

# 서울특별시 상권 침체 조기경보 지수 개발

AIEYES INSIGHT TRACK Team soul 김지현, 박연후, 채유정

2025 HUFS  
LINGUATECH  
EXPO

## 1 Intro

지역 상권은 지역 주민의 일상적 생활 기반이자, 소상공인 생계와 지역경제 순환을 떠받치는 핵심 축으로 기능하고 있다. 전통적인 소비 구조에서 상권은 단순한 소비 공간을 넘어 지역 공동체의 사회적 연결망을 형성하고, 지역 고용과 창업 생태계를 지탱하는 역할까지 수행해 왔다.

그러나 최근 온라인·비대면 소비의 급격한 확대, 물가 상승 및 금리 변동, 상권 임대료 및 관리비 부담 증가 등 복합적인 외부 요인에 의해 상권 환경은 그 어느 때보다 빠르게 변하고 있다. 이러한 구조적 변화는 오프라인 중심의 지역 상권에 직접적인 충격을 주며, 유동 인구 감소, 평당 매출 하락, 빈 점포 증가 등 침체 신호가 단기간에 확산될 수 있는 높은 취약성을 낳고 있다. 특히 소비 패턴의 변동성이 커지고, 상권 간 경쟁이 심화되면서 지역 상권은 예측 불가능한 리스크에 지속적으로 노출되고 있다.

그럼에도 불구하고 현재 많은 상권 분석 체계는 여전히 분기·연 단위로 이루어지는 **사후적 현황 파악에 머물러 있다**. 분석 결과가 늦게 반영됨에 따라 실제로 상권이 쇠퇴 국면에 진입한 이후에야 대응이 이루어지는 경우가 많고, 이는 회복 비용 증가와 정책 효과 저하로 이어진다. 즉, 현행 분석 체계로는 변화 발생 이전의 위험 신호를 적시에 포착하고 선제적으로 대응하기 어려운 구조적 한계가 존재한다.

따라서 **상권 데이터를 정교하게 수집·해석하고, 다양한 요인 간 상호작용을 실시간 또는 준 실시간으로 파악할 수 있는 새로운 의사결정 체계의 필요성이 점점 더 강조되고 있다**. 이러한 체계가 구축된다면 단순 ‘사후 분석’을 넘어 다음과 같은 사회적 의의와 활용 가능성을 실질적으로 확보할 수 있다.

### [상권 침체 위험의 조기 탐지]

유동 인구·매출·공실률·지역 경제지표의 미세한 변화를 기반으로 위험 신호를 조기에 포착하여, 대응 시점의 지연으로 인한 피해를 최소화할 수 있다.

### [지역별·특성별 맞춤형 대응 전략 수립]

상권 규모, 업종 구성, 인구 특성 등 조건이 서로 다른 지역별 상권에 대해 차별화된 정책 개입과 마케팅 전략을 설계할 수 있다.

### [상권 회복 및 정책 지원의 효율적 배분]

한정된 공공 자원을 침체 가능성이 높은 상권에 우선적으로 투입하여 회복 가능성을 높이고, 예산 집행의 효율성과 효과성을 극대화할 수 있다.

### [데이터 기반 의사결정 체계 구축]

정량적 근거에 기반한 정책 수립·사업 평가가 가능해지며, 지방정부·소상공인·컨설팅 기업 등 다양한 이해관계자가 활용할 수 있는 지속 가능한 분석 기반을 마련할 수 있다.

이러한 선제적 상권 분석 체계는 단순한 분석 도구가 아니라, 지역경제가 급격한 환경 변화 속에서도 안정적으로 기능하고 위험을 최소화할 수 있도록 지원하는 핵심적 정책 인프라로 자리매김할 수 있다.

## 2 Project Architecture

### 데이터 수집 및 전처리

### 상권지수 산출

### 지표 산출

### 클러스터링

#### 유동인구 데이터

서울시 상권별 유동인구 수

서울시 버스정류장별 승하차 인원

서울시 지하철역별 승하차 인원

#### 매출 데이터

서울시 상권별 추정매출

#### 부동산 데이터

서울시 상업/업무용 부동산 면적 및 거래금액

### 유동인구 지수

### 매출 지수

### 부동산 지수

### 유동인구·매출·부동산 변화율 지표

### 유동인구·매출·부동산 레벨 지표

### 종합 상권지수 산출

### 가중합 피어슨 상관분석 반영한 가중치 부여

PCA  
4가지 성분으로 축소

클러스터링  
6가지 클러스터로 구분

### 시계열 예측

SARIMA  
전통적인 시계열 예측 모델

PROPHET  
통계 기반 머신러닝 모델

## 3 데이터 수집 및 전처리

### 유동인구 데이터

### 매출 데이터

### 부동산 데이터

- 서울시 상권분석서비스 (길단위인구-상권)
- 서울시 버스노선별 정류장별 시간대별 승하차 인원정보
- 서울시 지하철호선별 역별 승하차 인원정보

→ (1) 지하철역(버스정류장)-상권영역 매핑 (QGIS 사용)  
→ (2) 일별 승하차 승객수 → 분기별 승하차 승객수(PYTHON 사용)

- 서울시 상권분석서비스(추정매출-상권)

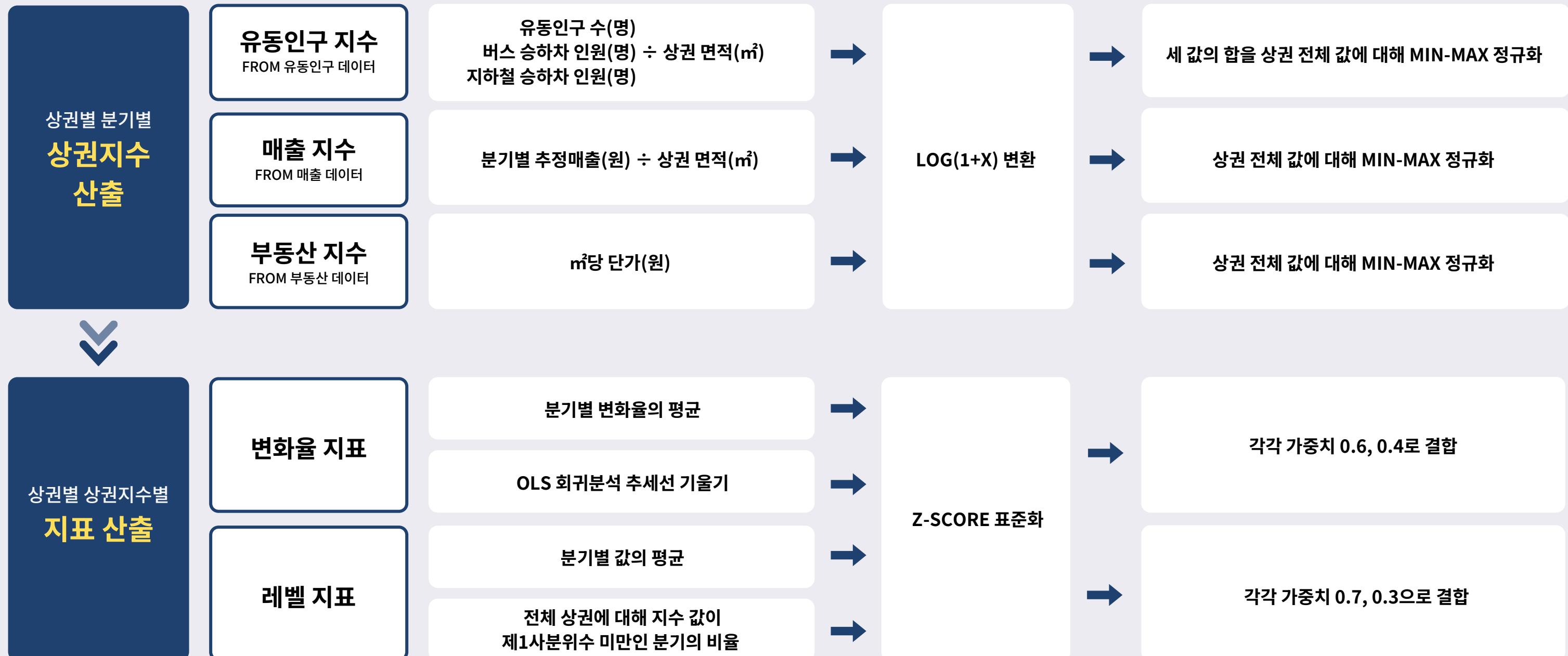
→ 매출 데이터, 위치 데이터를 상권코드 기준으로 결합

- 서울시 상권분석서비스(아파트-상권)

→ (1) 부동산 데이터-상권영역 매핑 (QGIS 사용)  
→ (2) 상권영역 기준으로 m<sup>2</sup>당 단가 평균 계산

→ 1,2,3의 데이터를 상권코드, 기준년분기 기준으로 통합

## 4 상권지수 및 지표 산출



- 상권지수 산출: 단위가 각기 다른 유동인구, 매출, 부동산 데이터를 정규화하여, 상권 간 절대적 규모를 직관적으로 비교 가능한 지수로 환산하였다.
- 지표 산출: 상권 내의 시계열 변화 추이를 파악해 동적 특성을 판단할 수 있도록 "변화율 지표", 상권 간의 절대 수준을 파악해 정적 특성을 판단할 수 있도록 "레벨 지표"를 구성하였다.

## 5 클러스터링

### a. 주성분 분석(PCA)

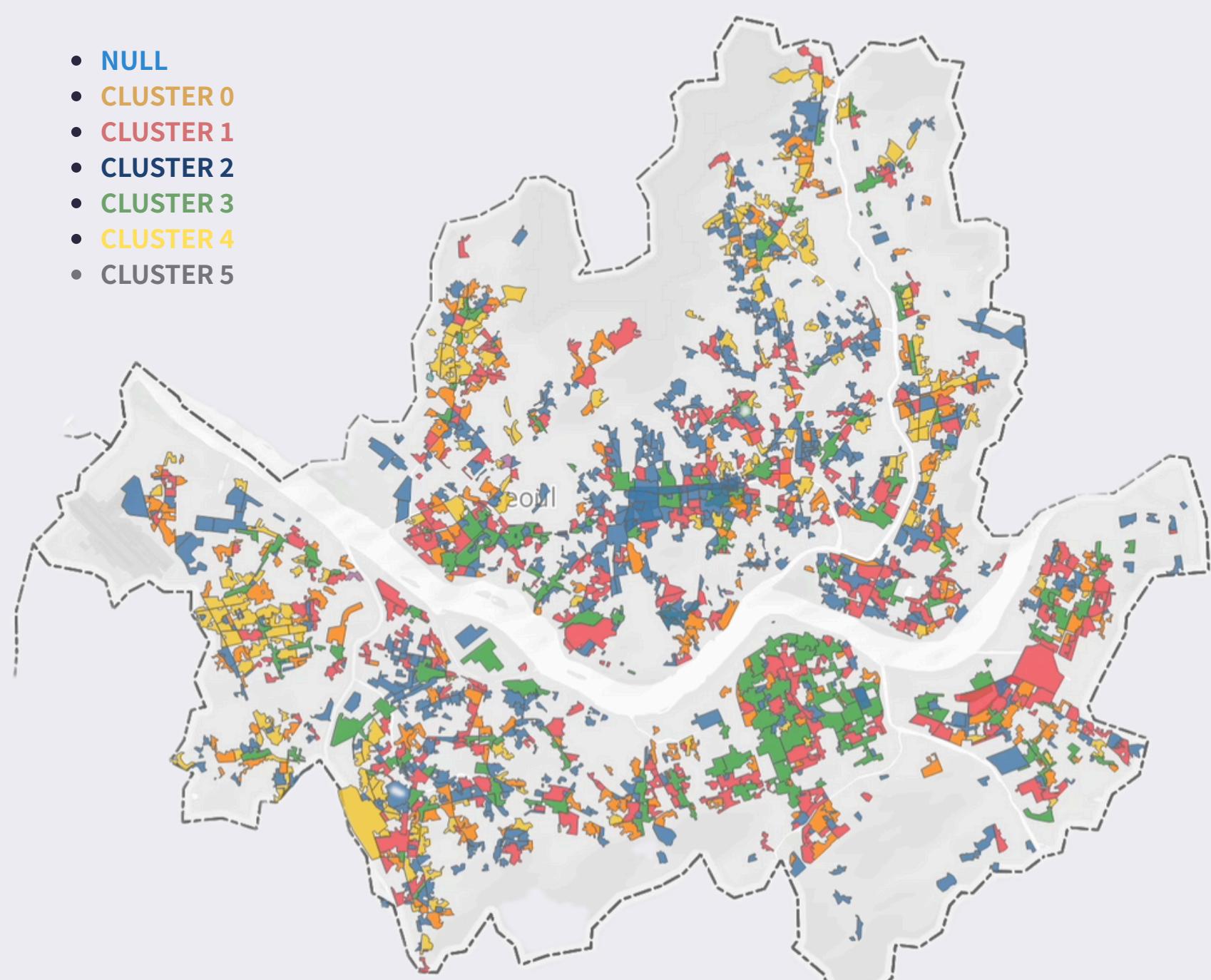
PC1 = “전반적인 상권 레벨(절대 수준)” 축  
“상권의 전체적인 절대 규모(매출/유동/부동산)가 큰가?”

PC2 = “자산·부동산 중심의 변화성” 축  
“부동산 기반 변동성이 심한가?”

PC3 = “유동(인구) 중심 성장·하락 추세” 축  
“유동인구 기반 변동성이 심한가?”

PC4 = “매출·부동산 단기적 변동성/특성 구분” 축  
“매출 기반 변동성이 심한가?”

### b. K-Means clustering



앞선 단계에서는 상권의 성격을 입체적으로 파악하기 위해 '매출, 유동인구, 부동산'이라는 지수를 각각 레벨(절대적 규모), 변화율(시계열적 추세)로 세분화하여 총 6차원 데이터를 구축하였다.

변수 간의 다중공선성 문제를 해결하고, 분석의 효율성을 높이기 위해 PCA(주성분 분석)을 수행하였고, 상권의 특성을 가장 잘 대변하는 4개의 핵심 잠재 변수를 도출하였다.

이 4가지 축을 기준으로 K-Means Clustering을 수행하여 유사한 상권들을 그룹화하였다. 이를 통해 단순 규모 차이를 넘어 '쇠퇴 정체 상권', '안정 성장 상권', '대형 핵심 상권', '부동산 급등 잠재 상권'이라는 정성적 특징이 뚜렷한 4개의 그룹과 특이 그룹(Outlier) 2개를 도출하였다.

### [Cluster 0] 쇠퇴 정체 상권

특징: [인구 감소 & 낮은 활력] 인구/매출 레벨이 평균 이하이며, 특히 인구 변화율이 마이너스로, 인구가 유출되고 있는 지역

상태: 매출은 소폭 오르고 있으나, 상권 자체의 활력은 떨어지는 주거 밀집지나 구도심일 수 있음.

주요 지역: 외대앞역 2번, 혜화역 1번, 상수역 3번, 마포역 4번, 영등포역 2번 등

### [Cluster 1] 안정 성장 상권

특징: [균형 잡힌 성장] 매출과 부동산 레벨이 평균 이상이며, 인구와 매출이 꾸준히 성장하고 있는 가장 건실한 상권 유형

상태: 부동산 지수도 높고 성장세도 양호하여 실질적인 투자 가치가 높은 지역들

주요 지역: 외대앞역 1번, 건대입구역 6번, 성수동 카페거리, 성신여대입구역 4번, 흥의대학교 등

### [Cluster 3] 대형 핵심 상권

특징: [압도적인 규모] 인구, 매출, 부동산 레벨 지표가 모두 최상위권 수준

상태: 이미 성숙한 상권으로 변화율(성장세)은 폭발적이기보다 안정적인 수준을 유지하고 있음.

주요 지역: 이태원 관광특구, 뚝섬역, 여의도역, 석촌역(석촌호수), 압구정 로데오 등

### [Cluster 4] 부동산 급등 잠재 상권

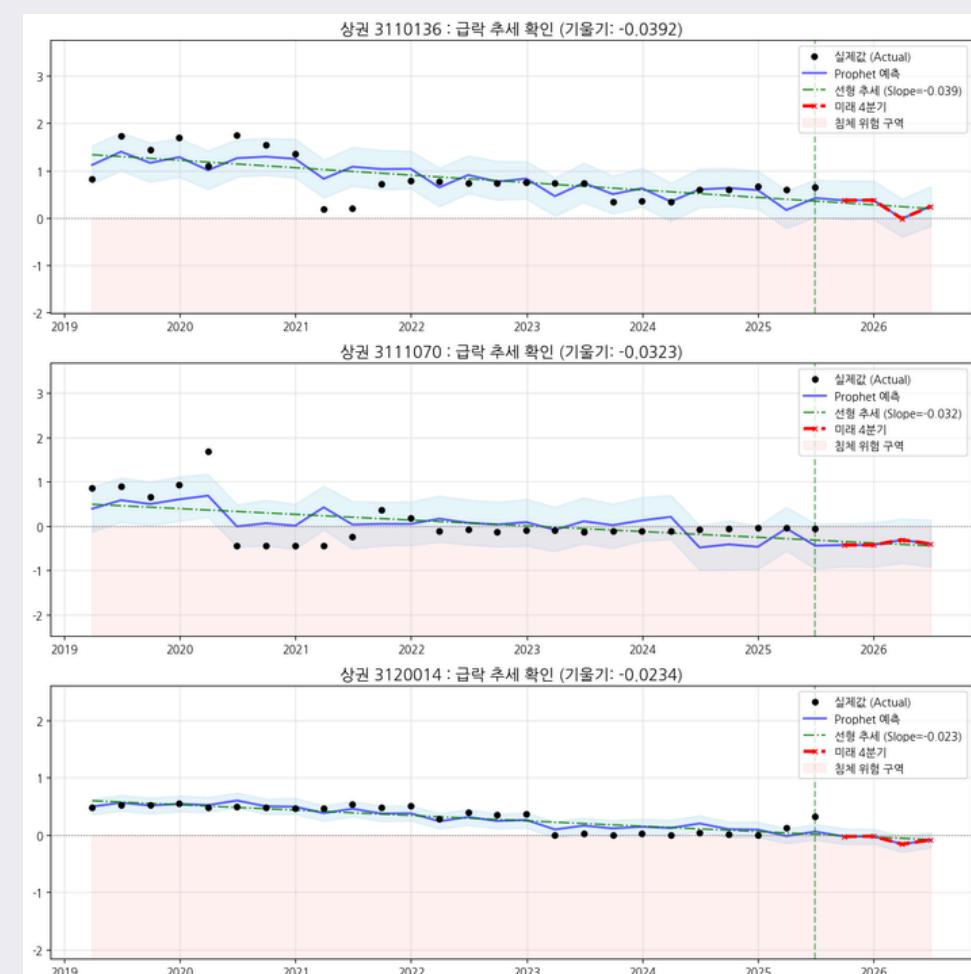
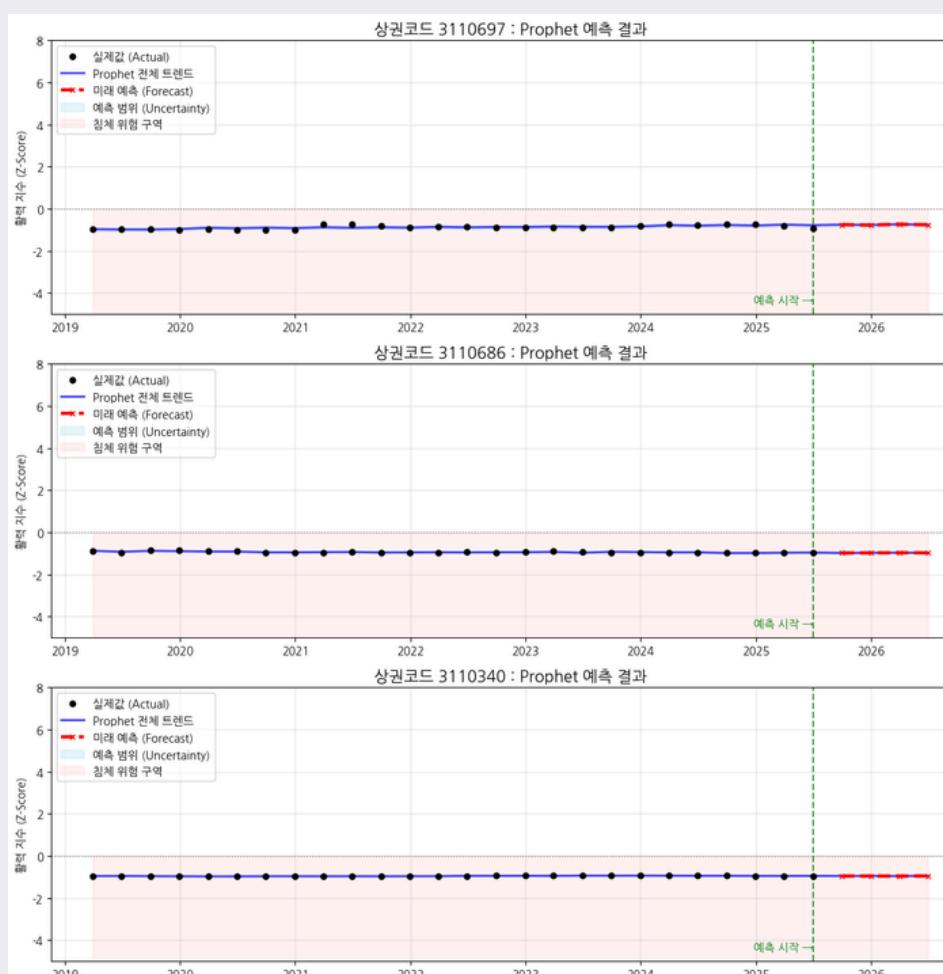
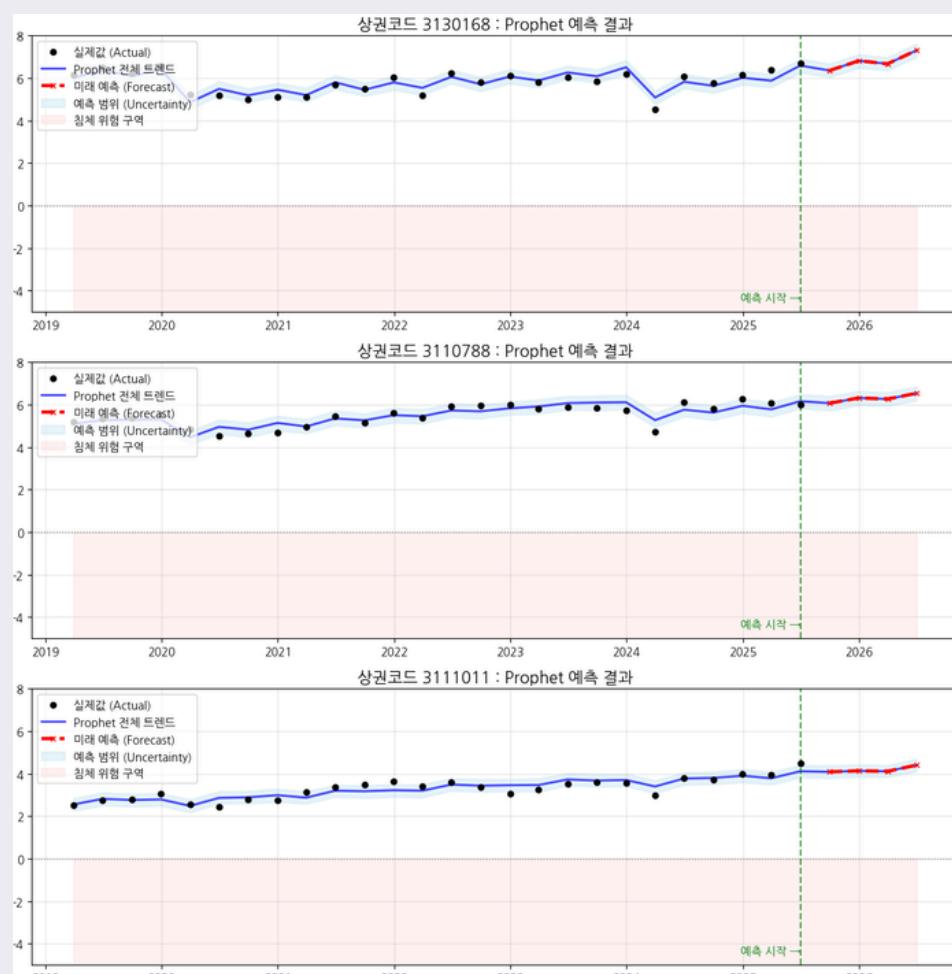
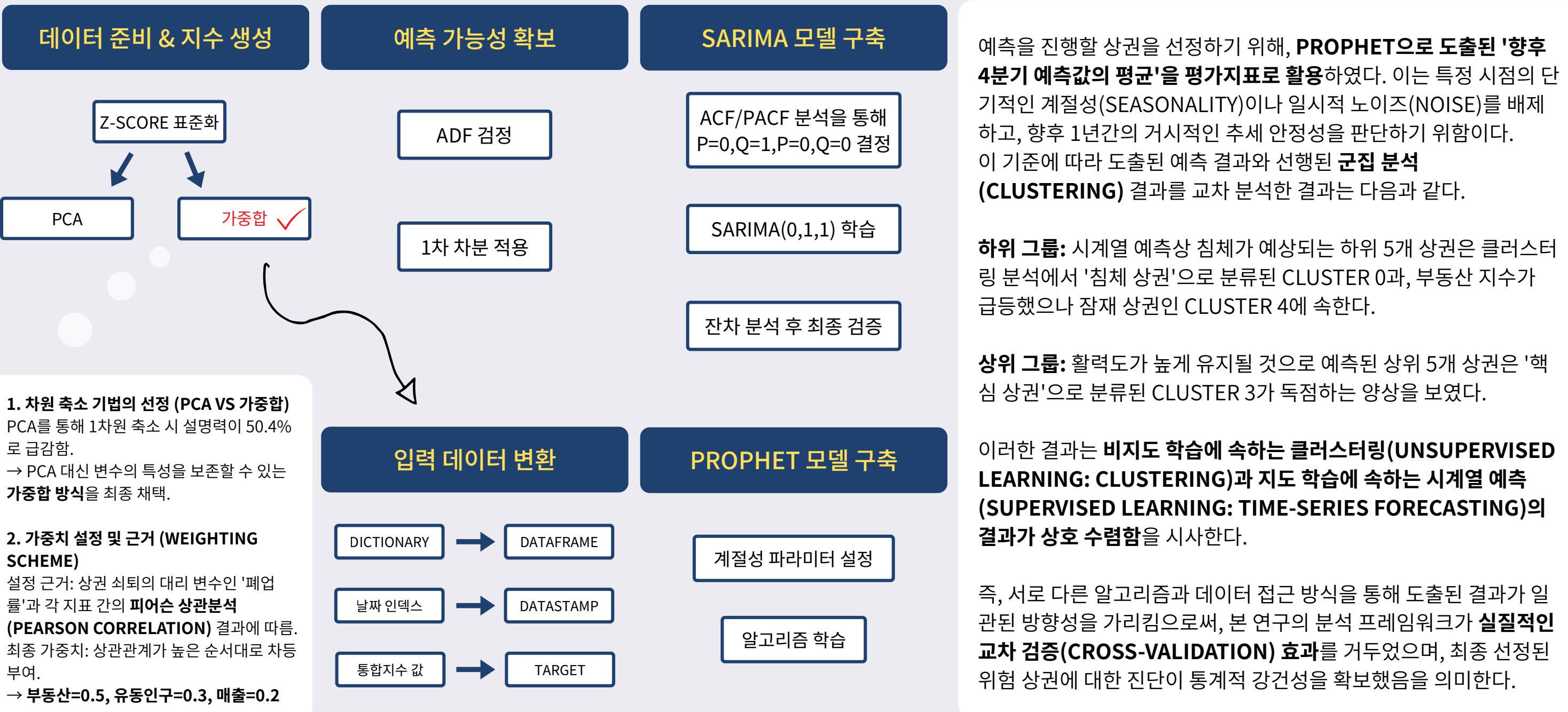
특징: [부동산 가치 급상승] 현재 부동산/매출의 절대 레벨은 평균 이하로 낮지만, 부동산 변화율(0.033)이 다른 그룹(0.010) 대비 3배 이상 높음.

상태: 상권 규모는 작지만 부동산 가치가 급격히 오르고 있는 '개발 예정지'나 '신흥 엔트리피케이션' 지역일 가능성이 큼.

주요 지역: 서래마을 카페거리, 목동 로데오, 동대문역 1번 등

## 6 위험 상권 시계열 예측

본 연구는 초기 모델링 단계에서 GRU와 같은 딥러닝 기반 모델을 고려하였으나, 24시점이라는 제한적인 데이터셋 환경에서는 파라미터 과다로 인한 과적합(Overfitting) 우려가 있어 제외하였다. 이에 대안으로 전통적 시계열 모델인 SARIMA와 통계적 기계학습 모델인 Prophet을 비교 분석하였다. 분석 결과, SARIMA는 데이터 부족으로 인해 계절성과 추세 변동을 충분히 학습하지 못하고 장기 평균으로 수렴하는 보수적인 예측(Conservative Prediction) 경향을 보였고, 이는 급격한 상권 침체 징후를 조기에 포착해야 하는 본 연구의 목적에 부합하지 않는다고 판단하였다. 반면, Prophet은 추세 변화점(Trend Changepoints) 탐지에 유연하여, 최근 관측된 상권의 활력 저하 추세를 미래 예측에 효과적으로 반영하였다. 따라서 상권의 하방 위험(Downside Risk)을 선제적으로 감지하기 위해 최종 모델로 Prophet을 선정하였다.



상권코드	상권코드명	자치구	클러스터
3130168	연서시장	은평구	3
3110788	영등포시장역 3번	영등포구	3
3111011	잠실나루역	송파구	3

상권코드	상권코드명	자치구	클러스터
3110697	개봉3동주민센터	구로구	4
3110686	수궁동주민센터	구로구	0
3110753	시흥동은행나무사거리	금천구	0

상권코드	상권코드명	자치구	클러스터
3110136	경수초등학교	성동구	1
3111070	천호초등학교	강동구	0
3120014	종로5가역	종로구	3

시계열 예측 결과, 상·하위 그룹의 전반적인 예측 수준은 기존 군집 분석 결과와 일치하여 분석 **프레임워크의 정합성을 확인**하였다. 다만, 추세선의 기울기가 급격히 하락하는 '고위험군'을 상세 분석한 결과, **특정 취약 군집뿐만 아니라 우수 군집(CLUSTER 1, 3)까지 존재된 양상**을 보였다. 이는 상권 유형(CLUSTER)과 무관하게 개별 상권의 쇠퇴가 진행될 수 있음을 시사하며, 군집 분석의 한계를 보완하는 시계열 모니터링의 실효성을 보여준다.

## 7 결론 및 시사점

### 1. 연구 요약

- 본 연구는 서울시 1650개 상권의 24분기 시계열 데이터를 기반으로 입체적인 상권 분석을 수행하였다.
- 지표 고도화:** 여러개의 데이터 단위를 통일하고 분포를 보정하기 정규화를 수행, 상권의 절대적 규모와 동적 변화를 동시에 설명하는 '상권지수'를 산출하였다.
  - 군집화(CLUSTERING):** PCA를 통해 도출된 핵심 인자를 바탕으로 상권을 정의하였다.
  - 미래 예측(FORECASTING):** PROPHET 알고리즘을 활용해 향후 4분기의 추세를 예측하여 선제적 위험을 진단하였다.

### 2. 분석 결과

분석 결과, 군집 분석과 시계열 예측 결과는 거시적으로 상호 수렴하며 연구 모형의 정합성을 입증하였다. 그러나 미시적 관점에서는 중요한 발견이 있었다. 미래 쇠퇴가 예측되는 고위험 상권들이 단순히 '침체 군집(CLUSTER 0)'에만 머무르지 않고, '건설한 성장 군집(CLUSTER 1, 3)' 내에도 다수 분포함을 확인하였다. 이는 현재의 정적인 상권 등급이 우수하더라도, 시계열적 하락 신호는 개별 상권마다 다르게 나타날 수 있음을 시사한다.

### 3. 정책적 시사점

- 본 연구는 기존 단편적인 상권 활성화 정책에 다음과 같은 변화가 필요함을 시사한다.
- 조기 경보 시스템(EARLY WARNING SYSTEM) 도입:** 쇠퇴가 고착화되기 전에 개입하는 조기 경보 체계가 필요하다.
  - 데이터 기반의 '핀셋(PINSET)' 지원:** 건설한 상권 군집에 속더라도 하락 징후가 뚜렷한 '숨은 위험 상권'을 발굴해야 한다.
  - 입체적 모니터링 체계 구축:** 절대적인 규모와 변화의 속도를 동시에 모니터링하는 방식을 서울시 상권 분석 시스템에 적용하여, 정책 의사결정의 정교함을 높여야 한다.