

더 나은 삶 더 나은 내일
함께 만드는 LH

COMPAS

데이터 중심의 도시문제 해결
The 혁신적인 문제해결

跳躍

도약

데이터 기반의 도시문제 해결 플랫폼

COMPAS



CONTENTS

I _ 추진배경

II _ COMPAS

III _ 향후계획

IV _ 협업사례



CHAPTER

I

추진배경

다양한 도시문제 발생

도시 문제에 대한 다양한 분석방법의 접근으로 해결책 제시

급격한 도시화로 도시는 다양한 문제에 직면

교통문제

환경문제

취업문제

주택문제

치안문제

일자리
복지
예방
고용
저소득

범죄
복지
연금
고용
저소득

금수저
비부적차
흙수저
기후변화
공포
전기자동차
위협
스마트홈
지구온난화
머신러닝
빅데이터
불안
배터리
실시간
불평등
노인빈곤
소득불평등
중산층
한염병
김염병
은퇴
출산감소
인구절벽
소득절벽
양극화
정년연장
독거노인
고령화
육아휴직
경력단절
맞벌이
제조업
휴머노이드

기존 인프라 중심에서 데이터중심의 문제해결 방식이 대두

도시에서 발생하는 각종 데이터를 수집, 활용함으로써 보다 효율적으로 문제에 대응 가능



도시에서 생산되는 다양한 데이터를 활용한 도시문제 해결

기존도시 대응



교통혼잡

도로확장 or 신규도로 건설



주차문제

신규 주차장 건설



방범혼잡

경찰 인력 전 지역 투입



상하수도

누수 지점 정보 취득 불가

스마트시티 대응

- 혼잡도로 정보 실시간 제공을 통해 우회 유도
- 실시간 교통량에 따른 교통신호 제어

- 빈 주차공간 정보 실시간 제공하여 주차 유도
- 카셰어링 등의 서비스 활용 도심진입 최소화

- CCTV 복합화로 경찰인력의 즉각적 투입
- 스마트 범죄 관련 앱 활용을 통한 도움요청

- 누수지점 센서 감지를 통한 즉각적 조치
- 노후도 추정에 따른 누수 가능지역 추정



적용 효과 사례

영국 시스템 도입 후
소요시간 25%, 교통사고 50%,
대기오염 10% 감소

향후 전 세계 410억 달러 이상 수익
스마트주차에서 발생 예상

국내 시스템 도입 후
20% 범죄 발생률 감소

카타르 도하/브라질 상파울로
40~50% 누수예방 효과 발생

지자체 정보화 조직 현황



지자체의 도시문제 해결을 위한 도움이 필요

데이터의 중요성은 높아지고 있으나 데이터 분석을 위한 조직은 부족한 실정

광역 자치단체(광역시·도) : 독립된 정보화 전담 조직 구성

기초 자치단체(시·군·구) : 일부 선도 자치단체를 제외한 대부분 지자체는 인력·예산 등의 한계로 일반부서에서 업무 담당

빅데이터 전담조직 : 서울, 부산 등 9개 광역 자치단체와 성남시, 창원시 등 17개 기초자치단체에서 운영

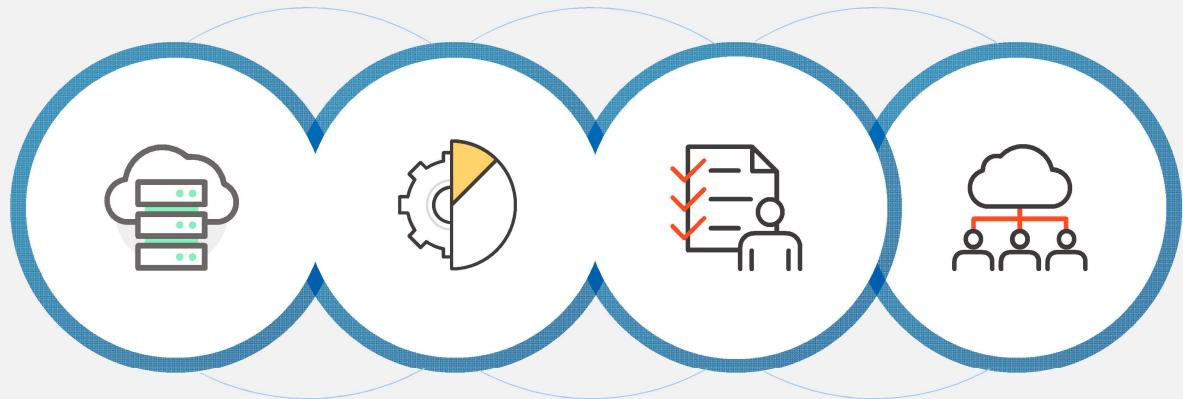
구분	광역시·도	시	군	자치구	합계
정보화 전담	17	51	2	15	85
정보화 부분 담당	-	12	2	18	32
일반부서	-	14	78	36	128
빅데이터 전담조직	9	13	1	3	26

출처 : 한국지역정보개발원(2019) 2018~2019 지역정보화 백서

도시데이터 문제

실무부서 담당자들의 데이터 중요성 인식 부족 및 전문조직 미비로 활용에 어려움

데이터 수집 등 관리체계, 개인정보 보호 등 기술적, 제도적 문제로 데이터 활용 애로



데이터 수집체계 미흡

데이터 가공 미흡

개인정보 이용제한

공공데이터 활용 저하



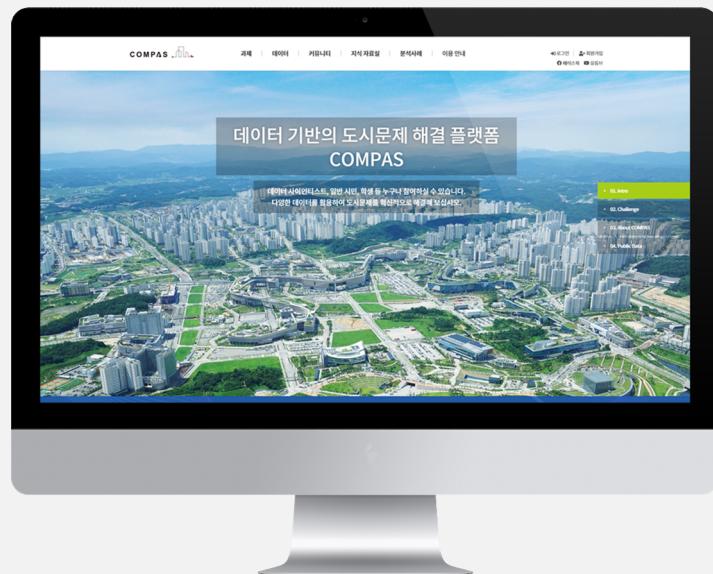
CHAPTER

II

데이터 기반의 도시문제 해결 플랫폼

COMPAS

COMPAS는 지자체의 도시문제를 다양한 전문가들이 참여하여
데이터 분석을 통해 해결방안을 경쟁하는 솔루션 경쟁형 플랫폼



나침반과 같이 도시문제 해결의 방향성을 제시하는 데이터 분석 플랫폼

COMPAS

나침반(Compass)과 같이 도시문제 해결의
방향성을 제시하는 데이터 분석 플랫폼



- C** Citizen, Crowdsourcing
- O** Opportunity, Offer
- M** Management
- P** Problem, Public
- A** Analysis, AI, Advance
- S** Solution, Share, Sustainable

문제를 먼저 정의하고 그 문제를 해결하기 위해
필요한 데이터를 수집하는 문제 해결 플랫폼

COMPAS만의 접근방법은?

COMPAS는
데이터를 쌓아놓고 문제를 찾기보다
도시문제를 해결하기 위해
우선 도시문제를 정의(Question First)하고,
문제 해결에 필요한 데이터 수집

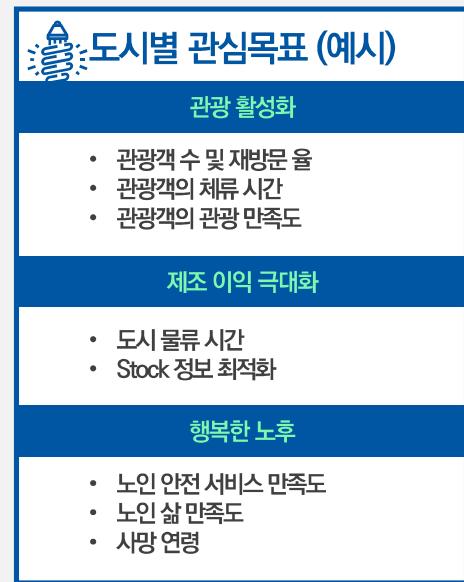
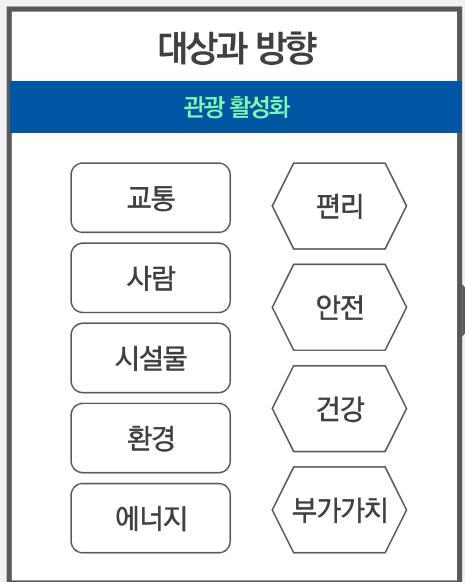
“APPROACH”

데이터를 보면
문제가 보이는가?

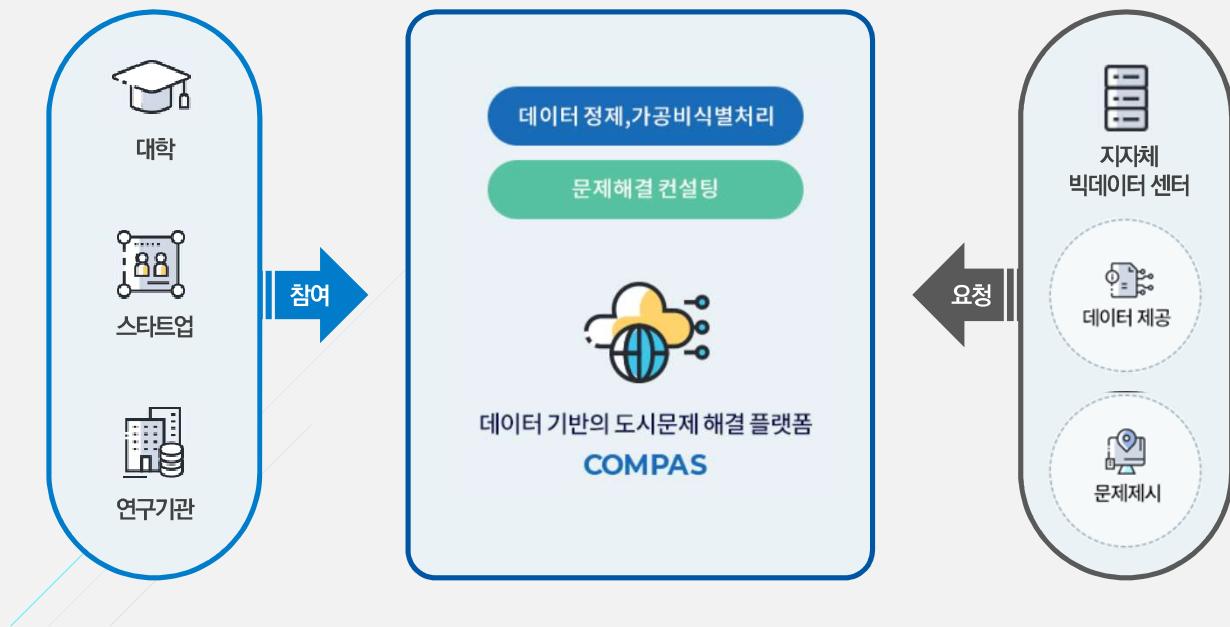
VS

문제를 해결하기
위해서 데이터를
보는가?

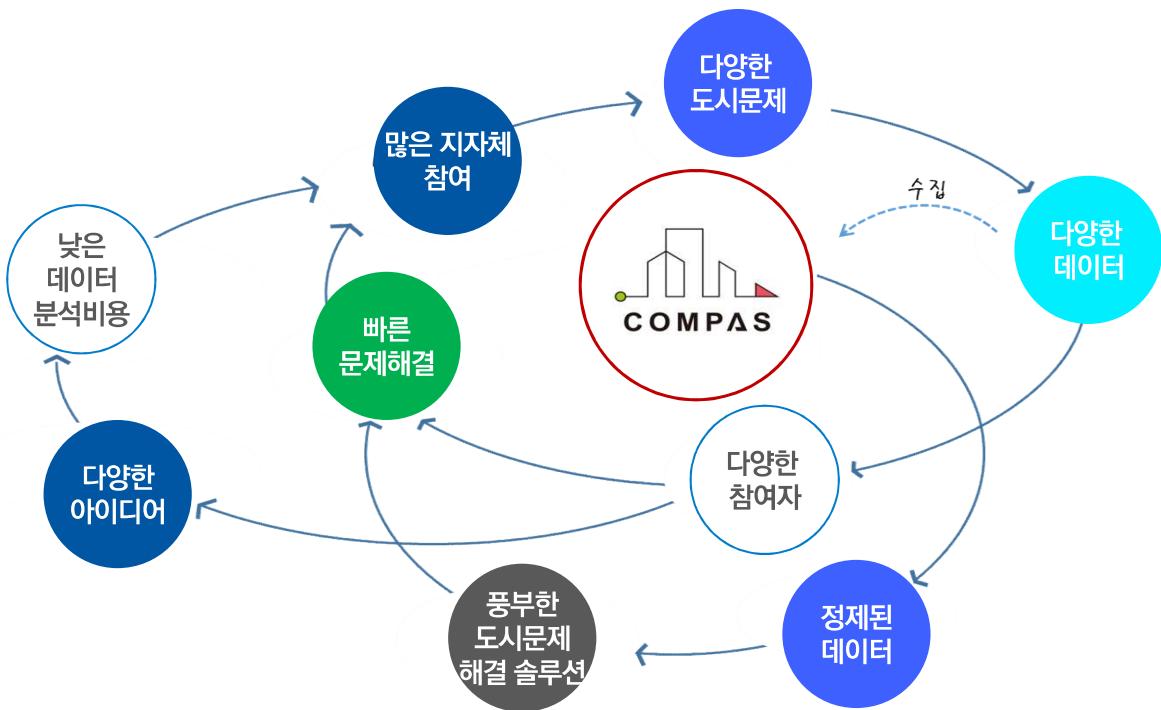
도시마다 관심 목표와 당면한 도시문제가 상이
도시가 필요로 하는 관심 목표를 선정하고 그에 필요한 데이터 수집



- 지자체는 도시문제와 데이터 제공, COMPAS는 도시문제를 구체화하고 필요한 데이터를 수집하여 참가자들에게 제공
- 참가자들은 다양한 도시문제를 탐색하고 관심있는 분야의 도시문제에 참여, 다양한 데이터 분석을 통해 솔루션 도출
- 참가자들은 이 과정에서 실제 도시데이터를 활용할 수 있으며, 지자체는 다양한 전문가들이 분석한 도시문제 해결 솔루션을 활용



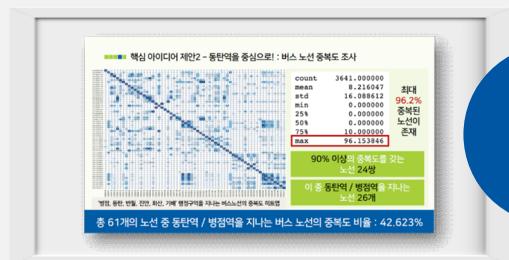
지자체가 참가하면 할수록 다양한 데이터와 솔루션을 발굴할 수 있으며,
솔루션이 증가하면 할수록 문제 해결 시간 및 비용이 감소하는 데이터 선순환 모델



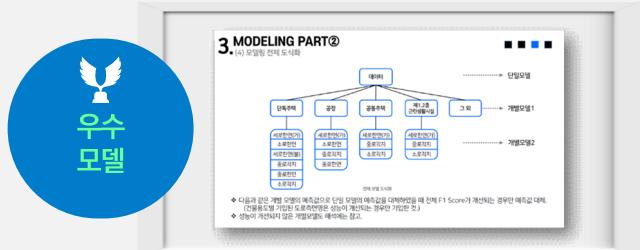
수요자 중심의 문제 발굴 및 집단지성을 이용한 도시문제의 혁신적 해결



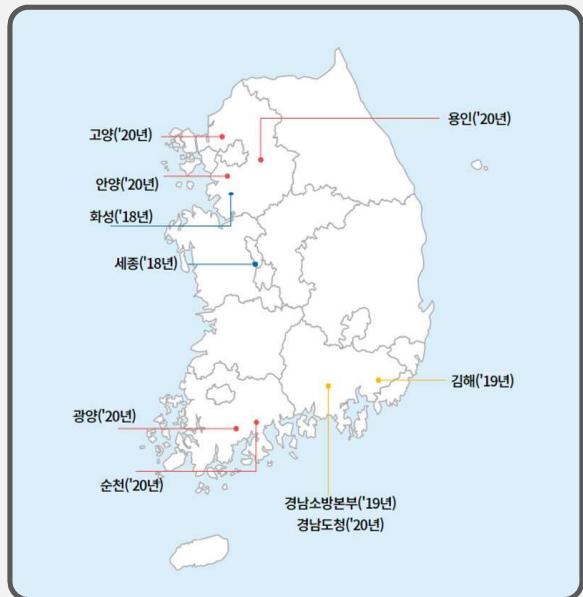
김해시, 화성시 2개 과제에 총 497개 팀 914명이 참여하여 총 145개의 모델이 제출



우수
모델



고양시, 광양시, 세종시, 안양시 등 10개 지자체 과제로 확대 향후 상시적으로 과제를 등록하고 해결할 수 있는 상용 플랫폼으로 운영



지자체	분야	협업 과제
고양시	교통	공공자전거 스테이션 위치 선정 및 배치 모델 개발
용인시	일자리	청년 창업 실태 분석
	관광	체류형 관광객 도출 모델 개발 및 현황 분석
안양시	환경	미세먼지 집중관리구역 위치 선정
	안전	안전 및 환경을 고려한 맞춤형 가로등 조도 분석
	교통	환승 패턴 분석을 통한 환승 노선 최적화
광양시	에너지	전기차 충전소 우선 설치위치 분석



CHAPTER

III

향후 계획

기대효과

- 향후 플랫폼을 상시 오픈하여 도시문제를 해결하고, 우수 인재 채용 기회 제공
- 실제 도시데이터를 이용한 데이터 분석 학습
- 연구기관, 학교 등과 협업하여 데이터 기반의 도시문제 해결과제 발굴

기대효과 #1

도시문제 해결

집단지성 기반으로 도시문제
해결을 위해 모델을 개발하고 경쟁



기대효과 #2

인력양성

실제 도시데이터를 이용한
다양한 데이터 분석모델 학습



기대효과 #3

산학협력

COMPAS를 통해 수집된 아이디어와 데이터를
이용하여 관련 연구 및 협력과제를 수행





데이터 수집, 활용, 확산 체계를 완성하여 데이터로 관리되고 가치가 창출되는 LH만의 스마트시티 플랫폼으로 발전

빅데이터 수집·활용·확산 체계



데이터 수집

지자체가 **데이터를 개방할 수 있는**
환경 조성



데이터 활용

대중의 주도적 참여를 통한
활용체계 다양화



데이터 확산

양질의 데이터 개방을 통한
신 산업 창출

데이터로 관리되고 가치가 창출되는
LH 스마트시티 플랫폼

LH 스마트시티 플랫폼 구축 로드맵

3

데이터 확산

도시 플랫폼('23-)

2

데이터 활용

분석서비스('20-'22)

1

데이터 기반

도시문제해결('19)

도시문제
데이터 분석

솔루션 도출

DATA



리빙맵
크라우드
소싱



디지털
트윈
콤파스



CHAPTER

IV

협업사례

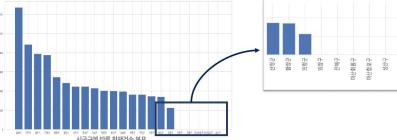
김해시 화재발생 예측

김해시는 공장지대의 화재가 빈번히 발생하고 있어
데이터 분석을 통해 화재위험이 높은 건물을 예측하여 방재활동을 추진하고자 함

1. DATA PREPROCESSING

(3) 창원 통합시 데이터 삭제

다음은 시군구에 따른 화재 발생 분포이다.

창원 통합시인 진해구, 마산회원구, 마산합포구, 성산구는 화재발생이 0건이다.
이 지역들의 경우 화재발생에 대한 정보를 구하지 못했다고 판단하여 해당 Data는 삭제하였다.

한편 김해시의 화재 발생건수가 암도적인 것을 확인할 수 있었다.

4. INTERPRETATION

(4) 단독주택에 대한 해석

다음은 단독주택의 경우 '화재발생'과 양의 관계인 변수들과 음의 관계인 변수들이다.



한편 단독주택의 경우 경우에 대한 고찰이 더욱 필요하다.

단독주택의 경우 경우에 대한 고찰이 더욱 필요하다.

4. INTERPRETATION

2. MODELING PART①

(1) 단일 모델 사용: LGBM

LGBM을 이용하여 변수 중요도를 파악하고자 하였다.

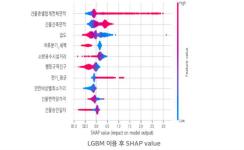
① Feature importance

일반적으로 사용하는 방법인 'node가 많이 나눠온 중요도가 높게 나눠지는 단점 존재. 이것으로 변수 중요도를 파악하는 것은 무리가 있다고 한다.'



② SHAP

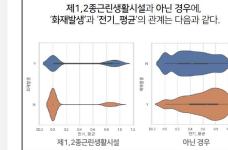
Feature importance의 단점을 보완하고자 SHAP value를 이용하였다. '검증용법개별변경', '검증검증변경' 등을 그 값이 높을수록, '낮을수록 화재발생'에 중요한 영향을 미친다는 것 확인.



3. MODELING PART②

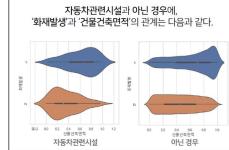
(1) 건물용도에 따른 데이터 분포 담색

제 1.2종근린생활시설



주요 건물용도에 따라 데이터의 분포가 다르다는 것을 가시적으로 확인할 수 있다.

④ 자동차관련시설



4. INTERPRETATION

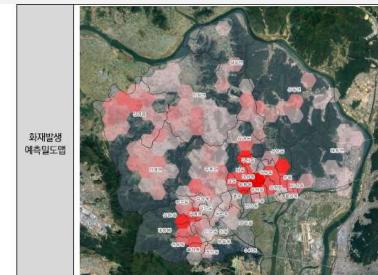
4. INTERPRETATION

(5) 결론

모델에 대한 질문

- 데이터 자체를 최대한 보존하는 모델이다. 차원 축소와 같은 방법을 진행하면 대체로 정보의 손실과 악영향이라는 한계점을 수반하기 때문에, 데이터 구조에 변화를 주는 방법을 사용하지 않았다.
- LGBM 모델을 이용하여 전반적인 모델링을 진행했다. 그 후, '건물용도', '도로측면면' 등 주요 변수에 따라 데이터의 분포, 즉 성질을 다른 기준에 따라 주어져서 설명할 수 있도록 각각의 개별모델을 생성했다. 이는 TreeBoosting 기법의 장점을 살리면서 단점은 보완하는 모델이다.
- EDA 및 상식으로서 여러 모델을 만들기 때문에 모델 구조의 논리가 합리적이어야 쉽게 재현이 가능하다.
- 또한 LGBM 모델은 사용함으로써 구조의 심도와 파라미터로 제한적으로 사용되었다. 파라미터의 경우 세개만 사용하였다. 특히 모델에서는 해당 파라미터를 자동화하지 않았으며, Grid Search Algorithm을 통해 자동화를 한지만 더욱 쉽게 재현이 가능 할 것이다.

5



고양시 공공자전거 스테이션 위치(종료)

고양시는 공공자전거 문제를 개선하고 보다 효율적인 공공자전거 서비스를 제공하기 위하여 데이터 분석을 통한 개선방안 마련

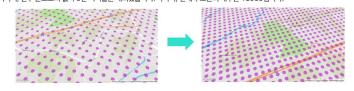
1. 자전거 스테이션 입지 및 이용 현황



3. 학습용 데이터와 예측용 데이터 구축

본적으로 예측용 데이터와 학습용 데이터를 구축합니다. 먼저, 예측용 데이터 구축 과정입니다.

STEP1. 고양시 전역 민족인 50m X 50m 격자의 중심점을 후보군으로 설정.
민족인은 주민번호로 등록된 1000만명 이상의 주민으로서 자전거 이용 패턴과 밀접한 관련이 있는 민족인을 대상으로 한다. 예상되는 자전거 이용자 수로 추정되는 경우는 15969명입니다.



STEP2. 각자리의 후보군에 대해 일별 여객수를 위해, 월(1~12)과, 요일(월~금) 변수를 추가. 대여 사용량 번호 month曜일

또, 차량 회수량 단위 8x8x12x7일짜리 대여량을 계산하는 것입니다.
이후, 각 후보군별로 대여량을 얻은 날, 상관 25% 날 등을 이용합니다.

STEP3. 학습용 데이터와 공통의 변수들을 추구.

예측용 데이터
0 0 1 Friday
1 0 0 1 Monday
2 0 0 1 Saturday
3 0 0 1 Sunday
4 0 0 1 Thursday

22

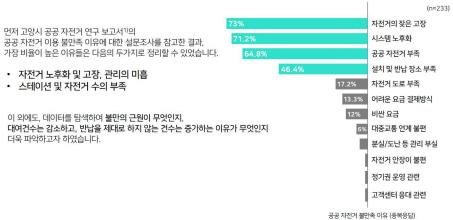
2. 불만 상황 파악

설문조사 참고

먼저 고양시 공공 자전거 이용자(보고서*)의
공공 자전거 이용 만족도에 대한 설문조사를 참고한 결과,
가장 비율이 높은 이용자들은 다음의 두 가지로 정리할 수 있었답니다.

- 자전거 노후화 및 교체, 관리의 미흡
- 스테이션 및 자전거 수의 부족

이 외에도, 대여료를 탐색하며 불편한 균형이 무엇인지,
대여관련 정보는 간소하고, 벤딩이 제대로 하지 않는 간수는 증가하는 이유가 무엇인지
등은 이야기하고 있었답니다.



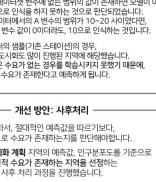
4. 모델 학습

예측값과 학습



5. 최종 결과와 평가

예측값과 실제값 차이



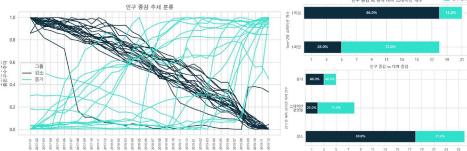
2. 불만 상황 파악

설문조사 참고

증가중인 사이 사용자 수와 함께, 그리고 기본 시스템 개선 개선의 관리를 더욱 비교해보았습니다. 고양시의 현재 출입국수는 증가 추세입니다. 특히 자전거 시스템이 1km당 10대 이상 설치되어 있는 등을 통해 86.7%의 인구가 접속하는 추세였습니다.

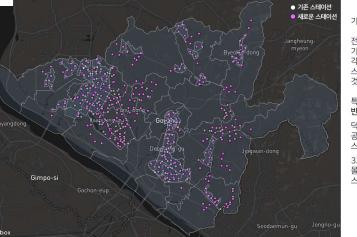
1km당 당일 이용 대여량은 2019년 대비 2020년에는 31%가 증가하는 추세로 재현되었습니다. 또 한 2019년 대여량은 60,678건, 2020년 대여량은 100,327건으로, 기관에 고려되어야 하는 출입국수는 75.5%가 인구 증가 추세에 재현되었습니다.

즉, 고양시에서 많은 도시계획적, 인구통계학적 변화가 일어나고 있었습니다.



5. 최종 결과와 평가

예측값과 실제값 차이



기존 스테이션의 비교집합

전체 스테이션을 평가할 때,

기존 스테이션의 위치와 새롭게 배치된 디자인을 비교해보았습니다. 특히 주변 대중교통망과 같은 환경을 고려해 소리마이너 총합이 배제되어 있는 것을 볼 수 있습니다.

FDA 경쟁력 분석 결과 요소별

판별되었답니다.

당면 고민과 인구가 증가하는 추세 지역들, 특히 경기도 북부(부평, 광교) 및 경기 남부(시흥, 오산) 지역들은 대중교통망과 같은 환경을 고려해 소리마이너 총합이 배제되어 있는 경우, 풍성한 전용 보도로 조성된 경우, 스테이션들이 추가 배치되었습니다.

안양시는 미세먼지의 효율적인 저감대책 수립을 위해 주요 구간에 대한 미세먼지 센서 설치 위치 도출



15 팀 참가

[안양시] 도시데이터 수집을 위한 IoT센서 ...

상금 : 1,000 만원

기간 : 2020-09-15 ~ 2020-11-03

파일 크기 (336.16 MB) API GetCompasData('SBJ_2009_002', '1', '01.안양시_격자별인구현황(전체).geojson')		
데이터 목록	데이터 설명	컬럼 목록
01.안양시_격자별인구현황(전체) 공개 2 col	2020년 4월 기준의 안양시 격자별 인구수(전체) 공간 정보입니다. (100M X 100M)	* gid 격자별 고유 분류코드입니다.
02.안양시_격자별인구현황(고령) 공개 2 col		* val 격자내의 총 인구수입니다.
03.안양시_격자별인구현황(생산가능인구) 공개 2 col		
04.안양시_격자별인구현황(유소년) 공개 2 col		
05.안양시_월별_시간대별_유동인구 비공개 27 col		
06.안양시_학교현황 공개 6 col		
07.안양시_유치원현황 공개 7 col		
08.안양시_복지회관_보건소현황 공개 4 col		

광양시는 전기차 보급 및 충전인프라 확보를 위하여 과학적인 방법을 통해 충전소의 입지를 선정하고자 함



19 팀 참가

[광양시] 전기자동차 충전소 최적입지 선정

상금: 800 만원

기간: 2020-09-10 ~ 2020-10-30

데이터 목록	데이터 설명	컬럼 목록
01.광양시_충전소설치현황_공개_10col	현재 기준의 광양시 전기차 충전소(급속, 완속)에 대한 위치 정보입니다.	▲ 충전소명 전기자동차 충전소 명칭입니다.
02.광양시_주차장_공간정보_공개_8col		▲ 충전소위치 전기자동차 충전소 위치입니다.
03.광양시_자동차등록현황_격자(100X100)_공개_2col		▲ 충전소 운영기관 충전기의 운영기관 명칭입니다.
04.광양시_대중집합시설_생활체육시설_공개_5col		▲ 급속/완속 충전기의 충전방식 구분입니다.
05.광양시_대중집합시설_아영장_공개_5col		▲ 충전기용량 충전기의 용량입니다.
06.광양시_전기차보급현황(연도별,읍면…_공개_3col		* 이용대수 충전기의 이용가능 차량 대수입니다.
07.광양시_행정구역별_인구현황(읍면동리)_공개_6col		▲ 충전기타입 충전기의 타입입니다.
		* 요금정보(원/kw) 충전기의 요금정보입니다 (원/kw)
		* lon 초저스이 거드이 lr



감사합니다

한국토지주택공사