

## 제 11 회 「2023 빅콘테스트」 결과보고서

\* 해당란에 ☒ 표시

참가분야	<input type="checkbox"/> 생성형 AI 분야 <input checked="" type="checkbox"/> 데이터신기술 분야 <input type="checkbox"/> 정형데이터 분석 분야 <input type="checkbox"/> 비정형데이터 분석 분야 <input type="checkbox"/> 빅데이터플랫폼 활용 분야		
세부리그 <small>*해당시 체크</small>	<input type="checkbox"/> 어드밴스드 리그 <input type="checkbox"/> 스타터 리그 <small>*정형데이터 분석분야에 한함(선택)</small>		
	<input type="checkbox"/> 지정주제 리그 <input type="checkbox"/> 자유주제 리그 <small>*빅데이터플랫폼 활용분야에 한함(선택)</small>		
개인/팀여부	<input type="checkbox"/> 개인 <input checked="" type="checkbox"/> 팀(총 4 명)	개인/팀명	데이벳
지도교사명	-		
대표 ID	minnie3120@naver.com		

# **목 차**

1. 개요.....	3
가. 배경 및 필요성 .....	3
1) 젠트리피케이션 이슈 .....	3
2) 젠트리피케이션과 상업젠트리피케이션 개념 및 양상.....	3
3) 국내 사례 .....	4
나. 연구 목적.....	4
다. 연구 범위 및 방법.....	5
1) 데이터 설명.....	5
2) 연구 범위 및 방법 .....	5
나. 선행연구와 본 연구의 차별성.....	6
1) 선행연구 현황 .....	6
2) 선행연구와 본 연구와의 차별성.....	6
2. 문제 수행 내용 .....	7
가. 젠트리피케이션 상권 분류 .....	7
1) 변수 선정 및 계산 .....	7
2) 클러스터링 .....	9
3) 최적 군집 개수 결정 .....	9
4) 상권 클러스터링 결과 .....	9
나. 젠트리피케이션과 경영위기의 가설검정 .....	13
1) 귀무가설의 설정 .....	13
2) 귀무가설의 검정 .....	13
다. 젠트리피케이션 여부에 유의미한 영향 미치는 변수 선택 .....	16
1) 고려 가능한 모든 변수 구하기.....	16
2) Anova를 통해 평균 차이 살펴보기 .....	16
3) Boruta 알고리즘을 통한 변수 선택.....	17
라. 젠트리피케이션 예상 지역 예측.....	18
1) 랜덤포레스트를 활용한 젠트리피케이션 예측 모델.....	18
2) 새로운 상권 적용 .....	18
마. 젠트리피케이션 군집 내 경영위기 .....	20
1) 경영위기에 따른 anova 검정 .....	20
2) 최종 변수 산정 .....	21
3. 활용방안 .....	23
가. 예비 창업주 및 사업주를 위한 젠트리피케이션 대시보드 제작 .....	23
나. 청년 사업주 법률 자문 및 상담 지원 .....	24
4. 주요 결과 및 시사점 .....	25
가. 주요 결과 요약 .....	25
나. 의의 .....	25
다. 한계점 및 향후 발전 방향 .....	26
5. 참고문헌 .....	26

## 1. 개요

### 가. 배경 및 필요성

#### 1) 젠트리피케이션 이슈

지난 2016 년에 서울 종로구 서촌의 한 맛집으로 유명했던 음식점의 사장이 건물주에게 갈등으로 인해 망치로 공격한 사건이 발생하였다. 해당 사건은 임대료 문제와 도심 지역에서의 부동산 가치 상승으로 인한 문제가 격화된 사례로, 젠트리피케이션으로 인한 원주민의 비자발적 이주라는 극단적인 결과를 보여주는 사례로 꼽히고 있다. 초기에는 임대료가 낮았으나, 주변 지역의 인기 상승으로 인해 임대료가 크게 인상된 것이 갈등의 시작점이 되었다.

점포 가치 상승에도 불구하고 권리금을 받지 못했다고 퇴거 요청을 받았던 임차인은 해당 퇴거명령이 부당했다고 주장하며 항의하였고, 이후 법적 분쟁으로 이어지게 되었다.

#### 2) 젠트리피케이션과 상업젠트리피케이션 개념 및 양상

도시 취약계층의 주거 및 상업 공간이 중산층 및 상위계층에 의해 다른 지역으로 대체되는 현상을 일컫는 개념이다. 1964 년 영국의 사회학자 루스 글래스가 저소득층 주거 공간의 구조개편과 관련된 용어로 처음 사용하였지만, 최근 들어 상업, 문화, 예술 등 점차 광범위한 영역에서 사용 되고 있다.

젠트리피케이션의 긍정적인 의미로는 낙후된 구도심에 상류층 인구의 유입으로 고급주거지역, 고급 상점가가 새롭게 형성되는 현상으로 여겨지고 있다. 그러나, 부정적인 의미에서는 젠트리피케이션으로 인해 재산가치 상승, 지가 상승이 일어나면 저소득층 가구가 지역사회 이동을 겪게 되는 문제점이 있어 젠트리피케이션에 대한 의문의 끊임없이 제기되고 있다.

특히, 2000 년대 들어 우리나라에서는 주거보다는 상업지역의 재활성화와 함께 급격한 임대료 상승에 따른 문화 예술인과 소상공인들의 비자발적 이주 문제가 발생하고 있다. 이는 재개발이나 대규모 택지조성으로 인해 나타나는 젠트리피케이션 이주와 달리, 상업 젠트리피케이션은 지나친 상업화로 야기된 문제로서 선제적으로 대응할 필요성이 대두되고 있다.

국내 상업 젠트리피케이션이 심화된 지역 상권에서는 임대료 급등으로 인해 기존 상권의 정체성을 형성하던 문화 예술인과 소상공인 등은 이탈하고 있다. 이는 공동체의 정체성을 확립하고 만들어가던 이들을 퇴출시키고 지역의 매력을 저하시키는 결과를 초래하고 있다. 이는 활성화된 지역을 다시 쇠퇴하게 한다는 점에서 문제로 나타나고 있다.

### 3) 국내 사례

#### [성동구의 사례]

국내에서는 서울의 주요 상권인 경리단길, 연남동, 성수동 등에서 투기세력이나 외부 자본의 유입으로 문제가 발생하기 시작했다. 지난 2015 년 부터 서울 성동구에서는 지속가능발전구역을 지정하여 젠트리피케이션 방지 정책을 시행중에 있다. 그러나 구역에 지정되지 않는 성수역 과 연무장길 일대가 젠트리피케이션 위험에 빠져 있다. 이 지역의 특징은 MZ 세대를 중심으로 유동인구가 증가하고 있는 핫플레이스로 여겨지고 있다는 점이다. 또한 경기침체에도 불구하고 유명 브랜드 입점 등으로 공실률이 0 에 가까운 지역들이며, 유동인구의 증가로 매출이 지속적으로 성장하고 있으나, 임대료도 함께 급격하게 증가하고 있는 것으로 나타나고 있다.

#### [대구 중구 대봉동의 도시재생 사례]

대구광역시에는 김광석 거리로 일컬어지는 방천시장이 있다. 방천시장은 도시재생사업의 일환으로 정부와 지방자치정부의 지원을 받아, 지역 예술인들에게 빈 점포를 저렴하게 임대하여 상권의 활력을 불어넣었다. 김광석 거리 조성 이후 이동인구가 증가하고 도심 내 관광명소로 소개되며 상권활성화를 이룬 지역이다. 그러나 2011 년에 보증금 500 만원, 월세 30~45 만원 정도였으나, 2014~15 년에 보증금은 3,000 만원, 월세는 90~100 만원까지 오르면서 영세 점포 상인들의 임차료 부담이 가중되어, 자리를 떠나게 되었다. 그 빈 자리에 음식점과 커피숍이 대체하게 되었다. (송지은 외,도심의 상업적 젠트리피케이션 대응 방안 연구 2016)

### 나. 연구 목적

이 연구는 "한국신용데이터 업장 매출입 및 기타 특성 정보"를 기반으로 경영위기

(사업장의 추정영업이익이 마이너스이고, 전월대비 10% 이상 감소한 경우, 추정영업이익 = ('카드 매출' + '매출 세금계산서' + '배달 매출') - ('매입 세금계산서' + '매입 카드' + '매입 현금'+월 임대료'))

의 요인 중 하나는 젠트리피케이션 현상이라는 가설을 세웠고, 이에 미치는 영향을 탐구하는 목적으로 시작되었다. 젠트리피케이션은 도시 및 지역 사회에서의 중요한 사회 경제적 변화 중 하나로, 경영위기와 연관성을 조사하기 위한 연구를 하였다. 이 연구의 핵심 목표는 서울시 상권을 젠트리피케이션 여부에 따라 분류할 수 있는 새로운 지표를 개발하는 것이었다. 이 지표는 서울시의 \_\_개의 상권을 젠트리피케이션의 여부에 따라 분류하고, 이를 통해 경영위기에 미치는 영향을 더 자세하게 이해하는 데 기여하는 것을 목적으로 하였다. 뿐만 아니라, 개발된 젠트리피케이션 지표를 활용하여 경영위기와 젠트리피케이션 간의 관계를 조사했다. 이를 통해 젠트리피케이션과 경영위기 간에 가장 큰 영향을 미치는 요인들을 도출하였다. 이러한 요인들을 식별함으로써 경영위기를 예방하거나 경영 상황을 개선하는 데 도움을 줄 수 있는 방안을 모색하였다. 이러한 연구 과정은 도시 및 지역 개발에 관련된 정책 및 전략 수립에 중요한 정보를 제공할 것으로 기대된다.

## 다. 연구 범위 및 방법

### 1) 데이터 설명

본 연구에서는 서울시 열린데이터 광장에서 제공하는 서울시 상권분석 서비스 공공 데이터와 한국신용데이터에서 제공하는 업장 메타 정보와 업장 매출업 정보 등을 활용한다. 해당 데이터는 2022년 1월부터 각각 16개월, 12개월동안 데이터를 수집하였다.

상권 분석 서비스 데이터는 상권 기본 정보, 점포 현황, 인구통계학 및 시계열 추정 매출 정보, 상권변화지표, 집객시설, 아파트, 상주인구, 직장인구, 생활인구, 소득소비 범주에서 총 200여 특성 정보를 포함하고 있다.

한국 신용 데이터는 사업장 정보는 서울 시내 3,067개 사업장과 관련하여 암호화 처리된 가명 정보를 제공합니다. 업장 메타 정보와 관련한 8개의 특성과 매출입 및 기타 특성정보와 관련된 시간 정보, 임대료, 고용인원, 고객인원, 매입 매출액과 연속형 변수에 대한 z-score 표준값 등 30개의 특성 정보를 제공하고 있다.

특히 한국 신용 데이터를 제공하는 디사일로는 동형암호 기반의 데이터 공유 및 협업 솔루션을 제공하고 있다. 이는 개별 식별이 가능한 민감정보를 노출시키지 않으면서 많은 데이터를 공유 및 결합하여 분석을 수행할 수 있도록 연산 결과 데이터를 제공하고 있다.

### 2) 연구 범위 및 방법

#### 2.1) 연구 범위

본 연구는 서울시 상권별 젠트리피케이션 여부를 특정하고, 정책적인 활용방안을 모색하는 것을 목적으로 한다. 이에 따라 한국 신용데이터 상권목록에서 정의하고 있는 상권 113개를 기준으로 분석하였다. 연구의 시간적 범위는 데이터 테이블 정의서에서 한정된 2022년 1월 ~ 2023년 4월 (총 16개월)로 한정하였다.

#### 2.2) 연구 방법

##### 2.2.1) 문헌연구

젠트리피케이션에 대한 분석을 수행하기 전, 기존의 선행 연구에 기반하여 체계적이고 포괄적인 문헌 검토를 실시하였다. 이 과정에서 다양한 학술 논문, 보고서, 정부 발간 자료, 통계 데이터 등 다양한 정보원을 참고하였다. 우선, 젠트리피케이션에 관한 선행 연구를 탐색하기 위해 학술 데이터베이스와 도서관 자료를 활용하여 관련 논문들을 검색하였다. 이 검색은 젠트리피케이션 현상의 정의, 특성, 원인 및 영향 등을 다룬 선행 연구를 특히 중점적으로 찾아보았다. 뿐만 아니라, 서울시와 다른 도시에서의 젠트리피케이션에 대한 사례 연구와 지역별 특성을 비교하는 연구를 찾아내어 비교 분석의 기초를 마련하였다. 이를 통해 서울시 상권의 젠트리피케이션 현상을 이해하는 데 필요한 맥락과 지식을 구축하였다. 마지막으로, 분석에 사용될 데이터와 변수를 결정하기 위해 관련 자료들 중에서 필요한 정보를 추출하고 정리하였다.

체계적이고 포괄적인 문헌 검토를 통해 연구의 방향성을 설정하고 연구 프레임워크를 갖추었으며, 선행 연구를 바탕으로 연구의 핵심 질문을 명확히 하여 연구의 타당성을 높였다.

## 2.2.2) 국내 외 사례 검토

서울시의 젠트리피케이션 분석을 시작하기 전, 국내 및 해외에서 이미 수행된 다른 지역의 젠트리피케이션 연구와 사례를 검토하였다. 국내 사례로는 서울시 이태원과 망원동, 미국의 사례로는 뉴욕, 샌프란시스코의 사례를 확인하였다. 이러한 선행 연구 검토를 통해 서울시에서의 젠트리피케이션 현상을 다른 지역의 경험과 비교하고, 서울시의 독특한 문화, 경제, 지리적 특성을 고려하여 분석 및 해석할 수 있는 기반을 마련하였다. 또한, 이러한 선행 연구를 참고하여 연구의 목표와 방향을 보다 명확히 설정하였다.

## 나. 선행연구와 본 연구의 차별성

### 1) 선행연구 현황

송윤아 송정은 이준기 (2023)는 일상 회복 시점 이후 상권이 회복하면서 임대료가 급상승한 점에 주목했고, 거리두기 전후 젠트리피케이션의 차이를 분석하였다. 젠트리피케이션과 관련한 텍스트 데이터를 수집 후 LDA 모델링을 거쳐 젠트리피케이션 지역의 특징을 분석했다. 또한 클러스터링을 통한 젠트리피케이션 지수를 만들었고 각 상권을 자체 지수로 평가하였다.

김영재 박인권 (2023)은 서울시 젠트리피케이션 종합대책(서울시,2015)에서 지정한 17개의 젠트리피케이션 상권을 과거 발생지역, 발생지역, 예상지역으로 분류하였다. 상권별 특징 분석을 위해 네이버 게시글 크롤링을 통해 키워드를 추출하였으며, 위계적 군집분석, 군집별 토픽 모델링(LDA)를 통해 상권마다 특징을 분석하였다. 젠트리피케이션의 가장 큰 위기 요인을 정통성의 부재라고 언급한다.

박재희 김태형 (2020)은 전통적 관점에서의 Freeman(2005)의 연구를 참고해 젠트리피케이션 발생지역을 측정하였다. 또한 국내 선행연구를 바탕으로 젠트리피케이션 지수를 책정하였고 젠트리피케이션 발생지역과 미발생지역을 이항로지스틱을 통해 분석하였다. 이를 통해 젠트리피케이션 분포를 파악하고 젠트리피케이션의 입지 요인 및 영향 요인을 분석하였다

### 2) 선행연구와 본 연구와의 차별성

본 연구는 경영 위기에 영향을 미치는 다양한 변수들을 바탕으로 젠트리피케이션 발생지역과 미발생지역을 이진 분류한다. 단순히 분류에서 그치는 것이 아니라 젠트리피케이션 발생지역에 결정적인 영향을 미치는 요인들을 추출하고 경영위기와도 연관성이 있는지 분석한다. 또한 행정구역 별로 분석을 한 많은 선행 연구들과는 달리 상권별 분석을 진행해 상권별 특징과 젠트리피케이션 여부를 연관 지을 수 있다. 젠트리피케이션을 상권별, 요인별로 나눠 상세히 분석함으로써 해결 방법을 좀 더 구체적이고 현실적으로 만들어 낼 수 있도록 하였다. 또한 단순히 평균과 표준편차를 사용해서 젠트리피케이션을 구분한 것이 아닌 클러스터링 방식을 사용하여 상권들의 유사성을 고려할 수 있었다. 마지막으로 예측 모델링을 구축하여 새로운 임대료와 경영위기에 대한 정보를 모르는 지역에도 젠트리피케이션 여부를 나눌 수 있는 새로운 모델을 제시하였다.

## 2. 문제 수행 내용

### 가. 젠트리피케이션 상권 분류

#### 1) 변수 선정 및 계산

국토 연구원에 따르면 젠트리피케이션을 진단하기 위해 행정동 단위의 상주인구, 유동인구, 창폐업, 영업기간, 프랜차이즈 업체, 매출액 변수를 사용한다.(이진희, 2019, p.11). 하지만 행정동 단위 안에서도 여러 상권이 존재하여 자세한 분석이 불가능하다는 점을 보완하고자 본 연구에서는 상권 단위로 변수를 계산하였다. 상권별로 계산할 경우 정확한 상주인구 데이터를 구할 수 없었으며, 해당 변수의 우선순위가 가장 낮다는 점을(이진희 외, 2018, p.102) 참고하여 제외하였다. 또한 기존 연구에서 데이터 수집의 어려움으로 사용하지 못했던 임대료 변수와 경영위기 변수를 추가하여 최종 변수를 산정하였다. 산정된 변수와 젠트리피케이션에 미치는 영향에 대한 설명은 아래와 같다. 새로 선정한 임대료와 경영위기 외 나머지 변수는 기존 연구의 방법으로(이진희, 2019, p.12) 연구하였다.

#### [유동인구]

유동인구는 상업 젠트리피케이션에서 해당 지역의 활성화 정도를 측정할 수 있는 중요한 변수이다. 따라서 젠트리피케이션이 진행되면서 방문객이 증가하면서 유동인구가 증가하는 양상을 띄게 된다. 상권별로 유동인구를 확인할 수 있는 정확한 데이터 역시 구할 수 없기 때문에 “특정 지역, 특정 시점에 존재하는 모든 인구”를 나타내는 생활인구로 대체하여 사용하였다.

$$\text{생활인구 증가율} = \frac{(\text{현재 변수값} - \text{이전 변수값})}{\text{이전 변수값}}$$

$$\text{현재 변수값} = \frac{(2022\text{년 } 3\text{분기} + 2023\text{년 } 1\text{분기})}{2}, \text{이전 변수값} = \frac{2021\text{년 } 3\text{분기} + 2022\text{년 } 1\text{분기}}{2}$$

#### [창폐업]

창폐업은 신규 창업자의 증가나 빈번한 업종 변경 현상을 살펴볼 수 있는 중요한 변수이다. 따라서 창폐업 비율이 증가할수록 상업활동이 증가하는 것으로 해석할 있으며 젠트리피케이션의 중요 징후 중 하나로 고려할 수 있다.

$$\text{전체 점포 수 대비 창폐업 수 (비율)} = \frac{2022\text{년 } 2\text{분기} \sim 2023\text{년 } 1\text{분기 전 구간 창업, 폐업 횟수 합산}}{\{(2022\text{년 } 3\text{분기} + 2023\text{년 } 1\text{분기 영업 중인 점포 수}) \div 2\}}$$

#### [영업기간]

영업기간이 짧을수록 상권이 활성화되면서 업종 변경이 자주 일어나고 있다는 것을 의미한다. 또한 젠트리피케이션이 진행될수록 임대료 상승 등의 이유로 빈번한 업종 변경이 발생하기 때문에 영업기간은 감소하는 양상을 띄게 된다.

$$\text{평균 영업 기간 (개월)} = \frac{\text{2022년 2분기} \sim \text{2023년 1분기에 영업 중인 점포의 영업기간 합}}{\text{2022년 2분기} \sim \text{2023년 1분기에 영업 중인 전체 점포 수}}$$

#### [프랜차이즈 업체]

특정 상권 내에서 일반 점포 대비 프랜차이즈 수가 증가한다는 것은 대규모 자본이 유입되고 있다는 것을 의미한다. 따라서 상권의 프랜차이즈 비율을 살펴보기 위해 해당 변수를 사용한다.

$$\begin{aligned} &\text{전체 점포수 대비 프랜차이즈 업체 수 (프랜차이즈 비율)} = \\ &\frac{\text{2022년 2분기} \sim \text{2023년 1분기 프랜차이즈 업체 수 합산}}{\{(2022년 3분기 영업 중인 점포 수 + 2023년 1분기 영업 중인 점포 수) \div 2\}} \end{aligned}$$

#### [매출액]

매출액 역시 상권의 활성화 정도를 살펴볼 수 있는 변수이다. 따라서 매출액 증가율을

$$\text{평균 매출액 증가율} = \{(\text{2022년 3분기 매출액 합산} \div \text{2023년 1분기 영업 중인 점포 수}) + (\text{2022년 3분기 매출액 합산} \div \text{2023년 1분기 영업 중인 점포 수})\} \div 2$$

#### [임대료]

임대료 상승은 해당 상권의 수요 증가를 의미하며 과한 임대료 상승은 젠트리피케이션을 측정할 수 있는 가장 중요한 변수이다(이진희 외, 2018, p.102). 기존 연구에서는 구할 수 없었지만, 해당 연구에서는 한국 신용 데이터의 사업장별 임대료 데이터를 활용하여 임대료 상승률을 계산하였다.

$$\text{임대료 상승률} = \frac{(\text{2023년 1분기 임대료} - \text{2022년 1분기 임대료})}{\text{2022년 1분기 임대료}}$$

#### [경영위기 증가율]

젠트리피케이션 현상은 지역 정체성 약화, 급격한 임대료 상승으로 인한 비자발적 이주 등으로 인한 장기적인 상권 쇠퇴와 같은 부정적인 문제도 있지만, 젠트리피케이션이 진행됨에 따라 지역 활성화 및 개발, 중산층 유입에 따른 지역 불균형 조정 등과 같은 긍정적인 효과도 존재한다. 따라서 젠트리피케이션 상권을 분류하되, 경영위기가 증가하고 있는 상권을 분리하여 보기 위해 경영위기 증가율 변수를 추가하였다.

$$\text{경영위기 증가율} = \frac{\text{현재 경영위기 비율} - \text{이전 경영위기 비율}}{\text{이전 경영위기 비율}}$$

$$\begin{aligned} \text{현재 경영위기 비율} &= (\text{2022년 1분기 경영위기 비율} + \text{2022년 2분기 경영위기 비율}) \\ \text{이전 경영위기 비율} &= (\text{2022년 4분기 경영위기 비율} + \text{2023년 1분기 경영위기 비율}) \end{aligned}$$

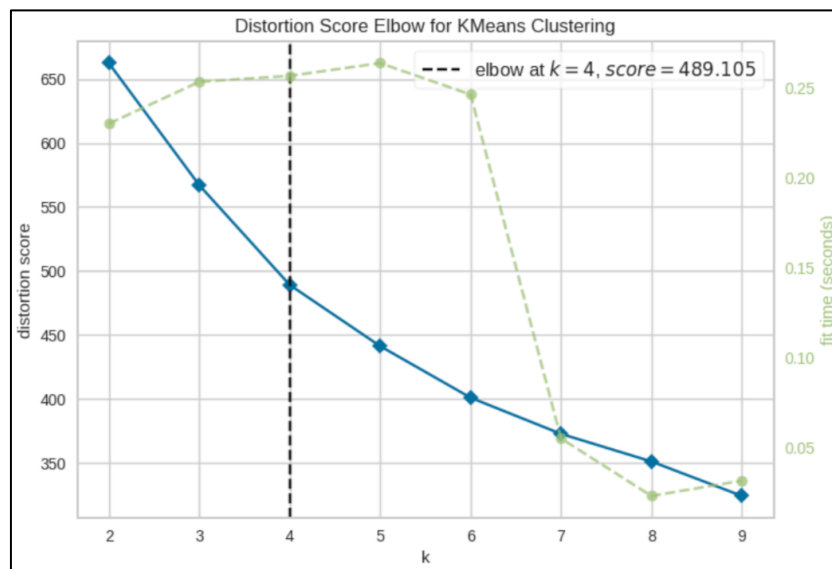


## 2) 클러스터링

기존 연구의 경우 각 변수값의 평균과 표준편차 값을 사용하여 단계별 점수를 구분하고 가중치를 곱하여 나온 값을 기준으로 젠트리피케이션 단계를 구분한다.(이진희 외, 2018, p.105) 하지만 이와 같은 경우 각 상권 사이의 유사성을 살펴보기 어렵고, 전문가를 통해 가중치를 산정해야 된다는 단점이 존재한다. 따라서 본 연구에서는 선정된 변수를 활용하여 K-mean 클러스터링을 수행하여 유사한 상권끼리 군집화 후 젠트리피케이션 상권을 식별하여 정의하고자 한다.

## 3) 최적 군집 개수 결정

총 113개의 상권에 대해 선정된 변수를 계산한 후 다양한 군집 수의 결과를 Elbow Method를 사용하여 비교한 후 최적의 군집 수를 결정한다. 아래의 그림을 통해 k=4일 때를 최적 군집 수로 선정하였다.



## 4) 상권 클러스터링 결과

113개의 상권을 4개의 군집으로 나눈 후 각 군집의 변수에 대한 평균을 나타낸 표는 아래와 같다.

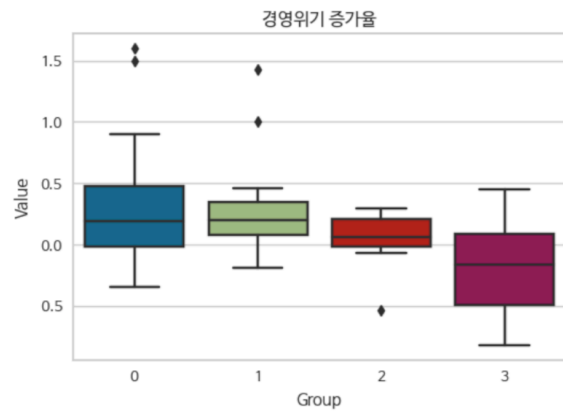
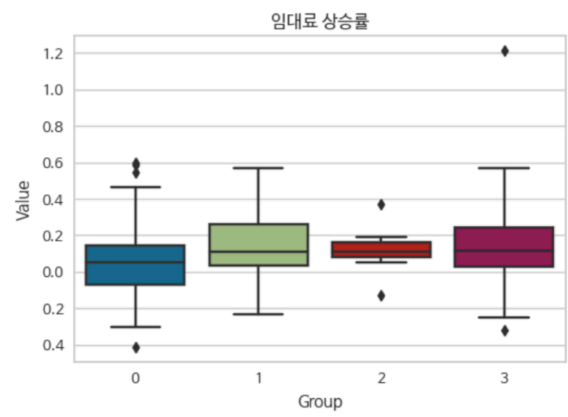
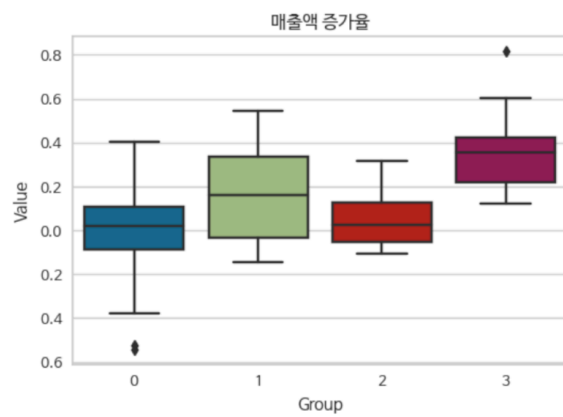
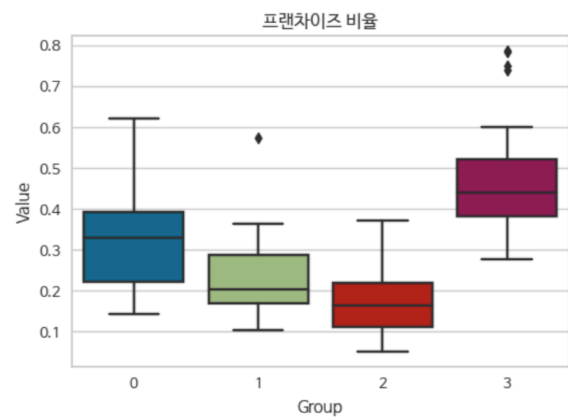
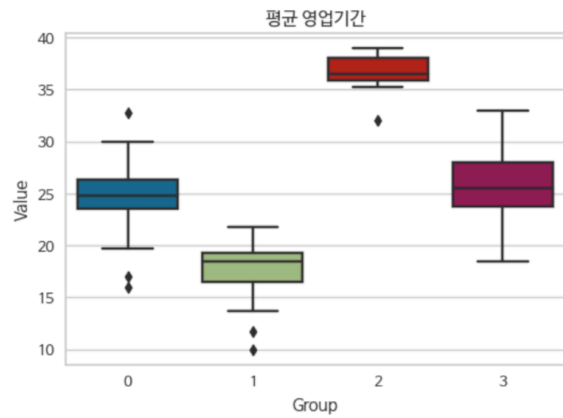
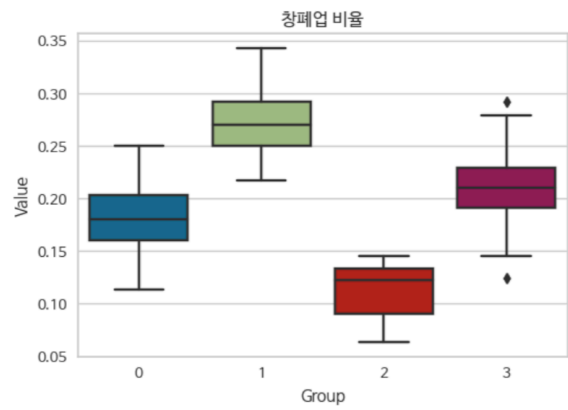
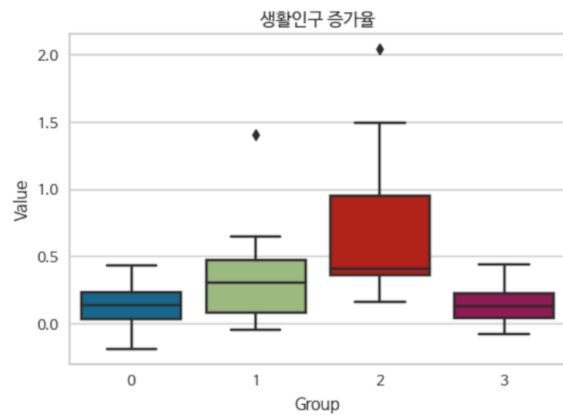
군집	생활인구 증가율	창폐업 비율	영업기간 (개월)	프랜차이즈 비율	매출액 증가율	임대료 상승률	경영위기 증가율
0	0.13	0.18	24.79	0.32	-0.01	0.08	0.26
1	<b>0.34</b>	<b>0.27</b>	<b>17.31</b>	<b>0.24</b>	<b>0.17</b>	<b>0.15</b>	<b>0.30</b>
2	0.75	0.11	36.44	0.18	0.06	0.12	0.03
3	0.15	0.21	25.73	0.46	0.35	0.14	-0.19

분류된 4개의 군집 중 창폐업 비율이 높고, 영업기간이 짧으며, 임대료 상승률과 경영위기 증가율이 가장 높은 군집을 젠트리피케이션 지역으로 설정하였다. 해당 군집을 선정한 가장 큰 이유는 현재 임대료 상승률이 가장 높다는 것은 앞으로 젠트리피케이션 심화 지역이 될 확률이 매우 높

으며, 동시에 경영위기 증가율도 가장 높기 때문에 상권에 대한 주의가 필요한 지역이라 판단하였기 때문이다. 또한 점포의 창폐업 비율과 영업기간이 다른 군집에 비해 몹시 낮기 때문에 소상공인들의 점포 유지와 경영위기에 어려움을 겪고 있는 지역임을 알 수 있다. 따라서 젠트리피케이션 예상 지역으로 분류된 17개의 지역은 아래와 같다. 그 중 실제 샬로수길, 상수역, 성수동카페거리, 연남동, 경의선책거리, 망리단길 등은 젠트리피케이션 현상에 언급되는 지역으로 해당 클러스터링이 젠트리피케이션 여부를 잘 반영하도록 수행되었다는 것을 알 수 있다. 또한 연트럴파크의 경우 임대료 상승률이 음수로 감소했던 지역이지만, 젠트리피케이션이 심화되는 지역임을 수기로 확인할 수 있었다. 따라서 임대료가 상승하지 않더라도 젠트리피케이션 주의 지역일 수 있으며 해당 클러스터 방식은 이를 각 변수를 복합적으로 반영하여 젠트리피케이션 여부 지역 파악한다는 것을 알 수 있다. 또한 젠트리피케이션 지역으로 가장 유명한 가로수길 등이 뽑히지 않은 이유는 생활인구, 매출액 임대료, 경영위기 변수 등이 상승률을 기준으로 보았기 때문에 이미 젠트리피케이션 심화 지역보다는 현상이 두드러지고 있거나 예상 지역 위주로 뽑힌 것으로 판단할 수 있었다.

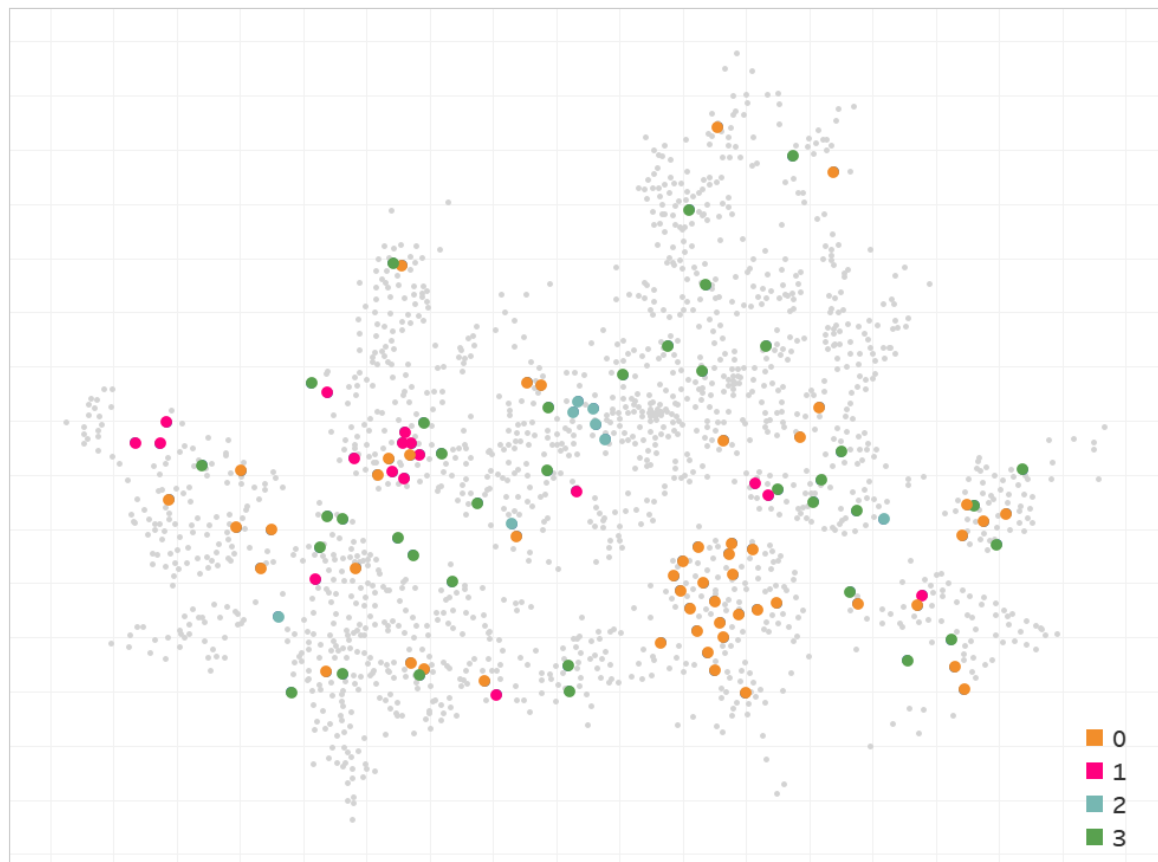
	trdar_nm	생활인구	창폐업	영업기간	프랜차이즈	매출액	임대료	경영위기	predict
72	송파나루역 1번(송리단길)	0.091094	0.240143	19.75	0.184588	0.299274	0.567416	0.047619	1
20	홍대입구역(홍대)	0.302025	0.255896	20.25	0.288008	0.448037	0.140158	-0.037361	1
93	연트럴파크(연남동주민센터)	0.011478	0.291971	15.00	0.102190	-0.143310	-0.173228	0.456790	1
67	샬로수길	0.058422	0.267890	19.25	0.300917	0.077740	0.307420	0.243914	1
66	상수역(홍대)	0.392552	0.325815	17.00	0.203843	0.349072	0.056044	0.200473	1
76	성수동카페거리	0.029604	0.275109	19.00	0.168122	0.318192	0.250000	-0.000904	1
17	연남동(홍대)	0.153126	0.290606	17.25	0.169829	-0.031629	0.042752	0.081904	1
15	경의선책거리	0.194908	0.271090	18.50	0.174408	0.052417	-0.232804	0.393566	1
19	홍대입구역 3번	0.405738	0.250531	16.50	0.186837	0.335789	0.500000	0.153996	1
57	문래예술촌(문래창작촌)	1.402145	0.264121	21.75	0.103421	0.160995	0.082746	0.349367	1
1	상암동상점가	0.642989	0.217166	18.50	0.320579	0.388002	0.399576	0.276336	1
50	마곡역(마곡)	0.551661	0.343149	10.00	0.277189	-0.063950	0.110901	0.334435	1
51	발산역(마곡)	0.440055	0.302266	13.75	0.363350	0.139036	-0.022567	0.178128	1
52	양천향교역 7번	0.467946	0.338259	11.75	0.574713	0.542920	0.030769	-0.186254	1
47	독섬역	0.621926	0.226221	20.75	0.282776	0.247440	-0.008787	1.000000	1
111	해방촌 남동측	-0.044992	0.243993	18.50	0.218115	-0.092345	0.257994	1.430233	1
53	망리단길	0.078818	0.270243	16.75	0.163934	-0.086997	0.185835	0.129283	1

또한 4개 군집의 특징은 아래와 같다.



해당 군집들을 지도에 시각화한 결과는 아래와 같다.

<젠트리피케이션 군집 4개 분류>



## 나. 젠트리피케이션과 경영위기의 가설검정

젠트리피케이션 발생 가능 군집 내 경영위기에 영향을 미치는 요인을 확인하기 위해서는 젠트리피케이션과 경영위기의 관계성 확인이 먼저 필요하다. 젠트리피케이션 발생 가능 여부가 경영위기에 영향을 미치는 파생변수로서 작용할 수 있음을 가설 검증하고, 젠트리피케이션 발생이 경영위기에 부정적인 영향으로 미치는 확인한다.

### 1) 귀무가설의 설정

모집단의 모수에 대하여 검정하고자 하는 집단에 대해 가설을 설정하는 것은 가설검정에서 가장 기본적인 사항이다. 가설은 귀무가설(Null hypothesis)과 대립가설(Alternative hypothesis)으로 설정하며, 가설검정에서 표본관찰 또는 실험을 통하여 귀무가설과 대립가설 중에서 하나를 선택하는 과정이라 할 수 있다(Washington, et. al., pp.28-37).

본 연구에서 검토하게 될 귀무가설을 설정하면 다음과 같다.

귀무가설 I : 젠트리피케이션 여부는 경영위기에 영향을 미치지 않는다.

이러한 귀무가설을 설정은 젠트리피케이션에 따른 경영위기의 차이를 검증하는데 있다. 귀무가설이 채택되면, 젠트리피케이션에 따른 경영위기의 차이는 없다는 의미가 되는 반면, 귀무가설이 기각되면, 젠트리피케이션에 따라 경영위기가 차이가 있다는 것으로 다음 단계의 분석을 진행할 수 있는 근거가 된다.

가설이 성립되기 위해서는 가설검정 전 다음과 같은 가정이 필요하다.

가정 : 군집 1을 젠트리피케이션 발생 가능 상권으로 구성되어있고, 젠트리피케이션의 부정적 영향만을 고려한다.

가설검정에서 등장하는 것은 유의확률(P-value)이며, 이를 통해 가설의 판단 여부를 정량적으로 확인할 수 있다. 유의확률이 유의수준 0.05보다 적으면 95%의 신뢰수준에서 귀무가설을 기각하고, 유의확률이 유의수준 0.05보다 크면 95%의 신뢰수준에서 귀무가설을 채택한다. 즉, 유의확률은 귀무가설을 기각할 수 있는 최소한의 관측 유의수준이라고 말할 수 있다. 이러한 검정결과는 젠트리피케이션과 경영위기 사이의 관계성 파악과 소상공인 경영위기 결정 요인 선택에 중요한 의미를 갖게 된다.

### 2) 귀무가설의 검정

귀무가설 I 검정

‘젠트리피케이션 여부는 경영위기에 영향을 미치지 않는다.’라는 귀무가설은 클러스터링 결과를 기반으로 젠트리피케이션 발생 가능 군집을 1, 아닌 군집을 0으로 구분하여 t-test 가설 검정하였다. 이때 1과 0을 종속변수로, 경영위기비율과 경영위기증가율을 독립변수로 검정을 실시하였고,

젠트리피케이션 여부와 경영위기비율과 경영위기증가율은 상권 단위로 계산되었다.

젠트리피케이션 발생 가능 군집의 경영위기비율 평균은 0.18로 젠트리피케이션 해당 지역이 아닌 군집 0.178과 큰 차이가 없다. 하지만, 경영위기비율이 아닌 경영위기증가율로 볼 때 젠트리피케이션 발생 가능 군집은 0.297로 해당 지역이 아닌 군집 0.07 값과 큰 차이를 보인다.

이는 종속변수를 경영위기비율로 보면 95%의 신뢰수준에서 젠트리피케이션 여부에 따라 경영위기에 차이가 없다는 귀무가설을 기각할 수 없지만, 종속변수를 경영위기증가율로 본다면 귀무가설은 기각될 수 있다.

하지만, 젠트리피케이션은 장기적인 현상으로 보아 경영위기를 경영위기 변화율로 확인하는 것이 더 올바르다고 판단하였고, 2022년 1분기 시점을 기준으로 2023년 1분기 시점까지의 경영위기증가율로 젠트리피케이션과 경영위기의 관계성을 판단하였다.

즉, 젠트리피케이션에 따라 경영위기증가율에 차이를 보이며 젠트리피케이션은 경영위기에 영향을 미친다고 결론 내릴 수 있으며, 전체 평균과 boxplot 시각화 결과를 살펴보면 젠트리피케이션 발생 가능 지역(군집1)이 아닌 지역(군집0)보다 경영위기증가율 평균이 더 높아지는 모습을 확인할 수 있었다. 이는 젠트리피케이션이 발생할수록 경영위기에 부정적인 영향을 미친다고 볼 수 있다.

<표 1> 귀무가설 I (젠트리피케이션과 경영위기증가율)의 검정 결과

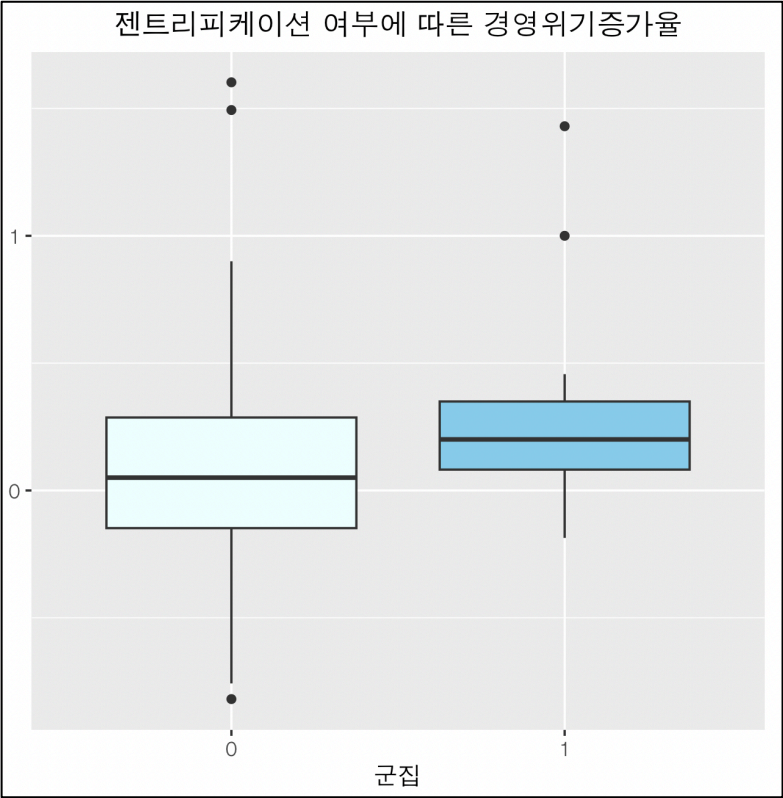
(a) 젠트리피케이션 여부에 따른 경영위기비율과 경영위기증가율

젠트리피케이션 구분	전체 평균	
	경영위기 비율	경영위기 증가율
1	0.18	0.297
0	0.178	0.07

(b) 가설 검정결과

젠트리피케이션 구분	전체 평균	
	경영위기 비율	경영위기 증가율
t	-0.072	-2.186
자유도	22	22
유의확률(P-value)	0.942	0.039

(c) 시각화 – boxplot



## 다. 젠트리피케이션 여부에 유의미한 영향 미치는 변수 선택

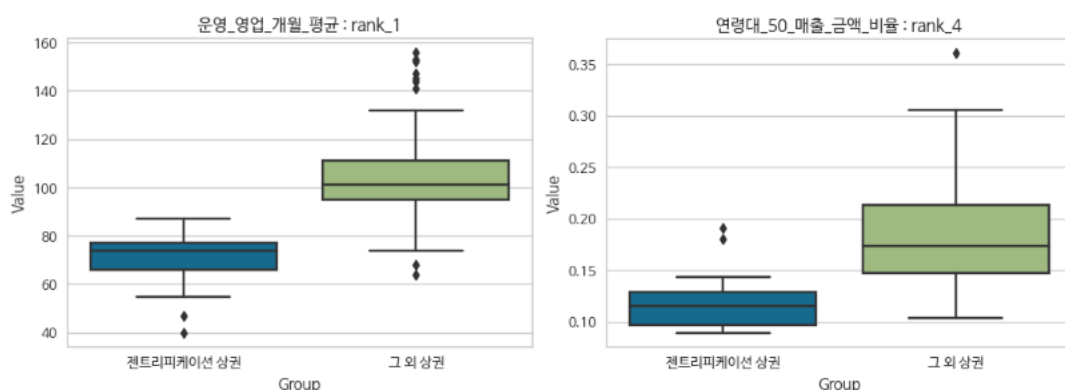
### 1) 고려 가능한 모든 변수 구하기

위의 과정을 통해 분석하고자 하는 경영위기 증가율이 높은 젠트리피케이션 예상 상권들을 선정하였다. 다음으로 해당 상권들과 나머지 상권들이 추가적으로 어떤 특징을 보이는지 알아보기 위해 한국 신용 데이터, 10개의 상권 특성 데이터와 클러스터링 구축 시 사용한 변수를 모두 결합하여 데이터를 구성하였다. 또한 상권 특성 데이터를 변환하여 2022년 1분기에서 2023년 1분기의 증가율을 볼 수 있도록 변환하여 추가하였다. 따라서 113개의 상권에 대한 569개의 변수를 구할 수 있었다.

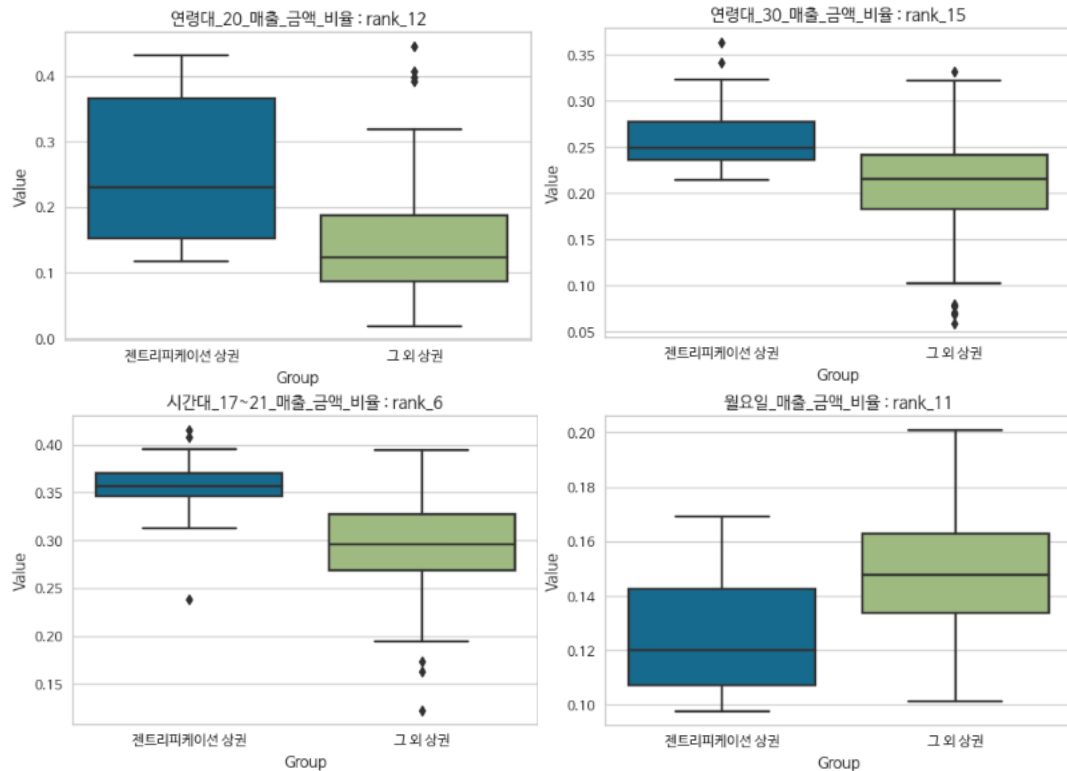
### 2) Anova를 통해 평균 차이 살펴보기

총 569개의 변수에 대해 젠트리피케이션으로 선정된 상권들과 아닌 상권들의 평균 차이를 검정하였다. 그 결과 평균 차이에 영향을 미치는 124개의 변수를 확인할 수 있었다. 아래는 유의성이 높은 일부 변수이며, 그림을 통해 간단히 특징을 파악할 수 있다. 예시로,

아래의 변수들을 통해 간단히 특징을 파악할 수 있다. 예시로, “운영 영업 개월 평균”이 유의미하게 작은 것을 보아 다른 상권에 비해 업종 변경과 창폐업이 활발히 일어나는 지역임을 알 수 있다. 또한 “연령대 50 매출 금액 비율”이 낮고, “연령대 20, 30 매출 금액”이 높은 것을 보아 젊은 세대가 주로 방문하는 상권임을 알 수 있다. 또한 월요일 평균값이 몹시 낮고, 시간대 17~21시 사이의 매출 비율이 높은 것을 통해 평일보다는 주말에, 점심보다는 저녁에 소비가 높다는 것을 알 수 있다.







### 3) Boruta 알고리즘을 통한 변수 선택

위의 과정을 통해 켄트리피케이션 여부에 유의미한 변수 124개를 확인할 수 있었지만, 설명력을 높이기 위해서는 변수의 개수를 제한할 필요가 있다. 따라서 Boruta feature selection을 사용하여 변수의 개수를 줄이고자 한다. Boruta feature selection이란 변수 중요도를 기반으로 원래 변수와 그림자 변수 차이를 비교 반복하여 변수 축소에서 안정적인 결과를 얻을 수 있다는 장점이 있다.

Boruta 알고리즘을 통한 변수 선택 결과, 선정된 변수는 총 14개이며 설명은 아래와 같다. 예시로, 운영\_영업\_개월\_평균은 켄트리피케이션 여부를 구분할 수 있는 유의미한 변수이며, 켄트리피케이션 예상 군집은 다른 상관들의 군집에 비해 평균 개월 수가 짧다.

변수명	설명	변수명	설명
운영_영업_개월_평균	낮음	월요일_매출_금액_비율	낮음
평균_영업기간	낮음	여성_매출_금액_증가율	큼
창폐업 비율	큼	목요일_매출_금액_증가율	큼
연령대_50_매출_금액_비율	낮음	시간대_17~21_매출_금액_증가율	큼
age	낮음	주중_매출_금액_증가율	큼
시간대_17~21_매출_금액_비율	큼	연령대_60_이상_매출_금액_비율	낮음
폐업_영업_개월_평균	낮음	시간대_14~17_매출_금액_증가율	큼

## 라. 젠트리피케이션 예상 지역 예측

본 연구의 목표 중 하나는 소상공인들에게 자신이 속한 상권에 대한 젠트리피케이션에 대한 정보를 제공하고 운영 관리를 위한 정보로 활용할 수 있도록 하는 것이다. 따라서 단순히 젠트리피케이션 여부에 미치는 변수를 파악하여 해석하는 것뿐만 아니라, 위험 지역을 빠르게 식별할 수 있는 예측 모델을 제안하고자 한다.

### 1) 랜덤포레스트를 활용한 젠트리피케이션 예측 모델

앞서 과정에서 도출된 변수 14개를 활용하여 모델을 구축하였다. 스케일링과 같은 기본적인 전처리를 수행하였으며, 행의 개수가 적고 타겟 분포가 불균형하다는 한계점을 극복하기 위해 오버샘플링과 같은 방법을 적용하며 최적의 모델을 선정하였다. 모델 수행 결과 0.96으로 높은 정확도를 가진 젠트리피케이션 예측 모델을 구축할 수 있었다. 예측 결과는 아래와 같으며, 7개의 젠트리피케이션 예상 상권에서 1개를 제외한 6개의 상권에서 올바르게 예측한 것을 확인할 수 있다.

	0	1
0	38	1
1	1	6

### 2) 새로운 상권 적용

실제로 위의 모델을 범용적으로 사용하기 위해서는 주어진 상권이 아닌 새로운 상권에 대한 예측 결과를 검토해야 한다. 따라서 새로운 상권에 대한 정보를 얻을 수 없는 한국 신용 데이터로 이루어진 변수를 제외하고 상권 특성 데이터로만 변수를 구성하였다. 먼저 기존의 젠트리피케이션 여부 라벨링을 부여한 113개의 상권을 사용하여 기존의 방식과 동일하게 boruta feature selection을 진행하였다. 하지만, 한국 신용 데이터를 포함하지 않게 되면서 기존의 선택된 변수들과 개수와 달리 총 15개의 변수가 선택되었다. 모델 수행 결과 정확도는 0.975로 총 6개의 젠트리피케이션 상권 중 한 개의 상권에 대해 오분류한 것을 확인할 수 있다. 만약 전체 상권에 대한 한국 신용 데이터를 적요할 수 있다면, 더 높은 예측 결과를 얻을 수 있을 것으로 예상된다.

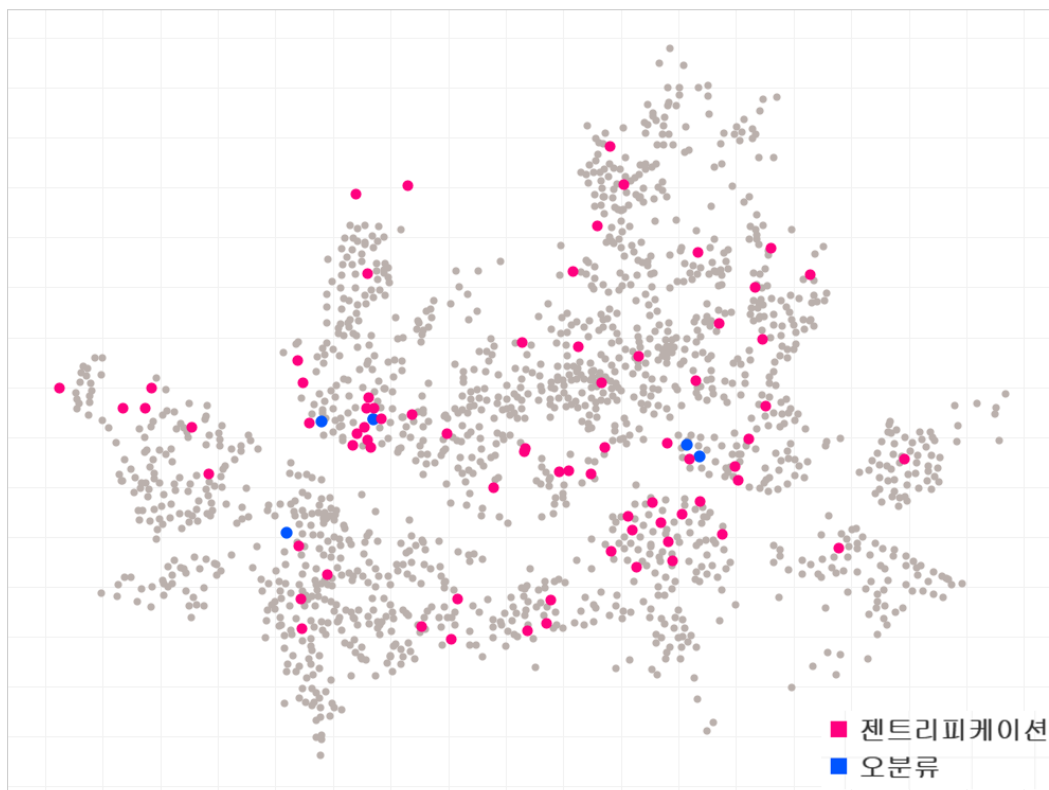
	0	1
0	34	0
1	1	5

다음으로, 새로 구축된 모델을 통해 젠트리피케이션 예상 지역을 예측된 상권은 아래와 같다. 총 1620개의 상권 중 75개의 상권에 대해 젠트리피케이션 지역으로 예측하였으며, 추출된 상권의 예시는 아래와 같다. 서울숲 카페거리, 상수역 등은 젠트리피케이션으로 언급되는 지역으로 새롭게 예측한 지역이 젠트리피케이션 위험 지역을 잘 반영한다는 것을 확인할 수 있다.

새롭게 예측된 젠트리피케이션 상권	기존에 정의한 젠트리피케이션 지역을 잘못 예측한 상권	
서울 숲 카페거리 상수역 3번 경의중앙 신촌역 독섬 유원지역 1번 성주 1동 주민센터 강남언주로 . . (총 75개)	맞는 지역을 아니라고 예측	아닌 지역을 맞다고 예측
	독섬역 망리단길 문래예술촌(문래창작촌) 성수동 카페거리	홍대 걷고 싶은 거리

예측한 젠트리피케이션 예상 지역과 오분류 지역에 대한 시각화는 아래와 같다.

<젠트리피케이션 예측>



## 마. 젠트리피케이션 군집 내 경영위기

### 1) 경영위기에 따른 anova 검정

본 연구에서는 앞서 젠트리피케이션 상권이 경영위기에 영향을 미친다는 가정을 검정하였다. 또한 선택된 변수에 의해 젠트리피케이션 예상 상권과 이외의 상권들 간의 특성 차이가 존재하는 것을 확인하였다. 이는 젠트리피케이션 예상 상권 내 경영위기에 영향을 미치는 변수를 파악할 수 있는 근거가 되며 다음 단계에서 젠트리피케이션에 영향을 미치는 동시에 경영위기에도 영향을 미치는 변수를 확인할 수 있다.

젠트리피케이션 예상 상권에서 경영위기에 따른 각 변수 별 평균차이를 살펴보기 위해 디사일로에서 클러스터링 군집별로 anova 검정을 수행하였다. 유의확률(P-value)이 0.05보다 작아 경영위기에 유의미하다고 선정된 변수의 우선순위 및 결과는 아래와 같다.

모든 군집에 영향	(젠트포함) 3개 군집에 영향	(젠트포함) 2개 군집에 영향	젠트에만 영향	젠트군집에만 영향 안 미침	경영위기에 안미치는 젠트	젠트리피케이션 여부
z_sum_sales_card2 z_sum_weekend_sales_card sum_weekend_sales_card sum_sales_card z_monthly_rental_fee z_sum_purchase_invoice sum_new_customer_cnt z_sum_purchase_cash sum_customer_cnt 매입_영업_개월_평균 아파트_면적_165_제곱미터_세대_수 버스_정거장_수 regular_employees_count 연령대_30_매출_비율 연령대_40_매출_비율 소득_구간_코드 class_1_name (입종대분류) class_2_name (입종중분류) trdar_nm (상권_코드_명) rental_type 남성연령대_20_상주인구_수 아파트_가격_1_억_미만_세대_수 아파트_면적_66_제곱미터_세대_수 아파트_평균_면적 시간대_14~17_매출_비율	남성_매출_비율 운영_영업_개월_평균 연령대_60_이상_매출_비율 관공서_수 age 월요일_매출_비율 시간대_21~24_매출_비율 여성_매출_비율 z_sum_sales_delivery 목요일_매출_비율 z_sum_weekend_sales_deliv ery2 상권_구분_코드_명 상권_변화_지표_명 연령대_50_매출_비율 sum_sales_delivery 화요일_매출_비율 z_sum_sales_invoice monthly_rental_fee z_sum_purchase_card2 sum_weekend_sales_delivery 아파트_가격_2_억_세대_수 sum_purchase_card 개업_점포_수	수요일_매출_비율 개업_율 아파트_가격_6_억_이상_세대_수 여성연령대_50_직장_인구_수 매입_율 소득_구간_코드 sum_purchase_cash 연령대_50_직장_인구_수 아파트_가격_5_억_세대_수 business_square_size 시간대_06~11_매출_비율 남성연령대_60_직장_인구_수	대학교_수 시간대_11~14_매출_비율 남성연령대_30_직장_인구_수 남성연령대_40_직장_인구_수 총_직장_인구_수	연령대_20_상주인구_수 아파트_면적_132_제곱미터_세대_수 집객시설_수 금요일_매출_비율 남성연령대_10_직장_인구_수 sum_purchase_invoice 연령대_10_상주인구_수 is_franchise (프랜차이즈여부) 남성연령대_10_상주인구_수 z_rental_deposit 연령대_10_매출_비율 여성연령대_20_상주인구_수 월요일_매출_비율 아파트_가격_3_억_세대_수 rental_deposit 여성연령대_50_상주인구_수 연령대_20_매출_비율 숙박_시설_수	시간대_17~21_매출_금액_비율	운영_영업_개월_평균 영원가관 창매입 연령대_50_매출_금액_비율 age 시간대_17~21_매출_금액_비율 매입_영업_개월_평균 월요일_매출_금액_비율 여성_매출_금액_증가율 목요일_매출_금액_증가율 시간대_17~21_매출_금액_증가율 주중_매출_금액_증가율 연령대_60_이상_매출_금액_비율 시간대_14~17_매출_금액_증가율

초반 클러스터링 결과를 이용하여 군집별 경영위기에 영향을 미치는 변수를 확인하였고, 최종적으로 모든 군집 혹은 3개 군집에 영향을 미치는 변수를 일반적인 경영위기 요인으로 여겼다. 모든 변수에 대한 요인분석을 수행하기 앞서 군집 간 경영위기에 미치는 요인과 그 요인의 효과크기가 다를 수 있음을 먼저 파악하였고, 이는 군집 별 요인분석에 대한 수행 이유로서 타당한 근거가 될 수 있다.

예를 들어, 연령대\_30\_매출\_비율은 모든 군집(0, 1, 2, 3)에서 경영위기에 영향을 미치고, 연령대\_30\_매출\_비율이 작아질수록 경영위기 평균 비율은 커지는 쪽으로 방향성은 같지만, 경영위기에 미치는 변수의 효과크기는 군집마다 다르다.

변수명	군집명	설명	효과크기
연령대_30_매출_비율	0	낮을수록	0.338
	1	낮을수록	0.305
	2	낮을수록	0.531
	3	낮을수록	0.272

\*효과크기는 그룹들 사이의 차이를 얼마나 의미있는지 나타내는 지표로 Cohen's *d*을 이용하여 측정하였다.

$$\text{Cohen's } d = \frac{(\text{그룹1의 평균} - \text{그룹2의 평균})}{\text{전체 표준편차}}$$

즉, 군집마다 경영위기에 미치는 변수가 다를 수 있으며, 경영위기에 영향을 미치는 동일한 변수 여도 군집 간 요인 효과 크기는 다르기에 군집별로 보는 것이 정확한 경영위기 결정 요인을 찾는 데 필요한 과정이라고 판단된다.

그 결과, 4개 군집 혹은 3개 군집과 같이 다수 군집에서 경영위기에 영향을 미치는 요인을 일반적인 경영위기 요인으로 보고 일반적인 경영위기 요인과 젠트리피케이션에 영향을 미치는 요인을 다음 단계에서 비교해보고자 한다.

## 2) 최종 변수 선정

본 연구에서는 젠트리피케이션과 경영위기에 모두 영향을 미치는 변수를 우선 선정하고 해당 변수를 통해 젠트리피케이션 내 경영위기를 확인해 보고자 한다.

변수	젠트리피케이션이다	일반적인 경영위기이다	F값	P값
운영_영업_개월_평균	낮을수록	낮을수록	85.46478854	2.48E-20
age	낮을수록	낮을수록	27.27307351	1.80E-07
폐업_영업_개월_비율	낮을수록	낮을수록	107.5442262	3.65E-25

변수	젠트리피케이션이다	일반적인 경영위기이다	F값	P값
연령대_50_매출_금액_비율	낮을수록	클수록	11.21859515	0.0008105981
월요일_매출_금액_비율	낮을수록	클수록	-	-

변수	젠트리피케이션 여부	F값	p값
여성_매출_금액_증가율	높을수록	12.294834	6.565401e-04
시간대_14~17_매출_금액_증가율	높을수록	7.306777	7.944968e-03
창폐업 (비율)	높을수록	61.290362	3.154783e-12
시간_17~21_매출_금액_비율	높을수록	22.162339	7.285122e-06

운영\_영업\_개월\_평균, age, 폐업\_영업\_개월\_비율, 여성\_매출\_금액\_비율, 시간대\_14~17\_매출\_금액\_증가율은 젠트리피케이션과 경영위기에 영향을 미치는 변수로서 모두 낮을수록 젠트리피케이션 발생 가능 군집에 가까워지며, 경영위기 또한 가까워진다. 이는 해당 변수가 낮아질수록 젠트리피케이션 발생 가능 상권에 해당될 수 있으며, 일반적인 경영위기에 속한 상권으로도 해석될 수 있다.

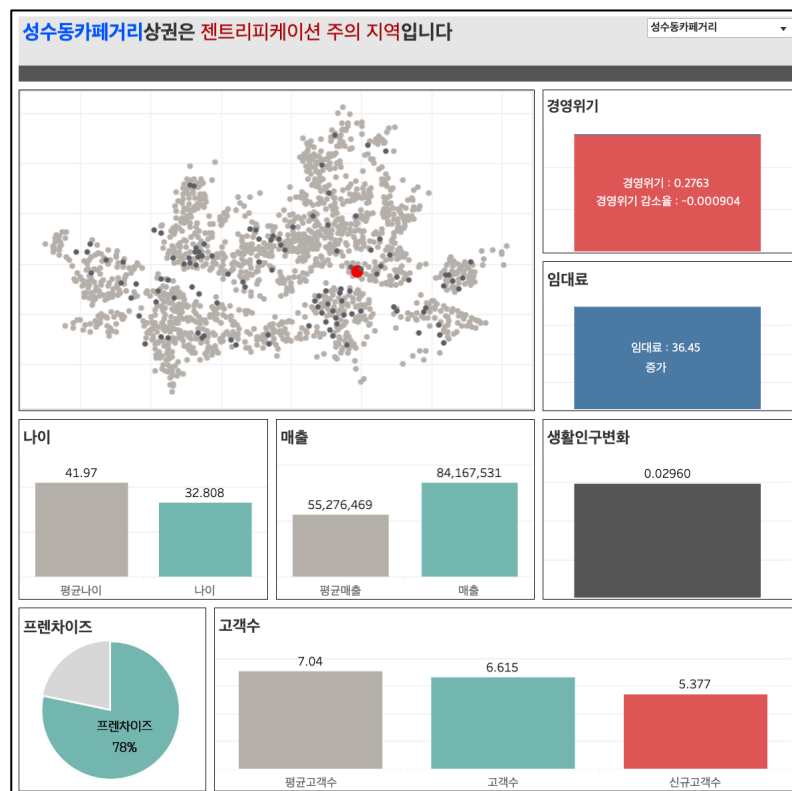
반면에, 젠트리피케이션과 경영위기에 영향을 미치지만, 다른 방향성을 가지는 변수도 있다. 예를 들어, 연령대\_50\_매출\_금액\_비율은 낮을수록 젠트리피케이션에 속하지만, 일반적인 경영위기에는 해당되지 않을 수 있다.

마지막으로 4개의 변수는 젠트리피케이션에 유의미한 영향을 미치는 변수로 설정하였다. 이는 젠트리피케이션의 여부가 당장의 경영위기에는 영향을 미치지 않을 수 있어도 장기적으로는 상권을 경영위기로 만들 수 있기 때문이다. 따라서 경영위기에 영향을 미치는 변수와 젠트리피케이션을 분류하는데 영향을 미치는 변수를 합하여 총 9개의 변수를 제시하였다.

### 3. 활용방안

#### 가. 예비 창업주 및 사업주를 위한 젠트리피케이션 대시보드 제작

현재 서울 신용 데이터에서 소상공인을 위해 제공하고 있는 “캐시노트” 서비스에서는 업장의 매출과 비용을 관리할 수 있는 서비스를 제공하고 있으며, 자신의 지역 내 다른 사업장의 평균 데이터를 살펴볼 수 있다. 또한 소상공인 진흥공단에서 제공하고 있는 상권현황 서비스에서는 총인구수, 주요시설수, 집객시설수 등의 정보를 통해 상권에 대한 정보를 제공하고 있다. 하지만 개별로 구성된 변수는 초기 창업자에게는 복합적인 해석에 어려움이 있을 수 있으며, 해당 변수들로 인한 파생되는 여러가지 상황을 고려하지 못한다는 단점이 존재한다. 따라서 소상공인들이 창업을 고려하고자 하는 상권이나 현재 운영 중인 점포를 운영 및 관리하는데 젠트리피케이션 정보를 활용할 수 있도록 대시보드 제작하는 것을 제안하고자 한다. 아래 그림은 본 연구에서 분석된 결과를 기반으로 구성한 대시보드이다.

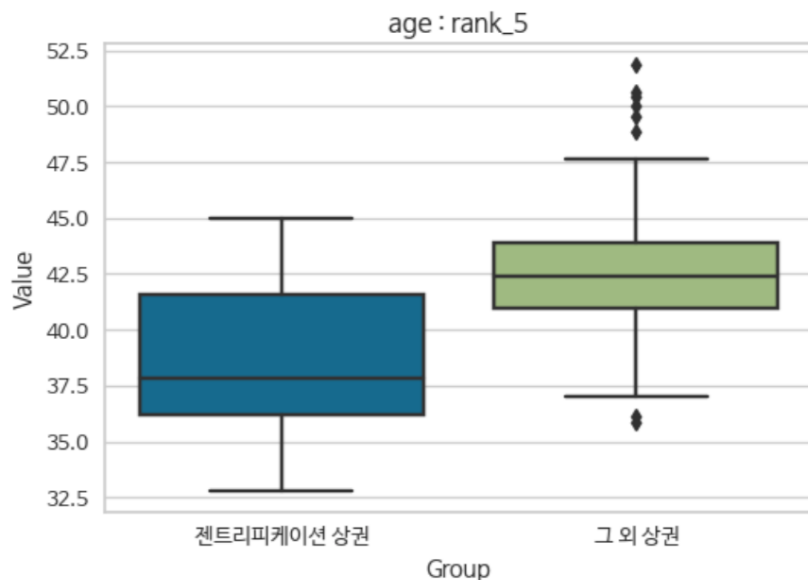


대시보드 링크 : [https://public.tableau.com/app/profile/.63303526/viz/\\_16957916214770/2?publish=yes](https://public.tableau.com/app/profile/.63303526/viz/_16957916214770/2?publish=yes)

위의 대시보드처럼 사업주가 현재 속한 상권이 젠트리피케이션인지 아닌지, 추후 더 나아가 젠트리피케이션 단계까지 구분하여 보여준다면, 해당 정보를 활용하여 창업 여부나 위기에 대응하는 행동을 취하는데 도움을 줄 것이다. 또한 젠트리피케이션 정보와 해당 사실에 대한 인지는 지역 상권 당사자들끼리의 협력을 강화하는데 기초로 사용될 수 있다. 현재는 기존의 상권에 대해서만 표시하였지만, 예측 모델을 활용하여 예측된 상권의 정보를 활용한다면 더 범용적으로 사용할 수 있을 것이다.

## 나. 청년 사업주 법률 자문 및 상담 지원

젠트리피케이션 예상 상권으로 지정된 지역에서의 점주의 평균 연령은 37.4세로 그 외 상권인 41.6세보다 유의미하게 낮다. 또한 젠트리피케이션 상권 내에서 점주의 연령이 낮을수록 경영위기가 높아진다는 것을 앞서 확인하였다. 이는 젠트리피케이션 상권 내에 상대적으로 어린 나이의 점주들이 많이 존재하고 경영위기를 겪을 확률이 높다는 것을 의미하며, 미래의 어린 창업주에게도 문제를 야기하게 될 것이다.



또한 임대료 상승의 원인으로 임차인 정보 부족으로 인한 보호제도 미활용과 상가 임대차법 미적용 대상인 신규 임대차계약으로 조사되었다.(성동저널, 2023) 이는 현재 상권에 존재하고 있는 사업주의 제도에 대한 인지와 법률 자문이 필요하다는 것을 의미하며, 젠트리피케이션 예상 상권 집입을 위한 신규 창업주에게도 어려움이 임대료에 관한 법률적인 어려움이 존재한다는 것을 의미한다. 따라서 어린 나이로 법률적인 지식이 없고 이에 대한 컨설팅 및 멘토링이 필요한 소상공인을 대상으로 무료로 법률 자문 및 컨설팅을 진행하는 지원사업을 제안하고자 한다.

구제 대상의 정확한 범위는

- A. 법률 자문 및 지원금 보조 : 젠트리피케이션 상권에 속해 있으면서 임대료 증가율이 양수인 점포의 점주
- B. 법률 자문 및 컨설팅 : 젠트리피케이션 상권에 속해 있으면서 점주의 연령이 평균보다 어린 경우
- C. 법률 자문 : 젠트리피케이션 상권의 예비 창업자

현 아이디어는 평균 나이보다 어린 점주를 대상으로 하였지만, 젠트리피케이션이 심화되고 있는 상권에서 법률 자문 및 컨설팅은 모두에게 필요할 것이다. 따라서 금액과 대상을 조절하면서 최대한 지원 대상을 늘리고자 하는 노력이 필요하다.

현재 사업주와 예비 창업자들을 대상으로 하는 컨설팅과 멘토링 지원 등은 흔히 찾아볼 수 있다.



하지만 해당 지원은 특정 사업자에게만 집중되어 해당 상권의 전체적인 변화를 꾀하기 어렵다는 단점이 존재한다. 따라서 넓은 대상의 법률적인 자문과 법률 교육을 제공함으로써 사업주의 법적 문제 및 제도에 대한 관심을 함양 시키고 자생적으로 비즈니스를 관리하도록 도움을 줄 것이다. 이는 잠재적인 법정 문제를 줄일 수 있을 뿐만 아니라 비즈니스의 장기적인 지속 가능성도 향상시킬 수 있을 것이다.

## 4. 주요 결과 및 시사점

### 가. 주요 결과 요약

본 연구는 기존의 젠트리피케이션 진단에 사용되는 변수를 활용하여 K-means clustering을 진행하고 젠트리피케이션 가능 지역을 선정하였다. 다음으로 기존 변수 외에 젠트리피케이션 상권과 다른 상권 사이에 유의미한 차이를 보이는 추가적인 변수를 살펴봄으로써 해당 상권의 특징을 파악하고 정의하였다. 또한 124개의 후보 변수 중 Boruta feature selection 알고리즘을 사용하여 가장 유의미한 영향을 미치는 14개의 변수를 통해 예측 모델을 구축하였고, 0.96이라는 높은 정확도로 젠트리피케이션 지역을 예측할 수 있었다. 또한 기존의 상권 외에 새로운 상권에 해당 모델을 적용해 보고 검토함으로써 모델의 타당성을 입증할 수 있었다. 이를 통해 젠트리피케이션 대시보드 아이디어를 실현 가능성을 높일 수 있을 것이며, 소상공인의 상권에 대한 이해 증진과 상권에 도움을 줄 것으로 예상된다. 또한 젠트리피케이션 상권 내의 경영위기 지표에 대한 유의성 검정을 통해 해당 지역의 어린 점주들이 경영위기를 겪는 비율과 확률이 높은 것을 확인하였다. 이를 해결하기 위한 “어린 점주 법률 자문 및 컨설팅” 지원 사업을 제안하였으며, 젠트리피케이션의 고질적인 임대료 상승 문제를 장기적으로 완화할 수 있을 것이라 기대한다.

### 나. 의의

본 연구는 상권 별 단위의 분석, 클러스터링을 통한 상권 분류, 높은 성능의 예측 모델링 구축면에서 기존 연구와의 차별화된 의의를 갖는다.

먼저 기존의 연구(이진희, 2019, p.11)는 상권을 분류하지 않고 행정동 단위의 데이터로 젠트리피케이션을 진단하였다. 하지만 상권 데이터 상 제기동의 경우 동 안에 17개의 상권이 존재하므로 행정동 단위의 진단은 부정확할 수 있다. 따라서 본 연구는 상권별로의 데이터를 수집 후 변수를 계산하여 좁은 범위의 상권에 대해 개별적으로 살펴볼 수 있다.

다음으로 클러스터링을 통한 상권 분류이다. 기존 연구의 경우 평균과 표준편차로 계산하는 방식을 사용하여 변수 간의 유사성을 살펴보기 어렵고, 전문가를 통해서 가중치를 산정해야 된다는 한계점이 존재하였다. 하지만 본 연구에서는 선정된 변수에 K-means 클러스터링 방식을 사용하였으며, 변수의 특성이 유사한 상권끼리 분류되어 더 쉽게 해석하고 파악할 수 있다는 장점이 있다.

마지막으로 예측 모델링의 구축이다. 기존의 경우 임대료 상승이나 매출액 증가 등과 같이 특정 변수만을 중점적으로 고려하였다. 하지만 젠트리피케이션에 영향을 새로운 변수 (ex:여성 매출 증가율)등과 같은 숨겨진 변수를 찾고 해당 변수로 복합적으로 고려함으로써 젠트리피케이션이 심

화되기 전 단계의 상권을 예측할 수 있는 모델을 구축할 수 있다. 또한 이는 한국 신용 데이터의 “캐시노트”나 소상공인 진흥공단의 “상권분석” 서비스에 도입하여 상권의 특성을 보여주는 새로운 정보로 활용할 수 있다.

## 다. 한계점 및 향후 발전 방향

본 연구는 코로나 시기로 인한 데이터 오염, 유동인구와 상주인구 데이터 수집의 어려움으로 인한 2가지 한계점을 가지고 있다.

첫번째로, 해당 시기의 데이터는 2022년 1월부터 2023년 4월까지의 데이터로 코로나가 완화되며 유동인구와 상점의 매출액이 자연스럽게 증가할 수 밖에 없는 시기이다. 따라서 재택 근무 등이 풀리면서 직장 밀집 지역 혹은 중심지의 유동인구와 매출액이 지나치게 상승하는 현상을 보였다. 따라서 젠트리피케이션 상권의 고유한 매출액 증가나 임대료 상승률을 볼 수 없었다는 한계점을 갖고 있다.

두번째로는 상주인구, 생활인구에 대한 데이터 수집의 어려움이다. 상권별로 데이터를 수집하는 과정에서 상주인구 데이터는 부정확한 경우가 많았다. (우선순위가 높지 않아 클러스터링 과정에서 제외함.) 또한 유동인구 데이터는 개인 수준에서 구할 수 없기 때문에 생활인구로 대체하여 분석을 진행하였다. 만약 이후 정확한 상주인구와 유동인구 데이터를 사용한다면 젠트리피케이션이 심화되면서 기존 거주민이 이탈하는 현상까지도 살펴볼 수 있을 것이다.

하지만 본 연구에서는 데이터 수집의 어려움에도 상권 별 클러스터링을 통해 새로운 젠트리피케이션 분류 방식과 예측 모델을 구축하는 프레임워크를 제시하였다는 점에서 의의가 있다. 또한 단순히 당장의 경영위기에 크게 영향을 미치는 변수를 분석하는 것이 아닌 젠트리피케이션이 경영위기를 미칠 수 있다는 점을 고려하여 대시보드나 법률 상담 아이디어를 제시하며 장기적으로 해당 아이디어가 젠트리피케이션 현상 완화에 긍정적인 효과를 미칠 것으로 예상한다.

## 5. 참고문헌

서울시 상권분석 시스템을 통한 젠트리피케이션 정책적 대응 (이진희, 2019)  
도시경제기반형 젠트리피케이션 지표개발 및 활용방안 연구 (이진희 외, 2018)  
소상공인과 지역상권 상생 및 활성화에 관한 법률의 주요 이슈 (장은정, 2023)  
[기획] 지속가능한 도시... 성동구, 젠트리피케이션 정책 재추진(2023)  
<https://www.seongdongnews.com/news/articleView.html?idxno=28653>