

---

# 경북 道 홈페이지 검색어 데이터 분석 결과 보고서

---

2024. 08

(주) 새올정보기술

# < 목 차 >

<b>1. 분석 개요 .....</b>	<b>1</b>
- 1. 분석배경 .....	1
- 2. 분석목적 .....	1
- 3. 분석방법 .....	1
- 4. 분석도구 .....	2
- 5. 분석범위 .....	2
- 6. 활용 데이터 .....	2
<b>2. 분석 내용 .....</b>	<b>3</b>
- 1. 2023년도 검색량 추이 .....	3
- 2. 2023년도 전체, 분기별 상위 30개 검색어 분석 .....	4
- 3. 2023년도 월별 검색어 상세 분석 .....	7
- 4. 2023년도 전체 상위 검색어의 연관검색어 분석 .....	9
<b>3. 분석 결과 및 활용방안 .....</b>	<b>11</b>
- 1. 종합 분석 결과 .....	11
- 2. 미흡 및 보완사항 .....	11
- 3. 분석 결과 활용방안 .....	12

## < 표 · 그림 목차 >

[표-01] 분석 데이터 리스트 .....	2
[표-02] 2023년도 전체 상위 30개 검색어 및 검색량 .....	5

\*\*\*

[그림-01] 2023년 道 홈페이지 연간 검색량 추이 .....	3
[그림-02] 2023년 道 홈페이지 연간 검색어수 추이 .....	3
[그림-03] 2023년 총 검색량 누적치 중 상위 30개 워드클라우드 및 순위 그래프 .....	4
[그림-04] 2023년 분기별 상위 30개 순위 그래프 .....	6
[그림-05] 2023년 1월, 2월, 3월 상위 30개 워드클라우드 .....	7
[그림-06] 2023년 4월, 5월, 6월 상위 30개 워드클라우드 .....	8
[그림-07] 2023년 7월, 8월, 9월 상위 30개 워드클라우드 .....	8
[그림-08] 2023년 10월, 11월, 12월 상위 30개 워드클라우드 .....	9
[그림-09] 연관검색어 시각화-1 .....	9
[그림-10] 연관검색어 시각화-2 .....	10

## 1. 분석배경

- 정보공개 서비스의 목적은 공공기관이 보유한 정보를 국민에게 제공하여 국민의 알 권리를 보장하고, 행정의 투명성을 높이기 위함임
  - 효과적 서비스 운영을 위해서는 국민이 원하는 정보가 무엇인지 파악하는 것이 필수적임
- 검색어 데이터를 통해 국민이 가장 많이 찾는 정보와 관심 있는 주제 파악하고, 국민의 실제 관심사를 반영한 정보를 제공하여 국민 요구에 선제 부응이 가능토록 함

## 2. 분석목적

- 도 홈페이지 검색어 및 검색 누적량을 체계적으로 분석 및 분석 결과 시각화하여 도정에 대한 도민들의 관심사 확인
  - 분석 결과를 토대로 도민의 요구를 반영한 정보공개 서비스 제공하여 도민의 만족도를 높이고, 도정 관련 정보를 원하는 도민의 편의성 향상을 도모함

## 3. 분석 방법

- 데이터의 구조적 특성을 알아내기 위한 통계적 기법인 탐색적 데이터 분석(Exploratory Data Analysis : EDA) 기법을 사용
- 데이터 시각화(Visualization) 분석에 따른 이슈 도출

## 4. 분석 도구

- Python(데이터프레임, 데이터 시각화 라이브러리 등), Tableau(데이터 시각화), Excel, Powerpoint 등

## 5. 분석범위

- 시간적 범위 : 2023년 1월 ~ 2023년 12월
- 공간적 범위 : 경북도청 홈페이지 내 검색어

## 6. 활용 데이터

데이터명	세부내용	형태	기간	출처
월별 검색어 및 검색 누적량	2023년도 道 홈페이지 검색어 및 검색 누적량	CSV	'23.01 ~ '23.12	경북도청 자치행정과

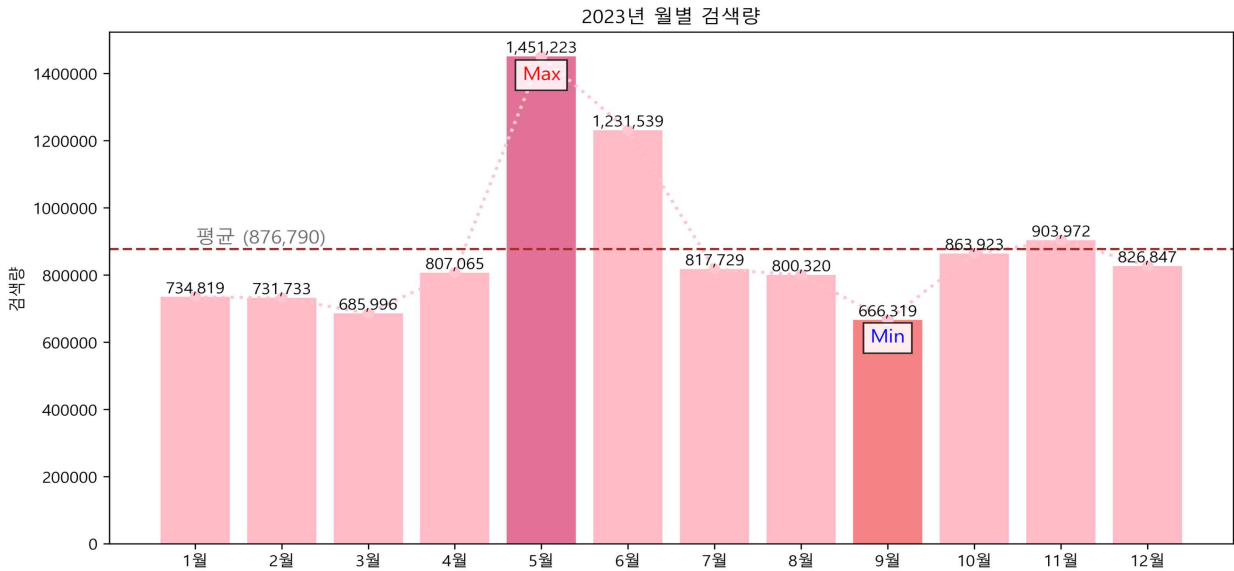
[표-01] 분석 데이터 리스트

- (진행 과정) 원본 데이터의 형식은 검색어, 검색량을 월별로 구성, 월별 총 209,534행으로 이루어짐
  - 1차 전처리 : 각 열의 이상치(빈 값, 의미를 알 수 없는 특수문자, 숫자, 자음 등) 삭제
  - 2차 전처리 : 최대 관심사 분석을 위해 검색 누적량이 전체 검색 누적량의 상위 1%에 해당하는 값을 기준으로 해당 값 이상 검색된 단어들만 분석 진행 (상위 1% 값\* : 399)

\* 1년('23.01.1 - '23.12.31) 동안 399번 이상 검색된 단어

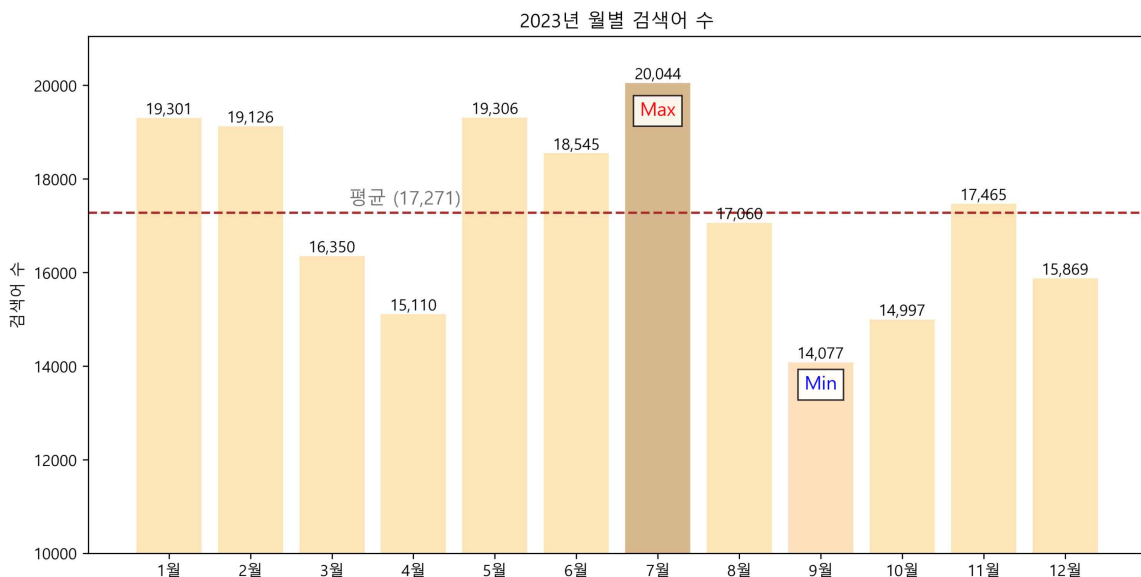
## 1. 2023년도 검색량 추이

- 전체 검색량 추이는 빈 값과 이상치를 제거한 데이터 기반으로 분석 진행



[그림-01] 2023년 道 홈페이지 연간 검색량 추이

- '23년 道 홈페이지 연간 검색량은 총 10,521,485건, 월평균 검색량은 876,790건이며,
- 최대 검색량은 5월 1,451,223건, 최소 검색량은 9월 666,319건임



[그림-02] 2023년 道 홈페이지 연간 검색어 수 추이

- '23년 道 홈페이지 연간 검색어를 살펴보면 총 207,250개(중복 검색어 포함), 평균 검색어 수는 17,271개며,
- 가장 다양한 검색어가 검색된 7월은 20,044개, 가장 적은 종류의 검색어가 검색된 9월은 14,077개임
- (검색량 및 검색어 추이) 전반적으로 검색량과 검색어 수는 유사한 방향으로 변동하고 5월에는 두 변수 모두 큰 폭으로 상승, 7월부터 9월까지의는 모두 감소

## 2. 2023년도 전체, 분기별 상위 검색어 분석

### ① 2023년도 전체 상위 30개 검색어 분석

- 중복 검색어를 제거하고, 각 검색어의 누적치를 기준으로 상위 30개 검색어 분석



[그림-03] 2023년 총 검색량 누적치 중 상위 30개 워드클라우드 및 순위 그래프

- 2023년 가장 많이 검색된 단어는 '대구경북통합신공항'(702,153회)이며, '자치경찰위원회'(682,991회), '경북도보'(238,209회), '토지이용계획'(232,497회), '팔공산'(230,057회) 순으로 검색됨

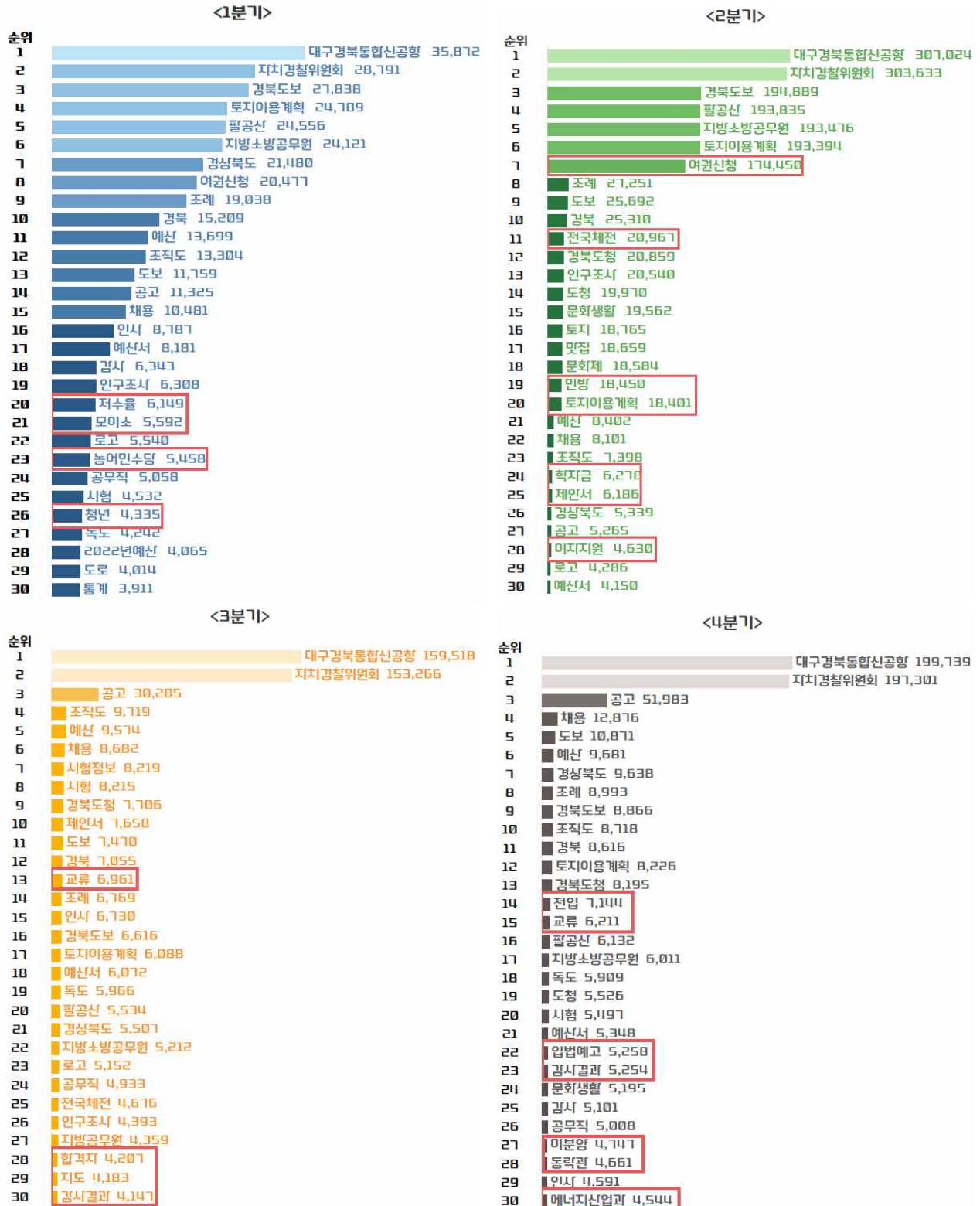
순위	검색어	검색량(23.01.01-23.12.31합산)
1	대구경북통합신공항	702,153
2	자치경찰위원회	682,991
3	경북도보	238,209
4	토지이용계획	232,497
5	팔공산	230,057
6	지방소방공무원	228,820
7	여권신청	201,619
8	공고	98,858
9	조례	62,051
10	경북	56,190
11	도보	55,792
12	경상북도	41,964
13	예산	41,356
14	채용	40,140
15	경북도청	39,758
16	조직도	39,139
17	인구조사	35,380
18	전국체전	31,135
19	도청	30,837
20	문화생활	27,575
21	토지	24,971
22	인사	23,906
23	예산서	23,751
24	맛집	21,869
25	시험	21,036
26	독도	20,060
27	문화제	19,469
28	로고	19,452
29	감사	18,942
30	공무직	18,789

[표-02] 2023년도 전체 상위 30개 검색어 및 검색량

- (지역 관련) '대구경북통합신공항, 팔공산, 경북, 독도' 등 대구 경북 지역과 관련한 검색어
- (공직 관련) '자치경찰위원회, 경북도보, 공고, 조례, 예산, 조직도' 등 지방정부의 업무 및 행정에 대한 검색어
- (구직자 관련) '채용, 인사, 공무직, 시험' 등 채용 및 인사와 관련 검색어
- (문화생활 관련) '전국체전, 문화생활, 문화제' 등 행사 및 문화 관련 검색어

## ② 2023년도 분기별 상위 30개 검색어 분석

- 중복 검색어를 제거하고, 검색어의 누적치를 기준으로 분기별 상위 30개 검색어 분석

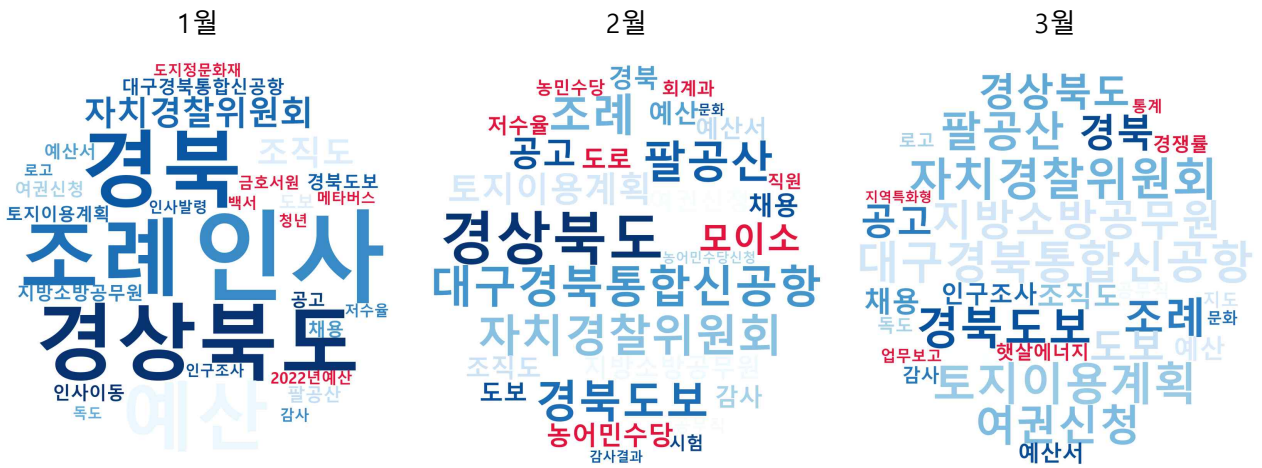


[그림-04] 2023년 분기별 상위 30개 순위 그래프

- 연간 상위 검색어와 분기별 상위 검색어는 전반적으로 유사하지만, 분기별 특징적 검색어가 확인됨
  - 1분기 : 저수율, 모이소, 농어민수당, 청년
  - 2분기 : 학자금, 제안서, 이자지원
  - 3분기 : 교류, 합격자, 지도, 감사결과
  - 4분기 : 전입, 입법예고, 미분양, 동락관, 에너지산업과
- 연간 누적 검색량 7위 검색어인 '여권신청' 의 전체 검색량의 약 84%가 2분기에 검색됨

### 3. 2023년도 월별 상위 30개 검색어 상세 분석

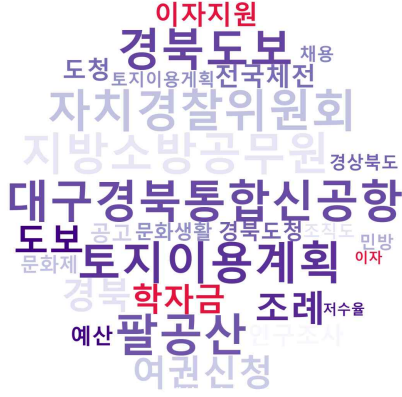
- 연간 및 분기별 상위 검색어와 월별 상위 검색어는 전반적으로 유사하지만, 월별 특징적 검색어가 확인됨



[그림-05] 2023년 1월, 2월, 3월 상위 30개 워드클라우드

- 1월 : 2022년예산, 도지정문화재, 금호서원, 청년, 백서, 메타버스
- 2월 : 모이소, 농어민수당, 도로, 저수율, 농민수당, 회계과, 직원, 농어민수당신청
- 3월 : 경쟁률, 업무보고, 업무보고, 통계, 지역특화형

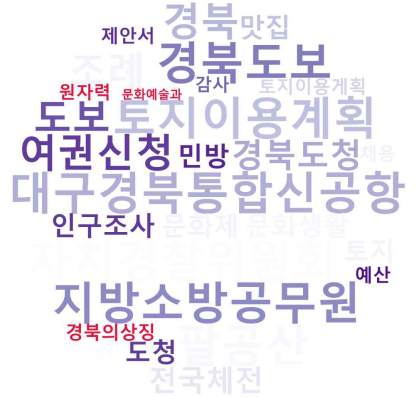
4월



5월



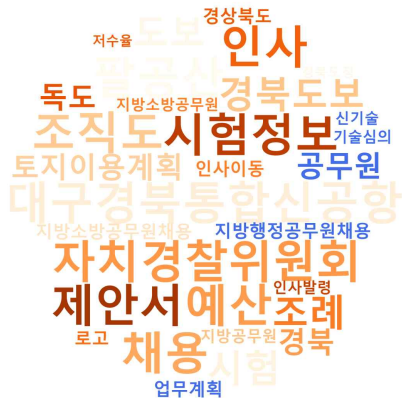
6월



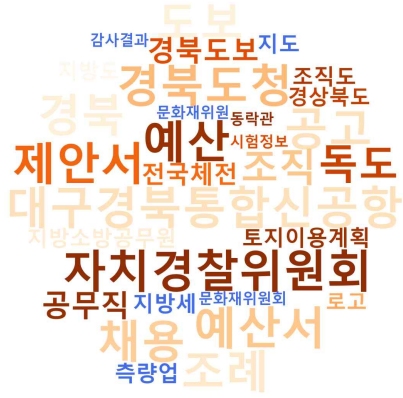
[그림-06] 2023년 4월, 5월, 6월 상위 30개 워드클라우드

- 4월 : 2022년예산, 도지정문화재, 금호서원, 청년, 백서, 메타버스
- 5월 : 모이소, 농어민수당, 도로, 저수율, 농민수당, 회계과, 직원, 농어민수당신청
- 6월 : 경쟁률, 업무보고, 업무보고, 통계, 지역특화형

7월



8월



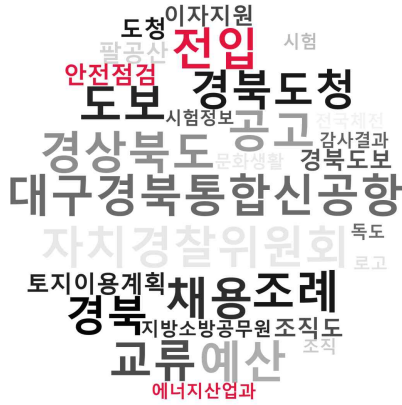
9월



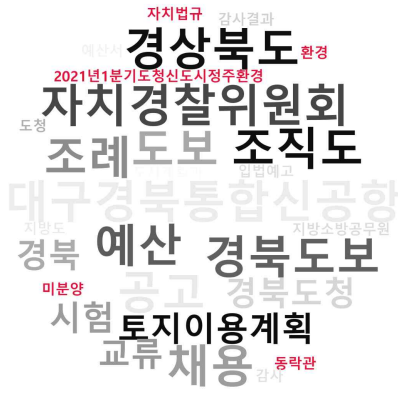
[그림-07] 2023년 7월, 8월, 9월 상위 30개 워드클라우드

- 7월 : 공무원, 업무계획, 지방행정공무원채용, 기술심의, 신기술
- 8월 : 지방세, 지도, 측량업, 문화재위원회, 문화재위원, 감사결과
- 9월 : 교류, 합격자, 지방하천

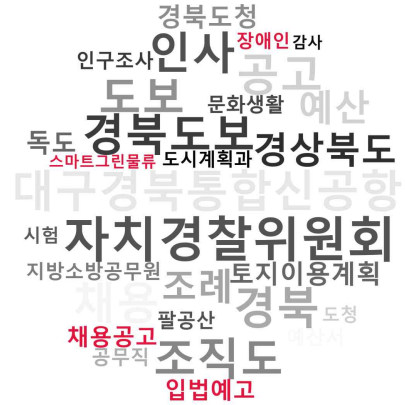
10월



11월



12월



[그림-08] 2023년 10월, 11월, 12월 상위 30개 워드클라우드

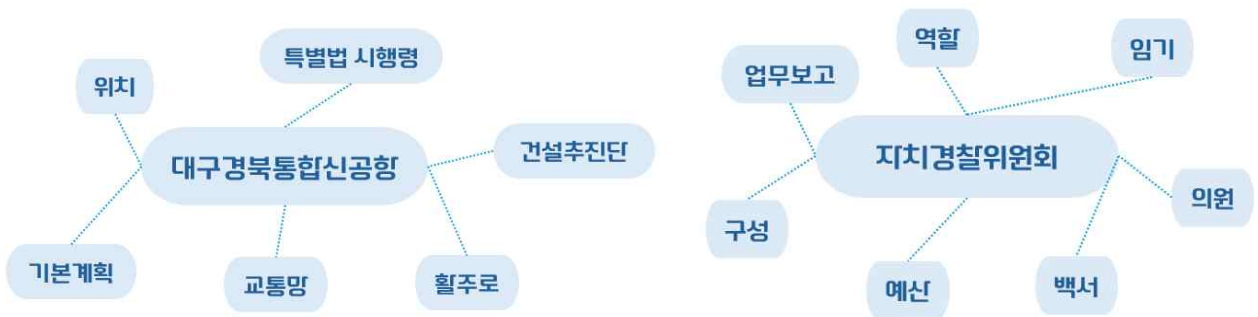
- 10월 : 전입, 안전점검, 에너지산업과
- 11월 : 2021년1분기도청신도시정주환경, 동락관, 미분양, 환경, 자치법규
- 12월 : 채용공고, 입법예고, 장애인, 스마트그린물류

#### 4. 2023년도 전체 상위 검색어의 연관검색어 분석

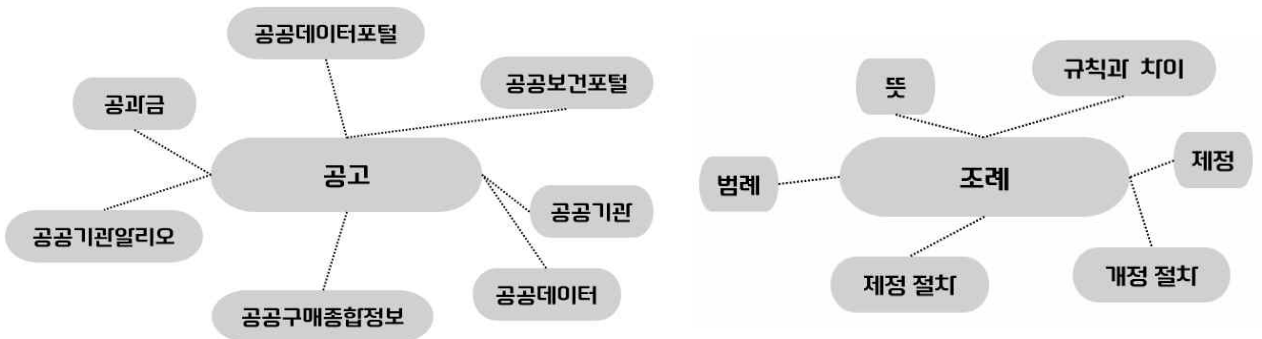
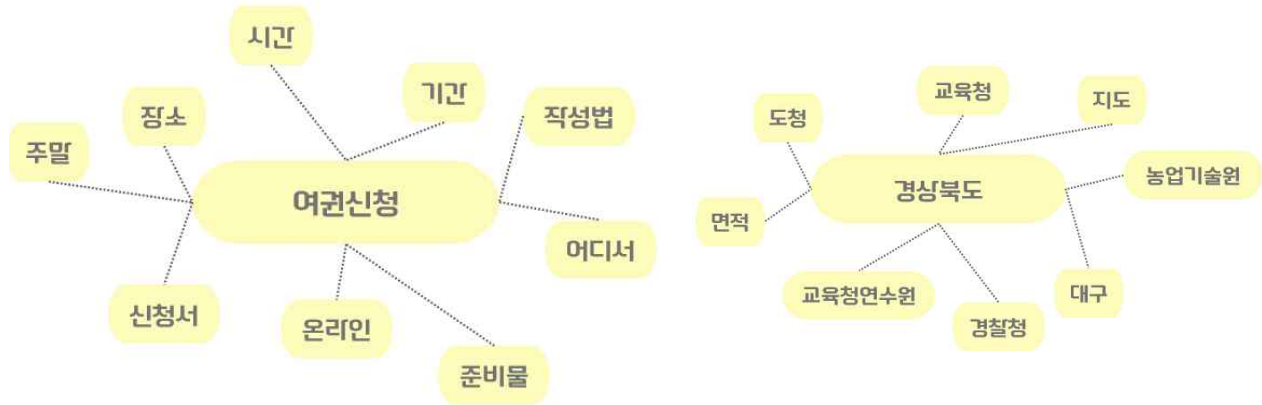
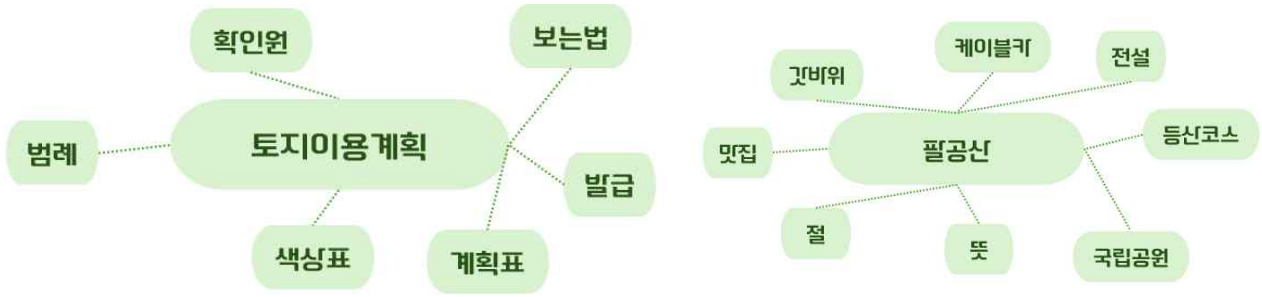
○ 단어별 빈도수 데이터의 한계를 보완하기 위한 연관검색어 분석

- 구글의 연관 검색 기능을 활용하여 2023년도 전체 누적치 상위 10개\* 검색어의 연관 검색어 분석

\* 상위 10개 검색어 중 연관검색어 분석이 없는 단어 제외(제외된 검색어 : 경북도보, 지방소방공무원)



[그림-09] 연관검색어 시각화-1



[그림-10] 연관검색어 시각화-2

- '대구경북통합신공항', '자치경찰위원회' 에 대한 연관 검색어의 경우 위치, 법적 기반, 조직에 관한 정보가 연관 검색어로 확인됨
- '토지이용계획', '여권신청'의 경우 발급·신청 과정, 준비물, 방법에 관한 정보가 연관 검색어로 확인됨
- '공고', '조례'의 경우 뜻, 절차, 관련기관 등 검색어에 대한 구체적·세부적인 단어가 연관 검색어로 확인됨

## 1. 종합 분석 결과

- (연간 상위 검색어) '대구경북통합신공항', '자치경찰위원회', '경북도보', '토지이용계획' 등 경북지역 발전, 도청 업무·역할 등이 도민의 최대 관심사로 나타남
- (분기·월별 상위 검색어) 분기·월별 정책 지원·시기와 맞물려 관련한 검색어가 나타남
  - 1분기 : 농어민수당 신청 기간과 맞물려, 관련 농업 지원 정책 및 정보 관심 증가
  - 2분기 : 경북 대학생 학자금 대출 이자 지원 신청 기간과 맞물려, 대학 학자금 대출 정보 관심 증가
  - 3·4분기 : 제안 사업, 경북 지방 공무원 채용 시기와 맞물려, 관련 정보 관심 증가

## 2. 미흡 및 보완사항

- 분석에 활용한 '단어별' 빈도수 데이터의 한계
  - 문장이나 구문을 단순히 단어 단위로 분리 시, 문맥·단어별 중요도 상실로 사용자의 구체적 검색 의도·의미 파악 어려움
  - 중복 및 불필요 단어의 누적 검색량이 많아져 누적 신뢰도가 떨어짐('경북, 경상북도'의 상위권 차지)
- 문장, 구문 분석이 가능한 원천 데이터 필요
  - 실제 검색된 문장, 구문 전체 데이터 분석 시 분석의 정확성과 유용성 증가

### 3. 분석 결과 활용방안

- 정보 요청 및 수요 예측을 통한 정보 공개 전략 수립
  - 상위 검색어 및 구글 연관 검색어 결과를 활용, 도민 주요 관심사와 필요를 반영한 정보 공개 가능
- 정보 공개 투명성 및 신뢰성 강화
  - 분석 결과를 활용, 공개 정보 사전 준비 가능 및 도민의 관심과 요구 충족으로 신뢰성 및 만족도 증가