

데이터 공학 중간발표

서울시 축제 대백과

6조

민형규

박상우

정민지

최호경

서울시 축제 대백과란?

- 서울시의 축제를 정리하고 분석하여 사용자들이 보다 직관적으로 서울시 축제를 파악하게 해주는 상품
- 주변 지역 정보를 함께 전달함으로써 외부 방문자들에게 지역 홍보 계기 마련

개요

1. 문제정의
2. 활용데이터 및 프로그램
3. 레이아웃 및 구현방안
4. 프로젝트 진행 일정
5. 현재 진행상황

1. 문제 정의

01

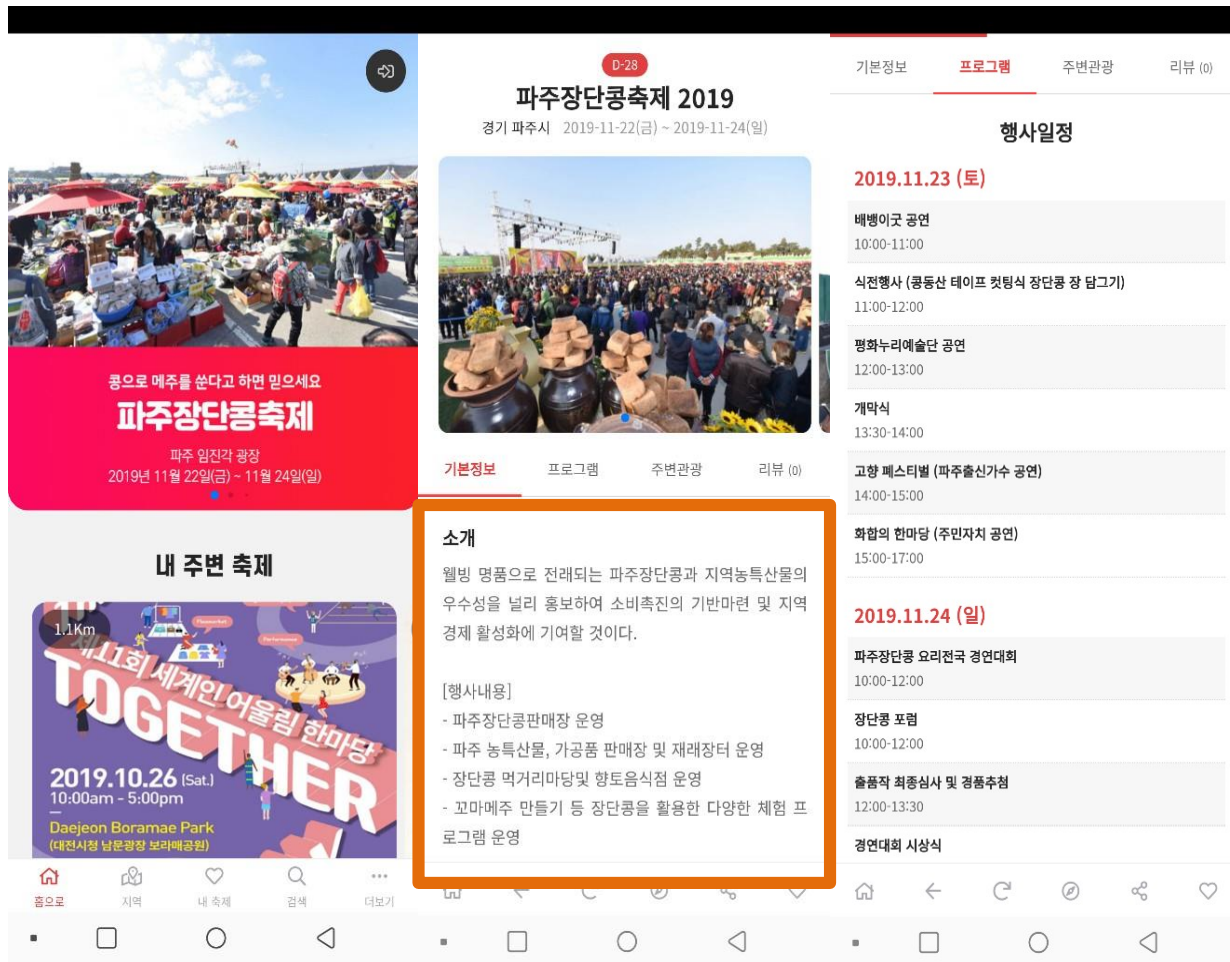
기존의 축제 추천 앱/사이트의 문제

02

03

04

05



문제점

전국 대상, 큼직큼직한 축제만 보여줌
➔ 서울 집중

기타 분석없이 관련 정보만을 제공
➔ 정보 + 데이터 분석 제공

모두의 축제

01

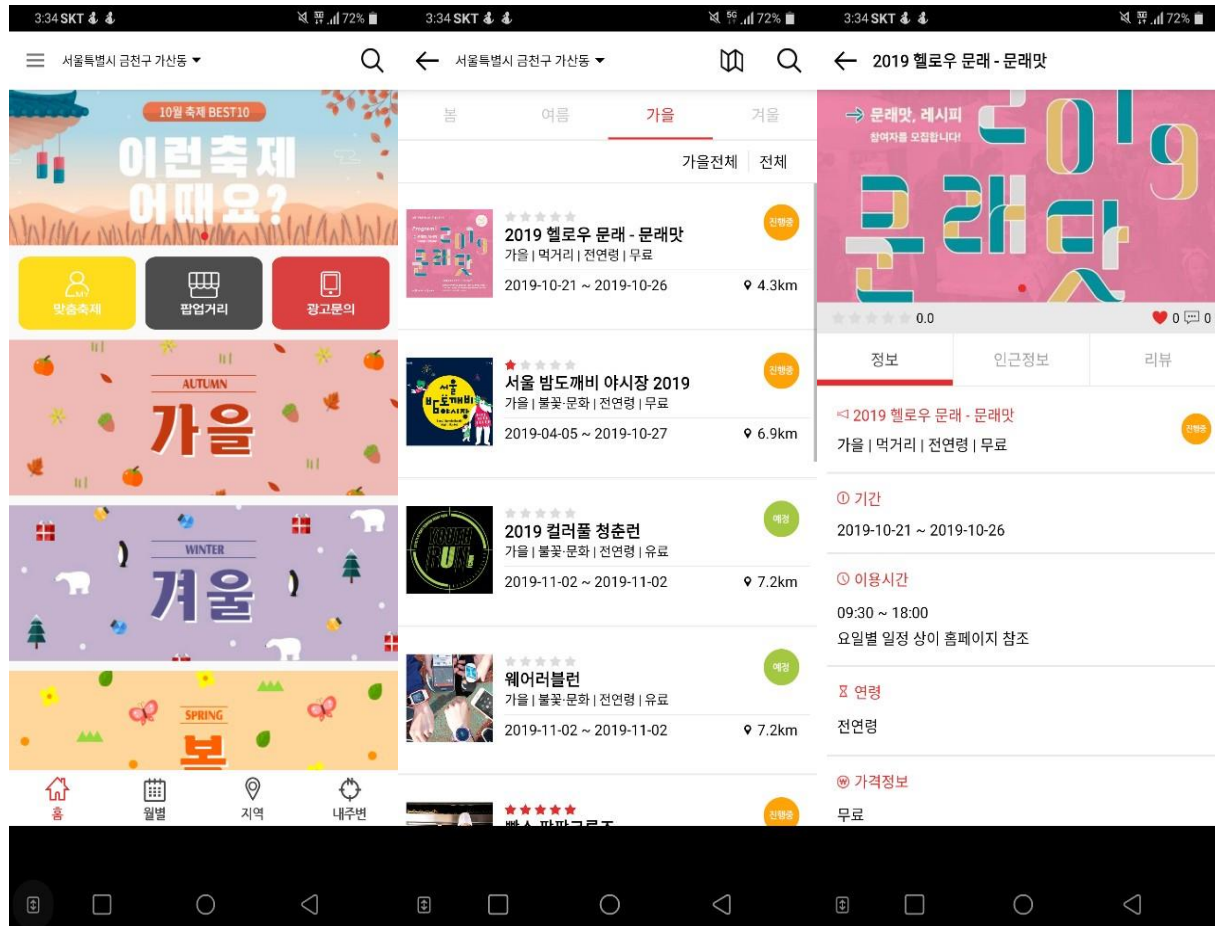
기존의 축제 추천 앱/사이트의 문제

02

03

04

05



장점

지역 기준으로 주변지역 축제 전달
➔ 벤치마킹

로그인 정보를 토대로 맞춤축제 제공
➔ 벤치마킹

단점

너무 기계적이고 피상적인 정보 제공
➔ 축제에 대한 직관적인 설명 제공

축제인

01

02

03

04

05

그래서?

1. 더 직관적인 축제 정보 및 키워드
2. 이용자 분포 분석
3. 축제 주변지역 정보

2. 활용데이터 및 프로그램

01

02

03

04

05

1. 2015/2016/2019 서울시 축제 리스트

- 문화체육관광부 제공

2. 유동인구(일별 유동인구, 거주지)

- 서울시 빅데이터 캠퍼스 제공
- 2015년 8월 ~ 2016년 7월 (일단위)

3. 크롤링

- naver, kakao API 활용

01

02

03

04

05



지역축제



추천여행지

문화예술공연

추천도서

지역축제

- 축제일정은 지자체의 사정에 따라 변경될 수 있습니다.
- 연도별 지역축제 및 문화관광축제 정보 받아보기
2016년 ▼ [내려받기](#)
- 하단의 축제정보는 한국관광공사가 운영하는 대한민국 구석구석 누리집 정보가 연
대한 자세한 사항은 한국관광공사에 문의하여 주시기 바랍니다.

축제명, 기간, 축제설명, 축제종류 등 축제에 관한 정보 제공

2016년 지역축제 총괄표

														※. 2016년 신규
연번	시도명	시군구명	축제명	2016년 개최기간	축제 주요내용	주최/주관 (담당자 연락처)	최초개최년도 및 횟수	2016년 축제 예산(안) (단위 : 백만원)	국비	시·도비	구·군 비	기타	축제종류 (택1)	축제 소재
1	서울특별시	시자체	서울드림페스티벌	5.27~5.28 (2일간)	○ 메인공연부대공연 ○ 타악관련 전시/체험/교육행사	서울시 문화예술과 (02-2133-2571)	1999년도 (17회)	400	-	400	-	-	문화예술	드림, 음악
2	서울특별시	시자체	서울거리예술축제	9.28~10.2 (5일간)	○ 국내외 초청 거리예술 공연 및 시민 참여 프로그램 등	서울시 문화예술과 (02-2133-2573)	2003년도	1,600	-	1,600	-	-	문화예술	예술, 거리
3	서울특별시	시자체	서울문화의 밤	8.26~8.27 (2일간)	○ 시민의 광장(서울광장 캠프) ○ 올빼미잔치, 서울의 발견(문화탐방 프로그램) 등	서울시 문화예술과 (02-2133-2574)	2008년도 (8회)	522	-	522	-	-	문화예술	밤, 캠프
4	서울특별시	시자체	서울김장문화제	11.4~11.6 (3일간)	○ 사랑의 김장나눔 ○ 김장문화 관련 주제전시 및 체험행사, 김치 및 부재료 판매, 문화 프로그램 등	서울시 문화예술과 (02-2133-2570)	2014년도 (3회)	1,800	-	1,800	-	-	문화예술	김치, 나눔, 문화
5	서울특별시	시자체	국악한마당	10.8~10.11 (4일간)	○ 메인공연(전통국악, 민요, 퓨전국악, 무용 등) ○ 체험 프로그램(민속놀이, 전통음식체험)	서울시 문화예술과 (02-2133-2571)	2005년도	150	-	150	-	-	전통역사	국악, 예술
6	서울특별시	시자체	서울사진축제	11.1~11.30 (30일간)	○ 전시 : 사진작가 전시, 시민참여 전시 ○ 부대행사 : 사진가 재능나눔사진관, 사진가의 열린서재, 심포지엄, 사진의 달 등	서울시 문화예술과 (02-2133-2561)	2010년도 (7회)	300	-	300	-	-	문화예술	사진
7	서울특별시	시자체	서울아리랑페스티벌	10.21~10.23 (3일간)	○ 개막공연, 아리랑 퍼레이드, 일렉트로닉 DJ공연, 아리랑 난장 등	서울시 문화예술과 (02-2133-2571)	2013년도 (3회)	1,300	-	-	-	1,300	전통역사	국악, 예술
8	서울특별시	시자체	한강 서래섬 꽃 축제	5월, 10월 중 각 2일간	○ 5월 유채꽃, 10월 메밀꽃밭 즐기기, 체험, 공연 등	서울특별시 한강사업본부 (02-3780-0796)	2004년도 (13회)	70	-	70	-	-	생태자연	꽃

01

02

자료목록

03

- SKT 일별 거주지별 성별 연령별 유동인구수 : SKT의 기준일자별 셀별 유입지별 성별 연령대별 유동인구수

04

- SKT 일별 거주지별 시간대별 유동인구수 : SKT의 기준일자별 셀별 유입지별 시간대코드별 유동인구수

05

필드구성

- SKT 일별 거주지별 성별 연령별 유동인구수

: 년월, 년월일, 블록코드, X좌표, Y좌표, 유입지 코드, 성별(2)/연령대(15)별 유동인구 30개 필드

- SKT 일별 거주지별 시간대별 유동인구수

: 년월, 년월일, 블록코드, X좌표, Y좌표, 유입지 코드, 시간대(24)별 유동인구 24개 필드

01

02

03

04

05



NAVER Developers

1. Kakao 지도 API

- 축제명 검색을 통한 축제 위치 파악
- 사용자의 현재 위치 파악

2. NAVER 검색 API

- 축제 관련 블로그/카페/네이버 웹문서 크롤링을 통해 축제 특성 파악

01

02

03

04

05



3. 레이아웃 및 구현방안

최종 결과 예시

이미지

이구현TT

01

02

03

04

05

서울 축제 대백과

서울드럼페스티벌

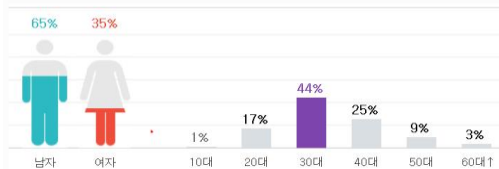


기간: 5.27~5.28 장소: 서울광장

키워드: 공연, 프로그래밍, 참여, 시민, 예술

축제설명: ○ 메인공연 부대공연 ○ 특약관련 전시/체험/교육행사

이용자 분석정보



주변지역정보



미구현 TT

1. 관련 이미지

축제 리스트

2. 정보 및 키워드

유동인구 데이터

3. 이용자 분석

4. 주변지역 정보

카카오, 구글 API

01

02

03

04

05

텍스트 마이닝

키워드 탐색

클러스터 분석

축제 성향 분류

예측 모델 구축

이용자 분포 예측

4. 프로젝트 진행 일정

PROJECT SCHEDULE 2019년 11월

SUN	MON	TUE	WED	THR	FRI	SAT
		1	2	3	1	2
		자료 수집				
3	4	5	6	7	8	9
	데이터 전처리 및 분석					
10	11	12	13	14	15	16
	시각화					
			어플리케이션 구현			
17	18	19	20	21	22	23
	어플리케이션 구현					
24	25	26	27	28	29	30
	피드백 및 수정					

- 01
- 02
- 03
- 04
- 05

PROJECT TIMELINE 2019년 11-12월

Week	11월 1주차	2주차	3주차	4주차	12월 1주차	2주차	3주차
데이터 전처리 및 분석							
시각화 과정							
어플리케이션 구현							
마무리 및 발표준비							
발표							

5. 현재 진행 상황

01

02

03

04

05

역할분담

기한 : 10월 25일 전까지

1. 최호경
 - 개발자
 - 구현할 서비스의 레이아웃 구성
2. 정민지
 - 데이터 수집
 - 축제 지역정보 크롤링
3. 박상우
 - 데이터 수집
 - 주변 지역 맛집 및 명소 추천 등의 정보 수집(크롤링)
4. 민형규
 - 데이터 분석
 - 지역 축제 데이터 설명을 활용해 축제 키워드 분석

01

02

03

04

05

축제 별 키워드 이항변수화

```
In [170]: festag19.head()
```

```
Out[170]:
```

	축제명	공연	체험	문화	프로그램	운영	전시	행사	예술	테마	...	거리	특별	개막식	포토	마을	동아리	마켓	전통	뮤지컬	퍼레이드
0	영등포 여의도 \n봄꽃축제	1	1	0	1	1	1	0	1	1	...	0	0	0	0	0	0	0	0	0	0
1	2019 이태원 지구촌 축제	1	0	1	0	0	0	0	0	0	...	1	0	0	0	0	0	0	0	0	0
2	2019 드래곤 IT 페스티벌	1	1	1	0	0	1	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	4.19혁명국민문화제2019	1	1	1	1	1	1	1	0	1	...	0	0	0	0	0	0	0	0	0	0
4	서울장미축제	1	0	1	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	1	0

5 rows × 31 columns

축제 장소 크롤링

	festival_name	x	y	address_name	시작일	종료일
0	서울드림페스티벌	126.977964	37.565691	서울 중구 태평로1가 54-3	2016-05-27	2016-05-28
1	서울사진축제	127.066776	37.640852	서울 노원구 중계동 508	2016-11-01	2016-11-03
2	한강 서래섬 꽃 축제	126.989929	37.508043	서울 서초구 반포동 1335-1	2016-05-14	2016-05-15
3	서울빛초롱축제	126.977782	37.569190	서울 중구 태평로1가 1	2016-11-04	2016-11-20
4	서울 북 페스티벌\nSeoul Book Festival\n도서관아 놀자	37.566295	126.977945	서울 중구 세종대로 110 (태평로1가, 서울시청)	2016-09-02	2016-09-03
5	한강뽕땡 여름축제	127.041625	37.538944	서울 성동구 성수동1가 642-25	2016-07-15	2016-08-21

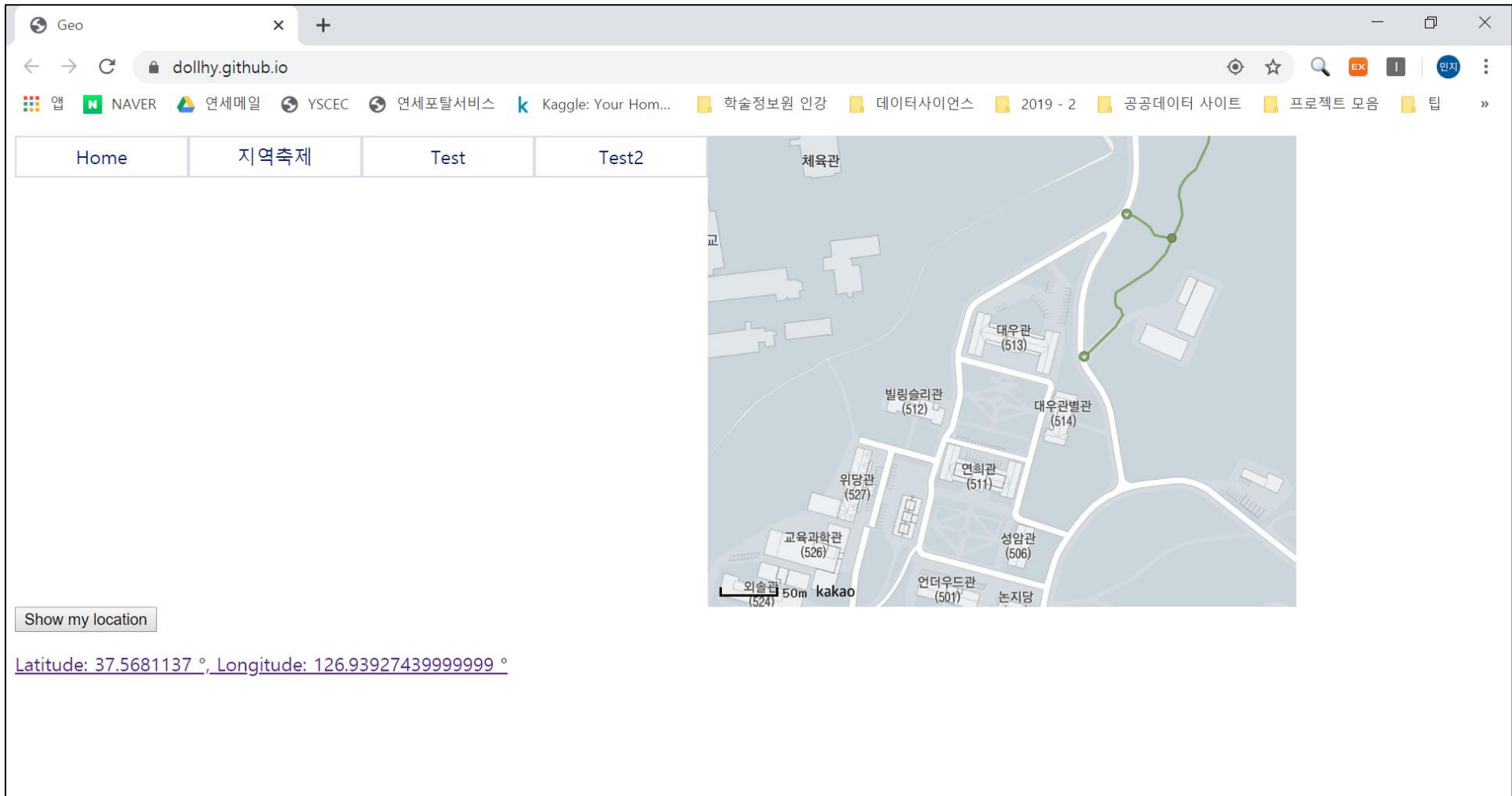
01

02

03

04

05



감사합니다