

서울시 공공자전거 효율적인 운영을 위한 빅데이터 분석

[따릉따릉]
정승원 박지용
김준국 장채연
김명회 정해명

목차

I. 분석 배경 및 방향

- 분석 배경
- 분석 방향

II. 데이터 분석

- 재배치 분석 및 결과
- 고장데이터 분석 및 결과

III. 프로젝트 결론

- 활용방안 및 한계점
- 참고 문헌

I. 분석 배경 및 방향

- 분석 배경
- 분석 방향



밤마다 서울도심에는 따릉이가 없다

퇴근 직장인들 빌려타고 외곽으로
외곽 → 도심 옮기기 왕복 2시간
시내 보관소 10곳뿐...시스템 열악

변재현 기자 | 2018-12-14 17:42:38 | 사회일반



가

가

퇴근 시간 이후 밤마다 도심에 따릉이가 없는 '따릉이 공동화 현상'이 이어지고 있다. 회사가 밀집한 도심 지역에서 직장인들이 따릉이를 빌려 외곽으로 향하는 탓에 동이 나기 때문이다. 서울 외곽에서 따릉이를 도심

이용대수가 특정 시간 및 특정 장소에 밀집되는

현상으로 운영의 어려움

서울시청 앞에서 서울시시설관리공단 차량관리장소를 만들 시간은 부족하다. 35대를 주차할 수 있는
정류소에 남아있는 따릉이는 고장 4대, 그마저도 채워져 버렸던 거이지 반쯤까지는 한 대도 얻었다고 했다. 초 30대

따릉이, 고장으로 사고 급증 '안전대책' 시급

4년 간 고장 15만6803건...“고장 및 사고방지위해 노력해야”

조성우기자(jsw5655@skyedaily.com)

기사입력 2019-10-13 15:43:07

[기자의 다른 기사 보기](#)

많은 시민들이 이용하고 있는 따릉이의 고장이 늘어나고 있는 가운데 고장으로 인한 사고도 급증하고 있는 추세다.

13일 이후삼 더불어민주당 의원(국회 국토교통위원회)가 서울특별시부터 제출받은 자료에 따르면 최근 4년간 따릉이 고장 건수는 15만6803건에 달했다. 이는 지난해 1년 동안 총 발생한 고장 건수인 5만9571건과 비슷한 수준인 것으로 나타났다.

급격한 이용객 증가로 인한
서울시 공공자전거 고장률 증가 및 사고 증가

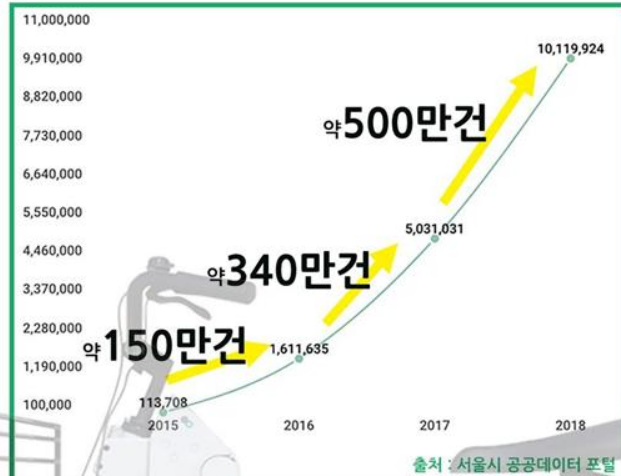
세부적으로 올해 발생한 5만1658건의 고장 원인은 △휠날아 7307건(15%) △타이어 5373건(11%) △브레이크 5296건(10%) △단말기 5650건(9%) 등으로 나타났다.

I. 분석 배경

따릉이 연도별 대여 건수



따릉이 연도별 고장 건수



따릉이의 연도별 대여건수가 늘어남에 따라
연도별 고장건수도 급속하게 늘어나는 추세

I. 분석 방향

✓ 이용 대수가 특정 시간 및 특정 장소에 **집중되는 현상**으로 인한 관리 운영의 어려움

- ➔ 계절별 요일별 시간대별로 데이터를 분석하여 집중되는 패턴을 파악
- ➔ 분석데이터를 기반으로 과수요 군집을 분류하여 **재배치 시간과 방법**을 제시



✓ 매년 이용객 증가로 인해 서울시 공공자전거 고장률 **증가**하는 추세

- ➔ 대여 이력 데이터를 분석하여 고장 자전거 **이용패턴** 파악
- ➔ 고장 신고 및 정비 내역을 분석하여 자전거 **상태를 진단, 군집화** 적용



II. 데이터 분석

- 재배치 분석
- 고장데이터 분석

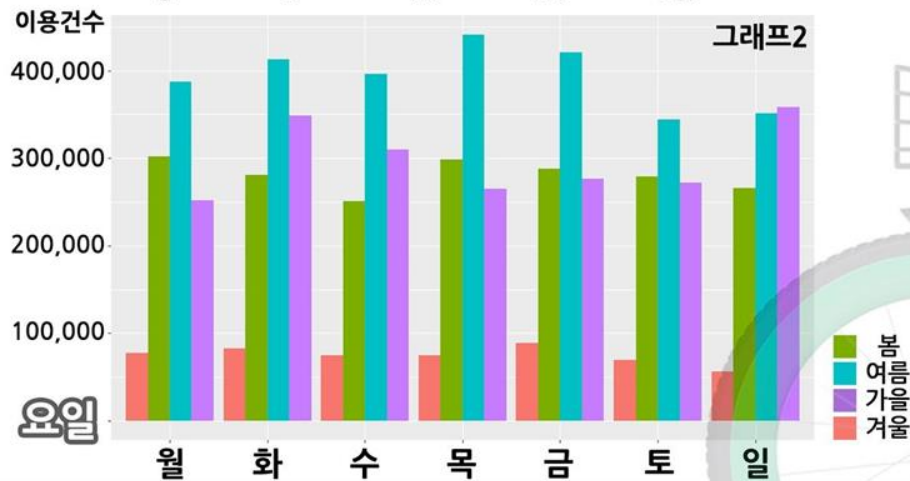
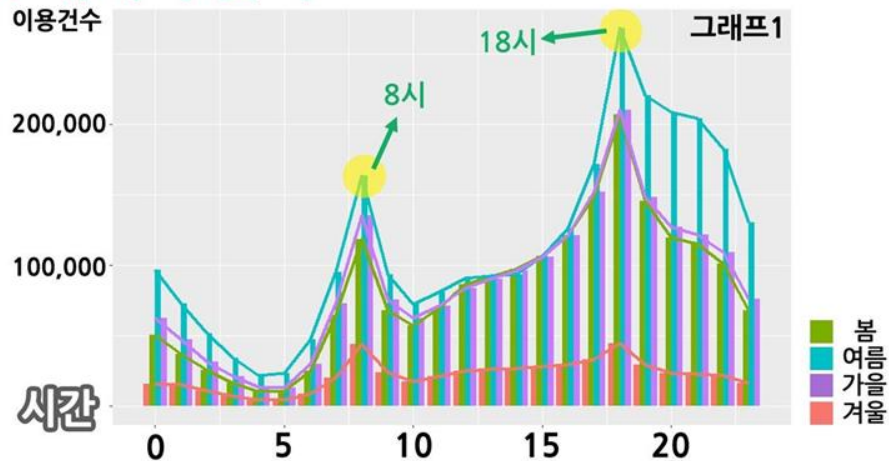
II. 재배치 분석



- ✓ 과거 데이터를 이용해 대여량에 영향을 주는 변수를 분석한 후 분석 결과에 따른 **재배치 방법**을 제안
- ✓ 출퇴근 시간에 **상위 대여 대여소**에 재배치 방법을 적용

II. 재배치 분석

여름 : 18.06 ~ 18.08 가을 : 18.09 ~ 18.11
겨울 : 18.12 ~ 19.02 봄 : 19.03 ~ 19.05



- ✓ 주말보다는 평일에
오전 8시와 오후 6시에 이용이 가장 많음
- ✓ 이용률에 가장 많은 영향을 미치는 요인은
평일 출퇴근 시간대라는 점을 알 수 있음
- ➔ 출근 시간 (7~ 9시) 와 퇴근시간 (17 ~ 20시)
해당 시간에 어느 대여소에서 대여와 반납을
하는지 상위 30개 대여소를 분석 및 해결책 제시

시간별, 요일별 Anova 검증을 실시 출퇴근 시간 (7~9시, 17~20시)의 유의성을 확인함
Anova 분석이란? 독립된 집단이 셋 이상인 경우, 집단간의 평균의 차이가 있는지 확인하기 위한 검증 방식

07-09 시

이동경로 이동방향

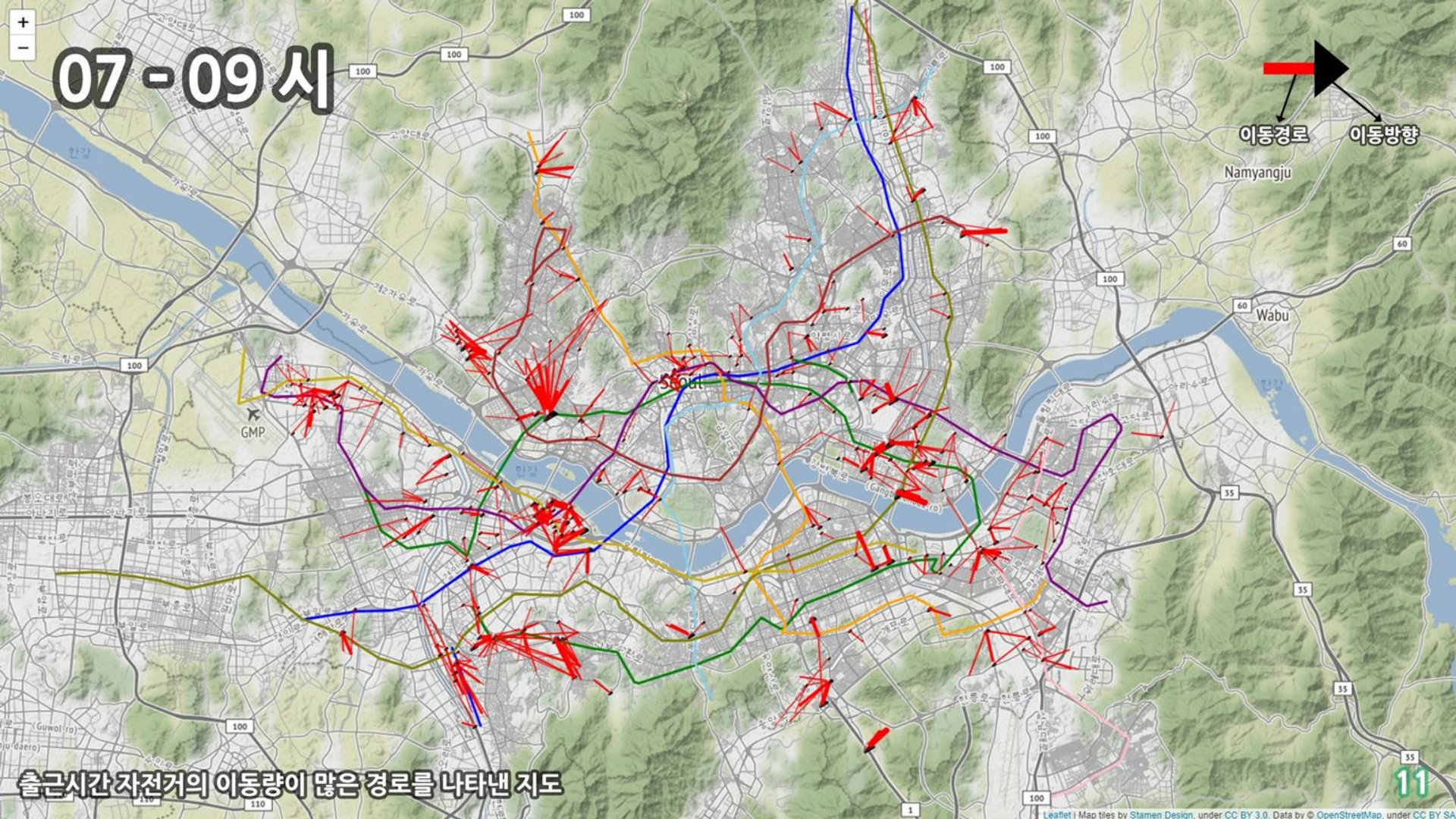
Namyangju

Wabu

GMP

11

출근시간 자전거의 이동량이 많은 경로를 나타낸 지도



07-09 시

이동경로 이동방향

Namyangju

Wabu

DMC
9번출구

홍대입구역
2번출구 앞

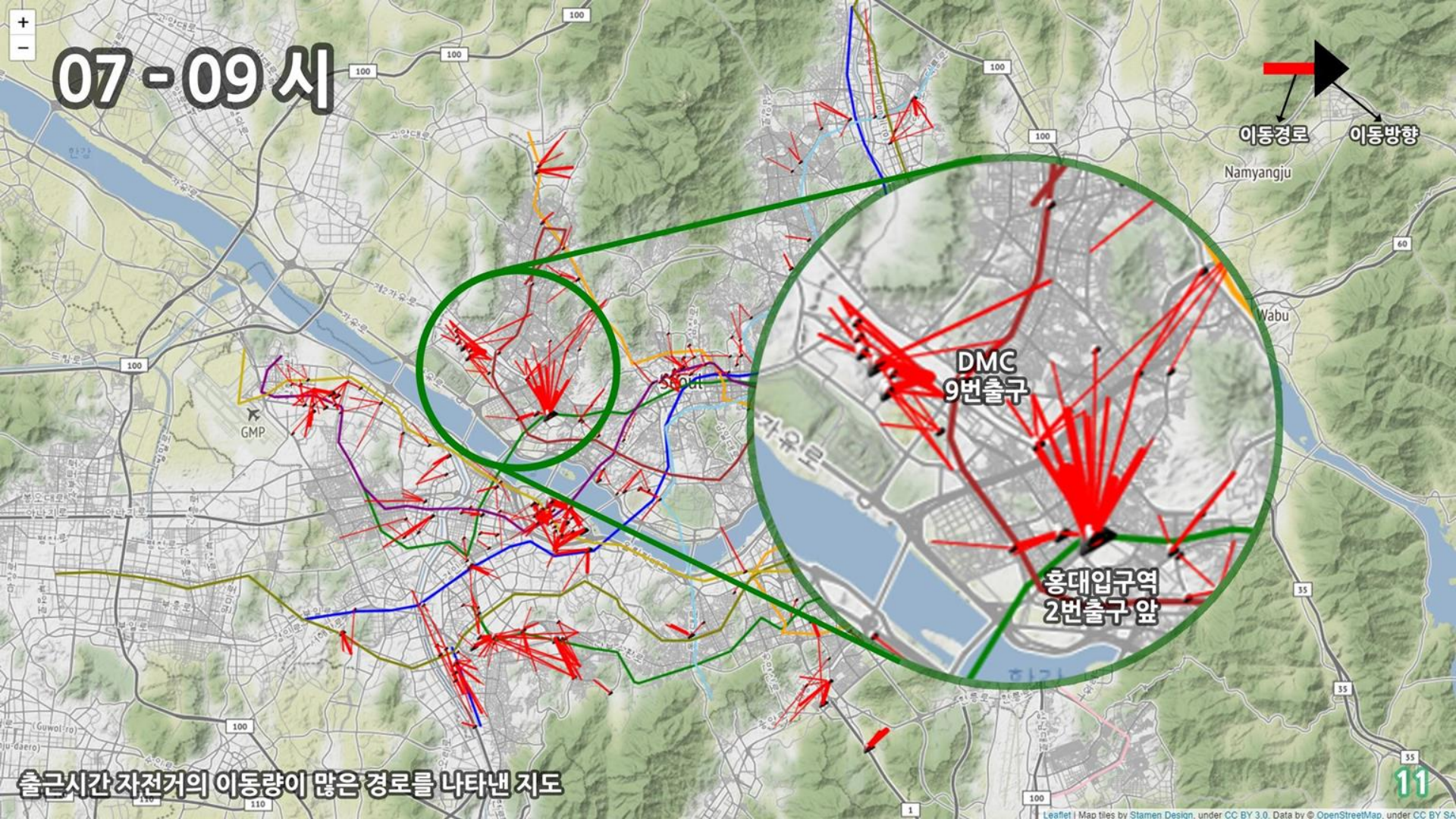
GMP

Seoul

Guwon

11

출근시간 자전거의 이동량이 많은 경로를 나타낸 지도



17 - 20 시

이동경로 이동방향

Namyangju

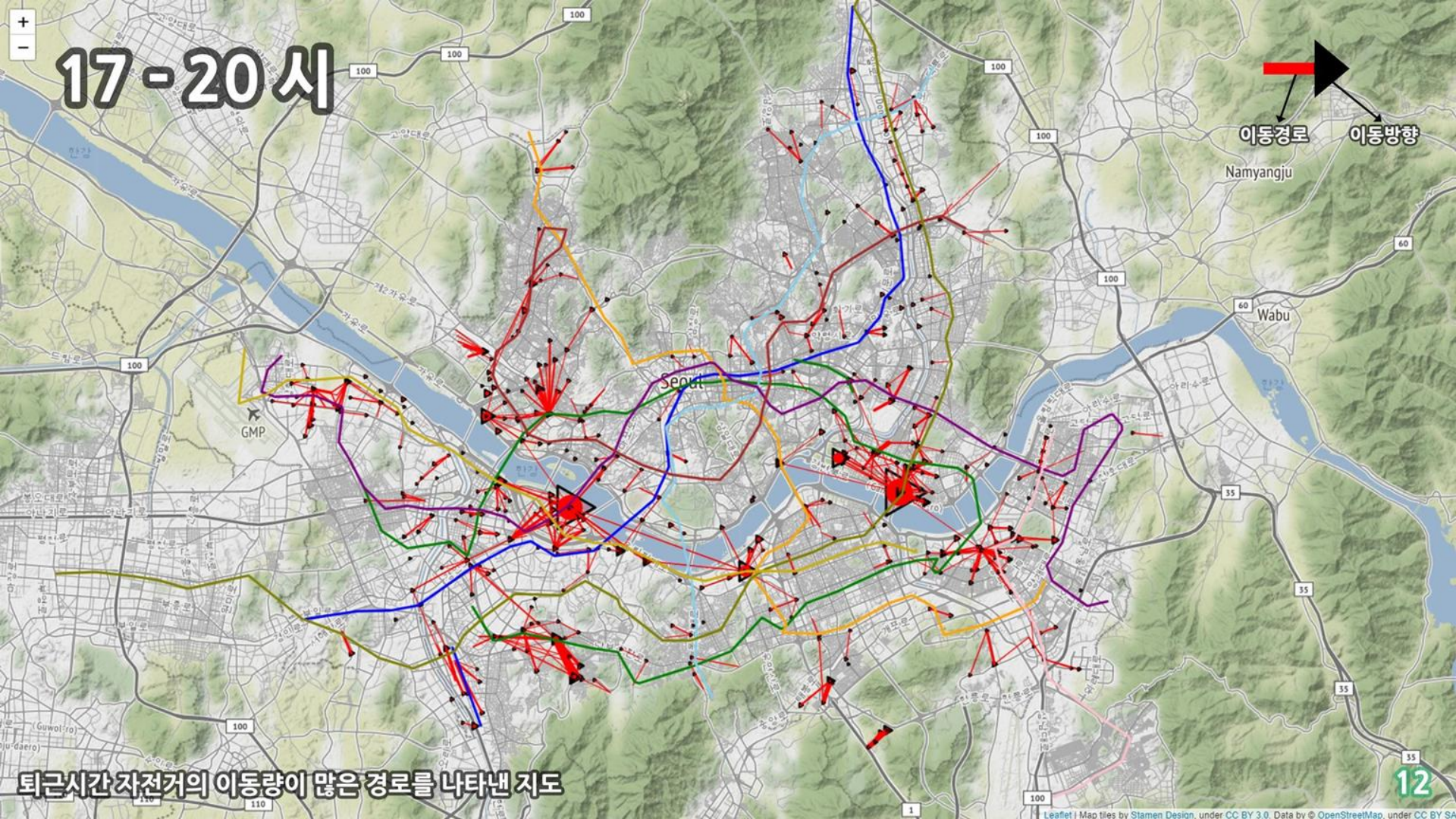
Wabu

Seoul

GMP

12

퇴근시간 자전거의 이동량이 많은 경로를 나타낸 지도



17 - 20 시

이동경로 이동방향

Namyangju

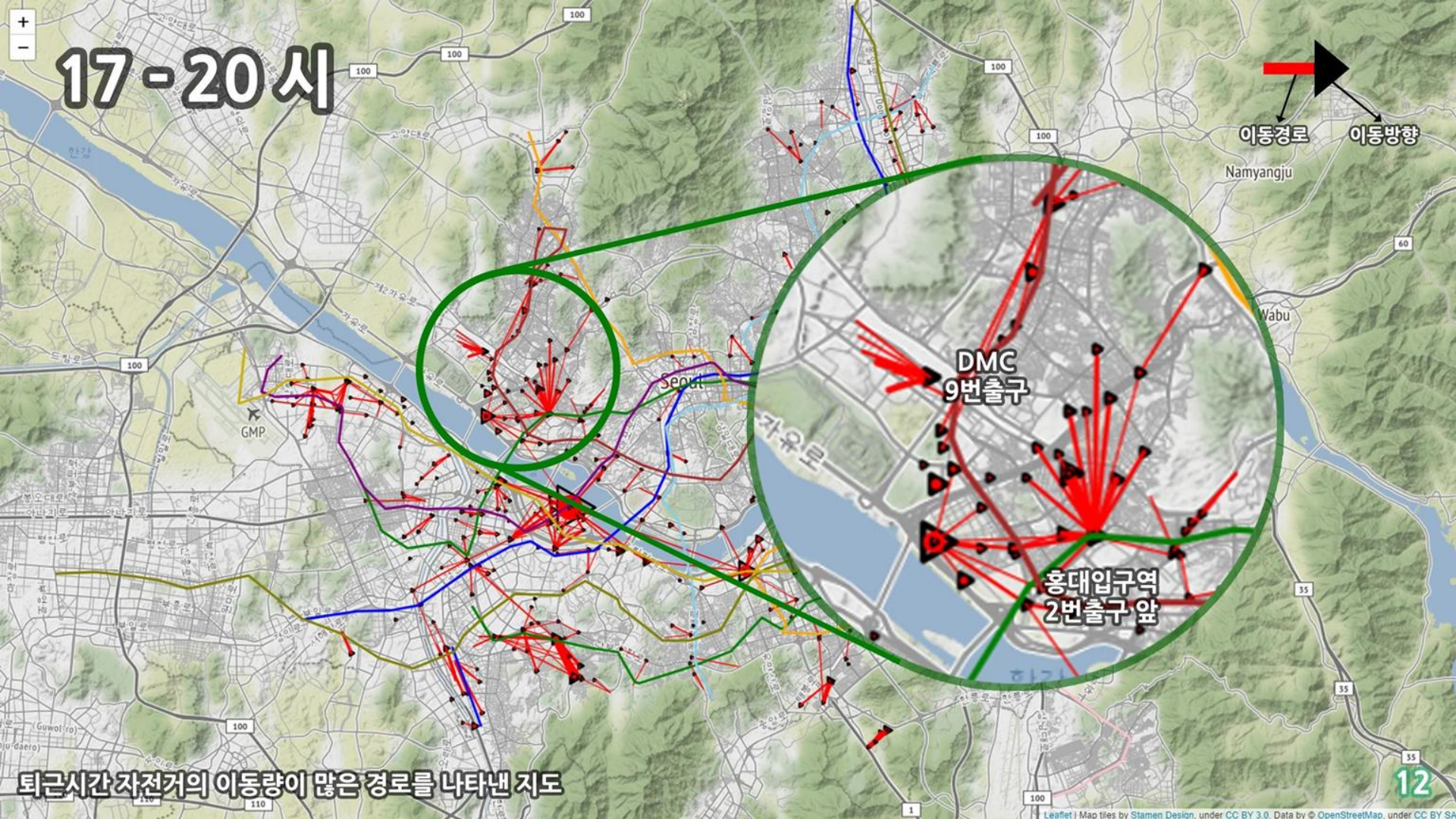
Wabu

DMC
9번출구

홍대입구역
2번출구 앞

12

퇴근시간 자전거의 이동량이 많은 경로를 나타낸 지도





현재 서울 공공자전거의 재배치 방법

- ✓ 전체적인 거치율을 80%로 유지
- ✓ 출근시간 전 보관소에 있는 자전거로 미리 정해놓은 특정 대여소에 배치
- ✓ 보관소에 있는 자전거로 수요가 충족되지 않을시 반납률이 많은 대여소에서 임의로 배치



- ✓ 출퇴근시간에 앞서 수요를 충족시킬 대여소의 리스트가 분석을 통해서 명확해져야함
- ✓ 랜덤한 대여소에서 가져오는게 아닌 더 효율적인 방법 필요
- ✓ 부족한 인력 대비 많은 대여소에 수요를 충족시키기 위해서는 최적화된 경로 필요

II. 재배치 분석 - 출근시간



- ✓ 상위 대여 지역구는 영등포구, 마포구, 강서구, 관악구, 동작구순으로 나타나며 상위 반납 지역구는 마포구, 구로구, 송파구, 영등포구, 성동구 순으로 나타난다. 위치는 주로 지하철역에 분포

번호	대여대여소명	횟수
1	대방역 6번출구	3,165(회)
2	여의나루역 1번출구 앞	2,669(회)
3	DMC역 9번출구 앞	2,097(회)
4	마곡나루역 5번출구 뒤편	2,064(회)
5	동방1교	1,995(회)

출근 시간 대여 장소 TOP 5

II. 재배치 분석 - 최근 3개월 상위 대여소 비교 (출근)

19년 3월

번호	대여대여소명	횟수
1	대방역6번출구	748(회)
2	DMC역 9번 출구 앞	588(회)
3	여의나루역 1번출구 옆	533(회)
4	여의도역 1번 출구 앞	527(회)
5	성수역 2번출구 앞	453(회)
6	도림천 신화교	440(회)
7	하늘채코오롱아파트 건너편	426(회)
8	광진광장 교통섬	408(회)
9	동방1교	399(회)
10	구로디지털단지역 앞	382(회)

19년 4월

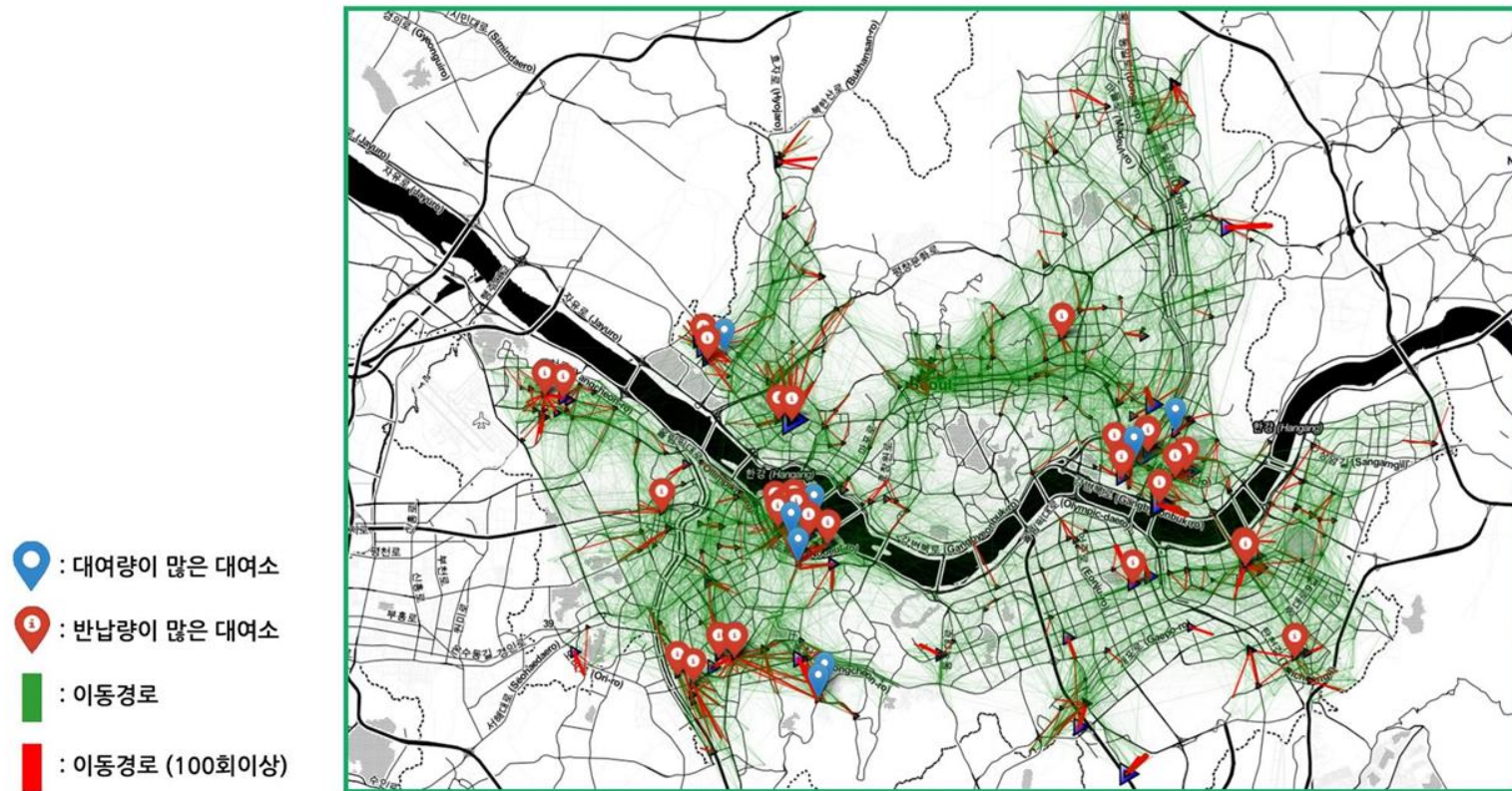
번호	대여대여소명	횟수
1	대방역6번출구	1,072(회)
2	여의나루역 1번출구 앞	878(회)
3	DMC역 9번출구 앞	682(회)
4	여의도역 1번출구 옆	676(회)
5	마곡나루역 5번출구 뒤편	650(회)
6	동방1교	646(회)
7	도림천 신화교	625(회)
8	구로디지털단지역 앞	591(회)
9	성수역 2번출구 앞	587(회)
10	세종사이버대학교	523(회)

19년 5월

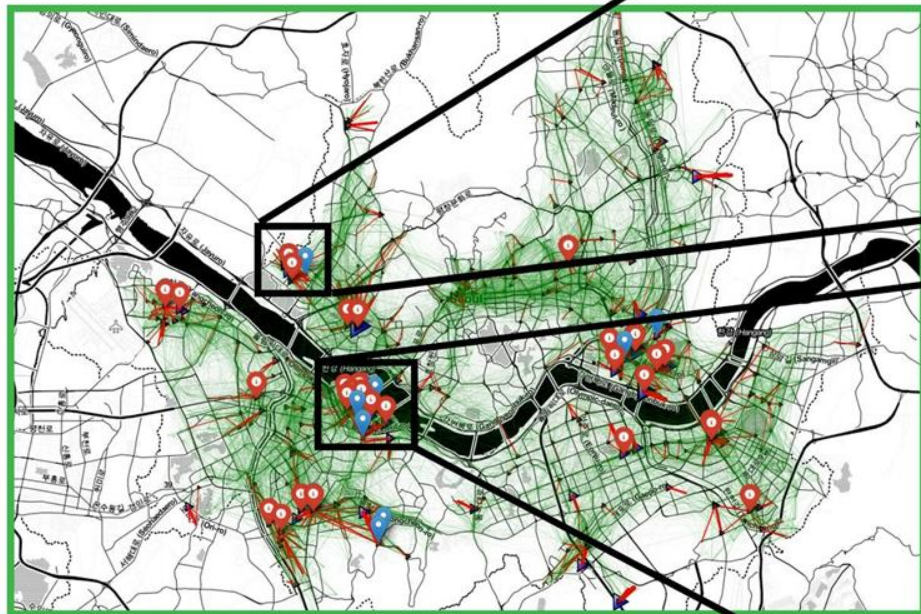
번호	대여대여소명	횟수
1	대방역6번출구	1,345(회)
2	여의나루역 1번출구 앞	1,258(회)
3	마곡나루역 5번출구 뒤편	1,092(회)
4	동방1교	950(회)
5	도림천 신화교	892(회)
6	구로디지털단지역 앞	836(회)
7	DMC역 9번출구 앞	827(회)
8	성수역 2번출구 앞	825(회)
9	여의도역 1번출구 옆	785(회)
10	양천향교역 7번출구 앞	751(회)

➡ 월별 상위 대여소를 비교해본 결과 대체로 일치함을 확인

II. 재배치 분석 - 현황 및 방향성 설정



II. 재배치 분석 - 현황 및 방향성 설정

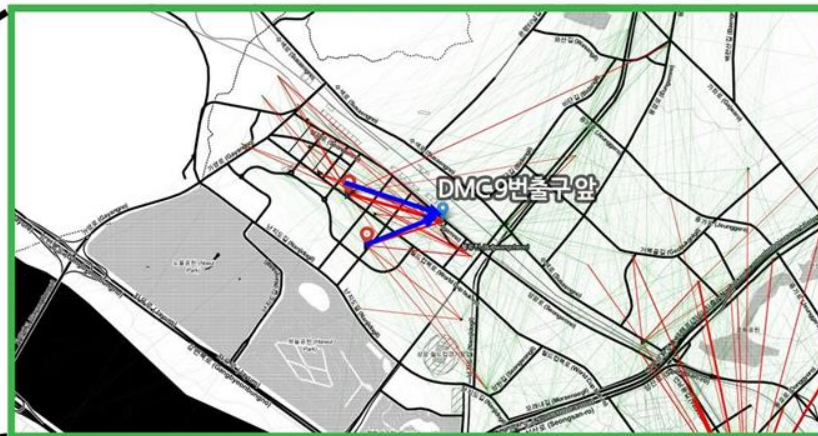


 : 대여량이 많은 대여소

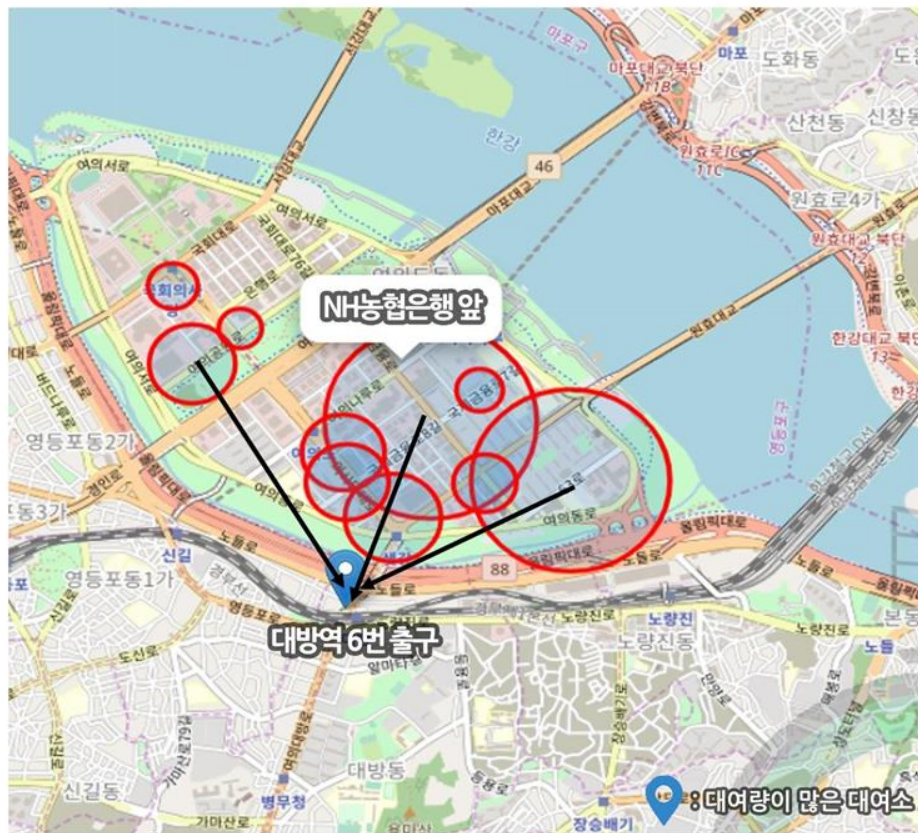
 : 반납량이 많은 대여소

 : 이동경로

 : 이동경로 (100회이상)



II. 재배치 분석 결과 - 출근시간



대방역6번 출구에서 반납하러 가는 대여소 수 상위 시각화

✓ 출근시간에 가장 많이 대여하는 대여소 : 대방역 6번출구

➔ 대방역 6번 출구에서 대여 후 반납 예정인 대여소의 자전거를 출근시간 전에 거치율을 50%로 낮추고 수요가 많은 대방역 6번 출구의 거치율을 200%로 높일 수 있도록 재배치

번호	대여소명	횟수
1	NH농협은행 앞	487
2	시범아파트버스정류장 옆	455
3	샛강역 1번출구 앞	222
4	KBS 앞	199
5	롯데캐슬엠피아어 옆	188
6	미성아파트 A동 앞	186
7	진주아파트상가 앞	142
8	국회의사당역 5번출구 옆	113
9	삼부아파트1동 앞	107
10	산업은행 앞	98

대방역6번 출구에서 반납하러 가는 대여소 수 상위 TOP10

II. 대방역 6번출구 재배치 경로 추천



추천경로 A
재배치시 우선순위



추천경로 B
추천경로 A로 수요가 충족이 되지 않을 시

II. 재배치 분석 결과 - 출근시간

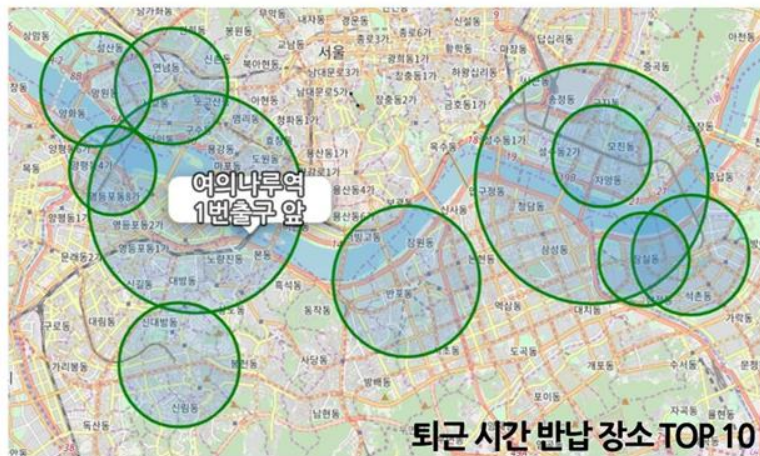
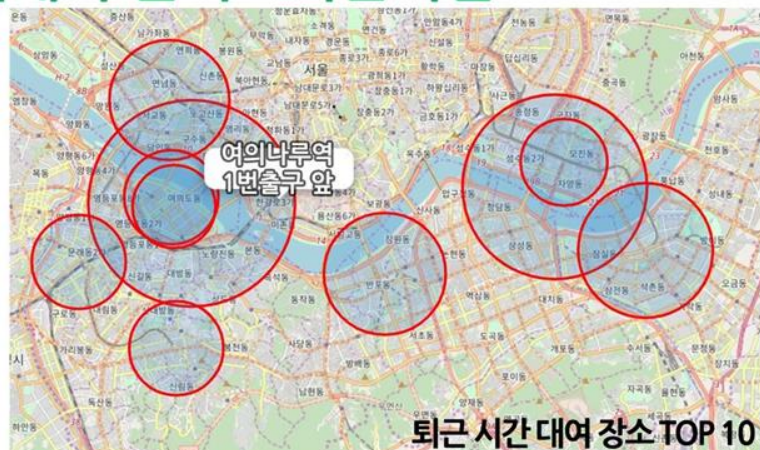


여의나루역 1번출구 앞에서 반납하러가는 상위대여소



DMC역 9번출구 앞에서 반납하러가는 상위대여소

II. 재배치 분석 - 퇴근시간



- ✓ 상위 대여 지역구는 영등포구, 광진구, 송파구, 서초구, 마포구 순으로 나타나며 상위 반납 지역구는 광진구, 영등포구, 서초구, 관악구, 송파구 순으로 나타남
위치는 주로 지하철역에 분포

번호	대여대여소명	횟수
1	여의나루역 1번출구 앞	10,168(회)
2	독섬유원지역 1번출구 앞	9,048(회)
3	롯데월드타워(잠실역2번출구 쪽)	6,628(회)
4	고속터미널역 8-1번, 8-2번 출구	5,982(회)
5	홍대입구역 2번 출구 앞	5,903(회)

퇴근 시간 대여 장소 TOP 5

II. 재배치 분석 - 최근 3개월 상위 대여소 비교 (퇴근)

19년 3월

번호	대여대여소명	횟수
1	여의나루역 1번출구 앞	1355(회)
2	롯데월드타워 (잠실역2번출구 쪽)	1163(회)
3	홍대입구역 2번출구 앞	1152(회)
4	독성유원지역 1번출구 앞	1099(회)
5	국민일보 앞	884(회)
6	봉림교 교통섬	734(회)
7	안암로터리 버스정류장 앞	734(회)
8	구로디지털단지역 앞	711(회)
9	고속터미널역 8-1번, 8-2번 출구 사이	705(회)
10	신도림역 1번 출구 앞	705(회)

19년 4월

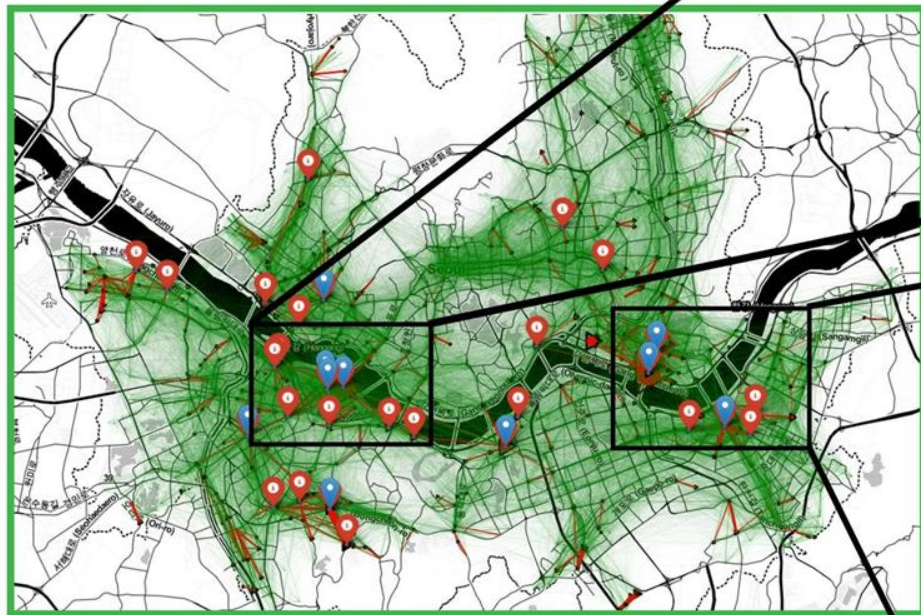
번호	대여대여소명	횟수
1	여의나루역 1번출구 앞	3,366(회)
2	독성유원지역 1번출구 앞	2,736(회)
3	롯데월드타워 (잠실역2번출구 쪽)	2,089(회)
4	고속터미널역 8-1번, 8-2번 출구 사이	1,871(회)
5	홍대입구역 2번출구 앞	1,793(회)
6	국민일보 앞	1,603(회)
7	신도림역 1번 출구 앞	1,519(회)
8	IFC몰	1,469(회)
9	봉림교 교통섬	1,396(회)
10	잠실역 8번출구	1,348(회)

19년 5월

번호	대여대여소명	횟수
1	여의나루역 1번출구 앞	5,447(회)
2	독성유원지역 1번출구 앞	5,213(회)
3	고속터미널역 8-1번, 8-2번 출구 사이	3,406(회)
4	롯데월드타워 (잠실역2번출구쪽)	3,376(회)
5	홍대입구역 2번출구 앞	2,958(회)
6	봉림교 교통섬	2,461(회)
7	건대입구역 사거리(롯데백화점)	2,445(회)
8	신도림역 1번 출구 앞	2,397(회)
9	국민일보 앞	2,120(회)
10	마포구민체육센터 앞	2,088(회)

➡ 월별 상위 대여소를 비교해본 결과 대체로 일치함을 확인

II. 재배치 분석 - 현황 및 방향성 설정



: 대여량이 많은 대여소



: 반납량이 많은 대여소



: 이동경로



: 이동경로 (100회이상)



II. 뚝섬유원지 1번출구앞 재배치 경로 추천



추천경로 A

재배치 우선순위 1



추천경로 B

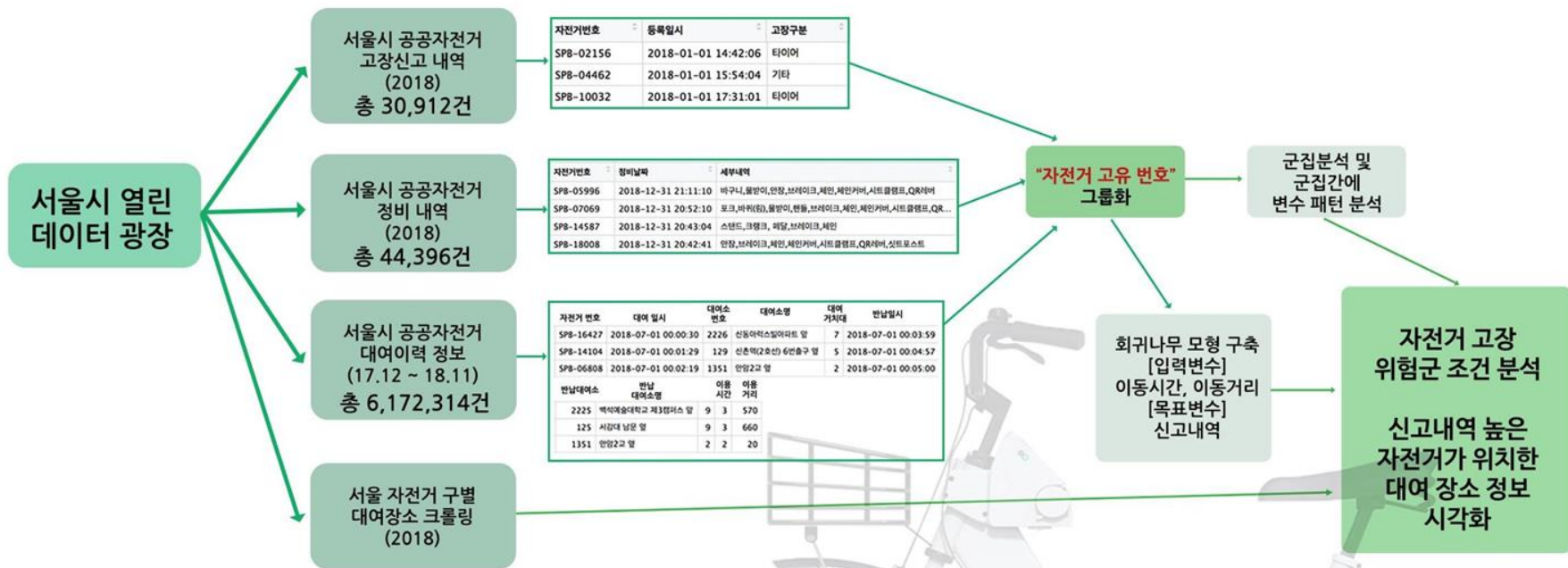
재배치 우선순위 2



추천경로 C

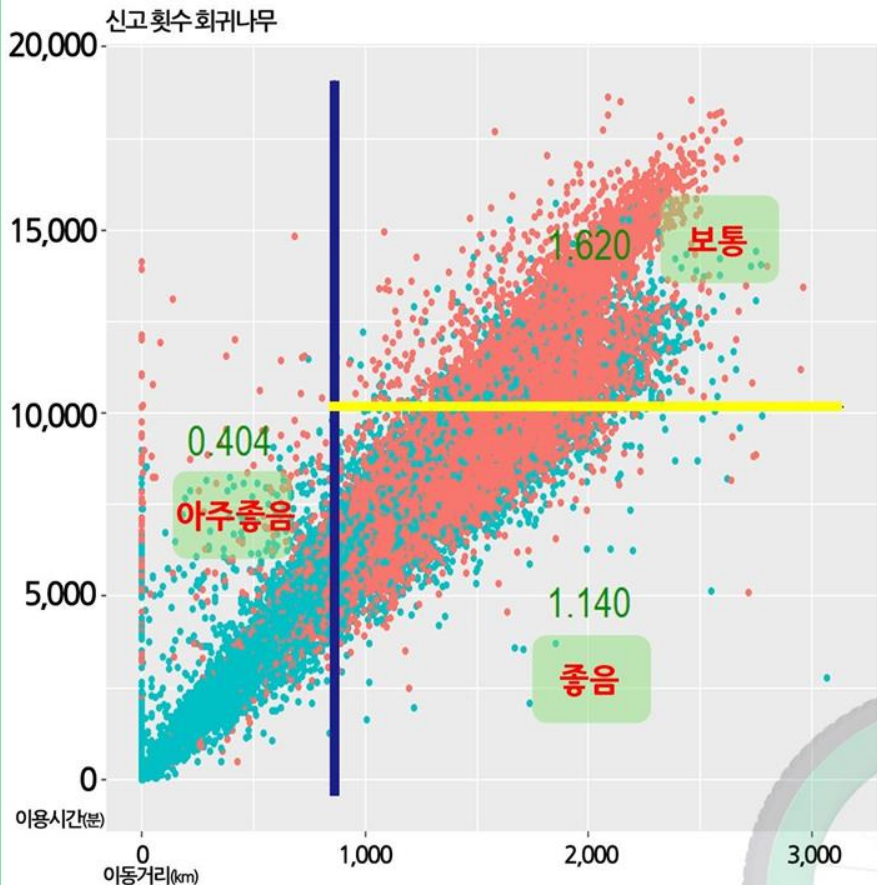
추천경로 A와 B로
수요가 충족되지 않을 경우

II. 고장데이터 분석



✓ 고장신고 및 정비 내역 데이터를 이용해 자전거를 군집화한 후 분석 결과에 따른 자전거 등급 적용

II. 고장데이터 분석



입력변수 : “이동거리”, “이용시간”
목표변수 : “신고횟수”

회귀 나무 모형 구축

- 모형구축 기계학습을 통해 2가지 기준선으로 분류결과가 나타났다.
- 첫번째 기준 : 862 km 이동거리로 보라색선
- 두번째 기준 : 10,172분 이용시간으로 노란색선

회귀 나무 분석 결과

- 이동시간과 이동거리가 높을 수록 신고횟수가 높을 것이라고 예측되었다.

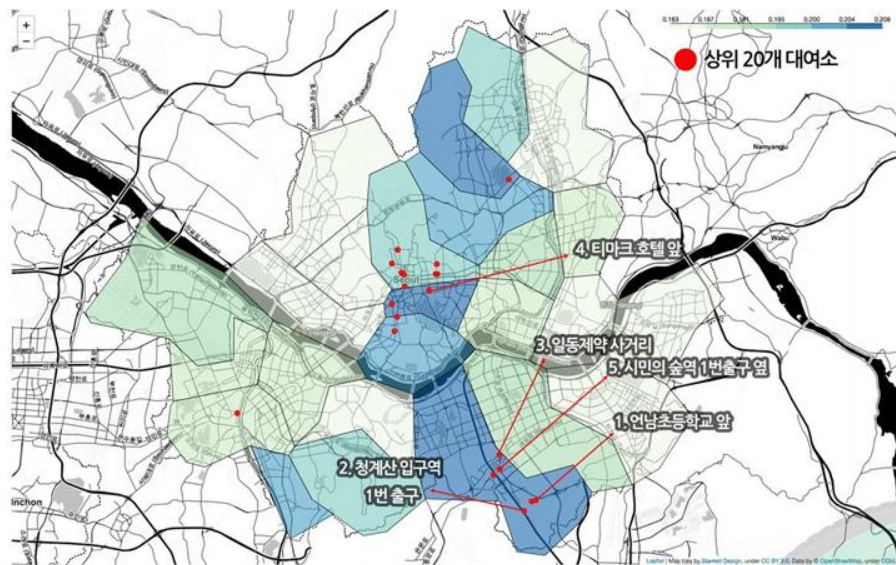
종합 결과

- 회귀나무 예측 분류 모형과 K-means cluster 군집의 결과가 비슷한 패턴을 나타냈다. 2가지 모형 바탕으로 **아주 좋음**, **중음**, **보통**으로 정의를 했다.

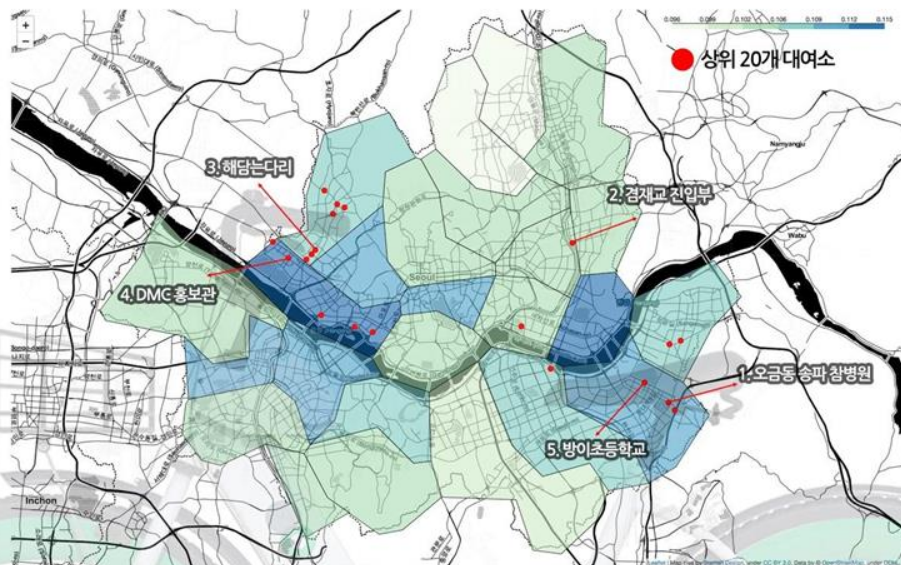
II. 데이터 분석 및 시각화 - 자전거 지역분포

“아주 좋음” 비율: “아주 좋음” 자전거 반납 수 / 전체 반납수
 “보통” 비율: “보통” 자전거 반납 수 / 전체 반납수

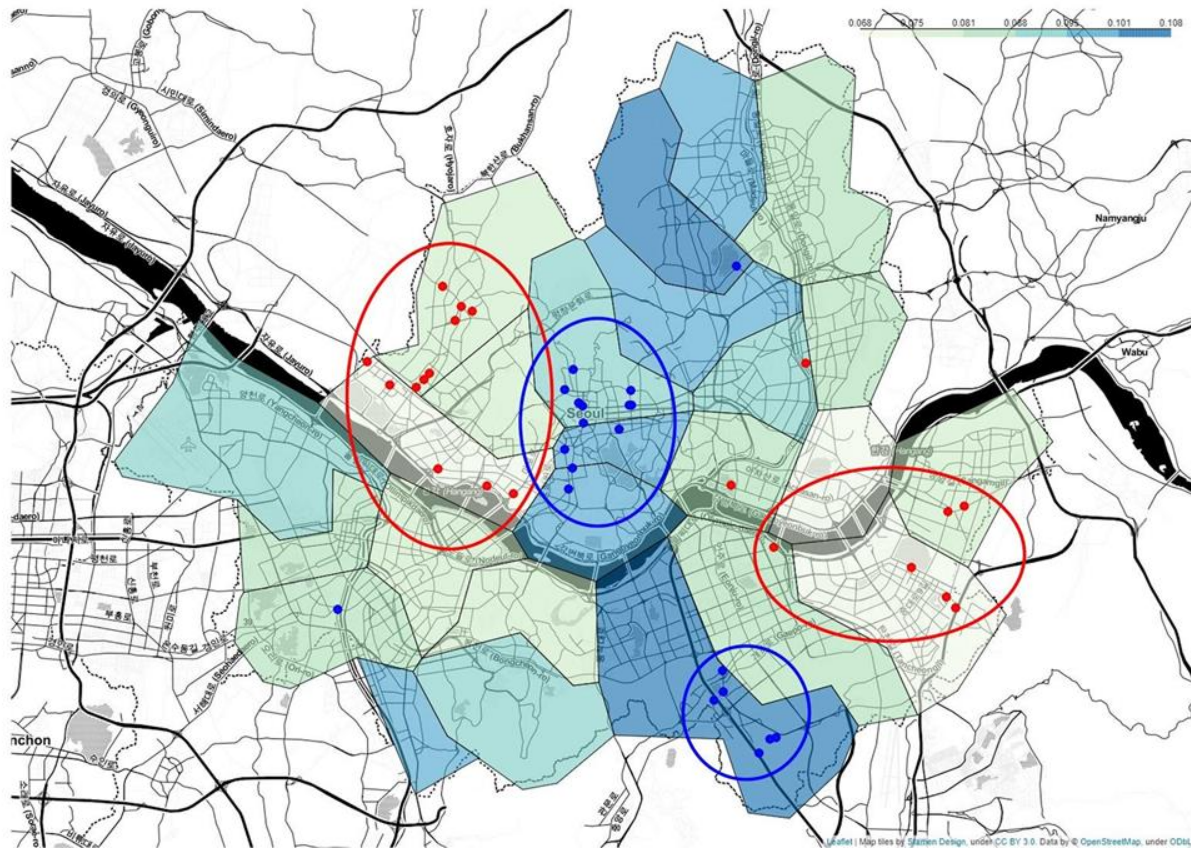
“아주 좋음” 자전거 분포



“보통” 자전거 분포



II. 고장데이터 분석 결과



TOP 20

- - “아주 좋음” 자전거의 반납 대여소
- - “보통” 자전거의 반납 대여소

✓ 해결책

➡ **“아주좋은”** 자전거와 **“보통”** 자전거를 적절히 교체하여 전체 자전거의 이용시간과 이동거리를 평준화시킨다.

IV. 프로젝트 결론



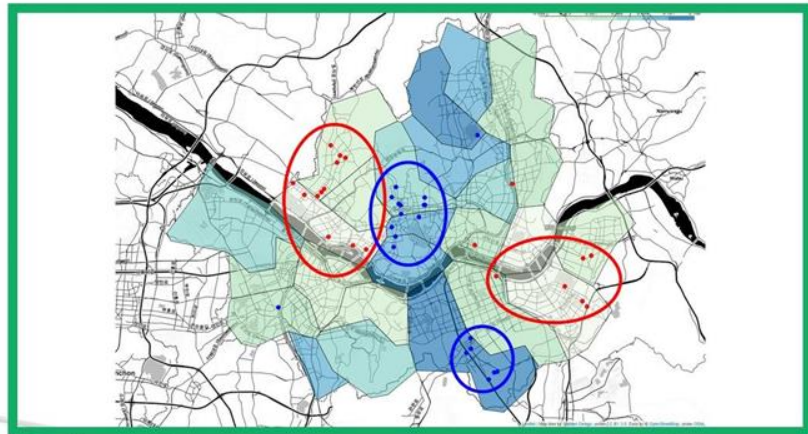
Ⅳ. 프로젝트 결론

재배치 분석



➔ 출퇴근 시간대에 자전거가 대여 후 어느 대여소로 반납이 이루어지는지 미리 알고 있으면, 반납이 이루어지는 대여소의 거치울을 낮춰도 자연스럽게 공급이 될 것이기 때문에 전체적인 거치울은 안정적으로 유지될 것이다.

고장데이터 분석



➔ 고장 날 확률이 많은 자전거가 밀집된 장소에 있는 자전거들과 고장 날 확률이 비교적 적은 자전거가 밀집된 장소에 있는 자전거들을 적절히 교체시킨다면 자전거의 기대 수명을 높일 수 있을 것이다.

IV. 프로젝트 결론 - 활용방안 및 한계점

기대효과



이용률이 높은 시간과 장소를 파악하여 물리는 수요를 충족 시킴으로써 이용 만족도 증가



효율적인 운영 재배치 방법을 제시하여 재배치 시간 단축



특정 자전거에 집중된 이용량을 분산시켜 자전거의 기대수명을 늘림으로써 관리 비용 절감

한계점 및 아쉬운점

- ➔ 거치율이 0%인 대여소의 초과 수요량에 대한 데이터 필요
- ➔ 재배치 동선 제안에서 교통상황을 반영하지 못함
- ➔ 자전거에 대한 데이터 부족 (자전거ID, 도입날짜, 폐기날짜)

IV. 참고문헌

출처	데이터명	기준연도	건수
서울열린데이터 광장	서울시 공공자전거 고장신고 내역	2018	30,912(건)
서울열린데이터 광장	서울특별시 공공자전거 대여이력 정보	2017.12 ~ 2018.11	6,172,314(건)
서울열린데이터 광장	서울시 공공자전거 실시간 대여정보(오픈 API)	2019	
서울자전거 따릉이·무인대여시스템	서울시 자전거 대여소 정보	2018.06~2019.05	13,004,020(건)

뉴스기사 인용

<https://www.sedaily.com/NewsView/1S8GNLEJSE>

http://www.skydaily.com/news/news_view.html?ID=91837

관련 논문

Russell V Lenth (2001) Some Practical Guidelines for Effective Sample Size Determination, The American Statistician, 55:3, 187-193, DOI: 10.1198/000313001317098149

분석도구



THANK YOU

