

I. 주제 설명

본 프로젝트의 주제는 <방학에 떠나는 나의 제주도 여행, 완벽한 여행을 위해선 어떤 숙소가 좋을까?>이다. 7월 31일부터 8월 2일까지 떠나는 제주도 여행을 위한 숙소를 고르는 여정으로, 수업에서 활용된 다양한 기법을 사용하여 최적의 숙소 찾아보는 것을 목표로 한다. 이때 최적이란, 예약 가능한 숙소들 중 여러가지 조건들을 포괄적으로 파악했을 때 가장 나에게 맞는 것으로 정의한다. 나에게 맞는 것에 대한 구체적인 설명은 아래 프로젝트 개요에서 진행할 예정이다. 목표를 위해 다양한 방법의 비교와 분석을 진행할 예정이며, 더 나아가 다른 숙소 예약 웹 사이트, 어플리케이션과의 차별점을 둔 숙소 예약 프로그램 가이드라인을 제공해볼 예정이다.

II. 주제 선정 이유

가장 큰 이유는 수업에서 학습한 내용들을 단순히 사용해보는 것에서 그치지 않고 결과물을 활용하여 실제 상황에 적용해보는 것이 나에게 정확한 목표와 성취감을 제공하는 새로운 경험이 될 것이라고 생각했기 때문이다. 마침 방학에 떠날 여행을 위한 숙소를 찾는 것에 힘을 쓰는 중이었기에 보다 완벽한 여행을 위한 프로젝트를 진행해보고자 다짐하였고, 현재 실제로 해당 프로젝트의 결과로 도출된 숙소를 예약한 상황이다.

또한 숙소를 찾아보며 답답함을 느꼈는데, 해당 부분을 해소하기 위한 분석을 진행해보고 싶었다. 여러 웹 사이트와 어플리케이션을 통해 숙소를 찾아보며 느꼈던 점은 다음과 같다. 첫째, 대부분의 브랜드는 여러가지 옵션을 선택하여 원하는 조건의 숙소를 확인하게 해주는 시스템을 제공한다. 그러나 해당 기능은 이미 정확히 원하는 조건이 있는 사람들에게 장점으로 작용한다. 만약 아직 선택한 장소와 시기에 대한 지식이 부족한 상태라면 적당한 조건을 선택하기 힘들다. 특히 "여행"이라는 것은, 익숙하지 않은 공간으로 떠나는 것이기에 한 번에 원하는 조건을 파악하기 쉽지 않다. 그러므로 원하는 조건을 알아보기 위해서는 결국 수많은 숙소들을 많이 찾아보고, 직접 비교해보는 수밖에 없다. 둘째, 사람들은 여행에서 추구하는 바가 모두 다르다. 같은 숙소를 이야기할 때에도 누구는 가성비 좋게 다녀왔다는 긍정적인 리뷰를 작성하고, 또 다른 누군가는 청결하지 못하다는 부정적인 리뷰를 작성하기도 한다.

이들에 대한 보완을 위해 본 프로젝트에서는 다음과 같은 분석을 진행하도록 한다. 첫째, 선택한 위치/날짜에서의 가격대 분포, 숙소 유형 분포, 숙소 유형별 가격대 분포, 평점 분포 등 많은 숙소들을 하나하나 확인해야만 확인할 수 있는 전체적인 분석 정보를 제공하여 원하는 조건을 논리적으로 선택할 수 있도록 돕는다. 둘째, 개개인이 추구하는 바를 적용한, 이른바 Personal Parameter를 적용하여 만족스러운 여행에 한발 더 다가갈 수 있도록 돕는다. 해당 내용의 자세한 이야기는 프로젝트 개요에서 진행할 예정이다.

III. 프로젝트 개요

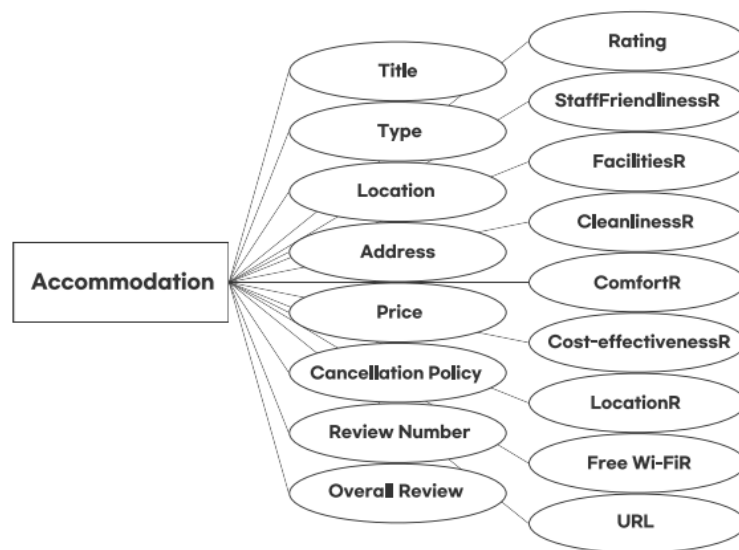
1. 데이터 수집

분석을 위한 데이터에 대한 조건은 다음과 같다.

- 1) 다양한 동네, 다양한 유형, 다양한 가격대의 숙소의 데이터가 존재해야 한다.
- 2) Personal Parameter 를 위해 상세한 판단의 척도 데이터가 존재해야 한다.

해당 조건을 만족시키는 데이터를 얻고자 찾아본 결과, 부킹닷컴(<https://www.booking.com>)이라는 웹 사이트를 발견할 수 있었다. 해당 사이트는 다양한 조건의 숙소들을 제공하며 직원 친절도, 시설, 청결도, 편안함, 가성비, 위치, 무료 Wi-Fi 의 7 가지 상세한 리뷰 점수를 제공한다. 위 사이트의 다양한 정보들을 웹 크롤링을 통해 수집하여 원하는 분석을 진행해볼 예정이다.

데이터베이스 구축을 위해 필요한 정보의 다이어그램은 다음과 같다.



<사진 1 - 숙소 관련 데이터 다이어그램>



<사진 2 - Personal Parameter 관련 데이터 다이어그램>

이때, Personal Parameter 관련 데이터는 숙소 관련 데이터의 분석을 통해 추가로 생성되는 값들을 포함하도록 한다.

2. 데이터베이스 구축

Booking 데이터 베이스를 생성하여 수집한 데이터를 기반으로 테이블들을 csv 파일로 불러올 예정이다. 수집을 통해 구축할 테이블은 숙소 기본 정보 테이블, 숙소 리뷰 테이블, 숙소 상세 테이블로 구성할 것이며, 추가적으로 분석에 필요한 동네 유형 테이블과 숙소 유형 테이블을 구축하고, Personal Parameter 를 적용하기 위한 선호 숙소 테이블을 구축한다. 이때, 크롤링한 데이터는 결측치와 이상치가 존재할 수 있기에 Python 을 사용하여 이들을 처리한 후 활용하고자 한다.

3. 데이터 분석

첫째로 수행할 "나에게 맞는 조건 찾기"를 위한 분석 내용은 다음과 같다.

- 1) Location 분석: 동네별 숙소 수 파악
- 2) Price 분석: 가격 분포 파악
- 3) Type 분석: 숙소 유형별 숙소 수 파악
- 4) Location-Price: 동네별 가격대 파악
- 5) Type-Price: 숙소 유형별 가격 분포 파악
- 6) Type-Location: 동네별 숙소 유형 분포 파악
- 7) Rating: 별점 분포 파악
- 8) Review Number: 리뷰 수 분포 파악
- 9) Rating-Location: 지역별 별점대 파악
- 10) Rating-Type: 숙소 유형별 별점대 파악
- 11) Rating-Price: 별점과 가격간의 상관관계 파악

둘째로 수행할 "Personal Parameter 를 이용한 최적 숙소 찾기"를 위한 분석 내용은 다음과 같다. (이때, 위 분석을 통해 정해진 조건에 모두 해당하는 숙소들의 집합인 선호 숙소 테이블에 한하여 분석이 이루어진다.)

- 1) Rating, Review Number, 7 가지 상세 별점 중 우선시하는 조건 선정
- 2) 선정된 조건들로 각각 선호 숙소 테이블에 존재하는 숙소들끼리 순위 정보 저장
- 3) 해당 순위들을 모두 합하여 더한 값을 최종 Personal Parameter 정보 저장

해당 분석을 진행한 후, 선호 숙소 테이블에서 최소 Personal Parameter 를 가지는 값을 최적의 숙소로 선정한다.

IV. 프로젝트 수행 내용

프로젝트 수행 내용은 대부분 코드에서 비롯되지만, 코드 같은 경우 추가적으로 첨부하기에 결과 도출 쿼리문을 제외하고는 수행 과정과 결과 위주로 수행 내용을 설명하고자 한다. (코드 파일 또한 주석으로 상세 내용이 정리되어 있다.)

1. 데이터 수집 수행 내용

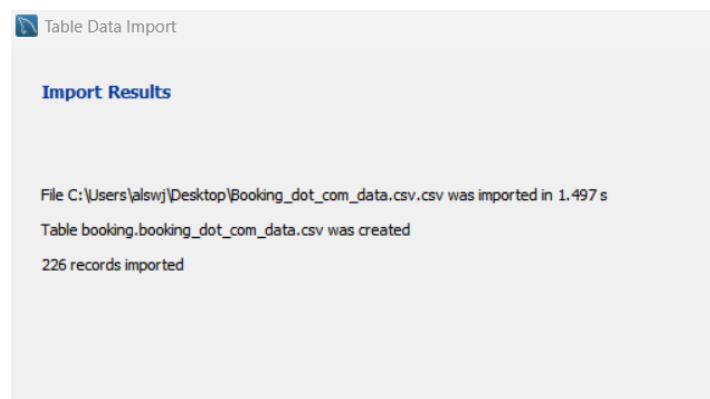
구글 Colab 환경에서 Python 을 활용하여 웹 크롤링을 진행한다. <http://www.booking.com> 에서 제주도, 2023.07.31~2023.08.02, 성인 2명, 객실 1개를 검색하여 얻은 정보들을 수집했으며, 크롤링 일자 는 2023.06.08이다.

	Title	Type	Location	Address	Price	Cancellation Policy	Review Number	Overall Review	Rating	Staff	Friendliness	Facilities	Cleanliness	Comfort	Cost-effectiveness	Location	Free Wi-Fi	URL
0	롯데시티호텔 제주	호텔	제주 시내, 제주	제주, 도항로 83	395808	무료취소	1677	제주 좋음	8.4		9.1	8.8	8.9	9.0	8.3	9.0	8.9	https://www.booking.com/hotel/kr/lotte-city-jeu-
1	오션스위츠 제주 호텔	호텔	제주 시내, 제주	제주, 탑동 해안로 74	209622	무료취소	1131	좋은	7.9		8.8	8.1	8.2	8.3	8.2	9.0	8.7	https://www.booking.com/hotel/kr/ocean-suites-jeju-
2	라마다 제주시티호텔	호텔	제주 시내, 제주	제주, 중앙로 304	218736	무료취소	2696	제주 좋음	8.0		8.6	8.3	8.4	8.3	8.7	8.6	8.8	https://www.booking.com/hotel/kr/ramada-jeju-city-
3	호텔 시리우스	호텔	제주 시내, 제주	제주, 도항로 133	208320	무료취소	1045	제주 좋음	8.2		8.6	8.5	8.6	8.5	8.5	8.8	8.8	https://www.booking.com/hotel/kr/sirius-ko.html
4	제주 아일라 호텔 & 스파	호텔	한림읍, 제주	제주, 한림읍 용포길 10	359352	무료취소	618	제주 좋음	8.0		8.3	8.2	8.3	8.4	7.9	8.4	8.3	https://www.booking.com/hotel/kr/agulia-ko.html
253	전통리조트	리조트	한림읍, 제주	제주, 한림읍 일주서로 4010-2	242172	무료취소	외부 사이트 716개	최고 9.0										https://www.booking.com/hotel/kr/gentle-resort-
254	아프리카 게스트하우스	게스트하우스	조천읍, 제주	제주, 조천읍 신동로2길 33 (조천읍)	82442	무료취소	22	좋은	7.9		9.1	7.0	7.5	7.3	8.6	8.7	9.1	https://www.booking.com/hotel/kr/apeurillage-jeju-
255	엑스온 게스트하우스	게스트하우스	제주 시내, 제주	제주, 서해안로 232	108066	무료취소	28	좋은	7.9		8.4	7.9	7.8	7.5	8.1	8.3	5.0	https://www.booking.com/hotel/kr/yejun-guesthouse-jeju-

<사진 3 – 초기 수집 데이터>

수집한 데이터의 열은 숙소 이름, 숙소 유형 등을 포함한 17 가지 정보이며 최종 예약을 돕기 위한 해당 숙소의 예약 페이지 URL 을 포함한다. 웹 크롤링 결과로 조건에 맞는 257 개의 숙소를 수집할 수 있었다.

그러나 해당 데이터프레임을 MySQL 로 연동해보니, 226 개의 행만 import 되는 것을 확인할 수 있었다.



<사진 4 – MySQL Import 창>

이상치와 결측치 때문인데, 이를 수정하고 보다 분석에 용이한 데이터를 생성해내고자 구글 Colab 환경에서 Python 을 활용하여 이들을 처리해주는 과정을 수행하였다. 첨부된 코드를 확인하면, 다음과 같은 내용을 수행한 것을 확인할 수 있다.

1) 데이터 파악

- : Review 관련 / 신규 숙소의 경우 결측치가 많음
- : Cancellation Policy / 무료 취소 정책이 적용되지 않을 경우, 값 자체가 존재하지 않음
- : Review Number / object 형으로 존재 (리뷰 개수를 의미하므로 int 형이 적절)

2) 결측치 처리

: Review 별점 관련 / 신규 숙소의 경우 0 삽입

: Overall Review / “신규 숙소” 추가

: Cancellation Policy / 무료 취소 가능 여부 O 혹은 X로 수정

3) 데이터 타입 수정

: Review Number / object 형을 알맞게 처리하여 모두 int 형으로 수정

4) Location 일반화

: 지역, 제주로 저장된 Location의 내용을 지역 명만 남도록 수정

5) Overall Review 일반화

: 값이 이상한 것을 확인할 수 있음

: Overall Review의 범위를 파악하고, 해당 범위에 맞는 Overall Review 대입

6) 중복 데이터 처리

: URL이 동일하므로 중복 데이터 삭제

위 사항을 모두 적용한 데이터 프레임을 생성하여 최종적으로 데이터 수집을 마쳤다. 아래 테이블 설계 내용에 따라 숙소 기본 정보 테이블, 숙소 리뷰 정보 테이블, 숙소 상세 정보 테이블의 세가지 csv 파일로 저장하여 import 해오는 방식으로 활용하였다.

2. 데이터베이스 구축 수행 내용

수집한 데이터들을 체계적으로 활용하기 위한 테이블 설계 내용은 다음과 같다. (상위 5개의 내용을 포함한다.)

- 숙소 기본 정보 테이블 : basic_info

Aid	Title	Type	Location	Price
0	롯데시티호텔 제주	호텔	제주 시내	395808
1	오션스위츠 제주 호텔	호텔	제주 시내	209622
2	라마다 제주시티호텔	호텔	제주 시내	218736
3	호텔 시리우스	호텔	제주 시내	208320
4	제주 아길라 호텔 오시아노 스위트	호텔	한림읍	359352

<사진 5 – 숙소 기본 정보 테이블>

숙소의 기본 정보를 확인하기 위한 테이블이며, 숙소 id, 숙소 명, 숙소 유형, 동네, 가격을 포함한다. 숙소 id를 의미하는 Aid를 기본키로 한다.

- 숙소 리뷰 정보 테이블 : review_info

Aid	Title	Review Number	Overall Review	Rating	StafffriendlinessR	Facilities	Cleanliness R	ComfortR	Cost-effectivenessR	LocationR	Free Wi-FiR
0	롯데시티호텔 제주	1677	매우 좋음	8.4	9.1	8.8	8.9	9	8.3	9	8.9
1	오션스위츠 제주 호텔	1131	좋음	7.9	8.8	8.1	8.2	8.3	8.2	9	8.7
2	라마다 제주시티호텔	2696	매우 좋음	8	8.6	8.3	8.4	8.3	8.7	8.6	8.8
3	호텔 시리우스	1045	매우 좋음	8.2	8.6	8.5	8.6	8.5	8.5	8.8	8.8
4	제주 아길라 호텔 오시아노 스위트	618	매우 좋음	8	8.3	8.2	8.3	8.4	7.9	8.4	8.3

<사진 6 – 숙소 리뷰 정보 테이블>

숙소의 리뷰와 관련된 정보들을 확인하기 위한 테이블이며, 숙소 id, 숙소 명, 리뷰 개수, 총평, 총 별점, 직원 친절도 별점, 시설 별점, 청결도 별점, 편안함 별점, 가성비 별점, 위치 별점, 무료 Wi-Fi 별점을 포함한다. 숙소 id를 의미하는 Aid를 기본키로 하며, 동시에 basic_info와 연결되기 위한 외래키로 한다.

- 숙소 상세 정보 테이블 : detail_info

Aid	Title	Address	Price	Cancellation Policy	Review Number	Overall Review	URL
0	롯데시티호텔 제주	제주, 도령로 83	395808	O	1677	매우 좋음	https://www.booking.com/hotel/kr/...
1	오션스위츠 제주 호텔	제주, 탐동 해안로 74	209622	O	1131	좋음	https://www.booking.com/hotel/kr/...
2	라마다 제주시티호텔	제주, 중앙로 304	218736	O	2696	매우 좋음	https://www.booking.com/hotel/kr/...
3	호텔 시리우스	제주, 도령로 133	208320	O	1045	매우 좋음	https://www.booking.com/hotel/kr/...
4	제주 아길라 호텔 오시아노 스위트	제주, 한림읍 옹포 2길 10	359352	O	618	매우 좋음	https://www.booking.com/hotel/kr/...

<사진 7 - 숙소 상세 정보 테이블>

선정된 숙소의 URL 을 포함한 상세 정보를 제공하기 위한 테이블이며, 숙소 id, 숙소 명, 상세 주소, 가격, 무료 취소 정책 여부, 리뷰 개수, 총평, URL 을 포함한다. 숙소 id 를 의미하는 Aid 를 기본키로 하며, 동시에 basic_info 와 연결되기 위한 외래키로 한다.

- 동네 유형 테이블 : location_info

Location
제주 시내
한림읍
기타 도시
애월읍
구좌읍

<사진 8 - 동네 유형 테이블>

데이터 분석을 위한 테이블이며, basic_info 의 동네열에 해당하는 종류를 포함한다.

- 숙소 유형 테이블 : type_info

Type
호텔
리조트
게스트하우스
홀리데이 홈
레지던스

<사진 9 - 숙소 유형 테이블>

데이터 분석을 위한 테이블이며, basic_info 의 숙소 유형열에 해당하는 종류를 포함한다.

- 선호 숙소 테이블 : want_info

Aid	Review Number Rank	Rating Rank	StaffFriendliness Rank	Facilities Rank	Cleanliness Rank	Comfort Rank	Cost-effectiveness Rank	Location Rank	Free Wi-Fi Rank	Final Parameter
11	1	3	0	0	3	0	0	2	0	9
50	2	4	0	0	4	0	0	5	0	15
87	5	6	0	0	6	0	0	7	0	24
103	6	2	0	0	2	0	0	4	0	14
104	7	8	0	0	7	0	0	3	0	25

<사진 10 - 선호 숙소 테이블>

조건 설정 후 Personal Parameter 의 값을 도출해내기 위한 테이블이며, 숙소 id, 리뷰 개수 순위, 별점 순위, 직원 친절도 순위, 시설 순위, 청결도 순위, 편안함 순위, 가성비 순위, 위치 순위, 무료 Wi-Fi 순위를 포함한다. 숙소 id 를 의미하는 Aid 를 기본키로 하며, 동시에 Review_info, Detail_info 와 연결되기 위한 외래키로 한다.

3. 데이터 분석 수행 내용

"나에게 맞는 조건 찾기"를 위한 분석 수행 내용은 다음과 같다.

- Location 분석 (basic_info)

동네별 숙소 수 파악을 진행하였다.

숙소의 양이 많지 않은 한경면, 우도를 원하는 숙소에서 제외하기로 한다.

- Price 분석 (basic_info)
 - Price 의 최솟값, 최댓값, 평균 및 5 만원 구간의 가격대 분포 파악을 진행하였다.
 - 대부분의 숙소가 존재하는 40 만원 이하의 가격 구간을 정하기로 한다.
 - Type 분석 (basic_info)
 - 숙소 유형별 숙소 수 파악을 진행하였다.
 - 숙소의 양이 적은 비앤비, 빌라, 모텔은 원하는 숙소에서 제외하기로 한다.
 - Location-Price (basic_info)
 - 동네별 가격대 파악을 진행하였다.
 - Type-Price (basic_info)
 - 숙소 유형별 가격 분포 파악을 진행하였다.
 - 공동 생활이 필요한 호스텔과 게스트하우스는 상대적으로 저렴한 것으로 파악된다.
 - 그러나 이번 여행에서 공동 생활을 바라지는 않으므로 홀리데이 홈을 포함한 이들은 원하는 숙소에서 제외하기로 한다.
 - Type-Location (basic_info)
 - 동네별 숙소 유형 분포 파악을 진행하였다.
 - 제주 시내가 전체적으로 다양한 숙소가 존재하는 것을 확인할 수 있었다.
 - Rating (review_info)
 - 별점 분포 파악을 진행하였다.
 - 분포를 보아, 무난할 경우 6 정도의 점수를 지닌다고 파악되었다.
 - 별점이 6 이상이라는 구간을 정하기로 한다.
 - Review Number (review_info)
 - 리뷰 수 분포 파악을 진행하였다.
 - 분포를 보아 0~100 사이가 가장 많은 것을 확인할 수 있었다.
 - 그러나 최소한으로 숙소를 판단할 수 있는 10 이상정도의 구간을 정하기로 한다.
 - Rating-Location (review_info-basic_info)
 - 지역별 별점대 파악을 진행하였다.
 - 별점이 좋지 못한 한경면, 구좌읍을 원하는 숙소에서 제외하기로 한다.
 - Rating-Type (review_info-basic_info)
 - 숙소 유형별 별점대 파악을 진행하였다.
 - 별점이 좋지 못한 컨트리하우스, 아파트를 원하는 숙소에서 제외하기로 한다.
 - 리뷰가 존재하지 않는 빌라, 비앤비를 원하는 숙소에서 제외하기로 한다.
 - Rating-Price (review_info-basic_info)
 - 별점과 가격간의 상관관계 파악을 진행하였다.
 - 가격대가 어느정도 있어야 숙소의 질이 좋아진다는 것을 확인할 수 있었다.
 - 일정한 증감을 통해 상세 리뷰의 신뢰도가 높은 편임을 확인할 수 있었다.
- 전체적인 분석을 통해 정해진 조건들을 모두 만족하는 숙소들만 담긴 want_info 테이블을 생성한다. 총 105 개의 숙소가 존재하였다.

```
CREATE TABLE want_info AS
SELECT b.Aid
FROM basic_info b
JOIN review_info r ON b.Aid = r.Aid
WHERE b.Location IN ('제주 시내', '애월읍', '조천읍', '한림읍', '기타 도시')
AND b.Type IN ('호텔', '리조트', '레지던스')
AND b.Price <= 600000
AND r.Rating >= 6
AND r.`Review Number` >= 10;
```

<사진 11 – 선호 숙소 테이블 생성 쿼리문>

105 개의 숙소 중 "Personal Parameter 를 이용한 최적 숙소 찾기"를 위한 분석 내용은 다음과 같다.

- Rating, Review Number, 7 가지 상세 별점 중 우선시하는 조건 선정
여행을 가는 나와 친구의 취향을 반영하여, 우선시하는 조건을 리뷰 개수, 별점, 청결도, 위치로 선정한다. (이때, 해당 분석에서는 필요한 열만 want_info 에 추가하기로 한다.)
- 선정된 조건들로 각각 선호 숙소 테이블에 존재하는 숙소들끼리 순위 정보 저장
- 해당 순위들을 모두 합하여 더한 값을 최종 Personal Parameter 정보 저장
해당 내용들을 수행한 후, 최종 결과 순 상세 정보를 제공(출력)한다.

```
SELECT d.*
FROM detail_info AS d
INNER JOIN want_info AS w ON d.Aid = w.Aid
ORDER BY w.FinalParameter ASC;
```

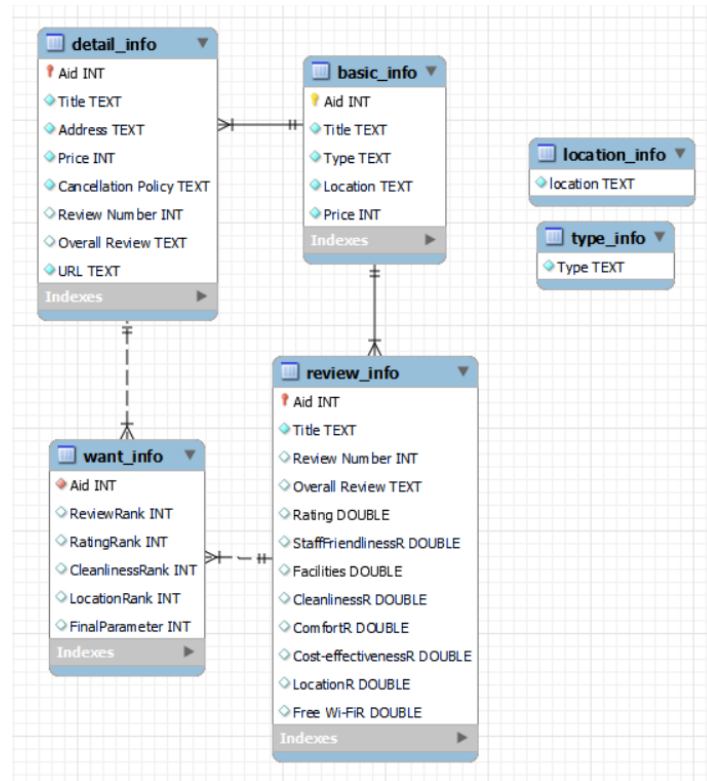
<사진 12 – 최종 결과 순 상세 정보 제공 쿼리문>

	Aid	Title	Address	Price	Cancellation Policy	Review Number	Overall Review	URL
▶	12	벤티모 호텔 앤 레지던스 제주	제주, 연상로 14	217434	O	1080	우수함	https://www.booking.com/hot
	16	호텔 레오	제주, 삼우로 14	217434	O	845	우수함	https://www.booking.com/hot
	0	롯데시티호텔 제주	제주, 도령로 83	395808	O	1677	매우 좋음	https://www.booking.com/hot
	7	호텔 리전트 마린 더블루	제주, 서부두2길 20	239568	O	1946	매우 좋음	https://www.booking.com/hot

<사진 13 – 사진 11 출력 예시>

직접 선정한 기준을 바탕으로 선정된 숙소들의 상세 정보를 확인할 수 있었다.

최종적으로 분석에 사용된 데이터베이스의 EER 다이어그램은 다음과 같다.



<사진 14 – EER Diagram>

V. 프로젝트 결과

다른 웹 사이트, 어플리케이션과 차별점을 두어 선택한 위치/날짜의 전반적인 정보를 분석함으로써, 최소한의 숙소 확인(상세 페이지 확인)으로 숙소를 찾을 수 있었다. 또 직접 고안한 Personal Parameter 의 방식을 활용하여 나의 원하는 바를 충족시키는 최적의 숙소를 찾을 수 있었다. 선정된 숙소는 "벤티모 호텔 앤 레지던스 제주"로, 제주 연삼로에 위치한 숙소이다. 상세 페이지에 접속하여 확인해보니, 깔끔하고 넓은 시설을 보유하고 있으며, 위치가 매우 좋은 것으로 판단되어 해당 숙소로 선정하여 예약을 마친 상태이다.

또한 프로젝트를 통해 수업에서 학습한 웹 크롤링 기법과 MySQL 의 다양한 문법(foreign key, pivot, 집계 함수, join, procedure, dynamic query 등)들을 활용하며 내용들에 대해 친숙해지는 경험이 되었다. 특히나 내가 원하고자 하는 결과를 위해 수정 및 보완의 과정을 진행하였기에 재미를 가지고 프로젝트를 진행할 수 있어 실무적 경험 학습에 도움이 크게 되었다.

VI. 한계점 및 개선점

가장 먼저, 데이터를 다뤄본 경험이 많지 않아 분석에 있어 더디게 진행되었다는 것이 아쉬웠다. 그러나 본 프로젝트를 통해 성장함을 직접 느끼기에, 웹 크롤링 혹은 SQL 을 활용한 다음 프로젝트를 진행한다면 훨씬 수월하게 진행할 수 있을 것으로 예상된다. 또한 제공받은 데이터가 아니라 직접 웹 크롤링을 통해 수집한 데이터임과 동시에, 실제 특정 기간의 예약 가능한 제주도의 숙소 데이터이다 보니 다루는 데이터 자체의 수가 적어 아쉬웠다.

추가적으로 이번 프로젝트에서는 아이디어를 바탕으로 한 데이터 수집/분석에 초점을 두었지만, 동일한 가이드라인을 따르는 숙소 예약 시스템을 구축하여 웹 사이트를 개발하는 형식의 프로젝트를 진행해보아도 좋을 것 같다.

VII. 참고 자료

세종대학교 2023-1 데이터베이스(001), 강의 교안

부킹닷컴, <https://www.booking.com>

Tstory 블로그, [MySQL] Workbench 다이어그램 그리는 방법(EER Diagram), 2020.05.14 13:49, 2023.06.24 15:32, <https://narup.tistory.com/52>