**[2차 아이데이션]**

**1. 팀 이름** : 취업 중급 1조

**2. 팀장 이름** : 최윤석

**3. 프로젝트의 최종 주제를 작성해 주세요.**

- 어떤 강의콘텐츠가 잘 팔리는지 파악후 금액과 판매 시기에 대한 분석을 통해 마케팅 전략 수립

**4. 팀의 데이터 분석 방향과 인사이트 도출 방안에 대해서 설명해주세요.**

- Fastcampus 21년도 판매데이터를 중심으로, 결제 완료가 된 사람들과 환불을 한 데이터가 유의미하다고 생각하기 때문에 state가 completed이며 type이 payment, refund인 데이터들 각각 pay, ref로 지정해 사용했습니다.

어떤 콘텐츠가 구매율이 높은지, 어떤 시기에 구매율이 높은지, 어떤 금액대가 판매율이 높은지, 기존의 마케팅방법이 효과가 있었는지 pandas, numpy 등 다양한 시각화를 통해 분석하였습니다.

**5. 데이터를 어떻게 활용하셨는지 설명해 주세요.**

- 결제 완료가 된 사람들과 환불을 한 데이터가 유의미하다고 생각하기 때문에 state가 completed이며 type이 payment, refund인 데이터들 각각 pay, ref로 지정해 사용했습니다.

- 보안상 데이터 공유가 불가한 컬럼들은 삭제했으며, 불필요한 row는 삭제했습니다. 빈 결측치를 처리하기 위해 replace, fillna 등을 활용했습니다.

-결제가 이루어진 시간( completed\_at) 컬럼을 월별, 일별, 시간별로 나누어 결제량을 파악했습니다. 또한 환불 시간대나, 일별 환불량, 시간별 환불량 등 환불에 관련된 데이터를 분석했습니다.

- 강의판매금액, 쿠폰할인금액, 최종결제금액을 임의로 구간을 나누어 어떤 구간에 많이 분포되어있는지 파악했습니다. 또한 기존에 발행한 쿠폰 이외에 할인이 들어간 데이터나, 할인액에 따른 월별 판매량도 분석했습니다.

- 강의별로 할인을 받아 결제한 사람들을 분석해보고, 강의결제시 쿠폰 사용비율이 얼마나 되는지, 여러개의 강의를 들은 사람이 있는지, 어떤 강의가 수강생이 많은지, 환불율이 높은지 분석했습니다.

**6. 앞으로 프로젝트를 마무리하면서 추가적으로 해야 하거나 아쉬운 점이 있다면 작성해 주세요**.

**김주호 :** EDA를 하면서 시각화할 때 어떤 데이터에 어떤 그래프를 활용할지 판단하기 어려웠다.

**송희연 :** 데이터에서 보안상 공개할 수 없는 부분이 type카테고리의 order 등인데 이 부분을 더 파고들어서 분석 해보고 싶었는데 할 수 없어서 아쉬웠다. 구현하고싶은 아이디어가 있어도 나의 프로그래밍 실력이 부족하여 표현을 못하는데 라이브러리와 함수를 더 공부해서 EDA를 더 다양하게 하고 싶다.

**안희수 :** 데이터 자체가 한정적이고, 가격과 시간으로만 분석해야해서 드라마틱한 시각화 연습을 못해본게 아쉽다.

**조세이 :** 여러 데이터를 다양한 시각화 툴을 사용해 분석해보고 싶다.

**최윤석 :** 세분화해서 분석해보고 싶은 데이터가 있었음에도, 불구하고 세분화된 데이터 정보가 없었다는 점이 아쉬웠고, Python 공부를 좀 더 열심히 해야할 필요성을 느꼈다.

**허재호 :** 데이터 분석을 하기에는 공유된 정보가 한정적이었고, 시각화를 다양하게 하기에도 제한적인 데이터였다는게 아쉬웠던거 같다. EDA는 처음이라 아주 깊이있게 분석을 할수가 없다는것을 느꼈고, 더욱더 많은 데이터를 보고 분석을 해야겠다고 생각을 했다.