

등록외국인을 중심으로 한 서울시 행정구역의 군집분석 연구

Cluster Analysis of Seoul's Administrative Districts Based on the Distribution of Registered Foreign Residents: Insights for Multicultural Policy Development

Sungjun Eom

January 12, 2025

Abstract

This study conducts a cluster analysis of Seoul's administrative districts focusing on the distribution and characteristics of registered foreign residents. Using the latest data from the third quarter of 2024, the research employs K-means and hierarchical clustering methods to categorize the 25 districts into distinct groups based on variables such as population structure and the ratio of foreign residents. Principal Component Analysis (PCA) is used to reduce dimensionality and enhance interpretability. The analysis identifies four clusters with unique demographic and multicultural characteristics, ranging from areas with high foreign resident concentrations and strong multicultural infrastructure (e.g., Yongsan, Jongno) to regions with predominantly local populations and minimal foreign influence (e.g., Eunpyeong, Gangbuk). The findings highlight significant regional disparities and provide insights into the factors driving these differences, such as economic stability, cultural diversity, and access to infrastructure. This study offers practical implications for policymakers by suggesting tailored multicultural policies, resource allocation strategies, and urban development plans that align with the specific needs of each cluster. It also underscores the importance of integrating data-driven insights into the formulation of inclusive and sustainable multicultural policies. Future research directions include incorporating longitudinal and qualitative data to capture dynamic changes and deepen the understanding of Seoul's multicultural landscape.

The study is done as a term project for an undergraduate class.

Contents

1	서론	3
1.1	연구 목적	3
1.2	문헌 연구	4
1.3	데이터 설명	5
1.4	분석 방법	5
1.4.1	K-평균 군집분석	6
1.4.2	계층적 군집분석	6
1.4.3	K-평균 군집분석과 계층적 군집분석의 결합	7
1.4.4	분석 결과의 해석	7
1.4.5	분석 결과의 해석	7
1.5	결과 활용 및 기대 효과	8
2	본론	9
2.1	분석 방법	9
2.2	데이터 분석 및 결과	9
2.3	분석의 타당성	16
3	결론	17
3.1	분석 결과 요약	17
3.2	분석의 장점 및 한계점	17
3.3	추가 연구사항 제안	19

1 서론

1.1 연구 목적

서울특별시(서울)는 대한민국의 수도이자 경제, 문화, 행정의 중심지로서 다양한 국적과 배경을 가진 외국인 거주자들이 지속적으로 증가하고 있다. 이러한 외국인 거주자들의 증가는 서울의 다문화적 특성을 강화하며, 세계적 도시로서의 정체성을 공고히 하는 동시에 새로운 사회적, 경제적, 행정적 도전을 제기한다. 외국인 거주자의 지역적 분포와 특성을 파악하는 것은 이들에 대한 정책적 대응과 행정적 지원을 효과적으로 수립하는 데 필수적이다.

현재 외국인 거주자들은 서울의 25개 행정구역 전반에 걸쳐 다양한 비율과 패턴으로 분포되어 있으나, 이들의 분포와 각 지역적 특성 간의 관계에 대한 체계적이고 구체적인 연구는 부족한 실정이다. 기존 연구들은 외국인 거주자의 사회적, 경제적 역할이나 특정 지역 내에서의 문화적 적응에 초점을 맞춘 경우가 많았으나, 외국인 거주자와 지역 간의 상호작용을 통합적으로 조망하는 데 한계가 있었다. 특히, 지역별로 외국인 거주 비율이나 특성이 다르다는 점을 감안할 때, 이를 정량적으로 분석하고 군집화하여 지역 간의 차이와 유사성을 밝혀내는 연구는 매우 필요하다.

본 연구는 현재까지 발표된 연구 중 가장 최신의 데이터 (2024년 3분기)를 기반으로 서울시 등록외국인의 지역별 분포를 탐구한다. 최신 데이터를 활용함으로써, 서울시 외국인 거주 현황에 대한 가장 정확하고 시의적절한 통찰을 제공한다는 점에서 본 연구는 선도적인 위치를 점한다. 이를 통해 기존 연구들의 한계를 극복하고, 데이터 기반의 새로운 분석 방법론을 제시하고자 한다.

본 연구의 목적은 서울시 등록외국인을 기준으로 각 행정구역의 분포와 특성을 분석하고, 이를 군집분석 기법을 통해 유사한 특성을 가진 지역들을 그룹화하는 것이다. 이를 통해 서울시의 외국인 거주 특성을 보다 구체적으로 이해하고, 지역 간의 차이와 유사성을 체계적으로 밝혀낼 수 있을 것으로 기대된다. 또한, 군집화된 행정구역별 주요 특성을 분석함으로써, 외국인 거주자들을 대상으로 한 맞춤형 정책과 행정서비스를 설계하는 데 필요한 실질적 자료를 제공하고자 한다.

특히, 본 연구는 다음의 세 가지 목표를 달성하는 데 중점을 둔다: 1) 서울시 등록외국인의 지역별 분포 및 특성을 최신 데이터를 활용해 정량적으로 파악하고, 데이터에 기반한 명확한 군집을 형성한다. 2) 군집별 행정구역의 공통점과 차이점을 도출하여 지역별 다문화적 특성에 대한 심층적인 통찰을 제공한다. 3) 도출된 군집을 바탕으로 서울시의 다문화 정책 수립에 기여할 수 있는 구체적인 시사점을 제시한다.

이를 통해 본 연구는 학문적 기여와 실질적 기여를 동시에 추구한다. 학문적으로는 최신 데이터와 다변량 분석 기법 중 군집분석을 활용하여 서울시의 다문화적 특성을 데이터 기반으로 체계적으로 분석하는 데 의의가 있다. 실질적으로는, 외국인 거주자의 특성을 고려한 행정구역별 맞춤형 정책을 설계하는 데 필요한 근거 자료를 제공함으로써 서울시가 더욱 효과적이고 통합적인 다문화 정책을 개발하는 데 기여할 수

있을 것이다.

결론적으로, 본 연구는 서울시의 지속 가능한 다문화 사회로의 발전과 외국인 거주자의 삶의 질 향상을 동시에 달성하기 위한 중요한 기초 자료를 제공하며, 글로벌 도시로서의 서울의 경쟁력을 강화하는 데 기여할 것이다. 특히, 최신의 연구 결과를 바탕으로 서울시 다문화 정책 설계의 근거를 제시함으로써, 정책 결정자와 학계 모두에게 유용한 지침을 제공할 것으로 기대된다.

1.2 문헌 연구

군집분석은 다양한 맥락에서 지역 특성과 구조를 분석하는 강력한 도구로 활용되어 왔다. 기존 연구들은 인구, 경제적 요인, 환경적 요인 등 다양한 변수를 기반으로 도시 및 행정구역의 특성을 탐구해 왔으며, 본 연구는 이러한 선행연구들을 기반으로 서울시의 등록외국인 분포와 특성을 분석하는 데 초점을 맞추고자 한다.

[5]은 1인 가구와 주간 인구 지수를 활용하여 대한민국의 지방자치단체를 군집화하는 연구를 수행하였다. 이 연구는 지역적 특성을 정밀하게 분석하기 위해 다변량 데이터를 활용하였으며, 이를 통해 도출된 군집은 지역 맞춤형 정책 설계의 기초 자료로 활용되었다. 이는 본 연구에서 등록외국인 데이터 기반으로 서울시의 행정구역을 군집화하는 방법론적 기반을 제공한다. [3]은 서울과 경기권을 포함한 수도권 도시를 대상으로 군집분석을 활용하여 도시 유형을 분류하였다. 이 연구는 다양한 경제적, 인구학적 데이터를 통합하여 도시 간의 유사성을 파악하고, 지역별 정책적 차별화를 제안하였다. 본 연구는 서울이라는 특정한 지리적 범위에서 유사한 접근법을 채택하여, 등록외국인을 중심으로 행정구역 간 차이를 분석하고자 한다. [4]은 다중적 지표를 활용하여 인구 감소 지역의 유형별 특성을 분석하고, 군집분석을 통해 지역적 차이와 유형을 시각화하였다. 이 연구는 특히, 다양한 변수를 통합적으로 분석함으로써 지역적 특성을 도출한 점에서 본 연구와 유사하다. 본 연구에서도 등록외국인 비율과 인구 구조, 지역적 특성을 통합적으로 분석하여 군집을 도출하고자 한다.

국내 연구뿐만 아니라 국제적인 연구들도 군집분석 기법의 유용성을 보여주고 있다. [2]은 군집분석을 활용하여 중국 서부 지역 도시들의 탄소 배출 상태를 평가하고, 지역별 특징을 고려한 분석 기법을 제안하였다. 이 연구는 환경적 변수라는 특정 요인을 중심으로 군집화를 수행하여, 지역적 차이를 체계적으로 시각화하였다. 이는 본 연구에서 등록외국인이라는 특정 변수를 중심으로 행정구역의 특성을 탐구하는 접근과 유사한 기법을 제공한다.

이와 같은 선행연구들은 군집분석 기법이 도시 및 행정구역의 다양한 특성을 규명하고, 데이터 기반의 정책 설계를 지원하는 데 있어 강력한 도구임을 입증한다. 특히, 행정구역 간의 차이를 분석하고, 지역별 맞춤형 전략을 제안하는 데 있어 군집분석의 활용 가능성은 매우 높다. 본 연구는 이러한 선행연구의 방법론적 틀을 바탕으로 서울시의 등록외국인 데이터를 심층적으로 탐구하며, 기존 연구에서 다루지 못한

외국인 거주 특성과 행정구역 간의 관계를 구체적으로 규명하고자 한다. 이러한 점에서 본 연구는 학문적 기여뿐만 아니라, 실질적인 정책적 활용 가능성을 동시에 추구하며, 서울시의 지속 가능한 다문화 사회 발전에 기여할 것으로 기대된다.

1.3 데이터 설명

본 연구에서 사용된 데이터는 서울시 등록인구 (구별) 통계[1]로, 특히 등록외국인의 분포와 특성을 분석하기 위해 다음과 같은 주요 변수로 구성되었다.

- **구:** 서울특별시의 25개 행정구역명
- **한국인남자:** 각 구에 거주하는 한국인 남성의 인구수
- **한국인여자:** 각 구에 거주하는 한국인 여성의 인구수
- **외국인남자:** 각 구에 등록된 외국인 남성의 인구수
- **외국인여자:** 각 구에 등록된 외국인 여성의 인구수

해당 데이터는 아래의 주요 특성을 가지고 있다.

지역 간 이질성 서울의 각 구는 인구 규모와 구성에서 큰 차이를 보이며, 특히 외국인 거주자의 비율은 지역마다 상당히 다르다. 예를 들어, 외국인 남성과 여성의 수는 중심지역(예: 용산구, 영등포구)과 외곽 지역(예: 도봉구, 강북구) 간에 뚜렷한 차이가 나타난다.

성별 분포 한국인과 외국인 모두 남성과 여성의 분포에서 지역적 특성이 확인된다. 특정 지역에서는 남성 비율이 높거나 반대로 여성 비율이 높은 경향이 관찰된다.

다문화적 특성 외국인 인구의 분포는 각 구의 다문화적 특성과 밀접하게 연관되어 있으며 이러한 특성은 지역별 군집화를 통해 명확히 드러난다.

이 데이터는 서울시 각 구의 인구 통계를 기반으로 하며, 다변량 분석 기법 중 군집분석을 통해 구별로 유사한 특성을 가진 그룹을 도출하는 데 활용된다. 특히, 등록외국인의 인구수와 성별 분포를 통해 서울시의 행정구역 간 차이를 규명하고, 지역적 특성을 심층적으로 이해할 수 있는 토대를 제공한다.

1.4 분석 방법

본 연구는 서울시 행정구역별 인구 통계 데이터를 활용하여 행정구역 간 유사성을 도출하고, 서울시의 다문화적 특성을 분석하기 위해 군집분석 기법을 적용하였다. 군집분석은 데이터 간의 유사성 또는 거리 기반으로 데이터를 그룹화하여 숨겨진 패턴을 발견하는 데 유용한 방법이다. 특히, 군집분석은 여러

변수들 간의 복잡한 관계를 파악할 수 있어, 지역별 인구 특성이나 외국인 거주 특성을 분석하는 데 강력한 도구로 사용된다. 본 연구에서는 K-평균 군집분석과 계층적 군집분석을 결합하여 서울시 행정구역을 분석하였다.

우선, 군집분석의 정확성을 높이기 위해 데이터의 정규화를 실시하였다. 군집분석은 거리 기반 알고리즘에 의존하기 때문에, 변수 간의 크기 차이가 분석 결과에 영향을 미칠 수 있다. 예를 들어, 어떤 변수는 수백 단위로 표현되고 다른 변수는 수십 단위로 표현될 수 있다. 이는 분석 결과를 왜곡시킬 수 있으므로, 모든 변수를 평균 0, 표준편차 1로 변환하여 정규화하였다. 이를 통해 각 변수는 동일한 척도를 가지게 되어, 변수 간의 크기 차이가 분석에 미치는 영향을 최소화할 수 있었다. 본 연구에서 사용된 주요 변수는 각 구의 한국인 남성과 여성 인구, 외국인 남성과 여성 인구로, 서울시의 인구 분포 및 다문화적 특성을 반영할 수 있는 핵심 지표들이었다.

1.4.1 K-평균 군집분석

K-평균 군집분석은 사전에 설정된 군집 수를 기준으로 데이터를 분류하는 알고리즘이다. 이 방법은 데이터 포인트를 군집 중심점에 할당한 후, 각 군집의 중심점을 반복적으로 갱신하여 최적의 군집을 형성한다. 본 연구에서는 군집 수를 4개로 설정하였다. 군집 수 설정은 사전 문헌 검토와 데이터 분석 결과를 바탕으로 결정되었으며, 서울시 행정구역의 특성과 외국인 거주 비율을 고려한 결과였다.

K-평균 군집분석의 과정은 다음과 같다. 먼저, 각 행정구역을 무작위로 군집에 배정한 뒤, 각 군집의 중심을 계산한다. 그런 다음, 모든 행정구역을 가장 가까운 중심에 할당하고, 이를 반복하여 군집의 중심을 점차 갱신한다. 반복 과정은 군집의 변화가 최소화될 때까지 진행된다. 최종적으로 각 군집은 유사한 특성을 가진 행정구역들로 구성되며, 이를 통해 서울시 내에서 지역적 차이를 명확히 도출할 수 있다.

1.4.2 계층적 군집분석

계층적 군집분석은 데이터를 계층적으로 군집화하여 군집 형성 과정을 시각화하는 방법이다. 이 방법은 유클리드 거리나 다른 거리 척도를 기반으로 각 데이터 간의 유사성을 계산하고, 점진적으로 군집을 합쳐가면서 데이터를 분류한다. 본 연구에서는 유클리드 거리를 사용하여 행정구역 간의 유사성을 측정하고, Ward 방법을 통해 군집 내 분산을 최소화하는 방식을 적용하였다. Ward 방법은 각 군집 간의 거리뿐만 아니라 군집 내 분산이 최소가 되도록 군집을 병합하는 방식으로, 군집의 내부 일관성을 강화하는 데 유리하다.

계층적 군집분석의 결과는 덴드로그램으로 시각화된다. 덴드로그램은 행정구역 간의 유사성 및 군집 간의 관계를 트리 구조로 보여주어, 행정구역 간의 계층적 관계를 직관적으로 이해할 수 있게 돕는다. 덴드로그램을 통해 행정구역들이 어떻게 결합되고, 어떤 특성을 가진 지역들이 군집을 형성하는지를 파악할 수

있다. 또한, 덴드로그램에서 특정 높이에서 가지를 잘라 군집 수를 결정할 수 있어, 최적의 군집 수를 찾는 데 유용한 도구로 활용된다.

1.4.3 K-평균 군집분석과 계층적 군집분석의 결합

본 연구에서는 K-평균 군집분석과 계층적 군집분석을 상호 보완적으로 사용하여 분석의 신뢰성을 높였다. K-평균 군집분석은 각 군집의 구성과 군집화된 데이터의 특성을 명확히 도출하는 데 강점을 가진다. 반면, 계층적 군집분석은 데이터 간의 관계를 시각적으로 표현하고, 군집 간의 유사성을 직관적으로 이해할 수 있게 돕는다. 이러한 방법론적 결합을 통해 서울시 행정구역 간의 유사성과 차이를 보다 명확히 분석할 수 있었으며, 데이터 분석 결과의 신뢰성을 확보할 수 있었다.

1.4.4 분석 결과의 해석

1.4.5 분석 결과의 해석

K-평균 군집분석과 계층적 군집분석을 통해 도출된 군집들은 서울시 내 행정구역 간 외국인 거주 비율, 인구 구조, 경제적 특성 등 주요 변수에 따라 분포되는 패턴을 효과적으로 반영하였다. 이러한 결과는 단순한 데이터의 나열에서 벗어나, 지역별 특성과 그 배경에 대한 구체적이고 체계적인 이해를 가능하게 하였다. 예를 들어, 외국인 거주 비율이 높은 지역은 특정 산업의 고용 특성, 주거 환경, 또는 사회적 네트워크와 밀접하게 연관되어 있음을 나타냈으며, 반대로 외국인 비율이 낮은 지역은 비교적 안정된 내국인 중심의 인구 구조를 보여주는 경향이 있었다.

군집 분석 결과는 지역별로 차별화된 다문화 정책 및 지역 개발 전략 수립에 실질적으로 기여할 수 있는 정보를 제공하였다. 예를 들어, 외국인 비율이 높은 군집은 다문화 지원 센터 설립, 언어 교육 프로그램 확대, 사회 통합을 위한 지역 커뮤니티 활성화와 같은 구체적인 정책적 시사점을 도출할 수 있었으며, 외국인 비율이 낮은 지역은 기존 주민 중심의 주거 안정 정책과 공공 서비스 강화에 초점을 맞출 수 있는 기반을 마련하였다.

또한, 본 연구에서 제안된 분석 방법론은 서울시에 국한되지 않고, 다른 도시나 국가의 유사한 연구에도 적용 가능한 유용한 접근법으로 평가된다. 군집분석은 특정 도시의 사회적, 인구학적, 경제적 특성을 심층적으로 이해하고, 지역 간 차이를 바탕으로 맞춤형 정책을 수립하는 데 있어 핵심적인 역할을 할 수 있다. 이러한 점에서 본 연구는 도시 계획 및 다문화적 포용성 강화를 위한 데이터 기반 정책 연구의 모범 사례로 자리매김할 수 있다.

결론적으로, 본 연구는 군집분석 기법을 활용하여 서울시 행정구역의 인구 및 외국인 거주 특성에 대한 심층적이고 다각적인 분석을 수행하였다. K-평균 군집분석과 계층적 군집분석을 결합한 방법론적 접근은 단일 기법의 한계를 보완하며, 지역 간 차이를 정확히 파악하는 데 유용한 도구로 기능하였다. 이를 통해

도출된 연구 결과는 서울시의 다문화 정책 수립 및 지역 개발 전략에 있어 실질적인 기초 자료를 제공하며, 나아가 글로벌화된 도시 환경에서 지속 가능한 발전을 위한 중요한 시사점을 제시하였다.

1.5 결과 활용 및 기대 효과

본 연구는 서울시 행정구역별 등록외국인 분포와 특성을 이해하기 위해 가장 최신 자료(2024년 3분기)를 바탕으로 수행되었으며, 이는 본 연구가 기존 연구 대비 최신 데이터를 활용한 가장 최근의 연구라는 점에서 의의를 갖는다. 최신 데이터를 기반으로 군집분석을 수행함으로써, 현재의 외국인 거주 분포 및 지역적 특성을 가장 정확히 반영할 수 있으며, 이를 통해 변화하는 사회적·문화적 환경에 적합한 정책적 시사점을 제시할 수 있다.

우선, 행정적 차원에서 군집화된 결과는 맞춤형 정책 개발과 자원 배분 최적화에 기여할 수 있다. 외국인 거주 비율이 높은 지역에서는 다문화 지원 센터 확대나 언어 교육 프로그램 강화와 같은 정책이 효과적일 수 있으며, 이러한 정책은 군집별 특성을 반영하여 설계될 수 있다. 최신 데이터를 활용한 분석 결과는 서울시가 변화하는 외국인 거주자들의 요구를 신속히 반영할 수 있도록 돕는 데 중요한 역할을 할 것이다. 또한, 각 군집의 인구 구조와 요구 사항을 기반으로 행정 서비스와 자원을 효율적으로 배분함으로써 지역 간 균형 발전을 도모할 수 있다.

도시 계획 및 지역 발전 측면에서도 연구 결과는 유용하다. 군집 간 차이를 분석함으로써 지역 간 불균형을 해소하고, 발전이 상대적으로 더딘 지역에 대한 전략을 수립하는 데 활용할 수 있다. 특히 외국인 거주 비율이 낮은 지역에서는 외국인 유입을 촉진하기 위한 방안을 모색할 수 있다. 또한, 외국인 거주자가 많은 지역에서는 문화적 다양성을 기반으로 지역 커뮤니티의 포용성을 강화하는 프로그램을 기획하여 주민 간의 상호작용을 증진시키는 데 기여할 수 있다.

본 연구는 최신 자료를 기반으로 한 연구라는 점에서 학문적으로도 큰 의의를 가진다. 이는 서울시 외국인 거주 패턴에 대한 최신 통찰을 제공할 뿐만 아니라, 다문화 사회로의 전환이 진행 중인 국내 다른 도시나 유사한 상황을 겪고 있는 국가와의 비교 연구에도 활용 가능한 자료를 제공한다. 특히, 과거 연구에서 관찰되지 못한 최근의 동향과 변화를 분석함으로써, 다문화 사회 연구의 중요한 기반 자료가 될 것이다. 연구의 기대 효과로는, 외국인 거주자 분포에 대한 데이터 기반의 이해를 통해 서울시가 글로벌 도시로서의 다문화적 포용성을 높이는 데 기여할 수 있다는 점을 들 수 있다. 또한, 유사한 군집 특성을 가진 지역 간 네트워크를 구축함으로써 외국인 주민의 지역 사회 참여를 확대하고, 지역 커뮤니티를 활성화하는 데도 도움이 될 것이다. 더 나아가 군집별 주요 특성과 요구 사항을 파악함으로써 행정 서비스 제공의 효율성을 높이고, 외국인과 한국인 간의 사회적 통합을 강화하는 데 기여할 수 있을 것이다.

본 연구의 결과는 단기적인 행정 지원뿐만 아니라, 서울시가 장기적으로 외국인 거주자와 상생할 수 있는 정책 방향성을 설정하는 데 기초 자료로 활용될 수 있다. 나아가, 가장 최신 데이터를 기반으로 한 분석

이라는 점에서 서울시가 더 나은 다문화 도시로 성장하고, 각 구의 지역적 특성을 기반으로 한 균형 잡힌 발전을 이루는 데 중요한 역할을 할 것으로 기대된다.

2 본론

2.1 분석 방법

본 연구에서는 서울시 행정구역별 등록외국인 데이터를 활용하여 행정구역 간 유사성을 분석하기 위해 군집분석 기법을 적용하였다. 군집분석은 데이터 간의 유사성을 바탕으로 그룹을 형성하여 숨겨진 패턴과 구조를 발견하는 데 유용한 분석 방법으로, 본 연구에서는 K-평균(K-means) 군집분석과 계층적 군집분석을 병행하여 사용하였다.

우선, 분석의 정확성을 높이고 변수 간 크기 차이에 따른 왜곡을 방지하기 위해 데이터의 정규화 과정을 수행하였다. 정규화된 데이터는 한국인남자, 한국인여자, 외국인남자, 외국인여자의 네 가지 주요 변수로 구성되며, 각 변수는 행정구역별 인구 분포 및 다문화적 특성을 반영한다. 이러한 변수 설정은 각 지역의 인구학적 특성과 다문화적 다양성을 보다 체계적으로 비교하고 분석하는 데 기초가 되었다.

K-평균 군집분석은 사전에 설정된 군집 수($K=4$)를 기준으로 데이터를 분류하여 행정구역 간 유사한 특성을 가진 그룹을 도출하였다. 이는 군집 내 데이터의 응집성을 극대화하고, 군집 간의 분리도를 극대화함으로써 서울시 행정구역의 다문화적 특성의 차이를 효과적으로 식별할 수 있도록 하였다. 한편, 계층적 군집분석은 행정구역 간의 계층적 관계를 시각화하고, 군집화 과정에서 데이터의 유사성이 어떻게 변하는지 파악하는 데 활용되었다. 덴드로그램(dendrogram)을 통해 군집 간의 연결 관계를 명확히 나타내고, K-평균 군집분석의 결과를 보완하여 더욱 심층적인 해석을 가능하게 하였다.

이와 같은 분석 방법론은 서울시 행정구역 간의 특성과 유사성을 도출하는 데 효과적으로 기여하였다. 특히, 등록외국인의 지역적 분포에서 나타나는 주요 패턴과 군집의 내적 특성을 식별함으로써, 각 행정구역의 다문화적 특성과 인구 분포에 대한 새로운 시각을 제공하였다. 이러한 분석 결과는 서울시 행정구역별 정책 수립 및 실행에 실질적인 자료를 제공할 뿐만 아니라, 다문화적 포용성과 지역 간 균형 발전을 촉진하는 데 중요한 기초 자료로 활용될 수 있을 것이다.

2.2 데이터 분석 및 결과

군집분석 결과, 서울시 25개 행정구역은 4개의 군집으로 분류되었다. 이는 데이터가 가진 다변량 특성을 효과적으로 요약하여 지역별로 유사한 특성을 가진 그룹을 도출한 결과이다. Figure 1는 principal component analysis (PCA)를 활용하여 96.1%의 총 분산을 설명하는 주요 변수 2개를 선택하여 시각화한 결과로, 각 군집이 명확히 구분된 모습을 보여준다. 이 시각화는 분석 결과가 데이터 내의 구조적

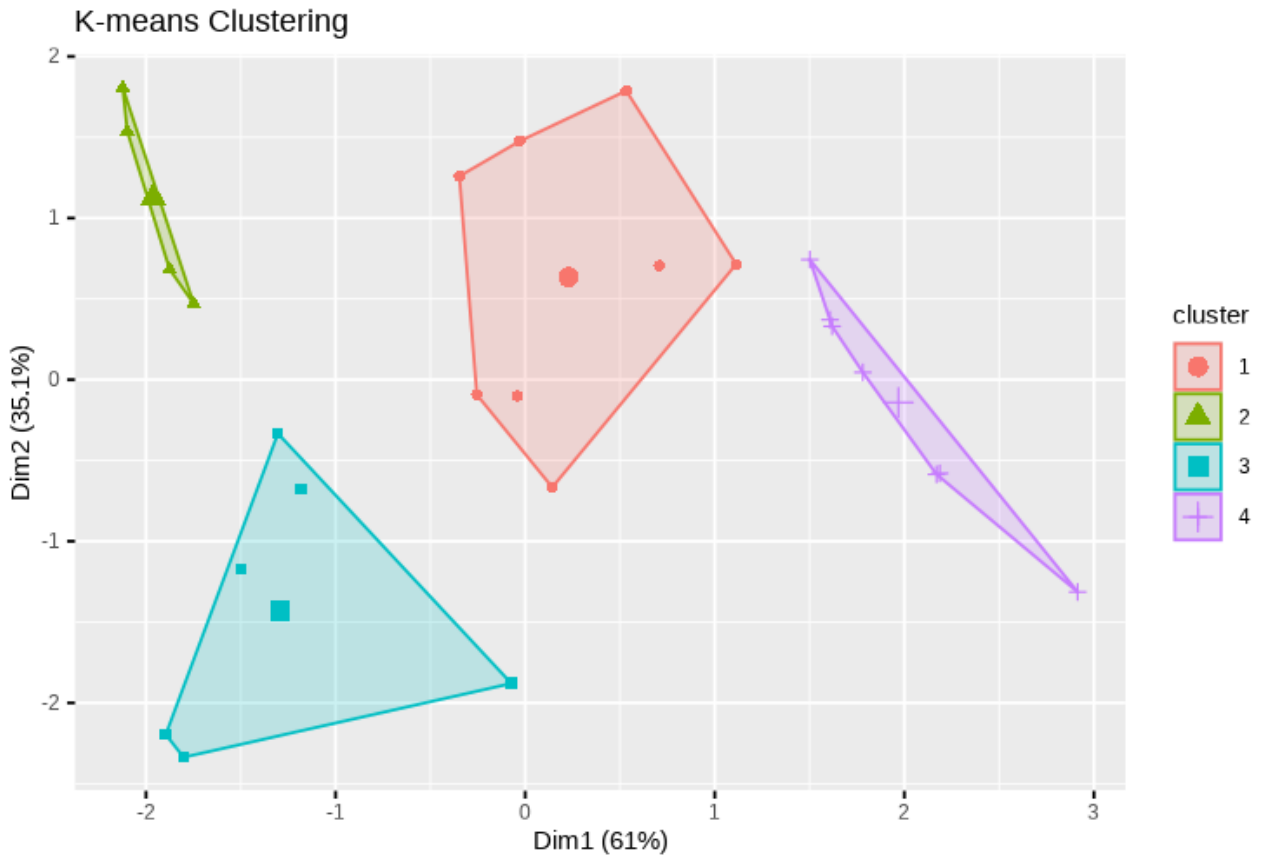


Figure 1: K-means 클러스터링 결과

군집 구

1	성동구, 중랑구, 성북구, 강북구, 도봉구, 마포구, 동작구, 서초구
2	종로구, 중구, 용산구, 금천구
3	광진구, 동대문구, 서대문구, 구로구, 영등포구, 관악구
4	노원구, 은평구, 양천구, 강서구, 강남구, 송파구, 강동구

Table 1: 군집별 행정구역

차이를 잘 반영하고 있음을 시사한다. 특히, 총 4개의 그룹은 겹쳐진 영역이 없이 분리되어 있으며, 이는 K-means 군집분석 기법이 본 데이터에 적합하다는 점을 뒷받침한다. 또한, 군집 간의 경계가 명확히 구분됨에 따라, 군집별 특성이 뚜렷하게 드러나고, 분석의 신뢰성을 높이는 데 기여한다.

Figure 2는 PCA를 통해 각 주성분이 데이터의 분산을 얼마나 설명하는지를 나타낸 scree plot이다. 이를 통해 첫 번째와 두 번째 주성분이 전체 데이터의 96.1%를 설명한다는 점을 확인할 수 있다. 특히, 세 번째 주성분부터는 설명력이 급격히 감소하여, 추가적인 주성분이 분석에 기여하는 바가 미미하다는 점이 드러난다. 이러한 결과는 PCA를 사용하여 차원을 축소하고, 분석의 단순성과 명확성을 유지하는 데 적절한 접근법이었음을 보여준다. PCA를 통해 데이터의 차원을 효과적으로 축소함으로써, 군집분석 결과의 해석 가능성을 높이고 시각화를 용이하게 하였다.

이러한 결과는 서울시 행정구역 간의 유사성을 파악하는 데 있어 K-means 군집분석이 데이터의 특성을

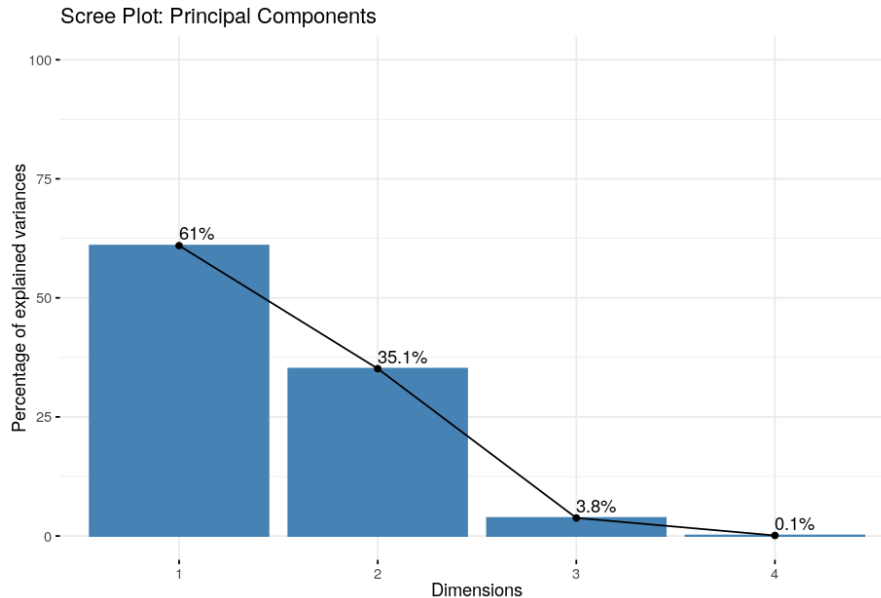


Figure 2: Scree Plot

잘 반영한 적합한 분석 방법임을 확인시켜준다. 각 군집은 서로 다른 특성을 가진 행정구역들로 구성되어 있으며, 이를 통해 서울시의 외국인 거주 및 인구 분포와 관련된 숨겨진 패턴과 구조를 효과적으로 도출할 수 있었다.

각 군집은 유사한 특성을 가진 행정구역들로 구성되었으며, 군집별 주요 특징은 다음과 같다.

첫 번째 군집은 성동구, 중랑구 등으로 구성되었으며, 이 지역들은 등록외국인 비율이 중간 수준인 특징을 보인다. 이러한 지역들은 주로 서울 외곽에 위치하며, 비교적 안정적인 인구 구조를 유지하고 있다. 이 군집은 주거 지역 중심으로 형성된 곳들이 많아, 외국인 비율은 높지 않지만 지역적 안정성과 내국인의 비율이 조화롭게 유지되는 경향을 보인다. 또한, 이 군집의 지역들은 상대적으로 저렴한 주거 비용과 교통 접근성이 외국인 거주자에게도 매력적으로 작용했을 가능성이 있다.

두 번째 군집은 외국인 거주 비율이 높은 지역들로, 종로구, 용산구와 같은 외국인 커뮤니티가 발달한 중심 지역들이 포함된다. 이 군집은 외국인 남성과 여성의 비율이 균등하며, 다양한 국적의 외국인들이 분포하는 특징을 보인다. 특히, 외국인 밀집 지역으로 알려진 이들 행정구역은 다문화적인 환경과 상업적 중심지로서의 역할을 하고 있다. 이 군집은 외국인 전용 상점, 다문화 지원 센터 등 다문화 관련 인프라가 발달해 있으며, 외국인 거주자들의 경제 활동과 문화 교류가 활발하게 이루어지고 있는 것으로 보인다. 또한 지리적 측면에서도 서울의 중심부에 위치하고 있다.

세 번째 군집은 첫 번째 군집과 유사하지만, 상대적으로 외국인의 비율이 조금 더 높은 지역들로, 광진구, 동대문구 등이 포함된다. 이 군집에 속한 지역들은 외국인과 내국인이 비교적 조화롭게 분포되어 있으며, 외국인 거주 비율이 3% 이상으로 나타났다. 다문화적 특성이 뚜렷하게 드러나며, 경제적 안정성 또한 높은 것으로 분석되었다. 이 지역들은 외국인 유학생, 단기 체류자, 그리고 다양한 배경의 외국인 근로자들이 거주하기에 적합한 환경을 제공하고 있다. 또한, 이들 지역은 주거 지역뿐만 아니라 상업과 교육 시설이

잘 발달되어 있어 외국인과 내국인의 공존이 비교적 원활한 것으로 평가된다.

네 번째 군집은 외국인 거주자 수가 극히 적은 지역들로, 노원구, 은평구 등이 포함된다. 이 군집은 외국인 거주 비율이 내국인 대비 약 0.9%로 계산되었으며, 서울시 내에서 외국인의 영향력이 상대적으로 미미한 지역들로 구성되어 있다. 이 지역들은 내국인을 중심으로 한 대규모 주거 단지와 안정적인 인구 구조를 특징으로 하며, 외국인과 관련된 다문화적 특성은 거의 나타나지 않는다. 또한 내국인 학생들을 중심으로 한 학원가가 밀집되어 있다. 이 군집은 내국인과 외국인을 합친 총 인구수가 크고, 외국인 밀집도는 낮은 경향을 보인다. 이는 외국인 거주를 위한 다문화적 인프라 부족이나 상대적으로 내국인 위주의 커뮤니티 형성이 주요 원인일 가능성이 있다.

군집별 외국인 거주 비율을 비교해 보면, 군집 2가 가장 높고, 그 다음으로 군집 3, 군집 1, 군집 4 순으로 나타났다. 이러한 군집 간 차이는 각 지역의 다문화적 특성과 경제적 특성, 그리고 외국인의 거주 환경에 영향을 미치는 정책 및 인프라 차이에 기인할 수 있다. 이와 같은 결과는 서울시 행정구역별 외국인 거주 특성을 심층적으로 이해하고, 지역 간의 차이를 반영한 맞춤형 정책 설계를 위한 중요한 기초 자료를 제공한다.

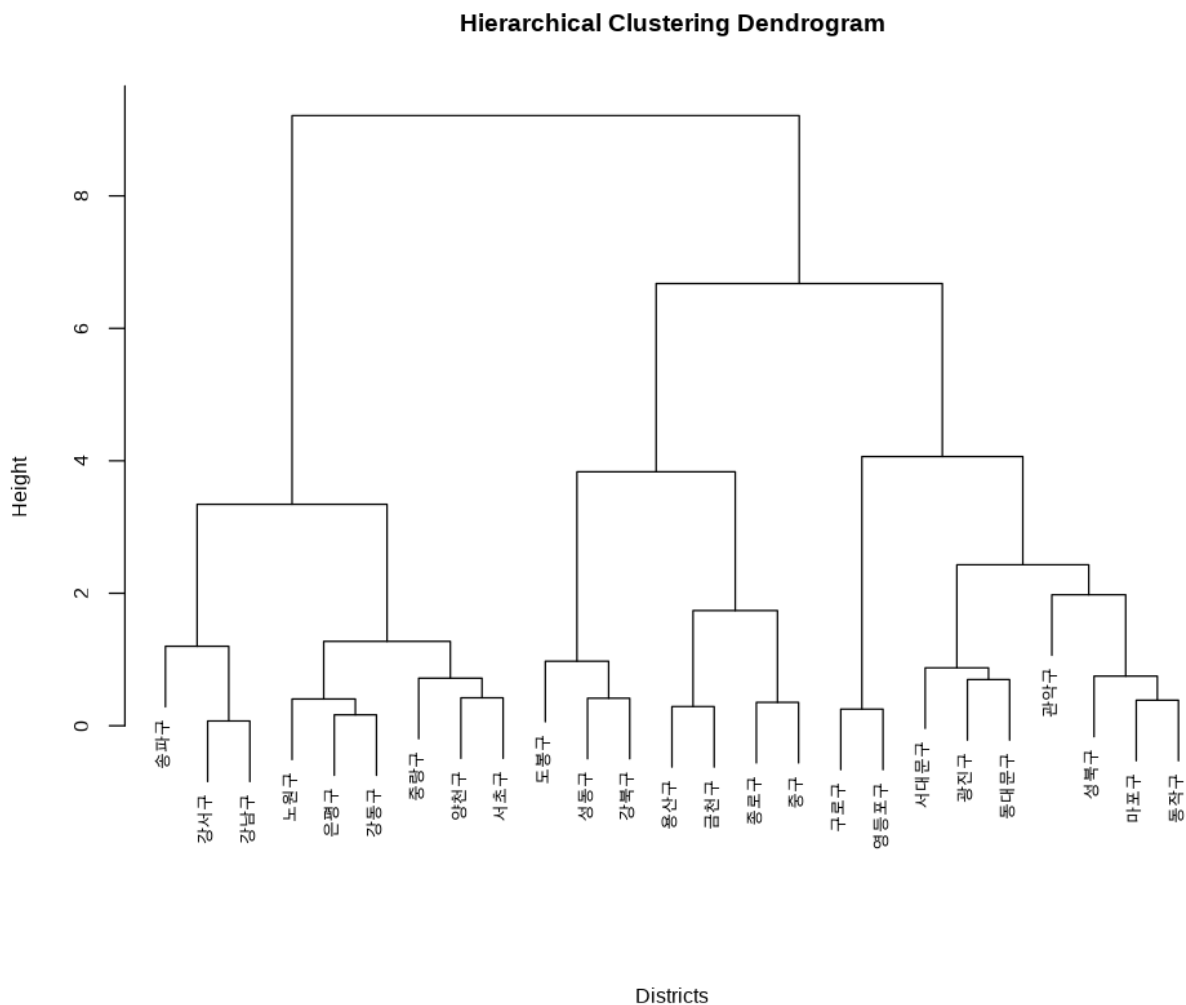


Figure 3: 계층적 군집분석 결과

계층적 군집분석 결과는 서울시 25개 행정구역 간의 유사성과 차이를 시각적으로 나타내는 덴드로그램으로 확인할 수 있다. 계층적 군집분석은 데이터 간의 유사성을 기반으로 군집을 형성하는 과정을 단계적으로 보여주는 분석 기법으로, 각 데이터 포인트가 개별 군집으로 시작하여 점차적으로 유사한 데이터끼리 병합되면서 최종적으로 하나의 군집으로 합쳐지는 과정을 시각적으로 표현한다. 본 연구에서는 Ward 방법을 활용하여 행정구역 간의 유사성을 측정하였다. Ward 방법은 군집 내 분산을 최소화하는 방식으로 작동하며, 이를 통해 유사도가 높은 두 지역부터 순차적으로 병합하여 최적의 군집 구조를 형성하였다. 덴드로그램은 이러한 군집 형성 과정을 트리 구조로 시각화한 결과물로, 행정구역 간의 관계를 직관적으로 이해할 수 있도록 돕는다. 덴드로그램의 각 가지(branch)는 행정구역 간의 유사성을 나타내며, 가지의 높이는 병합 과정에서 두 군집 간의 거리(유사성)를 반영한다. 높이가 낮을수록 두 군집 간의 유사성이 높고, 높이가 클수록 유사성이 낮음을 의미한다. 이를 통해 특정 지역들이 왜 군집화되었는지, 그리고 군집 간의 상호 관계는 어떠한지를 명확히 파악할 수 있다.

분석 결과에 따르면, 강서구와 강남구는 가장 낮은 거리에서 결합되었으며, 이는 두 구가 매우 유사한 인구 구조와 다문화적 특성을 공유하고 있음을 시사한다. 강서구와 강남구는 모두 서울시 내에서 경제적 안정성이 높고, 대규모 주거 지역과 상업 중심지를 포함하고 있는 특징을 가지고 있다. 이 지역들은 외국인 거주 비율에서 비슷한 수준을 보이며, 외국인과 내국인이 조화롭게 공존할 수 있는 환경을 제공한다. 예를 들어, 강남구는 고급 주거 단지와 국제적 비즈니스 중심지로 알려져 있으며, 강서구는 상대적으로 신도시 개발이 이루어진 지역으로, 외국인의 접근성이 뛰어난 지역이다. 이러한 공통점은 두 구가 초기 단계에서 군집화된 원인으로 볼 수 있다. K-means 결과에서도 강서구와 강남구는 동일한 군집(군집 4)으로 분류되었으며, 이는 두 지역 간의 유사성이 분석 전반에서 일관되게 나타났음을 보여준다.

반면, 관악구는 다른 군집과 결합되기 전에 비교적 높은 거리를 유지하며 독립적으로 존재하였다. 관악구는 계층적 군집분석 결과에서 성북구, 마포구, 동작구와 결합되었지만, K-means 분석에서는 이들 지역과 다른 군집으로 분류되었다. 이는 관악구가 해당 지역들과는 이질적인 특성을 가지고 있음을 나타낸다. 관악구의 이러한 특성은 외국인 거주 비율이 비교적 높고, 학군 중심지로서의 특성이 두드러지기 때문일 수 있다. 특히, 관악구는 대학가와 근접해 있는 지리적 위치로 인해 유학생이나 단기 체류 외국인이 밀집하는 경향이 있으며, 이러한 특성은 주거 구조나 인구 구성에서 다른 지역들과 차별화된다.

구로구와 영등포구는 계층적 군집분석에서 초기 단계에서 같은 군집으로 묶였으며, 오랫동안 다른 군집과 결합하지 않았다. 이들 두 구는 외국인 거주 비율이 높은 산업 지역으로, 주로 외국인 노동자들이 많이 거주하는 곳으로 알려져 있다. 구로구와 영등포구의 산업단지 및 공업지대는 외국인 근로자들에게 중요한 경제적 기회를 제공하며, 이러한 환경적 요인이 두 지역을 밀접하게 연결하는 주요 원인으로 작용했을 가능성이 높다. K-means 분석에서도 구로구와 영등포구는 동일한 군집(군집 3)에 속하였으며, 이는 두 구가 공유하는 인구 구조 및 경제적 특성이 군집 결과에 반영되었음을 보여준다.

덴드로그램의 높이를 기준으로 적절한 군집 수를 설정한 결과, 서울시 행정구역은 크게 세 개의 군집으로 나눌 수 있었다. 이러한 군집화는 행정구역 간의 유사성과 차이를 기반으로 지역별 특성을 체계적으로 파악하는 데 기여하며, 군집별로 지역적 특징이 뚜렷하게 드러난다.

첫 번째 군집은 송파구, 강남구, 강서구 등 경제적 안정성이 높고, 세대 수가 많은 지역들로 구성되었다. 이들 지역은 대규모 주거 단지와 상업 중심지를 포함하며, 비교적 높은 소득 수준과 안정적인 인구 구조를 특징으로 한다. 특히, 강남구와 송파구는 고급 주거지와 국제적 비즈니스 중심지로 잘 알려져 있으며, 강서구는 최근 신도시 개발과 교통 인프라 개선으로 인해 경제적 안정성이 강화된 지역이다. 이러한 지역들은 외국인 거주 비율이 낮거나 중간 수준이지만, 외국인 거주자를 위한 편의시설과 다문화적 인프라가 꾸준히 확대되고 있다. 따라서 이 군집은 경제적 안정성과 인구 밀집도를 바탕으로 지역적 경쟁력을 갖춘 행정구역들로 분류된다.

두 번째 군집은 도봉구, 용산구, 종로구, 중구 등으로 구성되며, 외국인 거주 비율이 매우 높은 지역과 그렇지 않은 지역이 혼재되어 있는 독특한 특징을 보인다. 용산구와 종로구는 외국인 커뮤니티가 잘 발달한 지역으로, 다문화적 특성이 강하게 나타난다. 용산구는 외국인 대사관 및 다문화 커뮤니티 센터와 같은 국제적 시설들이 집중되어 있어 다양한 국적의 외국인들이 밀집해 있다. 반면, 도봉구는 상대적으로 외국인 거주 비율이 낮지만, 특정 산업이나 주거지 형태로 인해 외국인 거주 비율이 증가할 가능성을 보인다. 이러한 군집의 특징은 지역 내의 인구 구성과 다문화적 특성이 조화롭게 공존하는 데 있다.

세 번째 군집은 구로구, 서대문구, 관악구 등으로 구성되며, 외국인 거주 비율이 중간에서 높은 수준을 유지하는 지역들이 포함되었다. 이 군집은 외국인 커뮤니티가 발달한 지역들로, 특히 구로구와 관악구는 외국인 노동자 및 유학생의 주요 거주지로 알려져 있다. 구로구는 대규모 산업단지와 공업지대가 위치해 있어 외국인 근로자들에게 경제적 기회를 제공하며, 관악구는 대학가와 근접한 지리적 이점으로 인해 유학생과 단기 체류 외국인들에게 적합한 주거 환경을 제공한다. 서대문구는 상대적으로 내국인 거주 비율이 높지만, 외국인 거주 비율도 꾸준히 증가하고 있으며, 다문화적 특성이 점차 강화되는 추세다. 이 군집은 외국인과 내국인이 조화롭게 공존하는 특징을 가지며, 외국인 거주자들에 대한 지원 정책과 커뮤니티 형성이 활발한 지역으로 볼 수 있다.

군집 간의 병합 과정은 행정구역 간의 관계를 명확히 보여주며, 지역 간 유사성과 차이를 직관적으로 이해할 수 있도록 돕는다. 덴드로그램에서 확인된 병합 순서는 서울시 행정구역의 인구 구조와 다문화적 특성에 대한 심층적인 통찰을 제공한다. 예를 들어, 강서구와 강남구는 가장 먼저 결합되었는데, 이는 이들 지역이 경제적 안정성과 인구 구성, 그리고 외국인 거주 환경에서 매우 유사하다는 것을 나타낸다. 강서구와 강남구 모두 대규모 주거 단지와 상업 중심지를 포함하며, 비교적 높은 소득 수준과 안정적인 인프라를 갖춘 지역들이다. 특히, 외국인 거주 비율이 낮은 편이지만, 국제적 비즈니스와 고급 주거 환경을 제공하는 강남구와 최근 교통 인프라와 신도시 개발로 주목받고 있는 강서구는 다문화적 요소를 수용할

준비가 잘 되어 있다. 이러한 특성은 두 지역이 초기 단계에서 군집화된 원인으로 볼 수 있다.

이와 같은 병합 과정은 특정 지역 간의 유사성을 확인하는 것뿐만 아니라, 각 지역이 가지는 특수성과 이질성을 명확히 드러낸다. 강서구와 강남구의 경우처럼 유사한 특성을 공유하는 지역들이 초기에 군집화되는 반면, 관악구와 같이 독특한 특성을 가진 지역은 다른 군집과 결합되기 전에 오랜 기간 독립적으로 존재할 수 있다. 이러한 결과는 계층적 군집분석이 단순히 데이터를 군집화하는 것을 넘어, 행정구역 간의 관계를 구조적으로 이해할 수 있는 강력한 도구임을 보여준다.

이 결과는 정책적 시사점을 제공하는 데 매우 유용하다. 군집화된 행정구역의 특성을 바탕으로 지역별로 맞춤형 정책을 설계할 수 있는 기초를 제공하며, 유사한 특성을 가진 지역 간에는 공통된 정책을 적용하고, 서로 상이한 특성을 가진 군집 간에는 차별화된 전략을 수립함으로써 보다 효과적인 정책 집행이 가능하다. 이러한 접근은 지역 간 불균형을 해소하고, 서울시 전체의 조화로운 발전을 도모하는 데 기여할 수 있다.

예를 들어, 외국인 거주 비율이 높은 지역에서는 다문화 지원 프로그램을 확대하고, 외국인 커뮤니티의 발전을 도모할 수 있는 정책을 강화할 필요가 있다. 이들 지역에서는 다문화 가정과 외국인 근로자를 위한 복지 혜택, 언어 지원 서비스, 그리고 다문화 행사 개최 등을 통해 다문화적 포용성을 증대시키는 정책이 요구된다. 또한, 외국인 거주자가 많은 지역의 사회적 통합을 촉진하기 위해 내국인과 외국인이 함께 참여할 수 있는 지역 커뮤니티 프로그램을 개발하는 것도 효과적일 것이다.

반면, 외곽 지역에서는 외국인 거주 비율이 상대적으로 낮고, 인구 유입이 필요한 특성이 있으므로, 이러한 지역에는 인프라 개발과 주거 환경 개선을 통해 내국인과 외국인 모두에게 매력적인 생활 조건을 제공할 수 있는 정책이 요구된다. 예를 들어, 교통망 확충, 공공 서비스 개선, 저렴한 주택 공급 등의 정책을 통해 인구를 유치하고 지역 경제를 활성화할 수 있다. 이러한 외곽 지역에서는 외국인의 정착을 장려하기 위한 프로그램도 병행하여 다문화적 균형을 강화할 수 있다.

더 나아가, 외국인과 내국인의 비율이 조화롭게 유지되는 지역에서는 이 조화를 유지하면서도 다문화적 이해를 심화시킬 수 있는 교육 프로그램과 문화 교류를 장려하는 정책이 필요하다. 이러한 지역은 다문화적 환경에서 이미 어느 정도 안정성을 확보한 만큼, 내외국인 간의 사회적 통합과 상호 이해를 강화하는 데 중점을 둘 수 있다. 이를 통해 다문화적 조화와 지역의 정체성을 동시에 유지할 수 있을 것이다.

이와 같은 정책적 시사점은 서울시의 균형 잡힌 발전과 다문화적 포용성을 높이는 데 중요한 기초 자료로 활용될 수 있다. 군집화 결과를 바탕으로 한 맞춤형 정책은 서울시가 단순히 행정적 효율성을 증대시키는 것을 넘어, 시민의 삶의 질을 향상시키고, 다문화 사회로의 전환을 촉진하는 데 기여할 것이다. 또한, 이 분석은 다른 도시나 국가의 다문화 정책 수립에도 벤치마킹 사례로 활용될 수 있으며, 글로벌 도시로서의 서울의 경쟁력을 한층 강화할 수 있는 전략적 자료로 의미를 가진다.

2.3 분석의 타당성

본 연구에서 사용된 분석 방법은 최신 데이터를 기반으로 하였으며, 이를 통해 서울시 행정구역 간의 유사성과 차이를 체계적으로 분석하고자 하였다. 데이터 분석 과정에서 정규화를 적용하여 변수 간 크기 차이로 인해 발생할 수 있는 왜곡을 최소화하였으며, 군집분석 기법을 통해 데이터의 숨겨진 패턴과 구조를 효과적으로 도출하였다. K-평균 군집분석과 계층적 군집분석을 병행하여 각 기법의 강점을 상호보완적으로 활용함으로써, 분석 결과의 신뢰성과 타당성을 높였다. K-평균 군집분석은 데이터의 군집 구조를 명확히 도출하는 데 유리하며, 계층적 군집분석은 행정구역 간의 계층적 관계를 시각적으로 표현하여 데이터 간의 상호작용을 보다 직관적으로 이해할 수 있도록 하였다.

또한, 변수 선정과 군집 수 설정 과정에서 기존 연구와 문헌 검토를 참고함으로써 분석의 타당성을 확보하였다. 변수로는 한국인 남녀 인구와 외국인 남녀 인구를 포함하여 서울시 인구의 다문화적 특성을 포괄적으로 반영하고자 하였다. 군집 수는 사전 문헌 검토와 데이터의 분포 특성을 바탕으로 설정하여 군집화 결과가 서울시 행정구역의 실제 특성을 충실히 반영할 수 있도록 하였다. 이를 통해 각 군집이 뚜렷한 특성을 가지면서도, 지역 간의 유사성과 차이를 효과적으로 나타낼 수 있었다.

다만, 본 연구는 몇 가지 한계를 지니고 있다. 첫째, 정량적 데이터에 기반하여 분석이 이루어졌으므로, 지역적 맥락이나 정성적 요인을 충분히 반영하지 못했을 가능성이 있다. 예를 들어, 각 지역의 문화적 배경, 주민들의 주관적 경험, 지역 고유의 정책적 요소 등은 분석 과정에서 충분히 고려되지 못했다. 이는 정량적 데이터만으로는 포착하기 어려운 지역적 특성을 간과할 수 있는 한계를 야기할 수 있다. 둘째, 군집분석에서 사용된 데이터가 특정 시점에 국한되어 있어 시간적 변화에 따른 다이내믹스를 반영하지 못하였다. 이는 지역의 인구 구조와 외국인 거주 특성이 시간이 지남에 따라 어떻게 변화하는지를 이해하는데 한계를 줄 수 있다.

이를 보완하기 위해 추가적인 정성적 데이터 수집과 분석이 필요하다. 지역 주민 인터뷰, 설문조사, 전문가 의견 수렴 등 다양한 정성적 접근 방식을 통해 각 행정구역의 고유한 맥락과 정책적 배경을 보다 풍부하게 이해할 수 있을 것이다. 또한, 다변량 시계열 분석이나 패널 데이터를 활용하여 시간적 변화와 지역 간의 상호작용을 분석하는 것도 유용할 것이다. 이러한 추가적인 분석은 본 연구의 결과를 보완하고, 지역별 정책 수립과 실행을 위한 보다 구체적이고 현실적인 기초 자료를 제공할 수 있을 것으로 기대된다.

결론적으로, 본 연구는 최신 데이터를 활용하고 정교한 분석 기법을 적용하여 서울시 행정구역 간의 다문화적 특성과 인구 구조를 체계적으로 분석하는 데 의의를 둔다. 그러나 이러한 결과를 더욱 심화하기 위해서는 정량적 분석과 정성적 분석을 결합하고, 시간적 변화를 반영한 추가 연구가 필요하다. 이는 서울시의 지속 가능한 발전과 다문화적 포용성을 증진하기 위한 정책 수립에 보다 포괄적이고 신뢰할 수 있는 기초 자료를 제공할 것이다.

3 결론

3.1 분석 결과 요약

본 연구에서는 서울시 25개 행정구역을 대상으로 등록외국인 인구 비율 및 인구 구조를 기반으로 한 군집분석을 수행하였다. K-평균 군집분석과 계층적 군집분석을 병행하여 데이터를 분석한 결과, 4개의 군집으로 행정구역을 분류하였다. 군집별로 외국인 비율, 경제적 안정성, 다문화적 특성 등의 차이를 도출하였으며, 이를 통해 지역별 특성과 주요 특징을 파악하였다.

PCA를 활용하여 데이터 차원을 축소하고 주요 변수의 설명력을 평가하였으며, 이를 통해 시각적으로 이해 가능한 2차원 공간에서 K-평균 군집 결과를 확인하였다. Scree Plot(Figure 2)을 통해 2개의 주요 변수가 전체 데이터의 96.1%를 설명함을 확인하였으며, 이 변수를 기반으로 한 클러스터링 결과는 Figure 1에서 확인할 수 있다. 계층적 군집분석에서는 Ward 방법을 사용하여 행정구역 간의 유사성을 도출하고, 덴드로그램(Figure 3)을 통해 군집 간 관계를 시각적으로 나타냈다. 이를 통해 K-평균 분석과 계층적 분석 간의 결과를 비교할 수 있었다.

K-평균 군집 결과, 첫 번째 군집은 외국인 비율이 중간 수준이며 외곽 지역에 위치한 행정구역들로 구성되었고, 두 번째 군집은 외국인 커뮤니티가 발달한 고밀도 지역들로 나타났다. 세 번째 군집은 외국인과 내국인이 조화롭게 분포된 지역으로, 다문화적 특성과 경제적 안정성이 높은 특징을 보였다. 네 번째 군집은 외국인 비율이 낮은 지역들로 구성되었으며, 내국인과 외국인을 합친 총 인구수가 큰 특징을 보였다. 계층적 군집분석 결과에서는 세 개의 주요 군집으로 분류되었으며, 강남구와 강서구가 가장 먼저 결합되는 등 지역 간 유사성을 확인할 수 있었다. 또한, 관악구는 다른 군집과 결합되기 전 독립적으로 존재하여 고유한 특성을 가진 지역임을 시사하였다.

이러한 분석 결과는 서울시 행정구역의 인구 및 외국인 거주 특성을 파악하는 데 중요한 정보를 제공하며, 지역별 특성에 맞는 정책적 시사점을 제시한다.

3.2 분석의 장점 및 한계점

본 연구는 서울시의 행정구역을 군집화하여 지역별 특성을 체계적으로 분석한 점에서 중요한 학술적 및 실질적 의미를 가진다. 데이터 정규화와 변수 선택 과정을 통해 데이터의 왜곡을 최소화함으로써 분석의 신뢰도를 높였다. 특히, K-평균 군집분석과 계층적 군집분석을 병행하여 각 기법의 강점을 상호 보완적으로 활용하였다. 이러한 접근은 단일 기법의 한계를 극복하고 데이터의 숨겨진 패턴을 다각도로 탐구할 수 있도록 하였다. 또한, PCA를 활용하여 변수 간 중복성을 제거하고 데이터의 설명력을 극대화함으로써, 주요 변수들이 분석 결과에 기여하는 정도를 명확히 시각화하였다. 이를 통해 도출된 군집은 서울시 내 특정 지역의 인구 구조와 외국인 거주 특성을 기반으로 한 정책 설계에 실질적인 기초 자료를 제공할

수 있다. 이러한 분석은 서울시의 다문화적 포용성을 증진시키고 지역 간 균형 잡힌 발전을 도모하는 데 기여할 수 있다.

특히, 본 연구는 단순히 데이터를 군집화하는 것을 넘어, 군집 결과를 통해 서울시 행정구역 간의 차별화된 특성을 도출하고, 이를 기반으로 한 정책적 시사점을 제공하였다. 예를 들어, 외국인 거주 비율이 높은 지역과 낮은 지역의 특성을 구체적으로 분석함으로써, 맞춤형 정책 설계가 가능하도록 하였다. 이러한 기여는 서울시뿐만 아니라 다른 지역이나 국가의 다문화 정책 수립에도 벤치마킹 사례로 활용될 수 있다. 그러나 본 연구는 몇 가지 한계를 지닌다. 첫째, 분석에 사용된 데이터가 등록외국인 인구에 국한되어 있어, 비등록 외국인이나 단기 체류자의 특성과 영향은 분석에 포함되지 않았다. 이는 서울시 내 외국인 거주 특성의 전반적인 파악에 있어 일부 한계로 작용할 수 있다. 예를 들어, 단기 체류자나 비등록 외국인은 특정 지역의 인구 밀도나 문화적 특성에 중요한 영향을 미칠 수 있으나, 본 연구에서는 이를 충분히 반영하지 못하였다. 향후 연구에서는 비등록 외국인 데이터를 보완하거나, 단기 체류 외국인을 포함한 분석을 통해 보다 포괄적인 이해를 도모할 필요가 있다.

둘째, 본 연구는 정량적 데이터 분석에 집중하여 지역적 맥락이나 정성적 요인을 충분히 고려하지 못하였다. 이는 지역별로 주민들의 주관적 경험, 사회적 분위기, 그리고 행정적 정책적 맥락과 같은 중요한 요소들이 분석에 반영되지 않았음을 의미한다. 예를 들어, 특정 지역의 문화적 배경이나 주민들의 다문화적 수용 태도는 정량적 데이터로는 포착하기 어려운 중요한 변수일 수 있다. 이러한 요소를 보완하기 위해서는 정성적 데이터 수집 및 분석이 필요하다. 예를 들어, 지역 주민 인터뷰, 설문조사, 또는 사례 연구를 통해 각 지역의 고유한 맥락을 반영할 수 있을 것이다.

셋째, 군집 수 설정과 변수 선정 과정에서 연구자의 주관적 개입될 가능성이 있다. 군집 수를 결정하거나 주요 변수를 선정하는 과정에서 연구자의 판단이 결과에 영향을 미쳤을 수 있으며, 이는 분석 결과의 객관성에 일부 영향을 줄 수 있다. 이러한 문제는 데이터 기반의 자동화된 클러스터링 방법이나 최적 군집 수를 제안하는 통계적 기법을 추가적으로 활용함으로써 해결할 수 있다. 또한, 다양한 변수 조합과 군집 수를 시도하여 결과의 일관성을 검증하는 과정이 필요할 것이다.

결론적으로, 본 연구는 서울시 행정구역 간의 다문화적 특성과 인구 구조를 체계적으로 분석하여 정책 수립에 기여할 수 있는 기초 자료를 제공하는 데 성공하였다. 그러나 분석의 한계를 보완하기 위해 추가적인 정성적 데이터 수집, 비등록 외국인과 단기 체류자의 포함, 그리고 보다 객관적인 변수 선정 방법론이 필요하다. 이러한 보완이 이루어진다면, 본 연구의 결과는 더욱 신뢰할 수 있는 기초 자료로 활용될 수 있을 것이다.

3.3 추가 연구사항 제안

본 연구의 결과를 바탕으로 다음과 같은 추가 연구 방향을 제안한다. 이는 본 연구의 한계를 보완하고, 서울시 행정구역 내 등록 외국인을 중심으로 한 군집 분석의 적용 가능성을 확장하는 데 기여할 것이다.

1) **정량적 분석과 정성적 분석의 결합:** 기존 연구에서 다루지 못한 지역적 맥락을 보다 심층적으로 이해하기 위해 정성적 데이터를 활용한 연구를 제안한다. 이를 위해 지역 주민 인터뷰, 설문조사, 포커스 그룹 인터뷰(FGI) 등 다양한 방법론을 통해 외국인 거주와 관련된 사회적, 문화적 요인을 탐구할 수 있다. 이러한 접근은 정량적 데이터로는 설명하기 어려운 심층적인 통찰을 제공할 것이다.

2) **시간적 변화 분석:** 다변량 시계열 분석을 통해 외국인 인구의 변화 추이와 행정구역별 지역 특성의 동태적 변화를 평가하는 연구를 제안한다. 이는 외국인 인구 유입 및 분포의 변화가 지역 사회에 미치는 영향을 이해하고, 정책적 대응 방안을 도출하는 데 도움을 줄 것이다. 특히, 외국인 관련 주요 변수의 시간적 상관성과 경향성을 분석하여 정책적 인사이트를 제시할 수 있다.

3) **다차원적 변수 확장:** 기존 군집 분석의 변수에 더해 교육 수준, 소득 수준, 고용률, 주거 형태 등의 다양한 변수를 추가로 고려한 연구를 제안한다. 이를 통해 군집의 특성을 보다 다차원적으로 파악하고, 각 군집의 내적 다양성과 복합적 특성을 규명할 수 있을 것이다. 이러한 확장은 정책 수립 시 더 세부적이고 구체적인 방향성을 제시할 수 있는 근거를 제공할 것이다.

4) **네트워크 분석:** 서울시 행정구역 간의 상호작용과 연계성을 반영하는 네트워크 분석을 통해 군집 간 관계와 영향을 분석하는 연구를 제안한다. 행정구역 간 자원, 정보, 인구의 흐름을 파악함으로써 지역 간 협력과 갈등의 구조를 이해하고, 보다 통합적인 정책 방안을 설계할 수 있을 것이다. 이는 단순한 군집 특성 파악을 넘어 지역 간의 상호작용을 반영한 동태적 분석을 가능하게 할 것이다.

5) **정책 효과성 분석:** 각 군집에 특화된 기존 정책의 효과를 분석하고, 정책적 개선점을 도출하는 연구를 제안한다. 군집별로 시행된 정책의 목표 달성 여부와 그 영향을 평가하여 실증적 데이터를 기반으로 정책 설계를 개선할 수 있다. 이를 통해 서울시 내 다양한 외국인 커뮤니티의 특성을 반영한 맞춤형 정책 개발에 기여할 수 있다.

이와 같은 후속 연구는 본 연구의 결과를 더욱 풍부하게 하고, 서울시의 지속 가능한 발전과 다문화적 포용성을 증진시키는 데 중요한 기여를 할 것으로 기대된다.

References

- [1] 서울시 등록인구 (구별) 통계. <https://data.seoul.go.kr/dataList/419/S/2/datasetView.do>.
Accessed: 2024-12-02.

- [2] Huaxing Lin, Ziqian Zhou, Shun Chen, and Ping Jiang. Clustering and assessing carbon peak statuses of typical cities in underdeveloped western china. *Applied Energy*, 329:120299, 2023.
- [3] 송민경 and 장훈. 군집분석을 이용한 수도권 도시의 유형화에 관한 연구. *대한공간정보학회지*, 18(1):83–88, 2010.
- [4] 이형석, 양한울, and 김승희. 다중적 지표를 활용한 인구감소지역 유형별 특성 분석. *고 문 오진모 한국관광개발원 원장*, 2023.
- [5] 정재준. 1 인 가구와 주간 인구 지수를 이용한 지방자치단체 군집분석. *국토지리학회지*, 56(4):435–449, 2022.