

제 21대 국회의원선거 정책선거 연구용역 결과보고서

-제21대 국회의원선거 공약이슈 파악을 위한 민원빅데이터 분석

디비디스커버코리아

2019. 12. 31.



CONTENTS

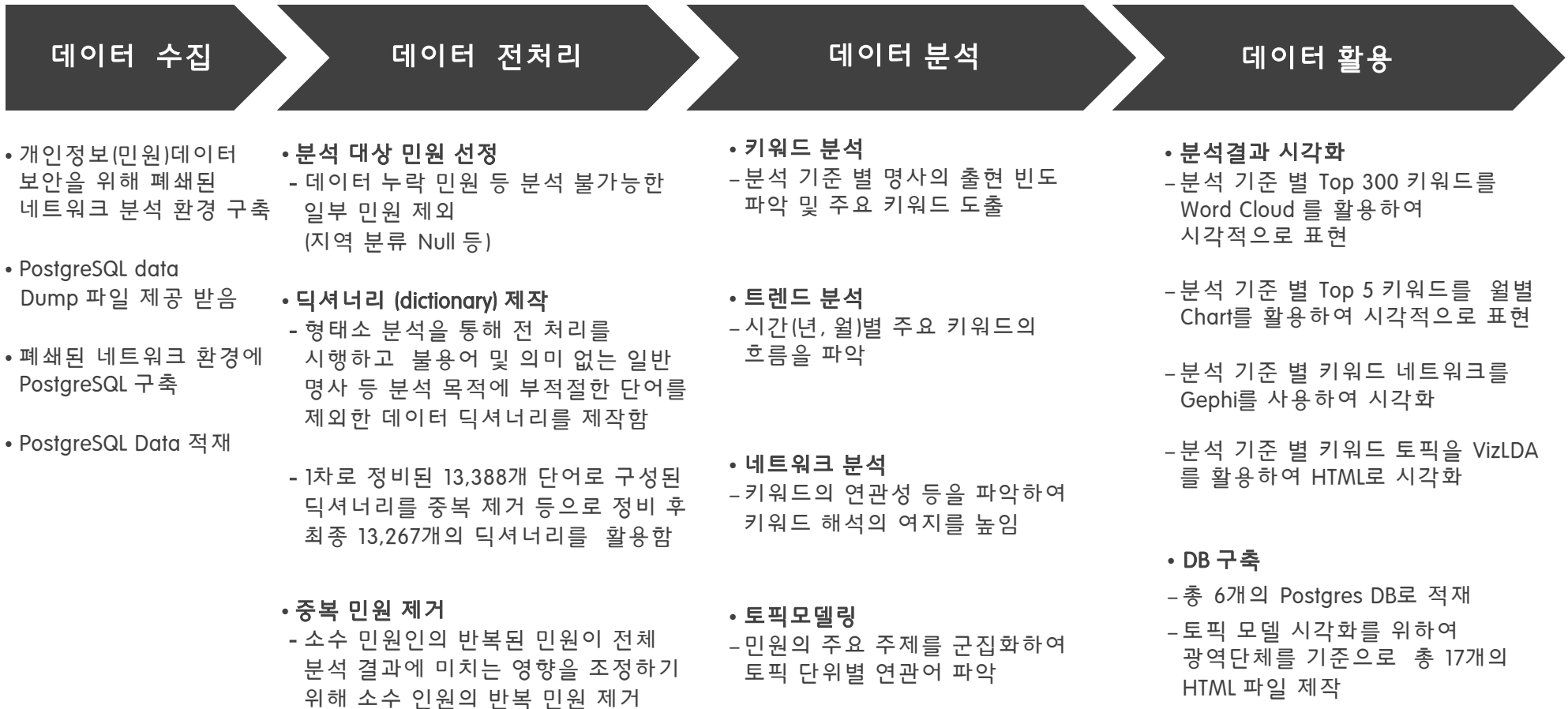
- I. 사업 개요
- II. 분석 개요
 - 1. 분석 프로세스
 - 2. 분석 대상 데이터 현황
 - 3. 분석 대상 데이터 검토 결과
- III. 분석 방법
 - 1. 분석 기준
 - 2. 키워드 분석
 - 3. 트렌드 분석
 - 4. 네트워크 분석
 - 5. 토픽 모델링
- IV. 분석 결과
 - 1. 민원 현황
 - 2. 대한민국 분석 결과
 - 3. 지역별 분석 결과 – 17개 광역단체
 - 4. 심층분석사례 – 부산시 해운대구

- 주제 : 제21대 국회의원선거 정책공약이슈 파악을 위한 빅데이터 분석
- 기간 : 2019 11.18 ~ 2019. 12. 31
- 대상: 제20대 국회의원 임기 중 국민신문고 민원 데이터
 - 연구시점을 감안하여 2016. 6. ~ 2019. 9. 로 데이터 한정
- 의의: 국민신문고 민원은 국민들이 공적 해결이 필요하다고 생각하는 사회문제에 대해 적극적으로 자신의 의견을 제시하는 것으로 민원 빅데이터를 분석하여 국민의 관점에서 사회문제가 무엇인지를 '빈출 키워드' 중심으로 정리하고 키워드의 맥락 파악을 돕는 추가 분석 자료를 제공하여 정당·후보자의 정책·공약 개발에 반영하는 한편, 유권자도 우리 사회의 현안에 대한 정보를 공유하고 이를 바탕으로 정책공약을 평가하는 자료로 활용할 수 있음.
연구결과는 공약이슈지도 홈페이지를 통해 공개될 예정이며 학회·언론 등의 해석과 보도를 통해 정책선거 실현을 위한 자료로서의 가치를 제고할 수 있음.

1. 분석 프로세스

II. 분석개요

- 텍스트 전처리는 데이터 사전 검출 방식을 활용하여 수행하여 데이터 프레임을 개발
- 해당 데이터 프레임을 활용하여, 키워드 분석, 트렌드 분석 및 네트워크 분석의 시각화 및 DB 구축 시행
- 별도로 토픽 모델을 개발하여 HTML 시각화 및 DB 구축



2. 분석 대상 데이터 현황

II. 분석개요



- 대한민국의 대표적인 민원 접수신청 경로는 국민신문고, 안전신문고, 생활불편신고, 스마트 국민제보, 한국바로알림서비스, 규제개혁신문고, 복지로 등이 있음.
- 이 중 안전신문고 등 다양한 범정부포털은 별도의 홈페이지를 가지고 있으나 신청화면은 모두 국민신문고를 쓰고 있어 모두 국민신문고에 포함되는 개념임.

국민신문고에 포함된 정부포털	운영기관
안전신문고	행정안전부
생활불편신고	행정안전부
복지로	보건복지부
규제개혁신문고 (규제정보포털)	국무조정실

- 국민권익위 민원정보분석시스템은 국민신문고 민원과 지자체 홈페이지를 통해 수집되는 민원을 포함하여 관리하고 있으며 이번 분석에서는 그 중 국민신문고 민원만을 분석대상으로 함.
- 2016년 6월부터 2019년 9월까지 약 1300만 개가 넘는 민원 데이터를 분석대상으로 하고, 성별 연령 지역 등 청원인의 속성이 함께 관리되는 데이터의 특성을 활용하여 청원인 속성별 민원 분석을 실시함.

3. 분석 대상 Data 검토 결과

II. 분석개요

- 6개 테이블 중 민원신청 테이블을 기준으로 문서 및 발생지, 신청지를 붙여서 사용
- 발생지 및 신청지의 EMD_CODE는 T_CO_ANC_INFO 테이블에 tot_anc_name_v(명칭)을 붙여서 사용

	테이블	테이블명	건수	비고
1	T_PC_PETITION_DAMO	민원신청	30,370,368	민원신청일자, 민원인생년월일,
2	CRWL_DOC	문서	67,975,926	민원 내용,성별
3	TFPC0409	민원발생지	15,000,000	민원발생지
4	TFPC0910	민원신청지	24,662,013	민원신청지
5	T_TO_CODE	코드정보	2,312	코드 정보
6	T_CO_ANC_INFO	기관정보	376,606	지역코드 이름

T_PC_PETITION_DAMO(민원신청)

PETI_NO_C	OPEN_YN_C	REG_D	BIRTHDAY_C	PA_PETI_NO_C	PA_PETI_RELATION_C	PA2_PETI_NO_C
1AA-1611-047683	N	2016-11-08 22:12:12.0	19870627	0000000000000000	[null]	[null]
1AA-1611-121798	N	2016-11-21 10:01:35.0	[null]	0000000000000000	[null]	[null]
1AA-1701-008260	Y	2017-01-02 21:50:24.0	19900515	0000000000000000	[null]	[null]
1AA-1702-118737	Y	2017-02-21 12:59:35.0	19871005	0000000000000000	[null]	[null]
1BA-1702-121557	N	2017-02-21 18:14:13.0	[null]	0000000000000000	[null]	[null]

CRWL_DOC(문서)

DOC_SEQ	SITE_ID	TASK_ID	SITE_CATEGORY_ID	TASK_CATEGORY_ID	DOC_ID	URL	TITLE	CONTENT	EXTEND_FD4	EXTEND_FD14
137971548	1	1	1001	1	1AA-1611-047683	/	불법 파헤친 신고	최근다 불동 토대오 아...	2	00010001
138357569	1	1	1001	1	1AA-1611-121798	/	11월12일 도로 외측 난간...	먼저 도로 관리로 불합...	-	00010001
139599921	1	1	1001	1	1AA-1701-008260	/	미성년자 대상 성매매 신고...	공공화점이 있어서 필...	-	00010001
140932411	1	1	1001	1	1AA-1702-118737	/	선행 자살이 위협 운전으로...	안녕하세요	1	00010001
140950319	1	1	1001	1	1BA-1702-121557	/	안전관리대행업 변경등록...	안전관리대행업 변경등록...	-	00010001

TFPC0409(민원발생지)

CVPL_REQST_NO	EMD_CODE	BRTC_CODE	LOAD_DT	CVPL_REQST_DE	RNUM	TOTCNT
1AA-1611-121798	Z070020	Z070020	2016-11-21...	20161121	7813839	15007058

TFPC0410(민원신청지)

CVPL_REQST_NO	EMD_CODE	BRTC_CODE	LOAD_DT	CVPL_REQST_DE	RNUM	TOTCNT
1AA-1611-047683	3330030	6260000	2016-11-08...	20161108	13313949	24667827
1AA-1611-121798	3330000	6260000	2016-11-21...	20161121	13503855	24667827
1AA-1701-008260	3330000	6260000	2017-01-02...	20170102	13638781	24667827
1AA-1702-118737	3330000	6260000	2017-02-21...	20170221	13887822	24667827
1BA-1702-121557	3330000	6260000	2017-02-21...	20170221	13837161	24667827

T_CO_ANC_INFO(기관정보)

anc_code_v	tot_anc_name_v
3330000	부산광역시 해운대구
3330030	부산광역시 해운대구 좌동

(부산광역시 해운대구 예시)

3. 분석 대상 Data 검토 결과

- 6개 테이블을 통해 민원 키워드분석에 필요한 컬럼 13개를 다음과 같이 선정하였음

컬럼명	컬럼명	테이블	비고
doc_id	민원신청번호	CRWL_DOC	
minwon_type	민원유형	CRWL_DOC	
gender	성별	CRWL_DOC	
reg_d	신청일	T_PC_PETITION_DAMOr	
city	시도	TFPC0410,TFPC0409	신청지가 null인 경우 발생지로 대체
county	시군구	TFPC0410,TFPC0409	신청지가 null인 경우 발생지로 대체
birtyday_c	생년월일	T_PC_PETITION_DAMOr	
age_code	연령대	T_PC_PETITION_DAMOr	
year	년	T_PC_PETITION_DAMOr	
Month	월	T_PC_PETITION_DAMOr	
ym	년월	T_PC_PETITION_DAMOr	
title	제목	CRWL_DOC	
content	내용	CRWL_DOC	

민원신청번호	민원유형	신청일	년	월	년월	생년월일	연령대	성별	시도	시군구	제목	내용
doc_id	minwon_type	reg_d	year	month	ym	birthday_c	age_code	gender	city	county	title	content
1AA-1702-118737	10001	2017-02-21 12:59:35	2017	2	201702	19871005	3	1	부산광역시	해운대구	선형 차량이 위법 운전	안녕하세요. 도로에서 선형차
1AA-1611-121798	10001	2016-11-21 10:01:35	2016	11	201611		9999	-	부산광역시	해운대구	11월12일 도로 외측 난	먼저 도로 관리로 불철주야 고
1AA-1701-008260	10001	2017-01-02 21:50:24	2017	1	201701	19900515	2	-	부산광역시	해운대구	미성년자 대상 성매매	궁금한 점이 있어서 질문드립
1AA-1611-047683	10001	2016-11-08 22:12:12	2016	11	201611	19870627	2	2	부산광역시	해운대구	불법 좌회전 신고	해운대 중동 로데오 아울렛 오
1BA-1702-121557	10001	2017-02-21 18:14:13	2017	2	201702		9999	-	부산광역시	해운대구	안전관리대행업 변경	안전관리대행업 변경등록 신

1. 분석 기준

- 정책 결정에 필요한 유용한 관점과 실제 분석 가능한 자원을 고려하여 내부 협의
- 각 분석 내용 별로 필요한 기준(지역분류, 성별, 연령, 기간별)을 고려하여 분석 및 모델 개발을 시행

분석 기준	분석 값	합계
성별	남성 :1 , 여성:2	총 3개 분류
연령	10대: 1 , 20대: 2, 30대: 3, 40대: 4, 50대: 5, 60대: 6, 70대: 7	총 8개 분류
지역	광역단체 _ 시도 17개, 기초단체 _ 시군구 226개	총 243개 분류

분석 내용	기준
키워드 분석	지역 분류 (243개)X 성별(3개) X 연령(8개)
트렌드 분석	지역 분류 (243개)X 성별(3개) X 연령(8개)
네트워크 분석	지역 분류 (광역단체 _ 시도 17개, 기초단체 _ 시군구 226개)
토픽 모델링	지역 분류(광역단체 _ 시도 17개)

2. 키워드 분석

- 특정 기간 및 특정 그룹에 출현 빈도가 높은 키워드를 추출하여, 가장 출현 빈도가 높은 단어를 직관적으로 이해 가능한 분석 방법.
- 직관적인 이해를 위해 보통 워드 클라우드 형태의 시각화를 활용함

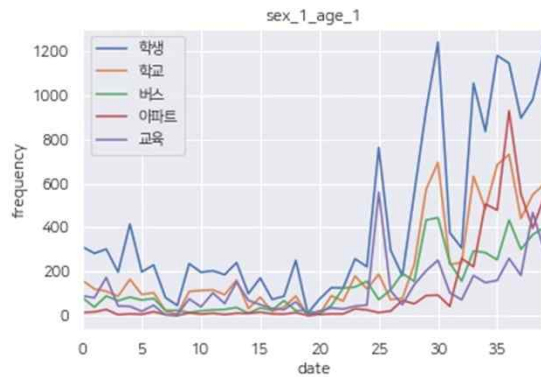
순위	단어	빈도	순위	단어	빈도
1	학생	16,603	11	입주민	2,201
2	학교	8,637	12	차별	2,195
3	버스	5,487	13	은계지구	2,050
4	아파트	4,554	14	대학교	2,010
5	교육	4,548	15	병역	1,901
6	고등학교	3,991	16	신체검사	1,883
7	분양	3,427	17	특수교사	1,592
8	시험	3,236	18	CCTV	1,423
9	교사	3,028	19	하자	1,420
10	특수교육	2,408	20	건축물	1,318
11	입주민	2,201	20	건축물	1,318

- 키워드(keyword) 분석은 단어, 문장, 문단, 문서 등으로 분석 단위를 지정하고, 단위 별 텍스트를 분석하는데 쓰이는 방법
- 특정 기간 및 특정 분석 단위에 단어가 출현한 빈도를 색상과 크기로 시각화하여 가장 크게 언급된 단어를 직관적으로 파악할 수 있음.
- 이를 위해 분석 단위별로 형태소(形態素, morpheme) 분석을 수행해야 함. 형태소란, 의미를 가지는 최소 단위로 정의되는데, 여기에서 의미는 어휘적인 부분 뿐만 아니라 문법적인 부분까지 포함. 형태소 분석을 위해서 자연어 처리 기술이 필요. 자연어 처리(Natural Language Processing, NLP)는 텍스트에서 의미 있는 정보를 분석, 추출하고 이해하는 일련의 기술 집합.

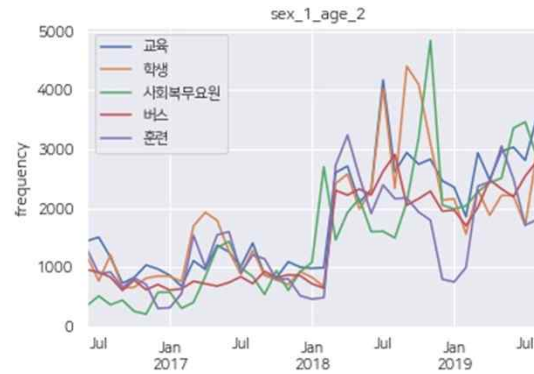


3. 트렌드 분석

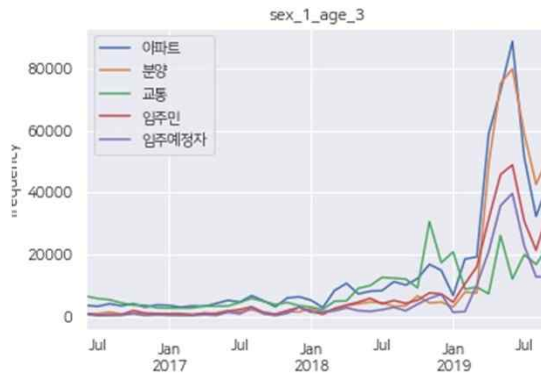
- 특정 기간 및 특정 그룹에 출현 빈도가 높은 키워드를 추출하여, 특정 기간별로 출현 빈도를 시간의 흐름에 따라 분석이 가능한 방법론임
- 시간의 흐름으로 빈도를 분석하는 시계열 분석을 많이 활용함



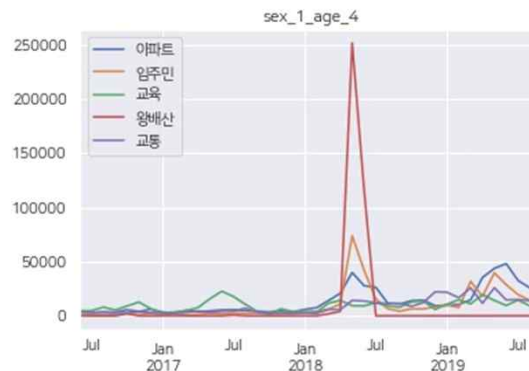
트렌드 10대



트렌드 20대



트렌드 30대 이상



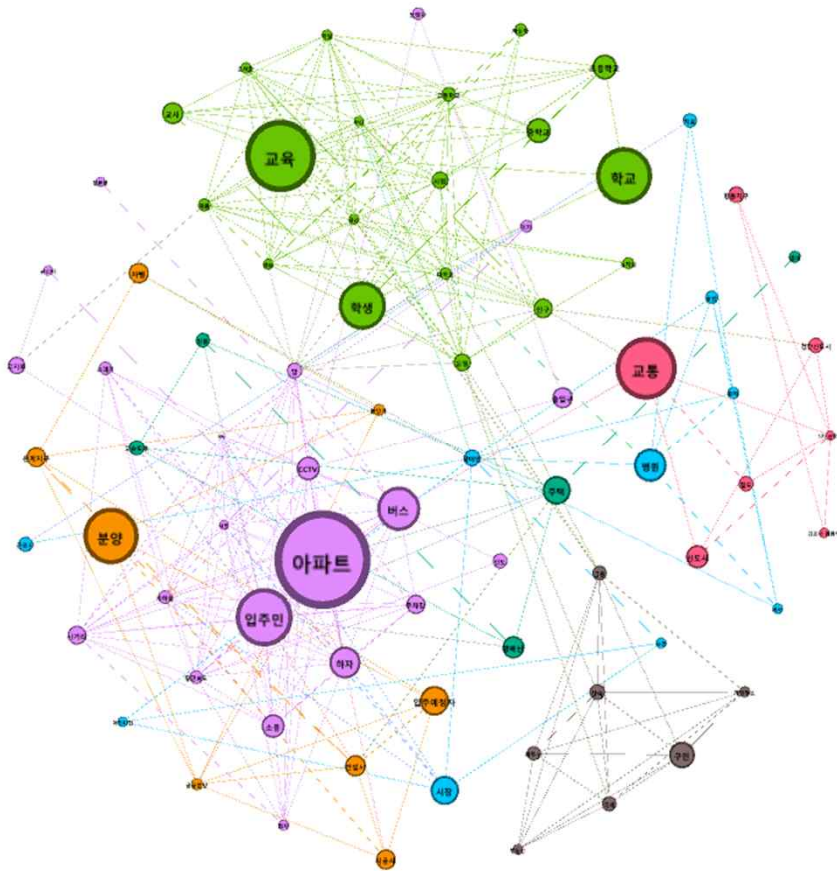
트렌드 40대

트렌드 (trend) 분석이란, 시간(년, 월, 일, 시간 등)의 흐름에 따라서 특정 키워드가 얼마나 많이 출현했는지 파악하고, 이를 바탕으로 어느 시점에 이슈가 되었는지 확인하는 일련의 과정임.

- 특정 키워드의 장기간 검색 추이 정보를 제공하며 특정 기간에 특이하게 발생한 사건이나 이슈를 확인 하는 분석 기법
예를 들어 연령별로 키워드를 분석 할 경우 실제 연령별 관심사 및 관심 시기를 개별적으로 파악 가능한 분석 기법임

4. 네트워크 분석

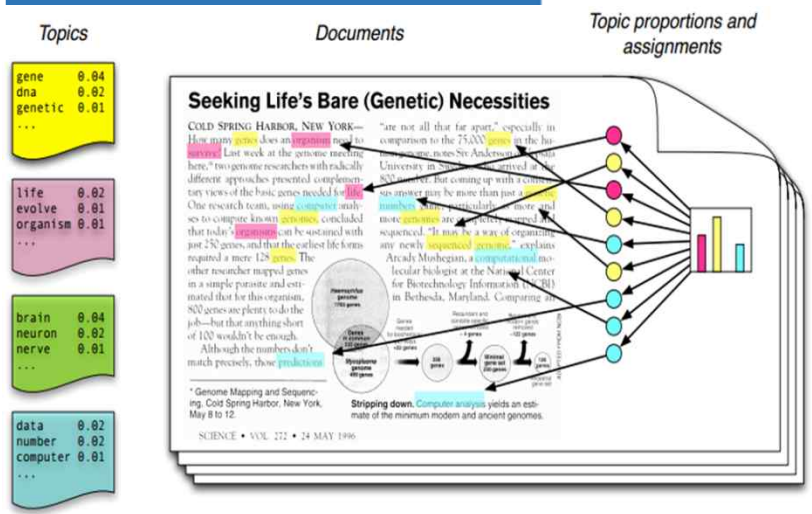
•개체(Node)간의 상호 유사성을 파악하여 관심사를 몇가지 대분류로 분류하고 상호 연관성 높은 관계를 시각적으로 확인하는 분석 방법.



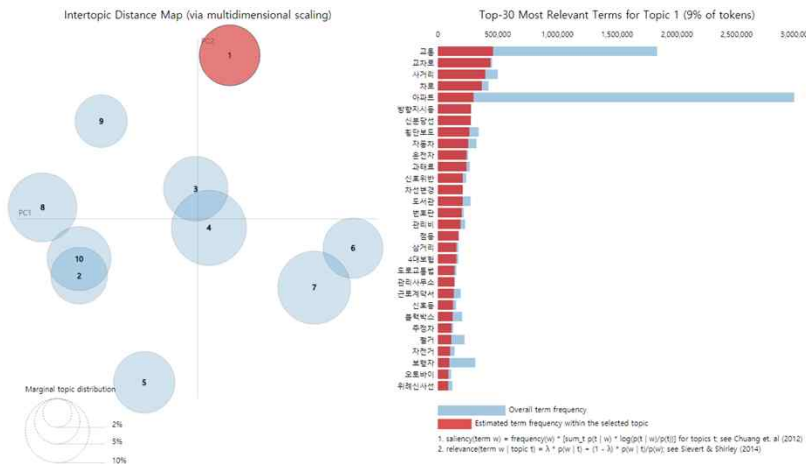
- 네트워크 분석 (Network analysis)은 개체간의 상호 관계에 초점을 맞춘 방법으로 구조나 확산 및 진화과정을 계량적으로 분석하는 방법으로 텍스트의 내용을 가시적, 구조적으로 이해할 수 있음.
- 노드 (Node)와 링크 (Link)라는 두개의 요소로 이루어진 네트워크 그래프 (Network graph)로 표현. 이러한 기본적인 요소를 각 분야에서 응용할 때 표현되는 노드와 링크의 이름과 형태, 속성을 추가 혹은 변형해서 사용.
- 예를 들어, 현재 대한민국 국민의 주요 관심사는 아파트, 분양, 교통, 교육 의 크게 넷의 영역에 관심사가 집중하고 있음

5. 토픽 모델링

• 비구조화된 텍스트 자료들의 뭉치로부터 의미 있는 주제(토픽)를 추출해주는 확률 모델



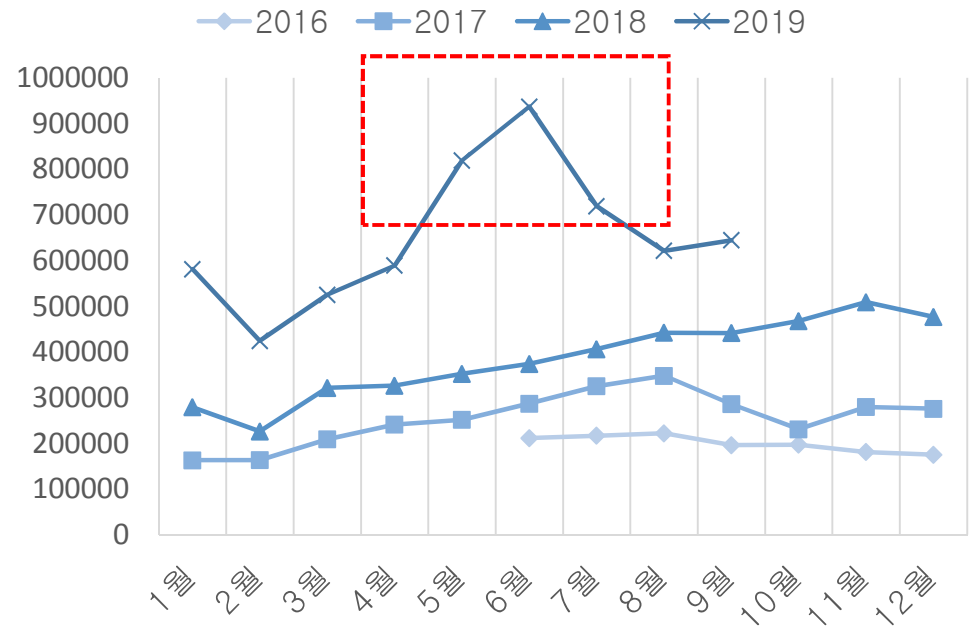
- 토픽 모델링(Topic Modeling)은 데이터 마이닝 기법 중의 하나로 비구조화된 텍스트 자료들의 뭉치로부터 의미 있는 주제(토픽)들을 추출해주는 확률모델 알고리즘.
- 방대한 텍스트 자료로부터 맥락과 관련된 단서들을 이용하여 해석가능성이 높은 주제들을 추출해 주는 특성이 있음.
- 토픽 모델링 기법 가운데 잠재디리클레할당(Latent Dirichlet Allocation, LDA)은 주어진 문서에 대하여 각 문서에 어떤 주제들이 존재하는지에 대한 확률모형으로 토픽별 단어의 분포, 문서별 토픽의 분포를 모두 추정.
- 잠재적인 주제(topic)들이 무작위 혼합으로 문서를 이루고 있으며, 그 주제의 특성이 텍스트에 분포해 있는 키워드(word)들의 패턴을 통해 추론되는 것



1. 전체 연도별/월별 민원 현황

- 2016년 6월부터 2019년 9월까지의 데이터
- 해마다 민원 건수가 꾸준히 증가해 왔으며, 2019년은 6,7월에 민원이 큰 폭으로 증가했다가 다시 예전의 상황으로 복귀하는 추세.

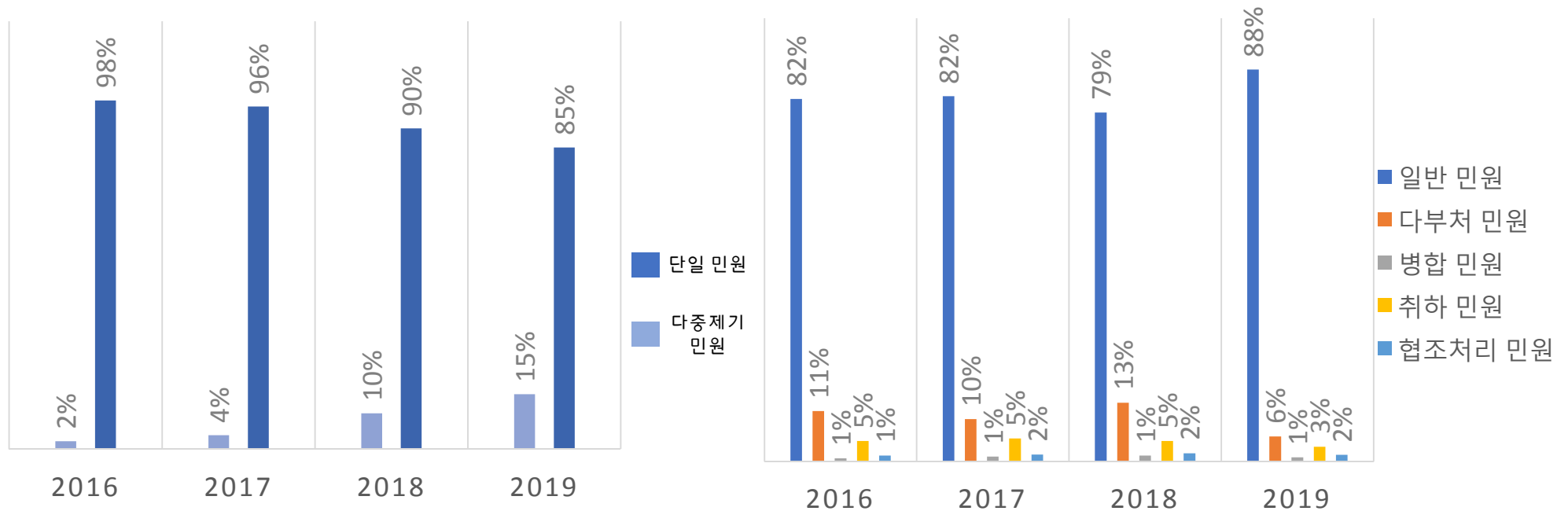
	2016	2017	2018	2019
1월		162993	278999	581184
2월		163375	226029	424256
3월		209021	321324	525074
4월		241340	326326	589188
5월		251806	352148	819186
6월	211515	287105	373959	937162
7월	216866	325164	406122	719025
8월	221837	347988	442105	621640
9월	196044	285866	441545	644652
10월	197084	231163	467912	
11월	181131	279580	509215	
12월	174891	275933	476825	
총 건수	1399368	3061334	4622509	5861367



2. 유형별 민원 현황

IV. 분석결과_민원현황

- 해마다 전체 민원에서 다중제기민원의 비율이 증가하는 추세
- 일반 민원이 전체 민원의 대부분이며, 2019년에는 다부처 민원이 소폭 감소 추세임



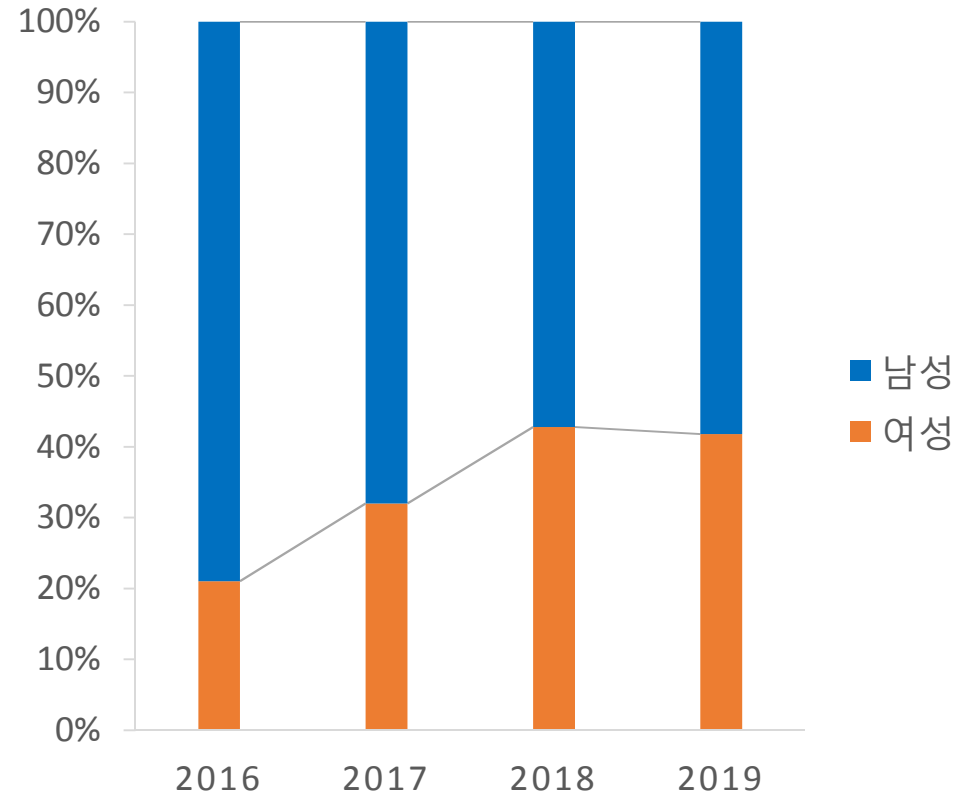
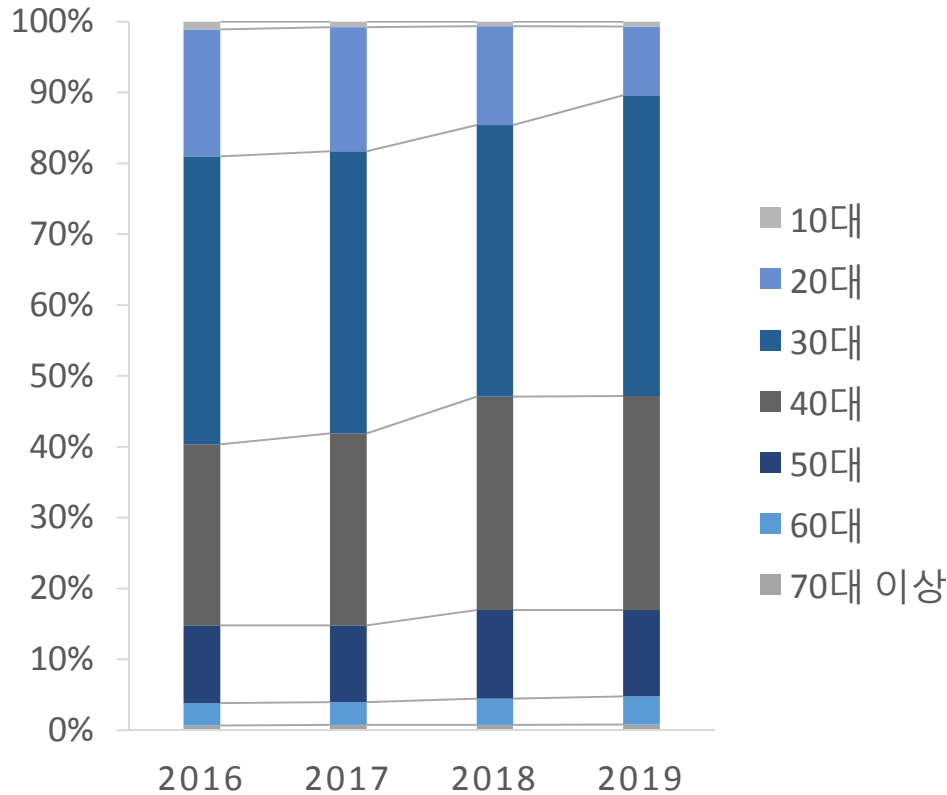
- 단일민원 : 일반적인 민원
- 다중제기민원: 소수에 의해서 동일한 내용을 반복적으로 제기한 민원

- 일반민원 : 신청된 하나의 민원을 하나의 기관(=업무 소관기관)에서 처리하는 민원
- 다부처민원 : 신청된 민원에 여러 가지 내용이 포함되어 있어 각각의 소관기관이 접수하여 회신하는 민원
- 병합민원 : 같은 내용의 민원이 여러 건 접수되어 하나로 묶어서 처리한 민원으로 중·반복민원과 유사
- 취하민원 : 민원 신청 이후 민원인이 취하한 건
- 협조처리민원 : 민원을 처리하는 주무기관 또는 부서에서 관계 기관·부서의 협조 또는 의견을 받아 처리하는 민원

3. 성별/연령별 인원 현황

IV. 분석결과_인원현황

- 30대와 40대 인원 비중이 크며, 20대 인원 비율의 감소 추세
- 남성 인원 비중이 전통적으로 크나, 2018년 이후 여성 인원이 증가하는 추세임



4. 지역별 민원 데이터 현황

- 경기도, 서울특별시, 인천광역시 순으로 민원이 많이 발생하며, 연도별로 꾸준히 민원이 증가하는 추세임
- 특히 2019년도는 9월까지 산정되었음에도 5,800,000건으로 연도별로 민원이 증가하는 추세임

시도별 민원 수

시도	민원 수
경기도	5,190,612
서울특별시	2,564,164
인천광역시	780,376
부산광역시	609,629
대구광역시	425,799
경상북도	387,596
경상남도	370,896
대전광역시	313,773
전라북도	291,589
충청남도	275,982
광주광역시	270,041
충청북도	249,372
전라남도	241,487
강원도	189,032
울산광역시	173,142
제주특별자치도	79,545
세종특별자치시	23,814
기타	2,507,729
총합계	14,944,578

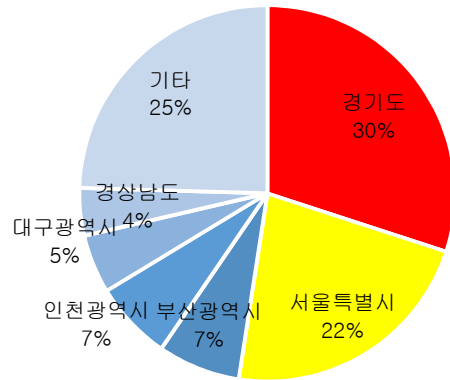
시도,연도 별 민원 수

시도	2016	2017	2018	2019	총합계
경기도	404,127	820,440	1,557,317	2,408,728	5,190,612
서울특별시	295,590	535,916	758,043	974,615	2,564,164
인천광역시	89,406	151,573	237,702	301,695	780,376
부산광역시	93,686	172,408	172,789	170,746	609,629
대구광역시	66,394	113,358	117,258	128,789	425,799
경상북도	56,999	122,974	106,063	101,560	387,596
경상남도	53,963	97,500	107,616	111,817	370,896
대전광역시	48,398	79,128	92,469	93,778	313,773
전라북도	35,580	76,279	83,457	96,273	291,589
충청남도	43,830	71,564	79,641	80,947	275,982
광주광역시	53,655	68,423	74,717	73,246	270,041
충청북도	33,909	64,936	67,715	82,812	249,372
전라남도	33,897	63,650	70,741	73,199	241,487
강원도	26,736	50,249	56,193	55,854	189,032
울산광역시	26,529	46,519	50,251	49,843	173,142
제주특별자치도	12,030	19,875	24,129	23,511	79,545
세종특별자치시	2,774	5,051	6,084	9,905	23,814
기타	21,865	501,491	960,324	1,024,049	2,507,729
총합계	1,399,368	3,061,334	4,622,509	5,861,367	14,944,578

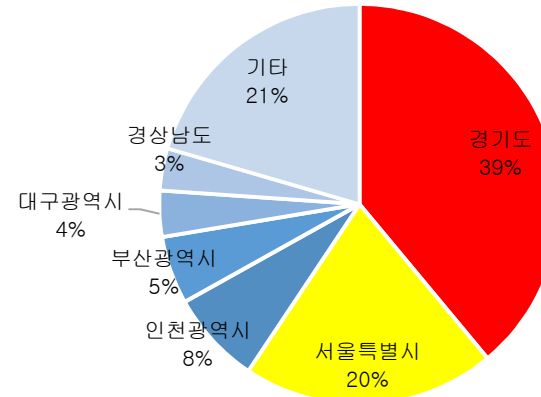
4. 지역별 민원 데이터 현황

IV. 분석결과_민원현황

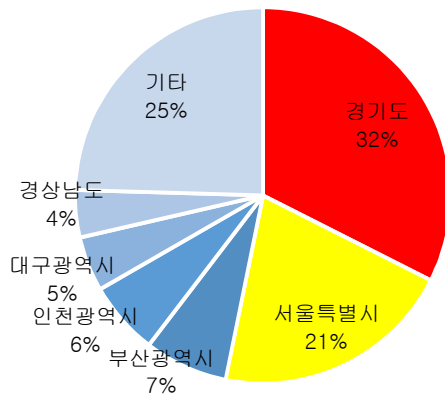
- 서울, 경기 지역이 절반 이상으로 2018, 2019년도로 갈수록 민원 점유율이 높아지는 추세임.
- 부산, 인천 지역의 점유율이 서울, 경기도 다음으로 높음



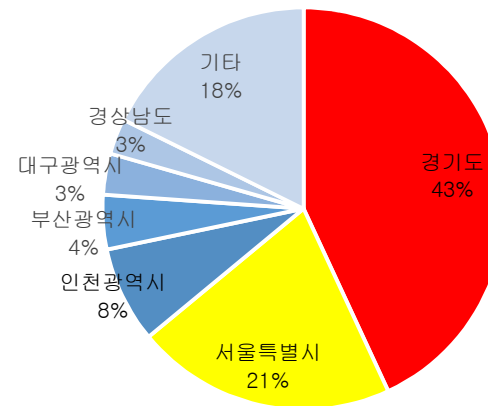
2016



2018



2017

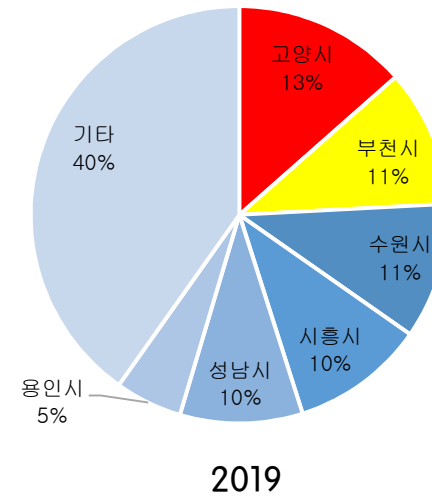
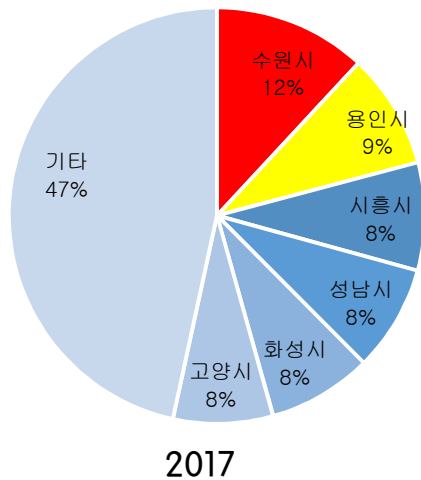
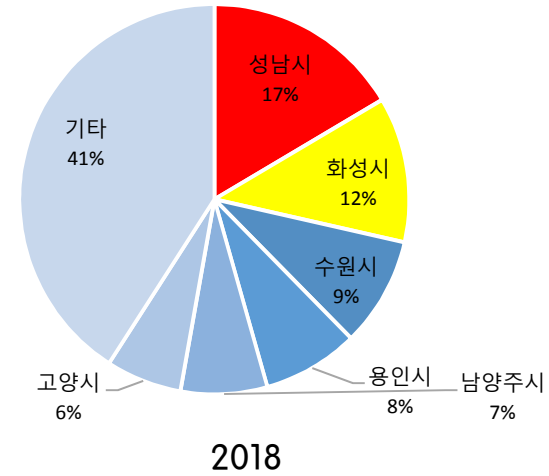
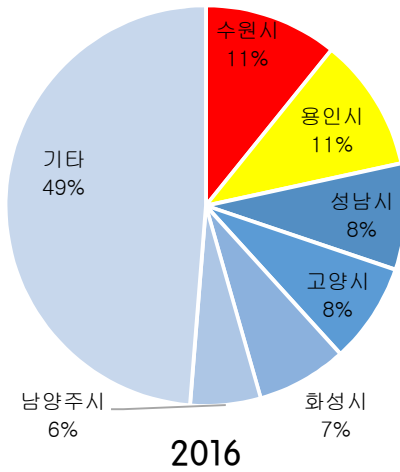


2019

4. 지역별 민원 데이터 현황_경기도

IV. 분석결과_민원현황

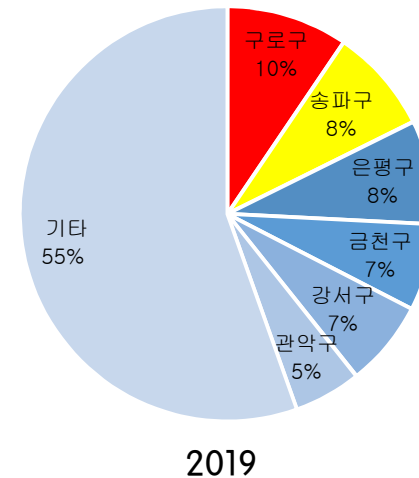
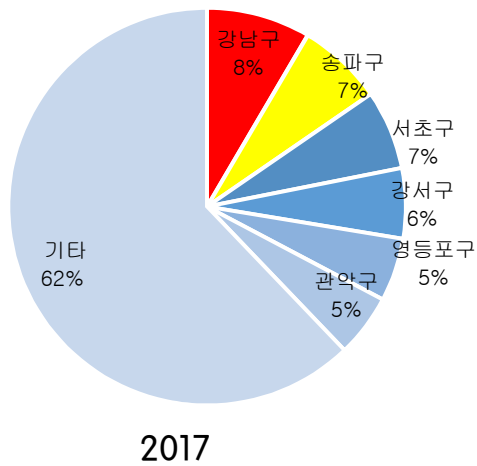
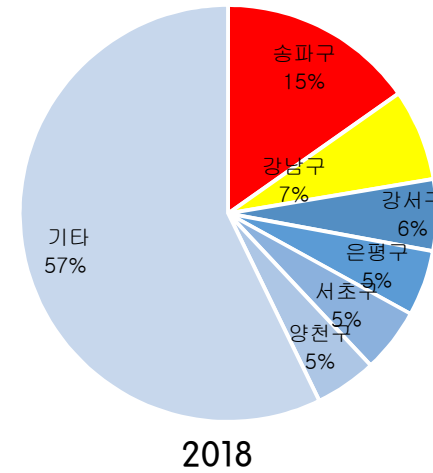
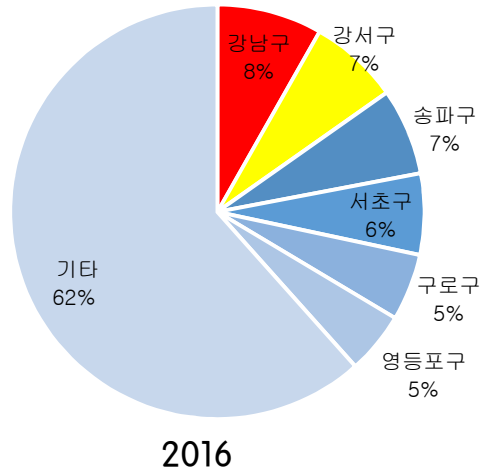
2019년 고양시 및 부천시의 민원이 증가했으며, 상대적으로 수원시와 성남시의 민원은 감소



4. 지역별 민원 데이터 현황_서울시

IV. 분석결과_민원현황

2018년 구로구 민원이 가장 커졌으며, 강남구 민원은 상대적으로 감소



대한민국

대한민국 민원키워드

IV. 분석결과_대한민국

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

대한민국 민원키워드

여성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상

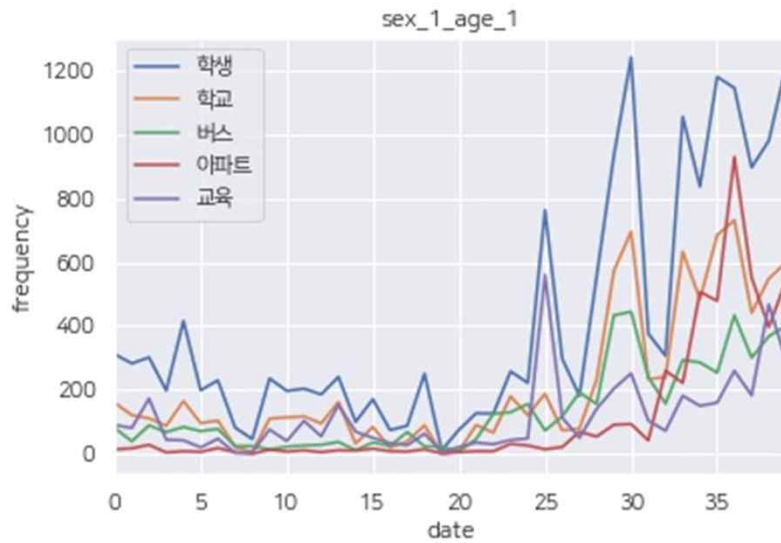


워드 클라우드 기타

대한민국 민원키워드

IV. 분석결과_대한민국

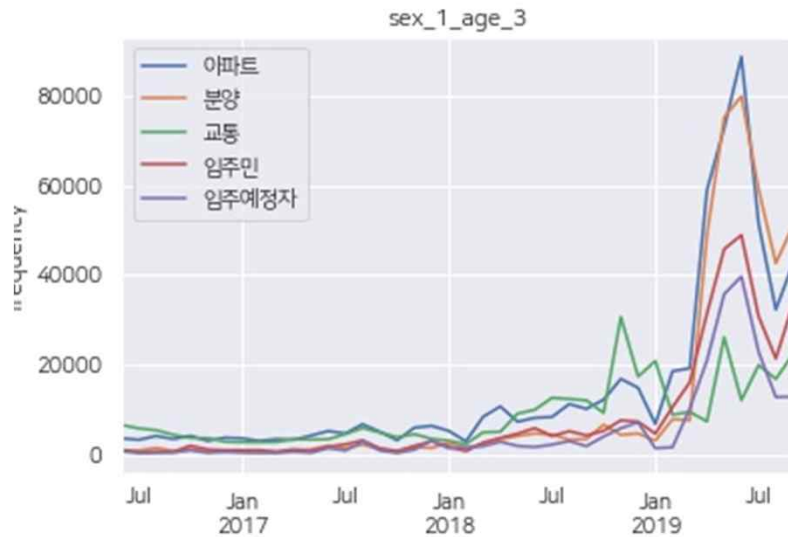
남성



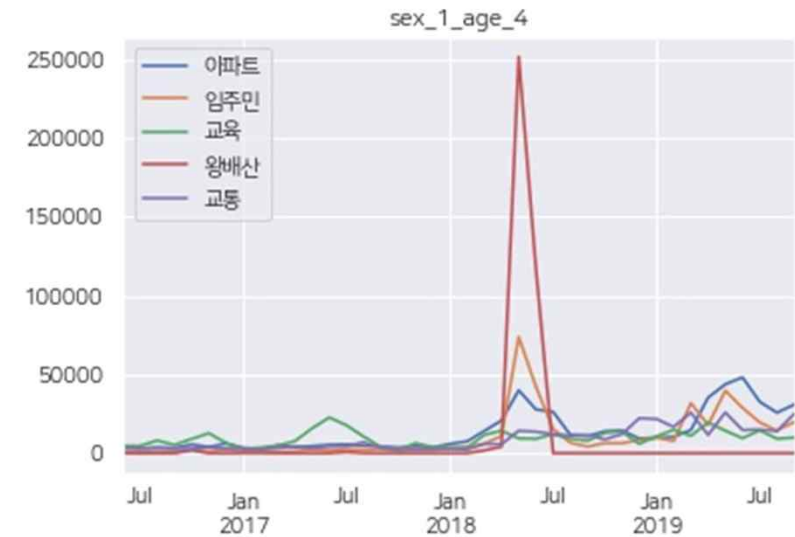
트렌드 10대



트렌드 20대



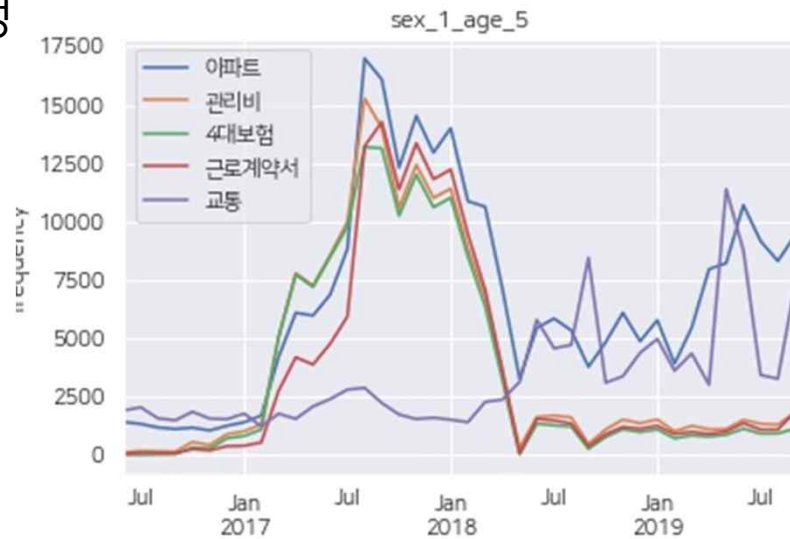
트렌드 30대 이상



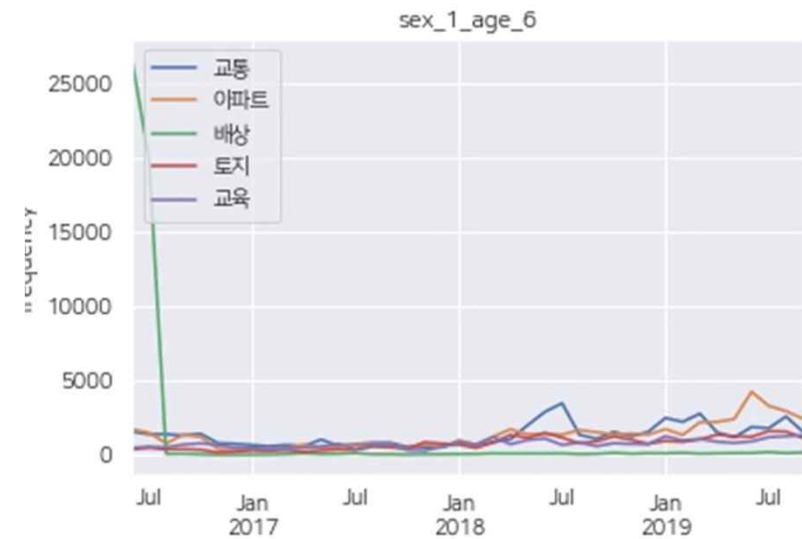
트렌드 40대 DBDISCOVER

대한민국 민원키워드

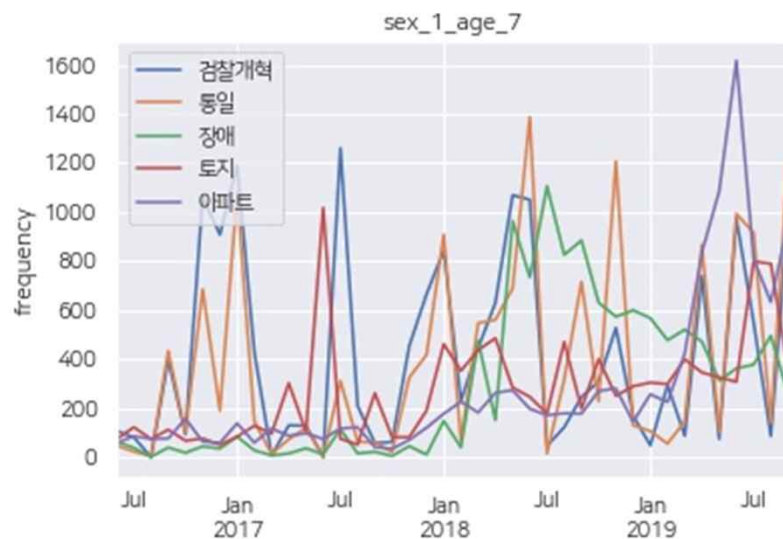
남성



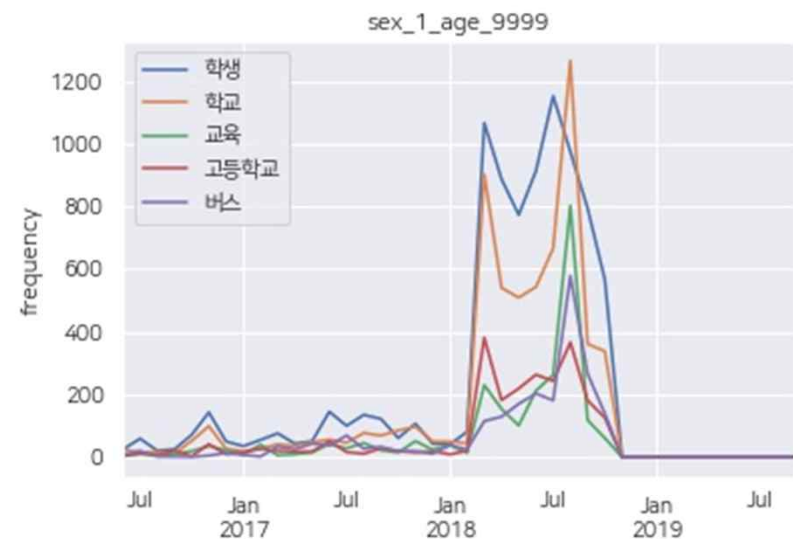
트렌드 50대



트렌드 60대



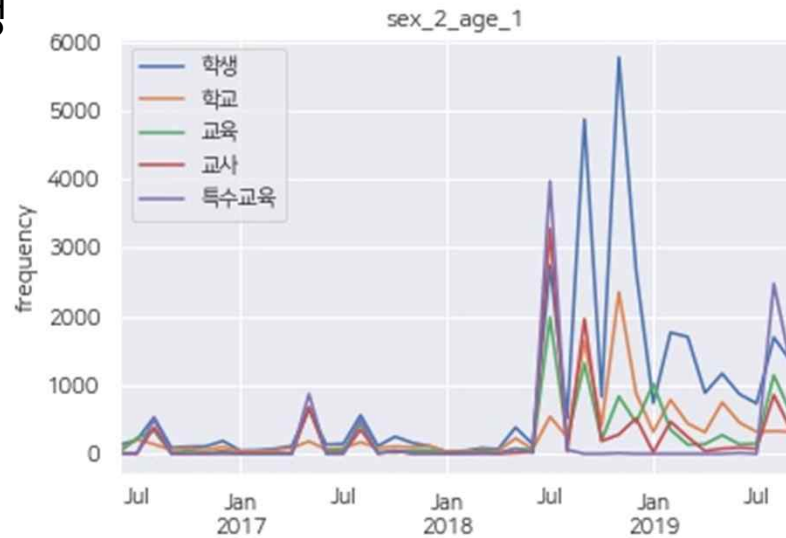
트렌드 70대 이상



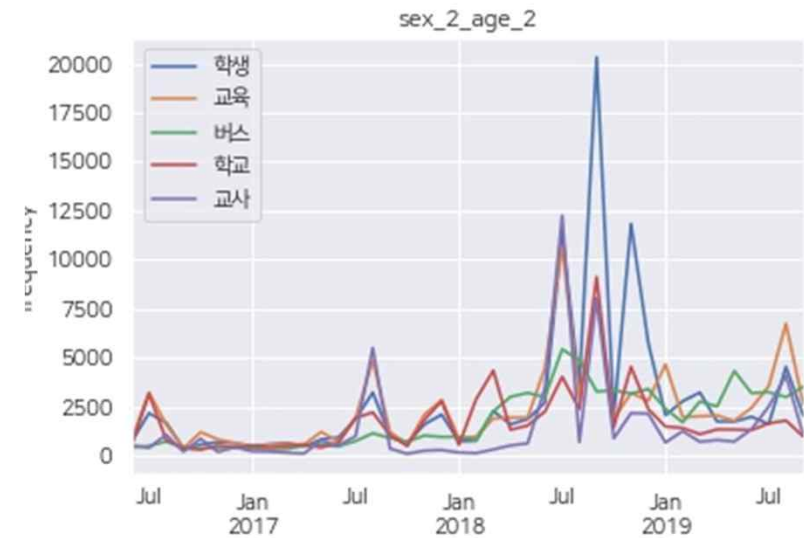
트렌드 기타 **DBDISCOVER**

대한민국 민원키워드

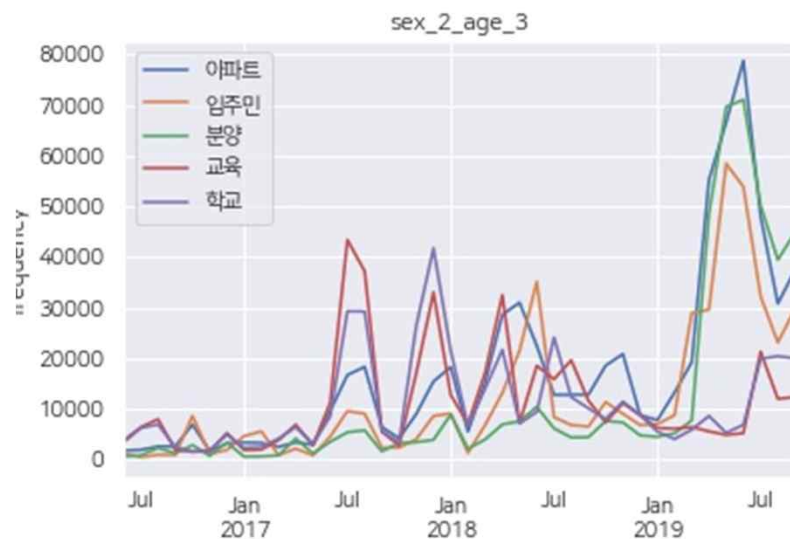
여성



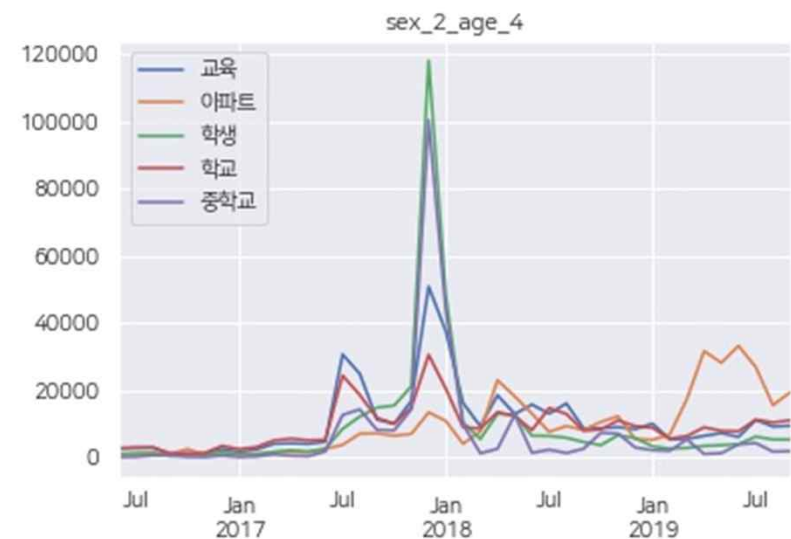
트렌드 10대



트렌드 20대

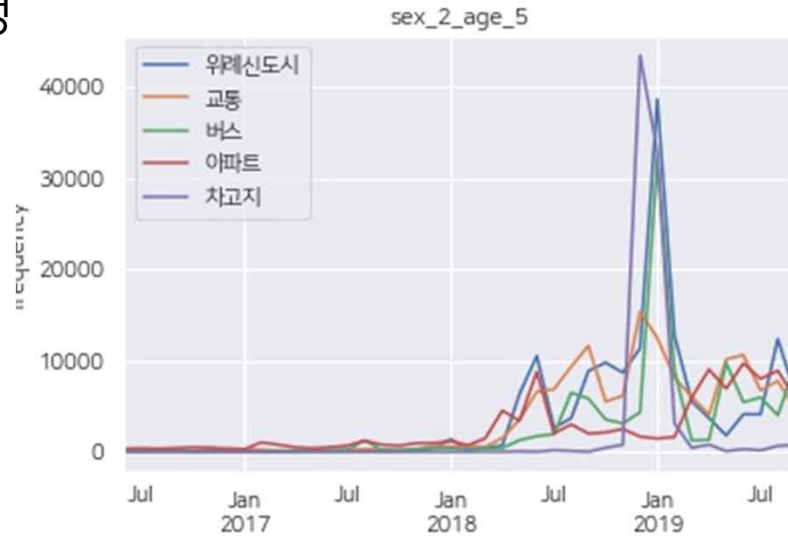


트렌드 30대 이상

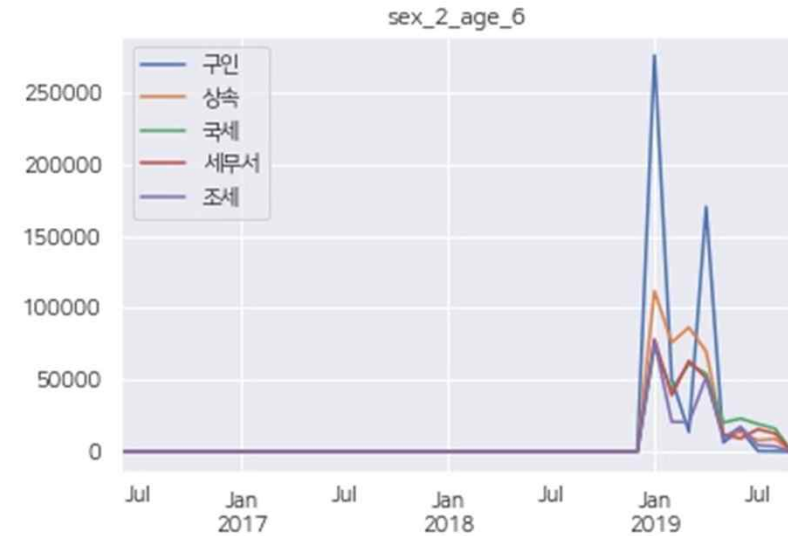


트렌드 40대 **DBDISCOVER**

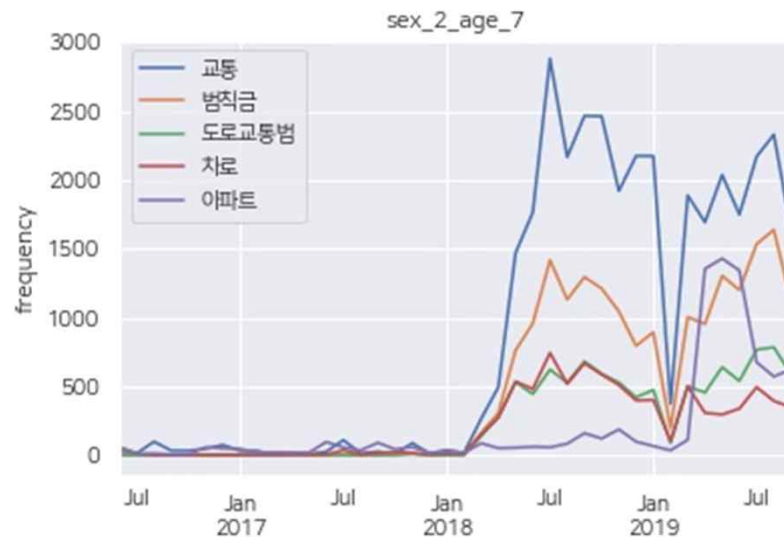
여성



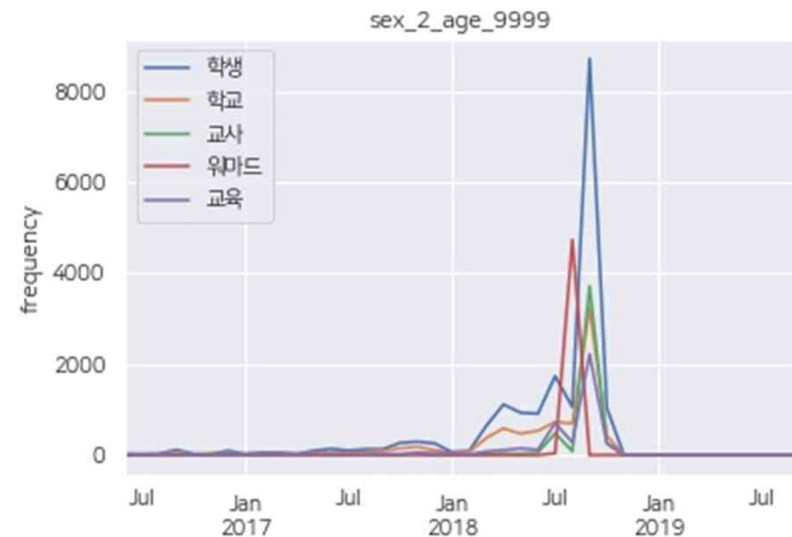
트렌드 50대



트렌드 60대



트렌드 70대 이상

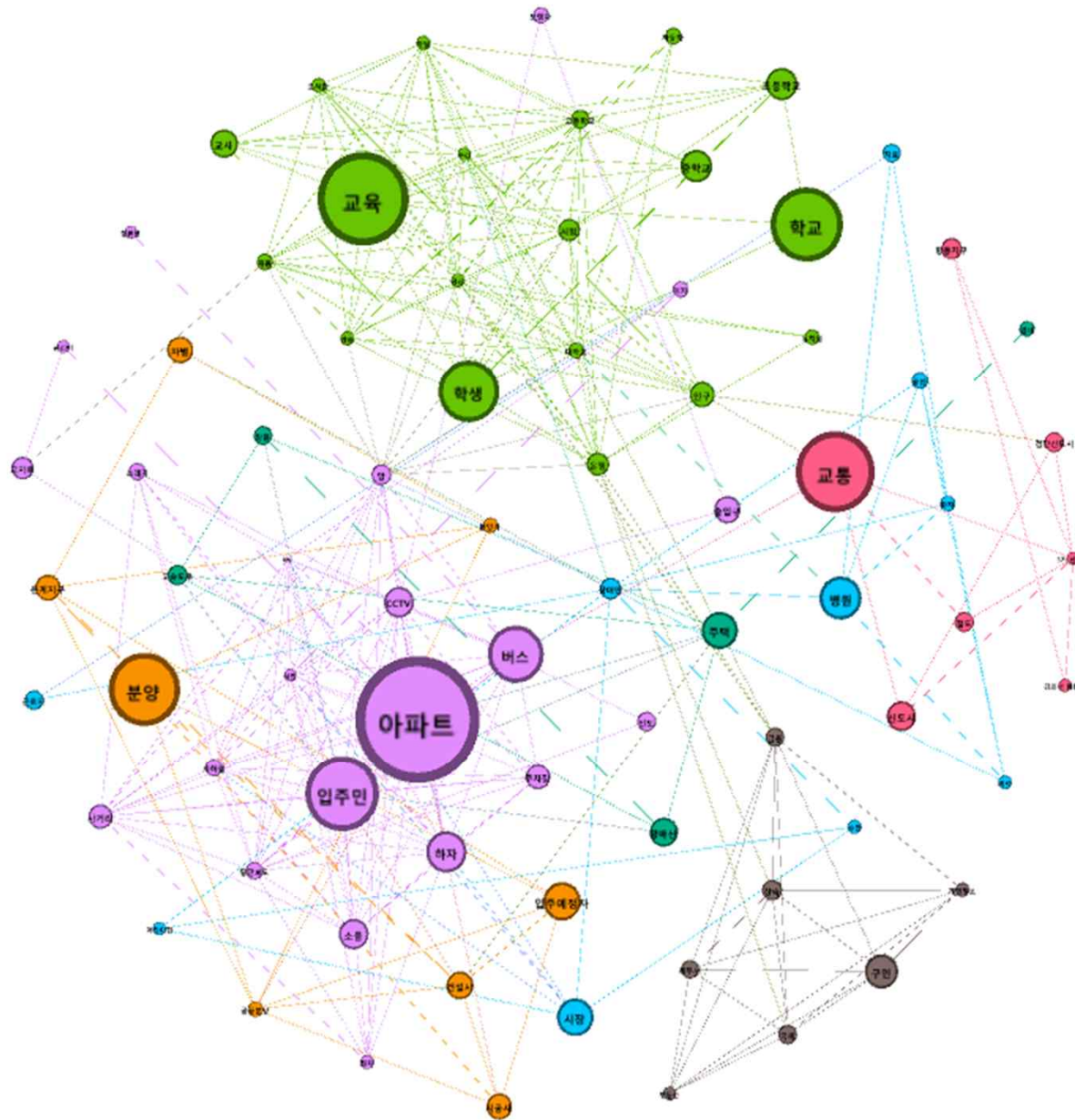


트렌드 기타 **DBDISCOVER**

대한민국 민원키워드

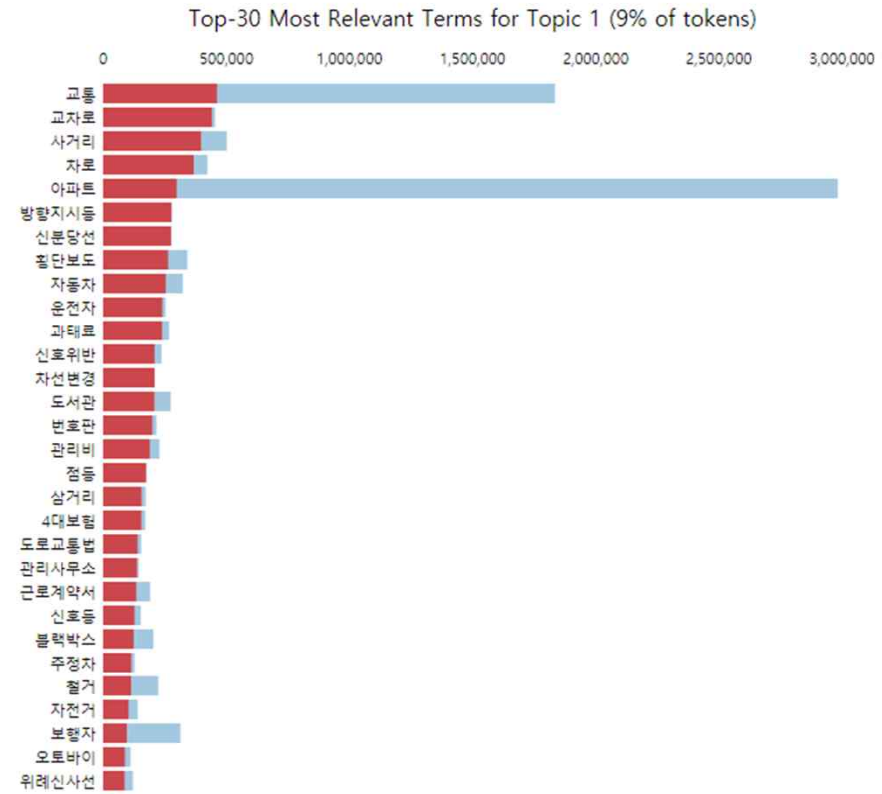
IV. 분석결과_대한민국

네트워크



대한민국 민원키워드

토픽모델링 - 1번 토픽 구성 키워드는 교통, 교차로, 방향지시등, 삼거리 등 도로 및 교통과 관련된 키워드 중심



Overall term frequency
 Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
 2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

서울시

서울특별시 민원키워드

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

서울특별시 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대



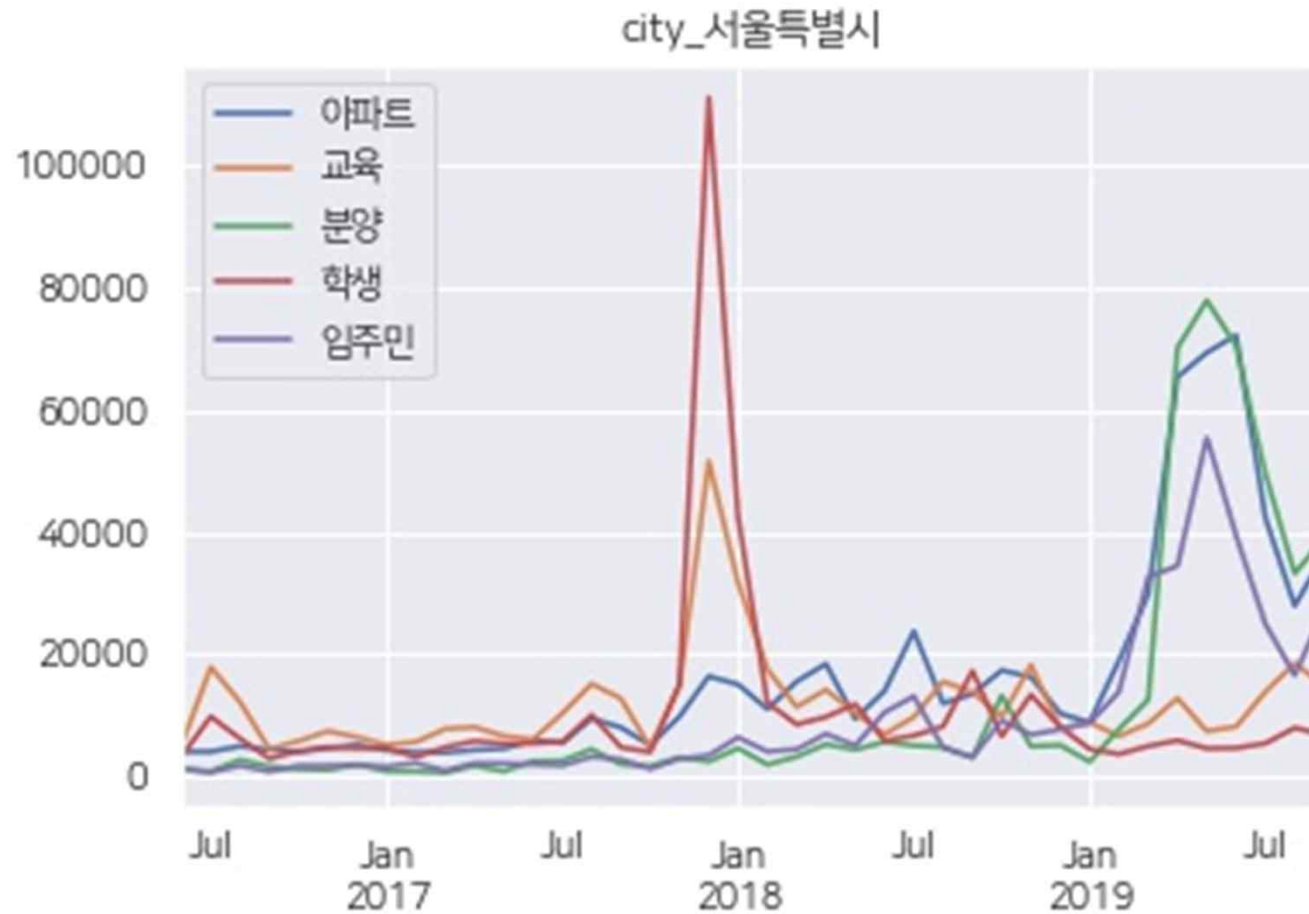
워드 클라우드 70대 이상



워드 클라우드 기타

서울특별시 민원키워드

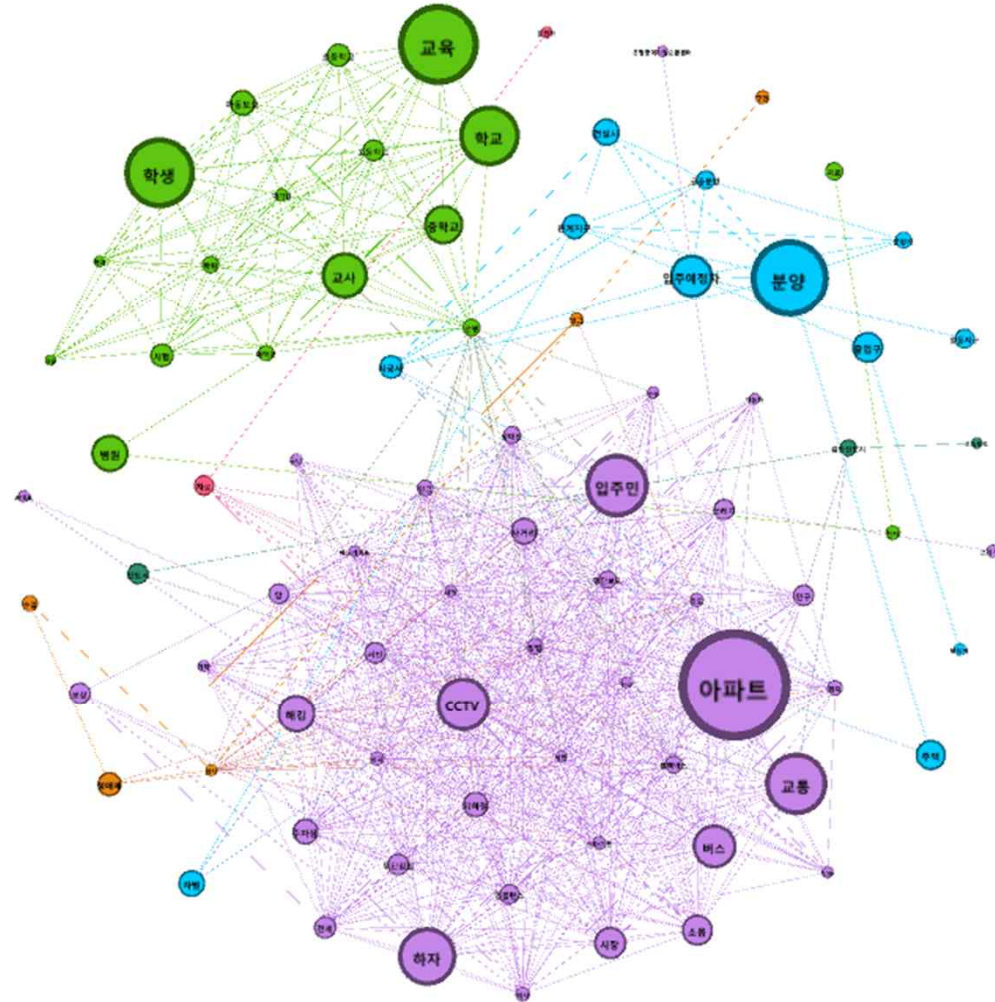
트렌드



서울특별시 민원키워드

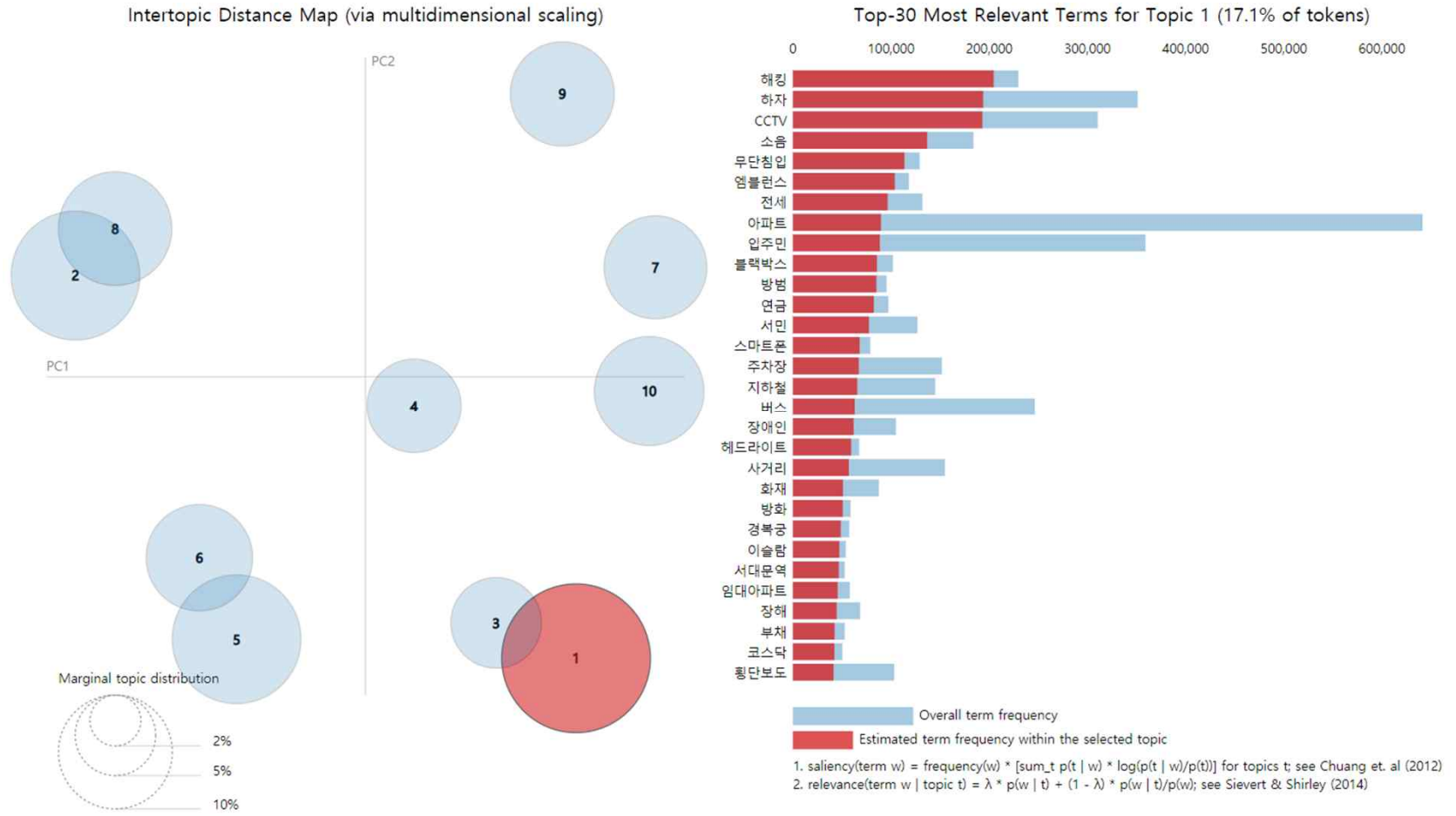
IV. 분석결과_지역별

네트워크



서울특별시 민원키워드

토픽모델링 1번 토픽 : cctv, 무단침입, 소음 등의 키워드들이 1번 토픽을 구성하고 있음



부산광역시

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대

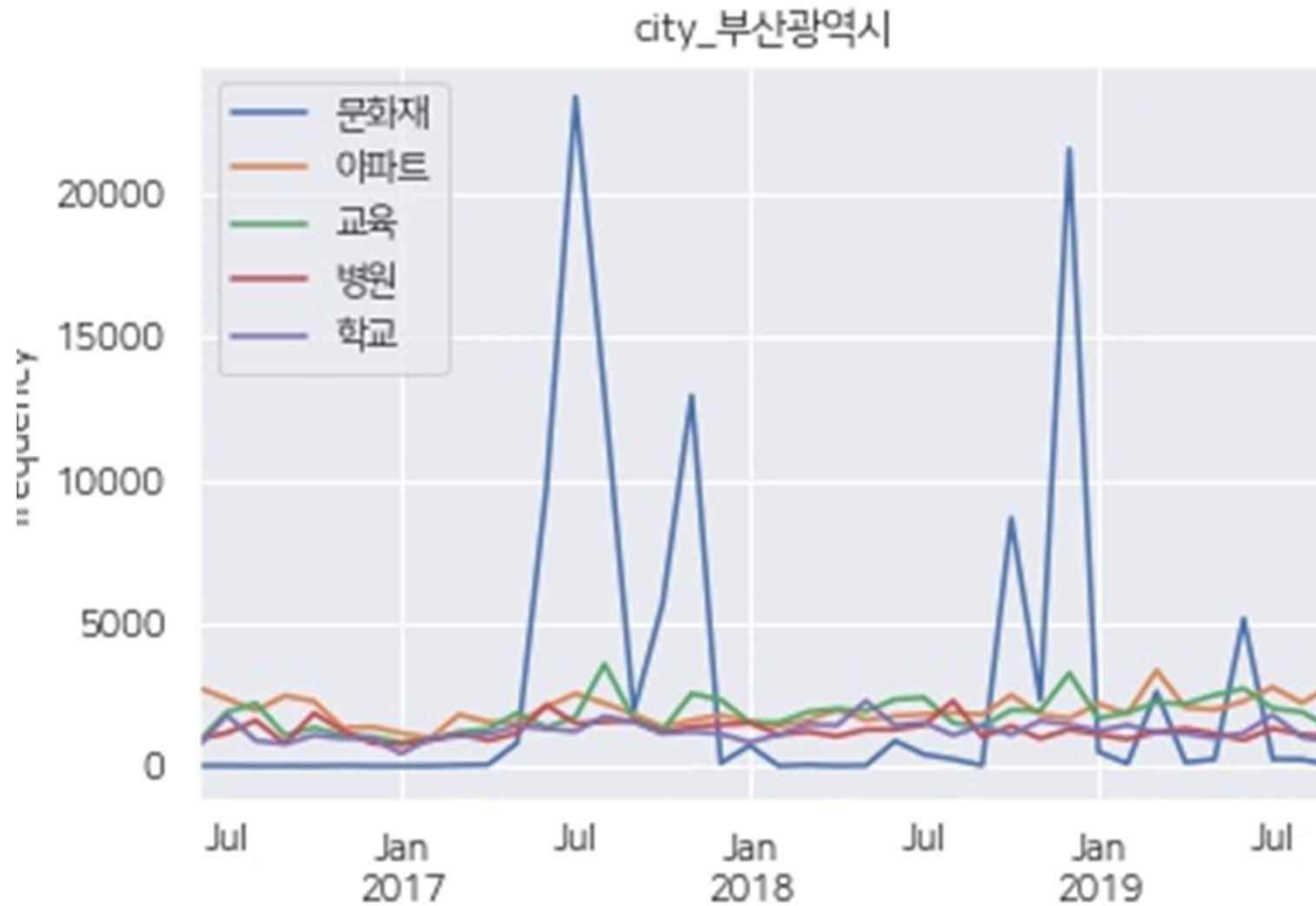


워드 클라우드 30대

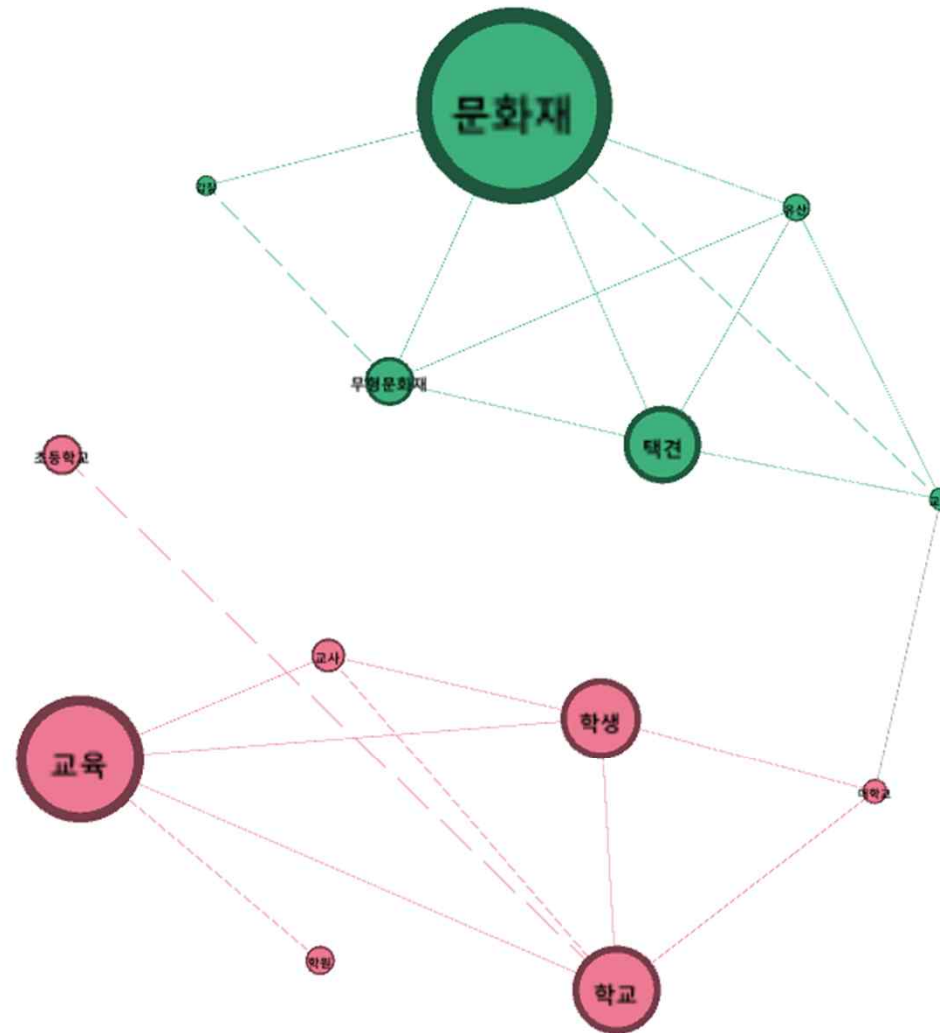


워드 클라우드 40대

트렌드



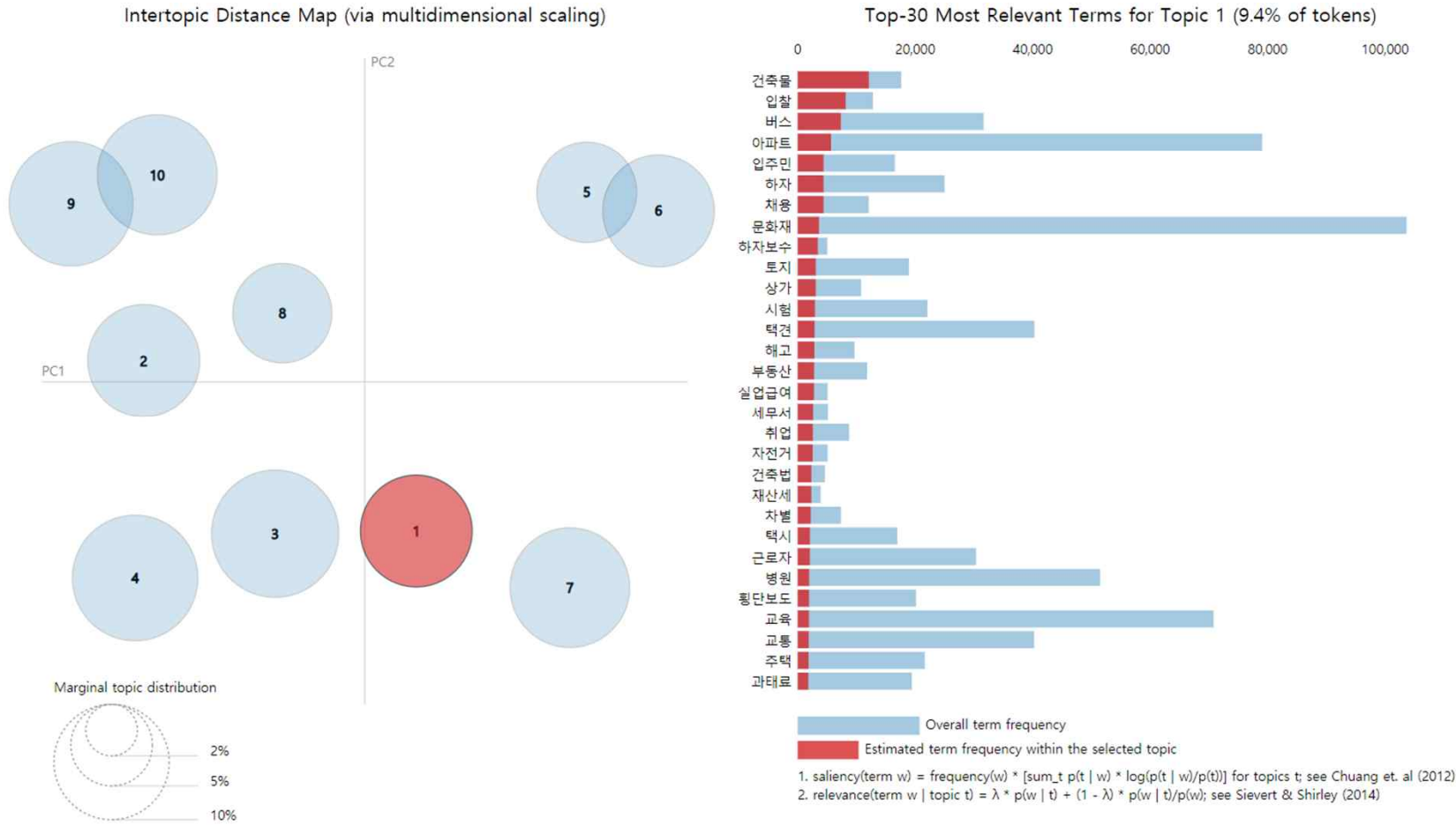
네트워크



부산광역시 민원키워드

IV. 분석결과_지역별

토픽모델링 : 1번 토픽은 건축물, 문화재, 하자보수 등의 토픽으로 구성되어 있음.



대구광역시

대구광역시 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



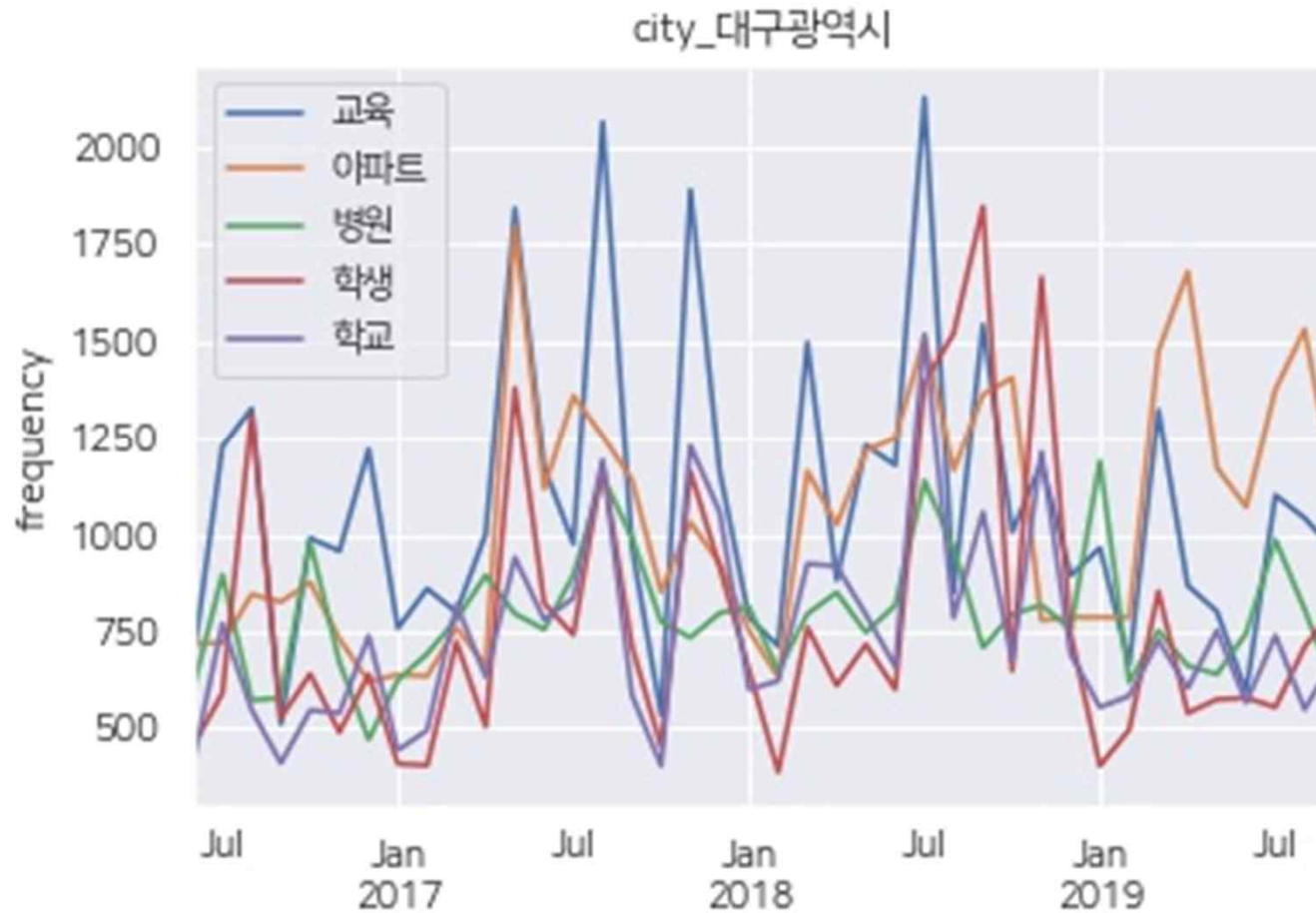
워드 클라우드 30대



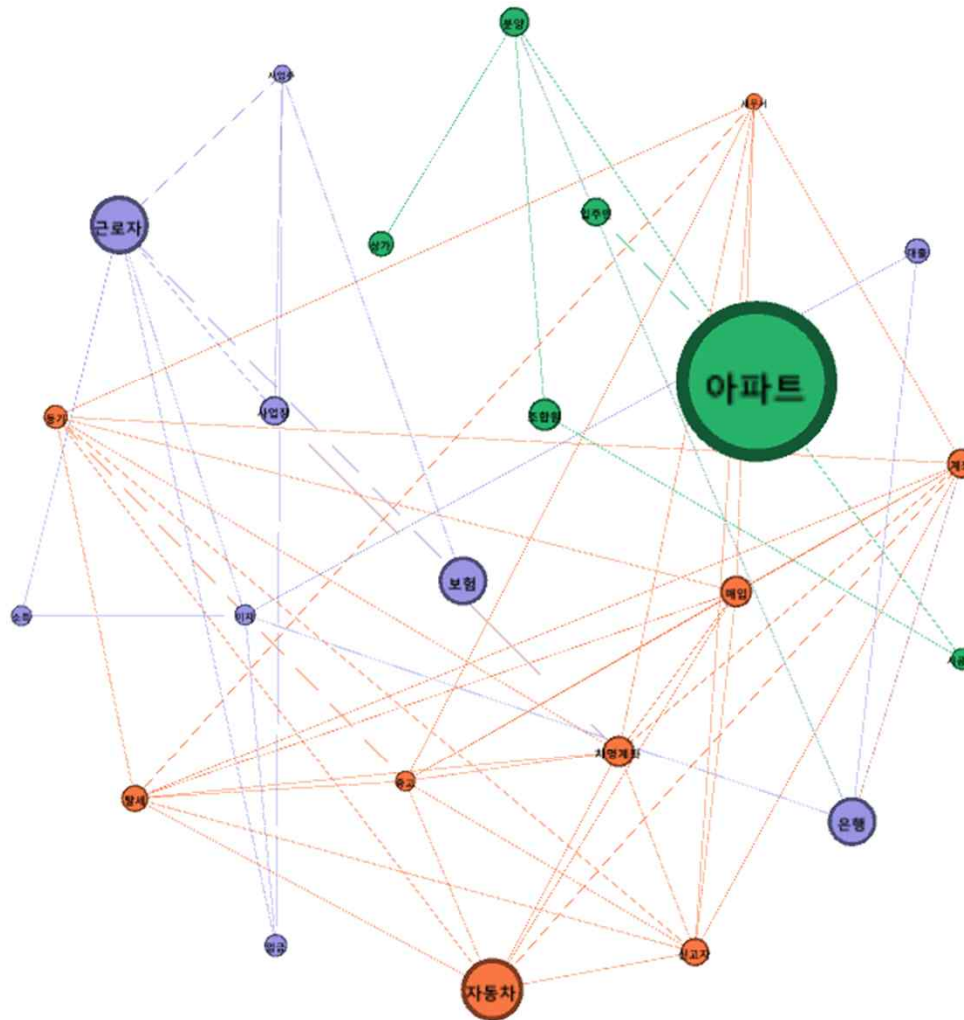
워드 클라우드 40대

대구광역시 민원키워드

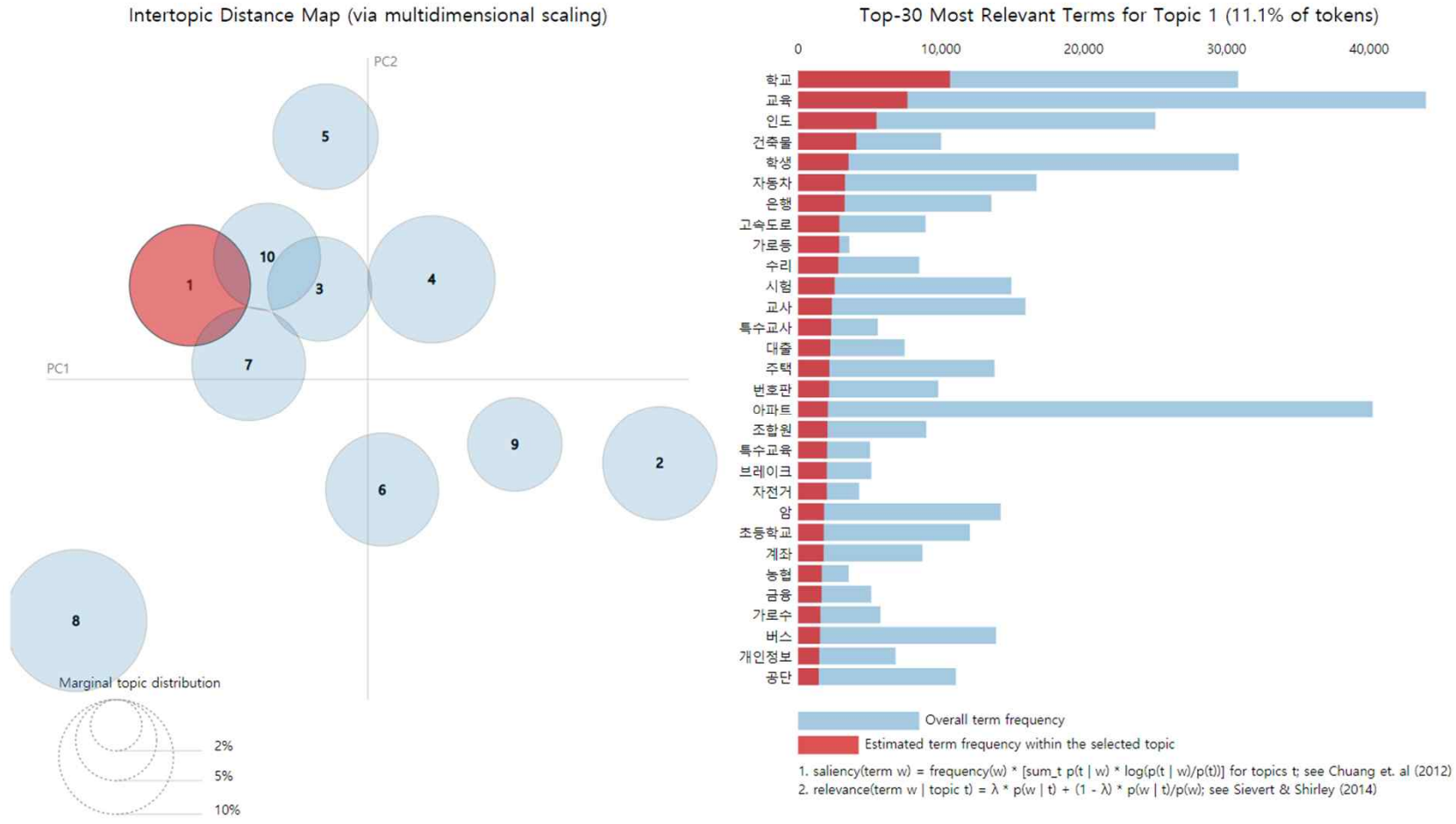
트렌드



네트워크



토픽모델링 : 1번 토픽은 학교, 교육, 학생, 교사 등의 교육 관련 연관어들로 구성되어 있음



인천광역시

인천광역시 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

인천광역시 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

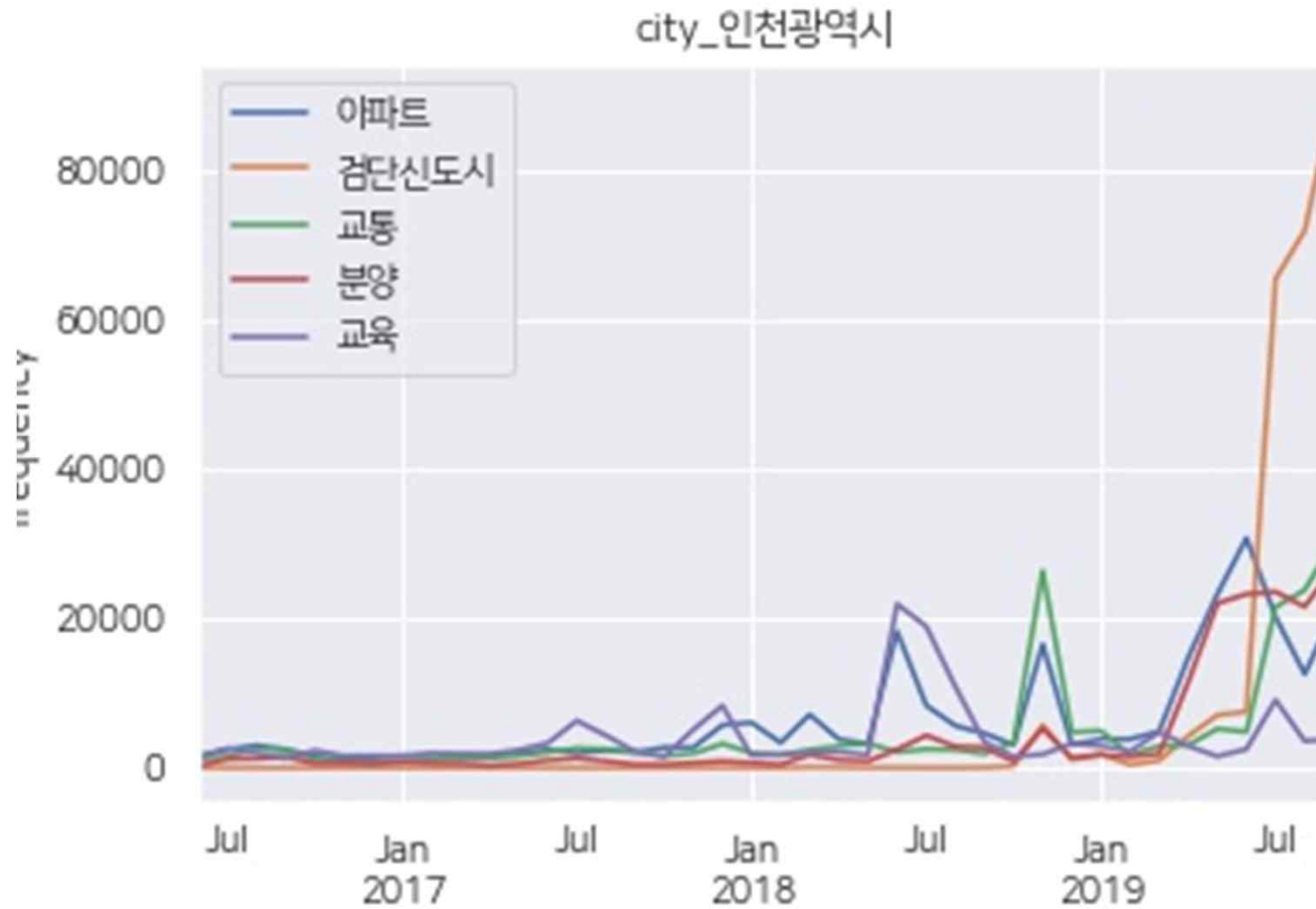


워드 클라우드 70대 이상

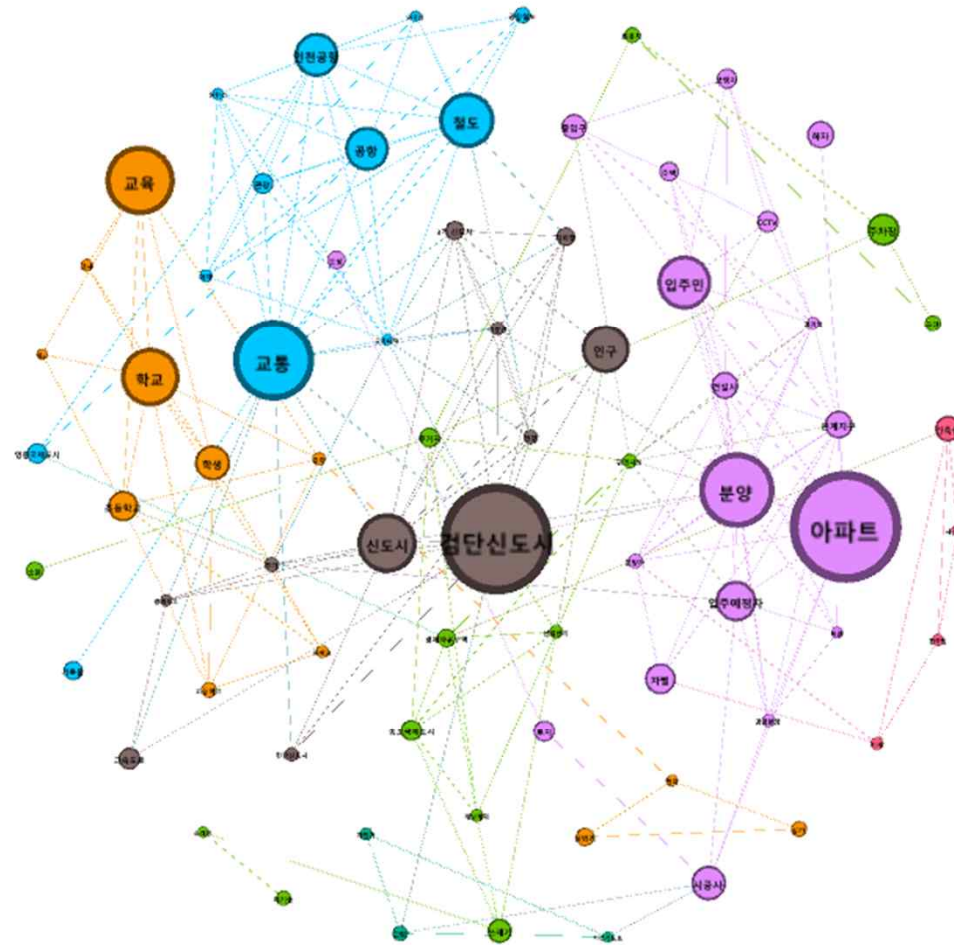


워드 클라우드 기타

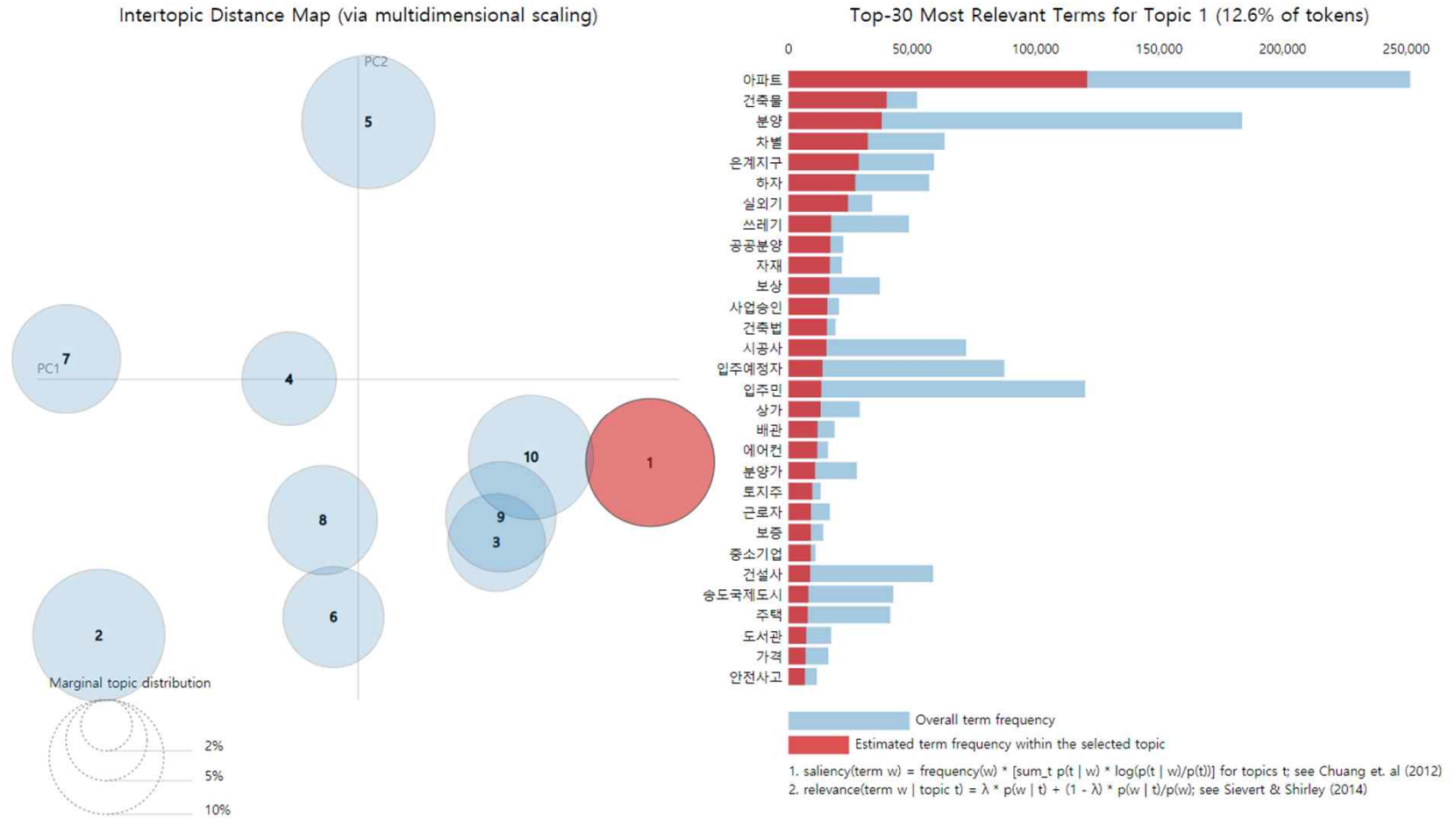
트렌드



네트워크



토픽모델링 : 1번토픽은 아파트, 건축물, 분양, 시공사 등 아파트 분양 및 입주와 관련된 키워드 들로 구성됨



광주광역시

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

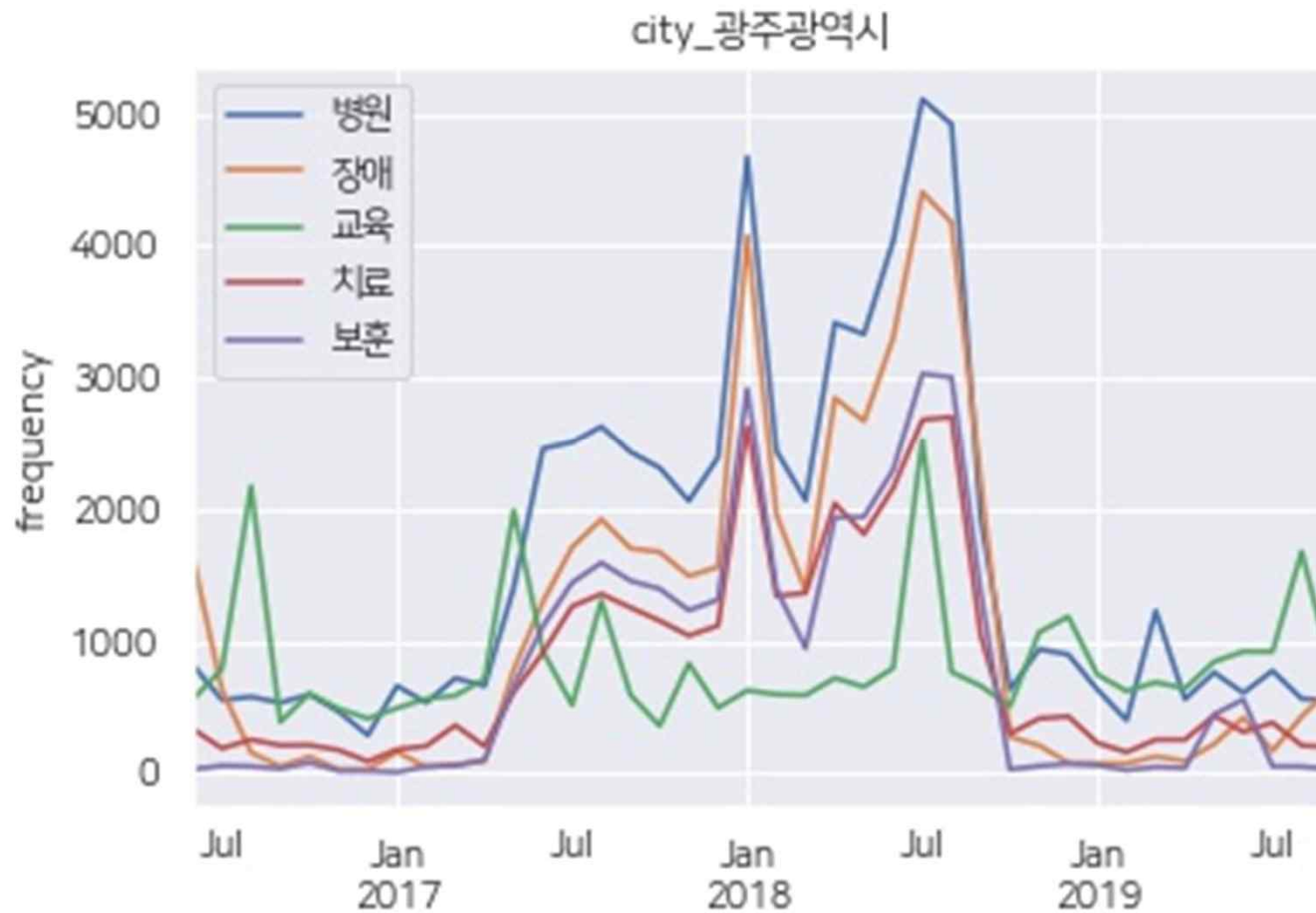


워드 클라우드 70대 이상

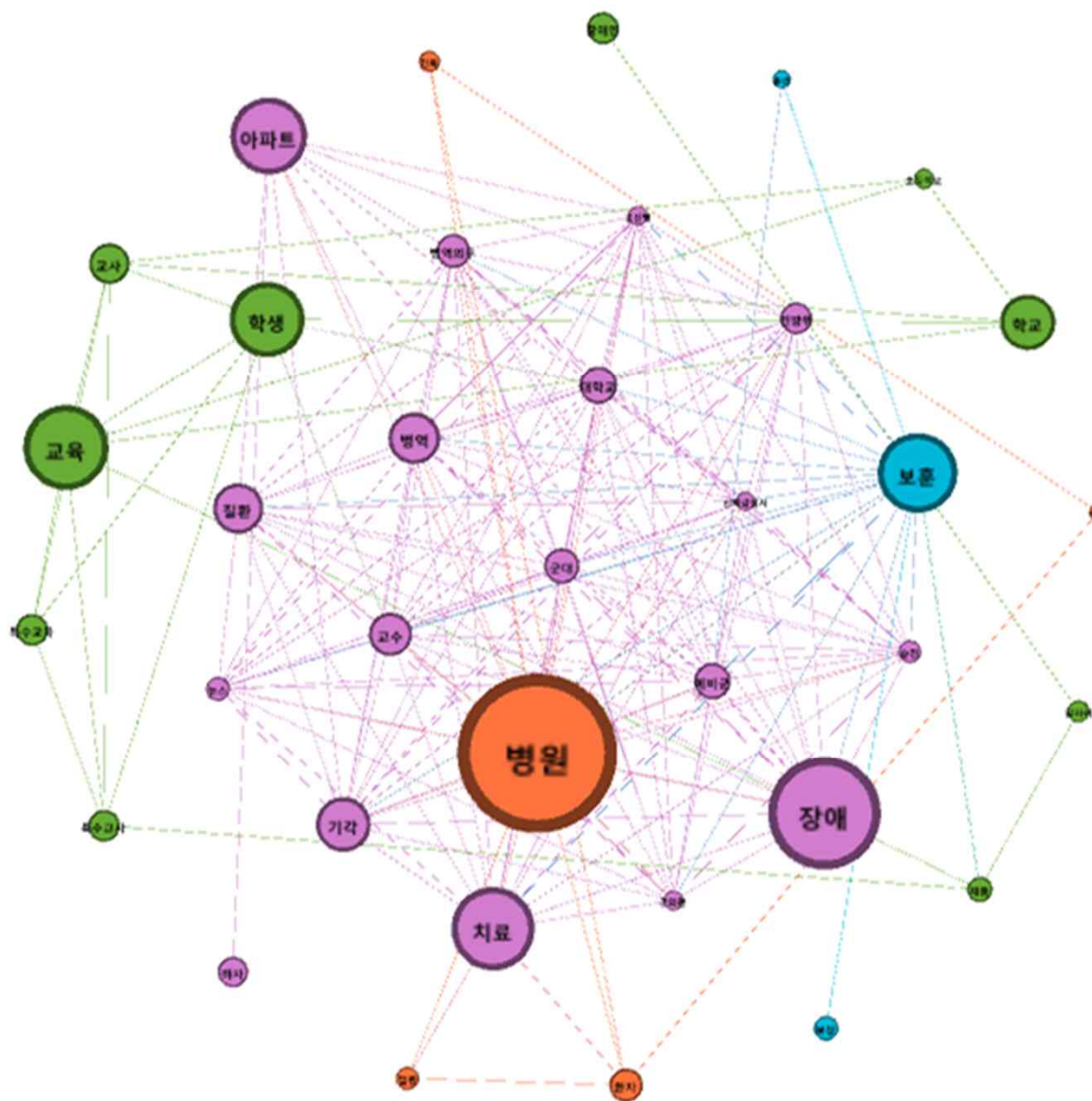


워드 클라우드 기타

트렌드

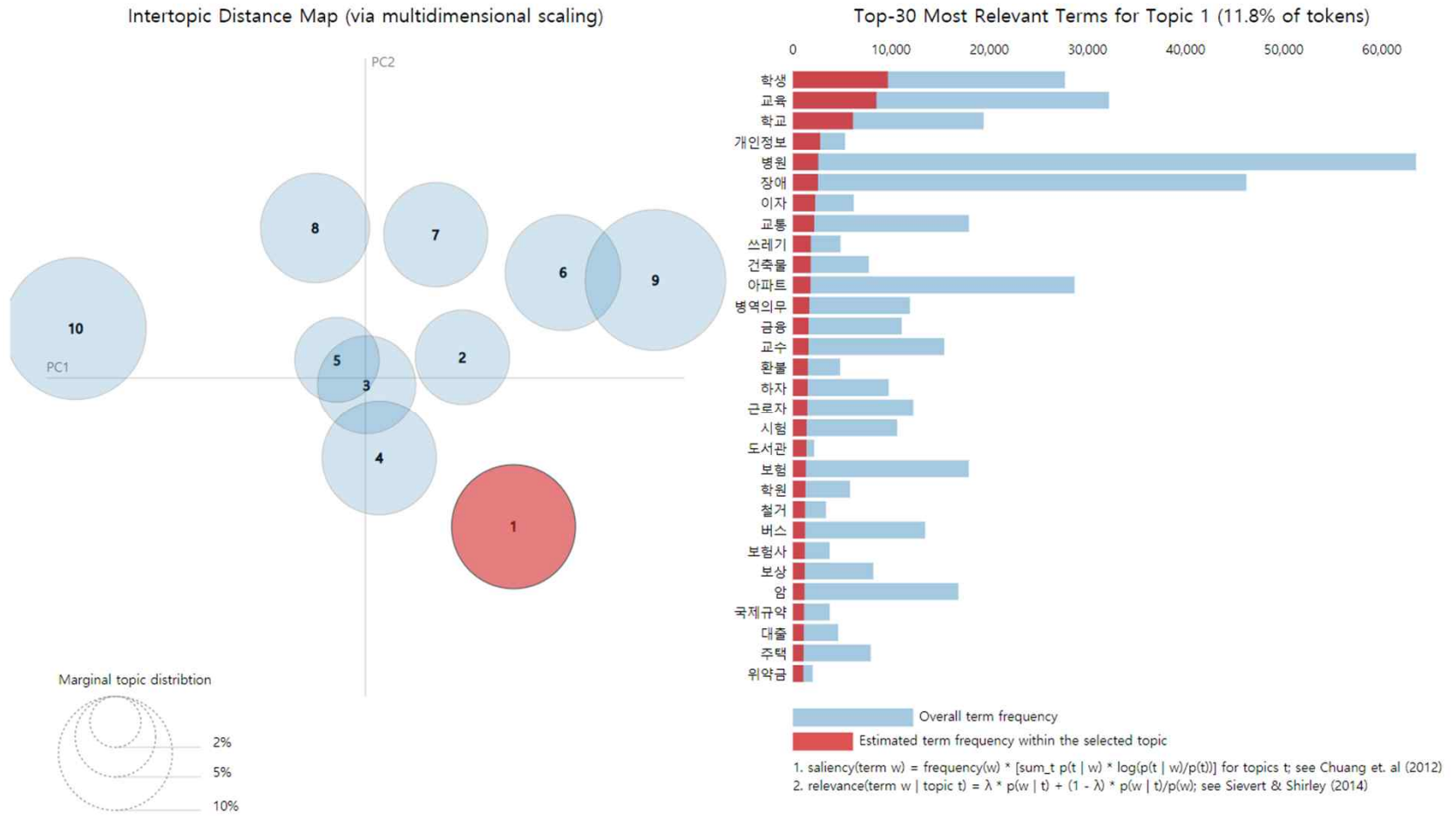


네트워크



광주광역시 민원키워드

토픽모델링 : 광주의 1번 토픽은 학생, 교육, 학교, 시험, 도서관 등 교육 관련 키워드가 많이 포함됨.



대전광역시

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

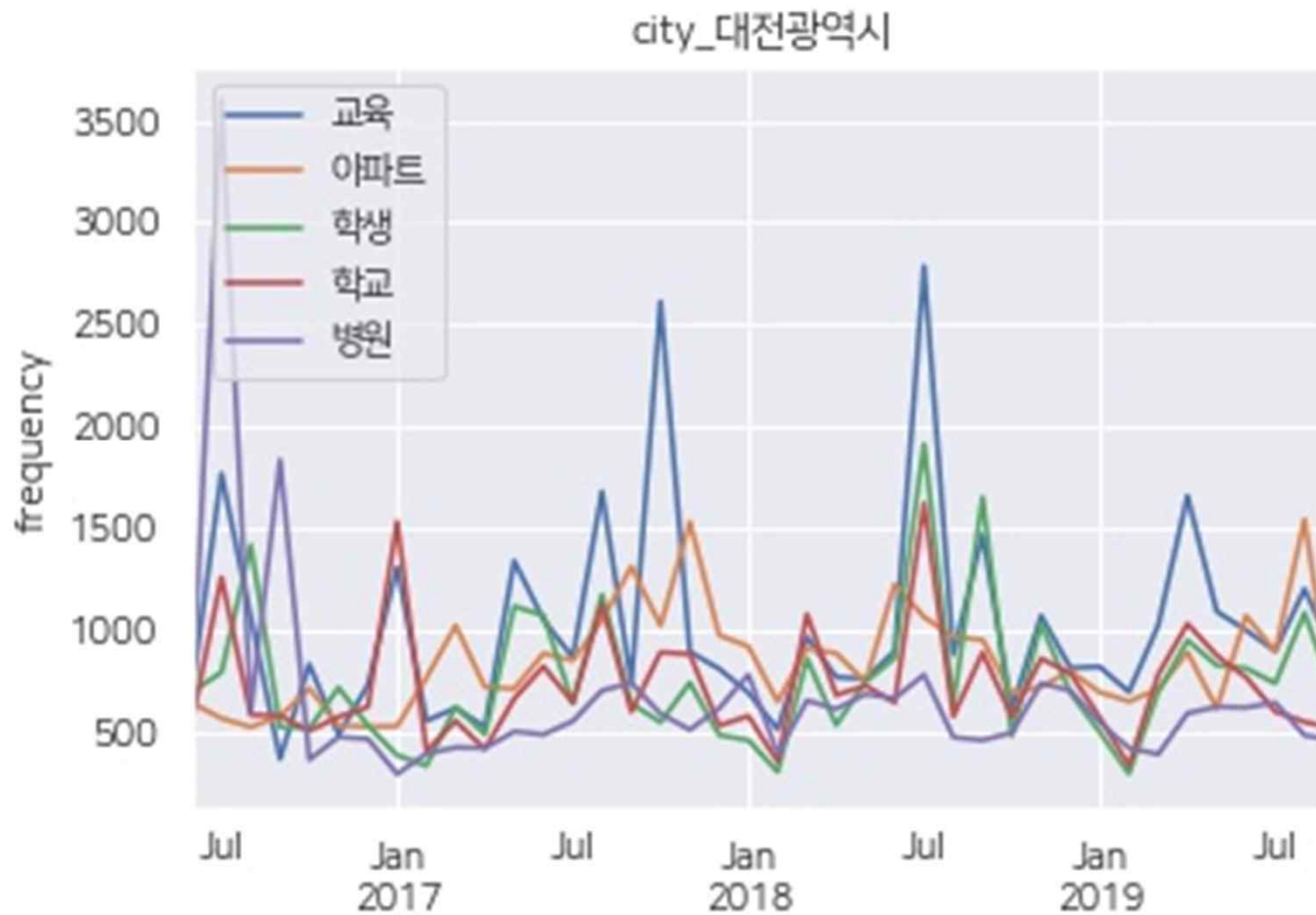


워드 클라우드 70대 이상

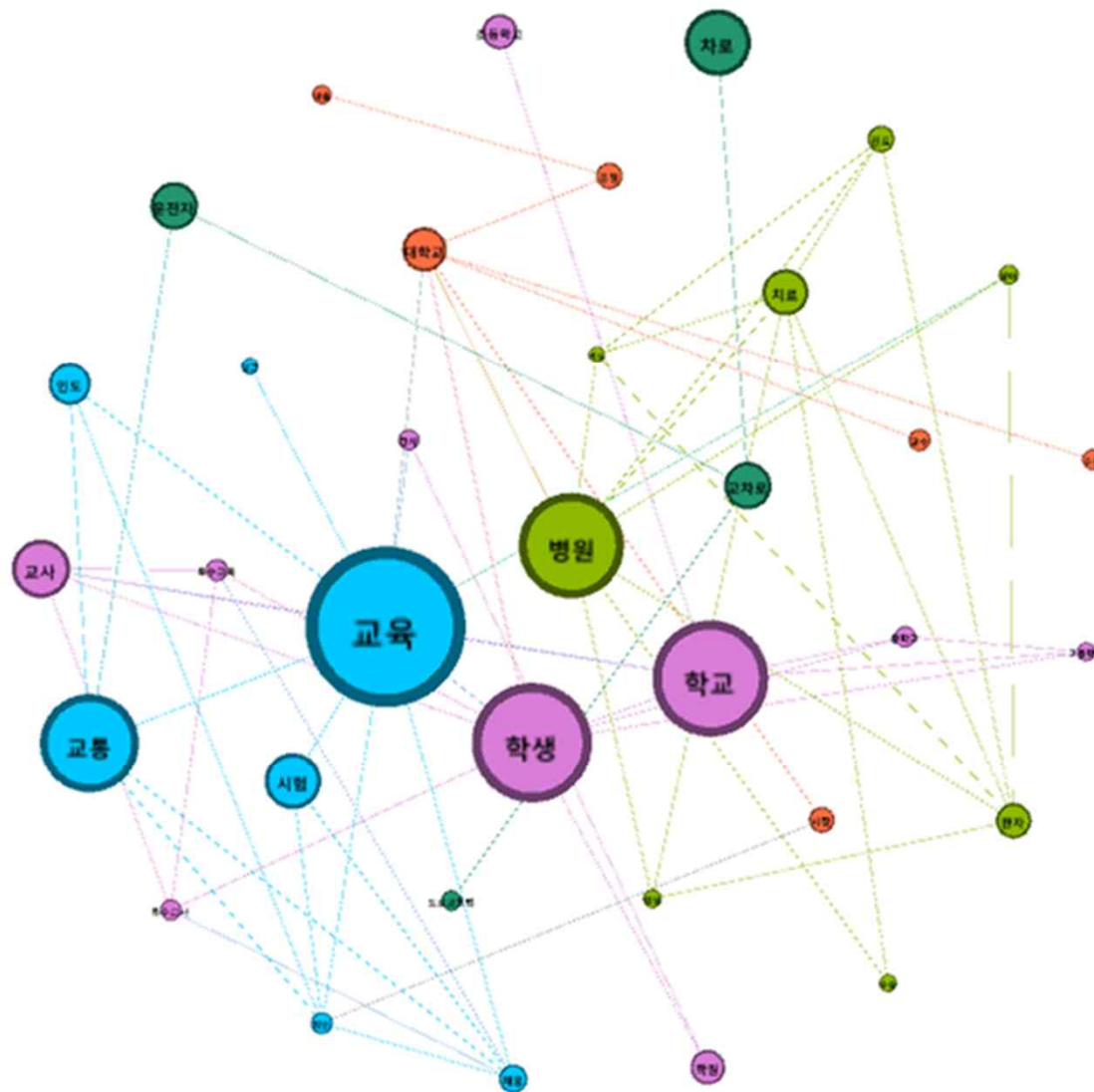


워드 클라우드 기타

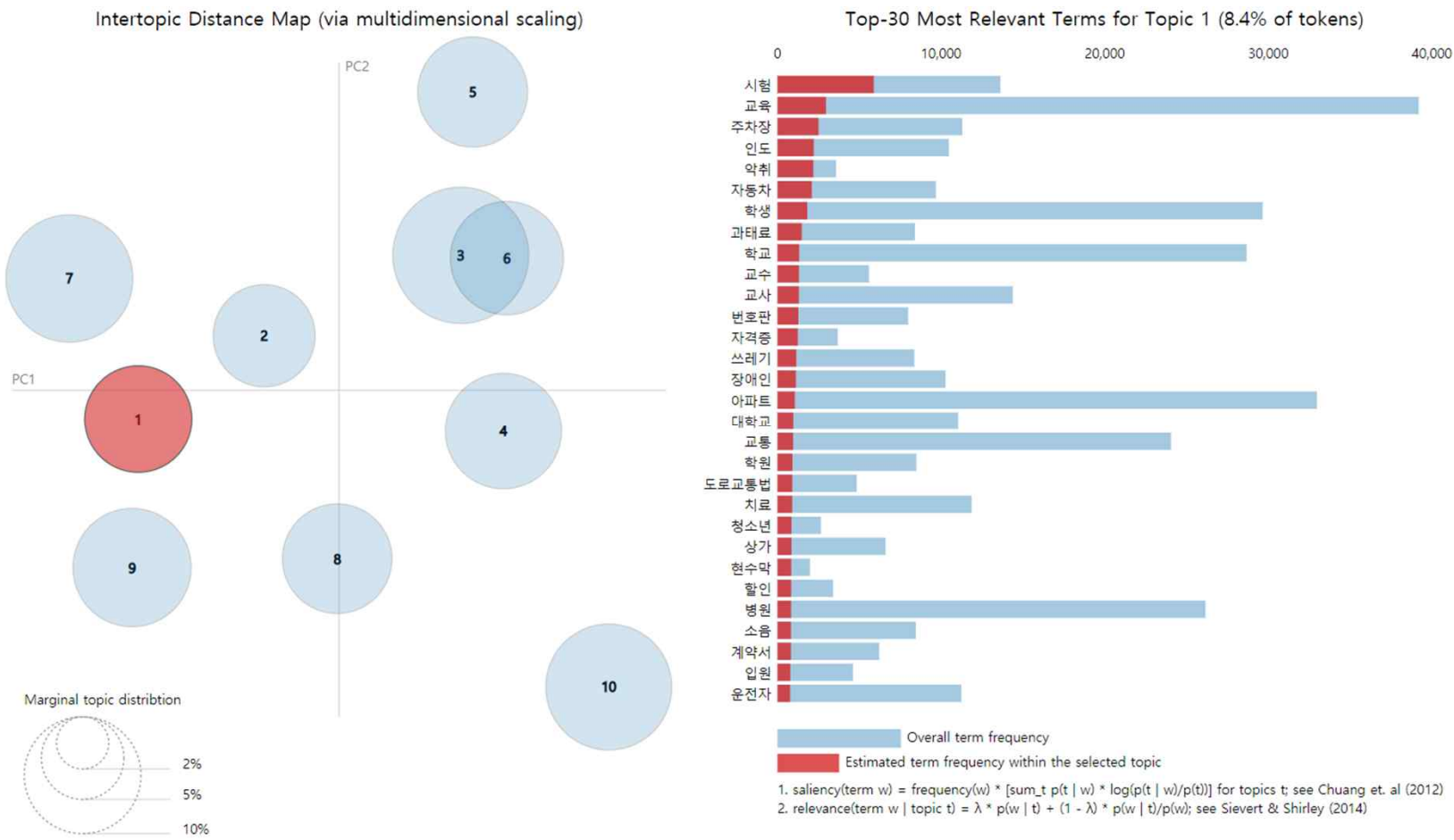
트렌드



네트워크



토픽모델링: 1번 토픽은 시험, 교육, 학교, 교사 등 교육에 관한 키워드가 비교적 많이 포함되어 있는 토픽임.



울산광역시

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

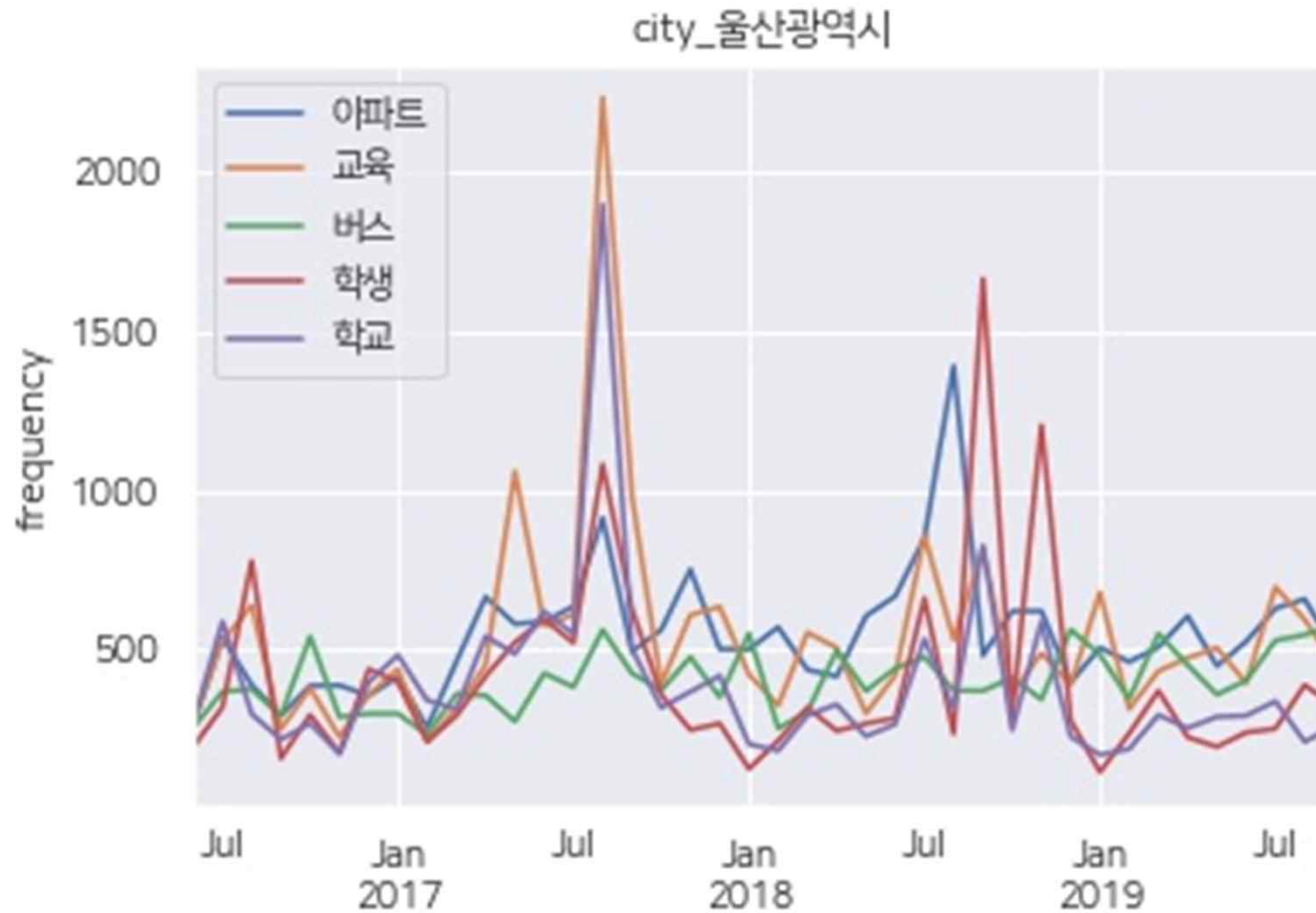


워드 클라우드 70대 이상

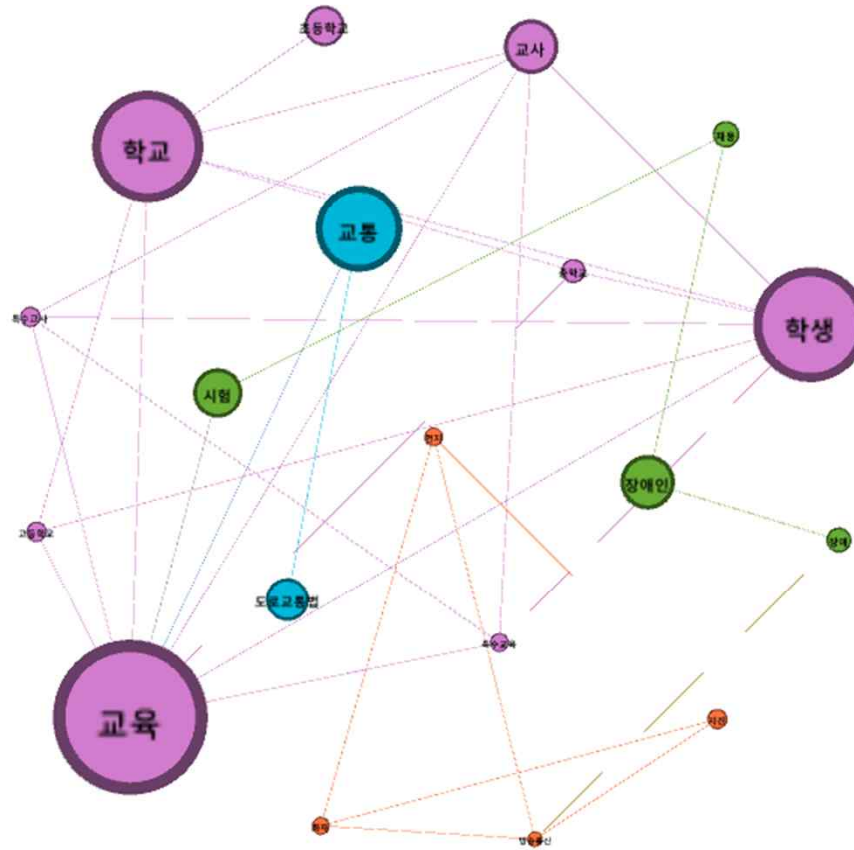


워드 클라우드 기타

트렌드



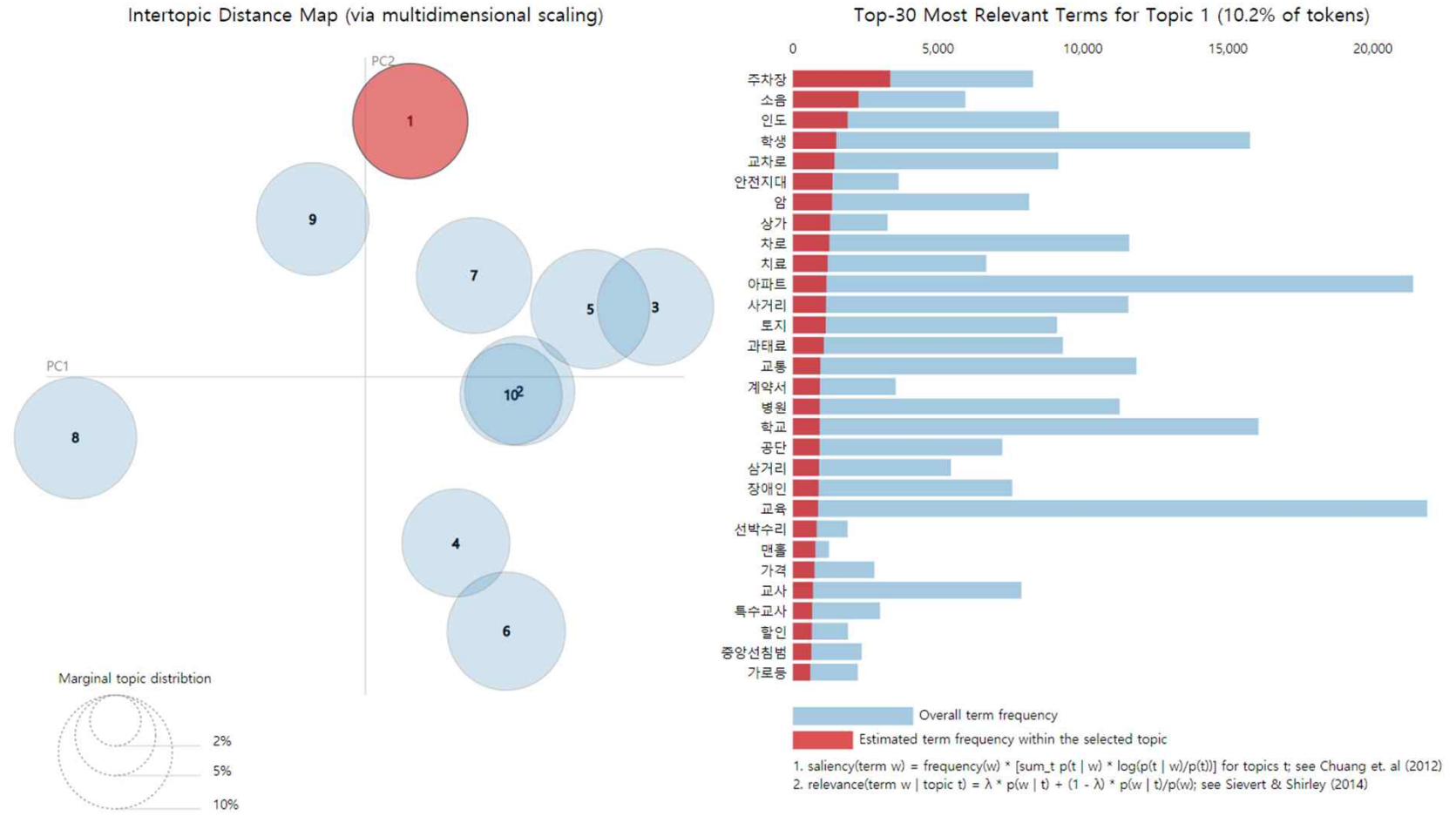
네트워크



울산광역시 민원키워드

IV. 분석결과_지역별

토픽모델링 : 울산의 1번 토픽은 주차장, 교차로, 안전지대, 차로, 사거리 등 교통 관련 키워드가 주로 포함됨.



세종특별자치시

세종특별자치시 민원키워드

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

세종특별자치시 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

세종특별자치시 민원키워드

IV. 분석결과_지역별

여성



워드 클라우드 50대



워드 클라우드 60대

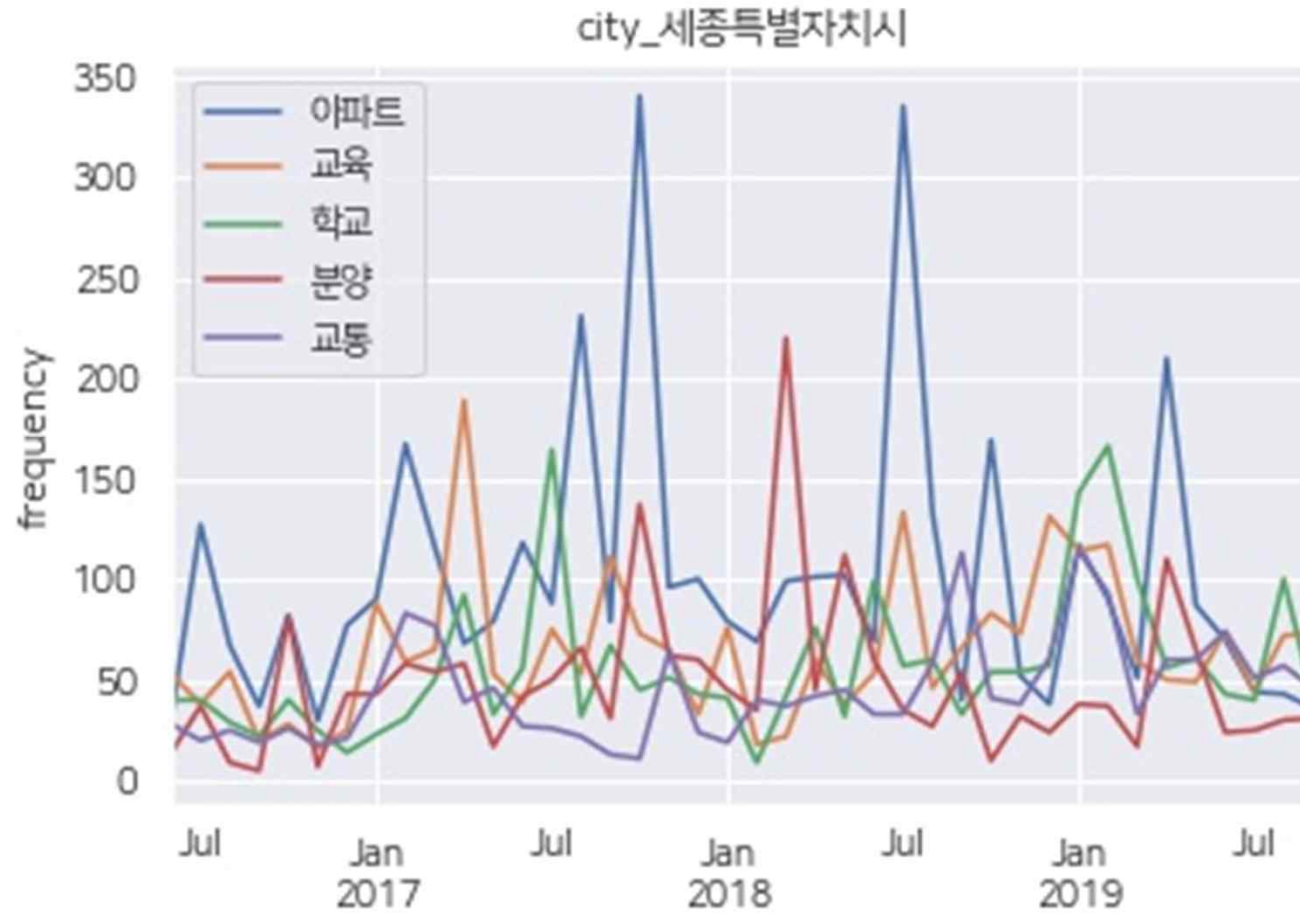


워드 클라우드 70대 이상



워드 클라우드 기타

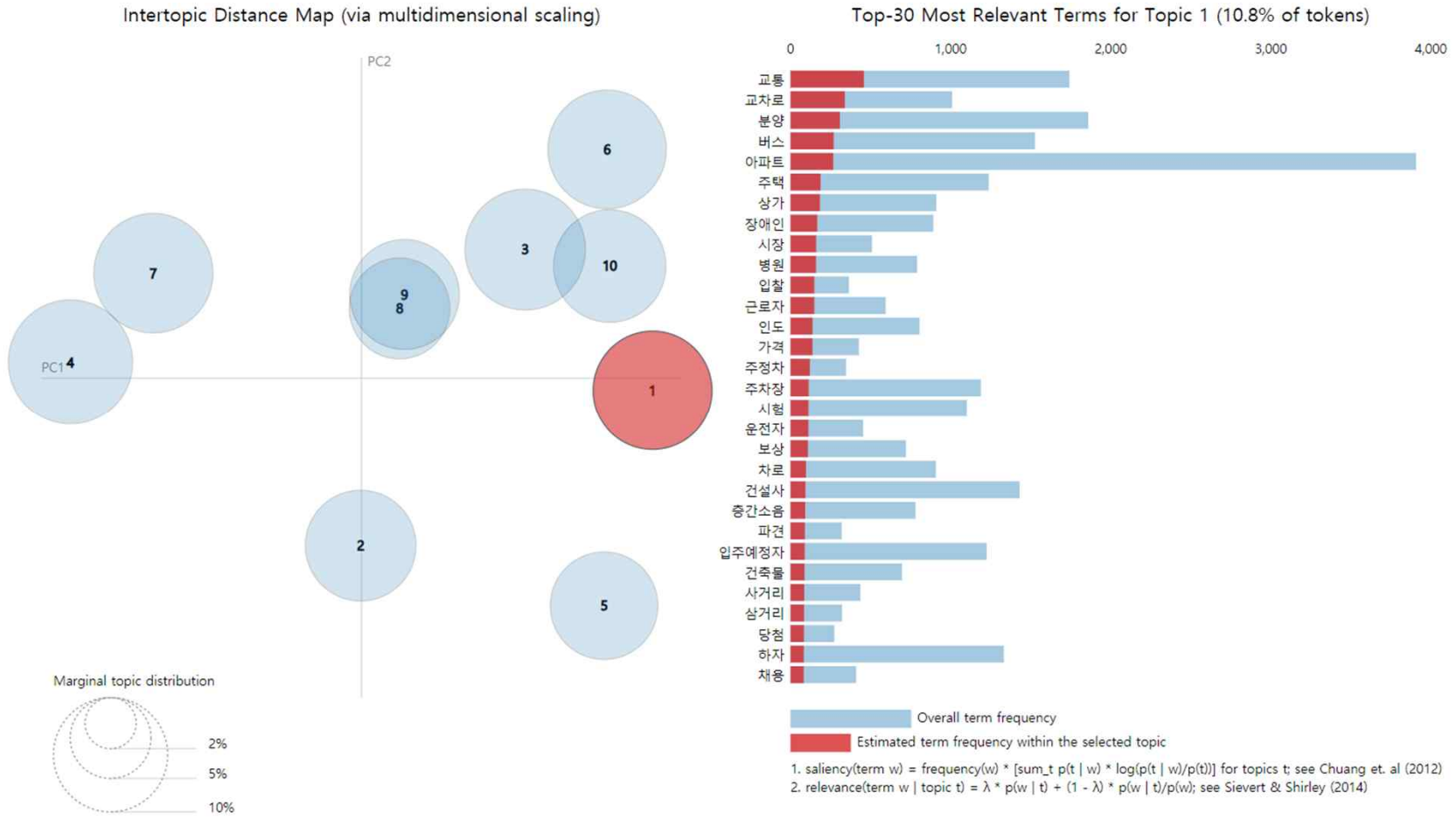
트렌드



세종특별자치시

IV. 분석결과_지역별

토픽모델링 :세종시 1번토픽은 교통, 교차로, 버스, 주정차, 주차장 등 교통 관련 키워드가 주로 포함되어 있음.



경기도

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대

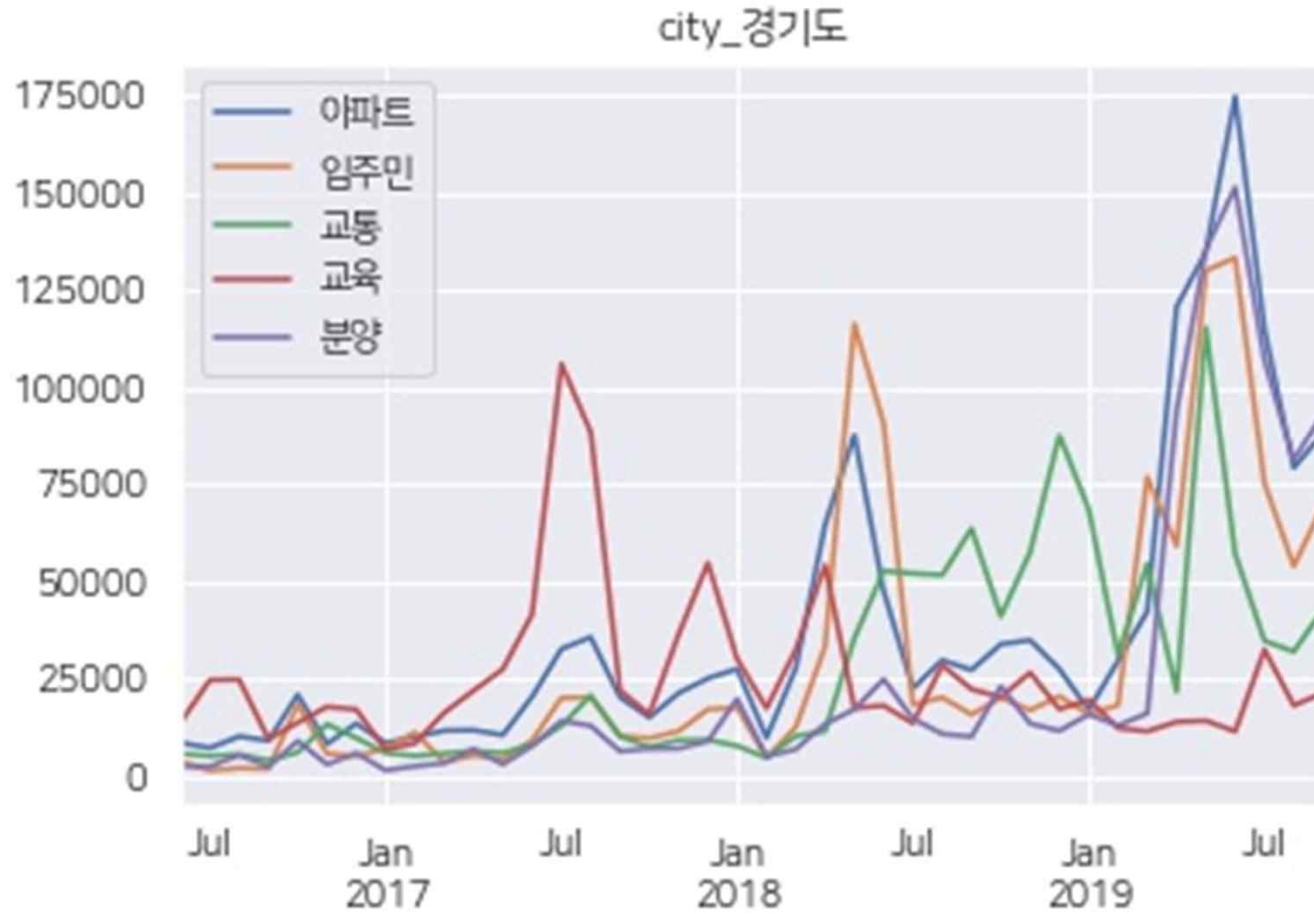


워드 클라우드 30대

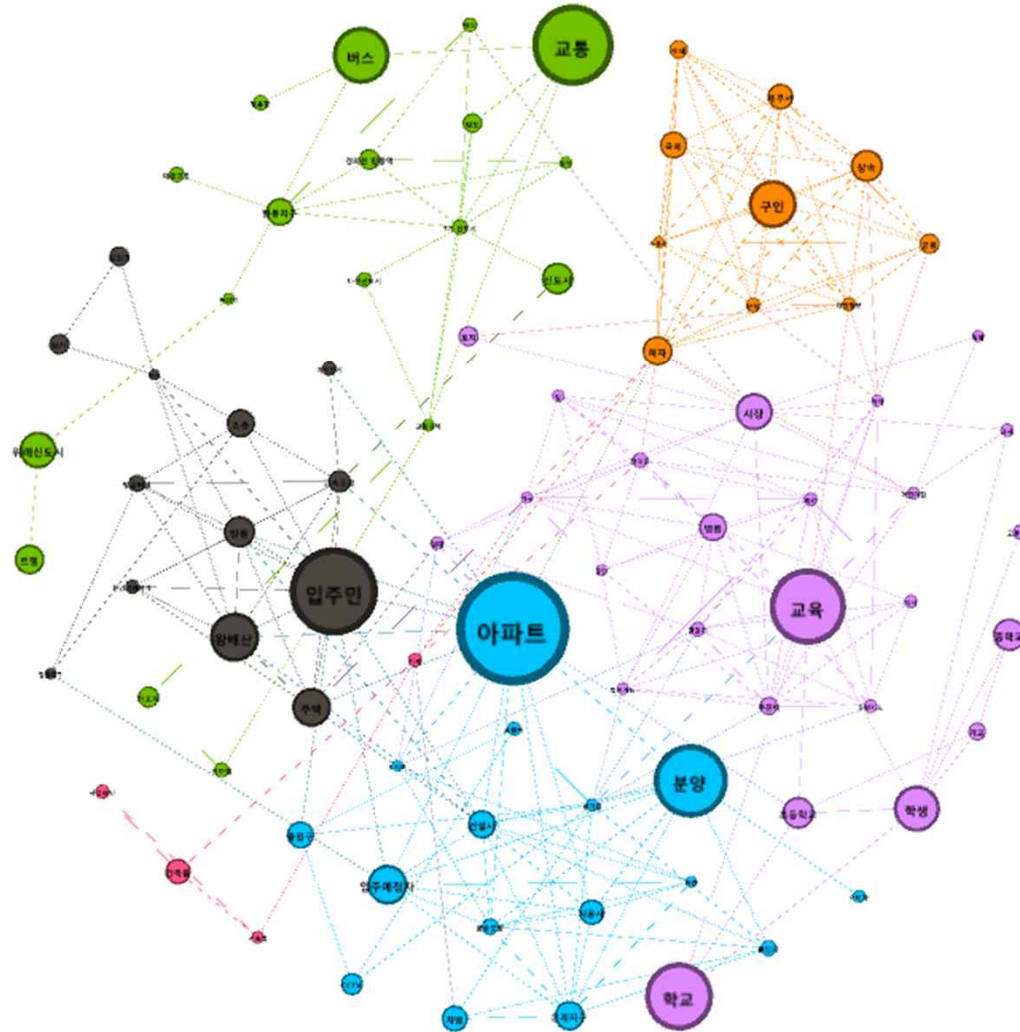


워드 클라우드 40대

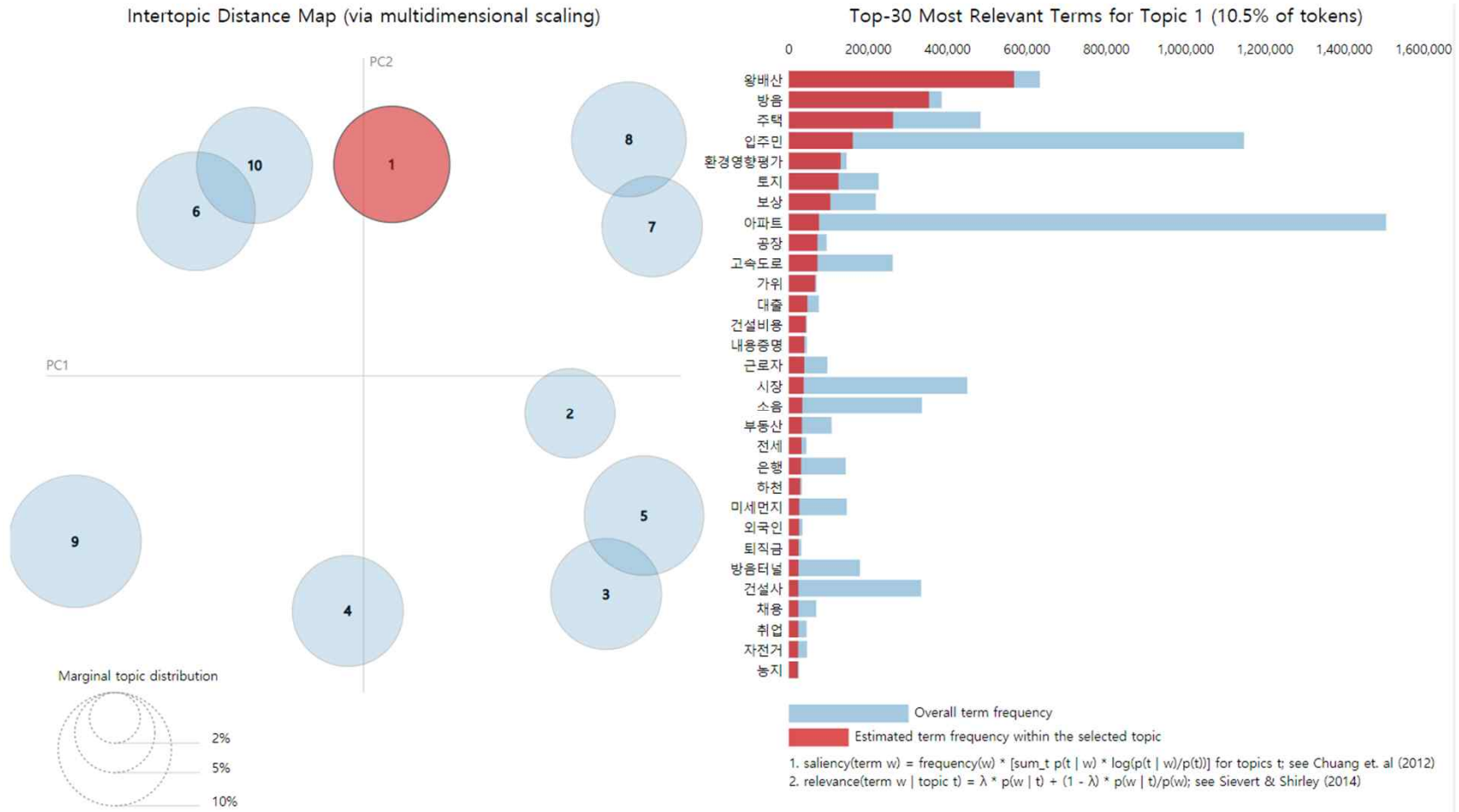
트렌드



네트워크



토픽모델링 : 경기도의 1번 토픽은 왕배산, 방음, 주택, 입주민, 아파트 등 거주여건과 관련된 키워드가 주로 포함



강원도

강원도 민원키워드

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

강원도 민원키워드

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

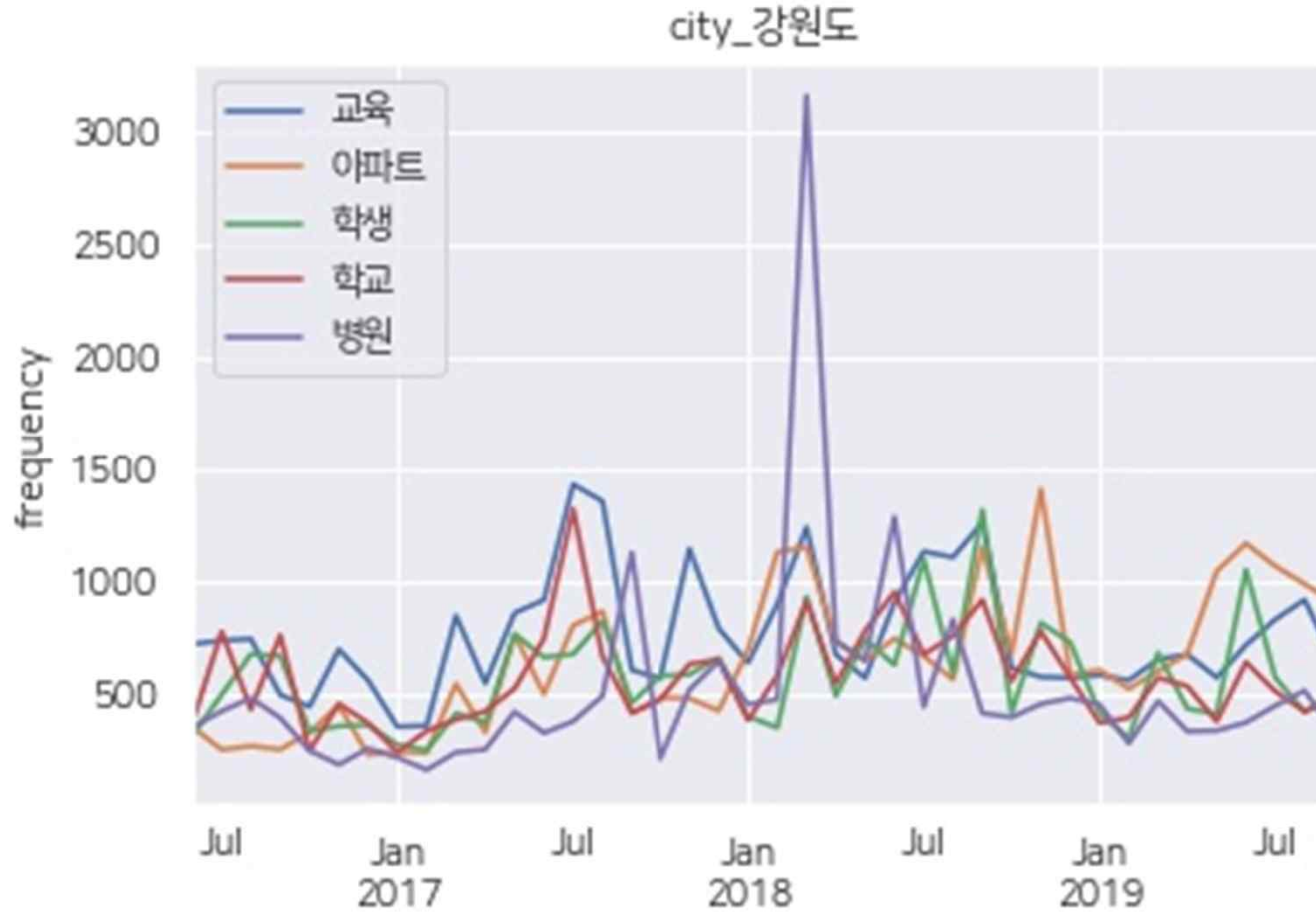


워드 클라우드 70대 이상

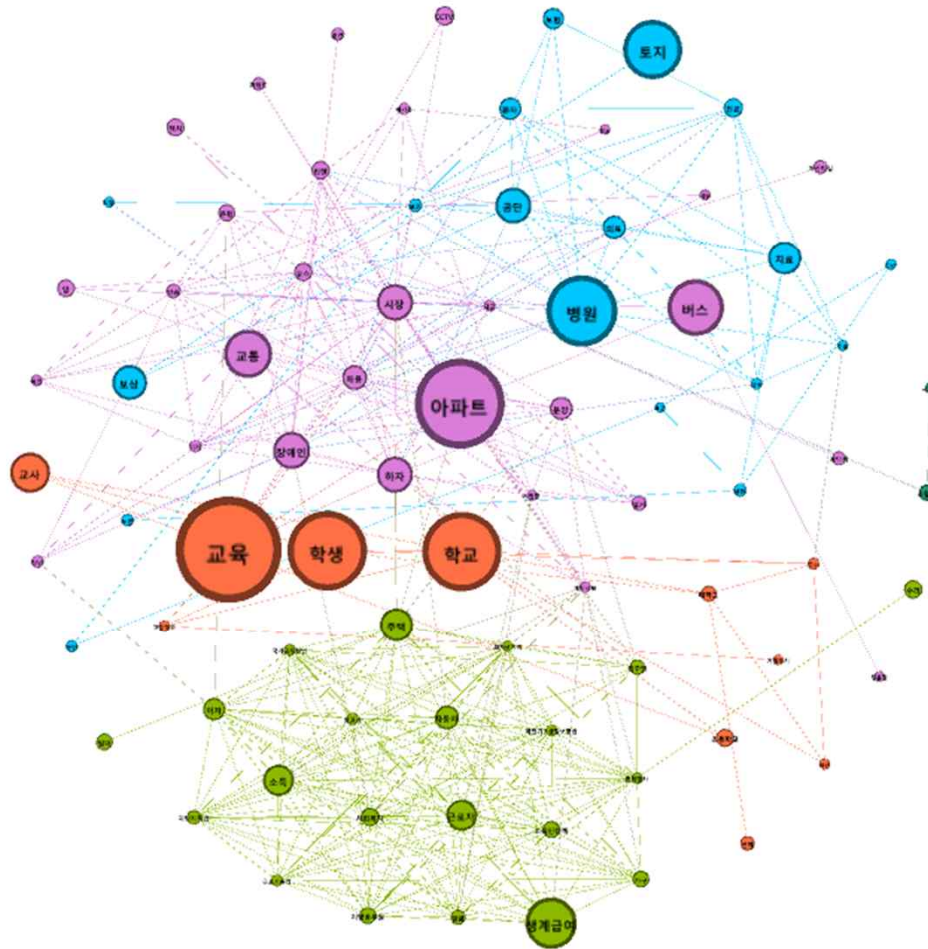


워드 클라우드 기타

트렌드

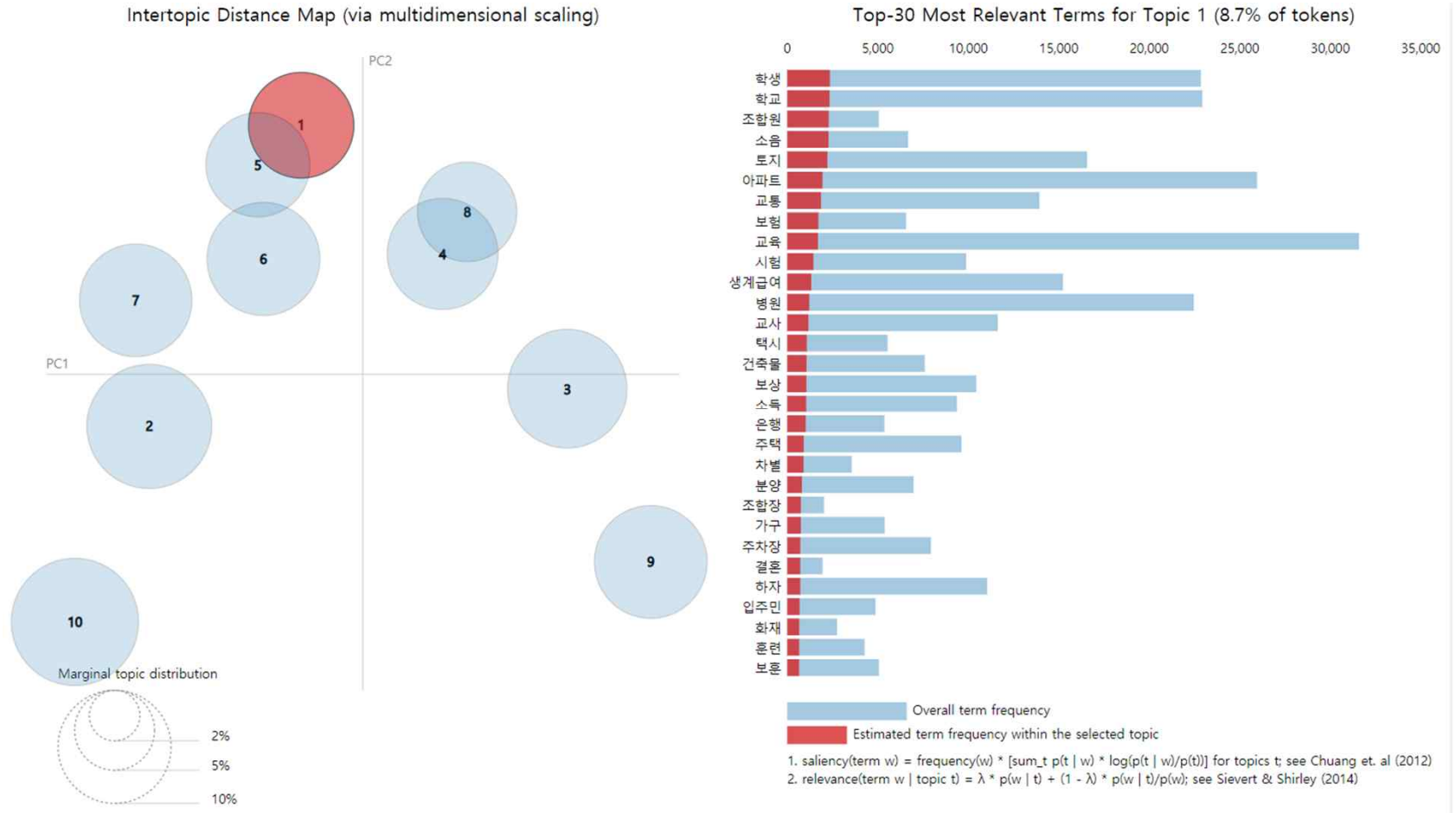


네트워크



강원도 민원키워드

토픽모델링 : 강원도의 1번 토픽은 학생, 학교, 교육, 시험 등 교육 관련 연관어와 여타 키워드 들이 골고루 포함



총청복도

충청북도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

충청북도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 50대



워드 클라우드 60대

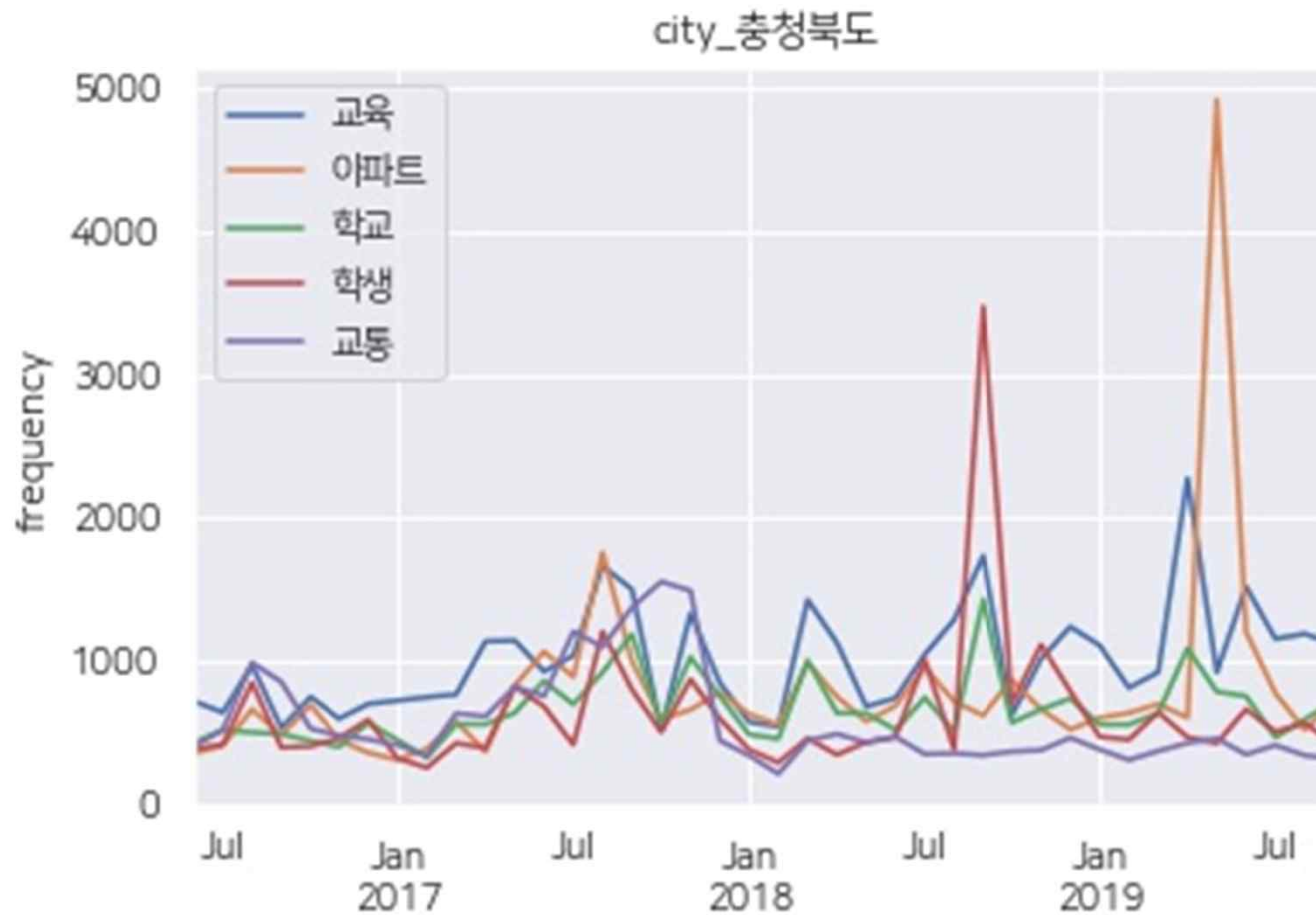


워드 클라우드 70대 이상

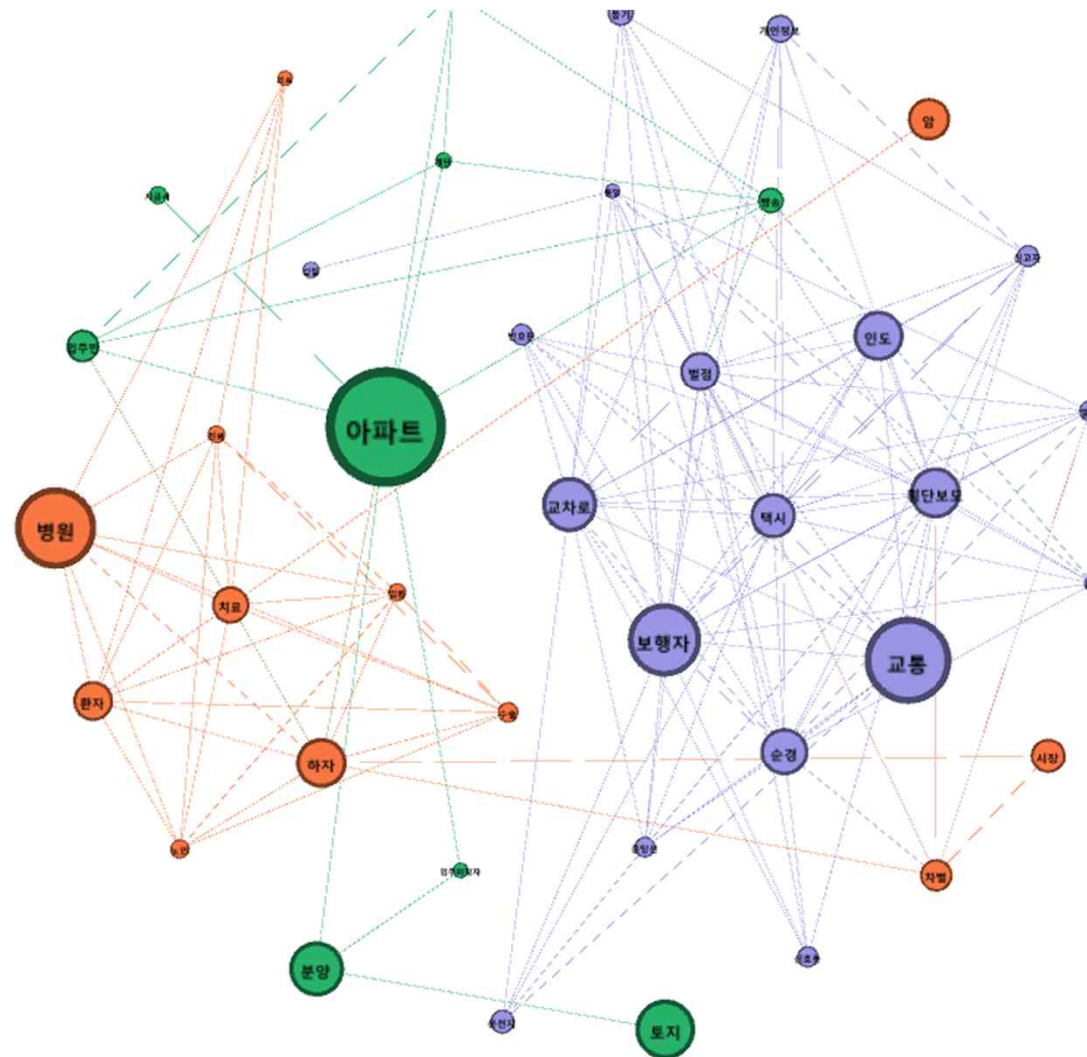


워드 클라우드 기타

트렌드



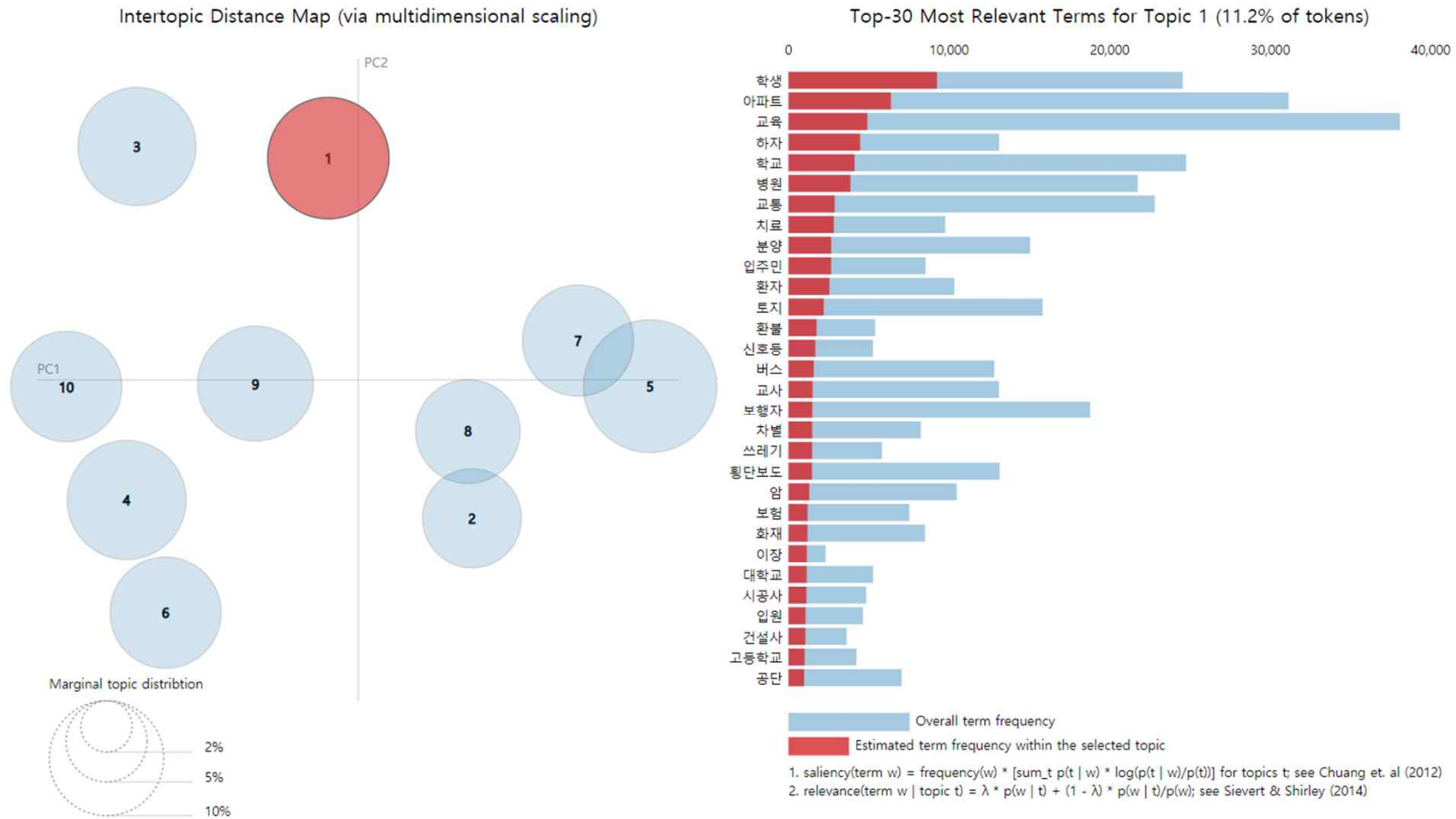
네트워크



충청북도 민원키워드

IV. 분석결과_지역별

토픽모델링 : 충북 1번 토픽은 학생, 교육, 학교 등 교육 관련 키워드와 아파트, 하자, 분양, 입주민 등 거주 관련 키워드가 동시에 포함되어 있음.



충청남도

충청남도 민원키워드

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

충청남도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

충청남도 민원키워드

IV. 분석결과_지역별

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

충청남도 민원키워드

여성



워드 클라우드 50대



워드 클라우드 60대



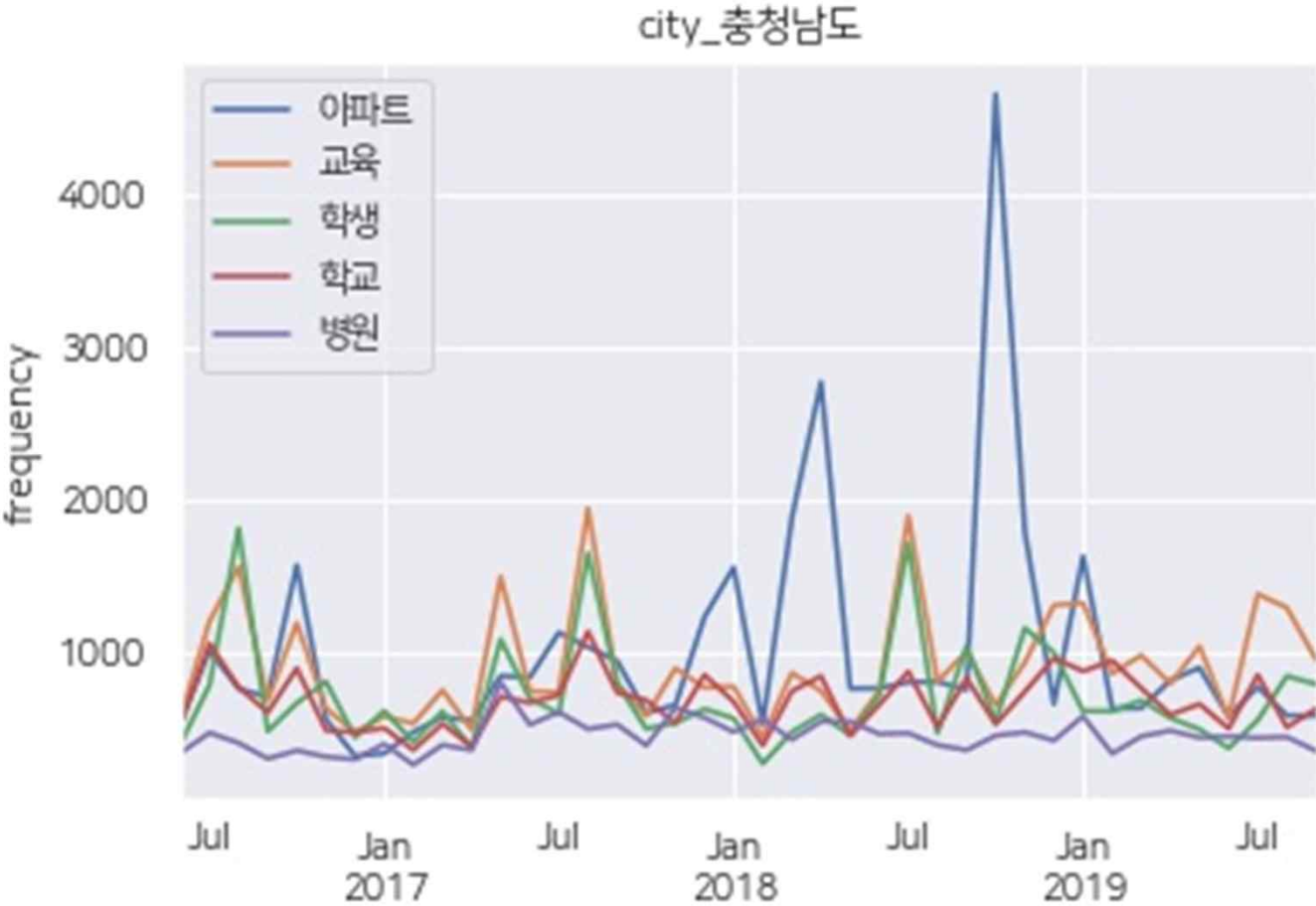
워드 클라우드 70대 이상



워드 클라우드 기타

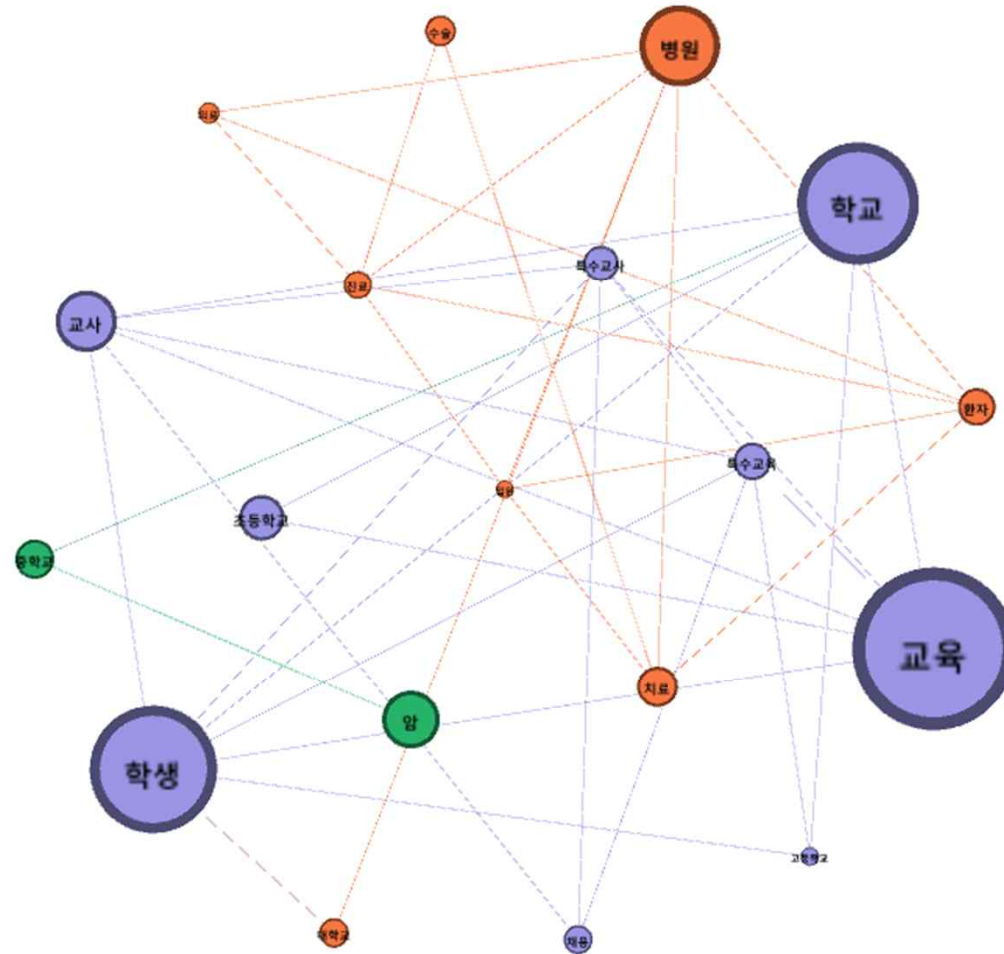
충청남도 민원키워드

트렌드



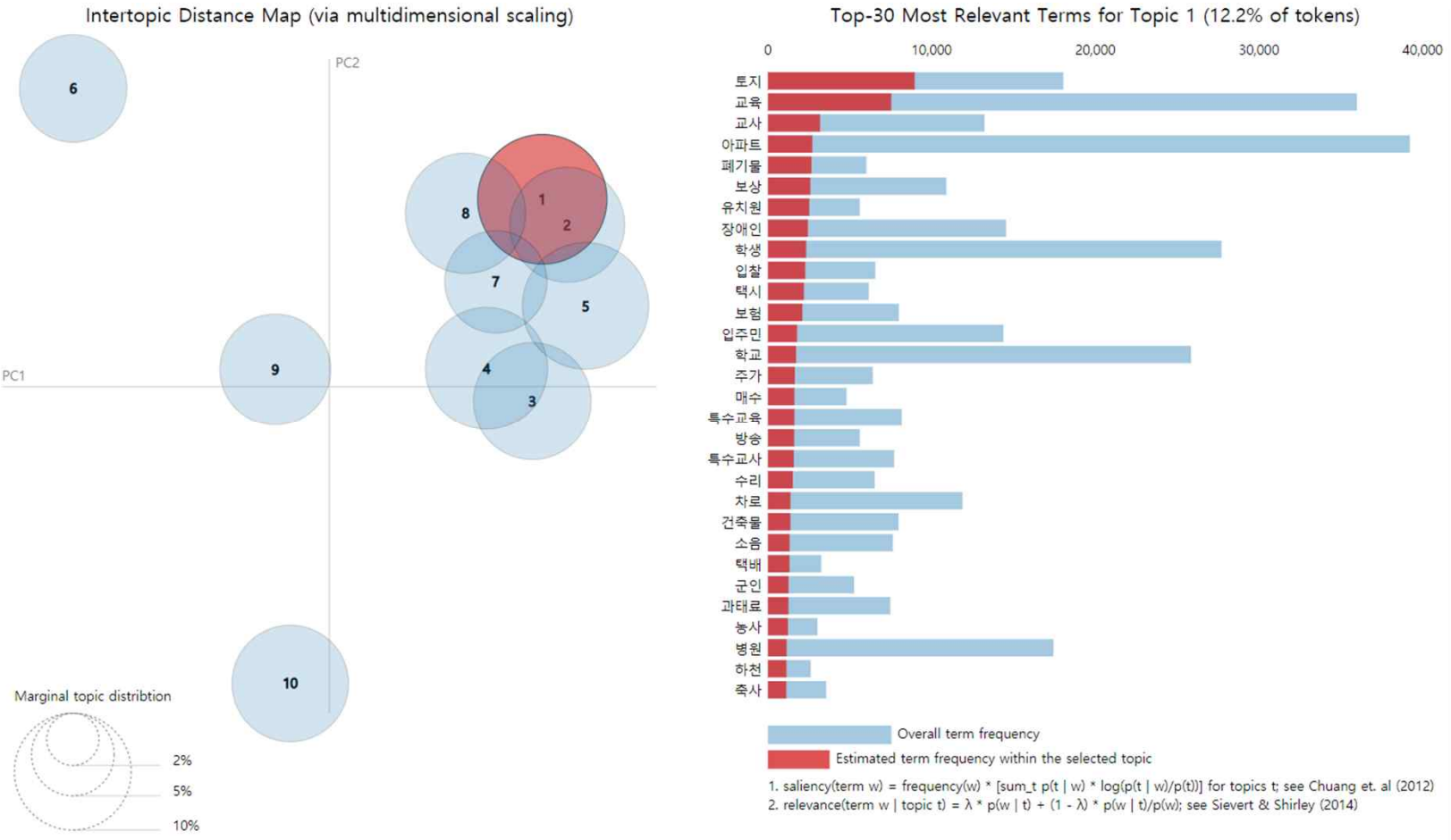
충청남도 민원키워드

네트워크



충청남도 민원키워드

토픽모델링 : 충남 1번 토픽은 교육, 교사, 유치원, 학생 등 교육 관련 키워드가 많이 포함되어 있음.



전라북도

전라북도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

전라북도 민원키워드

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

전라북도 민원키워드

여성



워드 클라우드 50대



워드 클라우드 60대



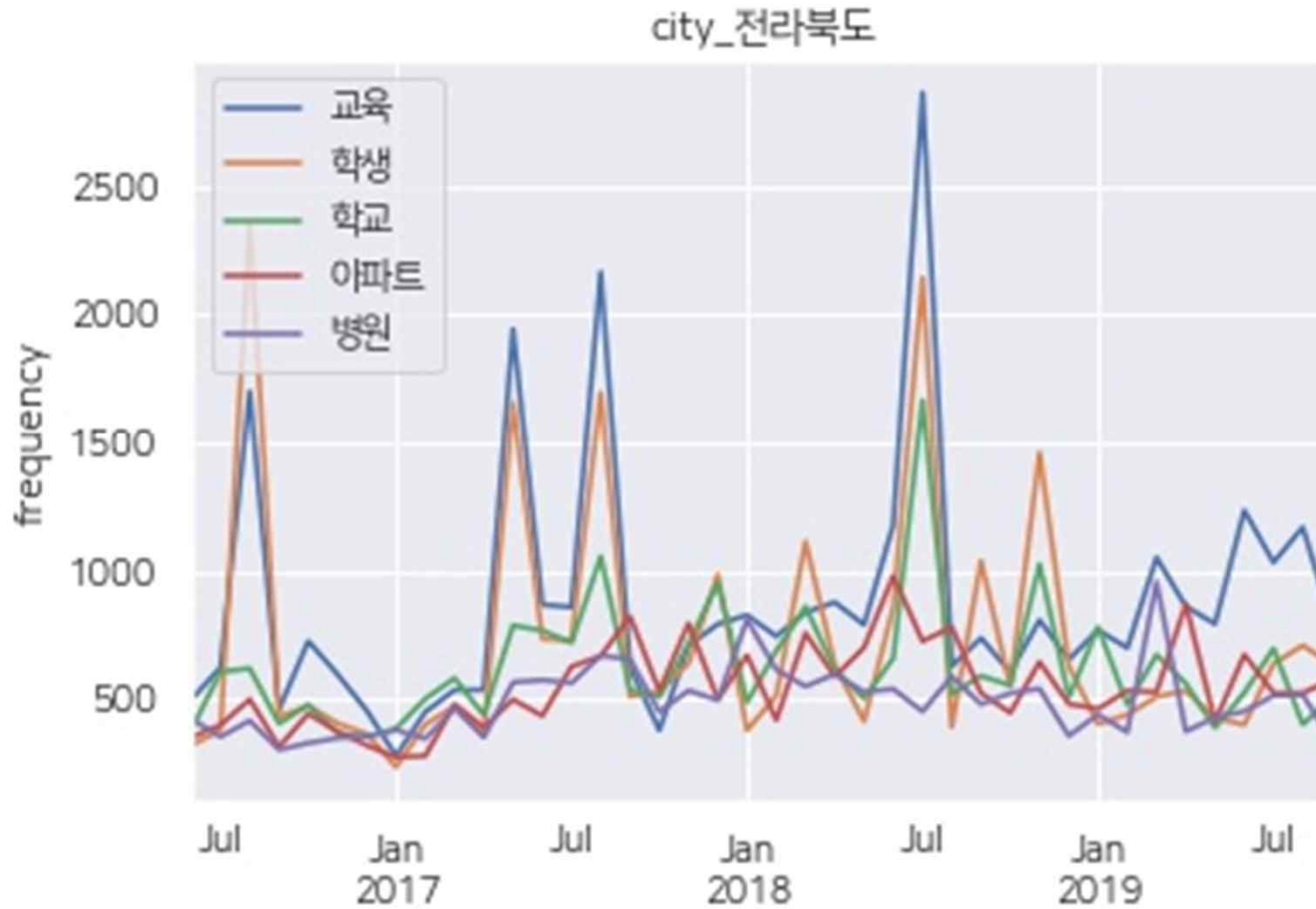
워드 클라우드 70대 이상



워드 클라우드 기타

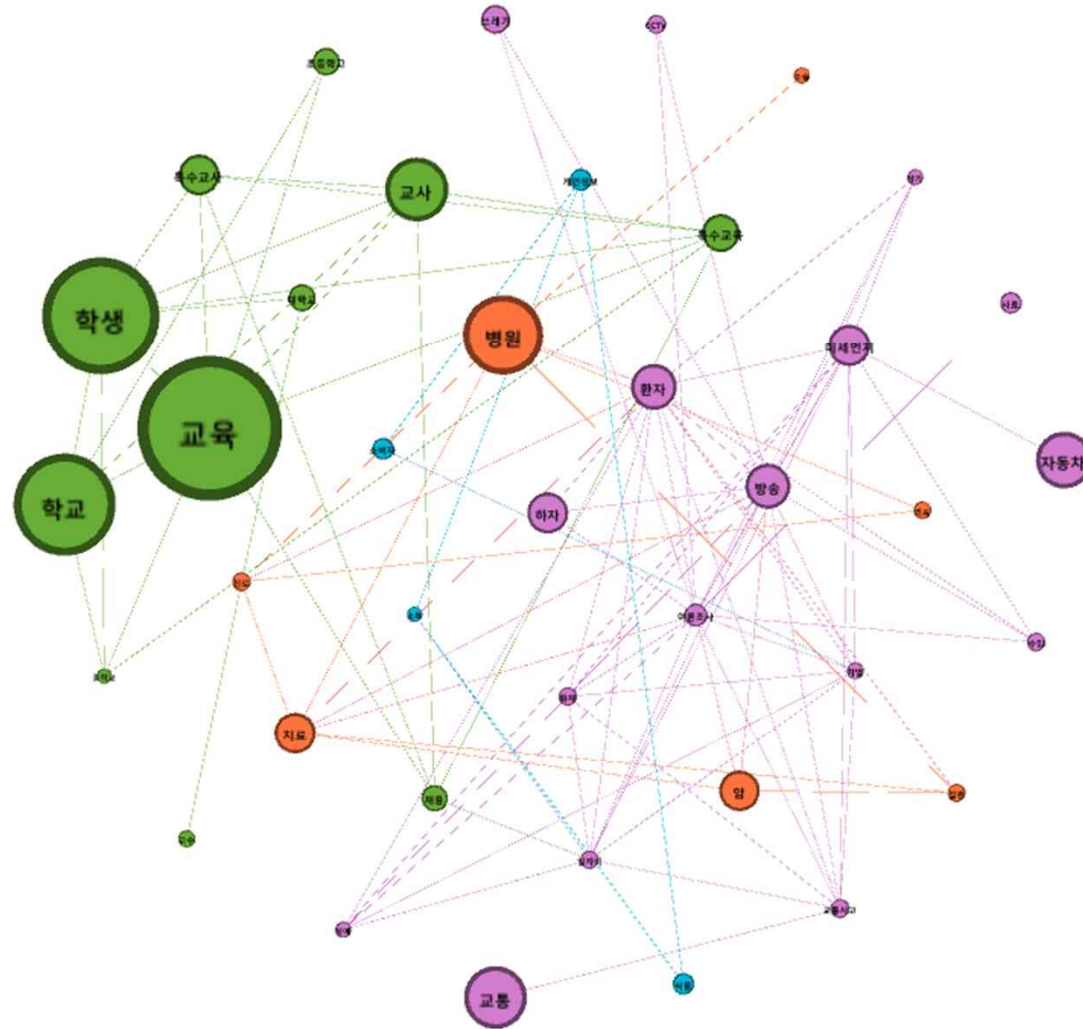
전라북도 민원키워드

트렌드



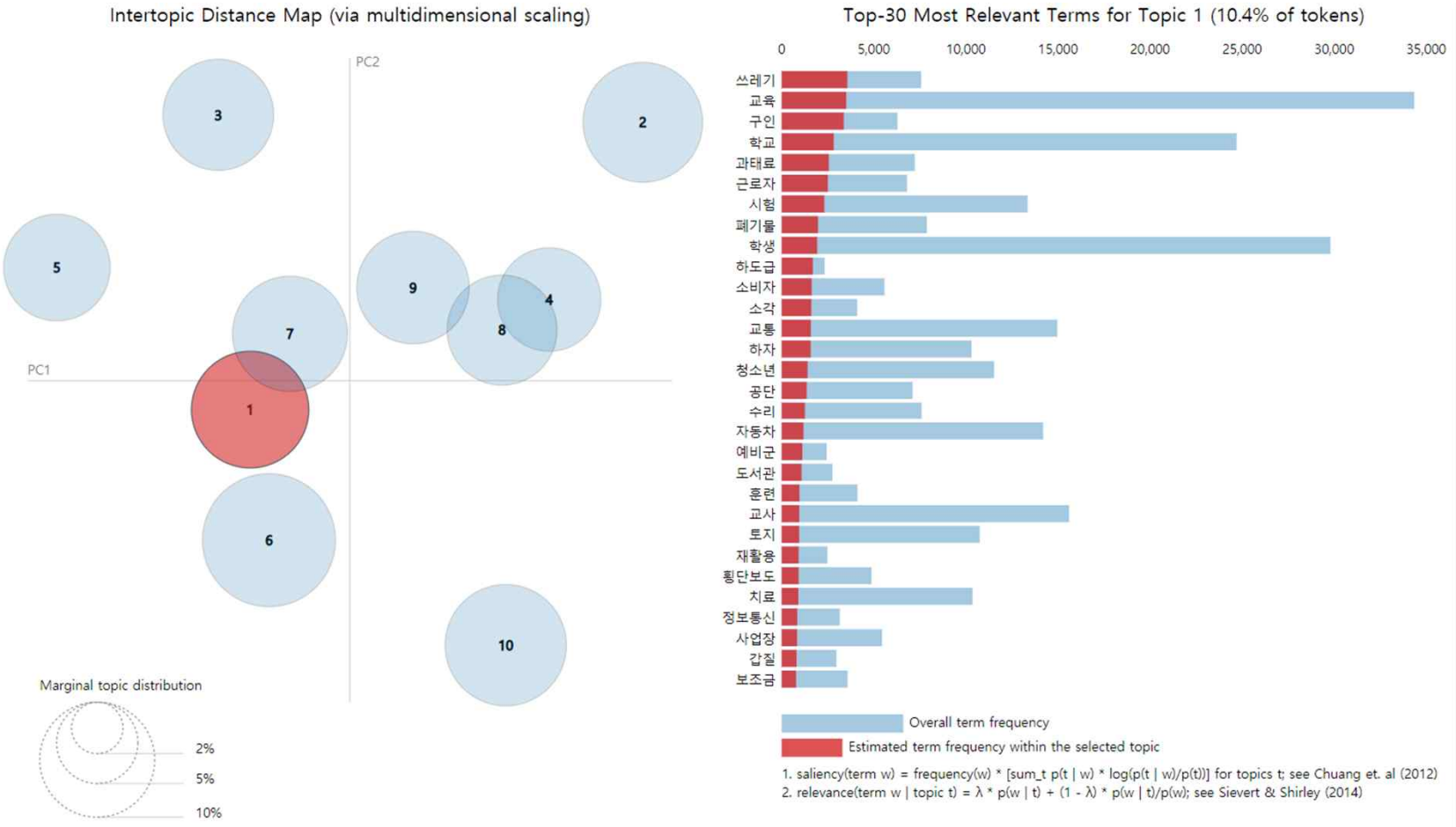
전라북도 민원키워드

네트워크



전라북도 민원키워드

토픽모델링 : 전북의 1번 토픽은 쓰레기, 과태료, 폐기물, 소각, 재활용 등 폐기물 관련 키워드가 다수 포함됨.



전라남도

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

전라남도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상

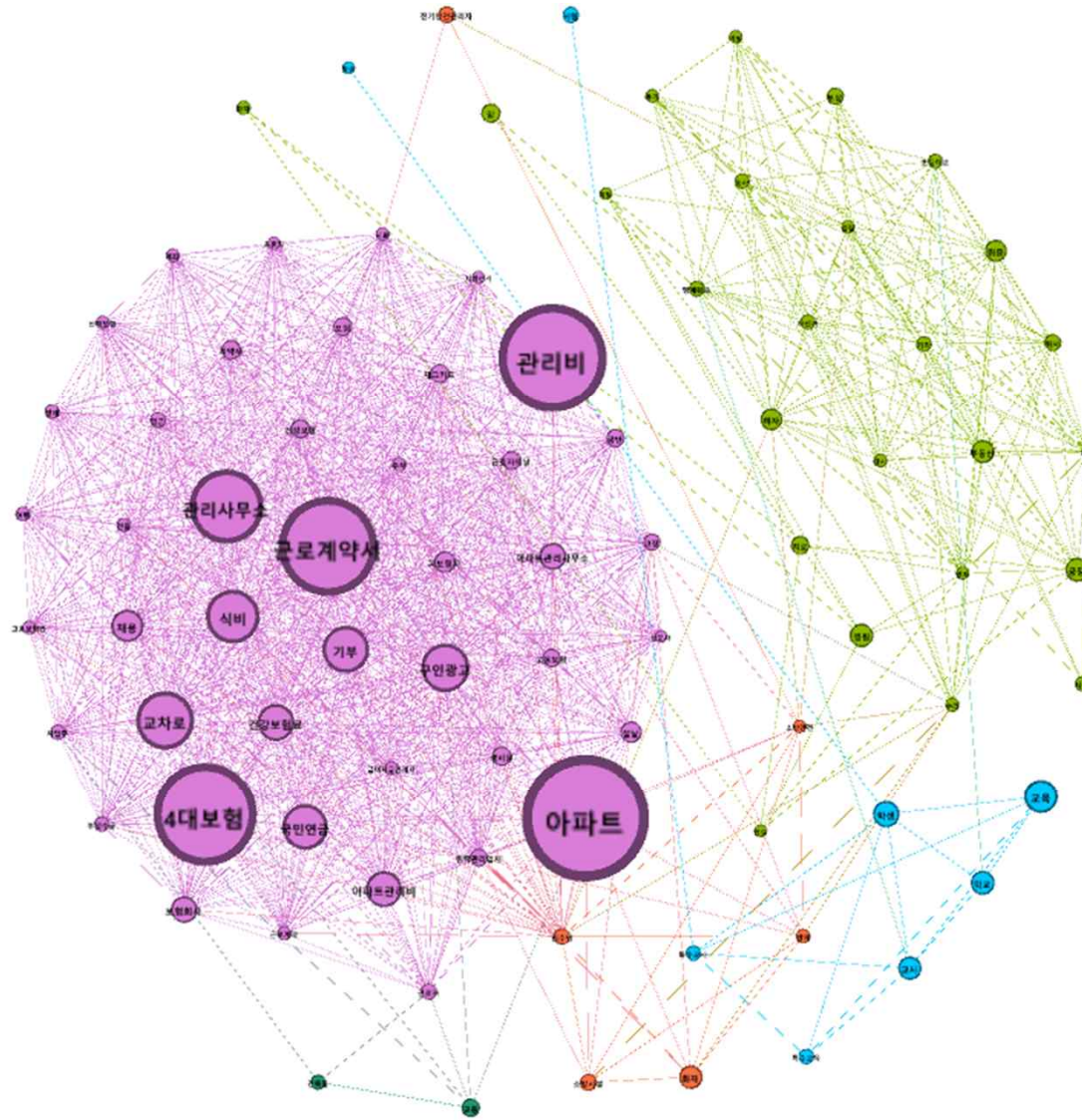


워드 클라우드 기타

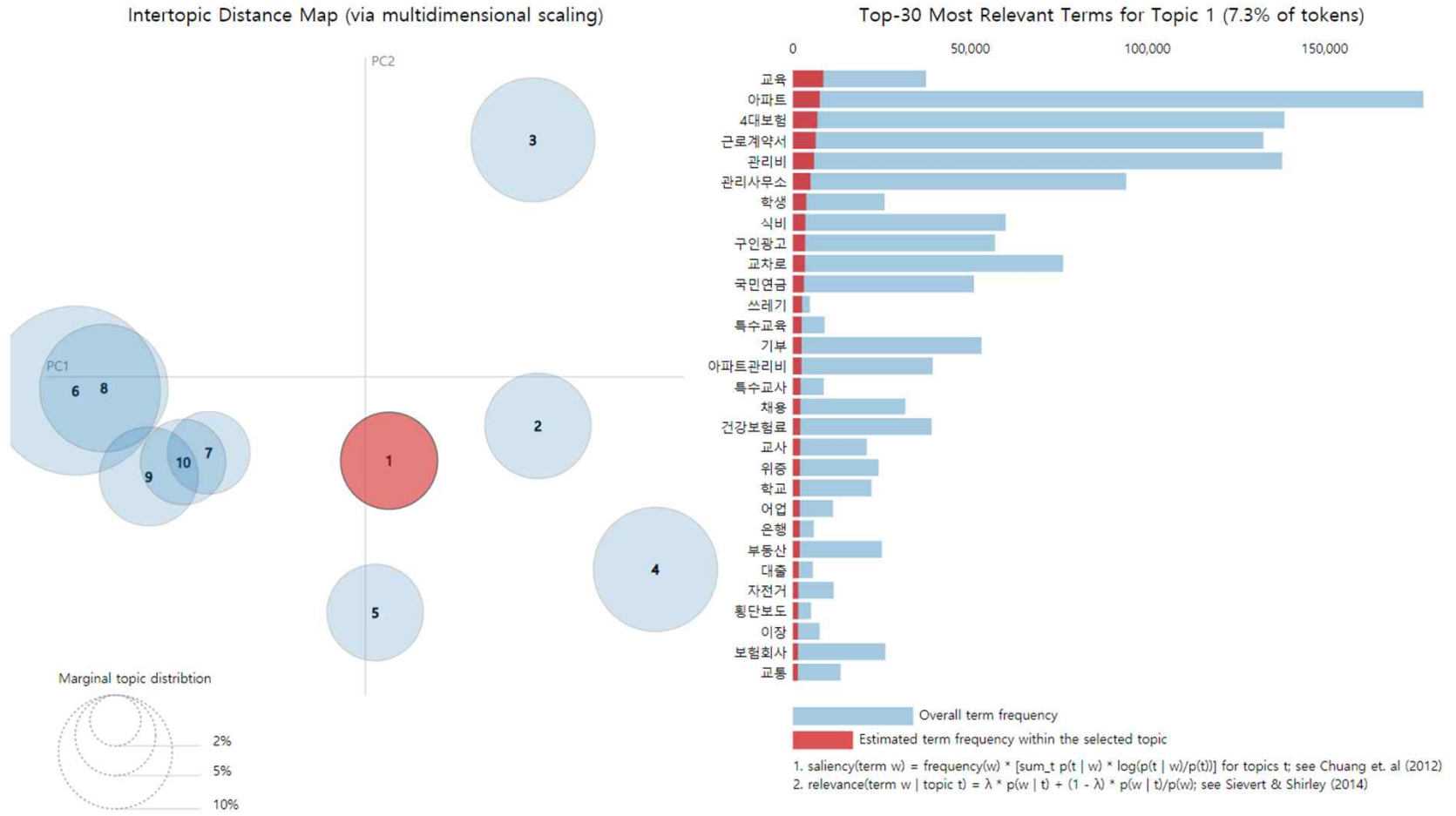
트렌드



네트워크



토픽모델링: 전남의 1번 토픽은 매우 다양한 키워드들이 포함되어 특별한 연관성 파악은 어려움.



경상북도

경상북도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

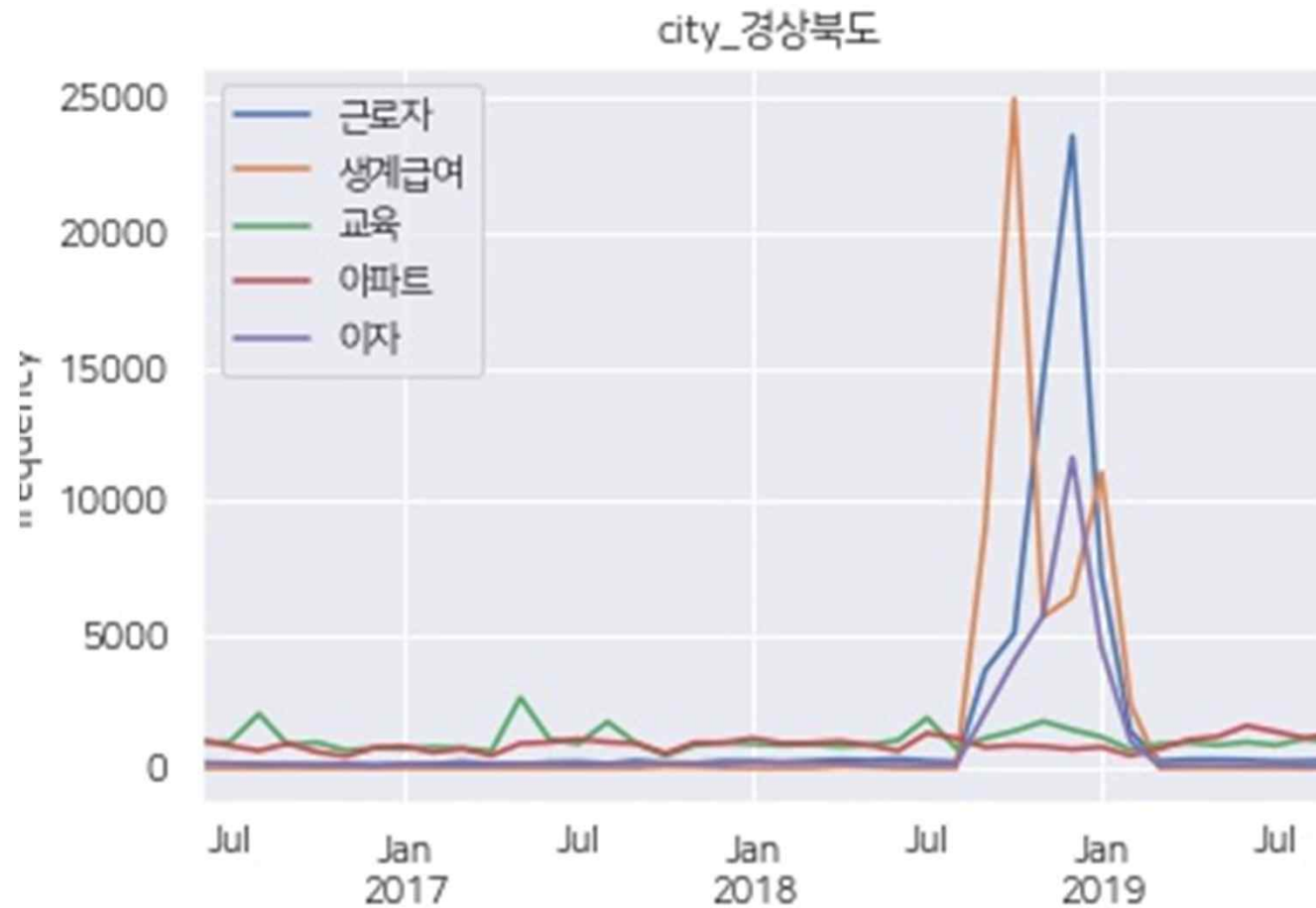


워드 클라우드 70대 이상

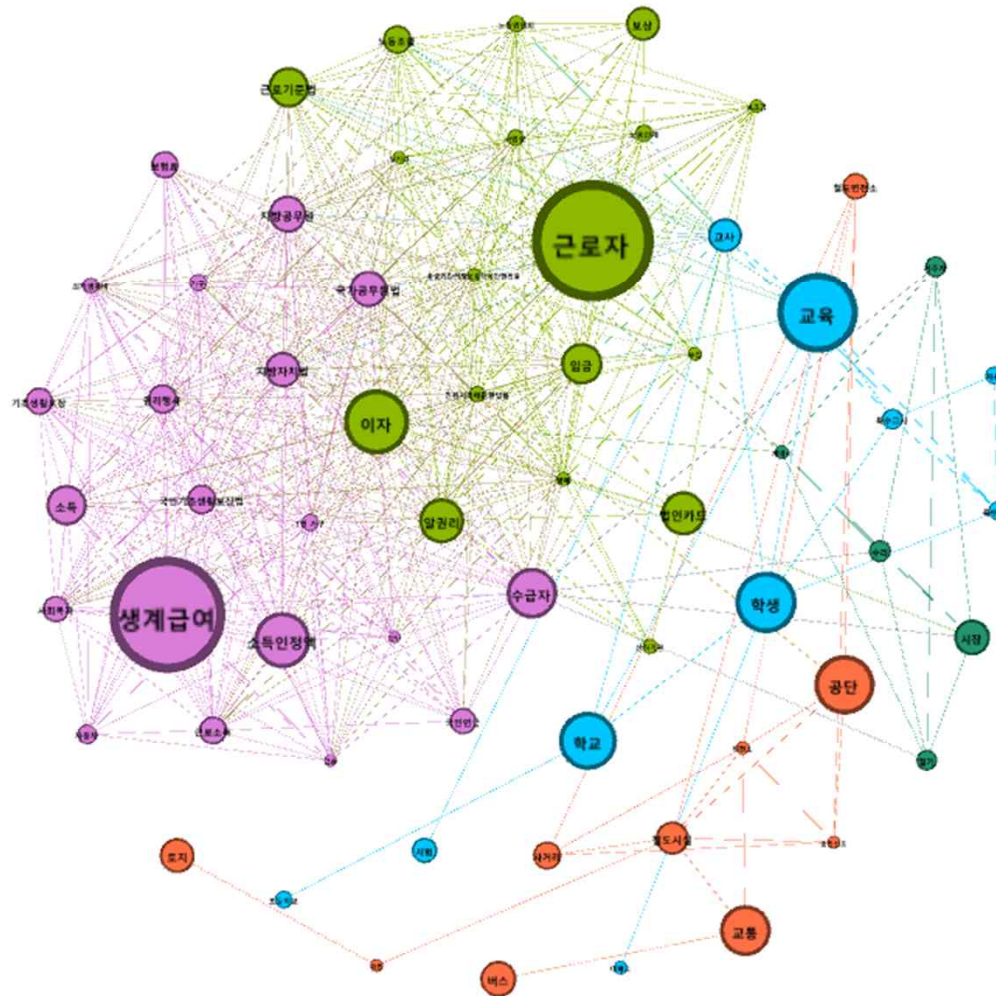


워드 클라우드 기타

트렌드



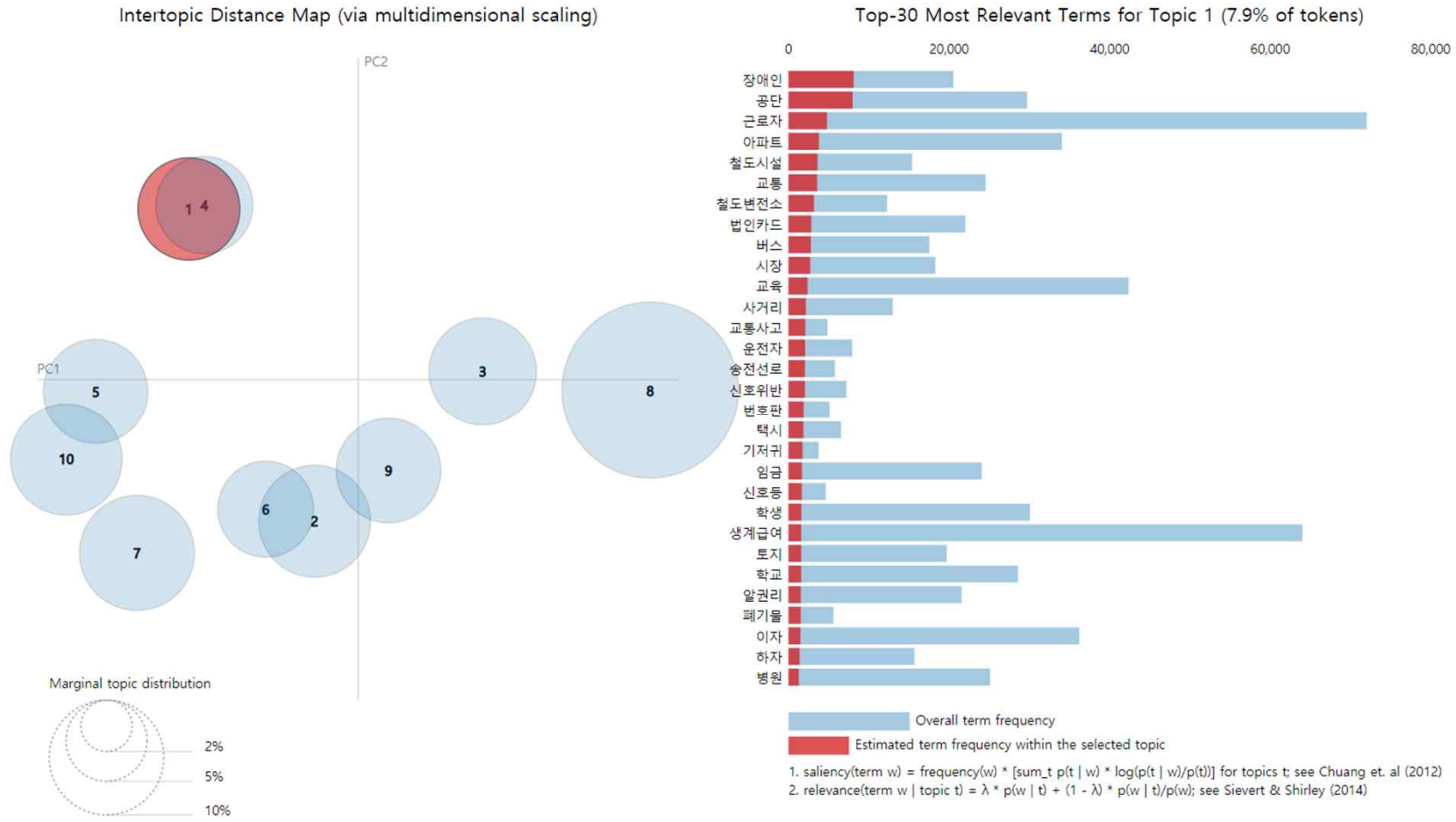
네트워크



경상북도 민원키워드

IV. 분석결과_지역별

토픽모델링 : 경북의 1번 토픽은 장애인, 공단, 근로자, 생계급여 등 복지와 관련된 키워드가 다수 포함됨.



경상남도

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

여성



워드 클라우드 50대



워드 클라우드 60대

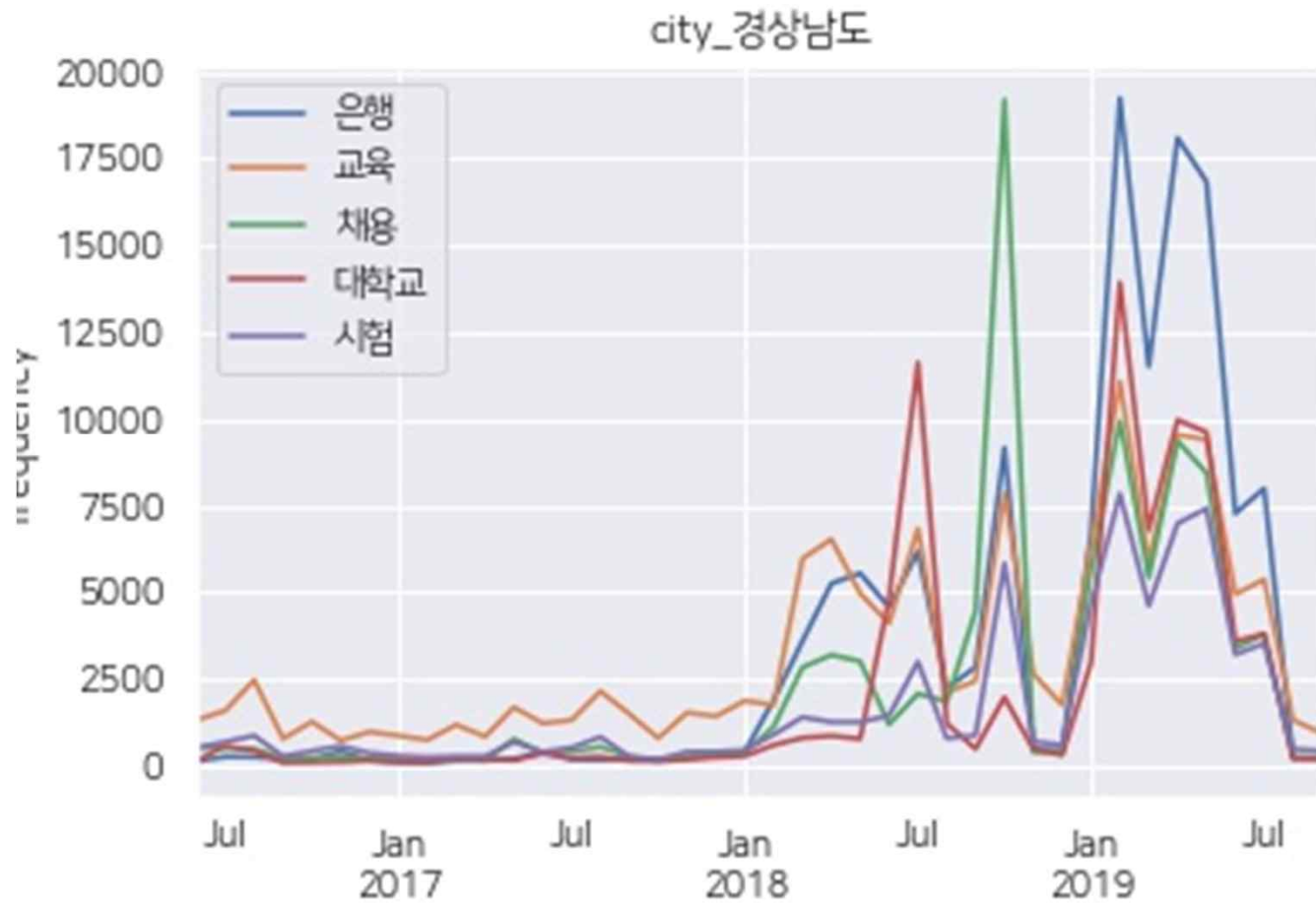


워드 클라우드 70대 이상

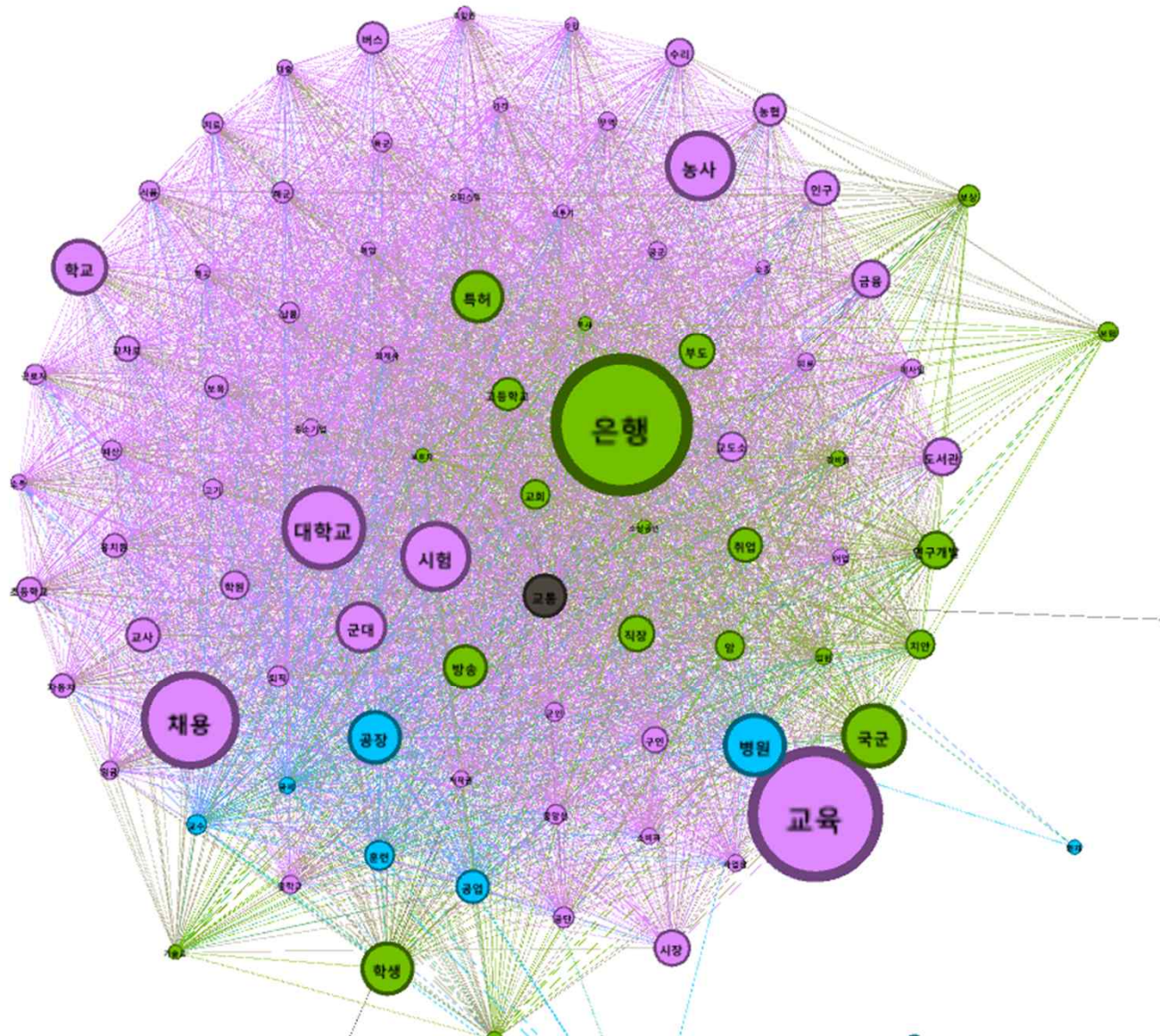


워드 클라우드 기타

트렌드



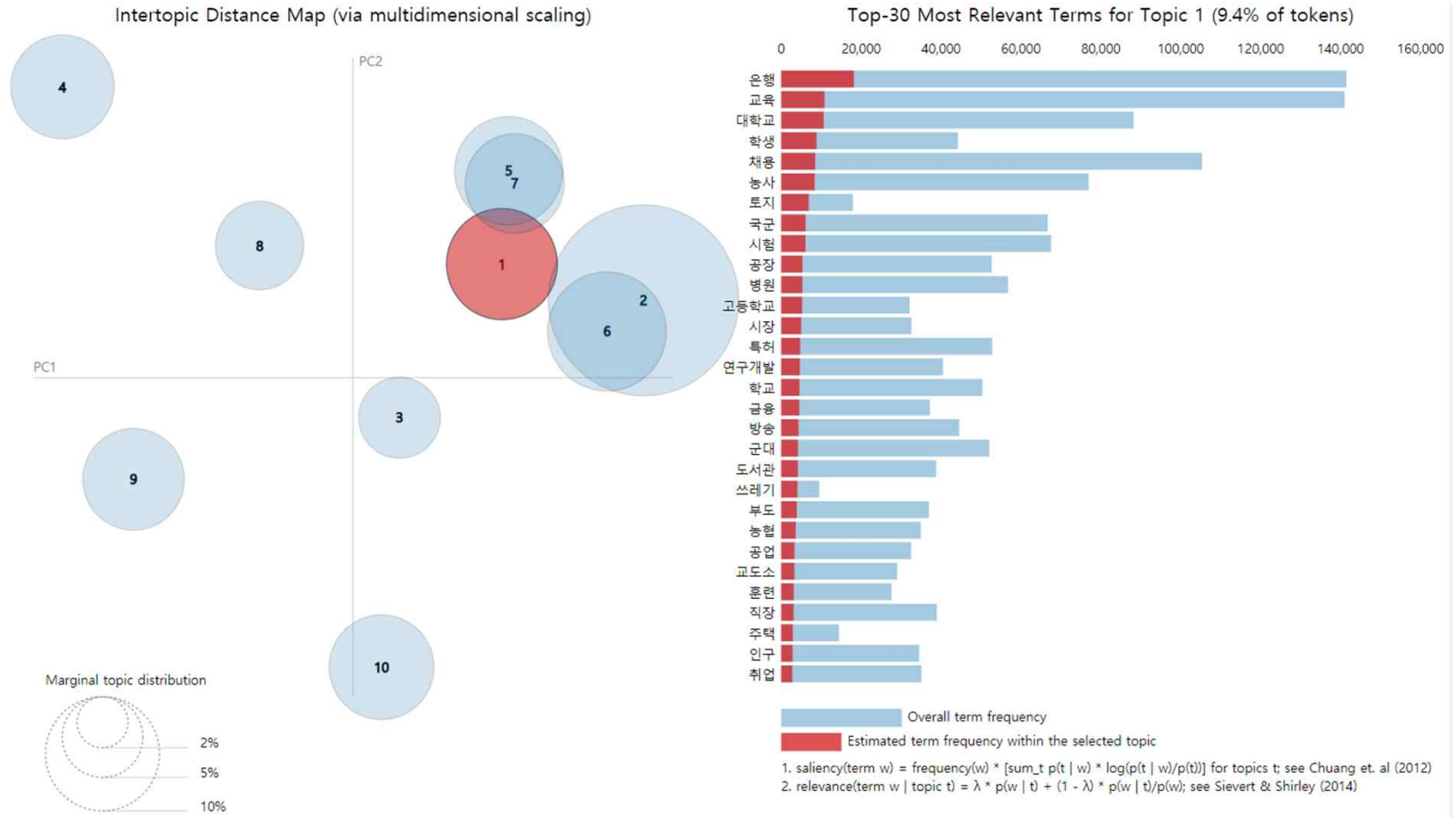
네트워크



경상남도 민원키워드

IV. 분석결과_지역별

토픽모델링 : 경남의 1번 토픽은 교육, 대학교, 학생, 시험, 고등학교 등 교육 관련 키워드가 주로 포함되어 있음.



제주특별자치도

제주특별자치도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

제주특별자치도 민원키워드

IV. 분석결과_지역별

남성



워드 클라우드 50대



워드 클라우드 60대



워드 클라우드 70대 이상



워드 클라우드 기타

여성



워드 클라우드 10대



워드 클라우드 20대



워드 클라우드 30대



워드 클라우드 40대

제주특별자치도 민원키워드

여성



워드 클라우드 50대



워드 클라우드 60대

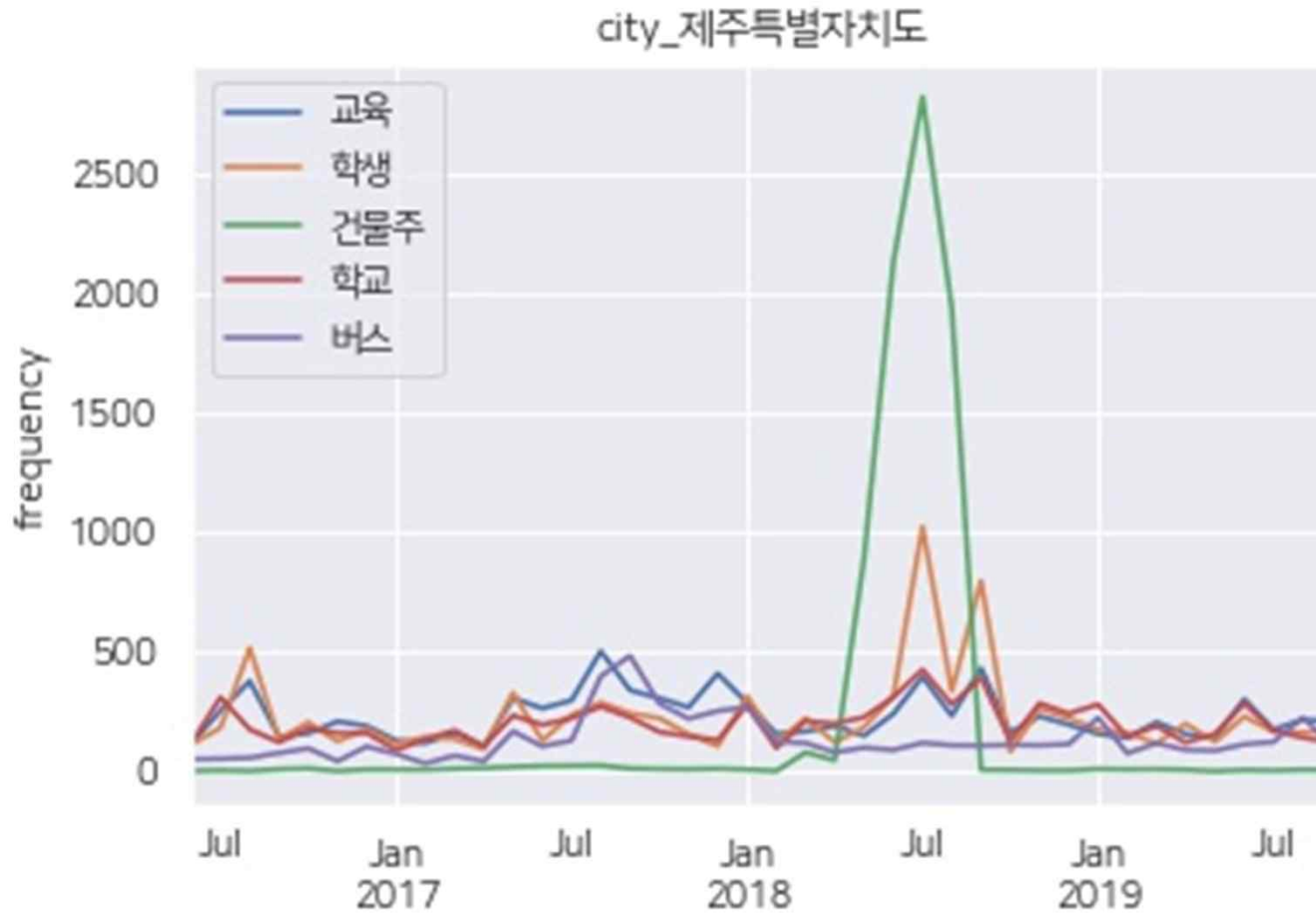


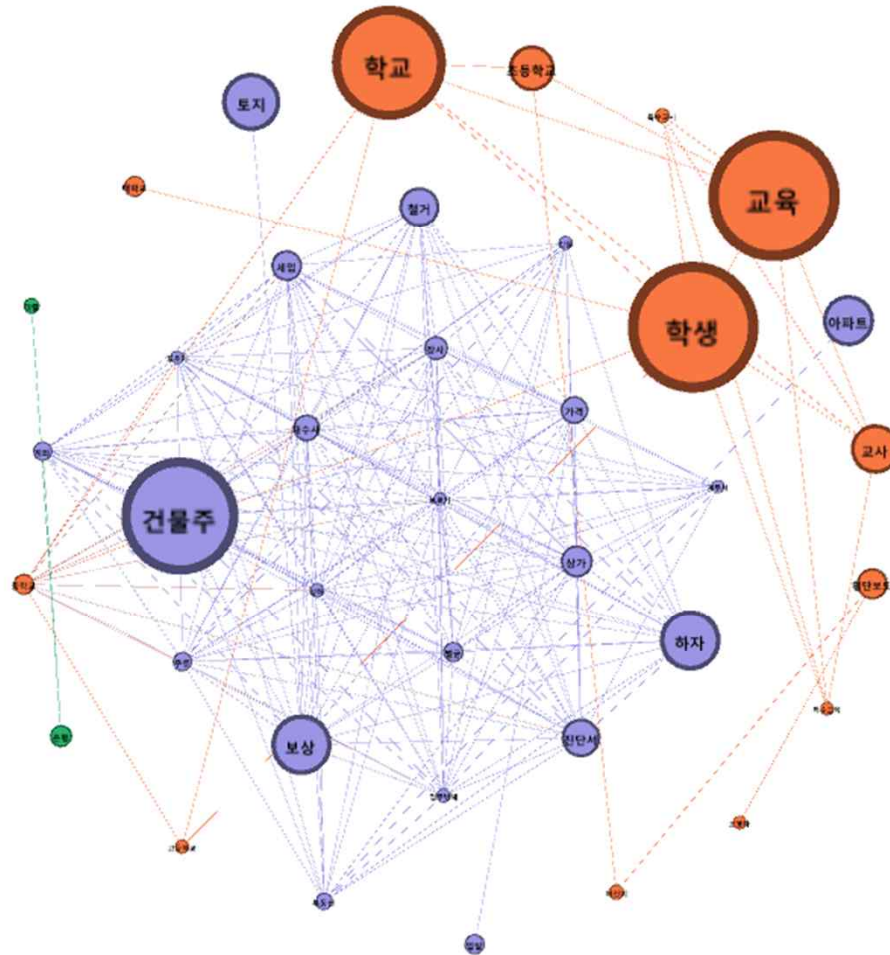
워드 클라우드 70대 이상



워드 클라우드 기타

트렌드

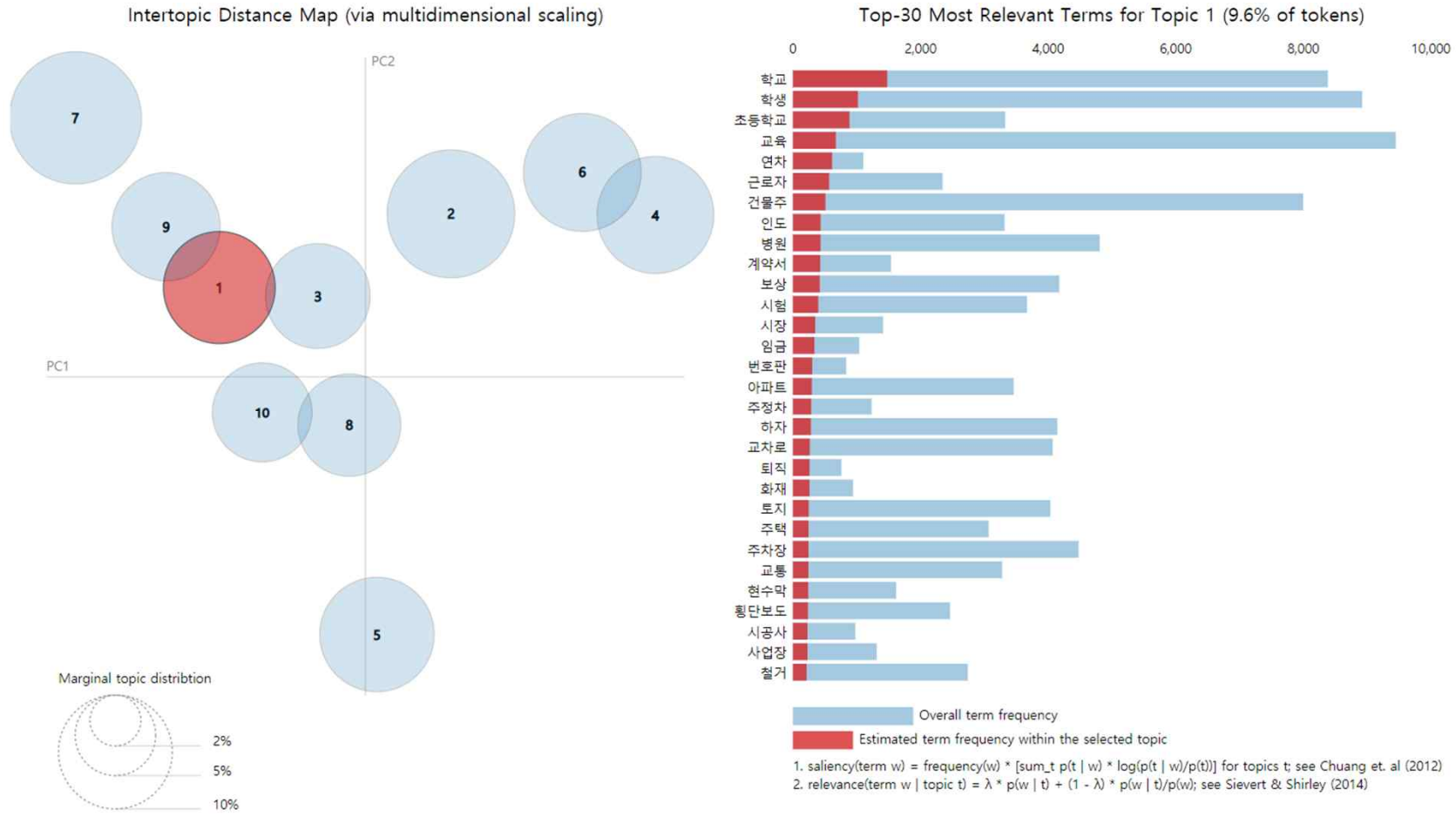




제주특별자치도 민원키워드

IV. 분석결과_지역별

토픽모델링 : 제주도의 1번 토픽은 학교, 학생, 초등학교, 교육 등 교육과 관련된 키워드가 다수 포함됨.



부산광역시 - 해운대구

심층분석 사례 : 부산시 해운대구

IV. 분석결과_해운대구

"부산광역시 해운대구" 의 민원데이터를 분석한 결과 차량, 위반, 관리, 촬영, 학교, 사람, 변경, 전화 등의 키워드가 높은 빈도수를 차지하고 있음.

키워드 분석 결과

단어	빈도 수
차량	5,524
위반	5,325
관리	2,860
촬영	2,845
학교	2,435
사람	2,373
변경	2,367
전화	2,365
휴대폰	2,229
사용	2,184
시설	2,165
차선	2,157
국민	1,934
아파트	1,906
사진	1,724
본인	1,681
감면	1,656
공사	1,631
경찰	1,630
운영	1,623
...	...

워드 클라우드



심층분석 사례 : 부산시 해운대구

남성은 차량, 위반, 차선 등 교통과 관련된 민원이 많은 것을 알 수 있고 여성은 학교, 학생 등 일상 생활과 관련된 민원이 많은 것을 알 수 있음. 이것은 남성과 여성 간 민원의 내용이 뚜렷하게 구분되는 것을 보여줌.

워드 클라우드(남)



차선 바꾸는 차량 굴라 예쿠스로 '중'...2억 뺏어낸 일당 검거

부산뉴스 2019년 05월 06일 09:00

자신을 변장하여 차량을 굴라 고급 승용차로 고위사고를 낸 뒤 탈락...으로 2억여가 넘는 자금을 뺏아낸 일당이 경찰에 붙잡혔다.

부산 해운대로 중앙버스전용차로 위반차량, '과태료' 부과

부산 해운대로 중앙버스전용차로 위반차량, '과태료' 부과

부산시는 9일 보험사기방지특별법 위반 혐의로 이모씨(21)등 [부산=뉴스1]이상진 기자 = 부산시는 지난 6월말 개통한 해운대로 중앙버스전용차로-운전자에 대한 중앙버스전용차로 위반차량(BRT) 위반 차량을 단속해 오는 12월부터는 과태료를 부과할 것이라고 6일 밝혔다.

시는 시민이승객과 함께 한승객에를 유치하고 안전강화체계(BRT)의 중앙버스전용차로 위반차량을 단속한다. 또 이 구간을 운행하는 141대 노선(노선 315와 144번 노선)에 3대에도 단속장비를 장착, 이동 단속을 병행할 계획이다.

이해 유은 지난 2017년 6월18일 부산 해운대구 연제동-풍동IC 방향으로 전교를 바꾸는 것을 발견하고 일과 후 속도를 올려 중앙 버스 전용차로와 보행길과 중앙버스전용차로를 등 운전 수법으로 최근까지 16차례에 걸쳐 2억 3000만원 상당의 이득을 장기간 탐탁을 받고 있다.

중앙버스전용차로 실선구간을 주행하는 차량에 대해 매일(토요일을 포함) 24시간 단속한다.

경찰조사 결과 이씨는 보험금을 가로채기 위해 예쿠스를 타고 해운대구 연제동 일대를 돌다다니다가 진로변경이 불가능한 도로에서 자신을 바꾸는 라 고의로 사고를 낸 것으로 나타났다.

이달말까지 후보 등 시정 운영기간에는 개도장을 발급 한 뒤 내달부터 중앙버스전용차로 위반차량에 대해 승용차와 4이하 화물차에 5만원, 승합차와 4과외 화물차에 6만원의 과태료를 각각 부과할 예정이다.

경찰은 이들에 대한 보강수사를 벌인 뒤 범행 기밀 정도에 따라 구공판유치로 운영에 적기 동참해 주실 것'을 당부했다.

부산시 집계자는 '간선급행체계(BRT)의 중앙버스전용차로 운영이 정착되면 버스 통행속도가 현중 빨라지게 될 것'이라며 '대중교통 활성화를 위해 시민들이 중앙버스전용차로 운영에 적극 동참해 주실 것'을 당부했다.

한편 부산시는 지난해 12월 30일 풍동IC-올림픽교차로 구간(3.7km)에 간선급행체계(BRT)와 함께 중앙버스전용차로를 운영한 이래 지난 4월30일 올림픽교차로에서 운전자까지 1.3km구간에 중앙버스전용차로 확장 개통했다.

부산시 집계자는 '간선급행체계(BRT)의 중앙버스전용차로 운영이 정착되면 버스 통행속도가 현중 빨라지게 될 것'이라며 '대중교통 활성화를 위해 시민들이 중앙버스전용차로 운영에 적극 동참해 주실 것'을 당부했다.

BRT(Bus Rapid Transit)는 버스전용전선로를 통해 외곽에서 도심으로 급행으로 이동할 수 있는 교통수단으로 지상의 지하철로 불릴 만큼 효율적인 교통체계로 각광받고 있다.

교통

워드 클라우드(여)



해운대 신축 아파트 '통학 안전권' 갈등

10일 입주할 무렵 단지 주민들, 초·중·고·대·보통학교 설치 촉구

국회신문 2019년 05월 06일 09:00

부산 해운대구의 한 신축 아파트 입주 예정자들이 자신의 통학 안전권 확보를 위해 투쟁에 나서자 이를 바라보는 시선이 엇갈린다. 이들의 요구가 현실을 무시한 억지라는 비판이 있는가 하면, 부프로써 당연한 요구라며 공감하는 의견

해운대구 우동 A아파트 입주예정자들의 해운대구에 통학 안전 대책을 마련해줄 것을 요구하며 아파트 초등학교가 1인도가 없다. 이에 따라 학생들은 찾았을다. 협의하는 인도가 없는 구간에 보통돌 나차할 볼라도를 설치해 달라고 구에 요구

"여중 주차장 24시간 개방 안돼"...해운대구-여중 대립

부산뉴스 2019년 05월 06일 09:00

부산 해운대구가 관내 여중학교 앞 통학길에 인도 설치 추진하는 과정에서 해당 여중과 대립을 빚고 있다. 구가 새로 조성할 인도 구역에는 한계 주민들이 주장한 부속으로 불법주차를 하고 있는데, 이 차량들을 수용하기 위해 구가 해당 여중 운동장을 24시간 주차장으로 개방하라는 방침을 세웠기 때문이다.

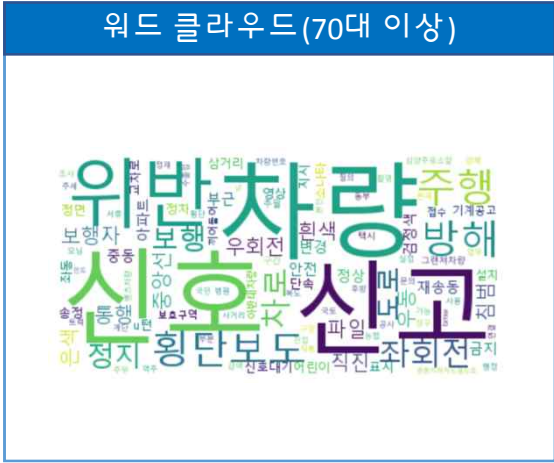


일상생활

심층분석 사례 : 부산시 해운대구

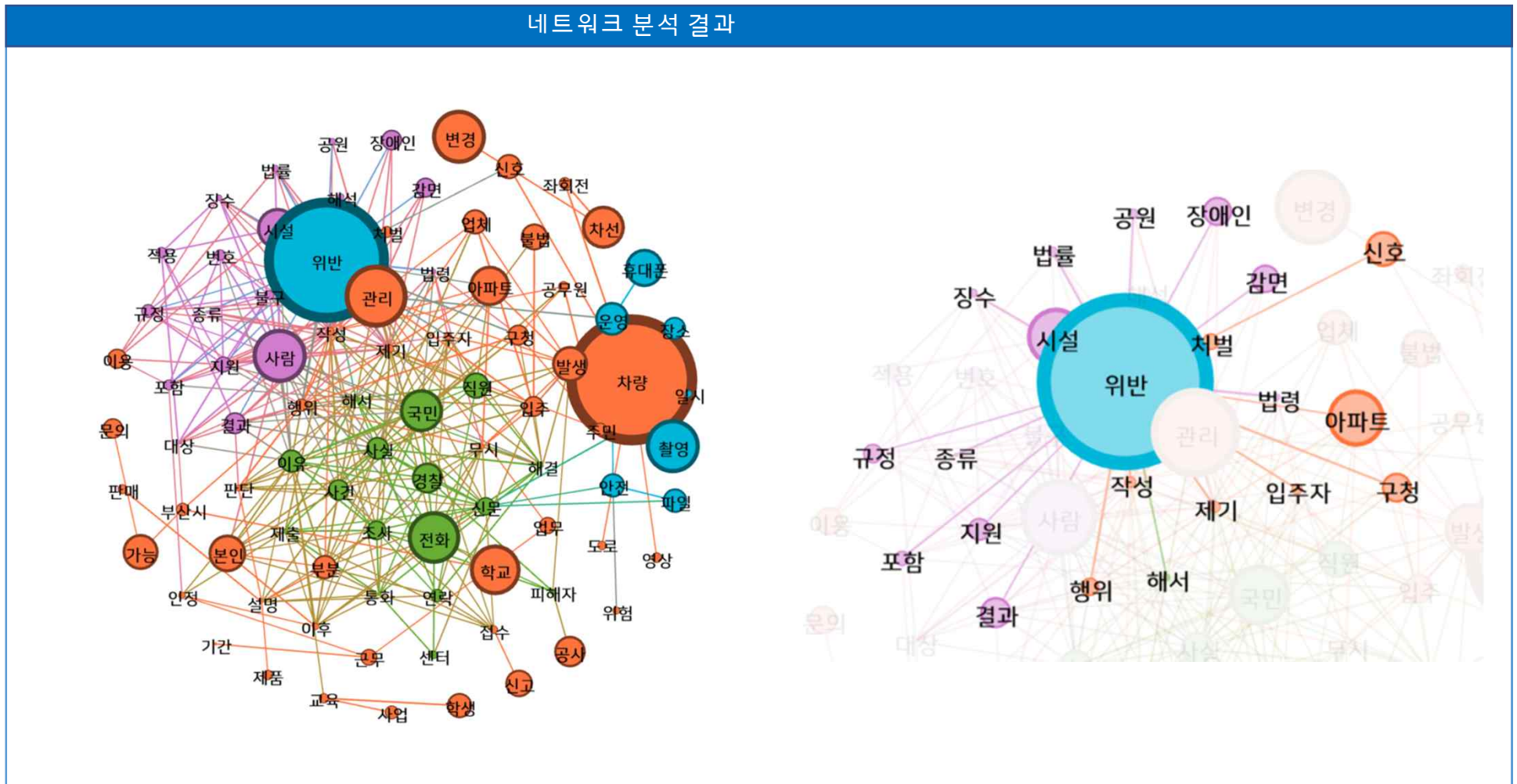
IV. 분석결과_해운대구

연령대를 기준으로 10대는 문화 20대는 학교, 30대는 교통과 학교, 40대는 생활 전반, 50대는 거주 관련 키워드의 출현 빈도가 높은 것이 확인되며 연령별로 확연히 다른 종류의 민원을 제기하고 있음.

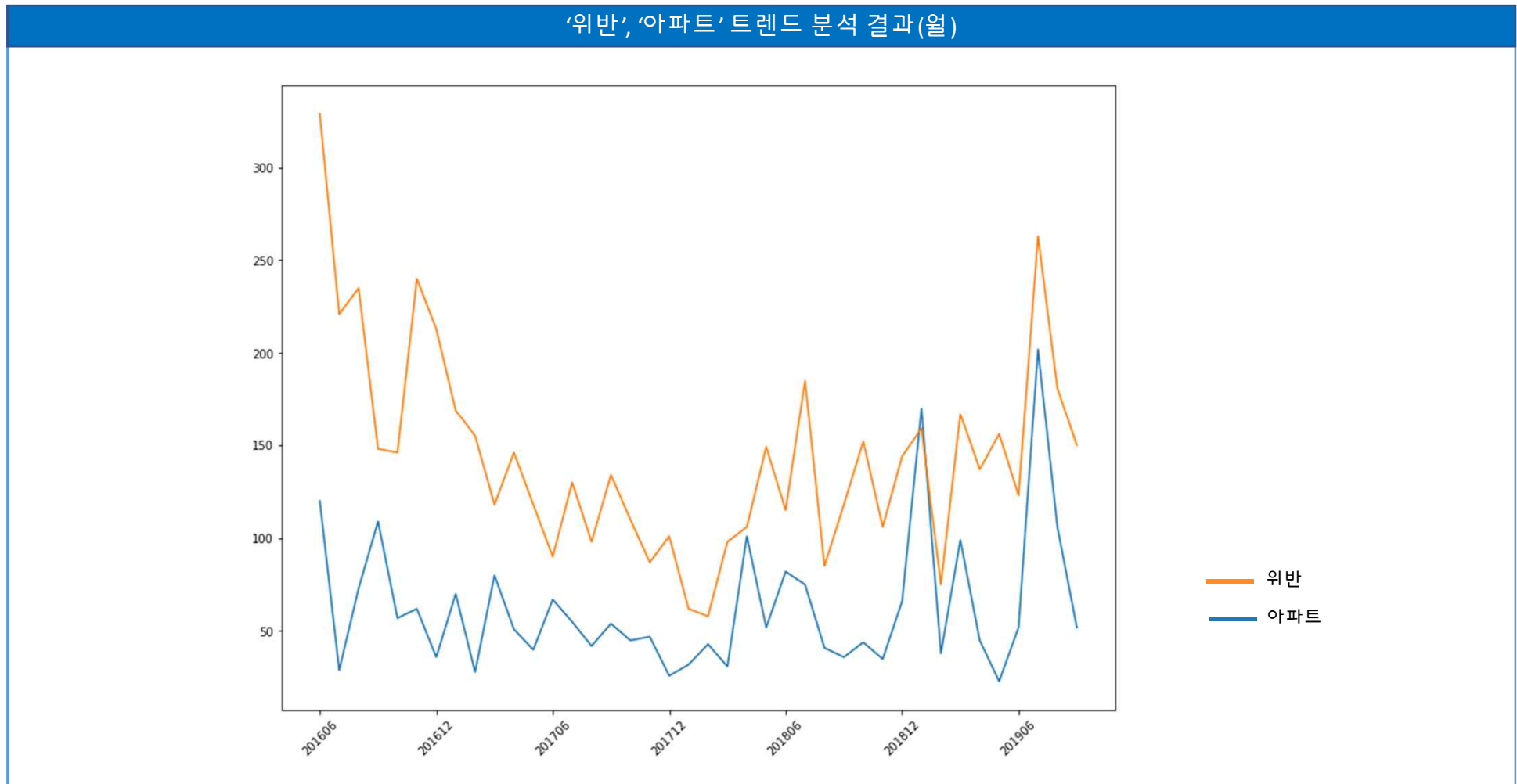


심층분석 사례 : 부산시 해운대구

민원데이터를 기반으로 네트워크 분석 결과 아파트 등 거주 관련 민원 간, 차량 등 교통 관련 민원간 관련성이 높게 나타남.



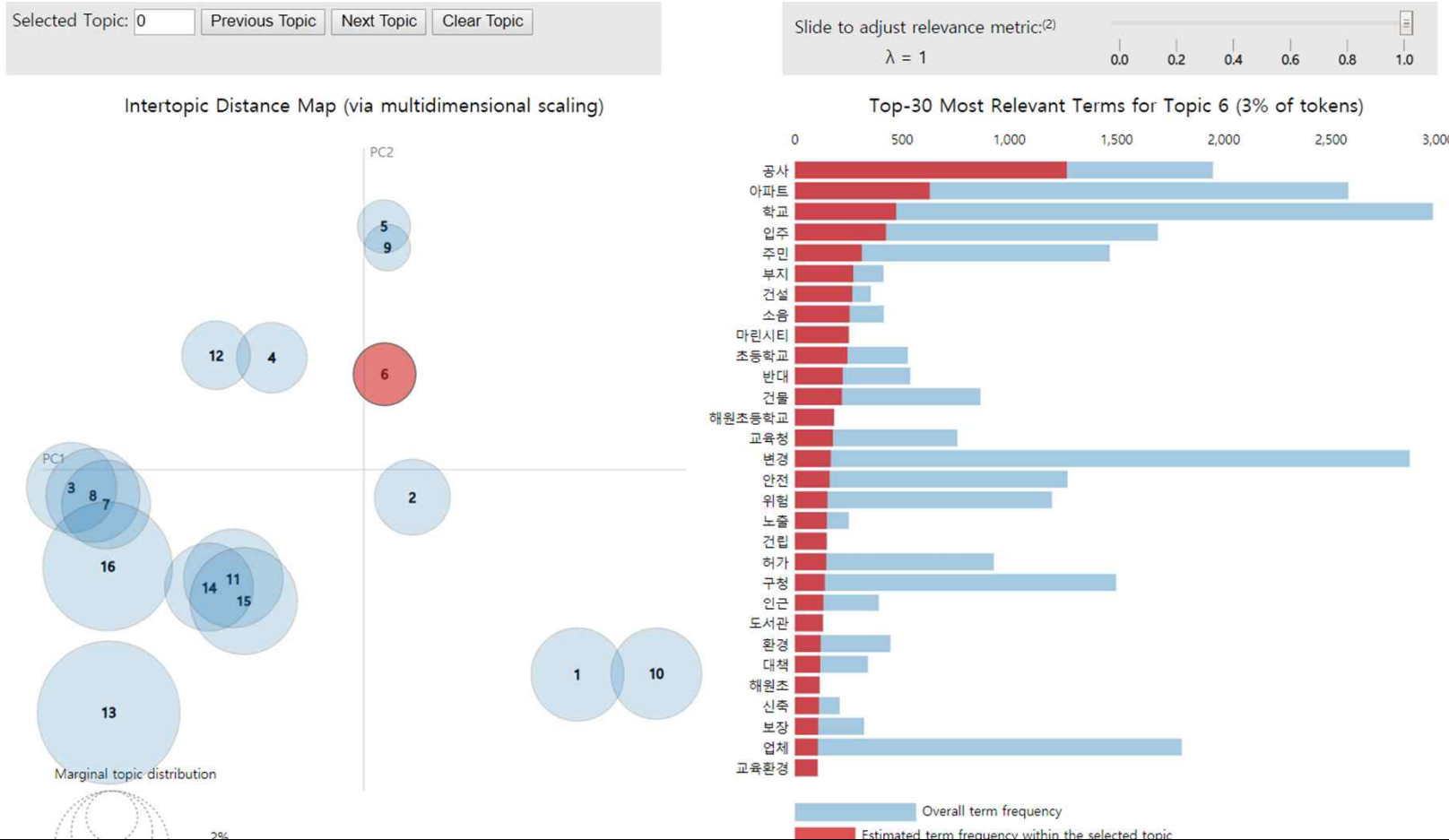
주요 키워드인 '위반' 과 '아파트' 라는 키워드가 출현한 빈도를 시간의 흐름에 따라 분석하면 아파트의 경우 계절별 순환의 형태가 나타나며 아파트 키워드의 출현은 점점 많아지고 있음을 알 수 있음.



심층분석 사례 : 부산시 해운대구

IV. 분석결과_해운대구

토픽 모델링 결과, 6번 토픽은 공사,아파트,학교,입주,주민,건설,소음,마린시티 등 아파트와 관련된 키워드가 많으며 아파트 관련 토픽임을 알 수 있으며 해운대구의 마린시티 인근 아파트와 관련된 민원이 다수 존재한다는 해석도 가능하며, 해당지역 후보자나 유권자의 경우 토픽에 대한 해석을 더욱 정확하게 할 수 있을 것임.



6번 토픽: 공사,아파트,학교,입주,주민,부지,건설,소음,마린시티,초등학교,반대,건물,해원초등학교,교육청,변경 등의 단어가 포함

아파트 민원 문제

End of Document

Thank You for your Time and Consideration