



Team 짜장

최종발표

윤대*

- 데이터 전처리
- 데이터 분석
- 데이터 시각화
- 발표

권나*

- 웹 크롤링
- 데이터 분석
- 데이터 시각화
- ML, DL 모델링
- 회의록 작성

김동*

- 웹 크롤링
- 데이터 전처리
- 상관관계 분석
- 데이터 시각화

김선*

- 데이터 전처리
- 데이터 분석
- 데이터 시각화

이하*

- 데이터 전처리
- 데이터 분석
- 데이터 시각화
- PPT 작성

이효*



1 주제 선정

2 데이터 분석

- 변수 설명
- EDA
- Machine Learning
- Deep Learning

3 결론 및 아쉬운점

Part 1, 주제 선정

↑ ВЪЕЗД ↑

'주차전쟁+원인'으로 검색, 뉴스 title과 밑에 preview 내용 크롤링



전체 내용 모두



전체 단어에서 검색어인 주차, 주차장, 전쟁, 원인을 빼 것

1. 주차 전쟁 Word Cloud

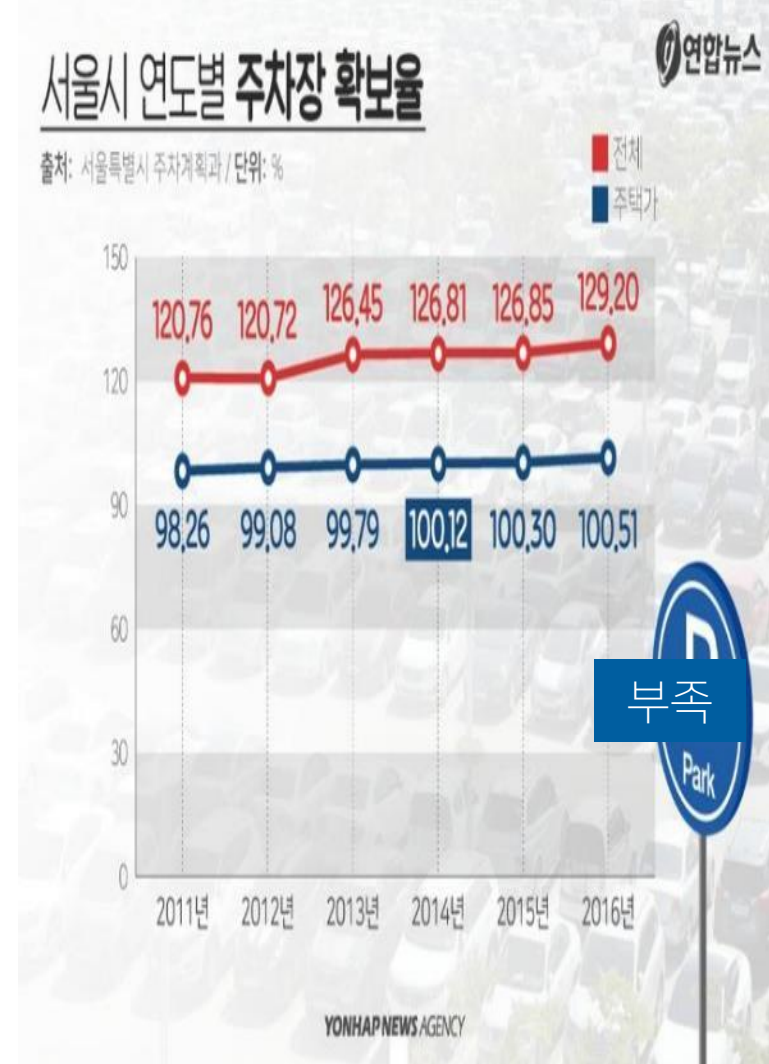
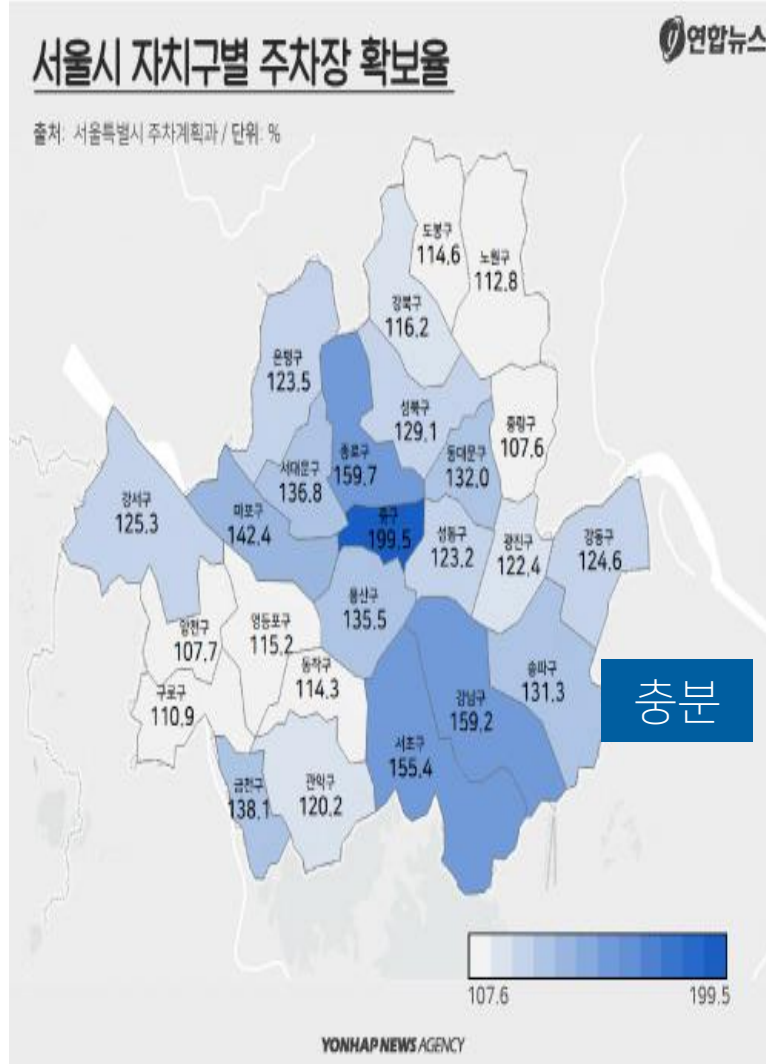
'주차전쟁+원인'으로 검색, 뉴스 title과 밑에 preview 내용 크롤링



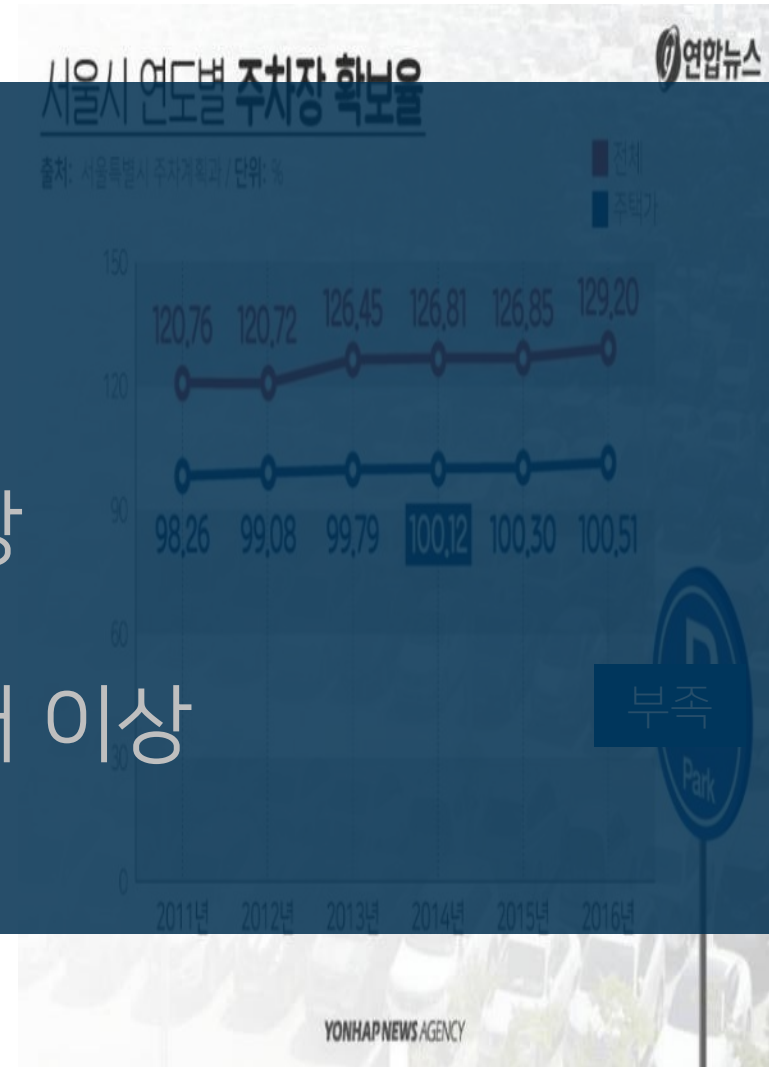
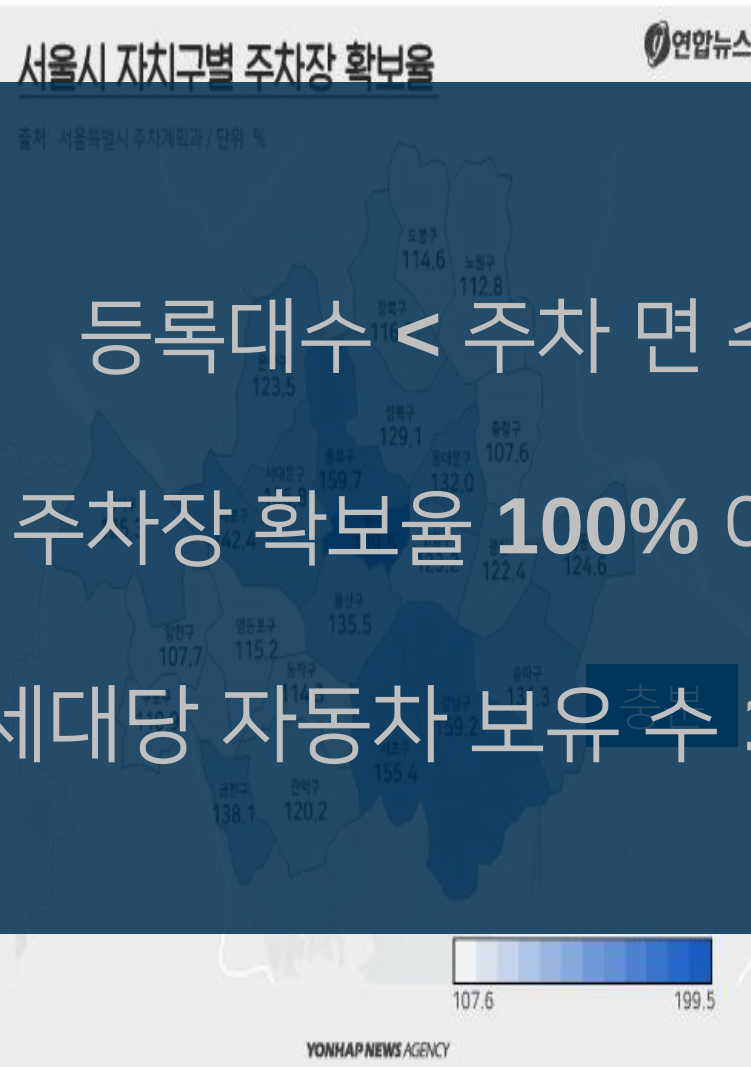
전체 내용 모두

전체 단어에서 검색어인 주차, 주차장, 전쟁, 원인을 빼 것

2. 충분 VS 부족



2. 충분 VS 부족



그러나 세대당 자동차 보유 수 1대 이상

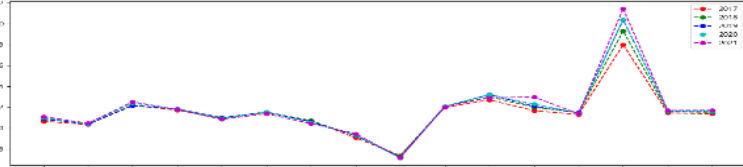
3. 왜 부족한가?

①세대별 등록차량수

세대수 차량등록대수 세대당등록수

시도명	세대수	차량등록대수	세대당등록수
강원도	746220	828146	1.110
경기도	8172214	8554726	1.047
경상남도	1957911	2451903	1.252
경상북도	1509047	1776477	1.177
광주광역시	645712	700783	1.085
대구광역시	1063893	1211095	1.138
대전광역시	664417	692702	1.043
부산광역시	1544663	1464608	0.948
서울특별시	4426007	3176743	0.718
울산광역시	482650	581429	1.205
인천광역시	1298647	1675405	1.290
전라남도	903108	1172282	1.298
전라북도	1142208	1298250	1.137
제주특별자치도	307529	658594	2.142
충청남도	1299902	1519895	1.169
충청북도	1147688	1344394	1.171

세대당 등록차량수



②건설비용 절감

아시아경제

PICK

‘갈등 유발’ 주차 분쟁...서울·인천 아파트, 세대당 1대도 못대

입력 2022.08.16. 오후 2:20 · 수정 2022.08.16. 오후 2:21 기사원문

유병돈 기자 TALK

1 18

더 보기

서울 세대당 0.89대... 인천 세대당 0.93대
전국 세대당 차량 보유대수는 2.5대 가량



사진은 기사 내용과 관련 없음. 사진=아시아경제 DB

③주차대수 선정

방법	장점	단점
과거추세 연장법	• 이해가 쉽고 적용이 간단 • 교통패턴변화가 적은 도심 • 중소도시의 단기간 주차수요 추정	• 너무 개괄적이므로 신뢰성이 떨어짐 • 장래불확실성 고려하지 못함
원단위법	주차발생 원단위법	• 적용변수가 간단하고 단기간 주차수요 예측에 적합 • 장래주차 발생원단위 변화시 신뢰성 저하 • 주차이용 효율 산정이 어려움
	건물연면적당 원단위법	• 일반적인 주차수요 예측시 주로 사용 • 중복통행량 고려 필요
	교통량 원단위법	• 교통여건이 안정적이고 지구경계가 분명한 곳에 적용 • 통행실대 O/D 조사 활용
P 요소법 Parking Space Factor Method	• 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 • 지구나 도심지와 같은 특정한 지역의 주차수요 추정에 적합	• 추정계념은 명확하나 각종요소를 얻기위한 조사가 필요 • 계절주차 집중계수, 지역주차 조정계수 등이 미비하여 현 단계에서 적용성이 미약함
누적주차수요 추정법	• 원단위법 단점을 보완하고 단일·복합 용도시설에 적용시 신뢰성이 높은 • 시간대별 특성별수를 독립변수로하여 시간대별 주차수요변화의 고려가 가능	• 용도별로 개별추정이 필요하므로 시간과 경비가 과다 소요됨 • 출발/도착 분포를 잘못가정시 신뢰성이 저하
사람통행조사 Person Trip Survey	• 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 • 가구설문조사와 가운 방법으로 얻어진 기종점조사표에 의해 통행발생량을 예측하고 이를 각 교통수단으로 분류하여 승용차의 유입통행량을 토대로하여 추정하는 방법	• 추정계념은 명확하나 각종요소를 얻기위한 조사가 필요
자동차 기종점에 의한 방법	• 지역의 총체적 수요추정(광역) 및 토지이용변화가 큰 도심지역 주차수요 추정에 적합	• 조사비용 및 시간 과다 소요 (노상 및 노외주차장의 차량 O/D 조사)

3. 왜 부족한가?

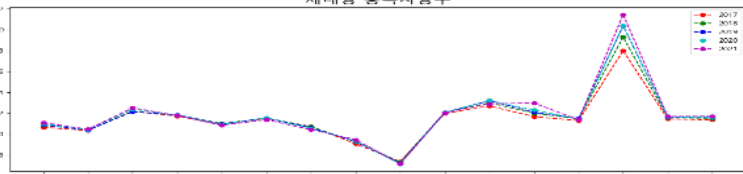
①세대별 등록차량수

세대수 차량등록대수 세대당등록수

시도명

강원도	746220	828146	1.110
경기도	8172214	8554726	1.047
경상남도	1957911	2451903	1.252
경상북도	1509047	1776477	1.177
광주광역시	645712	700783	1.085
대구광역시	1063893	1211095	1.138
대전광역시	664417	692702	1.043
부산광역시	1544663	1464608	0.948
서울특별시	4426007	3176743	0.718
울산광역시	482650	581429	1.205
인천광역시	1298647	1675405	1.290
전라남도	903108	1172282	1.298
전라북도	1142208	1298250	1.137
제주특별자치도	307529	658594	2.142
충청남도	1299902	1519895	1.169
충청북도	1147688	1344394	1.171

세대당 등록차량수



②건설비용 절감

아시아경제

PICK

‘갈등 유발’ 주차 분쟁...서울·인천 아파트, 세대당 1대도 못대

입력 2022.08.16 오후 2:20 수정 2022.08.16 오후 2:21 기사원문

유병돈 기자

1 18

1 18

서울 세대당 0.89대... 인천 세대당 0.93대
전국 세대당 차량 보유대수는 2.5대 차량

세대별 등록 차량 수 1대 이상

③주차대수 선정

방법	장점	단점
과거주세 연장법	• 이해가 쉽고 적용이 간단 • 교통패턴변화가 적은 도심 • 중소도시의 단기간 주차수요 추정	• 너무 개괄적이므로 신뢰성이 떨어짐 • 장래불확실성 고려하지 못함
원단위법	- 주차발생 원단위법 - 건물연면적당 원단위법 - 교통량 원단위법	- 주차발생 원단위법 - 건물연면적당 원단위법 - 교통량 원단위법
누적주차수요 추정법	• 원단위법 단점을 보완하고 단일·복합 용도시설에 적용시 신뢰성이 높은 • 시간대별 특성별수를 독립변수로하여 시간대별 주차수요변화의 고려가 가능	• 추경계념은 명확하나 각종요소를 열거위한 조사가 필요 • 용도별/도차 분포를 잘못가정시 신뢰성이 저하
사람통행조사 Person Trip Survey	• 사회경제지표와 트라이앵글, 교통수요, 주차특성 등을 포괄적으로 고려 • 가구설문조사와 가운 방법으로 얻어진 기종점조사표에 의해 통행발생량을 예측하고 이를 각 교통수단으로 분류하여 승용차의 유입통행량을 토대로하여 추정하는 방법	• 추경계념은 명확하나 각종요소를 열거위한 조사가 필요
자동차 기종점에 의한 방법	• 지역의 총체적 수요추정(광역) 및 토지이용변화가 큰 도심지역 주차수요 추정에 적합	• 조사비용 및 시간 보다 소요 (노상 및 노외주차장의 차량 O/D 조사)

3. 왜 부족한가?

①세대별등록차량수

시도명	세대수	차량등록대수	세대당등록수
강원도	746220	828146	1.110
경기도	8172214	8554726	1.047
경상남도	1957911	2451903	1.252
경상북도	1509047	1776477	1.177
광주광역시	645712	700783	1.085
대구광역시	1063893	1211095	1.138
대전광역시	664417	692102	1.043
부산광역시	1446007	1461113	0.948
서울특별시	442007	3176143	0.718
울산광역시	482650	581429	1.205
인천광역시	1298647	1675405	1.290
전라남도	903108	1172282	1.298
전라북도	1142208	1298250	1.137
제주특별자치도	307529	658594	2.142
충청남도	1299902	1519895	1.169
충청북도	1147688	1344394	1.171

수익성



②건설비용 절감

아시아경제

PICK

‘갈등 유발’ 주차 분쟁...서울·인천 아파트, 세대당 1대도 못대

입력 2022.08.16. 오후 2:20 · 수정 2022.08.16. 오후 2:21 기사원문

유병돈 기자 TALK

1 18

다 가

서울 세대당 0.89대...인천 세대당 0.93대
전국 세대당 차량 보유대수는 2.5대 가량



사진은 기사 내용과 관련 없음. 사진=아시아경제 DB

③주차대수 선정

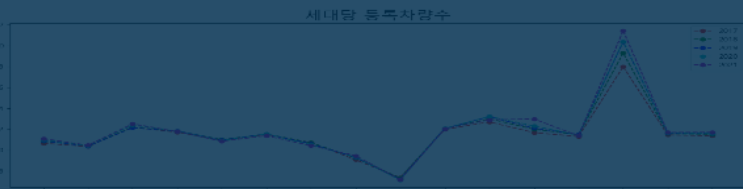
방법	장점	단점
과거주세 연장법	• 이해가 쉽고 적용이 간단 • 교통패턴변화가 적은 도심 • 중소도시의 단기간 주차수요 추정	• 너무 개괄적이므로 신뢰성이 떨어짐 • 장래불확실성 고려하지 못함
주차발생 원단위법	• 적용변수가 간단하고 단기간 주차수요 예측에 적합	• 장래주차 발생원단위 변화시 신뢰성 저하 • 주차이용 효율 선정이 어려움
진출입면적당 원단위법	• 건물면적당 주차수 예측 가능	• 건물면적당 주차수 예측 가능
교통수요 원단위법	• 교통수요 예측 가능	• 교통수요 예측 가능
P 요소법 Parking Space Factor Method	• 사회경제지표와 트라이앵글, 교통수요, 주차수요 등을 포괄적으로 고려 • 지리적 특성, 교통수요, 주차수요 등을 고려하여 주차수요를 예측	• 추정계수는 경험이나 각종요소를 얻기위한 조사가 필요
누적주차수요 추정법	• 원단위법 단점 보완하고 단일·복합 등 도시원에 적용시 신뢰성이 높음 • 시간대별 특성변수를 독립변수로하여 시간대별 주차수요변화의 고려가 가능	• 용도별 개별추정이 필요하므로 시간과 경비가 과다 소요됨 • 출발/도착 분포를 잘못가정시 신뢰성이 저하
사람통행조사 Person Trip Survey	• 사회경제지표와 트라이앵글, 교통수요, 주차수요 등을 포괄적으로 고려 • 가구설문조사와 가운 방법으로 얻어진 기종점조사자료에 의해 통행발생량을 예측하고 이를 각 교통수단으로 분류하여 승용차의 유입통행량을 토대로하여 추정하는 방법	• 추정계수는 경험이나 각종요소를 얻기위한 조사가 필요
자동차 기종점에 의한 방법	• 지역의 총체적 수요추정(광역) 및 토지이용변화가 큰 도심지역의 주차수요 추정에 적합	• 조사비용 및 시간 과다 소요 (노상 및 노외주차장의 차량 O/D 조사)

주차공간
최소화

3. 왜 부족한가?

①세대별등록차량수

시도명	세대수	차량등록대수	세대당등록수
강원도	746220	828146	1.110
경기도	8172214	8554726	1.047
경상남도	1957911	2451903	1.252
경상북도	1509047	1771477	1.177
광주광역시	645712	701113	1.085
대구광역시	1063893	1211113	1.138
대전광역시	664417	692702	1.043
부산광역시	1544663	1464608	0.948
서울특별시	4426007	3176743	0.718
울산광역시	482650	581429	1.205
인천광역시	1298647	1675405	1.290
전라남도	903108	1172182	1.298
전라북도	1142208	1238250	1.084
제주특별자치도	307529	658594	2.142
충청남도	1299902	1519895	1.169
충청북도	1147688	1344394	1.171



②건설비용 절감

아시아경제

‘갈등 유발’ 주차 분쟁...서울·인천 아파트, 세대당 1대도 못대

입력 2022.08.16 오후 2:20 수정 2022.08.16 오후 2:21 기사원문

유병돈 기자

서울 세대당 0.89대, 인천 세대당 0.93대
전국 세대당 차량 보유대수는 2.5대 가량



사진은 기사 내용과 관련 없음. 사진=아시아경제 DB

③주차대수 선정

방법	장점	단점
과거주세 연장법	• 이해가 쉽고 적용이 간단 • 교통패턴변화가 적은 도심 • 중소도시의 단기간 주차수요 추정	• 너무 개괄적이므로 신뢰성이 떨어짐 • 장래불확실성 고려하지 못함
원단위법	- 주차발생 원단위법 - 건물연면적당 원단위법 - 교통량 원단위법	- 주차발생 원단위법 - 건물연면적당 원단위법 - 교통량 원단위법
P 요소법 Parking Space Factor Method	• 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 • 지구나 도심지와 같은 특정한 지역의 주차수요 추정에 적합	• 추정계념은 명확하나 각종요소를 얻기위한 조사가 필요 • 계절주차 집중계수, 지역주차 조정계수 등이 미비하여 현 단계에서 적용성이 미약함
누적주차수요 추정법	• 원단위법 단점을 보완하고 단일·복합 용도시설에 적용시 신뢰성이 높은 • 시간대별 특성별수요를 독립변수로하여 시간대별 주차수요변화의 고려가 가능	• 용도별로 개별추정이 필요하므로 시간과 경비가 과다 소요됨 • 출발/도착 분포를 잘못가정시 신뢰성이 저하
사람통행조사 Person Trip Survey	• 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 • 가구설문조사와 가운 방법으로 얻어진 기종점조사표에 의해 통행발생량을 예측하고 이를 각 교통수단으로 분류하여 승용차의 유입통행량을 토대로하여 추정하는 방법	• 추정계념은 명확하나 각종요소를 얻기위한 조사가 필요
자동차 기종점에 의한 방법	• 지역의 총체적 수요추정(광역) 및 토지이용변화가 큰 도심지역 주차수요 추정에 적합	• 조사비용 및 시간 과다 소요 (노상 및 노외주차장의 차량 O/D 조사)

4. 주차대수는 어떻게 선정할까?

① 법정주차대수

구분	법정 주차 대수 계산
건물 건축	시설 면적 / 설치 기준 = 주차 대수
건물 증축	증축 면적 / 설치 기준 = 주차 대수
용도 변경	시설 면적 / 설치 기준 = 주차 대수

구분	법정 주차 대수 계산
공동 주택	소수점 올림
그 외	소수점 반올림

EX) 전용 85㎡ 100세대 아파트 지었을 때,

85㎡ * 100세대 / 75㎡(특별시 기준 면적)

= 113.3333대

- 주택단지 소수점 올림 => 114대

② 주차수요확인

방법	장점	단점
과거추세 연장법	- 이해가 쉽고 적용이 간단 - 교통패턴변화가 적은 도심 - 중소도시의 단기간 주차수요 추정	- 너무 개관적이므로 신뢰성이 떨어짐 - 장래불확실성 고려하지 못함
원단위법	- 주차발생 원단위법 - 건물연면적당 원단위법 - 교통량 원단위법	- 장래주차 발생원단위 변화시 신뢰성 저하 - 주차이용 효율 산정이 어려움
P 요소법 Parking Space Factor Method	- 일반적인 주차수요 예측시 주로 사용 - 교통여건이 안정적이고 지구경계가 분명한 곳에 적용	- 중복통행량 고려 필요 - 통행실대 O/D 조사 활용
누적주차수요 추정법	- 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 - 지구나 도심지와 같은 특정한 지역의 주차수요 추정에 적합	- 추정개념은 명확하나 각종요소를 열거위한 조사가 필요 - 계절주차 집중계수, 지역주차 조정계수 등이 미비하여 현 단계에서 적용성이 미약함
사람통행조사 Person Trip Survey	- 원단위법 단점을 보완하고 단일·복합 용도시설에 적용시 신뢰성이 높은 - 시간대별 특성별수를 독립변수로하여 시간대별 주차수요변화의 고려가 가능 - 사회경제지표와 토지이용, 교통수요, 주차특성 등을 포괄적으로 고려 - 가구설문조사와 가온 방법으로 얻어진 기점조사표에 의해 통행발생량을 예측하고 이를 각 교통수단으로 분류하여 승용자의 유입통행량을 토대로하여 추정하는 방법	- 용도별로 개별추정이 필요하므로 시간과 경비가 과다 소요됨 - 출발/도착 분포를 잘못가정시 신뢰성이 저하
자동차 기종점에 의한 방법	지역의 총체적 수요추정(광역) 및 토지이용변화가 큰 도심지역 주차수요 추정에 적합	추정개념은 명확하나 각종요소를 열거위한 조사가 필요
		조사비용 및 시간 과다 소요 (노상 및 노외주차장의 차량 O/D 조사)

주차원이 주변 주차대수를 수작업으로 계산
-> 좀더 객관적인 자료가 필요



Part 2, 데이터 분석

1. 변수 설명

컬럼명	전처리 방향	
단지코드	삭제	
총 세대수		
임대건물구분	Label Encoding	아파트 : 0, 상가 : 1
지역		노인청소년 비율로 대체
공급유형	삭제	
전용면적	Label Encoding	전용면적 별 카테고리 구분
전용면적별 세대수		공가수 제외한 세대수 컬럼 추가
공가수		
자격유형	삭제	A-Z 자격유형 기준 확인 불가
임대보증금	삭제	상가 및 매매 결측치 대체 불가
임대료	삭제	상가 및 매매 결측치 대체 불가
지하철역		지하철 + 버스 = 대중교통 컬럼 추가
버스정류장		
단지내주차면수		
등록차량수		

411 rows × 16 columns



	단지코드	총 세대수	임대 건물 구분	지역	전용 면적	전용 면적 별 세대수	공가 수	자격 유형	임대보증금	임대료	도보 10분 거리 내 지하철역 수(환승노선 수 반영)	도보 10분 거리 내 버스정류장 수	단지내 주차면수	등록차량 수
0	C2483	900	아파트	강남구	39.72	134	38.0	A	15667000	103680	0.0	3.0	1425.0	1015.0
1	C2483	900	아파트	강남구	39.72	15	38.0	A	15667000	103680	0.0	3.0	1425.0	1015.0
2	C2483	900	아파트	강남구	51.93	385	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0
3	C2483	900	아파트	강남구	51.93	15	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0
4	C2483	900	아파트	강남구	51.93	41	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0

2952 rows × 15 columns

	총 세대수	임대 건물 구분	빈 집	노선 수	버스	area10	area20	area30	area40	area50	area60	area70	area80	area100	주차면 수	지역변수
0	545	0	17.0	0.0	3.0	0	0	356	57	132	0	0	0	0	624.0	0.289602
1	1216	0	13.0	1.0	1.0	0	0	390	702	124	0	0	0	0	1285.0	0.283527
2	755	0	6.0	1.0	3.0	0	0	240	0	515	0	0	0	0	734.0	0.304165
3	696	0	14.0	0.0	2.0	0	0	254	196	246	0	0	0	0	645.0	0.320675
4	566	0	9.0	0.0	6.0	0	0	271	209	86	0	0	0	0	517.0	0.320675

1. 변수 설명

컬럼명	전처리 방향	
단지코드	삭제	
총세대수		
임대건물구분	Label Encoding	아파트 : 0, 상가 : 1
지역		노인청소년 비율로 대체
공급유형	삭제	
전용면적	Label Encoding	전용면적 별 카테고리 구분
전용면적별 세대수		공가수 제외한 세대수 컬럼 추가
공가수		
자격유형	삭제	A-Z 자격유형 기준 확인 불가
임대보증금	삭제	상가 및 매매 결측치 대체 불가
임대료	삭제	상가 및 매매 결측치 대체 불가
지하철역		지하철 + 버스 = 대중교통 컬럼 추가
버스정류장		
단지내주차면수		
등록차량수		



	단지코드	총세대수	임대건물구분	지역	공급유형	전용면적	전용면적별 세대수	공가수	자격유형	임대보증금	임대료	도보 10분거리 내 지하철역 수(환승노선 수 반영)	도보 10분거리 내 버스정류장 수	단지내주차면수	등록차량수
0	C2483	900	아파트	경상북도	민	39.72	134	38.0	A	15667000	103680	0.0	3.0	1425.0	1015.0
1	C2483	900	아파트	경상북도	민	39.72	15	38.0	A	15667000	103680	0.0	3.0	1425.0	1015.0
2	C2483	900	아파트	경상북도	민	51.93	385	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0
3	C2483	900	아파트	경상북도	민	51.93	15	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0
4	C2483	900	아파트	경상북도	민	51.93	41	38.0	A	27304000	184330	0.0	3.0	1425.0	1015.0

2952 rows × 15 columns



	총세대수	임대건물구분	빈집	노선수	버스	area10	area20	area30	area40	area50	area60	area70	area80	area100	주차면수	지역변수
0	545	0	17.0	0.0	3.0	0	0	356	57	132	0	0	0	0	624.0	0.289602
1	1216	0	13.0	1.0	1.0	0	0	390	702	124	0	0	0	0	1265.0	0.263527
2	755	0	6.0	1.0	3.0	0	0	240	0	515	0	0	0	0	734.0	0.304165
3	696	0	14.0	0.0	2.0	0	0	254	196	246	0	0	0	0	645.0	0.320675
4	566	0	9.0	0.0	6.0	0	0	271	209	86	0	0	0	0	517.0	0.320675

411 rows × 16 columns

1. 변수 설명

컬럼명	전처리 방향	
단지코드	삭제	
총 세대수		
임대건물구분	Label Encoding	아파트 : 0, 상가 : 1
지역		노인청소년 비율로 대체
공급유형	삭제	
전용면적	Label Encoding	전용면적 별 카테고리 구분
전용면적별 세대수		공가수 제외한 세대수 컬럼 추가
공가수		
자격유형	삭제	A-Z 자격유형 기준 확인 불가
임대보증금	삭제	상가 및 매매 결측치 대체 불가
임대료	삭제	상가 및 매매 결측치 대체 불가
지하철역		지하철 + 버스 = 대중교통 컬럼 추가
버스정류장		
단지내주차면수		
등록차량수		



단지코드	총 세대수	임대건물구분	지역	공급유형	전용면적	전용면적별 세대수	공가수	자격유형	임대보증금	임대료	지하철역	버스정류장	단지내주차면수	등록차량수
0	03403	500	17.0	0.0	3.0	0	0	356	57	132	0	0	0	624.0
1	03403	500	13.0	1.0	1.0	0	0	390	702	124	0	0	0	1285.0
2	03403	500	6.0	1.0	3.0	0	0	240	0	515	0	0	0	734.0
3	03403	500	14.0	0.0	2.0	0	0	254	196	246	0	0	0	645.0
4	03403	500	9.0	0.0	6.0	0	0	271	209	86	0	0	0	517.0

2952 rows × 15 columns



	총 세대수	임대 건물 구분	빈 집	노 선 수	버 스	area10	area20	area30	area40	area50	area60	area70	area80	area100	주차면 수	지역변수
0	545	0	17.0	0.0	3.0	0	0	356	57	132	0	0	0	0	624.0	0.289602
1	1216	0	13.0	1.0	1.0	0	0	390	702	124	0	0	0	0	1285.0	0.283527
2	755	0	6.0	1.0	3.0	0	0	240	0	515	0	0	0	0	734.0	0.304165
3	696	0	14.0	0.0	2.0	0	0	254	196	246	0	0	0	0	645.0	0.320675
4	566	0	9.0	0.0	6.0	0	0	271	209	86	0	0	0	0	517.0	0.320675

411 rows × 16 columns

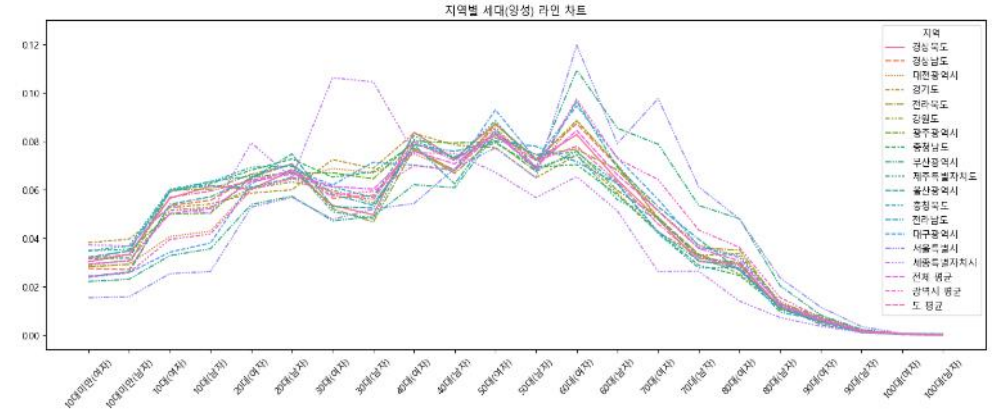
1. 변수 설명

지역별 인구비율 - 한국 토지주택공사

지역	10대미만 (여자)	10대미만 (남자)	10대(여 자)	10대(남 자)	20대(여 자)	20대(남 자)	30대(여 자)	30대(남 자)	40대(여 자)	40대(남 자)	...
경상북도	0.030158	0.0333195	0.056346	0.061360	0.060096	0.067859	0.053433	0.049572	0.083660	0.072613	...
경상남도	0.027400	0.026902	0.053257	0.055568	0.064920	0.070618	0.056414	0.057550	0.077092	0.067600	...
대전광역시	0.028197	0.029092	0.040490	0.042793	0.060834	0.064247	0.068654	0.066848	0.074667	0.067925	...
경기도	0.038030	0.039507	0.052546	0.053990	0.058484	0.059894	0.072331	0.068704	0.083208	0.078355	...
전라북도	0.028089	0.029065	0.059685	0.060080	0.066262	0.070322	0.052027	0.046596	0.077005	0.066645	...

5 rows × 22 columns

지역	노인청소년
경상북도	0.306899
경상남도	0.289602
대전광역시	0.283527
경기도	0.304165
전라북도	0.320675
강원도	0.300243
광주광역시	0.294664
충청남도	0.298290
부산광역시	0.323989
제주특별자치도	0.321988
울산광역시	0.296999
충청북도	0.306738
전라남도	0.307037
대구광역시	0.265155
서울특별시	0.327542
세종특별자치시	0.255091



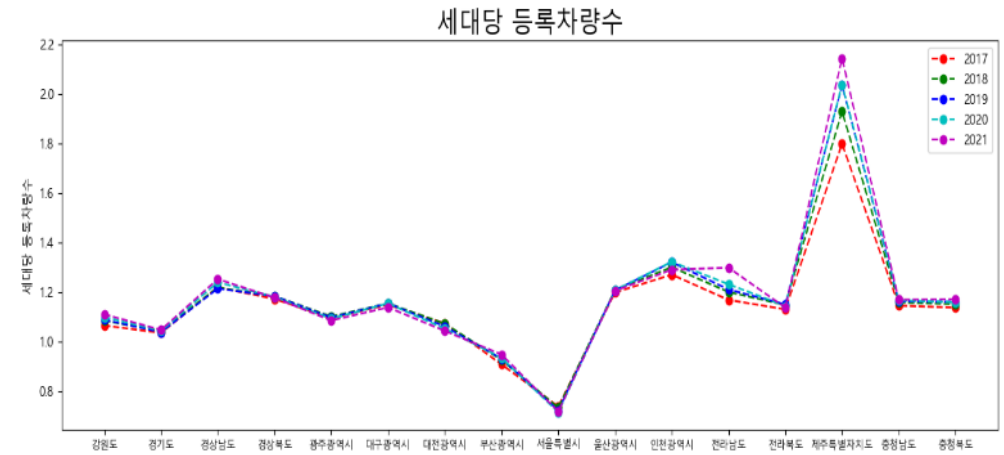
지역 변수 -> 노인청소년 비율로 대체

지역별 세대수 및 자동차 등록대수 - 공공 데이터 포털

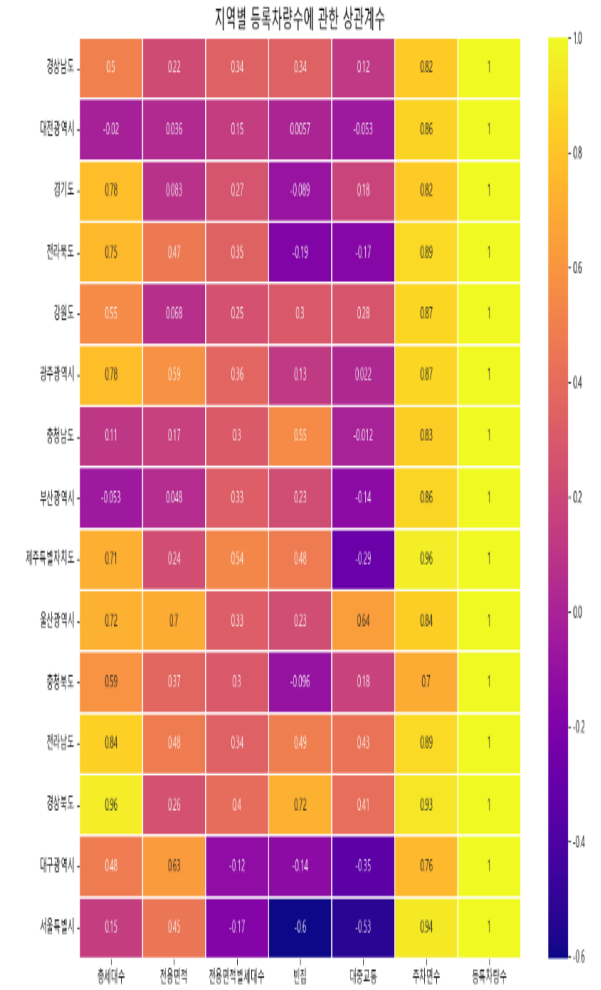
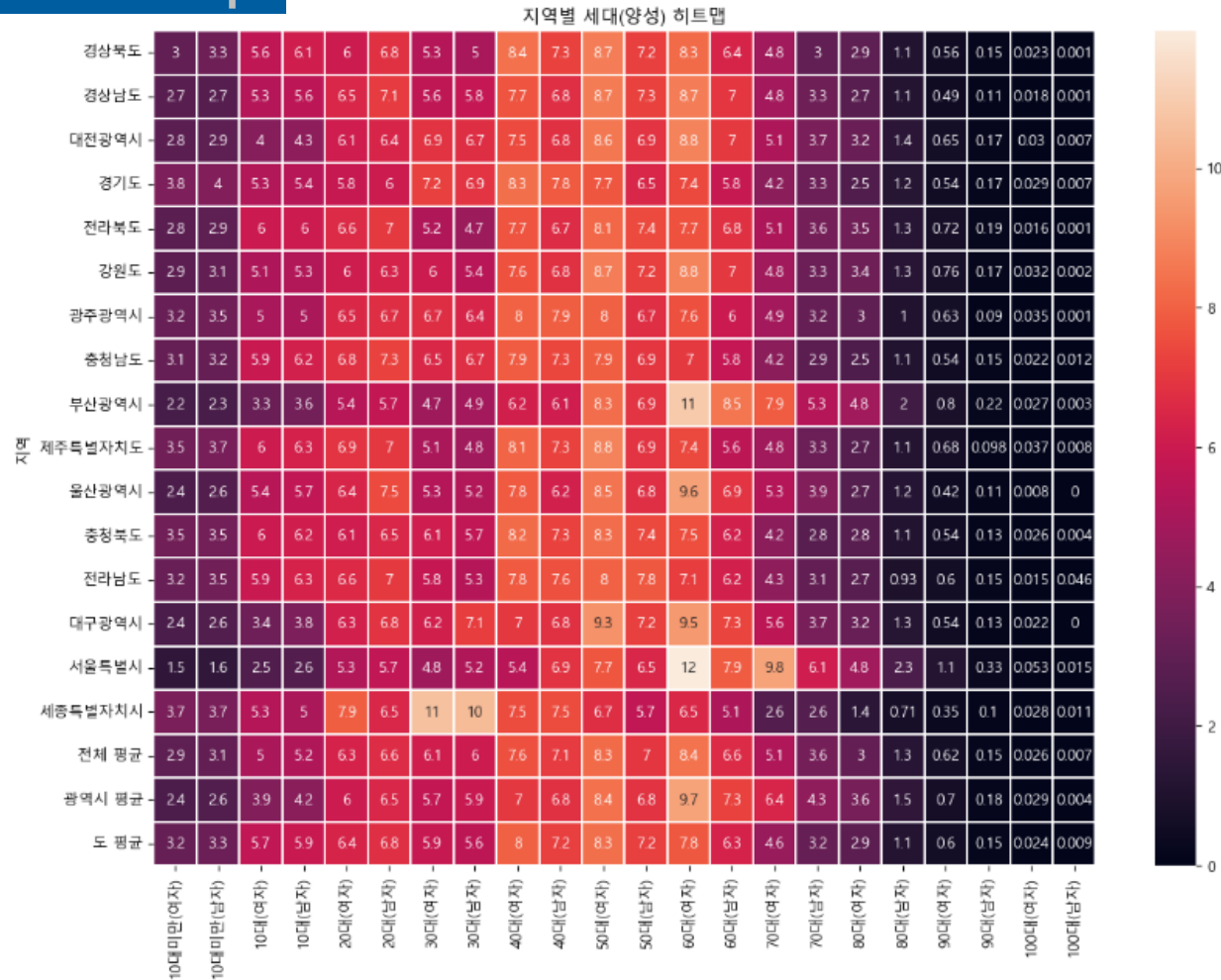
통계연도		시도명	시군구명	세대당 자동차등록(대)	주민등록세대수(명)	자동차등록대수(대)
0	2017	강원도	강릉시	1.09	95963	104693
1	2017	강원도	고성군	0.91	15519	14076
2	2017	강원도	동해시	1.07	40730	43647
3	2017	강원도	삼척시	0.91	33863	30889
4	2017	강원도	속초시	0.97	37772	36601
...	...	...	...	...	...	...
1295	2021	인천광역시	연수구	1.34	155984	208616
1296	2021	인천광역시	옹진군	1.05	12061	12666
1297	2021	인천광역시	미추홀구	0.88	193561	169809
1298	2021	제주특별자치도	제주시	2.50	219978	550432
1299	2021	제주특별자치도	서귀포시	1.24	87551	108162

1300 rows × 6 columns

시도명	세대수	차량등록대수	세대당등록수
강원도	746220	828146	1.110
경기도	8172214	8554726	1.047
경상남도	1957911	2451903	1.252
경상북도	1509047	1776477	1.177
광주광역시	645712	700783	1.085
대구광역시	1063893	1211095	1.138
대전광역시	664417	692702	1.043
부산광역시	1544663	1464608	0.948
서울특별시	4426007	3176743	0.718
울산광역시	482650	581429	1.205
인천광역시	1298647	1675405	1.290
전라남도	903108	1172282	1.298
전라북도	1142208	1298250	1.137
제주특별자치도	307529	658594	2.142
충청남도	1299902	1519895	1.169
충청북도	1147688	1344394	1.171

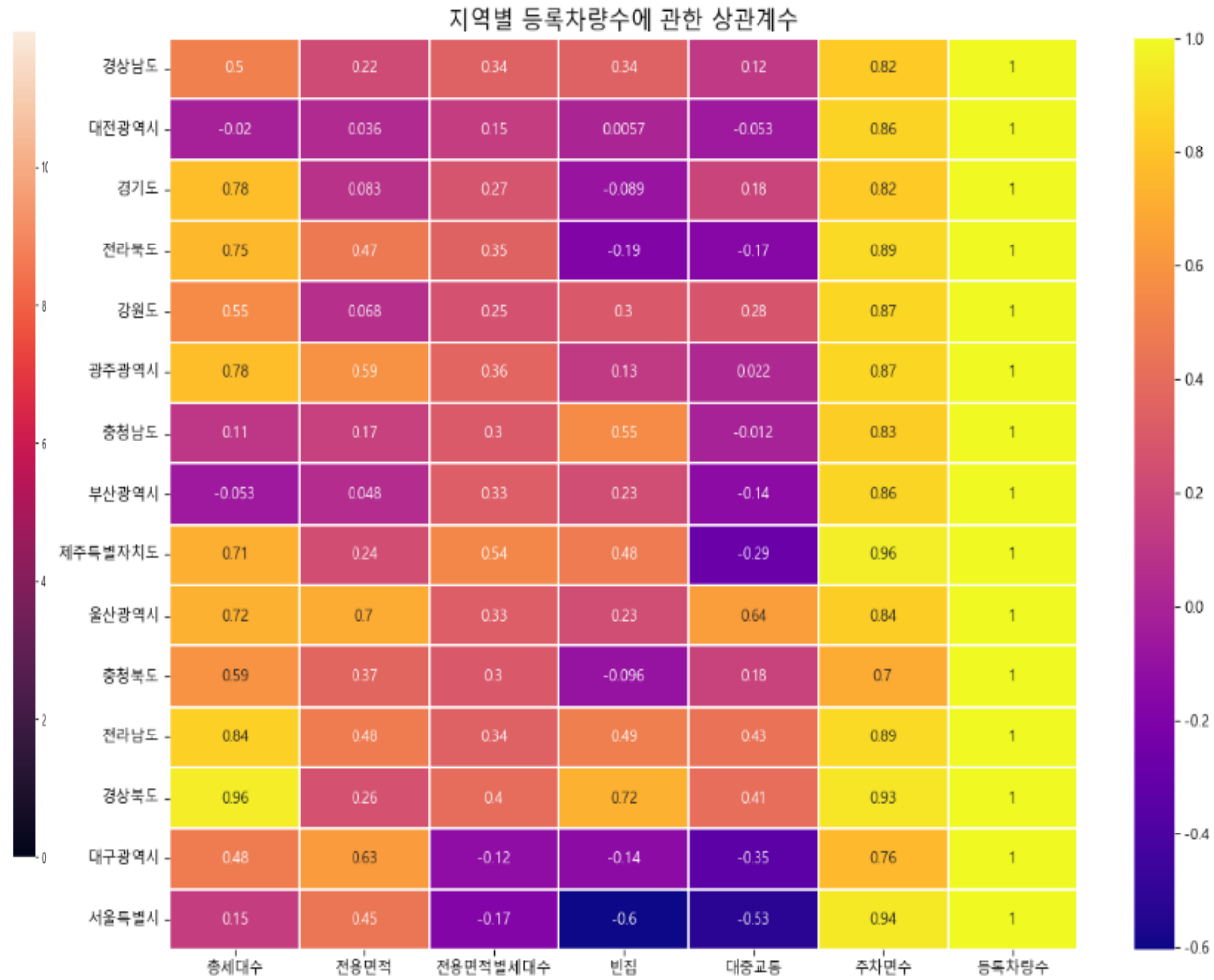
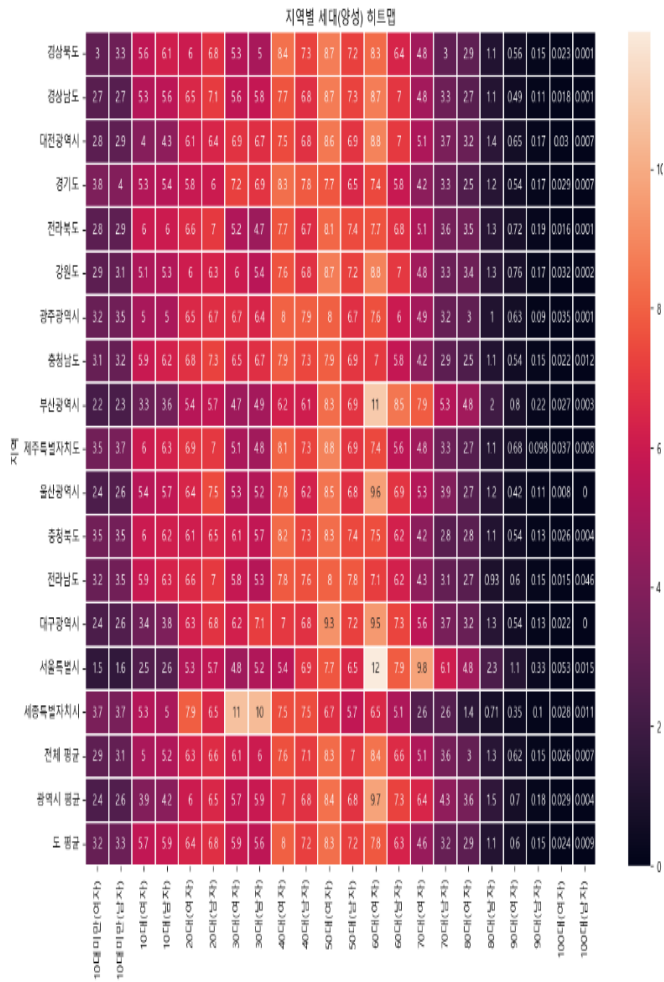


지역별 나이 Heat Map

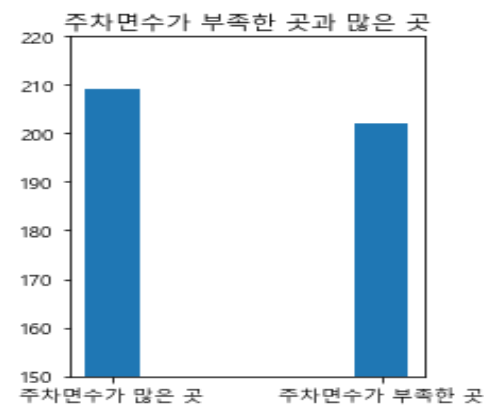
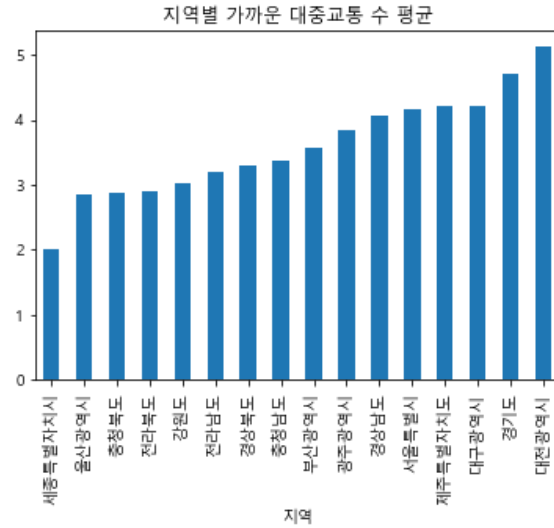
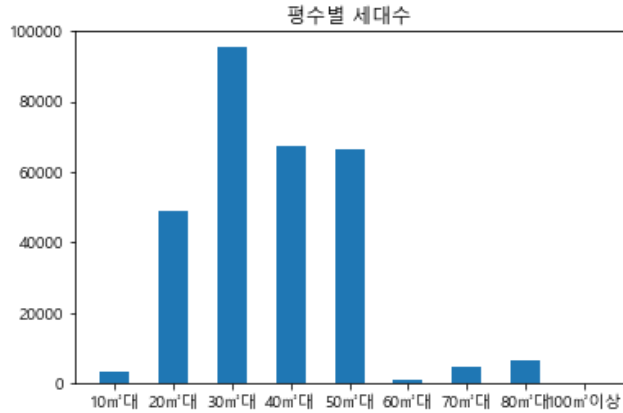


2. Data EDA

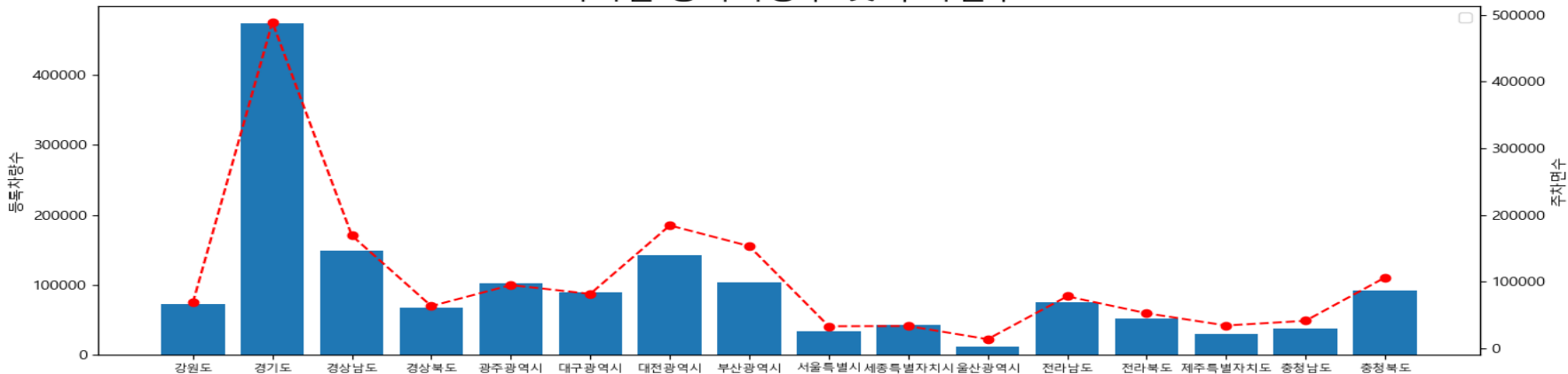
지역별 등록 차량 수에 관한 상관관계수



Part2 2. Data EDA – 데이터 분포 확인



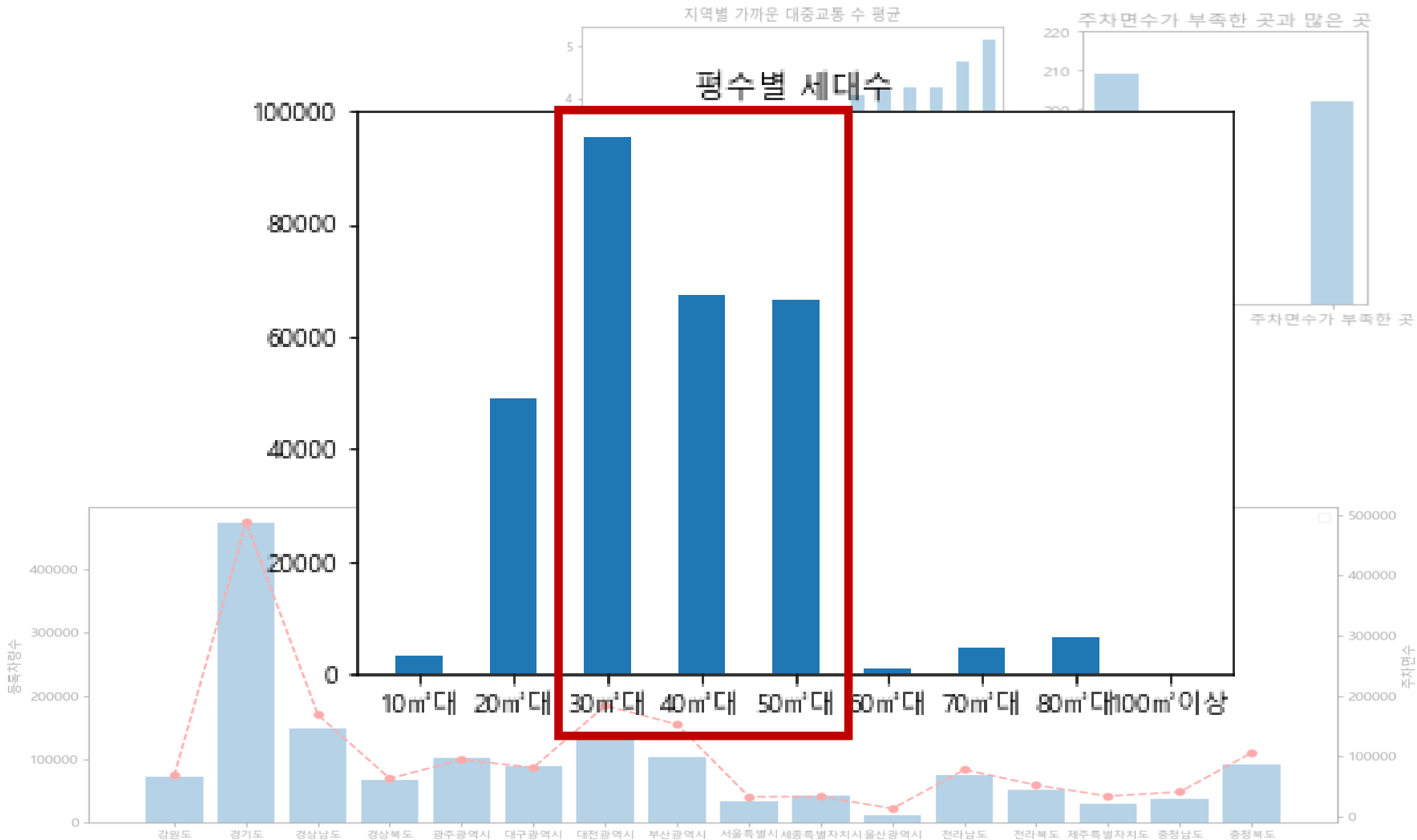
지역별 등록차량수 및 주차면수



데이터 분포 확인

- 30m²~50m² 평수가 가장 많다.
- 지역별 대중교통 수 분포 확인.
- 등록 차량 수 대비 주차 면 수 많은 곳 적은 곳 비슷하게 분포.
- 지역별 등록차량이 많으면 주차 면수도 많아짐.

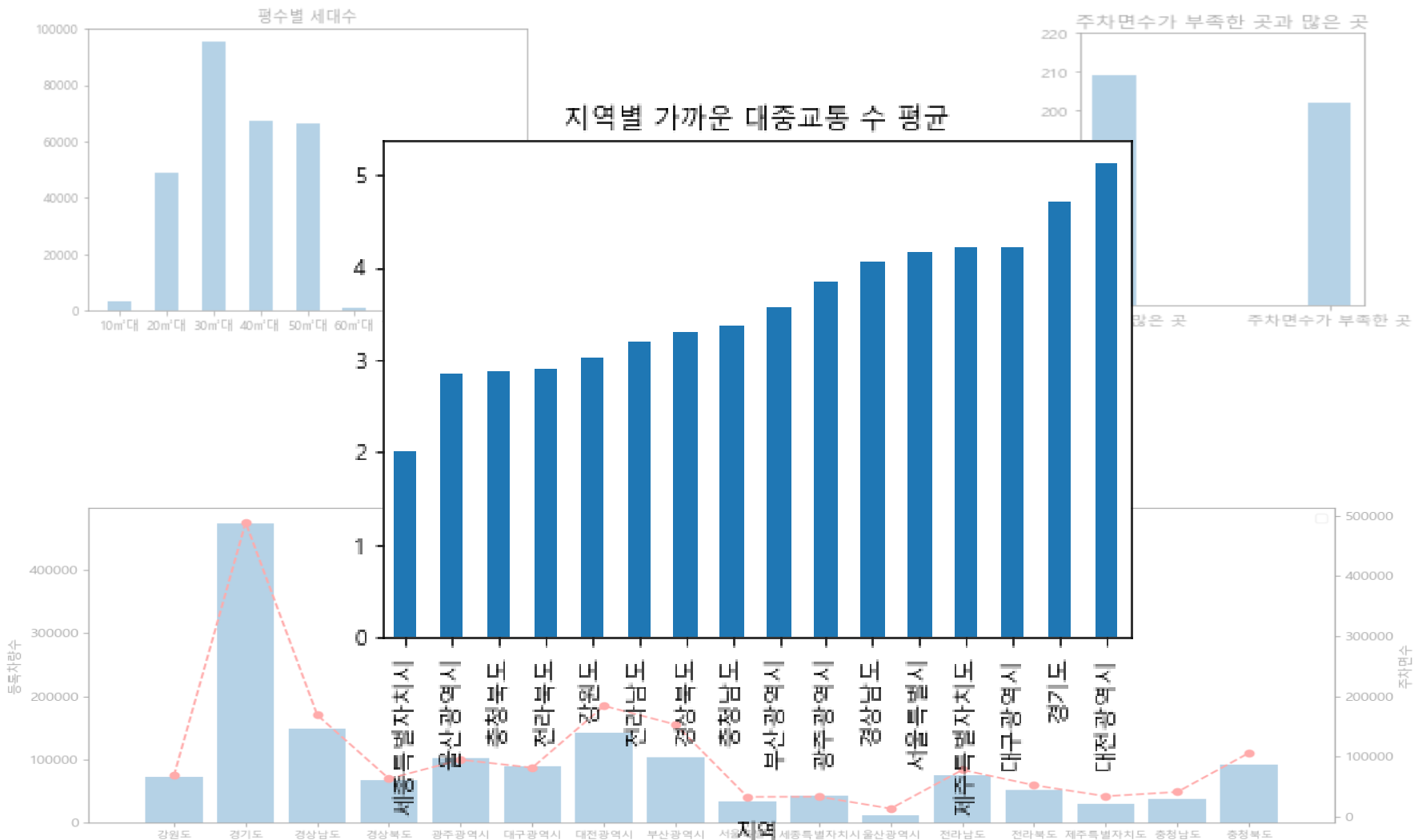
Part2 2. Data EDA – 데이터 분포 확인



데이터 분포 확인

- **30m²~50m² 평수가 가장 많다.**
- 지역별 대중교통 수 분포 확인.
- 등록 차량 수 대비 주차 면 수 많은 곳 적은 곳 비슷하게 분포.
- 지역별 등록차량이 많으면 주차 면 수도 많아짐

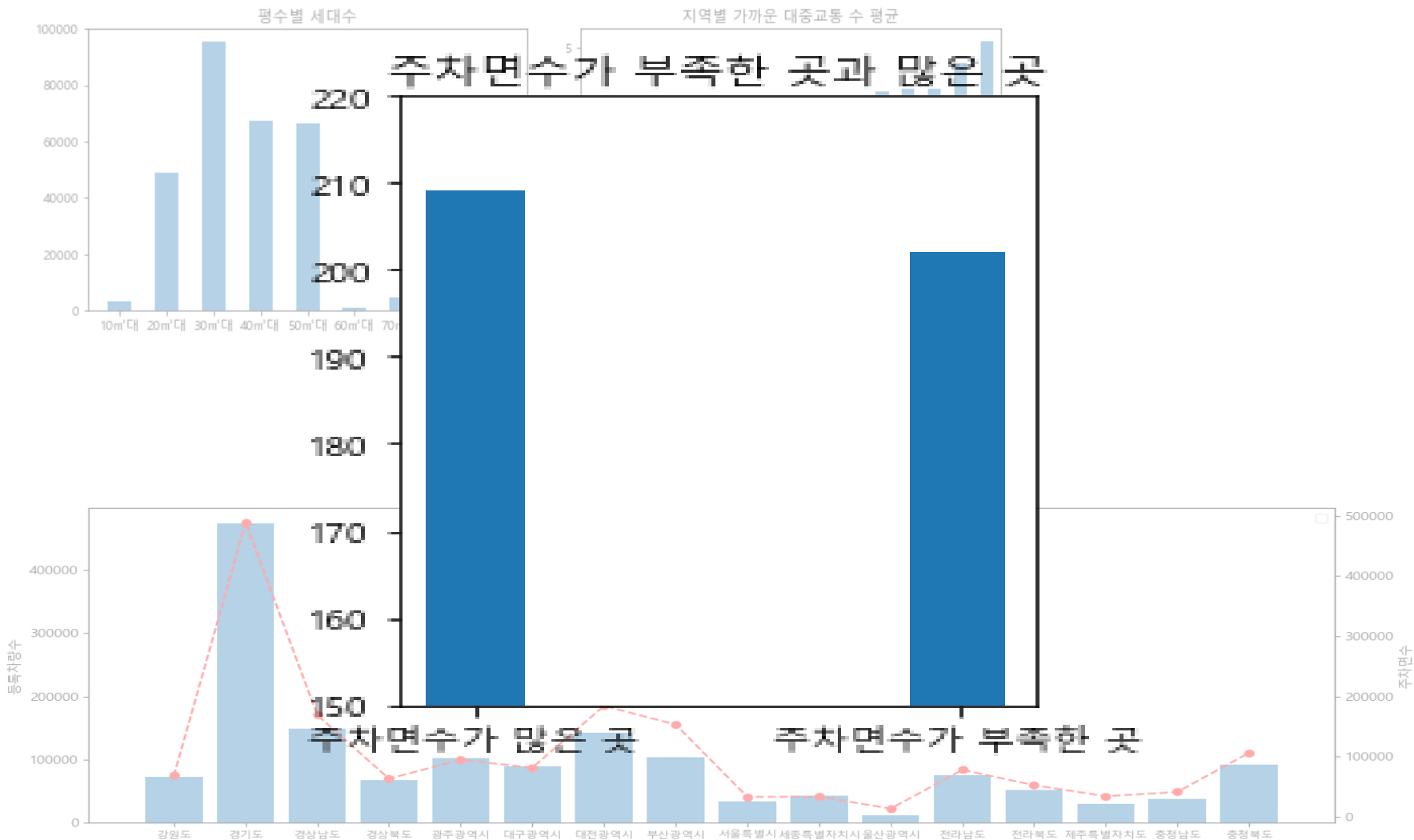
Part2 2. Data EDA – 데이터 분포 확인



데이터 분포 확인

- 30㎡~50㎡ 평수가 가장 많다.
- 지역별 대중교통 수 분포 확인.
- 등록 차량 수 대비 주차 면 수 많은 곳 적은 곳 비슷하게 분포.
- 지역별 등록차량이 많으면 주차 면수도 많아짐.

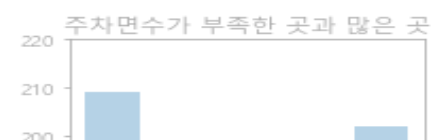
Part2 2. Data EDA-데이터 분포 확인



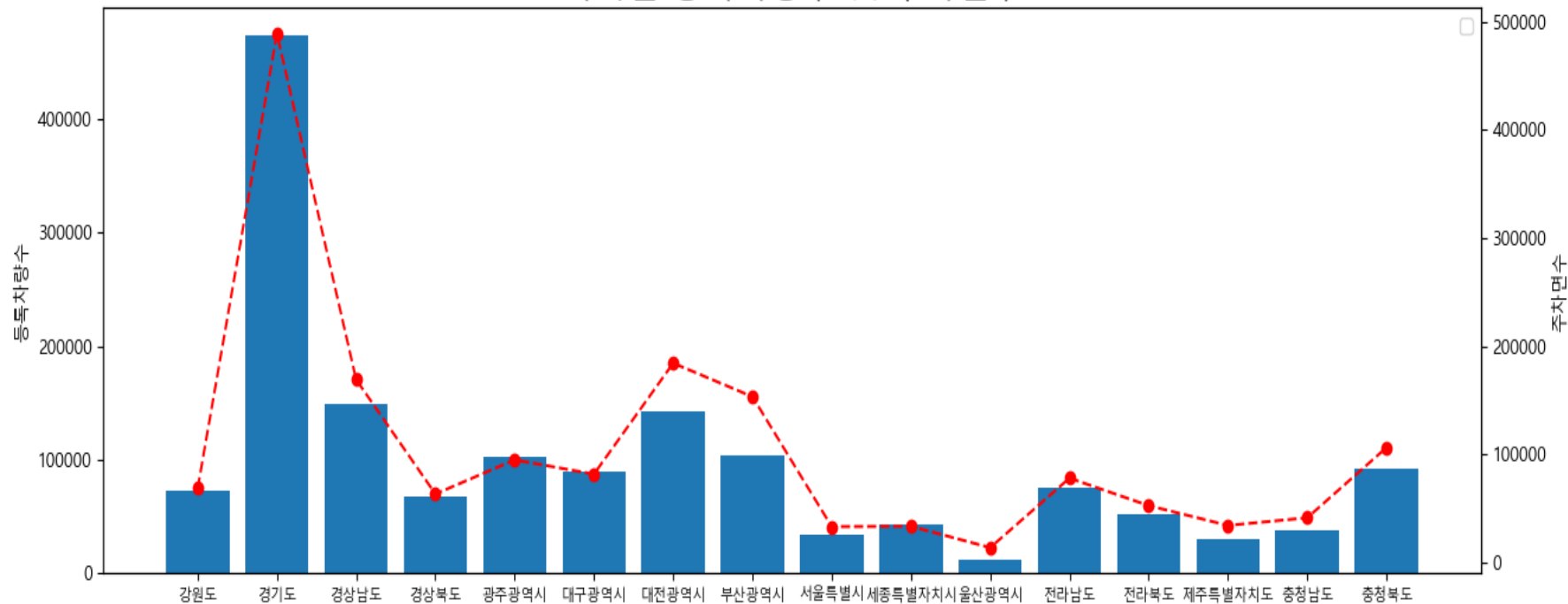
데이터 분포 확인

- 30㎡~50㎡ 평수가 가장 많다.
- 지역별 대중교통 수 분포 확인.
- 등록 차량 수 대비 주차 면 수 많은 곳 적은 곳 비슷하게 분포.
- 지역별 등록차량이 많으면 주차 면수도 많아짐.

Part2 2. Data EDA-데이터 분포 확인



지역별 등록차량수 및 주차면수

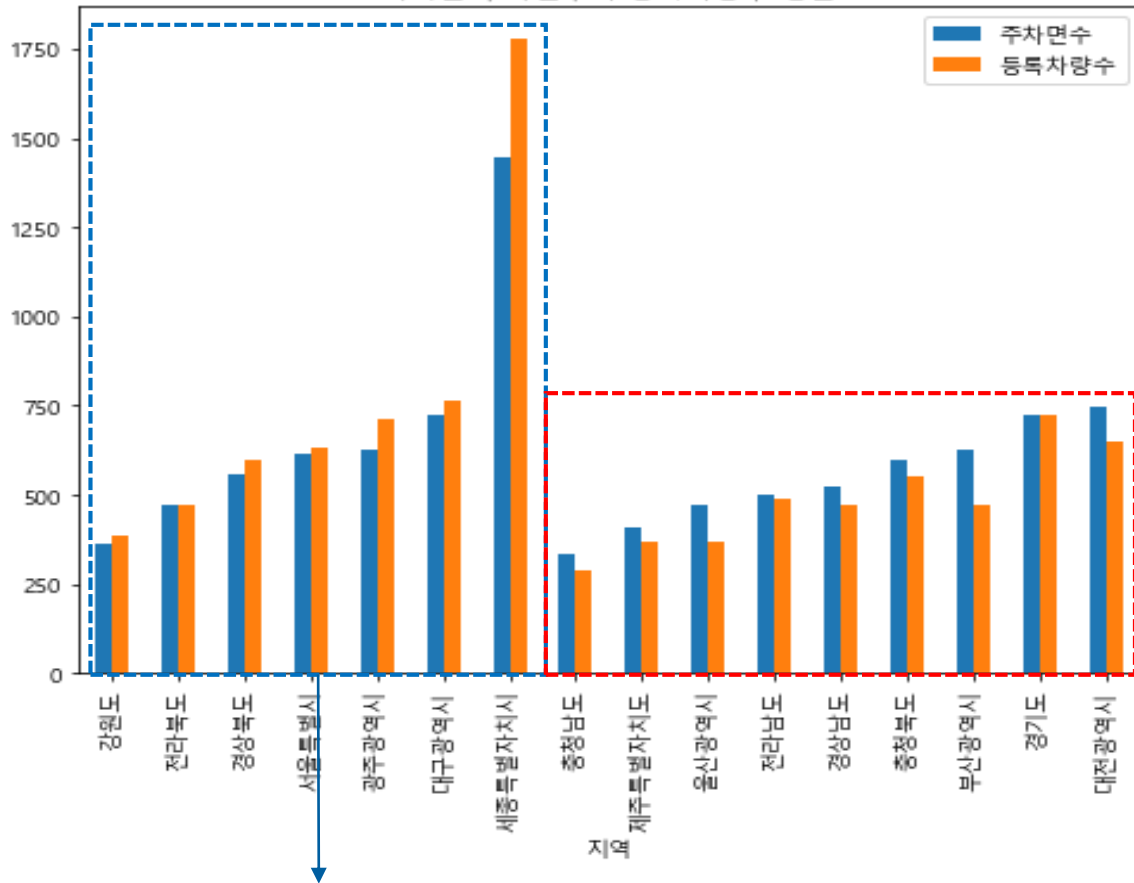


데이터 분포 확인

- 30㎡~50㎡ 평수가 가장 많다.
- 지역별 대중교통 수 분포 확인.
- 등록 차량 수 대비 주차 면 수 많은 곳 적은 곳 비슷하게 분포.
- 지역별 등록차량이 많으면 주차 면수 도 많아짐.

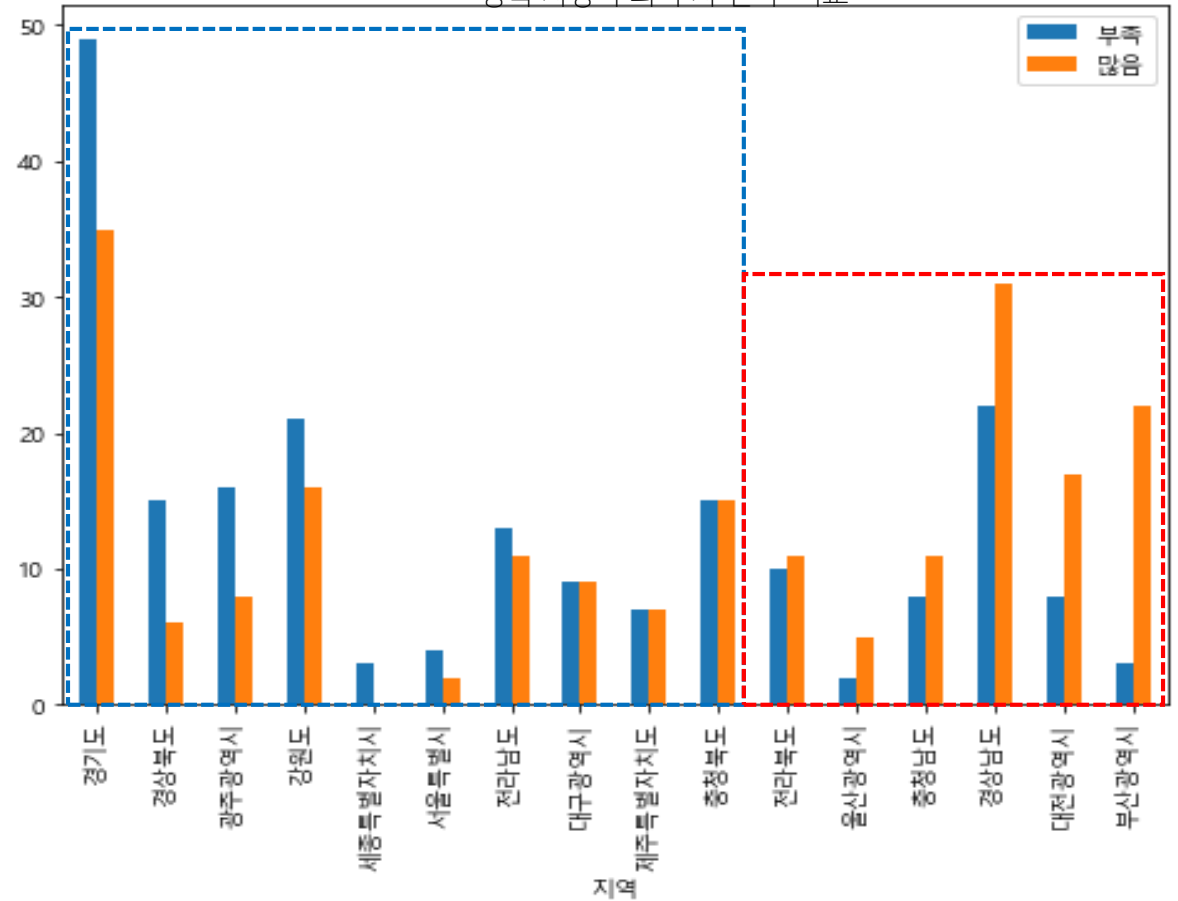
2. Data EDA – 지역별 주차 면 수 vs 등록 차량 수

지역별 주차면수와 등록차량수 평균

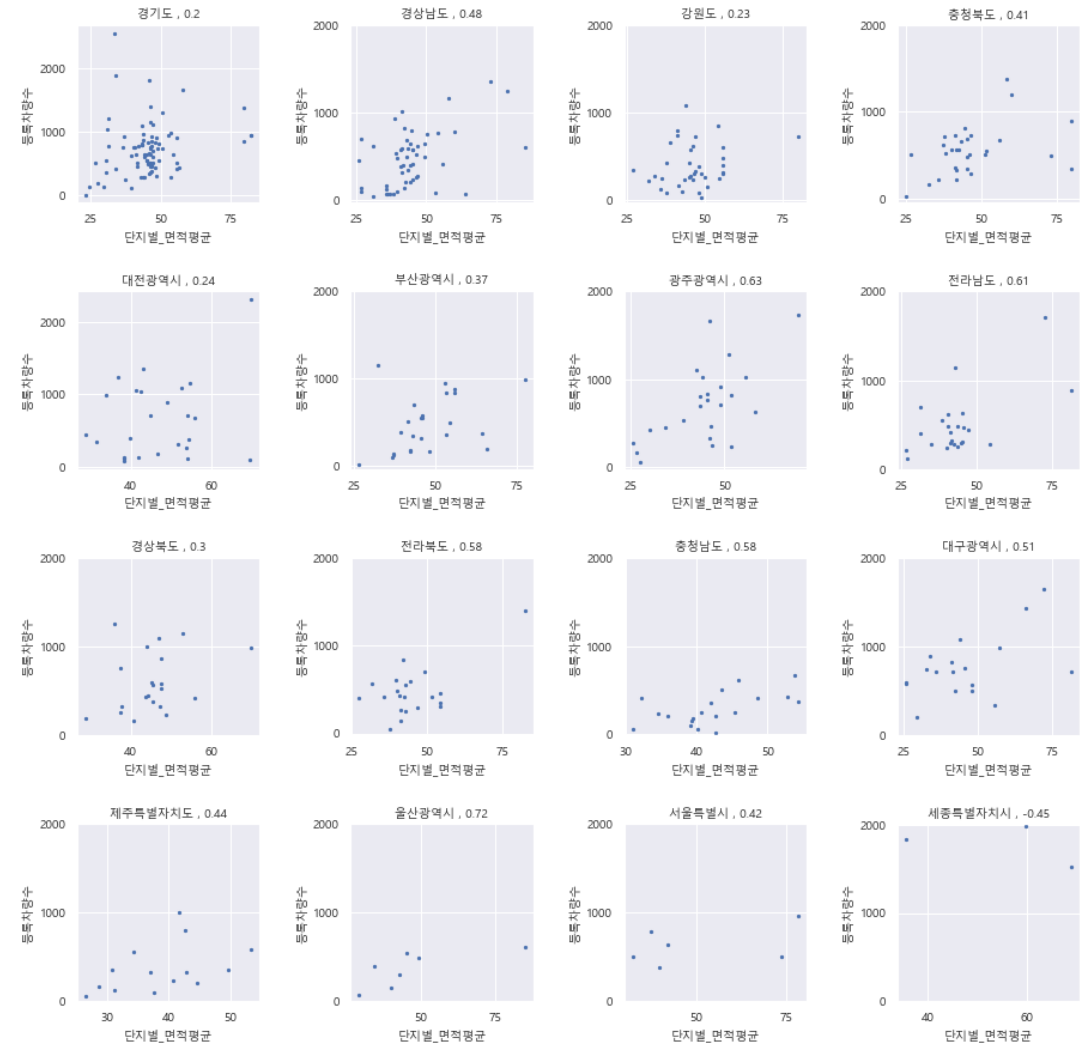
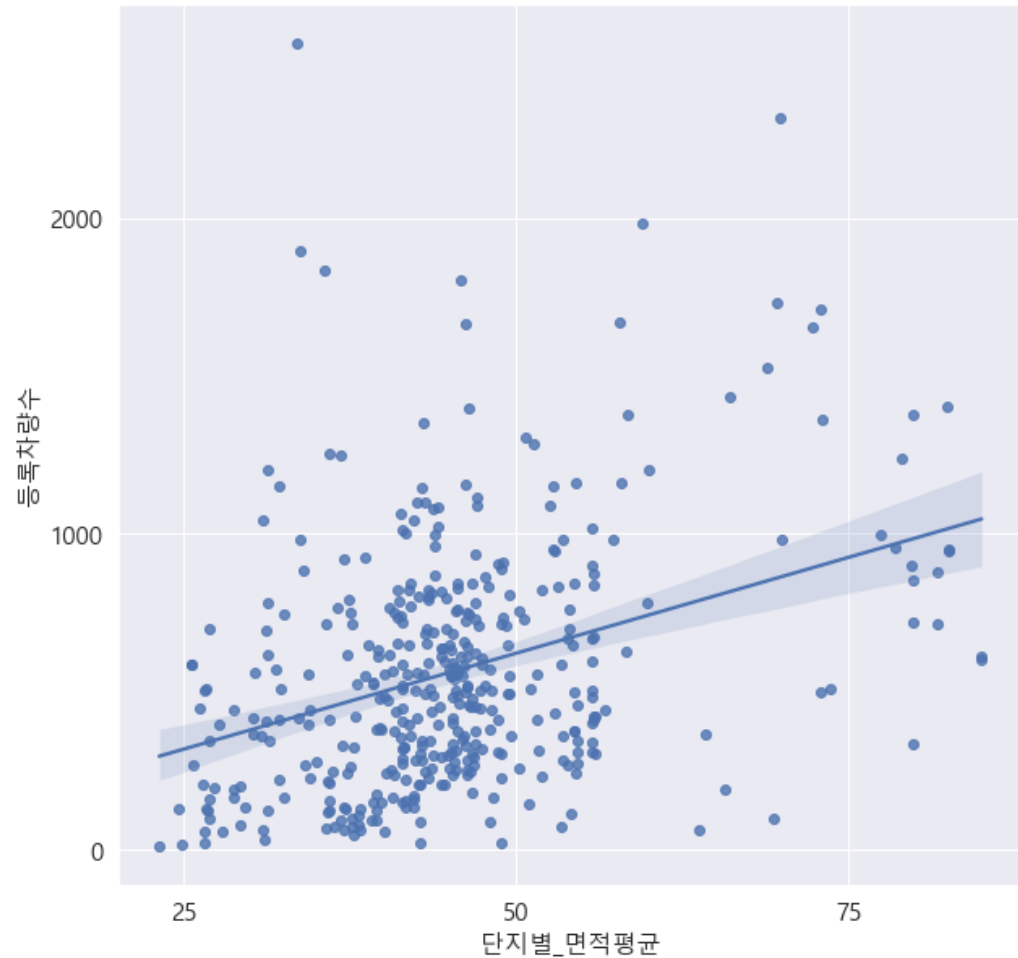


주차 면 수 < 자동차 등록대수 (주차 면 수 부족한 지역 44%)

등록 차량 수와 주차 면 수 비교



2. Data EDA – 지역별 면적 – 등록대수 상관관계



Part2 3. Machine Learning

훈련 데이터 변수 설정

구분		전처리
1	지하철 + 버스	대중 교통 변수 추가
2	지역	노인 청소년 인구비율로 대체
3	소형 평수 구분	면적 40㎡ 미만 세대수 비율
4	법정 주차 대수	전용면적 * 세대수 / 기준면적

훈련 모델 :

- Linear Regression
- Random Forest
- Ridge
- Lasso

오차 및 점수 :

- RMSE (Mean Squard Error)
- Score

1+2 : 대중교통 칼럼 추가 / 지역 변수 수정

구분	RMSE	Train_SCORE	Val_SCORE
Linear Regression	217.8458	0.7887	0.7675
Random Forest	236.2836	0.9621	0.6097
Ridge	183.1654	0.8215	0.8087
Lasso	180.7804	0.8275	0.8151

3 : 소형 평수 구분 추가

구분	RMSE	Train_SCORE	Val_SCORE
Linear Regression	173.8274	0.7811	0.7675
Random Forest	210.1943	0.9568	0.6799
Ridge	178.1784	0.8159	0.7856
Lasso	180.5819	0.8204	0.8141

4 : 법정 주차 대수 변수 추가

구분	RMSE	Train_SCORE	Val_SCORE
Linear Regression	166.8274	0.8235	0.7875
Random Forest	157.1412	0.9621	0.7759
Ridge	177.1784	0.8170	0.7849
Lasso	176.8833	0.8167	0.7856

Part2 4. Deep Learning – Modeling

Model 1

flatten_input	input:	[(None, 15)]
InputLayer	output:	[(None, 15)]

flatten	input:	(None, 15)
Flatten	output:	(None, 15)

입력층

dense	input:	(None, 15)
Dense	output:	(None, 64)

은닉층

dense_1	input:	(None, 64)
Dense	output:	(None, 64)

Hidden Layer : 2*2

Node : 64

Optimizer : SGD

Epoch : 500

dropout	input:	(None, 64)
Dropout	output:	(None, 64)

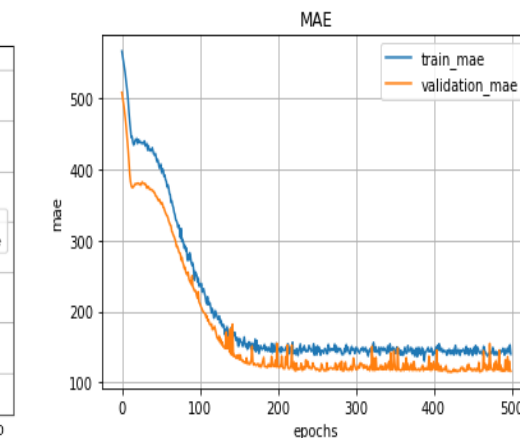
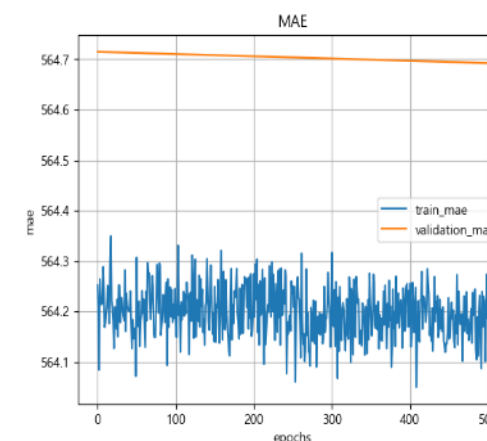
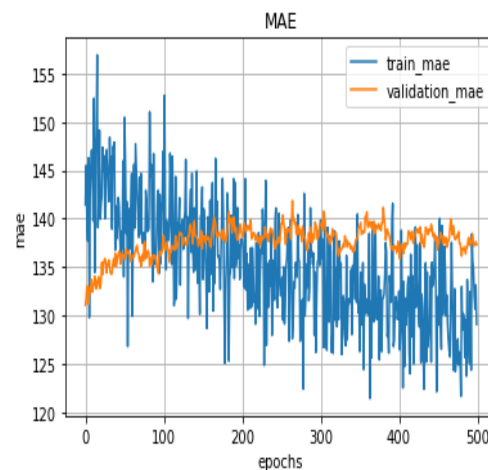
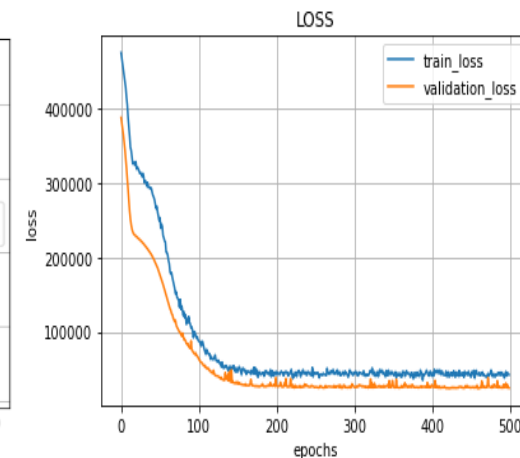
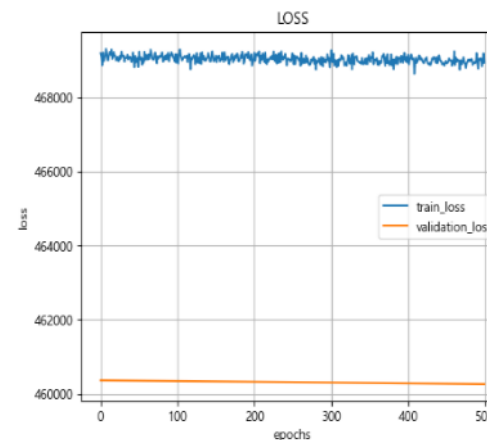
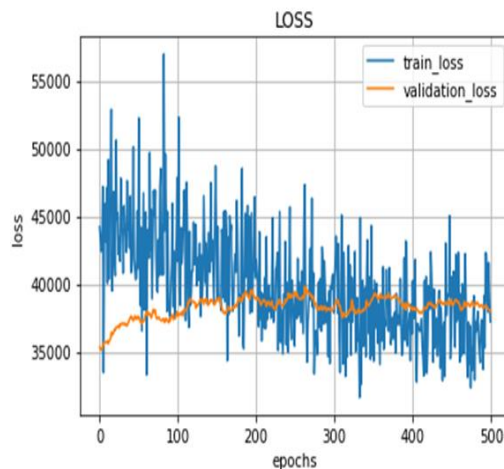
dense_2	input:	(None, 64)
Dense	output:	(None, 64)

dense_3	input:	(None, 64)
Dense	output:	(None, 64)

dropout_1	input:	(None, 64)
Dropout	output:	(None, 64)

dense_4	input:	(None, 64)
Dense	output:	(None, 1)

출력층



Part2 4. Deep Learning – Modeling

Model 2

flatten_12_input	input:	[(None, 15)]
InputLayer	output:	[(None, 15)]

flatten_12	input:	(None, 15)
Flatten	output:	(None, 15)

dense_78	input:	(None, 15)
Dense	output:	(None, 128)

dense_79	input:	(None, 128)
Dense	output:	(None, 128)

dense_80	input:	(None, 128)
Dense	output:	(None, 128)

dropout_24	input:	(None, 128)
Dropout	output:	(None, 128)

dense_81	input:	(None, 128)
Dense	output:	(None, 128)

dense_82	input:	(None, 128)
Dense	output:	(None, 128)

dropout_25	input:	(None, 128)
Dropout	output:	(None, 128)

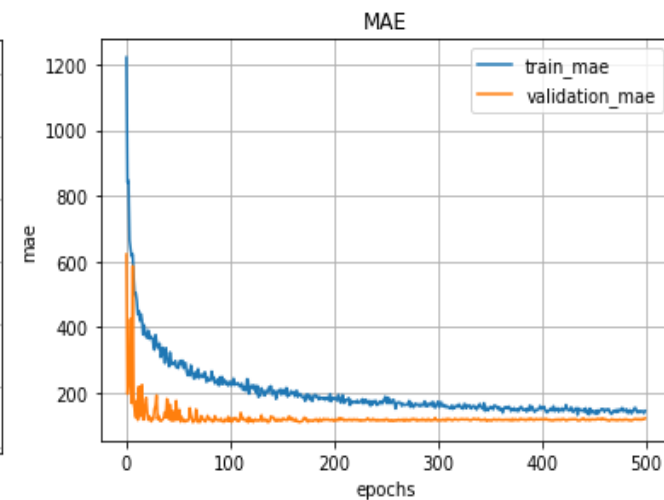
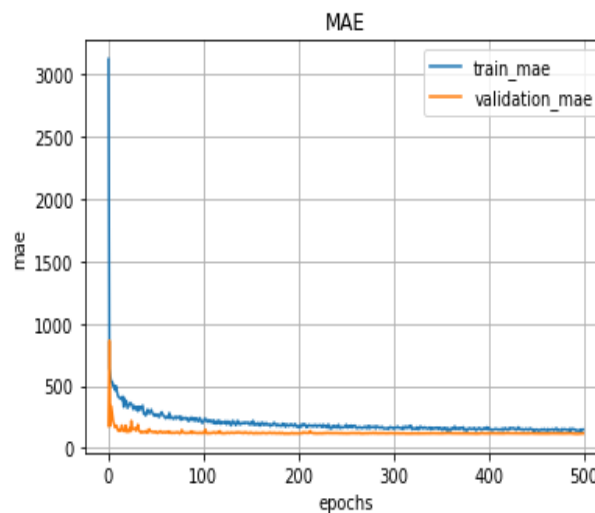
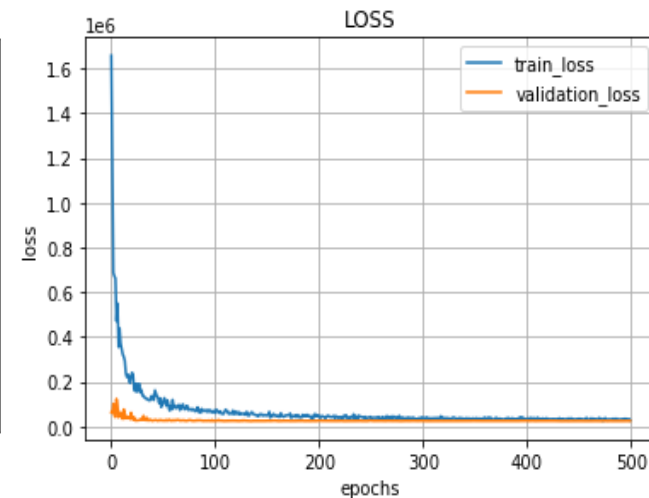
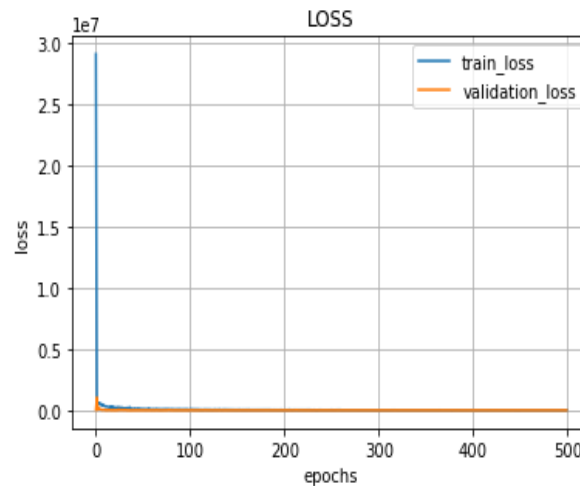
dense_83	input:	(None, 128)
Dense	output:	(None, 1)

입력층

은닉층

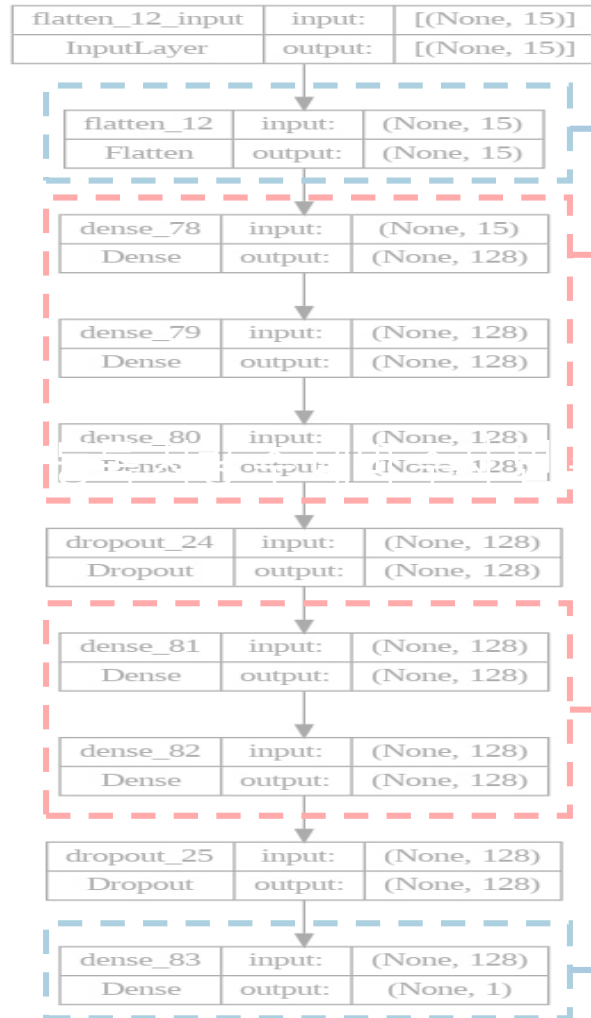
Hidden Layer : 3*2
Node : 128
Optimizer : SGD
Epoch : 500

출력층



Part2 4. Deep Learning – Modeling

Model 2

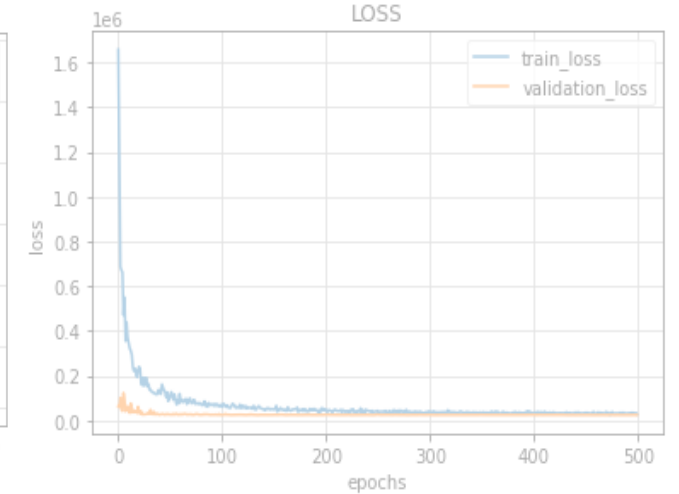
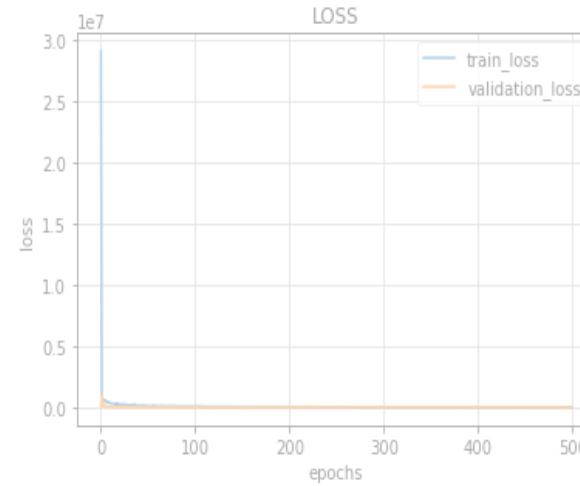


입력층

은닉층

Hidden Layer : 3*2
Node : 128
Optimizer : SGD
Epoch : 500

출력층

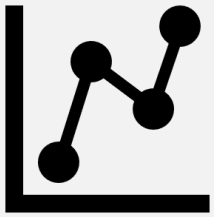


	train_RMSE	val_RMSE	train_MAE	val_MAE
Model1	189.2599	156.6007	163.8760	129.3510
model2	170.8627	141.1320	139.8752	114.3767

Part 3, 결론 및 아쉬운 점



객관적 수요예측



검사원님 이제 편해지세요!

건축허가기준



○○동네 신축아파트 주차장은
이정도는 돼야 합니다!

수익사업 응용가능



주차정보 필요하신 분?!

데이터수집



전국의 모든 아파트의 주차장을 알고 싶다!

손실함수제한적사용

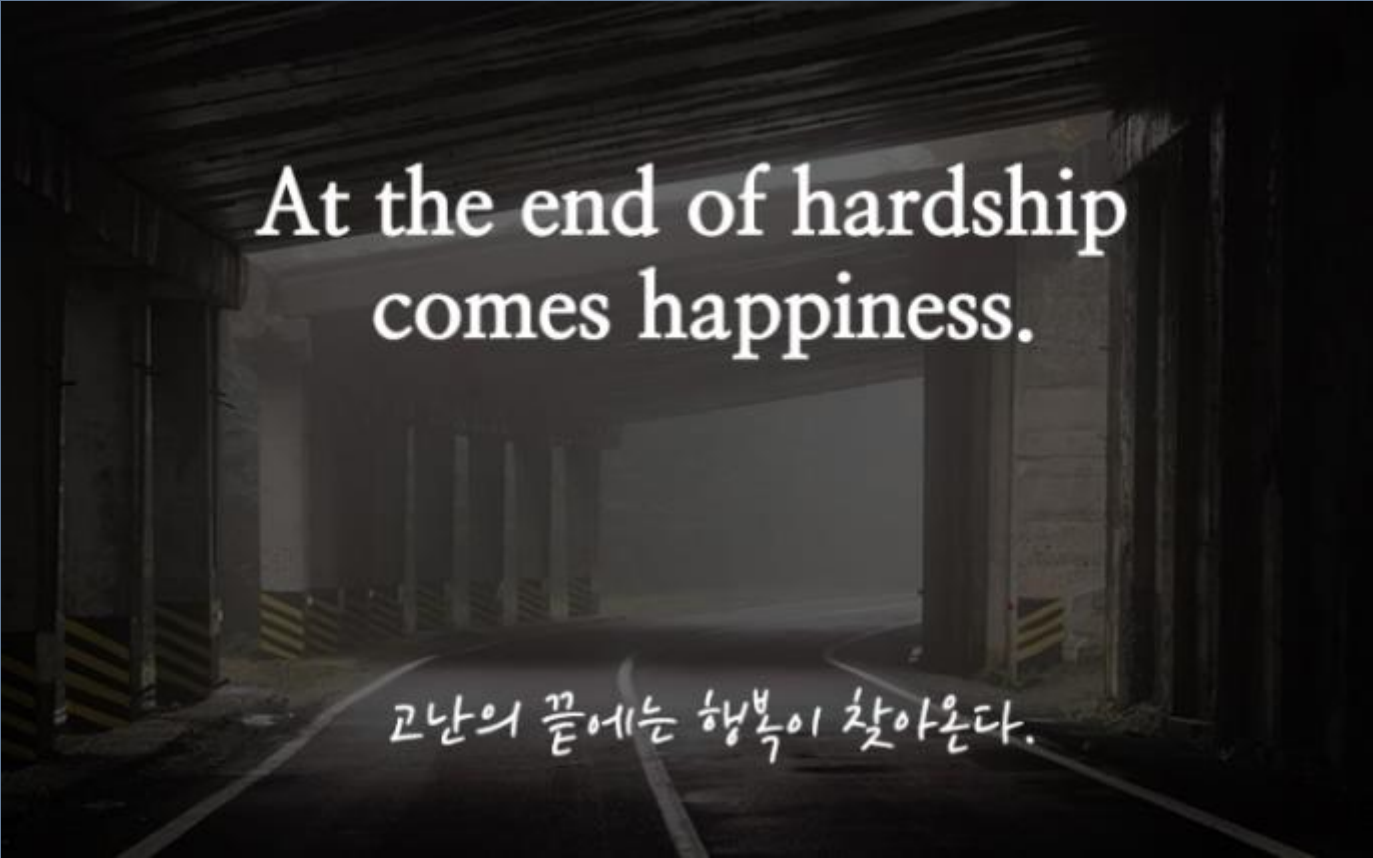


RMSE vs MAE

만나지못한팀원



목소리도 못 들어본 그...
잘가라!

A dark, misty tunnel with concrete pillars and a road leading to a bright exit. The text "At the end of hardship comes happiness." is overlaid in white serif font.

At the end of hardship
comes happiness.

고난의 끝에는 행복이 찾아온다.

감사합니다