



# 뉴노멀 시대의 가족체험학습

국민여행조사 데이터 분석을 통한 뉴노멀 가족체험학습 활성화 정책

DATA220

# CONTENTS

## 1. 주제 선정

- 문제 제기 및 연구배경
- 주제 선정 과정
- 뉴노멀 시대의 체험학습

## 2. 정책 소개

- 정책 아이디어
- 정책 아이디어 구체화

## 3. 데이터 설명

- 분석 데이터 선정
- 데이터 전처리
- 외부 데이터 사용

## 4. 데이터 분석

- 여행객 분석
- 체험학습 장소 분석
- 여행자 패턴 분석

## 5. 데이터 분석 결과

- 결론
- 모델 작동 방식
- 활용 예상도

## 6. 결론 및 제언

- 정책제안
- 정책의 활용방안



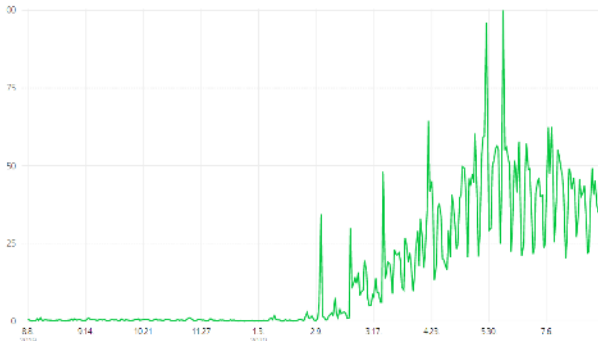
# 1. 주제 선정

## 문제 제기 및 연구배경



사회적 거리두기

**요약** 코로나19 확진자가 급증하면서 지역사회 감염 차단을 위해 실시된 정부의 권고 수칙으로, 많은 사람들이 모이는 행사 및 모임 참가 자제, 외출 자제, 재택근무 확대 등이 이에 해당한다. 사회적 거리두기는 코로나19 확산세가 계속된 3월 22일부터 시작돼 한 차례 연장됐고, 4월 20일부터 5월 5일까지 다소 완화된 형태로 16일간 연장돼 실시됐다. 그리고 안정화 추세가 겹쳐진 5월 6일부터 '생활 속 거리두기' 체계로 전환해서 시행했으나, 6월 28일부터는 각종 거리두기의 명칭을 '사회적 거리두기'로 통일하고, 코로나19 유행의 심각성과 방역조치의 강도에 따라 1~3단계로 구분해 시행하기로 했다.



▲ 포털 사이트 “언택트” 검색어 증가

[문 대통령 “언택트 시대…인터넷신문 역할 더 중요해질 것”](#)

코메디닷컴 **PICK** | 4일 전 | 네이버뉴스 | [🔗](#)

또한 "언택트 문화와 초연결 시대가 앞당겨지면서 인터넷신문에 대한 기대가 새로워지고 있다. 단순히 인터넷



[자치구도 언택트 열풍… 행정·교육 등 구정 혁신 의지 엿보여](#)

전남일보 | 18시간 전 | [🔗](#)

포스트코로나 시대를 맞아 하루하루 뻘뻘한 일정을 온라인·비대면으로 소화해내느라 바쁜 대상이라고 하면... 광산구는 지난달 29일 구청에서 언택트 화상회의로 '행복시책 추진상황 중간보고회'를 개최했다. 김삼호...

### 언택트(Untact)

[Untact]

코로나 19로 인해 새롭게 떠오르는 문화 '접촉하다'라는 의미의 '콘택트(contact)'에 부정적 의미인 '언(un)'을 합성한 단어. “ 언택트 ”

▶ 코로나 19로 인해 사회적 거리두기가 중요한 이슈로 부각 되어 새로운 사회 문화로 자리잡고 있음

# 1. 주제 선정

## 주제 선정 과정

무의미한 단체관광, 패키지여행 NO...관광 트렌드 바뀌었다

파이낸셜뉴스 입력 2020.05.06 18:30 수정 2020.05.06 18:30

김정섭 시장은 27일 시청 대백제실에서 열린 중간업무계획 보고회에서 "코로나19 장기화로 비대면 문화가 우리 생활 속에 자리 잡았고, 이에 따른 여행 문화도 크게 변화하고 있다"며 선제적 대응 필요성을 강조했다.

이어 "해외여행 대신 국내여행, 단체여행보다는 개별 소규모 가족여행으로 변화하는 추세이다. 특히 언택트 힐링 여행이 각광을 받고 있다"며, "변화하는 여행추세에 맞춰 신규 관광콘텐츠를 지속적으로 개발해 줄 것"을 당부했다.



## 2020~2024 변화 될 관광 트렌드

- 경험소비, 여행 액티비티 전성시대
- 소소한 여행- 소확행, 가심비, 가시비
- 여행 플랫폼 비즈니스의 진화
- 빅데이터가 여는 AI 관광 시대
- 일상에서 만나는 비일상 여행

코로나 19 이후 변화 될 관광의 모습은  
일상에서 소소하게 즐기는 여행, 데이터 기반의 여행  
이 증가할 것이다.

출처 : 한국문화관광연구원 문화.관광 인사이트 제 145호 (2020. 6.16)

▶ 여행, 관광에서도 생활 속 거리두기를 유지하면서  
자연에서 휴양할 수 있는 “언택트 관광”의 중요성 ↑

# 1. 주제 선정

## 뉴노멀 시대의 체험학습

### 한국교원대 '뉴노멀 시대 교육의 과제' 포럼 개최

(청주=뉴스1) 이성기 기자 | 2020-07-23 10:12 송고

뉴스홈 | 최신기사

### '학수고대한 수학여행도 코로나에 발목'...초중고 체험학습 차질

뉴스홈 | 최신기사

### 코로나 여파 부산 초중고 수학여행 전면 취소

송고시간 | 2020-07-13 14:40

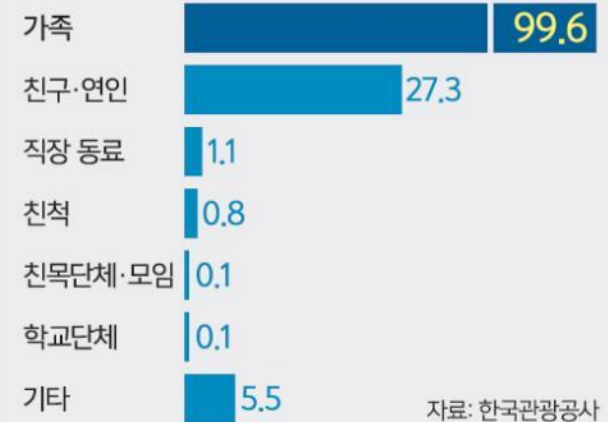
학교에서 진행하는 체험학습도 전면 취소 및 연기 되고 있음  
체험학습이 재개 시 사람이 많고 유명한 관광지 위주의 단체 체험  
학습에서 거리두기를 실천할 수 있는 뉴노멀 시대의 체험학습 필요

### 코로나19 국민 국내여행 영향조사 결과 (단위: %)

코로나19 이전과 이후  
여행패턴의 차이

있다	82.5
없다	17.5

### 국내여행 재개 시 여행 동반 의향자(중복응답)



응답자의 절대다수(99.6%)가 '국내여행 재개 시 여행  
동반자'로 가족을 꼽으면서, '2018년 국민여행조사'에서  
가족을 선택한 경우(49.4%)보다 배 넘게 증가

▶ **학교의 선생님, 가정의 부모님 들을 위한 새로운 방향의 체험학습을 제안**

## 2. 정책 소개

정책 아이디어

뉴노멀 시대,  
체험학습과 가족여행을 활성화 할 수 있는  
정책을 제안





## 2. 정책 소개

### 정책 실현을 위한 분석 아이디어 구체화

그동안 사람들이 많이 방문한 곳을  
분석하여 예상 관광객 수를 확인,  
사람들이 많이 몰리는 관광지를  
안내하여 밀집되지 않게 여행을  
떠날 수 있도록 정보 제공



지역 별 소비패턴을 분석하여  
체험여행 활성화가 낮은 지역을  
탐색, 지방 경제 활성화를 위해  
어느 지역에 체험 여행지를 늘려야  
하는지 정보 제공



가족단위 여행객의 취향과 여행시  
기를 분석하여 만족할 만한 여행지  
추천

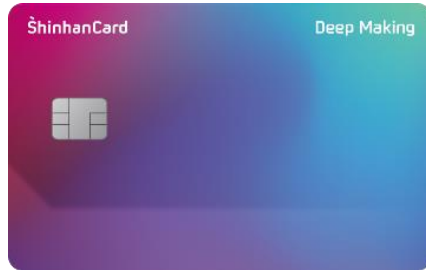


### 3. 데이터 설명

데이터 사용 - 데이터 소개

1. 필수 데이터 : 국민 여행 조사 데이터

2. 신한카드 결제 데이터  
(2019, 2020년)



**ShinhanCard**

3. 체험학습, 축제 관련 관광지 데이터

: 계절 테마여행 세부 코스정보 (한국 농어촌 공사)  
전국 문화축제 표준데이터 (공공데이터활용지원센터)  
전국 관광지 정보 표준데이터 (공공데이터활용지원센터)  
농어촌 체험 휴양마을 현황 (전라남도)



4. 주요관광지점 입장객 통계 (관광지식정보시스템)





### 3. 데이터 설명

분석 데이터 선정 : 국민 여행 조사 데이터

**데이터** : '국민여행조사' 에서 15세미만 동반자가 'NA' 가 아닌 가족단위 여행자 데이터 사용

2019 국민여행조사

**목적** : 가족단위 여행자들의 여행 행태를 파악하기 위함.

A7A | 국내 Q\_7.1\_여행동반자 중 15세미만 동반자수

	D_TRA1_1_SPOT1(B_여행1차_1번째방문 지~N차까지존재)	A2_1(Q_2_여행지선택이 유_1순위)	A2_2	A2_3	A3_1(여행 활동)	A3_11	A3_12	A12(Q_12_전반 적만족도)	A13	A14	DQ2	DQ2A(만 15세 이하 가족수 척도)	BAGE(연 령)	BMON(월 별)	A7B_1(가 족여행)	A7A(만 15세미 만)
0	31370.0	9	1.0	4.0	1.0	NaN	NaN	4	4	4	3	1.0	4	1	1.0	1
1	32380.0	5	12.0	NaN	1.0	NaN	NaN	4	4	4	3	0.0	5	1	1.0	1
2	32350.0	3	9.0	4.0	NaN	NaN	NaN	4	3	3	3	1.0	4	2	1.0	2
3	32340.0	5	8.0	4.0	1.0	NaN	NaN	4	4	4	3	1.0	3	2	1.0	2
4	23010.0	1	3.0	13.0	1.0	NaN	NaN	4	3	3	1	NaN	3	3	1.0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1272	32030.0	4	9.0	2.0	1.0	11.0	NaN	5	4	4	3	1.0	4	12	1.0	1
1273	11020.0	2	10.0	NaN	NaN	NaN	NaN	5	5	5	3	1.0	3	4	1.0	1
1274	11240.0	4	1.0	2.0	NaN	NaN	NaN	5	4	4	4	2.0	3	5	1.0	2
1275	31370.0	2	1.0	NaN	1.0	NaN	NaN	5	5	4	4	2.0	3	7	1.0	2
1276	25040.0	14	2.0	13.0	1.0	NaN	NaN	5	5	5	4	2.0	3	9	1.0	2

1277 rows x 16 columns

### 3. 데이터 설명

#### 데이터 전처리

#### 필요한 데이터 선정 및 변수 명 변경

##### 1. 시도 및 시군구 코드

시/도코드	시/도명	시/군/구명	시/군
11	서울특별시		
11	서울특별시	230	강남구
11	서울특별시	250	강동구
11	서울특별시	090	강북구
11	서울특별시	160	강서구

```
for x in range(0, len(traveler_2['1차방문지'])) :  
    if 11000 <= traveler_2['1차방문지'][x] < 12000:  
        traveler_2['1차방문지'][x] = '서울'  
    elif 21000 <= traveler_2['1차방문지'][x] < 22000:  
        traveler_2['1차방문지'][x] = '부산'  
    elif 22000 <= traveler_2['1차방문지'][x] < 23000:  
        traveler_2['1차방문지'][x] = '대구'  
    elif 23000 <= traveler_2['1차방문지'][x] < 24000:  
        traveler_2['1차방문지'][x] = '인천'  
    elif 24000 <= traveler_2['1차방문지'][x] < 25000:  
        traveler_2['1차방문지'][x] = '광주'  
    elif 25000 <= traveler_2['1차방문지'][x] < 26000:  
        traveler_2['1차방문지'][x] = '대전'  
    elif 26000 <= traveler_2['1차방문지'][x] < 27000:  
        traveler_2['1차방문지'][x] = '울산'  
    elif traveler_2['1차방문지'][x] == 29010:  
        traveler_2['1차방문지'][x] = '세종'  
    elif 31000 <= traveler_2['1차방문지'][x] < 32000:  
        traveler_2['1차방문지'][x] = '경기도'  
    elif 32000 <= traveler_2['1차방문지'][x] < 33000:
```

```
for x in traveler_2['1차방문지']:  
    print(x)
```

경기도  
서울  
경기도  
서울  
충청남도  
충청남도  
충청남도  
경기도  
전라남도

**전처리** : '시/도' 코드에 맞춰, 지역별 Category화

**목적** :

1. '지역 별' 가족 단위 여행 관광객 수 파악
2. '지역 별' 축제/체험 목적 여행객 파악
3. '지역 별' 축제/체험 만족도 파악 등

1차방  
문지

여행지선택이유  
\_1순위

0	경기도	9
1	강원도	5
2	강원도	3
3	강원도	5
4	인천	1
...	...	...
1272	강원도	4
1273	서울	2
1274	서울	4
1275	경기도	2
1276	대전	14

1277 rows x 16 columns

```
count_first = traveler_2['1차방문지'].value_counts()  
count_first
```

경기도 265  
경상북도 218  
강원도 206  
경상남도 173  
충청남도 75  
전라남도 54  
...

**그룹화** : '시/도' 별 그룹화

**목적** :

'시/도' 별 1차 여행지방문을 통해, 가족단위 여행객이 지역별 여행 우선도를 파악

### 3. 데이터 설명

**그룹화** : '시/도' 별 그룹화

**목적** :

'시/도' 별 1차 여행지방문을 통해, 가족단위 여행객 '지역' 우선도를 파악

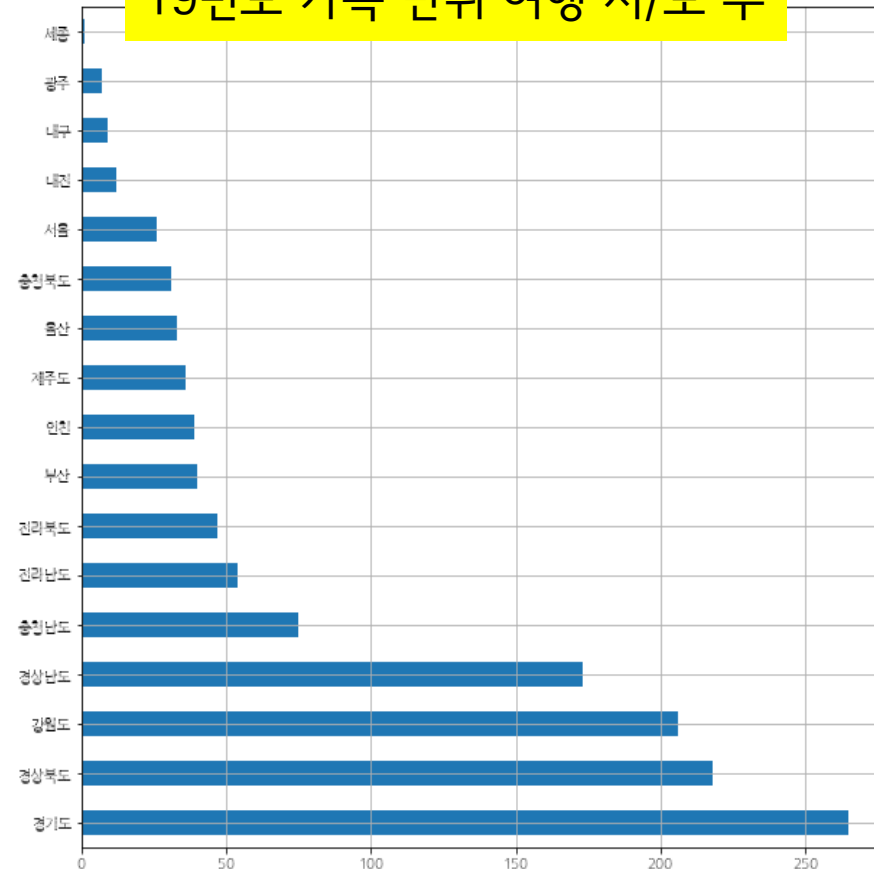
	1차방문지	여행지선택이유 __1순위
0	경기도	9
1	강원도	5
2	강원도	3
3	강원도	5
4	인천	1
...	...	...
1272	강원도	4
1273	서울	2
1274	서울	4
1275	경기도	2
1276	대전	14

1277 rows x 16 columns

```
count_first = traveler_2['1차방문지'].value_counts()
count_first
```

```
경기도    265
경상북도   218
강원도     206
경상남도   173
충청남도    75
전라남도    54
...
```

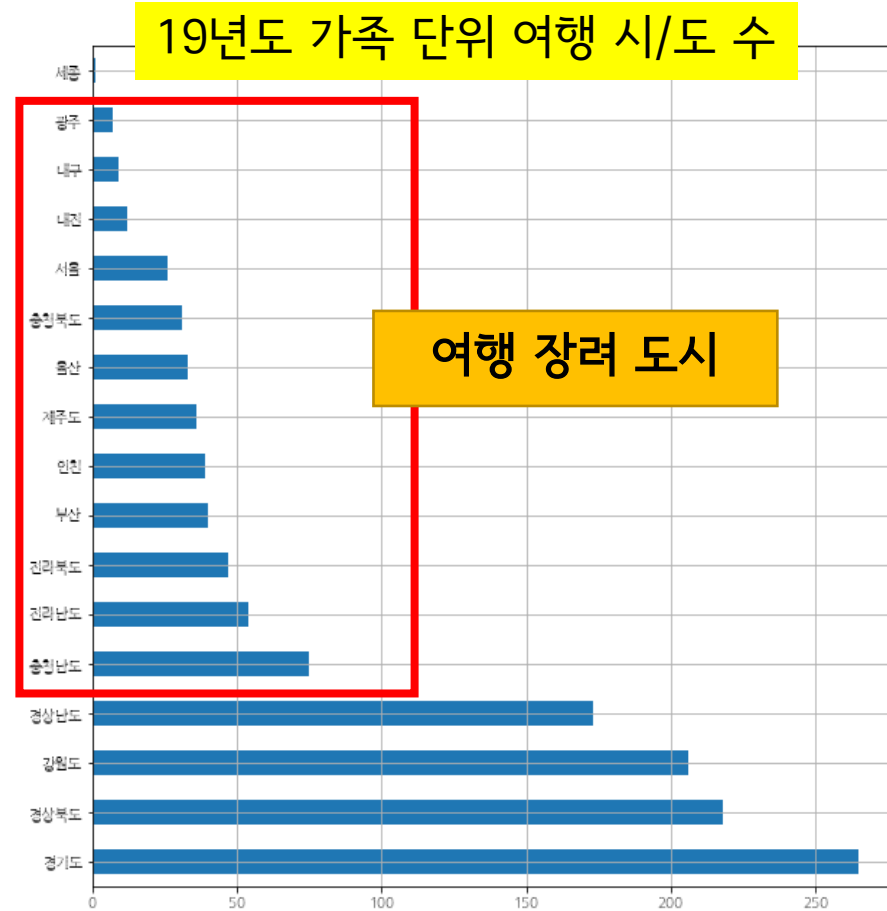
19년도 가족 단위 여행 시/도 수



**결론** : 가족단위 여행객이 19년도에 가장 많이 방문한 지역은 '경기', '경북', '강원', '경남' 순이다.

### 3. 데이터 설명

그동안 사람들이 많이 방문한 곳을 분석하여 예상 관광객 수를 확인, 사람들이 많이 몰리는 관광지를 안내하여 밀집되지 않게 여행을 떠날 수 있도록 정보 제공



결론 :

19년도 가족단위 관광객수는 '경기도', '경상북도', '강원도' 순으로 밀집하였다. 하여, 언택트시대에 맞게 '해당 관광객들을' 분포시켜 -> 광주, 대구, 대전 등에 여행을 장려할 필요가 있다.

# 4. 데이터 분석

## 여행객 분석

**목적** : 여행지 우선 선택 1순위, 2순위, 3순위 중 '11, 12(=지역축제, 체험학습)'을 목적으로 여행을 떠난 사람들 파악

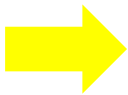
```
#여행 선택이유 1,2,3 순위가 11/12번인 사람들 = 지역축제/체험을 하기 위해 여행을 떠난 사람들

want_1 = traveler_2['여행지선택이유_1순위'] == 11
want_11 = traveler_2['여행지선택이유_1순위'] == 12
want_2 = traveler_2['여행지선택이유_2순위'] == 11
want_22 = traveler_2['여행지선택이유_2순위'] == 12
want_3 = traveler_2['여행지선택이유_3순위'] == 11
want_33 = traveler_2['여행지선택이유_3순위'] == 12

want_11_12 = traveler_2[want_1 | want_2 | want_3 | want_11 | want_22 | want_33]
```

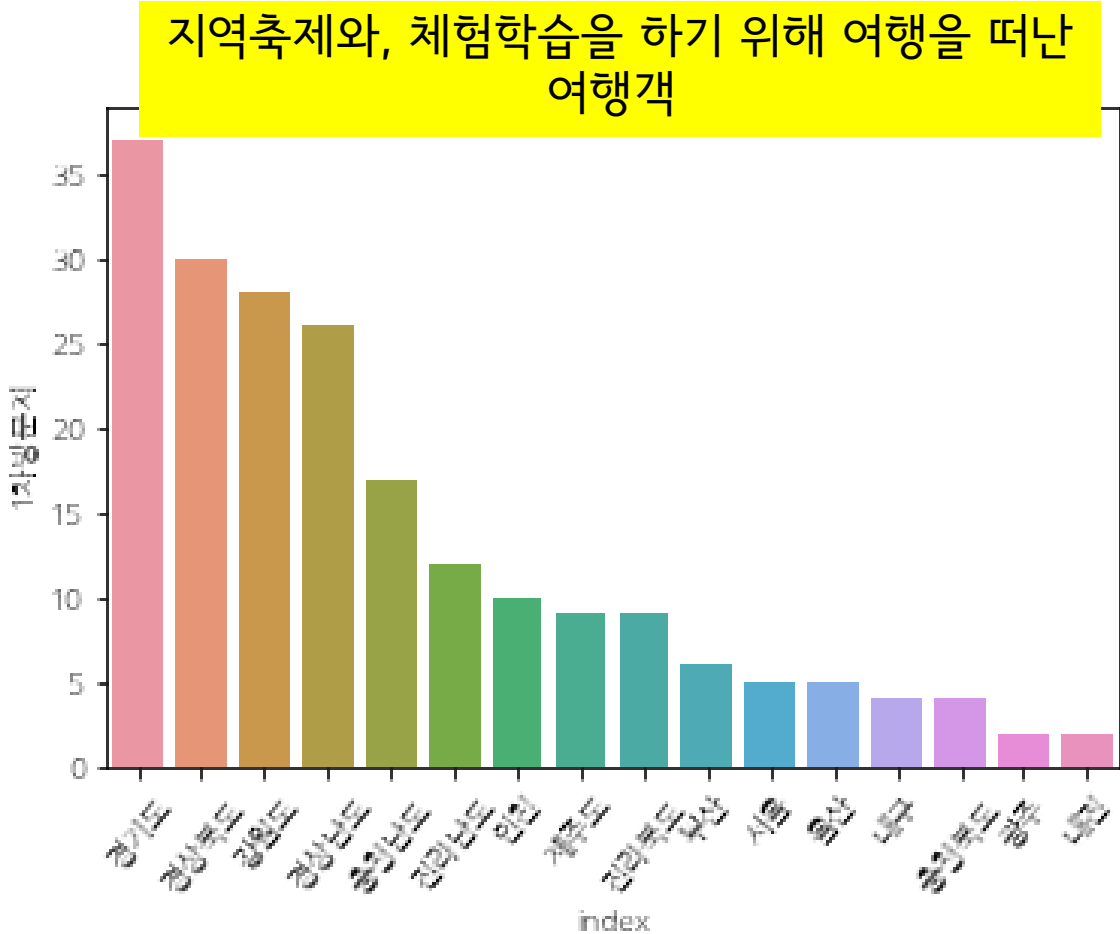
want\_11\_12  
#지역체험 축제 여행을 가고 싶었던 사람은 207명  
#그 중 실제로 11,12(지역축제 및 체험학습) 간 사람들의 명 수 파악

207명



	1차방문지	여행지선택이유_1순위	여행지선택이유_2순위	여행지선택이유_3순위
17	강원도	1	2.0	11.0
37	경기도	11	5.0	3.0
43	강원도	1	11.0	NaN
48	전라남도	12	1.0	8.0
57	강원도	2	12.0	11.0
...	...	...	...	...
1236	대전	1	11.0	4.0
1244	강원도	8	11.0	5.0
1251	서울	11	2.0	8.0
1256	충청남도	12	8.0	5.0
1265	경상남도	12	1.0	2.0

207 rows x 16 columns



## 4. 데이터 분석

### 여행객 분석

**목적** : 지역축제, 체험학습을 택한 사람들 중 실제로 '지역축제', '체험학습' 을 여행활동으로 경험 한 사람

```
i) really = real_want['1차방문지'].value_counts()
```

```
really
```

```
#지역체험/축제를 가기로 했으나, 실제로 지역/축제 체험한 사람들은 11명
```

```
#지역체험/축제를 목표로 한 사람들은 207 명
```

```
#그 간의 격차가 벌어짐
```

```
경상북도 4
```

```
충청남도 2
```

```
경기도 2
```

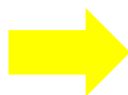
```
서울 1
```

```
경상남도 1
```

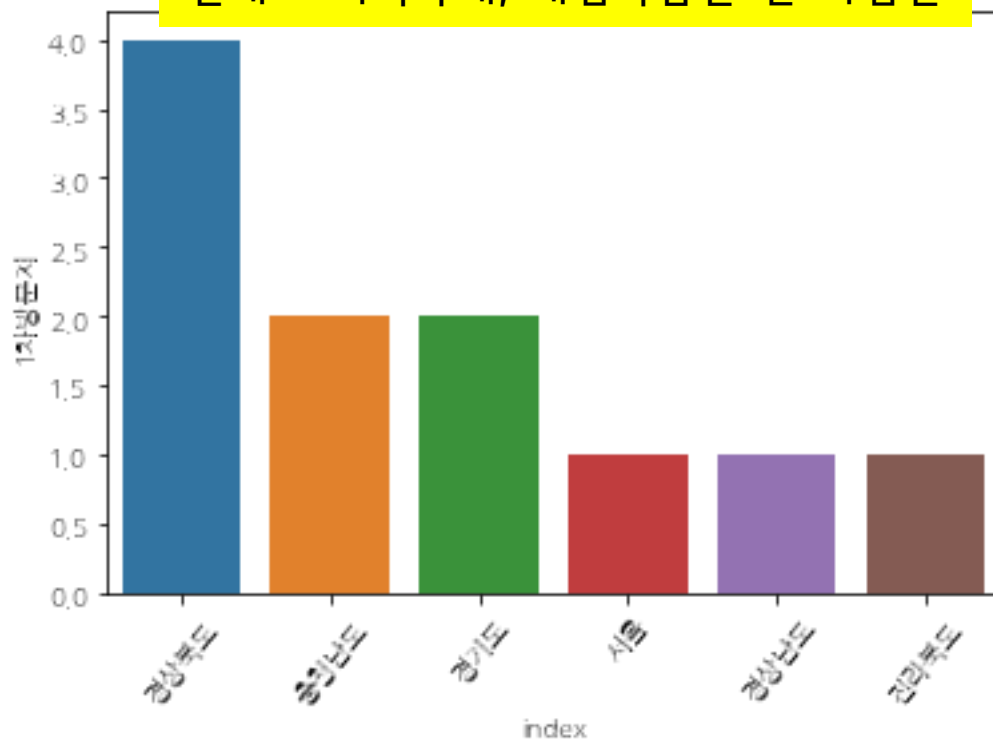
```
전라북도 1
```

```
Name: 1차방문지, dtype: int64
```

11명



실제로 지역축제, 체험학습을 한 사람들

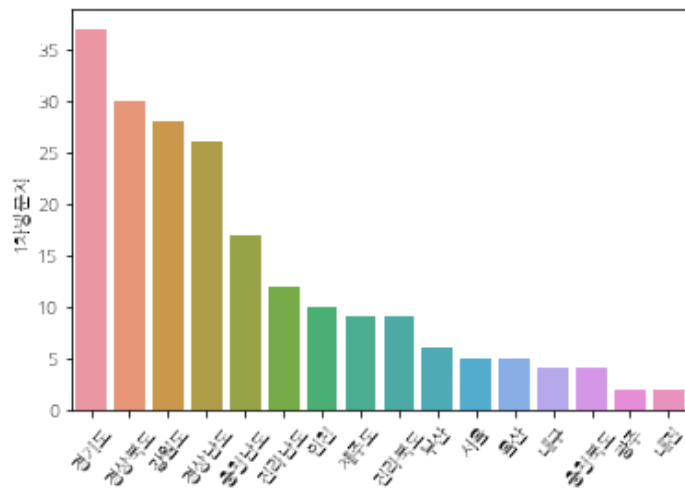


## 4. 데이터 분석

### 여행객 분석

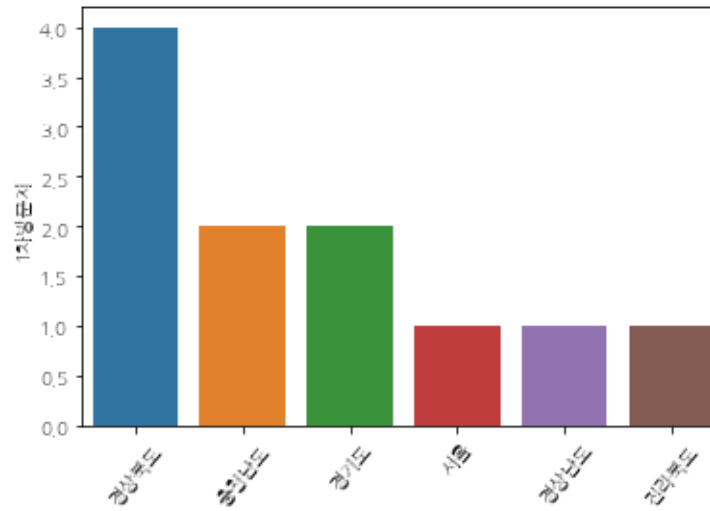
**목적** : 지역축제/체험 학습을 목적으로, 여행을 떠난 사람은 207명이다.  
그러나, 실제로 지역축제/체험 학습을 경험한 사람은 11명이다.  
이를 통해, 실질적으로 지역체험/학습을 목표로 하여도, 경험하지 못한 실태를 파악

207명



여행지선택 우선순위(지역축제,체험)

11명



실제로 지역축제, 체험학습을 한 사람들

196명

?

### 실태조사

11/207 명 약 5%정도의 여행객만 계획했던 '지역축제/체험학습'에 참가하였다.

**문제점** : 지역별 체험학습을 계획했어도, 실제로 체험하는 사람의 수는 적음.



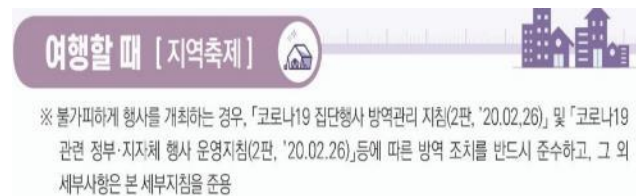
## 4. 데이터 분석

### 여행객 분석 **지역별 밀집도 문제**

**목적 :**

지역축제/체험학습의 문제는 '체험형 여행지 1곳' 별, 밀집도가 너무 높다는 것이다.

지역	주요관광지점	입장객통계	체험형 여행지 수	체험형 여행지 1곳 당 여행객 수
강원도		39938894	73	547108.137
경기도		72883450	97	751375.7732
경상남도		40002474	50	800049.48
경상북도		54516461	60	908607.6833
광주광역시		6263739	14	447409.9286
대구광역시		14075975	39	360922.4359
대전광역시		7987868	15	532524.5333
부산광역시		11931327	24	497138.625
서울특별시		24983320	35	713809.1429
세종특별자치시		700625	2	350312.5
울산광역시		5383106	7	769015.1429
인천광역시		7301559	9	811284.3333
전라남도		62422807	268	232920.9216
전라북도		38888678	34	1143784.647
제주특별자치도		14221336	22	646424.3636
충청남도		29707754	67	443399.3134
충청북도		30267942	50	605358.84



**뉴노멀 시대  
맞춤형  
축제/체험학습**

### 사례조사

1. 지역축제/체험학습은 시존성이 강하다.
2. 또한, 지역축제/체험학습은 체험지 별 사람 밀집도가 너무 높다는 것이다.  
(=>언택트 시대에 맞는, 체험/지역축제로 변화되어야 함)

## 4. 데이터 분석

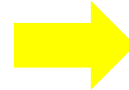
### 여행객 만족도 조사

**목적** : 지역축제/체험학습을 다녀온 사람들의 지역별, 만족도 조사

```
subset_df['전반적만족도'] + subset_df['재방문의도'] + subset_df['타인추천의도']
```

	1차방문지	여행활동_11. 지역 축제_이벤트 참가	여행활동_12. 교육_체험 프로그램 참가	전반적 만족도	재방문 의도	타인추 천의도
0	경기도	11.0	NaN	4	4	4
1	강원도	NaN	12.0	4	4	4
2	강원도	11.0	NaN	4	3	3
3	강원도	NaN	12.0	4	4	4
4	인천	NaN	12.0	4	3	3
...	...	...	...	...	...	...
1271	경기도	...	...	4	4	4
1273	서울	11.0	NaN	5	5	5
1274	서울	11.0	NaN	5	4	4
1275	경기도	NaN	12.0	5	5	4
1276	대전	NaN	12.0	5	5	5
		11.0	NaN			

1215 rows × 7



total_satis	
1차방문지	
경기도	301
경상북도	221
강원도	207
경상남도	151
전라북도	82
충청남도	70
전라남도	60
제주도	54
울산	39
서울	25
대구	24
부산	24
세종	13
인천	12
충청북도	12

지역축제 or 교육체험  
을 참가한 여행객들의  
'지역 별', '만족도' 를  
조사한다.

## 4. 데이터 분석

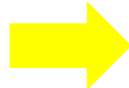
### 여행객 만족도 조사

**목적** : 지역축제/체험학습을 다녀온 사람들의 지역별, 만족도 조사

```
gu_join['만족지수'] = gu_join['total_satis'] / gu_join['응답자수']  
gu_join.sort_values('만족지수')
```

#그렇다면, 어떤 이유에서 지역별 지역 체험 차이가 날까?

	1차방문지	total_satis	응답자수	만족지수
3	경상남도	151	13	11.615385
5	충청남도	70	6	11.666667
4	전라북도	82	7	11.714286
6	전라남도	60	5	12.000000
10	대구	24	2	12.000000
11	부산	24	2	12.000000
13	인천	12	1	12.000000
14	충청북도	12	1	12.000000
0	경기도	301	25	12.040000
2	강원도	207	17	12.176471
1	경상북도	221	18	12.277778
9	서울	25	2	12.500000
8	울산	39	3	13.000000
12	세종	13	1	13.000000
7	제주도	54	4	13.500000



여행객들의 '지역 별' 체험/축제 만족도는 11~13 지수로  
동일함을 보였다.  
즉, 체험/축제를 목적으로 여행을 떠난 사람들의 '지역별'  
'만족도는 거의 동일하다.  
(=축제/체험 학습의 만족이 지역간 차이가 거의 없음)

\*만족지수 = 지역별 총 만족도 / 응답자수

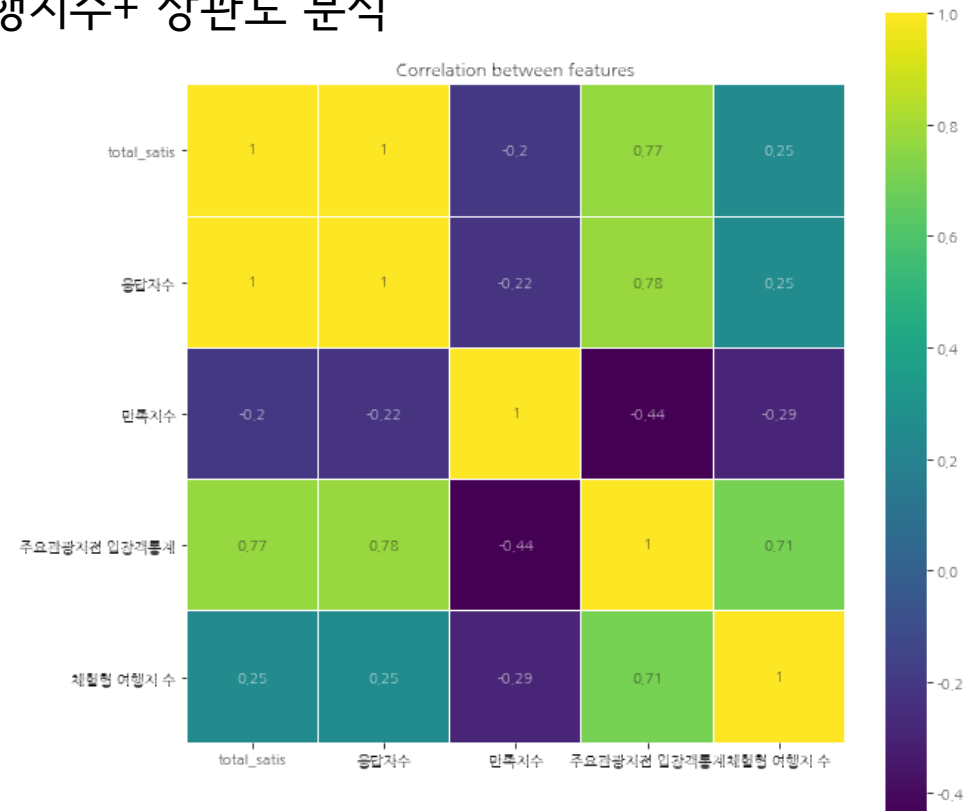
## 4. 데이터 분석

### 여행객 만족도 상관도 파악

**데이터** : 국민여행조사 보고서 + 관광데이터 + 신한카드 이용 데이터를 결합  
'지역 별' 관광객 통계, 지역별 '만족도'와 '체험형 여행지수+ 상관도 분석

```
gu_join_2 = pd.merge(gu_join, tour_select, on='1차방문지')  
gu_join_2
```

	1차방문지	total_satis	응답자수	만족지수	주요관광지점	입장객통계	체험형 여행지 수
0	경기도	301	25	12.040000		72883450	97
1	경상북도	221	18	12.277778		54516461	60
2	강원도	207	17	12.176471		39938894	73
3	경상남도	151	13	11.615385		40002474	50
4	전라북도	82	7	11.714286		38888678	34
5	충청남도	70	6	11.666667		29707754	67
6	전라남도	60	5	12.000000		62422807	268
7	울산	39	3	13.000000		5383106	7
8	서울	25	2	12.500000		24983320	35
9	대구	24	2	12.000000		14075975	39
10	부산	24	2	12.000000		11931327	24
11	세종	13	1	13.000000		700625	2
12	인천	12	1	12.000000		7301559	9
13	충청북도	12	1	12.000000		30267942	50



**결론** : 체험형 여행지(=지역별 총 체험 여행지 개수) 와 total satis(=지역별 총 만족도) 는 '약한 상관관계'  
여행지 개수의 차이는 만족도에 큰 영향을 주진 못한다.  
때문에, 소비자 데이터를 분석하여 이유를 찾아보기로 했다.

### 3. 데이터 설명

외부 데이터 사용 - 데이터 전처리

신한카드 결제 데이터

	v1	v2	v3	gb3	gb2	sex_ccd	cln_age_r	ta_ym	daw_ccd_r	apv_ts_dl_tm_r	vlm	usec
0	세종	세종	.	취미오락	종합쇼핑	M	40	201901	WHITE	활동	1.878919e+09	63579
1	대전	세종	.	취미오락	외식	M	20	201901	WHITE	휴식	6.945233e+07	2952
2	서울	세종	.	취미오락	외식	F	40	201901	RED	휴식	9.267020e+06	286
3	경남	세종	.	취미오락	종합쇼핑	M	50	201901	RED	휴식	6.034990e+05	20
4	충북	세종	.	취미오락	종합쇼핑	F	40	201901	WHITE	활동	1.076389e+08	2379

1. 국민 여행 조사에서 여행의 정의를 거주하는 시, 도에서 벗어나는 여행만을 여행으로 정의하고 있음  
→ 따라서 이용회원 거주지(v1)와 카드이용 가맹점 주소(v2)가 다른 결제건 만을 데이터로 사용한다.
2. 여행 중에서도 체험과 관련된 데이터를 분석할 것이므로,  
→ 가맹점 업종 대분류(gb3)가 여행이고, 가맹점 업종 소분류(gb2)가 체험인 데이터만을 사용한다.
3. 새벽 시간에는 체험학습과 관련된 결제가 이루어지지 않는다고 판단하기에,  
→ 결제가 이루어진 시간(apv\_ts\_dl\_tm\_r)을 휴식과 활동으로 한정한다.

### 3. 데이터 설명

#### 외부 데이터 사용 - 데이터 전처리

```
from sklearn.model_selection import train_test_split
```

```
s19_20_ex, imsi = train_test_split(s19[s19["연령대"] == 20], train_size=0.024, shuffle=True, random_state=13)
s19_30_ex, imsi = train_test_split(s19[s19["연령대"] == 30], train_size=0.462, shuffle=True, random_state=13)
s19_40_ex, imsi = train_test_split(s19[s19["연령대"] == 40], train_size=0.479, shuffle=True, random_state=13)
s19_50_ex, imsi = train_test_split(s19[s19["연령대"] == 50], train_size=0.028, shuffle=True, random_state=13)
s19_60_ex, imsi = train_test_split(s19[s19["연령대"] == 60], train_size=0.007, shuffle=True, random_state=13)
```

```
s19_total_ex = pd.concat([s19_20_ex, s19_30_ex, s19_40_ex, s19_50_ex, s19_60_ex])
```

#### 국민 여행 조사 데이터

각 연령 15세 미만 아이가 있는 인원 수

각 연령 전체 인원 수

	이용회원 거주지	카드이용 가맹점 주소	가맹점 업종 대분류	가맹점 업종 소분류	성별	연령대	이용년월	요일	이용시간대	취급액	이용건수
157950	강원	경북	여행	체험	F	20	201908	RED	휴식	243360.0	10
146625	인천	서울	여행	체험	M	20	201907	WHITE	활동	35420.0	5
265021	강원	충남	여행	체험	F	20	201912	WHITE	휴식	64375.0	5
120291	경북	경기	여행	체험	M	20	201906	WHITE	휴식	70140.0	5
275314	전남	제주	여행	체험	M	20	201912	RED	활동	1350330.0	30

1. 신한카드 데이터에서는 각 개인의 15세 미만 자녀의 유무를 알 수 없음

➔ 국민 여행 조사 데이터에서, 연령대별 15세 미만 자녀가 있는 사람의 비율을 계산

➔ 신한카드 데이터에서 해당 비율을 가중치로 사용하여,  
가중치 만큼의 결제 건을 각 세대별로 무작위 표본 추출

### 3. 데이터 설명

#### 외부 데이터 사용 - 데이터 전처리

##### 여행지 정보 데이터

체험명	주소	체험설명	위도	경도
가덕마을	전라남도			
가시리마을	제주특별자치도	▶ 체험 승마채꽃프라자	33.3831	126.7363
가맛고마을	경상북도	* 추천 체험   흙 등 * 소재	35.74176	128.2609
가정농촌체험휴양마을	전라남도	감 따기, 고사		
갯고을마을	전라남도	감자캐기, 깃		
갯고을마을	전라남도	* 추천 체험:	34.63989	127.7869
강변사리마을	전라북도	▶ 체험 산야이트, 글램핑	35.49574	127.1903
강변사리마을	전라북도	▶ 체험 산야이트, 글램핑	35.49574	127.1903
개령이고들빼기마을	전라남도	고들빼기 김		
개미들마을	강원도	* 추천 체험: 재지 : 강원도	37.31905	128.6927
개실마을	경상북도	체험 프로그램들기 5천원	35.66798	128.2298
거차벌레체험마을	전라남도	거차벌레 체험별미고름틀	34.8376	127.4475
계도어촌체험마을	경상남도	가조도의 서리덕 젓갈, (	34.97484	128.5175
고라데이마을	강원도	- 체험 : 고령폭포, 횡성	37.63122	128.1429

##### 주요관광지점 입장객 통계

	시도	총계_x	총계_y
0	서울특별시	24983320	2212997.0
1	부산광역시	11931327	1373010.0
2	대구광역시	14075975	1908574.0
3	인천광역시	7301559	NaN
4	광주광역시	6263739	391682.0
5	대전광역시	7987868	718373.0
6	울산광역시	5383106	NaN
7	세종특별자치시	700625	63499.0
8	경기도	72883450	4173590.0
9	강원도	39938894	1765063.0

- 여행지명, 주소, 설명과 위도, 경도 정보를 선택해 새로운 데이터 셋 구성  
→ 사용할 외부 데이터에서 '체험'을 키워드로 가지는 데이터만을 선별해 병합
- 위도, 경도가 같은 데이터를 제거함으로써, 같은 장소나 축제에 대한 데이터를 제외시킴
- 여행지 주소는 분류를 위해 17개 시도명으로 바꾸어줌

- 관광지점 입장객 통계에서 내국인 데이터만을 선택
- 개별 관광지 방문객을 합해 시도별 방문객 수의 총합을 계산
- 2020년, 5개 시도의 통계 조사 미완료로 결측치 존재

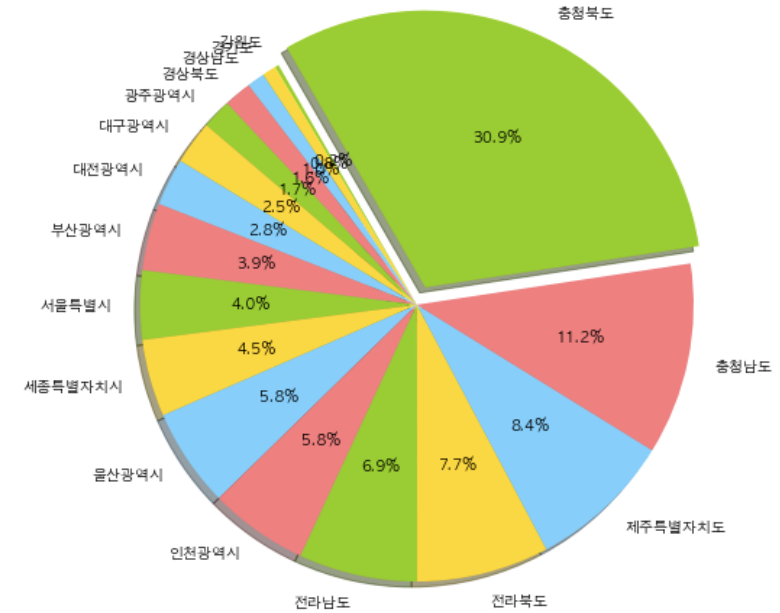


## 4. 데이터 분석

### 여행지 데이터 분석

지역	주요관광지점 입장객통계	체험형 여행지 수	총 여행 횟수	총 결제 인원 수	체
강원도	39938894	73	36699	13901	
경기도	72883450	97	36118	29363	
경상남도	40002474	50	24231	17814	
경상북도	54516461	60	24464	20755	
광주광역시	6263739	14	2884	2054	
대구광역시	14075975	39	4489	6735	
대전광역시	7987868	15	4181	4263	
부산광역시	11931327	24	15543	18458	
서울특별시	24983320	35	12846	23744	
세종특별자치시	700625	2	1029	731	
울산광역시	5383106	7	3651	1316	
인천광역시	7301559	9	10645	6535	
전라남도	62422807	268	24594	22210	
전라북도	38888678	34	18526	12868	
제주특별자치도	14221336	22	11841	8939	
충청남도	29707754	67	21955	18013	
충청북도	30267942	50	10589	10087	

### 전국 체험형 여행지 분포 비중

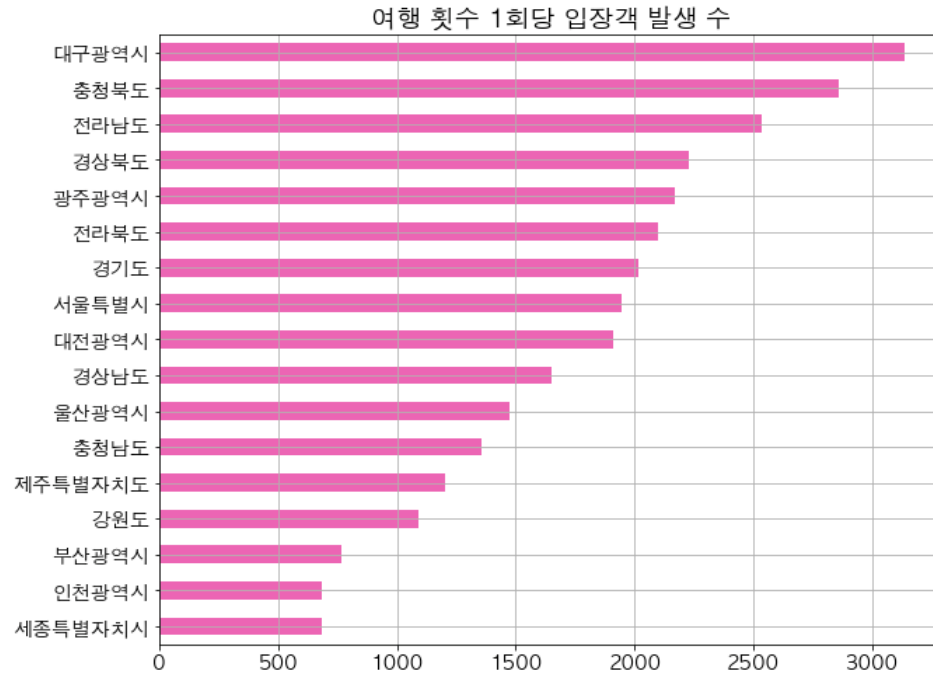


충북이 체험형 여행지 비중이 가장 높지만, 체험 관련 취급액에 있어서는 17개 시도중 중위권 이하. 체험형 여행지의 여행서비스 품질 이상을 의심해볼 수 있음.

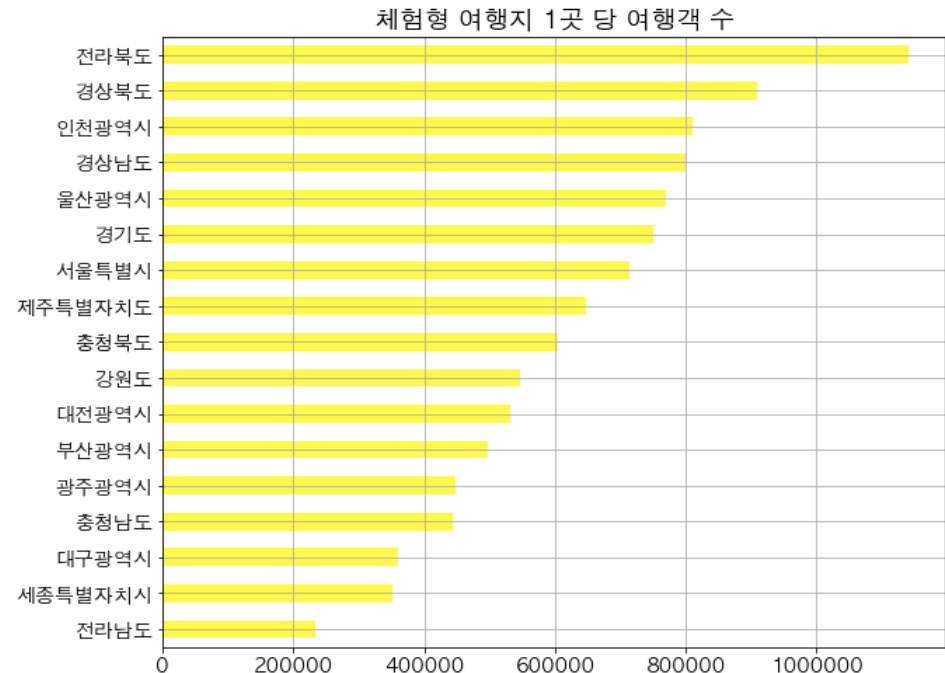
체험형 여행지 비중 높은 곳 : 충북, 충남, 제주, 전북, 전남  
 체험형 여행지 비중 낮은 곳 : 강원도, 경기도, 경남, 경북, 경남

## 4. 데이터 분석

### 여행지 데이터 분석



여행 횟수 1회당 입장객 발생 수가 가장 높은 지역은 대구. 이를 통해 대구, 충청북도, 전라남도 등 상위권 지역들은 다양한 여행 콘텐츠를 가진 것으로 해석할 수 있음.

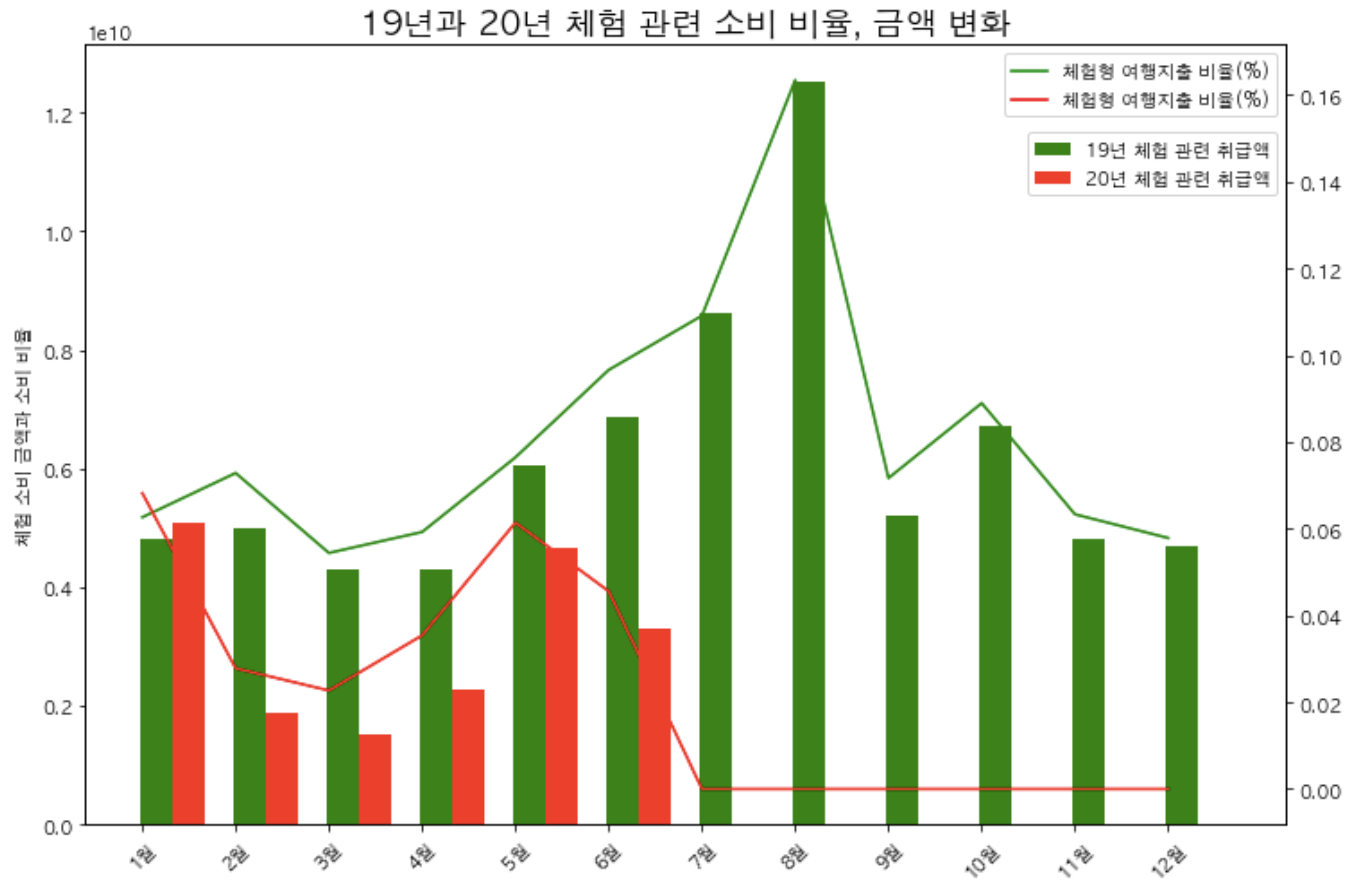


전북, 경북, 인천은 타 지역 대비 체험형 여행지 1곳 당 여행객 밀도가 높다. 이는 여행객을 위한 충분한 체험형 여행지가 제공되지 않고 있음을 의미한다. 거리두기를 통한 안전한 체험학습을 위해, 체험형 여행지를 충분히 확보함으로써 해당 수치를 낮출 필요가 있음.

높은 수준을 보이는 대구, 전남에 비해 전북은 3배 이상의 수치 차이를 보이고 있다. (세종 제외)

## 4. 데이터 분석

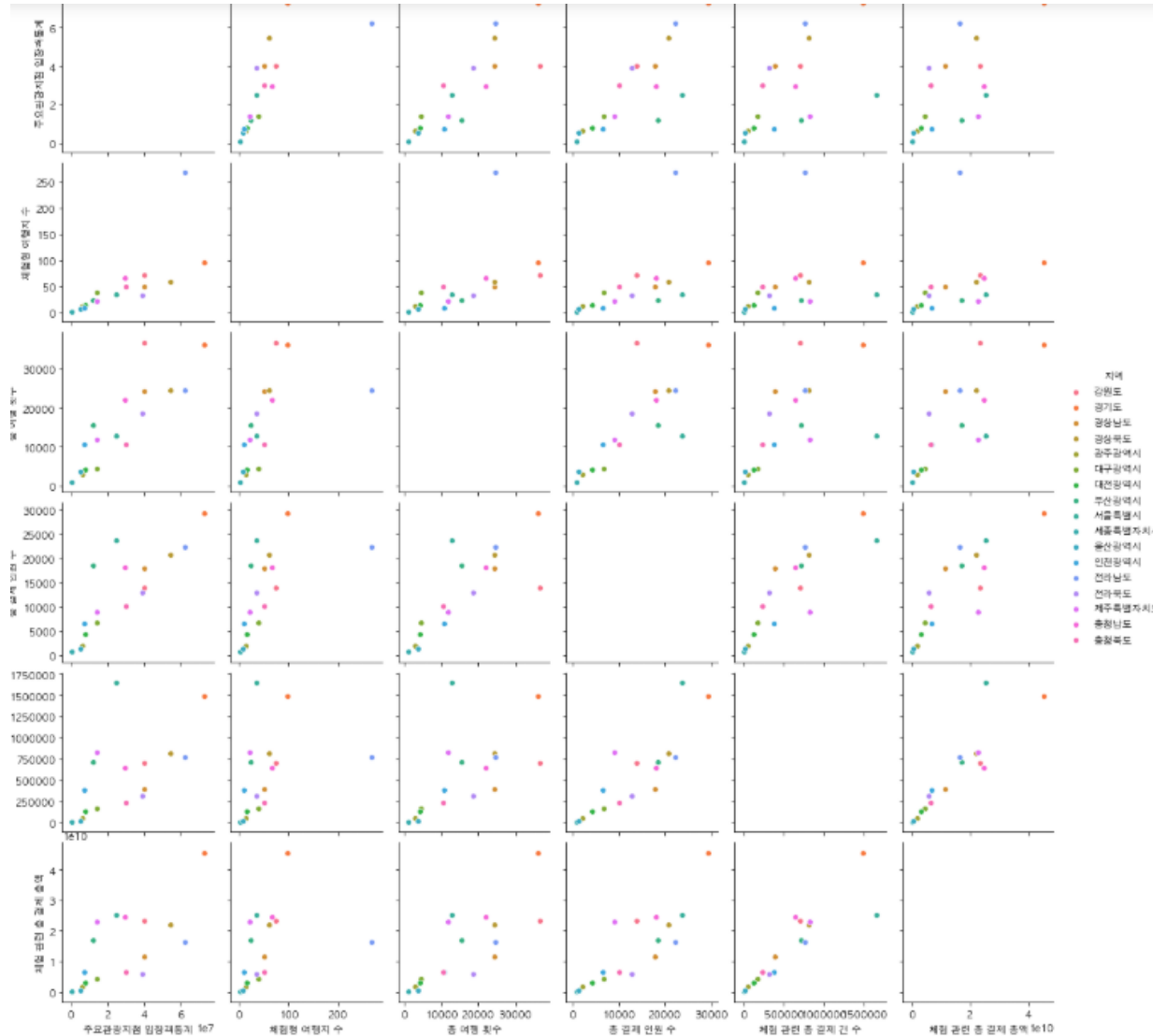
### 여행객 데이터 분석



예년의 추이 대로라면 1월까지 성장가능성 있었으나, 코로나로 인해 2월에 취급액의 절반 이상 급감. 코로나의 영향에도 불구하고 5월까지의 예년과 같은 동향을 보이고 있었으나, 6월부터 아예 반대의 동향(하락세)를 보이고 있음. 7, 8월 또한 장마가 오래 지속되어 상승세를 노리기 어려워 보임.

## 4. 데이터 분석

### 여행지 & 여행객 데이터 분석



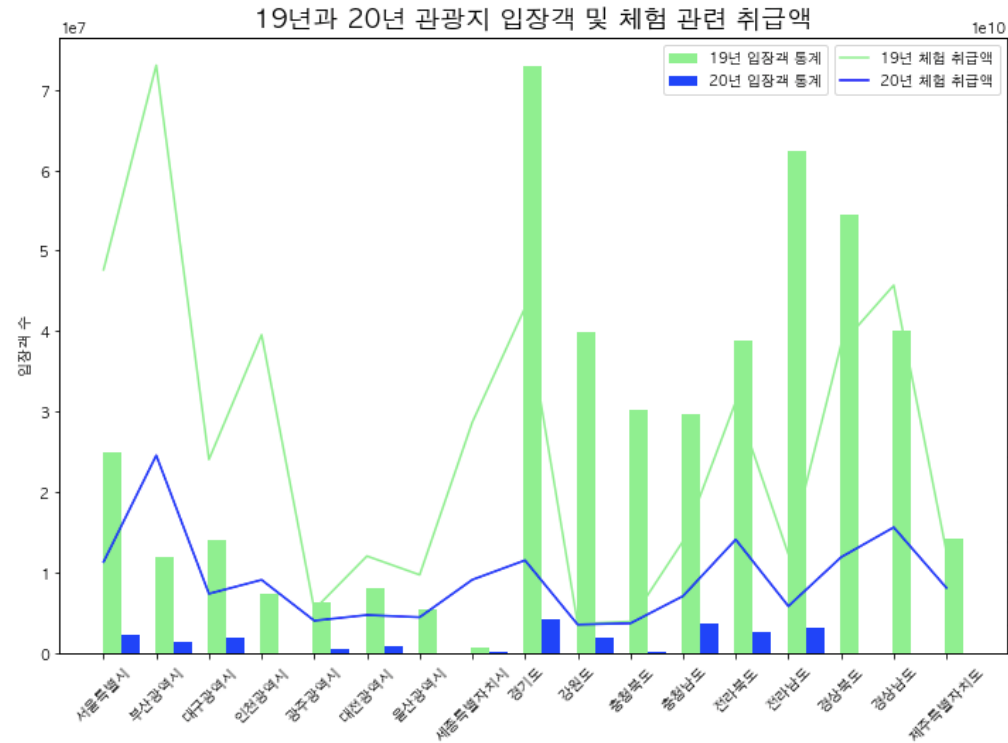
주요관광지점 입장객통계  
체험형 여행지 수  
총 여행 횟수  
총 결제 인원 수  
체험 관련 총 결제 건 수  
체험 관련 총 결제 총액



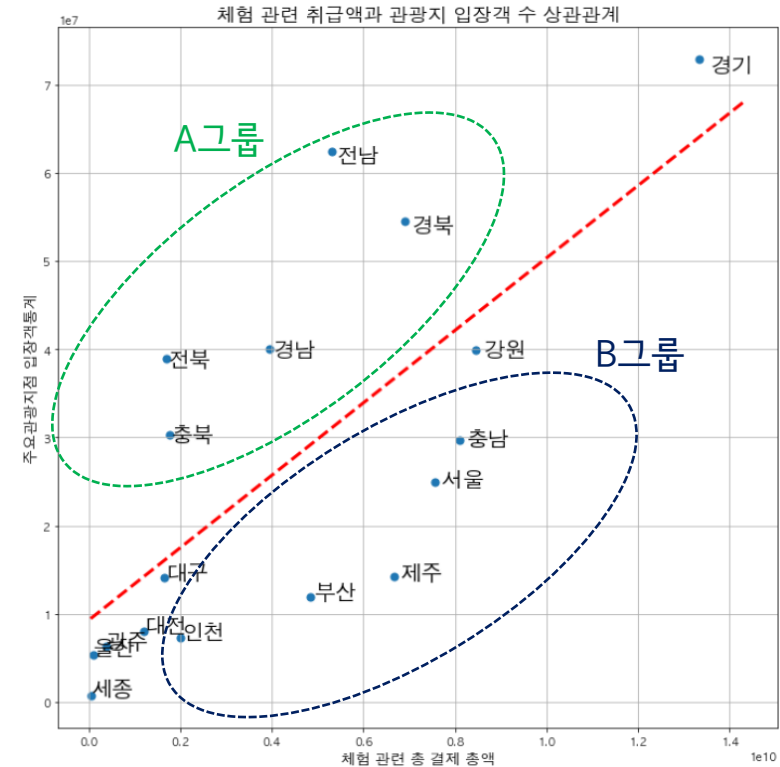
위의 변수들 사이의  
전반적인 선형관계 확인 가능  
(우상향, 비례관계)

## 4. 데이터 분석

### 여행지 & 여행객 데이터 분석



### 체험 관련 취급액과 입장객 수 상관관계 : 회귀선 도출



통계 기간은 반년 정도의 차이지만, 입장객은 작년 전체 대비 최소 6배 이상 차이가 나고 있음.

취급액에 크게 차이를 나타내지 않는 지역 : 광주, 강원도, 충청북도, 제주

- 입장객에 대한 체험 취급액 탄력성이 낮음

취급액 차이가 가장 뚜렷한 지역 : 부산, 서울, 경기, 경남, 경북, 인천

- 입장객에 대한 체험 취급액 탄력성이 높음

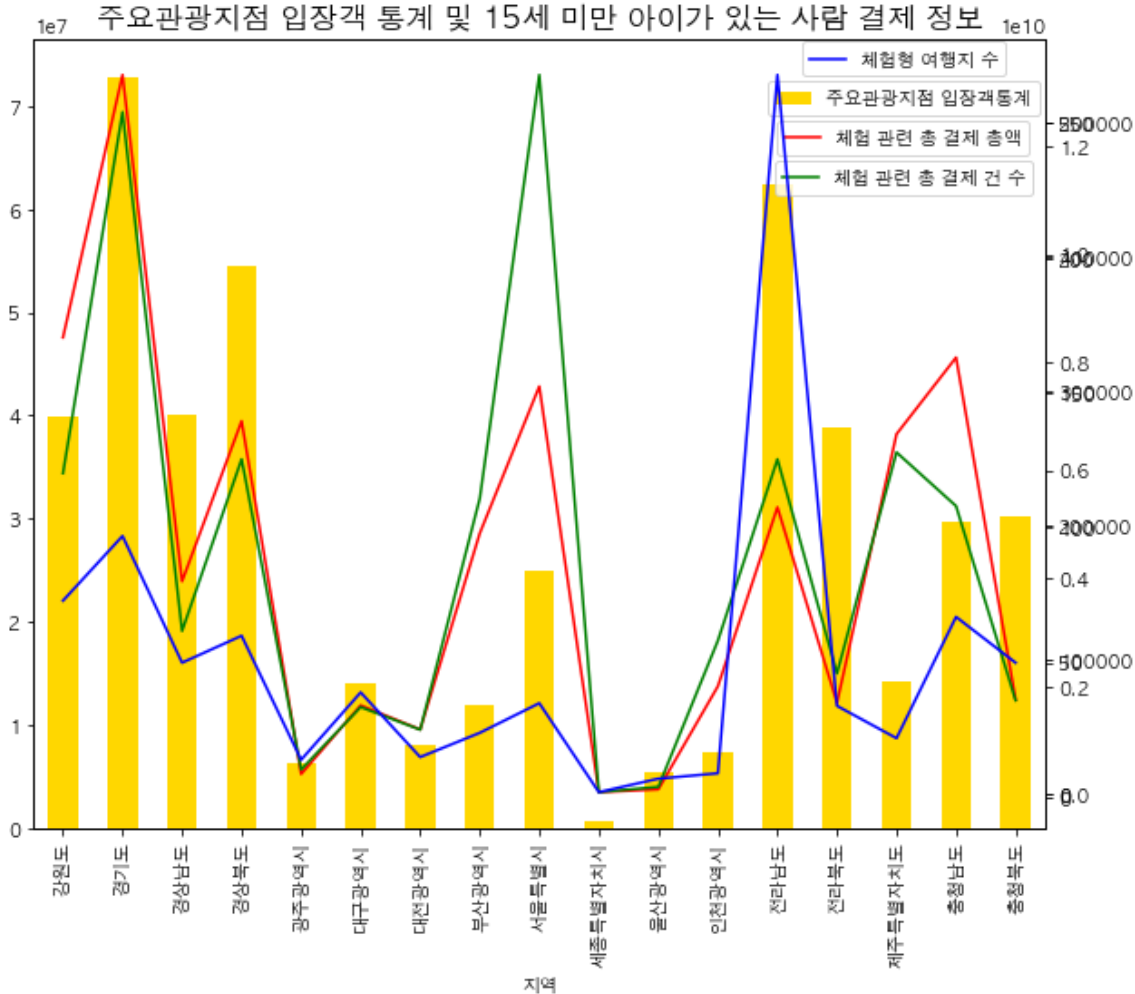
A그룹이 B그룹에 비해 입장객 수 대비 취급액이 적음

A그룹의 지역이 B그룹에 비해 체험 소비 지출 요인이 적음을 유추 할 수 있음

## 4. 데이터 분석

### 여행지 & 여행객 데이터 분석

#### 주요관광지점 입장객 통계 및 15세 미만 아이가 있는 사람 결제 정보



서울은 입장객 대비 체험 관련 취급액 총액과 총 결제 건수가 가장 높은 수준.

체험 관련 여행업이 잘 운영되는 지역  
: 강원도, 경기도, 경상북도, 부산광역시, 인천광역시, 제주특별자치도, 충청남도

체험 관련 여행업 개선 필요 파악이 필요한 지역  
: 광주, 대구, 세종, 전라남도, 전라북도, 충청북도

## 5. 데이터 분석 결과 결론

15세 미만 아이가 있는 사람의 통계치

체험 관련 1인당 평균 결제 건수	
1	제주 (129.46)
2	서울 (94.08)
3	인천 (69.66)
...	
15	전북 (26.9)
16	울산 (13.22)
17	세종 (9.48)

체험 관련 이용 건당 평균 취급액	
1	충남 (37,626)
2	강원 (35,339)
3	경남 (32,331)
...	
15	인천 (17,442)
16	울산 (15,775)
17	서울 (14,090)

체험 관련 1인당 평균 취급액	
1	제주 (3,386,652)
2	강원 (2,432,860)
3	충남 (1,837,958)
...	
15	전북 (504,453)
16	울산 (208,594)
17	세종 (201,301)

체험 관련 결제 평균 연령	
1	제주 (35.58)
2	전북 (35.41)
3	전남 (35.33)
...	
15	세종 (34.43)
16	울산 (34.20)
17	대구 (34.15)

평균 연령을 제외한 통계치에서 하위권 순위에  
**전북 2회, 울산 3회, 세종 2회** 등장 → 해당 지역의 **체험형** 여행지 경쟁력 부재 확인

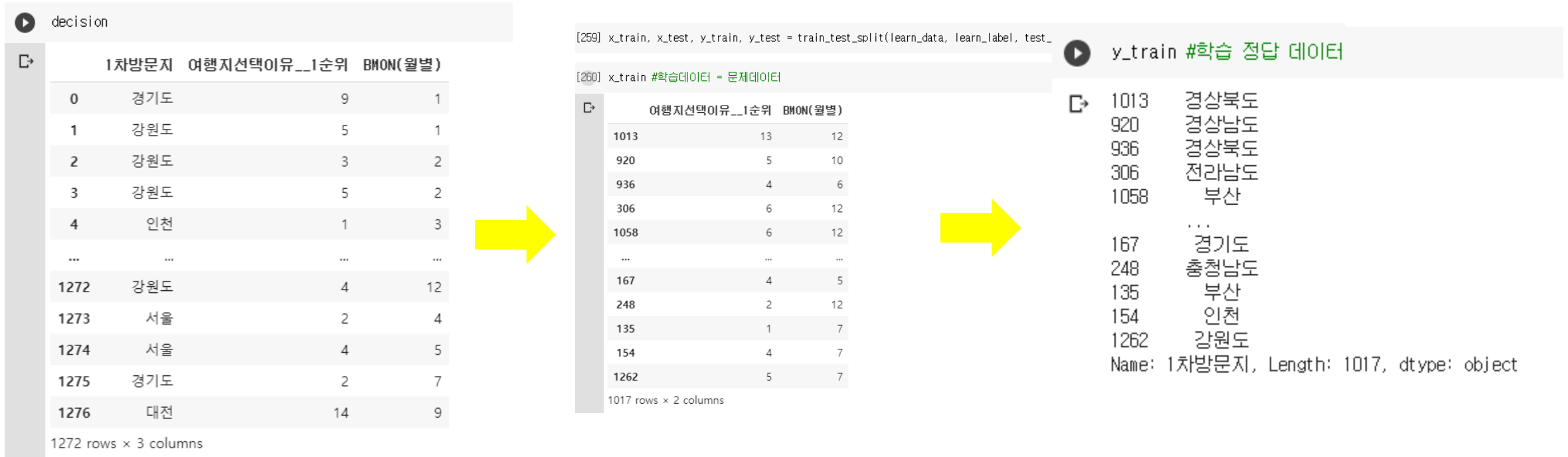




## 5. 데이터 분석 결과

### 모델 작동 방식

**모델** : '시즌성' 과 '여행지 우선순위'를 활용한 지역별 여행 추천 모델링



**설명** : Decision Tree를 활용하여 여행지 선택이유 1순위(ex : 11 = 지역축제) 와 월(ex : 9월) 로 ' 지역 여행지 추천'

인구밀도 과도화를 방지하며, 본인의 여행계획에 맞는 맞춤형 여행 모델링

## 5. 데이터 분석 결과

### 모델 활용

**모델** : '시즌성' 과 '여행지 우선순위'를 활용한 지역별 여행 추천 모델링

#### 여행자 조건 테스트

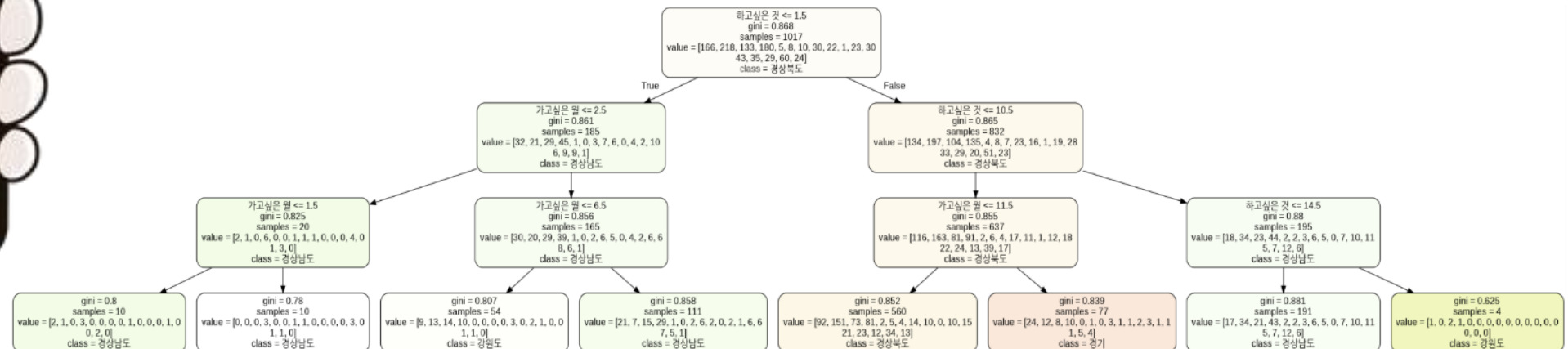
가족여행에 활용

가족 여행지 추천



```
[266] # 여행자 조건 테스트 : 지역축제 = 11번을 목적으로, 나는 9월달에 여행을 떠날것이다.  
jiheon = [11,9]
```

```
[267] tree_clf.predict_proba([jiheon])  
#각 인덱스 별로 맞는것을 정확도로 표현
```



## 6. 결론 및 제언

### 정책제안

### 결론

1. 15세 미만 아이가 있는 가족단위 여행객들의 [체험 관련 1인당 평균 카드 결제 건수], [체험 관련 이용 건당 평균 취급액], [체험 관련 1인당 평균 취급액] 에서 하위 3위 랭킹에 전북 2회, 울산 3회, 세종 2회 등장  
→ 해당 지역의 **체험형 여행지 경쟁력 부재 확인**
2. 충북, 전북, 전남은 전국 체험형 여행지 분포 비중에서는 상위권에 위치 하지만 실제 여행객 수 및 취급액이 낮음.  
→ 충북, 전북, 전남의 체험형 여행지 개선이 필요함



### 정책제안

1. 전북, 울산, 세종, 충북, 전북, 전남 지역의 체험형 여행지의 퀄리티를 높일 수 있도록 가족 여행객들이 선호하는 체험 여행을 조사하여 지역 체험 여행지에 반영 할 수 있도록 한다.
2. 지방 활성화를 위하여 비용을 많이 지불하더라도 효용가치가 높은 여행지를 개발한다.
3. 해당 지역에서 관광객 유치에 기대되는 민간 주도 사업 중 언택트 관련 사업에 지원금을 통해 효율적인 지방 개발이 될 수 있도록 한다.

## 6. 결론 및 제언

### 정책의 활용방안

#### 정부

지역 균등 활성화를 위한 소외 지역 보조  
영향력 있는 관광 관련 콘텐츠 제작 시  
소외 지역 위주의 홍보 활동 진행

#### 지자체

지자체 주도하에 다양한 가족 체험 관광객을  
모집할 수 있는 대책 마련  
뉴노멀 시대에 맞는 민간사업 분야 지원

#### 민간사업

사회적 거리두기 및 만족도를 높일 수 있는  
관광 콘텐츠 개발 및 운영

#### 가족 및 학교

사회적 거리두기가 가능한 여행지로 여행  
계획을 세워 뉴노멀 시대에 안전한 체험학  
습 여행을 할 수 있도록 함

▶ 2020년 이후 뉴노멀 시대의 새로운 체험학습 정책을 통해  
만족도 높은 국내 여행, 지역 균등 개발을 이룰 수 있다.

## 참고자료

1. 한국문화관광연구원 문화.관광 인사이트 제 145호 (2020. 6.16)
2. 뉴스자료  
<http://www.eduinnews.co.kr/news/articleView.html?idxno=31758>  
<https://www.yna.co.kr/view/AKR20200721018800057?input=1195m>  
<http://www.ifm.kr/news/284601>
3. 국민여행조사(엑셀,보고서,코드북) <https://know.tour.go.kr/stat/tourStatSearchDis19Re.do>
4. 계절 테마여행 세부 코스정보 (한국 농어촌 공사) <https://www.data.go.kr/data/15048443/fileData.do>
5. 전국 문화축제 표준데이터 (공공데이터활용지원센터) <https://www.data.go.kr/data/15013104/standard.do>
6. 전국 관광지 정보 표준데이터 (공공데이터활용지원센터) <https://www.data.go.kr/data/15021141/standard.do>
7. 농어촌 체험 휴양마을 현황 (전라남도) <https://www.data.go.kr/data/3068476/fileData.do>

감사합니다 ☺