**Report**

**{id\_name}**

**0. 데이터셋 구성**

https://www.kaggle.com/datasets/nelgiriyewithana/top-spotify-songs-2023/data

2023 spotify dataset (.csv) 에 존재하는 총 953곡에 대한 각 특징값인 열에 해당하는 값 중 의미 있다고 생각되는 열들을 선택하여 python기반 분석을 진행하였음. 선택한 열은 아래와 같이 총 13개임.

Artist(s)\_name, released\_year, released\_month, streams, key, mode, danceability\_%, valence\_%, energy\_%, acousticness\_%, instrumentalness\_%, liveness\_%, speechiness\_%

**1. 노래 취향 분석**

전체 곡들에 대하여 어떠한 특징들이 있는지 분석을 수행함.

(1) 노래 특성, 장르 통계적 분석

텍스트, 번호, 영수증, 폰트이(가) 표시된 사진

자동 생성된 설명

danceability\_ (%):

대부분의 곡들이 높은 댄스적인 성격(66.97) 을 가지고 있음.

valence\_ (%):

발란스는 트랙이 전달하는 음악적 긍정성을 의마하는데, 평균값이 약 51%로 다양한 곡들이 감정 스펙트럼을 지니고 있음을 알 수 있음

energy\_ (%):

댄스성과 유사하게, 에너지 수준은 평균적으로 중간에서 높은 수준(64.28) 이며, 많은 트랙들이 활기차고 강렬함을 알 수 있음

acousticness\_ (%):

어쿠스틱성에는 표준편차가 큰 것으로 보아(27.06) 곡마다 상당한 변동성이 있으며, 전자/합성적인 소리와 자연스러운 어쿠스틱 사운드가 혼합되어 있음

instrumentalness\_ (%):

매우 낮은 평균값 수치(1.58) 는 거의 대부분의 트랙들이 보컬 콘텐츠를 기반하고 있으며, 소수만이 악기 연주 곡임을 알 수 있음.

liveness\_ (%):

라이브성은 녹음에서 청중의 존재를 의미하는데 평균적으로 낮은 수치(18.21) 로 보아 대부분의 트랙들이 스튜디오 녹음임을 알 수 있음.

(2) 출시시기 분석

텍스트, 스크린샷, 폰트, 그래프이(가) 표시된 사진

자동 생성된 설명

유독 1월과 5월에 상대적으로 다른 달에 비해 비교적 많은 곡들이 출시됨을 알 수 있음.

(3) key 분석

**텍스트, 스크린샷, 그래프, 폰트이(가) 표시된 사진

자동 생성된 설명**

특히 C#의 곡들이 비교적 많은 것을 알 수 있음. 반대로 E,A#,D# 은 평균보다 적음을 확인할 수 있음.

**2. 많이 들은(Top100 이내 재생횟수) 노래 분석**

Streams 횟수기준 Top 100 이내 곡들에 대한 분석을 수행함

(1) 인기 있는 노래 특성, 장르 통계적 분석

**텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

자동 생성된 설명**

Top 100의 경우 전체 노래에 비하여 곡의 밸런스 적인 특성은 좀 떨어진다고 할 수 있다. 대신 그 외의 곡들에 비해 댄스곡 적인 성격이 더 많다 던가 energy 부분에서 높은 값을 받는다고는 할 수 없음

(2) 출시시기 분석

텍스트, 스크린샷, 그래프, 라인이(가) 표시된 사진

자동 생성된 설명

1월과 3월에 노래를 출시 할수록 top100에 들어갈 확률이 높다고 볼 수 있고 반대로 8월과 9월에 곡을 출시 하는 것은 좋지 않음

(3) key 분석

텍스트, 스크린샷, 그래프, 라인이(가) 표시된 사진

자동 생성된 설명

C#, A mode 에 해당하는 곡일수록 인기있는 노래가 될 수 있는 확률이 높고 반대로 D, G# 곡의 경우 비교적으로 낮은 재생횟수를 보임