

---

---

# 군집 분석 및 공간통계기법을 활용한 지하철-버스 환승 불편 정류장 탐색

Inconvenient Subway-Bus Transfer Station Search  
with Clustering Analysis and Spatial Analysis

---

---

- ▷ 학 생 이 름: 신제민 (2017-13717, 공과대학 건설환경공학부)  
이민규 (2017-17078, 공과대학 건설환경공학부)  
이에석 (2017-11443, 공과대학 건설환경공학부)
- ▷ 지 도 교 수: 김동규
- ▷ 지도교수 소속: 서울대학교 건설환경공학부 교통계획·물류연구실
- ▷ 대학원 도우미: 민진홍
- ▷ 도 우 미 소속: 서울대학교 건설환경공학부 교통계획·물류연구실

## - 목 차 -

I. 서론.....	4
1. 연구의 배경 및 목적 .....	4
2. 연구의 범위 및 방법론.....	4
3. 선행연구 고찰 .....	7

II. 부하량 산정 모델식을 통한 환승문제지점 예측 .....	8
1. 데이터 개요 .....	8
(1) 트립체인 데이터 .....	8
(2) 버스-노선 정류장 데이터 .....	8
(3) 지하철역 데이터 .....	9
2. 데이터 전처리 .....	9
(1) 최종도착역 기준 설정 .....	10
(2) 분석시간 .....	10
(3) 데이터 통합 .....	10
(4) 공급량 계산(노선수) .....	12
(5) 수요량 계산(인원수) .....	12
3. K-means clustering을 이용한 군집화 .....	13
4. 부하량 산정 모델 .....	14
(1) 하차정류장 부하량 계산식 .....	14
(2) 군집 부하량 계산식 .....	14
(3) 지하철 부하량 계산식 .....	14
5. 부하량 모델 적용 .....	15
(1) 모델 적용 .....	15
(2) 시각화 .....	16
6. 환승 문제지점 예측 .....	17
7. 선행연구와의 비교 .....	18
III. 상위 6개 역사에 대한 분석 .....	20
1. 분석방법 .....	20
(1) Kernel Density .....	20
(2) Getis Ord $G_i^*$ .....	20
2. 디지털미디어시티(8시~9시) .....	22
(1) 클러스터링 결과 및 부하량 모델 적용 .....	22
(2) Kernel Density .....	22
(3) 공간통계분석(Hotspot Analysis) .....	23
(4) 분석 결과 .....	23
3. 청담(8시~9시) .....	24
(1) 클러스터링 결과 및 부하량 모델 적용 .....	24
(2) Kernel Density .....	24
(3) 공간통계분석(Hotspot Analysis) .....	25
(4) 분석 결과 .....	26
4. 회기(8시~9시) .....	26
(1) 클러스터링 결과 및 부하량 모델 적용 .....	26
(2) Kernel Density .....	27
(3) 공간통계분석(Hotspot Analysis) .....	27
(4) 분석 결과 .....	28
5. 길음(18시~19시) .....	28
(1) 클러스터링 결과 및 부하량 모델 적용 .....	28
(2) Kernel Density .....	29
(3) 공간통계분석(Hotspot Analysis) .....	29
(4) 분석 결과 .....	30
6. 신림(19시~20시) .....	30
(1) 클러스터링 결과 및 부하량 모델 적용 .....	30
(2) Kernel Density .....	31

(3) 공간통계분석(Hotspot Analysis) .....	31
(4) 분석 결과.....	32
7. 낙성대(09시~10시) .....	32
(1) 클러스터링 결과 및 부하량 모델 적용 .....	33
(2) Kernel Density.....	33
(3) 공간통계분석(Hotspot Analysis) .....	34
(4) 분석 결과.....	34
IV. 사례분석 .....	34
1. 회기 .....	34
2. 낙성대.....	35
V. 결론 .....	36
VI. Appendix .....	38
1. 수도권 전체 역사별 부하량.....	38

## I. 서론

### 1. 연구의 배경 및 목적

도시 내 다양한 교통문제를 해결하고 도로 이용 효율을 높이기 위해 대중교통의 활성화가 매년 더욱 중요한 이슈로 대두되고 있다. 이에 따라 대표적인 대중교통 수단인 지하철과 버스 환승체계의 중요성은 점점 더 커지고 있지만 동일 교통수단 간 환승에 비해 타 교통수단으로의 환승서비스 수준은 아직 미흡한 점이 많은 실정이다. 국토교통부 대도시권광역교통위원회의 역사별 연계교통수단 환승서비스 수준 현황자료에서는 환승서비스 수준이 낮은 D등급 이하 역사가 69.2%에 이르렀으며 그 중에서도 최하위 등급인 F등급 역사 또한 9.3%를 차지하는 것으로 나타났다.

연계교통수단 환승서비스 수준이 낮은 원인은 크게 지하철 역과 버스 정류장 사이의 거리 같은 물리적인 문제와 버스 노선 배치 및 배차간격 등의 시스템적인 문제로 나눌 수 있다. 물리적 문제를 해결하기 위해선 상대적으로 많은 시간과 비용이 들기 때문에 본 연구에서는 시스템적인 문제에 집중하여 논의를 진행하고자 한다. 본 연구진은 지하철-버스 환승 시 대개 지하철이 간선, 버스가 지선으로서의 역할을 한다는 점을 고려하여, 지하철의 통행 수요가 특정 방향으로의 버스 통행으로 집중되었을 때 문제가 발생할 것이라고 판단하였다. 이에 본 연구에서는 지하철-버스 환승 시 통행이 특정 방향으로 집중되기 때문에 환승의 문제가 야기되는 지점을 예측 및 분석하고, 해당지점에서의 환승서비스 수준 개선방안을 제안하고자 한다.

### 2. 연구의 범위 및 방법론

통행이란 어떤 목적을 가진 사람이 이동하여 정지하기까지의 여행을 말한다. 이러한 통행은 일반적으로 수단통행과 목적통행으로 분류되는데, 하나의 통행목적 달성을 위해 대개 몇 개의 다른 교통수단을 이용하게 된다. 이때, 이들 하나하나를 수단통행이라 하고, 통행 전체를 목적통행이라 부른다. 하나의 목적통행을 고려했을 때, 지하철-버스 환승이 일어나는 경우를 생각해보면 지하철과 버스의 정류장 간 거리의 차이로 인해 지하철이 넓은 구간을 운영하면서 세부적으로 갈 수 없는 지역을 최종 수단통행으로서 버스가 분담하게 되므로 지하철은 간선으로서의 역할을 하게 되고, 버스는 지선으로서의 역할을 하게 된다.



그림 1. 지하철(간선)으로부터 버스(지선)환승 예시

지하철의 수송용량이 버스의 수송용량에 비해 압도적으로 크기 때문에 대부분의 지하철 역사에서 나타나는 지하철-버스 환승 수요는 지하철에서 하차한 승객들이 여러 버스 노선으로 분산되는 형태로 나타난다. 따라서 이러한 환승 수요가 특정 버스 노선으로 쏠리게 된다면 해당 버스 노선의 과부화로 인해 지하철에서 해당 버스로의 환승 서비스 수준이 크게 떨어지게 될 것임은 자명하다. 해당 버스 노선에 대한 수요예측 및 분석을 통해 이러한 문제를 해결할 수도 있겠지만, 본 연구진은 좀 더 근본적인 문제인 해당 버스 노선으로의 환승이 많은 이유와 어떤 경우에 문제가 될 것인지에 초점을 맞춰 연구를 진행하였다.

본 연구진은 특정 버스 노선의 환승 수요가 많을 때, 목적통행이 항상 해당 노선의 정류장으로 골고루 배분되는 것은 아니며 지하철로 한 번에 갈 수 없는 공간에 사람들을 끌어들이는 유인이 있을 것이고, 이러한 유인에 도착할 수 있는 노선이 한정되어 있는 경우 문제가 될 것이라고 판단하였다. 즉, 유인이 강력할수록 특정 방향으로 하차인원이 쏠리는 현상이 발생하고, 그 수요를 감당할 수 있는 충분한 버스 노선이 없다면 문제가 생긴다는 것이다. 또한 학교나 회사처럼 특정시간에만 유인으로서의 역할을 하는 경우 하루 전체 통행을 기준으로 쏠림을 판단할 경우 쏠림의 영향이 희석될 것이기 때문에 시간대 별로 나누어 분석을 진행하였다.

따라서 본 연구에서는 먼저 버스정류장 별 노선 수를 고려하여 각 정류장에 대한 공급량을 계산해 그 역수를 가중치로 설정하고, 시간대 별로 버스 각 정류장에서 내리는 하차인원을 산정하여 시간대 별 정류장의 수요량을 계산하였다. 이후 k-means clustering을 이용하여 지하철-버스 환승 최종목적지를 분류한 후 수요량만을 고려하여 군집 비율을 산정하였다. 앞서 구한 공급량과 수요량을 통해 최종하차역 별 부하량을 계산하고, 이를 군집별로 묶어 시간대별 군집 내 부하량의 합을 구하고, 군집 비율을 가중치로 하여 하나의 지하철역에 대해 군집 부하량을 구하였다. 이 중 가장 큰 값을 지하철 부하량으로 정의하고 다른 지하철 역의 값과 비교해 봄으로써 환승 문제지점을 예측하였다. 부하량 비교를 통한 환승 문제지점 예측의 전체 프로세스는 다음과 같다.



그림 2. 부하량 비교를 통한 환승 등

이후 환승 문제가 예상되는 상위 6개 역사(디지털미디어시티, 청담, 회기, 길음, 신림, 낙성대)에 대해 Kernel Density를 이용한 Heat Map을 제작하였고 이를 통해 어떠한 유인 때문에 환승 수요의 쏠림이 발생하는지 간략한 분석을 진행하였다. 또한 공간통계분석 기법 중 하나인 Hot Spot Analysis를 통해 쏠림의 유인에 대하여 구체적인 분석을 진행하였다. 최종적으로 회기, 낙성대 2개 역사에 대하여 설문조사를 통한 실제 사례분석을 진행하였고 본 연구에 대한 전반적인 프로세스는 다음과 같다.

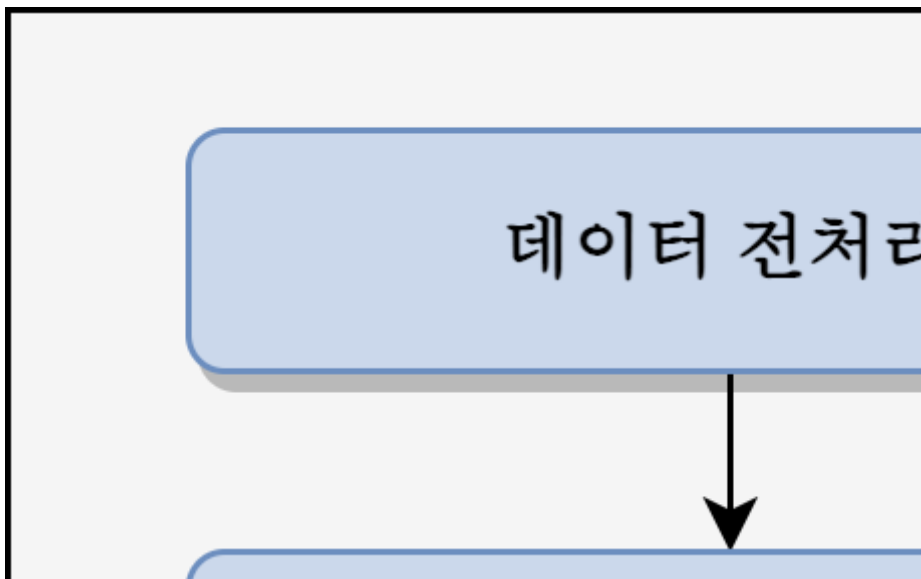


그림 3. 연구 프로세스

### 3. 선행연구 고찰

수도권 대중교통 환승행태별 평균 통행시간에 관한 연구(2010)에서 버스-지하철 환승시

간(4.49분)에 비해 지하철-버스 환승시간(7.39분)이 오래 나타나는 것으로 확인되었다.<sup>1</sup> 황보희, 노정현, 박영욱(2006)은 대중교통 환승이 경로선택에 미치는 영향을 토대로 환승시간 가치를 추정하고 환승행태별 효용의 차이를 분석하였고, 대중교통 이용객들의 환승시간가치가 매우 낮아 환승을 하더라도 전체 통행시간을 줄이려고 하지만 환승행태별 효용의 차이에서는 지하철-버스 환승이 가장 비효율적인 것으로 파악되었다.<sup>2</sup> 김대성, 엄진기, 문대섭, 최명훈, 송지영(2011)은 대중교통 카드 데이터를 이용하여 대중교통 이용자의 환승 비율이 쇼핑지구와 업무지구와 비교 주거지역에서 더 높은 것을 알아내었으며, 환승 버스 이용자의 최소 통행거리가 최대 보행거리와 유사할 것이라는 가정 하에 최대 보행거리를 300~400m로 추측하였다.<sup>3</sup> 최명훈, 엄진기, 이준, 김대성(2011)은 스마트카드 데이터를 활용하여 수도권 도시철도 역사에서 버스 정류장까지의 환승시간에 대한 분석을 수행하고 군집분석을 통해 A~F의 6단계 등급으로 분류하는 기준을 제시하였다.<sup>4</sup> 손무성, 김광태, 김세원, 오석문(2011)은 도시철도 환승시간 최소화를 위해 3개역, 40개 열차에 대해 열차시각표 기반 최적화 모형을 제시하고 각 환승역에서 노선간 환승시간을 최소화하도록 조정된 열차시각표를 제공하였다.<sup>5</sup> 이상혁, 김재석, 김민석, 우용한(2013)은 다항로짓모형을 이용하여 통행비용, 통행시간, 성별, 나이 등 7가지 설명변수의 변화가 환승교통 수단선택에 어떠한 영향을 미치는지 분석하였는데 통행 시간이 길어질수록 버스보다는 지하철-지하철 환승을 선호하는 것으로 나타났으며 지하철과 버스 간 환승이 전체 환승에서 가장 높은 비율을 차지하는 것으로 분석되었다.<sup>6</sup> 신호석, 이은학, 이호영, 고승영, 김동규(2017)는 수도권 지하철-버스 환승의 관한 연구에서는 초록 버스(지선)의 통행 수가 가장 높은 것으로 나타났으며 2호선 지하철의 환승객들을 거주지역으로 분산시켜주는 역할을 하는 것으로 나타났다.<sup>7</sup> 국토교통부 대도시권광역교통위원회에서 발표한 역사별 연계교통수단 환승서비스 수준 현황자료에 따르면 107개 주요 환승거점 역사 중 LOS A등급인 역사는 1개도 없었으며 LOS D등급 이하의 역사가 69.2%, LOS F등급 역사가 9.3%로 연계교통수단 환승 서비스의 심각성을 확인할 수 있었다.<sup>8</sup>

선행연구들로부터 지하철과 버스 간 환승이 전체 환승 중에서 가장 높은 비율을 차지하지만, 역설적이게도 가장 비효율적으로 운행되고 있음을 알 수 있었다. 반면 기존의 연구들은 주로 동일교통수단 환승을 최적화하는게 그치기 때문에 연계교통수단 환승에 관한 추가적인 연구의 필요성을 시사한다. 또한, 지하철과 버스 간 환승에 있어 버스가 주로 지선역할을 하는 것으로 나타나기 때문에 본 연구의 범위를 간선으로서의 지하철과 지선으로서의 버스로 좁히는 데에 타당성을 얻을 수 있었다.

## II. 부하량 산정 모델식을 통한 환승문제지점 예측

### 1. 데이터 개요

#### (1) 트립체인 데이터

교통카드 빅데이터<sup>1</sup>는 대중교통수단 이용자의 통행실태 파악을 위해 수집 및 공유되는 데이터로서, 티머니, 로카모빌리티, 마이비, 전국버스운송사업조합연합회 등으로부터 수집된다.

한 개인은 대중교통 이용 시, 하나의 수단에만 국한되지 않고 여러 교통수단을 혼용하여 목적지에 도달한다. 트립체인 데이터는 한 개인의 이러한 출발 지점과 도착 지점 사이의 통행 경로를 교통수단을 기준으로 나뉘어서 정리한 데이터이다. 카드번호는 각 개인의 고유한 식별자로서 존재하며, 행(레코드)은 카드 번호를 식별자로 하여 개인별 트립체인을 의미한다. 사용한 트립체인 데이터는 2017.05.22.일자 서울특별시에서 수집된 교통카드 데이터를

<sup>1</sup> <https://www.stcis.go.kr/wps/bizintro/trcrdbgdataintro/trcrdbgDataIntroView.do> 교통카드 빅데이터 통합정보시스템

기반으로 구성되어 있고, 14985884 rows x 60 columns로 구성된다.

	카드번호	노선ID1	노선ID2	노선ID3	차량ID1	차량ID2	차량ID3	승차일시1	하차일시1	승차역ID1	승차역ID2	하차역ID1	하차역ID2	최종하차역ID
0	+++PAYzWTdRTG	~	0	0	~	0	0	20170522094417	20170522095902	412	0	416	0	416
1	+++PAYzWTdR	~	0	0	~	0	0	20170522100648	20170522102019	416	0	412	0	412
2	+++PAYzWT	11110135	0	0	111742007	0	0	20170522190631	20170522192510	15047	0	8003053	0	8003053
3	+++rrVOsej	~	0	0	~	0	0	20170522104846	20170522113931	1806	0	1003	0	1003
4	+++rrVOsej+	~	0	0	~	0	0	20170522141301	20170522145125	1003	0	2748	0	2748
5	+++rrVOsej+h	11110265	0	0	111749731	0	0	20170522191111	20170522192123	71346	0	10503	0	10503
6	++/D2QQm	28206005	0	0	128715455	0	0	20170522112657	20170522113834	2805997	0	2800627	0	2800627
7	++/D2QQm	28010004	0	0	128702540	0	0	20170522130021	20170522130448	2800627	0	2800670	0	2800670
8	++/D2QQcZ	28030014	0	0	128705987	0	0	20170522163406	20170522164426	2800691	0	2800624	0	2800624
9	++/lqOVVxql	11110953	0	0	111741741	0	0	20170522214231	20170522215558	8001827	0	75399	0	75399
10	++/nWdvgc8E	~	0	0	~	0	0	20170522183323	20170522192318	1850	0	1030	0	1030
11	++/pDf79IY7	11110608	0	0	111758703	0	0	20170522082529	20170522083135	9012659	0	9009890	0	9009890
12	++/pDf79IY70	~	0	0	~	0	0	20170522140820	20170522150302	3117	0	2742	0	2742
13	++0+lcJIX3s	11110602	0	0	111758802	0	0	20170522182417	20170522182832	9009398	0	9012624	0	9012624
14	++0DIweEEP	41110040	0	0	111711469	0	0	20170522182405	20170522184249	9110	0	8598	0	8598
15	++0EqMfPC	~	0	0	~	0	0	20170522174243	20170522180052	150	0	157	0	157
16	++0OMibbg2	~	0	0	~	0	0	20170522161705	20170522164245	1024	0	2729	0	2729
17	++0XX59KeT	11110010	0	0	111743282	0	0	20170522225922	20170522234406	72572	0	71922	0	71922
18	++0h3mwFT	~	0	0	~	0	0	20170522222357	20170522224955	2754	0	2756	0	2756
19	++0rN0HjC9	11110045	0	0	111749952	0	0	20170522085513	20170522092004	73772	0	8049	0	8049

표 1. 트립체인 데이터 (2017.05.22.기준)

## (2) 버스-노선 정류장 데이터

버스 노선이 지나는 정류장ID와 정류장 이름, 위치를 나타내는 데이터이다. 2017.07. 기준으로 존재하는 데이터를 사용하였다. 본 연구에서는 행정 구역을 기준으로 서울특별시에 대해서만 분석을 진행했다.

노선ID	정류장ID	버스노선명	버스정류장명	정류장X좌표	정류장Y좌표	시도구분
11110061	7046	606번(부천상동~종로1가)	자동차매매단지	126.448493	37.293272	경기도
41110067	7046	673번(상동기점~이대부고)	자동차매매단지	126.448493	37.293272	경기도
41110211	7046	661번(부천상동~영등포역,신세계백화점)	자동차매매단지	126.448493	37.293272	경기도
11110061	7086	606번(부천상동~종로1가)	상도중학교	126.455783	37.296058	경기도
41110067	7086	673번(상동기점~이대부고)	상도중학교	126.455783	37.296058	경기도
11110061	7097	606번(부천상동~종로1가)	상도중학교	126.455997	37.296126	경기도
41110067	7097	673번(상동기점~이대부고)	상도중학교	126.455997	37.296126	경기도
41110079	7117	7728번(대화동~신촌)	탄현역	126.457578	37.414343	경기도
41110079	7118	7728번(대화동~신촌)	탄현역	126.457672	37.414765	경기도
41110078	7138	7727번(설문동~신촌)	건영아파트	126.460135	37.418135	경기도
41110078	7140	7727번(설문동~신촌)	탄현동문아파트	126.46038	37.419789	경기도
41110078	7146	7727번(설문동~신촌)	탄현효성주은아파트	126.460989	37.414371	경기도
41110079	7184	7728번(대화동~신촌)	일산터미널	126.46394	37.411317	경기도



41110079	7203	7728번(대화동~신촌)	북일산전화국.일산2동주민센터	126.465742	37.410946	경기도
11110061	7206	606번(부천상동~종로1가)	부천대학입구.KT부천지사	126.466799	37.291905	경기도
41110211	7206	661번(부천상동~영등포역,신세계백화점)	부천대학입구.KT부천지사	126.466799	37.291905	경기도
11110061	7230	606번(부천상동~종로1가)	부천상상거리	126.468113	37.2917	경기도
41110211	7230	661번(부천상동~영등포역,신세계백화점)	부천상상거리	126.468113	37.2917	경기도
11110061	7238	606번(부천상동~종로1가)	노마즈하우스.우리은행	126.47012	37.293127	경기도
41110211	7238	661번(부천상동~영등포역,신세계백화점)	노마즈하우스.우리은행	126.47012	37.293127	경기도

표 2. 버스-노선 정류장 데이터 (2017.07. 기준)

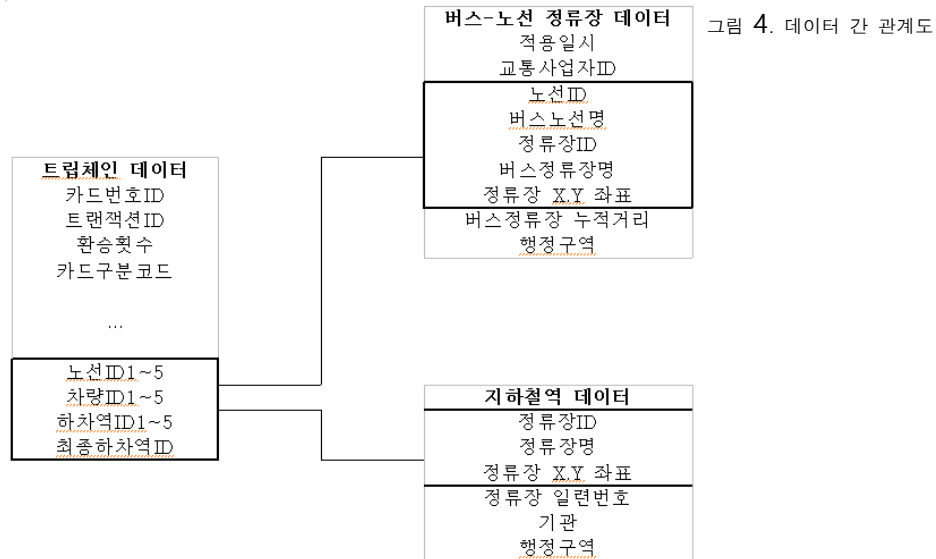
### (3) 지하철역 데이터

버스-노선 정류장 데이터에 지하철역에 대한 정보는 없어서 지하철역 데이터를 추가적으로 사용했다. 지하철역 데이터는 지하철역 ID, 역명, 역 위치에 관한 정보를 담고 있다. 본 연구에서는 행정 구역을 기준으로 서울특별시에 대해서만 분석을 진행했다.

기관	정류장ID	정류장 일련 번호	정류장명	행정구역				X	Y
				대 존	중 존	소 존	행정코드		
전철/지하철	150	1	서울역	서울특별시	중구	회현동	1102054	126.971133	37.554383
전철/지하철	151	2	시청역	서울특별시	중구	소광동	1102052	126.97714	37.565343
전철/지하철	152	3	종각역	서울특별시	종로구	종로1-2-3-4가동	1101061	126.983102	37.570164
전철/지하철	153	4	종로3가역	서울특별시	종로구	종로1-2-3-4가동	1101061	126.99268	37.570462
전철/지하철	154	5	종로5가역	서울특별시	종로구	종로5-6가동	1101063	127.001549	37.570964
전철/지하철	155	6	동대문역	서울특별시	종로구	종로5-6가동	1101063	127.010613	37.571638
전철/지하철	156	7	신설동역	서울특별시	종로구	송인2동	1101071	127.024509	37.576044
전철/지하철	157	8	제기동역	서울특별시	동대문구	제기동	1106082	127.034924	37.578165
전철/지하철	158	9	청량리역	서울특별시	동대문구	청량리동	1106080	127.044727	37.580019
전철/지하철	159	10	동묘앞역	서울특별시	종로구	창신1동	1101067	127.016346	37.573198
전철/지하철	201	11	시청역	서울특별시	중구	소광동	1102052	126.975539	37.56365
전철/지하철	202	12	율지로입구역	서울특별시	중구	명동	1102055	126.982161	37.565996
전철/지하철	203	13	율지로3가역	서울특별시	종로구	종로1-2-3-4가동	1101061	126.990967	37.566277
전철/지하철	204	14	율지로4가역	서울특별시	종로구	종로1-2-3-4가동	1101061	126.997271	37.566587
전철/지하철	205	15	동대문역사문화공원역	서울특별시	중구	광희동	1102059	127.008977	37.565669
전철/지하철	206	16	신당역	서울특별시	중구	신당동	1102069	127.019467	37.565656
전철/지하철	207	17	상왕십리역	서울특별시	성동구	왕십리도선동	1104071	127.029239	37.56441
전철/지하철	208	18	왕십리역	서울특별시	성동구	행당1동	1104056	127.037088	37.561248
전철/지하철	209	19	한양대역	서울특별시	성동구	사근동	1104055	127.043672	37.555735
전철/지하철	210	20	독섬역	서울특별시	성동구	성수1가2동	1104066	127.047317	37.547254

표 3 버스-노선 정류장 데이터 (2017.05. 기준)

## 2. 데이터 전처리



트립체인 데이터의 노선ID, 차량ID, 하차역ID, 최종하차역ID를 버스-노선 정류장 데이터 및 지하철역 데이터와 연결하여 노선명, 정류장명과 그 위치를 얻는다.

### (1) 최종도착역 기준 설정

트립체인 데이터는 한 개인의 전체 대중교통 이용 로그를 보여주기 때문에 목적 통행을 위한 환승이 아닌 출발 지점으로 복귀하는 통행도 기록되어 있다. 본 연구에서는 지하철역에서 버스로 환승하여 목적지로 도달하고자 하는 목적 통행만을 분석 대상으로 한다. 또한 본 연구의 최종적인 목적이 환승 서비스 수준 개선이고, 대중교통 환승에 소요되는 평균 시간이 13.3분<sup>2</sup>임을 고려했을 때, 기준 환승시간 30분을 초과한 이후의 탑승객은 곧바로 환승했다기 보다, 개인적인 불일을 보고 재탑승한 것이기 때문에 환승 서비스 수준 개선과 거리가 멀다고 판단했다. 따라서 환승으로 간주하지 않고 해당 탑승 이전 지점을 최종도착역으로 보았다. 또한 본 연구의 분석 대상은 지하철을 간선으로, 버스를 지선으로 이용하는 승객들이기 때문에 최종목적지가 최종도착역으로부터 반경 3.3km(위도 0.03°) 이내에 위치한 승객 데이터만을 고려하였다.

### (2) 분석시간

본격적인 데이터 분석에 앞서, 연구진들의 예상이 실제로 잘 들어맞는지 확인해 보기 위하여 일반적인 등교시간인 07:00 ~ 09:00에 대한 데이터를 먼저 분석해 보았다. 이후 06시 ~ 24시까지의 트립체인 데이터에 대해 1시간 간격으로 모든 시간대에 대해 적용하여 서비스 수준을 평가한다.

### (3) 데이터 통합

<sup>2</sup> 국토교통부 대중교통현황조사(2015)

이 과정에서 데이터 간 관계도에 따라 최종하차역ID를 공통키로 하여 데이터 통합이 수행된다. 설명을 용이하게 하기위해, 서울대입구역 하차데이터를 예시로 들었다. 위 기준으로 07시 ~ 09시에 서울대입구역에 하차한 데이터 4707 rows가 추출되었다. 이 데이터는 다시 1) 서울대입구역이 환승 지점 2) 서울대입구역이 최종하차역 2가지로 나뉜다.

### 1) 서울대입구역 환승

총 1557rows이며, 하차역ID1이 0228(서울대입구역)인 경우, 하차역ID2가 서울대입구역에서 환승해서 도착한 정류장을 나타낸다. 최종하차역ID가 0228이 아님을 확인할 수 있다.

	카드번호	노선ID1	승차일시1	하차일시1	최종하차일시	승차역ID1	승차역ID2	하차역ID1	하차역ID2	최종하차역ID
5635784	//y3mMQNnAhM7	41221001	20170522074306	20170522074509	20170522085642	4100606	1854	4100611	228	60068
1993515	/2K9KvOruJLgxJGV	~	20170522081739	20170522083102	20170522085223	232	70641	228	8003149	8003149
5636247	/3IR+dEjy7CEcJF	11110024	20170522064620	20170522065725	20170522074919	8000297	418	73875	228	9035960
1994792	/3mnPmMGcs8l	~	20170522080752	20170522083244	20170522084343	326	70665	228	8502044	8502044
5637178	/9TNqMXTU0	11110194	20170522081717	20170522082422	20170522090316	11888	216	9514254	228	9014545
2024158	/aV2DumxFuY5Y	~	20170522080627	20170522084341	20170522085918	423	70665	228	60068	60068
5642054	/gnkflzcmJ	11110027	20170522071610	20170522071739	20170522082059	72011	418	72012	228	9003721
5638252	/HHaIE/G6D24Z	11110536	20170522071944	20170522073244	20170522074320	9010212	230	9012698	228	9103556
2031328	/iO1ze0uTMGJfq	~	20170522071945	20170522074830	20170522075701	217	70665	228	9003721	9003721
5642338	/iZBoGE/XDmpN	11110287	20170522063950	20170522064704	20170522074104	70771	2519	10111	228	9035960

표 4. 서울대입구역 환승 데이터

### 2) 서울대입구역 최종하차

총 3150rows이며, 최종하차역ID가 0228(서울대입구역)임을 확인할 수 있다.

	카드번호	노선ID1	승차일시1	하차일시1	최종하차일시	승차역ID1	승차역ID2	하차역ID1	하차역ID2	최종하차역ID
1919775	zZM/okJ0E~	~	20170522081553	20170522085757	20170522085757	1803	0	228	0	228
1919770	zZKdrq7Cipl~	~	20170522073354	20170522084400	20170522084400	3223	0	228	0	228
14929542	zZDamD0fzd8rSe~	~	20170522071730	20170522072241	20170522072241	230	0	228	0	228
5625048	zyXYyYtrGlp03L~M=	41050001	20170522055630	20170522063904	20170522070050	4130405	222	4113600	228	228
1931723	zyjNx7lkH4atC9LMJaB~	~	20170522065843	20170522081049	20170522081049	409	0	228	0	228
11446568	ZYeLnkEbGgU5h~	~	20170522065949	20170522074840	20170522074840	1815	0	228	0	228
4099519	ZY4kzVxLioI0BT~	41072002	20170522075634	20170522080543	20170522085425	4101030	2754	4100861	228	228
11443963	ZXNcdqkK2pNd~	~	20170522071911	20170522074903	20170522074903	217	0	228	0	228
5623812	zx64UQxZtYSd/ZtYUa~	41062101	20170522065622	20170522073001	20170522082043	4199580	2812	4103184	228	228
1930938	zx1vBBNsWC3KDO9~	~	20170522081450	20170522085040	20170522085040	249	0	228	0	228

표 5. 서울대입구역 최종하차 데이터 (상위 10항목)

본 연구에서는 지하철-버스 환승 문제 지점을 특정 짓고자 하므로 표4. 서울대입구역 환승 데이터를 사용한다. 이후 서울대입구역 지하철-버스 환승 트립체인 데이터의 최종하차역 ID를 데이터 간 관계도에 기반하여 최종 하차역명, 경도, 위도를 연결한 “최종하차역 데이터”를 얻어낸다. 즉, 표4. 서울대입구역 환승 데이터와 표2. 버스-노선 정류장 데이터를 최종하차역ID를 key로 하여 연결한 데이터이다. 따라서 중복값이 있을 수 있고, 각 행은 한 승객의 최종하차 정보를 나타낸다.

	최종하차역ID	최종하차역명	경도	위도	X	Y
1939407	7000164	상도3차삼성래미안후문	126.9539	37.50077	0.0012	0.019523
1944083	70761	서울산업정보학교.삼성교	126.9411	37.47097	-0.0116	-0.01028
1945234	60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103
1946317	9035960	관악경찰서	126.9526	37.47525	-0.00011	-0.006
1952647	9530	백산아파트1단지.호암사입구	126.9208	37.45279	-0.03184	-0.02847
1955732	8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135
1961103	60077	공동기원	126.9545	37.45513	0.001848	-0.02612
1974493	9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122
1974754	9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122
1975238	16343	삼성동주민센터	126.933	37.47001	-0.01966	-0.01125
1979709	14384	공대입구	126.9497	37.45535	-0.00303	-0.0259
1985519	70762	고시촌입구	126.9379	37.47024	-0.01483	-0.01101
1985940	70675	신소재	126.9502	37.45364	-0.00244	-0.02761
1986705	60028	신성초등학교	126.9374	37.47108	-0.01526	-0.01017
1987861	60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103
1993515	8003149	중앙대후문	126.9536	37.50451	0.000915	0.023255
1994792	8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135
2018144	70672	관악사삼거리	126.9567	37.46053	0.003975	-0.02072
2024158	60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103

표 6. 최종하차역 데이터 (상위 20항목)

#### (4) 공급량 계산(노선수)

관심 대상 지하철역을 지나치는 버스 노선만이 각 최종 하차역에 대한 환승 수요를 감당한다고 할 수 있다. 이에 관심 대상 지하철역(서울대입구역)을 중심으로 환승 공급의 정도를 노선수를 통해 계산하였다. 최종 하차역명을 보면 중복되는 이름이 있는데 노선수는 서로 다르다. 이는 7시~9시에 대한 시간 단위 분석을 진행하였을 때 대상 시간에 실제로 노선이 운행되는 것만 공급량이라고 판단하였기 때문이다. 즉, 같은 최종 하차역이라도 분석된 시간대가 다르면, 노선수가 달라질 수 있다.

최종하차역ID	최종하차역명	경도	위도	X	Y	노선수
9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122	11
9003721	서울여상.서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	10
9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122	10
9003721	서울여상.서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	10
9003721	서울여상.서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	9
9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122	7
9003558	관악푸르지오아파트	126.9582	37.48776	0.005485	0.006505	7
9003558	관악푸르지오아파트	126.9582	37.48776	0.005485	0.006505	6
70760	신림중.삼성고.관악문화관도서관	126.9443	37.47032	-0.00843	-0.01093	5
9003438	송실대입구역	126.9537	37.49651	0.001053	0.015262	5
8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135	5
9003438	송실대입구역	126.9537	37.49651	0.001053	0.015262	5

9034813	봉천고개, 관악현대아파트	126.9566	37.49178	0.003915	0.010533	5
70761	서울산업정보학교, 삼성교	126.9411	37.47097	-0.0116	-0.01028	4
8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135	4
70760	신림중, 삼성교, 관악문화관도서관	126.9443	37.47032	-0.00843	-0.01093	4
9003558	관악푸르지오아파트	126.9582	37.48776	0.005485	0.006505	4
70761	서울산업정보학교, 삼성교	126.9411	37.47097	-0.0116	-0.01028	4
60028	신성초등학교	126.9374	37.47108	-0.01526	-0.01017	3
70761	서울산업정보학교, 삼성교	126.9411	37.47097	-0.0116	-0.01028	3
60028	신성초등학교	126.9374	37.47108	-0.01526	-0.01017	3
9003447	서울대학교	126.9487	37.46683	-0.00396	-0.01442	3
9034813	봉천고개, 관악현대아파트	126.9566	37.49178	0.003915	0.010533	3
70760	신림중, 삼성교, 관악문화관도서관	126.9443	37.47032	-0.00843	-0.01093	3

표 7. 노선수 포함된 최종하차역 데이터 (상위 20항목)

### (5) 수요량 계산(인원수)

전처리 한 표 5. 서울대입구역 최종 하차데이터에 중복제거법을 사용하여 인원수를 계산한다. 인원수도 마찬가지로 분석되는 시간대에 따라 차이가 있다.

최종하차역ID	최종하차역명	경도	위도	X	Y	노선수	population
9003721	서울여상, 서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	9	215
9014545	서울대학교치과병원, 동물병원	126.9519	37.47003	-0.00075	-0.01122	11	157
9014545	서울대학교치과병원, 동물병원	126.9519	37.47003	-0.00075	-0.01122	10	144
70669	경영대	126.951	37.46586	-0.00167	-0.01539	1	107
60068	자연대, 행정관입구	126.9489	37.46022	-0.00382	-0.02103	2	80
60068	자연대, 행정관입구	126.9489	37.46022	-0.00382	-0.02103	2	79
9003721	서울여상, 서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	10	69
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	68
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	65
70669	경영대	126.951	37.46586	-0.00167	-0.01539	1	60
9003721	서울여상, 서울문영여중고앞	126.9528	37.473	0.00016	-0.00825	10	53
70638	서울대학교	126.948	37.46683	-0.0047	-0.01442	2	50
70670	수의대입구, 보건대학원앞	126.9547	37.46577	0.001987	-0.01548	1	44
70761	서울산업정보학교, 삼성교	126.9411	37.47097	-0.0116	-0.01028	3	38
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	35
60068	자연대, 행정관입구	126.9489	37.46022	-0.00382	-0.02103	1	32
70761	서울산업정보학교, 삼성교	126.9411	37.47097	-0.0116	-0.01028	4	31
8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135	4	30
60066	법대, 사회대입구	126.949	37.4628	-0.00369	-0.01845	1	29
9014545	서울대학교치과병원, 동물병원	126.9519	37.47003	-0.00075	-0.01122	7	28

표 8. 인원수 포함된 최종하차역 데이터 (상위 20항목)

## 3. K-means clustering을 이용한 군집화

k-평균 알고리즘(k-means clustering algorithm)은 주어진 데이터를 k개의 군집으로 묶는 알고리즘으로, 각 군집과 거리 차이의 분산을 최소화하는 방식으로 동작한다. parameter로서 최종하차역ID의 X좌표, Y좌표만을 사용한다.

Elbow Method은 k를 1부터 차례로 늘려 결과를 확인한다. k가 증가할수록 결과 지표가 좋아지지만 군집화 효과가 절감된다. 만약 k를 하나 늘렸을 때 이전보다 본 연구에서는 결과의 기준으로서 ‘distortion’을 선택하였다.

적용결과, 서울대입구역에 대해서는 최적의 군집수 k가 3으로 나타났다.

calinski_harabasz	score computes the ratio of dispersion between and within clusters
silhouette	score calculates the mean Silhouette Coefficient of all samples
distortion	the sum of square distances from each point to its assigned center

표 9. Elbow Method 결과 지표 기준

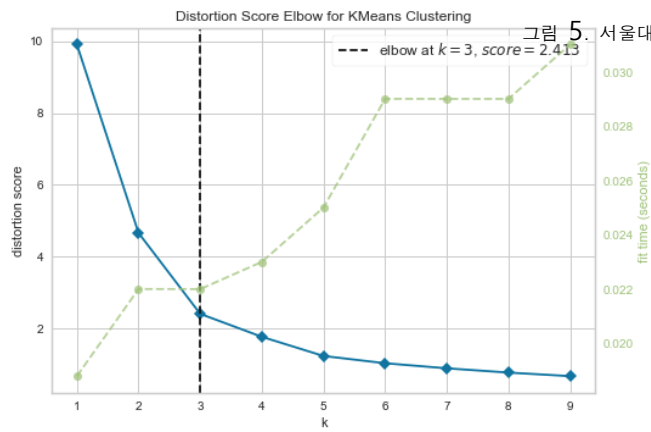


그림 5. 서울대입구역 Elbow Method 적용 결과

#### 4. 부하량 산정 모델

“최종하차역 데이터”를 바탕으로 좌표 상에 최종 하차지점을 표현할 수 있다. 기존 지하철역을 중심으로 하여 최종 하차지점의 빈도가 고르게 분포하지 않고 쏠리게 나타난다면 해당 방향으로 교통 수요가 증가하며 서비스 수준은 저하된다. 이에 기존 지하철역을 중심으로 산란된 하차 지점에 대해 군집화를 진행하여 해당 방향의 서비스 수준을 부하량 기준으로 예측하는 모델을 제안한다.

##### (1) 하차정류장 부하량 계산식

$$Load_{[ijkl]} = \frac{Demand_{[ijkl]}}{cap_{[acity_{[ijkl]}]}}$$

$Load_{[ijkl]}$  : 지하철역  $i$ 의  $j$ 시간대에서  $k$ 번째 군집의  $l$ 번째 하차정류장 부하량

$Demand_{[ijkl]}$  : 지하철역  $i$ 의  $j$ 시간대에서  $k$ 번째 군집의  $l$ 번째 하차정류장 수요량

$cap_{[acity_{[ijkl]}]}$  : 지하철역  $i$ 의  $j$ 시간대에서  $k$ 번째 군집의  $l$ 번째 하차정류장 공급량

이때,  $Demand_{[ijkl]}$  은 하차 인원수이며  $cap_{[acity_{[ijkl]}]}$  은 기존 지하철역을 경유하여 하차정류장에 도달하는 버스 노선수이다.

앞서 데이터 전처리에서 구한 공급량, 수요량 계산 결과를 통해 모든 최종 하차정류장 각 각에 대해 새로운 부하량 Field를 만들었다. 하차정류장 부하량은 해당 정류장의 교통 부담량을 의미하며 공급 대비 교통 수요에 비례한다.

##### (2) 군집 부하량 계산식

$$Weight_{[ijk]} = \frac{\sum_{l \in Stop_{[ik]}} Load_{[ijkl]} \times Demand_{[ijk]}}{\sum_{k \in T_{[ij]}} Demand_{[ijk]}}$$

$Weight_{[ijk]}$  : 지하철역  $i$ 의  $j$ 시간대의  $k$ 번째 군집의 군집부하량

$Stop_{[ik]}$  : 지하철역  $i$ 의  $k$ 번째 군집에 포함된 전체 하차정류장

$Load_{[ijkl]}$  : 지하철역  $i$ 의  $j$ 시간대의  $k$ 번째 군집의  $l$ 번째 하차정류장 부하량

$T_{[ij]}$  : 지하철역  $i$ 의  $j$ 시간대 전체 군집

$Demand_{[ijk]}$  : 지하철역  $i$ 의  $j$ 시간대의  $k$ 번째 하차정류장 수요량

특정 방향에 대한 쏠림의 정도를 가중치로 하여 쏠림의 영향이 결과에 반영되도록 군집 부하량을 계산하였다. Load는 위에서 계산한 하차 정류장 부하량이고, 여기에 전체 군집에 대한 해당 군집의 수요 부담률을 곱함으로써 쏠림의 영향을 반영하였다.

##### (3) 지하철 부하량 계산식

$$Subway_{[i]} = \max \{ Y_{[ij]} \mid j = 6, 7, 8, \dots, 24 \}$$

$Y_{[ij]}$  : 지하철역  $i$ 의  $j$ 시간대의 부하량

지하철 간의 환승 서비스 수준을 비교하기 위해 지하철 부하량의 기준을 정하였다. 단순히 전체 부하량의 평균값으로 비교하면 특정 시간대의 쏠림의 영향이 평균에 가려지는 결과가 나타날 수 있다. 따라서 쏠림이 가장 심했던 시간대의 부하량이 해당 지하철 부하량을 대표하도록 한다. 해당 내용을 위의 식으로 나타내었다. 한편, 쏠림이 가장 심했던 시간대를 정하는 기준은 아래의 식으로 하였다.

$$Y_{[ij]} = \max \{ Weight_{[ijk]} \mid k \in T_{[ij]} \}$$

$Y_{[ij]}$  : 지하철역  $i$ 의  $j$ 시간대의 부하량

$Weight_{[ijk]}$  : 지하철역  $i$ 의  $j$ 시간대의  $k$ 번 군집의 부하량

$T_{[ij]}$  : 지하철역  $i$ 의  $j$ 시간대의 전체 군집

$j$  시간대의 지하철 부하량은  $j$  시간대의 모든 방향의 군집 부하량 중 가장 큰 부하량으로 하였다. 이 역시 쏠림의 영향이 평균에 가려지지 않도록 하기 위함이다.

## 5. 부하량 모델 적용

### (1) 모델 적용

최종하차역ID	최종하차역명	경도	위도	X	Y	노선수	population	cluster	부하량
70669	경영대	126.951	37.46586	-0.00167	-0.01539	1	107	2	78.50481
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	68	2	49.89091
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	65	2	45.14741
70669	경영대	126.951	37.46586	-0.00167	-0.01539	1	60	2	41.67453
70670	수의대입구.보건대학원앞	126.9547	37.46577	0.001987	-0.01548	1	44	2	32.28235
60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103	2	80	2	29.34759
60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103	2	79	2	27.43573
71978	농생대	126.9492	37.45716	-0.00352	-0.02409	1	35	2	23.58075
60068	자연대.행정관입구	126.9489	37.46022	-0.00382	-0.02103	1	32	2	21.55954
60066	법대.사회대입구	126.949	37.4628	-0.00369	-0.01845	1	29	2	21.27701
70638	서울대학교	126.948	37.46683	-0.0047	-0.01442	2	50	2	18.34225
70669	경영대	126.951	37.46586	-0.00167	-0.01539	1	25	2	16.84339
9003721	서울여상.서울문명여중고앞	126.9528	37.473	0.00016	-0.00825	9	215	2	16.0948
70670	수의대입구.보건대학원앞	126.9547	37.46577	0.001987	-0.01548	1	20	2	13.89151
14384	공대입구	126.9497	37.45535	-0.00303	-0.0259	1	18	2	13.20642
70671	국제대학원	126.9549	37.46371	0.002242	-0.01754	1	17	2	12.47273
70670	수의대입구.보건대학원앞	126.9547	37.46577	0.001987	-0.01548	1	18	2	12.12724
60066	법대.사회대입구	126.949	37.4628	-0.00369	-0.01845	1	16	2	11.11321
9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122	10	144	2	10.56513
9014545	서울대학교치과병원.동물병원	126.9519	37.47003	-0.00075	-0.01122	11	157	2	9.913486
14384	공대입구	126.9497	37.45535	-0.00303	-0.0259	1	12	2	8.334906
70672	관악사삼거리	126.9567	37.46053	0.003975	-0.02072	2	21	2	7.703743
70671	국제대학원	126.9549	37.46371	0.002242	-0.01754	1	11	2	7.64033
60066	법대.사회대입구	126.949	37.4628	-0.00369	-0.01845	1	11	2	7.411093
70638	서울대학교	126.948	37.46683	-0.0047	-0.01442	2	16	2	5.556604
8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135	4	30	2	5.209316
60065	서울대정문	126.9484	37.46573	-0.00429	-0.01552	1	7	2	4.862028
60077	공동기원	126.9545	37.45513	0.001848	-0.02612	1	7	2	4.862028
9003721	서울여상.서울문명여중고앞	126.9528	37.473	0.00016	-0.00825	10	69	2	4.792571
14384	공대입구	126.9497	37.45535	-0.00303	-0.0259	1	7	2	4.71615
70674	유전공학연구소.반도체공동연구소	126.9536	37.45375	0.000868	-0.0275	1	7	2	4.71615
70762	고시촌입구	126.9379	37.47024	-0.01483	-0.01101	1	20	1	4.5677
70674	유전공학연구소.반도체공동연구소	126.9536	37.45375	0.000868	-0.0275	1	6	2	4.402139
70762	고시촌입구	126.9379	37.47024	-0.01483	-0.01101	1	26	0	4.087701
60077	공동기원	126.9545	37.45513	0.001848	-0.02612	1	6	2	4.042414
8502044	서울대학교	126.9497	37.46775	-0.003	-0.0135	5	27	2	3.961925
9003721	서울여상.서울문명여중고앞	126.9528	37.473	0.00016	-0.00825	10	53	2	3.888556
70762	고시촌입구	126.9379	37.47024	-0.01483	-0.01101	1	23	0	3.742925
60065	서울대정문	126.9484	37.46573	-0.00429	-0.01552	1	5	2	3.668449
70672	관악사삼거리	126.9567	37.46053	0.003975	-0.02072	2	10	2	3.472877

표 10. 부하량 포함된 최종하차역 데이터 (상위 40항목)



데이터 전처리를 거친 후 얻은 노선수와 population 필드의 값들을 앞서 제시한 모델에 적용함으로써 최종하차정류장에 대한 부하량을 얻을 수 있었다. 이 부하량은 군집 비율을 가중하기 전의 부하량이며, 하차 정류장의 교통 부담량을 점 단위로 정량적으로 나타내었다. 이 값을 군집별로 합하고 군집 비율을 가중하여 각 방향에 대한 군집 부하량을 구하여 비교한다.

지하철역	cluster	부하량	실제 방면
서울대입구역	0	35.17681	대학동 방면
서울대입구역	1	43.1057	행운동 방면
서울대입구역	2	628.3697	서울대학교 방면

표 11. 서울대입구역 군집별 부하량 (07시 - 09시)

부하량 모델 적용 결과, 서울대입구역 기준으로 서울대학교 방면(Cluster 2)으로 부하량이 628.370로 군집 0의 35.177, 군집 1의 43.106에 비해 각각 17.9배, 14.6배로 편향되어 나타났다.

### (2) 시각화

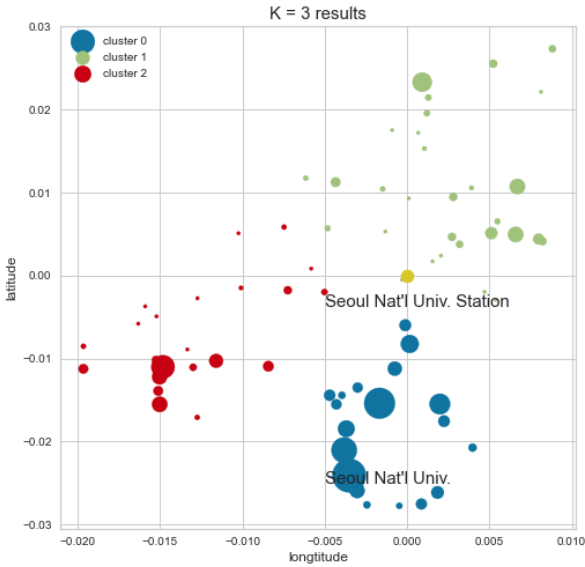


그림 6. 서울대입구역 기준 군집화 결과

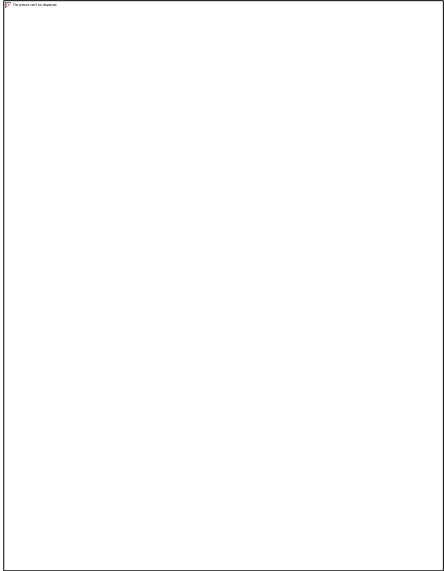


그림 7. 군집 부하량 (07~09시)

서울대입구역을 기준으로 각 하차역에 대해 군집화를 수행하고, 하차역 별 부하량의 크기를 원의 크기에 비례하여 시각화하였다. 시각화 결과 07시 - 09시에 서울대입구역에서 서울대학교 방면으로 인원의 쏠림이 크게 발생했음을 직관적으로 확인할 수 있었다. 특히 최종 목적지점을 지도에 나타내보면 어느 요인에 의해 인원이 해당 시간에 쏠린 것인지 파악할 수 있었는데, 서울대입구역의 경우에 서울대학교로 통학하는 목적 통행이 주된 원인이라고 할 수 있었다. 이제 모든 지하철역에 대해 방법론을 적용해 봄으로써 유인을 찾아보도록

하겠다.

6. 환승 문제지점 예측

부하량 산정 모델을 적용한 결과 중 상위 50개 역사의 최대 쏠림시간, 최대 부하량, 최대 부하 군집은 다음과 같다. 전체 결과에 대해서는 Appendix에 첨부하였다.

Rank	역명	최대쏠림시간	최대부하량	최대부하군집	Rank	역명	최대쏠림시간	최대부하량	최대부하군집
1	디지털미디어시티역	8	739.73	0	26	미아사거리역	19	270.28	0
2	청담역	8	699.40	1	27	연신내역	19	240.91	1
3	회기역	8	671.91	0	28	건대입구역	19	231.56	1
4	논현역	8	585.36	0	29	종합운동장역	18	223.24	1
5	길음역	18	567.81	0	30	신대방삼거리역	8	223.08	0
6	신림역	19	560.23	0	31	내방역	8	219.87	0
7	낙성대역	9	544.90	1	32	합정역	19	206.43	0
8	쌍문역	18	501.27	0	33	성신여대입구역	19	203.20	1
9	수유역	19	463.79	0	34	대방역	7	199.96	0
10	상일동역	19	457.68	1	35	신정네거리역	19	198.63	0
11	양재역	8	452.30	0	36	홍제역	19	192.65	1
12	구로디지털단지역	17	440.27	0	37	대림역	19	185.97	0
13	가산디지털단지역	8	438.72	1	38	독산역	19	185.58	0
14	북정역	19	390.78	0	39	문래역	8	184.76	1
15	개봉역	18	335.10	1	40	장지역	19	181.19	0
16	봉천역	18	333.63	0	41	신도림역	18	181.14	2
17	수서역	19	325.77	0	42	서울역	8	176.39	0
18	노량진역	18	320.26	2	43	노원역	19	175.78	1
19	당산역	22	316.44	0	44	삼성역	8	175.32	1
20	장한평역	19	305.03	1	45	여의나루역	8	174.98	0
21	신대방역	22	302.43	0	46	신논현역	7	168.85	2
22	온수역	7	297.95	1	47	석계역	18	168.59	0
23	서울대입구역	9	296.97	2	48	홍대입구역	7	164.44	1

24	매봉역	8	277.19	0	49	광흥창역	8	164.14	1
25	남부터미널역	8	274.02	0	50	상봉역	19	163.84	0

표 13. 환승 문제 발생 예측 지점(상위 50개)

부하량이 600~800 인 역사는 3개, 400~600인 역사는 10개, 200~300인 역사는 20개로, 부하량의 값이 작아질수록 속하는 역사의 수가 가파르게 증가하는 것을 볼 수 있다. 나아가 수도권 285개 역사들의 부하량 분포 추이를 알아보기 위해서 히스토그램으로 도시한 결과는 다음과 같다.

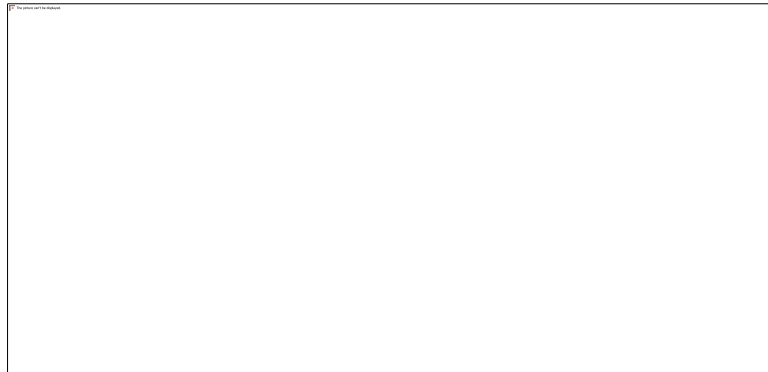


그림 8. 전체 지하철 역사 부하량 히스토그램

앞서 표를 분석했던 결과와 마찬가지로 부하량이 값이 커질수록 역사의 수가 기하급수적으로 줄어드는 것을 확인할 수 있다. 문제가 되는 부하량의 기준을 찾기 위해 간격을 10으로 줄여 더 세밀하게 나누고 누적 히스토그램을 그린 후, Knee Point Detection 알고리즘을 적용한 결과, 아래와 그림과 같이 부하량 200 근방에서 Knee Point가 나타남을 알 수 있었다.

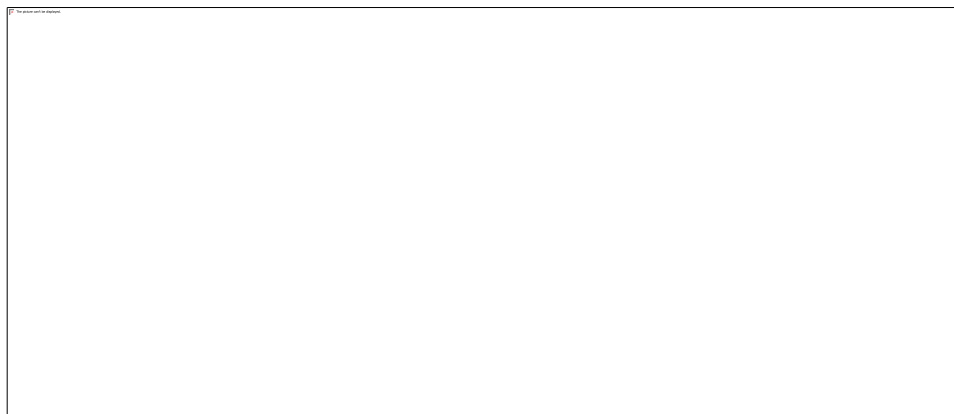


그림 9. 부하량 분포와 knee point

이후 본 연구가 실제로 환승이 문제되는 지점을 예측하는데 실효성이 있는지 판단하기 위해 앞서 구한 부하량 상위 6개 역사 중 LOS가 A인 역사(디지털미디어시티, 청담, 회기, 길음, 신림, 낙성대)에 대해 대략적으로 분석해보고, 비교적 설문조사가 용이한 3개 역사(회기, 신림, 낙성대)에 대하여 설문조사를 통한 사례분석을 진행하였다.

## 7. 선행연구와의 비교

앞서 구한 부하량 상위 50개 역사에 대하여 제3차 환승센터 및 복합환승센터 구축 기본 계획(국토교통부 대도시권광역교통위원회, 2021)에 실린 가중평균 환산거리 기반 주요 환승 지점의 환승서비스 수준과 Proceedings of the KSR Conference(한국철도학회:학술대회논문집, 2011)에 실린 역사별 연계 환승시간 서비스평가(최명훈 외 3인)를 비교한 결과는 다음과 같다

Rank	역명	환승거리 LOS	환승시간 LOS	Rank	역명	환승거리 LOS	환승시간 LOS
1	디지털미디어시티역	D	A	26	미아사거리역	-	A
2	청담역	-	A	27	연신내역	C	B
3	회기역	D	A	28	건대입구역	D	B
4	논현역	-	B	29	종합운동장역	C	A
5	길음역	-	A	30	신대방삼거리역	-	A
6	신림역	-	A	31	내방역	-	A
7	낙성대역	-	A	32	합정역	D	B
8	쌍문역	-	A	33	성신여대입구역	-	B
9	수유역	-	A	34	대방역	-	A
10	상일동역	-	A	35	신정네거리역	-	B
11	양재역	D	A	36	홍제역	-	B
12	구로디지털단지역	-	A	37	대림역	-	A
13	가산디지털단지역	-	A	38	독산역	-	A
14	북정역	E	B	39	문래역	-	B
15	개봉역	-	A	40	장지역	-	B
16	봉천역	-	A	41	신도림역	-	A
17	수서역	C	A	42	서울역	D	B
18	노량진역	-	A	43	노원역	-	A
19	당산역	C	A	44	삼성역	D	A
20	장한평역	-	A	45	여의나루역	-	A
21	신대방역	-	A	46	신논현역	-	B
22	온수역	-	A	47	석계역	C	A
23	서울대입구역	-	B	48	홍대입구역	C	A
24	매봉역	-	A	49	광흥창역	-	A
25	남부터미널역	-	A	50	상봉역	F	B

표 14. 부하량 상위 50개 역사 LOS 비교

환승거리나 환승시간에 기반한 서비스 수준의 경우 대부분의 역사별로 비슷한 정도로 측정되었으나, 부하량에 기반한 문제지점 예측 기법의 경우 역사별 환승 서비스 수준의 차이가 극명하게 드러날 뿐만 아니라 수요가 쏠려서 문제가 되는 방향까지도 알아낼 수 있다는 점에서 일반적인 비교는 어려운 것을 볼 수 있었다.

### III. 상위 6개 역사에 대한 분석

#### 1. 분석방법

##### (1) Kernel Density

Kernel Density Estimation(KDE)은 non-parametric 밀도 추정 방법 중 하나로서 커널 함수와 데이터를 바탕으로 연속성 있는 확률 밀도 함수를 추정하는 것이다.

이때 커널함수란 다음의 세 가지 조건을 모두 만족하는 함수를 의미한다.

$$\int_{-\infty}^{\infty} K(u) du = 1$$

$$K(u) \geq 0 \text{ for all values of } u$$

$$\text{Non-negative}$$

KDE에서는 다음의 수식을 이용하여 밀도를 추정한다.

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

$$= \frac{1}{nh} K\left(\frac{x - x_i}{h}\right)$$

$$x_i = \text{관측된 데이터 포인트}$$

$$h = \text{bandwidth parameter}$$

본 연구진은 이처럼 각 데이터의 포인트를 중심으로 하는 연속적인 커널함수를 생성하여 밀도를 추정하는 KDE의 특징이 환승 인원의 최종 목적지 예측에 유의미 할 것이라고 판단 하였고, 앞서 구한 부하량에 근거한 KDE를 통해 환승 인원들의 최종 목적지가 어디인지 예측해 봄으로써 쏠림 현상의 대략적인 유인을 파악해 보고자 하였다.

본 연구에서는 Q-gis에서 제공하는 Heat map(Kernel Density Estimation)을 분석에 이용 하였으며, 사용된 parameter는 다음과 같다.

Point layer	역사 별 최대 쏠림 시간 데이터
Radius	0.01
Pixel size X	0.0005
Pixel size Y	0.0005
Weight from field	부하량
Kernel shape	Epanechnikov

표 15. Heat map parameters

##### (2) Getis Ord Gi\*

공간적 자기 상관이란 토블러(W.Tobler)의 지리학 제 1법칙인 “모든 것은 다른 모든 것들과 관련이 되어 있으며 특히 가까운 것은 먼 것보다 더 많이 관련되어 있다.”에 근거, 어느 특정한 변수가 공간적으로 상관 관계를 맺고 있는가를 말해주는 지표이다. 본 연구진은 이러한 공간적 자기 상관을 파악하기 위한 공간통계 기법 중 하나인 Getis Ord Gi\*를 이용한 Hot spot Analysis를 통해 쏠림 현상의 유인을 보다 명확하게 파악하고자 하였다.

Getis-Ord Gi\*에서 사용하는 수식은 다음과 같다.

$$G_{-}[i]^{[*]} = \frac{(\sum_{j=1}^n w_{-}[i,j] x_{-}[j] - \bar{[X]} \sum_{j=1}^n w_{-}[i,j])}{S \sqrt{(\sum_{j=1}^n w_{-}[i,j]^2 - (\sum_{j=1}^n w_{-}[i,j])^2 / (n-1))}}$$

$$\bar{[X]} = (\sum_{j=1}^n x_{-}[j]) / [n]$$

$$S = \sqrt{(\sum_{j=1}^n x_{-}[j]^2 / [n] - (\bar{[X]})^2)}$$

$x_{-}[j]$  = 관측값  $j$ 의 속성 값

$w_{-}[i,j]$  = 관측값  $i$ 와  $j$  사이의공간 가중치

$n$  = 총 관측값의 개수

Hot Spot Analysis를 통해 각 관측값에 대해 계산되는 Getis Ord  $G_i^*$  값은 Z-score와 P-value이고, Z 값이 클수록 강한 상관관계를 나타내며, Hot spot으로 볼 수 있고, Z값이 낮을수록 공간적 상관관계가 없는 Cold spot으로 볼 수 있다. P-value는 통계값의 유의수준을 의미한다.

본 연구에서는 Hot Spot Analysis를 진행하기 전 어떠한 정류장을 이웃으로 보고 분석을 진행할지 결정하기 위해 Arc map에서 제공하는 Calculate Distance Band from Neighbor Count tool을 이용하여 각 정류장에서 1개의 이웃 정류장까지의 거리(Distance Band)의 최대값과 평균값을 계산하였다. 이후 그림에서 보는 바와 같이 Incremental Spatial Autocorrelation tool을 이용하여 공간적 자기상관성이 최대가 될 때의 이웃 정류장으로 판단 가능한 거리 값을 계산하였다. 이때 모든 정류장을 통계에 포함시킬 수 있어야 하기 때문에 앞

서 구한 최대값을 이웃 정류장으로 판단하기 위한 거리의 초기값으로 두고 평균값만큼 거리를 늘려나가면서 공간적 자기상관성을 계산하도록 하였다. 최종적으로 구해진 거리값을 이웃 정류장으로 판단 가능한 거리로 두고, 정류장 별 부하량을 가중치로 설정하여 Hot Spot Analysis를 진행하였다.



그림 10. 디지털 미디어 시티(전체)  
Spatial Autocorrelation

## 2. 디지털미디어시티(8시~9시)

### (1) 클러스터링 결과 및 부하량 모델 적용

- cluster0
- cluster1
- cluster2
- ◆ 디지털미디어시티역
- 클러스터별 부하량

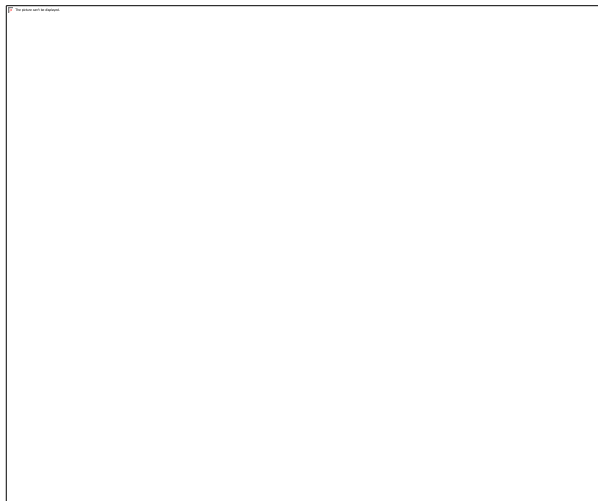


그림 16. 디지털미디어시티역 클러스터링 결과

디지털미디어시티역에 대한 클러스터링 결과 cluster0은 북서쪽의 수색동 ~ 상암동 방면, cluster 1은 동남쪽의 남가좌1동 ~성산2동 방면, cluster 2는 북동쪽의 북가좌동 ~ 증산동 방면으로 구분되었다. 각각의 클러스터에 대해 부하량 모델을 적용한 결과, cluster 0은 739.73, cluster 1은 0.65, cluster 2는 1.79로 각각 1, 138배, 413.2배의 큰 차이가 나타났고, 디지털미디어시티역에서의 환승 문제는 주로 수색동 ~ 상암동 방면으로의 환승에서 기인할 것이라는 것을 확인할 있다.

ID	부하량
cluster0	739.73
cluster1	0.65
cluster2	1.79

표 17. 디지털미디어시티역 부하량 모델 적용 결과

### (2) Kernel Density



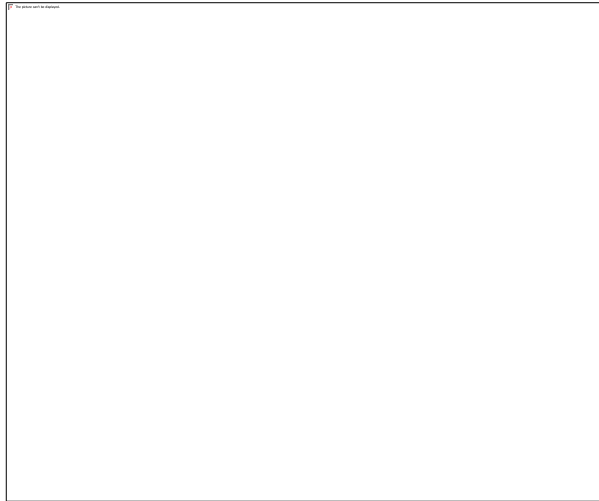


그림 17. 디지털미디어시티역 Kernel Density 적용결과

디지털미디어시티역의 어떠한 유인이 이렇게 편향된 환승 행태를 야기하는지 확인하기 위해 Kernel Density를 적용한 결과 JTBC빌딩, 중앙일보빌딩, KBS미디어센터등의 방송국 시설, 상암IT타워, LG유플러스 상암사옥, 삼삼 SDS 상암 센터등의 대규모 기술 회사의 사옥, 서울 일본인 학교, 서울드와이트 외국인학교 등의 대규모 교육 시설, 상암월드컵파크 2~12단지 아파트, 상암DMC 푸르지오시티, DMC 롯데캐슬 더퍼스트 등의 주거 시설이 환승 수요에 기여했을 것으로 파악되었다. 특히 08~09시라는 시간대를 감안했을 때 출근을 위해 방송국 사옥, 기술 회사의 사옥이 집중적으로 분포되어 있기때문에 대규모 교통 수요가 편향되어 나타났을 것으로 예상된다.

### (3) 공간통계분석(Hotspot Analysis)

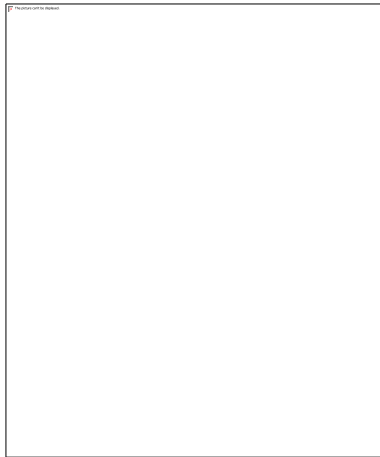


그림 18. 디지털미디어시티역 전체 대상 Hotspot 분석 결과

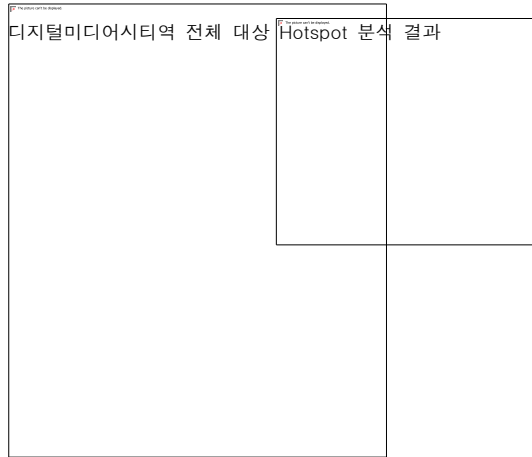


그림 19. 디지털미디어시티역 Hotspot 분석 결과

calculate distance band from neighbor count			
디지털미디어시티(전체)		디지털미디어시티(Cluster0)	
average distance	229	average distance	163
maximum distance	916	maximum distance	916
incremental spatial autocorrelation			
디지털미디어시티(전체)		디지털미디어시티(Cluster0)	
1374		1046	

표 18. 디지털미디어시티역 spatial autocorrelation 계산 결과

디지털미디어시티역 전체 정류장에 대한 핫스팟 분석 결과, 앞서 확인했던 바와 같이 북서쪽의 수색~상암동 방향으로 핫스팟이 형성됨을 알 수 있다. 그 중에서 더 정밀한 유인을 찾기 위해 cluster0에 대해 추가적으로 핫스팟 분석을 시행하였고, 그 결과 99% 신뢰수준에서 핫스팟인 정류장은 수색동 주민센터, 95% 신뢰수준에서 수색역 앞, MBC, 상암DMC 홍보관.YTN, 월드컵파크4단지, 월드컵파크4단지후문.JTBC, 월드컵파크2단지.에스플렉스센터, 월드컵파크3단지정문, 월드컵파크3단지.난지천공원으로 나타났다.

(4) 분석 결과

디지털미디어시티역은 오전 8~9시가 첨두시간인 것으로 조사되었으며, 역사 북서쪽의 수색~상암동 방면으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 보았을 때, 대규모 주거지역과 밀집된 사옥 단지가 형성되어 있음을 알 수 있었다. 정류장 단위의 분석을 해보았을 때,

수색역앞의 경우 대규모 주거단지 DMC 롯데캐슬 더퍼스트가 있고, 수색동 주민센터와 수색 감리교회 및 주변 상업지구가 있었는데, 출근 시간임을 고려했을 때, 수색동 주민센터 및 교회, 상업 지구로의 출근을 위한 통행 수요가 크게 쏠리는 것을 예상할 수 있었다.

MBC, 상암DMC홍보관.YTN의 경우 MBC, YTN, CJ등의 대규모 사옥이 밀집하여 위치하였다. 출근 시간임을 고려했을 때, 해당 시설에 대한 출근 수요가 매우 클 것으로 예상할 수 있었다.





월드컵파크4단지의 경우 주거 단지 앞에 위치해있었으나, 등교 시간임을 고려했을 때, 상암 중학교에 등교하려는 학생들의 수요가 집중되었을 것으로 예상할 수 있다.

월드컵파크 4단지 후문.JTBC, 월드컵파크2단지.에스플렉스센터의 경우 중앙일보, JTBC 대규모 사옥, 에스플렉스 등의 사옥 단지에 대한 출근 수요가 매우 클 것으로 예상된다.

종합해보면, 디지털미디어시티역에서의 환승 불편이 가장 크게 야기되는 유인으로 예상되는 것은 중앙일보 ~ MBC에 걸쳐서 집중 분포되어 있는 방송국 대규모 사옥이며, 이에 학교 등교 수요까지 겹쳐져 수요가 중첩적으로 집중되었다. 정류장 단위로 출근 수요 분담량을 알 수 있기 때문에 이를 바탕으로 정류장 및 노선을 신설할 수 있다.

3. 청담(8시~9시)

(1) 클러스터링 결과 및 부하량 모델 적용

-  cluster0
-  cluster1
-  청담역
-  클러스터별 부하량

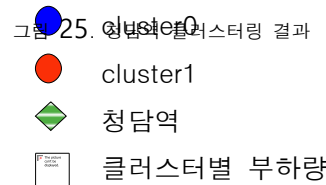
ID	부하량
cluster0	6.72
cluster1	699.40

표 20. 청담역 부하량 모델 적용 결과

종합해보면, 디지털미디어시티역에서의 환승 불편이 가장 크게 야기되는 유인으로 예상되는 것은 중앙일보 ~ MBC에 걸쳐서 집중 분포되어 있는 방송국 대규모 사옥이며, 이에 학교 등교 수요까지 겹쳐져 수요가 중첩적으로 집중되었다. 정류장 단위로 출근 수요 분담량을 알 수 있기 때문에 이를 바탕으로 정류장 및 노선을 신설할 수 있다.

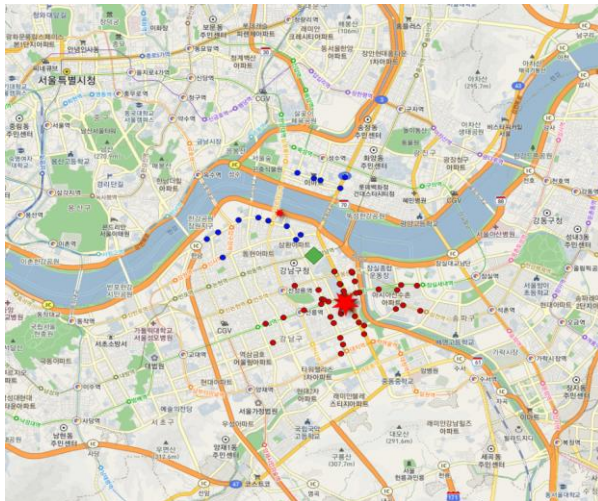
### 3. 청담(8시~9시)

#### (1) 클러스터링 결과 및 부하량 모델 적용



ID	부하량
cluster0	6.72
cluster1	699.40

표 20. 청담역 부하량 모델 적용 결과



청담역에 대한 클러스터링 결과 cluster0은 북쪽의 청담동 방면, cluster1은 남쪽의 삼성동 방면으로 구분되었다. 각각의 클러스터에 대한 부하량 모델을 적용한 결과 cluster0은 6.72, cluster1은 699.40로 104배 정도의 큰 차이가 나타난 것으로 보아, 청담역에서의 환승 문제는 주로 삼성동 방면으로의 환승에서 기인한다.

#### (2) Kernel Density

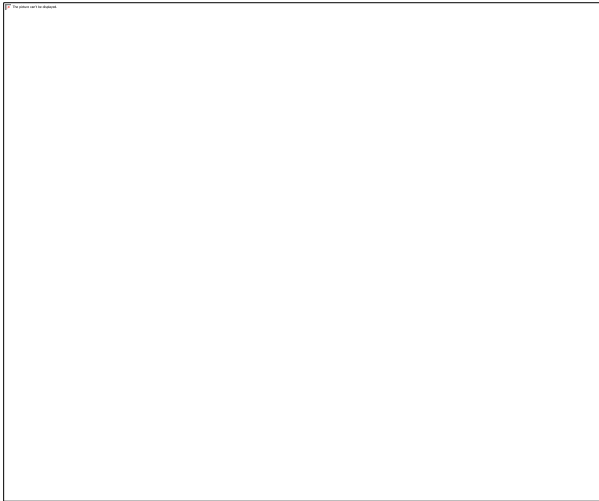


그림 26. 청담역 Kernel Density 적용결과

청담역의 편향된 환승 해태의 유인을 확인하고자 Kernel Density를 적용한 결과 코엑스, 포스코 센터 인근 상업단지, 경기고등학교 등의 유인이 환승수요에 크게 기여한 것으로 나타났다. 특히, 8~9시라는 시간대를 감안했을 때, 코엑스에 출근하는 직원들이나 인근 가게들의 자영업자들, 등교하는 학생들이 삼성동 방향으로 집중되기 때문에 편향된 환승행태가 나타나는 것으로 예상된다.

### (3) 공간통계분석(Hotspot Analysis)

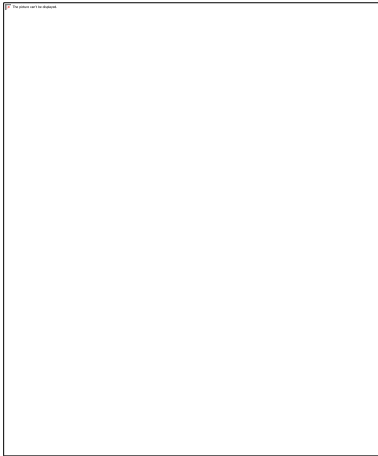


그림 27. 청담역 전체 대상 Hotspot 분석 결과

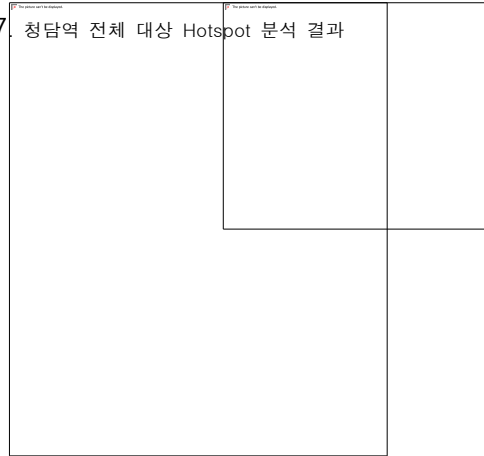


그림 28. 청담역 쏠림 군

calculate distance band from neighbor count			
청담(전체)		청담(Cluster1)	
average distance	227	average distance	190
maximum distance	670	maximum distance	648
incremental spatial autocorrelation			
청담(전체)		청담(Cluster1)	
1351		1028	

표 21. 청담역 spatial autocorrelation 계산 결과

청담역 전체 정류장에 대한 핫스팟 분석 결과 앞서 확인했던 바와 같이 남쪽의 삼성동 방

향으로 핫스팟이 형성됨을 알 수 있다. 그 중에서 더 정밀한 유인을 찾기 위해 cluster1에 대해 핫스팟 분석을 시행하였고, 그 결과 99% 신뢰수준에서 핫스팟인 정류장은 봉은사·삼성1파출소앞, 무역센터, 삼성역, 포스코사거리, 강남경찰서·강남운전면허시험장, 삼성역, 포스코사거리, 잠실종합운동운동장으로 나타났다.

#### (4) 분석 결과

청담역은 오전 8~9시가 첨두시간인 것으로 조사되었으며, 역사 남동쪽의 삼성동 방면으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 봤을 때 종합무역센터와 상가들이 늘어난 골목들이 곳곳에 위치해있었고, 이를 통해 삼성동 방면으로의 환승편향은 근처 상가에서 장사를 하는 자영업자들이나 코엑스로 출근하는 직원들의 출근에서 기인했음을 알 수 있다. 그러나 본 연구에서 2017년 05월 22일의 트립체인 데이터를 사용하였고, 2020년 9월 12일에 수인분당선이 개업하면서 문제가 예측되는 지점 근처에 수인분당선과 연결되는 역사들(강남구청역, 선정릉역 등)으로 인해 최근의 트립체인 데이터를 사용하여 다시 분석한다면 편향된 환승수요에 변화가 생겼을 수 있다.

### 4. 회기(8시~9시)

#### (1) 클러스터링 결과 및 부하량 모델 적용

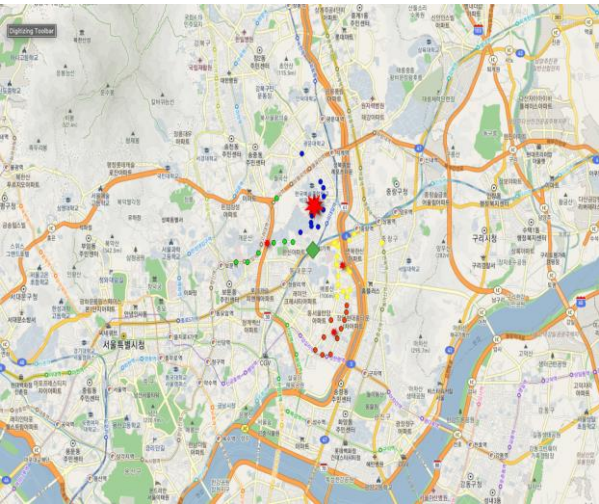


그림 36. 회기역 클러스터링 결과

회기역에 대한 클러스터링 결과, cluster0은 북이문동 방면, cluster1은 남쪽의 장안동 방면, cluster2은 회기동~청량리동 방면, cluster3는 남동쪽의 동 방면으로 구분되었다.	cluster0	671.91	쪽의 ster2 회경
	cluster1	4.18	
	cluster2	0.36	
	cluster3	37.76	

각각의 클러스터에 대해 부하량 모델을 적용한 결과 가장 큰 클러스터인 cluster0은 671.91, 가장 큰 클러스터인 cluster1은 0.36로 1800배 정도의 큰 차이가 나타났고, 회기역에서의 환승문제는 주로 이문동 방면으로의 환승에서 기인할 것이라는 것을 확인할 수 있다.

(2) Kernel Density

표 23. 회기역 부하량 모델 적용 결과

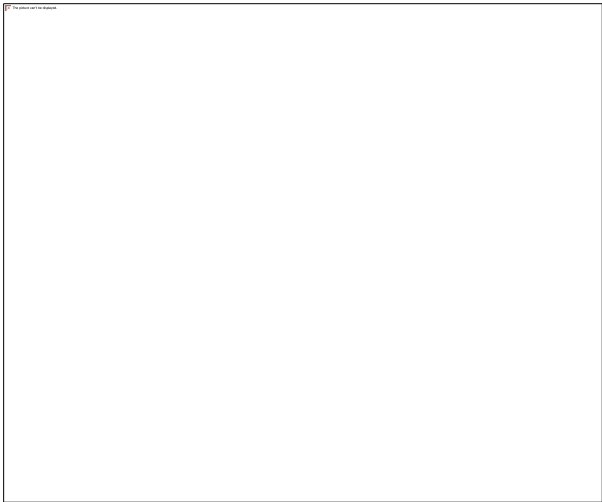


그림 37. 회기역 Kernel Density 적용결과

회기역의 편향된 환승 행태를 야기하는 유인을 확인하기 위해 Kernel Density를 적용한 결과 위와 같은 결과가 나타났으며, 특히 8~9시라는 시간대를 감안했을 때 KAIST 서울캠퍼스, 경희대학교 서울캠퍼스, 한국외국어대학교 서울캠퍼스, 한국예술종합학교 석관동캠퍼스 등의 대학교육시설이 이문동 방향에 집중되어 있기 때문에 편향된 환승 행태가 나타나는 것으로 예상된다.

(3) 공간통계분석(Hotspot Analysis)

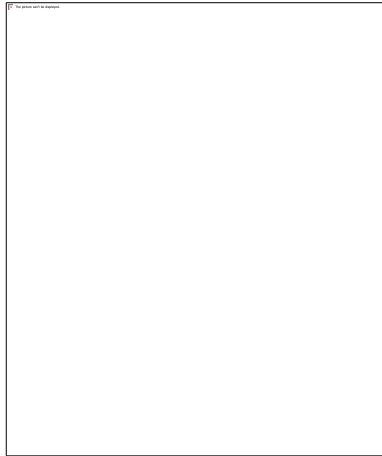


그림 38. 회기역 전체 대상 Hotspot 분석 결과

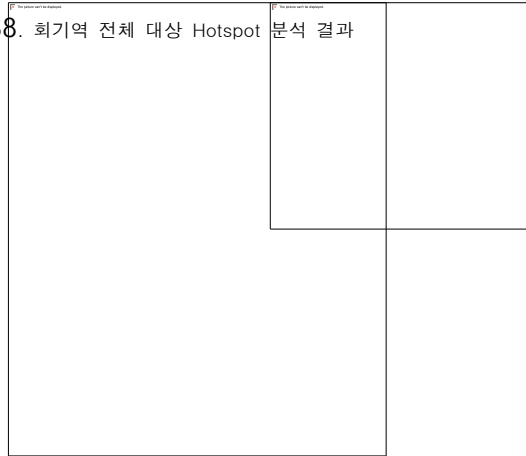


그림 39. 회기역 쏠림 군

calculate distance band from neighbor count			
회기(전체)		회기(Cluster0)	
average distance	95	average distance	308
maximum distance	1064	maximum distance	1056
incremental spatial autocorrelation			
회기(전체)		회기(Cluster0)	
1254		No peak (분석 시 1056대입)	

표 24. 회기역 spatial autocorrelation 분석 결과

회기역 전체 정류장에 대한 핫스팟 분석 결과 앞서와는 달리 회기동 방향으로 핫스팟이 형성되었다. 게다가 최대 부하량 군집인 cluster0에 대해 핫스팟 분석을 시행한 결과, 90% 이상의 신뢰수준에서 핫스팟이 존재하지 않는 것으로 나타났다.

#### (4) 분석 결과

회기역은 8~9시가 첨두시간인 것으로 조사되었으며, 역사 북쪽의 이문동 방면으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 보았을 때, 교육시설이 형성되어 있음을 알 수 있었고, 이를 통해 회기역에서 북쪽의 이문동 방면으로의 환승편향은 각 정류장 근처에 대학교로 등교하는 학생들로부터 기인했음을 알 수 있다.

### 5. 길음(18시~19시)

#### (1) 클러스터링 결과 및 부하량 모델 적용

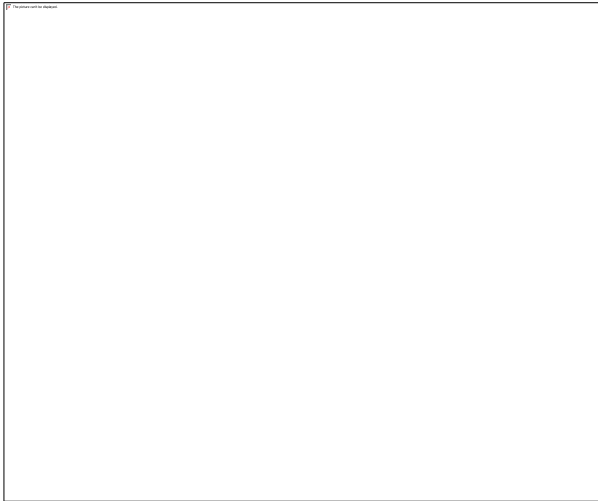


그림 45. 길음역 클러스터링 결과

길음역에 대한 클러스터링 결과 cluster0은 의 길음동 방면, cluster 1은 동쪽의 돈암동~하동 방면으로 구분되었다. 각각의 클러스터에 대하여 하량 모델을 적용한 결과 cluster0은 567.81, cluster 1은 5.01로 110배 정도의 큰 차이가 나타났고, 역에서의 환승문제는 주로 길음동 방면으로의 환승에서 기인할 것이라는 것을 확인할 수 있다.

● cluster0

● cluster1

◆ 길음역

□ 클러스터 별 부하량

ID	부하량
cluster0	567.81
cluster1	5.01

서쪽  
월곡  
해부  
cluster  
길음  
승에

표 26. 길음역 부하량 모델 적용 결과

## (2) Kernel Density

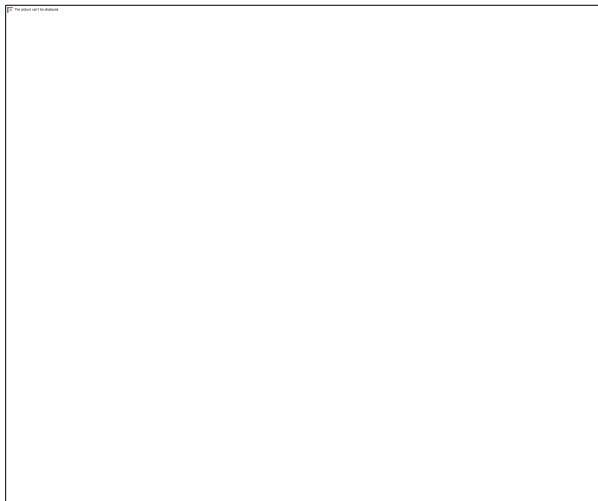


그림 46. 길음역 Kernel Density 적용 결과

길음역의 어떠한 유인이 이렇게 편향된 환승 행태를 야기하는지 확인하기 위해 Kernel Density를 적용한 결과 서경대학교 및 대일외국어고등학교의 교육시설과 길음뉴타운푸르지오 2단지아파트, 길음뉴타운e편한세상4단지아파트, 미아뉴타운래미안트리베라1차아파트등의 유인이 환승수요에 기여했을 것으로 보인다. 특히 18~19시라는 시간대를 감안했을 때 퇴근 후 귀가하는 사람들이 거주하는 주거단지가 길음동 방향에 집중되어 때문에 편향된 환승 행태가 나타나는 것으로 예상된다.

## (3) 공간통계분석(Hotspot Analysis)



### (3) 공간통계분석(Hotspot Analysis)

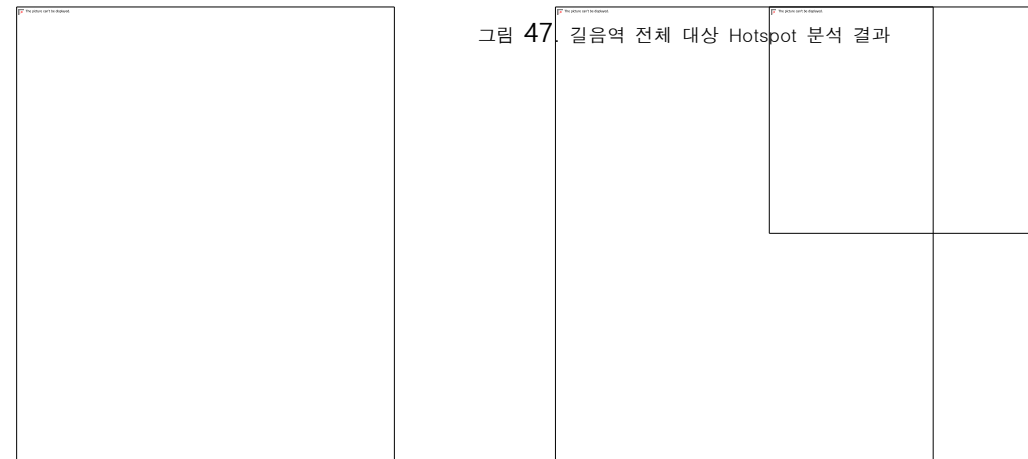


그림 47. 길음역 전체 대상 Hotspot 분석 결과

그림 48. 길음역 쏠림 군

calculate distance band from neighbor count			
길음(전체)		길음(Cluster0)	
average distance	191	average distance	140
maximum distance	1846	maximum distance	768
incremental spatial autocorrelation			
길음(전체)		길음(Cluster0)	
No peak (분석 시 1846 대입)		1188	

표 27. 길음역 spatial autocorrelation 분석 결과

길음역 전체 정류장에 대한 핫스팟 분석 결과 앞서 확인했던 바와 같이 서쪽의 길음동 방향으로 핫스팟이 형성됨을 알 수 있다. 그 중에서 더 정밀한 유인을 찾기 위해 cluster0에 대해 핫스팟 분석을 시행하였고 그 결과 99% 신뢰수준에서 핫스팟인 정류장은 봉국사, 정릉우체국앞, 정릉동태영아파트, 길음뉴타운8단지, 열매교회, 래미안501동앞, 미아초교, 미아뉴타운입구역으로 나타났다.

### (4) 분석 결과

길음역은 18~19시가 첨두시간인 것으로 조사되었으며, 역사 서쪽의 길음동 방면으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 보았을 때, 주거지역이 형성되어 있음을 알 수 있었고, 이를 통해 길음역에서 서쪽의 길음동 방면으로의 환승편향은 각 정류장 근처에 거주하는 사람들의 귀가에서 기인했음을 알 수 있다. 다만, 본 연구에서 2017년 05월 22일의 트립체인 데이터를 사용하였고, 이후 2017년 9월 2일에 위신설선이 개업하면서 문제가 예측되는 지점 근방에 정릉역이나 북한산보문국역 등의 지하철역이 들어왔음을 감안했을 때 최근의 트립체인 데이터를 사용하여 분석한다면 편향된 환승수요가 분산되었을 것임을 예측할 수 있다.

## 6. 신림(19시~20시)

## (1) 클러스터링 결과 및 부하량 모델 적용

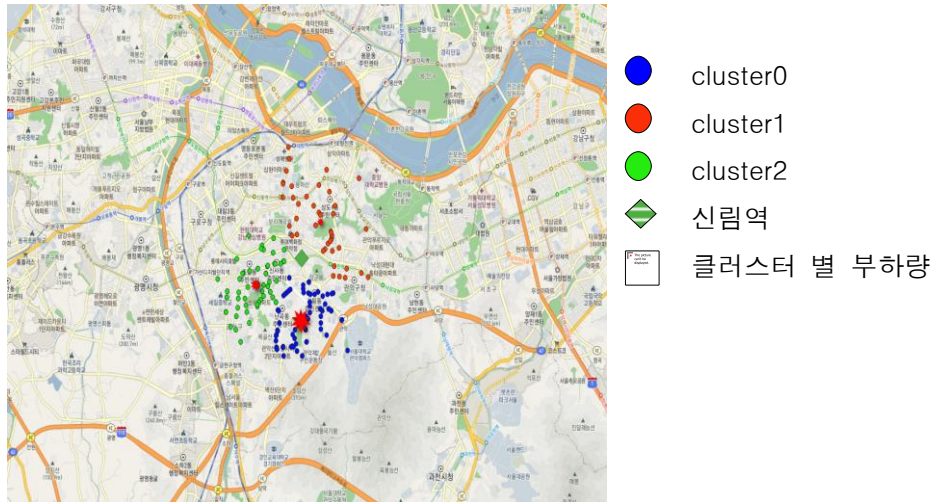


그림 55. 신림역 클러스터링 결과

신림역에 대한 클러스터링 결과 cluster0은 남쪽의 서림동, 대학동, 삼성동 방면, cluster 1은 북동쪽의 보라매동, 봉천동 방면, cluster2는 남서쪽의 신원동 방면으로 구분되었다. 각각의 클러스터에 대해 부하량 모델을 적용한 결과 cluster0은 560.23, cluster1은 38.95 cluster2는 144.78로 cluster0의 부하량이 임계부하 200의 2.8배 정도로 나타났고, 신림역에서의 환승은 주로 서림동, 대학동, 삼성동 방면으로의 환승이 기인할 것이라는 것을 확인할 수 있다.

ID	부하량
cluster0	560.23
cluster1	38.95
cluster2	144.78

량인  
승문  
승에

## (2) Kernel Density

표 29. 신림역 부하량 모델 적용 결과

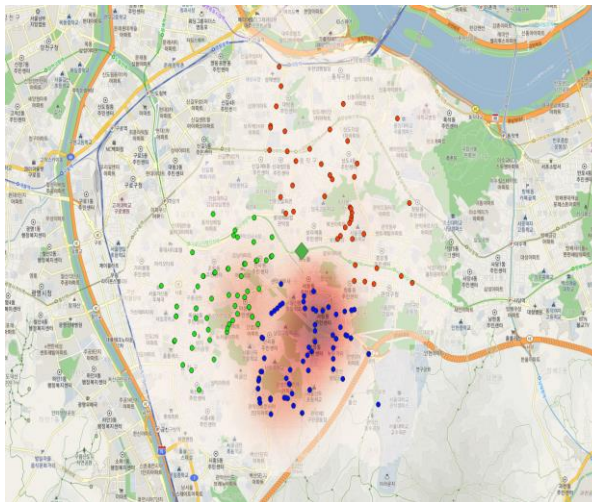


그림 56. 신림역 Kernel Density 적용 결과

신림역의 어떠한 유인이 이렇게 편향된 환승 행태를 야기하는지 확인하기 위해 Kernel Density를 적용한 결과 서림동 주민센터와 삼성동 주민센터 근방에 환승수요에 기여한 유인이 있을 것으로 보인다. 특히 19~20시라는 시간대를 감안했을 때 퇴근 후 귀가하는 사람들

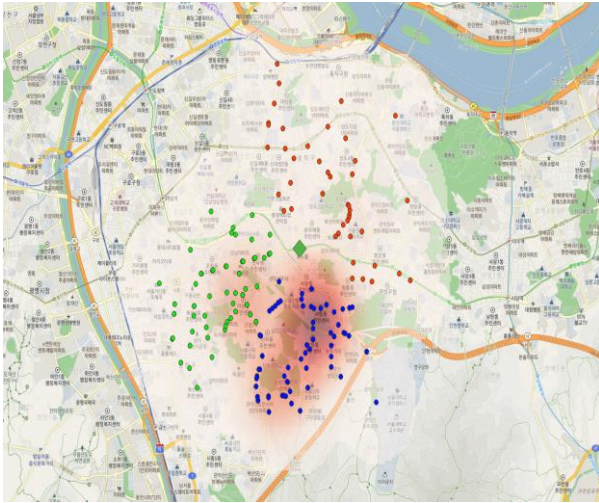


그림 56. 신림역 Kernel Density 적용 결과

신림역의 어떠한 유인이 이렇게 편향된 환승 행태를 야기하는지 확인하기 위해 Kernel Density를 적용한 결과 서림동 주민센터와 삼성동 주민센터 근방에 환승수요에 기여한 유인이 있을 것으로 보인다. 특히 19~20시라는 시간대를 감안했을 때 퇴근 후 귀가하는 사람들이 거주하는 주거단지가 서림동, 대학동, 삼성동 방향에 집중되어 때문에 편향된 환승 행태가 나타나는 것으로 예상된다.

### (3) 공간통계분석(Hotspot Analysis)

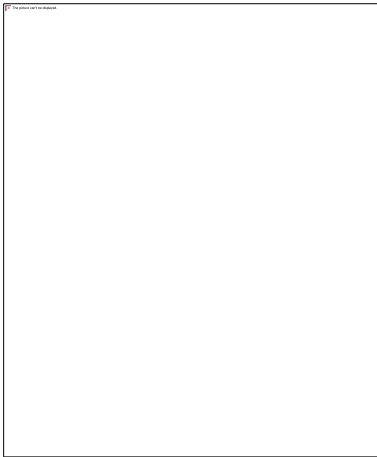


그림 57. 신림역 전체 대상 Hotspot 분석 결과

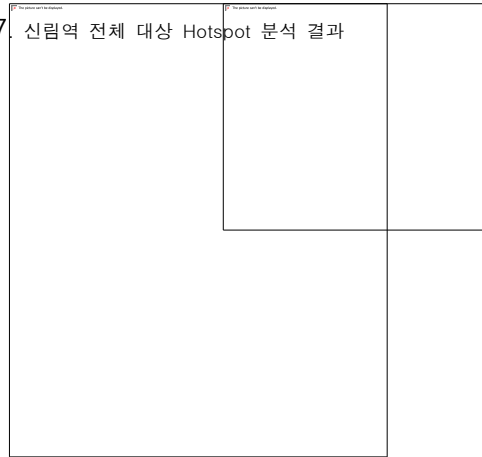


그림 58. 신림역 솔림 군...

calculate distance band from neighbor count			
신림(전체)		신림(Cluster0)	
average distance	126	average distance	56
maximum distance	1167	maximum distance	903
incremental spatial autocorrelation			
신림(전체)		신림(Cluster0)	
1797		1183	

표 30. 신림역 spatial autocorrelation 분석 결과

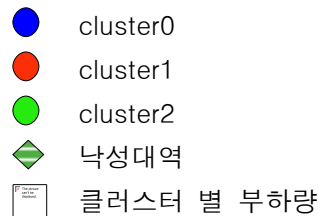
신림역 전체 정류장에 대한 핫스팟 분석 결과 앞서 확인했던 바와 같이 남쪽의 서림동, 대학동, 삼성동 방향으로 핫스팟이 형성됨을 알 수 있다. 그 중에서 더 정밀한 유인을 찾기 위해 cluster0에 대해 핫스팟 분석을 시행하였고 그 결과 99% 신뢰수준에서 핫스팟인 정류장은 관악세무서별관, 서원동문화교앞, 서림동주민센터, 현대아파트, 삼성동시장입구, 신성초등학교, 삼성동주민센터, 신성초등학교, 대학동고시촌입구, 대학동치안센터, 대학동노인정 등으로 나타났다.

#### (4) 분석 결과

신림역은 19~20시가 첨두시간인 것으로 조사되었으며, 역사 남쪽의 서림동, 대학동, 삼성동 방향으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 보았을 때, 주거지역이 형성되어 있음을 알 수 있었고, 이를 통해 신림역에서 남쪽의 서림동, 대학동, 삼성동 방향으로의 환승편향은 각 정류장 근처에 거주하는 사람들의 귀가에서 기인했음을 알 수 있다. 다만, 본 연구에서 2017년 05월 22일의 트립체인 데이터를 사용하였고, 이후 2022년 5월 28일에 경전철인 신림선이 개업하면서 문제가 예측되는 지점 근방에 서원역이나 서울대벤처타운역 등의 지하철역이 들어왔음을 감안했을 때 최근의 트립체인 데이터를 사용하여 분석한다면 편향된 환승수요가 분산되었을 것임을 예측할 수 있다.

### 7. 낙성대(09시~10시)

#### (1) 클러스터링 결과 및 부하량 모델 적용



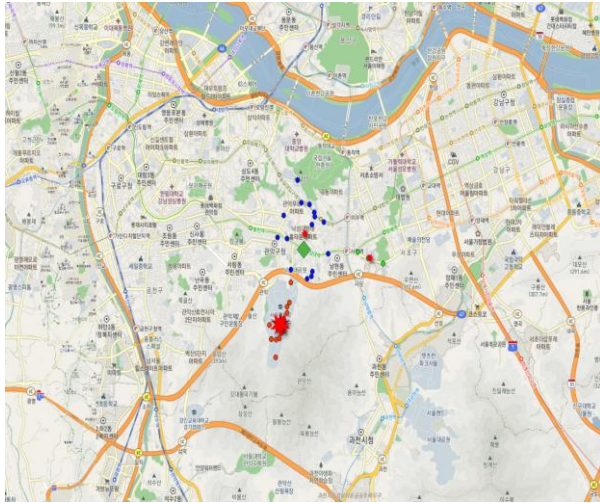


그림 65. 낙성대역 클러스터링 결과

낙성대역에 대한 클러스터링 결과 cluster0은 북쪽의 사당5동 방면 및 남동쪽의 인현동 방면, cluster 1은 남쪽의 서울대학교 방면, cluster2는 동쪽의 사당역 방면으로 구분되었다. 각각의 클러스터에 대해 부하량 모델을 적용한 결과 cluster0은 3.12, cluster1은 544.90, cluster2는 0.01로 cluster1의 부하량이 임계부하 200의 2.7배 정도로 나타났고, 낙성대에서의 환승은 주로 서울대학교 방면으로의 환승에서 기인한다는 것을 확인할 수 있다.

## (2) Kernel Density

ID	부하량	량인 승문 할 것
cluster0	3.12	
cluster1	544.90	
cluster2	0.01	

표 32. 낙성대역 부하량 모델 적용 결과

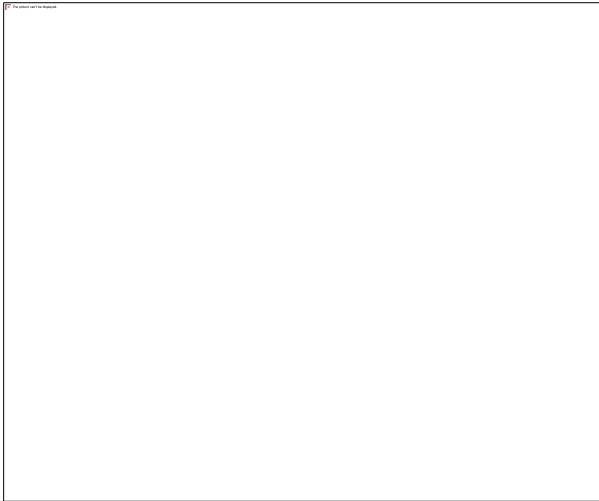


그림 66. 낙성대역 Kernel Density 적용 결과

낙성대역의 어떠한 유인이 이렇게 편향된 환승 행태를 야기하는지 확인하기 위해 Kernel Density를 적용한 결과 서울대학교 근방에 환승수요에 기여한 유인이 있을 것으로 보인다. 특히 09~10시라는 시간대를 감안했을 때 서울대학생들의 등교시간과 맞물려 서울대학교 방향으로 편향된 환승 행태가 나타나는 것으로 예상된다.

### (3) 공간통계분석(Hotspot Analysis)

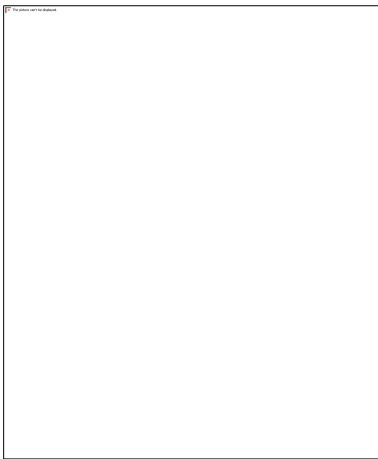


그림 67. 낙성대역 전체 대상 Hotspot 분석 결과

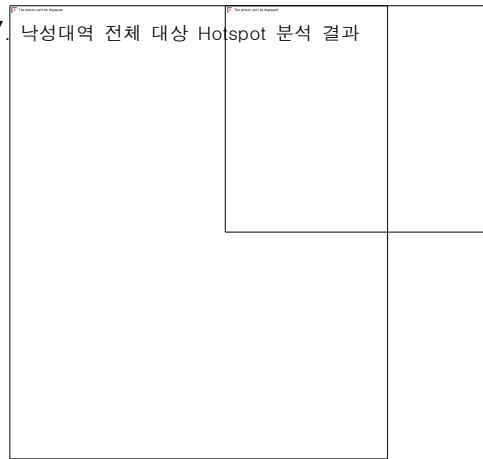


그림 69. 낙성대역 쏠림

calculate distance band from neighbor count			
낙성대(전체)		낙성대(Cluster1)	
average distance	464	average distance	220
maximum distance	569	maximum distance	589
incremental spatial autocorrelation			
낙성대(전체)		낙성대(Cluster1)	
2209		No peak (분석 시 589대입)	

표 33. 낙성대역 spatial autocorrelation 분석 결과

낙성대역 전체 정류장에 대한 핫스팟 분석 결과 앞서 확인했던 바와 같이 남쪽의 서울대

학교 방향으로 핫스팟이 형성됨을 알 수 있다. 그 중에서 더 정밀한 유인을 찾기 위해 cluster1에 대해 핫스팟 분석을 시행하였고 그 결과 99%나 95% 신뢰수준에서 핫스팟인 정류장은 나타나지 않았고, 90%의 신뢰수준에서 핫스팟인 정류장은 기숙사신관, 기숙사구관, 기숙사삼거리로 나타났다.

#### (4) 분석 결과

낙성대역은 09~10시가 첨두시간인 것으로 조사되었으며, 역사 남쪽의 서울대학교방면으로 수요가 쏠려있는 것을 확인할 수 있었다. 핫스팟으로 선정된 각 정류장의 근방을 확인해 보았을 때, 인문대, 도서관, 아랫공대 등의 건물이 위치해 있었고 이를 통해 낙성대역에서 남쪽의 서울대학교 방면으로의 환승편향은 인문대나, 공대생들의 통학수요에서 기인함을 확인할 수 있었다.

### IV. 사례분석

#### 1. 회기

본 연구가 실제로 환승이 문제되는 지점을 예측하는데 실효성이 있는지 판단하기 위해 회기역 근방 거주자 17명에 대해 설문조사를 통해 사례분석을 진행하였다.

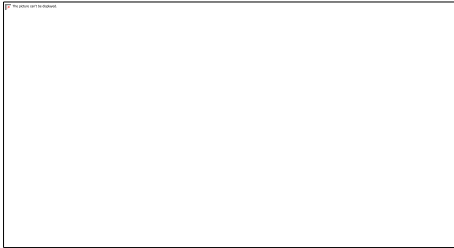
	질문	설문조사 및 결과	차트
1	회기역에서 버스로 환승 시 불편함을 느끼는 정도 (17)	① 매우 편하다 (1) ② 다소 편하다 (6) ③ 보통이다 (3) ④ 다소 불편하다 (6) ⑤ 매우 불편하다 (1)	
2	버스로 환승한다면 최종 하차역 (13)	- 경희대의료원 (5) - 경희대입구 (6) - 기타 (2)	
3	환승이 편한 이유 (7)	- 정류장까지 거리가 짧음 (3) - 배차간격이 짧음 (1) - 정류장까지 거리가 짧고 배차간격도 짧음(3)	
4	환승이 불편한 이유 및 불편함에도 환승을 하는 이유 (10)	- 기다리는 줄 위치가 위험함 (1) - 도보로 10분 이상 걸어야 캠퍼스 정문에 도착 (4) (건기에는 조금 멀고 택시 타기에는 가까운 거리) - 버스 외에 학교까지 갈 수 있는 다른 수단이 없음 (4) - 버스를 타 교통수단보다 좋아함 (1)	

표 34. 회기역 설문조사 결과

설문조사 결과 회기역에서 버스로 환승 시 불편함을 느끼는 정도에 대해 ③ 보통이다 ④



다소 불편하다 ⑤ 매우 불편하다 라고 응답한 인원은 총 10명으로 전체의 58.8%에 달하였으며 이를 통해 회기역에서 버스로의 환승에 어느 정도 불편이 있음을 실제로 확인할 수 있다. 또한 최종 하차역에 대한 질문에 응답한 13명 중 11명인 84.6%의 인원이 경희대 방향으로 환승하는 것을 통해 쏘림 현상이 버스 환승으로의 문제에 영향을 준다는 것을 알 수 있었다. 최종적으로 환승이 불편하다고 답한 10명을 대상으로 그 이유를 조사한 결과 도보로 이동하기 불편한 거리에 목적지가 위치해 있음 및 버스 외에 목적지 까지 갈 수 있는 다른 수단이 없음이 각각 40% 를 차지했다. 따라서 경희대 근처에 쏘림을 분담할 수 있는 새로운 지하철 역을 신설하거나, 버스 외에 경희대로 갈 수 있는 다른 수단(전동 킥보드 등)을 확충한다면 회기역에서의 환승 문제를 어느 정도 완화시킬 수 있을 것이다.

## 2. 낙성대

회기역과 마찬가지로 낙성대역 근방 거주자 17명에 대해 설문조사를 통해 사례분석을 진행하였다.

	질문	설문조사 및 결과	차트
1	낙성대역에서 버스로 환승 시 불편함을 느끼는 정도 (52)	① 매우 편하다 (9) ② 다소 편하다 (21) ③ 보통이다 (14) ④ 다소 불편하다 (7) ⑤ 매우 불편하다 (1)	
2	버스로 환승한다면 최종 하차역 (33)	- 기숙사 (4) - 노천강당 (5) - 기숙사 삼거리(6) - 서울대학교 내부(14) - 서울대입구역 (2) - 대학동 (1) - 상도역 (1)	
3	환승이 편한 이유 (28)	- 정류장까지 거리가 짧음 (11) - 배차간격이 짧음 (8) - 정류장까지 거리가 짧고 배차간격도 짧음(9)	
4	환승이 불편한 이유 및 불편함에도 환승을 하는 이유 (22)	- 버스에 사람이 많음 (6) - 버스 외에 학교까지 갈 수 있는 다른 수단이 없음 (14) - 대기시간이 김 (2)	

표 35. 낙성대역 설문조사 결과

설문조사 결과 낙성대역에서 버스로 환승 시 불편함을 느끼는 정도에 대해 ③ 보통이다 ④ 다소 불편하다 ⑤ 매우 불편하다 라고 응답한 인원은 총 22명으로 전체의 58.8%에 달하였으며 이를 통해 낙성대에서 버스로의 환승에 어느 정도 불편이 있지만 환승객이 불편을



느끼는 정도가 생각보다 크지 않음을 알 수 있었다. 이는 본 연구에서 2017년의 트립체인 데이터를 사용하는데 반해 최근 코로나로 인해 비대면 수업이 늘어나면서 환승 수요가 줄어들어 환승객들이 느끼는 불편이 줄어들었기 때문인 것으로 생각된다.

또한 최종 하차역에 대한 질문에 응답한 33명 중 29 명인 87.9%의 인원이 서울대학교 방향으로 환승하는 것을 통해 쏠림 현상이 버스 환승으로의 문제에 영향을 준다는 것을 알 수 있었다. 최종적으로 환승이 불편하다고 답한 22명을 대상으로 그 이유를 조사한 결과 버스 외에 목적지 까지 갈 수 있는 다른 수단이 없음이 63.6% 를 차지했다. 따라서 낙성대역 근처에 서울대학교로 갈 수 있는 다른 수단(전동 킥보드 등)을 확충한다면 낙성대 역에서의 환승 문제를 어느 정도 완화시킬 수 있을 것이다.

## V. 결론

본 연구에서는 지하철의 간선으로서의 역할, 버스의 지선으로서의 역할에 주목하여 지하철에서 버스로의 환승 시 특정방향으로의 수요 쏠림이 심하게 일어난다면 해당 방향으로의 통행에 과부하가 발생할 것이라는 가설을 수립한 후 클러스터링에 기반한 부하량 산정식을 모델링 하였다. 이를 통해 지하철-버스 환승 문제지점을 예측하고, 이후 문제 발생의 유인을 공간통계 기법을 적용함으로써 특정 짓고자 하였다.

각각의 하차정류장이 교통량을 분담하는 정도를 부하량으로서 정량적으로 계산했고, 공간 좌표를 기준으로 군집화를 수행하여 지하철역을 중심으로 각 방향이 교통량을 분담하는 정도를 군집 부하량으로서 계산하였다. 이 군집 부하량 값에 대해 모든 시간대에 대한 비교를 수행하였고 교통 분담의 쏠림이 가장 큰 시간대를 특정하여 지하철 부하량을 계산하였다. 이를 바탕으로 모든 지하철 역사에 대해 비교를 수행하여 환승 문제지점을 예측하였다. 한편, 쏠림의 유인이 있는 지역을 특정하고 그에 대한 분석을 위해 상위 6개 역사에 대해 공간통계 분석을 진행하였다. 먼저 앞서 구한 부하량 값에 대한 Kernel Density Estimation을 통한 시각화를 통해 최종 하차 정류장으로부터 환승객들의 최종 목적지를 예측하여 대략적인 수요 발생의 유인을 특정하였다. 이후 Hotspot Analysis를 수행하여 수요 발생의 유인을 시설 단위까지 예측함으로써 지하철-버스 환승 개선 시 고려해야 할 대상에 대해 보다 정밀하게 찾아 낼 수 있었다.

정리하자면, 본 연구에서는 클러스터링에 기반한 부하량 산정 모델을 통해 지하철-버스 환승 문제지점을 예측하였고, 이를 개선하기 위해 공간통계분석을 적용하여 유인을 분석하였다. 본 연구의 방법론을 적용함으로써, 교통 수요가 야기하는 지하철-버스 환승 문제 지점을 예측할 수 있고, 해당 문제의 유인을 시설 단위까지 상세하게 분석할 수 있어, 개선해야 할 대상을 선정하고 그것을 적용했을 때의 결과를 예상해보는데 활용할 수 있다.

한편, 본 연구에는 크게 3가지 제한 사항이 있다. 첫 번째로, 본 연구에서 사용한 트립체인 데이터는 17년도 데이터이므로 현 시점의 문제를 제대로 파악하지 못할 수 있다. 새로운 시설이 들어왔거나, 버스 노선이 변경된 사항, 지하철 노선 추가, 코로나로 인한 통행패턴 변화 등이 반영되지 않아 현 시점에서 발생하는 환승 문제 양상과 차이를 보일 수 있다. 이 문제는 최근에 수집된 데이터로 분석을 진행한다면 해소할 수 있을 것이다.

두 번째로, 교통 공급량을 버스 노선수로 정의했는데, 실제로는 노선수 뿐만 아니라 각 노선의 배차 간격까지 고려해야 보다 정확한 공급량을 계산할 수 있다. 본 연구에서는 데이터 확보의 한계 및 전처리 과정의 복잡함으로 인해 고려하지 못하였으므로, 해당 부분은 후속 연구에서 개선 해야할 사항이다.

마지막으로, 본 연구에서는 환승 거리, 환승 대기 시간 등 보편적인 환승 서비스 수준 분석에서 고려되는 사항들을 고려하지 않았기 때문에 실제 사용자가 느끼는 불편의 정도와 차이를 보일 수 있다. 수요가 많이 쏠리더라도 환승 거리가 매우 짧다면 환승에서 느끼는 불편함은 줄어들 여지가 있다. 그러나 이는 교통 수요를 집중시키는 유인을 찾기위해 의도적으로 설계된 사항이다. 따라서, 본 연구의 적용 범위를 목적에 맞게 유의해서 설정할 필요

가 있다.

VI. Appendix

1. 수도권 전체 역사별 부하량

역명	최대쏠림시간	최대부하량
디지털미디어시티역	8	739.729053
청담역	8	699.398161
회기역	8	671.907692
논현역	8	585.362698
길음역	18	567.813998
신림역	19	560.22967
낙성대역	9	544.900156
쌍문역	18	501.270431
수유역	19	463.792777
상일동역	19	457.684322
양재역	8	452.301335
구로디지털단지역	17	440.271278
가산디지털단지역	8	438.722814
북정역	19	390.782811
개봉역	18	335.102158
봉천역	18	333.62569
수서역	19	325.766082
노량진역	18	320.260265
당산역	22	316.439951
장한평역	19	305.027858
신대방역	22	302.434535
온수역	7	297.950156
서울대입구역	9	296.973191
매봉역	8	277.186801
남부터미널역	8	274.023968
미아사거리역	19	270.28223
연신내역	19	240.914988
건대입구역	19	231.555899
종합운동장역	18	223.238416
신대방삼거리역	8	223.079485
내방역	8	219.872615

합정역	19	206.434583
성신여대입구역	19	203.197785
대방역	7	199.96426
신정네거리역	19	198.634612
홍제역	19	192.647476
대림역	19	185.974026
독산역	19	185.578125
문래역	8	184.762774
장지역	19	181.191772
신도림역	18	181.144335
서울역	8	176.394576
노원역	19	175.776856
삼성역	8	175.324379
여의나루역	8	174.982906
신논현역	7	168.851085
석계역	18	168.587709
홍대입구역	7	164.442879
광흥창역	8	164.137652
상봉역	19	163.836512
옥수역	19	160.911628
강남역	8	160.473353
흑석역	8	159.585859
신촌역	19	157.769762
석수역	19	156.746253
잠실역	19	155.226284
금천구청역	18	152.531697
구파발역	18	148.960358
충신대입구역	8	146.822755
종로3가역	8	145.816901
한성대입구역	19	142.26087
왕십리역	18	140.201389
상계역	19	139.808007
증산역	9	136.152941
우장산역	19	135.692308
서초역	8	134.742964
천호역	19	133.878514
까치산역	19	132.121934
영등포역	8	130.582701
신용산역	8	124.343026
구로역	8	124.057915

새절역	9	123.459893
교대역	8	120.661017
사당역	18	119.05674
경복궁역	7	116.445328
시청역	8	111.968553
하계역	19	111.450851
광명사거리역	19	108.696695
창동역	19	108.639397
면목역	9	104.387755
청량리역	18	104.154811
상왕십리역	8	100.140476
중계역	19	97.8611898
동대문역	8	97.0212766
영등포구청역	19	95.030303
공덕역	19	94.4137931
가락시장역	22	92.9966887
오목교역	19	91.3901515
구의역	19	91.0292887
화곡역	18	90.4539326
암사역	19	89.3023256
압구정역	9	88.5457958
안국역	8	87.6939502
선릉역	8	85.9175532
신내역	7	84.4615385
미아역	19	84.1388889
학동역	8	83.3596273
광화문역	8	83.0231982
신이문역	19	81.8514056
강남구청역	9	81.5321152
응암역	19	81.1240741
신정역	18	81.046875
충정로역	18	79.3165468
돌곶이역	19	79.0828221
녹번역	19	77.2173913
방배역	8	76.8174988
고속터미널역	8	76.744186
사가정역	8	75.8293076
답십리역	19	75.7894737
고려대역	19	75.5002302
봉화산역	19	74.6385983

남영역	8	74.6178862
혜화역	19	74.4615385
태릉입구역	11	74.4567901
성수역	22	74.297619
목동역	18	73.16961
월곡역	19	71.770073
염창역	20	71.6274414
김포공항역	7	71.4119522
여의도역	7	69.7423174
강변역	18	68.9777155
강동구청역	18	68.1428571
도곡역	20	67.8258145
이수역	18	67.585997
신설동역	19	67.1612022
종로5가역	8	66.75
신당역	18	66.0357143
중화역	19	65.7630719
신길역	18	64.5216049
종각역	8	62.7269504
양천향교역	19	61.5416667
공릉역	13	61.4931507
녹사평역	18	59.7439024
금호역	19	59.0582524
이대역	8	58.5576923
가양역	22	58.0578035
아현역	8	57.8275862
숙대입구역	8	57.8033958
동묘앞역	18	57.6352941
광운대역	19	57.5650407
용산역	7	56.4971751
둔촌동역	18	54.2440945
수락산역	19	53.6022099
역삼역	7	53.443299
마들역	19	53.1702128
삼각지역	8	52.744186
언주역	8	52.0438596
월계역	7	50.969697
선정릉역	8	50.502924
신사역	8	49.8157895
남구로역	8	49.4478528

동대문역사문화공원역	8	47.4313725
송실대입구역	20	47.1855288
발산역	19	45.8229167
이촌역	8	44.7944444
오금역	19	44.0282258
미장역	8	43.9636364
녹천역	19	43.5125
상도역	9	42.7179487
망원역	19	42.6393443
한티역	8	42.2395161
보문역	18	41.6802469
석촌역	8	41.5686275
효창공원앞역	10	41.25
선유도역	18	41.0098039
제기동역	8	40.7140152
화랑대역	19	40.3333333
양재시민의숲역	7	39.6134141
신답역	18	39.0963855
역촌역	19	38.5285714
불광역	19	38.367713
신천역	20	38.1034483
오류동역	18	37.7064677
등촌역	19	37.1304348
문정역	19	36.75
월드컵경기장역	8	35.6435644
가좌역	9	35.1724138
개화역	8	34.8426724
도봉산역	19	34.8021489
군자역	19	34.5
마포역	19	34.3225806
남성역	18	33.9846154
독립문역	8	33.970674
서대문역	18	33.509434
송파역	8	33.2444444
안암역	8	32.890411
독섬역	19	32.8493151
봉은사역	9	32.6545359
독섬유원지역	18	32.3269231
한양대역	8	32.1315789
이태원역	8	31.6967871

천왕역	19	30.5344828
약수역	19	30.3868313
서울숲역	19	30.1162791
광나루역	8	30.0042194
한남역	8	29.6722222
강동역	21	28.6637681
잠실나루역	19	28.4761905
먹골역	19	28.3928571
장승배기역	18	27.6129032
압구정로데오역	8	27.0833333
신금호역	19	26.8837209
대청역	19	26.28125
반포역	8	26.1061644
외대앞역	18	24.7352941
대모산입구역	8	24.6551724
신풍역	19	23.5202703
일원역	18	23.0588235
개포동역	8	22.8118812
동대입구역	8	22.1875
개롱역	18	21.25
아차산역	18	19.9460317
상월곡역	19	19.862069
몽촌토성역	8	19.4791667
용마산역	19	19.24
충무로역	10	19.1842105
마포구청역	18	19.125
대치역	8	19.1037736
보라매역	8	18.974359
노들역	18	18.7717391
도봉역	19	18.5
회현역	8	18.1956522
명일역	20	17.474359
신목동역	19	17.3063725
서빙고역	8	17.1428571
망우역	19	16.4615385
당고개역	19	16.1904762
고덕역	18	15.9574468
한강진역	9	15.6216216
양천구청역	19	15.5454545
대흥역	19	15.2236842

어린이대공원역	8	15.1775701
방학역	18	14.8536585
영등포시장역	8	14.2258065
무악재역	8	13.8888889
서강대역	8	13.4354839
종량역	19	13.2
을지로입구역	8	12.9930314
송정역	20	11.5
올림픽공원역	7	11.4545455
중곡역	18	10.7142857
샛강역	8	10.5945946
마곡역	18	9.81818182
동작역	19	8.75
청구역	8	8.71212121
신반포역	8	8.625
경찰병원역	19	8.33333333
용답역	18	8.25
공항시장역	19	8.22954545
방이역	8	7.35632184
수색역	8	6.125
용두역	18	5.84375
증미역	19	5.78571429
구반포역	17	5.71428571
행당역	18	4.85294118
을지로4가역	8	4.5
학여울역	8	3.97058824
명동역	16	3.6
굽은다리역	16	3.5
양원역	19	3.23076923
을지로3가역	17	3
길동역	19	2.76923077
구산역	18	1.94444444
국회의사당역	17	1.45454545
도림천역	0	0
잠원역	0	0
남태령역	0	0
응봉역	0	0
구룡역	0	0
구일역	0	0
방화역	0	0



개화산역	0	0
양평역	0	0
애오개역	0	0
거여역	0	0
마천역	0	0
독바위역	0	0
상수역	0	0
버티고개역	0	0
창신역	0	0
신방화역	0	0
마곡나루역	0	0
사평역	0	0
삼성중앙역	0	0
청계산입구역	0	0

- 
- <sup>1</sup> 국토교통부 국토해양부 종합교통정책과 보도자료(2010), “수도권 대중교통 환승행태 조사”
  - <sup>2</sup> 황보희, 노정현, 박영욱(2006), “서울시 대중교통 이용객들의 환승시간가치 추정에 관한 연구”, 대한교통학회 학술대회지 51
  - <sup>3</sup> 김대성, 엄진기, 문대섭, 최명훈, 송지영(2011), “대중교통 card data를 이용한 통행행태 분석”
  - <sup>4</sup> 최명훈, 엄진기, 이준, 김대성(2011), “스마트카드 데이터를 활용한 역사별 연계 환승시간 서비스 평가”, 한국철도학회 학술발표대회논문집
  - <sup>5</sup> 손무성, 김광태, 김세원, 오석문(2011), “도시철도 환승시간 최소화를 위한 최적화 모형”, 한국철도학회 학술발표대회논문집
  - <sup>6</sup> 이상혁, 김재석, 김민석, 우용한(2013), “대중교통 이용자의 환승교통수단선택 행태분석에 관한 연구”, 한국ITS학회논문지 제 12권, 제 1호
  - <sup>7</sup> 신호석, 이은학, 이호영, 고승영, 김동규(2017), “수도권 지하철-버스의 환승분석을 통한 버스역할”, 대한교통학회 제 76회 학술발표회
  - <sup>8</sup> 대도시권광역위원회(2021), “제3차 환승센터 및 복합환승센터 구축 기본계획”, 국토교통부