< 2021-09-25 회의록 😊>

1. 데이콘

- 총 63개의 예측값 필요(21개 품목 \* 3가지 기간)

- private제출 후 모델 변경 불가능

- 28일까지 모델을 완성해야함

-> 데이콘 대회 포기

-> 해당 데이터 이용해서 응통 공모전에 집중하기로 결정

2. EDA 결과 발표 ( 코드 파일 첨부 )

- 데이터를 정규화하고 시각화 해보는 것이 좋을 것 같음

3. 회의 내용

- 의미 있는 피처 탐색

: 시계열 데이터임을 고려하여, 현재 컬럼을 가지고 기존 raw data에서 새로운 컬럼 도출

Ex) 판매지-원산지 간의 거리, 등급에 따른 거래량 count

- 추가 데이터 탐색 ( 1순위 )

- 품목, 품종 간의 관계를 확인해서 그룹화 하면 좋을 것 같음 (이유 : 21개를 다 고려하기 어려움)

- 모델 baseline : lgbm (gbm은 성능이 떨어졌음)

- 예측력이 좋은 모델을 해석하여 의미 있는 insight를 도출하면 좋을 것 같음

- 매주 토요일 20:00를 정기회의 시간으로 결정

4. 과제 (2021-10-02 20:00까지)

- 김권호, 유승희 : 피처 탐색

- 서진슬, 장효정 : 추가 데이터 탐색