Q4. 출근 시간 지하철 승객 수 분석

202011906 최해원

본 분석은 티머니(pay.tmoney.co.kr)에서 제공하는 대중교통 데이터를 기반으로 진행되었으며, 2023년 3월 지하철 시간대변 이용 현황 외의 데이터는 제거하고 또한 분석에 필요 없는 시간대와 행들도 삭제하였다.

텍스트, 스크린샷, 다채로움, 그래프이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 다채로움, 그래프이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 다채로움, 그래프이(가) 표시된 사진

자동 생성된 설명

‘q4.py’로 구현한 그래프 3가지로 승객수가 많은 상위 30개의 역을 확인할 수 있었다. 승차 승객 수 상위 30개역 중 신림역이 압도적으로 높은 분포를 보였지만, 하차 승객 수 상위 30개역에 들지 못해 승하차 승객 수 상위 30개 역에선 9위에 그쳤다. 그리고 서울역, 가산디지털단지, 여의도는 승차 승객 수 상위 30개역에 상대적으로 낮은 순위에 들거나 들지 못하였지만 하차 승객수가 약 6만 이상으로 몹시 큰 분포를 가져 전체 승객 수에서 상위 3개역으로 측정되었다.

텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

자동 생성된 설명

정리하면 승차 승객과 하차 승객이 골고루 많은 역은 3개의 집단에 모두 들어간 여의도, 서울역, 구로디지털단지로 나타났고, 구로디지털단지는 승차 승객 수와 하차 승객 수가 비슷하게 나온 반면 여의도와 서울역은 하차 승객 수가 훨씬 많은 것을 확인할 수 있었다.

추가적으로, 최대 승차 30개역과 최대 하차 30개역을 비교해봤을 때 서울역, 고속터미널, 구로디지털단지, 잠실을 제외하고는 두 집단의 교집합이 존재하지 않았다. 따라서 특정 역을 제외하면 승차 승객이 많은 역은 하차 승객이 상대적으로 적고, 하차 승객이 많은 역은 승차 승객이 상대적으로 적다고 볼 수 있다.