



WIT ZO LIGH
DRIE



강지원 이정현 최재영 박서현 이동욱

목 차



1. 주제 선정



1. 따릉이…?



이름: 따릉이

나이: 6세

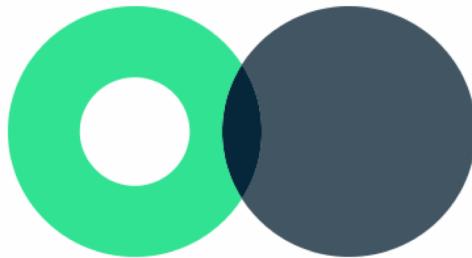
주소: 서울시

직업: 서울시 공영 자전거

탄소 배출 감량을 통한 녹색 성장 프로젝트의 일환으로
2014년, 서울시에서 처음 시행된
무인 공영자전거 대여 시스템



1. 따릉이…?



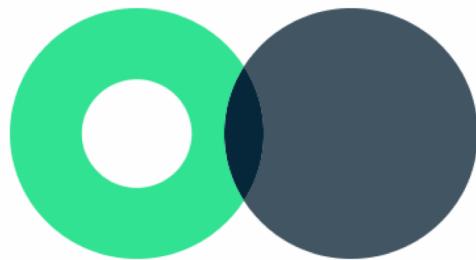
서울자전거
SEOUL BIKE 따릉이



- 2014년에 서울시에서 시범 운영
- 2015년 10월부터 정식 운영
- 2015년 2천대의 따릉이로 시작
- 2019년 현재 2만대의 따릉이 배치



1. 따릉이…?



서울자전거
SEOUL BIKE 따릉이

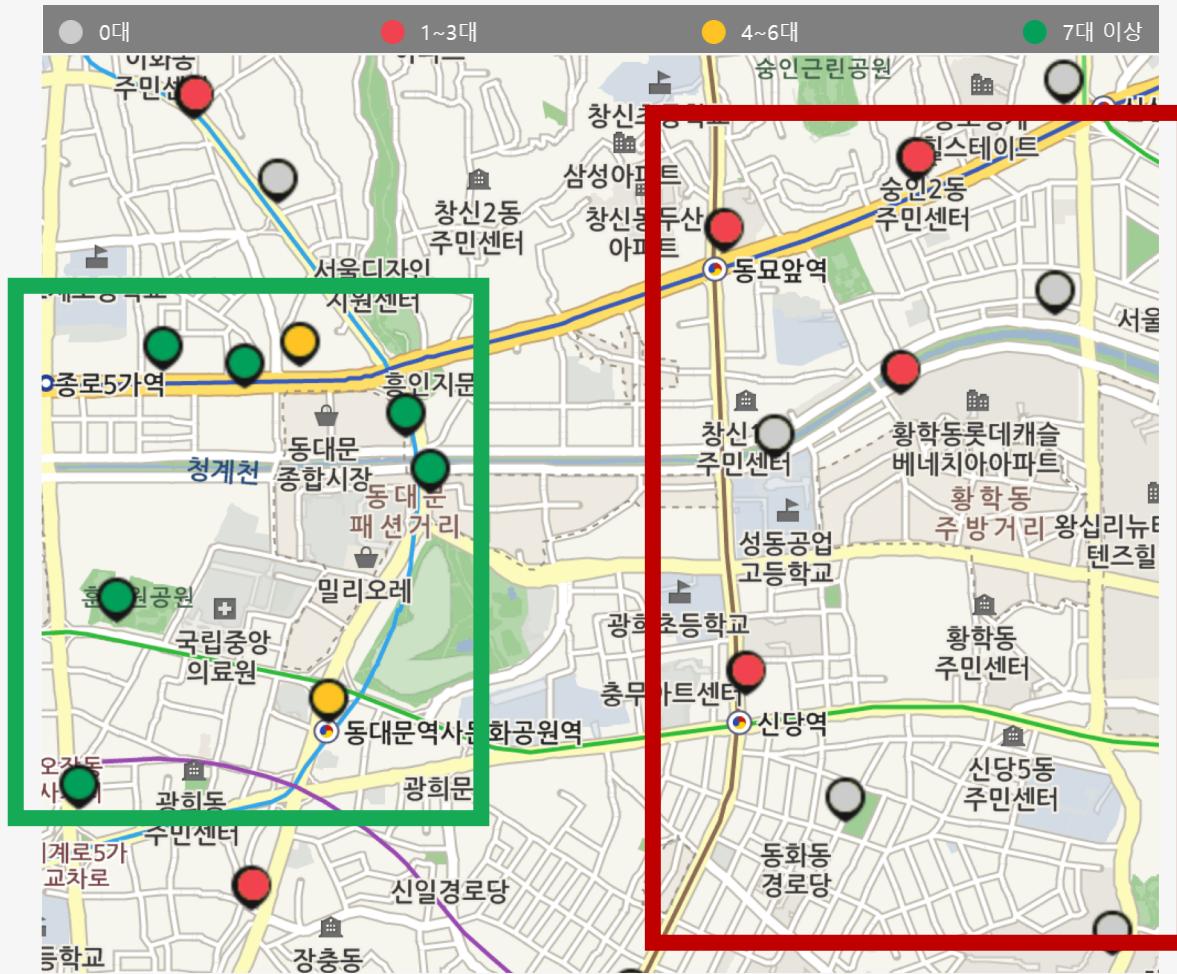


- 2014년에 서울시에서 시범 운영
- 2015년 10월부터 정식 운영
- 2015년 2천대의 따릉이로 시작
- 2019년 현재 2만대의 따릉이 배치

??



2. 따릉이 이용 실태



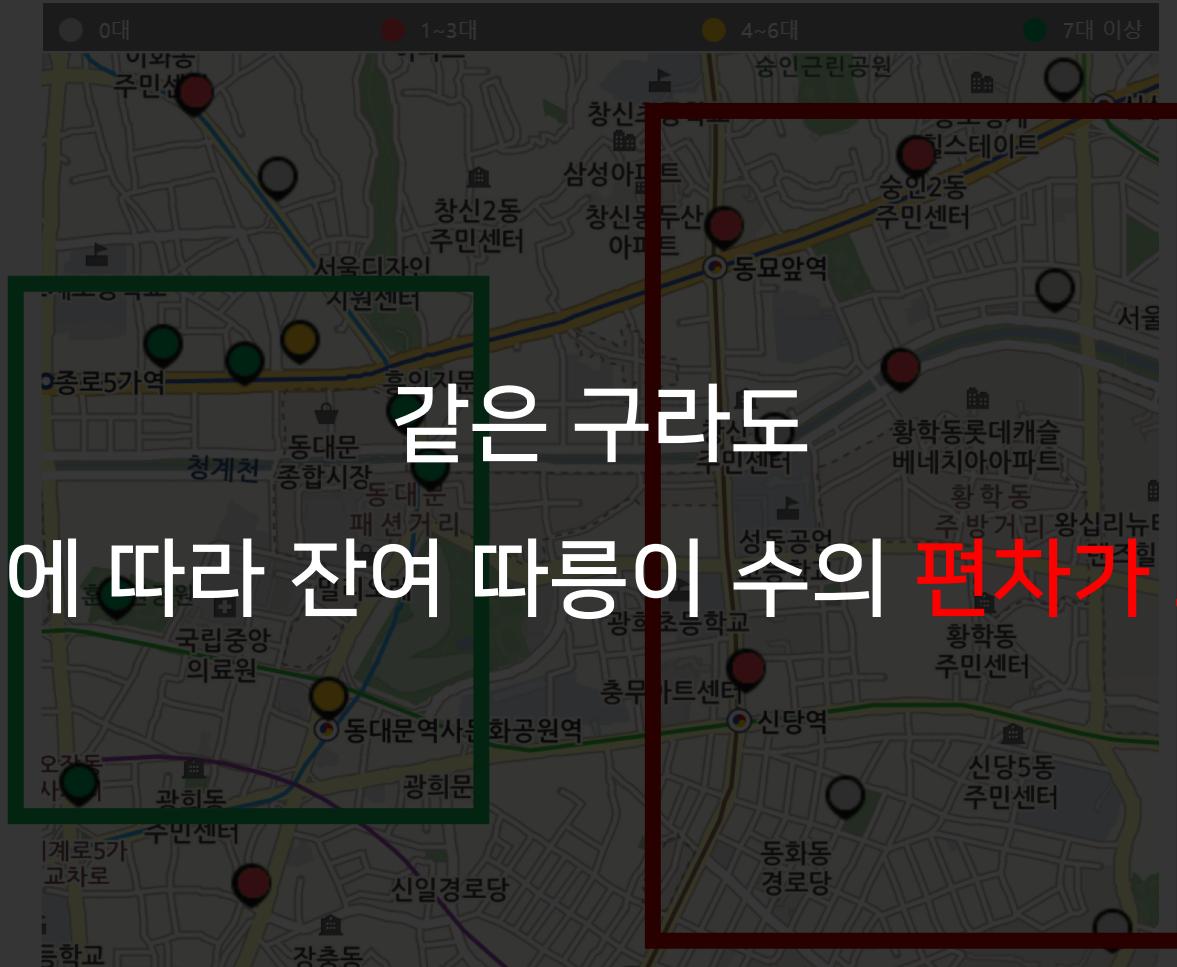
주제 선정

분석 방향

파생변수1

파생변수2

2. 따릉이 이용 실태



2. 따릉이 이용 실태

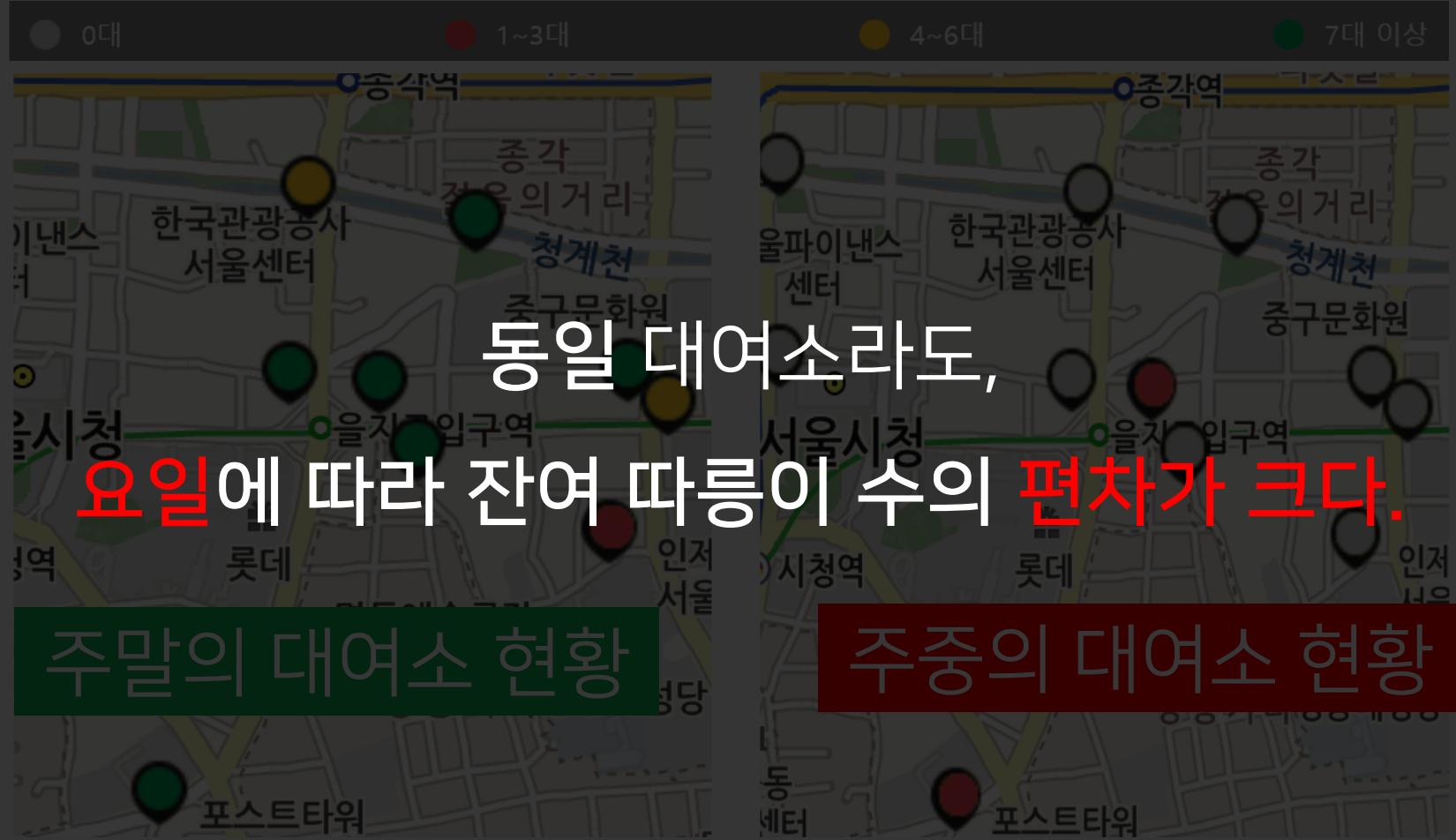


분석 방향

파생변수1

파생변수2
10

2. 따릉이 이용 실태



주제 선정

분석 방향

파생변수 1

파생변수2 11

2. 따릉이 이용 실태



사업소개

대여소 조회

이용권 구매

시민의견수렴

강동구 암사도서관앞에 설치해주세요

승실대와 상도역사이에 대여소가 하나더 필요합니다.

구로 디지털단지 우체국 근처에 대여소가 있으면 좋겠습니다.

왕십리쪽에 자전거 더 설치해주시십시오

지속적인 대여소 설치에도 끊이질 않는 **자전거 부족 현상!**



주제 선정

분석 방향

파생변수1

파생변수2

2. 따릉이 이용 실태



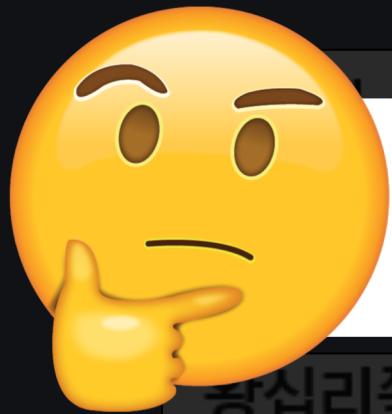
사업소개

대여소 조회

이용권 구매

시민의견수렴

강동구 암사도서관앞에 설치해주세요



현재 있는 대여소의 수요에 대한 공급을
맞추려면 언제 어디에 몇 대를 더 배치해야 될까?

광십리쪽에 자전거 더 설치해주십시오

지속적인 대여소 설치에도 끊이질 않는 추가 대여소 설치 요청!



주제 선정

분석 방향

파생변수1

파생변수2



일별 따릉이 대여 수요 예측

2. 분석 방향



1. 분석 방향

날짜 + 대여소

2019-01-01-101

2019-01-02-101

.

2019-03-30-993

2019-03-31-993



특정 대여소의,

일별 대여 건수를 예측하려면,

어떤 변수가 필요할까?

대여건수

14

45

.

3

14



1. 분석 방향

| 날짜 + 대여소 |
|----------------|
| 2019-01-01-101 |
| 2019-01-02-101 |
| . |
| . |
| 2019-03-30-993 |
| 2019-03-31-993 |

| 날짜 INFO | 대여소 INFO |
|---------|----------|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

| 대여건수 |
|------|
| 14 |
| 45 |
| . |
| . |
| 3 |
| 14 |



1. 분석 방향

2주차 FOCUS

| 날짜 + 대여소 |
|----------------|
| 2019-01-01-101 |
| 2019-01-02-101 |
| . |
| . |
| 2019-03-30-993 |
| 2019-03-31-993 |

| 날짜 INFO | 대여소 INFO |
|---------|----------|
| . | . |
| . | . |
| . | . |
| . | . |
| . | . |

| 대여건수 |
|------|
| 14 |
| 45 |
| . |
| . |
| 3 |
| 14 |



1. 분석 방향

| A | B | C | D | E | F | G | H | I | J | K | L |
|----|--------------|--------|---------|----------------|----------|------|---------|--------|--------|-------|-----------|
| 1 | '대여일자' | '대여시간' | '대여소번호' | '대여소' | '대여구분코드' | '성별' | '연령대코드' | '이용건수' | '운동량' | '탄소량' | '이동거리(M)' |
| 2 | '2017-08-01' | '00' | '538' | '답십리역 8번출구 앞' | '정기' | 'F' | '~10대' | 1 | 34.06 | 0.42 | 1830 |
| 3 | '2017-08-01' | '00' | '1006' | '롯데캐슬 115동앞' | '정기' | 'F' | '~10대' | 1 | 63.32 | 0.57 | 2460 |
| 4 | '2017-08-01' | '00' | '1323' | '월곡역 입구 육교 밑' | '정기' | 'F' | '~10대' | 1 | 161.37 | 1.89 | 8150 |
| 5 | '2017-08-01' | '00' | '1438' | '홈플러스 신내점 앞' | '정기' | 'F' | '~10대' | 1 | 29.11 | 0.24 | 1050 |
| 6 | '2017-08-01' | '00' | '115' | '사루비아 빌딩 앞' | '정기' | 'F' | '20대' | 1 | 13.9 | 0.18 | 780 |
| 7 | '2017-08-01' | '00' | '359' | '원남동사거리' | '정기' | 'F' | '20대' | 1 | 22.31 | 0.3 | 1310 |
| 8 | '2017-08-01' | '00' | '382' | '약수역 10번출구 앞' | '정기' | 'F' | '20대' | 1 | 38.02 | 0.37 | 1600 |
| 9 | '2017-08-01' | '00' | '114' | '홍대입구역 8번출구 앞' | '정기' | 'F' | '20대' | 1 | 0 | 0 | 0 |
| 10 | '2017-08-01' | '00' | '146' | '마포역 1번출구 뒤' | '정기' | 'F' | '20대' | 1 | 34.37 | 0.36 | 1550 |
| 11 | '2017-08-01' | '00' | '150' | '서강대역 2번출구 앞' | '정기' | 'F' | '20대' | 1 | 138.6 | 1.16 | 5000 |

대여건수

14

45

3

14

: Y(대여 건수)에 대한 양질의 변수가 다수 존재



성별 코드



대여 구분 코드



연령대 코드



파생변수1

파생변수2

2. 데이터 선별

자전거 별 데이터 .csv

- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_2분기_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_2분기_2.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_3분기_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_3분기_2.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_4분기_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2017년_4분기_2.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2018년_1분기_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2018년_2분기_1.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2018년_2분기_2.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2018년_2분기_3.csv
- [CSV] 서울특별시 공공자전거 대여이력 정보_2018년_2분기_4.csv
- [CSV] 서울특별시 공공자전거 대여정보_201807_01.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201807_02.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201808.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201809_1.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201809_2.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201810_02.xlsx
- [CSV] 서울특별시 공공자전거 대여정보_201811.xlsx

2018

대여소 별 데이터 .csv

- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_1.csv
- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_2.csv
- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_3.csv
- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_4.csv
- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_5.csv
- [CSV] 서울특별시 공공자전거 이용정보(시간대별)_2017년_6.csv

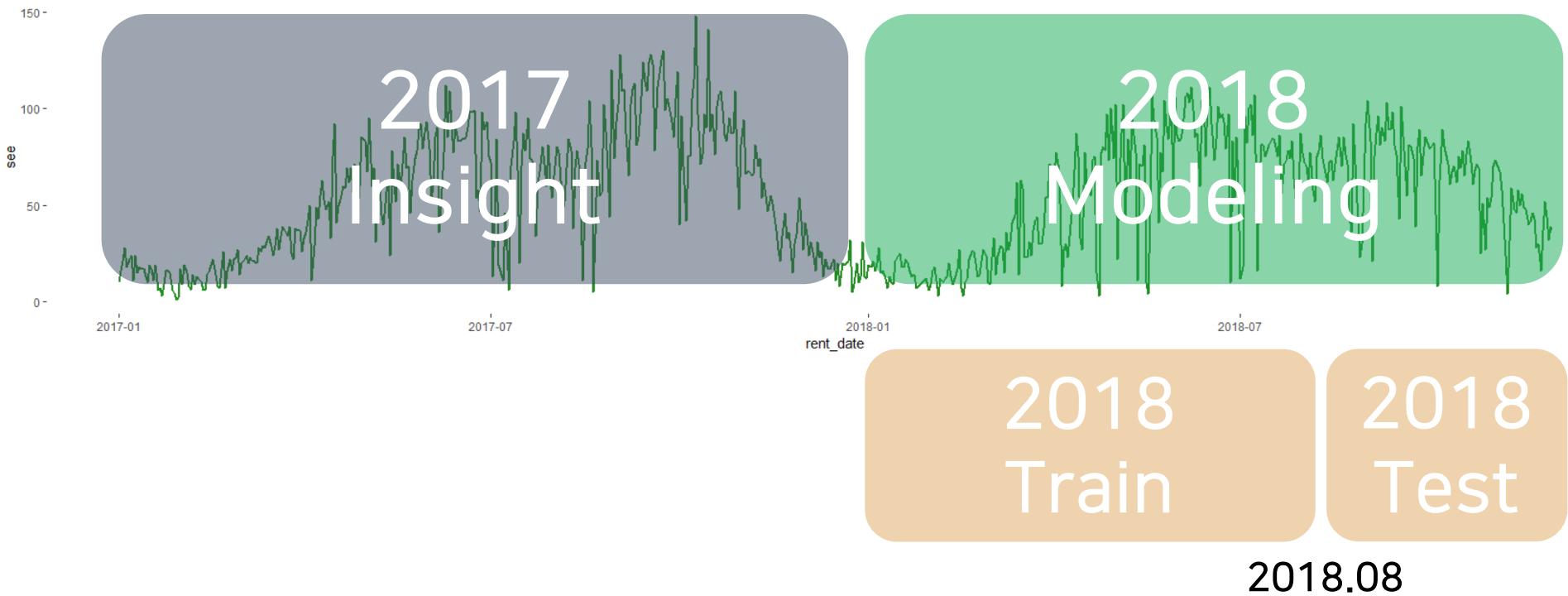
2017

- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_2018년_01_02.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_2018년_03_04.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201805.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201806.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201807.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201808.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201809.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201810.xlsx
- [XLSX] 서울특별시 공공자전거 시간대별 대여정보_201811.xlsx

2018

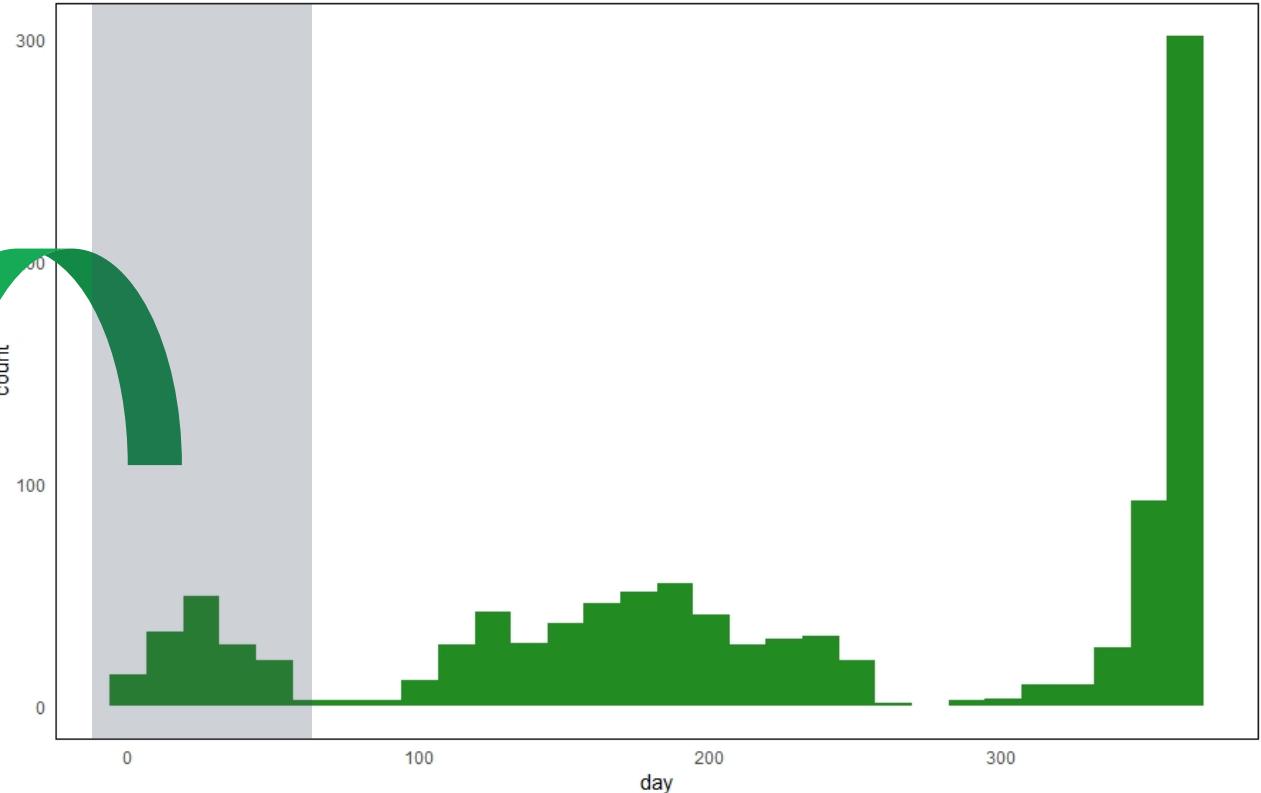


2. 데이터 선별



2. 데이터 선별

1년간 각 대여소 사용 일수



일별 대여 건수 정보가 충분하지 않은 대여소가 존재

(즉, 2017년 중에 새로 생긴 대여소)



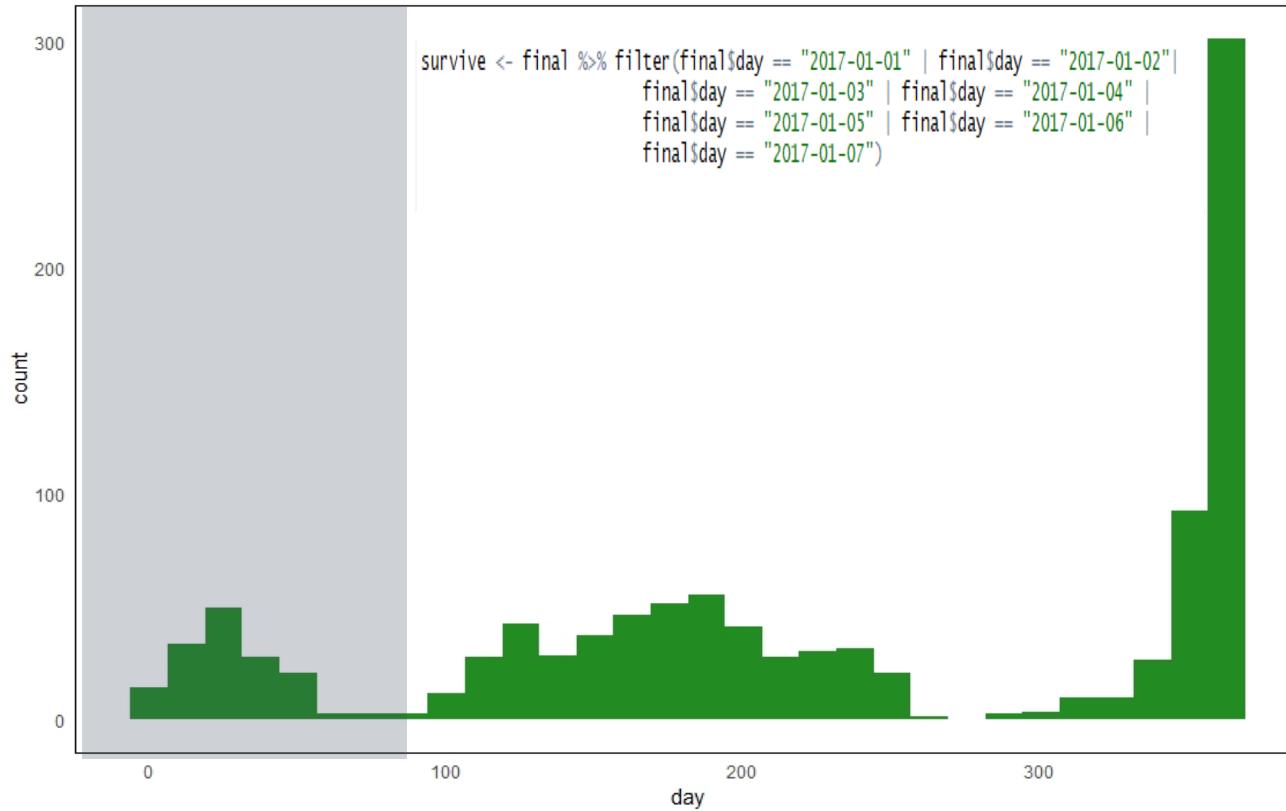
주제 선정

분석 방향

파생변수1

파생변수2

2. 데이터 선별



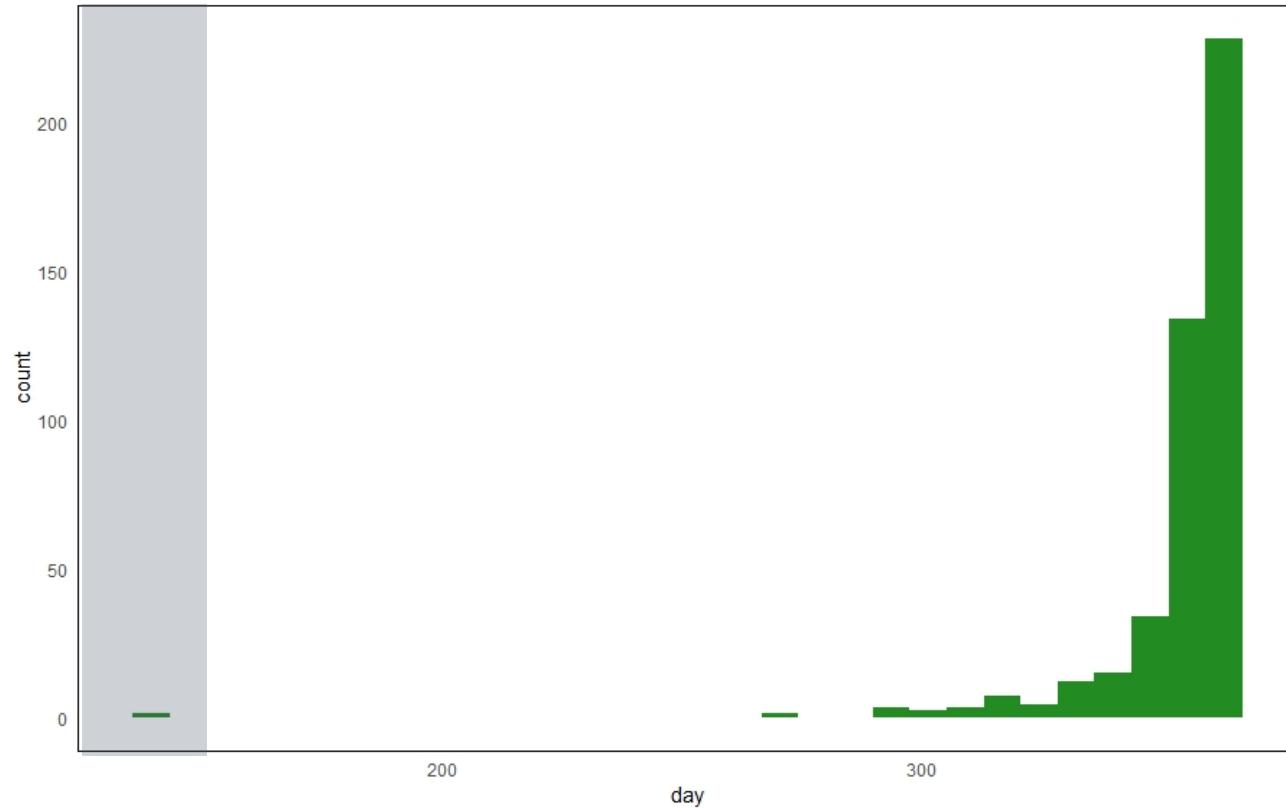
2017년 1월 첫째 주를 기준으로,



첫 주에 대여 건수 정보가 존재하지 않은 대여소 삭제



2. 데이터 선별



2017년 1월 첫째 주에는 사용되었지만,



이후 폐쇄된 대여소 삭제



2. 데이터 선별

```
> final[final$loc_num %in% final_survive,]  
  loc_num          loc_name  
1:    101  (구)합정동 주민센터  
2:    102      망원역 1번출구 앞  
3:    103      망원역 2번출구 앞  
4:    104      합정역 1번출구 앞  
5:    105      합정역 5번출구 앞  
---  
439:    803      한남초교 앞 보도육교  
440:    552      대림마크로리버 앞  
441:    564      금호역 3번출구  
442:    902      진관동 은빛초등학교  
443:    523      옥수동성당 옆
```

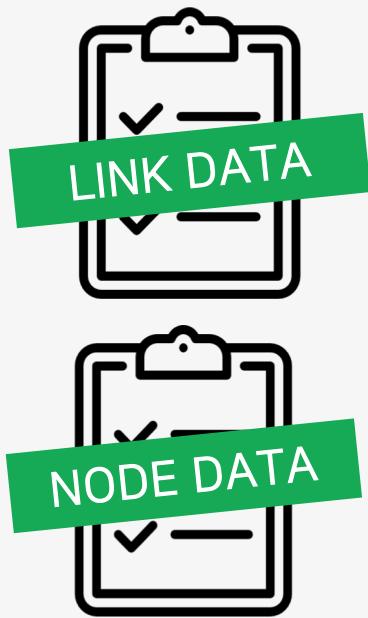


최종 443개의 대여소로 확정!



3. 이번 주 분석

2주차 FOCUS



* 2017년도 데이터

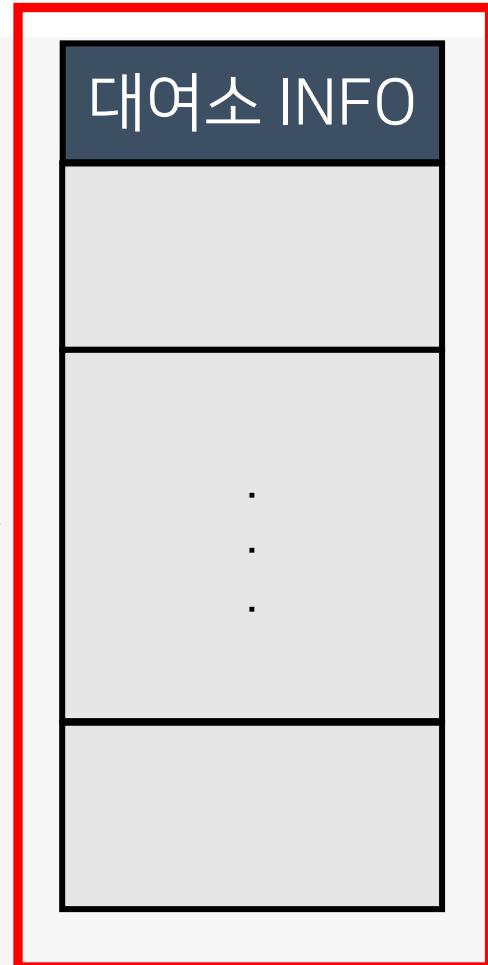


3. 이번 주 분석

2주차 FOCUS



PCA를 이용한
index 변수 생성



* 2017년도 데이터



3. 파생 변수1



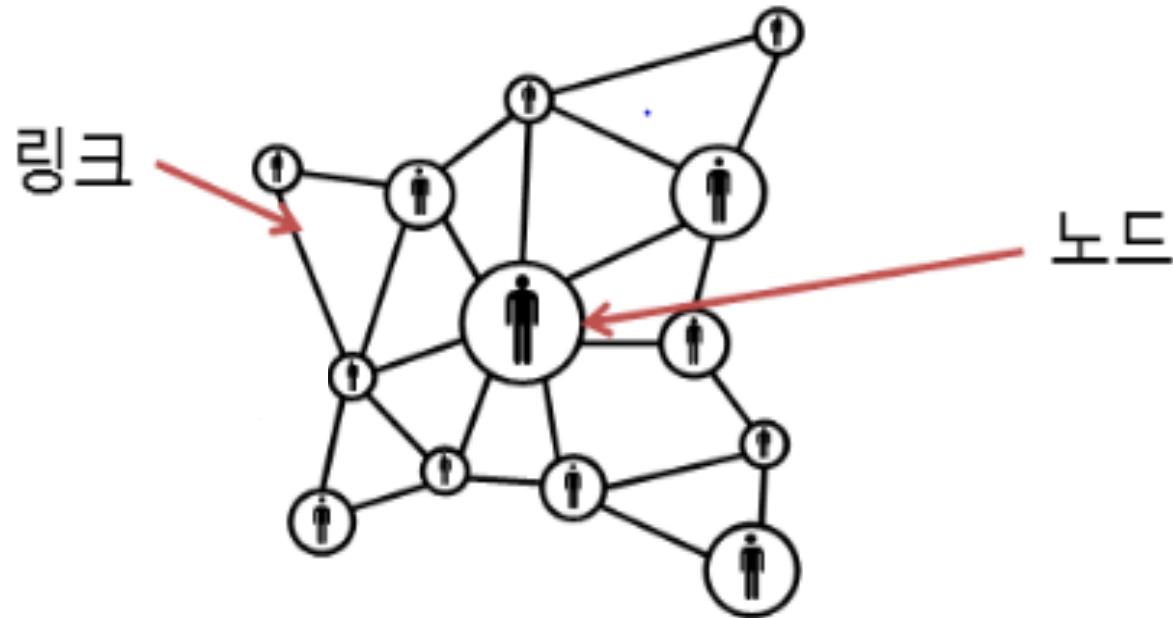
네트워크 분석을 통한 중심성 산출

PCA를 통한 지수 산출

네트워크 분석을 통한 중심성 산출

PCA를 통한 지수 산출

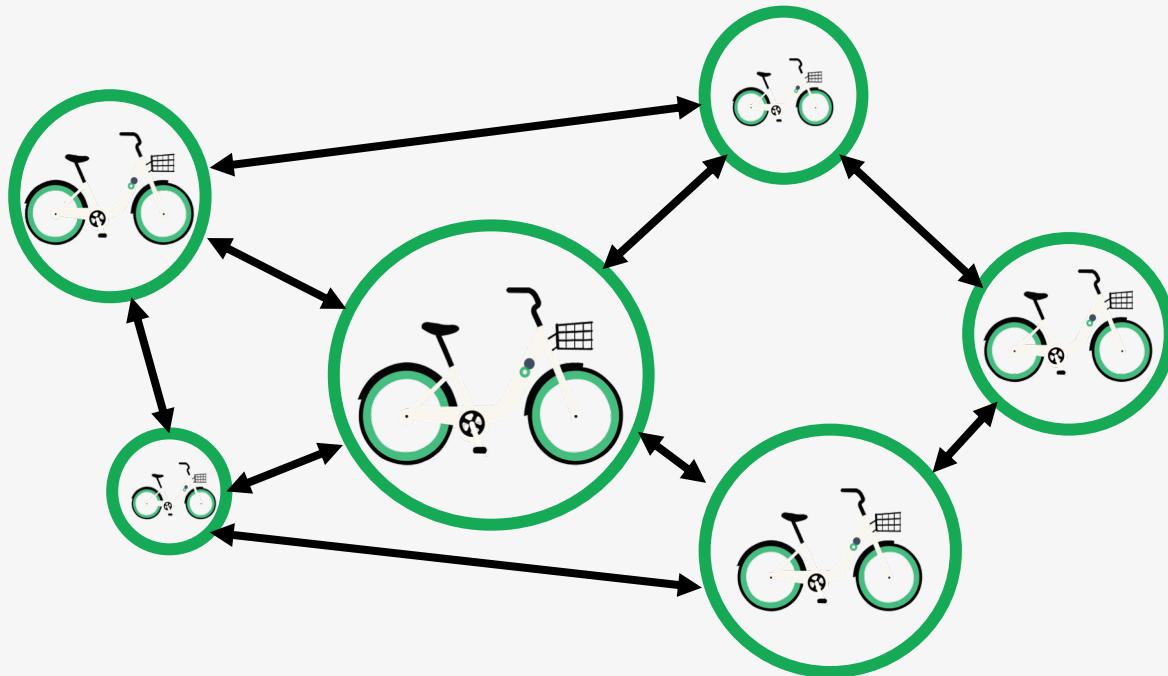
1. 네트워크 분석이란?



개인과 집단들 간의 관계를 노드와 링크로서 표현하여,

그것의 위상구조와 확산 및 진화과정을 계량적으로 분석하는 방법

2. 왜 네트워크 분석을 사용할까?



대여소별 따릉이의 이동량 파악 가능
교통 데이터를 다룰 때 빈번히 쓰임

3. 네트워크 분석 개념 알아보기

1. 노드

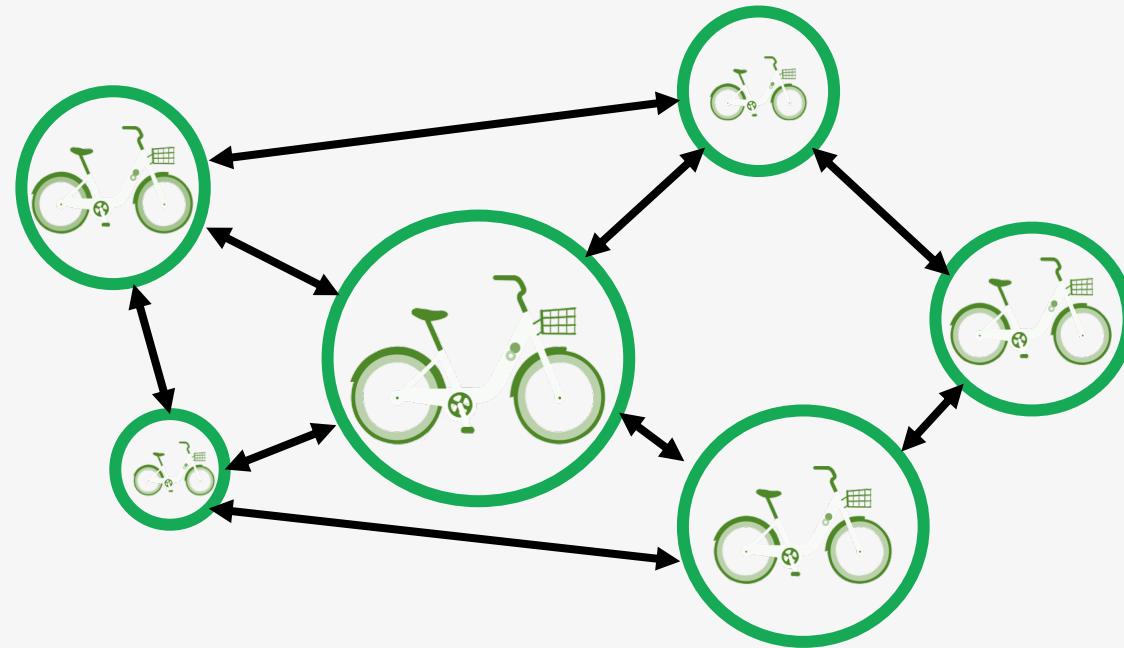
2. 링크

3. 중심성



3. 네트워크 분석 개념 알아보기

1. 노드

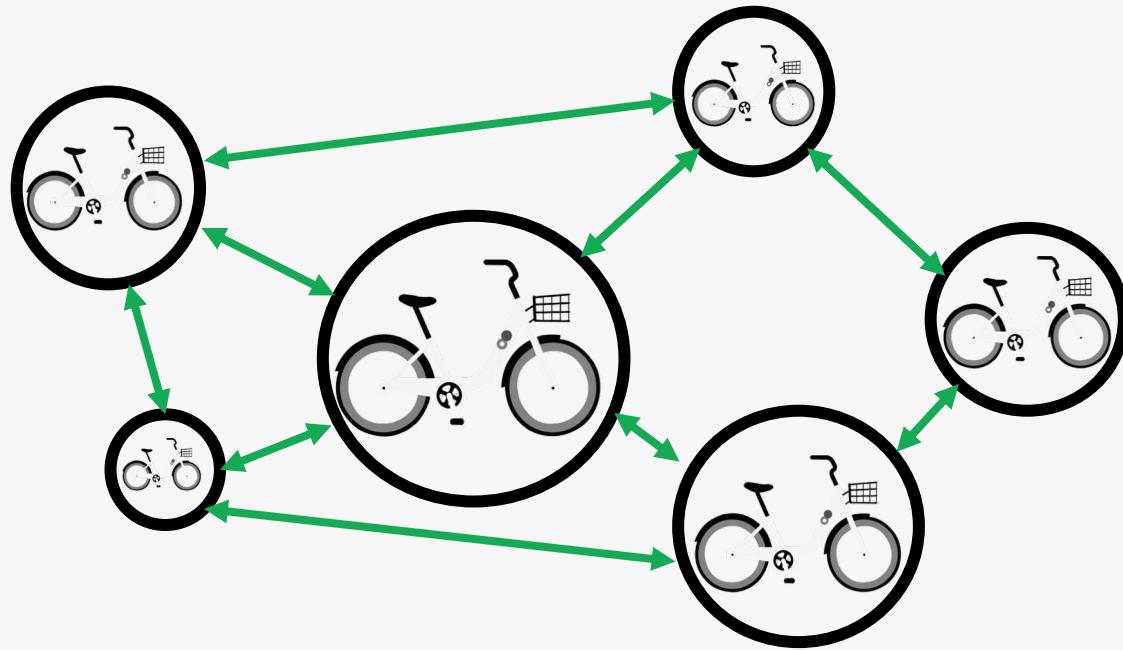


분석하고자 하는 객체들의 집합 = 각 대여소



3. 네트워크 분석 개념 알아보기

2. 링크

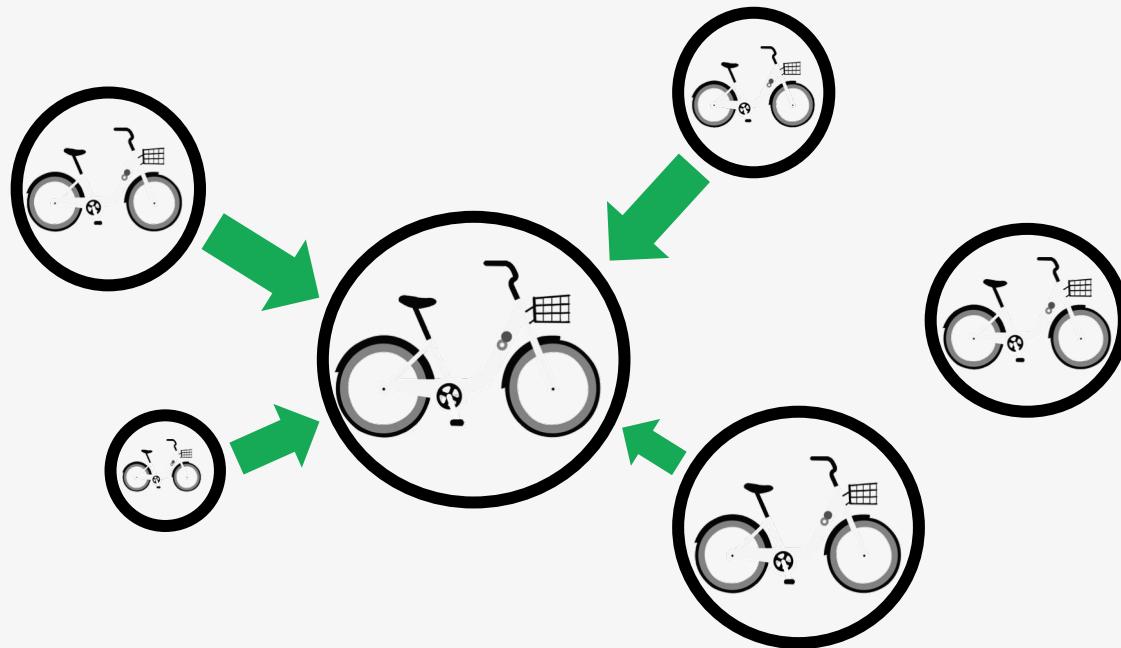


각 노드들의 관계= 따름이 이동량과 방향



3. 네트워크 분석 개념 알아보기

3. 중심성



각 노드들의 관계의 정도 =

(네트워크 상에서) 대여소들의 위치에 따른 상대적 중요성

3. 네트워크 분석 개념 알아보기

3. 중심성



연결정도 중심성 (Degree centrality)

링크의 연결정도를 기반으로 한 노드의 중심성



근접 중심성 (Closeness centrality)

각 노드 간의 거리를 기반으로 한 노드의 중심성



매개 중심성 (Betweenness centrality)

개별 노드들의 관계를 중개하는 정도를 기반으로 한 노드의 중심성



위세 중심성 (Eigenvector centrality)

연결된 노드의 중요성에 가중치를 두는 방식을 기반으로 한 노드 중심성



4. 파생변수1: 전처리

1. 노드 데이터

| 대여소 그룹 | 대여소 명 | 반납 일자 / 월 | 반납 건수 |
|--------|---------------------|------------|-------|
| 그룹명 없음 | 대여소명 없음 | 2017-01-01 | 10 |
| 광진구 | 500. 어린이대공원역 3번출구 앞 | 2017-01-01 | 12 |
| 광진구 | 501. 광진구의회 앞 | 2017-01-01 | 10 |
| 광진구 | 502. 뚝섬유원지역 1번출구 앞 | 2017-01-01 | 13 |
| 광진구 | 503. 더샵스타시티 C동 앞 | 2017-01-01 | 8 |
| 광진구 | 504. 신자초교입구교차로 | 2017-01-01 | 13 |
| 광진구 | 505. 자양사거리 광진아크로텔 앞 | 2017-01-01 | 10 |

| 대여소 그룹 | 대여소 명 | 대여 일자 / 월 | 대여 건수 |
|--------|---------------------|------------|-------|
| 그룹명 없음 | 대여소명 없음 | 2017-01-01 | 0 |
| 광진구 | 500. 어린이대공원역 3번출구 앞 | 2017-01-01 | 20 |
| 광진구 | 501. 광진구의회 앞 | 2017-01-01 | 10 |
| 광진구 | 502. 뚝섬유원지역 1번출구 앞 | 2017-01-01 | 9 |
| 광진구 | 503. 더샵스타시티 C동 앞 | 2017-01-01 | 9 |
| 광진구 | 504. 신자초교입구교차로 | 2017-01-01 | 6 |
| 광진구 | 505. 자양사거리 광진아크로텔 앞 | 2017-01-01 | 11 |
| 광진구 | 515. 광양중학교 앞 | 2017-01-01 | 0 |



4. 파생변수1: 전처리

1. 노드 데이터

| loc_num | rent_sum_tot | return_sum_tot | total_num_tot | loc_name |
|---------|--------------|----------------|---------------|-------------|
| 101 | 4728 | 5132 | 9860 | (구)합정동 주민센터 |
| 102 | 19707 | 18090 | 37797 | 망원역 1번출구 앞 |
| 103 | 13940 | 13684 | 27624 | 망원역 2번출구 앞 |

<최종 노드 데이터> ^{SUM}

| loc_num | total_num_tot | loc_name |
|---------|---------------|-------------|
| 101 | 9860 | (구)합정동 주민센터 |
| 102 | 37797 | 망원역 1번출구 앞 |
| 103 | 27624 | 망원역 2번출구 앞 |

노드 기준 : loc_num

노드 크기: total_num_tot



4. 파생변수1: 전처리

2. 링크 데이터

| '자전거번호' | '대여일시' | '대여대여소번호' | '대여대여소' | '대여거치대' | '반납일시' | '반납대여소번호' | '반납대여소' |
|-------------|-----------------------|-----------|---------------|---------|-----------------------|-----------|--------------------|
| 'SPB-00516' | '2017-04-01 00:00:37' | '259' | ' 대방역6번출구' | 8 | '2017-04-01 00:27:20' | '202' | ' 국민일보 앞' |
| 'SPB-02969' | '2017-04-01 00:00:42' | '516' | ' 광진메디칼 앞' | 5 | '2017-04-01 00:08:35' | '540' | ' 군자역 7번출구 베스트샵 앞' |
| 'SPB-01628' | '2017-04-01 00:00:43' | '316' | ' 종각역 1번출구 앞' | 11 | '2017-04-01 00:09:58' | '349' | ' 사직동주민센터' |
| 'SPB-00178' | '2017-04-01 00:01:59' | '233' | ' YP 센터 앞' | 7 | '2017-04-01 00:10:34' | '232' | ' 양평우림 이비즈센타 앞' |
| 'SPB-04518' | '2017-04-01 00:02:28' | '334' | ' 종로3가역 2번출구' | 7 | '2017-04-01 00:14:55' | '340' | ' 혜화동 로터리' |
| 'SPB-01681' | '2017-04-01 00:02:33' | '602' | ' 장안동 사거리' | 7 | '2017-04-01 00:13:02' | '604' | ' 답십리초등학교 옆 공원' |
| 'SPB-04584' | '2017-04-01 00:03:17' | '145' | ' 공덕역 5번출구' | 2 | '2017-04-01 00:12:01' | '142' | ' 아현역 4번출구 앞' |
| 'SPB-02268' | '2017-04-01 00:03:24' | '573' | ' 구의문주차장 앞' | 7 | '2017-04-01 00:14:36' | '554' | ' 아차산역 사거리' |
| 'SPB-03699' | '2017-04-01 00:03:43' | '361' | ' 동묘앞역 1번출구 두 | 1 | '2017-04-01 00:09:23' | '344' | ' 성균관대 E하우스 앞' |
| 'SPB-04965' | '2017-04-01 00:03:52' | '525' | ' 한양대병원사거리' | 18 | '2017-04-01 00:11:53' | '512' | ' 뚝섬역 1번 출구 옆' |
| 'SPB-04585' | '2017-04-01 00:04:07' | '529' | ' 장한평역 8번 출구' | 15 | '2017-04-01 00:13:44' | '617' | ' 청솔우성아파트 앞' |
| 'SPB-01816' | '2017-04-01 00:04:17' | '114' | ' 홍대입구역 8번출구' | 9 | '2017-04-01 00:18:07' | '135' | ' 명물길 원형무대 앞' |

4. 파생변수1: 전처리

2. 링크 데이터

| rent_loc_num | return_loc_num |
|--------------|----------------|
| 101 | 101 |
| 101 | 101 |
| 101 | 101 |
| 101 | 101 |
| 101 | 110 |
| 101 | 110 |
| 101 | 110 |
| 101 | 119 |

COUNT

| rent_loc_num | return_loc_num | tot_use |
|--------------|----------------|---------|
| 101 | 101 | 332 |
| 101 | 110 | 138 |
| 101 | 119 | 221 |

링크 방향:

loc_num -> return_loc_num

링크 크기: tot_use



4. 파생변수1: 전처리

노드 데이터



링크 데이터

네트워크 그래프



노드-링크 정보를 igraph 패키지를 통해 네트워크 그래프로 연결

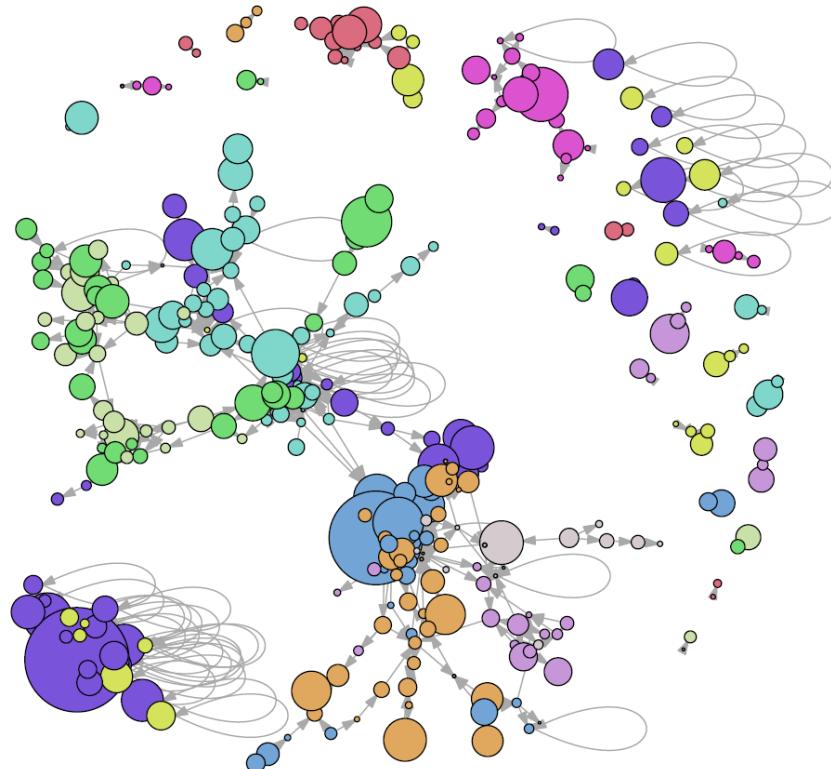
library(igraph)

```
net = graph_from_data_frame(link, node, directed = T)
```



5. 네트워크 시각화

네트워크 그래프



* Link > 300 기준으로 시각화

마포구
성동구
광진구
종로구
은평구
영등포구

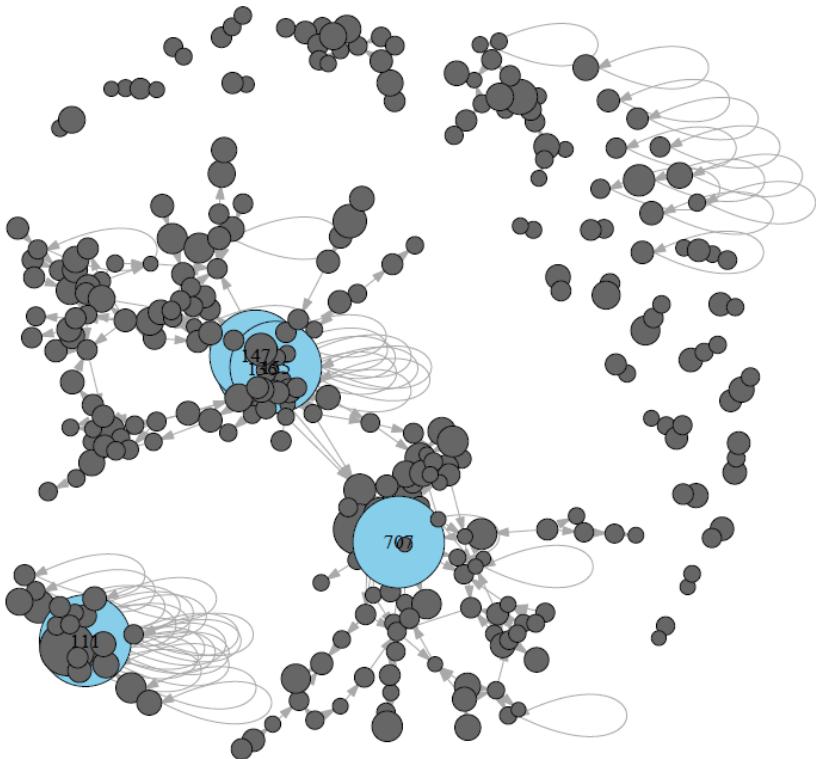
서대문구
중구
양천구
용산구
동대문구



5. 네트워크 시각화

연결 정도 중심성 TOP 5

몇개의 대여소와 연결되어 있는가?



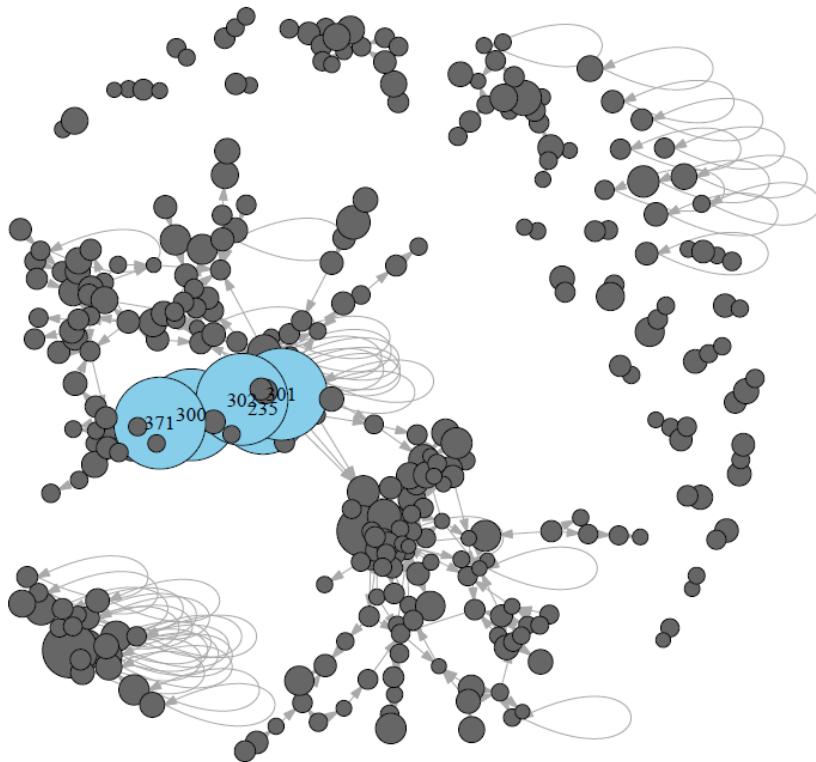
* Link > 300 기준으로 시각화

| 순위 | 대여소 번호 | 대여소 구 | 대여소 이름 |
|----|--------|-------|-------------|
| 1 | 155 | 마포구 | 가좌역 1번 출구 뒤 |
| 2 | 111 | 마포구 | 상수역 2번 출구 앞 |
| 3 | 147 | 마포구 | 마포역 4번 출구 뒤 |
| 4 | 136 | 마포구 | 대흥동 주민센터 |
| 5 | 707 | 양천구 | 신정 3동 주민센터 |

5. 네트워크 시각화

근접 중심성 TOP 5

전체 대여소와의 거리가 얼마나 가까운가?



* Link > 300 기준으로 시각화

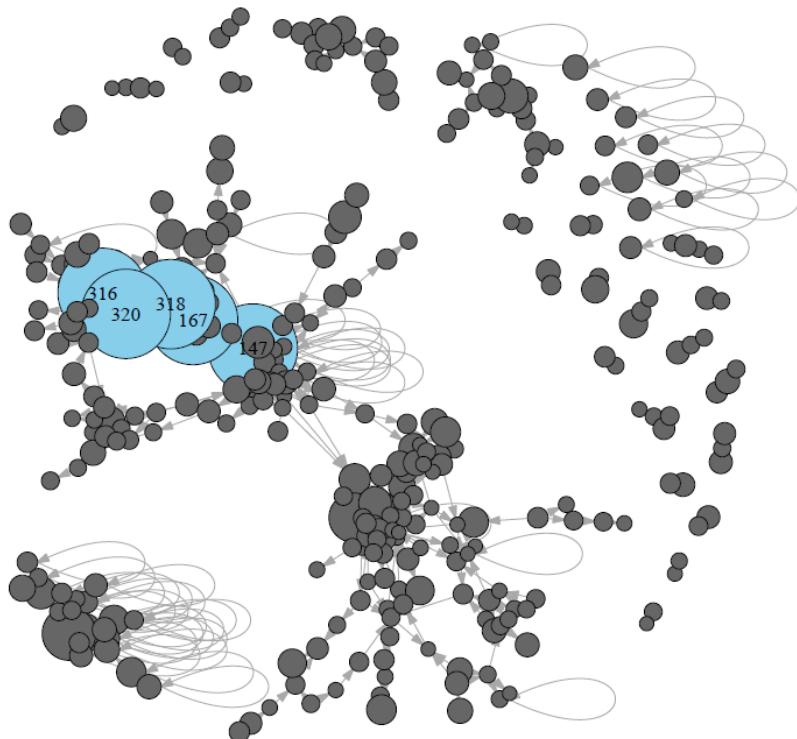
| 순위 | 대여소 번호 | 대여소 구 | 대여소 이름 |
|----|--------|-------|-----------------|
| 1 | 371 | 중구 | 동대입구역 6번출구 뒤 |
| 2 | 302 | 종로구 | 경복궁역 4번출구 뒤 |
| 3 | 300 | 중구 | 정동사거리 |
| 4 | 235 | 영등포구 | 신길동 우리은행 옆 |
| 5 | 301 | 종로구 | 경복궁역 7번출구 앞 |



5. 네트워크 시각화

매개 중심성 TOP 5

다른 대여소를 얼마나 잘 연결시켜주는가?



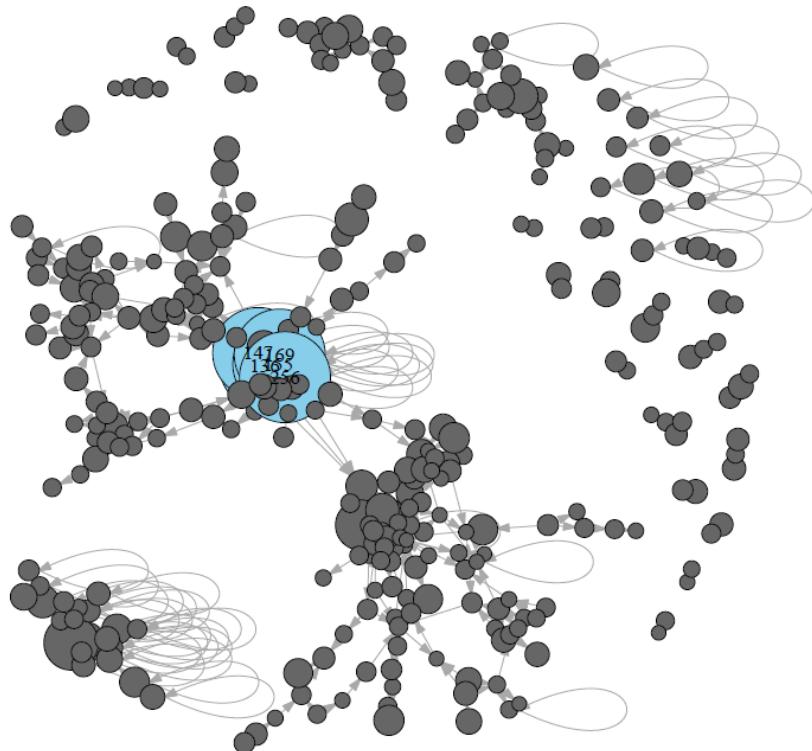
| 순위 | 대여소 번호 | 대여소 구 | 대여소 이름 |
|----|--------|-------|------------------|
| 1 | 167 | 서대문구 | 연가초등학교 옆 |
| 2 | 147 | 마포구 | 마포역 4번출구 뒤 |
| 3 | 316 | 종로구 | 종각역 1번출구 앞 |
| 4 | 318 | 중구 | 광교사거리 남측 |
| 5 | 320 | 중구 | 을지로입구역 4번출구 앞 |

* Link > 300 기준으로 시각화

5. 네트워크 시각화

위세 중심성 TOP 5

중요한 대여소와 연결되어 있는가?



| 순위 | 대여소 번호 | 대여소 구 | 대여소 이름 |
|----|--------|-------|---------------|
| 1 | 155 | 마포구 | 가좌역 1번출구 뒤 |
| 2 | 136 | 마포구 | 대흥동 주민센터 |
| 3 | 256 | 영등포구 | 동아에코빌 입구 |
| 4 | 169 | 서대문구 | 북가좌 삼거리 |
| 5 | 147 | 마포구 | 마포역 4번출구 뒤 |

* Link > 300 기준으로 시각화



6. 파생변수1: 군집화

중심성 산출

| btw | close | degree | eigen |
|-------------|--------------|--------|------------|
| 110.210152 | 0.0002233666 | 115 | 0.33637099 |
| 10.635965 | 0.0002213077 | 19 | 0.04755840 |
| 9.281161 | 0.0002232602 | 19 | 0.04990792 |
| 78.884304 | 0.0001715171 | 146 | 0.37186261 |
| 0.750000 | 0.0002191033 | 16 | 0.04979088 |
| 175.939734 | 0.0002812145 | 141 | 0.39729732 |
| 278.963849 | 0.0002702725 | 120 | 0.32125661 |
| 276.895491 | 0.0002999132 | 189 | 0.50894613 |
| 1942.525995 | 0.0002978306 | 315 | 0.83627831 |
| 1305.444845 | 0.0002800900 | 253 | 0.66723780 |

노드-링크 정보를 통해 구한
443개의 대여소 별,
4가지 중심성 수치 산출

✓중심성 변수 추가

| 날짜 + 대여소 |
|----------------|
| 2018-01-01-101 |
| 2018-01-02-101 |
| . |
| . |
| 2018-03-31-993 |
| 2018-03-31-993 |

| 날짜 INFO | 대여소 INFO |
|---------|---|
| | btw close degree eigen |
| | 110.210152 0.0002233666 115 0.33637099 |
| | 10.635965 0.0002213077 19 0.04755840 |
| | 9.281161 0.0002232602 19 0.04990792 |
| | 78.884304 0.0001715171 146 0.37186361 |
| | 0.750000 0.0002191033 16 0.04979088 |
| | 175.939734 0.0002812145 141 0.39729732 |
| | 278.963849 0.0002702725 120 0.32125661 |
| | 276.895491 0.0002999132 189 0.50894613 |
| | 1942.525995 0.0002978306 315 0.83627831 |
| | 1305.444845 0.0002800900 253 0.66723780 |

| 대여건수 |
|------|
| 14 |
| 45 |
| . |
| . |
| 3 |
| 14 |



4. 파생 변수2



네트워크 분석을 통한 중심성 산출

PCA를 통한 지수 산출



1. PCA란?

PCA: (Principal Component Analysis _ 주성분 분석)

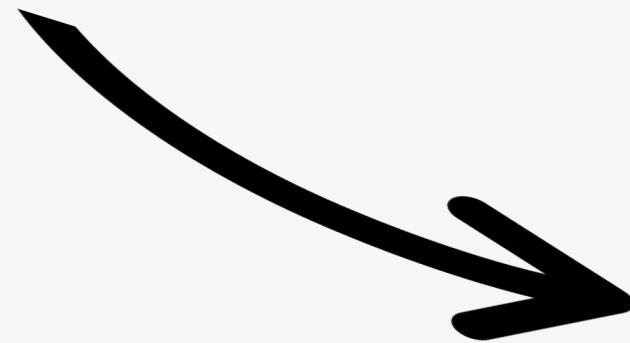


차원 축소를 통해,
고차원 정보를 유지하면서,
다변량 데이터 처리하는 방식



2. 왜 PCA를 사용할까?

주제분석 참여도
뒤풀이 참여 횟수
세미나 집중도
적극성



팀장 지수

고차원 변수들의 정보를 작은 차원의 변수로 표현 가능
(index를 만드는데 빈번히 쓰임)



3. 파생변수2 : 전처리



(1) 따릉이 이용자 정보

| rent_date | loc_num | loc_name | rent_code | sex | age | use_num | excersie | use_distance(m) | use_time(min) |
|------------|---------|----------|-----------|-----|-----|---------|----------|-----------------|---------------|
| 2017-01-01 | 605 | 신설동역8번출구 | 정기 | M | 40대 | 1 | 3198.61 | 115390 | 34 |
| 2017-01-01 | 258 | 신길역3번출구 | 정기 | M | 20대 | 4 | 1254.61 | 33620 | 170 |

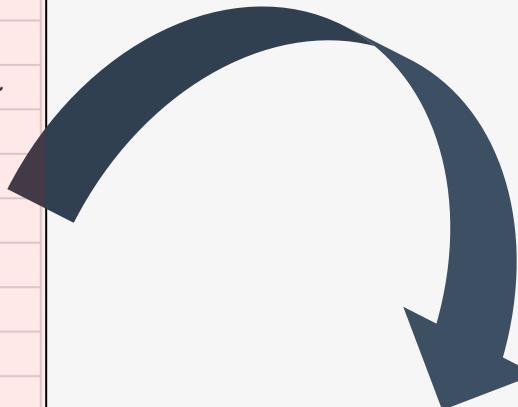
Rent_data , loc_num 을 기준으로 merge key 생성
(2017-01-01-101) … (2017-12-31-933)



3. 파생변수2 : 전처리

(1) mergekey별, 나이/성별 변수 생성

| mergekey | use_num | sex | age |
|----------------|---------|-----|------|
| 2017-01-01-101 | 1 | F | 20대 |
| 2017-01-01-101 | 1 | M | 20대 |
| 2017-01-01-101 | 1 | M | 70대~ |
| 2017-01-01-101 | 1 | F | 20대 |
| 2017-01-01-102 | 1 | F | 20대 |
| 2017-01-01-102 | 3 | M | 20대 |
| 2017-01-01-102 | 3 | M | 30대 |
| 2017-01-01-102 | 1 | M | 40대 |
| 2017-01-01-102 | 1 | M | 20대 |
| 2017-01-01-102 | 1 | M | 30대 |



| mergekey | female_10 | female_20 | female_30 | female_40 | female_50 | female_60 | female_70 | male_10 | male_20 | male_30 | male_40 | male_50 | male_60 | male_70 |
|----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------|---------|---------|---------|---------|---------|---------|
| 2017-01-01-101 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 2017-01-01-102 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 | 1 | 0 | 0 | 0 |



3. 파생변수2 : 전처리

(2) mergekey 별, 이용권 종류 변수 생성

| mergekey | rent_code |
|----------------|-----------|
| 2017-01-01-101 | 정기 |
| 2017-01-01-101 | 정기 |
| 2017-01-01-101 | 정기 |
| 2017-01-01-101 | 일일(회원) |
| 2017-01-01-102 | 정기 |
| 2017-01-01-102 | 일일(회원) |
| 2017-01-01-102 | 일일(회원) |



| mergekey | 정기 | 단체 | 일일_비회원 | 일일_회원 | 일일_2시간 | 정기_2시간 |
|----------------|----|----|--------|-------|--------|--------|
| 2017-01-01-101 | 3 | 0 | 0 | 1 | 0 | 0 |
| 2017-01-01-102 | 4 | 0 | 0 | 2 | 0 | 0 |



3. 파생변수2 : 전처리

(3) mergekey 별 주말 변수 생성

월 화 수 목 금 토 일

주 중 (weekday)

주 말 (weekend)

| mergekey | total | weekend |
|----------------|-------|---------|
| 2017-01-01-101 | 4 | weekend |
| 2017-01-01-102 | 10 | weekend |
| 2017-01-01-103 | 4 | weekend |
| 2017-01-01-104 | 9 | weekend |
| 2017-01-01-105 | 7 | weekend |
| 2017-01-01-106 | 12 | weekend |
| 2017-01-01-107 | 10 | weekend |
| 2017-01-01-108 | 11 | weekend |



| loc | weekday | weekend |
|------|---------|---------|
| 1001 | 4532 | 2279 |
| 1002 | 1965 | 913 |
| 1003 | 1990 | 848 |



3. 파생변수2 : 전처리

(4) mergekey 별, 시간대 변수 생성



(시간대별 이용 정보 활용)

| |
|------------------------------------|
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_1.csv |
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_2.csv |
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_3.csv |
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_4.csv |
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_5.csv |
| 서울특별시 공공자전거 이용정보(시간대별)_2017년_6.csv |



| mergekey | morning_0612 | afternoon_1218 | night_1800 | dawn_0006 |
|----------------|--------------|----------------|------------|-----------|
| 2017-01-01-101 | 2 | 1 | 1 | 0 |
| 2017-01-01-102 | 1 | 1 | 5 | 3 |



3. 파생변수2 : 전처리

대여소에 대한 **특징**을 살펴보기 위해,

| mergekey |
|----------------|
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |

대여소별 누적 도수

| loc_num | female_10 | female_20 | female_30 | female_40 |
|---------|-----------|-----------|-----------|-----------|
| 101 | 5 | 1135 | 576 | 124 |
| 102 | 128 | 5189 | 2538 | 675 |

(각 대여소에 대해 수치 합산)



3. 파생변수2 : 전처리

| mergekey |
|----------------|
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-101 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |
| 2017-01-01-102 |

대여소별 누적 도수

| 대여소명 | 누적 대여수 |
|------|--------|
| 101 | 101 |
| 102 | 128 |
| 103 | 129 |
| 104 | 130 |
| 105 | 131 |
| 106 | 132 |
| 107 | 133 |
| 108 | 134 |
| 109 | 135 |
| 110 | 136 |
| 111 | 137 |
| 112 | 138 |
| 113 | 139 |
| 114 | 140 |
| 115 | 141 |
| 116 | 142 |
| 117 | 143 |
| 118 | 144 |
| 119 | 145 |
| 120 | 146 |
| 121 | 147 |
| 122 | 148 |
| 123 | 149 |
| 124 | 150 |
| 125 | 151 |
| 126 | 152 |
| 127 | 153 |
| 128 | 154 |
| 129 | 155 |
| 130 | 156 |
| 131 | 157 |
| 132 | 158 |
| 133 | 159 |
| 134 | 160 |
| 135 | 161 |
| 136 | 162 |
| 137 | 163 |
| 138 | 164 |
| 139 | 165 |
| 140 | 166 |
| 141 | 167 |
| 142 | 168 |
| 143 | 169 |
| 144 | 170 |
| 145 | 171 |
| 146 | 172 |
| 147 | 173 |
| 148 | 174 |
| 149 | 175 |
| 150 | 176 |
| 151 | 177 |
| 152 | 178 |
| 153 | 179 |
| 154 | 180 |
| 155 | 181 |
| 156 | 182 |
| 157 | 183 |
| 158 | 184 |
| 159 | 185 |
| 160 | 186 |
| 161 | 187 |
| 162 | 188 |
| 163 | 189 |
| 164 | 190 |
| 165 | 191 |
| 166 | 192 |
| 167 | 193 |
| 168 | 194 |
| 169 | 195 |
| 170 | 196 |
| 171 | 197 |
| 172 | 198 |
| 173 | 199 |
| 174 | 200 |
| 175 | 201 |
| 176 | 202 |
| 177 | 203 |
| 178 | 204 |
| 179 | 205 |
| 180 | 206 |
| 181 | 207 |
| 182 | 208 |
| 183 | 209 |
| 184 | 210 |
| 185 | 211 |
| 186 | 212 |
| 187 | 213 |
| 188 | 214 |
| 189 | 215 |
| 190 | 216 |
| 191 | 217 |
| 192 | 218 |
| 193 | 219 |
| 194 | 220 |
| 195 | 221 |
| 196 | 222 |
| 197 | 223 |
| 198 | 224 |
| 199 | 225 |
| 200 | 226 |
| 201 | 227 |
| 202 | 228 |
| 203 | 229 |
| 204 | 230 |
| 205 | 231 |
| 206 | 232 |
| 207 | 233 |
| 208 | 234 |
| 209 | 235 |
| 210 | 236 |
| 211 | 237 |
| 212 | 238 |
| 213 | 239 |
| 214 | 240 |
| 215 | 241 |
| 216 | 242 |
| 217 | 243 |
| 218 | 244 |
| 219 | 245 |
| 220 | 246 |
| 221 | 247 |
| 222 | 248 |
| 223 | 249 |
| 224 | 250 |
| 225 | 251 |
| 226 | 252 |
| 227 | 253 |
| 228 | 254 |
| 229 | 255 |
| 230 | 256 |
| 231 | 257 |
| 232 | 258 |
| 233 | 259 |
| 234 | 260 |
| 235 | 261 |
| 236 | 262 |
| 237 | 263 |
| 238 | 264 |
| 239 | 265 |
| 240 | 266 |
| 241 | 267 |
| 242 | 268 |
| 243 | 269 |
| 244 | 270 |
| 245 | 271 |
| 246 | 272 |
| 247 | 273 |
| 248 | 274 |
| 249 | 275 |
| 250 | 276 |
| 251 | 277 |
| 252 | 278 |
| 253 | 279 |
| 254 | 280 |
| 255 | 281 |
| 256 | 282 |
| 257 | 283 |
| 258 | 284 |
| 259 | 285 |
| 260 | 286 |
| 261 | 287 |
| 262 | 288 |
| 263 | 289 |
| 264 | 290 |
| 265 | 291 |
| 266 | 292 |
| 267 | 293 |
| 268 | 294 |
| 269 | 295 |
| 270 | 296 |
| 271 | 297 |
| 272 | 298 |
| 273 | 299 |
| 274 | 300 |
| 275 | 301 |
| 276 | 302 |
| 277 | 303 |
| 278 | 304 |
| 279 | 305 |
| 280 | 306 |
| 281 | 307 |
| 282 | 308 |
| 283 | 309 |
| 284 | 310 |
| 285 | 311 |
| 286 | 312 |
| 287 | 313 |
| 288 | 314 |
| 289 | 315 |
| 290 | 316 |
| 291 | 317 |
| 292 | 318 |
| 293 | 319 |
| 294 | 320 |
| 295 | 321 |
| 296 | 322 |
| 297 | 323 |
| 298 | 324 |
| 299 | 325 |
| 300 | 326 |
| 301 | 327 |
| 302 | 328 |
| 303 | 329 |
| 304 | 330 |
| 305 | 331 |
| 306 | 332 |
| 307 | 333 |
| 308 | 334 |
| 309 | 335 |
| 310 | 336 |
| 311 | 337 |
| 312 | 338 |
| 313 | 339 |
| 314 | 340 |
| 315 | 341 |
| 316 | 342 |
| 317 | 343 |
| 318 | 344 |
| 319 | 345 |
| 320 | 346 |
| 321 | 347 |
| 322 | 348 |
| 323 | 349 |
| 324 | 350 |
| 325 | 351 |
| 326 | 352 |
| 327 | 353 |
| 328 | 354 |
| 329 | 355 |
| 330 | 356 |
| 331 | 357 |
| 332 | 358 |
| 333 | 359 |
| 334 | 360 |
| 335 | 361 |
| 336 | 362 |
| 337 | 363 |
| 338 | 364 |
| 339 | 365 |
| 340 | 366 |
| 341 | 367 |
| 342 | 368 |
| 343 | 369 |
| 344 | 370 |
| 345 | 371 |
| 346 | 372 |
| 347 | 373 |
| 348 | 374 |
| 349 | 375 |
| 350 | 376 |
| 351 | 377 |
| 352 | 378 |
| 353 | 379 |
| 354 | 380 |
| 355 | 381 |
| 356 | 382 |
| 357 | 383 |
| 358 | 384 |
| 359 | 385 |
| 360 | 386 |
| 361 | 387 |
| 362 | 388 |
| 363 | 389 |
| 364 | 390 |
| 365 | 391 |
| 366 | 392 |
| 367 | 393 |
| 368 | 394 |
| 369 | 395 |
| 370 | 396 |
| 371 | 397 |
| 372 | 398 |
| 373 | 399 |
| 374 | 400 |
| 375 | 401 |
| 376 | 402 |
| 377 | 403 |
| 378 | 404 |
| 379 | 405 |
| 380 | 406 |
| 381 | 407 |
| 382 | 408 |
| 383 | 409 |
| 384 | 410 |
| 385 | 411 |
| 386 | 412 |
| 387 | 413 |
| 388 | 414 |
| 389 | 415 |
| 390 | 416 |
| 391 | 417 |
| 392 | 418 |
| 393 | 419 |
| 394 | 420 |
| 395 | 421 |
| 396 | 422 |
| 397 | 423 |
| 398 | 424 |
| 399 | 425 |
| 400 | 426 |
| 401 | 427 |
| 402 | 428 |
| 403 | 429 |
| 404 | 430 |
| 405 | 431 |
| 406 | 432 |
| 407 | 433 |
| 408 | 434 |
| 409 | 435 |
| 410 | 436 |
| 411 | 437 |
| 412 | 438 |
| 413 | 439 |
| 414 | 440 |
| 415 | 441 |
| 416 | 442 |
| 417 | 443 |
| 418 | 444 |
| 419 | 445 |
| 420 | 446 |
| 421 | 447 |
| 422 | 448 |
| 423 | 449 |
| 424 | 450 |
| 425 | 451 |
| 426 | 452 |
| 427 | 453 |
| 428 | 454 |
| 429 | 455 |
| 430 | 456 |
| 431 | 457 |
| 432 | 458 |
| 433 | 459 |
| 434 | 460 |
| 435 | 461 |
| 436 | 462 |
| 437 | 463 |
| 438 | 464 |
| 439 | 465 |
| 440 | 466 |
| 441 | 467 |
| 442 | 468 |
| 443 | 469 |
| 444 | 470 |
| 445 | 471 |
| 446 | 472 |
| 447 | 473 |
| 448 | 474 |
| 449 | 475 |
| 450 | 476 |
| 451 | 477 |
| 452 | 478 |
| 453 | 479 |
| 454 | 480 |
| 455 | 481 |
| 456 | 482 |
| 457 | 483 |
| 458 | 484 |
| 459 | 485 |
| 460 | 486 |
| 461 | 487 |
| 462 | 488 |
| 463 | 489 |
| 464 | 490 |
| 465 | 491 |
| 466 | 492 |
| 467 | 493 |
| 468 | 494 |
| 469 | 495 |
| 470 | 496 |
| 471 | 497 |
| 472 | 498 |
| 473 | 499 |
| 474 | 500 |
| 475 | 501 |
| 476 | 502 |
| 477 | 503 |
| 478 | 504 |
| 479 | 505 |
| 480 | 506 |
| 481 | 507 |
| 482 | 508 |
| 483 | 509 |
| 484 | 510 |
| 485 | 511 |
| 486 | 512 |
| 487 | 513 |
| 488 | 514 |
| 489 | 515 |
| 490 | 516 |
| 491 | 517 |
| 492 | 518 |
| 493 | 519 |
| 494 | 520 |
| 495 | 521 |
| 496 | 522 |
| 497 | 523 |
| 498 | 524 |
| 499 | 525 |
| 500 | 526 |
| 501 | 527 |
| 502 | 528 |
| 503 | 529 |
| 504 | 530 |
| 505 | 531 |
| 506 | 532 |
| 507 | 533 |
| 508 | 534 |
| 509 | 535 |
| 510 | 536 |
| 511 | 537 |
| 512 | 538 |
| 513 | 539 |
| 514 | 540 |
| 515 | 541 |
| 516 | 542 |
| 517 | 543 |
| 518 | 544 |
| 519 | 545 |
| 520 | 546 |
| 521 | 547 |
| 522 | 548 |
| 523 | 549 |
| 524 | 550 |
| 525 | 551 |
| 526 | 552 |
| 527 | 553 |
| 528 | 554 |
| 529 | 555 |
| 530 | 556 |
| 531 | 557 |
| 532 | 558 |
| 533 | 559 |
| 534 | 560 |
| 535 | 561 |
| 536 | 562 |
| 537 | 563 |
| 538 | 564 |
| 539 | 565 |
| 540 | 566 |
| 541 | 567 |
| 542 | 568 |
| 543 | 569 |
| 544 | 570 |
| 545 | 571 |
| 546 | 572 |
| 547 | 573 |
| 548 | 574 |
| 549 | 575 |
| 550 | 576 |
| 551 | 577 |
| 552 | 578 |
| 553 | 579 |
| 554 | 580 |
| 555 | 581 |
| 556 | 582 |
| 557 | 583 |
| 558 | 584 |
| 559 | 585 |
| 560 | 586 |
| 561 | 587 |
| 562 | 588 |
| 563 | 589 |
| 564 | 590 |
| 565 | 591 |
| 566 | 592 |
| 567 | 593 |
| 568 | 594 |
| 569 | 595 |
| 570 | 596 |
| 571 | 597 |
| 572 | 598 |
| 573 | 599 |
| 574 | 600 |
| 575 | 601 |
| 576 | 602 |
| 577 | 603 |
| 578 | 604 |
| 579 | 605 |
| 580 | 606 |
| 581 | 607 |
| 582 | 608 |
| 583 | 609 |
| 584 | 610 |
| 585 | 611 |
| 586 | 612 |
| 587 | 613 |
| 588 | 614 |
| 589 | 615 |
| 590 | 616 |
| 591 | 617 |
| 592 | 618 |
| 593 | 619 |
| 594 | 620 |
| 595 | 621 |
| 596 | 622 |
| 597 | 623 |
| 598 | 624 |
| 599 | 625 |
| 600 | 626 |
| 601 | 627 |
| 602 | 628 |
| 603 | 629 |
| 604 | 630 |
| 605 | 631 |
| 606 | 632 |
| 607 | 633 |
| 608 | 634 |
| 609 | 635 |
| 610 | 636 |
| 611 | 637 |
| 612 | 638 |
| 613 | 639 |
| 614 | 640 |
| 615 | 641 |
| 616 | 642 |
| 617 | 643 |
| 618 | 644 |
| 619 | 645 |
| 620 | 646 |
| 621 | 647 |
| 622 | 648 |
| 623 | 649 |
| 624 | 650 |
| 625 | 651 |
| 626 | 652 |
| 627 | 653 |
| 628 | 654 |
| 629 | 655 |
| 630 | 656 |
| 631 | 657 |
| 632 | 658 |
| 633 | 659 |
| 634 | 660 |
| 635 | 661 |
| 636 | 662 |
| 637 | 663 |
| 638 | 664 |
| 639 | 665 |
| 640 | 666 |
| 641 | 667 |
| 642 | 668 |
| 643 | 669 |
| 644 | 670 |
| 645 | 671 |
| 646 | 672 |
| 647 | 673 |
| 648 | 674 |
| 649 | 675 |
| 650 | 676 |
| 651 | 677 |
| 652 | 678 |
| 653 | 679 |
| 654 | 680 |
| 655 | 681 |
| 656 | 682 |
| 657 | 683 |
| 658 | 684 |
| 659 | 685 |
| 660 | 686 |
| 661 | 687 |
| 662 | 688 |
| 663 | 689 |
| 664 | 690 |
| 665 | 691 |
| 666 | 692 |
| 667 | 693 |
| 668 | 694 |
| 669 | 695 |
| 670 | 696 |
| 671 | 697 |
| 672 | 698 |
| 673 | |

4. 지수 생성



한량 지수



노예 지수



4. 지수 생성

STEP_1 대여소 주 이용 용도

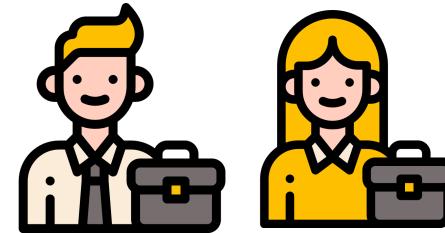
[한량지수]

[노예지수]



여가 즐기기

출퇴근용



4. 지수 생성

(각 지수의 특성을 잘 보여줄 수 있는 적절한 변수 선택)

STEP_2 지수 생성에 필요한 **변수 선정**

[한량지수] :여가



20대 남녀



12시~18시



일일권
일일권(2h)
단체권
비회원권



주말



4. 지수 생성

(각 지수의 특성을 잘 보여줄 수 있는 적절한 변수 선택)

STEP_2 지수 생성에 필요한 **변수 설정**

[노예지수] : 출퇴근



30대
40대
50대
남녀



6시~12시
(출근)
18시~24시
(퇴근)



정기권



주중



5. PCA를 통한 지수 산출

STEP_3 변수 간 상관관계 파악

[한량지수] :여가

| | female_20 | male_20 | group | oneday_2hour | oneday_no_id | oneday_id | weekend | afternoon_1218 |
|----------------|-----------|---------|-------|--------------|--------------|-----------|---------|----------------|
| female_20 | 1 | 0.82 | 0.71 | 0.78 | 0.79 | 0.93 | 0.95 | 0.89 |
| male_20 | 0.82 | 1 | 0.44 | 0.48 | 0.53 | 0.71 | 0.79 | 0.78 |
| group | 0.71 | 0.44 | 1 | 0.91 | 0.91 | 0.87 | 0.83 | 0.77 |
| oneday_2hour | 0.78 | 0.48 | 0.91 | 1 | 0.96 | 0.92 | 0.83 | 0.75 |
| oneday_no_id | 0.79 | 0.53 | 0.91 | 0.96 | 1 | 0.92 | 0.85 | 0.8 |
| oneday_id | 0.93 | 0.71 | 0.87 | 0.92 | 0.92 | 1 | 0.96 | 0.88 |
| weekend | 0.95 | 0.79 | 0.83 | 0.83 | 0.85 | 0.96 | 1 | 0.93 |
| afternoon_1218 | 0.89 | 0.78 | 0.77 | 0.75 | 0.8 | 0.88 | 0.93 | 1 |

[노예지수] : 출퇴근

| | female_30 | female_40 | female_50 | male_30 | male_40 | male_50 | morning_0612 | night_1800 | commutation_2 | commutation |
|-------------------|-----------|-----------|-----------|---------|---------|---------|--------------|------------|---------------|-------------|
| female_30 | 1 | 0.72 | 0.58 | 0.77 | 0.66 | 0.54 | 0.62 | 0.85 | 0.77 | 0.81 |
| female_40 | 0.72 | 1 | 0.64 | 0.55 | 0.66 | 0.56 | 0.53 | 0.63 | 0.59 | 0.61 |
| female_50 | 0.58 | 0.64 | 1 | 0.39 | 0.46 | 0.51 | 0.43 | 0.49 | 0.49 | 0.46 |
| male_30 | 0.77 | 0.55 | 0.39 | 1 | 0.82 | 0.65 | 0.73 | 0.8 | 0.69 | 0.87 |
| male_40 | 0.66 | 0.66 | 0.46 | 0.82 | 1 | 0.72 | 0.72 | 0.66 | 0.58 | 0.73 |
| male_50 | 0.54 | 0.56 | 0.51 | 0.65 | 0.72 | 1 | 0.65 | 0.51 | 0.46 | 0.62 |
| morning_0612 | 0.62 | 0.53 | 0.43 | 0.73 | 0.72 | 0.65 | 1 | 0.58 | 0.5 | 0.8 |
| night_1800 | 0.85 | 0.63 | 0.49 | 0.8 | 0.66 | 0.51 | 0.58 | 1 | 0.86 | 0.88 |
| commutation_2hour | 0.77 | 0.59 | 0.49 | 0.69 | 0.58 | 0.46 | 0.5 | 0.86 | 1 | 0.75 |
| commutation | 0.81 | 0.61 | 0.46 | 0.87 | 0.73 | 0.62 | 0.8 | 0.88 | 0.75 | 1 |



5. PCA를 통한 지수 산출

STEP_3 변수 간 상관관계 파악



[한량지수] :여가

두 지수 모두 PCA를 해도 괜찮겠어!

| | female_20 | male_20 | group | oneday_2hour | oneday_no_id | oneday_id | weekend | afternoon_1218 |
|----------------|-----------|---------|-------|--------------|--------------|-----------|---------|----------------|
| female_20 | 1 | 0.82 | 0.71 | 0.78 | 0.79 | 0.93 | 0.95 | 0.89 |
| male_20 | 0.82 | 1 | 0.44 | 0.48 | 0.53 | 0.71 | 0.79 | 0.78 |
| group | 0.71 | 0.44 | 1 | 0.91 | 0.91 | 0.87 | 0.83 | 0.77 |
| oneday_2hour | 0.78 | 0.48 | 0.91 | 1 | 0.96 | 0.92 | 0.83 | 0.75 |
| oneday_no_id | 0.79 | 0.53 | 0.91 | 0.96 | 1 | 0.92 | 0.85 | 0.8 |
| oneday_id | 0.93 | 0.71 | 0.87 | 0.92 | 0.92 | 1 | 0.96 | 0.88 |
| weekend | 0.95 | 0.79 | 0.83 | 0.83 | 0.85 | 0.96 | 1 | 0.93 |
| afternoon_1218 | 0.89 | 0.78 | 0.77 | 0.75 | 0.8 | 0.88 | 0.93 | 1 |

[노예지수] : 출퇴근

두 지수 모두 PCA를 해도 괜찮겠어!

| | female_30 | female_40 | female_50 | male_30 | male_40 | male_50 | morning_0612 | night_1800 | commutation_2hour | commutation |
|-------------------|-----------|-----------|-----------|---------|---------|---------|--------------|------------|-------------------|-------------|
| female_30 | 1 | 0.72 | 0.58 | 0.77 | 0.66 | 0.54 | 0.62 | 0.85 | 0.77 | 0.81 |
| female_40 | 0.72 | 1 | 0.64 | 0.55 | 0.66 | 0.56 | 0.53 | 0.63 | 0.59 | 0.61 |
| female_50 | 0.58 | 0.64 | 1 | 0.39 | 0.46 | 0.51 | 0.43 | 0.49 | 0.49 | 0.46 |
| male_30 | 0.77 | 0.55 | 0.39 | 1 | 0.82 | 0.65 | 0.73 | 0.8 | 0.69 | 0.87 |
| male_40 | 0.66 | 0.66 | 0.46 | 0.82 | 1 | 0.72 | 0.72 | 0.66 | 0.58 | 0.73 |
| male_50 | 0.54 | 0.56 | 0.51 | 0.65 | 0.72 | 1 | 0.65 | 0.51 | 0.46 | 0.62 |
| morning_0612 | 0.62 | 0.53 | 0.43 | 0.73 | 0.72 | 0.65 | 1 | 0.58 | 0.5 | 0.8 |
| night_1800 | 0.85 | 0.63 | 0.49 | 0.8 | 0.66 | 0.51 | 0.58 | 1 | 0.86 | 0.88 |
| commutation_2hour | 0.77 | 0.59 | 0.49 | 0.69 | 0.58 | 0.46 | 0.5 | 0.86 | 1 | 0.75 |
| commutation | 0.81 | 0.61 | 0.46 | 0.87 | 0.73 | 0.62 | 0.8 | 0.88 | 0.75 | 1 |



5. PCA를 통한 지수 산출

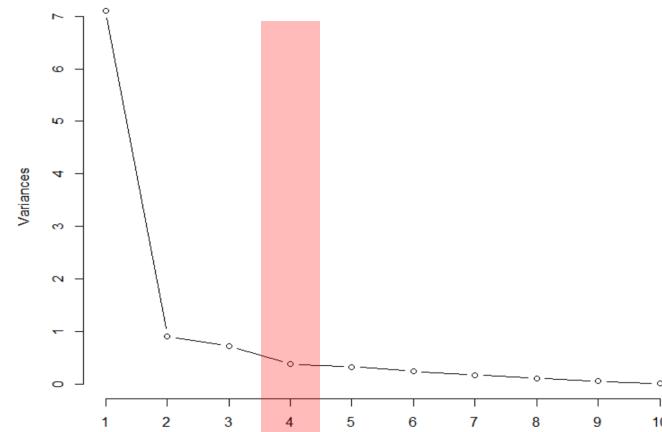
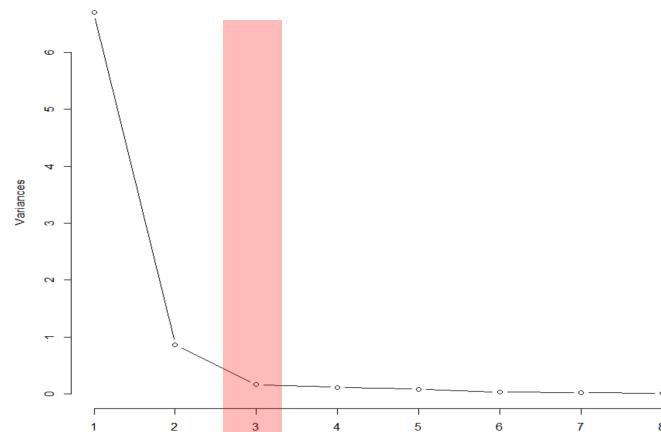
STEP_4 주성분 개수 선택

[한량지수] :여가

[노예지수] : 출퇴근

(Elbow Point : Scree Plot에서 보이는 팔꿈치 점)

Scree Plot에서 Elbow Point를 찾아 결정한다



4. PCA를 통한 지수 산출

STEP_4 주성분 개수 선택

[한량지수] :여가

[노예지수] : 출퇴근

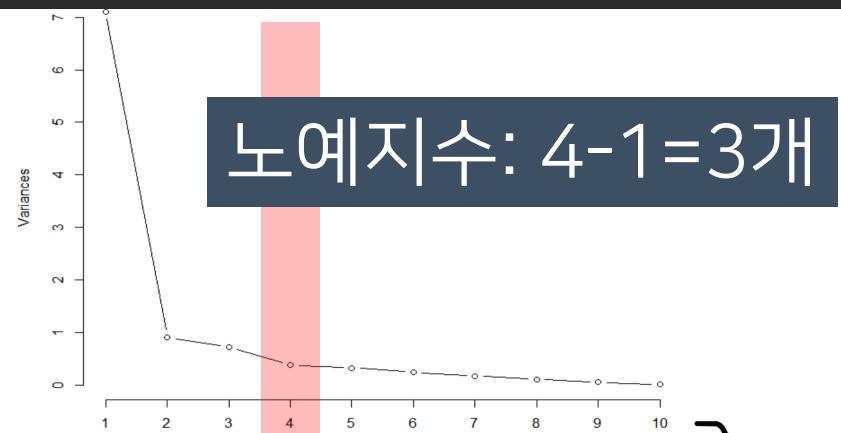


(Elbow Point : Scree Plot에서 보이는 팔꿈치 점)

주성분 개수 : [Elbow Point - 1]



한량지수: $3-1=2$ 개



노예지수: $4-1=3$ 개



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[한량지수]:여가

선정한 변수 Scaling



PC 1&2의
eigenvector

| | female_20 | male_20 | group | oneday_2hour | oneday_no_id | oneday_id | weekend | afternoon_1218 |
|----|--------------|--------------|--------------|---------------|--------------|-------------|--------------|----------------|
| 1 | -0.406046405 | -0.610313159 | -0.269257006 | -0.4434185471 | -0.348480581 | -0.58605387 | -0.595958125 | -0.67449814 |
| 2 | 2.404772724 | 1.112315609 | 0.368967781 | 0.5985379223 | 0.751481141 | 1.27969739 | 1.886587586 | 1.50909554 |
| 3 | 0.929335352 | 0.977698008 | 0.284990836 | 0.4490769533 | 0.428513657 | 0.72924923 | 1.098334020 | 0.73822408 |
| 4 | 1.389022842 | 0.905028329 | 0.293388530 | 0.5131316543 | 0.543970632 | 0.96839525 | 1.239336275 | 1.23479107 |
| 5 | 0.410713569 | 0.365366620 | -0.160086977 | 0.2611831638 | 0.135190531 | 0.33960615 | 0.308947907 | 0.19069935 |
| 6 | 2.591975774 | 2.550817847 | 0.872829456 | 2.0205522843 | 1.255435235 | 2.53109078 | 3.008376603 | 2.04198347 |
| 7 | 0.920321872 | 1.274928905 | -0.336438563 | -0.1829294297 | -0.005230114 | 0.21453553 | 0.989043117 | 1.18979646 |
| 8 | 0.656850913 | 1.080745994 | -0.059314642 | -0.0377387742 | 0.261569112 | 0.41451108 | 0.589253593 | 0.48939849 |
| 9 | 0.586129760 | 1.202854880 | 0.486535505 | 0.5003207141 | 0.406670445 | 1.02130974 | 0.943174914 | 0.37338830 |
| 10 | -0.677837500 | -0.720508983 | -0.521187843 | -0.3409310255 | -0.410889756 | -0.67882603 | -0.685995709 | -0.87995544 |

| | PC1 | PC2 |
|----------------|-----------|-------------|
| female_20 | 0.3620494 | -0.25743708 |
| male_20 | 0.2897697 | -0.66889303 |
| group | 0.3418233 | 0.41972913 |
| oneday_2hour | 0.3515802 | 0.38430577 |
| oneday_no_id | 0.3587788 | 0.32195523 |
| oneday_id | 0.3811167 | 0.03795655 |
| weekend | 0.3770062 | -0.13323921 |
| afternoon_1218 | 0.3583240 | -0.19898802 |



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[한량지수]:여가

obs*PC축 곱한 값



해당 PC의 설명력

| | PC1 | PC2 |
|----|-------------|-------------|
| 1 | -1.38654577 | 0.30852461 |
| 2 | 3.53884773 | -1.23935290 |
| 3 | 1.98534847 | -0.72861585 |
| 4 | 2.51976815 | -0.84155563 |
| 5 | 0.65441667 | -0.33963834 |
| 6 | 5.96724551 | -1.53753164 |
| 7 | 1.40241552 | -1.66330533 |
| 8 | 1.16677464 | -1.00734972 |
| 9 | 1.92749056 | -0.58925612 |
| 10 | -1.73227536 | 0.41511258 |

Importance of components:

| | PC1 | PC2 |
|------------------------|--------|--------|
| Standard deviation | 2.5891 | 0.9258 |
| Proportion of Variance | 0.8379 | 0.1071 |
| Cumulative Proportion | 0.8379 | 0.9451 |



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[한량지수]:여가

PC1 *설명력



PC2 * 설명력

각 대여소 obs마다 한량 지수 산출

| loc_num | HanLyang |
|---------|--------------|
| 101 | -1.128743718 |
| 102 | 2.832465818 |
| 103 | 1.585488730 |
| 104 | 2.021183122 |
| 105 | 0.511960458 |
| 106 | 4.835285370 |
| 107 | 0.996943965 |
| 108 | 0.869753317 |
| 109 | 1.551935009 |
| 110 | -1.407014968 |



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[노예지수] : 출퇴근

선정한 변수 Scaling



PC 1&2&3 의
eigenvector

| female_30 | female_40 | female_50 | male_30 | male_40 | male_50 | morning_0612 | night_1800 | commutation | weekday |
|--------------|--------------|---------------|---------------|--------------|--------------|--------------|--------------|---------------|--------------|
| -0.276765094 | -0.784590561 | -0.3316335128 | -0.1729657990 | -1.162162444 | -1.345764084 | -0.482964699 | -0.694970191 | -0.6975263703 | -0.696194762 |
| 3.622458636 | 1.338308177 | 2.0450525948 | 2.6364495393 | 0.482755910 | 0.312322966 | 1.408975378 | 2.062413954 | 2.3500760271 | 2.055016426 |
| 1.174016416 | 0.822031533 | 0.4330394088 | 1.0763252917 | -0.149432860 | 1.075043009 | 0.387998387 | 0.984545523 | 1.0553465750 | 0.918957063 |
| 1.889470312 | 0.174759322 | -0.1249651556 | 1.5112930960 | 0.384551829 | 0.399372536 | -0.109694805 | 1.510937125 | 1.2010070738 | 1.130069351 |
| 0.641400739 | -0.422426348 | -0.0216309770 | -0.0478380745 | -0.483940511 | -0.587189259 | -0.250001928 | 0.133240498 | 0.0280278121 | 0.055032606 |
| 3.268706432 | 0.687183156 | 0.3572610111 | 2.1140909434 | 1.265319679 | 1.460548248 | 0.472712122 | 2.939600563 | 2.3168214604 | 2.248211435 |
| 1.903381915 | 1.315191313 | 0.2745936683 | 1.6503239010 | 2.061693397 | 0.577616894 | 1.213074866 | 1.165716612 | 2.0936585831 | 1.637465922 |
| 0.263800072 | -0.588097212 | -0.6209692129 | 0.4417489746 | -0.373460920 | -0.578898824 | -0.892414417 | 0.676753762 | 0.5213496900 | 0.434671186 |
| 1.140231093 | -0.487924132 | -0.0905204294 | 0.6214959439 | 0.069991883 | -0.408944901 | -0.294123665 | 0.808948930 | 0.5023863420 | 0.568401602 |
| -0.606668834 | -0.233638620 | -0.7380812819 | -0.8095282705 | -0.890566783 | -1.047308415 | -0.848292680 | -0.745140646 | -0.8736931245 | -0.905749027 |

| | PC1 | PC2 | PC3 |
|--------------|-----------|-------------|-------------|
| female_30 | 0.3317988 | 0.05868355 | 0.36330096 |
| female_40 | 0.2877528 | 0.46257264 | 0.10754113 |
| female_50 | 0.2345605 | 0.71908529 | 0.10171187 |
| male_30 | 0.3384515 | -0.28822278 | -0.05448291 |
| male_40 | 0.3213432 | -0.03034252 | -0.35009958 |
| male_50 | 0.2822287 | 0.16653022 | -0.60104032 |
| morning_0612 | 0.3061180 | -0.15285358 | -0.36898824 |
| night_1800 | 0.3316932 | -0.15206840 | 0.42939483 |
| commutation | 0.3505380 | -0.24358527 | 0.10030376 |
| weekday | 0.3577514 | -0.21897715 | 0.16947379 |



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[노예지수] : 출퇴근

obs*PC축 곱한 값



해당 PC의 설명력

| | PC1 | PC2 | PC3 |
|--------------|-----------|-------------|-------------|
| female_30 | 0.3317988 | 0.05868355 | 0.36330096 |
| female_40 | 0.2877528 | 0.46257264 | 0.10754113 |
| female_50 | 0.2345605 | 0.71908529 | 0.10171187 |
| male_30 | 0.3384515 | -0.28822278 | -0.05448291 |
| male_40 | 0.3213432 | -0.03034252 | -0.35009958 |
| male_50 | 0.2822887 | 0.16653022 | -0.60104032 |
| morning_0612 | 0.3061180 | -0.15285358 | -0.36898824 |
| night_1800 | 0.3316932 | -0.15206840 | 0.42939483 |
| commutation | 0.3505380 | -0.24358527 | 0.10030376 |
| weekday | 0.3577514 | -0.21897715 | 0.16947379 |

Importance of components:

| | PC1 | PC2 | PC3 |
|------------------------|--------|---------|---------|
| Standard deviation | 2.6662 | 0.95075 | 0.84787 |
| Proportion of Variance | 0.7109 | 0.09039 | 0.07189 |
| Cumulative Proportion | 0.7109 | 0.80127 | 0.87316 |



5. PCA를 통한 지수 산출

STEP_ 5 주성분 점수 생성

[노예지수] : 출퇴근

PC1 *설명력



PC2 *설명력



PC3 *설명력

각 대여소 obs마다 노예 지수 산출

| loc_num | Noye |
|---------|--------------|
| 101 | -1.450935360 |
| 102 | 4.332511832 |
| 103 | 1.800316922 |
| 104 | 1.906950173 |
| 105 | -0.139965474 |
| 106 | 4.009767908 |
| 107 | 3.173926883 |
| 108 | -0.066687150 |
| 109 | 0.643999410 |
| 110 | -1.694856991 |



5. PCA를 통한 지수 산출

최종: 완성된 한량지수 & 노예지수

| loc_num | HanLyang_index | slave_index |
|---------|----------------|--------------|
| 101 | -1.128743718 | -1.450935360 |
| 102 | 2.832465818 | 4.332511832 |
| 103 | 1.585488730 | 1.800316922 |
| 104 | 2.021183122 | 1.906950173 |
| 105 | 0.511960458 | -0.139965474 |
| 106 | 4.835285370 | 4.009767908 |
| 107 | 0.996943965 | 3.173926883 |
| 108 | 0.869753317 | -0.066687150 |
| 109 | 1.551935009 | 0.643999410 |
| 110 | -1.407014968 | -1.694856991 |



5. PCA를 통한 지수 산출

최종: 완성된 한량지수 & 노예지수



각 대여소에 대한, 사용자의 특성이
파악 가능해짐



이 대여소는 한량 지수가 높다!

| loc_num | HanLyang_index | slave_index |
|---------|----------------|--------------|
| 101 | -1.128743718 | -1.450935360 |
| 102 | -0.82020318 | 4.037573832 |
| 103 | 1.585466130 | 1.80316922 |
| 104 | 0.211821 | 0.5650073 |
| 105 | 0.511960458 | -0.139965474 |
| 106 | 4.835285370 | 4.009767908 |
| 107 | 0.996943965 | 3.173926883 |
| 108 | 0.869753317 | -0.066687150 |
| 109 | 0.996943965 | 0.643999410 |
| 110 | -1.407014968 | -1.694856991 |



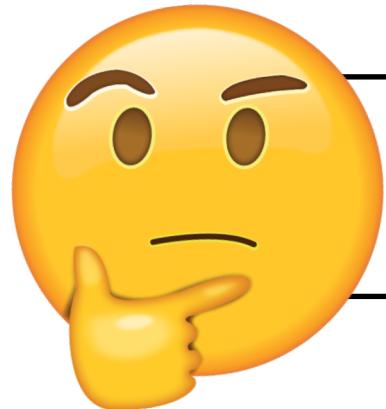
여기로 자주 사용되는 곳이다!

5. PCA를 통한 지수 산출

산출된 지수가 대여소를 잘 표현하는가 ?

[한량지수]

[노예지수]



각 지수의 대표적인 예시로 알아보자!



5. PCA를 통한 지수 산출

예시1_서울숲 남문 버스정류소 옆

| HanLiang_index | slave_index | loc_name |
|----------------|-------------|-----------------|
| 18.2126024 | 6.8875175 | 뚝섬 유원지역 1번 출구 앞 |
| 8.9440935 | 1.8974070 | 서울숲역 4번 출구 옆 |
| 3.4249387 | 1.3337781 | 외국어대 정문 앞 |
| 2.6983850 | -0.9929635 | 서울숲 남문 버스정류소 옆 |
| 1.9139273 | 5.6713292 | 마포구청 앞 |
| 1.4768766 | 3.7778088 | 경복궁역 4번 출구 뒤 |
| -0.1591175 | 2.6409344 | DMC산학협력연구센터 앞 |
| -0.4193508 | 2.9480187 | 정동사거리 |



① 해당 대여소 : 공원 옆으로, 여가 생활과 관련 있음

② 한량 지수는 2.7 | 노예 지수는 -1로 유의미한 차이



5. PCA를 통한 지수 산출

예시2_외국어대 정문 앞

| HanLyang_index | slave_index | loc_name |
|----------------|-------------|-----------------|
| 18.2126024 | 6.8875175 | 뚝섬 유원지역 1번 출구 앞 |
| 8.9440935 | 1.8974070 | 서울숲역 4번 출구 옆 |
| 3.4249387 | 1.3337781 | 외국어대 정문 앞 |
| 2.6983850 | -0.9929635 | 서울숲 남문 버스정류소 옆 |
| 1.9139273 | 5.6713292 | 마포구청 앞 |
| 1.4768766 | 3.7778088 | 경복궁역 4번 출구 뒤 |
| -0.1591175 | 2.6409344 | DMC산학협력연구센터 앞 |
| -0.4193508 | 2.9480187 | 정동사거리 |



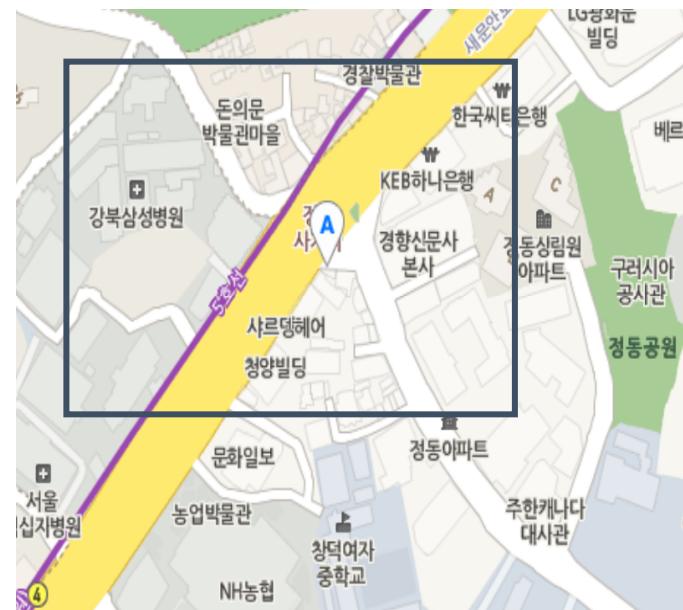
- ① 해당 대여소 : 대학가로, 20대를 한량 지수에 잘 반영
- ② 한량 지수는 3.4 | 노예 지수는 1.3로 유의미한 차이



5. PCA를 통한 지수 산출

예시3_정동 사거리

| HanLyang_index | slave_index | loc_name |
|----------------|-------------|-----------------|
| 18.2126024 | 6.8875175 | 뚝섬 유원지역 1번 출구 앞 |
| 8.9440935 | 1.8974070 | 서울숲역 4번 출구 옆 |
| 3.4249387 | 1.3337781 | 외국어대 정문 앞 |
| 2.6983850 | -0.9929635 | 서울숲 남문 버스정류소 옆 |
| 1.9139273 | 5.6713292 | 마포구청 앞 |
| 1.4768766 | 3.7778088 | 경복궁역 4번 출구 뒤 |
| -0.1591175 | 2.6409344 | DMC산학협력연구센터 앞 |
| -0.4193508 | 2.9480187 | 정동사거리 |



① 해당 대여소 : 회사 주변에 존재 (한국씨티은행, 경향신문)

② 한량 지수는 -0.42 노예 지수는 3 으로 유의미한 차이



✓ 지수 추가

| 날짜 + 대여소 |
|----------------|
| 2018-01-01-101 |
| 2018-01-02-101 |
| . |
| . |
| 2018-03-31-993 |
| 2018-03-31-993 |

| 날짜 INFO | 대여소 INFO | | | | |
|-------------|--------------|--------|------------|----------------|--------------|
| btw | close | degree | eigen | HanLiang_index | slave_index |
| 110.210152 | 0.0002233666 | 115 | 0.33637099 | -1.128743718 | -1.450935360 |
| 10.635965 | 0.0002213077 | 19 | 0.04755840 | 2.832465818 | 4.332511832 |
| 9.281161 | 0.0002232602 | 19 | 0.04990792 | 1.585488730 | 1.800316922 |
| 78.884304 | 0.0001715171 | 146 | 0.37186261 | 2.021183122 | 1.906950173 |
| 0.750000 | 0.0002191033 | 16 | 0.04979088 | 0.511960458 | -0.139965474 |
| 175.939734 | 0.0002812145 | 141 | 0.39729732 | 4.835285370 | 4.009767908 |
| 278.963849 | 0.0002702725 | 120 | 0.32125661 | 0.996943965 | 3.173926883 |
| 276.895491 | 0.0002999132 | 189 | 0.50894613 | 0.869753317 | -0.066687150 |
| 1942.525995 | 0.0002978306 | 315 | 0.83627831 | 1.551935009 | 0.643999410 |
| 1305.444845 | 0.0002800900 | 253 | 0.66723780 | -1.407014968 | -1.694856991 |

| 대여건수 |
|------|
| 14 |
| 45 |
| . |
| . |
| 3 |
| 14 |



5. 다음 주 예고



1. 다음 주 예고

날씨(기온/미세먼지) 변수 추가

모델 평가



1. 다음 주 예고

날씨(기온/미세먼지) 변수 추가

모델 평가



2. 날씨 변수 추가

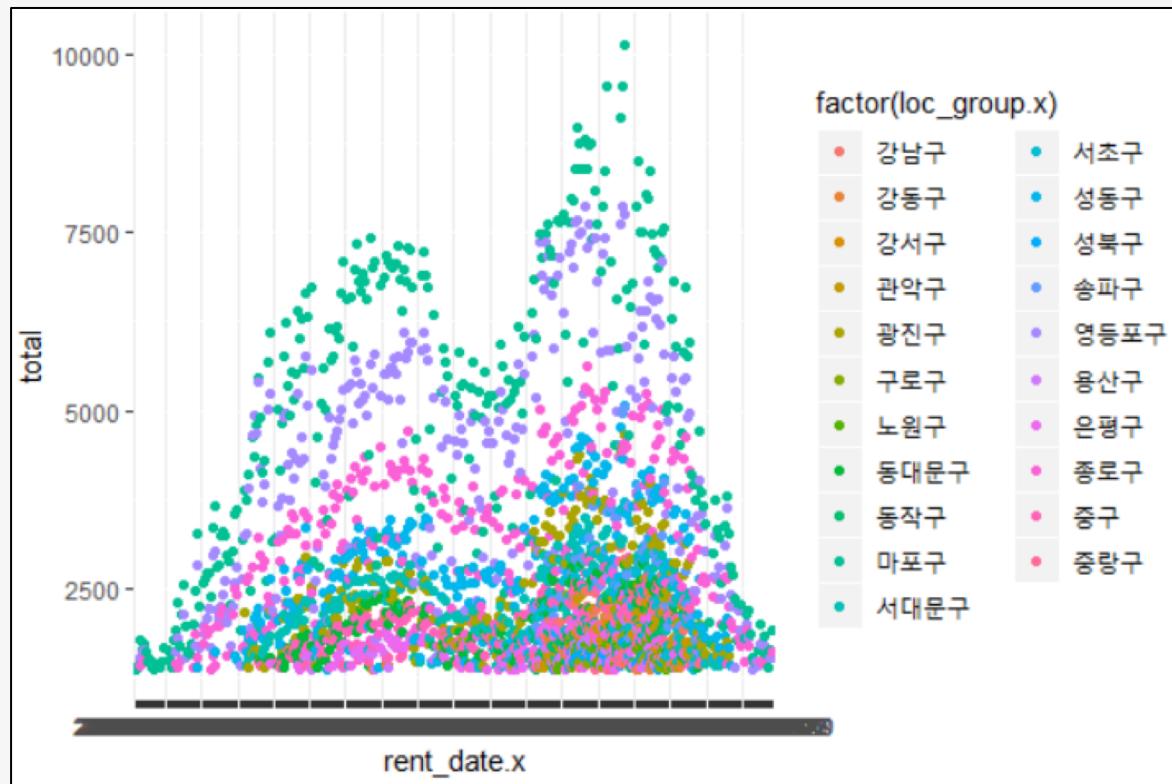
| 날짜 + 대여소 |
|----------------|
| 2018-01-01-101 |
| 2018-01-02-101 |
| . |
| . |
| . |
| 2018-03-31-993 |
| 2018-03-31-993 |

| 날짜 INFO | 대여소 INFO |
|-----------|----------|
| | |
| | |
| | |
| 4주차 FOCUS | . |
| . | . |
| . | . |
| | |
| | |

| 대여건수 |
|------|
| 14 |
| 45 |
| . |
| . |
| 3 |
| 14 |



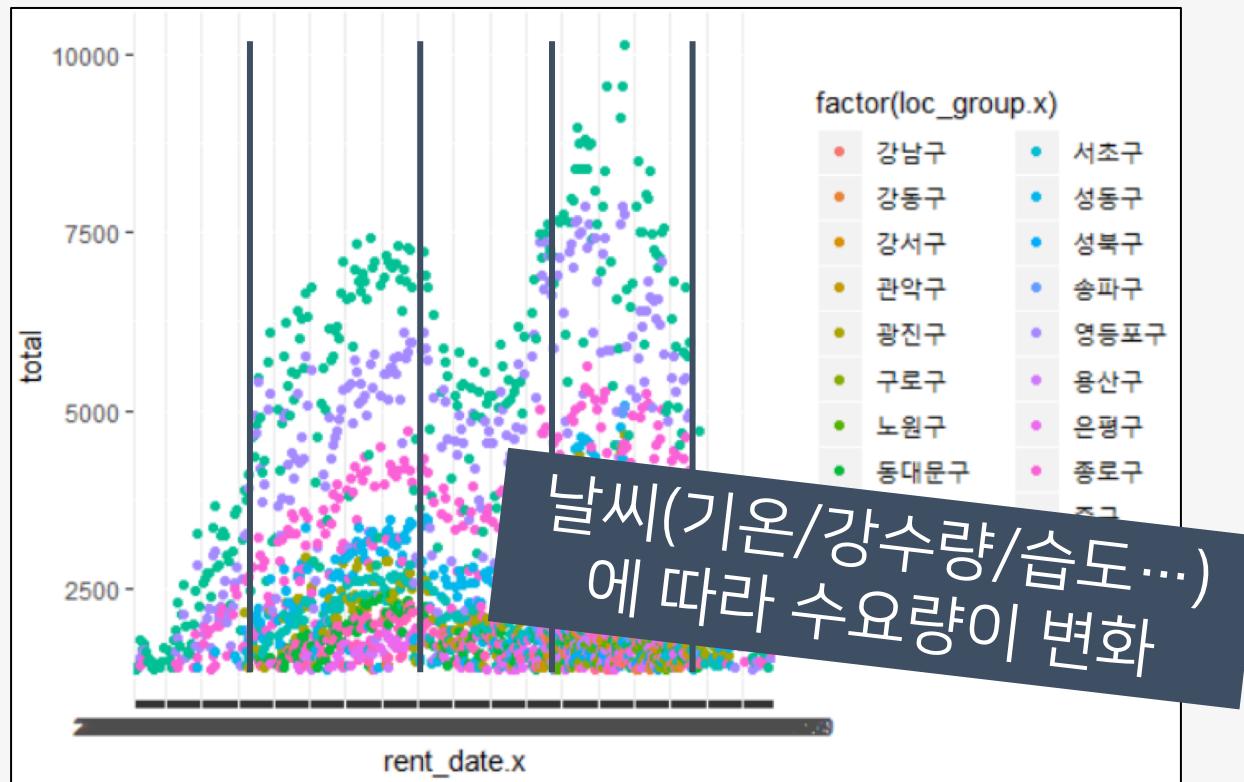
2. 날씨 변수 추가



X: 2017년의 365일 / Y: 구 대여소의 총 대여량



2. 날씨 변수 추가



X: 2017년의 365일 / Y: 구 대여소의 총 대여량



2. 날씨 변수 추가



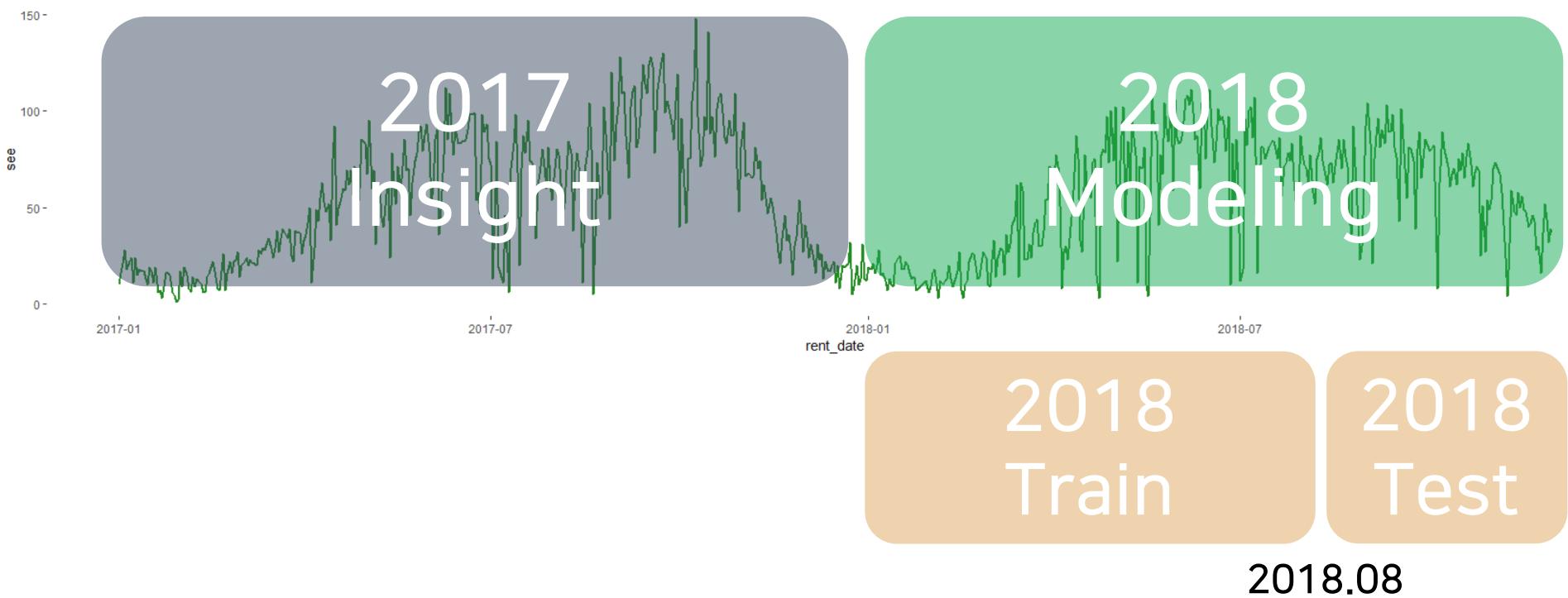
3. 모델 평가

날씨 변수 추가

모델 평가



3. 모델 평가



3. 모델 평가



