

데이터 중심의 도시문제 해결  
The 혁신적인 문제해결

# 跳躍 도약

데이터 기반의 도시문제 해결 플랫폼

COMPAS



데이터 기반의 도시문제 해결 플랫폼 COMPAS 

# CONTENTS

I \_ 추진배경

II \_ COMPAS

III \_ 향후계획

IV \_ 협업사례

CHAPTER

I

# 추진배경

---



도시 문제에 대한 다양한 분석방법의 접근으로 해결책 제시

## 급격한 도시화로 도시는 다양한 문제에 직면



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
- 카 셰어링 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 수의  
스마트주차에서 발생 예상

국내 시스템 도입 후  
**20%** 범죄 발생을 감소

카타르 도하/브라질 상파울로  
**40~50%** 누수예방 효과 발생

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

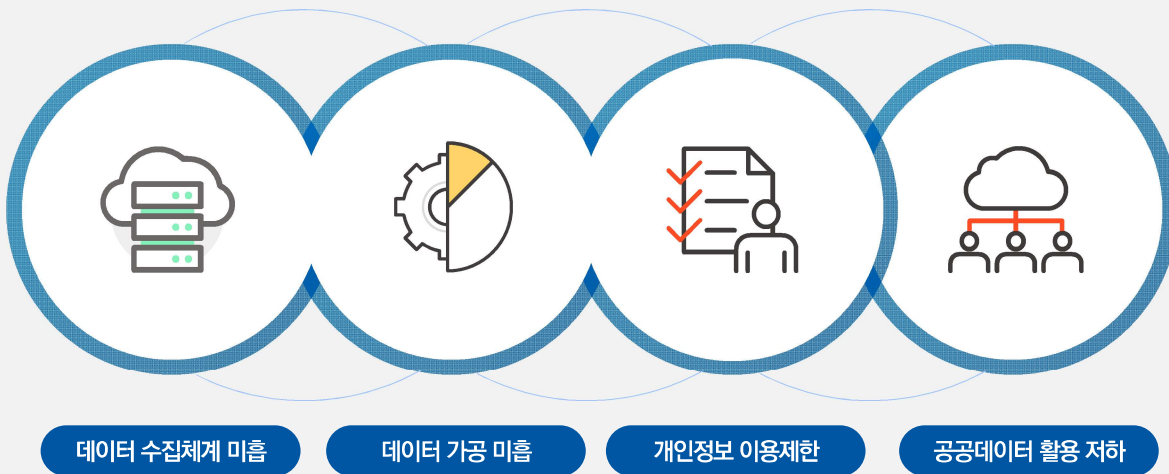
빅데이터 전담조직 : 서울, 부산 등 9개 광역 자치단체와 성남시, 창원시 등 17개 기초자치단체에서 운영

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

실무부서 담당자들의 데이터 중요성 인식 부족 및 전문조직 미비로 활용에 어려움

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



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 찾기 보다

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?

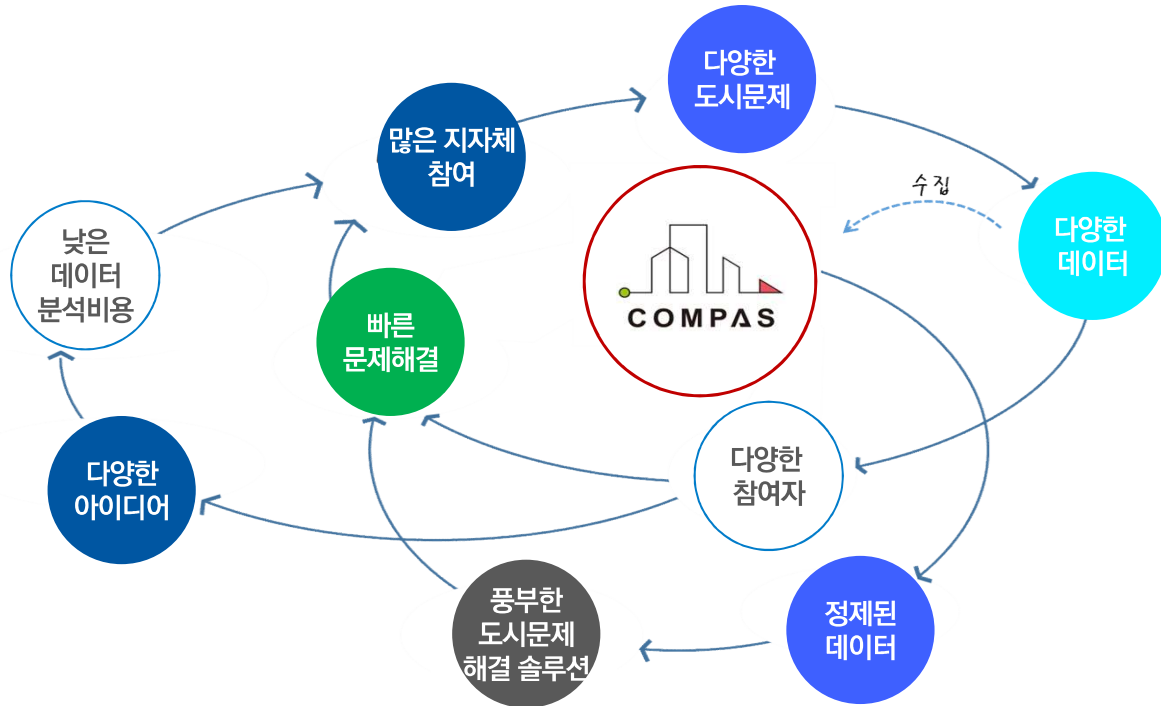
## 도시마다 관심 목표와 당면한 도시문제가 상이 도시가 필요로 하는 관심 목표를 선정하고 그에 필요한 데이터 수집



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

## 국토도시 데이터 분석 경진대회

COMPAS

주요 일정

- 신청 접수 (1월 15일 ~ 1월 25일)
- 선정 발표 (1월 26일)
- 진행 기간 (1월 26일 ~ 2월 15일)
- 결과 발표 (2월 16일)
- 대상 시상 (2월 17일)

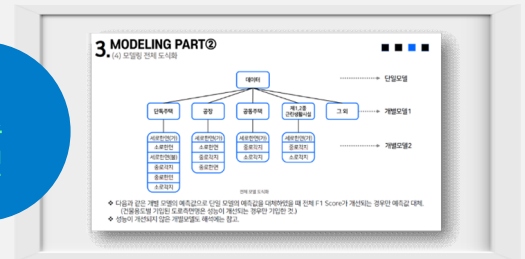
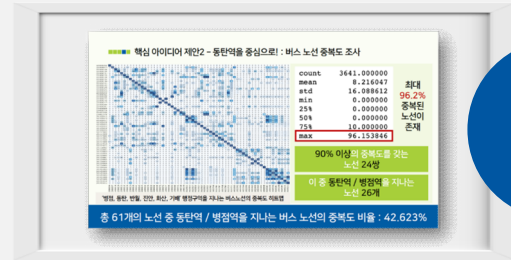
\* 신청서 접수 후 시상이 유첨된 팀은 주초수 시점에 따라 변경될 수 있습니다.  
\* 시상식은 시상장비 및 시상식에 대한 별도 통보 예정입니다.

시상

구분	대상	상금	주요 혜택
최우수상	300만 원 (현금)	300만 원 (현금)	국토도시 데이터 분석과제 특별 시상식
우수상	200만 원 (현금)	200만 원 (현금)	국토도시 데이터 분석과제 특별 시상식
장려상	100만 원 (현금)	100만 원 (현금)	국토도시 데이터 분석과제 특별 시상식

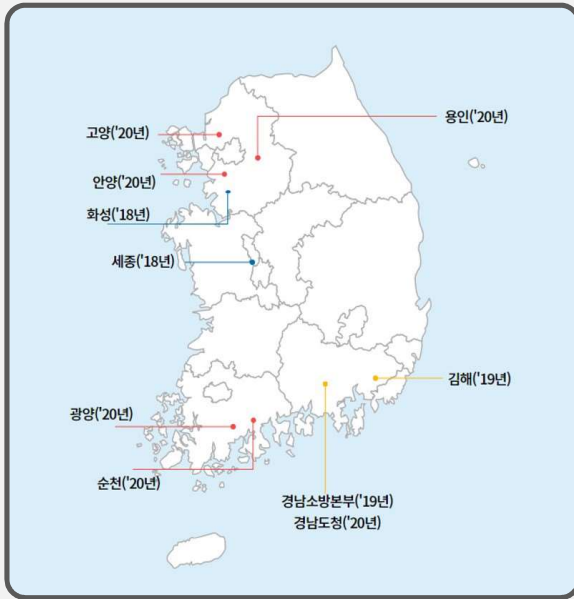
문의처

주최: LH 국토도시데이터 플랫폼사업부 | 후원: LH 국토도시데이터 플랫폼사업부, SK telecom





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



교통



미세먼지



화재 예측



도시 녹지



도시환경

교통사고  
예방

트래픽

황단보도  
신호체계

신호등 체계

어린이 보호구역  
사고 예측

버스노선

 불법 주정차 단속  
카메라 설치 위치 분석

지 자 체	분 야	협 업 과 제
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
광양시	교통	환승 패턴 분석을 통한 환승 노선 최적화
	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

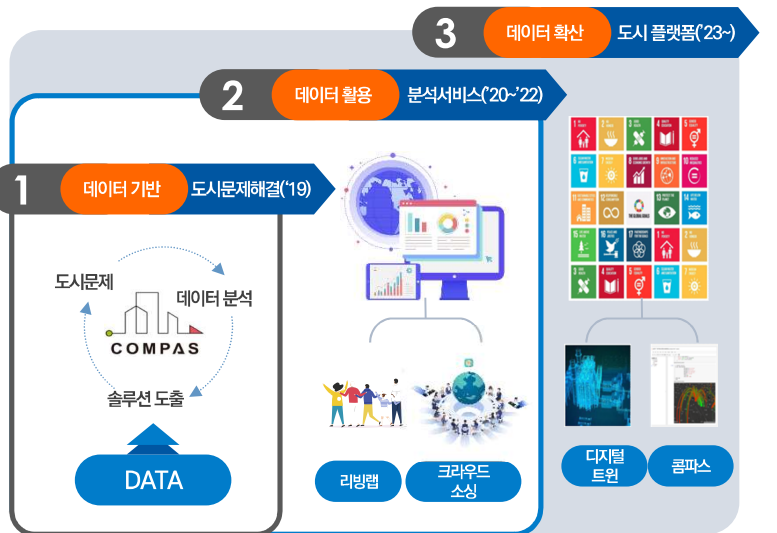


#### 데이터 확산

양질의 데이터 개방을 통한  
신산업 창출

데이터로 관리되고 가치가 창출되는  
LH 스마트시티 플랫폼

### LH 스마트시티 플랫폼 구축 로드맵



CHAPTER

IV

# 협업사례

---

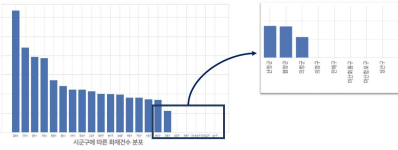


## 김해시는 공장지대의 화재가 빈번히 발생하고 있어 데이터 분석을 통해 화재위험이 높은 건물을 예측하여 방재활동을 추진하고자 함

### 1. DATA PREPROCESSING

(3) 창원 통합시 데이터 식재

다음은 시군구에 따른 화재 발생 분포이다.

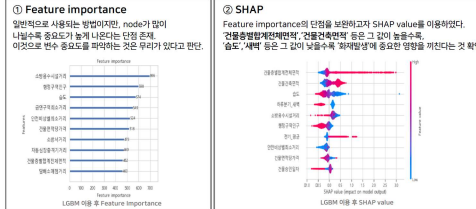


장원 통합시인 진해구, 마산회원구, 마산합포구, 성산군은 화재발생이 0건이다. 이 지역들의 경우 화재발생에 대한 정보를 구하지 못했다고 판단하여 해당 Data는 삭제하였다. 한편 김해시의 화재 발생건수가 압도적인 것을 확인할 수 있었다.

### 2. MODELING PART①

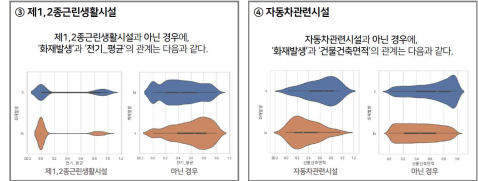
(1) 단일 모델 사용: LGBM

LGBM을 이용하여 변수 중요도를 파악하고자 하였다.



### 3. MODELING PART②

(1) 건물용도에 따른 데이터 분포 탐색

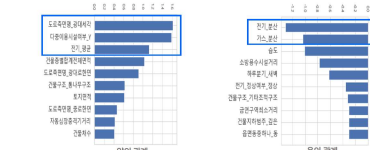


→ 주요 '건물용도'에 따라 데이터의 분포가 다르다는 것을 가시적으로 확인할 수 있다.

### 4. INTERPRETATION

(4) 단독주택에 대한 해석

다음은 단독주택의 경우 화재발생과 양의 관계인 변수들과 음의 관계인 변수들이다.



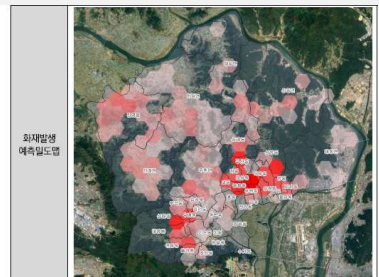
→ 한편, 단독주택의 경우는 화재발생에 영향을 주는 요인이 상성이 다른 것을 확인할 수 있다. 단독주택의 경우 양과 음에 대한 고찰이 더욱 필요하다.

### 4. INTERPRETATION

(5) 결론

- 모델에 대한 결론**
- I. 데이터 자체를 최대한 보존하는 모델이다. 차원 축소와 같은 방법을 진행하면 대용량 정보의 손실과 해석의 어려움이라는 한계점을 수반하기 때문에, 데이터 구조에 변화를 주는 방법을 사용하지 않았다.
  - II. LGBM 모델을 이용하여 전반적인 모델링을 진행했다. 그 후, '건물용도', '도로연면적' 등 주요 변수에 따라 데이터의 분포, 즉 성질이 다르기 때문에 이를 추가적으로 탐색할 수 있도록 각각의 개별모델을 생성했다. 이는 Tree 기반 Boosting 기법의 장점을 살리면서 단점을 보완하는 모델이다.
  - III. EDA 및 상식선에서 여러 모델을 만들었기 때문에 모델 구조의 논리가 합리적이거나 쉽게 재현이 가능하다.
  - IV. 또한 LGBM 모델만 사용함으로써 구현이 쉽고 파라미터도 제한적으로 사용하였다. 파라미터의 경우 4개만 사용하였다. 특히 모델에서는 해당 파라미터도 지정을 사용하지 않았지만, Grid Search Algorithm을 통해 자동화할 한다면 더욱 쉽게 재현이 가능 할 것이다.

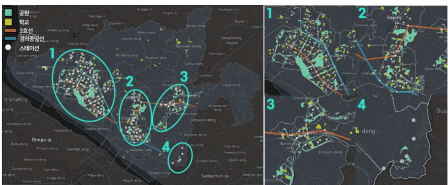
### 5



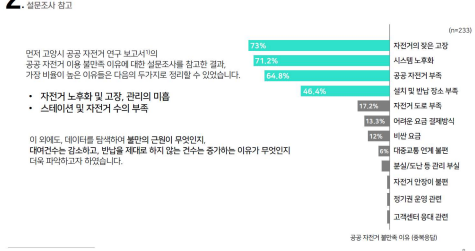
## 고양시는 공공자전거 문제를 개선하고 보다 효율적인 공공자전거 서비스를 제공하기 위하여 데이터 분석을 통한 개선방안 마련

### 1 자전거 스테이션 입지 및 이용 현황

현재 현재 자전거 스테이션은 주로 4개의 지역에 집중적으로 설치되어 있는 것을 확인할 수 있었습니다. 해당 지역들은 **공원, 학교, 지하철역**에 접근성이 매우 높다는 점에서, 고양시의 현재 인구가 집중되어 있는 지역들이라는 것을 짐작할 수 있었습니다.



### 2 불만 상황 파악



### 3 학습용 데이터와 예측용 데이터 구축

학습용 데이터 구축

본격적으로 예측을 위해 ED와 학습용 데이터를 구축합니다. 먼저 **예측용 데이터** 구축 과정입니다.

**STEP1.** 고양시 전체 면적의 50m x 50m 픽셀의 중심점을 후보점으로 설정.

데이터 인공위성(보통 100Mx100M, polygon)에서 중심점에 대한 삼각형, 각 Polygon의 꼭짓점의 중심점으로 후보군을 생성합니다. (예: 50mX50m로 픽셀) 이는 특정 지역의 수리 예측 사례, 공간 데이터 예측을 지원합니다.

이것이, 인공 분포도 없이 어떤 지역은 해당합니다. 따라서, 현재 후보군을 제거는 15963입니다.



**STEP2.** 각각의 후보군에 대해 일별 예측을 위해, 월(1~12)과, 요일(월~금) 변수를 추가.

즉, 하나의 후보군 당, 84=12x7개의 데이터를 예측하는 것입니다. 이후, 각 후보군별로 데이터를 평균 년 값, 상위 25% 값 등을 이용합니다.

**STEP3.** 학습용 데이터와 공통의 변수들을 추가.

데이터 스테이션 번호	month	요일
0	0	Friday
1	0	Monday
2	0	Saturday
3	0	Sunday
4	0	Thursday

예측용 데이터

### 4 모델 학습

해상도 300m 한계점

예측값과 한계점

이제, 학습 시킨 모델로 각 후보군의 예측 값을 출력하여 한계점을 넘어서는 범위에서 100만개의 데이터를 이룰 때까지 남은 값으로 인수를 하지 못하는 것으로 판단되었습니다. 예를 들어, 학습 데이터에서 A 변수의 범위가 10~20 사이라면, 예측 값에서 A 변수 값이 20보다, 10보다 큰 값은 없습니다.

그런데, 학습 데이터의 생활기반 스테이션의 경우, 모두 수치가 있고, 도시에도 많이 분포된 지역이 예상되었습니다.

따라서, 해당 모델은 수치가 없는 경우를 예상하지 못하여, 대부분의 후보군에 수치가 존재하고 예측값이 됩니다.

개선 방안, 사용처

따라서, 실제적인 예측값은 미흡합니다.

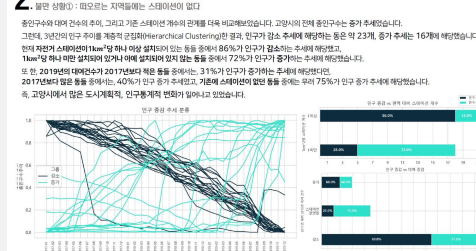
실질적으로 수치가 존재하는 경우를 판단하였습니다.

이후, **도시의 계획** 지역의 예측값, 인구분포도를 기준으로

실용성 수요가 존재하는 지역을 선정하는

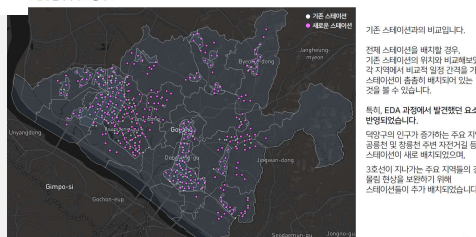
시행, 수리, 관리등은 진행되었습니다.

### 2 불만 상황 파악



### 5 최종 결과 및 평가

불만 해결이 수월, 충족도



가운 스테이션까지 비교합니다. 기존 스테이션을 폐지할 경우, 기존 스테이션의 위치를 비교해보았을 때, 각 지역별 비교의 일정 간격을 가지고 스테이션의 충실도를 평가되어 있는 것을 볼 수 있습니다.

특히, ED 과정에서 발견했던 요소들이 반영되었습니다.

국민의 인구가 증가하는 주요 지역, 공원과 밀접한 수변 지역, 밀접한 스테이션이 새로 배치되었습니다.

3곳은 기존에는 주요 지역들의 경우, 밀접한 밀접한 수변 지역이 스테이션들이 추가 배치되었습니다.

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



**15팀 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') 전체 내려받기

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. * val 격자내의 총 인구수 입니다.
02.안양시_격자별인구현황(고령) 공개 2 col		
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
<ul style="list-style-type: none"> <li>01. 광양시_충전소설치현황 공개 10 col</li> <li>02. 광양시_주차장_공간정보 공개 8 col</li> <li>03. 광양시_자동차등록현황_격자(100X100) 공개 2 col</li> <li>04. 광양시_대중집합시설_생활체육시설 공개 5 col</li> <li>05. 광양시_대중집합시설_아영장 공개 5 col</li> <li>06. 광양시_전기차보급현황(연도별,읍면... 공개 3 col</li> <li>07. 광양시_행정구역별_인구현황(읍면동리) 공개 6 col</li> </ul>	<p>현재 기준의 광양시 전기차 충전소(급속, 완속)에 대한 위치 정보입니다.</p>	<ul style="list-style-type: none"> <li>△ 충전소명 전기자동차 충전소 명칭입니다.</li> <li>△ 충전소위치 전기자동차 충전소 위치입니다.</li> <li>△ 충전소 운영기관 충전기의 운영기관 명칭입니다.</li> <li>△ 급속/완속 충전기의 충전방식 구분입니다.</li> <li>△ 충전기용량 충전기의 용량입니다.</li> <li>* 이용대수 충전기의 이용가능 차량 대수입니다.</li> <li>△ 충전기타입 충전기의 타입입니다.</li> <li>* 요금정보(원/kw) 충전기의 요금정보입니다 (원/kw)</li> <li>* loc 충전소의 경도입니다.</li> </ul>



감사합니다

한국토지주택공사