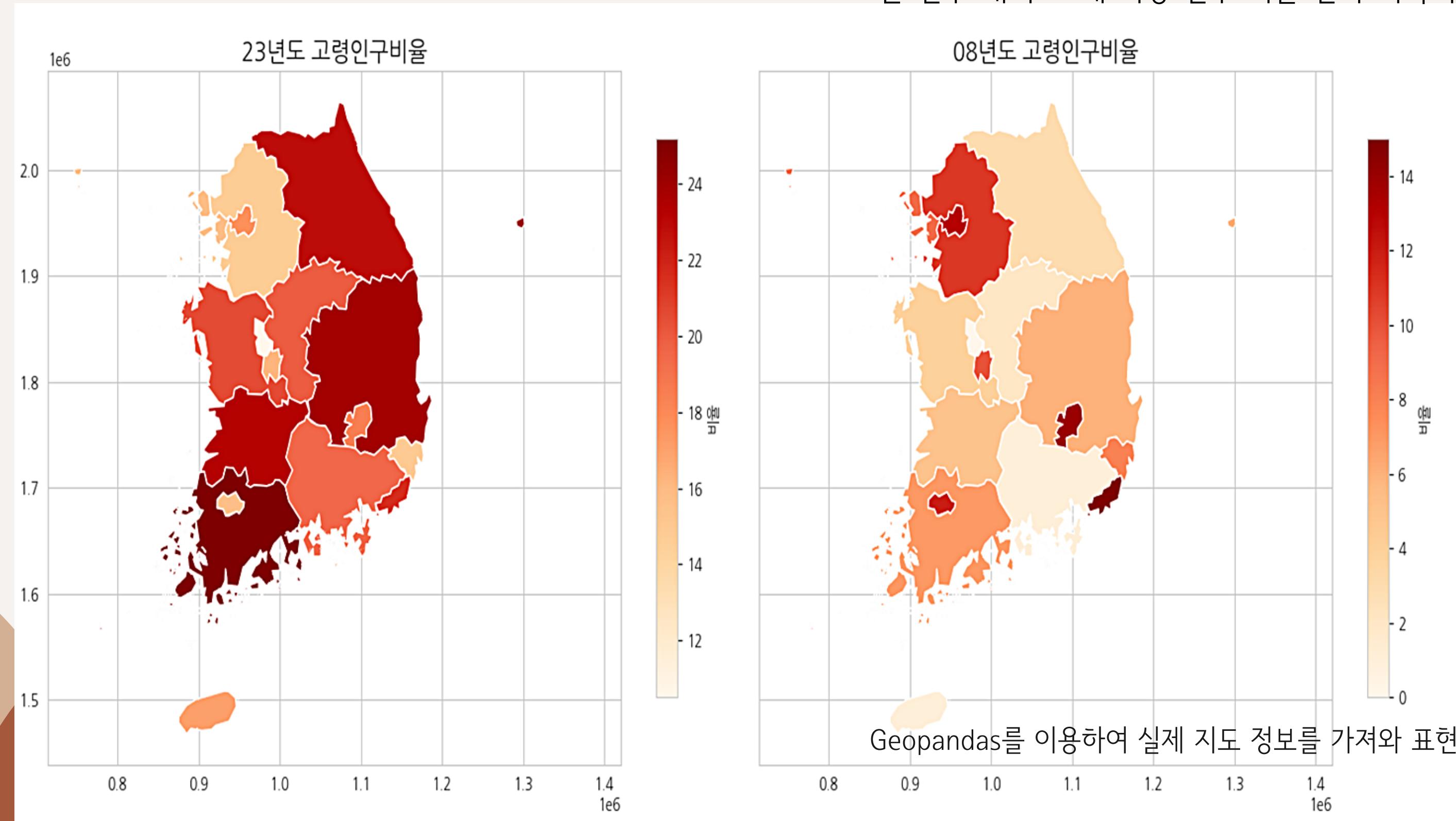
대한민국 노인들은 지금

시군구를 기점으로 각 지역에 거주하는 노인수를 시각화



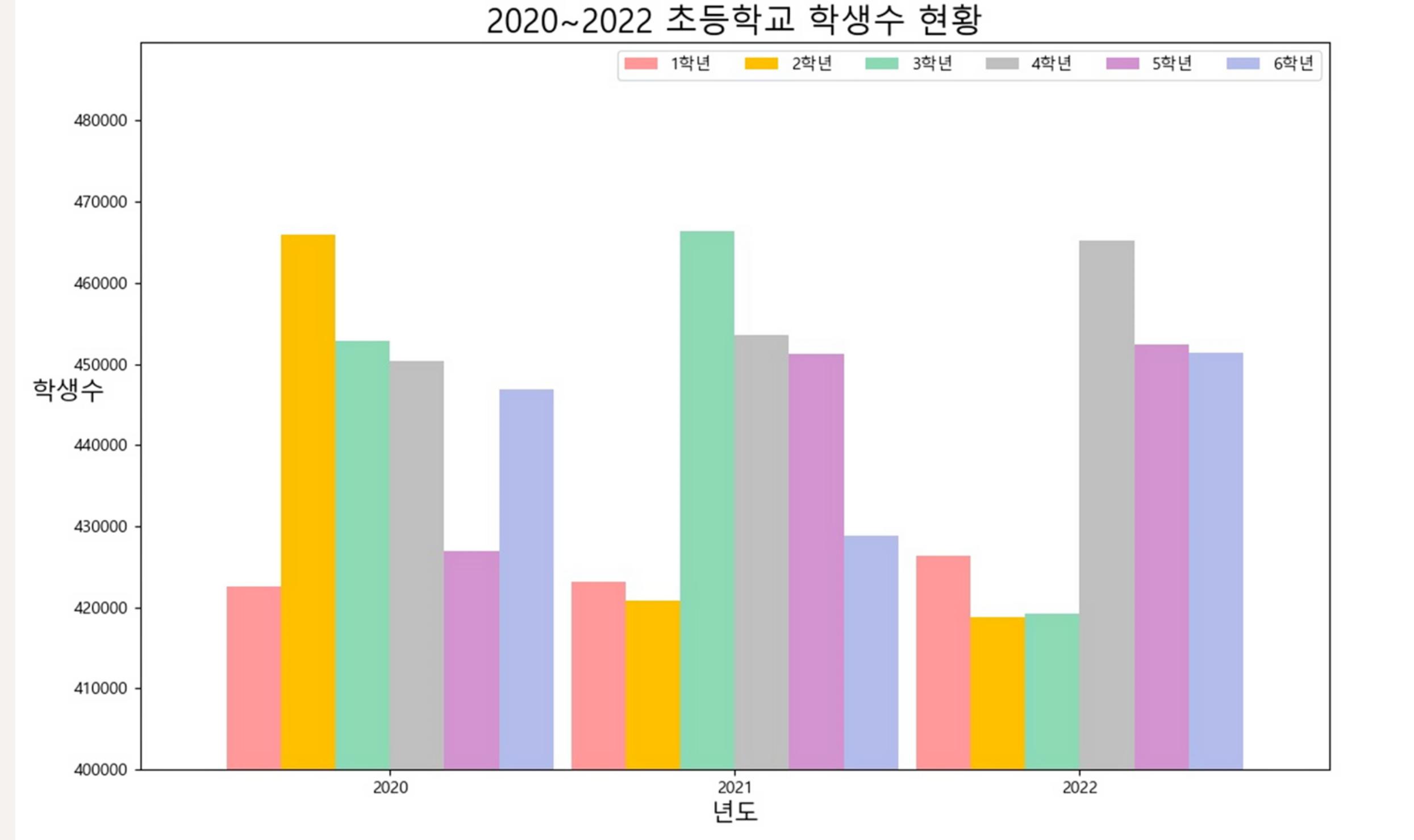
그래서 얼마나 늘어난 걸까?

모든 인구 대비 65세 이상 인구 비율 변화 시각화



2020~2022 초등학교 학생수 현황

공공데이터를 활용한 초등학교
학생수 감소 현황 시각화



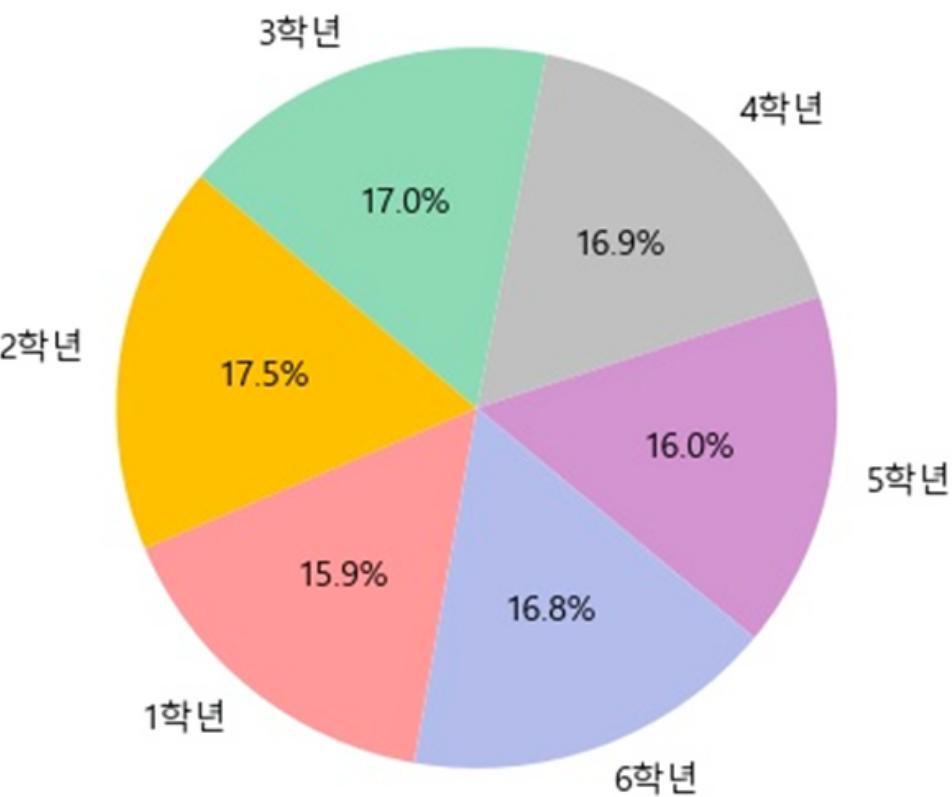
자료: 공공 데이터/ 학교 알리미/ 성별 학생수/ 2020, 2021, 2022년도

출처: data.go.kr, <https://www.schoolinfo.go.kr>

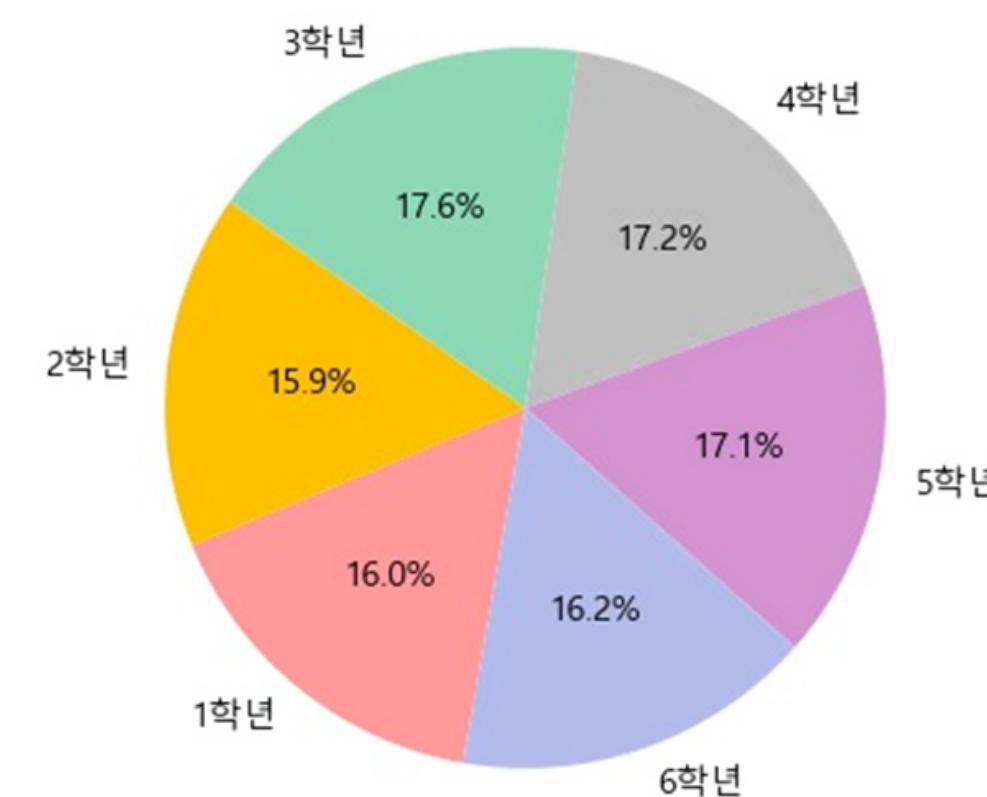
2020~2022 초등학교 학생수 현황

2020~2022년 초등학교 학생수 현황

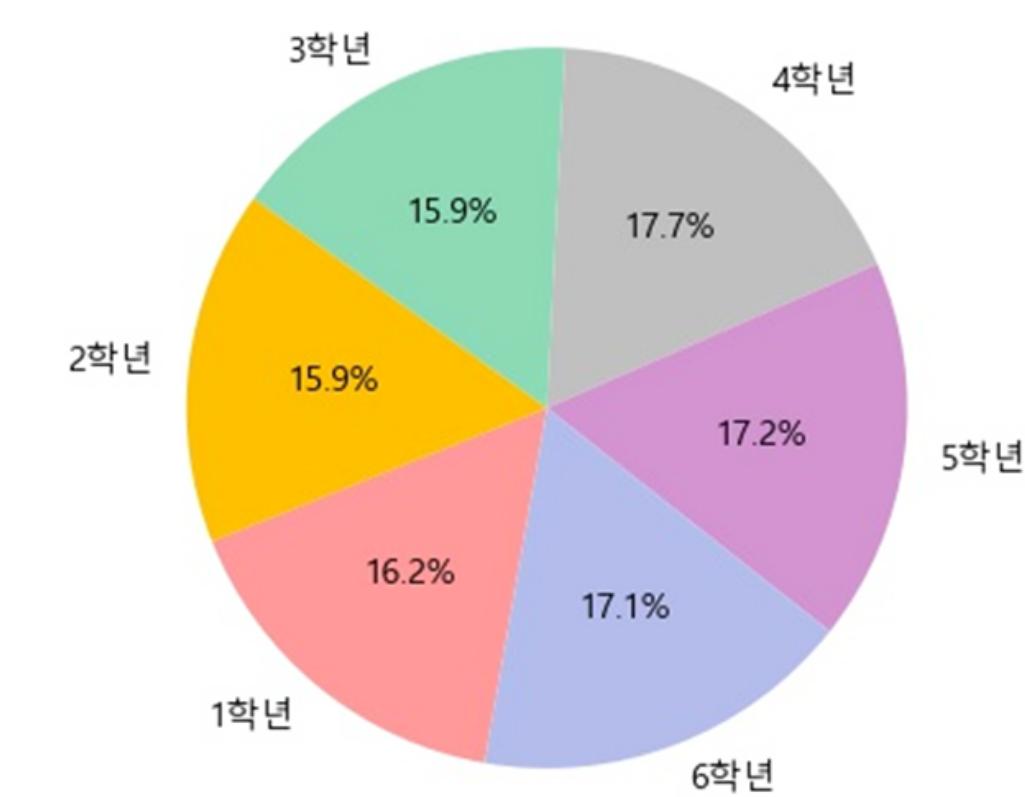
2020년



2021년



2022년



■ 1학년 ■ 2학년 ■ 3학년 ■ 4학년 ■ 5학년 ■ 6학년

자료: 공공 데이터/ 학교 알리미/ 성별 학생수/ 2020, 2021, 2022년도
출처: data.go.kr, <https://www.schoolinfo.go.kr>

저출산과 고령화 현황을 데이터 시각화를 통해 확인

자료

- 1.XML-행정안전부 법정동별 성별 출생 등록자 수 (2022.12-2023.1)
출처: data.go.kr
- 2.CSV-출생아 수 합계 출산율 자연 증가 (1970, 2013, 2014, 2015)
출처: <https://kosis.kr/index/index.do>

목표

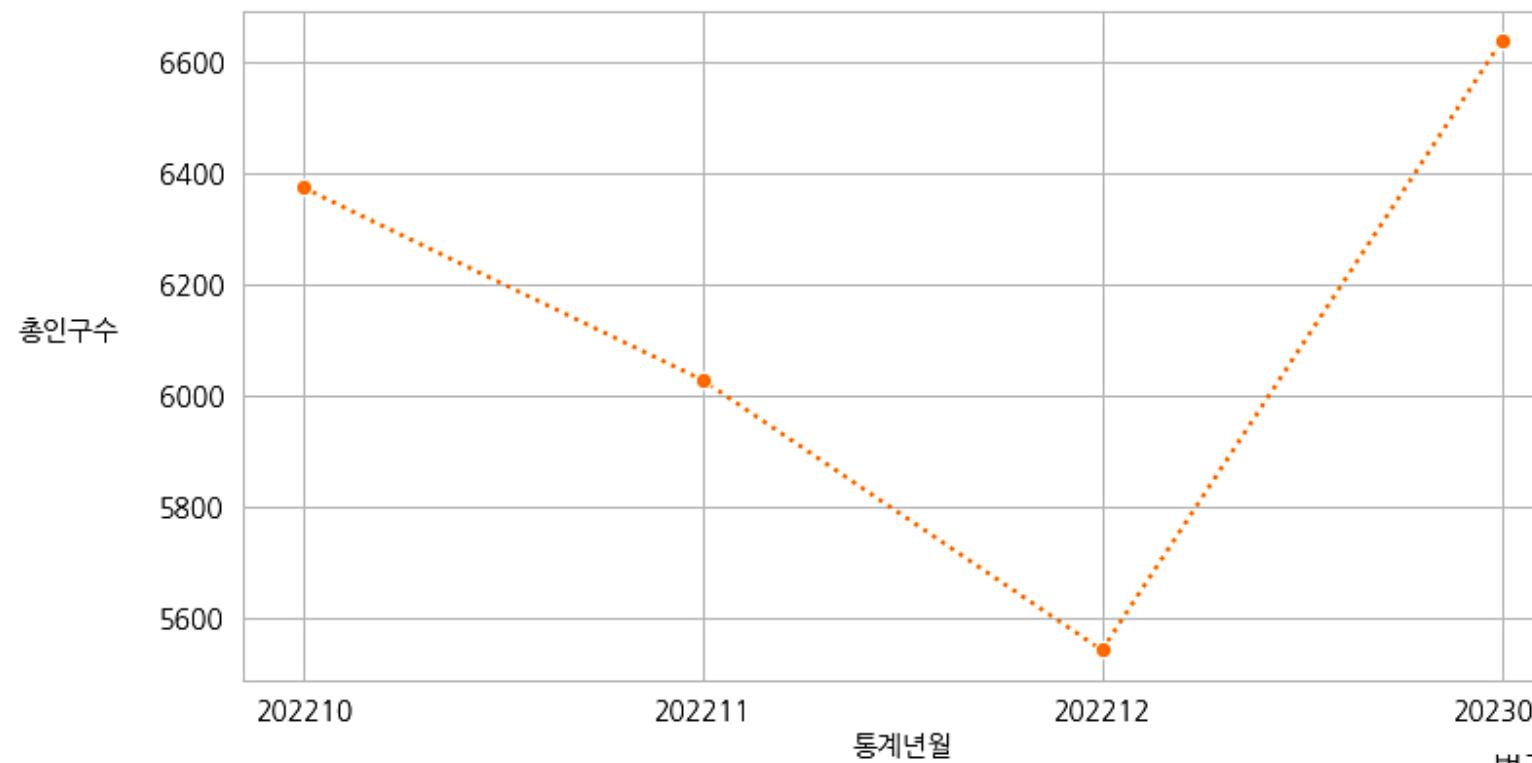
- 1.현재 어른, 초등학교 신입생, 신생아의 출생년도 출생아수(명) 비교를 통한 고령화 가시화
- 2.다양한 시각화 도구 활용

과정

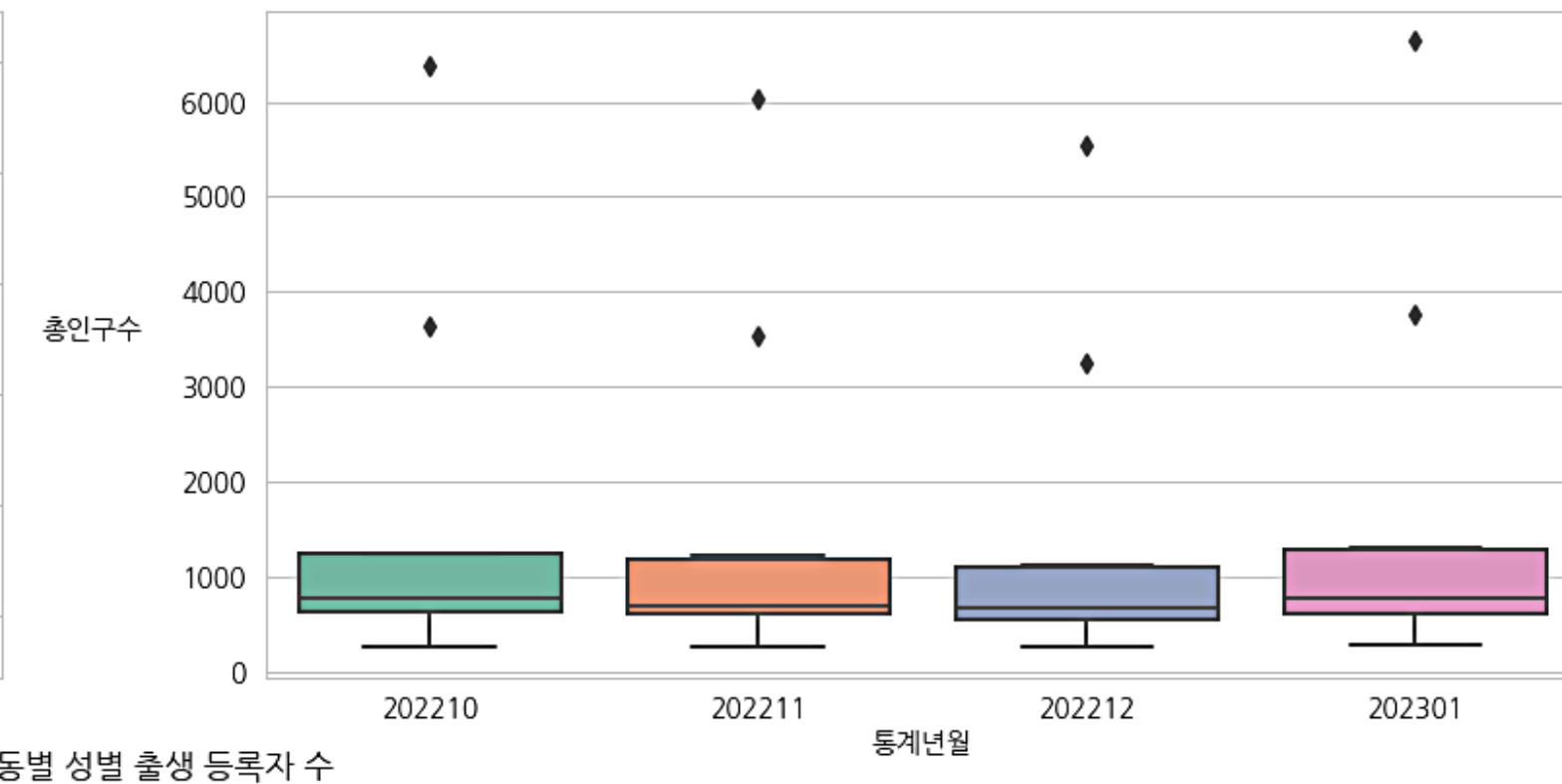
- 1.seaborn을 활용한 데이터 시각화
- 2.GridSpec을 활용하여 1행 2열 크기에 1개 그래프 시각화

법정동별 성별 출생 등록자 수 (2022.10-2023.1)

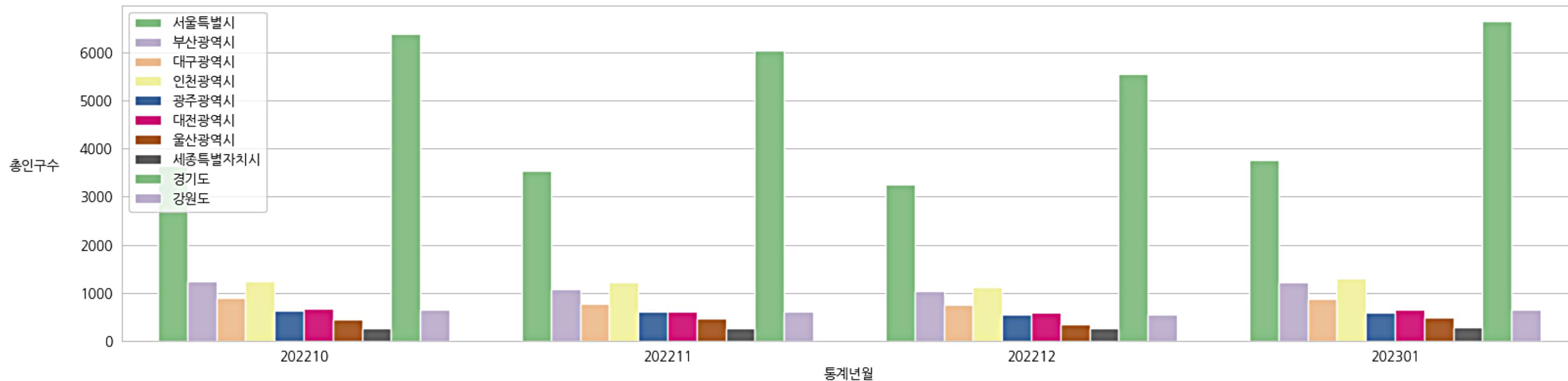
경기도 월별 출생 등록자 수



법정동별 성별 출생 등록자 수



법정동별 성별 출생 등록자 수



lineplot

경기도 지역만 데이터 시각화

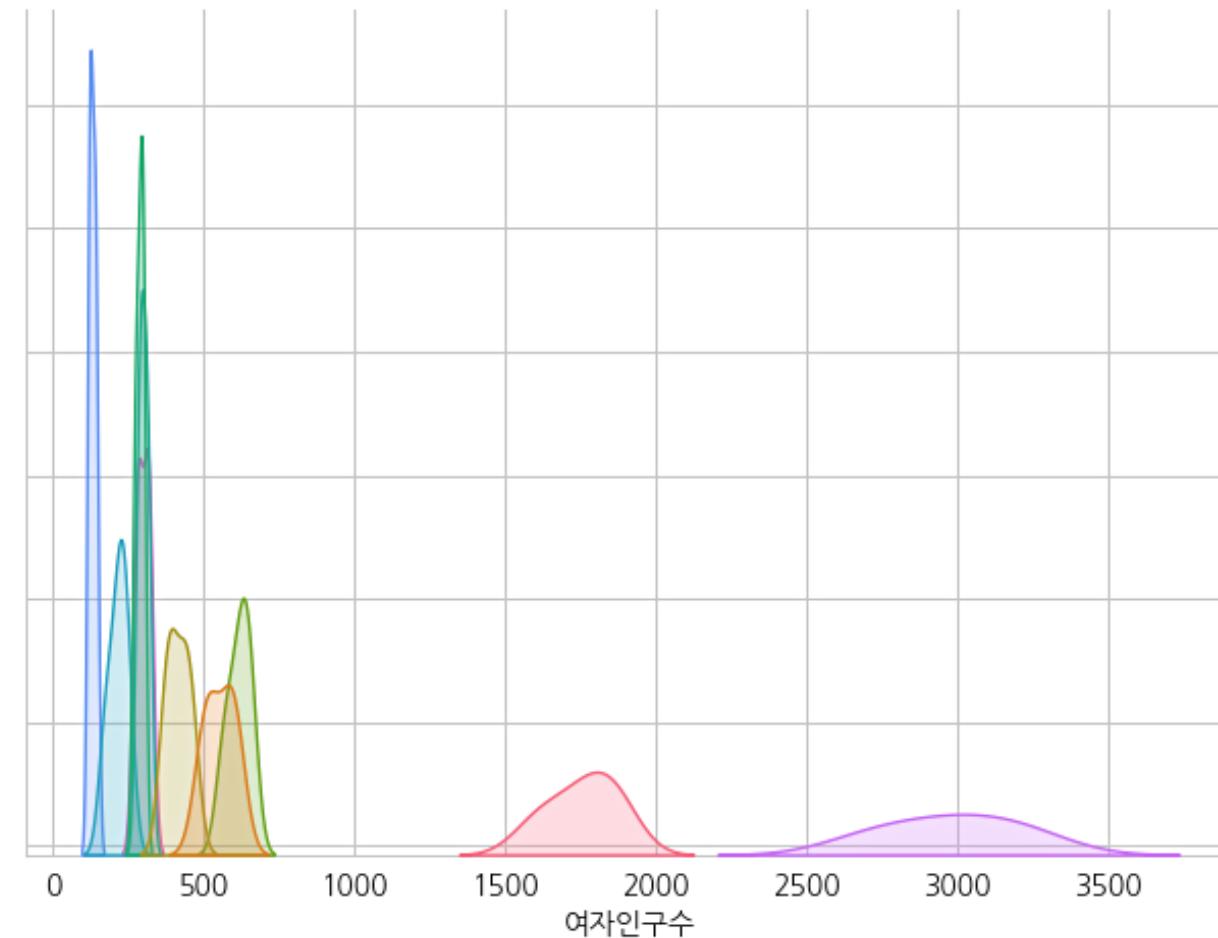
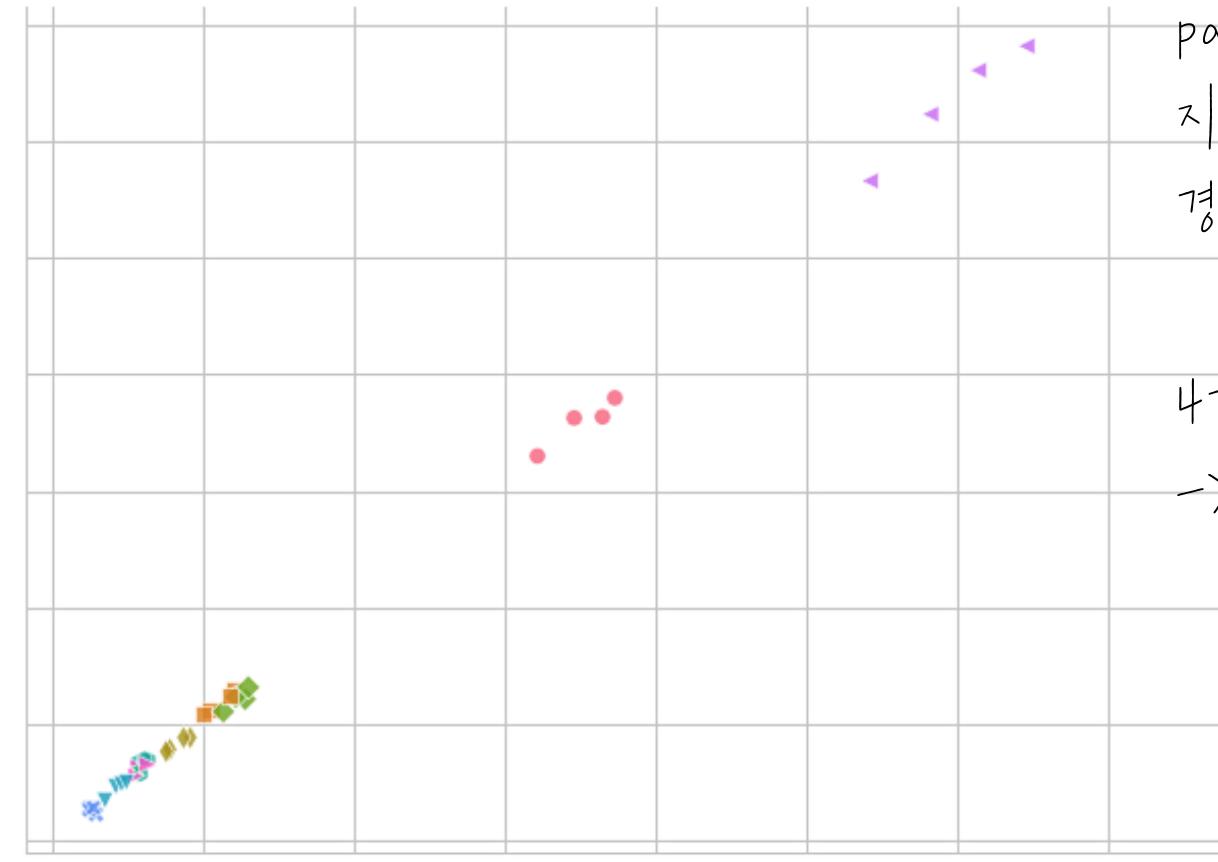
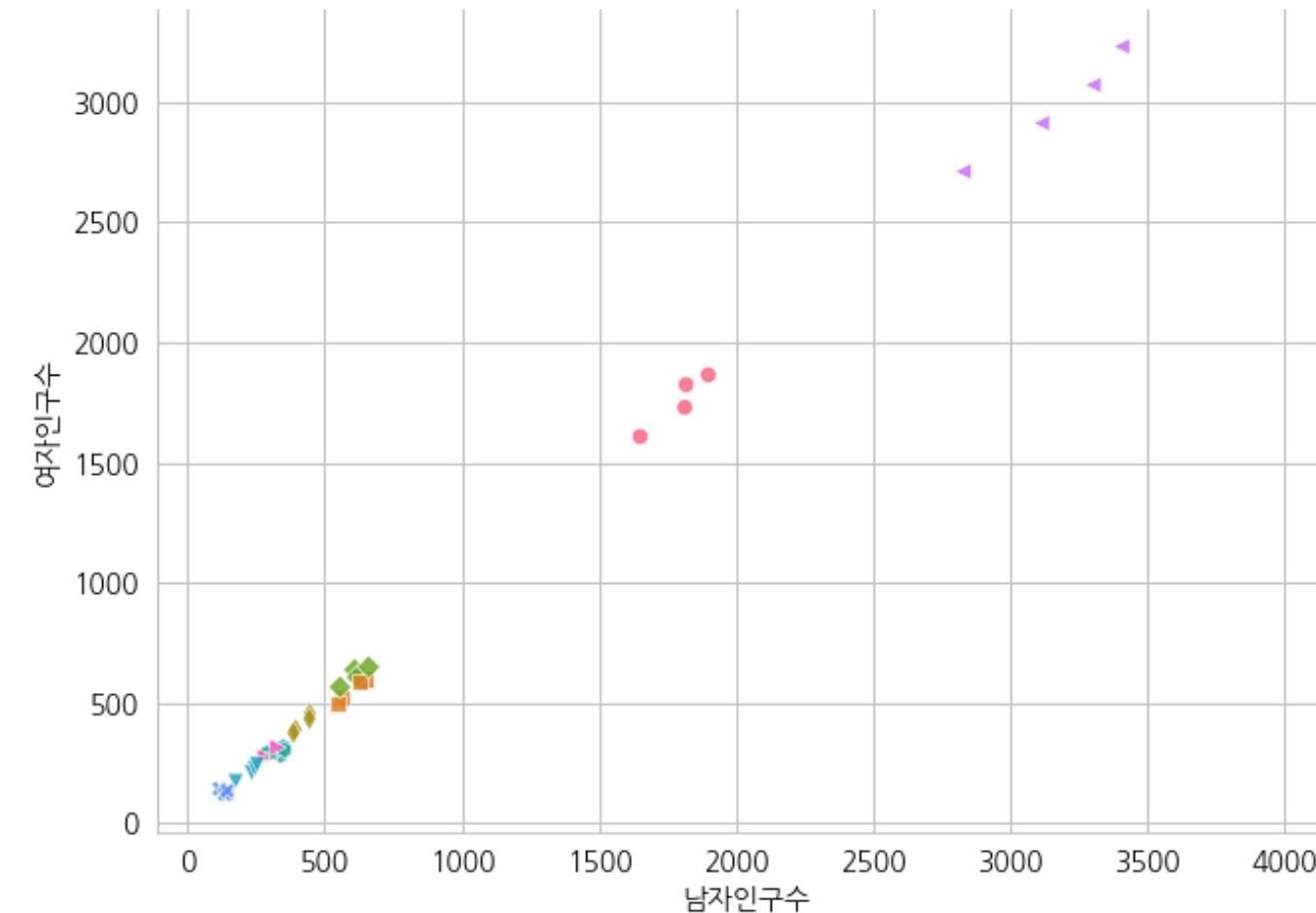
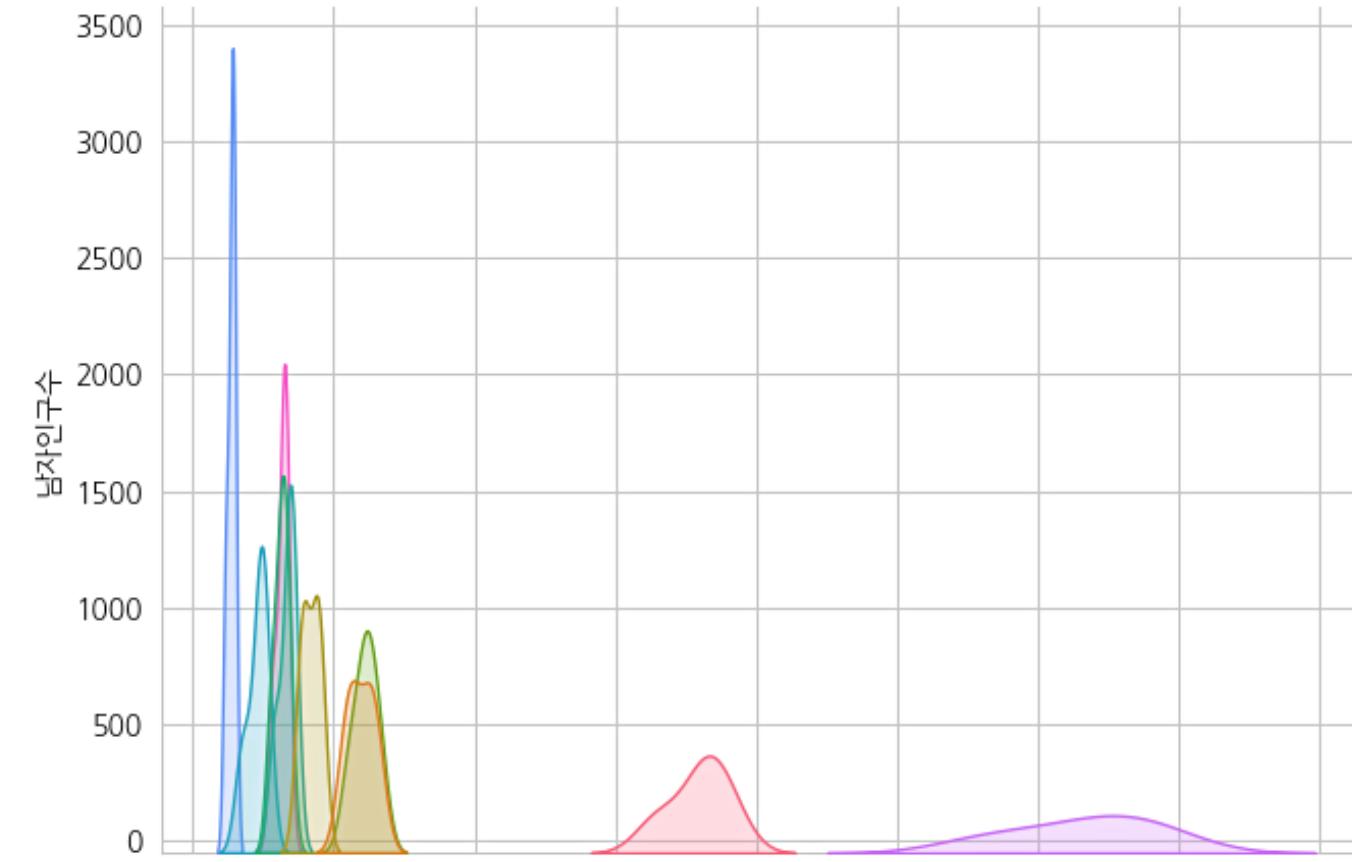
boxplot

모든 지역 데이터 시각화

barplot

지역별 월별 출생 등록자 수 시각화

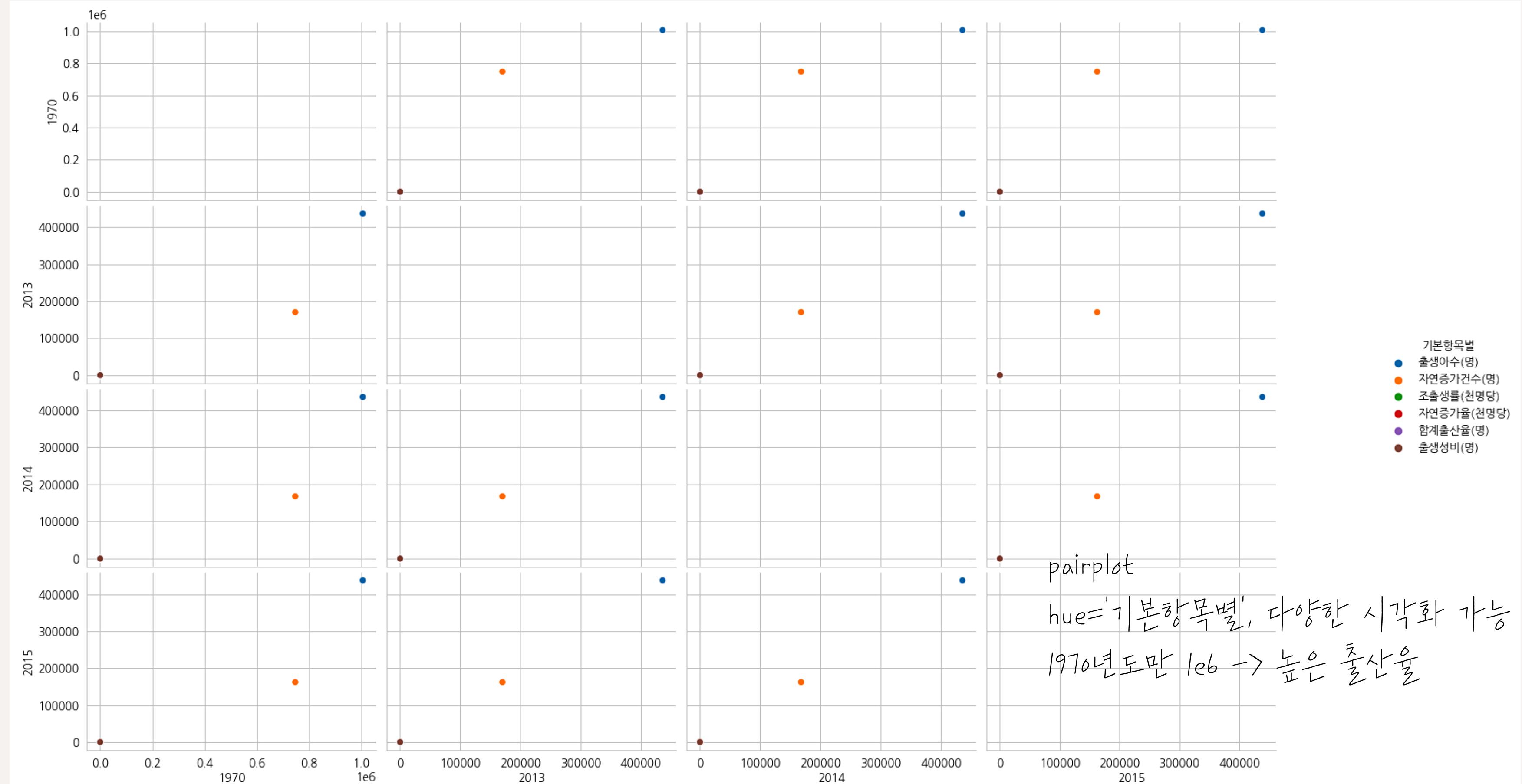
법정동별 성별 출생 등록자 수 (2022.10-2023.1)



시군구명

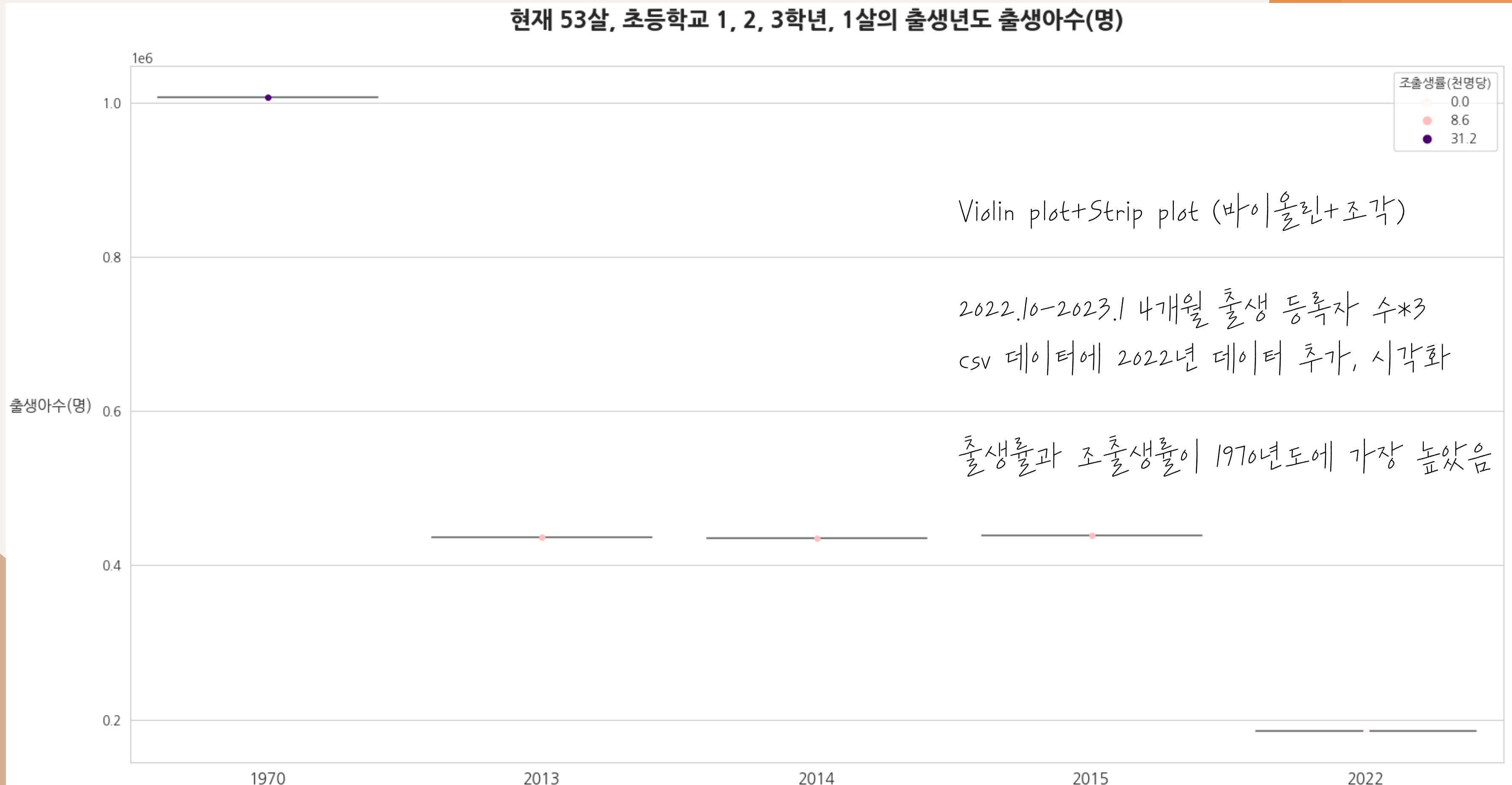
- 서울특별시
- 부산광역시
- ◆ 대구광역시
- ▲ 인천광역시
- ★ 광주광역시
- 대전광역시
- ▼ 울산광역시
- ✖ 세종특별자치시
- ▶ 경기도
- ▶ 강원도

출생아 수 합계 출산율 자연 증가 (1970, 2013, 2014, 2015)



출생아 수 합계 출산율 자연 증가 (1970, 2013, 2014, 2015)

현재 53살, 초등학교 1, 2, 3학년, 1살의 출생년도 출생아수(명)

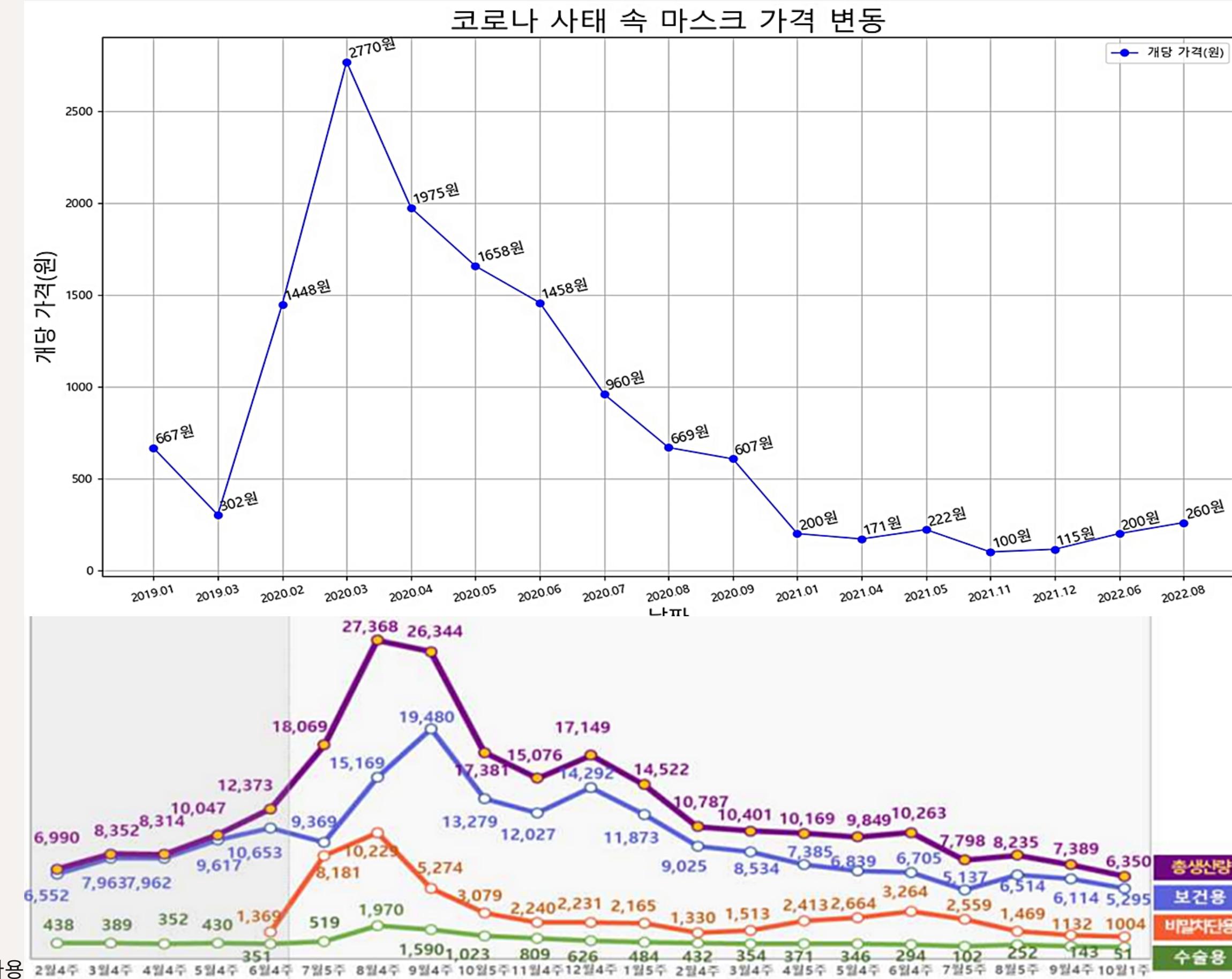


확진자 수 변동에 따른 마스크 가격 변동



출처: <https://www.ppomppu.co.kr-'kf94' 검색-'1p-18p' 사용>

출처: <http://m.ecomedia.co.kr/news/newsview.php?ncode=1065589583448217>



프로젝트 시연

5. 팀원 자체 평가



김종욱

서경녹

작업 중반에 알게 된 것이지만, 너무 좁은 범위의 데이터를 선택했었고 그 결과 변화의 폭이 적었으며, 그래도 최대로 시각화를 하려다 보니 시각화하는 것에 많은 시간이 소모되었습니다.



박수인



김성혜



권진우

다양한 방법으로 데이터를 크롤링하고 이를 파싱하면서 사용자가 필요한 정보만 골라서 추출할 수 있다는 것이 너무 신기하였고, 이를 기반으로 정보를 저장하고 시각화하면서 유의미한 결과물을 사용자에게 전달할 수 있음이 직접적으로 느껴져서 좋았습니다.

다양한 시각화 도구들이 있어서 ppt 자료, 정보 수집, 근거 자료로서 활용도가 높습니다. 얻고자 하는 데이터와 데이터 타입에 따라 필요한 도구가 달라서 데이터를 뽑아오는 연습을 해야겠단 생각을 했고 조원들과 협력이 잘 되었습니다.

기존에 크롬이나 여러 가지 인터넷 서핑 사이트에 html 이 기반이 되어 작동된다는 것을 알고 있었는데, 이번 기회로 html에 대해 여러 가지를 알 수 있는 기회가 됐던 거 같다



감사합니다