

개요

본 분석은 부산시 사이버보안 콘퍼런스를 지속적으로 개최해왔으나, 행사 참가자에 대한 분석이 부족하여 참가자 유치를 위해 지역별 인구 특성과 교통 접근성을 데이터 기반으로 분석하여, 참가자 유치 가능성이 높은 지역을 식별하여 효과적인 행사 개최하고자 한다.

데이터수집

- 인구데이터: KOSIS 부산시 구별 주민등록 인구통계를 활용하여 각 구군별 인구수 확보
- 교통데이터: 부산교통공사 지하철역 정보를 수집하여, 구별 지하철역 수를 산출

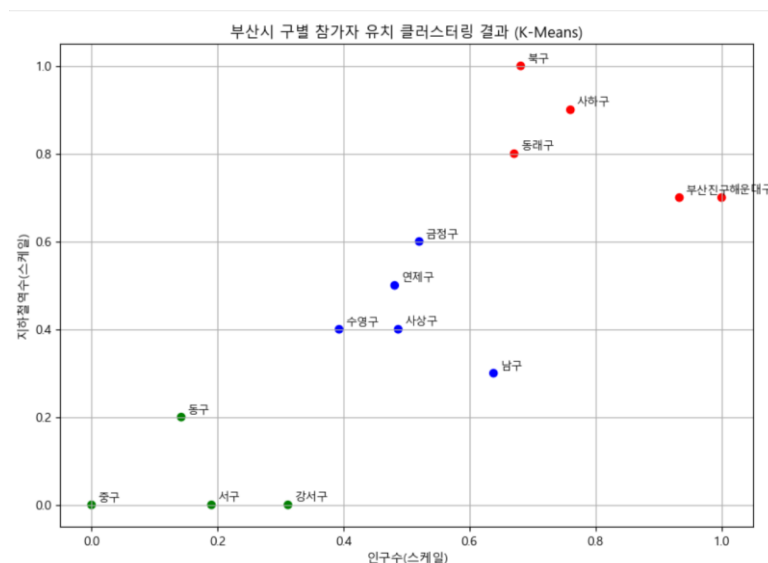
전처리

각 구별 총 인구수와 지하철역 수를 추출하여 데이터를 통합한 다음, 변수 간의 스케일 차이를 보정하기 위해 Min-Max Scaling 기법을 적용해 모든 값을 0~1 사이로 정규화시켜, 모든 변수의 중요도를 비슷하게 만든다. 정규화된 인구수와 지하철역 수를 곱하여 참가자 유치 스코어를 산출하여 이를 유치 가능성의 지표로 활용하였다.

분석방법

참가자 유치 스코어를 기반으로 유사한 특성을 가진 지역을 그룹화하기 위해 K-Means 클러스터링 기법을 적용하였다. 분석에 사용된 변수는 정규화된 인구수와 지하철역 수이며, 군집 수는 데이터의 분포를 고려해 3개로 설정하여 상위권, 중간권, 하위권 지역으로 구분하였다.

분석결과



클러스터링 결과, 다음과 같은 그룹이 도출되었다.

- 클러스터 0 (상위권 지역): 해운대구, 사하구, 북구, 부산진구, 동래구
- 클러스터 1 (하위권 지역): 중구, 서구, 동구, 강서구
- 클러스터 2 (중간권 지역): 남구, 금정구, 연제구, 수영구, 사상구

핵심 타겟 지역은 인구수와 교통 접근성 측면에서 모두 우수한 지표를 보였으며, 하위권 지역은 상대적으로 인구수와 교통 접근성 모두 낮은 수치를 기록하였다. 이 결과를 인구수와 지하철역 수를 축으로 하여 각 구군을 시각화하고, 클러스터링 결과에 따라 색상을 구분하여 군집별 특성을 표현하였다.

분석결과

분석 결과, 해운대구가 인구수와 교통 접근성 면에서 가장 높은 참가자 유치 스코어를 기록하며 기존 행사 개최지로서의 타당성이 데이터적으로 검증되었다. 또한 사하구, 북구 등 상위권에 위치한 구군들도 높은 유치 가능성을 보임에 따라, 향후에는 이들 지역을 대상으로 한 참가자 저변 확대 및 순환 개최 전략을 병행할 필요가 있음을 확인하였다.

대중교통 편의성까지 고려한 접근성 평가를 위해 구별 버스 접근성 지표를 추가 수집·분석함으로써, 교통 인프라 전반에 대한 종합 평가 체계를 구축할 수 있다는 점에서 이번 분석은 의미를 갖는다.