

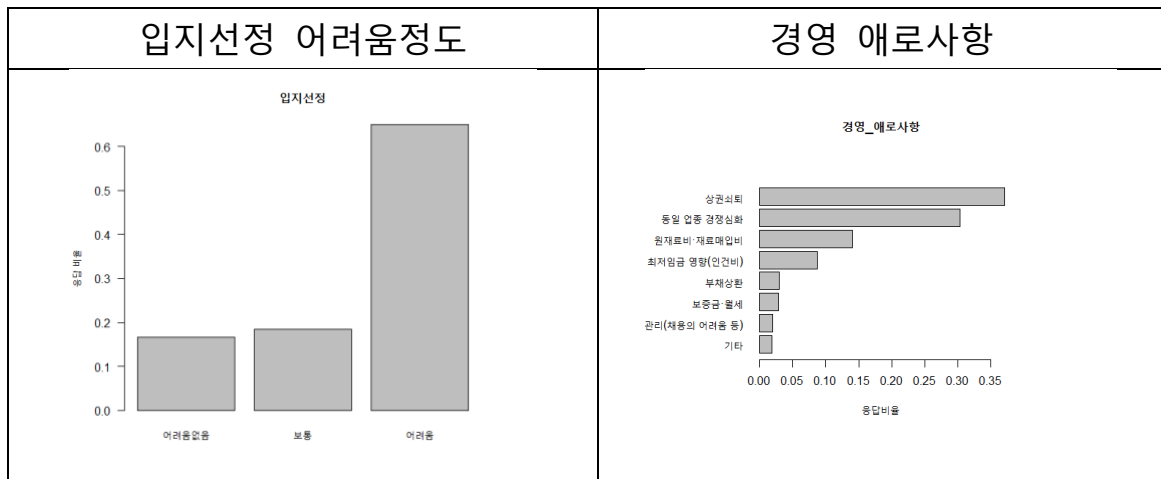
연령대별 매출패턴을 활용한 골목상권 활성화 방안

□ 주제 선정(소제목)

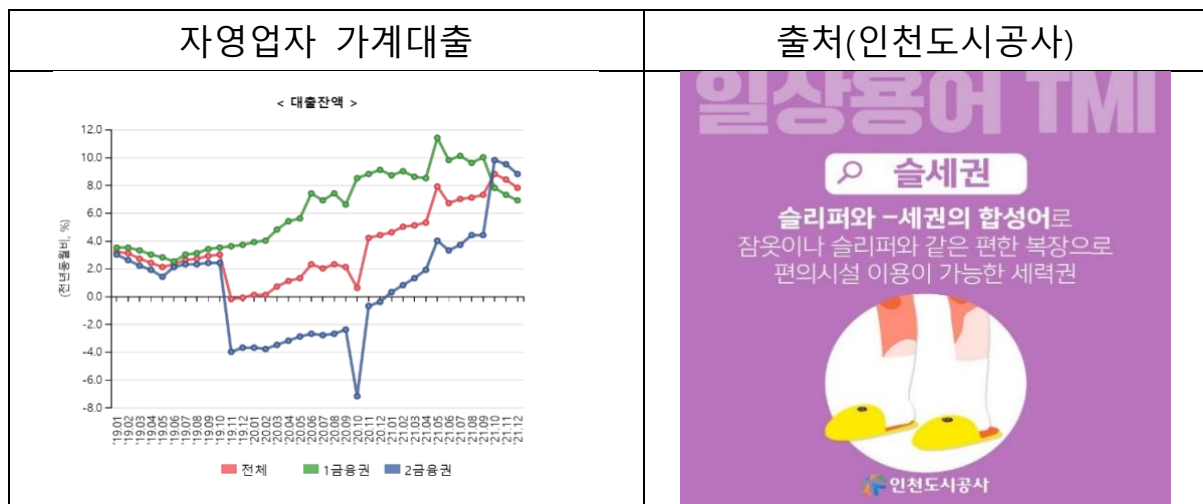
- 서울시 내 골목 상권들에 대해 연령대별 매출 모델을 만들어 매출에 유의미한 영향을 끼치는 지표들을 발견하고 이를 해석, 매출패턴을 통해 골목상권들을 군집화 한다. 해당 골목들에 대해 앞서 발견한 연령대 별 매출 지표들로 활성화가 필요한 골목을 선정하여 맞춤 서비스를 제안한다.

□ 분석 필요성(문제점) 및 전략

상권은 상업활동이 이루어지는 공간으로 우리의 삶에 다양한 형태로 존재하고 있다. 상권에서 입지, 동종업계와의 경쟁, 임대료 등 고려할 점이 많다. 2019년도 소상공인 실태조사 설문 분석 결과, 현재 사업하는 소상공인들의 가장 큰 애로사항은 상권의 쇠퇴임을 알 수 있다. 최근 거리두기가 전면 해제되어 일상회복이 진행중이지만, 이로 인한 경기침체로 누구보다 큰 피해를 입었던 상권 내 소상공인들에게는 아직 역부족인 상태이다. 자영업자 가계대출 동향을 보면 코로나 전 매달 1~2퍼센트의 사이의 대출고객 상승률이 코로나 이후 약 4~8퍼센트로 2~4배 정도 꾸준히 증가한 것을 볼 수 있다.¹ 특히 발달된 대형 상권이 아닌 골목 내 상권에는 코로나로 인한 타격이 매우 컸으며, 정부에서는 이를 지원하기 위해 다방면으로 정책을 펼치고 있다. 물론 코로나로 인해 먼 거리에 위치한 대형 유통업체 및 마트보다 ‘슬세권’과 같이 MZ세대의 거주지 근방에서의 소비가 증가하는 소비패턴의 변화는 골목상권에 긍정적인 영향을 줄 수 있다. 위 두가지 현상을 바탕으로 본 조는 골목상권 활성화에 대한 분석의 필요성을 시사하였다.



¹ 임현정, 최상범, 코로나19가 서울시 상권 매출 변화에 미치는 영향 분석. 서울도시연구 제23권(제1호),2022



<표 1>

물론 코로나로 인한 피해는 골목, 대형 상권 구분없이 동일하지만, 빠르게 회복중인 발달상권과는 달리 골목상권은 더딘 회복세를 보인다. 또한, 코로나 전부터 서울시 내에서는 강남, 이태원과 같은 특정 인기 상권에서는 점포가 몰려 경쟁이 심화되고, 건물주들은 임대료를 올리는 현상이 반복되며 이로 인해 대형상권 인근으로 상권이 이동하는 ‘젠트리피케이션’ 현상은 고질적인 문제점 중 하나이다.

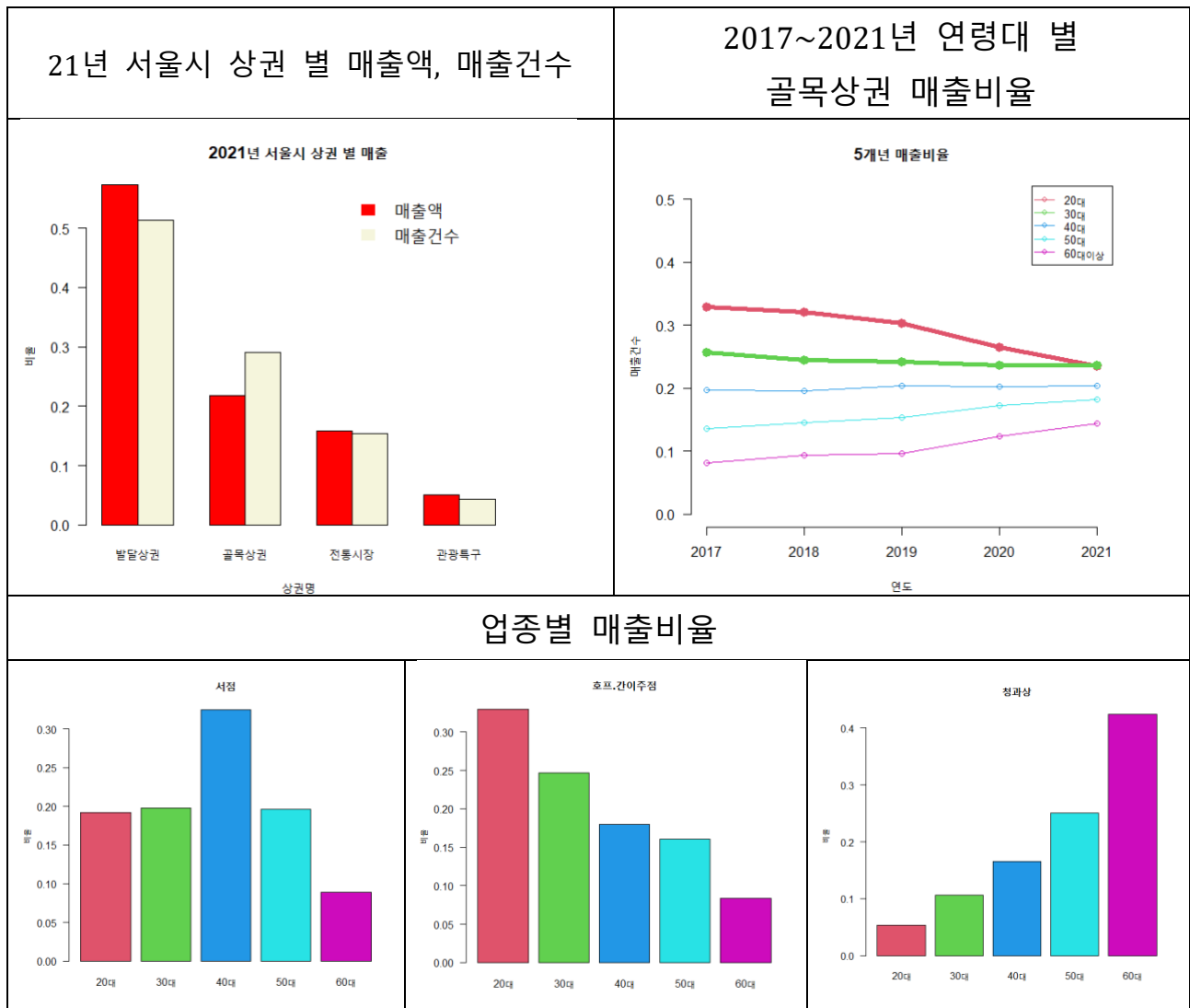
이는 특정 대형 상권 인근에서만 상권이 발달되어 지역 내 상권의 균형 잡힌 발달을 저해하며 이로 인해 소외된 상권은 점점 쇠퇴하는 경향이 가속화되는 것이다. 위와 같은 현상을 토대로 본 보고서는 분석 대상 상권을 골목상권으로 한정하였다.

앞서 언급했듯이 서울시 내 상권에 대한 정부의 지원과 연구는 활발히 진행 중이다. 특히 상권의 한 종류인 전통시장에 대한 정책과 분석은 대중들에게도 많이 알려져 긍정적인 영향력을 끼치고 있다.

이와 마찬가지로 골목상권에 대한 다수의 연구들이 존재하며 최근 서울시에서도 골목상권 육성사업 대상 골목을 선정하였다. 2019~2021년 총 3년간 서울시 내 모든 상권에 대해 매출액의 추이를 확인한 결과 발달상권 다음으로 골목상권이 전체 매출액의 약 20퍼센트, 매출건수의 30퍼센트의 비중을 차지하며 골목상권의 중요성을 다시한번 환기시켜준다.

또한 골목상권 내에서 연령대 별 매출 추이를 보면 1,2위인 2030의 비중이 전체 50퍼센트 이상임을 확인할 수 있으며, 연령대 별 업종마다 매출비율을 보면 연령대 별 편차가 있음을 확인할 수 있다. 흔히 얘기하는 MZ세대는 기존과 차별화된 것을 선호하며 소비 트렌드를 이끄는데, 이들로 인해 단순 주거지역 내 골목상권 이었던 망원동 망리단길, 노후 공장지역 이었던

성수동 카페거리 상권 등이 부활한 선례가 있다. 따라서 골목상권을 분석할 때 다양한 연령대의 특성을 고려한다면, 기존의 연구에선 단순히 골목 내 총 매출액에 영향을 끼치는 지표를 찾은 것과 다르게 더 다양한 분석을 할 수 있다고 판단하였다.



<표 2.>

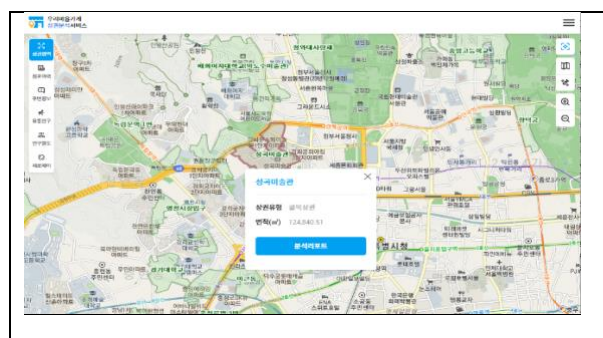
기존 연구들에서 상권발달의 척도로서 해당 상권의 매출액에 영향을 미치는 지표로는 유동인구, 상주인구와 같은 인구지표, 상권 내 동일업종을 고려한 경쟁지표, 인근 집객 시설이나 교통과 관련된 접근성 요인 등을 활용했다. 본 보고서에서는 위 3가지 지표를 세부적으로 나누고 골목특성을 나타내는 선호업종 점포수와 점포 밀집도를 추가하여 쇠퇴한 골목상권을 파악하고 다양한 연령대의 특성을 고려한 솔루션을 제안할 것이다.

2. 데이터 분석

□ 데이터 선정

서울특별시 골목상권관련 데이터는 서울 신용보증재단에서 제공하는 '우리마을가게 상권분석서비스' 2021 년 1~4 분기 데이터를 이용하였다.

1. 생활인구: 서울시와 KT 가 통신데이터를 이용하여 추계한 서울의 특정지역, 특정 시점에 존재하는 모든 인구에 대한 데이터이다. 상권영역 내의 유동인구파악을 위해 사용.
2. 추정매출: 골목상권 점포내에서 결제된 금액에 대해 카드사의 승인금액을 보정계수와 곱한 뒤 평균으로 산출된 매출 데이터. 데이터 내 총 매출, 연령대 별 매출, 시간대별 매출, 요일 별 매출 4 가지를 사용하였다
3. 점포 수: 업종에 대한 점포수의 경우 상권_점포데이터와 추정매출에서의 점포수가 달라 문의한 결과 기존 데이터에 확인되지 않아도 카드매출로 확인되는 경우가 있어 다를 수 있어 따라서 추정매출 데이터 내 산정된 점포수로 사용.
- 4(5). 집객시설(배후지): 상권의 접근성, 관공서 수 등을 파악하기 위해 집객시설 데이터를 사용하였다. 이때 지하철이나 은행, 관공서등은 상권 배후지에 집객시설로 서 상권에 주는 영향력이 타당하다고 볼 수 있으나, 버스정류장의 경우 반경 200m 이내로 정해진 배후지의 기준보단 상권 내 정류장의 수를 사용하는 것이 더 적합하다고 판단하였다. 따라서 버스정류장 수는 상권 내 기준, 지하철, 은행, 관공서 수 등은 상권 배후지(상권 반경 200M 이내)를 기준으로 데이터를 수집하였다.
6. 상권영역의 면적데이터의 경우 면적변수로 제공되지 않고 상권 중심점 좌표(X,Y)로 제공되었다. 다만 '우리마을가게 상권분석서비스' 홈페이지에서 지도안에서 면적을 확인할 수 있었기에 크롤링을 통해 면적데이터를 수집하였다.



<그림 1>

인구, 매출, 점포 데이터는 분기 평균값으로, 배후지의 경우 4 분기 데이터를 사용하였다. 면적의 경우 분기별로 나뉘지 않았기에 그대로 사용하였다.

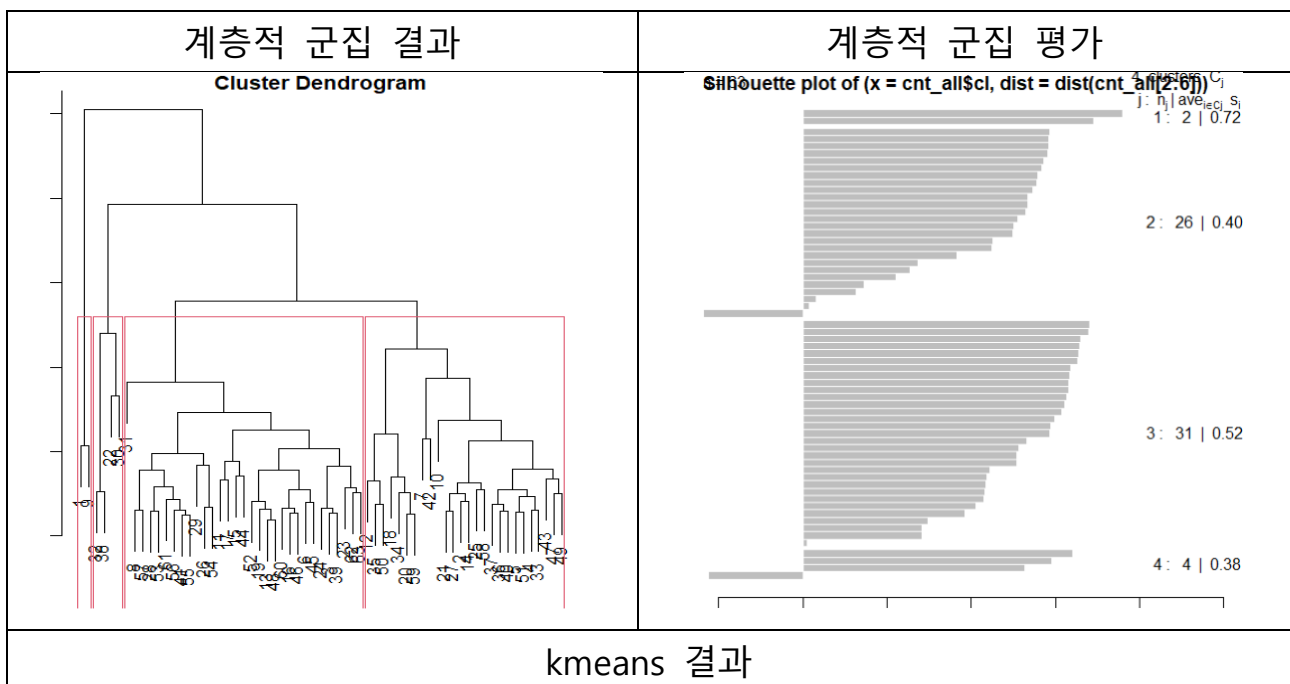
□ 데이터 분석

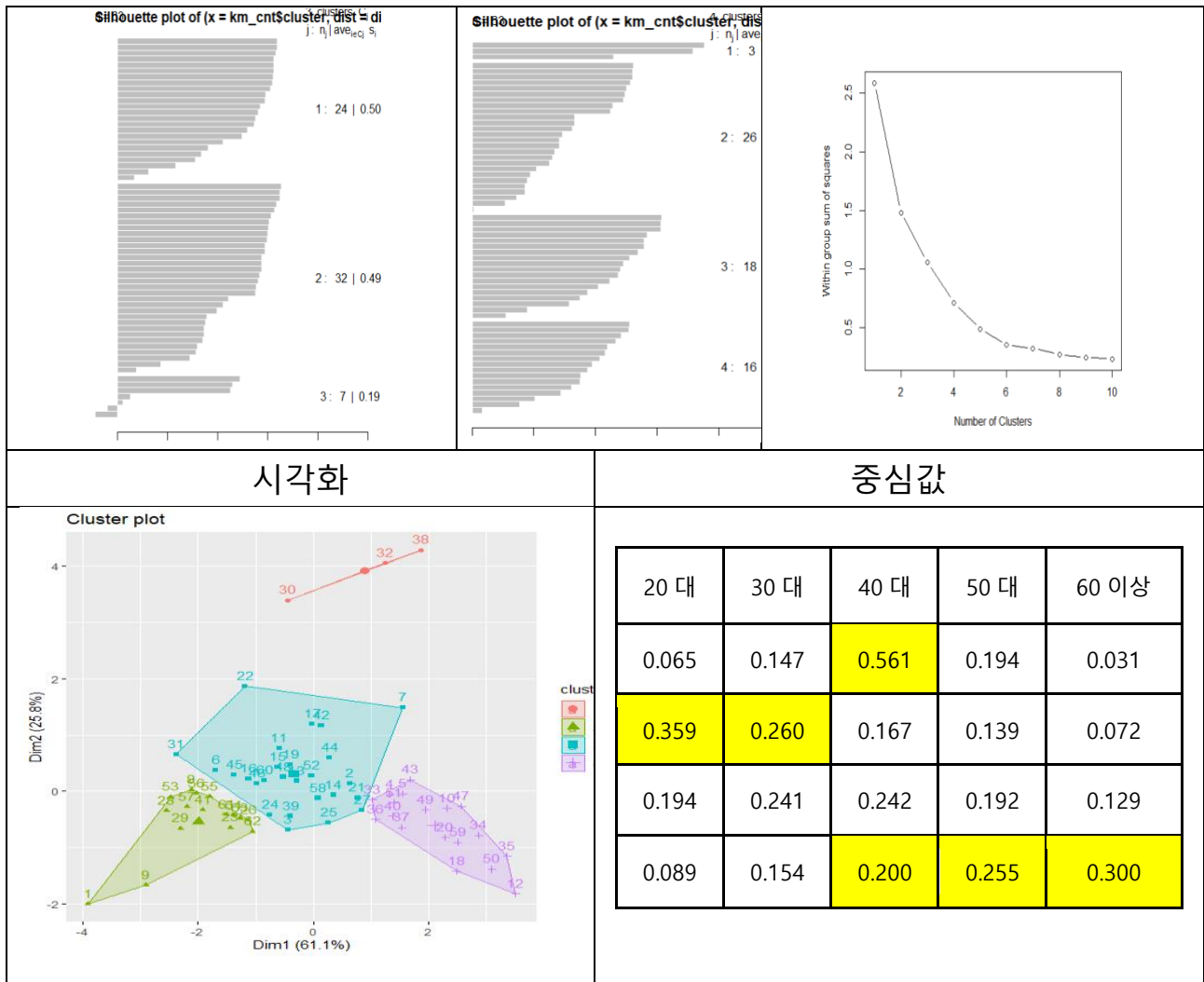
(1) 변수생성

• 연령대별 주요업종 점포수

서울신용보증재단에서는 100 개의 생활밀접업종을 표준산업분류와 연계하여 정의했으며, 이 중 골목상권에는 60 여개의 업종이 존재한다. 앞서 시각화로 본 것과 같이 연령대별로 주로 소비하는 업종은 다르다. 이를 고려하지 않고 단순히 상권 내 전체 혹은 특정 점포수와 매출액의 영향을 파악하는 것은 좋은 결과가 나오지 않을것이라 판단했다. 따라서 60 여개의 업종을 연령대별 매출비중을 통해 특정 연령대가 선호하는 업종을 찾아보았다. 이를 위해 업종별 20, 30, 40, 50 ,60 대 이상의 매출건수를 전체 대비 비율로 변환하여 군집화를 시도했다.

군집방법으로는 계층적 군집방법 hierarchical clustering 과 비계층적 군집방법인 K-means clustering 을 시도하였으며 결과는 아래와 같다.





<표 3.>

계층적 비계층적 군집결과 모두 최적의 군집수는 4 개이며, kmeans 에서 3 개의 군집과 4 개의 군집을 비교했을 때 군집 간 편차는 더 높지만, 높은 점수의 군집들과 해석의 용이성을 고려하여 최종적으로 4 개의 군집을 형성하였다.

군집 1 은 40 대 매출 비중이 높은 군집으로 외국어, 일반교습학원, 예술학원으로 구성되어있다.

군집 2 는 2030 의 매출 비중이 높은 군집으로 PC 방, 양식음식점, 커피-음료 등으로 구성되어있다.

군집 3 은 매출비중의 차이가 크지 않은 평이한 군집으로 미용실, 분식점, 슈퍼마켓 등으로 구성되어있다.

군집 4 는 405060 연령대의 매출 비중이 높은 군집으로 당구장, 가전제품, 철물점, 한의원 등으로 구성되어있다.

군집 결과를 확인해보면, 30 대는 단독으로 뚜렷하게 매출비중이 높은 업종을 찾을 수 없었으며 대부분 20 대와 비슷한 매출패턴을 보였다. 20 대와 40 대는 몇몇 업종에서 눈에 띄는 차이를 확인할 수 있었으며, 이를 제외하면 비슷한 연령대로 같이 묶이는 것을 볼 수 있다. 계층적 군집분석의 결과와 연계하여 최종 군집 결과는 다음과 같다.

2030 선호업종(18 개)		일반 업종(26 개)		405060 선호업종(21 개)	
“PC 방”	“컴퓨터 및 주변기기”	“중식음식점”	“가구”	“한의원”	“가전제품”
“네일숍”	“여관”	“핸드폰”	“가방”	가전제품수리	“당구장”
“노래방”	“피부관리실”	“미용실”	“고시원”	“미곡판매”	“섬유제품”
“스포츠클럽”	“호프-간이”	“반찬가게”	“골프연습장”	“수산물판매”	“운동/경기용품”
“양식음식점”	“주점”	“부동산중개업”	“문구”	“육류판매”	“의료기기”
“커피-음료”	“화초”	“분식전문점”	“서적”	“의약품”	“인테리어”
“안경”	“장치판매”	“세탁소”	“슈퍼마켓”	“일반의원”	“자동차수리”
“일식음식점”	“패스트푸드점”	“스포츠 강습”	“시계 및 귀금속”	“조명용품”	“철물점”
“편의점”	“화장품”	“신발”	“애완동물”	“청과상”	“치과의원”
		“완구”	“일반의류”	“예술학원”	“외국어학원”
		“자동차미용”	“자전거 및 기타운송장비”	“일반 교습학원”	
		“전자상거래업”	“제과점”		
		“치킨전문점”	“한식음식점”		

<표 4.>

- 집객시설

대학교 수의 경우 학교내 건물의 수로 되어있어 캠퍼스 수가 많아도 더 적은 값을 가진 것이 확인되었다.



<그림 2.>

예시로 두 학교 모두 상권 배후지영역에 캠퍼스는 동일하나 대학교 수는 9 개 차이가 남을 확인했다. 따라서 대학교 유무로 구분하여 factor 변수로 처리하였다. 또한 백화점 수의 경우 1 개이상 위치한 경우가 없었으며 지하철역 수는 1 개만 있는 곳이 대부분이기에 연속형 변수로 처리하기보단, 대학교와 마찬가지로 factor 변수로 처리하였다.

- 생활인구변수

생활인구 데이터는 '남성연령대_20_월요일시간대_1_생활인구_수'와 같이 변수가 세분화되었기 때문에 한 연령대에 대해 '남성연령대_00_생활인구_수', '연령대_00_시간대_1_생활인구_수', '연령대_00_주중_생활인구_수'와 같은 방식으로 재분류하였다.

- 면적관련 변수

골목마다 매출이나 생활인구 수에 편차가 큰 이유는 골목의 면적이 넓을수록 인구수나 매출액이 높기 때문이다. 따라서 이를 고려하지않은 채 매출액을 기준으로 골목상권의 활성화 필요 유무를 판단한다면, 면적이 좁은 상권만 취약상권으로 될 우려가 있다. 따라서 밀집도(점포수/면적)변수를 추가로 생성하였다.

따라서 최종 변수는 다음과 같다.

종속변수	독립변수	
매출액	생활인구 (전체, 남녀, 시간대, 요일)	연령대 별 선호 업종의 점포수
	대학교, 지하철, 백화점 유무	배후지 의료기관 수
	상권 내 버스정류장 수	배후지 관공서 수
	밀집도	

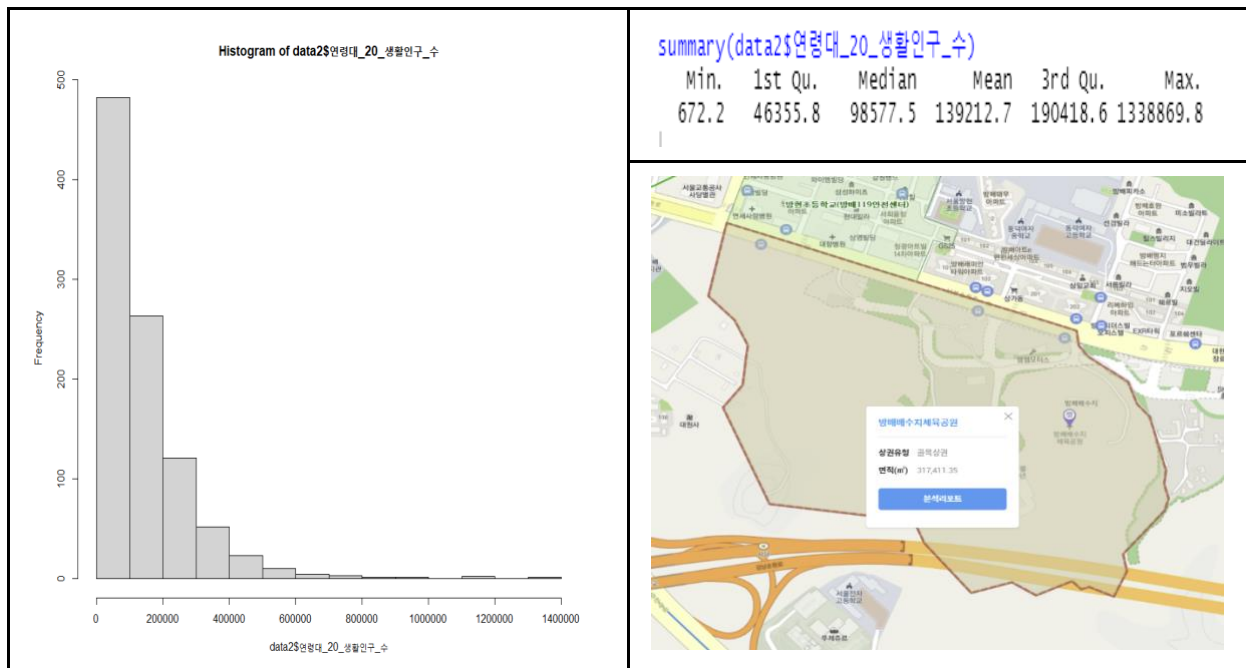
<표 5.>

(2) 다중회귀분석

- 이상치 확인

골목상권 활성화라는 분석목적에 부합하지 않는 발달상권의 범주에 포함되는 망리단길, 샤로수길과 같은 6 개의 골목상권은 분석에서 제외하였다.

대부분의 독립변수에서는 뚜렷한 이상치가 보이지 않는다. 하지만 생활인구 변수에서는 분석결과에 영향을 줄 수 있는 이상치가 있다고 판단하였다.

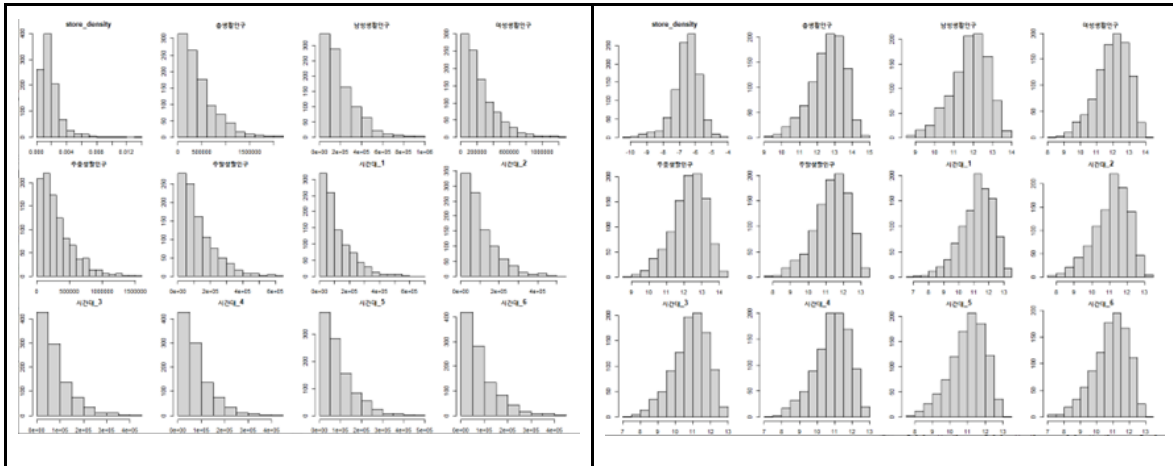


<그림 3.>

분기 평균 생활인구수가 1000 명 이하인 지역을 살펴보았을 때 데이터 집계가 잘못되었다고 판단하였다. 특정 연령대가 아닌 전 연령대에서 위와 같은 모습을 보이는 지역 총 12 곳을 제외하였다. 최종 데이터내 골목상권의 수는 951 개이다.

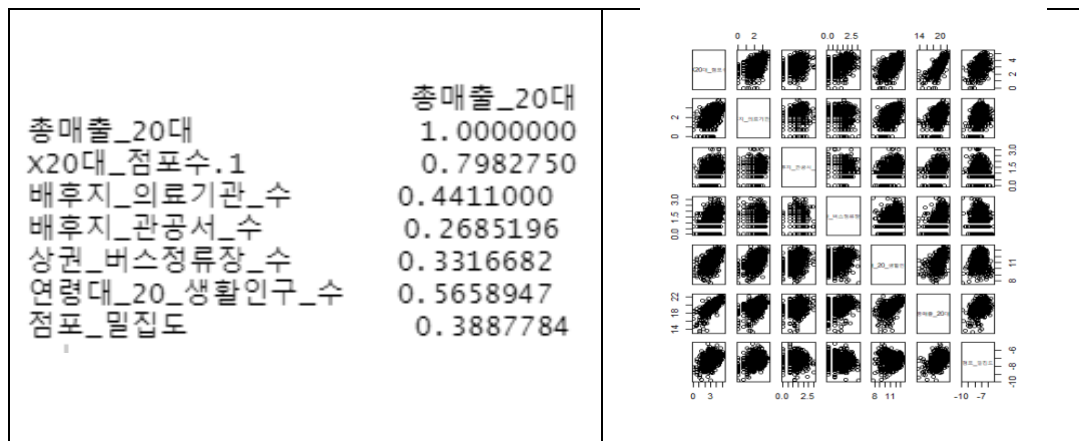
- 모델 구축

앞서 업종 분류에 따라 모델을 20 대, 30 대, 40~60 대이상으로 나눠 다중회귀분석을 진행하였으며 30 대는 2030 선호 업종 내에서 20 대의 매출비중이 매우높은 PC 방과 노래방을 제외하였다.



<그림 4.>

왼쪽 히스토그램에서 보이듯이 각 변수들이 오른쪽으로 꼬리가 긴 right skewed 한 분포를 보이므로 해당 변수 그대로 회귀모형을 돌릴 시 잔차의 정규성이 위배될 수 있다. 따라서 이를 방지하고자 factor 변수를 제외한 연속형 변수들에 대해 log 변환을 하였다. 변환 후 대부분의 변수들의 분포가 바뀐 것을 확인할 수 있다.



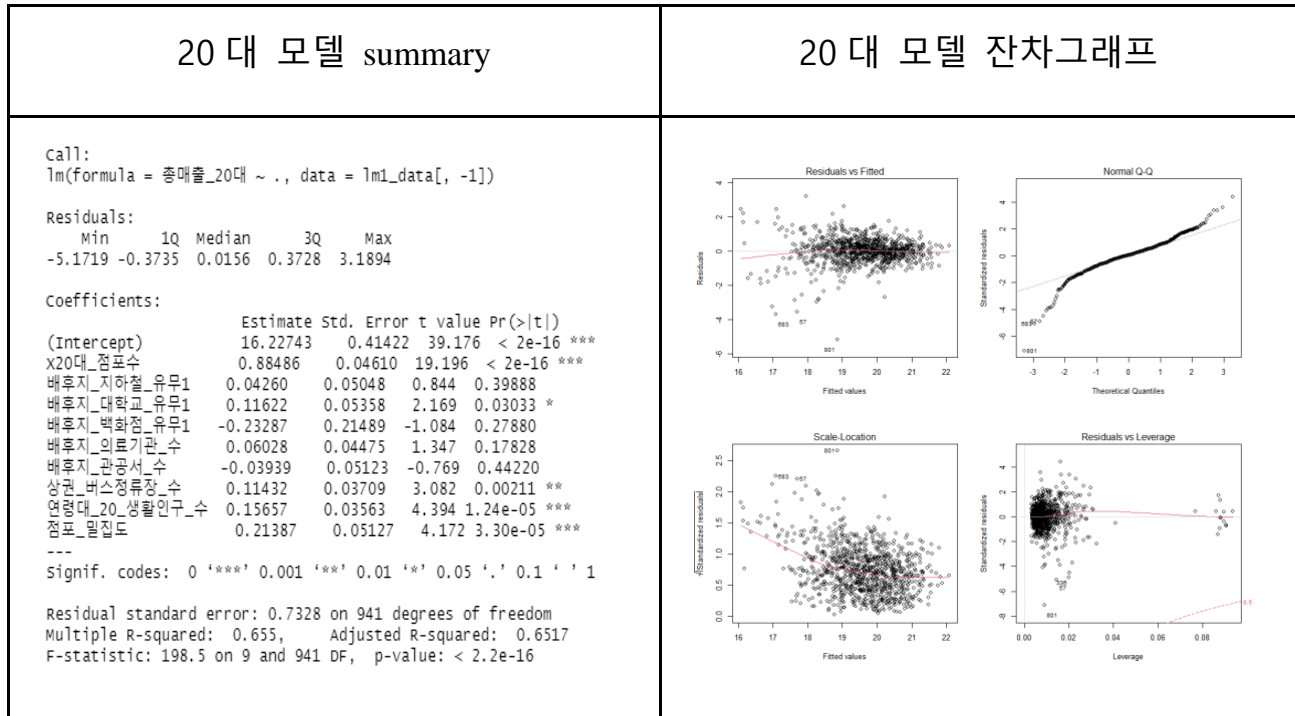
<그림 5.>

모델 별 매출액과 독립변수들간의 상관계수와 산점도를 확인 한 결과, 모든 모델이 선형성을 만족한다고 판단하여 그대로 분석을 진행하였다.

- 모델 결과

R^2_{adjust}	총 생활인구	남녀 생활인구	시간대 생활인구	요일별 생활인구
20 대	0.6516	0.6514	0.6537	0.6515
30 대	0.606	0.6063	0.6102	0.6069
고연령 (40 50 60)	0.701	0.713	0.7124	0.7133

<표 6.>



vif(lm00)	x20대_점포수	배후지_대학교_유무	상권_버스정류장_수	연령대_20_생활인구_수	점포_밀집도
	2.769766	1.024851	1.221519	1.962354	1.506016

<표 7.>

전반적으로 생활인구 별 모델들의 설명력에는 큰 차이가 없으며 모든 모델에서 잔차그래프 또한 적합한 것으로 나타났다. 따라서 20 대에서는 해석이 용이한 총 생활인구 모델을 사용하였으며 유의한 변수로는 배후지 대학교 유무, 20 대 선호 업종 점포 수, 버스정류장 수, 총 생활인구 수, 밀집도로 6 가지가 나타났다. 설명력은 0.6517 이며 잔차그래프를 확인했을 때 회귀가정을 만족하는 유의한 모델이라고 판단하였다.

30 대에서도 마찬가지로 결정계수 0.606 인 총 생활인구 모델을 선정하였으며 유의한 변수로는 30 대 점포 수, 버스정류장 수, 30 대 총 생활인구 수, 밀집도 변수 4 가지로 나타났다.

고연령대에서 회귀계수가 0.713 으로 가장 높은 주중, 주말생활인구 모델을 선정하였다. 유의한 변수는 고연령대 점포 수, 버스_정거장_수, 점포 밀집도, 주중 및 주말 생활인구로 확인되었다

(3) 상권 군집화

선형 회귀 모델을 구축함으로써 각 연령대 별 매출액에 유의한 영향을 미치는 변수들을 도출하였다. 매출액이 상권의 발달정도를 나타내는 지표일때, 이 변수들은 매출에 영향을 미치며 이를 통해 골목상권을 활성화요인을 세분화할 수 있다고 판단하였다. 따라서 앞서 분석한 951 개의 골목상권에 대해 위 변수들로 다시 군집화를 진행하여 활성화 필요 지역을 찾아보려 한다.

- 주성분 분석

앞서 선정된 20 대 선형회귀모델의 유의한 독립변수 5 개로 주성분 변수를 생성한다. 누적 설명력비율 70% 기준으로 변수 2 개를 선정하였으며 PC1 은 모든 변수에서 양의 값을 갖는 상권 발달지수로 정의했다, PC2 는 인구변수와 다르게 밀집도가 높은 음의 값을 가진다이는 상권의 반 밀집지수로 정의했다

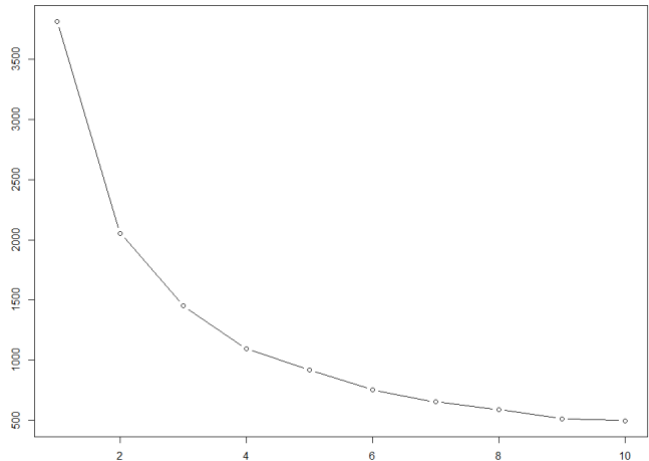
20 대	comp1	comp2	comp3
store_4060	0.661	-0.185	0.130
버스_정거장_수	0.443	0.248	-0.839
store_density	0.128	-0.924	-0.128
총 생활인구	0.590	0.222	0.511
cumulative P	0.4624	0.7332	0.9318

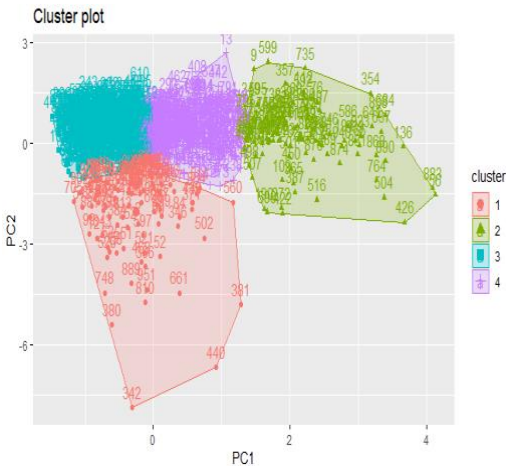
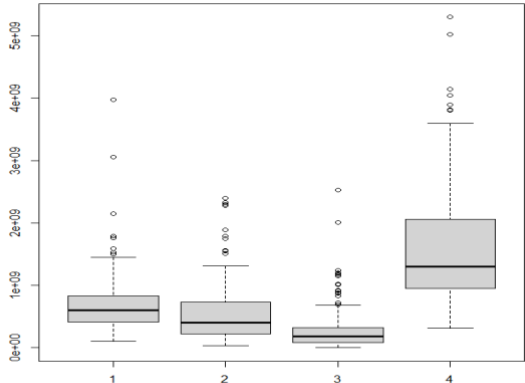
<표 8.>

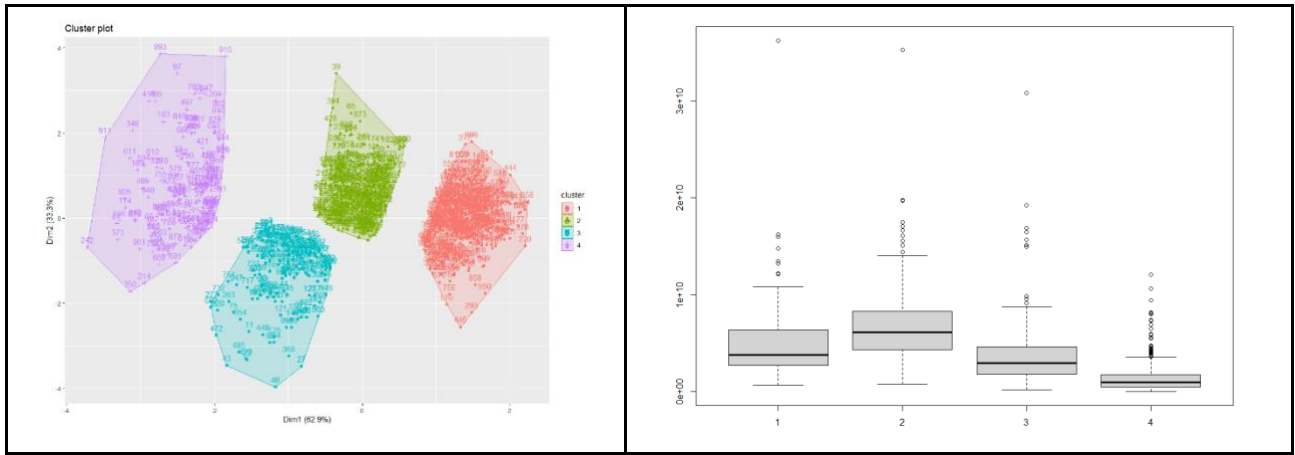
다른 두 연령대 모델에서도 20 대와 유사한 주성분 해석 결과가 나온다.

PC2 에서 해석이 반대로 나오는 것을 제외하면 두 주성분이 유사하므로 똑같이 발달지수와 밀집지수로 해석하였다.

- 군집화

20 대 모델 실루엣 계수	elbowplot
<p>1 : 293 0.32</p> <p>2 : 103 0.34</p> <p>3 : 100 0.28</p> <p>4 : 455 0.49</p>	

20 대 모델 시각화	집단별 평균차이 boxplot
	
고연령대 모델 시각화	집단별 평균차이 boxplot



<그림 6.>

20 대 군집화 결과 총 4 개의 군집이 유의하다고 나왔으며 군집 내 중심값을 기준으로 해석해보면 발달지수에 따라 매출액의 평균차이가 크게 나는 것을 확인할 수 있다. 또한 발달지수와 반 밀집지수의 차이가 큰 1 번군집과 2 번군집의 평균차이가 크지 않은 것을 보아 반 밀집지수도 평균에 영향을 주는 것을 확인할 수 있다. 3 번 군집은 매출액 평균이 가장 낮으므로 활성화 필요 지역으로 선정하였다.

30 대와 고연령대에서 주성분으로 골목상권을 군집화한 결과도 마찬가지로 각 군집별 매출차이가 있음을 확인했고 그 중 매출액이 가장 낮게 분포한 군집이 발달이 잘 안되고 밀집정도가 적은 활성화 필요 지역으로 선정했다.

-최종 활성화 필요 지역 선정

3 개의 모델에서 평균매출액이 제일 낮은 군집들을 활성화 필요 군집으로 선정하였다. 군집들은 각각 452 개, 460 개, 412 개의 골목들이 있으며 위 3 군집에 모두 포함되는 골목은 324 개이다. 해당 골목들에 대한 심층 분석의 필요성이 있지만, 본 연구는 연령대 별 군집을 통해 활성화 방안을 제시하는 것이므로 20 대의 매출액 하위 군집에는 포함되지만 30 대와 고연령대(405060)의 매출액 하위 군집에는 포함되지 않는, 다른 연령대 대비 특정 연령대의 매출액이 낮은 지역을 최종 활성화 지역으로 선정했다.

20 대 한정 취약지역은 3 곳으로 그 중 천호지구대, 천호초등학교 골목상권이 위치한 **천호 2 동** 지역으로 선정했다. 해당지역은 서울 유명 성매매 집결지가 있던 곳으로 최근에는 재개발로 인해 철거중이며, 아래로는 발달상권인 천호로데오거리, 위로는 암사동으로 주택단지가 많으며 풍남토성, 암사유적지와 같은 역사 유적지가 위치하고 있다. 인근 지하철역이 잘 발달되어 있으며,

천호동 내에서도 한강과 가장 인접한 지역으로 한강공원을 즐겨 찾는 젊은 세대들에게 최적의 입지가 될 수 있다. 따라서 재개발이 끝난다면 제 2의 망원이 될 가능성이 높은 지역이라고 판단된다. 활성화 방안으로는 자연-역사를 연계시킨 도심 속 힐링 역사여행과 같은 정책을 통해 20 대의 역사의식을 고취시킬수 있을 것이라고 생각된다. 유사 정책으로는 최근 다시 재개발 예정인 한강 역사탐방 내 광나루길 코스가 있다.

30 대 한정 취약지역은 19 곳이며 위 지역들은 30 대의 매출만 낮기보단, 대체적인 연령대 매출이 낮은 곳이 대부분이다. 이는 30 대의 뚜렷하지 않은 소비특성으로 인한 결과라고 판단하였다.

고연령대(405060) 한정 취약지역은 80 곳이며 위 지역들에서 특이한 점은 서울대입구역, 서울숲역과 같이 평균매출이 높은 지역도 선정된 것이다.

이는 골목 내 매출액에 대해 연령대 별 매출비율을 고려하였기 때문에 2030 연령대 대비 고연령층의 매출비율이 낮으며 젊은 연령층의 소비가 주가 된, 소위 MZ 세대의 핫플레이스가 취약지역으로 선정되었기 때문이다. 또한 30 대외 유사하게 해당되는 지역을 제외한 곳 중에서는 모든 연령대에서 낮은 매출액을 보이므로 해당되는 지역들에 대해서는 추가적인 분석이 필요하다고 판단된다.

3. 분석 활용 전략

□ 기대효과

특정 연령대를 겨냥한 활성화 필요 상권 발견으로 인해 소외되는 연령대 없이 균형 잡힌 골목상권의 발전이 가능하다.

100 여개의 업종 종류를 연령대 별 주요 소비 업종 군집으로 구분함으로써 현재 골목상권내 주요 타겟 연령층을 확인할 수 있으며, 창업 예정인 소상공인에게 특정 연령대를 위한 업종을 추천할 수 있다.

현재 상권의 주 소비층을 파악할 수 있으며 필요에 따라 새로운 소비층 유입 혹은 기존 소비층 강화와 같은 활성화가 가능하다.

□ 방향제시

더 세분화된 연령대 매출 모델을 만든다면 앞서 나온 결과와 같이 20 대 뿐만 아니라 더 다양한 연령대 소비활성화에 기여할수 있다. 연령대로 나누는 기준을 매출액과 업종만이 아닌 SNS 데이터와 같은 세대 별 특징을 파악할 수 있는 데이터를 사용한다면 더욱 좋은 군집결과를 얻을 것이라고 생각한다

참고문헌

임현정, 최상범(2022). 코로나19가 서울시 상권 매출 변화에 미치는 영향 분석. **서울도시연구 제23권(제1호)**

강현모(2018) 골목상권의 매출액과 성장 결정요인, **가천대학교 일반대학원 석사학위논문**

이동현 외(2020) 서울시 성장상권과 쇠퇴상권 내 외식산업의 생존율 비교. **국토연구 제105권**