

DB 2차 프로젝트

# 코로나19 전후 카드이용건수 변화 분석

데이터분석과 구분성

# CONTENTS

**01** 배경 및 주제  
프로젝트 배경 및 주제

**02** 분석 - ①  
1차 대유행 소비 변화

**03** 분석 - ②  
일별 확진자와  
업종간 상관분석

**04** 한계 및 개선  
분석 한계 및 개선 방향

CONTENTS

## 프로젝트 배경

- ✓ 코로나19 팬데믹이 1년이상 장기화되는 현재,  
코로나19 전후로 사회 속에서의 소비는 많은 변화가 있었음
- ✓ 특히, 외부활동이 지양되면서 일부 산업군은 큰 타격을 입었음
- ✓ 이번 프로젝트를 통해 코로나 전후로 어떻게 소비량이 달라졌는지 확인하고,
- ✓ 또, 어떤 산업군이 큰 피해를 입었는지 확인하고자 함

## 프로젝트 주제

- ✓ 분석 - ①
  - ✓ 1차 대유행 전후, 업종별 카드 소비량 변화 분석
  - ✓ 탐색적 자료분석 (EDA)
- ✓ 분석 - ②
  - ✓ 일별 확진자의 수와 업종별 카드 소비량의 관계
  - ✓ 상관분석

## 사용 데이터

- ✓ 사용 데이터 - ①
  - ✓ 코로나 19에 따른 카드소비동향(결제건수)
  - ✓ 13개의 업종 대분류
  - ✓ 금융데이터거래소, 신한카드 제공
  - ✓ name : CARD
- ✓ 사용 데이터 - ②
  - ✓ 대한민국 코로나19 일별 확진자
  - ✓ data.europa.eu 제공
  - ✓ name : COVID

이용일자	업종대분류	카드결제건수(천건)
20190101	가전/가구	149
20190101	가정생활/서비스	189
20190101	교육/학원	142
20190101	미용	464
20190101	스포츠/문화/레저	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

dateRep	day	month	year	cases	deaths	countriesAndTerritories
2020-12-14	14	12	2020	718	7	South_Korea
2020-12-13	13	12	2020	1030	2	South_Korea
2020-12-12	12	12	2020	950	6	South_Korea
2020-12-11	11	12	2020	688	8	South_Korea

## 분석 절차

- ✓ 카드소비동향 데이터에 대해 1차 대유행이 있었던 20년 2월~3월과  
코로나19 이전 같은 기간인 19년 2월~3월의 소비동향 비교
- ✓ 19년 2월~3월을 '코로나 유행 이전', 20년 2월~3월을 '코로나 유행 이후'로 지정,  
각 업종별 평균카드결제건수가 얼마나 달라졌는지 확인
- ✓ 최대한 ORACLE SQL 내에서 데이터 전처리 수행

## QUERY

```
SELECT 업종대분류, PERIOD, ROUND(AVG(카드결제건수),2) AS 평균카드결제건수 FROM(
SELECT 업종대분류, 카드결제건수,
CASE
    WHEN DATEREPR < TO_DATE('19/03/31') THEN '코로나 유행 이전'
    ELSE '코로나 유행 이후' END AS PERIOD FROM CARD
WHERE DATEREPR BETWEEN TO_DATE('19/02/01') AND TO_DATE('19/03/31') OR
    DATEREPR BETWEEN TO_DATE('20/02/01') AND TO_DATE('20/03/31'))
GROUP BY PERIOD, 업종대분류
ORDER BY PERIOD, 업종대분류;
```

	업종대분류	PERIOD	평균카드결제건수
1	가전/가구	코로나 유행 이전	296.05
2	가정생활/서비스	코로나 유행 이전	535.66
3	교육/학원	코로나 유행 이전	391.62
4	미용	코로나 유행 이전	565.24
5	스포츠/문화/레저	코로나 유행 이전	1602.69
6	여행/교통	코로나 유행 이전	209.9
7	요식/유흥	코로나 유행 이전	10746.86
8	유통	코로나 유행 이전	12983
9	음/식료품	코로나 유행 이전	1416.72
10	의료	코로나 유행 이전	2352.28
11	자동차	코로나 유행 이전	370.19
12	주유	코로나 유행 이전	2067.48
13	패션/잡화	코로나 유행 이전	440
14	가전/가구	코로나 유행 이후	252.84
15	가정생활/서비스	코로나 유행 이후	525.52
16	교육/학원	코로나 유행 이후	274.3
17	미용	코로나 유행 이후	400.62
18	스포츠/문화/레저	코로나 유행 이후	1211.98

## 코로나 전후 요약통계량 변화

✓ 위는 발생 전

아래는 발생 후

✓ 대체로 모든 업종에 대한

평균 값이 감소하는 성향

✓ 데이터 시각화를 통해

보다 자세히 알아보고자 함

### 〈코로나 대유행 전〉

	가전_가구	가정생활_서비스	교육_학원	미용	스포츠_문화_레저	여행_교통	요식_유흥	유통	음_식료품	의료	자동차	주유	패션_잡화
count	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00	59.00
mean	294.54	530.44	388.81	566.71	1607.64	210.24	10747.69	12989.95	1413.75	2320.88	368.08	2066.59	441.36
std	166.32	286.27	167.69	137.43	339.50	39.92	1237.13	1103.66	299.20	1059.48	84.54	168.90	110.79
min	84.00	85.00	37.00	176.00	1171.00	161.00	5332.00	8553.00	788.00	238.00	117.00	1659.00	162.00
25%	224.50	426.00	283.50	473.50	1309.00	178.00	10360.50	12349.00	1253.00	2171.00	374.00	2012.00	370.50
50%	260.00	504.00	403.00	540.00	1476.00	200.00	10876.00	13131.00	1399.00	2677.00	401.00	2096.00	400.00
75%	276.00	566.50	445.00	654.50	1928.50	233.00	11415.00	13642.50	1529.50	2953.00	415.00	2147.00	526.50
max	1020.00	1536.00	1010.00	881.00	2348.00	315.00	12732.00	15207.00	2398.00	3764.00	453.00	2477.00	714.00

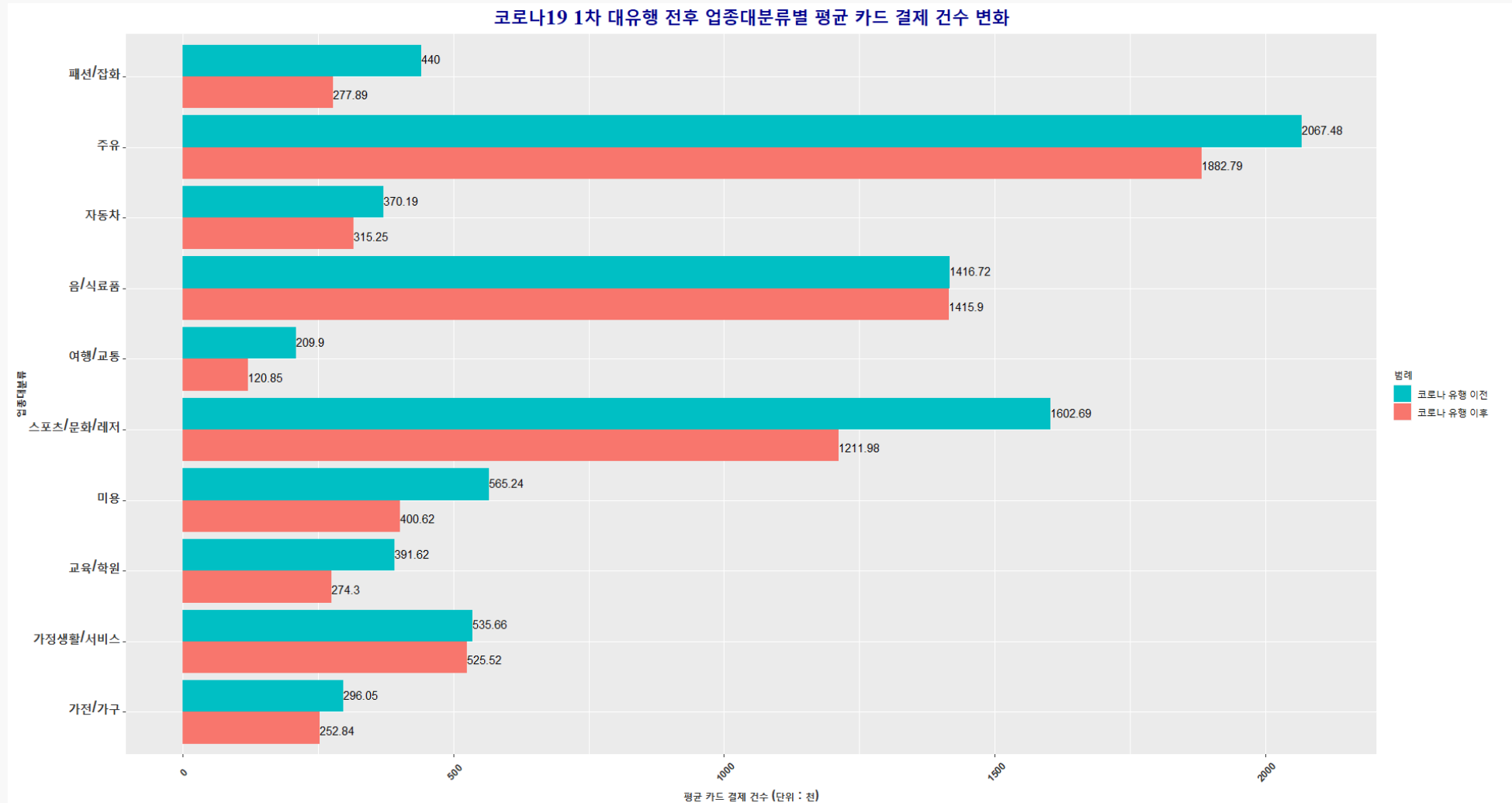
### 〈코로나 대유행 후〉

	가전_가구	가정생활_서비스	교육_학원	미용	스포츠_문화_레저	여행_교통	요식_유흥	유통	음_식료품	의료	자동차	주유	패션_잡화
count	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00	43.00
mean	299.42	569.98	433.67	563.35	1594.47	204.93	11113.81	13279.47	1429.53	2455.26	386.77	2087.33	449.00
std	148.19	280.84	157.66	129.12	350.65	39.87	777.14	795.55	127.84	952.10	60.52	135.20	112.38
min	201.00	221.00	202.00	402.00	1171.00	161.00	9645.00	11422.00	1195.00	370.00	234.00	1694.00	302.00
25%	233.00	480.00	339.00	465.50	1283.00	170.50	10510.00	12715.00	1346.50	2386.00	388.50	2027.50	373.00
50%	266.00	520.00	419.00	512.00	1476.00	194.00	11015.00	13353.00	1437.00	2692.00	406.00	2106.00	400.00
75%	277.00	576.00	489.00	654.50	1946.00	228.50	11766.50	13819.00	1529.50	2953.00	425.00	2164.00	536.50
max	1020.00	1536.00	1010.00	842.00	2348.00	315.00	12732.00	15207.00	1660.00	3764.00	447.00	2338.00	714.00



## 02 분석 - ①

### 1차 대유행 소비 변화

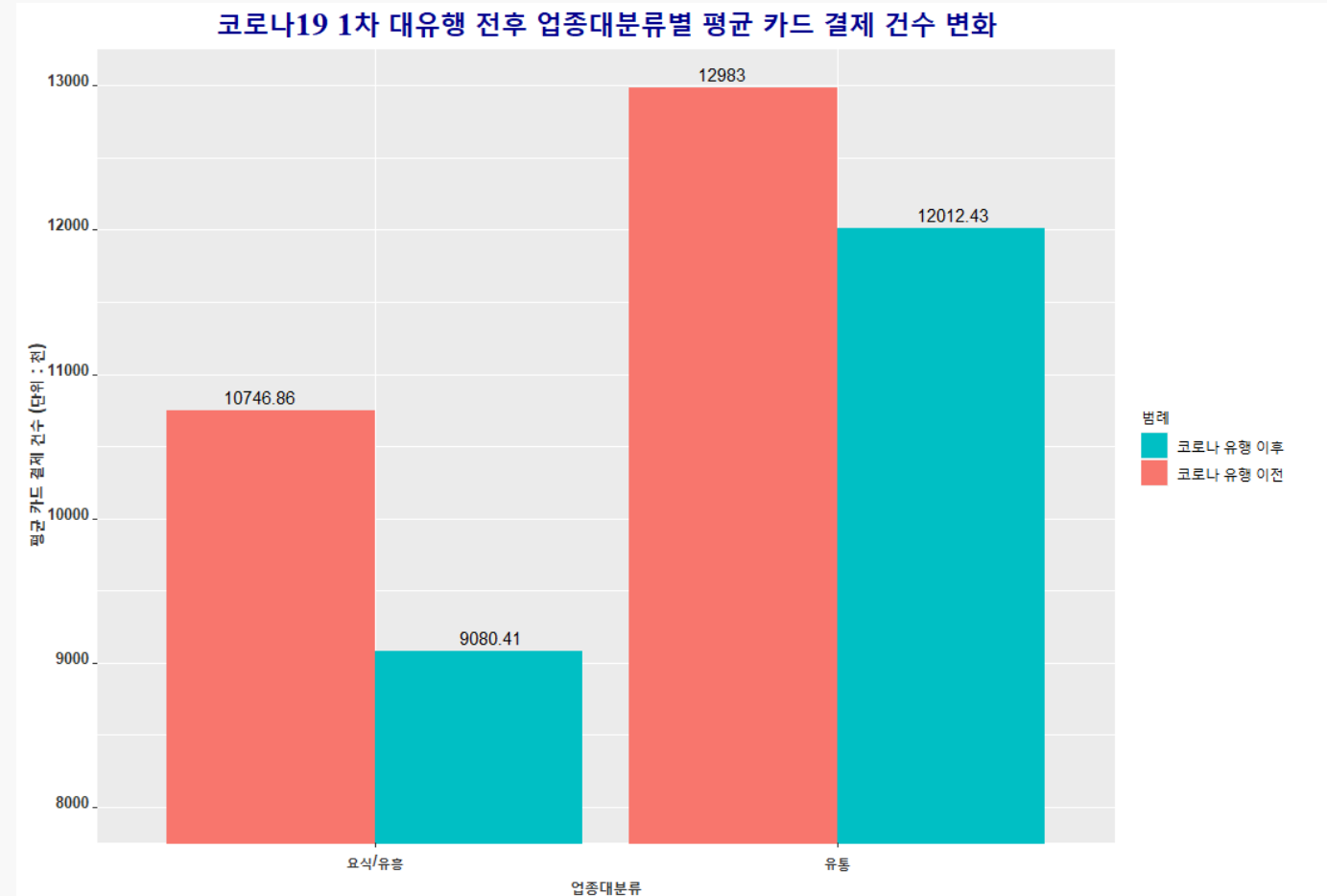


## 분석 결과

- ✓ 1차 대유행의 여파로 거의 모든 업종의 소비 감소  
(패션/잡화, 주유, 자동차, 음/식료품, 여행/교통, 스포츠/문화/레저, 미용, 교육/학원, 가정생활/서비스, 가전/가구)
- ✓ 스포츠/문화/레저, 패션/잡화는 매우 큰 폭으로 카드 결제건수가 감소한 것을 알 수 있음  
(각 39만, 16만의 데이터 카드 결제건수 감소)
- ✓ 여행/교통은 거의 절반 수준으로 감소  
(21만에서 12만 건으로 감소)
- ✓ 음/식료품과 가정생활/서비스는 거의 감소하지 않음  
(각 1천 건, 1만 건 감소)

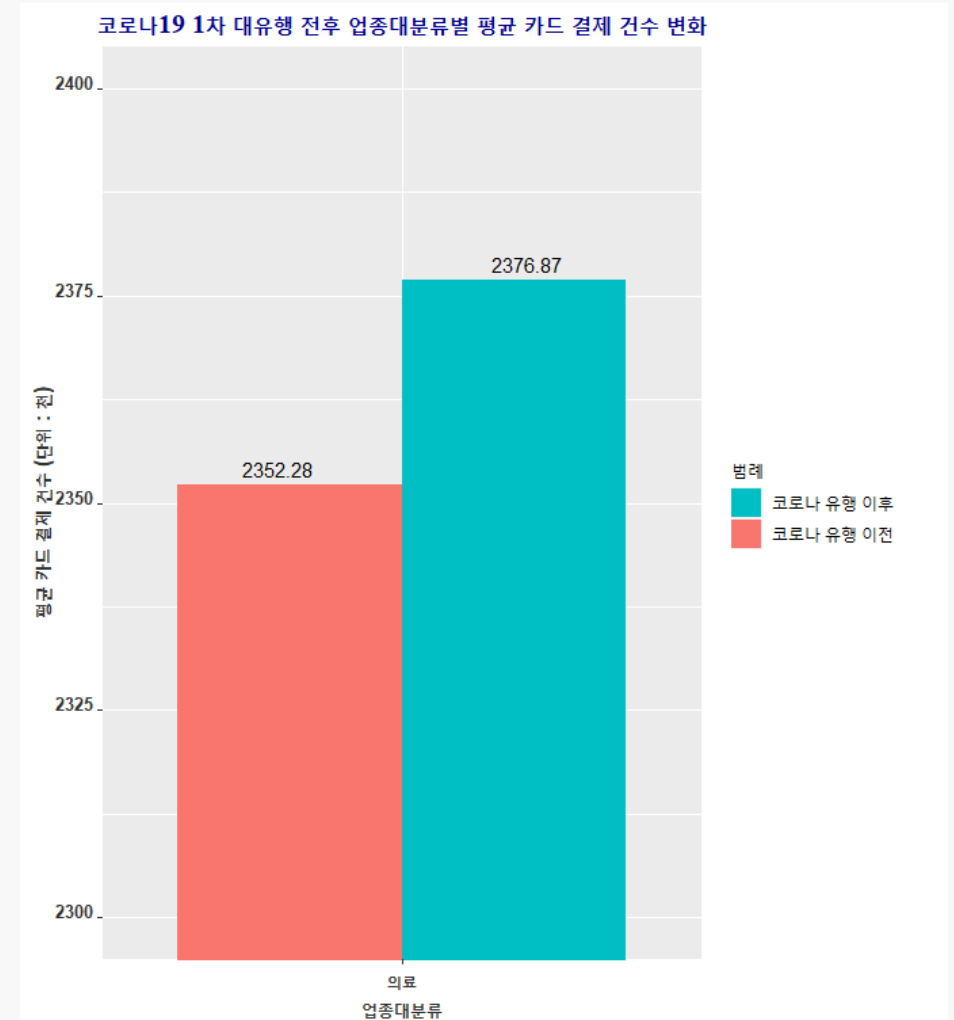
## 분석 결과

- ✓ 카드 결제 건수의 대부분을 차지하는  
요식/유흥, 유통분야도 감소
- ✓ 두 분야 모두 약 100만 건 감소



## 분석 결과

- ✓ 유일하게 의료분야만  
카드 결제 건수 증가하였음
- ✓ 코로나 유행 이전 대비 약 25만 건 증가
- ✓ 의료분야를 제외하고  
거의 모든 업종이 타격을 받았다고 할 수 있음



## 분석 절차

- ✓ 일별 확진자의 변화가 업종간 카드 결제 건수에 얼마나 영향을 주는지 알아보고자 함
- ✓ 확진자의 수와 13개 대분류 결제건수 간 상관분석 진행
- ✓ 상관계수 추출, 검정 및 해석
- ✓ 최대한 ORACLE SQL 내에서 데이터 전처리 시행

## QUERY

---

```
CREATE TABLE CARD_PIVOT AS (SELECT *  
FROM (  
    SELECT * FROM CARD  
)  
PIVOT (  
    SUM(카드결제건수) FOR 업종대분류 IN ('가전/가구' AS 가전_가구, '가정생활/서비스' AS 가정생활_서비스, '교육/학원' AS 교육_학원,  
        '미용' AS 미용, '스포츠/문화/레저' AS 스포츠_문화_레저, '여행/교통' AS 여행_교통, '요식/유흥' AS 요식_유흥,  
        '유통' AS 유통, '음/식료품' AS 음_식료품, '의료' AS 의료, '자동차' AS 자동차, '주유' AS 주유, '패션/잡화' AS 패션_잡화)  
));  
  
SELECT * FROM CARD_PIVOT;
```

-- 피벗테이블을 통해 대분류가 열이 될수 있도록 함

# 03 분석 - ②

일별 확진자와 업종간 상관분석

## QUERY

	DATEREP	가전_가구	가정생활_서비스	교육_학원	미용	스포츠_문화_레저	여행_교통	요식_유흥	유통	음_식료품	의료	자동차	주유	패션_잡화
1	19/02/12	223	462	370	468	1264	184	10153	11865	1191	2831	395	2081	358
2	19/02/14	226	446	329	544	1368	205	10796	12816	1256	2678	390	2114	382
3	19/02/22	241	558	401	558	1543	209	11595	12738	1382	2867	434	2244	417
4	19/02/25	292	1518	433	512	1420	207	10305	12865	1316	3386	430	2106	389
5	19/02/27	235	533	437	540	1504	204	11126	12692	1339	2773	425	2083	407
6	19/03/21	264	504	418	465	1257	177	11129	13085	1481	2846	405	2093	378
7	19/03/30	254	489	301	772	2072	260	12310	14342	1606	2415	406	2183	610
8	19/01/13	192	214	161	645	1844	278	10051	11356	1257	398	223	1769	474
9	19/01/14	229	481	439	457	1193	196	9478	12205	1302	3459	381	2034	330
10	19/01/18	239	454	412	551	1379	215	11383	12829	1472	2868	401	2243	404
11	19/01/21	597	795	511	465	1208	202	9667	12068	1360	3368	383	2074	333
12	19/01/24	236	450	337	490	1294	204	10502	12403	1378	2657	384	2085	361
13	19/06/19	279	541	497	453	1317	185	11263	14189	1565	2731	396	2157	434
14	19/07/06	260	473	259	733	2418	294	13118	16332	1774	2179	450	2403	731
15	19/07/08	333	533	374	431	1396	190	10509	14168	1514	3542	418	2213	418
16	19/07/15	610	919	571	427	1320	198	10226	13301	1442	3423	395	2176	387
17	19/07/28	200	217	162	506	2001	291	10748	12491	1312	388	230	1903	514
18	19/07/29	283	570	364	484	1686	256	11502	15284	1557	3374	464	2432	477

## QUERY

```
CREATE TABLE CARD_COVID AS (
SELECT C.DATEREPI, TO_NUMBER(SUBSTR(C.DATEREPI, 1,2)) AS YEAR, TO_NUMBER(SUBSTR(C.DATEREPI, 4,2)) AS MONTH,
TO_NUMBER(SUBSTR(C.DATEREPI, 7,2)) AS DAY,
V.CASES, C.가전_가구, C.가정생활_서비스, C.교육_학원, C.미용, C.스포츠_문화_레저, C.여행_교통, C.요식_유흥, C.유통, C.음_식료품, C.의료, C.자동차, C.주유, C.패션_잡화
FROM CARD_PIVOT C, COVID V
WHERE C.DATEREPI = V.DATEREPI(+));
```

```
SELECT * FROM CARD_COVID WHERE CASES IS NOT NULL ORDER BY DATEREPI;
```

-- RIGHT JOIN을 통해 COVID 데이터에 존재하는 날짜를 중심으로 서로 공통이 되는 데이터를 추출

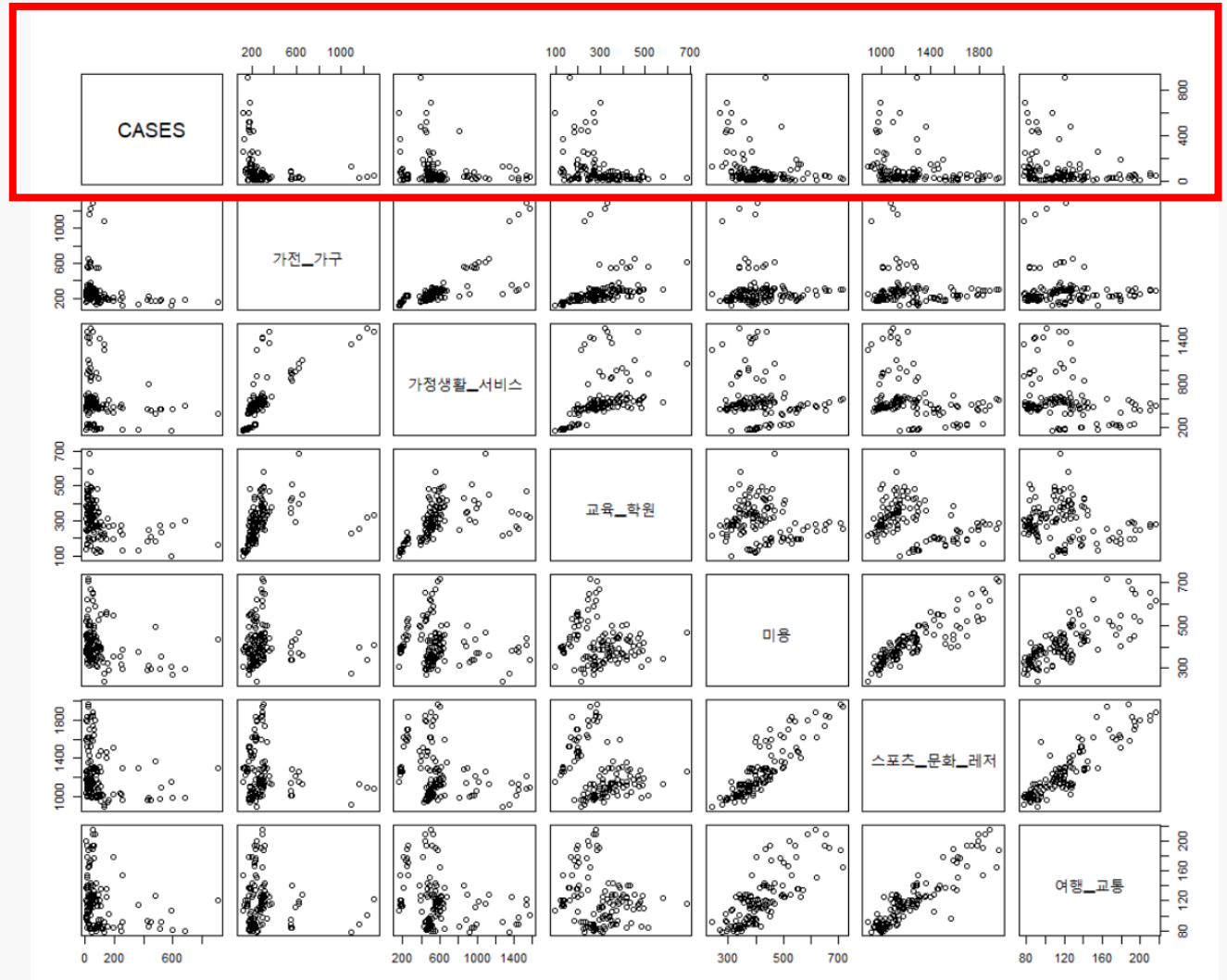
-- 일별 확진자 수와 업종간 결제건수가 어떤 영향을 가지는지 분석

DATEPI	YEAR	MONTH	DAY	CASES	가전_가구	가정생활_서비스	교육_학원	미용	스포츠_문화_레저	여행_교통	요식_유흥	유통	음_식료품	의료	자동차	주유	패션_잡화
1 20/01/01	20	1	1	0	158	197	145	464	1969	255	9695	11911	1291	494	214	1823	406
2 20/01/02	20	1	2	0	250	541	537	523	1581	200	10229	12348	1356	3778	435	2193	368
3 20/01/03	20	1	3	0	256	519	505	544	1640	201	11580	12784	1504	3539	445	2306	383
4 20/01/04	20	1	4	0	247	471	312	748	2166	265	11826	13537	1632	2559	426	2191	547
5 20/01/05	20	1	5	0	194	236	201	606	1920	241	10053	12292	1198	534	243	1834	436
6 20/01/06	20	1	6	0	288	898	504	418	1305	185	9731	12143	1308	3969	399	2034	282
7 20/01/07	20	1	7	0	228	443	388	381	1236	176	9796	11247	1214	3115	381	1913	260
8 20/01/08	20	1	8	0	310	523	395	470	1362	184	10770	12639	1443	3070	412	2168	334
9 20/01/09	20	1	9	0	246	496	352	489	1367	194	10972	12469	1441	3080	425	2242	356
10 20/01/10	20	1	10	0	1153	1447	348	551	1520	205	11889	13151	1552	3279	440	2342	397



## 분석 과정

- ✓ CASES에 대해 13개 중 6개 분류를 우선 알아본 결과, 대체로 감소하는 성향을 띄고 있음
- ✓ 특히, 여행/교통, 스포츠문화/레저는 대부분의 데이터의 수치가 낮은 값에 분포되어 있음



## 분석 과정

- ✓ CASES에 대해 13개 중 7개 분류를 알아본 결과, 역시 대체로 감소하는 성향을 띄고 있음
- ✓ 그렇지만, 대부분의 데이터가 일별 확진자 수가 매우 적을때에 몰려있으므로, 확진자 수가 6명 이상일 때로 데이터 축소



## QUERY

```
SELECT * FROM CARD_COVID WHERE CASES > 5 ORDER BY DATEREP;
```

-- 일별 확진자 수가 6명 이상일 때의 데이터만 추출

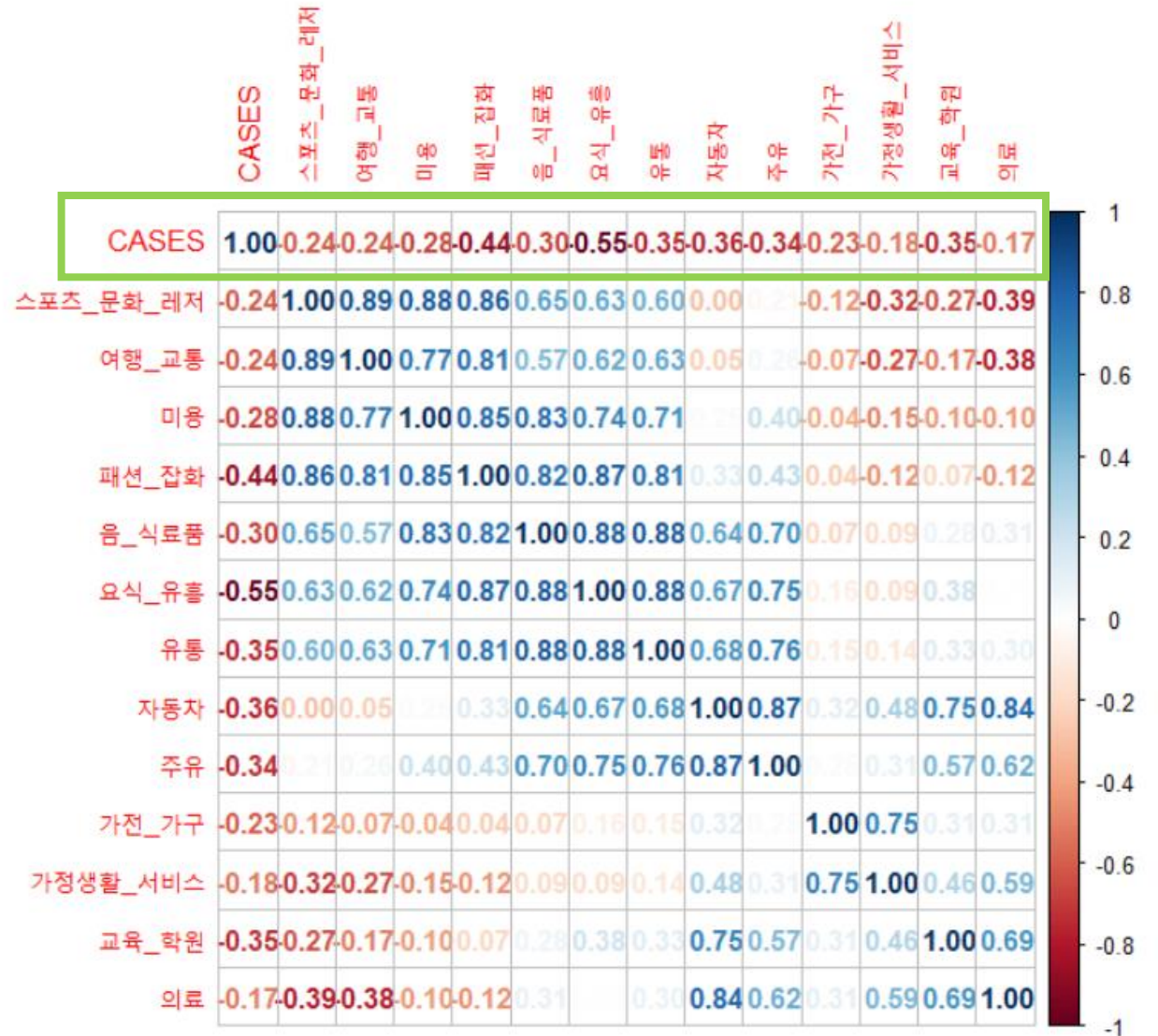
## 분석 결과

- ✓ CASES (확진자)에 대해 의료를 제외한 모든 대분류의  $p\text{-value} < 0.05$
- ✓ 상관관계는 의료를 제외하고 유의 (0.0576은 0.05에 상당히 근접하므로 어느 정도 참고하는 것이 가능)
- ✓ 유의하다는 것을 확인했으므로, 상관계수를 시각화하고자 함

CASES	NA
가전_가구	0.0096
가정생활_서비스	0.0417
교육_학원	0.0001
미용	0.0013
스포츠_문화_레저	0.0076
여행_교통	0.0065
요식_유흥	0.0000
유통	0.0000
음_식료품	0.0007
의료	0.0576
자동차	0.0000
주유	0.0001
패션_잡화	0.0000

## 분석 결과

- ✓ CASES(확진자)에 대한 상관계수는 상관행렬의 첫 줄에 표시되어 있음
- ✓ 요식/유통, 패션/잡화 순으로 상관계수 절댓값이 큼
- ✓ 상관계수의 절댓값이 가장 작은 업종은 의료, 가정생활/서비스 업종



## 분석 결과

- ✓ 대부분의 업종이 모두 큰 타격을 받았음  
(모든 업종이 음의 상관관계를 가짐)
- ✓ 그 중 요식/유흥 업종은 확진자 수에 따라 가장 큰 감소 추세를 보임 (강한 음의 상관관계)  
(-0.55로 가장 강한 음의 상관계수를 가짐)
- ✓ 패션/잡화 업종 역시 요식/유흥의 뒤를 이어 확진자 수에 따라 큰 감소 추세를 가짐  
(-0.44로 요식/유흥에 이어 2번째)
- ✓ 의료, 가정생활/서비스는 확진자 수에 따라 약간의 감소 추세가 있지만,  
타 업종에 다소 미미한 감소 추세를 보임

## 분석 한계

- ✓ 카드 결제 데이터에 대해 카드 결제 금액이 아닌 카드 결제 건수 데이터를 이용한 점  
(한 건당 소비금액은 늘어났을 수도 있기 때문에, 총 소비량에 대한 변화는 알 수 없음)
- ✓ 코로나 대유행 이후 최신 데이터가 아닌 20년 1월 ~ 6월 데이터를 사용한 점  
(21년 5월 현재도 코로나는 진행형이며, 최근 소비 동향에 대해 알 수 없음)
- ✓ 1차 대유행은 대부분 대구/경북지역에서 발생했지만, 카드 결제 건수 데이터는 전국 대상  
(당시 소비에 대한 지역구별 소비 변화를 정확하게 반영하지 못함)

## 개선 방향

- ✓ 카드 결제 금액 데이터를 이용해 총 소비량에 대해 분석하여  
정확한 인사이트를 도출할 수 있도록 함
- ✓ 최신 코로나 현황을 반영한 카드 결제 데이터 사용  
(20년 1월 ~ 21년 현재)
- ✓ 코로나는 1차 대유행 이후로 전국으로 확산되었으므로,  
소비 결제 데이터에 대한 지역성은 크게 없어졌을 것으로 추정