

# 비인기 지역의 장소 혹은 축제 활성화

혼자 열심히 하조

홍주오

# INDEX

01

- 분석 배경 및 목적

- 1) 분석배경
- 2) 분석 주제설정
- 3) 분석 목표설정

02

- 데이터 전처리 및 분석

- 1) 분석도출 과정
- 2) 데이터 전처리/분석

03

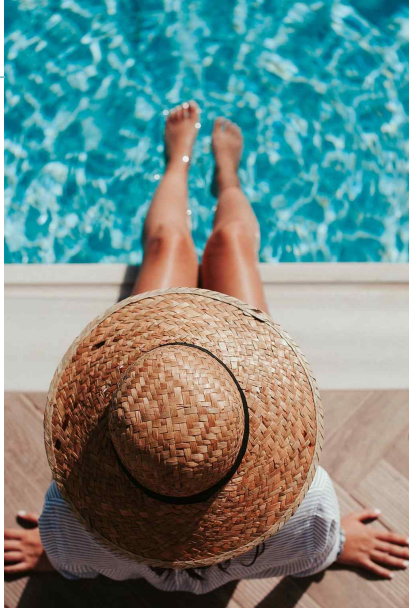
- 결론

- 1) 분석결과 및 아이디어

04

- 기대효과

- 1) 기대효과



**Part 1,**

**분석 배경 및 목적**

# 1) 분석배경- 국내여행

**학부모 71%,  
여름방학 계획 세웠다...  
1위는 국내 여행**

Y5 윤선생  
YOUNG ENGLISH SCHOOL

조사대상: 고등학교 이하 자녀를 둔 학부모 693명  
조사기간: 2023년 6월 16일~6월 27일

계획한 여름방학 활동 1~5위 \*복수응답



국내 여행  
61.5%



영화, 공연 관람  
42.4%



가족, 친지 방문  
34.6%



아웃도어 활동  
33.4%



박물관 등 관람  
28.3%

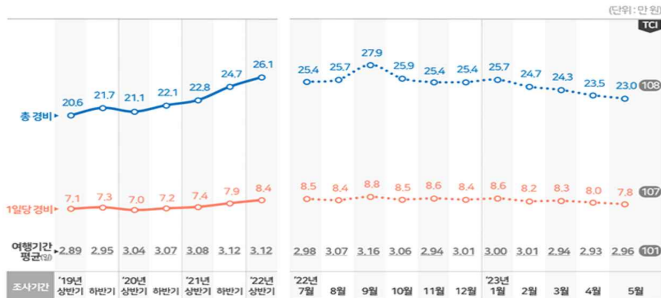
방학 계획을 세울 때 가장 중요한 점

**1위** 48.6% 여행, 체험학습 등으로 새로운 경험을 쌓는 것

**2위** 14.3% 자녀의 학습에 도움이 되는 것

**3위** 14.1% 평소보다 많은 대화를 통해 유대관계를 형성하는 것

[그림] 국내여행 기간 및 비용



\* 자녀가 있는 학부모들은 여름방학 계획 1위로 국내여행을 뽑혔다.

\* 국내로 여행을 자주 간다고 하지만 마스크 착용 의무가 풀렸는데도 불구하고  
고물가 시대인 만큼 여행기간도 비용도 줄어들면서 여행하는 추세이다.

# 1) 분석배경- 문화축제

“펜데믹 여파로 지친 일상...문화축제로 치유하세요”

전혜원 기자

기사승인 2023. 06. 29. 06:00

다녀오길 잘했다! 더위도 잊게 만든 '세계도시 문화축제'

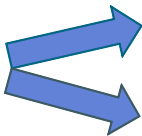
시민기자 박지영

발행일 2023.06.20. 11:32

수정일 2023.06.20. 11:32

조회 380

국내여행 뿐만아니라  
국내 문화축제에 대한 관심도 많아졌다.



국내여행 지역 베스트 5곳과 워스트 5곳을 파악한  
후 **베스트지역과 워스트 지역의 차이점을 분석**

국내 문화축제 베스트 5곳과  
워스트 5곳을 파악한후 **베스트와 워스트의 차이점 분석**

## 2) 분석 주제 설정

### (1) 국내 지역별 문화축제 검색량 데이터 파악

```
import pandas as pd
df1=pd.read_csv('./문화축제 검색량 데이터.csv')
df1.head()
```

	일련번호	검색량년월	출처ID	대표키워드명	검색어명	검색량값	출처분류명
0	31848	Mar-23	997	2014 안양시민축제	안양시민축제	170	문화축제
1	31849	Mar-23	992	2015 안양 시민축제	안양시민축제	170	문화축제
2	31850	Mar-23	396	2016 대가야체험축제	대가야체험축제	530	문화축제
3	31851	Mar-23	945	2016 대구약령시한방문화축제	대구약령시한방문화축제	190	문화축제
4	31852	Mar-23	994	2016 안양 시민축제	안양시민축제	170	문화축제

## 2) 분석 주제 설정

### (2) 국내지역별 관광명소 데이터 파악

```
import pandas as pd
import matplotlib.pyplot as plt
data1=pd.read_csv('./국내지역별 관광명소 데이터.csv')
data1.head()
```

	ID	대분류명	중분류명	POI_ID	POI명	지점명	보조명	분류코드	분류명	지번고유코드	...	도로명주소코드	도로명주소명	건물번호	위치경도	위치위도	격자코드	최종변경일자	출처명	파일명	기준일자
0	KCLANPO22N000000001	장소	관광지	817379	운문포	NaN	NaN	110208	대형토수/저수지	NaN	...	NaN	NaN	NaN	128.932090	35.729306	마마295490	20220809	KT	KC_495_LLR_ATTRCTN_2022	20220809
1	KCLANPO22N000000002	장소	관광지	5497426	유마사	NaN	NaN	180301	유명사찰	4.680000e+18	...	4.680000e+11	유마로	603	127.171274	35.018297	다라700693	20220809	KT	KC_495_LLR_ATTRCTN_2022	20220809
2	KCLANPO22N000000003	장소	관광지	5497607	연정사	NaN	NaN	180301	유명사찰	4.820000e+18	...	4.820000e+11	안정1길	363	128.395593	34.943969	라라817613	20220809	KT	KC_495_LLR_ATTRCTN_2022	20220809
3	KCLANPO22N000000004	장소	관광지	5545868	염불암	NaN	NaN	180301	유명사찰	4.120000e+18	...	4.120000e+11	예술공원로245번길	150	126.934391	37.428729	다사499367	20220809	KT	KC_495_LLR_ATTRCTN_2022	20220809
4	KCLANPO22N000000005	장소	관광지	332239	귀신사	NaN	NaN	180301	유명사찰	4.520000e+18	...	4.520000e+11	청도6길	40	127.046958	35.745312	다마590499	20220809	KT	KC_495_LLR_ATTRCTN_2022	20220809

5 rows × 27 columns



분석주제

관광명소 하위(위스트)5곳 과 문화축제 하위(위스트) 5곳 에 대한 해결방안

### 3) 분석 목표 설정



목표 1) 관광명소 베스트5 장소와  
문화축제 베스트5 장소에 공통점 파악



목표 2) 관광명소 워스트5 장소와  
문화축제 워스트5 장소에 대한 원인 탐색



목표 3) 관광명소 워스트5 장소와  
문화축제 워스트5 장소가 베스트  
되기 위해 필요한 방안 및 정책 제안



**Part 2,**

데이터 전처리 및 분석

# 1) 분석 도출 과정



첫 번째

국내지역별 관광명소 데이터와  
문화축제 검색량 데이터에서 각  
각 **베스트와 워스트 지역, 축제**  
파악



네 번째

**베스트 지역에서 축제 유형**  
**워드클라우드, 그래프로**  
**시각화**



두 번째

국내 지역별 관광명소가  
**많은 지역과 적은 지역이 어**  
**디인지 파악**



세 번째

베스트 지역에서 **축제 유형 탐색**

## 2) 데이터 전처리- 문화축제

(1) 검색량년월과 출처분류명의 열을 삭제하였다.

```
df2=df1.drop(['검색량년월','출처분류명'],axis=1) # 검색량년월이 3월 23일로 동일하여 데이터 삭제하였다. 또한 출처분류명도 문화축제로 동일하므로 삭제하였다.
df2.head()
```

	일련번호	출처 ID	대표키워드명	검색어명	검색량값
0	31848	997	2014 안양시민축제	안양시민축제	170
1	31849	992	2015 안양 시민축제	안양시민축제	170
2	31850	396	2016 대가야체험축제	대가야체험축제	530
3	31851	945	2016 대구약령시한방문화축제	대구약령시한방문화축제	190
4	31852	994	2016 안양 시민축제	안양시민축제	170



## 2) 데이터 분석- 문화축제

```
df4=df2["검색어명"].value_counts() # 검색어명에서 가진 단어들의 개수
df4
```

```
대구약령시한방문화축제    10
안양시민축제              5
구리코스모스축제          5
향촌동수제화골목          5
낙동강세계평화문화대축제    5
```

```
반딧불축제                1
바다로세계로              1
밀양아리랑대축제          1
민동산악새꽃축제          1
흥부제                    1
Name: 검색어명, Length: 529, dtype: int64
```

(2) 대표키워드명 기준으로 groupby 함수  
를 사용하여 검색량값의 평균값 출력



(1) “검색어명”의 중복된개수를 파악

```
df2.groupby("대표키워드명").mean()
```

```
<ipython-input-27-7c7fcf90e397>:1: FutureWarning: The default
df2.groupby("대표키워드명").mean()
```

	일련번호	출처ID	검색량값
대표키워드명			
2014 안양시민축제	31848.0	997.0	170.0
2015 안양 시민축제	31849.0	992.0	170.0
2016 대가야체험축제	31850.0	396.0	530.0
2016 대구약령시한방문화축제	31851.0	945.0	190.0
2016 안양 시민축제	31852.0	994.0	170.0
...	...	...	...
횡성한우축제	32445.0	1045.0	1880.0
효문화뿌리축제	32446.0	585.0	15.0
흑산도홍어축제	32447.0	155.0	47.0
흑석산 철쭉제	32448.0	467.0	80.0
흥부제	32449.0	320.0	80.0

602 rows x 3 columns



## 2) 데이터 분석- 문화축제

```
df4=df2.sort_values("검색량값", ascending=False) # 검색량값 기준으로 내림차순 정렬
df4
```

	일련번호	출처ID	대표키워드명	검색어명	검색량값
517	32364	654	진해군항제	진해군항제	951100
123	31971	1124	광양매화축제	광양매화축제	673600
159	32007	186	논산딸기축제	논산딸기축제	421400
350	32198	757	원동매화축제	원동매화축제	131900
365	32213	407	이천백사산수유꽃축제	이천백사산수유꽃축제	55310
...	...	...	...	...	...
545	32392	215	태안국제모래조각 페스티벌	태안국제모래조각페스티벌	10
543	32390	1072	크리스마스트리문화축제	크리스마스트리문화축제	10
110	31957	617	고싸움놀이축제(2차)	고싸움놀이축제	10
278	32125	1111	설화예술제	설화예술제	10
210	32058	709	방아다리벚꽃축제	방아다리벚꽃축제	10

602 rows × 5 columns

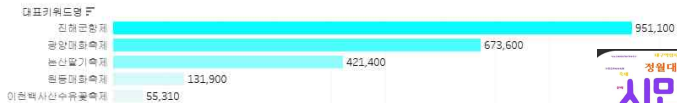
→ 검색량값을 기준으로  
상위 5개와 하위 5개 추출



(1) 검색량이 가장 많은 축제의 지역은  
순서대로 **경남, 전남, 충남, 경남, 경기도** 순이다.

(2) 검색량이 가장 적은 축제의 지역은  
순서대로 **서울, 충남, 광주, 경남, 충남** 순이다.

<대표키워드 검색량이 높은순서>



<대표키워드 검색량이 낮은순서>



대표키워드 검색량 상위, 하위 순서 파악



## 2) 데이터 전처리- 국내지역별 관광명소

(1) 분석에 필요없는 변수들을 삭제하였다.

```
data2=data1.drop(['대분류명','중분류명','지점명','보조명','분류코드','지번고유코드','번지번호','법정동코드','행정동코드','도로명주소코드',
                  '위치경도','위치위도','격자코드','최종변경일자','출처명','파일명','기준일자'],axis=1)
data2.head()
```

	ID	POI_ID	POI명	분류명	시도명	시군구명	법정동명	리명	도로명주소명	건물번호
0	KCLANPO22N0000000001	817379	운문호	대형호수/저수지	경상북도	청도군	운문면	대천리	NaN	NaN
1	KCLANPO22N0000000002	5497426	유마사	유명사찰	전라남도	화순군	사평면	유마리	유마로	603
2	KCLANPO22N0000000003	5497607	안정사	유명사찰	경상남도	통영시	광도면	안정리	안정1길	363
3	KCLANPO22N0000000004	5545868	염불암	유명사찰	경기도	안양시 만안구	석수동	NaN	예술공원로245번길	150
4	KCLANPO22N0000000005	332239	귀신사	유명사찰	전라북도	김제시	금산면	청도리	청도6길	40

## 2) 데이터 분석- 국내지역별 관광명소

```
data4=data2["분류명"].value_counts() # 분류명의 각 지역의 개수
data4 # 각 관광명소의 열 개수 출력
```

```
지역호수/저수지      12658
지역사찰             11209
일반관광지           2573
지역축제             2549
토속/특산물/기념품매장 1988
먹거리/패션거리      1401
폭포/계곡            955
관광농원/허브마을     931
서원/향교/서당       858
휴양림/수목원        557
해수욕장             549
일반유원지/일반놀이공원 445
유명사찰             395
유명관광지           310
테마공원/대형놀이공원 268
식물원              146
대형호수/저수지      47
동물원              44
숲                  39
아쿠아리움/대형수족관 29
카지노              9
파라다이스카지노     4
세븐럭카지노         3
Name: 분류명, dtype: int64
```



(1) 각 관광명소의 개수 출력

```
data7 = data2["시도명"].value_counts().sort_values(ascending=False)
data7 # 각 시도명의 열 개수 출력
```

```
경상남도      5657
경상북도      5163
전라남도      4798
경기도        3899
전라북도      3109
강원도        2658
충청남도      2381
서울특별시    2159
충청북도      2083
부산광역시    1397
제주특별자치도 1155
대구광역시    995
울산광역시    828
인천광역시    611
광주광역시    482
대전광역시    434
세종특별자치시 158
Name: 시도명, dtype: int64
```

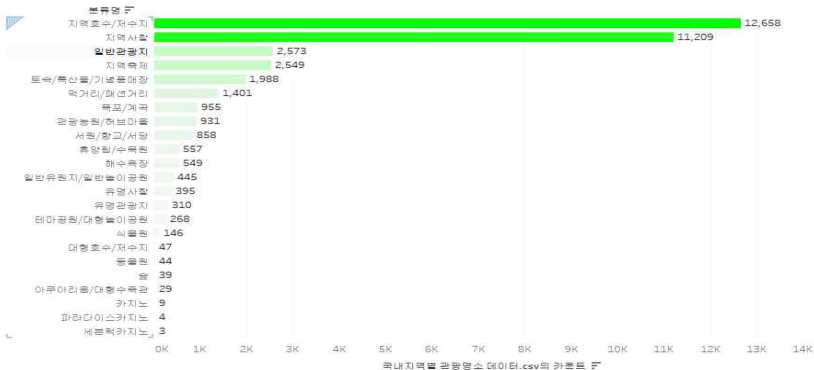
(2) 각 관광명소의 시도명 출력





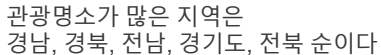
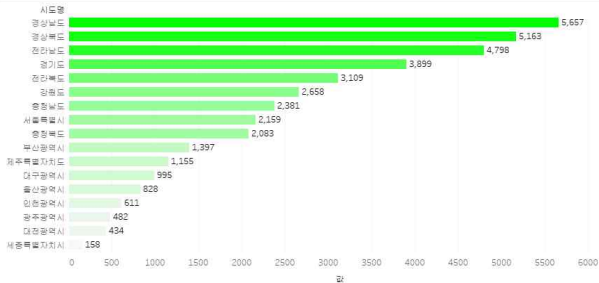
## 2) 데이터 분석- 문화축제

<국내 관광명소가 많은순>



▶ 관광명소가 많은 순서는 호수, 지역사찰, 일반관광지, 지역축제 순이다

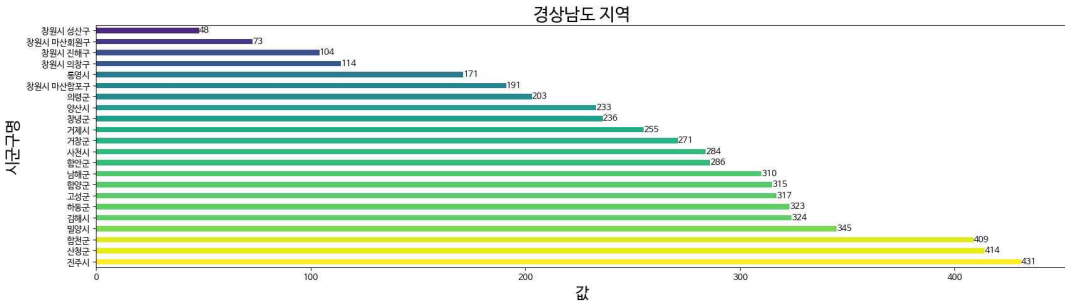
<국내지역 중 관광명소가 많은 순>



적은 지역은 세종, 대전, 광주, 인천, 울산순이다



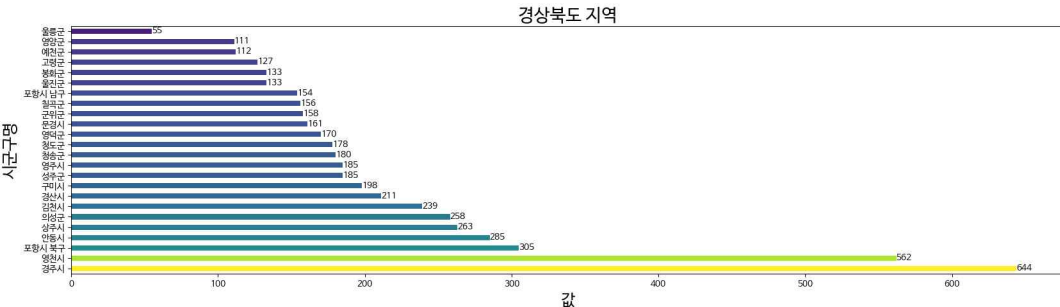
## 2) 데이터 분석- 문화축제



관광명소가 많은 경남지역의 시군구명이다.

경남의 관광명소가 많은 시군구명은 전주시, 산청군, 함천군, 밀양시 순이다.

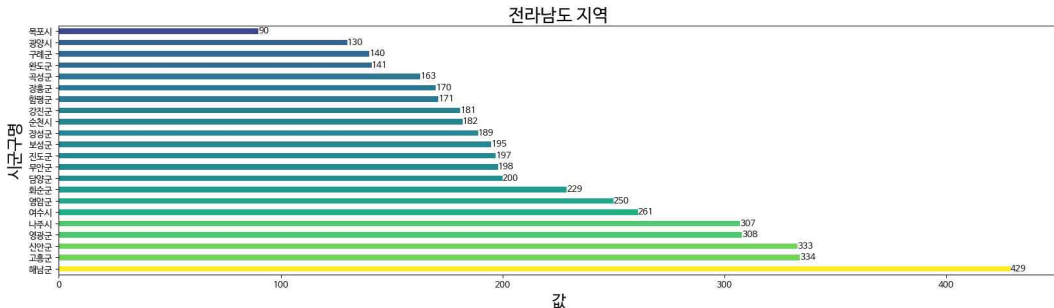
## 2) 데이터 분석- 관광명소



관광명소가 많은 경북지역의 시군구명이다.

경북의 관광명소가 많은 지역은 경주시, 영천시, 북구 순이다

## 2) 데이터 분석- 문화축제

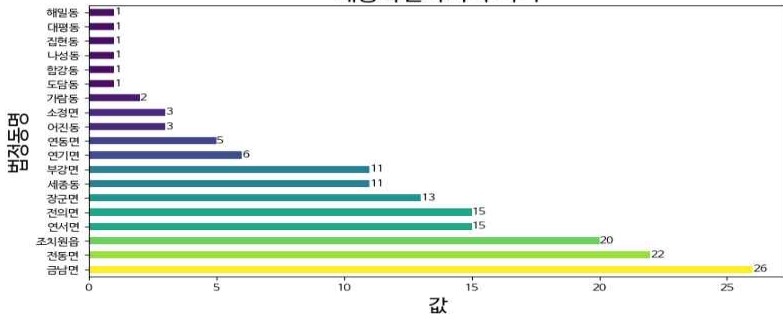


전남지역의 관광명소가 많은 시군구명이다.

전남의 관광명소가 많은 시군구명은 해남군, 고흥군, 신안군, 영광군 순이다.

## 2) 데이터 분석- 문화축제

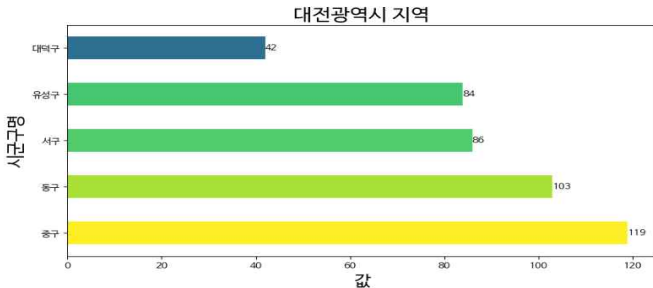
세종특별자치시 지역



관광명소가 적은 세종시의 동명이름이다.

세종시중에서도 관광명소가 많은 지역은 금남면, 전동면, 조치원읍 순이다.

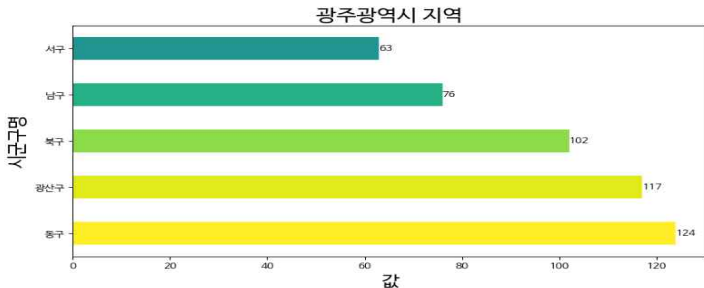
## 2) 데이터 분석- 문화축제



관광명소가 적은 대전광역시의 시군구명이다.

대전광역시중에서 관광명소가 많은 지역은 중구, 동구, 서구 순이다.

## 2) 데이터 분석- 문화축제



관광명소가 적은 광주광역시의 시군구명이다.

광주광역시 중에서 관광명소가 많은 지역은 동구, 광산구, 북구 순이다.



**Part 3,**

**결론**

# 1) 분석결과 및 개선아이디어

## - 베스트 5장소

(1) 관광명소 베스트 5장소의 지역 : 경남, 경북, 전남, 경기도, 전북

(2) 문화축제 베스트 5장소의 지역 : 경남, 전남, 충남, 경기도

(3) 이 장소들의 공통점 : 뚜렷한 축제 혹은 이목을 끌만한 장소

## - 위스크 5장소

(1) 관광명소 위스크 5장소 지역 : 세종, 대전, 광주, 인천, 울산

(2) 문화축제 위스크 5장소의 지역 : 서울, 충남, 광주, 경남

(3) 이 장소들의 대한 원인 탐색 : 관광명소 혹은 문화축제로 이목이 쏠림

## - 정책제안

관광명소 위스크와 문화축제 위스크 5장소가 베스트되기 위해 필요한 방안 및 정책 제안 :

위스크지역을 시민들에게 각 지역의 '시' 또는 '군' 의 소식을 카톡으로 전달 하거나 홍보

또한 비인기지역의 축제 혹은 장소에 오면 할인해주는 서비스 병행

개선 아이디어

베스트 장소  
와 위스트  
장소 파악

베스트 장소와  
위스트 장소  
차이점 파악

위스트 장소  
을 알리기 위  
해 정부 지원

위스트 장소의  
각 지역 시민들  
에게 홍보

위스트 장소가  
베스트 장소로  
전환

**Part 4,**

기대효과

# 1) 기대효과

빅데이터를 활용한 국내 여행/소비 트렌드 리포트

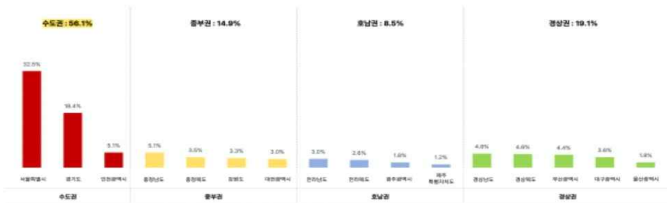


## 지역별 국내 여행객 비율

- 수도권이 여행의 중심으로 떠오름
- 중부권과 호남권은 거리두기 해제 이후에도 여행 수요를 흡수하지 못한 것으로 보임

권역별 국내 여행객 비율

(대한민국 지역 여행객 기준 Q2.1~Q2.4 누적 / %)



## - 기대효과(1)

맨 왼쪽부터 수도권, 중부권, 호남권, 경상권 이고 22년 1월부터 12월까지 조사한 그래프이다.

수도권의 국내 여행객 비율과 비슷하게 중부권, 호남권, 경상권의 비율이 올라갈것이다.

# 1) 기대효과

빅데이터를 활용한 국내 여행/소비 트렌드 리포트



## 국내 대표 여행지 숙박여행 증감률

- 22년 한 해 동안 국내 대표 여행지 전반에서 숙박여행이 증가
- 그중에서도 특히 22년에는 부산과 서울의 숙박여행 증가가 두드러지게 나타남

(전년 대비 여행지 방문 (22. 1-12 누적 숙박여행) / 2022)

서울

숙박여행 증감률 : ▲95.7% (2위)

광주

숙박여행 증감률 : ▲84.0% (3위)



강원

숙박여행 증감률 : ▲17.6% (이하위)

부산

숙박여행 증감률 : ▲98.2% (1위)

제주

숙박여행 증감률 : ▲20.4% (15위)

## - 기대효과(2)

정부지원을 통한 비인기 지역을 활성화 한다면, 국내여행과 함께 숙박여행도 증가할 것이다.

감사합니다.