「제5회 대구 빅데이터 분석 경진대회」 분석 결과 보고서

-

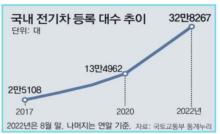
성명(팀명)	남지원
분석과제명	전기차 충전소 최적입지 선정

※ 20장 내외 자유 형식으로 작성, 목차는 필요시 변경 가능

I. 분석개요

□ 분석목적





환경 문제의 중요성이 강조되면서 전기차와 수소차 같은 대체 에너지 자동차의 등장과 관련 정책이 진행되고 있습니다. 전기차는 기존 화석 연료를 사용하는 자동차보다 환경적으로 우수하다고 인정받고 있지만, 전기차 이용을 위해서는 충전 인프라의 확충과 효율 개선이 필수적임.

따라서, 전기차의 수요를 예측하고 이를 바탕으로 전기차 충전소의 위치 분석을 진행하려고 합니다. 전기차 수요를 예측하기 위해서는 관련한 다양한 변수를 고려해야 하며, 이를 통해 예상 수요를 도출할 수 있습니다. 그리고 예측된 수요를 바탕으로 어디에 전기차 충전소를 설치해야 하는지 분석할 것임. 이를 통해 전기차 수요예측과 충전소 위치 분석을 통합적으로 고려하여, 전기차 확산을 지원하는 정책과 충전 인프라 구축에 관한 결정을 돕고자 합니다. 이는 미래의 환경과 교통 시스템에 대한 지속 가능한 발전을 위한 중요하기에 이 주제를 선정하였음

□ 분석요약

- 활용데이터 : 대구시 전기차 충전 이력 데이터 2020~2022,
- o 분석도구 : R,python,QGIS
- o 분석기법 : Random Forest, 역거리 가중법(IDW)
- ㅇ 분석결과
 - Random Forest 분석을 통해 변수 중요도를 구하였으며 이에 따라 각 가 중치를 부여하였고, 이를 통해 격자별 전기차 충전소 입지 점수를 부여하였음
 - 역거리 가중법(IDW)으로 국지적인 값에 대해 불연속데이터에 대한 보간을 함
 - 전기차 충전소 입지 점수를 통해 가장 충전수요 대비 점수가 높은 격자를 전기차 충전소 최적입지로 선정함
- 활용방안: 전기차 등록대수가 증가함에 따라 전기차 충전소의 설치의 필요성은 커지므로 추후에 새로운 충전소 입지를 선택할 때 보다 효율적인 장소를 하기위한 의사결정에 사용할 수 있음

□ 독창성 및 차별성

2023년 7월 대구로 편입된 군위군을 포함하여 최적의 입지를 선정함.

Ⅱ. 분석방법

□ 활용데이터

※ 활용 데이터명, 시계열, 사용된 변수 등 기재 ※ 양식은 필요시 변경 가능

데이터명	형식	대상기간	사용변수	출처	<u>ה</u> [u
전기차		2020년1월~2	Csv,파일에	대구공공시설관리	https://dipbigd
충전이력	CSV	2022년12	있는 변수	공단	ata.kr/
데이터		월	전부	01:	ata.ki/
대구교통공사_					
월별승하	CSV	1997~2023		대구교통공사	
<i>차인원</i> 국토통계_인구					
국도공계=한구 정보-생산					
가능 인구					
수(전체)-(
구(전세) (격자)	shp	2023.7월기준		국토정보플랫폼	
250M_ []					
구광역시					
797 					
국토통계_인구					
<i>정보-총</i>					
인구					
수(전체)-(7-	2023 4월		그ㅌ저ㅂ프레프	
격자)	shp	기준		국토정보플랫폼	
250M_ □					
구광역시					
구9개					
국토통계_건축					
물-건축물					
수 합계					
통계-(격	shp	2023 4월		국토정보플랫폼	
<i>\times_\text{T}</i>)	onp	기준		1-0-275	
250M_ □					
구광역시					
<i>구9개</i> 국토통계_공시					
<u> 국도공세-</u> 공시 지가-공시					
지가-중시 지가-(격					
ハバー(ゴ <i>자)</i>	shp	2023 6월기준		국토정보플랫폼	
250M_ []	\mathfrak{S}^{HP}	2020 0 <i>巨八</i> 七		공ᅩᆯᆺద	
797 1					
T3/11					

<i>국토통계_국토</i>			
정책지표-			
병원		2022 7월	
접근성	shp	기준	국토정보플랫폼
79711-25		7 / 12	
0M			
<i>국토통계_국토</i>			
7-8 #-			
응급의료			
시설	shp	2021년 기준	 국토정보플랫폼
<i>작근성</i>	SHP	2021 년 7 / 년	
79711-25			
<i>OM</i> 국토통계_국토			
<i>정책지표-</i> <i>소방서</i>			
	shp	2021년 기준	국토정보플랫폼
접근성			
79711-25			
0M			
<u>マモ통계_マ토</u>			
정책지표-			
주차장	shp	2021년 기준	국토정보플랫폼
접근성	•		
<i>₹97∦-25</i>			
OM .			
<i>국토통계_국토</i>			
정책지표-			
공연문화			
시설	shp	2021년 기준	국토정보플랫폼
접근성			
7 97#-25			
ОМ			
<i>국토통계_국토</i>			
정책지표-			
공공체육			
시설	shp	2021년 기준	국토정보플랫폼
접근성			
79711-25			
OM			
대구시 전기차			FII 그 고고 사사기기
충전소	CSV	2020~2022	대구공공시설관리
충전 이력			공단
	미기기타	lulol 호키 i	확용하 민가데이터 있을 시 반드시 기재

[※] 제공한 민간데이터 외 추가 활용한 민간데이터 있을 시 반드시 기재

□ 분석방법

※ 분석 결과를 도출하기 위한 알고리즘 및 방법론 (자료분석의 차별성)

- o 분석도구(툴) : R , Python
- ㅇ 데이터 전처리 내용 및 검증 여부

*전기차 충전 이력 데이터

-결측값 제거 : 전기차 충전 이력 데이터에서 NA로 표시된 데이터나 - , 0 등으로 표시된 관측값들을 확인할 수 있었고 확인 결과 평균값 대체등이 불가능하고 관측 치의 개수가 적으므로 제거하였음

-단위변환(, - > .) : 위탁 운영중인 충전소의 데이터에서 충전량 값(이하 kwh) 숫자사이 소수점을 소수점을 나타내는 .대신 ,으로 나타나 데이터상 1000이상의 값을 가지는 걸로 나타나였고 충전량이 100kwh이상 되는다는 것은 현실적으로 불가능하기 때문에 충전시간과 완속/급속 여부를 구분하여서 값이 유사하고 kwh 값이 정상적으로 입력된 데이터를 참고하여 값을 10000kwh이상의 값은 1000으로 나누는 등적당한 값을 나누어서 정상화 시켰음

-주소 처리 :



1.충전소명에서 ()안에 들어가있는 글자 때문에 같은 위치더라도 다른 충전소로 인식하여 ()의 문자를 제거하고 위탁과 직영 데이터에서 주소(이하 add)의 형식이 달랐으므로

통일하였고, 또한 gis활용을 위해 위도와 경도를 찾아 새롭게 입력하였음 -위탁 직영 변수 통일

충전이력 데이터에는 위탁과 직영 데이터가 따로 존재하였는데 입력된 데이터의 변수명이 서로 달랐으나 계산을 통해서 도출이 가능하였기 때문에 직영과 위탁의 변수를 만들고,제거하여 통일하였음.

데이터 결합 : 데이터셋을 살펴보던 중 , 위탁에서 직영으로 바뀐 관측치가 존재하여 위탁 직영데이터를 결합하고 3개년치 데이터를 결합하였음

*격자 데이터

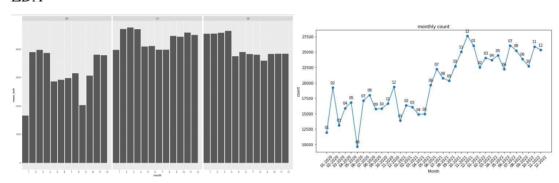
1.결측값 제거: NULL값을 0으로 변환하고 전체 격자데이터에서 군위군을 포함한

대구광역시의 데이터만을 추출하였음

2.데이터 결합 : 분석에 사용할 변수(총인구, 생산인구 , 공시지가 , 건축물 수 ,주차 장접근성 , 소방서 접근성 등등에 대한 250x250격자 데이터를 결합하고 전기차 충전 이력 데이터와 병합하였음

3.데이터 변환 폴리곤 내 속하는 충전소의 개수를 통해 두 개 이상의 충전소를 가 진 격자에 대해서 충전량의 합을 계산하고 총인원 생산인구 공시지가 데이터를 정 규화 하였음

EDA



데이터를 시각화하던 중 시간대에 따른 충전수요는 새벽에 낮고 또한 오후 12시이후부터 퇴근시간(오후6시)까지 낮고 퇴근시간대 이후 점차 올라가는 직관적인 추이를 보이지만 월별로 통계를 내보았을때 봄,여름보단 가을,겨울에 평균적으로 충전수요가 늘어난 것을 보아 기온이 충전수요에 영향을 준다는 생각을 하게 되었음

ㅇ 분석기법 및 분석절차

전처리한 격자 데이터를 한 데이터로 병합하였고, 이를 통해 격자별 속성값으로 모든 데이터를 입력하였음.

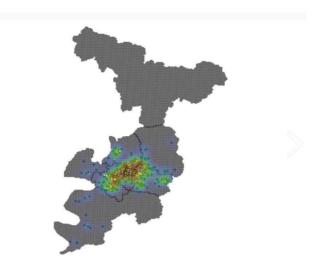
또한 역거리 가중법을 사용하여 국지적인 값을 가지는 충전량에 대해 보간을 실시 하여 충전소를 포함한 격자의 인근 지역에도 충전량을 가지는 연속적인 데이터로 변환하였음.

또한 RANDOM FOREST 기법을 통해 주요 변수들의 중요도를 산출하였고, 이를 기반으로 가중치를 선정하여 충전소 입지 선정 점수를 산출하였음

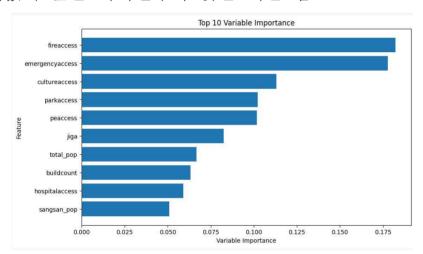
ㅇ 시각화 방법 등

R, PYTHON, QGIS를 통해 가공한 SHP파일에 대해 QGIS를 활용하여 시각화하였음.

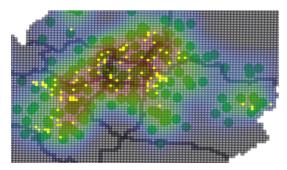
Ⅲ. 분석결과



전기차 충전소 충전 이력 데이터를 통해 구한 충전소별 충전전력 기댓값을 역거리 가중법을 사용하여 다음과 같이 지도에 나타냈음. 충전전력 기댓값이 높은 지역일수록 붉은 색을 띔



RANDOM FOREST 기법을 통해 기존 전기차 충전소의 충전전력 기댓값에 대한 변수별 중요도를 구하였고 다음과 같이 시각화 하였다. 위 결과값을 통해 가중치를 사용하여 충전소 입지 선정 점수를 산정하였음.



충전소 입지 선정 점수가 높은 상위 100개의 격자를 위와같이 시각화 하였음



구 별로 점수가 높은 격자를 선정하였고, 다음과 같이 행정구역 별 1개 씩 10개의 후보군을 선정하였음

최종 선택 입지

취저그어	OLE	거ㄷ
행정구역	위도	경도
군위군	36.24365	128.4808
<u>남구</u>	35.85067	128.5926
<u>달서구</u>	35.82169	128.5446
달성군	35.79719	128.4647
달성군 다사	35.86064	128.4652
동구 북구	35.87787	128.6461
북구	35.91166	128.5457
서구	35.88566	128.5621
수성구	35.85731	128.6186
중구	35.8707	128.5953

구별로 분석 스코어값을 기준으로 하여 높은 스코어를 가진 위치로 선 정하였음

Ⅳ. 활용방안

전기차 수요 예측 이를 통해 기업들은 시장 수요에 맞춘 제품 개발 및 마케팅 전략을 수립할 수 있습니다. 충전소의 수, 용량, 충전 방식 등을 고려하여 충전 인프라를 효율적으로 구축하고 전기차 사용자들의 편의성을 높일 수 있음. 정확한 수요 예측과 적절한 입지 선정은 전기차 시장의 성장과 확산을 촉진하고, 전기차의 사용성과 편의성을 향상시킬 수 있고 또한 추가적인 충전소 설치를 통하여 피크타임때 전력이 엄청나게 사용되는 것을 분산시킬 수 있음.

Ⅴ. 기타

- 1. 홍준기, 김순태, 김정아 (2023) 전기자동차 충전소 수요 예측 데이터 전처리 기법 및 서비스 운영 아키텍처
- 2. 최황규*·김광호 † 전기차 충전설비에 있어서 도로교통량 데이터를 기반으로 한 충전수요예측 모델 연구
- 3. 권상협, 손다혁, 전승찬, 박혜리, 배성우부 Support Vector Regression을 이용한 전기자동차 충전전력 예측 및 분석
- 4. 김여진*·허 진† 전기차 충전수요 분산을 위한 신규 전기차 충전소 위치 선정에 관한 연구