

한국 관광 트렌드

장수경 김범모 최정안 유수빈



목차



문제정의

Problem definition

저장

Data storage

결과 도출

Result presentation



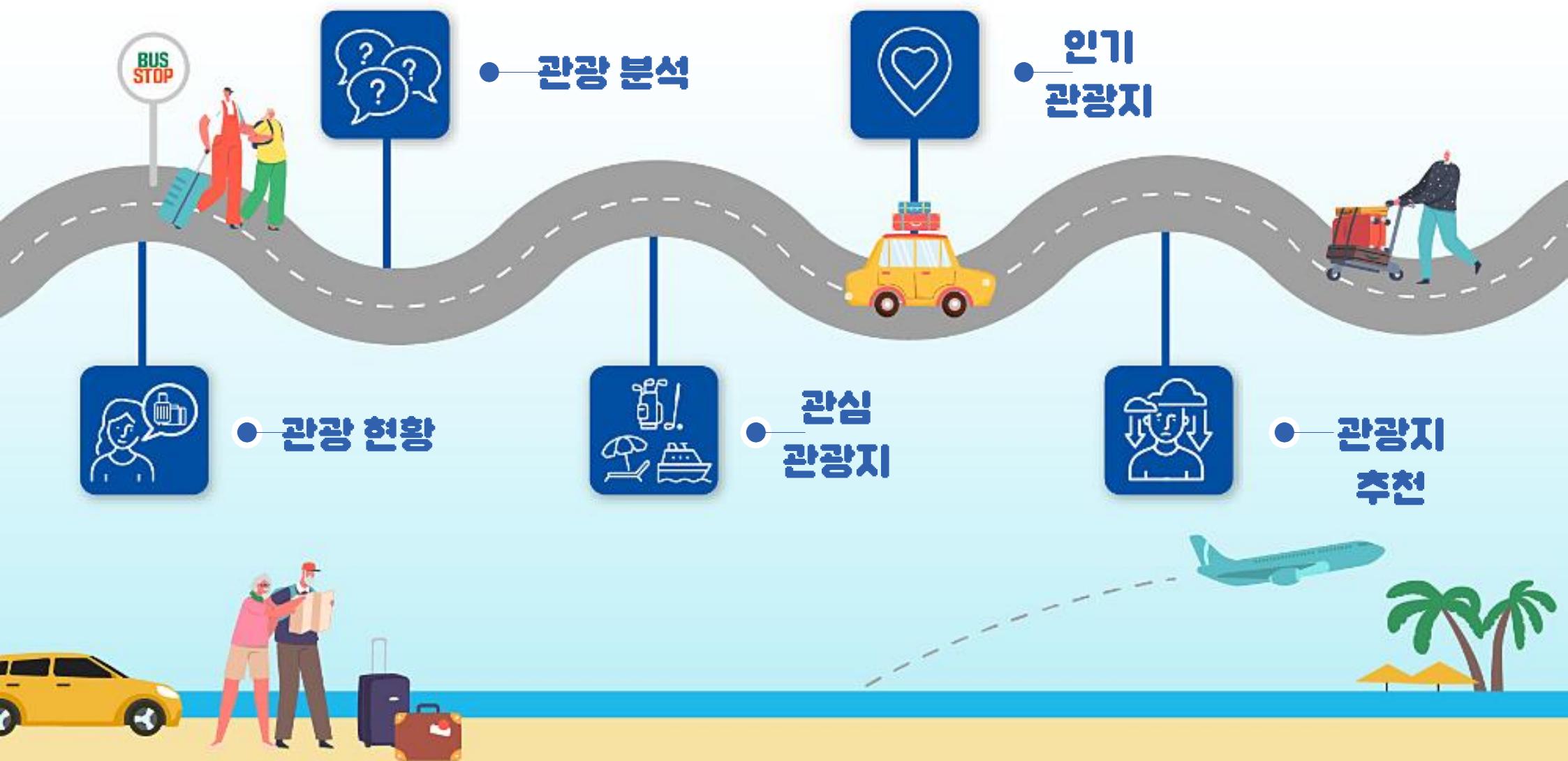
수집 및 요약

Data collection

처리

Data processing

문제 정의



데이터 수집



방문자
특성



차량
이동



관광업종
소비



관심
키워드



데이터 수집

관심
키워드
데이터

설명	필드명	자료형
도	province	STRING
도시	city	STRING
언급량	sns_reference	BIGINT
친구	companion_friend	BIGINT
기타가족	companion_other_family	BIGINT
부모님	companion_parents	BIGINT
학생	companion_student	BIGINT
회사동료	companion_colleague	BIGINT
자녀	companion_children	BIGINT
연인	companion_couple	BIGINT
싱글	companion_single	BIGINT
배우자	companion_spouse	BIGINT
미식	travel_gourmet	BIGINT
자연경관	travel_sightseeing	BIGINT
휴식/힐링	travel_healing	BIGINT
문화/예술	travel_culture	BIGINT
역사유적	travel_history	BIGINT
체험	travel_experience	BIGINT
반려동물	travel_pet	BIGINT
레포츠/운동	travel_exercise	BIGINT

차량 이동
데이터

설명	필드명	자료형
도	Province	STRING
도시	City	STRING
관광지	Tourist_Destination	STRING
주소	Address	STRING
중분류	Category	STRING
소분류	SubCategory	STRING
검색수	Num_Of_Searches	BIGINT

방문자
특성
데이터

설명	필드명	자료형
도	province	STRING
도시	city	STRING
방문자수	Num_of_visitors	BIGINT
전년방문자수	yoy_num_of_visitors	BIGINT
방문자수증감율	percentage_visitors	FLOAT
무박	no_night	BIGINT
1박	one_night	BIGINT
2박	two_nights	BIGINT
3박	three_nights	BIGINT
평균숙박일수	avg_stay_day	BIGINT
숙박율	avg_stay_ratio	FLOAT
평균체류시간(분)	avg_detention_period_minute	INT
외국인방문자	foreigner_visitors	BIGINT

데이터 수집

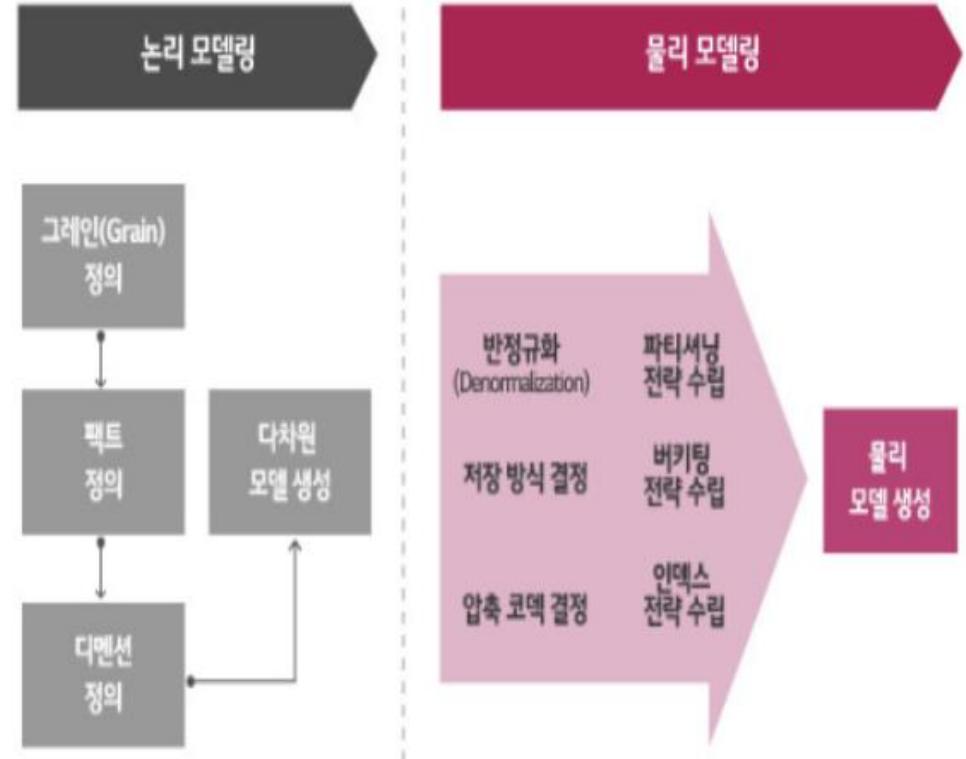
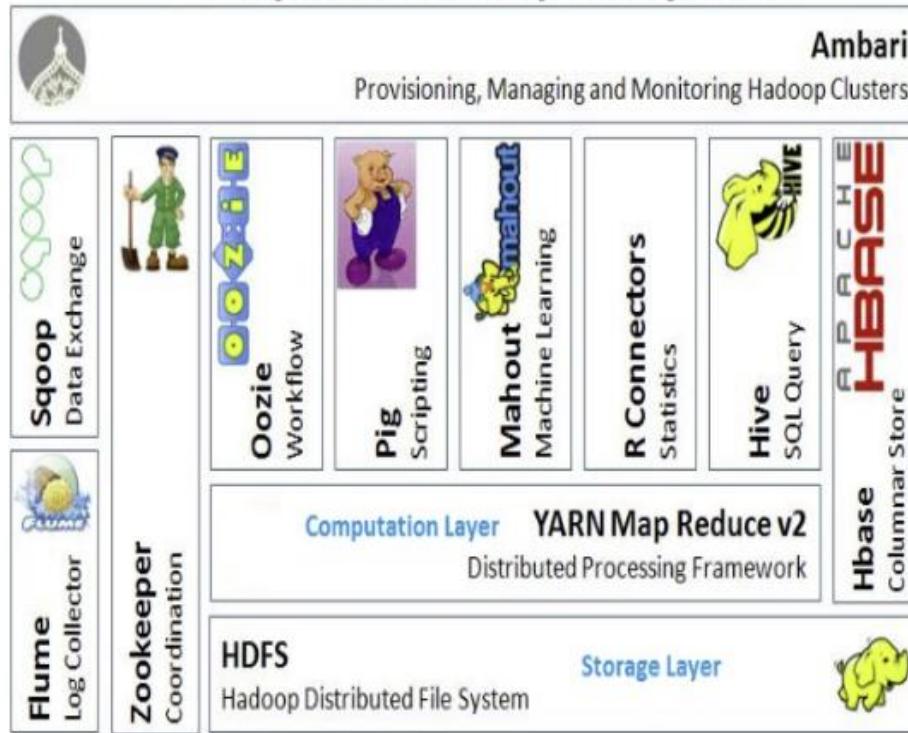
관광 업종
소비
데이터

설명	필드명	자료형
연도	year	INT
광역지자체	province	STRING
기초지자체	city	STRING
내국인광역지자체지출액	kor_pro_exp	BIGINT
내국인광역지자체지출액비율	kor_pro_exp_ratio	FLOAT
내국인기초지자체지출액	kor_cty_exp	BIGINT
내국인기초지자체지출액비율	kor_cty_exp_ratio	FLOAT
외국인광역지자체지출액	fg_pro_exp	BIGINT
외국인광역지자체지출액비율	fg_pro_exp_ratio	FLOAT
외국인기초지자체지출액	fg_cty_exp	BIGINT
외국인기초지자체지출액비율	fg_cty_exp_ratio	FLOAT

설명	필드명	자료형
연도	year	INT
대분류	category	STRING
중분류	subcategory	STRING
내국인대분류지출액	kor_cg_exp	BIGINT
내국인대분류지출액비율	kor_cg_exp_ratio	FLOAT
내국인중분류지출액	kor_scg_exp	BIGINT
내국인중분류지출액비율	kor_scg_exp_ratio	FLOAT
외국인대분류지출액	fg_cg_exp	BIGINT
외국인대분류지출액비율	fg_cg_exp_ratio	FLOAT
외국인중분류지출액	fg_scg_exp	BIGINT
외국인중분류지출액비율	fg_scg_exp_ratio	FLOAT

데이터 저장

Apache Hadoop Ecosystem



빅데이터 분석을
위한 환경 구성

데이터 웨어 하우스
모델링 및 적재

데이터 저장

자료형 결정

Year	INT
province	STRING
city	STRING
pro_search_cnt	BIGINT
pro_search_ratio	FLOAT
cty_search_cnt	BIGINT
cty_search_ratio	FLOAT

스키마 생성

```
CREATE TABLE if not exists search_location(
    year INT,
    province STRING,
    city STRING,
    pro_search_cnt BIGINT,
    pro_search_ratio
    FLOAT,
    cty_search_cnt BIGINT,
    cty_search_ratio
    FLOAT)
ROW FORMAT
DELIMITED
FIELDS TERMINATED BY ","
LINES TERMINATED BY
"\n"
STORED AS TEXTFILE;
```

작업

```
insert overwrite
directory
'/user/hive/y_pop_touri
st_destination.txt'
select k.year,
k.tourist_destination,
k.city, k.category,
k.num_of_searches
from (select year,
tourist_destination, city,
category,
max(num_of_searches)
as num_of_searches
from search_keyword
group by year,
tourist_destination,
.
.
```

데이터 처리

```
insert overwrite directory  
'/user/hive/output.txt'  
select s.year, s.province, s.city,  
s.pro_search_cnt, s.cty_search_cnt,  
sub.pro_visitants, sub.cty_visitants,  
sub.kor_pro_exp, sub.kor_cty_exp, sub.fg_cg_exp,  
sub.fg_cty_exp  
from search_location s FULL OUTER JOIN  
    (select v.year, v.province, v.city,  
v.pro_visitants, v.cty_visitants, e.kor_pro_exp,  
e.kor_cty_exp, e.fg_cg_exp, e.fg_cty_exp  
     from visitant_KT v FULL OUTER JOIN  
expenditure e  
      ON v.year = e.year and a.city = e.city) sub  
ON s.year = sub.year and a.city = sub.city
```

>> search_location, visitant_KT, expenditure 3가지 테이블

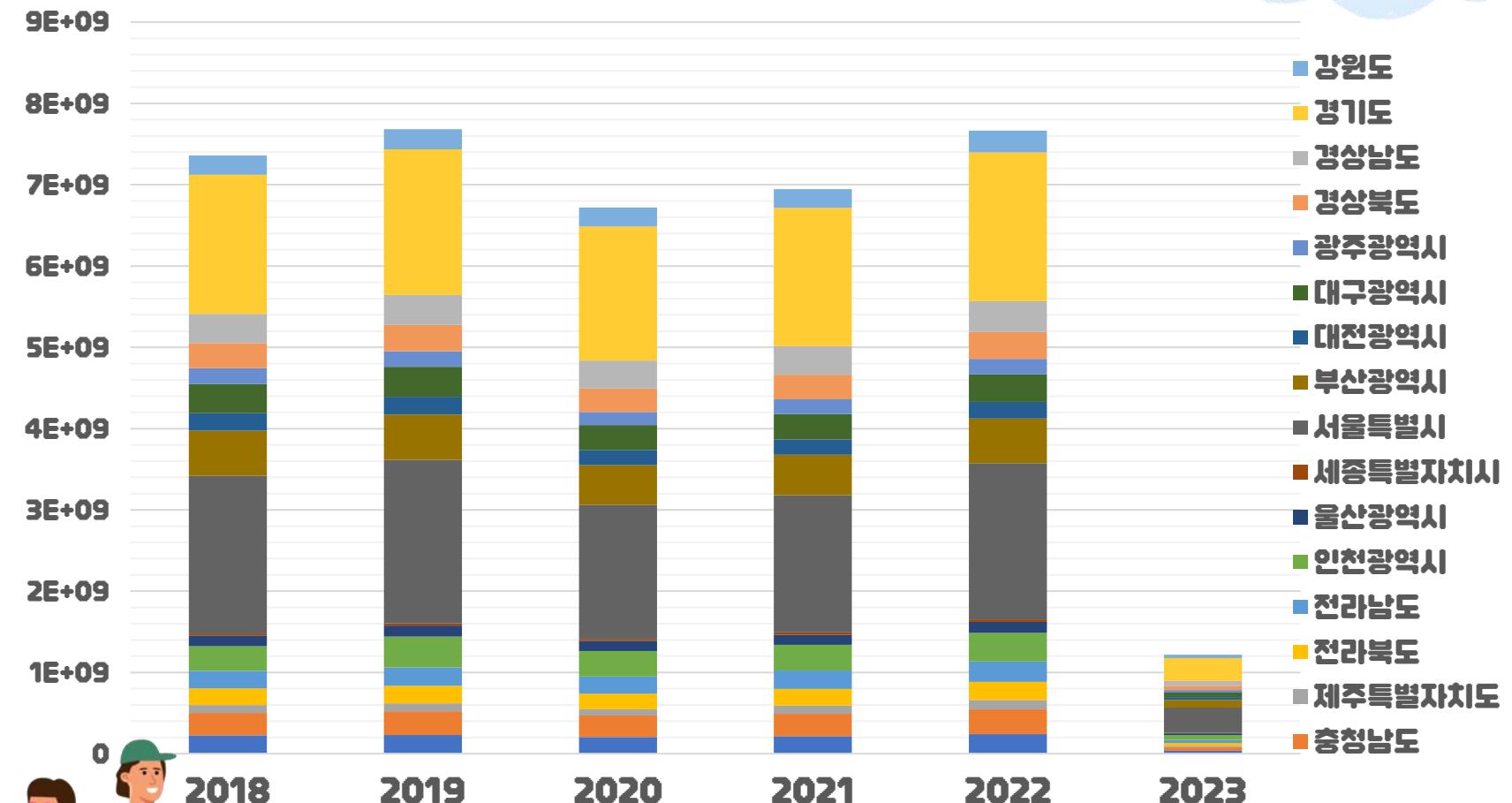
```
insert overwrite directory  
'/user/hive/output.txt'  
select s.year, s.province, s.city,  
s.pro_search_cnt, s.cty_search_cnt,  
sub.pro_visitants, sub.cty_visitants,  
sub.kor_pro_exp, sub.kor_cty_exp, sub.fg_cg_exp,  
sub.fg_cty_exp  
from search_location s FULL OUTER JOIN  
    (select v.year, v.province, v.city,  
v.pro_visitants, v.cty_visitants, e.kor_pro_exp,  
e.kor_cty_exp, e.fg_cg_exp, e.fg_cty_exp  
     from visitant_KT v FULL OUTER JOIN  
expenditure e  
      ON v.year = e.year and a.city = e.city) sub  
ON s.year = sub.year and a.city = sub.city
```

관심
관광지
데이터

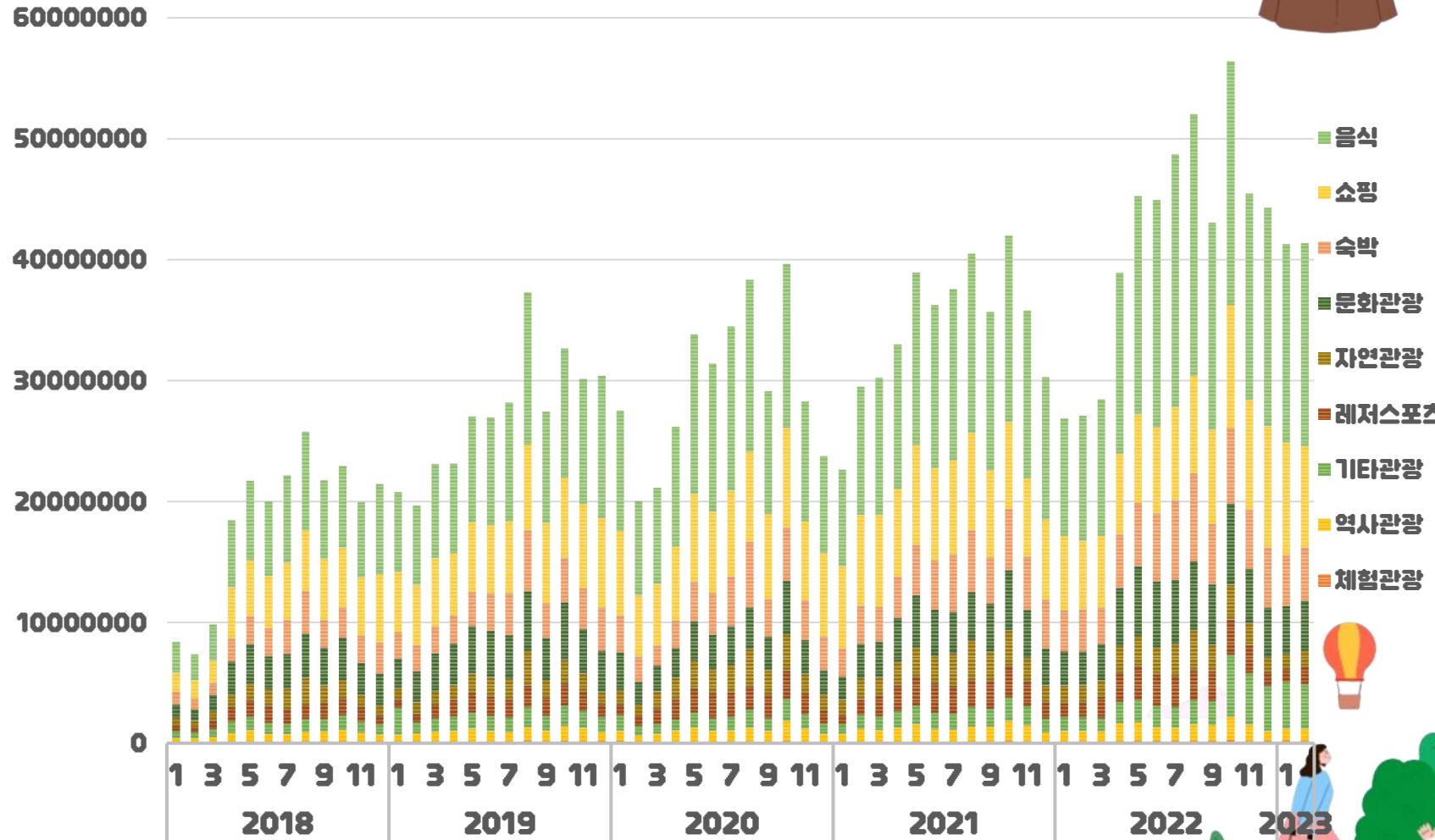


관광 현황

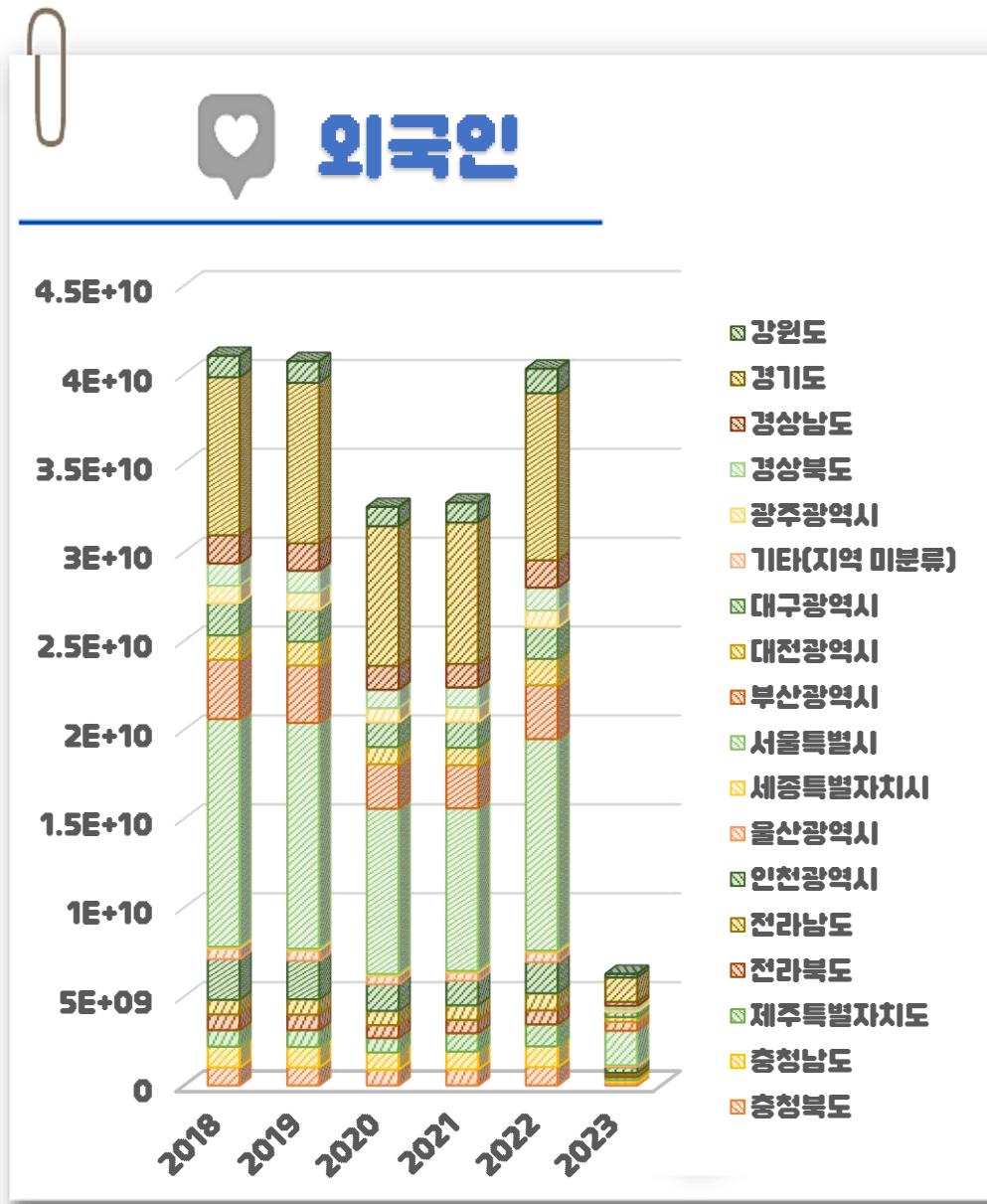
연도별 지자체 방문자 수



관광 현황



관광 현황



연도별 광역 지자체 지출액

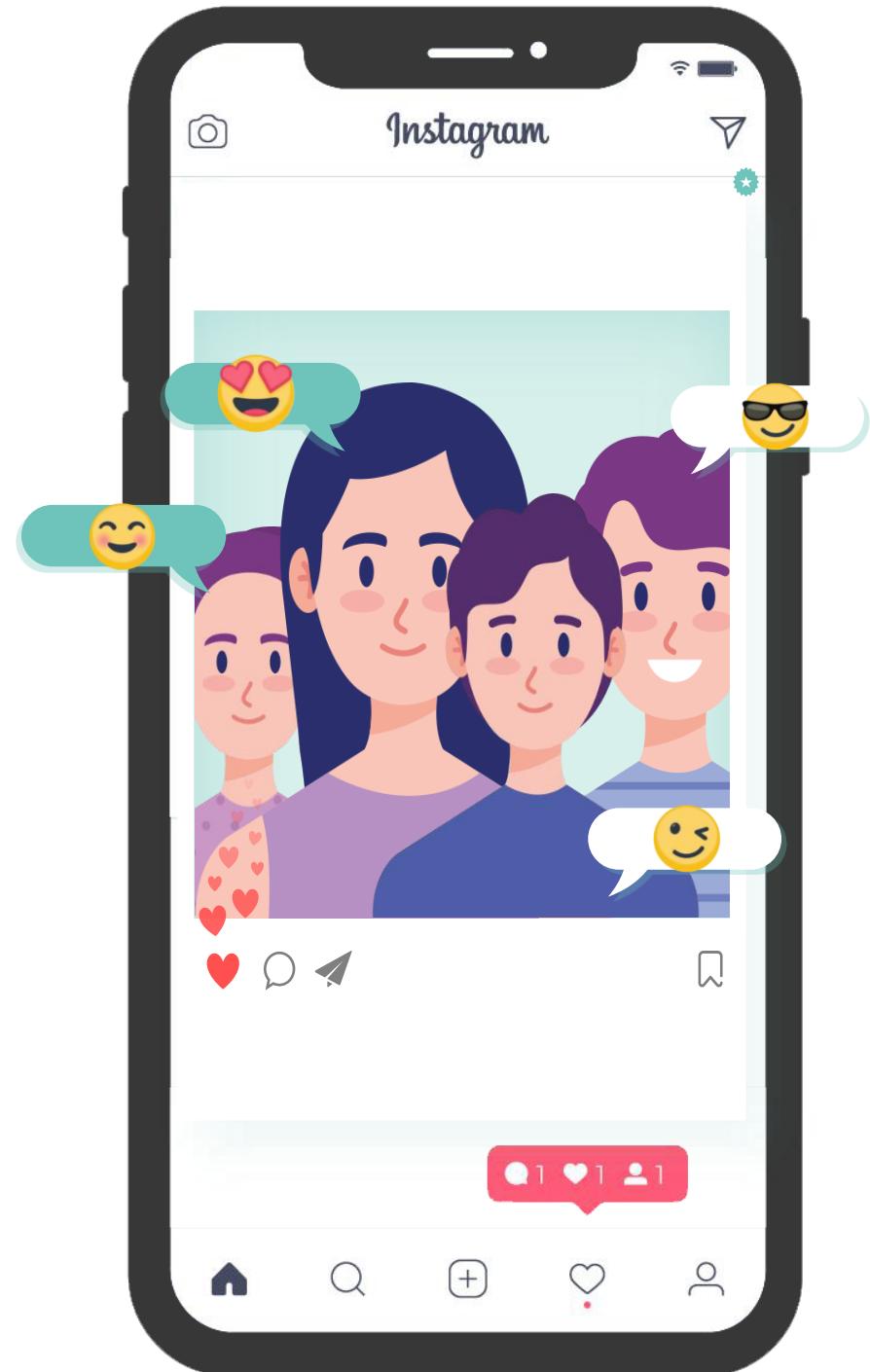


관심 관광지

SNS 언급량
상위 10개 도시

- 1위 서울
- 2위 부산
- 3위 수원
- 4위 강릉
- 5위 공주
- 6위 전주
- 7위 청주
- 8위 경주
- 9위 대구
- 10위 인천

“서울이 모든 부분에서 상위권이네!”



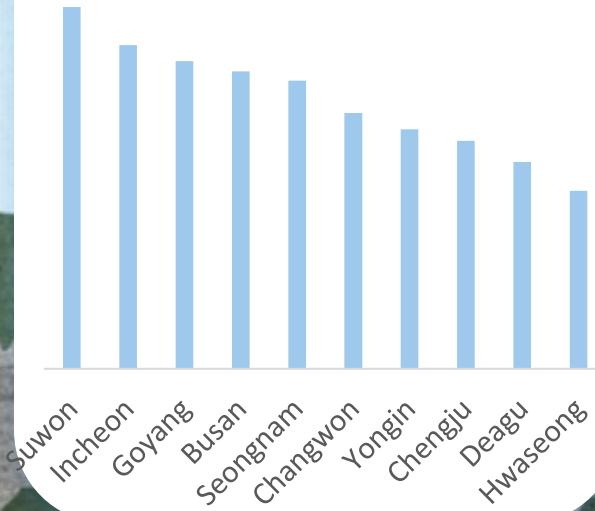
관심 관광지



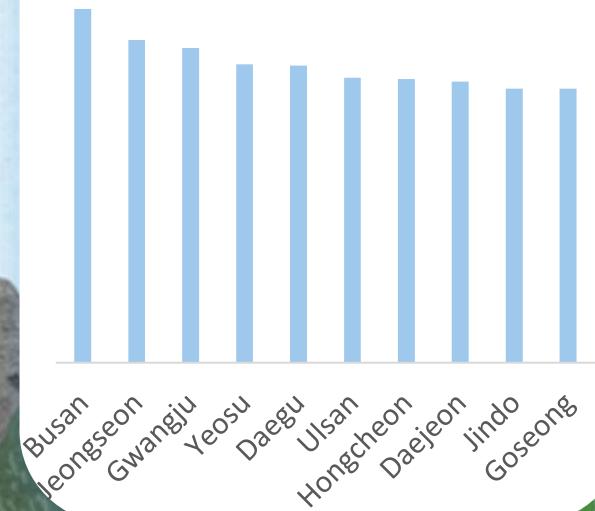
인기 관광지



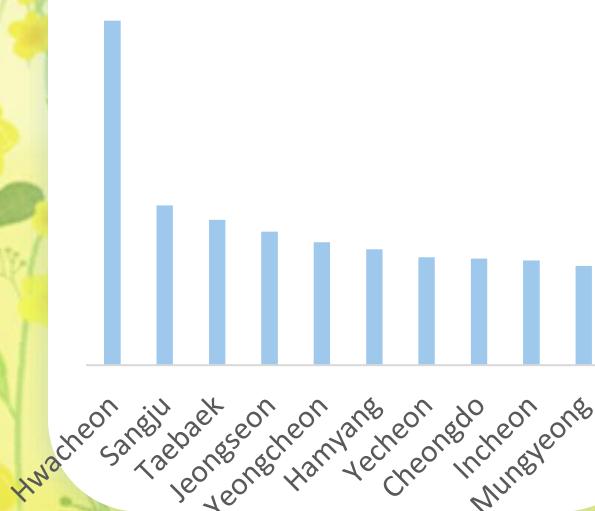
방문자가 많은 상위 10개 도시



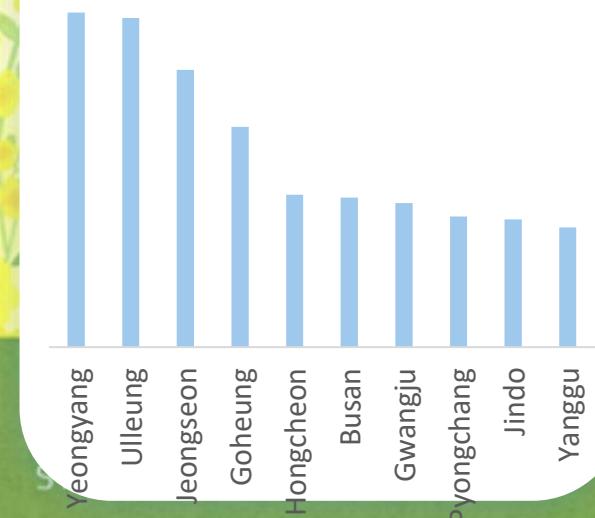
숙박률이 높은 상위 10개 도시



전년대비 방문자가 많이 늘어난 상위 10개 도시



머문 시간이 긴 상위 10개 도시



인기 관광지



배우자

충청남도 공주

Gongju-si, Chungcheongnam Province



친구

부산 광역시

Busan Metropolitan City



5·18 자유공원



연인

광주 광역시

Gwangju Metropolitan City



가족

전라북도 전주

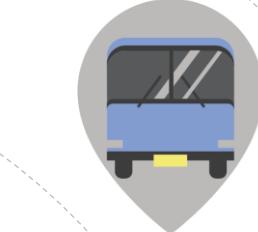
Jeonju-si, Jeollabuk Province



싱글

서울 특별시

Seoul Metropolitan City



회사동료

인천 광역시

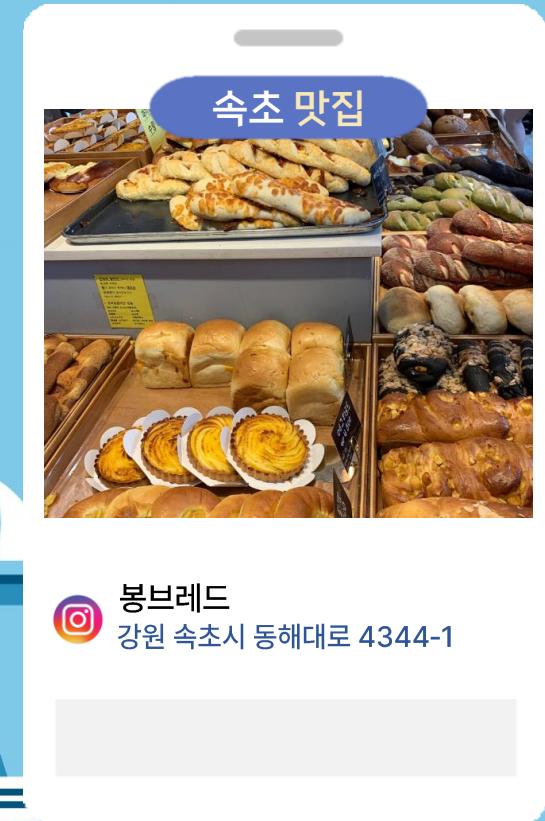
Incheon Metropolitan City



관심 관광지

동반유형

관광지 추천



데이터 출처

데이터



KOSIS
국가통계포털



소셜 미디어



사용 툴



