

## 논문

# 서울시 상권의 입지 특성과 유동인구에 따른 점포유형에 관한 연구: 명동, 강남대로, 홍대합정 상권을 중심으로\*

A Clustering Analysis of Major Retail Area in Seoul: Focused on the  
Myeongdong, Gangnam-daero, Hongik Univ.-Hapjeong Retail Area

김세환\*\* · 김대원\*\*\* · 김호철\*\*\*\*

Kim, Se-Hwan · Kim, Dai-Won · Kim, Ho-Cheol

## Abstract

When it comes to retail as a commercial real estate, selecting a proper location is of importance because it could be a biggest factor for sales and existence of itself. Nevertheless, in practice level, since a lot of decision makers tend to rely on subjective opinions of expertise's with abundant experience, it is obvious that the need for objective measurement and empirical analysis comes to fore.

In this regard, the purpose of this study lies on the deduction of meaningful and objective indices for retail location selection using Space Syntax values and floating population extracted from the real business district in Seoul. In order to pursue the purpose, this study collected the retail market data such as floating population, physical characteristics, etc., in Myeongdong, Gangnam, and Hongik Univ.-Hapjeong retail area of Seoul and measured Space Syntax values. With these data, we examined the feasibility of empirical analysis on the characteristics of the retail types based on the space syntax traits and floating population.

From the multinomial logistic regression analysis, we found the statistically meaningful relationship between retail types and space syntax characteristics and floating population. Succeedingly, we classified the types of retail based on the space syntax characteristics and floating population level using  $k$ -means clustering method and location quotient. From this, we found that the retail can be classified into four clusters and each clusters tend to be determined by the location characteristics and the level of floating population.

주 제/어: 점포입지, 공간구분, 유동인구,  $k$ -평균 군집분석, 입지계수

Keyword: Store Location, Space Syntax, Floating Population,  $k$ -means Clustering, Location Quotient

\* 본 논문은 제1저자의 석사학위논문 내용 중 일부를 수정·보완한 것이다.

\*\* 인천도시공사 차장(주저자: megaori@imcd.co.kr)

\*\*\* 인천도시공사 차장(공동저자: dw2613@gmail.com)

\*\*\*\* 단국대학교 교수(교신저자: khc2000@dankook.ac.kr)

## I. 서론

### 1. 연구의 배경 및 목적

소매 점포의 운영에는 다양한 요소가 영향을 미친다. 소상공인을 대상으로 한 조사에 따르면 자영업 창업 과정에서 가장 큰 애로사항은 자금마련이었으며, 입지의 선정, 업종의 선택, 적절한 정보의 획득 등이 그 뒤를 따르는 것으로 확인된다(중소기업청, 2013). 특히, 어떠한 위치에 어떤 업종으로 점포를 개설할 것인가에 대한 문제는 소매업자에게 가장 선택하기 어려운 요인으로 꼽힌다. 고객들의 접근성이 용이한 지점에 위치한 점포는 높은 매출액을 창출할 수 있는 반면, 반대의 경우라면 매출액 감소로 인하여 점포의 존립 자체를 위협할 수 있기 때문이다. 이러한 입지의 중요성과 관련하여, 이호병(2012)은 점포 개설 및 운영에 있어 입지의 문제는 잘못 선택할 경우 극복하기가 쉽지 않은 장기적인 고정 투자의 관점으로 보아야 함을 강조한 바 있다.

입지 선정의 중요성에 대한 인식이 만연해지는 추세 속에서, 최근 프랜차이즈 업계에서는 새로운 점포의 개설을 위해 다양한 입지 분석을 실시하고 있다. 주변 유동인구, 인구 유입시설의 유무, 임대료 추세, 동종 경쟁 업종 점포 수 등을 조사하여 예상 매출액과 수익률을 추정함으로써, 최적의 입지를 선정하기 위하여 많은 시간과 비용을 지불하고 있다. 이러한 추세는 개인 소매 창업자에게도 예외는 아니다. 개인들도 다양한 정보원을 동원하여 최적의 입지를 선정하기 위하여 다각적인 노력을 기울이고 있는 상황이다. 다만, 이러한 노력이 측정이 가능한 객관적인 지표나 검증이 가능한 방법으로 이뤄지고 있는지에 대해서는 많은 의문이 제기되고 있다. 실제 입지 선정 과정에서 일부 전문가의 경험과 주관적 예측에 근거한 의사결정이 자주 목격되고 있기 때문이다. 즉, 소매업에서 입지 선정이 중요한 요인으로 인지되고 있음에도 불구하고 입지 선정 실무 과정에서는 의사결정자들의 통찰력과 경험에 지나치게 의존하고 있고, 정성적인 판단을 근거로 최종 결정이 내려지고 있는 것이다.

이 같은 맥락에서 소매점의 입지 선정에 있어 객관적인 방법에 근거한 입지 유형의 분류와 현상 예측에 대한 필요성이 대두되고 있다. 그동안 시장에서 이루어진 특정 전문가 위주의 경험과 권위에 근거한 주관적인 판단으로는 적정 입지 선정의 여부를 판단하기 어렵다는 문제가 있기 때문이다. 동일한 관점에서 이미 다수의 연구자는 오래전부터 입지 유형별 상가 유형에 대한 지대한 관심을 가져왔다. 그럼에도 불구하고 점포의 입지에 대하여는 여전히 실무자 중심의 경험적 해석에 의존하는 비중이 높아 객관적인 입증이 어렵다는 한계를 갖고 있는 실정이다.

이에, 본 연구는 측정이 가능하고 객관적으로 계량화할 수 있는 입지 분석의 기초 작업을 위하여 입지 유형별 점포의 형태를 확인하는 연구를 도모하였다. 즉, 어떠한 길목에 어떠한 종류의 상점이 위치하는지를 측정하고 분류함으로써 입지 분석 연구의 기초적인 초석 생성에 일조하고자 하였다.

이를 위하여 본 연구는 서울시 대형 상권을 중심으로 공간구문론적 입지 특성과 유동인구 특성이 점포유형 결정에 영향을 미치는지 여부를 확인하였으며, 이러한 요소에 의해 유형별 점포의 분류 가능 여부에 초점을 맞춰 연구를 수행함으로써 객관적 점포유형의 분류 가능성을 가늠하고자 하였다.

## 2. 연구의 범위 및 방법

본 연구는 길목의 상대적 가치와 유동인구에 따른 점포유형 간의 관계를 도출하는 것으로 개별 점포에 대한 입지 특성을 계량화하기 위해 서울시 내의 주요 상권 중에서 명동, 강남대로, 홍대합정 상권 3곳의 개별 점포를 연구의 대상으로 설정하였다. 이는 한국감정원의 상업용 부동산임대 동향 대상지 전국 17개 시·도, 245개의 상권 중에서 38%에 이르는 93개 상권이 서울, 경기, 인천 등 수도권에 위치하고 있으며, 그 중 절반 정도가 서울에 위치하고 있다는 점 및 서울 주요 상권 중 유동인구가 가장 많은 지역이 명동과 강남대로 상권인 것으로 밝혀진 연구결과(‘2015 서울시 유동인구 조사 보고서’) 등을 연구대상 선정에 반영한 결과이다. 이에 본 연구는 우리나라에서 대표적인 상권을 보유하고 있는 서울지역의 상권 중 유동인구가 많은 명동, 강남대로 상권과 서울 서부지역의 핵심 상권인 홍대합정 상권 3개소를 분석대상으로 설정하였으며, 분석대상 상권에 대한 세부 현황은 <그림 1> 및 <표 1>과 같다.

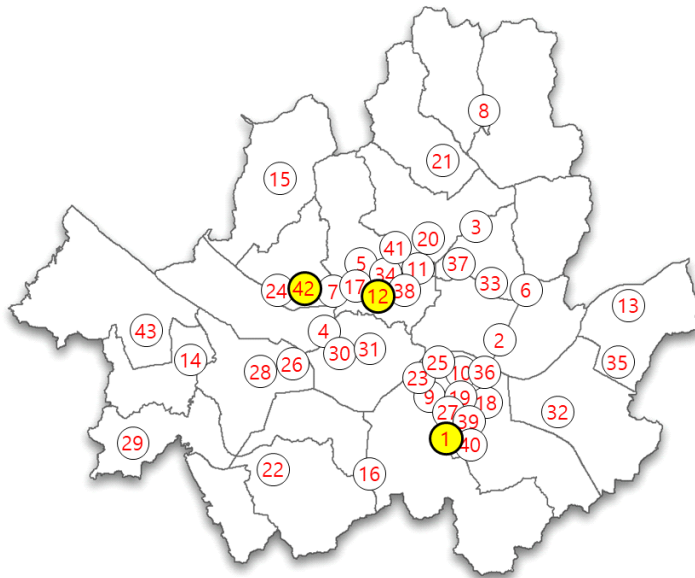


그림 1. 서울시 주요상권 43개소 중 분석대상 상권 현황(1. 강남대로, 12. 명동, 42. 홍대합정)

표 1. 분석대상 상권 세부 현황

그림1 표시	상권명	위치	면적(m <sup>2</sup> )	상위 2개 업종분포(%)
①	강남대로	강남구, 서초구	798,358	음식(33.7), 생활서비스(15.0)
②	명동	종구	338,877	음식(33.9), 소매(20.2)
④	홍대합정	마포구	973,935	음식(53.9), 생활서비스(9.8)

점포의 유형 결정에 영향을 미치는 변수를 측정하기 위해, 본 연구는 선행적으로 개별 점포가 접한 길목의 상대적 가치를 나타내는 지표인 공간구문론(space syntax)의 통합도(integration)를 측정하고, 개별 점포의 유동인구, 층수, 도로 폭, 개별공시지가 자료들을 수집하여 다항 로지스틱 회귀분석(multinomial logistic regression)을 함으로써 통합도와 유동인구가 점포의 유형과 맺는 인과관계를 확인하였다.

그리고 *k*-평균 군집분석(*k*-means clustering)을 사용하여 점포의 유형을 통합도와 유동인구를 기반으로 4개의 군집으로 분류한 후 입지계수(location quotient) 산정방법을 적용하여 각 군집별 특화 업종을 확인함으로써, 통합도와 유동인구에 따른 점포유형을 실증 분석하고자 한다. 연구의 분석과정은 <그림 2>와 같다.

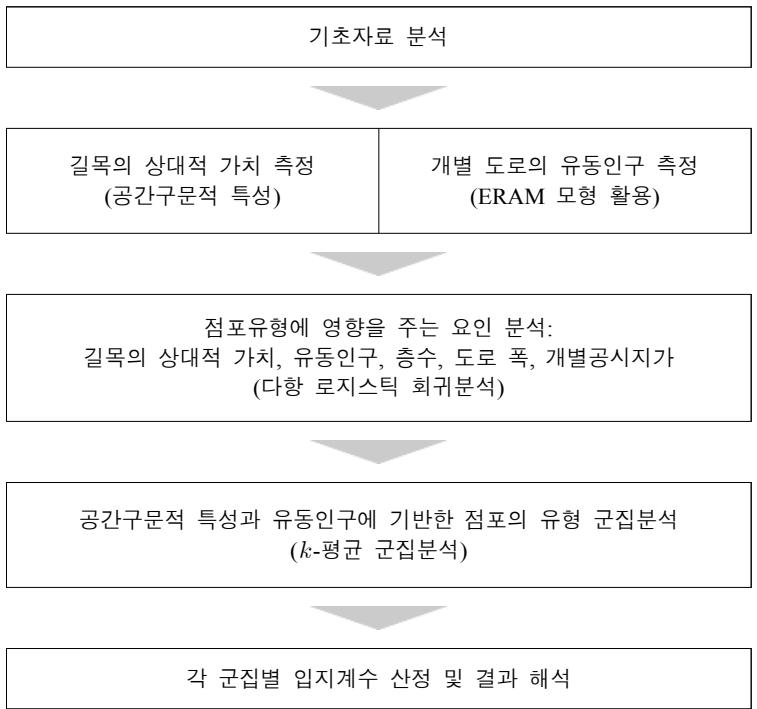


그림 2. 연구의 분석과정

## II. 선행연구 검토

### 1. 점포 입지 분석에 관한 연구

수익을 목적으로 하는 점포에 있어 입지, 규모, 서비스 수준 등은 중요한 요소이며, 입지를 제외한 나머지 요소들은 주변 환경에 따라 쉽게 변동되는 마케팅 요소이기에, 점포의 입지는 수익 창출을 위한 중요한 요소라 할 수 있다(Ghosh and Craig, 1983). 이러한 맥락에서 다수의 선행연구는 다양한 입지적 요인들이 점포의 매출액 결정에 영향을 미치고 있음을 다음과 같이 보여주고 있다.

황보윤·하규수(2008)는 미용서비스업에서는 가시성·접근성이 월평균 매출 및 디자이너 1인당 매출에 양(+의 영향을 미치는 것으로 확인하였다. 신유진·최승두(2014)는 프랜차이즈 커피전문점의 입지를 결정할 때 가장 중요한 요인은 접근성, 그 다음으로 도로 여건 및 도로와의 인접성, 주변 상업시설의 밀집도, 유동인구라고 분석하였다. 박주영(2015)은 프랜차이즈 패밀리레스토랑의 경우 접근특성 요인인 전면도로의 폭과 인구특성요인인 유동인구 수가 월평균 매출액과 양(+의 관계를 가지고 있음을 확인하였다. 김성문 외(2014)는 백화점 매출액의 영향요인은 입지요인 중 공시지가만이 매출액과 정(+의 관계를 가지는 것으로 분석하였다. 김성문 외(2014)는 대형할인점의 매출액에 영향을 주는 요인은 입지요인 중에서 세대 수, 세대당 지방세 징수액이 정(+의 관계를 가지는 것을 확인하였다.

상기 연구결과를 통해 접근성, 밀집도, 유동인구, 주요 시설 인접성 등 다각적인 입지요인이 점포 운영 성과에 중요한 요소로서 영향을 미치고 있음을 확인할 수 있다.

### 2. 공간구문론을 활용한 연구

1984년 Hillier의 공간구문론이 소개된 이후 도시의 공간구조 분석에 대한 중요성에 사회적 관심이 나타나기 시작하였으며, 상업용 부동산과 관련하여 다음과 같은 선행연구를 찾을 수 있었다.

권영환 외(2001)는 공간구문론의 도시유형 분석방법의 적용이 도시 공간구조의 특징에 따라 서로 다른 양상의 보행자 움직임이 나타난다는 것을 발견하였으며, 도시의 유형학적 분류의 방법론으로서의 가능성도 찾을 수 있었다고 하였다. 김진식(2008)은 도심 내 대규모 복합시설물을 공간구문론을 이용하여 분석한 결과 위상학적 접근성이 높을수록 상가의 입점객이 증가하는 것을 확인하였다. 김선호 외(2011)는 소매점포가 밀집된 대표적 상업지역 중 서울 압구정동을 분석한 결과 전체통합도와 부분통합도는 보행자 수와 양(+의 상관관계가 있으며, 특히 부분통합도는 보행자 수와의 관계를 잘 설명한다고 하였다. 김대원·유정석(2014)은 서울 중구 명동상가를 대상으로 공간구문적 요소와 임대료의 관련성을 분석한 결과 공간구문적 특성인 통합도와 사람의 공간 이용 가능성을 나타내는 ERAM 값 모두 상가의 임대료와 양(+의 상관관계를 가지고 있는 것으로 확인하였다.

### 3. 유동인구와 입지 특성 간의 관계에 대한 연구

유동인구는 점포의 입지 특성과 매출액에 영향을 주는 요소 중의 하나로 최근 점포의 입지와 관련된 연구들을 살펴보면 유동인구 특성을 통행패턴, 연령대, 시간 등으로 세분화하여 분석하는 연구가 활발히 진행되고 있음을 알 수 있다. 유동인구와 점포 간의 관계에 대한 선행연구를 요약하면 다음과 같다.

이경민·정창무(2014)는 업종의 주 수요층 비율이 높은 시간대의 유동인구가 점포의 입지에 양(+)<sup>1)</sup>의 영향을 주는 것으로 확인하였으며, 유동인구 특성에 맞는 점포들이 해당 지역에 입지한다고 하였다. 김수현 외(2015)는 보행량에 매출액이 가장 영향을 많이 받는 업종은 편의점과 커피전문점인 것으로 확인하였다. 노은빈(2015)은 이용인구가 대부분의 음식점 업종의 매출에 정(+)<sup>2)</sup>의 영향을 미치는 것으로 분석하였으며, 특히 동양식 전문 업종의 매출이 이용인구에 가장 민감하게 반응한다고 해석하였다.

### 4. 기존 선행연구와의 차별성

점포의 입지 분석과 관련된 선행연구에서 살펴본 바와 같이, 입지의 특성과 점포 운영 간의 관계에 관한 연구, 공간구문론을 통한 계량화된 길목의 상대적 가치와 경제적 효과와의 관계에 관한 연구, 그리고 유동인구와 입지 특성 간의 관계에 관한 연구 등 입지 분석을 둘러싼 다수의 연구가 수행되었음을 알 수 있다. 다만, 이들 연구는 입지 특성과 기타 변수와의 관계 확인에 초점이 맞춰져 있음을 확인할 수 있다. 반면, 본 연구에서는 상가 운영에 중요한 변수로 간주되는 입지 특성과 유동인구에 따른 점포유형을 분류하고, 길목의 상대적 가치와 유동인구의 규모에 따라 어떠한 종류의 점포가 입지하는지를 확인함으로써 다음과 같이 선행연구와의 차별성을 도모하였다.

첫째, 본 연구는 점포의 입지요인 중 길목의 상대적 가치와 유동인구에 따라 개별 점포의 입지 특성을 계량화하여 객관적 지표로 실증분석하였다는 점에서 차별성을 갖는다. 기존 선행연구는 점포의 입지에 영향을 주는 특정 요인을 분석하였다면, 본 연구는 점포의 입지 현황을 공간적 특성과 유동인구의 두 가지 요인으로 계량화한 점에서 차별성을 갖는다.

둘째, 개별 점포의 입지 특성을 분석하기 위해 공간적 특성과 유동인구에 따라 점포유형을 군집분석하고 군집별 입지계수를 측정하여 다각적인 측면에서 점포의 입지 특성을 실증분석하였다는 점에서 차별성을 추구하였다.

## Ⅲ. 분석자료와 방법론

### 1. 분석자료 및 기초통계

본 연구의 대상은 서울시 내 주요 상권 중 유동인구가 많으면서 대표적인 상권으로 분류되는 명동

상권 338,877㎡, 강남대로 상권 798,358㎡, 그리고 홍대합정 상권 973,935㎡를 대상으로 하였으며, 분석을 위한 자료는 다음과 같이 구성하였다.

먼저 점포 현황 기초 자료는 소상공인시장진흥공단에서 제공하는 2015년 12월 상가 업소 정보를 활용하여 서울시 내 408,714개의 점포 현황을 수집하였다. 소상공인시장진흥공단에서 제공하는 자료에는 점포명, 점포유형 분류, 지번, 층수 등의 정보가 포함되어 있으며, 집합건물 또는 구분상가 등으로 인해 1개의 필지(지번)에 여러 개의 점포 현황이 포함되어 있다.

본 연구는 점포의 공간구문적 특성인 통합도 산출과 점포별 유동인구를 통해 점포유형 군집분석을 실시하게 되므로, 분석에 있어 중복된 자료로 인한 분석결과와 오류를 감소시키기 위해 1개의 필지에 1개의 형태로 점포 현황 자료를 구축하였다. 따라서 1개의 필지에 여러 개의 점포 현황이 있는 경우 연구자의 주관으로 점포를 선택하는 것을 배제하기 위해 랜덤함수를 사용하여 점포를 무작위로 추출함으로써 140,781개의 점포 현황 자료를 구축하였다.

그리고 본 연구의 대상지인 명동, 강남대로, 홍대합정 3개 상권의 점포 현황을 구축하기 위해 GIS 소프트웨어를 이용하여 상권 내 필지와 일치되는 점포 1,902개의 분석대상 점포 현황을 작성하였으며, 이는 <표 2> 및 <그림 3>에 제시된 바와 같다.

표 2. 분석대상 점포 현황

구분	합계	관광/여가/오락	부동산	생활서비스	소매	숙박	음식	의료	학문/교육
개수	1,902	35	76	288	492	26	821	69	95
(%)	(100)	(1.8)	(4.0)	(15.1)	(25.9)	(1.4)	(43.2)	(3.6)	(5.0)

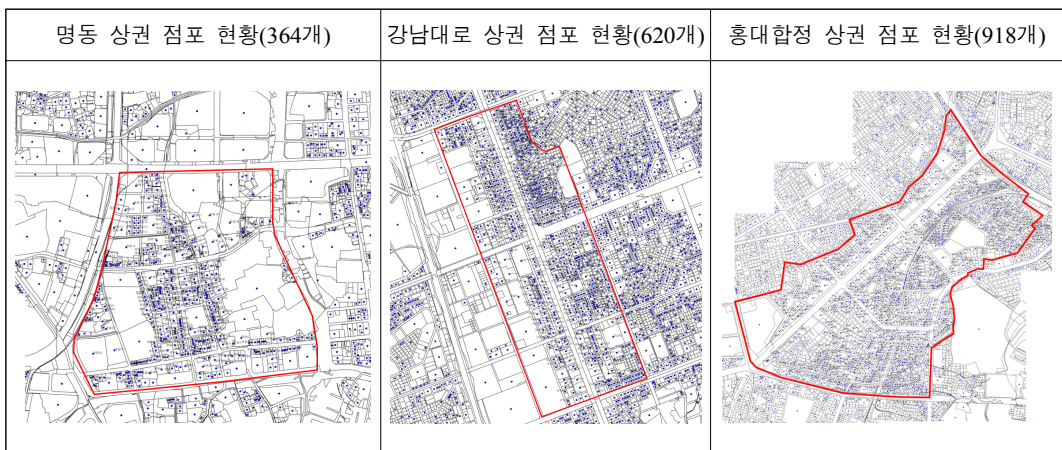


그림 3. 상권 점포 현황

3개 상권의 1,902개 개별 점포가 위치한 공간의 상대적인 가치를 측정하기 위해 공간구문적 특성인 통합도(integration)는 <그림 4>와 같이 서울대학교 도시건축공간분석 연구실에서 개발한 공간구문론 분석 프로그램인 S3 Axial 분석도구를 이용하여, 각 상권별 지적도 및 현황도를 기반으로 Autodesk사의 CAD프로그램으로 작성한 도로 축선도를 기초로 측정하였다.

개별 점포의 유동인구는 2015 서울시 유동인구 조사 보고서에 측정된 유동인구 데이터를 명동과 강남대로 상권에 적용하였고, 홍대합정 상권의 경우 해당 보고서에 조사되지 않아, 서울 열린데이터에서 제공하는 유동인구 자료를 기초로 하였으며, 축선별 유동인구는 공간 내 이동의 확률 지표를 산정하는 ERAM<sup>1)</sup>을 이용하여 자료를 변환하였다.

또한, 점포의 입지에 영향을 주는 요인을 분석하기 위해 분석대상 개별 점포의 층수, 점포가 접한 도로의 폭, 점포의 개별공시지가 자료를 조사하였으며, 소상공인시장진흥공단에서 제공하는 상가 업소정보와 서울시 열린데이터에서 제공하는 자료를 활용하였다.

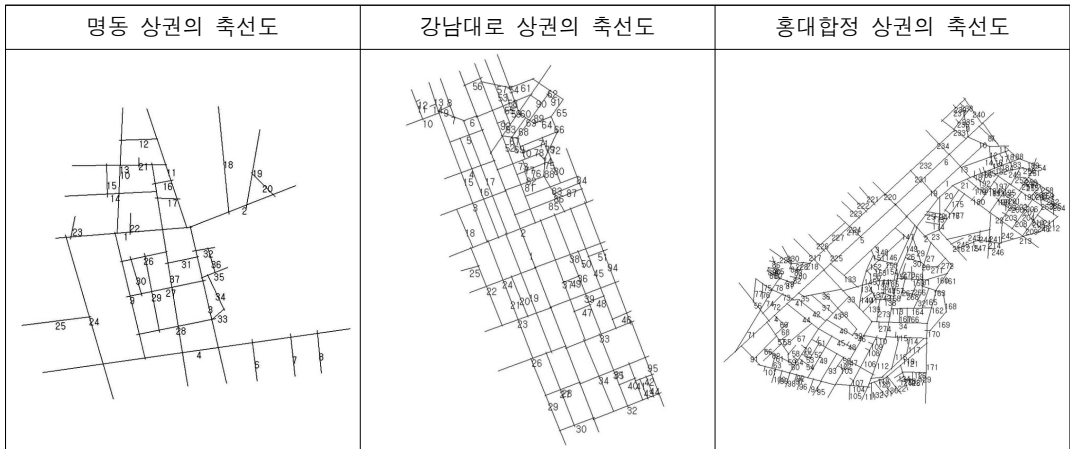


그림 4. 상권별 축선도

분석대상지인 3개 상권의 점포에 대한 점포유형(class), 통합도(integration), 유동인구(f\_pop), 층수(floor), 도로 폭(road\_w), 개별공시지가(p\_price)를 각 1,902개의 분석자료로 구성하였다. 분석자료에 대한 기초통계는 <표 3>과 같다. 기초통계 확인 결과, 점포유형은 8가지 범주로 분류되며, 통합도는 최소값 0.599부터 최대값 3.297까지 측정되었으며 평균값은 1.68이다. 유동인구는 최소값 0.026명부터 최대값 422.476명이며, 평균은 102.54명이다. 층수는 지하 2층부터 지상 14층까지 분포되며, 도로 폭은 최소값 1m부터 최대값 50m 까지, 개별공시지가는 최소값 2,851,000원/m<sup>2</sup>, 최대값 79,800,000원/m<sup>2</sup>이고 평균값은 16,800,000원/m<sup>2</sup>이다. <표 2>에서 확인할 수 있는 것과 같이, 각 변수별 단위값 상이하여 수준변수를 그대로 사용하여 분석할 경우 산출된 계수 간 격차가 클 것으로 우려된다. 이에

1) ERAM 모형은 최재필 외(2003) ‘인접행렬의 고유벡터 성분비를 이용한 공간 분석’ 에서 제안된 분석방법이다.



본 연구에서는 수치가 상대적으로 큰 유동인구 및 개별공시지가에 로그변환을 취함으로써, 변수 간 수치 차이를 줄일 수 있도록 하였다. 다만, 점포의 층수는 지하와 지상을 표현하기 위한 명목변수 속성을 가지고 있고, 도로 폭은 도로 폭 증가율에 따른 종속변수 간 관계를 확인하는 것보다 수준변수를 사용하는 것이 해석에 용이할 것으로 판단되어 별도의 변환은 실시하지 않았다.

로그변환된 유동인구의 경우 최소값 -1.539에서 최대값 8.943으로 확인되며, 로그변환된 개별공시지가는 최소값 14.863에서 최대값 18.915로 산정되었다. 이에 각 변수의 범위 값에 대한 격차가 줄어들게 되어 독립변수에 의한 종속변수 간의 상관관계 분석 시 계수 간 격차가 줄어들 것을 예상할 수 있다.

표 3. 분석자료의 기초통계

stats	class	integration	f_pop		floor	road_w	p_price	
			수준변수	로그변환			수준변수	로그변환
N	1,902	1,902	1,902	1,902	1,902	1,902	1,902	1,902
mean		1.685	1481.6	6.582		12.68	16,800,000	16.268
sd		0.616	1854.5	1.311		11.51	16,300,000	0.823
min	1	0.599	0.215	-1.539	-2	1	2,851,000	14.863
max	8	3.297	7654.8	8.943	14	50	79,800,000	18.195

## 2. 분석모형의 설정

### 1) 다항 로지스틱 회귀분석

본 연구에서는 점포유형이 갖는 통합도와 유동인구와의 관련성을 분석하기 위해 다항 로지스틱 회귀분석을 통하여 점포유형 결정에 미치는 변수들의 영향력을 검증하였다.

널리 알려진 바와 같이, 로지스틱 회귀분석은 종속변수가 이항분포 형태일 때, 로지스틱 모형에 매개변수를 추정하는 방식을 취한다. 로짓변환은 독립변수에 대한 선형성을 나타내므로 회귀모형은 식(1)과 같은 형태를 나타낸다.

$$\ln \frac{P}{1-P} = \beta_i \cdot X_i \tag{1}$$

한편 종속변수가 K개의 결과값을 가질 경우 하나의 결과값을 기준으로 선택하고 이를 제외한 K-1개의 로지스틱 회귀 모형을 만들어 낸다. K번째 종속변수를 기준으로 하면 식(2)와 같이 다항 로지스틱 모형이 된다.

$$\ln \frac{P(Y_i = 1)}{P(Y_i = K)} = \beta_1 \cdot X_i$$

$$\ln \frac{P(Y_i = 2)}{P(Y_i = K)} = \beta_2 \cdot X_i$$

.....

$$\ln \frac{P(Y_i = K-1)}{P(Y_i = K)} = \beta_{K-1} \cdot X_i \quad (2)$$

이에, 각  $\beta$  계수 값의 해석은 독립변수 1단위의 변화에 따른 종속변수 오즈비(odds ratio)의 변화율로 해석할 수 있다.

## 2) $k$ -평균 군집분석

군집분석은 동일한 그룹의 개체가 다른 그룹의 개체보다 서로 더 유사하도록 그룹화하는 작업이며, 이 분석방법은 기계학습, 패턴인식, 이미지 분석, 정보검색, 생물 정보학, 데이터 압축 등 여러 분야의 통계 데이터 분석 시 적용되는 기법이다.

알려진 바에 따르면, 군집분석에는 100개가 넘는 군집분석 알고리즘이 있으며 객관적으로 어떠한 정립된 표준 알고리즘이 정해진 바 없어, 분석 실무에서는 특정 문제에 가장 적합한 군집분석 알고리즘을 실험적·경험적 결과에 근거하여 선택하고 있다. 본 연구의 목적은 공간구문적 특성과 유동인구에 따른 점포유형을 군집별로 분류하여 점포유형별 입지 특성을 분석하는 것으로, 군집분석에 사용한 방법은 데이터마이닝에서 널리 활용되는  $k$ -평균 군집분석을 사용하였다.

$k$ -평균 군집분석은 MacQueen(1967)에 의해 처음 사용된 클러스터링 분석방법으로서, 이 기법은 군집분석 문제를 해결하는 데 가장 기초적인 비지도 학습 알고리즘이다.  $k$ -평균 군집분석은 기계학습(machine learning) 또는 인공지능 분야의 기초가 되는 비지도 학습(unsupervised learning) 분석기법의 하나로서 연구자의 주관을 배제할 수 있다는 장점을 가지고 있다.

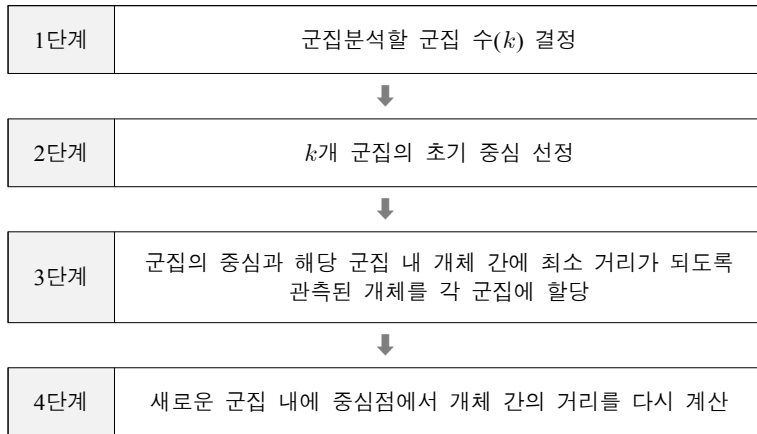
$k$ -평균 군집분석을 실시하기 위해서는 먼저 데이터를 분류할 군집 수  $k$  값을 사전적으로 결정하여야 한다.  $k$  값을 정하는 것은 분석결과에 중요한 영향을 주며,  $k$  값을 제한 없이 무작정 증가시키면 군집분석 시 오류의 양이 줄어들으나, 각 데이터 자체 군집으로 되는 경우 극단적인 오류가 발생하므로, 군집분석에 앞서 적정한  $k$  값을 결정하는 것은 매우 중요하다.

군집분석 수행을 위해  $k$  값을 정하였다면, 2단계 추정 과정을 통해 군집의 중심을 결정하는 절차를 거치게 된다. 그 첫 번째 단계는 할당단계로 군집의 중심과 해당 군집 내 개체 간에 최소거리가 되도록 각 관측값을 군집에 할당하는 과정으로서, 식(3)에 제시된 방법을 사용하게 된다.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (3)$$

$\|x_i^{(j)} - c_j\|^2$  :  $x_i^{(j)}$ (객체의 점)와  $c_j$ (군집의 중심) 사이의 측정된 거리값

두 번째는 업데이트 단계로 새로운 군집 내의 중심점에서 개체 간 거리를 다시 계산하는 것이다. 할당과 업데이트의 2개의 단계를 거치면서 더 이상 군집의 중심점이 움직이지 않게 될 때까지 반복 (iterative)하여 계산을 수행한다. 군집의 중심이 변경되지 않을 때  $k$ 개의 군집을 만들어 내며,  $k$ -평균 군집분석의 절차는 <그림 5>와 같이 나타낼 수 있다.



\* 군집의 중심이 변경되지 않을 때까지, 즉 군집 내 동일한 개체가 할당될 때까지 3, 4단계 반복

그림 5.  $k$ -평균 군집분석 절차

## IV. 실증분석 결과

### 1. 다항 로지스틱 회귀분석 결과

본 연구는 공간구문적 특성인 통합도와 유동인구에 따른 점포유형을 군집별로 분류하는 것으로, 이는 통계적 검정 과정을 거치지 않는 분석방법이다. 이에,  $k$ -평균 군집분석을 적용하기에 앞서 점포유형 결정에 통합도 및 유동인구 등 변수들이 미치는 영향력을 통계적인 유의수준 내에서 검정할 필요가 있다.

따라서 통합도와 유동인구에 따른 점포유형의 관련성 검정을 위해 다항 로지스틱 회귀분석을 선행적으로 실시하였다. 회귀분석을 위한 변수의 구성은 <표 4>와 같다.

다항 로지스틱 회귀분석은 각 종속변수가 발생할 로그 오즈비를 추정하는 방법으로 종속변수 중 기준 종속변수를 정한 후 이에 대한 상대적 계수 값을 확인하는 방식으로 해석한다. 이에 본 연구에서는 다항 로지스틱 회귀분석 시 종속변수 중 가장 높은 비중을 차지하는 음식점(43.2%)을 기준 종속변수로 상정하였다.

표 4. 다항 로지스틱 회귀분석 변수 설정

구분	변수		단위	변수설명	비고
독립변수	통합도	integration	지수	개별 점포의 통합도	
	유동인구	f_pop	명	개별 점포의 유동인구, 로그변환	로그변환
	층수	floor	층	개별 점포의 층수,	
	도로 폭	road_w	m	개별 점포가 접한 도로 폭, 로그변환	
	개별 공시지가	p_price	원/m <sup>2</sup>	개별 점포의 공시지가, 로그변환	로그변환
종속변수	점포유형	class	업종	8개 업종 (관광/여가/오락, 부동산, 생활서비스, 소매, 숙박, 음식, 의료, 학문/교육)	

다항 로지스틱 회귀분석 결과는 <표 5>와 같으며, 전반적으로 통합도 및 유동인구가 각각의 점포유형 결정에 통계적 유의수준 내에서 영향력을 미치고 있는 것으로 확인된다. 다만, 숙박 업종의 경우 통합도와 유동인구 모두 통계적 유의성이 없는 것으로 나타나 숙박업 결정에 이들 변수들은 별다른 의미를 갖지 않는 것으로 확인된다.

한편 통제변수로 사용된 층수, 도로폭, 개별공시지가는 일부 종속변수에서 통계적 유의성을 갖는 것으로 나타났으나, 전체적으로 유의성이 일관되지는 않는 것으로 확인된다.

결국 통합도와 유동인구가 점포유형에 미치는 통계적 영향력이 다항 로지스틱 회귀분석 결과를 통해 확인되기 때문에, 통합도와 유동인구에 따른 점포유형 분석을 위한 k-평균 군집분석 타당성은 확보된 것으로 판단된다.

표 5. 다항 로지스틱 회귀분석 결과

종속변수	독립변수	Coeff.	표준오차	Z	p-value
음식	(base outcome)				
관광_여가_오락	integration	0.837	0.505	1.66	0.098*
	ln_f_pop	-0.076	0.212	-0.36	0.720
	floor	-0.292	0.389	-0.75	0.453
	road_w	-0.013	0.020	-0.63	0.529
	ln_p_price	0.061	0.360	0.17	0.866
	cons	-4.560	4.960	-0.92	0.358
부동산	integration	1.036	0.345	3	0.003***
	ln_f_pop	-0.369	0.148	-2.5	0.012**
	floor	-0.080	0.228	-0.35	0.726
	road_w	0.010	0.012	0.86	0.388

	ln_p_price	0.666	0.261	2.55	0.011**
	cons	-12.644	3.640	-3.47	0.001***
생활서비스	integration	0.767	0.205	3.74	0.000***
	ln_f_pop	-0.148	0.082	-1.81	0.071*
	floor	0.113	0.094	1.21	0.227
	road_w	0.026	0.006	3.96	0.000***
	ln_p_price	0.046	0.152	0.3	0.763
	cons	-2.538	2.131	-1.19	0.234
소매	integration	-0.172	0.178	-0.97	0.334
	ln_f_pop	0.150	0.072	2.08	0.037**
	floor	-0.383	0.136	-2.81	0.005**
	road_w	0.014	0.006	2.31	0.021
	ln_p_price	0.545	0.122	4.46	0.000***
	cons	-9.838	1.680	-5.86	0.000***
숙박	integration	0.540	0.634	0.85	0.394
	ln_f_pop	-0.103	0.235	-0.44	0.661
	floor	-0.093	0.362	-0.26	0.797
	road_w	-0.067	0.043	-1.57	0.116
	ln_p_price	0.089	0.407	0.22	0.826
	cons	-4.362	5.647	-0.77	0.440
의료	integration	1.299	0.377	3.45	0.001***
	ln_f_pop	-0.668	0.165	-4.05	0.000***
	floor	0.265	0.117	2.26	0.024**
	road_w	0.022	0.012	1.8	0.071*
	ln_p_price	1.403	0.305	4.61	0.000***
	cons	-24.230	4.311	-5.62	0.000***
학문_교육	integration	1.532	0.322	4.76	0.000***
	ln_f_pop	-0.450	0.120	-3.77	0.000***
	floor	0.122	0.132	0.92	0.357
	road_w	-0.007	0.012	-0.59	0.555
	ln_p_price	0.039	0.233	0.17	0.867
	cons	-2.498	3.301	-0.76	0.449

주: p-value: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Number of obs=1,902, Pseudo  $R^2=0.0537$ .

LR  $\chi^2(35)=312.64$ , Prob> $\chi^2=0.0000$ , Log likelihood=-2752.0584.

## 2. k-평균 군집분석 결과

분석자료인 1,902개의 점포에 대해 통합도와 유동인구에 따른 8개 업종의 점포유형의 분포는 <그림 4>에 제시된 바와 같다. <그림 6>을 살펴보면, 통합도(integration)와 유동인구(f\_pop)가 높을수록 분포된 점포의 수가 줄어들었으며, 반대로 통합도와 유동인구가 낮은 경우에는 다수의 점포가 입지한 것을 확인할 수 있다. 따라서 산포도로만 확인할 경우, 길목의 상대가치가 높고 유동인구가 많은 지역의 점포 입지보다는 반대의 경우에 점포가 입지하는 경향이 뚜렷함을 알 수 있다. 한편, 산포도를 통한 시각적 분석으로는 통합도와 유동인구에 따른 점포유형을 구분하기는 불가능해 보이므로, 본 연구에서는 상술한 바와 같이 k-평균 군집분석 기법을 사용하여 통합도와 유동인구에 따른 점포유형을 분류하고자 한다.

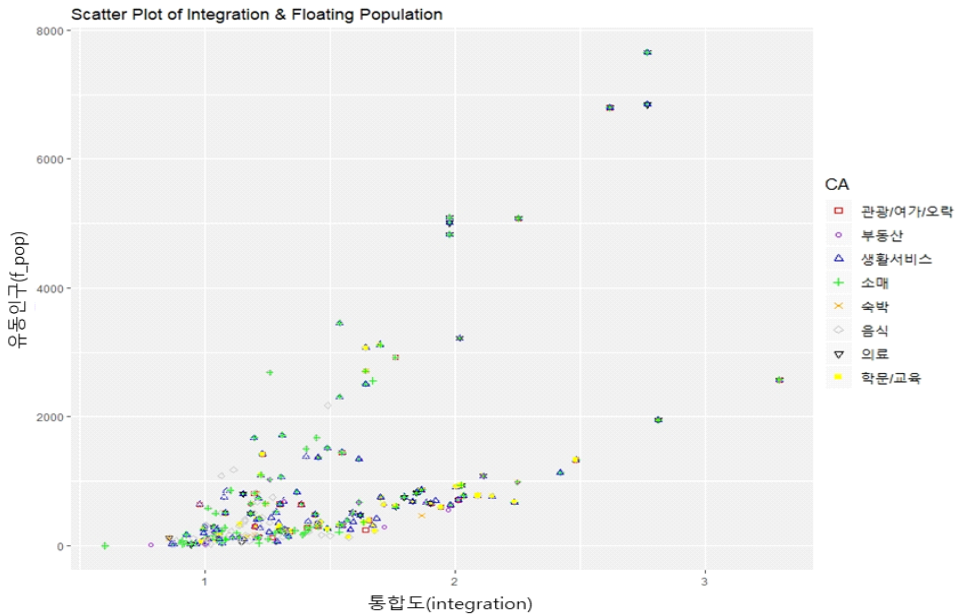


그림 6. 점포유형과 통합도, 유동인구 간의 산포도

상술한 바와 같이, k-평균 군집분석은 군집 수 k값의 결정이 분석결과에 중요한 영향을 주는 것으로 알려져 있다. 가령, 연구자가 군집 수 k를 자의적으로 결정할 경우, 분석결과와 신뢰성이 저하될 수 있다. 따라서 k값의 결정은 임의성이 배제된 계량 방법에 의할 때 분석결과와 신뢰도가 향상될 수 있다.

k값의 결정과 관련하여 여러 가지 방법이 알려져 있으나, 표준으로 확정된 방법은 존재하지 않는다.

일반적으로 Elbow Method(Thorndike, 1953)와 Silhouette Method(Rousseuw, 1987) 등이 사용되며 그 외 17여 개의 다양한 방법론이 존재한다. 다만, 각 방법론에 따른  $k$ 값 산출방법이 다르고, 이에 따른  $k$ 값 또한 상이하여 본 연구에서는 가장 널리 활용되는 분석기법인 Elbow Method를 사용하여  $k$ 값을 구하였다.

Elbow Method는 군집 수의 함수로 설명된 분산 비율을 조사하며, 군집 수에 따른 분산 비율<sup>2)</sup>을 군집 수와 비교하여 그래프를 도출하였다. 확인 결과, 특정 군집 수에서 그래프의 각도가 변하게 되는데, 이 변화가 발생하는 지점의 군집수가  $k$ 값의 선택 지점인 ‘Elbow Criterion’이라 한다.

따라서 분석자료인 1,902개의 점포 현황을 Elbow Method를 사용하여 표시한 scree plot은 <그림 7>과 같으며, 그래프의 꺾이는 부분이 둔감해지는 지점인 Elbow Criterion은 4인 지점으로 나타나므로, 군집분석을 위한 군집 수  $k$ 값은 4로 결정하였다.

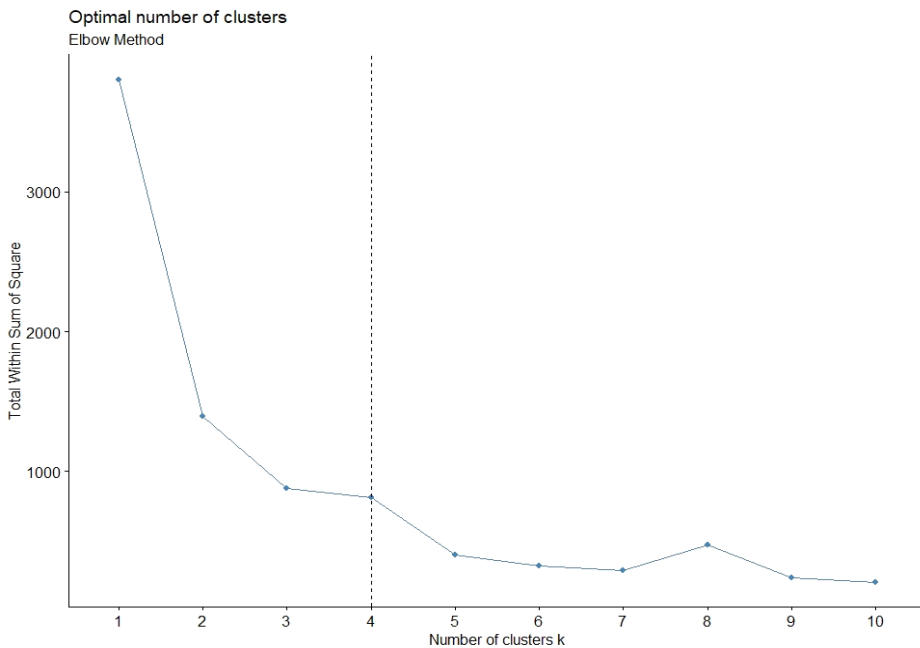


그림 7. Elbow Method 적용 scree plot

이에, 군집 수  $k$ 에 4를 적용하여  $k$ -평균 군집분석을 실시한 결과는 <그림 8>과 같다.

분류된 4개의 군집은 통합도(integration)가 가장 낮고 유동인구( $f_{pop}$ )가 가장 낮은 군집(cluster) ‘1’, 통합도와 유동인구 모두 중간 이하인 군집 ‘2’, 통합도와 유동인구가 중간 이상인 군집 ‘3’, 그리고 통합도와 유동인구가 가장 높은 군집 ‘4’로 분류되었다. 특히 군집 ‘1’과 군집 ‘2’의 경우는 중첩되는

2) 집단분산과 전체분산 간의 비율.

영역이 발생하였으며, 산포도상 대부분의 점포들이 군집‘1’과 군집‘2’에 할당되어 있어 군집 간의 뚜렷한 차이점을 발견하기에는 어려움이 있다.

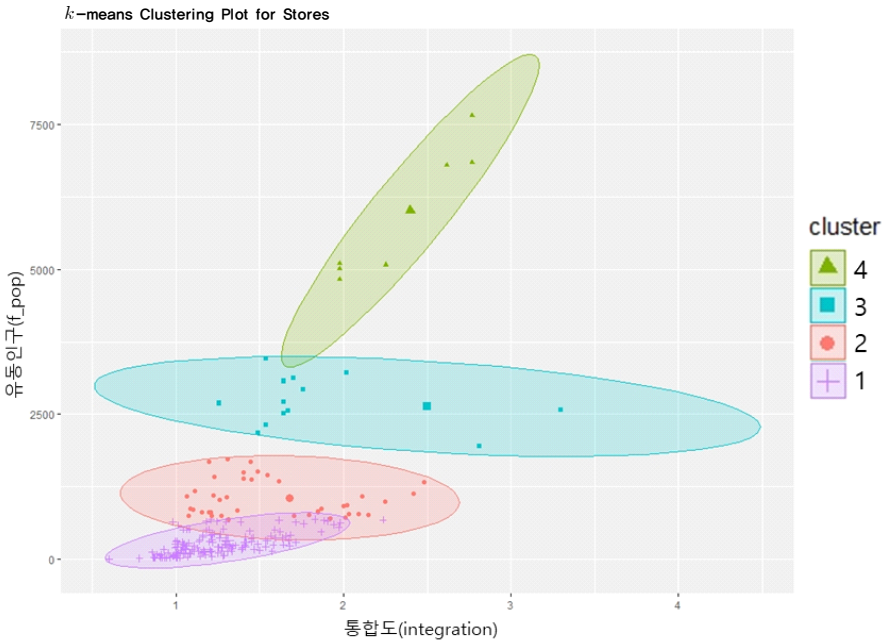


그림 8. k-평균 군집분석 결과

군집별 점포유형의 특성을 분석하기 위해 각 군집을 구성하는 개체(점포)의 현황을 자세히 살펴보면 <표 6>과 같다.

각 군집별로 군집‘1’은 877개, 군집‘2’는 586개, 군집‘3’은 215개, 군집‘4’는 224개의 점포가 할당되었으며, 각 군집에 할당된 점포의 유형 중 비중이 높은 점포유형을 확인하면, 모든 군집에서 소매 업종과 음식 업종의 비중이 높은 것으로 확인되었다. 이러한 결과는 분석자료인 1,902개의 점포 현황에서도 소매 업종과 음식 업종의 비중이 높은 것으로 보아, 군집분석 결과에서도 동일한 현상이 나타난 것으로 확인된다. 따라서 각 군집별 어떠한 점포유형이 어느 정도 특화되어 있는지 <표 6>의 내용으로는 확인하기 어렵다.

이에, 본 연구는 지역별 특화산업을 변별하는 데 오랜 기간 동안 널리 활용되고 있는 입지계수(LQ, Location Quotient)<sup>3)</sup>를 군집분석 결과에 적용함으로써 군집별 특화 점포를 확인하고자 하였다.

3) 입지계수(Location Quotient)는 1920년대 Robert Murray Haig의 경제기반이론(EBT, Economic Base Theory)에서 처음으로 제안한 분석도구로서 전체 지역의 특정 산업 종사자 비율과 해당 지역의 특정 산업 종사자를 비교하여 특정 산업의 지역 간 상대특화 정도를 확인하는 데 사용된다. LQ 계수는  $LQ = \frac{c_i/c_t}{C_i/C_t}$ 로 구할 수 있으며, 이때  $e_i$ 는 군집의  $i$ 업종의 점포 수,  $e_t$ 는 군집의 총 점포 수,  $C_i$ 는 분석대상 점포 중  $i$ 업종의 점포 수, 그리고  $C_t$ 는 분석대상 전체의 점포



표 6. 4개 군집별 점포유형 현황

구분	군집(cluster)				
	합계	1	2	3	4
합계	1,902	877 (100%)	586 (100%)	215 (100%)	224 (100%)
관광/여가/오락	35	19 (2.2%)	11 (1.9%)	2 (0.9%)	3 (1.3%)
부동산	76	30 (3.4%)	22 (3.8%)	12 (5.6%)	12 (5.4%)
생활서비스	288	125 (14.3%)	98 (16.7%)	42 (19.5%)	23 (10.3%)
소매	492	202 (23.0%)	113 (19.3%)	58 (27.0%)	119 (53.1%)
숙박	26	13 (1.5%)	9 (1.5%)	1 (0.5%)	3 (1.3%)
음식	821	418 (47.7%)	290 (49.5%)	62 (28.8%)	51 (22.8%)
의료	69	18 (2.1%)	11 (1.9%)	27 (12.6%)	13 (5.8%)
학문/교육	75	52 (5.9%)	32 (5.5%)	11 (5.1%)	-

표 7. 군집별 점포유형(업종) 입지계수(LQ)

구분	군집(cluster)			
	1	2	3	4
관광/여가/오락	1.177	1.020	0.506	0.728
부동산	0.856	0.940	1.397	1.341
생활서비스	0.941	1.104	1.290	0.678
소매	0.890	0.745	1.043	2.054
숙박	1.084	1.124	0.340	0.980
음식	1.104	1.146	0.668	0.527
의료	0.566	0.517	3.462	1.600
학문/교육	1.187	1.093	1.024	0.000

수를 의미한다.

군집별 특정 점포유형(업종)의 입지계수(LQ)를 산정한 결과는 <표 7>에 제시된 바와 같으며, 이를 시각적으로 표현한 결과는 <그림 9>와 같다.

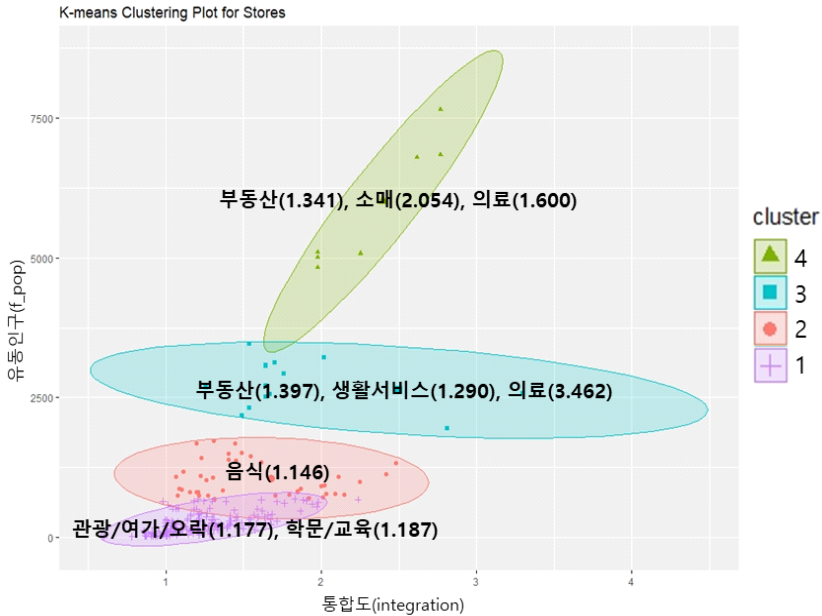


그림 9. 군집별 입지계수 특화업종

분석자료 전체 1,902개의 점포유형에 대한 각 군집별 점포유형의 특화 정도를 나타내는 입지계수(LQ) 결과는 최소값 0.000부터 최대값 3.462까지 나타났으며, 따라서 각 군집별 특화된 점포유형이 있는 것으로 확인됐다. 또한, 앞에서 살펴본 바와 같이 일반적으로 입지계수가 1.25보다 큰 경우 특화된 산업으로 볼 수 있으나, 본 연구에서 도출된 점포유형의 입지계수 결과 해석에서는 각 군집을 대표할 수 있는 업종들을 확인하기 위해 상위 70%에 해당하는 입지계수 1.125를 기준으로 정하였다.

먼저 군집‘1’의 경우, 관광/여가/오락, 숙박, 음식, 학문/교육 업종의 입지계수가 1보다 크며, 70% 이상의 상위 업종은 관광/여가/오락, 학문/교육으로 확인되었다. 군집‘1’의 입지는 상권 내 통합도와 유동인구가 낮은 지역으로 중심 상권에 입지하는 소매, 의료, 생활서비스 업종보다는 어느 정도 거리가 있는 지역, 통합도, 유동인구, 임대료가 저렴한 지역에 입지하는 현실이 반영된 결과로 볼 수 있다.

군집‘2’의 경우, 입지계수가 1보다 큰 점포유형이 관광/여가/오락, 생활서비스, 숙박, 음식, 학문/교육 업종으로 나타났으며, 입지계수 값의 상위 70%에 해당하는 1.125보다 높은 업종은 음식 업종으로 확인되었다. 군집‘2’는 통합도와 유동인구가 중간 이하로 형성되는 지역으로 음식 업종이 주를 이루는 것은 대부분의 음식 업종이 소규모 자영업자가 운영하는 경우가 많아, 통합도와 유동인구가 높아 임대료가 비싼 중심 상권지역보다는 임대료가 저렴한 지역에 산재된 현실을 반영한 결과로 추정할 수 있다.

군집‘3’의 경우, 부동산, 생활서비스, 소매, 의료, 학문/교육 업종의 입지계수가 1보다 크며, 이 중 입지계수 상위 70%에 해당하는 업종은 부동산, 생활서비스, 의료 업종이다. 군집‘3’의 경우 통합도와 유동인구가 중간 이상의 입지 특성을 가진 지역으로 어느 정도 배후지역을 가져야 입지하는 업종인 생활서비스 업종과 주로 지역 거주자가 방문하는 부동산, 의료 업종의 입지 특성을 보여주는 것으로 해석할 수 있다.

군집‘4’의 경우, 입지계수가 1보다 큰 점포유형이 부동산, 소매, 의료 업종으로 나타났으며, 3개 업종 모두 입지계수 상위 70% 값보다 높게 나타남을 확인했다. 이는 통합도와 유동인구가 높은 상권 내 중심지역으로 특히 소매 업종의 입지계수가 2.054로 높은 것은 소비자와 직접적 관련이 있는 업종의 특성을 반영한 것으로 해석된다.

## V. 결론

공간구문론적 입지 특성과 유동인구에 따른 점포의 입지유형을 확인하기 위해 본 연구는 다음과 같이 실증분석을 실시하였다. 먼저 본 연구의 분석대상인 서울시의 대표 상권으로 불리는 명동, 강남대로, 홍대합정 상권을 대상으로 개별 점포 현황자료를 구성하였다. 두 번째로 통합도와 유동인구가 점포의 입지에 영향을 주는 요소인지 확인하기 위해 다항 로지스틱 회귀분석을 활용하였다. 세 번째로, 통합도와 유동인구를 기반으로 점포유형 군집분석을 실시하여 유사한 속성을 나타내는 4개의 군집으로 분류하고, 군집별 특화 업종을 파악하기 위해 입지계수를 적용하였다. 실증분석 결과, 다음과 같은 결과를 확인할 수 있었다.

첫째, 통합도와 유동인구에 따른 점포유형(업종)과의 인과관계를 확인하기 위한 다항 로지스틱 회귀분석 결과, 대부분의 업종에서 공간구문적 특성인 통합도와 유동인구가 점포의 유형 결정에 영향을 미치고 있음을 확인할 수 있었다. 가령, 관광/여가/오락, 부동산, 생활서비스, 의료, 학문/교육 업종의 입지 결정은 통합도가 상승함에 따라 양(+)의 영향력을 미치는 것으로 확인된 반면, 유동인구의 증가는 부동산, 생활서비스, 의료, 학문/교육 업종의 입지 결정에 음(-)의 영향을 미치는 것으로 확인되었다.

둘째, 공간구문적 특성과 유동인구를 좌표 평면으로 구성된  $k$ -평균 군집분석 결과, 점포의 유형은 4개의 군집으로 분류되었으며, 통합도와 유동인구의 규모에 따라 통합도와 유동인구가 가장 낮은 곳에 입지하는 군집‘1’(점포 877개 할당), 통합도와 유동인구가 중간 이하 지역에 입지하는 군집‘2’(점포 586개 할당), 통합도와 유동인구가 중간 정도인 군집‘3’(점포 215개 할당), 통합도와 유동인구가 높은 지역에 입지하는 군집‘4’(점포 224개 할당)로 구분되었다. 각 군집에 할당된 점포유형을 보면 모든 군집에서 소매업종과 음식업종이 비중이 높은 것으로 나타났으며, 이는 분석대상인 전체 1,902개의 점포유형과 동일한 현상으로 보인다.

상술된 일련의 분석결과는 공간의 상대적 가치와 유동인구의 정도에 따라 점포의 유형이 특정될 수 있음을 시사하고 있다. 다만 본 연구에서 측정된 변수가 공간의 상대적 가치와 유동인구로 한정되지만 분석 대상을 좀 더 확장할 때, 더욱 정교한 유형 분류가 도출될 수 있음을 의미하는 것으로 해석된다. 본 연구의 결과는 비록 객관적 상가 유형 분류에 대한 기초 연구를 수행한 것에 불과하나, 충분한 자료의 수집과 보다 세분화된 분석이 수반된다면 현실이 보다 잘 반영된 결과의 도출이 가능함을 시사하는 결과로 해석된다. 이에, 본 연구의 연구결과를 초석으로 반복적이고 향상된 연구결과가 수집된다면 소규모 상공인에 대한 정책 입안 수립 시 객관적 기초분석 자료로써 충분히 활용될 수 있으리라 기대된다.

## 인용문헌

- 권영환·강동균·정소영·최윤경, 2001. “도시의 공간구조와 움직임에 관한 연구”, 『대한건축학회 학술발표대회 논문집 계획계』, 21(1): 431-434.
- 김대원·유정석, 2014. “공간구문적 특성이 상가 임대료 형성에 미치는 영향”, 『국토계획』, 49(2): 71-93.
- 김선호·신근창·양승우, 2011. “공간구문론을 통한 가로보행특성과 인터넷사진의 상관관계연구”, 『도시설계』, 12(4): 127-135.
- 김성문·안지상·심교언, 2014. “백화점 매출액이 영향요인에 관한 연구”, 『도시설계』, 15(1): 51-66.
- 김성문·안지상·심교언, 2014. “대형할인점의 매출액 영향요인에 관한 연구”, 『도시행정학보』, 27(2): 19-38.
- 김수현·김태현·임하나·최창규, 2015. “소매업의 매출액을 결정하는 보행량 및 건조 환경 요인에 관한 연구”, 『국토계획』, 50(3): 299-318.
- 김진식, 2008. “건축물 내부 상가의 위상학적 분석을 통한 입지가치 산정에 관한 연구”, 한양대학교 석사학위논문.
- 노은빈, 2015. “이용인구와 공간특성이 음식점 업종별 매출에 미치는 영향”, 가천대학교 석사학위논문.
- 박주영, 2015. “패밀리레스토랑의 입지요인이 매출액에 미치는 영향 분석”, 건국대학교 석사학위논문.
- 소상공인진흥원, 2013. 『전국 소상공인 실태조사 보고서』, 서울
- 신유진·최승두, 2014. “프랜차이즈 커피전문점의 상가 입지요인 분석 연구”, 『한국주거환경학회』, 12(3): 15-25.
- 이경민·정창무, 2014. “시간대별 유동인구가 업종별 점포 입지에 미치는 영향에 관한 연구”, 『대한건축학회 논문집 계획계』, 30(8): 47-55.
- 이호병, 2012. 부동산입지 분석론, 서울: 형설출판사
- 최재필·조형규·최현철·황용하, 2003. “인접행렬의 고유벡터 성분비를 이용한 공간 분석”, 『대한건축학회』, 19(11): 61-69.
- 황보운·하규수, 2008. “미용서비스업체의 입지요인과 경영성과에 관한 실증연구”, 『한국콘텐츠학회논문지』, 8(11): 296-305.
- Bill Hillier and Julienne Hanson, 1984. *The Social Logic of Space*, Cambridge University Press.
- Ghosh, A. and Craig, C.S., 1983. “Formulating Retail Location Strategy in a Changing Environment”, *Journal*

- of Marketing*, 47(3): 56-68.
- MacQueen, J.B., 1967. "Some Methods for Classification and Analysis of Multivariate Observations", *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, 281-297.
- Peter J. Rousseeuw, 1987. "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis", *Computational and Applied Mathematics*, 20: 53-65.
- Robert L. Thorndike, 1953, "Who Belongs in the Family?", *Psychometrika*, 18(4): 267-276.

<투고 2019.11.07, 심사 2019.12.01, 게재확정 2019.12.13>