아이디어 제안서

│. 참가자 정보

공 도	그 명	스쿨존 내 사고분석 및 예방					
성	명	세이프스쿨(유다현)	연 락 처	(전화) 010-6387-6255 (전자우편)ouleesun12@inu.ac.kr			

Ⅱ. 세부 내용

○ 제안서 요약

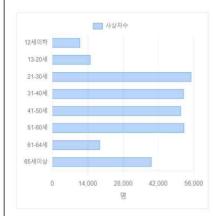
현재 전국의 차량 보유율이 늘어나면서 그만큼 교통사고율 역시 증가하고 있는 추세이다. 또한, 최근 스쿨 존 내 교통사고가 끊임없이 발생하고 있다는 뉴스를 쉽게 접할 수 있기 때문에 스쿨존이 과연 이름만큼 안 전한가에 대해 고찰하기 위해 스쿨존 내 안전시설에 따른 사고 발생 유무를 주제로 택하였다.

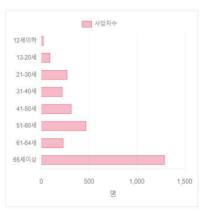
우리의 가설은 옐로우 카펫 존, 방호 울타리 길이, 안전표시판의 개수, 과속방지턱의 개수가 교통사고를 예방하고 있으며 이 안전시설물들의 차이가 사고 유무를 결정한다는 것이었다. 이를 위해 성남시 부근의 4 개의 시인 수원시, 하남시, 광명시, 구리시의 스쿨존을 이용해 학습과 테스트을 하고, 성남시 스쿨존 각각의 특징들을 이용해서 성남시 스쿨존 내의 교통사고 발생 유무를 예측하려 한다. 독립 변수로는 각 스쿨존의 신호등 개수, 옐로우카펫 존 유무, 울타리 길이, 횡단보도 폭, 횡단보도 개수, CCTV 설치 개수, 스쿨내 어린이 놀이시설의 개수, 안전표시판의 개수, 과속방지턱의 개수를 선택했다. 더욱 정확한 예측을 위해 StratifiedK-Fold, GridSearchCV를 이용하여 다른 모델들보다 AdaBoost 모델이 가장 높은 정확도와 F-value를 가지고 있다는 것을 확인하였고, AdaBoost를 사용해 성남시의 스쿨존 내에서 발생한 사고 발생 유무를 예측하였다.

분석 결과로 옐로우 카펫 존 유무와 울타리 길이는 교통사고 예방에 큰 영향을 끼친다는 것을 깨달았다. 교통사고 예방을 위해 스쿨존 내 옐로우 카펫과 방호 울타리를 더욱 추가로 설치하여 어린이들이 안전하게 다닐 수 있는 보행 도로를 확보해야 한다. 그러나 우리의 가설과 달리 과속방지턱과 안전표지판은 미미한 영향을 끼치고 있었다. 이러한 결과가 나온 이유로 데이터를 수집하고 위와 같은 결과를 도축하는 과정에서 과속방지턱과 안전표지판과 같은 안전시설물들이 그저 설치 의무를 위해서 설치되고 있는 것을 확인했으며, 사고 예방에 집중적으로 목적을 두지 않으며 설치되고 있다는 것을 깨달았다. 이 때문에 분명 교통사고 예방에 도움을 줄 수 있는 안전시설물의 설치가 미미한 영향을 끼친다는 분석 결과를 초래했다고 생각한다. 따라서 스쿨존의 이름대로 어린이 보행자에게 안전한 통학로를 제공하기 위해서는 안전시설들의제대로 된 지속적인 감시가 필요하다. 또 다른 분석 결과로는 성남시의 분당구, 수정구, 중원구 중에서 분당구에 사고가 일어날 수 있는 스쿨존의 수가 가장 많다는 것이었다. 특히 분당구의 서현동, 정자동, 야탑동에는 사고 위험성이 높은 스쿨존들이 집중되어 있으니 큰 주의가 필요하다. 더 나아가 이러한 행동으로성남시의 주민, 즉 성남시의 스쿨존을 통행하는 운전자와 보행자가 이러한 시의 점검과 관심으로 더욱 스쿨존 내 보행자 안전에 대해 경각심을 가질 수 있게 노력해야 한다.

○ 배경 및 필요성

현대 사회는 급속하게 경제가 성장함에 따라서, 더불어 교통 환경도 급격하게 변화하게 되었다. 산업화가 이뤄짐에 따라 자동차 산업이 급속도로 발전하였다. 22년 기준 전국 자동차 누적등록 대수는 2500만대를 돌파하였고, 2명당 1대 보유 수준을 보인다. 이렇게 많아진 자동차 수에 따라 관련된 사고 역시 가파르게 증가하였다. 2021년도 기준 교통사고분석시스템에서 보여준 교통사고 사상자 수는 294,524명 사망자 수는 2,916명으로 높은 수치를 보였다. 여기서 교통약자에 대한 사상자수와 사망자수 비율이 앞도적으로 높다는 것을 확인할 수 있다.





출처: 교통사고분석시스템

이렇듯 교통사고 문제는 우리가 관심을 가지고 해결해야 할 사회 문제이다. 여기서 우리는 12세 이하 사 고에 집중하기로 했다. 어린이 교통사고는 그들이 보호받아야 할 어린이 보호구역에서도 끊임없이 일어나 고 있다. 도로교통공단에 따르면, 어린이보호구역 교통사고는 2019년 567건에서 민식이법의 시행 첫 해인 2020년 483건으로 줄었다. 그러나 2021년 다시 523건으로 늘어났으며 최근에도 계속 증가하는 추세이다. 2021년 스쿨존(학교 정문 기준 300m 이내로 운영 중) 내 어린이(12세 이하) 교통사고의 건수를 따져보았 다. taas 교통사고분석시스템에 의하면, 스쿨존 내 어린이(12세 이하) 교통사고는 부상자 수 563명, 사망자 수 2명으로 어린이가 가장 안전해야 할 스쿨존 내에서도 어린이 교통사고가 발생하는 것을 알 수 있다. 이 러한 사고 발생은 성남시의 스쿨존 안에서도 충분히 발생할 수 있다고 생각했으며, 따라서 우리는 성남시 의 스쿨존을 분석하고 미래의 어린이 교통사고를 예방하기 위해 성남시 내 스쿨존을 분석하기로 하였다.

○ 세부 추진내용







출처: 옐로우 카펫존 (성남피플)

출처: 보행자 안전펜스(성남시 금상초등학교)

스쿨존 내 보행자를 보호할 수 있는 가장 보편적인 방법으로는 스쿨존 내 보행자 안전펜스 설치가 있다. 그리고, 부가적인 방법으로는 보행자는 통행에 주의를 기울일 수 있고, 운전자는 보행자에 대한 경각심을 가질 수 있는 옐로우 카펫 존이 있다. 위와 같은 보행자 안전펜스는 운전자의 부주의한 사고로 차량의 보 도 침범을 막아주고, 충돌사고가 발생했다 하여도 1차적으로 충격을 흡수하여 피해가 커지는 것을 막을 수 있다. 또한, 옐로우 카펫 존은 어린이나 보행자를 위해서 횡단보도의 대기 공간을 노란색으로 도색한 안전 시설물이다. 노란색 특유의 색으로 사람들의 시야 구분을 확실하게 해주기 때문에 운전자는 옐로우 카펫 존에 있는 보행자를 쉽게 인식할 수 있게 된다. 따라서 옐로우 카펫 존은 차량 속도의 감속, 보행자 주의 를 이끌어주는 효과가 있다. 보행자 역시 차량이 지나다니는 곳이라는 경각심을 가지게 해주어 좀 더 안전

에 주의하게 되는 효과를 준다. 또한, 보행자와 운전자를 감시할 수 있는 CCTV 설치, 차량의 속도를 감속 시켜주는 과속방지턱 역시 보행자의 보호에 도움을 준다. 우리는 이와 같은 시설들을 이용하여 성남시의 스쿨존 내 사고 발생의 유무를 예측하였다.

1) 데이터 전처리과정

먼저 횡단보도, 안전표시판, 과속방지턱과 같은 원본 데이터의 위도, 경도 좌표계를 거리 계산을 위해 'EPSG:4326'로 변환해주었다. 변환한 위도, 경도와 스쿨존 중앙의 위도, 경도를 이용해 haversine으로 거리를 계산해준 후, 거리가 300m 이내에 포함된 시설물들만 스쿨존마다 개수를 합산했다. 이 방법으로 스쿨존 내의 횡단보도 개수, 안전표시판 개수, 과속방지턱 개수를 독립변수로 만들 수 있었다.

성남시 스쿨존 내에서의 사고 발생 유무를 예측하기 위해서 성남시 스쿨존의 사고 발생 유무를 예측값으로 두었다. 그리고 모델을 학습시키기 위한 데이터로 성남시 근방의 4개의 시를 선정하였는데, 광명시, 하남시, 수원시, 구리시로 결정하였다. 독립변수는 아래와 같이 설정했다.

- 신호등 개수 : 스쿨존 300m 이내에 설치된 신호등 개수
- 옐로우카펫 : 스쿨존 300m 이내에 설치된 옐로우 카펫 존의 유무 (0 : 무, 1 : 유)
- 울타리 길이 : 스쿨존 300m 이내의 방호 울타리의 길이
- 횡단보도폭 : 스쿨존 300m 이내의 횡단보도들의 폭 평균
- 횡단보도개수 : 스쿨존 300m 이내의 횡단보도 개수
- CCTV설치대수 : 스쿨존 300m 내에 설치된 CCTV 개수
- 어린이놀이시설 개수 : 스쿨존 300m 이내의 어린이놀이시설 개수
- 안전표시개수 : 스쿨존 300m 이내에 설치된 안전표지판 개수
- 과속방지턱 : 스쿨존 300m 이내에 설치된 과속방지턱 개수

	A	В	C	D	E	F	G	Н	1	J	K	L	M	N
1	대상시설명:	소재지도로	위도	경도	신호등개수	옐로우키펏	울타리길이	횡단보도족	횡단보도?	CCTV설치	(어린이놀0	안전표시기	과속방지트	사고횟수
2	안서초등학	경기도 광당	37.40278	126.8706	1	0	133	5.7	7	5	30	8	4	0
3	구름산초등	경기도 광당	37.45299	126.8842	8	0	1989	5.7	7	3	12	28	4	0
4	하일초등학	경기도 광당	37.46909	126.8764	2		1332	6.1	7	4	6	4	6	0
5	철산초등학	경기도 광당	37.4698	126.8746	9	0	2368	5.6	17	5	4	10	6	0
6	하안남초등	경기도 광당	37.46217	126.8853	2	0	819	5.7	7	4	10	11	8	1
7	광명남초등	경기도 광당	37.47621	126,8531	6	0	228	5.8	14	5	11	5	9	2
8	광명동초등	경기도 광당	37.48395	126.8634	10	0	181	6.2	16	5	2	6	11	2
9	소하초등학	경기도 광당	37.44928	126,8874	8	.0	1749	5.7	7	5	10	39	11	1
10	연서초등학	경기도 광당	37.46596	126.8824	6	0	1083	5.9	4	4	8	4	14	0
11	서면초등학	경기도 광당	37.43864	126.8796	8	1	1086	5.7	7	5	18	6	14	1
12	광성초등학	경기도 광당	37.47885	126.8731	6	.0	672	6.1	2	4	2	7	14	0
13	하안북초등	경기도 광당	37.46513	126.869	6	0	1511	4.2	4	4	. 2	7	14	0
14	충현초등학	경기도 광당	37.43314	126.8865	9	1	556	5.7	7	5	24	8	14	1
15	광명초등학	경기도 광당	37.48086	126.8574	6	.0	192	6.6	12	5	2	2	16	9
16	하안초등학	경기도 광당	37.46306	126.874	6	0	1374	7.4	15	5	22	8	20	0
17	광일초등학	경기도 광명	37.47162	126.8469	9	0	1383	6.2	3	4	7	13	20	0

광명시의 데이터

종속변수로는 사고횟수를 선택했지만 사고 유무를 예측하기 위해 0과 1로 변환해주었다(0 : 무, 1 : 유). 그리고 성남시 스쿨존 데이터도 4개의 사와 마찬가지로 스쿨존 내 사고 발생 유무를 예측하기 위해 보행자 안전펜스 길이, 신호등 개수, 옐로우 카펫 존 유무, 횡단보도 폭, 횡단보도 개수, cctv 설치 대수, 어린이 놀이시설의 개수, 안전표시판 개수, 과속방지턱 개수를 입력했다. 또한, Sickit-Learn의 StandardScaler을 이용하여 데이터들의 특성을 표준화하였다.

2) 분석 tool

먼저 데이터 분석을 위한 프로그래밍 언어로는 파이썬(python)을 택하였다. 상호 대화식, 즉 코드를 작성하고 바로 실행해 결과를 볼 수 있으며, 그래프나 이미지 출력 등의 시각화를 위하여 주피터 노트북 (jupyter notebook)이란 웹 플랫폼과 아나콘다(anaconda)를 사용하였다. 또한, 편리한 사용을 위하여, 엑셀도 함께 사용하였다. 사용 라이브러리는 다음과 같다. dataframe을 사용하고 계산을 하기 위해서 Pandas와 Numpy를, 위치 데이터를 제공해주는 googlemaps, 위도 경도를 통해 거리를 계산하는 라이브러리 haversine, 지도를 시각화해주는 folium, 다양한 그래프로 데이터를 시각화하기 위한 파이썬 라이브러리 Matplotlib를 이용하였습니다. 과속 방지턱, 옐로우 카펫존 유무, 보행자 울타리 길이를 데이터로 사용하기 위해 'daum'의 'kakaomap'을 이용하였다. 데이터 분석을 위해서 Sickit-Learn을 사용하였다.

3) 데이터 분석

정확한 예측을 위해서 다양한 학습 모델들을 선정하고 그중에서 정확도가 높은 것을 선택하였다. 일단, Sickit-Learn의 train_test_split을 사용하여 train 세트 30%, test 세트 30%로 분할하였다. 학습을 위해 사용할 모델로 Logistic Regression, RandomForest, SVM(Support Vector Machine), AdaBoostClassifier를 선택하였다. 각 모델들의 최고 성능을 낼 수 있는 매개변수를 알아내기 위해 GridSearchCV와, 불균형하고 적은 양의 데이터의 정확도를 더욱 향상시키기 위해 K-Fold와 K-Fold 교 차검증의 종류인 StratifiedK-Fold와 RepeatedStratifiedK-Fold를 이용하였다. 그리고 각 모델들의 성능비교를 위해 정확도인 Accuracy score를 사용하였다.

```
from sklearn.svm import LinearSVC
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.svm import SVC
from sklearn.ensemble import AdaBoostClassifier
from sklearn.metrics import accuracy_score
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import StratifiedKFold
from sklearn.model_selection import KFold. cross_val_score
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.pipeline import Pipeline
import pandas as pd
import numpy as np

import warnings
warnings.filterwarnings('ignore')
```

분석 사용라이브러리

```
params ={
    'n_estimators': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100],
    'learning_rate': [0.0001, 0.001, 0.01, 0.1, 1.0]
}

adaboost = AdaBoostClassifier(random_state=10)
grid_adaboost = GridSearchCV(adaboost, param_grid=params, cv=skf, scoring='accuracy')
best_model = grid_adaboost.fit(X_train,y_train)

# 테스트 세트 점수
test_score = grid_adaboost.score(X_test, y_test)

print("테스트 세트 점수: {:.2f}".format(test_score))
print("최고 교차 검증 점수: {:.2f}".format(grid_adaboost.best_params_))
print("최고 교차 검증 점수: {:.2f}".format(grid_adaboost.best_score_))

테스트 세트 점수: 0.73
최적 매개변수: {'learning_rate': 0.1, 'n_estimators': 50}
최고 교차 검증 점수: 0.61

#위의 결과로 나온 최적 하이퍼 파리미터로 다시 모델을 확습하여 테스트 세트 데이터에서 예측 성능을 측정
adaboost = AdaBoostClassifier(learning_rate = 0.1, n_estimators = 50, random_state=10)
adaboost.fit(X_train,y_train)
pred = adaboost.predict(X_test)
accuracy = accuracy_score(y_test,pred)
accuracy
```

0.7254901960784313

데이터 분석의 예 (AdaBoost)

	Logistic Regression	SVM	Random Forest	AdaBoost	
test 세트 정확도	0.63	0.69	0.69	0.73	
최고 교차검증 점수	0.63	0.69	0.64	0.61	
최적 매개변수	C= 0.01, penalty= 'l2', solver= 'newton-cg'	C=50, gamma=0.1, kernel='sigmoid'	max_depth= 6, min_samples_leaf=8, min_samples_split=8, n_estimators= 100	learning_rate = 0.1, n_estimators = 50	
train 세트 정확도	train 세트 정확도 0.63		0.71	0.72	

모델들의 정확도 비교

train 세트를 이용하여 모델의 정확도를 살펴보았을 때 AdaBoost가 가장 높은 정확도가 나왔지만, test 세트를 예측했을 때에 RandomForest와 AdaBoost가 서로 비슷한 정확도가 나와 더 좋은 모델을 선정하기 위한 비교가 필요했다. Sickit-Learn의 classification_report를 사용하여 F1-score를 비교하였다. F1-score를 기준으로 정한 이유는 recall과 precision의 조화평균이므로 분류 모델에서 많이 사용되는 평가지표이기 때문이다. RandomForest보다 AdaBoost의 F1-score가 더 높게 나왔으므로 성남시의 사고 유무를 예측할 모델로 AdaBoost를 택하였다.

RandomForest	precision	recall	f1-score	support	AdaBoostClass	ifier precision	recall	f1-score	support
0 1	0.79 0.61	0.71 0.70	0.75 0.65	31 20	0 1	0.77 0.65	0.77 0.65	0.77 0.65	31 20
accuracy macro avg weighted avg	0.70 0.72	0.70 0.71	0.71 0.70 0.71	51 51 51	accuracy macro avg weighted avg	0.71 0.73	0.71 0.73	0.73 0.71 0.73	51 51 51

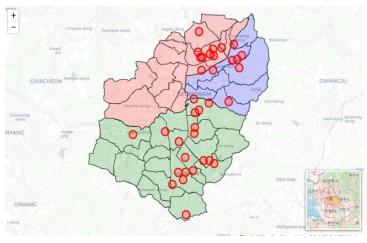
최종적으로 AdaBoost가 예측한 성남시의 스쿨존 내 사고 유무의 결과로 성남시의 73개 스쿨존 중에서 35개의 스쿨존이 사고가 발생할 것이라 예측되었다. 분당구에서는 19개, 수정구에서는 9개, 중원구에서는 7개의스쿨존이 위험했다.

횡단보도폭 횡단보도개수 CCTV설치대수 어린이놀이시설 개수 안전표지판 개수	5.342857 0.914286 49.028571 6.234286 5.342857 4.428571 7.314286 5.628571 19.514286	[사고가 안일어날 2 신호등개수 옐로우카펫 울타리길이 황단보도폭 황단보도개수 CCTV설치대수 어린이놀이시설 개 안전표지판 개수 과속방지턱	3.184211 0.921053 871.342105 6.110526 3.157895 3.631579
과속방지턱 dtype: float64	19.514286	과속방지턱 dtype: float64	14.500000

그리고 우리는 사고가 발생할 수 있는 스쿨존의 시설들과, 그렇지 않은 시설들을 분석 및 분석했다. 예상외의 결과가 나왔는데, 상식적으로 우리의 가설은 안전표시판 개수와 CCTV설치 대수, 과속방지턱의 개수가 사고를 줄이는 것에 영향을 줄 것이라고 가정했었다. 그러나 분석 결과, 사고가 일어날 스쿨존에 앞서 말한 특징의 개수가 사고가 안 일어날 지역보다 높은 것을 보였다. 따라서, 안전표시판 개수, CCTV설치 개수, 과속방지턱의 개수는 사고 예방에 큰 도움이 되지 않는 것을 알 수 있었다.

나머지 특성들을 살펴보면, 사고가 일어난다고 예측되는 스쿨존에서는 일어나지 않을 것이라 예측되는 스쿨 존보다 울타리의 길이가 현저하게 짧았다. 또한, 옐로우 카펫 존 유무(0:무, 1:유)로 계산했기 때문에 사고가 일어날 스쿨존은 그렇지 않은 스쿨존에 비해 옐로우 카펫 존이 없는 스쿨존이 더 많은 것을 알게 되었다. 또 한, 신호등의 개수, 횡단보도 개수가 사고가 일어날 지역이 더 많은 것을 보여주었다.

Folium과 GeoPandas를 이용하여 사고가 일어날 가능성이 높은 스쿨존을 시각화해보았다. 아래의 이미지는 그 결과이다.



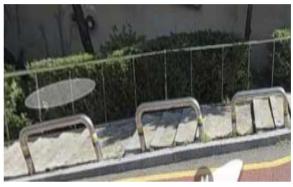
수정구(빨간색), 중원구(파란색), 분당구(초록색)

또한, 사고가 일어날 수 있는 스쿨존을 3개 이상 포함하는 동을 선정하였다. 수정구는 태평동과 신흥동, 중원구는 금광동, 분당구는 서현동, 정자동, 야탑동이 스쿨존 내 교통 사고 위험성이 높다는 것을 알 수 있었다.

○ 활용방안 및 기대효과

상식적으로 안전표시판 개수와 CCTV설치 대수, 과속방지턱의 개수가 사고를 줄이는 것에 영향을 줄 것이라고 가설을 세웠었다. 하지만, 분석 결과 영향이 미미한 것이 보여 우리는 좀 더 조사하게 되었다. 거리뷰로 표지판을 확인하니 스쿨존 내 표지판이 존재하지만, 운전자가 쉽게 알아볼 수 없는 위치에 설치되어 있었으며, 야간에는 보이지 않는 경우가 많았다. 도로환경 등 세부 사항을 고려하지 않고 단순히 의무를 위해서 설치한 표지판은 운전자들에게 경각심을 주지 못해 사고 예방에는 큰 도움이 되지 않았던 것이다. 그리고 우리는 거리뷰로 울타리 길이를 재는 과정에서 보행자 안전 울타리 또한 일정한 규격으로 설치되어 있지 않은 것을 알게 되었다. 따라서 스쿨존의 이름대로 어린이 보행자에게 안전한 통학로를 제공하기 위해서는 안전시설들의 제대로 된 환경 실태 조사와 지속적인 감시가 필요하며, 규격을 맞추어서 어린이들이 안전하게 다닐 수 있는 보행 도로를 확보하여 실질적으로 사고예방에 도움이 되도록 해야 한다.





금상초등학교의 보행자 안전펜스

상원초등학교의 보행자 안전펜스

앞선 분석 결과에서 보행자 안전펜스의 길이와 옐로우 카펫 존의 유무가 사고 예방에 큰 영향을 주는 요인인 것을 알 수 있었다. 그러므로 사고가 일어날 것이라 예측된 스쿨존에 보행자 안전펜스와 옐로우 카펫 존을 추가로 설치하는 것을 건의한다. 특히 사고 위험성이 높다고 판단되는 분당구의 서현동, 정자동, 야탑 동에는 사고 위험성이 높은 스쿨존들이 집중되어 있으니 큰 주의가 필요하다. 더 나아가 이러한 행동으로 성남시의 주민, 즉 성남시의 스쿨존을 통행하는 운전자와 보행자가 이러한 시의 점검과 관심으로 더욱 스쿨존 내 보행자 안전에 대해 경각심을 가질 수 있게 노력해야 한다.

○ 활용데이터 및 참고 문헌 출처 등

EJ.	0	F-	출기	ŧ.
-	_		<u> </u>	١,

데이터명	데이터 출처
국토교통부 국토지리정보원_읍면동 경계	공공데이터 포털(행정안전부)
경기도 성남시 옐로카펫 설치현황	공공데이터 포털(행정안전부)
경기도 스쿨존 위도경도	공공데이터 포털(행정안전부)
스쿨존내 어린이(12세 이하)교통사고	도로교통공단 Tass
전국 교통사고 다발 지역 표준데이터	공공데이터 포털(행정안전부)
경기도 어린이 놀이시설 현황	공공데이터 포털(행정안전부)
경기도 성남시_울타리 및 무단횡단 방지펜스	공공데이터 포털(행정안전부)
전국어린이보호구역표준데이터(cctv 개수)	공공데이터 포털(행정안전부)
전국신호등표준데이터	공공데이터 포털(행정안전부)
전국도로안전표지표준데이터	공공데이터 포털(행정안전부)
경기도_어린이 놀이시설 현황	공공데이터 포털(행정안전부)
행정안전부_과속방지턱정보	공공데이터 포털(행정안전부)
경기도 광명시_어린이보호구역 내 안전표지 정보	공공데이터 포털(행정안전부)
전국횡단보도표준데이터	공공데이터 포털(행정안전부)
경기도 수원시_어린이보호구역별 교통안전시설	공공데이터 포털(행정안전부)