

무지컬 곡마다.

혜화역4번출구

서영우 신진섭 신지민 한재선

목차

· 분석 배경 ·

분석배경
분석의 필요성

· 분석 과정 ·

분석개요
분석과정(군집분석)
분석내용(추천시스템)

· 활용 가능성 ·

활용가능성 및 기대효과

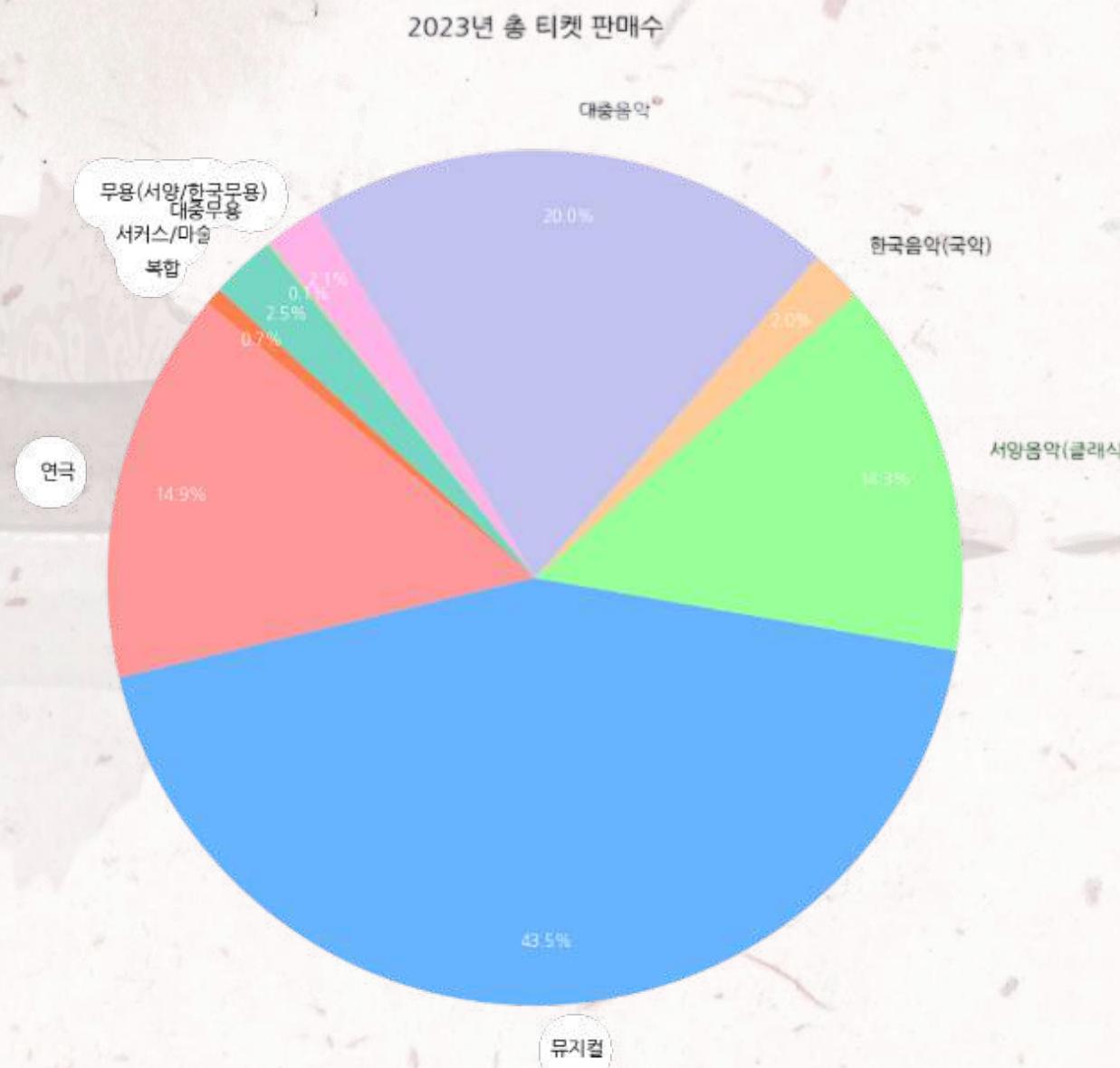
01.

제 1 장

분석 배경

- 국악장르 활성화 방안 모색
- 상호 추천 전략 필요성

분석배경



대부분을 차지하는 뮤지컬

가장 큰 비중을 차지하는 뮤지컬과 대중공연.

-> 뮤지컬 장르와 관객 분석을 통해
공연예술 분야 발전 방안 모색

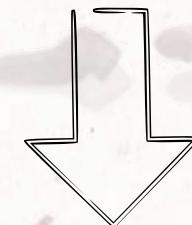
-> 비인기 장르 중 국악 장르 활성화

| 문석배경



어떻게 국악장르를 활성화시킬 수 있을까?

+ 뮤지컬 관객의 특성 활용



뮤지컬, 국악뮤지컬, 국악의 상황 분석

상호 연관성을 통한 관객 유치 극대화



분석배경

뮤지컬/국악 장르 특성

〈표 2-2〉 관람장르 중복 이용 현황

| | 단일 관람 | 중복 관람 | 계 |
|-------------|----------------------|---------------------|----------------|
| 연극 | 8,868 (53.0) | 7,849 (47.0) | 16,717 (100.0) |
| 뮤지컬 | 57,220 (79.5) | 14,731 (20.5) | 71,951 (100.0) |
| 서양음악(클래식) | 12,088 (57.8) | 8,824 (42.2) | 20,912 (100.0) |
| 한국음악(국악) | 1,227 (53.8) | 1,052 (46.2) | 2,279 (100.0) |
| 무용(서양/한국무용) | 4,090 (47.8) | 4,472 (52.2) | 8,562 (100.0) |
| 복합 | 515 (57.6) | 379 (42.4) | 894 (100.0) |

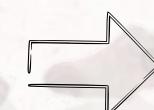
〈그림 2-1〉 관람장르 중복 이용 현황



뮤지컬 단일 관람 비율 79.5 % 중복관람비율 20.5%

국악 단일 관람 비율 53.8% 중복관람비율 46.2%

-> 국악은 다른 장르와의 교차 관람이 상대적으로 활발



국악뮤지컬과 국악 간의 상호 시스템 도입

"국악 뮤지컬"

- 뮤지컬 관객에게 국악뮤지컬을 추천하고, 국악뮤지컬 관객에게는 전통 국악 공연을 추천함으로써 장르 간의 자연스러운 확장과 균형 잡힌 성장을 유도

분석의 필요성

비인기 장르 활성화

비인기 장르,
특히 국악의 활성화 필요성

뮤지컬 강세

새로운 고객을 설득하는 것
보다 뮤지컬 고객을 국악으
로 유입하는 것이 효과적

국악장르 특징

국악 고객들은 다양한 장르
를 중복해서 관람하는 비율
이 상대적으로 높음 -> 타
장르에도 긍정적인 영향

뮤지컬관객에 국악뮤지컬을 추천하고,
국악뮤지컬 관객에게는 순수 국악 공연을 추천함으로써
관객들이 자연스럽게 다양한 전통 예술에 관심을 가질 수 있도록 유도

02

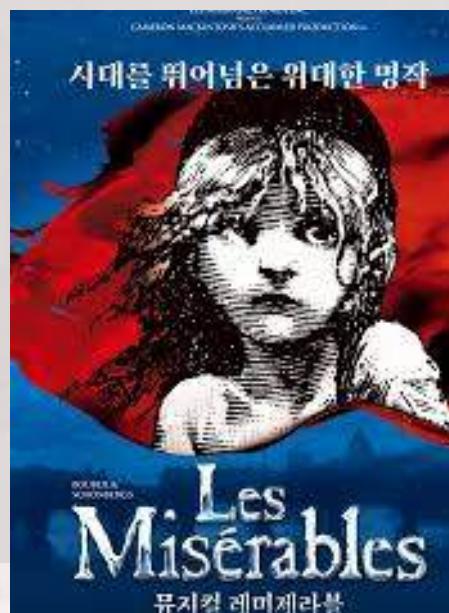
제 2 장

분석 과정

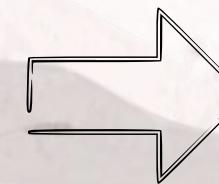
- 분석개요 및 데이터 전처리 과정
- 분석 내용 및 결과

| 분석 개요

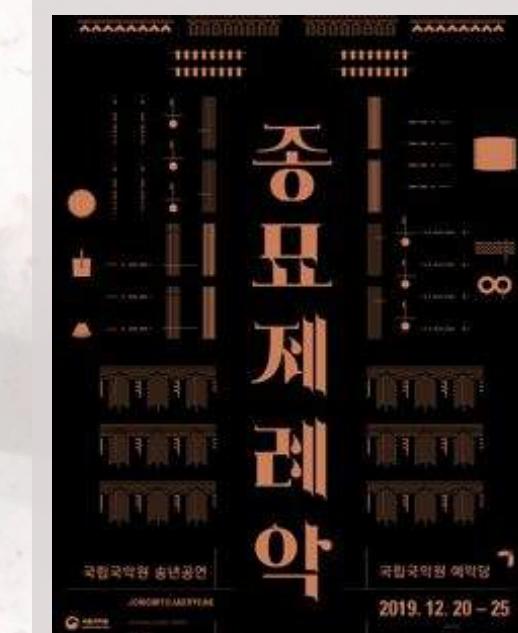
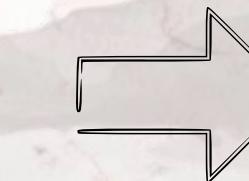
분석 개요



· 뮤지컬 ·



· 국악뮤지컬 ·



· 국악 ·

뮤지컬 관람객이 바로 국악을 관람하도록 유도하는 것은
현실적으로 진입장벽이 높다는 점을 고려해,
국악뮤지컬을 매개로 하는 단계적 유도 방식

분석과정

데이터 전처리

1. 데이터의 복잡성과 규모를 고려 -> '뮤지컬', '국악' 데이터

2. '뮤지컬'에서 세부장르명으로 표기된 '악극'

+

2023년 국악 뮤지컬 목록을 조사하여 데이터를 함께 추출 -> '국악뮤지컬' 데이터

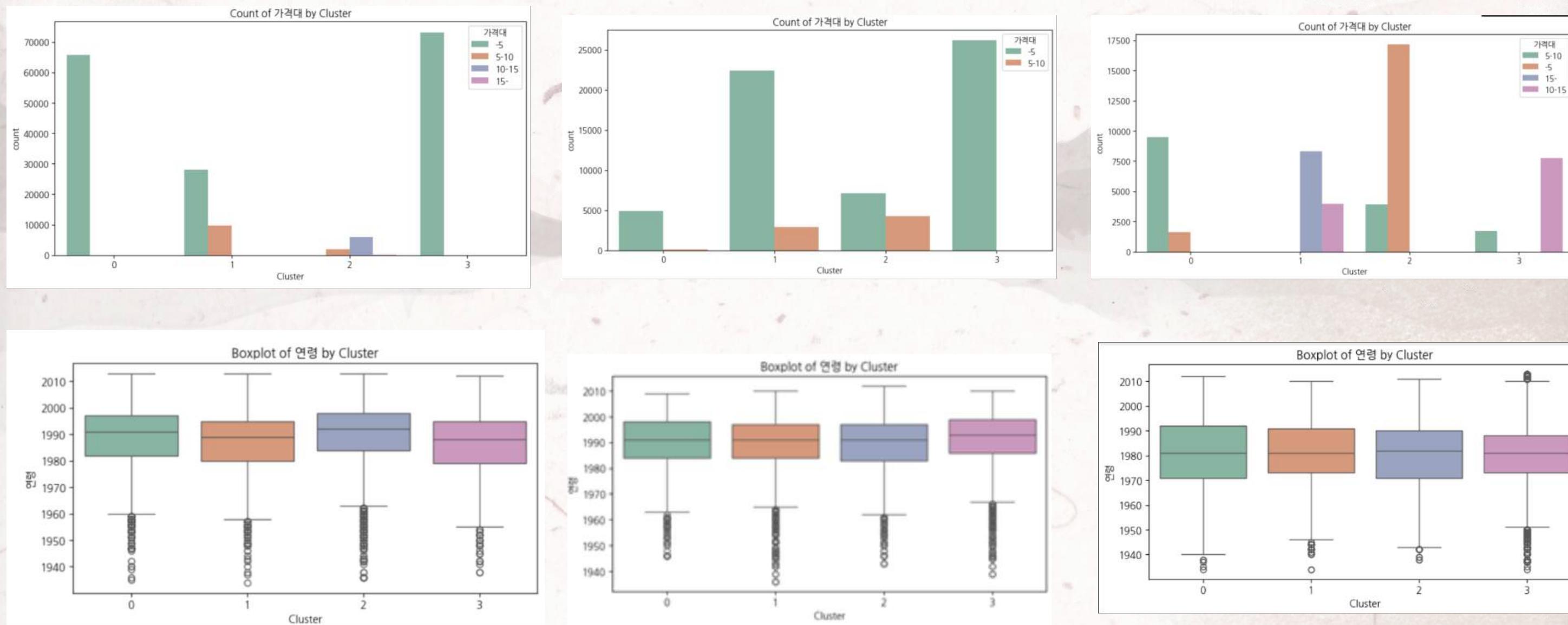
3. 새로운 column 생성 -> 범주형 변수를 카테고리화한 '공연지역명_코드', '예매/취소금액'을 그룹화한 '가격대'

4. 결측치 처리와 이상치 제거를 통해 데이터의 신뢰성 확보

분석과정

군집분석(k-prototype)

뮤지컬, 국악뮤지컬, 국악 장르 개별 군집 분석 후 공통점 파악



분석과정

군집분석(k-prototype)

뮤지컬 - 국악뮤지컬

두 장르 모두 티켓 가격이 주로 ‘5’와 ‘5-10’ 구간에 집중 -> 가격에 민감

평균 연령 20대 - 30대

대규모 공연장, 수도권 선호

-> 두 집단이 거의 비슷한 형태

국악뮤지컬 - 국악

중극장과 대극장을 선호 그중에서도 오케스트라 피트를 활용한 연주 선호

낮은 가격대의 티켓, 할인이 적음

평균연령 30대 - 40대, 여성

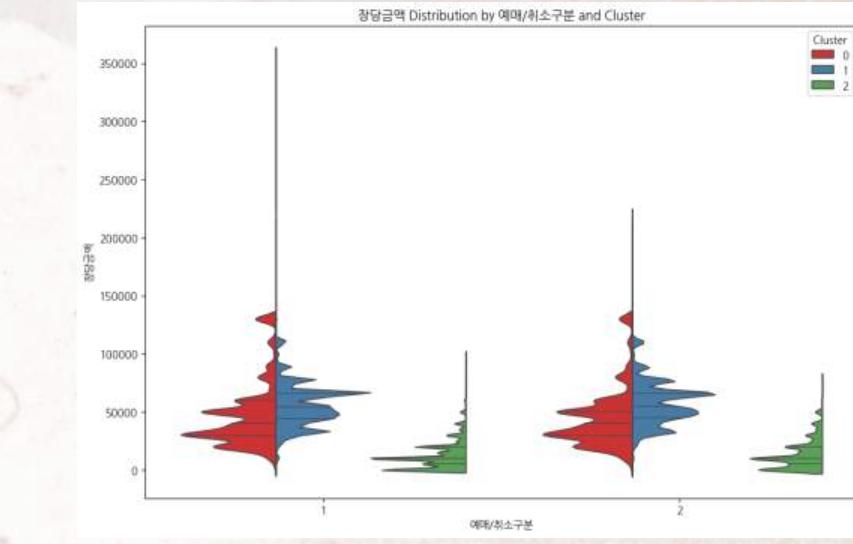
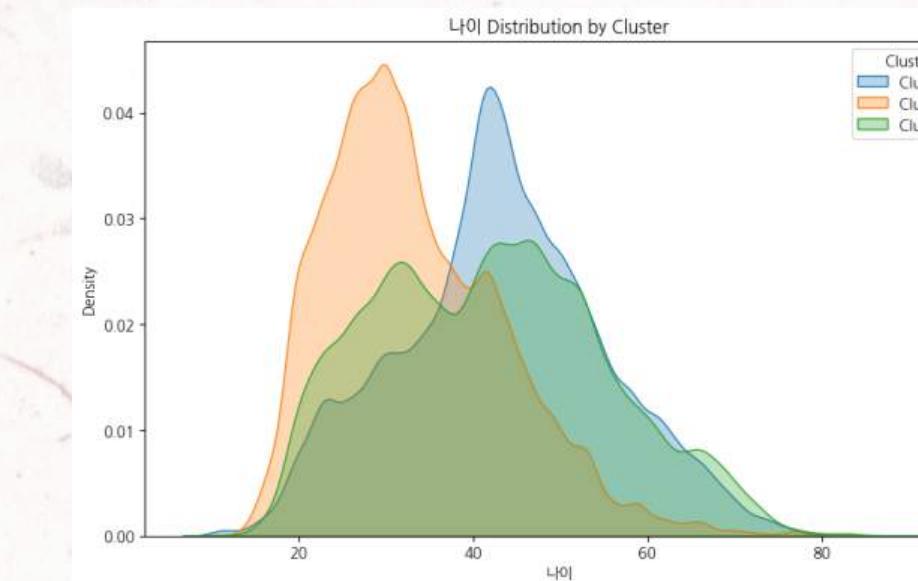
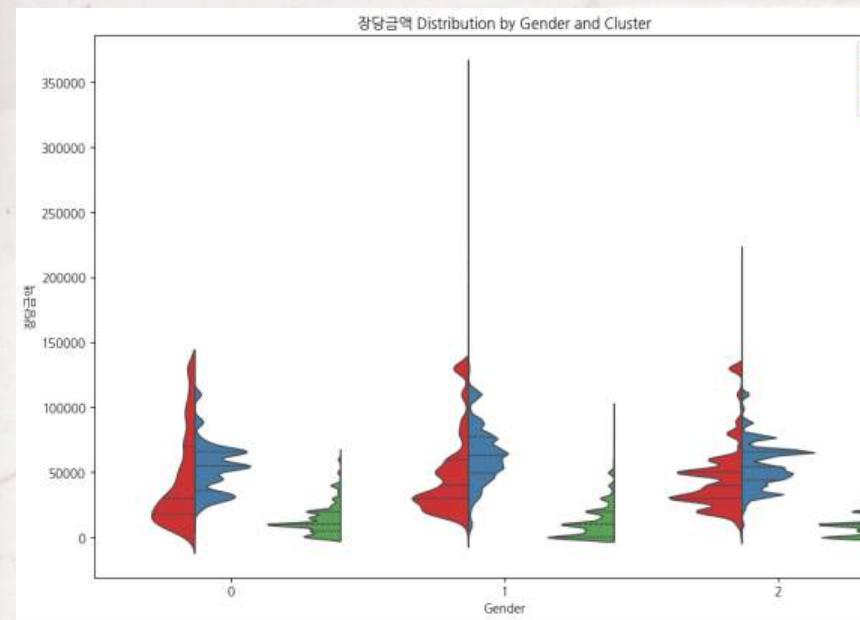
-> 오케스트라 포함 공연을 선호하며 뮤지컬보다 평균연령이 높음

분석과정

군집분석(k-means)

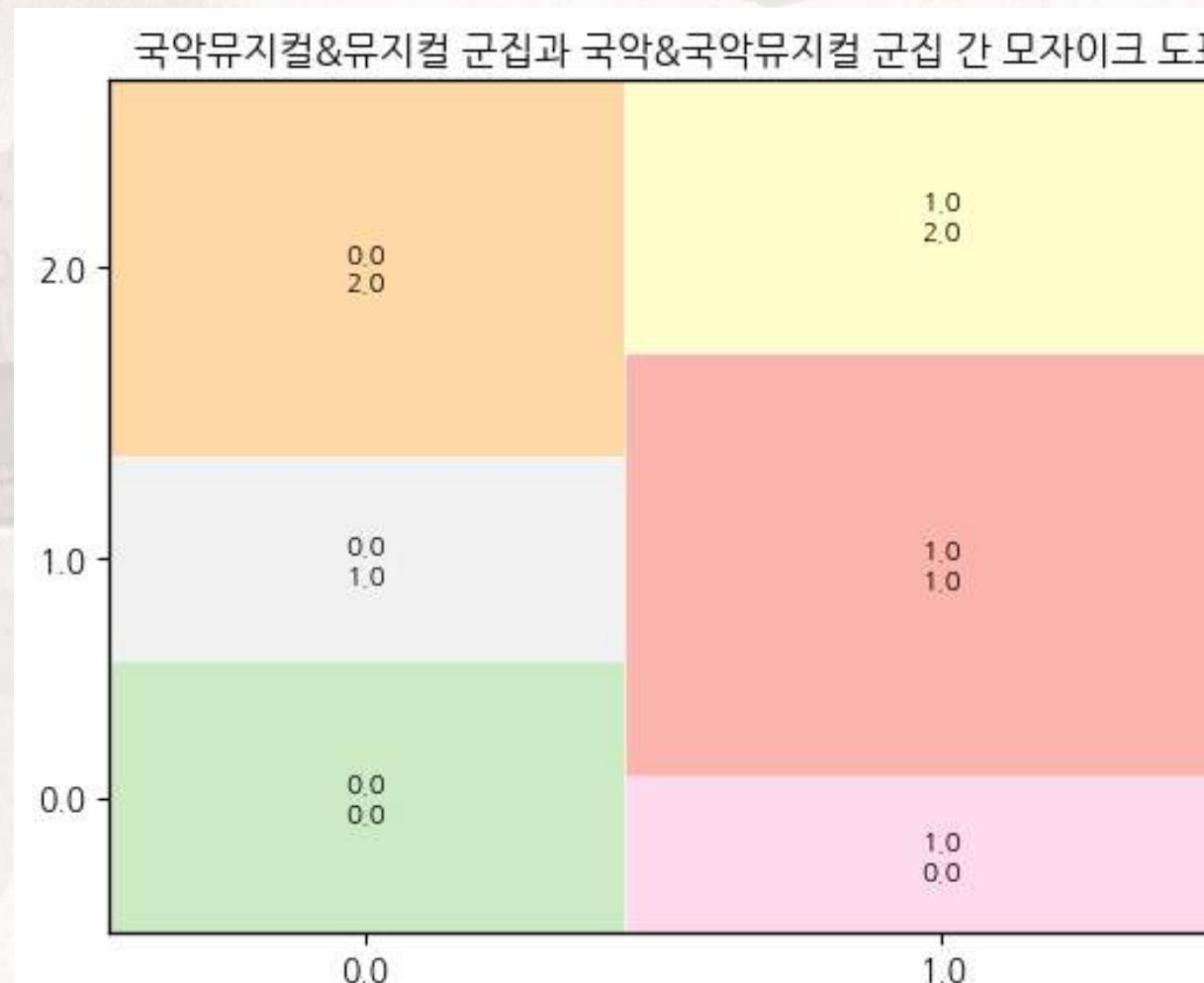
뮤지컬+국악뮤지컬, 국악뮤지컬+국악 군집 분석 후 교집합 도출

- K-평균(K-Means) 알고리즘을 사용하여 장르별로 유사한 관람 패턴을 가진 그룹을 식별
- 각 장르 간의 관람 패턴 차이를 이해, 이를 고려한 군집화를 통해 적절한 관람 유도 전략을 도출
- 국악뮤지컬이 뮤지컬과 국악을 연결하는 다리 역할을 할 수 있도록 데이터간 교집합을 찾고자 노력



분석과정

군집분석(k-means)



뮤지컬 군집만 존재하는 클러스터를 제외하고
국악뮤지컬의 고유 코드를 임의로 생성하여 연결한 후
모자이크 도표를 분석

-> 국악뮤지컬&뮤지컬의 군집 0은
국악뮤지컬&국악의 군집 0, 군집 2와 높은 연관성

국악뮤지컬&뮤지컬의 군집 1은
국악뮤지컬&국악의 군집 1과 강한 연관성

두 단계를 거쳐 뮤지컬, 국악뮤지컬, 국악을
유기적으로 연결할 수 있는 가능성 확인

문석내용

추천시스템

뮤지컬, 국악뮤지컬, 국악 공연의 간의 유사성을 기반으로 한 추천시스템

공연의 다양한 속성을 특징 벡터로 표현 -> 코사인 유사도를 계산하여 유사한 공연 추천
코사인 유사도는 벡터의 방향성에 중점을 두기에 더 높은 유사도

12개의 칼럼 (성별, 나이 구분, 가격대, 규모, 시설특성 코드, 공연지역명 코드, 분기, 공연일수, 할인금액, 장당금액, 소요시간, 판매기간) -> 공연의 특성과 관객의 성향을 모두 고려 가능

뮤지컬 코드



국악뮤지컬 코드



국악 코드

문석내용

뮤지컬 - 국악뮤지컬

뮤지컬 마니아를 국악 뮤지컬로!

코사인 유사도 계산후
상위 결과 3~5개 도출



PF435838

뮤지컬 공연코드 PF435838와 유사한 공연 추천:

- 국악 뮤지컬 공연코드 PF439804 (유사도: 0.9993)
- 국악 뮤지컬 공연코드 PF445217 (유사도: 0.9992)
- 국악 뮤지컬 공연코드 PF446339 (유사도: 0.9988)
- 국악 뮤지컬 공연코드 PF430092 (유사도: 0.9984)
- 국악 뮤지컬 공연코드 PF430826 (유사도: 0.9983)



PF445217

뮤지컬 공연코드 PF435838와 국악 뮤지컬 공연코드 PF445217의 차이점:

- 할인금액: 뮤지컬(0) vs 국악뮤지컬(0) => 차이: 0.00
- 장당금액: 뮤지컬(80000) vs 국악뮤지컬(66000) => 차이: 14000.00
- 성별: 뮤지컬(2) vs 국악뮤지컬(2) => 차이: 0.00
- 소요시간: 뮤지컬(130) vs 국악뮤지컬(90) => 차이: 40.00
- 시설특성_코드: 뮤지컬(0) vs 국악뮤지컬(2) => 차이: -2.00
- 공연지역명_코드: 뮤지컬(0) vs 국악뮤지컬(0) => 차이: 0.00
- 분기: 뮤지컬(1) vs 국악뮤지컬(3) => 차이: -2.00
- 공연일수: 뮤지컬(82) vs 국악뮤지컬(82) => 차이: 0.00
- 판매기간: 뮤지컬(35.0) vs 국악뮤지컬(29) => 차이: 6.00
- 나이_구분 비교: 다른 (뮤지컬: 20대_후반, 국악뮤지컬: 40대_초반)
- 가격대 비교: 같음 (뮤지컬: 5-10, 국악뮤지컬: 5-10)
- 규모 비교: 다른 (뮤지컬: 대극장, 국악뮤지컬: 중극장)
- 국악 뮤지컬 공연코드 PF430826 (유사도: 1.0000)

문석내용

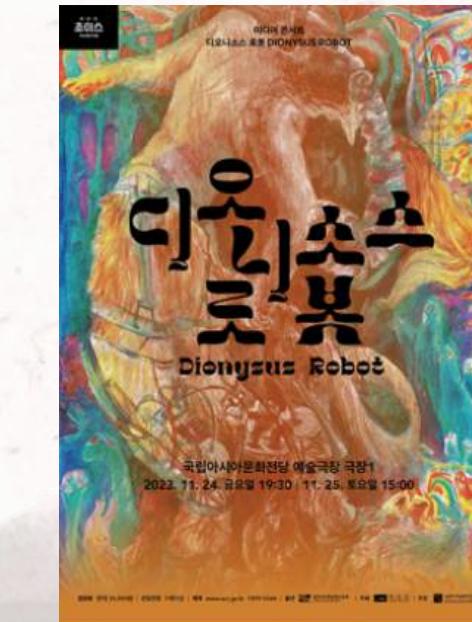
국악뮤지컬 - 국악

뮤지컬 마니아를 국악으로!

코사인 유사도 계산후
상위 결과 3~5개 도출



PF445217



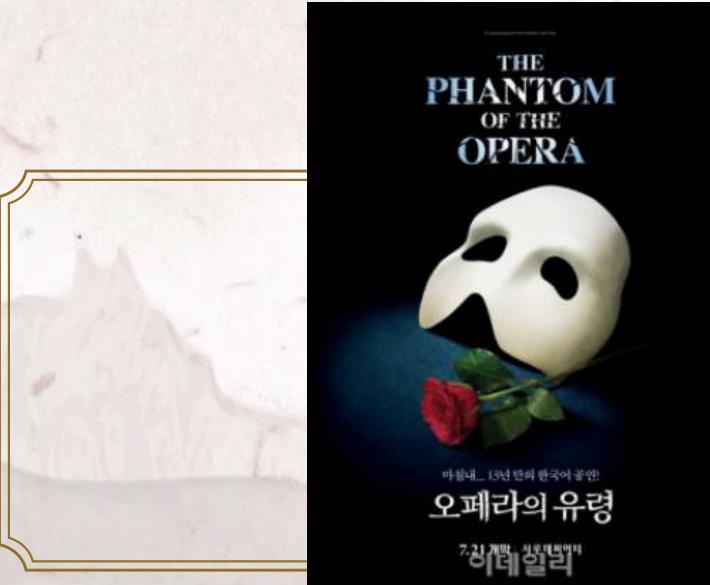
PF442432

국악뮤지컬 공연코드 PF445217와 국악 공연코드 PF442432의 차이점:

- 할인금액: 국악뮤지컬(33000) vs 국악(20000) => 차이: 13000.00
- 장당금액: 국악뮤지컬(66000) vs 국악(40000) => 차이: 26000.00
- 성별: 국악뮤지컬(2) vs 국악(2) => 차이: 0.00
- 소요시간: 국악뮤지컬(90) vs 국악(60) => 차이: 30.00
- 시설특성_코드: 국악뮤지컬(2) vs 국악(0) => 차이: 2.00
- 공연지역명_코드: 국악뮤지컬(0) vs 국악(0) => 차이: 0.00
- 분기: 국악뮤지컬(3) vs 국악(4) => 차이: -1.00
- 공연일수: 국악뮤지컬(82) vs 국악(2) => 차이: 80.00
- 판매기간: 국악뮤지컬(29) vs 국악(nan) => 차이: nan
- 나이_구분 비교: 다른 (국악뮤지컬: 30대_초반, 국악: 50대_초반)
- 가격대 비교: 같음 (국악뮤지컬: -5, 국악: -5)
- 규모 비교: 같음 (국악뮤지컬: 중극장, 국악: 중극장)
- 국악 공연코드 PF444722 (유사도: 1.0000)

문석내용

뮤지컬 - 국악뮤지컬



PF421016

뮤지컬 마니아를 국악 뮤지컬로!

코사인 유사도 계산후
상위 결과 3~5개 도출



PF439804

뮤지컬 공연코드 PF421016와 국악 뮤지컬 공연코드 PF439804의 차이점:

- 할인금액: 뮤지컬(9500) vs 국악뮤지컬(0) => 차이: 9500.00
- 장당금액: 뮤지컬(190000) vs 국악뮤지컬(110000) => 차이: 80000.00
- 성별: 뮤지컬(2) vs 국악뮤지컬(1) => 차이: 1.00
- 소요시간: 뮤지컬(150) vs 국악뮤지컬(150) => 차이: 0.00
- 시설특성_코드: 뮤지컬(0) vs 국악뮤지컬(2) => 차이: -2.00
- 공연지역명_코드: 뮤지컬(2) vs 국악뮤지컬(0) => 차이: 2.00
- 분기: 뮤지컬(1) vs 국악뮤지컬(2) => 차이: -1.00
- 공연일수: 뮤지컬(85) vs 국악뮤지컬(72) => 차이: 13.00
- 판매기간: 뮤지컬(60.0) vs 국악뮤지컬(35) => 차이: 25.00
- 나이_구분 비교: 다른 (뮤지컬: 40.0대_후반, 국악뮤지컬: 20대_중반)
- 가격대 비교: 다른 (뮤지컬: 15~, 국악뮤지컬: 10-15)
- 규모 비교: 다른 (뮤지컬: 대극장, 국악뮤지컬: 중극장)
- 국악 뮤지컬 공연코드 PF430826 (유사도: 0.9988)

문석내용

국악뮤지컬 - 국악

뮤지컬 마니아를 국악으로!

코사인 유사도 계산후
상위 결과 3~5개 도출



PF439804 Q



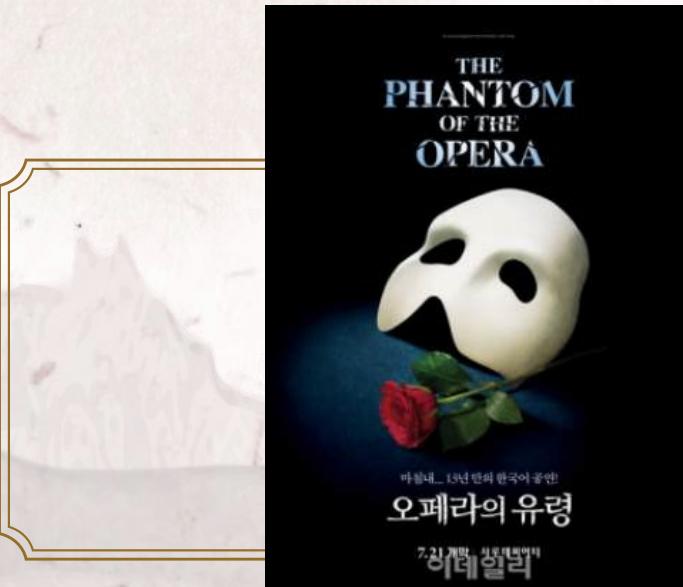
PF433691

국악뮤지컬 공연코드 PF439804와 국악 공연코드 PF433691의 차이점:

- 할인금액: 국악뮤지컬(0) vs 국악(0) => 차이: 0.00
- 장당금액: 국악뮤지컬(45000) vs 국악(20000) => 차이: 25000.00
- 성별: 국악뮤지컬(2) vs 국악(1) => 차이: 1.00
- 소요시간: 국악뮤지컬(150) vs 국악(80) => 차이: 70.00
- 시설특성_코드: 국악뮤지컬(2) vs 국악(4) => 차이: -2.00
- 공연지역명_코드: 국악뮤지컬(0) vs 국악(0) => 차이: 0.00
- 분기: 국악뮤지컬(2) vs 국악(1) => 차이: 1.00
- 공연일수: 국악뮤지컬(72) vs 국악(21) => 차이: 51.00
- 판매기간: 국악뮤지컬(27) vs 국악(16.0) => 차이: 11.00
- 나이_구분 비교: 다른 (국악뮤지컬: 40대_중반, 국악: 30대_중반)
- 가격대 비교: 같음 (국악뮤지컬: -5, 국악: -5)
- 규모 비교: 같음 (국악뮤지컬: 중극장, 국악: 중극장)
- 국악 공연코드 PF447908 (유사도: 1.0000)

문석내용

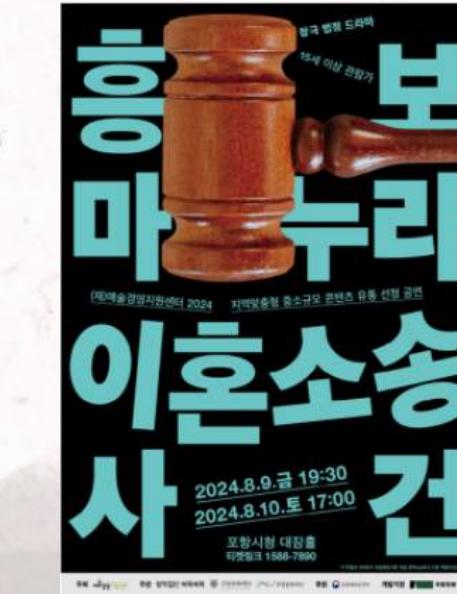
뮤지컬 - 국악뮤지컬 - 국악



PF421016

뮤지컬 마니아를 국악 뮤지컬로!

코사인 유사도 계산후
상위 결과 3~5개 도출



PF437398

- 국악 공연코드 PF437398 (유사도: 0.9976)

뮤지컬 공연코드 PF421016와 국악 공연코드 PF437398의 차이점:

- 할인금액: 뮤지컬(0) vs 국악(0) => 차이: 0.00
- 장당금액: 뮤지컬(160000) vs 국악(0) => 차이: 160000.00
- 성별: 뮤지컬(1) vs 국악(1) => 차이: 0.00
- 소요시간: 뮤지컬(150) vs 국악(80) => 차이: 70.00
- 시설특성_코드: 뮤지컬(0) vs 국악(4) => 차이: -4.00
- 공연지역명_코드: 뮤지컬(2) vs 국악(0) => 차이: 2.00
- 분기: 뮤지컬(1) vs 국악(2) => 차이: -1.00
- 공연일수: 뮤지컬(85) vs 국악(8) => 차이: 77.00
- 판매기간: 뮤지컬(45.0) vs 국악(19.0) => 차이: 26.00
- 나이_구분 비교: 다른 (뮤지컬: 20대_초반, 국악: 50대_초반)
- 가격대 비교: 다른 (뮤지컬: 15-, 국악: -5)
- 규모 비교: 다른 (뮤지컬: 대극장, 국악: 소극장)
- 국악 공연코드 PF421620 (유사도: 0.9963)

문석내용

추천시스템

1. 뮤지컬 - 국악뮤지컬 / 국악뮤지컬 - 국악

단일 단계 구조.

-> 뮤지컬 관객을 국악뮤지컬로, 국악뮤지컬 관객을 국악으로 유도.

2. 뮤지컬 - 국악뮤지컬 - 국악

복합적인 두 단계 구조.

장르간의 자연스러운 연결 촉진

-> 세 장르간의 유기적인 연결을 통해 뮤지컬 관객을 국악으로 유도

03

제 3 장

활용 가능성

- 추천 시스템 활용 방향성.
- 기대 효과.

활용가능성

활용가능성

가격 할인

뮤지컬 보면 국악뮤지컬 할인
국악뮤지컬 보면 국악 할인

회전문 관객 문화 활용

같은 클러스터 내에 속하는 공연들 합쳐서 적립
추천받은 국악뮤지컬, 국악 보면 적립

공연 중심의 추천 시스템

공연 중심의 추천시스템 (사용자 정보 제한 극복)

기대 효과

기대 효과

국악 활성화

국악뮤지컬이 뮤지컬 관객의 니즈를 충족 가능하고 이는 국악으로 이어질 것
국악 성장을 위해 새로운 사람들 유입시키는 것보다 원래 있던 사람들을 유입시키는게 효과적

건강한 공연 생태계

이 추천시스템 쓰면 국악 국악뮤지컬 뮤지컬 다 잘 될 수 있음.

KOPIS 데이터 수집 방향성 제시

공연 중심의 추천 시스템을 발전 -> KOPIS 공연 데이터 수집의 방향성 제시

무지털
국면들다.

감사합니다

혜화역 4번출구