회의 시간 : 오전 9시 ~ 오후 6시

참석자 : 전원. 오전 불참 - 윤보람

작성자 : 윤보람

다음 회의 :

* 3시 30분 : 기획서 초안 작성
* 분석 주제 및 데이터 선정

1. 도메인
2. 분석 대상

**< 데이터 선정 기준 >**

1. 지금 진행되는 현실적인 데이터, 도메인 반영
2. 학습용 데이터 : feature 개수가 많은, feature 정의가 잘 되어있는 것, 전처리 되어 있는,
   1. EDA(변수 처리, 기초 통계)
   2. 결측치 처리
   3. 파생 변수
   4. 스케일링
   5. 배웠지만 사용해보지 않은 기법들 사용해보기
3. 웹 로그 : 가공 필요.

* **HR** -> 채용, 선발, 배치, 보상, 퇴사
  + 더존비즈온(빅콘테스트) : 퇴사 여부 Classification. 내일 데이터  
    <https://www.bigcontest.or.kr/points/content.php#ct04>
* **e-commerce** -> 이윤 창출 (매출, 마케팅-클릭, 구매전환 / 예측 구매자 분류)
  + Ebay, Amazon(kaggle)  
    e-commerce 플랫폼 매출 증대를 위한 클릭률(CTR)과 구매전환율(CVR) 분석을 통한 마케팅 타깃 분류(캐글 데이터, naver api 등)
* **Kaggle**
  + <https://www.kaggle.com/mirbektoktogaraev/should-this-loan-be-approved-or-denied?select=SBAnational.csv>
  + <https://www.kaggle.com/c/house-prices-advanced-regression-techniques>
* Game -> 이윤 창출 (결제, 이탈)
  + ‘Flood it’

모바일 게임 서비스 유저 이탈률 개선을 위한 로그 데이터 분석(google)  
<https://cloud.google.com/blog/topics/developers-practitioners/churn-prediction-game-developers-using-google-analytics-4-ga4-and-bigquery-ml>

* [링크](https://docs.google.com/document/d/1YYzaPi1aih3PkPwbq5PF6B2WM46MpeUqNKx19qO6Gyw/edit)
  + 데이터셋 사이트  
    <http://insightcampus.co.kr:9090/insightcommunity/?pageid=4&mod=document&uid=12908>

<https://velog.io/@sukong/%EC%A0%95%EB%A6%AC-%EB%8D%B0%EC%9D%B4%ED%84%B0%EC%85%8B-%EC%82%AC%EC%9D%B4%ED%8A%B8>

* 업무 분담
  + AWS : 형림님
  + 문서 작성 :
    - 보고서 작성 : 형준님
    - 회의록 작성 : 보람님
  + 코드 정리 :
    - 코드 정리(AWS, Colab 취합) : 우창님
  + 참고 문헌 내용 정리 : 형림님
* 다음 시간에
  + AWS 사용 방법
  + 주제 선정
  + 내일 오전 (형림님 - 상담)