

코로나 시대 비즈니스 분석

–산업별 회복탄력성을 중심으로 –



CONTENTS

Analysis

데이터, 문제 정의 과정

코로나 클러스터 모델

클러스터 분석

비즈니스 인사이트

Data

카드 데이터

유동인구 데이터

미세먼지 데이터

온라인 쇼핑 데이터





미세먼지의 매출 영향은 여전할까?

회귀분석 모델의 adjusted R² 값이 모두 0.2~4 사이이다
미세먼지의 변화가 유동인구, 판매액, 판매 건수에 영향을 미친다는 직접적인 설명력을
가진다고 보기 어렵다

회귀분석 adjusted R²

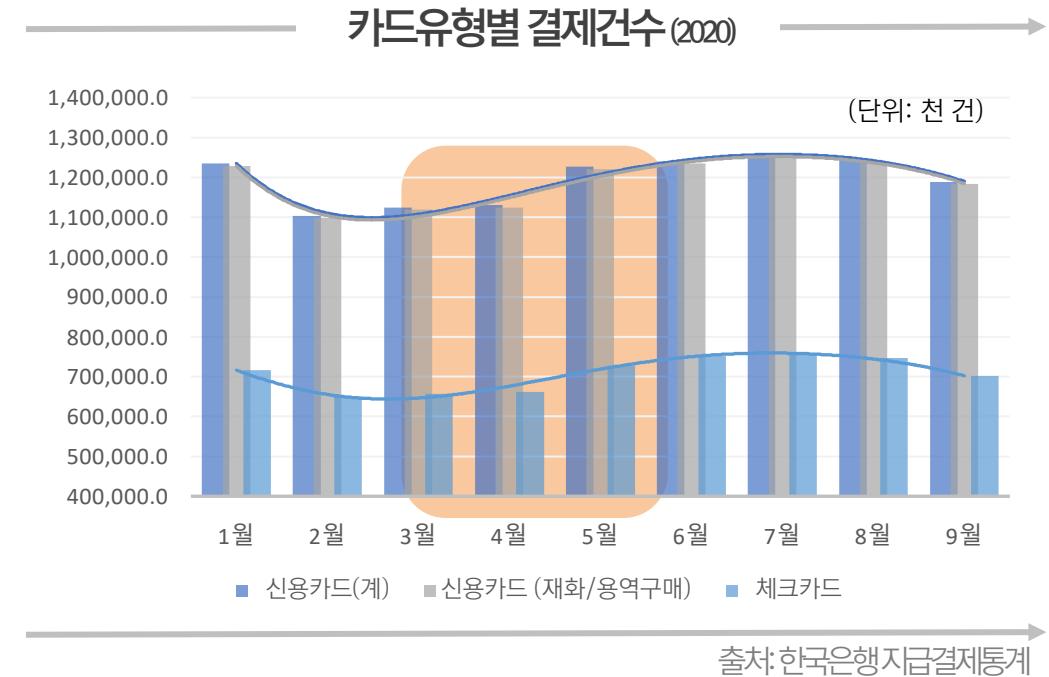
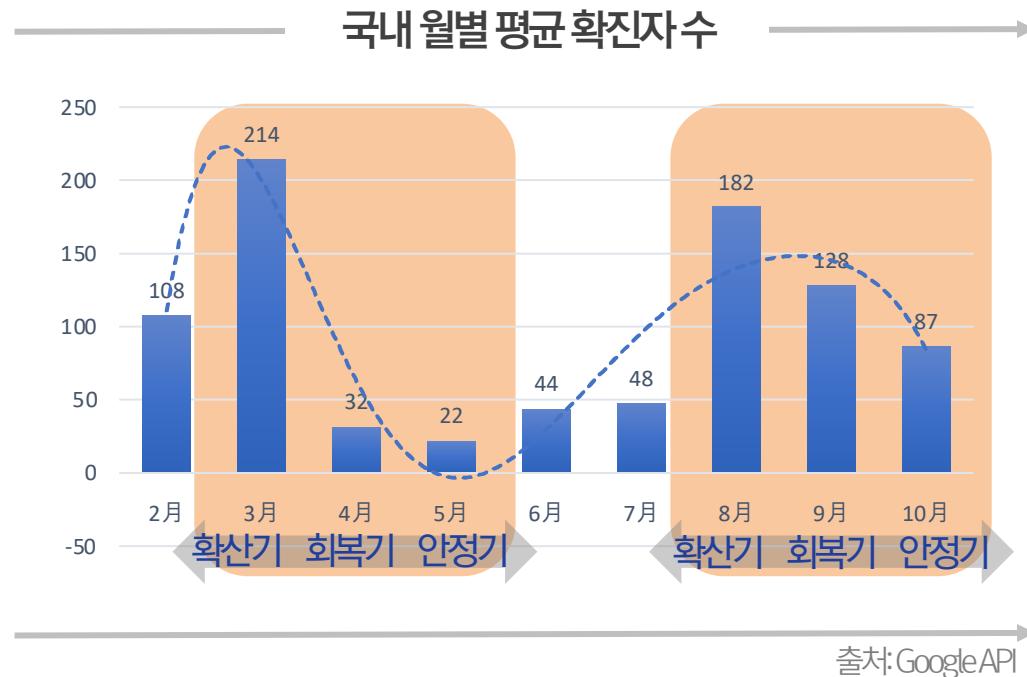
R ²	판매건수	판매액	유동인구
일일 확진자수	0.03	0.002	0.002
미세먼지 농도	0.025	0.002	0.04

상관관계 분석

구분	판매건수	판매액	유동인구
일일 확진자수	-0.16	-0.05	-0.06
미세먼지 농도	-0.17	-0.13	-0.02

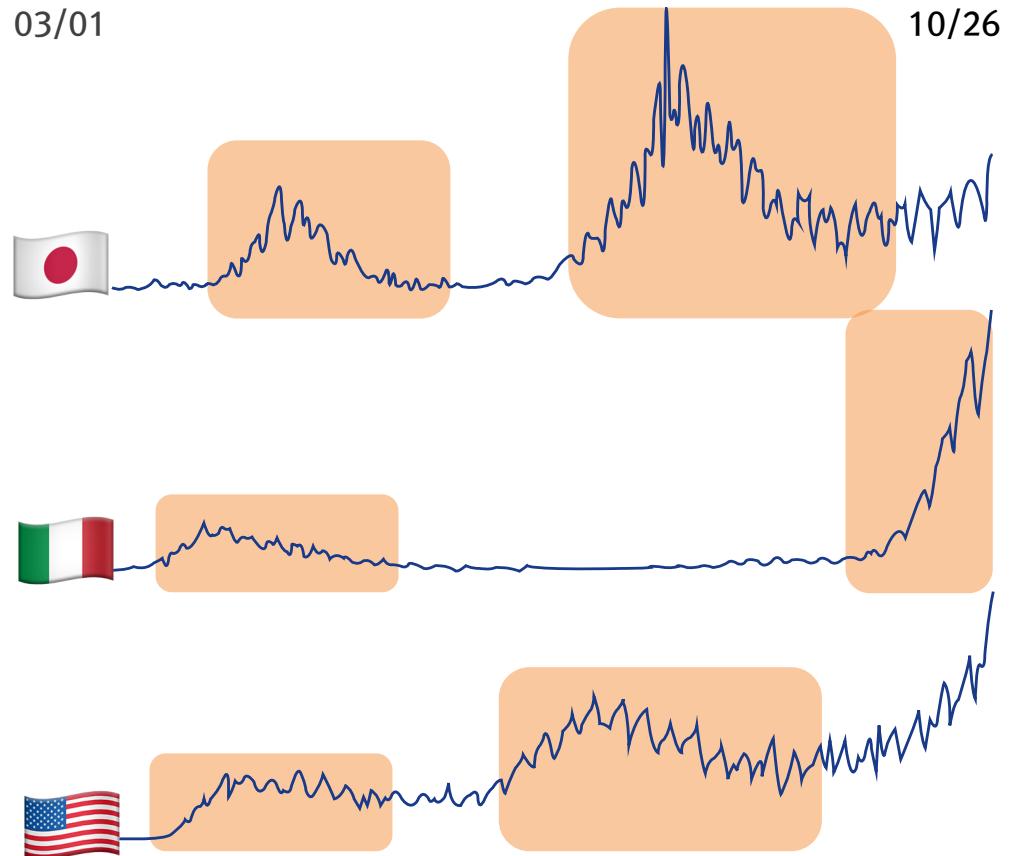
주기성 (periodicity) 이란?

국내 코로나 확진자 추이에서 확진자가 증가하고 감소하는 주기적인 패턴을 발견하였고
 패턴 내 기간을 코로나 확산기 – 회복기 – 안정기로 구분했다.
 국내 카드 결제 건수도 이러한 주기의 변화에 맞물려 증감하고 있다.





확진자수 추이의 주기성은 전세계적인 현상이며, 확산기-회복기-안정기로 구분하여 분석하였다.



주차별 코로나 확진자 수와 주차별 이슈

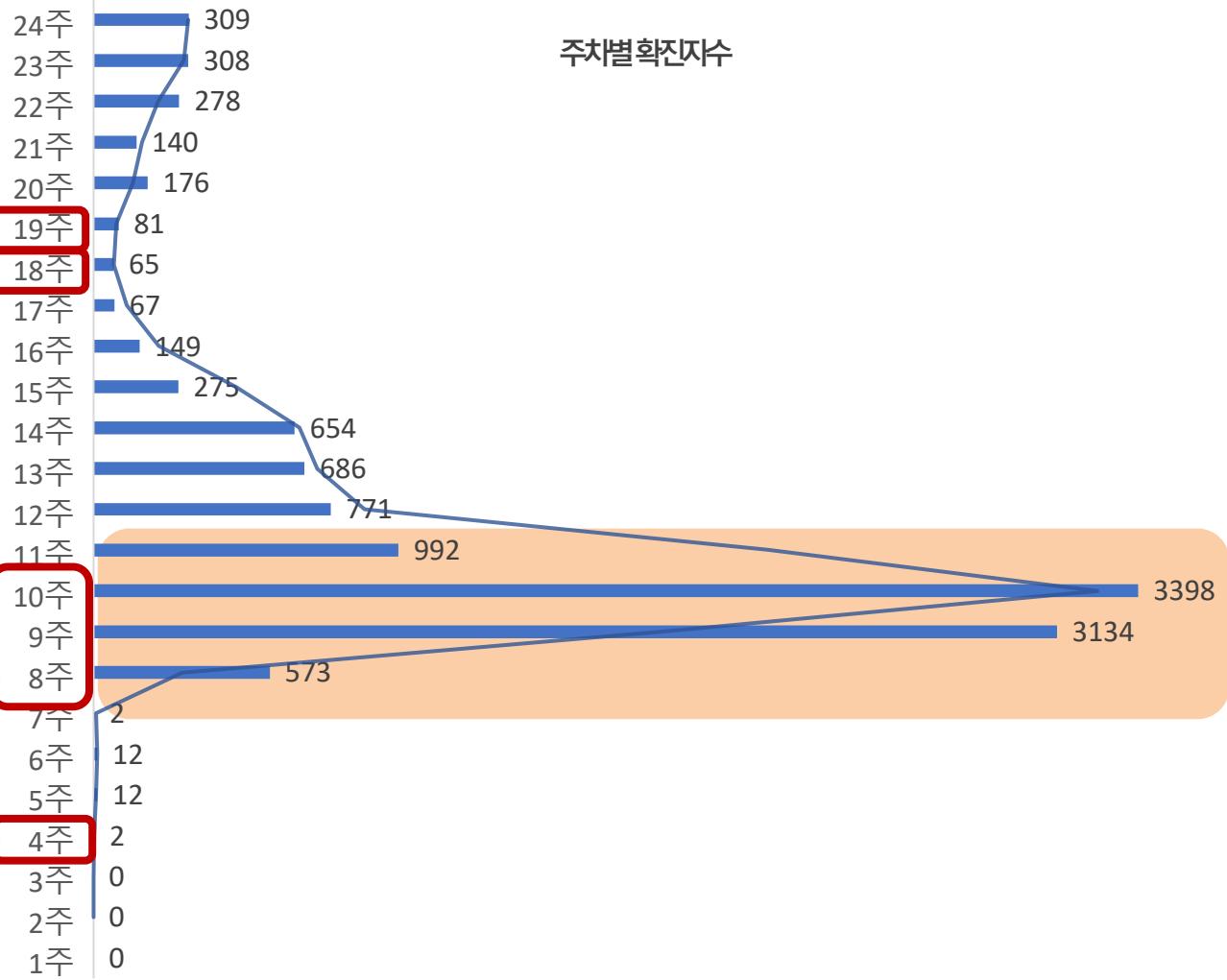
5월

재난지원금 지급
(5월 12일)

3월

이태원클럽발 확산
(5월 첫째주)

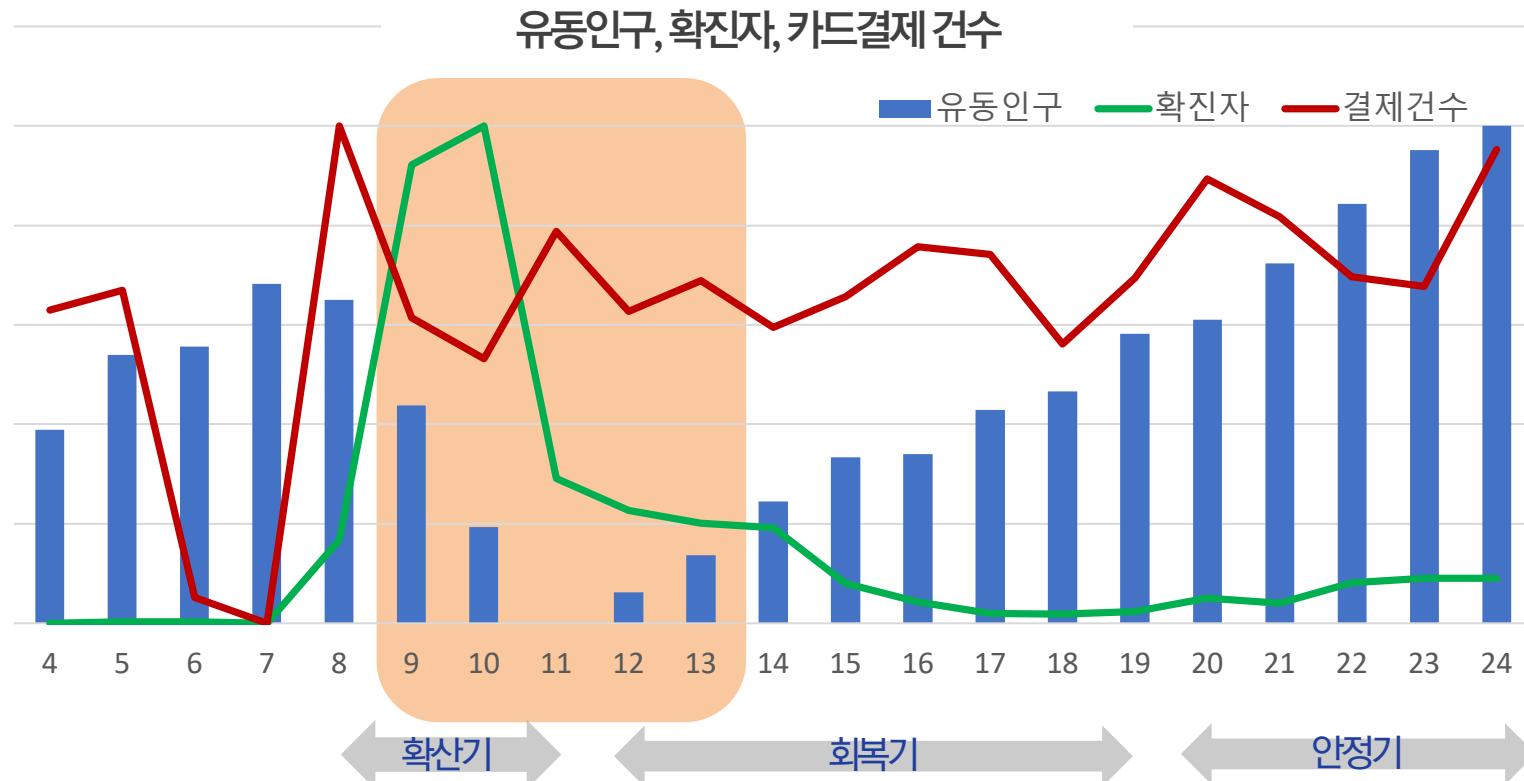
1월

코로나 대확산기
(2월 말~3월)설날 연휴
(1월 24~27일)



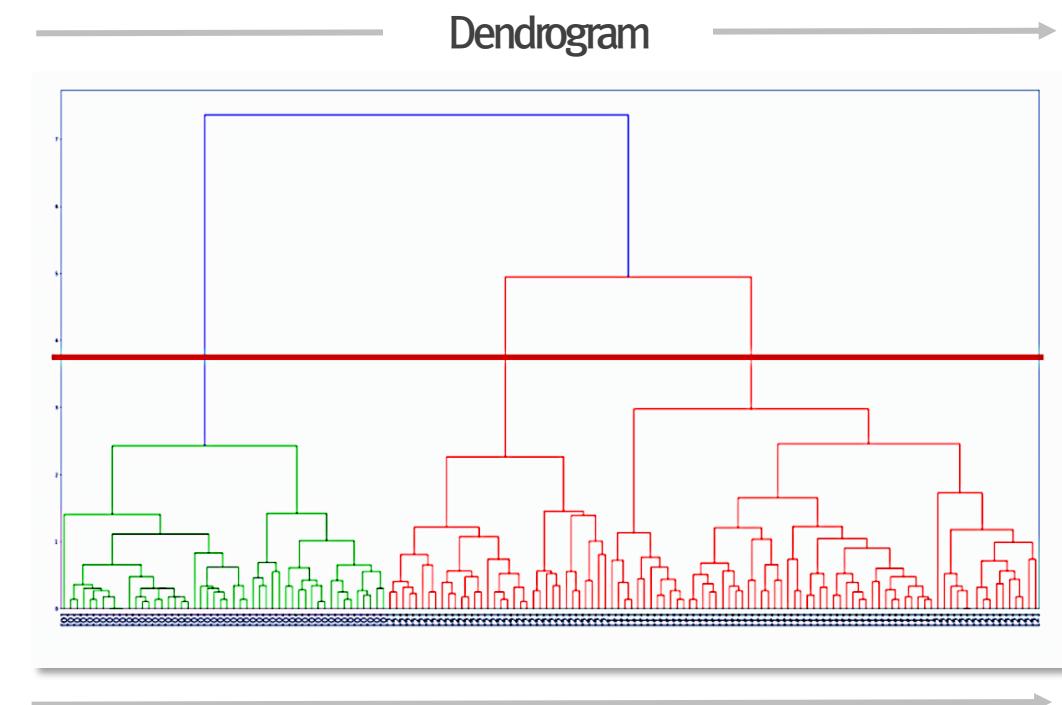
코로나 확진자 수에 따른 유동인구와 매출 영향

대확산기(8~10주)에 급증한 코로나는 유동인구에 시간차를 두고 영향을 주었다.
유동인구는 11주에 최저점을 찍고 그 이후로 회복되는 모습을 보인다.
다만, 유동인구의 증가와 감소한 매출증가는 함께 상승하지 않는다.



계층적 군집분석(Hierarchical clustering)이란?

계층적 군집 분석은 비슷한 군집끼리 묶어 가면서 최종적으로 하나의 케이스가 될 때까지 군집을 묶는 클러스터링 군집수를 미리 정하지 않아도 된다는 장점으로 K-means가 아닌 본 알고리즘을 선택하였다.





데이터 전처리 과정

주(week)기준으로 데이터를 합병한 데이터들에 추가적인 전처리 기법을 적용하였다.

이동평균 기법(window=3) 적용

업종명	4주	5주	...	24주
1급 호텔	508942652	592265204	...	231747690
2급 호텔	472445334	652130542	...	308605414
...
CATV 홈쇼핑	5645239480	10303560646	...	10597606171

날짜기준정렬된카드데이터 260rows × 21columns



업종명	4주	...	24주
1급 호텔	5.089427e+08	...	2.430602e+08
2급 호텔	4.724453e+08	...	3.292284e+08
...
CATV 홈쇼핑	5.372550e+07	...	1.019362e+08

이동평균적용된카드데이터 260rows × 21columns

Min-max 스케일링 적용

업종명	4주	5주	...	24주
1급 호텔	5.089427e+08	5.506039e+08	...	2.430602e+08
2급 호텔	4.724453e+08	5.622879e+08	...	3.292284e+08
...
CATV 홈쇼핑	5.372550e+07	7.974400e+09	...	1.019362e+08

이동평균적용된카드데이터 260rows × 21columns



업종명	4주	...	24주
1급 호텔	0.893965	...	0.217249
2급 호텔	0.769976	...	0.403297
...
CATV 홈쇼핑	0.051887	...	0.829887

0~1범위의스케일링된데이터 260rows × 21columns

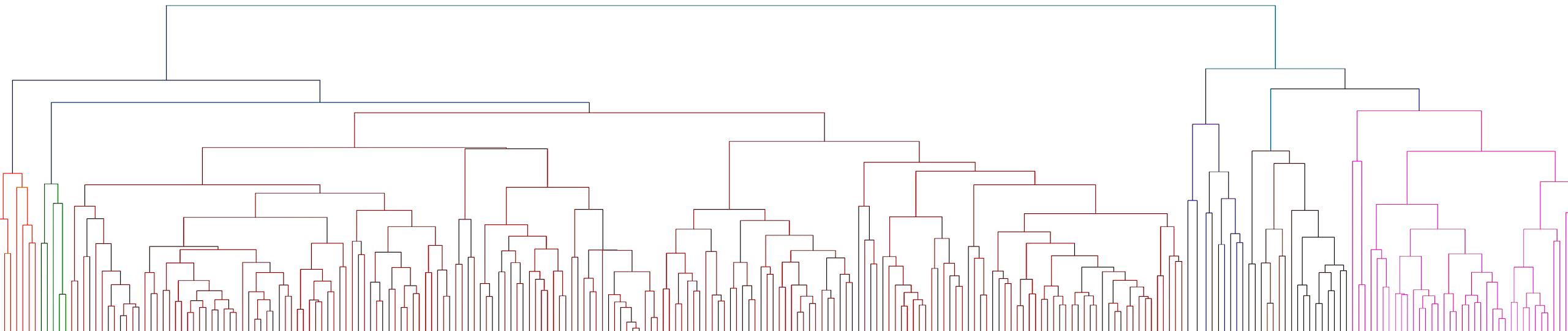


클러스터 모델 설명

- 1) 총 24주간의 매출액 데이터를 가진 260개 업종으로 모델 인풋을 설정하였다
- 2) Mergings(군집 매개변수) = 2로 설정하였다
- 3) Complete distance(군집 간 최대 거리)로 클러스터 방식을 설정하였다.

모델 결과로 17개의 클러스터와 최빈 클러스터 값은 63가 나왔다.

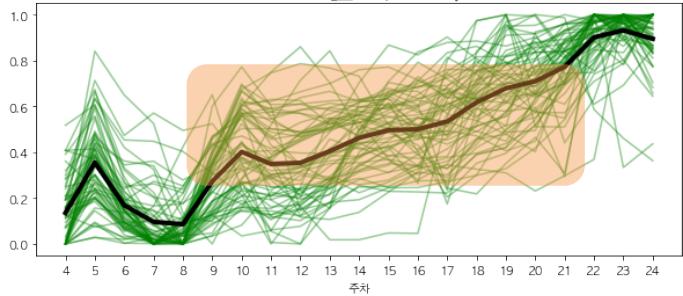
count:260 , unique: 17 , top:3 , freq:63



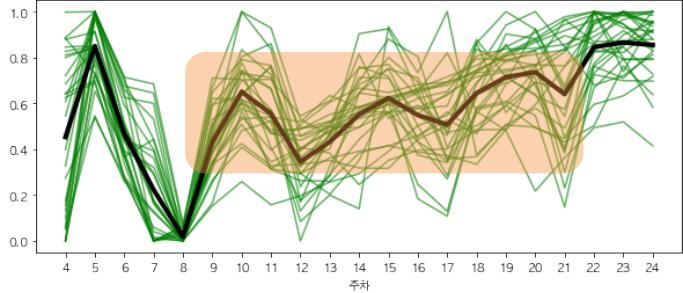
증가 클러스터

회복기(11~18주를 기준으로 매출회복세가 증가)

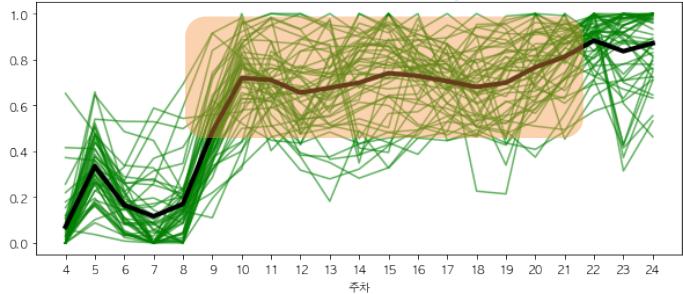
3번 클러스터



5번 클러스터



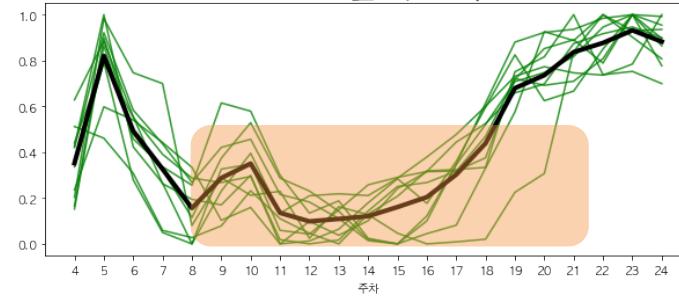
8번 클러스터



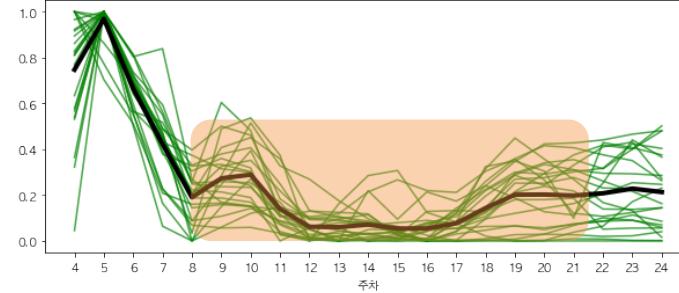
감소 클러스터

회복기(11~18주를 기준으로 매출회복세가 감소)

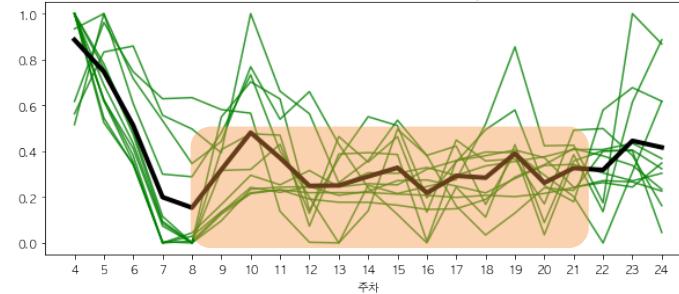
13번 클러스터



16번 클러스터



17번 클러스터

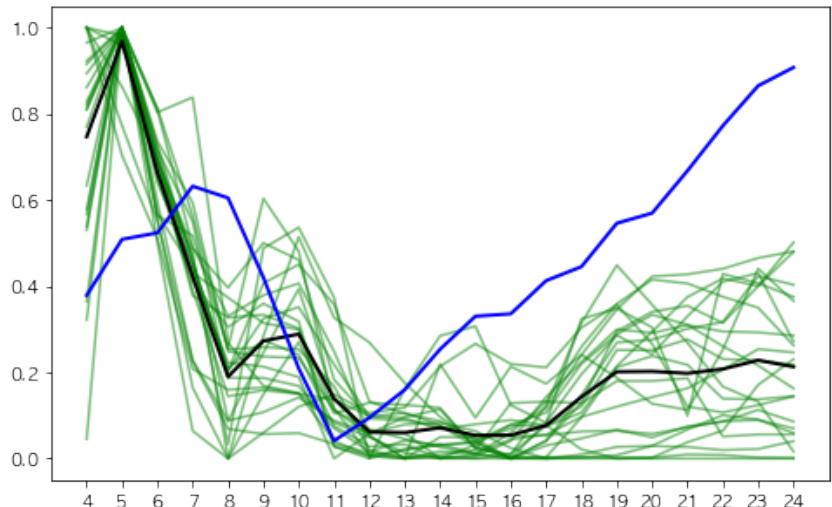


감소 클러스터 16번 INSIGHT

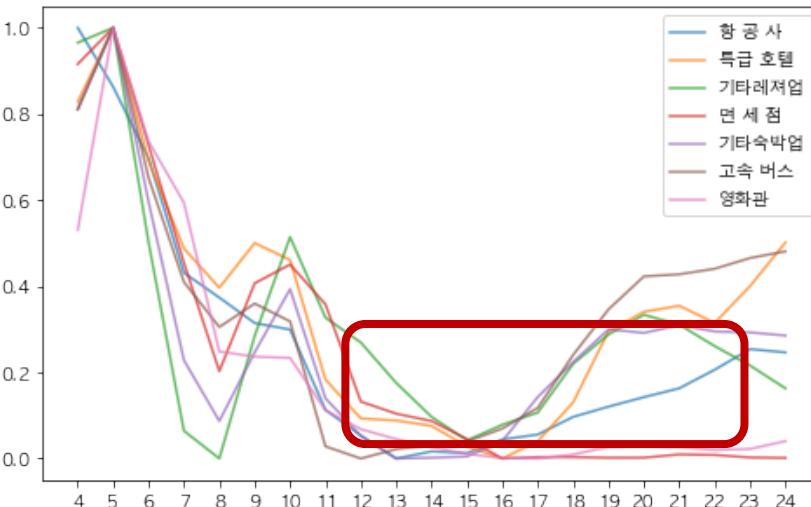
항공사, 호텔숙박, 레져업, 면제점, 버스, 영화관, 화장품은 코로나 안정기에서도
언텍트 라이프로 인해 매출을 기준만큼 회복하지 못하고 있다

‘문화, 여행, 교통’ 클러스터

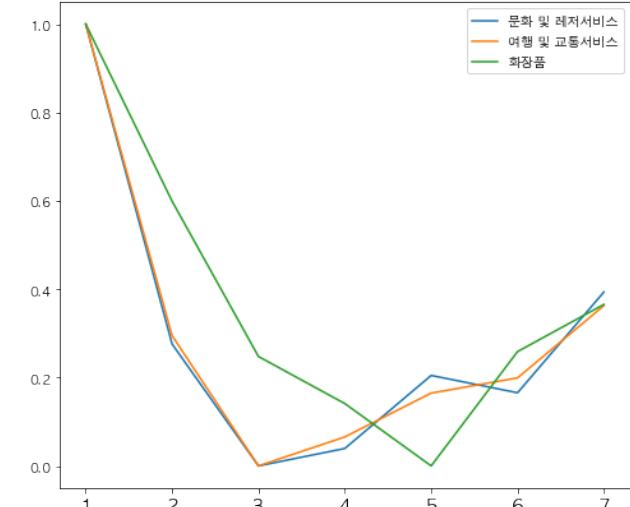
유동인구
추세선



상위품목



온라인쇼핑
클러스터

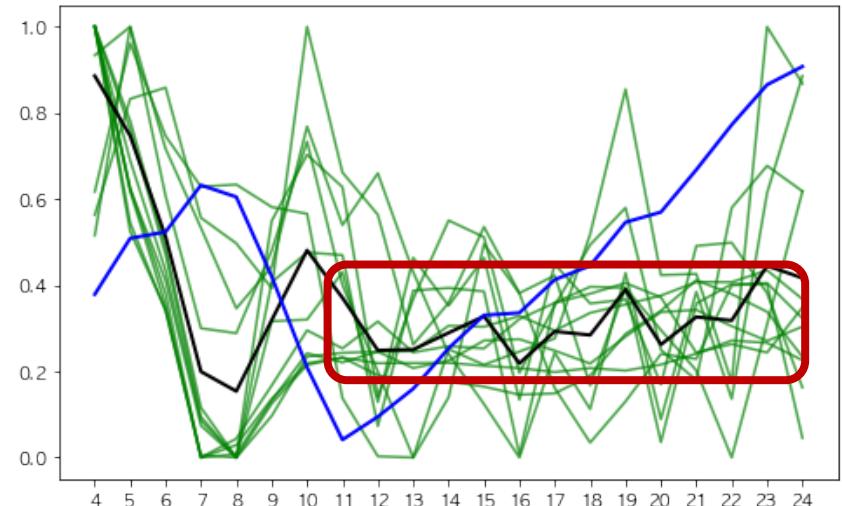


감소 클러스터 17번 INSIGHT

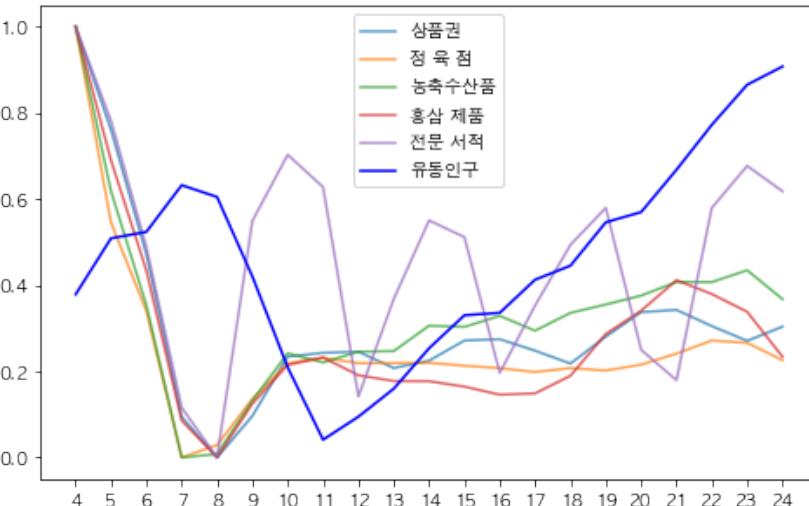
상품권, 정육점, 농축수산물, 홍삼제품, 전문서적은 코로나로 인하여 매출이 거의 변화하지 않았다.
오히려 4주차 설날의 영향(oulier)으로 보인다.
또한 온라인에서 음식서비스 매출은 38% 증가하였다.

‘음식료품’ 클러스터

유동인구
추세선



상위품목



온라인쇼핑
클러스터

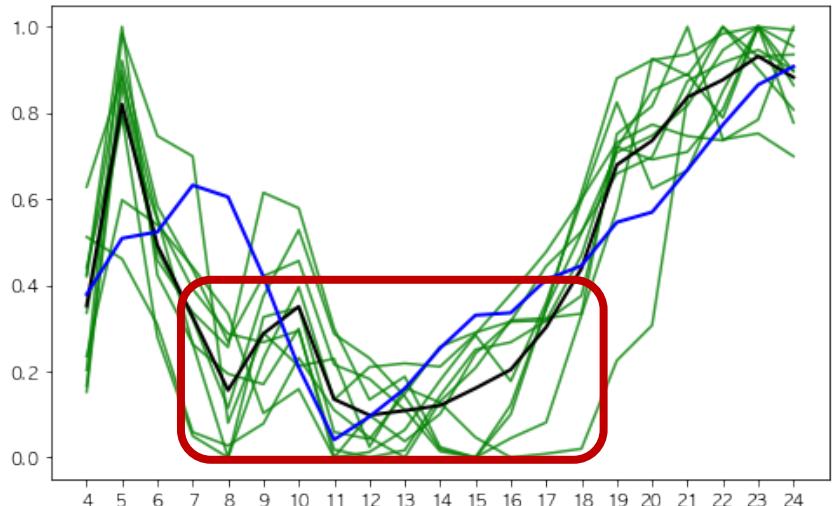


감소 클러스터 13번 INSIGHT

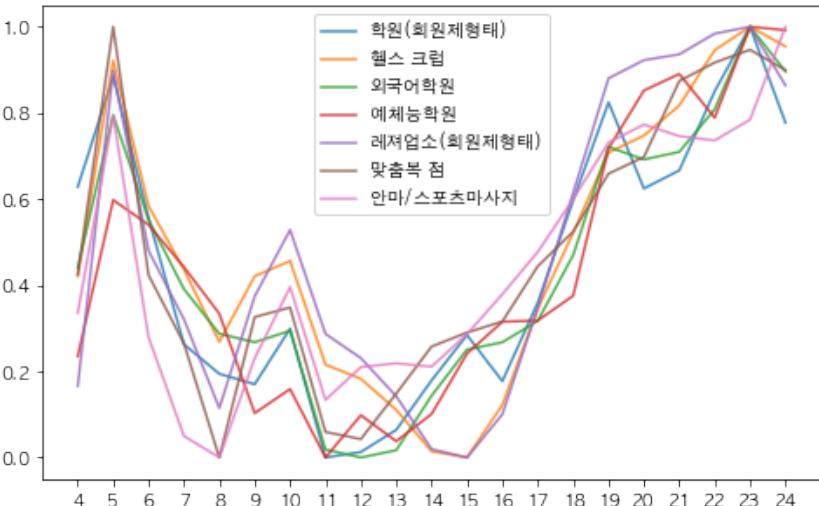
학원, 헬스클럽, 피트니스센터, 맞춤복, 패션 액세서리는 유동인구의 증감과 거의 증감이 일치한다.
다만, 현재(11월) 코로나 확산기로 매출 타격이 예상된다.

‘학원, 헬스’ 클러스터

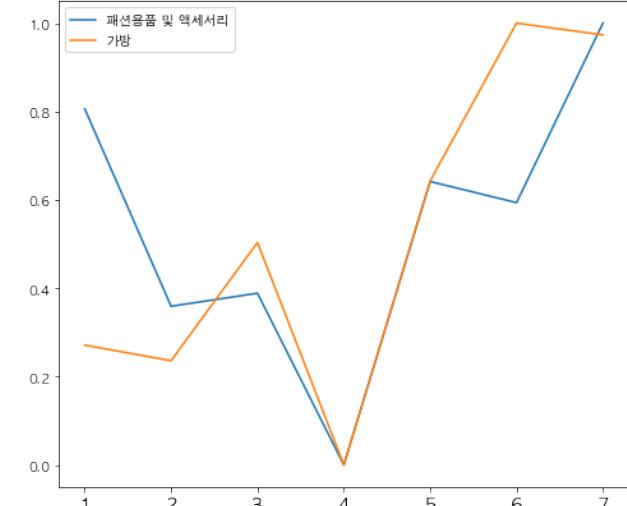
유동인구
추세선



상위품목



온라인쇼핑
클러스터

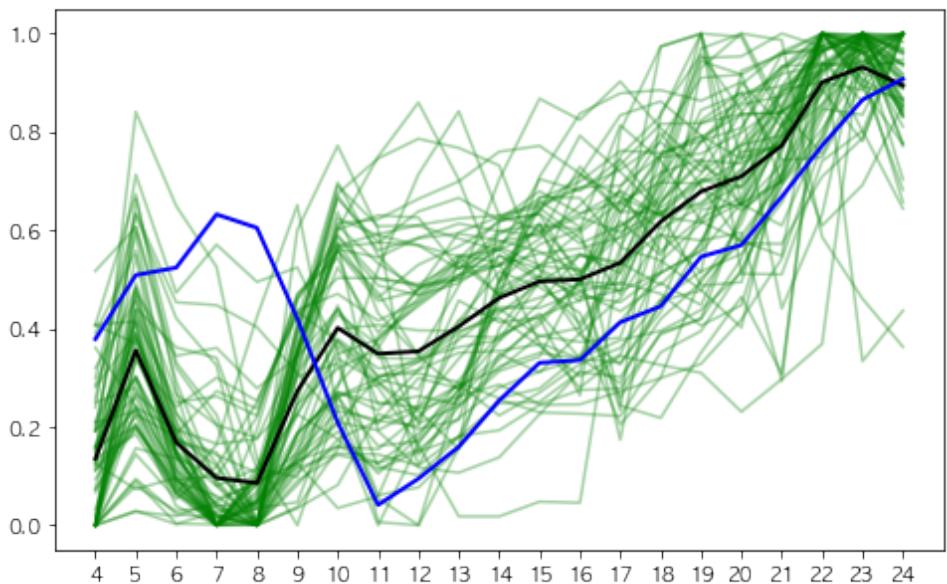


증가 클러스터 3번 INSIGHT

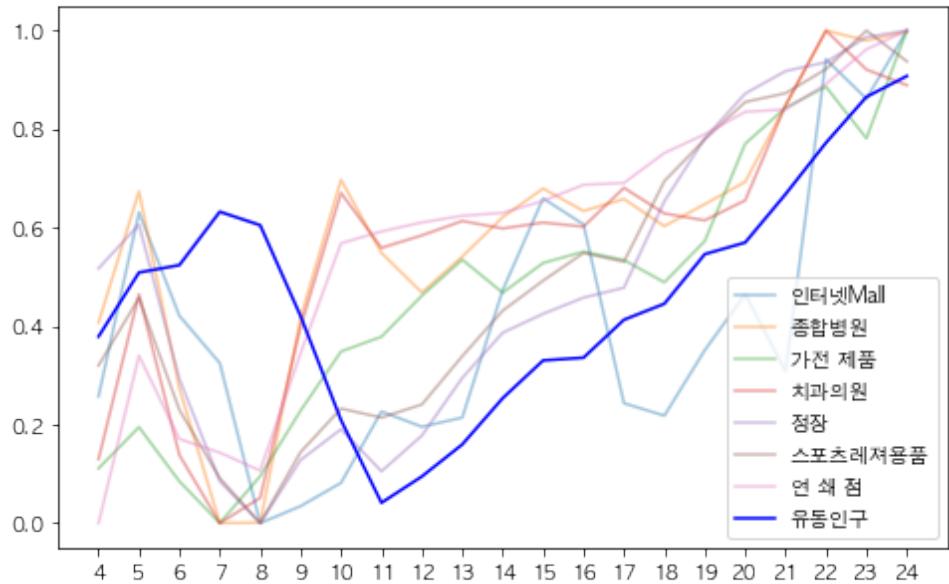
인터넷몰, 병원, 가전, 스포츠레저용품, 프렌차이즈는 일시적인 매출에 영향을 받았다.
최다 빈도 3번 클러스터의 대다수가 기업고객대상 업종들이다.

‘온라인, 병원, 스포츠용품’ 클러스터

유동인구
추세선



상위 품목

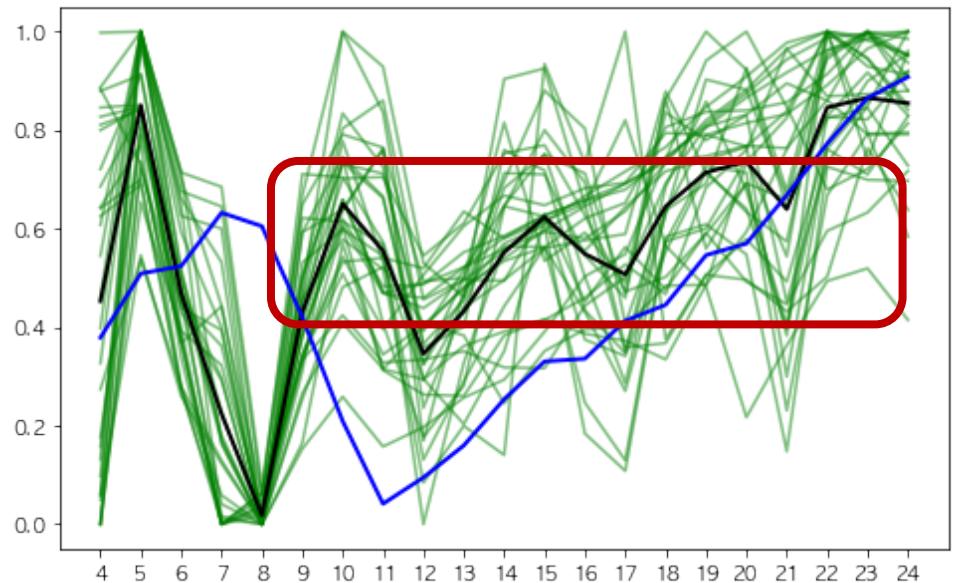


▶ 증가 클러스터 5번 INSIGHT

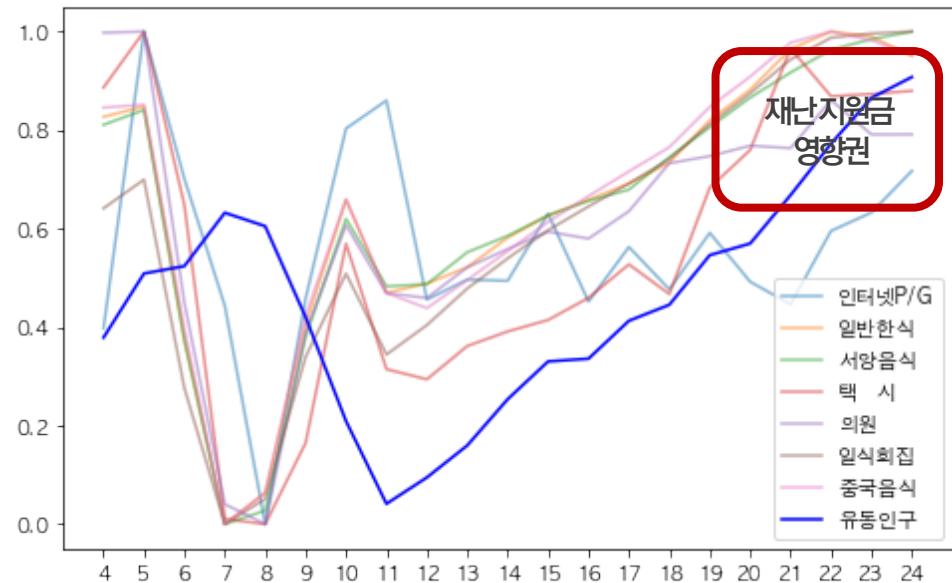
대행 결제, 음식점은 코로나 대확산기에 직접적인 영향을 받았지만
매출이 바로 회복되는 모습을 보이고, 재난지원금으로 5월에 회복세가 두드러졌다.

‘음식점’ 클러스터

유동인구
추세선



상위 품목

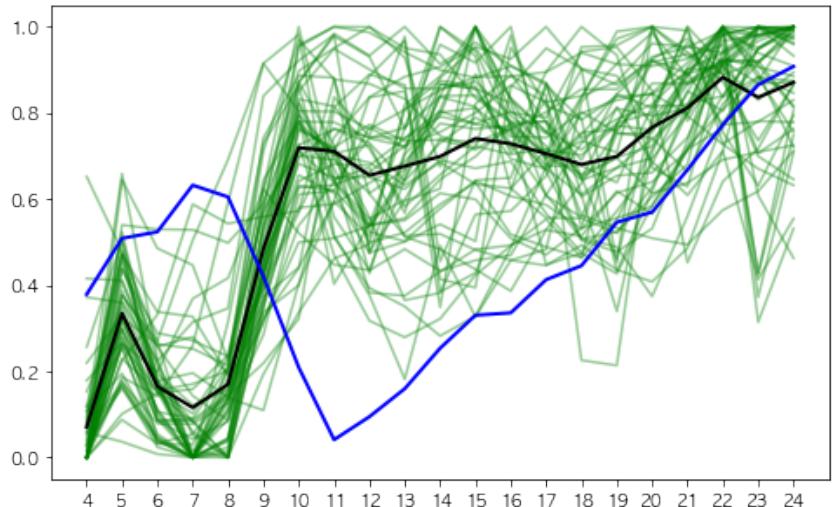


증가 클러스터 8번 INSIGHT

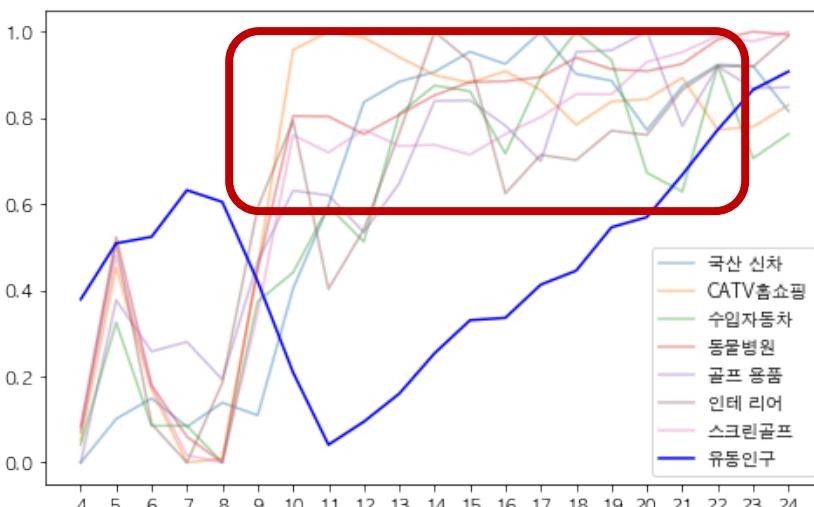
자동차, 홈쇼핑, 동물병원, 골프, 인테리어는 오히려 코로나 이전보다 이후 매출이 증가하였다.
언텍트 라이프스타일의 소비스타일을 대변하는 클러스터이다.

‘코로나 특수’ 클러스터

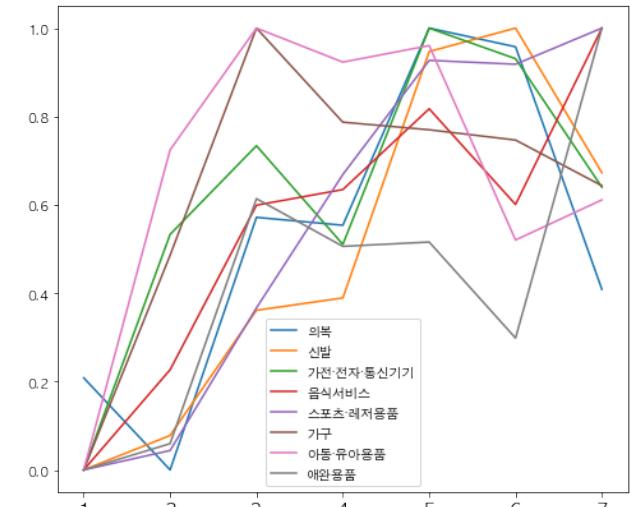
유동인구
추세선



상위품목



온라인쇼핑
클러스터

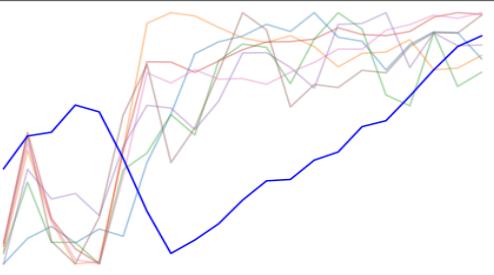




프라이빗 이코노미

비용이 좀 더 들더라도 언택트에 맞추어
오프라인 회원제(프라이빗) 서비스를
강화하여 운영하고 있다.

증가(8번)클러스터



코로나 특수 업종

자동차
골프용품
스크린골프
인테리어



회원제

기타회원제형태업소
자동차서비스(회원제)
건강식품(회원제)
서적출판(회원제)

회원제 서비스 제안



#바(bar)

불특정다수에게 공개되지 않고 회원제 바(bar)



#호텔

대중 호텔이 아닌, 회원제 중심의 호텔로 변경



#영화관

단체관람에서 소수 관람 형태로 전환