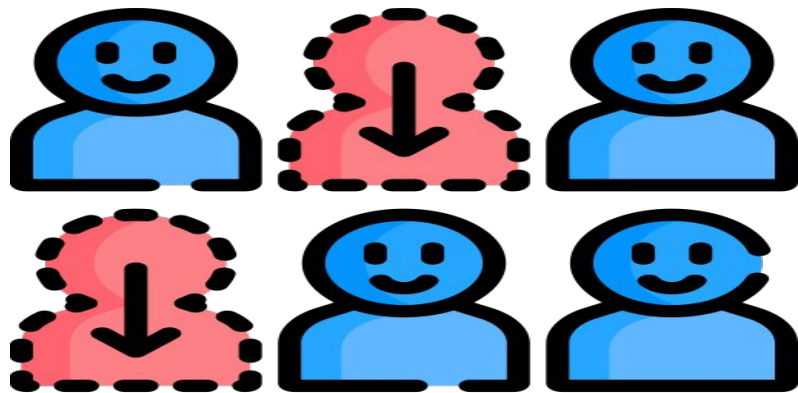


# 3조 ML 프로젝트 중간발표

제로베이스 데이터 스쿨



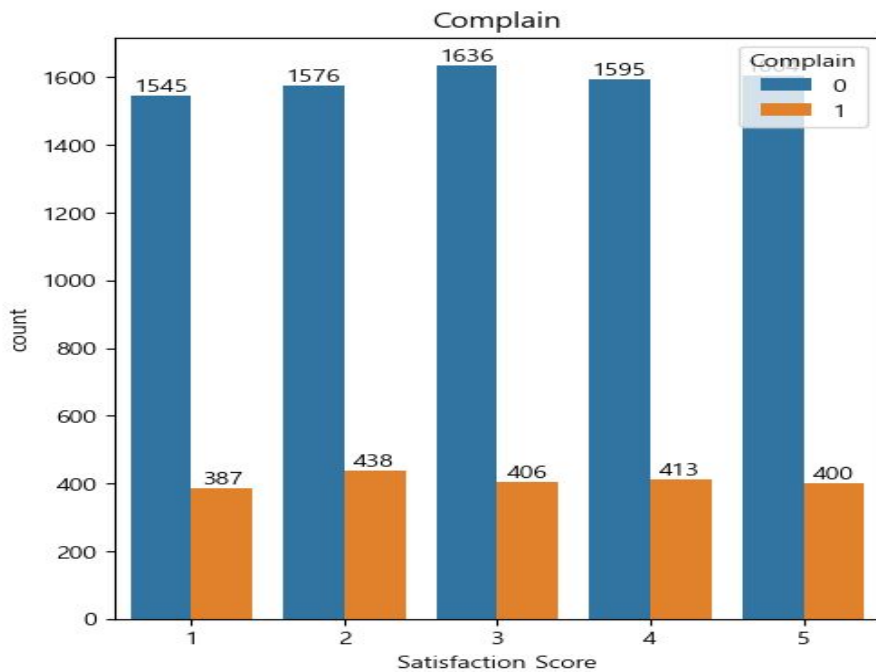
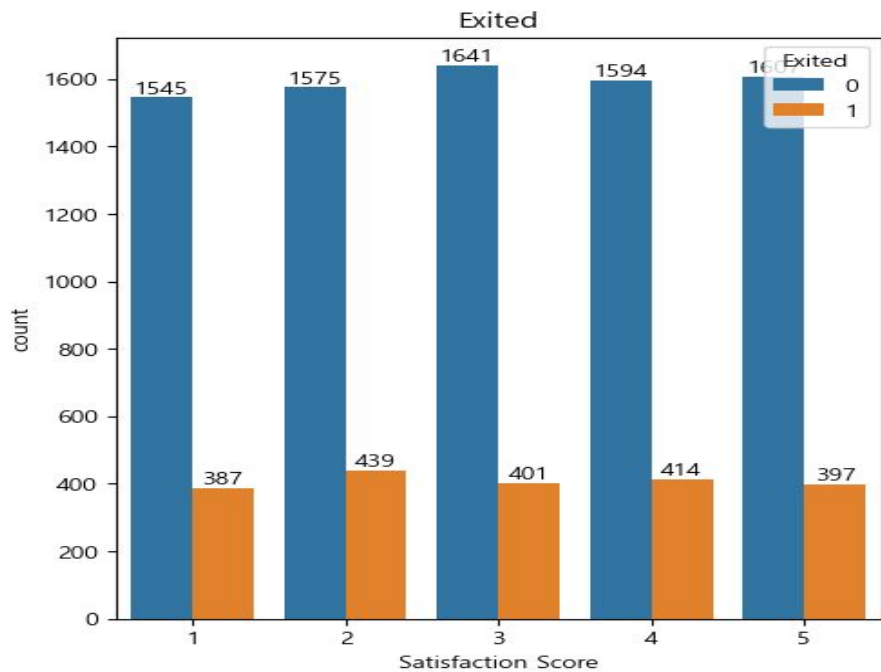
# 목차

1. 프로젝트 주제
2. EDA분석 진행 사항
3. 앞으로의 진행 계획

# 프로젝트 주제

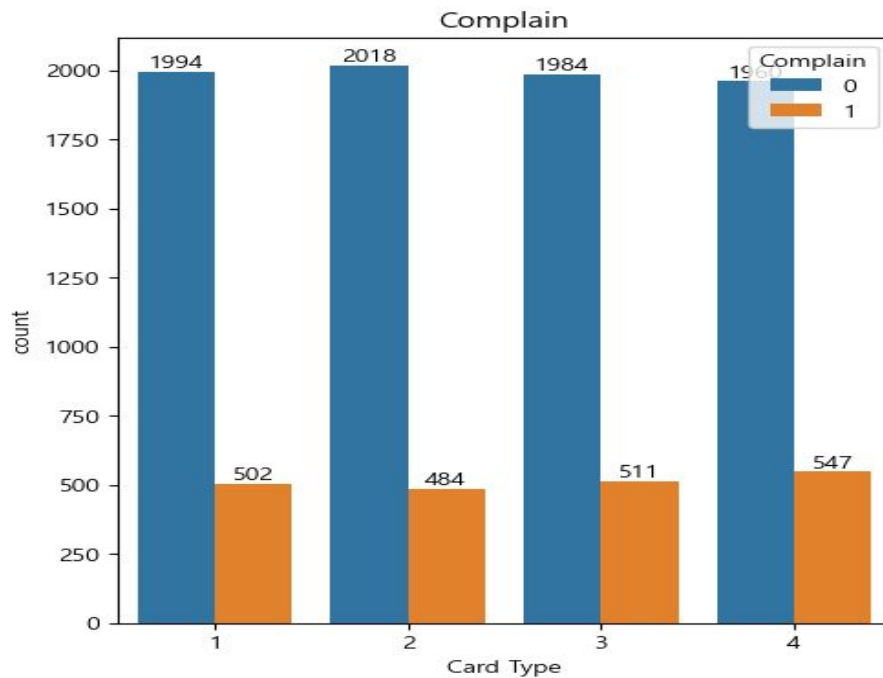
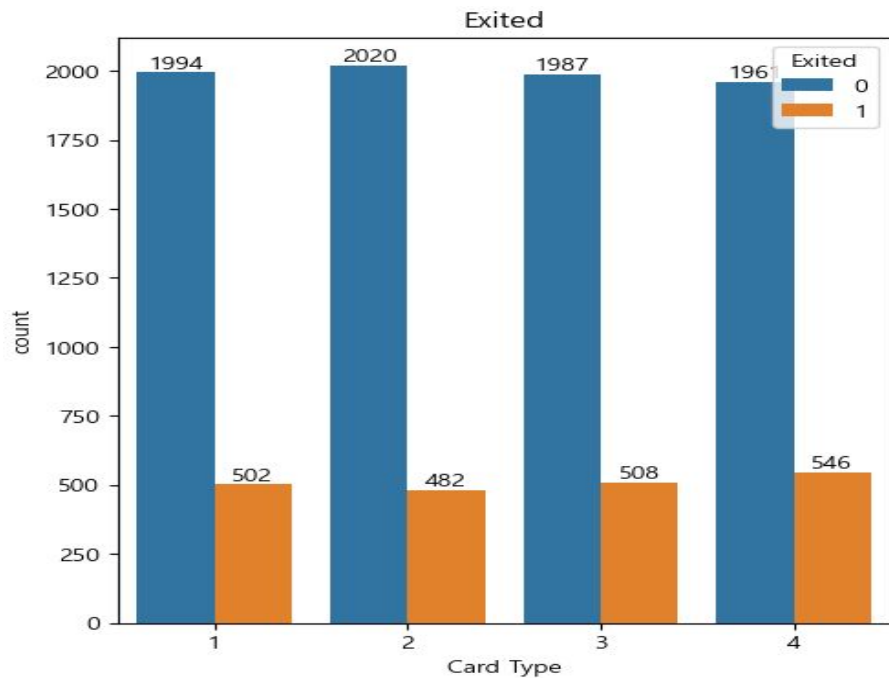
은행 고객 데이터를 이용한  
이탈 원인 및 고객 이탈 예측

# 만족도



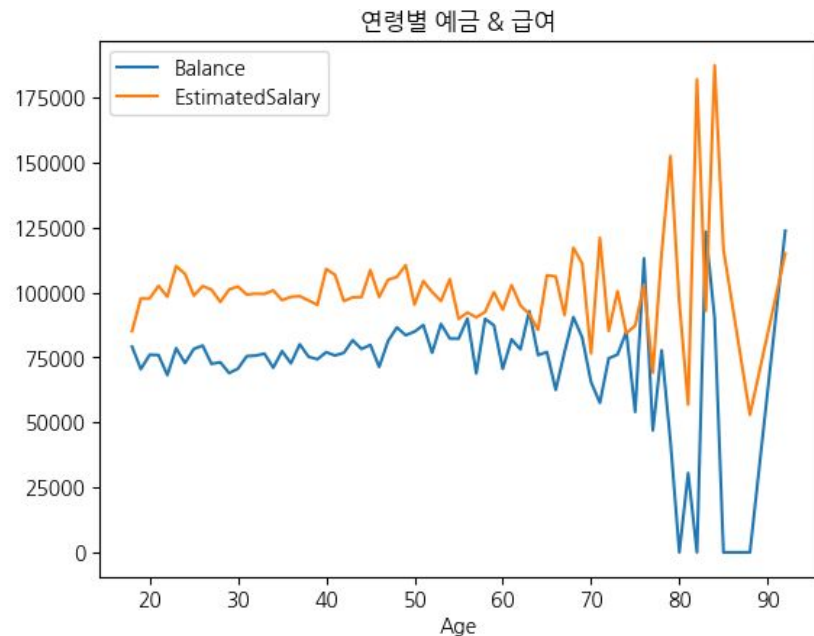
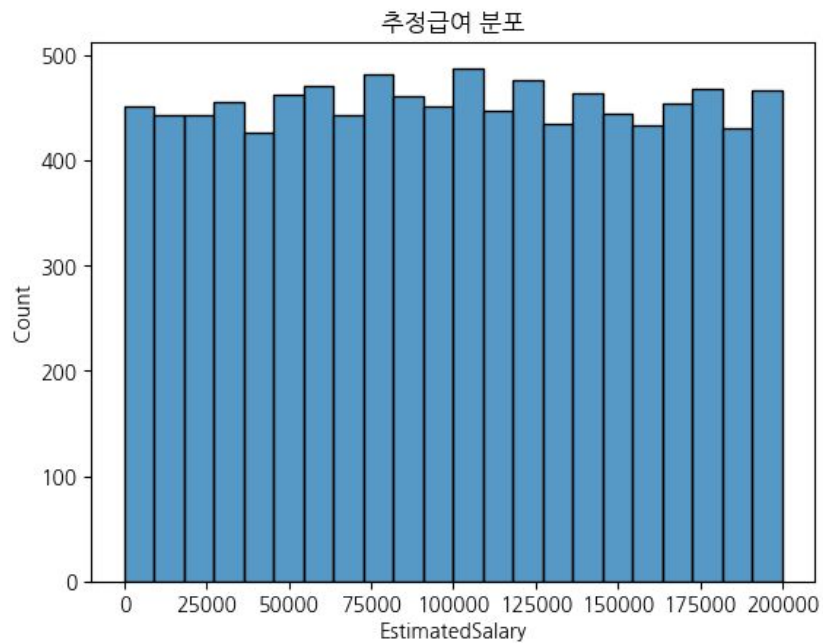
· 컴플레인과 이탈자를 만족도점으로 파악하려했지만 만족도 분포가 고르게 나타남

# 카드유형



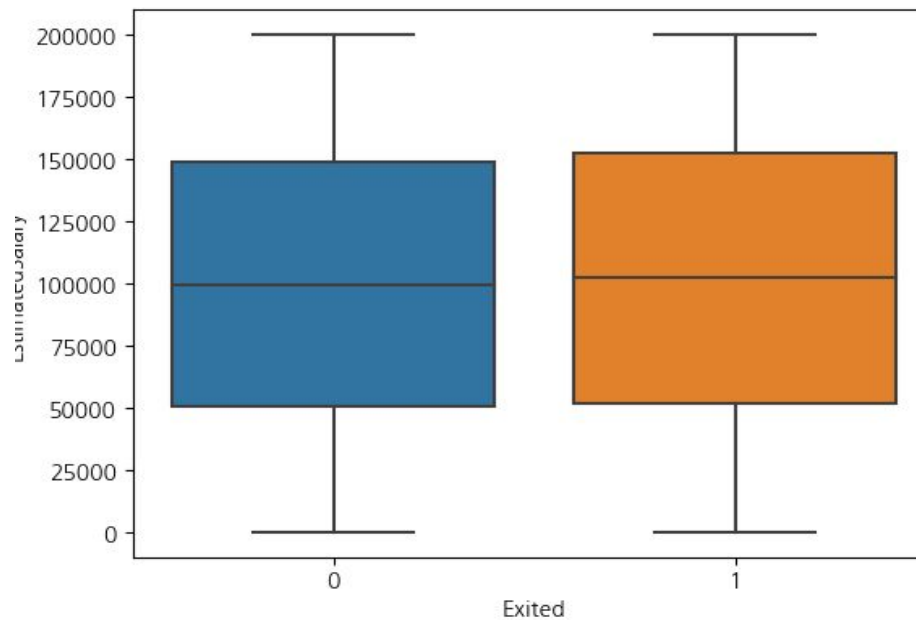
• 1 : Silver / 2 : Gold / 3 : Platinum / 4 : Diamond 또한 고른 분포로 나타남

# 추정급여



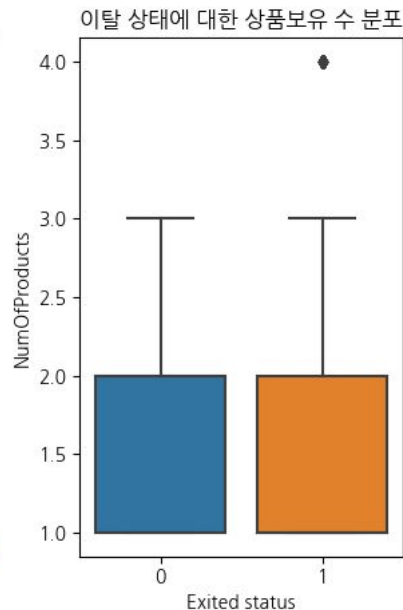
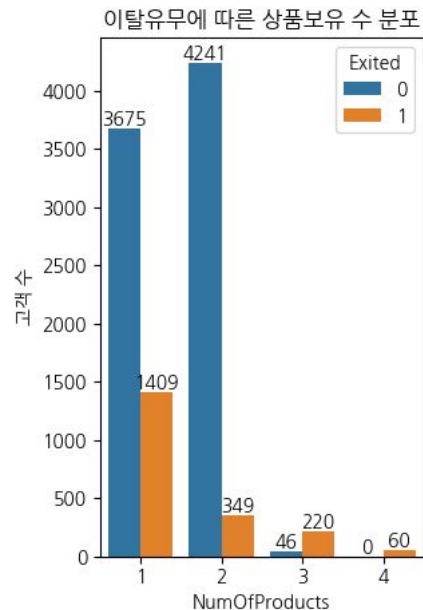
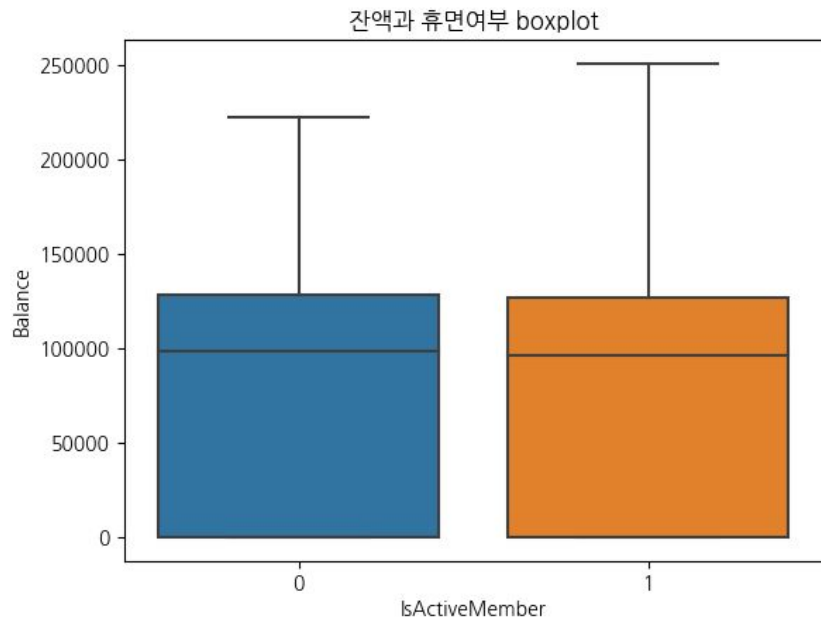
· 70대 후반 높은 추정급여에 비해 잔액이 굉장히 적음

# 추정급여



· 이상치를 확인하자고 했으나 이런 Boxplot이 그려짐

# 잔액과 휴면여부 / 이탈유무와 상품보유 수



· Boxplot 결과가 이렇게 나옴



# 변경된 데이터 소개

- 수집한 데이터 :
  - 카디용 고객에 대한 정보와 이탈 여부를 포함한 데이터
  - 총 23개 변수, 약 1만개의 샘플

- 출처 : 캐글(kaggle)

<https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attribution-with-m>

- Numeric Variables (15):  
Attrition\_Flag, Customer\_Age, Dependent\_count, Months\_on\_book, Total\_Relationship\_Count, Months\_Inactive\_12\_mon, Contacts\_Count\_12\_mon, Credit\_Limit, Total\_Revolving\_Bal, Avg\_Open\_To\_Buy, Total\_Amt\_Chng\_Q4\_Q1, Total\_Trans\_Amt, Total\_Trans\_Ct, Total\_Ct\_Chng\_Q4\_Q1, Avg\_Utilization\_Ratio
- Text Variables (5):  
Gender, Education\_Level, Marital\_Status, Income\_Category, Card\_Category
- Other (1): CLIENTNUM

# 앞으로의 진행 계획

- 변경된 데이터 EDA를 다양하게 분석해보기
- 기존에 존재하는 분석도 분석해보기
- 데이터 전처리 및 모델 학습 준비

# End of Document

