자율성을 더 드리지 못해서 유감입니다. 시간도 많이 없고 확고한 주제를 선정하지 않으면 세부적인 부분에 집중을 하여 통합을 하지 못할 것 같은 두려움에 이 문서를 만들게 되었습니다.

읽어 보시고 회의에 참여해주세요 그리고 이해가 안되는 부분이나 추가했으면 하는 부분 꼭꼭 말해주세요 피드백 주세요요!!! (ps.새벽에 만든거라 맞춤법과 글이 조금 어눌할 수 있습니다.)

**주제: 패스트 캠퍼스는 왜 우수고객을 만들지 못했는가**

선정이유: 우수고객을 위해서 고객군 분류를 진행해보았지만, 일반 고객에 비해서 너무 적은 수치를 나타냅니다. 그래서, 마케팅과 연결 짓기가 어렵다고 생각했습니다.

그리하여 생각해낸 것이 “이러한 마케팅으로 우수고객을 만들었다가 아니라 만들지 못한 이유를 가설로 설정하여 그것에 타당한 근거를 찾아서 이야기를 하는게 낫다고 생각이 들었습니다.(데이터를 활용할 수 있을 것 같더라고요)”

[간략한 데이터 전처리 설명]

자세한 것은 [Google Colab](https://colab.research.google.com/drive/1nZg6Sa7dYqJkPHgtzUIG9yzk7KUikjcL?usp=sharing) 참고해주세요

여러 고객 분류을 위한 데이터의 컬럼은 customer\_id, type=’PAYMENT’, state=’COMPLETED’, transaction\_amount만 사용합니다.(이 외 나머지 type와 state는 삭제했습니다.)

데이터 제거 이유: 그 데이터의 의미를 파악할 수 없기에 우수고객 선별에 지장을 주는 이상치 데이터이므로 제거를 감행했습니다. 저희는 주제에 맞게 해석할 수 있는 부분만 뽑아서 데이터 분석을 실시해야 합니다.

제거된 데이터: type 속 payment 와 refund, state속 completed 와 requested의 나머지

고정 컬럼: 고객 분류는 구매가 결정이 된 이후부터 분류가 가능하기에 type와 state는 고정입니다.

**<이 아래는 월요일 회의 전까지 진행이 안료 되었으면 합니다. 그래야 화요일에 모여서 정리하고 수,목 발표자료 준비할 수 있습니다. 맨 아래 대강 타임라인 잡을게요>**

Coustemer\_id = 1 :1개만 구매한 고객 = 일반 고객이라고 제가 편의상 정의한 것입니다.

[고객군분석]

**저희의 목적은 고객군 분류이지 고객의 분류가 아닙니다.!!**

**저의 생각: 75%정도가 1개만 구매했습니다. 그래서, 먼저 고객군분류(vip~일반 고객)을 나누고 나서 왜? 일반 고객이 이렇게 많은가? 어떤 이유가 있을까? 마케팅 문제인가? 패캠 강의 정책 문제인가? 기타 등등 인사이트를 잡을 수 있을 것 같아서 1개 구매한 고객이 중점으로 합니다.**

**데이터:**

* 기업이 포함(transaction\_level\_B2B.csv)
* 기업 미포함인 일반 데이터(transaction\_level\_normal.csv)

기준: 중복된 구매 수(customer\_id), 총 구매 금액(transaction\_amount)

방식: 구매 수와 총 구매 내역을 4가지로 구분하여서 등급별 점수 부여

컬럼 사용: course\_id, type=’payment’ , state = ‘completed’, transaction\_amount,

역할 분담

두 분 다 데이터는 기업이 미포함이 된 것으로 진행해주세요 이유는 압도적으로 기업보다 일반 고객 데이터가 많기 때문입니다.

- 수현님: 중복된 customer\_id를 바탕으로 구매 수 추출, 구매 수 기반으로 4구간 나누고 점수 부여(기존에 하셨던 거 진행), 구간별 인원에 대한 시각화 진행, 나누는 기준 공유 부탁드립니다.

- 창헌님: 총 구매 금액(sum of transaction\_amount )을 4구간으로 나눈 후 점수 부여 , 구간별 인원에 대한 시각화 진행, 나누는 기준 공유 부탁드립니다.

- 수현님 & 창헌님: Customer\_id를 나눈 구간의 기준을 바탕으로 총 구매 금액과 비교해주세요

각자의 기준에서 나온 점수를 총합합니다.

VIP~일반 고객 분류(비율에 대한 시각화)

[마케팅분석]

냉정하게 1개만 산 고객이 75%정도가 됩니다. 그래서, 마케팅 효과로 인해서 우수 고객을 만들었다는 것은 어불성설입니다. 허나 이 과정을 거치는 이유는 우수고객을 만들지 못했다라는 결론을 도달하기 위함입니다. 왜냐하면, 주제가 **패캠이 왜 우수고객을 만들지 못했나** 이니까요

**데이터:**

* 기업이 포함(transaction\_level\_B2B.csv)
* 기업 미포함인 일반 데이터(transaction\_level\_normal.csv)

기준: 쿠폰 이름(coupon\_title), 구매 확정 시간(completed\_at)

컬럼 사용: type=’payment’ , state = ‘complted’, 고객분류된 데이터, coupon\_title, completed\_at

[역할 분담]

두 분 다 데이터는 기업이 미포함이 된 것으로 진행해주세요 이유는 압도적으로 기업보다 일반 고객 데이터가 많기 때문입니다

-채영님: 수현님과 창헌님이 분류하신 고객군을 기반으로 쿠폰 마케팅[%,금액,0원 할인] (2021.03이전 데이터)이 효과가 있었는지 확인합니다. (만약 분류가 늦어진다면, type & state를 고정으로 잡고해도 괜찮을 것 같아요), 관련 시각화도 진행해주세요 무슨 쿠폰이 많은지 파악하실 때 창헌님의 워드 클라우드를 사용하면 조금 더 좋을 것 같긴해요.!!

-석민님: 수현님과 창헌님이 분류하신 고객군을 기반으로 주간, 월별로 패스트 캠퍼스의 마케팅(2021.03~2021.09)이 효과가 있었는지 파악해주세요. (만약 분류가 늦어진다면, type & state를 고정으로 잡고해도 괜찮을 것 같아요), 관련 시각화도 진행해주세요

-채영님&석민님: 효과에 대한 결론을 구간별로 나눠서 설명해주세요! 있다/없다

**[데이터의 아쉬운 점과 함께 우수고객을 만들지 못한 이유를 찾을 것입니다.]**

현재 상황: 데이터가 너무 쏠려있습니다. 왜냐하면, 패캠은 마케팅 정책을 실패해서 실제 협력과 후원 기업사에게 좋지 않은 평가를 받은 상태이고 회사 재정도 나빠지고 있습니다. 더 나아가, 일반 고객마저 패캠 강의 패키지 정책을 선호하지 않고 있습니다.

무슨 정책을 펼치고 있기에 이런 사태가 날까요? 가설 설정 혹은 추론의 단계입니다.

이건 기존 데이터로는 구분이 불가능하기에 여론 즉 VOC데이터를 기반으로 우리의 생각에 대한 근거를 잡을 것입니다.

**[VOC데이터 수집(크롤링)] – 재근님, 세현님**

이때 포인트는 수집량보단 수집 시간을 중점으로 이야기할 것입니다. 그래야, 조금 길게 한 것처럼 보이니까요

수집량: 각 번호당 10개씩 진행합니다.(생각보다 많이 없을 수 있고, 애매하면 저희가 추가적으로 글 작성해서 넣어버리죠 ㅎㅎ)

데이터 수집 장소: 강의 회사 비교가 담긴 글을 중점으로 해야할 것 같아요. 왜냐하면 저희가 원하는 것이 , 특정 강의에 내용보다는 패캠 정책에 대한 비판인 것 같아서요

+)추가적으로 저희의 가설과 유사한 내용을 담은 블로그가 있다면 더 좋겠죠

[우수고객을 만들지 못한 이유(가설) 혹은 재구매율 증가 못 시키는 이유](글을 좀 이쁘게 다듬어야할 것 같아요)

1. 여러 강의가 들어있는 패키지 형식보단 특정 강의를 원하는 사람은 필요없다.
2. 하나만 듣고 싶은 사람에겐 모든 강의를 사야하는 비용적 부담
3. 여러 개의 강의가 패키지에 있기에 한 번만 구매해서 자신의 니즈를 충족할 수 있게 되어서 재구매 할 이유가 사라짐(여러 개의 강의를 나눠서 못 팔기에 여러 번 구매한 고객이 현저히 적은 이유)
4. VIP를 만들기 위해서 쿠폰과 d%할인은 전혀 경쟁력이 없다.(다른 인강 회사들이 하는 정책)

**[가설 정리] – 모두**

저희가 다른 회사의 정책과 비교해서 이렇게 하면 늘 것 같다는 기대효과를 몇 개 제시하면서 마무리를 하면 좋을 것 같아요

그러면, 데이터 분석으로 일반 고객이 많은 것을 판단했고 크롤링을 이용하여 VOC 데이터 기반으로 패캠이 우수 고객을 만들지 못한 이유에 대한 근거가 갖춰졌고 이를 기반으로 다른 회사의 정책을 벤치마킹을 한 후의 기대효과를 말하면 분석은 끝난 것 같아요.

즉, 원인- 결과 그리고 해결책과 기대효과를 제시했으니 FLOW는 괜찮은 것 같네요

|  |  |
| --- | --- |
| 2022.05.12 ~ 2022.05.16 오전 회의까지 | 각자의 해당파트를 마무리해서 발표해주세요 |
| 2022.05.12~2022.05.13 | 1차 아이덴티 제출 |
| 2022.05.16 저녁 잠깐..? | 해당 파트의 연결점 잡기 및 가설 체크 |
| 2022.05.17~2022.05.18 오전 회의까지 | 추가 크롤링 작업(인원 투입예정) / 벤치마킹 회사 찾아서 가설 체크 22  혹은 발표자료 정리 |
| 2022.05.18 ~ 2022.05.20 오전 회의까지 | 발표자료 정리 및 제출!! |

[핵심 요약]

고객군 분류 : customer\_id와 transaction\_amount, vip~ 일반 고객 나누기(일반고객과 그 외 고객의 편차가 너무 심하다라는 문제의식 제공)

마케팅 분류; vip만들기보단 왜 만들지 못했는지 초점을 맞추는 인사이트 제공(일반 고객 중심)

크롤링: vip를 만들지 못한 이유를 크롤링을 통해서 voc데이터 수집 그리고 이것을 기반으로 우수고객을 만들지 못한 근거 확립(마케팅의 잘못)

[당부]

3~5일 정도는 팀끼리 나눠서 소통을 하고 정기 회의 때는 그 인사이트를 모으는 행위로 진행할 것 같습니다. 다들 피드백 많이 많이 주세요 이번주에 틀과 방향성을 잡아야 다음주에 고생 안합니다아!!