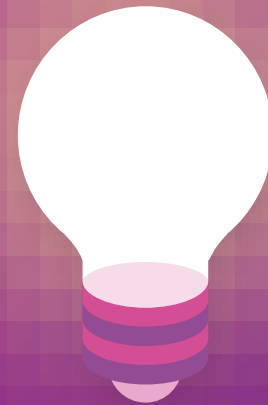


# Travel Trends

on Instagram

2021.07. 30

박종헌



# Contents



## Part 01

Background



## Part 02

Data Collection



## Part 03

Data Analysis  
&  
Recommender System



## Part 04

Dashboard Development

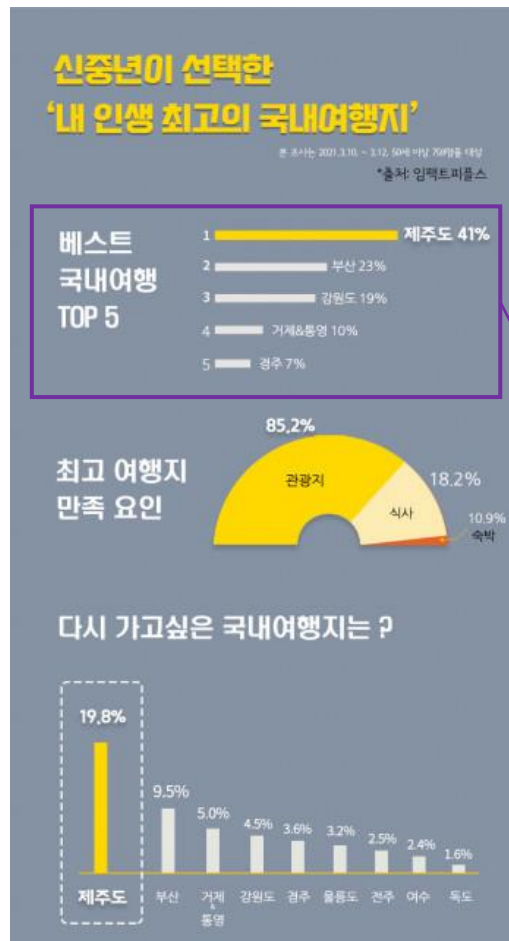
# Part 01

## Background

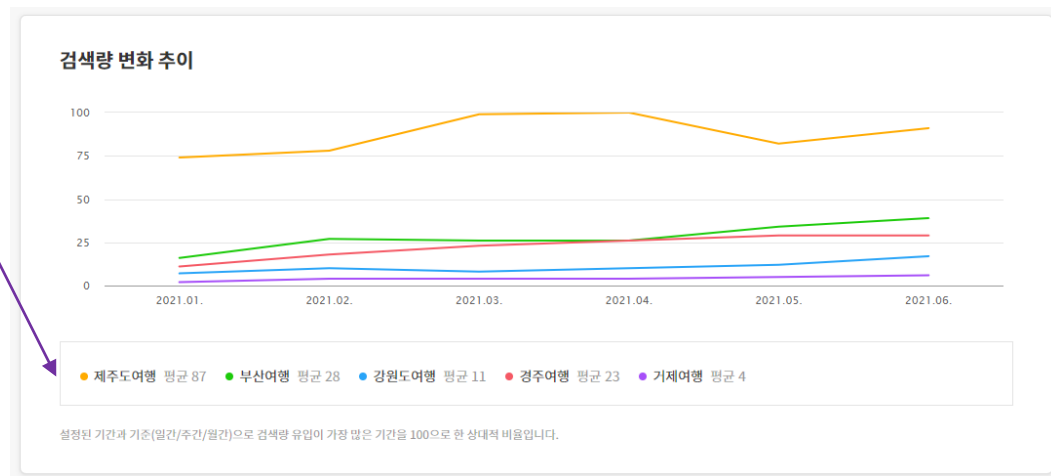


# Background

## 제주도 여행



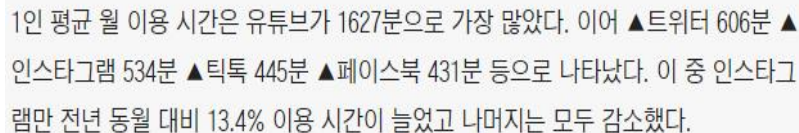
## Kakao datatrend 검색량 변화 추이





## 한국, 세계에서 두번째로 SNS 많이 한다

전 세계 국가별 소셜 미디어 이용률



연령별 가장 많이 이용하는 소셜 미디어 Top3

Note: 월 평균 이용자 수는 2021년 1분기(1~3월) 내 월별로 발생한 모바일(Android App+iOS App) 이용자 수의 산술평균값  
Source: 님스 코리아플릭(2021.04), 2021년 1분기 포털&SNS 보고서, DMC미디어 제공

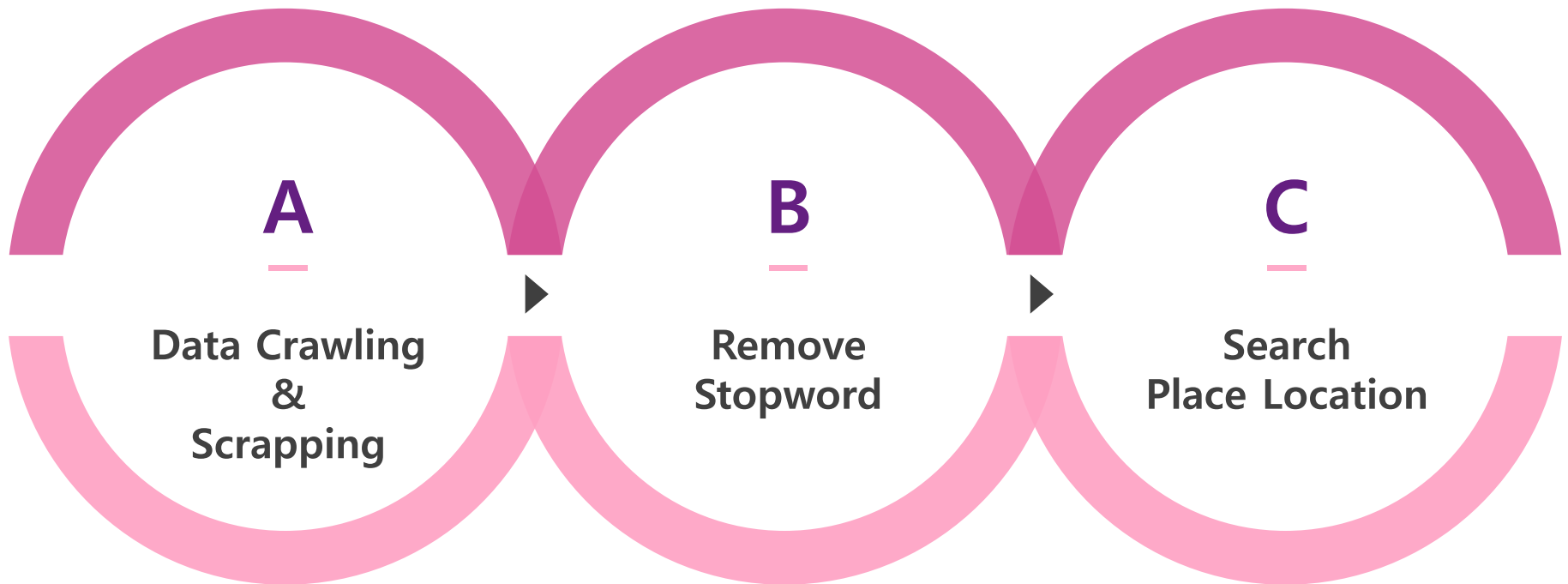
# Part 02

## Data Collection





# Data Collection

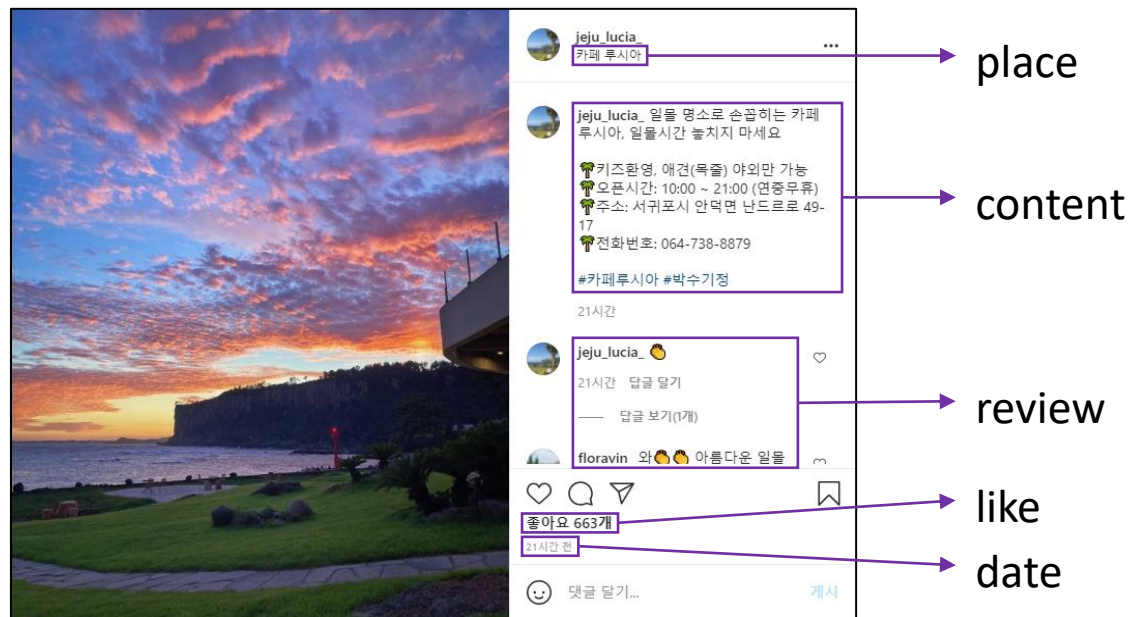


# Data Collection

## Data Crawling & Scrapping – Selenium

자연어 데이터 활용

- 비교적 많은 NLP 분야의 경험
- 적은 메모리 비용





# Data Collection

## Remove Stopword

- 자주 사용되는 해시태그 제거
- 제주도 관련 해시태그 제거



song\_hoon\_park  
송훈파크

...

#렛츠고우투더송훈팍 🌿

✓크라운돼지 12:00~21:00

✓하이드브레드 10:30~20:00

#송훈세프#송훈파크#제주도카페 #제주베이커리카페#애월카페#베이커리  
맛집#베이커리카페#빵#빵지순례#하  
이드브레드#금오름#제주도여행#커피  
#커피맛집#새별오름#새별오름카페#  
제주흑돼지#애월카페#송훈랜드#제주  
흑돼지맛집#제주도여행#제주도맛집#

### 인스타그램 한글 해시태그 랭킹

인스타그램의 게시된 해시태그별 게시물 수를 표시합니다.  
해시태그를 클릭하면 연관 해시태그가 표시됩니다.)  
더 많은 기능을 제공해드릴 수 있도록 노력하겠습니다. #코코태그 #COCOTAG #한글해시태그

<< 이전 100개

1위부터 100위까지

다음 100개 >>

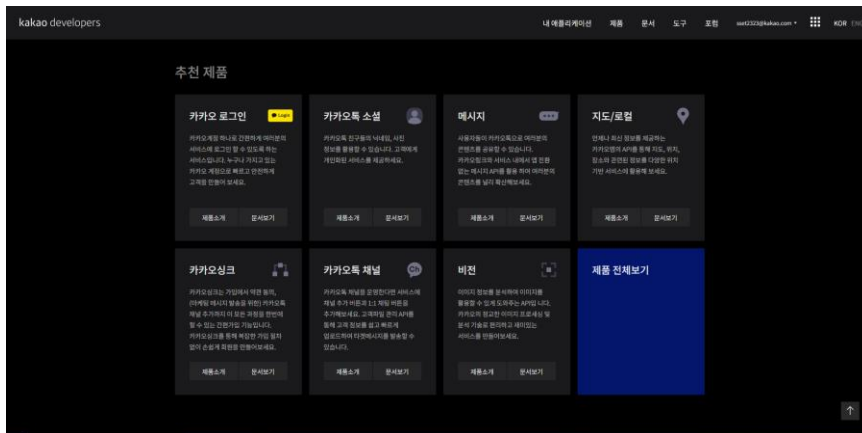
| #순위 | #해시태그  | #피드수   |
|-----|--------|--------|
| 1   | #일상    | 21,298 |
| 2   | #데일리   | 12,678 |
| 3   | #맛집    | 12,598 |
| 4   | #소통    | 12,48  |
| 5   | #좋아요   | 9,368  |
| 6   | #선물    | 8,438  |
| 7   | #셀카    | 8,288  |
| 8   | #인스타그램 | 8,148  |
| 9   | #먹스타그램 | 7,398  |
| 10  | #팔로우   | 6,498  |
| 11  | #셀피    | 6,268  |
| 12  | #데일리룩  | 5,158  |
| 13  | #여행    | 4,858  |
| 14  | #인스타그램 | 4,398  |
| 15  | #맛스타그램 | 4,138  |
| 16  | #인친    | 4,078  |
| 17  | #웃웃티디  | 3,888  |

<http://cocotag.kr/index.html>

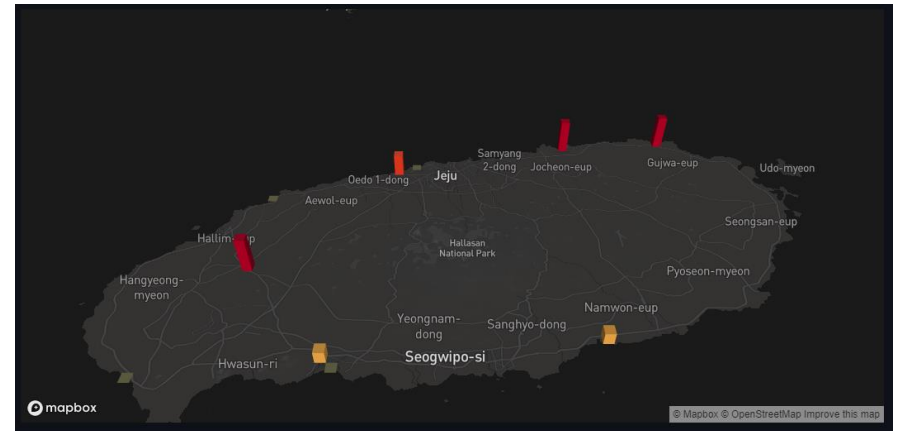
# Data Collection

## Search Place location

- 대시보드에 각 장소를 표시하기 위해 카카오맵 API를 활용하여 위도, 경도 검색



<https://developers.kakao.com/>



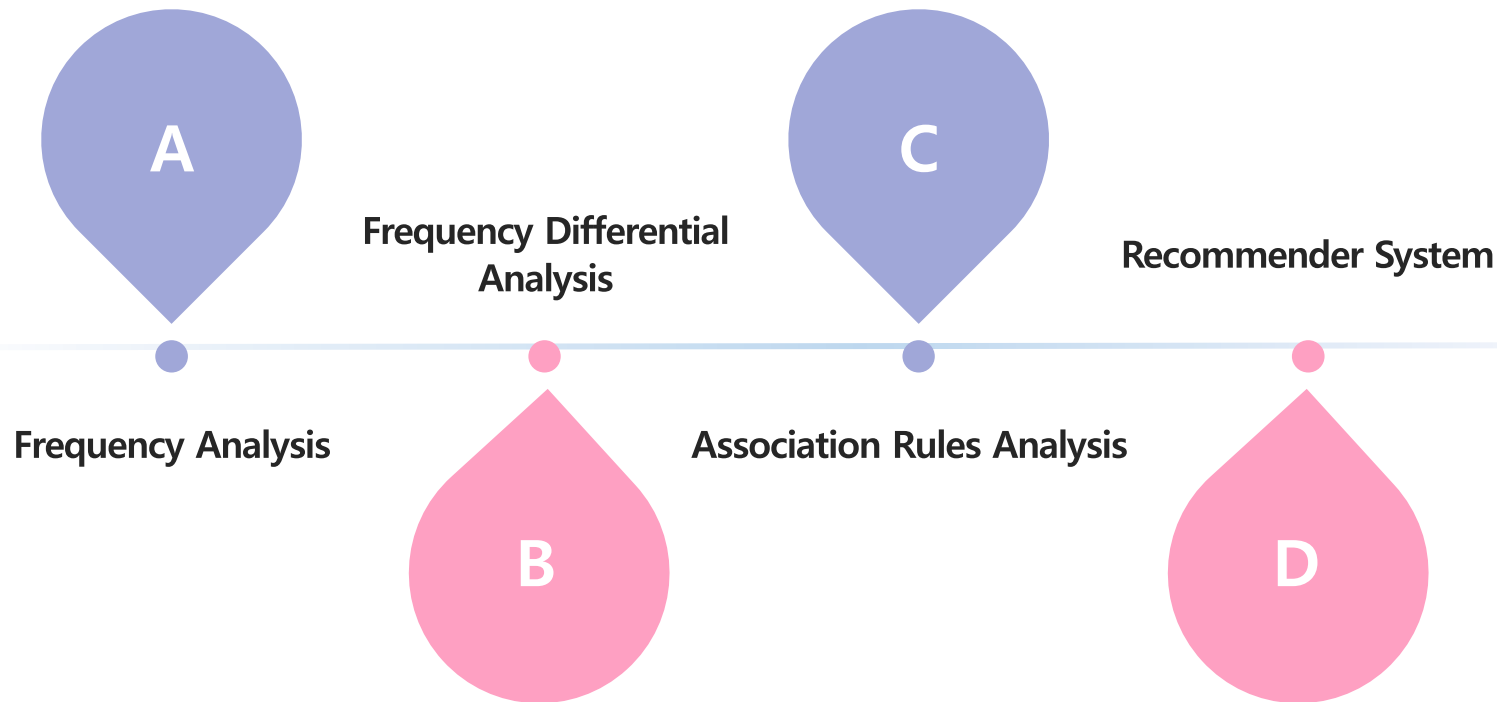
# Part 03

## Data Analysis & Recommender System





# Data Analysis & Recommender System



# Data Analysis & Recommender System

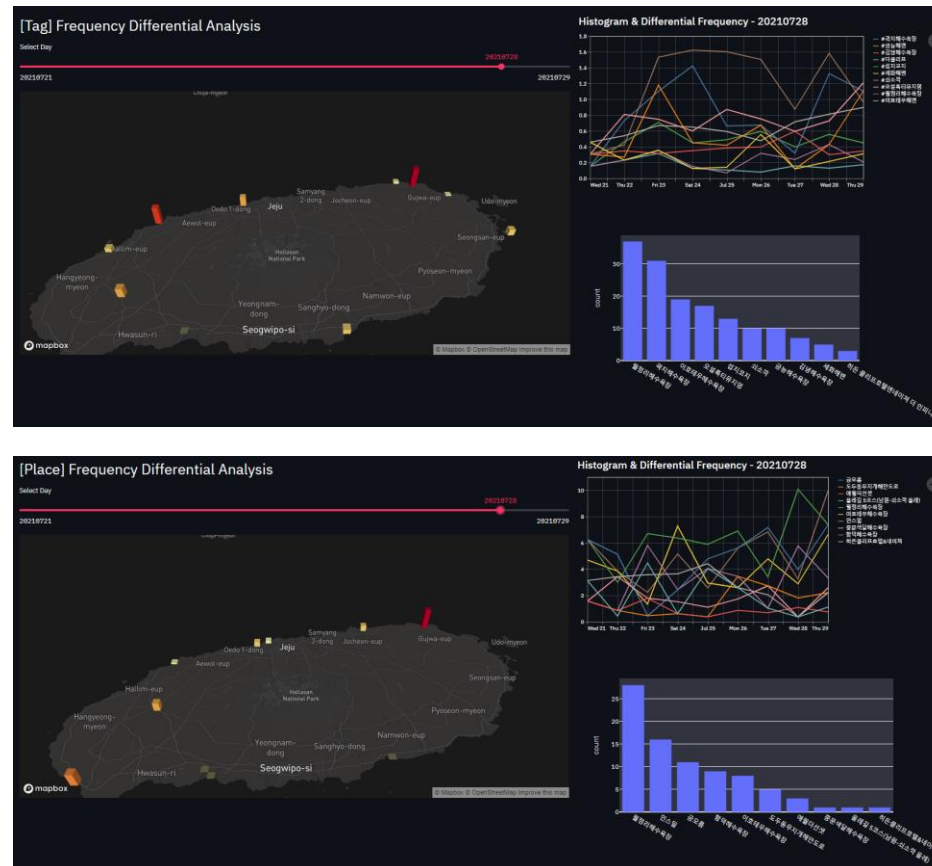
## Frequency Analysis

- Tag, Place 별 빈도수를 통해 최근 주목받고 있는 지역 분석
- 2개의 결과가 동일할수록 신뢰도 증가





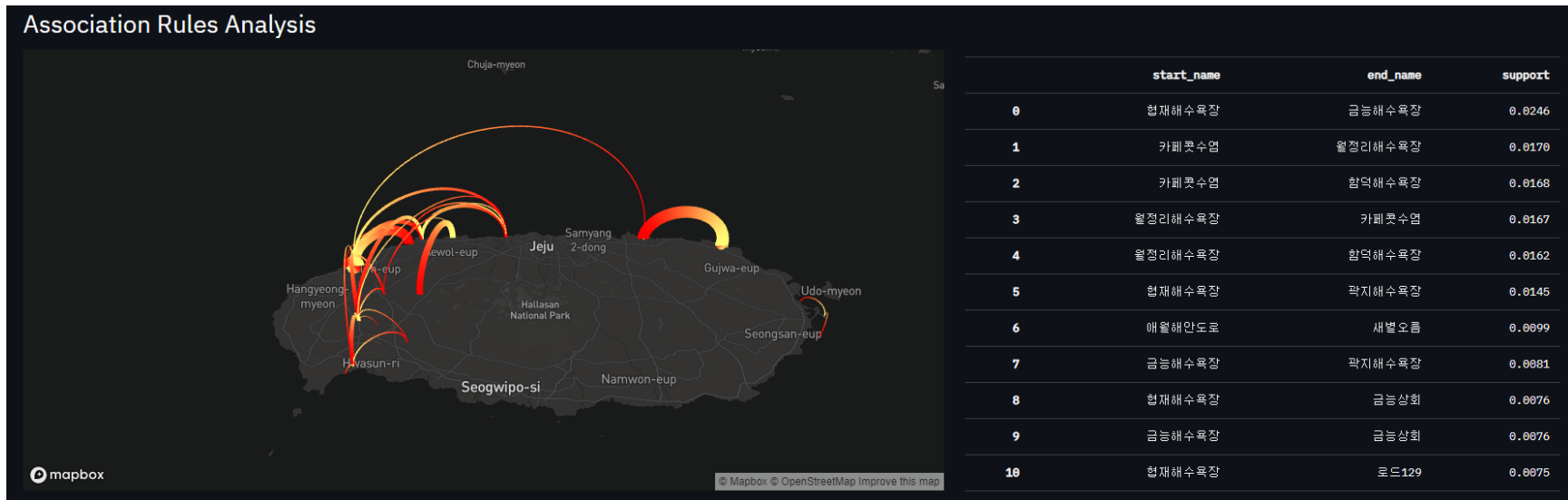
- Tag, Place 별 빈도수의 변화량을 통해 최근 점점 주목받고 있는 지역 분석



# Data Analysis & Recommender System

## Association Rules Analysis

- Apriori 라이브러리를 활용한 게시물 별 tag의 연관규칙 분석을 통해 같이 여행하면 좋은 지역 분석
- 지도 상에 그래프를 나타내어 동일 지역의 장소들로 군집 형성 확인



# Data Analysis & Recommender System

## Recommender System

- Sparse 검색 엔진인 Elastic Search를 활용하여 최신 트렌드를 고려한 추천 시스템 구현
- Elastic search에서 지원하는 형태소 분석기인 nori tokenizer 사용
- Association rules analysis를 통해 연관성, 거리 기반의 함께 여행하면 좋은 지역 추천

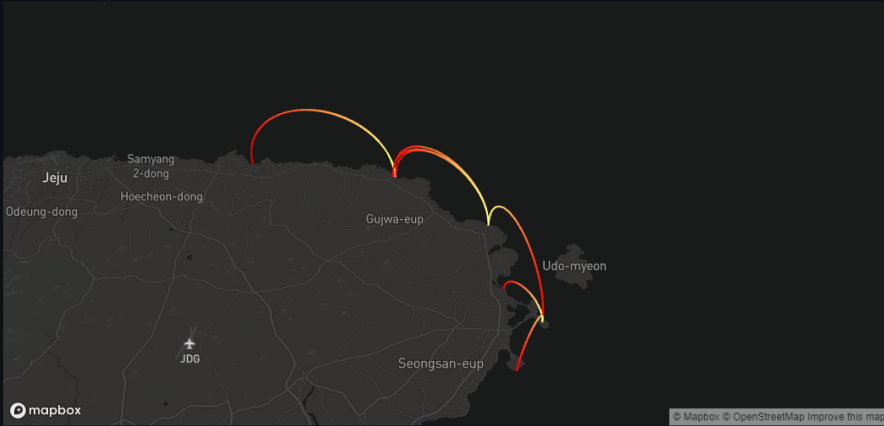
Recommender System

Enter keyword

돈까스 맛집

Submit

살편고등어



Good places to go with 살편고등어

|    | start_name | end_name | distance |
|----|------------|----------|----------|
| 0  | 살편고등어      | 카페꽃수업    | 0.0972   |
| 1  | 살편고등어      | 월정리해수욕장  | 0.0938   |
| 2  | 살편고등어      | 픽미팜클럽    | 0.0961   |
| 3  | 살편고등어      | 성산일출봉    | 0.0875   |
| 4  | 카페꽃수업      | 월정리해수욕장  | 0.0038   |
| 5  | 카페꽃수업      | 함덕해수욕장   | 0.1250   |
| 6  | 월정리해수욕장    | 카페꽃수업    | 0.0038   |
| 7  | 월정리해수욕장    | 함덕해수욕장   | 0.1273   |
| 8  | 픽미팜클럽      | 월정리해수욕장  | 0.0025   |
| 9  | 월정리해수욕장    | 픽미팜클럽    | 0.0025   |
| 10 | 성산일출봉      | 브라보비치    | 0.0364   |
| 11 | 성산일출봉      | 성지코지     | 0.0366   |



# Part 04

## Dashboard Development



# Dashboard Development

## Dashboard

- Streamlit을 활용한 dashboard 구현



# Thank you

