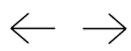


TEAM.5MYGOE\_PORTFOLIO



# 취약 요인 별 화재발생의 다면성



참여자

민경원

박정원

전다인

조희은

## ▼ 취약 요인 별 화재발생의 다면성



INDEX

### 1. 분석 배경

- ▼ 기획 배경
- ▼ 각자의 가설

### 5. 최종 결론

- ▼ 결론
- ▼ 기대효과
- ▼ 한계점 및 보완점
- ▼ 기타 정보

### 2. 데이터 가공

- ▼ 데이터 선정
- ▼ 데이터 사용 컬럼
- ▼ 데이터 전처리
- ▼ 사용 라이브러리

### 3. 가설

- ▼ 가설\_1
  - ▼ 시각화
  - ▼ 가설 검증
- ▼ 가설\_2
  - ▼ 시각화
  - ▼ 가설 검증
- ▼ 가설\_3
  - ▼ 시각화
  - ▼ 가설 검증

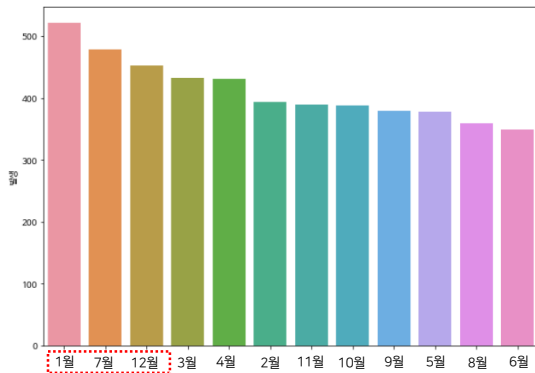
### 4. 머신러닝

- ▼ EDA
- ▼ DecisionTree
- ▼ RandomForest

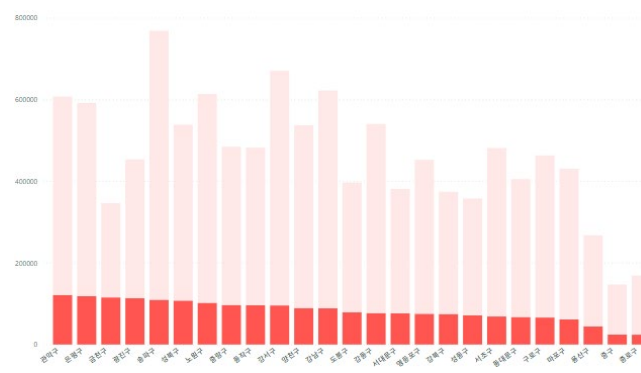
## ▼ 1. 분석 배경

### ▼ 기획 배경

서울시 월별 화재발생 비율



119안전센터 1개소당 담당 인구와 자치구 인구 비교

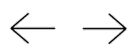


2021년 소방재난본부 주요업무계획의 상반기 중점사업

1	'시민이 '안심'하고 생활할 수 있는 신속한 대응	3
1.	재난을 압도하기 위한 선제적 과잉대응	3
2.	빈틈없는 상황관리 및 취약시설 관리 강화	4
3.	감염병 환자 이송 등 대응체계 강화	5
4.	동절기·해빙기 신속한 재난수습을 위한 유관기관 협치	6
5.	부모의 아동학대 방지를 위한 소방안전대책	7

- 서울시 월별 화재 발생 비율에 따르면 '1월' '7월' '12월'로 예상 밖에 여름에 화재 발생 비율이 높다는 것을 볼 수 있음
- 119안전센터 1개 소당 담당 인구와 자치구 인구를 비교했을 때 '관악구'가 높은 비율을 보였으나 송파구가 인구수와 소방서당 담당 인구가 많은 것을 확인 할 수 있음
- 2021년 소방재난본부 주요업무계획을 살펴보면 재해 발생시 대응 및 복구에 초점을 맞추는 방향성을 취하고 있음

시각화를 통해 겨울과 여름에 화재 발생 빈도가 높고 하나의 소방서에 담당 인구도 자치구마다 다른 만큼 시기나 구마다 각기 다른 소방 정책을 시행해야 된다고 생각 또한 예방이나 대비에 철저하면 적은 예산으로도 더 큰 효과를 나타낼 수 있기에 본문에서 강조하고 있는 취약시설 관리에서 아이디어를 얻어 지역의 화재발생성을 예측하고 소방 자원을 최대한으로 활용할 수 있도록 화재취약지도 제작 (각 팀원이 화재취약요인을 선정)



## ▼ 1. 분석 배경

### ▼ 각자의 가설



DEAR . 조 희 은

환경에 따라 화재 원인과 빈도가 다를  
것이기에 계절별 대책이 필요할 것이다.



DEAR . 박 정 원

소방용수시설 및 소화 비상벨 설치로  
초기 진압이 가능 할 것이기에 추가적인  
배포로 화재 피해를 줄일 수 있을 것이다.



DEAR . 민 경 원

화재 발생의 피해 정도가 소방서 위치에 따라  
피해액이 높을 것이다.





DEAR . 전 다 인



주택의 종류나 연식에 따라 화재 취약도가  
다르기에 취약 계층은 화재에 더 쉽게  
노출되고 이후 피해에도 크게 영향을  
받는다.

## ▼ 2. 데이터 가공



### ▼ 데이터 선정 및 사용 컬럼

 소화전\_비상소화장치\_통계.xls 



[기간], [자치구], [화재건수합계], [이재가구수(가구)], [면적(m²)],  
[부동산피해합계], [평단가], [인명피해합계], [인명피해비율]

 서울시화재발생현황(구별)통계.xls 



[자치구], [소화전], [비상소화장치]

 서울시\_노후기간별\_주택현황\_통계.xlsx 



[자치구], [단독주택], [아파트], [연립주택], [다세대주택]

 서울시 자치구별 고령인구 (추계인구) 통계.xls 



[지역], [기간], [합계], [남자], [여자]

 서울시 도시가스 이용현황 (용도별동별) 통계.xls 



[자치구], [동], [도시가스이용가구], [도시가스이용가수 순위]

 서울시 국민기초생활보장 수급자 (구별) 통계.xls 



[자치구], [동], [총 수급자], [총 수급자 순위]

 서울시 소방서관할 위치정보 (좌표계\_WGS1984).csv 

[서소이름], [위도], [경도]

 서울시 119 안전센터 1개소당 시민수 통계.xls 

[자치구], [인구], [119안전센터], [119안전센터 1개소당 담당인구]

 서울시 인구밀도 (동별) 통계.xls 

[기간], [자치구], [동], [인구], [면적], [인구밀도]

## ▼ 2. 데이터 가공

### ▼ 데이터 전처리

2021년 화재발생 현황																
기간	자치구	발생	발생.1	발생.2	발생.3	소실	소실.1	소실.2	피해액	피해액.1	피해액.2	재산피해경감액	인명피해	인명피해.1	인명피해.2	
0	기간	자치구	합계	실화	방화	기타	동수(동)	이재가구수(가구)	면적(m <sup>2</sup> )	합계	부동산	동산	재산피해경감액	합계	사망	부상
1	2021	합계	4951	4399	87	465	468	282	21723	18268346	6872172	11396174	336039670	317	37	280
2	2021	종로구	192	169	1	22	-	1	390	465499	169588	295911	21075683	12	-	12
3	2021	중구	171	136	3	32	47	6	1350	2780374	686265	2094109	19184807	16	-	16
4	2021	용산구	176	160	1	15	6	40	422	296793	123439	173354	3767112	5	-	5

- 0번째 행을 컬럼명으로 올리고 0번째 행 삭제
  - 데이터전처리 합계 행 제거
  - 합계라는 동일한 컬럼명이 많아서 컬럼 이름 수정
  - 인명피해 비율 넣고 컬럼 '인명피해비율' 생성
- ▶ (인명피해 합계 / 화재건수 합계 \* 100)

기간		자치구	화재건수합계		실화	방화	기타	동수(동)	이재가구수(가구)	면적(㎡)	부동산피해합계	부동산	동산	재산피해경감액	인명피해합계	사망	부상	인명피해비율
2	2021	종로구	192	169	1	22	0	1	390	390	465499	169588	295911	21075683	12	0	12	6.25
3	2021	중구	171	136	3	32	47	6	1350	1350	2780374	686265	2094109	19184807	16	0	16	9.36
4	2021	용산구	176	160	1	15	6	40	422	422	296793	123439	173354	3767112	5	0	5	2.84
5	2021	성동구	191	163	2	26	14	9	720	720	458412	156128	302284	10949317	5	0	5	2.62
6	2021	광진구	205	187	3	15	33	6	554	554	889681	184833	704848	5924256	12	2	10	5.85

주거환경에 대한 화재 취약도를 나타내기 위해 취약도에 대한 컬럼을 선정 ▶

각 수치데이터를 범주화 데이터로 바꾸기 위해 데이터 describe를 통해 통계를 조회 ▶ 점수를 5단계로 구분되기 위해 `percent = [0.20, 0.40, 0.60, 0.80]` ▶

`nohu.describe(percentiles = percent)` 써서 min,20,40,60,80,max로 나누어 구분 기준을 정함 ▶ 범주화 데이터로 바꿈 ▶ 선정된 데이터 모두 진행

`nohu["단독주택"].describe()`

```
count    25.000000
mean     10935.680000
std      4147.358564
min      4541.000000
25%      7869.000000
50%     10855.000000
75%     13166.000000
max     20556.000000
```

단독주택20    단독주택30    단독주택

count	25.000000	25.000000	25.000000
mean	6629.400000	4306.280000	10935.680000
std	2715.183803	1775.026796	4147.358564
min	2093.000000	1464.000000	4541.000000
20%	4170.800000	3100.800000	7402.800000
40%	5605.000000	3707.200000	9593.200000
50%	6487.000000	3919.000000	10855.000000
60%	7700.400000	4338.000000	11878.000000
80%	8654.600000	5817.400000	13848.400000
max	13598.000000	8248.000000	20556.000000

## ▼ 2. 데이터 가공

### ▼ 데이터 전처리

#### X 서울시 화재발생 현황 (구별) 통계.xls 전처리

['기간', '자치구', '발생', '발생.1', '발생.2', '발생.3', '소실', '소실.1', '소실.2', '피해액', '피해액.1', '피해액.2', '재산피해경감액', '인명피해', '인명피해.1', '인명피해.2', '이재민수', '구조인원']

	자치구	화재건수	실화	방화	기타	피해액	부동산	동산	인명피해	사망	부상	이재민수	구조인원
24	강남구	391	356	6	29	1354949	612282	742667	17	2	15	35	36
25	송파구	263	237	4	22	788289	228466	559823	6	1	5	7	19

변수(f\_gu.columns)를 지정하여 쓰는 컬럼만 따로 추출하고 중복 이름 바꿈

#### 서울시\_노후기간별\_주택현황\_통계.xlsx 전처리

20년 이상 30년 미만에 대한 주택의 분류와 노후도 30년 이상의 주택 분류되어 있는 데이터를 '20' 컬럼과 '30'컬럼을 새로 생성하고 모두 더해 input함  
그래서 '20'과 '30'을 더해 하나의 변수에 저장하고 컬럼화 함

```
ob['20'] = ob['단독주택20']+ob['아파트20']+ob['연립주택20']+ob['다세대주택20']+ob['비거주용 건물내 주택20']
ob['30'] = ob['단독주택30']+ob['아파트30']+ob['연립주택30']+ob['다세대주택30']+ob['비거주용 건물내 주택30']
ob['20_30'] = ob['20']+ob['30']
```

### ▼ 사용 라이브러리

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

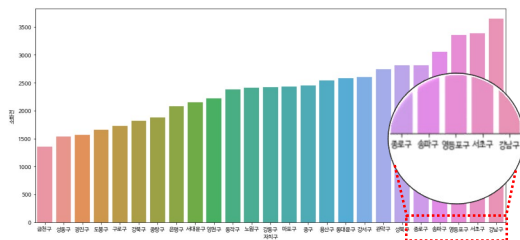
import folium
from folium import plugins
```

로고도 있음

## ▼ 2. 가설\_1 박정원 소방용수시설 및 소화 비상벨 설치로 초기 진압이 가능 할 것이기에 추가적인 배포로 화재 피해를 줄일 수 있을 것이다.

### ▼ 데이터 시각화

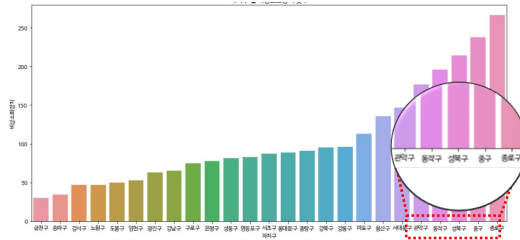
2018년 서울시 자치구별 소화기 배치 개수



◀ 강남구 > 서초구 > 영등포구 > 송파구 > 중로구 순으로 많은 소화기 개수 확인

주택보다 건물이 많은 곳이 소화기 개수가 많다는 것을 확인

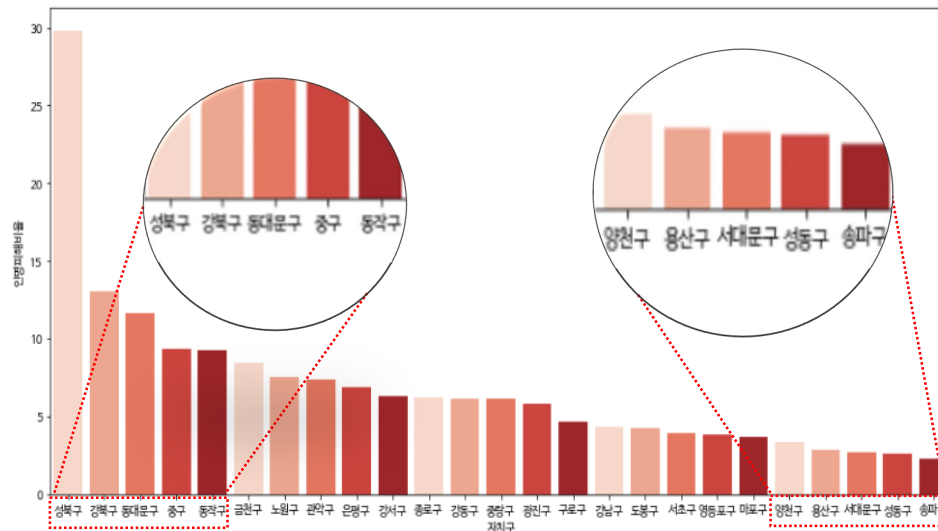
서울시 자치구별 비상소화장치 배치 개수



◀ 종로구 > 중구 > 성북구 > 동작구 > 관악구 > 순으로 비상소화장치 개수 확인

인구가 많은 자치구가 비상소화장치가 많은 곳이 많다는 것을 확인

2021년 서울시 자치구 별 화재발생 시 인명 피해율



◀ 송파구 > 성동구 > 서대문구 > 용산구 > 양천구 순으로 화재발생시 인명피해가 가장 적었다.

◀ 성북구 > 강북구 > 동대문구 > 중구 > 동작구 순으로 화재발생시 인명피해가 많이 발생된 것 확인

송파구를 제외한 자치구 인구가 상대적으로 적은 곳이 인명피해가 적게 일어나고 성북구는 인명피해율이 이상치를 보일 정도로 압도적인 수치를 확인

### ▼ 가설 검증

비상소화장치와 소화기 등 소방용수시설이 설치 개수가 많은 곳 자치구와 자치구별 화재발생시 인명 피해율의 상관관계를 분석

그러나 시설이 많으면 인명피해가 적게 나타날 것으로 예상했지만 소화기 배치 개수가 많은 송파구를 제외한 나머지 자치구인 **성동구, 서대문구, 용산구, 양천구**는 소화기 개수가 전체 평균 범위에 속하고 비상소화장치는 하위에 속하여 소방시설과 화재발생시 인명피해 상관관계는 뚜렷하게 나타나지 않음



### ▼ 3. 가설\_2 전다인 조회은 주거환경에 따라 화재 취약도와 때문에 부주의로 인한 화재 지역구와 상관관계가 있을 것이다.

#### ▼ 데이터 시각화

주거환경에 취약부분을 선정하여 범위 별 점수화하여 상위 10개 자치구만 추출

	자치구	단독주택	아파트	연립주택	다세대주택	도시가스 이용 가구	기초 수급자	고령자	합계
1	관악구	5	2	5	5	5	4	5	31
2	강남구	4	5	3	3	5	5	5	30
3	송파구	2	5	4	4	5	3	5	28
4	중랑구	5	4	3	2	3	5	3	25
5	서대문구	3	3	4	4	2	4	5	25
6	마포구	3	3	2	4	4	2	3	21
7	서초구	1	4	5	3	2	2	3	20
8	강동구	2	3	2	2	3	3	4	19
9	용산구	4	2	2	3	1	2	1	15
10	종로구	3	1	4	2	1	1	1	13

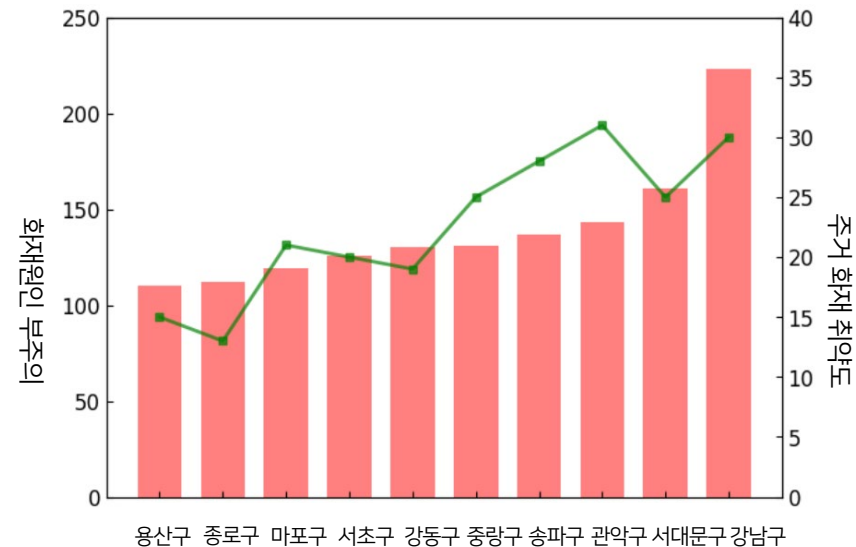
단독주택, 아파트, 연립주택, 다세대 주택 : 서울시 노후기간 별 주택현황 데이터

연식 20년 이상의 주택이 많음5~적음1

도시가스 이용 가구 : 도시가스 이용현황 데이터를 통해 도시가스 이용 가구가 많음5~적음1

기초 수급자, 고령자 : 서울시 기초 수급자 데이터와 고령인구 데이터 통해 많음5~적음1

화재 취약도와 화재원인 부주의에 대한 자치구별 수치



#### ▼ 가설 검증

주택의 노후도와 화재위험성이 높은 도시가스 이용 가수, 화재발생시 인명피해로 이어질 수 있는 고령자, 기초생활 수급자가 많은 지역구를 확인하고 점수를 부여하여 환산하고 합산하여 화재원인 중 부주의로 인한 통계를 따로 정제하여 그래프로 비교한 결과 '화재 취약도가 높을 수록 부주의로 인한 화재발생 수치도 비슷하거나 높아진다'는 것을 확인할 수 있음

## ▼ 4. 가설\_3

전다인

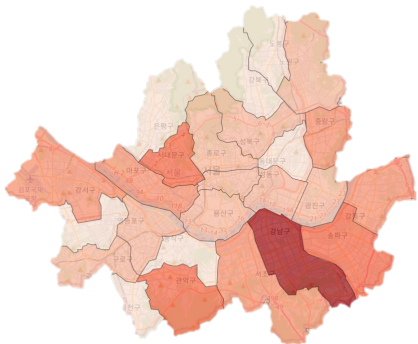
민경원

자치구별 화재 취약도에 대한 상관관계를 분석하고 소방서의 골든 타임 안에 들어가는지 확인하고 화재 피해지역과 비교

### ▼ 데이터 시각화

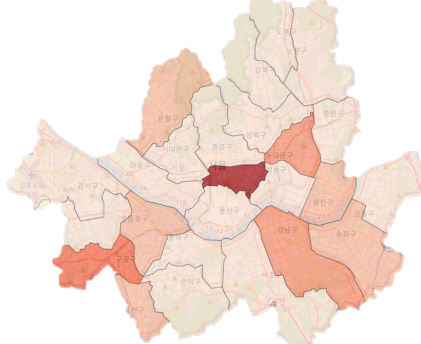
화재 발생 건수

강남구>송파구>서대문구>관악구>마포구



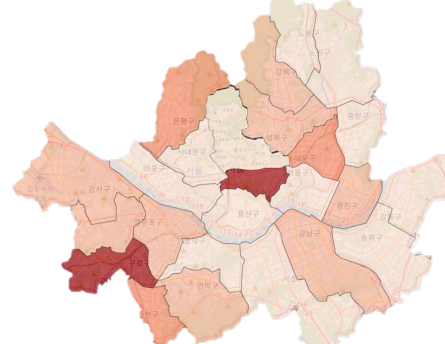
재산 피해액

중구>구로구>강남구>동대문구>영등포구



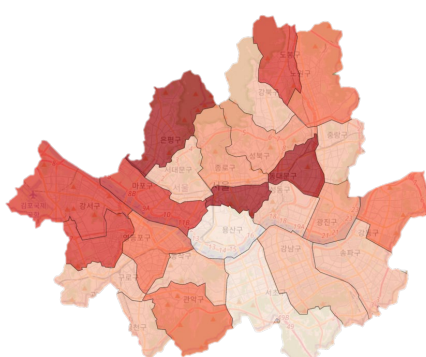
평단가 적용 재산 피해액

중구>구로구>은평구>동대문구>금천구



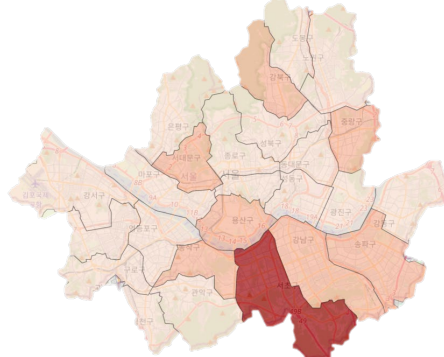
인구 밀도

양천구>동작구>중랑구>구로구>광진구



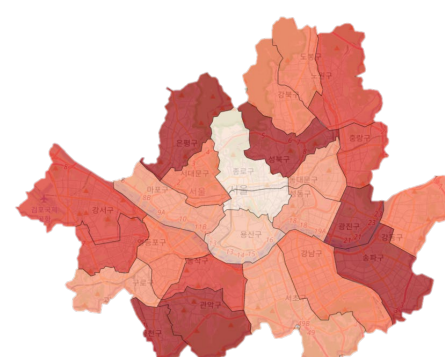
자치구 면적

서초구>강서구>강남구>노원구>송파구



1개 소방서당 담당 인구

관악구>은평구>금천구>광진구>송파구



화재 취약도를 분석

6가지로 기준을 나누어 시각화한 결과 '강남구'가 가장 높은 순위 차지 '구로구, 송파구, 강서구, 은평구'가 비슷한 수치로 그 뒤를 차지

서울시 자치구 중 상위 5개 자치구가 화재 취약도가 높게 나옴을 알 수 있음

## ▼ 4. 가설\_3

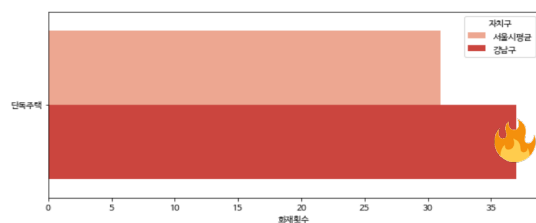
전다인

민경원

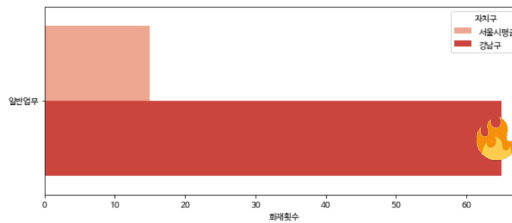
자치구별 화재 취약도에 대해 분석하고 소방서와 안전센터 개수 대한 화재발생과 비교

### ▼ 데이터 시각화

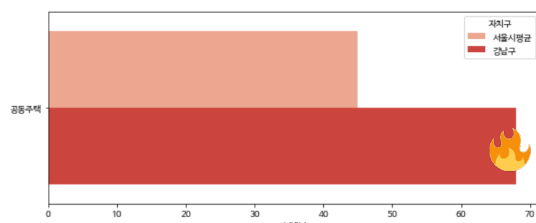
서울시 평균과 강남구  
'단독주택' 화재 횟수 비교



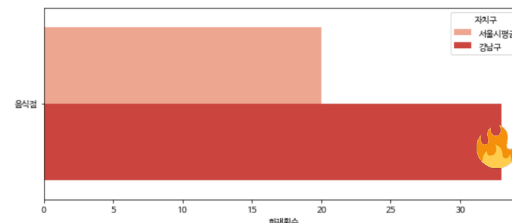
서울시 평균과 강남구  
'일반업무' 화재 횟수 비교



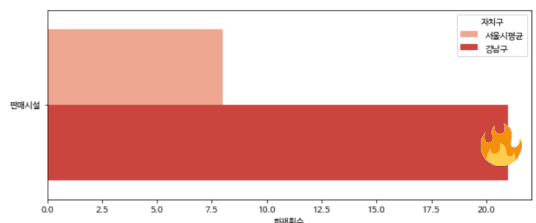
서울시 평균과 강남구  
'공동주택' 화재 횟수 비교



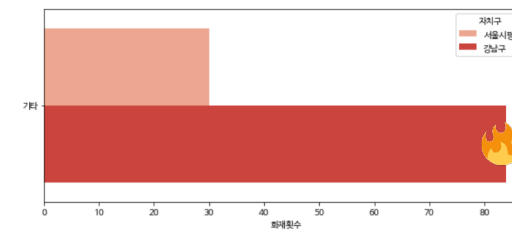
서울시 평균과 강남구  
'음식점' 화재 횟수 비교



서울시 평균과 강남구  
'판매시설' 화재 횟수 비교

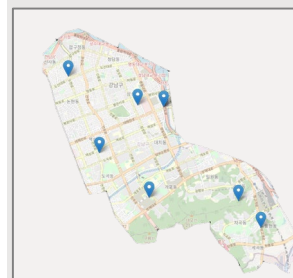


서울시 평균과 강남구  
'기타' 화재 횟수 비교



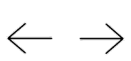
### ▼ 가설 검증

6가지 기준을 통한 화재 취약도를 통해 '강남구'가 제일 높은 순위를 차지하여 서울시 장소 별 화재 평균 수치와 강남구 수치와 비교하여 더 자세히 분석  
주택 뿐만 아니라 업무시설과 음식점, 판매시설에서도 서울시 화재 평균 수치를 웃도는 것을 확인



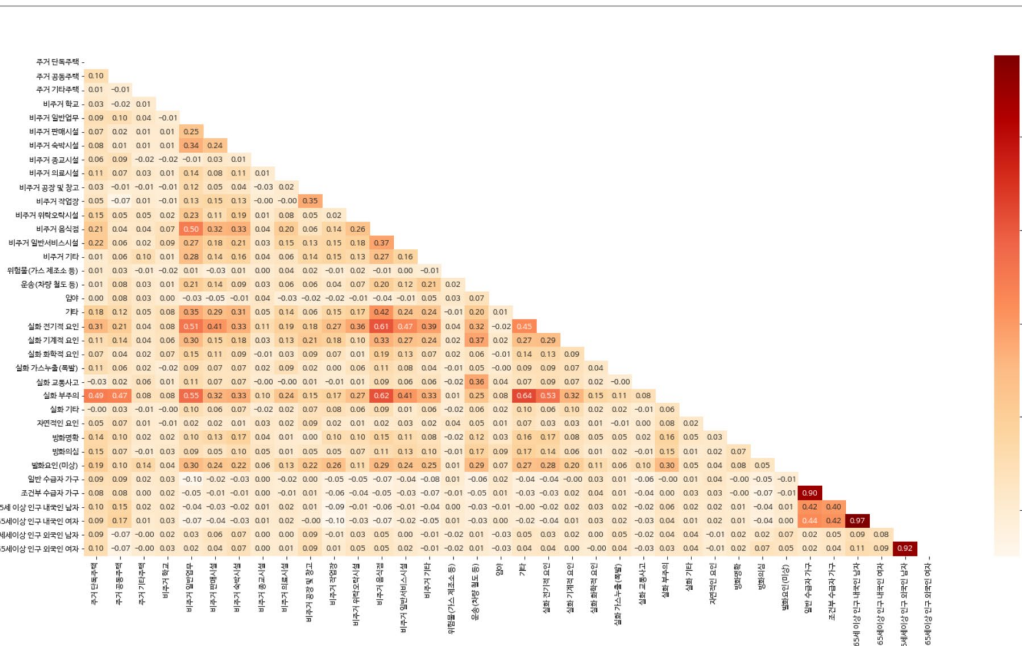
강남구는 소방서 및 안전 센터가 다른 자치구에 비해 고루 분포되어 있어 안전센터&소방서의 개수와 화재의 횟수는 상관관계가 없는 걸로 확인

강남구의 경우를 통해 안전 센터와 소화기 설치가 많은 것은 화재 발생시 도움이 될 수 있지만 화재 자체를 줄이는 방법이 되지 못한다는 것을 보여줌



## ▼ 4. 머신러닝

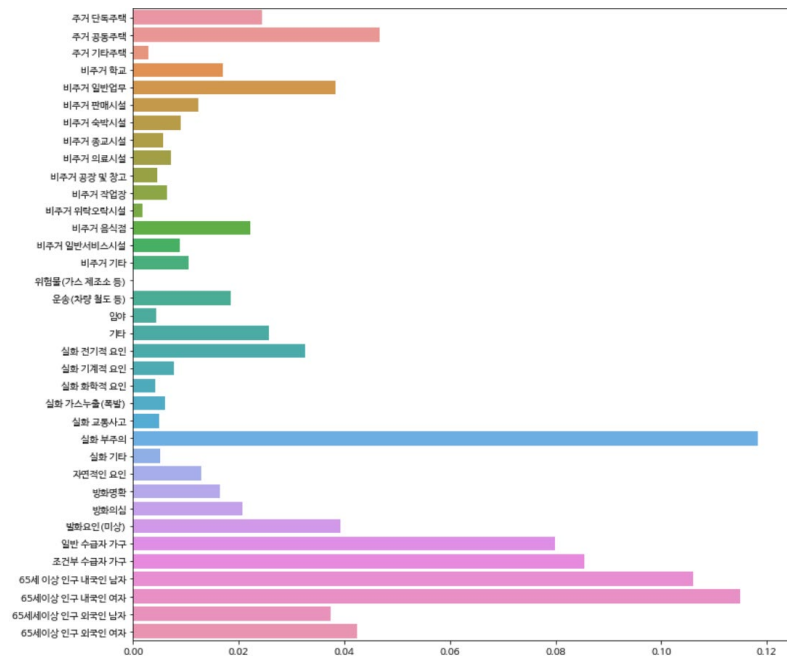
### ▼ EDA



히트맵을 구현하여 화재발생 시 인명피해가 난 원인 관련성이 높은 부분을 확인

사람으로는 65세 이상 남자와 외국인, 그리고 기초수급대상자가 높은 관련성 보임

장소로는 비주거음식점과 일반업무 건물이 높은 관련성을 보임



실제 막대그래프를 통해 화재로 인한 인명피해가 난 원인으로 '실화 부주의'가 가장 높았으며 인구는 히트맵 분석과 동일하게 나타났다.

## ▼ 4. 머신러닝

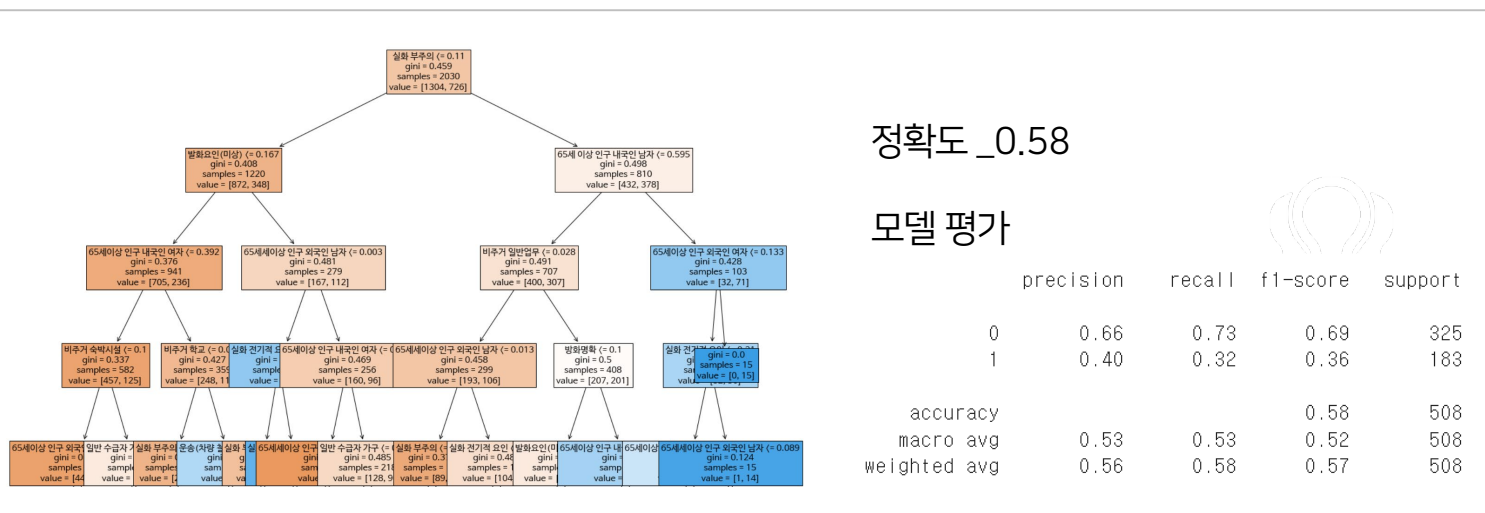
### ▼ Data Shape

train.shape    test.shape  
(2538, 36)    (2538, 1)

x\_train.shape,   x\_valid.shape,   y\_train.shape,   y\_valid.shape  
(2030, 36)    (508, 36)    (2030, 1)    (508, 1)

### ▼ DecisionTree

어떤 항목에 대한 관측값과 목표값을 연결시켜주는 예측 모델



### 모델 설명

-Accuracy  
올바르게 예측된 데이터의 수를 전체 데이터의 수로 나눈 값

-Recall  
실제로 True인 데이터를 모델이 True라고 인식한 데이터의 수

-Precision  
모델이 True로 예측한 데이터 중 실제로 True인 데이터의 수

F1 score  
precision 과 recall의 조화평균

## ▼ 4. 머신러닝

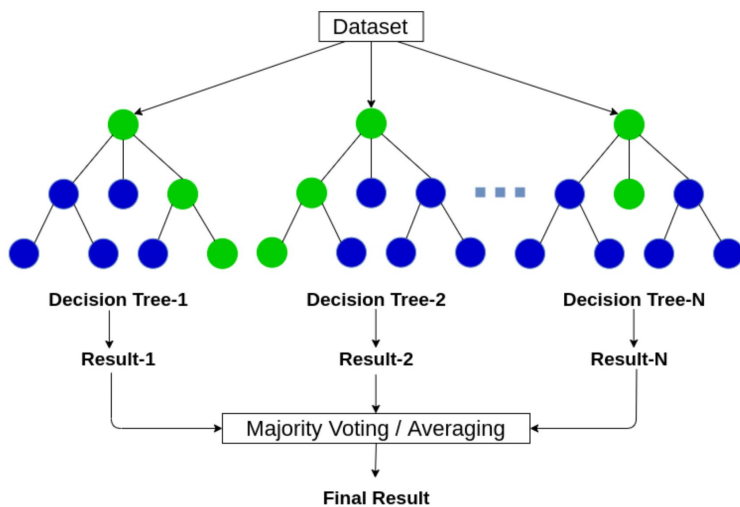
### ▼ Data Shape

train.shape    test.shape  
(2538, 36)    (2538, 1)

x\_train.shape, x\_valid.shape, y\_train.shape, y\_valid.shape  
(2030, 36)    (508, 36)    (2030, 1)    (508, 1)

### ▼ Random Forest

오버피팅을 방지하기 위해, 최적의 기준 변수를 랜덤 선택하는 머신러닝 기법



정확도 \_0.63

모델 평가

	precision	recall	f1-score	support
0	0.67	0.85	0.75	325
1	0.49	0.26	0.34	183
accuracy			0.64	508
macro avg	0.58	0.55	0.54	508
weighted avg	0.61	0.64	0.60	508

### 모델 설명

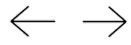
-Accuracy  
올바르게 예측된 데이터의 수를 전체 데이터의 수로 나눈 값

-Recall  
실제로 True인 데이터를 모델이 True라고 인식한 데이터의 수

-Precision  
모델이 True로 예측한 데이터 중 실제로 True인 데이터 수

F1 score  
precision 과 recall의 조화평균





TEAM. 5MYGOD



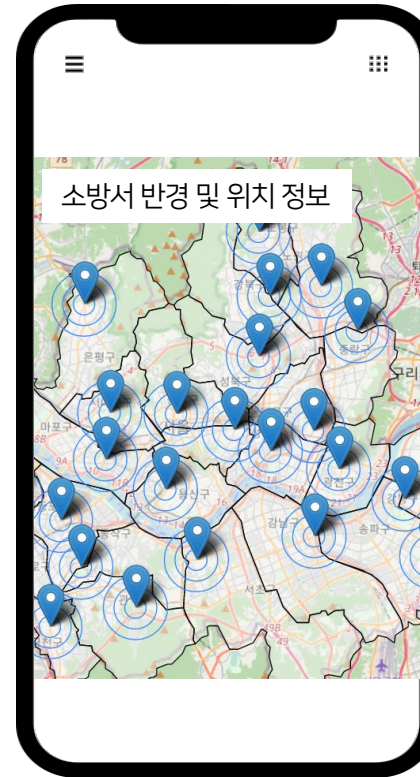
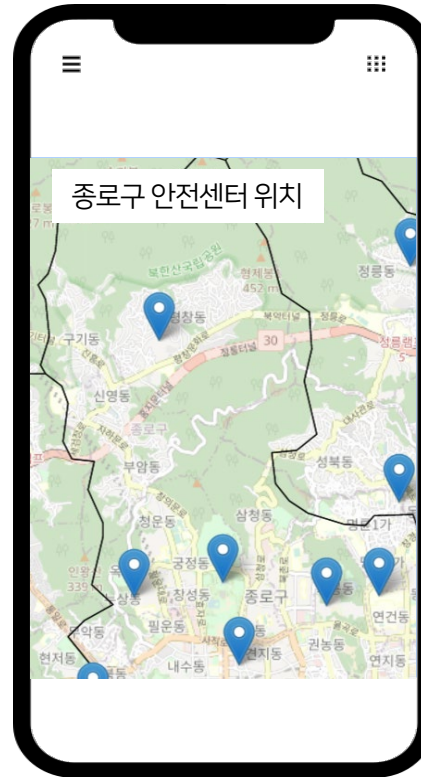
## ▼ 5. 📄 총정리

### ▼ 결론(정보\_시각화)

가설2의 주거환경에 대한 취약도와 가설3의 장소별 화재 취약도 부분 모두 강남구 1위

예상 외로 고령인구, 기초생활 수급자, 가스사용 인구도 자치구 중 높은 비율을 차지  
특히 노후화된 아파트의 비율이 높아 제대로 된 아파트 자체에 소화 시설이 없고 오래된 가스관 등 다른 자치구보다 화재 발생시 피해 금액이 클 수 밖에 없음

국민비서 '구뵐' 처럼 강남구 자체 알람서비스를 만들어 강남구 주민에게 화재 발생 전 다양한 위치 정보를 선제적 발송을 통해 재난 전에도 인지 할 수 있도록 함



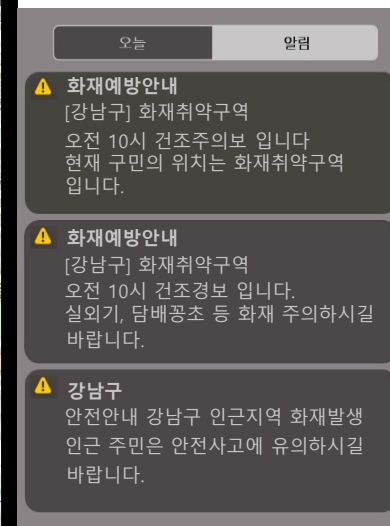
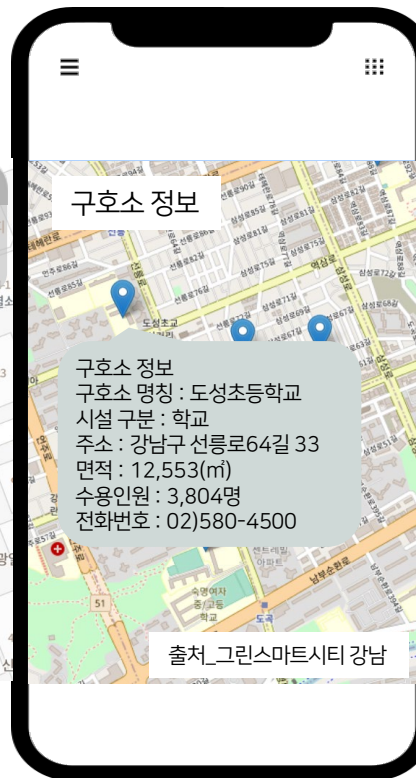
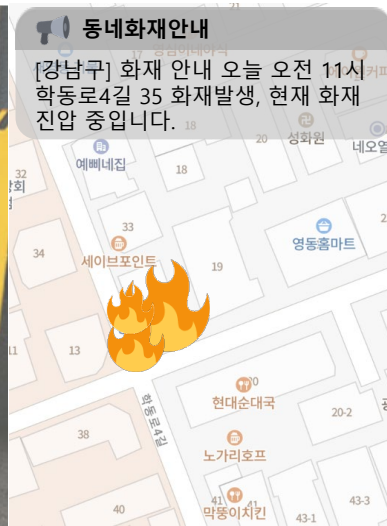
## ▼ 5. 결과

### ▼ 결론(정보\_시각화)

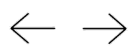
소방차 전용 주차 구역 알림

동네 화재 발생시 알림을 통해  
구민에게 안전사고 문자 발송

화재 피해가 발생할 경우 주변  
구호소 위치 알림 서비스 제공







## ▼ 5. 총정리

### ▼ 기대효과

- (1) 화재 취약도와 환경의 상관관계를 분석하여 지역별 예방 및 대비책을 수립할 수 있다.
- (2) 화재 취약도와 화재 진압력의 상관관계를 분석하여 인구 및 관할 면적을 기준으로 단순 배치됐던 소방력을 보다 효율적으로 사용할 수 있다.
- (3) 취약지역을 체계적으로 정비하고 위험 요소를 줄여가며 궁극적으로 안전한 도시를 만들고자 하는 계획에 활용할 수 있고, 그 과정에서도 점검 기준으로서 작용할 수 있다.
- (4) 서울시 외의 지역이나 새로운 데이터에도 머신러닝 모델을 확장 적용할 수 있다.

### ▼ 한계점 및 보안점

- (1) 자치구 기준으로 분석을 진행하였기에 하위 분류인 동 기준으로는 추가적인 분석이 필요하다.
- (2) 2021년의 데이터를 위주로 분석을 진행하였기에 당장의 상황은 반영이 되지 않는다.
- (3) 도로 상황이나 불법 주차 등의 데이터는 빠져있어 모델과 실제 현상 간 오차가 발생할 수 있다.
- (4) 각 팀원이 제시한 가설을 기준으로 취약 요인을 설정하였기에 모델의 설명력이 다소 제한적일 수 있다.
- (5) 지역구별 화재 취약 요인에 대해서는 세부적인 분석이 필요하다

## ▼ 5. 최종 결론

### ▼ 기타 정보

#### 참고 논문

강영옥 (2004)	서울시 방재 지도 작성 방안 연구: 화재위험도를 중심으로. 서울시정개발연구원.
임현석 외 5명(2019)	SPSS를 활용한 화재 요인과 영향인자간의 연관성 분석. 한국방재학회논문집, 19(5), 103-112.
황지은 외 3명(2015)	지역 특성이 화재에 미치는 영향에 관한 연구. 3. 한국방재학회 학술대회논문집, 2015(1), 168-168.
최진만 외 2명 (2009)	국민소득과 화재발생 건수의 상관관계에 관한 연구. 대한안전경영과학회지, 11(4), 147-152.
황종아, 구자훈(2019)	서울시 취약계층 밀집지역 분포와 주거지 특성 분석 : 민간임대주택 거주 기초생활수급자를 중심으로. 국토연구, 102, 99-116.

#### 참고 사이트

스마트서울맵	<a href="https://map.seoul.go.kr/smgis2/">https://map.seoul.go.kr/smgis2/</a>	그린 스마트시티 강남	<a href="https://www.gangnam.go.kr/main.do">https://www.gangnam.go.kr/main.do</a>
서울시 빅데이터 캠퍼스	<a href="https://bigdata.seoul.go.kr/main.do">https://bigdata.seoul.go.kr/main.do</a>		
소방 안전 빅데이터 플랫폼	<a href="https://www.bigdata-119.kr/">https://www.bigdata-119.kr/</a>		
공공데이터 포털	<a href="https://www.data.go.kr/">https://www.data.go.kr/</a>		
통계지리정보서비스	<a href="https://sgis.kostat.go.kr/view/index">https://sgis.kostat.go.kr/view/index</a>		