

민원

실시간 민원 분석 및 외부요인에 따른 민원 수 예측

- 발표 일시: 2020년 12월 24일 (목)
- 김혜빈*, 신지수, 이상은, 최윤형, 최현정

목차



팀원 소개

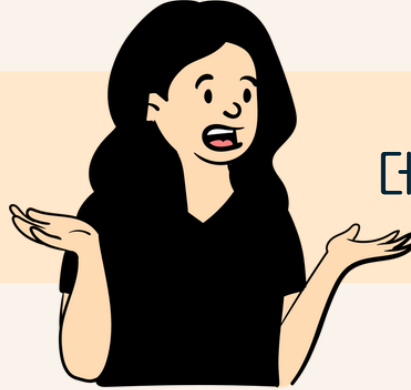
다양한 백그라운드

김혜빈(데이터정보학)



데이터 분석

이상은 (환경학)



데이터 분석/발표자료

최현정(컴퓨터공학)



데이터 분석/크롤링

최윤형(컴퓨터공학)



웹 구조 및 연동

신지수(전자공학)



웹 상세 페이지

프로젝트 계획

11일 이내에 하나의 웹 서비스를 구현해야 한다.

14	15	16	17	18	19	20	21	22	23	24
주제 변경 >										
		데이터 크롤링								
			서비스 설계							
			분석 계획							
			데이터 수집 및 전처리							
			웹 구조화 및 DB 연동							
				데이터 분석 및 시각화						
							웹 모듈 통합			
								사이트 빌드		+
									발표자료	+
										발표*

민원 데이터 활용

민원 접수와 피드백이 원활하게 진행되나?



실시간?

문제?

키워드?

주민이 행정기관에 대하여 원하는 바를 요구하는 일

주제 선정 이유

궁금증 네 가지

어떤 **집단**이 민원을 가장 많이 낼까?

민원 발생의 **흐름**은 어떠한가?

많이 나온 민원 **키워드**는 무엇일까?

외부요인 (코로나19확진자, 기온)이 민원 발생에 영향을 끼치지 않을까?

데이터 수집

2018년부터 2020년까지 (단위)



- 민원 데이터 API

- 1> 일일 발생량 (개)
- 2> 주간 상위
키워드 5 (단어)



- 기온 데이터

->전국 일 평균 (℃)



- 코로나 데이터

->일일 확진자 수 (명)



API Crawling

실시간 데이터 연동을 위한 API

일일 민원 건수 (회)

- 기간: 2018년, 2019년, 2020년
- 그룹: 10대부터 50대까지 | 남녀 각각

어떤 **집단**이 민원을 가장 많이 낼까?

민원 발생의 **흐름**은 어떨까?

주간 키워드 5

- 기간: 2019년, 2020년
- 그룹: 주간 상위 5개 분류기

많이 나온 민원 **키워드**는 무엇일까?

전국 일 평균 기온 (°C)

코로나 일 확진 수 (명)

외부요인이 민원 발생에 영향을 끼치지 않을까?

(1)

어떤 집단이 민원을 가장 많이 낼까?

민원 빈도가 많은 `성별`과 `연령`은 무엇일까?

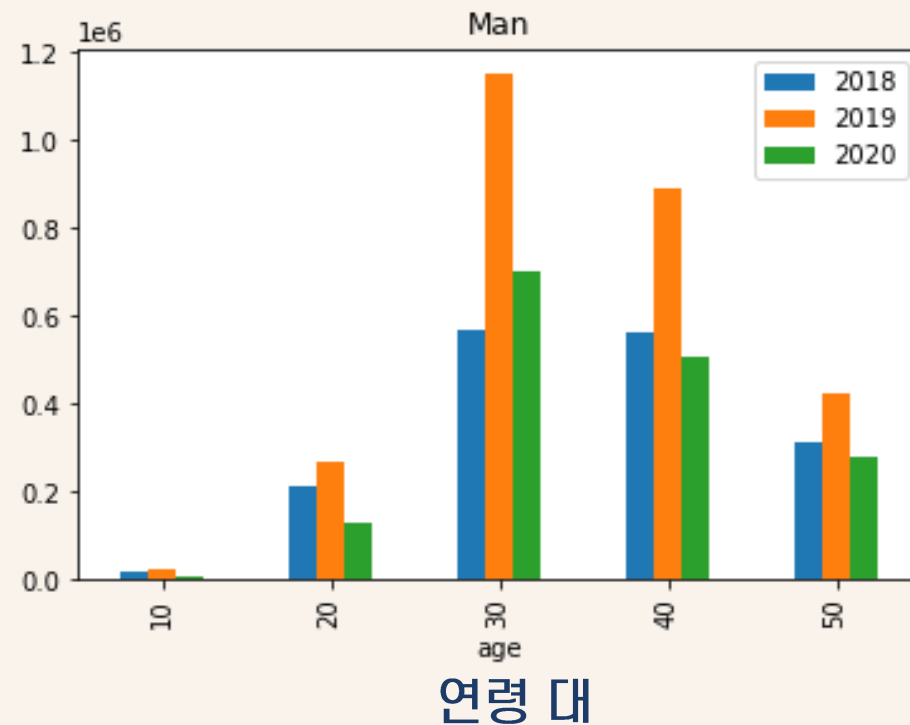
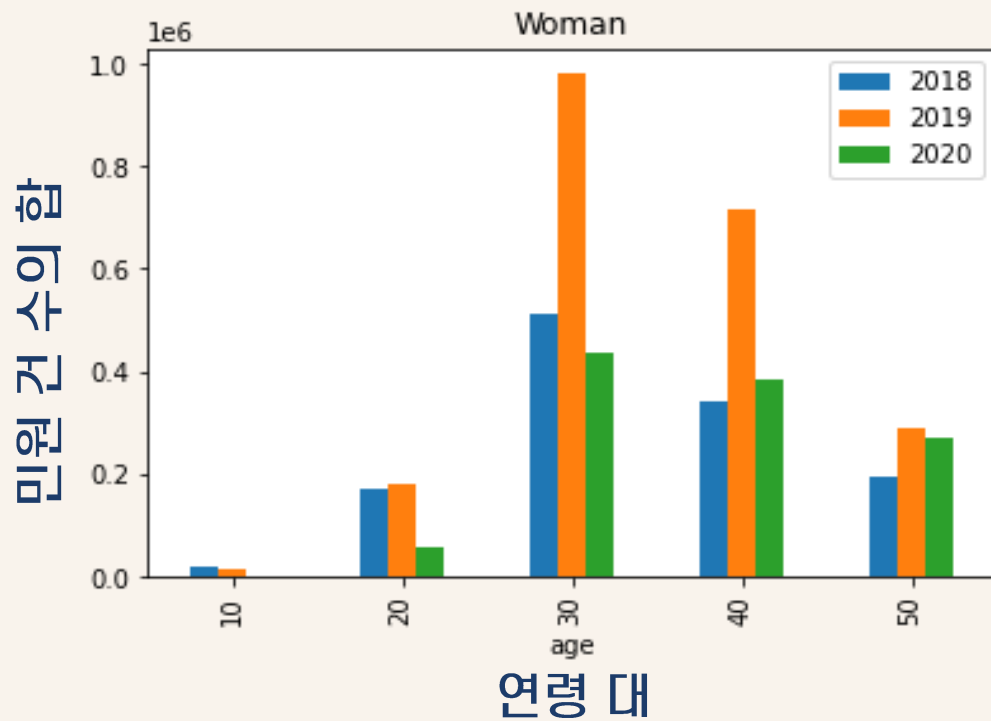
성별 | ? 성이 민원을 더 많이 낼 것이다.

어떤 집단이 민원을 가장 많이 낼까?

연령 | ? 연령에서 많이 나타날 것이다.

민원 데이터 집단 빈도

- 최근 3년 동안의 민원 데이터 총합계를 비교



성별 | (남)성이 민원을 더 많이 내었다.

연령 | (경제 활동) 연령에서 많이 나타났다.

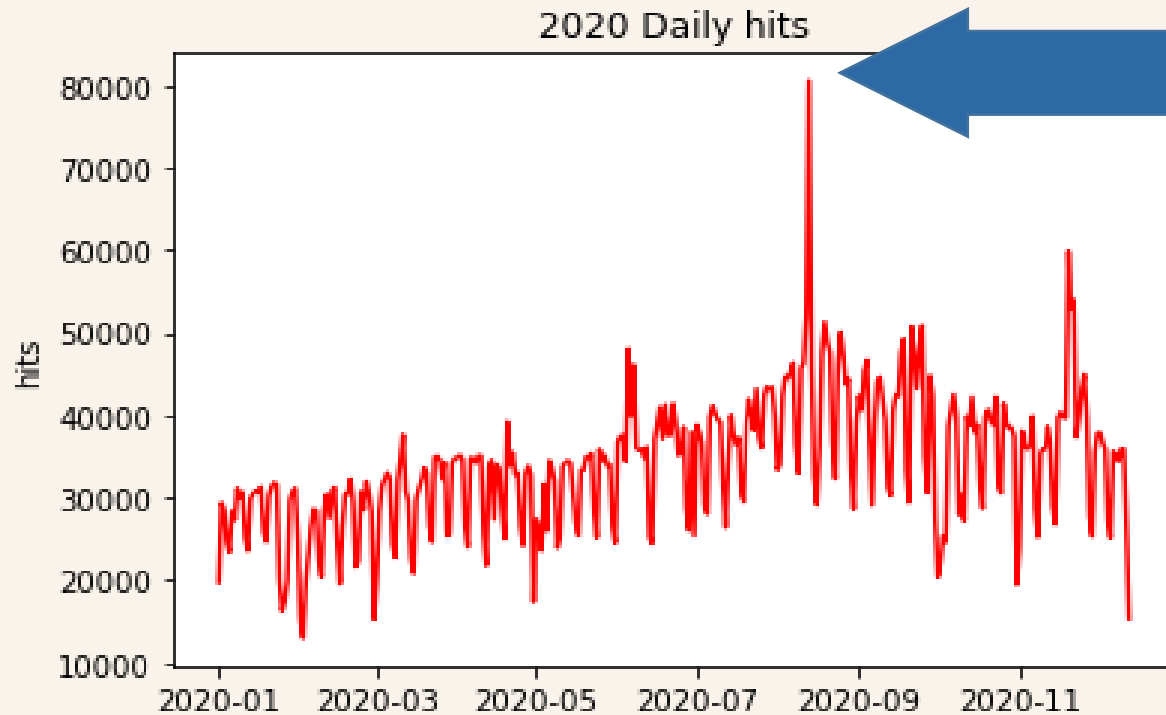
(2)

민원 발생의 흐름은 어떠할까?

2020년이 가기 전 알아보는 2020년의 민원발생 흐름

2020년 전체 민원 건 수 변화

• 데이터: 2020년 1월 1일부터 2020년 12월 13일까지



2020년 08월

#하남_연장

#직결연장_원안

#교산지구_지하철

#하남시_범시민연대

#하남연장_원안

소프트 스킬 분석

• 데이터: 2020년 1월 1일부터 2020년 12월 13일까지

N **하남시 범시민연대**

통합 **VIEW** 이미지 지식iN 동영상 쇼핑 뉴스 실시간검색 어학사전 ...

정렬 ▾ 2020-08-12 - 2020-08-13 ▾ 옵션유지 ☐ 상세검색 ▾

- 전체
- 블로그
- 카페

VIEW VIEW에서 #일상을 검색해보세요!

 하남감일 한양수자인 입주자모임 | 2020.08.12.

[민원릴레이] 총연합회카페에서 각 카페마다 동참을 요하셔서 내용을 올립니다~!!

하남 시민의 분노는 **하남시 범시민연대** 결성과 3호선 직결연장 원안 사수라는 한 목소리로 이어지고, 20.07.17 최종윤 의원, 김상호 하남시장, 방미숙 하남시의회 의장 및 시·도의원들로 하여금 '3호선 하남 연장' 관련...

 감일 신혼희망타운 입주자모임 :) | 2020.08.12.

[총연공지] 3호선 관련 민원 릴레이 진행

하남 시민의 분노는 **하남시 범시민연대** 결성과 3호선 직결연장 원안 사수라는 한 목소리로 이어지고, 20.07.17 최종윤 의원, 김상호 하남시장, 방미숙 하남시의회 의장 및 시·도의원들로 하여금 '3호선 하남 연장' 관련...

2. 민원내용은 아래와 같습니다.

제목 : 국토부님 **지하철 3호선 복선 하남 직결연장 원안**을 조속히 이행하라.

내용 : 지난 2018년 12월19일, 국토교통부 김현미 장관이 <3기 신도시 건설계획과 수도권 광역교통망 개선방안>을 발표하면서 교산지구 지하철 3호선 직결연장을 교통대책으로 내놓았으나, LH가 3억 썩이나 들여 용역 발주한 <송파-하남 간 도시철도건설 사전타당성 용역 결과보고서>는 당초 연장하기로 한 3호선 중전철이 아니라 경전철에 유리하게 맞춰진 엉터리 보고서이기에 하남 시민의 분노는 매우 격하게 고조된 지경에 이르렀습니다.

하남 시민의 분노는 하남시 범시민연대 결성과 3호선 직결연장 원안 사수라는 한 목소리로 이어지고, 20.07.17 최종윤 의원, 김상호 하남시장, 방미숙 하남시의회 의장 및 시·도의원들로 하여금 '3호선 하남 연장' 관련 원팀 체제로의 원안추진 결의를 이끌어 내었으며, 20.08.06 하남시의회로부터 '제3기 교산신도시 조성관련 지하철 3호선 하남연장 원안 추진 촉구 성명서'를 이끌어 내었습니다.

이처럼 하남 시민들은 3호선 직결 연장 원안이 실행되는 그날까지 끊임없이 강력한 촉구를 이어갈 것입니다.

• 출처: <https://cafe.naver.com/soojain1/867>

(3)

많이 나온 민원 키워드는 무엇일까?

- 일주일 단위 빈출 상위 5개의 민원 키워드 분석



2019



2020

민원 키워드 TOP20

• 상위 5개 키워드의 총합

#시흥_은계지구

#불법_주정차

#경의선_향동역

#착공_요청

#이재준_시장

#불법_주정차

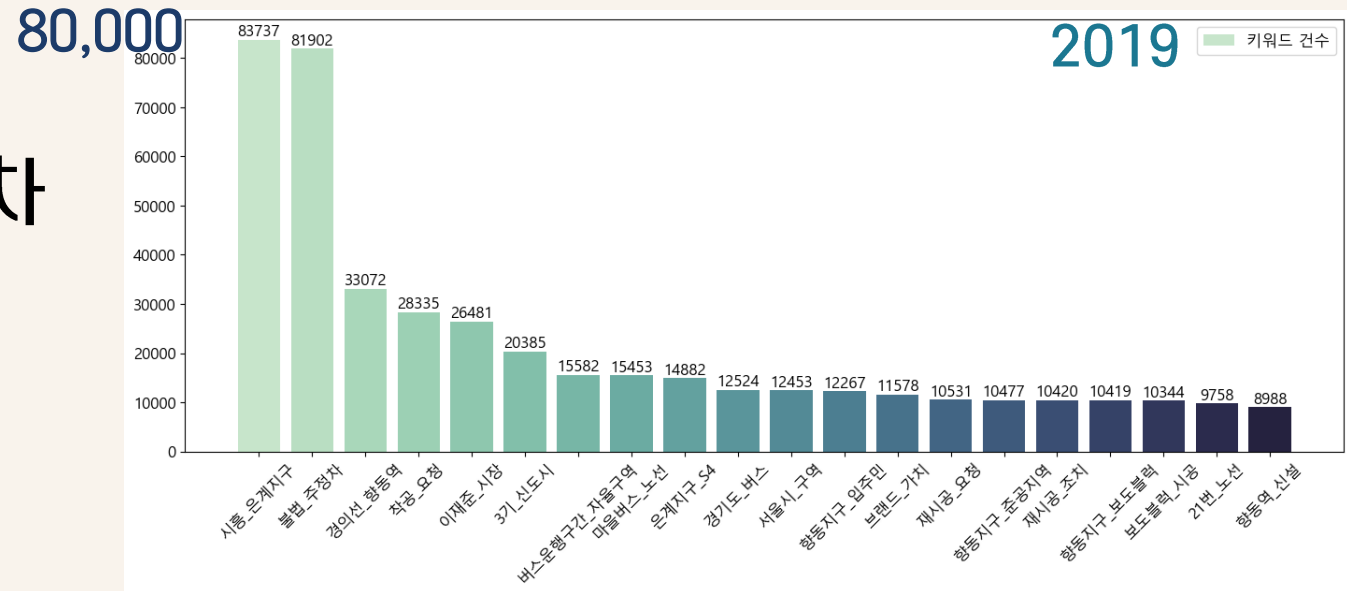
#주택_분양가

#녹지공간_설계원안

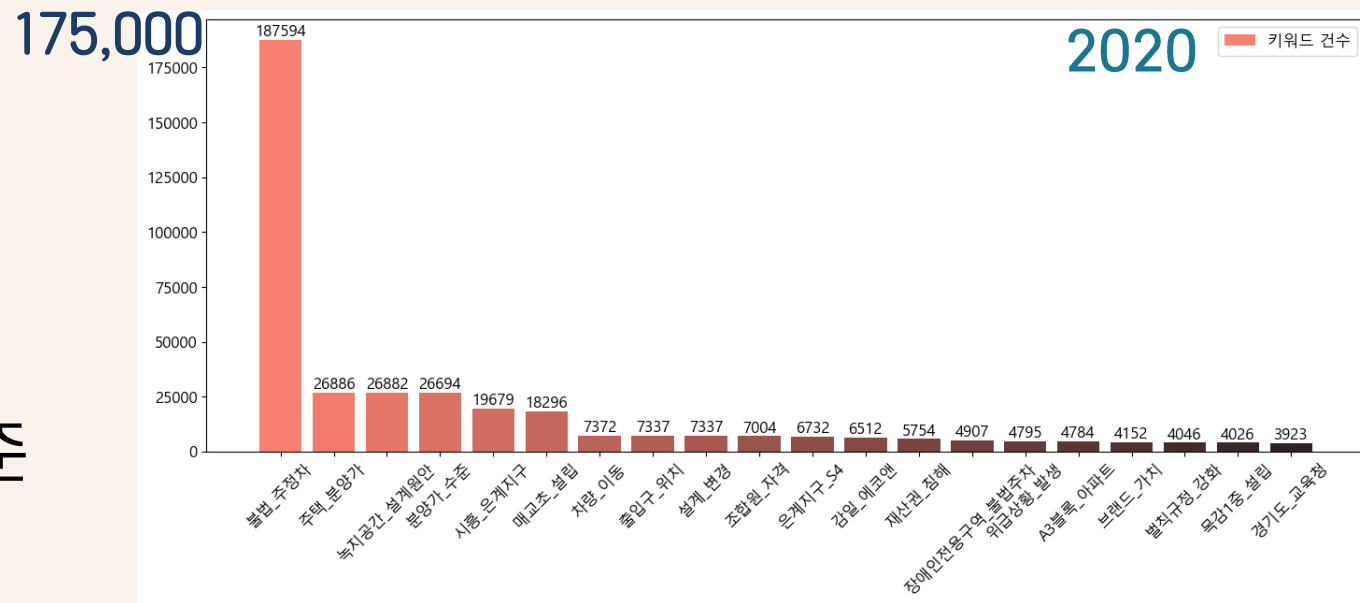
#분양가_수준

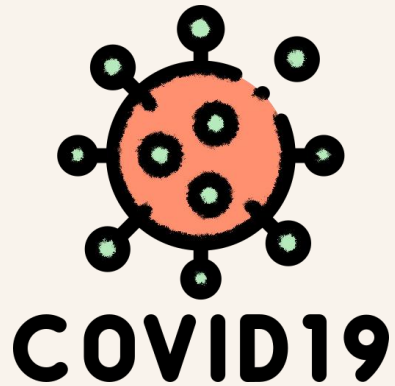
#시흥_은계지구

80,000

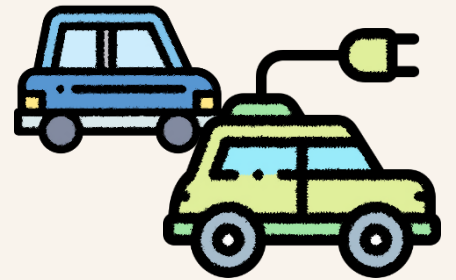


175,000





왜 2020년에는
불법 주정차 키워드가
2배 이상 증가했을까?

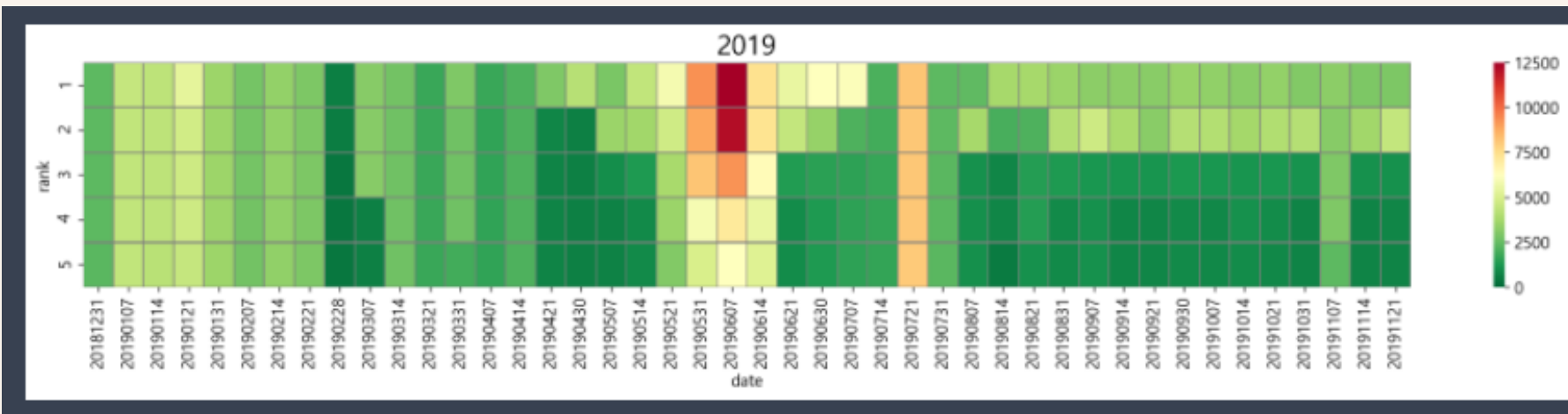


연중 언제 가장 집중 되었나?

- 1년 동안 일주일 단위로 상위 5개 키워드의 회수

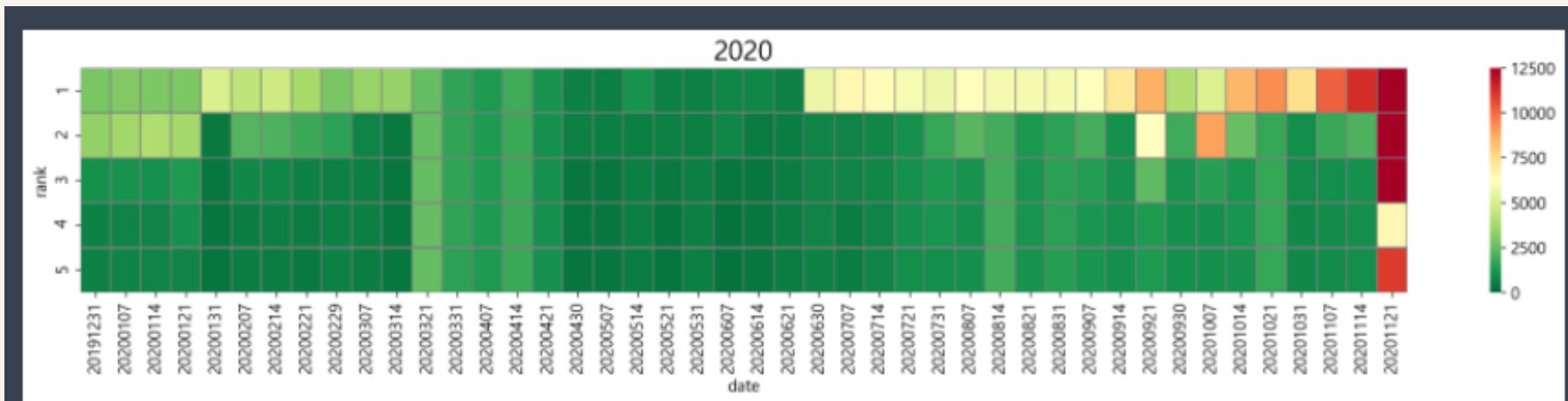
2019

순위
5 4 3 2 1



2020

순위
5 4 3 2 1

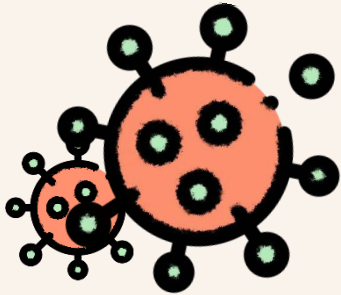


<Figure size 2200x300 with 0 Axes>

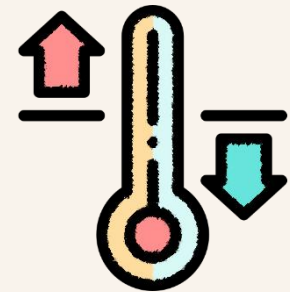
날짜 (일주일)

(4)

외부 요인이 민원 발생에
영향을 끼치지 않을까?



코로나 팬데믹 COVID-19 || `기온`에 따른 민원 빈도수



회귀 분석 민원 X 외부 요인

- 스케일링은 통한 변수 표준화

신뢰할 만한 `코로나 확진자 수(명)` 데이터

=> 2020년 2월 20일부터

*최초 발생 => 2020년 1월 19일

데이터

2020년 2월 20일- 2020년 12월 13일

코로나 확진자 수

X_1

전국 평균 기온

X_2

민원발생 수

y

민원 확진자 기온

	gunsu	confirmed	평균
day			
2020-02-20	32293.0	73.0	5.0
2020-02-21	29517.0	100.0	7.1
2020-02-22	21682.0	229.0	6.4
2020-02-23	22296.0	169.0	3.5
2020-02-24	30747.0	231.0	7.2
...
2020-12-09	34419.0	686.0	1.5
2020-12-10	35914.0	682.0	5.1
2020-12-11	35971.0	689.0	5.6
2020-12-12	26760.0	950.0	4.2
2020-12-13	15316.0	1030.0	1.3

298 rows × 3 columns

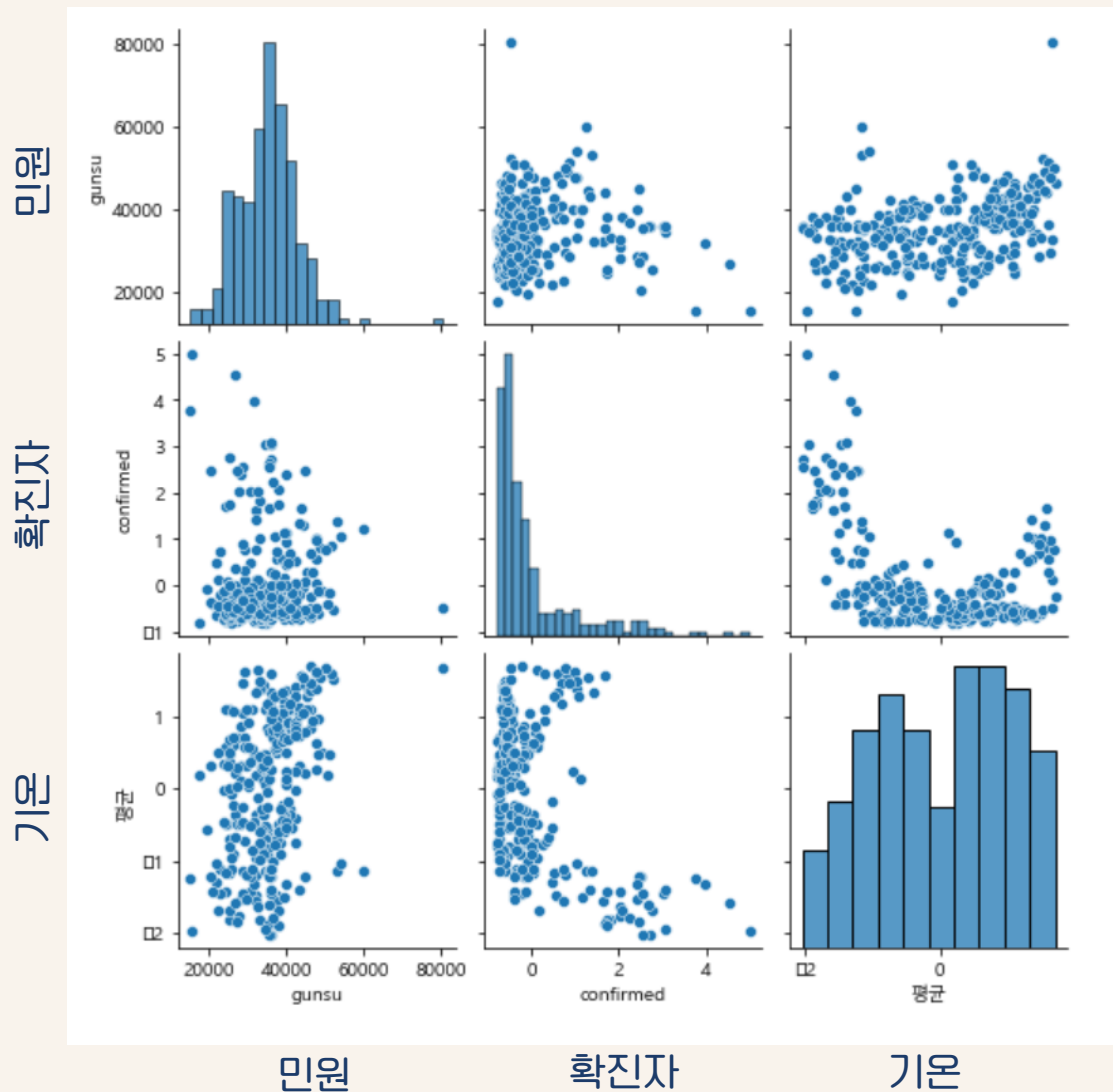
확진자 기온

confirmed	평균
-0.399194	-1.475190
-0.246971	-1.191067
0.480319	-1.285774
0.142045	-1.678135
0.491595	-1.177537
...	...
3.056843	-1.948729
3.034291	-1.461660
3.073756	-1.394012
4.545250	-1.583428
4.996283	-1.975789

15

코로나 확진자 수(x1), 기온(x2), 민원발생 수(y)

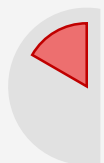
- 산점도(Scatter) 그래프로 데이터 간의 상관 관계를 확인



	VIF Factor	features
0	1.219215	confirmed
1	1.219215	평균

변수들 간의 상관성

- 코로나 확진자 수(x1), 기온(x2), 민원건수(y)의 피어슨 상관계수



상관관계 없음

확진자수 X 민원건수



약한 상관관계
 $0.3 < \text{corr} < 0.7$

평균기온 X 민원건수



음의 상관관계

확진자수 X 기온

		민원 gunsu	확진자 confirmed	기온 평균
민원	gunsu	1.000000	-0.057029	0.367170
확진자	confirmed	-0.057029	1.000000	-0.452494
기온	평균	0.367170	-0.452494	1.000000



예측 모델

- 코로나 확진자 수(x1), 기온(x2) 입력에 따른 민원발생 수(y) 예측 모델

그래프는 선형에 가까워야한다

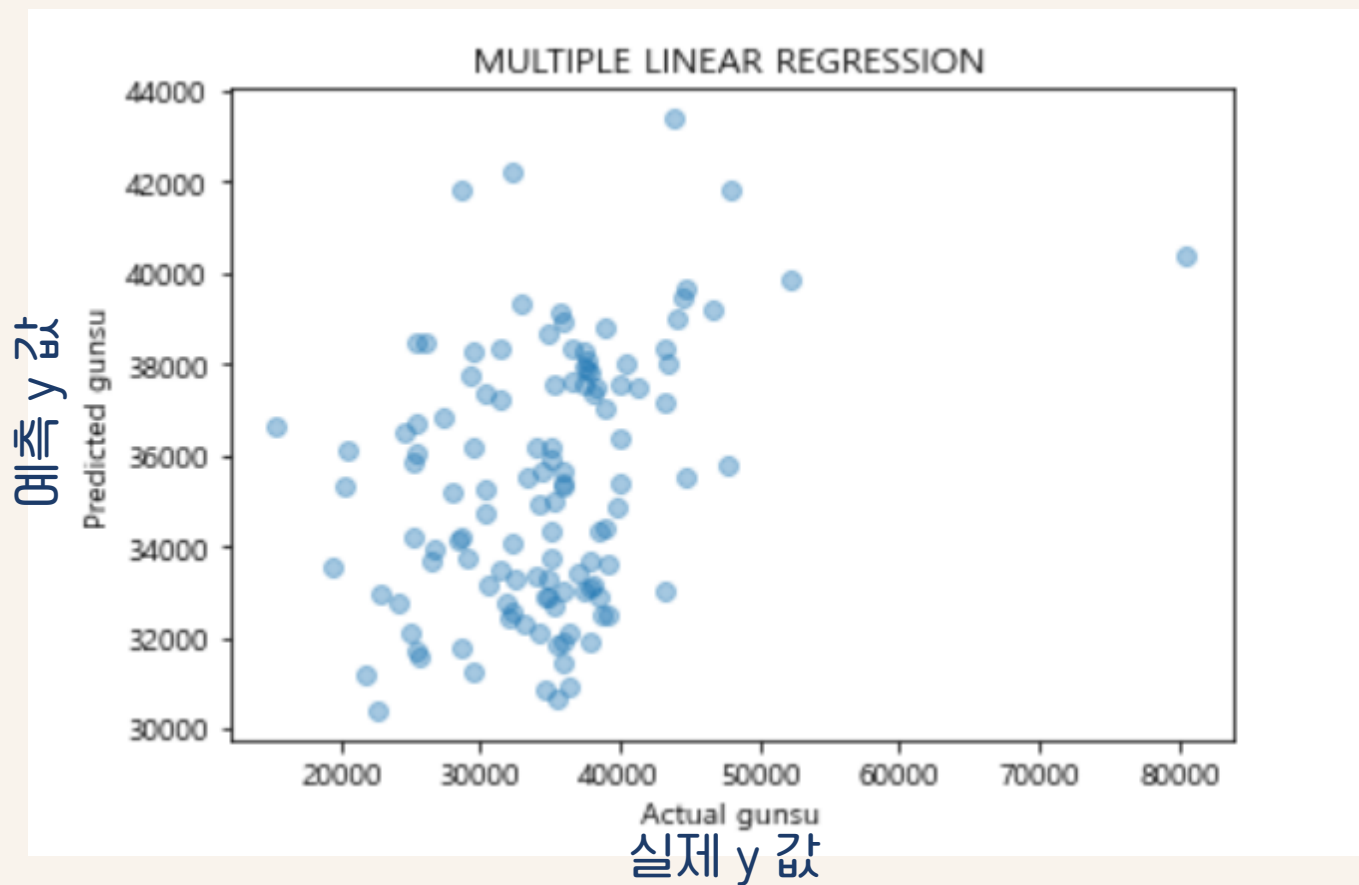
그러나 선형성이 강하지도 않고

모델의 결정계수도 매우 낮음

->적합한 회귀 모델은 아니다.

```
#모델의 결정계수  
print(mlr.score(x_train, y_train))
```

0.16244420052245212



추가 발전 방향

- '민원 라이브' 서비스의 추가 발전 가능성



실시간 연동

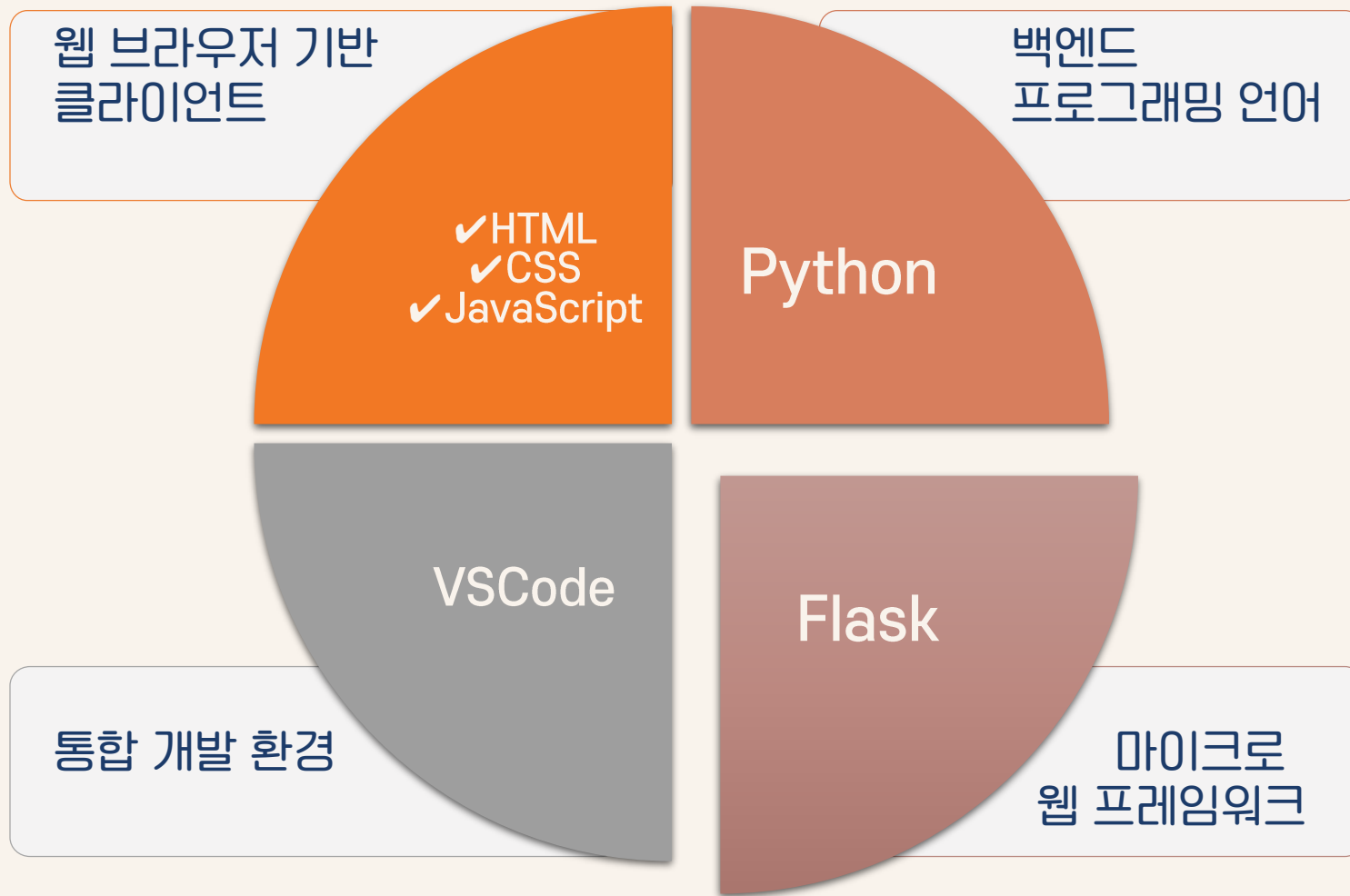
- 실시간으로 처리되고 있는 민원 알림판
- 전화 민원 && 트위터 등 민원 키워드 크롤링
- 실시간 집계 및 분석

시스템 연결

- 민원 전화가 몰릴 때 웹 페이지로 자동 연결
- 전화 연결 전까지 민원 현황을 파악할 수 있게 서비스 제공

웹 구현

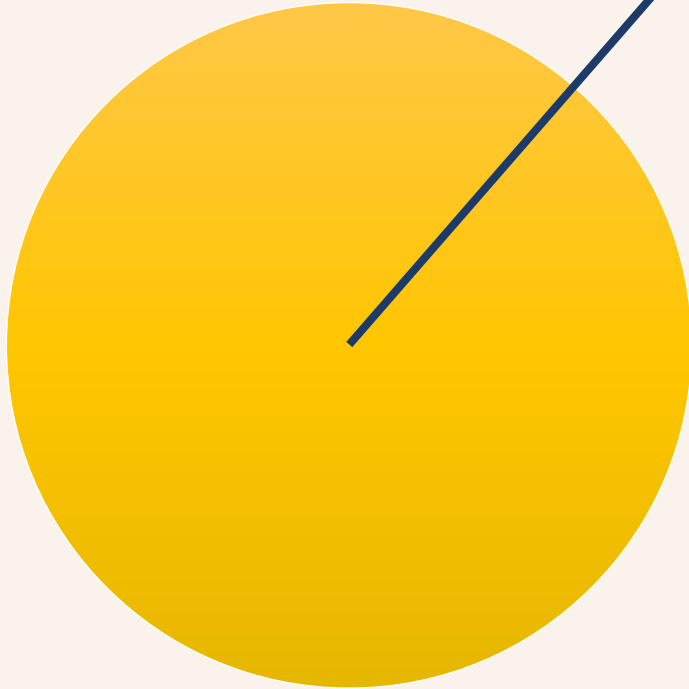
- 웹 서비스를 구현하기 위한 사용 언어 및 프레임



웹 Web 시연

- 통합 민원 웹 서비스

바로가기



출처 및 참고 자료

<https://www.edaily.co.kr/news/read?newsId=01374326625933576&mediaCodeNo=257&OutLnkChk=Y>

<https://cafe.naver.com/soojain1/867>

- Icon, <https://www.flaticon.com/kr/free-icon>
 - Eucalyp
 - Freepik

Thank you for listening.

