



장기체류 외국인 관광객을 고려한 지역관광 활성화 전략

-포항시와 거제시를 중심으로

0. 목차

01
문제발굴

02
문제제기

03
해결방안분석

04
전략적 제언

05
시사점

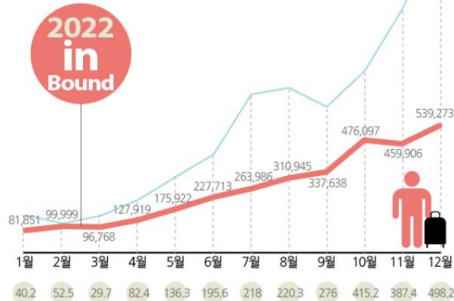
1. 문제 발굴

포스트 코로나와 K 콘텐츠의 세계화와 같은 호재로 인바운드 관광객이 크게 증가할 것으로 예상되며, 관광대국으로의 성장에 있어 이들은 매우 중요하므로 인바운드 관광객에 대한 분석 시작

포스트 코로나

2022년 출입국자수 단위=명, %

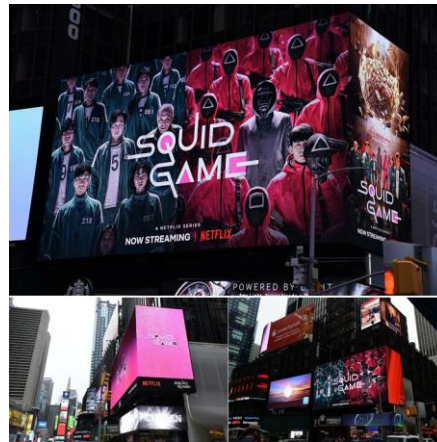
외국인 입국자수 합계= 3,198,017명
증감률 전세=230.7%



* Source: 2022년 출입국자 수, 12월 정점 찍었다... 회복세 뚜렷

- 포스트코로나 시대 이후 출입국자 수는 지속적으로 증가하며 뚜렷한 회복세를 보임
- 실제로 문화체육관광부의 발표에 의하면 2023년 1분기 외국인 관광객은 코로나 이전인 2019년에 비해서도 약 44% 증가

K 콘텐츠의 세계화



* Source: 글로벌 OTT 전쟁, K-콘텐츠 어디로 가야 할까

- 실제로 최근에는 K-POP, K-드라마 등과 같은 K-콘텐츠들의 전례 없는 인기로 많은 외국인 관광객들이 국내 관광에 큰 호기심을 갖고 있음

관광대국으로의 성장

尹 "고려청자 넘어 순대·떡볶이·어묵 먹으러 올때 K관광 성공"(종합)
여행에 K콘텐츠 결합... '야놀자' 5천만 관광객 시대' 연다



* Source: 여행에 K콘텐츠 결합... '야놀자' 5천만 관광객 시대' 연다

- 특히 정부와 국내의 관광기업 모두 인바운드 관광객 유치를 위해 힘쓰고 있는 상황이며, 다양한 기회 요인으로 국내 인바운드 관광은 큰 호황기를 맞을 것으로 예상됨

1. 문제 발굴

인바운드 관광객을 분석하기 위해서 2022년 외래관광객 조사를 활용하였고,
잠재적계층분석(LCA 분석)을 할 수 있는 범주형 변수의 형태로 전처리

2022 외래관광객 조사

질문사항

RQ1: 방한횟수 (4개범주)

Q3: 방한목적 (6개범주)

가장 만족한 활동(17개범주)

RDAY평균: 여행 지출 경비 (7개범주)

여행 중 방문 지역 1순위 2순위 (17개범주)

D_NAT_CLUST :국가 (21개범주) ->
대륙(7개범주로 축소)

:

2022 외래 관광객 조사 전처리

1
60일 이상의
체류 외국인
제거

- 유학생, 장기 출장 등의 요인을 제외하고 관광객들만을 포함시키기 위해 60일 이상 체류 외국인 제거

2
범주형 변수
축소

- 국가(21개범주) 와 같이 축소 가능한 변수 축소

3
중복 가능
문항 전처리

- 중복 가능 문항 중 범주형 변수의 처리가 가능했던 방문지역을 체류 기간에 따라 1순위 2순위로 전처리

잠재적 계층분석(LCA)을 위한
여러 전처리 과정이 필요

For example

```
tour['체류2순위'] = selected_columns.apply(lambda x: x.nlargest(2).index[-1]  
if not np.isnan(x.nlargest(2).iloc[-1]) else '없음', axis=1)
```

id	RQ1	Q3	Q4	RQ2A2	C통역1인	MS	RDAY평균	Q7	D_BUN	D_SEX	D_AGE	체류1순위	체류2순위	D_NAT_CLUST	만족한활동
0	289	4	2	1	12	3	2	5	1	1	2	서울일TOT	없음	유럽	4
1	306	1	3	1	2	3	4	4	1	1	5	부산일TOT	없음	기타	14
2	324	1	2	1	11	1	1	5	1	1	2	서울일TOT	없음	동아시아	2
3	327	1	1	1	12	3	2	5	1	1	5	서울일TOT	없음	기타	4
4	328	1	3	1	2	3	4	4	1	1	5	부산일TOT	없음	기타	1
...
15431	41149	4	3	1	11	3	3	4	4	1	6	서울일TOT	인천일TOT	북아메리카	14
15432	41152	3	1	1	1	2	4	5	4	1	6	부산일TOT	없음	동아시아	3
15433	41156	4	1	1	1	2	5	5	4	2	6	부산일TOT	없음	동아시아	2
15434	41159	4	3	1	12	1	1	4	4	1	3	부산일TOT	없음	유럽	14
15435	41161	1	1	1	2	3	4	5	4	1	3	서울일TOT	부산일TOT	유럽	4

15436 rows x 15 columns

15436개의 열과
15개의 행을
갖는 데이터 도출

1. 문제 발굴

범주형 변수에 대한 통계적 처리가 가능한 잠재적계층분석(LCA분석)을 활용해 관광객들을 계층(Class)으로 유형화 할 수 있었음

잠재적계층분석(LCA분석)

$$P(Y = y) = \sum_{l=1}^C P(Y = y, L = l) \\ = \sum_{l=1}^C \gamma_l \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{mk|l}^{I(y_m=k)}$$

- $Y = (Y_1, \dots, Y_M)$: 범주형 변수
- $y_m = 1, \dots, r_m$: r_m 개의 응답범주를 갖는 범주형 변수
- $\gamma_l = P(L = l)$: l 번째 잠재계층의 소속 확률
- $\rho_{mk|l} = P(Y_m = k | L = l)$: 범주형 변수의 각 항목에 대한 응답확률

- 잠재적계층분석(LCA 분석)은 응답결과의 확률적 분포를 가정해 통계적으로 추정하므로 주어진 자료에 의한 통계적 근거를 바탕으로 집단의 분류를 설명할 수 있어 논리적인 분석 방법임
- 특히 잠재적계층분석(LCA 분석)은 범주형 변수를 통해 분류를 하는데 효과적임

잠재적계층분석 코드 구현



poLCA 패키지를 활용해 최적 군집수 도출 및 LCA 분석 진행

```
foreign_lca_list <- c()
set.seed(1234)

for (k in 1:10){foreign_lca_list[[k]] <- poLCA(foreign_formula, foreign_fac, nclass = k,
verbose = FALSE, na.rm = TRUE)}

optimal_k_result <- data.frame(k = NULL, aic = NULL, bic = NULL)

for (k in 1:10){result <- data.frame(k = k,
                                     aic = foreign_lca_list[[k]]$aic,
                                     bic = foreign_lca_list[[k]]$bic)

optimal_k_result <- rbind(optimal_k_result, result)}

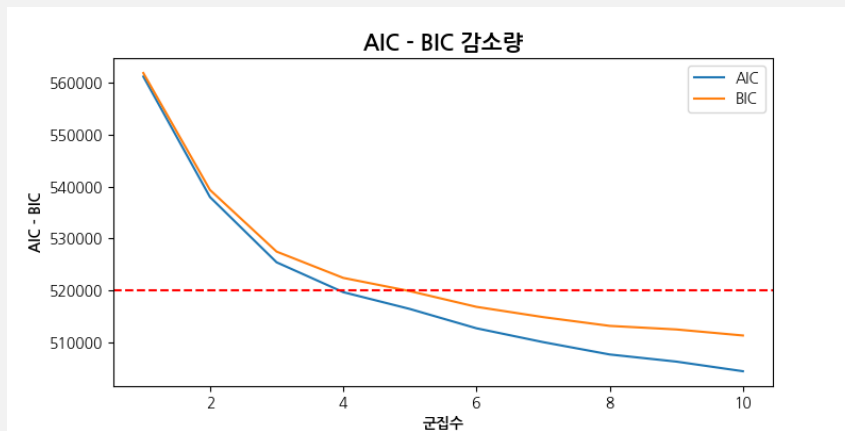
foreign_optimal <- foreign_lca_list[[4]]
```

- 잠재적계층분석(LCA 분석) 에서 추정할 두 가지 모수인 계층의 크기(class prevalence)와 범주형 변수의 각 항목에 대한 응답확률(Item-response probability)를 추정하기 위해 최적 세분시장의 개수를 결정해야 함
- 최적 계층의 수를 결정하는 기준으로 AIC와 BIC가 주로 활용됨

1. 문제 발굴

AIC BIC의 감소량을 기준으로 최적의 계층수(Class)를 4개로 설정할 수 있었으며,
LCA 분석의 항목 반응 행렬은 다음과 같음

AIC BIC 감소량 시각화



- 군집수를 1개에서 10개로 한정하여 분석을 한 결과 나타난 AIC와 BIC의 변화량은 위 그래프와 같음
- AIC - BIC 감소량이 급격하게 줄어드는(기울기가 완만해지는) 지점인 520000를 기준으로 하여 군집수 4개를 최적 모형으로 결정

LCA 분석의 항목반응행렬

Unnamed: 0	class 1:	class 2:	class 3:	class 4:
방한횟수_1회	0.695000	0.439000	0.331000	0.354000
방한횟수_2회	0.156000	0.172000	0.196000	0.172000
방한횟수_3회	0.066000	0.098000	0.124000	0.112000
방한횟수_4회이상	0.083000	0.291000	0.349000	0.362000
주요방한목적_여가위락	0.722000	0.145000	0.842000	0.004000
주요방한목적_친구친지방문	0.180000	0.450000	0.067000	0.004000
주요방한목적_사업	0.029000	0.190000	0.069000	0.988000
주요방한목적_교육	0.069000	0.090000	0.007000	0.001000
주요방한목적_종교및순례	0.000000	0.009000	0.000000	0.001000
주요방한목적_기타	0.000000	0.116000	0.015000	0.001000
여행형태_개별여행	0.964000	0.926000	0.884000	0.857000
여행형태_AIRTEL	0.006000	0.022000	0.036000	0.115000
여행형태_단체	0.029000	0.052000	0.079000	0.027000
체류기간_1-3일	0.000000	0.008000	0.172000	0.090000
체류기간_4-7일	0.126000	0.127000	0.667000	0.541000
체류기간_8-10일	0.229000	0.175000	0.122000	0.148000
체류기간_11-30일	0.583000	0.491000	0.039000	0.184000
체류기간_30-60일	0.061000	0.199000	0.000000	0.038000
1인경비_\$1000이하	0.018000	0.090000	0.329000	0.044000
1인경비_\$1000~2000	0.223000	0.241000	0.403000	0.160000

⋮

*항목반응행렬:군집 별 속성이 속한 확률을 나타내는 행렬

1. 문제 발굴

LCA 분석 결과 도출된 클래스 별로 EDA 하기 위해 클래스라는 새로운 변수로 추가 하였으며, 간편한 EDA를 위한 전처리 작업 진행

전처리 이전의 데이터

ID	방한횟수	목적	형태	방문기간	1인자출경비	1day경비	만족도	분기별	성별	나이	채류1순위	채류2순위	대륙	만족활동	클래스	
0	289	4	2	1	12	3	2	5	1	1	2	서울일TOT	없음	유럽	4	2
1	306	1	3	1	2	3	4	4	1	1	5	부산일TOT	없음	기타	14	4
2	324	1	2	1	11	1	1	5	1	1	2	서울일TOT	없음	동아시아	2	2
3	327	1	1	1	12	3	2	5	1	1	5	서울일TOT	없음	기타	4	1
4	328	1	3	1	2	3	4	4	1	1	5	부산일TOT	없음	기타	1	4
...
15431	41149	4	3	1	11	3	3	4	4	1	6	서울일TOT	안전일TOT	북아메리카	14	4
15432	41152	3	1	1	1	2	4	5	4	1	6	부산일TOT	없음	동아시아	3	3
15433	41156	4	1	1	1	2	5	5	4	2	6	부산일TOT	없음	동아시아	2	3
15434	41159	4	3	1	12	1	1	4	4	1	3	부산일TOT	없음	유럽	14	4
15435	41161	1	1	1	2	3	4	5	4	1	3	서울일TOT	부산일TOT	유럽	4	1

15436 rows x 16 columns

전처리 이후의 데이터

ID	방한횟수	목적	형태	방문기간	1인자출경비	1day경비	만족도	분기별	성별	나이	채류1순위	채류2순위	대륙	만족활동	클래스	
0	289	4회 이상	친구/친지 방문	개별여행	31-60일	\$2000-3000이하	\$50-100이하	매우만족	1분기	남성	21-30세	서울일TOT	없음	유럽	역사문화체험	2
1	306	1회	비즈니스	개별여행	4-7일	\$2000-3000이하	\$300-5000이하	대체로만족	1분기	남성	51-60세	부산일TOT	없음	기타	업무수행	4
2	324	1회	친구/친지 방문	개별여행	11-30일	\$1000이하	\$50이하	매우만족	1분기	남성	21-30세	서울일TOT	없음	동아시아	쇼핑	2
3	327	1회	여가/워킹/휴식	개별여행	31-60일	\$2000-3000이하	\$50-100이하	매우만족	1분기	남성	51-60세	서울일TOT	없음	기타	역사문화체험	1
4	328	1회	비즈니스	개별여행	4-7일	\$2000-3000이하	\$300-5000이하	대체로만족	1분기	남성	51-60세	부산일TOT	없음	기타	식도락관광	4
...
15431	41149	4회 이상	비즈니스	개별여행	11-30일	\$2000-3000이하	\$100-300이하	대체로만족	4분기	남성	61세이상	서울일TOT	안전일TOT	북아메리카	업무수행	4
15432	41152	3회	여가/워킹/휴식	개별여행	1-3일	\$1000-2000이하	\$300-5000이하	매우만족	4분기	남성	61세이상	부산일TOT	없음	동아시아	자연경관감상	3
15433	41156	4회 이상	여가/워킹/휴식	개별여행	1-3일	\$1000-2000이하	\$500초과	매우만족	4분기	여성	61세이상	부산일TOT	없음	동아시아	쇼핑	3
15434	41159	4회 이상	비즈니스	개별여행	31-60일	\$1000이하	\$50이하	대체로만족	4분기	남성	31-40세	부산일TOT	없음	유럽	업무수행	4
15435	41161	1회	여가/워킹/휴식	개별여행	4-7일	\$2000-3000이하	\$300-5000이하	매우만족	4분기	남성	31-40세	서울일TOT	부산일TOT	유럽	역사문화체험	1

15436 rows x 16 columns

For Example

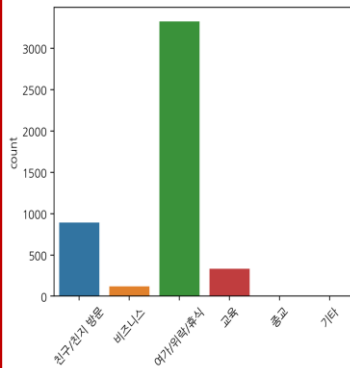
```
df['방한횟수'] = df['방한횟수'].map({1: '1회', 2: '2회', 3: '3회', 4: '4회 이상'})
df['목적'] = df['목적'].map({1: '여가/워킹/휴식', 2: '친구/친지 방문', 3: '비즈니스', 4: '교육', 5: '종교', 6: '기타'})
df['형태'] = df['형태'].map({1: '개별여행', 2: 'AIR-TEL', 3: '단체여행'})
df['방문기간'] = df['방문기간'].map({1: '1-3일', 2: '4-7일', 3: '8-10일', 11: '11-30일', 12: '31-60일'})
df['1인자출경비'] = df['1인자출경비'].map({1: '$1000이하', 2: '$1000-2000이하', 3: '$2000-3000이하', 4: '$3000-4000이하', 5: '$5000-10000이하', 6: '$10000초과', 98: '초과'})
df['1day경비'] = df['1day경비'].map({1: '$50이하', 2: '$50-100이하', 3: '$100-300이하', 4: '$300-5000이하', 5: '$500초과'})
df['만족도'] = df['만족도'].map({1: '매우불만족', 2: '대체로불만족', 3: '보통', 4: '대체로만족', 5: '매우만족'})
df['분기별'] = df['분기별'].map({1: '1분기', 2: '2분기', 3: '3분기', 4: '4분기'})
df['성별'] = df['성별'].map({1: '남성', 2: '여성'})
df['나이'] = df['나이'].map({1: '15-20세', 2: '21-30세', 3: '31-40세', 4: '41-50세', 5: '51-60세', 6: '61세이상'})
```

1. 문제 발굴

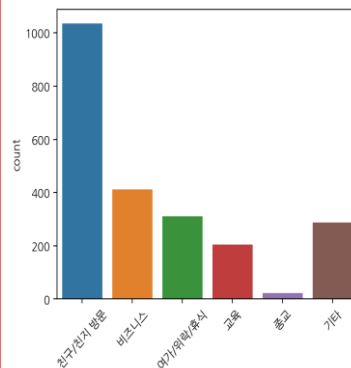
각각의 클래스를 활용해 EDA를 진행한 결과 다음과 같은 결과가 나왔으며,
장기체류 외국인 관광객인 Class 1과 단기체류 외국인 관광객 Class3에 집중하여 이후 분석 진행

방문
목적

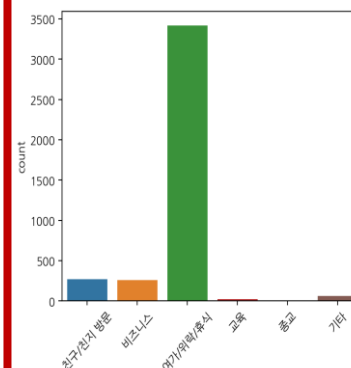
Class 1



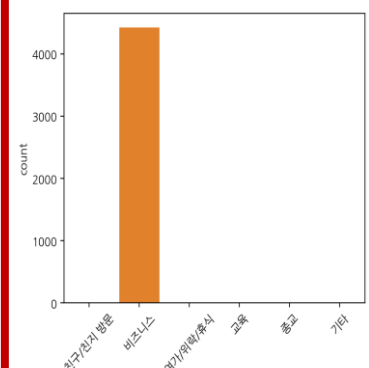
Class 2



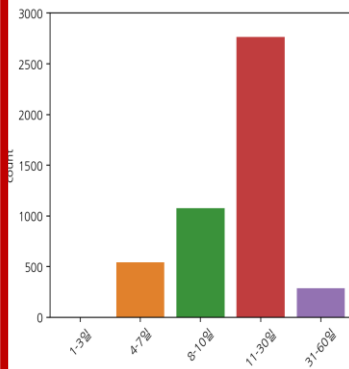
Class 3



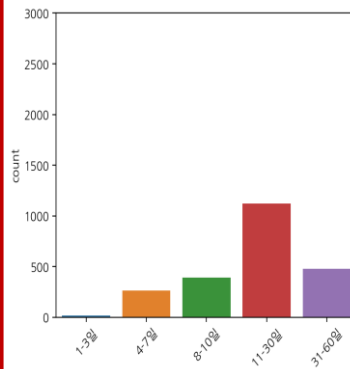
Class 4



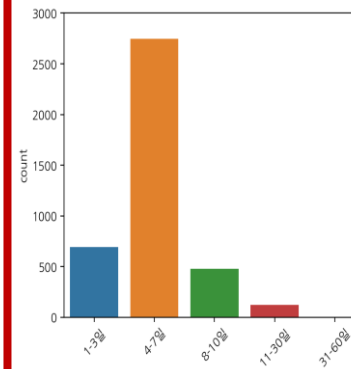
체류
기간



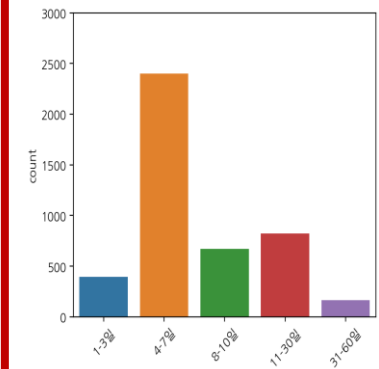
여행을 위해
방문하였으며
상대적으로 **장기간 체류**



다양한 목적의 방문이
혼재되어 있으며,
상대적으로 장기간 체류



여행을 위해
방문하였으며
상대적으로 **단기간 체류**

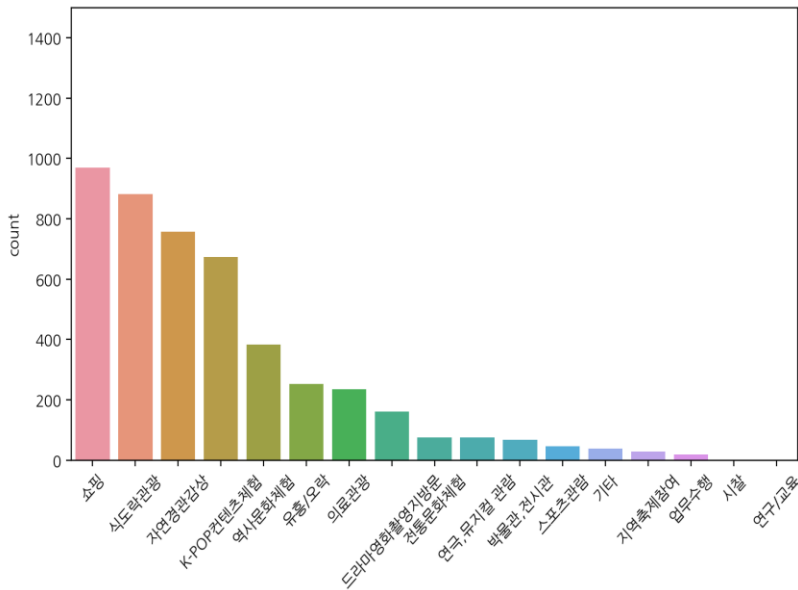


비즈니스를 위해
방문하였으며
상대적으로 단기간 체류

1. 문제 발굴

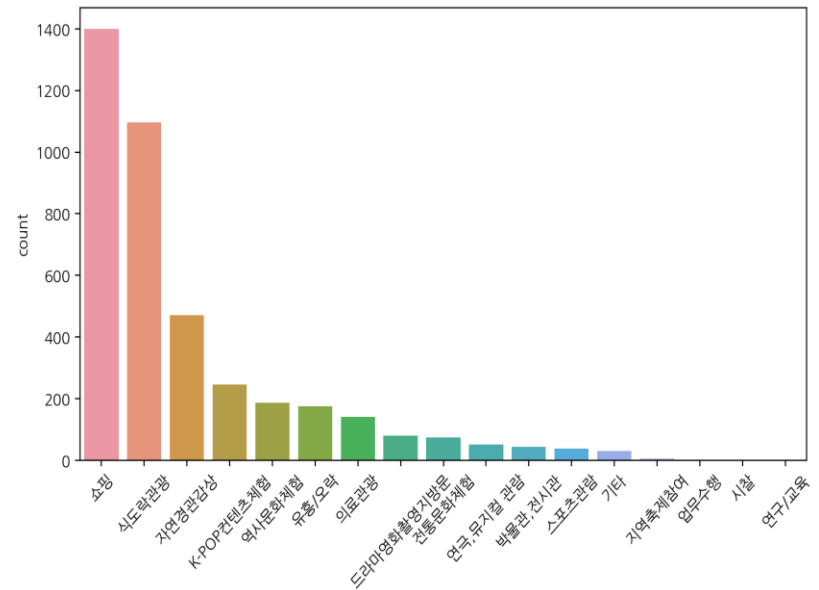
Class1과 Class3를 바탕으로 가장 만족한 활동을 EDA 한 결과 다음과 같으며,
더 다양하게 국내 관광 콘텐츠를 향유하는 Class1을 최종 타겟으로 선정

Class 1(장기체류 외국인 관광객)



- Class1(장기체류 외국인 관광객)은 다양한 관광콘텐츠들에서 만족경험이 높으며 Class3(단기체류 외국인 관광객)에 비해 더 역동적으로 국내 관광을 즐김
- 특히 '자연경관감상' 'K-POP컨텐츠체험' 에서 Class3(단기체류 외국인 관광객)에 비해 가장 만족한 비율이 압도적으로 높음

Class 3(단기체류 외국인 관광객)



- Class3(단기체류 외국인 관광객)은 쇼핑과 식도락관광이 압도적으로 높은 비율을 차지하며, Class1(장기체류 외국인 관광객)에 비해 즐기는 관광콘텐츠가 제한적임

가장
만족한
활동

2. 문제 제기

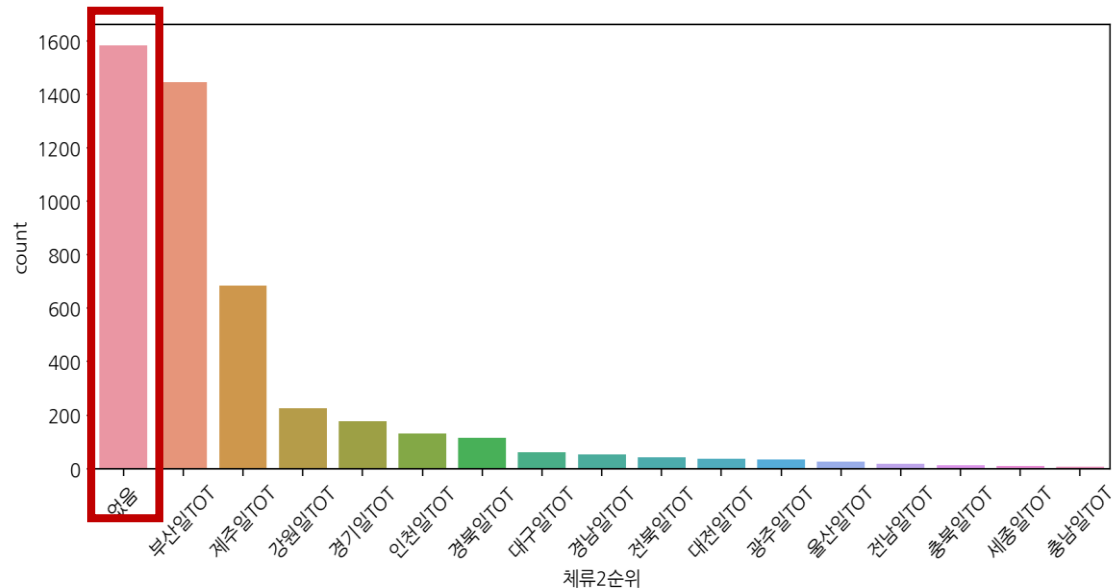
Class1(장기체류 외국인 관광객) 내에서 체류1순위 체류2순위 지역을 EDA 한 결과 방문하는 지역이 서울, 부산, 제주, 강원 등 일부 지역으로 한정되는 문제가 존재함

Class1(장기체류 외국인 관광객)의 지역별 체류현황

체류1순위 지역

	체류1순위	count
0	서울일TOT	4646
1	제주일TOT	8
2	대구일TOT	7
3	경남일TOT	3
4	대전일TOT	2
5	충남일TOT	2
6	울산일TOT	1
7	경기일TOT	1

체류 2순위 지역



- *체류1순위: 외국인들의 지역별 체류 기간 중 가장 오래 머문 지역
- *체류2순위: 외국인들의 지역별 체류 기간 중 두번째로 오래 머문 지역

- Class1(장기체류 외국인 관광객)은 대체적으로 한국을 방문하면 서울에서 가장 오랜 시간을 보냄

- 국내에서 체류기간이 평균 16.2일인 Class1(장기체류 관광객)들도 체류2순위 지역은 '없음'이 가장 많을 만큼, 서울 외의 지역은 잘 방문하지 않으며, 방문을 한다고 하더라도 부산광역시, 제주특별자치도, 강원특별자치도에 많이 한정되어 있음

2. 문제 제기

지역관광활성화는 다음의 장점을 띄며 장기체류 외국인 관광객의 특징을 고려했을 때 효과적일 수 있음.
이후 지역관광 후보지를 선정하여 이들을 대상으로 관광지가 일부지역으로 한정되는 문제를 해결하고자 함.

지역관광 활성화의 장점

장기체류 외국인 관광객의 특징

지역관광 활성화 지역 제안

국내 관광 경쟁력 제고

- 국내 관광 콘텐츠의 다양화를 꾀할 수 있음

지역경제 활성화

- 관광산업은 파급효과가 큰 산업이므로 해당 지역 외에 인근 지역까지 지역경제 활성화에 효과가 있음

오버투어리즘 예방

- 현재 국내 여행은 수도권, 제주, 부산, 강원도에 많이 치중되어 있는 상태인데, 다양한 지역관광 활성화를 통해 오버투어리즘을 예방할 수 있음

지역소멸 예방

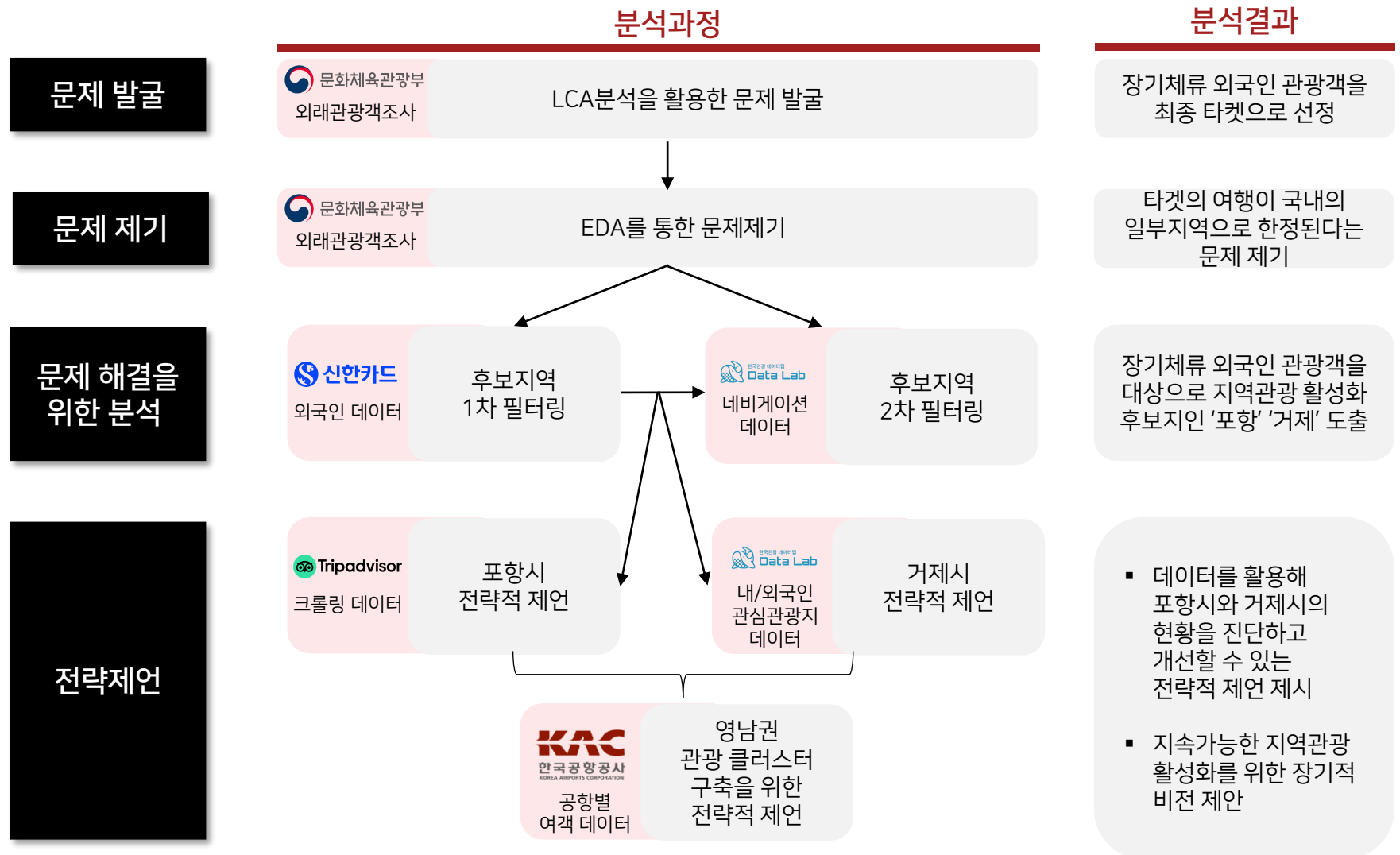
- 지역 관광활성화를 통해 지역에 기반시설과 산업 강화를 통해 지역 소멸 문제를 예방할 수 있음

- 평균 체류기간이 약 16.2일
- 자연경관 감상, K POP 콘텐츠 체험 등 다양한 관광 콘텐츠 향유
- 국내 여행 시에 방문하는 관광지가 한정됨

이후의 내용에서는
장기체류 외국인 관광객 대상
지역관광 활성화 후보지역 제안

3. 해결 방안 분석

데이터를 활용한 분석 로드맵은 다음과 같음



- 데이터를 활용해 포항시와 거제시의 현황을 진단하고 개선할 수 있는 전략적 제언 제시
- 지속가능한 지역관광 활성화를 위한 장기적 비전 제언

3. 해결 방안 분석

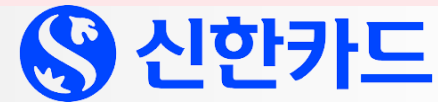
앞선 내용과 이후의 내용에서 활용한 데이터는 다음과 같음

2022 외래관광객조사



- 활용내용 : '문제발굴' 및 '문제제기' 파트에서 잠재적 계층분석 (LCA 분석)을 사용하기 위해 활용
- 전처리방법: 잠재적 계층분석 (LCA 분석)을 사용할 수 있는 변수 출력 및 변수 범주화
- 활용결과: 장기체류 관광객도 국내 여행 시 방문하는 지역이 한정적이라는 문제 도출

신한카드 외국인데이터



- 활용내용 : '문제 해결을 위한 분석' 파트에서 외국인들이 많이 찾는 국내 관광지를 찾기 위해 활용
- 전처리방법: 수도권, 부산 등 이미 관광이 활성화 되어 있는 지역 제외 및 관광과 관련된 지출만을 활용하기 위해 '숙박'에 해당하는 데이터만 확인
- 활용결과 : 수도권과 부산 제주 등을 제외한 인바운드 관광이 활성화된 상위 20개 시도 도출

한국관광데이터랩 데이터



- 활용내용 : '문제 해결을 위한 분석' 과 거제와 관련된 '전략제언' 파트에서 활용
- 전처리방법: 타겟으로 하는 지역만을 선정해 분석진행
- 활용결과 :
 - 1) '문제 해결을 위한 분석' : 장기 체류 외국인 관광객이 좋아할만한 관광 콘텐츠가 있는 지역 도출
 - 2) '전략제언' : 거제시 현재 관광 문제점 파악

트립어드바이저 리뷰 크롤링 데이터



- 활용내용 : '포항과 관련된 '전략제언' 파트에서 활용
- 전처리방법: 토큰화, 불용어 처리, 표제어 추출 등을 활용한 전처리 진행
- 활용결과 : 포항시 현재 관광 문제점 파악

활용
데이터

3. 해결 방안 분석

신한카드 외국인 데이터를 활용해 지역관광 활성화 목적에 부합하는 1차 후보지 20개 지역 선정

신한카드 데이터 전처리 및 1차 후보지역 선정



전처리 내용

기존 관광
활성화 지역 제외

- '지역관광 활성화' 목적에 부합하는 지역을 선정하기 위해 수도권(서울, 경기, 인천), 부산, 제주, 강원 지역 등 상대적으로 외국인 방문객이 많은 지역은 제외

관광 목적의 카드
사용으로 한정

- 관광을 목적으로 지역에 방문한 외국인들로 분석을 한정하기 위해 카드 사용 내역 중 '업종(대분류)'가 '숙박'인 것으로 필터링

이용 건수
기준 상위 20개
지역 추출

- 카드 이용 건수가 많은 상위 20개 지역을 추출하여 1차 후보 지역으로 선정
- 1차 후보 지역은 이후 2차 후보 선정 작업에 활용됨



전처리과정 코드 구현

```
1 # 카드 데이터(외국인) 로드
2 for_df = pd.read_csv("/content/drive/MyDrive/2023년 문화관광 데이터 #
3 분석대회(카드데이터_외국인).txt", encoding = "utf-8", sep = '¶')
4
5 # 컬럼명 변경
6 new_col_name = ["국적", "가맹점(시도)", "가맹점(시군구)",
7                 "업종(대분류)", "업종(소분류)", "이용년월",
8                 "요일", "이용시간대", "취급액", "이용건수"]
9 for_df.columns = new_col_name
10
11 # '가맹점(시도)' 기준 '서울', '경기', '인천', '제주', '부산', '강원'
12 county = for_df[(for_df["가맹점(시도)"] != "서울") &
13                 (for_df["가맹점(시도)"] != "경기") &
14                 (for_df["가맹점(시도)"] != "인천") &
15                 (for_df["가맹점(시도)"] != "제주") &
16                 (for_df["가맹점(시도)"] != "부산") &
17                 (for_df["가맹점(시도)"] != "강원")]
18
19 # 카드 사용 내역 중 '업종(대분류)'이 '숙박'인 것으로 필터링
20 tourist = county[ (county["업종(대분류)"] == "숙박") ]
21
22 # 카드 사용 내역 중 '이용건수' 함께 기준 상위 20개 지역 추출
23 toursit = tourist.groupby(["가맹점(시도)",
24                             "가맹점(시군구)"])["이용건수"].agg(["sum"]).reset_index()
25 toursit.columns = ["가맹점(시도)", "가맹점(시군구)", "이용건수"]
26 toursit = toursit.sort_values(by = "이용건수", ascending = False)
27 toursit = toursit.reset_index(drop = True).iloc[:20,:]
28 toursit
```

3. 해결 방안 분석

관광지의 특성을 파악할 수 있는 네비게이션 데이터를 활용해 장기체류 외국인 관광객이 선호하는 관광 유형인 '자연관광', '문화관광'에 부합하는 최종 후보지 선정

한국관광 데이터랩 데이터 전처리 및 최종후보 선정



장기체류 외국인 관광객의 특성

LCA 및 EDA

Class1에 해당하는 외래 관광객들 다수가 '자연경관감상', 'K-POP컨텐츠 체험'에 대한 만족도가 높음

'자연'과 '문화' 관련 콘텐츠가 풍부한 지역을 제안할 필요가 있음



각 후보지의 관광지 특성 파악

네비게이션 데이터

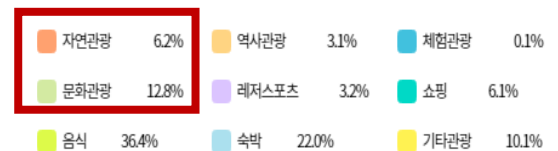
네비게이션 검색 데이터를 통해 각 후보지의 관광지 유형 비중을 파악할 수 있음

'자연관광'과 '문화관광' 비중이 높은 관광지를 도출할 필요가 있음

For example

전라남도 여수시 : 유형별 검색건수

[데이터 다운로드 시 세부유형별 검색건수를 확인하실 수 있습니다.]



"자연관광 문화관광 비율의 합을 최종 후보 선정 기준으로 활용"

3. 해결 방안 분석

앞선 과정을 거쳐서 최종후보로 포항시와 거제시를 선정 하였으며,
이후의 내용에서는 포항시와 거제시의 장기체류 외국인 관광객을 위한 전략 제언을 진행

가맹점(시도)	가맹점(시군구)	취급액sum	이용건수sum	자연_비율	문화_비율	자연+문화_비율
경남	거제시	5866996599	19105	13.4	12.4	25.8
경북	포항시	7257935194	24239	11.25	9.05	20.3
전남	여수시	6112014189	25082	6.2	12.8	19
대구	중구	5166700794	28957	0	16.8	16.8
전북	군산시	6764843739	23090	9.2	6.6	15.8
전북	전주시	4619811945	30123	2.35	13.3	15.65
충남	부여군	2220222805	10857	1.2	14.2	15.4
울산	남구	11333991837	56560	1.1	14.2	15.3
울산	동구	3267056780	8432	6.5	8.3	14.8
대구	수성구	11852899146	75452	4.6	10.2	14.8
경북	경주시	25367293521	95986	2.7	12	14.7
경남	창원시	10468658090	33841	2.8	9.9	12.7
광주	서구	8567108864	46371	0.4	10.6	11
충북	청주시	11769461741	30835	0.75	9.25	10
대전	유성구	19810638950	68716	0.1	9.8	9.9
경북	구미시	4887867179	18909	0.8	9	9.8
경남	통영시	6364732410	26688	4.1	5.3	9.4
충남	아산시	2071203229	13555	2.7	5.8	8.5
대구	동구	10091014422	19274	0.2	7.7	7.9
충남	천안시	12236706509	22504	0.4	6	6.4

poohang
포항시

거 제 시
GEOJE-SI

장기체류 외국인 관광객을
대상으로, 지역관광 잠재력이 있는
포항시와 거제시를 위한 전략제언

4. 전략적 제언 - 포항시를 중심으로

지금까지 포항을 여행한 외국인 관광객들의 관광 패턴을 파악하기 위해 대표적인 여행 플랫폼 사이트인 Tripadvisor의 고객 리뷰 데이터를 크롤링 후 전처리 진행

리뷰 데이터를 통한 포항시 관광 패턴 분석



```
0    It's a full quater of fish shops ,hundreds and...
1    Really peculiar , a mix between an aquarium an...
2    I was there in an evening , they were closing ...
3    The market is a fascinating place. There is a ...
4    If you wish to eat seafood, you can pick your ...
    ...
164    Interesting place.
165    Has a statue of a dumbass near the bow. Read t...
166    Stumbled across this location after seeing it ...
167    It is one bus away from Pohang Terminal maybe ...
168    first time to visit inside a warship, well org...
    :
```

IDEA	소셜미디어 분석을 통한 외국인관광객들의 관광패턴 파악
Reason	가장 많은 외국어 리뷰 보유
Mechanism	키워드로 'Pohang' 검색 시 나오는 영문 리뷰 200여개 크롤링

텍스트 전처리

```
def my_tokenizer(r):
    r1 = word_tokenize(r)
    r2 = [w.lower() for w in r1 if w.isalpha()]
    r3 = [w for w in r2 if not w in new_stopwords]
    r4 = [lem.lemmatize(w, get_pos(w)) for w in r3]
    return(r4)
```

1 토큰화

- 단어 단위의 분석을 하기 위해 토큰화 작업 진행

2 불용어처리

- 불용어처리를 통해 중요도가 떨어지는 단어 삭제
 - scikit-learn의 불용어를 통한 처리
 - 리뷰 내에서 빈번히 등장하는 단어 삭제

3 표제어 추출

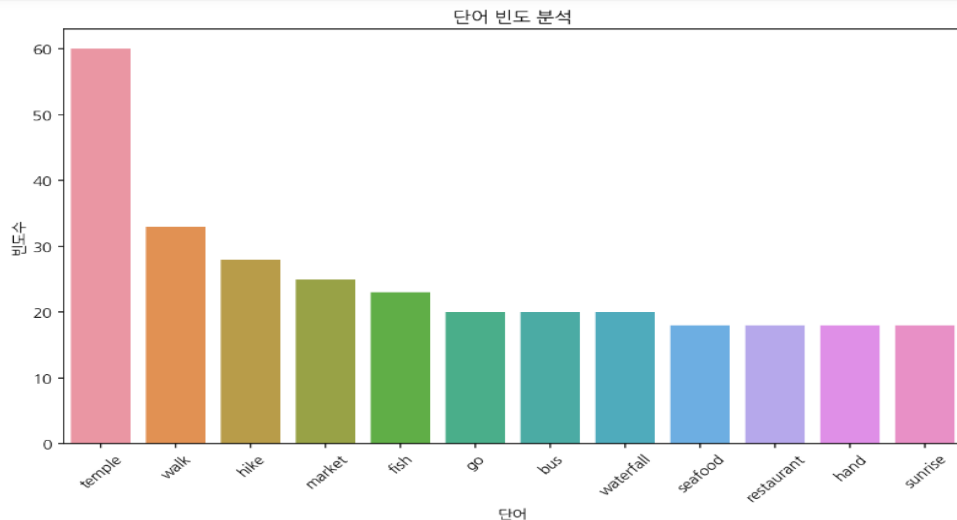
- 단어의 원형으로 처리하기 위해 사전형으로 단어 추출
 - NLTK 패키지의 Lemmatizer 모듈을 이용하여 normalization 진행

4. 전략적 제언 - 포항시를 중심으로

텍스트 분석 결과 외국인 관광객들 대부분이 포항의 한정된 관광지와 여행 콘텐츠에 국한되어 있음을 확인했고 이러한 문제를 개선해야 함

빈도 분석을 통한 상위 빈도 키워드 추출

Counter 함수를 통한 빈도 출력 결과



상위 빈도 30개 단어 기준 해석 결과

- 주요 방문 관광지 : temple , museum, hand, sunrise
- 자연관련 액티비티 : Walk, hike, trail, waterfall, water, beach
- 죽도시장 : Market, fish, crab, seafood

빈도 분석 시사점

- 추출된 단어의 빈도 분석 결과 포항을 방문한 외국인 관광객들의 여행 패턴들을 확인할 수 있었음
- 특히, temple , market , hand, sunrise 등 포항의 관광지인 보경사, 죽도시장, 호미곶 등을 연상할 수 있는 지역들이 많이 추출됨
- 이를 통해 외국인들의 관광패턴이 일부 관광지와 액티비티에 집중되어 있음을 확인할 수 있었음

**“한정된 관광지와
여행 콘텐츠에서 벗어나
다양한 선택지를 제공해야 하고,
적극적으로 홍보할 필요가 있음”**

4. 전략적 제언 - 포항시를 중심으로

최근 포항에서는 K-드라마 촬영지를 관광지로서 활성화하고자 하는 노력을 하고 있으나, 글로벌한 홍보 및 외국어 정보 부족으로 이를 해결할 필요가 있음

K컨텐츠를 활용한 외래 관광객 유치 전략

포항, K-드라마 촬영지 눈길...한류 관광지 발길 이어져

‘한류 드라마가 이 곳에서?’ 인기 촬영명소 포항, 대만 국제관광박람회 사로잡아

포항시, K-드라마 촬영지로 눈길 한류 관광지로 발길

‘동백꽃 필 무렵’, ‘갯마을 차차차’ 이어 최근 ‘꼭두의 계절’ 배경 눈길

매력적인 포항의 배경 녹아든 드라마 감동과 여운 즐기려는 관광객 발길 지속

한국관광공사, 포항 청하공진시장과 구룡포 ‘한류 성지순례 여행지’로 최근 선정

드라마 물결 타고 나들이 인파 몰려 지역 홍보와 경제 활성화 두 마리 토끼 잡아

걸음마다 드라마 같은, 모든 순간이 여행인 관광도시 포항 조성 매진

문제점

현재 K-드라마 촬영지를 활용한
관광지는 **국내 여행객들을**
중심으로 홍보가 이루어져 있음

해외는 동아시아(일본, 중국,
대만)에 **관련 홍보가 국한**되어
있음 글로벌 시장을 대상으로 한
홍보도 필요함

관광지 내 **외국어 정보 부족**
문제도 해결되어야 함

해결방안

해당 드라마가 방영되는 해외
국가들을 대상으로 관광상품
패키지를 개발 공급

해당 드라마가 공개된 OTT
사이트를 통해 관광 상품
패키지를 홍보하는 방안이 있음

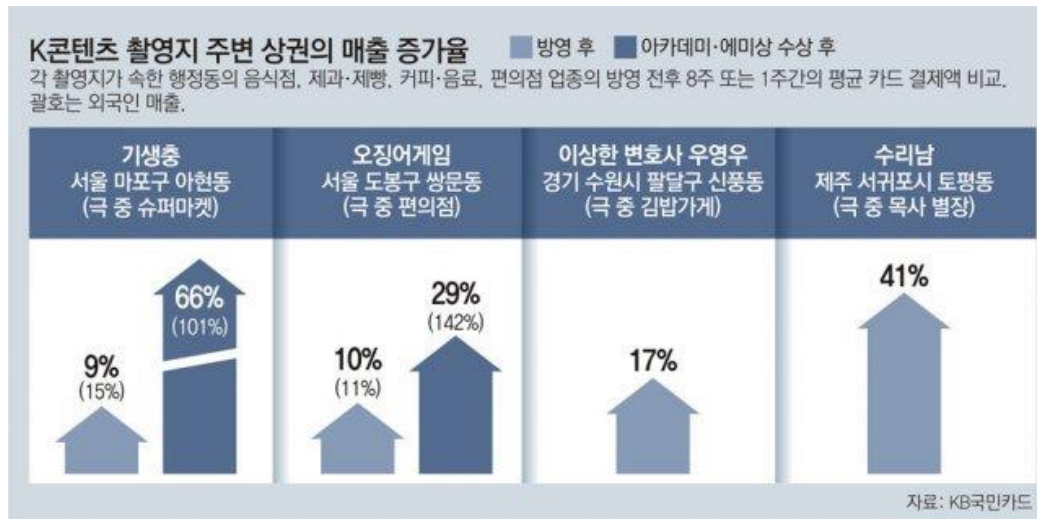
외국인들이 불편함 없이
K-드라마 촬영지를 방문할 수
있도록 외국어 가이드 시설과
충분한 정보를 해외 여행
플랫폼을 통해 공개할 필요가
있음

4. 전략적 제언 - 포항시를 중심으로

K-드라마 관광지 활성화 전략은 경제적 효과를 기대할 수 있지만, 지속가능 하기 위해서는 관광객을 유인할 수 있는 추가적인 요인이 필요

K콘텐츠를 활용한 외래 관광객 유치 전략

기대 효과



- 일부 콘텐츠(수산물시장, 자연 등)에 한정되어 있던 포항의 관광 콘텐츠를 다각화 할 수 있음
- 촬영지 주변 상권이 활성화됨으로서 지역 경제에 긍정적인 효과를 기대할 수 있음

한계

- K-콘텐츠를 활용한 관광지들은 시간이 지날수록 해당 콘텐츠의 **화제성이 떨어지게** 되어 관광객들의 일시적인 방문으로만 그칠 위험성이 있음
- 따라서, 지속가능한 관광지 활성화를 위해서는 K-콘텐츠에 더하여 **관광객들을 유인할 수 있는 매력적인 요소**를 추가 보완할 필요가 있음

4. 전략적 제언 - 거제시를 중심으로

거제시를 검색한 내 / 외국인의 관심 관광지를 비교하기 위해
한국관광 데이터랩의 원본 데이터를 결합하여 하나의 데이터셋으로 전처리를 진행

거제시 내/외국인 관심 관광지 전처리 과정



원본 데이터

내국인 관심 관광지		외국인 관심 관광지	
이름	구분	이름	구분
이수도(거제)	자연관광지	Oedo-Botania (외도 보타니아)	자연관광지
바람의 언덕	자연관광지	Windy Hill (바람의 언덕)	자연관광지
외도 보타니아	자연관광지	Geogje Haengumgang Island (거제도 해금강)	자연관광지
매미성	자연관광지	Historic Park of Geogje POW Camp (거제도 포로수용소 유적공원)	역사관광지

출처 : VisitKorea 사이트 로그를 기반으로 하여 조회가 많은 관광지

영어권, 일본어, 중간체, 중번체 데이터를 결합하여 단일 데이터로 통합

출처 : '대한민국 구석구석' 사이트 로그를 기반으로 하여 조회가 많은 관광지

결합 데이터

➤ 내/외국인 관심 관광지 순위를 한눈에 비교하기 위해 두 데이터셋을 전처리 후 left join 결합

pd.merge(geogje_visit_kor, geogje_visit_for, how = 'left', on = '관광지명')

시군구명_x	관광지명	KTO가이드고리대분류명_x	방문자 수	순위_x	시군구명_y	KTO가이드고리대분류명_y	사용자방문건 수	순위_y
0 거제시	이수도(거제)	자연	8737	1	NaN	NaN	NaN	NaN
1 거제시	바람의 언덕	자연	8582	2	거제시	자연	3200	2
2 거제시	외도 보타니아	자연	4898	3	거제시	자연	7493	1
3 거제시	저구항	자연	3043	4	NaN	NaN	NaN	NaN
4 거제시	거제 관광모노레일	인문(문화/예술/역사)	3021	5	거제시	인문(문화/예술/역사)	27	22
5 거제시	매미성	자연	2834	6	거제시	자연	54	20
6 거제시	거제맹종죽테마공원	자연	2577	7	거제시	자연	132	14
7 거제시	거제맹종죽테마공원	자연	2577	7	거제시	인문(문화/예술/역사)	28	21
8 거제시	거제도 포로수용소 유적공원	인문(문화/예술/역사)	2529	8	거제시	인문(문화/예술/역사)	1467	4
9 거제시	거제도 해금강	자연	2435	9	거제시	자연	1936	3

시군구명	관광지명	내국인 순위	외국인 순위
거제시	이수도(거제)	1	
거제시	바람의 언덕	2	2
거제시	외도 보타니아	3	1
거제시	저구항	4	
거제시	거제 관광모노레일	5	22
거제시	매미성	6	20
거제시	거제맹종죽테마공원	7	14
거제시	거제맹종죽테마공원	7	21
거제시	거제도 포로수용소 유적공원	8	4

For example

4. 전략적 제언 - 거제시를 중심으로

관광지 별 내국인, 외국인의 관심도에는 큰 차이가 존재하며 이에 대한 원인은 정보력의 차이, 트렌드를 반영하지 못한 관광지 추천에 기인한 것으로 추측됨

거제도 내/외국인 관심 관광지 분석 결과



시군구명	관광지명	유형	내국인 순위	외국인 순위
거제시	이수도(거제)	자연	1	
거제시	바람의 언덕	자연	2	2
거제시	외도 보타니아	자연	3	1
거제시	저구항	자연	4	
거제시	거제 관광모노레일	인문(문화/예술/역사)	5	22
거제시	매미성	자연	6	20
거제시	거제맹종죽테마공원	자연	7	14
거제시	거제도 포로수용소 유적공원	인문(문화/예술/역사)	8	4
거제시	거제도 해금강	자연	9	3
거제시	거제식물원(거제정글돔)	자연	10	19
거제시	산방산 비원	자연	11	
거제시	내도(거제)	자연	12	
거제시	학동흑진주몽돌해변	자연	13	7
거제시	지심도	자연	14	12
거제시	거제씨월드	인문(문화/예술/역사)	15	23
거제시	거제 대우조선해양 옥포조선소	인문(문화/예술/역사)	16	
거제시	근포동굴	자연	17	18
거제시	숲소리공원	자연	18	
거제시	소노캄 거제 마리나베이 요트투어	레포츠	19	
거제시	칠천도	자연	20	
거제시	해금강유람선	인문(문화/예술/역사)	21	
거제시	한화리조트 거제 벨버디어	인문(문화/예술/역사)	22	
거제시	바람꽃우체국	인문(문화/예술/역사)	23	
거제시	함목몽돌해변	자연	24	
거제시	거제도	자연	25	5
거제시	거제 계도어촌체험마을	인문(문화/예술/역사)	26	
거제시	백만석	음식	27	26
거제시	학동자동차야영장	레포츠	28	

현상 분석

- 내/외국인의 관심 관광지를 서로 비교한 결과 내국인과 외국인이 선호하는 관광지가 상이한 것을 확인할 수 있음
- 내국인 기준 상위 28개 관광지 중 절반인 14개의 관광지는 외국인의 검색 영역에 들어와 있지 않음 (파란색 부분)
- 또한, 내국인들에게 관심을 받고 있는 상위 관광지(ex. 거제 관광모노레일, 매미성 등) 중 일부는 외국인들의 관심도에서는 현저히 낮은 순위를 차지하고 있음

원인 분석

1. 내국인과 외국인의 여행 스타일 차이에서 발생한 문제
 2. 내국인과 외국인이 여행 계획 시 접근할 수 있는 정보력의 차이
 3. 최근 트렌드를 반영한 관광지 추천이 이루어지지 않고 있는 문제
- 2, 3에 집중하여 내국인과 외국인의 선호하는 관광지가 상이한 현상을 밝혀내고자 함

4. 전략적 제언 - 거제시를 중심으로

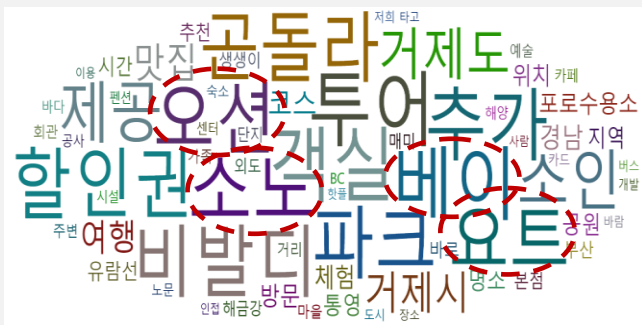
국내 SNS 키워드 분석을 통해 현재 내국인의 관심을 받고 있는 관광 콘텐츠를 중심으로 외국인 관광객들에게 추천 또는 정보를 제공해 앞서 언급한 현상을 해결할 수 있음

트렌드를 반영한 관광지 추천

SNS 분석을 통한 거제 관련 국내 키워드



1. 검색 키워드: 거제
2. 수집일자:
2022-07-25 ~ 2023-07-25
3. 채널: 블로그, 카페, 트위터



국내 SNS 키워드 검색을 통해 최근 거제의 해양 레저
관련 키워드들이 검색되고 있음

2023년 해외 여행 플랫폼에 소개된 거제 관광지



- **Gujora Beach**
- **Oedo Paradise Island**
- **Marine Cultural Center**
- **Geoje POW Camp**
- **Okpo Daecheop Ginyeom Park**
- **Gonggoji**
- **Maengjongsjuk Theme Park**
- **Oedo Botania**
- **Geoje Haegeumgang**
- **Windy Hill**
- **Maemiseong Castle**
- **Singseondae**
- **Hakdong Mongdol Beach**
- **Gyeryongsan**

하지만 2023년 기준 해외 여행 플랫폼을 통해 확인할 수 있는 거제 관광지들은 현재 트렌드를 반영하지 못하고 있음

위의 문제를 개선해 내국인들의 최근 트렌드를 반영한 해양 액티비티를
거제에 방문한 외국인 관광객들에게 추천할 수 있음

4. 전략적 제언 - 영남권 관광 클러스터 구축

포항시, 거제시에 대한 전략적 제언에 덧붙여 국제 공항을 중심으로 한 영남권 관광 클러스터를 형성하여 지역 관광 활성화를 증진할 수 있음

지역 관광을 활성화 하기 위한 관광 클러스터 구축

국제 공항을 중심으로 한 영남권 관광 클러스터

공항명	운항(편)			여객(명)		
	도착	출발	계	도착	출발	계
김포	3,228	3,229	6,457	522,678	525,338	1,048,016
김해	11,054	11,087	22,141	1,828,044	1,821,167	3,649,211
제주	1,331	1,331	2,662	199,716	190,431	390,147
대구	1,545	1,555	3,100	246,398	241,687	488,085
무안	74	73	147	11,186	11,592	22,778
청주	231	233	464	33,380	35,291	68,671
양양	400	404	804	35,898	35,515	71,413
인천	94,598	94,630	189,228	16,510,575	16,422,244	32,932,819
합계	112,461	112,542	225,003	19,387,875	19,283,265	38,671,140

* Source 한국공항공사 공항별 여객데이터

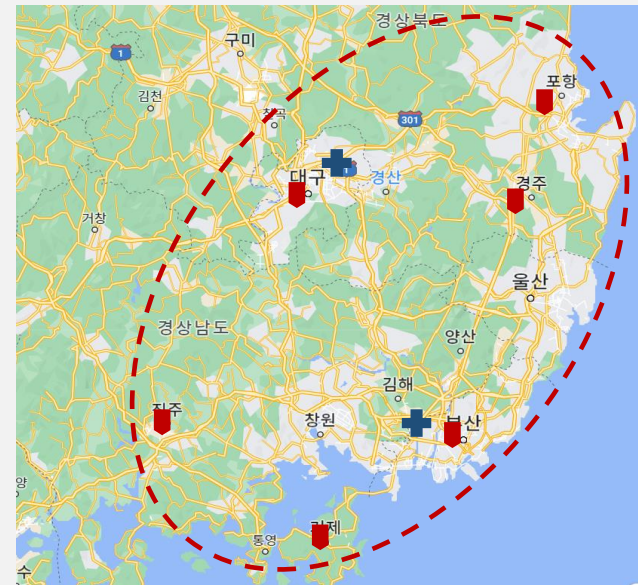
Opportunity

- 지역 관광을 활성화하기 위해서는 각 권역별 국제공항이 활성화되어야 함
- 국제 공항 2곳(김해, 대구)을 보유 중인 영남 지역은 해외 관광객을 유치하기 위한 충분한 조건을 갖추고 있음

Weakness

- 하지만 여전히 인천국제공항을 통한 입국객이 많으며, 지역 국제공항의 노선 또한 한정적이기 때문에 인프라 차원의 해결 또한 지역 관광에서 시급히 해결되어야 함
- 지자체와 항공사 상호 간의 노력으로 해소 가능함

현상분석



- 대구 국제 공항, 김해 국제 공항을 중심으로 한 영남권 관광 클러스터를 구축할 것을 제안함
- 특히, 영남권은 부산 경주 등 관광이 활성화 된 도시들이 존재하며, 거제 포항과 같이 관광 잠재력이 높은 도시들이 존재하기에 영남권 관광 클러스터의 구축은 지역관광에 큰 파급효과를 불러올 수 있음

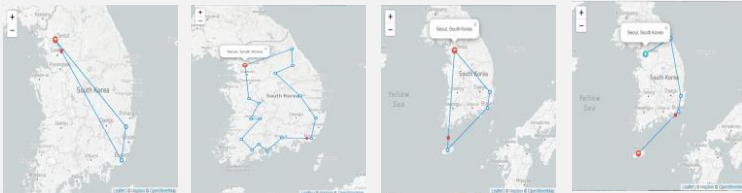
4. 전략적 제언 - 영남권 관광 클러스터 구축

위를 실현할 수 있는 구체적 전략으로 패키지 여행, 투어패스 등 다양한 관광 상품들을 결합할 수 있으며, 영남권 관광 클러스터의 구축은 영남권 관광 역량 제고에 큰 기여를 할 것이라 생각됨

지역 관광을 활성화 하기 위한 관광 클러스터 구축

구체적 실현 방안

1. 외국인 관광객 대상 영남권 패키지 여행 유치



*<https://www.tourradar.com/d/south-korea>

- 패키지 여행을 판매하는 tourradar에서 우리나라에서 가장 인기있는 패키지 관광객의 동선은 다음과 같음
- 대부분 서울, 부산, 제주 등 유명 관광지 만을 포함
- 유명 관광지 외에도 새로운 지역을 포함 시키거나, 영남권의 공항으로 입국 후 영남권 집중 패키지 여행과 같이 클러스터 내에서 새로운 시도가 가능함

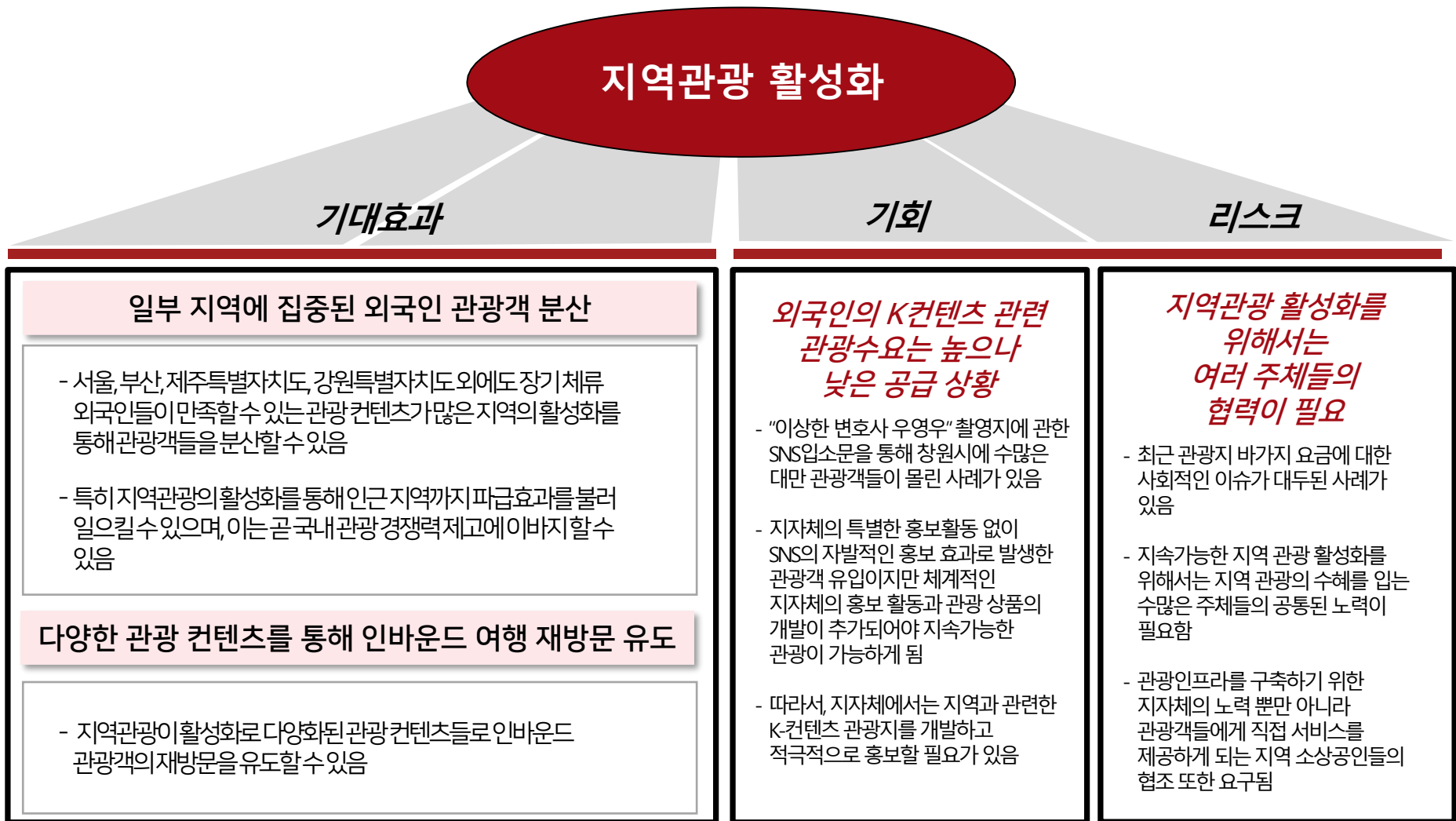
2. 외국인 관광객 대상 영남권 투어 패스 유치

- 영남권 관광 클러스터의 집적성장을 위해 영남권의 다양한 관광지를 저렴한 가격에 이용할 수 있는 투어패스를 구축해 영남권 관광 클러스터의 집적 성장을 도모할 수 있음

거제시, 포항시 외에도 영남
지역에는 **관광 잠재력이
높은 지역이 많기에**
영남권 관광 클러스터의
구축은 큰 집적 효과를
불러일으킬 것임

5. 시사점

지역 관광 활성화는 국내 관광의 양적인 성장과 질적인 성장에 모두 기여할 수 있으며, 지속가능한 지역관광 활성화를 위해서는 여러 주체들의 협력이 필요



End of Document