



# 서울 소재 전통시장 군집화를 통한 전통시장 활성화 방안 모색

팀원: 김재경, 맹준영, 신유승



# CONTENTS

01 개요

02 데이터 구축

03 군집화

04 군집 해석

05 정리

# 01 개요



- **서울 소재 전통시장 371개의 군집화**

- 정책을 특정 전통시장을 중심으로 수행해야할 경우에, 비슷한 성격의 다른 전통시장을 고려 가능
- 같은 군집내에서 일어난 이슈 혹은 사건은 비슷한 군집내의 다른 전통시장에서도 해당 이슈가 일어날 가능성이 높음
- 전통시장의 군집화를 통해, 전통시장의 종류를 세분화할 수 있음



# 01 개요

## 서울소재 전통시장



- 서울소재 전통시장의 경우, 서울 중심부에 몰려있는 경향을 보임
- 5일장, 10일장, 일장 등 다양한 형태의 전통시장이 존재함

02

## 데이터 구축

## 02 데이터 구축

### 데이터 출처 : 금융빅데이터 플랫폼



- BC카드의 금융빅데이터 플랫폼 내에서 데이터를 찾아보았으나, 해당 분석에 필요한 데이터는 적었음
  - 상권 지역별 업종별 데이터가 유용할거라 생각했으나 서울이 아닌 부산 대상
  - 사용자 이용데이터 기반의 상권분류 데이터는 해당 과제의 목표라고 할 수 있음



맛집

활용:외식업 상권 분석

#### 사용자 이용데이터 기반의 상권 분류

상권 분석 데이터

지역별 사용자의 활동 기록을 기반으로 축정한 데이터

최종등록일자 :2022-06-07 조회수 : 3,010 다운로드수 : 66



zip(CSV)

여기어때컴퍼니

3,300,000원



유동인구

상권 지역별 업종별 데이터

#### 상권 지역별 업종별 데이터

상권 지역별 업종별 데이터

상권 지역별 업종별 데이터

최종등록일자 :2021-12-21 조회수 : 1,089 다운로드수 : 82

카드

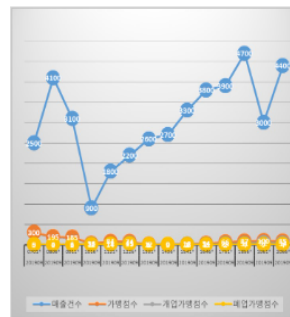
창업 준비

#### 상권소비데이터

상권소비데이터

지역별 가맹점 기반 상권데이터 및 지역, 업종, 시간, 고객 속성별 매출 정보

최종등록일자 :2022-10-21 조회수 : 586 다운로드수 : 3



카드

활용:창업 준비

#### 지역별 업종별 매출상세 내역 및 유효/개·폐업 가맹점 수

상권집계

전국 지역별(우편번호) 업종별 매출상세 내역, 유효/개·폐업 가맹점수

최종등록일자 :2022-11-20 조회수 : 3,485 다운로드수 : 119



## • 서울 특별시 홈페이지 소상공인지원 메뉴

- 해당 전통시장의 면적, 점포수, 형태(건물형, 골목형, 지하도상가) 위치에 대한 정보
- 2022년 9월 기준의 서울 전통시장은 371개

서울특별시 서울소식 시민참여 분야별정보 서울소개 부서안내 정보공개 응답소 로그인 나의서울 전자우편

**경제** 검색어를 입력해 주세요

새소식 정책소개 서울경제 일자리 투자·창업 도농·지역상생 거점성장 노동정책 소상공인지원 사회적경제 소상공인 디지털 경제 자료실

**소상공인지원**

공경경제 정책  
공경경제 사업  
소기업·소상공인 종합지원  
**전통시장 조성**  
전통시장 정책  
전통시장 행사안내  
**서울시 전통시장 소개**  
공경경제 서비스  
서울중대시장

**서울시 전통시장 소개**  
경제 > 소상공인지원 > 전통시장 조성 > 서울시 전통시장 소개

**서울시 전통시장 현황**

2022.09.05

전화 02-2133-5483

노동공경상생정책관  
상권활성화담당관

정렬	자치구	시장명	형태	신주소 (구주소)	전화번호	건물형 (연면적)	점포수 (빈점포제외)
	종로구						8,874
1	종로구	광장시장	건물형	종로구 창경궁로 88 (종로구 예지동 6-1)	2272-0967	18,975	688
2	종로구	동대문종합시장	건물형	종로구 종로 266 (종로구 종로6가 289-42)	2262-0211	61,476	3,270
3	종로구	동대문종합시장 신관	건물형	종로구 종로272 (종로구 종로6가 289-57)	2262-0211	13,533	286
4	종로구	동대문종합시장D동상가	건물형	종로구 종로 266 (종로구 종로6가 262-1)	2279-8751~ 2	12,452	1,048
5	종로구	산설종합시장	건물형	종로구 난계로27길 51 (종로구 송인동 206-9)	2234-7151	1,691	141



01 개요

데이터 출처 : 스마트치안 빅데이터 플랫폼



- 서울 특별시 홈페이지 소상공인지원 메뉴
  - 해당 전통시장의 장형태(2/5/10/일장) 위도, 경도 등에 대한 정보 포함

스마트 치안 빅데이터 플랫폼

Smart Policing Big Data Platform

검색어를 입력해 주세요.

9 빅데이터

三

빅데이터마켓

분석/시각화

혁신서비스

커뮤니티

스마트치안

로그인

회원가입

빅데이터 상세

빅데이터마켓 > 데이터 검색

지역환경

2022-10-18 최초등록 / 2022-10-18 업데이트됨

전통시장

#전통시장 위치

#잠재 취약시설

유형 CSV

가격 무료

조회 37

다운로드 5

수집주기 매년

데이터 미리보기

샘플다운로드

EXCEL

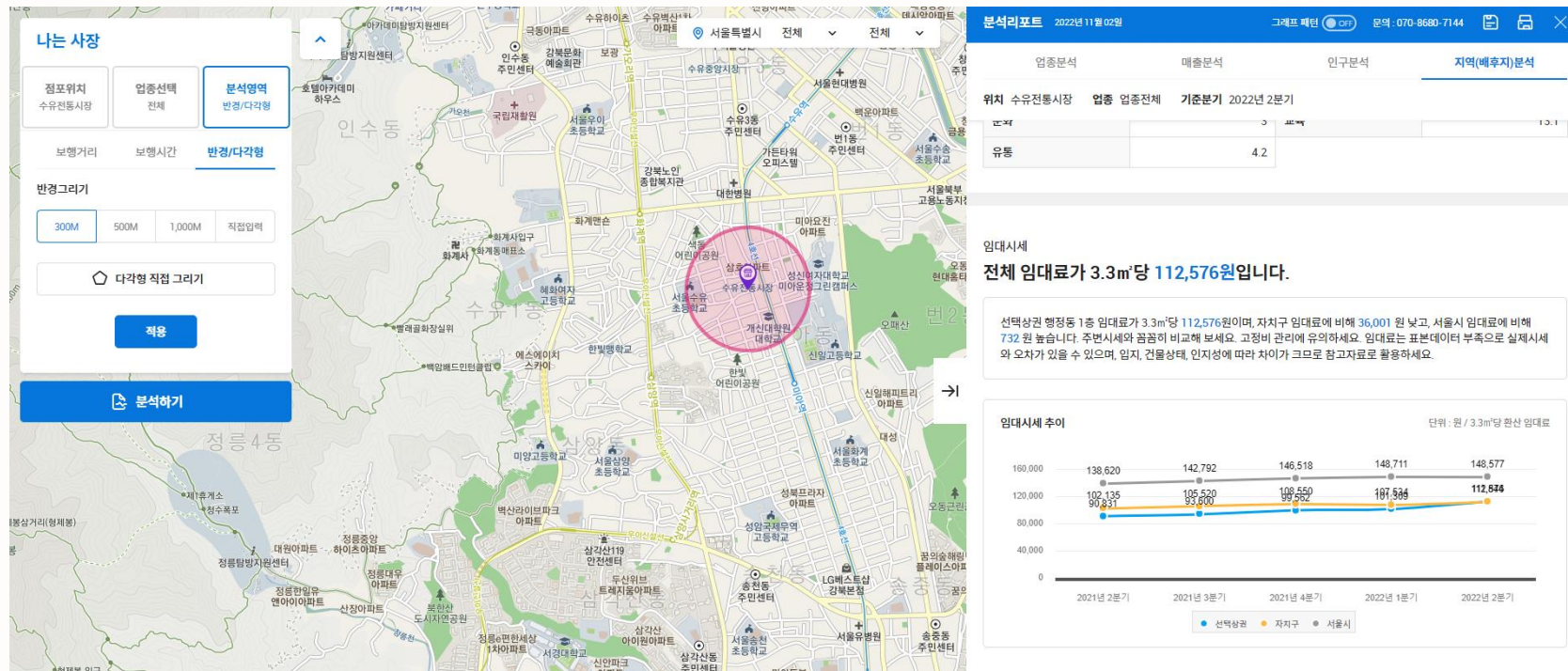
CSV

키ID	시군구	전통시장명	소재지주소	오픈일	전화번호	행정동명	법정동명	행정동코드	법정동코드	위도	경도
2063060	합천군	삼가시장	삼가면 일부5길 1	277일장	-	삼가면	일부리	3840043	4889043024	35.4128146125189	128.1222829606
2063059	합천군	초계시장	초계면 초계중앙로 61	5/10일장		초계면	초계리	3840036	4889036022	35.5593148749891	128.2668344987
2063058	합천군	(주)합천시장	합천읍 합천리 473-1	3/8일장		합천읍	합천리	3840011	4889025021	35.568614755381	128.1606640346

## 01 개요

# 데이터 출처 : 우리마을가게 상권분석 서비스

- 가장 많은 정보를 포함하고 있는 플랫폼은 우리마을가게 상권분석 서비스
  - 특정 지점을 중심으로 반경 300m 상권에 대한 종합적인 통계값 크롤링
  - 40여개 이상의 다양한 특징들을 추출





- 40개 이상의 특징값들 추출

컬럼명	샘플데이터
장소	광장시장
분기	2022년 2분기
면적	279,926m <sup>2</sup>
점포수	2056
신생기업생존률	0.69
평균영업기간	3.7
개업수	31
폐업수	21
증가하는업	외식업
점포당월평균매출액(만원)	2637
월평균매출건수	492
매출많은요일	화요일
남성비율	0.552
여성비율	0.448
일평균유동인구수	13109
주거인구수	1569
직장인구수	7362

컬럼명	샘플데이터
가구세대수	859
부동산평수대	66m <sup>2</sup> 대
부동산가격대	5억대
관공서수	4
유통점수	1
금융기관수	7
영화관수	0
병원수	16
숙박수	2
학교수	0
교통시설수	12
소득분위수	8
행정동소득분위수	7
소비트렌드	음식
임대료수준	259160
latitude	37.5700398
longitude	126.9996036

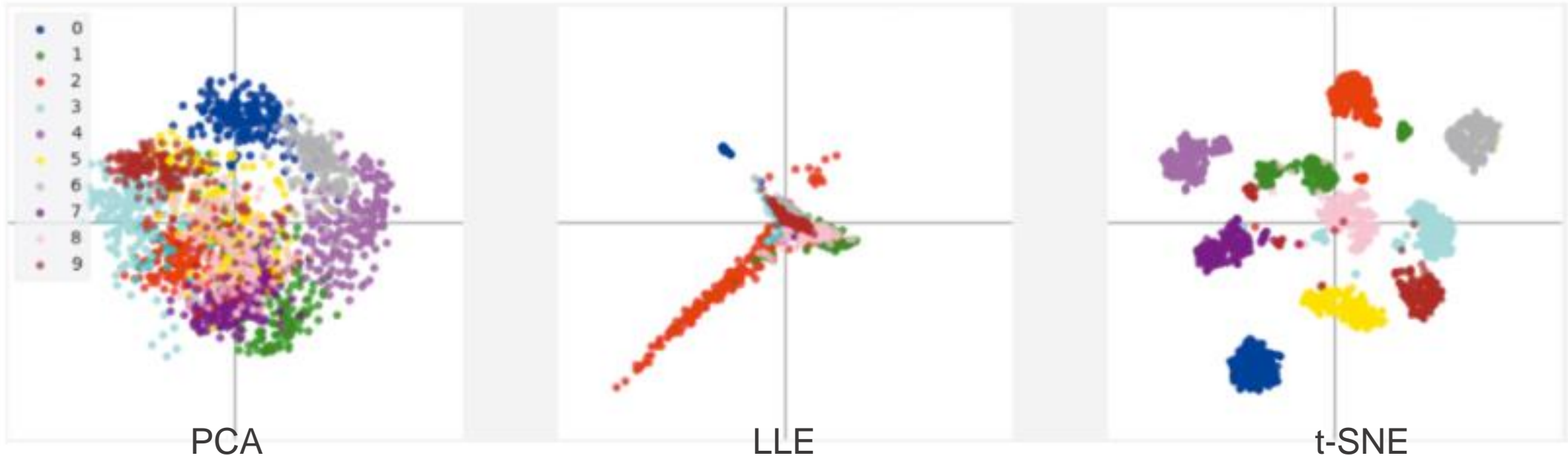
03

군집화



- 군집화를 하기에 앞서 다양한 방법으로 차원을 축소 (의미없는 축의 삭제)

- PCA, LLE, t-SNE 등의 방법을 사용하여, 차원을 다차원에서 2차원으로 축소할 수 있음
- 일반적으로 t-SNE의 차원축소 성능이 가장 좋다고 알려짐



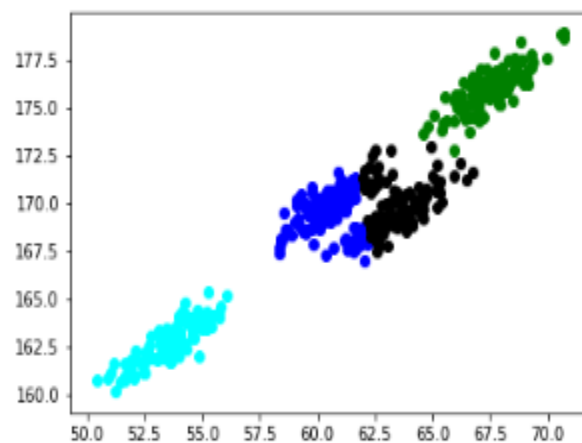
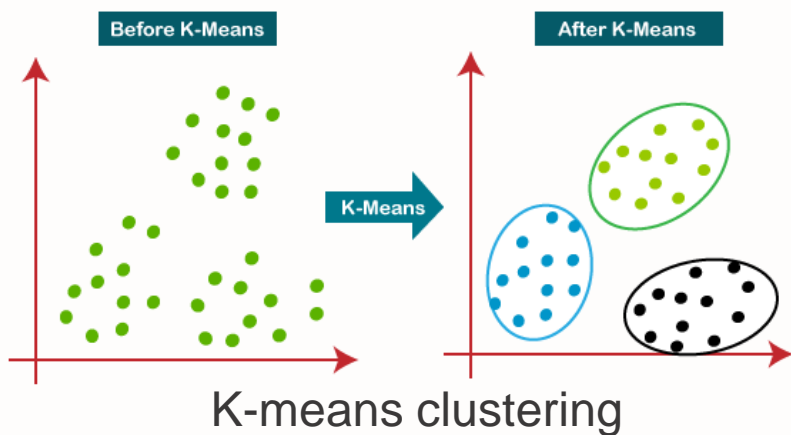
**주성분 분석(PCA, Principal Component Analysis)** : 가장 대표적인 차원 축소 알고리즘이다. PCA는 먼저 데이터에 가장 가까운 초평면(hyperplane)을 구한 다음, 데이터를 이 초평면에 투영(projection)시킨다.

**t-SNE(t-distributed stochastic neighbor embedding)** : 높은 차원의 복잡한 데이터를 2차원에 차원 축소하는 방법

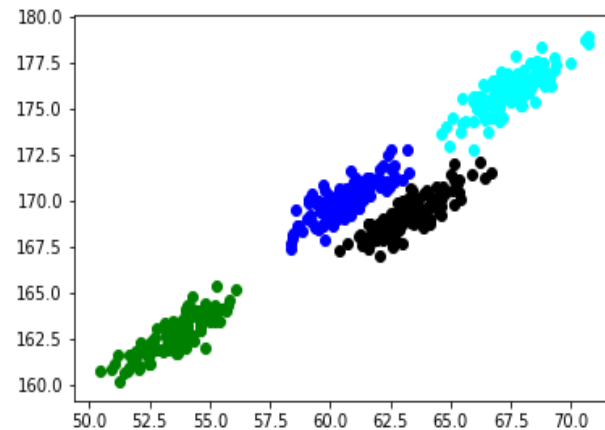


- 다양한 방법으로 군집화

- 최적의 군집 개수를 파악하는 알고리즘 수행
- K-means clustering, GMM(Gaussian Mixture Model), DBSCAN 등의 방법을 활용해 군집화 수행
- K-means clustering의 한계를 극복한 모델이 GMM
- 각 군집에 대한 특징 파악



K-means



GMM

**K-means clustering** : 주어진 데이터를 k개의 클러스터로 묶는 알고리즘으로, 각 클러스터와 거리 차이의 분산을 최소화하는 방식

**GMM** : Parametric 모델로서 EM(Expectation Maximization) 알고리즘을 이용한 대표적인 군집화 모델이다. 개별 데이터가 가우시안 분포에 속한다고 가정을 한 상태에서 특정 정규분포에 속할 확률을 추정하는 것

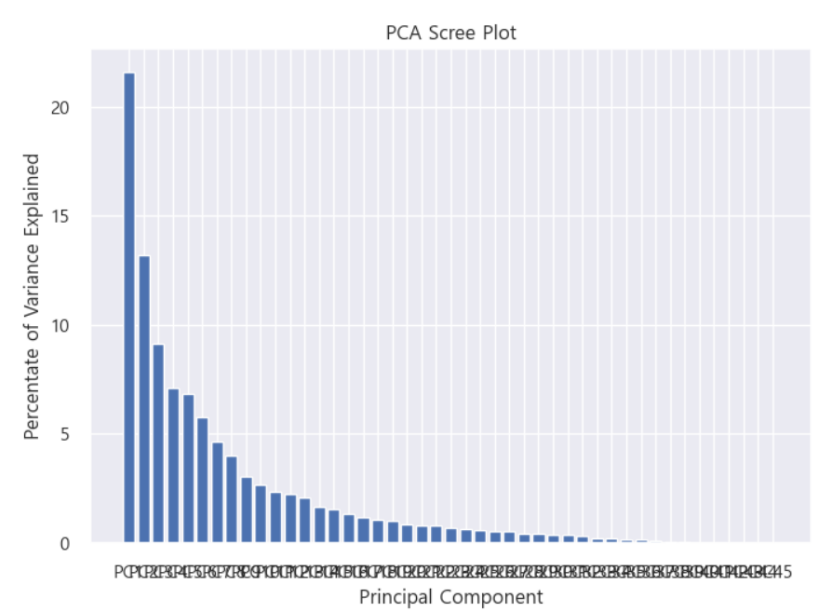
**DBSCAN(Density-based spatial clustering of applications with noise)** : 밀도 기반의 클러스터링은 점이 세밀하게 몰려 있어서 밀도가 높은 부분을 클러스터링 하는 방식



- 주성분 개수 설정
  - 최적의 군집 개수를 파악하는 알고리즘 수행
  - Screeplot을 통해 16개의 주성분을 사용시 전체 데이터의 90%를 표현할 수 있음
  - 주성분의 개수를 16개로 설정



주성분 개수2개의 Plot



Scree plot

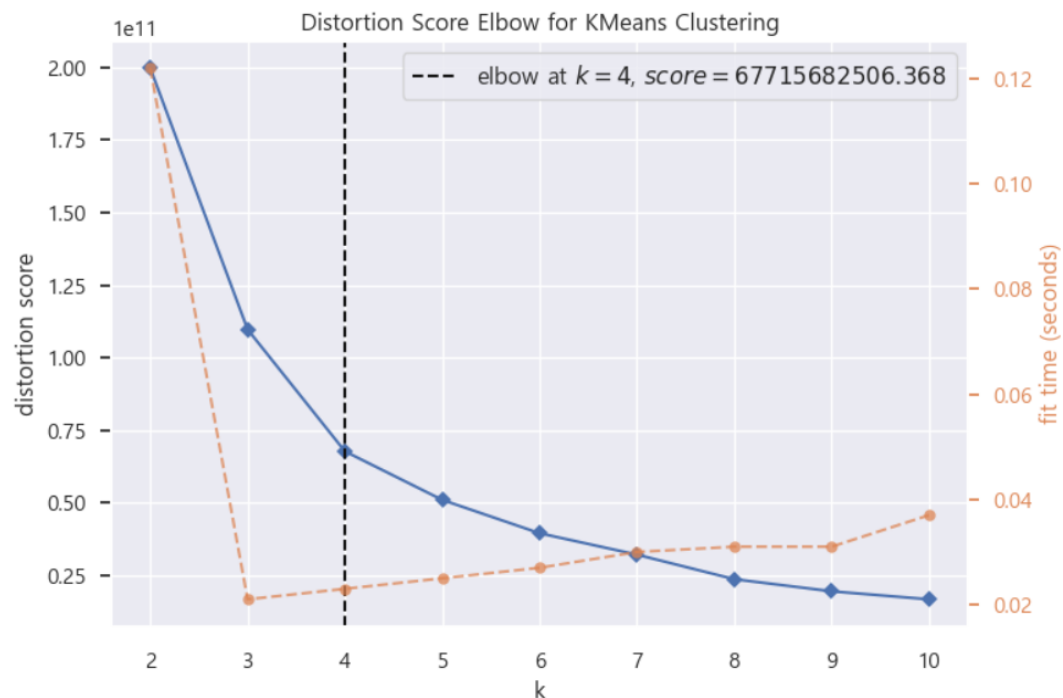
주성분수	cumulative variance	주성분수	Cumulative variance
0	0.2160	11	0.8243
1	0.3479	12	0.8447
2	0.4390	13	0.8610
3	0.5098	14	0.8764
4	0.5781	15	0.8895
5	0.6359	16	0.9010
6	0.6824	17	0.9114
7	0.7225	18	0.9213
8	0.7527	19	0.9296
9	0.7792	20	0.9373
10	0.8024	21	0.9448

Percentate of Variance Explained

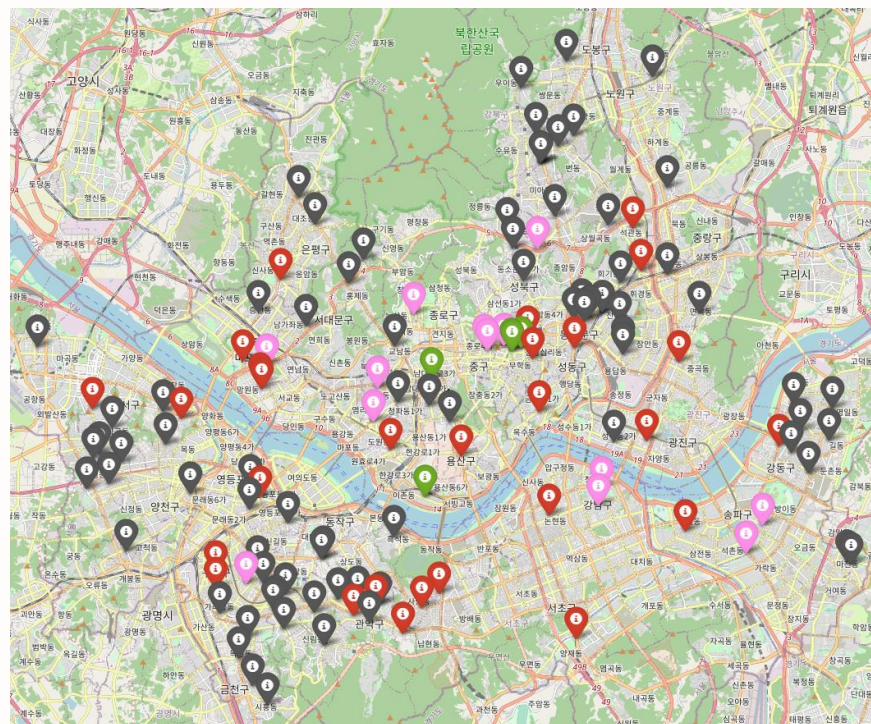




- Kmeans를 통한 군집화
  - 최적의 군집 개수를 파악하는 알고리즘 수행
  - Elbow 알고리즘을 사용해서 최적의 K=4로 결정
  - 총 4개의 군집으로 구별하여 시각화



Elbow 알고리즘



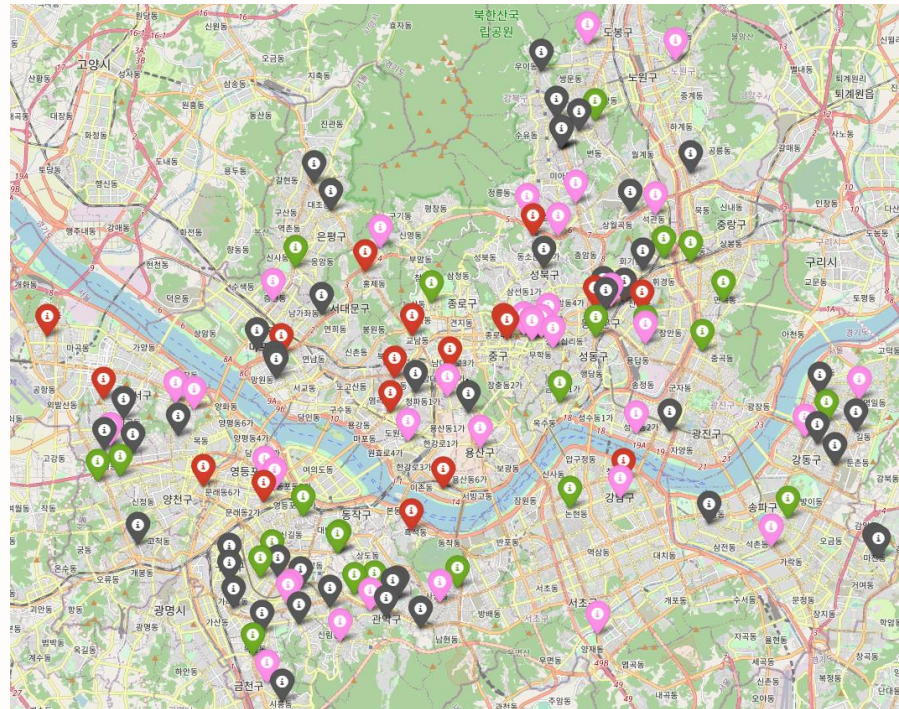
군집결과 시각화





- **GMM을 통한 군집화**

- 앞에서 구한 주성분 개수(16)과 Elbow 알고리즘으로 구한 최적의 군집개수(4개) 적용
- K-means clustering의 한계를 극복한 모델이 GMM
- 총 4개의 군집으로 구별하여 시각화

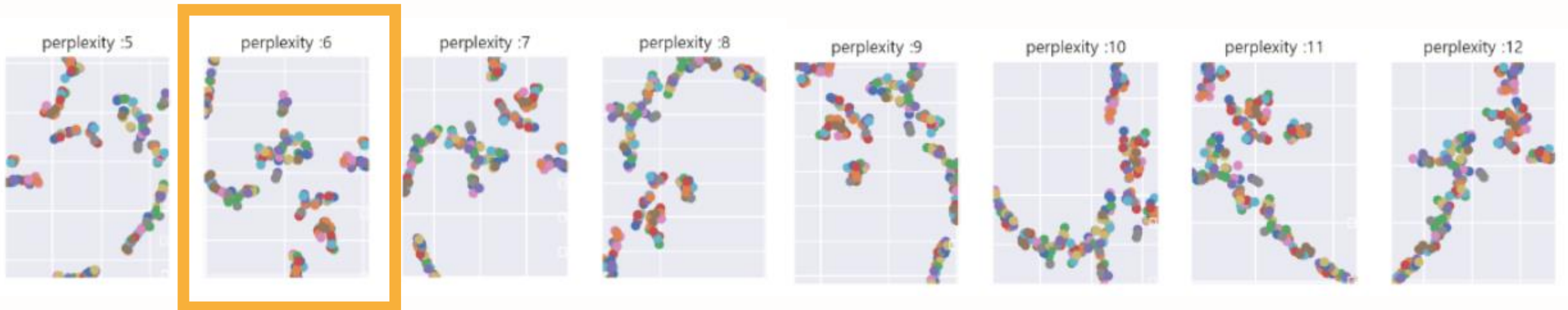


군집결과 시각화



- **t-SNE를 통한 차원축소 후 DBSCAN을 통한 군집화**

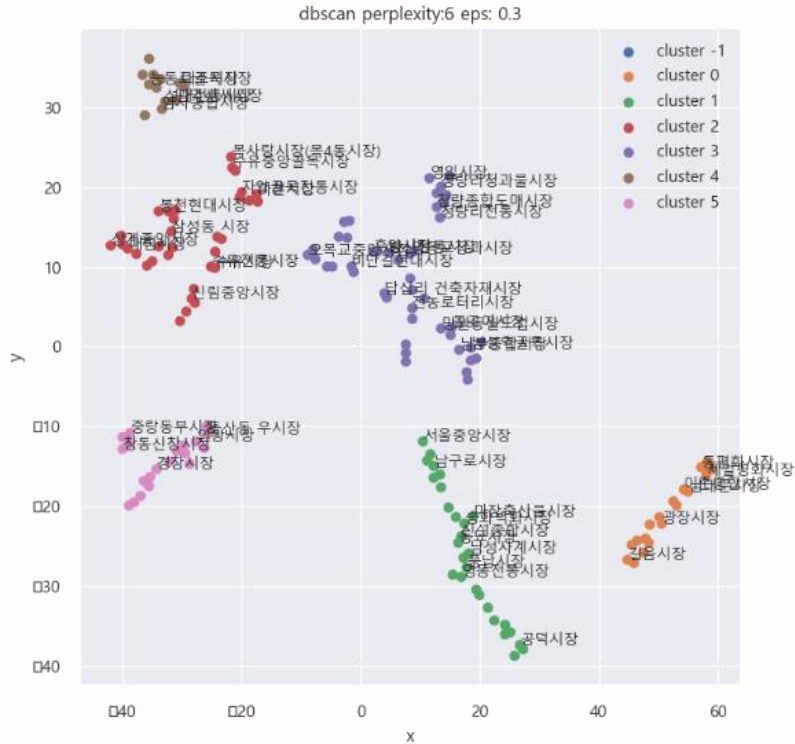
- t-SNE를 통해 2차원으로 표현
- t-SNE의 Perplexity 파라미터값에 따라 2차원축소의 모양 달라짐, 6으로 설정
- DBSCAN은 군집의 수를 설정하지 않고, 최대 탐색 거리를 의미하는 Epsilon(EPS), 클러스터로 구성되기 위한 최소 샘플 개수인 Min\_samples(2로 설정)을 하이퍼파라미터로 사용
- EPS이 너무 낮으면 비교적 가까운 점들도 같은 군집으로 형성되기 어려웠고, EPS가 높은 경우는 대부분의 점들이 모두 같은 군집으로 포함.





- 군집화 성능평가 방법 : 실루엣 지수(Silhouette index)
  - DUNN, 실루엣 지수, Rand 지수, AMI (Adjusted mutual information) 등의 다양한 방법들 중 실루엣 지수 선택
  - T-SNE의 perplexity =6, DBSCAN의 epsilon =0.3일때 성능이 가장 좋았음

클러스터링 방법	t-SNE Perplexity	Epsilon	실루엣지수
K-Means	-	-	0.1954
GMM	-	-	0.1814
DBSCAN	6	0.3	0.3244
DBSCAN	6	0.2	0.1995
DBSCAN	9	0.3	0.2195
DBSCAN	9	0.2	0.2109
DBSCAN	12	0.3	0.0106
DBSCAN	12	0.2	0.2119



군집번호	시장수
0	18
1	28
2	33
3	45
4	12
5	20

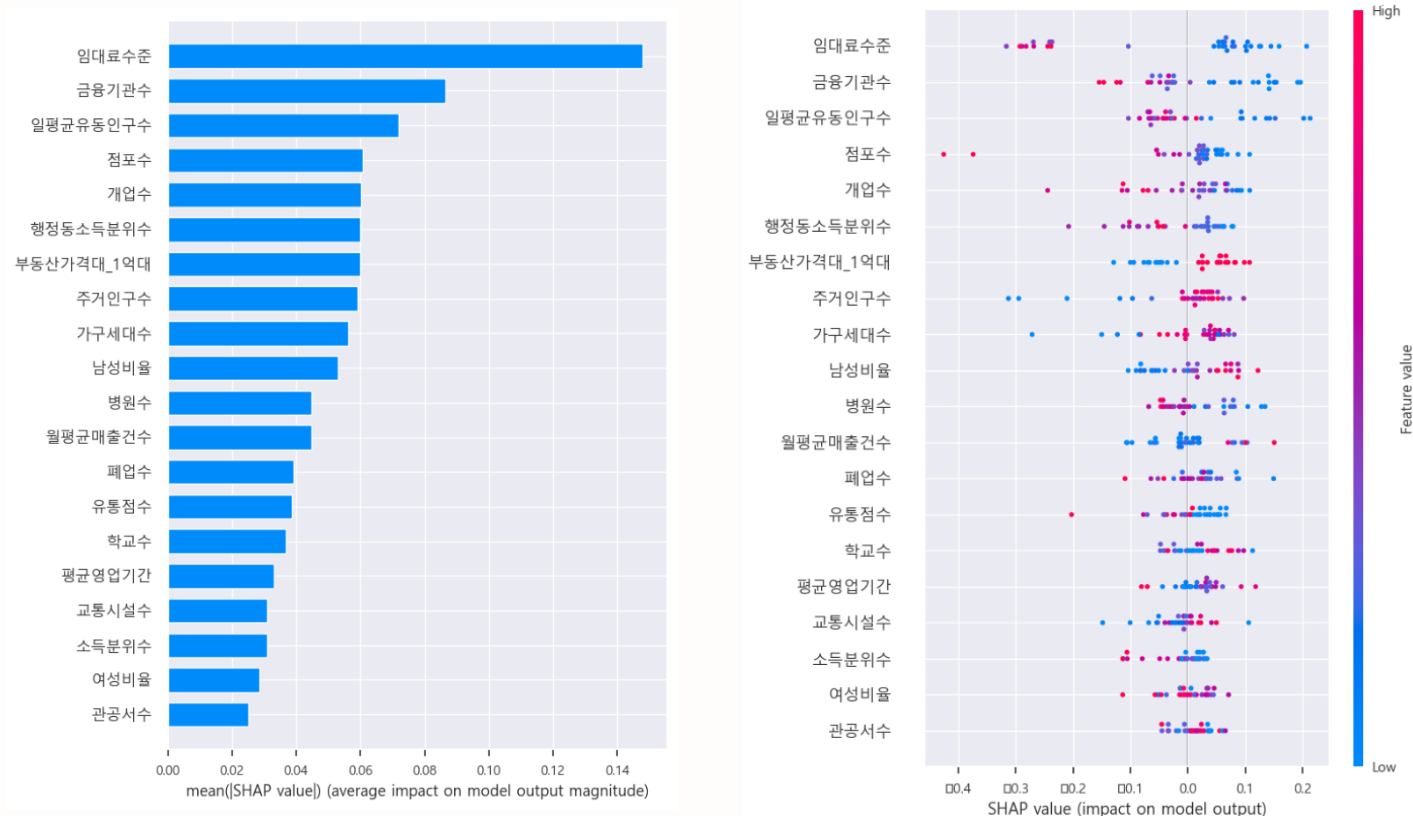
Perplexity 6, eps 0.3

04

군집 해석



- 기존 X 변수와 클러스터링 예측값을 Y로 사용한 모델을 통한 중요 변수 파악
  - SHAP(SHapley Additive exPlanations) 알고리즘 사용
  - 군집화를 함에 있어서 임대료 수준, 금융기관수, 유동인구수, 점포수, 개업수, 행정동소득분위, 주거인구수, 가구세대수, 남성비율, 병원수, 유통점수, 학교수, 평균영업기간, 교통시설, 소득분위수 순으로 중요





- **0번 군집 : 직장 특화 지역의 전통시장**

- 광장시장, 동대문종합시장, 동문시장, 낙원시장, 남대문시장, 중부시장, 방산종합시장, 평화시장, 제일평화시장, 이촌종합시장, 길음시장, 강남시장, 방이시장 등
- 신생 기업 생존률 가장 높음, 평균 영업기간 가장 김, 개업수 높음, 폐업수도 압도적으로 많음, 월평균매출건수 가장 낮음, 여성 비율 가장 높음, 유동 인구가장 적음, **직장 인구수 가장 많음**, 가구수 가장 적음, 유통점 가장 많음, 금융 기관수 가장 많음, **토요일보다 평일 특히 화요일에 매출 높음**, 점포수 가장 많음, **임대료 가장 비쌈**

- **1번 군집 : 주거+직장 임대료 높은 지역의 전통시장**

- 신설종합시장 통인시장 동묘시장 서울중앙시장 이태원시장 금남시장 마장축산물시장 이경시장 공덕시장 마포시장 망원시장 아현시장 구로시장 남성사계시장 봉천제일종합시장 봉천중앙시장 인현시장 등
- 6억이상 부동산의 비율이 가장 높음, **직장인수 두번째로 많음**, **가구세대수 평균 이상**, 교통시설 가장 적음, **소득분위 가장 높음**, 행정동 소득 분위도가장 높음, **임대료 두번째로 비쌈**



- **2번 군집 : 서민 주거지역의 소규모 전통시장**

- 창신골목시장, 신흥시장, 자양골목전통시장, 노론산골목시장, 신성골목시장, 동원전통종합시장, 사가정시장, 수유시장, 송인시장, 강북북부시장, 수유전통시장, 수유재래시장, 수유중앙골목시장 등
- 평균영업기간 가장 낮음, **유동인구와 가구세대수가 두번째**, 월평균매출액 낮음, 직장인 가장 적음, **작은 평수대 부동산이 가장 많음**

- **3번 군집 : 주거지역 활발한 상권의 전통시장**

- 만리시장 후암시장 용산용문시장 뚝도시장 중곡제일골목시장 동서시장 전농로터리시장 동부시장 청량리청과물시장 청량리종합시장 서울약령시장 청량리전통시장 회기시장 청량종합도매시장 등
- 금~토 매출이 골고루 높음, **월평균 매출액 가장 높음, 월평균매출건수 가장 높음**, 숙박수 가장 많음, 학교가장 많음, 낮은임대료, 주거 인구많고, 직장인 적은 주거지역
- 가장 많은 시장들이 해당되는 군집(45개)



- **4번 군집 : 레드오션 상권에 위치한 전통시장**

- 자양골목전통시장, 노론산골목시장, 신성골목시장, 수유시장, 백련시장, 우리시장, 성대시장, 신림중앙시장, 조원동 펭귄시장, 관악중부시장, 마천시장 능동로골목시장 우림골목시장 돈암시장 등
- **신생기업생존률 가장 낮음, 개업수 높음, 월평균매출액 낮음, 1억대 부동산 비율이 다른 군집보다 상대적으로 많음, 유동인구 가장 많음, 주거 인구 가장 많음, 가구수 가장 많음, 관공서 수 가장 많음, 병원수 가장 많음, 교통 시설 많음, 토요일에 매출이 몰려있음**

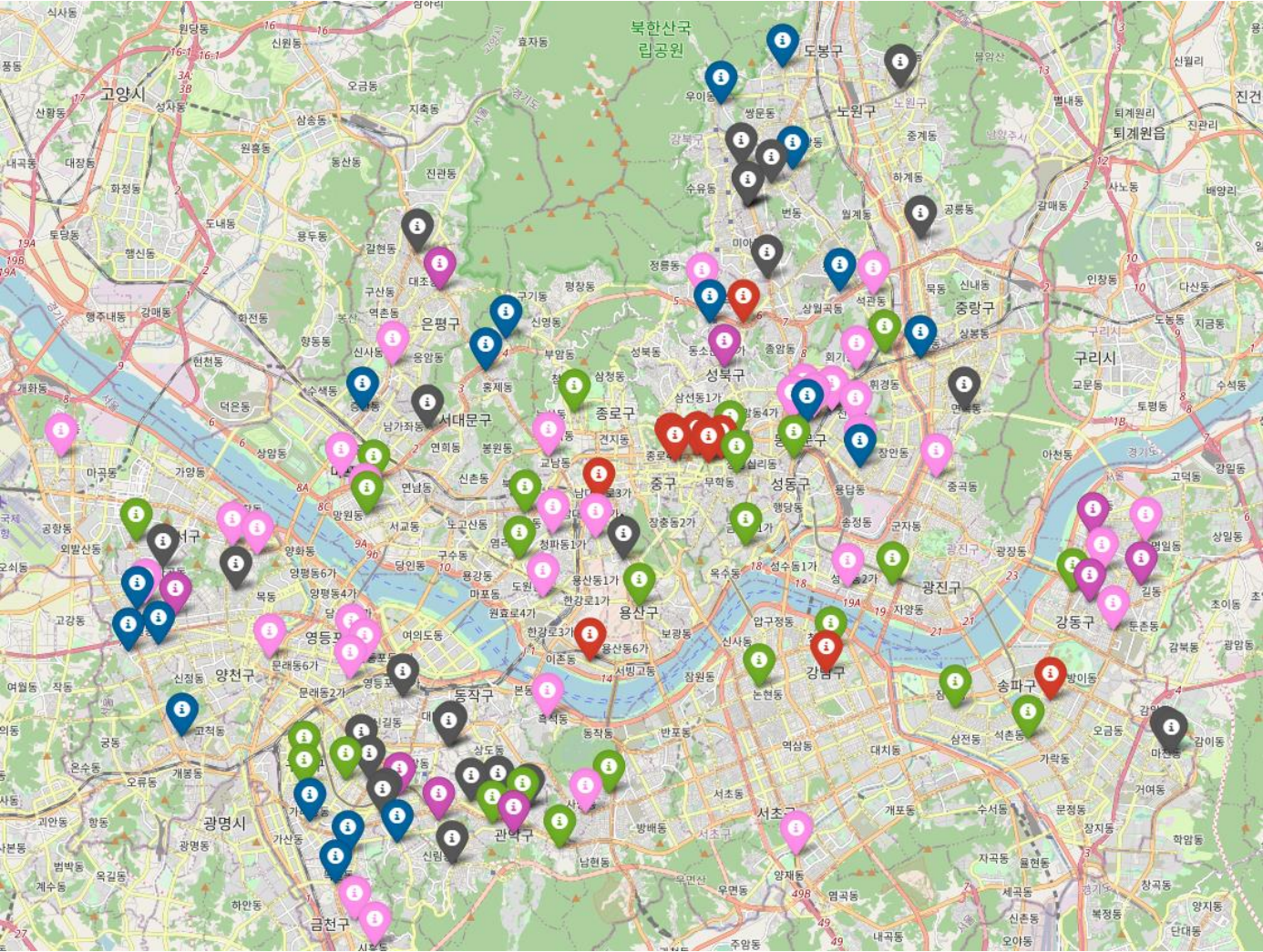
- **5번 군집 : 임대료 가장 저렴한 상권의 전통시장**

- 답십리현대시장 청량리수산시장 중랑동부시장 장위전통시장 정릉아리랑시장 백운시장 증산종합시장 인왕시장 포방터시장 경창시장 신영시장 신월3동 골목시장 고척 근린시장 가리봉시장 등
- 소매업 비율이 다른 군집보다 상대적으로 높음, **개업수 가장 낮음, 폐업수 적음, 금융기관수 가장 적음, 병원수 가장 적음, 소득 분위 가장 낮음, 행정동 소득분위도 가장 낮음, 임대료 가장 싼**



### 03 군집화

## 군집화 시각화



- 0번 군집(빨강) : 직장 특화 지역의 전통시장
- 1번 군집(초록) : 주거+직장 임대료 높은 지역의 전통시장
- 2번 군집(검정) : 서민 주거지역의 전통시장
- 3번 군집(핑크) : 주거지역 활발한 상권의 전통시장
- 4번 군집(보라) : 레드오션 상권에 위치한 전통시장
- 5번 군집(남색) : 임대료 가장 저렴한 상권의 전통시장

군집번호	시장수
0	18
1	28
2	33
3	45
4	12
5	20

05

정리



### [활용방안]

- 비슷한 성격의 시장들에게 비슷한 문제나 이슈가 생길 수 있음
- 전통시장의 군집화를 통한 현재 전통시장 현황 파악 가능
- 직장인/직장인+주거 특화지역에 위치한 전통시장 (군집0~1)
  - 직장인들 혹은 직장인+주거인의 니즈를 충족할 수 있는 아이디어 필요
- 주거지역 활발한 상권의 전통시장(군집3)보다는 레드오션 상권에 위치한 전통시장(군집4)를 고려한 정책의 필요성
  - 다양한 프랜차이즈와 경쟁에서 밀려 망할 가능성이 다른 지역보다 높을 수 있음

### [결론]

**전통시장은 주변 상권 특성에 따라 차별적인 접근법이 필요하다!**



Thank you