

3기 신도시 집값 예측

2021. 8. 11

신진우 박태윤 김범중



목 차



PROLOGUE

프로젝트 추진 동기

프로젝트 목표와 방향설정



데이터 집계 및 분석

각 데이터의 그래프와 추세

및 취합데이터 출처, 전처리 과정



1.2기 신도시 시각화

인구 및 실거래가

유익시설현황 및 유해시설현황



1분당신도시 & 판교신도시 집값회귀분석

1기 신도시, 2기 신도시 대표 지역

공원과 역과의 거리, 아파트별 위치와 환경에 따른 집값분석



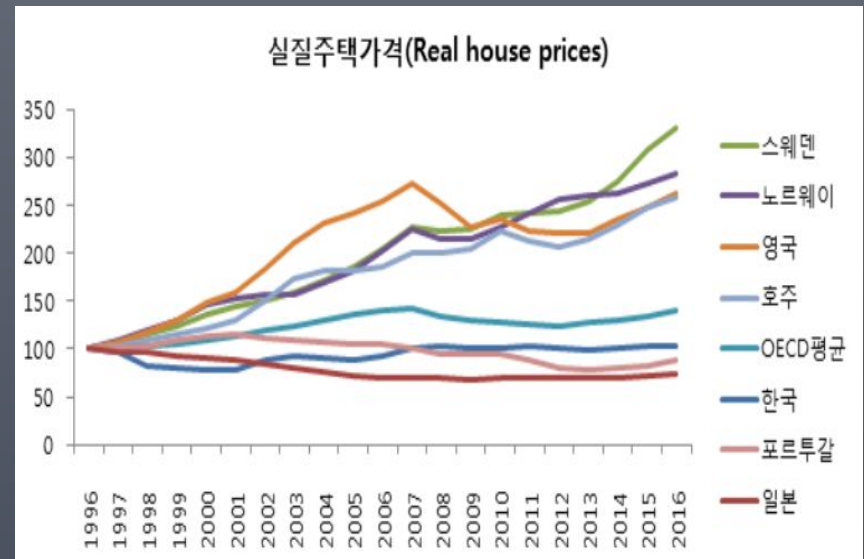
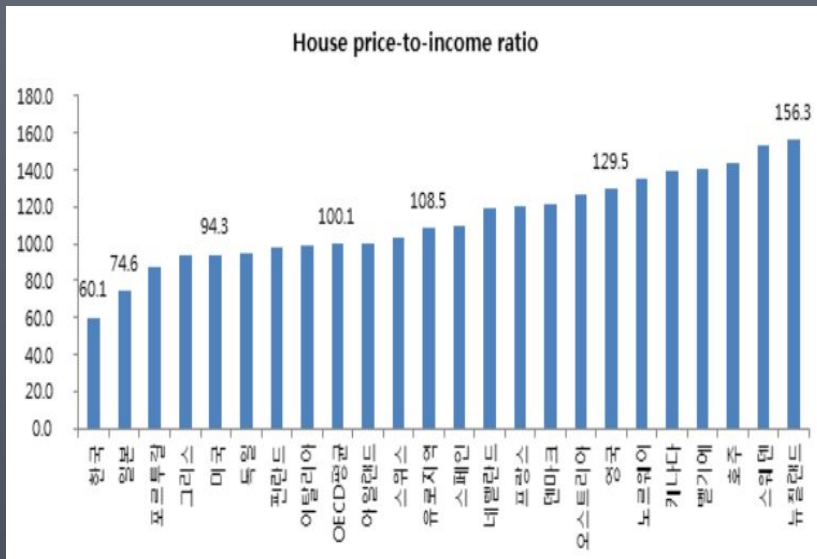
EPILOGUE

보완점 및 향후 발전 계획



프로젝트 추진 동기

- 2017년 부터 시작된 집값 상승
- 전국 평균 집값상승은 타국가에 비해 낮은편
- 소득대비 집값비율도 낮은편
- 실질주택가격 (나라별 물가지수에 따른 가격)
OECD평균 이하



프로젝트 방향 및 목표

○ 앞의 조사기간과는 상이하지만
주요도시의 집값상승을 상위에 서울
(수도권)이 랭크됨(5위)

○ 수도권 집중화와 구조적인 문제로
집값 상승 지속

○ 집값 안정화를 목표로 했던
1.2기 신도시 중 분당, 판교신도시
의 집값에 영향을 주는 요인분석

○ 추진중인 3기 신도시의 집값 예측
및 집값에 미치는 요인 분석

THE KNIGHT FRANK GLOBAL RESIDENTIAL CITIES INDEX, Q1 2021

RANKED BY ANNUAL % CHANGE (Q1 2020-Q1 2021)

CITY	12-MONTH CHANGE (%)	CITY	12-MONTH CHANGE (%)	CITY	12-MONTH CHANGE (%)
1 IZMIR	33.9%	51 PERTH	8.9%	101 SANTIAGO	3.9%
2 ANKARA	30.3%	52 HAMBURG	8.9%	102 THESSALONIKI	3.8%
3 WELLINGTON	30.1%	53 NAGOYA	8.8%	103 BEIJING	3.6%
4 ISTANBUL	28.8%	54 GUANGZHOU	8.6%	104 HANGZHOU	3.5%
5 SEOUL	26.1%	55 SYDNEY	8.6%	105 SHENZHEN	3.4%
6 HALIFAX	22.5%	56 SOFIA	8.6%	106 SKOPJE	3.2%
7 MOSCOW	22.4%	57 LYON	8.6%	107 ROME*	3.1%
8 ST. PETERSBURG	22.1%	58 QUEBEC	8.1%	108 TIANJIN	3.1%
9 HAMILTON	20.9%	59 MUNICH	8.1%	109 EDMONTON	2.9%
10 PHOENIX	20.0%	60 XI'AN	8.0%	110 TAIPEI CITY	2.8%
11 AUCKLAND	19.6%	61 VANCOUVER	7.9%	111 HAIFA	2.7%
12 SAN DIEGO	19.1%	62 EXETER	7.9%	112 BUCHAREST	2.3%
13 OTTAWA GATINEAU	19.0%	63 WINNIPEG	7.8%	113 BOGOTA	2.2%
14 SEATTLE	18.2%	64 OXFORD	7.0%	114 DUBLIN	2.2%
15 LUXEMBOURG	16.6%	65 GENEVA	6.8%	115 HONG KONG**	2.2%
16 MONTREAL	16.1%	66 SINGAPORE	6.6%	116 RIO DE JANEIRO	2.1%
17 CANBERRA	15.7%	67 BERN	6.6%	117 CALGARY	1.8%
18 BRATISLAVA	15.6%	68 HELSINKI	6.5%	118 BANGKOK	1.4%
19 BOSTON	14.8%	69 LILLE	6.5%	119 HYDERABAD, IN	1.3%
20 STOCKHOLM	14.4%	70 MELBOURNE	6.4%	120 ZHENGZHOU	1.2%
21 COPENHAGEN	14.4%	71 CHONGQING	6.4%	121 JAKARTA	1.1%
22 GLASGOW	14.3%	72 ZURICH	6.4%	122 VALENCIA	1.0%
23 HOBART	13.8%	73 NANJING	6.3%	123 LIMASSOL	1.0%
24 DENVER	13.4%	74 PARIS	6.3%	124 DELHI	0.9%
25 DALLAS	13.4%	75 AMSTERDAM	6.1%	125 ABU DHABI	0.8%
26 LOS ANGELES	13.4%	76 CHANGSHA	5.9%	126 LJUBLJANA	0.7%
27 MALMO	12.9%	77 BERLIN	5.9%	127 NICOSIA	0.6%
28 OSLO	12.9%	78 ZAGREB	5.9%	128 ABERDEEN	0.4%
29 NEW YORK	12.3%	79 WUXI	5.8%	129 LIMA	-0.1%
30 MIAMI	12.2%	80 BRUSSELS	5.7%	130 MALAGA	-0.3%
31 WASHINGTON	12.2%	81 OSAKA	5.6%	131 RIGA	-0.4%
32 SAN FRANCISCO	12.1%	82 TOKYO	5.6%	132 PALERMO*	-0.5%
33 DETROIT	12.1%	83 WUHAN	5.5%	133 BARCELONA	-0.9%
34 MANCHESTER	11.7%	84 ATHENS	5.5%	134 PRAGUE	-0.9%
35 VIENNA	11.6%	85 LISBON	5.5%	135 BUDAPEST	-0.9%
36 MILAN	11.5%	86 VIENNA	5.5%	136 BUDAPEST	-0.9%
37 AMSTERDAM	11.4%	87 VIENNA	5.5%	137 BUDAPEST	-0.9%
38 VIENNA	11.4%	88 VIENNA	5.5%	138 BUDAPEST	-0.9%
39 VIENNA	11.4%	89 VIENNA	5.5%	139 BUDAPEST	-0.9%
40 VIENNA	11.4%	90 VIENNA	5.5%	140 BUDAPEST	-0.9%

신도시 개발 현황

1989

2004

2008

2011

2021

1기 신도시

분당 성남시
일산 고양시
평촌 안양시
산본 군포시
중동 부천시

2기 신도시 1차

화성 동탄 1기
성남 판교
파주 운정

2기 신도시 2차

수원 광교
김포 한강

2기 신도시 3차

화성 동탄 2기
양주 옥정
성남 위례
인천 검단 (2017)
고덕 신도시 (2018)

3기 신도시

남양주 왕숙
하남 교산
과천 과천
안산 장상
광명 시흥
부천 대장
인천 계양
고양 창릉

데이터 집계 및 분석

국가 교통 DB : 교통관련 빅데이터

<https://www.ktdb.go.kr/www/selectBbsNttList.do?bbsNo=18&key=301>

교통데이터 거래소 : 사람, 차량, 공간정보 분석

<https://www.bigdata-transportation.kr/>

국가공간정보포털 : 도로, 환경, 기후, 도시개발

<http://www.nsdi.go.kr/lxportal/>

통계지리정보 서비스 : 인구, 주택, 행정구역, 교통, 통계 및 분석지도

<https://sgis.kostat.go.kr/view/index?param=0>

교통정보 연계시스템 : 교통정보, CCTV

<http://www.roadplus.co.kr/main/main.do>

국가 대중교통 정보 : 시내고속버스, 철도, 항공, 해운 정보

<https://www.tago.go.kr/>

교통카드 빅데이터 : 정류장, 노선별 차내인원 등 통합정보 시스템

<https://stcis.go.kr>

경기도 교통정보 센터 : 경기도 교통 지표, 교통량

https://gits.gg.go.kr/gtdb/web/trafficDb/newzone/HCO01/4/VOLUMN/VOL_LANE.do

통계청

<https://kosis.kr/oneid/cmmn/login/MemberType.do>

데이터 집계 및 분석

경기 데이터 드림 : 경기도 실거래가 현황 등

<https://data.gg.go.kr/>

행정표준코드 관리시스템 : 아파트, 기관, 도로, 유동인구 데이터 검색

https://www.code.go.kr/etc/index_iframe.do

데이터 집계 및 분석

1. Geojson과 json 데이터 수집과 사용방법 숙지 3일 소요 (데이터 수집 및 전처리 5일)

- 시/군/구/읍/면/동 별 맵 차트를 그리기 위해 가장 중요한건 사실 지도 데이터
- 원천데이터는 SHP File(Shape file 이하 SHP)
- SHP 포맷은 GIS(지리정보시스템) 에서 제공하는 파일
- (<http://www.gisdeveloper.co.kr/?p=2332>)

2021-01-15 글쓴이 김 행준

대한민국 최신 행정구역(SHP) 다운로드

대한민국의 행정구역에 대한 시도, 시군구, 읍면동, 리에 대한 공간 데이터입니다. 읍면동의 동은 법정동입니다. 도로명주소 DB로 제공되는 행정구역도를 일반 사용자들이 쉽게 사용할 수 있도록 병합하여 제공하고 있습니다. 본 데이터를 사용하실 경우 다른 분들도 널리 사용하실 수 있도록 출처를 언급하여 알려 주시면 감사하겠습니다.



[2021년 1월 업데이트 다운로드](#)
[2020년 5월 업데이트 다운로드](#)
[2019년 5월 업데이트 다운로드](#)
[2019년 2월 업데이트 다운로드](#)
[2018년 4월 업데이트 다운로드](#)
[2017년 3월 업데이트 다운로드](#)
[2016년 2월 업데이트 다운로드](#)
[2015년 6월 업데이트 다운로드](#)
[2014년 7월 업데이트 다운로드](#)
[2014년 5월 업데이트 다운로드](#)
[2013년 11월 업데이트 다운로드](#)



[2021년 1월 업데이트 다운로드](#)
[2020년 5월 업데이트 다운로드](#)
[2019년 5월 업데이트 다운로드](#)
[2019년 2월 업데이트 다운로드](#)
[2018년 4월 업데이트 다운로드](#)
[2017년 3월 업데이트 다운로드](#)
[2016년 2월 업데이트 다운로드](#)
[2015년 6월 업데이트 다운로드](#)
[2014년 7월 업데이트 다운로드](#)
[2014년 5월 업데이트 다운로드](#)

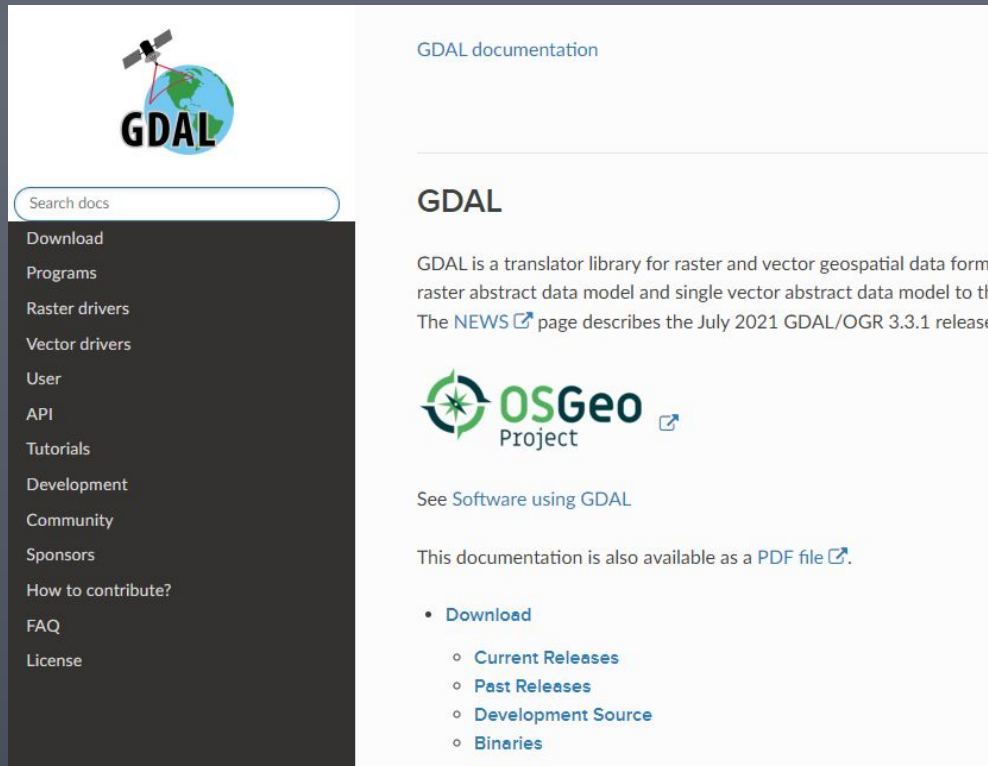
Name

[TL_SCCO_CTPRVN.dbf](#)
[TL_SCCO_CTPRVN.prj](#)
[TL_SCCO_CTPRVN.shp](#)
[TL_SCCO_CTPRVN.shx](#)

데이터 집계 및 분석

1. Geojson과 json 데이터 수집과 사용방법 숙지 3일 소요 (데이터 수집 및 전처리 5일)


- 다운로드 받은 SHP 파일을 직접 핸들링하기에는 힘들어 맵을 표현하기 위해 GeoJSON 또는 Json 이라는
- 지리 정보 표시를 위한 표준 Json포맷으로 변환해야 함
- SHP to GeoJSON에는 여러가지 방법들이 있었고 이를 적용하는데 많은 시간을 할애했
- GDAL이라는 지리데이터 변환 라이브러리에서 제공하는 ogr2ogr 이라는 tool을 사용하여 변환
- (<https://gdal.org/index.html>)



GDAL documentation

GDAL

GDAL is a translator library for raster and vector geospatial data formats to a range of raster abstract data model and single vector abstract data model to the software. The [NEWS](#) page describes the July 2021 GDAL/OGR 3.3.1 release.

 **OSGeo** Project

See [Software](#) using GDAL

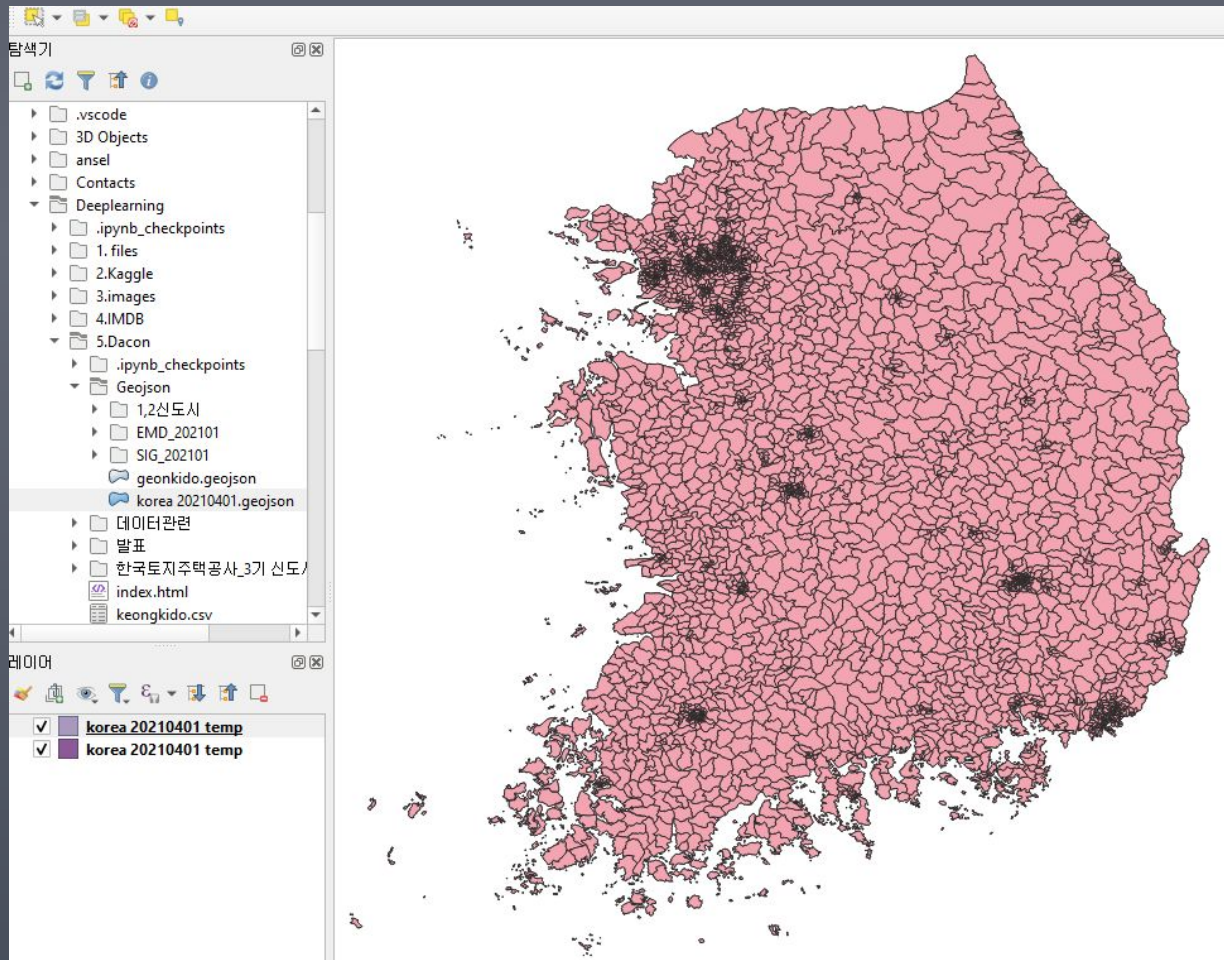
This documentation is also available as a [PDF file](#).

- **Download**
 - [Current Releases](#)
 - [Past Releases](#)
 - [Development Source](#)
 - [Binaries](#)

데이터 집계 및 분석

1. Geojson과 json 데이터 수집과 사용방법 숙지 3일 소요 (데이터 수집 및 전처리 5일)

- 변환을 하였으나 여러 이유로 사용이 힘들었고 결국에 Github에 올라와있는 행정동기준 Geojson파일을 받음
- Json에 표현되는 벡터 정보는 폴리곤들의 사이즈가 크기도 하고 원하는 지역만 가져오기 위해 QGIS와 Mapshaper 사용

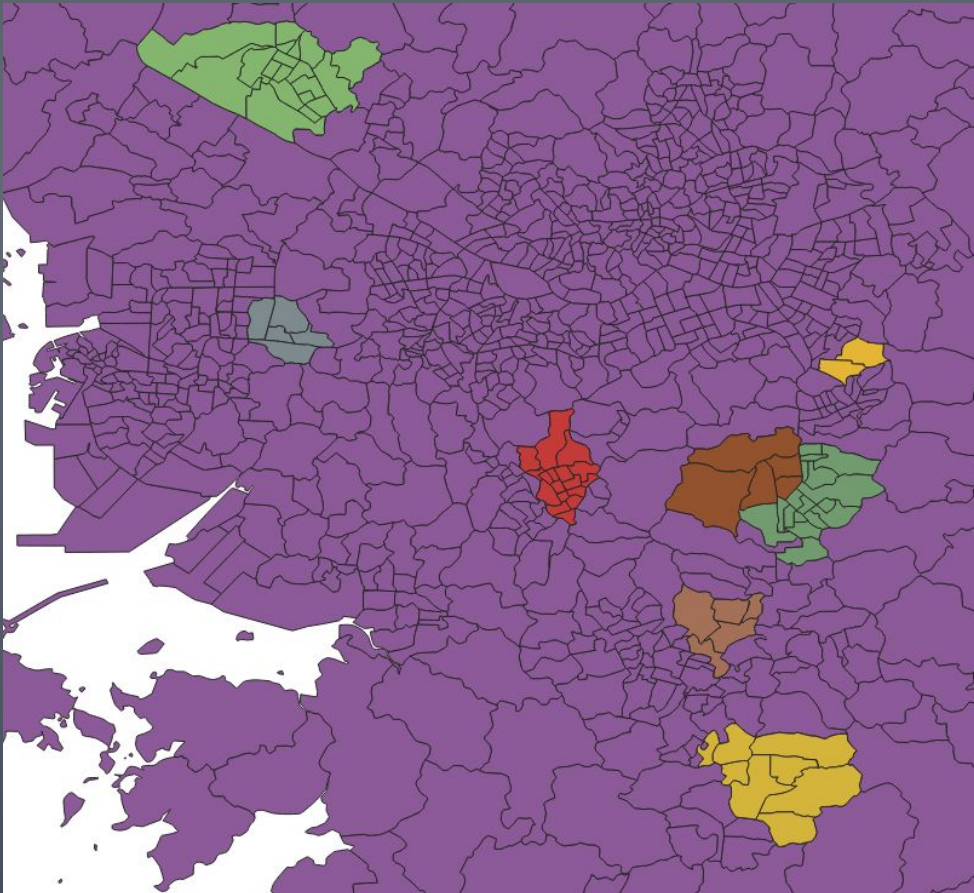


데이터 집계 및 분석

1. Geojson과 json 데이터 수집과 사용방법 숙지 3일 소요 (데이터 수집 및 전처리 5일)

- 법정동기준 json파일로 변환작업(Mapshaper)으로 각 시도시별 Geojson파일 생성
- 지도에 매핑할 동별 실거래가 평균계산 작성 - colab

https://drive.google.com/file/d/185b78G185dNp8Qw7g_a_HXMocm8YWCFm/view?usp=sharing

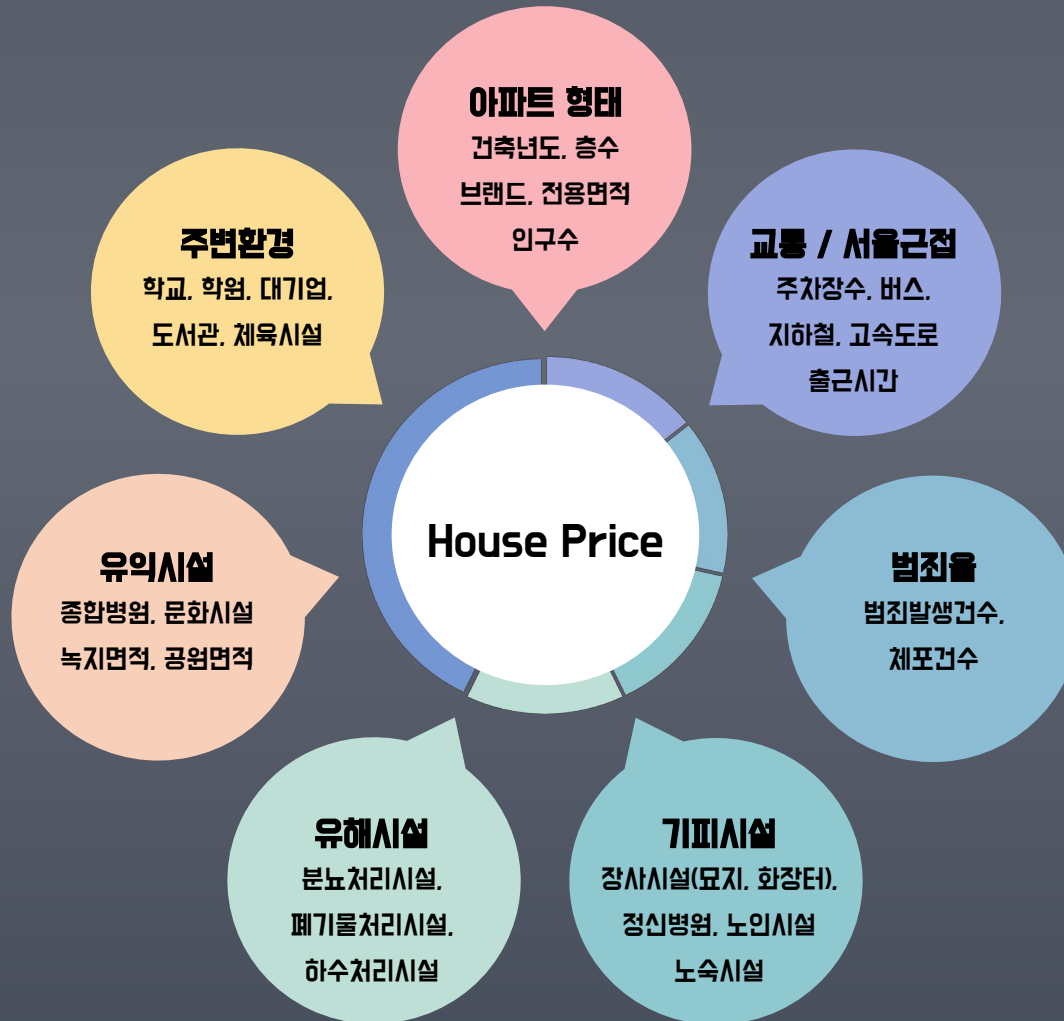


- boondang2.geojson
- boondang2.qmd
- dongtan.geojson
- dongtan.qmd
- gangkyo.geojson
- gangkyo.qmd
- ilsan.geojson
- ilsan.qmd
- joongdong.geojson
- joongdong.qmd
- pankyo.geojson
- pankyo.qmd
- pyeongchon.geojson
- pyeongchon.qmd
- pyeongchon2.geojson
- pyeongchon2.qmd
- Wiryegeojson
- Wiryeqmd

데이터 집계 및 분석

2. 집값에 영향을 미치는 관련자료 수집 및 전처리 2일 소요 (데이터 수집 및 전처리 5일)

- 집값에 영향을 미칠것으로 생각되는 자료 수집



1. 27기 신도시 시각화

1. 경기도 아파트 실거래가 시각화

- 2020. 7 ~ 2021. 7 / 70 ~ 85m2 / 법정동 기준

경기도 아파트 실거래

매매 평단가(만원)

과천시(4,659만원)

2018년 아파트 평당(실거래가) 최상위지역

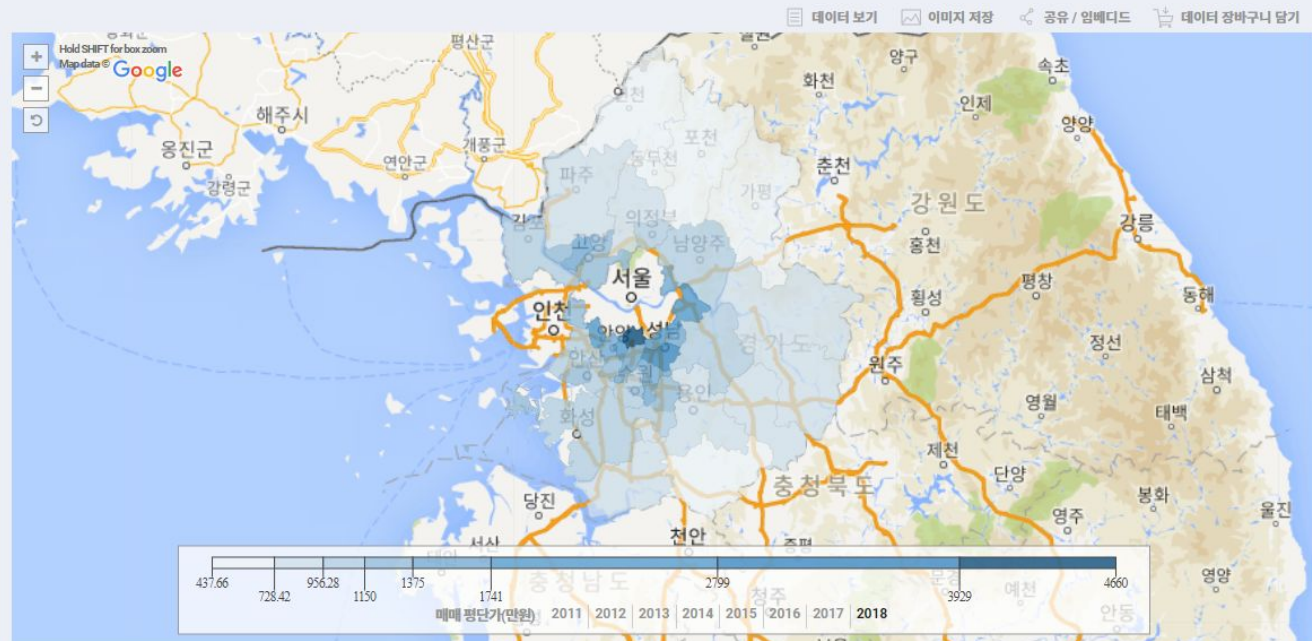
연천군(485만원)

2018년 아파트 평당(실거래가) 최하위지역

2018년 경기도에서 아파트 평당(실거래가) 최상위지역은 과천시(4,659만원)로 나타났고, 최하위지역은 연천군(485만원)로 나타났습니다.



Dataset: 국토교통부
Source: Ministry of Land, Infrastructure and Transport

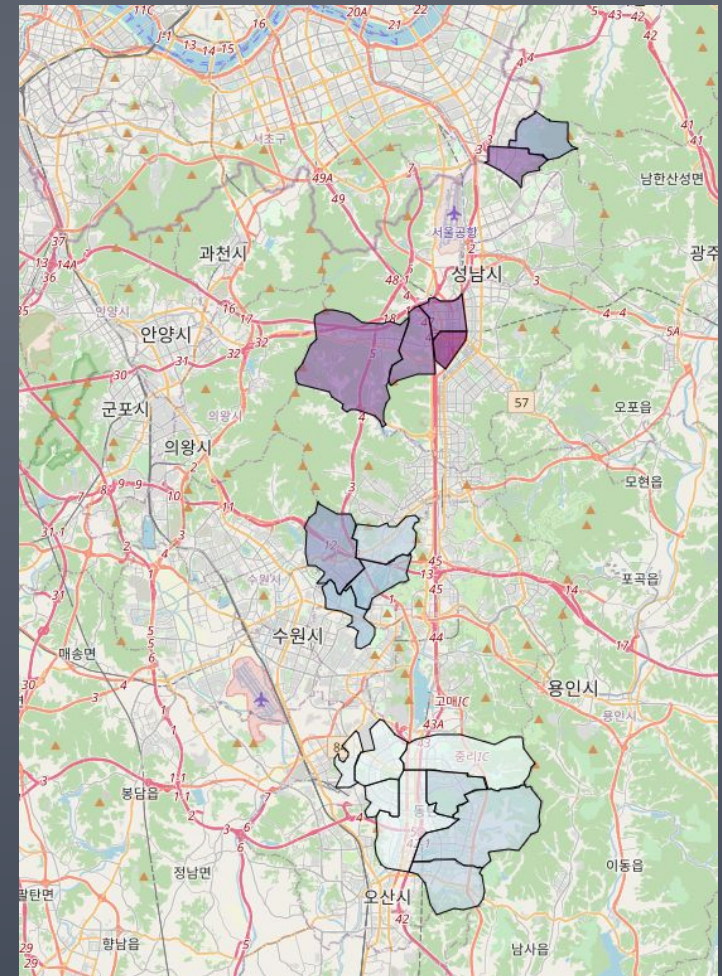
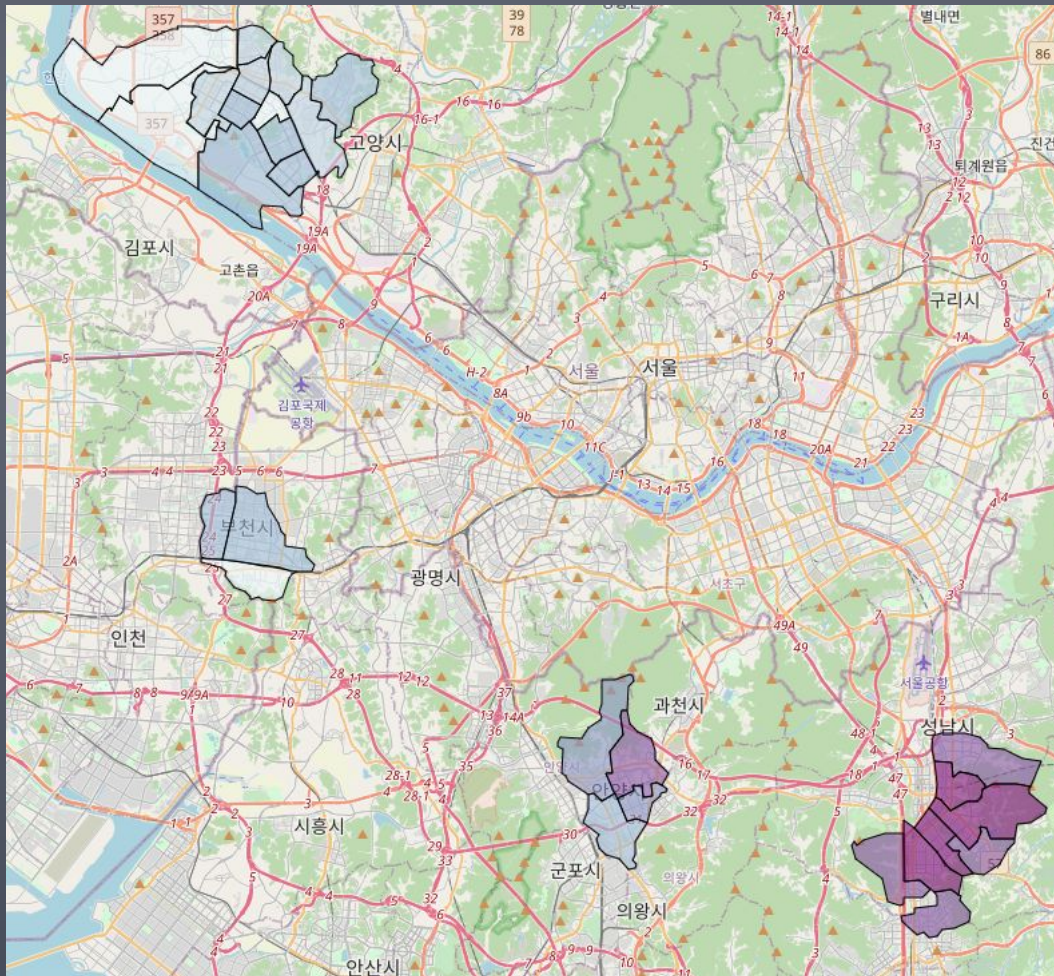


1. 2기 신도시 시각화

1. 경기도 아파트 실거래가 시각화 - colab

<https://colab.research.google.com/drive/1dG6oa5mdFEPon8w3yxRVeAgpeWQYKbVI?usp=sharing>

- 2020. 7 ~ 2021. 7 / 70 ~ 85m² / 법정동 기준

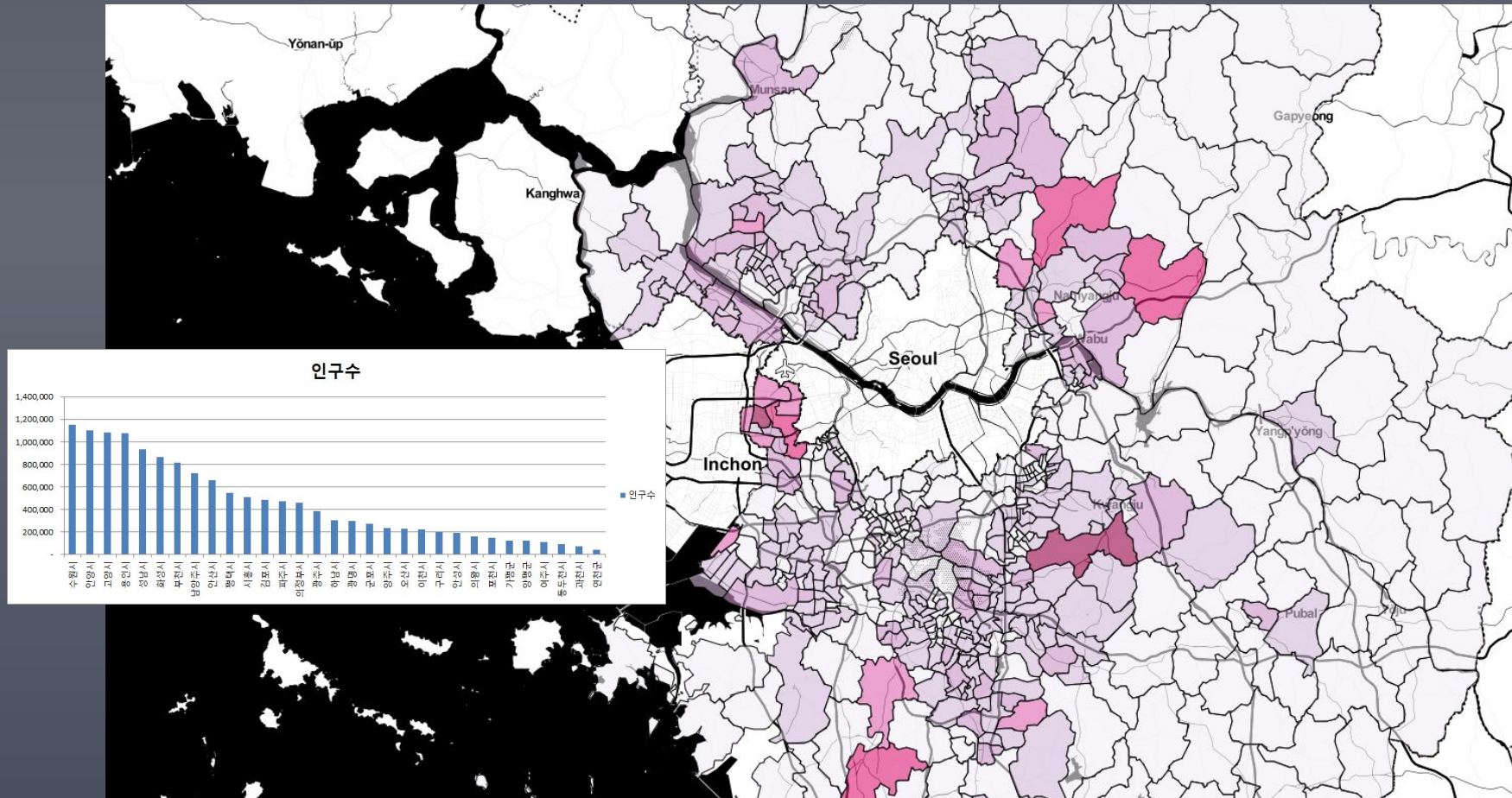


1. 2기 신도시 시각화

2. 경기도 주민등록상 인구수 시각화 - colab

<https://drive.google.com/file/d/1wcz9IR-p2Ly9EXBn-j60udEhWJzDURd4/view?usp=sharing>

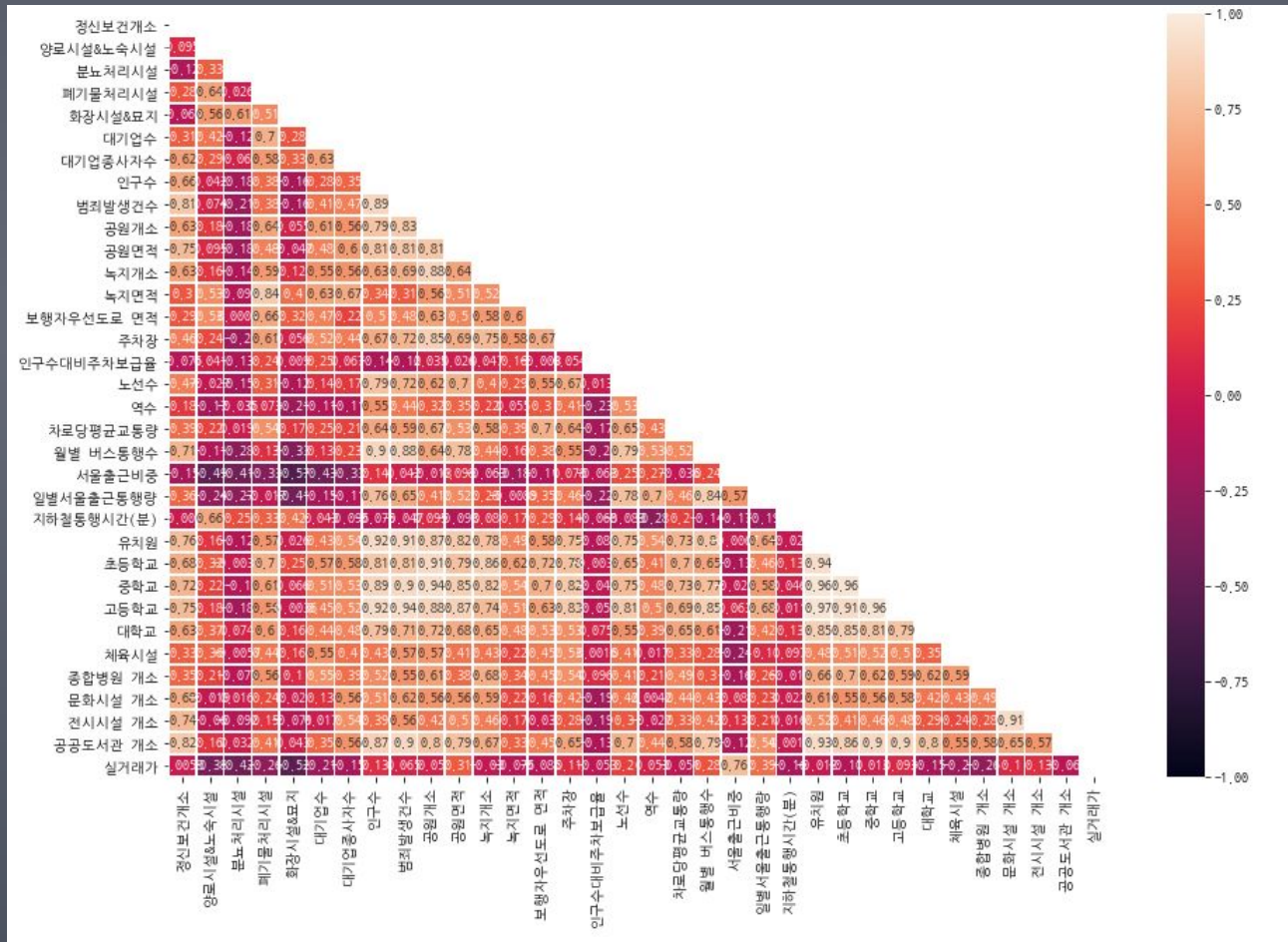
- 2021. 4 / 행정동 기준



1. 27기 신도시 시각화

3. 집값관련 요인 전처리 및 시각화 - colab

- 2020년 / 시군구 기준
- 집값과 연관성 있는 분야 : 공원, 출퇴근, 기피시설, 인구수, 문화시설



분당신도시 & 판교신도시 집값회귀분석

1. 분당 판교의 집값 영향요소 선정 및 모델링

- 변수설정 : 동별인구, 전용면적, 층수, 건축년도, 역과의거리, 공원과의 거리
 - ※ 브랜드명은 단지명으로 분류가능하나 금번 모델링에 적용 안함
- 거리측정 - https://drive.google.com/file/d/1js7N7Dq4D_qiMOp0fz4WAEoo_ge7hTVw/view?usp=sharing
- 시각화 - https://drive.google.com/file/d/1WxNpKW60PCYS1CUMIKk3FUKeUBE__Umv/view?usp=sharing
- 모델링 - <https://drive.google.com/file/d/13--RT4NwaS5GFfKup2b-R00FBaxNPUIw/view?usp=sharing>
-

결론

1. 시사점 및 한계점

- 집값요인과 관련된 여러 요인에 대한 데이터는 여러 사이트와 기관에서 제공하고 있지만 오픈API나 자료요청으로 받을 수 있는 한계가 있고, 취합하는데 시간이 오래걸려 실질적으로 프로젝트 기간의 절반을 할애함
- 많은 데이터 취합해도 연관성이 있는지 없는지 여부를 일부 데이터만 가지고 판단할수는 없어서 그에 따른 각 데이터간 연관여부에 좀더 명확한 구분과 연구가 필요할것으로 사료
- 지도에 맵핑하여 시각화를 공부하면서 다양한 Visualization 툴이 존재함과 시각화의 매력을 알게되었음
- 분당과 판교로 압축하여 집값과 연관된 요소로 모델링을 하였을때 나오는 결과값으로 향후 집값 또는 제 3기 신도시의 집값을 예측할 수 있겠는가 라는 의문이 남음
- 이를 통해 집값에 미치는 요인들의 시계열 데이터를 뽑아서 분석하면 더 정확한 예측값을 구할수 있을것으로 예상

2. 향후 발전 계획

- DAICON에서 진행하는 국토교통 빅데이터 온라인 해커톤 대회에 3기 신도시 실거래가예측으로 참여하려고 했으나 이번 프로젝트를 통해 실거래가 예측은 금리와 미국주요도시의 집값변동추이, 전세, 월세가격의 변동추이 등의 많은 변수 중 연관성 있는 데이터를 어떻게 가져와야되는지 고민해봐야될 것 같고, 시계열데이터(일정한 시간동안 수집된 일련의 정해진 데이터셋)의 분석기법 여러가지를 연구해보고 최적화된 모델을 활용하면 좋을것 같습니다.

감 사 합 니 다