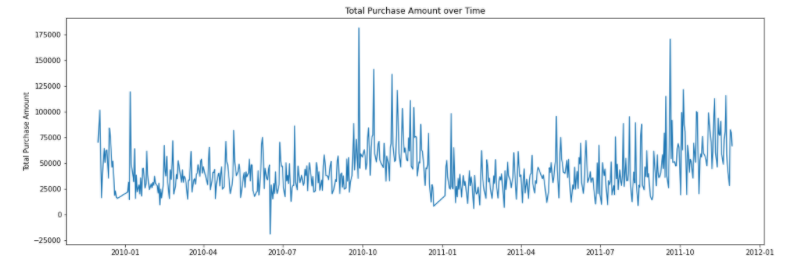
<해커톤 주제> 04.23일 오후 5시

최근 온라인 거래를 이용하는 고객들이 많이 늘어나고 있어 고객들의 log 데이터가 많이 늘어나고 있습니다. 온라인 거래 고객 log 데이터를 이용하여 **고객들의 미래 소비를 예측** 분석프로젝트를 진행하려 합니다.

고객들의 월별 총 구매 금액을 확인했을 때 연말에 소비가 많이 이루어지고 있는 것으로 확인이 되었습니다. 그리하여 **12월을 대상으로 고객들에게 프로모션**을 통해 성공적인 마케팅을 하기 위해 모델을 만들려고 합니다.

온라인 거래 log 데이터는 2009년 12월부터 2011년 11월까지의 온라인 상점의 거래 데이터가 주어집니다. 2011년 11월 까지 데이터를 이용하여 2011년 12월의 고객 구매액 300초과 여부를 예측해야 합니다.

평가 방법 : AUC Score

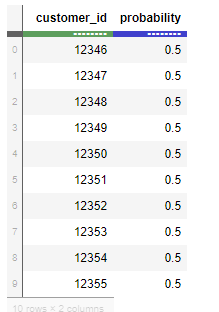
[11:55](https://sd-si52796.slack.com/archives/C01TF9M21FW/p1618887359001800)

<데이터 개요>

고객들의 구입이 이루어질 때 1개의 데이터가 생성이 되며 데이터에는 고객에 대한 정보(고객 id / 고객 거주 국가), 상품 정보(상품 id / 상품 설명 / 상품 가격)와 거래 정보(주문 일자 / 상품 수량)로 2009년 12월 부터 2011년 11월 데이터 까지 학습 데이터셋으로 주어집니다.

데이터 컬럼 설명

* order\_id : 주문 번호. 데이터에서 같은 주문번호는 동일 주문을 나타냄
* product\_id : 상품 번호
* description : 상품 설명
* quantity : 상품 주문 수량
* order\_date : 주문 일자
* price : 상품 가격
* customer\_id : 고객 번호
* country : 고객 거주 국가
* total : 총 구매액(quantity X price)



평가 매트릭 : ***from*** sklearn.metrics ***import*** roc\_auc\_score

score = roc\_auc\_score(labels,preds)

Submission sample에 결과를 입혀서 제출하시오