

2024년 3월 20일

프로그래머스- SQL 활용한 EDA수행 프로젝트

Spotify User 행동 데이터 EDA 수행

Spotify 사용자 파악, 음악 추천 시스템 개선 및 구독 증진
마케팅 전략

- 02기 2팀 1조 -

김희섭, 김연우

손준영, 이지윤, 한승윤

<목차>

1. 서론

- 주제 선정 및 프로젝트 목적
- 데이터 설명
- 활용 기술 및 프레임워크
- 참여자 정보 및 각 역할

2. 데이터 전처리

- 결측값 확인
- 중복값 처리

3. 주요 분석 결과

- USER EDA 통한 음악 추천 시스템 개선
- 사용자 경험 지표 분석을 통한 음악 추천 시스템 개선
- 프리미엄 구독자 증진 계획

4. 결론과 향후 방향(= 기대효과)

- 분석 결과 요약 및 방향 제시

1. 서론

1.1. 주제 선정 및 프로젝트 목적

● 주제 선정의 배경

디지털 시대의 도래와 함께 음악 스트리밍 서비스는 급속도로 발전하고, 사용자들의 생활 속 깊숙이 자리 잡았다. Spotify는 이러한 변화의 최전선에서 사용자 경험을 혁신하고, 개인화된 서비스를 제공하기 위해 노력하고 있다. 그러나 스트리밍 시장의 경쟁 심화와 사용자의 기대치 상승은 Spotify에게 지속적인 서비스 개선과 차별화된 마케팅 전략의 필요성을 강조한다. 이러한 상황에서 "Spotify 사용자 행동 데이터를 기반으로 한 음악 추천 시스템의 개선 및 구독 증진 마케팅 전략 수립"이라는 주제는 Spotify가 시장에서의 경쟁 우위를 확보하고, 사용자 기반을 확장하는 데 중요한 역할을 할 것이다.

● 프로젝트 목적

1. Spotify 사용자 행동의 심층 분석

- a. 사용자의 연령, 성별, 가입 기간은 물론, 음악과 팟캐스트에 대한 선호도 등까지 다양한 차원에서 분석을 수행한다. 이를 통해 사용자들의 선호와 행동 패턴의 미묘한 차이를 포착하고, 이들이 서비스에서 무엇을 가장 가치 있게 여기는지 이해하고자 한다.
- b. 사용자 세그먼트별로 구체적인 분석을 실시하여, 각 그룹의 고유한 특성과 요구 사항을 식별한다. 이 정보는 사용자 경험을 개인화하고, 각 사용자 그룹에 맞는 맞춤형 서비스를 제공하는 기반을 마련한다.

2. 음악 추천 시스템의 체계적 개선

- a. 사용자의 피드백과 서비스 사용 데이터를 분석하여, 음악 추천 시스템의 현재 성능을 평가한다. 이를 통해 시스템의 정확도와 사용자 만족도를 높일 수 있는 개선 사항을 도출한다.
- b. 추천 알고리즘의 개선뿐만 아니라, 사용자가 새로운 음악을 발견하고 탐색하는 방식에 혁신을 도모함으로써, 더욱 풍부하고 개인화된 음악 청취 경험을 제공한다.

3. 효과적인 구독 증진 마케팅 전략 개발

- a. 분석 결과를 기반으로, 프리미엄 구독 전환 가능성이 높은 사용자 그룹을 명확히 식별한다. 이들에게 특화된 마케팅 메시지와 프로모션을 설계하여, 구독 전환율을 최대화한다.
- b. 기존 프리미엄 구독자의 서비스 사용 패턴과 만족도를 분석하여, 구독 연장을 유도하고 충성도를 높이는 전략을 수립한다. 이는 장기적으로 구독자 기반의 안정성과 수익성을 강화하는 데 기여한다.

이 프로젝트는 Spotify가 사용자 경험을 극대화하고, 지속 가능한 성장을 이루는 데 필수적인 전략적 통찰력을 제공할 것이다. 사용자 데이터에 근거한 분석과 이를 통해 도출된 마케팅 전략은 Spotify가 스트리밍 시장에서 지속적인 경쟁 우위를 확보하는 데 중추적인 역할을 할 것으로 기대된다.

1.2. 데이터 설명

- **데이터:** Spotify User Behavior Dataset (행: 520 , 열: 20)

<https://www.kaggle.com/datasets/meeraaiyavakumar/spotify-user-behavior-dataset>

- **변수:**

데이터 항목	설명	예시
Age	사용자의 연령대	'6-12', '12-20', '20-35', '35-60'
Gender	사용자의 성별	'Male', 'Female', 'Others'
spotify_usage_period	스포티파이 사용 기간	'Less than 6 months', '6 months to 1 year', '1 year to 2 years', 'More than 2 years'
spotify_listening_device	스포티파이 청취 장치	'Smartphone', 'Computer or laptop', 'Smart speakers or voice assistants'
spotify_subscription_plan	스포티파이 구독 유형	'Free (ad-supported)', 'Premium'
premium_sub_willingness	프리미엄 구독 의향	'Yes', 'No'

preffered_premium_plan	선호하는 프리미엄 플랜	'Family Plan-Rs 179/month', 'Individual Plan- Rs 119/month', 'Student Plan-Rs 59/month'
preferred_listening_content	선호하는 청취 콘텐츠	'Music', 'Podcast'
fav_music_genre	선호하는 음악 장르	'Pop', 'Rock', 'Rap', 'Melody'
music_time_slot	음악 청취 시간대	'Morning', 'Afternoon', 'Evening', 'Night'
music_Influencial_mood	음악 청취가 영향을 미치는 기분	'Relaxation and stress relief', 'Sadness or melancholy', 'Social gatherings or parties'
music_lis_frequency	음악 청취 빈도	'leisure time', 'Study Hours', 'While Traveling', 'Office hours', 'Workout session'
music_expl_method	음악 탐색 방법	'Playlists', 'recommendations'
music_recc_rating	음악 추천 평가	1부터 5까지의 숫자
pod_lis_frequency	팟캐스트 청취 빈도	'Daily', 'Several times a week', 'Once a week', 'Rarely', 'Never'
fav_pod_genre	선호하는 팟캐스트 장르	'Comedy', 'Sports', 'Lifestyle and Health'
preffered_pod_format	선호하는 팟캐스트 형식	'Interview', 'Story telling'
pod_host_preference	팟캐스트 호스트 선호도	'Well known individuals', 'Both', 'None'
preffered_pod_duration	선호하는 팟캐스트 지속 시간	'Short', 'Long', 'Both', 'None'
pod_variety_satisfaction	팟캐스트 다양성에 대한 만족도	'Satisfied', 'Ok', 'Unsatisfied'

1.3. 분석 방법

본 프로젝트에서는 데이터 분석 및 시각화 과정을 통해 Spotify 사용자 행동에 대한 깊이 있는 이해를 목표로 했다. 이 과정에서 사용된 주요 기술과 방법론을 자세히 설명하겠다.

● 데이터 분석 - SQL

○ SQL 활용

데이터 추출 및 분석: 데이터 분석 과정에서는 SQL 데이터베이스 관리 시스템(DBMS)인 MySQL, MariaDB 등을 사용하여 데이터를 불러왔다. 이러한 시스템은 대용량의 데이터를 효율적으로 관리하고, 복잡한 쿼리를 실행할 수 있는 강력한 기능을 제공한다. 프로젝트에서는 데이터베이스에 저장된 Spotify 사용자 행동 데이터에 접근하기 위해 SQL 쿼리를 작성하고 실행했다.

쿼리 작성: 분석을 위해 필요한 정보를 추출하기 위한 쿼리를 작성했다. 이 쿼리들은 사용자의 연령대, 성별, 스트리밍 기간, 음악 장르 선호도 등 다양한 차원에서 데이터를 조합하고 분석하기 위해 사용되었다. SQL의 집계 함수, 조인, 서브쿼리 등의 고급 기능을 활용하여, 분석에 필요한 정교한 데이터 집합을 준비했다.

데이터 저장: 쿼리의 결과는 CSV 파일 형태로 저장되었다. 이 과정은 분석 결과를 다른 소프트웨어로 쉽게 전송하고, 다양한 시각화 도구에서 활용할 수 있도록 함으로써, 분석 과정의 유연성을 크게 향상시켰다.

1. 데이터 시각화 - EXCEL & POWER BI

- **Excel 사용:** 분석 결과로 얻어진 CSV 파일은 Microsoft Excel에 가져와졌다. Excel의 다양한 데이터 분석 및 시각화 기능을 활용하여, 데이터를 요약하고, 통계적 분석을 수행했다. 또한, Excel에서 제공하는 차트, 피벗 테이블 및 그래프 도구를 사용하여, 데이터의 분포, 경향성 등을 시각적으로 표현했다. Excel은 사용자 친화적인 인터페이스와 폭넓은 기능 덕분에 데이터 분석 및 시각화 과정에서 매우 유용하게 사용되었다.

- **Power BI** 사용: 더 복잡하고 동적인 시각화가 필요한 경우, Microsoft Power BI를 사용했다. Power BI는 대용량 데이터셋에 대한 고급 분석 기능, 인터랙티브한 대시보드 제작, 데이터 시각화 옵션을 제공한다. 이를 통해 사용자 행동의 다차원적 분석 결과를 직관적으로 이해할 수 있는 시각화를 제작했다. Power BI의 강력한 데이터 모델링 및 시각화 기능은 복잡한 데이터 관계를 명확하게 표현하고, 인사이트 도출을 위한 도구로서 큰 역할을 했다.

1.4. 참여자 정보 및 각 역할

- 김희섭: 음악 추천 시스템 개선을 위한 USER EDA (20%)
- 김연우: 평점을 바탕으로 개선해야할 추천시스템 분야 분석 (20%)
- 손준영: spotify 프리미엄 구독자 증진 - 현구독자 (20%)
- 이지윤: spotify 프리미엄 구독자 증진 - 비구독자 (20%)
- 한승윤: 사용자 경험 지표 분석을 통한 음악 추천 시스템 분석 (20%)

2. 데이터 전처리

2.1 결측값 확인

COLUMN_NAME	null_existence
Age	No
fav_music_genre	No
fav_pod_genre	No
Gender	No
music_expl_method	No
music_Influencial_mood	No
music_lis_frequency	No
music_recc_rating	No
music_time_slot	No
pod_host_preference	No

pod_lis_frequency	No
pod_variety_satisfaction	No
preferred_listening_content	No
preffered_pod_duration	No
preffered_pod_format	No
preffered_premium_plan	No
premium_sub_willingness	No
spotify_listening_device	No
spotify_subscription_plan	No
spotify_usage_period	No

20개의 칼럼 중 각 칼럼에 null 값이 하나라도 존재한다면 'Yes'를 반환하고, 그렇지 않으면 'No'를 반환하도록 했다. 확인 결과 모든 칼럼에 null값이 존재하지 않는다.

2.2 중복값 처리

[다중응답 전처리 전]

spotify_listening_device	count
Smartphone	300
Smartphone, Computer or laptop	48
Smart speakers or voice assistants	43
Smart speakers or voice assistants, Wearable d...	2

[다중응답 전처리 후]

single_device	count
Smartphone	391
Computer or laptop	144
Smart speakers or voice assistants	92
Wearable devices	44

spotify_listening_device, music_Influencial_mood, music_lis_frequency, music_expl_method 칼럼에서 사용자들의 응답을 받을때 다중응답을 받아 행마다 중복되는 값들이 존재했다. 따라서 각 값들이 개별적으로 카운트 되도록 하기 위해, SUBSTRING_INDEX 함수를 사용하여 문자열을 쉼표(,)를 기준으로 분리하고, TRIM 함수를 사용하여 각 추출된 문자열에서 앞뒤 공백을 제거하여 각 칼럼들의 다중응답을 개별적으로 추출할 수 있도록 전처리 하는 과정을 거쳤다.

다른 칼럼들과 연관성을 찾기 위한 분석 시 여러 칼럼 중 다중응답 칼럼을 포함했을 때에는 다중응답 컬럼들의 값을 기준으로 새로운 테이블을 생성하여 진행하였다.

3. 주요 분석 결과

3.1. USER EDA 통한 음악 추천 시스템 개선

1. 분석 개요

최근 디지털 음악 서비스의 성장과 함께 사용자의 음악 경험을 개인화하는 것이 중요한 경쟁요소가 되었다. 해당 분석은 사용자의 음악 선호, 시간대, 감정 등 다양한 데이터를 종합적으로 분석함으로써, 사용자에게 더욱 맞춤화 되고 맥락에 부합하는 음악 추천을 제공하는 방법을 모색하고자 한다.

해당 분석을 통해 사용자 데이터에서 의미 있는 사용자 패턴 및 인사이트를 도출하고, 기존 추천 알고리즘을 개선한다. 이 과정을 통해 사용자의 만족도를 높이고 서비스 이탈률을 감소시키며, 장기적으로 Spotify 수익성을 증진시킬 것으로 기대된다.

또한 음악 추천 시스템의 효율성을 향상시키고 사용자 경험을 극대화하며 서비스 제공자 측면에선 사용자의 행동에 대한 이해를 제공함으로써 지속 가능한 서비스 개선을 가능하도록 한다.

참고로 팟캐스트의 경우 ‘preferred_listening_content’에서 확인했을 때 약 4:1 비율을 보여줘 데이터가 충분하지 않다고 판단하여 범위를 음악으로 세분화하여 진행하였다.

2. 사용 변수

변수는 음악 추천 시 고려되어 질 수 있는 변수들을 사용하였다.

- 1) **Age** : 사용자의 나이는 음악 취향 및 선호도에 중요한 영향을 미칠 수 있다.
- 2) **Gender** : 성별에 따른 취향 차이가 있을 수 있다.
- 3) **spotify_usage_period** : 사용 기간은 사용자의 서비스에 대한 경험이나 음악 감상 습관을 반영할 수 있다.
- 4) **fav_music_genre** : 사용자의 선호 장르는 매우 중요한 변수로, 개인화된 음악 추천에 도움이 되는 변수이다.
- 5) **music_time_slot** : 사용자가 음악을 주로 듣는 시간대는 그들의 일상 패턴 및 상황을 반영할 수 있고, 음악 선택에 중요한 영향을 끼칠 수 있다.

- 6) **music_influential_mood** : 사용자가 어떤 기분일 때 음악을 듣는지 이해함으로써 기분 개선 및 유지에 어떤 음악이 사용되는지 알 수 있다.
- 7) **music_list_frequency** : 사용자가 어떤 상황에 자주 음악을 듣는지 감상 습관 및 관심도를 나타낸다.

3. 분석 전략

분석 진행은 먼저 기준 대상 범위를 선정한다. 전체 범위를 기준으로 진행하기보단, 음악 추천 시스템을 개선하려는 것이 목적이므로 분석 대상을 먼저 음악을 선호하는 사용자로 추린다. 그 다음 사용자들 중에서도 개선을 필요로 하는 사용자는 평점 기준 비교적 낮은 평점을 수여한 사람들이라고 생각된다.

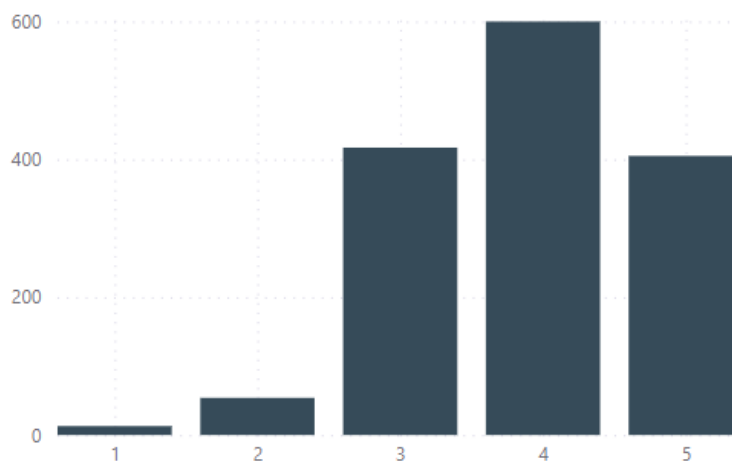
해당 분석 대상을 기준으로 1차적으로 컬럼별 분포를 확인한다. 그 결과에 따라 가설을 세워가며, 이를 증명하고 확인하는 방식으로 진행한다.

이 과정을 반복하며 사용자들을 일정 기준에 따라 그룹화하고, 그룹별 적절한 추천 전략을 수립하는 것이 주된 목적으로 진행된다.

4. 분석 결과

1) 기준 대상 범위 선정

선정을 위해 음악 선호 사용자들의 평점을 확인하였다.



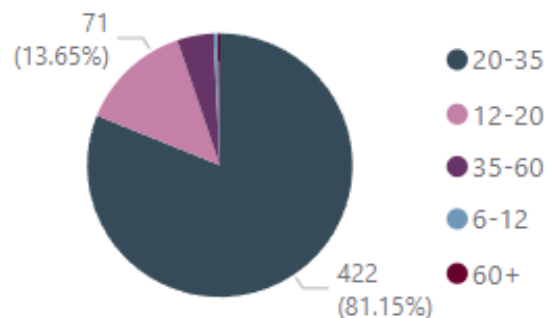
4가 가장 많고, 3, 5 순으로 이어진다.

보통 1 2 3 4 5 를 기준으로 선택 시 아주 나쁨, 나쁨, 보통, 좋음, 아주 좋음으로 생각되어진다.

낮은 평점을 기준 삼을 때, 보통의 경우 개인적으로 좋지도 나쁘지도 않은 애매한 상태라고 생각된다. 따라서 이 집단을 낮은 평점을 준 집단에 포함해 좋은 평점이 나오도록 개선하는 방향이 맞다고 생각되어 범위는 1 ~ 3의 평점을 준 사용자들을 대상으로 진행하였다.

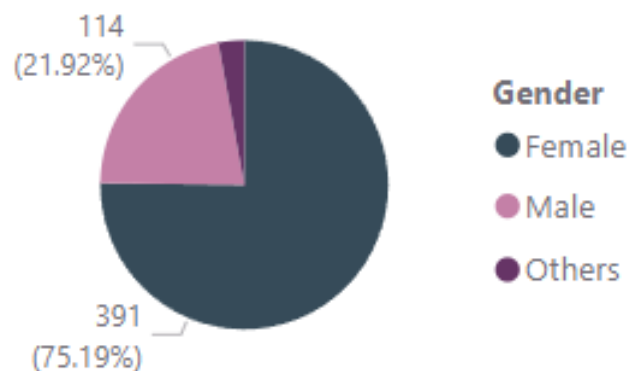
2) 사용자 분포

2-1) Age



20-35세에 해당하는 젊은 청년층이 약 81.15%로 가장 많이 사용함을 알 수 있다. 다른 나이 범위의 경우 집계 데이터 수가 적어 젊은 청년층을 기준으로 진행하는 것이 좋아보인다.

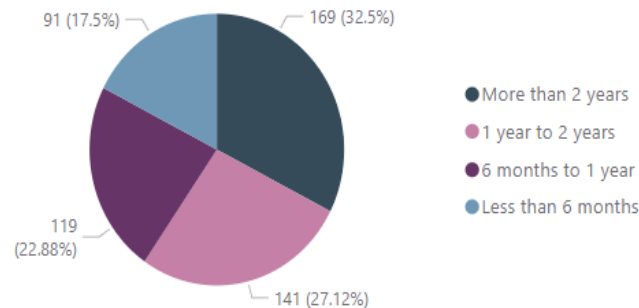
2-2) Gender



여성의 경우 약 75.19%의 비율을 보여주며 가장 많은 비중을 차지하고 있음을 알 수 있다. Others라는 특이한 성별 값을 가진 집단이 있지만, 그 수가 적어 무시하고 넘어가도 될 것 같다.

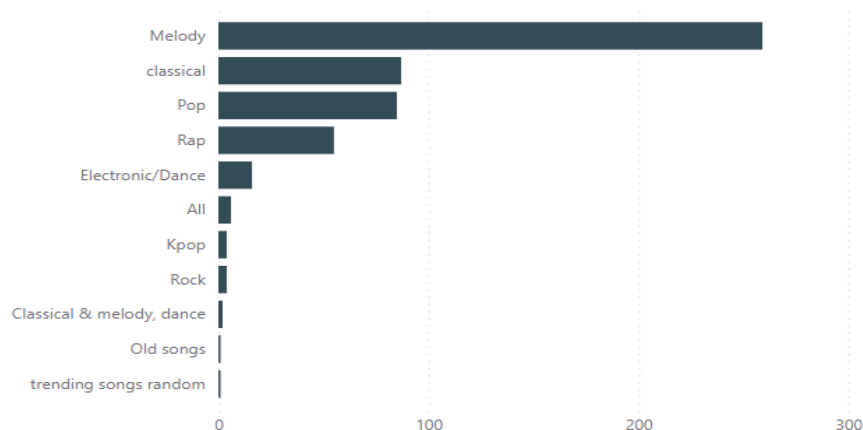
남성과 여성의 경우 감정에 있어 느끼는 정도의 차이가 있다는 연구결과가 있으므로, 이 점을 참고하여 음악 감상 시 감정과 연관지어 분석을 해보는 것이 좋아보인다.

2-3) spotify_usage_period



사용 기간이 짧을 수록 사용자 별 수집된 데이터가 적기에 만족도가 낮을 것이라고 예상했지만, 오히려 사용 기간이 길수록 만족도가 낮게 나왔다. 이를 통해 수집된 데이터를 분류하는 시스템에 문제가 있다고 예상해 볼 수 있다.

2-4) fav_music_genre

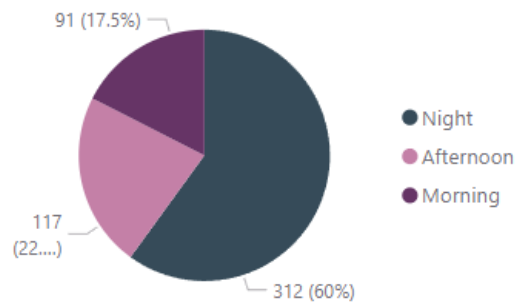


장르의 경우 Melody 장르를 가장 선호한다. 집계된 선호 장르 수에 따라 고려 시 Melody, Pop, classical, Rap, Electronic/Dance, Rock 까지만 고려를 하는 것이 좋아보인다.

장르의 경우 mood / frequency와 연관성을 찾아보는 것도 좋아보인다.

Melody의 장르의 경우 정의를 가사가 없는 멜로디만 있는 형식으로 음악으로 진행하고자 한다.

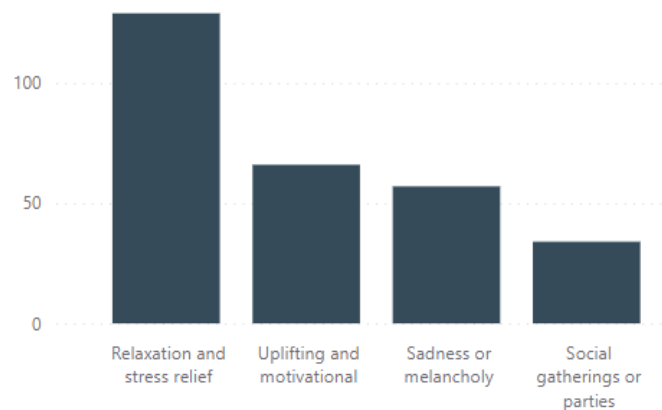
2-5) music_time_slot



음악 감상 시간대의 경우 Night 시간대에 약 60%로 가장 많은 사용자가 있음을 알 수 있다.

이는 퇴근 시간, 여가 시간, 취침, 파티등의 이유로 Night에 가장 많이 듣는다고 예상된다. 이 부분도 mood / frequency와 연관 지어 보는 것이 좋아보인다.

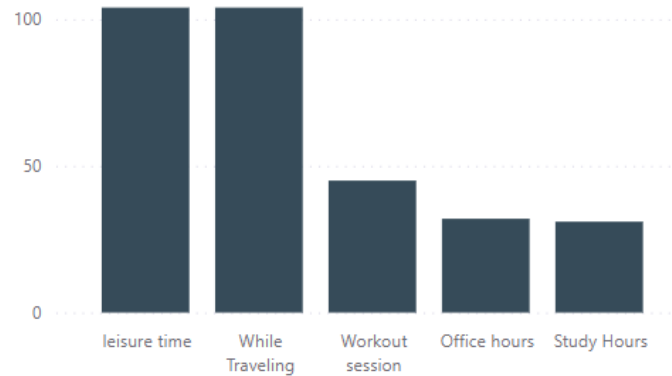
2-6) music_Influential_mood



Relaxation and stress relief가 가장 많은 수를 보여준다. Uplifting and motivational, Sadness or melancholy, Social gatherings or parties 순으로 이어진다.

각 상황 별 frequency / genre를 연관지어 본다면 상황별 어울리는 음악 장르와 분위기를 알 수 있을 것 같다.

2-7) music_lis_frequency



leisure time, While Traveling 이 가장 많이 동등하게 집계되었다. leisure time 의 경우 여가 시간을 뜻하기에 일반적으로 음악을 가장 많이 들을 수 있는 시간이기때 가장 많은 수로 집계된 것으로 보인다.

이점도 자세히 알기 위해선 각 상황에 맞는 mood를 알아야 명확하게 구분지을 수 있을 것 같다.

3) 가설 증명 및 추가분석

앞선 분포들을 기반으로 설정한 가설들은 총 4가지 이다.

1. 젊은 층 대상의 성별에 따른 감정

여성의 경우 더 감정에 대해 민감하게 반응한다는 연구 결과를 바탕으로 'Sadness or melancholy'의 경우가 높게 나타날 것으로 보인다.

2. 감정에 따른 선호 장르

음악은 감정 상태를 반영하고 영향을 끼칠 수 있는 매체이기에 감정에 따라 선호하는 장르에 영향을 끼칠 것으로 보여진다.

3. 상황 별 선호되는 감정 및 장르

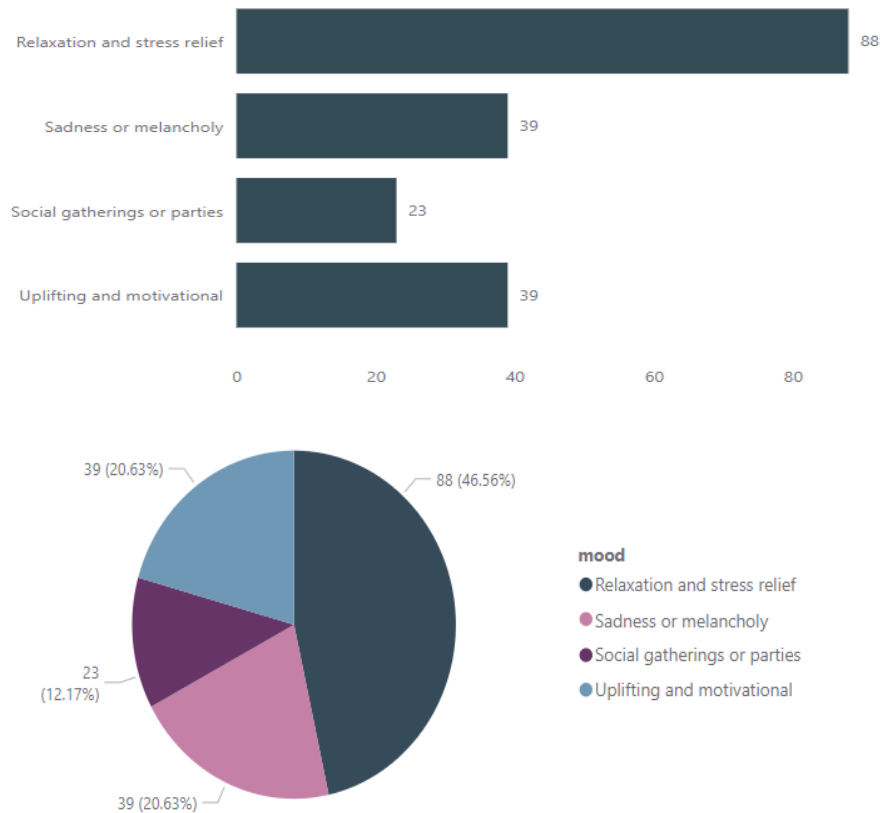
사용자가 상황에서 느끼는 감정은 각자 다르기 때문에 앞선 감정에 따른 선호 장르에 차이가 있다면 이 또한 차이점이 있을 것으로 보여진다.

4. 음악 감상 시간대 별 상황 및 감정

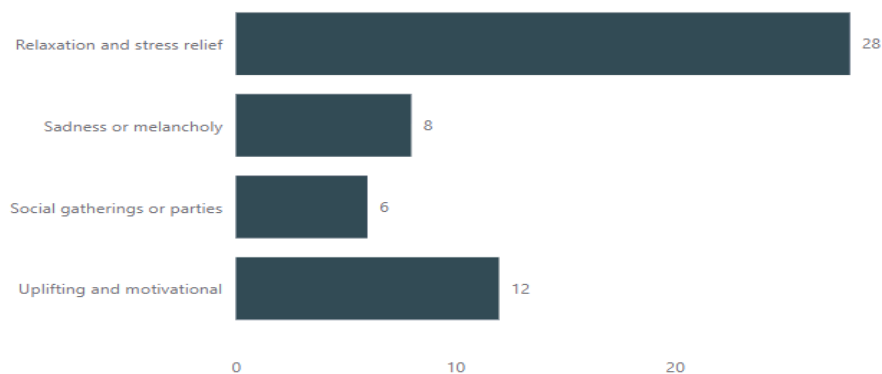
시간대 별로 출근, 운동, 여행 등 특정 상황이 많이 나타날 수 있기 때문에
영향이 있을 것으로 보이고, 또한 시간대 별 느낄 수 있는 감정의 차이가 있기에
이를 연관지어 볼 수 있다고 생각된다.

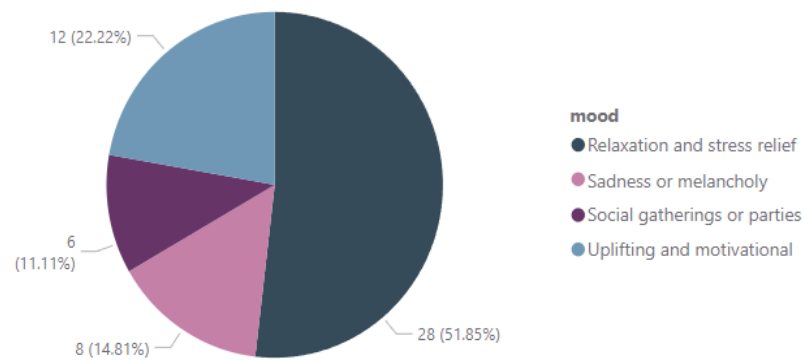
3-1) 젊은 층 대상의 성별에 따른 감정에 대한 증명

[여성]



[남성]





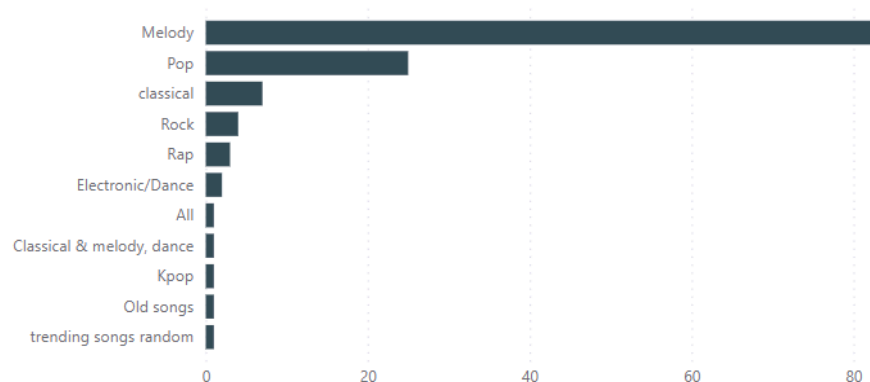
여성 사용자들이 슬픔이나 우울함, 고무적이고 동기 부여가 되는 기분일 때 음악을 듣는 것을 선호하는 이유는 음악이 감정적인 치유 및 긍정적인 자극을 제공하기 때문일 수 있다.

반면, 남성 사용자들이 이완 및 스트레스 해소를 위해 음악을 듣는 것은 스트레스를 완화하고 마음을 진정시키는 데 도움을 줄 수 있기 때문이다.

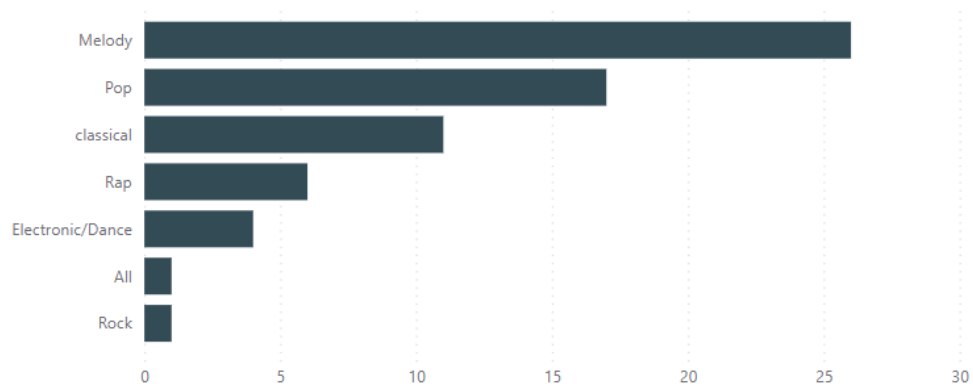
이 점은 연구 중 여성이 남성보다 행복감 및 우울감을 좀 더 잘 느낀다는 점이 반영되어 남성과 여성의 음악 감상 시 감정에 차이를 보여주는 것으로 보여진다.

3-2) 감정에 따른 선호 장르에 대한 검증

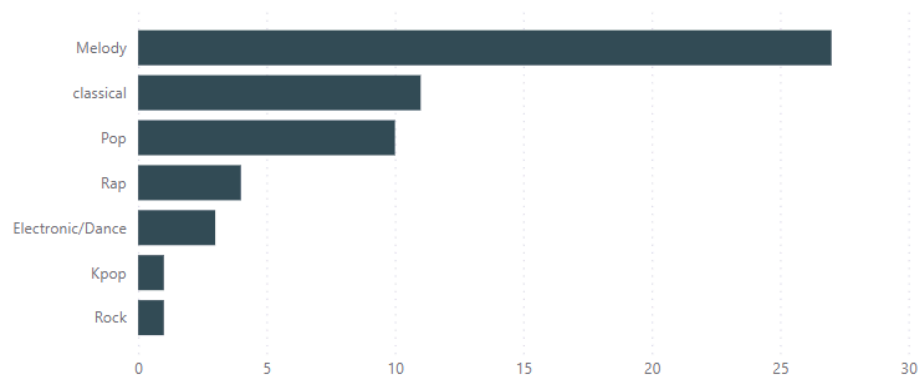
[Relaxation and stress relief]



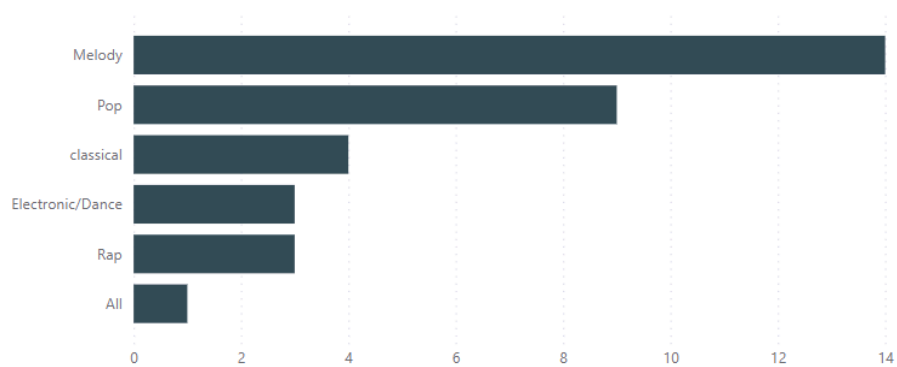
[Uplifting and motivational]



[Sadness and melancholy]



[Social gatherings and parties]



처음 가설을 세울 때 예상했던 결과는 Relaxation and stress relief의 경우 심신 안정이 목적일 경우 차분한 음악, 스트레스 해소 목적일 경우 신나거나 차분한 음악 (ex) Melody, Pop, classical, Rock 등)

Uplifting and motivational의 경우 동기부여 목적으로 신나고 도전적인 음악 (ex) Pop, Rap, Electronic/Dance, Rock 등)

Sadness and melancholy의 경우 비교적 어둡고 슬픈 선율의 음악 (ex) Melody, Pop, classical 등)

Social gatherings and parties의 경우 파티 등의 목적으로 신나는 음악 (ex) Pop, Rock, Electronic/Dance 등)

이렇게 가설을 세우고 진행했지만, 모두 Melody, Pop, classical 장르가 많이 나타났다. 물론 Sadness and melancholy 때 classical이 Pop보다 더 많이 선호되는 차이가 있지만, 그 차이가 미미하고, 모든 감정 상황에 동일하게 Melody가 많은 비중을 차지하는 것을 보여주고 있다. 이는 앞서 전체적으로 살펴보았을 때 Melody의 비중이 다른 값들에 비해 많은 비중을 차지하고 있기 때문으로 보인다.

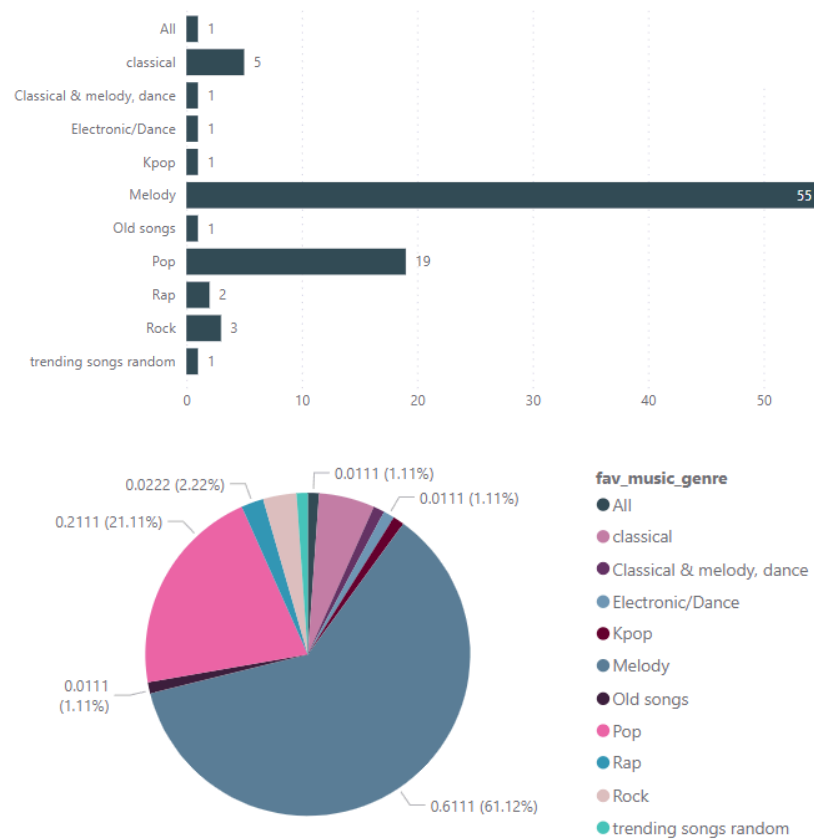
추가적으로 앞서 장르를 살펴볼 때 Melody라는 장르의 정의가 불분명한 문제가 있었고, 또한 Pop이라는 장르에도 곡 분위기가 즐겁거나 슬플 수 있기 때문에 값 자체가 연관성을 찾기 힘든 값으로 이루어져 있음을 알 수 있다.

이 결과로 보면 현 데이터에서 장르보단 감정이 음악 분류에 더 적합해 보이지만, 신나는 분위기의 곡일 경우 Uplifting and motivational, Relaxation and stress relief, Social gatherings or parties 이 3가지에 모두 속할 수 있기에 적합하다고 보기 어렵다.

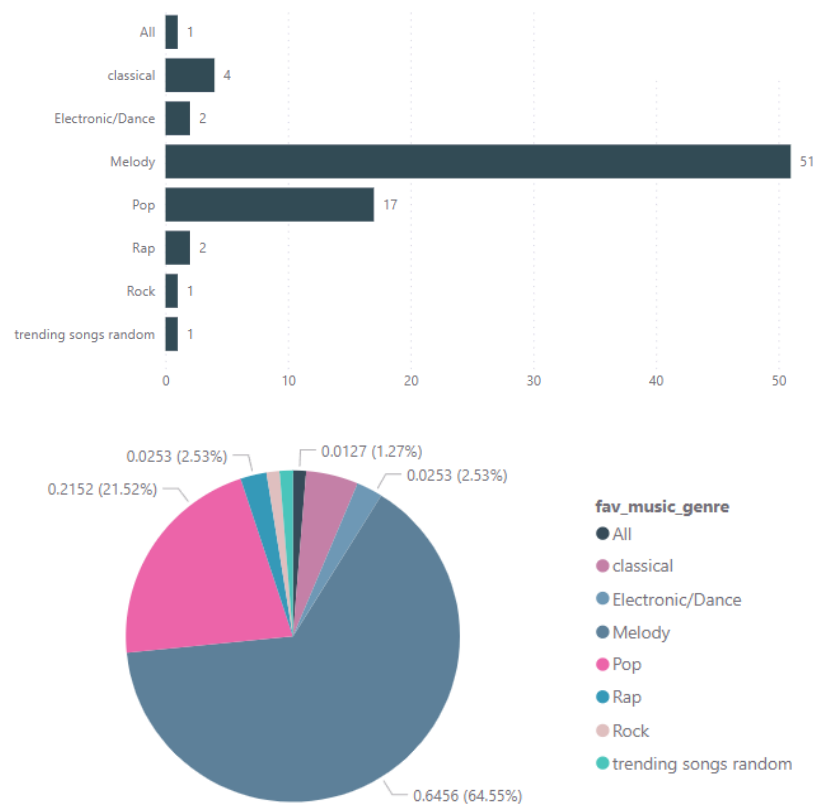
3-3) 상황 별 선호되는 감정 및 장르에 대한 검증

앞선 3-2의 결과를 바탕으로 특정 컬럼의 값이 한 쪽으로 치우쳐져 있다면, 연관지어 다른 컬럼과 대조했을 때도 특정 값으로 치우쳐진 결과가 나온다는 것을 알 수 있었다. 따라서 상황 별 장르에 대해 주로 진행, 감정의 경우 특정 상황에서 현재 데이터의 감정 값들을 기준으로 살펴보았을 때 여러 대조적인 감정이 나타날 것으로 보여지지 않기 때문에 (ex) 동기부여 vs 우울감) 해당 상황 별 감정 집계 값 중 평균 이상인 감정들을 기준으로 장르를 구분하고자 하였다. 평균 이상의 감정의 경우 해당 상황에 가장 많이 느낄 수 있는 감정이고, 해당 감정을 곡의 분위기로 가정, 장르만 구분한다면 그룹화를 할 수 있을 것으로 보인다.

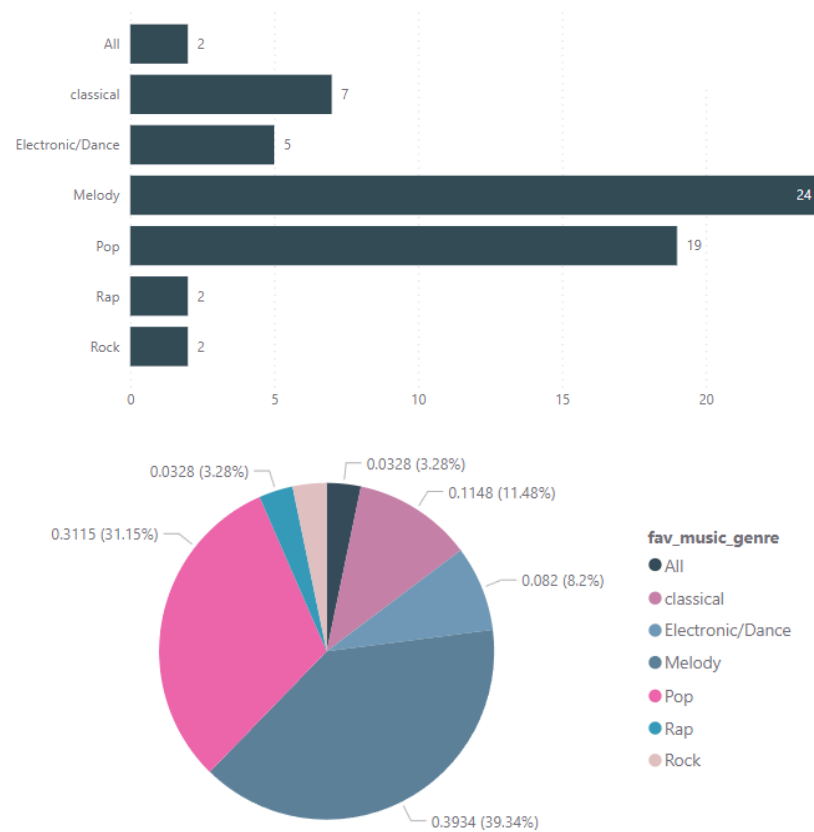
[leisure time – Relaxation and stress relief]



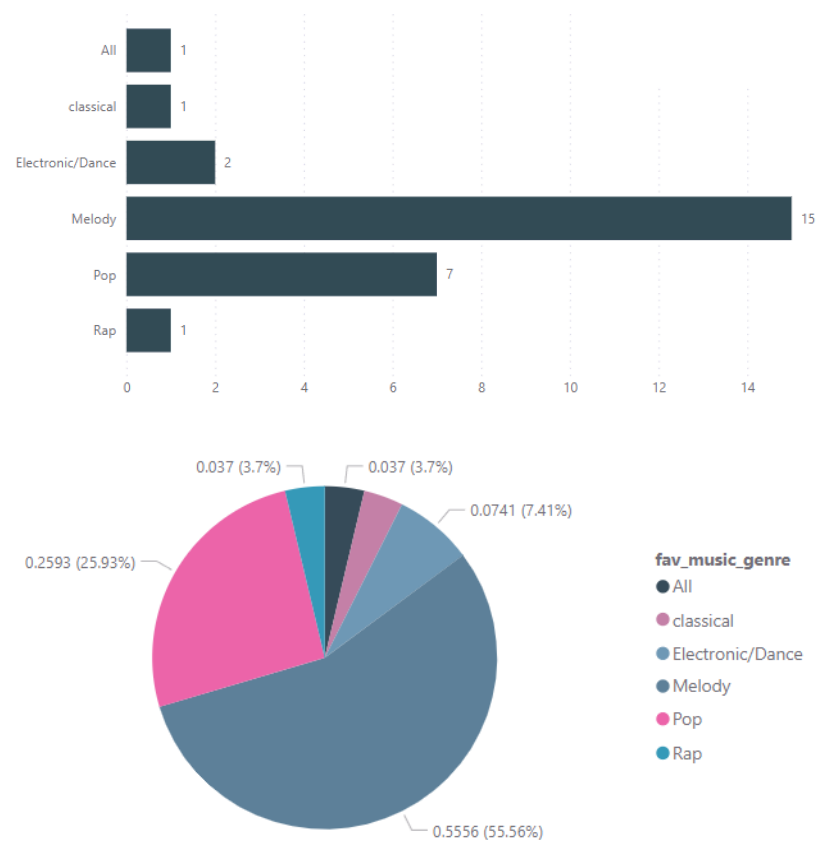
[While Traveling – Relaxation and stress relief]



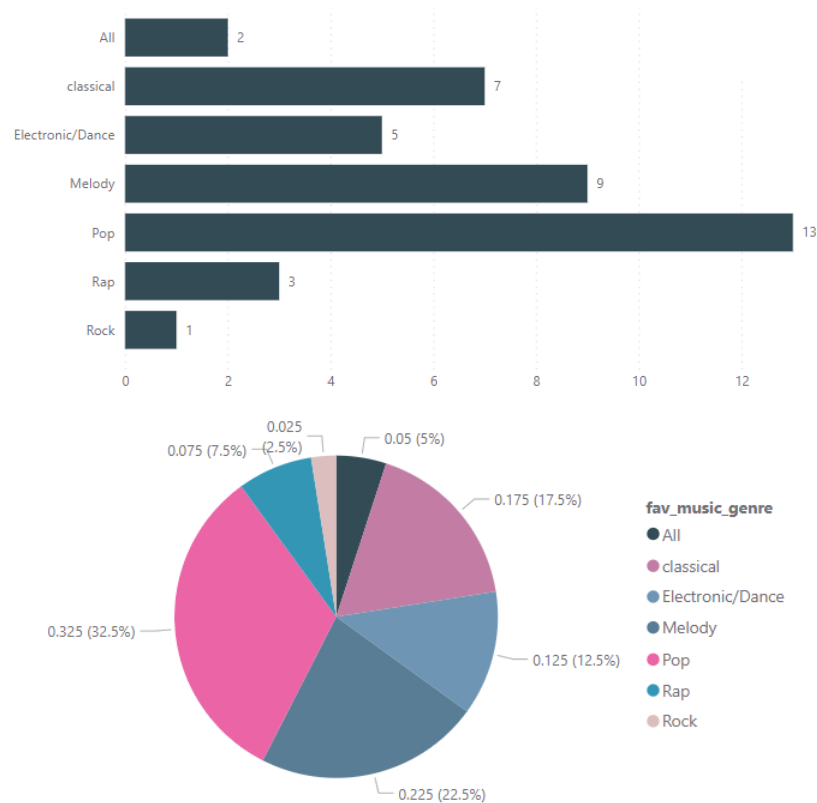
[Workout session – Uplifting and motivational, Relaxation and stress relief]



[Office hours – Relaxation and stress relief]



[Study hours – Relaxation and stress relief]



전체 기준에서 가장 많이 나타났던 Melody, Pop 제외하여 각 상황 별로 살펴 볼 경우, leisure time에서는 dance 음악 류를 선호하는 사람이 적다. while traveling에서는 락, 유행 음악류를 선호하는 사람이 적다. workout session에서는 의외로 랩, 락 음악류를 선호하는 사람이 적다. office hours 에서는 클래식, 랩 음악류를 선호하는 사람이 적다. study hours에서는 락을 선호하는 사람이 적다.

이것을 기반으로 간단한 분류를 한다면 상황 및 장르를 사용자에게 입력받고 추천 조건에 ‘Workout session – Pop’을 선택한다면 신나는 분위기의 팝을 추천하는 방식을 생각해 볼 수 있다.

따라서 락, 랩, dance 음악 장르들은 비교적 매니아 층이 선호하는 것을 알 수 있고, 대다수의 사람들은 Melody, Pop 장르를 선호하는 것을 알 수 있다.

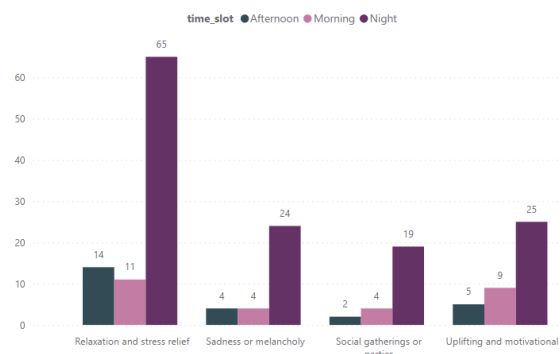
또한 이를 통해 장르의 구체화 혹은, 곡의 분위기를 기준으로 삼아 분류하는 개선이 필요해 보인다.

3-4) 음악 감상 시간대 별 상황 및 감정에 대한 검증

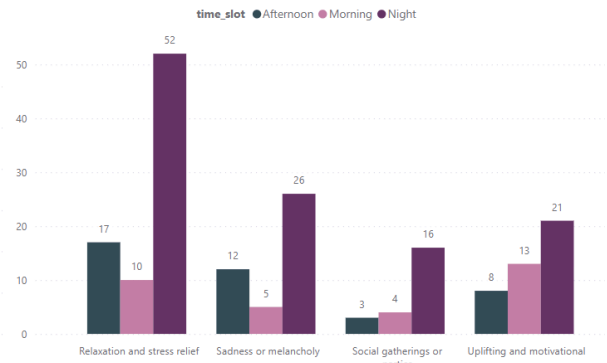
마지막으로 시간대를 고려한 상황 및 감정에 대한 검증이다. 앞선 분석을 통해 상황 별 장르 혹은 감정 상태 별 장르의 연관성을 확인할 수 있었다. 그러나 앞선 두 경우 특정짓기 어려웠고 시스템의 개선보단 자료 수집의 개선이 필요함을 느꼈다.

이번 가설은 장르를 배제시키고 상황 별 감정을 시간대 별로 분석하여 검증하였다.

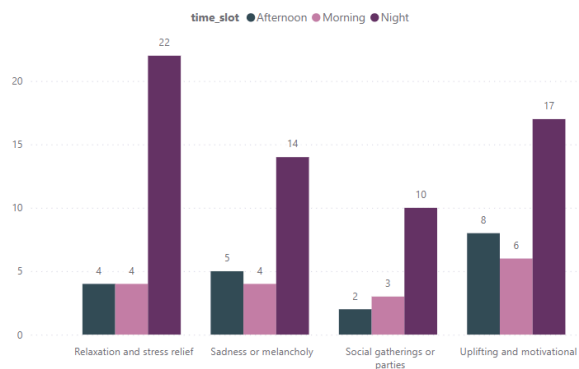
[leisure time]



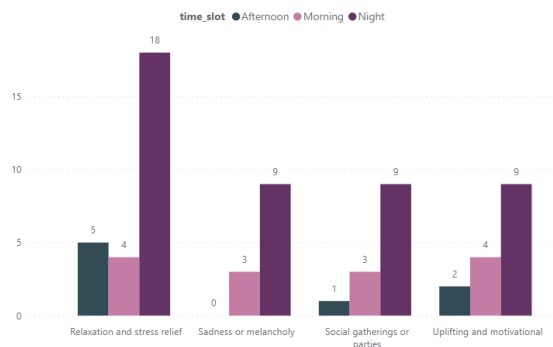
[While Traveling]



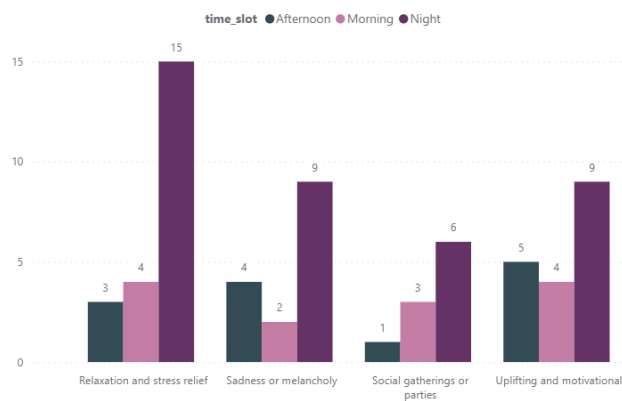
[Workout session]



[Office hours]



[Study hours]



상황 별 시간대에 따라 음악 감상 시 조금씩 차이점을 보여주고 있다.

아침

leisure time : Relaxation and stress relief

While Traveling : Uplifting and motivational

Workout session : Uplifting and motivational

Office hours : Uplifting and motivational , Relaxation and stress relief

Study hours : Uplifting and motivational, Relaxation and stress relief

오후

leisure time : Relaxation and stress relief

While Traveling : Relaxation and stress relief

Workout session : Uplifting and motivational

Office hours : Relaxation and stress relief

Study hours : Uplifting and motivational

저녁

leisure time : Relaxation and stress relief

While Traveling : Uplifting and motivational

Workout session : Relaxation and stress relief

Office hours : Relaxation and stress relief

Study hours : Relaxation and stress relief

5. 결론

Age, Gender 변수들을 통해 성별에 따라 여성의 경우 음악을 감정적 치유 및 긍정적 자극을 위해 음악을 듣는 경향이 있고, 남성의 경우 스트레스 완화 및 마음을 진정시키는 목적으로 음악을 듣는 경향이 있음을 알 수 있었다.

음악 감상 상황 및 감정 상태, 시간대, 장르를 통해 사용자들은 현재 활동이나 감정 상태를 반영하고 이를 감정적으로 강화하는 데 도움을 주는 용도로 음악을 듣는다는 것을 알 수 있었다. 예를 들어, 아침에 운동하는 사용자들은 신나는 노래로 운동 시 동기부여를 하는 목적으로 음악을 듣는다는 것을 알 수 있었다.

추천 시스템을 개선하는 방안은 먼저 성별 기준으로

1. 여성 : 감정에 따른 추천 시스템
2. 남성 : 안정 및 스트레스 해소, 동기부여 위주의 추천 시스템

상황 / 시간대 / 감정 기준으로

1. leisure time 의 경우 아침 ~ 저녁까지 힐링을 위한 차분한 노래
2. While Traveling 의 경우 아침, 저녁에는 신나는 노래, 오후의 경우 차분한 노래로 여행의 즐거움을 만끽
3. Workout session 의 경우 아침, 오후 운동 시에는 신나는 노래로 운동 효율 증진, 저녁에는 일과 후 지친 몸을 무리하지 않으며 운동하도록 차분한 노래
4. Workout session 의 경우 아침에는 출근을 알리며 힘을 내자는 느낌의 신나는 노래, 오후, 저녁의 경우 업무 스트레스를 감당하기 위한 차분한 노래
5. Study hours 의 경우 아침에는 동기부여를 위해 신나거나 집중을 위한 차분한 노래, 오후에는 식곤증으로 빠지지 않기 위한 신나는 노래, 저녁엔 마지막 집중을 위한 차분한 노래

이런식으로 사용자의 음악 추천 시스템 이용 전 상황 / 장르를 입력 받고 해당 분위기에 맞는 장르의 곡들을 추천하는 방향으로 개선하는 것이 좋아보인다.

추가적으로 이번 분석에서 장르, 분위기 등 장르의 값의 구분이 충분하지 않은 변수들에 대해서 추가 수집이 필요해 보인다. 추천 시스템 자체의 개선 보다는 곡에 대해 장르 뿐 아니라 분위기에 따라서 분류를 추가적으로 진행하고 데이터를 수집한다면 보다 정확한 음악 추천 시스템을 구축할 수 있을 것으로 예상된다.

3.2. 사용자 경험 지표 분석을 통한 음악 추천 시스템 개선

1. 분석 개요

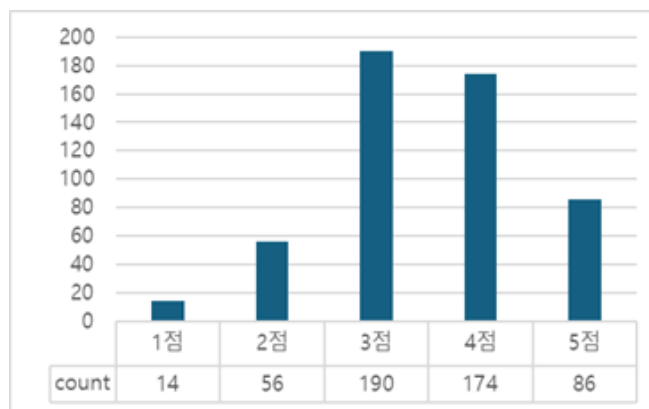
스포티파이에서는 사용자 개인에게 맞춤형 플레이리스트(recommendations), 스포티파이에서 직접 큐레이션한 플레이리스트(Playlists), 특정 아티스트 또는 선택한 음악과 접점이 있는 음악을 자동으로 재생해주는 라디오(Radio)를 제공한다.

본 분석 과제에서는 개인 맞춤형 플레이리스트(recommendations), 즉 스포티파이의 음악 추천 시스템에서 개선 방안 및 개선이 필요한 분야를 찾고자 한다.

스포티파이의 음악 추천 시스템은 규칙 기반 추천(Rule-Based), 내용 기반 필터링(Content-Based Filtering), 협업 필터링(Collaborative Filtering)이 혼합된 하이브리드 시스템일 것이라 가정하고, 3.1. 절에서 USER EDA 분석을 통해 규칙 기반 추천에 대한 개선 방안을 제시했다면, 본 절에서는 사용자 경험 지표 및 피드백을 통해 사용자를 세분화한 뒤 분석을 수행하여 필터링 모델에 대한 개선 방안을 제시하는 데 집중한다.

2. 분석 전략

사용자들의 음악 추천 시스템에 대한 평점(이하 평점) 분포는 다음과 같다.



평점 분포

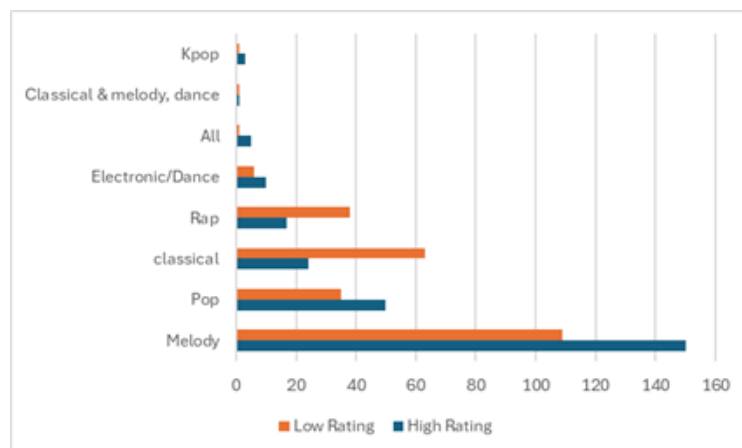
평균과 사분위수는 각각 평균 3.5, Q3 4, Q2 3.5, Q1 3이다. 평균과 중앙값이 일치함에 따라 해당 값인 3.5점을 기준으로 4점과 5점을 준 집단을 High Rating 집단, 1점, 2점, 3점을 준 집단을 Low Rating 집단으로 분류하였다.

High Rating과 Low Rating 집단에 대해 사용자의 기초 정보에 대한 칼럼 Age, Gender, 사용자의 음악 선호도 및 추천 시스템 사용 경험에 대한 칼럼 fav_music_genre, music_time_slot, music_Influential_mood, music_lis_frequency을 분석하여 경향성을 파악하고 개선 지점을 도출하고자 한다.

3. 분석 결과

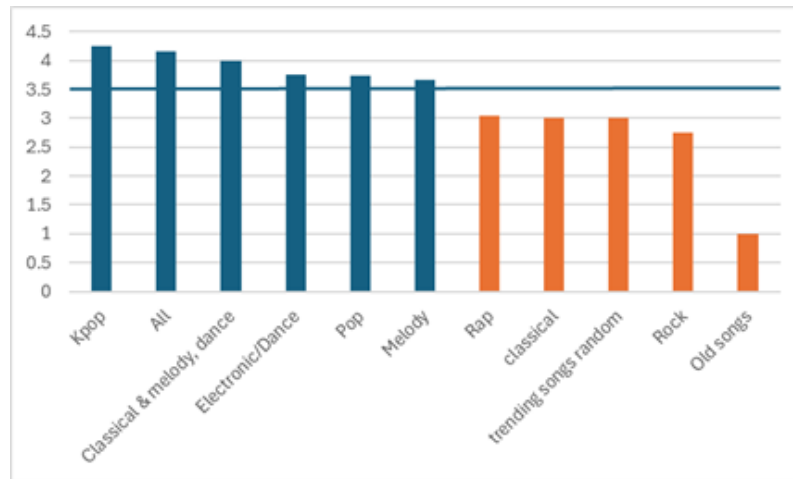
분석 전략에 따른 분석 결과 유의미한 결과를 보인 칼럼에 대한 설명은 다음과 같다.

1) 선호하는 음악 장르(fav_music_genre)



평점과 선호하는 음악 장르

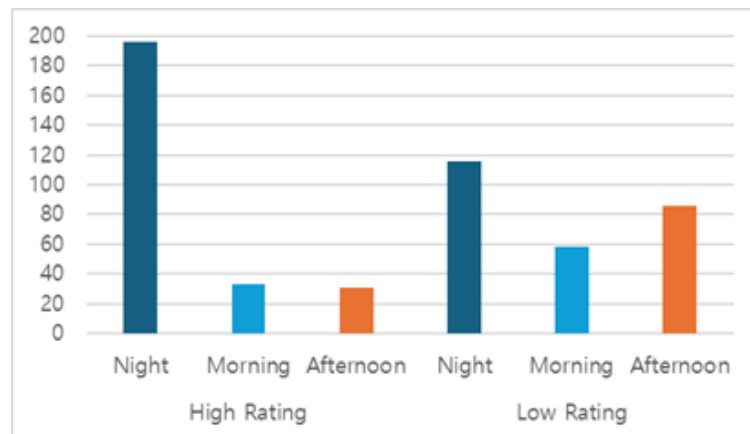
classical, Rap 집단에서 Low Rating의 비율이 유의미하게 높았다.



선호하는 음악 장르와 평점 평균

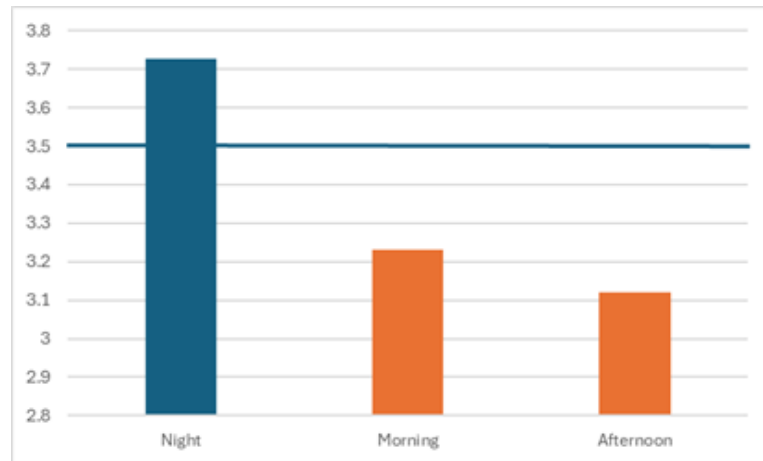
한편 해당 칼럼을 기준으로 평점의 평균을 분석한 결과 평점이 낮은 것으로 분류할 수 있는 장르는 Rap, classical, Rock, Old songs로 나타났다. Rock, Old songs의 경우 자료가 매우 적지만 Rap과 classical의 경우 유의미한 결과라고 할 수 있다. 따라서 Low Rating 집단과 classical, Rap 선호 집단의 관계가 검증되었다고 할 수 있다.

2) 음악 감상 시간대(music_time_slot)



평점과 음악 감상 시간대

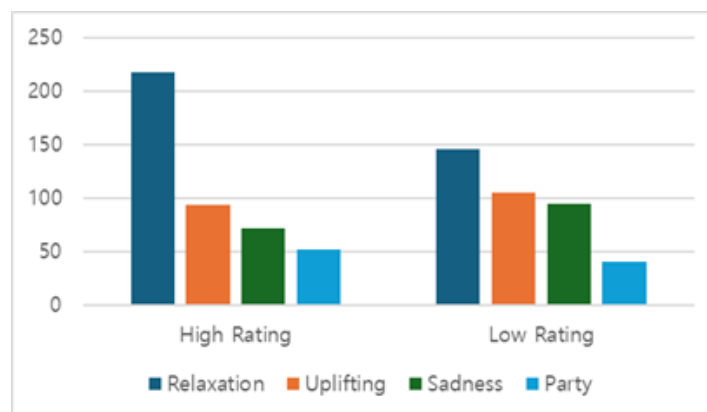
Low Rating 집단에서 High Rating 집단에 비해 Morning, Afternoon 사용자 비율이 유의미하게 높았다.



음악 감상 시간대와 평점 평균

실제로 해당 칼럼을 기준으로 평점 평균을 분석한 결과, Morning과 Afternoon 사용자들의 평점 평균이 3.5 미만으로 낮았다. 따라서 Low Rating 집단과 Morning, Afternoon 사용자 집단의 관계가 검증되었다고 할 수 있다.

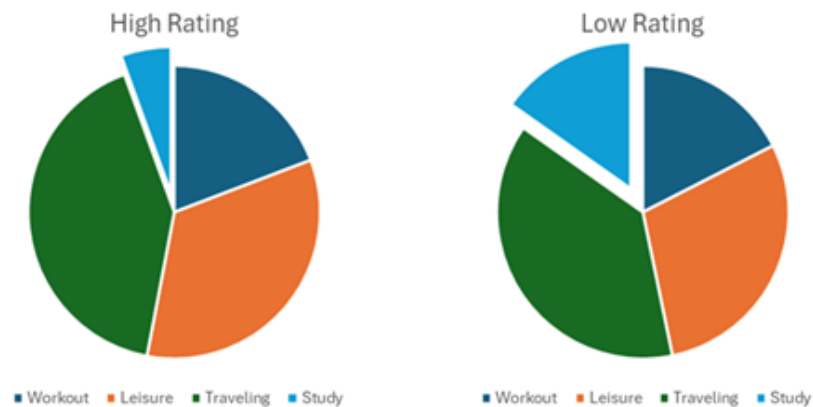
3) 음악 감상 시 감정(music_Influential_mood)



평점과 음악 감상 시 감정

Low Rating 집단에서 High Rating 집단에 비해 Uplifting and motivational(그래프상 표기는 Uplifting), Sadness or melancholy(그래프상 표기는 Sadness)의 비율이 높았다. 해당 칼럼의 경우 중복 선택이 가능했기 때문에 칼럼 기준 평점 분석을 통한 검증 단계는 제외하였다.

4) 음악 감상 시 상황(music_lis_frequency)

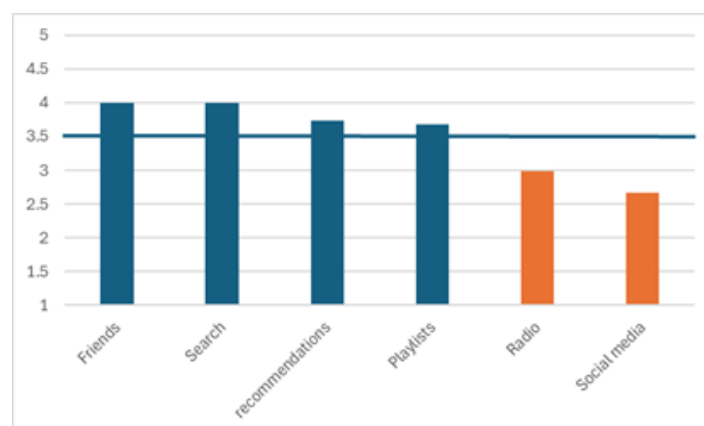


Low Rating 집단에서 음악 감상 시 상황이 Study인 비율이 유의미하게 높았다. 해당 칼럼의 경우 중복 선택이 가능했기 때문에 칼럼 기준 평점 분석을 통한 검증 단계는 제외하였다.

다른 칼럼 Age, Gender는 전체 사용자 집단, High Rating 집단, Low Rating 집단 간에 유의미한 분포 차이를 보이지 않았다.

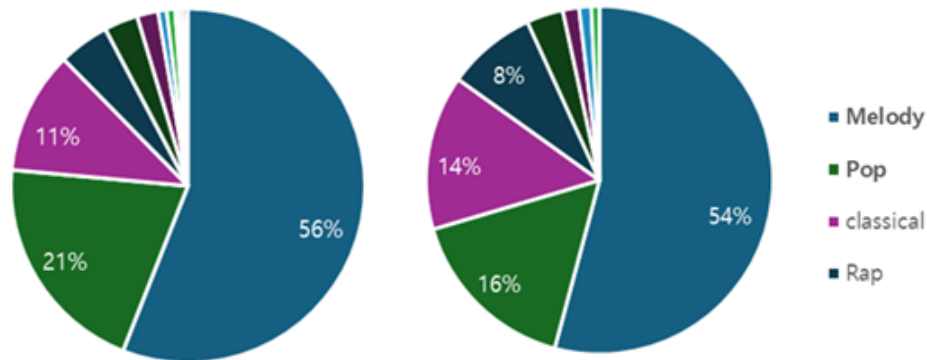
5) 추가 분석

추가적으로, 음악 탐색 방법(music_expl_method)에 따른 평점과 선호하는 음악 장르 분포는 다음과 같았다.

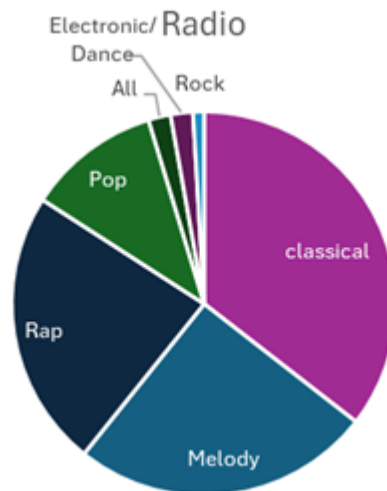


음악 탐색 방법과 평점 평균

먼저 칼럼 기준으로 평점의 평균을 분석한 결과 Radio를 사용하는 사용자들의 평균 평점이 3점 미만으로 낮은 것이 검증되었다. 이를 통해 Radio 사용자들은 추천 시스템 서비스에 만족하지 못해 Radio를 사용하고 있다고 할 수 있다.



좌) recommendations (우) Playlists 사용자의 선호 음악 장르



Radio 사용자의 선호 음악 장르

음악 탐색 방법으로 Radio를 사용하는 사용자의 경우 선호하는 음악 장르가 classical인 경우의 비율이 높고, recommendations와 Playlists를 사용하는 사용자 집단에 비해 Rap을 선호하는 경우의 비율도 높았다. 이는 선호하는 음악 장르(fav_music_genre)에 대한 분석과 일맥상통하는 결과로서, classical, Rap 선호 집단이 스포티파이 음악 추천 시스템에 만족하지 못하는 경향이 있고, classical 선호 집단이 추천 시스템을 활용하지 않는 주요 집단인 것으로 나타났다.

3. 추천 시스템 필터링 모델 개선 제안

평점을 기준으로 나눈 각각의 집단에서 현저한 분포 차이를 보인 지점이 스포티파이 음악 추천 시스템의 필터링 모델 개선 지점이라 할 수 있다. 스포티파이는 추천 시스템을 경쟁력으로 내세우고 있으며, 사용자들에게 다양한 사용 경험을 제공하기 위해 위 분석 내용에 따라 추천 시스템을 개선할 필요가 있다.

High Rating 집단은 추천 시스템에 충분히 만족하고 있는 것으로 판단하고, High Rating 집단에 대비되는 Low Rating 집단의 특성에 주목하여 추천 시스템 불만족 사용자들의 평점을 높일 수 있는 개선 방안을 제안하고자 한다.

1. 선호하는 음악 장르가 classical, Rap으로 분류되는 사용자의 경우 사용자 기반 협업 필터링보다 내용 기반 필터링에 중점을 두도록 하이브리드 시스템의 비중을 조절한다.
2. 음악 감상 시간대가 Morning, Afternoon일 경우 사용자 협업 필터링 과정에서 시간대 부문에 가중치를 둔다.
3. mood를 Uplifting and motivational, Sadness or melancholy 선택 시 사용자 협업 필터링 및 내용 기반 필터링 과정에서 mood 부문에 가중치를 둔다. 즉 같은 mood를 선택한 사용자, 음악의 mood 분류 등을 집중적으로 고려하도록 한다.
4. 음악 감상 상황을 Study 선택 시 사용자 협업 필터링 과정에서 상황 부문에 가중치를 둔다.

특히 클래식 음악 시장은 약 1%의 사용자가 존재하는 작은 시장으로서, 각종 스트리밍 앱에서 음악 제목이 잘려 나오거나, 아티스트 정보가 제대로 표기되지 않는 문제가 잦다. 한편 본 데이터의 기초 분석 결과에서는 선호하는 음악 장르가 classical인 사용자가 17%에 달했는데, 따라서 스포티파이 측에서 클래식에 대한 내용 기반 필터링 성능 강화에 투자할 가치가 있다.

3.3. spotify 프리미엄 구독자 증진

1. 분석 개요

Spotify는 다양한 사용자 경험을 제공하여 음악과 팟캐스트의 소비 방식을 혁신하고 있다. 그 중심에는 프리미엄 구독 서비스가 있으며, 이는 사용자에게 광고 없는 청취 경험, 오프라인 청취 기능, 높은 오디오 품질 등의 혜택을 제공한다. Spotify의 프리미엄 구독자 수 증가는 회사의 장기적인 수익성과 사용자 만족도를 향상시키는 핵심적인 전략이다.

본 분석의 목적은 Spotify의 프리미엄 서비스에 대한 사용자의 구독 의향을 극대화하기 위한 인사이트를 도출하는 것이다. 이를 위해, 프리미엄 구독자와 비구독자 간의 행동 패턴과 선호도를 비교 분석하여, 프리미엄 구독 의향에 영향을 미치는 주요 요인들을 파악한다. 이 과정에서, 데이터 분석을 통해 얻어진 통찰력을 기반으로, 특정 사용자 세그먼트에 초점을 맞춘 타겟 마케팅 전략을 제안한다. 이는 Spotify가 프리미엄 구독자 베이스를 확장하고, 사용자 경험을 개선함으로써 전반적인 서비스의 가치를 증진시킬 수 있는 방안을 모색하는 데 도움이 될 것이다. 분석을 통해 도출된 전략은 Spotify가 시장 내 경쟁력을 강화하고, 사용자에게 더 나은 음악 청취 경험을 제공하는 데 필수적인 요소가 될 것이다. 또한, 이러한 전략적 접근은 Spotify의 수익성 향상과 지속 가능한 성장을 이끌어낼 핵심 요인으로 작용할 것으로 기대된다.

2. 분석 방법 (프리미엄 구독 여부에 따라 차별화된 전략)

- 프리미엄 구독을 하고 있지 않은 경우(spotify_subscription_plan = Free), 프리미엄 구독 의향 (premium_sub_willingness)이 Yes인 경우와 No인 경우분석
- 프리미엄 구독을 하고 있는 경우(spotify_subscription_plan = Premium), 프리미엄 구독 지속 의향 Yes인 경우와 No인 경우 분석

3. 분석 가설

- 비구독자 그룹

원래 프리미엄 구독을 하지 않던 사용자들 중, 프리미엄 구독 의향이 '예'인 그룹(A)을 특정하여 이들의 연령, 성별, 가입 기간, 음악 추천 시스템에 대한 만족도 점수 등을 분석함으로써, Spotify가 프리미엄 구독 전환을 위해 어떤 사용자 세그먼트에 집중해야 하는지 알아낼 수 있다는 것(A는 프리미엄 구독 의향이 있는 비구독자 그룹을 지칭함)

- 구독자 그룹

현재 프리미엄 구독을 하고 있는 사용자들 중, 구독 연장 의향이 '예'인 그룹(B)을 특정하여, 이들의 연령, 성별, 가입 기간, 서비스 만족도(음악 추천 시스템 포함) 등을 분석함으로써, Spotify가 프리미엄 구독 연장을 위해 어떤 사용자 세그먼트에 집중해야 하는지 알아낼 수 있다는 것.(B는 프리미엄 구독 연장 의향이 있는 현 구독자 그룹을 지칭함)

4. 분석에 사용된 변수

Age / Gender / spotify_usage_period / preferred_premium_plan / music_recc_acting / pod_variety_satisfaction / music_time_slot

→ 단순히 유저의 음악적 취향만 알 수 있는 변수는 제외시킴

→ 해당 변수 선택 이유

- **Age:** 연령대는 사용자의 음악 취향, 구독 가능성, 그리고 서비스 사용 습관에 중요한 영향을 미친다. 연령대별로 다른 마케팅 전략과 커뮤니케이션 방법이 필요하기 때문에, 이 변수는 타겟팅 전략을 세울 때 핵심적인 역할을 한다. 예를 들어, 젊은 층은 새로운 기술과 서비스를 더 빠르게 받아들이며, 노년층은 보수적일 수 있으므로, 이에 맞춘 전략을 세우기 위해 연령 정보가 필요하다.
- **Gender:** 성별 역시 사용자의 콘텐츠 선호도와 구독 결정에 영향을 줄 수 있다. 특정 성별이 특정 유형의 음악이나 팟캐스트에 더 관심을 보일 수 있으며, 이는 마케팅 메시지와 프로모션 전략을 세분화하는 데 유용하다.
- **spotify_usage_period:** 사용 기간은 사용자의 서비스 충성도와 만족도를 반영할 수 있으며, 프리미엄 구독으로의 전환 가능성과도 관련이 있다. 사용 기간이 긴 사용자는 Spotify 생태계에 더 익숙하며, 이들을 대상으로 한 구독 전환 전략이 다를 수 있다.
- **preferred_premium_plan:** 사용자가 선호하는 프리미엄 요금제는 가격 민감도, 가족 구성원 수, 그리고 사용 패턴을 반영한다. 이 정보는 맞춤형 프로모션을 제공하거나 새로운 요금제를 개발하는 데 도움이 된다.
- **music_recc_acting:** 사용자가 Spotify의 음악 추천 시스템을 얼마나 활용하는지는 서비스 만족도와 직접적인 연관이 있다. 추천 시스템에 대한 긍정적인 반응은 프리미엄 구독의 가치를 높이는 요소가 될 수 있다.

- **pod_variety_satisfaction**: 팟캐스트 콘텐츠에 대한 다양성 만족도는 사용자가 프리미엄 구독을 고려할 때 중요한 요소이다. 다양하고 풍부한 콘텐츠 제공은 구독 결정에 긍정적인 영향을 미칠 수 있다.
- **music_time_slot**: 사용자들이 어떤 시간대에 음악 감상을 즐기는지 안다면 해당 시간대를 타겟한 효율적인 마케팅 전략을 생각해볼 수 있는 중요한 요소이다.

5. 데이터 분석 및 세분화 마케팅 전략

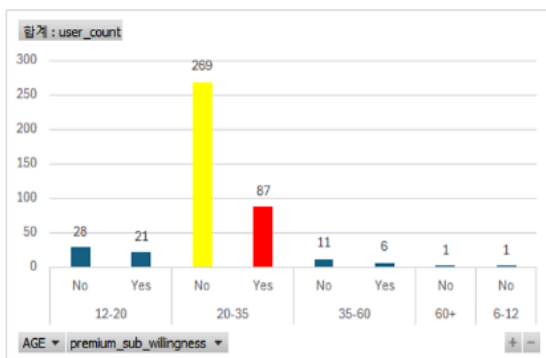
- **프리미엄 비 구독자** (spotify_subscription_plan = Free (ad-supported))

■ 특정 변수에 따른 구독 연장 여부

Age / Gender / spotify_usage_period / preferred_premium_plan / music_recc_acting / pod_variety_satisfaction / music_time_slot V.S premium_sub_willingness

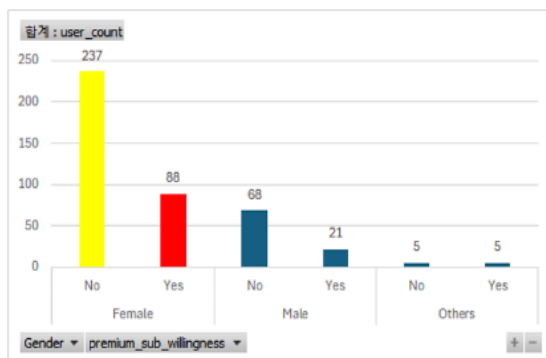
> 시각화

[Age]



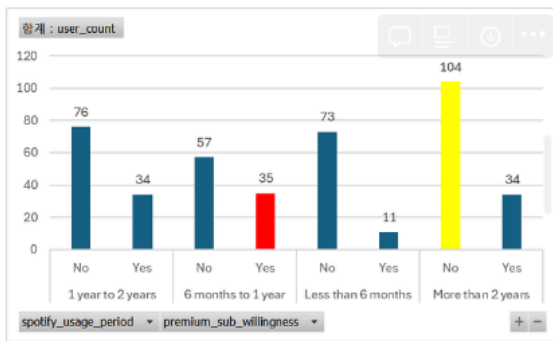
프리미엄 구독 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 20-35 나이대의 비중이 가장 높다.

[Gender]



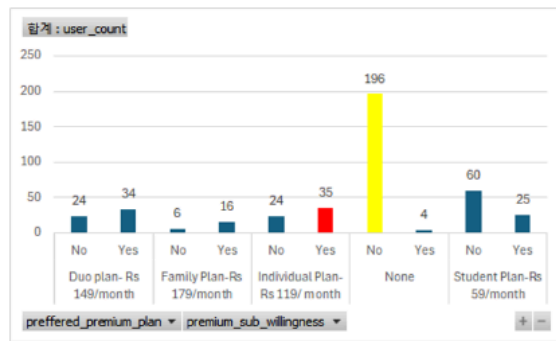
프리미엄 구독 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 여성의 비중이 가장 높다.

[spotify_usage_period]



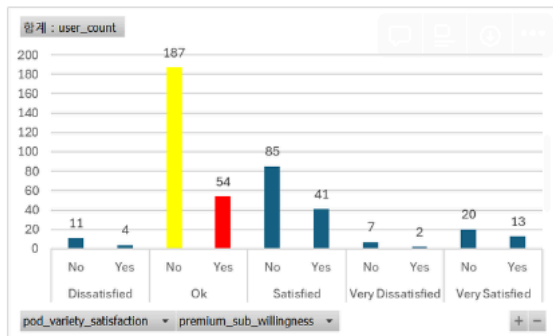
프리미엄 구독 의향이 있는 그룹에서는 가입 기간이 6개월~1년인 사람들의 비중이 가장 높고, 구독 의향이 없는 그룹에서는 2년 이상인 사람들의 비중이 가장 높다.

[preffered_premium_plan]



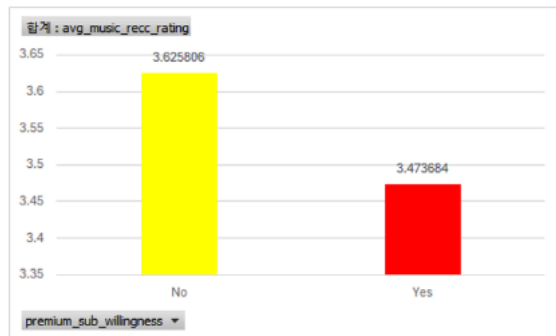
프리미엄 구독 의향이 있는 그룹에서 가장 선호하는 구독료는 'Individual Plan- Rs 119/month'이고, 구독 의향이 없는 그룹은 선호하는 구독료가 없는 사람의 비중이 가장 높다.

[free_pod_variety_satisfaction]



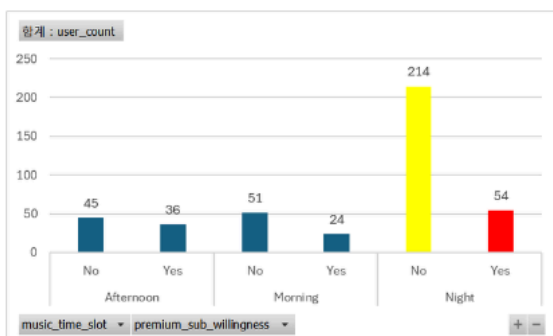
프리미엄 구독 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 팟캐스트 다양성에 대한 만족도는 'ok'다. 두 그룹 모두 매우 만족하지도 매우 만족하지 않는 것도 아니다.

[free_music_recc_rating]



프리미엄 구독 의향이 없는 그룹의 만족도가 더 높다.

[music_time_slot]



프리미엄 구독 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 가장 좋아하는 음악 감상 시간대는 밤이다.

> 시각화 결과 해석

추후에 프리미엄 구독 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 팟캐스트 다양성에 대한 만족도, 음악 감상 시간대, 연령대, 성별에서 유사한 패턴을 보인다.

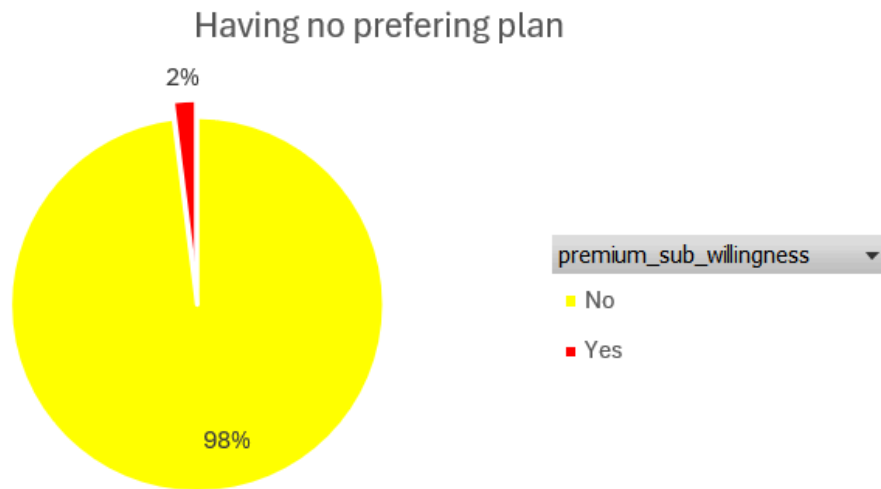
가입 기간이 6개월~1년에 속하는 사용자가 추후 프리미엄 구독 의향이 가장 높았고, 2년 이상인 경우 추후 구독 의향이 가장 낮았다. 이를 통해 가입 기간이 6개월~1년인 경우 구독 의향이 가장 높아지는 시기이며 가입 기간이 이미 2년이 지난 이후까지 프리미엄 가입이 되어 있지 않다면 앞으로도 계속 구독 의향이 없다고 추측할 수 있다. 또한, 가입 기간이 6개월~1년이 되어가는 사용자들을 대상으로 프리미엄 구독료에 관한 정보와 이벤트 알람을 주기적으로 전송해 프리미엄 구독자를 증폭시키는 마케팅 전략을 제안해 볼 수 있다.

‘Individual Plan- Rs 119/ month’ 구독료를 선호하는 사용자들이 프리미엄 구독 의향이 가장 높았고, 특별히 선호하는 구독료가 없는 사용자들의 경우 추후 구독 의향이 가장 낮았다. 이는 개인 사용자가 추후 구독 의향이 가장 높고, 추후 구독 의향이 없기 때문에 구독료에 관심을 가지지 않아 선호하는 구독료가 없는 사용자가 구독 의향이 가장 낮았다고 생각할 수 있다. 또한, ‘Individual Plan- Rs 119/ month’ 구독료를 선호하는 경향이 있으므로, 해당 요금에 대한 프로모션을 제공해 이 그룹을 연장할 수 있는 마케팅 전략을 제안해 볼 수 있다.

추후 프리미엄 구독 의향이 없는 그룹이 spotify 음악 추천 시스템에 대한 만족도가 더 높았다. 이를 통해, 프리미엄 비 구독자 그룹에서는 음악 추천 시스템은 구독 의향에 큰 영향을 미치지 않는다고 볼 수 있다.

따라서, spotify가 프리미엄 구독자를 증가시키기 위해 비구독자 그룹에서 집중해야 하는 사용자는 가입 기간이 6개월~1년인 경우, 선호하는 구독료가 'Individual Plan- Rs 119/ month'인 경우이다. 또한, 두 그룹 간 차이를 보이는 두 변수에 집중해 마케팅 전략을 고안해 볼 수 있다.

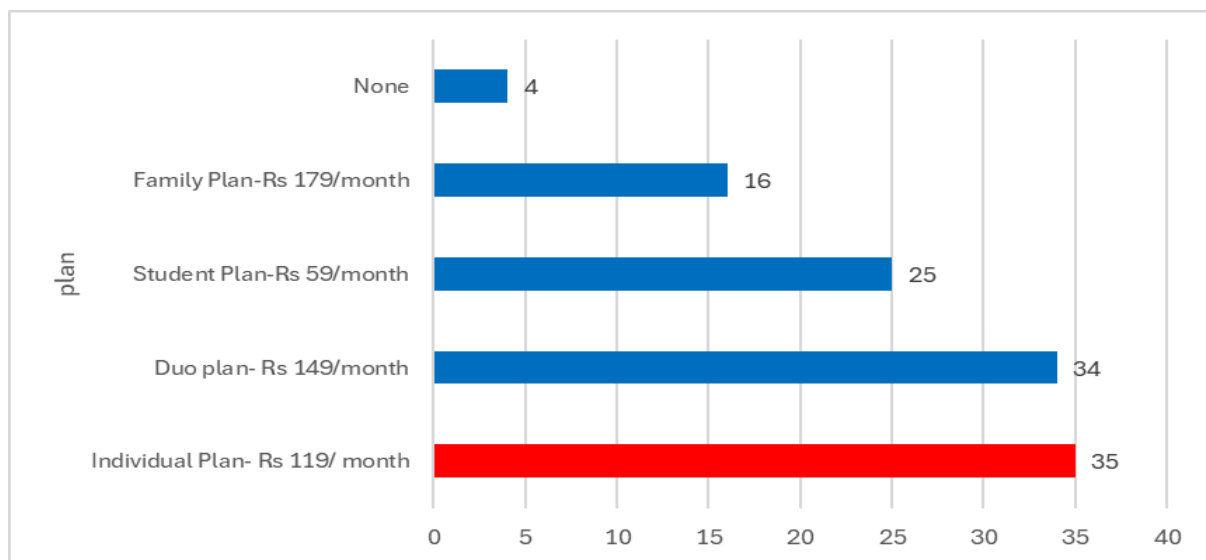
- 비 구독자 대상 마케팅 전략
 - 구독료 plan 의 영향력 파악



비 구독자를 대상으로 선호하는 플랜이 없는 사용자일수록 재 구독 의사가 없는 것으로 나타난다. 따라서, 새로운 구독료 플랜을 제시하거나 기존의 구독료 플랜 개선을 고려할 수 있다.

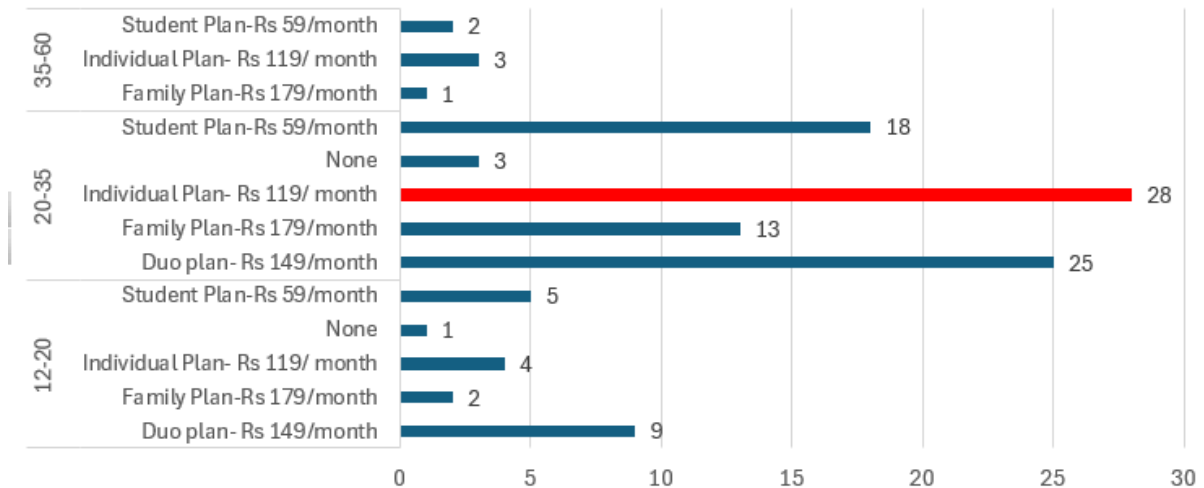
- 플랜의 사용자 분석

현재 무료(무료 구독)을 사용하고 있으며 프리미엄 구독 의향이 'Yes'인 Spotify 사용자들의 선호하는 프리미엄 플랜별로 사용자 수를 분석한다.



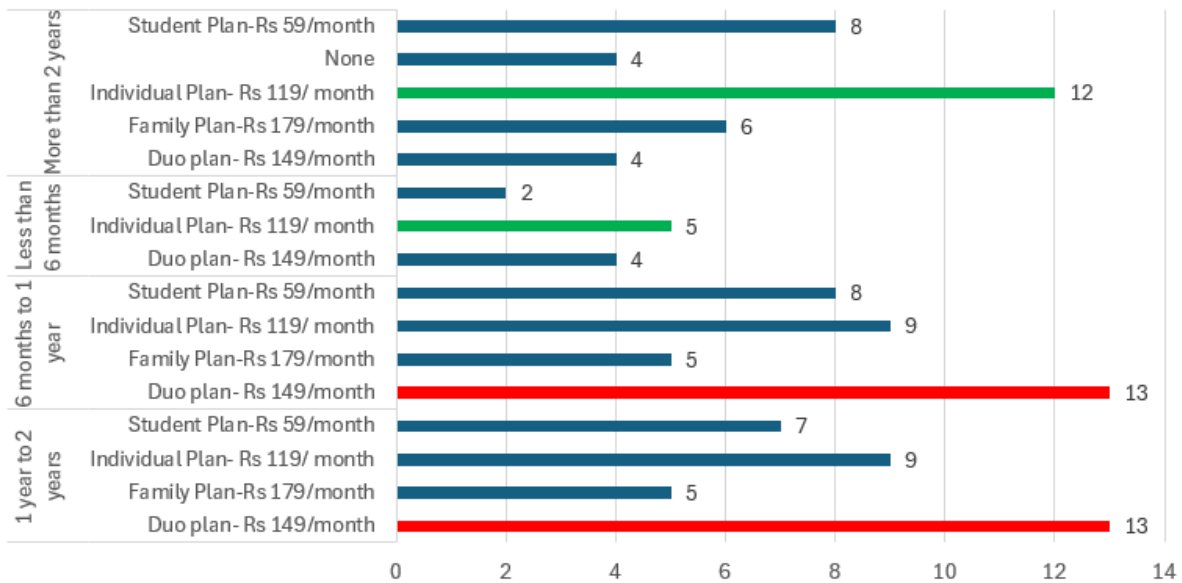
비 구독자를 대상으로 구독 의사가 있는 사람을 대상으로 분석해 보면 Individual plan을 가장 선호하는 그룹이 제일 많으며 근소한 차이로 Duo plan이 차등 순위에 위치한 것으로 보인다.

- 연령대를 타겟할 경우



12-20세 사용자는 Duo plan을 선호하는 경향이 있으며, 이는 개인 금액을 부담하기 어렵기 때문에 고려한 선택으로 보인다. 20-35세 사용자는 개인 소득이 생기는 연령대이기에 Individual plan을, 동시에 금액 부담을 줄이기 위해 반 씩 부담하는 Duo plan을 선호한다. 12-20세 사용자들은 금액상 59/Month으로 Duo plan보다 더싼 Student plan이 있지만 선호하지 않은 것으로 보아, Student plan에 아쉬운 점이 있다고 추측된다. 이러한 분석을 바탕으로 12-20세 사용자에게는 Student plan을 개선시켜 합리적인 가격임을 강조해 12-20세 사용자를 더욱 유입시키고, 20-35세 사용자에게는 Individual plan 혹은 Duo plan에 프로모션을 적용해 프리미엄 구독자를 증진시킬 수 있을 것이다. 35-60세 사용자 그룹에 대해서는 데이터가 상대적으로 부족해 결과의 신뢰성을 우려하여 결과를 도출하지 않기로 한다.

- 이용 기간으로 타겟할 경우



6개월에서 1년, 1년에서 2년 사이의 사용자에게는 Duo plan에 대한 promotion을 진행하고, 6개월 미만과 2년 이상의 사용자들 대상으로는 Individual plan에 대해 promotion을 진행한다면 구독 전환율을 증폭 시킬 수 있다. 혹은, 2년 이상의 사용자들이 가장 오랜 기간 동안 Spotify를 이용하고 있으므로, 그들은 이미 서비스에 만족하고 지속적으로 사용하고 있는 중요한 고객 세그먼트로 분류할 수 있다. 따라서 그들이 가장 선호하는 Individual plan을 spotify에서 가장 주력으로 홍보하고 권장함으로써 이들 고객 세그먼트를 유지하거나 확장할 수 있을 것으로 기대된다.

* 현 구독자와 다르게 비구독자 그룹에서 시간대별 변수는 추후 구독 의향이 있는 그룹과 없는 그룹의 패턴이 동일하고, 다른 변수와 조합해서 진행하는 분석이 아니기 때문에 비구독자를 대상으로 한 시간대별 분석은 하지 않는다.

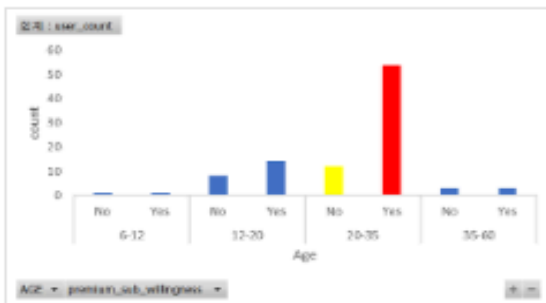
○ 프리미엄 구독자 (spotify_subscription_plan = Free (ad-supported))

■ 특정 변수에 따른 구독 연장 여부

Age / Gender / spotify_usage_period / preferred_premium_plan / music_recc_acting / pod_variety_satisfaction / music_time_slot V.S premium_sub_willingness

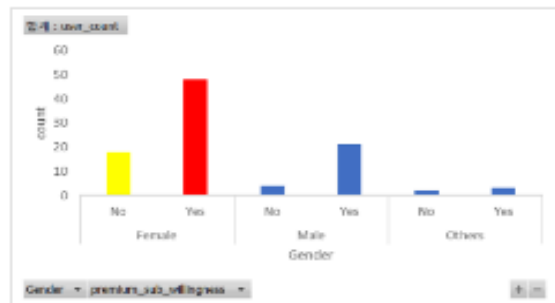
> 시각화

[Age]



프리미엄 구독 연장 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 20-35 나이대의 비중이 가장 높다.

[Gender]



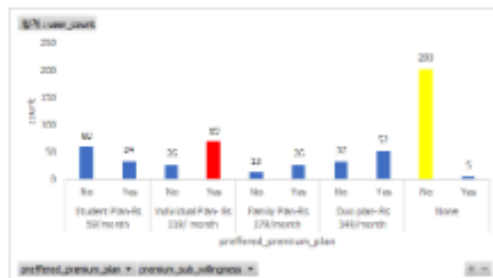
프리미엄 구독 연장 의향이 있는 그룹과 구독 의향이 없는 그룹 모두 여성의 비중이 가장 높다.

[spotify_usage_period]



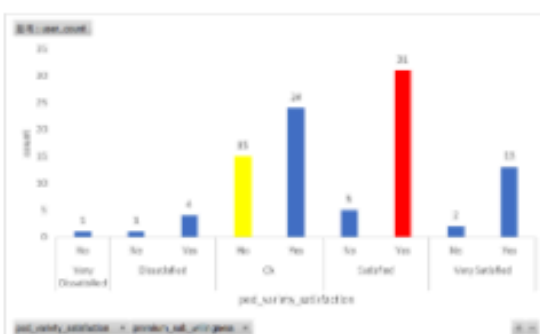
프리미엄 구독 의향이 있는 그룹에서는 가입 기간이 2년 이상인 사람들의 비중이 가장 높고 구독 의향이 없는 그룹에서는 1년에서 2년 사이가 가장 높다.

[preferred_premium_plan]



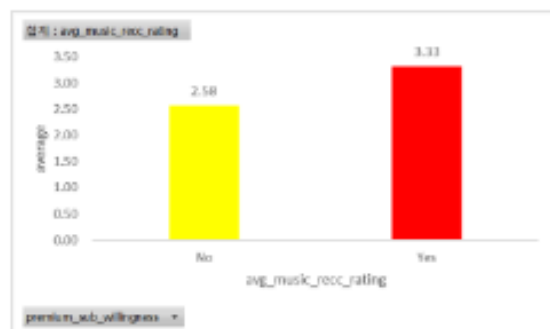
프리미엄 구독 연장 의향이 있는 그룹에서 가장 선호하는 구독료는 'Individual Plan- Rs 119/ month'이고, 구독 의향이 없는 그룹은 선호하는 구독료가 없는 사람의 비중이 가장 높다.

[pod_variety_satisfaction]



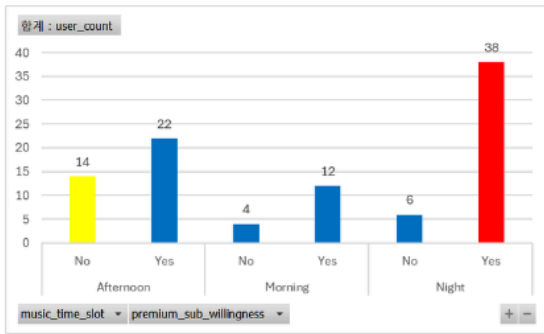
프리미엄 구독 연장 의향이 있는 그룹은 Satisfied 비율이 제일 많으며 프리미엄 구독 의향이 없는 그룹은 Ok인 비율이 가장 많다.

[avg_music_recc_rating]



프리미엄 구독 연장 의향이 있는 그룹이 음악의 추천 시스템의 만족도가 평균적으로 더 높다

[music_time_slot]



프리미엄 구독 연장 의향이 없는 그룹이 가장 선호하는 음악 감상 시간대는 오후이며, 밤 시간대에 사용하는 사용자는 거의 프리미엄 구독 연장 의향이 있다.

> 시각화 결과 해석

가입 기간이 증가함에 따라 프리미엄 구독을 계속하려는 사용자의 비율이 높아진다는 점을 확인할 수 있다. 이는 사용자의 서비스 충성도와 직접적인 연관이 있을 수 있으며, 장기 사용자를 대상으로 한 마케팅 전략이 중요함을 시사한다.

프리미엄 플랜에 대한 선호도가 없는 사용자들 사이에서 구독 연장 의향이 낮은 경향이 관찰된다. 반면, 'Individual Plan- Rs 119/ month'를 선호하는 사용자들 사이에서는 프리미엄 구독 연장 의향이 높은 것으로 나타났다. 이는 특정 가격대의 플랜이 사용자에게 매력적임을 나타내며, 이러한 선호도를 기반으로 한 마케팅 전략 개발이 유익할 수 있음을 제시한다.

팟캐스트의 다양성에 대해 'Satisfied'라고 응답한 사용자들이 프리미엄 구독을 연장할 경향이 더 높은 것으로 나타났다. 이는 콘텐츠의 질과 다양성이 사용자의 구독 결정에 중요한 요소임을 시사하며, 고품질의 다양한 콘텐츠 제공이 중요함을 강조한다.

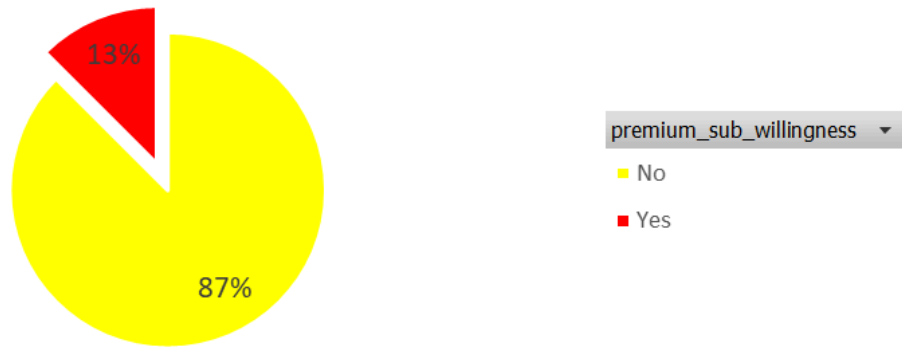
■ 현 구독자 대상 마케팅 전략

- 구독료 plan 의 영향력 파악

Spotify 데이터를 사용하여 현재 프리미엄(유료 구독) 사용자들의 선호하는 프리미엄 플랜과 구독 연장 의향 (premium_sub_willingness) 간의 관계를 분석. 구체적으로, 사용자들이 특정 프리미엄 플랜을 선호하는지, 아니면 특정 플랜에 대한

선호가 없는지(None)를 조사하고, 이러한 선호가 구독 연장 의사에 어떤 영향을 미치는지를 파악하고자 한다.

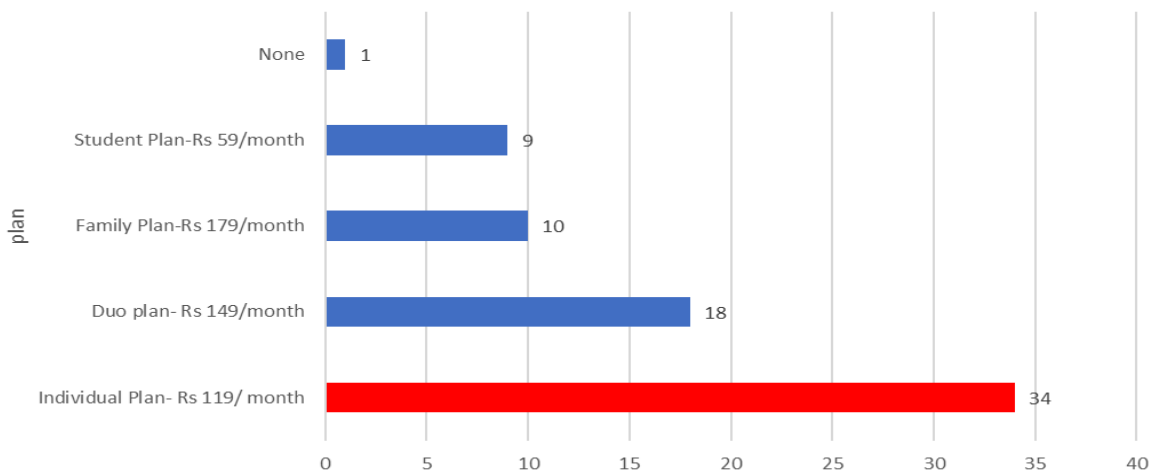
Having no preferring plan



데이터의 양이 현저히 적지만, 구독자를 대상으로 선호하는 구독료 플랜이 없는 사용자 일수록 13%로 재 구독 의사가 없는 것으로 나타난다. 따라서, 새로운 구독료 플랜을 제시하는 거나 기존의 구독료 플랜 개선을 고려 할 수 있다.

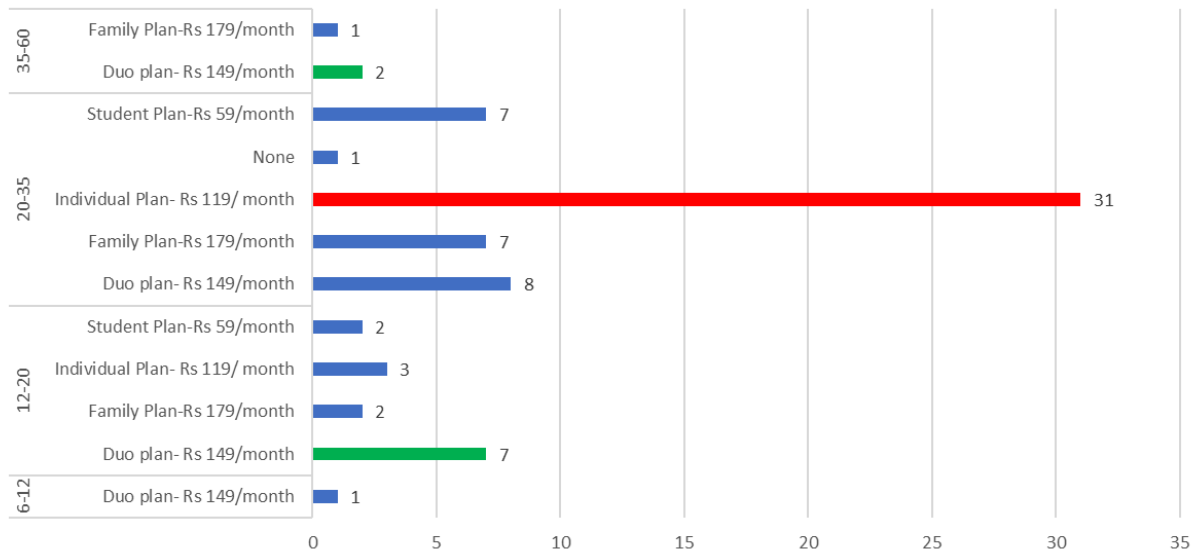
- 플랜의 사용자 분석

현재 프리미엄(유료 구독)을 사용하고 있으며 프리미엄 구독 의향이 'Yes'인 Spotify 사용자들의 선호하는 프리미엄 플랜별로 사용자 수를 분석한다.



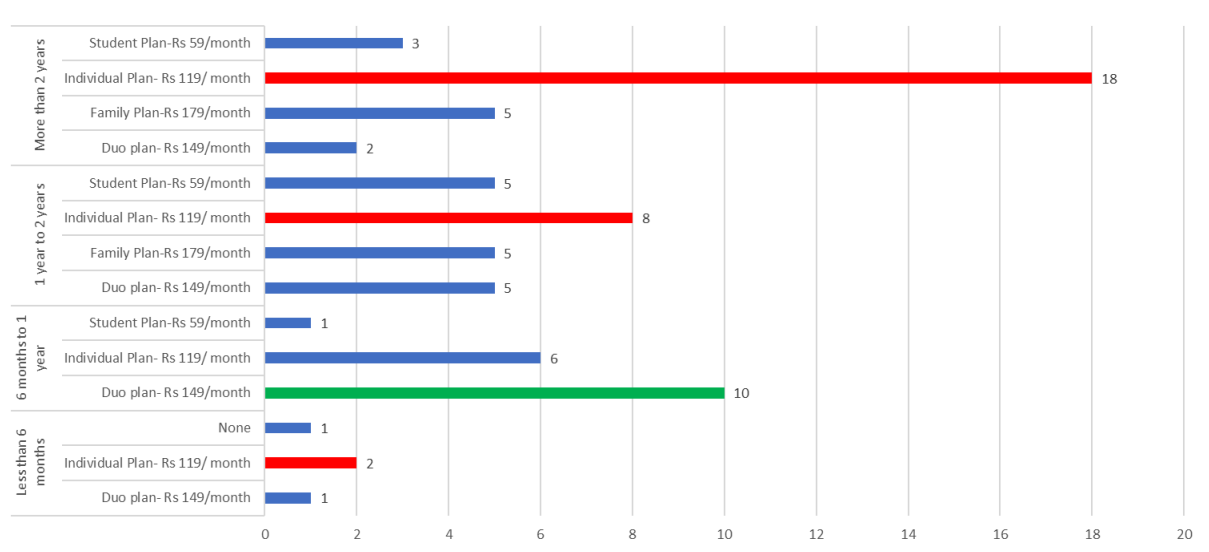
구독자를 대상으로 재 구독 의사가 있는 사람을 대상으로 분석해 보면 Individual plan을 가장 선호하는 그룹이 72명중 34명으로 제일 많으며 차등적으로 18명으로 Duo plan 인 것으로 보인다.

- 연령대를 타겟할 경우



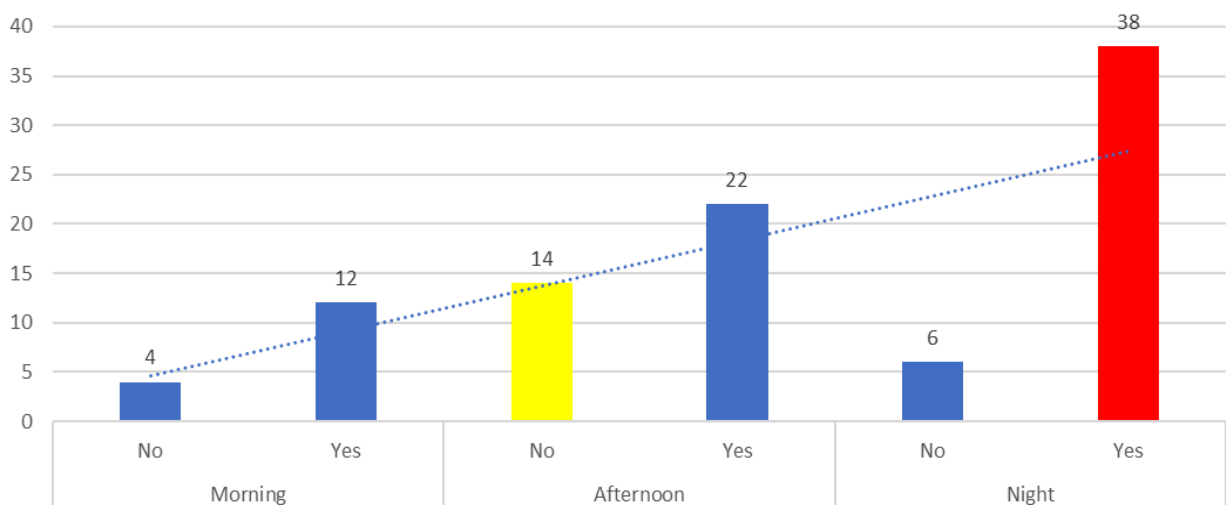
12-20세 사용자 그룹은 14명 중 7명이 Duo plan을 선호하는 경향이 있으며, 이는 주로 개인의 금전적 부담을 고려한 선택으로 보인다. 반면, 20-35세 사용자 그룹은 소득이 발생하기 시작하는 생애 단계에 있어 47명 중 31명이 Individual Plan을 선호한다. 특히 12-20세 그룹에서는, Duo Plan의 가격 대비 가치를 높이 평가하는 것으로 해석될 수 있으며, 학생을 위한 비교적 저렴한 Student Plan에도 불구하고 그 선호도가 낮은 것으로 나타나, Student Plan에 대한 플랜에 대한 아쉬움이 있을 가능성을 시사한다. 이러한 분석을 바탕으로, Student plan을 더욱 개선하거나, 12-20세 대상으로 Duo Plan, 20-35세 대상으로는 Individual Plan에 대한 프로모션을 강화함으로써 구독자의 연간 유지율을 증진시킬 수 있는 전략적 기회가 있음을 제시한다. 35-60세 및 6-12세 사용자 그룹에 대해서는 데이터가 상대적으로 부족하여 명확한 결론을 도출하기 어려운 상황이다.

- 이용 기간으로 타겟 할 경우

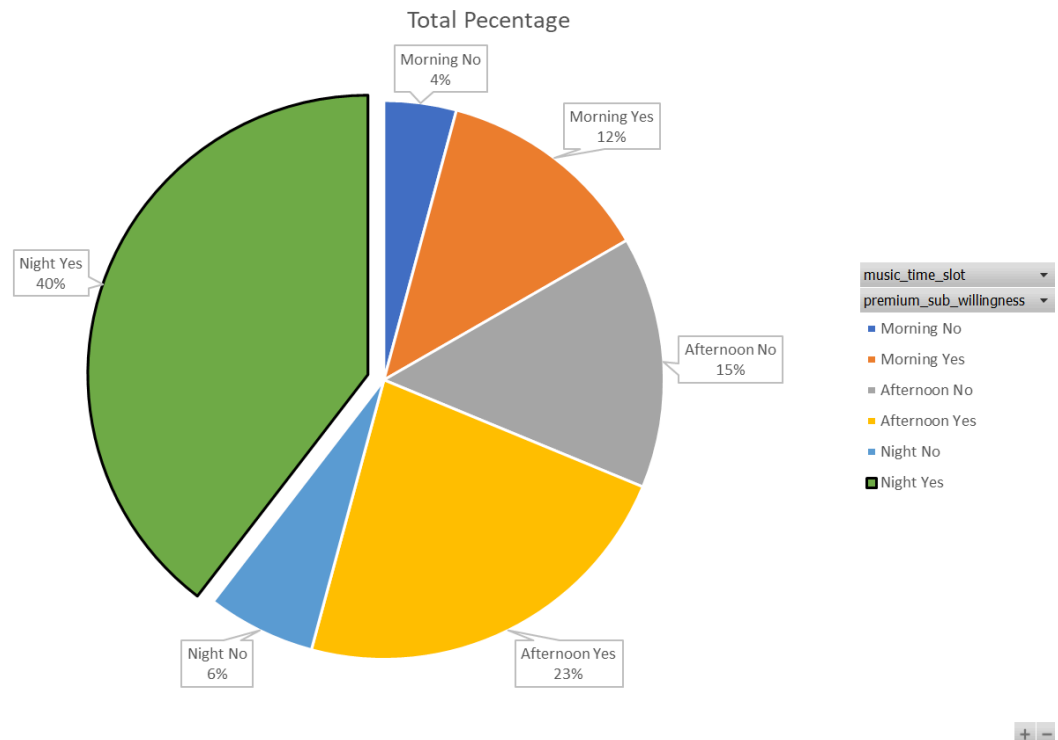


2년 이상의 장기 사용자들에게는 'Individual Plan'을 중심으로 한 마케팅 활동을 강화하고, 해당 요금제의 특별 혜택을 강조하여 충성도를 높이고 구독 연장을 장려해야 한다. 1년에서 2년 사이의 사용자들에게는 다양한 요금제 옵션을 제공하여, 'Individual Plan' 이외의 다른 플랜들도 적극적으로 홍보하는 전략이 필요하다. 6개월 미만의 신규 사용자들에게는 'Individual Plan'과 'Duo plan'을 대상으로 하는 시험적인 프로모션을 제공하여, 서비스의 가치를 경험할 수 있는 기회를 마련해야 한다.

- 프로모션 시간대 선정



합계 : willingness_percentage_total



Afternoon (오후) 시간대에는 사용자 중 약 61.11%가 프리미엄 구독 의향이 'Yes'이며, 이는 해당 시간대에 22명의 사용자가 해당된다. 'No' 응답을 한 사용자는 38.89%로 14명이다. Morning (오전) 시간대에는 사용자 중 75%가 프리미엄 구독 의향이 'Yes'로 나타나며, 이는 12명의 사용자가 해당됩니다. 'No' 응답은 25%로 4명의 사용자이다. Night (밤) 시간대에는 사용자 중 86.36%가 프리미엄 구독 의향이 'Yes'이며, 이는 전체의 약 39.58%에 해당하는 38명의 사용자이다. 'No' 응답은 13.64%로, 6명의 사용자가 해당된다.

이 데이터를 통해 알 수 있는 것은 밤 시간대에 프리미엄 구독에 대한 의향이 'Yes'인 사용자의 비율이 (40%) 다른 시간대에 비해 가장 높다는 것이다. 이는 밤 시간대가 Spotify 사용자에게 프리미엄 구독 서비스의 가치를 전달하기에 가장 유리한 시간임을 시사한다.

따라서, 마케팅 전략으로서는 밤 시간대에 프로모션 활동을 집중하는 것이 효과적이다. 이 시간대에 사용자들이 Spotify를 더 많이 사용하고 구독에 대한 긍정적인 태도를 보이는 만큼, 다음과 같은 마케팅 전략을 고려할 수 있습니다

타겟 광고 캠페인: 밤 시간대에 소셜 미디어, 모바일 앱, 웹 플랫폼을 통해 타겟 광고를 실시한다.

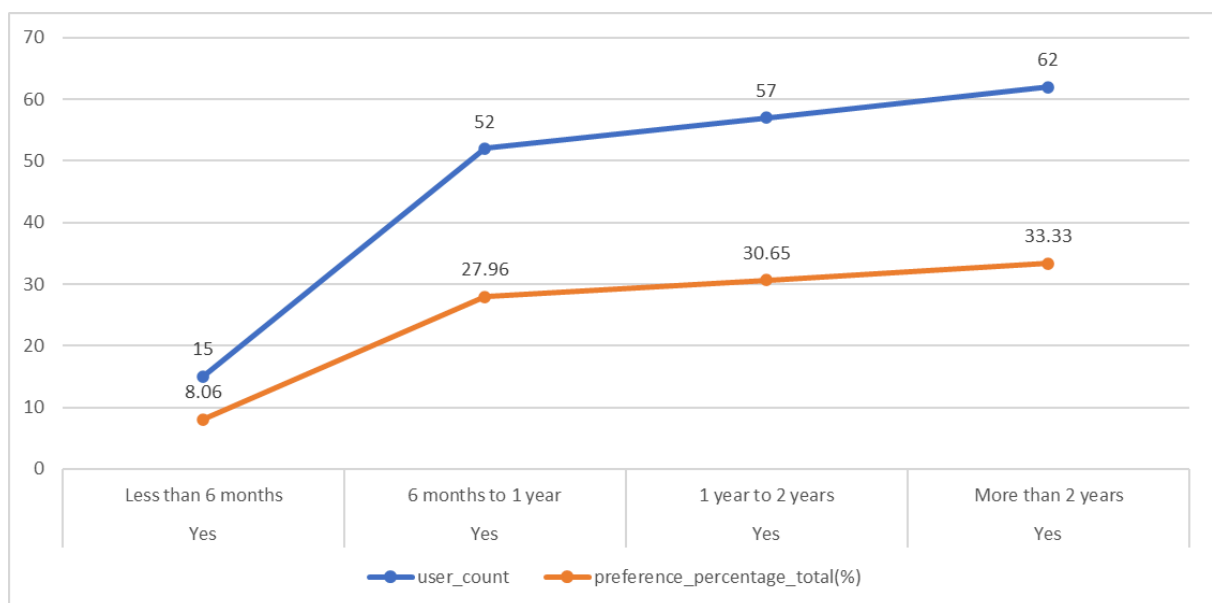
특별 할인 및 프로모션: 밤 시간대에만 사용 가능한 특별 할인이나 프로모션 코드를 제공하여 프리미엄 구독 전환을 유도한다.

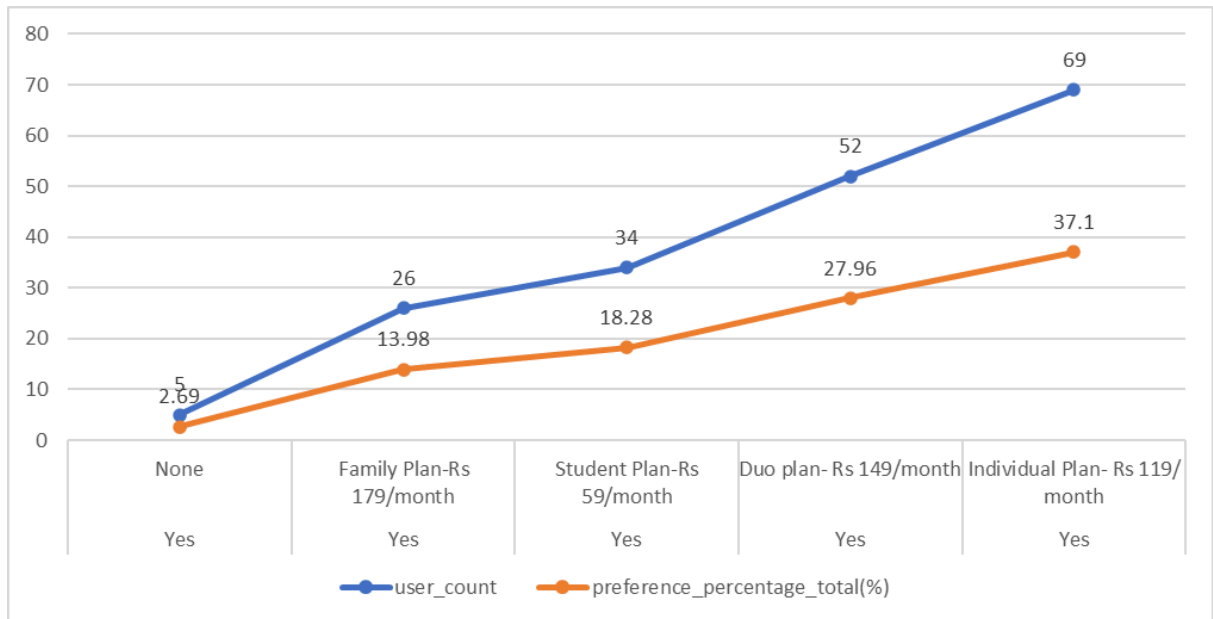
개인화된 추천 플레이리스트: 밤 시간대 사용자에게 특별히 제작된 개인화된 플레이리스트를 추천하여 프리미엄 구독의 가치를 강조한다.

○ 잠재 고객 탐색 및 통합 마케팅 전략 형성

스포티파이의 '가입 기간' 및 '선호하는 구독료' 변수 분석을 통해 얻은 인사이트는 향후 프리미엄 구독 의향이 있는 사용자 그룹과 그렇지 않은 사용자 그룹 사이에 분명한 패턴이 나타나지 않았음을 보여준다. 이러한 발견은 스포티파이에게 특정 변수에 초점을 맞춘 차별화된 프리미엄 구독 마케팅 전략을 개발할 수 있는 독특한 기회를 제공한다.

프리미엄 구독에 대한 관심이 있는 신규 사용자 그룹과 구독 연장 의사가 있는 기존 프리미엄 사용자 그룹을 '가입 기간'과 '선호하는 구독료'를 기반으로 세분화함으로써, 각 그룹에서 상대적으로 가장 높은 비율을 차지하는 주요 값을 배제한 후, 그 다음으로 높은 비율을 나타내는 사용자들을 잠재 고객으로 식별하였다. 이 잠재 고객군을 대상으로 한 타겟 마케팅 전략은 프리미엄 구독 전환율을 향상시킬 수 있는 잠재력을 가지고 있다.





가입기간에 대해서는 1~2년이 가입기간에 속하는 사용자를 30.65%로 잠재 고객으로, 선호하는 구독료에서는 27.96%로 Duo plan- Rs 149/month 구독료를 선호하는 사용자들을 잠재 고객으로 볼 수 있다. 따라서, 잠재 고객에게 특히 집중한다면 spotify의 프리미엄 구독율을 상승 시킬 수 있을 것이라고 기대된다.

이러한 전략적 접근은 스포티파이가 사용자 선호와 행동 패턴을 보다 세밀하게 분석하고, 이를 통해 맞춤형 프리미엄 구독 옵션을 제공함으로써 고객 만족도를 높이고 구독자 기반을 확대하는 데 중요한 역할을 할 것이다. 잠재적인 통합 고객 탐색의 중요성은 단순히 사용자 기반을 확장하는 것을 넘어서, 지속 가능한 성장과 시장 내 경쟁력 강화에 기여한다는 점에서 매우 크다. 이러한 분석과 전략 형성 과정을 통해 스포티파이는 다가오는 변화에 능동적으로 대응하며, 사용자들에게 더욱 개인화되고 만족스러운 경험을 제공할 수 있을 것이다.

4. 결론과 향후 방향(= 기대효과)

본 프로젝트를 통해 Spotify 사용자 행동 데이터를 심층적으로 분석함으로써, 사용자들의 선호도와 행동 양상에 대한 깊은 이해를 바탕으로 한 다양한 인사이트를 도출하였다. 이러한 분석을 통해, Spotify 음악 추천 시스템의 개선 방향과 프리미엄 구독 증진을 위한 마케팅 전략을 제안하였다. 본 결론과 향후 방향은 Spotify의 서비스 개선과 시장 경쟁력 강화에 중요한 가이드라인을 제공할 것으로 기대한다.

● 결론

Spotify 사용자 행동 데이터 분석은 사용자의 연령, 성별, 사용 기간 등의 인구 통계학적 정보와 음악 취향, 청취 장치 선호도 등 행동 데이터를 통합 분석함으로써 명확한 사용자 세그먼트를 확인했다. 이러한 세분화된 사용자 그룹의 존재는 Spotify가 타겟 마케팅 전략을 수립하는 데 있어 핵심적인 기반이 되며, 마케팅 메시지와 캠페인을 더욱 효과적으로 설계할 수 있도록 한다. 또한, 사용자 피드백과 상세한 분석 데이터를 바탕으로 음악 추천 시스템에서 개선할 수 있는 다양한 방안이 존재한다는 것이다. 특히 사용자의 개별 선호도, 감정 상태, 청취하는 환경을 세심하게 고려하여 개인화된 추천을 제공하는 알고리즘의 개발이 중요하다고 강조되었다. 이는 사용자 개인의 경험을 풍부하게 만들고, 서비스에 대한 만족도를 높이며, 최종적으로 사용자의 장기적인 충성도를 증진시키는 데 기여할 것이다. 더불어, 비 구독자와 현재 구독자를 명확하게 구분함으로써, 프리미엄 구독으로의 전환 가능성이 높은 사용자 그룹을 효과적으로 식별할 수 있었다. 이 그룹에 대한 맞춤형 프로모션 전략을 개발하고 실행함으로써, 프리미엄 구독자 베이스를 확장하는 데 중점을 두었다. 예를 들어, 특정 사용자 그룹의 청취 패턴과 선호도에 기반한 맞춤형 구독 혜택을 제공하거나, 프리미엄 구독에 대한 독점 콘텐츠 접근성을 강화하는 등의 전략이 제안되었다.

● 향후 방향 및 기대효과

1. 데이터 기반의 지속적인 서비스 개선: Spotify는 사용자 데이터를 지속적으로 분석하고, 이를 통해 서비스의 다양한 측면을 개선해야 한다. 예를 들어, 사용자의 감정 상태나 활동에 맞춘 음악 추천, 특정 콘텐츠에 대한 선호도가

높은 세그먼트를 위한 전용 플레이리스트 제공 등이다. 이러한 개선은 사용자 만족도를 높이고, 브랜드 충성도를 강화하는 효과를 가져올 것이다.

2. **고도화된 타겟 마케팅 전략의 실행:** 사용자 세분화에 기반한 맞춤형 마케팅 전략을 시행함으로써, 더 높은 프리미엄 구독 전환율을 달성할 수 있을 것이다. 특히, 프리미엄 구독에 대한 관심이 높은 사용자 세그먼트를 정확히 타겟팅하여, 이들에게 매력적인 구독 옵션과 혜택을 제공해야 한다.
3. **사용자 참여 및 커뮤니티 활성화:** 추천 시스템의 개선과 맞춤형 마케팅 전략은 사용자의 Spotify 플랫폼 내 참여를 증진시킬 것이다. 이는 사용자의 콘텐츠 소비를 늘리고, 커뮤니티 내에서의 상호작용을 촉진시키며, 장기적으로는 Spotify의 마켓 포지션을 강화할 것이다.
4. **신규 시장 및 기회 탐색:** 분석을 통해 식별된 사용자의 새로운 니즈와 행동 패턴은 Spotify에게 새로운 서비스 개발과 시장 확장의 기회를 제공할 것이다. 이는 Spotify의 지속 가능한 성장을 위한 중요한 기반을 마련할 것이다.

종합적으로 볼 때, 이번 프로젝트는 Spotify가 사용자 중심의 접근 방식을 통해 음악 추천 시스템을 개선하고, 프리미엄 구독 전환율을 높이며, 전반적인 사용자 경험을 향상시키는 데 필수적인 전략적 통찰력을 제공한다. 지속적인 데이터 분석과 사용자 피드백의 활용은 Spotify가 스트리밍 시장에서 지속적인 경쟁 우위를 확보하고, 사용자에게 더 나은 가치를 제공하는 데 중추적인 역할을 할 것으로 기대된다.