

2023년 1학기 숭실대학교 AI융합학부 데이터사이언스 과제 공지

- 데이터 전처리 작업을 수행하고 기술통계, 상관관계 분석, 시각화 방법을 통해 데이터에 담겨있는 유의미한 패턴을 발견하여 보고서 제출
- 포맷: Jupyter notebook (ipynb) 으로 작성하되 pdf로 변경하여 제출(최대 10장)
코드, 출력 (데이터 출력, 시각화 결과 등), 블록 단위 분석 설명(텍스트), 결과 해석(텍스트) 등 포함
- 평가 기준: soundness (6점, 기술적 타당성)
presentation (4점, 얼마나 효과적으로 표현하였는지)
- 유의사항
 - 5주차 강의까지 다루지 않은 방법을 사용할 경우 해당 내용은 평가 제외
 - 분석 및 시각화 방법을 적재적소에 사용하여야 하며 단순히 분량이 많으면 감점
- 과제 만점자의 경우 학점 등급 + 기회 획득 (공지 확인 바람)
- 데이터 명세
 - 4669 rows, 9 columns
 - ID_USER: user id
 - USER_STATE: user가 있는 주 이름
 - USER_TIMEZONE: user가 있는 곳의 time zone
 - ID_HOTEL: hotel id
 - HOTEL_CITY: hotel이 있는 도시 이름
 - HOTEL_STATE: hotel이 있는 주 이름
 - HOTEL_TIMEZONE: hotel이 있는 곳의 time zone
 - Trip Type: 여행 유형
 - Rating: 평가 점수

- PDF 변환 방법

