# 개별 탐색

손지호

특정 도시 개발할 때 어떤 시설물이 들어오면 좋을지 추천 (재개발 지역)

[건축연도별 주택현황 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00231_2017&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-문화예술관람활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028353&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-문화예술참여활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028357&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-스포츠관람활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028358&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-스포츠참여활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028359&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-관광활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028362&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-취미오락활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028364&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-휴식활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028365&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-사회 및 기타활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028381&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

2장 교통운수시설(주차장, 철도)

3장 도시공간시설(광장, 공원, 유원지, 관망타브 공공공지, 청소년수련시설)

4장 유통 및 공급시설(시장)

5장 공공문화복지시설(운동장, 학교, 도서관, 연구시설, 문화시설, 사회복지시설)

6장 도시방재시설(하천, 저수지)

7장 보건위생시설(장례식장, 종합의료시설)

[문화공간(공연장) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00164_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[문화공간(영화관) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00165_2009&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[문화공간(전시시설) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00166_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[문화공간(지역문화복지시설) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00167_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[문화재 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00399_2008&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[박물관 현황 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00401_2000&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[사설학원 및 독서실 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00195_2015&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[시립운동장 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00393_2016&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[공원 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_10052_2013&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

인구분포

[인구,가구 및 주택 – 읍면동(연도 끝자리 0,5), 시군구(그 외 연도) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1502&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[연령 및 성별 인구 – 읍면동(연도 끝자리 0,5), 시군구(그 외 연도) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1503&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

[성, 연령 및 세대구성별 인구 - 시군구 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1509&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

<https://www.data.go.kr/data/15108074/openapi.do>

<https://www.data.go.kr/data/15088190/openapi.do>

**국민여가활동 실태**

-> 추가 시설?

이 지역에는 이게 필요하다

이 지역 상권에 진출하려면 어떤 것이 가장 좋을지

* 이 지역의 인구분포에 따르면 이러한 여가시설이 필요(국민여가활동)
* 현재 이 지역의 여가시설 분포(데이터 있다면…)
* ⇒ 이 지역에서는 이게 성공할 것!

코로나 데이터

[COVID-19 in Korea dataset | Kaggle](https://www.kaggle.com/datasets/hongsean/covid19-in-korea-dataset)

→ 공공데이터 포털로

보건복지부 코로나19 시도 발생현황

<https://www.data.go.kr/data/15098776/openapi.do>

[COVID\_19 Datasets | Kaggle](https://www.kaggle.com/datasets/ognevdenis/covid-19-datasets)

드라마 데이터

[Top 100 Korean Drama (MyDramaList) | Kaggle](https://www.kaggle.com/datasets/chanoncharuchinda/top-100-korean-drama-mydramalist)

-> 연령대 별로 어떤 장르, 러닝타임, 캐스팅 등등을 분석한 후 새로운 드라마가 어느 연령대에 인기가 있을지 예측해보자!

⇒ 인기가 많을 것으로 예상되는 지역 선정 → 그쪽에 마케팅 비용 더 많이 사용

========================================

심장병이 전국적으로 모델링 하기도 좋을 것 같다. 일반 의료원, 동네 병원에서도 가능하지만

코로나는 지역별로 해서 조금 더 커져야 할 것이다.

코로나 중증은 지정된 병원이 있으니 전체적으로 봐야하는게 더 좋다. (지역과 연결하는건 전염병 특성상 조금 힘들 것이다.)

고위험군 백신 접종은 의사의 진단에 따라 판단되기 때문에 의사 진단 데이터가 없으면 조금 어려울 것이다.

지역 자체로 놓고 보자면 심장 질환이 더 효율적. (코로나, 심장병 둘 중으로 따졌을 때)

무엇을 할지 명확하다. 데이터도 명확하다.

여가활동

이 여가활동이 반드시 있어야 하느냐??

- 바로 옆 구에 있는데 굳이??? 인접 지역도 같이 생각 해봐야한다.

지역을 넓혀서 진행해보면 더 좋을 수도 있다!

지역을 대분류까지 넓히는 것도 생각해보면 좋다.

(여가활동 데이터는 서울시가 더 좋다.)

지하철

예산과 결부했을 때...가 문제 ㅎㅎ

분석 연습에는 좋은데 최종까지 갔을 때 효율성이 얼마나 있을지.

정치적으로도 연관되어 있기 때문에.. 프로젝트 주제로는 조금 어렵다.

(정치적 이슈가 좀 잦아지면은 가능하다)

범죄율

빈집과의 당위성을 보여줄 수 있는 논문자료가 필요할수도 있다.

이웃은 이용률과 연계해서 하면 좋다.

재미있으려면 마지막 의료데이터 두 개나 여가활동(중분류로 범위 넓혀서, 시간이 되면 지방과도 연관\_단순 비교. 분석은 서울에만!!)

1,2,3순위 정해보기

=================================================

[| CCJS 범죄와 형사사법 통계정보 (crimestats.or.kr)](https://www.crimestats.or.kr/portal/stat/easyStatCrimePage/800000.do)

범죄 관련 데이터 모여있는 곳

<https://www.news1.kr/articles/?3671243>

순찰 강화했더니 범죄율 낮아 졌다는 기사

[경찰청 서울특별시경찰청\_경찰서별 112 순찰차 현황](https://www.data.go.kr/data/15088384/fileData.do)

<http://www.riss.kr/link?id=T15829339>

한국 지역사회의 범죄율에 관한 연구 : 시민 참여가 범죄율에 미치는 영향을 중심으로

* 지방선거, 대통령선거, 국회의원 선거 투표율로 측정하는 공식적 시민참여 활동
* 장기기증률, 자원봉사율, 사회단체 참여율 등 비공식적 시민참여 활동

#### 

박민정

***221209***

미세먼지 농도에 따른 생활상의 차이 -> 미세먼지 농도가 높으면 어떤 행동을 하는 게 좋은지..??

* 미세먼지 농도: airkorea
* 생활상: 대중교통 이용도, 특정 가계 이용도, 배달 증가 등등
  + 서울 지하철 역별 승하차 인원 정보  
    <https://data.seoul.go.kr/dataList/OA-12914/S/1/datasetView.do>
  + 서울 버스노선별 정류장별 승하차 인원 정보  
    <https://data.seoul.go.kr/dataList/OA-12912/S/1/datasetView.do>
  + 한국도로공사 일자별 전국 교통량  
    <https://www.data.go.kr/data/15062049/openapi.do>

**지하철 이용 관련** - 교통 약자가 역을 얼마나 이용하는지 분석 -> 인원 배치, 시설 증설 등등

* 승하차 인원
  + 서울교통공사 역별/일별/시간대별 승객유형별 승하차 인원  
    <https://www.data.go.kr/data/15099330/fileData.do>
  + 서울교통공사 역별/월별 장애인 승하차인원 정보  
    <https://www.data.go.kr/data/15061870/fileData.do>
  + 서울교통공사 역별/월별 노인 승하차인원 정보  
    <https://www.data.go.kr/data/15062070/fileData.do>
  + 서울교통공사 역별/일별/시간대별 노인 승하차인원 정보  
    <https://www.data.go.kr/data/15101985/fileData.do>
* 편의 시설 현황
  + 서울교통공사 교통약자 이용시설 승강기 가동현황  
    <https://www.data.go.kr/data/15067770/openapi.do>
  + 서울교통공사 편의시설 현황  
    <https://www.data.go.kr/data/15044443/fileData.do>
* 대체 교통 수단
  + 전국교통약자이동지원센터정보표준데이터 <https://www.data.go.kr/data/15028207/standard.do>
  + 서울시설공단 장애인콜택시 일별 이용현황  
    <https://www.data.go.kr/data/15057705/openapi.do>
  + 서울시설공단 장애인콜택시 목적지 Best100  
    <https://www.data.go.kr/data/15057474/openapi.do>

결론

1. 역 시설 관련
   1. 도움 인원 배치
      1. 특정 시간대에 도움 인원을 더 배치
   2. 편의시설(엘리베이터, 휠체어리프트) 증설
      1. 편의시설 설치가 필요한 역 우선순위 계산
2. 대체 교통수단의 마련
   1. 셔틀 버스 노선
      1. 필요 근거: <https://www.yna.co.kr/view/AKR20220331051400502>
      2. 서울은 서울시설공단에서 장애인 콜택시, 서울시각장애인 생활이동지원센터에서 장애인 복지콜과 장애인 바우처 택시를 시행 중

건강기능식품 추천, 기능별로

* 식품의약품안전처 건강기능식품 <https://www.data.go.kr/data/15043011/openapi.do>
* 식품의약품안전처 건강기능식품정보 <https://www.data.go.kr/data/15056760/openapi.do>
  + 주된 기능성
* 식품의약품안전처 건강기능식품 이상사례 신고 현황 정보 <https://www.data.go.kr/data/15099480/fileData.do>
  + 신고접수현황 / 성별,연령별 / 제품 유형별 / 증상별 / 의료기관이용여부별 / 구입처구분별 정보
* 원재료 관련 DB도 있음 -> 알러지 유발

결과

1. 추천
2. 필터링
   1. 이상사례
   2. 원재료

Kaggle

1. 노보자임스사 효소 안정성  
   <https://www.kaggle.com/competitions/novozymes-enzyme-stability-prediction/overview/description>
   1. X: 단백질 순서(VPVNPEPDATSVENVAEKTGSGDSQS…)
   2. y: tm
2. 코로나19 데이터셋  
   <https://www.kaggle.com/datasets/meirnizri/covid19-dataset>
   1. X: 상태(성별, 연령, 임신 등), 기저질환(고혈압, 당뇨, 만성신부전 등)
   2. y: incubed(산소호흡기) 혹은 icu(중증환자) 혹은 death(사망)
   3. 국민건강보험
3. 버섯 먹을 수 있는지 여부  
   <https://www.kaggle.com/datasets/devzohaib/mushroom-edibility-classification>
   1. X: 버섯갓 지름, 모양, 표면, 색 등
   2. y: 식용 여부
4. 휴대폰 추천  
   <https://www.kaggle.com/datasets/meirnizri/cellphones-recommendations>
   1. X: 사용자 특성(연령, 성별, 직업), 평점
      1. 사용자의 특성 -> 휴대폰
      2. 사용자가 높게 평가한 휴대폰 -> 휴대폰
   2. y: 휴대폰
   3. 인구 특성과 결부해서 이 지역에서는 이것을 주력으로 하자 이런 느낌
5. Personal Key Indicators of Heart Disease  
   <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>
6. 화산 폭발  
   <https://www.kaggle.com/competitions/predict-volcanic-eruptions-ingv-oe>
   1. X: 10개의 센서 데이터
   2. y: 폭발 여부

공모전

1. 2022 유통/소비 도시가스 빅데이터 분석 해커톤  
   <https://linkareer.com/activity/113457> - 공고 링크   
   <https://kdx.kr/data/view/35568> - 데이터 링크

***221212***

범용 데이터 -> API

1. 인구 특성
   * 행정안전부 법정동별(행정동 통별단위) 성/연령별 주민등록 인구수 <https://www.data.go.kr/data/15108074/openapi.do>
   * 통계청 SGIS 오픈 플랫폼 인구통계 <https://www.data.go.kr/data/15088190/openapi.do>
2. 건강 정보
   * 국민건강보험공단 건강검진정보 <https://www.data.go.kr/data/15007122/fileData.do#tab-layer-openapi>
     + 시도(지역), 성별, 연령대, 신장, 체중, 허리둘레, 시력, 청력, 혈압, 혈당, 콜레스테롤, 요단백, 혈색소, 흡연, 음주, 치아 상태
     + 수치 데이터 -> 기준에 따라 고혈압 O/X 등으로 클래스 구분
   * 국민건강보험공단 2021 지역별 의료이용 통계연보 <https://www.nhis.or.kr/nhis/together/wbhaec06900m01.do?mode=view&articleNo=10829008&article.offset=0&articleLimit=10>
     + 만성질환: 시도, 고혈압, 당뇨, 치주질환, 관절염, 정신질환, 전염병, 간질환

Kaggle

1. 휴대폰 추천  
   <https://www.kaggle.com/datasets/meirnizri/cellphones-recommendations>
   1. X: 사용자 특성(연령, 성별, 직업), 평점
      1. 사용자의 특성 -> 휴대폰
      2. 사용자가 높게 평가한 휴대폰 -> 휴대폰
   2. y: 휴대폰
2. 코로나19 데이터셋  
   <https://www.kaggle.com/datasets/meirnizri/covid19-dataset>
   1. X: 상태(성별, 연령, 임신 등), 기저질환(고혈압, 당뇨, 만성신부전 등)
   2. y: incubed(산소호흡기) 혹은 icu(중증환자) 혹은 death(사망)
3. 심장 질환 <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>
   1. X: BMI, 생활 습관(흡연, 음주, 수면시간), 기저질환(당뇨, 천식), 상태(성별, 연령 등)
   2. y: 심장질환 여부

의견

1. 인구 특성 + 제품 정보 >> 추천
   1. 새로운 이용자의 인구 특성을 바탕으로 추천
   2. 우리나라 데이터를 찾아야 할 것 같음…
2. 건강 정보 + 질병 >> 위험군 분류
   1. 고위험군 분류해서 백신 우선순위 선정 & 알림
   2. 고위험군이 많이 위치한 곳에 의료기관 현황 분석 -> 지역별 의견?
      1. 전국의료기관표준데이터 <https://www.data.go.kr/data/15096293/standard.do>

Kaggle 데이터는 우리나라 데이터가 별로 없음. 국가별 차이 고려해야 함.

아니면 다른 사이트에서 크롤링하는 것도 생각해보기

국립암센터 암검진자 코호트 메타 정보

<https://www.data.go.kr/data/15034213/openapi.do>

암검진자 코호트 검사, 암과거력, 질병력, 수술력, 약물복용력, 가족력, 음주, 흡연, 암등록 및 사망, 운동, 감염, 일반사항, 검진일반, 검진특화, 검진, 문진, 남성문항, 여성문항 총 19개 정보에 대한 메타정보(제공 되어질 데이터 항목, 타입, 사이즈, 항목별건수, 샘플데이터 등)를 제공한다.

* 건강 검진 데이터와 엮어서?
  + 그렇기에는 암 환자가 아닌 사람의 정보가 없어서 모델링이 불가

성별, 연령, 임신

폐렴, 당뇨병, 만성폐색성폐질환,

==================

[심장 질환]

- 인구 특성과 결부

- 코로나보다는 심장

코로나로 한다면 중증 환자가 어느정도로 예측이 된다

지역과 연계는 전염병의 특성 상 불가능

중증환자 병상과 치료가 얼마가 더 필요하다

[여가]

- 반드시 있어야 하느냐? -> 바로 옆 구에 있다

- 인구 대비 부족하다

- 자치구를 동부권 등으로 나눠서

질문: 서울 자치구별 여가활동 데이터가 나을지?

https://data.seoul.go.kr/dataList/10557/S/2/datasetView.do

중분류로

시간이 되면 지방(5대광역시)과도 비교 -> 문화공간 현황을 단순 비교

시작: 서울에 시설이 집중되어 있음

-> 지방과의 격차를 넣으면 좋을 것 같음

시설현황은 현재에 맞추기

[지하철]

- 예산...

- 정치적 이슈 -> 피해야 함...

- 분석이 많음

[범죄율]

- 지금 얼마 전에 전국 범죄지도 나왔었음 -> 집값하고 연계

- 데이터 특성을 만들기 힘듦

- 특성에 대한 근거가 필요

- 문화공간 이용율

[할일]

1,2,3순위 정하기

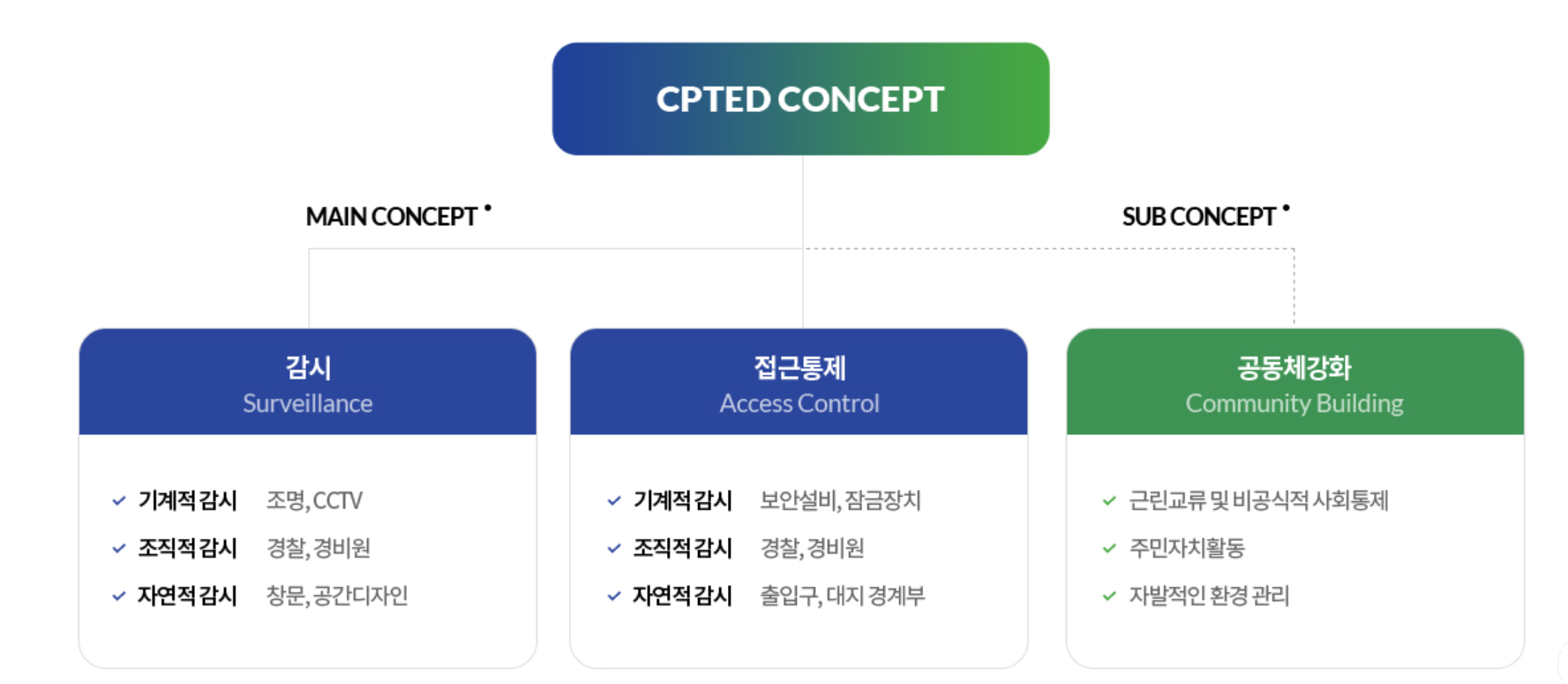
데이터 수집하기

다음주 화요일까지 최종 기획서

**CPTED(셉테드, Crime Prevention Through Environmental Design)**

*한국셉테드 학회*

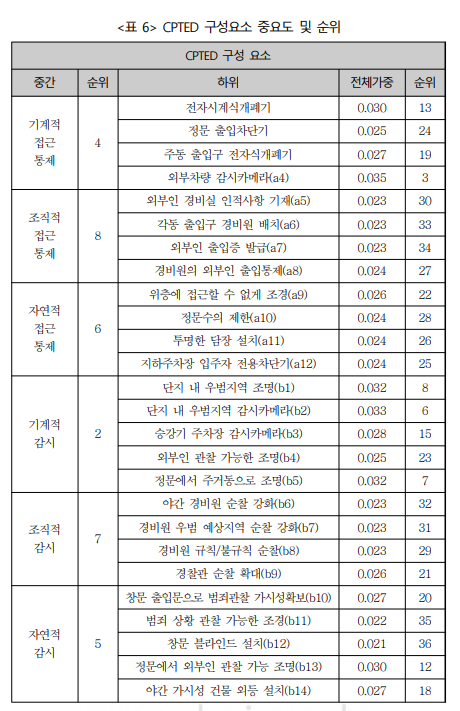
* “도시·건축환경의 적절한 설계와 효과적인 사용을 통해서 범죄 및 불안감을 감소시키고 삶의 질을 향상” 시키기 위한 다학제적(multi-disciplinary) 이론이자 실천전략입니다.
* 개발시대의 도시가 경제성, 기능성, 효율성 중심으로 성장함에 따라 사회 곳곳에 범죄 취약 환경이 양산되는 문제에 대한 반성으로 시작된 셉테드는 범죄예방의 패러다임을 변화시키고, 현대사회의 중요한 정책 및 연구과제로 자리매김하고 있습니다.
* 범죄가 “동기를 가진 범죄자, 취약한 대상, 환경”의 3요소가 갖추어질 때 발생한다는 점에 근거하여, 셉테드는 “감시와 접근통제, 공동체 강화”를 핵심 개념으로 정립하고 다양한 분야에서 응용되고 있습니다.
* 셉테드를 통한 범죄예방을 위해서는 건축, 도시, 경찰, 범죄, 사회, 심리 등 다양한 학문과 산업 분야의 협업이 중요하기 때문에, 관련 분야 종사자 및 연구자, 행정가 등의 활발한 교류와 소통이 필요하며 기본개념과 실천원리(디자인 원리 등)의 이해를 바탕으로 실효성 있는 범죄예방 전략을 도출하는 것이 매우 중요합니다.

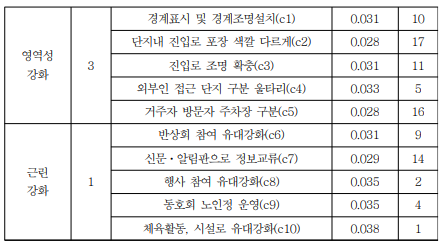


*정철우(2016)*

* 접근 통제
  + 일반적으로 아파트 단지에서 시행
  + 지역 기반에서는 하기 어려운 듯?
* 감시
  + CCTV
  + 가로등
  + 경찰 순찰
* 강화
  + 근린 강화
    - 반상회 참여 유대 강화
    - 행사 참여 유대강화
    - 동호회 노인정 운영
      * [노인 사회활동 참가](https://data.seoul.go.kr/dataList/47/S/2/datasetView.do) 2016년
      * [서울시 노인여가 복지시설 통계](https://data.seoul.go.kr/dataList/54/S/2/datasetView.do)
      * [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do) 2018년 자료
    - 체육활동, 시설로 유대 강화
      * [주로 이용하는 운동 장소](https://data.seoul.go.kr/dataList/10554/S/2/datasetView.do) 2019년

======================





**빈집**

염윤호(2019) 빈집과 절도/폭력 범죄의 관계에 대한 공간가중회귀모형의 적용

<https://doi.org/10.36889/KCR.2019.12.31.4.139>

| [서울시 하고 싶은 여가활동 통계](https://data.seoul.go.kr/dataList/10557/S/2/datasetView.do) | 서울특별시, 주민등록인구통계 | 여행야외나들이, 영상시청, 문화예술관람, 운동, 휴식, 종교활동, 문화예술참여활동, 컴퓨터게임인터넷검색 등, 운동경기관람, 사회봉사활동, 독서 |
| --- | --- | --- |
| [서울시 주말·휴일 여가활동 통계](https://data.seoul.go.kr/dataList/10260/S/2/datasetView.do) | 서울특별시, 서울시도시정책지표조사 | TV또는비디오시청, 여행야외나들이, 컴퓨터게임인터넷검색 등, 휴식, 사회활동, 운동, 문화예술관람, 창작적 취미활동, 운동경기관람, 종교활동, 독서 |
| [서울시 희망하는 여가활동 통계](https://data.seoul.go.kr/dataList/10261/S/2/datasetView.do) | 서울특별시, 주민등록인구통계 | TV및비디오(DVD) 시청, 문화예술관람, 문화예술참여, 스포츠관람, 스포츠활동, 관광활동, 컴퓨터게임인터넷검색 등, 취미/자기개발 활동, 휴식활동, 사회 및 기타활동 |
| [최근 1년 간 가장 많이 참여한 여가활동](https://data.seoul.go.kr/dataList/DT201013D025/S/2/datasetView.do)  5권역별 | 서울특별시, 서울특별시복지실태조사 | 휴식 활동, 취미/오락활동, 관광 여행, 문화예술관람활동, 사회봉사 및 단체 활동 참여, 스포츠 관람 활동, 문화예술 참여 활동, 스포츠 참여 활동 |

서대훈

1. data.go.kr/에 있는 서울특별시 동작구\_행정동별자동차등록대수와 서울특별시\_동작구\_등록인구 현황 데이터를 이용해 인구 대비 자동차 수 계산

2. kosis.kr/index/index.do년도별 추계인구, 합계출산율, 출생아수, 사망자수, 기대수명에 대해서 비교(출생아수, 사망자수 예측, 총인구수 예측)

3. 월별 미세먼지 평균

4. KOSIS 전국 범죄발생건수 대비 CCTV 대수(우범지대) 그래프로 비교(CCTV가 많을수록 범죄비율이 떨어진다는 전제를 증명하기 위함)

* 혹은 CCTV 대수 대비 범죄발생검거율(범죄 발생 및 검거 현황)(지역 경찰청별로)
* 결론: 인원 배치, CCTV, 파출소 등 제안
  + 경찰 인원 데이터, CCTV 데이터, 파출소 수 데이터 등
  + 인구 분포도 고려해야 함 -> 영향 미치는 요소이기에
* 범죄발생 및 검거 현황: <https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_2011_211&vw_cd=MT_ZTITLE&list_id=132_13204_GKIT659_dike256_eii6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE>
* CCTV  
  <https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do>
* 서울 지구대·파출소 1개소당 시민수  
  <https://data.seoul.go.kr/dataList/528/S/2/datasetView.do>
* 서울지역 경찰 공무원 현황  
  <https://data.seoul.go.kr/dataList/527/S/2/datasetView.do>
* 서울 도로시설물 통계(가로등 등)  
  <https://data.seoul.go.kr/dataList/261/S/2/datasetView.do>
* 반려동물 유무

(https://n.news.naver.com/mnews/article/015/0004721201?sid=104)  
<https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do>

5. LearnPlatform COVID-19가 디지털 학습에 미치는 영향

-디지털 학습 데이터를 사용하여 COVID-19가 학생 학습에 미치는 영향 분석

<https://www.kaggle.com/c/learnplatform-covid19-impact-on-digital-learning>

6. 지구 온난화

* 환경문제에 대한 인식변화(2008~2022):  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1SSEN051R&vw_cd=MT_TM2_TITLE&list_id=101_B7705&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_TM2_TITLE&path=%252FeasyViewStatis%252FcustomStatisIndex.do>
* 온실가스 배출량:  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_2KAAF02&vw_cd=MT_RTITLE&list_id=101_2022_O&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_RTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do>
* 연료 연소로 인한 온실가스 배출량:  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_2KAAD37&vw_cd=MT_RTITLE&list_id=101_2022_O&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_RTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do>
* 환경관련 세수:  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_2KAAF22&vw_cd=MT_RTITLE&list_id=101_2022_O&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_RTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do>
* 초미세먼지(PM2.5) 배출량(결측치 존재):  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_2KAAF26&vw_cd=MT_RTITLE&list_id=101_2022_O&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_RTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do>
* 대기오염 배출:  
  <https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_2KAAF23&vw_cd=MT_RTITLE&list_id=101_2022_O&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_RTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do>
* 결론 ->

7.

Kaggle

1. 세계 범죄 지수  
   : 전세계 주요 도시들의 범죄 지수와 안전지수(범죄지수+안전지수 = 100)  
   <https://www.kaggle.com/datasets/ahmadjalalmasood123/world-crime-index>
2. 미국 경찰 총격 사건 데이터(2015~2020)  
   <https://www.kaggle.com/datasets/ahsen1330/us-police-shootings>
3. 머신러닝을 이용한 질병 예측  
   <https://www.kaggle.com/datasets/kaushil268/disease-prediction-using-machine-learning?select=Testing.csv>
4. 성별 분류 데이터 셋  
   <https://www.kaggle.com/datasets/elakiricoder/gender-classification-dataset>
5. 진짜와 가짜 얼굴 감지  
   <https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection>
   1. 이미지 분석
6. 청소년&성인 식자율(비문맹률)  
   <https://www.kaggle.com/datasets/thedevastator/youth-adult-literacy-rates-in-2019>
7. 대학 전공 및 졸업생  
   : 직업 기회, 급여 및 성별 격차  
   <https://www.kaggle.com/datasets/thedevastator/uncovering-insights-to-college-majors-and-their>

- 코로나 고위험군 분류, 심장질환 위험군 분류

코로나와 심장 질환 위험군 분류가 비슷한 주제인데 코로나는 세계적인 것이지만 심장 질환 위험군 분류는 특정 지역에 한정하는 것이 가능하다고 함. 심장 질환 위험군 분류가 좋다고 칭찬 많이 하셨음.

- 서울시 자치구별 공공시설물 수요 예측

편의시설이 바로 옆 구에 있는데 꼭 여기도 있어야 하느냐. 거리로 따지면 강남구에서 송파구 가는 것 일도 아니다. 송파구에도 있는데 강남구에도 이걸 만들어야 하느냐. 인구대비 부족하다는 가능하지만 강남구에는 부족하지만 송파구에는 여유가 있다면 송파구에 가면 된다.

인접지역도 생각을 해볼 필요도 있다. 서울처럼 교통 잘 되어있고, 이동쉽고, 문화나 공공시설물이 집중되어있는 데. 수도권이나 지방으로 가면 굉장히 인구대비 많이 부족한데 인구가 많다고 서울에 이걸 계속 만들고 있기 때문에 그런 부분을 고려를 해야하지 않을까.

데이터는 충분해 보이는데 효율적으로 제시를 했을 때 그 구에서 이걸 받아들일만한 아이템인지는 잘 모르겠다.

- 지하철 교통약자

승하차 현황이 없는줄 알았는데 있다면 이것도 데이터는 정확하게 충분한 데이터가 있는거다. 그런데 문제가 예산하고 직결되는 거라 지금도 이슈다. 이런 것들이 앞으로 얼마나 예산이 반영되어있는지. 노인이나 장애인들은 지하철 같은 경우 무임승차다. 이런 편의를 이렇게 부족하니까 계속 늘려야한다고 이야기를 하기가 힘들다. 이건 (민감한 주제라)피해야하는 문제가 아닌가.

그리고 이것과 관련된 분석은 지금도 굉장히 많다. 분석시각화의 문제로 해보기는 괜찮은데 프로젝트로 하기는 힘들다. 정치적인 이슈가 잦아들면 문화편의측면, 인구적인 측면으로 봤을때도 필요한 내용이긴 하다. 앞으로 점점 인구가 불균형하게 가고 있다. 그래서 이 주제는 계속 분석해야하는 주제는 맞다.

결론: 정치적 이슈는 피하는 게 좋다.

- 서울시 자치구별 범죄율 분석

환경정비 항목 괜찮다고 함. 범죄율 분석할때 얼마전에 전국범죄지도가 하나 나왔다고 이슈가 됐었다. 생각보다 잘 안 만든다. 집값하고 관계가 되기 때문에 분석해놓고도 잘 발표 안 한다. 어쨌든 현재 상황에서는 있어야 되는. 검거율은 거의 100%라고 한다.

이웃간 교류는 데이터특성을 만들어내기가 힘들것 같긴 한데 만들어내면 기존의 분석보다 더 좋은 분석일 것 같다. 지금까지 있는건 CCTV, 유동인구, 경찰, 주민등록인구 정도인데 새로운 특성 2개가 들어가는 것인데 빈집같은 것이 범죄와 어떤 연관이 있는지. 가미성. 논문자료를 먼저 보여주고 시작을 하긴 해야할 거다. 이웃간 교류는 단순하게 얘가 있고 없고의 통계와 더불어서 이용율.여가와 연계해서 봐도 되긴 하겠다.

- 서울시 자치구별 공공시설물 수요 예측

분석은 서울에 국한되어서 하고 시작할때 인구가 서울이 많기 때문에 서울에 집중되어있다고 시작하고 인구대비 지방광역시는 얼마나 부족한지 하나로 보여주고 끝맺음 하면 괜찮을 것 같다.

전체적으로 데이터가 충분한 주제들을 가지고 왔다.

선정한 주제가 5개인데 1, 2, 3순위 정도 정해서 데이터 수집을 진행을 하고, 멘토님한테 1순위에 대한 것을 적극적으로 어필하고, 별로라 하면 2순위, 3순위를 보여주는 형태.

그리고 다 끝나고 나면 5개를 다 멘토링을 받아봐라.

질문들

1. 데이터에 따라 연도가 다른 경우 연도가 통일이 필요한가?

여가활동이 21년이면 21년에 맞춰야 한다. 인구분포는 22년 조사만 있나?

인구분포는 22년것을 써도 되긴 할 것 같다. 박물관이나 공원은 시설과 관련된 것이라 21년으로 맞춰야 한다. 21년을 못맞추겠으면 22년 현재를 기준. 여가활동은 21년을 써도 되는데 그 이전년도 데이터는 사용하면 안 된다. 차라리 이후로. 박물관 현황이나 공원현황은 크롤링을 하더라도 현재에 맞춰서 진행을 해야 할 것 같다.

주제 1, 2, 3순위 결정해놓고 멘토링받고 어떤 주제로 결정되었는지 최종계획서 써서 담주 화요일쯤에 구글 드라이브에 써놓아라.

* [코로나 19로 인해 전세계적으로 범죄발생률이 감소함](https://n.news.naver.com/mnews/article/028/0002546976?sid=105)
* [반려견 산책이 잦은 곳은 범죄율이 낮다.](https://www.thedailypost.kr/news/articleView.html?idxno=88187)
* [1인가구 밀집지역은 범죄율이 높다.](https://www.lawtimes.co.kr/Legal-News/Legal-News-View?serial=140865)
* [인플레 1% 오르면 살인률 1.75% 상승](https://n.news.naver.com/mnews/article/001/0013315609?sid=102)
* [CCTV 설치가 범죄 발생률과 검거율에 미치는 영향](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=8fa1103bb15bfe93ffe0bdc3ef48d419&outLink=N)
* [가로등만 설치해도 범죄 16%↓…공동잠금장치 설치하면 43%↓](https://www.yna.co.kr/view/AKR20200108078300004)  
  ([경찰청 보도자료 파일 원본](https://www.korea.kr/news/pressReleaseView.do?newsId=156370145))
* [1인 가구와 범죄 발생에 관한 연구](https://www.si.re.kr/si_download/60865/22904)
* [편의업종의 확산이 범죄발생률에 미치는 영향 연구](http://kpaj.or.kr/xml/03231/03231.pdf)
* [경제적 요인이 범죄율에 미치는 영향 : 전국 시·군·구를 중심으로](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=6ba54291892aeb15ffe0bdc3ef48d419&outLink=N)
* [커뮤니티 활성화가 주거지의 범죄 예방에 미치는 영향 : 서울시 동작구를 대상으로](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=c07bc070a53f5a2fffe0bdc3ef48d419&outLink=N)

# 221209 회의 정리

주제 후보

1. 지하철 관련
2. 여가활동과 인구 분포
3. CCTV 관련

계획서

조 이름 -> 주제 결정되고 정하기

팀장: 박민정

팀원: 서대훈, 손지호

주제 -> 주제 결정되고 정하기

수행방향 -> 주제 결정되고 정하기

수행 도구

* 파이썬 -> pandas, scikit-learn
* R -> ggplot2

프로젝트 목적

필수 기능

포함 기술

WBS 첨부 (OR 별도로 폴더 업로드)

프로젝트 1주차 수행 일지

프로젝트 타이틀

팀명

팀원

프로젝트 수행 계획 및 현황

| 금주 수행 내용 | 차주 수행 내용 |
| --- | --- |
| 주제 후보 선정 | 데이터 및 방향 구체화 |

# 221212 회의 정리

주제 후보

1. 지하철 관련
   1. 역 시설 관련
      1. 도움 인원 배치 → 특정 시간대에 도움 인원을 더 배치
      2. 편의시설(엘리베이터, 휠체어리프트) 증대 → 편의시설 설치가 필요한 역 우선순위 계산
   2. 대체 교통수단의 마련
      1. 셔틀 버스 노선
2. 여가활동과 인구 분포
   1. 수요 예측
   2. 법으로 명시된 기준에 부합한지 확인

<https://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=40062#AJAX>

* 2장 교통운수시설(주차장, 철도), 3장 도시공간시설(광장, 공원, 유원지, 관망탑 공공공지, 청소년수련시설), 4장 유통 및 공급시설(시장), 5장 공공문화복지시설(운동장, 학교, 도서관, 연구시설, 문화시설, 사회복지시설), 6장 도시방재시설(하천, 저수지), 7장 보건위생시설(장례식장, 종합의료시설)
* 의견 제시할 때 이 지역에 이게 필요하다는 것의 근거로 제시

1. CCTV 관련
   1. 지역별 비교 → 개선 사항 제시
2. 질병(코로나19, 심장질환)   
   → 새로운 개인이 고위험군에 속하는지 & 지역별로 고위험군 비율
   1. 고위험군 분류해서 알림 or 백신, 다른 검사 우선순위 선정
   2. 고위험군이 많이 위치한 곳에 의료기관 현황 분석 -> 지역별 의견?
      1. 전국의료기관표준데이터 <https://www.data.go.kr/data/15096293/standard.do>

**프로젝트 기획안**

1. 팀명 : 우리동네 행복 공간 지킴이
2. 프로젝트 주제
   1. 쾌적한 도시 생활을 위한 공공시설 수요 예측
3. 프로젝트 수행 방향
   1. 데이터 설명
      1. 인구분포현황  
         [건축연도별 주택현황 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00231_2017&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 자치구 칼럼 모두, 계, 단독주택, 아파트, 연립주택, 다세대주택, 2017-2020년  
[인구,가구 및 주택 – 읍면동(연도 끝자리 0,5), 시군구(그 외 연도) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1502&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 서울시 자치구 칼럼  
[연령 및 성별 인구 – 읍면동(연도 끝자리 0,5), 시군구(그 외 연도) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1503&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 서울시 자치구 칼럼, 나이대 별

[성, 연령 및 세대구성별 인구 - 시군구 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1509&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 서울시 데이터  
<https://www.data.go.kr/data/15108074/openapi.do>  
<https://www.data.go.kr/data/15088190/openapi.do>

* + 1. 문화공간 현황  
       [문화공간(공연장) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00164_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 공연장(2)(1)(공공or민간 공연장 / 규모별 공연장) (연도는 1993-2020 중 선택 가능)  
[문화공간(영화관) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00165_2009&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 개소수 (연도는 2009-2020 중 선택 가능)  
[문화공간(전시시설) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00166_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 박물관, 미술관, 화랑, 전시실 (연도는 1993-2020 중 선택 가능)  
[문화공간(지역문화복지시설) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00167_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 문예회관, 복지회관, 구민(다목적) 체육시설, 청소년수련시설 (연도는 1993-2019 중 선택 가능)  
[문화재 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00399_2008&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 문화재(2)(1)(국가or시 지정 문화재), 문화재(3)(1)(국보, 보물, 사적, 명승, 천연기념물) (연도는 2008-2020 중 선택 가능)  
[박물관 현황 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00401_2000&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 박물관(1)(국립공립, 대학, 사립) (연도는 2001-2020 중 선택 가능)  
[사설학원 및 독서실 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00195_2015&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 학원/독서실(1)(사설학원, 독서실), 학원/독서실(2)(1)(학원수) (연도는 2015-2019 중 선택 가능)  
[공원 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_10052_2013&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 공원수/면적(1)(공원수) (연도는 2013-2019 중 선택 가능)

* + 1. 여가활동 현황  
       [지난 1년 동안 한번이상 참여한 여가활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028353&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별), 영화관람, 박물관관람, 연극공연관람, (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동유형(복수응답)-문화예술참여활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028357&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-스포츠관람활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028358&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-스포츠참여활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028359&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-관광활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028362&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-취미오락활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028364&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-휴식활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028365&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)  
[지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답)-사회 및 기타활동 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028381&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)

-> 통계분류(1)(성별, 연령별, 동거가구원수별) (연도는 2016-2021 중 선택 가능)

* 1. 주요 분석 및 모델링 내용
     1. 여가활동 현황 데이터를 분석하여 성별/연령대별 주로 참여하는 여가활동을 파악한다. 이를 지역별 인구 분포에 적용하여 지역별로 주로 원하는 여가활동을 도출한다.
     2. 도시별, 지역별 공공시설의 설치 현황을 바탕으로 부족한 공공시설을 조사하여 추후에 시설을 설치해야 할 때 어떤 시설을 설치 해야할지 조사한다.
     3. 특정 지역, 도시에 새로운 시설을 설치할 때 그 시설이 설치 되기에 적절한 지역, 도시인지 점검해 볼 수 있다.
     4. 특정 인구의 분포가 많은 지역에는 어떤 시설을 설치하는 것이 적합한지 예측할 수 있다.

============================================================

1. 팀명 :
2. 프로젝트 주제
   1. 지하철 교통약자 편의를 위한 모델 구현
   2. 지하철 교통약자들의 조금 더 편리한 이용
3. 프로젝트 수행 방향
   1. 데이터 설명
      1. 교통약자 승하차 현황
         1. 서울교통공사 역별/월별 장애인 승하차인원 정보 <https://www.data.go.kr/data/15061870/fileData.do>
         2. 서울교통공사 역별/일별/시간대별 노인 승하차인원 정보 <https://www.data.go.kr/data/15101985/fileData.do>
      2. 편의 시설 증설
         1. 서울교통공사 편의시설 현황 <https://www.data.go.kr/data/15044443/fileData.do>
            1. 휠체어리프트 여부, 승강기 여부
         2. 서울교통공사 교통약자 이용시설 승강기 가동현황 <https://www.data.go.kr/data/15067770/openapi.do>
            1. 승강기 구분(ES vs. EV vs. WL), 운행상태(사용가능 vs. 보수중 vs. 점검중)
      3. 셔틀 노선 제시
         1. 서울시설공단 장애인콜택시 목적지 Best100 <https://www.data.go.kr/data/15057474/openapi.do>
            1. 목적지 시도, 시군구, 동, 최소단위(동)에 대한 건수
   2. 주요 분석 및 모델링 내용
      1. 교통약자(노인, 장애인)의 지하철 이용 빈도를 분석하여 교통 편의를 증진하고자 한다.
         1. 노인이 많이 이용하는 시간대를 기반으로 역에 도움 인원 배치를 효율화하고자 한다.
         2. 노인과 장애인이 많이 이용하는 역을 기반으로 승강기, 휠체어리프트 등의 편의 시설을 우선적으로 증설해야 하는 역을 선정하고자 한다.
         3. 장애인이 많이 이용하는 역의 데이터와 장애인콜택시 데이터를 바탕으로 대체 교통 수단으로 셔틀 노선을 제시하고자 한다.

=====================================================================

1. 팀명
2. 프로젝트 주제
   1. 서울지역 연도별 CCTV 대수 통계와 연도별 범죄발생 검거율 비교
3. 프로젝트 수행 방향
   1. 데이터 설명
      1. 서울 경찰청 범죄발생 및 검거 현황  
         <https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_2011_N213&vw_cd=MT_ZTITLE&list_id=132_13204_GKIT659_dike256_eii6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE>
      2. 서울시 자치구 연도별 밤범용 cctv 운영 현황  
         <https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do>
      3. 경기 북부 경찰청 범죄발생 및 검거 현황  
         <https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_2011_N231&vw_cd=MT_ZTITLE&list_id=132_13204_GKIT659_dike256_eii6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE>
      4. 경기 남부 경찰청 범죄발생 및 검거 현황  
         <https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_2011_N230&vw_cd=MT_ZTITLE&list_id=132_13204_GKIT659_dike256_eii6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE>
      5. 경기도 CCTV 현황  
         <https://data.gg.go.kr/portal/data/service/selectServicePage.do?page=1&sortColumn=&sortDirection=&infId=VIPK0N53968Q7DV5TT2312643570&infSeq=1&searchWord=cctv>
      6. 대구 경찰청 범죄발생 및 검거 현황  
         <https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_2011_N215&vw_cd=MT_ZTITLE&list_id=132_13204_GKIT659_dike256_eii6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE>
      7. 대구광역시 CCTV  
         <https://data.daegu.go.kr/open/data/dataView.do?dataSetId=15013094&dataSetDetailId=150130941dcf249b56908_201909241818&provdMethod=FILE#this>
      8. 서울시 5대범죄 발생 현황  
         <https://data.seoul.go.kr/dataList/316/S/2/datasetView.do>
      9. 서울시 5대범죄 발생장소별 현황  
         <https://data.seoul.go.kr/dataList/10943/S/2/datasetView.do>
      10. 서울시 지구대·파출소 1개소당 시민수(경찰서, 치안센터 제외)  
          <https://data.seoul.go.kr/dataList/528/S/2/datasetView.do>
      11. 서울시 경찰공무원 현황 통계  
          <https://data.seoul.go.kr/dataList/527/S/2/datasetView.do>
      12. 시내주요기관(경찰·소방관서)(경찰서, 경찰청, 치안센터 포함)  
          <https://data.seoul.go.kr/dataList/224/S/2/datasetView.do>
      13. 서울시 도로시설물 통계(가로등)(구별로도 조회 가능)  
          <https://data.seoul.go.kr/dataList/261/S/2/datasetView.do>
      14. 경찰청 서울특별시경찰청\_경찰서별 경찰관수 현황(구별로도 조회 가능)  
          <https://www.data.go.kr/data/15036460/fileData.do>
      15. 경찰청 서울특별시경찰청\_경찰서별 지역경찰 현황(지구대, 파출소, 치안센터)  
          <https://www.data.go.kr/data/3074546/fileData.do>
      16. 반려동물 유무
      17. (https://n.news.naver.com/mnews/article/015/0004721201?sid=104)  
          https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do
   2. 주요 분석 및 모델링 내용

서울시 우리마을가게 상권분석서비스(상권-상주인구)

<https://data.seoul.go.kr/dataList/OA-15584/S/1/datasetView.do>

서울시 우리마을가게 상권분석서비스(상권-생활인구)

<https://data.seoul.go.kr/dataList/OA-15568/S/1/datasetView.do>

=====================================================================

1. 팀명
2. 프로젝트 주제
   1. 코로나19 고위험군 예측 모델 구현
   2. 심장 질환 위험군 분류 모델 구현
3. 프로젝트 수행 방향
   1. 데이터 설명(i 혹은 ii 둘 중 하나만 사용 예정)
      1. COVID-19 Dataset <https://www.kaggle.com/datasets/meirnizri/covid19-dataset>
         1. 멕시코 코로나 환자 1,048,576명에 대한 데이터
         2. X: 상태(성별, 연령, 임신 등), 기저질환(고혈압, 당뇨, 만성신부전 등), 생활습관(담배)
         3. y: icu(중증환자) 혹은 death(사망)
      2. Personal Key Indicators of Heart Disease <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>
         1. X: BMI, 생활 습관(흡연, 음주, 수면시간), 기저질환(당뇨, 천식), 상태(성별, 연령 등)
         2. y: 심장질환 여부
      3. 국민건강보험공단 건강검진정보 <https://www.data.go.kr/data/15007122/fileData.do#tab-layer-openapi>
         1. 시도(지역), 성별, 연령대, 신장, 체중, 허리둘레, 시력, 청력, 혈압, 혈당, 콜레스테롤, 요단백, 혈색소, 흡연, 음주, 치아 상태
         2. 수치 데이터 -> 기준에 따라 고혈압 O/X 등으로 클래스 구분
      4. 전국의료기관표준데이터 <https://www.data.go.kr/data/15096293/standard.do>
   2. 주요 분석 및 모델링 내용
      1. 코로나19 환자의 데이터를 분석하여 중증 혹은 사망과 관련된 특성을 파악하고, 분류 모델을 작성한다.
      2. 심장 질환과 관련된 특성을 파악하고, 분류 모델을 작성한다.
      3. 앞서 작성한 모델에 건강검진정보 데이터를 적용하여 고위험군을 분류한다.
      4. 고위험군에 해당하는 개인에게 개별 추가 검사를 안내할 수 있으며 지속적으로 관찰을 시행할 수 있을 것이다.
      5. 지역별로 위험군으로 분류된 인원을 계산하여 위험군으로 분류된 인원과 의료기관 수준을 분석하여, 의료기관이 부족한 부분을 찾는다.

======================================================================

1. 역할 분담
2. 프로젝트 수행도구
   1. 전처리 기술
      1. pandas
   2. 시각화 기술
      1. ggplot2
   3. 모델링 기술
      1. scikit-learn

# 221213 1차 기획안

| **서울시 자치구별 공공시설물 수요 예측**  **팀명 : 우리동네 여유 지킴이** | |
| --- | --- |
| 데이터 | ● 인구 분포 현황 [통계청\_SGIS오픈플랫폼\_인구통계](https://www.data.go.kr/data/15088190/openapi.do)   * 인구주택총조사 정보 중 전체적인 인구 통계 정보를 조회 * 요청: 기준년도, 행정기관코드 * 결과: 총인구, 평균나이, 인구밀도, 노령화지수, 노년부양비, 유년부양비, 총가구, 평균가구원수, 총주택   [행정안전부\_행정동별(통반단위) 성/연령별 주민등록 인구수](https://www.data.go.kr/data/15108072/openapi.do)   * 요청: 행정기관코드, 조회시작년월(2022년 10월) * 결과: 통계년월, 시도명, 시군구명, 총인구수, 남자인구수, 여자인구수, 만0~9세남자, 만10~19세남자, …, 만100세이상남자,  만0~9세여자, 만10~19세여자, …, 만100세이상여자   ● 여가활동 현황  [지난 1년 동안 한번이상 참여한 여가활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028353&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)   * 통계분류(1): 성별, 연령별, 동거가구원수별 * 2016년 ~ 2021년  | [문화예술참여활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028357&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 영화관람, 박물관관람, 연극공연관람 | | --- | --- | | [스포츠관람활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028358&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 사진촬영, 악기연주/노래교실, 미술활동, 문화행사참여 | | [스포츠참여활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028359&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 배드민턴/줄넘기, 헬스/에어로빅, 농구/배구, 당구/포켓볼, 볼링/탁구, 요가/필라테스, 육상/조깅, 사이클링, 수영, 격투스포츠, 테니스/스쿼시, 댄스스포츠, 인라인스케이트 | | [관광활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028362&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 자연명승, 문화유적지방문, 삼림욕, 테마파크가기, 소풍/야유회 | | [취미오락활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028364&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 쇼핑/외식, 음주, 미용, 게임, 등산, 독서, 노래방가기, 어학/기술/자격증 취득 공부 | | [휴식활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028365&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 모든 열 다 가능할 듯. (집에서 가능한 것들. 그러면 아예 빼도 좋을 듯..??) | | [사회 및 기타활동](https://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028381&vw_cd=MT_ZTITLE&list_id=113_11314_001&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 친구만남, 잡담/통화 하기, 가족 방문, 사교모임, 종교활동, 동호회 모임 |   ● 문화공간 현황   |  | 연도 | 활용 예정 컬럼 | | --- | --- | --- | | [문화공간(공연장)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00164_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 93 - 20 | 공연장(2)(1)(공공or민간 공연장 / 규모별 공연장) | | [문화공간(영화관)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00165_2009&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 93 - 20 | 개소수 | | [문화공간(전시시설)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00166_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 93 - 20 | 박물관, 미술관, 화랑, 전시실 | | [문화공간(지역문화복지시설)](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00167_1993&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 93 - 19 | 문예회관, 복지회관, 구민(다목적) 체육시설, 청소년수련시설 | | [문화재](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00399_2008&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 08 - 20 | 문화재(2)(1)(국가or시 지정 문화재), 문화재(3)(1)(국보, 보물, 사적, 명승, 천연기념물) | | [박물관 현황](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00401_2000&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 01 - 20 | 박물관(1)(국립공립, 대학, 사립) | | [사설학원 및 독서실](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00195_2015&vw_cd=MT_ZTITLE&list_id=201_20104_M&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 15 - 19 | 학원/독서실(1)(사설학원, 독서실), 학원/독서실(2)(1)(학원수) | | [공원](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_10052_2013&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do) | 13 - 19 | 공원수/면적(1)(공원수) | | [서울시 유통업체 현황](https://data.seoul.go.kr/dataList/175/S/2/datasetView.do) | 95 - 21 | 대형마트, 백화점, 전문점, 쇼핑센터, 복합쇼핑 | | [서울시 식품위생업 현황](https://data.seoul.go.kr/dataList/33/S/2/datasetView.do) | 97 - 21 | 유흥주점, 단란주점 | | [서울시 공공위생업 현황](http://data.seoul.go.kr/dataList/32/S/2/datasetView.do) | 11 - 21 | 미용업 | | [서울시 공공도서관 통계](https://data.seoul.go.kr/dataList/11064/S/2/datasetView.do) | 07 - 21 | 시.시도 도서관, 교육청도서관, 사립도서관 | | [서울시 도시계획시설 통계](https://data.seoul.go.kr/dataList/245/S/2/datasetView.do) | 04 - 21 | 공간시설: 광장, 공원, 녹지, 유원지, 공공공지  공공문화체육시설: 운동장, 체육시설, 공공청사, 학교, 도서관, 연구시설, 문화시설, 사회복지시설, 청소년수련시설) |   ● 공간 관련 정보  [건축연도별 주택현황](https://kosis.kr/statHtml/statHtml.do?orgId=201&tblId=DT_201_00231_2017&vw_cd=MT_ZTITLE&list_id=201_20104_I&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=MT_ZTITLE&path=%252FstatisticsList%252FstatisticsListIndex.do)   * 2017년 ~ 2020년 * 자치구별 계, 단독주택, 아파트, 연립주택, 다세대주택 * 어떤 주택 현황이 많은지에 따라 무슨 여가 시설을 추가할지 의견 |
| 주요 분석 및 모델링 내용 | 여가활동 현황 데이터를 분석하여 성별/연령대별 주로 참여하는 여가활동을 파악한 후, 이를 서울시 구별 인구 분포에 적용하여 해당 지역 사람들이 주로 원하는 여가활동을 도출하고자 한다.   1. 특정 인구(성별, 연령)의 분포가 많은 지역에는 어떤 시설을 설치하는 것이 적합한지 예측할 수 있다. 2. 서울시 구별 공공시설의 설치 현황과 비교하여 수요가 높은 공공시설을 파악하고자 한다. 3. 특정 구에 새로운 시설을 설치할 때, 그 시설이 수요가 있는지 점검하고자 한다. |

| **지하철 교통약자 편의 증진**  **팀명 : 함께 편하자? 동행?** | |
| --- | --- |
| 데이터 | ● 교통약자 승하차 현황  [서울교통공사 역별/월별 장애인 승하차인원 정보](https://www.data.go.kr/data/15061870/fileData.do)  [서울교통공사 역별/일별/시간대별 노인 승하차인원 정보](https://www.data.go.kr/data/15101985/fileData.do)  ● 역 내 편의 시설 증설  [서울교통공사 편의시설 현황](https://www.data.go.kr/data/15044443/fileData.do)   * 편의시설: 문화공간, 휠체어리프트, 만남의장소, 환승주차장, 자전거보관소, 엘리베이터, 기차예매역, 무인민원발급기, 환전키오스크, 수유실   [서울교통공사 교통약자 이용시설 승강기 가동현황](https://www.data.go.kr/data/15067770/openapi.do)   * 승강기 구분(ES/EV/WL), 운행상태(사용가능/보수중/점검중) |
| 주요 분석 및 모델링 내용 | 교통약자(노인, 장애인)의 지하철 이용 빈도를 분석하여 교통 편의를 증진하고자 한다.   1. 노인이 많이 이용하는 시간대를 기반으로 역에 도움 인원 배치를 효율화하고자 한다. 2. 노인과 장애인이 많이 이용하는 역을 기반으로 승강기, 휠체어리프트 등의 편의 시설을 우선적으로 증설해야 하는 역을 선정하고자 한다. 3. 장애인이 많이 이용하는 역을 바탕으로 대체 교통 수단으로 셔틀 노선을 제시하고자 한다. |

| **서울시 자치구별 범죄율 분석**  **팀명 : 서울 안전 지킴이** | |
| --- | --- |
| 데이터 | ● 범죄 발생 및 검거  [경찰청 서울특별시경찰청\_경찰서별 5대범죄 발생 검거 현황](https://www.data.go.kr/data/15054738/fileData.do)   * 2000년 ~ 2021년 서울경찰청 관할 경찰서별 살인, 강도, 강간및추행, 절도, 폭력 발생 검거 현황   [범죄발생지 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=135&tblId=TX_13501_A048&vw_cd=MT_ZTITLE&list_id=135_001_002_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)   * 1994년 ~ 2020년 범죄별 시/구별 범죄발생지   ● 경찰 관련  [경찰청 서울특별시경찰청\_경찰서별 경찰관수 현황](https://www.data.go.kr/data/15036460/fileData.do)   * 2011년 ~2021년 서울특별시경찰청 경찰서의 경찰관 인원 현황   [경찰청 서울특별시경찰청\_경찰서별 지역경찰 현황](https://www.data.go.kr/data/3074546/fileData.do)   * 2018년 ~ 2021년 서울특별시 관서별 지역경찰 * 31개 경찰서, 31개 경찰서 관내 파출소, 31개 경찰서 관내 지구대, 31개 경찰서 치안센터   ● 이웃 간 교류  [서울시 반려동물 유무 및 취득 경로 통계](https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do)   * 반려 동물 여부, 취득 경로(동물판매업소, 입양 등)   [서울시 문화공간 (지역문화복지시설) 통계](https://data.seoul.go.kr/dataList/167/S/2/datasetView.do)   * 문예회관, 종합사회복지관, 구민(다목적)체육시설, 청소년수련시설   ● 환경 정비  [서울 빈집 현황](https://data.seoul.go.kr/dataList/11005/S/2/datasetView.do)  [서울시 자치구 연도별 밤범용 cctv 운영 현황](https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do)  [서울시 도로 시설물 통계](https://data.seoul.go.kr/dataList/261/S/2/datasetView.do)   * 2021년 가로등   ● 주민등록인구 [통계청\_SGIS오픈플랫폼\_인구통계](https://www.data.go.kr/data/15088190/openapi.do)   * 인구주택총조사 정보 중 전체적인 인구 통계 정보를 조회 * 요청: 기준년도, 행정기관코드 * 결과: 총인구, 평균나이, 인구밀도, 노령화지수, 노년부양비, 유년부양비, 총가구, 평균가구원수, 총주택   [행정안전부\_행정동별(통반단위) 성/연령별 주민등록 인구수](https://www.data.go.kr/data/15108072/openapi.do)   * 요청: 행정기관코드, 조회시작년월(2022년 10월) * 결과: 통계년월, 시도명, 시군구명, 총인구수, 남자인구수, 여자인구수, 만0~9세남자, 만10~19세남자, …, 만100세이상남자,  만0~9세여자, 만10~19세여자, …, 만100세이상여자   ● 유동 인구  [자치구 단위 서울 생활인구(내국인)](https://data.seoul.go.kr/dataList/OA-15439/S/1/datasetView.do)   * 생활인구: 통신데이터로 특정 시점에 개인이 위치한 지역을 집계한 현주인구 * 시간대 구분 → 상주인구와 주간인구 |
| 주요 분석 및 모델링 내용 | 서울지역 25개 자치구별 범죄율과 관련된 특성을 파악하고자 한다. CCTV, 가로등, 빈집과 같은 환경적 요인, 반려동물, 문화시설과 같은 이웃 간의 관계, 경찰 인력 및 경찰서와 범죄율의 관계를 분석하여 자치구별로 의견을 제시하고자 한다. |

| **코로나 고위험군 분류**  **팀명: 건강 지킴이** | |
| --- | --- |
| 데이터 | ● 코로나19 데이터  [Kaggle COVID-19 Dataset](https://www.kaggle.com/datasets/meirnizri/covid19-dataset)   * 멕시코 코로나 환자 1,048,576명에 대한 데이터 * X: 구분(성별, 연령, 임신), 만성질환(고혈압, 심혈관질환, 당뇨, 만성신부전), 기존 폐질환(폐렴, 만성폐색성폐질환, 천식, 면역억제), 생활습관(비만, 흡연) * y: 환자상태(입원여부, 산소호흡기 사용여부, 중증, 사망)   ● 개인 건강 정보  [국민건강보험공단 건강검진정보](https://www.data.go.kr/data/15007122/fileData.do#tab-layer-openapi)   * 시도(지역), 성별, 연령대, 신장, 체중, 허리둘레, 시력, 청력, 혈압, 혈당, 콜레스테롤, 요단백, 혈색소, 흡연, 음주, 치아 상태 * 수치 데이터 → 기준에 따라 고혈압 O/X 등으로 클래스 구분   ● 지역별 의료기관 현황  [전국의료기관표준데이터](https://www.data.go.kr/data/15096293/standard.do)   * 의사가 환자를 진찰 및 진료하는데 필요한 설비가 갖춰있으며, 30명 이상의 환자를 수용할 수 있는 시설을 갖춘 의료기관(병원) 현황 정보 * 전국 의료기관별 소재지, 의료기관종류, 의료인수, 입원실수, 병상수, 진료과목 등 |
| 주요 분석 및 모델링 내용 | 코로나19 환자의 데이터를 통해 중증 혹은 사망과 관련된 특성을 파악한 뒤, 건강검진정보에 적용하여 고위험군에 대해 대비하고자 한다.   1. 고위험군에 해당하는 개인에게 백신 우선 순위를 주거나 관련 의료 기관에서 지속적으로 관찰하는 등 우선적으로 대비하고자 한다. 2. 지역별로 고위험군으로 분류된 인원과 의료기관 수준을 분석하여, 의료기관이 고위험군 인원에 비해 부족한 지역을 선정하고자 한다. |

| **심장 질환 위험군 분류**  **팀명: 건강 지킴이** | |
| --- | --- |
| 데이터 | ● 심장 질환  [Personal Key Indicators of Heart Disease](https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease)   * 미국질병통제센터 401,958명에게 진행한 설문 조사 300개 정도 문항 중 심장 질환과 관련된 특성을 발췌한 데이터 * X: 구분(성별, 연령대, 인종), 병력(뇌졸중, 당뇨, 천식, 신장병, 피부암), 건강 상태(신체, 정신, 일반, 걷기 어려운 정도), 생활 습관(BMI, 흡연, 음주, 운동, 수면시간) * y: 심장질환 여부   ● 개인 건강 정보  [국민건강보험공단 건강검진정보](https://www.data.go.kr/data/15007122/fileData.do#tab-layer-openapi)   * 시도(지역), 성별, 연령대, 신장, 체중, 허리둘레, 시력, 청력, 혈압, 혈당, 콜레스테롤, 요단백, 혈색소, 흡연, 음주, 치아 상태 * 수치 데이터 → 기준에 따라 고혈압 O/X 등으로 클래스 구분   ● 지역별 의료기관 현황  [전국의료기관표준데이터](https://www.data.go.kr/data/15096293/standard.do)   * 의사가 환자를 진찰 및 진료하는데 필요한 설비가 갖춰있으며, 30명 이상의 환자를 수용할 수 있는 시설을 갖춘 의료기관(병원) 현황 정보 * 전국 의료기관별 소재지, 의료기관종류, 의료인수, 입원실수, 병상수, 진료과목 등 |
| 주요 분석 및 모델링 내용 | 설문 데이터를 분석하여 심장 질환과 관련된 특성을 파악한 뒤, 건강검진정보에 적용하여 심장 질환 위험군에 대해 대비하고자 한다.   1. 고위험군에 해당하는 개인에게 개별 추가 검사를 안내할 수 있으며 지속적으로 관찰을 시행하고자 한다. 2. 지역별로 위험군으로 분류된 인원을 계산하여 위험군으로 분류된 인원과 심장 질환을 진료하는 의료기관 수준을 분석하여, 의료기관이 부족한 지역을 선정하고자 한다. |

# 221214 팀미팅

**질문**

1. 데이터에 따라 연도가 다른 경우, 연도 통일 필요?
   1. 예시: 공공시설물 관련해서 인구분포는 22년, 여가활동 현황은 21년, 박물관 현황은 20년, 공원 현황은 19년
2. 지도 시각화 Python Folium 활용
3. 문화체육관광부에서 진행한 국민여가활동조사 데이터를 활용하고자 하는데, 아래의 문항 중에서 어떤 문항을 활용하는 것이 좋을지

| **문항** | **분류** | **표본수** |
| --- | --- | --- |
| 지난 1년 동안 한번이상 참여한 여가활동 유형(복수응답) | 소분류 | 10,049 |
| 지난 1년 동안 가장 많이 참여한 여가활동 (1순위) | 중분류 | 10,049 |
| 지난 1년 동안 가장 많이 참여한 여가활동 (1+2+3+4+5순위) | 중분류 | 10,049 |
| 지난 1년 동안 지속적으로 참여하는 여가활동 분야 및 활동 | 소분류, 중분류 | 3,540 |

* 중분류(8개): 문화예술관람활동, 문화예술참여활동, 스포츠관람활동, 스포츠참여활동, 관광활동, 취미오락활동, 휴식활동, 사회및기타활동
* 소분류: 문화예술관람활동 → 영화관람, 전시회관람, 박물관관람 등등

**강사님 피드백**

1. 여가 활동 & 공공시설
   1. 분석 방법
      1. 자치구를 동부권, 서부권 등으로 나눠서 대분류
         1. 서울 권역 ([출처](https://urban.seoul.go.kr/view/html/PMNU2030000000))

| 도심권 | 종로구, 중구, 용산구 | 3 |
| --- | --- | --- |
| 동북권 | 성동구, 성북구, 광진구, 강북구, 동대문구, 도봉구, 중랑구, 노원구 | 8 |
| 서북권 | 은평구, 서대문구, 마포구 | 3 |
| 서남권 | 양천구, 영등포구, 강서구, 동작구, 구로구, 관악구, 금천구 | 7 |
| 동남권 | 서초구, 강남구, 송파구, 강동구 | 4 |

* + 1. 지방(5대광역시 등)과 문화공간 현황을 단순 비교(분석은 서울만) → 지방과의 격차가 이러하다라는 것 언급
  1. 데이터
     1. 여가활동 현황: [서울시 데이터](https://data.seoul.go.kr/dataList/10557/S/2/datasetView.do)
        1. 선호, 실제 – 2021년
     2. 시설 데이터는 여가활동 데이터와 같은 연도 or 현재로 해야함 → 기준이 여가활동 데이터
        1. 2021년 시설 데이터 – 서울시 여가 데이터와 맞춰서 시설도 다시 구성
        2. 지방지역의 시설 데이터 – 5대광역시로
     3. 인구는 현재 인구 괜찮음

1. ~~지하철~~
   1. 분석 연습으로는 좋지만 최종까지 갔을 때 효율성이 얼마나 있을지 생각해봐야 한다.
   2. 지금 당장으로는 정치적으로 연관이 되어 있기 때문에 어려울 수 있다.
2. 범죄
   1. 빈집과의 당위성을 보여주기 위한 논문자료가 필요할 수 있다.
   2. 이웃 간 교류, 환경 조성과 범죄 발생 간의 관계에 대한 근거가 필요 – 왜 이 특성을 선정하여 분석했는가
   3. 이웃 간 교류에서 문화시설 현황을 넣었는데, 현황보다는 이용율이 더 좋음
   4. 관련 자료
      1. 범죄 발생과 관련된 특성 도출
3. ~~코로나~~
   1. 코로나 중증은 코로나 지정 병원이 있기 때문에 전체적으로 보는 것이 좋다. (지역과 연결하는건 점염병 특성상 조금 힘들 수 있다.)
   2. 고위험군의 백신 접종은 의사의 진단에 따라 판단이 되기 때문에 의사 진단 데이터가 없으면 어려울 수 있다.
   3. 방향: 현재 시점에서 중증 환자가 어느 정도 될 것 같음. 중증 병상이 전국적으로 몇 개가 있음 → 의료시설 부족? 충분? 등의 의견 제시
   4. 코로나 관련 의료시설 – 음압병동, 최근 코로나 현황
4. 심장병
   1. 지역별로 인구 특성과 연관지을 수 있을 것 같음

**우선순위**

1. 심장병
2. 범죄
3. 여가

# 221214 기획안 수정 방안

**범죄 관련**

**수정 사항**

* 범죄 발생과 관련된 특성과 관련된 자료

| **특성명** | **근거 자료** | **관련 데이터** |
| --- | --- | --- |
| CCTV | [정철우(2016)](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002171600) | [서울시 자치구 연도별 방범용 cctv 운영 현황](https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do) ~ 2022년 |
| 가로등 | [정철우(2016)](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002171600)  [경찰청(2020)](https://www.police.go.kr/user/bbs/BD_selectBbs.do?q_bbsCode=1002&q_bbscttSn=20200108130713336) | [서울시 도로 시설물 통계](https://data.seoul.go.kr/dataList/261/S/2/datasetView.do) 1997년 ~ 2021년 |
| 경찰 순찰 | [정철우(2016)](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002171600) | [경찰청 서울특별시경찰청\_경찰서별 112 순찰차 현황](https://www.data.go.kr/data/15088384/fileData.do) 2021년 |
| 근린 강화 - 유대 활동 | [정철우(2016)](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002171600) | [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do) 2008년 ~ **2018년**  [노인 사회활동 참가](https://data.seoul.go.kr/dataList/47/S/2/datasetView.do) 2002년 ~ 2016년  [주로 이용하는 운동 장소](https://data.seoul.go.kr/dataList/10554/S/2/datasetView.do) 2012년 ~ 2019년 |
| 시민 참여 | [여개명(2021)](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=69786938a803f3ebffe0bdc3ef48d419) | [서울시 투표현황 통계](https://data.seoul.go.kr/dataList/285/S/2/datasetView.do) 1997년 ~ 2020년 각 선거별  [서울시 장기등 기증희망등록자 현황 통계](https://data.seoul.go.kr/dataList/10928/S/2/datasetView.do) 2011년 ~ 2019년  [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do) 2008년 ~ 2018년  [서울시 자원봉사활동 경험여부 및 참여 횟수 통계](https://data.seoul.go.kr/dataList/284/S/2/datasetView.do) 2011년 ~ 2021년 |
| 빈집 | [염윤호(2019)](https://doi.org/10.36889/KCR.2019.12.31.4.139) | [서울 빈집 현황](https://data.seoul.go.kr/dataList/11005/S/2/datasetView.do) 2000년 ~ 2021년 |
| 공원밀도 | [오미진(2011)](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=030f7df246982962ffe0bdc3ef48d419) | [서울시 공원 (공원율) 통계> 데이터셋> 공공데이터 | 서울열린데이터광장](https://data.seoul.go.kr/dataList/10817/S/2/datasetView.do)  2010년 ~ 2021년 |
| 1인 가구 | 장진희(2018) | [서울시 1인가구(연령별) 통계](https://data.seoul.go.kr/dataList/10995/S/2/datasetView.do) 2010년 ~ 2021년 |
| 반려동물 수 | [Pinchak et al (2022)](https://www.google.com/url?q=https://doi.org/10.1093/sf/soac059&sa=D&source=docs&ust=1671001359944021&usg=AOvVaw0x14p8Sd40asSAET4MRWtM) | [서울시 반려동물 유무 및 취득 경로 통계](https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do) 2014년 ~ 2021년 |
| 실업율 | [권성은](http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=6ba54291892aeb15ffe0bdc3ef48d419&outLink=N) | [서울시 고용지표 통계](https://data.seoul.go.kr/dataList/59/S/2/datasetView.do) 1999년 ~ 2022년 |
| 1인당 GRDP | `` | [서울시 자치구별 1인당 지역내총생산 및 수준지수 통계](https://data.seoul.go.kr/dataList/11063/S/2/datasetView.do) 2015년 ~ 2019년 |

수정 필요 데이터 (20년도까지 찾아보고 없으면 18년도), 자치구별

* 1인당 GRDP
  + [1인당 GRDP(시도) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1C86_02&vw_cd=MT_GTITLE01&list_id=109&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=E1) 85-20년도 (시도로만 나뉘어 있음ㅠ)
  + [1. GRDP(시/군/구) (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1C65_03E&vw_cd=MT_GTITLE01&list_id=109&scrId=&seqNo=&lang_mode=ko&obj_var_id=&itm_id=&conn_path=E1) 15-19년도 (시군구 나뉘어져 있음!)
  + [서울 구별 세금](https://kosis.kr/statHtml/statHtml.do?)
* 단체 활동 관련 데이터
  + [최근 1년 간 참여 경험 - 사회활동](https://data.seoul.go.kr/dataList/DT201013D029/S/2/datasetView.do) 2020년도, 5권역별
    - 반상회/부녀회/주민회의, 지역봉사활동, 마을공동체사업, 시민운동단체/사회단체활동, 선거 때 투표 참여
    - 조사 표본은 서울시 5개권역(도심권, 동북권, 서북권, 서남권, 동남권)을 기준으로 비례배분되었기 때문에 제공자료의 행정구역은 5개 권역만 제공됨
  + [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do) 2008년 ~ **2018년**
* 반려동물 시설 혹은 산책 빈도

**특성 정리**

● 환경

* CCTV [서울시 자치구 연도별 방범용 cctv 운영 현황](https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do)
* 가로등 [서울시 도로 시설물 통계](https://data.seoul.go.kr/dataList/261/S/2/datasetView.do)
* 공원율 [공원 밀도 서울시 공원(공원율) 통계](https://data.seoul.go.kr/dataList/10817/S/2/datasetView.do)

● 이웃 간 관계

* 유대 활동 [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do)
* 반려동물 [서울시 반려동물 유무 및 취득 경로 통계](https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do)

● 경제적 요인

* 실업율 [서울시 고용지표 통계](https://data.seoul.go.kr/dataList/59/S/2/datasetView.do)
* 1인당 GRDP [서울시 자치구별 1인당 지역내총생산 및 수준지수 통계](https://data.seoul.go.kr/dataList/11063/S/2/datasetView.do)

● 인구 특성

* 주민등록인구 [행정안전부\_행정동별(통반단위) 성/연령별 주민등록 인구수](https://www.data.go.kr/data/15108072/openapi.do)
* 생활인구 → 상주인구와 주간인구 [자치구 단위 서울 생활인구(내국인)](https://data.seoul.go.kr/dataList/OA-15439/S/1/datasetView.do)
* 1인 가구 [서울시 1인가구(연령별) 통계](https://data.seoul.go.kr/dataList/10995/S/2/datasetView.do)

● 경찰 관련

* 경찰관 수 [경찰청 서울특별시경찰청\_경찰서별 경찰관수 현황](https://www.data.go.kr/data/15036460/fileData.do)
* 지역경찰 현황 [경찰청 서울특별시경찰청\_경찰서별 지역경찰 현황](https://www.data.go.kr/data/3074546/fileData.do)
* 순찰차 현황 [경찰청 서울특별시경찰청\_경찰서별 112 순찰차 현황](https://www.data.go.kr/data/15088384/fileData.do)

● 범죄 현황

* 검거 현황 [경찰청 서울특별시경찰청\_경찰서별 5대범죄 발생 검거 현황](https://www.data.go.kr/data/15054738/fileData.do)
* 발생 현황 [범죄발생지 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=135&tblId=TX_13501_A048&vw_cd=MT_ZTITLE&list_id=135_001_002_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)

[의견]

* 인구 특성과 경제적 요인을 자치구의 특성으로 묶기 – 인구 특성과 경제적 요인은 의견 제시하기 애매… 인구랑 경제를 인위적으로 제어하기는 어렵기 때문.
* 환경, 이웃 간 관계, 경찰 관련 특성을 기반으로 최종 의견 내기 – 이 지역에 경찰서가 더 필요하다. 이런 환경 조성이 필요하다 등등
  + 즉, 중점으로 봐야 할 분석해야 할 사항은 이 세 변수
  + 위에 인구 특성이랑 경제적 요인은 범죄 현황과의 상관관계를 보는 정도만

여가활동

**수정 사항**

* 여가활동 설문 데이터
* 여가시설 데이터 → 서울, 지방

**서울 권역 (**[**출처**](https://urban.seoul.go.kr/view/html/PMNU2030000000)**) - 권역별로 분류**

| 도심권 | 종로구, 중구, 용산구 | 3 |
| --- | --- | --- |
| 동북권 | 성동구, 성북구, 광진구, 강북구, 동대문구, 도봉구, 중랑구, 노원구 | 8 |
| 서북권 | 은평구, 서대문구, 마포구 | 3 |
| 서남권 | 양천구, 영등포구, 강서구, 동작구, 구로구, 관악구, 금천구 | 7 |
| 동남권 | 서초구, 강남구, 송파구, 강동구 | 4 |

**여가활동 현황**

[서울시 문화활동의 어려움](https://data.seoul.go.kr/dataList/10555/S/2/datasetView.do) (12-16년도) → 서론에 이용

| [서울시 하고 싶은 여가활동 통계](https://data.seoul.go.kr/dataList/10557/S/2/datasetView.do)  자치구별, 성별, 연령대 | 주민등록인구통계  2012년 ~ 2021년 | 여행야외나들이, 영상시청, 문화예술관람, 운동, 휴식, 종교활동, 문화예술참여활동, 컴퓨터게임인터넷검색 등, 운동경기관람, 사회봉사활동, 독서 |
| --- | --- | --- |
| [서울시 주말/휴일 여가활동](https://data.seoul.go.kr/dataList/10260/S/2/datasetView.do)  자치구별, 성별, 연령대 | 서울특별시, 서울시도시정책지표조사 | TV 또는 비디오시청, 여행야외나들이, 컴퓨터게임인터넷검색 등, 휴식, 사회활동, 운동, 문화예술관람, 창작적취미활동, 운동경기관람, 종교활동, 독서 |
| [최근 1년 간 가장 많이 참여한 여가활동](https://data.seoul.go.kr/dataList/DT201013D025/S/2/datasetView.do)  5권역별, 성별, 연령대  1순위, 2순위, 1+2순위 | 서울특별시복지실태조사  2020년 | 휴식 활동, 취미/오락활동, 관광 여행, 문화예술관람활동, 사회봉사 및 단체 활동 참여, 스포츠 관람 활동, 문화예술 참여 활동, 스포츠 참여 활동 |

* **2020년 기준으로 합치기**
* 열 병합 과정

| **새 분류** | **하고 싶은 여가활동** | **주말/휴일** | **최근 1년 간** |
| --- | --- | --- | --- |
| 문화예술 관람 및 참여 | 문화예술관람, 문화예술참여활동 | 문화예술관람, 창작적취미활동 | 문화예술관람활동, 문화예술 참여 활동 |
| 스포츠 관람 및 참여 | 운동경기관람, 운동 | 운동, 운동경기관람 | 스포츠 관람 활동, 스포츠 참여 활동 |
| 사회봉사 및 단체활동 | 사회봉사활동 | 사회활동 | 사회봉사 및 단체 활동 참여 |
| 종교활동 | 종교활동 | 종교활동 |  |

* 연령대

| **하고 싶은 여가활동** | **주말/휴일** | **최근 1년 간** |
| --- | --- | --- |
| * 10대 * 20대 * 30대 * 40대 * 50대 * 60대 이상 | * 10대 * 20대 * 30대 * 40대 * 50대 * 60대 이상 | * 20대 이하 * 30~40대 * 50~60대 * 70대 이상 |

⇒ 멘토링

* 여가활동 현황 데이터 어떤 것으로?

**집 주변 여가활동 분류**

* 문화예술 및 참여
* 스포츠 및 참여
* 사회봉사 및 단체활동

**여가시설 데이터**

**무조건 2020년!!!!!!!!!!! → 안되면 21년…**

| **데이터 이름** | **서울 or 지방** | **연도** | **여가활동 분류** |
| --- | --- | --- | --- |
| [문화공간(지역문화복지시설)](https://data.seoul.go.kr/dataList/167/S/2/datasetView.do)  [문화공간(전시시설)](https://data.seoul.go.kr/dataList/166/S/2/datasetView.do)  [청소년 수련시설](https://data.seoul.go.kr/dataList/392/S/2/datasetView.do) | 서울 | 93 - 21 | 문예회관 → 문화예술 참여  종합사회복지관 → 사회봉사 및 단체활동  구민(다목적)체육시설 → 스포츠 참여  청소년수련시설 → 사회봉사 및 단체활동  박물관, 미술관 → 문화예술 관람 |
| [서울시 도서관 현황 통계](https://data.seoul.go.kr/dataList/387/S/2/datasetView.do) | 서울 | 99 - 21 | 도서관 => 문화예술 관람 |
| [서울시 공공체육시설](https://data.seoul.go.kr/dataList/390/S/2/datasetView.do) | 서울 | 95 - 21 | → 스포츠 참여 |
| [영화관](https://data.seoul.go.kr/dataList/165/S/2/datasetView.do) | 서울 | 93 - 21 | 영화관 -> 문화예술관람 |
| [국도교통부\_도시계획통계](http://www.eum.go.kr/web/cp/st/stUpisStatDet.jsp) | 전국(시군구 단위) | 05 - 21 | 10-1-5. 문화시설 > 공연장, 박물관, 미술관, 과학관, 문화산업단지, 문화산업진흥시설, 지방문화원, 문화예술진흥시설, 전시시설, 공공도서관, 전문도서관 → 문화예술  10-1-6. 사회복지시설 > 계(국가, 지방자치단체, 사회복지법인, 비영리법인) → 사회봉사 및 단체활동  10-1-8. 청소년수련시설 > 계 → 사회봉사 및 단체활동  10-1-9. 체육시설 > 공공체육시설, 골프장(9홀 이상), 스키장, 종합운동장 → 스포츠 |
| [서울시 문화공간 (공연장) 통계](https://data.seoul.go.kr/dataList/164/S/2/datasetView.do) | 서울 | (93-21) | 공연장 -> 문화예술 관람 |

* [주말, 휴일 여가활동](https://data.seoul.go.kr/dataList/10260/S/2/datasetView.do)
* [종교시설 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=408&tblId=DT_408003_A006&vw_cd=MT_ZTITLE&list_id=U_14_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE) 18-21년도

[국도교통부\_도시계획통계](http://www.eum.go.kr/web/cp/st/stUpisStatDet.jsp)

[종교시설 (kosis.kr)](https://kosis.kr/statHtml/statHtml.do?orgId=408&tblId=DT_408003_A006&vw_cd=MT_ZTITLE&list_id=U_14_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)

# 221216 질문

* 희망하는 여가활동과 실제 하고 있는 여가활동 데이터를 결합하여 사용하고자 함. 실제 하고 있는 여가활동 데이터로 “주말/휴일 여가활동”(1개 선택)과 “최근 1년 간 가장 많이 참여한 여가활동”(2개 선택)이 있음. 후자가 실제 여가활동 현황을 더 많이 대표할 것으로 예상되긴 하지만, 희망하는 여가활동 데이터와 연령대가 맞지 않음.

| **하고 싶은 여가활동** | **주말/휴일** | **최근 1년 간 가장 많이** |
| --- | --- | --- |
| * 10대 * 20대 * 30대 * 40대 * 50대 * 60대 이상 | * 10대 * 20대 * 30대 * 40대 * 50대 * 60대 이상 | * 20대 이하 * 30~40대 * 50대 이상 * 50~60대 * 70대 이상 |

* 반려동물을 소유와 사회단체활동에 참여를 활용하고자 함. 찾은 데이터는 반려동물 유무, 사회단체활동 참여도가 표본 중에서 몇 퍼센트인지로 나타남. 해당 데이터를 **그 당시 가구 수에 곱해서** 구해서 활용해도 괜찮을지 아니면 이 비율 자체를 활용하는 것이 좋을지

| **자료 출처** | **자치구별 표본배분 방법** |
| --- | --- |
| 서울시도시정책지표조사 | 최소 400표본  제곱근 비례배분 |

* 서울 자치구별 실업률이 2021년도부터 산정되었음. 실업률만 현재로 두고 해도 괜찮을지

| 자치구별 실업률 | 2021년 ~ |
| --- | --- |
| 1인당 GDRP | ~ 2019년까지 |
| 사회단체활동 참여도 | ~ 2018년까지 |

* 2021년으로

→ 방향성 위주로

# 221217 멘토링 정리

| **심장 질환 위험군 분류**  **팀명: 건강 지킴이** | |
| --- | --- |
| 데이터 | ● 심장 질환  [Personal Key Indicators of Heart Disease](https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease)  ● 개인 건강 정보  [국민건강보험공단 건강검진정보](https://www.data.go.kr/data/15007122/fileData.do#tab-layer-openapi)  ● 지역별 의료기관 현황  [보건의료빅데이터개방시스템 의료통계정보](https://opendata.hira.or.kr/op/opc/selectMedInfoSvcList.do) |
| 주요 분석 및 모델링 내용 | ● 심장 질환 분류 모델   1. 심장 질환이 일어나는 특성 파악 2. 1의 결과를 바탕으로 심장 질환 분류 모델 작성   ● 시도별 분석   1. 건강검진정보를 심장 질환 분류에 적용    1. 개인별로 심장 질환 위험군을 분류한다    2. 지역별로 심장 질환 위험군으로 분류된 사람의 수를 도출한다 2. 의료 수준 분석    1. 1-b를 바탕으로 지역별 의료 수준과 심장 질환 위험군 수를 분석한다    2. a를 토대로 의료 수준과 관련된 의견을 제시한다 |

| **건강검진** | **변환 기준** | **Kaggle** |
| --- | --- | --- |
| 성별 |  | 성별 |
| 연령대 코드(5세 단위) |  | 연령 |
| 신장(5Cm 단위) |  | BMI |
| 체중(5Kg 단위) |
| 시력(좌) |  | BLIND |
| 시력(우) |
| 청력(좌) |  | DEAF |
| 청력(우) |
| 수축기 혈압 |  | 고혈압 |
| 이완기 혈압 |
| 식전혈당(공복혈당) |  | 당뇨 |
| 총 콜레스테롤 |  | 콜레스테롤 |
| 요단백 |  | 신장질환 |
| 혈청크레아티닌 |
| (혈청지오티) AST |  | 간 |
| (혈청지오티) ALT |
| 감마 지티피 |
| 흡연상태 |  | Smoking |
| 음주여부 |  | Alcohol |

* 미국 – 우리나라
  + 성별, 연령, 당뇨병, 고혈압, 콜레스테롤, BMI, 간, 신장, 흡연, 음주 → X
* X가 안 맞는 부분
  + 얼추 맞쳐뒀음
* 왜?라는 전제를 깔고 가는 게 좋다
  + 우리나라 사망율의 몇 퍼센트로 굉장히 높다 등등
* 예측 방향으로 가야 하지 않나…
  + 처음에 위험군으로 분류된 사람에게 미리 안내하거나 검사를 요청하는 것도 생각했었는데 이게 여기에 해당할까?
  + 응급 센터, ~~심장제세동기~~
  + ~~관련 교육 프로그램(추운 날 조심하라 등등????)~~
* ~~계절적 요인~~
  + ~~기온 차, 시간대 → 응급센터 오는 사람이 얼마냐…~~
* ~~안전 지수 등등 → LinearRegression 계수?? 예측 확률~~
* 건강검진정보 + 코시스 심장질환인구분포도 추가해라
  + [국민건강보험공단\_심장판막질환의 진료년월별 연령별 진단건수 및 진단인원](https://www.data.go.kr/data/15106272/fileData.do)

| [기관수현황](http://opendata.hira.or.kr/op/opc/olapYadmStatInfo.do) |
| --- |
| 지역별 종별 요양기관수 → 상급종합병원, 종합병원, 병원 – 연도/분기 |
| 지역별 의원 표시과목별 기관수 → 심장혈관흉부외과 |
| 지역별 종별 의료인력Ⅰ(의사 등) → 일반의, 전문의 |
| 지역별 종별 의료인력Ⅱ(간호사 등) → 간호사, 임상병리사 |
| 지역별 종별 전문과목별 전문의수 → 심장혈관흉부외과 |
| [시설 및 장비 현황](http://opendata.hira.or.kr/op/opc/olapInfraEquipmentStatInfo.do) → 상급종합병원, 종합병원, 병원 |
| 입원실(일반입원실 개수, 중환자실 개수) → 상급종합병원, 종합병원, 병원 |

[전국응급의료기관표준데이터](https://www.data.go.kr/data/15096291/standard.do)

[민간구급차 검색 | E-GEN통합홈페이지](https://www.e-gen.or.kr/egen/search_commercial_ambulance.do)

[종합상황판 (nemc.or.kr)](https://portal.nemc.or.kr:444/medi_info/dashboards/dash_total_emer_org_popup_for_egen.do)

[데이터 상세 | 공공데이터포털 (data.go.kr)](https://www.data.go.kr/data/15065413/fileData.do)(2018기준)

[구급차 수 및 배치된 응급구조사 수(시도별)](https://kosis.kr/statHtml/statHtml.do?orgId=411&tblId=DT_41104_422&vw_cd=MT_ZTITLE&list_id=D1_41101_04&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)(2014~2020)

[119 구급 운영 현황(2004~2021)](https://kosis.kr/statHtml/statHtml.do?orgId=156&tblId=TX_15602_A018&vw_cd=MT_ZTITLE&list_id=156_15602_A001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)

[119 구급대 현황](https://kosis.kr/statHtml/statHtml.do?orgId=156&tblId=TX_15602_A002&vw_cd=MT_ZTITLE&list_id=156_15602_A001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)(2002~2021)

119 구급대 현황

보유구급차 (대) 운영구급차 (대) 구급대원 (명)

소계 소계 소계

1: 윤석용 멘토님

| 심장 질환은 미국 데이터 → 우리나라의 지역적 데이터를 반영하지 못함  → 심장 질환에 있어 우리나라와 미국이 동일한 특성을 갖는다? 지역이 상관 없다? 등의 자료가 있으면 그대로 사용 가능  X가 안 맞는 부분이 있음 → 갭을 맞추는 것이 필요  ⇒ 이 작업 진행 중이었음  예측과 학습은 동일한 데이터의 사용을 전제로 함  의료 수준 분석 → EDA 군집분석 |
| --- |
| 심장지로한은 미국 데이터로 우리나라의 특성을 반영하지 못한다.  독립변수를 다 맞춰줘야 하는데 안 맞는 부분이 있을 수도 있다. 얼마까지의 독립변수를 쓸 수 있는지가 사전에 검토되어야 한다.  예측과 학습은 동일한 데이터의 활용을 전제로 한다. 모델을 만든다는 것은 무슨 말이냐하면 데이터가 갖고있는 특징을 찾아낸다는 말이 들어가있는 것이다. 미국의 케글 데이터와 우리나라는 분명 차이가 있을 것이다. 이걸 가지고 도전하려면 학술논문, 참고도서를 인용해서 이 문제를 어떻게 해결할건지 찾아야 한다. 이걸 해결하지 못하면 분명 지적이 나올 것. |

2: 최인호 멘토님

| 예방/예측 방면으로 결론이 나는 게 좋음 → 응급센터와의 연계  ⇒ 처음에 위험군으로 분류된 사람에게 미리 안내하거나 검사를 요청하는 것도 생각했었는데 이게 여기에 해당할까?  전제를 더 깔아야 함 – 심장 질환 늘어나고 있음 등 왜 심장 질환을 예측하는지  현재는 모델 + 건강검진결과를 활용하는데, 심장질환관련인구분포 데이터를 추가하면 더 좋을 것 같음  예방/예측을 통한 이 시점에 어느 지역에 발생할 가능성이 높다 |
| --- |
| 심장질환은 예측이 중요한데 지금은 치료목적에 맞춰져있는 것 같음.  심장질환은 계졀적 요인이 크다. 계절적인 온도차도 생각하면 좋다. |

3: 이정숙 멘토님

| 결론이 조금 더 명확해야 함 → 의료 안전 지수, 심장 질환 안전 지수 등등  배경 → 왜?와 같은 당위성 필요 |
| --- |
| 뭘 도출할건지가 명확해야된다. 심장질환자 위험군을 분류했다면 그걸 가지고 뭘 어떻게 할 것인지를 선정해야 한다.  뭘 하던지 분석을 할 때 ‘왜’가 앞부분에서 당위성을 확보하지 못하면 공감대 형성이 안 된다. 그 부분을 먼저 잡아줘야 한다. |

| **서울시 자치구별 범죄율 분석**  **팀명 : 서울 안전 지킴이** | |
| --- | --- |
| 데이터  기준: 2021년 | ● 범죄 현황  [서울시 5대 범죄 발생현황 통계](https://data.seoul.go.kr/dataList/316/S/2/datasetView.do)  ● 인구 분포  [서울시 주민등록인구(구별) 통계](https://data.seoul.go.kr/dataList/419/S/2/datasetView.do)  [자치구 단위 서울 생활인구(내국인)](https://data.seoul.go.kr/dataList/OA-15439/S/1/datasetView.do)  [서울시 1인가구(연령별) 통계](https://data.seoul.go.kr/dataList/10995/S/2/datasetView.do)  ● 경제적 특성  [시군구 경제활동인구 총괄](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1ES3A01S&vw_cd=MT_ZTITLE&list_id=B19_EQ13_1&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)  ● 환경  [서울시 자치구 연도별 방범용 cctv 운영 현황](https://data.seoul.go.kr/dataList/OA-21097/F/1/datasetView.do)  [서울시 도로 시설물 통계](https://data.seoul.go.kr/dataList/261/S/2/datasetView.do)  [공원 밀도 서울시 공원(공원율) 통계](https://data.seoul.go.kr/dataList/10817/S/2/datasetView.do)  ● 이웃 간 관계  [서울시 사회단체활동 참여도 통계](https://data.seoul.go.kr/dataList/10300/S/2/datasetView.do)  [서울시 반려동물 유무 및 취득 경로 통계](https://data.seoul.go.kr/dataList/10970/S/2/datasetView.do)  ● 경찰 관련  [경찰청 서울특별시경찰청\_경찰서별 경찰관수 현황](https://www.data.go.kr/data/15036460/fileData.do)  [경찰청 서울특별시경찰청\_경찰서별 지역경찰 현황](https://www.data.go.kr/data/3074546/fileData.do) |
| 주요 분석 및 모델링 내용 | ● 범죄와 관련 변수 상관관계 분석   1. 자치구 특성    1. 범죄 현황과 관련 있는 특성을 파악한다 2. 환경, 이웃 간 관계, 경찰력    1. 범죄 현황과 관련 있는 특성을 파악한다    2. 1의 결과와 상관관계가 있는 특성에 대해서는 정규화를 진행한다    3. a와 b를 토대로 자치구에서 범죄를 줄일 수 있는 의견을 제시한다 |

* 범죄 범위 줄여라
  + 최근 증가폭이 가장 높은 것 → 얘가 요즘 중요하다
* 변수
  + 추가
    - 신기한 변수
    - 외국인 관련
  + 줄여라!
    - 다 따로 논다 – CPTED 환경 조성 이런 느낌으로
* 기존 이론을 증명해라..
  + CPTED!
* ~~예측 모형, 지수 등등~~
* 25개 다 하지 말고, 발생 건수가 높은 곳과 낮은 곳 뽑아라(높은 곳과 낮은곳 특징)
  + 25개 전부 + 높은 곳 vs. 낮은 곳
* 왜 범죄 현황을 분석하는가?

1: 윤석용 멘토님

| 자치구는 큰 범위 → 행정동?  더 많은 변수 필요 – 특히, 신기한 변수(예: 지하철 입구 방향)  예측 모형을 제시하는 것도 좋음(예: 이런저런 변수를 바탕으로 범죄가 일어날 확률 계산)  경찰청 범죄 백서 참고하면 좋을 듯  깨진 유리창 이론 등 기존 이론을 증명하는 방식도 고려  5대 범죄 말고 특정 범죄에 대한 변수를 탐색하는 것도 추천 |
| --- |
| 범죄관련 데이터분석은 하다보면 읍면동으로 깊게 내려가고 싶어진다. 같은 구 안에서도 차이가 크기 때문이다. 멘토님이 옛날에 해보니 지하철 출구별로도 범죄율이 다르더라.  그리고 정책제안을 하는 것은 남들이 보지 않았던 것을 찾아서 내 주장을 데이터로 증명하는 것이다.  깨진 유리창 이론 같은 관련된 이론들을 찾아서 증명하는 쪽도 생각해보자.  아이템을 너무 크게 보지말고 성범죄만 하겠다던지 절도만 하겠다던지 이렇게 확 줄여서 하는 것도 추천한다. |

2: 최인호 멘토님

| 가장 큰 변수 중 하나가 외국인 거주자 혹은 외국인 노동자 수  하나의 범죄를 깊게... -> 표본이 높은 것으로(증가폭이 가장 높은 범죄를 선정했다는 전제)  25개 자치구 중에서 범죄 발생이 많은 곳과 적은 곳의 특징 비교 -- 25개 전체를 돌리면, 특징이 잘 안 나타날 수 있음 |
| --- |
| 외국인 노동자 수가 중요하다. 외국인 노동자 수가 많은 안산, 구로쪽에 범죄가 많이 발생한다. 상관관계가 생각보다 높게 나타난다. 외국인 노동자수에 대한 데이터가 추가적으로 고려가 됐으면 좋겠다.  5대 범죄를 다 갖고가는 것도 방법이지만 한 가지 범죄만 갖고 깊게 가는 것도 방법이다.  범죄 발생이 많은 곳과 적은 곳의 적은 곳의 특징을 분석해서 상관관계를 보여줬으면 좋겠다. |

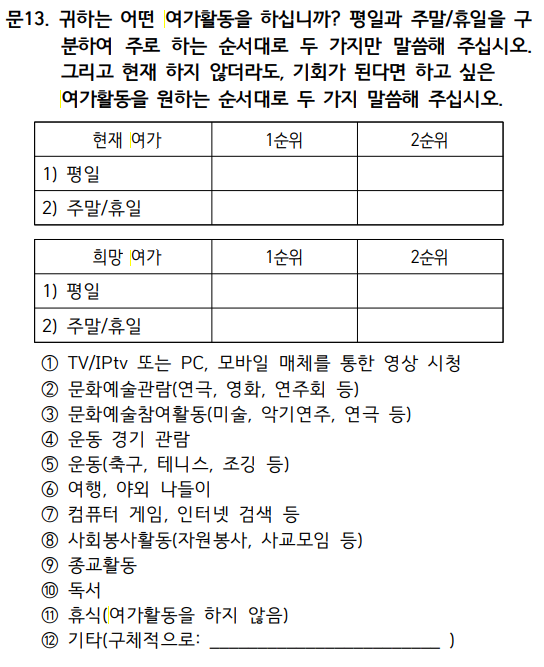
3: 이정숙 멘토님

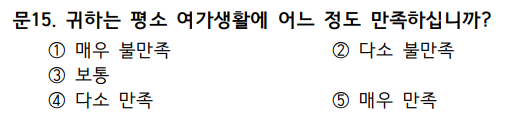
| 범죄 범위가 너무 넓다  결론 예시: CCTV 최적지 |
| --- |
| 이걸 왜 하려고 했는지  5대 범죄라 하면 너무 넓다.  셉티드라는 게 있다. 범죄가 일어나지 않을(범죄를 일으키지 못할) 환경을 조성하는 거다. 달동네도 가로등을 설치하고 벽화를 그려서 환경조성을 가져간다. |

| **서울시 자치구별 공공시설물 수요 예측**  **팀명 : 우리동네 여유 지킴이** | |
| --- | --- |
| 데이터  기준: 2020년 | ● 인구 분포  [행정안전부\_행정동별(통반단위) 성/연령별 주민등록 인구수](https://www.data.go.kr/data/15108072/openapi.do)  ● 여가활동  [서울시 하고 싶은 여가활동 통계](https://data.seoul.go.kr/dataList/10557/S/2/datasetView.do)  [최근 1년 간 가장 많이 참여한 여가활동](https://data.seoul.go.kr/dataList/DT201013D025/S/2/datasetView.do)  ● 여가 시설  [국도교통부\_도시계획통계](http://www.eum.go.kr/web/cp/st/stUpisStatDet.jsp) |
| 주요 분석 및 모델링 내용 | ● 서울 권역별 분석   1. 여가활동과 인구 분포의 상관관계 분석    1. 성별/연령대별로 자주 하는 여가활동과 원하는 여가활동이 어떤 것인지 파악한다    2. 해당 권역에서 자주 하는 여가활동과 원하는 여가활동이 어떤 것인지 파악한다 2. 여가시설 분포 조사 3. 1-b에서 조사한 결과를 바탕으로 여가활동 수요와 여가활동의 관계를 파악한다 4. a를 토대로 해당 권역에 필요한 여가시설을 제안한다   ● 광역시 분석   1. 여가시설 분포와 인구 분포 조사    1. 서울 권역별 분석 1의 결과를 바탕으로 해당 광역시의 여가활동 수요를 예측한다    2. a를 토대로 해당 광역시에 필요한 여가시설을 제안한다 |

* 생활인구 이용해라! [자치구 단위 서울 생활인구(내국인)](https://data.seoul.go.kr/dataList/OA-15439/S/1/datasetView.do)
  + 주간생활인구/야간생활인구
  + 주중/주말
* 서울만!
* 왜 이 주제?
  + 코로나 유행하면서 멀리 못 간다 – 근처 여가시설을 알아야 한다?

| 문13-1-1. 현재 여가 활용 유형\_평일(1순위)  문13-1-1. 현재 여가 활용 유형\_평일(2순위)  문13-1-2. 현재 여가 활용 유형\_주말(1순위)  문13-1-2. 현재 여가 활용 유형\_주말(2순위)  문13-2-1. 희망 여가 활용 유형\_평일(1순위)  문13-2-1. 희망 여가 활용 유형\_평일(2순위)  문13-2-2. 희망 여가 활용 유형\_주말(1순위)  문13-2-2. 희망 여가 활용 유형\_주말(2순위)  문15. 여가활동 만족도 | 1.TV/IPtv 또는 PC, 모바일 매체를 통한 영상 시청  2.문화예술관람  3.문화예술참여활동  4.운동경기 관람  5.운동  6.여행, 야외 나들이  7.컴퓨터 게임, 인터넷 검색 등  8.사회봉사활동  9.종교활동  10.독서  11.휴식  12.기타  21.사우나/찜질방  22.쇼핑  23.산책  24.텃밭농사  25.반려동물과 시간보내기  26.자기계발  27.레져활동(낚시 등)  28.친목 도모  29.만들기  30.명상  31.음악감상  32.웹툰/만화보기  33.맛집탐방  34.블로그 활동  35.취미활동 |
| --- | --- |





1: 윤석용 멘토님

| 정책 제안 → 뻔한 이야기  EDA – 다른 데이터 많이 필요 → 인사이트 도출이 필요 |
| --- |
| 아침에 멘토링 시간에 EDA라는 과정을 말했었는데 그게 대단히 중요하다. 여가라는 이 부분을 데이터를 갖고 다양한 입체적인 탐색적 데이터분석을 많이 해줘야 한다.  3개 중에 완벽한 것은 없다. KT관련된 멘토님이 있는데 부탁해서 KT와 관련된 데이터를 받을 수도 있을 것이니 물어보라. |

2: 최인호 멘토님

| 여가 활동 현황 관련 데이터를 SNS 데이터로(인스타그램 해시태그 등등)  광역시와 시군을 같이 비교 |
| --- |
|  |

3: 이정숙 멘토님

| 주민등록인구 대신 생활인구를 활용하는 것은? |
| --- |
| 생활인구, 관심인구 |

# 221220 전제

[사망원인 2위 '심장병'… 심장 위협하는 4대 요인 - 당신의 건강가이드 헬스조선 (chosun.com)](https://health.chosun.com/site/data/html_dir/2020/12/11/2020121101784.html)

[매년 뇌경색 환자 2% ‘심근경색’…“AI로 심장병 찾아낸다” (kbs.co.kr)](https://news.kbs.co.kr/news/view.do?ncd=5521938)

[전 세계 사망원인 1위, 국내 사망원인 2위 심혈관질환](http://keydr.kr/mobile/article.html?no=3782)

[심장질환 증상과 심혈관 건강의 중요성](https://m.blog.naver.com/mrrah/221403794663)

["사망률 1위 심혈관 질환, 무엇보다 예방이 중요"](http://www.dailypharm.com/Users/News/NewsView.html?ID=274837)

[사망원인(104항목)/성/시도별 사망자수](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B34E09&vw_cd=MT_ZTITLE&list_id=F_27&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)

# 참조) 성공적인 프로젝트 위한 17가지 질문들

(이 중 저희 주제에 맞는 질문들만 뽑아서 데이터 찾아봐요!)

**사업에 대한 이해 Business Understanding**

1. 머신러닝 시스템으로 얻고자 하는 것은 무엇인가요?

2. 개발하고자 하는 어플리케이션의 정확한 시나리오는 무엇이며, 어떤 비즈니스 임팩트를 기대하나요?

3. 머신러닝을 활용했을 때 예상되는 성능과 한계를 제대로 이해하고 있나요?

4. 머신러닝 모델의 성능을 어떻게 모니터링하고 측정할 계획인가요?

**모델링 Modeling**

5. 어떤 머신러닝 모델을 활용할 계획인가요?

6. 컴퓨팅 속도(inference speed), 정확도(accuracy), 정밀도(precision), 재현율(recall) 면에서 성능을 담보할 수 있는 필수 조건(requirements)은 무엇인가요?

7. 학습을 위한 필수 조건은 무엇인가요? 클라우드 컴퓨팅 서버에 의존할 계획인가요? 모델은 지속적으로 업데이트 되어야 하나요?

**데이터 확보 & 이해 Data Acquisition & Understanding**

8. 모델을 훈련시키기 위한 충분한 데이터를 확보하였나요? 그렇지 않다면, 추가적인 데이터를 어떻게 수집할 것인가요? 크라우드 소싱? 웹 크롤링? 또는 미리 만들어진 데이터를 구매할 예정이신가요?

9. 데이터 소스를 법적으로 어디까지 활용할 수 있나요? 데이터에 대한 권리를 보유했나요?

10. 데이터증강 기술을 사용할 예정인가요?

11. 서비스 유저들이 만들어 내는 데이터를 활용할 수 있나요?

**데이터 라벨링 Data Labeling**

12. 누가 데이터 라벨링 작업을 진행하나요? 회사 내부에 인하우스(in-house) 데이터 라벨링 팀이 있나요? 외부의 데이터 라벨링 업체를 용역으로 활용하나요?

13. 어떤 데이터 라벨링 도구를 사용할 예정인가요? 활용하고자 하는 라벨링 도구나 시각화, 통계, 버전 관리, 팀 협업 등 원하는 기능을 제공하고 있나요?

14. 라벨링 속도를 높여줄 수 있는 이미 학습된 모델을 사용할 예정인가요? 그렇다면, 그 모델을 계속 더 훈련시킬 수 있나요? 얼마나 자주 이 모델을 추가적으로 훈련 시킬 수 있나요?

**모델 훈련 Model Training**

15. 데이터 세트를 구축한 다음에, 모델을 훈련시키기 위해서는 어떤 인프라를 사용하나요? 내부에 GPU 서버를 사용하나요? Cloud를 서버를 사용하나요?

16. TPU와 같은 더 고도의 하드웨어도 필요한가요?

17. 내부에서 모델을 훈련시킬 예정인가요, 아니면 외부의 모델 훈련 서비스를 활용할 예정인가요? 하이퍼 파라미터 튜닝(hyper-parameter tuning)과 신경 구조 검색(architecture search) 등을 자동화 할 예정인가요? (Auto-ML을 활용하나요?)