



# **서울시 장애인 셔틀버스 동선탐색**

**2016122015**  
**응용통계학과**  
**김선우**

# Research Q.

그림1

그림 1. 교통수단 이용시 어려움을 느끼는 주된 이유 (단위 %)

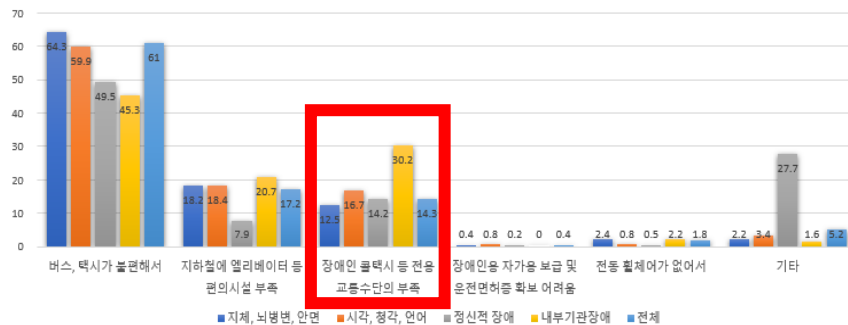
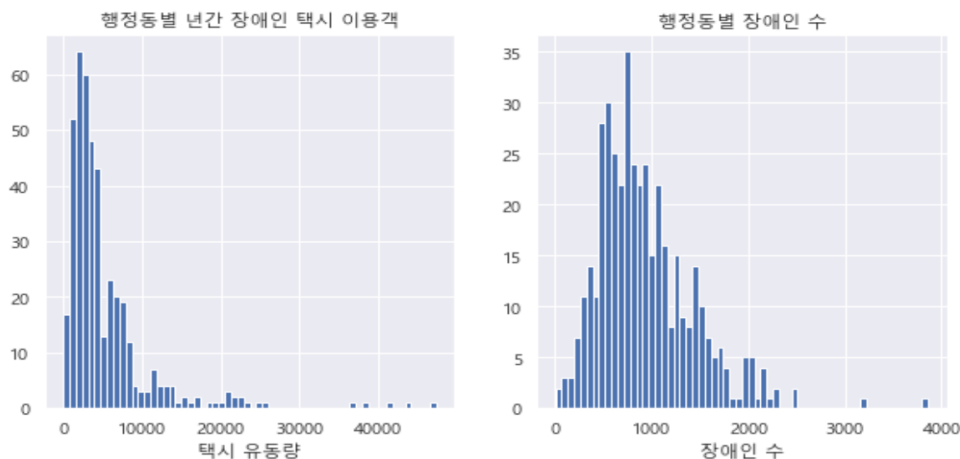


그림2



교통에 대한 장애인들의 불편은 심각(그림 1.)

코로나 사태까지 겹치며, 도움을 받기 더욱 어려워짐

이러한 가운데, 19년 하반기부터 **장애인 셔틀버스는 폐지**

→ 노약자 + 장애인 + etc.. 등의 **교통약자 셔틀버스로 대체**

관련부처(서울시 장애인 자립 지원과) 문의 결과, **저조한 이용** 때문!

그렇다면, **Why??**

10조 통합 데이터 EDA 결과, 행정동 별 장애인 수와

유동인구는 확연한 차이를 보임! 하지만 과거 장애인 셔틀버스는

**장애인이 많은 곳, 장애인이 많이 다니는 곳을 고려하지 않았다.**

장애인들이 많이 다니고, 장애인들이 많이 살며,

장애인 관련 시설이 많은 곳으로 동선을 설정해,

**'수요가 많을' 셔틀버스 동선을 탐색**하는 것이 본 프로젝트의 목적

본 프로젝트의 한계 :

- [유동량/거주 장애인/관련 시설이 많음] == 장애인들의 수요가 많을 것이라는 **굉장히 직관적인 가정**  
**\*\* 해당 부처 문의결과 관련하여 조사된 바가 없다고 함. 관련한 조사가 필요**
- 장애인 유동인구 데이터가 시중에 공개되지 않았기에, 콜택시 호출량이 많은 지역일수록 유동량이 많을 것이라는 또 다른 **강력한 가정** / 이에 근거하여 콜택시 호출 데이터를 유동인구의 Proxy로 사용
- 장애인 셔틀버스 이용률이 저조했던 원인에 대한 선행 연구가 없어, 기타 시설적 측면의 원인에 대해서 유추하기 힘들.

# Task Flow

## 장애인 콜택시 이용 API 데이터

	total_date	startpos1	endpos1	startpos2_	endpos2_
0	20190101	종로구	성북구	삼청동	월곡1동
1	20190101	성북구	강북구	보문동	삼각산동

## 10조 통합 데이터

	si	gu	dong	number_of_disabled	welfare_facility
0	서울시	종로구	사직동	295	4.0
1	서울시	종로구	삼청동	112	1.0
2	서울시	종로구	부암동	316	0.0

동선의 선택 기준은(수요가 많은 동선!)

1. 장애인 유동량이 많고.(콜택시 데이터 기준)
2. 거주 장애인이 많으며.(10조 통합 데이터 기준)
3. 보유 복지시설이 많다!(10조 통합 데이터 기준)

인접한 3개의 자치구 중,  
동선 선택 기준에 가장 부합하는  
인접 3개 자치구 선정!

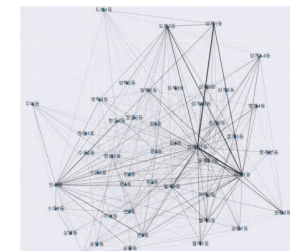
## 순위 리스트

	first	second	third	moving	number_of_dis	facility
0	노원	도봉	강북	130341.0	60309	186.0
1	노원	강북	성북	120562.0	62546	162.0
2	노원	중랑	도봉	119064.0	63153	160.0
3	노원	중랑	강북	118508.0	65220	155.0

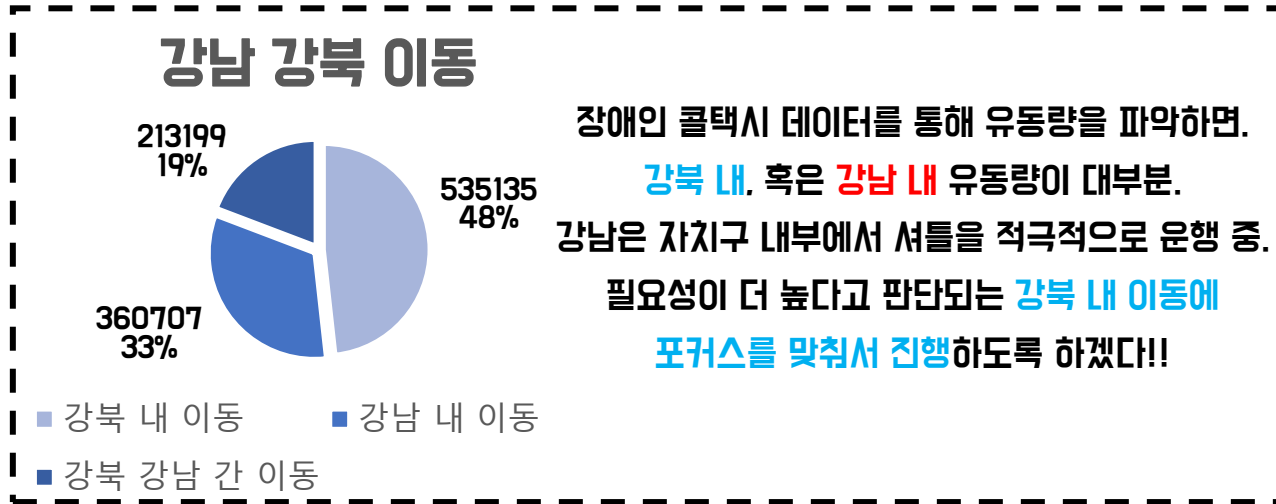
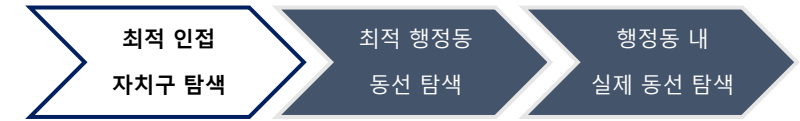
복지 및 편의시설 위치를 고려해  
행정동 내 실제 버스 동선 선정



선정된 인접 3개 자치구 내,  
동선 선택 기준에 가장 부합하는  
행정동 연결 동선 탐색



# 최적 인접 자치구 탐색



### 1.1. 인접한 자치구 목록 생성

인접의 개념 : Adjacency Matrix(A)를 사용  
 두 자치구가 맞닿으면 1 아니면 0  
 $A^2$ 을 통해 연결된 세개의 자치구 탐색 가능

A =

	mapo	eunpyung	seodae	jongro	gangbook	sungbook	joonggu	yongsan	dobong	dongdae	sungdong	nowon	joongrang	gwangjin
mapo	0	1	1	0	0	0	1	1	0	0	0	0	0	0
eunpyung	1	0	1	1	0	0	0	0	0	0	0	0	0	0
seodae	1	1	0	1	0	0	1	0	0	0	0	0	0	0
jongro	0	1	1	0	0	1	1	0	0	1	0	0	0	0
gangbook	0	0	0	0	0	1	0	0	1	0	0	1	0	0
sungbook	0	0	0	1	1	0	0	0	0	0	1	0	1	1
joonggu	1	0	1	1	0	0	0	1	0	0	1	0	0	0
yongsan	1	0	0	0	0	0	1	0	0	0	1	0	0	0
dobong	0	0	0	0	1	0	0	0	0	0	0	1	0	0
dongdae	0	0	0	1	0	1	0	0	0	0	1	0	1	1
sungdong	0	0	0	0	0	0	1	1	0	1	0	0	0	1
nowon	0	0	0	0	1	1	0	0	1	0	0	0	1	0
joongrang	0	0	0	0	0	1	0	0	0	1	0	1	0	1
gwangjin	0	0	0	0	0	0	0	0	0	1	1	0	1	0

이를 통하여  
 46개의  
 인접 3개  
 자치구 탐색

[['마포', '은평', '서대문'],  
 ['마포', '중구', '서대문'],  
 ['마포', '은평', '종로'],  
 ['마포', '종로', '서대문'],  
 ['마포', '중구', '종로'],  
 ['마포', '중구', '용산'],  
 ['마포', '중구', '성동'],  
 :  
 :  
 :]

1.2. 이렇게 만들어진 인접 3개 자치구 목록에  
 다음과 같은 항목 추가 (ex. 마포, 종로, 서대문)

[교통량]

- 해당 3개 자치구 내에서의 유동량
- 택시 데이터에서 (마포 - 마포, 마포 - 서대문 ...) 유동량 추출
- (마포 종로 서대문)의 경우 총 6개의 pair가 생성됨. 이를 합산

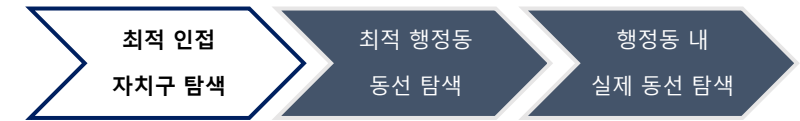
[거주 장애인 현황]

- 마포, 종로, 서대문 내에 거주중인 장애인의 수를 합산

[장애인 시설 수 현황]

- 마포, 종로, 서대문 내에 위치한 장애인 시설 수를 합산

# 최적 인접 자치구 탐색



## 1.3. 어느 인접 자치구를 선택할 것인가?

### 생성된 인접 자치구 데이터

	first	second	third	유동량 moving	장애인 수 number_of_dis	복지시설 수 facility
0	노원	도봉	강북	130341.0	60309	186.0
1	노원	강북	성북	120562.0	62546	162.0
2	노원	중랑	도봉	119064.0	63153	160.0

이 세 지표는 높을수록 좋다.

하지만, 이 세 지표에 대한

3개자치구 별 Rank가 다름!

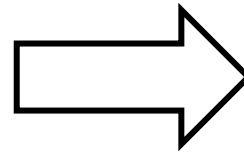
Ex. 유동량은 [노원 도봉 강북]이 1위

하지만 장애인 수는 [노원 중랑 성북]이 1위

Then, 이들을 어떻게 비교/평가할 것인가?

PCA를 통해,

이 정보를 잘 보존하는  
변수 선형 결합 탐색  
세 변수를, 하나의 변수로  
축약할 수 있음!



Eigen Vector들이 모두 양수 = PC값이 클수록 좋다!!

First Eigen Vector : [0.64720141 0.55742599 0.52000635]  
Explained Variance By First PC : 0.7707153573697966

첫 PC가 전체 변동량의 77.1% 설명 가능하므로,

첫 PC만으로 충분히 설명력이 있다!

Z : PC값 / X : Scaled된 [유동량 장애인수 복지시설수] / U : Eigen Vector

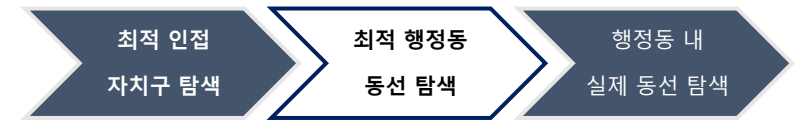
$Z = Xu^T$  에 따라 계산된 PC 값을 Score로 하여 정렬

	first	second	third	moving	number_of_dis	facility	Score
0	노원	도봉	강북	130341.0	60309	186.0	3.714567
1	노원	강북	성북	120562.0	62546	162.0	3.242748
3	노원	중랑	강북	118508.0	65220	155.0	3.214448
2	노원	중랑	도봉	119064.0	63153	160.0	3.208358

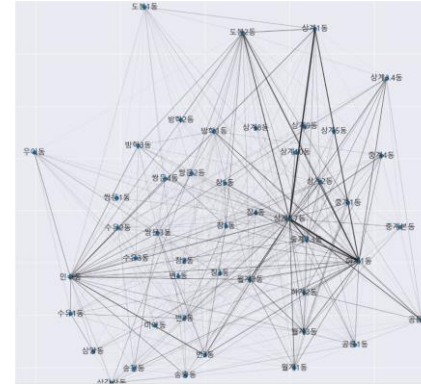
2위와 꽤 큰 격차로 1위를 한

**노원 - 도봉 - 강북** 을 최종 구 동선으로 선정!

# 최적 행정동 동선 탐색



이를 토대로 네트워크 모델링

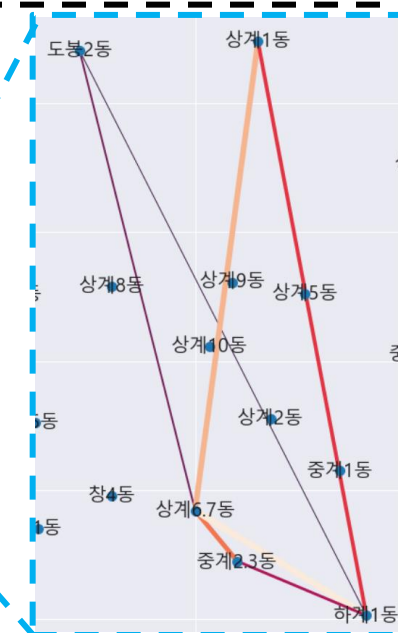
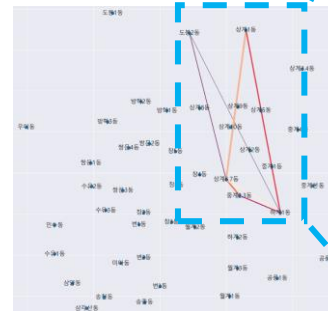


모든 [노원 도봉 강북] 내 유동성을 시각화 하는 것은 해석이 거의 불가능한 수준

유동량이 많은 상위 k개 동선만 시각화

노원 도봉 강북 내 46개의 행정동  
행정동 내부의 이동은 고려 X  
행정동 간 총 1035개의 조합 가능  
해당 조합 내 유동량을 파악하여 데이터 생성

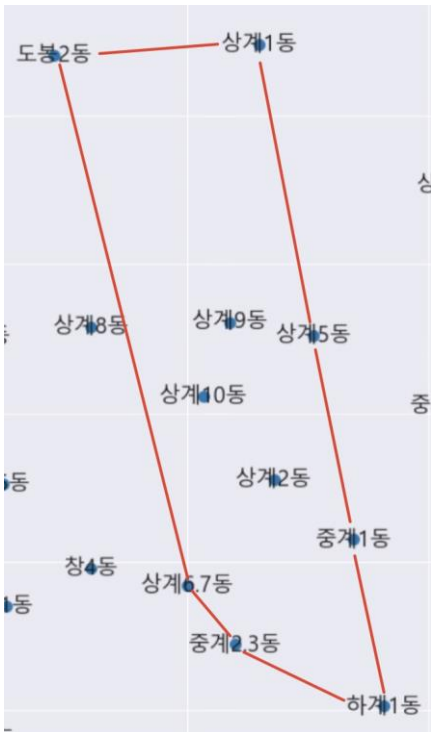
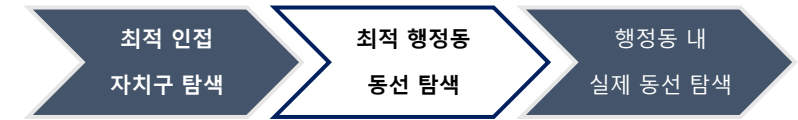
	dest1	dest2	moving
682	상계6.7동	하계1동	3311.0
542	상계1동	상계6.7동	3171.0
674	상계6.7동	중계2.3동	2080.0
568	상계1동	하계1동	1854.0
997	중계2.3동	하계1동	1679.0
146	도봉2동	상계6.7동	1673.0
172	도봉2동	하계1동	1634.0
598	상계2동	하계1동	1422.0
428	번3동	인수동	1364.0
61	공릉2동	상계6.7동	1331.0



**\*\* 각 Node의 위치는 실제 지리적 위치 \*\***  
Weight가 높은 상위 7개의 Edge.  
(= 유동량이 많은 상위 7개의 동선)  
이 동선만 시각화 하면 어떻게  
연결된 큰 두개의 삼각형이 보여진다!

도봉2동 - 하계1동 - 중계2.3동  
- 상계6.7동 - 상계1동이 연결됨!  
그 동선에 있는 다른 동까지 연결 가능!  
다음 슬라이드에 이를 정리!

# 최적 행정동 동선 탐색



이와 같이 다음 5개의 행정동을  
순환하는 노선 구축 가능!  
(동선에서 겹쳐지는 중계1동, 상계 5동도  
추후 추가 가능!)

도봉2동 - 상계6.7동 - 중계2.3동 - 하계1동 - 상계1동

현재 이 지역이 유동량은 가장 많다!  
이어서, 이 동선에 거주하는 장애인 수 및  
이 동선에 존재하는 복지시설 현황을 확인

행정동의 경우 인접 행정동의 정의와,  
개수를 선택하는 것이 매우 복잡한 문제.  
따라서, 앞선 PCA Score와 다른 방법을 채택

행정동 별 장애인 수 및 시설 수

dong	number_of_disabled	welfare_facility
도봉2동	1480	7.0
하계1동	2129	4.0
상계1동	1901	4.0
상계6.7동	1275	8.0
중계2.3동	3215	7.0
계	10,000	30

해당 지역 장애인 수

Rank	si	dong
Rank	si	dong
0	서울시	중계2.3동
3	서울시	하계1동
6	서울시	상계1동
13	서울시	도봉2동
22	서울시	상계6.7동

해당 지역 복지시설 수

Rank	si	dong
Rank	si	dong
3	서울시	상계6.7동
5	서울시	중계2.3동
6	서울시	도봉2동
18	서울시	하계1동
21	서울시	상계1동

이 지역들의 장애인 수 & 복지시설 수가 타 지역보다 유의미하게 많은가?

데이터의 분포를 모르기 때문에, 각각이 동일 분포에서

생성되었다고 가정을 하고, 비모수 방법 사용!

노원구, 도봉구, 강북구 내 46개의 행정동 중

임의의 3개의 행정동의

a. 장애인 수 합 및 그들의 Rank 합을 계산

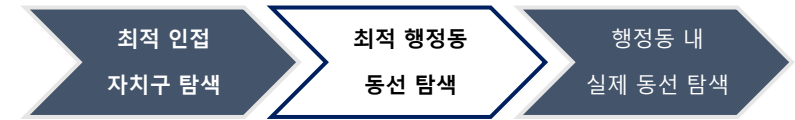
b. 복지시설 수 합 및 그들의 Rank 합을 계산

총  $\binom{46}{3} = 1370754$ 개 의 경우의 수 발생

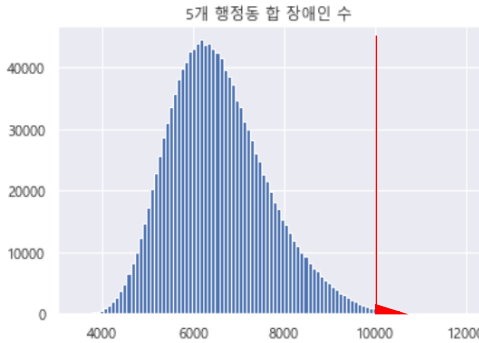
= 실제 값의 합과 Rank의 합에 대한 분포를 그려서,

해당 지역의 Quantile을 확인하자!

# 최적 행정동 동선 탐색



실제 복지시설 수  
및 장애인 수  
Quantile



동선으로 선택한 5개 행정동(도봉2동, 상계6.7동, 중계2.3동, 하계1동, 상계1동)의 장애인 수(10,000명), 복지시설 수(30)는 임의의 5개 행정동 합인 분포에서 붉은 색 삼각형에 해당하므로, 직관적으로 매우 높은 Quantile임을 확인 가능.

\*\* 컴퓨팅 파워의 문제로 Kernel Density Estimation을 시행하지 못함.

\*\* CLT에 의한 정규 근사가 잘 되지 않아 Approximate Distribution을 구하지 못하여 구체적인 Quantile 값 제시 불가.

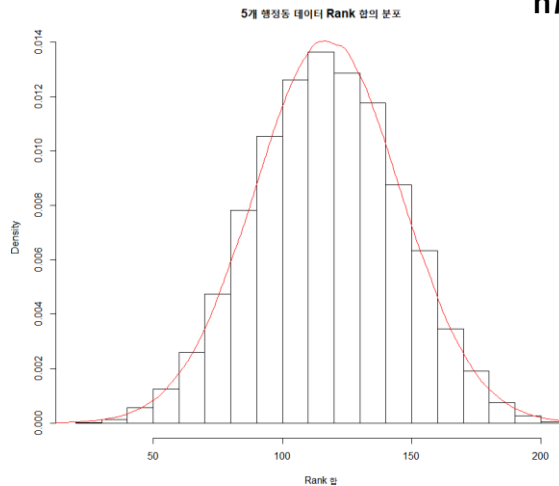
복지시설 수  
및 장애인 수  
Rank Quantile

해당 지역 장애인 수 Rank

Rank	si	dong
0	서울시	중계2.3동
3	서울시	하계1동
6	서울시	상계1동
13	서울시	도봉2동
22	서울시	상계6.7동

해당 지역 복지시설 수 Rank

Rank	si	dong
3	서울시	상계6.7동
5	서울시	중계2.3동
6	서울시	도봉2동
18	서울시	하계1동
21	서울시	상계1동



n개 Rank의 합에 대해서는 Approx 분포와 Quantile 제시 가능.

$$Y \sim N\left(\frac{n(N+1)}{2}, \frac{n(N+1)(N-n)}{12}\right) \sim N(117.5, 28.336^2)$$

\*\* n = 5 표본의 수, N = 46 전체 데이터 수  
이는 좌측의 히스토그램을 통해 확인 가능.

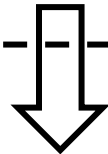
동선으로 선택한 5개 지역의  
장애인의 수에 대한 Rank 합 : 49  
복지시설 수에 대한 Rank 합 : 58

\*\* Pandas는 0부터 index를 출력하므로 이전 슬라이드의  
Index에서 + 1씩 진행해서 Rank를 구했다. \*\*

```

> print(paste('장애인 수 랭크합 Quantile', pnorm(49,117.5,28.336)))
[1] "장애인 수 랭크합 Quantile 0.00781549471816914"
> print(paste('복지시설 수 랭크합 Quantile', pnorm(58,117.5,28.336)))
[1] "복지시설 수 랭크합 Quantile 0.0178731147770434"
  
```

장애인의 수에 대한 Rank 합의 Quantile은 0.0078  
복지시설 수에 대한 Rank 합의 Quantile은 : 0.018



두 결론을 종합했을 때,  
이에 대한 통계적 지식 부족으로 정확한 가설  
검정 진행 불가. 따라서 직관적 해석 진행.  
직관적으로 두 Quantile을 종합하여, 이 지역의  
장애인의 수와 복지시설 수가  
유의미하게 크다고 결론을 내림.  
따라서, 동선에 들어갈 행정동은

**도봉2동 상계6.7동 중계2.3동  
하계1동 상계1동**

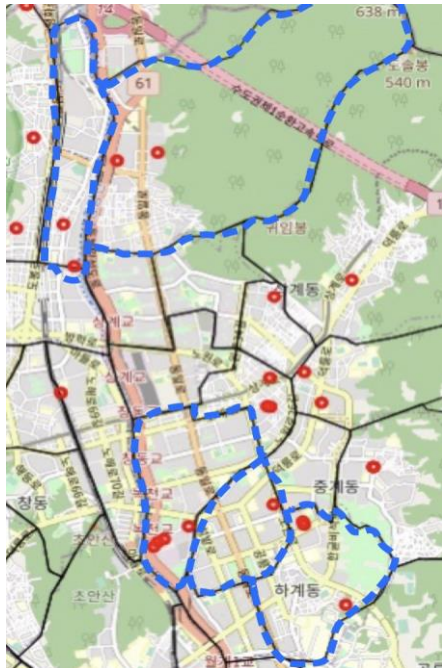


# 행정동 내 실제 동선 탐색

최적 인접  
자치구 탐색

최적 행정동  
동선 탐색

행정동 내  
실제 동선 탐색



- 붉은색 지점이 장애인의 복지 시설의 위치
- 파란색 점선이 동선으로 선정된 5개 행정동

장애인들이 해당 '행정동'에 거주하는 것은 알지만, 구체적으로 어느 아파트, 어느 빌라에 거주하는지는 알 수 없음.

**\*\* 관련부처에 문의한 결과 개인정보때문에 제공 불가**

보다 구체적인 거주지 정보를 획득하기 힘드므로.

1. 실제 거주지는 동선에 반영 불가
2. 우선 장애인 관련시설을 순회하는 셔틀버스를 구축해야 한다.



**하지만** 시설을 바로 앞까지 버스가 운행되는 것은 제한적.  
따라서 우선 대로변으로 정류장을 설정하도록.

본 프로젝트에서는, 동선을 이와 같이 구축했으나,  
더욱 적합한 동선 선택을 위해서는, 해당 지역 거주 장애인들을 대상으로 **수요조사를 진행하는 것도 하나의 방법!**

행정동 단위까지 그 범위를 좁혔으니, 해당 행정동에 거주중인 장애인들의 의견을 반영하는 것은 충분히 현실성이 있음.  
이들의 **주요 이용 시설을 토대로 동선을 구축하면 더욱 효과적!**

실제 중랑구에서는 이러한 공청회를 진행중!  
본 프로젝트에서 제시한 행정동은 두 개 자치구가 혼합되어 있으므로 서울시에서 주최하는 것이 바람직

## 지자체 소식

노약자 등 교통약자를 위한 무료셔틀버스 노선 변경에 따른 주민의견 제출 안내

서울특별시 중랑구

우리구에서는 노약자 등 교통약자의 이동편의 증진 및 사회참여 확대를 위한 무료셔틀버스의 효율:

대한 주민의견을 접수받고 있습니다.

☐ 노선변경 개요

○ 변경일자: 2020. 10. 19(월)부터

# 행정동 내 실제 동선 탐색

최적 인접  
자치구 탐색

최적 행정동  
동선 탐색

행정동 내  
실제 동선 탐색

해당 노선은 사회복지시설인근의 대로로 설정

★ 는 실제 복지시설의 위치

이 두 노선을 순환하는 노선을 다양한 시간대에 운영하는 것도 가능!

\* N 포털사이트 지도 App 활용

1. 한국봉사회 북부종합사회복지관
2. 경희동인한의원
3. 도봉2동 주민센터
4. 자운초등학교
5. 상계사회체육센터
6. 롯데우성아파트
7. 대진고등학교

- 총 거리 : 11.4km  
- 예상소요시간 : 1시간



1. 대진고등학교
2. 서라벌고등학교
3. 그린코아빌딩
4. 중계중학교
5. 대명하우스
6. 한국봉사회 북부종합사회복지관

- 총 거리 : 7.7km  
- 예상소요시간 : 50분



# 결론 및 개선사항

## 1. 최종적으로 제시한 동선과 그 실효성은 다음과 같다.

한국봉사회 북부종합사회복지관 - 경희동인한의원 - 도봉2동 주민센터 - 자운초등학교 - 상계사회체육센터 - 롯데우성아파트 - 대진고등학교 / 총거리 11.4km / 예상 소요시간 1시간  
대진고등학교 - 서라벌 고등학교 - 그린코아빌딩 - 중계중학교 - 대명하우스 - 한국봉사회 북부종합사회복지관 / 총 거리 7.7km / 예상 소요시간 50분

1.1. 관련부처 문인결과, 현재 코로나로 인하여, 서울시에서는 아직 장애인 셔틀버스의 재운영을 검토하지 않고 있음.

1.2. 그러나 장애인 전용 교통수단의 부족을 서울시에서도 인지하고 있으므로, 하나의 서브자료로 참고 가능할 것으로 기대됨.

## 2. 본 프로젝트는 다음과 같은 한계를 존재한다.

2.1. 행정동 탐색 과정에서 비모수 방법을 통한 직관적 해석이 아닌, Quantile을 활용한 가설 검정을 실시하는 것이 더욱 설득력이 있음.

2.2. 행정동 내부의 구체적 동선 탐색 시, 해당 지역 장애인들을 대상으로 설문조사를 실시하면 더욱 실질적인 수요를 반영할 수 있을 것으로 예상됨.

# 감사합니다.

