

2023년 1학기 P-SAT 주제분석

# 서울특별시 안심택배함 재배치 분석

-안심택배함 이용 건수 예측을 바탕으로





데이터마이닝팀 | 김수빈 조건우 김보현 이지원 조성우



## 안심택배함이란?

#### 무인 택배함 기능

바쁜 직장인들이 시간에 구애 받지 않고 편리하게 택배를 이용 -

범죄 예방 기능

1인 가구를 대상으로 하는 택배 관련 및 각종 범죄가능성 예방

~~ (사학 인심택배활

1인 가구 밀집지역 및 다가구 주택 주변으로 현재 총 250여 개의 안심 택배함 운영 중 택배 도난으로 인한 개인정보 유출 방지 및 위급 상황 시 사용할 수 있는 비상벨과 CCTV를 설치





안심택배함





## 안심택배함 관련 수요

#### 비대면 수령에 대한 수요 증가

1인가구와 맞벌이 가구의 증가로 낮에 직접 택배 수령 어려운 주민 증가

#### 택배 도난 범죄 증가

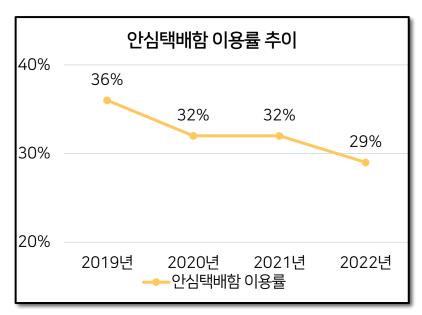
코로나 19로 인한 비대면화 이후 택배 도난 사건 증가 추세

#### 1인 가구 대상 범죄 증가

1인 가구를 대상으로 하는 범죄 위험 증가



## 안심택배함 이용 현황



그러나, 아이러니하게도 안심택배함의 이용률은 매년 **하락세** 



안심택배함 이용률 하락의 주된 원인은 택배함을 설치할 때 **접근성**과 **주위환경**을 고려하지 않았기 때문…



안심택배함 위치의 **재조정** 작업 필요성 대두





## 분석 방향성

## 안심 택배함의 필요성 증가

1인가구와 맞벌이 가구의 증가로 낮에 직접 택배 수령 어려운 주민 증가

## 감소하는 안심 택배함 이용률

비효율적인 배치로 인해 점차 줄고 있는 안심택배함 이용률





비효율적인 안심택배함의 위치를 더욱 효율적인 위치로 재조정하여 안심택배함의 이용률을 높이고 사회문제를 해결하자

## 주제 분석 흐름

안심택배함과 관련된 **데이터를 수집하고 EDA**를 통한 인사이트 도출



다양한 모형을 통해 현재 안심택배함이 설치되지 않은 곳에 설치 시의 이용건수 예측



예측된 이용건수를 바탕으로 비효율적인 입지의 재배치 제안



예측된 이용건수와 현재의 이용건수를 비교하여 정책의 결론 도출



## 데이터 수집 방법

#### 공공 데이터







서울 열린데이터 광장

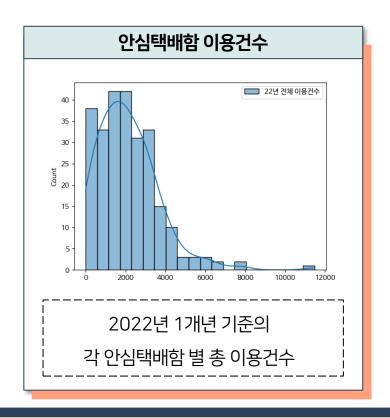
설치 장소 종류, 1인가구수, 직장인구수, 인구밀도, 국민기초생활보급자수, 생활인구수 ···

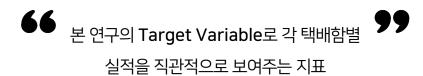
#### 카카오 API



반경 내 지하철역 수, 은행 수, 주차장 수, 숙박시설 수, 경찰서 수, 공원 수, 도서관 수, 학교 수 …

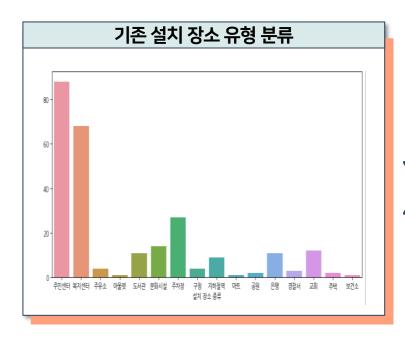
## 안심택배함 이용건수

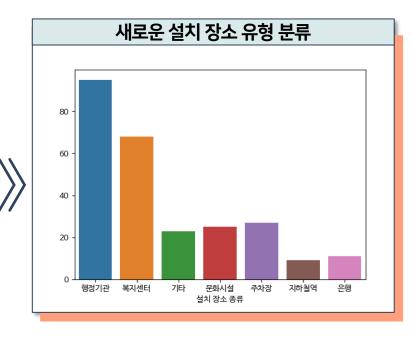




- 상봉1동 주민센터는 2022년에 미운영이므로 제외
- 2022년 12월 말에 운영을 시작하여 전체 연도의 이용건수를 반영하지 못하는 택배함 존재
- → 2022년 기준 직전 년도 미운영 택배함 모두 제거

총 245개 택배함에 대해 분석 진행

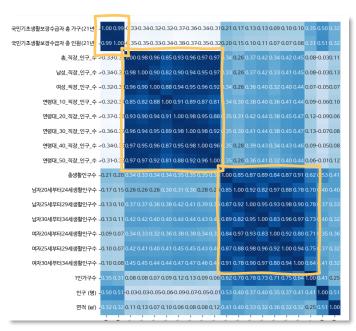




기존 분류의 종류가 너무 많다고 생각되어 유사한 특성을 지니는 장소끼리 묶어줌

#### $\square X$

## 인구 특성 변수



cut-off 0.8로 설정하여 그룹끼리 높은 상관관계를 갖는 변수들 중 하나만 선택

#### 국민기초생활보장수급자

국민기초생활보장수급자 총 가구

국민기초생활보장수급자 총 인원

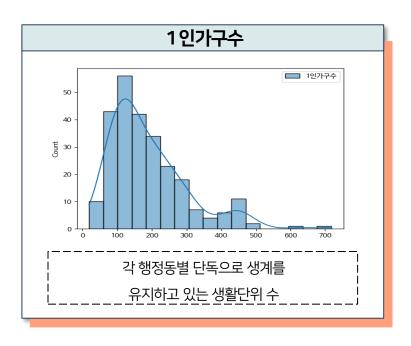
#### 직장인구

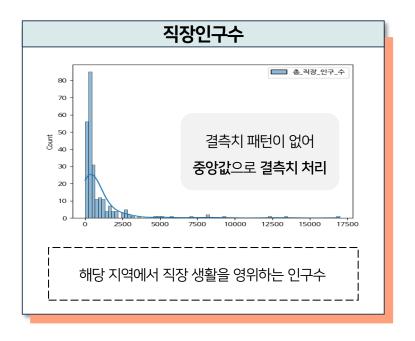
총직장인구수	
남자직장인구수	
여자직장인구수	

#### 생활인구

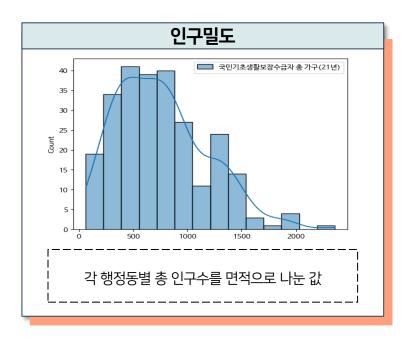
총생활인구수
남자20~24생활인구수
여자20~24생활인구수

상관관계를 확인했을 때, 성별과 나이를 고려할 필요가 없어 보여 각각 <mark>총 인구 수</mark> 사용





1인 가구 대상 사업임을 감안하여 **1인 가구 수** 필요 유동인구 분석 시, 거주인구와 **직장인구**를 나누어 고려하는 것이 필요하다 판단

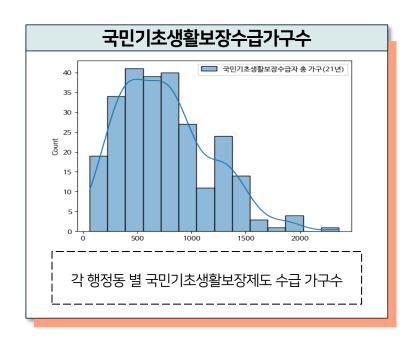


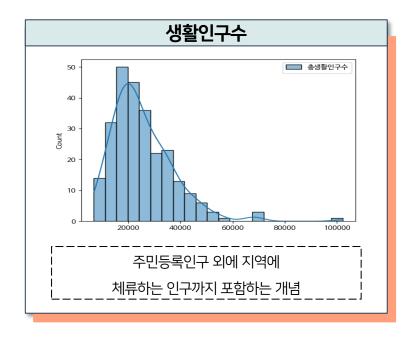
#### 김태일 외. (2023) 서울시 여성안심택배함 증설 최적 입지 분석 中

집지역을 분석에 반영하였다. 또한 서울시 인구밀도 (동 별) 통계 데이터를 통해 인구밀도를 데이터에 포함시켜 주거인구가 밀집한 지역을 데이터에 반영하였다. 참고로 한국교통연구원의 택배화물 발생량과인구밀도가 양의 상관관계를 가진다는 연구 결과 기반으로 택배 수요 측정에 인구밀도 데이터 사용하였다[3].

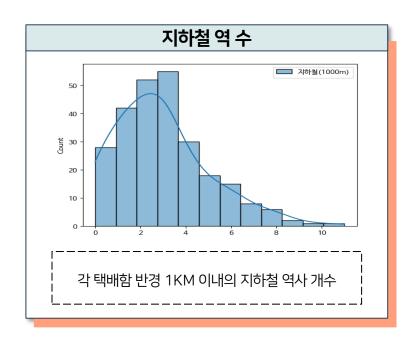
택배화물 발생량을 대체하는 변수로서의 역할 가능

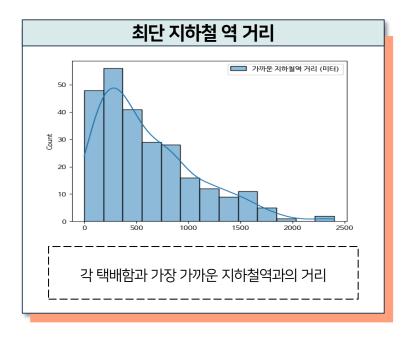
지역의 면적을 고려한 인구수를 살펴보기 위한 파생변수





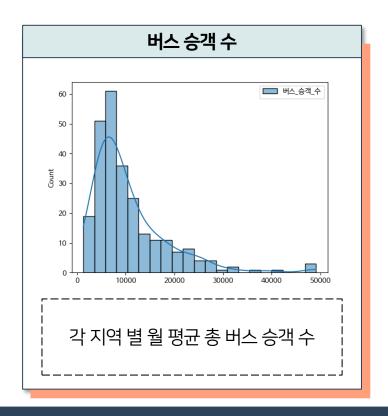
**국민기초생활보장수급수**는 각 지역별 취약계층의 정도를 파악하기 위한 변수로 추가 생활인구수는 해당 지역에서 생활하는 전체 인구수를 살펴보기 위해 선정





지하철역과 걸어서 15분 정도의 거리면 접근성이 높은 편이라 판단하여 반경 1km로 설정 지하철역과 가까운 정도로 접근성의 정도 및 유동인구의 지표로 사용될 수 있을 것이라 판단

## 유동인구/접근성 관련 변수





유동인구를 살펴보기 위한 변수로 교통 변수 중 버스 승객 수 선택



## 편의시설 변수

#### 이향숙, 진무위, 추상호. (2017). 무인택배함의 최적입지 선정을 위한 방법론 개발 中

무인택배함 적정 위치를 선정하기 위해서는 명확한 기준을 설정해야 한다. 입지선정 관련 국내외 연구를 종합적으로 요약해보면 공공시설의 입지조건으로는 주로 접근성 및 교통의 편리성, 상가지역 및 지역의 중심지가 언급되었으며, 그 외 확장 가능한 부지와 주차공간, 자연환경 등이 거론되었다. 본 연구는 Han(2002)

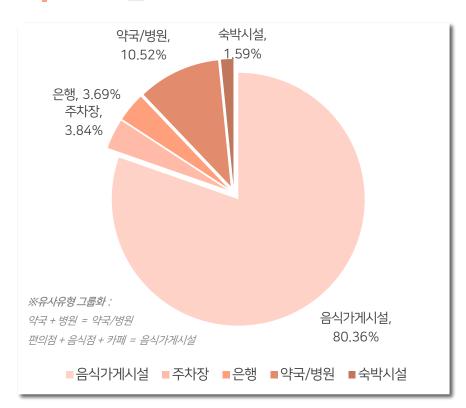


안심택배함의 적정 위치를 선정하기 위해 상가지역 및 지역의 중심지인지 확인



따라서 설치된 안심택배함 주변 편의시설과 공공시설의 개수를 수집

## 편의시설 변수

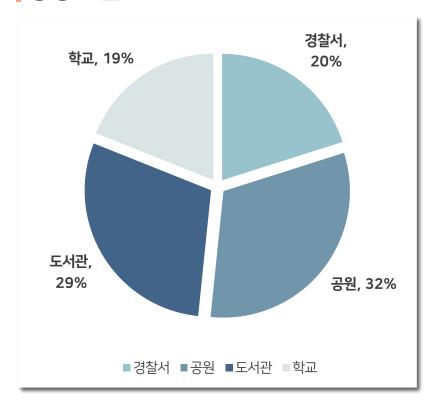


변수	설명	
약국/병원 수	반경 500m 내 약국/병원 개수	
은행 수	반경 500m 내 은행 개수	
주차장 수	반경 500m 내 주차장 개수	
음식가게시설 수	반경 500m 내 음식가게시설 개수	
숙박시설 수	반경 500m 내 숙박시설 개수	

반경 500m: 가까운 편의시설의 기준을 도보 5분으로 지정

공통적인 특징으로 **유동인구**가 많은 지역이 시설이 밀집되어 있음

## 공공시설 변수

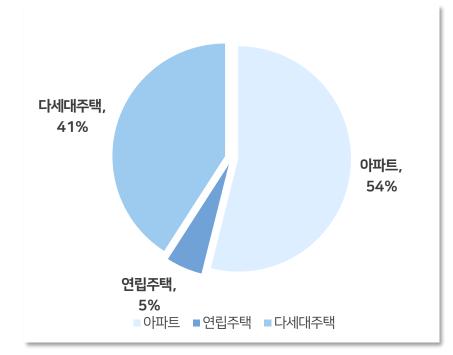


변수	설명
경찰서 수	반경 1000m 경찰서 개수
공원 수	반경 500m 내 공원 개수
도서관 수	반경 500m 내 도서관 개수
학교 수	반경 500m 내 학교 개수

반경 1000m: 유사한 결과를 피하기 위해

**학교가 많을수록 생활인구수가 높다**는 연관성 확인 할 수 있음

## 주변 거주환경 변수



변수	설명	
아파트	행정동별 아파트에 거주하는 가구 수 (층수가 5개 층 이상)	
다세대주택	행정동별 다세대주택에 거주하는 가구 수 (층수가 4개 층 이하 1개 동의 바닥면적의 합이 660㎡ 초과)	
연립주택	행정동별 연립주택에 거주하는 가구 수 (층수가 4개 층 이하 1개 동의 바닥면적의 합이 660㎡ 이하)	



## 지도 시각화



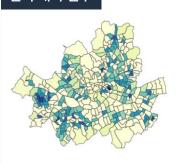
생활 인구



국민기초생활보장수급자



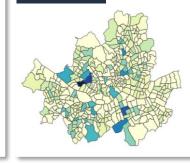
면적 대비 인구



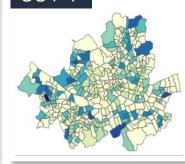
1인 가구



버스 승객 수



<del>공동주</del>택





## 클러스터링

## Why Clustering?



안심택배함이 설치된 장소들을 군집화하여 설치된 지역의 내재적 특성 파악



안심택배함 위치의 <mark>재조정</mark>을 목표 클러스터링을 통해 이용건수가 높은 곳과 그렇지 않은 곳의 특징 추출

지표	Best Subsets	
Adj. R sq.	1인 가구 수, 인구면적, <mark>총생활인구 수,</mark>	
	숙박시설 수, 주차장 수, 아파트 수,	
	최단지하철역거리, 경찰서 수,	
	음식가게시설 수	
Ср	인구 면적, 총생활인구 수,	
	<u>최단지하철역거리,</u> 경찰서 수	
BIC	최단지하철역거리	

3개의 지표 중 가장 많이 나오는 변수로 최단지하철역거리과 총생활인구를 확인



## 클러스터링

#### 출처- 안심택배함 과업지시서

- 3. 설치 입지조사 및 안심택배함 설치
  - 가, 일반주택가 지역
  - 싱글여성, 맞벌이가구 등이 많은 원룸, 빌라, 다가구주택 지역 등
  - <u>택배물량이</u> 많이 유통되는 지역 중 임시보관소(편의점, 슈퍼 등)의 보관 물량이 많은 지역
  - 기존 공동주택, 기숙사 등 경비원 상주시설은 입지요건에서 제외
  - 나, 유통인구가 많고, 주거지에서 접근이 용이한 지역
  - 유동인구가 많고 교통이 편리한 상권
  - 재래시장, 대학가 등과 밀접한 주거지역
  - ㅇ 공단 등 근로자 밀집 거주지역
  - 여성안심주택, 임대주택 등
  - 기타 <u>택배수요가</u> 많을 것으로 예상되는 지역
  - 다. 보안 및 편의시설 이용이 가능한 지역
  - 방범 cctv 설치 구역, 주차장 이용이 편리한 지역, 시설관리자 상주지역 등
  - 눈·비 가리 및 전력·통신 <u>장비설치가 용이한</u> 지역(필요시 캐노피 설치 보완 등)

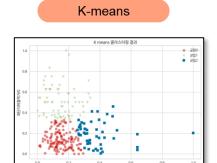
정부가 제시한 과업지시서를 참고한 결과 생활인구수과 최단 지하철역 거리가 안심택배함의 입지를 살펴보기 좋은 변수임을 확인

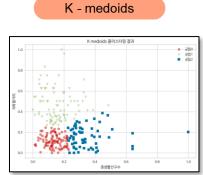


**총생활인구수**과 **최단 지하철역 거리**를 최종 클러스터링 변수로 확정



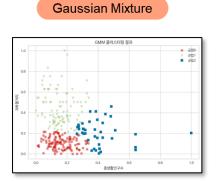
## 클러스터링







Hierarchical





모든 클러스터링 모델에서 가장 가까운 지하철 역과의 거리가 가까울수록 실적이 높고 PAM과 GMM의 경우 생활인구가 많은 군집일수록 실적이 높음을 확인



## 모델링을 위한 변수 처리

#### **One-Hot Encoding**

X 변수들 중 유일한 범주형 변수인 '설치 장소 종류'에 대해 원핫인코딩 진행

#### **Min-Max Scaling**

거리 기반 모델링을 수행할 때 상대적으로 범위가 넓은 변수들에 대해 영향력이 더 높은 변수로 인식되는 문제를 방지하기 위해 범위를 재조정

#### 베이지언 최적화

베이지안 최적화 기반 오픈 소스 라이브러리인 OPTUNA를 이용해 하이퍼파라미터 튜닝 진행



## KNN regression

#### **Full Model**

One-Hot Encoding
MinMax Scaling
OPTUNA



**ALL Features** 

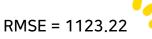
RMSE = 1128.53

#### **Model with Best Subset Selection**

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 인구밀도, 총생활인구수, 숙박시설, 주차장, 아파트, 지하철역거리, 경찰서, 음식가게시설



#### Model with SHAP selection

One-Hot Encoding
MinMax Scaling
OPTUNA



행정기관, 문화시설, 복지센터, 주차장, 지하철역거리, 아파트, 지하철, 인구밀도, 기초생활수급자, 경찰서, 도서관,다세대주택

RMSE = 1195.95



## Support Vector Regression

#### **Full Model**

One-Hot Encoding MinMax Scaling **OPTUNA** 



**ALL Features** 



MinMax Scaling \



변수 선택은 진행하지 않음 기계시설





## Random Forest Regressor

#### **Full Model**

One-Hot Encoding
MinMax Scaling
OPTUNA



**ALL Features** 

RMSE = 984.50

#### **Model with Best Subset Selection**

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 인구밀도, 총생활인구수, 숙박시설, 주차장, 아파트, 지하철역거리, 경찰서, 음식가게시설

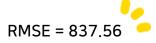
RMSE = 900.22

#### Model with SHAP selection

One-Hot Encoding
MinMax Scaling
OPTUNA



지하철역거리, 행정기관, 총생활인구수, 다세대주택, 버스승객수, 인구밀도, 음식가게시설, 1인가구수, 연립주택, 경찰서, 총직장인구수, 기초생활수급자





## XGBoost

#### **Full Model**

One-Hot Encoding
MinMax Scaling
OPTUNA



**ALL Features** 

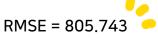
RMSE = 936.279

#### **Model with Best Subset Selection**

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 인구밀도, 총생활인구수, 숙박시설, 주차장, 아파트, 지하철역거리, 경찰서, 음식가게시설



#### Model with SHAP selection

One-Hot Encoding
MinMax Scaling
OPTUNA



버스승객수, 지하철역거리, 1인가구수, 다세대주택, 경찰서, 인구밀도, 아파트, 기초생활수급자, 음식가게시설, 의료시설, 직장인구, 숙박시설

RMSE = 937.629



## LightGBM

#### **Full Model**

One-Hot Encoding
MinMax Scaling
OPTUNA



**ALL Features** 

RMSE = 965.317

#### **Model with Best Subset Selection**

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 인구밀도, 총생활인구수, 숙박시설, 주차장, 아파트, 지하철역거리, 경찰서, 음식가게시설

RMSE = 885.492

#### Model with SHAP selection

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 총생활인구수, 도서관, 학교, 지하철역거리, 버스승객수, 다세대주택, 경찰서, 복지센터, 행정기관

RMSE = 996.960



## 최종 모델 선정

#### **XGBoost with Best Subset Selection**

One-Hot Encoding
MinMax Scaling
OPTUNA



1인가구수, 인구밀도, 총생활인구수, 숙박시설, 주차장, 아파트, 지하철역거리, 경찰서, 음식가게시설



RMSE = 805.743

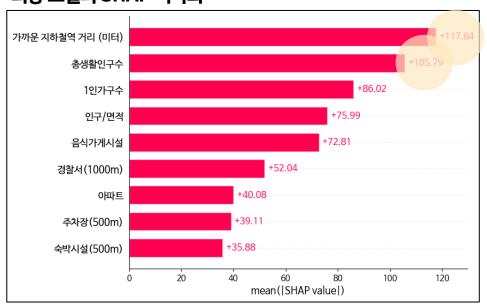
(축하의 댄스)





## 최종 모델 선정

#### 최종 모델의 SHAP 시각화



최단 지하철역 거리와 총생활인구수가 가장 강한 영향력을 보이는 것을 확인





1주차 때 사용한

클러스터링 변수와 동일





## 재배치 대상 택배함 선정



재배치 대상 개수

다른 공공시설물 재배치 사례를 살펴본 결과, 대체적으로 전체의 10% 가량 재배치되는 것을 확인



총 245개의 안심택배함 중 10%인

24개의 이용건수가 낮은 안심택배함을

재배치 대상으로 선정



재배치 장소의 중요성

더욱 효율적일 것으로 판단되는 장소로 재배치하여야 하므로 후보지를 잘 선정하는 것이 중요



하나의 기준으로 재배치 장소를 선택하는 것은 위험하다 판단하여 각기 다른 기준으로 **두 개**의 후보지를 선정하기로 결정

## 재배치 장소

#### 후보 1

기존 장<mark>소와 가장</mark> 가까운 지하철역

가장 가까<mark>운 지하철역은 기존 사용자들의</mark> 편의도를 고<mark>려하기 위해 첫번째 후</mark>보지로 설정

#### 후보2

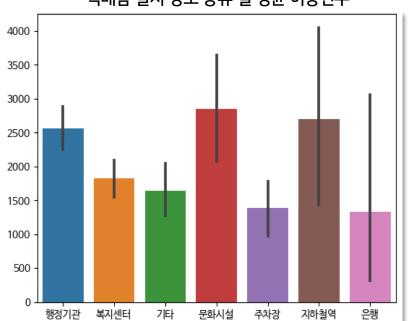
기존 장소 행정동 내의 공공시설 한 곳

안심택배함은 지자체에서 관리하여 대부분 공공시설에 설치되어 있으며, 공공시설 외 사유지에 설치할 경우 추가적인 절차가 필요

#### 

## 재배치 장소: 후보 2

택배함 설치 장소 종류 별 평균 이용건수



장소 유형별 평균 이용건수를 살펴본 결과,지하철역을 제외하여 **동일 행정동 내의 행정기관/문화시설**을 고려

편차가 더 작은 **행정기관**을 두번째 후보로 선정

만일 이미 행정기관에 배치되어 있는 경우, **문화시설을** 후보로 선정

배치가능한 문화시설 마저 미존재할 경우, **복지센터를** 후보로 선정

## 재배치 장소

No.	시설명	장소유형	행정동		최단지하철역거리	
1	효창공원역	지하철역	효창동		0	
2	독산3동주민센터	행정기관	독산3동	•••	1100	
3	독산역	지하철역	독산3동		0	
	<u>:</u>					
46	상계1동주민센터	행정기관	상계1동		102	
47	석계역	지하철역	장위3동		0	
48	장위3동주민센터	행정기관	장위3동		592	

앞서 언급한 기준을 바탕으로 총 48행의 최종 Test 데이터 셋 완성



## 예측 결과

#### 재배치 후보 결과 비교



No.	No. 시설명	
1	효창종합사회복지관	
2	2 청춘삘딩	
:		
23	상계문화정보도서관	
24	성북청소년수련관	





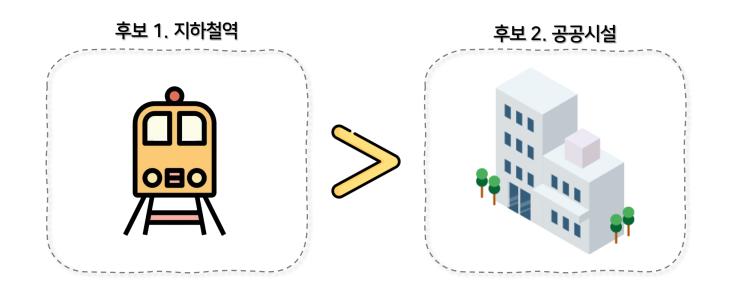
1 1. 1-1-			
시설명	예측값		
효창공원역	2351.60		
독산역	2424.24		
:			
수락산역	2237.64		
석계역	2444.88		



시설명	예측값
용산수시설관리공단	1995.78
독산3동주민센터	2135.91
:	
상계1동주민센터	2174.39
장위3동주민센터	2050.53

결과 비교했을 때, 모든 재배치 대상 안심택배함에서 후보 1(지하철역)이 후보 2(공공시설)보다 예측 이용건수가 높게 나오며, 후보 1(지하철역)이 후보 2(공공시설)보다 평균이 높은 것을 확인

## 결론 및 정책제언



분석 결과 기존의 택배함 위치보다 EDA과정을 통해 도출해낸 입지인 지하철역과 가까운 곳에 설치하는 것이 이용건수를 높이는 방법으로 보임



## 결론 및 정책제언

#### 서울시 안심택배함 과업지시서 中

- 3. 설치 입지조사 및 안심택배함 설치
- 가. 일반주택가 지역
- 싱글여성, 맞벌이가구 등이 많은 원룸, 빌라, 다가구주택 지역 등
- 택배물량이 많이 유통되는 지역 중 임시보관소(편의점, 슈퍼 등)의 보관 물량이 많은 지역
- 기존 공동주택, 기숙사 등 경비원 상주시설은 입지요건에서 제외
- 나, 유동인구가 많고, 주거지에서 접근이 용이한 지역
- 유동인구가 많고 교통이 편리한 상권
- 재래시장, 대학가 등과 밀접한 주거지역
- 공단 등 근로자 밀집 거주지역
- 여성안심주택, 임대주택 등
- 기타 택배수요가 많을 것으로 예상되는 지역
- 다. 보안 및 편의시설 이용이 가능한 지역
- 방범 cctv 설치 구역, 주차장 이용이 편리한 지역, 시설관리자 상주지역 등
- 눈·비 가리 및 전력·통신 장비설치가 용이한 지역(필요시 캐노피 설치 보완 등)

하지만 이미 과업지시서 상에서 접근성과 유동인구를 주요 입지조건으로 고려 중

#### 출처 - 머니투데이

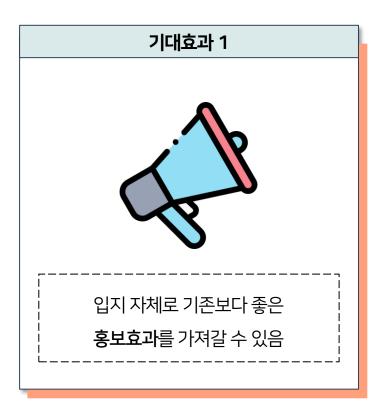
"안심택배함이 눈에 잘 보이는 곳에 있어야 많이 사용하고 그 수도 많아질텐데 아쉽다"고 했다.

1인가구인 김모씨(36)는 "집에서 가장 근처에 있는 안심택배함이 도보로 10분 넘는 거리에 있어 이용하기 어렵다"며 "집 근처에 없더라도 지하철역처럼 접근성이 좋은 곳에 있다면 꼭 사용하고 싶다"고 했다.

접근성 즉 유동인구도 중요하지만 **잘 보이는 곳**에 설치하는 것 역시 중요해 보임

주요 입지 요건으로 **눈에 잘 띄는 입지**를 고려하는 것이 **이용건수를 높이는데** 주요할 것으로 보임

## 기대효과



# 기대효과 2 높아진 인지도로 인해 더욱 많은 사람이 이용하게 되는 **선순환** 발생



## 의의 및 한계

#### 연구 의의

 ○ 각 택배함의 입지를 파악할 수 있는 다양한 변수를 활용하여 분석을 진행

 ○ 행정동 단위의 분석에서 벗어나 택배함 단위의 분석을 진행

 ○ 여론에 부합하는 안심택배함의 이용을 활성화하기 위한 인사이트 도출

#### 연구 한계

✓ 정량화된 데이터만을 활용하여 정성적인 측면을 고려하지 못함
 ✓ 이용건수에만 집중하여 안심택배함의 다른 기능에는 초점을 두지 못함
 ✓ 장소 유형보다 더욱 세세한 입지 조건의 데이터를 활용하지 못함