

PROJECT REPORT			
MID-PROJECT			
팀 개요			
팀명	사조참치	프로젝트 기간	
역할 분담			
번호	이름	분담 내용	역할
1	안지현	유사도 분석, 중간발표, 추천 알고리즘, 발표자료	팀장
2	김나리	유사도 분석, 추천 알고리즘, 최종발표	팀원
3	김동수	데이터 전처리, 텍스트 데이터 정제	팀원
4	최선빈	데이터 전처리, 유사도 분석, 시각화, 프론트엔드	팀원
5	허수정	데이터 전처리, 시각화(워드클라우드), 발표자료	팀원
6			팀원
프로젝트 계획(팀별)			
프로젝트 주제와 목적			
<p><주제> 호텔 리뷰 데이터 및 여행 특성 기반 프랑스 파리 호텔 추천 시스템</p> <p><목적> Booking.com에서 수집된 유럽 전역의 럭셔리 호텔에 대한 정보가 담긴 엑셀 데이터를 전처리하고, 고객의 여행 특성에 맞는 호텔을 추천해주는 알고리즘을 제작함. 2021년 10월 15일자로 전체국가에 대한 여행경보의 단계별 위험 수준에 관한 고시 일부가 '출국권고'로 개정되어 코로나로 인해 불가능했던 여행이 가능해짐. 관광지식정보시스템의 보고서에 따르면 2021년 3.4분기 기준 코로나로 인해 하지 못했던 여가활동 1순위가 '해외여행'이며, 유럽 중에서는 서유럽 방문 의향이 강함. 이에 따라 서유럽 국가 여행이 증가할 것으로 보아 그 중 프랑스 파리를 선정함</p> <p>활용 데이터 및 출처</p> <p><활용데이터> 유럽 럭셔리 호텔별 고객 리뷰 및 평점, 위도 및 경도, 고객의 이용형태별 태그, 2021~2022 1인당 평균 소비액(출처: Kaggle, Google Maps, 관광지식정보시스템, 통계청, 외교부 해외안전여행)</p>			
예상 결과물			
<ol style="list-style-type: none"> 1. INPUT으로 사용자의 선호 정보를 받는다. 친절도, 접근성, 조식 퀄리티 등을 생각중이다. 2. 개발한 알고리즘은 추천 호텔 리스트를 데이터프레임으로 반환한다. 3. TF / IDF, Cosine similarity를 통해 (2)의 리스트와 리뷰가 비슷한, 즉 투숙객이 생각하기에 비슷한 호텔 또한 계산하여 반환한다. 4. OUTPUT으로 호텔 리스트, 호텔 위치 등이 반환될 것이다. 			
프로젝트 결과(팀별)			
프로젝트 진행 내용			
<p>프로젝트 진행 flow)</p> <ol style="list-style-type: none"> 1) 김나리 : TF / IDF와 Cosine similarity를 통한 리뷰간 유사도 분석, 추천 알고리즘 개발 2) 김동수 : EDA를 통한 인사이트 도출 및 적용, 데이터 전처리 및 관리 3) 안지현 : TF / IDF와 Cosine similarity를 통한 리뷰간 유사도 분석, 추천 알고리즘 개발 4) 최선빈 : Folium을 활용한 지도 데이터 시각화, HTML을 이용한 데이터 시각화 5) 허수정 : 리뷰 속 단어의 빈도수를 통해 워드 클라우드 제작. 통계청을 통한 통계 자료 및 발표 자료 준비 			
결론(향후 방안 및 활용성)			
<p>여행특성, 객실, 청결, 직원, 친절, 위치 및 접근성, 조식, 인테리어 등으로 호텔을 추천하는 프로그램을 개발하였다. 질적인 측면에서, 결과는 꽤나 유의미하게 보인다. 특히 특정 호텔에서 긍정 리뷰 유사도가 높은 호텔들이 지역 근처에 위치해있거나, 가격 및 수준이 비슷한 호텔이 추천되었다. 따라서, 우리 팀은 추천 프로그램을 유용하다고 판단하고, 차후 다른 데이터 분석 과제나 프로그램 개발에 활용하려고 한다.</p>			
프로젝트 회고(개인별)			
기술력 성장에 대한 회고			
<p>협업능력 및 커뮤니케이션 성장에 대한 회고</p>			