

강원도 평창군 관광객 데이터 분석

과목명 : AI고급프로그래밍
학번 : 20201512
학과 : 소프트웨어학과
이름 : 함형빈

1. 데이터 전처리

```
data = pd.read_csv(file, encoding='cp949')
for i in range(4, 16):
    data.iloc[:,i] = data.iloc[:,i].astype(str).str.replace(',', '')
    data.iloc[:,i] = data.iloc[:,i].astype(str).str.replace('-', '0')
data.iloc[:,3:16] = data.iloc[:,3:16].apply(pd.to_numeric)
df = data.iloc[:, :-1]
```

* 연산에 필요한 열(4~16)열에 형변환을 해주기 위해 , 와 – 값을 각각 공백, 0으로 대체 했다.

```
df["평균"] = df.iloc[:,4:16].mean(axis=1) # 평균 방문자 수
df['봄'] = df.iloc[:,6:9].sum(axis=1) #봄
df['여름'] = df.iloc[:,9:12].sum(axis=1) #여름
Df[ ' 가을 ' ] = df.iloc[:,12:15].sum(axis=1) #가을
Df[ ' 겨울 ' ] = df.iloc[:,[15, 4, 5]].sum(axis=1) #겨울
Local = df[df.iloc[:,2] == '내국인'] #데이터 나누기
Foreigner = df[df.iloc[:,2] == '외국인']
Local = local.reset_index(drop=True) #인덱스 초기화
Foreigner = foreigner.reset_index(drop=True)
```

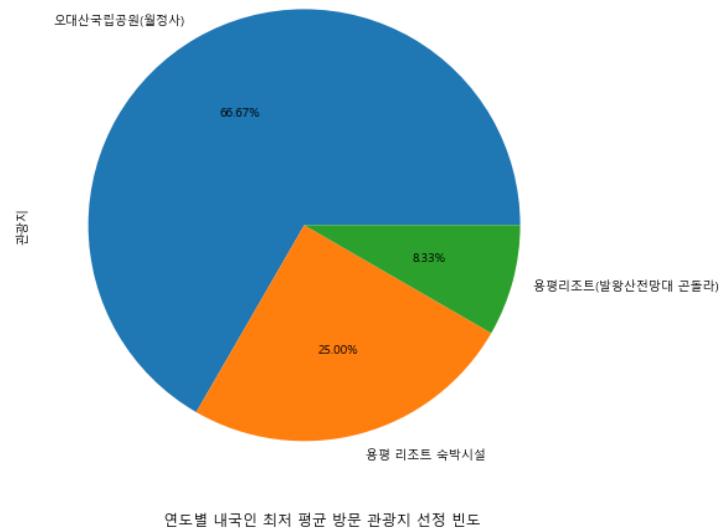
* 관광지 별 매월 평균 방문자 수, 계절별 방문자 수를 추가했다.

* 내국인, 외국인 데이터를 나눠 분석하기로 결정해 각각 local, foreigner 변수에 따로 저장 후 인덱스 초기화를 했다.

2. 데이터 분석

연도별 내국인 최고, 최저 평균 방문 관광지 선정 빈도를 원그래프

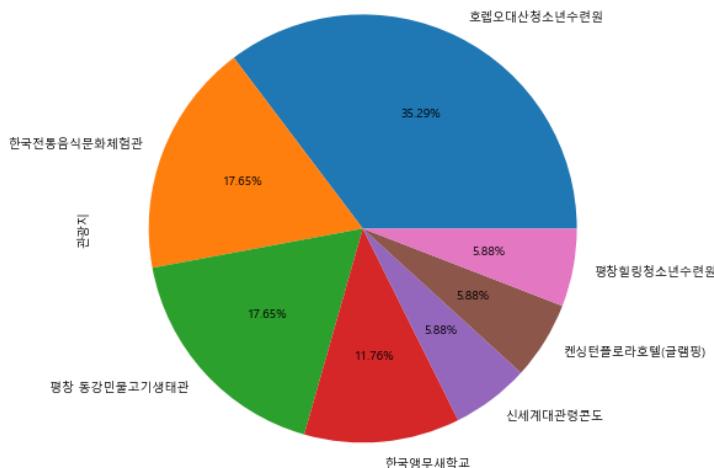
연도별 내국인 최고 평균 방문 관광지 선정 빈도



- 최고 그래프 : 월정사(67%), 용평리조트(25%, 8%)가 차지했다.

오대산국립공원(월정사)	9
용평리조트(피크아일랜드)	9
용평리조트(유스호스텔)	9
한화리조트(컨벤션홀)	9
용평리조트(골프장)	9
알펜시아리조트(오션700)	9
휘닉스파크(워터파크)	9
알펜시아리조트(스키장)	9
용평리조트(발왕산전망대 곤돌라)	9
용평리조트(스키장)	9
켄싱턴플로라호텔(컨벤션홀)	9
용평 리조트 숙박시설	3

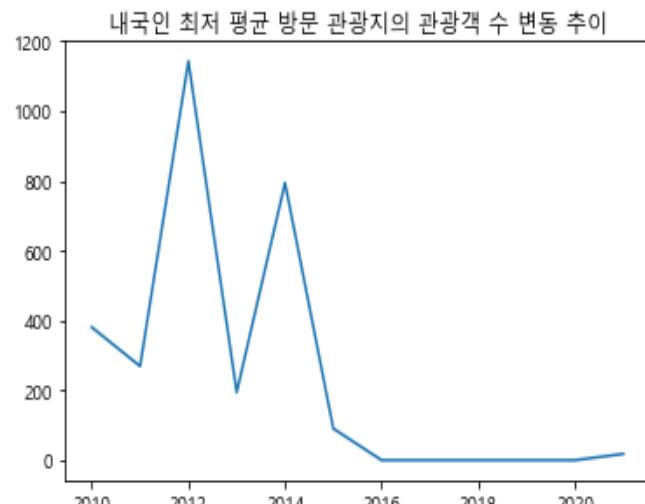
- 집계된 관광지의 빈도 수가 같지 않기 때문에 절대적으로 월정사가 1위라고 할 수 없을 것 같다.



- 최저 그래프 : 평창 동강민물고기생태관(18%), 오대산청소년수련원(35%), 대관령콘도(6%), 플로라호텔(6%), 평창힐링청소년 수련원(6%), 한국앵무새학교(12%), 한국전통음식문화체험관(18%)이 차지했다.
- 청소년수련원이 2곳, 동물관련 관광지, 체험관, 호텔 등 여러 분야 관광지가 선정되었다.

2. 데이터 분석

연도별 내국인 최고, 최저 평균 방문 수 선그래프



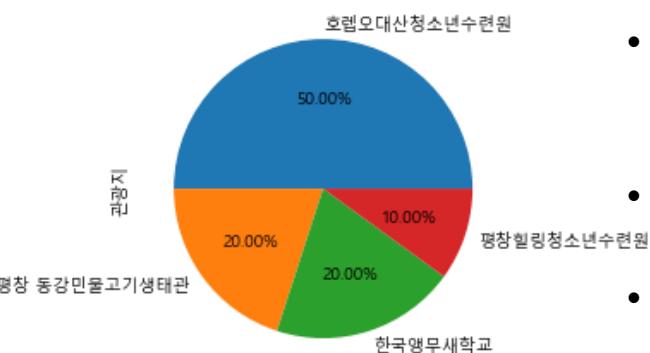
- 최고 그래프 : 2013년도 기준으로 첫 번째 급감하는 추세를 보인다. 또한 2020년도 두 번째 급감을하게 된다.

- 2020년도의 경우 코로나 사태가 있어 감소한 것 같지만 2013년도의 급감 이유를 알아보자.

관광지	유_무료	내국인_외국인	연도	평균
8 용평리조트 숙박시설	유료	내국인	2010	209253.416667
30 용평리조트 숙박시설	유료	내국인	2011	207453.666667
51 용평리조트 숙박시설	유료	내국인	2012	178624.416667
72 오대산국립공원(월정사)	유료	내국인	2013	104935.250000
74 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2013	9188.500000
100 오대산국립공원(월정사)	유료	내국인	2014	96059.583333
102 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2014	16201.750000
127 오대산국립공원(월정사)	유료	내국인	2015	112477.833333
129 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2015	19425.833333
154 오대산국립공원(월정사)	유료	내국인	2016	84835.583333
156 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2016	17852.250000
182 오대산국립공원(월정사)	유료	내국인	2017	111319.583333
184 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2017	19036.833333
210 오대산국립공원(월정사)	유료	내국인	2018	116151.916667
212 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2018	19993.750000
237 오대산국립공원(월정사)	유료	내국인	2019	100595.500000
239 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2019	46443.583333
264 오대산국립공원(월정사)	유료	내국인	2020	50578.833333
266 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2020	35636.416667
291 오대산국립공원(월정사)	유료	내국인	2021	26162.833333
293 용평리조트(발왕산천왕대 곤돌라)	유료	내국인	2021	31068.000000

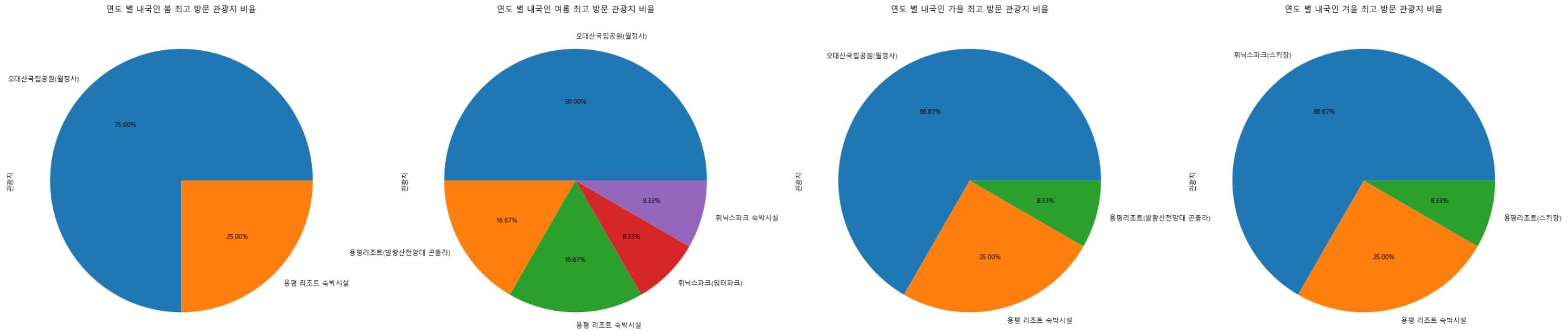
- 최저 그래프 : 2016년도부터 방문을 하지 않는 관광지가 생겼다.

방문이 없는 관광지의 빈도 수



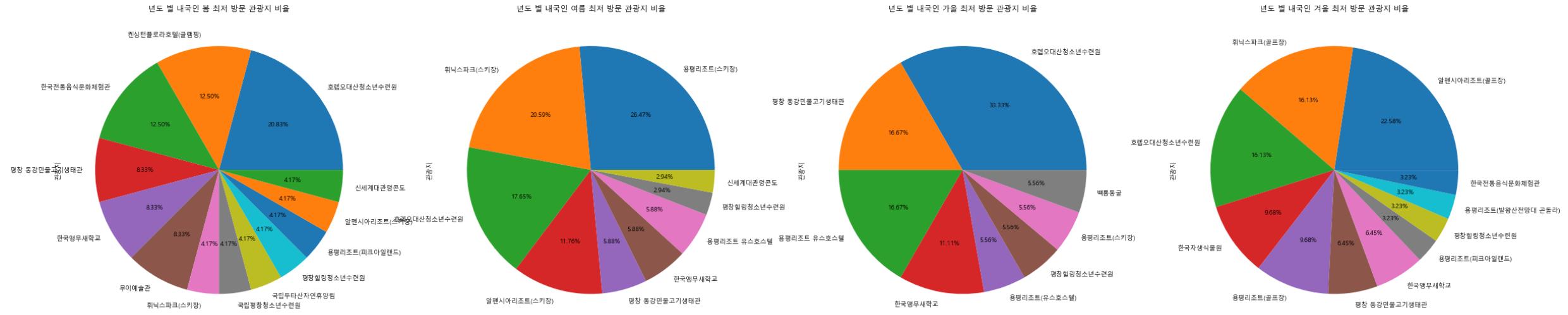
- 방문이 없는 관광지 중 : 오대산청소년수련원은 5년간 평창물고기생태관(2), 앵무새학교(2), 이외 1년 씩 정되었다.
- 오대산청소년수련원의 경우 학교에서 체험으로 가는 수가 줄어들어 그런 것 같다.
- 비인기 관광지에 대한 개편, 홍보가 더 이루어져야 할 것 같다.

2. 데이터 분석 시즌 별 내국인 최고 방문 관광지 비율 원그래프



- 전체적으로 월정사 1순위, 용평리조트 숙박시설이 2순위로 나타난다.
- 여름의 경우 휘닉스파크의 워터파크, 숙박시설이 순위에 들었다. 여름의 더위를 식히기 위해 워터파크를 찾은 관광객이 증가한 이유일 것이다.
- 겨울의 경우 용평리조트의 스키장이 순위에 들었다. 겨울의 묘미인 스키장을 즐기기 위한 관광객 때문일 것이다.
- 봄, 가을의 경우 방학 시즌이 아니다 보니 기존 최고 평균 방문지의 결과와 유사하게 나왔다.

2. 데이터 분석 시즌 별 내국인 최저 방문 관광지 비율 원그래프

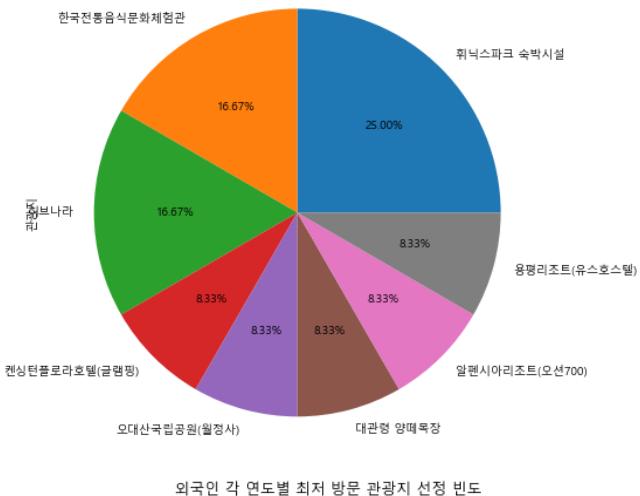


- 전체적으로 월정사 1순위, 용평리조트 숙박시설이 2순위로 나타난다.
 - 여름의 경우 용평리조트, 휘닉스파크, 알펜시아의 스키장이 비시즌이여 선정됐고 이외에 청소년수련원, 앵무새학교, 물고기 생태관 등 최저 평균 관광지와 유사하다.
 - 겨울의 경우 용평리조트, 휘닉스파크, 알펜시아의 골프장이 선정됐고 청소년 수련원, 한국자생식물원, 앵무새 학교 등 최저 평균 관광지와 유사하다.
 - 봄, 가을의 경우 방학 기존 최저 평균 방문지 결과와 유사하게 선정됐다.

2. 데이터 분석

연도별 외국인 최고, 최저 평균 방문 관광지 선정 빈도를 원그래프

외국인 각 연도별 최대 방문 관광지 선정 빈도



외국인 각 연도별 최저 방문 관광지 선정 빈도



- 최고 그래프 : 내국인의 그래프와 전혀 다른 모습을 보여주고 있다.
내국인 최저 그래프에 나온 한국전통음식문화체험관이 외국인 인기 관광지로 선정됐고, 휘닉스파크, 알펜시아, 허브나라 등 내국인 인기 관광지보다 다양하게 선정되는 모습이다.
 - 유일하게 월정사가 내국인 최대 방문 관광지의 교집합으로 선정됐다.

```
for i in foreigner.i.loc[f_top, :][['관광지']].unique():
    if i in local.i.loc[l_top, :][['관광지']].unique():
        print(i)
```

모대산국립공원(월정사)

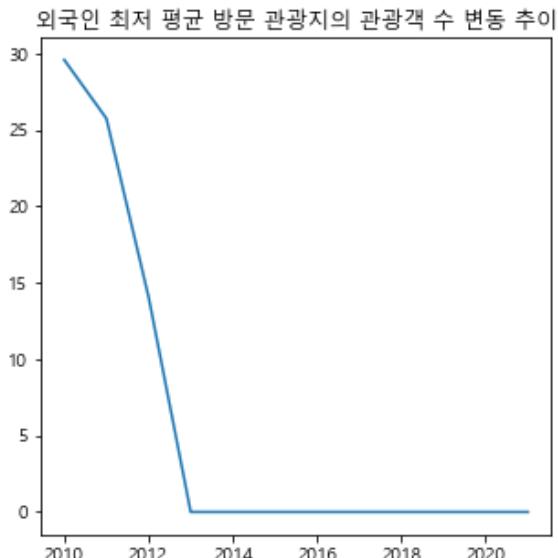
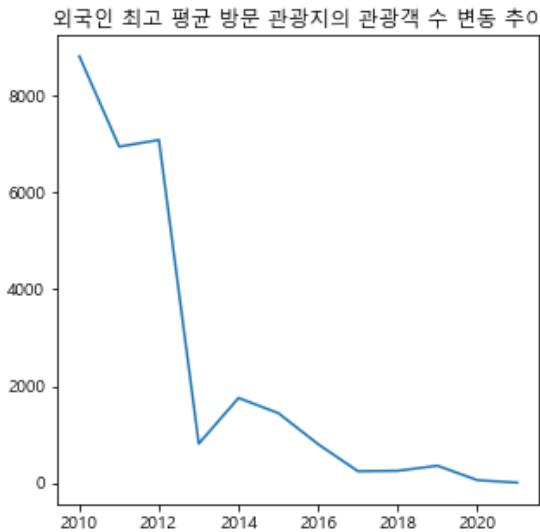
- 최저 그래프 : 총 관광지 수의 70% 이상이 선정됐다.
 - 외국인 관광객을 늘리기 위해 홍보와 관광지 프로그램 개선이 필요할 것 같다.

```
print("외국인 방문 총 관광지 수 : ", len(foreigner['관광지'].unique()))
print("외국인 최고 평균 방문 관광지 수 : ", len(foreigner.iloc[f_top, :]['관광지'].unique()))
print("외국인 최저 평균 방문 관광지 수 : ", len(foreigner.iloc[f_bottom, :]['관광지'].unique()))
```

외국인 방문 총 관광지 수 : 38
외국인 최고 평균 방문 관광지 수 : 8
외국인 최저 평균 방문 관광지 수 : 29

2. 데이터 분석

연도별 외국인 최고, 최저 평균 방문 수 선그래프

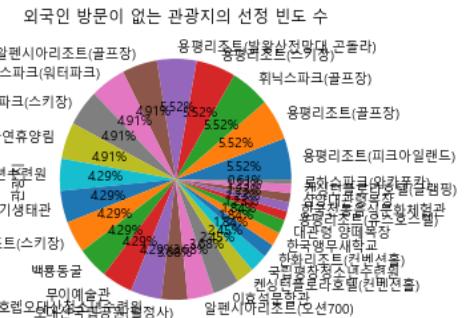


- 최고 그래프 : 2013년도 기준으로 급감하는 추세를 보이고 이후 외국인의 방문이 현격히 줄어든 모습이다.

32	휘닉스파크 속박시설	유료	외국인	2012	7087.333333
35	대관령 양떼목장	유료	외국인	2013	360.750000
38	알펜시아리조트(오션700)	유료	외국인	2013	0.000000
39	오대산국립공원(월정사)	유료	외국인	2013	0.000000
43	용평리조트(유스호스텔)	유료	외국인	2013	3.333333
45	캔성턴플로라호텔(글램핑)	유료	외국인	2013	0.000000
47	한국전통음식문화체험관	유료	외국인	2013	192.666667
49	허브나라	유료	외국인	2013	812.916667

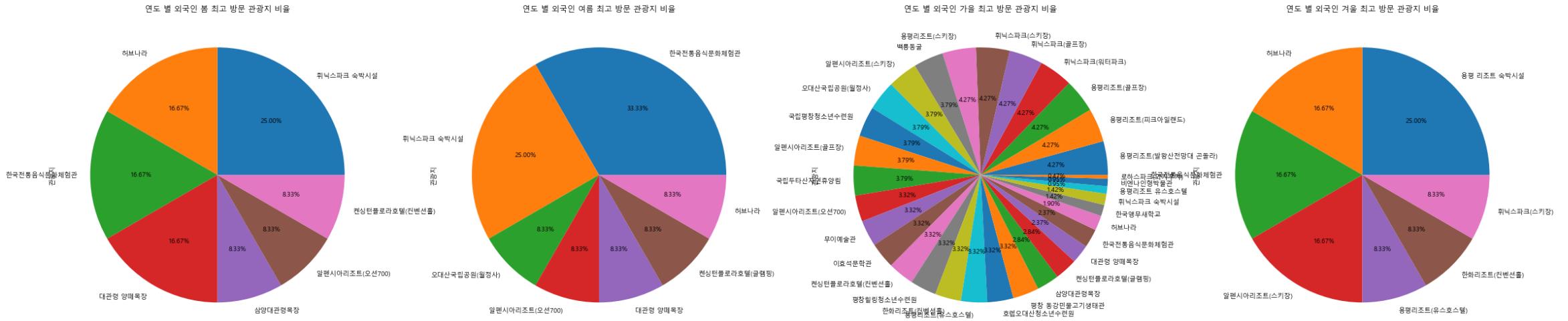
- 2012년까진 휘닉스파크 숙박시설이 집계가 되었지만 2013년 이후부터 집계되지 않았다.
- 내국인 집계와 비슷한 추세를 보이고 만약 휘닉스파크 숙박시설이 집계되었다면 결과가 달라졌을 것 같다.

- 최저 그래프 : 2016년도부터 방문을 하지 않는 관광지가 생겼다.



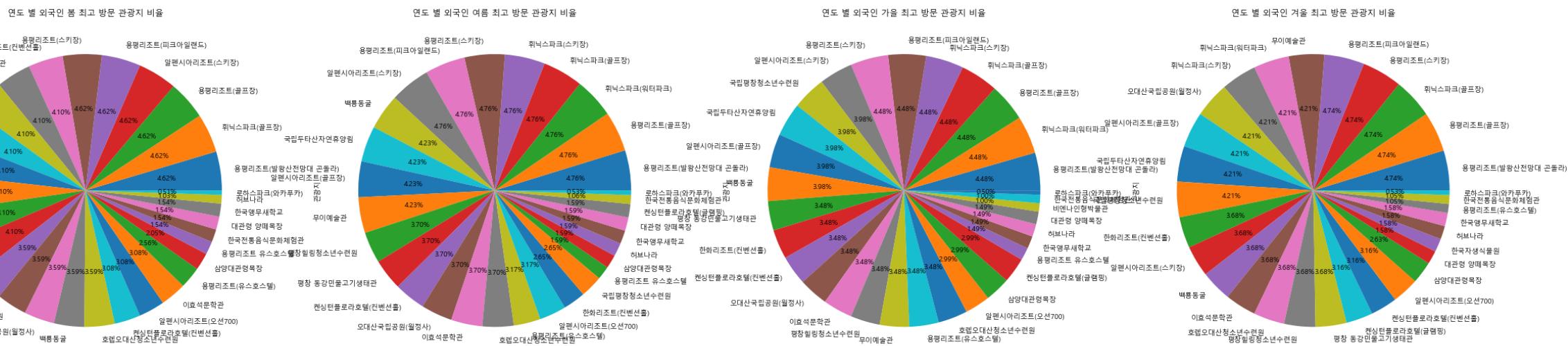
- 방문이 없는 관광지 중 : 오대산청소년수련원은 5년간 평창물고기생태관(2), 앵무새학교(2), 이외 1년 씩 선정되었다.
- 오대산청소년수련원의 경우 학교에서 체험으로 가는 수가 줄어들어 그런 것 같다.
- 비인기 관광지에 대한 개편, 홍보가 더 이루어져야 할 것 같다.

2. 데이터 분석 시즌 별 외국인 최고 방문 관광지 비율 원그래프



- 전체적으로 한국전통음식문화체험관, 허브나라, 휘닉스파크 숙박시설는 꾸준하게 순위에 든 모습이다.
- 여름의 경우 휘닉스파크 숙박시설, 한국전통음식문화체험, 오션700이 순위에 들었는데 여름의 묘미인 워터파크를 찾은 관광객이 증가했다.
- 겨울의 경우 용평리조트의 스키장이 순위에 들었고 겨울의 묘미인 스키장을 즐기기 위한 관광객이 증가했다.
- 봄의 경우 기존 최고 평균 방문지의 결과와 유사하게 나왔다.
- 가을은 다양한 관광지가 선정되었다.

2. 데이터 분석 시즌 별 외국인 최저 방문 관광지 비율 원그래프



- 전체적으로 최저 평균 그래프에 선정된 장소가 선정됐다.
 - 외국인 최저 방문 관광지는 시즌에 상관없이 대부분의 관광지가 선정됐다.
 - 외국인에 대한 관광지 홍보가 필요할 것 같다.

3. 결론

- 용평리조트, 월정사가 내국인에게 인기를 끌고 있다. 하지만 데이터 셋에 용평리조트가 연도 마다 존재하지 않아 절대적이지 않아 2013년 이후 방문 수가 급감한다.
- 동물 관련 관광지, 청소년 수련관, 한국전통음식문화체험관 등 내국인에게 인기가 없다.
- 한국전통음식문화체험관은 내국인에게 인기 없는 관광지이지만 외국인이 많이 찾는 관광지이다.
- 외국인은 허브나라, 한국전통문화체험관, 휘닉스파크, 등 다양한 관광지를 찾지만 내국인의 경우 월정사, 용평리조트가 큰 비중을 차지 한 것 같다.
- 내국인, 외국인 모두 여름에 워터파크 관광지를 많이 찾고, 겨울엔 스키장을 많이 찾는다.
- 내, 외국인의 봄에는 평균적으로 인기있는 관광지를 많이 찾는다.
- 내국인의 가을은 평균적으로 인기있는 관광지를 찾는 반면 외국인은 인기, 비인기 관광지 모두 선정되었는데 이는 외국인이 가을에 많이 관광을 하러 오지 않는다는 뜻이 될 수 있다.