

국내 주요 숙박업소 데이터를 이용한 여행자별 여행지 및 숙박시설 추천

목차

1. 서론

- 1.1. 주제 선정 및 프로젝트 목적
- 1.2. 데이터 설명
- 1.3. 활용 기술 및 프레임워크
- 1.4. 참여자 정보 및 각 역할

2. 데이터 전처리

- 2.1. 여행유형 분류
- 2.2. 새 컬럼 생성하기
- 2.3. 결측값 채우기

3. 주요 분석 결과

- 3.1. 숙박업소 데이터 중심의 접근
- 3.2. 여행 유형별 접근
- 3.3. 여행자 중심의 접근

4. 결론과 기대효과

1. 서론

1-1. 주제 선정 및 프로젝트 목적

(1) 주제선정의 배경

국내여행시장은 다양한 요인에 영향을 받습니다. 특히 최근 몇년간 코로나 팬데믹동안의 여행제약을 이유로 국내 여행 수요가 급증하면서 여행자들의 선호도와 소비패턴이 다양하게 변화하였습니다. 전염병이 잠잠해지고 여행관련 사업이 다시 활발해지고 있는 요즘의 여행 트렌드는 다양한 취향과 욕구를 반영하며, 개별 여행자의 선호도와 특성에 따른 맞춤형 여행 경험이 강조되고 있습니다. 이에 대응하여 저희는 국내 주요 숙박업소 데이터를 기반으로 여행자별로 최적화된 여행경험을 제공하기 위한 여행지 및 숙박시설 추천 프로젝트를 기획하였습니다.

(2) 프로젝트 목적

- 숙박업소 데이터 중심의 접근: 지역, 숙소가격, 숙소 규모 등 다양한 변수가 여행만족도에 미치는 영향을 파악하고 그에 따라 여행지별 숙박시설을 추천할 수 있다.
- 여행유형별 접근: 여행 유형별 국내여행 트렌드를 파악하여 예비 소비자의 패턴을 예측하고 그에 따른 마케팅 방안을 수립할 수 있다.
- 여행자 중심의 접근: 소득군, 직업군, 자녀여부, 반려동물 여부 등 여행자의 개별 특징에 따른 여행 경향을 파악할 수 있다.
- 여행산업 발전에 기여: 이 프로젝트를 통해 얻은 인사이트와 결과를 기반으로 여행산업의 서비스 품질 향상과 마케팅 전략 개발에 기여한다.

1-2. 데이터 설명

(1) 데이터: 국내 주요 숙박업소별 정보, 가격, 별점 및 이용자 특성 (행: 10255, 열: 23)

https://www.bigdata-culture.kr/bigdata/user/data_market/detail.do?id=354c0cc0-6403-11ee-9c85-79328b353134

(2) 숙소 유형 설명

- 호텔업: 관광객의 [숙박]에 적합한 시설을 갖추어 이를 [관광객]에게 제공하거나 숙박에 딸리는 음식, 운동, 오락, 휴양, 공연 또는 연수에 적합한 시설 등을 함께 갖추어 이를 이용하게 하는 업
- 휴양 콘도미니엄업: 관광객의 [숙박과 취사]에 적합한 시설을 갖추어 이를 그 [시설의 회원이나 공유자, 그 밖의 관광객]에게 제공하거나 숙박에 딸리는 음식, 운동, 오락, 휴양, 공연 또는 연수에 적합한 시설 등을 함께 갖추어 이를 이용하게 하는 업

구분		등록기준
호텔업	관광호텔업	(1) 욕실이나 샤워시설을 갖춘 객실 30 실 이상. (2) 외국인에게 서비스를 제공할 수 있는 체제 (3) 대지 및 건물의 소유권 또는 사용권을 확보하고 있을 것. (회원을 모집하는 경우에는 소유권이어야함) (4) 부대시설 의무설치
	가족호텔업	(1) 취사시설이 객실별로 설치되어 있거나 층별로 공동취사장이 설치되어있을 것 (2) 욕실이나 샤워시설을 갖춘 객실 30 실 이상. (3) 외국인에게 서비스를 제공할 수 있는 체제 (4) 객실별 면적이 19 제곱미터 이상. (5) 대지 및 건물의 소유권 또는 사용권을 확보하고 있을 것. (회원을 모집하는 경우에는 소유권이어야함)
	일반호텔업	일반 호텔업. 부대시설 의무설치 아님.

(3) 변수

순서	컬럼영문명	컬럼한글명	데이터타입	길이	PK여부	NOT NULL여부
1	MCATE_CD_NM	분류코드명	VARCHAR	200	Y	Y
2	SEXDSTN_CD	성별코드	VARCHAR	30	N	Y
3	BIRTH_YEAR	출생년도	VARCHAR	4	N	Y
4	MRRG_AT_CD	결혼여부코드	VARCHAR	30	N	Y
5	CHLDRN_AT_CD	자녀여부코드	VARCHAR	30	N	Y
6	FAMILY_TY_NM	가족유형명	VARCHAR	200	N	Y
7	OC_AREA_CD	거주지역코드	VARCHAR	30	N	Y
8	ANSWRR_OC_AREA_NM	답변자거주지역명	VARCHAR	200	N	Y
9	OCCP_CL_NM	직업분류명	VARCHAR	200	N	Y
10	MT_AVRG_HSHLD_INCOME_DGREE_NM	월평균가구소득정도명	VARCHAR	200	N	Y
11	COM_AT	동행자여부	VARCHAR	1	N	Y
12	COM_TY	동행자유형	VARCHAR	4000	N	Y
13	TOUR_PD	여행기간	DECIMAL	38	N	Y
14	LDGS_NM	숙박업명	VARCHAR	200	N	Y
15	LDGS_ADDR	숙박업주소	VARCHAR	500	N	Y
16	LDGS_ROAD_NM_ADDR	숙박업도로명주소	VARCHAR	500	N	Y
17	GSRM_SCALE_CN	객실규모내용	VARCHAR	500	N	Y
18	LDGS_GRAD_VALUE	숙박업등급값	VARCHAR	200	N	Y
19	LDGMNT_TY_NM	숙박유형명	VARCHAR	200	N	Y
20	LDGS_AVRG_PRC	숙박업평균가격	DECIMAL	28,5	N	Y
21	LDGS_MIN_PRC	숙박업최저가격	DECIMAL	28,5	N	Y
22	LDGS_MXMM_PRC	숙박업최대가격	DECIMAL	28,5	N	Y
23	LDGS_AVRG_SCORE_CO	숙박업평점수	DECIMAL	38	N	Y

1-3. 활용 기술 및 프레임워크

(1) 데이터 분석

- MariaDB: HeidiSQL 을 사용하여 데이터베이스를 사용 하였다.
- MySQL: 분석을 위한 정보를 추출하기 위해 쿼리를 작성하였다.
- Excel: 쿼리의 결과는 CSV 파일 형태로 저장하여 Excel 로 쉽게 확인할 수 있도록 하였다. 표, 계산등은 Excel 에서 수행하였다.

(2) 데이터 시각화

- Power BI: 효과적인 데이터 시각화와 대시보드 제작을 위해 Power BI 를 사용하였다.
데이터를 PowerBI 내부에서 그룹화 하여 사용하기도 하였다.
- Python: 일부 데이터는 파이썬 패키지를 통하여 시각화 하였다.

1-4. 참여자 정보 및 각 역할

이름	역할
김도희	데이터 분석, PowerBI 를 통한 분석시각화, 보고서 작성
안세은	데이터 분석, 파워포인트 제작
오수연	데이터 전처리 & 분석 (나이대, 소득군)
윤재찬	데이터 분석, 파워포인트 제작

2. 데이터 전처리

2-1. 여행 유형 분류: COM_TY(동행자 유형)을 기준으로 분할을 진행했다.

여행 유형	설명
가족 여행	(a) 자녀를 포함하고 있으면 가족 여행으로 분류한다
효도 여행	(a) 부모님을 포함하고, 자녀가 포함되지 않을 시 효도여행으로 분류한다. (b) 자녀를 포함하면 '효도'보다는 '아이 케어'에 집중하게 될 것이라고 판단, 제외하게 되었다.
혼자 여행	(a) 아무도 포함하지 않고 혼자서 다니면 혼자 여행으로 분류한다.
커플 여행	(a) '배우자' 혹은 '연인'과 떠나는 여행을 커플 여행으로 분류한다. (b) 배우자 + 친구 + 기타 등으로 이루어진 더블 데이트가 있을 수 있지만, 연인과의 교류를 중심으로 하지 않기 때문에 친목 여행으로 분류했다.
출장	(a) 직장 동료만을 포함하고 있으면 출장 목적으로 판단, 분류한다.
친목 여행	(a) 상기 목적을 만족하지 않는 모든 종류의 여행을 친목 여행으로 분류한다.

2-2. 새 컬럼 생성하기

(1) 반려동물 여부 (PET) 컬럼 생성

- (a) SQL 을 활용해 COM_TY(동행자유형) 컬럼을 사용했다.
- (b) 여행 시 반려동물을 데려가는 경향성을 보기 위해 활용한다.

(2) 나이 (AGE) 컬럼 생성

- (a) BIRTH_YEAR(출생년도), TOUR_PD(여행기간) 컬럼을 사용했다.
- (b) 여행 기간 컬럼이 YYYY 로 구성되어 있어서 나이를 계산할 수 있었다.
- (c) 엑셀을 활용해 TOUR_PD - BIRTH_YEAR + 1 로 AGE 를 생성했다.

(3) 나이대 (AGE_RANGE) 컬럼 생성

- (a) 엑셀과 AGE 컬럼을 활용해 나이대를 계산했다.
- (b) CONCAT 을 통해 AGE 앞자리와 '대'를 붙여, '20 대' 형식으로 생성했다.
- (c) SQL 을 활용해 60 대 이상을 하나로 묶어 주었다.

(4) 고소득, 저소득 그룹 생성

- (a) PowerBI 에서 MT_AVRG_HSHLD_INCOME_DGREE_NM(월평균 가구 소득정도)컬럼을 사용하여 소득군을 그룹화 하였다.
- (b) 500 만원 이하는 저소득군, 이상은 고소득군으로 묶어주었다.

2-3. 결측값 채우기

(1) 숙박업 주소 컬럼 (LDGS_ADDR)

- (a) SQL 을 활용해 주소가 비어 있는 곳의 호텔 이름을 출력했다.
- (b) '롯데 호텔 울산'의 주소만 NULL 이었기 때문에, 울산으로 채워 넣었다.

(2) 업소 규모 (GSRM_SCALE_CN)

- (a) 엑셀을 활용해 GSRM_SCALE_CN 중 미분류 상태인 것을 선택했다.
- (b) 해당 호텔 홈페이지를 방문해 직접 객실 규모를 찾아 입력했다.

(3) 업소 등급 (LDGS_GRAD_VALUE) 컬럼

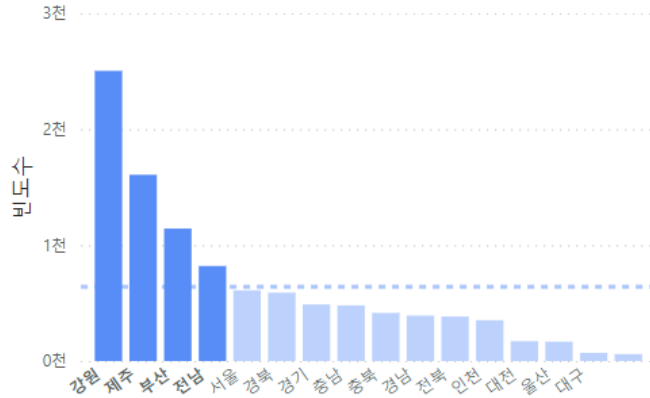
- (a) 엑셀을 활용해 업종 분류(LDGMNT_TY_NM)가 관광호텔인 것 중, 성급이 없는 업소를 선택했다.
- (b) 해당 호텔 홈페이지를 방문해 직접 성급을 찾아 입력했다.

3. 주요 분석 결과

3-1. 숙박업소 데이터 중심의 접근

(1) 여행유형: 여행유형별 지역선호도 분석

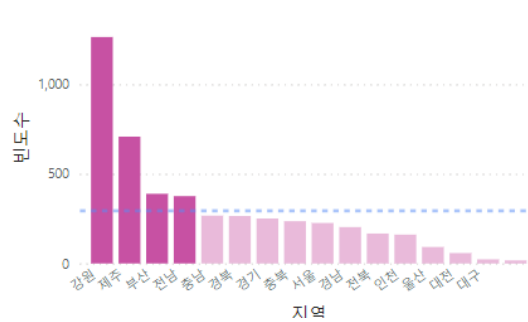
전체- 여행지역



- 위 그림은 전체 데이터에서 여행지역별 방문빈도를 시각화한 것이다. 평균빈도선을 기준으로 위에 있는 지역은 강원, 제주, 부산, 전남순이었다.

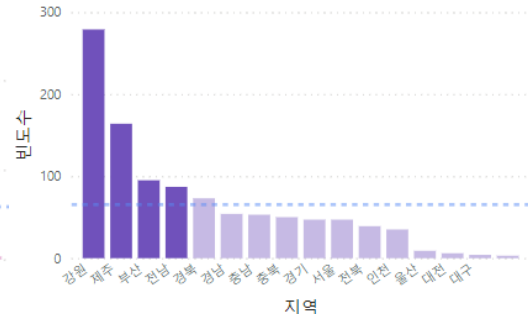
가족- 여행지역

TRAVEL_TYPE ●가족여행



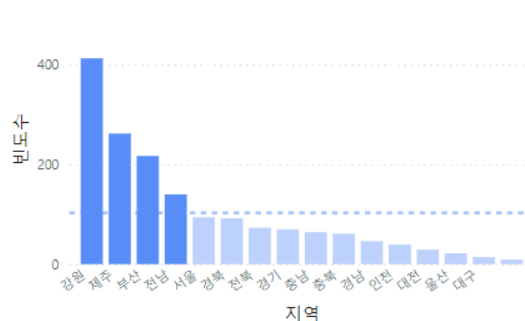
효도- 여행지역

TRAVEL_TYPE ●효도여행



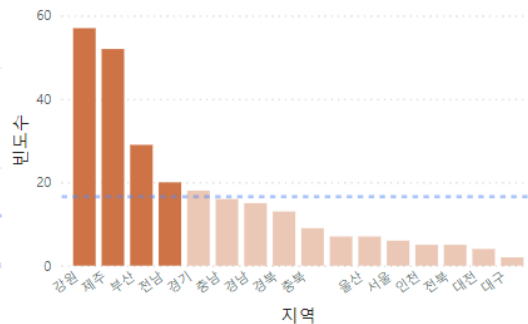
친목- 여행지역

TRAVEL_TYPE ●친목여행



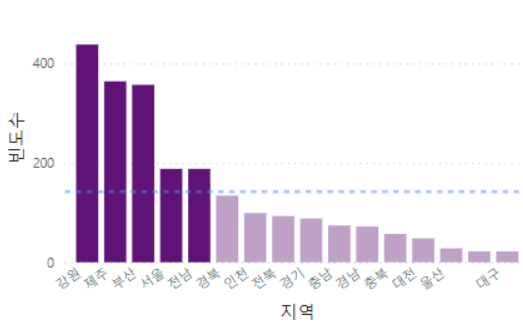
출장- 여행지역

TRAVEL_TYPE ●출장

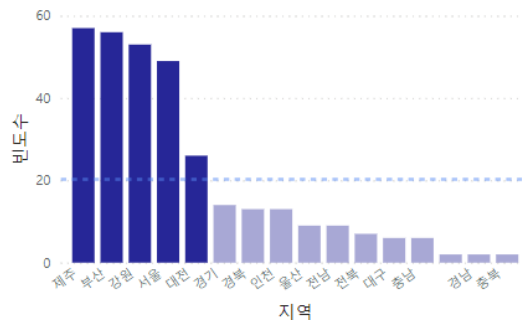


- 여행 유형별 지역선호도가 다른지 분석해보았다. 평균빈도선 위쪽에 있는 강원, 제주, 부산, 전남의 경우 [가족여행], [효도여행], [친목여행], [출장]에서 모두 동일한 순위를 나타내었다. 하위 항목은 순서에서 차이를 보였지만 격차가 작아 유의미해 보이지는 않았다.

연애- 여행지역
TRAVEL_TYPE ● 연애여행

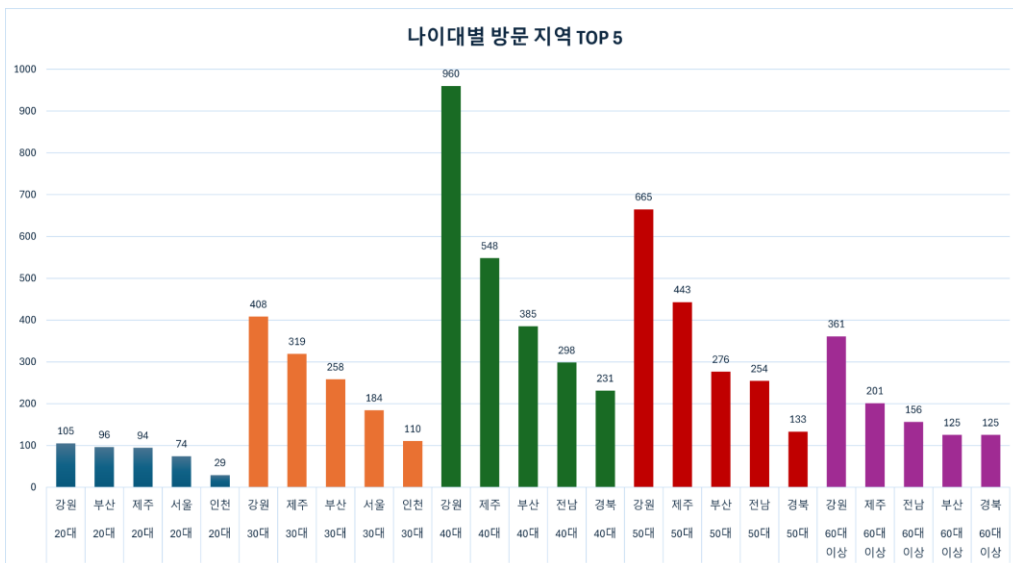


혼자- 여행지역
TRAVEL_TYPE ● 혼자여행



- [커플여행], [혼자여행]의 경우 다른 경향성을 보였다.
- [커플여행]의 경우, 평균선 위쪽에 서울이 전남보다도 위의 순위로 나타났다. 또한 제주와 부산의 방문빈도 격차가 매우 작은 것을 알 수 있었다. 이를 통해 커플 여행의 경우 부산과 서울의 선호도가 다른 유형의 여행에 비해 높음을 알 수 있었다.
- [혼자여행]의 경우 모든 여행 유형중 유일하게 전남 대신 대전이 평균빈도수 선 위쪽에 자리하였고, 강원도보다 제주와 부산의 인기가 높았다.
- 혼자여행 그래프를 살펴보았을 때 선호도 상위지역과 하위지역의 빈도수 격차는 뚜렷하게 나타나는데에 반해 상위지역간의 격차는 매우 작은 것을 알 수 있었다. 이를 통해 혼자여행 유형의 경우 확실하게 상위 4 개지역인 제주, 부산, 강원, 서울로의 여행을 선호한다는 사실을 알 수 있었다.

(2) 나이대: 나이대별 지역 선호도와 숙박유형 분석

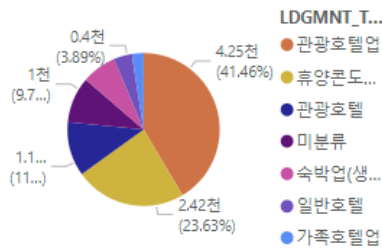


나이	인원수
10 대	21
20 대	540
30 대	1886
40 대	3822
50 대	2612
60 대	1229
70 대	131
80 대	12

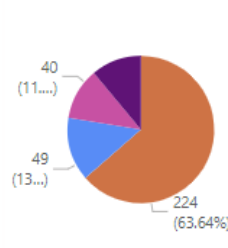
- 다른 나이대에 비하여 20~60 대의 나이에서 여행빈도가 유의미하게 나타났다. 따라서 10 대, 70 대, 80 대를 제외한 나이대에서 분석을 수행하였다.

- 강원, 제주, 부산의 경우 모든 연령대에서 선호도가 높았다. 이 세 지역을 제외하고는 나이대별로 차이를 보이는 것을 알 수 있었다. 20-30 대의 경우 강원, 제주, 부산을 제외한 TOP5 에 서울, 인천이 자리하였고, 40-60 대의 경우 전남과 경북이 자리하는 것을 알 수 있었다.
- 즉, 인천과 서울(수도권)은 젊은층에서, 전남, 경북(남부지역)은 중장년층과 노년층에서 인기있는 지역임을 알 수 있었다.

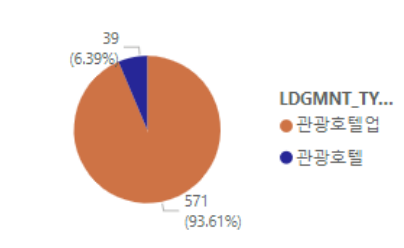
전체 숙소 유형



인천시 숙소 유형

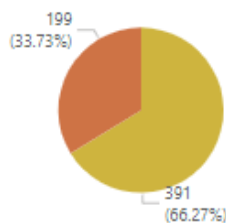


서울시 숙소 유형

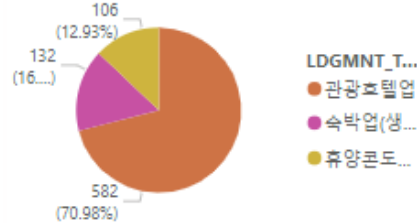


- 20-30 대의 인기지역인 인천과 서울의 숙소유형을 파이차트로 나타내 보았다. 관광호텔은 부대시설 설치가 의무인 고급호텔 계열이다.
- 인천의 숙소유형은 관광호텔, 관광호텔업의 비중이 매우 큰 것을 알 수 있었고, 서울의 경우 모든 숙소가 관광호텔, 관광호텔업이었다. 즉, 인천, 서울지역은 젊은층의 고급호텔로의 호강스가 여행트렌드임을 예상할 수 있었다.

경북 숙소 유형



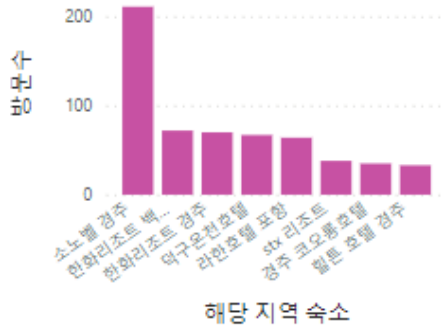
전남 숙소 유형



- 40-60 대의 인기지역인 경북과 전남의 숙소유형 또한 파이차트로 나타내었다. 경북의 경우 휴양콘도미니엄업이 우세하게 나타났지만 전남의 숙소유형은 뚜렷한 특징을 드러내지 않았다.

지역별 숙소 인기도

지역 ● 경북



지역별 숙소 인기도

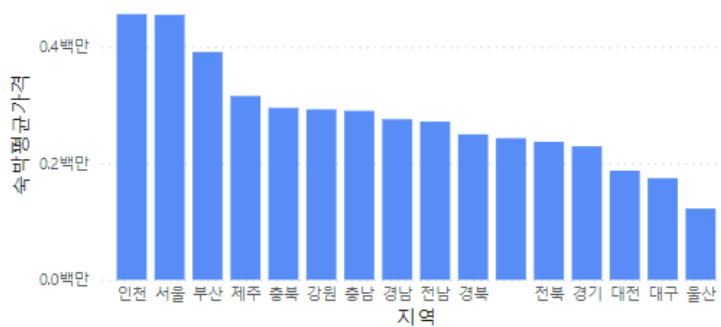
지역 ● 전남



- 더욱 자세한 분석을 위해 데이터에 있는 개별 숙소별 인기도 그래프를 그려보았다. 경북의 경우 “소노벨 경주”, 전남의 경우 “쑈비치 진도 호텔”이 다른 호텔에 비해 압도적인 방문수를 보였다. 두 숙소의 공통점은 “소노벨 경주”의 경우 보문호수, “쑈비치 진도 호텔”의 경우 바다가 보이는 아름다운 자연경관을 자랑하는 숙소라는 점이였다. 이를 통해 중장년-노년층은 여유롭게 자연경관을 즐기는 여행을 선호한다는 것을 예상할 수 있었다.

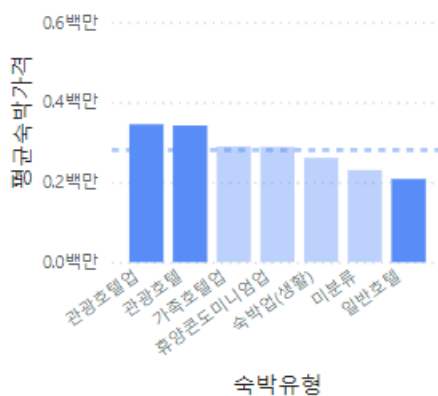
(3) 가격대: 지역별 숙박평균 가격과 가격대별 숙박유형 분석

지역별 숙박 평균 가격

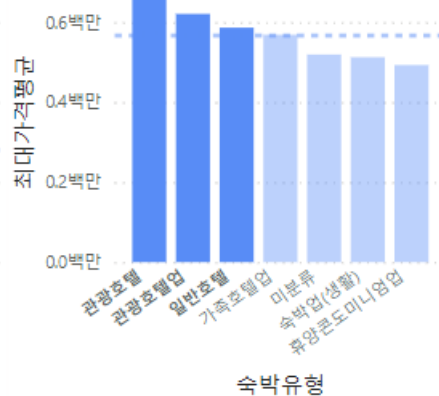


- 지역별 평균 가격은 인천, 서울이 유의미하게 높은 것을 확인할 수 있었다. 앞서 분석한것처럼 20-30 대의 럭셔리 호캉스 때문임을 예상할 수 있다.

숙박유형별 평균숙박가격



숙박유형별 최대 가격평균



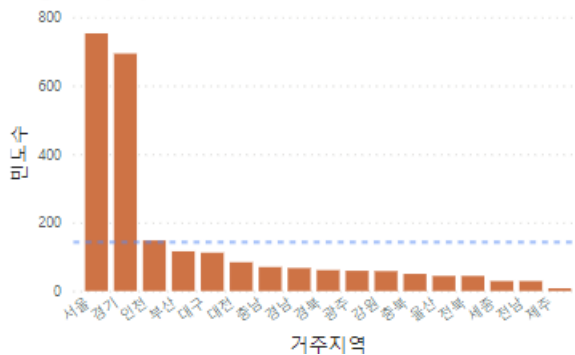
- 숙박 유형별 평균가격 그래프를 그려보았을 때, 역시 고급호텔계열인 관광호텔업과 관광호텔이 가장 높은 가격대를 나타내었다. 그 다음으로는 가족호텔업과 휴양콘도미니엄업이 비슷한 가격대로 나타났고 일반호텔이 가장 낮은 가격대를 보였다.
- 숙박가격은 객실수용 인원수에 따라 차이가 있으므로 우리는 최대 가격이 높을 수록 해당 숙소가 고급이라는 가설을 세워 그래프를 그려보았다. 숙소유형별 평균숙박가격 그래프와 최대가격평균 그래프를 그려보았을 때 [일반호텔]이 평균숙박가격은 가장 낮았지만 최대가격평균에서는 3 번째에 자리하는 것을 알 수 있었다. 이를 통해 [일반호텔]의 경우 적은 인원의 숙박이 많은 유형임을 예상할 수 있다.

(4) 거주지별: 여행지까지의 이동거리를 통한 선호도 분석

- 앞서 20-30 대와 40-60 대의 지역선호도 차를 알아보았으므로 젊은층과 중,노년층의 거주지에 유의미한 차이가 있는지 분석해보았다.

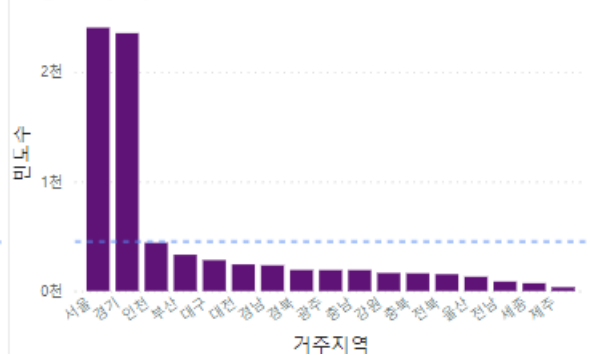
젊은층(20-30대)- 거주지역

AGE_RANGE(그룹) ● 20대 & 30대



중노년층(40-60대)- 거주지역

AGE_RANGE(그룹) ● 40대 & 50대 & 60대



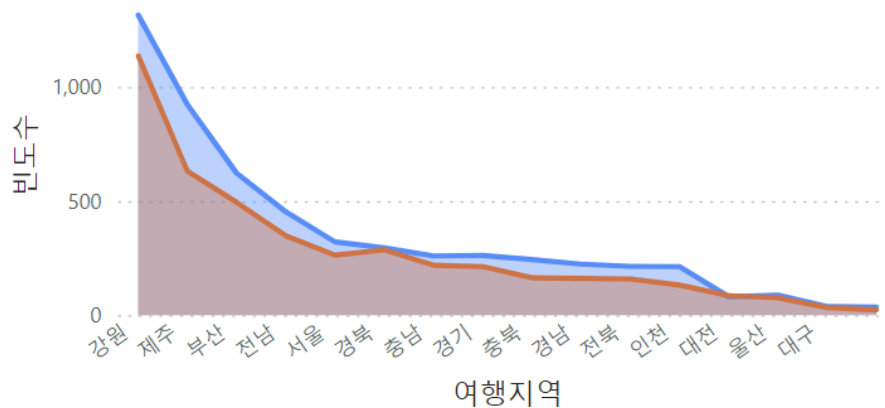
- 거주지역은 젊은층과 중노년층 모두 서울 경기지역이 압도적으로 많았다. 차이가 나타나지 않았지만 이를 통해 이동거리와 상관없이 해당 나이층이 여행지역으로서 젊은층은 서울, 인천을, 중노년층은 전남, 경북을 선호한다는 사실을 알 수 있었다.

(5) 소득군: 소득군별 선호 여행지와 숙박유형 분석

- '컨슈머인사이트'에서 조사한 '주례여행행태 및 계획조사'에 따르면, 국내여행지 선택시 경비를 주 요인으로 뽑은 비율이 증가했다. 이에 따라, '월평균가구소득정도' (MT_AVRG_HSHLD_INCOME_DGREE_NM) 컬럼에서 500 만원 까지를 저소득, 500 초과 부터는 고소득으로 분류하여 분석하였다.

소득군별 선호 여행지

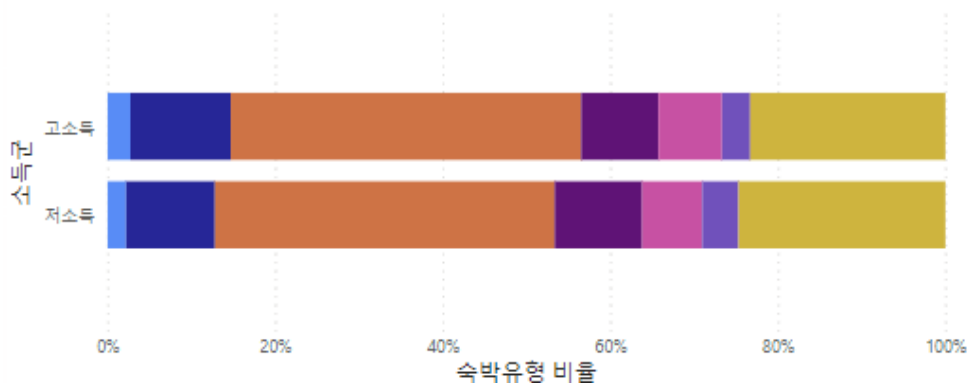
MT_AVRG_HSHLD_I... ● 고소득 ● 저소득



- 그래프에 따르면, 고소득과 저소득의 지역 선호도는 크게 차이 나지 않음을 확인할 수 있었지만 고소득층의 여행빈도가 저소득층보다 높음을 알 수 있었다.

소득군별 숙박유형 누적100 그래프

숙박유형 ●가족호텔업 ●관광호텔 ●관광호텔업 ●미분류 ●숙박업(생활) ●일반호텔 ●휴양콘도미니엄업

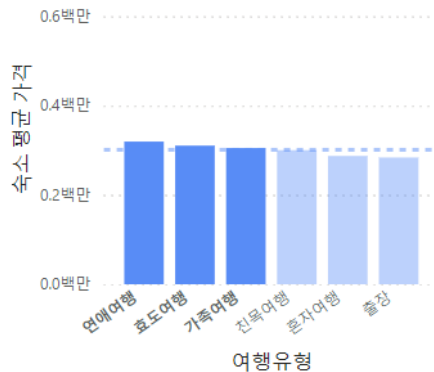


- 소득군별 숙박유형을 누적 100 퍼센트 그래프로 그려보았다. 역시 고급호텔 계열인 관광호텔, 관광호텔업은 고소득군에서 더 많은 비율로 나타났다.

3-2. 여행유형별 접근

(1) 숙소가격대별 여행유형 분석

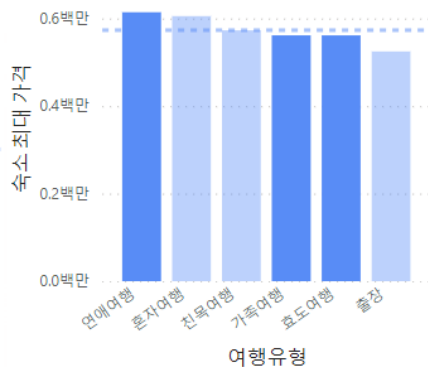
여행유형별 평균 숙소 가격



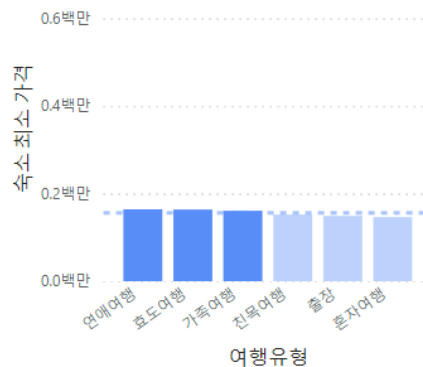
여행 유형	숙박평균가격 (원)
커플여행	319,580
효도여행	310,551
가족여행	305,165
전체평균	301,104
친목여행	300,339
혼자여행	287,422
출장	283,925

- 여행 유형별 평균 숙소가격대는 연애여행, 효도여행, 가족여행 순으로 높은것을 확인할 수 있었다. 하지만 여행의 경우 인원수에 따라 숙박금액이 차이 나기 때문에 해당 숙박업소의 최저가격과 최대가격또한 그래프로 만들어 비교해 보았다.

여행유형별 숙소 가격 - 최대값

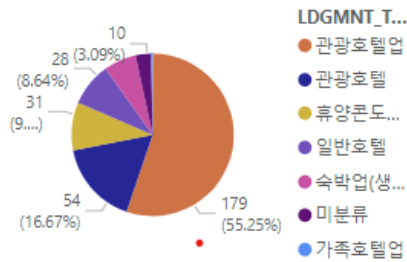


여행유형별 숙소 가격 - 최소값

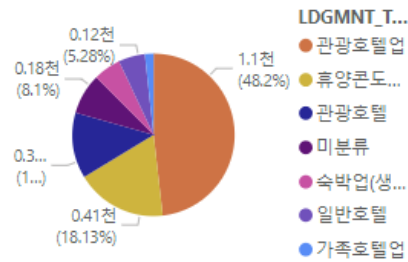


- 가장 비싼방의 가격을 그래프로 만들어 보았을 때 평균숙소가격 그래프의 Top3 와 달리 효도여행과 가족여행 대신 연애여행, 혼자여행이 들어온 것을 확인할 수 있었다. 최소가격그래프에서는 평균숙소가격과 비슷한 경향을 보였지만 출장보다도 혼자여행이 낮은 가격대를 나타내었다.
- 우리는 혼자여행의 평균 가격대가 낮은 이유가 낮은 급의 숙소에 묶어서가 아닌 1 인실이여서일것이라 예상하였고 이를 알아보기위해 혼자여행의 숙소유형을 나타내는 파이차트를 그려보았다. 또한 모든 자료에서 가장 높은 가격대로 나타난 연애여행과 낮은 가격대로 나타난 출장의 파이차트도 비교해 보았다.

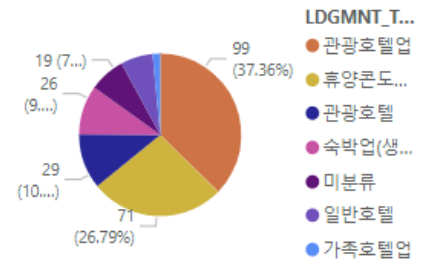
혼자여행 - 숙소유형



연애여행 - 숙소유형

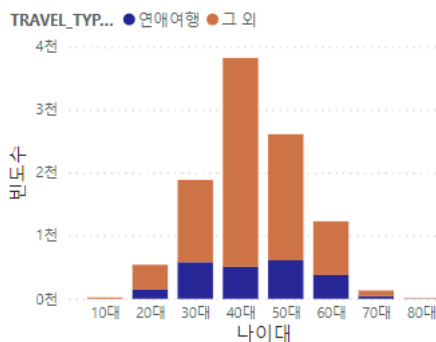


출장 - 숙소유형



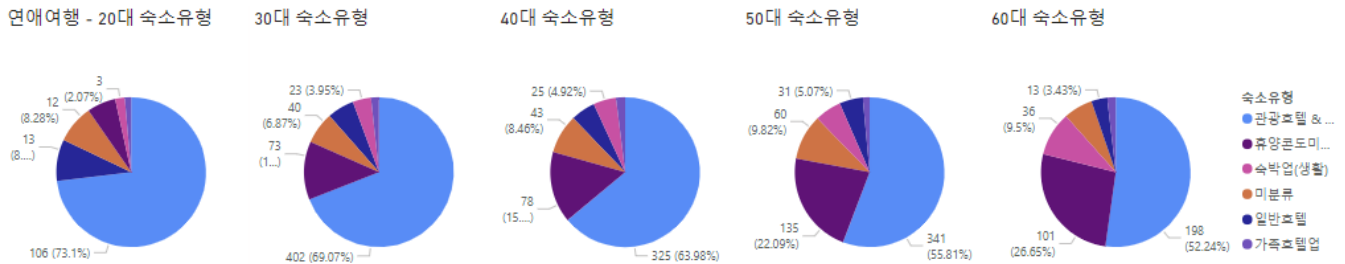
- 파이차트를 보았을 때 예상대로 혼자여행에서 고급숙소인 관광호텔, 관광호텔업의 비율이 높게 나타났다. 이와 반대로 가장 가격대가 낮았던 출장에서는 관광호텔의 비율이 낮게 나타나는 것을 알 수 있었다.
- 연애여행의 경우 가장 높은 가격대로 나타났지만 예상과 달리 관광호텔업의 비중이 압도적이지 않았다. 따라서 연애여행에서 나이대별 소비정도와 선호 숙소를 구체적으로 분석해보았다.

연애여행 비중



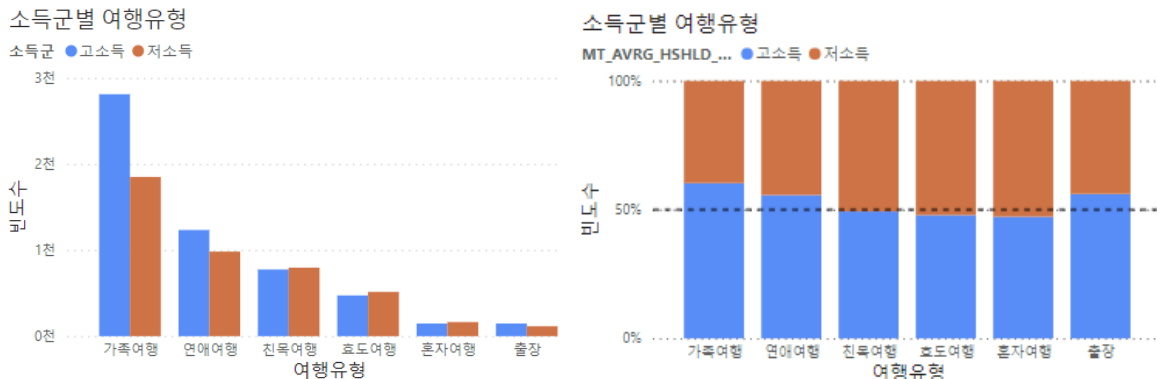
나이대	평균 가격 (원)	최소 가격 (원)	최대 가격 (원)
20 대	330,610	172,627	656,064
30 대	352,271	179,421	678,120
40 대	332,008	174,926	632,870
50 대	298,408	148,553	579,184
60 대 이상	285,708	150,149	540,401

- 연애 여행에서 숙박비용을 가장 많이 소비하는 나이대는 30 대로 나타났다. 또한 나이대별 전체 여행에서 연애여행의 비중 그래프를 그려보았을때 30 대에서 연애여행 비중이 가장 큰 것을 알 수 있었다. 30 대부터는 나이가 들수록 숙소비용이 낮아지는 것을 알 수 있었고 나이대별 숙소유형 파이차트를 그려 알아보았다.



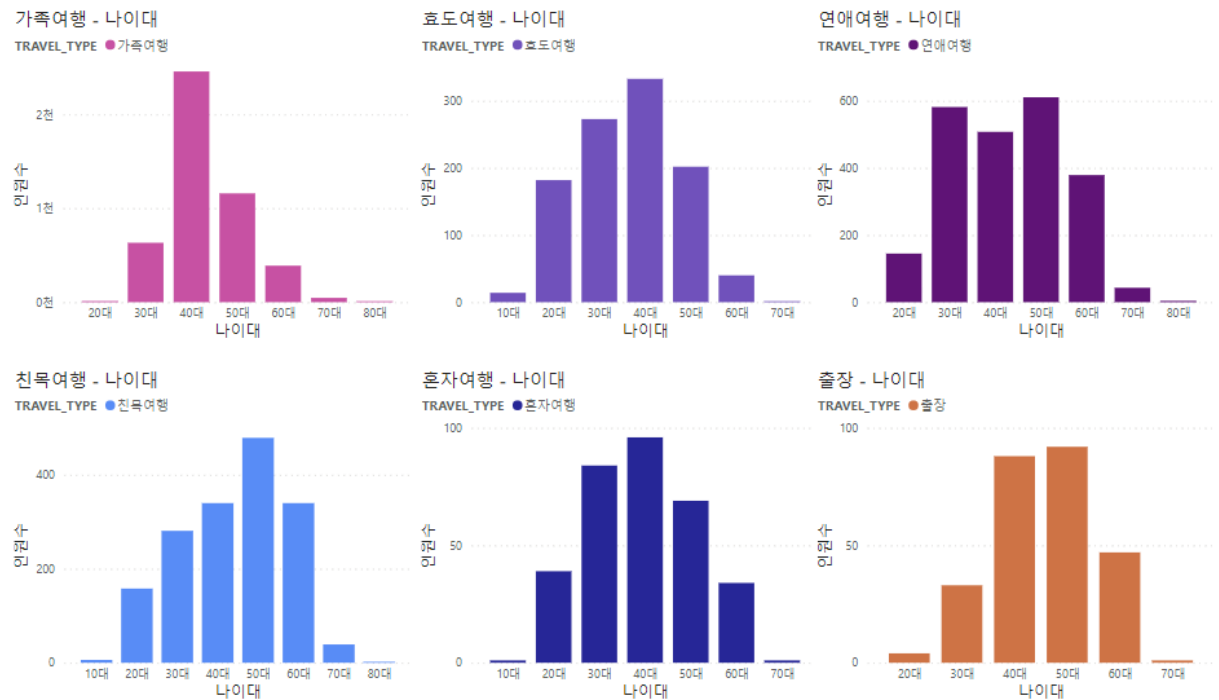
- 파이차트를 통해 나이가 들수록 숙소유형에서 관광호텔의 비율은 유의미하게 낮아지고 휴양콘도미니엄업의 비율이 올라감을 알 수 있었다.
- 따라서 숙박시설에서 커플 여행을 타겟으로 할 경우 젊은층은 럭셔리, 중노년층은 가성비를 즐길 수 있도록 해야한다는 사실을 알 수 있었고, 나이대 중에선 30 대가 여행중 연애여행이 차지하는 비율이 가장 크고, 소비금액도 크기 때문에 주요 마케팅 타겟층으로 설정하는 것이 좋겠다는 인사이트를 얻을 수 있었다.

(2) 소득군별 여행유형 분석



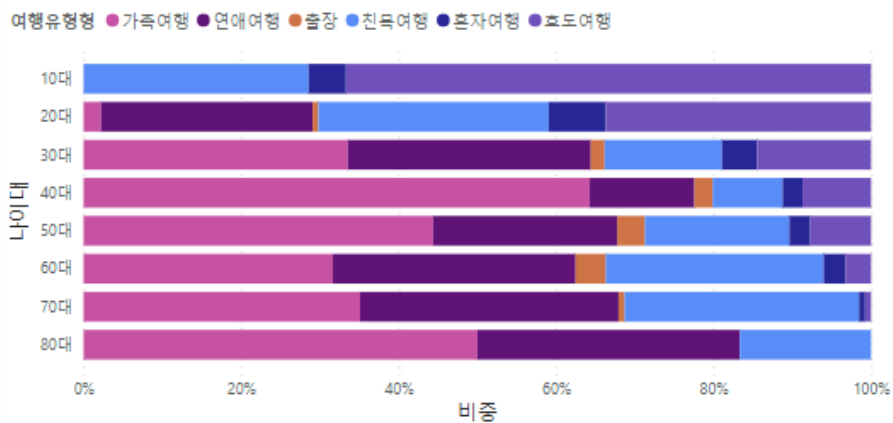
- 소득군을 고소득 (600 만원 이상을 버는 소득층)과 저소득(500 만원 이하를 버는 소득층)으로 나누어 여행유형을 분석해보았다. 고소득군이 선호하는 여행유형과 저소득군이 선호하는 여행유형을 누적그래프를 그려 알아보았다.
- 고소득군은 [가족여행, 연애여행, 출장]을, 저소득군은 [친목여행, 효도여행, 혼자여행]을 선호하는 경향을 알 수 있었다. 특이한 점으로는 평균숙박비용을 분석하였을 때 상위권에 존재했던 [효도여행]이 저소득군이 선호하는 여행유형으로 나타난 것을 알 수 있었다.

(3) 나이대별 여행유형 분석



- 나이대별 여행유형은 30-50 중장년층이 언제나 가장 여행을 많이 가는 경향을 보였다. 중장년층 비중이 많은 이유는 경제적으로 안정되어 있고, 시간적인 여유가 많기 때문이라고 예상할 수 있다.

나이대별 여행유형 비중



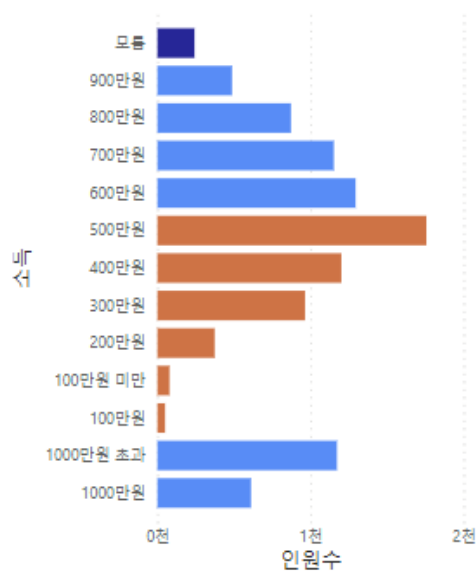
- 나이대별 여행유형선호도를 알아보기 위하여 100 퍼센트 누적그래프를 그려보았다.
- 가족여행의 경우 40대에서 압도적으로 많은 여행수를 보였다. 이는 자녀의 나이대와도 관련이 있을 것이라 생각하여 이후 분석하였다.
- 친목여행의 경우 젊은 나이에서 많이 갈 것으로 예상했지만 의외로 40-60 나이대에서 많이 여행했음을 알 수 있었다.
- 혼자여행과 효도여행의 경우 나이가 증가하면서 크게 비율이 내려감을 알 수 있었다. 혼자여행과 효도여행은 저소득군의 선호도가 높은 여행유형이었으므로 젊은 층의 소득군이 낮음을 짐작할 수 있었다.

3-3. 여행자 중심의 접근

(1) 소득군, 직업군별 경향성

소득군별 인원

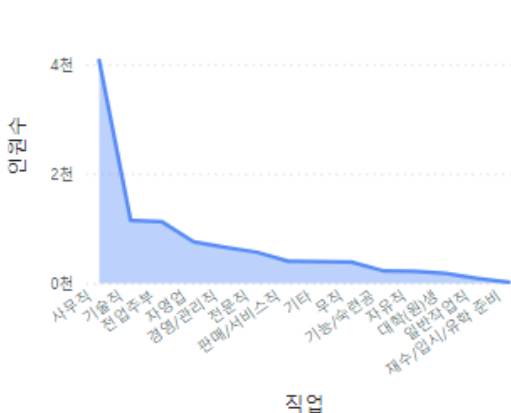
소득군 ● 고소득 ● 모름 ● 저소득



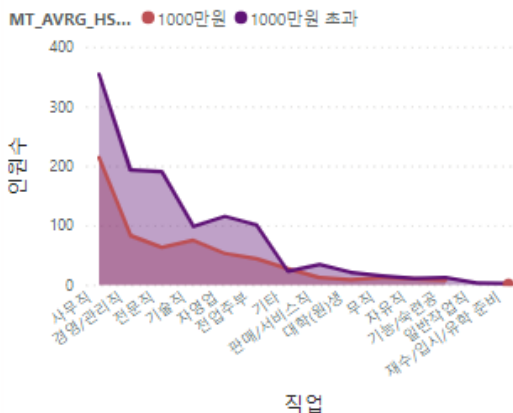
소득군		인원수
저소득	100 만원 미만	79
	100 만원	50
	200 만원	374
	300 만원	964
	400 만원	1201
	500 만원	1756
	600 만원	1294
고소득	700 만원	1152
	800 만원	872
	900 만원	487
	1000 만원	611
	1000 만원 이상	1172

- 저소득층의 경우 소득이 낮아질수록 여행빈도가 낮아지는 경향성을 보였고, 반대로 고소득층의 경우 1000 만원 미만까지는 소득이 높아질수록 여행빈도가 낮아지는 경향성을 보이다 그 이상의 소득군부터 다시 빈도가 증가하기 시작하였다. 이러한 특징의 이유를 알아보기 위하여 1000 만원 이상의 소득을 버는 사람들을 초고소득군으로 분류하여 직업군을 살펴보았다.

전체 - 직업군

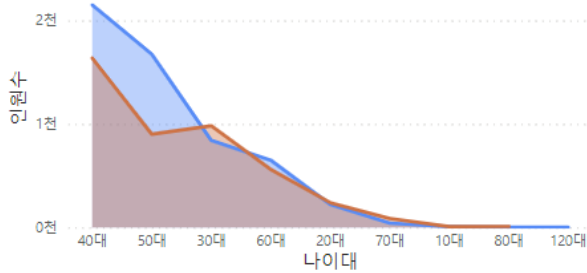


초고소득군 - 직업군

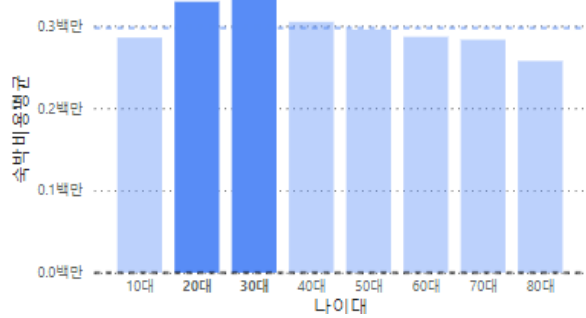


- 초고소득군의 직업은 전체직업군과 달리 [경영관리직, 전문직]의 비중이 큰 것을 알 수 있었다. 해당 직업은 휴가가 자유로운 편이므로 직업과 높은 여행빈도의 관계를 예상할 수 있었다.

소득군별 나이대
소득군 ● 고소득 ● 저소득



나이대별 숙박비용평균

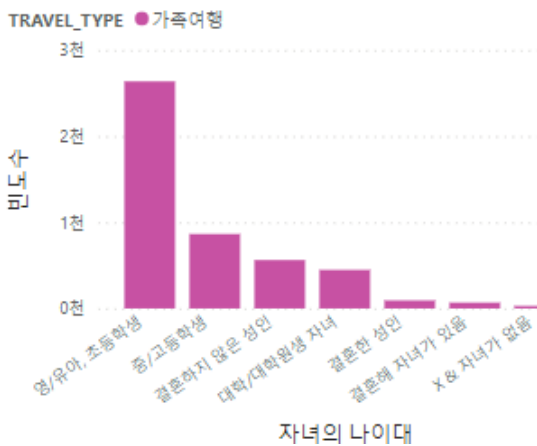


- 나이대별 소득 구간은 40-60 은 고소득이 20-30 은 저소득이 우세함을 알 수 있었다. 특히 40-50 은 고소득이 압도적으로 우세하였다.
- 하지만 나이대별 숙박비용 평균은 20-30 대가 다른 나이대에 비해 높은 것을 확인할 수 있었는데 이를 통해 20-30 대의 경우 벌이에 비해 여행에 많이 투자하는 경향이 있음을 알 수 있었다.

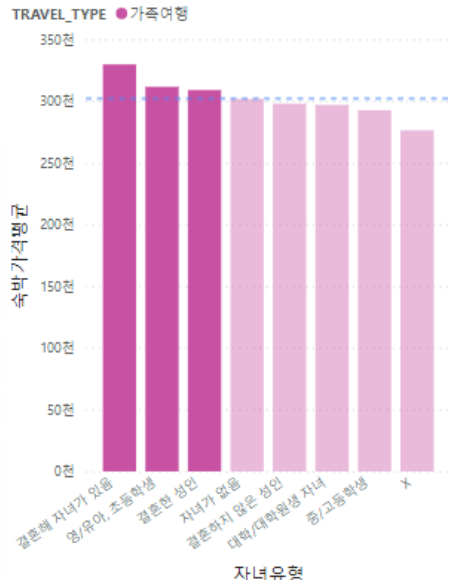
(2) 성별, 결혼여부 및 자녀의 나이대별 분석

a) 가족여행

가족여행 - 자녀의 나이대



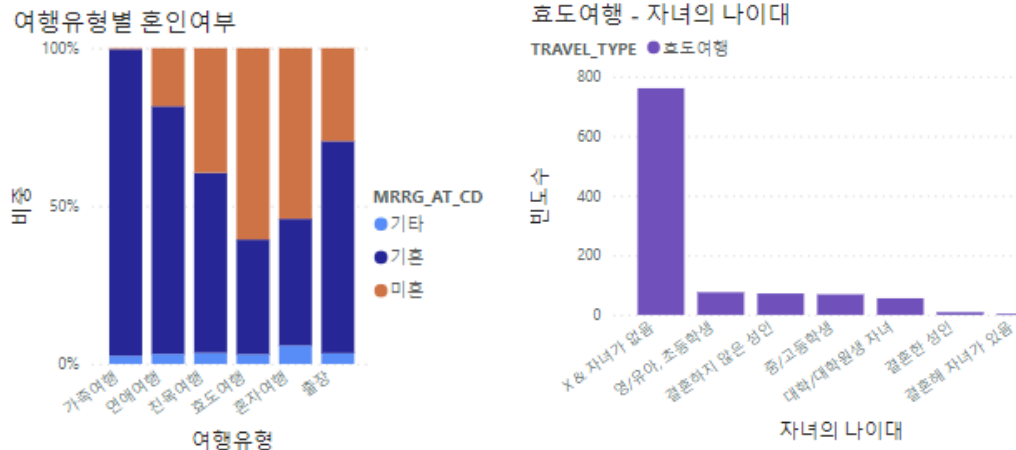
가족여행 - 자녀유형별 숙박가격평균



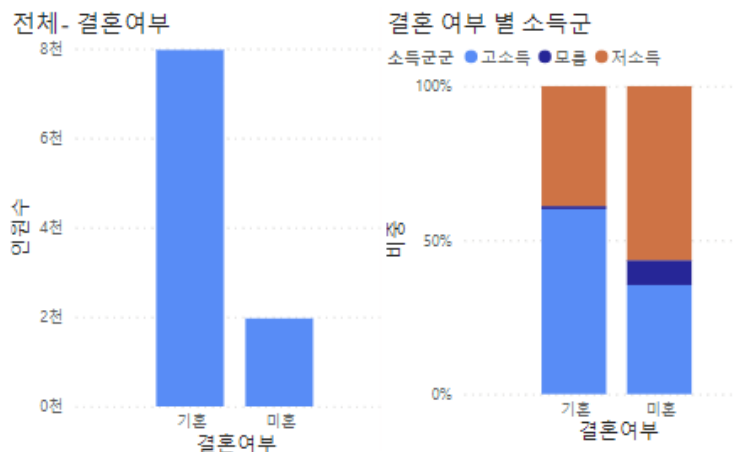
- 가족 여행에서 자녀의 나이대는 영/유아/초등학생, 중/고등학생, 결혼하지 않은 성인 순으로 많은 비중을 차지하였다. 특히 영/유아/초등학생 자녀가 많았는데 앞서 40대에서 가족여행을 압도적으로 많이 간 것과 같이 해석하면 40대의 부부의 경우 영/유아/초등학생 나이대의 자녀가 있는 것이 일반적일 것이고 그에 따라 가족여행을 많이 간 것으로 해석할 수 있었다.

- 평균 숙박 가격은 자녀가 있는 성인, 영/유아, 결혼한 성인 순으로 높은 것을 볼 수 있다. 자녀가 결혼한 경우 경제 활동을 하고 있을 가능성이 높아 여행에 투자를 많이할 수 있을것임을 예상할 수 있고, 자녀가 영/유아나 초등학생인 경우 추가적인 시설이나 서비스 등을 위해 비용을 더 지불하는 경향이 있음을 추측할 수 있었다.

b) 효도여행



- 앞서 보았듯 효도 여행은 고소득층보다 저소득층에서 많이 가는 모습이 보였다. 여행 유형별 혼인여부 그래프를 그려보니 효도여행의 경우 미혼의 비율이 가장 높았고, 자녀는 없는 경우가 제일 많았다.



- 기혼여부가 소득군과 관련이 있는지를 살펴보기 위해 그래프를 그려보니 전체 데이터에 있어 결혼 여부 별 소득군을 보았을 때, 기혼보다 미혼의 경우에 저소득 비중이 높아지는 모습을 볼 수 있다.
- 이를 바탕으로, 효도 여행에서 저소득 비중이 높은 이유는 미혼 인구가 많기 때문임을 이해할 수 있었다. 또, 미혼이거나 자녀가 없는 사람일수록 효도 여행을 떠나는 경우가 많다고 판단했다.

c) 출장

결혼 여부	성별	여행자 수
기타	M	5
기타	F	5
기혼	M	134
기혼	F	44
미혼	M	38
미혼	F	40

- 출장의 경우, 여성보다 남성의 수가 훨씬 많았고 기혼자의 수도 훨씬 많았다. 미혼의 경우에는 여성과 남성의 수가 비슷하지만, 기혼의 경우 남성의 수가 여성에 비해 약 3배 가량 되는 것을 보아 여성의 결혼, 출산을 통한 경력단절이 영향을 미친 것으로 보인다.

d) 혼자여행

결혼 여부	성별	여행자 수
기타	M	14
기타	F	5
기혼	M	110
기혼	F	20
미혼	M	87
미혼	F	88

- 혼자 여행의 경우 미혼, 기혼, 기타 순이며 기혼 남성의 여행자 수가 110 명으로 제일 많은 것을 확인할 수 있다.

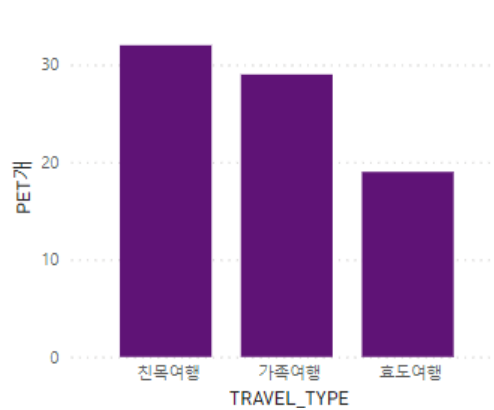
숙박업명	숙박업주소	숙박유형명	여행자 수
썰비치 호텔 삼척	강원	관광호텔업	17
유성호텔	대전	관광호텔	14
롯데 호텔 부산	부산	관광호텔업	12
신라스테이 삼성	서울	관광호텔업	12

- 방문 수가 많은 숙소는 썰비치 호텔 삼척, 유성호텔, 롯데 호텔 부산 순이었다. 혼자 여행은 바다와 가까운 곳이나 자연과 밀접한 곳의 숙소를 예상했지만 썰비치 호텔 삼척을 제외한 나머지 곳은 도심에 밀집한 곳에 위치하고 있다.
- 기혼 남성의 비중이 높은것과 숙소가 도심에 위치하고 있는 것으로 미루어 보았을 때, 혼자여행에 출장의 경우가 많이 포함됐음을 짐작할 수 있었다.

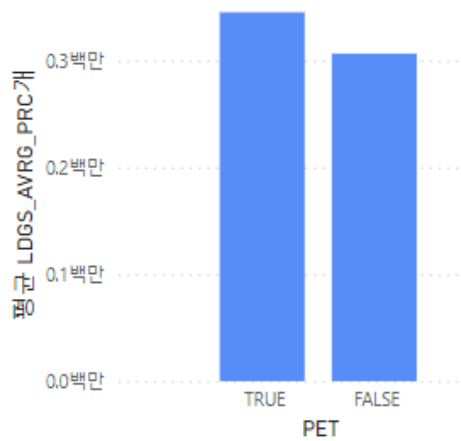
(3) 반려동물 유무에 따른 경향성

반려동물과 함께하는 여행

PET ● TRUE



반려동물 유무에 따른 숙박평균가격



- 주로 친목여행, 가족여행, 효도여행 순으로 반려동물을 동반하는 경우가 많은 것을 알 수 있다. 또한, 반려동물을 동반한 경우가 그렇지 않은 경우보다 평균 숙박가격이 높다고 볼 수 있다. 이는 반려동물과 머무를 수 있는 숙소가 더 비싼 편이거나 추가금액이 부과되는 것으로 설명할 수 있다.
- 하지만 반려동물과 함께 숙박한 경우는 전체 10,000 여 건 중 80 건밖에 해당되지 않았으므로 보아, 유의미한 인사이트를 도출하기엔 힘들어 보인다.

4. 결론과 기대효과

4-1. 결론

본 프로젝트는 국내 주요 숙박업소 데이터를 분석해 여행자별 맞춤형 숙박 시설 추천을 목적으로 수행했다. 분석 결과, 여행 유형, 나이대, 소득 군 등 다양한 요인이 여행지 선택과 숙박시설 선호도에 영향을 미치는 점을 확인할 수 있었다. 주요 결론은 다음과 같다.

1. 30 대를 중심으로 한 연애여행 마케팅

전체적으로 강원, 제주, 부산, 전남 등 바다와 가까운 주요 관광지에 대한 수요가 높은 모습을 볼 수 있다. 그러나 20-30 세대만 유독 서울과 인천에서 숙박하는 비율이 높았다. 서울과 인천의 숙박 업소는 고급 호텔 업종이 대다수를 차지한다는 점과, 30 대의 경우 다른 나이대에 비해 연애 여행에서 지불하는 금액이 가장 높은 것을 알 수 있다. 이를 바탕으로, 30 대 회원을 대상으로 기념일이나 특별 이벤트 패키지를 제공하는 마케팅 방안을 고안했다.

2. 나이와 소득 차이

500 만 원 이상의 고소득자 중 40 대와 50 대가 가장 큰 비중을 차지했다. 하지만 소득수준과 무관하게 여행에 많은 돈을 투자하는 것은 20-30 세대라는 사실을 발견할 수 있었다. 따라서 젊은 층엔 럭셔리, 중장년 층엔 자연경관을 무기로 한 가성비 여행을 중심으로 마케팅 방안을 마련해야 한다는 인사이트를 발견하였다.

4-2. 기대효과

1. 마케팅 효과 극대화

- 각 세대, 각 고객에 맞춘 서비스를 제공함으로써 고객의 만족도를 높이고 광고 효율성을 극대화할 수 있다.
- 청년 세대의 특성과 선호를 반영하면 소비를 촉진하고, 브랜드 인지도를 높일 수 있을 것이다.

2. 데이터 기반의 의사결정 강화

- 본 프로젝트를 통해 얻은 마케팅 전략의 효과성을 모니터링하고, 조정함으로써 지속적인 데이터 기반 의사결정 환경을 구축할 수 있다.
- 타지역, 타 세대를 대상으로 하는 마케팅에서 참고할 수 있는 성공 사례를 만들 수 있다.

3. 여행 산업 발전 기여

- 맞춤형 여행 경험 제공을 통해 국내 여행 수요를 활성화하고, 지역 경제 발전에 기여할 수 있을 것이다.
- 여행산업 전체의 서비스 품질 향상과 마케팅 전략 개발에 기여할 수 있을 것이다.