



4.5 Hz



구성원

정유진 유현준 이도형 최세인 조민진



목차



서비스
선정 이유



데이터



서비스
구성



모델링



웹 구성



개선점

음악 추천 서비스를 선정한 이유



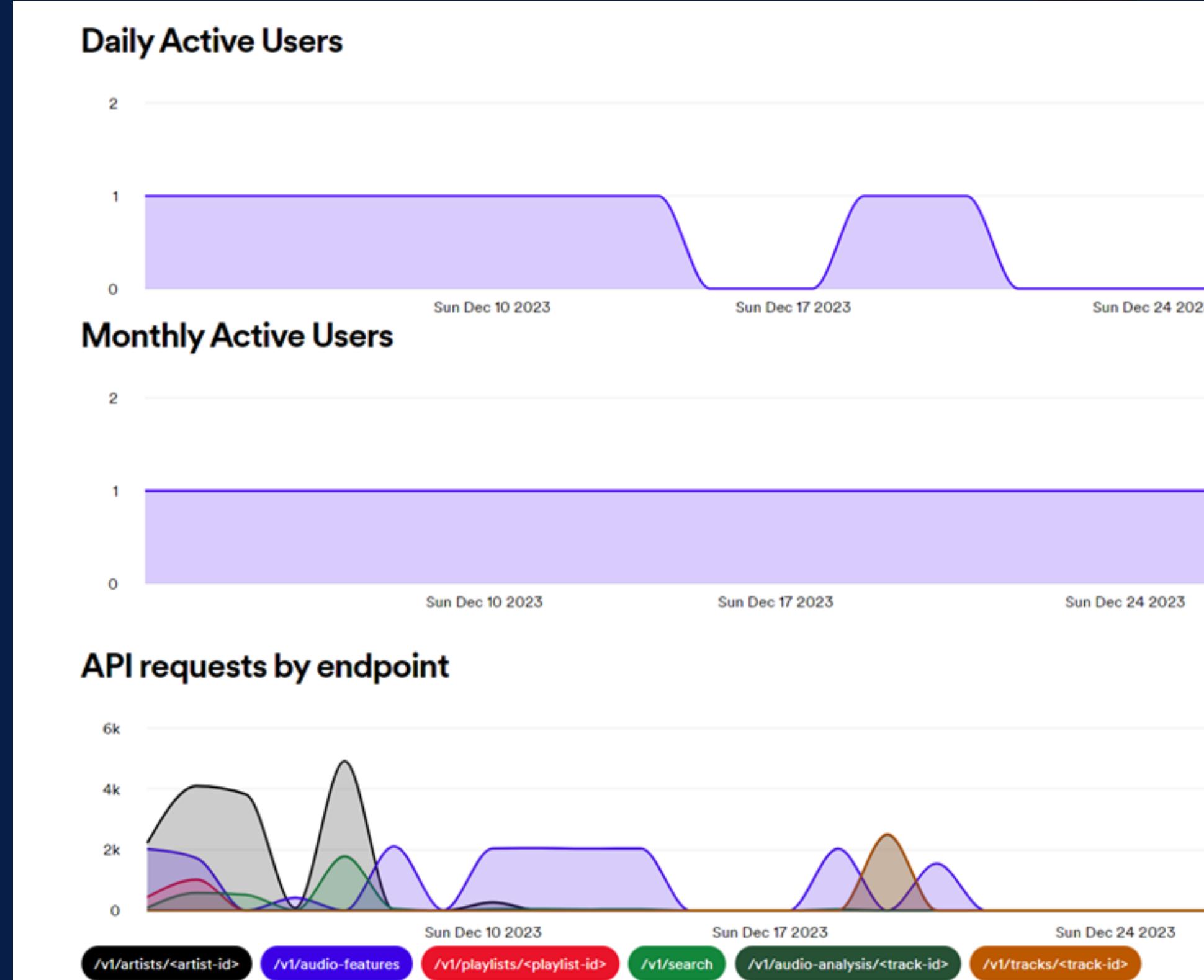
- 어떤 음악과 아티스트를 선호하는지 확인하고 싶다.
- 선호하는 곡과 유사한 곡들을 추천받고 싶다.
- 나와 비슷한 음악적 성향을 가진 사람들이 듣는 곡을 알고 싶다.

주요 서비스

노래 추천 서비스

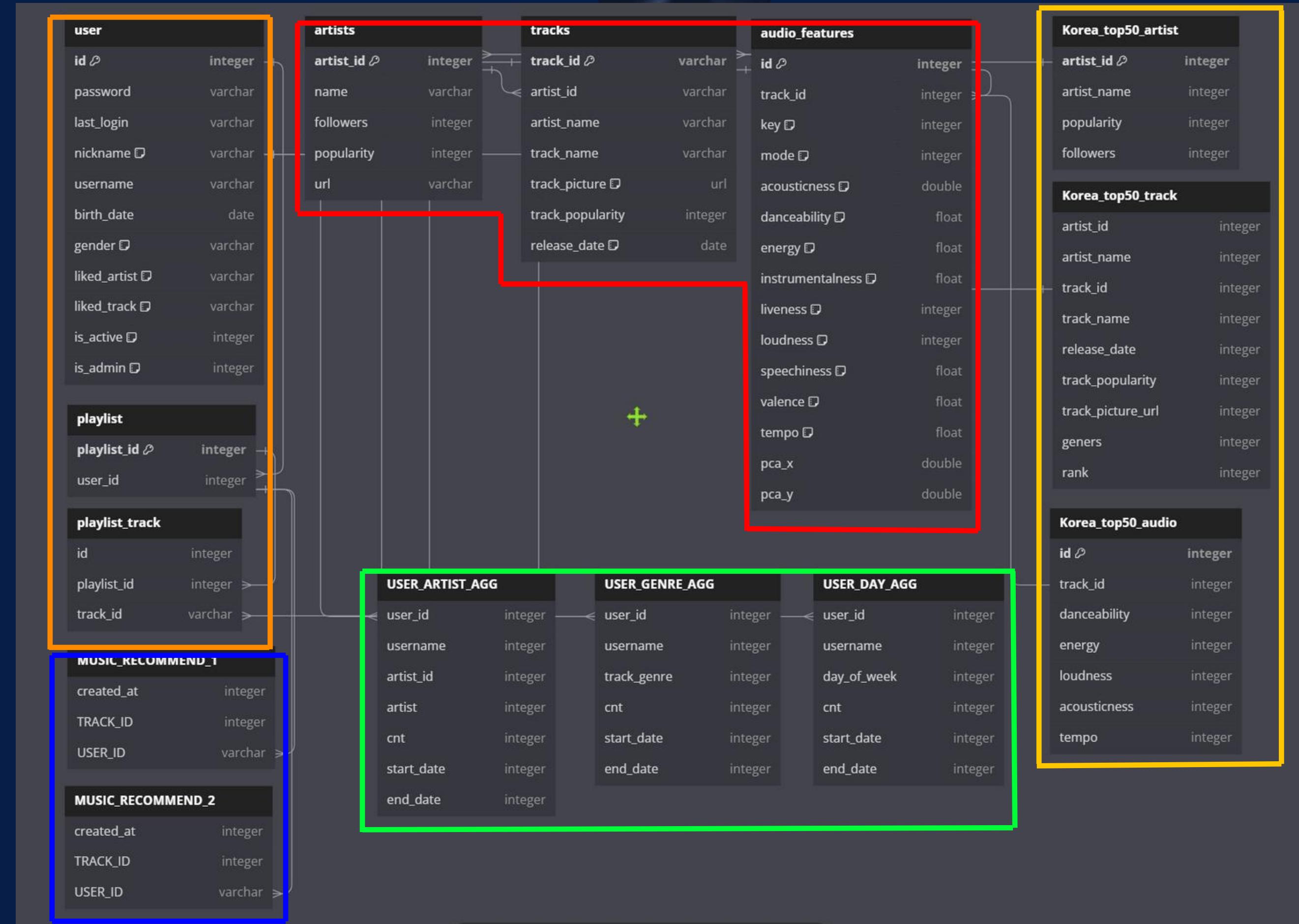
4.5Hz 이용 정보 제공 서비스

데이터 수집 방법



- 스포티파이 API
- 수집 장르 : 총 10개
- 'acoustic', 'dance', 'edm', 'hip-hop', 'pop', 'r-n-b', 'rock', 'romance', 'soul', 'k-pop'
- 누락 데이터 확인
- 중복 데이터 제거
- Audio features : tempo의 이상치 제거
- Artist : 5735
- Track : 56875
- Audio features : 56875
- 유저 : 10000명 생성
- 플레이리스트 : 10000개 생성

ERD

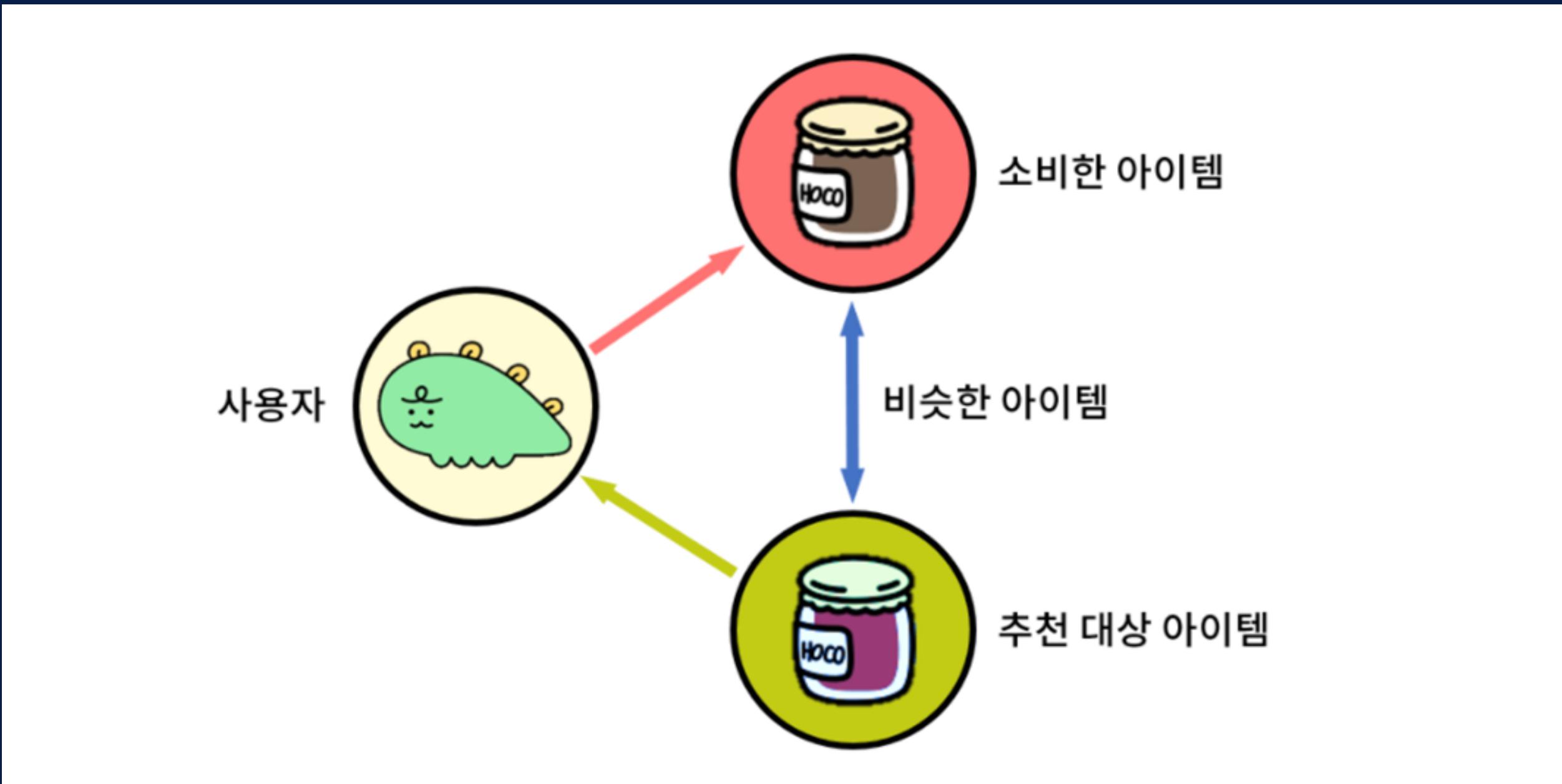




추천 모델링

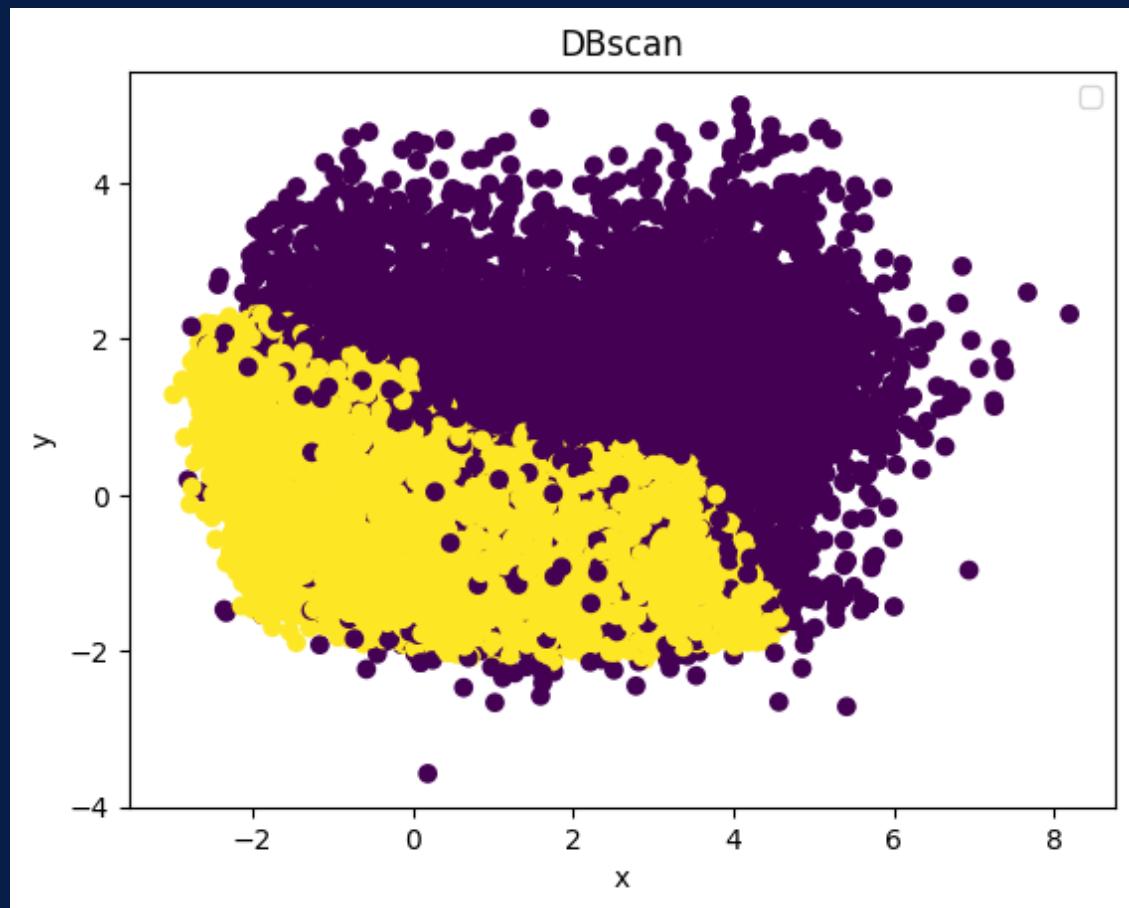
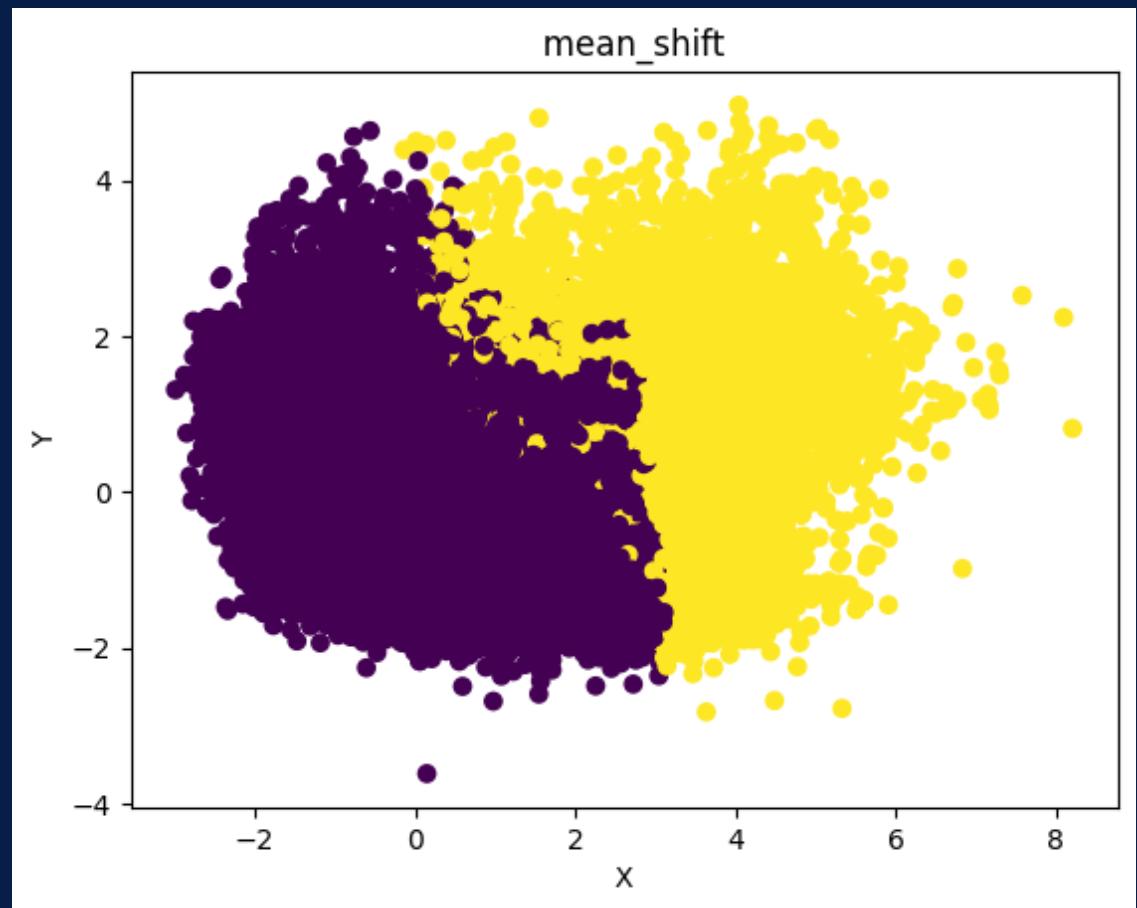
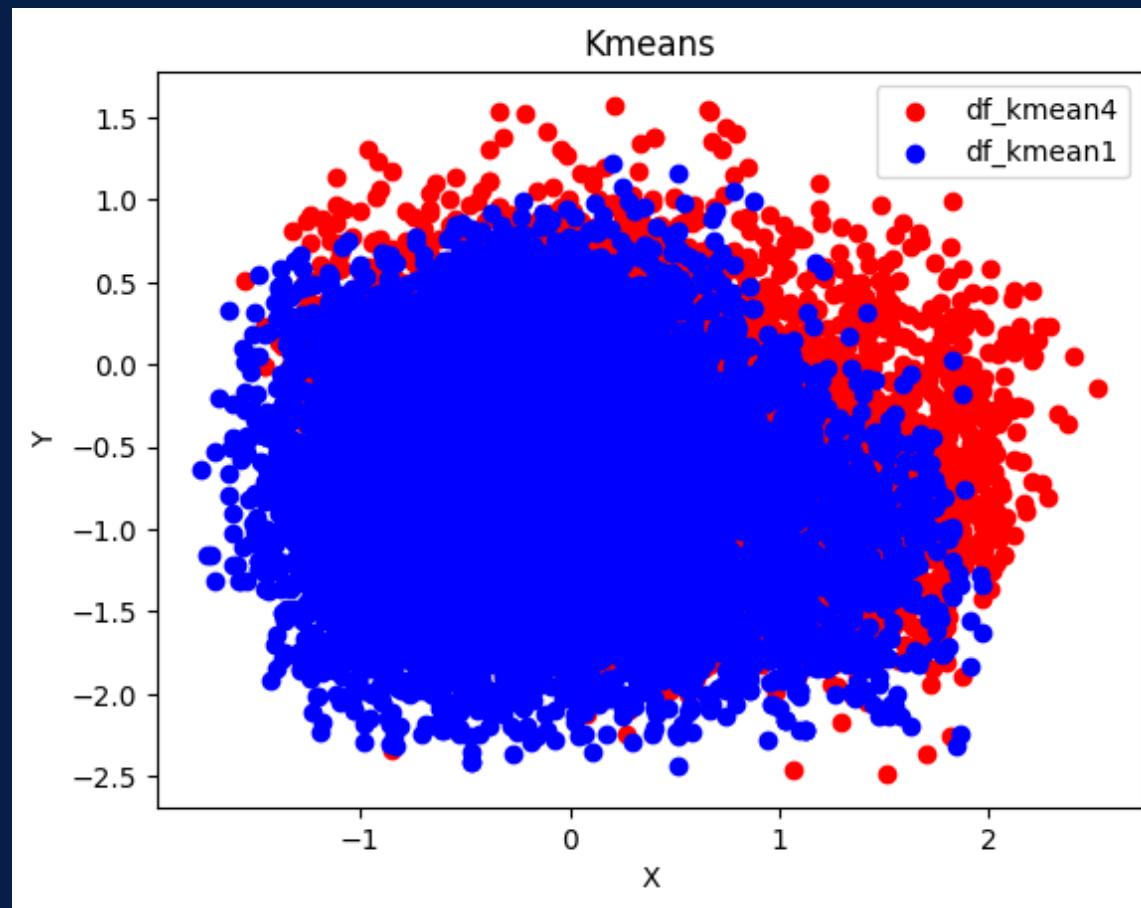
콘텐츠 기반 필터링

사용자가 소비한 아이템에 대해 아이템 내용이 비슷하거나 특별한 관계가 있는 다른 아이템을 추천하는 방법



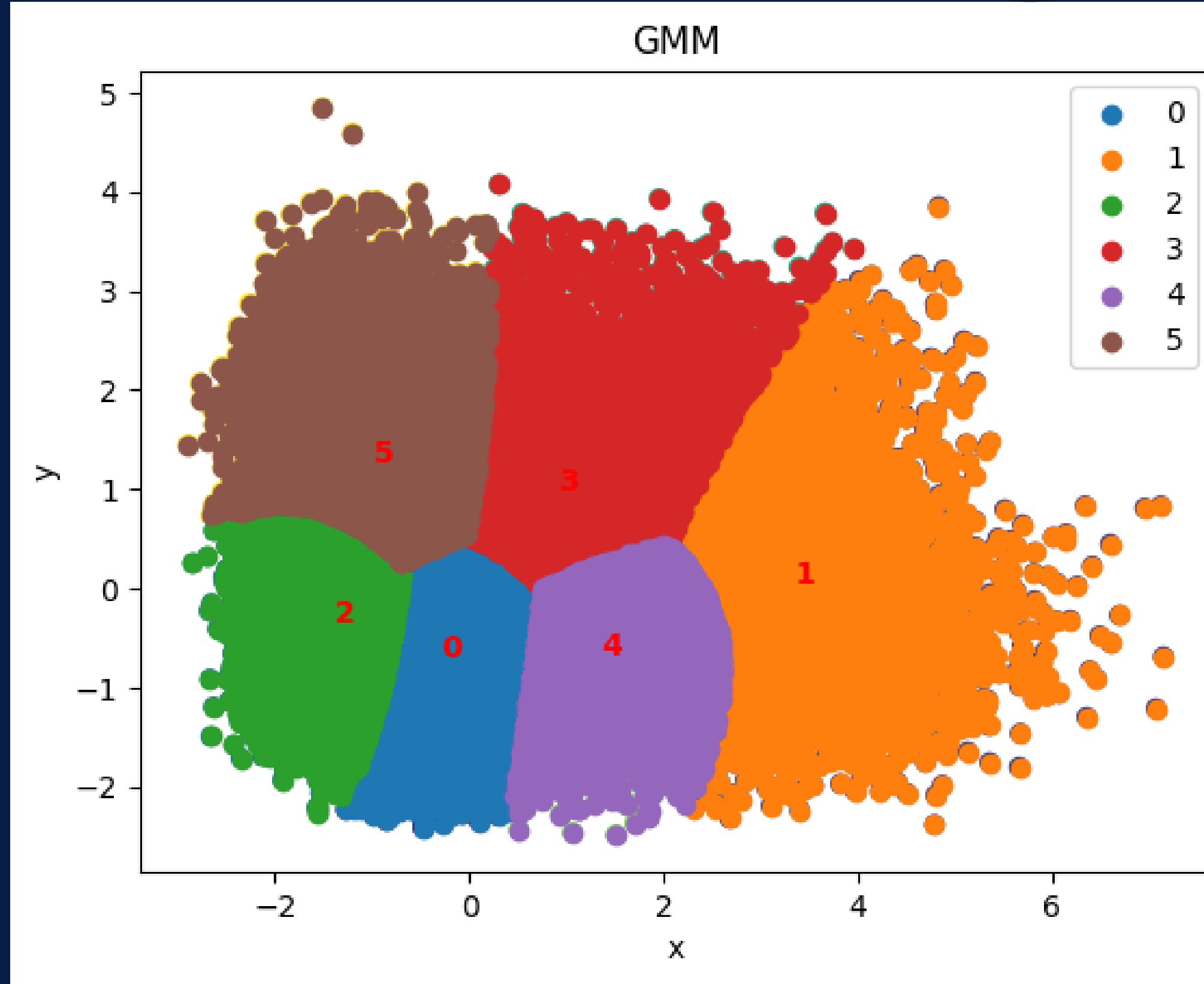
군집화 알고리즘 : Kmeans, Mean-shift, DbSCAN, GMM
코사인 유사도

일부 군집화 알고리즘의 한계



- Kmeans: 데이터 간 침범, 모호한 군집 경계
- Mean-Shift, DBscan: 군집 개수 지정 불가, 모호한 군집 경계

GMM

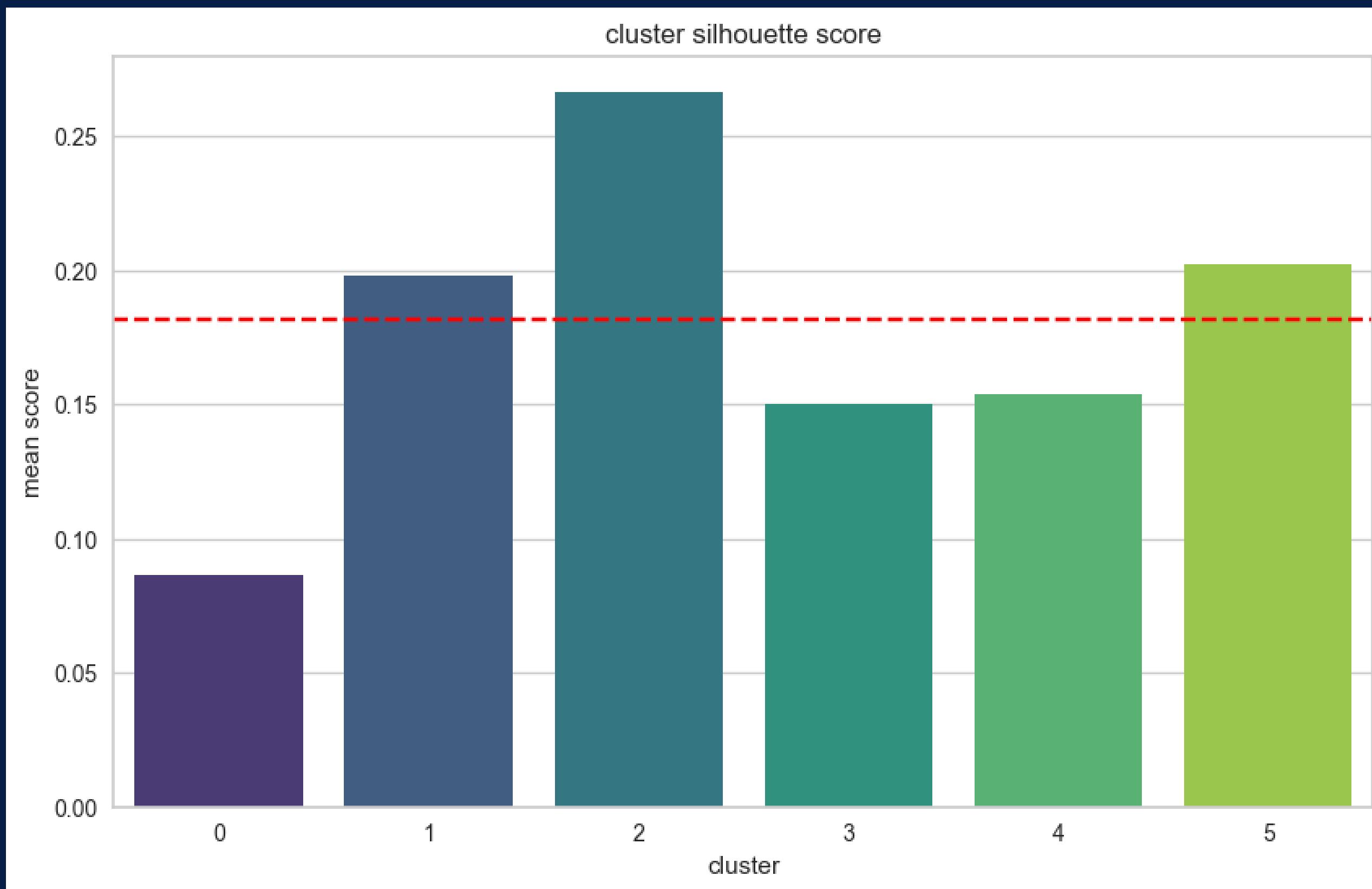


가장 뚜렷하고 명확한
군집화

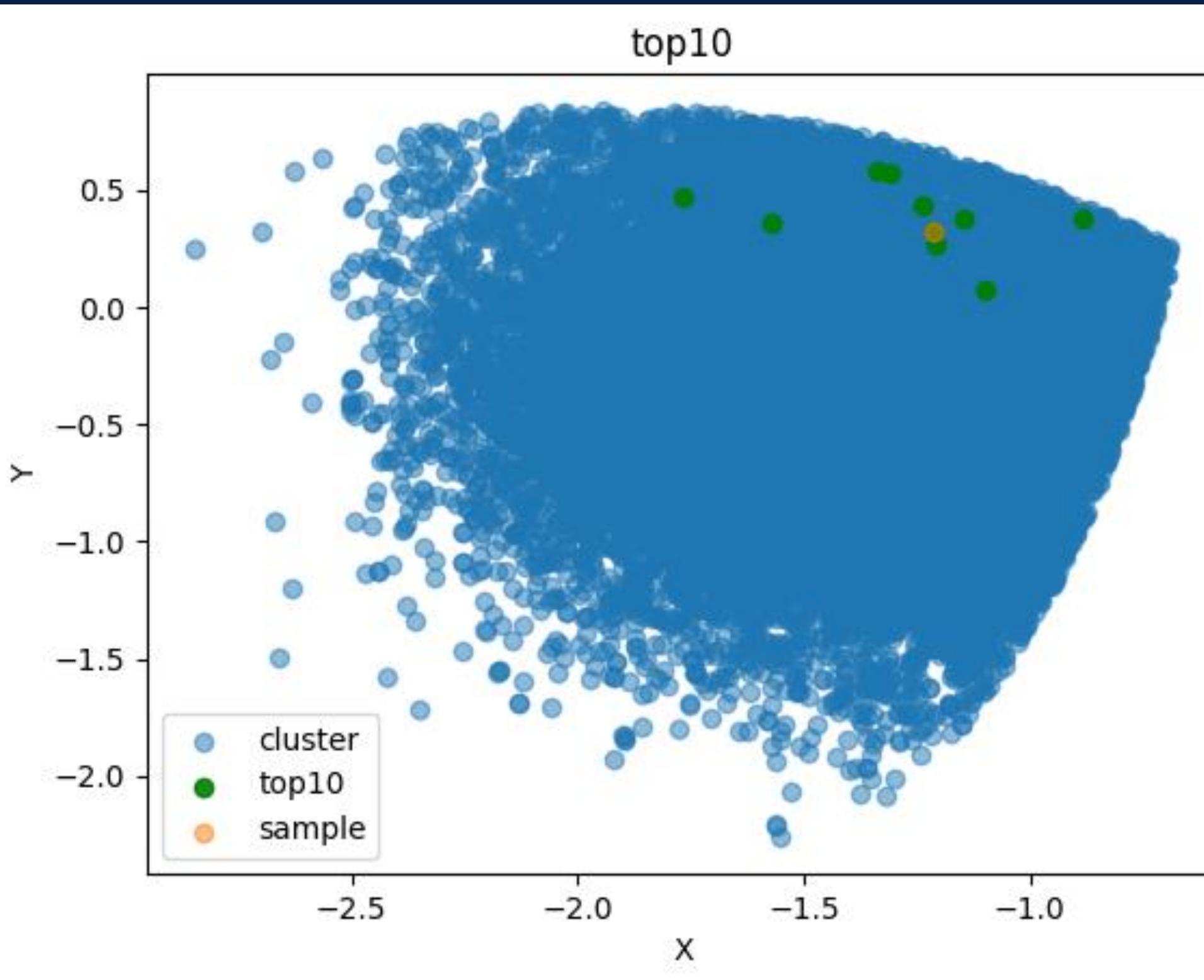
"군집화 모델로 선정"



cluster silhouette score



콘텐츠 기반 필터링



사용자의 최근 재생수가 가장 높은 2곡

첫번째 곡 군집에서
오디오 특성의
코사인 유사도가
높은 곡 15곡 선정

두번째 곡 군집에서
오디오 특성의
코사인 유사도가
높은 곡 15곡 선정

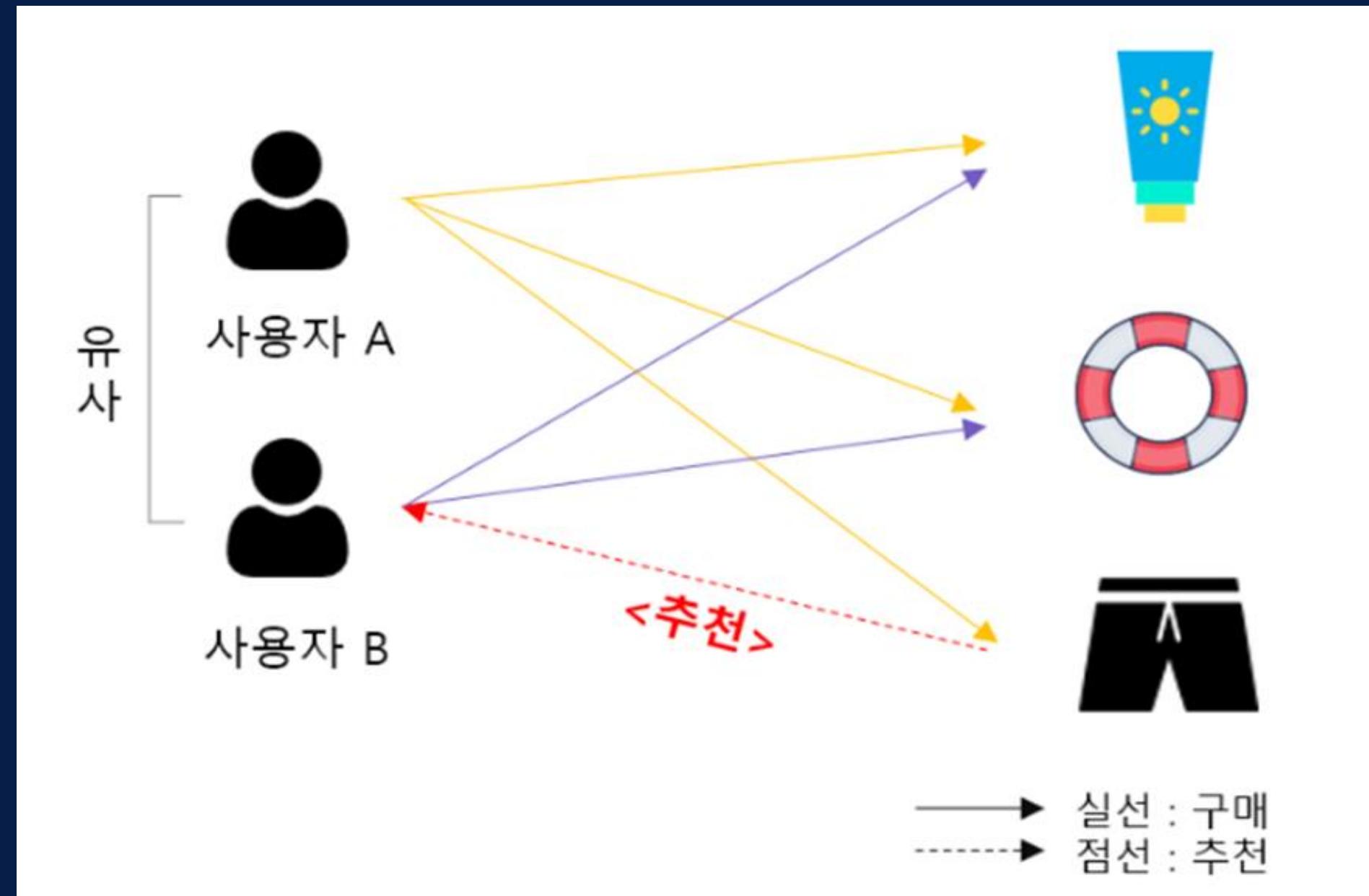
총 30곡에서 랜덤으로 4개 추천



협업 필터링

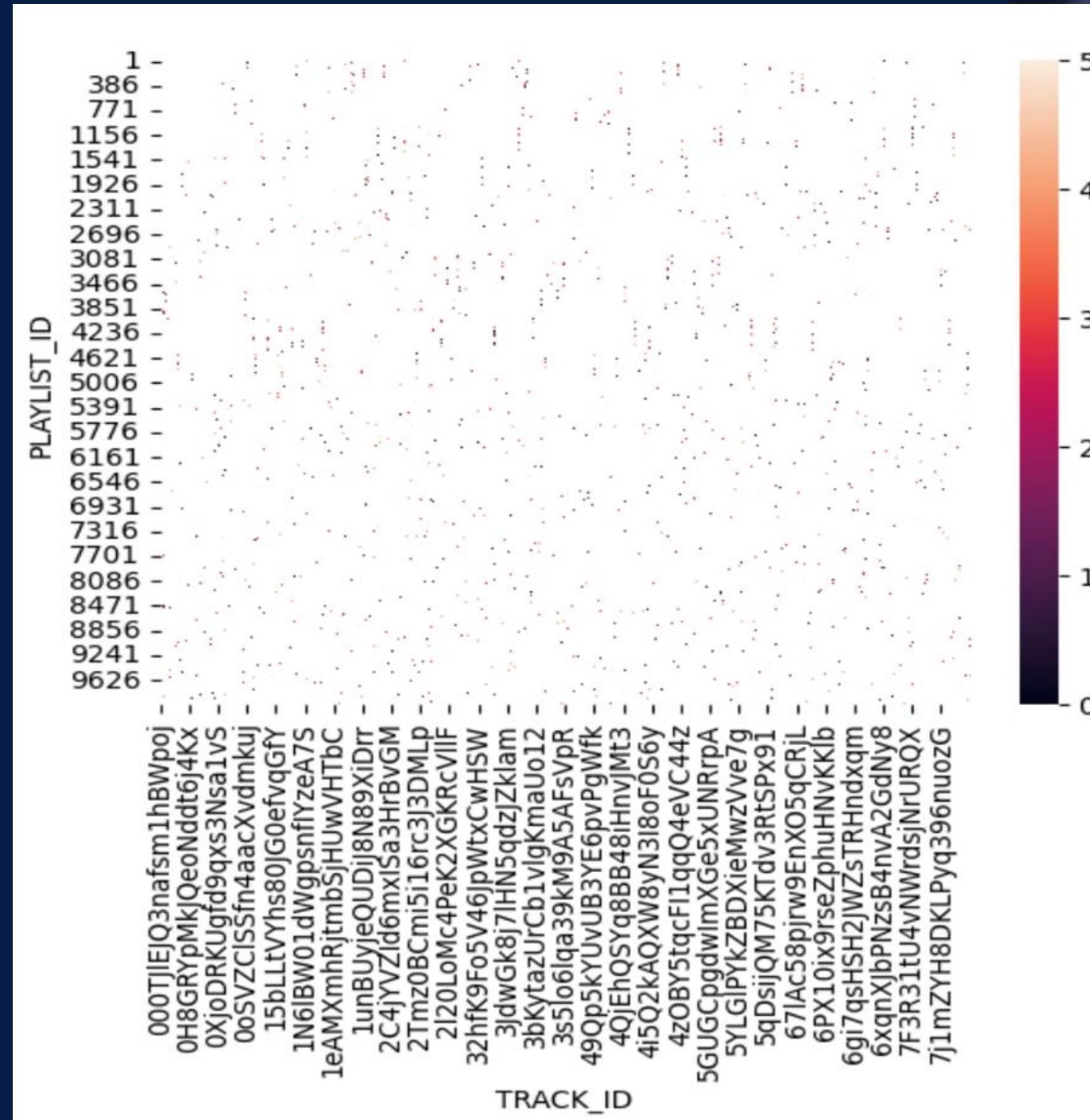
협업 필터링

특정 집단에서 발생하는 유사한 사용행동을 파악하여 비슷한 취향의 사람들에게 아이템을 추천하는 것



Surprise 라이브러리 SVD(Singular Value Decomposition) 모델
Matrix Factorization과 Deep Learning을 결합한 Neural Collaborative Filtering(NCF) 방식

협업 필터링을 위한 플레이리스트 생성



Cold-start problem

유저 : 약 10000명

플레이리스트 : 10000개 이상

플레이리스트 수록곡 : 500~2000곡

노래 평점 : 0 ~ 5점

SVD 모델

```
# rating 점수에 따라 평점 범위 0 ~ 5로 지정
reader = Reader(rating_scale=(0, 5))
# surprise 라이브러리가 사용할 수 있는 형태로 변환
data = Dataset.load_from_df(df[['PLAYLIST_ID', 'TRACK_ID', 'rating']], reader)
# 4. 모델 학습
model = SVD()
# cross_validate 함수를 이용해서 데이터를 5개로 구분하고 각각의 부분을 한번씩 테스트 세트로 사용해서 모델을 학습 및 평가
cross_validate(model, data, measures=['RMSE', 'MAE'], cv=5, verbose=True)

# 5. 모델 테스트
trainset, testset = train_test_split(data, test_size=.25)
model.fit(trainset)
predictions = model.test(testset)
accuracy.rmse(predictions)

-----  

Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean      Std
RMSE (testset)  1.6438  1.6435  1.6444  1.6436  1.6422  1.6435  0.0007
MAE (testset)   1.3731  1.3732  1.3744  1.3739  1.3721  1.3733  0.0008
Fit time        52.66   55.87   56.69   56.73   60.32   56.46   2.44
Test time        13.14   13.24   13.19   12.52   13.44   13.10   0.31
RMSE: 1.6364
1.636401097071551
-----  

Best params: {'lr_all': 0.001, 'n_epochs': 5, 'reg_all': 0.4}, Best score: 1.5019
RMSE: 1.5022
```

SVD 모델

나의 플레이리스트

Playlist #5

나를 위한 새로운 발견

Colorado
Milky Chance

Tonight (Mousse T. Remix)
Simply Red

Hazard
Georg Erixon

Tunechi Rollin'
Lil Wayne

Todo Menso
MC Davo

4.5Hz차트

실시간 2024-01-03 22:00:00 기준

시간대 선택

순위 가수명 제목 재생 start end

14 IM HERO If We Ever Meet Again Lim Young Woong

15 wherever u r (ft. V of BTS) UMI, V

16 GOLDEN Yes or No Jung Kook

17 Face-off Jimin

18 Alone Jimin

19 Slow Dancing V

20 Drama aespa

21 IM HERO Rainbow Lim Young Woong

22 STANDING YO U Standing Next to You (Usher Remix) Jung Kook, USHER

23 IM HERO Our Blues, Our Life Lim Young Woong

24 Interlude : Dive Jimin

25 To. X TAEYEON

26 IM HERO A bientot Lim Young Woong

27 Super Shy NewJeans

28 GOLDEN Hate You Jung Kook

29 IM HERO Rainy Days V

최종 5개 추천 음악:

Track ID: 4t7pE7Au1BciptoSBOQqLU, 예상 평점: 2.8804491260250558

Track ID: 6iFtVBq76rk1H2tkT6oESe, 예상 평점: 2.885936377037126

Track ID: 4eXFdrCO61JdnhKzf0z1n7, 예상 평점: 2.8657920936670696

Track ID: 4hEOBub4I6EGkIKCxrELgb, 예상 평점: 2.852668832747606

Track ID: 7up1kromWJY9ZoR747ngVV, 예상 평점: 2.8757844345372137

NCF·모델

```
# 필요한 컬럼만 사용
df = df[['USER_ID', 'TRACK_ID', 'rating', 'birth_date', 'gender']]

# 'birth_date'를 10대, 20대, 30대, 40대, 50대로 나이대를 나눔
df['birth_date'] = pd.to_numeric(df['birth_date'], errors='coerce')
df['age'] = (2023 - df['birth_date']) // 10
-----
# 모델 구성
user_input = Input(shape=(1,), name='user_input', dtype='int64')
user_embedding = Embedding(num_users, 10, name='user_embedding')(user_input)
user_vec = Flatten(name='flatten_users')(user_embedding)

track_input = Input(shape=(1,), name='track_input', dtype='int64')
track_embedding = Embedding(num_tracks, 10, name='track_embedding')(track_input)
track_vec = Flatten(name='flatten_tracks')(track_embedding)

age_input = Input(shape=(1,), name='age_input', dtype='int64')
age_embedding = Embedding(df.age.unique(), 5, name='age_embedding')(age_input)
age_vec = Flatten(name='flatten_ages')(age_embedding)

gender_input = Input(shape=(1,), name='gender_input', dtype='int64')
gender_embedding = Embedding(df.gender.unique(), 5, name='gender_embedding')(gender_input)
gender_vec = Flatten(name='flatten_genders')(gender_embedding)
-----
(RMSE) : 1.5053451900813966
(MAE): 1.3000865707350726
```

NCF·모델

```
Trial 10 Complete [01h 10m 23s]
val_loss: 2.2514332135518393
Best val_loss So Far: 2.2513275146484375
Total elapsed time: 11h 47m 56s
Epoch 1/5 99990/99990 [=====] - 280s 3ms/step - loss: 2.2529 - val_loss: 2.2514
Epoch 2/5 99990/99990 [=====] - 278s 3ms/step - loss: 2.2487 - val_loss: 2.2513
Epoch 3/5 99990/99990 [=====] - 278s 3ms/step - loss: 2.2487 - val_loss: 2.2514
Epoch 4/5 99990/99990 [=====] - 278s 3ms/step - loss: 2.2486 - val_loss: 2.2519
Epoch 5/5 99990/99990 [=====] - 278s 3ms/step - loss: 2.2486 - val_loss: 2.2517
테스트 손실: 2.250011920928955
-----
(RMSE): 1.5000104877109977
(MAE): 1.300105970542733
```

음악 검색

검색어 입력 검색

장르 모아보기

장르 고르기

30대 남성에게 가장 핫한 노래

-  **Softly**
Karan Aujla
-  **Hurricane (feat. SHIBUI)**
Blasterjaxx
-  **Lune - Mindchatter Remix**
Kasbo
-  **Nu Wave**
Boldy James
-  **AT THE PARTY (feat. Pharrell Williams & Travis Scott)**
Kid Cudi

나의 플레이리스트

 Playlist #5

Top50 차트

현재 가장 많이 재생된 트랙의 일일 업데이트

순위	제목	가수명
1	Like Crazy	Jimin
2	Like Crazy (English Version)	Jimin
3	Seven (feat. Latto) (Explicit Ver.)	Jung Kook, Latto
4	Closer Than This	Jimin
5	Set Me Free Pt.2	Jimin
6	3D (feat. Jack Harlow)	Jung Kook, Jack Harlow
7	Standing Next to You	Jung Kook
8	Do or Die	Lim Young Woong
9	Grain of Sand	Lim Young Woong
10	Love Me Again	V
11	London Boy	Lim Young Woong
12	Polaroid	Lim Young Woong

최종 5개 추천 음악:

Track ID: 4L8pKwvlbzwHKwgQM4Mz, 예상 평점: 2.542133331298828
 Track ID: 62D6bjz4gTKDJRLPD3gJY, 예상 평점: 2.529088258743286
 Track ID: 1ODqkQQkoP8kyXSpwAS60, 예상 평점: 2.5073540210723877
 Track ID: 1zKEyeeSWHOyd9csMC1WTO, 예상 평점: 2.5073540210723877
 Track ID: 7LJZyNjN5U0LdYSBQ9h1Lz, 예상 평점: 2.5073540210723877



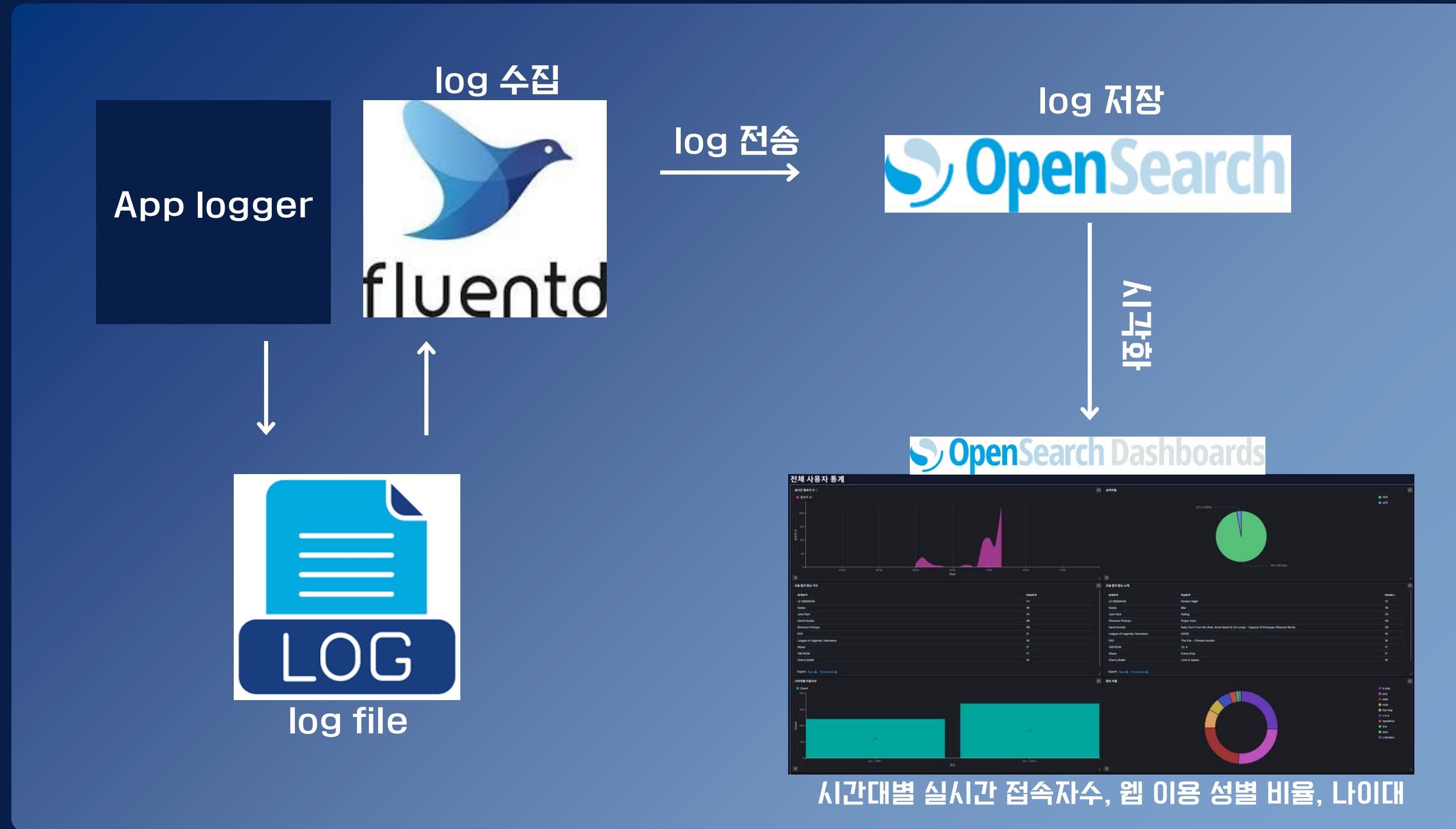
4.5Hz 이용 정보
집과 서비스

집계를 위한 logger 설계

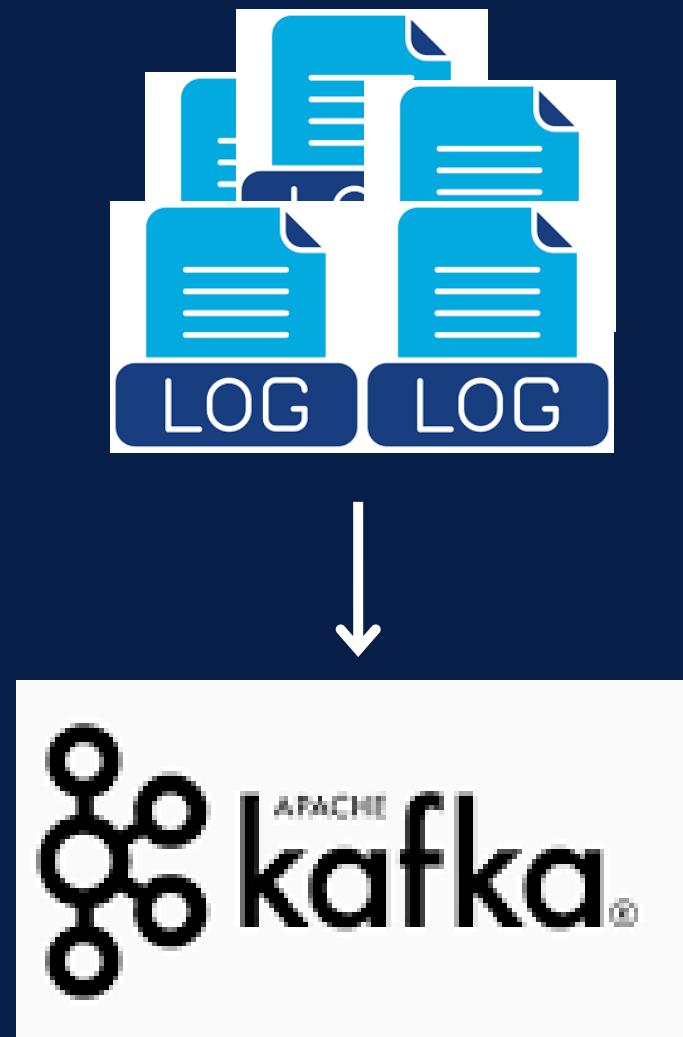
logger 설정

log file

전체집계



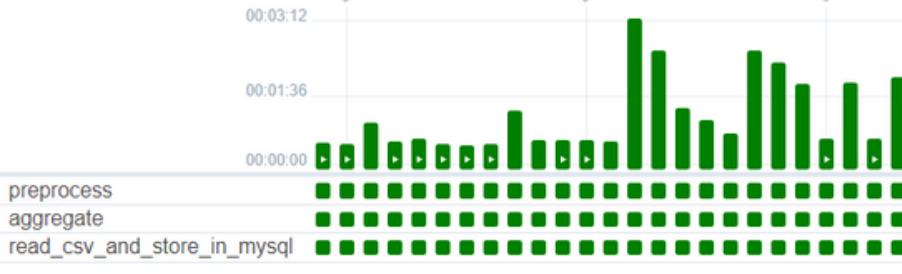
시간대별 이용 집계



순위	가수명	제목	재생
1	Mike Perry	Closer	45
2	K-391	Play - Niya & Alan Walker Remix	40
3	Gryffin	Nobody Compares To You (feat. Katie Pearlman)	36
4	MEDUZA	Playground (from the series Arcane League of Legends) - MEDUZA Remix	12
5	SHAUN	Way Back Home	11
6	Zedd	Clarity	10

• 매일 9시 30분 스케줄링

개인 집계-데이터



개인 집계 - 시각화

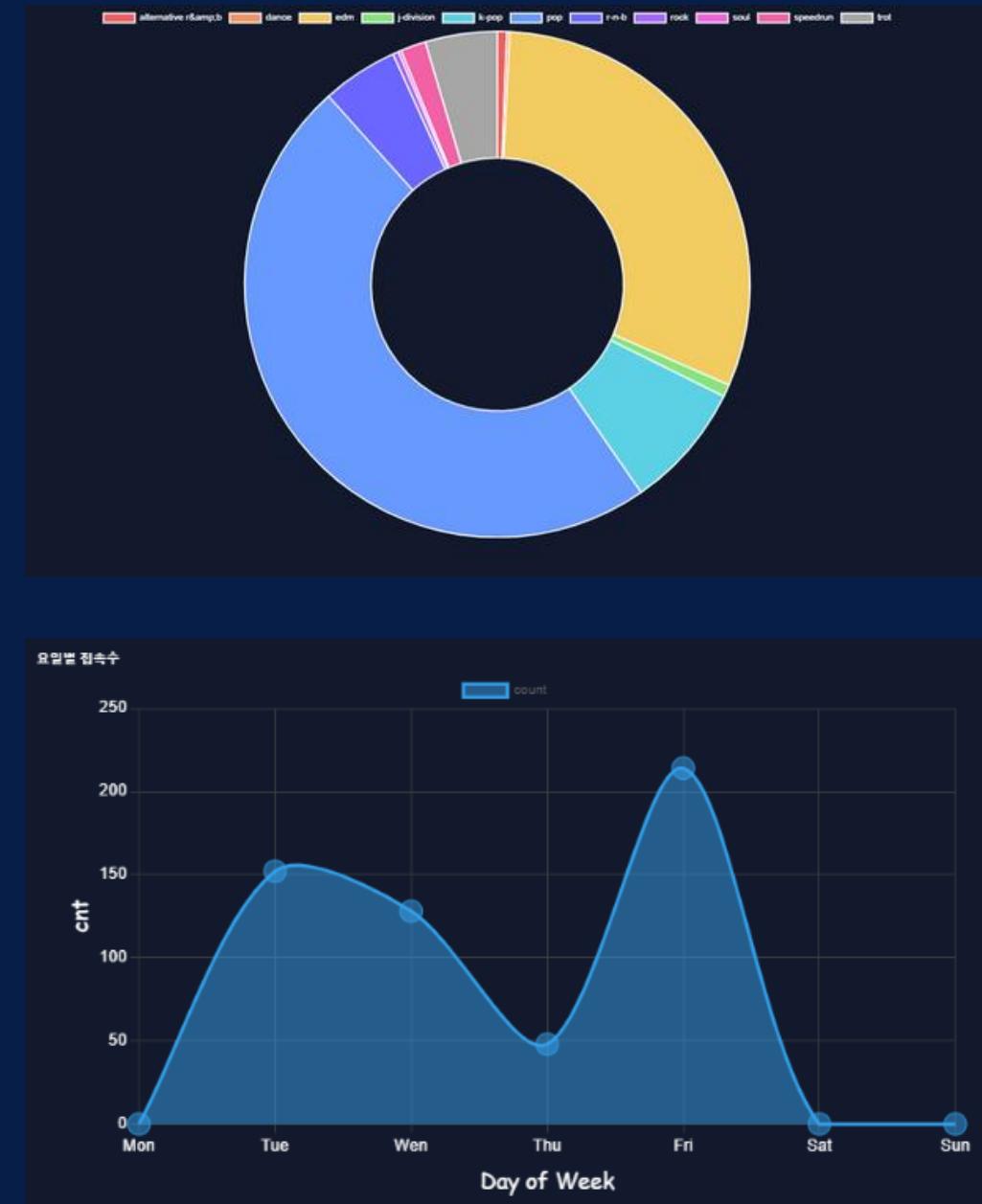
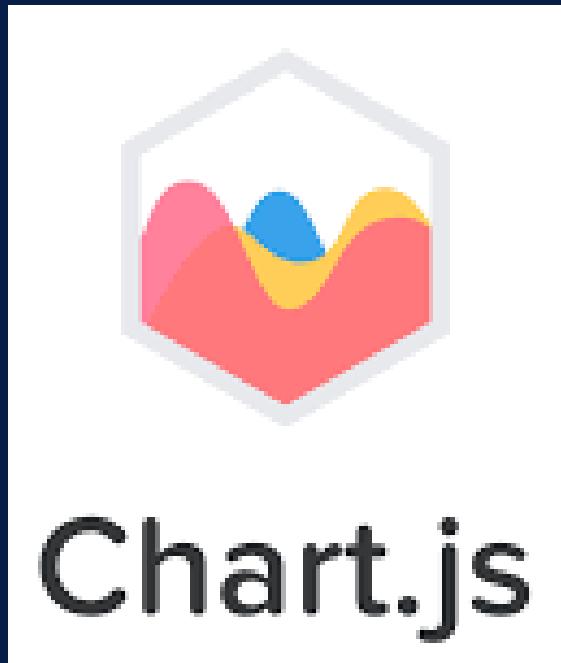


Chart.js는 8개의 차트 유형을 지원하는 데이터 시각화를 위한 자유의 오픈 소스 자바스크립트 라이브러리



4.5Hz 웹 구성

홈 화면

로그인 안 한 경우



로그인 한 경우



회원가입

nickname

이름

생년월일

성별

비밀번호

다른 개인 정보와 유사한 비밀번호는 사용할 수 없습니다.
비밀번호는 최소 8자 이상이어야 합니다.
통상적으로 자주 사용되는 비밀번호는 사용할 수 없습니다.
숫자로만 이루어진 비밀번호는 사용할 수 없습니다.

비밀번호 확인

확인을 위해 이전과 동일한 비밀번호를 입력하세요.

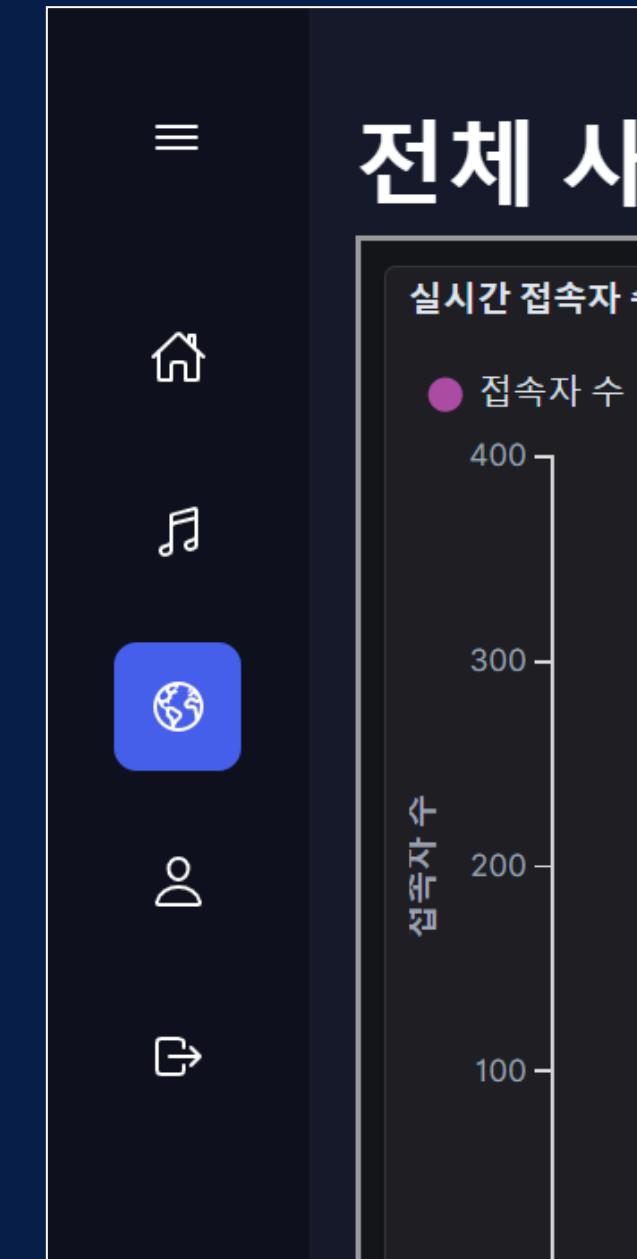
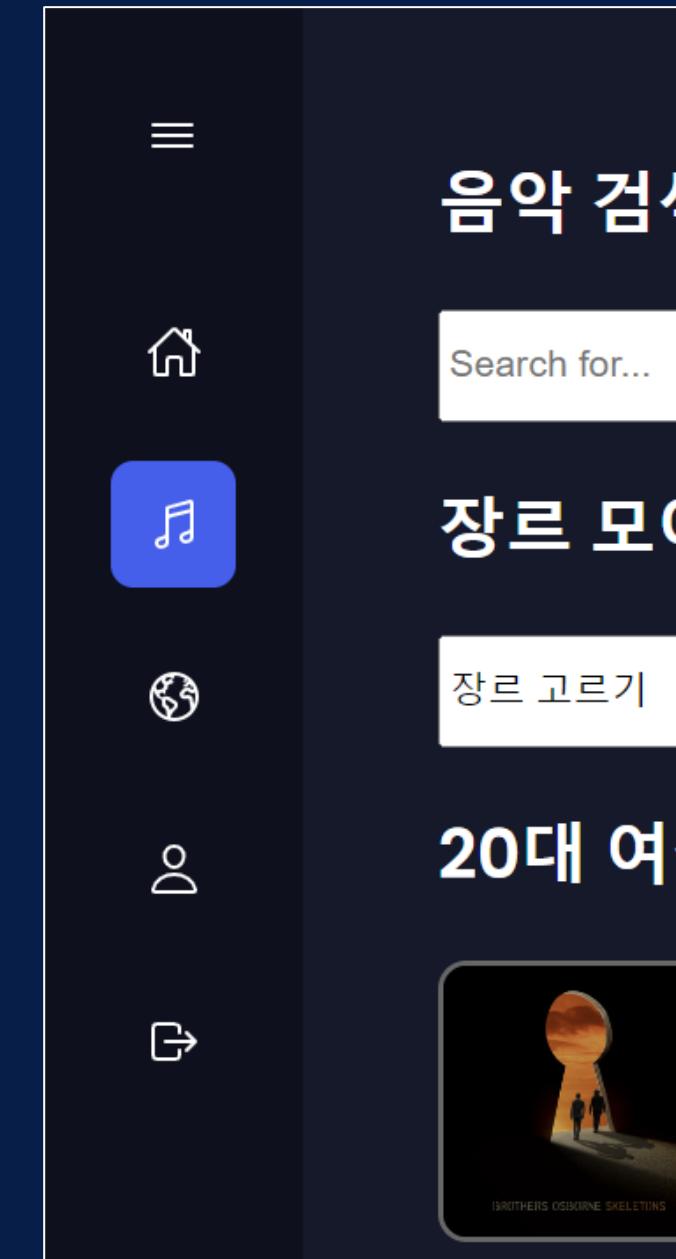
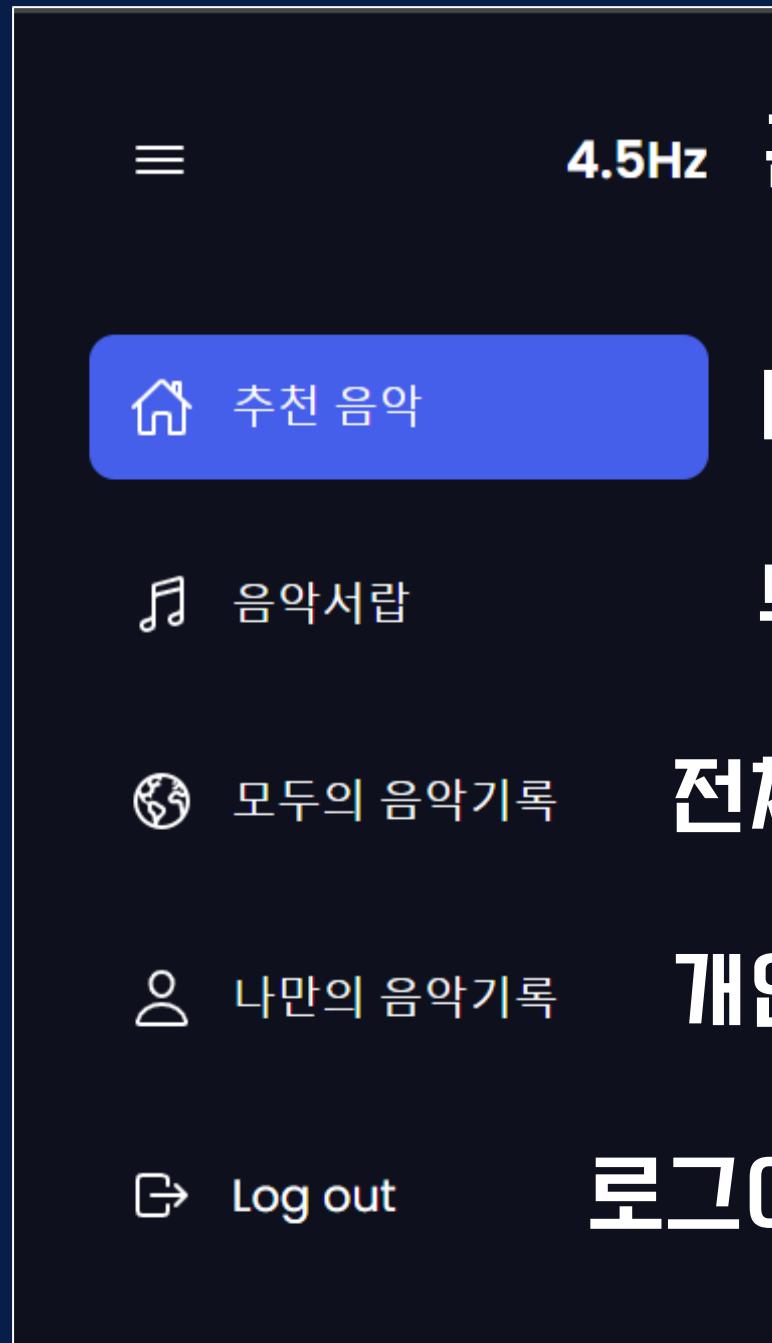
[제출](#)

로그인

Nickname

비밀번호

[제출](#)



현재 접속 페이지 확인

메인 페이지

사용자가 많이 들은 곡

Top Tracks		
순위	제목	Stream
1	The Eve	104
2	Eté 90	76

콘텐츠 기반 필터링을 사용해
사용자가 제일 많은 들은 2곡과
코사인 유사도가 높은 곡 랜덤 추천

최세인님에게 추천하는 노래



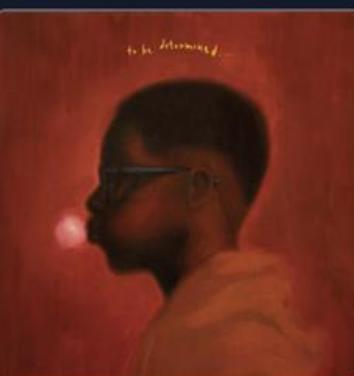
YUMEDONO

Kenji Omura



Amen (feat. Teezo
Touchdown)

Drake



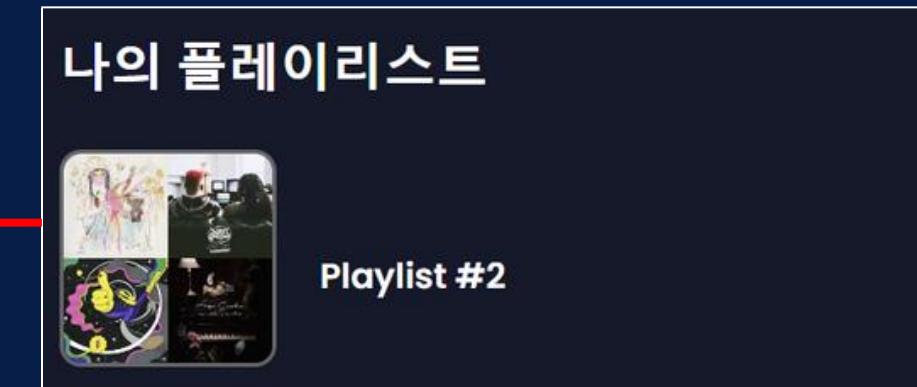
Believe
The Main Squeeze



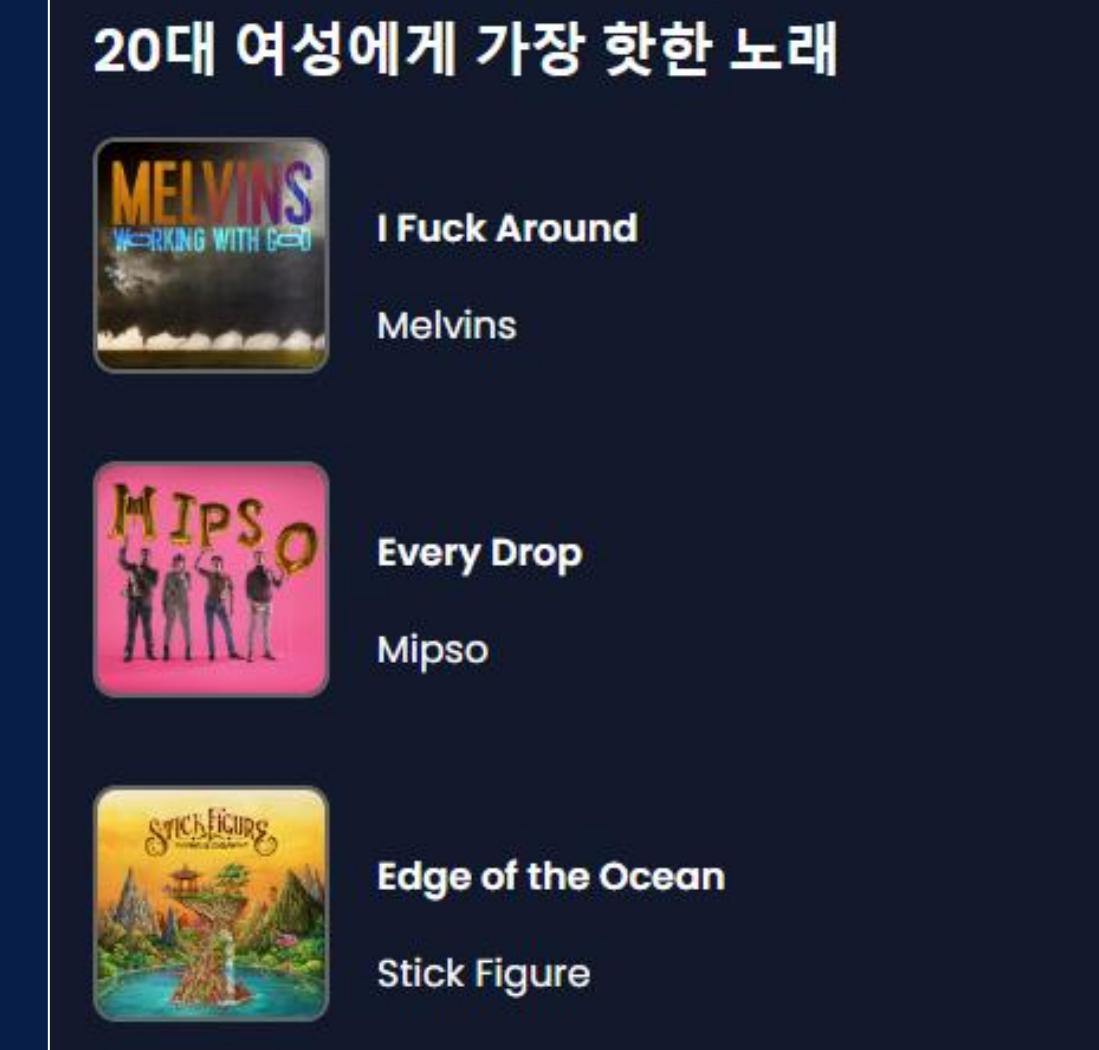
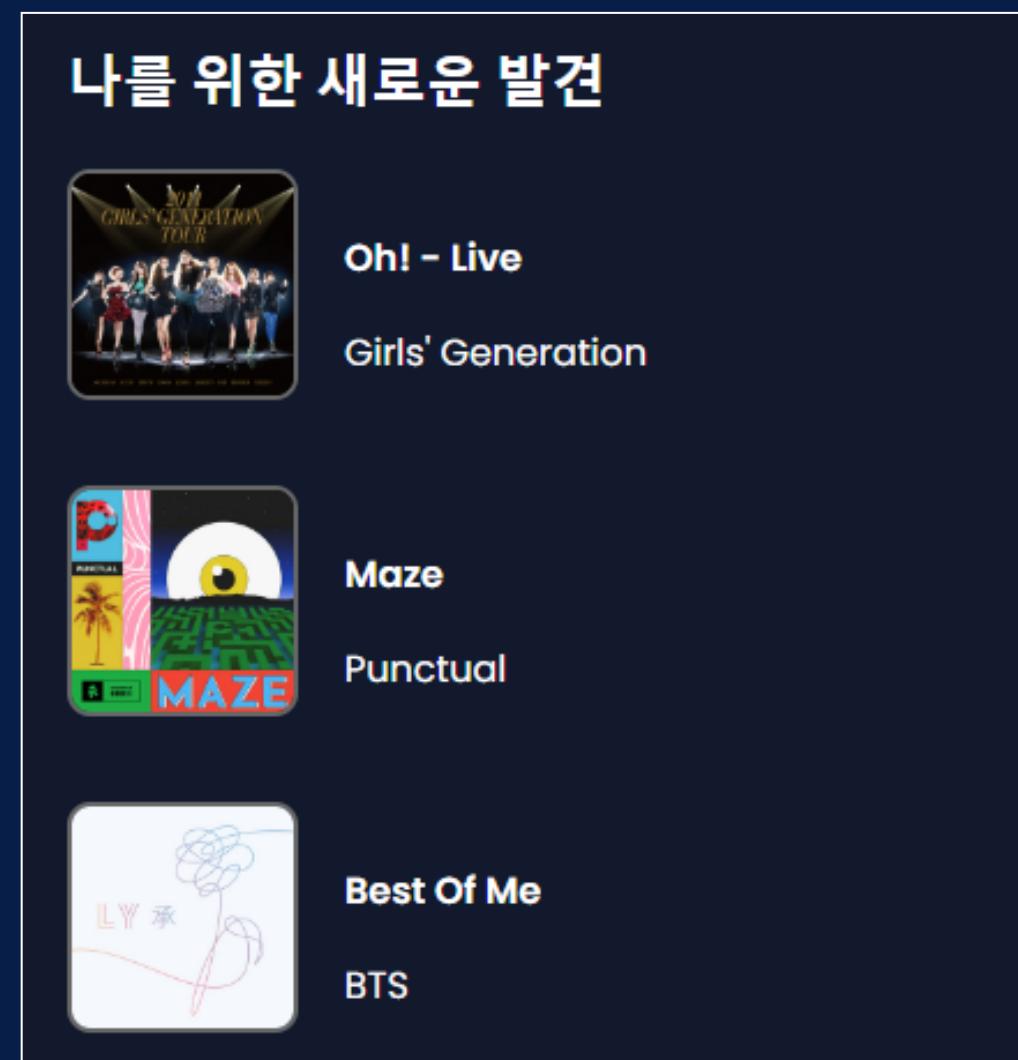
Night Away
TAEMIN

뮤직 페이지 - 사용자 기반 추천

나의 플레이리스트와
유사한 플레이리스트 내에서
내가 안 들어본 곡 추천



사용자와 같은 연령대, 성별을 가진
사람들이 많이 듣는 노래 추천



뮤직 페어 - 검색

음악 검색

검색어 입력 검색

→

"iu"에 대한 검색 결과

iu...im
IU
marshmallow

iu...im
IU
Tear drops in the morning

iu...im
IU
Last night story

iu...im
IU

iu...im
IU

iu...im
IU

"iu eight"에 대한 검색 결과

iu...im
IU
eight(Prod.&Feat. SUGA of BTS)

장르 모아보기

장르 고르기 검색

장르 고르기

- Acoustic
- Dance
- EDM
- Hip-hop
- K-pop
- POP
- R&B
- Rock
- Romance
- Soul

장르 모아보기

핫한 노래

Good Ones Are

Osborne

→

"hip-hop" 모아보기

LEMONHEAD (feat. 42 Dugg)
Tyler, The Creator

Hot Shower
Chance the Rapper

Stay Away (feat. Machine Gun Kelly & Goody Grace)
MOD SUN

Low Down
Lil Baby

Look Over My Shoulder
Kendrick Lamar

Busta Rhymes

곡 상세페이지



Every Drop
Mipso



Shake Your Ass-Smoke Your Grass
Rob Zombie



This Town's Been Too Good To Us
Dylan Scott

logger 실행

User 정보,
노래 정보,
가수 정보,
클릭 시간



Every Drop
Mipso
발매일 : 2021년 10월 29일
인기도 : 22
음악 장르 : r-n-b
별점 선택하기
★ ★ ★ ★ ★
플레이리스트에 추가

곡 상세페이지

플레이리스트에 없는 경우



127.0.0.1:8000 내용:
플레이리스트에 추가되었습니다.

확인

Every Drop

Mipso

발매일 : 2021년 10월 29일

인기도 : 22

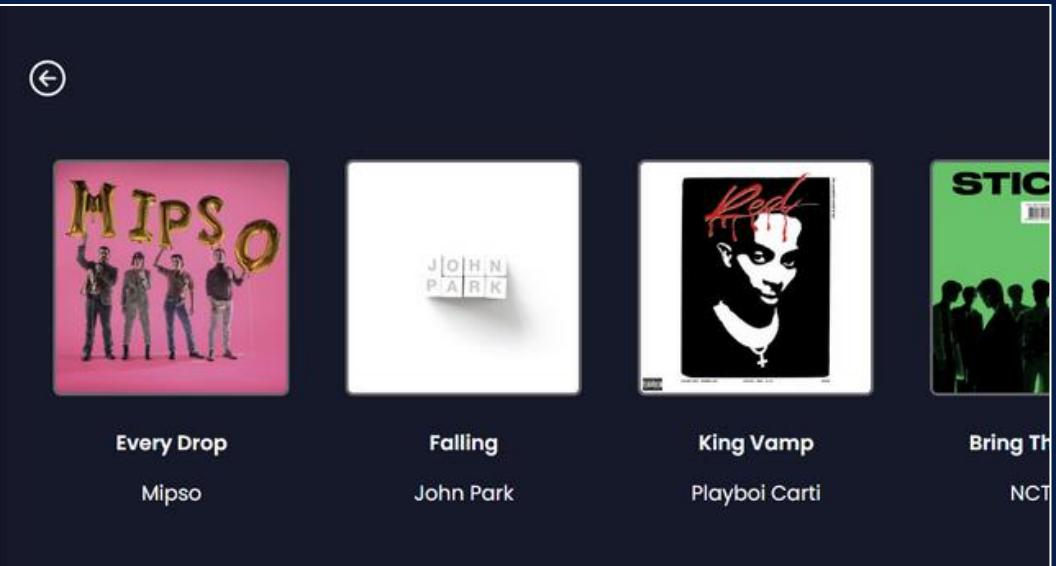
음악 장르 : r-n-b

별점 선택하기

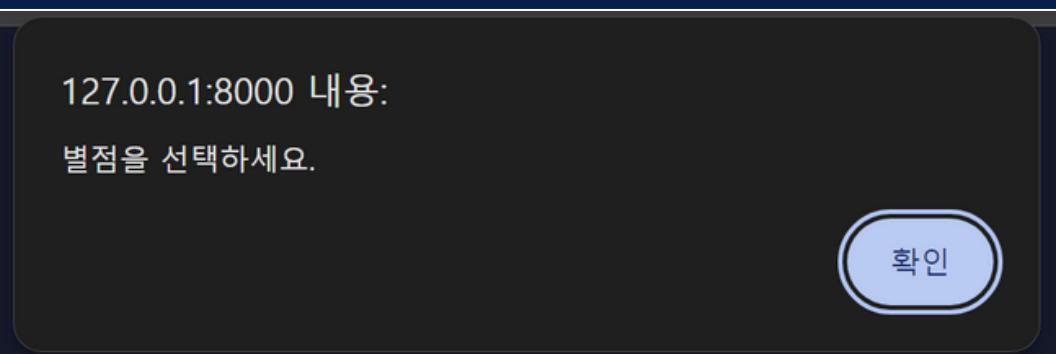
★★★★★

플레이리스트에 추가

평점 선택 후 버튼 클릭 시
평점과 함께 플레이리스트에 추가



평점 선택 안 할 시 :



곡 상세페이지

플레이리스트에 있는 경우



127.0.0.1:8000 내용:
플레이리스트에서 제거되었습니다.

확인

Every Drop

Mipso

발매일 : 2021년 10월 29일

인기도 : 22

음악 장르 : r-n-b

별점 : 5

[플레이리스트에서 빼기](#)

사용자가 지정한 평점 표시 &
버튼 클릭 시 플레이리스트에서 제거

↶



Falling
John Park

King Vamp
Playboi Carti

Bring The Noize
NCT 127

STICKER
NCT 127

TREASURE EP.
ATEEZ

Top50

Top 50 - 대한민국

#	제목	재생
1	Like Crazy	135,627
2	Like Crazy (English Version)	87,294
3	Seven (feat. Latto) (Explicit Ver.)	80,828
4	Closer Than This	78,614
5	Set Me Free Pt.2	63,663
6	3D (feat. Jack Harlow)	51,719

매일 업데이트 되는
대한민국 Top50 차트

Top50 차트

현재 가장 많이 재생된 트랙의 일일 업데이트

순위	제목	가수명
1	Like Crazy	Jimin
2	Like Crazy (English Version)	Jimin
3	Seven (feat. Latto) (Explicit Ver.)	Jung Kook, Latto
4	Closer Than This	Jimin
5	Set Me Free Pt.2	Jimin
6	3D (feat. Jack Harlow)	Jung Kook, Jack Harlow
7	Standing Next to You	Jung Kook
8	Do or Die	Lim Young Woong
9	Grain of Sand	Lim Young Woong
10	Love Me Again	V
11	London Boy	Lim Young Woong
12	Polaroid	Lim Young Woong
13	Perfect Night	LE SSERAFIM
14	If We Ever Meet Again	Lim Young Woong

- 매일 9시 30분 스케줄링

Top50

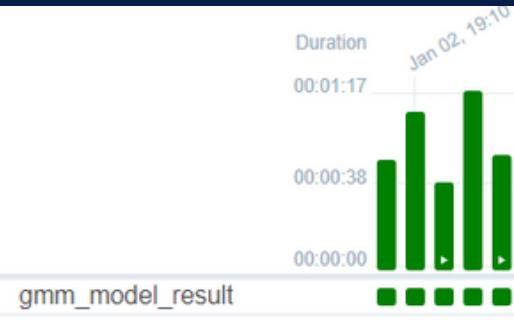


매일 업데이트를 위해 Airflow를 사용하여 스케줄링



Top50

Airflow를 사용하여 주기적으로 모델 재학습
이전 데이터와 신규 데이터를 통합해 군집화 후 DB에 갱신



```
mysql_hook = MySqlHook(mysql_conn_id='mysql_db') # MySQL Connection ID
connection = mysql_hook.get_conn()
cursor = connection.cursor()

sql_query = " SELECT id, danceability, energy, loudness, acousticness, tempo, track_id FROM AUDIO_FEATURES"

cursor.execute(sql_query)
results = cursor.fetchall()
```



DB에 있는 새로운 오디오 특성 수집



스케일링, 차원축소



GMM 모델 적용



DB 업데이트

```
for index, row in df_gmm.iterrows():
    update_query = f"""
        UPDATE AUDIO_FEATURES
        SET pca_x = {row['pca_x']}, pca_y = {row['pca_y']}, gmm_cluster = {row['gmm_cluster']}
        WHERE id = {row['id']}
    """
    try:
        cursor.execute(update_query)
        print(f"{index+1}")
    except Exception as e:
        raise

connection.commit()
```

뮤직 페이지 - 실시간 차트

4.5Hz차트

실시간 18:00 기준

순위	가수명	제목	재생
1	Zara Larsson	wow (feat. Sabrina Carpenter) - Remix	70
2	Glass Animals	Heat Waves	39
3	Conor Maynard	You Broke Me First	16
4	SEVENTEEN	Rock with you	11
5	DAY6	Love me or Leave me	10
6	LE SSERAFIM	ANTIFRAGILE	4
7	Cash Cash	All My Love (feat. Conor Maynard)	3
8	BLACKPINK	Typa Girl	2

시간대 선택

확인

시간대 선택

확인

2024-01-04 09:00:00

2024-01-04 10:00:00

2024-01-04 11:00:00

2024-01-04 12:00:00

2024-01-04 13:00:00

2024-01-04 14:00:00

2024-01-04 15:00:00

2024-01-04 16:00:00

2024-01-04 17:00:00

2024-01-04 18:00:00

뮤직 페이지 - 시간대별 차트



4.5Hz차트

2024-01-04 12:00:00 집계 확인

순위	가수명	제목	재생
1	TAEYEON	INVU	37
2	Gryffin	Cry	21
3	ILLENIUM	Eyes Wide Shut (feat. Travis Barker) - Sickick Remix	12
4	EVERGLOW	NO GOOD REASON	7
5	Gryffin	Remember (with ZOHARA) - Madison Mars Remix	6
6	Gryffin	Forever (feat. Elley Duhé)	3
7	JENNIE	You & Me	1

시간대 선택

확인

시간대 선택

2024-01-04 09:00:00

2024-01-04 10:00:00

2024-01-04 11:00:00

2024-01-04 12:00:00

2024-01-04 13:00:00

2024-01-04 14:00:00

2024-01-04 15:00:00

2024-01-04 16:00:00

2024-01-04 17:00:00

2024-01-04 18:00:00



전체 집계 & 개인 집계 페이지

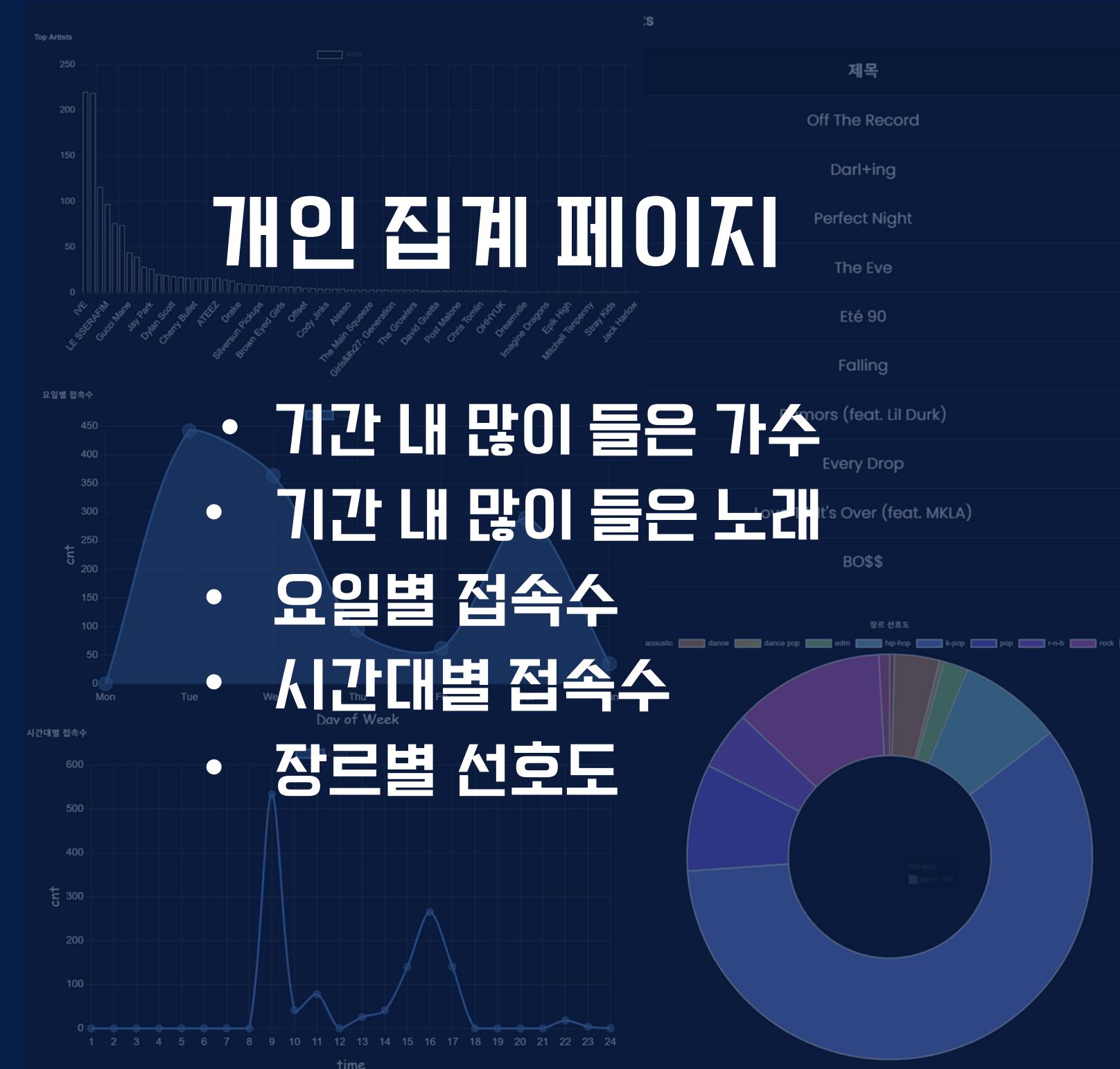
전체 집계 페이지

- 실시간 접속자수
 - 나이대별 이용자수
 - 남녀비율
 - 장르별 이용자율
 - 오늘 많이 들은 가수
 - 오늘 많이 들은 노래



개인 집계 페이지

- **기간 내 많이 들은 가수**
 - **기간 내 많이 들은 노래**
 - **요일별 접속수**
 - **시간대별 접속수**
 - **장르별 선호도**



아쉬운 점 / 추후 구현 예획



- 스포티파이API를 이용한 음악 재생 기능 구현하지 못함
- 음원 제공을 받지 않아, librosa 라이브러리를 이용한 오디오 특성 분석을 직접 하지 못함
- 노래 가사를 이용한 감성분석을 제외함

音楽

