크롤링 해서 데이터 확보한 뒤

판다스와 시각화 툴을 이용해서

기술 통계적으로 분석할 수 있는 게 어떤 게 있을까 고민해보기

기존 데이터 칼럼에서 피처 엔지니어링을 통해 새로운 칼럼을 생성하고 인사이트 발견하기

공모전/ 경진대회

COMPAS, 서울빅캠퍼스 / 캐글, 데이콘

뷰티풀숩 / 판다스 학습하고

프로젝트 아이디어 생각해보기

내일 오전 미팅

맷플롯립 시본

6 박정현 허권 한정현

오전 수업 오후 프로젝트 1주일 다음주 금요일 발표

추천 주제

로켓펀치

정글

멜론

트위터

인스타그램

네이버증권 + 기업정보

웹툰 댓글

맛집정보

공모전/ 경진대회

COMPAS, 서울빅캠퍼스 / 캐글, 데이콘

로켓펀치 정글 멜론 트위터

인스타그램은 조금 어려움(로그인해서 해야 함)

네이버 증권 데이터 / 유가 정보 / 기업 정보

웹툰 댓글 / 맛집 정보

판다스와 시각화 툴을 이용해서 기술 통계적으로 분석할 수 있는 게 어떤 게 있을까 고민해보기

크롤링도 연습하지만

가이드라인은 공모전과 경진대회 가이드라인을 참고하는 것이 좋다

기존 수상자들의 코드를 보자

컴파스 서울빅캠퍼스 데이콘 캐글

화합적인 태도로 잘 마치는 게 중요.

너무 어렵거나 쉬운 것은 피하기.

서로 맘상하지 않게 하기.

전제: 독자들의 만족도/관심도는 댓글 반응과 평점을 통해 알 수 있다.

1. 웹툰 헬퍼의 회차별 전체 댓글수 / 좋아요수 / 평점 / 평점 참여자수 /

베댓의 좋아요 싫어요 수를 크롤링해 데이터프레임을 만든다.

회차별 만족도 추이를 살펴보기 위해 필요한 분석들. > 그래프/시각화 요망

1. 평점 추이 > 웹툰 만족도 변화의 지표
2. 전체 댓글 수 추이 / 평점 참여자수 추이 > 관심도 변화의 지표
3. 베댓 좋아요 & 싫어요 수의 총합 > 관심도 변화의 지표
4. 베댓 싫어요 수 총합 추이 > 독자 만족/불만족 의견 일치 정도의 지표-1?
5. 베댓 싫어요/좋아요 비율의 추이 > 독자 만족/불만족 의견 일치 정도의 지표-2

> 가설 1: 웹툰이 흥하다 망하기 시작할 때는 베댓 싫어요/좋아요의 비율이 높다

> 가설 2: 웹툰이 흥하는 중이거나 망했을 때에는 긍정적인 베댓이든 부정적인 베댓이든   
 싫어요/좋아요의 비율이 낮다.

1. 등등

> 상기 작업이 끝나면 동일한 흥망을 이미 거친 웹툰 2개를 추가 선정하여 동일한 크롤링과 분석 시각화를 진행한다.

> 3개 정도의 웹툰에서 공통점을 뽑아낸 뒤 신의 탑과 같이 현재 논란이 불거지고 있는 웹툰의 향방에 대해 직관할 수 있는 비교 분석을 진행한다.