**2조 1차 프로젝트 회의록**

작성자 : 이재현(팀원)

**프로젝트 기간 (총 2주)**

**11월 04일 (월)**

**11월 05일 (화)**

**11월 06일 (수)**

**11월 07일 (목)**

**11월 08일 (금)**

**11월 11일 (월)**

**11월 12일 (화)**

**11월 13일 (수)**

**11월 14일 (목)**

**11월 15일 (금)**

**11월 18일 (월) 발표일**

**11월 04일 (월)**

매일 아침 1교시 그날 목표 공유하고 진행

[일정 및 제출 일자 정리]

11월 5일 기획서 마무리 후 제출

11월 8일 1차 주간보고서 제출

11월 15일 2차 주간보고서 제출

11월 18일 발표 (13시 예정), 발표자료 & Github 소스코드 정리

오늘 논의 사항 3가지

1. 데이터 시각화 어떻게 할 것인지?
2. 프로젝트 진행 상황 공유
3. 강사님과 저녁 회식 메뉴 선정

소상공인과의 연결점을 잡으려고 방향 처음 잡음

2030 신규 고객 유입이 목적으로 방향이 크게 바뀜

시각화를 진행해보았는데, 은행권 자체가 2030 고객 잡는 데 혈안이 되어 있기도 하고, iM뱅크 고객을 인구별로 시각화 해보면 2030 비율이 높지 않고 고객도 대구 경북에 집중되어 있음

21, 22, 23년에는 시중은행으로의 전환 이전이기도 했고 데이터의 75%가 대구 경북 중심

지정학 및 자영업자 키워드를 제외하고 신규유입에 초점으로 방향 전환

전국 단위로 분석 진행해봤자, 전국 단위의 패턴은 발견하기가 어려움 있을 것으로 예상

통계적 분석

“지역 상생”이라는 키워드가 대구 경북에 메인이 맞춰지는 느낌

다른 지역의 소비자도 이런 식의 분포를 가지고 있기 때문에 우리가 비슷하게 다른 지역 고객에게도 적용할 수 있을 것으로 판단됨

지역 상생 카드

지자체의 협력과 함께 카드 출시 가능 (인천 페이 등)

카드 혜택 등을 조사하여서 전국 지자체와 협력이 가능한 카드 조사 필요

상생 추구를 넣은 이유는 임시로 넣은 이유는 자영업자를 빼면 아예 주제가 바뀔 우려 있기 때문

2030을 유입하면서 지역 활성화하는 것이 목표

지역 상생 [지역 활성화, 지자체와의 협력 등이 포함된 개념]

일단 customer\_data 말고 card\_data 위주로 사용하는 방향으로 가야 할 듯

고객의 전체적인 동선 파악, 특성 파악을 할 때에는 merged data를 사용해야겠지만,

그래서 카드의 사용 패턴을 분석할 때에는 card csv만 가지고 해야 할 것 같음.

ppt는 이번주 수요일부터 천천히 제작 예정

프로젝트이다 보니까 미래캔버스보다는 formal하고 simple하고 깔끔한 템플릿 찾을 예정

분업 방안

1. 코드팀 2명 (전처리 및 시각화, 통계적 검정 → box plot 등으로 시각화)
2. 조사 및 기획팀 2명 (기획 및 자료조사팀, 우리에게 필요한 배경 데이터 다른 카드사는 이런식으로 하고 있더라 등) 우리의 카드 상품은 어떤 혜택을 주어야 할까?

코드팀은 앞에서부터, 조사팀은 뒤에서부터 시작해서 중간에 중간에서 만나는 것이 목표임

최현진 송현서 : 코드팀

박현이 이재현 : 조사 및 기획팀

완전 은행 기반으로만 리서치를 하였는데,

비은행권에서 나오는 신용카드 중에서도 좋은 것이 많을 것으로 에상

카드사에서 나오는 혜택과 은행에서 나오는 카드의 경우 약간 방향성이 다르다면 이러한 경향성도 고려하는 방향으로 진행해야 할 듯

다른 은행에서도 분명히 mz/청년/2030 대상으로 무엇인가 진행할 것임

층화 추출 데이터 : 원본 데이터를 그룹화 해서 연령대, 성별, 고객등급, 자택 시군구 붙임

경쟁사 카드 분석

<https://www.shinhancardblog.com/1313>

**11월 05일 (화)**

기획안 작성

기획안을 읽는 사람이 “2조에서 하고 싶은 이야기가 딱 한 가지이구나”라는 느낌이 들도록 작성

식생활 소비 패턴을 파악하고 나서, 2030 vs 4050 식생활 중분류끼리 서로 비교분석 진행 및 식생활 & 신규 고객 유입 방안 아이디어 토의

2030을 잡을지 vs 식생활 소비를 잡을지 큰 방향성 잡을 필요성 있음

분석 진행 시 y축은 이미 결정되어 있고, 성별 및 연령대 등 다양한 각도로 시각화 진행해서 결과 분석 진행할 필요 있음

타켓을 점점 좁혀서 20대

나이대로 나누어서 시각화를 진행한 후에, 가장 특징이 두드러진 집단을 언급

DIY 카드 아이디어 생각해보기

연령대, 성별, 승인금액 등 컬럼에서 승인금액은 합산금액으로 되어 있고 승인 건수는 범주화되어 있어서 우선 max 값으로 가정

혹시 시계열을 위해 분기 별로 group by까지 이중으로 걸면, 특정 그룹이 특정 분기에 평균 얼마 썼는지에 관한 정보까지 알 수 있음

나이, 분기, 업종명 까지 그룹 바이 하면 특정 그룹이 특정 분기에 특정 업종에서 평균 얼마 썼는지 알 수 있음

사용할 변수 (최종) : 성별, 연령대, 승인금액

y축은 식생활

x축은 성별, 연령대, 승인금액

지방 사시는 분들, 서울에 있는 핫플레이스에 와서 사가고 먹고 하는거 있음

대전 성심당 가서 빵 퀵턴 느낌으로 짧은 여행도 증가 추세에 있음

지역을 사용하고 싶으면, 특정 지역에 가서 이 카드로 결제를 하는 것에 대한 혜택도 고려할 수는 있음

강사님 조언

통계책 과거에 재밌었던 사례 중 전체적인 데이터는 우하향 하고 있으나, 몇 개의 그룹 데이터는 일부 몇 개의 그룹화된 데이터는 우상향 중이었음

전체적으로는 우하향 하는게 맞은데 그 중에서 일부 데이터는 우상향 하는 것을 찾는 것이 유의미함

대구에 있는 주요 기업을 조사해서, 주식이 상장되어 있는 회사들 중에서, 해당 회사들의 3년치 데이터 주가를 외부 데이터로 활용하는 것도 가능

일단 대구 경북으로 아예 한정하는 것은 어떨지 고민해볼 필요 있음

서울의 소비패턴까지 보겠다고 하면 갑자기 데이터가 섞여 버려서 노이즈가 생김

"가맹점업종명" 205개를 5개의 중분류로 분류

* 식생활 (Food-related)
* 소매 및 생활용품 (Retail & Consumer Goods)
* 서비스 및 여가 (Service & Leisure)
* 교통 및 물류 (Transportation & Logistics)
* 기타 (Miscellaneous)
* 식생활 (Food-related)

**11월 06일 (수)**

외부 데이터로써 소비자 물가 지수는 넣었으면 좋겠음

인플레이션이 구매력에 영향을 많이 끼치기 때문

구매력이 높은 시기임에도 불구하고 사람들이 특정 영역에서는 사람이 돈을 많이 사용하지 않았다 이런 현상이 보일 수 있음

소비자 심리 지수 : 카드 사용, 선택 등도 전부 심리와 관련이 되어 있기도 하고, 소비자동향지수라는 것도 있는데, 이것은 6개 정도의 변수를 계산해서 나오기 때문에 의미 있을 것으로 에상

소비자동향지수는 심리, 물가 등 다양한 변수를 포함하고 있기 때문에 이미 이거 하나로 많은 것을 포함할 수 있어서 좋을 듯

외부 데이터 4가지

소비자 관련

소비자동향지수

거시 경제 관련

금리

부동산 등 관련 아이디어

한국형 리츠 제대로 된 곳 아직 없음

요즘 부동산 조각 투자 관련해서 어플 나오고 있음, 옛날에 음반 투자 나왔던 것과 비슷한 현상임

한국에는 청약, 전세 같은 한국형 특이한 거주 형태가 있어서 한국형 리츠 얘기 나오는 이유도 이것을 통해서 전세 문제를 해결할 수 있지 않을까? 고민 필요

파생변수를 이용한 시각화! 시각화가 조금 우리가 예상한 반응과 다르거나 그러면 그 때 통계 검정을 진행하면 됨

**11월 07일 (목)**

목요일 예정 진행 내용 : EDA + 통계 분석

이전까지와 다르게 다 같이 이야기하면서 진행하는 것이 좋다는 의견 有

각자 데이터에서 보고자 하는 것 생각하는 것 다를 수 있다는 의견 有

오전에 EDA 각자 진행하고 오후에 EDA 진행 결과 공유

이후에 오전에 진행한 EDA를 기반으로 통계 분석 진행

EDA 진행 시 주요 변수

연령대, 업종별, 성별, 거주지

이후 개인 별 EDA 진행

**11월 08일 (금)**

1차 주간 보고서 완성해서 제출

클러스터링 + 다른 통계 기법 고민

소비 패턴이 성별 별로도 지역별로도 연령 별로도 다르다는 결론 도출됨

카드 혜택의 경우 타깃층을 먼저 정하고 이후에 고려 진행

DIY 카드 만드는데, 혜택은 여러 가지 중에서 고객이 직접 자신이 원하는 것 고르는 방향으로 진행

EX. 통신 특화 카드 - 통신 관련 혜택

대구 사람인데, 수도권으로 가서 살 사람들 (취업 등) 출장 때문이든 가족 때문이든 다양한 이유로 서울에 자주 경우 다양한 사례가 있을 것으로 EDA 결과 추정됨

중요 변수 4가지 [나이, 성별, 업종, 지역]

강사님 조언 : DIY 카드를 제안하더라도 가이드는 어느 정도 정해져 있을 필요

EX. 공통 혜택 + 선택 혜택

선택형으로 조건을 너무 복잡하게 선택하게 하면, 선택형을 막상 발급 받더라도 선택형 혜택을 잘 사용하지 않게 됨

어차피 일반적인 대부분의 신용카드를 분석해보면, 기본적인 혜택의 한도가 한 달에 3만원 정도가 최대임

그렇기 때문에 꼼꼼하게 세부 선택 혜택을 정할 할 바에 단순하지만 선명한 카드를 선택하는 경향 有

우선 나이 성별 + 지역 기반 총 12가지 그룹을 그룹화 하여 소비패턴 분석 진행

지역별 클러스터링 5가지 지역, 이렇게 종합하면 총 60가지 그룹 나올 것으로 예정

6 x 2 = 12가지, 12 x 5 = 60

**시각화할 때 기본 방향**

x축 : 2021~2023년 (시계열)

y축 : 승인금액 추이, 승인건수, 1인당 평균 승인금액, 가맹점업종\_범주화 비율

**11월 11일 (월)**

다른 팀과 회의 진행 결과 다른 팀의 경우 승인금액을 위주로 소비 패턴 진행 중

승인금액의 높음 유무를 통계 분석을 통해 증명하여, 연령대 별로 차이가 있는지 없는지 가설 세우고 사후 검정 진행 예정

오늘 우리팀 진행 목표 : 통계 분석

OBJECT 형태로 진행할 것이라면 카이 제곱 밖에 없고 INT형을 사용하려면 승인금액만 사용해야 한다는 한계 존재함

그래서 분석은 크로스탭을 사용해서 비율을 나타내낼 것

PCA는 연속형 변수에 특화되어 있고, TSNE 티에센이는 결과가 특별하게 안 나옴

대형 모 신용카드 본사가 중구에 위치해 있어서 “기타” 및 “인터넷 P/G” 관련 비율에서 중구가 혼자 너무 이상치로 튀어버림

중구를 제외하고 나면 서울특별시와 경기도 전체를 5개 정도로 클러스터링 결과 도출 가능

금리, 인구 등도 외부 변수와의 고려를 위하여 금리랑 cpi를 그냥 그래프로 최근 동향 설명 느낌으로 초반에 넣고 거시적인 시각화 진행

통계 분석 및 인사이트 도출에 있어서는 인터넷 PG 혜택을 포함해서 진행할 것

단, 시각화 할 때에는 인터넷 PG의 비중이 높기 때문에 이것은 시각화에서는 제외하도록 진행 예정

인터넷 P/G를 제외한 이유 : 다른 업종에 비해서 너무 압도적으로 승인 건수가 많아서 시각화 그래프가 눈에 보기 좋게 나오지 않고, 추가적으로 인터넷 P/G의 경우 정확하게 소비자가 어떤 업종에 돈을 사용하였는지 알기 어렵다는 이유 등

성별 비율 관련 에러 사항

남녀 인원수 차이가 심하게 났었음

고객 ID를 기준으로 해서 여자가 결제 건수가 더 많음

수도권 지역에서 고객 ID가 여러 번 결제한 사람은 여러 번 찍히니까... 그래서 진짜 머릿수를 보려면 DISTINCT로 봐야 함

**11월 12일 (화)**

목요일에 강사님과 팀별로 짧게 미팅 진행

PPT 완벽하게는 끝내지 못 하더라도 목차 부분이라도 간단하게 나와있어야 함

대시보드 시연을 해야 하는데, 배포는 저번에 데이터를 클라우드에 올리면 안 된다고 확답을 받았기 때문에 STREAMLIT 배포 등은 일단 안 하는 방향으로 잡음

마케팅 방안 내용 토의

현직자 입장에서 어떤 기법을 사용했는지 보다도 이런 것을 기반으로 어떤 인사이트 얻어서 이런 카드 제안하고 싶은지가 더욱 중요할 것으로 보임

데이터 관련해서 지금까지 고민을 많이 했으니까 이제 마케팅 적으로 어떻게 할 지 고민하는 것이 맞음

DIY카드도 사실은 카드 제안 보다는 혜택 제안을 하는 방향으로

소비 금액이랑 금리 기반으로 회귀분석 어제 진행했었는데, 유의미하지 않았음

모델의 설명력 부족

아쉬운 점 : 차원 축소를 사용해보았는데, 데이터가 충분히 크지 않아서 데이터의 차원이 낮아서 행은 많은데 열이 많이 않아서 제 성능을 잘 발휘하지 못 하였음

근데 우리가 보고 싶은 것은 소비 패턴이 비슷한 지역끼리 묶고 싶다는 의도가 있었음

차원 축소 - 이따가 현서님이 업종으로 205개 한 번 해볼 예정, 이제 와서 뭘 바꾸지는 말고 그냥 이러한 시도도 해보았다는 느낌으로 발표에 넣기

혜택 아이디어

커피 할인 혜택을 주자고 했으면 커피 소비가 많다는 결과가 필요함 (근거자료)

외부 카드 자료 조사 결과 공유

국민은행 국민카드

8가지 정도 혜택이 있는데, 국민카드는 모바일 카드가 실물에 비해서 발급비가 저렴

그리고 스탬프 카드가 인상적이었음 → 적립이 되거나 나중에 청구 할인 받음

한시적인 특징을 가진 경우 많았음

기업과 연계된 카드가 많았음

ev 전기차

쿠팡 이마트 갤러리아 cj 등등

모바일 신규 고객의 경우에 ott 할인은 기본적으로 가지고 있음

전월 실적을 충족하지 못 하였어도 채워주는 제도가 인상적

전월 실적이 대체적으로 높았음 50만원이 적당하다고 개인적으로 생각하는데, 국민카드의 경우 대부분 50~150 만원 등

플래티넘이 아니더라도 이렇게 실적이 높은 경우 많았음

삼성카드

삼성 10개 정도 보셨는데 기준 실적이 훨씬 낮았음 30만원 정도

DIY가 많았음 삼성 카드도 많았음

오픈 마켓, 소셜 커머스, 트렌드 숍 등 4가지 섹션을 두고 가장 많이 쓰는 것을 할인 받을 수 있음 본인이 선택할 수 있음

개인 의견 : 만원 한도로 할인이 들어가는 경우가 많은데, 할인을 정할 때 한도를 정하는 것이 좋겠다는 생각이 들었음

해외 관련된 것도 많았음 → 공항

호텔, 골프, 고급 패션, 고급 쇼핑

1년에 15만원 정도 기프티콘 형식으로 받을 수 있음

연회비는 보통 2~3만원

비싼 건 22만원 - 플래티넘 카드

탑 5를 보셨는데도, 플래티넘 연회비 높은 것이 있다는 것 자체가 인상적이었음

인터파크 티켓이랑 여행 갈 때 적립

고급화 시키고 싶은 부분 → 5가지 기프트 중에서 1년에 한 가지 선택해서 받을 수 있음

패션도 고급 패션

삼성이라서 그런지 삼성물산과 연계해서 패션 관련된 혜택이 많았음

자라 h&m 등 패션 브랜드와 연계된 것 많았음

생각보다 삼성페이 관련된 혜택은 딱히 들어가 있지 않았었음

현대카드

사람들이 많이 사용하는 카드는 국내외 가맹점,

혜택은 적립이 훨씬 더 많음

현대는 앱 몰이 따로 있어서 적립금을 기반으로 사람들이 현금처럼 사용 가능

처음 들어갈 때 캐시백해주는 것이 있었음

한 달 썼을 때 50만원 이상 쓰면 20만원 캐시백

할인 카드도 실적 제한 없이 0.8% 할인

이거 이외에 혜택이 없음에도

무제한 적으로 할인 해줘… 최대 금액 제한이 없음

현대는 테마가 굉장히 확실해

유치원, 학원, 병원 등 카테고리 확실해 → 3040 자식이 있는 엄마가 많이 사용

모비율 검정을 하는 이유

클러스터링에 대한 근거를 대기 위함

모비율 검정 → 비율을 확인하기 위해서 시행함 20대 남녀가 하나의 그룹으로 묶이면 카이제곱을 했을 때 차이가 없을 것이다라고 생각함 즉 귀무가설이 맞을 것이라고 생가했는데 → 결과는 둘이 다르다고 나왔음

어떤 결과에서든지 20대는 성별 상관 없이 무조건 묶이길래 차이 없을 거라고 생각했는데 우리의 예상과 달랐음

그러면 우리가 mz라도 다 같은 mz가 아니다! 그래서 6개 집단만 남겨놓고 → 마케팅 방향을 다르게 잡는 방향으로 가면 좋을 듯

지금 클러스터링을 다시 한다거나 이런거도 좋기는 하지만…

가맹점\_업종명\_중분류 → 비율로 낼 수 있어서 가능 “문자열”변수이긴 하지만 counting이 가능해서…

클러스터링 진행 시 : 가맹점\_업종명\_중분류

이게 지금 우리 예상과 달라서 카이제곱을 뺴든가 해야 하는데…

모비율 검정의 결과도 모집단 간의 분포가 (거의 다) 다르다! 라고 통계 분석 결과가 나왔음

203040으로 간다고 하면, 203040 남녀 다 다르다는 카이제곱 검정을 이야기할 수 있음

이미 아이엠뱅크에서 여행 카드만들어서 대구은행이 미리 정해둔 지역 가서 쓰면 혜택 받음

아이엠뱅크 현직자들도 이게 의미가 있으니까 관련 상품이 지금까지 없었던 것일 수도 있음

마지막에는 아쉬웠던 점으로 금리로도 해보려고 했는데 의미 없었더라, 차원 축소 해보려고 했는데 차원이 낮아서 잘 안 되어서 아쉬웠다는 내용을 발표 마지막에 언급하면 좋을 듯

**11월 13일 (수)**

이번주 주말에 ppt 작업 진행해야 할 수도 있는데, 가능하면 사실 가장 좋은 것은 금요일 안으로 끝내는 것

토스뱅크에 ADS 라고 하는 것이 있는데 케이뱅크랑 했던 것인데 신규고객을 잡고 이탈고객이 되는 것을 막기 위해서 주유비 지원 캠페인을 진행해서 주유비 지원이 되면 어플에 push 알람이 와서 캠페인에 참여할 수 있도록 하였음

그래서 iM뱅크에서도 카드를 긁으면 비슷한 알림이 오는 기능 적용하면 좋을 듯

PPT 템플릿 방향 2가지

1. 비즈니스 느낌의 옛날 느낌
2. IT회사 스러운 템플릿

관련 현업자에게 문의한 결과 요즘 토스 증권을 많이 쓴다고 하는데, 미래에셋 같은 이미 대형 증권사가 있음에도 불구하고 토스 쓰는 이유가 UI가 좋아서 사용성이 좋고 더 간단하게 사용할 수 있어서 토스 증권 사용한다고 함

**11월 14일 (목)**

현진님 삼촌 카드 회사 다니셔서 어제 전화하여 조언 구함

나이, 성별, 지역 패턴 파악을 진행해서 새로운 카드를 여러 가지 만드는 것이 아니라

우리가 바우처라는 개념을 가지고 가지 않았는데 바우처 개념을 넣는 방향으로 가는 것도 고려해보라는 피드백 받음

백화점, 호텔은 여가 및 레저로 들어가 있고, 식생활 등 이미 많은 소비가 이루어지고 있는 분야 이외에 백화점, 호텔 등 적은 것에 집중하는 방향도 고려

현대 백화점 카드는 이미 있음 그래서 이것을 바우처로 돌려서 업그레이드 시키는 방향으로 진행 중임

바우처 : 업체와 업체가 만나서 서로 대량구매로 가격 저렴하게 해서 상품권 느낌으로 진행해서 호텔은 뷔페, 조식 무료 느낌으로 연회비 자체를 높게 잡아서 프리미엄 카드 만들면 될 듯

백화점 쇼핑몰은 이미 클러스터링 5곳 모두 존재하기 때문에 혜택 제공에 어려움은 없음

호텔이 일반 호텔이 아니라, 중구 신라 호텔 등 굉장히 고급화 전략이었음

유입 목적 : 우리가 현직자도 아닌데, 현실적으로 득과 실을 우리는 따질 수 없음

우리한테 원하는 것은 아이디어일 수도 있음

우리는 지금까지 많은 것에 초점을 뒀는데, 너무 중구난방임

우리가 어리고 돈이 없어서 바우처라는 개념을 모르는 것일 뿐임

그리고 다른 팀에서도 이런 개념 안 나왔음

오늘 인사이트 정리

서비스를 만들면 “페르소나” 설정해야 함

실제 기업에서 이런 프로젝트를 진행할 때 페르소나를 굉장히 세세하게 정함

진짜 완전 세부적으로 설정함

EX 20대 대학생 여성 → 이 타겟 대상으로만 제품 제작

그래서 그냥 20대 전체, 30대 전체 이렇게만 묶기에는 굉장히 평이함

다른 팀의 경우에도 20대 중 여행 좋아하는 사람을 타겟으로 굉장히 세분화 해서 목표 타겟을 명확하게 정의하고 있음. 우리도 그럴 필요 있음

카드사 전체를 놓고 보면 프리미엄형 신용카드, 생활 밀착형 신용카드 카드 양분

차트 : matplotlib을 그대로 가져가면 안 예쁘니까 파워포인트에서 직접 차트를 (삽입 → 표 만들기) 넣을 것

발표자는 1명 고정인 편이 나음 (데모 시연 제외)

질문자는 다른 사람이어도 됨 (ex 발표자가 스트림릿을 담당하지 않았을 수도 있고…)

발표 시간 초과되는 것은 안 좋아함

**11월 15일 (금), 11월 16일 (토), 11월 17일 (일)**

PPT 제작

**11월 18일 (월)**

1차 프로젝트 발표일