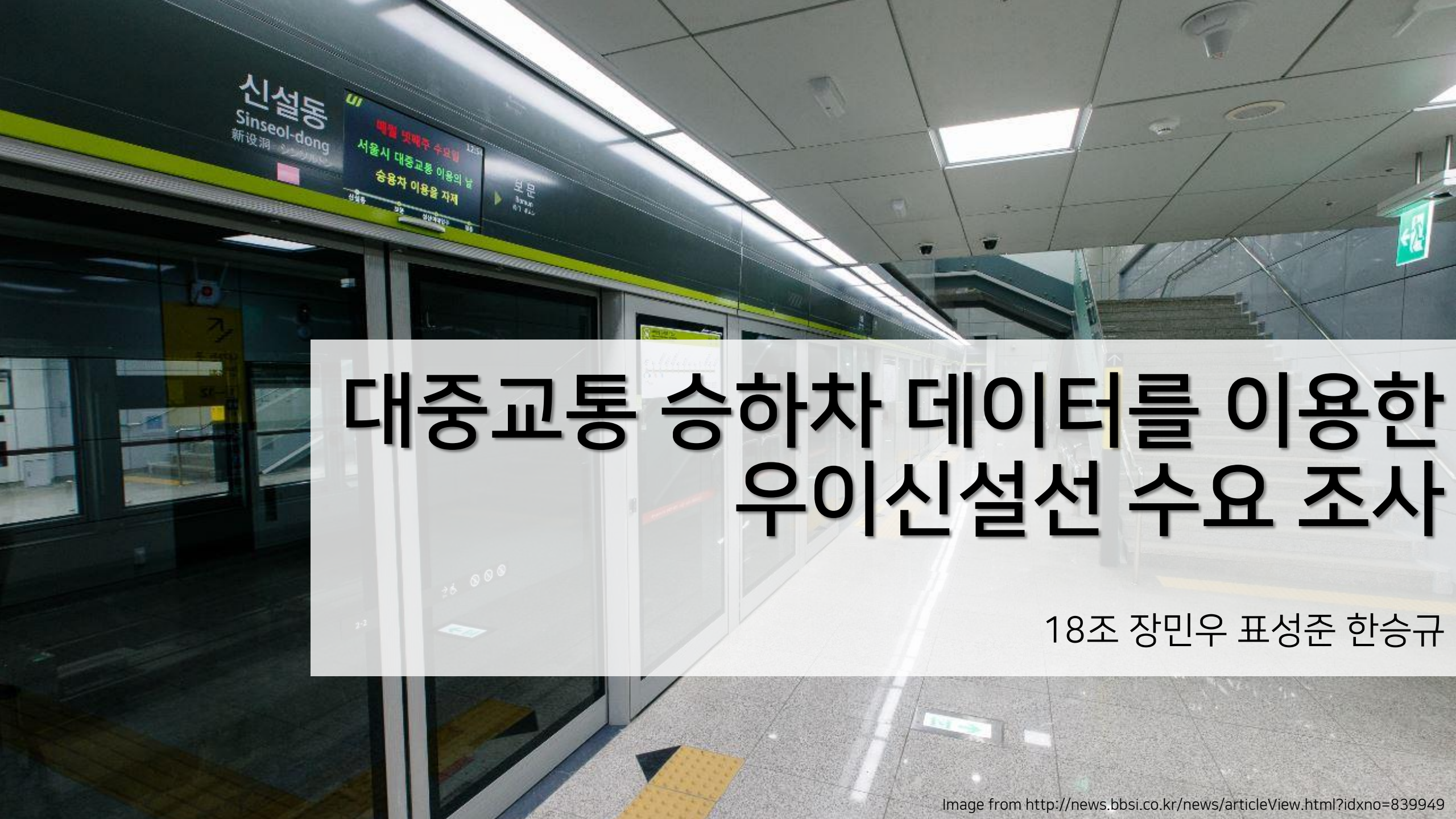


대중교통 승하차 데이터를 이용한 우이신설선 수요 조사

18조 장민우 표성준 한승규

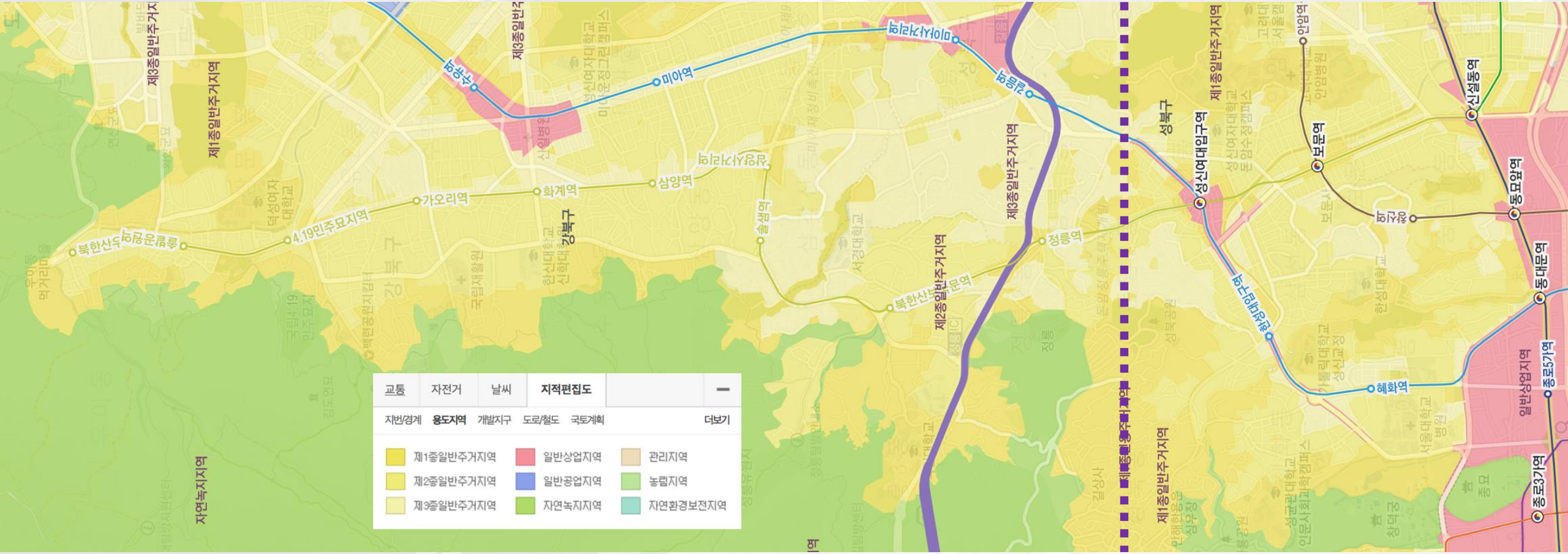
Image from http://news.bbsi.co.kr/news/articleView.html?idxno=839949

우이신설선 개요

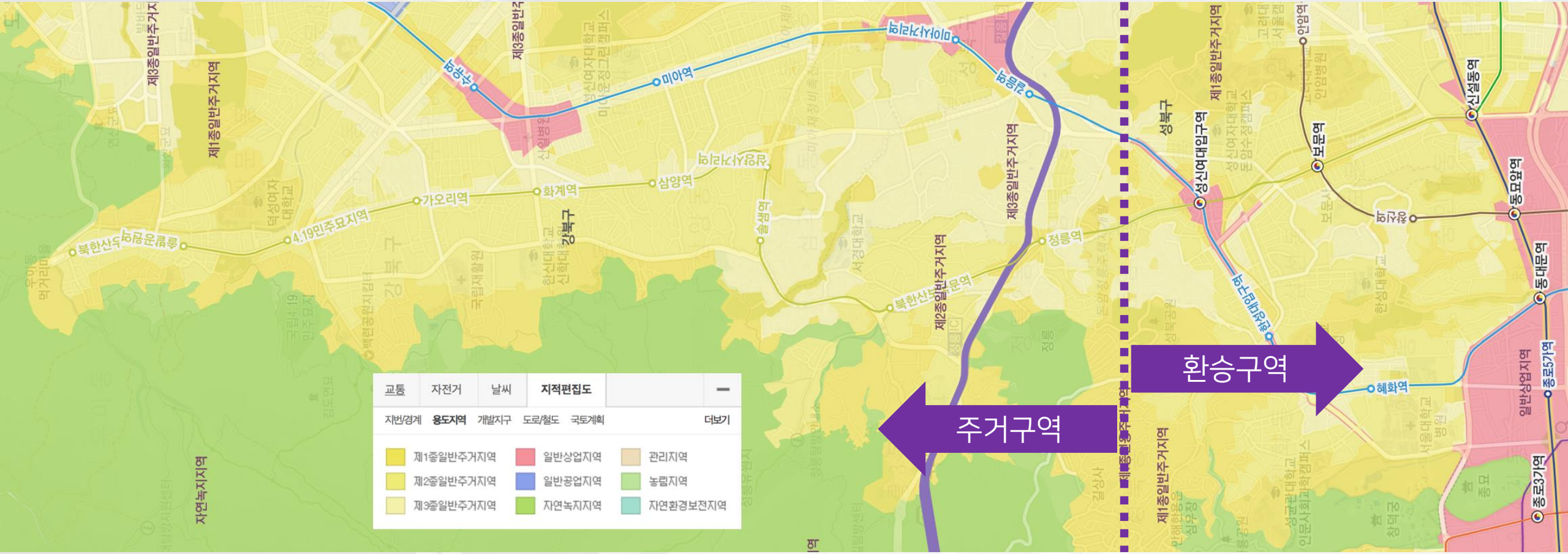
- 2017년 9월 개통
- 수도권 통합요금제
- 전철 음영지역 해소

http://www.hankookilbo.com/v/45ee35570b1c4325a6f6c4f7dd519ebd

우이신설선 노선의 특징


교통
자전거
날씨
지적편집도
지번/경계
용도지역
개발지구
도로/철도
국토계획
더보기
제1종일반주거지역
제2종일반주거지역
제3종일반주거지역
일반상업지역
일반공업지역
자연녹지지역
관리지역
농림지역
자연환경보전지역
북한산우이역
4.19민주묘지역
가오리역
화계역
삼양역
솔샘역
미아역
수유역
북한산보문역
정릉역
길음역
미아사거리역
성신여대입구역
보문역
안암역
신설동역
동묘앞역
동대문역
종로5가역
종로3가역
혜화역
성북구
강북구
자연녹지지역
일반상업지역

우이신설선 노선의 특징


교통
자전거
날씨
지적편집도
지번/경계
용도지역
개발지구
도로/철도
국토계획
더보기
제1종일반주거지역
제2종일반주거지역
제3종일반주거지역
일반상업지역
일반공업지역
자연녹지지역
관리지역
농림지역
자연환경보전지역
주거구역
환승구역
북한산우이역
4.19민주묘지역
가오리역
화계역
삼양역
솔샘역
미아역
수유역
북한산보문역
정릉역
길음역
성신여대입구역
보문역
안암역
신설동역
동묘앞역
동대문역
종로5가역
종로3가역
혜화역
강북구
성북구
자연녹지지역
일반상업지역

우이신설선 노선의 특징

주거구역
10개역

환승구역
3개역

데이터 수집

- 서울 열린데이터 광장
  - 서울특별시 버스 관련 정보 (승하차통계, 정류장일람 등)
  - 검증용 우이신설선 승하차 통계
  - http://data.seoul.go.kr

- 서울시 교통정보 시스템(TOPIS)
  - http://topis.seoul.go.kr

- 다음지도
  - 버스 정류장 정보 수집
  - http://map.daum.net

데이터 수집

- 서울시 지하철 주변 버스정류장 일람
  - 총 1829 row

	A	B	C	D	E	F	G	H
1	전철역코드	외부코드	전철역명	호선	정류장명	정류장ID	X좌표	Y좌표
902	321	331	충무로	3	대한극장앞	2153	199551	451304
903	423	423	충무로	4	퇴계로3가한옥마을한	2151	199270	451303
904	423	423	충무로	4	퇴계로3가한옥마을한	2152	199336	451265
905	423	423	충무로	4	초동	2155	199392	451539
906	423	423	충무로	4	매일경제	2242	199414	451345
907	423	423	충무로	4	대한극장앞	2153	199551	451304
908	2737	735	내방	7	방배보건분소	22231	199479	443025
909	2737	735	내방	7	방배보건분소	22230	199505	443042
910	2631	630	이태원	6	해밀턴호텔	3190	199382	448340
911	2631	630	이태원	6	해밀턴호텔	3191	199396	448314
912	2631	630	이태원	6	보광동입구	3283	199479	448241
913	2631	630	이태원	6	보광동입구	3284	199495	448242
914	4122	922	신반포	9	반포2동주민센터	22201	199506	444823
915	4122	922	신반포	9	반포2동주민센터	22364	199559	444842
916	4122	922	신반포	9	신반포	22791	199594	444888
917	4122	922	신반포	9	반포2동주민센터	22363	199765	444951

데이터 수집

- 2017년 8월 서울 버스 정류장별 노선별 승하차량 통계
  - 100만 row 이상: 하지만 노선 운행 순서대로의 배치가 아님

1	사용일자	노선번호	노선명	표준버스정류	버스정류장AF	역명	승차총승객수	하차총승객수	등록일자
288	20170801	2112	2112번(면목동	105000141	6227	답십리2동한양	46	51	20170804
289	20170801	2112	2112번(면목동	105000142	6228	답십리2동한양	68	57	20170804
290	20170801	2112	2112번(면목동	105000057	6143	휘경여중고휘	39	42	20170804
291	20170801	2112	2112번(면목동	105000161	6247	동경호텔앞	29	280	20170804
292	20170801	2112	2112번(면목동	105000149	6235	국민은행장안	310	162	20170804
293	20170801	2112	2112번(면목동	105000150	6236	국민은행장안	78	223	20170804
294	20170801	109	109번(우이동	108000075	9163	삼양초등학교	43	121	20170804
295	20170801	109	109번(우이동	108000073	9161	롯데마트삼양	170	307	20170804
296	20170801	109	109번(우이동	108000071	9159	빨래골입구	112	266	20170804
297	20170801	109	109번(우이동	108000069	9157	화계사입구	43	161	20170804
298	20170801	109	109번(우이동	108000047	9135	우이초등학교	80	149	20170804
299	20170801	109	109번(우이동	108000033	9121	인수동장미원	101	161	20170804
300	20170801	109	109번(우이동	108000014	9102	우이동도선사	82	3	20170804
301	20170801	109	109번(우이동	108000015	9103	우이동성원아	56	2	20170804
302	20170801	109	109번(우이동	108000019	9107	덕성여대솔밭	229	44	20170804
303	20170801	109	109번(우이동	108000021	9109	국립4.19민주.	35	198	20170804
304	20170801	109	109번(우이동	109000133	10218	북서울교회	13	113	20170804
305	20170801	109	109번(우이동	108000016	9104	우이동성원아	4	69	20170804
306	20170801	109	109번(우이동	990072600	~	우이동종점(7	0	332	20170804
307	20170801	109	109번(우이동	109000236	10325	동아운수앞	4	72	20170804
308	20170801	2211	2211번(면목동	105000127	6213	장안2동주민신	763	200	20170804
309	20170801	2211	2211번(면목동	105000066	6152	동아.위더스빌	189	252	20170804
310	20170801	2211	2211번(면목동	105000058	6144	장안동현대아	177	233	20170804
311	20170801	2211	2211번(면목동	105000288	6417	동일스위트리	152	155	20170804
312	20170801	2211	2211번(면목동	105000052	6136	동일스위트리	159	179	20170804

데이터 수집

- 서울시 교통정보 시스템 (TOPIS)
  - 버스 운행 정보 (경유 정류소 순차 일람)

topis.seoul.go.kr/map/openPopBusRoute.do?routId=100100031

152 번

운행정보

프린트하기

기점	화계사		종점	삼막사사거리		버스타입	간선	
버스회사	동아운수		전화	02-908-4551~4		E-mail	dong-a-2000@hanmail.net	
배차간격	평일	7분	첫차	평일	04:00	막차	평일	22:20
	토요일	7분		토요일	04:00		토요일	22:20
	공휴일	7분		공휴일	04:00		공휴일	22:20

순서	경유정류소
1	수유중학교혜화여고(09211)
2	서울정인학교(09213)
3	구세군강북종합사회복지관입구(09215)
4	미양중고등학교.팔팔·미송경로당(09218)

순서	경유정류소
45	장승배기역(20151)
46	상도초등학교입구(20150)
47	성대시장(20147)
48	신대방삼거리(20186)
49	보라매병원입구(20188)

순서	경유정류소
88	청화병원(20193)
89	동작구청.노량진초등학교앞(20191)
90	노량진역(20114)
91	노들역(20013)
92	노들섬(03341)

데이터 분석 과정

1. 우이신설선 각 역 근처의 모든 버스 정류장 정리
   - 환승구역: 서울시에서 제공한 데이터를 바탕으로, 오류가 있는 항목을 수정하여 작성
   - 주거구역: 각 역 기준 반경 350m 이내의 모든 버스 정류장 수집
2. 노선별 승하차량 통계를 통하여 두 구역을 잇는 노선들 정리
3. 해당 노선들의 정류장 목록을 운행순서대로 수집
4. 각 버스 노선 중 어느 구간이 우이신설선과 중복되는지 검출
5. 중복 구간의 승하차량을 기준으로 해당 구간을 이동한 승객 수 파악
6. 모든 구간에 대해 합산

데이터 분석 결과

```
In [17]: total = 0
         for line in onboardsection.keys():
             thisline = onboardsection[line]
             for i in thisline['board_u']:
                 for j in thisline['getoff_u']:
                     total += dic_inout[line][i[1]][j[1]]
             for i in thisline['board_d']:
                 for j in thisline['getoff_d']:
                     total += dic_inout[line][i[1]][j[1]]

         print(total)

         377020
```

예상 승(하)차 총 승객수
377,020

한달에 38만명이 이용하는 노선: 사실상 개통해서는 안 될 수준의 수요

(경기 광주시, 여주시, 판교시를 잇는 경강선의 11월 승하차량: 75만명)

데이터 분석 결과:실제 데이터와의 비교

```
In [17]: total = 0
         for line in onboardsection.keys():
             thisline = onboardsection[line]
             for i in thisline['board_u']:
                 for j in thisline['getoff_u']:
                     total += dic_inout[line][i[1]][j[1]]
             for i in thisline['board_d']:
                 for j in thisline['getoff_d']:
                     total += dic_inout[line][i[1]][j[1]]

         print(total)

         377020
```

예상 승(하)차 총 승객수
377,020

```
In [18]: novemberSubwayUsers = pd.read_csv('CARD_SUBWAY_MONTH_201711.csv')
         novemberSubwayUsers.head()
         UllineUsers11 = novemberSubwayUsers.where(novemberSubwayUsers['노선명'] == '우이신설선').dropna()
         board11 = UllineUsers11['승차총승객수'].sum()
         getoff11 = UllineUsers11['하차총승객수'].sum()
         print((board11 + getoff11) / 2.0)

         1172045.0
```

실제 승(하)차 총 승객수
약 1,000,000

```
In [19]: septemberSubwayUsers = pd.read_csv('CARD_SUBWAY_MONTH_201709.csv')
         septemberSubwayUsers.head()
         UllineUsers9 = septemberSubwayUsers.where(septemberSubwayUsers['노선명'] == '우이신설선').dropna()
         board9 = UllineUsers9['승차총승객수'].sum()
         getoff9 = UllineUsers9['하차총승객수'].sum()
         print((board9 + getoff9) / 2.0)

         1065903.5
```

최소
3배
차이

데이터 분석 고찰

- 승하차량 통계의 근본적인 한계
  - 승하차량인 정적인 통계량으로 동적인 경로를 분석하는 모델링의 부정확성
  - 대중교통 외 자가용이나 도보 중심의 승객 수는 집계가 안 됨
  - 내릴 때 카드를 반드시 태그하지 않아도 된다는 점에 따른 오차
  - 날짜 구분에 따른 승하차량 변동 (자정을 기준으로 분리)

- 데이터 자체의 불충분함
  - 서울시에서 제공하는 대부분의 데이터에 오류가 있었음
    - 중복 정류장 반복 경유에 대한 승하차량 통계 미분리 (자체 모델링으로 분리)
    - 지하철 역 근처 버스 정류장 API의 오류 (지도를 통해 수기로 정리)
    - 버스 기-종점의 불확실성 (순환형 노선 등에 대한 오차 필연적으로 발생)

데이터 분석 고찰

- 그 외의 오차 원인
  - 지하철은 무임승차가 존재: 노인들은 버스보다는 전철을 이용하려는 경향이 큼
  - 전철 개통에 따른 신규 노인 수요 창출의 효과를 무시하지 못함
    - (실제로 개통 초기 무임승차 비율이 40%에 가깝다는 기사가 등장)
  - 근처 대학의 통학 수요가 움직일 가능성
    - 분석에 집계되지 않은 경로로 통학하던 학생들이 개통 이후 우이신설선 쪽으로 통학경로를 변경할 가능성 (근처에 대학만 최소 3개)

추후 발전 방향

- 약 100만명 정도 탑승한 현 실적도 서울시의 예측의 50%가 안 되는 수요
  → 희망적이지 않은, 최소한의 보수적인 수요 예측으로써의 가치가 있음

- 앞에서 제시한 오차의 원인들 을 해결할수록 보수적인 예측량을 보다 현실화 할 수 있음

- 개인의 사생활이 침범되지 않는 비 추적형 데이터를 이용하므로, 정확도가 개선된다면 충분히 활용의 여지가 있음

감사합니다.