**2018.08.20 회의록**

1. 8/21(내일) 오전 10시 행아웃
2. 종은 : 코드 찾아서 공유하기(Matrix Factorization)
3. 결측치를 제거하고 집계한 것 vs 결측치를 변환하여 집계한 것
4. 내일 모레(8/22, 수)까지 완

→ 결측치를 추정한 것을 그룹화하여 MF(null=sth)

→ 결측치를 제거한 것을 그룹화하여 MF(null=0)

1. 8/31 금융 주제 완료 목표 → 9/1 ~ 9/14 게임 주제 참여

→ 8/22 : MF

→ 8/25 : 군집분석, Mapping, 백분율

→ 8/27 ~ 8/31 : 가산점항목, PPT

참고) 넷플릭스 추천알고리즘 https://car9380.blog.me/221175862829

**\*\* 시행착오 회의록**

1. 군집분석 개수 찾기
2. 임의의 열 값을 선택했을 때, 어떤 군집에 속하는 가(구현 및 확인 방법)
3. 14만개 예측시, 누적 예측? or 한번에 예측? (누적예측)

→ 열 하나당 머신러닝 모델 여러개

→ 14만개의 그룹(정보)와 가장 연관이 깊은 열 모색 : TOT\_SOBI

→ 결정계수가 가장 높은게 TOT\_SOBI(가장 설명력高)

→ 독립변수에 고객정보(범주형) + TOT\_SOBI(연속형) → log\_regression → 모델링

1. 기본정보(8개) 중 필요한 정보는 무엇인지 근거를 설명하세요. (가산점항목)

→ 군집의 특성 파악 필요