



서울시 지역구 치안 분석을 통한 예측모델 학습데이터 구축

불사조

안기용 맹재영 문준영 조래혁 조소진

목차

1. 프로젝트 배경

2. 팀 구성, 역할

3. 수행절차 및 방법



4. DB구축

5. 시각화

6. 결론 및 향후과제

7. 느낀점

프로젝트 배경



'서울시 지역구 치안 분석을 통한 예측모델 학습데이터 구축'





마이너리티 리포트

팀 구성 및 역할

설치, 데이터 전처리, 시각화,
DB구축, Django, 발표

안기용

맹재영

설치, 데이터 수집, 전처리,
시각화, DB구축

주제 기획, 설치, 데이터 수집,
전처리, 시각화

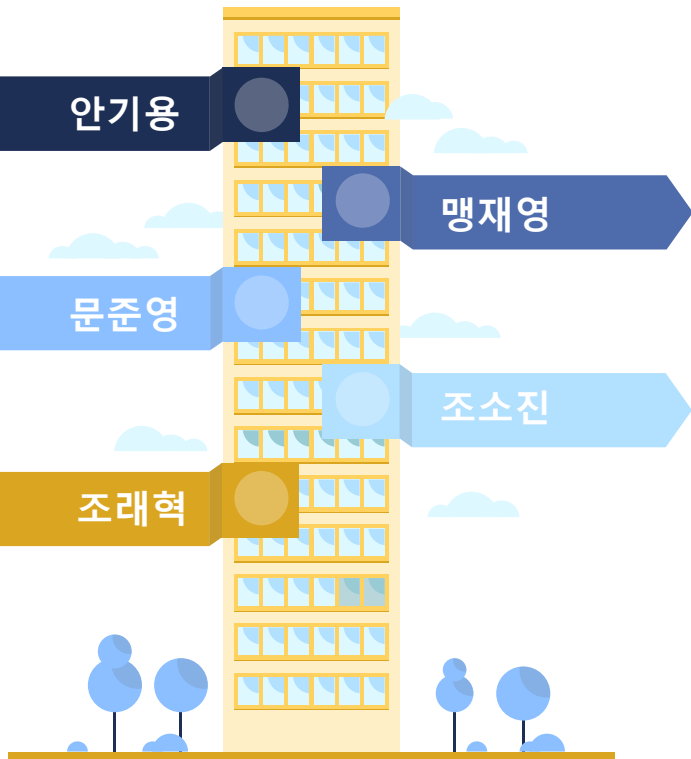
문준영

조소진

설치, 데이터 수집, 전처리,
시각화, ERD

설치, 데이터 수집, 전처리,
시각화, DB구축, ERD

조래혁



수행절차 및 방법

수행절차 및 방법

03-1 work-flow

구분	기간	활동	도구
주제 기획	9/09(목) - 8/10(금)	주제 선정 및 일정 수립	
데이터 수집 및 시스템 구축	9/11(토) - 9/13(월)	데이터 수집	Pandas, NumPy, bs4, 크롤링
	9/14(화)	설치	Spark, Hadoop, Mysql, Jupyter
데이터 처리	9/14(화) - 9/17(금)	데이터 전처리	Pyspark, Pandas, Numpy
데이터 베이스 구축	9/18(토) - 9/20(월)	ERD, DB 저장	Draw.io, Mysql
시각화	9/21(화) - 9/24(금)	데이터 시각화	Seaborn, Matplotlib, Zeppelin, Visual Analysis
장고	9/25(토) - 9/27(월)	Django 배포	Django
프로젝트 준비	9/28(화) - 9/29(수)	PPT준비 및 스크립트 작성	PowerPoint
프로젝트 발표	9/30(목)	프로젝트 발표	

수행절차 및 방법

03-2 데이터 수집

출처	데이터이름	제공형태	요약
서울특별시	서울시 건축물 정보	크롤링	서울시 건축물 정보 현황(611908x5)
공공데이터 포털	서울시 5대 범죄	CSV	서울시 5대 범죄 현황(79x14)
	서울시 CCTV	CSV	서울시 CCTV 현황(26x13)
	서울시 노후건물	CSV	서울시 건물 노후 현황(80x14)
	서울시 경찰관	API	서울시 경찰관 현황(11x4)
	서울시 치안센터	API	서울시 치안센터 현황(275x2)
	서울시 파출소	API	서울시 파출소 현황(116x3)
	서울시 외국인	API	서울시 외국인 현황(81x13)
서울 열린데이터 광장	서울시 유동인구	API	서울시 유동인구 현황(16199x14)
	서울시 유흥업소	CSV	서울시 유흥업소 현황(4877x25)
	서울시 인구밀도	CSV	서울시 인구밀도 현황(75x5)
	서울시 공시지가	API	서울시 공시지가 현황(376x3)
	서울시 여성 1인 가구	CSV	서울시 여성 1인 가구 현황(234x19)
	서울시 성비	CSV	서울시 성비 현황(28x11)
KOSIS(국가통계포털)	서울시 성비	CSV	서울시 성비 현황(28x11)

수행절차 및 방법

03-2 데이터 수집

따라서 이 연구는 최근 급증하는 1인 가구뿐 아니라 자치구별 특성과 범죄발생의 관계를 규명하고자 한다. 이를 위해 2014년부터 2016년까지의 서울시 25개 자치구의 5대 범죄 발생건수, 인구밀도, 1인 가구 수, 여성비율, 치안시설, 유흥시설 등의 자료를 활용하여 패널자료를 구축하고 패널회귀분석을 실시하였다. 이 연구는 향후 서울시 1인 가구 지원정책 또는 서울시 범죄예방 정책 수립 시 정책대상·사업지 선정 등에 1인 가구가 적극 반영될 수 있는 기초자료로 활용되는 토대를 제공할 것이다.

분야	위해지표	취약지표	경감지표
화재	화재 사망자 수	재난약자 수, 음식점 및 주점업 종사자 수, 창고 및 운송관련 서비스업 업체 수	병상수, 재정자주도, 도시지역면적
교통사고	교통사고 사망자 수	재난약자 수, 기초수급자 수, 의료보장사업장 수, 자동차등록대 수	의료기관수, 인구밀도, 구조구급대원수
범죄	5대 강력범죄 발생건수	총 전입자 수, 인구밀도, 기초수급자 수, 제조업 업체수, 음식점 및 주점업 업체수	경찰관서 수
안전사고	안전사고 발생건수	취청면적, 상리면적, 재난약자 수, 시군구외 전입자 수, 건설업 종사자 수, 제조업 종사자 수	의료보험료 수납액
자살	자살 사망자 수	고령인구 수, 혼인귀화자 수, 음식점 및 주점업 종사자 수, 기초수급자 수	보건업 및 사회복지 서비스업 종사자수
감염병	감염병 사망자 수	건강보험급여실적(임원), 고령인구 수, 기초수급자 수	의료기관 수, 인플루엔자 예방접종률
자연재해	지역안전도진단 결과 활용		

자료: 국민안전처(2015)

수행절차 및 방법

03-3 데이터 설명

5대 범죄?



절도

타인이 점유하고 있는 재물을 절취하는 범죄



강도

폭행,협박 등으로 다른 사람의 재산을 빼앗는 범죄



폭력

다른 사람에게 상해를 입히거나, 협박을 하는 등의 행위



살인

다른 사람을 죽이는 것



강간

폭행 또는 협박으로 사람을 간음하는 행위

수행절차 및 방법

03-4 데이터 정리



인구학적 요인

인구밀도
유동인구
외국인 비율
여성 1인 가구
성비



주거 환경 요인

건물 용도
파출소 개수
경찰관수
치안시설



범죄 발생요인

5대 범죄 발생
5대 범죄 검거



치안 환경 요인

CCTV 현황
유흥 업소 수
안심 지킴이 수



사회 경제적 요인

서울시 개별
공시 지가

수행절차 및 방법

03-5 데이터 파이프라인

1. 수집



2. 적재



3. 처리



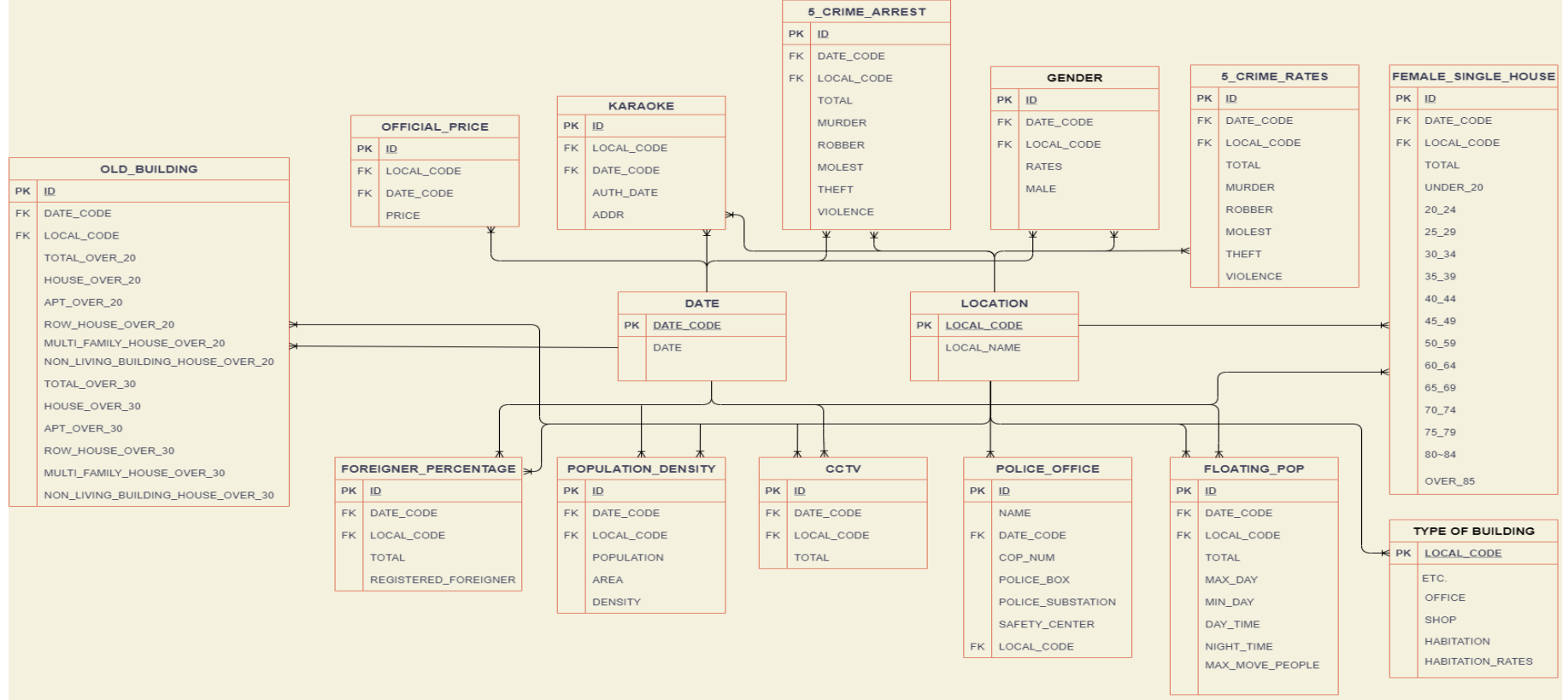
4. 시각화, 저장



DB구축

DB구축

04-1 ERD



DB구축

04-2 테이블 설정

Database changed
mysql> SHOW TABLES;

Tables_in_project
5_CRIME_ARREST
5_CRIME_RATES
BUILDING_TYPE
CCTV
DATE
FEMALE_SINGLE_HOUSE
FLOATING_POP
FOREIGNERS
GENDER
KARAOKE
LOCATION
OFFICIAL_PRICE
OLD_BUILDING
POLICE_OFFICE
POPULATION_DENSITY

15 rows in set (0.00 sec)

mysql> select * from information_schema.table_constraints where table_name = '5_CRIME_RATES';

CONSTRAINT_CATALOG	CONSTRAINT_SCHEMA	CONSTRAINT_NAME	TABLE_SCHEMA	TABLE_NAME	CONSTRAINT_TYPE
def	project	PRIMARY	project	5_CRIME_RATES	PRIMARY KEY
def	project	5_CRIME_RATES_ibfk_1	project	5_CRIME_RATES	FOREIGN KEY
def	project	5_CRIME_RATES_ibfk_2	project	5_CRIME_RATES	FOREIGN KEY

3 rows in set (0.00 sec)

시각화

시각화

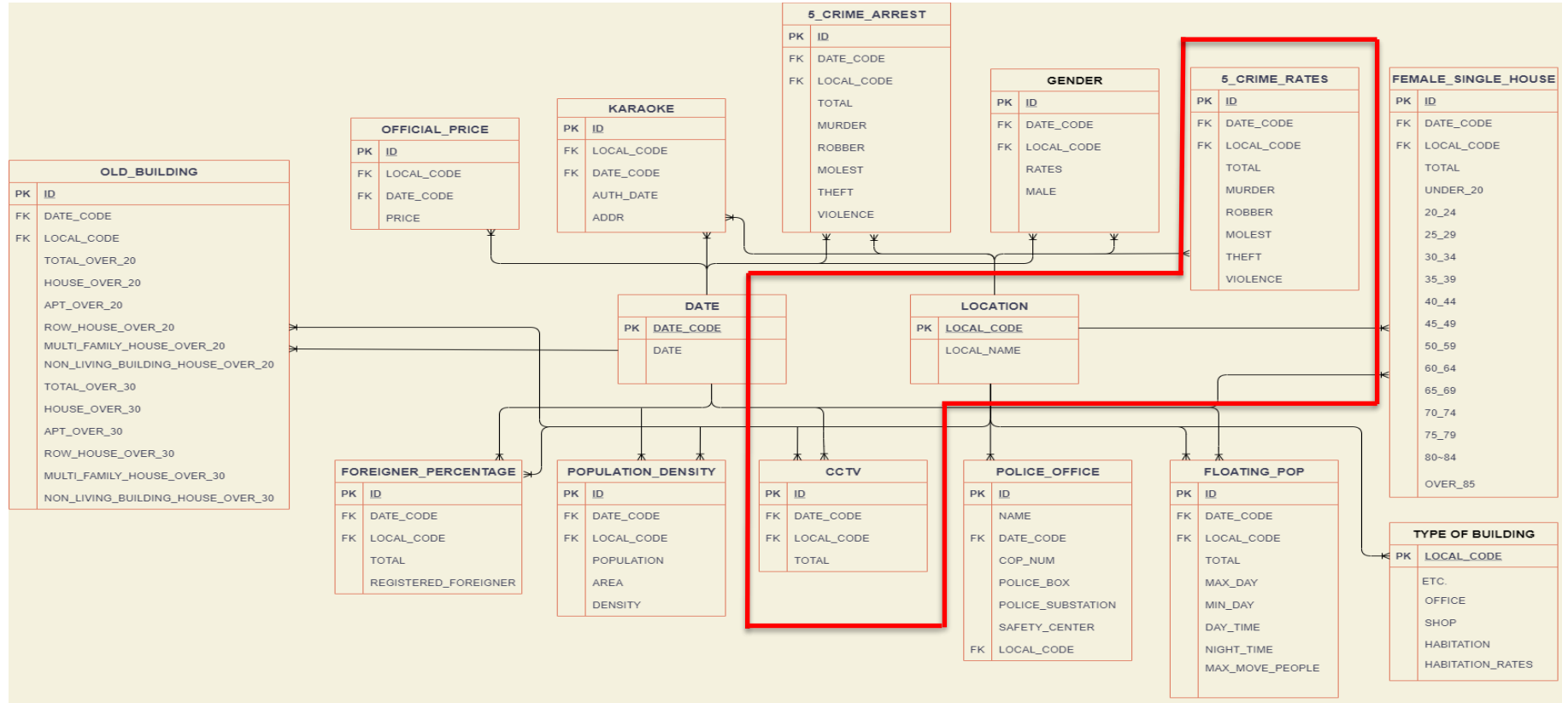
05-1 테이블 조인

```
SELECT LOCAL_NAME AS 지역구, A.TOTAL AS CCTV,  
C.TOTAL AS 5대범죄발생 FROM CCTV A  
JOIN LOCATION B  
ON A.LOCAL_CODE = B.LOCAL_CODE  
JOIN 5_CRIMES C  
ON A.LOCAL_CODE = C.LOCAL_CODE  
INTO OUTFILE '/val/lib/mysql-files/cctv5crime.csv'  
FIELDS TERMINATED BY ','  
LINES TERMINATED BY '\r\n';
```

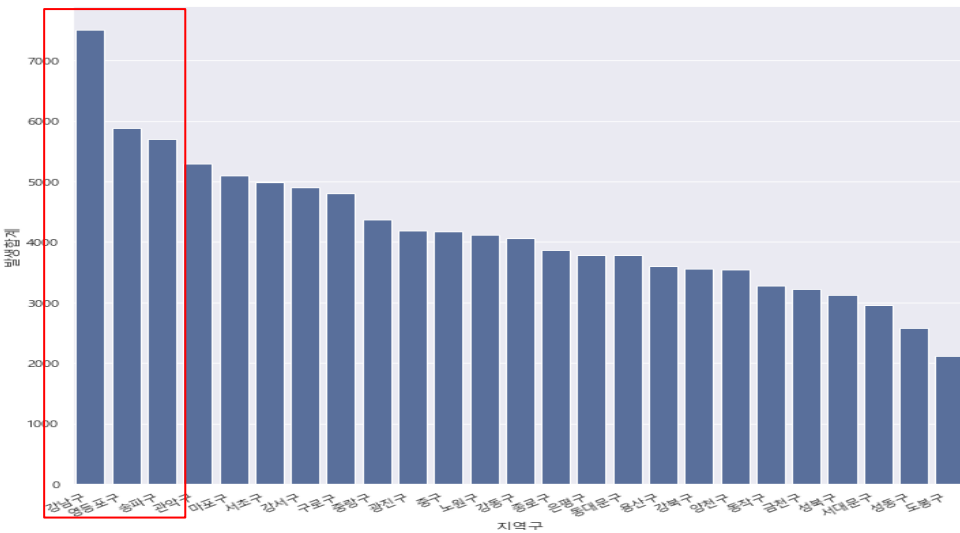
지역구	CCTV	5대범죄발
강남구	5559	7720
강남구	4770	7720
강남구	4022	7720
강남구	2560	5969
강남구	2495	5969
강남구	1837	5969
강남구	2518	5576
강남구	1450	5576
강남구	908	5576
강남구	4577	5525
강남구	3937	5525
강남구	3274	5525
강남구	2101	5278
강남구	1607	5278
강남구	1264	5278
강남구	2081	5135
강남구	1624	5135
강남구	1365	5135
강남구	3660	4895
강남구	3226	4895

시각화

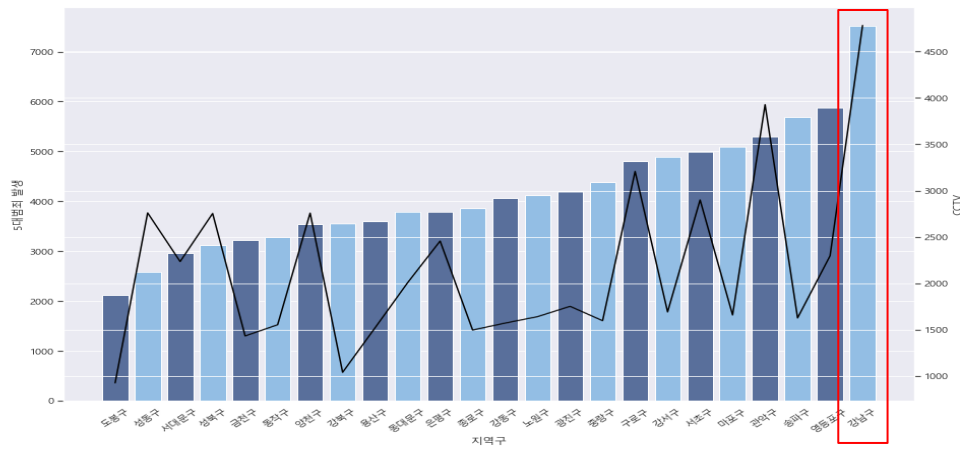
05-1 테이블 조인



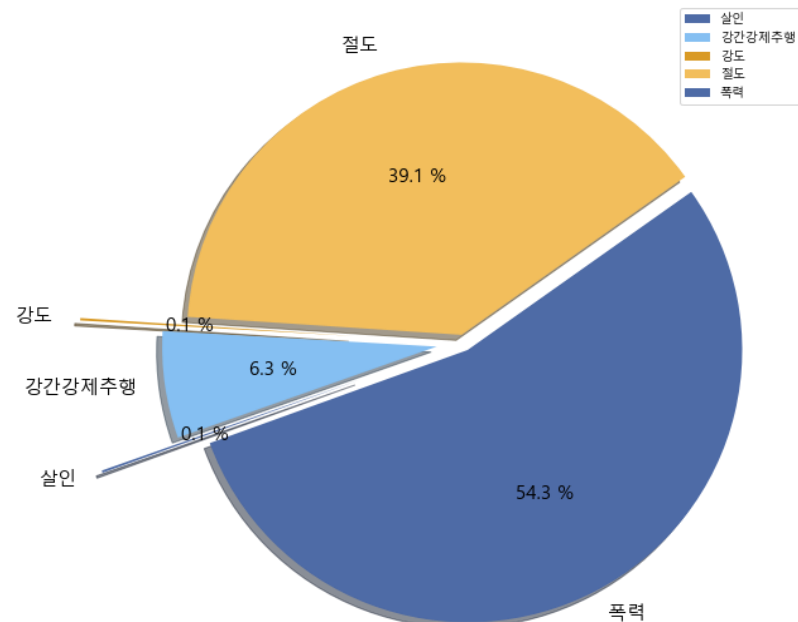
<서울시 지역구별 5대범죄발생건수>



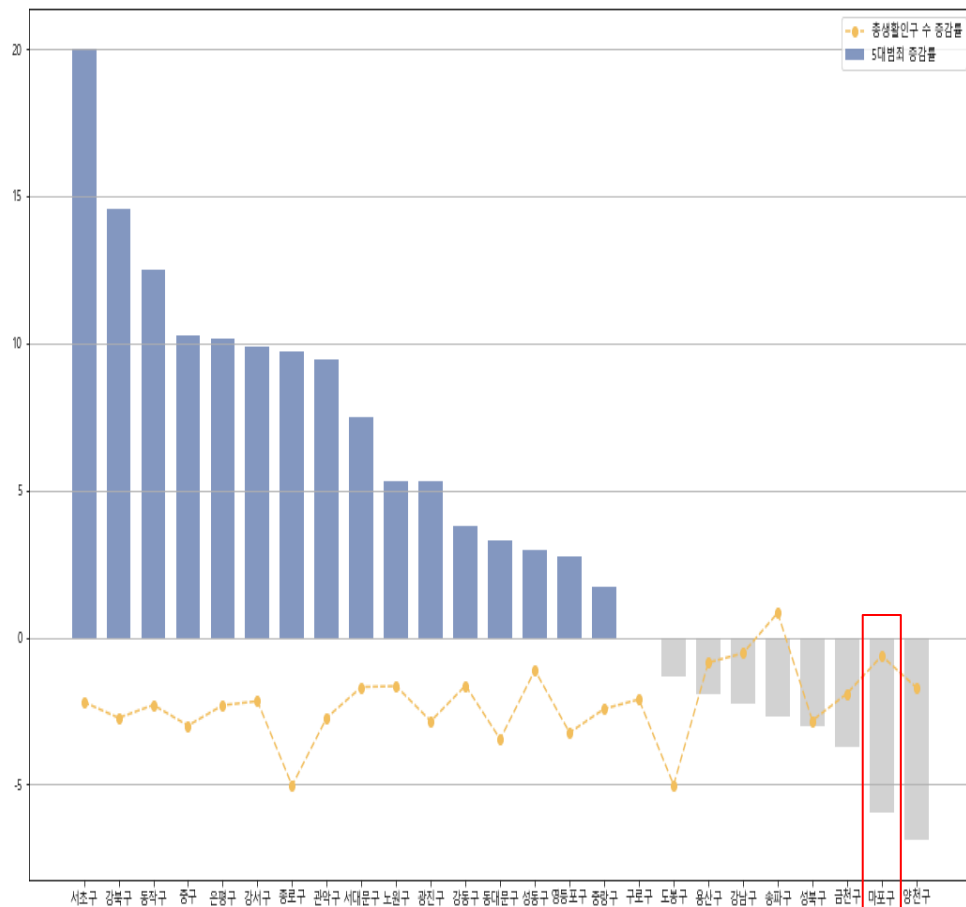
<서울시 CCTV개수와 5대범죄발생건수>



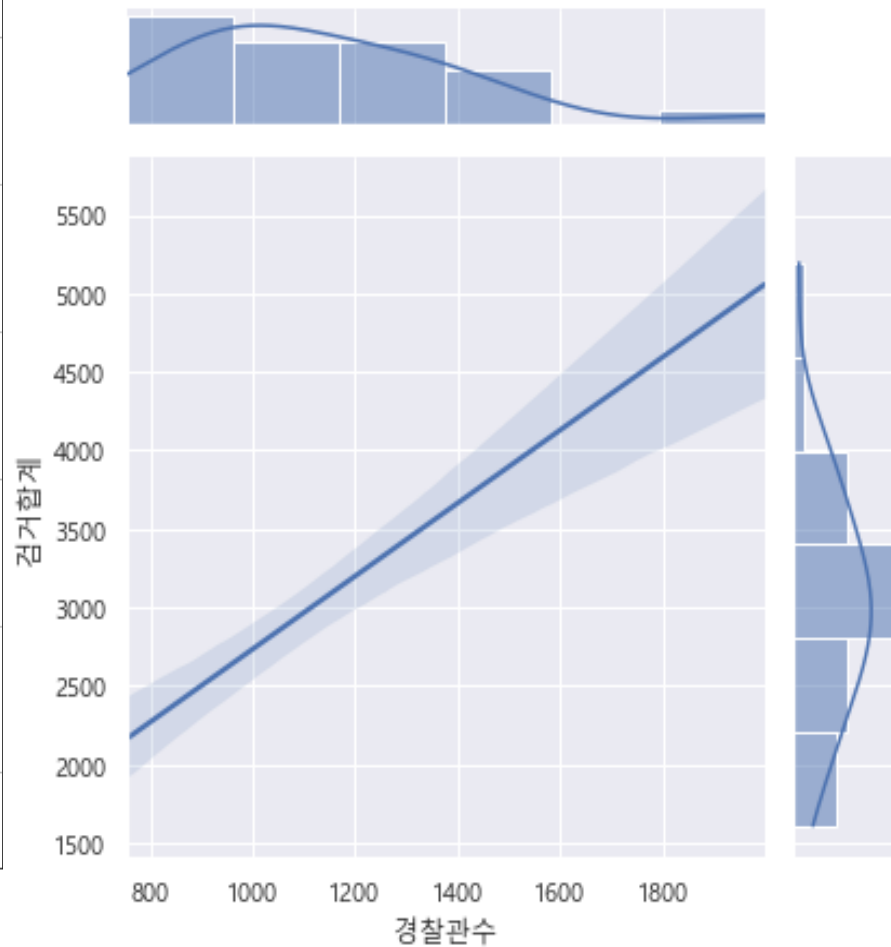
<2017~19년 서울 5대 범죄 비율>



<전년 대비 2019년 5대 범죄 및 총생활인구 수 증감률(>



<서울시 경찰관수대비 5대범죄 검거건수>





결론 및 향후과제

느낀점

감사합니다

