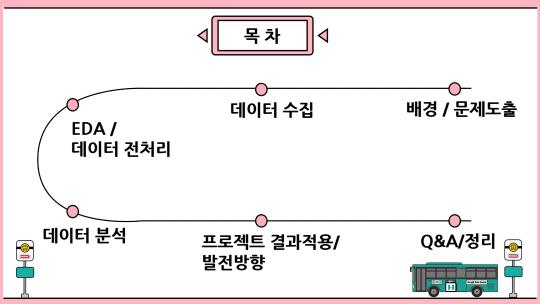
포스트 코로나 시대의 버스 산업 분석: 소비자 및 운영회사 관점으로

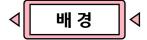
〈클러스터링과 시각화를 통한 버스 사용률 분석〉

7조 엄영민, 조규원, 김한슬









코로나에 대중교통 이용를 21.2% 감소

코르나 이저 수준 회복 주말 교통량 10.7% † 사고도 17.3% 늘어

서울시 "거리두기 2단계 격상 후 대중교통 이용객 급감"

조선비즈 민서연 기자

0124 2020 00 24 24 04

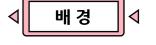
전년동기대비 대중교통 이용객 감소율 14.7%→30.5%로 급증 지하철이 버스보다 이용객 감소 더 커 이용자 유형별로는 청소년→노인→일반인 순

코로나 바이러스 감염증(코로나19)의 확산을 막기 위한 '사회적 거리두기' 2단계 격상으로 서울시 내 대중교통 이용률이 지난 2월 코로나19 심각 단계 발병 시점 수준으로 떨어진 것으로 집계됐다.









서울시 대중교통 이용량 변화추이









서정협 권한대행 "코로나 영향 버스·지하철 운영적자 6000억··· 요금인상은 최후의 수단"

[VT〒つ시]_/ **조선비**즈 | API관기자

[아유경제_사회] '코로나19 영향' 따름이 이용률 ↑ 대중교통 기피

의 조용비 기자 (1) 중인 2020,04,10 17:52

[아유경제-조은비 기자] 신종 코로나바이러스 감염증(이하 코로나19)의 확산으로 인해 대중교통 이용이 감소하고, 서울시 공공자전거 '따름 이'의 이용률이 상승했다.

서울시는 올해 2-3월 따름이의 총 이용건수가 229만5809건으로 전년 동기 대비 13만6330건보다 66.8% 증기했다고 이달 8일 밝혔다. 일 평균 이용 건수도 같은 기간 1만4863건에서 3만7877건으로 64.6% 늘었다.

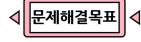
또한 전년 동기 대비 버스, 지하철 이용은 출근시간 이용률이 23%, 퇴근시간 이용률이 26.42% 각각 줄어든 반면 따름이는 출근시간 이용률이 20.46%, 퇴근시간 이용률이 93.33% 각각 증가했다.

서울시 관계자는 "따름이는 아외에서 '사회적 거리두기'를 실천하면서 개인 단위로 쓸 수 있어 밀폐된 환경을 기피하는 시민들이 많이 칭 은 것 같다"고 이용률이 상승한 이유를 설명했다.

서울시는 많은 시민들이 이용하는 만큼, 안심하고 이용할 수 있도록 매일 시내 전역에 있는 따름이 2만5000대와 1540여 개 대여소를 소독하는 방역체계를 진행하고 있다.







프로젝트 목적

- 버스회사 : 회사의 적자를 주이는 방안 제시
- 소비자 : 감염 우려가 적은 안전한 버스 제공

진 행 방 안

버스노선 분석

비교 분석

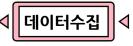
문제해결 방안 도출

- 사용률 감소 x 회복 더딘 노선
- 사용률 증가 노선









●●● 수집데이터셋 ●	0 0
--------------	-----

구분	출처	데이터 명
대중교통	서울시 열린 데이터 광장	서울시 버스노선별 정류장별 승하차 인원정보
대중교통	대중교통 서울시 열린 데이터 광장 서울시 버스노선별 정류장별 시간대별 승하차 인원정보	
대중교통	서울시 열린 데이터 광장	서울시 버스정류소 위치정보
대중교통	티머니	2020년 1월 ~ 10월 버스 노선별 승하차 인원정보
대중교통	공공데이터포털	코로나 전후 버스이용객수 변화
코로나-19	서울시 열린 데이터 광장	서울시 코로나-19 확진자



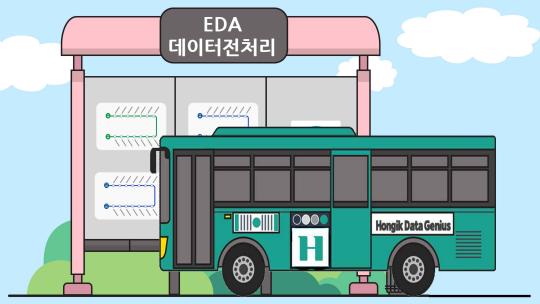




구분	출처	데이터 명
위치(지도)	서울 열린 데이터 광장	서울시 주요 공원 현황
위치(지도)	데이터셋	전국 초/중등학교 위치 표준 데이터
유동인구	Google	Google 모빌리티 데이터









원본자료 ••• 000

사용월		노선ID	노선번호	노선명	버스정류장ID	버스정류장명	승차승객수	하차승객수
	Jan.20	11110897		0017번(정암동~이존동)		청암자이아파트	605	7
	Jan.20	11110897	17	0017번(청암동~이촌동)	9003185	청암동강변삼성아파트	1,582	43
	Jan.20	11110897		0017번(청암동~이촌동)		청심경로당	2,924	61
	Jan.20			0017번(청암동~이촌동)		원효2동주민센터	5,048	98
	Jan.20	11110897	17	0017번(청암동~이촌동)	72985	산전동	6,235	274
	Jan.20			0017번(청암동~이촌동)		신창동세방아파트	5,857	270
	Jan.20	11110897	17	0017번(청암동~이촌동)	72987	남이장군사당	1,458	1,427
	Jan.20	11110897	17	0017번(청암동~이촌동)	72988	새마음금고	4,164	3,159
	Jan.20	11110897	17	0017번(청암동~이촌동)	33066	용문시장	5,203	1,574
	Jan.20	11110897	17	0017번(청암동~이촌동)	72990	용산전자삼가입구	2,252	966
	Jan.20	11110897	17	0017번(청암동~이촌동)	72991	신용산지하차도	1,074	2,362
	Jan.20	11110897	17	0017번(정암동~이촌동)	72992	용산역	30,229	20,332
	Jan.20	11110897	17	0017번(청암동~이촌동)	72314	용산푸르지오써밋	1,015	1,895
	Jan.20	11110897	17	0017번(청암동~이촌동)	72993	한강대교복단	6,139	4,568
	Jan.20	11110897	17	0017번(청암동~이촌동)	72994	서부이촌동입구	2,050	11,402
	Jan.20	11110897	17	0017번(청암동~이촌동)	72995	이촌2동대림아파트,새남터성	995	10,468
	Jan.20	11110897	17	0017번(정암동~이촌동)	72996	이촌2동주민센터	860	6,002
	Jan.20	11110897	17	0017번(청암동~이촌동)	72997	성촌공원.포르쉐센터용산	450	2,027
	Jan.20	11110897		0017번(청암동~이촌동)	73000	이촌119안전센터.성촌공원	316	2,260
	Jan.20	11110897	17	0017번(청암동~이촌동)	72998	원효2동산호아파트후문	148	2,667
	Jan.20	11110897	17	0017번(청암동~이촌동)	8003000	청암동강변삼성아파트	2,135	3,559
	Jan.20	11110897	17	0017번(청암동~이촌동)	9003186	원효2동산호아파트	4,546	976
	Jan.20	11110897	17	0017번(청암동~이촌동)	9000272	원효로풍전아파트	3,186	484
	Jan.20	11110897	17	0017번(청암동~이촌동)	9000367	성존공원.포르쉐센터용산	3,034	1,414
	Jan.20	11110897	17	0017번(청암동~이촌동)	9000269	이촌2동주민센터	7,835	1,117
	Jan.20	11110897	17	0017번(청암동~이촌동)	9000267	이촌2동대림아파트,새남터성	11,282	839
	Jan.20	11110897	17	0017번(청암동~이촌동)	9000266	이촌2동강변아파트	8,896	3,316
	Jan.20	11110897	17	0017번(청암동~이촌동)	9034958	한강대교북단LG유플러스	6,795	12,902
	Jan.20	11110897	17	0017번(청암동~이촌동)	9034956	신용산역	4,543	11,837









사용월

노선ID

노선명

노선번호

버스정류장명

버스정류장 ID

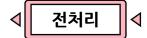


승차승객수

하차승객수







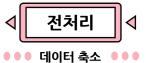
●●● 데이터 통합 **●●●**

노선먼프 1를	승자_1월	승지_2별	승지_3월	순자_4월	승차_5월	승차_6월	승차_7월
17	166568	142700	120347	122239	137603	146884	164389
01A	561	8909	11849	15884	18394	20309	22760
01B	625	10089	11517	14439	17082	19439	22052
2	81,953	53,652	49,361	71,837	65,117	52,645	60548
3	37045	2441	1703	2005	3081	3579	4758
4	1980	26089	24685	35230	36047	32787	44696
100	492717	411701	343980	375197	424889	453228	480577
101	521234	452759	408534	420056	463559	480832	504986
1014	235041	193375	175022	184331	203041	212439	223827
1017	137282	120283	101746	104012	117409	127271	134023
102	383943	333575	283808	303454	337249	347681	365659
1020	169126	140924	116518	135082	153373	161495	169542
103	349722	299477	268929	289915	323534	337059	356096
104	320739	266207	237685	254495	276297	279159	292542
105	334828	273710	242613	263884	290520	293589	308008
106	325226	274105	259483	268019	289474	286183	298475
107	469485	384093	347892	353798	393128	402951	422020
108	367691	331137	320304	330479	369630	371558	360708
109	243473	218475	193727	204256	222890	238696	254705
110A	387482	338313	295245	307703	325573	351798	377606
110B	425144	373581	328853	340738	361172	388254	417802
1111	323657	285819	255416	256394	288289	311413	324163
1113	59134	47044	40963	44574	50726	50369	53188
1114	44093	36679	31278	34033	37650	41822	44453
1115	64230	55609	43982	48002	53363	60023	66079
1116	48067	42063	36097	39004	43254	47929	49900









데이터를 불러오고 전처리를 진행한다 (data filitering, data transformation)

	HIO	터를 둘러오	고 선저리들 신	맹한다 (data	filitering,
In [8]:	bus['	노선번호_1월 ']		
Out[8]:	0	17			
	1	01A			
	2	01B			
	1 2 3 4 5 6 7 8 9	2			
	4	3 4			
	5	4			
	6	100			
	7	101			
	8	1014			
	9	1017			
	10	102			
	11	1020			
	12	103			
	13	104			
	14	105			
	15	106			
	16	107			
	17	108			
	18	109			
	19	110A			
	20	1108			
	21	1111			
	22	1113			
	23	1114			
	24	1115			
	25	1116			
	26	1119			
	27	1120			
	28	1122			
	29	1124			

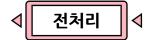
* 기준 : 2020년 1월 버스 노선

〈분석 대상 제외 노선 〉

N850, 9709, 8002, 8003, 8112, 마포18-1, 서대문15







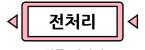
●●● 데이터 변환●●●

- scale을 고려한 정규화
- 데이터 분포를 고려한 표준 정규화 사용 (z-score)





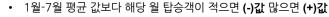
 $X - \mu_{n \text{th} \text{th} - \phi \text{s}^{\text{th}}}$



●●● 최종 데이터 ●●●

n [224]:	rati	o_data						
ut[224]:		승차_1월	승차 2월	승차 3월	승차_4월	승차 5월	승자_6월	승자 7월
	0	1.699753	0.032372	-1.529172	-1.397000	-0.323696	0.324660	1.547531
	1	-2.358331	-0.997514	-0.518261	0.139489	0.548646	0.860813	1.260353
	2	-2.426600	-0.775913	-0.526845	-0.017198	0.443788	0.854890	1.310642
	3	1.470965	-0.672061	-0.996986	0.704955	0.196099	-0.748313	-0.149878
	4	2.989075	-0.397295	-0.469516	-0.439962	-0.334664	-0.285929	-0.170552
	5	-2.318643	-0.503932	-0.609613	0.184121	0.245617	0.000233	0.896637
	6	1.634065	-0.112465	-1.572383	-0.899412	0.171840	0.782767	1.372353
	7	1.792782	0.012351	-1.137552	-0.837966	0.293163	0.742282	1.370315
	8	1.934411	-0.365580	-1.354933	-0.853114	0.155484	0.662100	1.275992
	9	1.688311	0.216514	-1.388445	-1.192252	-0.032321	0.821545	1.406142
	10	1.843209	0.093413	-1.635504	-0.952997	0.221049	0.583459	1.208018
	11	1.283129	-0.302544	-1.674785	-0.631014	0.397408	0.854072	1.306519
	12	1.345557	-0.405683	-1.470404	-0.738958	0.432800	0.904200	1.567716
	13	2.145852	-0.060616	-1.214670	-0.534506	0.347644	0.463446	1.004948
	14	2.074561	-0.331579	-1.555830	-0.718417	0.334146	0.451032	1.018690
	15	2.231047	-0.259756	-0.972193	-0.556288	0.489079	0.328729	0.927640
	16	2.243174	-0.100190	-1.093633	-0.931558	0.147753	0.417320	0.940619



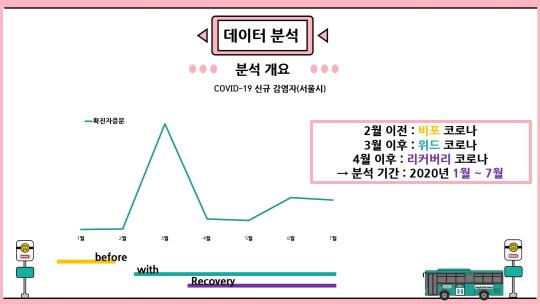


• 절대값이 클수록 많이 (증가/감소)됨을 의미



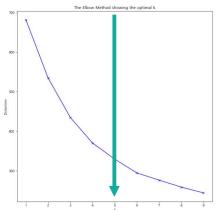






십 데이터 분석 △

●●● K값의 선정 ●●



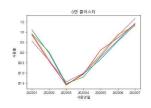
1.elbow method 사용 → 적정 클러스터의 개수 **5**개

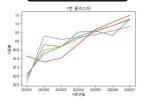
2. k-means 사용

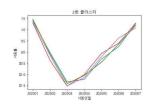


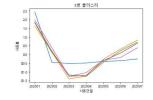


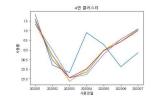
△ 데이터 분석 △









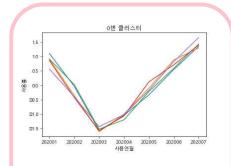






십 데이터 분석 ↓

●●● 군집분석 결과 ●●●



0번 클러스터

위드 코로나 시기(2020.03) 사용률 감소

V

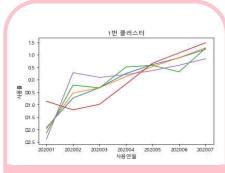
리버커리 코로나 시기(2020.04.-07.) 사용률 회복 - <mark>이전보다 사용률 증가</mark>





◁ 데이터 분석 ◁

▶●● 군집분석 결과 ●●●



1번 클러스터

2020.01-07 (위드코로나시기, 리커버리코로나시기)

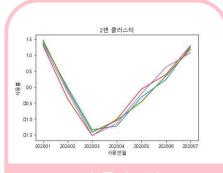
오히려 <mark>사용률이 꾸준히 증가</mark>한 노선





십 데이터 분석 ↓

●●● 군집분석 결과 ●●●



2번 클러스터

위드 코로나시기(2020.03) 사용률 감소

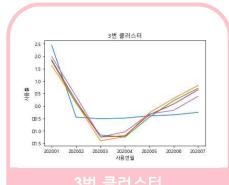
리버커리 코로나 시기(2020.04.-07.) 사용률 회복 - <mark>코로나 이전시기와 비슷한 수준</mark>





△ 데이터 분석 △

●●● 군집분석 결과 ●●●



위드 코로나시기(2020.03) 사용률 감소

V

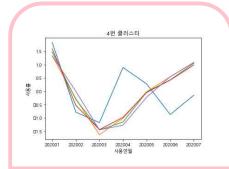
리버커리 코로나 시기(2020.04.-07.) 사용률 회복 - <mark>더디게 회복</mark>





십 데이터 분석 ↓

▶●● 군집분석 결과 ●●●



4번 클러스터

위드 코로나시기(2020.03) 사용률 감소

리버커리 코로나 시기(2020.04.-07.)

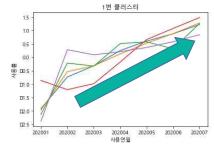
사용률 회복 - <mark>코로나 이전시기와 비슷한 수준</mark>



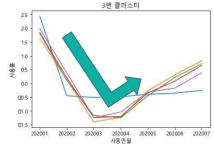


십 데이터 분석 ↓

●●● 정성적 분석 ●●●

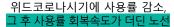


VS

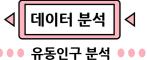


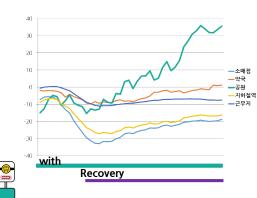


위드코로나시기에 꾸준히 사용률이 증가한 노선





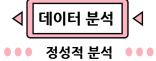




가설 1 -사용률이 증가하는 노선은 <mark>공원</mark>을 많이 지날 것이다.

가설 2 -사용률이 더디게 회복되는 노선은 공원이 아닌, <mark>학교</mark>를 많이 지날 것이다.





1번 클러스터	3번 클러스터
01A	3
01B	1119
4	1122
1120	1126
660	1129
8772	1142
8777	1144





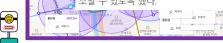
◁ 데이터 분석 ◁

●● 정성적 분석 ●●



버스 정류장 기준 반경 500m 내 학교,공원,산책로 수 확인

개정안은 현재 시행령으로 돼 있는 통학구역 관련 사항을 법률로 상향 신설해 초등학교 통학거리를 <u>1km</u>로 하고, 통학거리가 1km를 초과하는 학생에 대해서는 국가 및 지방자치단체가 통학에 필요한 비용을 보조할 수 있도록 했다.







△ 데이터 분석 △

●●● 정성적 분석 ●●●







*공원

*학교



◁ 데이터 분석 ◁

분석 결과

000









●●● 분석 결과

1번 클러스터						
노선	공원 / 강	학교				
01A	19	7				
01B	19	8				
4	20	7				
1120	17	10				
660	12	9				
8772	6	3				
8777	14	5				

3번 클러스터					
노선	공원 / 강	학교			
3	12	12			
1119	6	32			
1122	4	22			
1126	6	18			
1129	6	23			
1142	8	26			
1144	21	57			





◁ 데이터 분석 ◁

●●● 왜 그런가?



등아일보

수도권 모든 학교 선제적 3단계...15일부터 유치원-초중고 등교 중단

코로나19 확진자 급증세가 가팔라지면서 수도권뿐 아니라 전국적으로 등교 를 줄여야 한다는 의견도 나온다. 현재 비수도권은 등교 밀집도 3분의 1(...



👤 한국일보

일일 확진 1000명에... 수도권 학교 15일부터 전면 등교중 지

신종 코로나바이러스 감염증(코로나19) 일일 확진자가 1,000명을 넘으면서 서울지역 중·고등학교에 국한됐던 등교중단이 수도권 유·초·중·고교로 ...



☑ 조선에듀 코로나19 중단

코로나19 확진자 폭증에...15일부터 수도권 학교 등교수업 주다

서울시교육청은 '7일부터 등교수업을 중단한 중고등학교에 이어 15일부터 31일까지 유치원과 초등학교, 특수학교의 수업도 원격으로 전환하기로 ... 39분 전



-- 디지털타인스 등 13명 1단 6일 전 네이버뉴스

위기의 호텔·면세... 이부진 호텔신라 새판 짜나 호텔신라 역시 코로나19 여파가 이어지면서 3분기 연속 적자를 기록했다. 속원사 업인 한옥호텔 건립 사업도 1년 간 보류하기로 결정했다. 면세호텔 등 주요 관광...

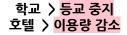


'변화vs안정' 갈림길...인사 앞둔 호텔신라, 이부진... 뉴스1 7일 전 네이버뉴스

Ø 연합인포맥스 6일 전

'**코로나**19 직격탄' 면세·호텔기업 신용등급 줄강동

신중 코로나바이러스 강염증(코로나19) 직격탄을 맞은 면세호텔기업들의 신용등급이 출출이 강동되고 있다. 영업 부진이 장기화하면서 올해 대규모 적자를 낸데다. 코로나19 이전 수준으로의 실적 회복 시점이...









03 (녹색순환)

1119

1126

운행지역

운행시간

배차간격

운행지역 서울 시청앞 ↔ 시청앞 ~

운행시간 첫차 06:30, 막차 22:45 배차간격 매일 22분

운행지역 서울 강북청소년수련관 ↔ 녹천역 ~ 운행시간

첫차 04:40, 막차 23:35

평일 5-9분, 토요일 5-9분, 일요일 5-10분

1142

1144

1129

운행지역 서울 상계8동 ↔ 창동역 ~

운행시간 첫차 05:00, 막차 23:30 배차간격 매일 18분

서울 중계본동 ↔ 창동역 ~ 첫차 05:00, 막차 00:00

배차간격

평일 8분, 주말 9분

서울 하계동 ↔ 우이동 ~ 운행시간

첫차 04:25, 막차 23:50

첫차 05:00, 막차 23:40 평일 20분, 토요일 20분, 일요일 22분

서울 강북청소년수련관 ↔ 안방학동 ~

평일 5-9분, 토요일 5-9분, 일요일 5-10분

저상예약 02-976-9523



3번 클러스터_노선 [<mark>배차간격</mark>을 늘린다]>자원(인원,장비) 절약



결과 적용

04 (녹색순환)

운행시간

1120

660

운행지역 서울 남산서울타워 ↔ 남산도서관 ~

첫차 06:30, 막차 22:30

배차간격 평일 12분, 토요일 17분, 일요일 19분 서울 하계동 ↔ 삼양동입구 ~

운행시간 첫차 04:30, 막차 23:00

평일 16분, 토요일 18분, 일요일 19분

위행지역 서울 온수통 ↔ 가양역 Y

운행시간 첫차 04:00, 막차 22:30

배차간격 평일 15분, 토요일 17분, 일요일 18분

저상예약 02-2613-6004

8777 (주말운행)

서울 난지캠핑장 ↔ 월드컵경기장남축 ~

운행시간 첫차 10:00, 막차 20:00

배차간격 주말 40-50분

8772 (맞춤버스(주말))

서울 진관공영차고지 ↔ 북한산성입구 >

첫차 08:00, 막차 18:00

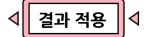
배차간격 주말 13분



1번 클러스터_노선

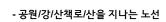
[배차간격을 줄인다]>밀집 완화/적자감소





●●● 버스 노선 제안 ●●●



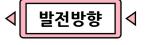




-Seoul Park Tour Bus







한계선

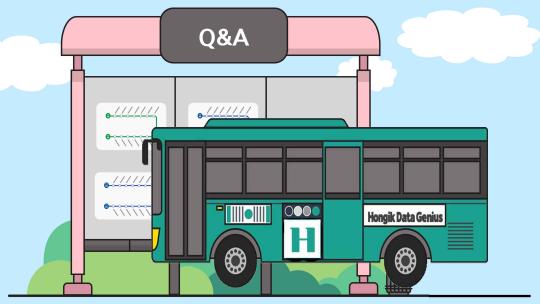
- 1. 클러스터링 결과분석 분야
- 2. 결론 활용 분야

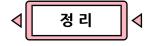
발전방향

- 1. 1번, 3번 클러스터 뿐 아니라, <mark>다른 클러스터와의</mark> 추가적인 상호 비교도 가능할 것
- 2. <mark>신규 버스 노선</mark>에 대한 구체적인 내용이 필요할 것으로 사료됨.
- → 사업 타당성 알아보기 위해 추가적인 분석 요구









목적달성방안

- 버스회사: 사용률 감소한 노선에 대해 배차간격을 줄여 회사의 적자를 줄이는 방안 제시
- 소비자 : 사용률이 증가한 노선에 대해 배차간격을 늘려 감염우려가 적은 버스 이용 방안 제시

프로젝트과정

버스노선 분석

비교 분석

문제해결 방안 도출

- 1번 클러스터 x 3번 클러스터 비교 배차간격 조정 / 신버스노선 제안
- 정량적 수치에 기반한 정성적 분석 (2가지의 가설 검증)





