



# 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

PAXBD : 유한별 문정승 탁근주 박미선 홍운표



# CONTENTS

## 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

### I. 연구배경 및 목적

### II. 정량적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 파생변수
- 4) 모델링
- 5) 결과

### III. 정성적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 감성분석
- 4) 결과

### IV. 결과

# CONTENTS

## 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

### I. 연구배경 및 목적

### II. 정량적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료전처리
- 3) 파생변수
- 4) 모델링
- 5) 결과

### III. 정성적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 감성분석
- 4) 결과

### IV. 결과

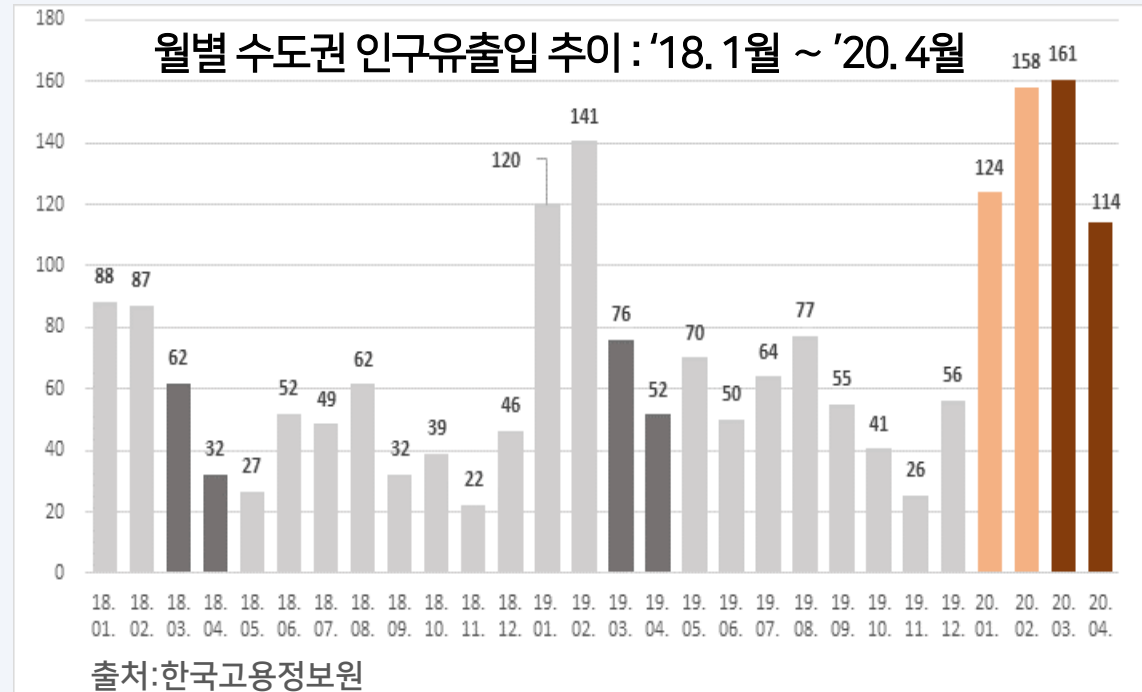
- 2020년 1월 국내 코로나 바이러스의 확산으로 자영업 및 중소기업 등의 영업에 타격
- 경기불황과 취업난이 심각해지면서, 고용상황이 더 나은 수도권으로 청년·중장년층 이동

➡ 지방소멸의 가속화를 불러일으킬 가능성이 높아지고 있음

## 코로나發 '줄폐업'에...무너지는 자영업 생태계



## 코로나 19로 인한 지방소멸 위험의 가속화



- 국회는 지방소멸 문제를 해결하고자 '인구소멸위기지역 지원 특별법 제정안' 등 법안 3건 추진 중
- 지방중소도시들에 대한 소멸 위험을 절감하고 자생력을 키우려는 시도를 하고 있음



지방소멸 위기를 인식하고 현 상태를 진단하기 위해, 지방자치단체별 지방소멸 위험지역을 분류하고  
동시에 지방소멸에 미치는 영향 요인들을 구체화할 필요성 있음

## [ 연구 목적 ]

- 국회는 지방소멸 문제를 해결하고자 '인구소멸위기지역 지정안' 등 법안 3건 추진 중
- 지방소멸방지 위원회를 설치해 소멸 지역을 관리하고, 청년 일자리를 지원하는 등의 특례 조항이 담겨 있음

지방소멸에 미치는 특정한 영향 요인들이 각 지방소멸위험지역에

어떻게 영향을 미치고 있는지를 구체적으로 파악



지방소멸위험지역들의 사회적 관심도와 사회적 자본을 측정하여

지방소멸의 극복가능성에 대한 논의



향후 지방소멸의 극복을 위한 사회적 자본 구축과

지역의 취약 요인 극복 방안에 대한 결론과 정책적 함의를 제공



지방소멸 위기를 인식하고 현 상태를 진단하기 위해, 지방자치단체별 지방소멸 위험지역을 분류하고  
동시에 지방소멸에 미치는 영향 요인들을 구체화할 필요성 있음



## [ 지방 소멸 ]

인구 유출의 지속으로 인구가 사라질 가능성이 높은 지역을 지칭

지방의 인구가 대도시, 중핵도시 등으로 유출되며 해당 지역의 인구 공동화 현상이 나타남



## [ 지방 매력도 ]

대도시, 중핵도시 집중화는 경제·정치·행정적 집중화를 야기하여, 다른 지역과 격차를 형성  
이러한 격차는 상대적으로 지방의 매력도를 감소시켜, 인구 유출로 이어짐

지방 매력도: 경제적인 요인 + 어메니티(amenity)

- 근린어메니티: 의료 서비스, 주택가격, 교육의 질, 범죄 예방 등
- 환경어메니티: 백화점·대형마트, 대학의 질, 도서관, 극장 등
- 도시어메니티: 대기의 질, 상수도 수질, 공원 및 오픈 스페이스 등

## [ 소멸위험지수 ]

20~39세 가임기 여성 인구

노인 인구



[ 지방 소멸 ] 지역 어메니티를 구성하는 요소들에 주목하여 이들이

[ 지방소멸지수 ]

인구 유출 지역 소멸 위험 가능성에 미치는 영향 유무 및 그 정도를 확인

20~39세 가임기 여성 인구

지방의 인구가 대도시, 중핵도시 등으로 유출되며 해당 지역의 인구 공동화 현상이 나타난다

COFOG 세출 10개 분야에 기초하여 소멸위험을 목표값으로 하여

지방정부의 기능적 요인이 소멸위험에 미치는 영향을 분석

[ 지방 매력도 ]

COFOG(정부기능분류)

대도시, 중핵도시 집중화는 경제·정치·행정적 집중화를 야기하여, 다른

이러한 격차는 상대적으로 지방의 매력도를 감소시켜, 인구 유출로 이어

지방 매력도: 경제적인 요인 + 어메니티(amenity)

- 근린어메니티: 의료 서비스, 주택가격, 교육의 질, 범죄 예방
- 환경어메니티: 백화점·대형마트, 대학의 질, 도서관, 극장
- 도시어메니티: 대기의 질, 상수도 수질, 공원 및 오픈 스페이스



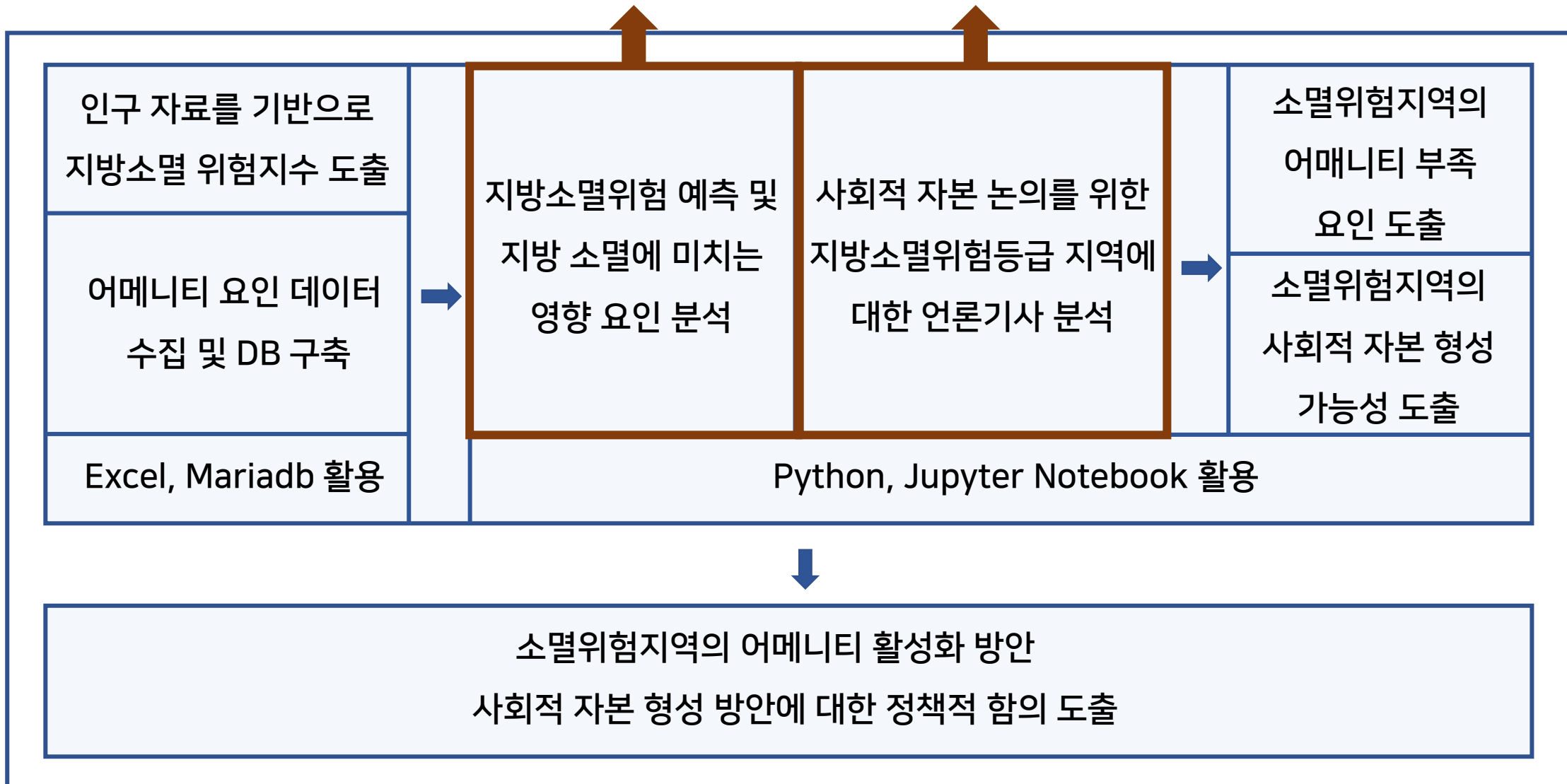
국제분류	통계청 (지역소득계정)	한국은행
01. 일반공공서비스	1. 일반행정 및 국방(01,02)	01. 일반공공서비스
02. 국방		02. 국방
03. 공공질서 및 안전	2. 공공질서 및 안전(03)	03. 공공질서 및 안전
04. 경제업무	3. 경제(04)	04. 경제업무
05. 환경보호	4. 보건복지 및 환경(05,07,10)	05. 환경보호
06. 주택 및 지역개발	5. 주택 및 지역사회개발(06)	06. 주택 및 지역개발
07. 보건	4. 보건복지 및 환경(05,07,10)	07. 보건
08. 오락, 문화 및 종교	6. 교육 및 오락문화(08,09)	08. 오락, 문화 및 종교
09. 교육		09. 교육
10. 사회보호	4. 보건복지 및 환경(05,07,10)	10. 사회보호



## [ 연구 과정 도식화 ]

### II. 정량적 분석

### III. 정성적 분석



# CONTENTS

## 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

### I. 연구배경 및 목적

### II. 정량적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 파생변수
- 4) 모델링
- 5) 결과

### III. 정성적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 감성분석
- 4) 결과

### IV. 결과

### 1) 데이터 탐색 및 수집



“마스다 보고서는 소멸위험지수에 대한 논의를 가임 여성의 감소를 제시하면서  
가임 여성의 감소가 지방소멸 가능성을 높인다고 제시했을 뿐  
소멸의 도달 시기와 해당 소멸위험지수의 정확성 등에 대한 논의는 거의 없음”  
(정성호, 2019).

- 지역의 인구이동을 반영하지 못하는 소멸위험지수의 한계를 인식하여, 기존 소멸위험지수에 대한 개선 진행

#### [ 소멸위험지수 ]

$$\frac{20\sim39\text{세 가임기 여성 인구}}{\text{노인 인구}}$$



#### [ 개선 소멸위험지수 ]

$$(\text{소멸위험지수}^2) + \log \frac{\text{전입인구}}{\text{전출인구}}$$

소멸위험지수의 가중치 부여

해당 값의 분포를 정규분포에 가깝게 구성

### 1) 데이터 탐색 및 수집



개선된 식을 사용하여 소멸위험지수 도출

“마스다 보고서는 소멸위험지수에 대한 논의를 가임 여성의 감소를 제시하면서

해당 소멸위험지수의 4분위수로 나누어 등급화<sup>뿐</sup>

소멸의 도달 시기와 해당 소멸위험지수의 정확성 등에 대한 논의는 거의 없음”

(정성호, 2019).

분류모델의 목표값으로 활용

- 지역의 인구이동을 반영하지 못하는 소멸위험지수의 한계를 인식하여, 기존 소멸위험지수에 대한 개선 진행

#### [ 소멸위험지수 ]

20~39세 가임기 여성 인구

노인 인구

#### [ 개선 소멸위험지수 ]

$$(\text{소멸위험지수}^2) + \log \frac{\text{전입인구}}{\text{전출인구}}$$

소멸위험지수의 가중치 부여

해당 값의 분포를 정규분포에 가깝게 구성

### 1) 데이터 탐색 및 수집

- 국가통계포털(KOSIS)을 통해 2015~2018년도의 전국 228개의 지방자치단체 데이터 수집

[ 국가통계포털(KOSIS) ] <http://kosis.kr/>



1) 교원 1인당 학생수(시도/시/군/구)

자료갱신일: 2020-03-25 / 수록기간: 년 2007 ~ 2019 / 자료문의처 : 042-481-3608

일괄설정 + 항목[3/3] 행정구역별[18/255] 시점[1/13]

(단위 : 명)

2) 행정구역별	2019 교원1인당 학생수 (A÷B)	재적학생수 (A)	교원수 (B)
전국	15.51	8,763,576	565,034
서울특별시	16.03	1,507,375	94,010
부산광역시	17.14	592,113	34,555
대구광역시	14.99	411,916	27,488
인천광역시	15.04	421,065	27,999
광주광역시	16.36	313,463	19,159
대전광역시	16.29	330,007	20,254
울산광역시	13.97	182,827	13,091
세종특별자치시	14.26	77,844	5,460
경기도	16.15	2,101,116	130,119
강원도	14.39	293,979	20,424
충청북도	15.78	307,046	19,453
충청남도	16.20	453,418	27,988
전라북도	14.16	343,400	24,256
전라남도	12.20	284,239	23,297
경상북도	15.05	490,243	32,580
경상남도	14.24	538,036	37,777
제주특별자치도	16.21	115,489	7,124

### 1) 데이터 탐색 및 수집

- COFOG(정부기능분류)를 기반으로 한 다양한 특성 변수들을 구성

#### COFOG(정부기능분류)

국제분류
01. 일반공공서비스
02. 국방
03. 공공질서 및 안전
04. 경제업무
05. 환경보호
06. 주택 및 지역개발
07. 보건
08. 오락, 문화 및 종교
09. 교육
10. 사회보호

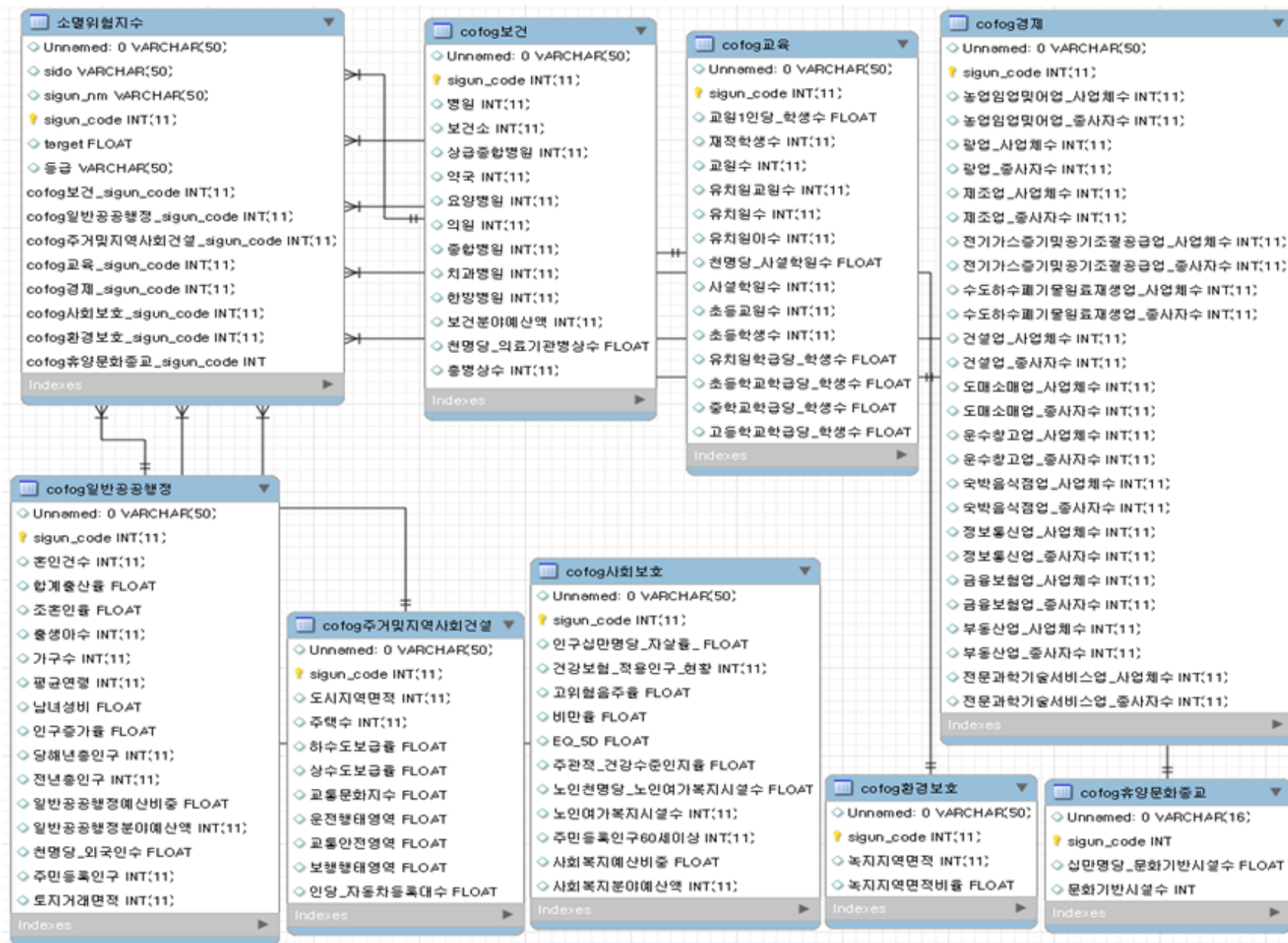


특성	자료 구성	특성	자료 구성
목표값	소멸위험지수 분류(A~D등급)		
주거 및 지역사회 건설	하수도보급률, 상수도보급률, 주택수, 교통문화지수, 운전행태영역, 교통안전영역, 보행행태영역 등	일반 공공행정	혼인건수, 조혼인율, 합계출산율, 평균 연령, 남녀성비, 인구증가율, 당해 년 총 인구, 전년총인구 등
보건	병원 수, 보건소 수, 상급종합병원 수, 약국 수, 요양병원 수, 의원 수, 종합병원 수, 치과병원 수, 한방병원 수 등	사회보호	고위험음주율, 비만율, EQ.5D(건강상태 표준화), 주관적건강수준인지율, 인구 십만 명 당 자살률 등
경제활동	천명당사업체수, 천명당종사자수, 종사자수(남자, 여자), 도소매업사업체수, 서비스업사업체수 등	교육	교원 1인 당_학생 수, 재적 학생 수, 교원 수, 유치원교원 수, 유치원 수, 유치원아 수, 천 명 당 사설학원 수 등
환경보호	녹지지역면적, 녹지지역면적비율	휴양·문화·종교	십만 명 당 문화기반시설 수, 문화기반 시설 수
공공질서 및 안전	교통사고, 화재, 범죄, 자연재해, 생활 안전, 자살, 감염병의 각 등급	국방	국가 단위에서 지출이 있는 특성으로 본 연구는 지방 정부를 대상으로 하여 특성 변수에서 제외함.

## 1) 데이터 탐색 및 수집

수집한 데이터를

Maria DB를 통해 구축





## 1) 데이터 탐색 및 수집

- Train Set: 2015~2017년 기간동안  
수집된 전국 시군구 기준 지방자치단체의  
COFOG 분야 변수

In [4]:

```
1 #기초 통계량 확인
2 train.describe()
```

Out [4]:

	sigun_code	병원	보건소	상급종합병원	약국	요양병원	의원	종합병원	치과병원	한방병원	...	숙박음식점업 _종사자수	정보통신 사업자
count	684.000000	684.000000	684.000000	684.000000	684.000000	684.000000	684.000000	684.000000	684.000000	684.000000	...	684.000000	684.000
mean	37610.293860	6.543860	1.059942	0.187135	94.220760	6.328947	132.628655	1.305556	0.975146	1.248538	...	7013.421053	146.824
std	11671.293746	7.334477	0.617631	0.481041	89.561431	6.478367	161.287598	1.580972	1.852772	3.512735	...	8008.737058	413.441
min	11110.000000	0.000000	0.000000	0.000000	2.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	158.000000	3.000
25%	28597.500000	2.000000	1.000000	0.000000	24.000000	2.000000	24.000000	0.000000	0.000000	0.000000	...	1307.000000	14.000
50%	42725.000000	4.000000	1.000000	0.000000	61.000000	4.000000	74.000000	1.000000	0.000000	0.000000	...	4312.000000	30.000
75%	46722.500000	9.000000	1.000000	0.000000	142.000000	9.000000	186.500000	2.000000	1.000000	1.000000	...	9796.000000	96.250
max	50130.000000	44.000000	6.000000	3.000000	471.000000	38.000000	1518.000000	11.000000	17.000000	40.000000	...	58385.000000	4290.000

8 rows × 93 columns

- Test Set: 2018년 기간에 수집된  
전국 시군구 기준 지방자치단체의  
COFOG 분야 변수

In [5]:

```
1 #기초 통계량 확인
2 test.describe()
```

Out [5]:

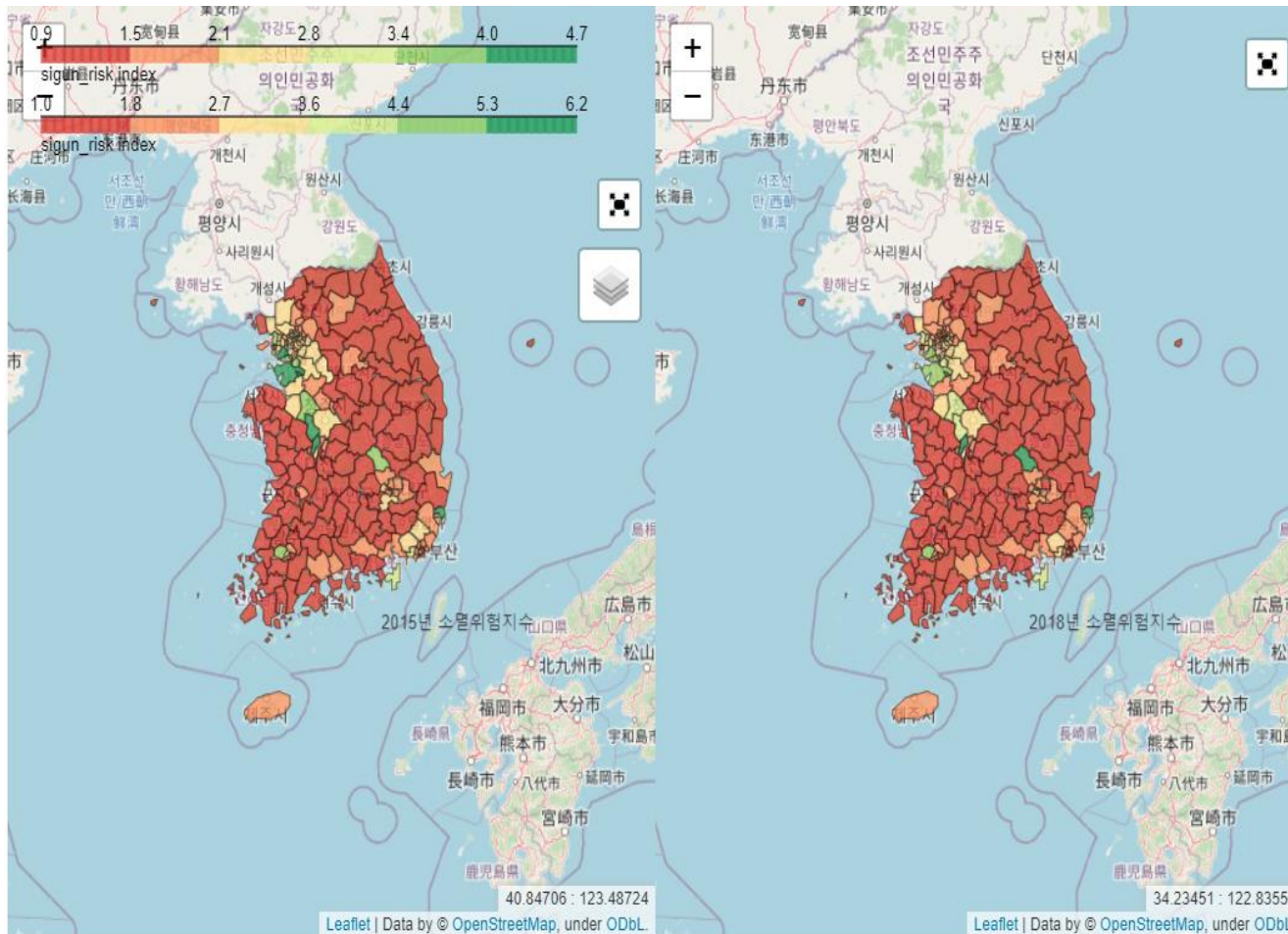
	sigun_code	병원	보건소	상급종합병원	약국	요양병원	의원	종합병원	치과병원	한방병원	...	숙박음식점업 _종사자수	정보통신 사업자
count	228.000000	228.000000	228.000000	228.000000	228.000000	228.000000	228.000000	228.000000	228.000000	228.000000	...	228.000000	228.000
mean	37610.293860	6.425439	1.057018	0.184211	96.850877	6.842105	139.114035	1.364035	1.039474	1.346491	...	7613.754386	149.381
std	11688.419644	7.163211	0.616437	0.479827	92.745565	7.034940	170.962435	1.626847	1.868643	3.520630	...	8580.580267	401.447
min	11110.000000	0.000000	0.000000	0.000000	2.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	233.000000	2.000
25%	28597.500000	1.750000	1.000000	0.000000	23.000000	2.000000	25.750000	0.000000	0.000000	0.000000	...	1347.000000	14.000
50%	42725.000000	4.000000	1.000000	0.000000	60.500000	5.000000	76.500000	1.000000	0.000000	0.000000	...	4982.500000	33.000
75%	46722.500000	9.000000	1.000000	0.000000	146.000000	9.000000	197.500000	2.000000	1.000000	1.000000	...	10456.250000	108.500
max	50130.000000	39.000000	6.000000	3.000000	469.000000	40.000000	1588.000000	11.000000	16.000000	32.000000	...	55341.000000	3974.000

8 rows × 93 columns

### 2) 자료 전처리

- 소멸위험지수 시각화 : 소멸위험지수가 가장 높은 지역은 A등급으로 **빨간색**으로 표시

[ 2015년 ]

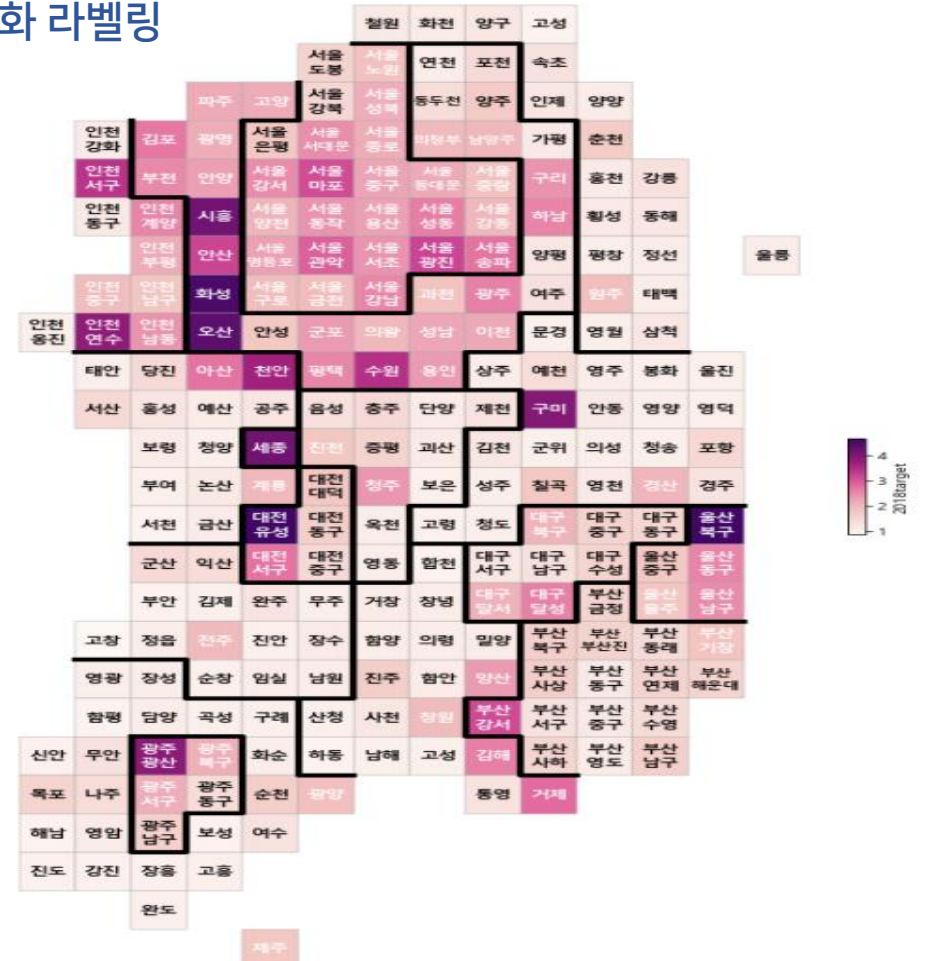


[ 2018년 ]



등급화 라벨링

[ 2018년 ]

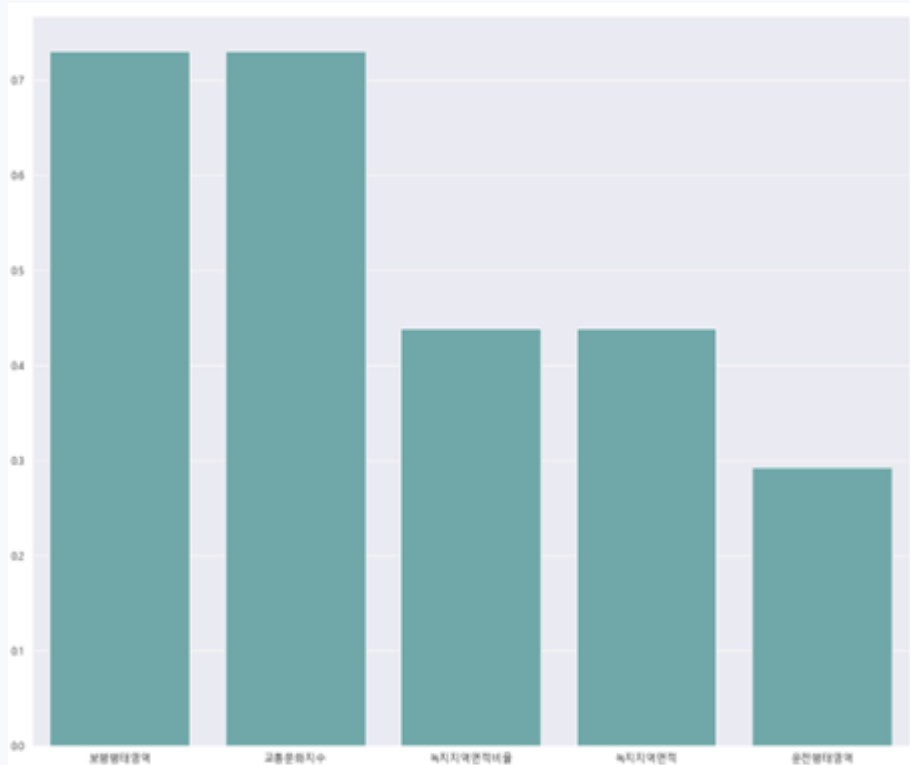


### 2) 자료 전처리

#### - 결측치 및 이상치 확인

전국 228개 지방자치단체에 대한 분석을 진행하기 위해 결측치가 있는 경우,  
행정 구역의 평균으로 대체하거나 지역의 특성을 고려하여 고려할 수 있는 값으로 대체

[Train Set 결측치 시각화]



[Train Set 결측치]

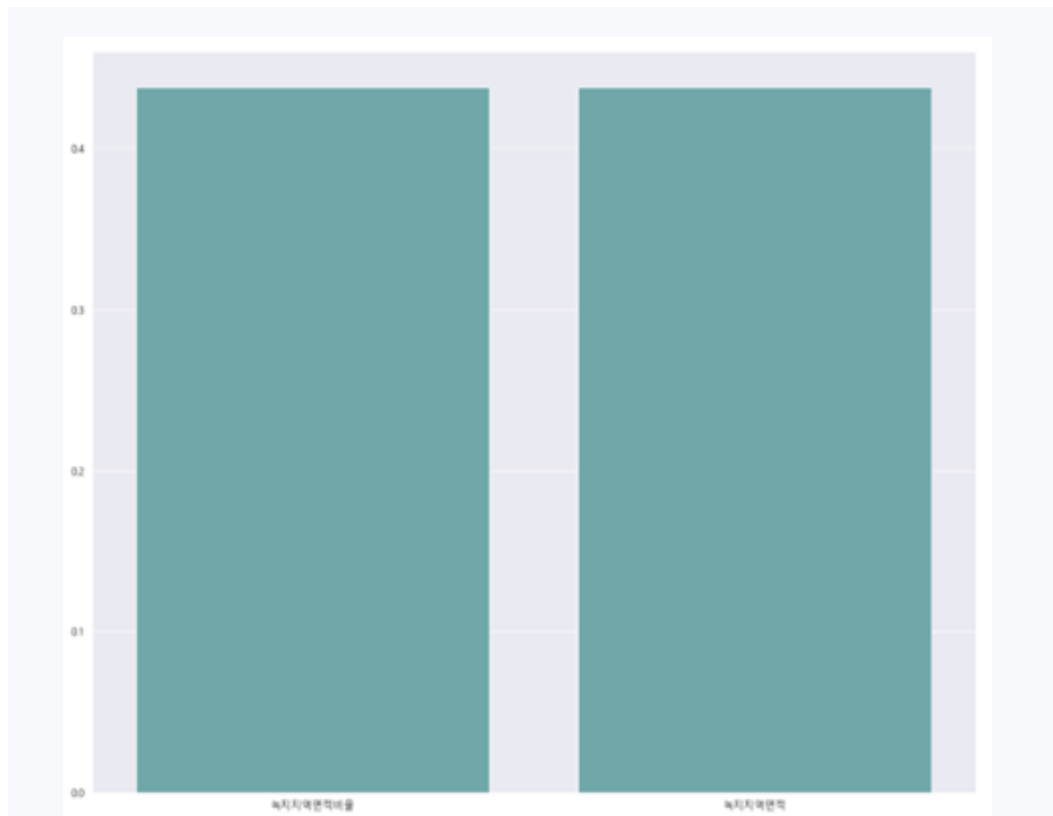
- 1) 보행행태영역, 교통문화지수, 운전행태영역, 녹지지역면적, 녹지지역면적비율의 결측치를 행정구역의 평균으로 대체
- 2) 경상남도 함양군의 2017년 제조업종사자수 결측치를 중분류별 제조업종사자 수인 943명으로 대체
- 3) 인천광역시 옹진군의 녹지면적은 통계 조사 시 누락된 것으로 판단하여 해당 시도의 평균으로 대체

### 2) 자료 전처리

#### - 결측치 및 이상치 확인

전국 228개 지방자치단체에 대한 분석을 진행하기 위해 결측치가 있는 경우,  
행정 구역의 평균으로 대체하거나 지역의 특성을 고려하여 고려할 수 있는 값으로 대체

[Test Set 결측치 시각화]



[Test Set 결측치]

- 1) 녹지지역면적과 녹지지역면적비율 변수의 결측치를 행정구역의 평균으로 대체
- 2) 인천광역시 옹진군의 자살자가 0명인 통계가 있어 자살률을 0으로 대체

### 2) 자료 전처리

#### - 지방소멸지수 연도별 변화 분석

1) 지방소멸지수가 높지만 감소폭이 큰 지역

경기도, 대전광역시, 울산광역시

2) 지방소멸지수가 낮지만 감소폭이 작은 지역

강원도, 전라북도, 전라남도

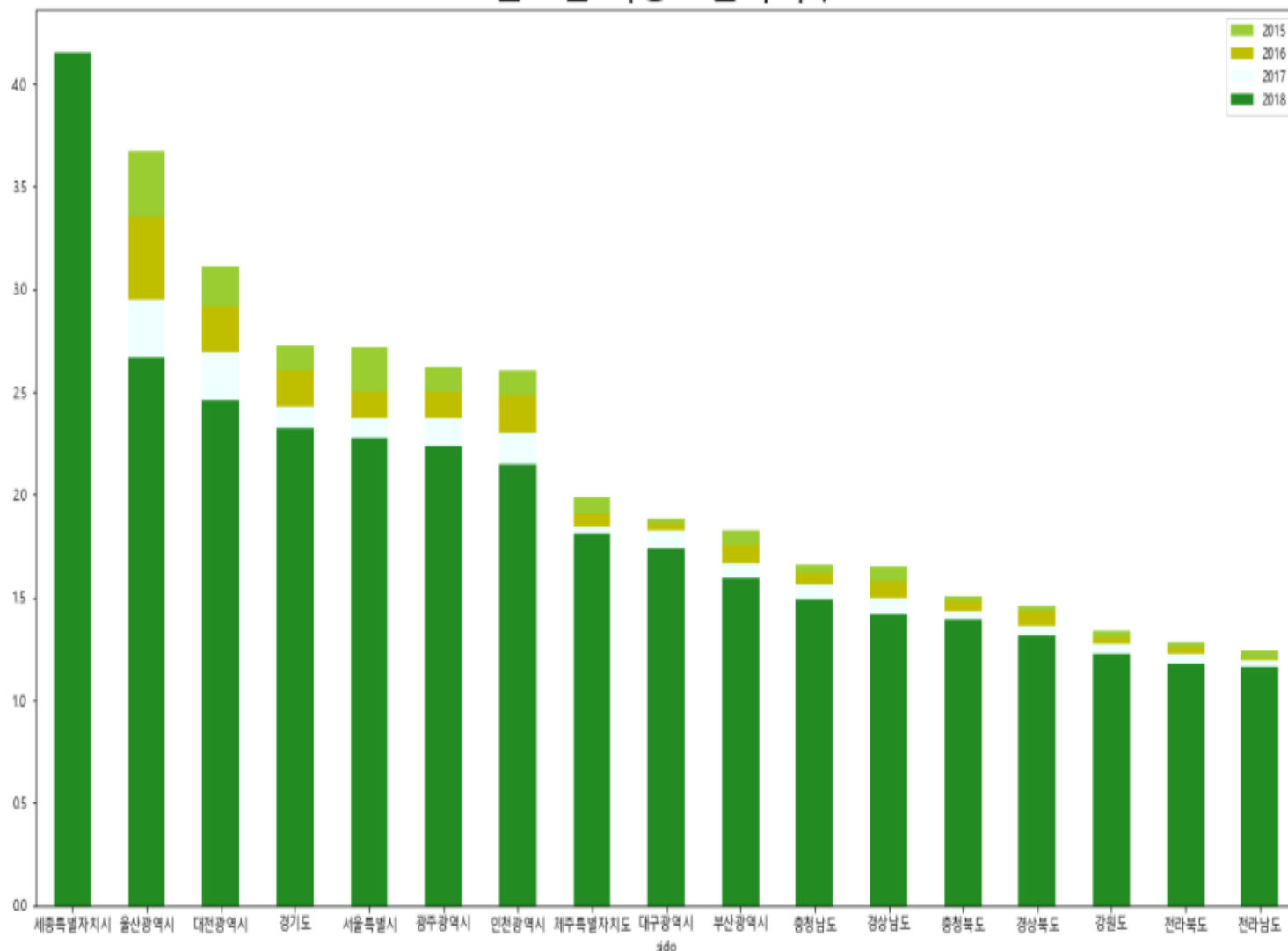


광역시 수준에서도 지방소멸지수가 높아질 수 있음을 확인

여성인구 감소와 지방 간 인구이동이 적어짐

시도 수준에서도 지방소멸에 대한 논의가 필요함

연도별 지방소멸화지수





## 3) 파생변수

특성 중요도가 높은 변수를 곱하거나, 변수 간 상관관계가 높은 변수 등을 합한 변수들을 생성하여 모델에 반영

```
# 파생변수 생성
# Importance가 가장 높은 두 변수를 곱해줄
dat1["평균연령x인구증가율"] = dat["평균연령"] * dat["인구증가율"]
# 의료, 복지 기관으로 분류할 수 있는 모든 feature들을 더해줄
dat1["의료복지기관"] = dat["요양병원"] + dat["병원"] + dat["의원"] + dat["보건소"] + dat["약국"]
# 가구수와 합계출산율을 곱해줄
dat1["가구수당x합계출산율"] = dat["합계출산율"]*dat["가구수"]
# 한 가구가 평균적으로 3명으로 이루어져 있다는 가정하에 각 가구별 평균연령
dat1["가구수당_평균연령"] = (dat["평균연령"]/dat["가구수"])*100
# 녹지지역면적과 토지거래면적의 곱
dat1["녹지지역비율x토지거래면적"] = (dat["녹지지역면적비율"]*dat["토지거래면적"])/100

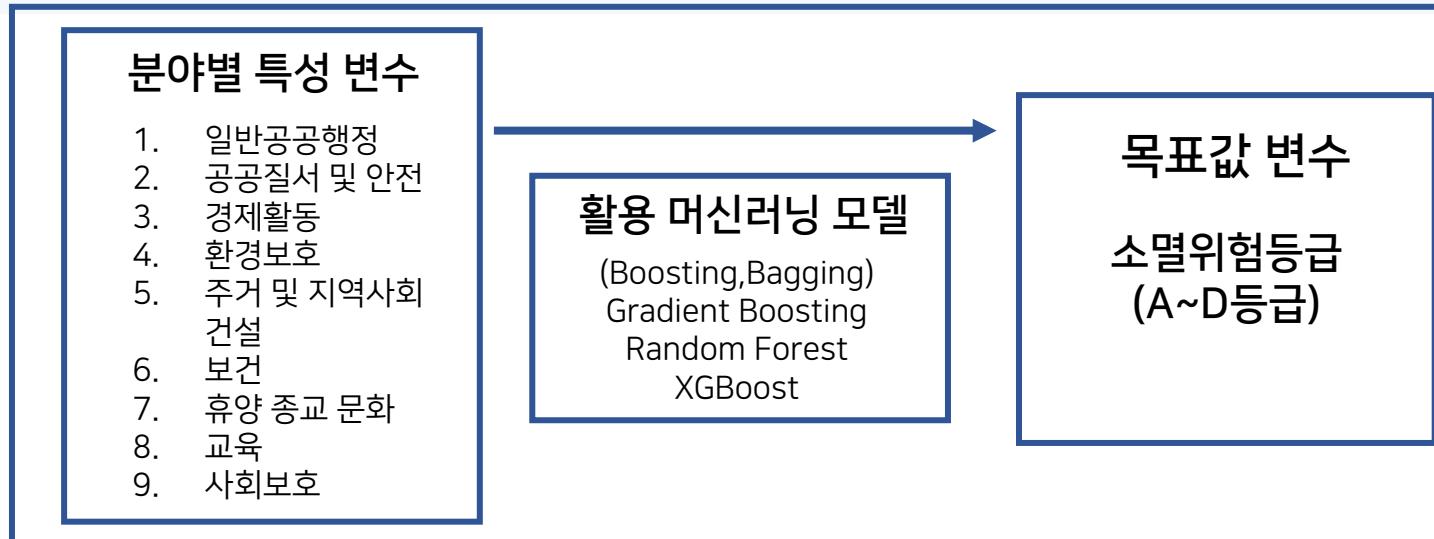
test_x1["평균연령x인구증가율"] = test_x["평균연령"]*test_x["인구증가율"]
test_x1["의료복지기관"] = test_x["요양병원"] + test_x["병원"] + test_x["의원"] + test_x["보건소"] + test_x["약국"]
test_x1["가구수당x합계출산율"] = test_x["합계출산율"]*test_x["가구수"]
test_x1["가구수당_평균연령"] = (test_x["평균연령"]/test_x["가구수"])*100
test_x1["녹지지역비율x토지거래면적"] = (test_x["녹지지역면적비율"]*test_x["토지거래면적"])/100
```

### 4) 모델링

자료 전처리를 통한 자료 분석과 정제가 완료된 후에 해당 특성 변수와 소멸위험지수로 도출한 등급 목표값을 활용한 분류 모델을 구성

#### [분석 모델 선정과 분석]

- 데이터 셋을 분할하여 기본 모델을 구축하고, RandomizedSearchCV 방법을 이용하여 변수 최적화
- K-fold 교차검증을 통해 모델의 검증 성능의 신뢰도를 높임
- Boosting, Gradient Boosting Machine(GBM), XGBoost, Random Forest 에 대한 학습과 검정을 진행
- 가장 적합한 분류모델 선택 및 지방소멸위험지역과 해당 지역의 특정 분야 취약성 도출

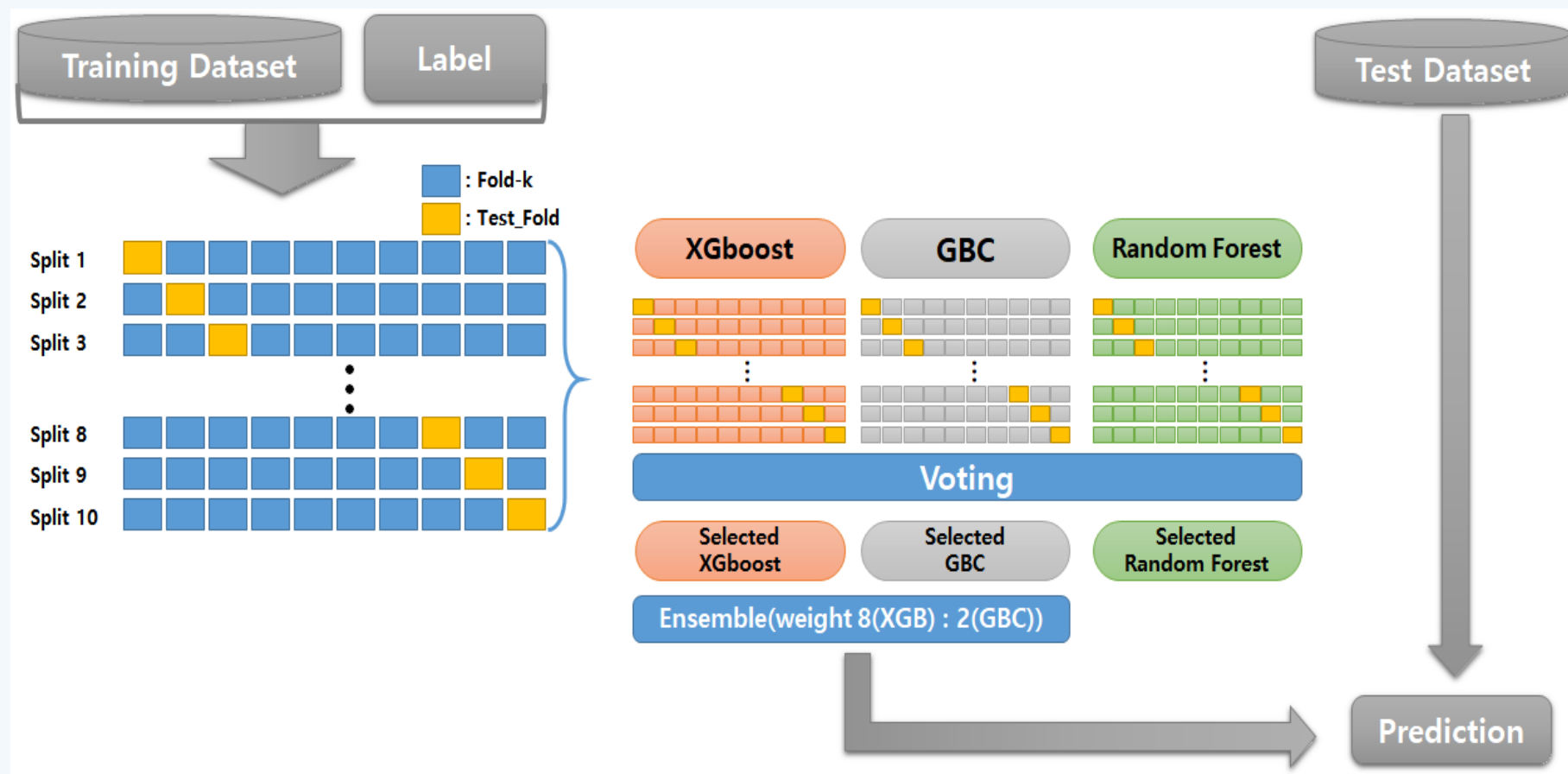




## 4) 모델링

자료 전처리를 통한 자료 분석과 정제가 완료된 후에 해당 특성 변수와 소멸위험지수로 도출한 등급 목표값을 활용한 분류 모델 구성

## [모델 구성 과정의 도식화]



## 4) 모델링

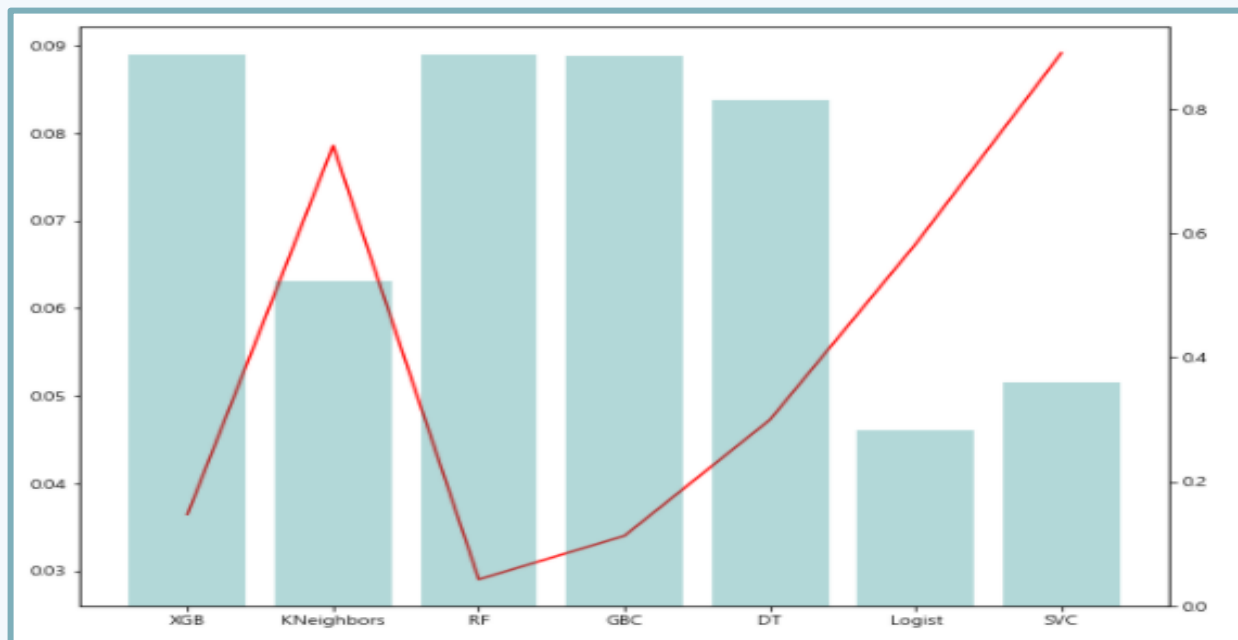
### K-Fold

- 데이터의 수가 적은 경우에는 이 데이터 중의 일부인 검증 데이터의 수도 적기 때문에 검증 성능의 신뢰도가 떨어짐
- 그렇다고 검증 데이터의 수를 증가시키면 학습용 데이터의 수가 적어지므로 정상적인 학습이 되지 않음

➡ 이러한 딜레마를 해결하기 위한 검증 방법으로 Stratified K -폴드(K-fold) 교차검증 방법이 있다

### [각 모델의 결과값과 그래프]

	CrossValMeans	CrossValerrors
XGB	0.888361	0.036470
KNeighbors	0.522065	0.078578
RF	0.887933	0.029035
GBC	0.885359	0.034012
DT	0.813907	0.047284
Logist	0.282241	0.067356
SVC	0.359241	0.089164



## 4) 모델링

## K-Fold

데이터 세트를 10개의 그룹으로 분할한 다음

10개의 모델에 대한 기본 모델구축

- 데이터의 수가 적은 경우에는 이 데이터 중의 일부인 검증 데이터의 수도 적기 때문에 검증 성능의 신뢰도가 떨어짐
- 그렇다고 검증 데이터의 수를 증가시키면 학습용 데이터의 수가 적어지므로 정상적인 학습이 되지 않음

→ 이러한 딜레마를 해결하기 위한 검증 방법으로 Stratified K-폴드(K-fold) 교차검증 방법이 있다

[각 모델의 결과값과 그래프]

RandomizedSearchCV를 통해서 변수 최적화

	CrossValMeans	CrossValerrors
XGB	0.888361	0.036470
KNeighbors	0.522065	
RF	0.887933	0.029035
GBC	0.885359	0.034012
DT	0.813907	0.047284
Logist	0.282241	0.067356
SVC	0.359241	0.089164

kfold 확인 결과 3개의 모델(XGB,GBC,RFC)을  
이용하여 앙상블 진행



#### 4) 모델링

##### - Gradient Boosting Machine(GBM)

- GBM은 회귀분석 또는 분류분석을 수행할 수 있는 모델로 예측모형의 앙상블 방법론 중 부스팅 계열에 속하는 알고리즘
- 트리 기반 부스팅 방법의 일종으로 현재 데이터 과학자들이 좋은 결과를 내는데 가장 많이 쓰는 머신러닝 방법

RandomizedSearchCV 통해 구한  
최적화된 파라미터 가지고  
Gradient Boosting Classifier 모델을 만듦

모델에서 test를 예측하고 f1_score점수화
F1 Cross_validate 0.8985238083498659
F1 Macro: 0.9006486882886648
F1 Micro: 0.9035087719298246
F1 Weighted: 0.9041772655286356



다중 분류기들을 Voting 하는 방법인  
단일 모델, Soft Voting, Hard Voting 사용  
=> F1\_score 0.9042로 가장 높게 나옴

- Validation score of a single GBC Classifier: 0.9042
- Validation score of a VotingClassifier on 3 GBC with hard voting strategy: 0.9042
- Validation score of a VotingClassifier on 3 GBC with soft voting strategy: 0.9042

#### 4) 모델링

##### - Random Forest

- Boosting 방법은 과적합의 가능성이 있기 때문에 앙상블 알고리즘의 다른 방법인 Bagging중 하나로 Random Forest 방법 선택
- 의사결정나무를 여러 개 구성 후 앙상블하여 학습성능을 높이는 방법

RandomizedSearchCV 통해 구한  
최적화된 파라미터 가지고  
Random Forest 모델을 만듦

모델에서 test를 예측하고

f1\_score점수화

F1 Cross_validate	0.8938804577096446
F1 Macro:	0.8659813097830549
F1 Micro:	0.868421052631579
F1 Weighted:	0.8699009261842074



다중 분류기들을 Voting 하는 방법인  
단일 모델, Soft Voting, Hard Voting 사용  
=> F1\_score 0.8829로 가장 높게 나옴

- Validation score of a single RF Classifier: 0.8829
- Validation score of a VotingClassifier on 3 RF with soft voting strategy: 0.8828
- Validation score of a VotingClassifier on 3 RF with hard voting strategy: 0.8828

## 4) 모델링

## - XGboost

- GBM은 residual을 줄이는 방향으로 weak learner를 결합해 강력한 성능을 자랑하지만,과적합 되기 쉽다는 문제점이 있음
- XGBoost는 GBM에 regularization term을 추가한 알고리즘으로 다양한 loss function을 지원해 유연한 튜닝이 가능

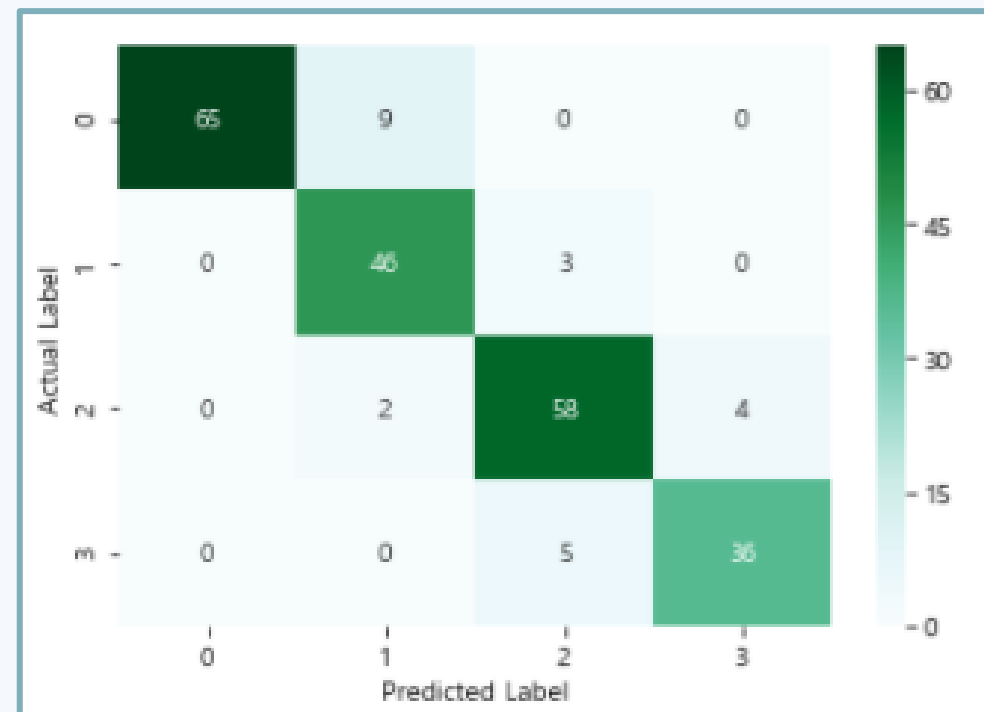
[모델 점수화]

f1 0.9003906654932211

	precision	recall	f1-score	support
0	1.00	0.88	0.94	74
1	0.81	0.94	0.87	49
2	0.88	0.91	0.89	64
3	0.90	0.88	0.89	41
accuracy			0.90	228
macro avg	0.90	0.90	0.90	228
weighted avg	0.91	0.90	0.90	228

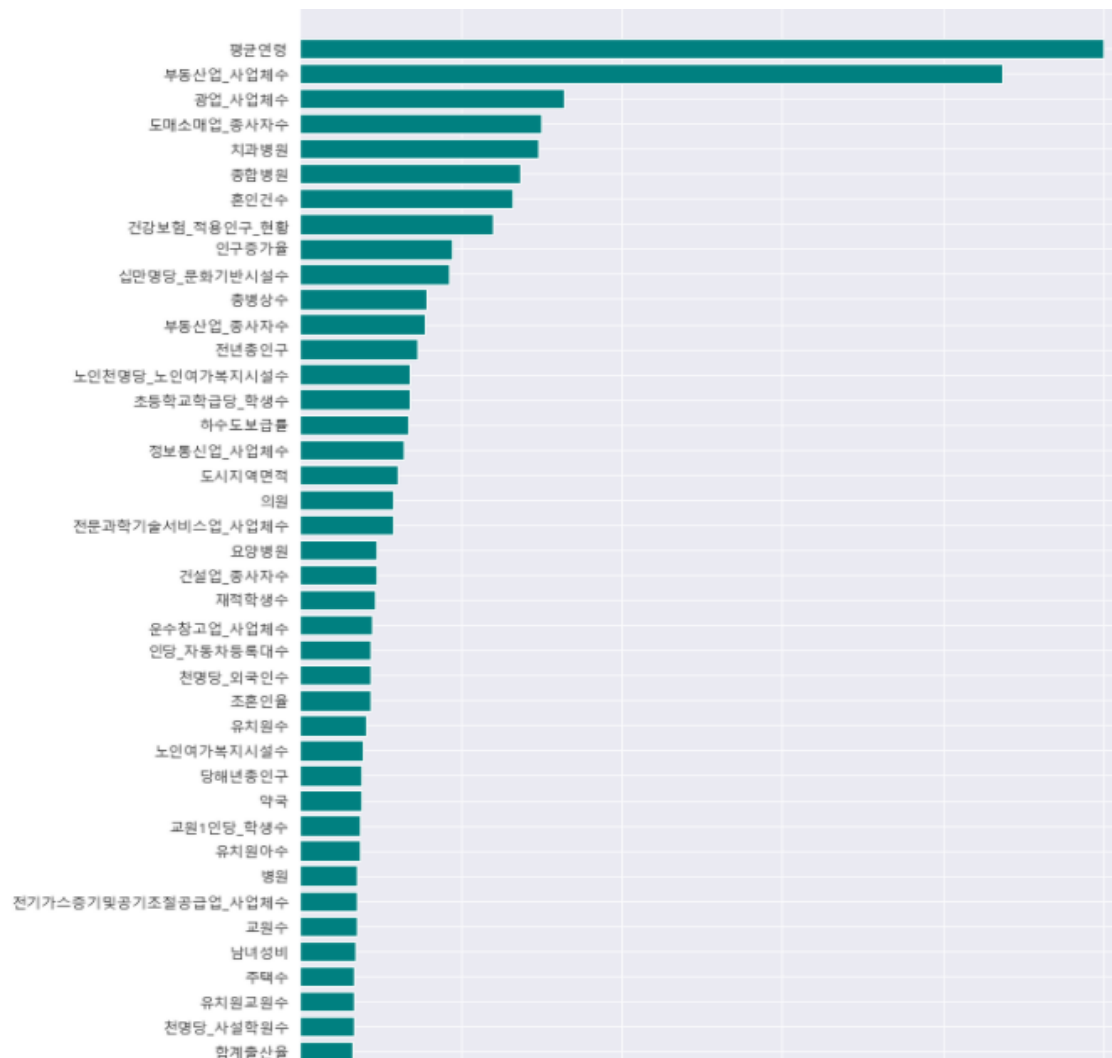


[모델 분류 시각화]



### 4) 모델링

#### [변수 중요도]



#### [순열 특성 중요도]

Weight	Feature
0.3049 ± 0.0232	평균연령
0.0459 ± 0.0078	실업률당_문화기반시설수
0.0386 ± 0.0092	인구증가율
0.0205 ± 0.0018	노인전명당_노인여가복지시설수
0.0023 ± 0.0023	혼인건수
0.0015 ± 0.0000	광업_사업체수
0.0003 ± 0.0012	조혼인율
0 ± 0.0000	교통안전영역
0 ± 0.0000	주민등록인구60세이상
0 ± 0.0000	교원1인당_학생수
0 ± 0.0000	재직학생수
0 ± 0.0000	교원수
0 ± 0.0000	교통문화지수
0 ± 0.0000	운전행태영역

#### [1차적 결론 도출]

소멸위험이 높은 지방자치단체들은 긍정적 요인인 지역의 매력도를 높일 수 있는 **의료시설, 문화시설과 같은 다양한 긍정적 수단**에 대한 확보가 필요하며, 경제적인 기반 또한 확보 및 관리가 필요하다



### 4) 모델링

#### - XGBoost 모델 성능 향상

- XGBoost 최적화 파라미터 적용을 위해 `n_estimators`, `max_depth`, `num_class` 등 파라미터를 조정 -> f1\_score 90%까지 상승
- XGB voting 실시하여 클래스 불균형 문제 해결 -> 모델 성능 향상

#### [XGB voting]

분류 가중 균형화된  $y$  값을 사용하여 입력 데이터의 클래스 빈도에 반비례하여 가중치를 자동으로 조정하고,  
다중 출력의 경우  $y$ 의 각 열의 가중치가 곱해진다.



#### [앙상블]

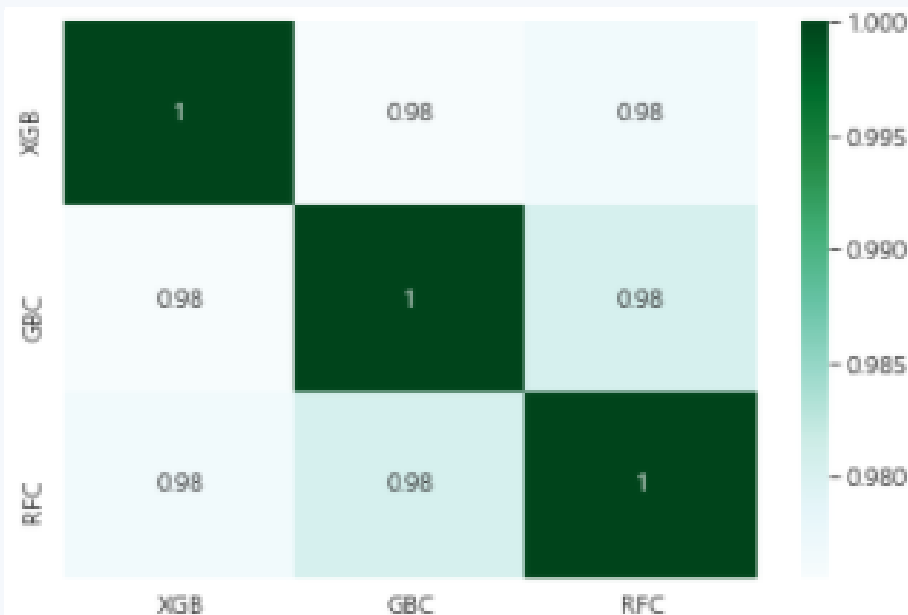
앙상블을 통해 여러 예측 모형을 만들어 낸 후 최종 예측 모형 정합

## 4) 모델링

## - 앙상블

- 여러가지 예측모형들을 만든 후 조합하여 하나의 최종 예측모형을 만드는 방법
- 성능이 많이 좋아지고, 과적합이 방지되지만 해석력이 많이 떨어지고 모델이 무거워진다는 단점이 있음

모델끼리의 상관관계를 확인하여  
XGB와 GBC를 통해서 앙상블 진행



최종 앙상블 모델의 성능은 91.3%

```
#xgb에는 0.8의 가중치를 gbc에는 0.2의 가중치를 줘서 앙상블  
combo_preds = combine_voters(test_x1, weights=[0.8, 0.2])  
global_combo_score_soft = f1_score(test_target, combo_preds, average='weighted')  
global_combo_score_soft  
0.913219839582179
```

## 5) 결과

정형 자료를 활용한 머신러닝 모델을 구축하여

최종적으로 **91.3% 모델 성능**을 도출하여

**66개의 지방자치단체의 소멸위험을 예측**



강원도, 경상도, 인천광역시, 전라도, 충청도

=> 소멸위험이 가장 높은 등급의 시군구 존재

## [ 공통점 ]

인구사회학적으로 노인 인구가 많고

여성 인구와 혼인건수가 적으며

해당 지방자치단체로의 전입이 적음

소멸위험이 높은 지방자치단체명(총 66개)

강원도 양양군	경상북도 고령군	인천광역시 옹진군	전라남도 해남군	충청남도 청양군
강원도 영월군	경상북도 군위군	전라남도 강진군	전라남도 화순군	충청남도 태안군
강원도 정선군	경상북도 문경시	전라남도 고흥군	전라북도 고창군	충청북도 괴산군
강원도 태백시	경상북도 봉화군	전라남도 곡성군	전라북도 김제시	충청북도 단양군
강원도 평창군	경상북도 상주시	전라남도 구례군	전라북도 남원시	충청북도 보은군
경상남도 거창군	경상북도 성주군	전라남도 담양군	전라북도 무주군	충청북도 영동군
경상남도 고성군	경상북도 영덕군	전라남도 보성군	전라북도 부안군	충청북도 옥천군
경상남도 남해군	경상북도 영양군	전라남도 신안군	전라북도 순창군	
경상남도 산청군	경상북도 영주시	전라남도 영광군	전라북도 임실군	
경상남도 의령군	경상북도 울릉군	전라남도 영암군	전라북도 정읍시	
경상남도 창녕군	경상북도 울진군	전라남도 완도군	전라북도 진안군	
경상남도 하동군	경상북도 의성군	전라남도 장성군	충청남도 금산군	
경상남도 함양군	경상북도 청도군	전라남도 장흥군	충청남도 부여군	
경상남도 합천군	경상북도 청송군	전라남도 진도군	충청남도 서천군	
	인천광역시 강화군	전라남도 함평군	충청남도 예산군	

# CONTENTS

## 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

### I. 연구배경 및 목적

### II. 정량적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 파생변수
- 4) 모델링
- 5) 결과

### III. 정성적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 감성분석
- 4) 결과

### IV. 결과

## 1) 데이터 탐색 및 수집

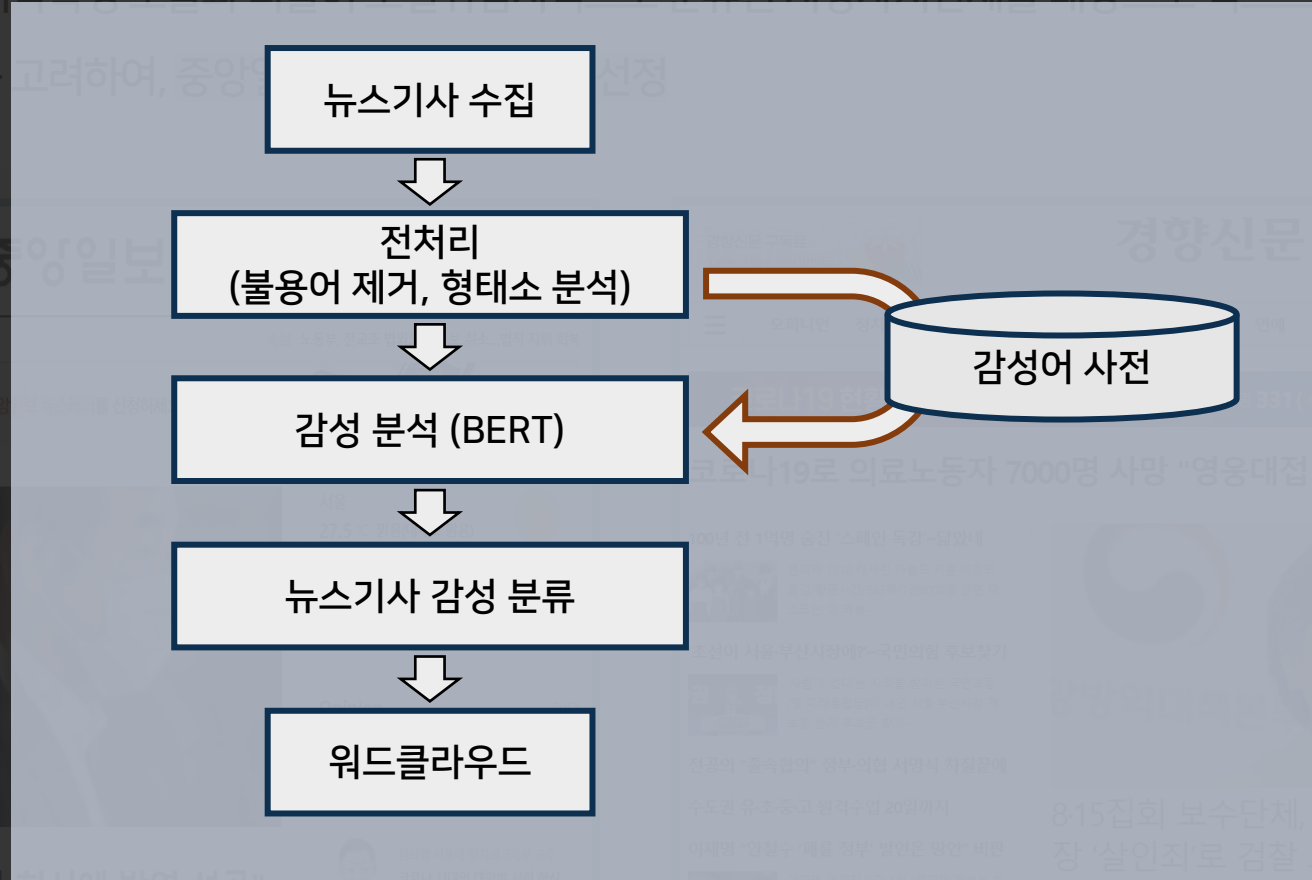
- 소멸위험지역 분류 및 취약특성 도출과 더불어 소멸위험지역으로 분류된 지방자치단체를 대상으로 텍스트 분석을 진행
- 신문기사를 이용한 텍스트마이닝 분석을 통해 사회적으로 형상되어 있는 소멸위험지방자치단체에 대한 인식 조사
- 신문사별 논조의 차이를 고려하여, 중앙일보와 경향신문 선정



#### 1) 데이터 탐색 및 수집

#### 정성적 분석 과정 도식화

- 소멸위험지역 분류 및 취약특성 도출과 더불어 소멸위험지역으로 분류된 지방자치단체를 대상으로 텍스트 분석을 진행
- 신문사별 논조의 차이를 고려하여, 중앙



#### 1) 데이터 탐색 및 수집

- 시군구 단위의 각 지방자치단체명과 '지방'이라는 키워드로 검색된 뉴스 기사를 수집
- 감성 분석을 위한 학습데이터로 사용될 기사는 언론사별 1만 개씩 총 2만개의 최근 뉴스 기사를 수집

- 뉴스기사의 수집기간: 1년으로 설정  
(2019년 8월 21일 ~ 2020년 8월 22일)
- 각 지방자치단체별로 수집된 기사의 수는  
상이하지만 약 40 ~ 100건의 기사수집

기사 제목	작성 시간	기사 내용
최북단 기차역 간 이인영 금강산 관광 재개방법 적극 찾을 것	2020-07-31	이인영 통일부 장관은 31일 금강산 관광이 재개될 수 있는 방법을 적극적으로 찾을 것
DMZ 찾아가는 피아니스트 임미정 음악으로 생명에 위로를	2020-07-31	PLZ 페스티벌의 임미정 감독 피아니스트로 연주에도 참여한다 PLZ 페스티벌
임종석 남북교류사업 도시로 고성군 선정 北 도시와 항구적 협력	2020-07-29	29일 강원 고성군청 대회의실에서 열린 고성군과 남북경제문화협력재단 경문협 간의
기록적 폭우 지나간 자리 5명 목숨 잃고 방방곡곡 침수 피해	2020-07-25	부산에 23일 오후 8시를 기점으로 호우 경보가 발효된 가운데 시간당 50 이상의 많
사망 속보때 음악방송 톤 KBS 수신료 아깝다 성난 부산	2020-07-24	부산 동구 동천 범람으로 24일 오전 부산 동구 범일동 주민들이 복구 작업을 하고 있
삼육대 음악학과 학생들 미래 정치지도자 의회연수 참가	2020-07-23	사진 삼육대 음악학과 학생들과 김철호 지도교수 삼육대 음악학과 성악
해변가에 자욱한 소독 안개 코로나 막자 마스크 쓴 채 물놀이	2020-07-11	강원도 동해안 속초해수욕장 개장 첫날인 10일 피서객들이 마스크를 쓴 채 물놀이를
동해안 해수욕장 개장 전 잇단 사고 수상안전요원 조기 투입	2020-07-02	지난달 20일 강원 속초해수욕장이 더위를 식히려고 찾은 관광객과 시민들로 붐비고
로컬 프리즘 임진각 곤돌라는 언제쯤 개장할까	2020-07-02	최모란 사회2팀 기자 접경 지역인 경기도 파주시엔 비무장지대 DMZ 를 한눈
산림 123 잣더미 고성 산불 원인은 부실시공 화목보일러	2020-06-29	강원 고성군 토성면 산불 주불이 진화된 지난달 2일 오전 화재 현장의 산림이 검게
미리보는 오늘 톱 차기 경찰청장 내정자 발표	2020-06-25	차기 경찰청장은 누구 오늘 윤곽이 나옵니다 민갑룡 경찰청장이 어린이 생명안
날 중대결심 앞두고 절 들어갔다 주호영으로 본 사찰정치	2020-06-24	주호영 미래통합당 원내대표가 약 열흘간 15일 24일 의 사찰 칩거 를 끝내고 24일 국회
강원 고성 북북동쪽 해역서 규모 2.3 지진 피해 없을 듯	2020-06-21	사진 기상청 강원 고성군 북북동쪽 해역에서 규모 2.3의 지진이 발생했다
대북전단 살포 막아라 강원도 전단 살포 금지 행정명령	2020-06-18	지난 17일 서울 중구 커뮤니티하우스 마실 라이브홀에서 열린 남북공동선언을 지키
강원도 접경지역 대북전단 살포 막겠다 위반시 처벌 검토	2020-06-16	16일 경기도 파주시 접경지역에서 경찰이 대북전단 살포를 막기 위해 검문소를 운영
돌이 8000마리 잡았다 오징어 컴백에 미소 번지는 동해안	2020-06-13	지난해 7월 27일 강원 고성군 봉포해수욕장에서 열린 봉포해변 오징어 조개잡이 체
석탄화력발전 있는 전국 지자체 10곳 실무협의회	2020-06-12	인천광역시 용진군 군수 장정민은 지난 10일 충남 당진시청 해나루홀에서 석
강원 속초 앞바다서 길이 13m 무게 35t 죽은 향고래 발견	2020-06-02	지난 1일 오후 7시35분쯤 강원 속초 대포 동방 16해리 해상에서 죽은 채 발견돼 2일
강찬수의 에코사이언스 최악의 동해안 산불 20년 숲 되살리는 자연의 놀라운 복원력	2020-06-02	강찬수 환경전문기자 지난달 27일 강원도 속초 시내에서 북쪽으로 약 20 를 통
강원 고성 경북 안동 울산 울주 산불 복구에 700억 투입	2020-06-01	올해 강원 고성과 경북 안동 강원 고성에서 발생한 산불 복구에 약 700억원을 쓴다 산
희망브리지 전국재해구호협회 재난안전연구소 강원 산불 이후 지역회복 정책포럼	2020-05-28	제2회 강원 산불재난 이후 지역회복을 위한 정책포럼에서 희망브리지 전국재해구호
고성군 산불재해지역에 희망공작소 운영	2020-05-25	강원도 고성군이 구상한 화목 Burning Tree 한 희망공작소 블랙우드 사업이
르포 코로나에 조선사 매출 2배 삼강엠엔티 틈새공략 대박	2020-05-25	지난 21일 삼강엠엔티가 제작한 해상풍력 하부구조물이 바지선에 실려 대만으로 출
민간인 임종석 유엔사 작심 비판에 미국무 이례적 논평	2020-05-24	2017년 11월 6일 국회 운영위원회에 출석한 임종석 당시 대통령 비서실장이 전회경



- Python에서 제공하는 Konlpy 라이브러리를 활용하여 기사명, 광고 등의 불용어와 특수문자 등을 제거
- 형태소 분석기를 사용하여 지방자치단체별 명사단어의 빈도를 계산

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840.

[illegible]

### 3) 감성분석

- 뉴스기사가 내포하는 긍정과 부정의 정도를 수치화하기 위하여 군산대학교 BI\_LSTM 기반의 한국어 감성사전을 활용
- 전처리된 뉴스기사의 단어들과 매칭시켜 감성 스코어가 -3 이하이면 부정, +3 이상이면 긍정으로 라벨링

Figure 10 shows the 'KNU 한국어 감성사전' (KNU Korean Sentiment Dictionary) application. The interface includes a search bar with the input '기다려다' and a button '입력'. Below the search bar, there are labels for '어근:' (Root) and '감성:' (Sentiment). A red box highlights the 'ㄱ' (G) consonant button in the top navigation bar. A red arrow points from this box to a list of sentiment words categorized by consonants. The list shows words like '기다려다' (waiting) with a sentiment of -2, '기다' (wait) with a sentiment of -2, '기다랗다' (long) with a sentiment of -2, and '기다랗게' (longly) with a sentiment of -2. Other words like '기다랗게' (longly) and '기다랗게' (longly) are also listed with their respective sentiments.

## 단어 매칭

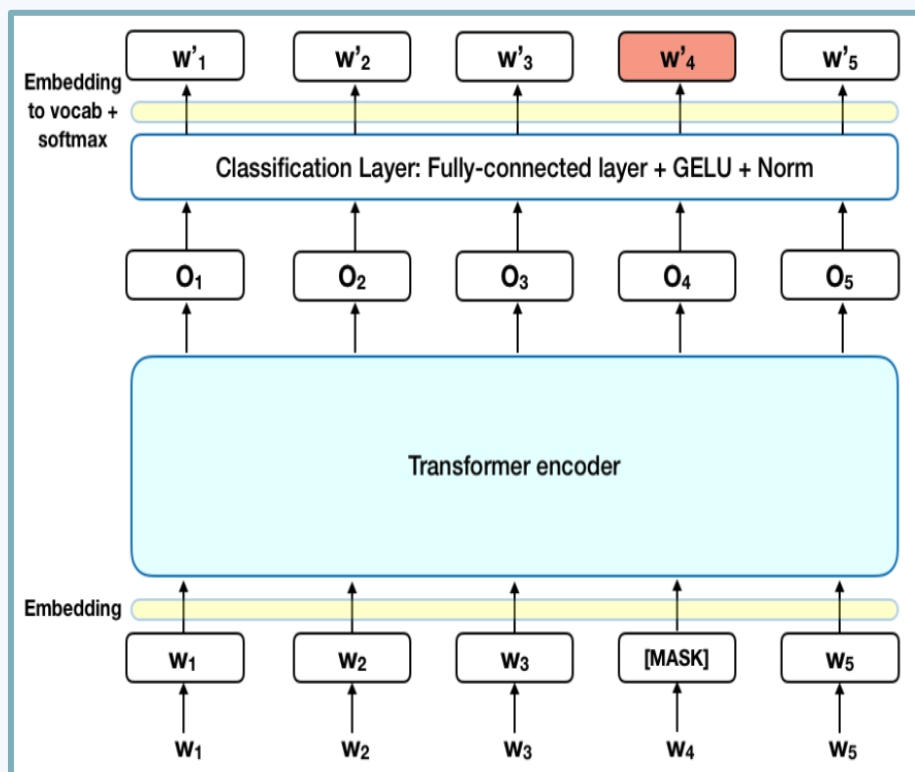


|   | score | text   | class_2 | class_3 |
|---|-------|--|---------|---------|
| 0 | -1    | 기고 치수의 백년대계 하천관리 일원화부터홍수와 같은 재해가 발생하면<br>통상적으로 천...  | 0       | 긍정      |
| 1 | -3    | 김종인 혼자 짊어진 통합당 개혁 어디까지5 18 유공자 예우 강화 법안 발<br>표서...   | 0       | 긍정      |
| 2 | -1    | 단독 값없는 노동 에 막막 2030 연구자들에 주거 연구 공간 준다정부 비정<br>규직...  | 0       | 긍정      |
| 3 | 0     | 경북서 21일 광화문 집회 관련 4명 등 6명 코로나19 추가 확진21일 경북<br>지역... | 0       | 긍정      |
| 4 | 0     | 원주 체조교실발 코로나19 확진자 급증 4명 추가 12명으로 늘어강원 원주<br>시의 ...  | 0       | 긍정      |

## 3) 감성분석

- 2018년 발표된 BERT(Bidirectional Encoder Representations from Transformers)모델을 이용하여 학습
- BERT모델: Transformer의 인코더를 여러 층으로 구성하고, 모델 내부에서 입력 단어의 일정부분을 제거한 후 출력 계층에서 예측하게 함으로써 단어의 임베딩을 학습시켜 기존의 언어학습모델에 비해 문맥을 고려한 단어의 임베딩을 학습시킬 수 있음

## [ BERT모델 도식화 ]






사전학습된 모델에 각 지방의 지역명을 키워드로 한  
새로운 기사를 입력값으로 넣어 긍정과 부정으로 분류














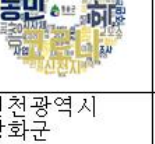






```
Epoch 1/5  
682/682 [=====] - 1144s 2s/step - loss: 0.6889 - accuracy: 0.5535 -  
val_loss: 0.6692 - val_accuracy: 0.5902  
Epoch 2/5  
682/682 [=====] - 1147s 2s/step - loss: 0.6559 - accuracy: 0.6112 -  
val_loss: 0.6439 - val_accuracy: 0.6253  
Epoch 3/5  
682/682 [=====] - 1146s 2s/step - loss: 0.6125 - accuracy: 0.6562 -  
val_loss: 0.6139 - val_accuracy: 0.6656  
Epoch 4/5  
682/682 [=====] - 1145s 2s/step - loss: 0.5266 - accuracy: 0.7336 -  
val_loss: 0.5960 - val_accuracy: 0.6761  
Epoch 5/5  
682/682 [=====] - 1149s 2s/step - loss: 0.3706 - accuracy: 0.8309 -  
val_loss: 0.5518 - val_accuracy: 0.7487
```

#### 4) 결과

- 각 지역의 긍정적인 기사와 부정적인 기사를 각각 워드클라우드 기법을 이용하여 시각화하여 최종 분석에 활용

| 소멸위험이 높은 지방자치단체명(총 67개)   |   |   |   |   |
|---|---|---|---|---|
| 강원도 양양군   | 경상북도 고령군  | 인천광역시 동진군   | 전라남도 해남군  | 충청남도 충청양군   |
|    |    |    |    |    |
| 강원도 영월군   | 경상북도 군위군  | 전라남도 강진군  | 전라남도 화순군  | 충청남도 태안군  |
|    |    |    |    |    |
| 강원도 정선군   | 경상북도 문경시  | 전라남도 고흥군  | 전라북도 고창군  | 충청북도 괴산군  |
|   |   |   |   |   |
| 강원도 태백시   | 경상북도 봉화군  | 전라남도 곡성군  | 전라북도 김제시  | 충청북도 단양군  |
|  |  |  |  |  |

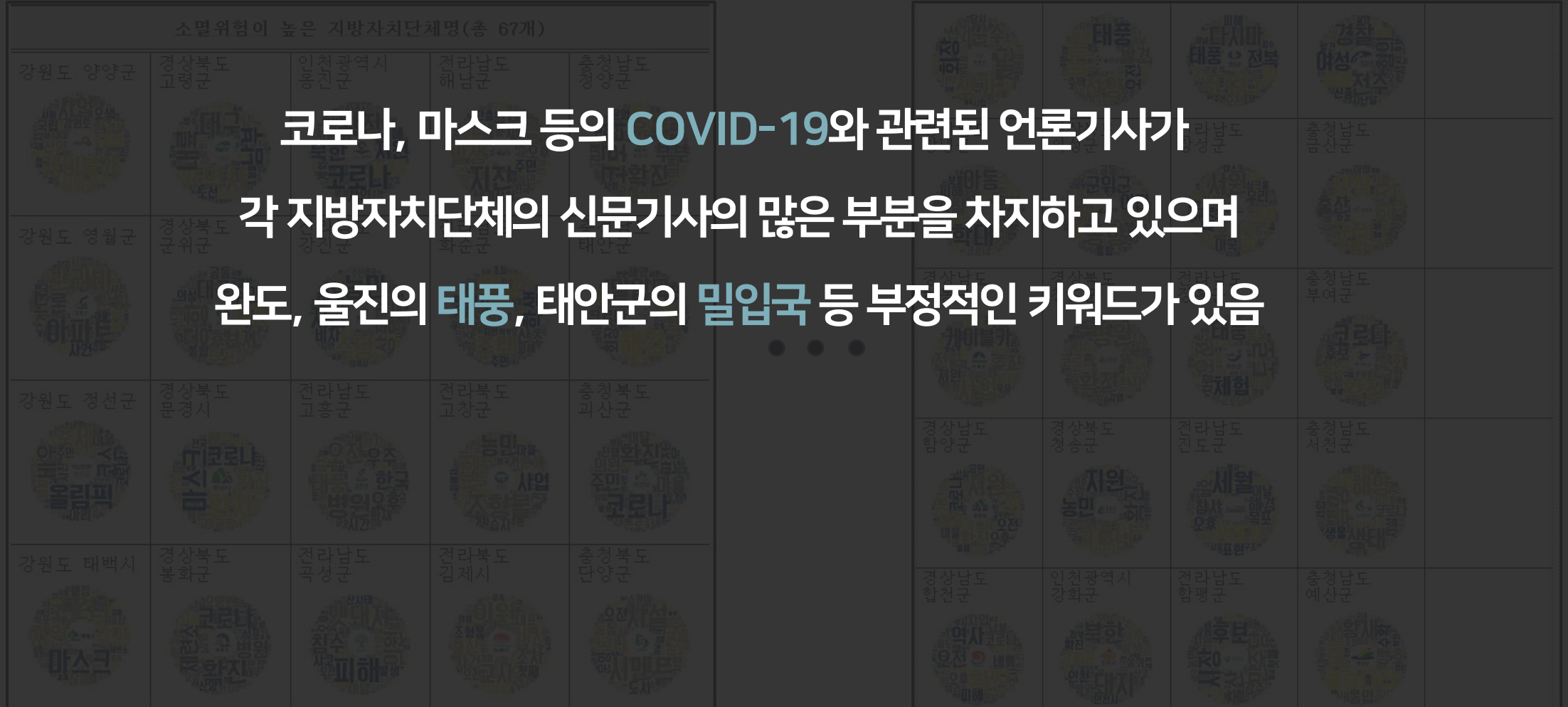
...

|   |   |   |   |  |
|---|---|---|---|--|
|    |    |    |    |  |
|    |    |    |    |  |
|    |    |    |    |  |
|  |  |  |  |  |
|  |  |  |  |  |



#### 4) 결과

- 각 지역의 긍정적인 기사와 부정적인 기사를 각각 워드클라우드 기법을 이용하여 시각화하여 최종 분석에 활용



#### 4) 결과

- 소멸위험 상위 3순위인 지방자치단체에 대한 긍·부정 기사를 워드클라우드화 하여 제시

##### [고성군]

부정적인 키워드:

산불, 수사



##### [서천군]

부정적인 키워드:

장항, 다방, 제련소



#### 4) 결과

- 소멸위험 최상위 지방자치단체 언론기사 긍·부정 감성분석 결과

[고성군]

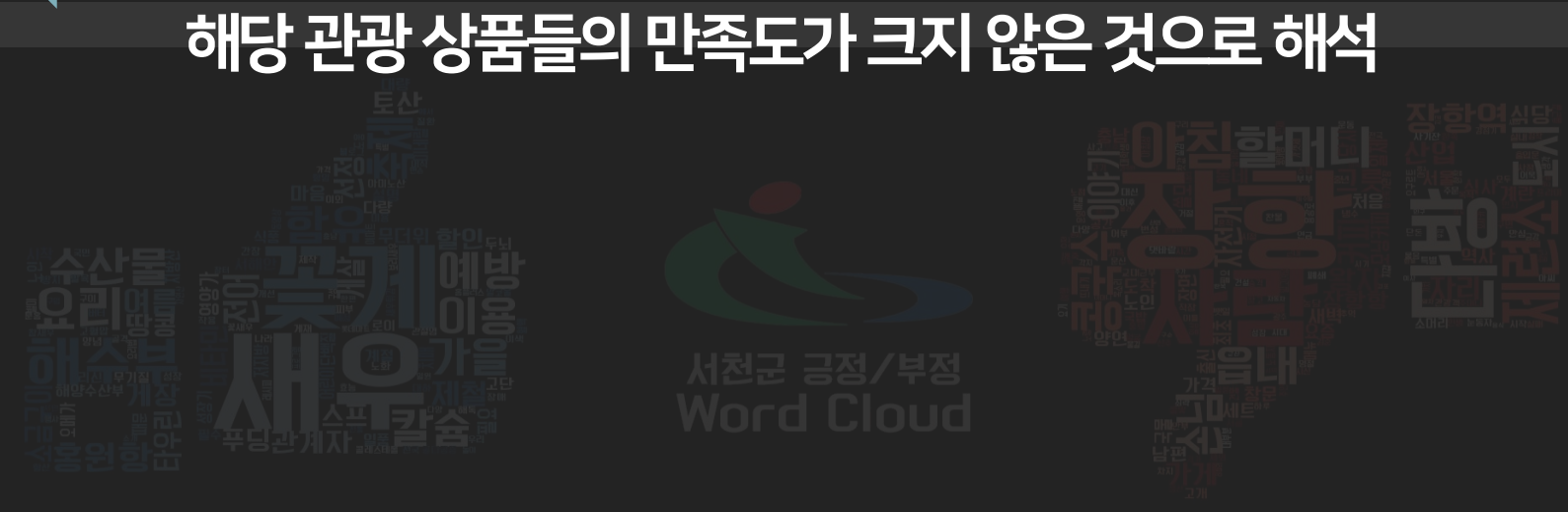
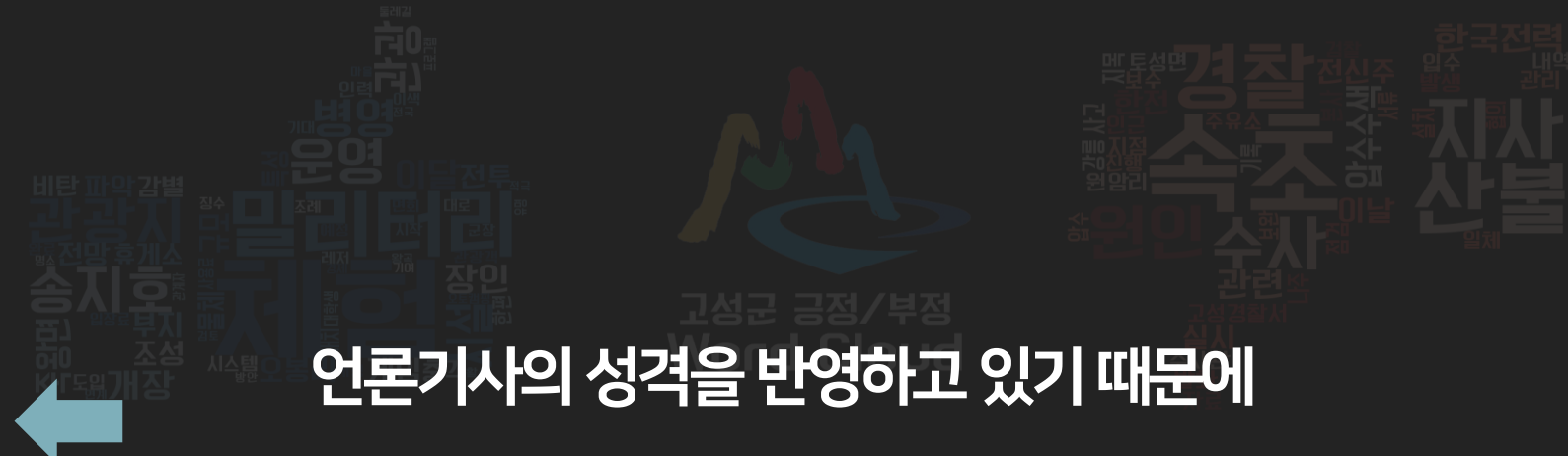
부정적인 키워드:

산불, 수사

[서천군]

부정적인 키워드:

장항, 다방, 제련소



언론기사의 성격을 반영하고 있기 때문에  
해당 관광 상품들의 만족도가 크지 않은 것으로 해석

#### 4) 결과

- 소멸위험 최상위 지방자치단체 언론기사 긍·부정 감성분석 결과

[평창군]

긍정 & 부정 키워드:

평창강, 관광, 체험 축제



고성, 서천, 평창 모두 체험, 관광, 상품 등에 대한 긍정적인 인식이 존재  
그러나 관광 상품의 차별성이 낮고 지역 간 표절도 일어나는 등  
지역 간 관광 상품의 중복이 부정적인 사회적 인식을 이끌어 낼 수 있음을 확인



#### 4) 결과

##### - 비정형 데이터를 이용한 텍스트마이닝 분석 결과

- 소멸위험이 큰 지방자치단체를 대상으로 자료 수집과 분석을 진행하였기 때문에 언론기사들의 긍·부정성이 대체로 **부정적인 쪽으로 기울어짐**
- 최근 COVID-19의 확산으로 관련 키워드가 여러 지방에서 도출
- 돼지 콜레라 등과 같은 **가축 감염병 키워드**가 크게 도출된 지역, **범죄(학대, 강력범죄 등) 관련 키워드**가 크게 도출된 지역  
태풍이나 홍수와 같은 **자연재해 키워드**가 크게 도출된 지역 등이 있었다.

##### ➡ 소멸위험 지방자치단체들의 **사회적 신뢰는 구축되기 어려워 보임**

- 긍정 키워드에 있어야 할 키워드가 부정 키워드에 위치 (EX. 체험, 관광)

##### ➡ **지역의 축제나 관광 상품 등이 무분별하게 생산되면서 해당 상품들이 국민들에게 긍정적인 경험만을 선사하지 않음**

# CONTENTS

## 한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

### I. 연구배경 및 목적

### II. 정량적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료전처리
- 3) 파생변수
- 4) 모델링
- 5) 결과

### III. 정성적 분석

- 1) 데이터 수집 및 탐색
- 2) 자료 전처리
- 3) 감성분석
- 4) 결과

### IV. 결과

## • 정책적 함의

### 1) 경제 기반 시설의 수도권 집중화 현상을 타파

| 소멸위험이 높은 지방자치단체명(총 66개) |           |           |          |          |
|-------------------------|-----------|-----------|----------|----------|
| 강원도 양양군                 | 경상북도 고령군  | 인천광역시 옹진군 | 전라남도 해남군 | 충청남도 청양군 |
| 강원도 영월군                 | 경상북도 군위군  | 전라남도 강진군  | 전라남도 화순군 | 충청남도 태안군 |
| 강원도 정선군                 | 경상북도 문경시  | 전라남도 고흥군  | 전라북도 고창군 | 충청북도 괴산군 |
| 강원도 태백시                 | 경상북도 봉화군  | 전라남도 곡성군  | 전라북도 김제시 | 충청북도 단양군 |
| 강원도 평창군                 | 경상북도 상주시  | 전라남도 구례군  | 전라북도 남원시 | 충청북도 보은군 |
| 경상남도 거창군                | 경상북도 성주군  | 전라남도 담양군  | 전라북도 무주군 | 충청북도 영동군 |
| 경상남도 고성군                | 경상북도 영덕군  | 전라남도 보성군  | 전라북도 부안군 | 충청북도 옥천군 |
| 경상남도 남해군                | 경상북도 영양군  | 전라남도 신안군  | 전라북도 순창군 |          |
| 경상남도 산청군                | 경상북도 영주시  | 전라남도 영광군  | 전라북도 임실군 |          |
| 경상남도 의령군                | 경상북도 울릉군  | 전라남도 영암군  | 전라북도 정읍시 |          |
| 경상남도 창녕군                | 경상북도 울진군  | 전라남도 완도군  | 전라북도 진안군 |          |
| 경상남도 하동군                | 경상북도 의성군  | 전라남도 장성군  | 충청남도 금산군 |          |
| 경상남도 함양군                | 경상북도 청도군  | 전라남도 장흥군  | 충청남도 부여군 |          |
| 경상남도 합천군                | 경상북도 청송군  | 전라남도 진도군  | 충청남도 서천군 |          |
|                         | 인천광역시 강화군 | 전라남도 합평군  | 충청남도 예산군 |          |

- 인구사회학적 변수(평균연령, 인구 증가율, 혼인건수)
  - 지방소멸화지수에 영향을 미치는 주요 변수
  - 경제적 기반인 사업체수와 취업률 등과 밀접하게 연관
  - 의료시설, 문화시설 등과 밀접하게 연관



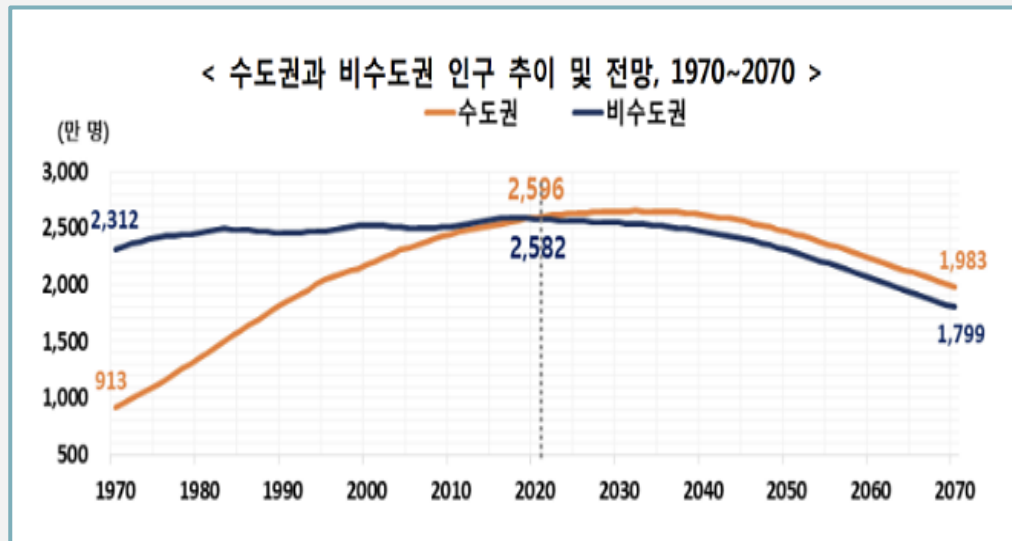
지방 소멸은 결국 인구감소로 인해  
해당 지방자치단체의 인구 공동화 현상이 나타나는 것  
이와 같은 공동화 현상은 기본적으로  
경제 기반 시설의 부재에서 나타날 수 있음을 반증

- 정책적 함의

- 1) 경제 기반 시설의 수도권 집중화 현상을 타파: 공공기관 지방 이전 & 취업 우대 정책

공공기관 지방이전과 청년들이 지방으로 돌아올 수 있도록 하는 취업 우대 정책의 확대를 통해 경제적 기반을 갖추고, 경제 활동인구를 증가시켜야 한다. 이는 상권 발달로 이어질 수 있다

**'수도권 집중화는 계속된다'...올해 첫 수도권 인구, 비수도권 추월**



- 정책적 함의

- 1) 경제 기반 시설의 수도권 집중화 현상을 타파: 산업 유치

공공기관 더불어 중소·대기업에는 세제 혜택이나 각종 세금을 탕감하여 개별 업체가 아닌 산업 전체를 지방으로 유치하는 규모의 경제를 확보해야 한다.

### '드론 산업은 태안이 최적지', 태안 UV 랜드가 뜬다

최형순 기자 | 승인 2020.09.01 09:39 | 댓글 0

28일 김철환 서울지방항공청장 '태안 UV 랜드' 방문

장애물 없고·현재 운영 시설 중 최대규모·수도권 접근성 유리 등 최적의 입지조건 평가 받아

충남 태안군이 4차 산업혁명의 핵심분야인 드론산업 등을 선도하는 지자체로 자리매김하고자 조성 중인 전국 유일의 무인조종 복합테마파크인 '태안 유브이(UV) 랜드'가 드론 산업의 최적지로 평가받으며 큰 관심을 끌고 있다.

### 영덕에 1조원 규모 에너지융합단지 짓는다

입력 : 2020-08-21 04:00:00 | 수정 : 2020-08-20 19:18:36

N  
●  
T  
f  
경북도 "정부 공모사업 선정"  
해상풍력·신재생혁신단지 등 조성  
운영·정비 플랫폼 개발 인력 양성

1조원대 사업비가 투입되는 에너지산업융복합단지가 경북 영덕군에 들어선다. 영덕군은 풍력발전 전에 탁월한 지리적 이점에서 높은 점수를 받아 사업 유치에 성공했다.

경북도는 20일 "산업통상자원부가 주관한 에너지산업융복합단지 지정 공모사업에 영덕군이 최종 선정됐다"고 밝혔다. 이번 사업은 '에너지산업융복합단지 지정 및 육성에 관한 특별법'에 따른 지역 특성에 맞는 에너지를 개발하는 데 목적을 둔다. 경북도는 2025년까지 1조312억원을 에너지산업융복합단지 조성사업에 투자한다.

## • 정책적 함의

### 1) 경제 기반 시설의 수도권 집중화 현상을 타파: 산업 유치

공공기관 더불어 중소·대기업에는 세제 혜택이나 각종 세금을 탕감하여 개별 업체가 아닌 산업 전체를 지방으로 유치하는 규모의 경제를 확보해야 한다.

**지방 소멸 타파의 핵심은 경제적 기반의 확보 후**

**문화, 의료, 관광 시설의 확충까지 이어지는 일련의 과정에 있다**

28일 김철환 서울지방항공청장 '태안 UV 랜드' 방문

장애물 없고·현재 운영 시설 중 최대규모·수도권 접근성 유리 등 최적의 입지조건 평가 받아

충남 태안군이 4차 산업혁명의 핵심분야인 드론산업 등을 선도하는 지자체로 자리매김하고자 조성 중인 전국 유일의 무인조종 복합테마파크인 '태안 유브이(UV) 랜드'가 드론 산업의 최적지로 평가받으며 큰 관심을 끌고 있다.

N

Y

T

F

경북도 "정부 공모사업 선정"

해상풍력·신재생혁신단지 등 조성

운영·정비 플랫폼 개발 인력 양성

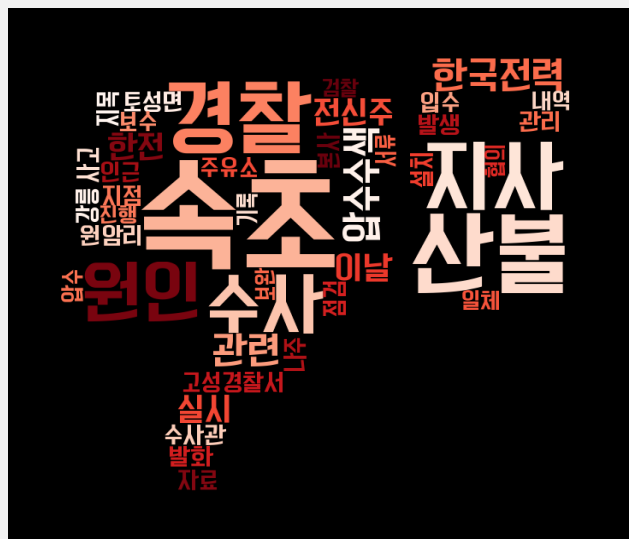
1조원대 사업비가 투입되는 에너지산업융복합단지가 경북 영덕군에 들어선다. 영덕군은 풍력발전에 탁월한 지리적 이점에서 높은 점수를 받아 사업 유치에 성공했다.

경북도는 20일 "산업통상자원부가 주관한 에너지산업융복합단지 지정 공모사업에 영덕군이 최종 선정됐다"고 밝혔다. 이번 사업은 '에너지산업융복합단지 지정 및 육성에 관한 특별법'에 따른 지역 특성에 맞는 에너지를 개발하는 데 목적을 둔다. 경북도는 2025년까지 1조312억원을 에너지산업융복합단지 조성사업에 투자한다.

## 2) 긍정적인 지역 이미지 구축

- 소멸 위험 지방자치단체들은 사회적 자본을 형성하지 못하고 있음
- 경제·산업 불안요소와 COVID-19 감염자 증가 등과 같은 안전 불안 요소 등이 언론 보도를 통해 노출되면 지역 신뢰도와 매력도가 떨어진다

[고성군/ 서천군 / 평창군 부정 키워드]



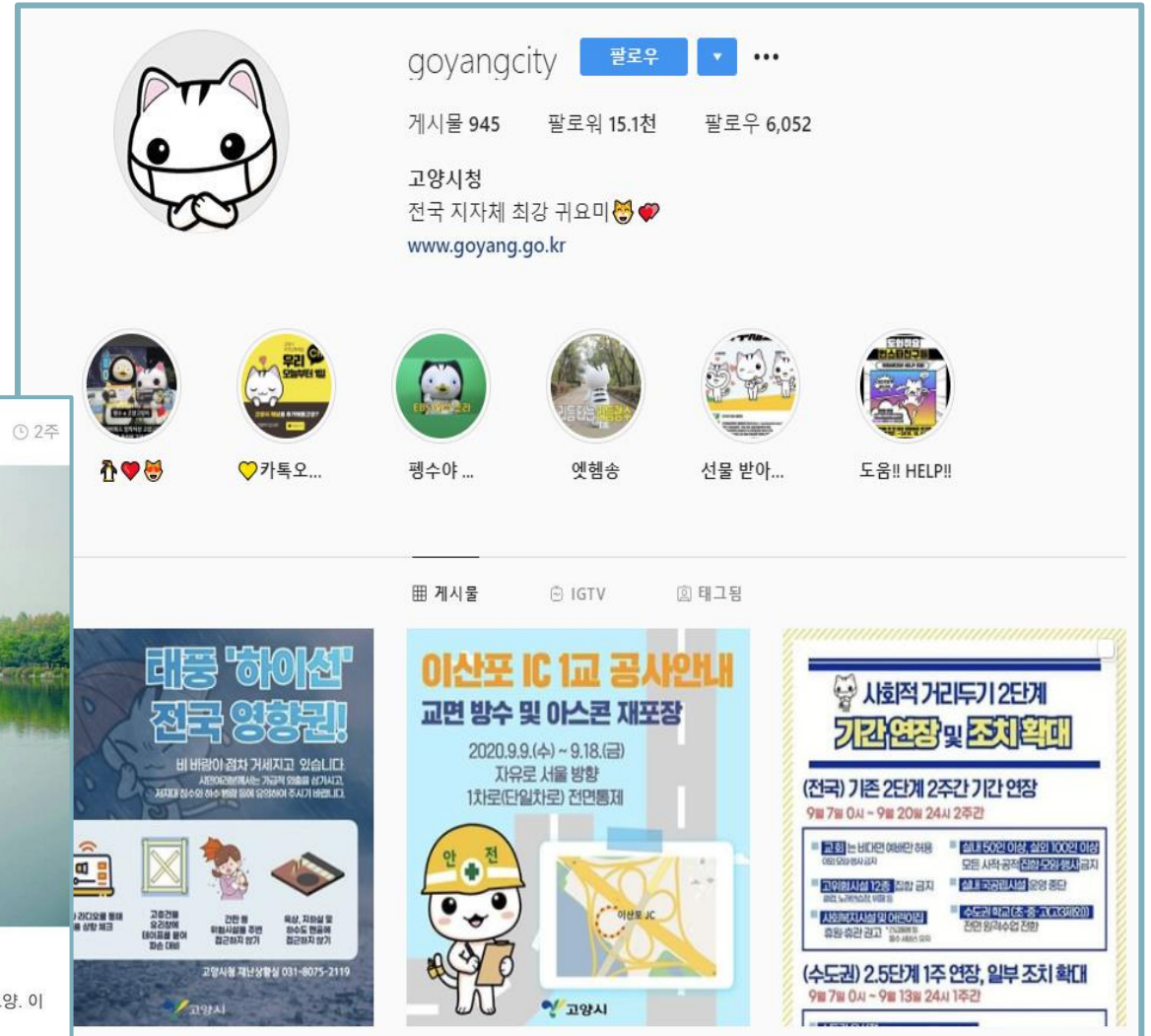
- 정책적 함의

- 2) 긍정적인 지역 이미지 구축 : SNS를 활용한 적극적인 홍보

- 지방자치단체의 긍정적 부분을 부각시키는 소셜 미디어 활동
- 지방자치단체의 관광이나 산업 등에 대한 적극적인 홍보
- 긍정적인 보도자료를 언론사에 지속적으로 제공

지방자치단체에 대한 긍정적 인식도를  
높일 수 있도록 적극적 대응해야 한다

## [ 고양시청 소셜 미디어 활동 ]





- 강병수. (2014). 지역어메니티와 주거이전과의 관련성에 관한연구. 한국도시행정학회 학술발표대회 논문집, 329-351.
- 기영화. (2013). 지방정부 노인일자리사업의 사회적 자본 효과 탐색: 근거이론. 지방정부연구, 17(1), 265-295.
- 김서인, 김동성, & 김중우. (2016). 국내 주요 10 대 기업에 대한 국민 감성 분석: 다범주 감성사전을 활용한 빅 데이터 접근법. 지능정보연구, 22(3), 45-69.
- 김성자, 문하은, & 정성호. (2016). 정부 기능별분류 (COFOG) 의 체계적 고찰 및 활용. [BOK] 국민계정리뷰, 2016(3).
- 김현호, & 오은주. (2007). 어메니티를 활용한 지역발전 방안. 한국지방행정연구원 기본연구과제, 2007, 1-179.
- 남민지, 이은지, & 신주현. (2015). 인스타그램 해시태그를 이용한 사용자 감정 분류 방법. 멀티미디어학회논문지, 18(11), 1391-1399.
- 소진광. (2004). 사회적 자본 형성을 통한 지방자치와 지역발전의 연계화 방안. 지방행정연구, 18(2), 67-91.
- 심재현. (2009). 주거환경과 어메니티: 아파트의 주거만족도를 중심으로. 한국행정과 정책연구, 7(1), 65-83.
- 안명준, 배정한, 주신하, 신지훈, & 이동근. (2008). 농촌 어메니티 경관의 평가체계 개발과 적용-[2007 농촌 어메니티 100 선] 을 중심으로. 농촌계획, 14(2), 77-84.
- 이희태. (2012). 지방정부의 경쟁력 강화를 위한 사회적 자본 확충 전략: 부산광역시 해운대구를 중심으로. 지방정부연구, 16(3), 69-89.
- 임형백. (2012). 농촌 어메니티를 이용한 농촌활성화 정책 방향. 지방행정연구, 26(3), 3-25.
- 진관훈. (2012). 지방정부의 사회적 자본 증대 방안 연구. 지방정부연구, 16(3), 395-412.
- 최길수, 정영윤, & 방정희(2013). 지역단위의 사회적자본 측정 및 관리방안에 관한 연구, 대전발전연구원
- 최예나. (2016). 사회적 자본이 지방정부 신뢰에 미치는 영향 연구: 주민들과 선출직 기관들간 소통의 조절효과를 중심으로. 지방정부연구, 20(3), 69-88.
- 최유진. (2017). 도시어메니티의 지역경제 활성화 효과 분석: 우리나라 기초지방자치단체를 중심으로. 지방정부연구, 20(4), 299-324.
- 하연섭. (2020). 한국 행정: 비교역사적 분석. 다산출판사
- 한국고용정보원. (2020). 지역고용동향브리프 2020년 여름호. 한국고용정보원
- 한상일, & 권소일. (2019). 사회적 기업 인지도와 사회적 자본의 사회적 기업 신뢰에 대한 효과 분석. 사회적경제와 정책연구, 9(2), 33-55.
- 현지연, 유상이, & 이상용. (2019). 평점과 리뷰 텍스트 감성분석을 결합한 추천시스템 향상 방안 연구. 지능정보연구, 25(1), 219-239.

- Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 370-374.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Choi, E. Urban amenities as determinants of selecting a logo type in Korea: the multinomial logit approach with the bootstrap sample. *Quality & Quantity* 46: 391-404.
- Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780), 1612.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
- Futures and Quality of Life. *Partners for Livable Places*, Washington (1984).
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in neural information processing systems* (pp. 3146-3154).
- Gottlieb, P. D.(1994). Amenities as an economic development tool: Is there enough evidence?. *Economic Development Quarterly* 8(3), 270-285.
- Logan, J. R. & Molotch, H. L.(1987). *Urban Fortunes*, Univ. of California Press.
- McNulty, R. H., Jacobson, D. R. & Penne, R. L.: *The Economics of Amenity: Community*

한국 지방소멸에 관한 탐색과 극복 방안에 관한 연구

---

발표를 들어주셔서 감사합니다

---

