

# 카드 소비 패턴 분석 시각화

정보통계종합설계\_FLEX



20190699 김지은 | 20190702 남수빈 | 20190704 박가온

# Contents

정보통계종합설계

## 01. 개요

- 주제 선정 이유
- 데이터 설명
- 데이터 전처리

## 02. 성별/연령별

- 성별 카드 소비 패턴
- 성별/연령별 카드 소비 패턴

## 03. 시간대별

- 카드이용금액
- 소액결제금액
- 소액 결제 top 5상품

## 04. 연도/월별

- 연도별 월 이용 건수 금액 추이

## 05. 업종별

- 카드이용금액
- 소액결제건수
- 성별 카드 소비 패턴

## 06. 지역별

- 카드 이용 건수
- 카드 이용 금액

# 01. 개요

## 주제 선정 이유 및 목적



카드 소비 패턴 분석을 통해  
기업 마케팅에 중요한 지표를 제시



카드 소비에 대하여 성별 및 연령, 지역, 시간대와 같이 다양한  
관점에서 비교하고 분석하여 새로운 인사이트를 도출

## 데이터 설명



국내 전 지역 가맹점을 대상으로 서울시민의 개인카드  
(신용카드와 체크카드 포함, 법인카드 제외) 사용  
패턴을 집계한 데이터



출처 : 서울특별시 빅데이터 캠퍼스  
엑세스 날짜 : 2022.09.19

# 01. 개요

## 데이터 불러오기

data1 : 블록별 성별연령대별 카드소비패턴

data2 : 블록별 시간대별소액결제건수 카드소비패턴

data3 : 집계구별 일별소비지역별 카드소비패턴

data4 : 집계구별 일별시간대별 카드소비패턴

data5 : 카드소비 업종코드

## 변수 이름 변경

```
1 data1.rename(columns = {'서울시민업종코드(UPJONG_CD)' : 'upjong_cd',
2                          '기준년월(YM)' : 'ym',
3                          '고객주소블록코드(BLOCK_CD)' : 'block_cd',
4                          '성별(GEDNER)' : 'gender',
5                          '연령대별(AGE)' : 'age',
6                          '카드이용금액계(AMT_CORR)' : 'amt_corr',
7                          '카드이용건수계(USECT_CORR)' : 'usect_corr'
8                          }, inplace = True)
9
10 data2.rename(columns = {'서울시민업종코드(UPJONG_CD)' : 'upjong_cd',
11                          '기준년월(YM)' : 'ym',
12                          '시간대구간(TIME)' : 'time',
13                          '고객주소블록코드(BLOCK_CD)' : 'block_cd',
14                          '카드이용금액계(AMT_CORR)' : 'amt_corr',
15                          '소액결제건수(MICRO_PYM)' : 'micro_pym'
16                          }, inplace = True)
17
18 data3.rename(columns={'가맹점주소광역시도(SIDO)' : 'sido',
19                      '가맹점주소시군구(SGG)' : 'sgg',
20                      '업종대분류(UPJONG_CLASS1)' : 'upjong_class1',
21                      '기준일자(YMD)' : 'ymd',
22                      '고객주소집계구별(TOT_REG_CD)' : 'tot_reg_cd',
23                      '카드이용금액계(AMT_CORR)' : 'amt_corr',
24                      '카드이용건수계(USECT_CORR)' : 'usect_corr'}, inplace = True)
25
26 data4.rename(columns = {'업종대분류(UPJONG_CLASS1)' : 'upjong_class1',
27                          '기준일자(YMD)' : 'ymd',
28                          '시간대구간(TIME)' : 'time',
29                          '고객주소집계구별(TOT_REG_CD)' : 'tot_reg_cd',
30                          '카드이용금액계(AMT_CORR)' : 'amt_corr',
31                          '카드이용건수계(USECT_CORR)' : 'usect_corr'
32                          }, inplace = True)
33
34 data5.rename(columns={'업종코드(UPJONG_CD)' : 'upjong_cd',
35                      '대분류(CLASS1)' : 'class1',
36                      '중분류(CLASS2)' : 'class2',
37                      '소분류(CLASS3)' : 'class3'}, inplace = True)
```

# 01. 개요

## 데이터 전처리

### 필요없는 변수 제거

```
1 # df의 block_cd의 의미를 파악해보려했습니다. 고유코드로 이용이 될까 싶어서
2 # 파헤쳐보니 같은 block_cd 를 가진 사람이 있었습니다.
3 # 즉 같은 고유코드를 가진 사람은 같은 주소의 사람 (가족) 으로 추측하였습니다.
4 # 카드소비내역을 설명해주기엔 다소 모호한 부분이 있어서 제외하였습니다.
5 temp = pd.DataFrame(data1.groupby('block_cd')['gender'].value_counts())
6 temp.rename(columns = {'gender' : 'count'}, inplace=True)
7 display(temp.query('count == 2'))
8 display(data1.query("block_cd == 152747"))
```

		count
block_cd	gender	
152747	M	2

upjong_cd	ym	block_cd	gender	age	amt_corr	usect_corr	
45	SS016	201911	152747	M	20대	13478242	5
251	SS008	201603	152747	M	30대	608831	10

## 데이터 merge

```
1 data5['upjong_cd']=data5['upjong_cd'].str.upper()
2 data15=pd.merge(data1, data5, left_on='upjong_cd',
3                 right_on='upjong_cd',how='inner')
```

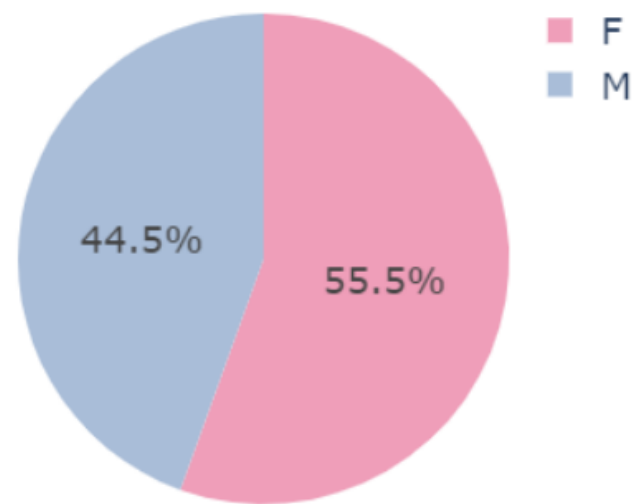
	upjong_cd	class1	class2	class3
0	SS001	요식/유흥	한식	한식
1	SS002	요식/유흥	일식/중식/양식	일식
2	SS003	요식/유흥	일식/중식/양식	양식
3	SS004	요식/유흥	일식/중식/양식	중식
4	SS005	요식/유흥	제과/커피/패스트푸드	제과점
...	...	...	...	...
70	SS084	스포츠/문화/레저	스포츠/문화/레저	스키
71	SS090	교육/학원	학원	입시보습학원
72	SS091	교육/학원	학원	외국어학원
73	SS092	교육/학원	학원	예체능학원
74	SS093	교육/학원	학원	취미/전문학원

75 rows × 4 columns

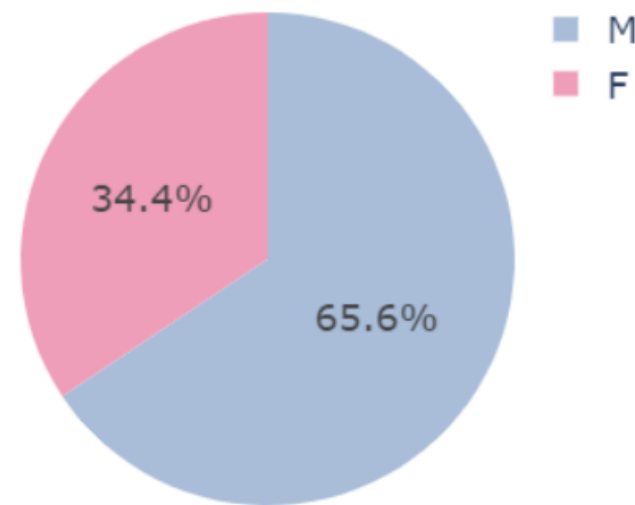
## 02. 성별/연령별

여자가 남자에 비해 카드 이용 건수는 많지만, 카드 이용 금액은 남자가 약 2배 높은 것으로 나타남  
생애 주기 별로 카드 사용 패턴 추측가능

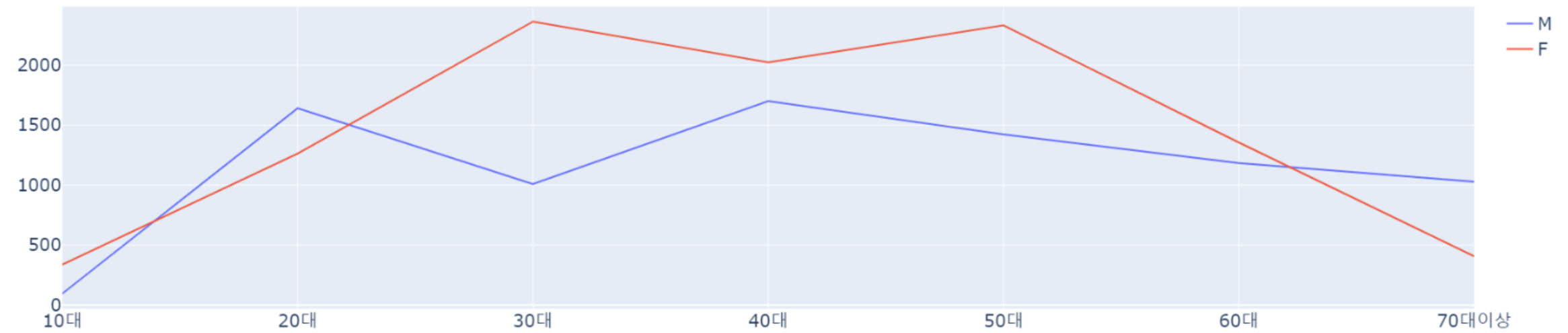
성별 카드이용건수



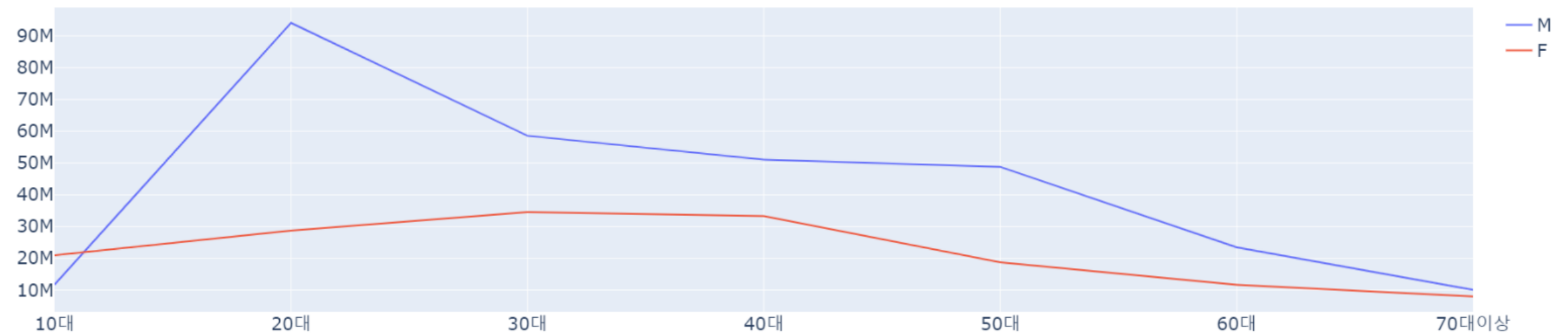
성별 카드이용금액



성별 연령별 카드이용건수



성별 연령별 카드이용금액계



## 03. 시간대별

시간대 : 1 (0~3), 2 (4~7), 3 (8~11), 4 (12~15), 5 (16~19), 6 (20~23)  
서울은 보통 8~15시까지 카드 이용, 소액결제 활발, 그 중에서 요식/유흥 소비가 많음

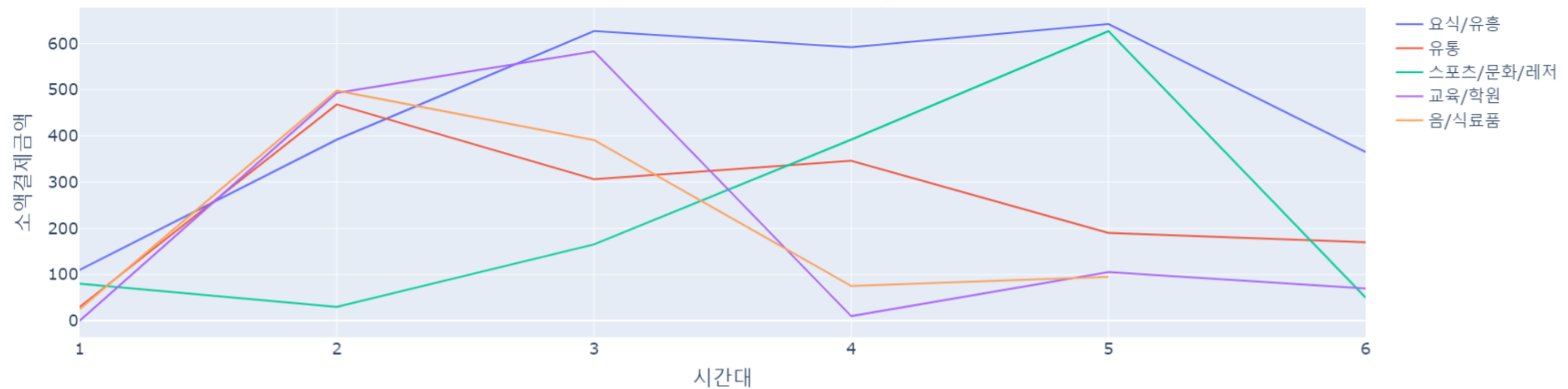
시간대별 카드이용금액



시간대별 소액결제금액



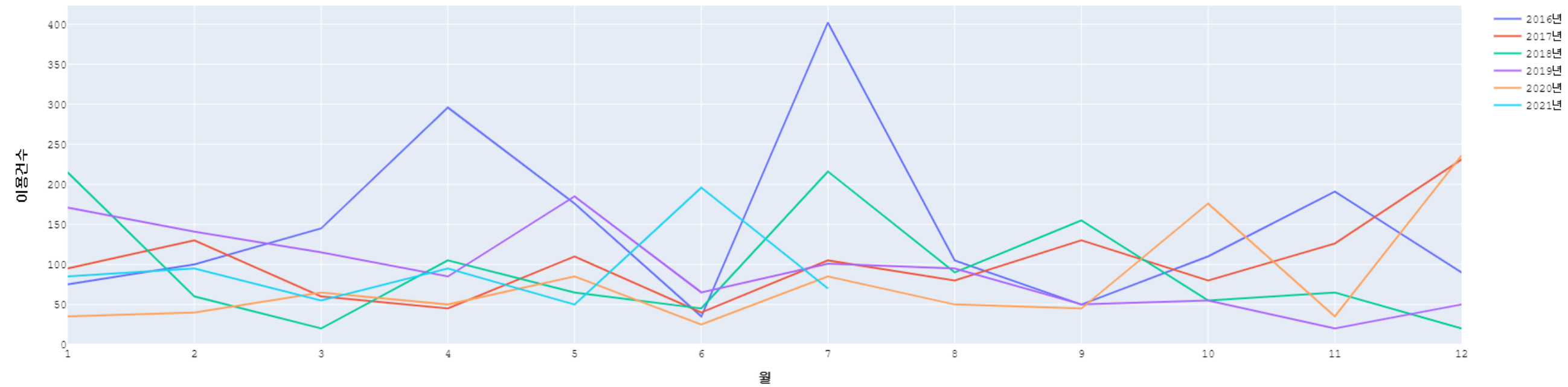
시간대별 소액결제금액 top5 상품



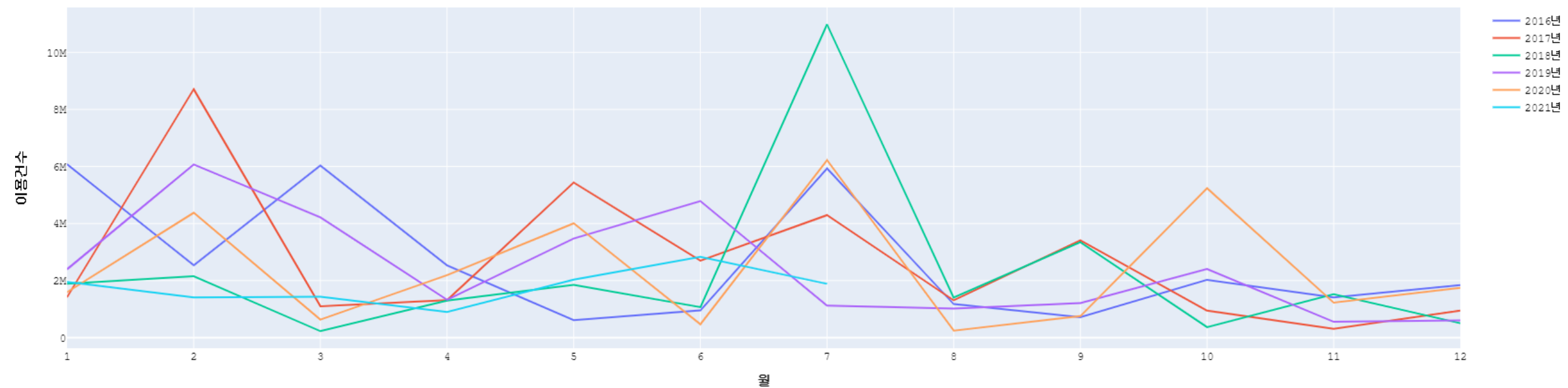
## 04. 연도 월별

연도별 월 이용건수와 이용 금액 추이를 살펴본 결과 7월에 소비가 많지만 추세가 크게 나타나지 않음

연도별 월 이용건수 추이



연도별 월 이용금액 추이

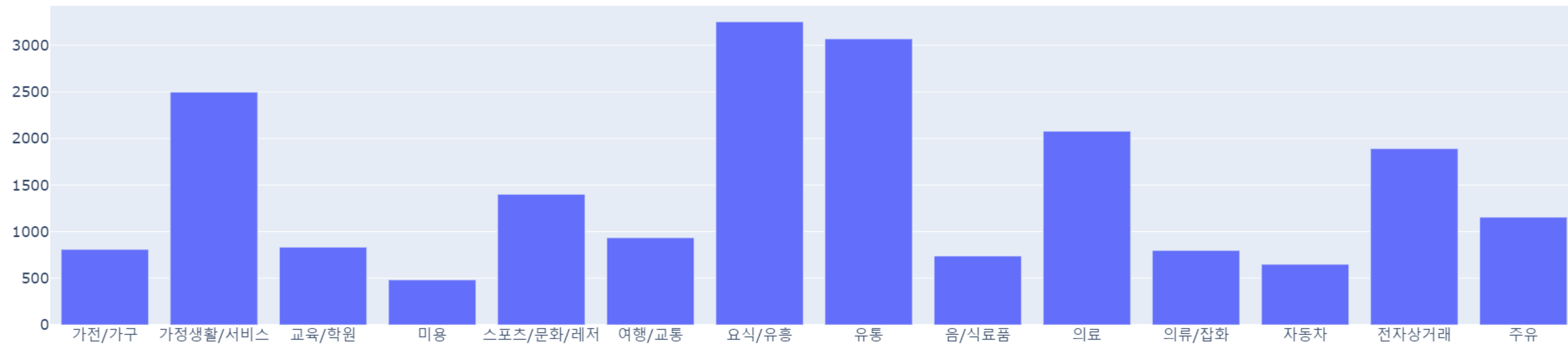




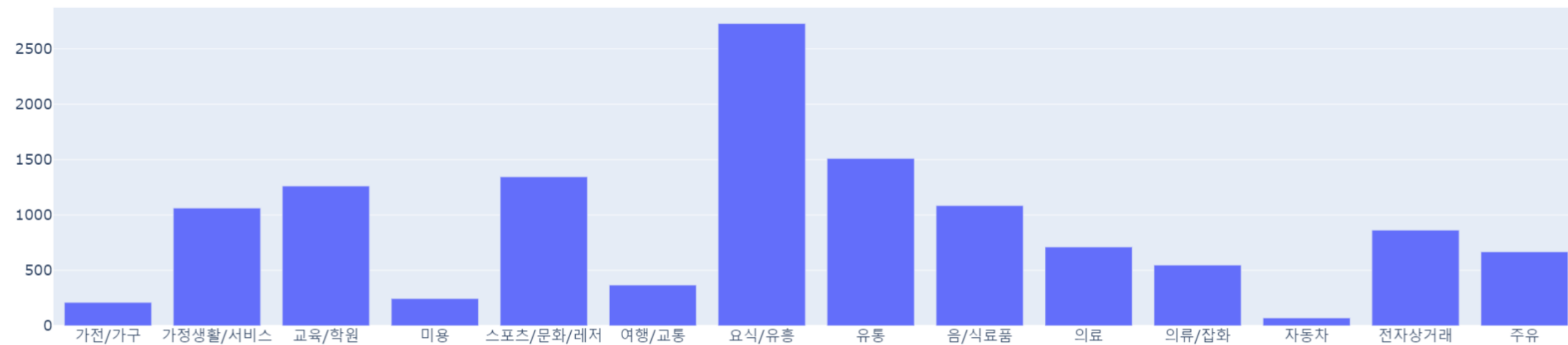
## 05. 업종별

업종별 카드이용금액과 소액결제건수를 살펴본 결과 카드 이벤트 시행시 추천 타겟 : 요식/유흥

업종별 카드이용금액



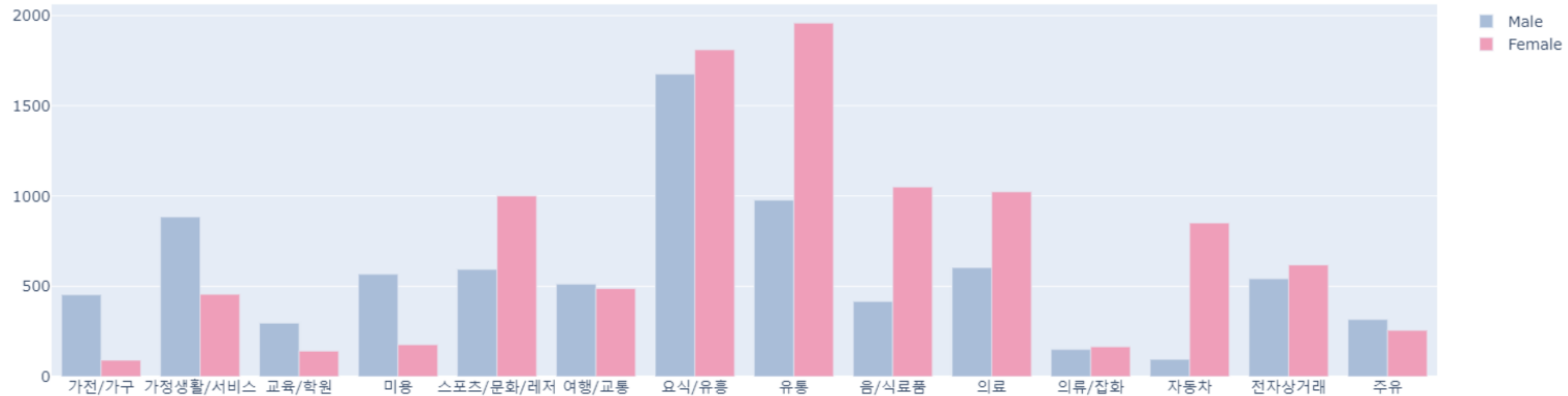
업종별 소액결제건수



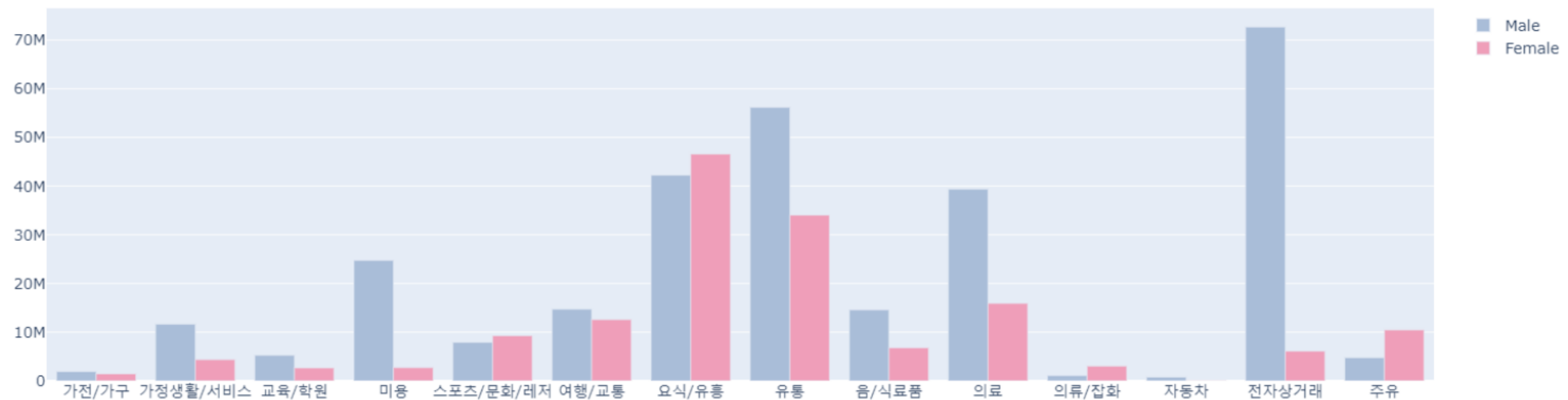
## 05. 업종별

업종별 카드이용건수와 금액을 살펴본 결과 카드 이벤트 시행시 여자 추천 타겟 : 요식/유흥  
업종별 카드이용건수와 금액을 살펴본 결과 카드 이벤트 시행시 남자 추천 타겟 : 전자상거래

업종별 성별 카드이용건수



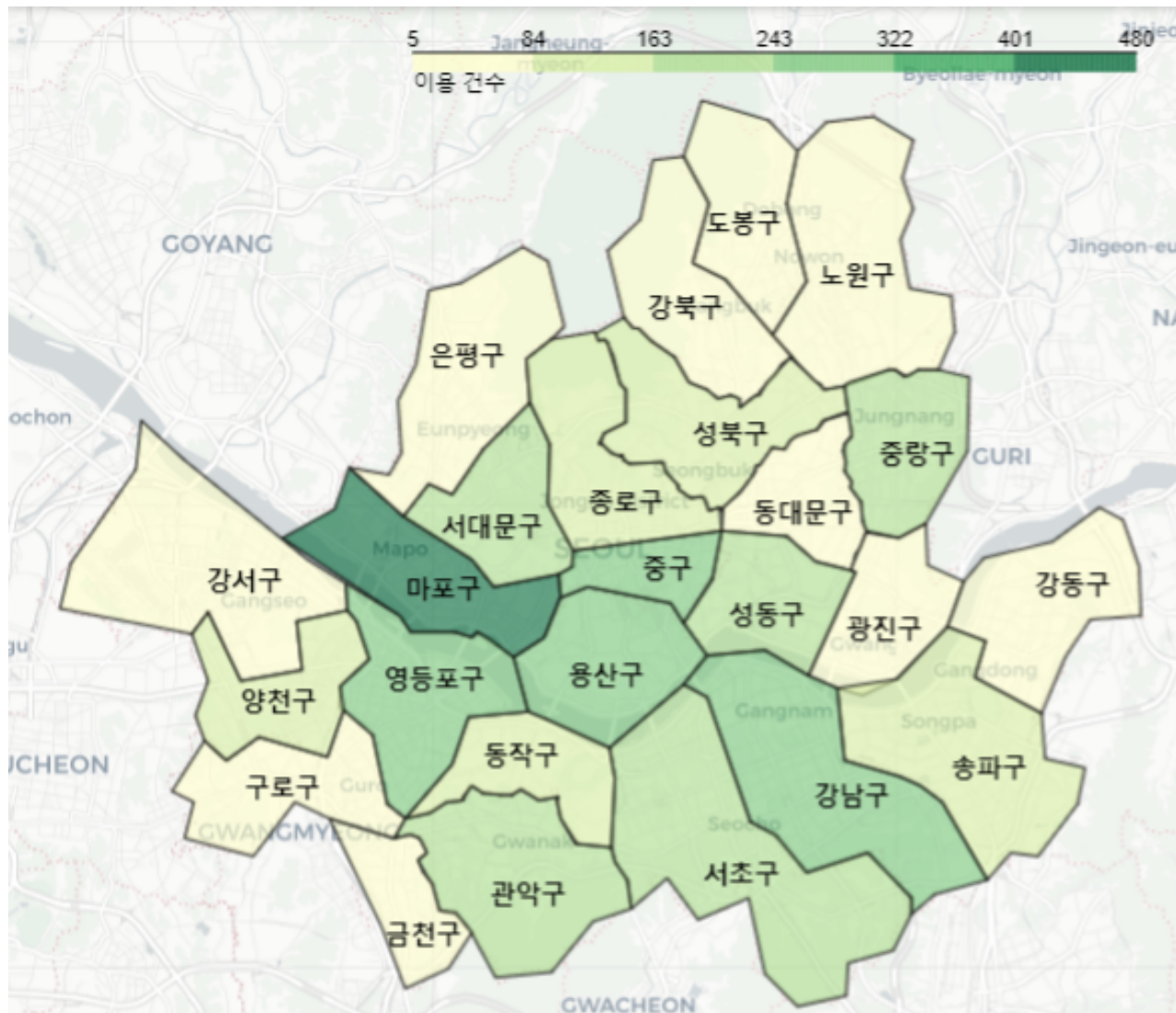
업종별 성별 카드이용금액



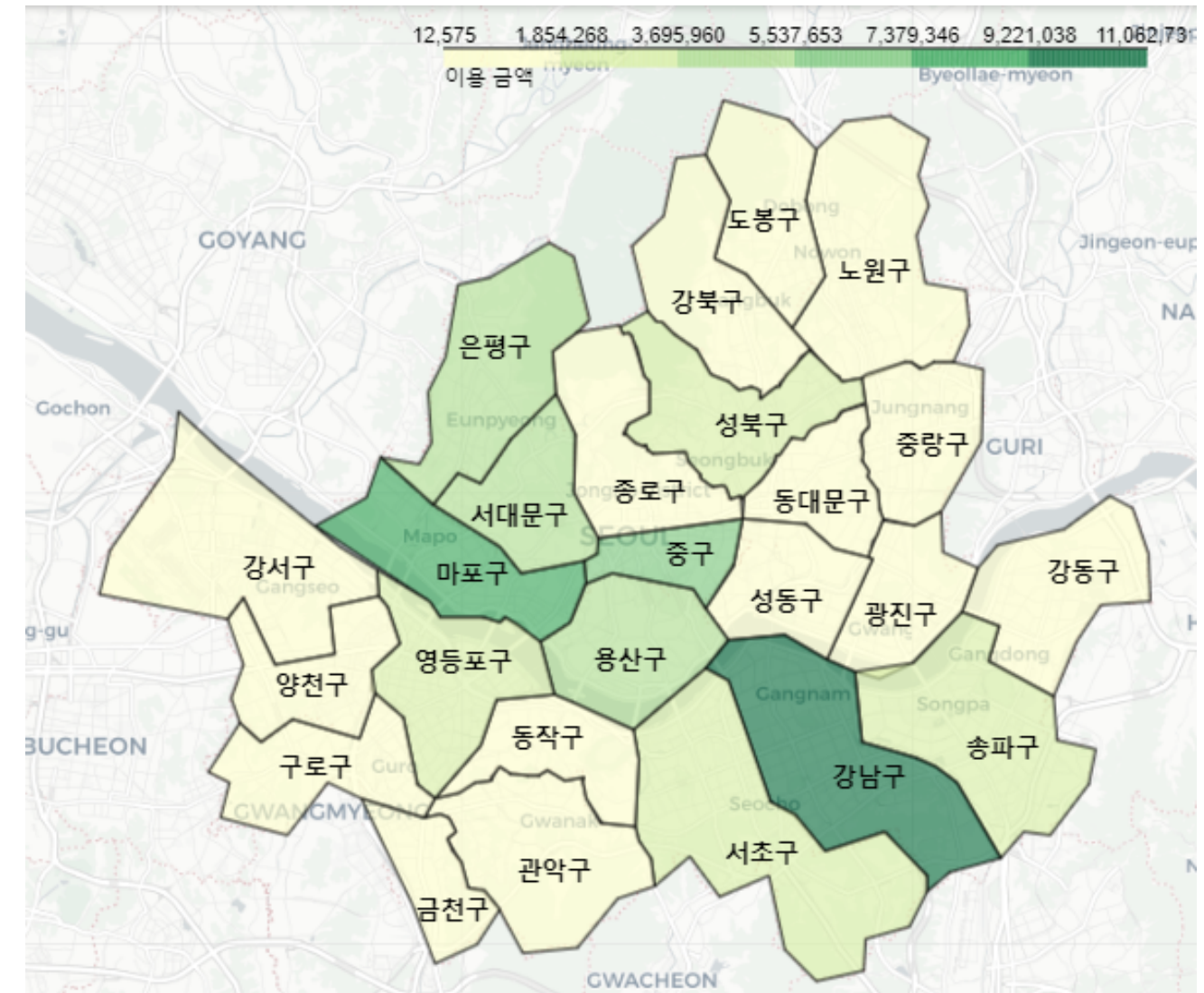
## 06. 지역별

서울 지역 내 카드 이용 건수가 많은 지역은 대학생들이 많이 분포해있는 마포구이고,  
금액이 높은 지역은 직장인(회사)들이 많이 분포하고 있는 강남구로 나옴

카드이용건수

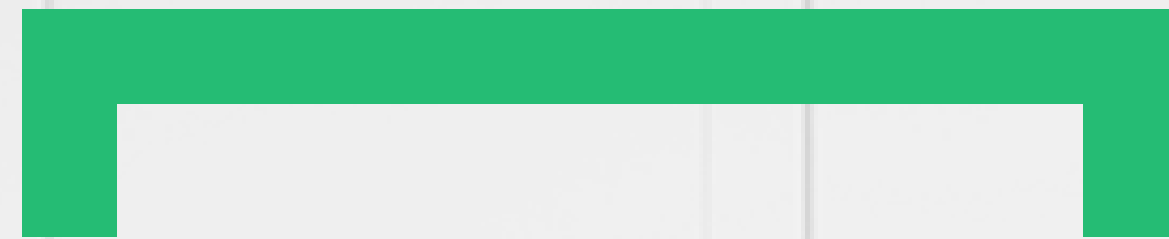


카드이용금액



# 결론

- ✔ 타겟 1 : 점심시간 직장인  
-> 결제 데이터나 위치 정보를 통해 맞춤형 혜택을 카드사 앱 푸시 알림 서비스 제공
- ✔ 타겟 2 : 20대 남성  
-> 눈에 띄는 전자상거래 이용 금액을 보유한 20대 남성에서 온라인 거래나 홈쇼핑 특화 상품 개발
- ✔ 타겟 3 : 30~50대 여성  
-> 소액 결제가 많은 여성을 대상으로 다양한 결합 혜택을 제공하고 지속적 노출 광고 마케팅 진행



# 감사합니다!

정보통계종합설계\_FLEX



20190699 김지은 | 20190702 남수빈 | 20190704 박가온