**프로젝트 기획안**

**2021년 8 월 9 일**

**과정명 : 데이터 사이언스/엔지니어링 전문가**

|  |  |
| --- | --- |
| 팀 명 | 사여일남 |
| 팀 원 | 팀장: 이새봄  팀원:강민정, 김선재, 방병제, 여정문 |
| 프로젝트 주제 | 사회적 거리두기 단계에 따른 연애에 대한 관심도와  넷플릭스 내 로맨스 장르 프로그램 시청률의 상관관계 분석 |
| 프로젝트 수행 방향 | 데이터 설명   * Flixpatrol * Google 검색량 trend 변화   주요분석 내용   * 넷플릭스 내 로맨스 장르의 시청률 순위 추이 * 구글 검색어 중 “연애”, “거리두기”의 검색률 변화 추이   기대 효과   * 사회적 거리두기 단계 변화에 따른 인기 장르 예측 가능 * 연애를 하기 위한 최적의 시기 예측 가능 |
| 역할 분담 및 일정 | ● 역할분담 (변경 가능)  강민정 : 데이터(엑셀)  김선재 : 데이터(엑셀)  방병제 : 웹크롤링  여정문 : 웹크롤링  이새봄 : 데이터(엑셀)  ● 일정  8/9 주제 수립  8/15자료 가공 및 정리  8/16~8/23 데이터 시각화 및 웹 형식 보고서 디자인  (8/19~8/23 중점적으로!)  8/24 프로젝트 발표 |
| 프로젝트 수행 도구 | ● 수집기술 패키지  - pandas  - numpy  ● 분석기술 패키지  - bs4  - selenium  - pandas  ● 시각화 기술 패키지  - matplotlib, seaborn(그래프 생성)  - oauth2 : Twitter 인증 Pakcage - pytagcloud (워드 클라우드 그림 생성)  - jpyepe1. KoNLP (자연어 처리)  - folium : 지도 |