

고객 세분화의 이해

CRM 분석

※ 러닝스폰즈 CRM Analysis 강의 중 01고객세분화 내용을 기반으로 작성된 자료임

2024.09.11

목차

1. 고객 세분화의 이해

2. 통신사 고객 가치 세분화 Case Study

3. RFM 고객 세분화 실습

#별첨. RFM 고객 세분화 실습 스크립트

고객 세분화란?

- 고객을 하위 유사한 그룹으로 나누는 마케팅 전략
- 지리적, 인구 통계적, 가치, 행동 기반 등 다양한 기준 적용가능
- 고객을 이해하고 선택과 집중을 할 수 있도록
- 1~3개 정도의 변수를 기준으로 기반으로 cutoff 정하여 집단을 나누는 방법

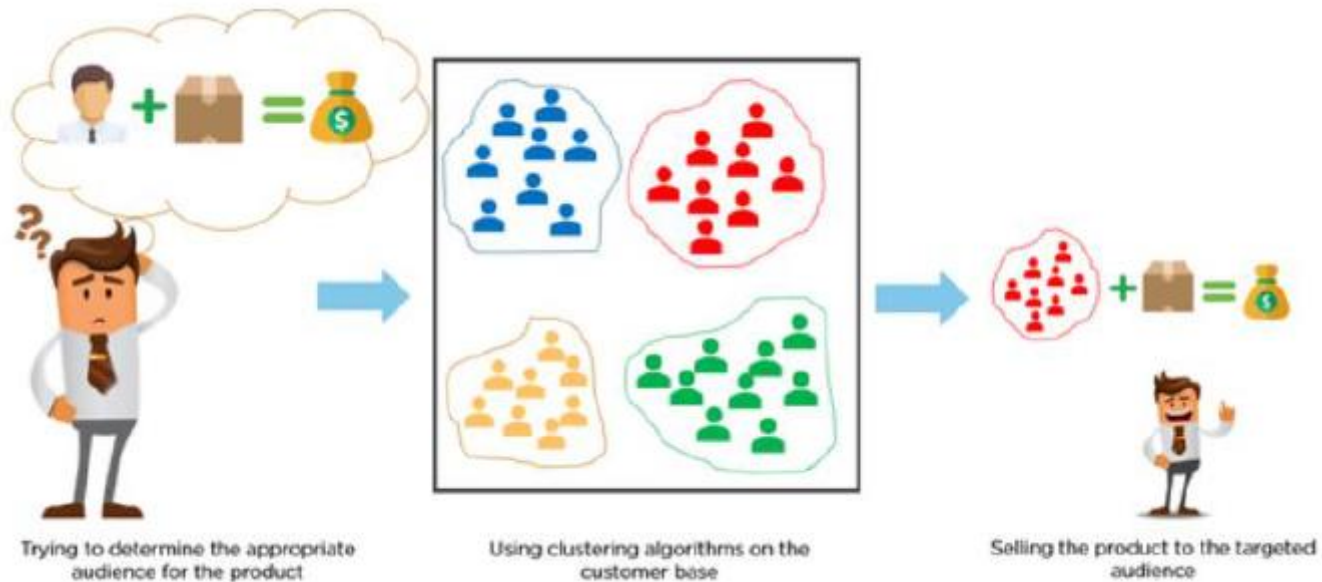
고객은 단일 집단으로 이루어져 있지 않다!

- 마케팅을 오래 하다 보면 고객이 하나의 거대한 집단으로 착각할 수 있음
- 과거 경험의 함정
 - › 우리 고객은 말이야~
 - › 내가 해봐서 아는데
 - › 과거에 이런 이벤트가 대박이었어!
- 평균 및 대표값의 함정
 - › ARPPU
 - › 평균 연령대
 - › 가장 많이 산 상품...

※ ARPPU (Average Revenue Per Paying User, 결제 유저당 평균 매출)

특정 기간 동안 결제 유저가 평균적으로 얼마나 많은 수익을 창출하는지 파악하기 위해 앱 마케터가 사용하는 측정 지표
ARPPU 계산 = 기간 동안 발생한 총 수익 / 기간의 총 결제 유저 수

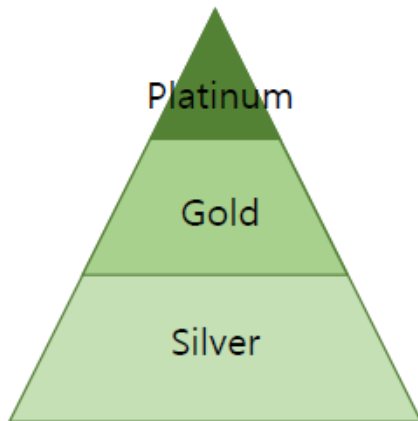
고객 세분화로 무엇을 얻을 수 있는가?



고객 세분화로 무엇을 얻을 수 있는가?

- 효과적인 캠페인
- 고객에 대한 좀 더 정확한 이해 → 고객만족도 향상 → 더 많은 이익
- 고객 니즈 파악 → 그에 부합하는 상품개발 기회
- 고객 성장 및 확장 유도
- 매출 증대

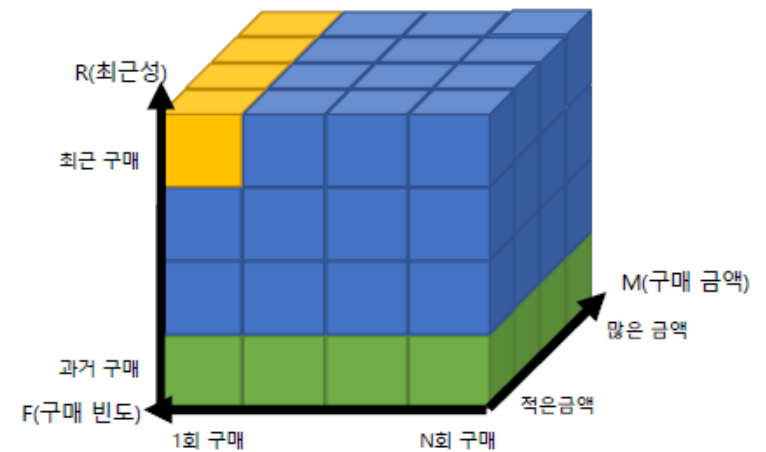
고객 세분화 대표적인 사례



고객 등급

50대 이상 여성	50대 이상 남성
40대 여성	40대 남성
30대 여성	30대 남성
20대 미만 여성	20대 미만 남성

인구 통계적 구분



RFM

어떻게 세분화를 할 것인가?

데이터 탐색 및 직관 But 논리적
알고리즘 사용 X

EDA를 통한 고객 세분화의 특징

- 목적에 맞게 분석가 직관에 따라 만듦
- 세그먼트 대한 이해가 쉬움
- 세그먼트 수 및 크기를 조절 가능
- 전략 등을 짜기 쉬움(고객 성장 및 확장)

※ EDA (Exploratory Data Analysis, 탐색적 데이터 분석)

수집한 데이터를 다양한 각도에서 관찰하고 이해하는 과정
데이터를 분석하기 전에 그래프나 통계적인 방법으로 자료를 직관적으로 바라보는 과정

EDA를 통한 고객 세분화 단계

- ① 세분화 목적 정의
- ② 목적에 맞는 개념 확인
- ③ 개념을 반영할 수 있는 데이터 확인
- ④ 변수의 분포 확인 후 분할 지점(Cut-off) 정의
- ⑤ 각 세그먼트 별 비율 확인
- ⑥ 각 세그먼트 별 프로파일링

잘 만든 세분화

- Mutually Exclusive Collectively Exhaustive (상호 배타적, 전체 포괄)
- 의미가 명확한 세분화
- 액션 가능한 세분군 (최소 개수)
- 환경의 변화에도 유지 될 수 있는 세분화
- 측정 변수가 단순한 세분화

세분화 종류 : 가치 기반

- 기업입장에서 고객의 가치를 기준으로 한 세분화
- 구매력을 기반으로 기간 내 실적을 기반으로 정의
- 현재 또는 과거의 구매력과 미래의 구매력(LTV)을 교차하여 사용 하기도 하며, RFM 등으로 다양한 축을 반영하여 고객 케어 프로그램 등의 우선순위 및 차등에 활용
- 고객 멤버십 등으로 공개하여 고객의 활동 동기부여에 활용
- 상위 등급으로 이동 전략을 세울 수 있음
- 금전적인 요소 외 비금전적인 요소도 반영 (활동력, 전파력, 지속력, 확장성)



고객 등급

세분화 종류 : 인구 통계학적

- 성별, 연령대별 관심사와 니즈가 다름
- 사람을 이해하는 가장 전통적이며 강력한 관점
- 구매력 (경제력) 및 결혼 유무 등 다양한 정보가 밑바닥에 깔려 있음
- 단, 수집을 못하는 경우도 많음
- 너무 큰 범위, 고정관념이 작용할 수 있음
- 유사 변수: 지역, 소득수준, 결혼여부, 교육수준, 종교, 주거 형태 등...

50대 이상 여성	50대 이상 남성
40대 여성	40대 남성
30대 여성	30대 남성
20대 미만 여성	20대 미만 남성

인구 통계적 구분

세분화 종류 : 행동기반

- 금전적인 요소보다 고객의 행동패턴 및 니즈 등 좀 더 근본적인 욕구 등을 파악하여 대응
- 고객의 행동은 지속적으로 유지될 가능성이 높음
- 불편한 점을 해소 하거나, 비 금전적인 보상에도 만족도와 충성도를 높일 수 있음
- 한명의 유저가 여러 행동을 보일 수 있음
- 비즈니스 로그설계와 수집 환경에 따라 크게 달리질 수 있음
- 이동정보, 특정 위치 체류시간, 사용 시간대, 결제 유형, 소비 Item 구성

세분화 종류 : 가치 기반 확장 (RFM)

RFM 분석이란?

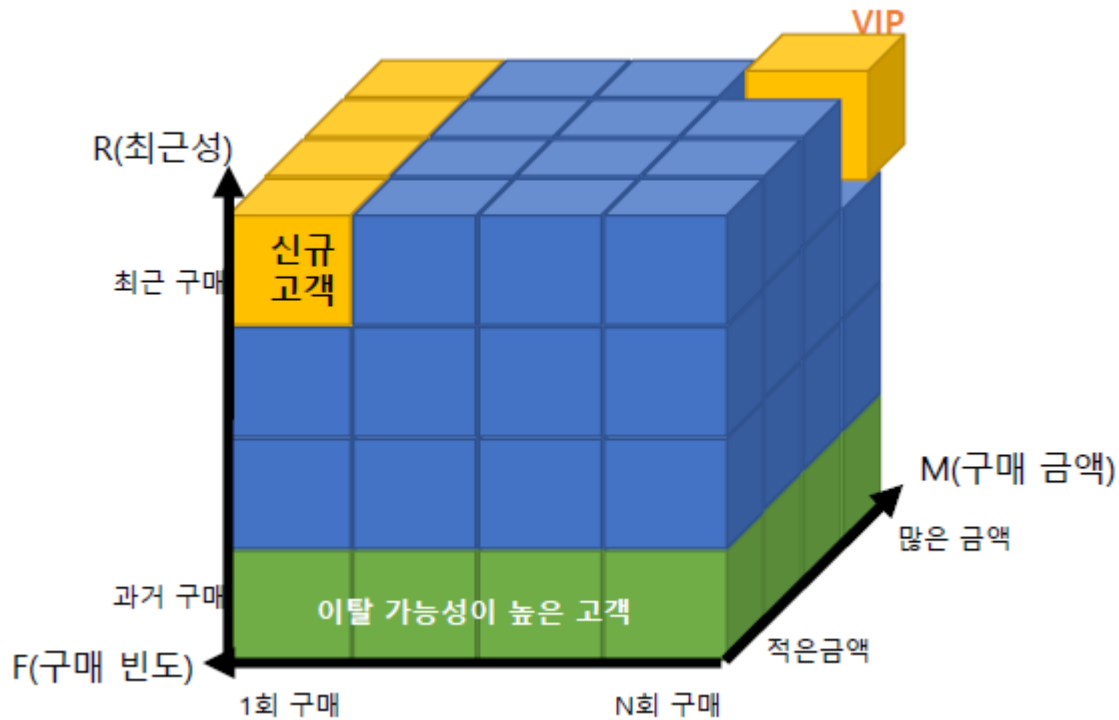
- 고객의 가치를 평가하거나 고객을 여러 세분군으로 나눠 특성을 파악
- 사용금액과 함께 CRM 관점에서 중요한 비금전적이지만 요소를 추가 하여 확장
- RFM은 Recency, Frequency, Monetary 개의 관점으로 고객의 가치를 평가
- 고객의 과거 구매 이력을 바탕으로 유사한 N개의 집단으로 분류하는 기법

Three Metrix

- Recency (거래의 최근성) : 고객이 얼마나 최근에 구입을 했는가?
- Frequency (거래 빈도) : 고객이 얼마나 자주 상품을 구입했는가?
- Monetary (거래 규모) : 고객이 얼마나 많은 금액을 지불했는가?

세분화 종류 : 가치 기반 확장 (RFM)

RFM 집단별 의미



RFM 세분화 시 주의할 점

- 각 차원별 너무 많은 숫자로 나눌 경우 각 세분별 유저수가 작아질 수 있음
 - › 주변의 집단과 특성상 큰 차이가 없을 수 있음
 - › 각 세분군 별 action을 하기 어려울 수 있음
- 3차원 데이터 실무에 사용하기 어려울 수 있음 (단순화 필요)

RFM 세분화 단순화 적용 사례

Recency 축 제거 : 기간 한정 (최근 3개월, 1년 등 기간)

- 티몬 고객 등급

멤버십 등급 조건 변경 안내

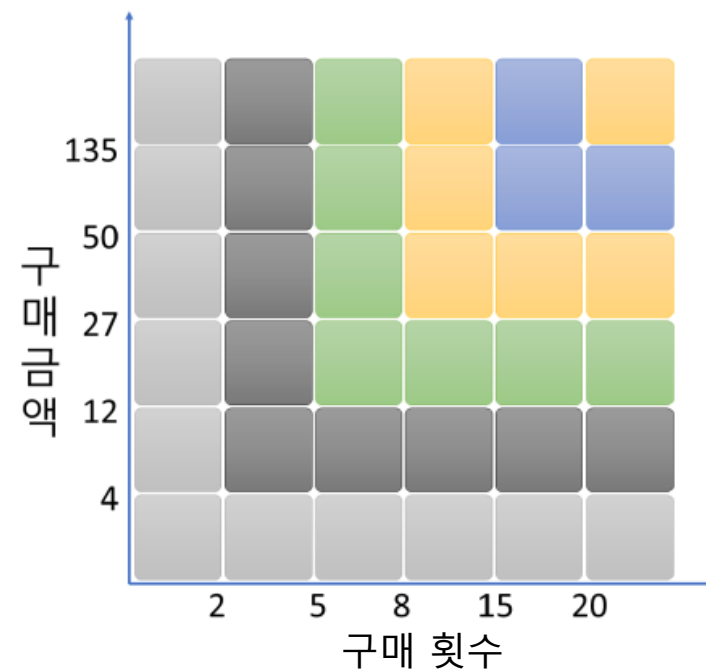
3월 10일부터 멤버십 등급 조건이 변경 적용됩니다.

변경내용	등급 산정기간 6개월 → 3개월로 변경
	등급별 구매조건(하단 표 참고)

변경 조건으로 산정한 누적/필요 구매 및 예상 등급은 기존과 동일하게 멤버십 페이지에서 확인하실 수 있습니다.

등급별 구매 조건




 1st THE FIRST 상품 구매 건수 20건 & 구매금액 135만원	 VVIP 상품 구매 건수 15건 & 구매금액 50만원	 VIP 상품 구매 건수 8건 & 구매금액 27만원
 GOLD 상품 구매 건수 5건 & 구매금액 12만원	 SILVER 상품 구매 건수 2건 & 구매금액 4만원	 NEW SILVER 등급 기준 미충족



RFM 세분화 단순화 적용 사례

Frequency 축 제거 & 기간 한정 : 상품별 가격의 차이가 크지 않은 경우

• Yes24

 플래티넘	최근 3개월간 순수 주문금액이 30만원 이상인 고객님	<ul style="list-style-type: none"> • 구매금액별 YES포인트 3% 추가 적립 (국내도서, eBook 제외) • 3,000원 포인트쿠폰 1장 (8만원 이상 주문시 / 잡지,외서,CD/LP,DVD/Blu-ray,GIFT 주문시) • 2,500원 포인트쿠폰 1장 (6만원 이상 주문시 / 잡지,외서,CD/LP,DVD/Blu-ray,GIFT 주문시) • 무료배송쿠폰 1장 (업체배송,해외배송,공연상품제외)
 골드	최근 3개월간 순수 주문금액이 20만원 이상인 고객님	<ul style="list-style-type: none"> • 구매금액별 YES포인트 2% 추가 적립 (국내도서, eBook 제외) • 2,500원 포인트쿠폰 1장 (6만원 이상 주문시 / 잡지,외서,CD/LP,DVD/Blu-ray,GIFT 주문시) • 무료배송쿠폰 1장 (업체배송,해외배송,공연상품제외)
 로얄	최근 3개월간 순수 주문금액이 10만원 이상인 고객님	<ul style="list-style-type: none"> • 구매금액별 YES포인트 1% 추가 적립 (국내도서, eBook 제외)

• L Point

회원등급별 혜택

본기별(개월) 은 :오프라인 구매실적(등급이 올라갈수록 높아지는 혜택 놓치지 마세요)

 PLATINUM 플래티넘	 GOLD 골드	 SILVER 실버	 FAMILY 패밀리
60만원 이상	35만원 이상 60만원 미만	15만원 이상 35만원 미만	15만원 미만
L.POINT 0.5% 적립	L.POINT 0.3% 적립	L.POINT 0.1% 적립	L.POINT 0.1% 적립
 3종	 3종	 3종	 2종

▶ 매 분기별 구매 실적에 따라 고객등급이 변경됩니다.

RFM 세분화 단순화 적용 사례

새로운 축 추가 : 상품 카테고리 추가 (cross-selling 목표)

- CJ ONE

— 2019년 VIP 선정 기준 —

VIP 구분	구매 금액	구매 횟수	구매 브랜드 수
SVIP	VVIP 중 내부 기준에 따른 최상위 0.1%		
VVIP	80만원	60회	4개
VIP	50만원	20회	4개

※ 2018년 연간 실적 반영 (기간 내 취소 건 제외)
※ 온라인몰은 상품 수령 후 포인트 받기 완료해야 실적 적용

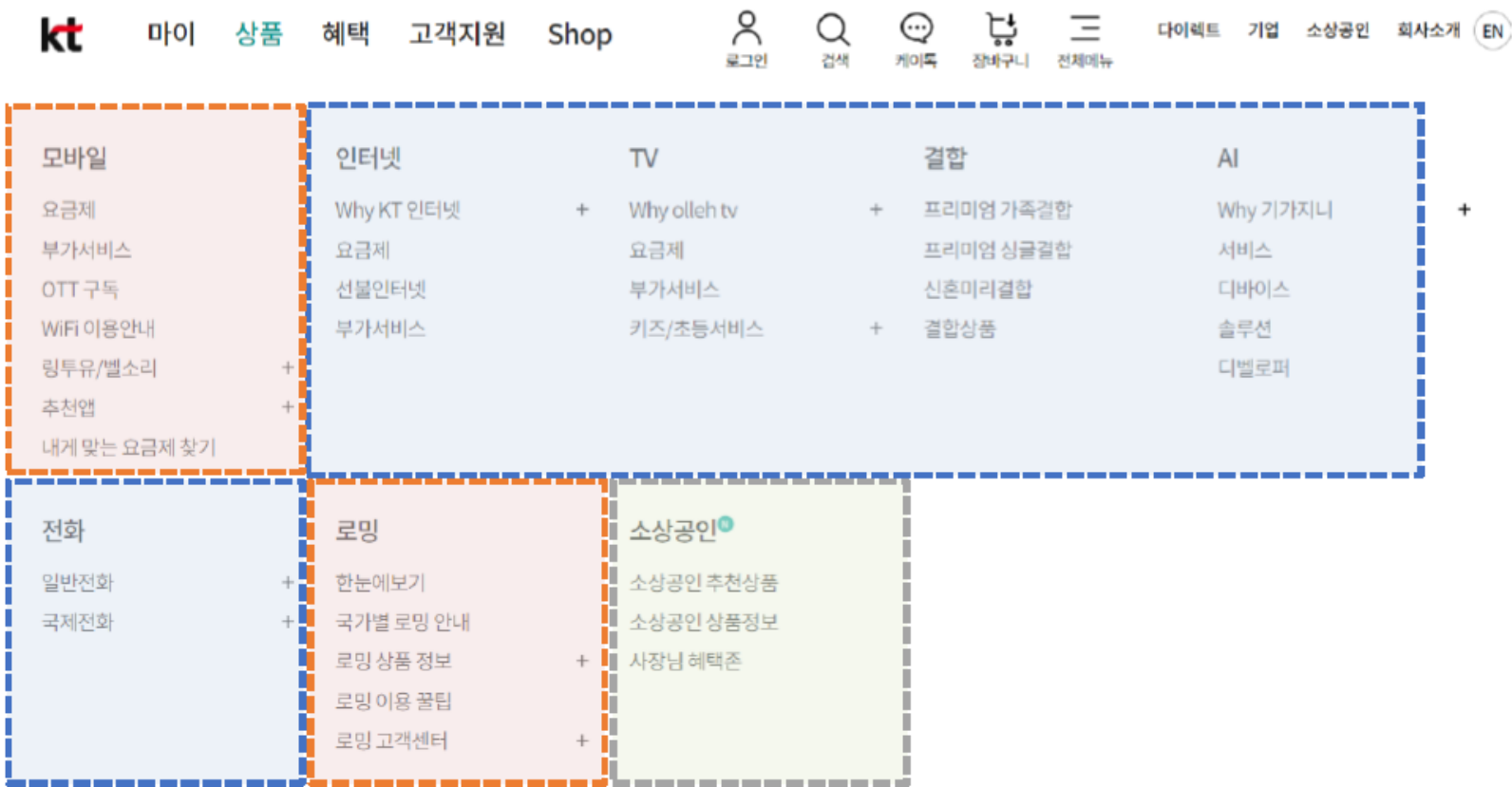
통신사 산업의 이해

- 통신사는 유무선 음성 및 데이터 통신서비스를 제공하는 기업으로 기간 통신 설비를 설치 및 운영하기 때문에 주로 국가 기간산업과 맞물려 있으며, 과점 형식으로 운영됨
- 통신 서비스는 계약 서비스 일정 기간의 사용 기간을 계약하고 회사로부터 통신 서비스를 받는 형식. 서비스는 주로 과거 유/무선 서비스로 나뉘었으나 현재 홈서비스(인터넷, IPTV, 유선전화, 보안 등...) 및 개인서비스(휴대폰)로 구분
- 과거 유선 사업부의 매출이 컸으나 무선 핸드폰 시장이 데이터 사업으로 커지며 지속 적으로 매출이 증가
- 상호 교차 판매를 통한 서비스 Lock-in 효과 객단가 (ARPU)증가
- 전 국민의 최소 20% 이상의 고객의 이동 및 통신 및 데이터 정보 보유

통신사별 주요 사업 및 서비스



통신사별 주요 사업 및 서비스



통신사별 주요 사업 및 서비스

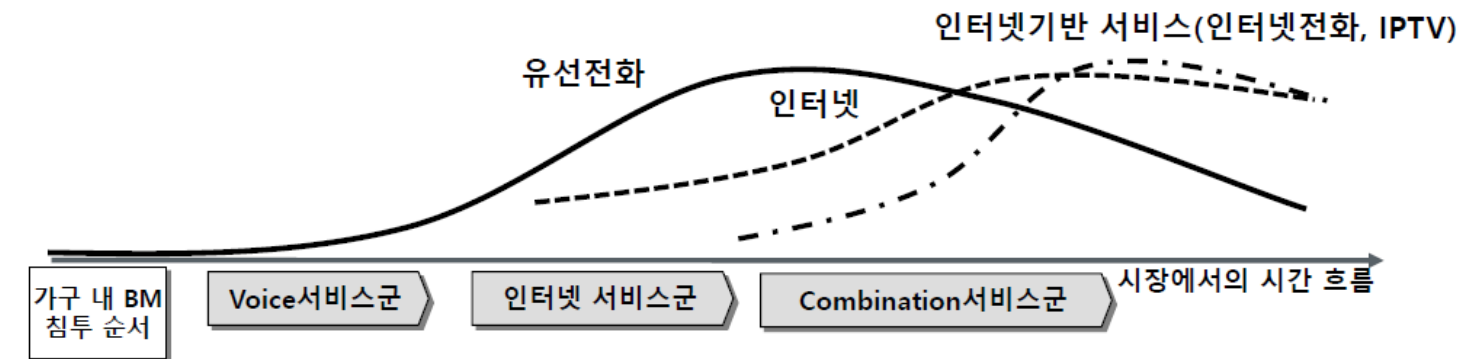
LGU+						
모바일 기기	모바일 요금제	인터넷/IPTV	마이페이지	혜택	고객지원	#유틸리티상
카테고리 - 5G - LTE	요금제 - 5G - LTE - 3G - 태블릿 및 스마트기기 - U+투게더 결합	부가서비스 - 전체 - 5G - 데이터 - 뮤직/영상 - 통화/메시지 - 생활편의 - 안심/안전	해외로밍 - 전체 - 국가별 요금검색 - 해외 로밍 요금제 - 해외 로밍 부가서비스 - 해외 로밍 상품권 - 이용안내 - 로밍센터 - 자주하는 질문 - 공지사항			
모바일 기기	모바일 요금제	인터넷/IPTV	마이페이지	혜택	고객지원	#유틸리티상
결합상품 - 요금제	인터넷 - 요금제 - 부가서비스	IPTV - 요금제 - 채널안내 - 주요 서비스 - 부가 서비스 - 공지사항	스마트홈 - 요금제 - 스마트홈 기기 - 스마트홈 소개	전화서비스 - 요금제 - 인터넷전화 부가서비스 - 인터넷전화 기기 - 수신자 부담전화 - 국제전화 002	가임사은품 - 전체 - 상품권 - TV - 컴퓨터/주변기기 - 주방/생활가전 - 에어컨/에어컨 - 키즈/맞춤품	소상공인상품 - 매장/사무실 인터넷 - 매장/사무실 인터넷전화 - 매장/사무실 지능형 CCTV - 매장/사무실 IPTV - 매장/사무실 패키지 상품 - 소상공인상품 혜택

비즈니스 상황

- 유선 사업부와 무선 사업부 합병이후 각각 운영 중이었던 서비스를 통합화해 가는 과정 진행
- 핸드폰 산업은 스마트폰의 등장으로 데이터 통신의 비율이 증가로 매출이 증가
- 통신사는 가구 단위 상품인 유선 서비스와 개인단위 상품인 무선 서비스 상품을 Mix하여 마케팅을 할 수 있는 고객 세분화를 필요한 상황

유선 상품의 가구 침투도

유선 서비스는 기존 유선전화 기반 Voice 서비스를 기반으로 가구 내 침투를 하였으며 이후 인터넷 기반의 서비스를 기반으로 그 외 추가 서비스를 제공. 이를 기반으로 3가지 Seg로 나눌 수 있음

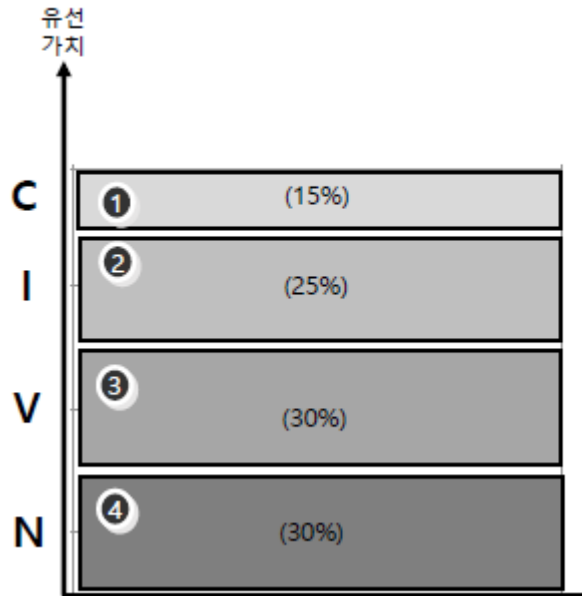


약자	가치 Seg. 명	기준 정의	BM조합
C	Combination 서비스군	음성과 인터넷 기반 서비스를 함께 사용하는 가구	인터넷 +IPTV (+ Voice서비스군)
I	Internet 서비스군	인터넷 중심 서비스를 사용하는 가구	인터넷 (+ Voice서비스군)
V	Voice 서비스군	유선전화나 인터넷전화 등 음성 서비스만 사용 가구	유선전화, 인터넷전화
N	유선 미사용	유선상품을 사용하지 않는 가구	유선 미 가입 가구

유선 고객 가치 세분화

Voice 군에서 Combination 군으로 갈수록 연령층이 낮아지고, 결합비율이 높아짐. 당연히 ARPPU 도 증가하는 모습을 보임

가구 단위 유선 가치 세분화 결과



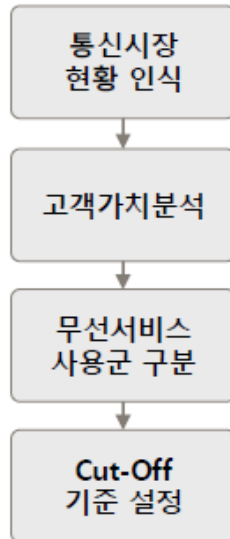
각 세분군 별 BM 이용 현황

구분	대표 고객	ARPPU	특징
① C	인터넷과 TV기반으로 사용하는 30~40대	약 5만원	<ul style="list-style-type: none"> 30~40대 : 50% 결합비율(90%)은 평균보다 4배 높음 유선전화 10년이상 가입비율이 60%
② I	유선전화와 인터넷을 사용하는 40~50대	약 4만원	<ul style="list-style-type: none"> 40~50대 : 50% 결합비율(60%)은 평균보다 3배 높음 유선전화 10년이상 가입비율이 70%
③ V	유선전화를 사용하는 60대이상	약 1.5만원	<ul style="list-style-type: none"> 60대 이상 : 40% 결합비율(5%)은 평균 ¼ 수준 유선전화 10년이상 가입비율이 80%
④ N	20~30대 유선 미 가입자	-	<ul style="list-style-type: none"> 무선 BM만 사용하는 20대 이하 가입자(50%)로 많음

무선 고객 세분화

무선 인터넷 사용량이 급격히 높아지는 기준을 기반으로 음성통화 고객과 무선 데이터 고객을 나눌 수 있는 지점을 도출

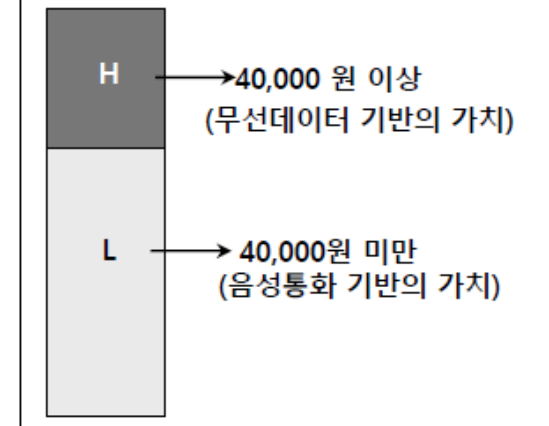
Cut-Off 도출 절차



내용

- 스마트폰 가입자 증가
- 무선데이터 데이터 사용자 증가
- 음성통화 이용 가치보다 무선데이터 이용 가치가 높음
- 음성통화 중심의 사용군과 무선데이터 중심의 사용군으로 정의
- 무선데이터 선호 고객의 ARPU를 이용한 Cut-Off 기준 설정
- 40,000원 이상은 무선데이터 중심의 사용 고객으로 40,000원 미만은 음성통화 중심의 사용 고객으로 설정함

Cut-Off 기준

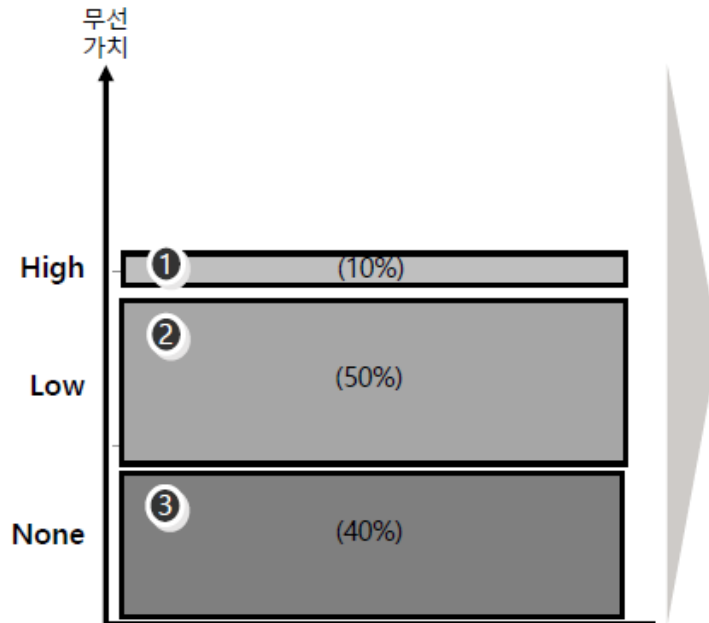


Seg. 명	기준 정의
H	무선데이터 중심의 사용 고객 모바일 ARPU 40,000원 이상
L	음성통화 중심의 사용 고객 모바일 ARPU 0~40,000원

무선 고객 세분화

무선 인터넷 사용량이 급격히 높아지는 기준을 기반으로 음성통화 고객과 무선 데이터 고객을 나눌 수 있는 지점을 도출

개인 단위 무선 가치 세분화 결과



각 세분군 별 BM 이용 현황

구분	대표 고객	무선 ARPU	특징
High	30~40대	약6만원	<ul style="list-style-type: none"> 30~40대 : 50% 스마트폰 가입율은(30%) 평균 이상 유무선결합율 : 15%
Low	20대 이하	약2만원	<ul style="list-style-type: none"> 20대 이하 : 40% 스마트폰 가입율은(15) 평균 이하 유무선결합율 : 20%
None	40~60대 이상 유선Only 고객	-	<ul style="list-style-type: none"> 40대 이상 : 80% 남성 비율이(60%)로 높음 유선전화 Only 고객이 50%

고객 세분화 도출

유선 가치와 무선 가치를 Mix 하여 11개의 세분군 도출 후 전략적인 관점에서 상호 연관성이 높은 고객군 통합

통합 고객 세분군 도출 Matrix

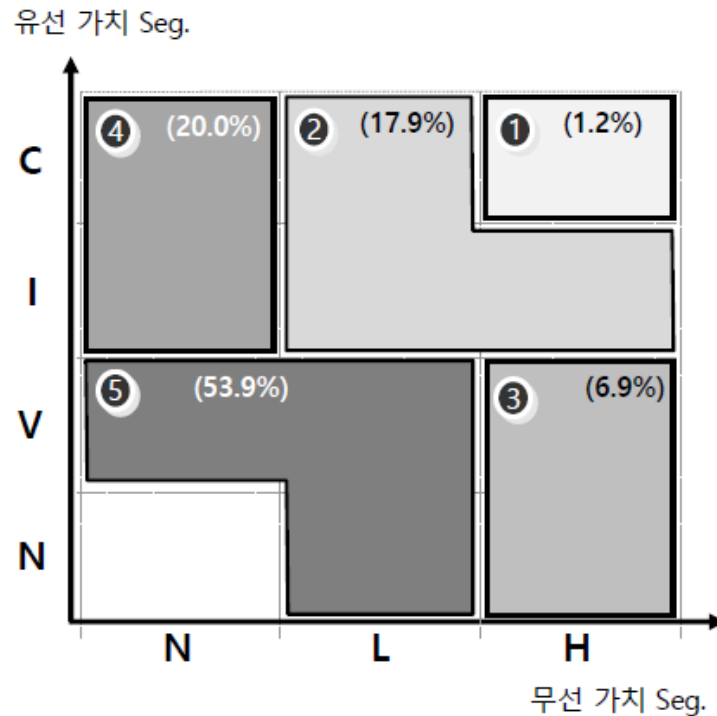
유선 가치 Seg.			
Combination (TV + 인터넷 + Voice)	CN	CL	CH
Internet (인터넷 + Voice)	IN	IL	IH
Voice (유선전화 or 인터넷전화)	VN	VL	VH
None (유선상품 미가입)	NN ¹⁾ (가망고객)	NL	NH
	None (무선 상품 미가입)	Low (무선 ARPU 4만원 미만)	High (무선 ARPU 4만원 이상)
	무선 가치 Seg.		

전략 고객군 도출 Matrix

유선 가치 Seg.			
Combination	CN	CN	CH
Internet	IN	IN	IH
Voice	VN	VL	VH
None	NN	NL	NH
	None	Low	High
	무선 가치 Seg.		

고객 세분화 도출

유선 가치와 무선 가치를 Mix 하여 11개의 세분군 도출 후 전략적인 관점에서 상호 연관성이 높은 고객군 통합

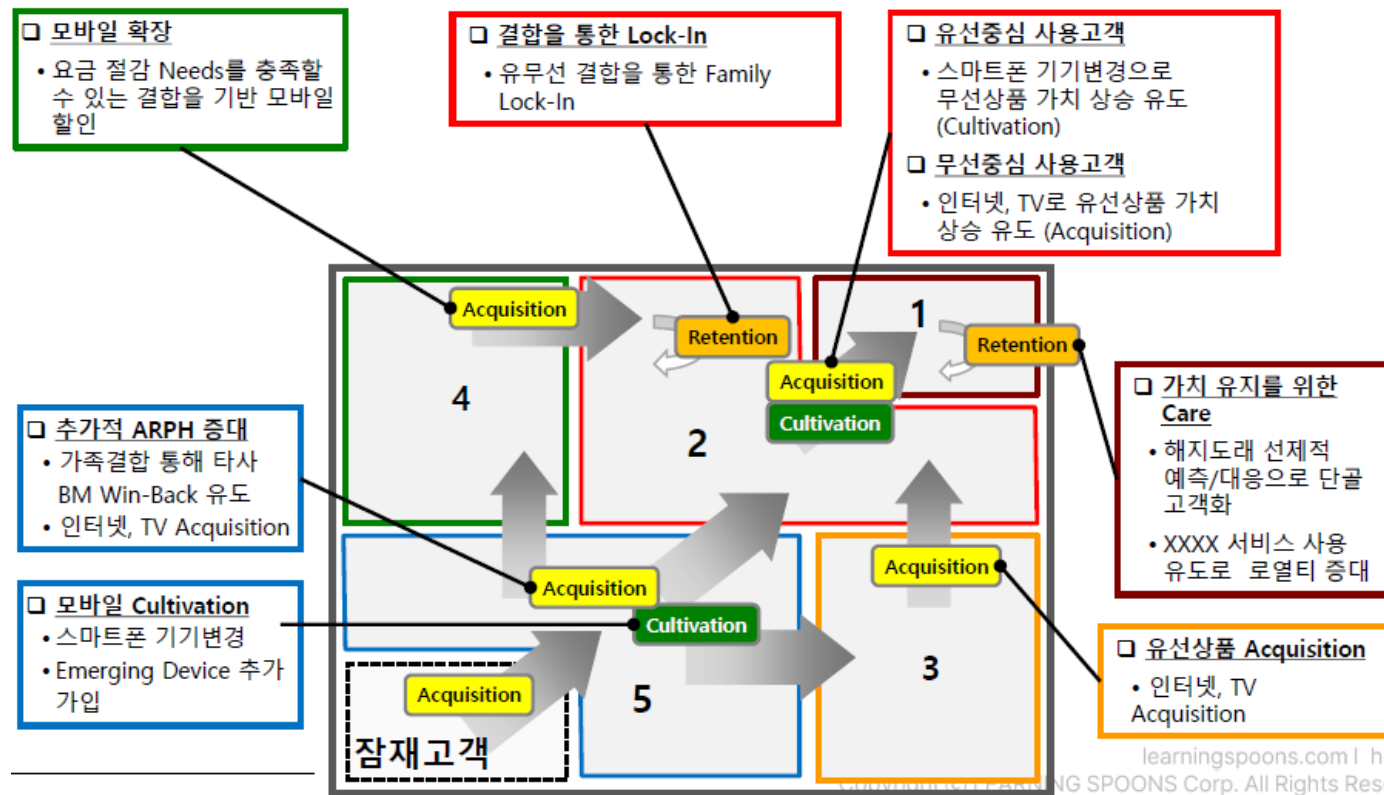


전략 세분군 별 대표고객, 평균ARPH 및 특징

구분	대표 고객	평균 ARPH	특징
1	모바일과 인터넷, TV 사용하는 20~40대	약12만원	<ul style="list-style-type: none"> 2인 이하 가구구성이 75% 평균 BM보유는 4.5개 가입기간은 2년 미만인 30% 스마트폰 보유 35%
2	모바일과 인터넷을 사용하는 30~50대	약8만원	<ul style="list-style-type: none"> 2인 이하 가구구성이 75% 평균 BM보유는 3.5개 무선과 유선의 ARPH 비중이 유사
3	모바일과 유선전화를 사용하는 20~30대	약7만원	<ul style="list-style-type: none"> 1인 가구구성이 75% 무선의 ARPH는 높으나 유선은 낮음 평균 BM보유는 1.5개 스마트폰 보유 35%
4	인터넷과 TV기반으로 사용하는 40~50대	약5만원	<ul style="list-style-type: none"> 남성의 비중이 60% 1인 가구구성이 높음 (60%) 무선과 유선의 ARPH 비중이 유사
5	모바일, 유선전화 단독으로 사용하는 60대 이상	약 2만원	<ul style="list-style-type: none"> 1인 가구 비율이 높음 (85%) 평균 BM보유는 1.1개 가입기간이 10년 이상이 40%로 길며 ARPH는 모바일 비중이 높음

고객 세분화 기반 전략 도출

각 고객군별 전략을 도출 하여 고객 성장 프로그램 실행



learningspoons.com | he
NG SPOONS Corp. All Rights Rese

데이터 세트 설명

- 이커머스 행동 데이터 (data.csv)
- <https://www.kaggle.com/allunia/e-commerce-sales-forecast/data?select=data.csv>
- UCI Machine Learning Repository 제공 (원본, <https://archive.ics.uci.edu/ml/datasets/online+retail>)
- 전체 약 54만건 UK 온오프라인 매장판매 데이터 (2010년12월 ~ 2011년12월), 45MB
- 고객 세분화(RFM)를 다루기에 매우 좋은 데이터

E-commerce Sales Forecast 데이터 설명

NO	변수명	한글설명	유형	영문설명
1	InvoiceNo	주문번호	숫자형	Invoice number. Nominal, a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation.
2	StockCode	상품 코드	문자형	Product (item) code. Nominal, a 5-digit integral number uniquely assigned to each distinct product.
3	Description	상품 설명	명목형	Product (item) name. Nominal.
4	Quantity	주문수량	연속형	The quantities of each product (item) per transaction. Numeric.
5	InvoiceDate	주문일시	명목형	Invoice Date and time. Numeric, the day and time when each transaction was generated.
6	UnitPrice	상품가격	연속형	Unit price. Numeric, Product price per unit in sterling.
7	CustomerID	고객 ID	연속형	Customer number. Nominal, a 5-digit integral number uniquely assigned to each customer.
8	Country	국가명	명목형	Country name. Nominal, the name of the country where each customer resides.

분석 프로세스

1. 데이터 품질 확인
2. EDA를 통한 주요 정보 확인
3. RFM 스코어 생성
4. RFM 모형 단순화
5. 프로파일링

1주차_리테일_고객세분화.ipynb ☆

파일 수정 보기 삽입 런타임 도구 도움말 오전 9:41에 마지막으로 저장됨

연결 Gemini

+

코드

+

텍스트

↻

0.118949 0.084371 0.034809 0.011387

↻

[] sns.heatmap(df_count_pivot, annot=True, fmt='.1%', linewidths=.5, cmap='Blues')

↻

<matplotlib.axes._subplots.AxesSubplot at 0x7f17e24ca650>

	0-1	0-2	0-3	0-4
1	2.5%	3.6%	6.7%	12.7%
2	4.7%	5.9%	7.2%	6.9%
3	5.9%	7.0%	7.6%	4.3%
4	11.9%	8.4%	3.5%	1.2%

↻

[] # 전체 데이터 수 확인
df.shape

↻

```

NameError                                Traceback (most recent call last)
<ipython-input-78-e111a67b56ed> in <module>()
      1 # 전체 데이터 수 확인
----> 2 df.shape

NameError: name 'df' is not defined

```

↻

[] temp2 = df_outlier_treat[['CustomerID', 'Country']]
temp2.groupby(['CustomerID', 'Country']).count()
temp2.

↻

더블클릭 또는 Enter 키를 눌러 수정

↻

[] RFM_data.CustomerID.unique()

LS Securities

36