

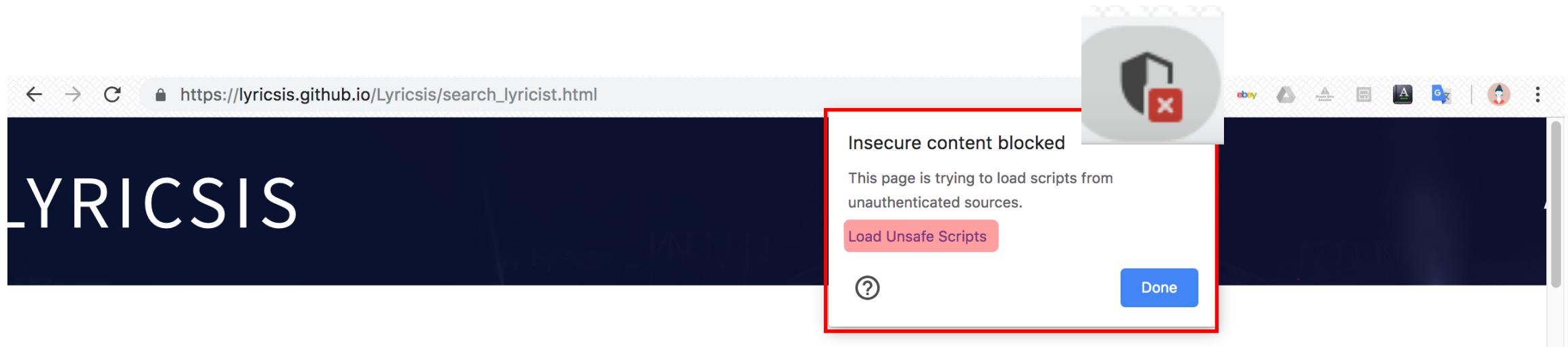
LYRICSIS

IDOL! Dejavu

아이돌 음악으로 분석하는 아이돌 시장의 젠더 불균형

•webpage

<https://lyricsis.github.io/Lyricsis>



- github repository

https://github.com/BLUENCE/M5_Idol_lyrics

●문제 정의



호르몬전쟁

방탄소년단

그녀 머리 바디 허리 다리
말 못하는 범위까지
관심 없단 말이
남자로선 많이 어이상실



뚜두뚜두

블랙핑크

착한 얼굴에 그렇지 못한 태도
가녀린 몸매 속
가려진 volume은 두 배로



표정관리

세븐틴

참 단정해 넌
내 눈에만 복장불량
발 무릎 발조차도 안보이게
치말 내려야 만족이 돼
목까지 단추 꼭꼭 잠궈
쇄골이 숨 못 쉬게



Likey

트와이스

BB크림 파파파
립스틱을 맘맘마
카메라에 담아볼까 예쁘게
이거 보면 웃어줘 그리고 꼭 눌러줘
저 밑에 앙증맞고 새빨간 Heart Heart

● 문제 정의

아이돌 음악 가사 속
성적 대상화, 외양 평가의 빈도

여성 > 남성 ?

● 편중된 외양 묘사가 왜 문제인가?

우리나라 **여성 아동 및 청소년**의 비만율은 14.1%로, **OECD 최하위** 수준이다.

하지만 2015년 청소년 건강행태 온라인 조사(교육부, 보건복지부, 질병관리본부)에 따르면 체질량 지수 85%미만인 마른 여학생들 중 34.7%가 자신이 뚱뚱하다고 생각하는 것으로 나타났다.
날씬한 여학생 10명중 3.5명이 자신의 신체를 과도하게 부정적으로 바라본다는 것이다.

많은 전문가들은 10대의 외모 집착과 **대중매체가 깊은 연관**이 있다고 말한다.

대중매체가 또래 연예인들의 **마른 몸매를 미의 상징으로 포장하기** 때문에
아름다움에 대한 잘못된 가치관이 형성된다는 분석이다.

출처

한국일보 <http://www.hankookilbo.com/News/Read/201712291765785898>
서울경제 <http://www.sedaily.com/NewsView/1L0BXPUJC7/GE02>

●왜 아이돌인가?



영향력

●왜 가사인가?



정확성

● 자료 수집

Crawling



Artists

333 그룹

Songs

10,000 곡

- 데이터 수집

1. 가수 정보

2. 곡 정보

●데이터 수집 - 가수 정보

●데이터 수집

- 가수 정보
- 이름
- 멤버 이름
- 성별
- 아티스트id
- 소속사(예전,현재)
- 데뷔 년도
- 데뷔 곡
- 팬 수(멜론 '팬 맷기')

https://www.melon.com/artist/detail.htm?artistId=228069

멜론 티켓 | 멜론 아지톡

이용권매 | VIP 혜택관 | 이벤트

Melon

바이브 'ABOUT ME' 앨범 발매 감상평 이벤트

급상승 7. 윤종신 5 [ZZZ]로 돌아온 Zion.T
'멋지게 인사하는 법' MV 촬영 현장

멜론차트 최신 장르 멜론DJ 멜론TV 스타포스트 매거진 뮤직어워드 멜론Hi-Fi 더보기 · 마이뮤직 피드

아티스트 채널



소녀시대 (GIRLS` GENERATION)
태연 (TAEYEON), 써니 (SUNNY), 유리 (YURI), 수영 (SOOYOUNG), 티파니 (TIFFANY),
서현 (SEOHYUN), 윤아 (YOONA), 효연 (HYOYEON)

데뷔 2007.08.05 | 다시 만난 세계 (Into The New World)
결성일 2007
활동유형 그룹
소속사 (주)SM엔터테인먼트
수상이력 2017 골든디스크 어워즈 | 음반본상 더보기

관리자 105,667명
HunnyJ, madlue, 장큰곰, 꼬오즈123

나와의 친밀도 더보기 | 이 아티스트로 스마트라디오 듣기

공유하기

상세정보 | 곡 | 앨범 | 비디오 | 포토/스토리 | 팬 | Hi-Fi

수상이력

●데이터 수집

	artist	artist_m	gender	idol_id	agency_old	agency_new	debut_y	debut_t	fan
0	서태지와 아이들	서태지, 양현석, 이주노	남	100161	NaN	요요기획	1992.0	난 알아요	4893
1	더 블루	손지창, 김민종	남	180144	NaN	SM	1992.0	너만을 느끼며	474
2	잼	NaN	혼	225786	NaN	DSP미디어	NaN	NaN	19
3	에스오에스	NaN	여	101645	NaN	삼포니레코드	NaN	NaN	39
4	솔리드	이준, 정재윤, 김조한	남	100037	NaN	비앤비	1993.0	이젠 나를 (Give Me A Chance)	1772

●데이터 수집

- 곡 정보
- 제목
- 아티스트
- 앨범 명
- 발매일
- 장르
- 타이틀곡 여부
- 좋아요 수
- 창작자(작곡, 작사, 편곡)
- 가사

NO	곡명	아티스트	앨범	좋아요	뮤비	다운
1	Girls Are Back	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	18,064		
2	All Night	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	40,410		
3	Holiday	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	68,346		
4	FAN	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	17,991		
5	Only One	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	16,107		
6	One Last Time	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	18,185		
7	Sweet Talk	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	15,284		
8	Love Is Bitter	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	16,323		
9	오랜 소원 (It's You)	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	15,547		
10	Light Up the Sky	소녀시대 (GIRLS` GEN…)	Holiday Night - The …	19,448		
11	그 여름 (0805)	소녀시대 (GIRLS` GEN…)	그 여름 (0805) - SM …	31,495		
12	그 여름 (0805) (Inst.)	소녀시대 (GIRLS` GEN…)	그 여름 (0805) - SM …	4,157		
13	GEE (DJ BEEJAY, NECK TOWL Remix)	소녀시대 (GIRLS` GEN…)	SBS MTV 매시업 K-P…	3,169		

●데이터 수집

- 곡 정보
- 제목
- 아티스트
- 앨범 명
- 발매일
- 장르
- 타이틀곡 여부
- 좋아요 수
- 창작자(작곡,작사,편곡)
- 가사

https://www.melon.com/song/detail.htm?songId=30565628

Girls Are Back
소녀시대 (GIRLS` GENERATION)

앨범 Holiday Night - The 6th Album
발매일 2017.08.04
장르 Dance
FLAC Flac 16bit

18,064 곡 다운 > FLAC 다운 > 선물하기 >

가사

Eye I know it's been a while
I don't think
we've got the girls out woo
Woo woo woo yeah
Uh you feel that
It feels good right
Dress and Heels

펼치기 ▼

작사 / 작곡

조윤경
작사
임정호
작사
디즈
작곡
Anne Judith Wik
작곡

●데이터 전처리 - 곡 정보

중복곡 제거

- 제목에 리믹스, 라이브, inst, 영일중, ver 인 행
- 띄어쓰기, 대소문자, 피처링에 의한 중복곡

+ 작사작곡 분석 용 전처리

creator가 없는 행 제거

방송 음원 제거

- 앨범에 나가수, 불명, 복면인 행

외국어 곡 제거

- 타이틀, 가사, 앨범에 히라가나/가타카나가 들어간 행
- 히라가나/가타카나를 제거한 후에도 일본어 가사가 한글로 포함되어 있는 경우
 - > contains로 확인한뒤 행제거 반복
- 가사가 모두 영어, 중국어인 경우
 - > 가사에 한글이 하나도 들어가지 않은 행 제거

기타

is_title이 nan인 행을 '수록곡'으로 변경

가사에 WrWrWn을 공백으로 변경

creator칼럼을 lyricist, composer, arranger로 나눔

●데이터 수집

		title	artist	album	release_date	song_genre	is_title	like	creator	lyrics	lyricist	composer	arranger
0	latata	(여자) 아이 들	I am	2018.05.02	Dance	타이틀 곡	85,374	소연 작사 곡 빅싼초 (Yummy Tone) 작 곡 빅싼...	기나긴 너와 이 밤을 너 와 이렇게 너와 기다린 너와 시간을 너와 이...	소연	소연, 빅싼초 (Yummy Tone)	빅싼초 (Yummy Tone), 소연	
1	달라(\$\$\$)	(여자) 아이 들	I am	2018.05.02	Dance	수록곡	4,229	Le'mon 작사 소연 작사 윤종성 작곡 Le'mon 작곡...	Knock knock 들어간다 Soyeon이 머리 위에 Spotlight...	Le'mon, 소연	윤종성, Le'mon	윤종성	
2	maze	(여자) 아이 들	I am	2018.05.02	Dance	수록곡	6,198	ZeroZine (MosPick) 작사 Ferdy (MosPick) 작사...	너의 깊은 그 눈빛이 날 데려가 저 멀리 누군가 맘에 들 때 더 신...	ZeroZine (MosPick), Ferdy (MosPick), 소연	ZeroZine (MosPick), Ferdy (MosPick)	ZeroZine (MosPick), Ferdy (MosPick)	
3	don'ttextme	(여자) 아이 들	I am	2018.05.02	Dance	수록곡	2,950	빅싼초 (Yummy Tone) 작사 박해일 (Yummy Tone) 작 사 ...	오늘따라 이상하리만큼 잠이 오지 않아 괜히 뒤척여 익숙한 듯 어떻 게 연락...	빅싼초 (Yummy Tone), 박해일 (Yummy Tone), 재 리포터, 소연	빅싼초 (Yummy Tone), 박해일 (Yummy Tone), 재 리포터	빅싼초 (Yummy Tone), 박해일 (Yummy Tone)	
4	알고싶어	(여자) 아이 들	I am	2018.05.02	Dance	수록곡	3,384	ARIN 작사 VINCENZO 작사 Fuxxy 작사 Any M...	What's in your house What's in your house ...	ARIN, VINCENZO, Fuxxy, Any Masingga, 소연	ARIN, VINCENZO, Fuxxy, Any Masingga	VINCENZO, Any Masingga	

●Index

1. LYRICS

2. INDUSTRY

LYRICS

● 가사 데이터 처리 - Tokenizing

```
from soynlp.tokenizer import RegexTokenizer
```

```
In [16]: def preprocessing(text):
    # 개행문자 제거
    text = text.strip('\t\n\r')
    pattern = re.compile(r'\s+')
    text = re.sub(pattern, ' ', text)
    # 특수문자 제거
    # 특수문자나 이모티콘 등은 때로는 의미를 갖기도 하지만 여기에서는 제거했습니다.
    # text = re.sub('[?.,;:\/\)*~`!^`-_+<>@\#$%&-=#\}*]', ' ', text)
    # 한글, 영문, 숫자만 남기고 모두 제거하도록 합니다.
    # text = re.sub('[^가-힣ㄱ-ㅎㅏ-ㅣa-zA-Z0-9]', ' ', text)
    # 한글, 영문만 남기고 모두 제거하도록 합니다.
    text = re.sub('[^가-힣ㄱ-ㅎㅏ-ㅣa-zA-Z]', ' ', text)
    return text
```

```
In [17]: def tokenize(df):
    sentences = df['lyrics'].apply(preprocessing)
    tokens = sentences.apply(tokenizer.tokenize)
    tokens = tokens.tolist()
    tokens_all = sum(tokens, [])
    return tokens_all
```

● EDA – word2vec

단어개수 : 142784

Window = 5

50차원

Mincount = 1

```
In [62]: all2vec.most_similar(positive=['여자'])

/Users/yunkim/anaconda3/lib/python3.6/site-packages/ipykernel
most_similar` (Method will be removed in 4.0.0, use self.wv.m
""Entry point for launching an IPython kernel.

Out[62]: [('남자', 0.6845423579216003),
('너랑 똑같은', 0.6748972535133362),
('뚜루뚜뚜뚜뚜뚜뚜', 0.6692284941673279),
('여자를', 0.6670475602149963),
('다를', 0.6649527549743652),
('위해주는', 0.6615994572639465),
('여자는', 0.6593049764633179),
('저기 끌리는', 0.6506280303001404),
('김당하긴', 0.6495503187179565),
('남자라고', 0.6484718322753906)]
```

```
In [63]: all2vec.most_similar(positive=['남자'])

/Users/yunkim/anaconda3/lib/python3.6/site-packages/ipykernel
most_similar` (Method will be removed in 4.0.0, use self.wv.m
""Entry point for launching an IPython kernel.

Out[63]: [('남자는', 0.7288351655006409),
('여자를', 0.7103898525238037),
('말고도', 0.6928809881210327),
('여잔', 0.6916475296020508),
('남자가', 0.6878726482391357),
('여자', 0.6845423579216003),
('남자들은', 0.6817284822463989),
('누구십니까', 0.6776272058486938),
('여잔_많아', 0.6703451871871948),
('나쁜_사람', 0.6700171232223511)]
```

```
In [64]: all2vec.most_similar(positive=['여자는'])

/Users/yunkim/anaconda3/lib/python3.6/site-packages/ipykernel_lau
most_similar` (Method will be removed in 4.0.0, use self.wv.most_
""Entry point for launching an IPython kernel.

Out[64]: [('여자를', 0.7716938257217407),
('밖에 모르는', 0.7122102975845337),
('여자들은', 0.7120750546455337),
('쪽을', 0.6951364278793335),
('여자지', 0.6948778629302979),
('무엇_하나도', 0.6918861865997314),
('부분이', 0.6909956932067871),
('너같이', 0.690981924533844),
('여잔', 0.6877371072769165),
('남자들은', 0.686608612537384)]
```

```
In [65]: all2vec.most_similar(positive=['남자는'])

/Users/yunkim/anaconda3/lib/python3.6/site-packages/ipykernel_lau
most_similar` (Method will be removed in 4.0.0, use self.wv.most_
""Entry point for launching an IPython kernel.
```

```
Out[65]: [('애가', 0.7781566381454468),
('남잔', 0.7589086294174194),
('사람은', 0.7495557069778442),
('남자', 0.7288351655006409),
('너같이', 0.7252179384231567),
('헤프니', 0.7149851322174072),
('아리속해', 0.7147759795188904),
('귀찮다고', 0.711052656173706),
('말고도', 0.706988513469696),
('짐승이라', 0.702533483505249)]
```

● EDA – 정규표현식으로 어미 보기

해줘요, 해줄게, 해줄래,
해주세요, 해줄까, 해줄거야.....

```
In [98]: end_of_word(female_t, r"해(줘|줄|주|요).")
```

```
Out[98]: [('해줘요', 44),  
          ('해줄게', 31),  
          ('해주고', 19),  
          ('해줄래', 18),  
          ('해주세요', 15),  
          ('해주는', 10),  
          ('해주고픈', 8),  
          ('해줄까', 8),  
          ('해줄거야', 8),  
          ('해주면', 6),  
          ('해주오', 6),  
          ('해줄래요', 6),  
          ('해줄께', 5),  
          ('해주길', 5),  
          ('해주던', 4),  
          ('해주기', 3),  
          ('해줘도', 3),  
          ('해주지', 3),  
          ('해주고만', 3),  
          ('해줘야', 2)]
```

```
In [97]: end_of_word(male_t, r"해(줘|줄|주|요).")
```

```
Out[97]: [('해줄게', 121),  
          ('해주고', 54),  
          ('해줄래', 31),  
          ('해줘요', 26),  
          ('해주는', 22),  
          ('해줘서', 17),  
          ('해주지', 12),  
          ('해주면', 12),  
          ('해줄께', 12),  
          ('해주길', 7),  
          ('해줄거야', 7),  
          ('해주겠니', 6),  
          ('해줘야만', 6),  
          ('해요말고', 6),  
          ('해줘야', 5),  
          ('해주던', 4),  
          ('해주오', 4),  
          ('해줄수', 4),  
          ('해줄꺼야', 4),  
          ('해주세요', 4)]
```

특정 단어의 어미를 보기
ex: 해-, [되돼]-

```
In [86]: end_of_word(female_t, r".{1,3}세요$")
```

```
Out[86]: [('주세요', 79),  
          ('해주세요', 15),  
          ('안아주세요', 10),  
          ('여보세요', 9),  
          ('받아주세요', 7),  
          ('보세요', 6),  
          ('오세요', 6),  
          ('놔두세요', 6),  
          ('안녕하세요', 6),  
          ('하세요', 5),  
          ('말해주세요', 5),  
          ('찾아오세요', 5),  
          ('생각하세요', 4),  
          ('지켜내세요', 4),  
          ('행복하세요', 4),  
          ('느껴보세요', 3),  
          ('알려주세요', 3),  
          ('비춰주세요', 3),  
          ('마세요', 3),  
          ('비켜나세요', 2)]
```

```
In [84]: end_of_word(male_t, r".{1,4}세요$")
```

```
Out[84]: [('여보세요', 63),  
          ('안녕하세요', 46),  
          ('마세요', 23),  
          ('주세요', 22),  
          ('누구세요', 13),  
          ('가세요', 13),  
          ('가지마세요', 7),  
          ('하세요', 7),  
          ('행복하세요', 7),  
          ('받아주세요', 5),  
          ('말해보세요', 5),  
          ('일어나세요', 4),  
          ('이겨내세요', 4),  
          ('느껴보세요', 4),  
          ('보세요', 4),  
          ('해주세요', 4),  
          ('어디세요', 4),  
          ('보여주세요', 3),  
          ('아껴주세요', 3),  
          ('안아주세요', 3)]
```

특정 어미가 쓰이는 단어를 보기
ex: -세요, -줘요, -요, -까, -래, -자

● EDA – 단어의 맥락 보기 (Keyword in Context)

```
In [53]: unique(before_word(female_t, '돼줄게'))
```

```
Out[53]: [['막아주고', '빛이', '돼줄게'],  
['너의', '여자가', '돼줄게'],  
['밝은', '빛이', '돼줄게'],  
['네', '편이', '돼줄게'],  
['너만의', '천사가', '돼줄게'],  
['걸', '빛이', '돼줄게'],  
['골목에', '등불이', '돼줄게'],  
['바다에', '등대가', '돼줄게'],  
['양큼한', '고양이', '돼줄게'],  
['너만의', '빛이', '돼줄게'],  
['될게', '그늘이', '돼줄게'],  
['반짝이는', '별이', '돼줄게']]
```

```
In [54]: unique(before_word(male_t, '돼줄게'))
```

```
Out[54]: [['사랑의', '포로가', '돼줄게'],  
['말해', '다', '돼줄게'],  
['너만의', '수퍼맨이', '돼줄게'],  
['bae', '복이', '돼줄게'],  
['이제는', '위로가', '돼줄게'],  
['style', '내가', '돼줄게'],  
['그런', '사람', '돼줄게'],  
['하루', '쉼표가', '돼줄게'],  
['아파해줄게', '날개가', '돼줄게'],  
['꽃', '길이', '돼줄게'],  
['해줄게', '힘이', '돼줄게'],  
['내가', '힘이', '돼줄게'],  
['좋은', '사람이', '돼줄게'],  
['마지막', '빈칸이', '돼줄게'],  
['되어', '빛이', '돼줄게'],  
['너의', '침대가', '돼줄게'],  
['너의', '불씨가', '돼줄게'],  
['내가', '빛이', '돼줄게'],  
['너의', '기적이', '돼줄게'],  
['너의', '편이', '돼줄게'],  
['아프다면', '그늘이', '돼줄게'],  
['괜찮은', '너의남자', '돼줄게'],  
['위해', '집이', '돼줄게'],  
['너만의', '별이', '돼줄게'],  
['포근한', '침대가', '돼줄게'],  
['밤에', '달이', '돼줄게'],  
['누구보다', '힘이', '돼줄게'],  
['탑', '스타', '돼줄게']]
```

너만의 천사가 **돼줄게**

너의 여자가 **돼줄게**

골목에 등불이 **돼줄게**

바다의 등대가 **돼줄게**

밝은 빛이 **돼줄게**

양큼한 고양이 **돼줄게**

●Keyword in Context

1

2

3

4

5

행복했어 정말 행복했어 세상 누구보다 예쁜 사랑을 했어 처음 만난 그 순간...

1

2

3

4

5

… 속 주인공처럼 맨 얼굴이 더 예쁜 그댄 나랑 결혼해 baby …

● EDA – Konlpy로 가사 속 품사 비교

konlpy Okt로
품사판별하여
여자/남자 아이돌 가사의
동사/형용사/부사 등 비교

동사 - 여성	동사 - 남성	형용사 - 여성	형용사 - 남성	부사 - 여성	부사 - 남성
다려지다	돕다	부다	아리깝다	같이	씩
봄타다	치키다	양큼하다	푸르다	어쩔	첨
잘라먹다	들이키다	꿀꿀하다	어떠하다	빙그르	걍
넘겨보다	겁니다	소란하다	타락하다	조아	휙휙
잡혀가다	재우다	억지스럽다	처참하다	찰랑찰랑	왜냐하면
엉클어지다	더디다	맥빠지다	거대하다	쩝쩝	우르릉
써리다	쑤다	말랑하다	머나멀다	쩝	휘청휘청
허우적이다	돋아나다	솔직담백하다	끼부리다	보들보들	당당
간질거리다	병들다	달짝지근하다	드높다	씰룩씰룩	찌그덕
남아돌다	지껄이다	화하다	비열하다	요렇게	나긋나긋
깔보다	알리다	맛깔나다	불과하다	우당탕	꽉꽉
나부끼다	뵈다	급격하다	강인하다	그렁그렁	왁
일구다	더러워지다	알록달록하다	정직하다	깜박깜박	건들건들
리듬타다	거슬리다	싱싱하다	수수하다	살짝살짝	서로서로
바꿔치다	굳히다	얼얼하다	안전하다	하늘하늘	저벅
갈아엎다	굽히다	보드랍다	삭막하다	콩당콩당	오래간만
출싹대다	꼴리다	느닷없다	팍팍하다	말랑말랑	방방
아뢰다	바라다보다	날렵하다	불공평하다	나풀나풀	해롱해롱
허둥거리다	틀어놓다	부탁드리다	고달프다	해롱	팍팍
얼쩡거리다	바로잡다	힘나다	현명하다	쪼로록	이왕

●가설 알아보기

[가설] 가사 속 성적 대상화, 외양 평가의 빈도는 여성 > 남성일 것이다.

● 본격적인 기시감 파헤치기

우리가 외양 평가를 할 때
가장 보편적으로 사용하는 형용사

“ 예쁘다, 잘생기다 ”

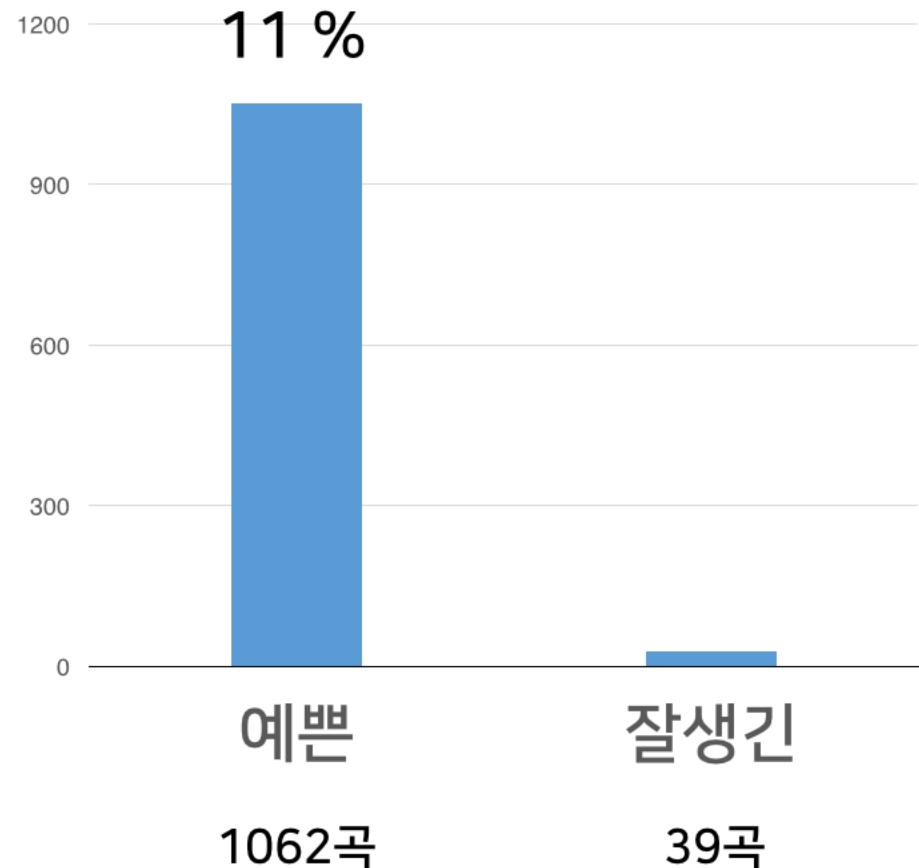
● 본격적인 기사감 파헤치기

konlpy Okt를 이용해
품사 판별해 많이 등장한
형용사 보기

	word	parts_of_speech	female	male	song
300	있다	Adjective	7046.0	16312.0	23358.0
48	없다	Adjective	6555.0	15829.0	22384.0
208	같다	Adjective	3173.0	5803.0	8976.0
43	그렇다	Adjective	2649.0	4981.0	7630.0
289	아니다	Adjective	1900.0	3611.0	5511.0
110	좋다	Adjective	2299.0	2987.0	5286.0
223	어떻다	Adjective	1811.0	2617.0	4428.0
379	아프다	Adjective	1292.0	2631.0	3923.0
917	이렇다	Adjective	1461.0	2293.0	3754.0
173	원하다	Adjective	1122.0	2160.0	3282.0
116	미치다	Adjective	939.0	1962.0	2901.0
1861	행복하다	Adjective	816.0	1881.0	2697.0
416	영원하다	Adjective	827.0	1841.0	2668.0
860	많다	Adjective	682.0	1510.0	2192.0
450	싫다	Adjective	878.0	1282.0	2160.0
293	예쁘다	Adjective	689.0	1424.0	2113.0
273	아름답다	Adjective	478.0	1532.0	2010.0
487	괜찮다	Adjective	588.0	1195.0	1783.0
375	힘들다	Adjective	523.0	1253.0	1776.0
337	변하다	Adjective	620.0	1119.0	1739.0

● 예쁘다 vs 잘생기다

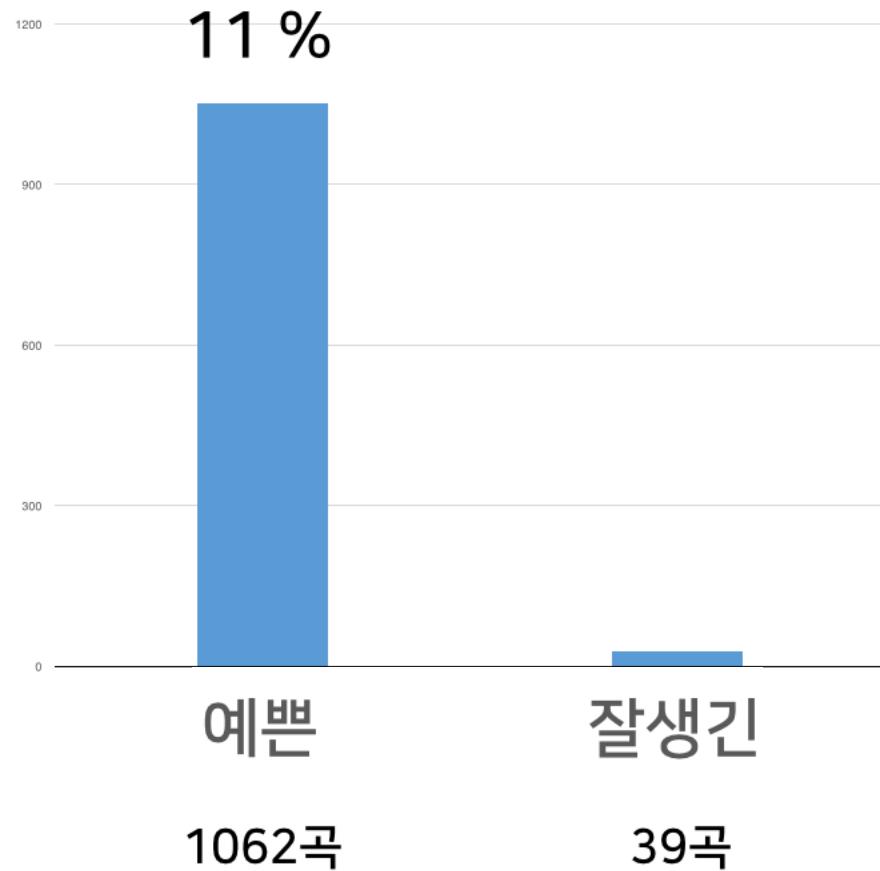
전체 곡의 수 : 9566



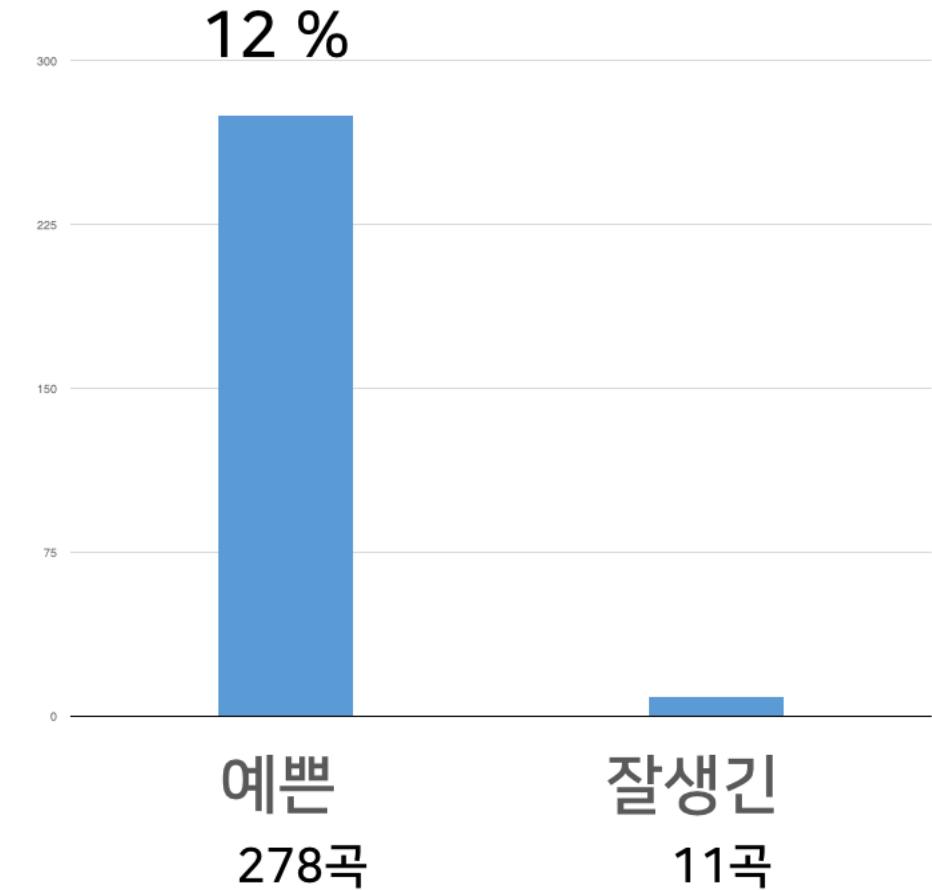
[이예][쁘쁘쁠쁠뽀뽀뽀] 잘생 잘생[기긴길겨겼]

● 예쁘다 vs 잘생기다

전체 곡의 수 : 9566



타이틀 곡의 수 : 2379



•Word Cloud

예쁜



잘생긴



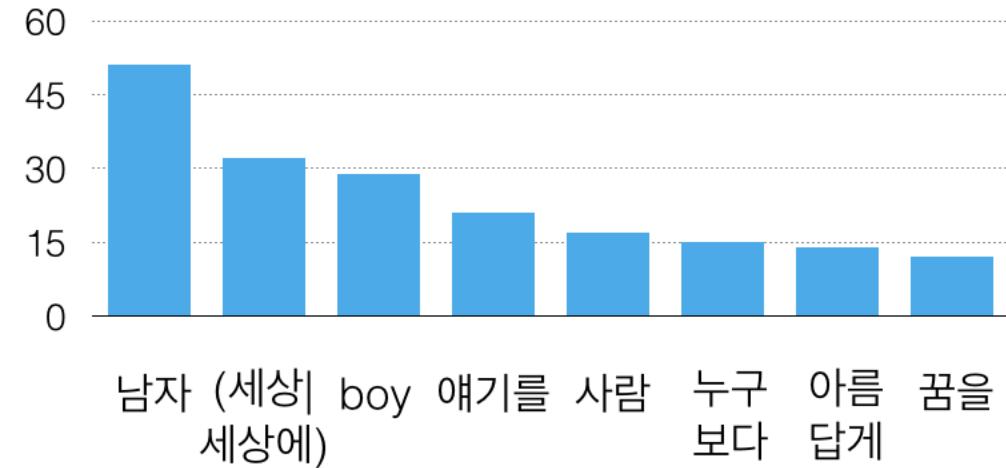
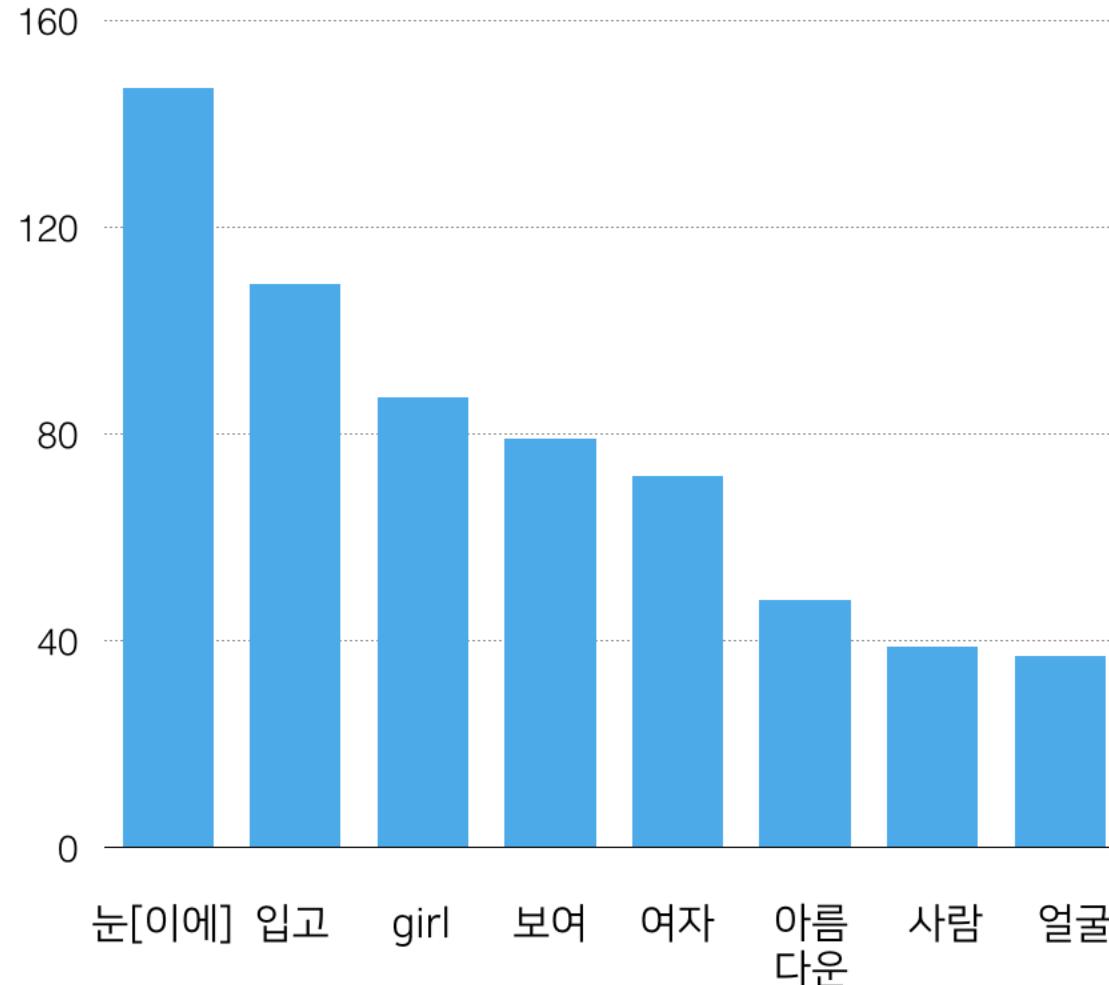
● '예쁘다'와 '멋지다' 비교하기 - 여자, 남자와 함께 나온 형용사

여자

남자

다른	다른
같은	그런
그런	나쁜
예쁜	같은
나쁜	멋진

● 예쁘다 vs 멋지다



• 한 세트처럼 등장하는 형용사

립스틱은 **빨간** 립스틱

치마는 **짧은** 치마

스타킹은 **새까만** 스타킹

하이힐은 **아찔한** 하이힐

●가사 속 등장하는 단어 : 하이힐

여자 아이돌

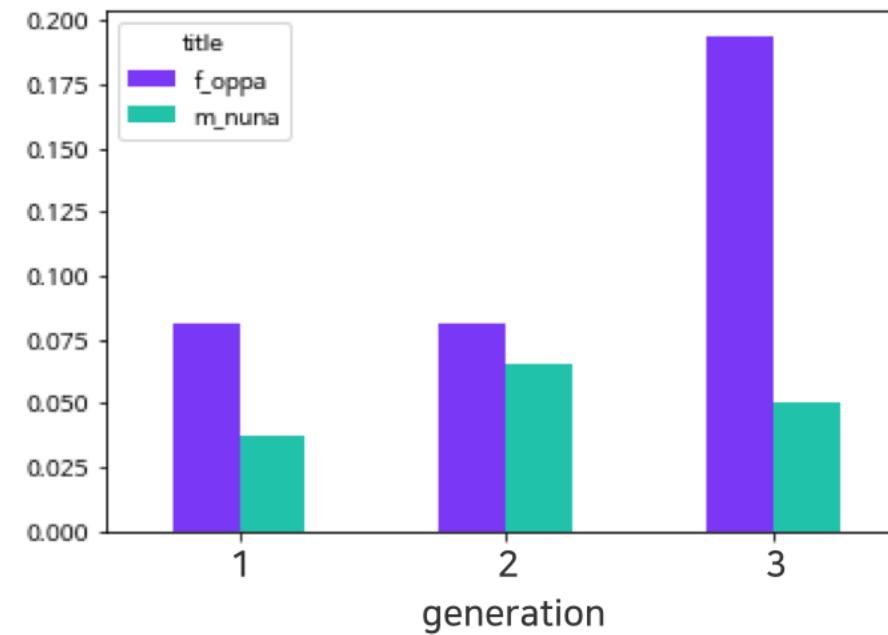
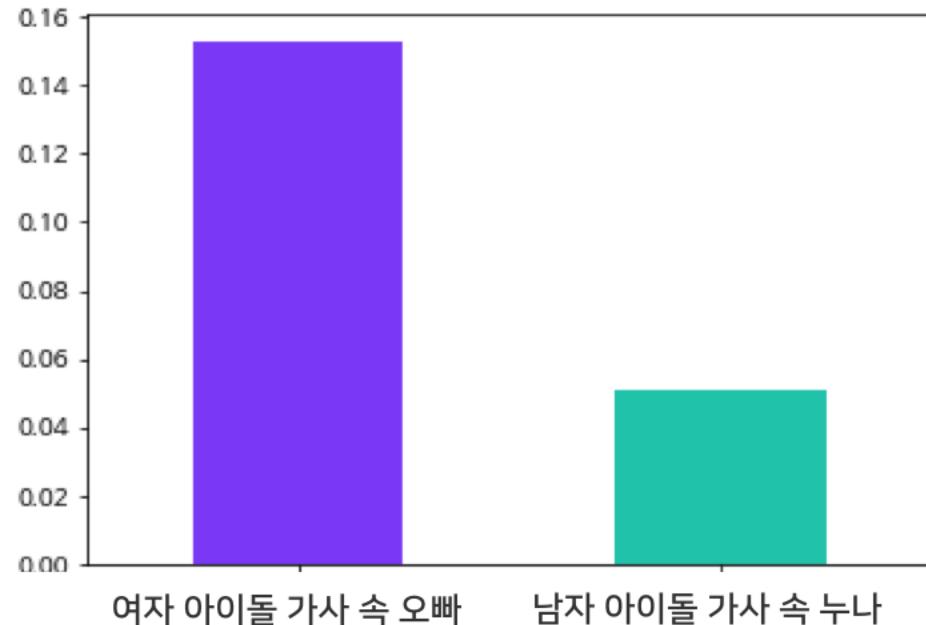


남자 아이돌

사랑해 너하나만 굽 높은 하이힐 날 좀 봐요 그대가 좋아한다 해서 신었잖아요 섹시한 하이힐
짧아진 치마 빨간 립스틱 풀어 해쳐진 머리도 길어진 손톱 길어진 하이힐 매일 취기에 살아
반짝거리는 불빛의 유혹 아찔한 내 하이힐 들썩거리는 Party Mama She told
때로는 bad girl Glam glam girl 높은 하이힐 꽤 진한 아이라인 머리부터 발끝이 bling bling
Oh Wait up girl 빠딱한 하이힐 걸음걸인 빠뚤빼뚤 Grow up girl Oh Grow up girl
귀찮아도 매일 화장을 하고 다리가 아파도 굽 높은 하이힐 신고 너에게 잘 보이려고 예쁜 옷 입고
내 모든 미모 각선미 비결이 뭐 그건 내가 신은 높은 하이힐 잘 봐 완벽해 보이는 비밀 날 봐
굳이 노출 안 해도 아찔한 나의 하이힐 새까만 스타킹 도저히 눈을 뗄 수 없을 걸
고민 고민 끝에 고른 옷을 빼입고 오늘따라 조금 높은 하이힐 어색하지만 들떠버린 걸음걸이
더 높은 하이힐 신고 진한 향수 하고 널 유혹하는 나를 좀 나를 봐줘요
너를 만날 때마다 치마도 입고 키를 맞추려 하이힐 신고 설레 설레
Oh 나랑 밀당하지마 짧은 치마를 입고 널 만나러 가고 하이힐을 신고 애교부려봐도 No Oh
힘준 내 하이힐도 공들인 립스틱도 못 알아보는 눈들이 참 따분해 그 때쯤 니가 걸어

아가씨 하이힐 그 끝으로 오르는 치마 차마 말로는 못할 아무도 모르는 나만의 상상
she 비친 실루엣 도도한 날카로운 하이힐 소리 뜨거워진 심장을 찌르고
hate a 컨버스 화려한 니 걸모습에는 말야 검정스타킹에 아찔한 하이힐 그래 그것은 사기야
하얀 하이힐 살짝 가린 레이스 스커트 Beautiful lady 네 어깨 위 아슬아슬 시스루
몰라보게 여자가 된 걸 핑크 색깔 운동화는 oh 아찔한 하이힐로 단발머리 교복 치만 생머리의 미니스커트
보내주긴 싫어 걸 벗어둬 니 하이힐 오늘 가기 전에 yeah now
니 하이힐 뒤에 넌 내 바로 맞은편 가까워진 얼굴 아름다운 너 오늘따라 심장이
이건 말실수 AW 그대로가 좋아 니 맨 얼굴이 좋아 화장하지 마 하이힐 신지 않아도 돼

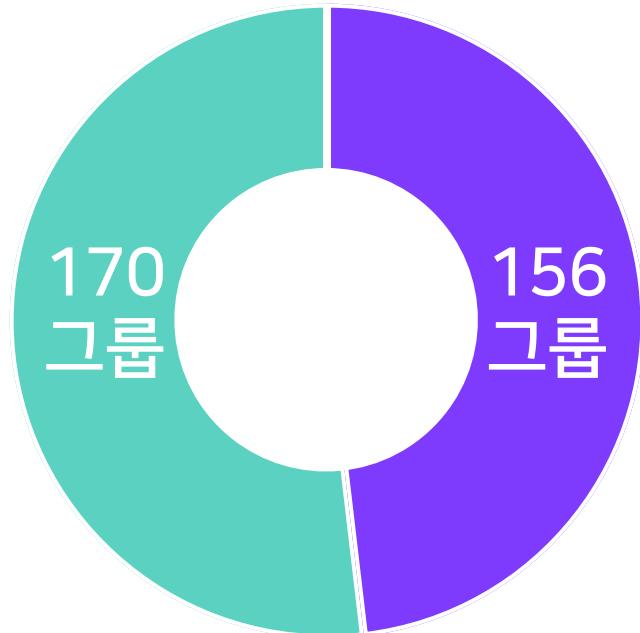
●번외 - 아이돌의 오빠와 누나는 몇 명일까?



INDUSTRY

● 그룹과 곡의 성비

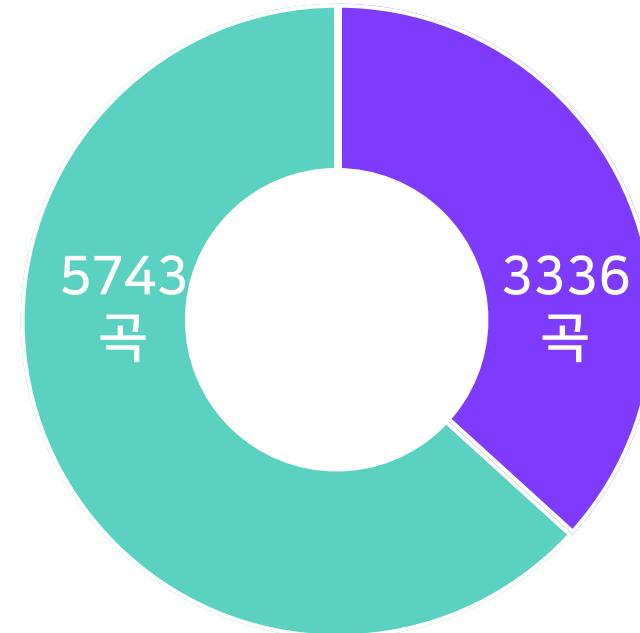
그룹 수



x 1.1

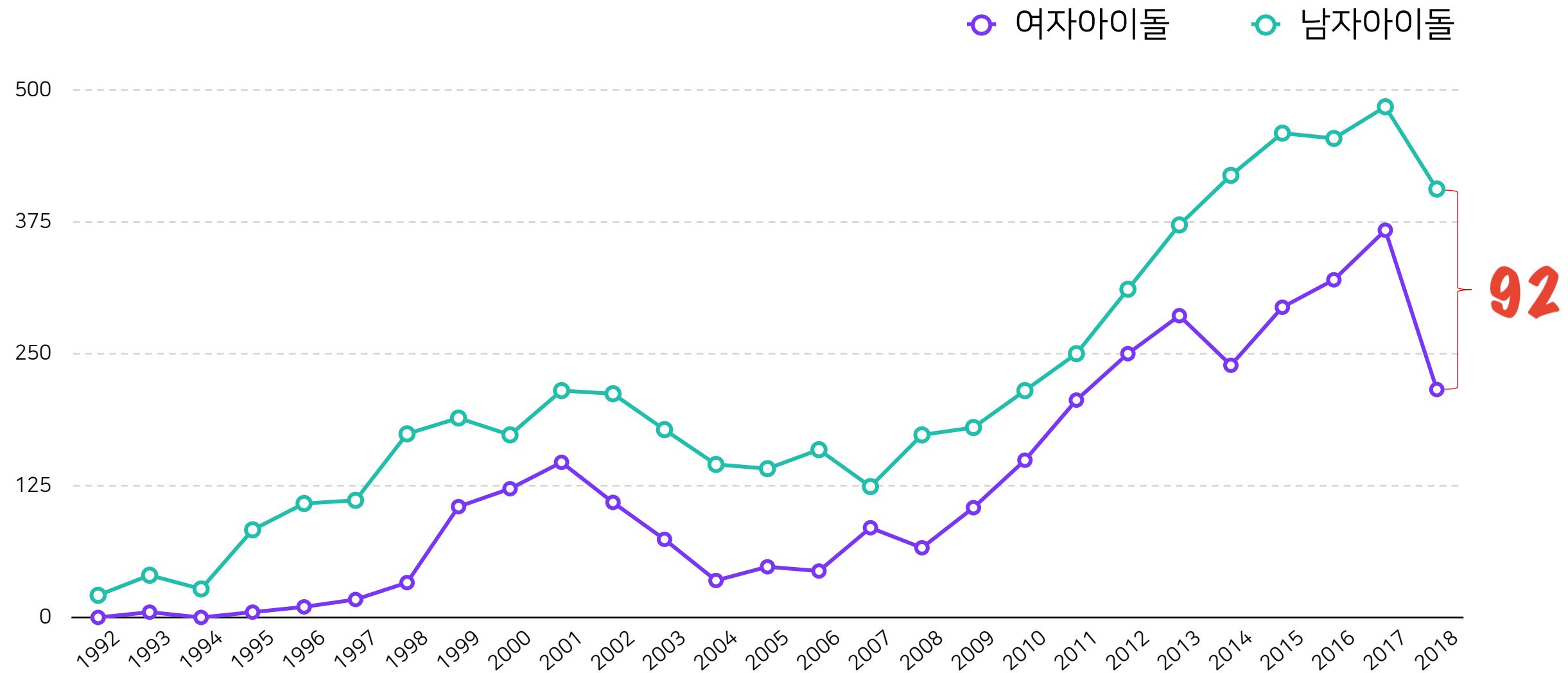
- 여성아이돌
- 남성아이돌

곡 수



x 1.7

● 발매년도 기준 성비



● 네트워크 시각화 – 데이터프레임 전처리

- 1세대 : 1992 ~ 2000년
- 2세대 : 2001 ~ 2009년
- 3세대 : 2010 ~ 2018년

	source	count
0	박진영	68
1	김영아	59
2	유영진	53
3	유건렬	40
4	이승호	32
5	M (이민우)	26
6	테디	25
7	이상민	23
8	강타 (KANGTA)	22
9	장대성	21
10	천명훈	19
11	송백경	17
12	조은희	16
13	서정환	16
14	문희준	14
15	박성호 (구피)	14
16	에릭	14
17	신동욱	13

● 네트워크 시각화 - 노드/링크 생성

```
# 노드 생성하기
nodes_list_gen1 = []
nodes_list_gen1.clear()
for index, row in generate_node_gen1.iterrows():
    nodes_list_gen1.append({"name":row['source'], "type": "lyricist", "impor":row['count']})
for idol in group_dict_gen1_artist:
    nodes_list_gen1.append({"name":idol, "type": "artist", "impor":0})
nodes_list_gen1
```

```
ame': '김영아', 'type': 'lyricist', 'impor': 69},
ame': '박진영', 'type': 'lyricist', 'impor': 69},
ame': '유영진', 'type': 'lyricist', 'impor': 54},
ame': '이승호', 'type': 'lyricist', 'impor': 43},
ame': '유건형', 'type': 'lyricist', 'impor': 40},
ame': 'M (이민우)', 'type': 'lyricist', 'impor': 28},
ame': '강타 (KANGTA)', 'type': 'lyricist', 'impor': 28},
ame': '이상민', 'type': 'lyricist', 'impor': 28},
ame': '장대성', 'type': 'lyricist', 'impor': 26},
ame': 'tedi', 'type': 'lyricist', 'impor': 25},
ame': '조은희', 'type': 'lyricist', 'impor': 23},
ame': '천명훈', 'type': 'lyricist', 'impor': 19},
ame': '송백경', 'type': 'lyricist', 'impor': 19},
ame': '서정환', 'type': 'lyricist', 'impor': 17},
ame': '장우혁', 'type': 'lyricist', 'impor': 16},
ame': '이재원', 'type': 'lyricist', 'impor': 15},
ame': '에릭', 'type': 'lyricist', 'impor': 15},
ame': '박성호 (구피)', 'type': 'lyricist', 'impor': 14},
ame': '지국현', 'type': 'lyricist', 'impor': 14},
```

```
1 # 링크 생성하기
2 links_list_gen1 = []
3 for link in temp_links_list_gen1:
4     record = {"value":link['value'], "source":unique_idol_gen1_all.get_loc(link['source']),
5     "target": unique_idol_gen1_all.get_loc(link['target'])}
6     links_list_gen1.append(record)
7 links_list_gen1
```

```
[{'value': 68, 'source': 0, 'target': 31},
{'value': 40, 'source': 1, 'target': 123},
{'value': 26, 'source': 2, 'target': 124},
{'value': 24, 'source': 3, 'target': 125},
{'value': 23, 'source': 4, 'target': 126},
{'value': 22, 'source': 5, 'target': 127},
{'value': 19, 'source': 6, 'target': 128},
{'value': 19, 'source': 7, 'target': 129},
{'value': 18, 'source': 3, 'target': 130},
{'value': 17, 'source': 8, 'target': 131},
{'value': 16, 'source': 9, 'target': 129},
{'value': 16, 'source': 10, 'target': 123},
{'value': 14, 'source': 11, 'target': 127},
{'value': 14, 'source': 12, 'target': 131},
{'value': 14, 'source': 13, 'target': 124},
{'value': 14, 'source': 14, 'target': 129},
{'value': 13, 'source': 15, 'target': 127},
{'value': 13, 'source': 7, 'target': 124},
{'value': 13, 'source': 16, 'target': 132},
```

● 네트워크 시각화 - json 파일 추출

```
1 json_prep_gen1 = {"links":links_list_gen1, "nodes":nodes_list_gen1}
2 json_prep_gen1.keys()
dict_keys(['links', 'nodes'])

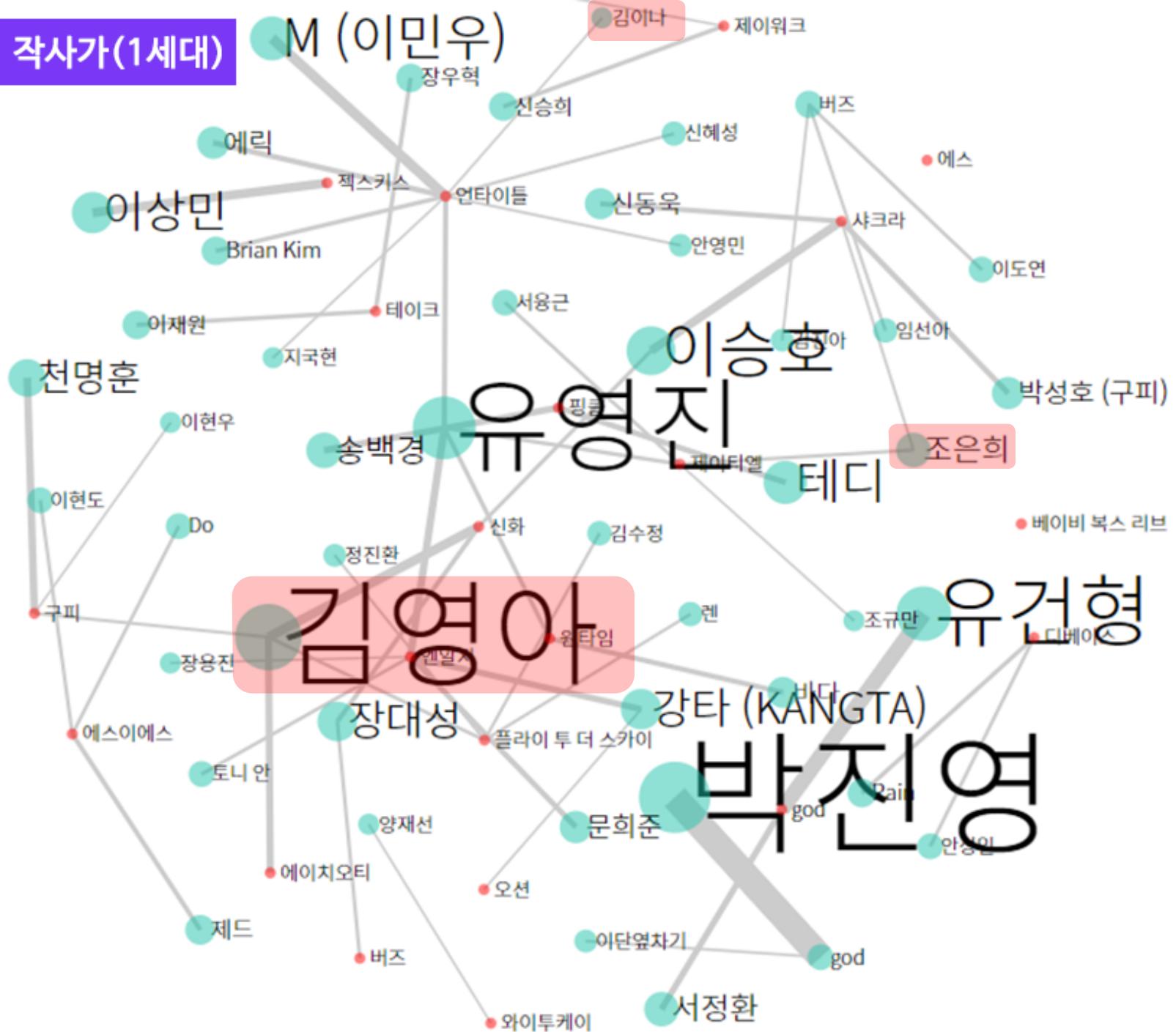
1 json_dump_gen1 = json.dumps(json_prep_gen1, indent=1, sort_keys=True, ensure_ascii=False)

1 print(json_dump_gen1)
{
  "links": [
    {
      "source": 0,
      "target": 31,
      "value": 68
    },
    {
      "source": 1,
      "target": 123,
      "value": 40
    },
    {
      "source": 2,
      "target": 124,
      "value": 26
    }
  ]
}
```

작사가

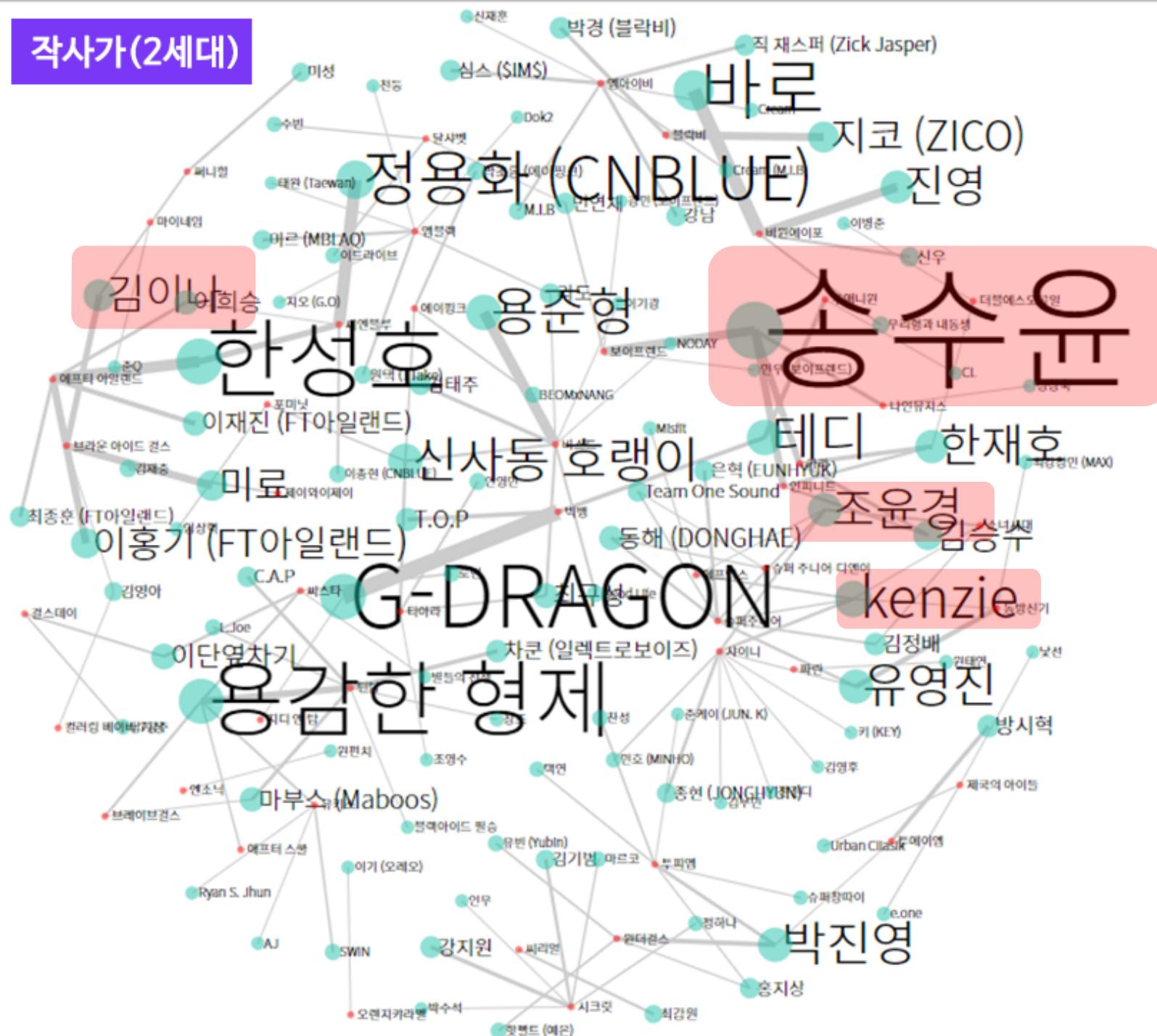
작사가(1세대)

M (이민우)

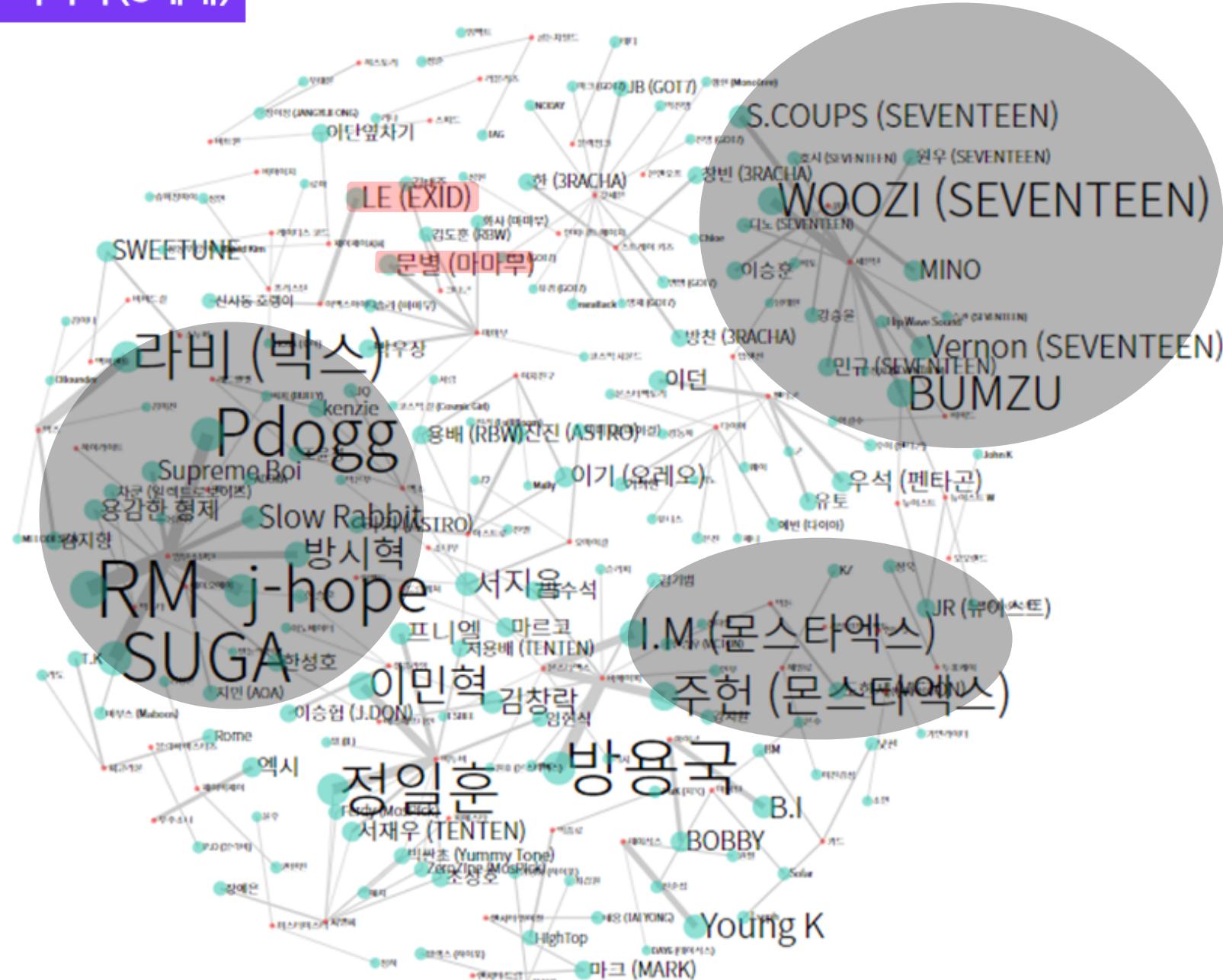


LYRICSIS

작사가(2세대)

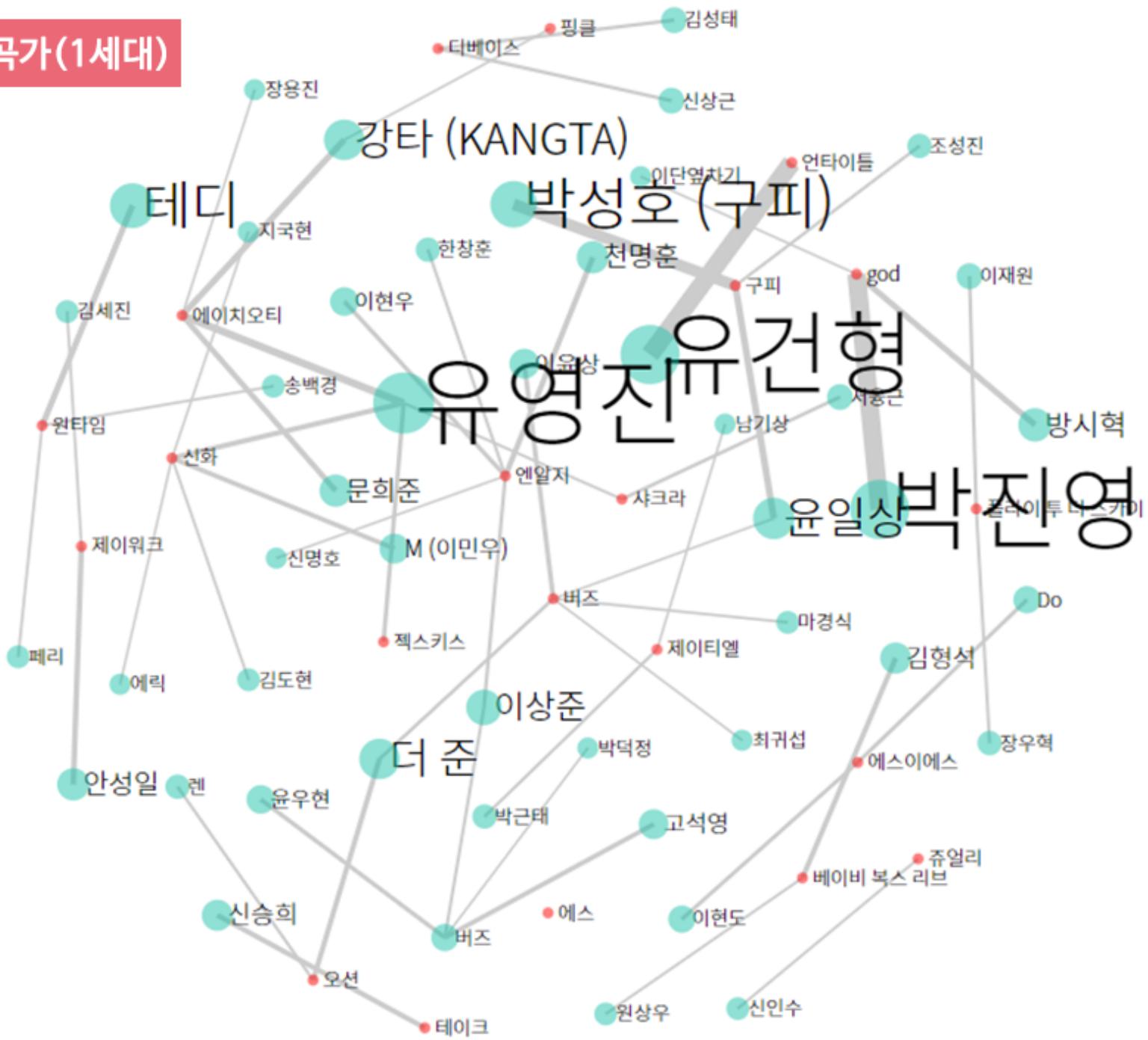


작사가(3세대)

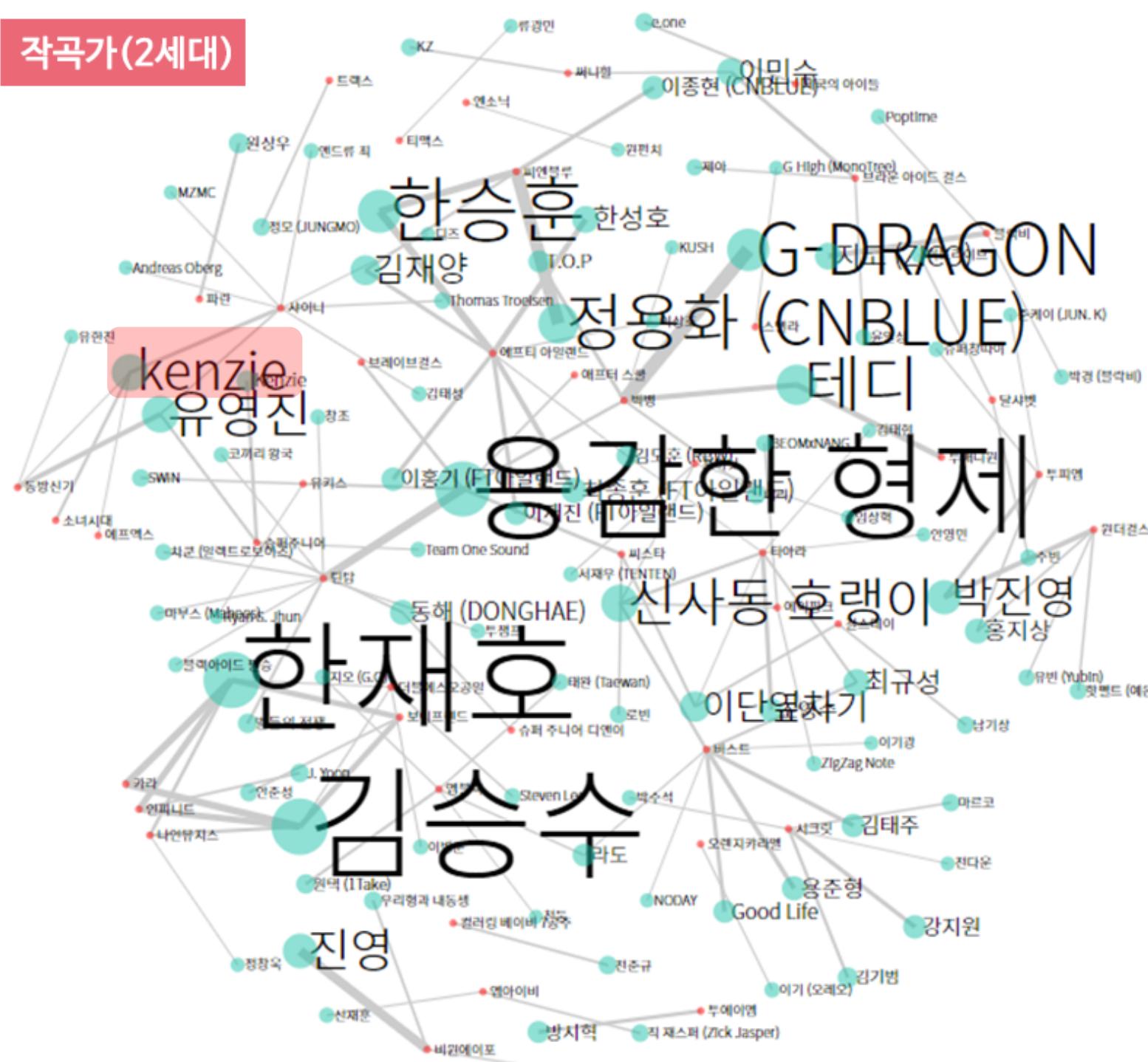


작곡가

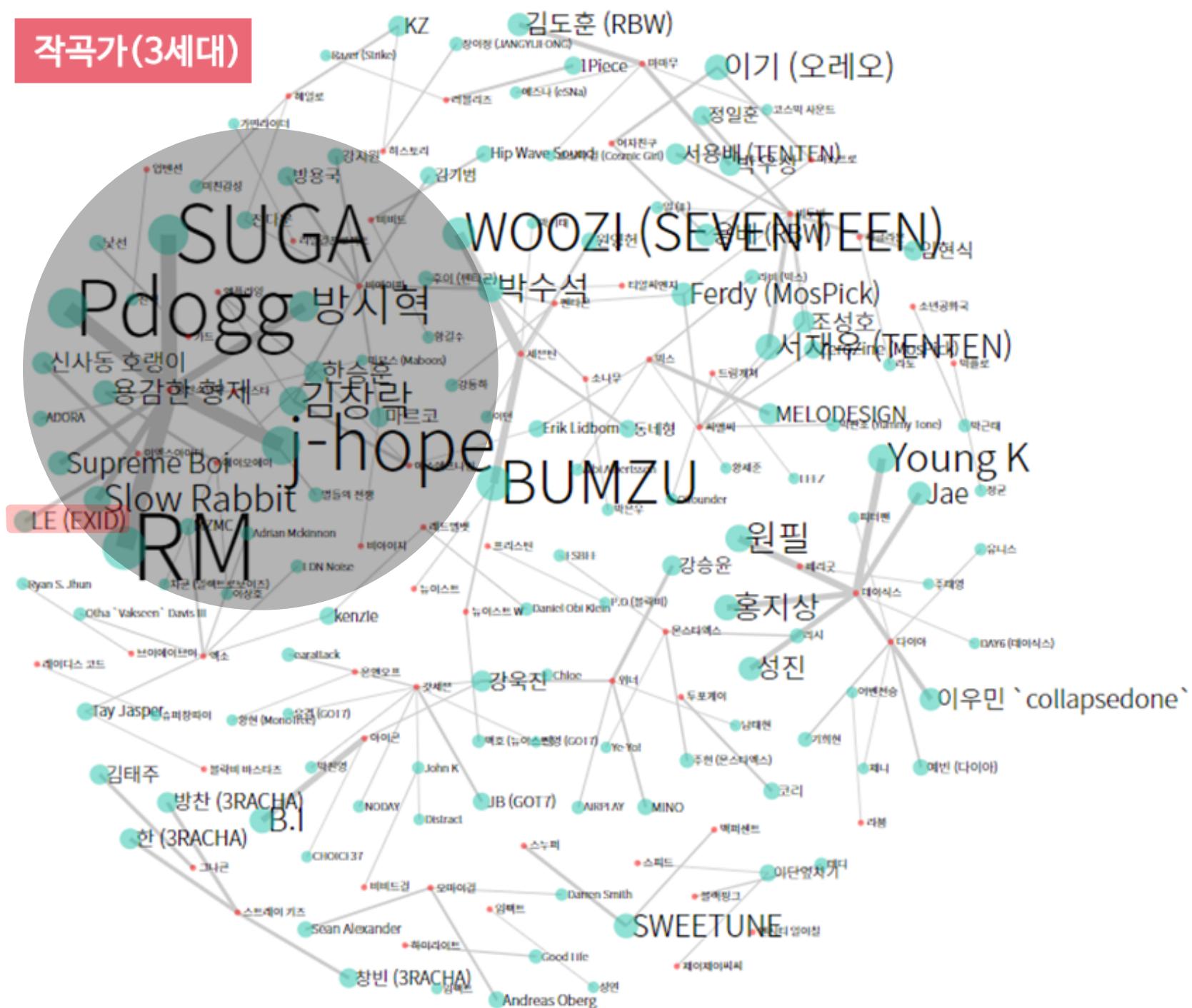
작곡가(1세대)



작곡가(2세대)

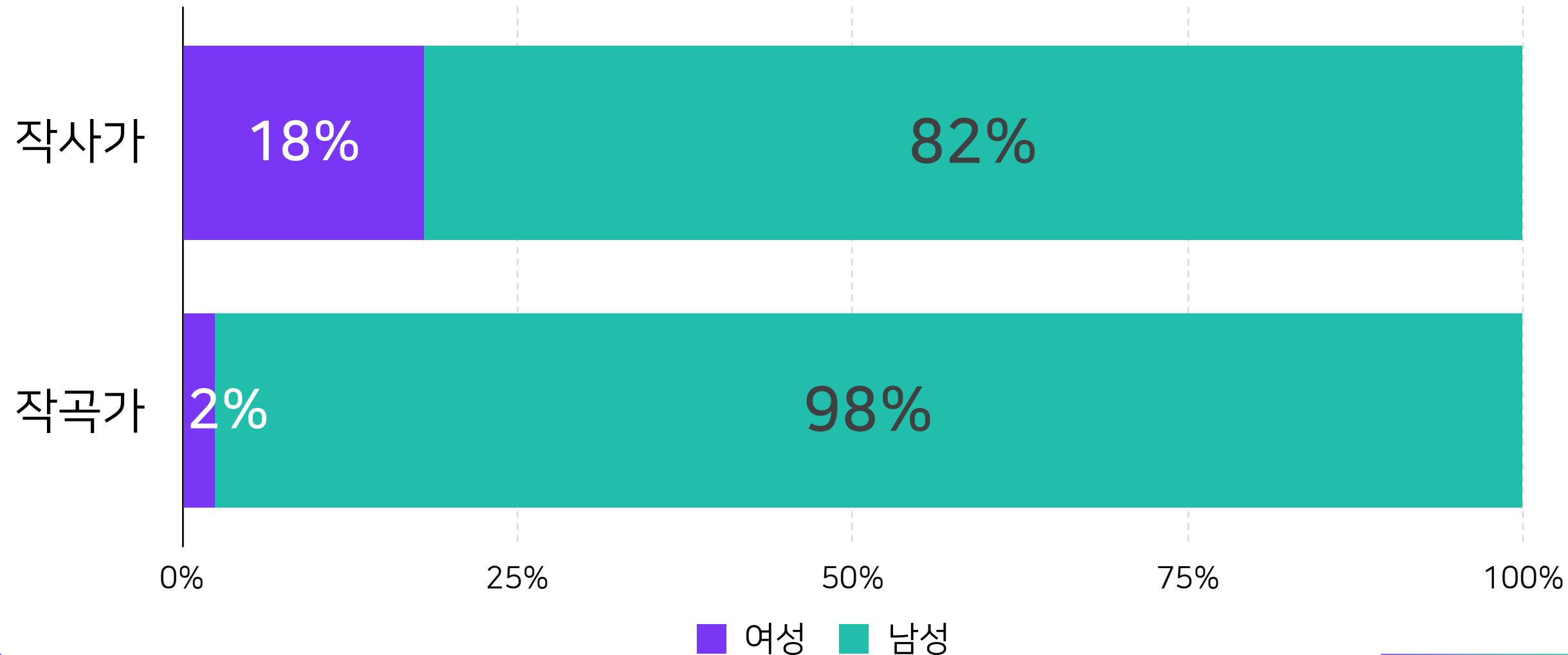


작곡가(3세대)



● 전체 작사/작곡가 Top 100 성별 참여율

(곡 수)



● Since 2012 누가 작사를 하나?

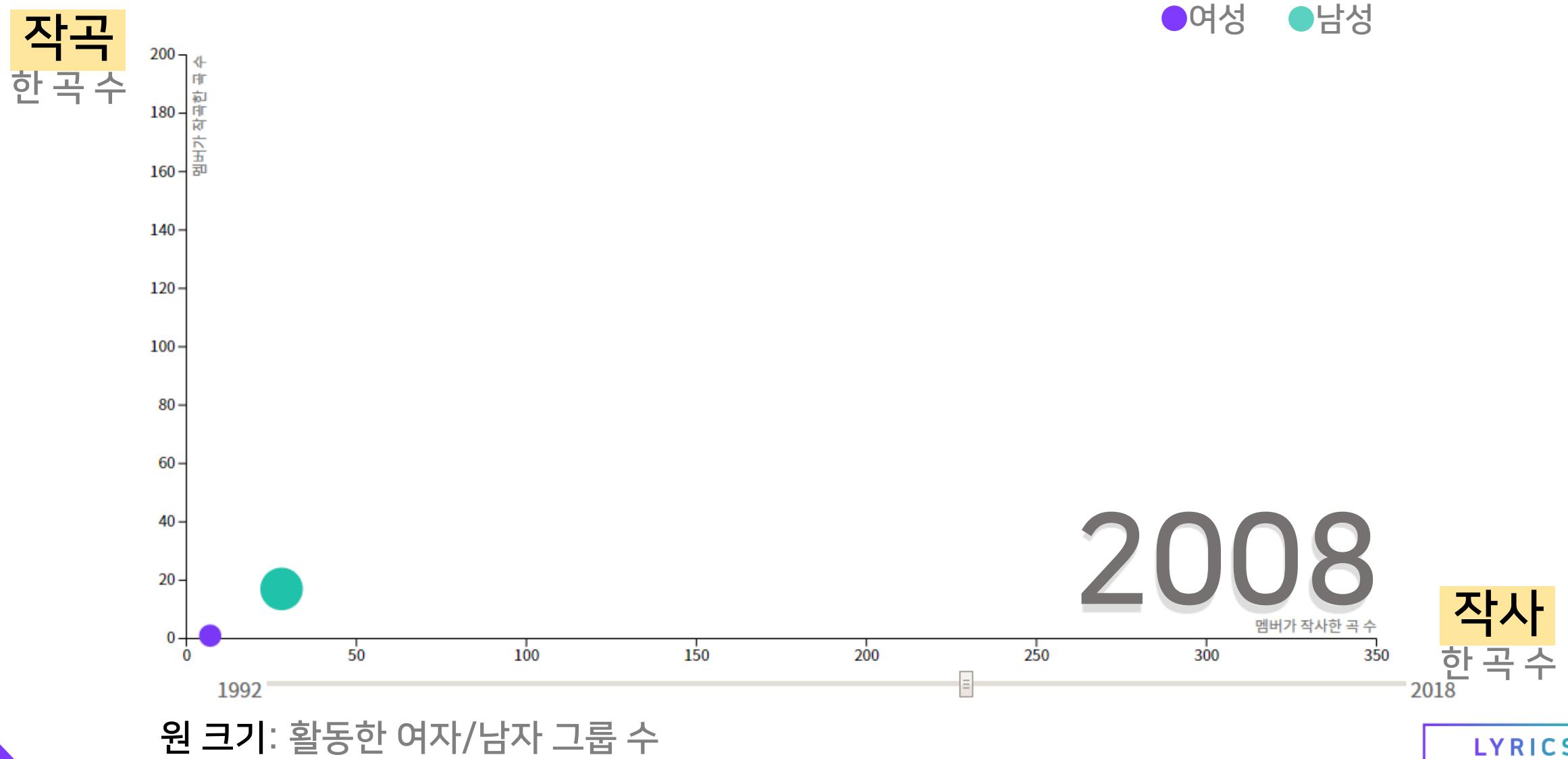


2012년~현재
작사가 랭킹100
총 2428곡 중

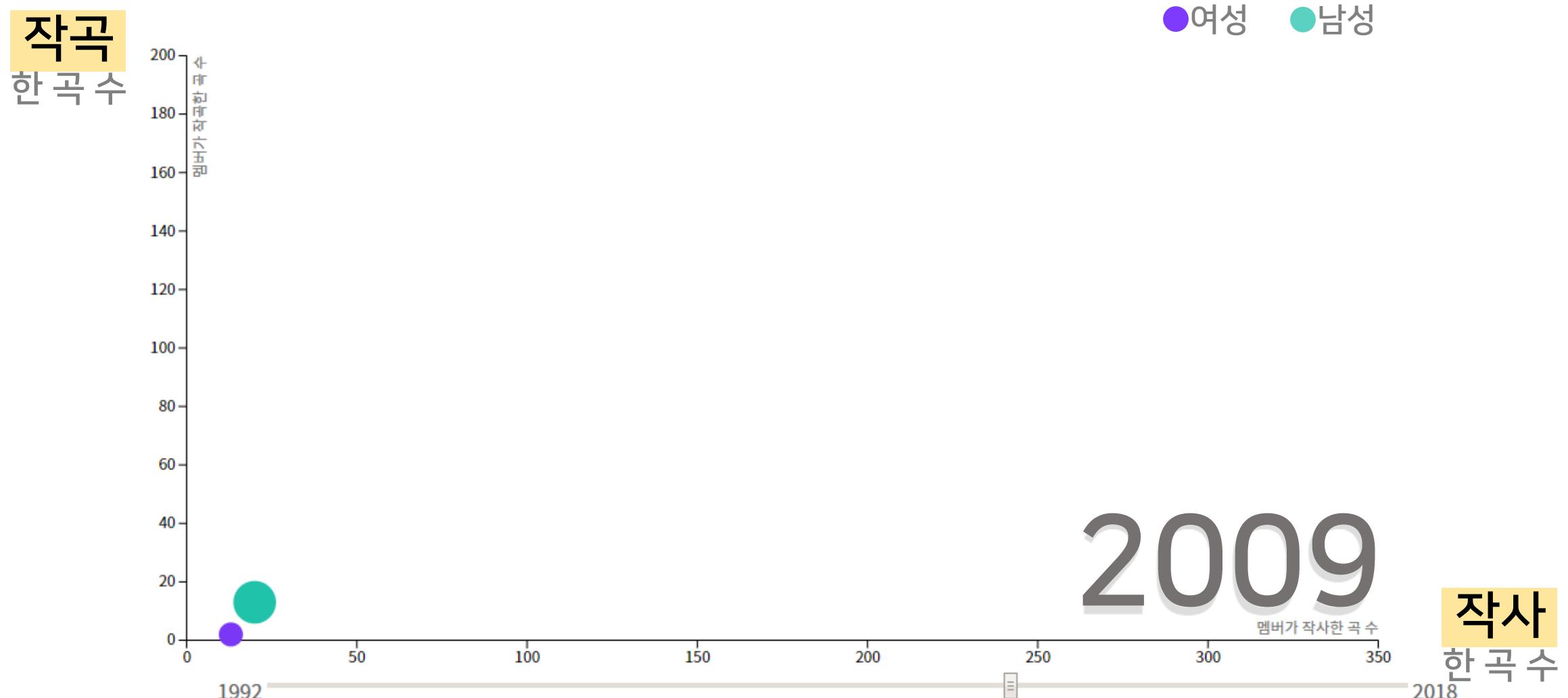
여자 아이돌
186곡



● Since 2012 누가 작사를 하나?

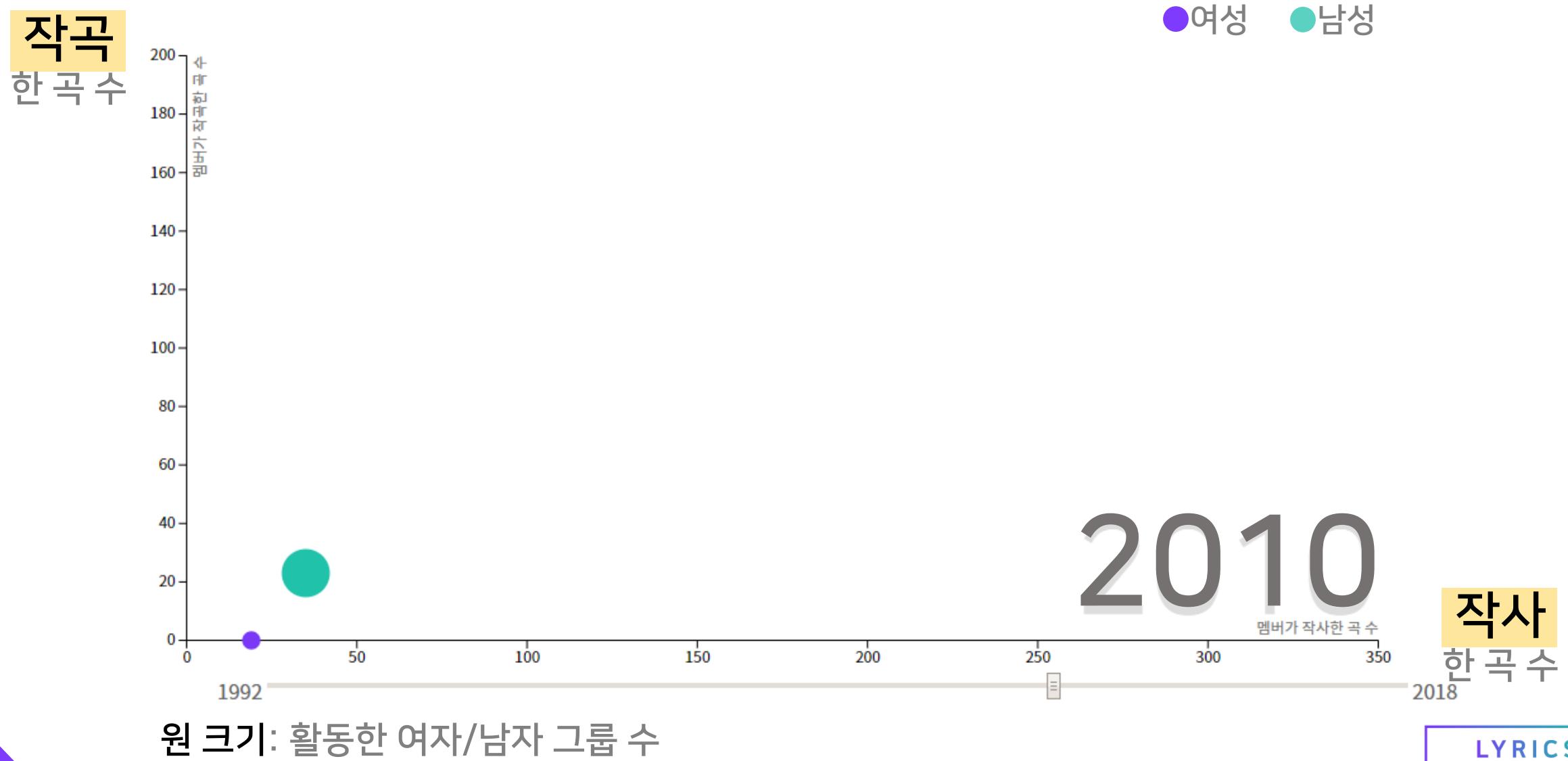


● Since 2012 누가 작사를 하나?

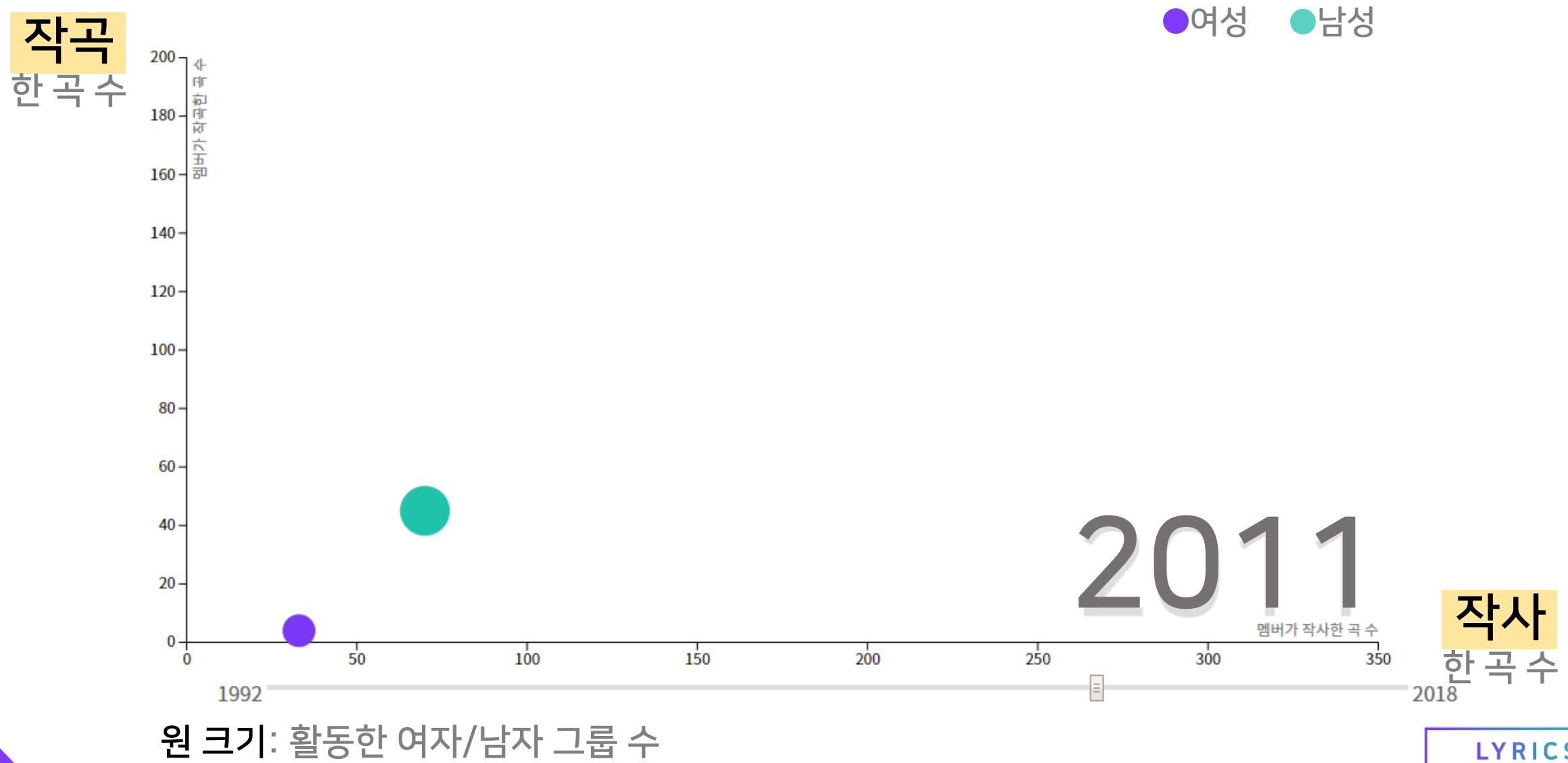


원 크기: 활동한 여자/남자 그룹 수

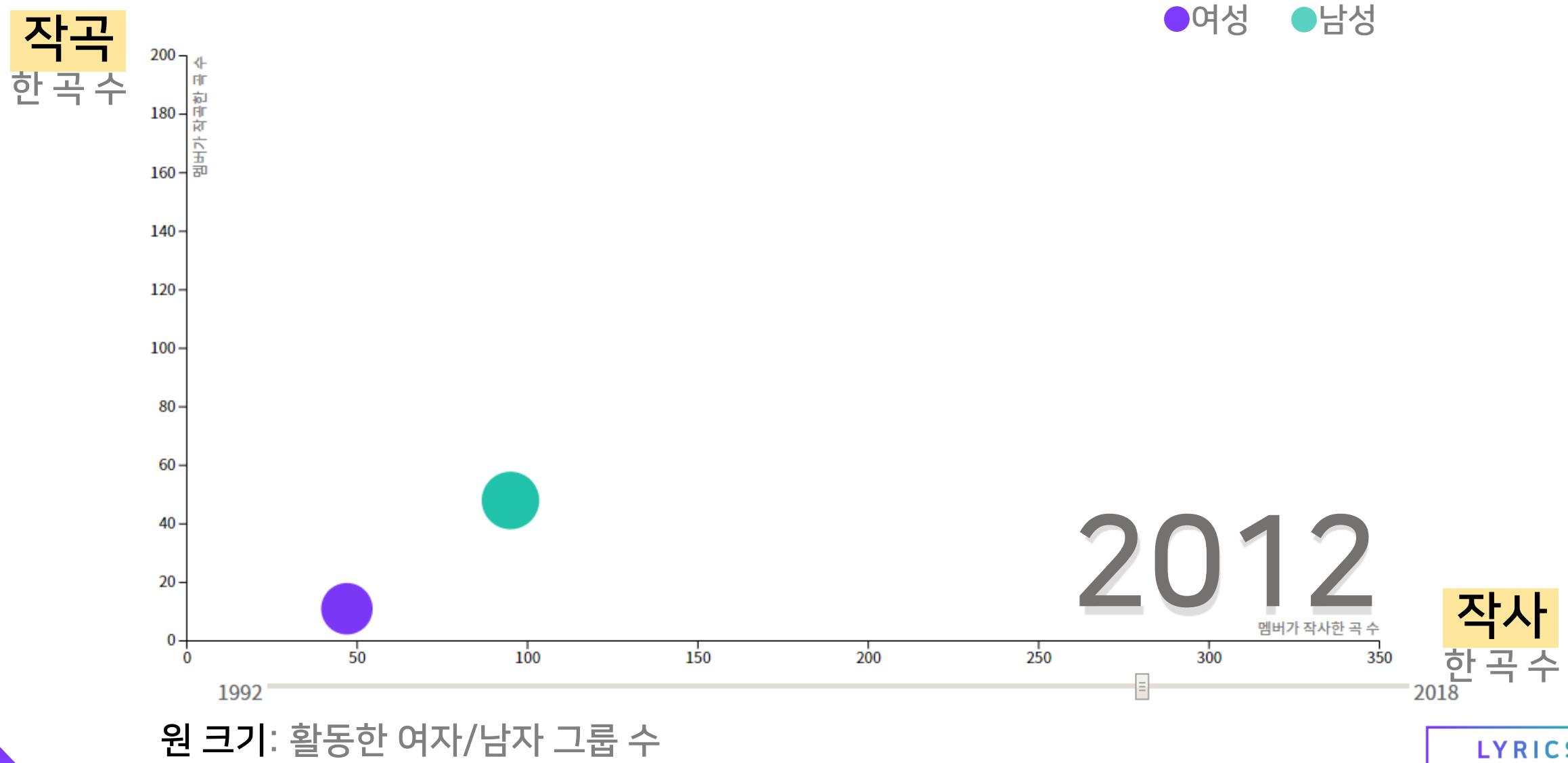
● Since 2012 누가 작사를 하나?



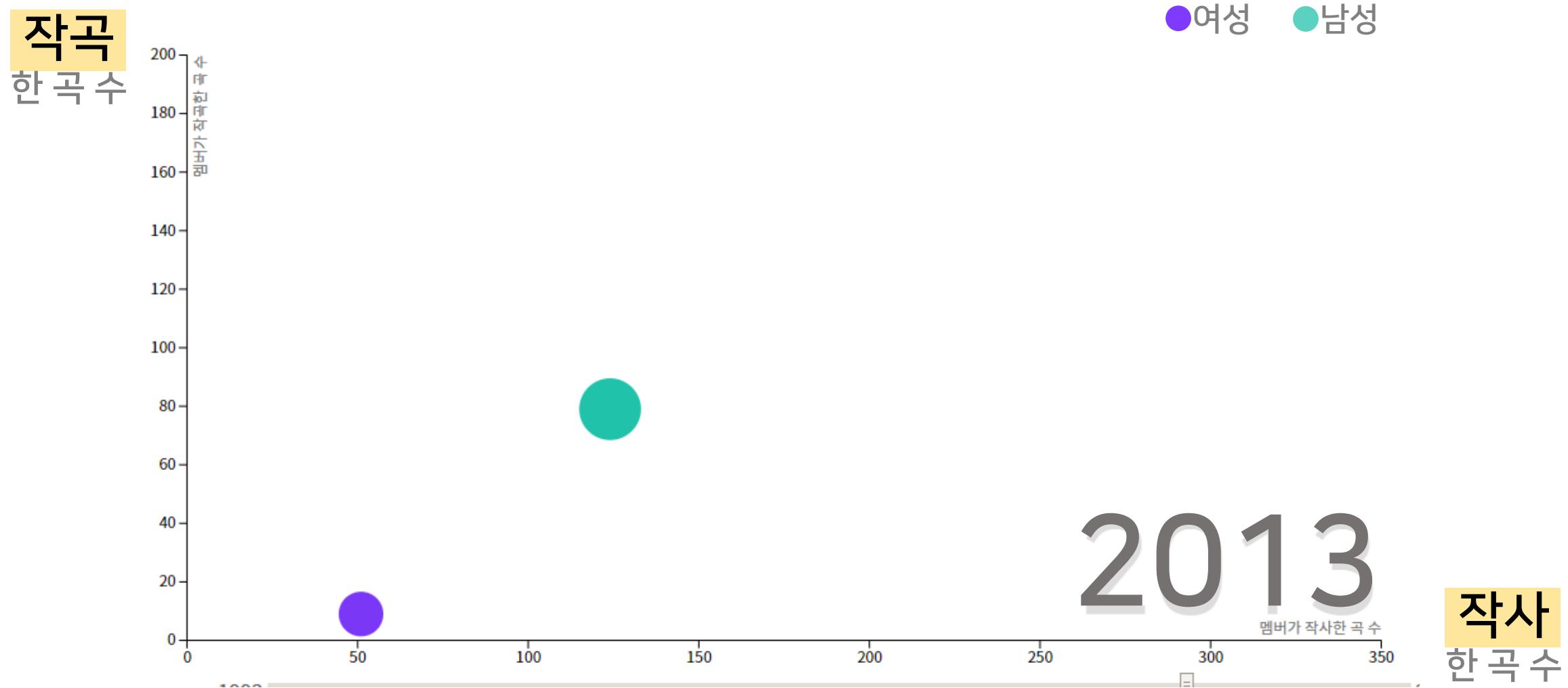
● Since 2012 누가 작사를 하나?



● Since 2012 누가 작사를 하나?



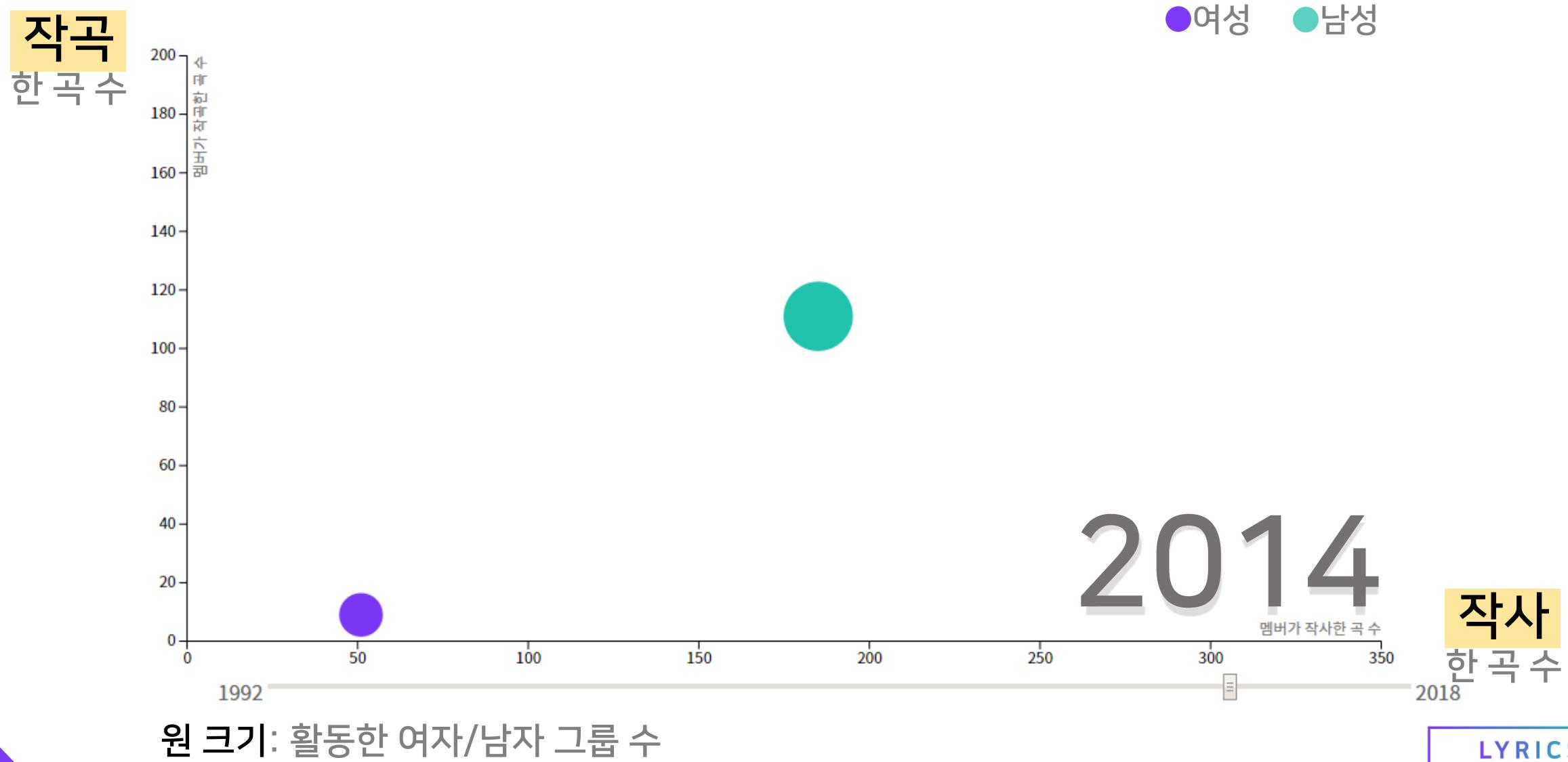
● Since 2012 누가 작사를 하나?



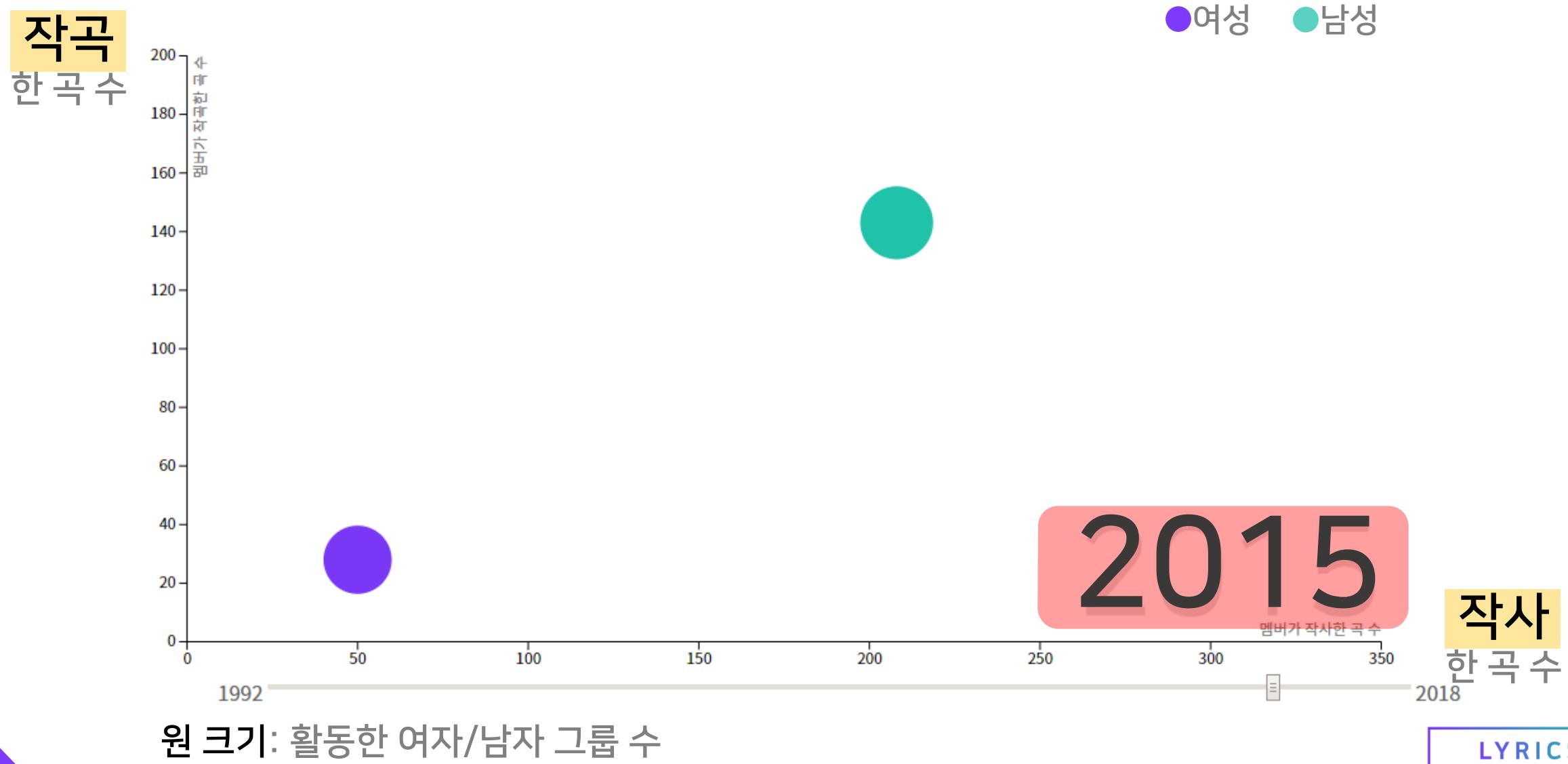
원 크기: 활동한 여자/남자 그룹 수

LYRICSIS

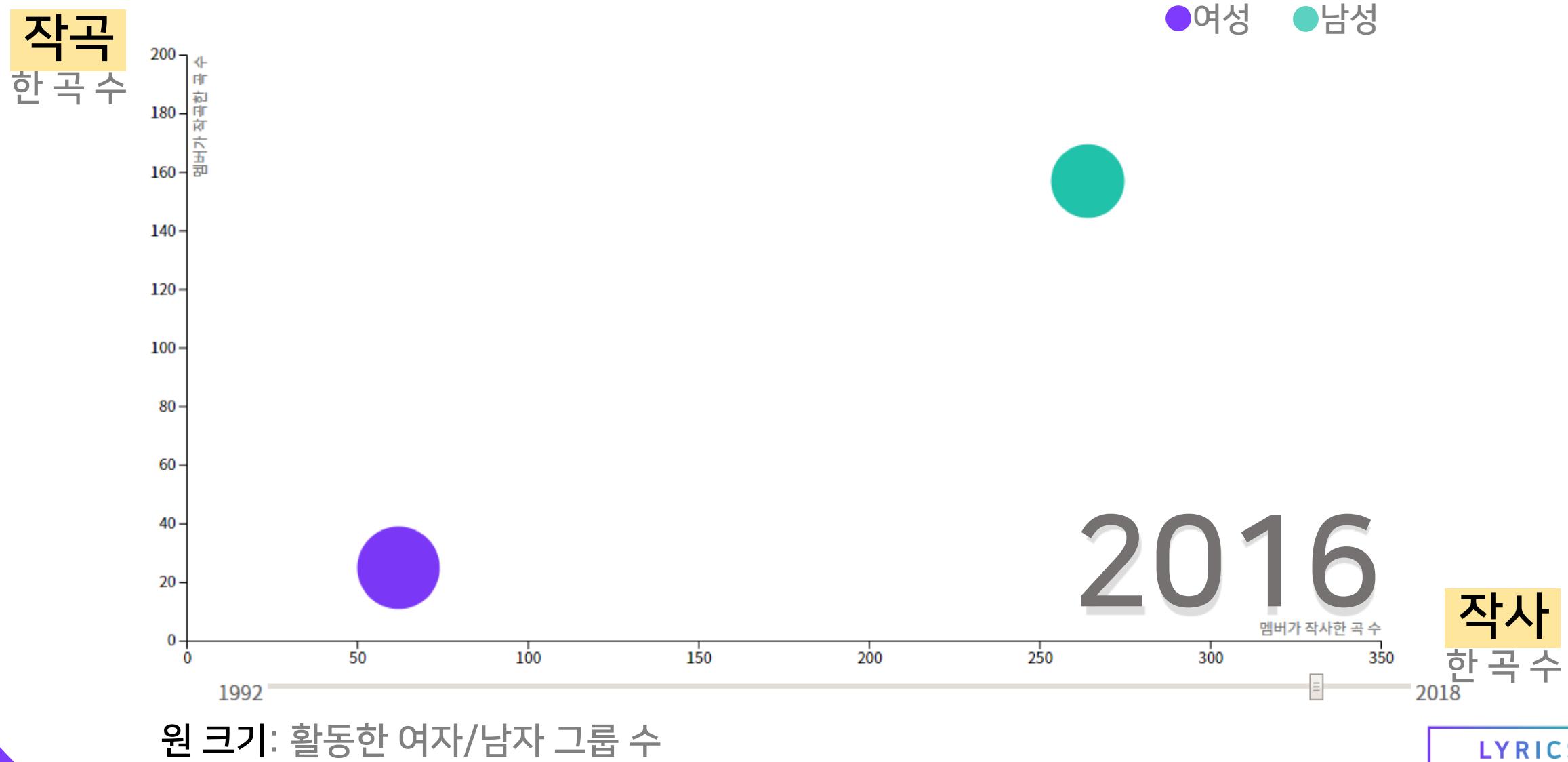
● Since 2012 누가 작사를 하나?



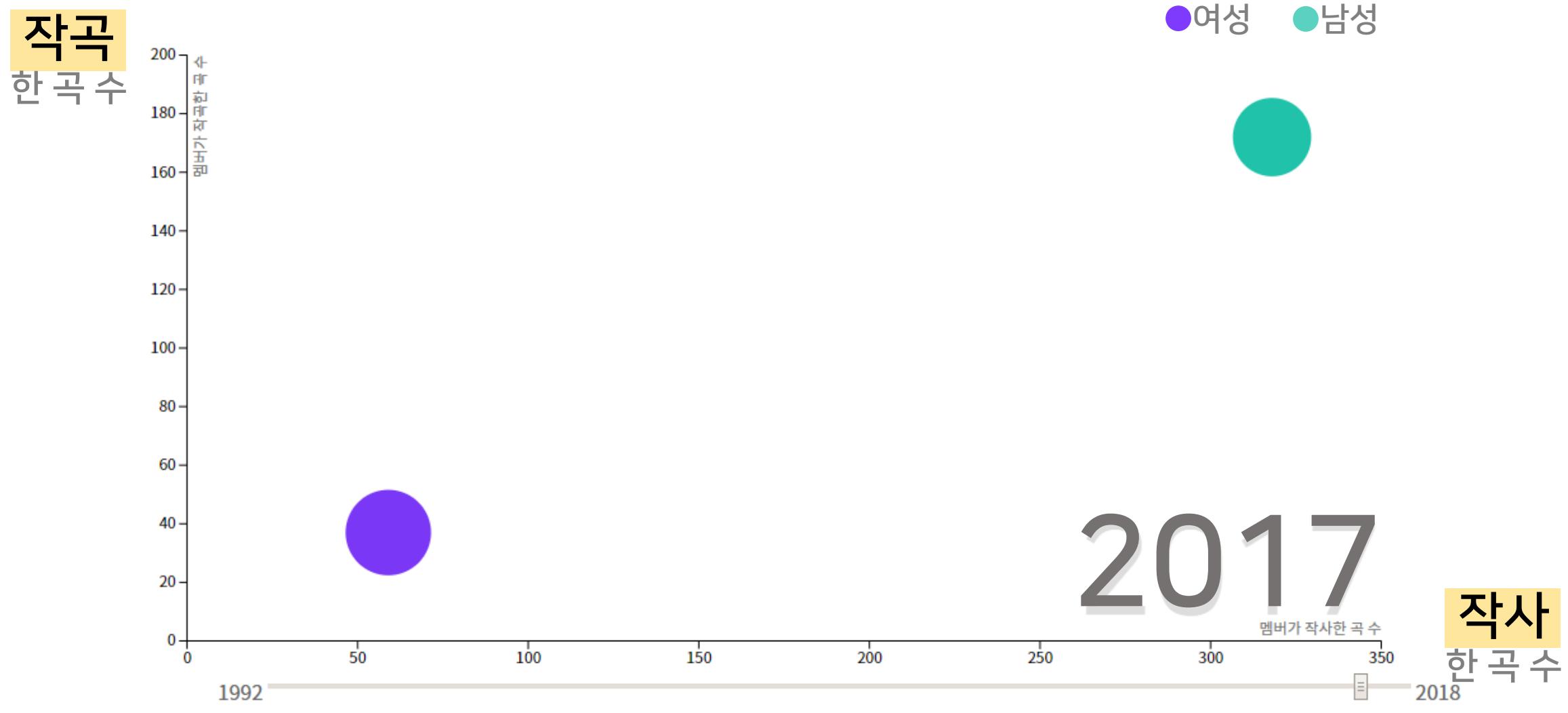
● Since 2012 누가 작사를 하나?



● Since 2012 누가 작사를 하나?



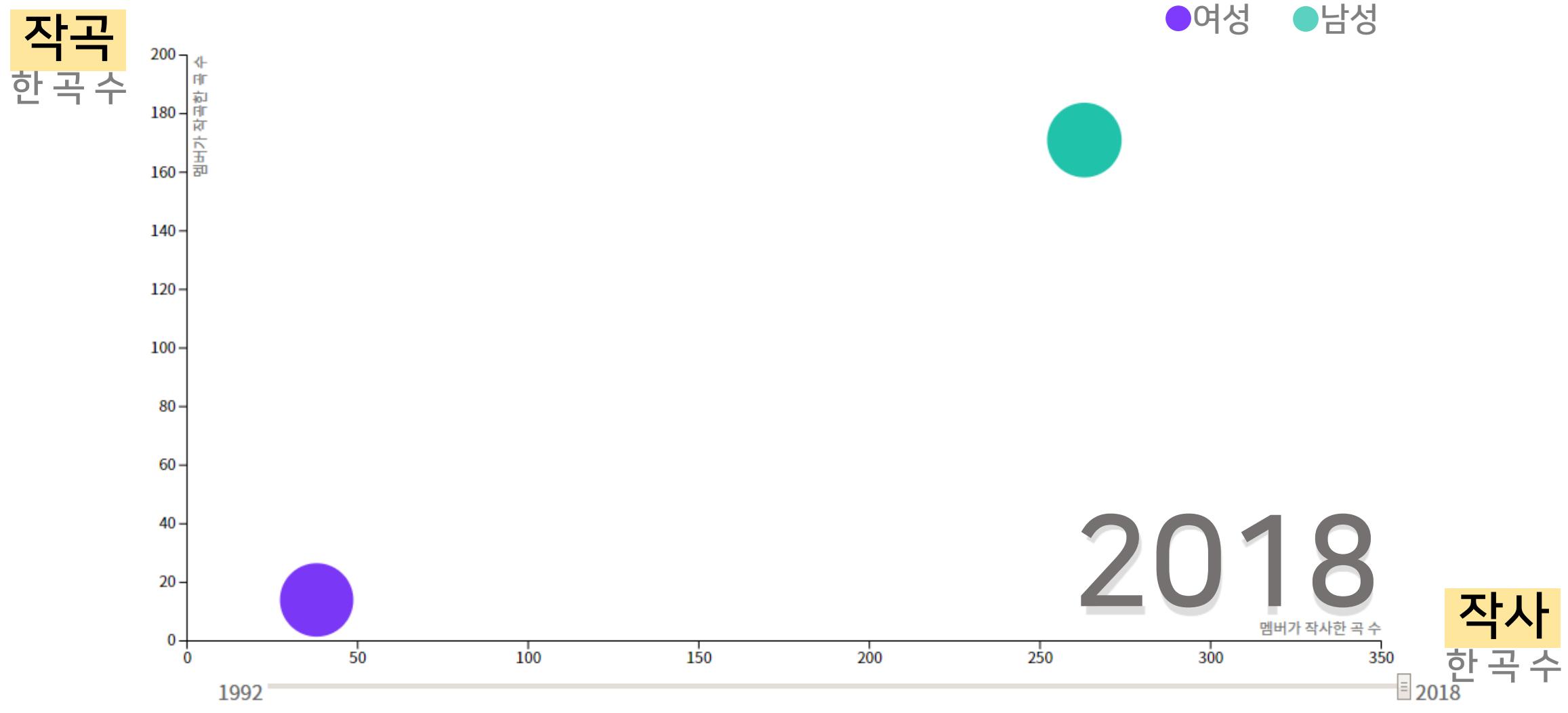
● Since 2012 누가 작사를 하나?



원 크기: 활동한 여자/남자 그룹 수

LYRICSIS

● Since 2012 누가 작사를 하나?



원 크기: 활동한 여자/남자 그룹 수

LYRICSIS

●아이돌 산업, 이대로라면..?

