

## 기업연계 멘토링 진행

차수명	4차시	일자/시간	2021.11.04/2pm
팀장명	이규정	기업명	신한은행
주제	중간 점검		
세부 내용			
[참여인원] 4명(김윤서, 방지환, 이규정, 정희진)			
[주요 진행 내용] - 웹 페이지 중간 점검 - Backend 진행 상황 공유 - 그 외 질문			
[질문/답변 내용]  • 웹 페이지 피드백 ◦ 여러 데이터 셋 삽입 가능인가요?  → 그럴 필요 없다  ◦ 데이터 삽입 시점은?  → 등록하면 바로 페이지 이동하면서 분석 수행  • Front- Back 협업 방식은?  → Notion 에 작성한 API 문서를 기반으로 설계  • 필터링할때마다 분석 or 아니면 한꺼번에 선택 후 분석  → 분석이 real time 가능하다면 선택 순간 바로 분석   • csv 확장자 문제?			

→ 완제품이 아닌 프로토타입을 만드는 것이기 때문에 굉장히 많은 기능을 많이 포용하는 것 보다는 적은 기능이라도 단단한 페이지가 좋다. (프로젝트 종료 시 제한사항으로 명시해주면 됨 - 추후 확장 방향으로 언급)

→ 대신 csv 와 UTF-8 파일은 완벽하게 될 수 있도록 하기

- 이상치 제거 기준에 대한 피드백

→ 현업에선 z-score 를 많이 안써

→ 실제 모델링에는 완전 이상치가 아닌데 빼야하는 경우가 생김 그럴 때 대비해서 z-score 나 IQR 을 자주 사용하지는 않지만 이거는 데이터 성격을 파악하기 위한 용도로 사용할 예정

- box plot 의 이상치를 그리기 위해 원본 데이터 넘기는거 어떤가요

→ 모든 데이터를 파악하기 위함이 아니기 때문에 굳이 이상치를 다 표기해야할까? q1, q3, iqr 정도만 표기해도 좋다.

- 필터링 순서 문제

→ 이상치 치환은 순서에 따라 달라질 수 있지만 삭제는 순서 고려 x

### [멘토 전달사항]

\*\* outlier 찾아내는 규칙이 중요하고 front-end 가 가장 중요할 수 있다. 겉으로 보이는 완성도가 중요하다@@@

\*\* django 도 backend api 작성된 거 활용해보자 / swagger 사용 추천!

### [건의사항]

