1.

안녕하십니까 제주 핫 플레이스에 대해 발표하게된 김동현입니다.

2.

발표 순서는 분석주제, 데이터구성, 분석 및 결과, 결과 순으로 발표하겠습니다.

3.

5 ~ 8월에 구좌읍을 방문하는 사람들은 어떤 곳을 선호하고 어디를 가는지 그들의 핫 플레이스는 어디일지 알아보기 위해

4.

카드 데이터와 sns데이터를 가지고 핫플레이스를 찾아보았습니다.

5.

카드 데이터 먼저 설명 드리겠습니다. 카드데이터는 2019년 3~8월까지의 카드 사용 데이터이고요 city는 제주시 서귀포시, dong은 조천읍, 구좌읍, 성산읍 grid\_id는 제주도를 격격자 나누었을 때 각 격자의 아이디 이고 biz\_type은 업종, user\_type은 사용자 제주도민, 내국인, 법인 등이 있고 age는 10~70대까지, 성별, 그리고 사용금액이 있습니다.

처음엔 카드를 많이 사용하는 곳이 사람들이 많이 찾는 곳이 다라고 생각하여 분석을 시작했는데 다시 확인해 보니 가격이 한사람의 것이 아니라 격자안에서의 20 남자의 총합 이런 식으로 되어있어서 사람들이 가장 많이 가는 장소를 가격의 총합으로 다시 분석을 시작했어요

6.

전체 카드사용 그래프를 보면 7월이 조금 떨어지지만 전체 적으로 상승선을 그리고 있고 사용자별로 봤을 때는 법인이 월등히 높고, 그 다음 내국인, 다음은 도민, 중국이나 일본 등의 외국인들은 큰 특징을 찾진 못했습니다.

7.

업종별로 보면 가장 카드 사용금액이 많은 업종이 골프장으로 나왔고 연령별로 봤을 때는 나이가 없는 na값이 가장 높게 나왔어요 그 이유가 법인이 회사 카드이다 보니 나이가 없는 것으로 되어있어서 높게 나온 것으로 나왔습니다. 다음으로 카드를 많이 사용하는 연령대는 20~50대가 카드사용이 많은 것을 확인할 수 있었습니다.

8.

법인의 소비금액이 많아 법인만 따로 확이 했을 때 골프장에서 소비금액이 가장 높게 나왔고 골프장에 법인만 오는지 살펴보니 내국인과 제주도민 모두 이용하는 것을 알 수 있었습니다.

9.

정리해서 말씀드리면 법인인 가는 곳은 핫플레이스가 아니라고 판단하여 법인을 제거하고, 골프장은 골프장을 찾는 사람들이 가는 장소이므로 골프장도 제거하고, 중국, 일본, 동남아 등 외국 국적을 가진 사용자들을 외국인으로 통합해 주었습니다. 내국인의 나이와 성별의 비어 있는 값이 있는 것은 나이의 중간 값인 40대로 채워주고, 성별은 남성의 비율이 많아 남성으로 채워주었습니다. 금액은 억 단위로 해주고 5~8월의 핫플레이스를 찾을 것이니 기간은 5~8월로 해서 구좌읍의 데이터 셋을 다시 만들어 주었습니다.

10.

제주도를 격자로 나누었을 때 구좌읍은 약 300개의 격자로 나누어지고, 그 중 카드를 사용하는 곳은 76개곳이 나왔습니다. 아까 말씀드렸다시피 사람들이 많이 가는 곳은 소비금액이 많은 곳 이다라고 했을 때 각 격자의 총합으로 했을 때 상위 25%가 사람들이 가장 많이 가는 곳 이다라고 생각해서 뽑았을 때

11.

총 17곳이 나왔습니다. 이 곳들을 사용금액으로 나누었을 때 7억이상은 월정, 6억이상은 세화, 5억이상은 평대, 4억이상은 없고 4억 밑으로 나머지지역들이 나왔습니다. 월정, 세화, 평대 이 세지역이 사람들이 많이 찾는 지역으로 나왔습니다.

12.

17곳을 세분화 해서 보면 성별로 봤을 때에 남자들은 주로 세화, 월정, 평대에서 소비가 많고 주로 음식과 술에 소비가 많은 것으로 나왔고

13.

여자들은 월정, 세화, 평대에서 소비가 많은 것으로 나왔고 주소 음식과 비알콜(커피)에 소비가 많은 것으로 나왔습니다.

14.

제주도민을 남자와 여자로 나누어서 봤을 때는 남녀 모두 세화, 월정에 소비가 많고 연령별로 봤을 때는 20대가 월정, 세화 순으로 소비가 많고 20대를 제외한 나머지는 세화, 월정 순으로 소비가 많은 것을 확인할 수 있었습니다.

15.

내국인을 남자와 여자로 나눠서 봤을 때는 남녀 모두 월정, 평대, 세화 순으로 소비가 많고 연령별로 봤을 때 남자는 월정, 세화, 여자는 월정, 평대 60~70대는 평대, 세화 순으로 소비가 많은 것을 확인할 수 있었습니다.

16.

외국인 같은 경우는 월정, 하도에서 소비가 많고 주로 음식에 소비가 많은 것을 알 수 있었습니다.

17.

정리하자면 성별 전체로 봤을 때 남자는 음식과 술, 여자는 음식, 커피, 제주도민에서 남자는 음식과 술, 여자는 음식, 커피 내국인은 남녀 모두 음식, 커피 외국인은 음식에 소비가 많은 것을 알 수 있었습니다. 가장 소비가 많은 지역이 월정, 세화, 평대로 나왔는데 월정과 평대에선 음식과 커피에 소비가 많고 세화에서는 음식과 술에 소비가 많은 것을 알 수 있었습니다.

18.

다음은 sns데이터입니다. sns데이터는 2019년 1월 ~ 12월 까지의 데이터이고 sns는 인스타그램, 네이버 블로그, 네이버 카페, 다음카페 title, view수, 댓글 수, contents로 구성 되어있습니다. 사람들은 sns에 글을 올릴 때 어디를 갔는지 언급을 하기 때문에 가장 많이 언급된 곳이 사람들이 많이 방문하는 곳이 다라고 생각하여 분석을 시작하였습니다.

19.

사람들이 sns에 글을 올릴 때 해시태그로 짧게 올리는 경우가 있습니다. contents에서 tag를 추출하여 새로운 속성을 만들어 주었고 구좌읍 내에 어디라고 적는 사람들이 많아 구좌읍 내에 지역명이 언급되었다면 언급된 지역명을 추출하여 새로운 속성을 추가해 주었습니다.

20.

정리해서 말씀드리자면 먼전 contents의 중복을 제거해주고, tag추출, 지역명 추출, 아이디에 jeju가 들어간 contents는 광고성의 글이 많아 제거를 해주었고 5~8월로 기간을 정해주었습니다.

21.

Tag, place, contents를 가지고 워드클라우드를 하여 많이 언급된 단어가 무었인지 살펴보았습니다. tag에서는 제주카페, 제주맛집, 한달살이 같은 단어들이 많이 언급이 되었고 place에서는 월정, 세화, 김녕, 종달, 평대 순으로 언급이 되었습니다. contents에서는 게스트하우스, 카페, 맛집, 바다 등의 단어가 많이 언급되었습니다.

22

카드데이터에서 뽑은 세 지역과 sns데이터에서 언급된 공통 지역으로 contents에서 언급이 많이 된 게스트하우스, 카페, 맛집을 가지고 크롤링을 하고 구좌읍 데이터셋과 비교를 한 후 네이버map api를 이용하여 지도 정보를 추출하고 사람들이 많이 가는 top 3를 선정해 보았습니다. Contents에서 언급된 단어를 사용한 이유는 Tag와 contents가 언급하는 단어들이 비슷 하여 contents에서 언급된 단어들을 사용하였습니다.

23.

자세히 설명 드리자면 월정리 카페라고 네이버 map에 검색을 했을 때 title인 상호명들을 크롤링을 하고

24.

그 상호명을 가지고 구좌읍 데이터셋의 contents에서 언급이 되었는지 비교를 하고 언급된 상호명을 가지고 네이버 map api를 이용하여 지도정보를 추출하였습니다.

25.

이렇게 해서 나온 것들이 게스트하우스는 월정에 3군데 세화, 평대는 언급된곳이 없었고

26.

카페는 월정의 9군데, 세화의 4군데, 평대의 1군대

27.

맛집 같은 경우 월정의 12군대, 평대에 5군데, 세화에 3군대가 나왔습니다.

28.

가장 많이 언급된 것으로 순위를 매길 때 게스트하우스는 1위 라비 게스트하우스, 2위 문게스트하우스, 3위 월정리썬시티 게스트하우스

29.

카페는 1위 월정에 구좌상회, 2위 평대에 르토아 베이스먼트, 3위 세화에 카페 공작소

30.

맛집은 1위 평대에 명진전복, 2위 월정에 월리, 만월당, 월정갈비밥, 3위 평대에 종가전복, 갈치공장이 나왔습니다.

31.

종합해 봤을 때 사람들이 가장 많이 찾는 지역은 월정, 세화, 평대 게스트하우스는 월정, 카페는 월정 평대, 세화, 맛집은 평대, 월정이 나왔습니다. 그래서 구좌읍에 방문하는 사람들의 핫플레이스는 게스트하우스의 핫플레이스는 월정에 라비, 카페의 핫플레이스는 월정의 구좌상회, 맛집의 핫플레이스는 평대의 명진전복으로 나왔습니다.

32.

여기까지입니다.