

제주도 관광객 소비 패턴 분석 프로젝트



목차

1. 프로젝트 개요

- 프로젝트 기획 배경 및 목표
- 구성원 및 역할

2. 프로세싱

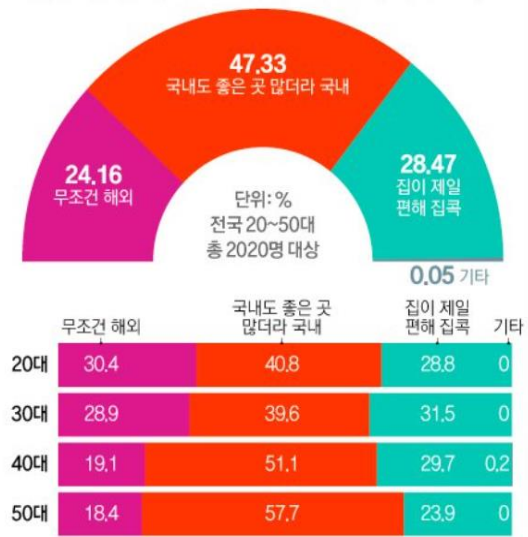
- 데이터 수집 / 전처리
- 데이터 분석
- 분석 결론

3. 종합 인사이트

4. 프로젝트 후기

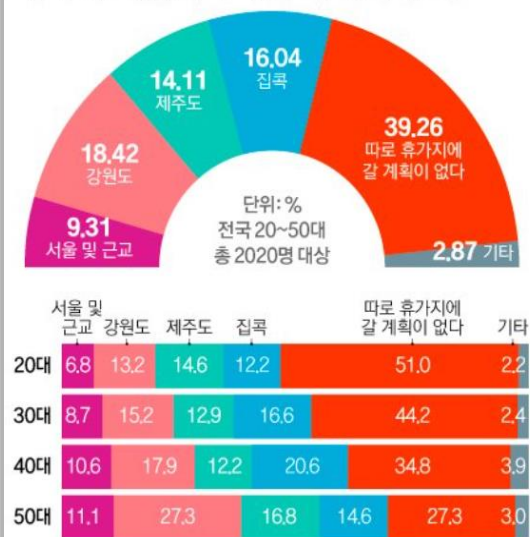
프로젝트 배경

내년 이맘때 코로나19가 극복됐다는 가정 하에 어디로 떠나고 싶으신가요?



자료: 중앙일보·SMC&C 톨리언 프로 공동 설문 조사 The JoongAng

올 여름 휴가지 계획하고 있는 지역은 어디인가요?



자료: 중앙일보·SMC&C 톨리언 프로 공동 설문 조사 The JoongAng

보복여행 몰린 제주, 카드 소비 사상최대

입력 2021-06-28 17:24:12 수정 2021.06.28 17:24:12 이태규 기자

제주로 몰리는 관광객...7~8월 렌터카 '품귀' 현상

파이낸셜뉴스 입력 2021.06.27 16:02 수정 2021.06.27 16:02

‘제주는 북적북적’ 관광객 3달 연속 100만여명...코로나19 확산 우려도 고조

박미라 기자 입력 : 2021.07.04 15:29



프로젝트 목표

- 빅데이터 분석으로 제주도 창업자들에게 인사이트를 제공한다.
- 제주도는 대한민국의 관광명소로 유동인구가 많은 지역이기 때문에 많은 사람들이 창업을 준비하고 있다. 하지만 여러 매체에서 창업에 관한 정보가 부족하다고 알고 있다.
- 이러한 문제를 해결하기 위해 제주도 관광객의 소비패턴을 분석하여 예비 창업자들에게 위치, 업종, 타겟고객, 마케팅 전략 등 데이터에 기반한 창업 전략을 제시한다.
- 실제로 빅데이터 분석이 다양한 곳에서 도움되는 만큼 이번 프로젝트는 제주도 창업을 보다 성공적으로 이끌 것으로 기대된다.



구성원 및 역할

김수진 : 데이터 전처리 , 분석, 개인 ppt 부분 작성 , 발표문 작성 , 코드리뷰 , 계획서 작성

조아진 : 데이터 전처리 , 분석 , 개인 ppt 부분 작성 , 발표문 작성

김하연 : 데이터 전처리 , 분석 , 개인 ppt , 디자인 , 발표문 작성

이동준 : 데이터 전처리 , 분석 , 개인 ppt, ppt 레이아웃 작성 , 발표문 작성

장은준 : 데이터 전처리 , 분석 , 개인 ppt , 발표 , 발표문 작성

데이터 수집 / 전처리

1. 데이터 수집

제주여비창업자를 위한 업종별카드매출과매장수데이터활용.csv

(출처: 공공데이터 포털 www.data.go.kr)

	년월	시도명	지역구분	읍면동명	업종코드	업종명	이용자 구분	관광구분	연령대	성별	이용금액	매장수	업종명 대분류	데이터기준일자
0	2017-01	서귀포시	도심	강정동	G47121	슈퍼마켓	제주도민	도민	50.0	여성	365160	214	소매업	2020-12-15
1	2017-01	제주시	도심	삼도일동	I55112	여관업	제주도민	도민	60.0	여성	40000	1610	숙박업	2020-12-15
2	2017-01	제주시	도심	삼양이동	I55112	여관업	제주도민	도민	20.0	여성	162000	1610	숙박업	2020-12-15
3	2017-01	제주시	도심	삼양이동	I55112	여관업	제주도민	도민	50.0	남성	70000	1610	숙박업	2020-12-15
4	2017-01	제주시	도심	삼도이동	I55112	여관업	내국인	개별	40.0	남성	3150460	1610	숙박업	2020-12-15

2. 데이터 정의

변수명	설명	비고
년월	2017-01~2020-04	
시도명	제주시, 서귀포시 중 1	
지역구분	도심, 읍면 중 1	
읍면동명	73개 읍면동 중 1	
업종명 대분류	소매업, 숙박업, 음식점업, 임대업, 서비스업, 운송업 중 1	결측치 있음
업종명	42개 업종명 중 1	
이용자 구분	제주도민, 내국인, 중국, 법인, 동남아, 일본, 기타외국 중 1	
관광구분	도민, 개별, 단체	결측치 있음
연령대	10~70, 10 단위	결측치 있음
성별	여성, 남성, 알수없음 중 1	결측치 있음
이용금액	10~3102095316	
매장수	1~7030	년월별 데이터

3. 데이터 전처리

I. 컬럼 추가

i. 연도, 월, 분기 추가

II. 사용하지 않는 컬럼 삭제

i. 년월, 업종코드, 데이터기준일자 삭제

III. 결측치 처리

i. 연령대 결측치 - 채워지지 않음

ii. 성별 결측치 - NaN으로 채움

iii. 업종명대분류 결측치 - 운송업으로 채움

4. 전처리 결과

	연도	월	분기	시도명	지역구분	읍면동명	업종명 대분류	업종명	이용자 구분	관광구분	연령대	성별	이용금액	매장수
0	2017	1	1	서귀포시	도심	강정동	소매업	슈퍼마켓	제주도민	도민	50.0	여성	365160	214
1	2017	1	1	제주시	도심	삼도일동	숙박업	여관업	제주도민	도민	60.0	여성	40000	1610
2	2017	1	1	제주시	도심	삼양이동	숙박업	여관업	제주도민	도민	20.0	여성	162000	1610
3	2017	1	1	제주시	도심	삼양이동	숙박업	여관업	제주도민	도민	50.0	남성	70000	1610
4	2017	1	1	제주시	도심	삼도이동	숙박업	여관업	내국인	개별	40.0	남성	3150460	1610

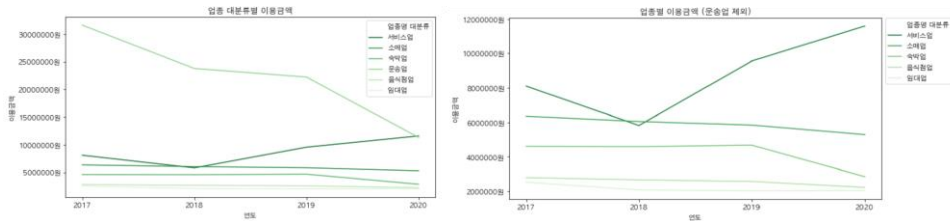
데이터 분석

I. 업종명대분류별 이용금액 변화

소매업, 숙박업, 음식점업, 임대업의 경우 이용금액이 감소하는 추세를 보인다.

운송업의 경우 이용금액이 가장 높으며 2017년 이후 급격히 줄어드는 추세를 보이며,

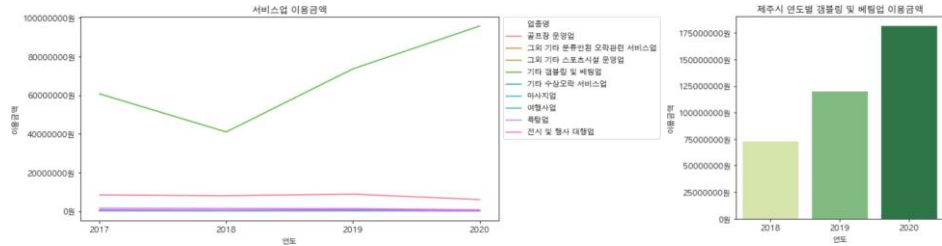
서비스업이 유일하게 점점 증가하는 추세를 보인다.



II. 서비스업 이용금액 변화

서비스업은 업종분류 중 유일하게 이용 금액이 계속해서 증가하는 업종이다. 이에 서비스업을

더 자세히 들여다 본 결과 **게임 및 베팅업**에서 뚜렷한 증가세를 보인다.



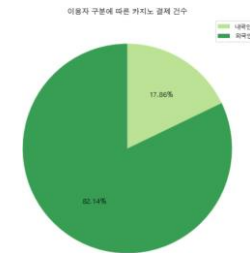
III. 게임 및 베팅업 위치 분석

서비스업은 업종분류 중 유일하게 이용 금액이 계속해서 증가하는 업종이다. 이에 서비스업을 더 자세히 들여다 본 결과 **게임 및 베팅업**에서 뚜렷한 증가세를 보인다.



IV. 게임 및 베팅업 이용자 분석

게임 및 베팅업의 이용 내역을 분석한 결과, 외국인 결제 비율이 82%로 높게 나타났다.



분석 결론

6. 분석을 통한 인사이트 도출

I. 복합리조트

복합리조트란 숙박, 식음료, 쇼핑시설 및 카지노 등 다양한 부대시설을 갖춘 복합시설을 이룬다. 한 곳에서 먹고 쉬고 즐길 수 있는 콘텐츠를 제공한다는 점에서 관광객들이 많이 찾는다. 그 중에서도 카지노는 리조트 공공 면적 중 5% 이하의 규모이지만 리조트 매출의 7~80%를 차지할 정도로 매출 비중이 높다.

I. 제주도 카지노 사업 전망

데이터 분석 결과 제주도의 서비스업 매출액이 꾸준히 증가하고 있다. 그 중에서도 캠핑 및 배팅업 업종에서 꾸준히 증가하고 있다.

I. 위치 선정

제주도 POI 데이터 분석 결과 제주도 카지노 업종은 제주시 연동, 삼도아동, 그리고 서귀포시 색달동에 집중적으로 분포되어 있다. 이러한 위치는 유동 인구가 높으며, 호텔업이 특히 발달한 지역으로 보인다. 따라서 카지노업을 시작한다고 한다면, 호텔이 많은 제주시 조천읍, 서귀포시 서귀동 등을 추천한다.

7. 향후 개선 사항

I. 코로나의 영향

제주의 카지노 업종 이용금액을 살펴보면 2020년 초 코로나 바이러스가 처음 한국에서 발견된 직후에도 계속해서 증가한다고 나타난다. 이는 제주도

에 관광 온 외국인 관광객들이 관광호텔 등에 머물러 있으면서 사용한 금액 인 것으로 보인다. 본 프로젝트에서 사용한 데이터는 2017년 1월부터 2020년 4월까지의 데이터로, 코로나 초기 데이터를 포함하지만 그 이후의 데이터는 누락되어 있다. 따라서 코로나 바이러스의 영향은 분석에 반영되어 있지 않다고 볼 수 있다.

II. 외국인 관광객에 대한 정보 부족

- 본 분석에 사용된 데이터에는 이용자의 국적, 성별, 연령대 정보가 들어 있다. 그러나 외국인에 대해서는 성별 및 연령대 데이터가 누락되어 있다. 이 에 외국인 이용객에 대한 더욱 자세한 분석은 진행하지 못했다. 외국인이 용객에 대한 데이터를 수집하면 외국인 관광객을 상대로 마케팅 전략을 세울 수 있을 듯 하다.

8. 기대 효과

I. 카지노계의 블루오션에서 매출 싹쓸이

- 본 프로젝트에서는 매출이 계속해서 증가하는 캠핑 및 배팅업 업종이 전 망이 좋다고 판단하며 해당 업종을 추천하였다. 위치는 기존에 관광호텔이 많지만, 카지노업이 활성화되어 있지는 않은 제주시 조천읍, 서귀포시 서귀동 등을 창업 위치로 선정하였다. 해당 위치의 관광객을 끌어옴으로써 상당한 매출 올릴 수 있을 것이라 전망된다.

데이터 수집 / 전처리

〈전처리 전 데이터〉

```
df_BC=pd.read_csv('C:\study\jeju_4.csv', encoding='cp949')
df_BC.head()
```

	기준년월	관광객 유형	제주 대분류	제주 중분류	업종명	성별	연령대별	카드이용금액	카드이용건수	건당이용금액	데이터기준일자
0	2014-09	내국인 관광객	서귀포시	대륜동	농축수산물	여	50대	14434000	67	215433	2017-02-16
1	2014-09	내국인 관광객	서귀포시	대륜동	농축수산물	남	50대	15119000	72	209986	2017-02-16
2	2014-09	내국인 관광객	서귀포시	대륜동	농축수산물	여	40대	7609500	41	185598	2017-02-16
3	2014-09	내국인 관광객	서귀포시	남원읍	농축수산물	남	50대	7092500	40	177313	2017-02-16
4	2014-09	내국인 관광객	서귀포시	대륜동	농축수산물	남	40대	9098500	59	154212	2017-02-16



〈전처리 후 데이터〉

df_BC5

	연도	월	분기	계절	시도명	읍면동명	업종명	업종분류	성별	연령대	연령대2	카드이용금액	카드이용건수	건당이용금액
0	2014	9	3	가을	서귀포시	대륜동	농축수산물	농축수산물	여	50대	50	14434000	67	215433
1	2014	9	3	가을	서귀포시	대륜동	농축수산물	농축수산물	남	50대	50	15119000	72	209986
2	2014	9	3	가을	서귀포시	대륜동	농축수산물	농축수산물	여	40대	40	7609500	41	185598
3	2014	9	3	가을	서귀포시	남원읍	농축수산물	농축수산물	남	50대	50	7092500	40	177313
4	2014	9	3	가을	서귀포시	대륜동	농축수산물	농축수산물	남	40대	40	9098500	59	154212
...
13141	2016	8	3	여름	서귀포시	남원읍	약국	의약/의료	남	30대	30	252900	43	5881
13142	2016	8	3	여름	서귀포시	대륜동	약국	의약/의료	남	20대	20	276100	49	5635
13143	2016	8	3	여름	서귀포시	대륜동	기타음료식품	식재료유통	남	30대	30	610850	118	5177
13144	2016	8	3	여름	서귀포시	대륜동	기타음료식품	식재료유통	여	20대	20	416900	85	4905
13145	2016	8	3	여름	서귀포시	대륜동	기타음료식품	식재료유통	여	50대	50	56200	12	4683

13146 rows × 14 columns

• 데이터 출처

- 제주특별자치도 개별관광(FIT) 증가에 따른 제주 관광객 소비패턴 변화 분석 BC 카드 빅데이터 내국인관광객.csv
- 공공데이터 포털(<https://www.data.go.kr>)

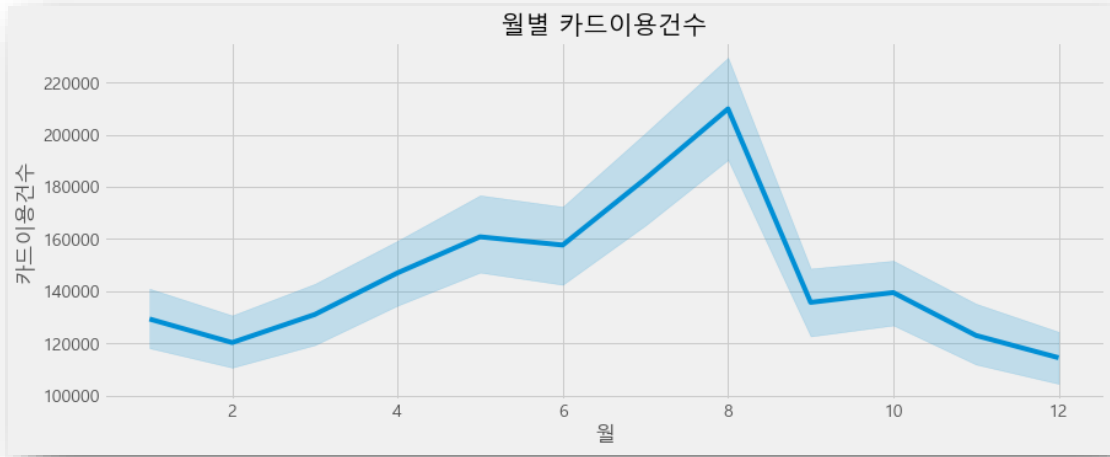
• 데이터 설명

- 2014~2016년도 제주도 내 내국인 관광객의 업종, 성별, 연령대별 카드이용금액, 카드이용건수, 건당이용금액

• 전처리 과정

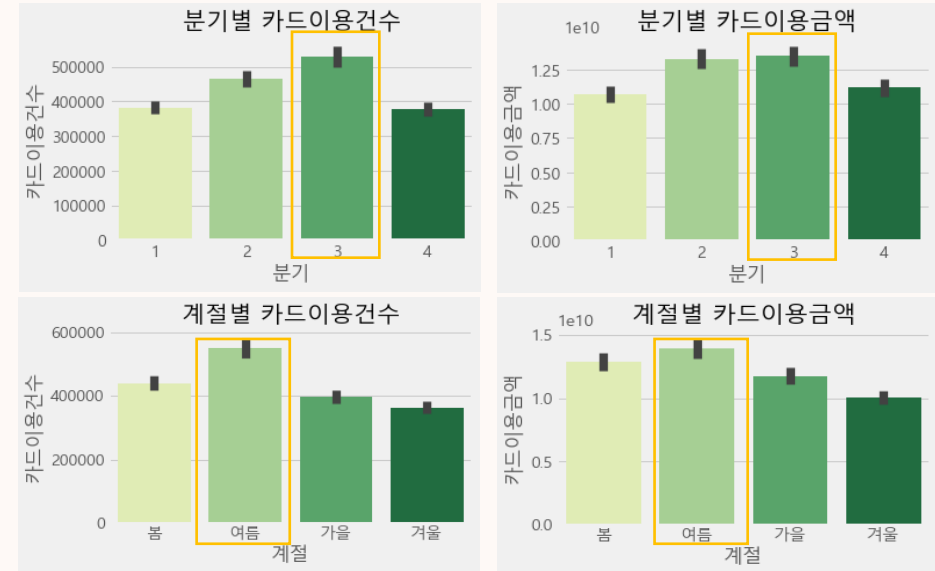
- 1) 변수 정보(Dtype) 확인, 데이터타입 알맞게 변경
- 2) 결측치 확인 및 제거(isnull() 메소드 활용)
- 3) 칼럼 분리, 추가, 제거
 - 기준년월 -> 연도, 월 분리, 분기, 계절 추가
 - 업종명 ->업종분류 추가
 - 칼럼명 변경(제주 대분류->시도명,제주 중분류->읍면동명)
 - 불필요한 칼럼 삭제 (관광객 유형, 데이터기준일자)
- 4) 값에 있는 불필요한 공백 제거

데이터 분석



제주도 관광객 집중 시기
(시각화 : seaborn lineplot, barplot)

- 여름 카드이용건수 약 55만 건 > 겨울 카드이용건수 약 36만 건
- 여름 카드이용금액 약 140억원 > 겨울 카드이용금액 약 100억원
- 7,8월 여름휴가철에 타 월 대비 카드이용건수가 높게 나타남.

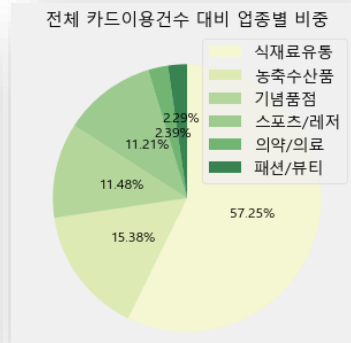
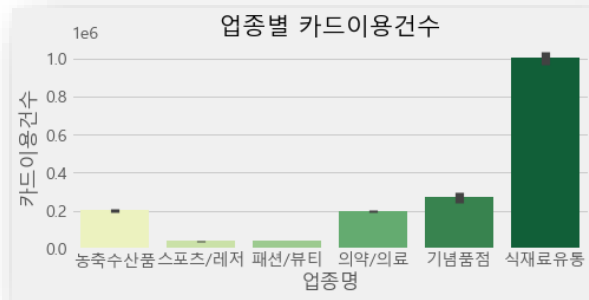


- '월' 기준 : 7,8월
- '계절' 기준 : 여름(6-8월)
- '분기' 기준 : 3분기(7-9월)
- 성수기/비수기 시기별 맞춤형 판매 전략 필요
 - 성수기 : 집중 홍보 기간으로 삼아 매출을 극대화하는 광고 전략 사용
 - 비수기 : 가격 할인 이벤트 등 실시, 수요 창출 및 고객 유입 효과를 기대

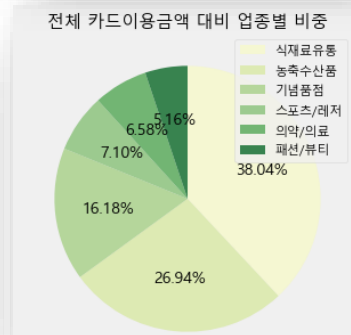
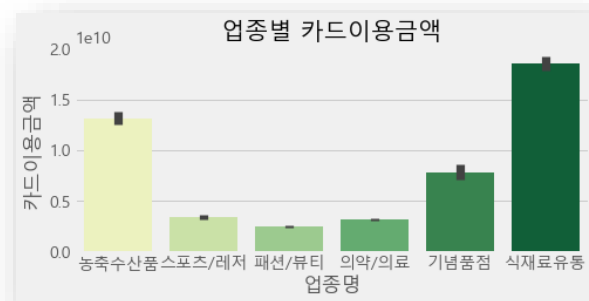
데이터 분석

추천 업종

*카드이용건수



*카드이용금액



(시각화 : matplotlib pie chart, seaborn barplot)

연도를 종합한 업종별 카드이용건수를 보면,

‘식재료유통’의 카드이용건수가 1,003,354건으로 가장 큰 비중(57.25%)을 차지하였고, 다음으로 ‘기념품점’이 269,606건, 15.38%, ‘농축수산물’이 201,118건, 11.48%, ‘의약/의료’ 196,500건, 11.21%, ‘패션/뷰티’ 41,888건, 2.39%, ‘스포츠/레저’ 40,002건, 2.28% 순으로 나타남.

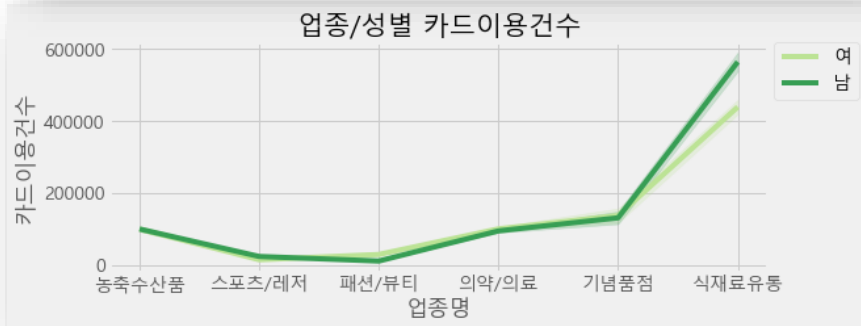
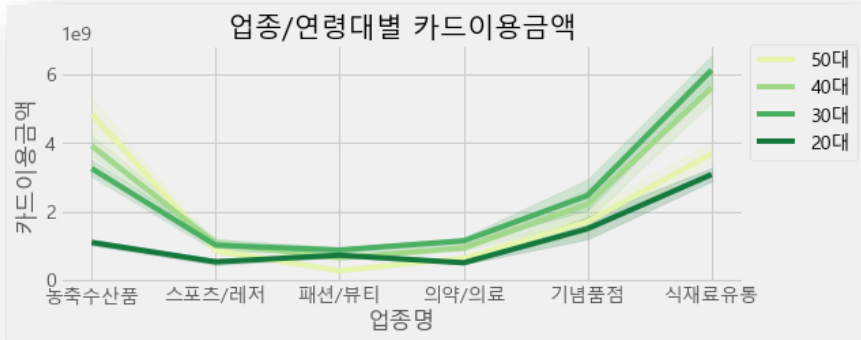
- [건수기준] 1순위 식재료유통, 2순위 기념품점, 3순위 농축수산물

연도를 종합한 업종별 카드이용금액을 보면,

‘식재료유통’의 카드이용금액이 18,489,834,266원으로 가장 큰 비중(38.04%)을 차지하였고, 다음으로 ‘농축수산물’이 13,092,560,386원, 26.94%, ‘기념품점’이 7,864,854,995원, 16.18%, ‘스포츠레저’가 3,450,652,118원, 7.10%, 의약/의료 3,199,275,729원, 6.58%, 패션/뷰티 2,506,003,780원, 5.16% 순으로 나타남.

- [금액기준] 1순위 식재료유통, 2순위 농축수산물, 3순위 기념품점

데이터 분석



타깃 고객

(시각화 : seaborn lineplot)

식재료유통, 농축수산업, 기념품점 TOP3 업종 창업 시,
타깃 고객의 연령대와 성별에 대한 정보

농축수산물

- 성별로 비교하면 남, 여 간에 큰 차이가 나타나지 않음. 성별의 영향력 작음.
- 50대에서 농축수산업 카드이용금액이 가장 높게 나타남.
- 다음으로 40대, 30대, 20대 순
- 50대, 40대를 타깃 고객으로 추천함.

식재료유통

- 성별로 비교하면 남성고객이 더 많음.
- 40대에서 슈퍼마켓 카드이용금액이 가장 높게 나타남.
- 다음으로 30대, 50대, 20대 순
- 40대, 30대를 타깃 고객으로 추천함.

기념품점

- 성별로 비교하면 남, 여 간에 큰 차이가 나타나지 않음. 성별의 영향력은 작음.
- 30대에서 기념품점 카드이용금액이 가장 높게 나타남.
- 다음으로 40대, 50대, 20대 순
- 30대, 40대를 타깃 고객으로 추천함.

데이터 수집 / 전처리

- 파일 출처: 제주데이터허브 (<https://www.jejudatahub.net/>)

- 파일 이름: 제주 관광을 위한 날씨 별 지역 추천 데이터.csv

- 원본 파일:

	일자	시도명	읍면동명	평균 기온	일강수량	최대 풍속	업종명	이용금액	데이터기준일자
0	2018-01-01	제주시	추자면	5.7	0.0	1.8	과실 및 채소 소매업	13569170	2020-12-15
1	2018-01-01	제주시	이도1동	2.9	0.0	2.4	종식 음식점업	18299000	2020-12-15
2	2018-01-01	제주시	이도1동	2.9	0.0	2.4	전시 및 행사 대행업	819550	2020-12-15
3	2018-01-01	제주시	이도1동	2.9	0.0	2.4	일식 음식점업	32160000	2020-12-15
4	2018-01-01	제주시	이도1동	2.9	0.0	2.4	일반유통 주점업	199833010	2020-12-15

• 일자: 평균 기온, 일강수량, 최대 풍속을 기록한 날입니다. 하루 단위로 되어 있습니다.

• 시도명: 제주시와 서귀포시가 있습니다.

- 변수 설명:

• 읍면동명: 제주시와 서귀포시 각각의 읍면동이 있습니다.

• 평균기온: 섭씨온도 (℃)

• 일강수량: 일일강수량 (mm)

• 최대 풍속: (m/s)

• 업종명: 다양한 업종이 있습니다.

• 이용금액: 한달치 금액입니다. 하루치 금액이 아님에 주의합니다.

• 데이터기준일자: 데이터를 등록한 날짜입니다.

- 데이터 전처리:

- 널 값 제거

- 필요 없는 컬럼 삭제

- 숫자형 데이터 확인

- 업종 확인

- 누락된 읍면동 없는지 확인

- 월(month)과 분기(quarter) 컬럼 추가

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1603367 entries, 0 to 1603366
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  -
0   일자        1603367 non-null object
1   시도명      1603367 non-null object
2   읍면동명    1603367 non-null object
3   평균기온    1591259 non-null float64
4   일강수량    1601446 non-null float64
5   최대 풍속   1577476 non-null float64
6   업종명      1603367 non-null object
7   이용금액    1603367 non-null int64
8   데이터기준일자 1603367 non-null object
dtypes: float64(3), int64(1), object(5)
memory usage: 110.1+ MB
```

```
data['업종명'].value_counts()
```

```
data.drop(['데이터기준일자'], axis=1, inplace=True)
```

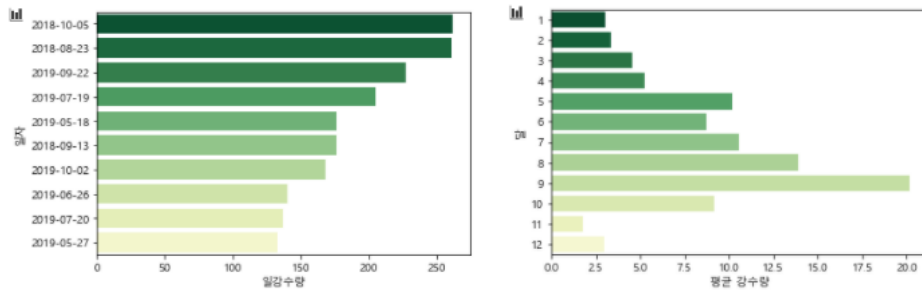
```
len(data['읍면동명'].unique())
```

```
data["quarter"] = data["datetime"].dt.quarter
data["month"] = data["datetime"].dt.month
```

데이터 분석

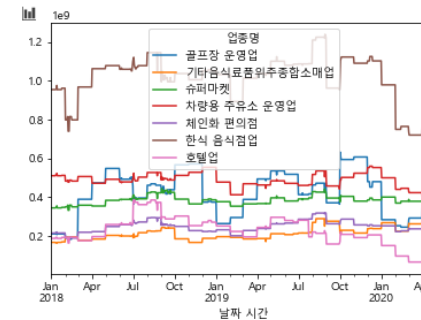
1. 제주도 강수량 정보

- 하루 동안 내린 비는 10월 5일이 가장 많았고 5월 또한 비가 많이 내렸다. 제주도에서는 여름(6, 7, 8월) 전후도 비에 대한 대비를 할 필요가 있다.
- 한 달을 기준으로 9월에 비가 가장 많았다. 비와 관련있는 업종은 9월을 주의해야 한다.



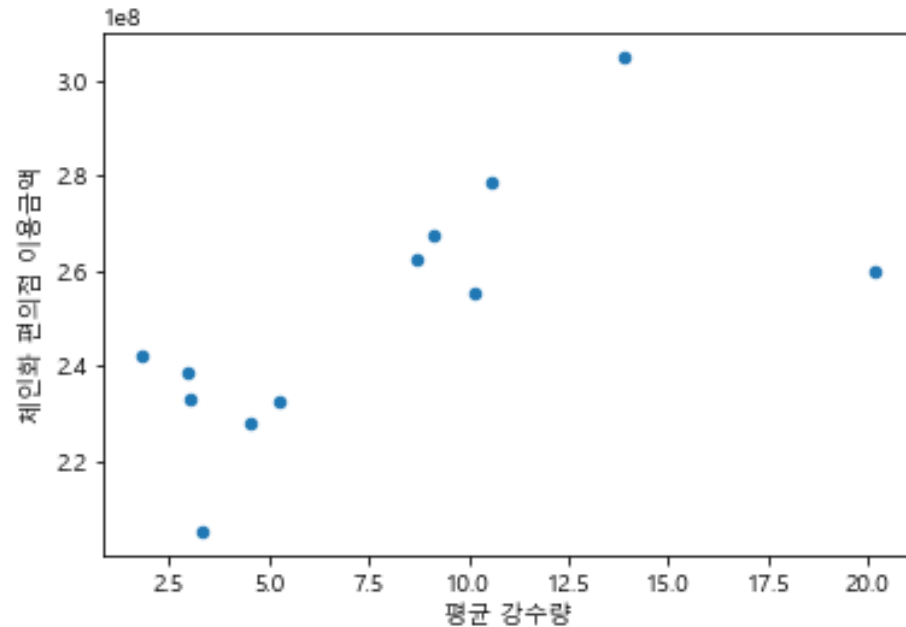
2. 개인이 창업하기 비교적 쉬운 업종

- 이용금액은 높지만 개인이 창업하기 어려운 업종들이 있다.
- 다음 업종을 제외한다. (정기 항공 운송업, 면세점, 기타 대형 종합 소매업)
- 그 후 상위 7개 업종의 이용금액을 분석한다.
- 그래프를 통해 체인화 편의점과 슈퍼마켓이 가장 안정적인 매출임을 알 수 있다. 또한 이들은 창업하기 비교적 쉬운 업종이다.



데이터 분석

편의점과 강수량의 상관 관계



	체인화 편의점 이용금액	평균 강수량
체인화 편의점 이용금액	1.000000	0.684596
평균 강수량	0.684596	1.000000

두 변수의 상관 관계(correlation) 크기는 .68 로 두 변수는 강한 관계가 있다. 결정 계수(coefficient of determination)는 .46으로 이용금액 분산(variance)의 46%를 비와 관련하여 설명할 수 있다.

● 분석 결론

- 분석결론:

- 과거에 비가 많이 내린 날과 달을 파악하여 이후를 대비할 수 있다.
- 체인화 편의점은 다른 업종에 비해 비교적 적은 금액으로 안정적인 사업을 할 수 있다.
- 비가 많이 내리면 체인화 편의점의 매출이 높아지는 상관 관계가 있다.
- 분석한 내용을 토대로 편의점 매출을 증가할 수 있는 인사이트를 제공한다.

- 인사이트:

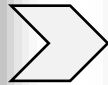
- 적용 대상: 비가 많이 내리는 제주도 지역에 편의점 창업을 원하는 고객
- 첫 번째 개선 사항: 비가 내리면 주류 매출이 증가하므로 편의점에 다양한 주류를 구비한다.
- 기대 효과: 손님이 선호하는 주류가 많을 수록 구매 욕구가 높아질 것이다.
- 관련 자료: [GS25의 장마가 시작된 지난 열흘간 막걸리 매출 분석](#)
- 두 번째 개선 사항: 편의점에 다양한 우산과 양말을 배치하고 고객이 잘 보이는 위치에 둔다.
- 기대 효과: 우산과 양말 판매로 높은 매출 상승을 기대할 수 있다.
- 관련자료: [비오면 양말 20배 팔려...](#)

데이터 수집 / 전처리

제주특별자치도_제주도장소(POI)데이터_20151231

	A	B	C	D	E	F	G
1	장소 POI	위치좌표	위치좌표	구분	장소명	소재지	데이터기준일자
2	3	126.5688	33.23655	교통시설	동방파제	제주특별지	2015-12-31
3	4	126.5626	33.23507	지명관련	새섬	제주특별지	2015-12-31
4	5	126.5997	33.23031	지명관련	섬섬	제주특별지	2015-12-31
5	6	126.5659	33.22603	지명관련	문섬	제주특별지	2015-12-31
6	7	126.5164	33.21841	지명관련	범섬	제주특별지	2015-12-31
7	8	126.2225	33.24152	공공기관	일과2리어	제주특별지	2015-12-31
8	9	126.2705	33.17016	지명관련	가파도	제주특별지	2015-12-31
9	10	126.2683	33.11767	지명관련	마라도	제주특별지	2015-12-31
10	11	126.2703	33.42584	교육기관	수원초등학교	제주특별지	2015-12-31
11	12	126.3129	33.42017	교육기관	어도초등학교	제주특별지	2015-12-31
12	13	126.2934	33.44242	교육기관	귀덕초등학교	제주특별지	2015-12-31
13	14	126.3305	33.4354	교육기관	남읍초등학교	제주특별지	2015-12-31
14	15	126.4379	33.47664	지명관련	참오랭이	제주특별지	2015-12-31
15	16	126.483	33.49187	주책관련	대진빌라	제주특별지	2015-12-31
16	17	126.4734	33.483	주책관련	노형맨션	제주특별지	2015-12-31
17	18	126.4494	33.48382	지명관련	도근천	제주특별지	2015-12-31
18	19	126.4807	33.48877	주책관련	대성맨션	제주특별지	2015-12-31
19	20	126.4805	33.49189	주책관련	로얄아트빌	제주특별지	2015-12-31
20	21	126.4929	33.4827	주책관련	백록빌라	제주특별지	2015-12-31
21	22	126.4482	33.46889	지명관련	눌동산	제주특별지	2015-12-31
22	23	126.4526	33.46948	기업	오성레미콘	제주특별지	2015-12-31
23	24	126.3731	33.46953	기업	삼양교통	제주특별지	2015-12-31
24	25	126.495	33.46818	지명관련	광이오름	제주특별지	2015-12-31
25	26	126.4527	33.46827	기업	서한산업	제주특별지	2015-12-31
26	27	126.4837	33.48364	주책관련	원노형주	제주특별지	2015-12-31
27	28	126.3935	33.47824	지명관련	귀일리	제주특별지	2015-12-31
28	29	126.4427	33.49587	지명관련	도근내	제주특별지	2015-12-31
29	30	126.4919	33.46588	지명관련	상여오름	제주특별지	2015-12-31
30	31	126.4952	33.48284	주책관련	삼창빌라	제주특별지	2015-12-31
31	32	126.496	33.48659	관광/숙박	코리야나호텔	제주특별지	2015-12-31
32	33	126.4489	33.46559	문화/종교	해성사	제주특별지	2015-12-31
33	34	126.4702	33.46251	도로시설	월산교	제주특별지	2015-12-31

<전처리 전 csv file>



	A	B	C	D	E	F	G
1	위치좌표	위치좌표	구분	장소명	소재지	시구분	동구분
2	126.5688	33.23655	교통시설	동방파제	제주특별지서귀포시	서귀동	서귀동
3	126.5626	33.23507	지명관련	새섬	제주특별지서귀포시	서귀동	서귀동
4	126.5997	33.23031	지명관련	섬섬	제주특별지서귀포시	보목동	보목동
5	126.5659	33.22603	지명관련	문섬	제주특별지서귀포시	서귀동	서귀동
6	126.5164	33.21841	지명관련	범섬	제주특별지서귀포시	법환동	법환동
7	126.2225	33.24152	공공기관	일과2리어	제주특별지서귀포시	대정읍	대정읍
8	126.2705	33.17016	지명관련	가파도	제주특별지서귀포시	대정읍	대정읍
9	126.2683	33.11767	지명관련	마라도	제주특별지서귀포시	대정읍	대정읍
10	126.2703	33.42584	교육기관	수원초등학교	제주특별지제주시	한림읍	한림읍
11	126.3129	33.42017	교육기관	어도초등학교	제주특별지제주시	애월읍	애월읍
12	126.2934	33.44242	교육기관	귀덕초등학교	제주특별지제주시	한림읍	한림읍
13	126.3305	33.4354	교육기관	남읍초등학교	제주특별지제주시	애월읍	애월읍
14	126.4379	33.47664	지명관련	참오랭이	제주특별지제주시	외도일동	외도일동
15	126.483	33.49187	주책관련	대진빌라	제주특별지제주시	연동	연동
16	126.4734	33.483	주책관련	노형맨션	제주특별지제주시	노형동	노형동
17	126.4494	33.48382	지명관련	도근천	제주특별지제주시	도평동	도평동
18	126.4807	33.48877	주책관련	대성맨션	제주특별지제주시	노형동	노형동
19	126.4805	33.49189	주책관련	로얄아트빌	제주특별지제주시	연동	연동
20	126.4929	33.4827	주책관련	백록빌라	제주특별지제주시	연동	연동
21	126.4482	33.46889	지명관련	눌동산	제주특별지제주시	해안동	해안동
22	126.4526	33.46948	기업	오성레미콘	제주특별지제주시	도평동	도평동
23	126.3731	33.46953	기업	삼양교통	제주특별지제주시	애월읍	애월읍
24	126.495	33.46818	지명관련	광이오름	제주특별지제주시	연동	연동
25	126.4527	33.46827	기업	서한산업	제주특별지제주시	해안동	해안동
26	126.4837	33.48364	주책관련	원노형주	제주특별지제주시	노형동	노형동
27	126.3935	33.47824	지명관련	귀일리	제주특별지제주시	애월읍	애월읍
28	126.4427	33.49587	지명관련	도근내	제주특별지제주시	내도동	내도동
29	126.4919	33.46588	지명관련	상여오름	제주특별지제주시	연동	연동
30	126.4952	33.48284	주책관련	삼창빌라	제주특별지제주시	연동	연동
31	126.496	33.48659	관광/숙박	코리야나호텔	제주특별지제주시	연동	연동
32	126.4489	33.46559	문화/종교	해성사	제주특별지제주시	해안동	해안동
33	126.4702	33.46251	도로시설	월산교	제주특별지제주시	해안동	해안동

<전처리 후 csv file>

- 데이터 설명

제주특별자치도_제주도장소(POI)데이터_20151231

출처 : 공공데이터포털 (<https://www.data.go.kr/>)

- 칼럼

장소 POI, 위치좌표 X , 위치좌표 Y, 구분, 장소명, 소재지, 데이터 기준일자

- 데이터 전처리 과정

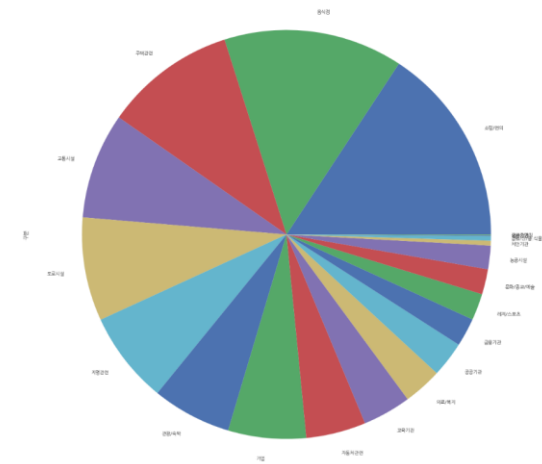
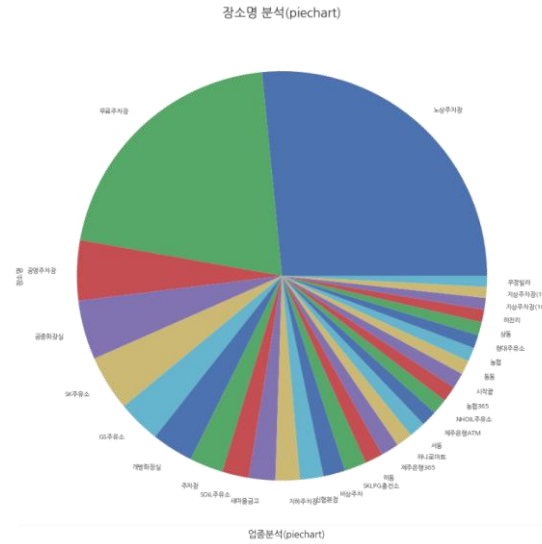
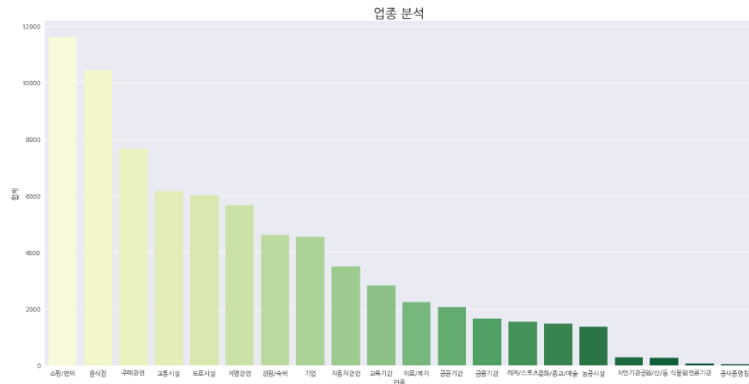
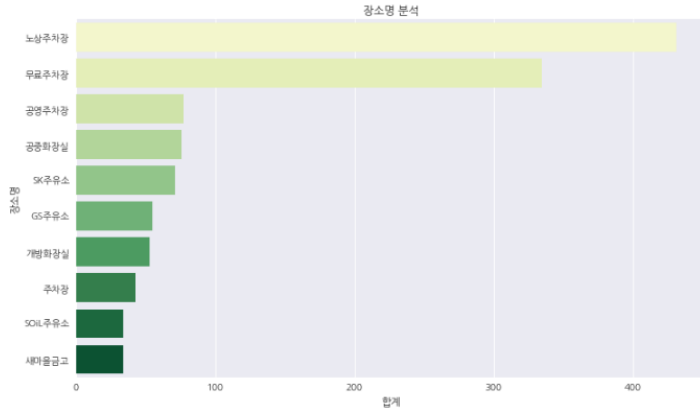
불필요한 칼럼 제거 (장소 POI 구분 코드, 데이터 기준일자)

소재지 -> 시구분, 동구분으로 나눠서 칼럼 추가

결측치 제거

이상치 검사

데이터 분석



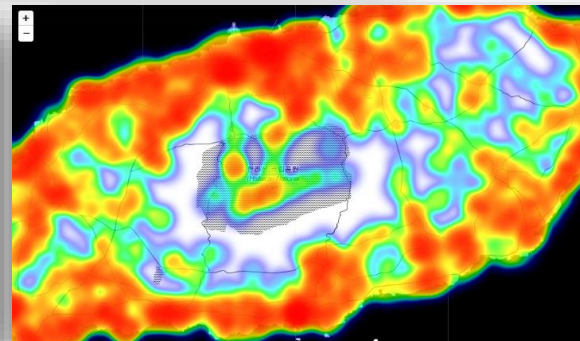
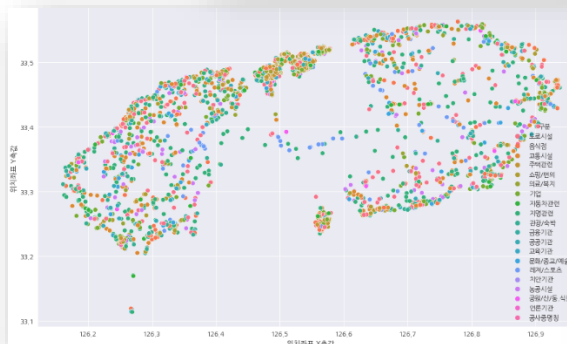
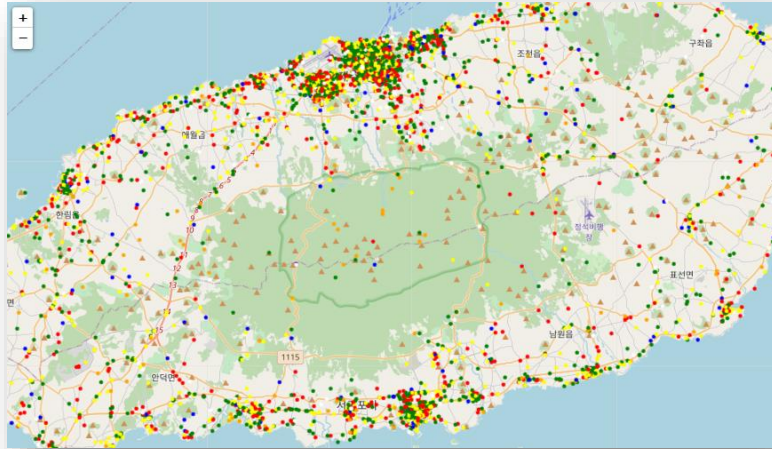
장소명 :

총 POI는 73,760개이며, 장소명 중 노상주차장이 464개, 무료주차장 325개로 장소명에서 주차장이 가장 큰 비율을 차지하고 있다. 제주도에 주차 수요가 많은 것을 알 수 있다.

업종 :

POI 데이터에서 업종 분석한 결과, 쇼핑/편의가 가장 많은 수를 차지하였고 그다음 음식점, 주택관련 순이다. 제주도의 대부분의 소비는 쇼핑과 음식 업종이 차지 할 것으로 예측된다.

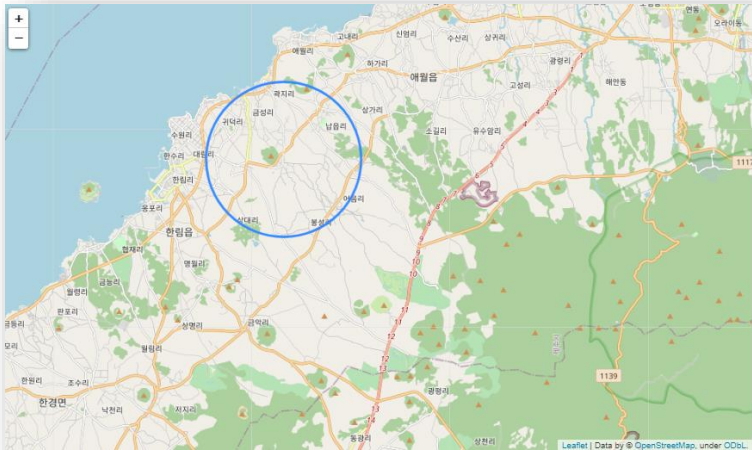
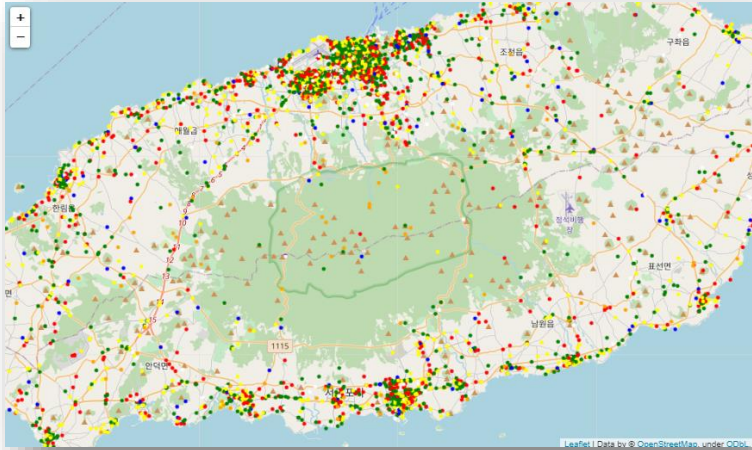
데이터 분석



- POI 시각화 한 결과 POI(Point of Interest) 는 주로 해안가 주변에 몰려 있는 것을 알 수 있음.
- 내륙 쪽은 오히려 산악지형 때문에 장소가 상대적으로 많지 않다.
- 음식점 분포, 쇼핑, 주택, 문화예술 분야가 많이 나타남
- POI 데이터를 업종별 색으로 구분하여 위치 분석.
- 쇼핑, 주택, 교통 등이 해안가 근처에 주로 위치하여 있음.
- 자동차 및 교육기관은 숫자가 전체비율상 상대적으로 적지만, 주로 제주도 중심 부분에 위치하고 있음.

● 분석 결론

〈업종 분포 지도〉



〈애월읍〉

1.

장소명과 업종명을 분석한 결과, 어떤 장소와 업종명이 제주도에서 많고 적은 수의 비율을 차지하고 있는지 알 수 있음.

상대적으로 장소가 많은 수의 업종은 경쟁에서 불리 할 것으로 분석되며,

반대로 적은 수의 업종은 유리할 것으로 판단됨.

만약 각 업종의 수익 및 매출과 접목하여 분석한다면 경쟁력에 대한 더 확실한 인사이트를 얻을 것이 기대됨.

2.

POI 위치 분석에서는 매장의 위치와 업종별 매장이 모여있는 것을 찾아 볼 수 있었음.

특정 업종 매장이 모여있는 곳은,

그로인해 기대수익이 높거나 매출을 올리기에 적합한 입지라고 판단되어,

창업주가 장소입지 선정을 고려할 때 활용 될 것을 기대 할 수 있음.

데이터 수집 / 전처리

년월 칼럼 전처리

```
year = []
for i in df5['년월'].str.split('-'):
    year.append(i[0])

month = []
for i in df5['년월'].str.split('-'):
    month.append(i[1])

df5['연도'] = year
df5['월'] = month
```



업종 대분류 전처리

#시리즈에 넣을 데이터를 만들기 위해서 리스트 안에 조건부 정보를 집어 넣기.

```
job_type = []
for i in range(len(df_cd['업종명'])):
    if df_cd['업종명'][i] in retail:
        job_type.append('소매업')
    elif df_cd['업종명'][i] in food_store:
        job_type.append('음식점')
    elif df_cd['업종명'][i] in vehicle:
        job_type.append('운송업')
    elif df_cd['업종명'][i] in drink:
        job_type.append('주류')
    elif df_cd['업종명'][i] in service:
        job_type.append('서비스업')
    elif df_cd['업종명'][i] in rooms:
        job_type.append('숙박업')
    else :
        job_type.append('기타')
```

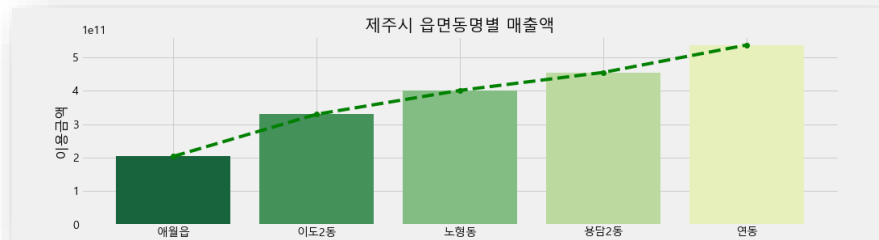
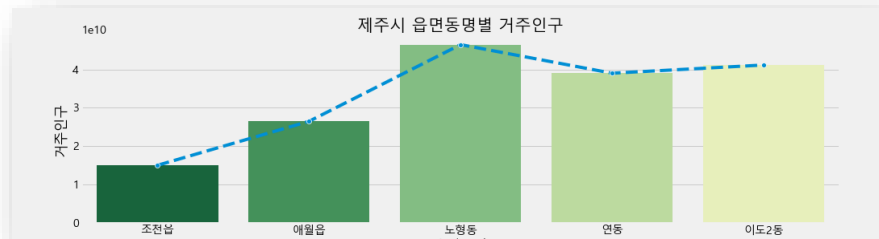
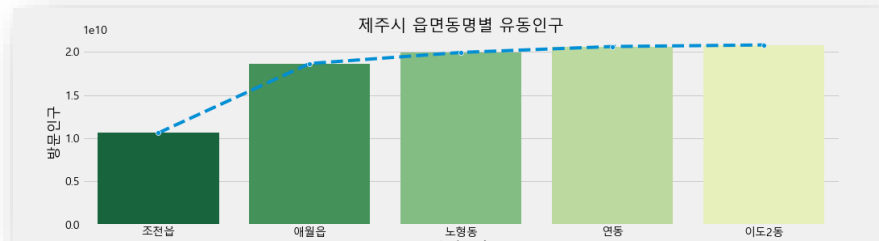
데이터 출처

- 제주특별자치도 개별관광(FIT) 증가에 따른 제주 관광객 소비패턴 변화 분석
BC카드 빅데이터 내국인관광객.csv
- 공공데이터 포털(<https://www.data.go.kr>)

데이터 설명

- “상권분석을 위한 유동인구에 따른 카드 매출액 변화 데이터“
- 전체 행 개수 9만여개, 칼럼 12개

데이터 분석



- 유동인구가 많은 지역을 찾기

-방문인구와 거주인구를 기준으로 해서 유동인구가 많은 지역

- - 이도2동 -> 연동 -> 노형동 -> 애월읍 -> 조천읍 (유동인구)

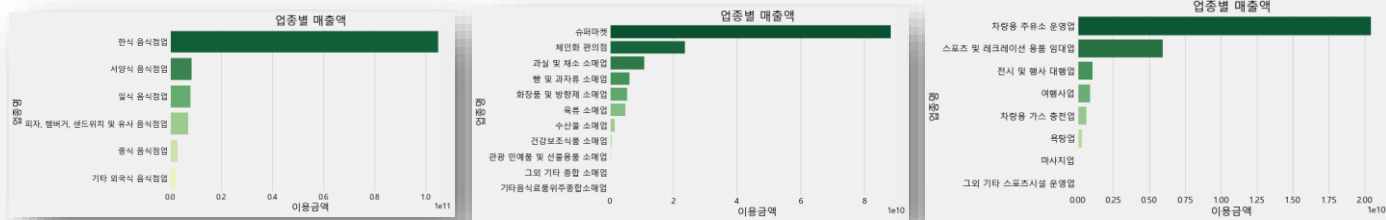
- 매출액이 많은 지역을 찾기

-이용금액을 기준으로 해서 매출액이 많은 지역

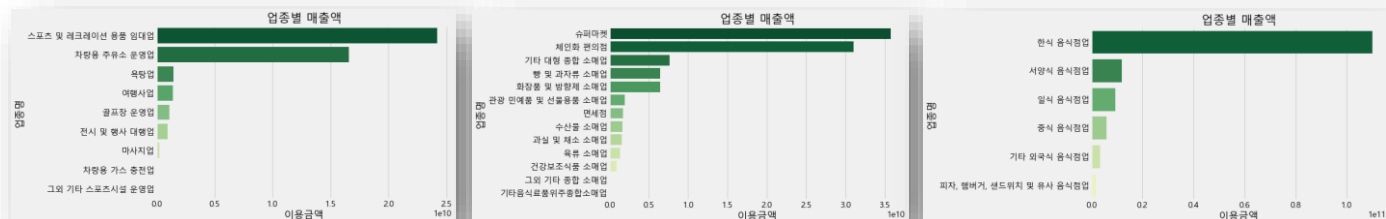
- - 연동 -> 용담2동 -> 노형동 -> 이도2동 -> 애월읍 (유동인구)

데이터 분석

• 유동인구 높은 지역의 업종 분석



연동 소매업, 기타업, 음식점업



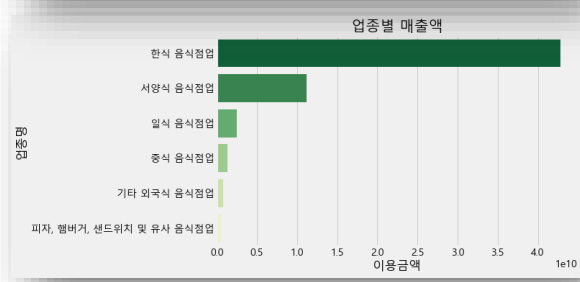
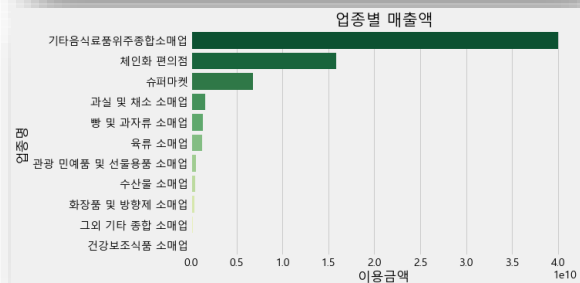
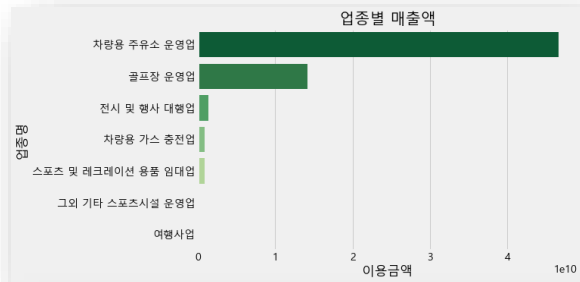
이도2동 소매업, 기타업, 음식점업

▪ 유동인구가 많은 이도2동과 연동 2동 살펴봄

▪ 이도2동과 연동의 업종분석 결과, 슈퍼마켓이나 주유소가 강세인 이유는, 두 지역 모두 아파트 단지이거나 시내이기 때문이다.

▪ 창업자 중에서는 관광지에 창업을 하려는 경우가 있을 것이다. 따라서 유동인구가 많으면서 관광지로 유명한 "애월읍" 을 따로 분석하려고 한다.

데이터 분석



애월읍 소매업, 기타업, 음식점업

• 골프장 운영업

- 애월읍의 기타업은 다른 지역과 다르게 골프장 운영업이 매출액의 2위를 차지하고 있다.
- 실제로 애월읍의 위치와 골프장 개수를 네이버를 통해서 확인해보니 많은 골프장들이 애월읍에 위치하고 있었다.
- 따라서 골프장 운영업을 추천하거나, 골프장 관련 장비 임대업을 추천한다.

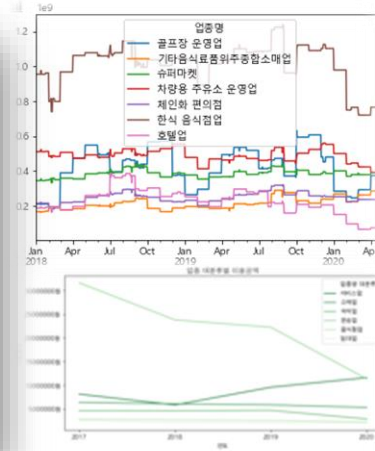
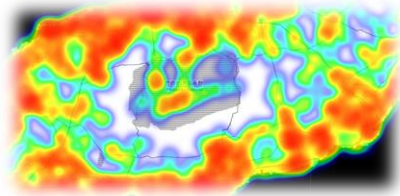
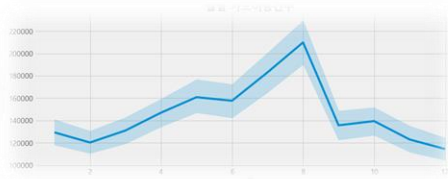
• 스포츠 및 레크레이션

- 스포츠 및 레크레이션의 매출액이 다른 지역에 비해서 굉장히 낮게 나왔다.
- 스포츠 및 레크레이션 매장수가 다른 지역과 동일함에도 불구하고 매출액이 낮게 나왔기 때문에, 애월읍에서 스포츠 및 레크레이션 사업을 하지 않는걸 추천한다.

• 서양 음식점

애월읍은 다른 지역과 달리 서양 음식점이 매출액에 높은 비율을 차지하고 있다. (한식제외) 따라서 애월읍에 음식점을 차리는 것을 추천한다.

● 종합 인사이드



추천 업종

추천 업종 : 편의점

편의점 매장 수는 계속해서 증가세
분석 결과 안정적인 매출 지속 중.

위치

유동인구가 높은 지역일수록 상권 매출도 높음.
유동인구가 높으면서 호텔업 (복합리조트)
이 발달한 지역인 노형동을 추천 위치.

타겟 고객

대부분의 매출이 4-50대 남녀 외국인에서
가장 높게 나타남.

마케팅 전략

복합 리조트와의 계약을 맺고 비가 오는 날 편의점
이용자에게 이용권 추천 행사를
진행한다라는 마케팅 전략

프로젝트 후기

김수진 : 데이터를 분석하는 과정에서 흥미로운 결과를 얻어 내고, 이로부터 도출한 결론이 다른 데이터에서도 검증될 때 가장 흥미를 느꼈다. 서로 다른 데이터에서 도출해낸 결론이 상이한 경우도 있었는데, 이게 심할 경우 데이터를 사용해도 되는가에 대한 의문이 들었다.

데이터를 분석하다 보면 흥미로운 부분을 찾아서 계속해서 파고들다가 본래의 분석 목표를 벗어나게 되는 경우가 잦았다. 분석 목표를 명확하게 정하고 분석을 해야겠다고 생각했다.

김하연 : 안녕하세요 멀티캠퍼스 데이터사이언스 / 엔지니어링 3기를 수강하고 있는 김하연입니다. 저는 데이터 정제, 분석이나 시각화가 처음이라, 파이썬의 문법적인 걱정이 많았었습니다.

하지만 분석을 진행하면서 늘어난 구글링 스킬로 극복했고, 즉각적으로 분석결과가 나오는 즐거움을 처음으로 느낀 시간이었습니다.
멀티캠퍼스 3, 4기 파이팅 !

조아진 : 데이터 분석이라는 게 자료수집-전처리-분석-결론도출 전반에 걸쳐 많은 시간과 노력이 드는 일이라는 것을 알게 되었다.

어느 한 과정이라도 소홀히 하면 다시 전 단계로 돌아가 작업해야 하는 일이 발생하는 것을 경험하면서 전 과정에 걸쳐 충분한 고민이 필요하다는 것을 알게 되었다.

장은준 : 전처리 과정에서 통계와 관련된 지식이 중요함을 느꼈다. 통계를 다음에 더 보충해서 전처리를 해야겠다. 더 심화된 분석을 하기 위해서는 시각화만으론 부족하다고 느꼈다. 결과를 분류하거나 예측하기 위해서는 머신러닝이 필요하다고 느꼈다.

다음 과정으로 머신러닝 수업을 듣는데, 집중해서 들어야겠다. 프로젝트에서 협업이 중요함을 느꼈다. 소통을 잘하는 법을 배워야겠다.

이동준 : 조원과 협력해서 작업 내용과 기한을 정하고 서로의 역할을 분배하는 과정이 재미있었다. 프로젝트를 진행해 나가면서 처음 기획했던 것보다 훨씬 많은 내용을 수정했고, 그 결과 멋진 결과물이 나온 것 같다.

특히 프로젝트 회의에서 조원 모두가 적극적으로 참여한 것이 매우 인상 깊었다. 협업을 하면서 여러모로 당황한 적도 많았지만 그때마다 조원분들이 친절하게 도와줘서 고맙다.

THANK YOU

제주도 관광객 소비 패턴 분석 프로젝트

