



경상남도 창원시 전기버스 노선추천

대 통 운 수

[팀원] 고태영 박현경 배수지 정현준

CONTENTS



1

공모배경

- 창원시 전기버스 도입
- 분석목표

2

데이터 정의

- 활용데이터 정의 및 세부내용

3

데이터 전처리

- 데이터 전처리 진행방향

4

데이터 분석

- 데이터 분석기법
- 데이터 분석결과

5

서비스 활용방안 및 기대효과

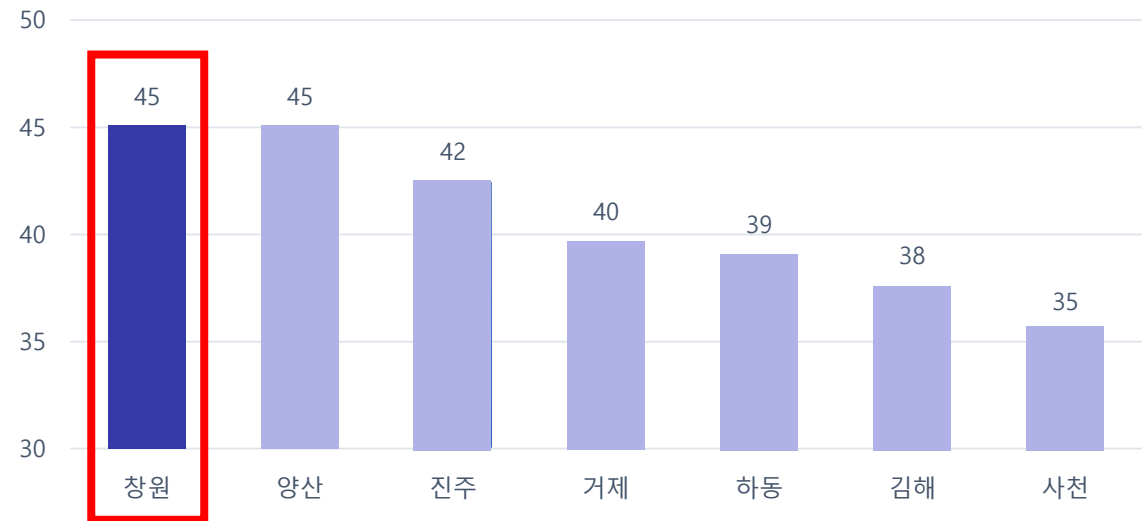
창원국가산업단지의 조성과 대기오염

경상남도 내에서 창원시의 미세먼지 농도가 가장 높다!



1974년 창원국가산업단지 위치도

- ✓ 산업기지개발구역으로 지정
- ✓ 국내 최초의 계획도시로 발전



2017~2018 경남 미세먼지 평균

미세먼지 저감을 위한 6+4 정책

“6+4정책”이란?

6개의 새로운 미세먼지 저감정책



4개의 강화된 미세먼지 저감정책

목표 : 2022년까지 미세먼지 50%저감

예산 : 5년간 시비 800억 포함 3000억원 투입예정

전기·수소 시내버스 도입

친환경 건설기계

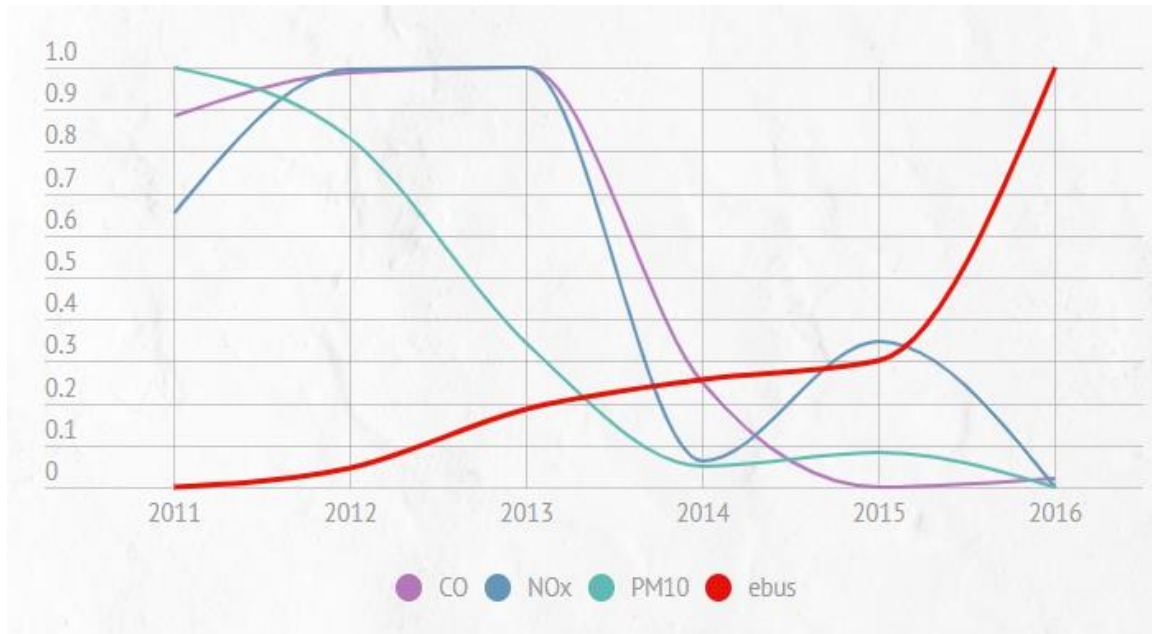
공원 및 운동장 미세먼지 억제제 살포

미세먼지 안심공간

버스정보시스템 미세먼지 안내판 설치

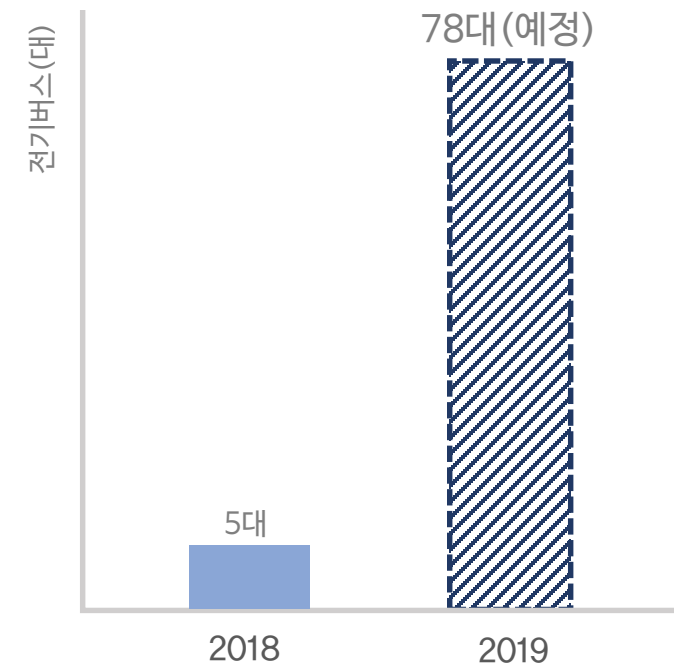
미세먼지 클린 시스템

전기버스 도입효과와 창원시의 전기 버스 도입



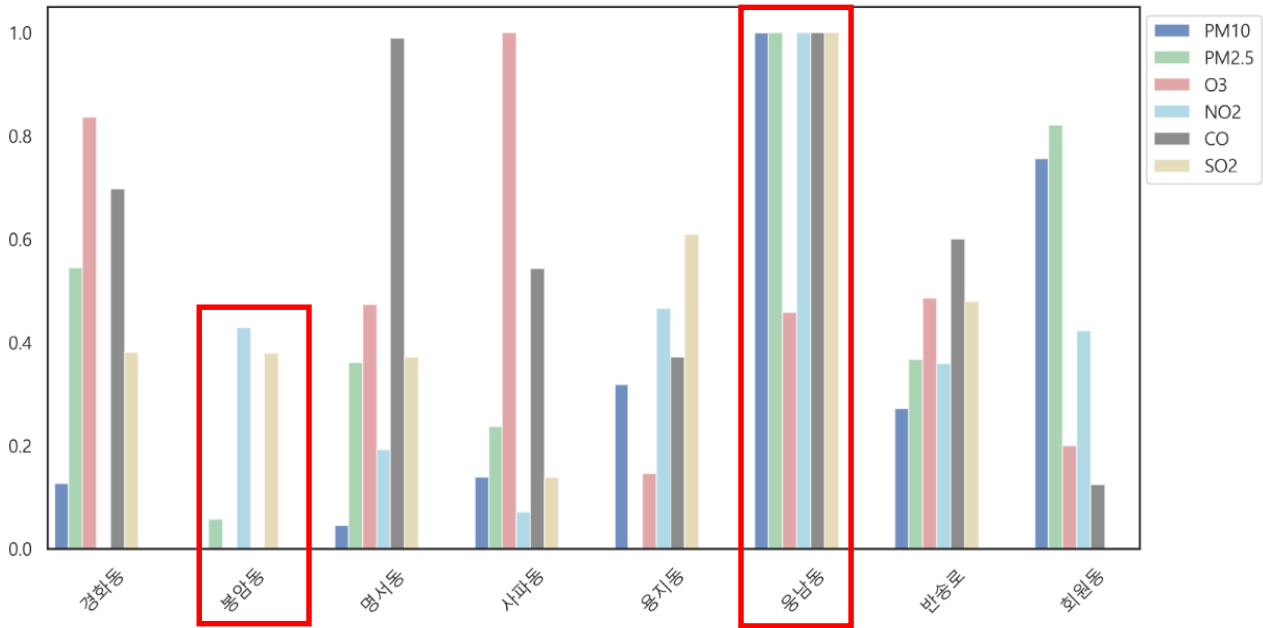
전기버스와 대기오염물질농도

✓ 전기버스증가에 따라 대기오염물질농도 감소



창원시의 전기버스도입

관측소별 대기오염물질 농도와 버스정류장 이용인원



창원시 관측소별 대기오염물질 농도

버스정류장명	월요일(명)
지산	2
창원우편집중국	9
경상대병원동문	11
...	...
창원종합터미널	2894
정우상가	4023
창원역	8150

버스정류장 요일별 이용인원수(월요일)



대기오염물질 **저감효과**를

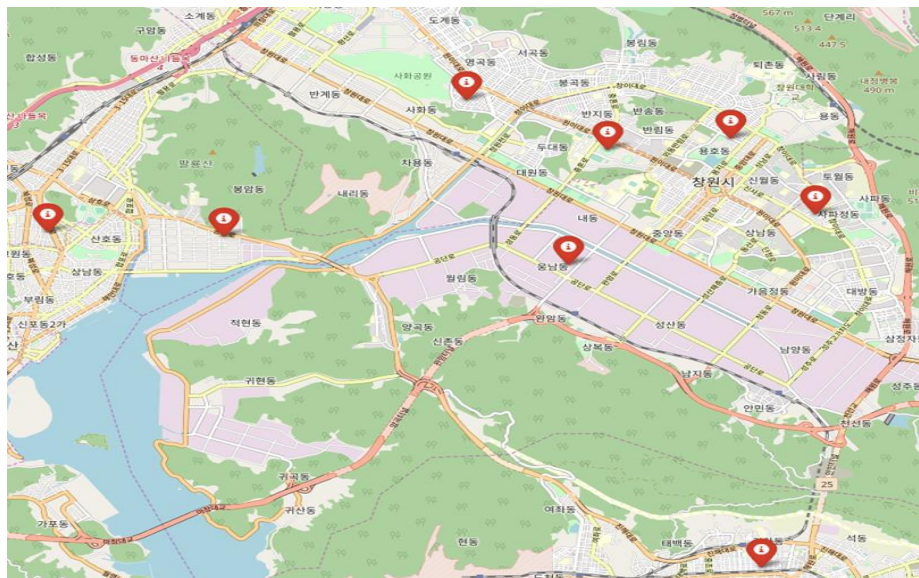
극대화하기 위한 **전기버스** 노선은?

활용 데이터

	활용 데이터			
출처				
기간	2018.01.01~2018.12.31	2019년도 기준	2019년도 기준	2018.01.01~2018.12.31
내용	대기오염 물질 농도	노선/정류소 정보	창원시_버스정류장 위치정보	대중교통 이용지표
변수	PM10, PM2.5, CO, O ₃ , NO ₂ , SO	정류소명, 버스번호, 첫차시간, 막차시간, 배차간격	관리번호, 정류장명, 위도, 경도	정류장별 이용인원 : 요일별, 정류장명

변수의 특징에 따라 결측치 대처

- PM10, PM2.5 : 일별에 따라 값에 차이가 큼
→ 주변의 다른 관측소들의 평균으로 값을 대체하는 **다중대치법** 사용
- CO, O₃, NO₂, SO₂ : 지역별에 따라 값에 차이가 큼
→ 관측소의 전날과 다음날 측정소에서 측정된 값을 평균으로 대체하는 **단순평균대치법** 사용



[창원시 8개 대기오염물질 측정소]

측정소	날짜	PM10	PM2.5	O ₃	NO ₂	CO	SO ₂
경화동	18.01.01	42.23924	21.719	0.460562	0.034299	0.017197	0.003902
봉암동	18.03.15	42.45745	21.81722	0.46233	0.03433	0.01725	0.003928
명서동	18.05.21	42.212	21.68087	0.462448	0.034166	0.01742	0.003954
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
사파동	18.07.27	42.4003	21.6046	0.469114	0.033675	0.018187	0.004035
용지동	18.10.05	42.46621	21.70327	0.461456	0.033631	0.017606	0.0039
회원동	18.12.31	43.08201	21.92245	0.47348	0.033573	0.018092	0.003948

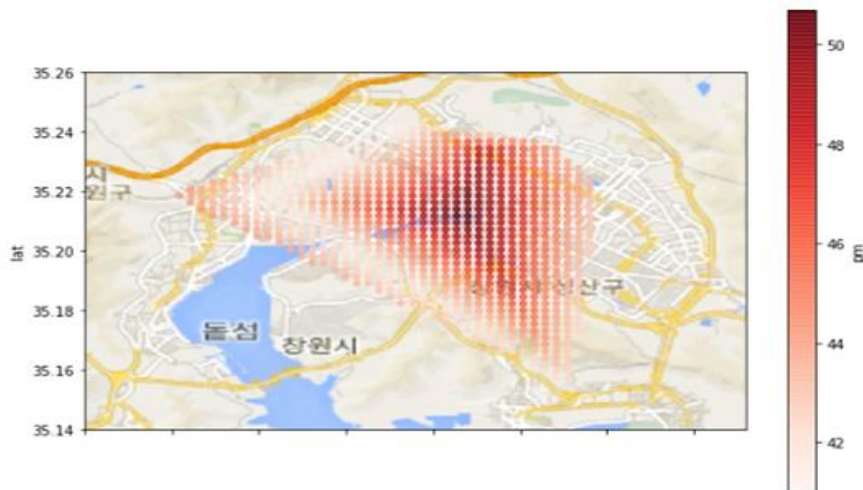
[365일 × 8개 관측소 × 6개 대기오염물질 → n=15,360]

공간 보간법 (GIS interpolation)

- IDW 보간법 (Inverse Distance Weighted, 역거리 가중법)
 - 거리가 가까울수록 높은 가중치 적용

$$\hat{v} = \frac{\sum_{i=1}^n \frac{1}{d_i} v_i}{\sum_{i=1}^n \frac{1}{d_i}}$$

- 8개 관측소에 대해 IDW 보간법 진행
 - 총 806개 지점의 대기오염물질정보 보간
 - 정류장별 위치정보와 보간된 지점 간의 유클리드 거리를 계산
 - 가장 가까운 거리의 대기오염정보 적용



[806개 지점에 대한 PM10 시각화]

정류장명	PM10	PM2.5	...	NO ₂	SO ₂
힐스테이트 4차	43.72718	19.35158	...	0.019612	0.00403
효성중공업	49.03538	23.56986	...	0.02331	0.004885
화천리	44.06726	18.83247	...	0.020332	0.004277
⋮	⋮	⋮	⋮	⋮	⋮
LG산전	48.94694	23.63359	...	0.022772	0.004818
KT동마산지점	43.23348	19.49838	...	0.018965	0.003828

[정류장별 대기오염정보 보간법 적용]

버스정류장과 버스노선

- 창원시 버스정보 시스템
 - 정류소별 버스번호 정보 : $n = 2564$
 - 버스번호별 첫차 막차 시간, 배차간격 : $n = 173 \times 3$
- 버스정류장별 이용인원
 - 버스정류장별 승하차 인원(2018.01~2018.12) : $n = 2075$
 - 버스정류장명이 동일한 경우, 같은 위치의 상/하행 정류장이라고 판단하여 총 이용인원으로 합침 : $2075 - 865 \rightarrow n = 1210$
 - 버스정류장 고유번호와 버스정류장 명이 일치하지 않아, 존재하지 않는 버스정류장과 이용인원이 없는 정류장 제거 : $1210 - 100 \rightarrow n = 1110$
- 창원시 버스정류장 위치정보
 - 창원시 버스정류장 위치정보 : $n = 2398$
 - 이용인원을 알 수 있는 버스정류장만 사용할 수 있기 때문에 버스정류장별 이용인원 데이터의 정류장명을 기준으로 합침 : $2398 - 1288 \rightarrow n = 1110$

[창원 버스정보 시스템]

버스번호	첫차 막차	평균 배차간격
1	평일05:40 ~ 23:00	평일 10~29분
2	평일06:35 ~ 21:50	평일115~145분
⋮	⋮	⋮
1000	평일06:00 ~ 22:00	평일 60~130분

[버스정류장별 이용인원]

정류장명	Mon	Tue	...	Sun
감계입구	339	336	...	166
노인병원	97	114	...	46
⋮	⋮	⋮	⋮	⋮
회산교	1732	1685	...	875

[창원시 버스정류장 위치정보]

정류장명	위도	경도
3.15아트센터	128.57874	35.225762
LG산전	128.67304	35.202916
⋮	⋮	⋮
STX아파트	128.7787	35.120663

버스 노선과 버스 정류장 데이터

파생변수 생성(정류장별 혼잡도)

- 정류장별 혼잡도는 경유하는 버스 수(count)와 이용인원 수(Week)로 생성
- 경유하는 버스 수(count) = 버스번호별 하루 경유횟수 + 정류장별 버스번호
- 버스번호별 하루 경유횟수 = (막차시간 - 첫차시간) / (평균 배차간격)

버스번호	첫차막차	평균배차간격	경유횟수
1	평일 05:40 ~ 23:00	평일 10~29분	52
2	평일 06:35 ~ 21:50	평일 115분~145분	7
⋮	⋮	⋮	⋮
1000	평일 06:00 ~ 22:00	평일 60~130분	10

[버스번호별 하루 경유횟수]

+

정류장명	버스번호
3.15아트센타	29 46 49 53 80 100 105 116 250 259
LG산전	150 504
⋮	⋮
STX아파트	305 305-1 305-2 315 315-1 315-2 757 860

[정류장별 버스번호]

→

정류장명	count
3.15아트센타	283
LG산전	38
⋮	⋮
STX아파트	147

[경유하는 버스 수(count) 생성]

- 버스정류장별 이용인원 수(Week) = 일주일 평균 이용인원 합

정류장명	Mon	Tue	...	Sun	week
감계입구	339	336	...	166	2167
노인병원	97	114	...	46	625
⋮	⋮	⋮	⋮	⋮	⋮
회산교	1732	1685	...	875	10915

[승하차 인원 수(Week)]

Min-Max Scaling

- 오버플로우(overflow)나 언더플로우(underflow)를 방지 위해 최종 데이터 **정규화** 실행
- **Min-Max Scaling**를 사용해 자료의 최소값이 0, 최대값이 1이 되도록 변환



$$x = \frac{x - x_{min}}{x_{max} - x_{min}}$$

[Min-Max Scaling 개념과 공식]

정류장명	count	Week	...	PM10	PM2.5
감계입구	52	2167	...	43.72718	23.56986
노인병원	7	625	...	43.23348	19.35158
⋮	⋮	⋮	⋮	⋮	⋮
LG산전	10	10915	...	49.03538	19.49838

[최종 데이터]



정류장명	count	Week	...	PM10	PM2.5
감계입구	0.04802	0.03245	...	0.345223	0.094832
노인병원	0.03294	0.05432	...	0.0234155	0.042532
⋮	⋮	⋮	⋮	⋮	⋮
LG산전	0.02115	0.02351	...	0.432512	0.043542

[Min-Max Scaling 적용 데이터]

유의미한 변수 선정

• 대기오염물질 데이터

PM10, PM2.5, CO, O₃, NO₂, SO

• 버스 정류장별 혼잡도 데이터

하루에 지나가는 버스대수 (Count)
일주일 평균 승하차 인원수 (Week)

VIF계수 확인을 통한 유의미한 변수 선정

종속변수 생성

군집분석 (K-means)

가중치 부여 방법

상관분석

회귀분석

요인분석

전기버스 도입이 시급한
버스정류장 Top10 선정

전기버스 도입이 시급한
버스노선 Top10 선정

다중공선성 존재 여부 확인

VIF (Variation Inflation Factor)

- VIF 값이 큰 변수를 제거 후,
남아있는 변수 각각의 VIF 변화를 확인

$$VIF_i = \frac{1}{1 - R_i^2}$$

features	VIF Factor
Week	2.05
count	2.97
pm10	19.72
pm2.5	9.84
CO	65.00
O3	22.35
NO2	22.89
SO2	28.11

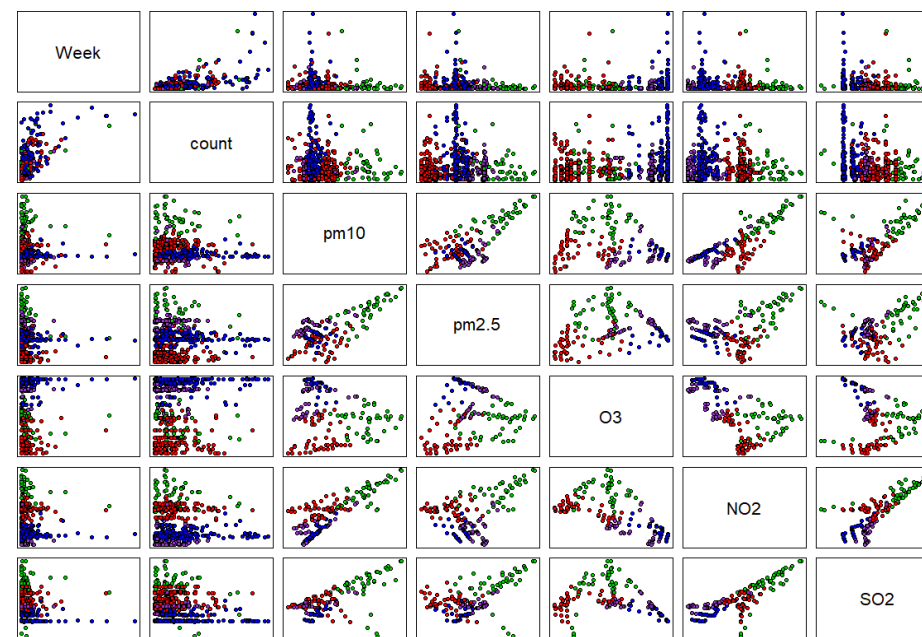
CO제거
→

features	VIF Factor
Week	2.05
count	2.97
pm10	18.51
pm2.5	8.36
O3	6.02
NO2	19.25
SO2	13.94

- 대기오염물질데이터(PM10, PM2.5, CO, O3, NO2, SO2)와 버스정류장별 데이터(Week, count)에서 **종속변수 생성**
- 변수 간 독립성을 확인하기 위해 다중공선성 존재 여부 확인이 필요
- 변수들의 **분산팽창요인(VIF)계수** 확인
- VIF계수가 **높은 변수(CO)**를 제거

■ 군집분석(K-means Clustering)

- 선정된 변수를 통해 군집분석을 적용
- 최적의 K값을 찾기 위해 실루엣 기법 사용
 - K값을 2부터 20까지 반복
 - **K=4**일 때, 실루엣 점수가 가장 높음을 확인
- 변수 간 군집이 잘 형성되지 못함
- 결과적으로 종속변수 생성이 어려움



[K-means Clustering (K=4)]

가중치 부여방법 선정

상관분석

- 종속변수와 독립변수들 간의 상관분석을 실시하여 상관계수 값을 가중치로 설정

회귀분석

- 가중치 부여 대상인 지표를 독립변수로 지정, 해당지표의 상위 항목을 종속변수로 지정
- 회귀분석을 실시하여 추정되는 회귀계수 값을 가중치로 설정

→ 종속변수의 값을 사전에 알고 있어야 함

요인분석

- 독립변수들 간의 요인분석을 실시하여 요인점수계수를 가중치로 설정

→ 종속변수가 필요 없음

가중치 부여방법 선정

상관분석

- 종속변수와 독립변수들 간의 상관분석을 실시하여 상관계수 값을 가중치로 설정

회귀분석

- 가중치 부여 대상인 지표를 독립변수로 지정, 해당지표의 상위 항목을 종속변수로 지정
- 회귀분석을 실시하여 추정되는 회귀계수 값을 가중치로 설정

→ 종속변수의 값을 사전에 알고 있어야 함 ❌

요인분석

- 독립변수들 간의 요인분석을 실시하여 요인점수계수를 가중치로 설정

→ 종속변수가 필요 없음 ○

요인분석

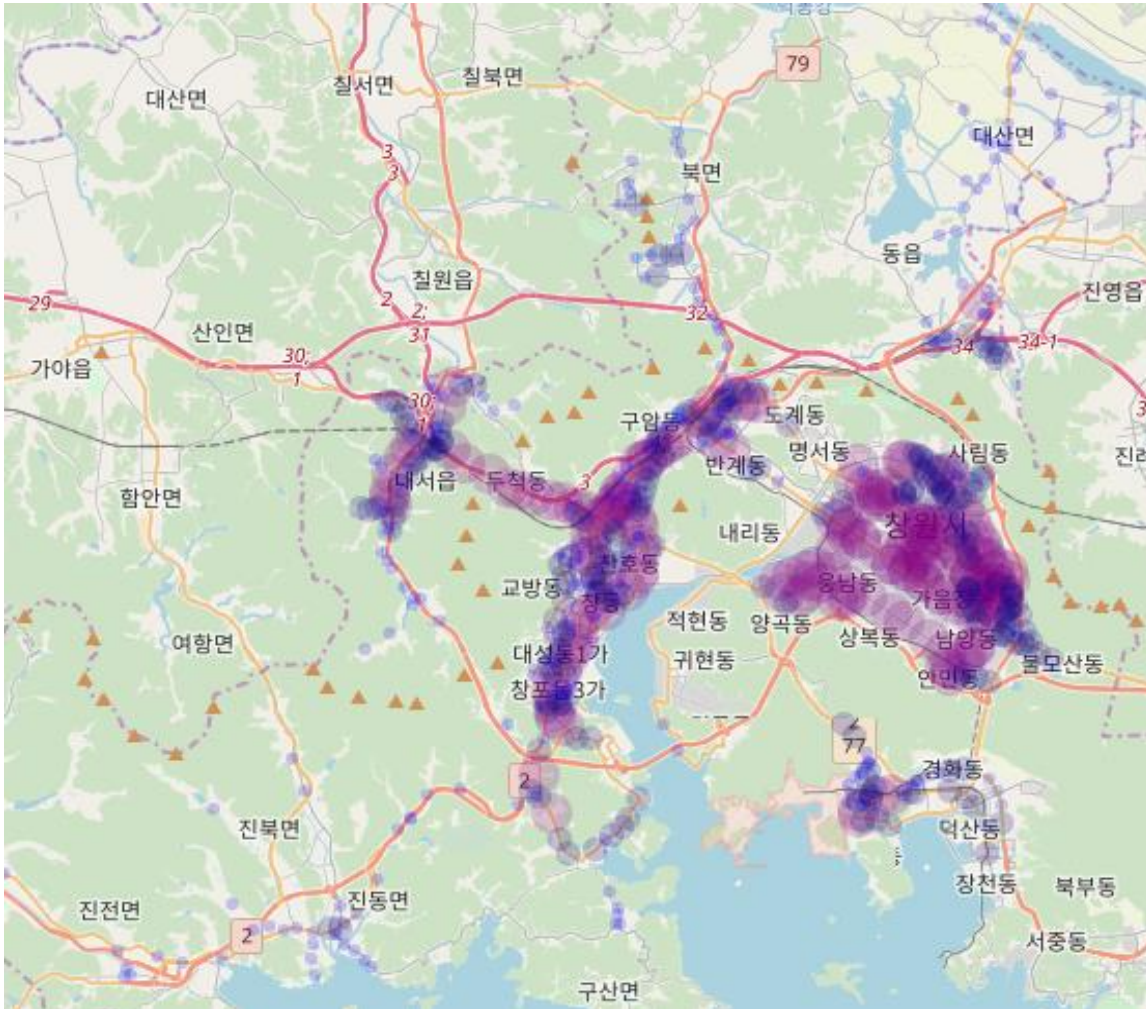
- 선정된 변수들의 상관관계를 분석하여 공통요인으로 묶어 설명하는 다변량 분석방법
 - 상관분석과 회귀분석과 달리, 종속변수가 필요 없음 → 새로운 척도의 개발이 목적인 **적합한 가중치 부여방법**
 - 선정된 변수에 가중치를 부여하여 **요인점수계수** 확인
 - Factor를 3개로 지정하였을 때, 가장 적절하게 Group이 형성됨
 - 정류장별 점수(종속변수) = ***Factor1* + *Factor2* + *Factor3***

features	Factor1	Factor2	Factor3
Week	0.003	-0.024	0.7
count	0.000	0.022	0.891
PM10	0.990	0.102	0.068
PM2.5	0.000	0.698	-0.032
O3	-0.457	0.838	0.051
NO2	0.801	-0.573	0.031
SO2	0.672	-0.358	-0.086

Group 1 PM10, SO2, NO2

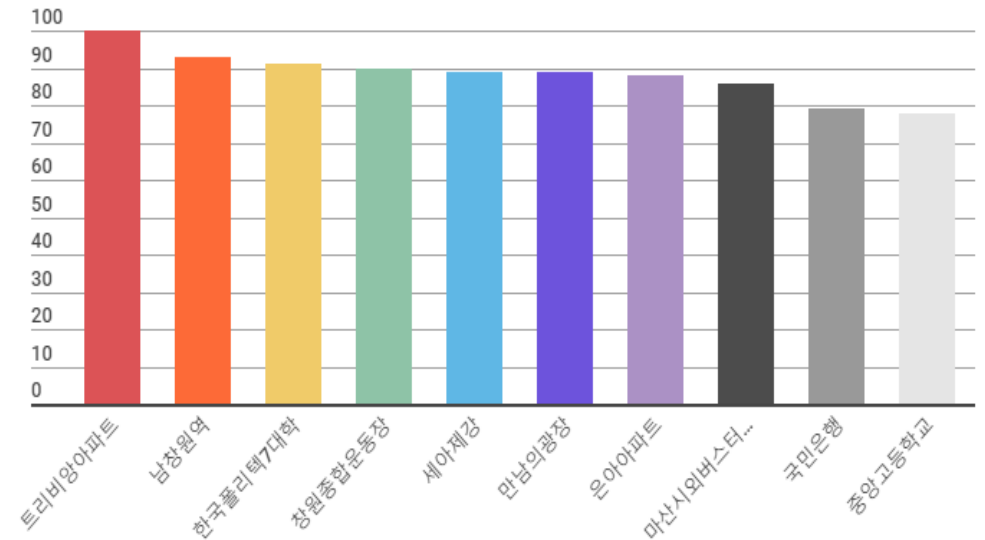
Group 2 PM2.5, O3

Group 3 Week, count

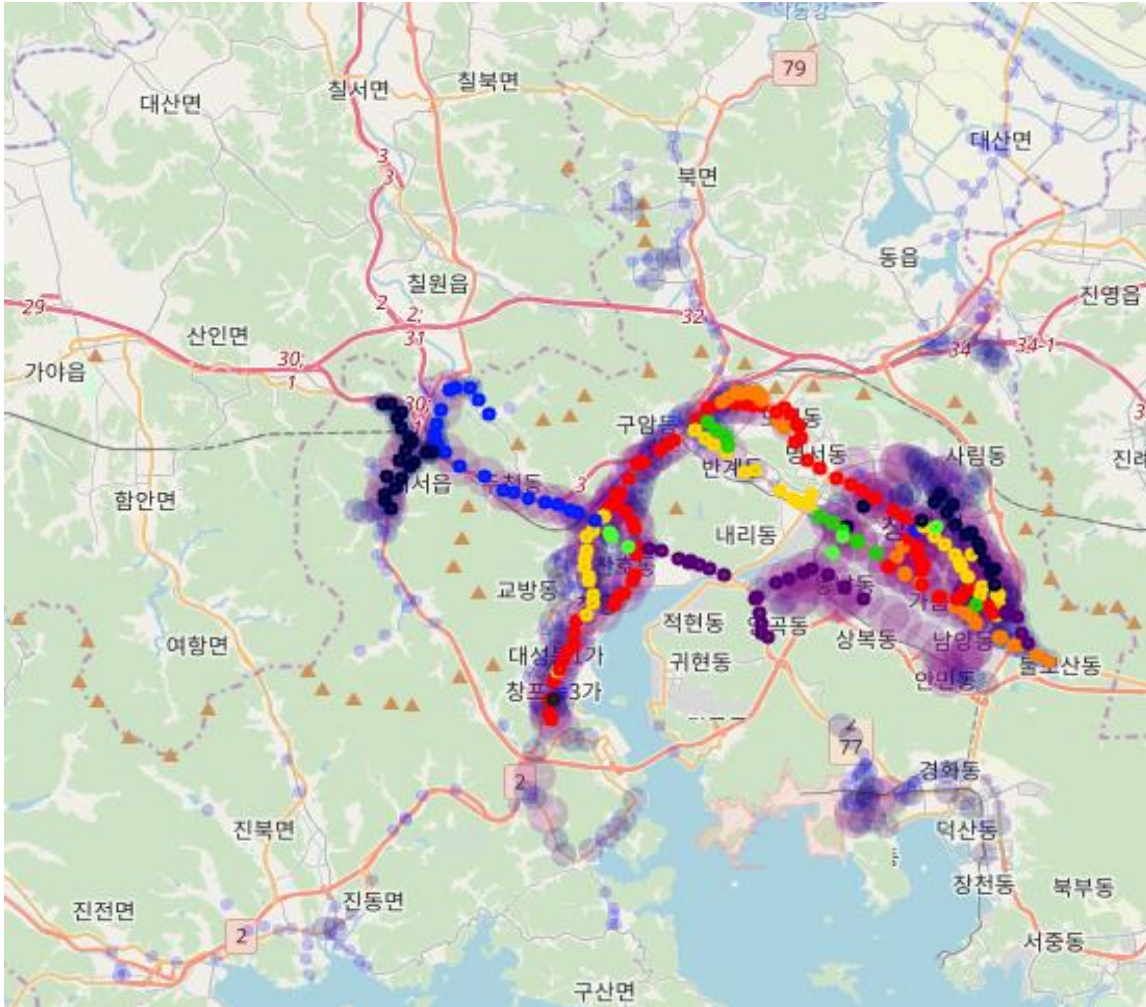


■ 전기버스를 우선적으로 도입해야 하는 정류장

- 대기오염물질 농도와 버스정류장별 혼잡도가 **높을수록**
→ 정류장별 점수(중속변수 값)가 높음
- 정류장 점수에 대한 지도 시각화
→ 전기버스를 우선적으로 도입해야 하는 정류장 Top10 도출

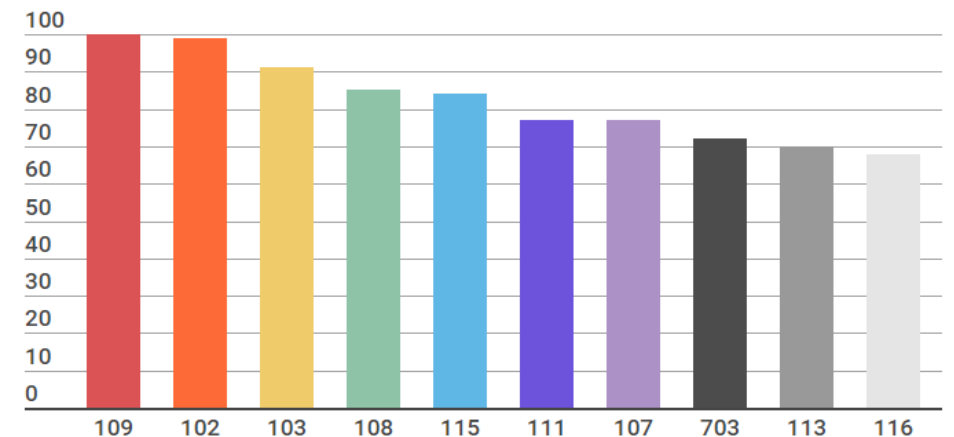


[전기버스 도입 정류장 Top10 도출]



■ 전기버스를 우선적으로 도입해야 하는 노선

- 본 분석의 목적은 전기버스 도입이 시급한 노선을 선정하는 것
- 정류장별 점수를 활용하여 버스노선 점수 생성
→ 전기버스를 우선적으로 도입해야 하는 노선 Top10 도출
- 대기오염물질 농도가 높은 운남동과
버스정류장 혼잡도가 높은 회원동을 아우르는 빨간색 노선
→ 109번 버스



[전기버스 도입 노선Top10 도출]

기대
효과

기대 효과

“ 전기버스로 기존 대형 버스에서 배출되는 대기오염물질 저감효과 ”

한정된 예산으로 투자대비
높은 효율 기대

기존 버스노선을 활용하여
시민들에게 친숙

전 기 버 스
노선추천

장기적인 전기버스 확대의
근거와 지표로 활용 가능

버스 노선 대체방안의
정책추진 분석모델로서 작용

전기에너지 신산업 창출

- ✓ 전기에너지 산업육성을 뒷받침할 맞춤형 제도도입
: 배터리, 충전기를 활용한 부가가치 신사업 창출
- ✓ 기존 엔진기반 대비에 취약한 전기버스의 임대·수리 등
전후방 연관 산업의 보강에 기여

충전소 설치와 연계하여 인프라 구축

- ✓ 전기버스노선에 적합한 충전소 입지 선정



- 창원국가산업단지, 창원시 기업경제포털
- 허성무 창원시장, '미세먼지 저감 6+4 대책' 발표, 중앙일보, 2018.08.16
- 창원시 '미세먼지 저감 6+4대책' 추진 효과, 경남신문, 2018.12.06
- 산업도시 창원의 초미세먼지(PM2.5)의 중금속 조성 and 위해성 평가 연구, 이창흡, 2015
- 올해 창원 시내버스 중 전기버스 비중 10% 넘길 듯, 매일경제, 2019.04.27