

# 프로젝트 보고서

**코로나 확산에 따른 산업,분야별 소비 변화 분석**

# 01

## 프로젝트 목적

코로나가 확산됨에 따라 소비 매출에 어떠한 변화가 있는지 파악해보고자 한다.

코로나 확산 이후에 업종 별로 소비에 변화가 있는지에 대한 분석이 필요하게 되었다. 소비 변화에 따라 예측 데이터의 필요성이 대두 되면서 통계적 의사결정을 통해 데이터를 분석해보고자 한다.



## 02 분석 방법

### 상관관계 분석(Corrleation Analysis)

상관분석이란 확률론과 통계학에서 두 변수간에 어떤 선형적 또는 비선형적 관계를 가지고 있는지 분석하는 방법이다.  
두 변수는 독립된 관계거나 상관된 관계일 수도 있다.

도매 및 소매업	0.510989
보건업 및 사회복지 서비스업	0.098901
예술, 스포츠 및 여가관련 서비스업	-0.376892
숙박 및 음식점업	-0.774725



상관계수 값의 범위는 **-1~1** 사이에 위치하며  
**양수**에 가까울수록 상관관계가 높다고 보며  
**음수**에 가까울수록 상관관계가 없다고 본다.

두 개의 변수에 사이의 상관분석 뿐 아니라  
두 개의 변수그룹 사이의 상관관계를 분석하는  
정준상관분석도 생각해볼 수 있다.



## 03

# 분석 프로세스

### 데이터 수집

코로나 확진자 현황 데이터  
업종, 산업 별 매출 데이터  
소비패턴 분석 동향 데이터  
소비자 물가지수 데이터

### 데이터 전처리

수집 된 데이터를 통해  
새로운 데이터프레임 생성  
업종,산업 별 데이터 셋 분리  
결측치 제거

### 데이터 시각화

matplotlib  
- barplot  
- boxplot  
- subplots 등을 활용  
  
seaborn  
- lineplot  
- heatmap

### 데이터 분석

상관관계 분석  
  
그래프와 분석 결과를 토대로  
인사이트 도출

# 04 참고 데이터

inent	location	date	total_cases	new_cases
	Afghanistan	2020-02-24		1
	Afghanistan	2020-02-25		1
	Afghanistan	2020-02-26		1
	Afghanistan	2020-02-27		1
	Afghanistan	2020-02-28		1
	Afghanistan	2020-02-29		1
	Afghanistan	2020-03-01		1
	Afghanistan	2020-03-02		1
	Afghanistan	2020-03-03		2
	Afghanistan	2020-03-04		4
	Afghanistan	2020-03-05		4
	Afghanistan	2020-03-06		4
	Afghanistan	2020-03-07		4
	Afghanistan	2020-03-08		5
	Afghanistan	2020-03-09		7
	Afghanistan	2020-03-10		8
	Afghanistan	2020-03-11		11
	Afghanistan	2020-03-12		12
	Afghanistan	2020-03-13		13
	Afghanistan	2020-03-14		15
	Afghanistan	2020-03-15		16

총지수	식료품 및 비주류음료	식료품 및 비주류음료	4
총지수	계	식료품	비
106.09	110.68	111.39	
106.06	110.31	111.1	
105.97	111.51	112.26	
105.55	111.94	112.63	
105.32	111.18	111.77	
105.51	111.61	112.28	
105.42	111.61	112.3	
106.05	115.19	116.21	
106.51	119.92	121.34	
105.96	119.67	120.94	
105.88	115.03	115.95	
105.91	114.42	115.38	
106.6	117.01	118.09	

## 소비자 물가지수

	세무항목	2021년 01분기	2020년 04분기	2020년 03분기	2020년 02분기
전체		86.4	90.8	99.6	9
주연령	20대	95.7	96.8	102.6	9
주연령	30대	83.7	85.7	96.9	9
주연령	40대	84.6	87.4	99.7	9
주연령	50대	83.3	91.2	99.6	8
주연령	60대 이상	90.6	95.9	100.5	9
원 수	1인 가구	90.5	93.9	100.5	9
원 수	2-3인 가구	87.7	91.4	99.3	9
원 수	4인 이상 가구	78.3	85.7	99.1	9
군 가구 소득	99만원 이하	90	96.4	102.2	9
군 가구 소득	100만원-199만원 미만	92.3	90.5	99.5	9
군 가구 소득	200만원-299만원 미만	87.2	89.7	97.1	9
군 가구 소득	300만원-399만원 미만	83.2	89.8	98.9	8
군 가구 소득	400만원-499만원 미만	83.8	87.8	102.5	
군 가구 소득	500만원 이상				
군 가구 소득	500만원-599만원 미만	81.6	91.9	99.7	9
군 가구 소득	600만원 이상	87.4	89.3	98.1	9
전체		103.2	104.3	103.2	10
주연령	20대	111.4	107.4	104.4	10
주연령	30대	111	108.1	103.4	10
주연령	40대	105.3	105.5	104.5	10
주연령	50대	99.2	103.8	103.5	9
주연령	60대 이상	96.2	99.4	100.6	
원 수	1인 가구	106	102.9	102.8	10
원 수	2-3인 가구	101.9	103.5	102.6	9
원 수	4인 이상 가구	103.7	107.9	104.8	10

## 소비 패턴 동향

20190101	가정생활/서비스	189
20190101	교육/학원	142
20190101	미용	464
	스포츠/문화/레	
20190101	저	1,752
20190101	여행/교통	260
20190101	요식/유흥	9,264
20190101	유통	11,307
20190101	음/식료품	1,192
20190101	의료	432
20190101	자동차	186
20190101	주유	1,782
20190101	패션/잡화	415
20190102	가전/가구	239
20190102	가정생활/서비스	566

## 업종별 결제 건수

	2019. 02	2019. 03
	2890	3339
	135	144
	935	1083
	37	40
	115	129
	370	390
	82	110
	145	170

## 산업별 결제 건수

## 세계 확진자 현황

출처 : OurWorldindata.org

출처 : 서울 열린데이터 광장

출처 : 공공데이터 포털

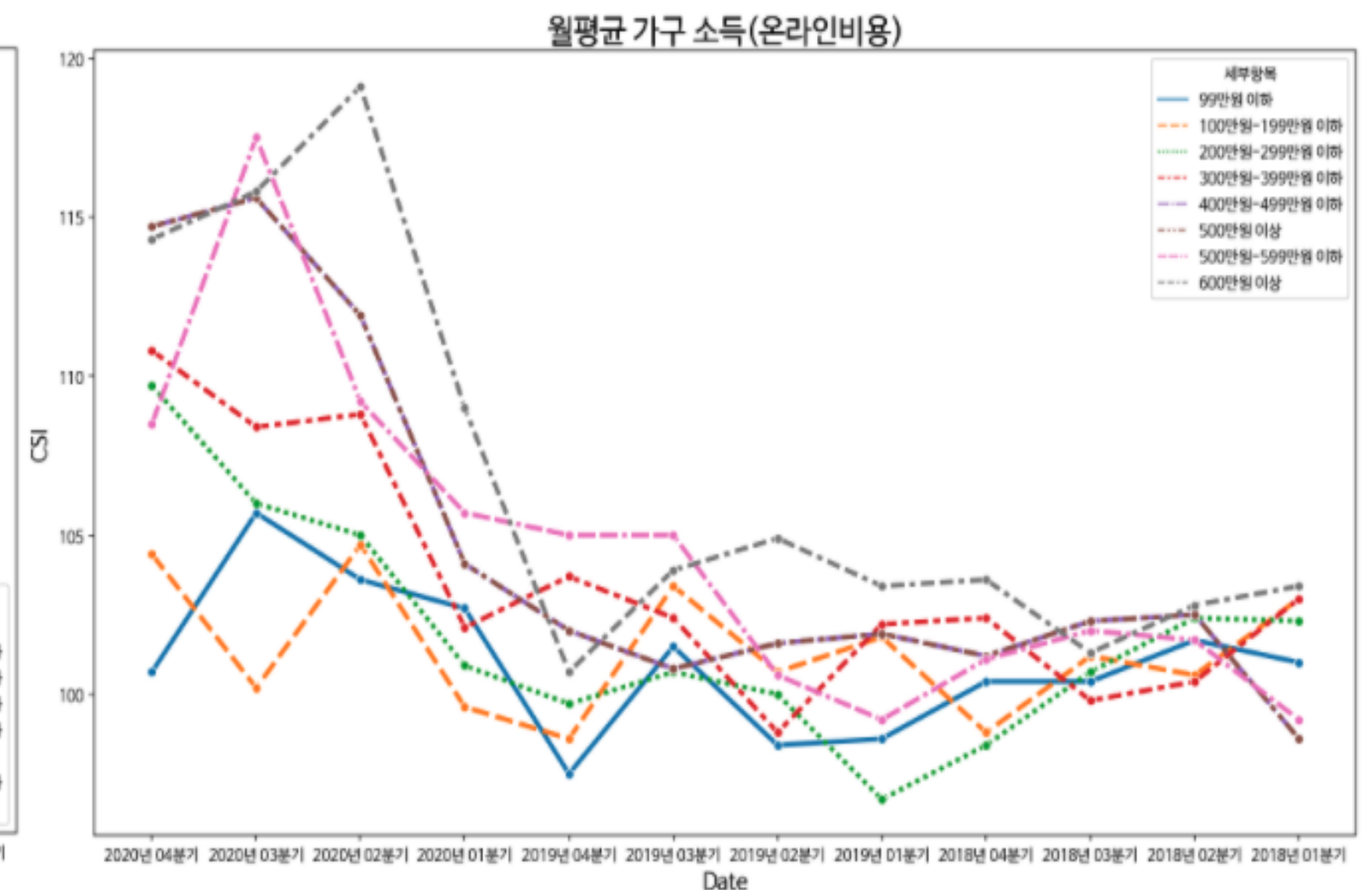
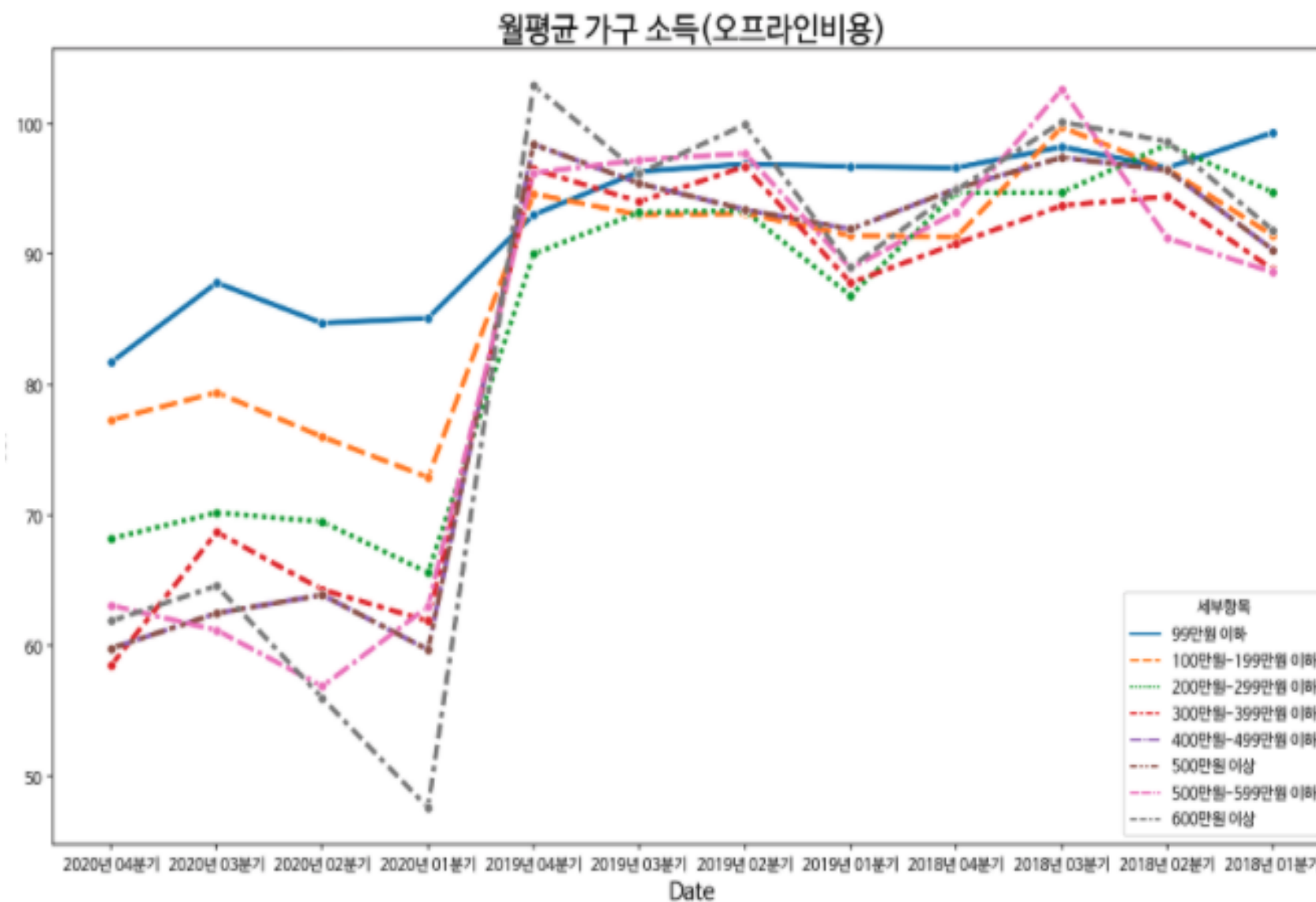
출처 : 금융데이터거래소

출처 : 통계청 '한국표준산업분류  
별 월간 국내카드승인실적'

# 05 온,오프라인

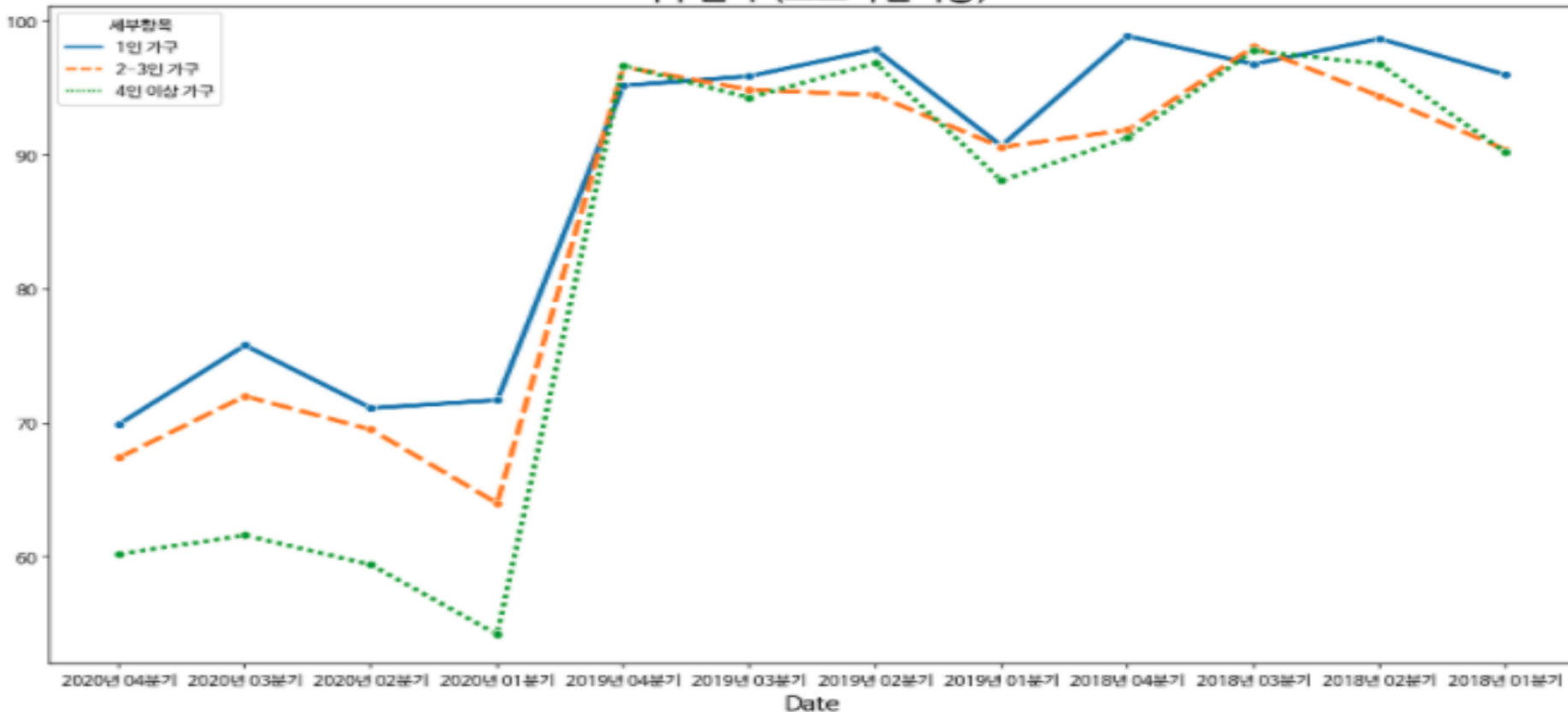
**가설 : 소비패턴으로 사회적 코로나 현황파악 및 거리두기 지표를 파악할 수 있다**

소비패턴 분석 동향 데이터를 사용하여 18년도 1분기부터 20년도 4분기까지 분기별 오프라인, 온라인 비용을 가구주연령, 가구원 수, 월평균 가구소득에 따라 소비패턴을 분석하여 시각화.

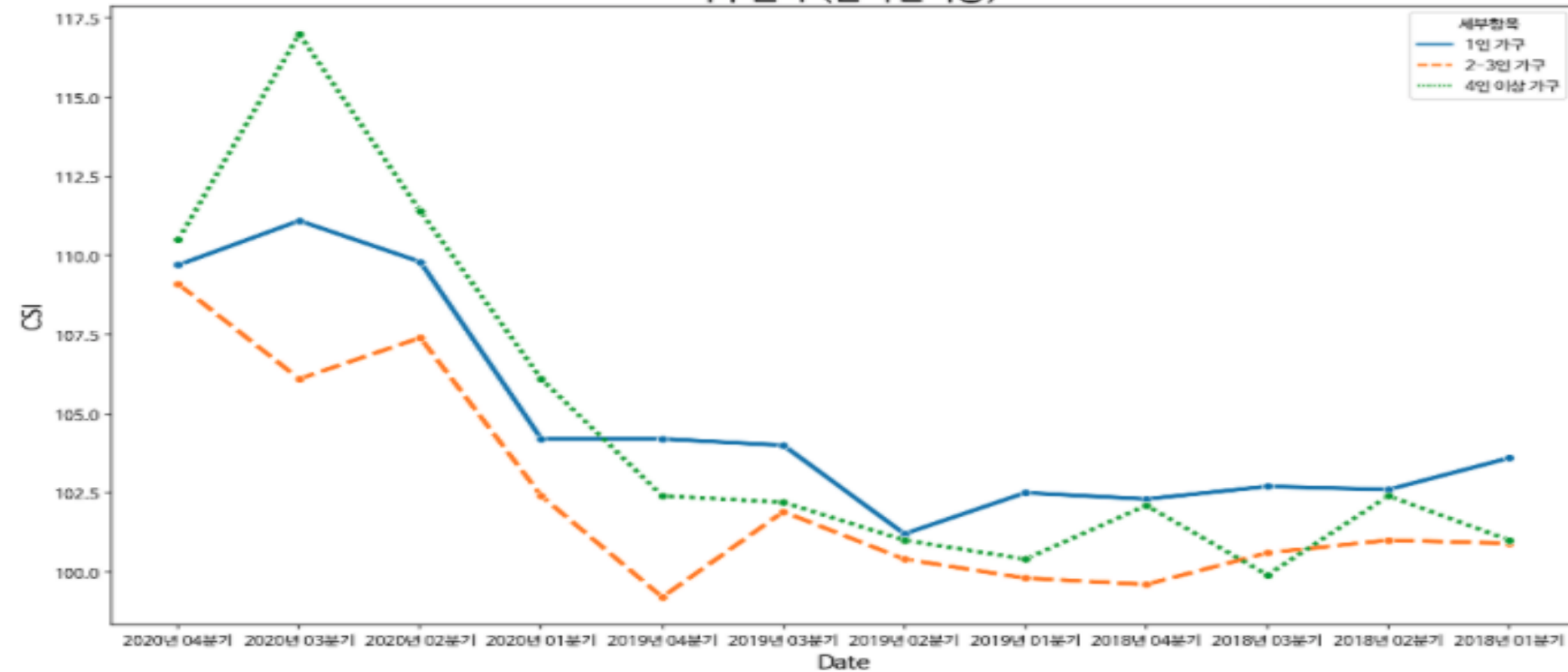


# 05 몬,오프라인

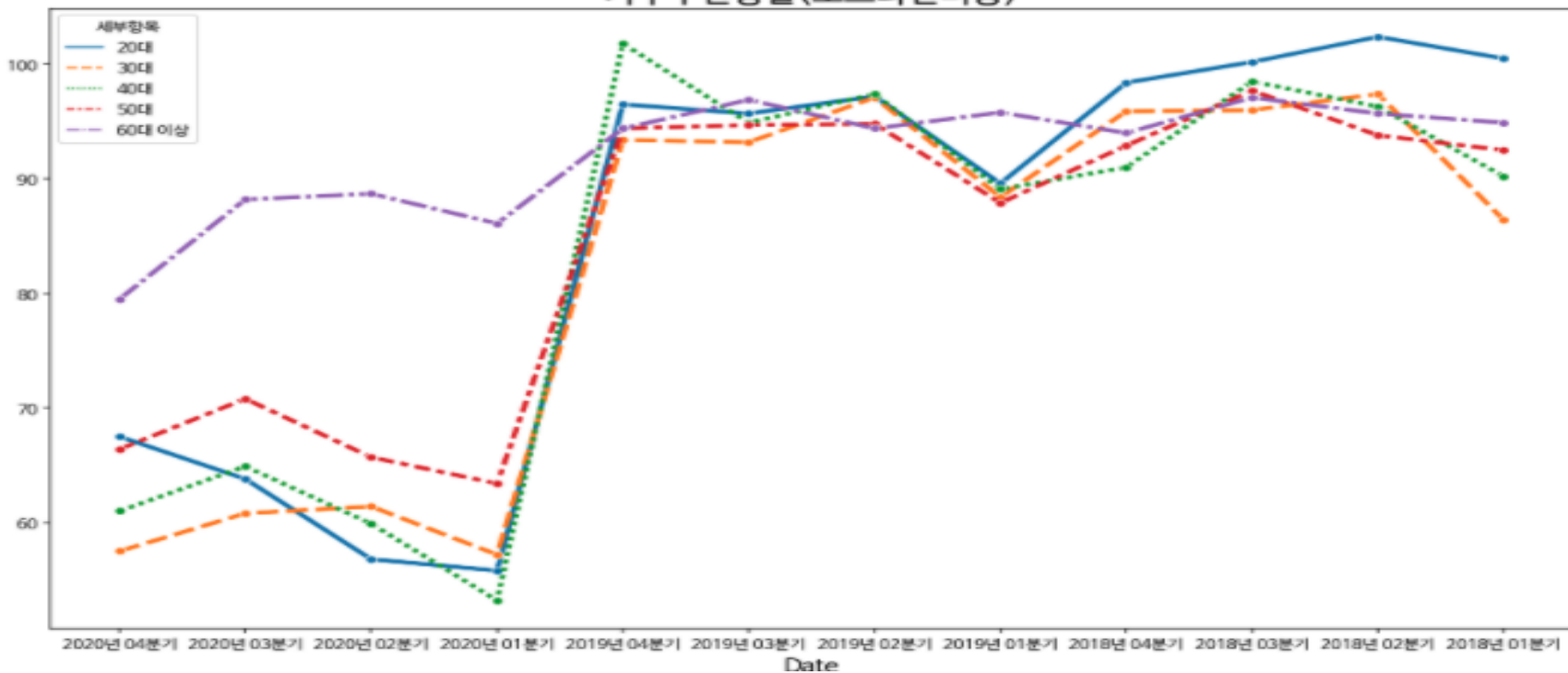
가구원 수(오프라인비용)



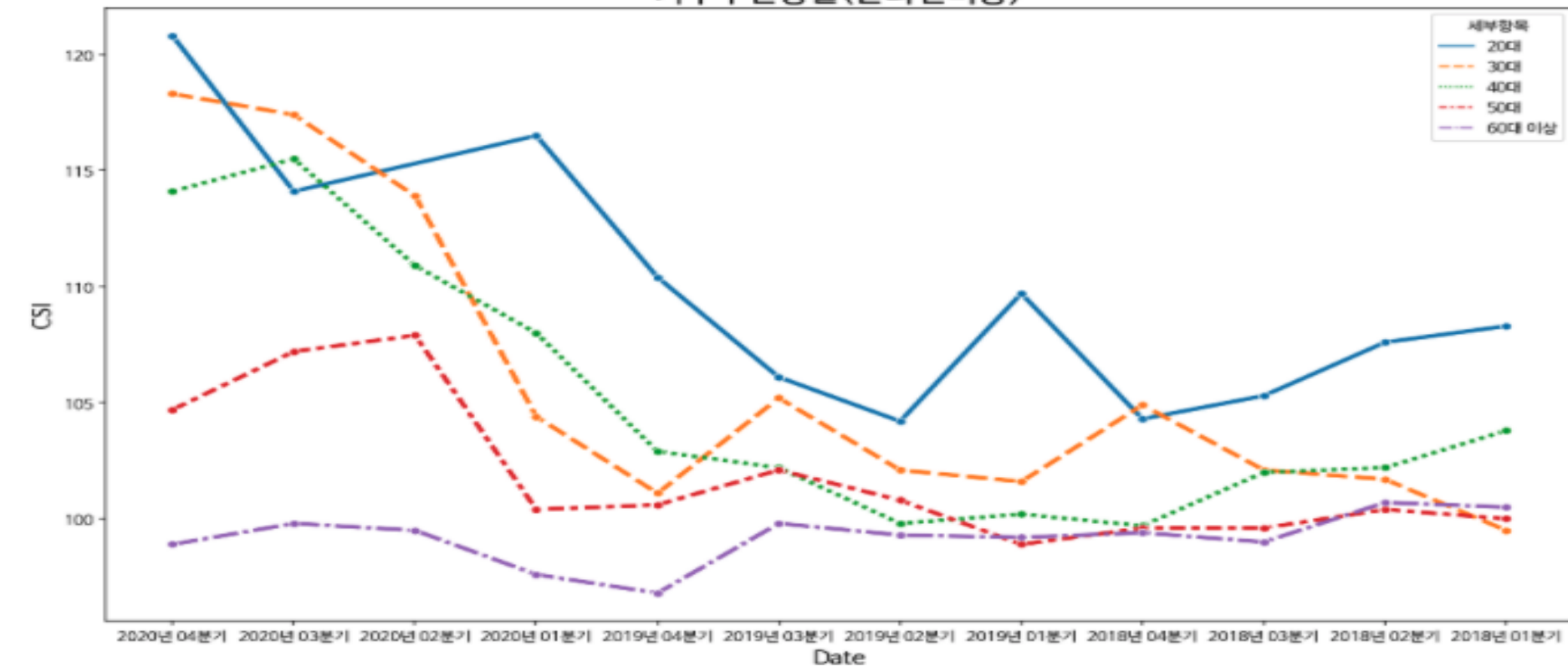
가구원 수(온라인비용)



가구주연령별(오프라인비용)



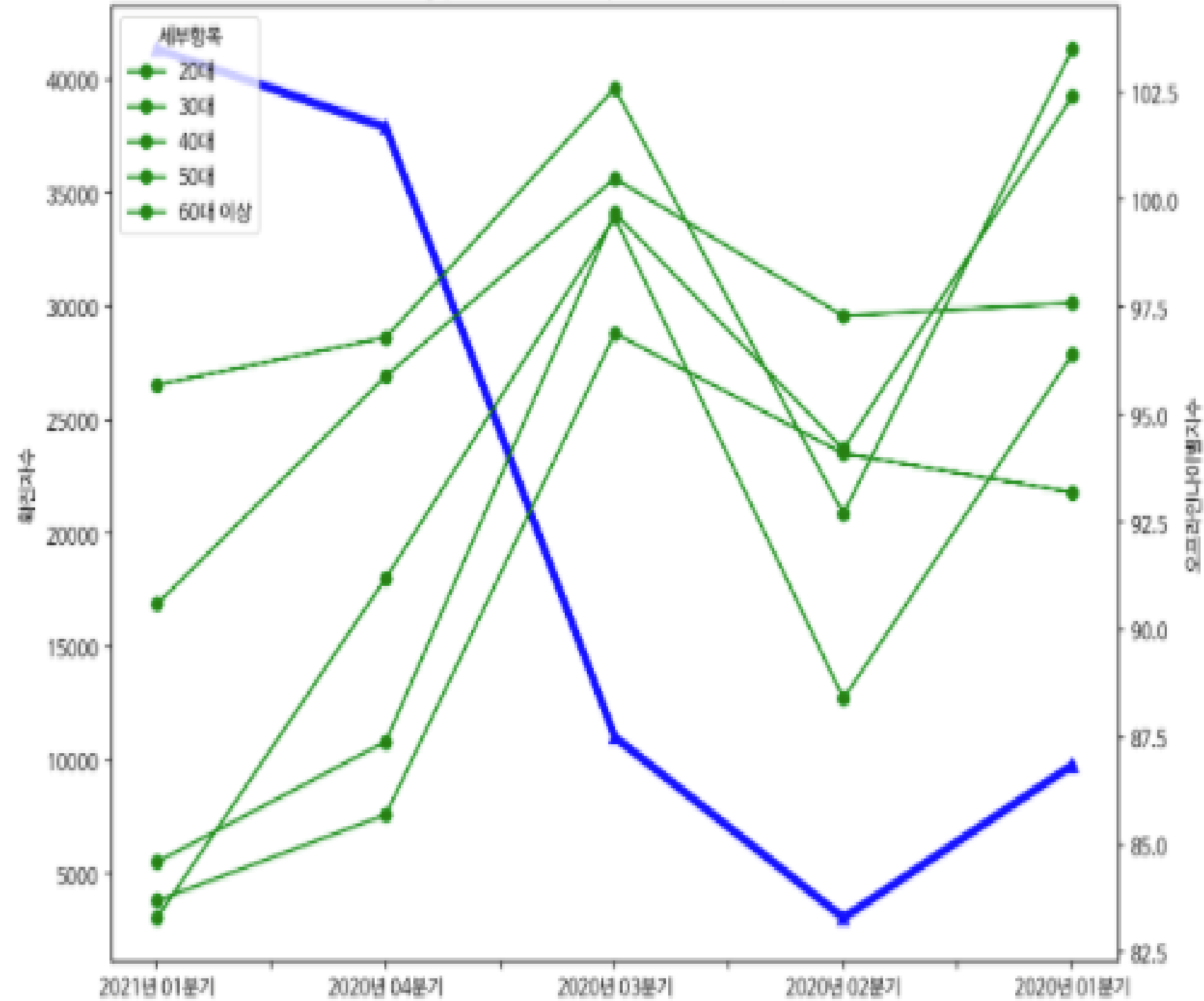
가구주연령별(온라인비용)



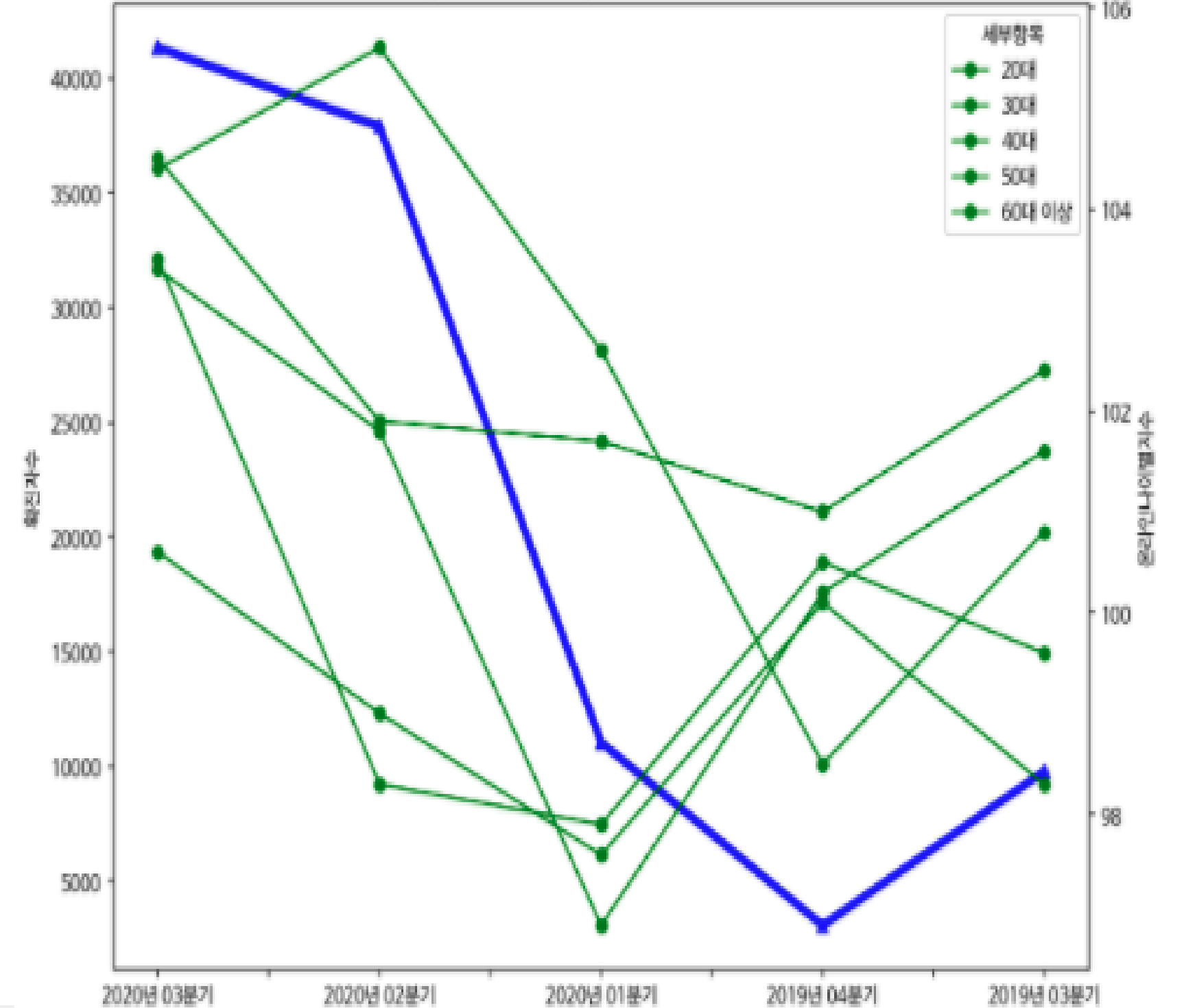


# 05 온,오프라인

분기별 가구주 연령(오프라인 문화생활)와 분기별 코로나 확진자 수 대조 그래프



분기별 가구주 연령(온라인 문화생활)와 분기별 코로나 확진자 수 대조 그래프





## 05 --- 온,오프라인

### 분석 및 결론

가구주연령, 가구원 수, 월평균 가구소득에 따른 소비패턴의 급한 변동이 있는 구간은 모두 2020년 1분기 이다. 이때 오프라인 비용은 급격히 줄고 온라인 비용은 급격히 늘어난다.

실제로 2020년 1분기는 코로나가 급격히 퍼지는 시점이자 이에 따라 오프라인 활동이 급격히 감소했으며 온라인 활동이 늘어났다.

월평균 가구소득 그래프에서 소득이 낮을수록 소비패턴의 변화가 적다.

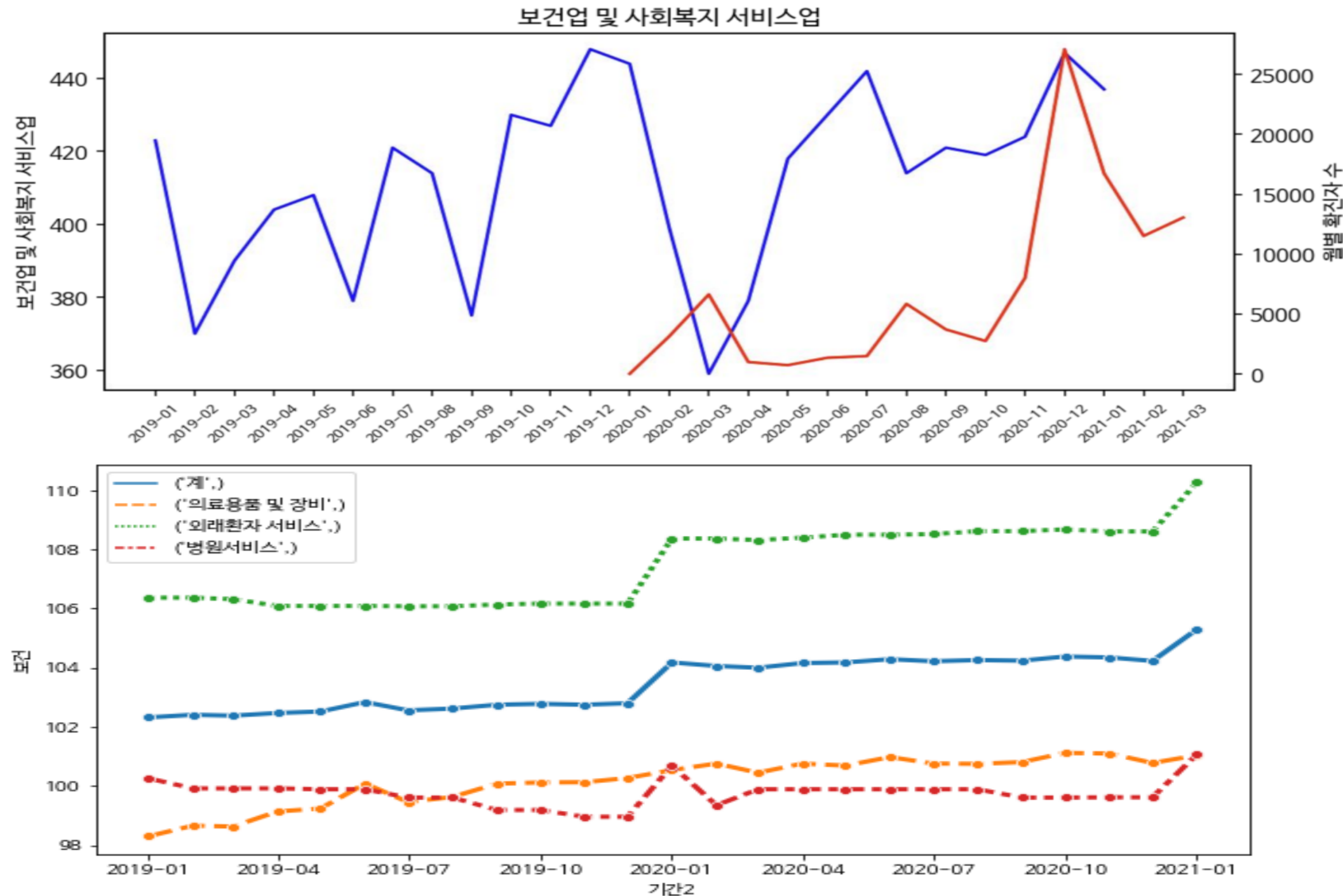
가구주 연령별 그래프에서 다른 연령층에 비해 60대 이상은 다른 연령대에 비해 변화도가 적으며 점진적으로 감소하거나 증가하는 양상을 보인다.

결론적으로, 코로나로 인해 소비패턴 변화의 폭이 크거나 작은 그룹을 파악할 수 있었고 이를 이용하여 전반적인 사회 분위기를 파악할 수 있다.

---

# 06 의료

가설 : 질병 확산 정도에 비례하는 변화가 있었을 것이다.



## 데이터 분석

- ‘코로나 확산’과 ‘보건업 및 사회복지 서비스업 카드 승인 매출’ 사이의 낮은 상관 관계
- 스피어만 상관 계수 : 0.098901
- 최초 유행 시 급격한 변화, 이후 코로나 확산에 따른 관계 확인 어려움
- 물가지수 역시 최초 코로나 대 유행 시기에 지수 급증 이후 큰 변화 없이 유지

## 결론 도출

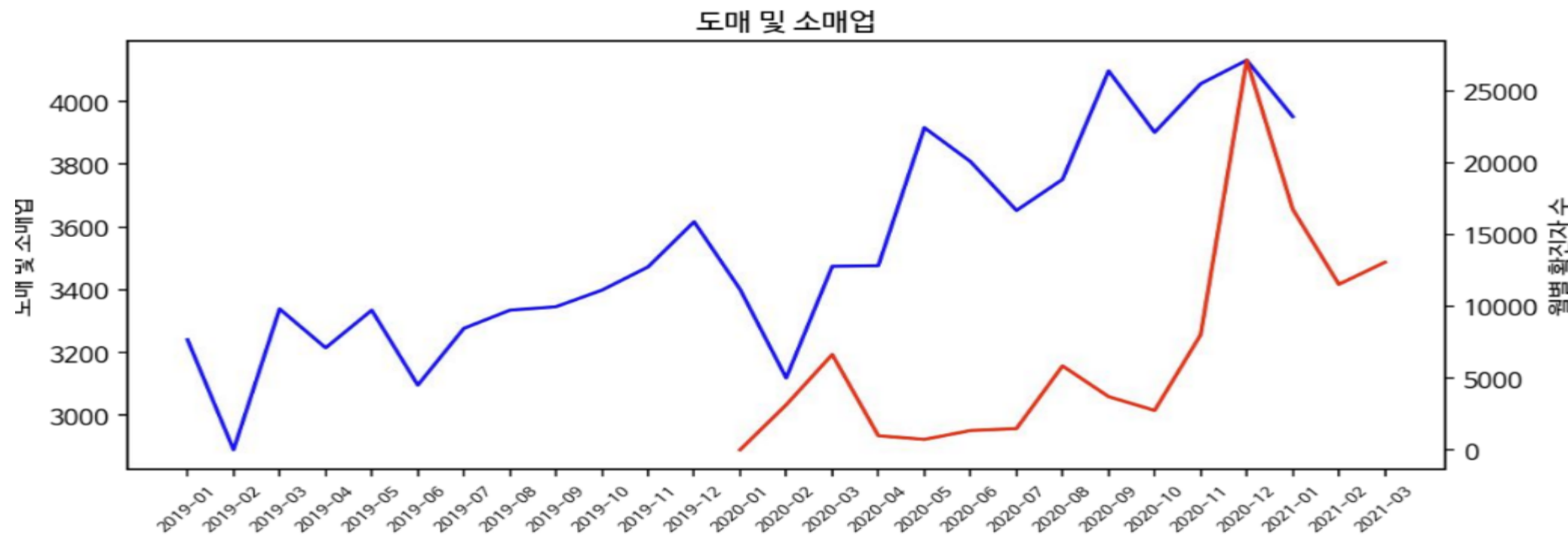
- 공공 개입 / 의료계 파업 등의 이슈로 정확한 상관관계 파악이 어려움

## 추가 연구

- 시장 외 요인 타임라인 분석 필요

# 07 도,소매업

가설 : 생필품 등의 소비와 재난 상황 발생에 따라 양의 상관 관계를 가질 것이다



## 데이터 분석

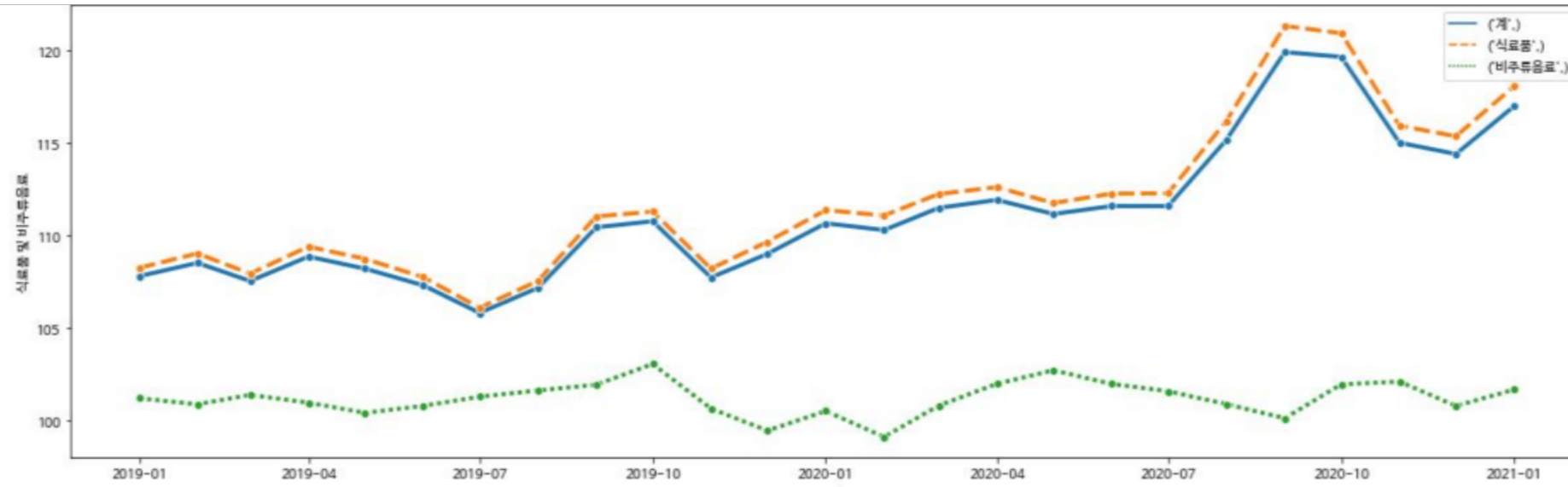
- '코로나 확산'과 '도매 및 소매업' 승인된 카드매출 사이의 '강한 양의 상관 관계'
- 스피어만 상관 계수 : 0.510989
- 특히 식료품에 대한 소비자 물가지수
- > 뚜렷한 우상향, 시기별 물가 변동 폭이 크게 증가 (7~10월 물가 상승 폭 증가)

## 결론 도출

- '거리두기' 등의 사회적 운동에 따른 배달식 및 간편식 소비 증가

## 추가 분석

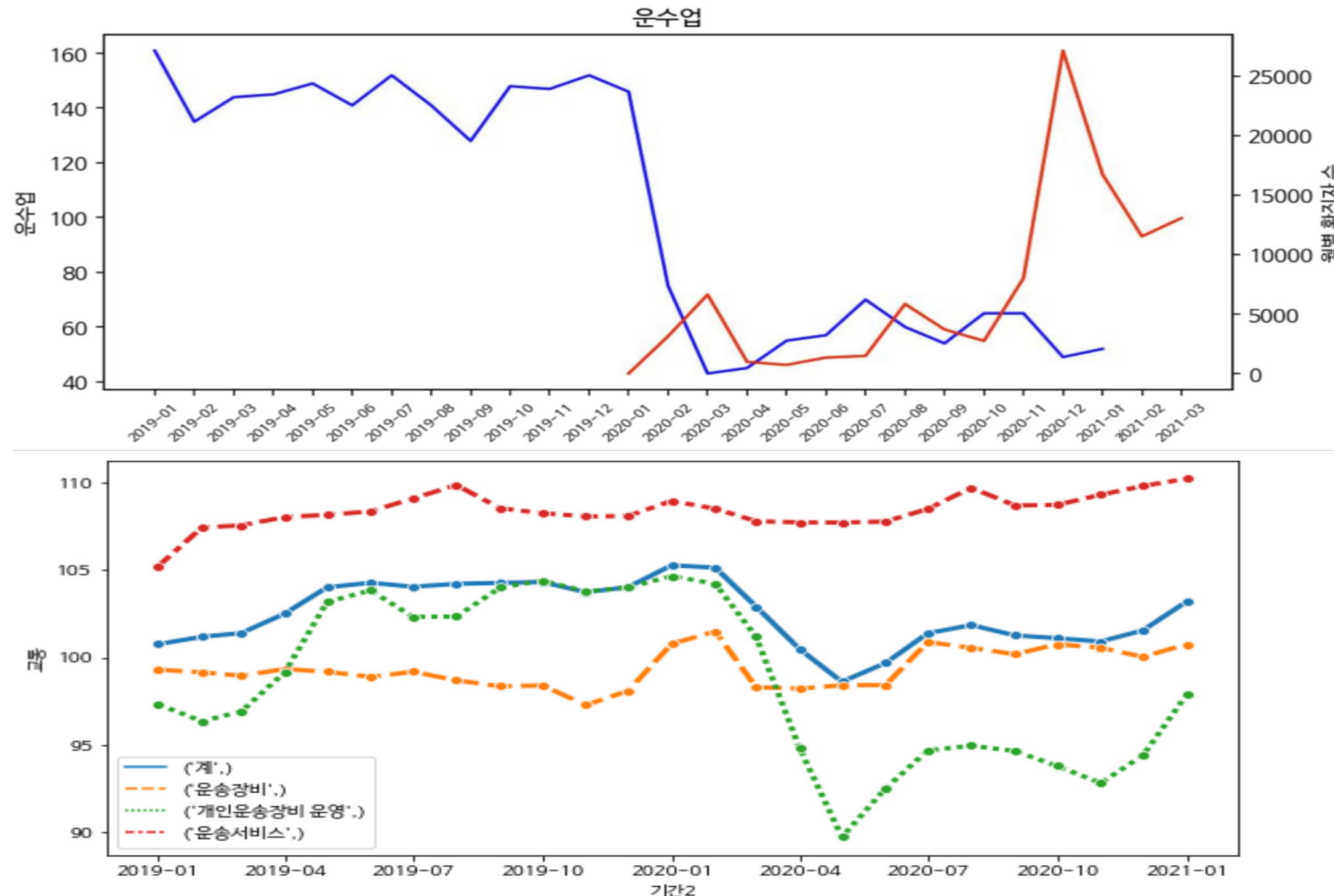
- 식료품 외 품목에 대한 매출 확인 필요



통계청 카드 승인 매출(상) / 소비자 물가지수(하)

## 08 여행,교통

가설 : 코로나 확산에 따른 비대면 / 거리두기 등 사회적 움직임에 따라 음의 상관 관계를 보일 것이다.



### 데이터 분석

- '코로나 확산'과 '운수업'의 강한 음의 상관 관계
- 스피어만 상관 계수 :  $-0.376892$
- 물류 유통에 비해 여객 운수 업종에 큰 타격
- 최초 유행 시 발생한 급락 이후 회복세 X

### 결론 도출

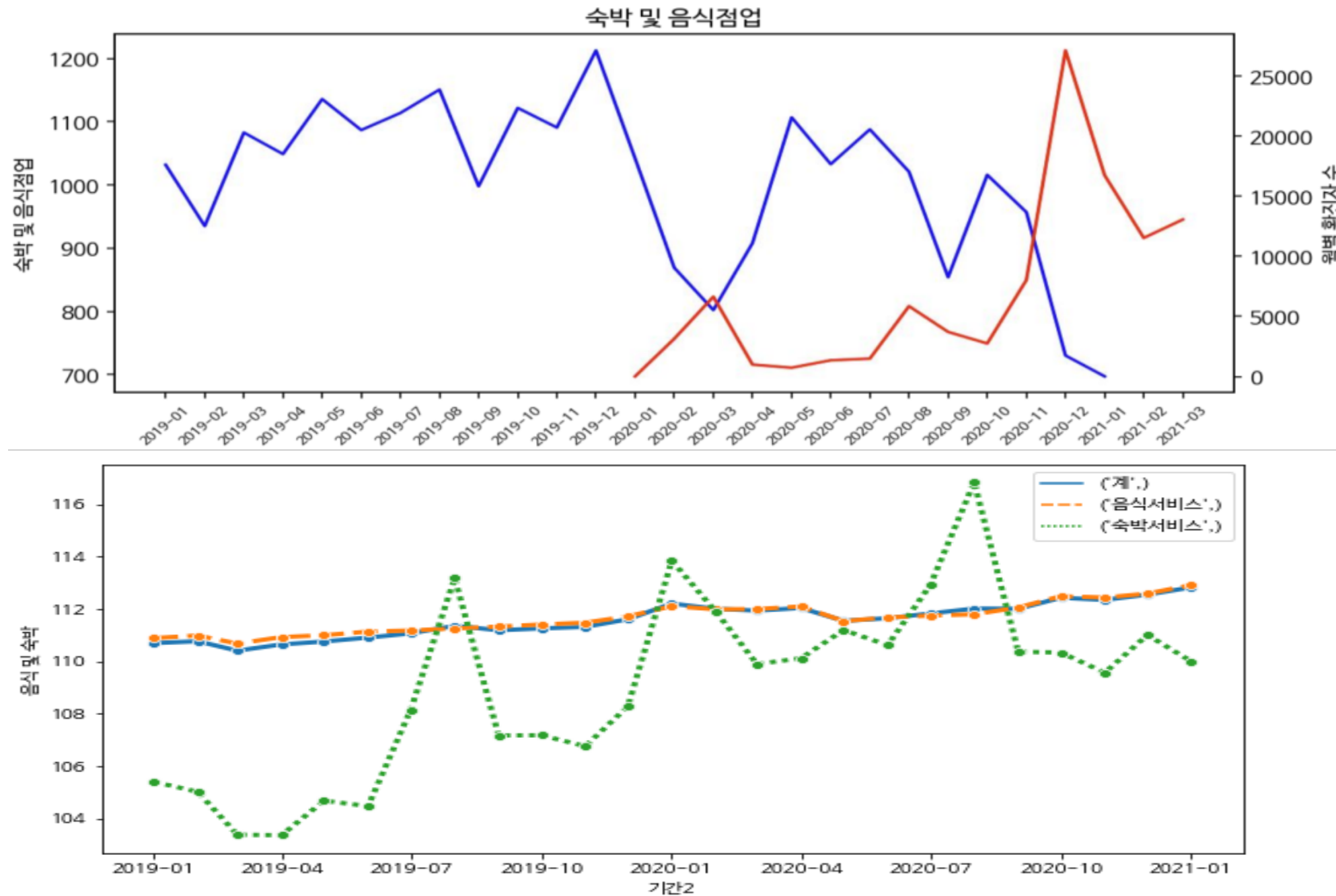
- 전염성이 높은 질병 확산에 대한 불안 및
- 국가적 통제로 여객/운수업에 대한 소비가 대폭 감소

### 추가 분석

- 여객 운수업과 물류 운수업을 분리하여 분석 필요

# 09 요식, 숙박

가설 : 코로나 확산으로 음식 및 숙박업이 영향을 받았을 것이다.



## 데이터 분석

- 숙박 및 음식점업 과 코로나 확진자수의 음의 상관관계
- 상관계수 :  $-0.774725$
- 2020년도 2분기까지,  
결제 건수와 코로나 확진자 수 의 음의 상관관계
- 물가지수는 코로나 확산에 대해  
상관없이 일정한 수준을 유지함
- 숙박서비스의 물가지수의 변화는  
코로나 확진자 수와는 상관없이 변동폭이 큼

## 결론 도출

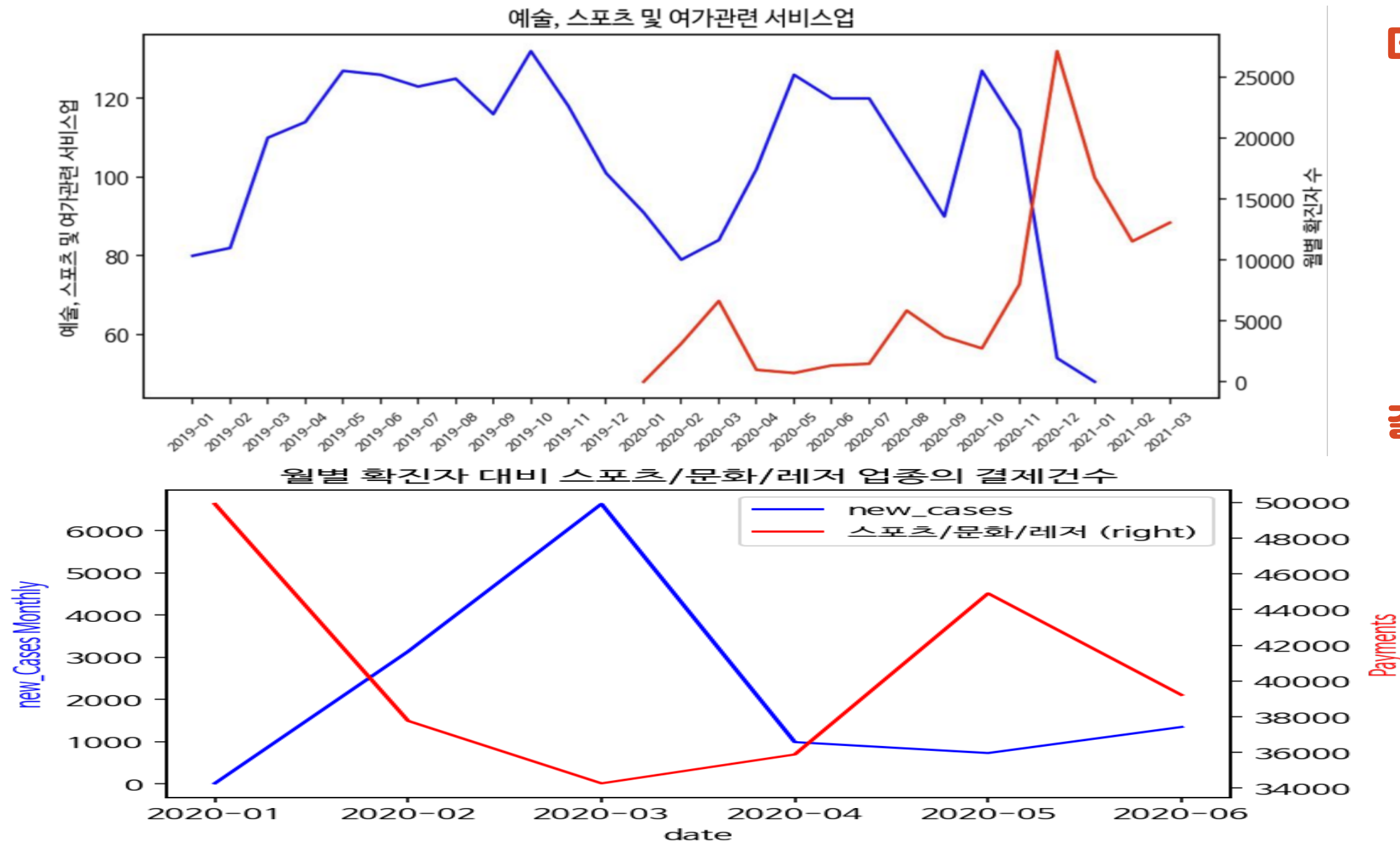
- 그래프와 상관계수를 많을 가지고 정확한 판단이 불가능



# 10

## 스포츠,여가

가설 : 코로나 확산으로 스포츠, 여가 산업이 영향을 받았을 것이다.



### 데이터 분석

- 스포츠,여가 산업과 코로나 확진자수의 음의 상관관계
- 상관계수 :  $-0.569464$
- 모든 업종이 확산이 급증하였던 3월에 정반대의 지수를 보여준다

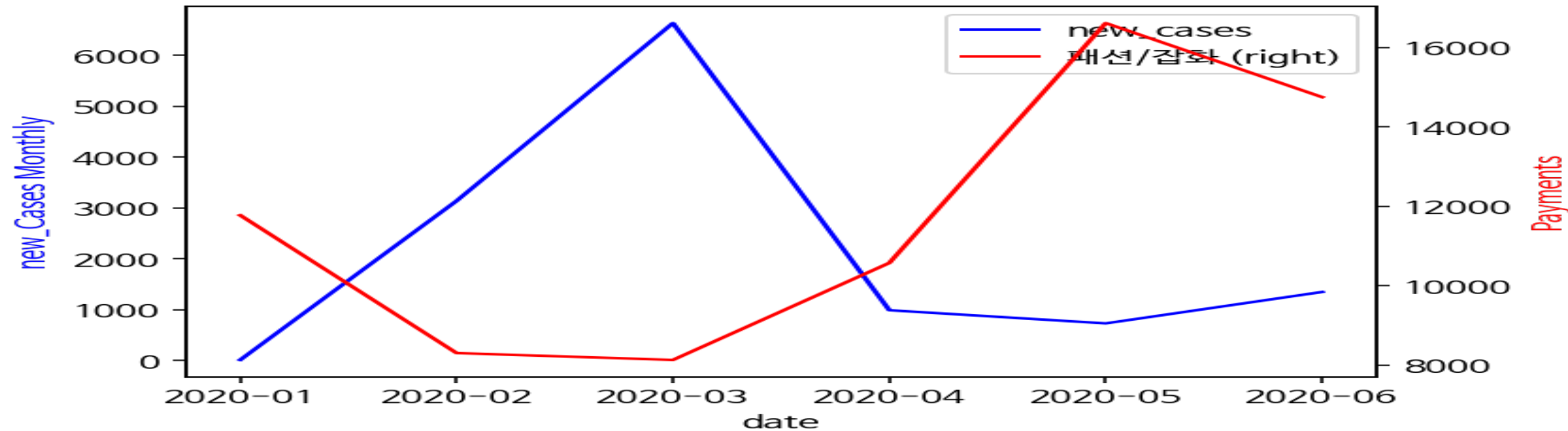
### 결론 도출

- 그래프와 상관계수를 많을 가지고 정확한 판단이 불가능
- 코로나 확산 후 그래프를 보고 영향을 받았을거라 추측

# 11 패션

가설 : 코로나 확산으로 패션,의복업이 영향을 받았을 것이다.

월별 확진자 대비 패션/잡화 업종의 결제건수

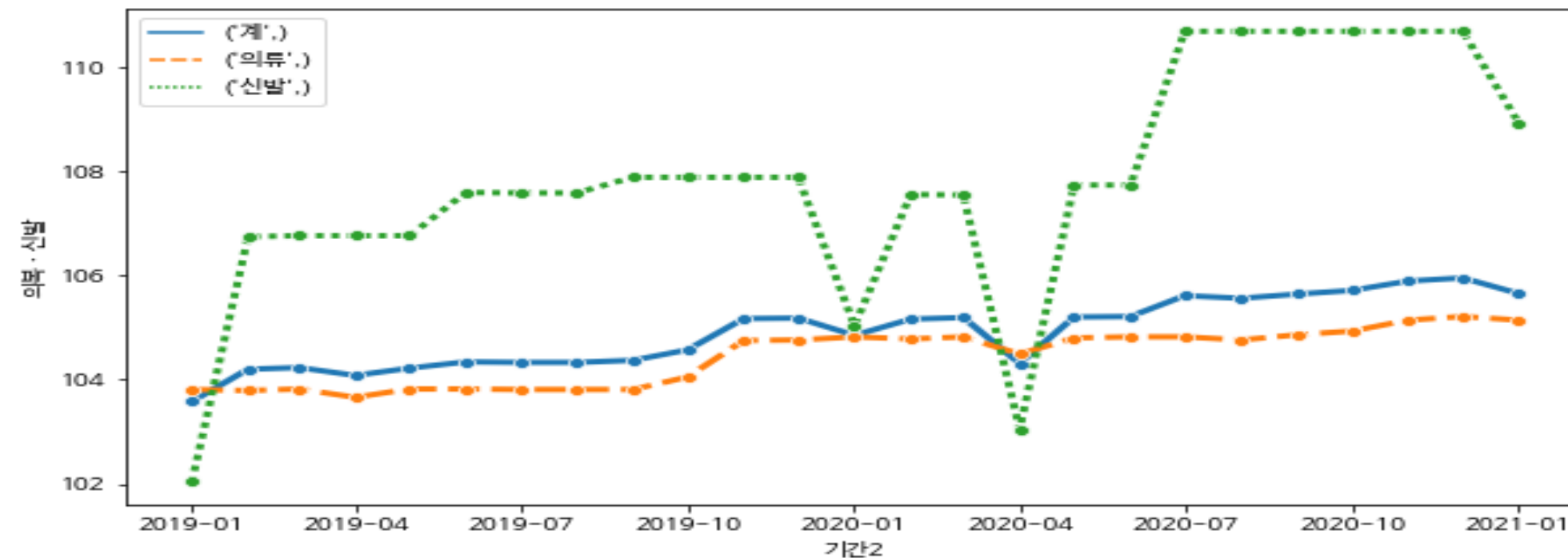


## 데이터 분석

- 신발의 물가지수의 변화는 코로나 확진자 수와는 상관없이 변동폭이 큼
- 모든 업종이 확산이 급증하였던 3월에 정반대의 지수를 보여준다

## 결론 도출

- 그래프와 상관계수를 많을 가지고 정확한 판단이 불가능





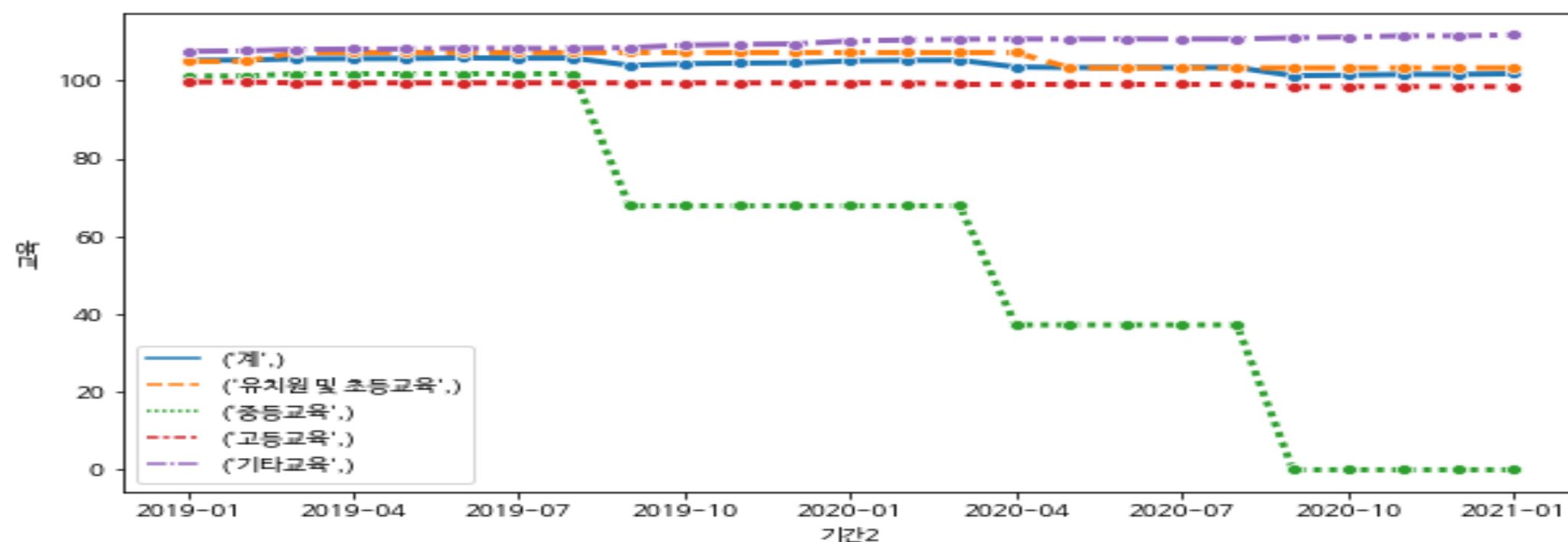
**가설 : 코로나 확산으로 음식 및 숙박업이 영향을 받았을 것이다.**



- 코로나 확산과 교육서비스의 경미한 음의 상관관계
- 상관계수 :  $-0.397796$
- 2020년도 2분기 까지, 코로나 확진자 수 월별 교육업종의 결제 건수 또한 음의 상관관계를 보임
- 물가지수 코로나 확진자 수의 증가 와 관련없이 일정한 수준을 유지함 (중등교육 제외)

## 결론 도출

- 특정 시기에 음의 상관관계를 보이고 있으나,  
전체적으로 파악하기에는 어려움**



## 한계점 및 애로사항

상관관계 만으로 산업 별 소비와 확진자 수 사이의 관계를 이끌어내기 어렵다.

카드 승인 매출로 소비 행태를 유추하여, 현금 소비가 강한 분야에서 오차가 발생할 수 있다.

데이터 수집 기관에 따라 분류가 상이하어, 세분화된 업종에 대한 분석이 어렵다.



# 감사합니다!

---

## 프로젝트 팀원

김채연, 김형원, 박찬호, 이종섭, 전해진