



CHAPTER 06.

고객 분석 2





AGENDA

01 전략적 고객 기반
가치 측정지표

02 일반적인 고객선정전략

03 대체 고객선정전략
평가 방법



01

전략적 고객 기반 가치 측정지표

전략적 고객 기반 가치 측정지표

• 목적

- 전략적 고객 기반 가치 측정은 **미래 지향적**이며 고객의 **장기적인 수익성**을 최대한 할 수 있도록 기업의 의사결정을 유도하는 목적을 가지고 있음.

• 종류

- RFM
- 과거 고객가치
- 고객생애가치
- 고객 순자산 가치

〈표 6-1〉 고객 분석 2에서 사용한 측정지표와 측정 방법

측정지표	방법
6.2 전략적 고객 기반 가치 측정지표	6.3 대중적인 고객 선택 전략
6.2.1 RFM	6.3.1 프로파일링
6.2.2 과거 고객가치	6.3.2 이진 의사결정나무
6.2.3 고객생애가치	6.3.3 로지스틱 회귀 분석
6.2.4 고객 순자산 가치	6.4 고객 선정 전략 평가
	6.4.1 오분류율
	6.4.2 리프트 분석

RFM (1)

- 정의

- RFM은 **최근성**(recency), **구매빈도**(frequency), 그리고 **구매금액**(monetary value)을 의미함.

1. 최근성은 고객의 마지막 주문으로부터 얼마나 오래 되었는지를 측정하는 방법
2. 구매빈도는 특정 기간 동안 고객이 얼마나 자주 주문했는지를 측정하는 방법
3. 구매금액은 고객이 평균 거래에 지출한 금액

- RFM의 일반적인 개념은 RFM을 기반으로 고객을 **분류**하는 것이며 분류된 고객 그룹은 **구매 행동**(예: 마케팅 캠페인에 반응할 가능성)과 관련됨.
- RFM은 상태 공간이라 불리는 시간에 걸쳐 고객의 행동을 추적한다는 점에서 **전환 매트릭스 접근법**(transition matrix approach)과 유사함.

RFM (2)

- 방법
 - 방법 소개를 위해 사용될 예시 사례

대상기업은 현재 **40만 명의** 고객을 보유하고 있다.

이 중에서 우선 테스트를 위해 **4만 명의** 고객 표본을 선정하였다. 구체적으로 전체 고객을 대표하는 4만 명의 고객으로 테스트 그룹을 구성하기 위해 40만 명의 고객 데이터베이스에서 **매 10번째 고객**을 선택하는 방식으로 이들을 선정하였다.

현재 대상기업은 **150달러 할인 쿠폰**을 고객에게 우편으로 발송하는 캠페인을 진행하려 하고 있다. 그래서, 4만 명의 테스트 그룹에 해당 쿠폰을 발송해 본 결과, 총 **808명의 고객**(4만 명의 2.02%)이 반응한 것으로 나타났다.

만약 이 사례에서 소개된 대상기업이 R(최근성), F(구매빈도), M(구매금액) 코딩을 수행한다고 하면...

RFM (3)

- 방법

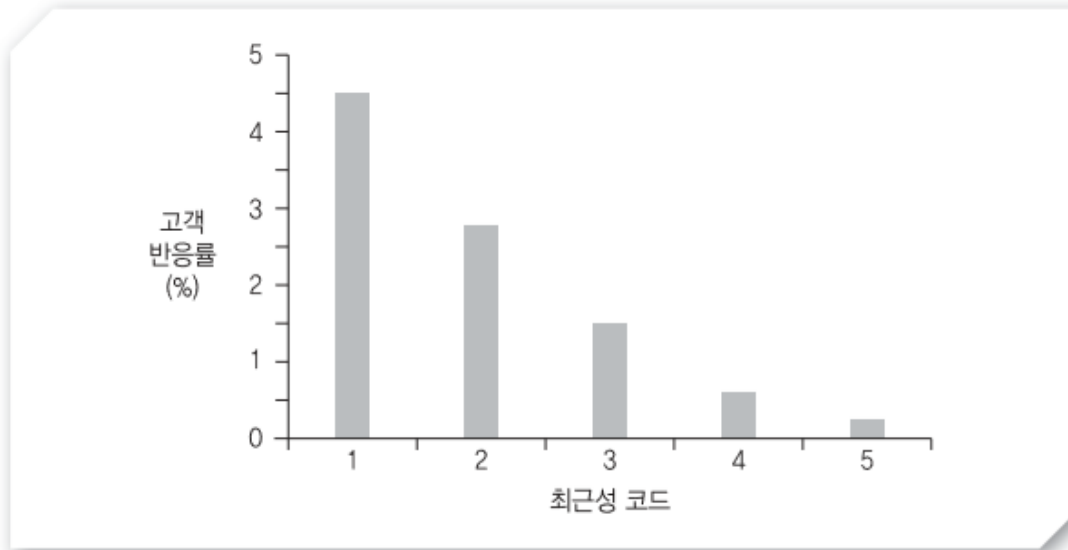
- 최근성 코딩

1. 4만 명 고객의 테스트 그룹에서 최근 구매일을 기준으로 내림차순으로 정리함.
 2. 정리된 데이터를 같은 크기의 5개 그룹(각 그룹의 20%)으로 나눔.
 3. 맨 위의 그룹에 최근성 코드 1을 부여하고 다음 그룹에게는 코드 2, 이런 순으로 맨 밑의 그룹이 코드 5를 부여 받을 때까지 진행.
- 캠페인과 최근성을 기반으로 한 고객 반응 분석에서 최근성 코드가 1인 고객 그룹이 캠페인에 대한 응답률이 가장 높으며, 다음은 코드 2를 가진 그룹이라는 것을 알 수 있음.
 - 최근성 코딩이 끝나면 고객이 속한 5분위 수에 따라 고객 그룹을 $r=1$ 부터 5까지의 값으로 최근성에 대한 값을 부여함.

RFM (4)

- 방법
 - 최근성 코딩

〈그림 6-1〉 고객 반응과 최근성



- 캠페인에 대한 반응률이 가장 높은(4.5%) 그룹은 최근성 코드가 1인 그룹에 속한 고객
- 5개의 그룹에 대한 평균 고객 반응률은 캠페인에 의해 달성된 실제 반응률인 **2.02%**
 $(4.50\% + 2.80\% + 1.50\% + 1.05\% + 0.25\%)/5 = 2.02\%$

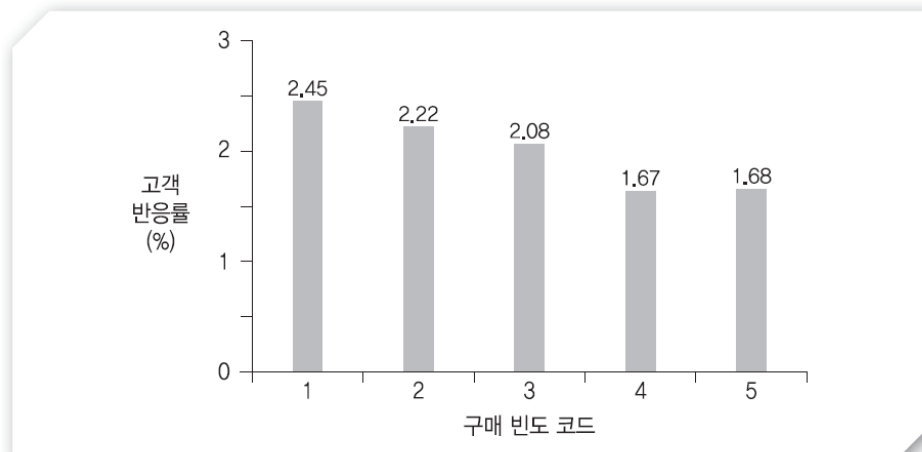
RFM (5)

• 방법

• 구매빈도 코딩

- 구매빈도 코딩 과정은 앞의 최근성과 유사하나, 기준이 **월 평균 구매횟수**가 됨.
- 5분위수를 기준으로 구매빈도가 상위 20% 범주에 속하는 고객에게 f=1을 부여하고, 그 다음 차상위 20%에 f=2... 이런 식으로 코드를 부여함.

〈그림 6-2〉 고객 반응과 구매빈도



- 캠페인의 가장 높은 반응률(2.45%)은 **가장 높은 구매빈도**(구매빈도 코드=1)을 가진 고객으로 나타남.

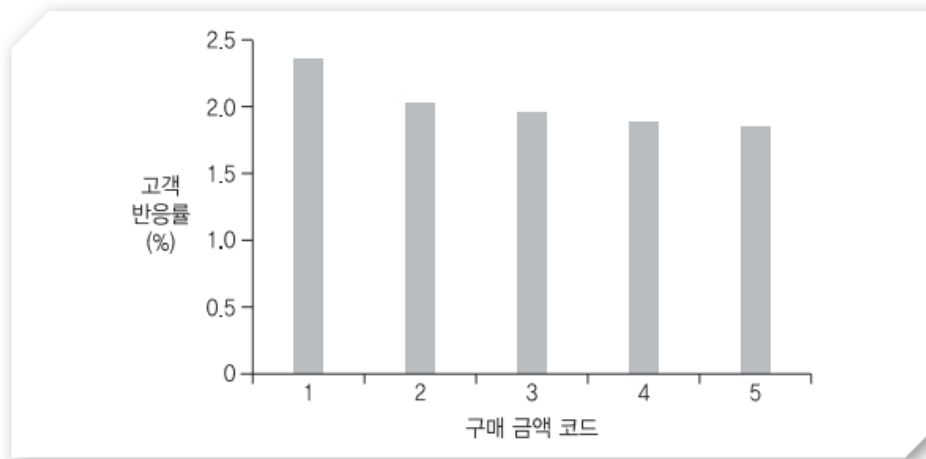
RFM (6)

• 방법

• 구매금액 코딩

- 구매금액 코딩 과정은 앞의 구매빈도와 동일하며, 기준이 **월 평균 구매금액**이 됨.
- 5분위수를 기준으로 구매금액이 상위 20% 범주에 속하는 고객에게 m=1을 부여하고, 그 다음 차상위 20%에 m=2... 이런 식으로 코드를 부여함.

〈그림 6-3〉 고객 반응과 구매금액



- 캠페인의 가장 높은 반응률(2.35%)은 **가장 높은 구매금액**(구매금액 코드=1)을 가진 고객으로 나타남.

RFM (7)

- 코딩 결과 종합

- 이러한 과정(R, F, M)을 수행한 후, **각 고객에 대한 R, F, M의 개별 점수**를 알 수 있음.
(〈그림 6.4〉 참조)
각 고객은 각각의 RFM 코드에 따라 125개의 그룹 중에 111, 233, 432, ..., 555 같은 그룹 중 하나에 배정됨.

- 한계

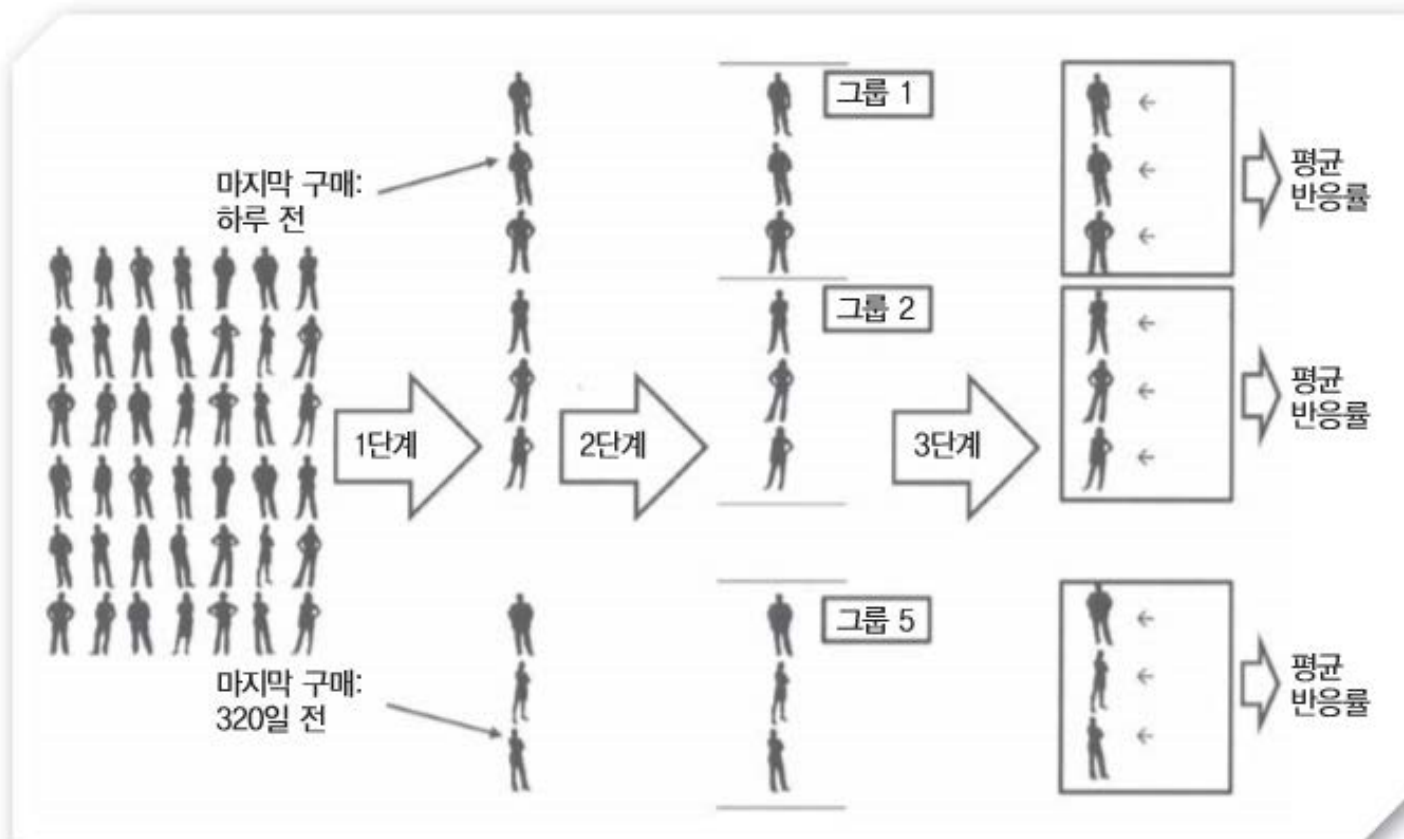
- 캠페인에 대한 고객 반응 데이터를 R, F, M값과 독립적으로 연결하여 특정 RFM 코드에 의해 고객을 그룹화하는 이 방법은 각 RFM 구성 단위에 **동일한 수의 고객**을 생성하지 않을 수 있음.(개별적인 R, F, M이 서로 연관되어 있기 때문)

ex) 더 많은 금액(높은 M)을 소비하는 사람은 평균적으로 더 자주 구매할 가능성이 있다.
그러나 현실적인 목적을 위해서 각 RFM 셀은 동일한 수의 고객을 갖는 것이 바람직하다.

RFM (8)

• RFM 셀 분류 기법

〈그림 6-4〉 RFM 절차



RFM (9)

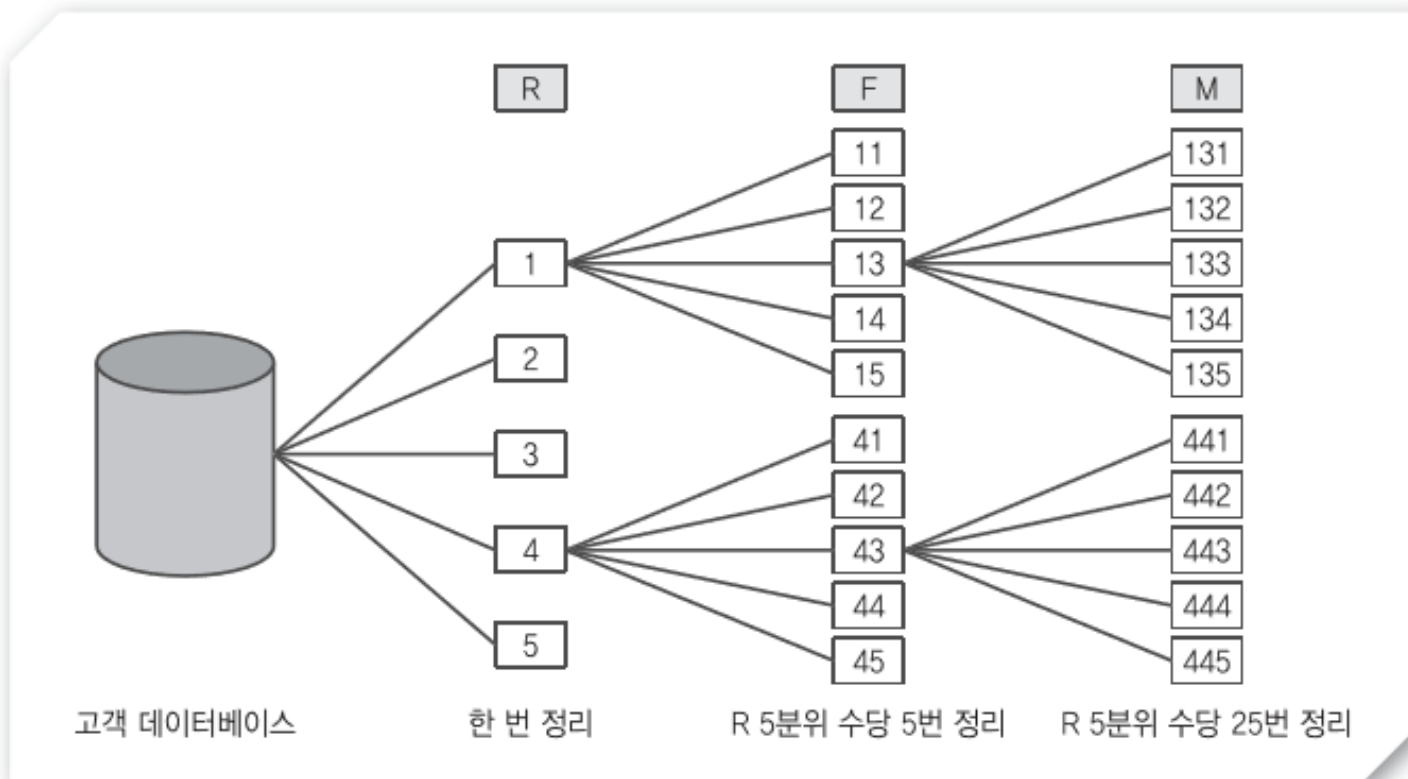
• RFM 셀 분류 기법

- 1) 4만 명의 테스트 그룹 고객 목록이 있을 때, 고객 목록을 먼저 날짜순으로 정렬하여 8,000명의 고객으로 구성된 **5개의 동등한 그룹으로 그룹화**함.
- 2) 각 그룹에 있는 8,000명의 고객을 대상으로 **구매빈도를 기준으로** 1,600명씩 5개의 그룹으로 나눔.
- 3) 각 그룹의 1,600명 고객은 11에서 55까지의 **RF 코드**를 가지게 됨.
- 4) 각각의 RF 그룹을 **구매금액을 기준으로** 다시 분류하여 320명의 고객으로 구성된 5개의 그룹으로 나눔. 즉, 111에서 555까지의 RFM 코드를 가지며, 각 320명의 고객을 보유하게 됨.
- 5) 각 RFM 코드를 셀로 간주하면 125개의 셀을 가짐
(**5개의 최근성 × 5개의 구매빈도 × 5개의 구매금액 = 125개의 RFM 코드**)

RFM (10)

• RFM 셀 분류 기법

〈그림 6-5〉 RFM 셀 정리



RFM (11)

• 손익분기 가치

• 정의

- 마케팅 문헌에서 손익분기란 마케팅 프로모션을 통해 얻은 **순수익**과 프로모션 제공과 관련된 **비용**이 **같음**을 의미.

$$\text{손익분기} = \frac{\text{단위당 비용}}{\text{단위당 순이익}} \quad \langle \text{식 6.1} \rangle$$

- 만약 특정한 프로모션에 대한 비율이 **1**이면 프로모션은 **어떠한 순이익도 발생시키지 않았음**을 의미함.
- 이 비율은 프로모션 비용을 상쇄하기 위해 필요한 최소 반응률을 산출할 수 있기에 손익분기(BE)를 **손익분기점 회수율**이라고도 함.

RFM (12)

• 손익분기 가치

• 예제

150달러 짜리 할인 쿠폰 메일링의 예시를 고려해 보자. 우편 한 개당 발송 비용이 **1달러**라고 가정하면 순이익은 **45달러**이다. 그리고 필요한 손익분기 혹은 손익분기 회수율은 **0.0222(=\$1/\$45)** 또는 **2.22%**로 계산할 수 있다.

이 값을 계산한 다음 각 RFM 셀의 실제 응답률과 비교할 수 있다.

비교를 단순화하기 위해 계산된 손익분기 회수율은 모든 RFM 셀을 위한 **손익분기점**(breakeven index: BEI)을 계산하는 데 사용될 수 있다.

$$\text{손익분기점} = \frac{\text{실제 응답률} - \text{손익분기 회수율}}{\text{손익분기 회수율}} \quad \langle \text{식 6.2} \rangle$$

- 손익분기점 값이 **양** - 거래로부터 수익이 발생됨
- 손익분기점 값이 **0** - 거래가 손익 평형을 이룸
- 손익분기점 값이 **음** - 거래가 손실을 야기했음

RFM (13)

• R, F, M의 중요도 순서

- 대부분의 비즈니스에서는 최근성, 구매빈도, 구매금액의 순서로 RFM 방법을 사용함.
- 그러나 순서는 **산업에 따라** 다르며 보다 정확한 코딩 순서는 **고객 반응률의 하락 속도**에 따름.
- 고객 반응률이 더 빨리 하락하는 측정 지표(R, F, M)는 향후 고객 반응을 예측하는 가장 좋은 지표가 될 수 있으므로 먼저 코딩해야 함.

〈표 6-3〉 타겟팅 캠페인 테스트를 위한 수익의 비교

구분	테스트	고객 전체 기반	RFM 선정
평균 응답률	2.02%	2.02%	15.25%
응답 수	808	8,808	2,372.8
평균 순이익/판매	\$45	\$45	\$45
순세입	\$36,360	\$363,600	\$122,976
메일을 보낸 수	40,000	400,000	17,920
메일당 비용	\$1.00	\$1.00	\$1.00
메일 발송 비용	40,000.00	400,000.00	17,920.00
수익	(-\$3,640.00)	(-\$36,400.00)	(-\$105,056.00)

모든 고객을 타겟팅할 때와, RFM을 적용하여 선정된 고객만 타겟팅 할 때의 수익을 비교한 표. 선택한 RFM 셀의 고객에게만 우편 발송을 했을 때, 전체 고객에게 우편을 발송하는 것보다 상당히 많은 이익을 얻을 수 있음을 알 수 있음.

RFM (14)

- R, F, M의 상대적 중요도
 - 예제

〈표 6-4〉 최근성 점수

고객	구매 횟수	최근성 (달)	할당된 점수	가중된 점수
John	1	2	20	100
	2	4	10	50
	3	9	3	15
Smith	1	6	5	25
Mags	1	2	20	100
	2	4	10	50
	3	6	5	25
	4	9	3	15

최근성 점수: 지난 2개월 내이면 20점, 지난 4개월 내이면 10점, 지난 6개월 내이면 5점, 지난 9개월 내이면 3점, 지난 12개월 내이면 1점. 상대적 가중치 = 5.

〈표 6-5〉 구매빈도 점수

고객	구매 횟수	구매 빈도	할당된 점수	가중된 점수
John	1	1	3	6
	2	1	3	6
	3	1	3	6
Smith	1	2	6	12
Mags	1	1	3	6
	2	1	3	6
	3	2	6	12
	4	1	3	6

구매빈도 점수: 12개월 내에 구매 횟수당 3점, 최대 15점. 상대적 가중치 = 2.

RFM (15)

- R, F, M의 상대적 중요도
 - 예제

〈표 6-6〉 구매금액 점수

고객	구매 횟수	구매 금액	할당된 점수	가중된 점수
John	1	40	4	12
	2	120	12	36
	3	60	6	18
Smith	1	400	25	75
Mags	1	90	9	27
	2	70	7	21
	3	80	8	24
	4	40	4	12

구매금액 점수: 12개월 내에 구매한 금액의 10%, 최대 25 점. 상대적 가중치 = 3.

〈표 6-7〉 RFM 누적 점수

고객	구매 횟수	가중된 점수	누적 점수
John	1	118	118
	2	92	210
	3	39	249
Smith	1	112	112
Mags	1	133	133
	2	77	210
	3	61	271
	4	37	308

RFM (16)

- R, F, M의 상대적 중요도

- 예제

세 고객은 지난 12개월 동안 계산된 구매 내역을 가지고 있다.

모든 고객에게 RFM 공식에 따른 각 거래에 대한 점수가 할당된다.

과거의 고객 거래에 대한 분석을 기반으로 R, F, M에 할당된 중요도를 기준으로 한 상대적 가중치는 다음과 같다.

최근성 = 5, 구매빈도 = 2, 구매금액 = 3

누적 점수 결과, **John 249점, Smith 112점, Mags 308점**으로 나타났다.

이 예시에서 John은 유망한 고객으로 보이지만, Smith에게 메일을 보내는 것은 적절하지 못한 마케팅 활동이 될 수 있다.

이 예시는 RFM의 간단한 적용을 보여주지만, 실제로는 분석해야 할 고객의 수가 수백만 명이 될 수 있다.

RFM (17)

- R, F, M의 상대적 중요도

- **회귀 방법**을 사용하여 R, F, M의 상대적 가중치를 계산하는 방법이 있으며, 이러한 상대적 가중치는 **각 고객의 누적 점수를 계산**하는 데 사용됨.
- 계산된 점수가 **높을수록** 미래에 **수익성이 더 높은 고객**이 될 가능성이 있음.

- 평가

- RFM 기법은 **구매 가능성이 높은 고객을 식별**하고, 타겟팅 할 뿐만 아니라 구매 확률이 낮은 고객에 대한 값비싼 캠페인 비용을 피할 수 있게 함.
- 구매 확률이 높은 고객만을 식별하여 기업의 **마케팅 전략 및 커뮤니케이션**을 제공할 수 있게 해 줌.
- RFM의 한계점은 가망 데이터가 아닌 **과거 고객의 데이터에만 적용 가능**하다는 것.

과거 고객가치 (1)

• 정의

- 과거 고객가치(past customer value: PCV)는 과거 거래의 결과가 고객의 **향후 공헌도**를 나타내는 지표라고 가정하는 측정기준
- 과거에 고객이 제공한 **총 공헌이익**에 따라 고객의 가치가 결정됨.
- 고객의 과거 성과가 **미래의 수익성 수준**을 나타내는 것으로 가정함.
- 고객의 생애 기간 동안 각기 다른 시점에 제품 또는 서비스가 구매되기 때문에 모든 거래는 **시간 가치에 맞게 조정되어야 함**.

$$\text{과거 고객가치} = \sum_{t=0}^T GC_{i(t_0-t)} \times (1 + \delta)^t \quad \langle \text{식 6.3} \rangle$$

i = 고객

t = 시간 지표

δ = 적용 가능한 할인율(예: 매달 1.25%)

t_0 = 현재 시점

T = 고려되어야 할 현재 기간 이전의 기간의 수

$GC_{it} = t$ 기간 동안 고객 i 의 거래의 총기여

과거 고객가치 (2)

• 예제

〈표 6-8〉 고객의 소비 패턴

	1월	2월	3월	4월	5월
구매량(\$)	800	50	50	30	20
GC	240	15	15	9	6

총공헌도(GC) = 구매량 × 구매 마진

일정 기간 동안 고객이 구매한 제품에 대해 구매 가치 및 공헌이익에 대한 데이터가 있다고 가정하면 모든 거래를 **현재 가치로 계산**하여 고객으로부터 생성된 가치를 계산할 수 있다. 공헌이익이 **0.3**이라고 하고, 월별 할인율 $r = 1.25\%$ 일 때, 표의 고객 소비 패턴에 대한 과거 고객가치(PCV)는 다음과 같이 계산된다.

$$\begin{aligned}
 PCV_t &= 6(1+0.0125)^0 + 9(1+0.0125)^1 + 15(1+0.0125)^2 + 15(1+0.0125)^3 \\
 &\quad + 240(1+0.0125)^4 = 302.01486
 \end{aligned}$$

이 고객의 5월에 대한 순현재가치는 **302.01달러**이다.

과거 고객가치 (3)

• 평가

- 고객의 과거 가치를 비교함으로써 향후 마케팅 노력을 위한 **우선순위**를 매길 수 있음.
- 높은 가치를 가진 고객은 일반적으로 더 많은 마케팅 자원을 누릴 자격이 있다고 여기는 이 방법은 매우 유용하지만 수익성 있는 고객을 선정하는 과정을 **개선하는 데 도움을 줄 수 있는 다른 정보**를 포함하지 않음.

ex) 고객이 미래에도 활동하는 고객일지에 대한 것은 고려하지 않음.

미래에 고객을 유지 관리하는 데 필요한 예상 비용을 포함하지 않음.

생애가치 측정기준 (1)

• 정의

- 고객에 대한 장기간의 경제적 가치를 설명하는 일반적인 용어는 **생애가치**(lifetime value: LTV)이며, **고객생애가치**(customer lifetime value: CLV)로도 사용됨.
- 기업에 대한 고객의 가치를 **순현재 가치**로 할인하여 평가함.

• LTV의 계산

〈그림 6-7〉 LTV 계산의 원리



- LTV를 계산하기 위한 방식은 다양하며, 요구사항에 따라 적용할 수 있는 방식으로 원칙을 제시하는 것이 중요함.

생애가치 측정기준 (2)

• LTV의 기본 모델

- 개별 고객 i 의 생애가치는 각각의 **관찰 기간 T 에 대해 할인된 총공헌의 합계**

$$LTV_i = \sum_{t=1}^T GC_{it} \left(\frac{1}{1 + \delta} \right)^t \quad \langle \text{식 6.4} \rangle$$

i = 고객

t = 기간

δ = 이율(혹은 할인율)

GC_{it} = t 기간 동안 고객 i 의 총공헌

T = 관찰 시점

LTV_i = 한 고객의 생애가치

시점 t 에서 고객 i 의 순현재가치 = 0

- LTV의 결과는 기업에 대한 **개별 고객의 가치**를 측정한 것
- 일반적으로 **과거 고객 행동**을 기반으로 하며, 향후 의사결정을 위한 가치에는 제한적일 수 있음.
- 만약 시간 단위가 연간 단위와 다른 경우 **이자율**은 이에 따라 조정되어야 함.
예를 들어, 연간 이자율이 15%인 경우 분기별 이자율은 3.56%이다.

생애가치 측정기준 (3)

• 분리된 수익과 비용을 가진 LTV

$$LTV_i = \sum_{t=1}^T \{ (S_{it} - DC_{it}) - MC_{it} \} \left(\frac{1}{1 + \delta} \right)^t \quad \langle \text{식 6.5} \rangle$$

i = 고객

t = 기간

T = 관찰 시점

δ = 이율(혹은 할인율)

S_{it} = t 시점에 고객 i 의 판매가치

DC_{it} = t 시점에 고객 i 가 구매한 제품의 직접 경비

MC_{it} = t 시점에서 고객 i 의 직접적인 마케팅 비용

LTV_i = 고객 i 의 생애가치

시점 t 에서 고객 i 의 순현재가치 = 0

- 이 예시에서 비용 요소는 제품과 관련된 **직접 비용**과 **마케팅 비용**으로 분류됨.
- 데이터의 가용성에 따라 서비스 관련 비용, 배송 비용 또는 기타 관련 비용 요소를 포함시킬 수 있음.

생애가치 측정기준 (4)

• 고객유지 가능성을 포함한 LTV

$$LTV_i = \left[\sum_{t=1}^T \left(\prod_{k=1}^K Rr_k \right) GC_{it} \left(\frac{1}{1+\delta} \right)^t \right] - AC_i \quad \langle \text{식 6.6} \rangle$$

i = 고객

t = 기간

T = 고려 중인 시점

δ = 이율(혹은 할인율)

Rr_t = 시점 t 의 평균 유지율(개인 수준의 유지 가능성 Rr_{it} 를 사용해도 되지만 대부분 얻기 힘들다. 5.4.2절 참조.)

GC_{it} = 기간 t 에서 고객 i 의 총공헌

AC_i = 고객 i 를 획득하기 위한 비용(획득비용)

- 고객유지 가능성을 고려하는 방법을 나타내며 이것은 고객이 **평균 유지율 Rr** 에 의해 근사화된 확률로 관계를 유지하는 경향이 있다는 사실과 관련이 있음.
- 이 수식에서 용어 $\prod_{k=1}^K Rr_k$ 는 실제 생존율 SR_t 와 같음.

생애가치 측정기준 (5)

• 지속적인 유지율과 총공헌을 가진 LTV

- ‘ $T \rightarrow \infty$ ’ 라는 가정에서 시간이 지남에 따라 유지율(R_r)은 **일정**하고, GC가 변하지 않는다는 가정하에 <식 6.6>은 다음과 같은 공식으로 단순화될 수 있음.

$$LTV_t = GC_t \left(\frac{R_r}{1 - R_r + \delta} \right) - AC_t \quad \langle \text{식 6.8} \rangle$$

- $\frac{R_r}{1 - R_r + \delta}$ 를 **마진 승수**라고 부름.
- 유지율이 매우 높지 않다면 보다 정교한 공식을 적용했을 때의 결과와 매우 비슷한 결과를 얻을 수 있음.

생애가치 측정기준 (6)

- 생애 기간은 얼마나 되나?

- 생애는 다양한 상황이 존재하기 때문에 그대로 받아들여서는 안됨.
생애라는 용어는 일회용 구매(예: 집)의 경우는 이치에 맞지만, 식료품 구매자의 LTV에 대해 이야기하는 것은 어색할 수 있음.
- 모든 현실적인 목적을 위해, 생애는 **관리자가 사용하는 장기간의 지속 시간**이라고 할 수 있음.

- LTV에 외부 효과 통합하기

- 고객이 기업에 제공하는 가치는 상품 및 서비스 구매로 인한 수익만으로 구성되지는 않음.
따라서 통신 기술과 온라인 소셜(예: 페이스북), 제품 평가 웹사이트 등의 제품 또는 브랜드에 대한 개인적인 견해가 고객의 평생 가치에 상당한 기여를 할 수 있음.
- 이러한 모든 활동을 **구전**(word-of-mouth: WOM)이라 함.

생애가치 측정기준 (7)

• 구전의 측정 및 통합

$$LTV_i = \left[\sum_{t=1}^T \left(\prod_{k=1}^K Rr_k \right) (GC_{it} + n_{it} ACS_t) \left(\frac{1}{1 + \delta} \right)^t \right] - AC_i \quad \langle \text{식 6.9} \rangle$$

i = 고객

t = 기간

T = 고려 중인 시점

δ = 이율(혹은 할인율)

Rr_t = 시점 t 의 평균 유지율

GC_{it} = 기간 t 에서 고객 i 의 총공헌

n_{it} = 시점 t 에서 고객 i 의 소개로 인한 신규고객의 수

ACS_t = 시점 t 에서 고객 i 의 소개를 통해 얻은 고객당 평균 획득비용 절감액

AC_i = 고객 i 를 획득하기 위한 비용(획득비용)

- 구전은 **고객의 획득비용을 감소**시킴으로써(혹은 부정적인 구전으로 인해 획득비용이 증가하는 경우) 신규고객획득에 **직접적인 영향**을 미칠 것으로 예상됨.
- <식 6-9>는 강한 영향력을 가지는 오피니언 리더와 개인 수준인 n_{it} 를 통해 고객의 소셜 네트워크에는 차이가 존재한다는 사실을 설명함.

생애가치 측정기준 (8)

• 외부 효과를 설명하기 위한 대체 방법

- 구전이 획득비용을 감소시킬 뿐만 아니라 **구매빈도, 구매량** 및 **교차 구매**에도 영향을 미치며 추천으로 인해 획득한 고객의 **LTV**에도 영향을 미친다면, 상황은 더욱 복잡해짐.
 - 고객 i의 구전으로 인해 고객 j의 총공헌이 상승한다면 고객 j에 대한 부가가치를 누구에게 고려해야 하는지 생각해야 함.
- ∴ 고객의 추천 가치를 각 고객에 대한 **별도의 고객 추천 가치**(Customer Referral Value: CRV)로 계산하여 LTV와 분리하는 것이 필요함!

〈표 6-9〉 고객가치 매트릭스

		1년 후 평균 CRV	
		낮음	높음
1년 후 평균 LTV	높음	부유층 (Affluents)	우승자 (Champions)
	낮음	구두쇠 (Misers)	옹호자 (Advocates)

출처: Kumar, Petersen, and Leone(2007).

생애가치 측정기준 (9)

• 평가

- LTV(혹은 CLV)는 **장기적인 의사결정**을 가능하게 하는 **미래 지향적**인 척도
- 이는 산업의 특정한 기업 환경에 맞게 조정하는 **유연한 방법**임.
- 개별 고객은 기업에 대한 예상 공헌도로 평가되지만, 모든 예측은 **불확실성의 대상**이 된다는 사실을 명심해야 함.
- LTV의 정확한 이해와 정확한 행동을 수행하기 위해서는 **가정**(특히 유지율과 총공헌에 대한 추정)이 중요.
- LTV는 다양한 상황에서 유용하게 사용될 수 있는데 예를 들어, **로열티 프로그램**(혹은 고객만족에 대한 투자)을 평가할 때에도 사용됨.

고객자산 (1)

• 정의

- LTV의 정의를 기반으로 LTV 측정값을 **고객**에게서 집계할 수 있는데, 그 결과로 나온 값이 **고객자산**(Customer Equity: CE)임.
- 이 측정기준은 기업의 고객관리에 대한 노력의 결과로 **특정 시점에 기업이 얼마나 가치가 있는지**를 나타냄.

$$\text{고객자산} = \sum_{i=1}^I LTV_i \quad \langle \text{식 6.10} \rangle$$

i = 고객

I = 기업의 모든 고객(특정한 세분집단이나 코호트)

LTV_i = 고객 i 의 생애가치

- 고객자산은 순현재가치에서 **개별 고객의 생애가치의 합**으로 나타냄.
- 고객자산의 측정값은 고객 전체 코호트 혹은 세분화 집단의 **경제적 가치**를 제공함.
- 고객자산이 LTV를 기반으로 하기 때문에 이는 개별 고객 수준에서 수익과 비용을 할당해야 함. **고객의 평균 LTV**를 계산하고 평균 LTV의 **합계**를 구함으로써 이러한 제약 조건을 완화시킬 수 있음.

고객자산 (2)

• 고객자산 점유율

- 고객의 **생애가치**를 고려한 시장점유율의 대체적인 척도는 **고객자산 점유율** (Customer Equity Share: CES)임.

$$CES_j = \frac{CE_j}{\sum_{k=1}^K CE_k} \quad \langle \text{식 6.11} \rangle$$

j = 중심 기업

K = 기업이 제공하는 모든 브랜드

CE_j = 브랜드 j 의 고객자산

• 평가

- 고객자산은 **기업에 대한 고객의 가치**를 나타내므로 기업의 **주주 가치**에 대한 연결로 볼 수 있음.
- 고객자산에 영향을 미치는 중요한 요소는 **LTV**(유지율 및 충공헌도)의 핵심 요소 외에도 **수익성 있는 고객**과 **수익성 없는 고객**의 비율임.
고객자산을 증가시키기 위해서는 수익성이 높은 고객의 수를 늘리는 동시에 수익성이 낮은 고객의 수를 줄이는 것이 중요함.

고객자산 (3)

• 고객자산 계산 예제

〈표 6-10〉 고객자산 계산 예제

획득 연도	고객당 판매액	제조사 이윤	제조사 총공헌	마케팅 및 서비스 비용	실제 유지율	생존율	예상 되는 활동 고객 수	제조사에게 돌아오는 기간당 고객 당 수익	제조사에게 돌아오는 기간당 고객당 감가된 수익	제조사에게 돌아오는 기간당 감가된 총수익
0	120	0.3	36	20	0.4	0.4	400	16	16	6,400
1	120	0.3	36	20	0.63	0.25	250	16	14	3,500
2	120	0.3	36	20	0.74	0.187	187	16	12	2,244
3	120	0.3	36	20	0.82	0.153	153	16	11	1,683
4	120	0.3	36	20	0.85	0.131	131	16	9	1,179
총 고객자산										15,006



02

일반적인 고객 선정 전략

일반적인 고객 선정 전략

• 고객 선정이 필요한 이유

- 프로모션을 보내기 위함일 수도 있고 혹은 특별한 이벤트에 초대하기 위함일 수도 있음.
- **올바른 타겟팅**은 기업으로 하여금 자원을 분별력 있게 사용할 수 있게 하며 해당 고객이 관련 메시지를 받을 수 있도록 함.
- **CRM을 성공적으로 구현**하기 위해서는 기업과 고객의 이익을 극대화하기 위한 방법론이 사용되어야 함.

• Tesco : 분석 기술, 현명한 판단, 신중한 타겟팅을 결합한 사례

영국 슈퍼마켓 체인인 **테스코(Tesco)**는 테스코 클럽카드(Tesco Clubcard)에서 매우 성공적인 **로열티 프로그램**을 운영하고 있다. 또한 테스코는 다양한 **분석 및 타겟팅 기능**을 구축하여 다양한 **고객 세분집단에 맞춤형 프로모션**을 설정할 수 있게 되었다. 로열티 프로그램을 사용하여, 테스코는 각 카드 소지자에 대한 장바구니 내역을 작성한다. 그런 다음 장바구니 내용물을 분석하여 현재 **장바구니를 기반으로 프로모션을 통한 특정 상품을 제안**할 수 있다. 그러나 항상 이 제안이 실행되는 것은 아니다. 테스코는 고기를 구입하지 않는 고객에게 고기를 사도록 유혹하는 쿠폰을 보낼 수 있지만 고객이 채식주의자일 경우를 염두하여 쿠폰을 보내지 않기도 한다.

프로파일링 (1)

- 프로파일링 접근법

- 기업은 현재 가장 수익성이 높은 고객과 **유사한 프로파일을 가진 고객을 타겟팅** 해야 함.
- 의도된 **목적**(예: 고객획득 혹은 기존 고객을 위한 다이렉트 메일 프로모션)에 따라 “가장 수익성이 높은”은 다른 의미를 가질 수 있음(예: 다이렉트 메일 프로모션에 반응 할 가능성이 가장 높은 고객)
- 프로파일링을 위해 최근성, 구매빈도, 구매금액 대신에 사용 가능한 **인구통계학적 정보**를 사용하게 되며, 수익성 있는 고객을 잘 나타내는 변수를 식별하기 위해 **분류나무** 또는 **회귀 모형**을 사용할 수 있음.

- 단점

- **기존 고객과 유사한 대상**에 대해서만 고객을 고려한다는 것.
즉, 기존 고객과 유사하지 않는 수익성 있는 고객집단을 놓칠 수도 있음.

프로파일링 (2)

• 예시

고가치(high-value) 가능성을 가진 가망고객을 확보하는 것이 목적이기 때문에, **은행**은 기존 고객(중요한 프로파일 설정의 기준)과 가망고객(고객의 프로파일을 기준으로 스코어링) 모두의 **공통적인 고객 특성**에 의존해야 한다.

기존 고객의 응답 변수가 **총공헌도**(필드 A)라고 하자.

기업은 고객을 총공헌도별로 분류하고 **상위 20%**를 선정하여 프로파일링 하기로 결정했다. 거래 정보(필드B)는 가망고객에게는 유효하지 않다. 이것은 왜 은행이 기존 고객과 가망고객 모두에게 유용한 정보에 의존해야 하는가에 대한 이유이다.

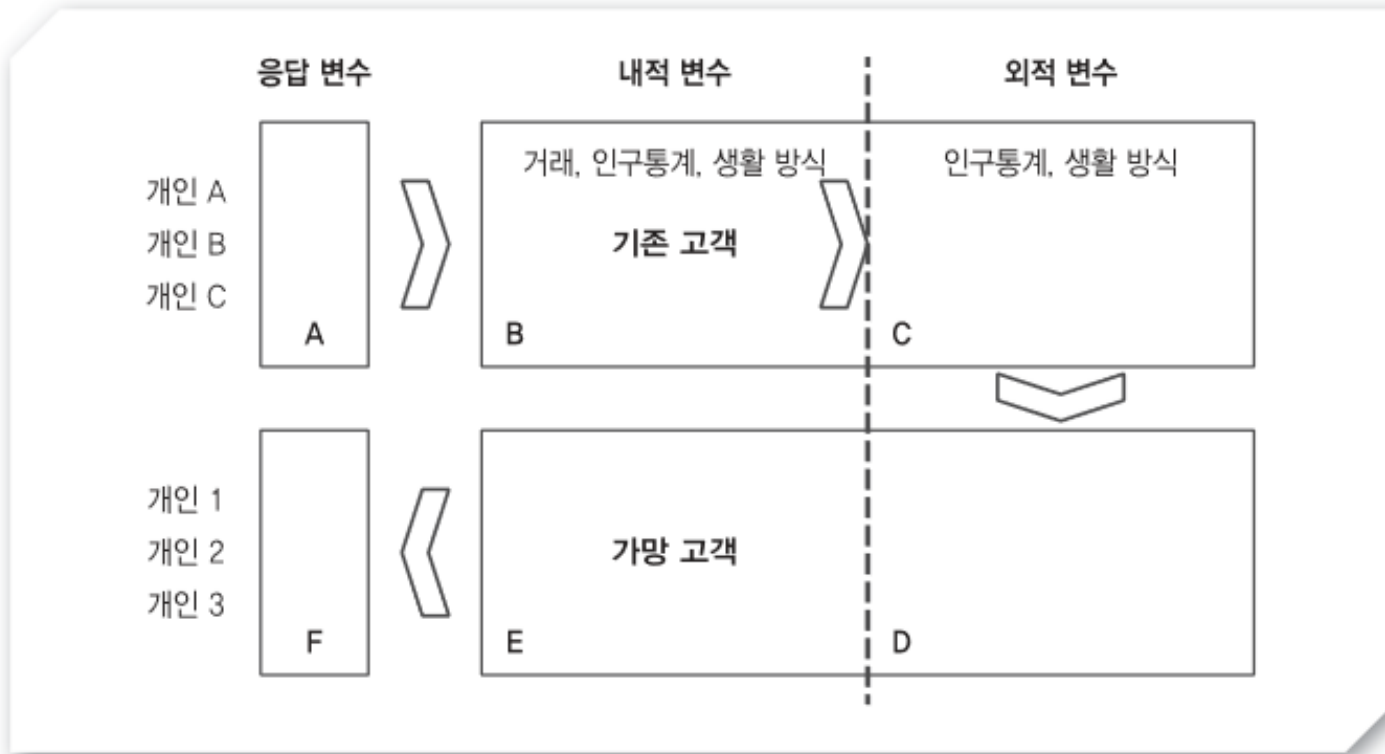
정보 유형은 지역의 사회경제적 지위, 평균 연령, 주택 유형 등과 같은 **지리적 인구통계 데이터**이다. 모형은 지리적 인구통계학 변수를 독립 변수로 사용하고 총 공헌도를 종속 변수로 하여 예측한다.

이 프로세스의 근거는 고가치 **고객을 가장 잘 특징짓는 프로파일**을 찾아 향후 가망 고객의 정보에 적용하는 것이다. 마지막으로 높은 기대 효과를 가진 가망고객은 획득을 위한 캠페인의 대상이 된다.

프로파일링 (3)

• 예시

〈그림 6-8〉 신규고객획득을 위해 프로파일링 사용



이진 분류나무 (1)

• 분류나무 알고리즘

• 정의

- 분류(혹은 결정) 나무를 사용하는 것은 **0또는 1**로 나타나는 종속 변수의 가장 좋은 예측 변수를 찾아내는 데 사용하는 방법론 중 하나임.
- 분류나무 알고리즘은 데이터를 **반복 검색**하여 **어떤 예측 변수**가 이진으로 나타나는 목표 변수의 두 범주를 가장 잘 **구분**하는지 확인할 수 있음.

ex) Y를 이진 결과 변수, 즉 설명변수 X_1, \dots, X_p 에 의해 설명되는 $Y \in \{0, 1\}$ 이라고 가정함.

- 1) 설명 변수 X_i 가 결과 Y를 가장 잘 설명하는지 알아보려면 **오분류**(정확하게 예측되지 않은 결과의 수)의 수를 계산
- 2) 고객층을 나누기 위해 가장 낮은 오분류율을 가진 변수 X_i 를 사용
- 3) 오분류율이 허용 가능한 임계점 이하로 떨어지거나 모든 예측 변수가 모형에 적용될 때까지 이 프로세스는 하위 세분집단에 반복됨

이진 분류나무 (2)

• 예제

〈표 6-11〉 잠재적인 하키 용품 구매자의 분류

	남성		여성		합계	
	하키 구매	하키 비구매	하키 구매	하키 비구매	하키 구매	하키 비구매
스쿠버 구매	60	1,140	50	1,550	110	2,690
스쿠버 비구매	1,540	2,860	80	1,320	1,620	4,180
합계	1,600	4,000	130	2,870	1,730	6,870

위 표에서는 간단하게 성별과 고객이 이전에 **스쿠버 용품을 구매했는지의 여부**에 대한 두 가지의 예측 변수만 주어졌다.

총 **8,600명**의 고객 중 **1,730명**은 하키 용품을 구매했고, **3,000명**의 여성과 **5,600명**의 남성으로 구성되어 있다.

또한 8,600명의 고객 중 **2,800명**은 과거에 스쿠버 용품을 구매했고 **5,800명**은 스쿠버 용품을 구매하지 않았다는 것을 알 수 있다.

이진 분류나무 (3)

- 예제

- 1단계

어떻게 고객을 분류할 것인지에 대한 최적의 접근법을 결정하기 위해
예측 변수(성별과 스쿠버 용품)에 대한 **오분류**의 수를 계산함.

하키 판매의 예측 변수로 **성별**을 사용한다면 모든 남성은 하키 용품을 구매한다고 가정할 수 있다. 반면에 여성 구매자는 없다.

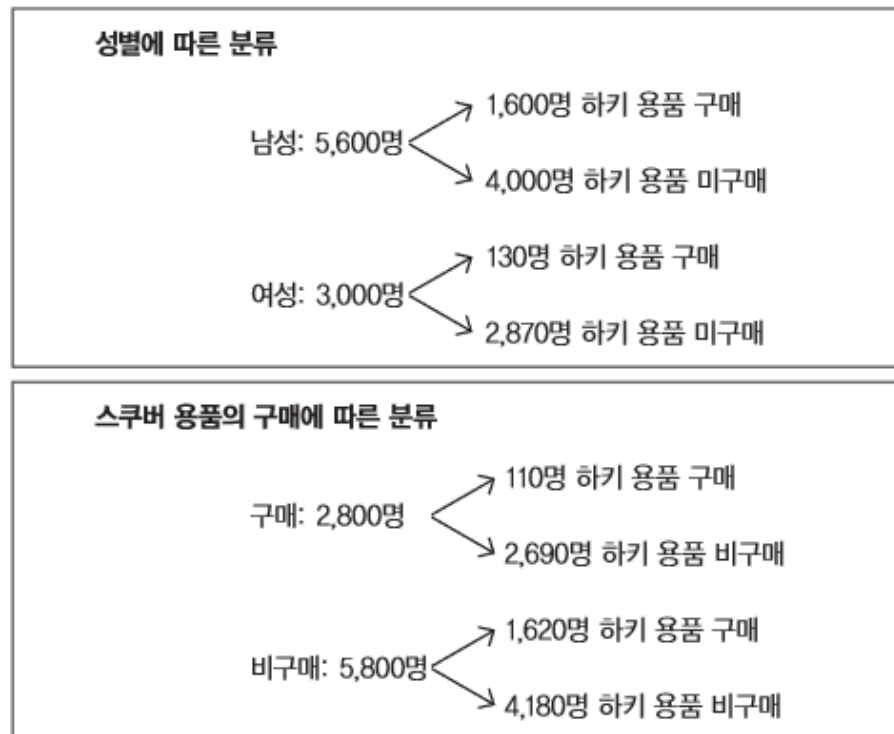
예측 변수인 **성별**에 대해 오분류율 $(4,000+130)/8,600=0.48$ 을 구할 수 있다.

스쿠버 용품의 **구매 여부**를 예측 변수(스쿠버 용품을 구입한 모든 사람이 하키 용품을 구입할 것이라고 가정)로 사용하면 $(2,690+1,620)/8,600=0.50$ 을 얻을 수 있다.

이진 분류나무 (4)

- 예제
 - 1단계

〈그림 6-9〉 잠재적인 하키 용품 구매자의 가능성 분류



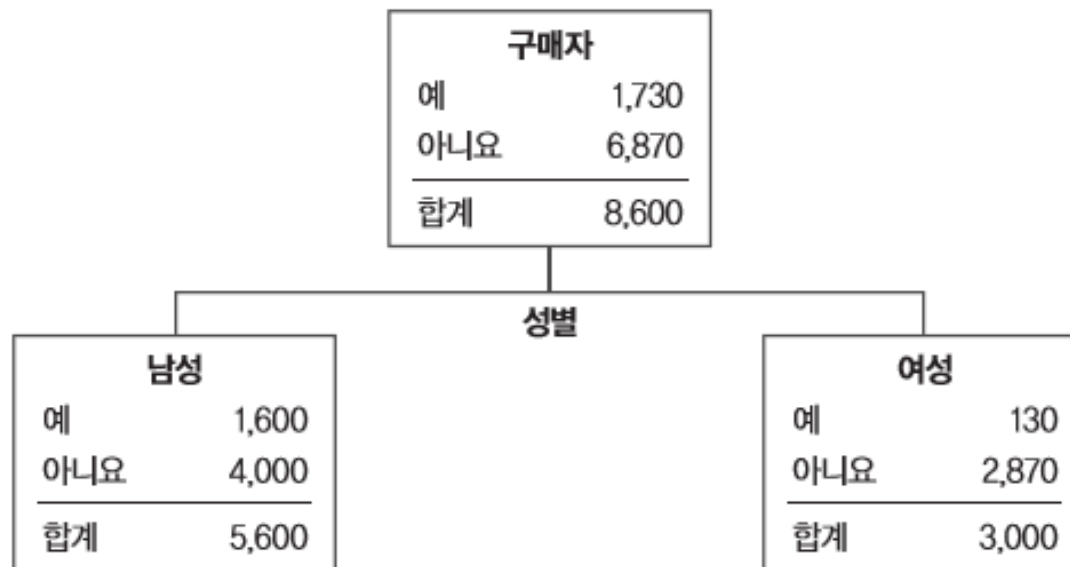
이진 분류나무 (5)

- 예제

- 2단계

성별로 고객을 분류하면 분류나무를 얻을 수 있다.

〈그림 6-10〉 성별에 따른 하키 구매자의 분류



이진 분류나무 (6)

- 예제

- 3단계

스쿠버 지표 이외의 추가 예측 변수가 있다면, 각 하위 집단에서 분리에 대한 최적의 예측 변수를 확인하고 1단계부터 다시 시작할 수 있다.

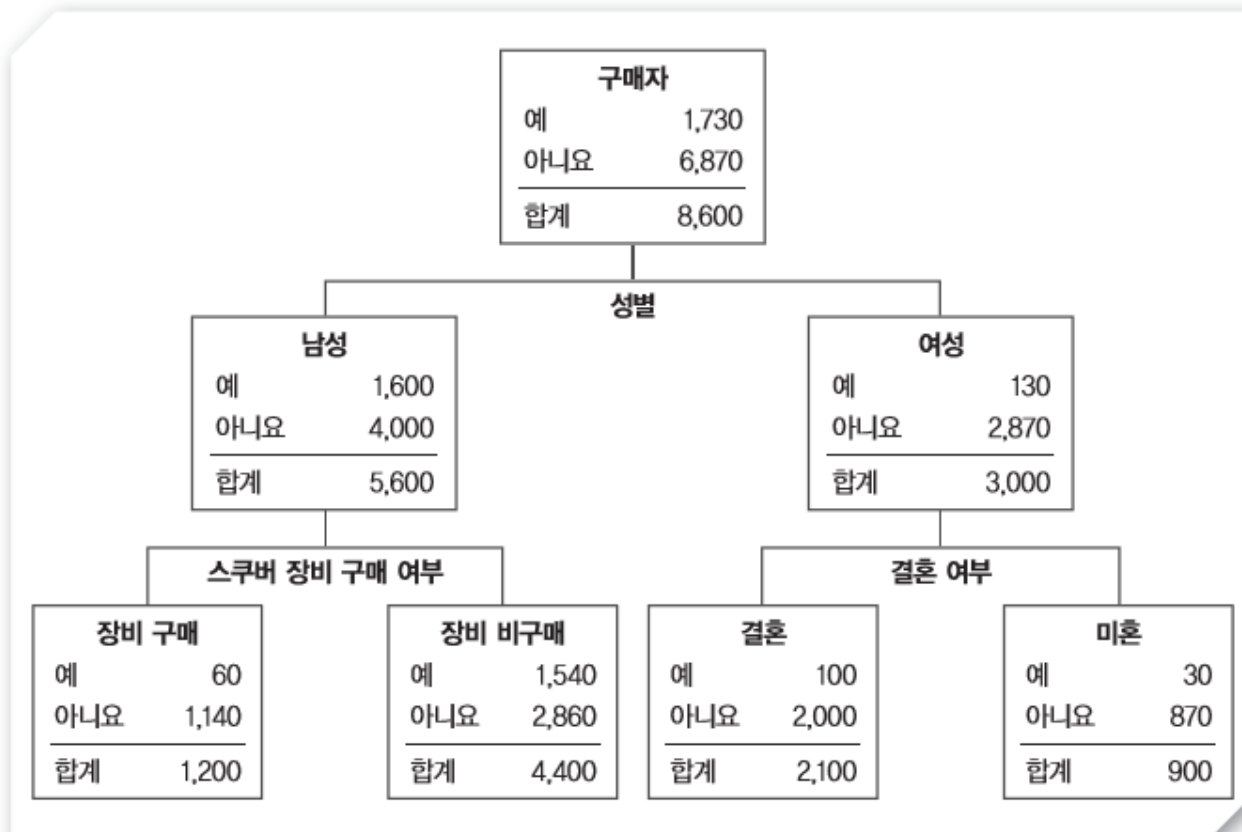
예를 들면, **여성 고객의 결혼 여부**와 **남성 고객의 스쿠버 용품의 구매 여부**에 따라 분류하는 것이 가장 좋다는 것을 알 수 있다.

이 프로세스가 완료되면 분류나무를 통해 고객은 세분화되고, 수익성 있는 세분집단을 식별하여 **타겟 대상**으로 선정할 수 있다.

이진 분류나무 (7)

- 예제
 - 3단계

〈그림 6-11〉 하키 용품 구매자를 위한 분류나무



이진 분류나무 (8)

- 평가

- 의사결정나무 접근법의 한 가지 문제는 세그먼트가 매우 작은 세그먼트(학습용 모형을 구축하는 데 사용된 데이터 셋을 기반)에 맞춰지는 **과적합화**(overfitting) 문제가 있다는 것. 이러한 문제를 해결하기 위해 **데이터 셋을 분리**하여 모형을 구축함.
- 만약 결과가 예상했던 것과 큰 불일치를 보인다면, 모형은 **재평가**될 필요가 있음. 결과가 모형에서 예상되는 범위 내에 있다면 모형의 **적합도가 높다**는 것을 뜻함.

로지스틱 회귀 분석 (1)

- 선형회귀

- 선형 회귀는 **종속 변수**와 **독립 변수**의 지정으로 시작됨.

- **종속변수** : 토요일에 상점에 입장하는 사람의 수
- **독립변수** : 광고를 위해 금요일에 상점에서 지출한 금액
- **결과** : 상점 트래픽에 대한 광고의 효과를 나타내는 계수의 추정치

- 로지스틱 회귀모형

- 정의

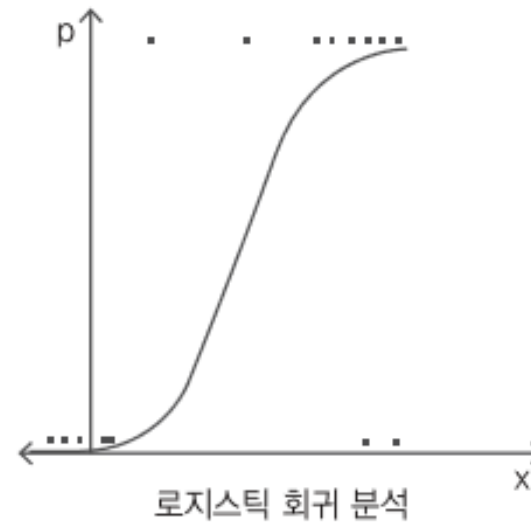
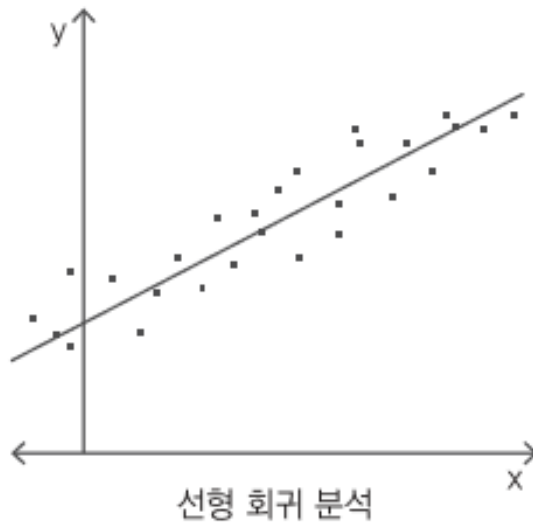
- 종속 변수가 **이진**이고 오직 두 개의 이산적인 값만 있는 **회귀모형**

- 예시

- 고객이 마케팅 캠페인에 반응할 것인가 아닌가
 - 고객이 차를 살 것인가 아닌가

로지스틱 회귀 분석 (2)

〈그림 6-12〉 선형 회귀와 로지스틱 회귀 분석의 비교



로지스틱 회귀 분석 (3)

- 예시

신용카드의 업그레이드를 고려해 보자.

로지스틱 회귀 분석은 은행의 기존 고객에 대해 마케팅 신용카드를 제공할 것인지 잠재적인 타깃을 식별하는 데 사용할 수 있다.

- **목표** : 골드카드를 제안 받은 고객 샘플에 대해 로지스틱 회귀를 추정하는 것
- **종속 변수**: 고객이 골드카드 제안에 가입할 것인지 여부
- **예측 변수**: 고객이 사용하는 다른 은행 서비스 및 재무 정보와 인구통계학적 정보
- **결과** : 잠재적인 목표고객을 위한 예측 변수값을 입력하여 로지스틱 모형을 적용하여 골드카드 제안에 가입할 것으로 예상되는 확률을 산출

로지스틱 회귀 분석 (4)

• 예시

선형 회귀의 경우 y 는 임의의 값을 취할 수 있는 반면,
로지스틱 회귀의 y 는 **0 또는 1**의 값을 가진다.
로지스틱 회귀 곡선은 주어진 x 에서 y 의 **예측 확률 p** 를 반영한다.

$$y = \alpha + \beta x + \varepsilon$$

y = 종속 변수

x = 예측 변수

α = 상수(선형 회귀 분석에 의해 추정되며 보통 절편이라 함)

β = y 에 대한 x 의 효과(또한 선형 회귀로 추정)

ε = 오차항

이러한 회귀 분석에서 y 는 **음의 무한대와 양의 무한대 사이의 값을** 가질 수 있다.
이벤트에 대한 발생 확률을 나타내기 위해 실제 관찰된 종속 변수를 0과 1 사이로 제한해야 하는 경우, **변환**이 필요하다.

로지스틱 회귀 분석 (5)

• 변수의 변환 단계

• 1단계

p가 이벤트의 발생 확률을 나타낸다면 비율 $\frac{p}{1-p}$ 를 고려해야 함.

P가 1보다 작은 양수이기 때문에 이 표현식의 범위는 **0부터 무한대**임.

• 2단계

이 비율에 **로그**를 취함.

$$\log\left(\frac{p}{1-p}\right)$$

이 변환은 표현식 값의 범위를 **음의 무한대와 양의 무한대** 사이에 있도록 해야 함.

• 3단계

$$z = \log\left(\frac{p}{1-p}\right)$$

이제 이 값은 종속 변수로 간주될 수 있음.

$z = \alpha + \beta x + \varepsilon$ 형식의 예측 변수와 이 값의 **선형 관계**를 작성할 수 있음.

로지스틱 회귀 분석 (6)

- 변수의 변환 단계
 - 4단계

예측 확률 p 를 얻기 위해 다음과 같은 **역변환**이 필요함.

$$\log\left(\frac{p}{1-p}\right) = z = \alpha + \beta x + \varepsilon = \left(\frac{p}{1-p}\right) = e^z$$

이것은 관심 있는 이벤트의 **발생 확률 p** 를 다음과 같이 계산할 수 있게 함.

$$p = \left(\frac{e^z}{1 + e^z}\right) = \left(\frac{1}{1 + e^{-z}}\right) \quad \langle \text{식 6.12} \rangle$$

로지스틱 회귀 분석 (7)

• 계수의 해석

• 오즈 비율(odds ratio)

- 이벤트가 **발생하지 않을 확률**을 나타냄.
- 만약 로지스틱 회귀 분석 계수가 $\beta=2.303$ 일 때, 오즈 비율 구하는 공식

$$e^{\beta} = e^{2.303} = 10$$

- 독립 변수 x가 하나의 단위를 증가시키면 종속 변수가 1인 확률은 **10배** 증가하고 다른 모든 변수는 일정하게 유지됨.
- 보통, 계수를 제외하고 오즈가 보고됨.

• 예제

B2B 환경에서 **서비스 계약의 연장**을 고려해 보자.

서비스 기업은 고객이 기존 계약을 연장할 가능성을 알고자 한다. 따라서

- 1) 영업 통화 수와 고객이 계약을 연장하는 지 추적한다.
- 2) 기업의 영업 통화 수가 계약 연장의 가능성에 영향을 줄 것이라고 예상한다.
- 3) 이를 기반으로 모형을 추정한다.

로지스틱 회귀 분석 (8)

- 계수의 해석
 - 예제

〈표 6-12〉 계약 연장과 판매를 위한 샘플 데이터

계약 연장	판매 수
1	2
1	4
0	6
0	2
1	4

계약 연장	판매 수
1	8
0	3
0	0
0	2
0	5

계약 연장	판매 수
0	0
1	2
1	8
1	4

로지스틱 회귀 분석 (9)

- 계수의 해석
 - 예제

〈표 6-13〉 로지스틱 회귀 분석 예시의 오즈

	영업 통화 = 0	영업 통화 = 1
오즈 $\{ \exp(\alpha + \beta \times \text{영업 통화}) \}$	0.253	0.375
계약 연장의 확률	0.202	0.273
확률의 차이	0.071	

결과 추정치는 $\alpha = -1.37$ 과 $\beta = 0.39$ 이다.

이 예시에서 **오즈**는 고객이 **계약을 연장하지 않을 것**이라는 상대적인 가능성을 나타낸다. 영업 통화가 없는 경우 오즈는 **0.253**이며 영업 통화가 있을 경우 오즈는 **0.375**이다. 로그 오즈는 $e^{0.39} = 1.477$ 이며 추가 영업 통화마다 계약 연장의 변경 확률은 **1.477배**가 된다.

계약 연장의 확률은 영업 통화가 없을 경우 **0.202**, 영업 통화가 있을 경우 **0.273**이다. 따라서 영업 통화를 할 때와 하지 않을 때의 계약 연장 확률의 차이는 **0.07**이다.

로지스틱 회귀 분석 (9)

- 계수의 해석

- 평가

- 로지스틱 회귀 분석에서는 예측 변수의 한 단위당 증가량이 회귀 곡선에 따라 달라지는데, 이것은 한 단위당 변화가 아주 미미한 경우에는 영향력이 거의 없지만, **중앙**에서는 한 단위의 변화가 상당히 큰 영향을 줌.

ex) 개인이 주택을 소유할 가능성의 차이는 소득이 10,000~30,000달러 혹은 1,000,000~1,020,000달러로 증가함에 따라서는 크게 변하지 않을 수 있지만, 소득이 50,000~70,000달러로 증가할 경우는 주택을 소유할 가능성이 상당히 증가



03

대체 고객 선정전략 평가방법

대체 고객 선정 전략 평가 방법

- 모형의 성능 평가

- 모형의 예측 성과를 비교하기 위해 일반적으로 데이터를 **훈련용 데이터**(데이터의 2/3)와 **검증용 데이터**(데이터의 1/3)로 나눔.
- 모형은 훈련용 데이터를 기반으로 구축되며 이를 통해 검증용 데이터에 대한 예측이 이루어짐.
- 일부 모형은 다른 모형보다 더 높은 예측력을 보이며, 일반적으로 검증용 데이터에 **가장 일반화된 모형**을 선택함.

- 평가방법

- 검증용 데이터에서 얻는 각각의 오분류율을 비교
- 리프트 분석

오분류율

• 정의

- 잘못된 예측의 수를 **총 데이터 수로 나눈 것**

• 예시

〈표 6-14〉 오분류표

		예상		합계
		1	0	
관찰	1	726	56	782
	0	173	504	677
합계		899	560	1,459

- 잘못된 예측 수는 비대각선 항목의 합계($56+173=229$).
따라서 그림에서의 오분류율은 **$229/1,459=15.7\%$** 로 계산될 수 있음.

리프트 분석 (1)

• 정의

- 모형이 적용되지 않은 경우의 예상 결과에 대해 현재 모형이 적용되었을 때 **얼마나 잘 수행하는지**를 보여 줌.
- 이는 **모형이 얼마나 좋은지에 대한 기준 척도**를 제공함.

• 예시

1,000명의 가망고객 중 100명이 구매했다고 하자.

좋은 예측 모형은 선택된 그룹에서 상대적 구매자의 수가 증가하는 것을 도와준다.

100명의 가망고객을 임의로 선정하면 **약 10명**의 구매자가 포함되지만 모형을 기반으로 하면 **30명**의 구매자가 포함될 수 있다. 이것은 리프트 차트를 **시각화** 하는데 도움이 된다.

가장 높은 리프트를 가진 모형이 최종적으로 선정된다.

리프트 분석 (2)

• 리프트 분석의 사용

- 1) 두 개 이상의 대체 모형을 **비교**할 때
- 2) 시간이 경과함에 따른 모형의 **성능을 추적**할 때
- 3) 다른 샘플에서 모형의 **성능을 비교**하기 위해

• 리프트를 계산하기 위해 필요한 정보

- **누적 고객의 수**: 10분위 수까지의 전체 고객 수
- **누적 고객 비율**: 10분위 수까지의 전체 고객 비율
- **누적 구매자의 수**: 10분위 수까지의 총 구매자 수
- **각 10분위 수의 실제 반응률**: 각 10분위 수에 대한 고객의 수로 구매자의 수를 나눈 값
- **각 10분위의 모형에 기초한 예측된 반응률**: 예측된 구매자의 수를 각 10분위 수의 고객 수로 나눈 값

리프트 비율 = (각 10분위의 반응률) ÷ (종합적인 반응률) × 100

누적된 리프트 비율 = (누적된 반응률) ÷ (종합적인 반응률) × 100

누적된 반응률 = 누적된 구매자의 수 ÷ 10분위 당 고객의 수

리프트 분석 (3)

• 리프트 성능 실례

처음에 평가하고자 하는 모형(예: RFM 또는 로지스틱 회귀 분석)을 실행하고 그에 따라 고객층을 분류해야 한다.

분류된 고객 목록을 기반으로, 고객은 10개의 동일한 크기의 그룹으로 분배된다.

〈표 6-15〉 리프트와 누적된 리프트

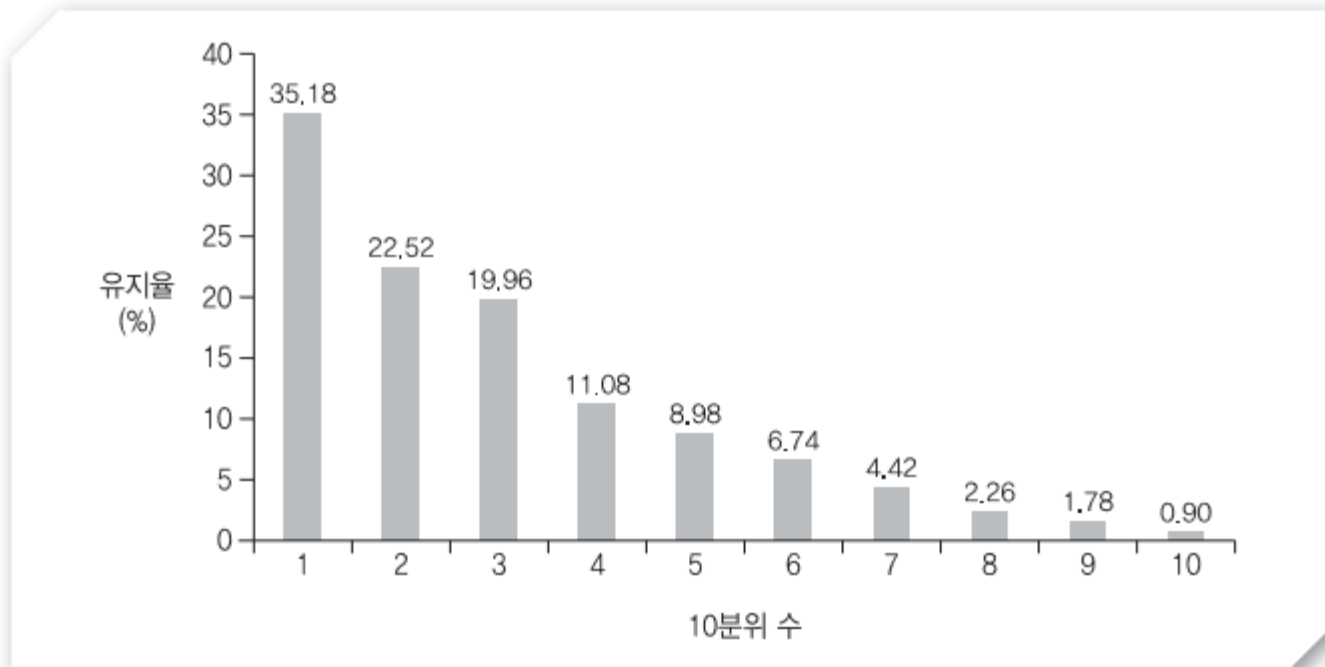
10분위 수	고객 수	구매자 수	반응률	리프트	누적 리프트
1	5,000	1,759	35.18	3.09	3.09
2	5,000	1,126	22.52	1.98	5.07
3	5,000	998	19.96	1.75	6.82
4	5,000	554	11.08	0.97	7.80
5	5,000	449	8.98	0.79	8.59
6	5,000	337	6.74	0.59	9.18
7	5,000	221	4.42	0.39	9.57
8	5,000	113	2.26	0.20	9.76
9	5,000	89	1.78	0.16	9.92
10	5,000	45	0.90	0.08	10.00
합계	50,000	5,691	11.38		

리프트 분석 (4)

- 리프트 성능 실례

성과가 좋은 모형에서 **첫 10분위 수**의 고객은 **가장 높은 반응률**을 보이고
다음 10분위 수로 진행될수록 반응률은 **지속적으로 하락**한다.

〈그림 6-13〉 10분위 수 분석



리프트 분석 (5)

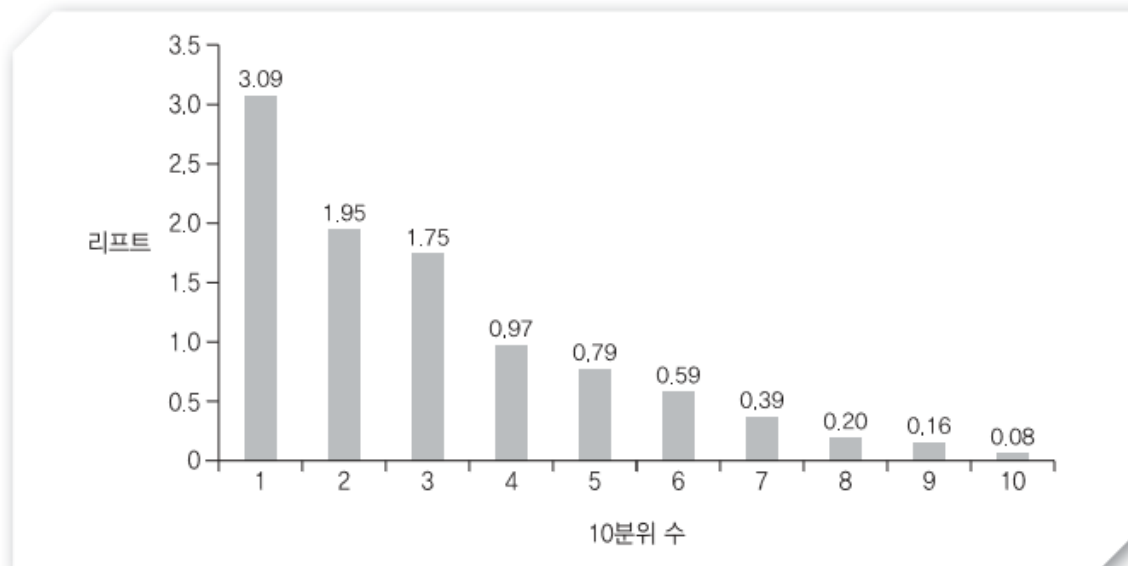
• 리프트 성능 실례

이 경우 상위 10분위 수의 **리프트**는 **3.09**이다.

이는 이러한 고객만을 타겟으로 하여 메일링 했을 때의 구매자 수가 동일 고객 수를 무작위로 메일링하여 구매한 구매자 수의 **3.09배**를 기대할 수 있음을 나타낸다.

그에 반해, 마지막 10분위 수(10 분위 수 10)는 동일한 크기의 무작위 표본에서 기대할 수 있는 구매자 수의 **0.08배**에 불과하다.

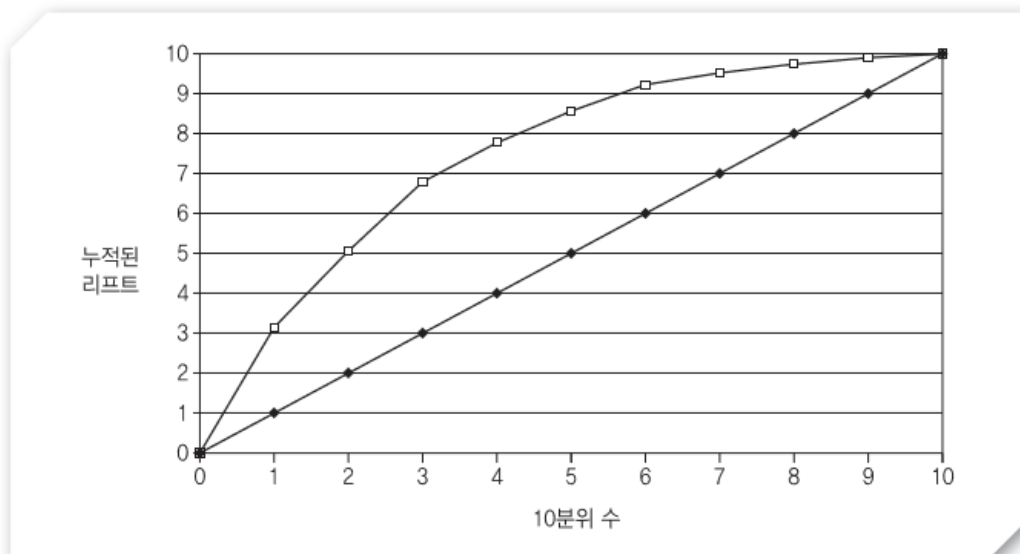
〈그림 6-14〉 리프트 분석



리프트 분석 (6)

• 리프트 성능 실례

〈그림 6-15〉 누적된 리프트 분석



고객의 **상위 30%**에서 고객을 선정하면 전체 응답자의 **68%**를 얻을 수 있다.

모형과 모형 사이의 **거리가 멀수록 모형은 더 강력**해지고, 누적 리프트 곡선의 **기울기**는 **리프트**를 반영한다.

간단히 말해서, 리프트를 사용하여 두 개 이상의 대체 모형을 **비교**하고, 시간의 경과에 따른 모형의 **성능을 추적**하거나, 다른 샘플에서 모형의 **성능을 비교**하는데 사용될 수 있다.

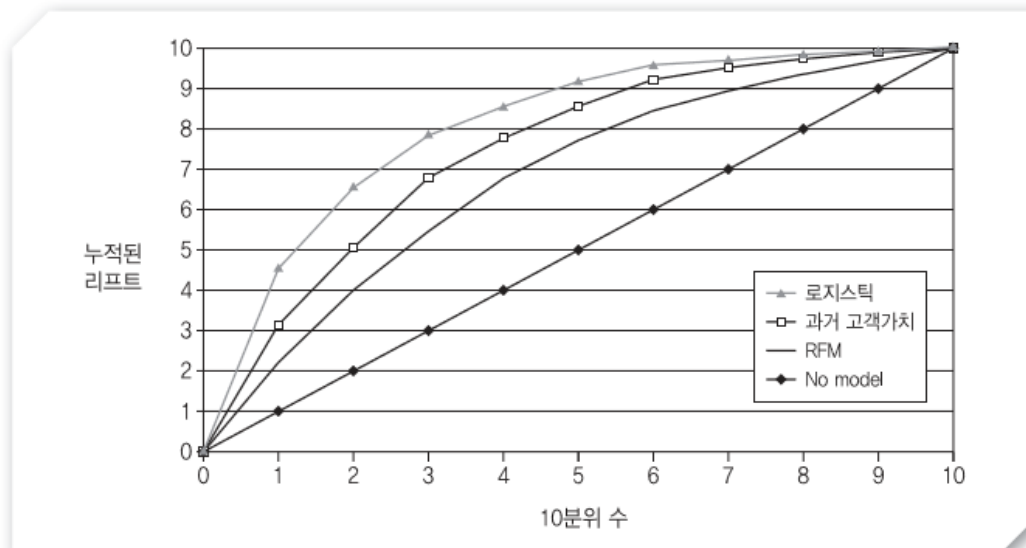
리프트 분석 (7)

• 리프트 성능 실례

과거 경험은 로지스틱 모형이 고객 선정을 위한 도구로 사용될 때 가장 좋은 리프트 성능을 제공하는 경향이 있음을 보여준다.

이것은 리프트 차트의 **가장 높은 곡선**에서 볼 수 있다.

〈그림 6-16〉 리프트 분석을 사용한 모델 비교



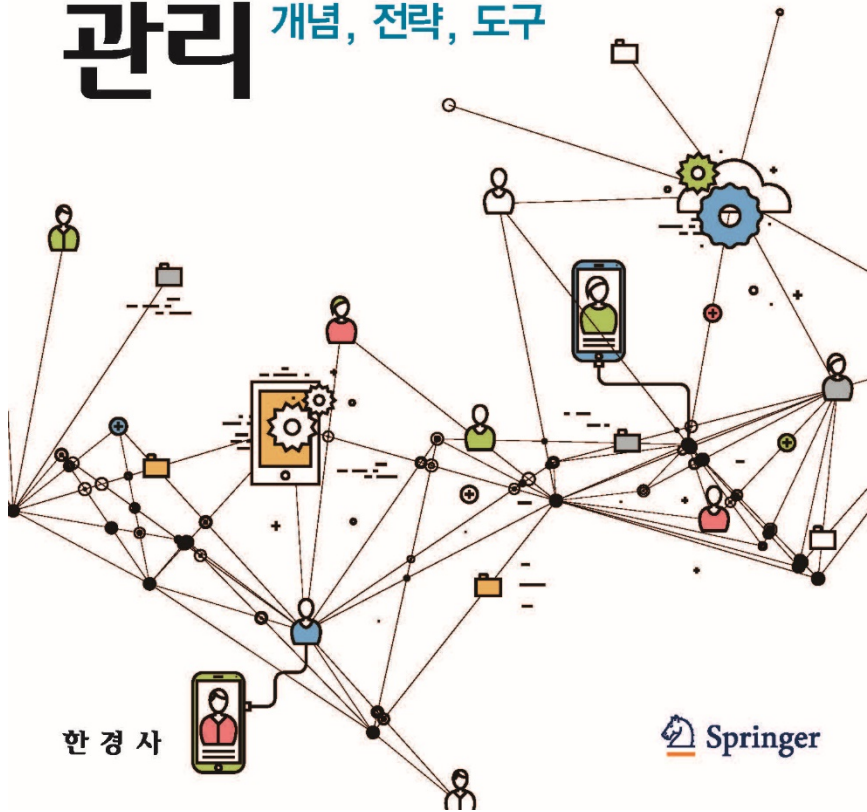
과거 고객가치 접근법은 차선의 성능을 제공하는 반면, **전통적인 RFM 기법**은 가장 낮은 성능을 보여 준다.



Q & A

고객관계 관리

개념, 전략, 도구



한경사

Springer

고객관계관리 개념, 전략, 도구 (제2판)

V. Kumar, W. Reinartz 공저

홍태호, 신태수, 안현철, 김은미 공역

한경사, 2018

본 강의보조자료는 고객관계관리 개념, 전략, 도구(제2판)의 한국어판 서적을 기초로 제작되었으며, 해당 서적의 저작권은 '도서출판 한경사'에 있습니다. 저작권법에 의하여 한국 내에서 보호를 받는 저작물이므로 무단전재와 복제를 금합니다.