



YOUTUBE ANALYSIS

어 떤 컨 텐 츠 가 좋 을 까 ?

키 워 드 분 석 을 통 한 엔 터 테 인 먼 트 유튜브 채 널 의 컨 텐 츠 선 호 도 분 석



한양대학교 전자컴퓨터통신공학과
로봇지능 및 제어연구실
석사과정 박지수

Contents

YOUTUBE CONTENTS ANALYSIS

01

주제 선정 및 개요

주제의 필요성과
문제를 정의합니다.



02

데이터 수집 및 분석

데이터 수집, 전처리
그리고 분석 과정을
소개합니다.



03

결론 및 한계

분석결과를 바탕으로
제안을 합니다.

1

주제 선정 및 개요

DALGONA Ent.

YOUTUBE CONTENTS ANALYSIS



'달고나' 종합 엔터테인먼트

셀럽들의 짧고 재미있는 컨셉 컨텐츠로 시청자들의 투표와 팬덤을 놓고 경쟁하는 별들의 리그



유튜브 채널 '달고나' 운영

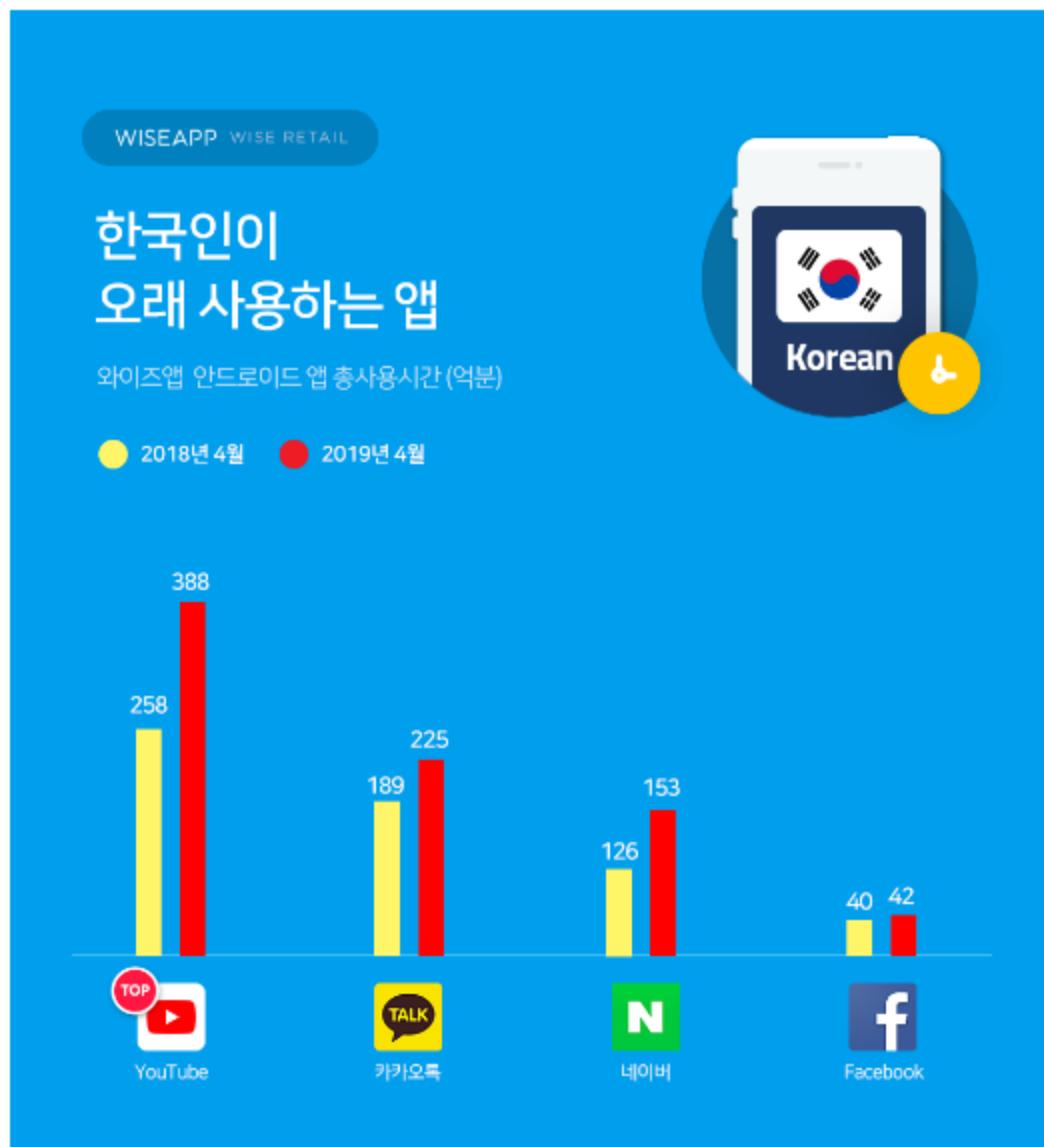
- 전 세계 최대 무료 동영상 공유 사이트



1

주제 선정 및 개요

YouTube 채널을 운영하는 이유



1 인지도 증가

기업 인지도와 이벤트 홍보 가능

2 유튜브 광고 수익

유튜브 게시물의 광고 수익

3 접근성 용이

사업 시작하기에 용이

1

주제 선정 및 개요

YouTube 채널을 운영하는 이유



1 인지도 증가

기업 인지도와 이벤트 홍보 가능

2 유튜브 광고 수익

유튜브 게시물의 광고 수익

3 접근성 용이

사업 시작하기에 용이



YOUTUBE 채널 활성화 필요

1

주제 선정 및 개요

YouTube 채널을 운영하는 이유



1 인지도 증가

기업 인지도와 이벤트 홍보 가능

2 유튜브 광고 수익

유튜브 게시물의 광고 수익

3 접근성 용이

사업 시작하기에 용이



YOUTUBE 채널 활성화 필요

관련 영상 소비자가
선호하는 컨텐츠 생산!!

1

주제 선정 및 개요

소비자가 **선호하는 콘텐츠 ??**

1

주제 선정 및 개요

소비자가 **선호하는 콘텐츠 ??**



데이터 분석!!

소비자가 선호하는 컨텐츠

선호도 지표 = 조회수, 좋아요

유튜브 사용 패턴을 생각할 때 관심가는 영상을 선택해 시청하므로 조회수가 높으면 해당 영상에 대한 관심도가 높다고 볼 수 있다.

컨텐츠 요약 = 제목

제목은 컨텐츠의 가장 핵심내용을 축약하므로 컨텐츠의 정보를 알 수 있다.

유튜브 사용 패턴상 대부분 제목과 썸네일을 보고 시청할 영상을 선택한다.

데이터 수집

크롤러를 이용한 데이터 수집

'달고나'채널을 사용하기에는 데이터가 충분하지 않아 엔터테인먼트 관련 영상을 올리는 '딩고' 채널을 타겟으로 하였다.

Selenium과 BeautifulSoup을 이용해 딩고 채널에 존재하는 모든 게시물에 접근해 제목, 조회수, 좋아요 수, 싫어요 수, 게시날짜를 수집하는 크롤러를 설계했다.

The screenshot shows the main page of the 'dingo' YouTube channel. At the top, there's a search bar and a navigation menu with links like EN LIVE, MOVIE, FACEBOOK LIVE, etc. Below the menu, a large red banner displays various video thumbnails. A specific video titled '세상에서 하나뿐인 인생샷스튜디오' (One-of-a-kind life shot studio) is highlighted. To the right of the banner, there's a sidebar with the channel's logo (a white bear), the name 'dingo / dingo', and a subscriber count of 94.9 million. Below the banner, there are sections for '업로드된 동영상' (Uploaded videos) and '최신 영상' (Latest video). The main content area shows a grid of video thumbnails with titles like '달고나' and '여행' (Travel).

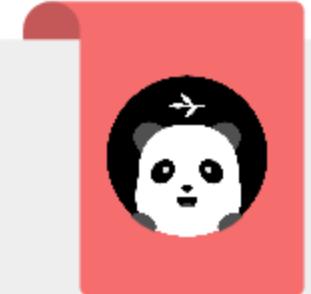


title	view	like	unlike	date
0 ENG SUB 우울한 나를 위로하는 방법에 대한 고민 여러분들은 어떤가? 조회수 30374회	3.3천	8	20시간 전 최초 공개	
1 이제 노래까지 노부부끼미 폭발하는 다비치의 나의 오랜연인에게를 조회수 28651회	1.4천	2	2019. 12. 6.	
2 비하인드/EVENT 세정SEJEONG X 딩고 펄영 현장 뮤직비디오 딩고: 조회수 11256회	913	0	2019. 12. 6.	
3 저세상 텐션으로 대원장파티 열고 간 아스트로 ASTRO - BLUE FLAN 조회수 106538회	1.4만	25	2019. 12. 6.	
4 ENG SUB 편찮아 너 잘 하고 있어 세정SEJEONG - 터널 TUNNEL 조회수 79773회	7.9천	18	2019. 12. 6.	
5 첫 눈 연금 들려 온 발리드 세순-정승환하심이월 이심오일의 고백My 조회수 54921회	3천	11	2019. 12. 6.	
6 ENG SUB 재생버튼 누르는 순간 계속 빙빙 티자는 역대급 벅본의 고 조회수 87865회	7.5천	15	2019. 12. 5.	
7 지치고 힘든 날 듣기 좋은 세정의 위로곡 BESTS 한술 하루의 꿀꽃 조회수 187481회	1.5만	40	2019. 12. 4.	
8 ENG SUB 꽉 막힌 하루 꿈에서 날 기다릴게 세정SEJEONG - 터널 조회수 533103회	3.2만	140	2019. 12. 2.	
9 ENG SUB 세정 SEJEONG '터널 Tunnel' Lyric Video Teaser리릭비디오 조회수 36567회	3천	5	2019. 12. 1.	
10 설현이 즐추다가 빵 터진 이유는에이오에이 - AOA 날 보러와요음향 조회수 352563회	8.8천	77	2019. 11. 30.	
11 ENG SUB 끌길 이후 3년만에 솔로곡 발표하는 김세정 세정SEJEONG 조회수 83982회	4.8천	19	2019. 11. 30.	
12 ENG SUB SOSO하게 백일장 할랬더니 아무말페스티벌 멜레버린 워너 조회수 53174회	5천	15	최초 공개: 2019. 11. 29.	
13 ENG SUB 데뷔 후 처음으로 손절 직전까지 치열하게 싸운 그들 빅: 조회수 94297회	7.1천	17	2019. 11. 29.	
14 AOA - 날 보러 와요 Come See Me Performance video 4K MOVE 조회수 330404회	2.3만	160	2019. 11. 28.	
15 ENG SUB 위로가 필요한 당신에게 전하는 세정의 따뜻한 한마디 자가조회수 47886회	3천	12	2019. 11. 28.	
16 노을 Noel - 늦은 밤 너의 집 앞 골목길에서 Late Night세로라이브링 조회수 174891회	2.5천	72	2019. 11. 26.	
17 유빈 Yubin - 무성영화 Feat윤미래로라이브링dingo Music 조회수 32207회	1.2천	20	2019. 11. 24.	
18 미술라이브 자이언티 - 5월의 밤 May 빛술 한 잔 땡기고 듣는 꺼내 조회수 150920회	1.5천	32	2019. 11. 23.	
19 ENG SUB 위너가 잠든 후에 여기 자면서도 미모 열일하는 꽃도령 조회수 100724회	7.7천	21	최초 공개: 2019. 11. 22.	
20 4K 아스트로 ASTRO - BLUE FLAME Performance video MOVE RE 조회수 325282회	3.2만	66	2019. 11. 21.	
21 ENG SUB 데뷔 후 역대급 모습으로 휴게소에 나타난 아이돌 빅데스 조회수 115051회	8.1천	11	2019. 11. 21.	
22 Crush - 그냥 Just크러쉬DF CAMDingo Music당고뮤직 조회수 75041회	3.5천	13	2019. 11. 20.	
23 에이틴 주하가 본인 중에 몇 때린 이유는골든차일드 Golden Child - 조회수 104459회	7.9천	13	2019. 11. 20.	
24 4K 우주소녀 WJSN - 이루리 As You Wish choreography MOVE 조회수 204584회	1.4만	82	2019. 11. 20.	
25 막내가 33살이라는 어느 걸그룹매력 터져버리는 브아걸 Zip터뷰 조회수 249258회	5천	37	2019. 11. 16.	
26 ENG SUB 힐링잔치 열어줬더니 웃음참기 헬리지풀 하는 아이들이 있 조회수 136500회	1만	33	최초 공개: 2019. 11. 15.	
27 ENG SUB 오랜만에 딩고 와서 명버들의 모든 비밀 탈탈 털고 간 아이 조회수 208371회	1.3만	31	2019. 11. 14.	

데이터 수집 및 분석
?

데이터 전처리

자연어를 포함한 데이터 전처리



Pandas | 데이터 분석 및 조작 Library

- 제목에 있는 이모티콘, 특수문자, 알파벳 등 불용어 제거
 - 다른 데이터 내의 불필요한 문자열 제거
 - 문자열 표기된 단위를 모두 숫자로 변환



KoNLPy | 한국어 자연어 처리 Library

- 형태소분석 패키지 Kkma(꼬꼬마)를 사용하여 명사(키워드) 추출
 - 한 제목에서 중복되는 키워드 제거

title	view	like	token
엔니드가 영화 보다 빅 빠진 이유는?※고힐임 주의※ BVNDIT 이연 습은 조진웅 이하늬	3775	55 ['고힐임', '영화', '하늬', '이연', '조진', '설랙', '성리', '미나', '습']	
대자연의 날을 구원해줄 구원자의 등장 feat생리 전 PMS 내 맘을 알아주는 웨스민어문하고	2487	72 ['구원자', '인아', '생리', '구원', '대자연', '등장']	
퇴사 후 제주 한 달 살이 브이로그* 소리 꼭 캐고 봐야함	62450	662 ['소리', '브이', '제주', '퇴사', '로그', '살이']	
캠핑 브이로그 일에 지쳐 술 속 캠핑을 떠나는 그녀의 해후	65091	182 ['하루', '브이', '캠핑']	
한국의 서원과 나는 디렉션 대비리듬 글작업금 부족세미나 개봉서시 미안디네이처와 함께 PMS	391503	1000 ['한국', '서원', '나는', '디렉션', '비교', '작업', '금지', '부족', '세미나', '개봉', '서시', '미안디', '네이처', '와', '함께', 'PMS']	
세상에 하나님뿐인 인생스타트디오에서 사진을 찍는다면 feat죽여내자대학교 What if I met my	428789	55 ['대학교', '여자', '세상', '스타트디오', '인생', '죽명', '사진']	
보사 과학전시회 당시의 그날은 그때였네요...feat생리 전 PMS 내 맘을 알아주는 웨스민어문	205306	204 ['웨이', '보산', '생리', '고백', '보스']	
서울 한강 공원 한복판에 고백부스가 설치됐다! feat생리 전 PMS 내 맘을 알아주는 웨스민어문	394583	1000 ['공원', '언어', '한복판', '서울', '설치', '생리', '고백', '보스']	
신조어 맞히셨더니 거의 단어 재정조한 권상우 우도환ㅋㅋㅋ 키워드 워리어 신의 한 수: 규	5897	193 ['수면', '키워드', '칭조', '단어', '상우', '신조어']	
특히 퀴 호강하는 이번주의 딩고 핫클립모음 딩고ZIP 10월 넷째 주	2719	50 ['모음', '이번', '클립', '호강', '주의']	
이번 주는 차들이나 나느끼의 니엘리 친멘일 모반 건나블리 그리고 라이크 까지 핫클립모음	2544	84 ['라이크', '놀이', '모음', '이번', '오반', '클립']	
위기 탈출 새로운 시작 씽남 EP 후반부 하이라이트 장기용 최우식의 브로맨스 웨드라마	3136	61 ['하이라이트', '후반부', '탈출', '시작', '위기', '드라마', '장기']	
생리 전 다중인격vs활潑하게 소리치는 상사_곁에는 feat생리 전 PMS 내 맘을 알아주는 웨스민어문	655835	580 ['결과', '살벌', '다중', '언어', '리딩', '인격', '생리', '상사', '다이']	
홀바비어브 티지는 기신 며이그오브 feat비스포크	4208	64 ['오버', '포크', '이브', '기신']	
3_꽃인연 연예인들의 회사 체험기 딩고 꽃 인연 모음 딩고 시리즈堂	2752	114 ['시리즈', '체험', '모음', '회사', '연예인', '인연']	
웹드라마 로봇이 아닙니다 속 남사친의 정색 정성을 로봇이 아닙니다 정성을 모아보기	3510	120 ['정색', '로봇', '사친', '정성', '드라마']	
강다니엘은 갑자기 고양이 소리를 듣 이어는 다넬이 한 번 웃이을 때마다 딩고 심장 꿩 하고 내	80180	9000 ['소리', '다니엘', '고양이', '심장']	
웨드 드라마 예능 다 하는 콤黢스 맛집 딩고의 핫클립모음 딩고ZIP 10월 넷째 주	1816	52 ['콤黢스', '유학', '모음', '드라마', '클립', '예능', '고의']	
꼭 한 쪽 가슴 잘라내야 돼한국유방건강재단필리브캡페인	3723	80 ['재단', '리본', '핑크', '가슴', '캡페인', '한국', '건강', '유방']	
이제는 스킨십이 너무 자인스러워진 두 남자 씽남 EP78 하이라이트 장기용 최우식의 브로맨	3707	79 ['남자', '하이라이트', '스킨십', '이제', '드라마', '자인', '장기']	
2_의리게임 카메라 밖에서 내 아이들의 의리는 딩고 시리즈堂	3917	208 ['시리즈', '아이', '게임', '의리', '카메라']	
로봇한테 이렇게 설레도 되는거야 로봇이 아닙니다 안도영 모아보기	3136	127 ['로봇']	
침연필로 칭 받고 시작하는 지난 주 핫클립모음 딩고ZIP 10월 첫째 주	2265	74 ['클립', '모음']	
이 콤黢스를 도대체 푸악→→→→→ 딩고가 준비한 별맛 모아주ZIP 딩고ZIP 별맛 편	3682	75 ['준비', '콤黢스']	
점점 길어지는 두 남자의 관계 씽남 EP56 하이라이트 장기용 최우식의 브로맨스 웨드라마	3001	71 ['남자', '관계', '하이라이트', '드라마', '장기']	
1_뒹고문고 눈과 귀의 즐거움을 한 번에 느낄 수 있는 강성 영상 딩고 시리즈堂	3122	70 ['시리즈', '강성', '고문', '영상', '즐거움']	
웹드에서도 시일한 도가풀과 도아 얼굴 보다가 시간 순삭 주의 로봇이 아닙니다 한여름 모아	12952	354 ['로봇', '한여름', '얼굴', '클라', '순삭', '주의', '시간']	
싣고 새 월드 그리고 칭연필의 칭끌듯 지난 주 딩고의 핫클립모음 딩고ZIP 9월 넷째 주	1999	41 ['클립', '모음', '고의']	
조회수 상승 소취 조회수 올려주ZIP 딩고ZIP 조회수 낮은 편	2918	71 ['회수', '소취', '상승']	
남사친한테 고백밖에이쁜 적 있는 시립 썰스데이4_레전드 썰	8134	179 ['고백', '시친', '시辱']	
친구들 앞에서 성형한 사실을 들키버렸다 필스데이3_레선드 썰	7475	179 ['성형', '사친', '친구']	
제4세로크리에이터 최종 스사자 뱃표 The last 세로연사	2825	57 ['최종', '크리', '스사자', '여자', '뱃표', '세로']	

"세상에 하나뿐인 인생샷스튜디오에서 사진을 찍는다면 feat숙명여자대학교"

[대학교, 여자, 세상, 스튜디오, 인생, 숙명, 사진]

명사 를 구분하여 키워드 추출

데이터 분석

문제점



KKMA의 한계

KoNLPy의 kkma 패키지는 사람이름과 같은 고유명사를 제대로 명사로 분류하지 못할 뿐만 아니라 비슷한 실제 단어를 잘못 추출함
Ex) '윤종신' -> '종신'



오류 키워드 제거

일상적 키워드

키워드가 등장하는 게시물의 조회수나 좋아요 수를 모두 합하여 선호도를 평가할 시 등장빈도가 높은 일반적인 키워드가 높게 평가된다.



키워드 선호도 평가
= 조회수[좋아요] 합 / 키워드 등장 횟수

단발성 키워드

굉장히 높은 조회수나 좋아요수를 얻은 게시물의 키워드들은 컨텐츠와의 연관성이 떨어지는 단발성 키워드더라도 높게 평가된다.



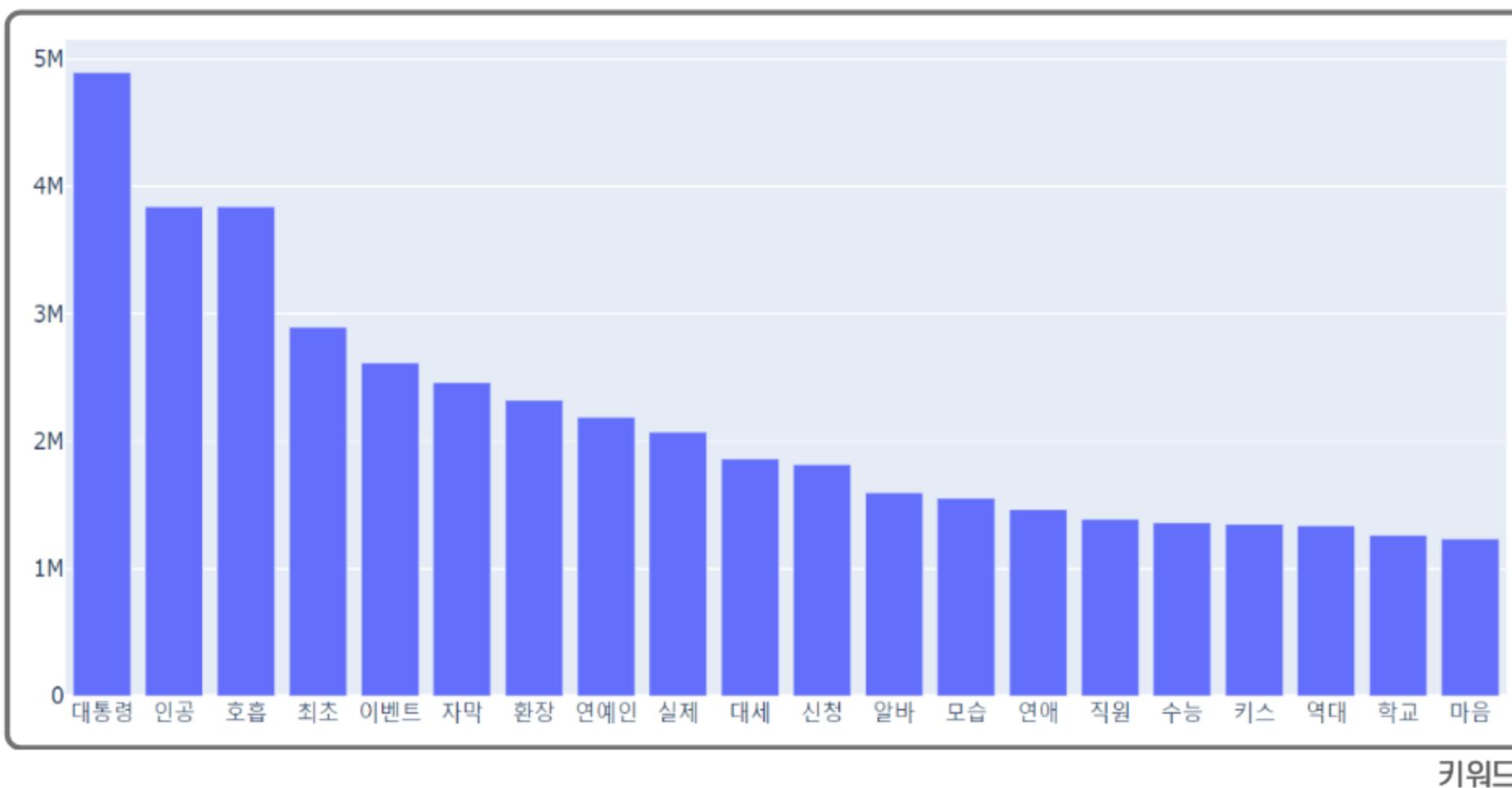
등장 횟수가 한 번 이상인 키워드에 대해서만 평가

데이터 분석 결과

조회수가 높은 상위 20개 키워드



조회수/등장 횟수



최초, 역대 등 강조성 키워드 조회수가 높음

대통령, 연예인, 대세 등 유명인 관련 키워드와
실제, 이벤트, 알바, 직원, 학교 등 일상적인 키워드가 동시에 높음

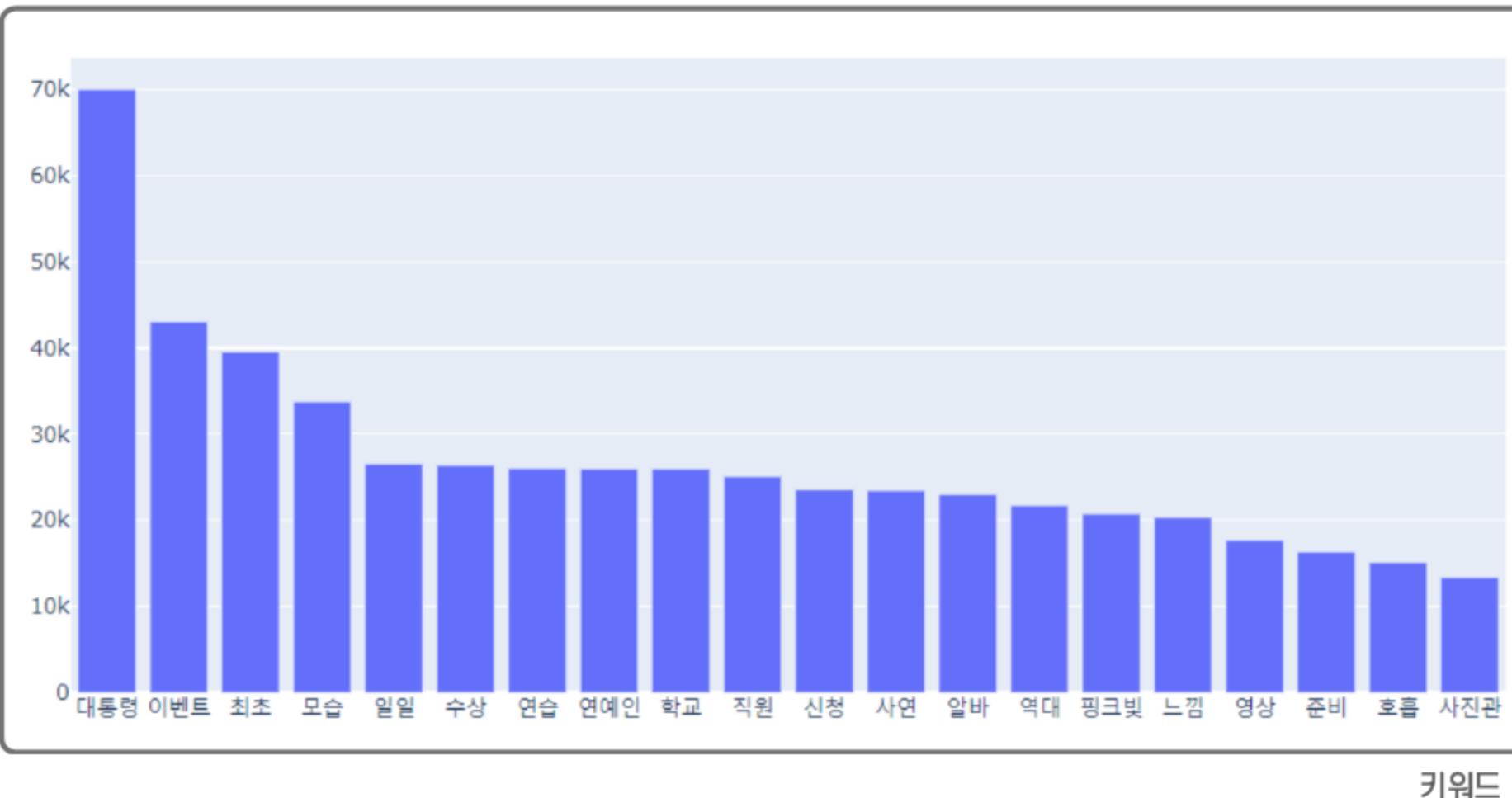
연애, 키스, 마음 등 애정에 대한 키워드도 조회수가 높음

데이터 분석 결과

좋아요 수가 높은 상위 20개 키워드



좋아요수/등장 횟수



최초, 역대 등 **강조성** 키워드 좋아요수가 높음

대통령, 연예인 등 **유명인** 관련 키워드와
이벤트, 알바, 직원, 학교 등 **일상적인** 키워드가 동시에 높음

핑크빛, 느낌 등 **애정**에 대한 키워드도 좋아요수가 높음

데이터 분석 결과

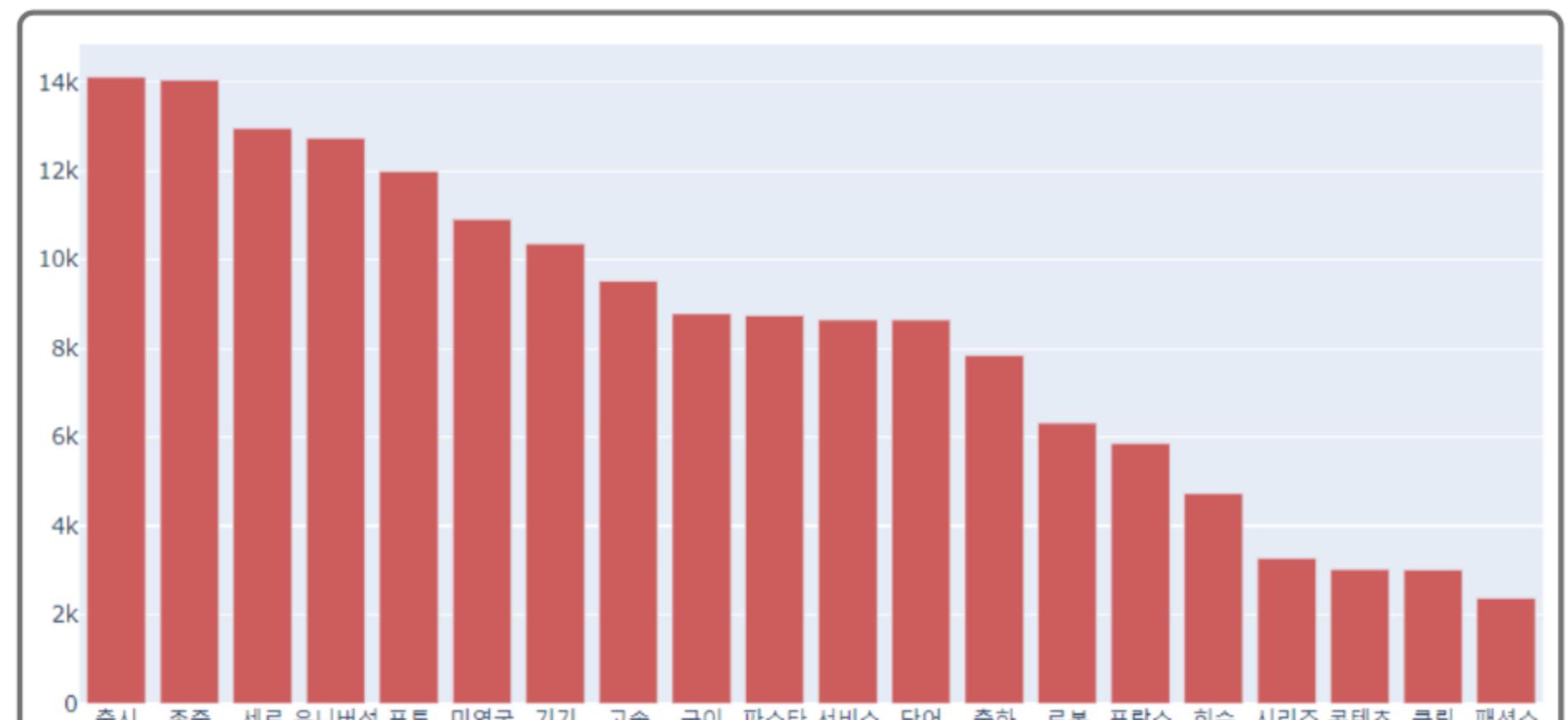
조회수가 낮은 하위 20개 키워드

데이터 수집 및 분석 ?

미역국, 구이, 파스타 등 음식관련 키워드 조회수가 낮음

클립이라는 키워드 조회수가 낮음

조회수/등장 횟수



키워드

데이터 분석 결과

좋아요 수가 낮은 하위 20개 키워드

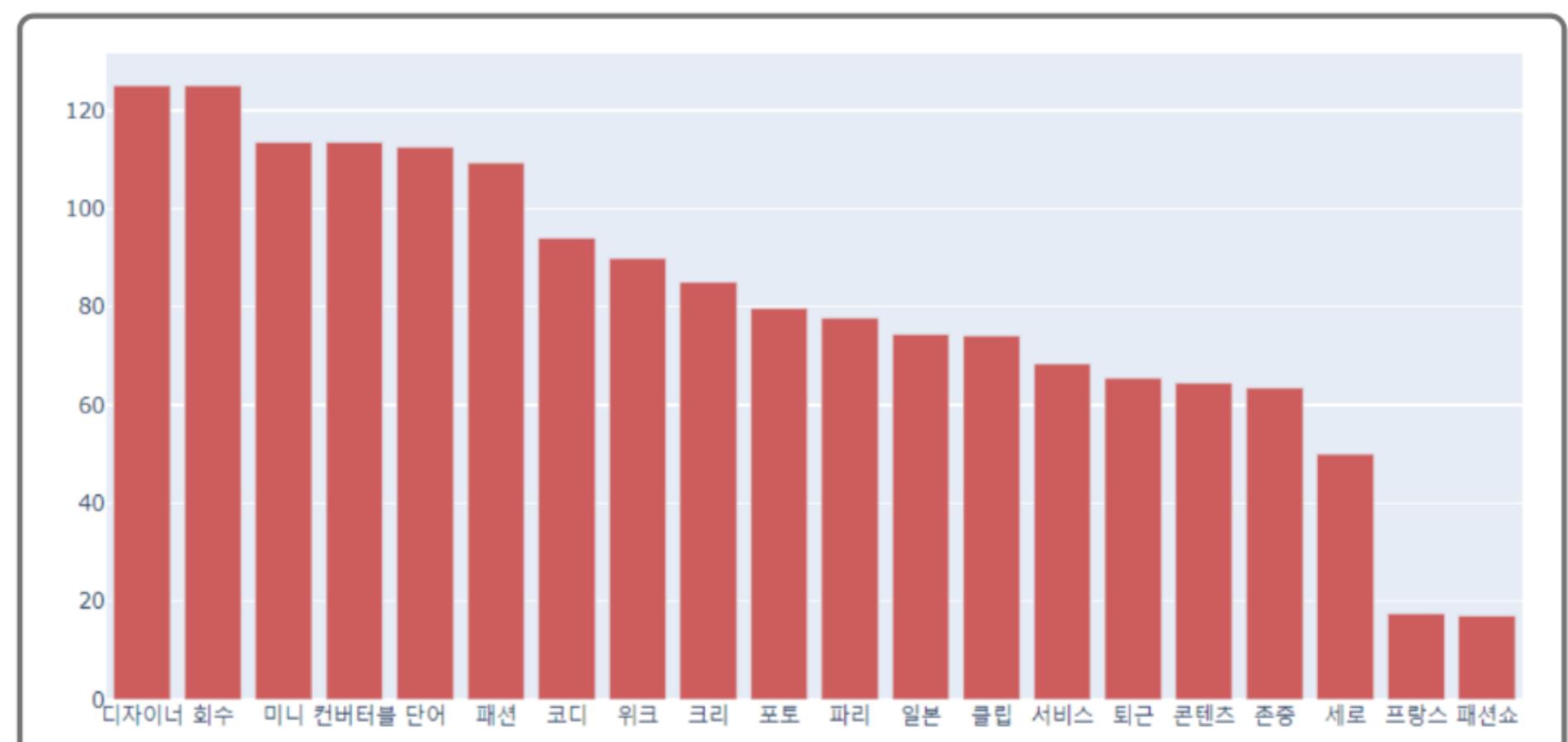


미역국, 구이, 파스타 등 음식관련 키워드 좋아요수가 낮음

클립이라는 키워드 좋아요수가 낮음

디자이너, 패션, 코디, 포토, 파리, 패션쇼 등
패션관련 키워드 조회수가 낮게 나타남

좋아요수/등장 횟수



키워드



GOOD

유명인이 일상적 상황에 등장하는 컨텐츠

유명인과 일상상황의 키워드가 동시에 높은 이유가 "내가 좋아하는 연예인이 알바를 도와준다면"과 같이 두 키워드가 연관될 때 매우 높은 관심을 보였다.

연애 관련 컨텐츠

연애와 관련된 키워드의 경우 관심도가 높았으며, 자극적일수록 조회수가 감성적일수록 좋아요수가 높게 나타났다.

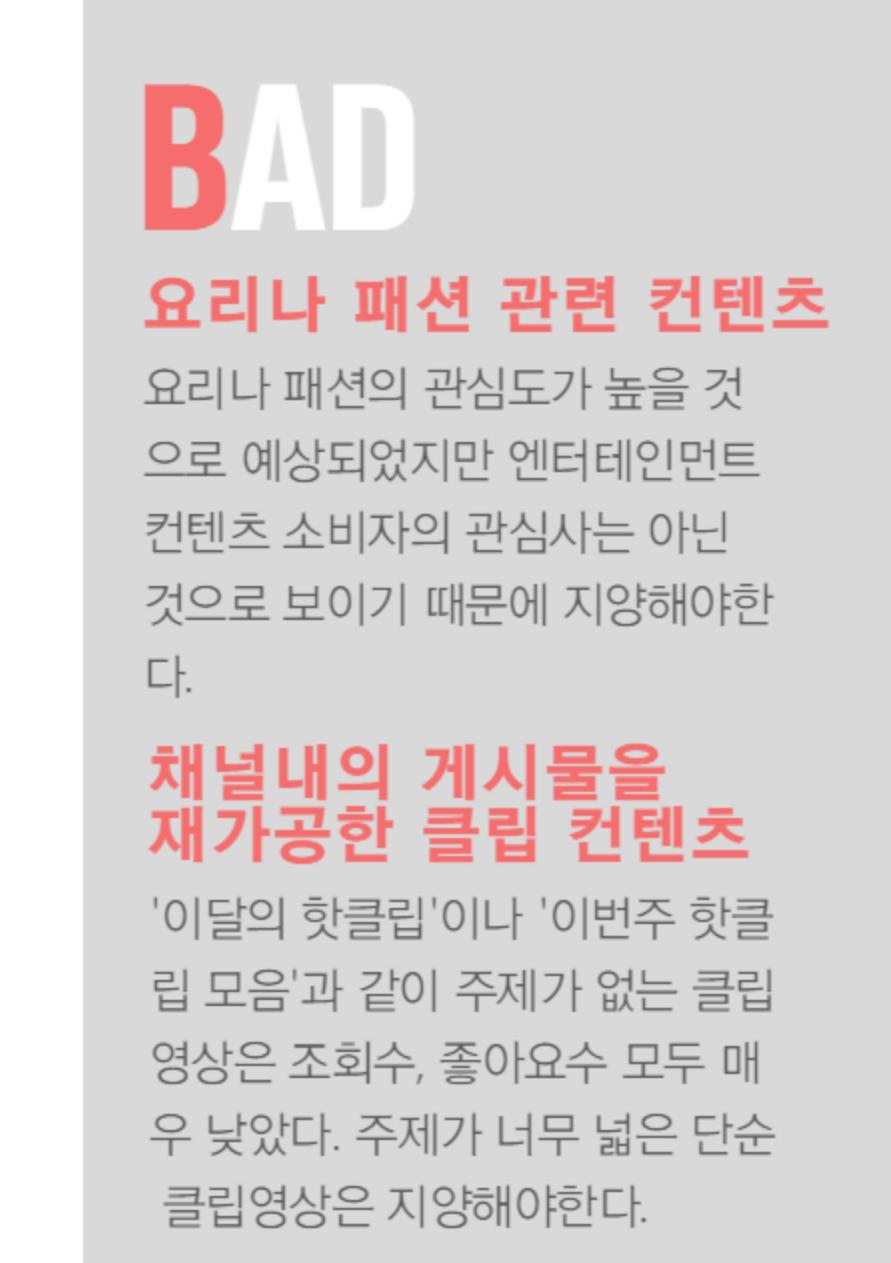
BAD

요리나 패션 관련 컨텐츠

요리나 패션의 관심도가 높을 것으로 예상되었지만 엔터테인먼트 컨텐츠 소비자의 관심사는 아닌 것으로 보이기 때문에 지양해야한다.

채널내의 게시물을 재가공한 클립 컨텐츠

'이달의 핫클립'이나 '이번주 핫클립 모음'과 같이 주제가 없는 클립 영상은 조회수, 좋아요수 모두 매우 낮았다. 주제가 너무 넓은 단순 클립영상은 지양해야한다.

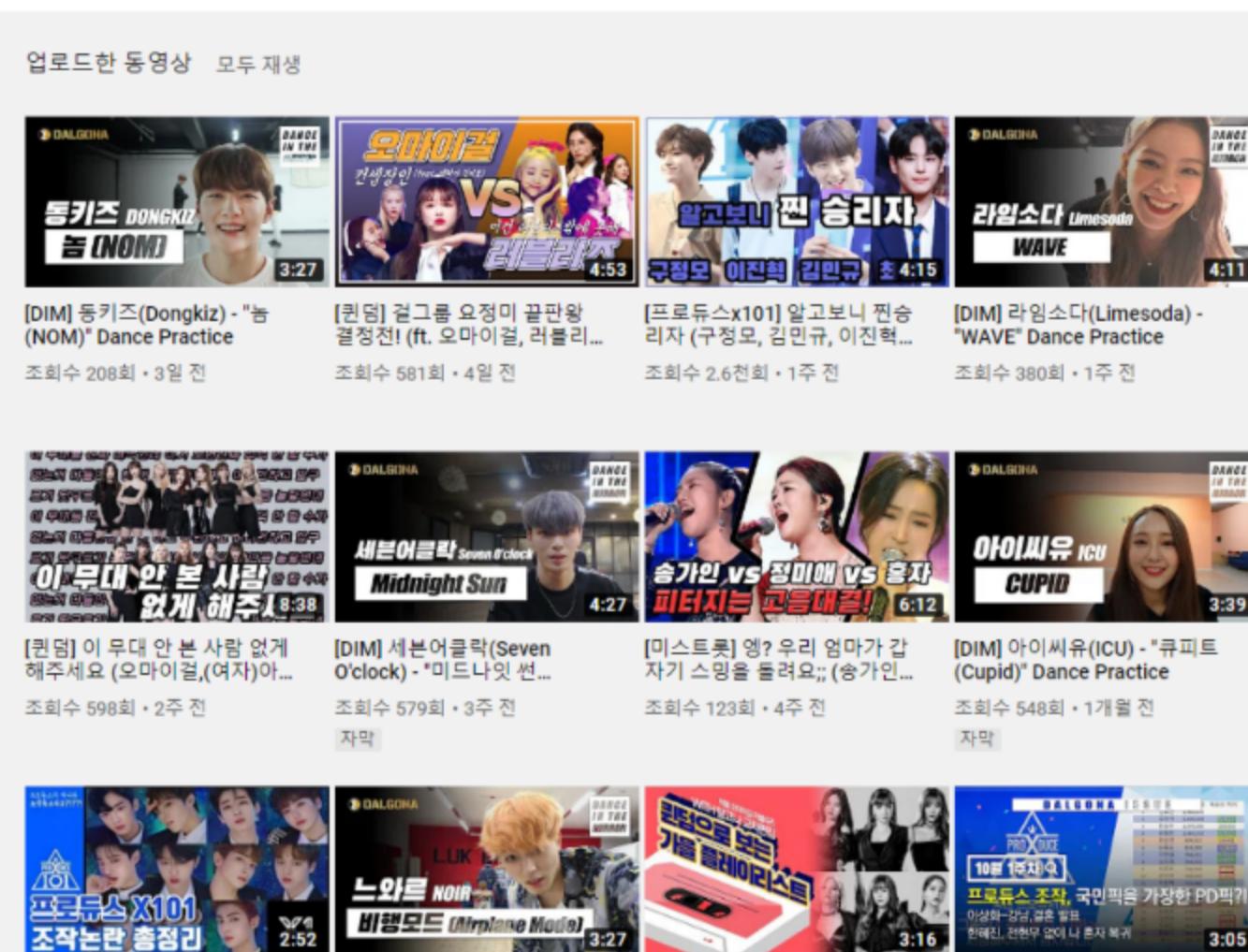


3

결론 및 한계

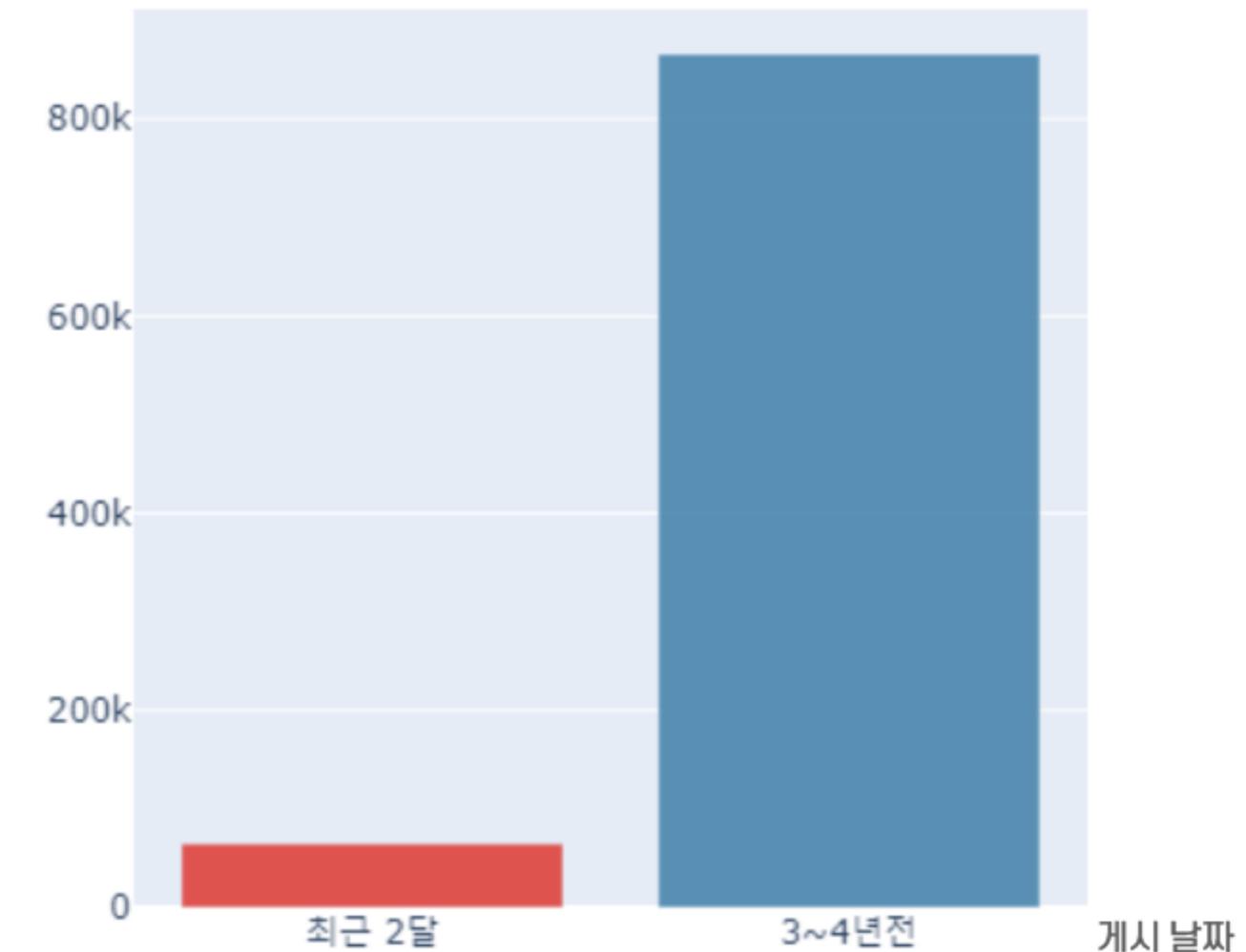
분석 한계점

더 육 정 확 한 분석 을 위 한 방 법



썸네일을 고려하지 않았다.

평균 조회수(40개)



기간을 고려하지 않았다.

로봇지능 및 제어연구실
인공지능팀
박 지 수
jijohn01@naver.com

THANK YOU.

YOUTUBE CONTENTS ANALYSIS