
2024년 제 12회 빅콘테스트 데이터분석 분야

축제 성과 지표 개발

: 지속 가능성, 화제성, 수익성, 외부 유입성 분석을 통한 접근

나도 축제꺼줘
김찬유, 이유송, 장은별, 정수빈

Contents

축제 성과 지표 개발

01

서론

연구배경

개요

연구 목적

아이디어

02

데이터

제공 데이터

추가 데이터

03

평가 지표

지속 가능성

화제성

수익성

외부 유입성

04

결론

결과 예시

기대 효과

사용 방안

나도 축제꺼줘

2024년 제 12회 빅콘테스트 데이터분석 분야

01. 서론

연구 배경
개요
연구 목적
아이디어

01 연구 배경

사회

"지역경제 활성화"... 축제 늘렸지만, 방문객 '저조'

장정복 장수군의원 "지역축제·행사, '선택과 집중' 발휘해야"

현재 지역축제의 경우 축제의 효용성 없이 축제의 개수만 많아지는 경향을 보이고 있음
지역축제의 선택과 집중에 대한 필요성이 대두 되는 중이다.

01 연구 배경

지역 축제의 증가

지역마다 개최되는 축제의 수가 증가하는 경향을 보임

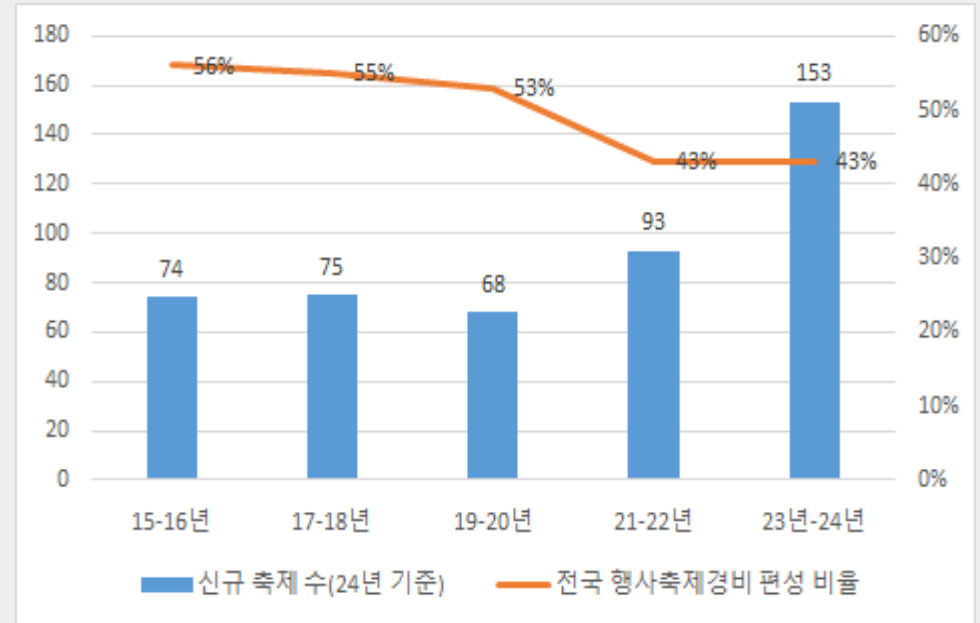
전국 지방자치단체가 개최하는 지역 축제가 5년 사이 32% 가량 증가
코로나 이후 지역 경제 활성화를 위해 지역축제를 늘렸으나, 성과는 크지 않음

지역 축제 경비 감소

지역의 비전보다 단체장 개인의 성과에 집착해 성과가 있던 지역 축제를 발전할 생각보다 새로운 사업을 확장하려는 단기적인 정책 행태를 보임

→ 기존의 것을 없애기에는 부담이 크고 새로운 일을 진행하려고 함

축제 수가 증가함에 따라 축제에 편성되는 경비는 감소



* 출처 : 나라살림연구소, 2024년 지역축제 현황 및 성과분석에 따른 성과관리 제도개선 방향성 제언, 2024.08.21

01 연구 배경

지역 축제 참여율의 감소

전반적인 축제 경비는 0.08% 감소

주민 참여율 평균 9.6% 감소

외부 방문객 1.58% 감소

관광 소비액 12.72% 감소

구분	지역주민의 지역축제 참가율		
	2019년(a)	2023년(b)	차이 (c=b-a, %p)
강원	38.7	35.7	-3.0
충북	56.6	47.0	-9.6
충남	55.7	63.2	7.5
전북	61.3	27.5	-33.8
전남	50.0	53.4	3.4
경북	44.3	31.8	-12.5
경남	56.7	35.7	-21.0
제주	51.8	18.9	-32.9
평균	45.0	35.4	-9.6

*출처 : 문화체육관광부 「19-23년 국민여가활동조사」, 재구성

1인당 관광소비액이 감소하고 있다는 것은 지역주민은 외면하고 외부 방문객에겐 소비를 유인할 만큼의 매력
이나 효과가 낮음을 시사함

무분별한 지역 축제의 증가보다 질적 개선이 필요함을 알 수 있다.

01 개요

축제의 성과 및 문제점 파악
축제의 성과를 화제성, 수익성,
지속 가능성, 외부 유입성 지표로
분석하여 성과 및 개선점을 파악함



데이터 기반 의사결정 지원
지표 기반 평가를 통해 축제에 투입되는
예산의 효율성을 높일 수 있음
축제별로 분석된 데이터를 활용하여
투자 대비 효과가 높은 축제를 선별해
전략적인 자원 배분이 가능해짐

지역 경제 및 이미지 제고
외부 유입성과 지역 특성을 활용해 지역
경제 활성화 및 이미지 향상에 도움을 줌.

01 연구 목적



새로운 성과 지표



이에 본 연구에서는 빅콘테스트 대회에서 제공하는 od 데이터, stay 데이터와 한국관광공사의 ‘한국관광데이터랩’에서 제공되는 데이터를 분석해 효과적인 축제 성과 지표를 제공하고자 한다.

01 연구 목적

본 연구의 목적을 위해 지속가능성, 화제성, 수익성, 외부 유입성의 네 가지 주요 지표를 설정하여 각 지역 축제를 종합적으로 분석하고자 한다. 이러한 분석을 통해 다음과 같은 효과를 부가적으로 기대할 수 있다.



지역 주민과 방문객의 참여를 증대시키기 위한 효과적인 방안을 모색

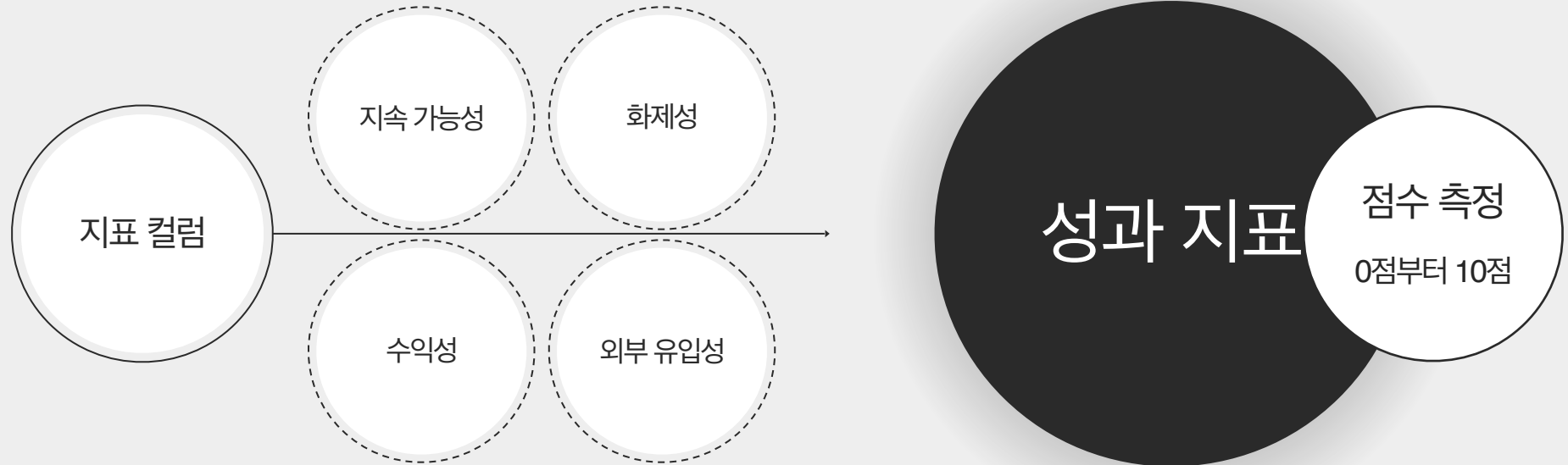


각 축제의 성과를 명확하게 측정할 수 있는 기준을 마련



지속 가능한 축제 운영을 위한 전략을 제안

01 아이디어 성과 지표의 구조



지표의 구성은 지속가능성, 화제성, 수익성, 외부 유입성 4가지 측면으로 구성되어 있으며
각 지표 별 0점부터 10점까지 매기어 평균으로 총 성과 지표 점수 측정함

2024년 제 12회 빅콘테스트 데이터분석 분야

02. 데이터

제공 데이터
추가 데이터

02

제공 데이터

OD 데이터 & STAY 데이터

- 데이터 구성

- 기간 별 OD 데이터 및 Stay 데이터를 한번에 관리하기 위하여 병합 시도
- 초기 데이터 병합 시 각각 약 1억 5천개 이상의 데이터 존재

- DB 활용

- 병합한 OD 데이터 및 stay 데이터를 활용하기 위해 DB에 저장 후 관리 및 분석 수행.

- 축제 데이터 추출

- 축제 일부를 기준으로 하기 위해 한국 관광 데이터 랩의 축제 지표 중 기간이 겹치는 38개의 축제를 선정
- 그 이후 38개의 축제가 이루어진 행정동 데이터를 모두 추출하여 활용

OD 데이터

od_data
origin_hdong_cd BIGINT
dest_hdong_cd BIGINT
date VARCHAR(45)
start_time VARCHAR(45)
end_time VARCHAR(45)
gender INT
age INT
modal INT
origin_purpose INT
dest_purpose INT
od_dist_avg INT
od_duration_avg INT
od_cnts INT

STAY 데이터

merged_st...
hdong_cd BIGINT
date VARCHAR(45)
time VARCHAR(45)
gender INT
age INT
purpose INT
stay_cnts INT

02 추가 데이터.

크롤링 데이터 & 한국관광 데이터랩 빅데이터

크롤링 데이터

네이버 블로그 크롤링

네이버 뉴스 크롤링

유동인구 데이터 기간(0901~1015)이 포함된 2023년 7월부터 12월 하반기의 뉴스와 블로그 언급 크롤링

제목 "횡성 한우축제탈락"...

날짜 2023-12-20

url <https://www.news1.kr>

내용 당초 지난 선정에 이어 ...

한국 관광 데이터 랩

지역별 관광 지출액

해당 기간 내(0901~1015) 축제에 대한 시군구별 관광 지출액 데이터

•사용 데이터 출처

- 한국관광데이터랩: 관광 및 매출 데이터로 시군구별 관광매출데이터 추출 [Visit Korea Data Lab¹](#)
- 주민등록 인구통계: 시도별 인구수 데이터 사용 [MOIS¹](#)
- 문화체육관광부: 연도별 지역 축제 정보 [MCST¹](#)

2024년 제 12회 빅콘테스트 데이터분석 분야

03. 평가지표

지속 가능성
화제성
수익성
외부 유입성

03

평가지표

평가지표 구성

지속 가능성

해당 축제에 대한 긍정적인 반응,
부정적인 반응을 분석하여 향후
축제가 지속될 수 있는지의 가능성을 지
표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

화제성

해당 축제가 얼마나 많이 언급이
되었는지 언급 수를 지표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

수익성

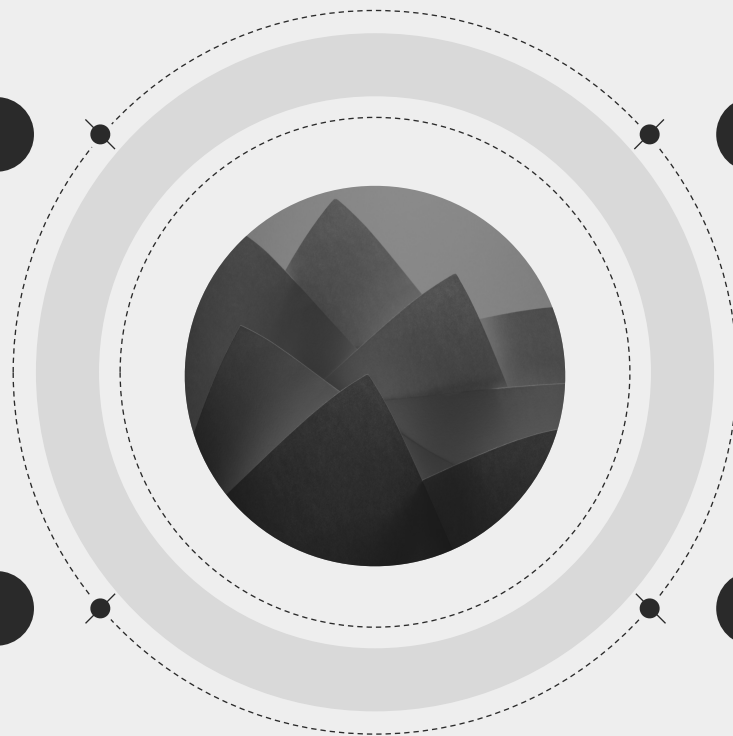
해당 축제의 수익성을 파악하여
평소 해당 지역의 수익성 대비
얼마나 증가했는지 지표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

외부 유입성

해당 축제 기간 동안 외부에서
얼마나 많은 외지인들이 방문하는지에 대
해 유입성을 계산하여 지표로
표현한다.

최소 0점에서 최대 10점 사이의 점수



03 평가지표

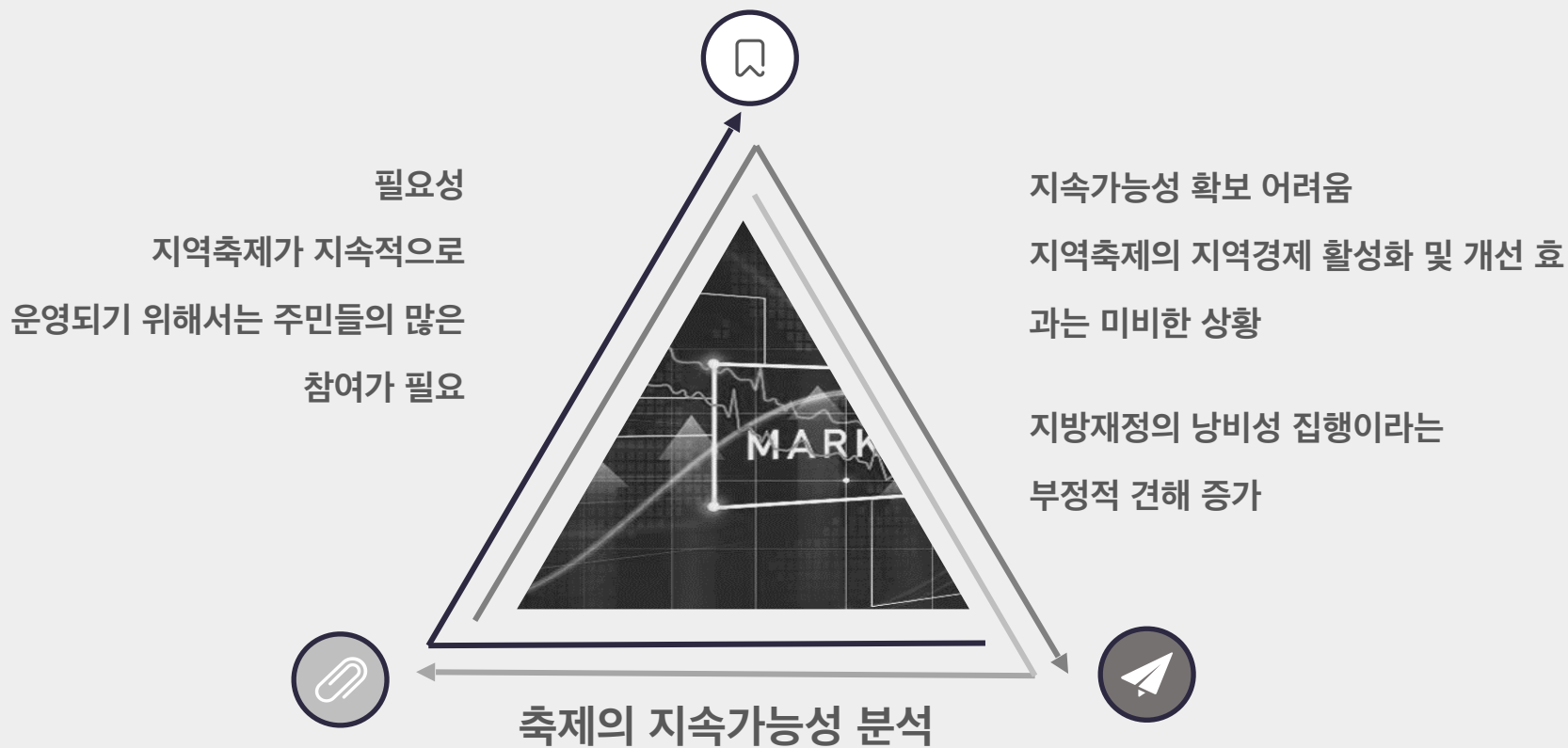
평가지표 구성



03

지속가능성

지속가능성 분석 개요



한 축제에 대한 후기를 발췌하여 긍정 분석 모델에 입력 후 예측한 결과, 99.55%의 확률로 부정 리뷰임을 판단

03 지속가능성

지표 산출 방법



1 블로그 수

- 10개 미만: 기본 점수 2점 부여
- 10개 이상: 기본 점수 4점 부여



2 점수 산출

$$5 + \frac{((\text{긍정비율} \times \text{블로그 수}))}{((\text{블로그 수} + 10) \times 10) \times (1 - \text{부정 비율})}$$

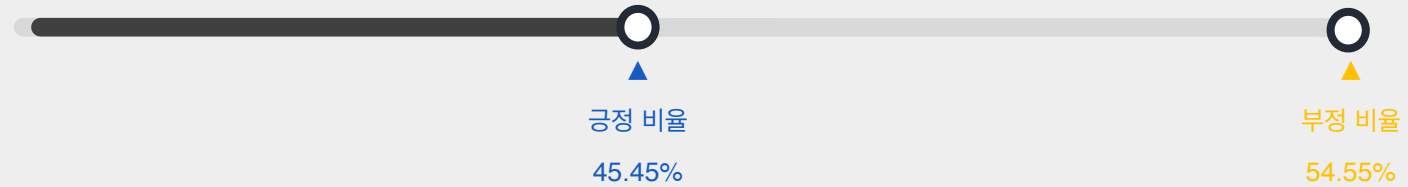
03

지속가능성

화제성 평가 예시 & 결론

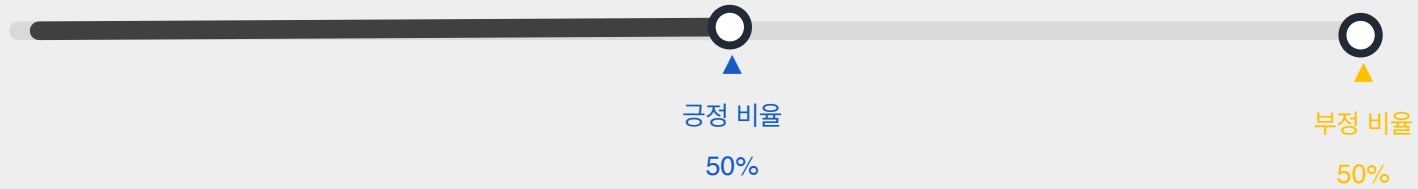
블로그 수
33개

5.6 임실치즈축제



블로그 수
20개

5.7 황성한우축제



→ 임실 치즈 축제의 경우 점수가 5.6, 황성한우축제의 경우 5.7점으로 두 축제의 지속가능성은 비슷한 수준이라고 할 수 있다.

03

화제성

화제성 분석 개요



03

화제성

지표 산출 방법



03

화제성

화제성 평가 예시 & 결론

9.0 임실치즈축제



“ 화제성 점수를 통해 축제의 사회적 관심도를 평가하고, 향후 마케팅 전략에 사용할 수 있다. ”

5.1 횡성한우축제



→ 임실 치즈 축제의 경우 점수가 9.0점으로 화제성이 매우 높고, 횡성한우축제의 경우 점수가 5.1점으로 화제성은 평균 수준이라고 할 수 있다.

03

수익성

수익성 분석 개요

중요성



지역 경제 활성화

관광객 유치로 숙박업, 요식업 등 지역 경제에서 매출 증대 효과 창출



지역 정체성 강화

지역 문화 보존 및 지역민과 관광객 간 교류 촉진



문화 및 관광 가치 증진

수익성 있는 축제는 홍보 및 콘텐츠 개선에 재투자 가능



재정적 안전성 강화

수익을 통한 향후 축제의 개최 시 예산 부담 감소

03

수익성

데이터 전처리



Step 01
결측치 제거



Step 02
STAY 데이터 전처리

- 날짜 및 시군구별 인구수를 행 단위로 저장
- 열단위 : 나이대, 목적, 성별 기준의 인구수 집계



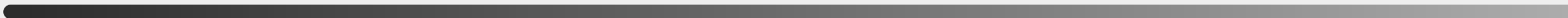
Step 03
추가 변수 생성

휴무 여부, 추석 여부
열 추가



Step 04
매출 데이터 결합

각 날짜 및 시군구 기준 매출 데이터와 통합



03

수익성

지표 산출 방법

평가 기준

📍 문화체육관광부의 지역 축제 정보 데이터를 바탕으로, 축제 예산 대비 매출 기준을 120%로 설정

📍 예산 대비 실제 매출이

- 120% 초과 : 수익성 높은 축제
- 190% 이상 : 최고점 부여

수익이 나지 않을 경우 : 0점

계산 공식

$$5 + \frac{\left(\left(\frac{\text{실제매출}}{\text{비축제예산매출}} \right) \times 100 \right) - 120}{14}$$

$$\frac{\left(\left(\frac{\text{실제매출}}{\text{비축제예산매출}} \right) \times 100 \right) - 100}{4}$$

03

수익성

데이터 모델링

Random forest

주요 변수로 시도명, 휴무 여부, 나이대, 목적 변수를 사용
랜덤포레스트 모델: 모든 변수를 포함하여 예측 성능 비교

다중공선성에 덜 민감하며 성능 향상 가능성 높음

Final Test RMSE: 89630.72455237304

Final Test R^2 : 0.9642129932904948

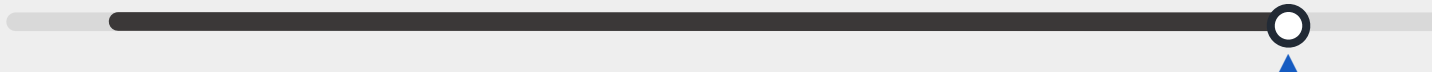
Test R^2 : 0.9619643830193861

결과 랜덤포레스트를 통해 축제 기간의 수익성 예측
최고 성과 축제 예시: 임실 축제가 비축제 대비
187.52%의 매출을 달성하여 높은 점수 기록

03 수익성

수익성 평가 예시 & 결론

9.8



임실치즈축제

- 주중예상 평균 관광액 27776.73(천원)
- 주말예상 평균 관광액 60917.37(천원)
- 실제값의 예상 평균액 대비 비율의 평균
→ 187.5874%

5.3



황성한우축제

- 주중예상 평균 관광액 100449(천원)
- 주말예상 평균 관광액 204413.4(천원)
- 실제값의 예상 평균액대비 비율의 평균
→124.8203%

→ 임실 치즈 축제의 경우 점수가 9.8점으로 수익성이 매우 높고, 황성한우축제의 경우 점수가 5.3점으로 수익성은 평균 수준이라고 할 수 있다.

03

외부 유입성

외부 유입성 분석 개요

외부 방문객의 기여

지역 전통 문화를 직접 체험하며 지역 정체성 홍보
지역민과 방문객 간의 상호작용으로 긍정적 이미지 형성



지표 개발의 목적

외부 방문객 유입 효과를 객관적으로 측정
축제 성과를 수치화하여 경제적 문화적 기여
도 비교

축제의 역할

지역 경제와 문화적 가치를 평가하는 기회.
외부 관광객 유치로 숙박, 요식업, 소매업 등
지역 경제 활성화

03

외부 유입성

데이터 전처리



1

유입성 지표 정의

- 특정 축제에 외부 방문자가 얼마나 많이 유입되는지 평가하기 위한 지표
- 점수는 0점에서 10점 사이로 소수 둘째 자리에서 반올림을 한다.



2

가중치 설정

- 주요 컬럼: 출발지/도착지 행정동 코드, 날짜, 체류 목적, 평균 이동거리
- 이동거리 기준 가중치(weight) 설정

87.6km 이상	2점
13.59km 이상 · 87.6km 미만	1점
13.59km 이하	0.5점

- 가중치는 해당 논문을 참고하여 산정하였다.

국민국내관광객의 이동시간거리에 대한 수요 탄력성 분석 P.7
‘국내 관광 수요는 출발지에서 멀어질수록 급감하고, 거리가 늘어날수록 지출 비용과 숙박 비율이 증가한다.’

→ 지출 비용과 숙박 비용이 증가한다는 점은, 관광객이 오랜시간 해당 축제에 참여한다는 의미로 해석되어 더 높은 가중치를 부여한다.



3

점수 산출 기준

- 현지인 데이터 제외 및 주말·평일 구분
 - 출발지와 도착지 시·군·구가 동일한 경우 현지인으로 판단하여 제외
 - 주말·공휴일과 평일을 구분하는 ‘dav_of_week’ 컬럼 추가
 - (0=평일, 1=주말/공휴일)

03 외부 유입성

지표 산출 방법

STEP 1

필터링

데이터 필터링

- 출발지와 도착지의 행정동 코드 앞 5자리를 비교하여 동일 시·군·구 내 데이터 제외
- 지역 간 이동이 이루어진 경우에만 데이터 반영

STEP 2

가중치

거리 간 가중치 반영

- $Weight \propto d$
두 컬럼을 곱해 거리의 영향을 반영한 방문자 수 산출
- 이동 거리가 더 긴 방문자에게 높은 가중치 부여

STEP 3

방문자 구분

축제 기간 방문자 구분

- 주말 및 공휴일, 평일 방문자 수를 각각 구분하여 가중 평균 계산
- 축제 기간 일별 방문자 수와 비축제 기간 일별 평균 방문자 수 차이로 외부 유입성 측정

STEP 4

Log - Scale

scale 적용

- 축제와 비축제 간 방문자 수 차이 완화 및 값의 분포 평탄화
- log-scale을 이용하여 방문자 수 차이 60%, 비율 40% 가중치 적용 후, min-max scale을 통해 점수 지표 생성

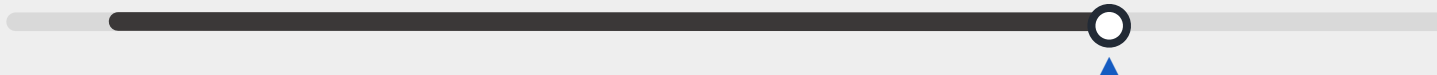
03

외부 유입성

외부 유입성 평가 예시 & 결론

“이 분석을 통해 각 축제의 외부 유입 효과를 객관적으로 평가하고, 향후 축제 기획에 활용할 수 있는 기초 자료를 제공합니다.”

7.4



6.6



임실치즈축제

- 평균 외부 방문객: 3271.5
- 축제 방문객 : 333.3
- 방문자 수 차이(log scaling) : 3.5
- 변환된 비율 값 : 0.99

횡성한우축제

- 평균 외부 방문객 : 3059.0
- 축제 방문객 : 920.8
- 방문자 수 차이(log scaling) : 3.3
- 변환된 비율 값 : 0.52

→ 임실 치즈 축제의 경우 점수가 7.4점으로 외부 유입성이 다른 축제들에 비해 높은 편이라고 할 수 있다. 반면 횡성한우축제의 경우 점수가 6.6점으로 외부유입성은 평균 수준이라고 할 수 있다.

2024년 제 12회 빅콘테스트 데이터분석 분야

04. 결론

결과 예시
기대 효과
사용 방안

결론. 평가지표

평가지표 구성

지속 가능성

해당 축제에 대한 긍정적인 반응,
부정적인 반응을 분석하여 향후
축제가 지속될 수 있는지의 가능성을 지
표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

화제성

해당 축제가 얼마나 많이 언급이
되었는지 언급 수를 지표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

최종 점수

각 지표 점수의 평균

최소 0점에서 최대 10점 사이의 점수

수익성

해당 축제의 수익성을 파악하여
평소 해당 지역의 수익성 대비
얼마나 증가했는지 지표로 표현한다.

최소 0점에서 최대 10점 사이의 점수

외부 유입성

해당 축제 기간 동안 외부에서
얼마나 많은 외지인들이 방문하는지에 대
해 유입성을 계산하여 지표로
표현한다.

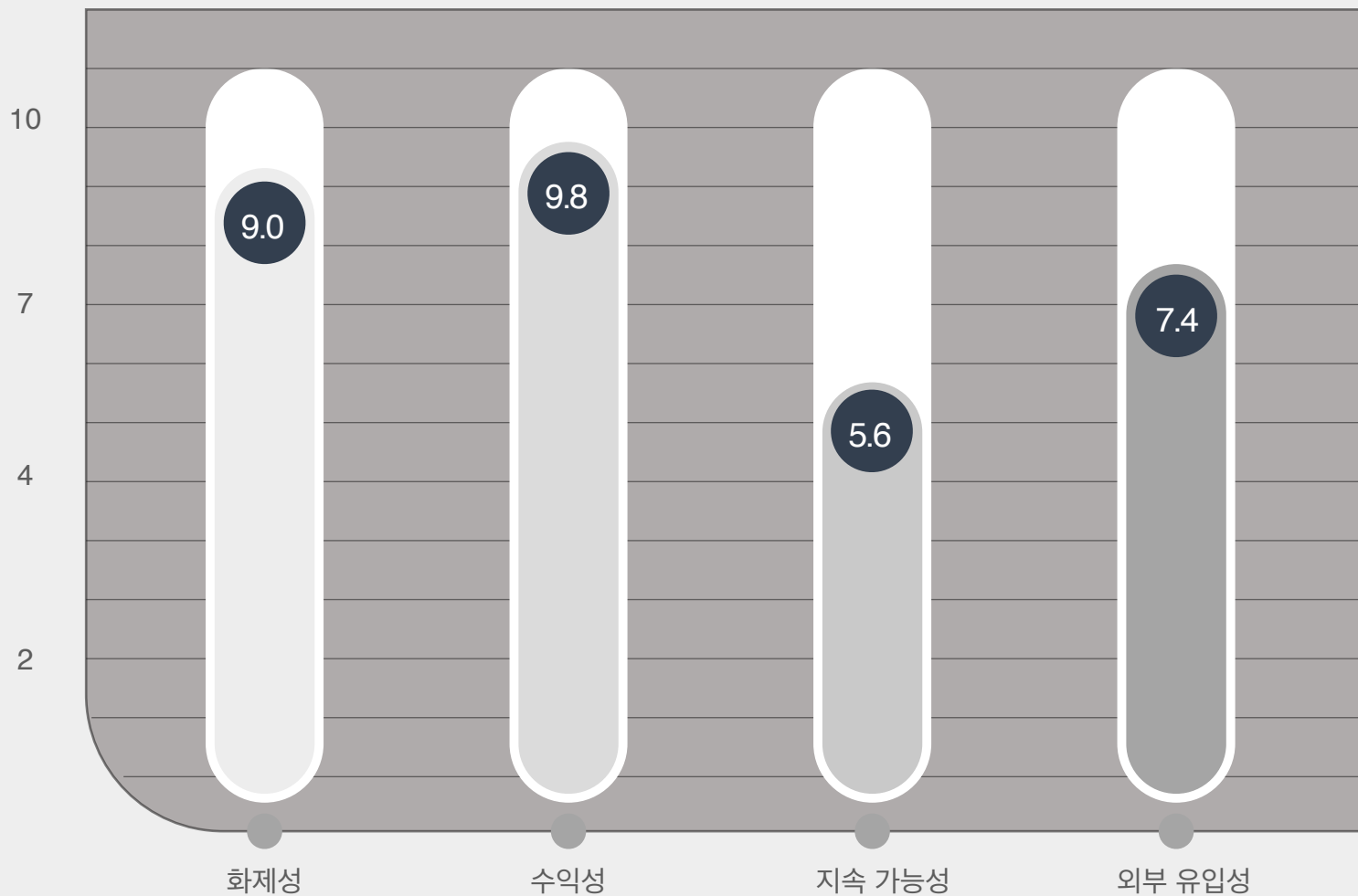
최소 0점에서 최대 10점 사이의 점수

결론. 결과 예시

임실치즈축제

7.9

임실치즈축제는 화제성과 수익성에서 매우 높은 점수를 기록하여, 경제적 성과와 사람들의 관심을 동시에 끌어내는 성공적인 축제로 평가됨.
다만, 지속가능성 측면에서 중간 수준에 머물러 있어 장기적인 발전을 위해 환경적, 사회적 측면에서의 개선이 필요할 수 있음.



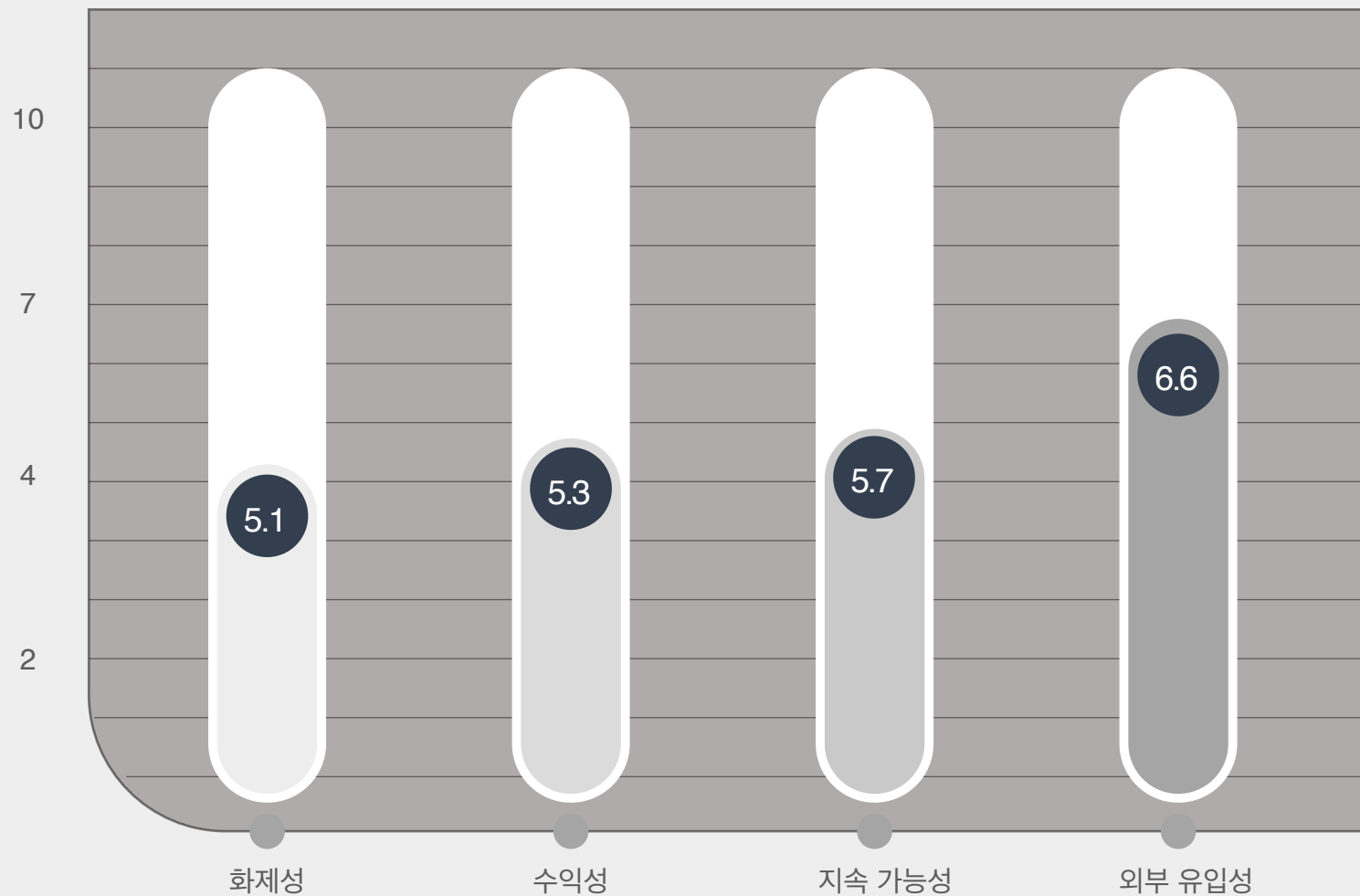
결론. 결과 예시

횡성한우축제

5.7

횡성한우축제는 전반적으로 10점 만점에서 중간 이상의 점수를 기록했으나, 모든 지표가 평균 수준을 크게 넘어서지 않는 점을 고려할 때, 지속적인 개선이 필요한 축제라 평가할 수 있다.

특히, 화제성과 수익성 강화를 위한 마케팅 및 콘텐츠 개선이 중요한 과제로 보임.



결론. 기대효과

? 효율성

국가예산의 효율성 증가

본 평가 지표를 통해 축제의 경제적 성과를 객관적으로 분석함으로써, 정부와 지방자치단체가 축제에 투입하는 예산을 보다 효율적으로 사용할 수 있음.

축제별로 수익성과 외부 유입성을 평가하여, 예산 투입 대비 높은 효과를 거둘 수 있는 축제에 집중함으로써 국가 예산의 효율성 증대

! 만족도

관광객의 만족도 향상

관광객의 관심을 끌고 만족도를 높이는 요소를 분석하여 축제 기획에 반영할 수 있음.

화제성과 외부 유입성 지표를 통해 관광객들이 기대하는 축제 요소를 파악하고, 이를 강화하여 방문객의 만족도를 극대화할 수 있음.

이로 인해 재방문율이 높아지고 지역 경제에도 긍정적인 영향을 미침

결론. 사용 방안

? 판단 지표

향후 축제 계획 시 판단 지표로써 사용

지표를 활용하여 화제성, 수익성, 지속 가능성, 외부 유입성 등의 다양한 측면에서 축제의 성과를 종합적으로 평가할 수 있음.

이를 통해 향후 축제 기획 시 각 지표를 기준으로 축제의 강점과 개선점을 파악하여 보다 전략적인 기획이 가능.

각 지역 특성에 맞는 축제를 설계하고, 평가된 지표를 바탕으로 지속 가능한 축제 운영 방안을 모색할 수 있음.

! 발굴 및 육성

유망 축제 발굴 및 육성

축제 평가 지표를 통해 우수한 성과를 보이는 유망 축제를 발굴하고, 이를 집중적으로 육성할 수 있는 기반을 마련함.

특히, 경제적 파급 효과가 크거나 높은 외부 유입성을 보이는 축제를 선별하여, 지역 경제 활성화와 문화 가치 제고에 기여하는 방향으로 지원 정책을 강화할 수 있음.

참고문헌.

데이터

- 한국관광데이터랩, 관광 및 매출 데이터 시군구별 관광매출데이터, [Visit Korea Data Lab](#)¹
- 주민등록 인구통계, 시도별 인구수 데이터, [MOIS](#)¹
- 문화체육관광부, 연도별 지역 축제 정보 [MCST](#)¹

뉴스 기사

- 경상일보, 무분별한 울산지역 축제, 질적 개선 시급하다, 2024.08.23, <https://www.ksilbo.co.kr/news/articleView.html?idxno=1007095>
- 장정복 장수군의원 "지역축제·행사, '선택과 집중' 발휘해야“, 김현호, 아주경제, 2024.10.25, <https://www.ajunews.com/view/20241025141218564>
- “지역경제 활성화”... 축제 늘렸지만, 방문객 ‘저조’, 박정훈, 서울신문, 2024.08.24, <https://www.seoul.co.kr/news/society/2024/08/23/20240823500196>

논문 및 보고서

- 국민국내관광객의 이동시간거리에 대한 수요탄력성 분석, 김의준, 서울대학교, 2015
- 온라인 고객 리뷰가 구매의도에 미치는 영향 -신뢰성을 매개변수로-, 박경옥, 부경대학교, 2008.07
- 나라살림연구소, 2024년 지역축제 현황 및 성과분석에 따른 성과관리 제도개선 방향성 제언, 2024.08.21, <https://www.narasallim.net/report/654>

END OF DOCUMENT

THANK YOU

나도 축제꺼줘
