



첨단물류 복합 주유소 입지 선정



20162245 김정준

20173601 정재현

20182472 김지섭

20182543 한지성

Contents.

서론

- 주제선정이유
- 감소원인
- 주유소의 현황

본론

- 데이터 수집 및 전처리
- 클러스터링
- K-Means
- K-medoids
- Hierarchical Clustering
- GMM Clustering
- Voting
- 입지 기준 도출

결론

- 최종 주유소
 - 기대 효과 / 한계점 / 발전 방향
 - 분석 활용 데이터 / 분석 Tool
-

탄소중립 추구

서울시 기후 변화 대응
2050 탄소 중립 종합계획 발표

온실가스 감축
서울시 2026년까지 온실가스
30% 감축 계획



30%

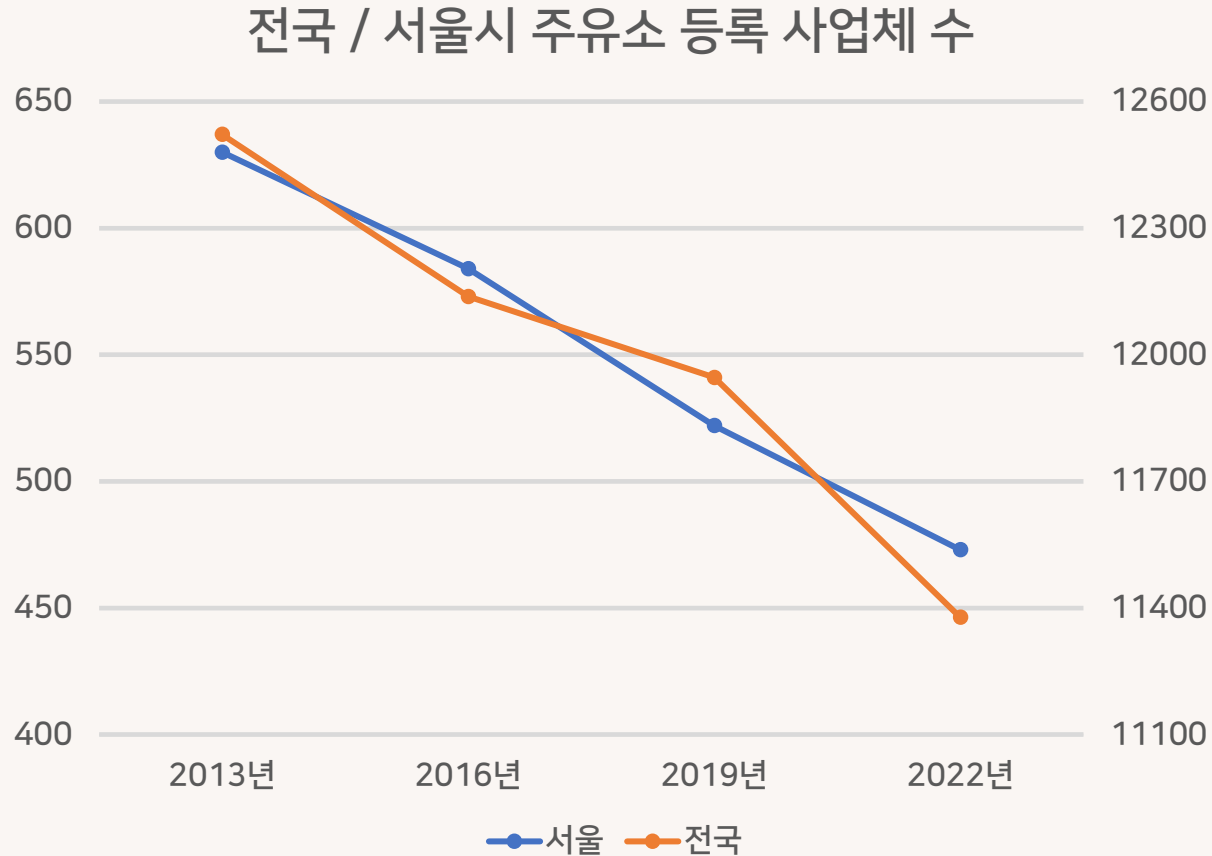


88%

교통+건물의 배출량
서울시 전체 온실가스 배출량
88%는 건물+교통부문

환경문제가 대두되면서 탄소중립을 향한 움직임 활발

주유소 사업체 현황



전국 주유소 개수 추이

- 2013년 12,522개 / 2022년 11,378개로 감소

서울시 주유소 개수 추이

- 2013년 630개 / 2022년 473개로 감소

선정 이유

현재 전국적으로 감소하는 추세
탄소 중립 등 다양한 원인으로 수익성 감소
따라서 대응 방법 모색 / 생존 전략 수립 필요

신차 시장에서의 친환경 차량의 전망

친환경 차량 종류 ▼

하이브리드차

전기차

수소차

<표 4> 국내 친환경자동차 전망

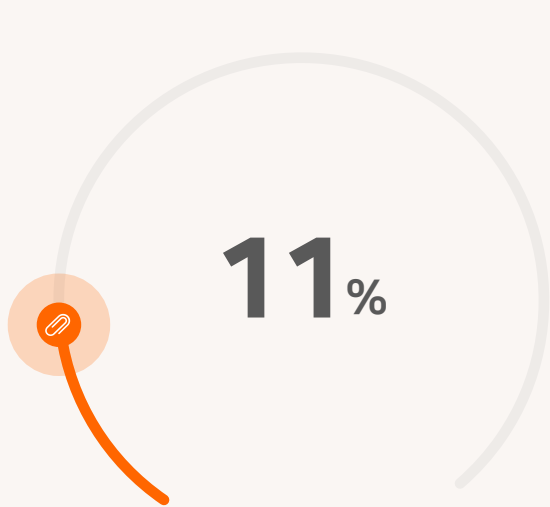
(단위 : 만 대)

구분		2015	2019	2020	2025	2030
친환경차	하이브리드차	3.9	10.4	17.3	78.0	105.0
	전기차	0.3	3.5	4.7	20.0	60.0
	수소차	0.0	0.4	0.6	7	15
	소 계	4.2	14.4	22.6	91	150
내연기관		180.8	165.6	169.4	89.0	30.0
총 계		185	180	192	180.0	180.0
비율	친환경차	2.3%	8.0%	11.8%	50.6%	83.3%
	내연기관	97.7%	92.0%	88.2%	49.4%	16.7%

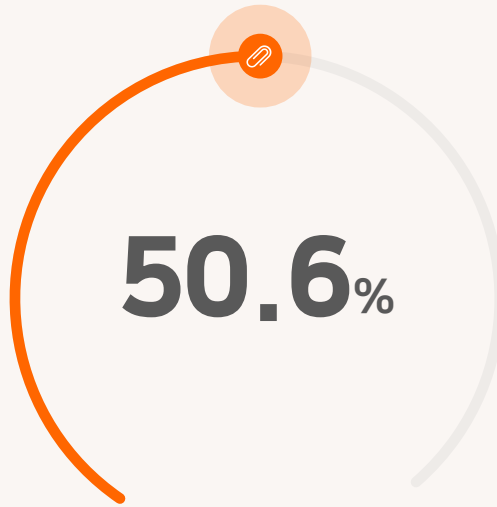
자료 : 2015~19(국토부 차량 등록대수 참조), 2020~30(한국자동차연구원 전망치).

한국노동연구원에 따르면 전체 신차 시장에서 친환경 차량의 비중은
2025년 약 91만 대 (50.6%) / 2030년 약 150만 대(83.3%) 달성할 것으로 전망

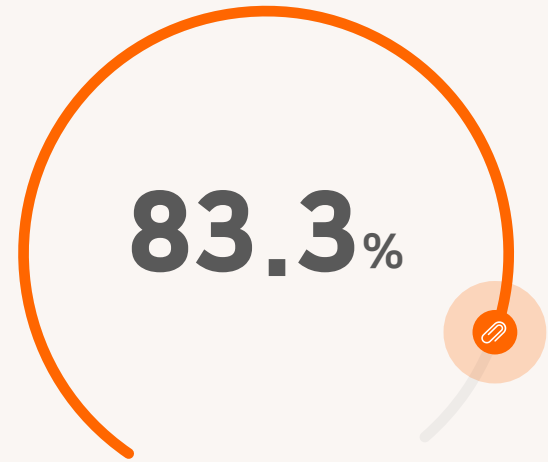
신차 시장에서의 친환경 차량의 전망



2020년 친환경 차량



2025년 친환경 차량



2030년 친환경 차량

▲ 신차 시장 내 친환경 자동차 비율
신규 친환경 차량 판매량 / 신차 총 판매량

친환경 차량 누적 판매량

<친환경차 등록 추이>

(단위 : 대, %)

구 분	'16말	'17말	'18말	'19말	'20말	'21.6월	'21말	'22.9월
전체 자동차	21,803,351	22,528,295	23,202,555	23,677,366	24,365,979	24,642,251	24,911,101	25,355,938
하이브리드	233,216	313,856	405,084	506,047	674,461	785,143	908,240	1,103,642
전기차	10,855	25,108	55,756	89,918	134,962	173,147	231,443	347,395
수소차	87	170	893	5,083	10,906	15,225	19,404	26,719
친환경차	244,158	339,134	461,733	601,048	820,329	973,515	1,159,087	1,477,756
친환경차 등록비중(%)	1.1	1.5	2.0	2.5	3.4	3.9	4.7	5.8

친환경 차량 누적 판매량 증가 대수

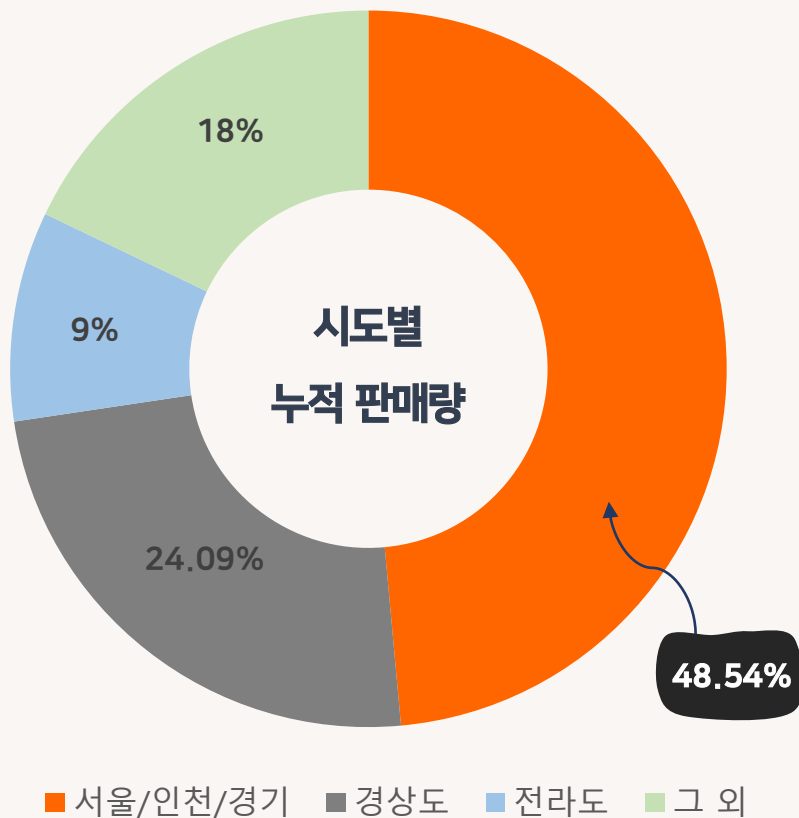
- 2013년 244,158대 / 2022년 1,477,756대

친환경 차량의 누적 판매량 증가율

- 16년도 비교 현재 83.5% 판매량 증가를 확인

16년도 말 1.1%의 비중밖에 안되던 친환경 차량의 등록비중은
22년도 9월 기준 5.8%로 성장하고 있음을 확인 할 수 있다.

친환경 차량 누적 판매량



구분	2022.9			
	전기	수소	하이브리드	친환경차
서울	53,798	2,790	162,401	218,989
부산	19,858	1,807	75,621	97,286
대구	22,329	476	56,928	79,733
인천	22,375	1,467	111,549	135,391
광주	7,887	1,032	28,235	37,154
대전	12,525	1,198	27,589	41,312
울산	4,651	2,391	22,496	29,538
세종	2,677	245	12,263	15,185
경기	67,941	5,671	289,391	363,003
강원	12,171	2,374	27,381	41,926
충북	13,704	1,515	29,834	45,053
충남	15,449	1,149	37,615	54,213
전북	12,162	1,507	29,522	43,191
전남	13,802	672	45,023	59,497
경북	17,057	175	42,129	59,361
경남	19,453	2,244	68,401	90,098
제주	29,556	6	37,264	66,826
합계	347,395	26,719	1,103,642	1,477,756
전체 대비 비율	(1.4%)	(0.1%)	(4.4%)	(5.8%)

서울 / 인천 / 경기 (수도권)
친환경 차량 누적 판매량 : **717,383 대**
수도권 친환경 차량 비율 : **48.54%**

Why 서울특별시?

친환경 차량의 대부분이 수도권에 분포
서울 생활권역으로 묶어서 진행

수익성 감소 및 출혈 경쟁

기업·CEO ▾

치솟는 기름값에 출혈경쟁...문 닫는 주유소들

오현길 기자

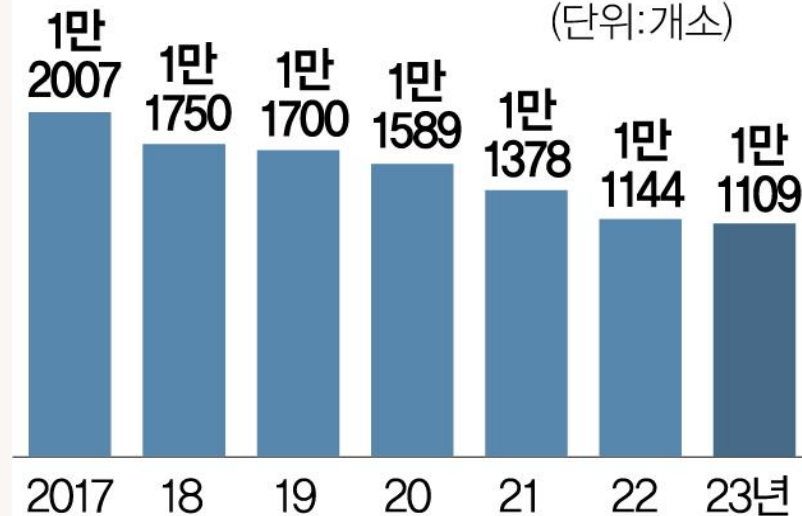
입력 2022.07.06 10:54 수정 2022.07.06 11:12

🕒 읽는 시간 1분 7초

한국석유관리원에 따르면 올 3월 말 기준 전국에 영업 중인 주유소는 1만 1109개소로 지난해 말(1만 1144개소) 대비 35개소 감소했다. 2017년 말만 해도 약 1만 2000개소였던 주유소는 올해 말 1만 1000개소를 밑돌 것으로 추정된다. 석유관리원은 “이런 주유소의 감소는 인구밀도가 높은 서울과 광역시 등에서 더 뚜렷하다”며 “업체 간 치열한 경쟁과 전기차 등 친환경 차량 보급 확대 등 경영상 어려움에 기인한 것으로 보인다”고 설명했다.

연도별 전국 주유소 현황

(단위:개소)

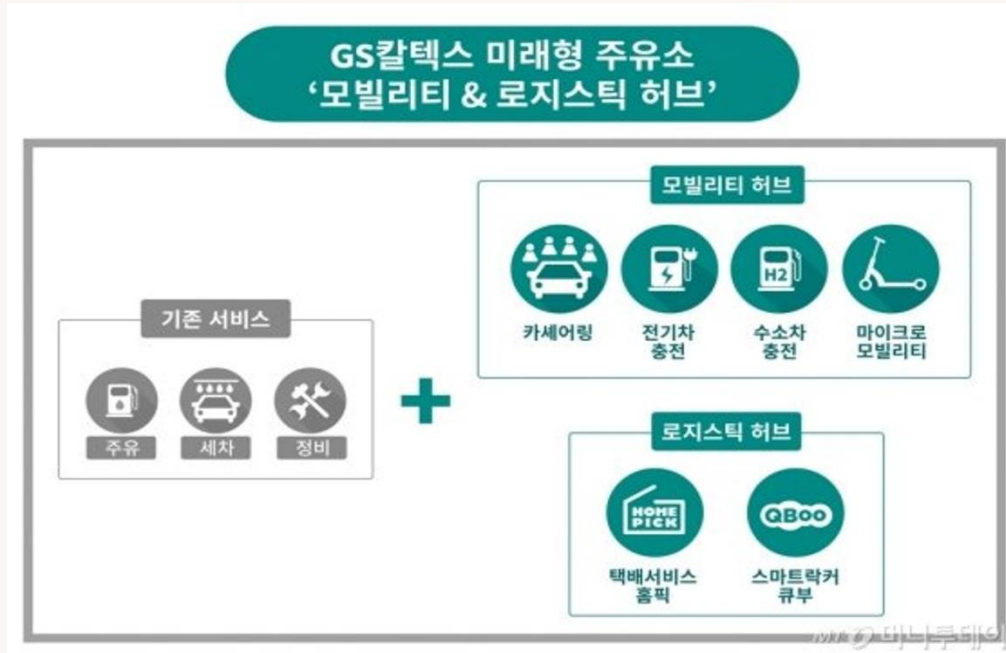


※2023년 3월 기준

자료: 석유관리원

기름값이 오를 수록 폐업 사례는 비례함 / 저렴한 주유소에 물리는 경향이 있어 불가피하게 출혈 경쟁 중
철거 비용과 환경개선부담금 등 폐업 부담이 높아 유령 주유소 또한 증가 중

물류 / 유통 HUB



주유소를 택배 집하장으로 활용하는 공유경제 모델 적용
연료 판매나 세차 뿐만 아니라 **지역 물류 유통 거점**이 되면서 **추가 수익 + 일자리 창출**

물류 / 유통 HUB



디지털, AI 등 강점을 가진 네이버와의 협력을 통해 진보한 주유소 거점 물류 서비스를 제공을 위해 노력 중
네이버는 이커머스 서비스의 물류기지로 활용하고 도심형 풀필먼트 물류센터로 구축 중

첨단 물류 복합 주유소 조성



업계 1·2위 사업자가 협력해 사업 추진 중

- 정유·유류 판매업과 무관해 **공정거래법상의 규제 X**

친환경 모빌리티 거점 구축

- **전기차** 충전시설 및 **공유 이동수단** 집약

서울시 도시 물류 환경 개선

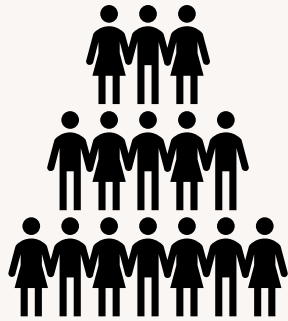
- 드론, 로봇 등 **미래 물류 기술**을 실증해 물류 경쟁력 상승

탄소 중립 정책 / 친환경 차량의 증가로 인해 수익성 감소해 첨단 물류 복합 주유소의 **필요성 증가**

2023년 상반기부터 첨단 물류 복합 주유소 선보일 예정

첨단 물류 복합 주유소 **입지 분석이 필요한 시점**

데이터 수집



인구



물류창고



충전소



택배함

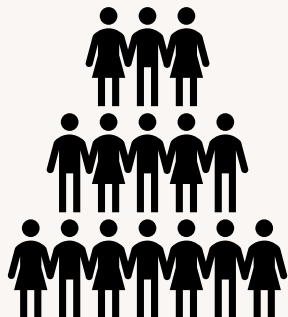


자동차



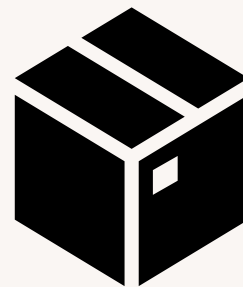
공공자전거

본론 데이터 수집 및 전처리



법정동

	법정동코드	기준연월	시도명	시군구명	읍면동명	리명	계
0	1111010100	2023-04-30	서울특별시	종로구	청운동	청운동	2339
1	1111010200	2023-04-30	서울특별시	종로구	신교동	신교동	1361
2	1111010300	2023-04-30	서울특별시	종로구	궁정동	궁정동	192
3	1111010400	2023-04-30	서울특별시	종로구	효자동	효자동	572
4	1111010500	2023-04-30	서울특별시	종로구	창성동	창성동	414
...
450	1174010600	2023-04-30	서울특별시	강동구	둔촌동	둔촌동	24160
451	1174010700	2023-04-30	서울특별시	강동구	암사동	암사동	66418
452	1174010800	2023-04-30	서울특별시	강동구	성내동	성내동	65051
453	1174010900	2023-04-30	서울특별시	강동구	천호동	천호동	85764
454	1174011000	2023-04-30	서울특별시	강동구	강일동	강일동	32855



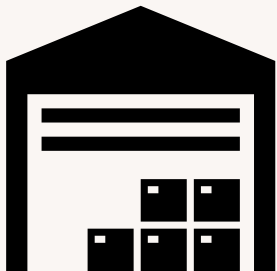
안심택배함.tail(5)

	동	택배함명	주소	비고	Unnamed: 4	Unnamed: 5
260	260.0	NaN	명일1동	강동구 평생학습관	구천면로 395	4.6.재설치
261	261.0	NaN	암사1동	강동구립암사도서관	고덕로20길 42	NaN
262	262.0	NaN	천호2동	강동구립해공도서관	올림픽로 702해공공원	NaN
263	263.0	NaN	길동	길동주민센터	길동 228-10	NaN
264	264.0	NaN	천호2동	천호청소년문화의집	천중로 61	4.6. 신규설치

행정안전부 지역별(법정동) 성별 연령별 주민등록 인구수

서울시 안심 택배함 현황

본론 데이터 수집 및 전처리



물류창고.head(3)

번호	개방서비스명	개방서비스아이디	개방자치단체코드	관리번호	인허가일자	인허가취소일자	영업상태구분코드	영업상태명	상세영업상태코드	...	냉동냉장고면적	보관장소면적	직원수	시설/장비현황	보관요율	범인여부명	업태보관및창고업	업태운송및택배업	업태판매업	업태제조업	
0	1	물류창고업체	11_45_02_P	4070000	2.02E+18	2018-12-03	NaN	1	영업/정상	1	...	0	0.0	12	랙설치, 지게차	1000	범인	1.0	NaN	NaN	NaN
1	2	물류창고업체	11_45_02_P	4050000	2.02E+18	2018-10-31	NaN	1	영업/정상	1	...	0	0.0	31	랙(rack)설치, 지게차	7400	범인	1.0	NaN	NaN	NaN
2	3	물류창고업체	11_45_02_P	4980000	2.02E+18	2018-11-02	NaN	1	영업/정상	1	...	0	0.0	3	창고 3층	NaN	개인	1.0	NaN	NaN	NaN

행정안전부 물류창고 현황

자동차_등록대수.tail(5)

	사용본거지시군구	읍면동(행정동)	연로계	관용승용	관용승합	관용화물	관용특수	자가용승용	자가용승합	자가용화물	자가용특수	영업용승용	영업용승합	영업용화물	영업용특수
3565	서울특별시강동구	강동구 둔촌2동	하이브리드	377	0	0	0	0	376	0	0	0	1	0	0
3566	서울특별시강동구	강동구 둔촌2동	수소	8	0	0	0	0	8	0	0	0	0	0	0
3567	서울특별시강동구	강동구 둔촌2동	기타연료	12	0	0	0	3	0	2	6	0	0	0	1
3568	서울특별시강동구	강동구 기타1	휘발유	28	0	0	0	0	25	1	2	0	0	0	0
3569	서울특별시강동구	강동구 기타1	경유	18	0	0	0	0	0	4	14	0	0	0	0

서울시 행정동 별 자동차 등록 현황

보론



주소.head(2)																				
순번		충전소_관리번호	충전소명	전화번호	도로명주소	지번주소	충전소유명	충전소유명비고	충전가능차량코드	...	토요일_시작	토요일_종료	일요일_시작	일요일_종료	공휴일_시작	공휴일_종료	휴식_시간	휴식_종료	경도	위도
0	1	2826020121HS2022033	하이브리드형충전소	070-7727-1719	인천서구장서동410	인천서구장서동1708	인천서구장서동1708	NaN	5	...	08:00	20:00	08:00	20:00	08:00	20:00	12:00	13:00	126.646836	37.568838
1	2	1150020121HS2022001	마곡에스피노사충전소	02-6081-8883	서울강남구양재동720	서울강남구양재동61-1	서울강남구양재동720	NaN	5	...	09:00	18:00	NaN	NaN	NaN	NaN	NaN	NaN	126.829152	37.573081

전처리

1. 각 시설의 주소를 위경도로 변환 후 카카오 API를 통해 법정동명으로 변환

```
from geopy.geocoders import Nominatim
geo_local = Nominatim(user_agent='South Korea')

def geocoding(address):
    try:
        geo = geo_local.geocode(address)
        x_y = [geo.latitude, geo.longitude]
        return x_y
    except:
        return [0,0]
```

```
latitude = []
longitude = []

for i in address:
    latitude.append(geocoding(i)[0])
    longitude.append(geocoding(i)[1])

친환경차['위도'] = latitude
친환경차['경도'] = longitude
```

```
API_KEY = KEY
URL = 'https://dapi.kakao.com/v2/local/geo/coord2regioncode.json'

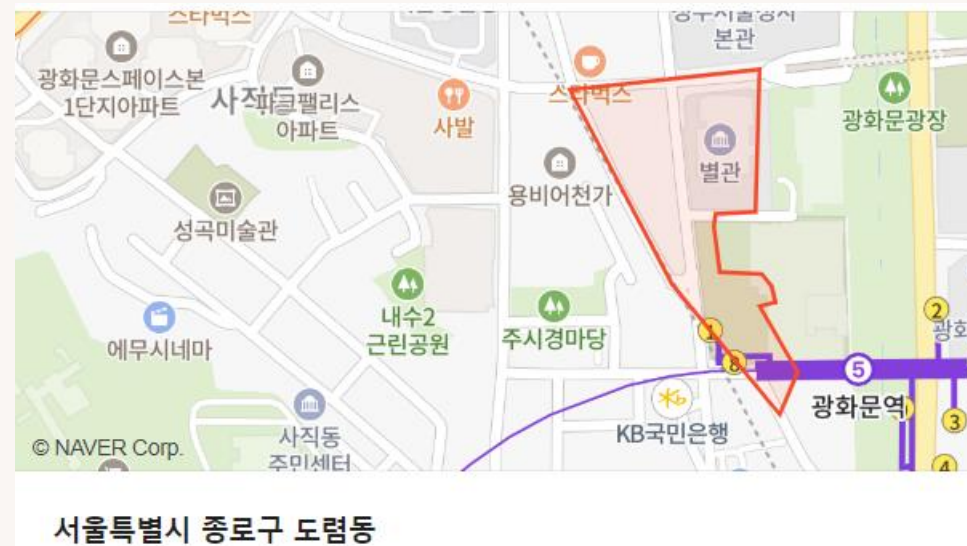
for i in range(len(친환경차)):
    x_crd = float(친환경차.loc[i, ['경도']])
    y_crd = float(친환경차.loc[i, ['위도']])
    address = get_address(x_crd, y_crd)

    친환경차.loc[i, ['법정동']] = address
```

2. 법정동 기준으로 데이터셋 만든 후, 인구수 하위 25%이며 택배함 & 물류창고 & 충전기 & 전기차 수 = 0인 데이터 제거

인구수	
count	455.000000
mean	20709.287912
std	34250.428984
min	1.000000
25%	844.000000
50%	4615.000000
75%	30061.500000
max	245562.000000

16	1.111E+9	#####	서울특별시 종로구	사직동	사직동	3193
17	1.111E+9	#####	서울특별시 종로구	도림동	도림동	1
18	1.111E+9	#####	서울특별시 종로구	당주동	당주동	301
19	1.111E+9	#####	서울특별시 종로구	내수동	내수동	2655
20	1.111E+9	#####	서울특별시 종로구	신문로1가	신문로1가	77
21	1.111E+9	#####	서울특별시 종로구	신문로2가	신문로2가	728
22	1.111E+9	#####	서울특별시 종로구	청진동	청진동	77
23	1.111E+9	#####	서울특별시 종로구	서린동	서린동	2
24	1.111E+9	#####	서울특별시 종로구	수송동	수송동	246
25	1.111E+9	#####	서울특별시 종로구	중학동	중학동	4
26	1.111E+9	#####	서울특별시 종로구	종로1가	종로1가	93
27	1.111E+9	#####	서울특별시 종로구	공평동	공평동	1



전처리

3. 파생변수 생성

	법정동	인구수	택배함 개수	물류창고 개수	공공자전거 거치대수(LCD + QR)	친환경차 등록대수(전기+수소)	충전소 개수(전기+수소)	친환경차 한대당 충전소 개수
0	종로구 청운동	2339	0	0	8	0	0	0.000000
1	종로구 신교동	1361	0	0	0	0	0	0.000000
2	종로구 궁정동	192	1	0	0	69	0	0.000000
3	종로구 누상동	3206	0	0	0	0	0	0.000000
4	종로구 옥인동	1880	0	0	0	0	0	0.000000
5	종로구 필운동	994	0	0	0	0	0	0.000000
6	종로구 사직동	3193	0	0	14	0	4	0.000000
7	종로구 내수동	2655	0	0	12	163	2	0.012270

“하이브리드차 이르면 2025년부터 친환경차 제외된다.”

홍남기 부총리, 혁신성장 빅3 추진 회의
“HEV, 2025·2026년부터 저공해차 제외”
하이브리드를 내연기관차로 보겠단 의미
전기·수소차 중심 보급 확대하려는 목적
충전 인프라 부족하면 ‘충전대란’ 불가피

[출처] 서울신문, 22년 2월 24일 뉴스 기사

* 친환경차 등록대수 (전기 + 수소)

* 친환경차 한대당 충전소 개수

* 충전소 개수 (전기 + 수소)

• 법정동 별로 충전소 구축이 얼마나 얼마나 잘되어 있는지 비율로 확인하기 위함

2025년부터 하이브리드가 친환경차에서 제외됨을 고려

전처리

4. 최종 데이터 셋

```
## 'data.frame':   363 obs. of  8 variables:
## $ 법정동           : chr  "종로구 청운동" "종로구 신교동" "종로구 궁정동" "종로구
## $ 인구수           : int   2339 1361 192 3206 1880 994 3193 2655 246 106 ...
## $ 택배함.개수      : int    0 0 1 0 0 0 0 0 0 0 ...
## $ 물류창고.개수    : int    0 0 0 0 0 0 0 0 0 0 ...
## $ 공공자전거.거치대수.LCD...QR.: int    8 0 0 0 0 0 14 12 0 26 ...
## $ 친환경차.등록대수.전기.수소. : int    0 0 69 0 0 0 0 163 0 140 ...
## $ 충전소.개수.전기.수소.        : int    0 0 0 0 0 0 4 2 1 0 ...
## $ 친환경차.한대당.충전소.개수   : num    0 0 0 0 0 ...
```

클러스터링

- 법정동 중 어느 곳을 우선적으로 '첨단 물류 복합 주유소'로 전환 시킬 것인지 판단하기 위해 클러스터링 진행

1. 첨단 물류 복합 주유소 특징을 고려

- 물류창고, 택배함, 전기차 및 수소차 충전소, 공공자전거 대여소 등 다양한 서비스 공간 존재
- 주된 사용자는 지역별(법정동) 주민등록 인구라고 판단

2. 상관관계 확인 `corrplot(corr_df, method = 'number', order = 'hclust', type = 'lower', diag = FALSE)`



물류창고는 대부분 없어서 클러스터링 변수로 사용하기엔 유의미하지 않음

인구수가 택배함, 공공자전거 거치대, 친환경차 충전소와 양의 상관관계를 보임

- “인구수가 많은 지역에서는 해당 시설들이 더 많이 필요함”으로 해석

결론

인구수와 해당 변수들을 클러스터링 변수로 사용
(택배함, 공공자전거 거치대, 친환경차 충전소)

인구수 대비 시설들이 부족한 지역들을 첨단 물류 복합 주유소 필요 지역으로 선정

해당 지역들을 Target Cluster로 정하고 Voting을 통해 최종 지역(법정동) 선정

클러스터링

- 법정동 중 어느 곳을 우선적으로 '첨단 물류 복합 주유소'로 전환 시킬 것인지 판단하기 위해 클러스터링 진행

3. 클러스터링 변수 선정

```
## 'data.frame':   363 obs. of  4 variables:
## $ 인구수                : int  2339 1361 192 3206 1880 994 3193 2655 246
## $ 택배함.개수           : int   0 0 1 0 0 0 0 0 0 0 ...
## $ 공공자전거.거치대수.LCD...QR.: int   8 0 0 0 0 0 14 12 0 26 ...
## $ 충전소.개수.전기.수소.   : int   0 0 0 0 0 0 4 2 1 0 ...
```

4. Scaling

```
min_max_scaling <- function(x) {
  (x - min(x)) / (max(x) - min(x))
}
```

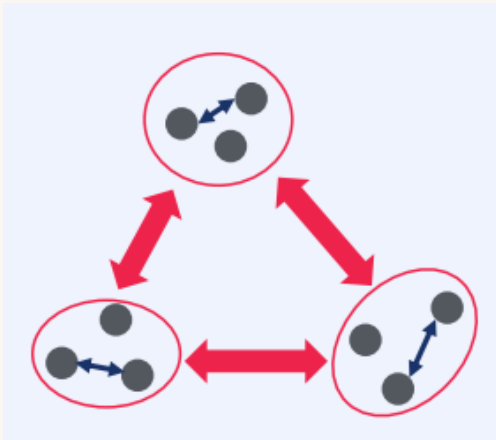
```
scaled_final <- as.data.frame(lapply(df5, min_max_scaling))
row.names(scaled_final) <- row.names(df5)
str(scaled_final)
```

```
## 'data.frame':   363 obs. of  4 variables:
## $ 인구수                : num  0.009521 0.005538 0.000778
## $ 택배함.개수           : num   0 0 0.1 0 0 0 0 0 0 0 ...
## $ 공공자전거.거치대수.LCD...QR.: num  0.00633 0 0 0 0 ...
## $ 충전소.개수.전기.수소.   : num   0 0 0 0 0 ...
```

클러스터링 평가

1. Voting을 하기 위해 K-means, K-medoids, Hierarchical, Gaussian Mixture Clustering을 진행
2. 군집의 수를 정하기 위해 기본적으로 많이 사용하는 Elbow Method, Silhouette Method, Gap Statistic Method 사용
3. 그럼에도 군집의 수를 결정하기 어려울 때 내부 평가 방법 중 하나인 Dunn Index를 확인하여 군집의 수 결정

Dunn Index

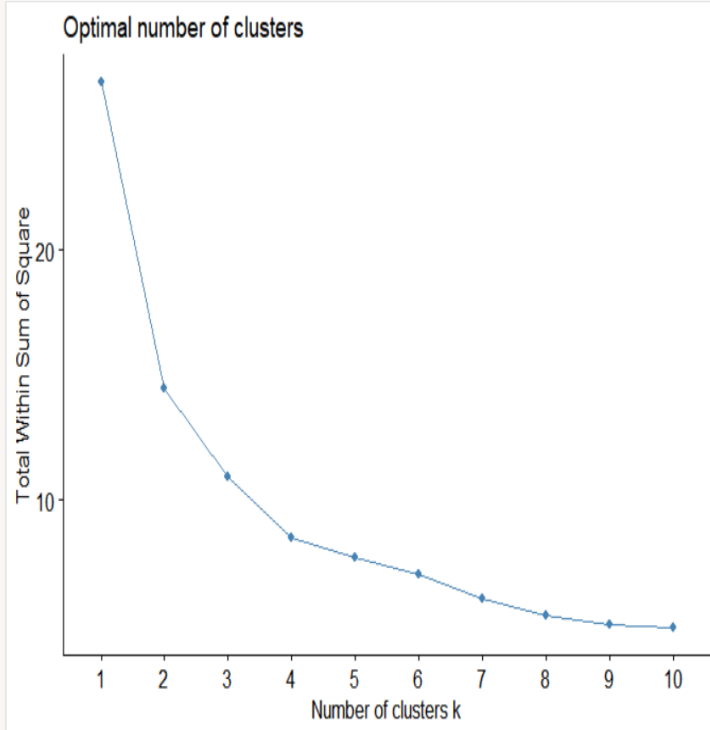


$$D = \frac{\min_{1 \leq i < j \leq n} d(i, j)}{\max_{1 \leq k \leq n} d'(k)},$$

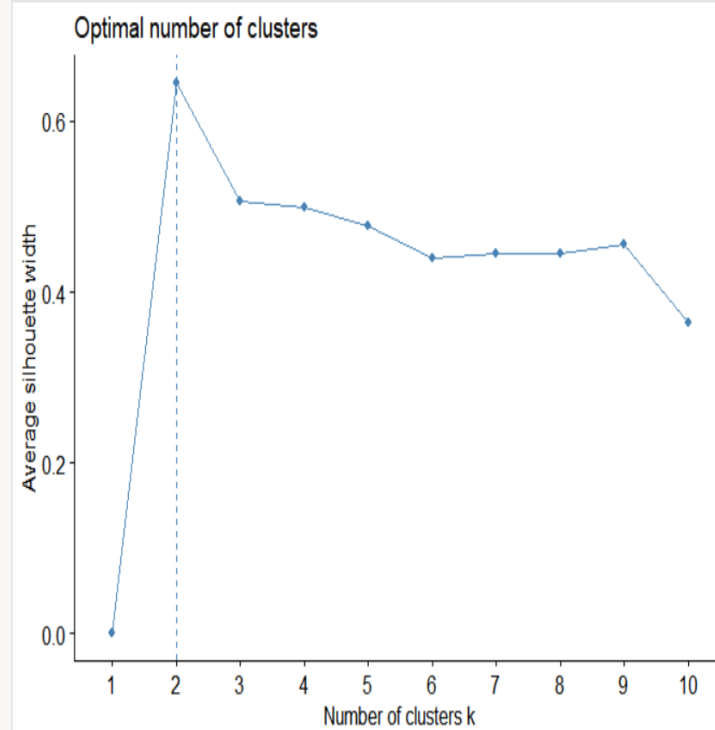
- $d(i, j)$: 클러스터 i, j 간 거리
- $d'(k)$: 클러스터 k 의 클러스터 내 거리

군집과 군집 사이의 거리가 클수록
군집 내 데이터 간 거리가 작을수록 좋은 모델
즉 **DI(Dunn Index)**가 큰 모델이 좋은 모델이다.

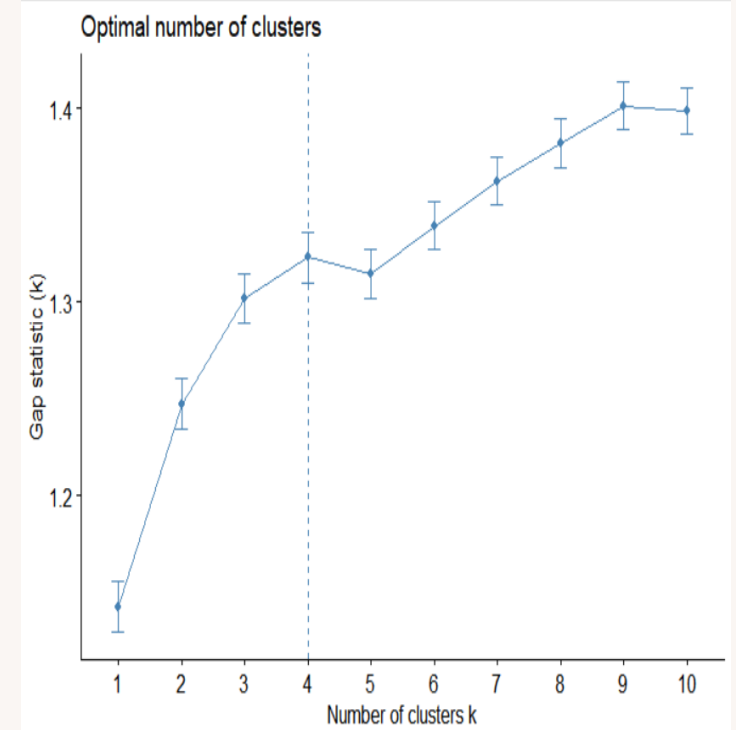
K-means 진행



Elbow Method



Silhouette Method



Gap Statistic Method

Elbow 결과: Total Within Sum of Square의 감소가 완화된, 즉 꺾이는 지점을 보았을 때 $k = 2$ 또는 $k = 4$ 로 생각해 볼 수 있음

Silhouette 결과: $k = 2$ 일 때 Average Silhouette width가 가장 넓어서 $k = 2$ 일 때 가장 안정적이라 할 수 있음

Gap Statistic 결과: $k = 4$ 이후로 기울기가 계속 증가하므로, Gap이 증가하기 직전인 $k = 4$ 를 최적이라 할 수 있음

본론 K-means

Between_SS ↔ Total_SS 값을 비교.
해당 값이 높을수록 잘 분류된 군집

클러스터 간의 거리를 나타내므로
값이 클수록 클러스터간 구별이 잘 된다는 것

```
km2 <- kmeans(scaled_final, centers=2)
km2
## Within cluster sum of squares by cluster:
## [1] 7.920463 6.548936
## (between_SS / total_SS = 45.8 %)
```

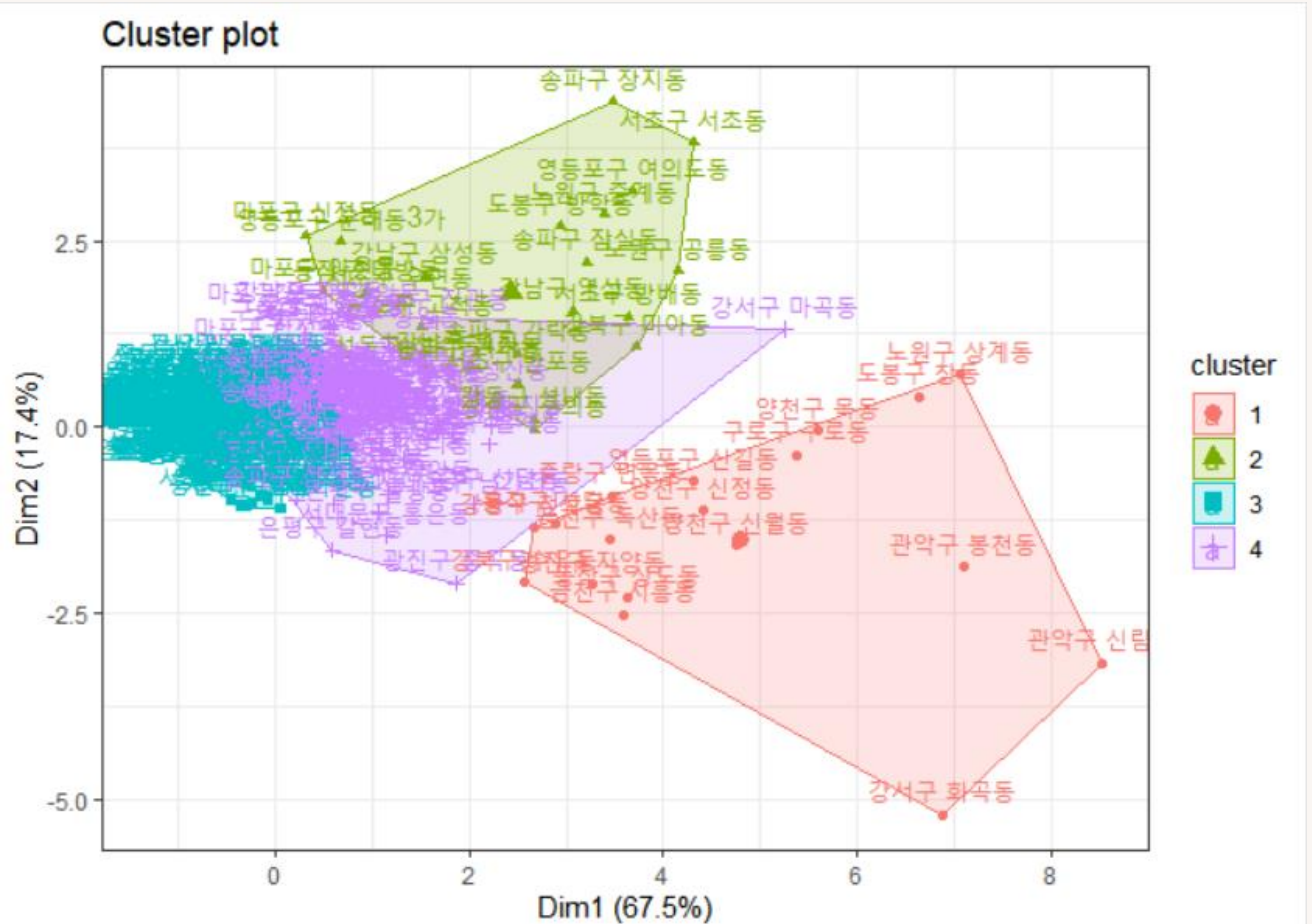
```
km <- kmeans(scaled_final, centers = 4)
km

## Within cluster sum of squares by cluster:
## [1] 2.341060 1.675096 1.464450 3.029937
## (between SS / total_SS = 68.1 %)
```

▲ K = 4로 클러스터링 진행

K = 4일 때 K-means 결과

```
fviz_cluster(km, data = scaled_final)+theme_bw()
```



본론 K-means

K = 4일 때 K-means 결과 프로파일링

```
## # A tibble: 4 x 8
##   clst_k4  인구수  택배함.개수  물류창고.개수  공공자전거...1  친환경...2  충전소...3  친환경...4
##   <int>    <dbl>    <dbl>        <dbl>        <dbl> <dbl>    <dbl>    <dbl>
## 1      1 138542.    5.22        0.0556        332.  498.    11    0.0332
## 2      2  61316.    1.17        0.5          227.  408.   21.6  0.0823
## 3      3   6716.    0.242       0.0932        28.1  29.2    1.04  0.00841
## 4      4  45298.    1          0.388        161.  258.    6.21  0.0568
## # ... with abbreviated variable names 1공공자전거.거치대수.LCD...QR.,
## # 2친환경차.등록대수.전기.수소., 3충전소.개수.전기.수소.,
## # 4친환경차.한대당.충전소.개수
```

↓

```
mean_4$한명당.택배함 <- mean_4$택배함.개수/mean_4$인구수
mean_4$한명당.물류창고 <- mean_4$물류창고.개수/mean_4$인구수
mean_4$한명당.친환경충전소 <- mean_4$충전소.개수.전기.수소./mean_4$인구수
mean_4$한명당.자전거거치대 <- mean_4$공공자전거.거치대수.LCD...QR./mean_4$인구수
```

↓

```
## # A tibble: 4 x 7
##   clst_k4  인구수  한명당.택배함  한명당.물류창고  한명당.친환경...1  한명당...2  친환경...3
##   <int>    <dbl>    <dbl>        <dbl>        <dbl>    <dbl>    <dbl>
## 1      1 138542.    0.0000377    0.000000401    0.0000794  0.00240  0.0332
## 2      2  61316.    0.0000190    0.00000815    0.000353  0.00371  0.0823
## 3      3   6716.    0.0000360    0.0000139    0.000155  0.00419  0.00841
## 4      4  45298.    0.0000221    0.00000857    0.000137  0.00356  0.0568
## # ... with abbreviated variable names 1한명당.친환경충전소, 2한명당.자전거거치대,
## # 3친환경차.한대당.충전소.개수
```

군집 1이 인구수 대비 시설들(택배함, 충전소, 공공자전거 거치대)이 부족

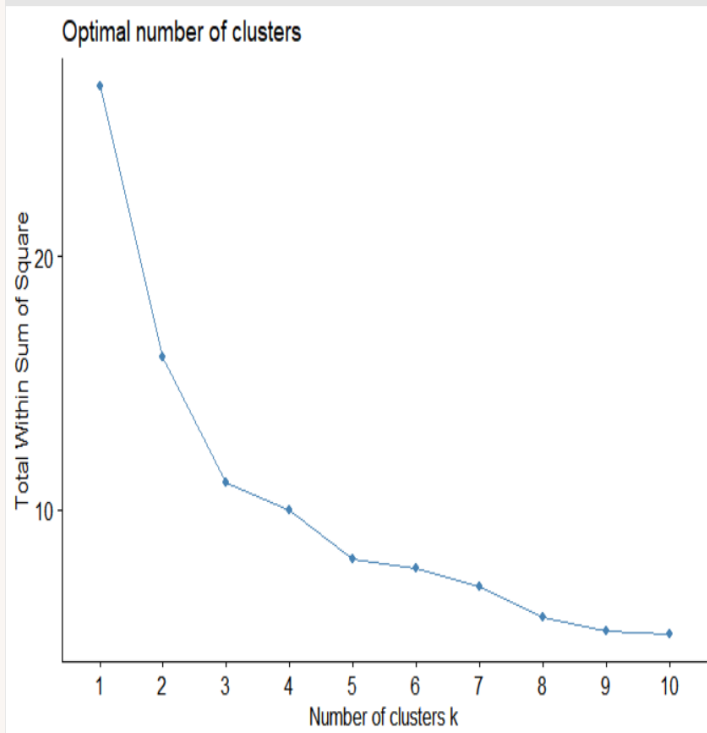
Cluster 1이 Target Cluster

```
k_means_result <- k_means[km$cluster == 1, "법정동"]
k_means_result
```

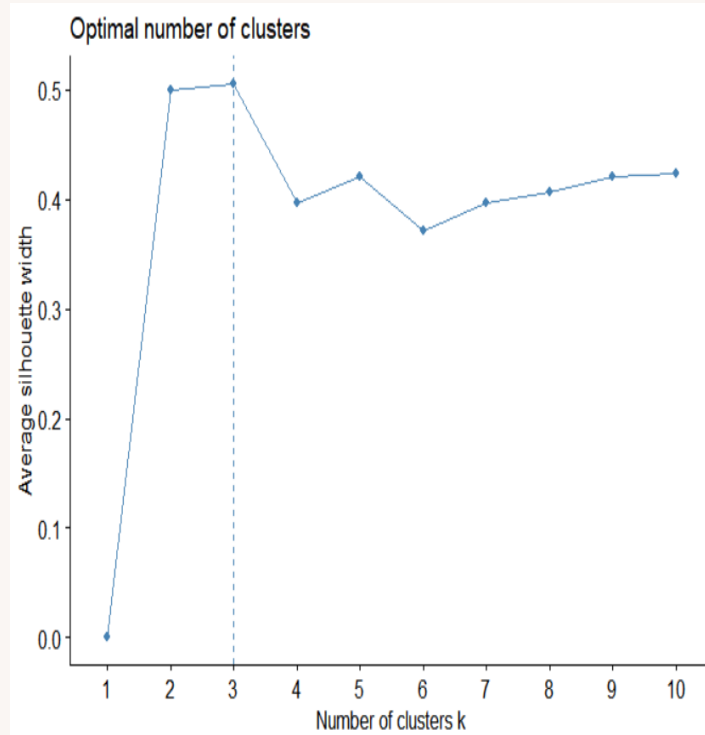
```
## [1] "광진구 자양동" "중랑구 면목동" "강북구 수유동" "도봉구 창동"
## [5] "노원구 상계동" "양천구 신정동" "양천구 목동" "양천구 신월동"
## [9] "강서구 화곡동" "구로구 구로동" "금천구 독산동" "금천구 시흥동"
## [13] "영등포구 신길동" "동작구 상도동" "동작구 사당동" "관악구 봉천동"
## [17] "관악구 신림동" "강동구 천호동"
```

K-means 결과 다음과 같은 법정동이 Target임을 확인할 수 있음

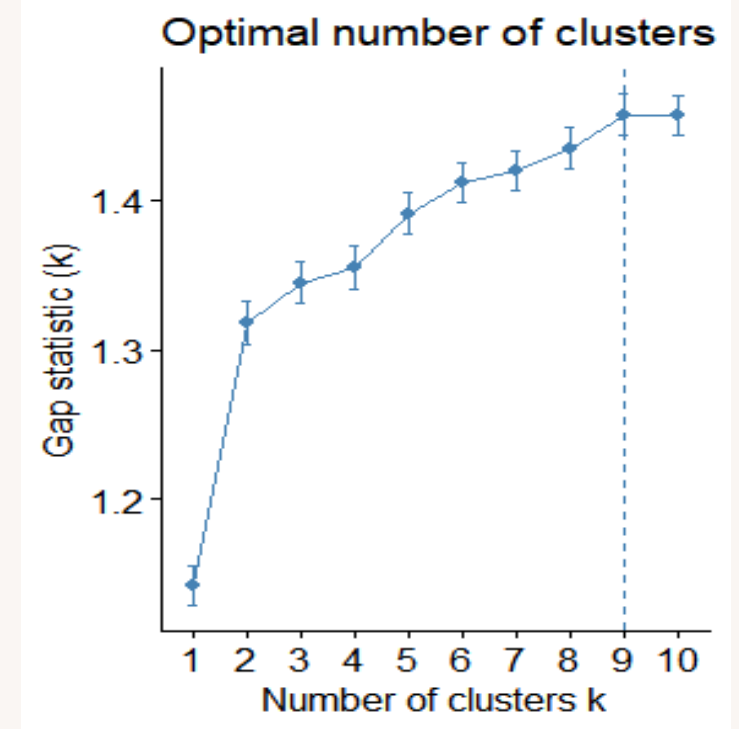
K-medoids 진행



Elbow Method



Silhouette Method



Gap Statistic Method

Elbow 결과: Total Within Sum of Square의 감소가 완화된, 즉 꺾이는 지점을 보았을 때 $k = 3$ 또는 $k = 5$ 로 생각해 볼 수 있음

Silhouette 결과: $k = 3$ 일 때 Average Silhouette width가 가장 넓어서 $k = 3$ 일 때 가장 안정적이라 할 수 있음

Gap Statistic 결과: $k = 10$ 이후로 기울기가 계속 증가하므로, Gap이 증가하기 직전인 $k = 9$ 를 최적이라 할 수 있음

군집의 수를 결정하기 어려움. Dunn Index를 확인하여 군집의 수 결정

본론 K-medoids

```
k_med_clvalid <- clValid(scaled_final,3:9, clMethods="pam",validation="internal", maxitems = nrow(scaled_final))
summary(k_med_clvalid)
```

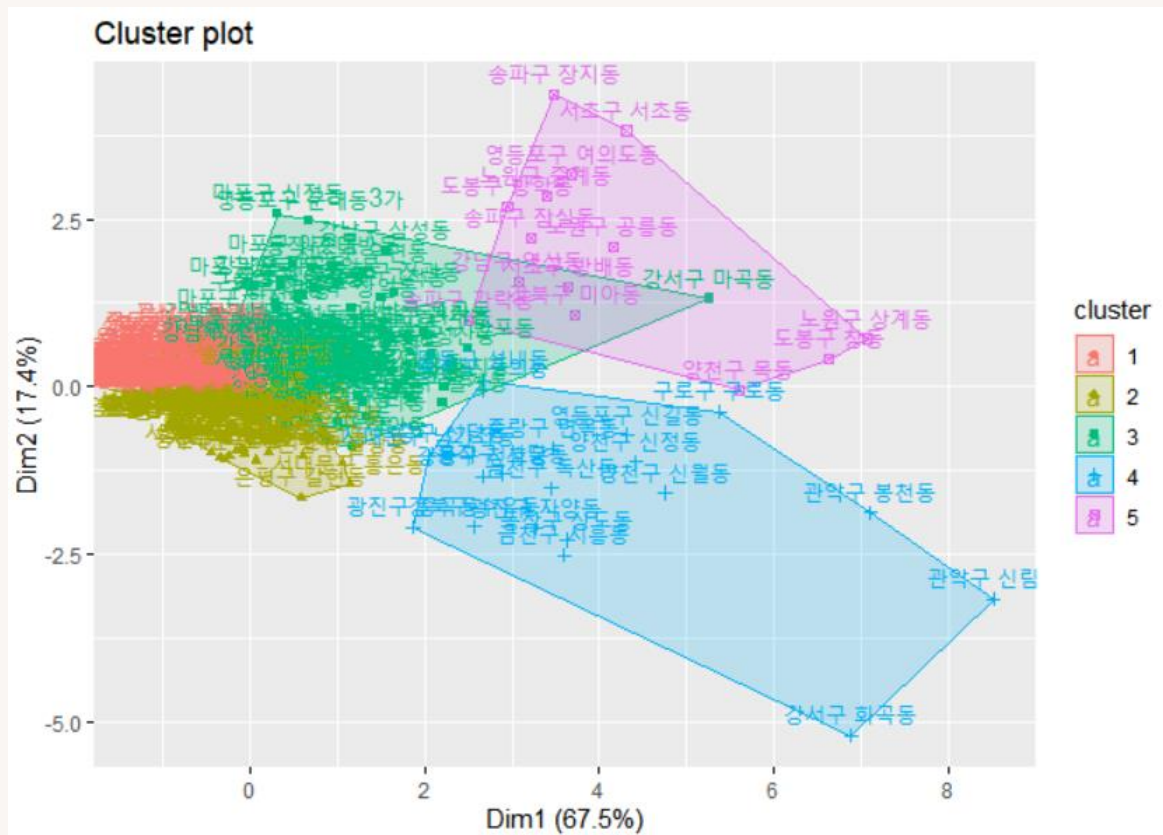
```
##
## Clustering Methods:
##  pam
##
## Cluster sizes:
##  3 4 5 6 7 8 9
##
## Validation Measures:
##           3           4           5           6           7           8           9
##
## pam Connectivity  55.3417  59.1071  66.7036  92.1206  93.2008  98.7587 102.7929
##      Dunn         0.0219   0.0219   0.0240   0.0168   0.0191   0.0202   0.0203
##      Silhouette   0.5046   0.3965   0.4201   0.3715   0.3969   0.4068   0.4211
##
## Optimal Scores:
##
##           Score  Method Clusters
## Connectivity 55.3417 pam      3
## Dunn         0.0240 pam      5
## Silhouette   0.5046 pam      3
```

앞의 3가지 방법의 결과를 고려해 k = 3부터 k = 9일 때까지 Dunn값을 확인한 결과, **Dunn값이 가장 높은 k = 5**로 결정
클러스터 간의 분리도가 높고, 클러스터 내의 응집도가 높을 수록 Dunn 지수는 더 큰 값을 가짐

본론 K-medoids

K = 5일 때 K-medoids 결과

```
fviz_cluster(k_med, data=scaled_final)
```



K = 5일 때 K-medoids 결과 프로파일링 Cluster 4가 Target Cluster

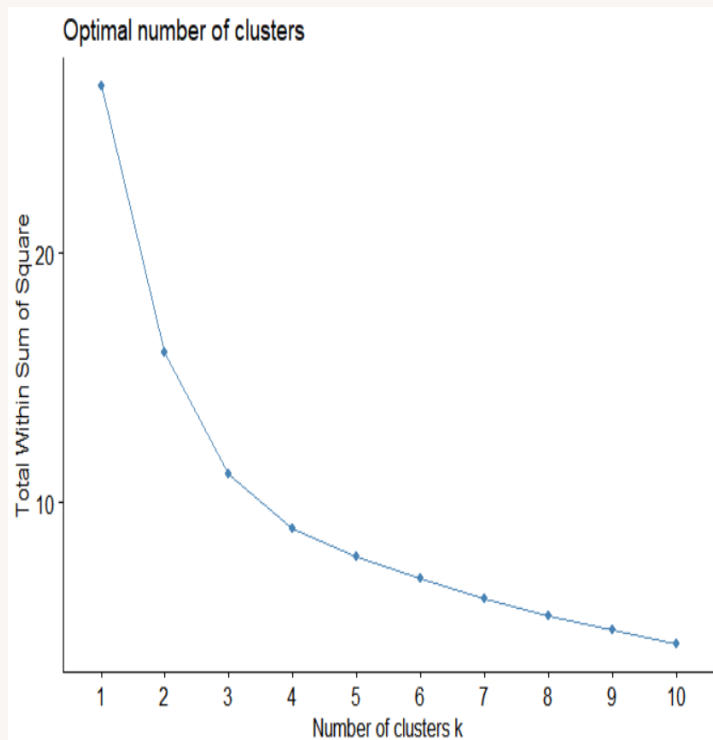
```
## # A tibble: 5 x 7
##   clst_k5  인구수 한명당.택배함 한명당.물류창고 한명당.친환경...1 한명당...2 친환경...3
##   <int>    <dbl>      <dbl>          <dbl>          <dbl>      <dbl>    <dbl>
## 1      1    4661.        0          0.0000196      0.000215  0.00487  0.00386
## 2      2   19916.     0.0000614    0.0000151      0.0000526  0.00343  0.0231
## 3      3   44041.     0.0000179    0.00000568     0.000200  0.00367  0.0682
## 4      4  118530.     0.0000409    0.00000422     0.0000709  0.00229  0.0319
## 5      5   98925.     0.0000195    0.00000794     0.000266  0.00355  0.0803
## # ... with abbreviated variable names 1한명당.친환경충전소, 2한명당.자전거거치대,
## # 3친환경차.한대당.충전소.개수
```

K-medoids 결과 다음과 같은 법정동이 Target임을 확인할 수 있음

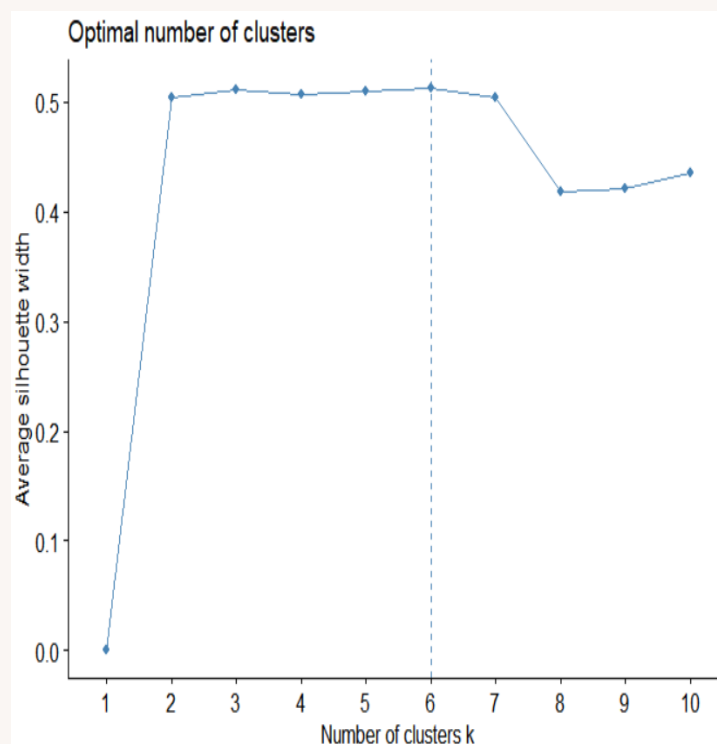
```
k_med_result <- k_medoids[k_med$cluster == 4, "법정동"]
k_med_result
```

```
## [1] "중구 신당동"      "광진구 중곡동"    "광진구 구의동"
## [4] "광진구 자양동"    "중랑구 면목동"    "강북구 수유동"
## [7] "서대문구 남가좌동" "양천구 신정동"    "양천구 신월동"
## [10] "강서구 화곡동"    "구로구 구로동"    "금천구 독산동"
## [13] "금천구 시흥동"    "영등포구 신길동"  "동작구 상도동"
## [16] "동작구 사당동"    "관악구 봉천동"    "관악구 신림동"
## [19] "강동구 성내동"    "강동구 천호동"
```

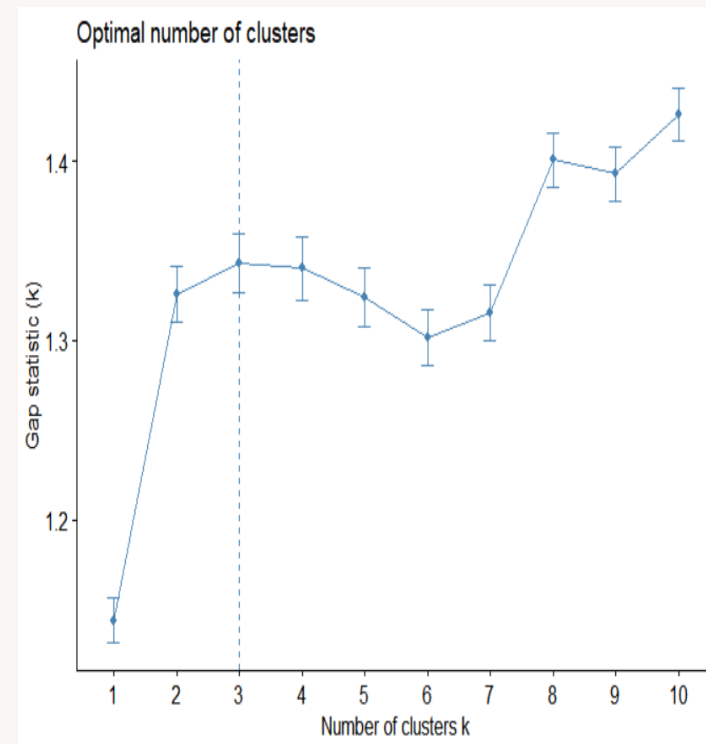
Hierarchical Clustering 진행



Elbow Method



Silhouette Method



Gap Statistic Method

Elbow 결과: Total Within Sum of Square의 감소가 완화된, 즉 꺾이는 지점을 보았을 때 $k = 3$ 또는 $k = 4$ 로 생각해 볼 수 있음

Silhouette 결과: $k = 6$ 일 때 Average Silhouette width가 가장 넓어서 $k = 6$ 일 때 가장 안정적이라 할 수 있음

Gap Statistic 결과: $k = 3$ 이 최적이지만, $k = 6$ 부터 Gap이 증가하므로 증가하기 직전인 $k = 3$ 또는 $k = 5$ 를 최적이라 할 수 있음

본론 Hierarchical Clustering

```
hcut_clvalid <- clValid(scaled_final, 3:30, clMethods="hierarchical", validation="internal", maxitems = nrow(scaled_final))
summary(hcut_clvalid)

## Validation Measures:
##
##
## hierarchical Connectivity      8.1603 18.1560 19.5349 25.9444
##                      Dunn      0.3203  0.1513  0.1513  0.1864
##                      Silhouette 0.7101  0.6511  0.6331  0.6375
##
## Optimal Scores:
##
##      Score Method      Clusters
## Connectivity 8.1603 hierarchical 3
## Dunn      0.3203 hierarchical 3
## Silhouette 0.7101 hierarchical 3
```

앞의 3가지 방법의 결과를 고려해 k = 3부터 k = 6까지 Dunn값을 확인한 결과,
Dunn값이 가장 높은 **k = 3**으로 결정

안정적인 군집을 만들기 위해 HC의 Alg을 비교

```
m <- c("average", "single", "complete", "ward", "weighted", "gaverage")
names(m) <- c("average", "single", "complete", "ward", "weighted", "gaverage")
ac <- function(x) {
  agnes(scaled_final, method = x)$ac
}
choose_best_alg <- map_dbl(m, ac)
print(choose_best_alg)
```

```
## average single complete ward weighted gaverage
## 0.9628766 0.9359720 0.9700834 0.9901746 0.9549834 0.9822914
```

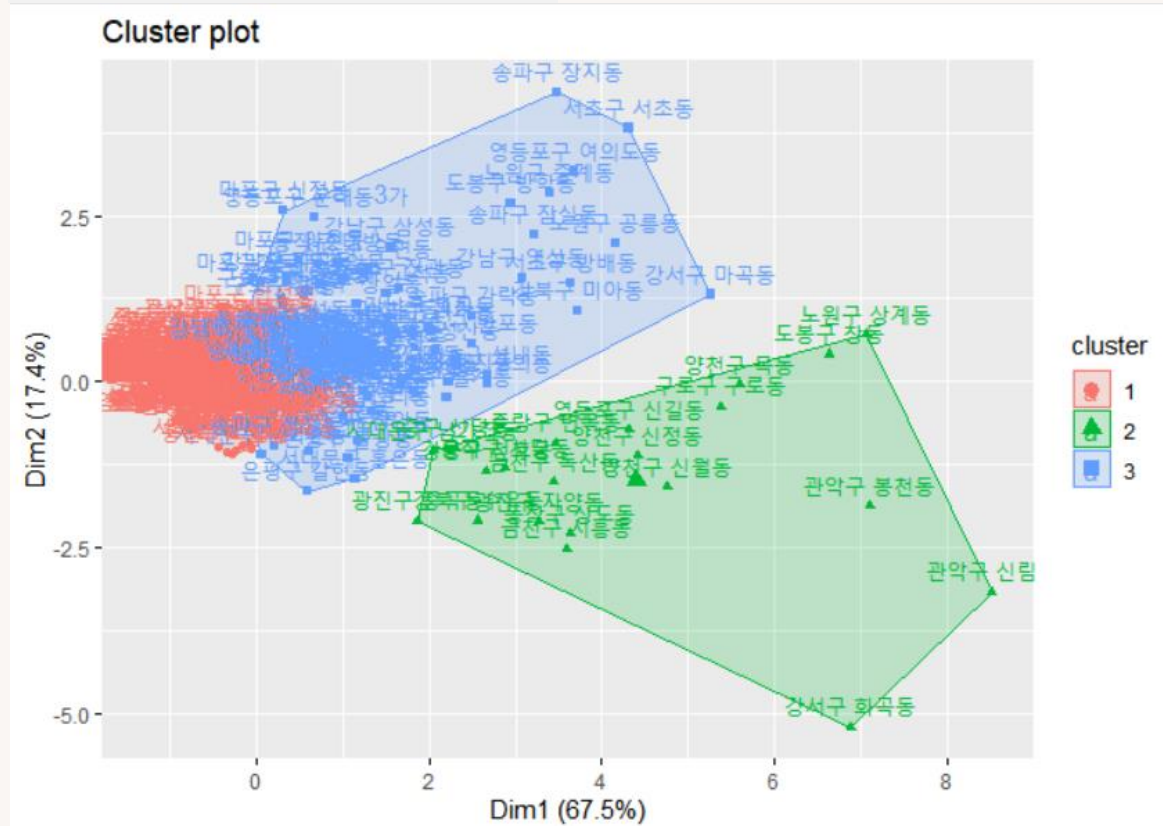
응집형 계수(ac)가 값이 **1에 가까울수록**
해당 Alg을 사용하면 군집이 가장 안정적임

Ward Alg이 ac값이 가장 높으므로, 해당 Alg을 사용하면
6가지 방법 중 **가장 강력한 클러스터링 구조를 식별**할 수 있다.

본론 Hierarchical Clustering

K = 3일 때 Ward Alg 적용한 HC 결과

`fviz_cluster(hc.cut, scaled_final)`



K = 3일 때 HC 결과 프로파일링

Cluster 2가 Target Cluster

```
## # A tibble: 3 x 7
##   clst_k3  인구수 한명당.택배함 한명당.물류창고 한명당.친환경...1 한명당...2 친환...3
##   <int>    <dbl>      <dbl>      <dbl>      <dbl>      <dbl> <dbl>
## 1      1  7073.      0.0000365    0.0000112    0.000160    0.00417 0.0126
## 2      2 127992.    0.0000394    0.000000372  0.0000807  0.00237 0.0346
## 3      3 49167.      0.0000191    0.00000957    0.000200    0.00366 0.0552
## # ... with abbreviated variable names 1한명당.친환경충전소, 2한명당.자전거거치대,
## # 3친환경차.한대당.충전소.개수
```

HC 결과 다음과 같은 법정동이 Target임을 확인할 수 있음

```
hcut_result <- hcut_cluster[hc.cut$cluster == 2, "법정동"]
hcut_result
```

```
## [1] "중구 신당동"      "광진구 중곡동"      "광진구 자양동"
## [4] "중랑구 면목동"    "강북구 수유동"      "도봉구 창동"
## [7] "노원구 상계동"    "서대문구 남가좌동"  "양천구 신정동"
## [10] "양천구 목동"      "양천구 신월동"      "강서구 화곡동"
## [13] "구로구 구로동"    "금천구 독산동"      "금천구 시흥동"
## [16] "영등포구 신길동"  "동작구 상도동"      "동작구 사당동"
## [19] "관악구 봉천동"    "관악구 신림동"      "강동구 천호동"
```

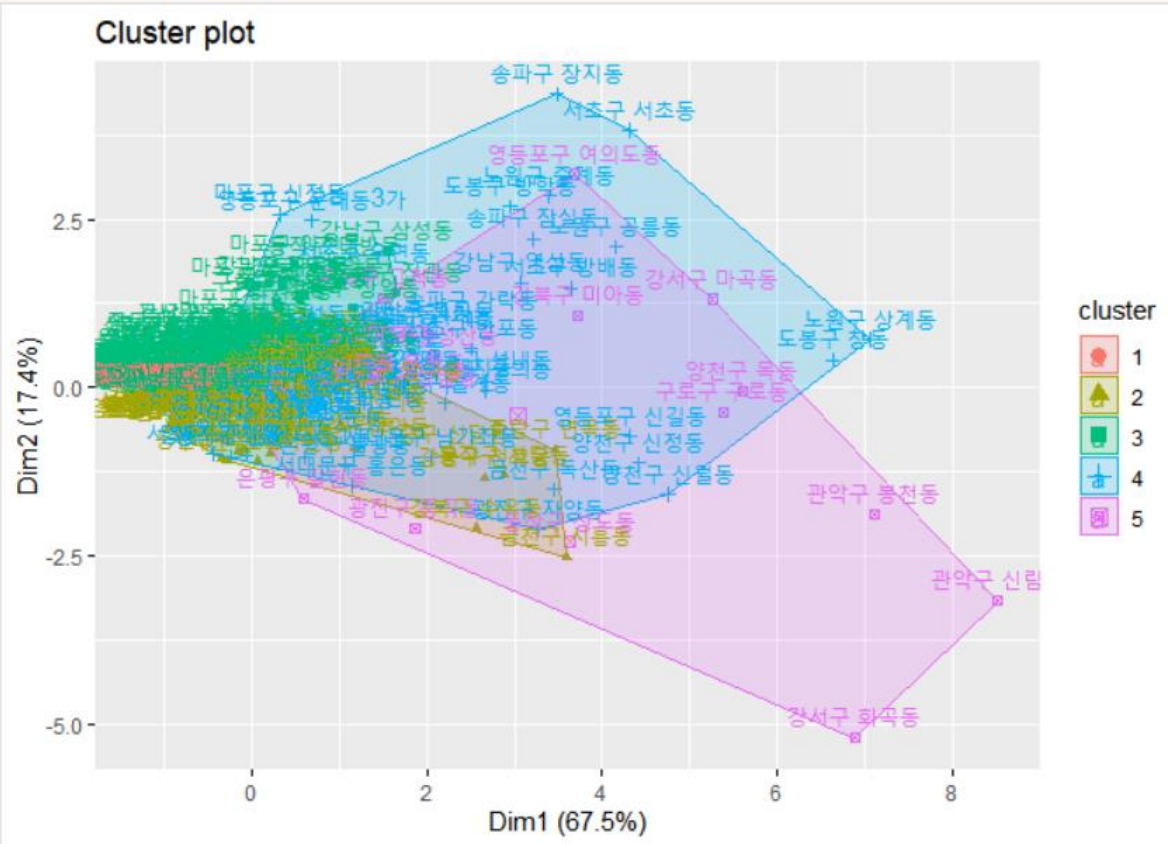

본론 GMM Clustering

mclust library의 Mclust 함수를 이용하여
GMM을 만든 결과 다음과 같이 5개로 분류됨을 확인할 수 있었음

```
gmm_model$G
## [1] 5
```

GMM 결과

fviz_cluster(gmm_model, scaled_final)



GMM 결과 프로파일링

Cluster 2가 Target Cluster

```
## # A tibble: 5 x 7
##   clst_k5 인구수 한명당.택배함 한명당.물류창고 한명당.친환경충전소 한명당.자전거거치대,
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1     1 2552.     0 0.0000218     0 0.00534 0
## 2     2 34862. 0.0000447 0.0000276 0.0000916 0.00194 0.0327
## 3     3 17737.     0 0.0000833 0.000230 0.00412 0.0234
## 4     4 50765. 0.0000359 0.0000676 0.000196 0.00341 0.0632
## 5     5 89102. 0.0000326 0.0000909 0.000102 0.00362 0.0362
## # ... with abbreviated variable names 1한명당.친환경충전소, 2한명당.자전거거치대,
## # 3친환경차.한대당.충전소.개수
```

GMM 결과 다음과 같은 법정동이 Target임을 확인할 수 있음

```
gmm_result <- gmm_cluster[gmm_model$classification == 2, "법정동"]
gmm_result
```

## [1] "종로구 궁정동"	"종로구 송인동"	"중구 남산동2가"
## [4] "중구 신당동"	"용산구 남영동"	"용산구 청파동1가"
## [7] "용산구 신창동"	"용산구 효창동"	"용산구 보광동"
## [10] "성동구 하왕십리동"	"성동구 사근동"	"성동구 금호동2가"
## [13] "동대문구 답십리동"	"동대문구 장안동"	"동대문구 회기동"
## [16] "중랑구 면목동"	"성북구 보문동6가"	"성북구 길음동"
## [19] "성북구 장위동"	"성북구 석관동"	"강북구 수유동"
## [22] "도봉구 쌍문동"	"은평구 구산동"	"은평구 신사동"
## [25] "은평구 중산동"	"서대문구 북아현동"	"서대문구 신촌동"
## [28] "마포구 도화동"	"마포구 대흥동"	"마포구 상수동"
## [31] "마포구 망원동"	"마포구 연남동"	"강서구 오곡동"
## [34] "구로구 가리봉동"	"구로구 개봉동"	"구로구 오류동"
## [37] "구로구 궁동"	"금천구 시흥동"	"영등포구 당산동1가"
## [40] "영등포구 당산동"	"영등포구 도림동"	"영등포구 대림동"
## [43] "동작구 사당동"	"관악구 남현동"	"강남구 일원동"
## [46] "송파구 송파동"	"송파구 석촌동"	"송파구 잠전동"
## [49] "송파구 마천동"	"강동구 길동"	"강동구 암사동"
## [52] "강동구 천호동"		

분석

각 클러스터링 기법에서 **두 번 이상 나온** 지역(법정동)을 선정하였을 때 다음과 같이 **21개의 법정동이 Target Cluster**

```
result <- unique(result_vector[duplicated(result_vector) | duplicated(result_vector, fromLast = TRUE)])  
print(result)
```

## [1] "광진구 자양동"	"중랑구 면목동"	"강북구 수유동"
## [4] "도봉구 창동"	"노원구 상계동"	"양천구 신정동"
## [7] "양천구 목동"	"양천구 신월동"	"강서구 화곡동"
## [10] "구로구 구로동"	"금천구 독산동"	"금천구 시흥동"
## [13] "영등포구 신길동"	"동작구 상도동"	"동작구 사당동"
## [16] "관악구 봉천동"	"관악구 신림동"	"강동구 천호동"
## [19] "중구 신당동"	"광진구 중곡동"	"서대문구 남가좌동"

본론

입지 기준 도출

급속 전기차 충전기 설치

Target Cluster로 선정된 법정동에 있는 주유소에 추가적으로 급속 전기차 충전기를 설치

입지 기준을 설정하기 위해 **Y = 급속충전기일 때** 이를 가장 잘 설명하는 변수 확인 (급속 충전기가 없는 법정동 제거)

	법정동	인구수	급속충전기 (대)	완속충전기 (대)	전기차 등록 수	전기차 충전소 개수	전기차 한대당 급속 충전소	전기차 한대당 충 전소	인구수 한명당 급속 충전기	친환경차_전기차 비율
0	강동구 천호동	85764	5	2	210	7	0.023810	0.033333	0.000058	0.878661
1	강북구 수유동	102828	1	0	185	1	0.005405	0.005405	0.000010	0.963542
2	강서구 화곡동	219202	2	0	399	2	0.005013	0.005013	0.000009	0.913043
3	관악구 봉천동	224184	7	6	439	13	0.015945	0.029613	0.000031	0.936034
4	관악구 신림동	245562	4	11	444	15	0.009009	0.033784	0.000016	0.942675
5	광진구 자양동	91273	6	10	252	16	0.023810	0.063492	0.000066	0.969231
6	구로구 구로동	115338	7	10	3313	17	0.002113	0.005131	0.000061	0.968430
7	금천구 독산동	103061	4	0	234	4	0.017094	0.017094	0.000039	0.947368
8	금천구 시흥동	101795	3	2	271	5	0.011070	0.018450	0.000029	0.960993
9	노원구 상계동	197531	10	16	462	26	0.021645	0.056277	0.000051	0.960499
10	도봉구 창동	116122	14	14	244	28	0.057377	0.114754	0.000121	0.949416
11	동작구 사당동	97424	2	6	153	8	0.013072	0.052288	0.000021	0.910714
12	동작구 상도동	114653	3	4	257	7	0.011673	0.027237	0.000026	0.908127
13	서대문구 남가 좌동	45963	1	7	105	8	0.009524	0.076190	0.000022	0.882353
14	양천구 목동	145972	8	9	351	17	0.022792	0.048433	0.000055	0.897698
15	양천구 신월동	119488	4	9	336	13	0.011905	0.038690	0.000034	0.965517
16	양천구 신정동	173388	9	15	486	24	0.018519	0.049383	0.000052	0.931034
17	영등포구 신길 동	104273	4	12	198	16	0.020202	0.080808	0.000038	0.916667
18	중구 신당동	68056	4	6	179	10	0.022346	0.055866	0.000059	0.957219
19	중랑구 면목동	135903	2	6	262	8	0.007634	0.030534	0.000015	0.939068

인구수 한명당 급속 충전기, 친환경차 중 전기차 비율, 전기차 한
대당 급속 충전소 등의 파생변수를 만들어서
급속 충전소 구축이 얼마나 잘 되어 있는지 비율로 확인

전기차 한 대 당 급속 충전소 $\frac{\text{급속 충전기 (대)}}{\text{전기차 등록 수}}$

인구수 한 명 당 급속 충전기 $\frac{\text{급속 충전기 (대)}}{\text{인구수}}$

전기차 한 대 당 충전소 $\frac{\text{전기차 충전소 개수}}{\text{전기차 등록 수}}$

친환경차 중 전기차 비율 $\frac{\text{전기차 등록 수}}{\text{친환경차 등록 대 수 (수소+전기)}}$

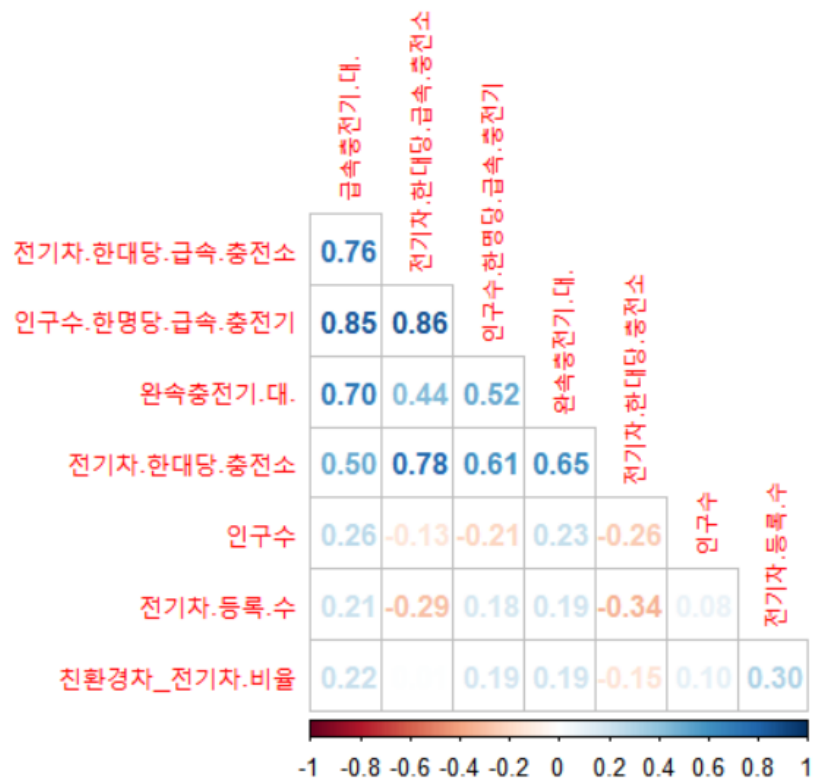
본론 입지 기준 도출

1. 다중공산성 제거

다중공산성은 회귀분석에서 사용된 모형의 일부 독립 변수가 다른 독립 변수와 상관 정도가 높아 데이터 분석 시 부정적인 영향을 미치므로 VIF를 통해 제거

```
df_cor <- df%>%  
  select(-법정동, -전기차.충전소.개수)
```

* 충전소 개수는 급속충전기.대. 변수와 완속충전기.대. 변수의 합이므로 제거



```
model <- lm(급속충전기.대.~, df_cor)  
vif(model)
```

##	인구수	완속충전기.대.	전기차.등록.수
##	4.719085	8.821511	7.254413
##	전기차.한대당.급속.충전소	전기차.한대당.충전소	인구수.한명당.급속.충전기
##	51.525154	15.823227	40.737862
##	친환경차_전기차.비율		
##	1.406167		

VIF가 10 이상인 변수를 제거



```
df2 <- df_cor%>%  
  select(인구수, 급속충전기.대., 완속충전기.대., 전기차.등록.수, 친환경차_전기차.비율)
```

본론 입지 기준 도출

2. 변수 선택 - Best Subset Selection

가능한 모든 변수들의 조합을 다 고려하는 방법인 Best Subset Selection을 진행

```
## best subset selection
response <- df2$급속충전기.대.
predictors <- df2[, c("인구수", "완속충전기.대.", "전기차.등록.수", "친환경차_전기차.비율")]
bestsub.model <- regsubsets(response ~ ., data = data.frame(response, predictors), nvmax = 4)
best_subset <- summary(bestsub.model)$which[which.min(summary(bestsub.model)$bic), ]
```

```
summary(bestsub.model)
```

```
## Subset selection object
## Call: regsubsets.formula(response ~ ., data = data.frame(response,
##      predictors), nvmax = 4)
## 4 Variables (and intercept)
##              Forced in Forced out
## 인구수              FALSE      FALSE
## 완속충전기.대.      FALSE      FALSE
## 전기차.등록.수      FALSE      FALSE
## 친환경차_전기차.비율 FALSE      FALSE
## 1 subsets of each size up to 4
## Selection Algorithm: exhaustive
##      인구수 완속충전기.대. 전기차.등록.수 친환경차_전기차.비율
## 1 ( 1 ) " "      "*"              " "              " "
## 2 ( 1 ) "*"      "*"              " "              " "
## 3 ( 1 ) "*"      "*"              " "      "*"
## 4 ( 1 ) "*"      "*"      "*"      "*"

```

Best Subset Selection 결과

```
cbind(
  RSS = summary(bestsub.model)$rss,
  Cp   = summary(bestsub.model)$cp,
  r2    = summary(bestsub.model)$rsq,
  Adj_r2 = summary(bestsub.model)$adjr2,
  BIC   =summary(bestsub.model)$bic
)
```

##		RSS	Cp	r2	Adj_r2	BIC
##	[1,]	108.0277	-0.382789	0.4904355	0.4621264	-7.4925122
##	[2,]	105.7772	1.291861	0.5010512	0.4423513	-4.9178373
##	[3,]	104.3850	3.090596	0.5076181	0.4152965	-2.1870842
##	[4,]	103.7583	5.000000	0.5105741	0.3800605	0.6882171

본론 입지 기준 도출

3. Factor Analysis – KMO 검정

변수들 간의 상관관계를 분석하여 공통차원들을 통해 축약하는 통계 기법으로서 변수들간 잠재요인 파악

Feature Selection에서 뽑힌 변수들 Scaling

```
predictors1 <- df2[, c("완속충전기.대.")]
predictors2 <- df2[, c("인구수", "완속충전기.대.")]
predictors3 <- df2[, c("인구수", "완속충전기.대.", "친환경차_전기차.비율")]
predictors4 <- df2[, c("인구수", "완속충전기.대.", "전기차.등록.수", "친환경차_전기차.비율")]

scaled_predictors1 <- scale(predictors1)
scaled_predictors2 <- scale(predictors2)
scaled_predictors3 <- scale(predictors3)
scaled_predictors4 <- scale(predictors4)
```

KMO 검정 결과

```
kmo_result2 <- KMO(scaled_predictors2)
print(kmo_result2)
```

```
## Call: KMO(r = scaled_predictors2)
## Overall MSA = 0.5
## MSA for each item =
##      인구수  완속충전기.대.
##      0.5      0.5
```

```
kmo_result3 <- KMO(scaled_predictors3)
print(kmo_result3)
```

```
## Call: KMO(r = scaled_predictors3)
## Overall MSA = 0.55
## MSA for each item =
##      인구수      완속충전기.대.  친환경차_전기차.비율
##      0.56      0.54      0.58
```

```
kmo_result4 <- KMO(scaled_predictors4)
print(kmo_result4)
```

```
## Call: KMO(r = scaled_predictors4)
## Overall MSA = 0.6
## MSA for each item =
##      인구수      완속충전기.대.      전기차.등록.수
##      0.59      0.61      0.59
## 친환경차_전기차.비율
##      0.59
```

본론

입지 기준 도출

3. Factor Analysis

변수가 인구수와 완속충전기.대.로 구성된 상태에서 친환경차_전기차.비율, 전기차.등록.수 변수를 추가할 수록 MSA값이 높아지므로, 인구수와 완속충전기.대.로 구성된 상태에서 변수를 하나씩 추가할 때 적합도가 향상되는지 anova() 함수를 통해 확인

```
lm2 <- lm(급속충전기.대.~인구수+완속충전기.대., data=df2)
```

```
lm3 <- lm(급속충전기.대.~인구수+완속충전기.대.  
+친환경차_전기차.비율, data=df2)
```

```
lm4 <- lm(급속충전기.대.~인구수+완속충전기.대.  
+전기차.등록.수  
+친환경차_전기차.비율, data=df2)
```

```
anova(lm2,lm3,lm4)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Model 1: 급속충전기.대. ~ 인구수 + 완속충전기.대.
```

```
## Model 2: 급속충전기.대. ~ 인구수 + 완속충전기.대. + 친환경차_전기차.비율
```

```
## Model 3: 급속충전기.대. ~ 인구수 + 완속충전기.대. + 전기차.등록.수 +
```

```
## 친환경차_전기차.비율
```

```
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
```

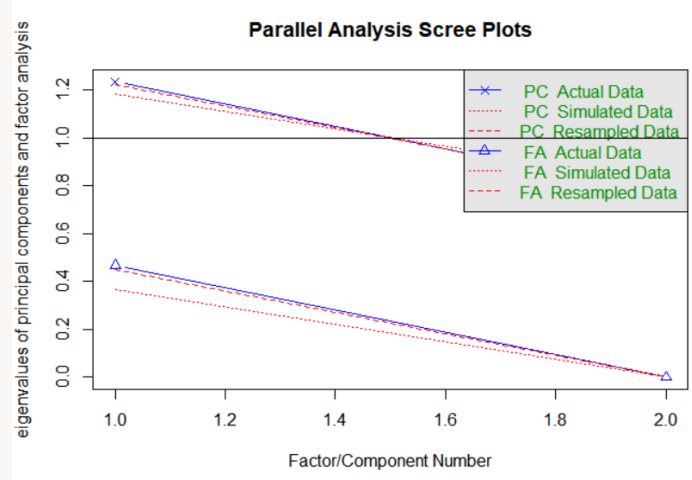
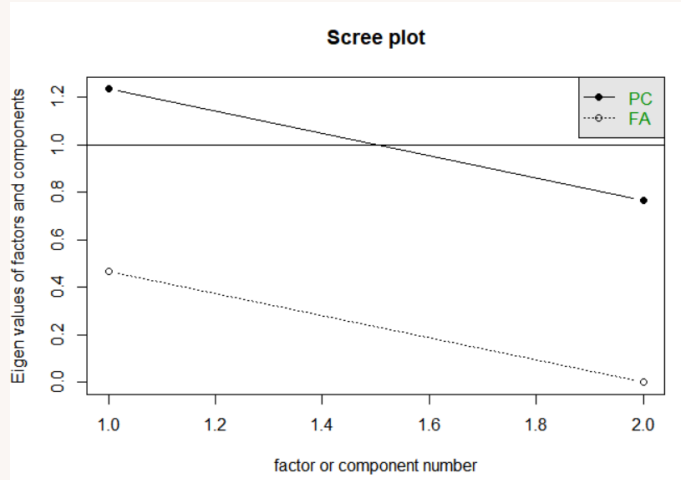
```
## 1      17 105.78
```

```
## 2      16 104.39  1   1.39219 0.2013 0.6601
```

```
## 3      15 103.76  1   0.62667 0.0906 0.7676
```

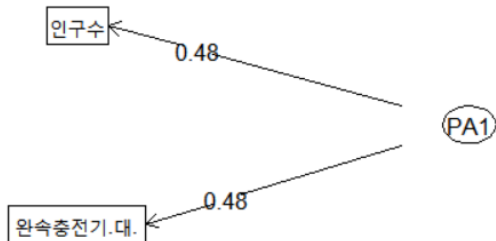
p값이 유의하지 않으므로
Model 1에 변수를 추가적으로 더해도
적합도 향상 X

3. Factor Analysis - 요인 개수 선택



요인 개수 1개로 선택

Factor Analysis



```
pa1 <- fa(final_df[-1], nfactors=1, fm="pa", max.iter = 100, rotate="oblimin")
fa.diagram(pa1, digits = 2)
```

요인 개수 1개로 인구수와 완속 충전기 변수를 설명

인구수 + 완속 충전기(대)를 통해 충전환경이라는 잠재요인 변수 생성

본론

입지 기준 도출

1. 입지 분석 - 최종 변수

- 충전환경이라는 급속 충전기 잠재요인 변수와 / 급속 충전기, 공공 자전거 거치대 개수가 상대적으로 적은 곳으로 결과 값을 도출하기 위해
각 시설의 결핍률 변수를 추가

충전 환경 : 급속 충전기 잠재 요인

급속 충전기 결핍률 : 현재 서울시에 설치된 급속 충전기 개수가 상대적으로 적은 곳으로 결과값을 도출하기 위함

공공자전거 거치대 결핍률 : 현재 서울시에 설치된 공공 자전거 대여소 개수가 상대적으로 적은 곳으로 결과값을 도출하기 위함

택배함 결핍률 : 현재 서울시에 설치된 택배함 개수가 상대적으로 적은 곳으로 결과값을 도출하기 위함

물류창고 결핍률 : 현재 서울시에 설치된 물류창고 개수가 상대적으로 적은 곳으로 결과값을 도출하기 위함

$$\text{급속 충전기 결핍률} = 1 - \frac{\text{서울시 급속 충전기 개수}}{\text{인구 수}}$$

$$\text{공공 자전거 거치대 결핍률} = 1 - \frac{\text{서울시 공공 자전거 대여소 (거치대) 개수}}{\text{인구 수}}$$

$$\text{택배 함 결핍률} = 1 - \frac{\text{서울시 택배 함 개수}}{\text{인구 수}}$$

$$\text{물류 창고 결핍률} = 1 - \frac{\text{서울시 물류 창고 개수}}{\text{인구 수}}$$

본론

입지 기준 도출

2. 주성분 분석 (PCA)을 활용한 평가 지표 개발

- PCA는 데이터의 분산을 최대한 보존할 수 있는 방향으로 변수를 변형시켜 주성분들을 찾는 차원 축소 기법
- **PCA의 Loading**을 평가지표의 **가중치**로 사용 *PCA의 Loading은 단위 척도화된 주성분 혹은 요인들이 한 변수를 정의 혹은 적재하는 선형 결합된 가중치
- Proportion of Variance(분산의 비율)가 충분히 높거나, Cumulative Proportion(누적 분산비)가 높은 PCA의 Loading을 평가 지표의 가중치로 사용

3. 평가 지표 적용

Scaling 적용한 PCA 결과

```
final_pca2 <- prcomp(final[-1],scale=T)
summary(final_pca2)
```

Importance of components:

##	PC1	PC2	PC3	PC4	PC5
## Standard deviation	1.4149	1.2011	0.8928	0.66912	0.55735
## Proportion of Variance	0.4004	0.2885	0.1594	0.08954	0.06213
## Cumulative Proportion	0.4004	0.6889	0.8483	0.93787	1.00000

주성분 3개일 때 전체 변동의 84.8% 정도를 설명
분산의 비율도 PC3에서 PC4로 갈때 급격히 감소하므로 주성분은 3개면 충분
PC3의 loading을 평가지표의 가중치로 사용

Loading Matrix

```
loadings <- final_pca2$rotation
round(loadings,2)
```

##	PC1	PC2	PC3	PC4	PC5
## 충전환경	0.50	-0.48	0.08	-0.32	0.65
## 급속충전기결핍률	0.46	0.22	-0.72	0.46	0.13
## 공공자전거거치대결핍률	0.52	0.40	-0.03	-0.54	-0.41
## 택배함결핍률	0.39	-0.60	0.15	0.32	-0.60
## 물류창고결핍률	0.36	0.45	0.68	0.43	0.18

Score = 0.08 x 충전환경 + 0.72 x 급속충전기결핍률 + 0.03 x 공공자전거거치대결핍률 + 0.15 x 택배함결핍률 + 0.68 x 물류창고결핍률

Loading은 PC에 대한 각 변수의 영향력을 의미
PCA는 Y 변수가 없으므로 음수가 나왔을 때 절댓값을 씌워 영향력 파악 가능

본론 입지 기준 도출

4. 입지 분석 - 주유소 선정

```
final0 <- final%>%  
  select(법정동)
```

```
final0$충전환경 <- final$충전환경*0.08  
final0$급속충전기 <- final$급속충전기결핍률*0.72  
final0$공공자전거거치 <- final$공공자전거거치대결핍률*0.03  
final0$택배함 <- final$택배함결핍률*0.15  
final0$물류창고 <- final$물류창고결핍률*0.68
```

```
final0$score <- final0$충전환경+final0$급속충전기+final0$공공자전거거치+final0$택배함+final0$물류창고
```

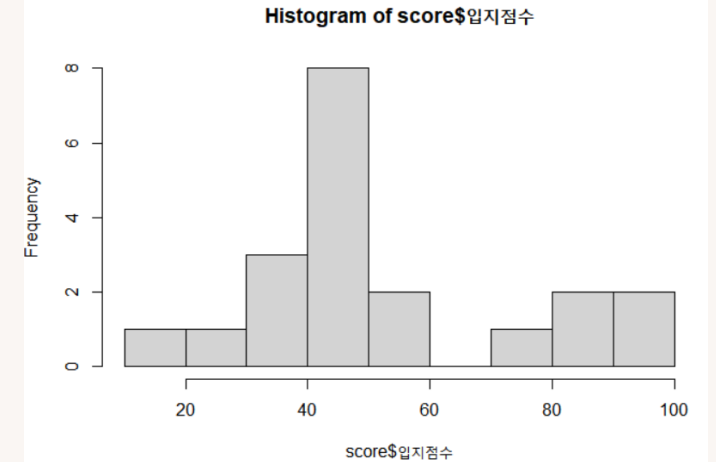
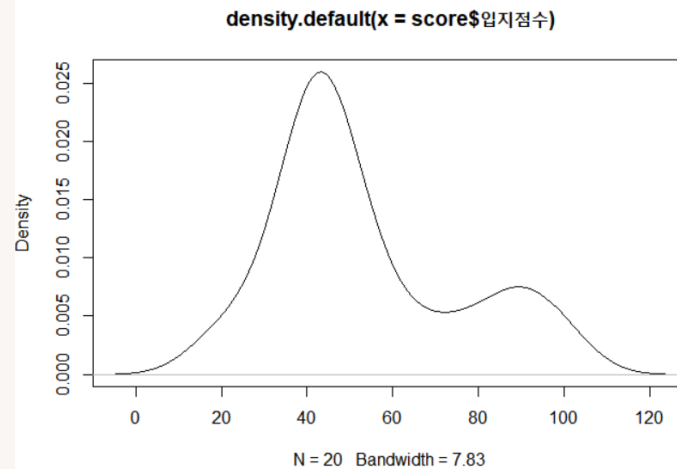
```
score <- final0%>%  
  select("법정동", "score")  
score$입지점수 <- (score$score/max(score$score))*100
```

$$\text{입지 지수} = \frac{\text{score}}{\max(\text{score})} \times 100$$

```
quantile(score$입지점수, probs=seq(0,1,0.01))
```

##	0%	1%	2%	3%	4%	5%	6%	7%
##	18.73310	20.44207	22.15105	23.86002	25.56899	27.27797	28.73682	30.10635
##	8%	9%	10%	11%	12%	13%	14%	15%
##	31.47588	32.84541	34.21494	35.13791	35.56469	35.99147	36.41825	36.84502
##	16%	17%	18%	19%	20%	21%	22%	23%
##	37.28206	37.75758	38.23309	38.70861	39.18412	39.65964	40.00470	40.34252
##	24%	25%	26%	27%	28%	29%	30%	31%
##	40.68034	41.01816	41.35597	41.51722	41.59698	41.67673	41.75649	41.83624
##	32%	33%	34%	35%	36%	37%	38%	39%
##	41.89001	41.90803	41.92606	41.94408	41.96210	41.99223	42.08692	42.18160
##	40%	41%	42%	43%	44%	45%	46%	47%
##	42.27629	42.37097	42.46566	43.19352	43.99586	44.79821	45.60056	46.40291
##	48%	49%	50%	51%	52%	53%	54%	55%
##	46.73227	46.78572	46.83918	46.89263	46.94608	47.00230	47.06326	47.12422
##	56%	57%	58%	59%	60%	61%	62%	63%
##	47.18517	47.24613	47.32804	47.58804	47.84804	48.10803	48.36803	48.62803
##	64%	65%	66%	67%	68%	69%	70%	71%
##	49.73820	51.00779	52.27737	53.54696	54.81654	55.80219	56.58133	57.36047
##	72%	73%	74%	75%	76%	77%	78%	79%
##	58.13961	58.91875	60.12173	62.24301	64.36430	66.48559	68.60687	70.71481
##	80%	81%	82%	83%	84%	85%	86%	87%
##	72.58252	74.45023	76.31794	78.18565	80.05336	81.76907	83.44423	85.11940
##	88%	89%	90%	91%	92%	93%	94%	95%
##	86.79457	88.46974	89.46632	89.85218	90.23804	90.62389	91.00975	91.72936
##	96%	97%	98%	99%	100%			
##	93.38349	95.03762	96.69175	98.34587	100.00000			

입지점수의 분포를 고려하여 입지점수가 80점 이상인 법정동을 최종 입지로 선정



본론 입지 기준 도출

4. 입지 분석 - 주유소 선정

```
score2 <- score%>%  
  arrange(desc(입지점수)) %>%  
  select(법정동, 입지점수)  
score2
```

##	법정동	입지점수
## 1	관악구 신림동	100.00000
## 2	관악구 봉천동	91.29407
## 3	강서구 화곡동	89.26324
## 4	노원구 상계동	80.44657
## 5	양천구 신정동	70.61651
## 6	양천구 목동	59.45185
## 7	중랑구 면목동	55.35111
## 8	양천구 신월동	48.66908
## 9	도봉구 창동	47.30067
## 10	구로구 구로동	46.97984
## 11	동작구 상도동	46.69851
## 12	영등포구 신길동	42.47563
## 13	금천구 독산동	41.97728
## 14	강북구 수유동	41.88242
## 15	금천구 시흥동	41.46265
## 16	동작구 사당동	39.68466
## 17	광진구 자양동	37.18196
## 18	강동구 천호동	34.93575
## 19	중구 신당동	27.72770
## 20	서대문구 남가좌동	18.73310

```
score <- score %>%  
  filter(입지점수 >= 80) %>%  
  arrange(desc(입지점수)) %>%  
  select(법정동, 입지점수)  
score
```

##	법정동	입지점수
## 1	관악구 신림동	100.00000
## 2	관악구 봉천동	91.29407
## 3	강서구 화곡동	89.26324
## 4	노원구 상계동	80.44657

입지점수가 80점 이상인 **관악구 신림동, 관악구 봉천동, 강서구 화곡동, 노원구 상계동** 우선적으로 첨단 물류 복합 주유소가 필요한 **최종 입지**

*고려사항

- 가장 가까운 **충전소까지의 거리**가 먼 주유소일수록 급속 충전기 설치 필요
 - Haversine Method를 이용해 각 주유소에서 해당 법정동에 위치한 모든 충전소까지의 거리 계산
 - 각 주유소에서 가장 가까운 충전소까지의 거리 확인, 거리가 먼 주유소 선정

- 첨단 물류 복합 주유소 드론 운영, **비행 금지 구역**인지 확인 필요
 - 대통령 집무실이 용산으로 바뀌면서 서울의 비행금지구역이 변경
 - 2023년 기준 비행금지구역 P73에 선정한 주유소가 있는지 확인

본론 입지 기준 도출

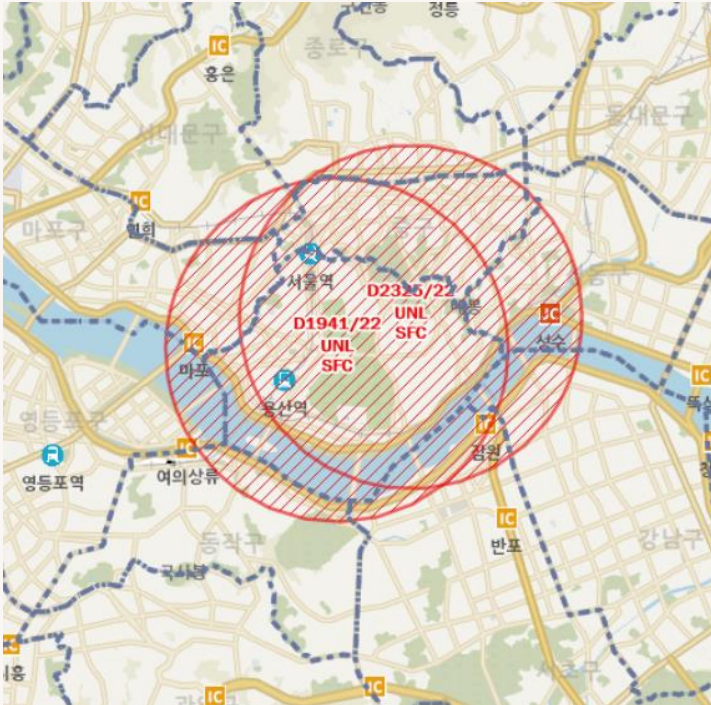
5. 최종 선택 주유소

주유소 데이터 로드

```
gas_station_df = pd.read_excel("score_gas_station_xy.xlsx")
elec_station_df = pd.read_excel("elec_station.xlsx")
```

비행 금지 구역 좌표 설정

```
forbidden_area_centers = [(37.3209, 126.5838), (37.3232, 126.5943)]
forbidden_area_radius = 3.7 # 단위: km
```



2023년 서울시 드론 P-73(비행금지구역)

두 지점 간의 하버사인 거리 계산 함수

```
def haversine_distance(lat1, lon1, lat2, lon2):
    R = 6371 # 지구 반지름 (단위: km)
    dlat = math.radians(lat2 - lat1)
    dlon = math.radians(lon2 - lon1)
    a = math.sin(dlat / 2) ** 2 + math.cos(math.radians(lat1)) * math.cos(math.radians(lat2)) * math.sin(dlon / 2) ** 2
    c = 2 * math.atan2(math.sqrt(a), math.sqrt(1 - a))
    distance = R * c
    return distance
```

비행 금지 구역 내에 있는 주유소 제외한 데이터프레임 생성

```
filtered_gas_station_df = gas_station_df.copy()
for center in forbidden_area_centers:
    lat_center, lon_center = center
    filtered_gas_station_df = filtered_gas_station_df[filtered_gas_station_df.apply(
        lambda row: haversine_distance(row["Latitude"], row["Longitude"], lat_center, lon_center) > forbidden_area_radius,
        axis=1
    )]
```

결과 출력

print(filtered_gas_station_df)

	법정동	주소	Latitude	Longitude
0	화곡동	서울특별시 강서구 공향대로 432(화곡6동 1125)화곡동	37.554668	126.854805
1	화곡동	서울특별시 강서구 화곡로 273(화곡6동 983-1)화곡동	37.549163	126.846996
2	화곡동	서울특별시 강서구 국회대로 225(화곡4동 843-4)화곡동	37.529621	126.859797
3	화곡동	서울특별시 강서구 국회대로 71(화곡1동 931-11)화곡동	37.527258	126.842890
4	화곡동	서울특별시 강서구 공향대로 207(화곡2동 862-16)화곡동	37.531580	126.856052
5	화곡동	서울특별시 강서구 국회대로 251(화곡4동 785-2)화곡동	37.530040	126.862983
6	화곡동	서울특별시 강서구 강서로 151(화곡1동 402-1)화곡동	37.540261	126.841198
7	화곡동	서울특별시 강서구 강서로 154(화곡본동 1076-14)화곡동	37.540719	126.841530
8	봉천동	서울특별시 관악구 남부순환로 1920 (인현동 1659-1)봉천동	37.476547	126.962648
9	봉천동	서울특별시 관악구 남부순환로 1691 (은천동 948-17)봉천동	37.486514	126.941524
10	봉천동	서울특별시 관악구 남부순환로 1753 (은천동 912-15)봉천동	37.486514	126.941524
11	봉천동	서울특별시 관악구 남부순환로 1883 (행운동 1679-30)봉천동	37.481060	126.960857
12	봉천동	서울특별시 관악구 보라매로 41 (보라매동 726-10)봉천동	37.493202	126.926737
13	봉천동	서울특별시 관악구 남부순환로 1880 (낙성대동 1663-11)봉천동	37.474160	126.958181
14	봉천동	서울특별시 관악구 보라매로 26 (보라매동 729-4)봉천동	37.491687	126.927160
15	신림동	서울특별시 관악구 난곡로 206 (난곡동 616-1)신림동	37.473410	126.918793
16	신림동	서울특별시 관악구 신림로 290 (서원동 1641-58)신림동	37.486790	126.927611
17	신림동	서울특별시 관악구 난곡로 108 (난향동 666-14)신림동	37.461990	126.918499
18	신림동	서울특별시 관악구 문성로 97 (미성동 1477-5)신림동	37.477794	126.914302
19	신림동	서울특별시 관악구 남부순환로 1520 (미성동 1571-8)신림동	37.476024	126.915989
20	신림동	서울특별시 관악구 신림로 176 (서림동 112-47)신림동	37.473627	126.938828
21	신림동	서울특별시 관악구 남부순환로 1515 (신사동 507-24)신림동	37.485691	126.918659
22	상계동	서울특별시 노원구 노원로26길 191(상계동)상계동	37.658394	127.071138
23	상계동	서울특별시 노원구 노원로 527(상계동)상계동	37.655939	127.066406
24	상계동	서울특별시 노원구 동일로 1772(상계동)상계동	37.686659	127.055818

본론 입지 기준 도출

5. 최종 선택 주유소

```
# 가장 가까운 충전소 거리 계산
gas_station_df["dis_elec"] = float('inf') # 초기값으로 무한대 설정
for _, elec_row in elec_station_df.iterrows():
    elec_lat, elec_lon = elec_row["Latitude"], elec_row["Longitude"]
    gas_station_df["Distance"] = gas_station_df.apply(
        lambda row: haversine_distance(row["Latitude"], row["Longitude"], elec_lat, elec_lon),
        axis=1
    )
gas_station_df["dis_elec"] = gas_station_df[["dis_elec", "Distance"]].min(axis=1)

# 가장 가까운 충전소와 매일 주소 거리 계산
gas_station_df["dis_address"] = float('inf') # 초기값으로 무한대 설정
for _, address_row in mail_address_df.iterrows():
    address_lat, address_lon = address_row["Latitude"], address_row["Longitude"]
    gas_station_df["dis_gas"] = gas_station_df.apply(
        lambda row: haversine_distance(row["Latitude"], row["Longitude"], address_lat, address_lon),
        axis=1
    )
gas_station_df["dis_address"] = gas_station_df[["dis_address", "dis_gas"]].min(axis=1)

result_df["sum_dis"] = result_df["dis_bike"] + result_df["dis_elec"] + result_df["dis_address"]
print(result_df)
```

	법정동	Latitude	Longitude	dis_bike	dis_elec	dis_address	sum_dis
0	화곡동	37.554668	126.854805	0.629078	2.392421	0.583660	3.605158
1	화곡동	37.549163	126.846996	0.295716	1.503003	0.450540	2.249258
2	화곡동	37.529621	126.859797	0.389992	0.971863	0.527042	1.888897
3	화곡동	37.527258	126.842890	0.493867	0.553404	0.533816	1.581088
4	화곡동	37.531580	126.856052	0.666669	0.704420	0.138499	1.509588
5	화곡동	37.530040	126.862983	0.160850	1.255868	0.551146	1.967864
6	화곡동	37.540261	126.841198	0.154077	0.391365	0.325201	0.870643
7	화곡동	37.540719	126.841530	0.138532	0.449855	0.274768	0.863156
8	봉천동	37.476547	126.962648	10.654848	0.258459	0.285818	11.199125
9	봉천동	37.486514	126.941524	8.493601	0.444271	0.514397	9.452269
10	봉천동	37.486514	126.941524	8.493601	0.444271	0.514397	9.452269
11	봉천동	37.481060	126.960857	10.242812	0.210314	0.340756	10.793883
12	봉천동	37.493202	126.926737	7.002997	0.717156	1.050450	8.770603
13	봉천동	37.474160	126.958181	10.493184	0.219920	0.638931	11.352035
14	봉천동	37.491687	126.927160	7.135752	0.546051	0.955026	8.636829
15	신림동	37.473410	126.918793	8.079294	0.124683	0.582609	8.786585
16	신림동	37.486790	126.927611	7.514963	0.000000	0.756466	8.271429
17	신림동	37.461990	126.918499	9.111208	0.127155	1.739584	10.977947
18	신림동	37.477794	126.914302	7.451469	0.503953	0.175620	8.131042
19	신림동	37.476024	126.915989	7.698220	0.257848	0.246384	8.202453
20	신림동	37.473627	126.938828	9.235526	0.227901	0.239308	9.702734
21	신림동	37.485691	126.918659	7.029246	0.057297	0.048775	7.135319
22	상계동	37.658394	127.071138	0.086559	0.480492	0.305428	0.872479
23	상계동	37.655939	127.066406	0.097205	0.060736	0.214360	0.372301
24	상계동	37.686659	127.055818	0.263639	0.401524	0.979209	1.644971

```
# 법정동 별로 sum_dis 값이 가장 큰 행 선택
max_sum_dis_row = result_df.groupby("법정동")["sum_dis"].idxmax()
max_sum_dis_df = result_df.loc[max_sum_dis_row]
```

```
# 맵핑을 위한 Latitude와 Longitude 추출
max_sum_dis_lat = max_sum_dis_df["Latitude"]
max_sum_dis_lon = max_sum_dis_df["Longitude"]
```

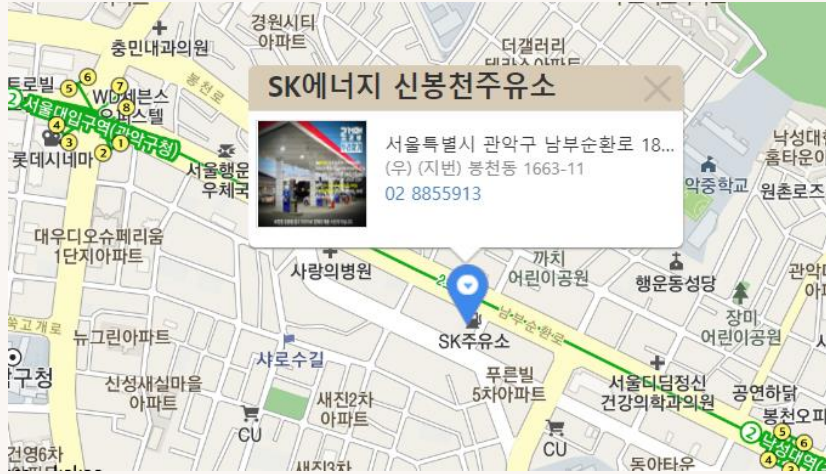
```
# 맵핑 결과 출력
mapping_result_df = pd.DataFrame({"법정동": max_sum_dis_df["법정동"], "Latitude": max_sum_dis_lat, "Longitude": max_sum_dis_lon})
print(mapping_result_df)
```

	법정동	Latitude	Longitude
13	봉천동	37.474160	126.958181
24	상계동	37.686659	127.055818
17	신림동	37.461990	126.918499
0	화곡동	37.554668	126.854805



법정동	주소	Latitude	Longitude
13 봉천동	서울특별시 관악구 남부순환로 1880 (낙성대동 1663-11)봉천동	37.474160	126.958181
24 상계동	서울특별시 노원구 동일로 1772(상계동)상계동	37.686659	127.055818
17 신림동	서울특별시 관악구 난곡로 108 (난향동 666-14)신림동	37.461990	126.918499
0 화곡동	서울특별시 강서구 공항대로 432(화곡6동 1125)화곡동	37.554668	126.854805

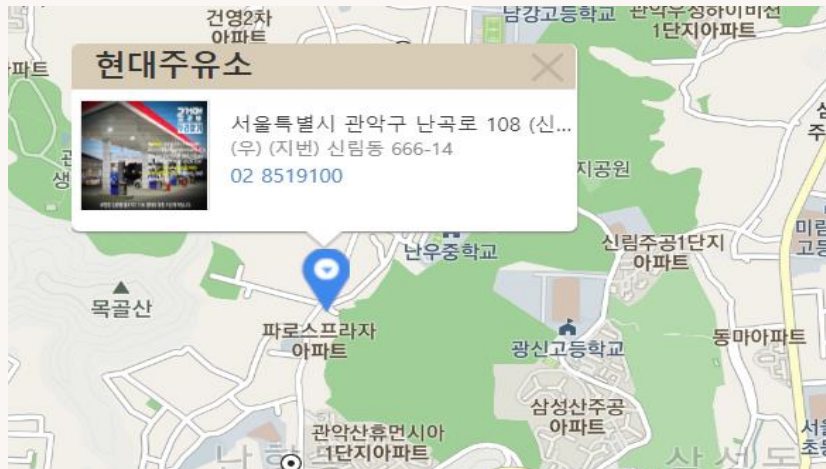
결론 최종 주유소



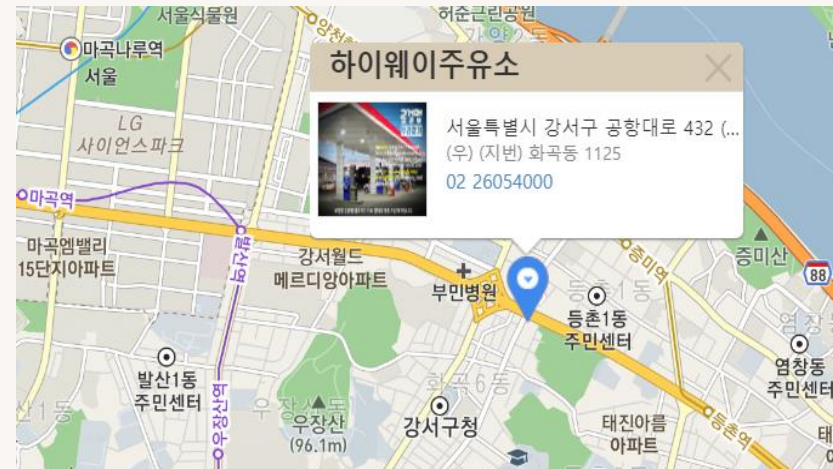
1.봉천동



2.상계동

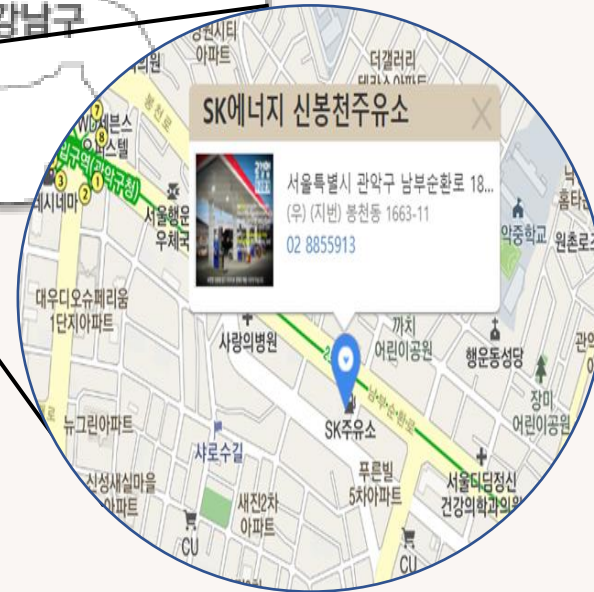
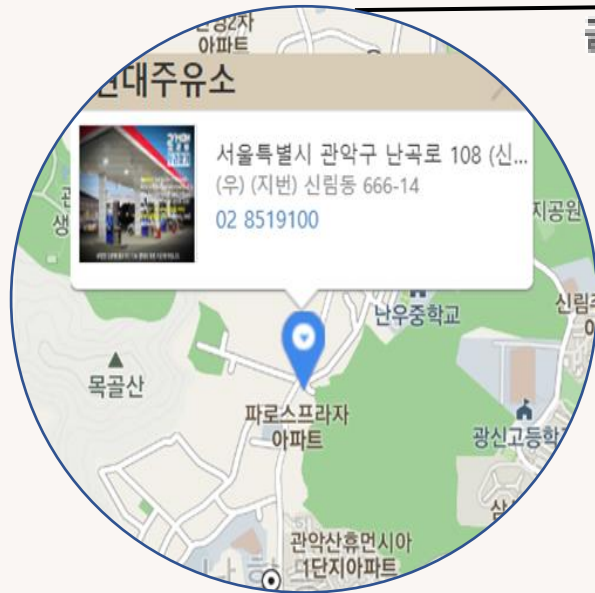
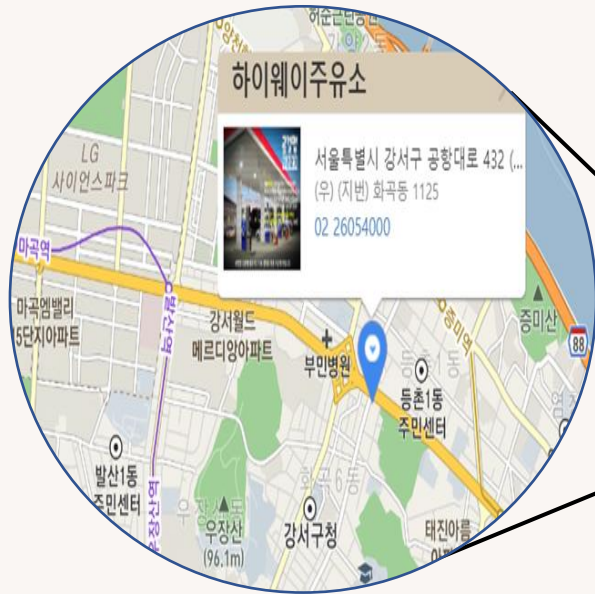


3.신림동



4.화곡동

결론 최종 주유소



기대 효과

탄소 배출 감소

- 서울시 기후 변화 대응에 따른 방향성과 일치해 결과적으로 탄소 배출 감소에 기여 가능성 상승
- 전기차 포함 **친환경 차량의 인프라 구축 확대**에 기여

현황에 맞는 입지 제시

- 행정동 별 인구 현황과 자동차 현황 등 **다양한 변수를 반영하여** 현재 행정동 상황에 맞는 **최적의 입지 제공**
- 전기차 신차 / 누적 판매량 증가에 따른 시장의 트렌드 반영

주유소 생존 방향 제시

- 기존의 주유 서비스 포함한 다양한 이윤 창출의 서비스를 제공
- 출혈 경쟁이 아닌 새로운 사업 제시로 **기존 석유 판매업자들의 경제 활동 지속 가능성 제공**
- 첨단 물류 복합 주유소로의 전환으로 주유소 폐업 위기 극복 가능

물류 서비스의 접근성 향상

- 드론 / 미래 물류 기술을 실증해 생활 물류 거점으로 활용
- 첨단 기술들을 통해 **도시 내 스마트 & 자동화 물류 서비스 구축 가능성 향상**
- 무인, 자동화 물류시설인 스마트MFC(마이크로 풀필먼트 센터) 조성을 통해 물류 서비스의 접근성 향상

한계점

정성적 측면 간과

- 사람들의 주유소 일별, 월별 이용량 / 가격에 따른 이용량 등 **정성적 요소**는 고려하지 못함

프로젝트적 한계

- 기존의 주유소를 첨단 물류 복합 주유소로 변환하는데 인적, 물적 자원 및 **추가적인 예산 확보**가 필요함
- 각 주유소 마다의 **재무 상태** 고려하지 않았으며, 첨단 물류 복합 주유소로 **변환을 강요할 수 없음**

알고리즘적 한계

- **이상치 (Outlier)**에 영향을 많이 받는다는 알고리즘적 한계가 존재함
- 도로명주소를 **경위도로 변환**함에 따라, 실제 위치와 미세한 차이가 발생

발전 방향

예산 추가 확보

- 첨단 물류 복합 주유소 전환 장려금, 기업의 투자 등 예산을 추가로 확보한다면 **전국적으로 시행 가능**

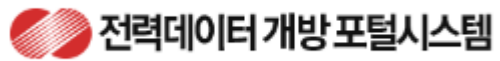
정성적 데이터 요소 추가

- 정성적 데이터 요소를 **수치로 구체화한 데이터를 추가**한다면 더욱 현실적인 입지를 선정 가능함

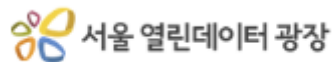
분석 활용 데이터



- 행정안전부 지역별 (법정동) 성별 연령별 주민 등록 인구 수 (2023)
- 한국가스공사 수소충전소 현황 (2023)
- 서울특별시 행정동 별 자동차 등록 대수 현황 (2022)
- 행정안전부 물류 창고 업체 등록 현황 (2023)



- 전기차 충전소 설치 현황 (2022)



- 서울특별시 공공자전거 대여소 정보 (2022)



- 서울특별시 안심 택배함 운영 현황 (2023)

분석 Tool



참고 문헌

- 21년말기준 자동차 등록대수 2,491만 대... 국토교통부 22.02.03
- 국토부, 전기차 등록대수 30만대 돌파. 한경 경제. 22.10.31
- 서울시 '복합주유소' 추진 로봇 물류, 드론 배송. 피플투데이. 22.11.29
- 전국 주유소 개수, 알뜰주유소 개수. MTN. 21.09.27
- 전국 주유소 개수 추이. 헤럴드 경제. 21.12.14
- 서울시 2030년까지 신재생 에너지 보급률 4%에서 21%로 확대한다. 경향신문 .22.03.21
- 서울시, 기후변화 '2050 탄소중립' 종합계획 발표 ... 전기차 10% 추진. 뉴스로. 22.01.20