

목차

2. 신설 노선 개통

3. 연세로 차 없는 거리 해제



1. 데이터 소개



1. 지하철 데이터 소개

1 기본 데이터

	사용일자	노선명	역명	승차총승객수	하차총승객수	등록일자
0	20150101	2호선	낙성대	14586.0	14889.0	20151217
1	20150101	2호선	사당	19233.0	20298.0	20151217
2	20150101	2호선	방배	5920.0	6065.0	20151217
3	20150101	2호선	서초	4379.0	4120.0	20151217
4	20150101	분당선	선정릉	1972.0	1828.0	20151217

- 변수: 사용일자(실제 이용 일자), 노선명, 역명, 승,하차 총 승객수, 등록일자(DB 등록 일자)
- 기간: 2015년 1월 1일 ~ 2022년 11월 30일
- 크기: 16000여 개의 행, 6 개의 열로 구성

1. 지하철 데이터 소개

2 시간별, 승객 유형별 데이터

연 번	수솜일자	호 선	고유역번호(외 부역코드)	역명	승하차 구분	승객 유형	06시간대 이전	06-07시 간대	07-08시 간대	•••	15-16시 간대	16-17시 간대	17-18시 간대	18-19시 간대	19-20시 간대	20-21시 간대	21-22시 간대	22-23시 간대	23-24시 간대	24시간 대이후
1	2022년 7월 1 일 금요일	1	150	서울 역	승차	일반	220	338	1385		2635	2887	4414	7 <mark>5</mark> 59	3081	2415	2400	1919	989	41
2	2022년 7월 1 일 금요일	1	150	서울 역	승차	어린 이	0	1	2		9	17	7	8	8	10	4	10	2	0
3	2022년 7월 1 일 금요일	1	150	서울 역	승차	중고 생	0	0	0		2	1	1	0	0	0	2	1	0	0
4	2022년 7월 1 일 금요일	1	150	서울 역	승차	청소 년	0	5	24	'n	76	48	34	53	18	36	26	22	19	1
5	2022년 7월 1 일 금요일	1	<mark>150</mark>	서울 역	승차	우대 권	98	102	216	***	626	688	594	489	308	252	210	128	43	2

● 변수: 수송일자(실제 이용 일자), 호선, 역명, 승하차 구분, 승객 유형, 시간대별 승,하차 인원 등

• 기간: 2022년 7월 1일 ~ 2022년 11월 30일

• 크기: 535000여 개의 행, 27 개의 열로 구성

2. 버스 데이터 소개

1 기본 데이터

	사용일자	노선번호	노선명	버스정류장ARS번호	역명	승차총승객수	하차총승객수	등록일자
0	20150101	100	100번(하계동~용산구청)	11374	서울시립북서울미술관	22	2	20151228
1	20150101	6631	6631번(강서공영차고지~영등포역)	ex	당산동진로아파트	32	32	20151228
2	20150101	105	105번(상계동~서울역)	11416	하계역	65	199	20151228
3	20150101	6631	6631번(강서공영차고지~영등포역)	~:	건양병원.김안과병원	143	28	20151228
4	20150101	102	102번(상계동~동대문)	10249	KEB하나은행창동지점	181	171	20151228

• 변수: 사용일자(실제 이용 일자), 노선번호, ARS번호, 승,하차 총 승객수, 등록일자(DB 등록 일자) 등

• 기간: 2015년 1월 1일 ~ 2022년 12월 31일

• 크기: 1,150,000 여 개의 행, 8 개의 열로 구성

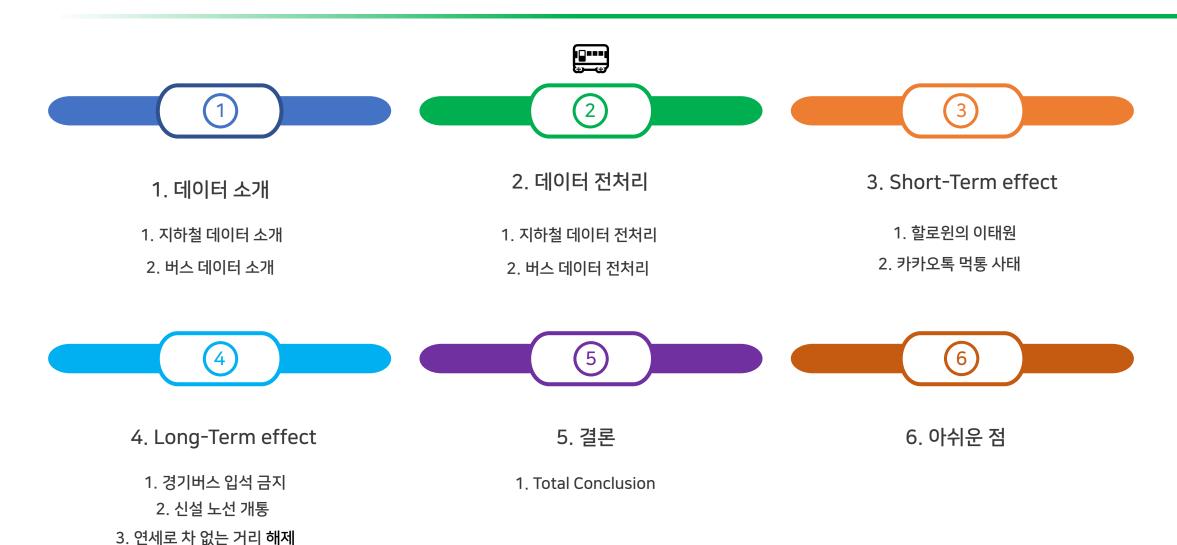
2. 버스 데이터 소개

② 경기 버스 데이터

	년도	월	지역	노선	시,종점 (출발,도착지점)	일시	시간	승차	하차	전체
0	2022	12	화성시	1006	수원대학교-강남역나라빌딩앞	평일	05시~06시	13	0	13
1	2022	12	화성시	1006	수원대학교-강남역나라빌딩앞	평일	06시~07시	62	19	81
2	2022	12	화성시	1006	수원대학교-강남역나라빌딩앞	평일	07시~08시	94	72	166
3	2022	12	화성시	1006	수원대학교-강남역나라빌딩앞	평일	08시~09시	90	90	180
4	2022	12	화성시	1006	수원대학교-강남역나라빌딩앞	평일	09시~10시	19	83	102

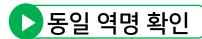
- 변수: 년도, 월, 지역, 노선, 시,종점, 시간, 승,하차 승객 수 등으로 구성
- 기간: 2022년 10월, 12월
- 크기: 600여 개의 행, 10개의 열로 구성

2. 데이터 전처리

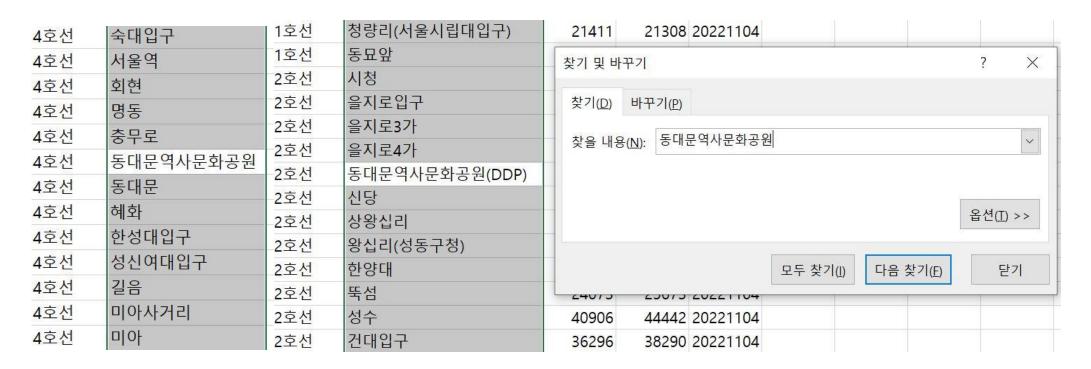


1. 지하철 데이터 전처리





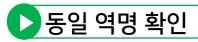




• 공개입찰을 통한 부역명 제외하고 주역명만 고려

1. 지하철 데이터 전처리





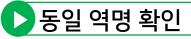


영문명 역단	한국어 역명
4.19 Democratic Cemetery Statio	4.19민주묘지
Sunseong Statio	가능
Garak Market Statio	가락시장
Gasan Digital Complex Statio	가산디지털단지
Gayang Statio	가양
Kaori Statio	가오리
Gajwa Statio	가좌
Gachondae Statio	가천대
Gapyeong Statio	가평
Ganseok Statio	간석

- 한국어 역명은 시각화 과정에서 폰트가 깨짐
- 동적 크롤링을 통해 한국어 역명을 파파고로 번역

1. 지하철 데이터 전처리







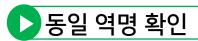
영문명 역명	한국어 역명
4.19 Democratic Cemetery Station	4.19민주묘지
Sunseong Station	가능
Garak Market Station	가락시장
Gasan Digital Complex Station	가산디지털단지
Gayang Station	가양
Kaori Station	가오리
Gajwa Station	가좌
Gachondae Station	가천대
Gapyeong Station	가평
Ganseok Station	간석



- 한국어 역명은 시각화 과정에서 폰트가 깨짐
- 동적 크롤링을 통해 한국어 역명을 파파고로 번역

2. 버스 데이터 전처리





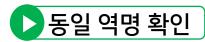


	사용일자	노선번호	노선명	버스정류장ARS번호	역명	승차총승객수	하차총승객수	등록일자
0	20150101	100	100번(하계동~용산구청)	11374	서울시립북서울미술관	22	2	20151228
1	20150101	6631	6631번(강서공영차고지~영등포역)	2 8	당산동진로아파트	32	32	20151228
2	20150101	105	105번(상계동~서울 역)	11416	하계역	65	199	20151228
3	20150101	6631	6631번(강서공영차고지~영등포역)	~	건양병원.김안과병원	143	28	20151228
4	20150101	102	102번(상계동~동대문)	10249	KEB하나은행창동지점	181	171	20151228

• 버스 데이터는 노선번호와 버스정류장 ARS번호를 사용하여 동일 역명이 존재하지 않았음

2. 버스 데이터 전처리







한국어 노선번호	영문명 노선번호
강남01	Gangnam 01
강남02	Gangnam 02
강남03	Gangnam 03
강남05	Gangnam 05
강남06	Gangnam 06
강남06-1	Gangnam 06-1
강남06-2	Gangnam 06-2
강남07	Gangnam 07
강남08	Gangnam 08
강남10	Gangnam 10

- 한국어 역명은 시각화 과정에서 폰트가 깨짐
- 동적 크롤링을 통해 한국어 역명을 파파고로 번역

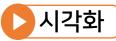
3. Short-Term effect







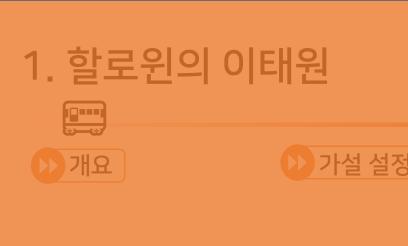








2022년 10월 29일 이태원에는 2020년 코로나 19 확산 이후 3년 만에 사회적 거리두기 없는 핼러윈을 즐기려는 수많은 인파가 몰려들었는데 오후 10시 15분경 서울특별시 용산구 이태원동 119-3번지 일대 해밀톤호텔 옆 보행로 폭이 4m 안팎인 골목에 핼러윈 축제를 즐기기 위해모인 다수의 인파가 뒤엉키면서 159명이 사망하는 등 300명이 넘는 사상자가 발생하였다.



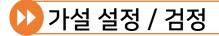




2022년 10월 29일 이태원에는 2020년 코로나 19 확산 이후 3년 만에 사회적 거리두기 없는 핼러윈을 즐기려는 수많은 인파가 몰려들었는데 오후 10시 15분경 서울특별시 용산구 이태원동 119-3번지 일대 해밀톤호텔 옆 보행로 폭어났을 가을 예 핼러윈 축제를 즐기기 위해모인 다수의 인파가 뒤엉키면서 159명이 사망하는 등 300명이 넘는 사상자가 발생하였다.













할로윈 직전 토요일에 승, 하차 승객 수는 유의미한 차이가 없을 것이다.

VS.

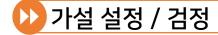


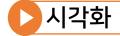
할로윈 직전 토요일에 승, 하차 승객 수가 더 많을 것이다.















<mark>할로윈 직전 토요일</mark>에 승, 하차 승객 수는 유의미한 차이가 없을 것이다.

VS.



할로윈 직전 토요일에 승, 하차 승객 수가 더 많을 것이다.

UsedDate	Line	Station	Board	Alight
20221029	6호선	이태원	48459.0	81362.0
20151031	6호선	이태원	37666.0	71712.0
20171028	6호선	이태원	39662.0	64310.0
20181027	6호선	이태원	40276.0	61902.0
20181013	6호선	이태원	43183.0	54732.0
20171014	6호선	이태원	42625.0	54618.0
20191026	6호선	이태원	38619.0	57844.0
20151017	6호선	이태원	40597.0	48892.0
20161015	6호선	이태원	39703.0	48687.0
20191012	6호선	이태원	38648.0	48537.0



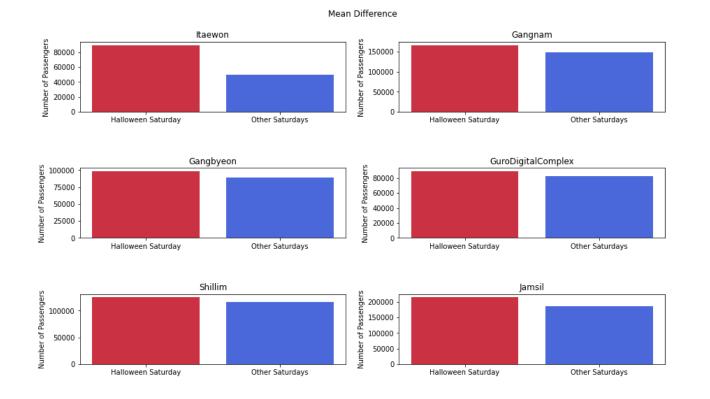






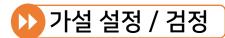


역명	검정통계량	p-value
이태원	3.85845	0.00023
강남	1.17170	0.12462
강변	0.82230	0.20824
구로디지털 단지	1.00514	0.16086
신림	1.03848	0.15308
잠실	1.83357	0.03762



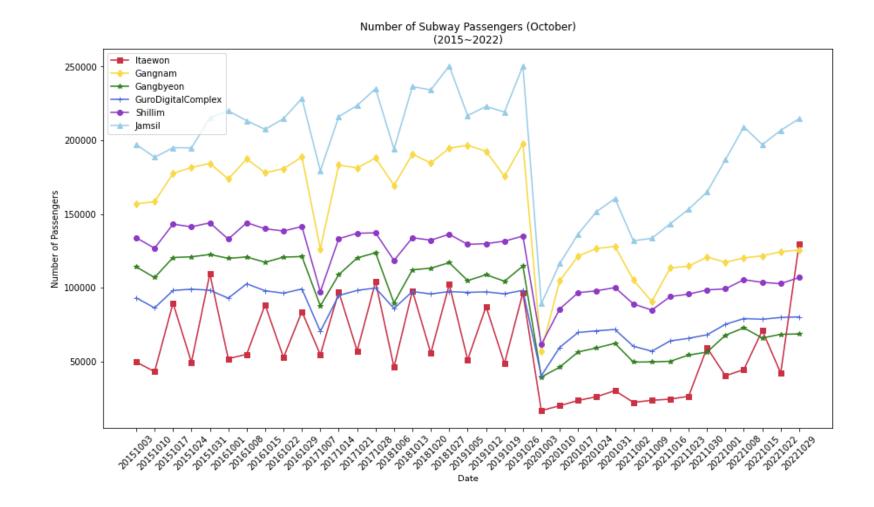








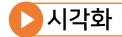




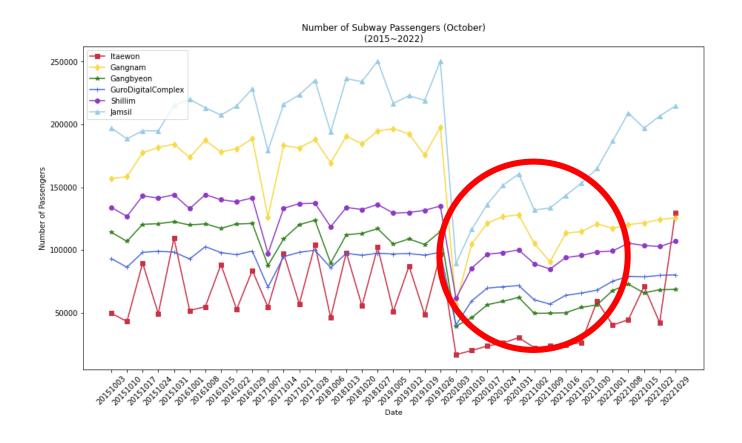


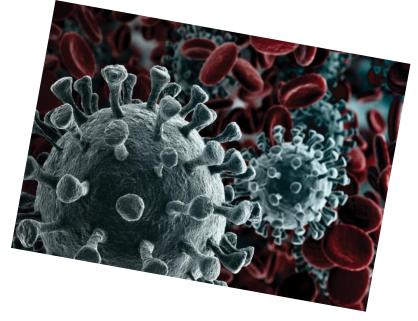












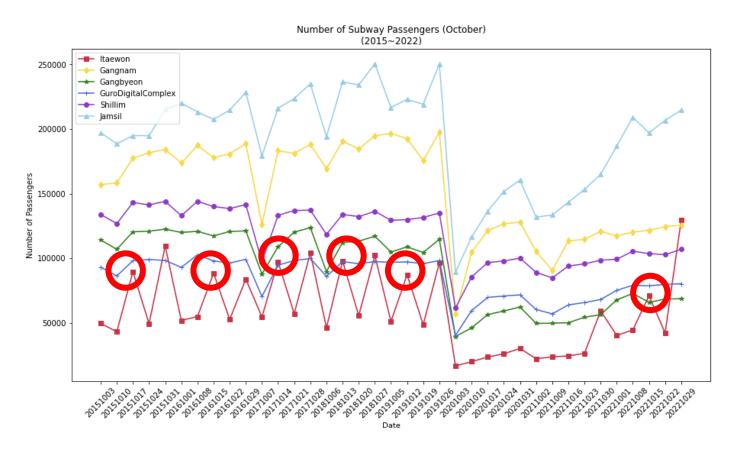






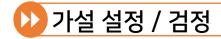
















• 결론 도출

주기적인 통행량 급증 예의 주시 필요



• 필요한 추가 분석

지역별 특징에 따른 밀집도 파악 필요 시간대별 분석 필요













2022년 10월 15일 오후 3시30분부터 카카오 톡 PC 및 모바일에서 서버 오류로 추정되는 문제 로 일부 이용자의 메시지 송·수신이 원활하지 않 은 문제가 발생했습니다. 아울러 카카오 네비·다 음카페·카카오페이지· 다음뉴스등 카카오에서 제공 중인 서비스 대부분 또한 이용 불가한 상황 입니다.













Kakao 서버 오류와 관계 없이 대중교통 사용량은 일정할 것이다.

VS.



Kakao 서버 오류로 인해 KaKao Taxi를 사용하지 못한 사람들로 인해 대중교통 사용량은 증가하였을 것이다.















- 1 Kakao 서버 오류 전 vs 후검정통계량 : -0.2334 / p- value : 0.5896
- (2)Kakao 서버 오류 일주일 전 vs Kakao 서버 오류 후검정통계량 : -0.4933 / p- value : 0.6277
- Kakao 서버 오류 후 vs Kakao 서버 오류 일주일 후 검정통계량 : -0.2383 / p- value : 0.8143

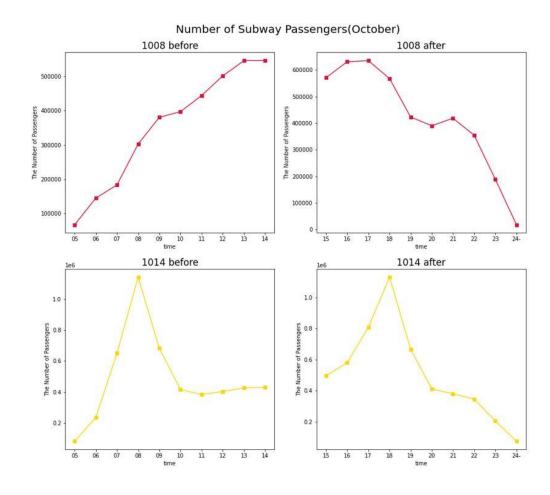












카카오톡 먹통 사태 일주일 전, 하루 전 총 승,하차 승객수

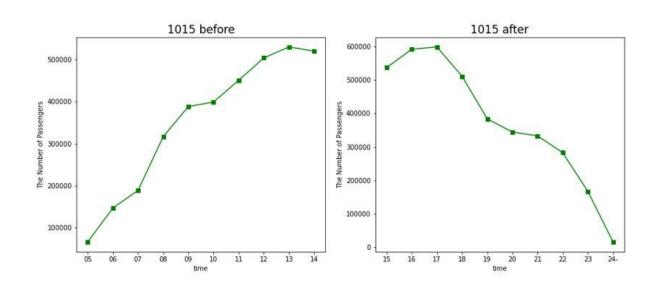




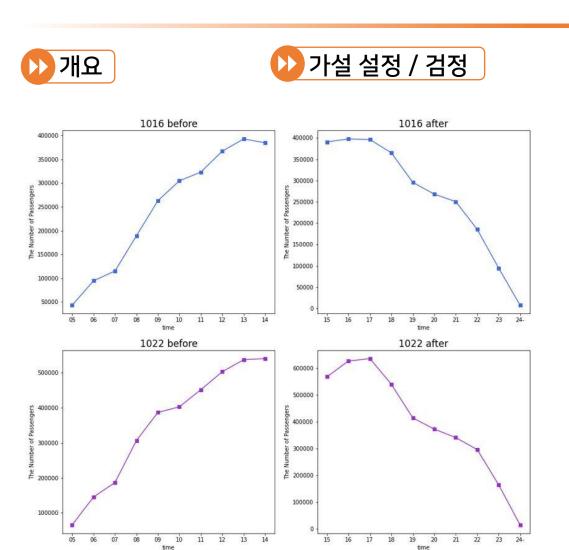








카카오톡 먹통 사태 당일 총 승,하차 승객수









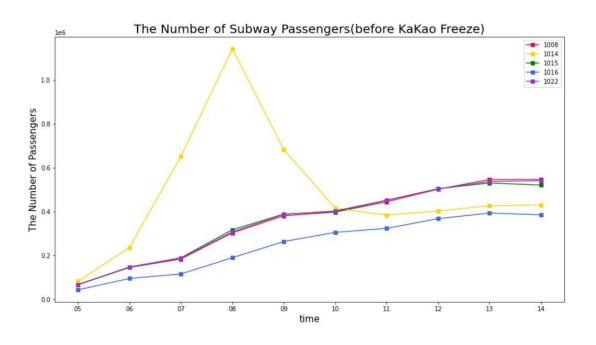
카카오톡 먹통 사태 다음날, 일주일 후 총 승,하차 승객수

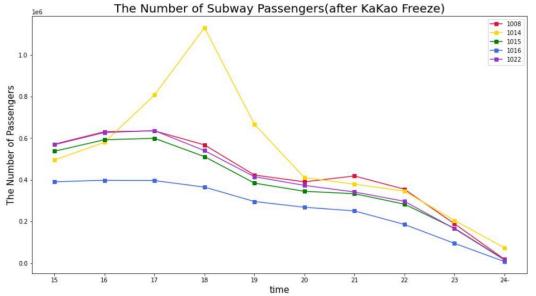




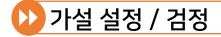
















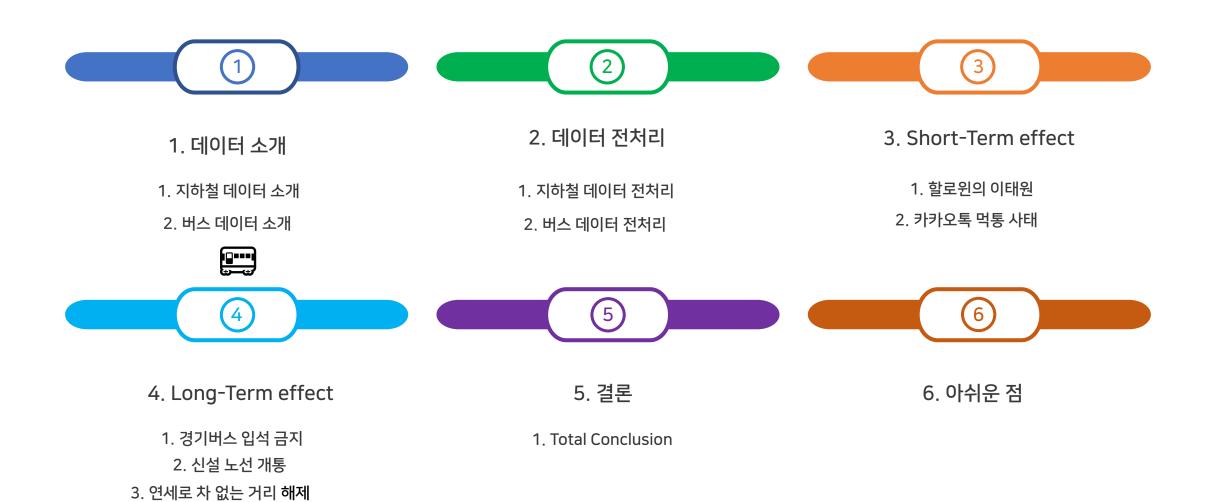
• 결론 도출

카카오톡 플랫폼의 교통분야 영향력 미미



카카오 택시를 대체하는 타 택시 어플 이용했는지 교통 분야에서의 카카오 택시 실제 점유율 분석 필요

4. Long-Term effect

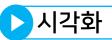










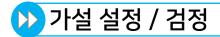


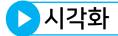


KD운송그룹 계열 14개 버스업체가 운영하는 경기지역 광역버스 입석 승차가 22년 11월 18일부로 전면 중단됐다. 광역버스 입석 탑승 자체가 애초에 금지된 사안이지만 대부분 버스업체는 그동안 출퇴근 때 이용 수요가 집중되는 현실을 고려해 입석 승차를 용인해 왔다. 그러나 최근 10.29 참사 이후 대중교통 등 밀집된 장소에서 안전에 대한 관심이 높아지자 업체 측이 입석 승차 금지를 결정했다.













경기도 버스 입석 금지와 관계 없이 출퇴근 시간대의 승차 인원은 변하 지 않았을 것이다.

VS.



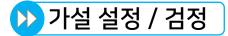
경기도 버스 입석 금지 이후 출퇴근 시 간대의 승차 인원은 유의미하게 증가하 였을 것이다.

	지역	시간	pvalue
0	남양주시	05시~06시	0.368831
1	남양주시	06시~07시	0.426036
2	남양주시	07시~08시	0.317057
3	화성시	05시~06시	1.000000
4	화성시	06시~07시	0.271924
5	화성시	07시~08시	0.025254
6	용인시	05시~06시	0.915333
7	용인시	06시~07시	0.788461
8	용인시	07시~08시	0.221403
9	수원시	05시~06시	0.567951
10	수원시	06시~07시	0.245695
11	수원시	07시~08시	0.963738

	지역	시간	pvalue
0	남양주시	17시~18시	0.706028
1	남양주시	18시~19시	0.011059
2	남양주시	19시~20시	0.272275
3	화성시	17시~18시	0.725444
4	화성시	18시~19시	0.013279
5	화성시	19시~20시	0.471055
6	용인시	17시~18시	0.038738
7	용인시	18시~19시	0.343223
8	용인시	19시~20시	0.676424
9	수원시	17시~18시	0.213287
10	수원시	18시~19시	0.147517
11	수원시	19시~20시	0.115458

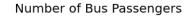


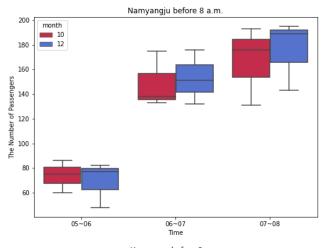


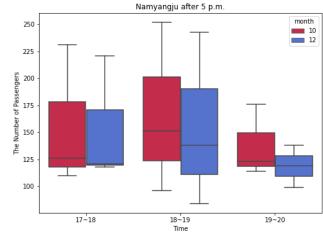


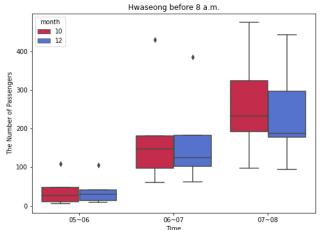


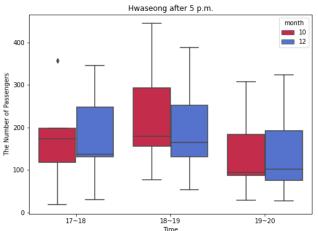








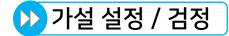




• 남양주시, 화성시에서 출발하는 버스의 출,퇴근 시간 승,하차 승객수

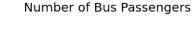


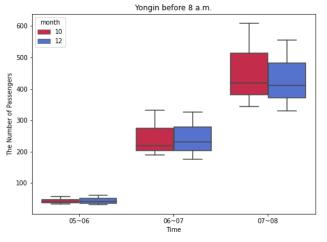


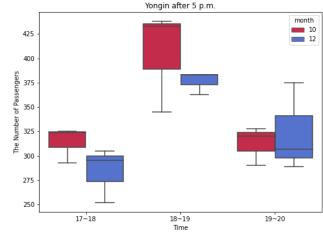


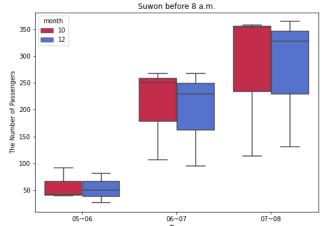


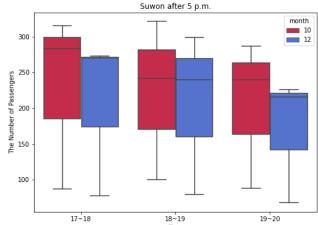












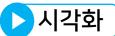
• 용인시, 수원시에서 출발하는 버스의 출,퇴근 시간 승,하차 승객수

1. 경기 버스 입석 금지





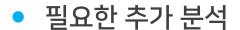






• 결론 도출

입석금지 이후 시간대별 버스 승객수의 통계적 유의미한 변화 확인 불가



버스 노선에 대한 대체재가 존재하는지 해당 버스의 배차 간격이 짧음











지하철 9호선 2단계 구간 노선도









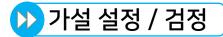












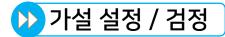


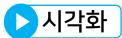


























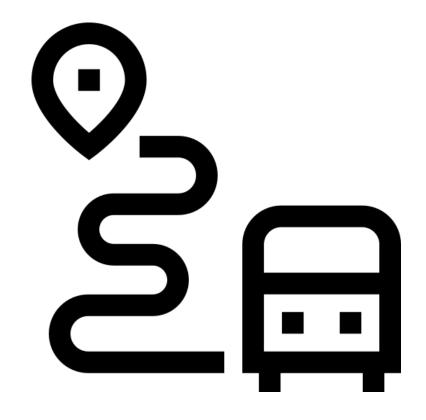


신설 노선의 개통은 장기적으로 주변 노선에 영향을 거의 미치지 못한다

VS.

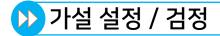


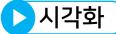
장기적 관점에서 신설 노선을 통해 유입된 승객으로 기존 노선의 이용승객 수가 변화하였을 것이다.













1 9호선 2단계

	Line 2	bundang line
t_stat_st	-2.805859	-0.750324
p_value_st	0.107014	0.531321
t_stat_lt	-1.177832	0.225614
p_value_lt	0.360034	0.842459

2 9호선 3단계

	Line 5	Line 8
t_stat_st	2.450678	2.757984
p_value_st	0.091624	0.070271
t_stat_lt	1.722022	1.339166
p_value_lt	0.183544	0.272940

3 우이신설선 개통

	Line 1	Line 2	Line 4	Line 6
t_stat_st	-3.896951	-3.622670	-0.214981	-7.795917
p_value_st	0.059985	0.068465	0.849712	0.016059
t_stat_lt	-16.828973	-2.457289	4.649190	- 7.229485
p_value_lt	0.003512	0.133288	0.043283	0.018601

4 신림선 개통

	Line 1	Line 2	Line 7	Line 9
t_stat_st	11.842601	31.485186	4.600240	2.808119
p_value_st	0.007055	0.001007	0.044148	0.106868
t_stat_lt	9.549298	12.214162	1.455086	-0.415976
p_value_lt	0.010789	0.006636	0.282893	0.717815





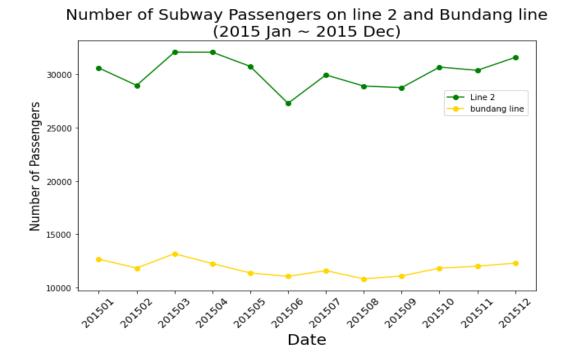


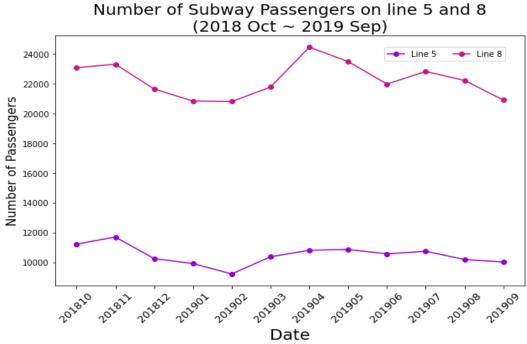




• 9호선 2단계 개통

• 9호선 3단계 개통

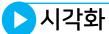








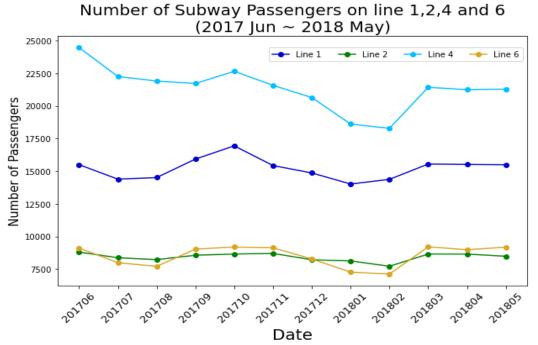


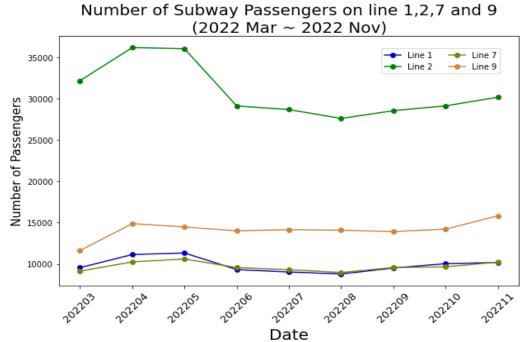




• 우이신설선 개통

• 신림선 개통

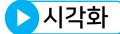














• 결론 도출

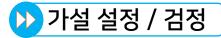
간선노선에 비해 큰 지선노선의 신설 영향력

• 필요한 추가 분석

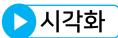
지선 노선의 경우 지하철 - 버스 환승객 신설 노선으로 일부 흡수













서대문구는 연세로 주변 상권 활성화 및 교통체 중 해소를 위해, 2022년 10월 9일부터 '차 없는 거리' 운영을 종료했다. 또한 지난 1월 20일부터 '대중교통전용지구'가 시범 해제되어, 이륜차를 제외한 모든 차량이 9월 30일까지 24시간 연세 로를 통행할 수 있게 되었다. 그러나 서울환경연합 등 일부 환경단체 및 일부 대학생 등이 신촌 일대의 대기 오염을 우려하고, 차량 운행이 상권 활성화에 영향을 주지 않을 것이라는 이유로 대중교통전용지구 해제를 반대하고 있다.











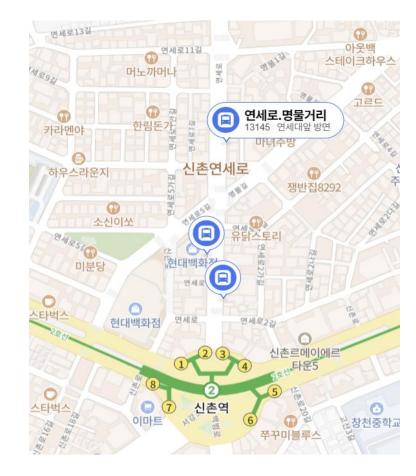


연세로 차 없는 거리 해제와 무관하게 연세로로 들어오는 사람의 수 변화는 유의미하지 않을 것 이다.

VS.



연세로 차 없는 거리 해제로 인해 연세로로 들어오는 사람의 수는 증가할 것이다.





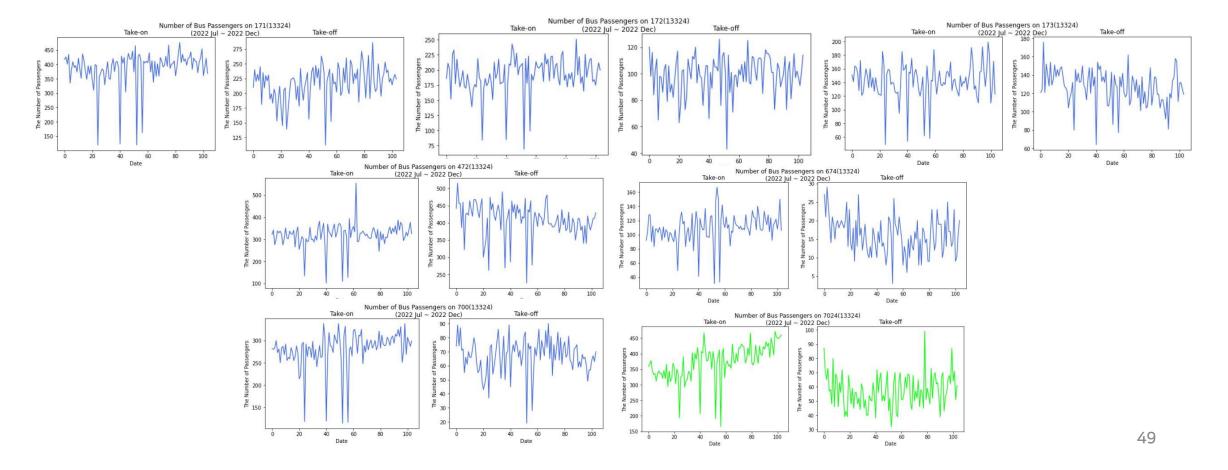








• 13324 (playground 앞 정류장) 버스 승객수 변화추이





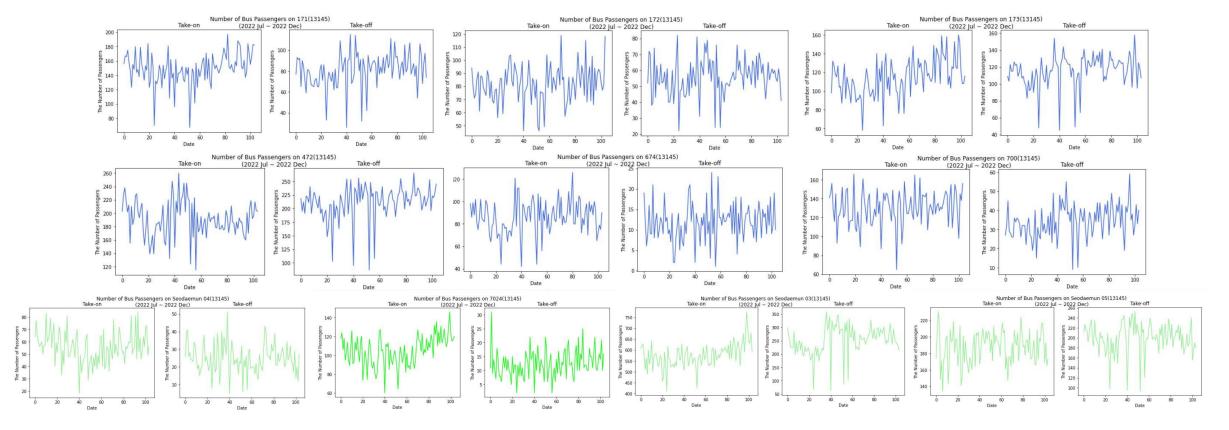






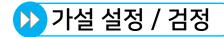


13145 (타코벨 앞 정류장) 버스 승객수 변화추이





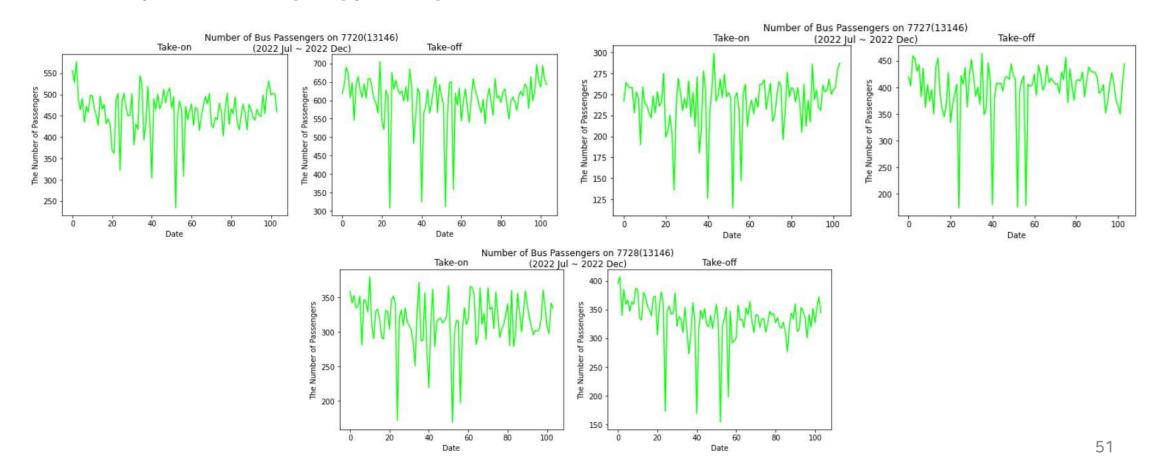






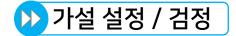


13146 (유플렉스 앞 정류장) 버스 승객수 변화추이





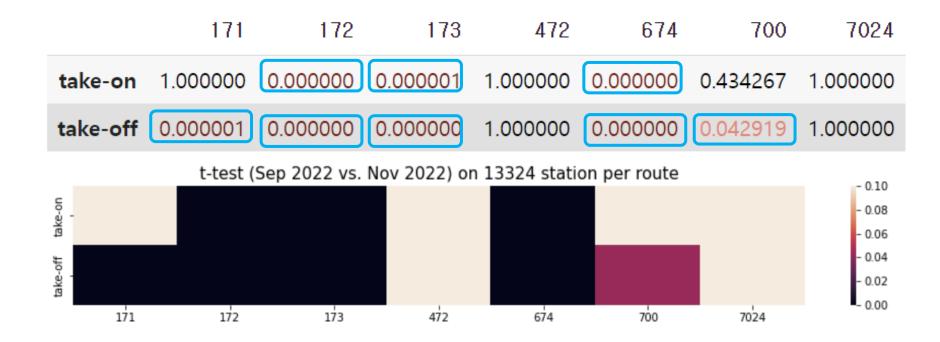






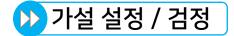


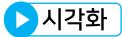
• 13324 (playground 앞 정류장) 버스 승객수 변화 t-검정 결과





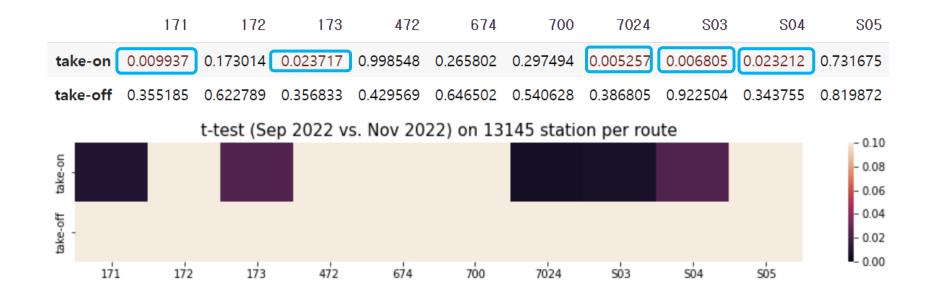








• 13145 (타코벨 앞 정류장) 버스 승객수 변화 t-검정 결과













• 13146 (유플렉스 앞 정류장) 버스 승객수 변화 t-검정 결과



3. 연세로 대중교통 전용지구 해제





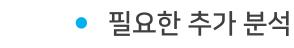






• 결론 도출

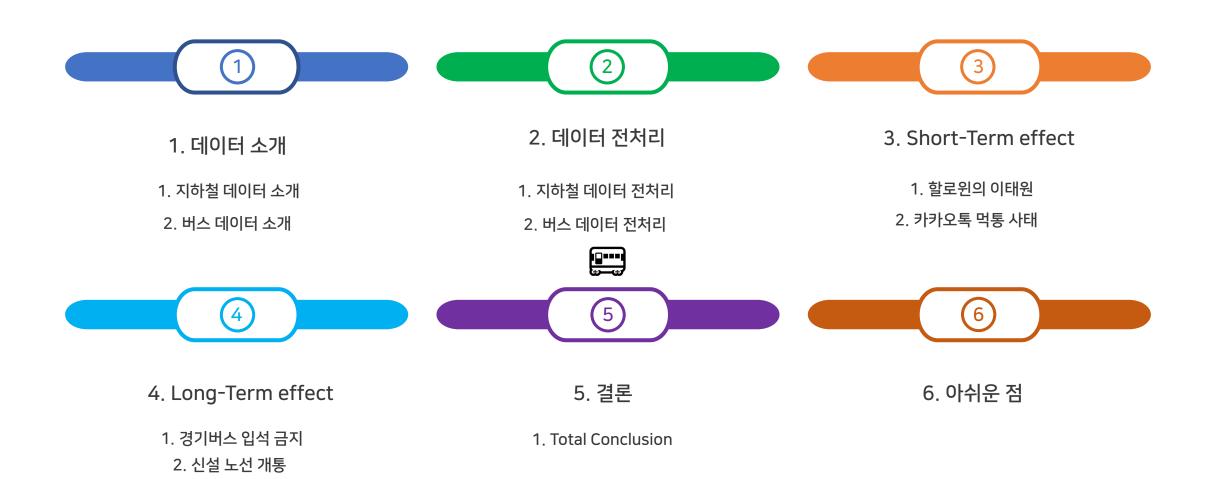
신촌 방문객 수의 통계적 유의미한 증가 존재



상권 활성화와의 연관성 파악 필요 올해의 변화 추이 지속적 주시 필요

5. Conclusion

3. 연세로 차 없는 거리 해제



3. Total Conclusion





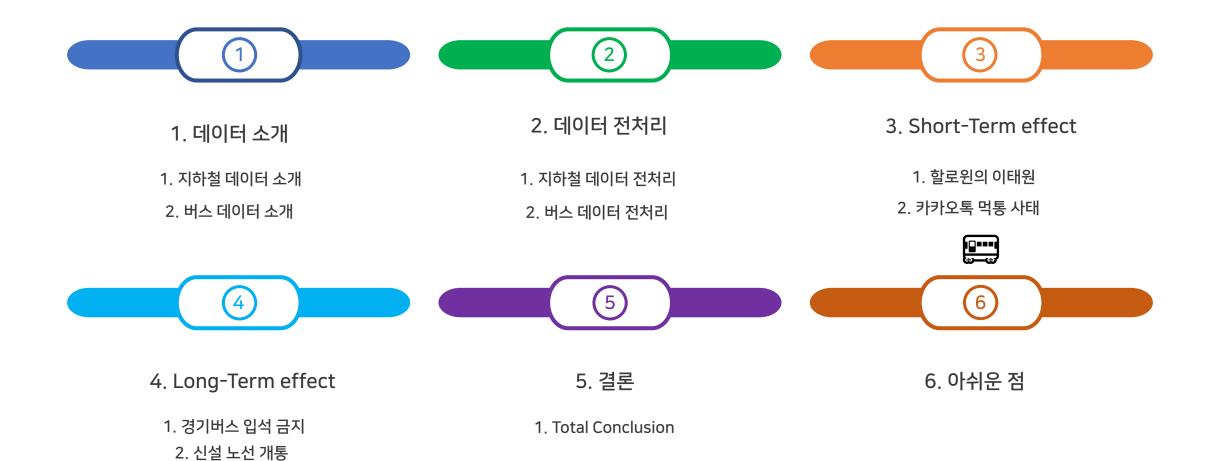


교통 데이터 EDA의 결론 및 의의

- 버스 또는 지하철의 승하차인원을 다양하게 조망하여, 피상적인 데이터에서도 유의미한 결론을 도출할 수 있다.
- 같은 자료를 활용하더라도 분석 목표 및 방법에 따라 다양한 결론을 도출할 수 있다.
- 교통은 우리 삶과 밀접한 만큼, 일상적인 데이터를 통해 편리하고 안전한 삶을 위한 방안을 찾아낼 수 있다.

6. 아쉬운점

3. 연세로 차 없는 거리 해제



1. 아쉬운점 - Period of Data



Period of Data

Direction of Data

10.29 사태 분석의 경우 시간대별로 승객수를 파악했다면 더 정확한 예측이 가능했을 것

- ② 경기도 버스 데이터의 경우 월별 통합 데이터만 제공되어 입석 금지 시행 직전과 직후의 일자별 승차인원 변화를 비교해볼 수 없었음.
- 지난주 대중교통전용지구 해제의 영향은 상권 활성화와의 연관성 파악은 어려우며 올해의 변화 추이 계속해서 살펴봐야 유의미한 판단 가능할 것

2. 아쉬운점 - Direction of Data



Period of Data

Direction of Data

환승역에서 하차해 버스로 환승하던 인원의 지하철 노선 환승 추정되나 버스데이터 미확인으로 입증 불가

2 버스 데이터만으로는 승하차 인원의 행방을 알 수 없어 상권활성화의 구체적 파악에 어려움이 있음

