

서울시 푸드트럭 활성화를 위한 입지 선정

2조

김성희 문성만 김준희 이새문

INDEX



01. 분석배경 및 목적

- 분석주제
- 분석배경
- 분석목적




02. 데이터 분석

- 분석방향(전략)
- 분석절차(단계)
- 분석기법
- 분석과정
(시각화/요약 데이터)
- 분석결과



03. 분석 환경

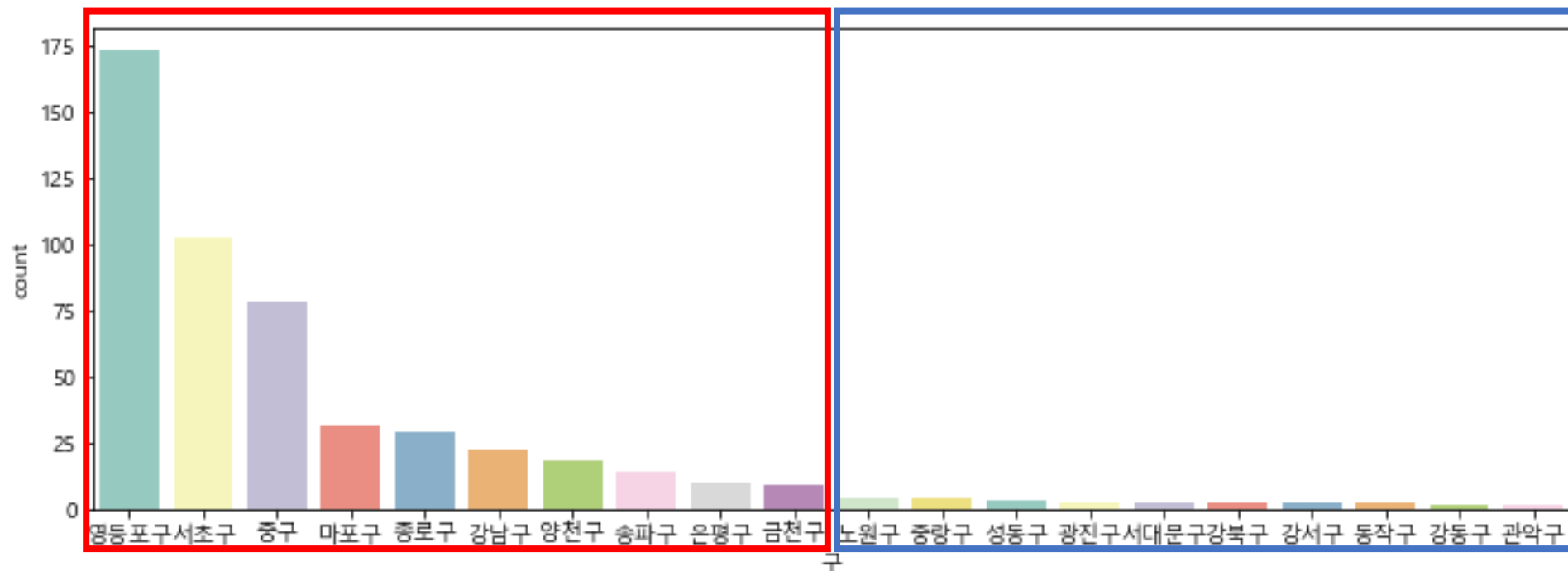
- 분석주제
- 분석배경
- 분석목적



01 분석 배경 및 목적



분석주제



현재 서울시의 푸드트럭은 20개 구에서 500여 곳이 운영 중이다. 푸드트럭의 운영대수를 구 별로 확인해보면 활성화된 구와 그렇지 않은 구가 명확하다. 이에 활성화되어 있지 않은 구를 대상으로, 앞으로 푸드트럭이 원활하게 운영될 수 있는 입지를 추천하고자 한다.



분식배경

최근 푸드트럭이 인기다. 신문 방송 등 미디어들은 너 나 할 것 없이 푸드트럭에 대해 이야기하고, 이러한 관심에 부응하듯 푸드트럭이 창업 아이템으로 급부상하고 있다.

그러나.....



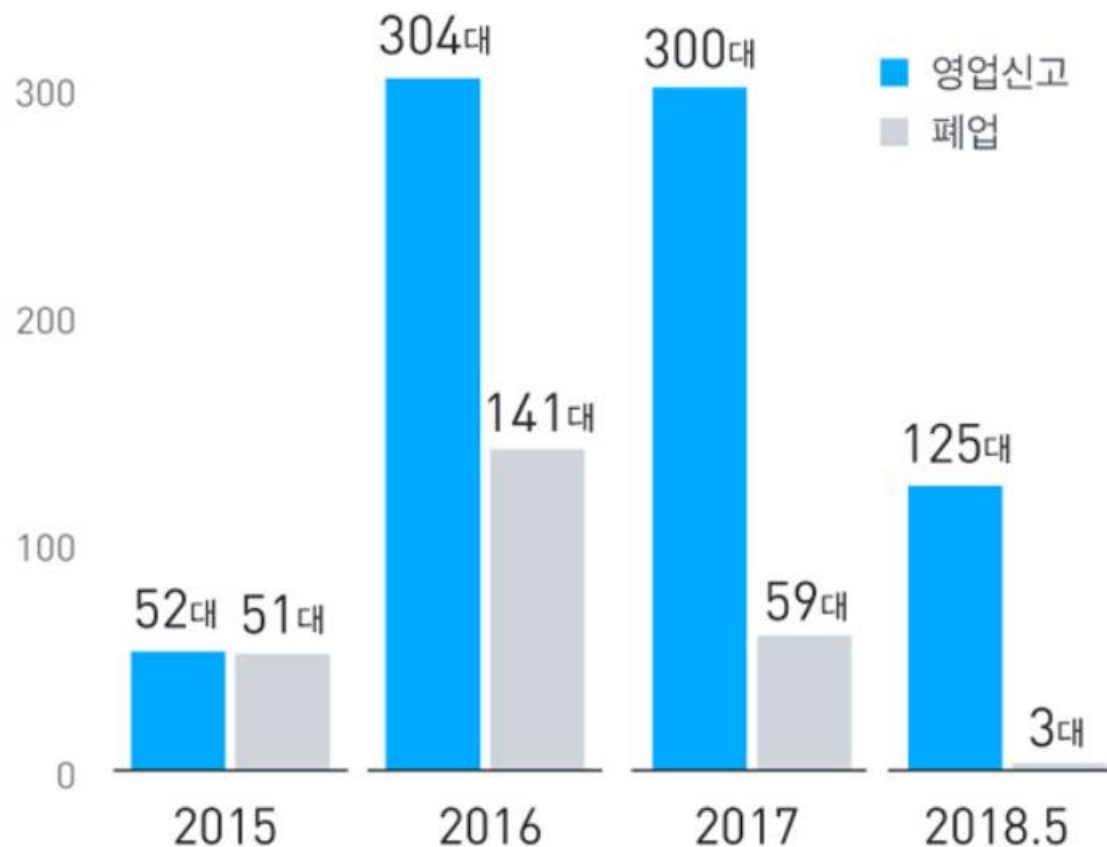
분식배경

창업하는 만큼 많이 폐업한다.

2015년부터 현재까지 700여대가 창업하고 250여대가 폐업해 약 36%의 폐업률

특히 푸드트럭을 본격적으로 장려하기 시작한 2016년은 거의 50%

서울시 푸드트럭 영업신고 및 폐업 현황 (단위: 대)





분식배경

장려만 할 뿐 어떻게 활성화시킬지 대책이 없었기 때문이다.

특히 입지적인 문제를 간과할 수 없다.


한국의 경우 정부(지방자치단체)가 허가한 구역에서만 푸드트럭을 운영할 수 있기 때문에, 푸드트럭 산업의 활성화를 위해서는 허가 구역을 정하는 정부의 역할이 중요하다.

다른 요식업들이 그렇듯 푸드트럭 또한 높은 매출을 올리기 위해선 입지가 매우 중요하다. 때문에 허가구역 입지에 대한 심도 있는 분석이 없다면 추가 지역을 선정하더라도 하나하나 한 결과를 불러올 수 있다.



분석목적

- 푸드트럭 활성화 지역 확대
 - 푸드트럭 산업이 활성화 되어 있지 않는 자치단체 중 잠재력 있는 곳을 탐색
 - 그 중에서도 입지적으로 가장 적합한 지역을 선정
 - 앞으로 어떤 지역 위주로 확대해 나갈 것인지를 추천



02 데이터 분석



분석 방향(전략)

■ 의문

- 어떤 지역이 푸드트럭 허가 구역을 선정하기에 적합한가?
 - ✓ 높은 매출을 올릴 만큼 유동인구를 보장하는 곳은 어디인가?
 - ✓ 주요시설들이 많은 곳은 어디인가?
 - ✓ 기존 상권과의 갈등이 적을 만한 곳은 어디인가?

■ 가정

- 해당 지역이 주요 시설을 얼마나 가지고 있는지, 생활 인구는 얼마인지 분석하면 푸드트럭 허가 구역을 선정하기에 최적의 위치를 알아낼 수 있다.



분석 방향(전략)

■ 기법

1. 스코어링 :

- 동 단위로 분석
- 각 지역의 생활인구(거주인구+유동인구)에 따른 점수 부여
- 각 지역에 있는 주요 시설의 규모에 따른 점수 부여



분석 방향(전략)

■ 기법

2. 시각적 분석 :

- 보조적 기법
- 해당 동 안에서는 어떤 지역을 피하고 어떤 지역을 선택해야 하는지 대략적으로 파악하기 위해
- 주로 Python의 Folium 패키지의 HeatMap과 ClusterMarker를 사용해 지도 위에 변수가 밀집된 정도를 표시.
- 이후 이미지화한 변수를 레이어로 겹쳐 색이 짙은 부분을 최종적으로 선택.

3. R(통계분석), QGIS(지리정보)를 보조적으로 사용



분석 방향(전략)

■ 분석

1. 특정 구에서 동 단위 스코어링을 통해 점수가 가장 높은 동을 선별
2. 시각화 자료를 통해 동 내부에서 어느 구역이 적합하고 어느 구역이 적합하지 않은지를 표시
3. 이후 서울시 전체 동에 적용



분석 절차(단계)

- 자료수집
 - 공공데이터포털 등에서 수집한 공공데이터
 - 원하는 정보가 누락된 경우가 있음
 - 좌표가 없거나, 지번주소와 도로명주소가 섞여있음
 - 네이버 Open API를 사용해 주소를 검색해 좌표를 찾거나, 좌표를 검색해 주소를 검색함



분식 절차(단계)

■ 전처리

이름	전처리 방법	이유
도시공원 정보	어린이 공원 및 소공원을 제외. 넓이 : 모집단의 평균 이상 좌표계 변환	푸드트럭을 설치하고도 여유가 있는 곳 위주로 선발 Folium에 표시가능한 좌표로 변환
대학 및 전문대학 정보	재학생 수 1000명 이상 크롤링으로 좌표 정보 수집	인원수가 너무 적은 곳은 유동인구가 많다고 볼 수 없다고 판단 지도에 놓기 위해 좌표 정보 필요
공공체육시설	넓이 : 모집단의 중위수 이상 학교 운동장은 제외	푸드트럭을 설치하고도 여유가 있는 곳 어느 정도 이상의 인원을 수용할 수 있어야 많은 유동인구를 기대해볼 수 있음
도로	폭 10m 이상 구 코드를 구 이름으로 변환	푸드트럭 설치 및 대기줄을 고려하면서 통행에 방해가 되지 않을 기준을 설정
호텔 정보	크롤링으로 좌표 정보 수집	지도에 놓기 위해 좌표 정보 필요



분석 절차(단계)

■ 전처리

이름	전처리 방법	이유
관광명소	위키피디아에서 수집한 정보를 토대로 크롤링으로 좌표 정보 수집	지도에 놓기 위해 좌표 정보 필요
지하철역 정보	각 지하철역의 행정구 표기(크롤링) 잘못찍힌 지하철역의 좌표 수정(크롤링)	각 지하철역을 행정구별로 구분하기 위함
생활인구 정보	행정동 코드를 행정동 명으로 변환	식별의 편의를 위해
문화공간 현황	주소 정보를 공백을 기준으로 쪼갬	주소 중 행정구만 추출하기 위해
초중고 정보	필요한 정보만 추출 (이름, 행정구, 좌표)	작업을 효율성을 위해



분석 절차(단계)

■ 전처리

이름	전처리 방법	이유
상가업소 정보	음식점들의 정보를 필터링	기존 자영업자들의 밀집지역을 고려해야 함
푸드트럭 정보	주소에서 행정구 정보를 추출	푸드트럭들을 행정구별로 분류하기 위함



분석 절차(단계)

■ 2차 탐색적 분석

- 1. 구별로 몇 대의 푸드트럭이 운영되고 있는지 집계
- 2. 상위 3개 구에 70%, 10개 구에 95%의 푸드트럭이 밀집되어 있음을 확인

운영구	운영대수	푸드트럭 분포율	누적분포		
영등포구	173	33.98821218	33.98821218		
서초구	102	20.03929273	54.02750491		
중구	78	15.32416503	69.35166994	69.35166994	1그룹
마포구	31	6.090373281	75.44204322		
종로구	29	5.697445972	81.13948919		
강남구	22	4.322200393	85.46168959		
양천구	18	3.536345776	88.99803536		
송파구	14	2.750491159	91.74852652		
은평구	10	1.964636542	93.71316306		
금천구	9	1.768172888	95.48133595	26.12966601	2그룹
노원구	4	0.785854617	96.26719057		
중랑구	4	0.785854617	97.05304519		
성동구	3	0.589390963	97.64243615		
강북구	2	0.392927308	98.03536346		
강서구	2	0.392927308	98.42829077		
광진구	2	0.392927308	98.82121807		
동작구	2	0.392927308	99.21414538		
서대문구	2	0.392927308	99.60707269		
강동구	1	0.196463654	99.80353635		
관악구	1	0.196463654	100	4.518664047	3그룹
용산구	0	0.00%	0%		
동대문구	0	0.00%	0%		
성북구	0	0.00%	0%		
도봉구	0	0.00%	0%		
구로구	0	0.00%	0%		
SUM	509	100.00%	100%		

하위5% 지역:

노원구, 중랑구, 성동구, 강북구, 강서구, 광진구, 동작구, 서대문구, 강동구, 관악구

0%지역

용산구, 동대문구, 성북구, 도봉구, 구로구



분석 절차(단계)

■ 2차 탐색적 분석

- 3. 기존에 푸드트럭이 활성화 되어있는 지역을 대조군으로 설정해 우선적으로 분석.
- 우리가 설정한 모델이 대조군과 비교해 신뢰도가 있는지 확인하기 위함

행정구명	푸드트럭 대수	전체 대수 대비 비율
영등포구	173	34%
서초구	102	20%



분석 기법

■ 1. 스코어링

- 주요시설과 생활인구의 점수 배분 비율 1.618 : 1
- 각 주요시설의 가중치(기존 푸드트럭의 8개 주요시설 설치 현황을 기준으로)

	한강공원	관광명소	문화시설	보행자도로	체육시설	지하철	공원개수	대학교	초중고	호텔
총 갯수	236	67	59	33	23	19	16	3	0	0
비율	51.416122	14.5969499	12.8540305	7.189542484	5.010893246	4.139433551	3.485839	0.653594771	0	0
순위	1	2	3	4	5	6	7	8	9	10
누적확률	51.416122	66.0130719	78.8671024	86.05664488	91.06753813	95.20697168	98.69281	99.34640523	99.34641	99.34640523
비율 반올림	51	15	13	7	5	4	3	1	1	1
가중치	0.5	0.15	0.13	0.07	0.05	0.04	0.03	0.01	0.01	0.01



분석 기법

▪ 2. Python의 Folium 패키지를 이용한 지도 분석

- Marker를 이용해 각 자료로부터 수집한 좌표를 표시
 - ✓ 각 자료들은 법률상으로 푸드트럭을 설치할 수 있는 장소들과 유동인구가 많은 곳의 좌표로 구성
 - ✓ 사용한 좌표들 : 유명 관광장소, 호텔, 공공체육시설, 문화공간, 초/중/고등학교, 대학교, 도시공원, 지하철역
- 좌표들이 밀집한 지역에 유동인구가 많을 것이라 추정
 - ✓ ClusterMarker를 이용해 좌표의 밀집된 정도를 지도 상에 표시
 - ✓ Folium의 starting zoom은 14로 고정
 - ✓ Blur = 20, radius = 30, min_opacity = 0.5



분석 기법

- R을 이용한 상가밀집도 분석
 - R의 커널 밀집도 분석 툴을 사용, 음식점들의 밀집도를 등고선 그래프로 출력
 - 밀집폭은 좌표 정보를 활용($\text{bandwidth} = 0.0015, 0.0015$)



분석 과정

■ 스코어링

- 점수를 매긴 기준이 기존에 푸드트럭이 활성화된 (구 단위)지역과 일치하는지 대조
- 대조 결과 높은 70%의 점유율을 가지고 있는 **영등포구**와 **서초구**가 포함됨
- 이를 바탕으로 **하위 5%** 동에 대해 스코어링 적용

Show 10 entries Search:

	group	x	rank
20	영등포구	0.265788183	1
18	송파구	0.256118455	2
15	서초구	0.194567822	3
23	종로구	0.15872246	4
13	마포구	0.152503526	5
6	광진구	0.143445247	6
1	강남구	0.141093624	7
2	강동구	0.140215514	8
21	용산구	0.137546091	9
24	중구	0.100131279	10

Showing 1 to 10 of 25 entries Previous 1 2 3 Next



분석 과정

■ 스코어링

- 송파구 높은 이유
 - ✓ 한강공원 多
 - ✓ 문화시설 多
- 중구 낮은 이유
 - ✓ 다른 구에 비해 동 수가 적음
 - ✓ 한강공원이 없음

Show	10	entries	Search:	
	group	x	rank	
20	영등포구	0.265788183	1	
18	송파구	0.256118455	2	
15	서초구	0.194567822	3	
23	종로구	0.15872246	4	
13	마포구	0.152503526	5	
6	광진구	0.143445247	6	
1	강남구	0.141093624	7	
2	강동구	0.140215514	8	
21	용산구	0.137546091	9	
24	중구	0.100131279	10	

Showing 1 to 10 of 25 entries

Previous 1 2 3 Next



분석 과정

- 스코어링

- 설정한 기준을 가지고 하위 5% 구에 있는 동을 스코어링

- 스코어링 결과와 순위

Show

10

entries

Search:

	행정 동	구	SCORE	rank
1	암사2동	강동구	0.06011271	1
2	자양3동	광진구	0.034616842	2
3	서빙고동	용산구	0.032149317	3
4	천호2동	강동구	0.031952374	4
5	이촌1동	용산구	0.03181528	5
6	방화3동	강서구	0.031050981	6
7	성수2가1동	성동구	0.030785731	7
8	자양4동	광진구	0.030476986	8
9	자양2동	광진구	0.029628266	9
10	신촌동	서대문구	0.017564899	10

Showing 1 to 10 of 236 entries

Previous

1

2

3

4

5

...

24

Next



분석 과정

- 상위 2개 선정 지역을 클러스터링으로 지도에 표시
 - 모든 데이터프레임에서 이름, 위도, 경도 column만을 추출해 하나의 데이터프레임으로 통합 후 새로운 지도 위에 클러스터링을 이용해 적용

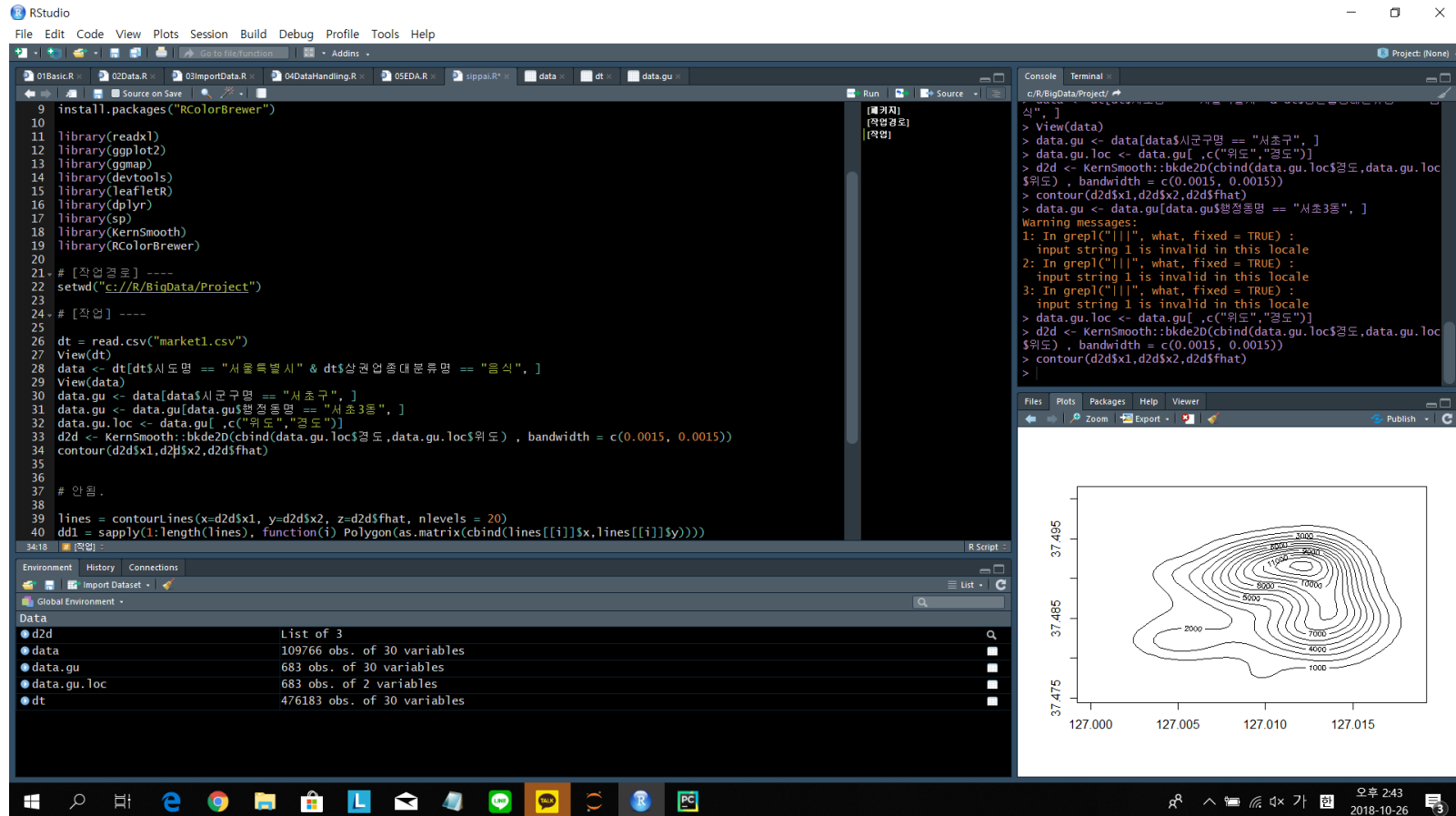
```
df = pd.concat([cul_gu[['이름', 'X', 'Y']], tra_gu[['이름', 'X', 'Y']], gym_gu[['이름', 'X', 'Y']], park_gu[['이름', 'X', 'Y']], hotel_gu[['이름', 'X', 'Y']])
```

```
heat_data = [[row.X, row.Y] for index, row in df.iterrows()]  
map3 = folium.Map(location = [37.5792250, 126.9368000], zoom_start = 13.5)  
HeatMap(heat_data, max_zoom = 200, blur = 20, radius = 25).add_to(map3)  
map3
```



분석 과정

- 상가 밀집 그래프를 R을 이용해 구현





분석 결과

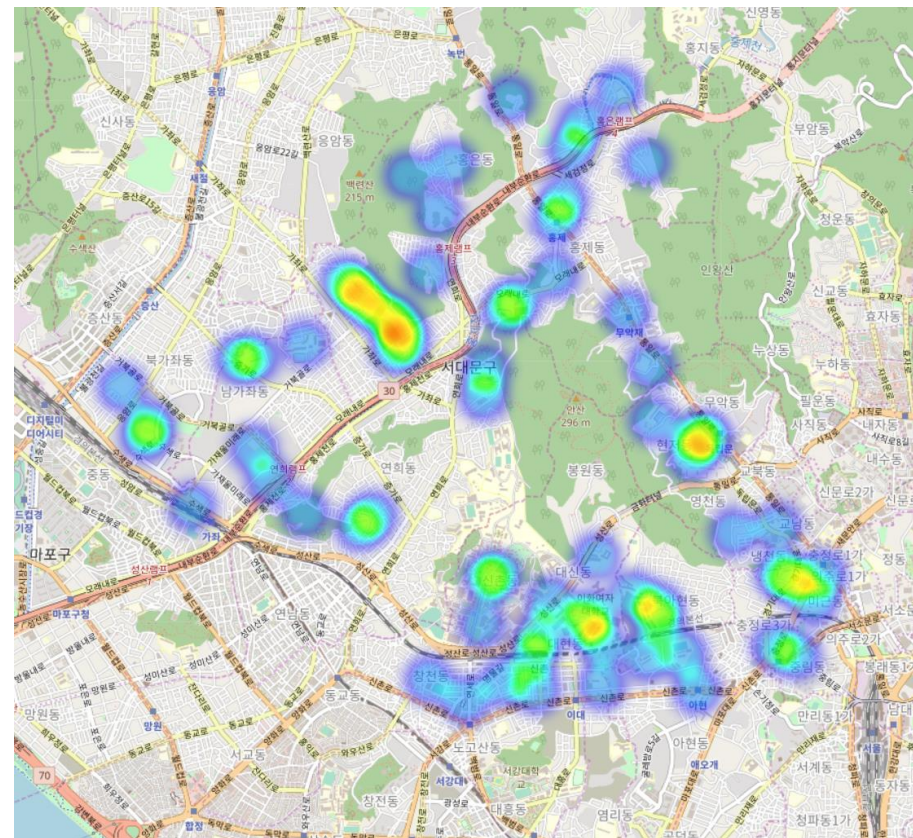
- 동 내 구체적인 위치 시각화
 - ❖ 광진구 자양 3,4 동
- ✓ 클러스터링이 높은 지역
 - 주요 시설 밀집화
- ✓ 등고선이 낮은 지역
 - 상권 밀집도가 낮은 지역
 - 밀집도가 높을 수록 등고선이 높아짐






분석 결과

- 한계점
 - 푸드트럭 매출액과 폐업률 자료 無
 - ✓ 스코어링 적용의 한계
 - 단위 면적당 밀집도 분석이 필요
 - ✓ 동을 선정한 후 시각적으로 위치를 선정
- 향후 과제
 - 푸드트럭 매출액과 폐업률 자료를 바탕으로 푸드트럭 활성화에 미치는 요인분석 정교화
 - 위치 기반 밀집도 분석을 이용하여 단위 면적당 밀집도 분석을 통해 정밀한 위치 선정





03 분석 환경



분석 도구

분석 도구	버전	용도
Python	3.7	데이터 분석 및 시각화
R Rstudio	10.3	SQL
QGIS	2.18	데이터 시각화
Excel	Office365	데이터 전처리, 자료 정리
Polaris Office	2017	데이터 전처리, 자료 정리



분석 도구

자료이름	생산년도	출력년도	생산자	출처	수집/생명주기
서울특별시 도시공원정보	2018.08	2018.08	서울특별시 각 구청	공공데이터포털	2017.08~2018.08
서울시 대학 및 전문대학 DB 정보	2016.02	2016.03	서울특별시	서울열린데이터광장	~2016.02
서울시 공공체육시설 현황	2018.10	2018.10	서울특별시	서울 정보소통광장	~2018.10
서울시 문화공간 현황	2018.02	2018.03	서울특별시	서울 정보소통광장	~2018.02
서울시 역코드로 지하철역 위치 조회	2018.02	2018.03	서울교통공사	서울 정보소통광장	~2018.02
행정동별 서울생활인구(내국인)	2018.03	2018.03	서울특별시	서울열린데이터광장	2018.10~2018.10
2018년2차 공시대상학교정보	2018	2018	한국교육학술정보원	학교알리미	2018
서울시 관광호텔 등록 현황 통계	2018.04	2018.04	서울특별시	서울 정보소통광장	~2018.04
서울시 관광명소	2018.09	2018.09	위키피디아	위키피디아	
서울시 도로구간 위치정보	2014.10	2014.10	서울특별시	서울열린데이터광장	~2014.10(비정기)
소상공인시장진흥공단 상가업소 정보	2018.10	2018.10	소상공인시장진흥공단	공공데이터포털	2018.06~2018.09
서울특별시 푸드트럭 정보	2018.05	2018.05	서울특별시	서울열린데이터광장	2017.04~2018.04