

# 전통시장 활성화지수 개발을 통한 DT 입지선정

지역사회 공헌 아이디어 경진대회 - 전통시장 활성화 프로젝트

박성희 팀

박성희  
강동현  
박선영  
최정원



고려대학교  
KOREA UNIVERSITY



## 1

### 분석 개요

분석 배경 및 목적  
모형 구축을 위한 이론적 검토  
flow chart

## 2

### 데이터 전처리

분석 데이터 설명  
변수 설명  
데이터 생성 과정  
데이터 전처리 및 정제 과정

## 3

### 분석 결과

모델 설명  
고객/시장/외부 요인 분석 결과  
요인별 표준화  
가중치 산출  
전통시장 활성화지수 산출 결과  
활성화지수 유의성 검정

## 4

### 활용방안 및 시사점

활성화지수 기반 DT 입지선정  
향후 활용방안  
본 과업의 한계점

# 분석 개요

분석 배경 및 목적  
모형 구축을 위한 이론적 검토  
flow chart

## 1. 분석 배경 및 목적

### Traditional Market Revitalization Analyzer (전통시장 활성화 분석)

고객 요인

시장 요인

외부 요인

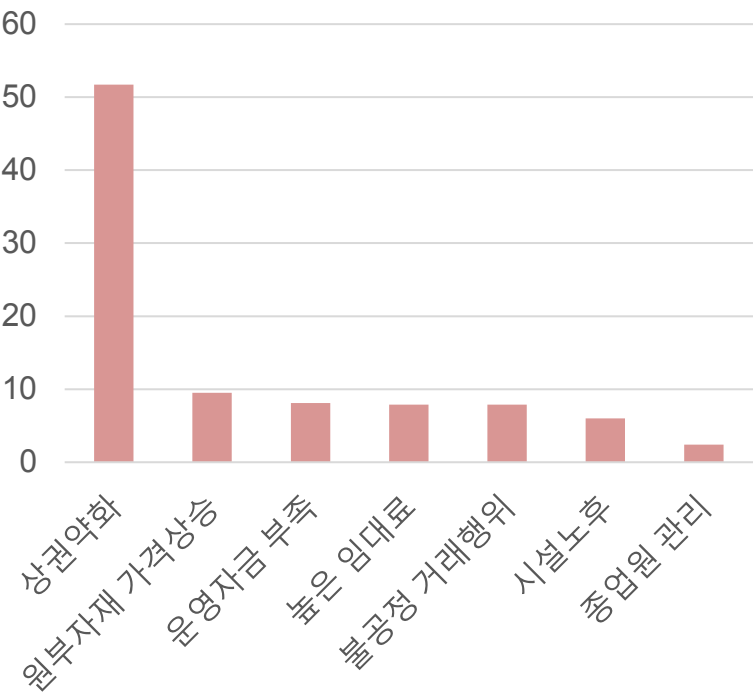


전통시장 활성화 지수 생성

# 1. 분석 배경 및 목적

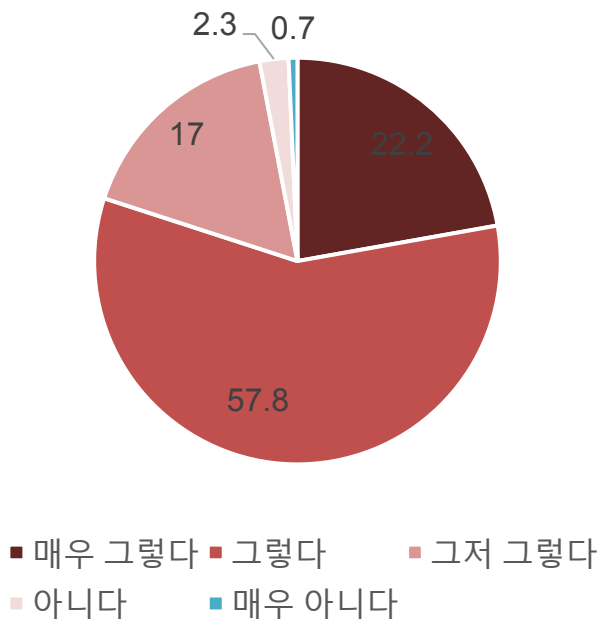
## ① 전통시장의 정부지원 필요성

[전통시장 애로사항 원인]



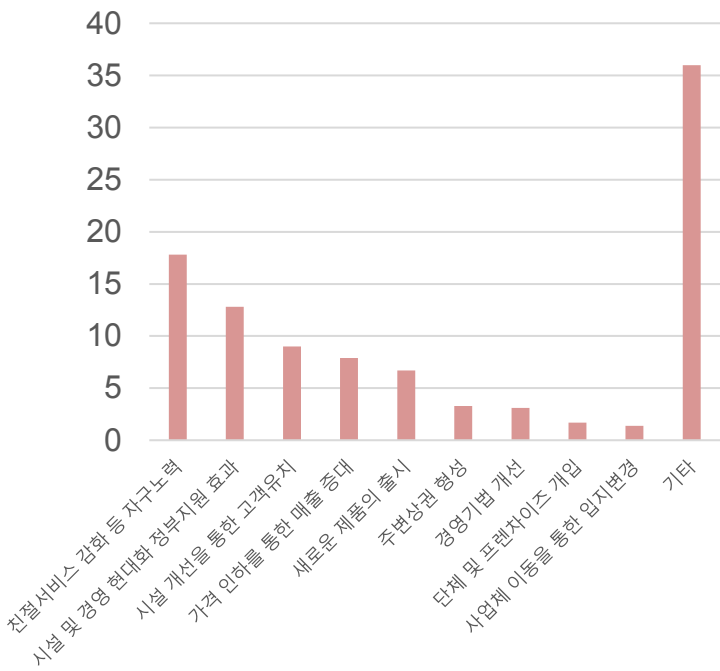
상권약화가 애로사항의 원인

[정부지원사업 고객 증가 기여도]



정부지원사업이 전통시장 고객 증가에 기여도가 높다

[영업이익 증가이유]



온라인 판매 및 택배 서비스가 영업이익에 큰 영향을 줌

# 1. 분석 배경 및 목적

## ② Drive Through 도입 이유

### 언택트 소비 지속...관련 상품 '붐물'

#### "집콕은 싫고 바깥은 위험해"...코로나 장기화에 '드라이브 스루' 전성시대

기사입력 : 2020년09월12일 07:40 | 최종수정 : 2020년09월12일 08:19



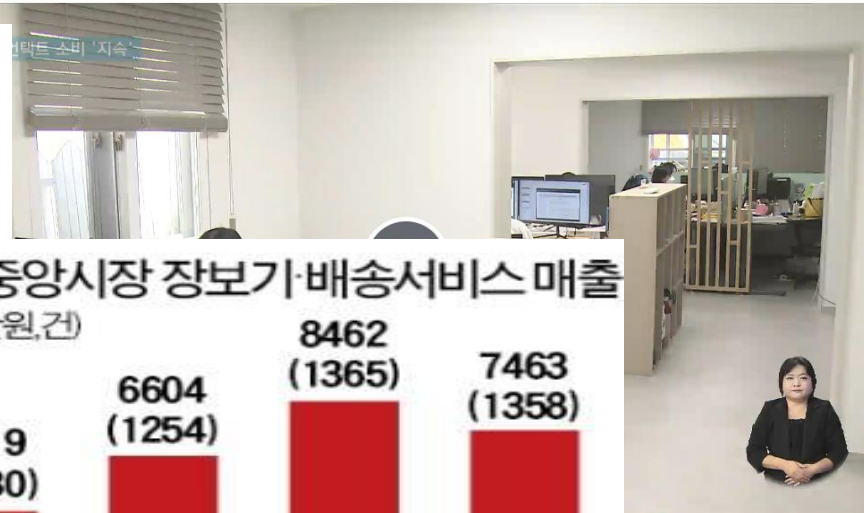
가 + 가 - 프린트

코로나19로 비대면 소비 증가...드라이브 스루 매장 손님 ↑  
프랜차이즈부터 마트·백화점도 가세...비용·소음 등 문제도

[서울=뉴스핌] 장주연 기자 = #1. 새신부 이신혜(33)씨는 지난 주말 남편과 오랜만에 외출에 나섰다. 로나바이러스 감염증(코로나19) 장기화로 집에 보내는 시간이 많아지면서 심신이 지친 까닭이다. 와 달리 자동차 데이트에 나섰다. 자동차 극장에 가기 전 부부가 들른 곳은 **스타벅스** 용인마북DT(루)점이다. 유동 인구가 많은 카페에 들어가지 않고도 간편하게 음료를 살 수 있는 방법을 택했다.

#2. 직장인 김지은(35)씨는 점심시간이 시작되자 출근길 타고 왔던 차에 다시 올라탔다. 목적지는 구에 위치한 **맥도날드** 고척DT점. 회사 근처에 **맥도날드**가 있긴 있지만 더 먼 곳을 선택했다. 차에 앉고 주문해 먹을 수 있는 '안전한' 맥드라이브 매장을 이용하기 위해서다.

코로나19 사태가 8개월 넘게 이어지면서 소비자들의 소비 형태에도 변화가 감지되고 있다. 매장 내에서 방문 언택트(비대면) 주문건은 증가세다. 특히 드라이브 스루(차량에서 내리지 않고 주문하기) 인기가 뜨겁다. 이에 유명 프랜차이즈 매장은 물론 편의점, 개인 식당까지 드라이브 스루 시스템 도입을 검토하는 분위기다.



- 코로나 시대 이후 언택트 소비의 증가
- 배송 서비스 및 드라이브 스루 이용 고객의 증가

## 2. 모형 구축을 위한 이론적 검토

### 고객요인이 전통시장 발달에 미치는 영향

- 시장의 **주변 인구수** 분석은 전통시장을 찾는 수요를 가늠할 수 있는 중요한 데이터
- 시장의 내·외부 여건이 비슷한 두 시장을 비교했을 때, 비교적 넓은 면적의 세대수가 많고 주택 가격이 더 비싸게 형성된 시장의 평균 매출액이 거의 두 배가량 높게 나왔음.

[Reference] 박상훈, 이희정(2018)

### 시장요인이 전통시장 발달에 미치는 영향

- 시장의 외적요소(**시설수준 및 접근성**)에 대한 인식이 높을수록 시장의 재방문 횟수가 많아지고, 그 효과는 큰 것으로 나타남

[Reference] 임정빈, 김덕준(2017)

- 근린시장과 중소시장의 경우, **축산물**을 취급하는 점포의 비중이 상대적으로 높을수록 재래시장은 활성화된 것으로 나타남. 기본적으로 점포 비중이 높은 **농산물 및 수산물** 점포 역시 시장 활성화에 매우 중요한 역할을 함

[Reference] 박주영(2009)

### 외부요인이 전통시장 발달에 미치는 영향

- 시장 **주변 유통상업시설**(대형마트, 편의점 등) 현황은 시장 활성화에 큰 영향을 미치는 요소

[Reference] 박상훈, 이희정(2018)

- 근처 **대형마트** 입점으로 인해 매출이 크게 감소한 시장을 대상으로 연구한 논문이 있음

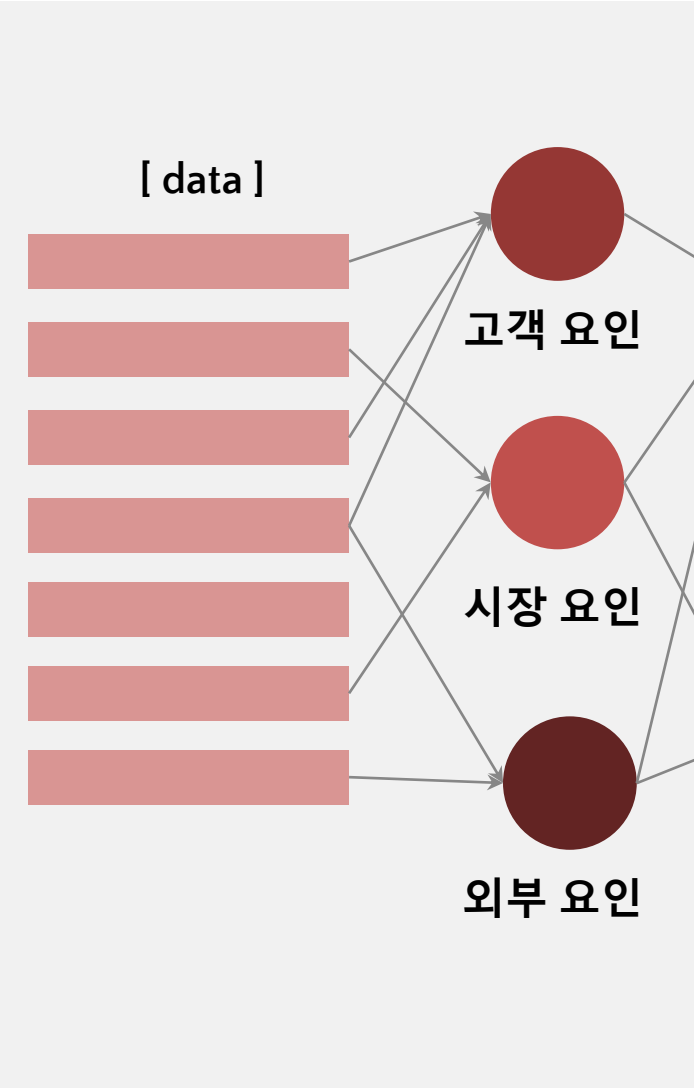
[Reference] 김민정, 류성민(2016)

- **유동인구**의 변화는 상권변화를 예측할 수 있을 뿐 아니라 유동인구의 감소는 상권의 위축과 소매점포의 폐업을 증가와 관련이 있음

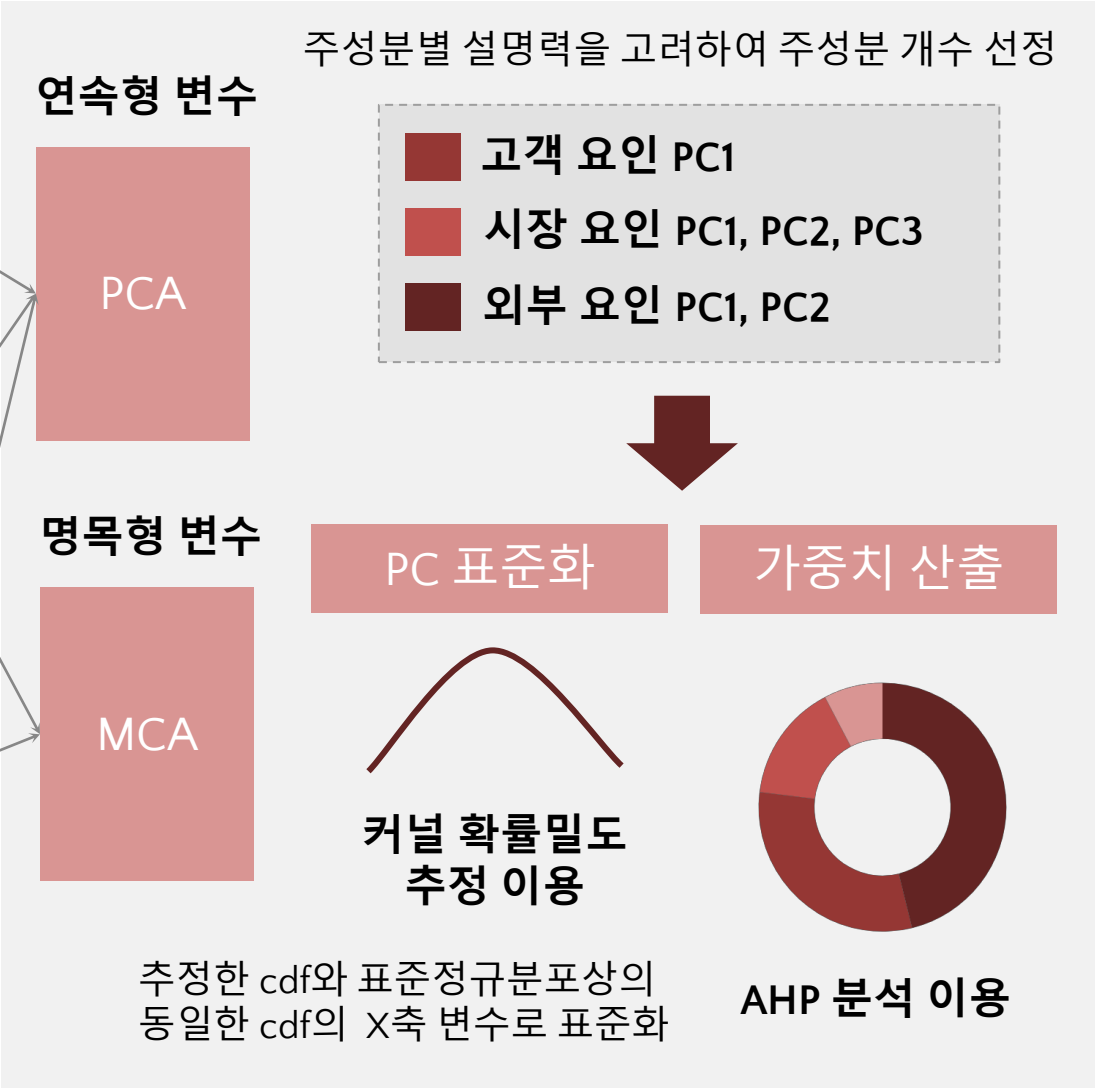
[Reference] 김홍관, 여성준(2007); 연상호, 김광현(2004)

### 3. Flow Chart

#### 데이터 합성 및 파생변수 생성



#### 주성분 분석 및 지수 산출



#### DT 입지 선정





# 데이터 전처리

- 분석 데이터 설명
- 변수 설명
- 데이터 생성 과정
- 데이터 전처리 및 정제 과정

## 1. 분석 데이터 설명

영역	데이터명	시점	사용용도	출처
고객 요인	자격 시구신고 평균소득월액	20.05.31	전통시장 위치의 이용객 소득수준 파악	공공데이터포털 - 국민연금공단
	행정구역별(시군구별), 성별 인구수	22.10.15	전통시장 위치의 인구수 파악	국가통계포털
	주민등록인구기타현황(평균연령)	22년 9월	전통시장 위치의 평균연령 파악	행정안전부
시장 요인	전국 전통시장 표준데이터	22.07.06	전통시장의 주요 편의시설 및 농축수산물 취급 여부 파악	공공데이터포털
외부 요인	카테고리별 장소 정보	22년 10월	전통시장 반경 1km에 위치한 시설 정보 파악	kakao developers
	행정구역별(시군구별) 일일 평균 유동인구	22년 9월	전통시장 근방의 유동인구 파악	소상공인시장진흥공단

**활성화 지수 생성을 위해 전통시장 데이터에 시장 내부와 외부, 이용 고객의 특성을 독립변수로 설정**

2. 변수 설명

고객 요인

1. 활성화 관련 요소

지역 주민들의 경제적 수준 ↑  
주변 인구 ↑  
- 시장 활성화 ↑

2. 변수 생성

평균소득월액

인구 수

연령별 비율

시군구별  
데이터  
사용

0-29세 / 30-59세 / 60세 이상

시장 요인

1. 활성화 관련 요소

전통시장 편의시설 ↑  
점포수 ↑ 품목의 다양성 ↑  
- 시장 활성화 ↑

2. 변수 생성

개설 연도

점포 수

장보기서비스

배송서비스

편의시설

농축수산물 취급

화장실 / 주차장

외부 요인

1. 활성화 관련 요소

전통시장 주변의 상권 ↑  
유동 인구 ↑  
- 시장 활성화 ↑

2. 변수 생성

편의점 수

대형마트 여부

지하철역 여부

음식점 수

유동 인구

전통시장 기준  
1km 반경 내의  
시설 여부/개수  
계산

### 3. 데이터 생성 과정 - 시장 요인

시장명	화장실 보유여부	주차장 보유여부	취급품목	점포수
강남시장	Y	Y	농산물,축산물,가공식품,의류/신발,기타소매업	70
신사상가	Y	N	농산물,수산물,가공식품,의류/신발	90
영동전통시장	Y	Y	축산물,가공식품,의류/신발,기타소매업	110
청담삼익시장	Y	Y	농산물,축산물,가공식품,의류/신발,기타소매업	183
고분다리전통시장	Y	Y	농산물,축산물,수산물,가공식품,가정용품	106



model input을 위해 dummy variable 생성

시장명	화장실 보유여부	주차장 보유여부	농산물 취급여부	축산물 취급여부	수산물 취급여부	점포수
강남시장	1	1	1	1	0	70
신사상가	1	0	1	0	1	90
영동전통시장	1	1	0	1	0	110
청담삼익시장	1	1	1	1	0	183
고분다리전통시장	1	1	1	1	1	106

3. 데이터 생성 과정 - 외부 요인

	address_name	category_group_code	category_group_name	category_name	distance	id	phone	place_name	place_u
0	대전 동구 가양동 430-14	CS2	편의점	가정,생활 > 편의점 > CU	196	625127705	042-631-5474	CU 대전가양스타점	http://place.map.kakao.com/62512770
1	대전 동구 가양동 288-1	CS2	편의점	가정,생활 > 편의점 > CU	232	17882352	042-672-0321	CU 가양플라워점	http://place.map.kakao.com/1788235
2	대전 동구 가양동 428-42	CS2	편의점	가정,생활 > 편의점 > GS25	254	1085514499	1644-5425	GS25 가양큰별점	http://place.map.kakao.com/108551449
3	대전 동구 가양동 438-26	CS2	편의점	가정,생활 > 편의점 > CU	285	74069596		CU 대전가양점	http://place.map.kakao.com/7406959



주변 시설 개수 및 존재 여부 파악하여  
새로운 데이터 생성

시장명	편의점 수	대형마트 여부	지하철역 여부	카페 수	음식점 수	학교 수	유동인구
가양시장	65	N	N	173	921	8	1267
강남시장	70	Y	Y	231	1177	8	9654
고분다리전통시장	142	N	Y	362	2294	11	7502
보성시장	74	Y	N	221	1436	7	1298
인제시장	12	Y	N	26	189	3	2873

## 4. 데이터 전처리 및 정제 과정

### 데이터 표준화

시장	평균소득월액	인구수
강남시장	1670051	530103
세종대평시장	1348116	381925
마천전통시장	1197359	37962
상계중앙시장	1381690	505503



시장	평균소득월액	인구수
강남시장	4.847622	0.7214982
세종대평시장	0.57644536	0.19564896
마천전통시장	-1.4236782	-1.02499568
상계중앙시장	1.02187839	0.63419852

데이터의 스케일에 따라 주성분의 설명 가능한 분산량이 달라질 수 있기 때문에 Standard Scaling 진행

### 결측치 처리

시장명	점포수
금곡시장	0
예산시장	0
...	...
회진시장	0



시장명	점포수
금곡시장	21
예산시장	400
...	...
회진시장	40

점포수가 누락된 시장 (총 8곳)에 데이터를 입력하여 결측치 보간

## 4. 데이터 전처리 및 정제 과정

### 결측치 처리

시장명	배송서비스	장보기서비스
흑석시장	0	0
에덴시장	0	0
...	...	...
동부시장	0	0



시장명	배송서비스	장보기서비스
흑석시장	N	N
에덴시장	N	N
...	...	...
동부시장	N	N

data merge 과정에서  
누락된 일부 변수를  
결측치 최빈값으로 대체

시장명	주소
금남시장	신축 공사중으로 도로 명주소 부존재
만수시장	인천광역시 강동구 만 수로27번길
...	...



시장명	주소
금곡시장	서울특별시 성동구 금 호산2길
만수시장	인천광역시 남동구 만 수로27번길
...	...

주소 표기 오류 수정

# 분석 결과

- 모델 설명
- 고객/시장/외부 요인 분석 결과
- 요인별 표준화
- 가중치 산출
- 활성화지수 산출 결과
- 활성화지수 유의성 검정



# FAMD : Factor Analysis of Mixed Data

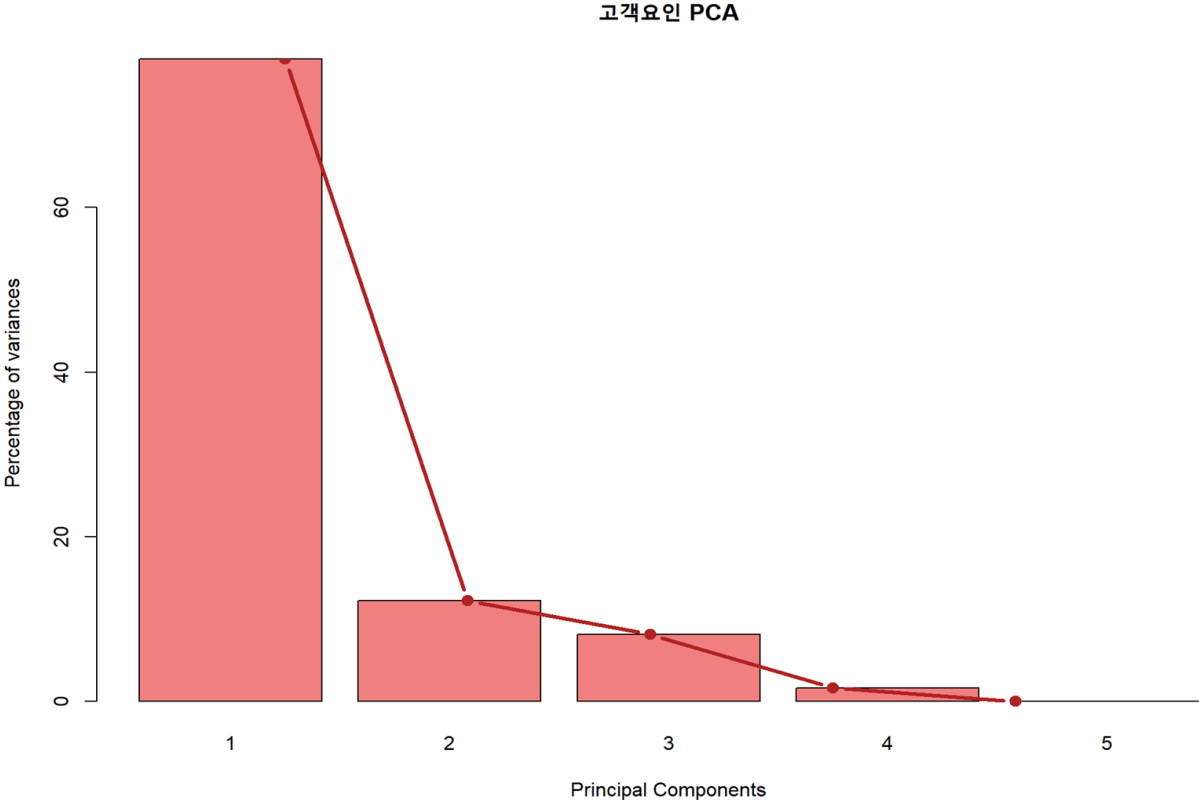
- 수치(numerical) 변수와 범주형(categorical) 변수를 모두 포함하는 다차원의 데이터(mixed data)를 효율적으로 저차원의 데이터로 요약하는 방법
- 각 저차원의 값은 하나의 지수 값이 되며, 각 지수를 나타내는 요소 간의 가중치를 도출할 수 있다.

## FAMD 분석법 활용 이유

- 본 연구는 전통시장 활성화 지수(Index)를 만드는 것을 목적으로 한다.
- 데이터에 범주형(categorical) 변수가 포함된 경우, PCA를 사용할 수 없으므로 본 방법을 사용하여 차원을 축소시켜야 한다.
- 주성분(Principal Component)은 각 지수를 나타내며, 이는 전통시장 별 지수 값을 각 주성분에 따라 도출해낼 수 있다.

## 2. 고객 요인 분석 결과

[ 주성분별 분산비율 ]



[ 고객요인 요소별 가중치 ]

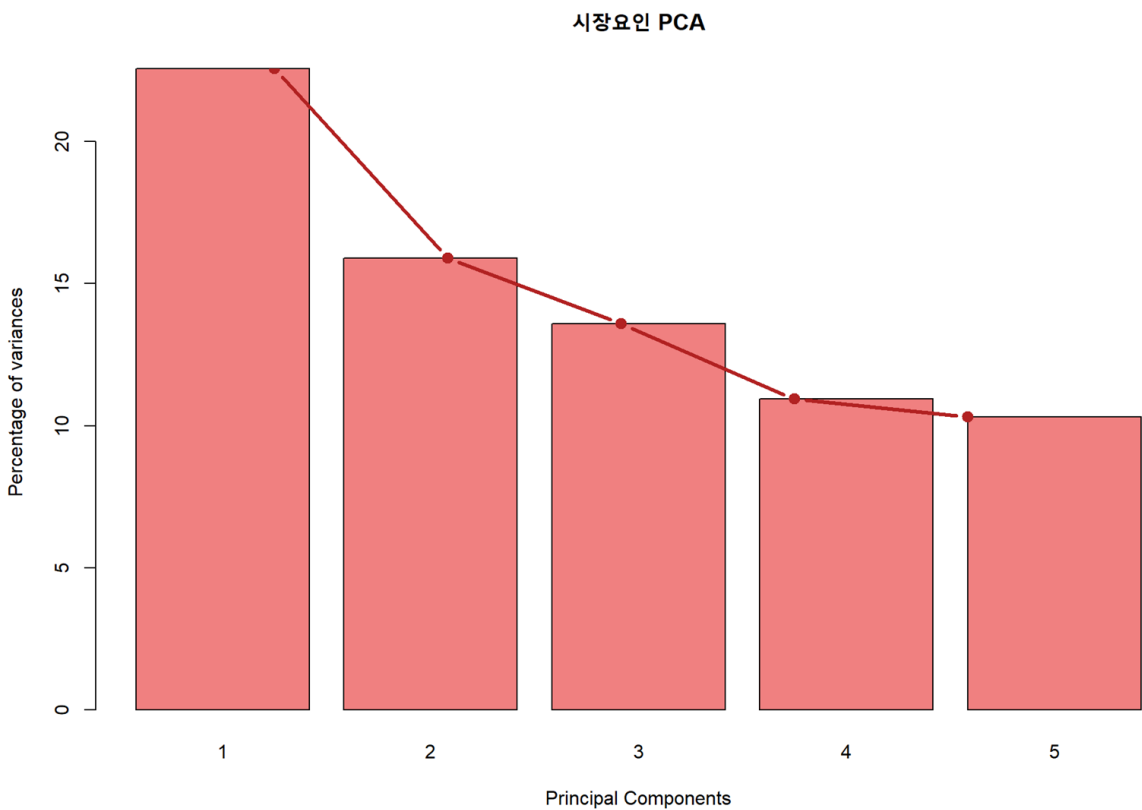
평균소득월액	1번 주성분	2번 주성분
시구신고 평균소득월액	0.751	-0.536
인구수	0.738	0.565
0-29세	0.944	0.035
30-59세	0.970	-0.044
60세 이상	-0.98	0.005



1번 주성분의 분산비율이 약 78%이므로 1번 주성분만 사용하기로 결정

### 3. 시장 요인 분석 결과

[ 주성분별 분산비율 ]



[ 시장요인 요소별 가중치 ]

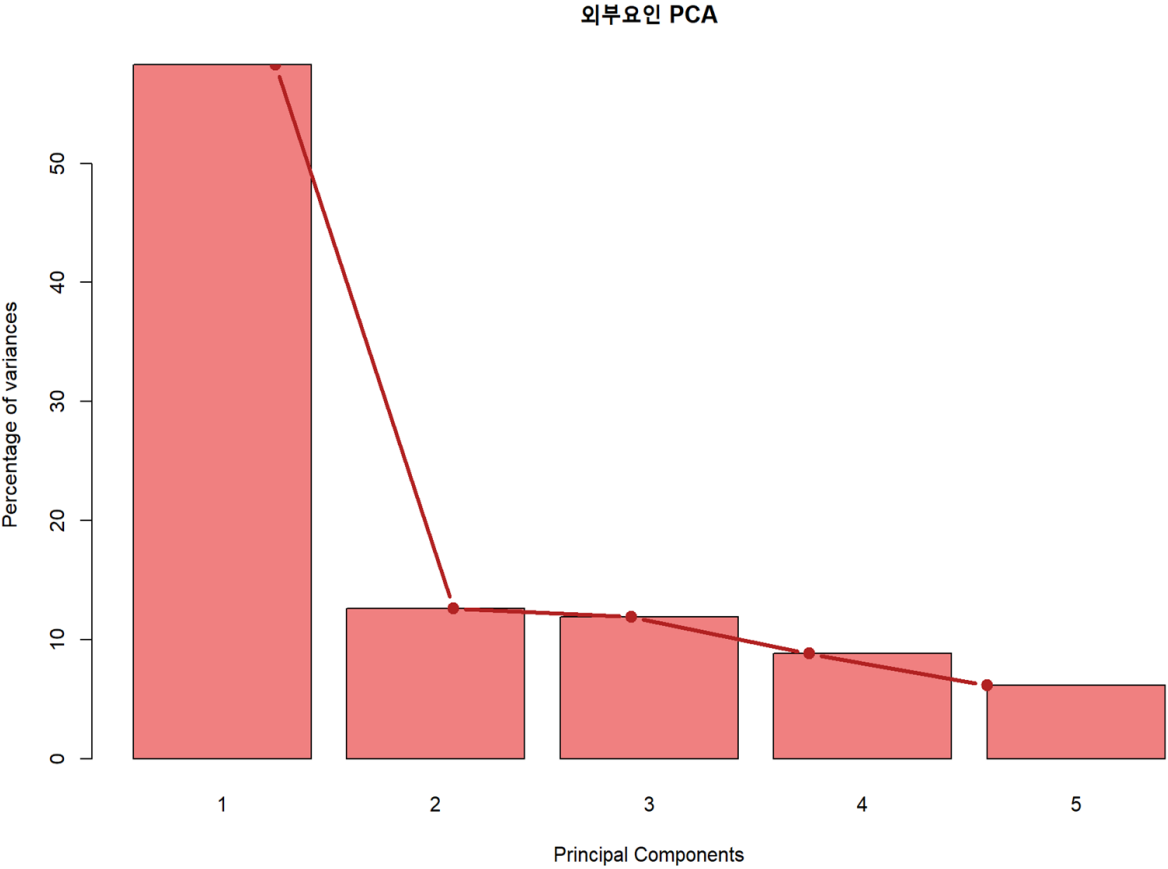
	1번 주성분	2번 주성분	3번 주성분
화장실 보유 여부	0.0111	0.3245	0.2841
주차장 보유 여부	0.0011	0.3801	0.0940
농산물 취급 여부	0.5626	0.0197	0.0849
축산물 취급 여부	0.5261	0.0348	0.0005
수산물 취급 여부	-0.091	0.0011	0.0740
점포수	0.0837	0.0340	0.0547
개설연도	0.0286	0.0290	0.2518
배송서비스 여부	0.1966	0.2793	0.1910
장보기서비스 여부	0.1197	0.3226	0.1881



1번, 2번, 3번 주성분의 분산 비율을 합하면 약 52%이므로  
3개의 주성분을 사용하기로 결정

## 4. 외부 요인 분석 결과

[ 주성분별 분산비율 ]



[ 외부요인 요소별 가중치 ]

	1번 주성분	2번 주성분
편의점 수	0.899	0.0003
대형마트 여부	0.803	0.0007
지하철역 여부	0.880	0.0010
카페 수	0.565	0.0217
음식점 수	0.166	0.7102
학교 수	0.204	0.1499
유동인구	0.562	0.0007



1번, 2번 주성분의 분산 비율을 합하면 약 71%이므로 2개의 주성분을 사용하기로 결정

## 5. 요인별 표준화

### 기여율 (explained variance ratio)

기여율 = 특정 주성분의 분산에 대한 비율  
특정 주성분 분산

=  
모든 주성분 분산의 합

따라서 2개 이상의 주성분이 선택된 경우,  
해당 주성분 기여율을 **가중치**<sup>1)</sup>로 사용하여  
하나의 값으로 재설정

### 표준화

각 지수의 스케일이 모두 다름  
→ 각 지수의 순수한 영향력 판단 불가

따라서 주성분별로 **커널 밀도 함수**를 추정하고,  
추정한 cdf와 표준정규분포상의 동일한 cdf의  
X축 변수로 표준화

고객, 시장, 외부 요인별 최종 값

$$= \sum_{i=1}^p (weight_i + variable_i)$$

$p$ 는 주성분 수,  $weight_i$ 는 표준화된 기여도,  $variable_i$ 는 표준화된 변수

1) 선택된 주성분들의 기여율 합이 1 미만이므로, 각 기여율을 기여율의 합으로 나누어 표준화

5. 요인별 표준화

커널 밀도 추정(Kernel Density Estimation)

“ 확률분포형에 대한 가정없이 비모수적인 확률분포를 추정하기 때문에  
자료에 가장 가까운 확률분포를 찾을 수 있음 ”

hyperparameter set

KDE는 hyperparameter 설정에 따라 결과가 매우 상이함

↓

분포 함수

Gaussian Kernel

bandwidth

GridSearchCV

데이터의 수가 많지 않으므로 LOOCV(Leave-One-Out Cross Validation) 이용하여 추정

KDE 추정 적절성 평가

KS test (Kolmogorov-Smirnov test)  
: 두 누적 확률 분포간의 차이를 거리로 산정한 통계량

	고객	시장	외부
PC1	0.482	0.171	0.663
PC2		0.947	0.204
PC3		0.358	

유의수준 0.05 하에서 추정된 cdf의 적절성 확인

cdf 추정 후 표준화

추정한 cdf

표준정규분포 cdf

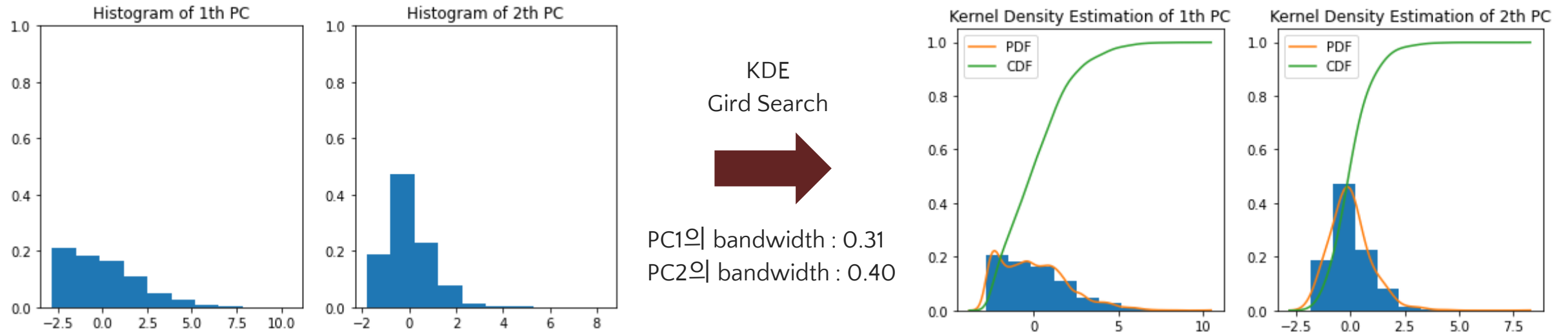


기존값

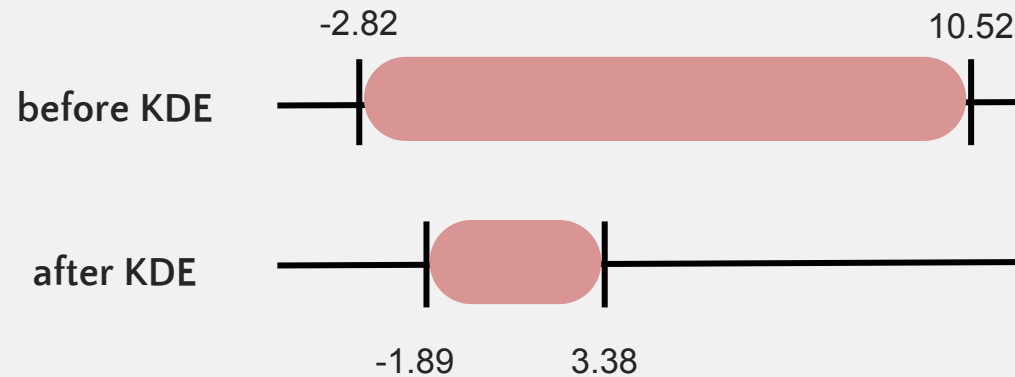
대체값

기존값의 추정된 cdf와 동일한 표준정규분포상 cdf의 x축 변수로 표준화

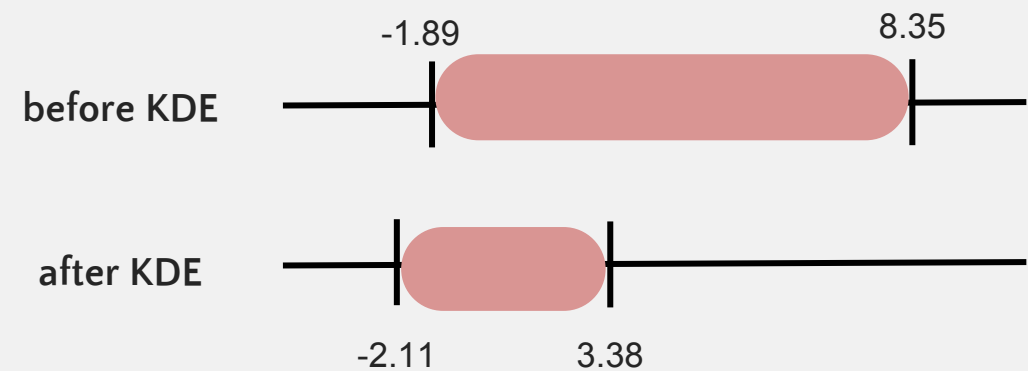
## 5. 요인별 표준화 : 외부요인 예시



[ PC1의 분포 변화 ]



[ PC2의 분포 변화 ]



## 6. 가중치 산출

- 30 명의 시장 이용자에게 요인별 상대 중요도 설문 조사 시행
- 설문 조사 결과를 바탕으로 AHP 분석을 통해 요인별 가중치 생성

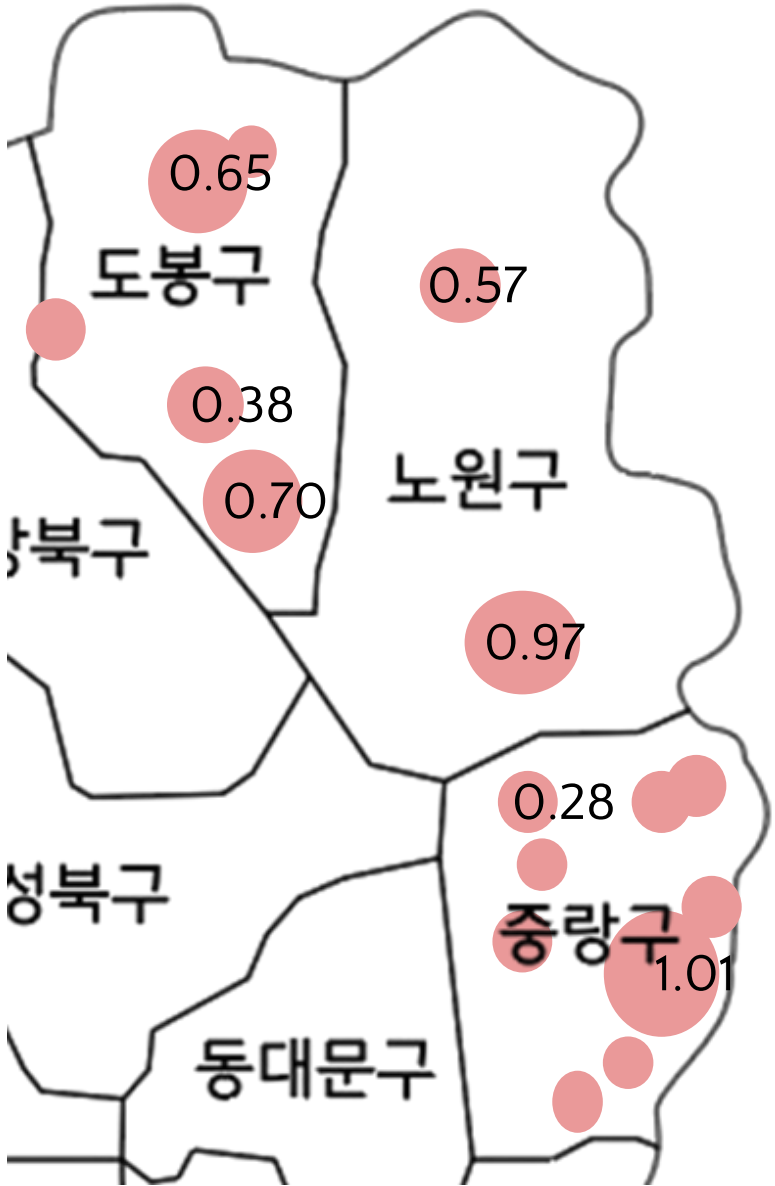
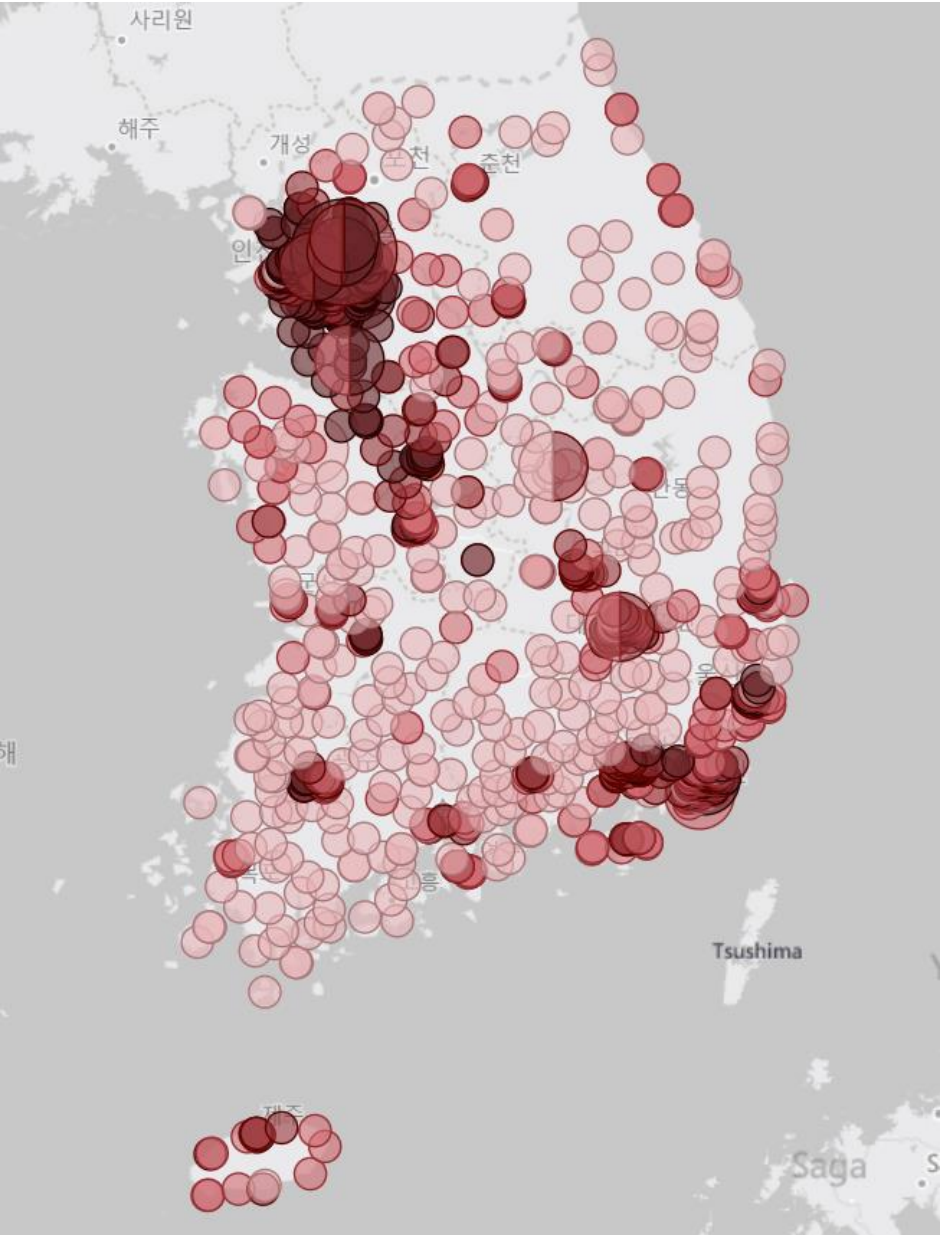
질문	매우 중요하지 않음 (1/4점)	중요하지 않음 (1/3점)	약간 중요하지 않음 (1/2점)	보통 (1점)	약간 중요 (2점)	중요 (3점)	매우 중요 (4점)
고객 요인이 시장 요인보다 전통시장 활성화에 더 영향이 있다고 생각하시나요?					V		
시장 요인이 외부 요인보다 전통시장 활성화에 더 영향이 있다고 생각하시나요?						V	
외부 요인이 고객 요인보다 전통시장 활성화에 더 영향이 있다고 생각하시나요?					V		



	고객 요인	시장 요인	외부 요인
Weight	0.3297	0.3271	0.3432



## 7. 활성화지수 산출 결과

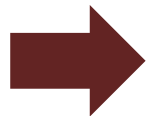


## 8. 활성화지수 유의성 검정

매출 / 점포수 와 활성화지수를 비교하여 지수가 유의한지 검정

모델 가정 :  $\log(\text{매출} / \text{점포수}) = \beta * \text{활성화지수}$

OLS Regression Results						
=====						
Dep. Variable:	index	R-squared (uncentered):	0.736			
Model:	OLS	Adj. R-squared (uncentered):	0.735			
Method:	Least Squares	F-statistic:	499.3			
Date:	Tue, 29 Nov 2022	Prob (F-statistic):	1.15e-53			
Time:	19:37:50	Log-Likelihood:	-64.794			
No. Observations:	180	AIC:	131.6			
Df Residuals:	179	BIC:	134.8			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
sales/n	0.0319	0.001	22.344	0.000	0.029	0.035
=====						



$\beta$ 가 양수이고 p-value가 매우 작으므로  
매출/점포수와 활성화지수는 양의 상관관계를 가진다고 볼 수 있음

# 활용방안 및 시사점

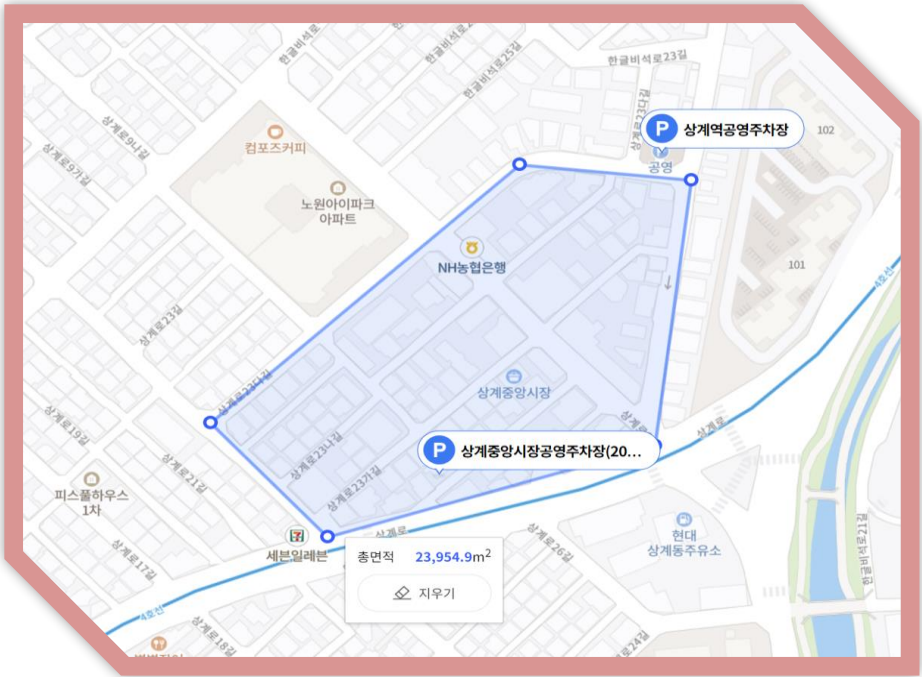
활성화지수 기반 DT 입지선정  
향후 활용방안  
본 과업의 한계점

## 1. DT 입지선정

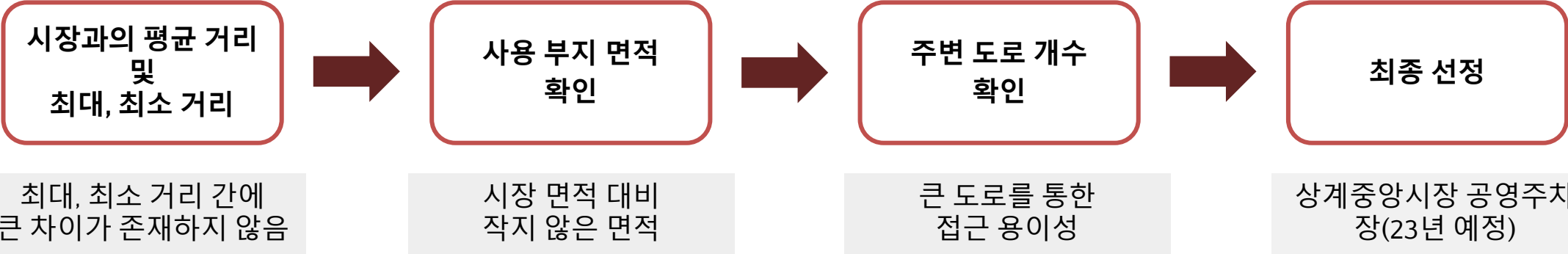


활성화 지수가 낮은 상계중앙시장을 DT 적용 시장으로 선정

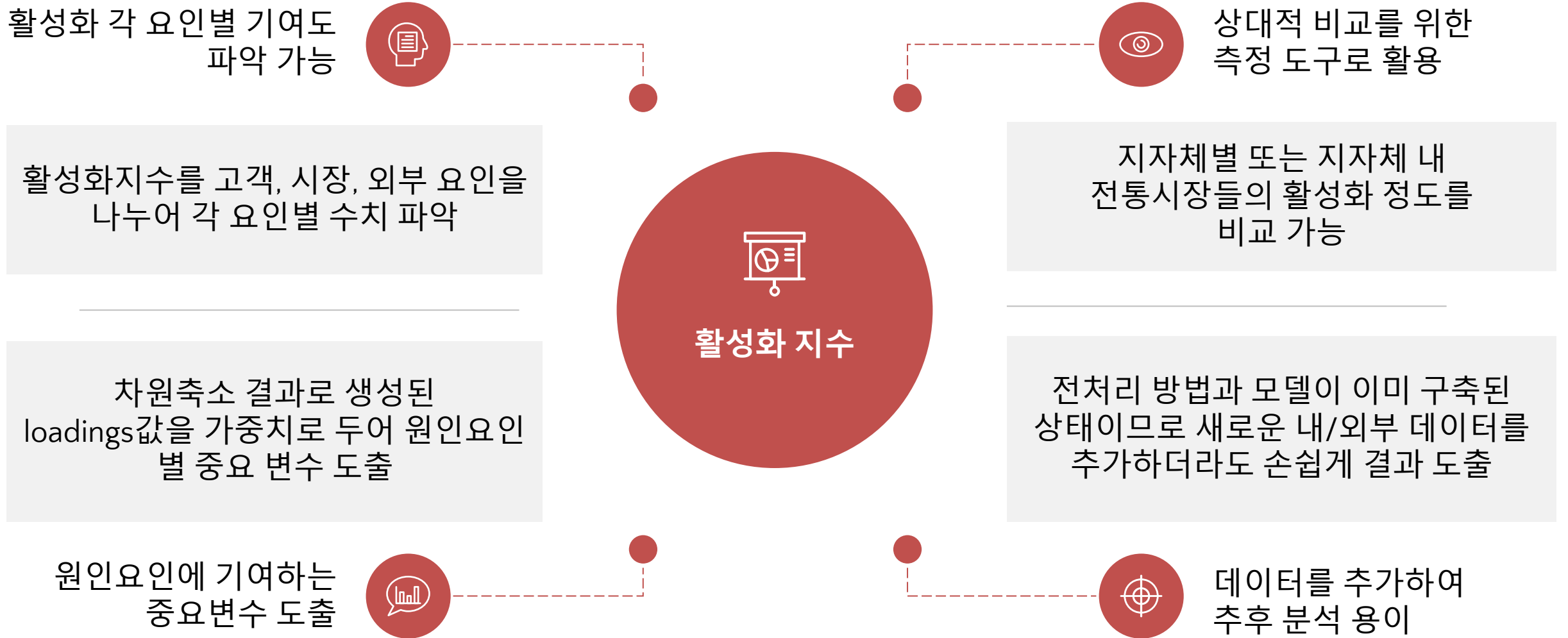
1. DT 입지선정



시장과의 평균거리 및 최대, 최소 거리	118.5m / 261.8 m, 64.2 m
부지 면적	약 200 m²
주변 대로변/소로변 개수	3개 / 4개



## 2. 향후 활용방안



### 3. 본 과업의 한계점



#### 데이터의 부족

- 시장별 데이터의 부족으로 행정구역별 데이터(외부요인 - 유동인구)를 시장별 데이터로 사용
- 시계열적으로 지수를 도출하지 못하여 시간의 흐름에 따른 비교 불가능



#### 수집한 데이터의 문제점

- 각 데이터 별 수집 시점이 다르기에 데이터 간의 관계를 불완전하게 반영
- 불완전한 설문지 응답자의 존재 가능성으로 인해 활성화 지수 요인별 가중치에 대한 신뢰도의 하락

### 참고 문헌

박상훈·이희정, 2018, “소셜 빅데이터를 이용한 전통시장 활성화 요인 도출 연구-서울시 망원시장과 수유시장을 대상으로-”, 「서울도시연구」 제 19권 제3호 논문 pp.1-18. 서울연구원

임정빈·김덕준, 2017, “전통시장 활성화에 미치는 영향요인에 관한 연구”, 「정책개발연구」 제 17권 제2호 pp. 1-28, 한국정책개발학회.

김홍관·여성준, 2007, “GIS를 활용한 부산시 도심재생에 관한 연구”, 「한국지리정보학회지」 제 10권 제1호, pp. 205-217. 한국지리정보학회.

연상호·김광현, 2004, “중소도시 상권변화 조사 및 분석을 위한 공간정보의 활용”, 「한국콘텐츠학회 추계종합학술대회 논문집」, 제2권 제2호, pp.165-169, 한국콘텐츠학회.

박주영, 2009, “재래시장의 상품구성이 활성화에 미치는 영향에 대한 분석과 재래시장 정책에 대한 시사점”, 「중소기업연구」, 31권2호, pp 137-155, 한국중소기업학회.

장정식·김홍순, 2019, “서울시 편의점의 입지요인 분석”, 「국토지리학회지」 제53권 3호, pp 283-293, 국토지리학회

김민정·류성민, 2016, “전통시장의 활성화 방안에 대한 정성적 연구”, 「유통연구」 제 21권 제4호, pp.157-177, 한국유통학회.

### 분석 도구





**감사합니다**

---