



토픽 모델링 결과는 (그림 1)과 같다. 3, 5, 10 번째와 2, 3, 4, 7 번째의 토픽의 경우 토픽끼리 서로 겹치는 부분이 많아 각각을 한 토픽으로 통합하였다. 결과적으로 인구 소멸과 관련된 총 다섯 개의 토픽을 발견할 수 있었고, 이에 대한 주제를 부여하였다. <표 1>에 주제와 주제를 형성하는 단어들을 선별하여 제시한다.

### 3. 인구 소멸 요인 선정

<표 1>을 보면 다섯 개의 주제는 교육, 경제, 문화, 지역, 제도이다. 우리는 인구 소멸과 관련된 주제별 요인으로 ‘교원 1 인당 학생 수(교육)’, ‘의료시설 수(지역)’, ‘폐기물 배출량(경제)’, ‘인구 십만 명당 문화기반시설 수(문화)’, ‘인구 순이동 수(지역)’, ‘혼인 건수(제도)’, ‘노인 천 명당 노인 여가시설 수(문화)’, ‘예산 중 일반 행정 비용의 비중(지역)’의 8 개를 후보로 선정하였다.

<표 1> 5개 토픽의 주제 및 주제와 연관된 단어들

토픽	주제	관련 단어
토픽 1	교육	대학, 교육, 정부, 교육부, 지방대, 경쟁력, 학생
토픽 2 (4개 통합)	경제	경제, 정책, 시장, 계획, 사회, 창출, 기업, 미래, 전략
토픽 3 (3개 통합)	문화	산업, 발전, 문화, 개선, 주민, 관광, 발전, 유치
토픽 4	지역	고향, 지자체, 농촌, 충북, 전남, 강원
토픽 5	제도	답례품, 기부, 고향사랑기부제, 홍보, 제도, 공모

선정한 후보 요인에 대한 시군구 공공 데이터를 수집하였고, 결측 값이 존재하는 인천 남구, 광주 서구, 세종시를 데이터에서 제외하였다. 또 제주시와 서귀포시는 제주특별자치도로 통합되어 있는 경우가 있어 모든 데이터에 대하여 두 개의 시의 평균값을 사용하였다. 정제된 데이터로 요인 간 상관관계를 확인한 결과 ‘혼인 건수’와 ‘의료시설 수’ 사이에 강한 양의 상관관계가 있음을 확인하였다. 이는 군집분석 시 다중공선성의 우려가 있어 ‘혼인 건수’를 제거한 경우와 ‘의료시설 수’를 제거한 경우에 대해 각각 실험을 한 결과, ‘혼인 건수’를 제거한 경우가 군집분석의 실루엣 계수가 높게 나와, ‘혼인 건수’를 제거한 총 7 개의 데이터를 군집분석 용 최종 데이터로 선정하였다.

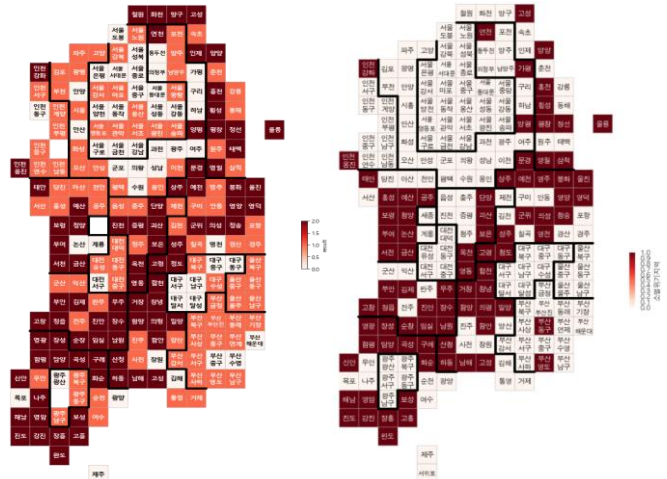
### 4. 군집 분석

군집분석에는 k-means, DBSCAN, Fuzzy C-means 의 세 가지 방법을 비교하였다. 이때 총 7 개의 데이터를 통합·표준화하고, 군집의 개수를 3 개부터 5 개까지 늘려가며 실험하였다. 실루엣 계수와 군집의 분포를 고려한 결과, 군집 개수가 3 개인 k-means 군집분석 결과가 가장 좋은 결과를 보여줬다.

(그림 2)는 군집분석으로 분류된 지역(왼쪽)과 지방 소멸위험지수로 분류된 지역(오른쪽)을 지도에 시각화하여 비교한 그림이다. 짙은 색으로 표시된 지역이 인구 소멸 위험 지역이다. 이처럼 다양한 요인을 고려한 인구 소멸 위험지역 분류 결과가 인구통계학적

요인을 고려한 지방소멸위험지수에 의한 분류 결과와 유사한 것을 확인할 수 있었다.

그러나 차이도 존재하였다. 군집분석 기준에서는 소멸 지역이지만 지방소멸위험지수 기준에서는 소멸 지역이 아닌 지역으로 철원, 화천, 양구, 인제, 태백, 진천, 증평, 부산 동구, 부산 영도, 나주가 있었고, 군집분석 기준에서는 소멸 지역이 아니지만 지방소멸위험지수에서는 소멸 지역으로 분류된 지역으로 삼척, 영주, 음성, 공주, 영천, 논산이 있었다.



(그림 2) 군집분석 결과(왼쪽)와 지방소멸위험지수 기준 인구 소멸 지역(오른쪽). 짙은 색이 인구 소멸 위험 지역임.

### 5. 결론 및 향후 연구

본 연구에서는 지방 소멸과 관련된 요인들을 ‘인구 소멸’을 다룬 신문 기사에 대한 토픽 모델링을 통해 발견하고, 추출된 토픽과 연관된 시군구 공공 데이터를 수집하여 군집분석을 수행함으로써 인구 소멸 위험 지역을 재발견하였다. 이를 통해 인구통계학적 요인 이외에 다양한 요인이 지역의 인구 소멸과 관련이 있음을 확인하였다. 향후에는 좀 더 다양한 공공 데이터를 수집하여 군집분석을 수행하고, 연도별 데이터를 입수하여 시계열로 인구 소멸 위험 지역의 변화를 살펴보고자 한다.

### 6. Acknowledgment

이 논문은 2017 년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017S1A6A3A01078538).

### 참고문헌

- [1] Hiroya Masuda, “Local Extinction: Population Decline Caused by Tokyo-Centric Overconcentration,” Chuko Shinsho, 2014.
- [2] 이상림, 이지혜, Bernhard Köppen, 임소정, 성백선. 지역 인구공동화 전망과 정책적 함의, 한국보건사회연구원 연구보고서, 131 쪽, 2018.
- [3] 유한별, 탁근주, 문정승. 한국 지방소멸 요인과 극복 방안에 관한 연구: 머신러닝 방법을 통한 탐색. 지방정부연구, 24(4), pp. 443-476, 2021.