

팀명	프로젝트 명	수강생명	담당 역할	제출 파일(발표자료 및 코드_데이터 포함)
JYP	당신은 우리 JYP 팝업스토어에 오게 될 것입니다!	제태성	발표1	발표자료: team_jyp_miniprj.pdf 코드_데이터: local_data_cp949.csv team_jyp_miniprj.ipynb
		오성진	발표2	
		이예지	자료 수집1	
		이승현	자료 수집2	
		김민정	알고리즘1	
		권성재	알고리즘2	
		황나겸	알고리즘3	
		안윤빈	알고리즘4	

📢 당신은 우리 **JYP** 팝업스토어에 오게 될 것입니다



1. 프로젝트 목표 수립

1) 목표 수립하기

목표 : 18년도 제주도 내국인과 관광객 카드 이용 데이터(지역, 업종, 성별, 연령대별 기준)를 기반으로, 포스트 코로나를 맞이하여 언제, 어떤 업종으로 제주도에서 팝업스토어를 열어야 좋을지 파악

목표 선정 이유 :

제주도에서 이벤트성 팝업스토어를 론칭한다는 가정.

제주도는 코로나 시기에 대표 여행지로 부각되었던 만큼, 포스트 코로나를 맞이하고 있는 현재는 코로나 시기 이전과 유사한 소비패턴으로 회복할 것이 예상된다. 이러한 예상을 바탕으로, 2018년 실제 제주도에서 이루어졌던 카드 소비 패턴을 확인하여 이벤트성 팝업스토어 전략을 수립하고자 한다.

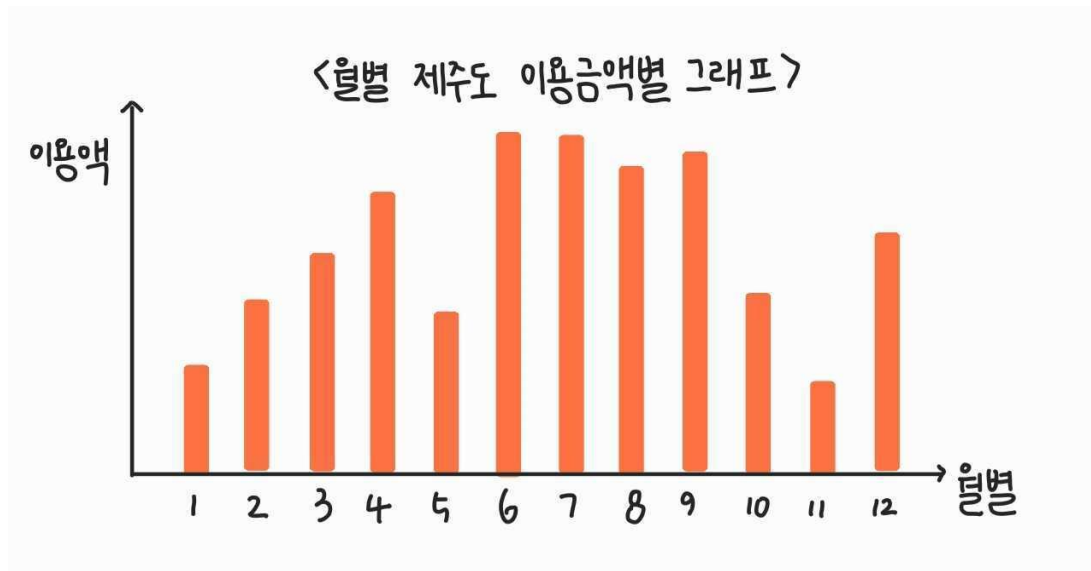
하위 목표

① 사람들이 언제 가장 소비를 많이 했을까?

- 월별 이용금액 비교

12개월 내국인 카드 사용금액을 한 눈에 보이도록 막대그래프로 시각화

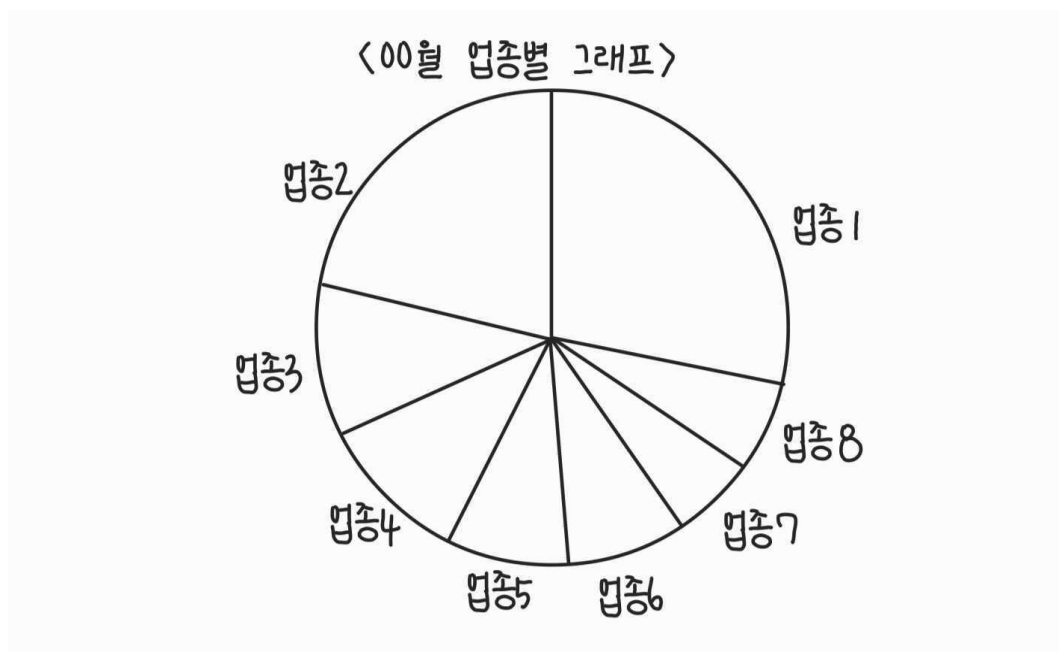
: x축 - 1월~12월 / y축 - 전체 이용금액



② 사람들은 어디서 가장 많은 돈을 썼을까?

- 업종(8개)별 이용금액

하위목표①에서 이용금액이 가장 높게 산출된 달을 기준으로, 8개 업종의 이용금액 비율을 파이그래프로 시각화

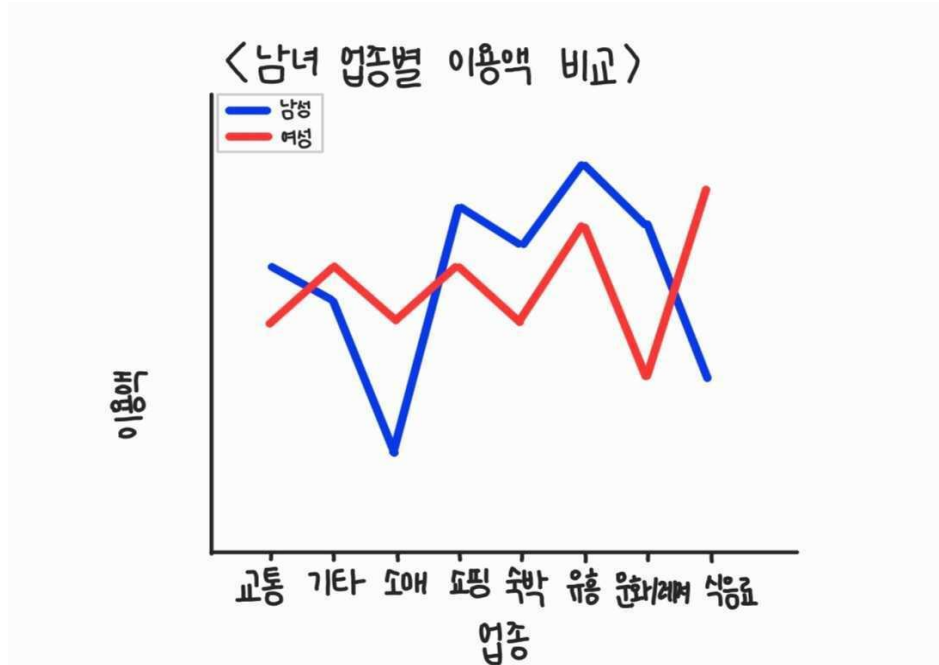


③ 누가 돈을 많이 썼을까 (1)?

- 남/녀 업종별 이용금액 그래프

업종별 남녀의 카드 이용금액을 두 개의 꺾은선 그래프로 표현

: x축 - 업종 / y축 - 이용금액



④ 누가 가장 많은 돈을 썼을까 (2)?

- 연령별 이용금액 그래프

식음료 업종에 해당하는 연령별 카드 이용 금액을 가로축 막대그래프로 시각화

: x축 - 연령 / y축 - 이용금액



2) 데이터 선정하기

데이터 : 제주데이터허브 - 내국인 관광객 지역, 업종, 성별, 연령대별 카드 이용 데이터

<https://www.jejudatahub.net/data/view/data/597>

2. 프로그램 구현

하위목표1

월별 제주도 이용금액 그래프

알고리즘

1. 월별 빈 리스트 만들기 `sum_mon = [0]*12`
2. 전체 이용 내역을 돌며 반복하기 `for row in data`
 - a. `if row[8][:4] == '2018'`
 - b. `month = int(row[8][4:]) - 1` # 1월의 인덱스 = 0 / 2월 = 1 / 3월 = 2 ...
 - c. `sum_mon[month]` 위치에 월별 이용금액(`row[-1]`) 더하기
3. 막대그래프 그리기

코드

```
def analysis1():  
    # 월별 카드이용금액 합계  
    sum_mon = [0]*12  
    for row in data:  
        if row[8][:4] == '2018':  
            month = int(row[8][4:]) - 1  
            sum_mon[month] += int(row[-1])  
    for i, sum in enumerate(sum_mon):  
        sum_mon[i] = sum/10000000000  
  
    # 그래프 그리기  
    plt.figure(figsize=(8, 4))  
    plt.rc('font', family='Malgun Gothic')  
    plt.title('제주도 내 월별 카드 이용금액 그래프')  
    plt.bar(range(1, 13), sum_mon, color='orange')  
    plt.xlabel('월별')
```

```
plt.ylabel('이용금액 (단위:100억)')
plt.xticks(range(0,13))
plt.ylim([10,15])
plt.plot(range(1, 13), sum_mon, color='green')
plt.savefig('graph.png')
plt.show()
```

하위목표2

제일 사용량이 많은 월 기준 업종별 그래프를 그려본다.

알고리즘

1. 이용금액이 가장 높은 달(max_sum_mon) 확인 max_sum_mon = max_mon()
2. 빈 딕셔너리(sum_shop)를 만들기 sum_shop = {}
3. 전체 이용 내역을 돌면 반복하기 for row in data
 - a. 이용금액이 가장 높은 달에 해당하는 row 구하기 if row[8][4:] == max_sum_mon:
 - b. shop(임시 변수)에 해당 row의 업종 할당 shop = row[4]
 - c. sum_shop의 키값이 shop에 있을때만 해당 데이터 합산
- if shop not in sum_shop.keys():


```
sum_shop[shop] = int(row[-1])
```
- else:


```
sum_shop[shop] += int(row[-1])
```
4. 파이 그래프 그리기

코드

```
# 금액이 가장 높은 달 확인하는 함수
def max_mon():
    sum_mon = [0]*12
    for row in data:
        if row[8][4:] == '2018':
            month = int(row[8][4:]) - 1
            sum_mon[month] += int(row[-1])
```

```
max_sum=max(sum_mon)
max_sum_mon=sum_mon.index(max_sum)+1
```

```
if len(str(max_sum_mon)) != 2:
    max_sum_mon='0'+str(max_sum_mon)
return max_sum_mon
```

```
# 해당 달의 업종별 이용금액 비중 분석 함수
```

```
def analysis2():
```

```
    max_sum_mon = max_mon()
    sum_shop = {}
```

```
for row in data:
```

```
    if row[8][4:] == max_sum_mon:
        shop = row[4]
        if shop not in sum_shop.keys():
            sum_shop[shop] = int(row[-1])
        else:
            sum_shop[shop] += int(row[-1])
```

```
# 그래프 그리기
```

```
plt.figure(figsize=(8, 4))
```

```
plt.title(max_sum_mon+"월 업종별 카드 이용금액 그래프")
```

```
sum_shop = dict(sorted(sum_shop.items(), key=operator.itemgetter(1), reverse = True))
```

```
plt.pie(sum_shop.values(), labels = sum_shop.keys(), autopct='%1.1f%%')
```

```
plt.savefig('graph.png')
```

```
plt.show()
```

하위목표 3

업종별 남녀의 카드 이용금액을 두 개의 꺾은선 그래프로 표현

알고리즘

1. male, female 각 딕셔너리 만들기 = {'교통': [0]*12, '기타': [0]*12, '소매': [0]*12, '쇼핑': [0]*12, '숙박': [0]*12, '유흥': [0]*12, '문화/레저': [0]*12, '식음료': [0]*12}
2. 전체 이용 금액을 돌며 반복하기 for row in data
 - a. shop(임시 변수)에 해당 row의 업종 할당 shop = row[4]
 - b. month(임시 변수)에 해당 row의 월 할당 month = int(row[8][4:])-1
 - c. if '여' : female[업종][달]에 int이용금액 추가하기
 - d. elif '남' : male[업종][달]에 int이용금액 추가하기
3. 각 업종 월별금액을 합산하기 위해 male_sum, female_sum에 빈 딕셔너리 생성
4. 업종 리스트를 for문으로 반복
 - a. 업종에 대한 월별 이용금액을 male_sum, female_sum에 저장
5. 남, 여 꺾은선그래프 생성
6. x축=업종, y축=이용금액 labels(남,여)

코드

```
def analysis3():
    # 빈 딕셔너리 생성
    male = {'교통': [0]*12, '기타': [0]*12, '소매': [0]*12, '쇼핑': [0]*12, '숙박': [0]*12, '유흥': [0]*12, '문화/레저': [0]*12, '식음료': [0]*12}
    female = {'교통': [0]*12, '기타': [0]*12, '소매': [0]*12, '쇼핑': [0]*12, '숙박': [0]*12, '유흥': [0]*12, '문화/레저': [0]*12, '식음료': [0]*12}

    for row in data:
        shop=row[4]
        month = int(row[8][4:])-1
        if row[7] == '여':
            female[shop][month]+=int(row[-1])          # 여성 데이터 합산
        elif row[7] == '남':
            male[shop][month]+=int(row[-1])            # 남성 데이터 합산

    male_sum = {}
```



```

female_sum = {}
for shop in female.keys():                # 단위 수정
    female_sum[shop] = sum(female[shop])/10000000000
for shop in male.keys():
    male_sum[shop] = sum(male[shop])/10000000000

# 그래프 그리기
plt.figure(figsize=(8, 4))
plt.rc('font', family='Malgun Gothic')
plt.title('남녀 업종별 카드 이용금액 비교 그래프')
plt.plot(male_sum.keys(),male_sum.values(), color='b', label='남성')
plt.plot(female_sum.keys(),female_sum.values(), color='r', label='여성')
plt.xlabel('업종')
plt.ylabel('이용금액 (단위:100억)')
plt.legend()
plt.savefig('graph.png')
plt.show()

```

하위 목표 4

알고리즘

1. `sum_age`(임시변수)에 빈 딕셔너리 생성
2. 전체 이용 금액을 돌며 반복하기 `for row in data`
 - a. 식음료 데이터만 사용 `if row[4] == '식음료'`
 - i. `sum_age[age : 각 연령대별 이용금액]` 딕셔너리로 저장
3. `sum_age`를 연령대 기준으로 내림차순(역방향) 정렬(단위 수정을 위해)
4. `age_li`에 `sum_age`의 키(업종)값들을 리스트로 저장
5. `age_li`(키값 리스트)를 `for` 문으로 반복
 - a. 각 나이별 이용금액을 10000000000으로 나누기
6. 가로 방향 막대그래프 생성 (`x축=연령`, `y축=이용금액`)

코드

```
def analysis4():
    sum_age = {}
    for row in data:
        if row[4] == '식음료':
            if row[6] == '20 미만' or row[6] == '20대미만':
                age = '10대이하'
            else:
                age = row[6]
            if age not in sum_age.keys():
                sum_age[age] = int(row[-1])
            else:
                sum_age[age] += int(row[-1])

    sum_age = dict(sorted(sum_age.items(), key=operator.itemgetter(0), reverse = True))
    age_li = list(sum_age.keys())

    for age in age_li:
        sum_age[age] = sum_age[age]/10000000000

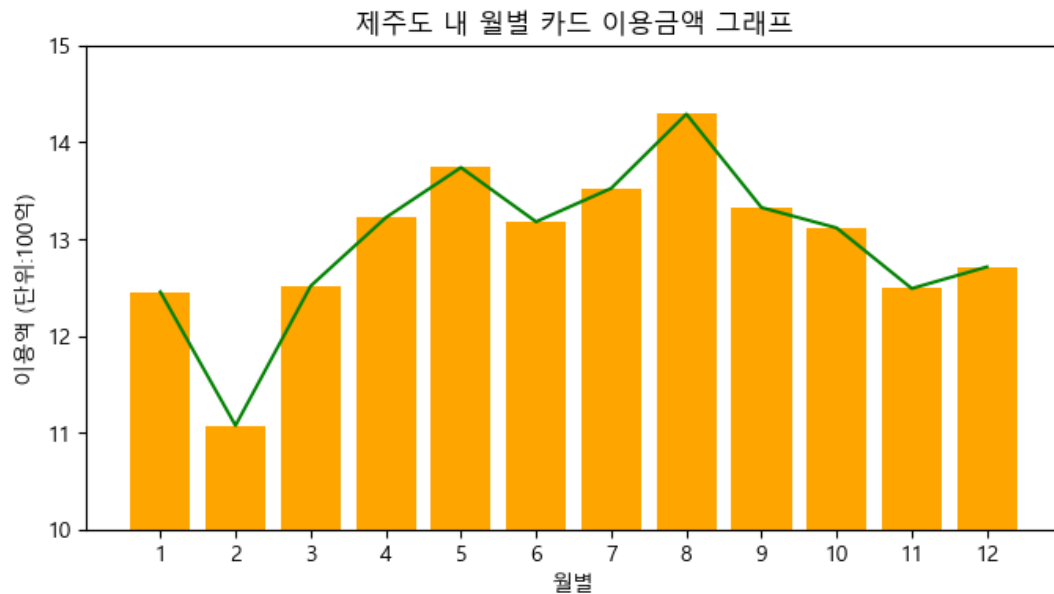
    # 그래프 그리기
    plt.figure(figsize=(8, 4))
    plt.rc('font', family='Malgun Gothic')
    plt.title('연령별 제주도 식음료 카드 이용 그래프')
    plt.xlabel('이용금액 (단위: 100억)')
    plt.barh(list(sum_age.keys()), list(sum_age.values()), color='orange')
    plt.savefig('graph.png')
    plt.show()
```

3. 결과 분석

하위목표1

➡ 8월, 5월, 7월, 9월, 4월, 6월, 10월, 12월, 3월, 11월, 1월, 2월 순으로 월별 제주도 이용금액이 높다.

따라서 8월에 팝업 스토어 개시를 목표로 설정한다.

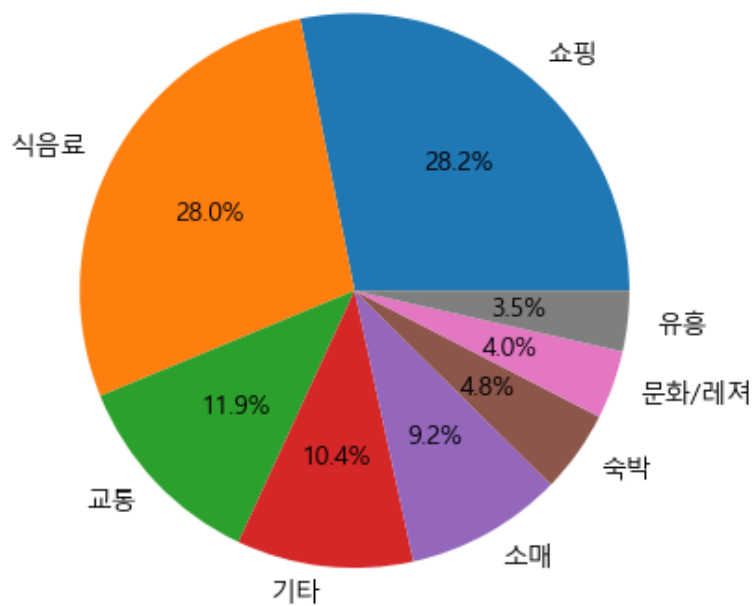


0

하위목표2

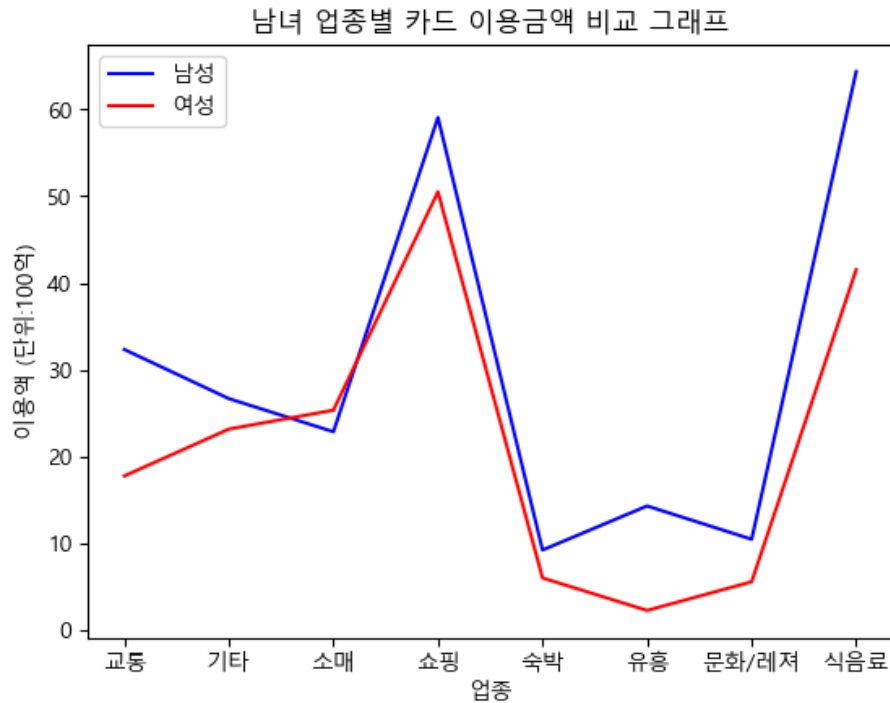
➡ 이용금액이 가장 높은 8월의 업종을 분석해보니 식음료와 쇼핑 업종의 이용금액이 높게 나왔다.

08월 업종별 카드 이용금액 그래프



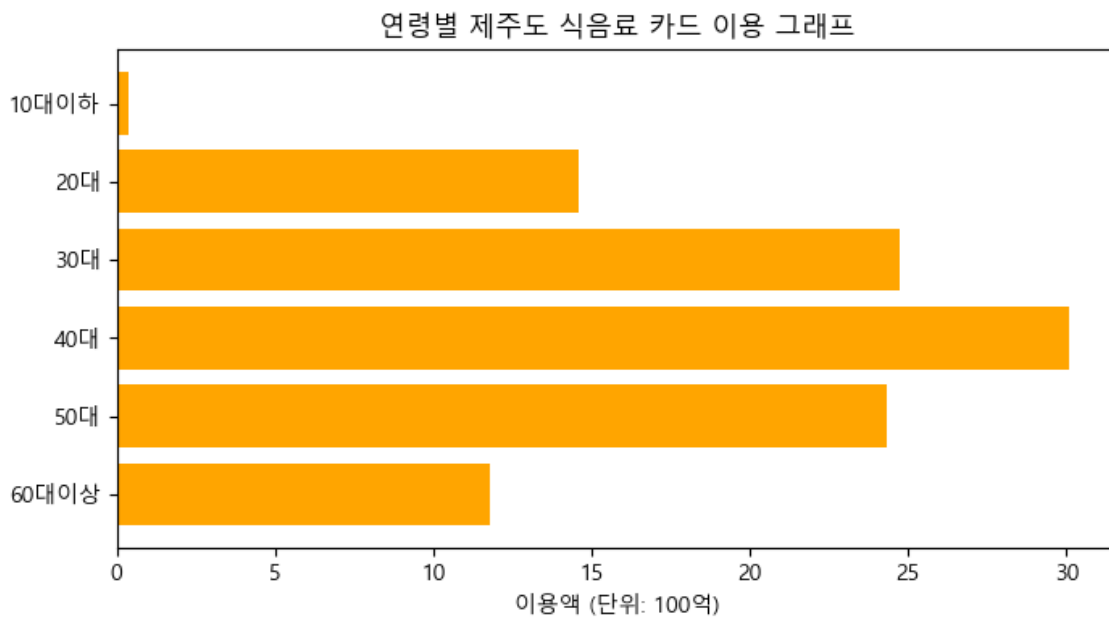
하위목표3

➡ 남성과 여성의 업종별 카드 이용금액을 비교해보니 거의 모든 업종에서 남성의 이용금액이 높았고, 특히 여성에 비해 쇼핑과 식음료, 유흥의 이용금액이 높았다. 여성은 남성에 비해 소매 부분이 높게 나왔다.



하위목표4

➡ 연령별 제주도 식음료 카드 이용 그래프를 보았을 때, 40대, 30대, 50대, 20대, 60대 이상, 10대 이하 순으로 카드 이용금액이 높음을 확인할 수 있다.



4. 결론

우리 JYP조는 제주도 공공데이터를 이용하여 카드이용액을 분석하였고, 코로나 시기에는 사람들의 소비패턴에 변화가 생겼을 것으로 추정, 코로나 이전 **2018년**을 분석하였다. 그리하여 분석결과를 통해 우리는 쇼핑과 식음료 부문에서 가장 많은 소비가 이뤄졌음을 알 수 있었다.

이번 분석은 우리가 제주도에서 이벤트성 팝업스토어를 론칭한다는 가정으로 시작되었다. 분석 결과, 쇼핑(백화점, 유통)은 **28.2%**, 식음료는 **28.0%**로 **0.2%**의 미미한 차이를 확인할 수 있고 팝업스토어 성격으로는 식음료가 더 적합하다고 판단하였다. 따라서 우리는 분석결과와 인사이트를 통해 성수기인 **8월**에 식음료 업종에 대한 비즈니스를 **40대** 남성이 포함된 가족을 타겟팅으로 준비하는 것이 좋을 것 같다고 판단하였다.