

## 1. 들어가며

숙취해소 음료에 있어서 효과의 유무, 맛의 유무는 모두 주관적 요소이다. 사람의 입맛이나 취향, 음주량, 컨디션 등에 따라 다르므로 맛이 있다고 한 사람이 좀 더 많았다고 해서 맛이 있는 제품이라 단정할 수 없고, 효과를 본 사람이 더 많았다고 해서 효과가 있다고 단정할 수 없다.

따라서 제품보다 고객 개개인에 포커스를 맞추어야 한다. 즉 검증해야 할 것은, 숙취해소 등 고객 각기 특정한 목적을 가지고 제품을 구매했을 때, 그 제품에 대해 개인적으로 느끼는 효과의 유/무나 맛의 유/무에 따라 해당고객이 제품의 평가에 얼마나 민감하게 반응하는가이다. 이에 대한 검증을 통해, 예를 들어, “생각한 것보다 효과가 부족하더라도 맛이 있을 경우” 고객이 이탈할 것인지 아닌지 (and vice versa) 등을, i.e., (1)각 제품 자체의 특성, (2)제품 특성에 따른 고객의 잠재적 충성도/이탈가능성과 (3)각 제품을 구매할 때 고객들이 중점적으로 고려하는 특성 등을 수치화하고, (제품간) 직접적으로 비교하며, 예측할 수 있다. 아래의 검정은 이러한 전제 하에 진행된다.

## 2. 연구 개요

이커머스에서 판매하는 기능성 음료에 대한 평가에 있어 고객후기란의 rating(1~5)은 제품의 효과, 맛, 가격 등 제품 자체에 대한 주관적 평가와 배송, 포장에 대한 만족도 등 externalities로 구성된 복합적인 고객반응 지표이다.

예를 들어 “숙취해소 효과는 좋았지만 맛이 역했다”며 3점의 rating을 남긴 후기가 있다면, 이는 효과(+)와 맛(-)의 (이 고객만의 주관적인)플러스-마이너스 상쇄 결과물이 3인 것으로 받아들일 수 있다.

제대로 활용할 수 있다면(i.e., 고객후기에 나타나는 복합적인 고객반응 자연어를 dissect하여 개별적으로 변수화한다면), 이 지표는 다음과 같은 몇 가지 측면에서 여타 지표에 비해 유용하며 정확하다.

첫째, 제품의 효과나 맛에 대해 알아보고자 제품무료제공/설문참가료 제공 등 인센티브를 제공한 설문조사자료에 비해 이커머스의 고객후기 rating은 고객이 (1)자발적으로, (2)특정한 목적(숙취해소 등)을 위해 각자 제품을 구매한 후 제품에 대한 평가를 토대로 rating을 남긴다는 점에서 여타 외부효과가 차단된 가장 순수한 고객반응 데이터이다. 둘째, 설문자료 등과 달리 비용이 발생하지 않는다. 셋째, standardized beta coefficient를 활용한 비교를 통해 각 제품에 대해 직접적으로 “수치화”된 비교가 가능하다.

따라서 다음과 같은 몇 가지 조건들이 충족된다면 rating을 활용한 모형은 유용하며 정확한 결과를 만들어낼 수 있다.

1. 데이터의 양이 정규성을 가정할 수 있을 만큼 충분히 많을 것
  - Ex) 주관적 지표이므로 어떤 고객은 “숙취해소 효과는 좋았지만 맛이 역했다”며 2점을 줄 수도 있고, 또 다른 고객은 마찬가지로 이유로 3점을 줄 수도 있다.
2. 제품을 제공받아 작성한 대가성 리뷰(협찬 등)를 필터링할 수 있을 것
3. Rating에 영향을 미치는 유의미한 독립변수(효과, 맛, 가격, 배송, 포장 등)만을 선별해낼 수 있을 것
4. 자연어처리를 통해 각 후기에서 선별된 독립변수들을 각각 개별추출해내고 수치화할 수 있을 것
5. 적용 모델에 따른 Robustness test를 통해 필요조건의 충족을 검정하고 추정된 모수의 타당성을 확보할 수 있을 것

조건을 충족하는 모델링 절차는 다음과 같다.

1. 전체 고객후기 데이터를 대상으로 빈도분석/감성분석 실시 후 EDA를 토대로 potentially effective variables 목록을 추출한다.
2. 광고/협찬 관련 후기를 제거한다 (광고 필터링 알고리즘 적용).
3. Feature selection을 통해 potentially effective variables로부터 설명력이 높은 변수들만을 선별한다.
4. 최종적으로 선별된 변수들에 대한, 그리고 종속변수-독립변수간 관계에 대한 EDA + 분석 목적에 맞는 모델 검토를 통해 적용할 모델을 선별한다.
  - A. EDA 및 모델 검토 결과 (1)OLS추정과 (2)ordered regression의 두 모델로 좁혔으며 최종적으로 OLS추정 모형을 사용하기로 결정함.
  - B. 근거1: 이번 분석의 목적은 맛의 유/무, 효과의 유/무 등에 대한 고객반응에 있으므로 비선형적 관계에 대한 고려는 배제한다. 따라서 (1)과 (2) 중 하나를 택한다.
  - C. 근거2: 다음으로 Rating을 ordinal variable과 interval variable 중 어느 것으로 취급할 것인가에 따라 최종 선택이 갈리는데, 이 경우에는 각각 장단점이 있다. 정석적인 방식대로 한다면 rating을 ordinal variable로 취급해야 할 것이나, 이 경우 해석력과 직관성에서 떨어지는 측면이 있다. 일반적으로 ordinal DV의 구분이 3~4 이상인 경우 데이터의 양이 충분하고 모형 결과에 대한 잔차의 정규성 가정이 충족된다면 interval variable 취급을 통한 ols추정치에 왜곡이 없을 것으로 기대할 수 있다.
5. 모델링
6. Robustness test를 통한 추정치 타당성 검정
7. 결과해석

### 3. 변수 선택

EDA 및 feature selection을 통해 최종적으로 효과(+, -), 맛(+, -), 가격(+, -), 배송/포장(+, -)을 설명변수로 채택

### 4. 분석 결과

#### 4.1. 컨디션 헛개

#### ESTIMATION RESULT

		<i>Beta Coefficients</i>	<i>Partial-R<sup>2</sup></i>	<i>R<sup>2</sup> (adj)</i>	<i>R<sup>2</sup> (adj)</i>
<b>Target Effect (+)</b>	효과	0.91*** (0.09)			0.716
	맛	0.33*** (0.11)			
	가격	0.32*** (0.10)			
<b>Target Effect (-)</b>	효과	-1.55*** (0.32)			
	맛	-1.36*** (0.38)			
	가격	0.00 (0.38)			
<b>Control</b>	배송/포장(+)	-0.04 (0.16)			
	배송/포장(-)	-1.96*** (0.11)			

(1) \*\*\* p<0.01, \*\*p<0.05, (2) SE in parentheses

#### ANOVA

F-value	Critical F
124.26	0.000

위 결과치 및 implications에 대해 말로 풀어서 설명하면:

만들어진 모형은 매우 유용하며 모형의 설명력은 매우 뛰어나다. 고객들은 효과, 맛, 가격만족도 순으로 제품에 대해 민감하게 반응하며, 특히 효과와 맛에 대해 매우 엄격하게 반응한다. 예를 들어, 생각한 것만큼 숙취해소 효과가 좋지 않으면 곧바로 이탈한다. 맛보다는 숙취해소 효과의 영향력이 훨씬 더 크며, 효과가 좋을 때 제품에 대한 충성도가 올라가는 영향보다 효과가 나쁠 때 이탈이 확고해지는 영향이 더 크다. 또한 맛이 좋을 때 제품에 대한 충성도가 올라가는 영향보다 맛이 맞지 않을 때 이탈하는 영향이 매우 크다. (또한 효과가 좋더라도 맛에 불만족할 경우, 또는 맛이 좋더라도 효과에 불만족할 경우 평균적으로 불만족하며, 이탈할 가능성이 높다). 쿠팡 가격을 기준으로, 가격이 상대적으로 낮다고 생각하여 만족하는 사람들은 많은 반면, 높다고 생각하는 경우 그것을 이유로 제품에 대해 불만족하는 경우는 거의 없다.

따라서 사람에 따라 효과나 맛 두가지 측면 모두에서 호불호가 강할 확률이 높다. 그러나 전체 rating 지수가 높은 편이라는 점을 동시 고려하면 효과와 맛의 두 가지 측면에서 모두 만족하는 고객의 비율이 매우 높다 할 수 있으며, 숙취해소음료의 기본인 효과와 맛에 모두 충실한 밸런스가 잘 잡힌 제품이라 할 수 있다.

배송이나 포장에 좋을 것은 디폴트로 당연하게 생각하는 경향이 있으며(i.e., 포장/배송이 좋다/빠르다고 해서 제품 만족도/충성도가 올라가는 것은 아니다), 반면 배송/포장이 부실한 경우 매우 민감하게, 부정적으로 반응하므로 제품의 포장/배송에 더 신경을 써야 할 필요가 있다.

## 4.2.레디큐

### ESTIMATION RESULT

		Beta Coefficients	Partial-R <sup>2</sup>	R <sup>2</sup> (adj)	R <sup>2</sup> (adj)
Target Effect (+)	효과	0.47*** (0.16)			0.525
	맛	1.29*** (0.25)			
	가격	0.32 (0.39)			
Target Effect (-)	효과	-0.86*** (0.17)			
	맛	-0.87*** (0.35)			
	가격	0.25 (0.60)			
Control	배송/포장(+)	-0.24 (0.50)			
	배송/포장(-)	-1.05*** (0.26)			

(1) \*\*\* p<0.01, \*\*p<0.05, (2) SE in parentheses

### ANOVA

F-value	Critical F
24.997	0.000

위 결과치 및 implications에 대해 말로 풀어서 설명하면:

만들어진 모형은 매우 유용하며 모형의 설명력은 뛰어나다. 고객들은 맛, 효과 순으로 제품에 대해 민감하게 반응하며 현재 형성된 가격풀은 제품 만족도에 영향을 미치지 않는다.

특히 맛에 대해 매우 긍정적으로 반응한다. 다른 모든 제품군에 비해 맛이 차지하는 비중이 가장 높으며, 평균적으로 효과가 없더라도 맛이 있어 제품에 대해 만족할 (또는 구매할) 확률이 높다. 반면 효과가 좋더라도 생각한 맛과 다르거나 맛에 대해 불만족할 경우 평균적으로 제품에 대해 불만족하고 이탈할 가능성이 높다.

고객들이 체감하는 숙취해소 효과는 다른 제품군에 비해 상대적으로 미미한 반면 맛에 대한 평가는 매우 뛰어나, 맛에 올인한 제품이라 할 수 있다.

컨디션의 경우와 마찬가지로 배송이나 포장의 좋은 것은 디폴트로 당연하게 생각하는 경향이 있

으며(i.e., 포장/배송이 좋다/빠르다고 해서 제품 만족도/충성도가 올라가는 것은 아니다), 반면 배송/포장이 부실한 경우 매우 민감하게, 부정적으로 반응하므로 제품의 포장/배송에 더 신경을 써야 할 필요가 있다.

### 4.3.여명808

#### ESTIMATION RESULT

		<i>Beta Coefficients</i>	<i>Partial-R<sup>2</sup></i>	<i>R<sup>2</sup>(adj)</i>	<i>R<sup>2</sup>(adj)</i>
<b>Target Effect (+)</b>	효과	0.98*** (0.11)			0.617
	맛	0.36** (0.24)			
	가격	0.48*** (0.16)			
<b>Target Effect (-)</b>	효과	-1.37*** (0.28)			
	맛	-1.10*** (0.25)			
	가격	0.04 (0.20)			
<b>Control</b>	배송/포장(+)	0.02 (0.36)			
	배송/포장(-)	-1.40*** (0.15)			

(1) \*\*\* p<0.01, \*\*p<0.05, (2) SE in parentheses

#### ANOVA

F-value	Critical F
53.325	0.000

위 결과치 및 implications에 대해 말로 풀어서 설명하면:

만들어진 모형은 매우 유용하며 모형의 설명력은 매우 뛰어나다. 여명808의 경우, 만족하는 경우와 불만족하는 경우에 대하여 민감하게 반응하는 요소가 다르다. 만족하는 경우에는 무엇보다 효과의 영향이 가장 크고, 쿠팡기준 상대적 가격에 대한 만족도가 높은 반면 맛의 영향력은 낮다. 불만족하는 경우에는 효과에 불만족하는 경우 가장 민감하게 반응하며, 맛에 불만족하는 경우 역시 매우 민감하게 반응하고, 가격의 영향은 없다.

효과에 관한 한 다른 모든 제품군과 비교할 때 고객반응 민감성이 두번째로 가장 높다. 이를 전

체 rating 지수와 동시 고려할 경우 효과 측면에서 뛰어난 제품이라 할 수 있다. 반면 효과에 대해 기대하고 구매하였는데도 생각한 만큼의 효과가 없을 경우 이탈할 가능성 역시 매우 높다.

또한 흥미로운 점은 여명808에 대해 만족하는 경우 맛으로 인해 만족하는 경우는 적다는 점으로, 순수하게 숙취해소 효과를 위해 이 제품을 구매한다는 점을 시사한다. 동시에 효과가 좋더라도 맛에 불만족하는 경우 부정적 영향력이 더 높다 (즉, 평균적으로 효과가 좋더라도 맛이 역하게 느껴진다면 이탈할 가능성이 높다). 이러한 점은 숙취해소음료도 결국 음료라는 점, 즉 숙취해소를 위해 구매한 음료라 하더라도 고객 만족도에 있어 맛이 차지하는 비중이 생각보다 높다는 점을 암시한다. (그도 그런게 보통 술마시면 안그래도 역한데 역한거 마시면 더 올라오지 않나?)

배송이나 포장에 좋을 것은 디폴트로 당연하게 생각하는 경향이 있으며(i.e., 포장/배송이 좋다/빠르다고 해서 제품 만족도/충성도가 올라가는 것은 아니다), 반면 배송/포장이 부실한 경우 매우 민감하게, 부정적으로 반응하므로 제품의 포장/배송에 더 신경을 써야 할 필요가 있다.

#### 4.4.개수강

##### RESULT

		<i>Beta Coefficients</i>	<i>Partial-R<sup>2</sup></i>	<i>R<sup>2</sup> (adj)</i>	<i>R<sup>2</sup> (adj)</i>
<b>Target Effect (+)</b>	효과	0.42** (0.19)			0.506
	맛	0.95*** (0.29)			
	가격	0.65* (0.36)			
<b>Target Effect (-)</b>	효과	-0.62*** (0.20)			
	맛	-0.91** (0.38)			
	가격	-0.20 (0.86)			
<b>Control</b>	배송/포장(+)	-0.25 (0.90)			
	배송/포장(-)	-1.54*** (0.31)			

(1) \*\*\* p<0.01, \*\*p<0.05, \*p<0.10, (2) SE in parentheses

##### ANOVA

F-value	Critical F
16.483	0.000

위 결과치 및 implications에 대해 말로 풀어서 설명하면:

만들어진 모형은 매우 유용하며 모형의 설명력은 뛰어나다.(유의수준의 차등을 고려할 경우) 고객들은 맛, 효과, 가격 순으로 제품에 대해 민감하게 긍정반응하며 효과, 맛 순으로 제품에 대해 민감하게 부정반응하고, 가격이 높다고 생각하더라도 가격이 제품 만족도에 부정적 영향을 미치지 않는다.

효과가 적더라도 맛이 좋아서 제품을 구매할 확률이 높으며, 효과가 있더라도 맛이 맞지 않으면 이탈할 확률이 높다. 평균적으로 맛에 대한 평가가 가장 좋고 다음이 효과, 가격 순이다.

긍정평가에 있어서도 효과의 영향력은 강하지 않으며, 부정평가에 있어서는 효과가 없다고 느낄 경우 가장 유의하고 강한 영향력을 가지는 만큼 효과는 떨어지는 제품이라 할 수 있다.

맛에 대한 특징을 제외한다면 효과, 맛, 가격 세가지 측면에서 어느 정도 밸런스가 잡힌 측면에서 컨디션의 하위 호환 제품으로 볼 수도 있다. 맛에 부정반응하는 경우 탄산에 대한 언급(빈도분석)이 가장 많다.

배송이나 포장에 좋을 것은 디폴트로 당연하게 생각하는 경향이 있으며(i.e., 포장/배송이 좋다/빠르다고 해서 제품 만족도/충성도가 올라가는 것은 아니다), 반면 배송/포장이 부실한 경우 매우 민감하게, 부정적으로 반응하므로 제품의 포장/배송에 더 신경을 써야 할 필요가 있다.



## 4.5.상패환

### RESULT

		<i>Beta Coefficients</i>	<i>Partial-R<sup>2</sup></i>	<i>R<sup>2</sup>(adj)</i>	<i>R<sup>2</sup>(adj)</i>
<b>Target Effect (+)</b>	효과	1.20*** (0.17)			0.618
	맛	0.31 (0.22)			
	가격	0.49*** (0.18)			
<b>Target Effect (-)</b>	효과	-0.69*** (0.23)			
	맛	-1.04** (0.47)			
	가격	0.44 (0.36)			
<b>Control</b>	배송/포장(+)	0.09 (0.34)			
	배송/포장(-)	-1.34*** (0.24)			

(1) \*\*\* p<0.01, \*\*p<0.05, (2) SE in parentheses

### ANOVA

F-value	Critical F
28.359	0.000

위 결과치 및 implications에 대해 말로 풀어서 설명하면:

만들어진 모형은 매우 유용하며 모형의 설명력은 매우 뛰어나다. 고객들은 효과, 가격 순으로 제품에 대해 민감하게 긍정반응하며, 맛은 제품의 평가에 영향을 미치지 않는다. 효과, 맛 순으로 민감하게 부정반응하며, 가격은 제품의 평가에 영향을 미치지 않는다.

맛은 버리고, 가성비 대비 효과가 가장 좋은 제품이라 할 수 있다. 맛이 좋지 않더라도 (가격도 저렴한데다가) 순수하게 숙취제거 효과를 위해 구매하게 되는 제품이다. 맛이 좋지 않더라도 효과를 보고 제품에 대한 충성도가 올라가는 경향이 있으며, 낮은 확률로 이탈할 경우에는 개인적인 효과도 나쁜데다가 맛까지 좋지 않아서 이탈할 확률이 높다.

배송이나 포장에 좋을 것은 디폴트로 당연하게 생각하는 경향이 있으며(i.e., 포장/배송이 좋다/빠르다고 해서 제품 만족도/충성도가 올라가는 것은 아니다), 반면 배송/포장이 부실한 경우 매우 민감하게, 부정적으로 반응하므로 제품의 포장/배송에 더 신경을 써야 할 필요가 있다.

## 5. 제품간 비교

변수 설명력과 standardized beta coefficient를 기준으로 rescaled score를 환산해보면, 다음과 같이 항목별 순위를 산출할 수 있다.

제품 \ 항목별 순위	효과 만족도	맛 만족도	쿠팡가 기준 상대적 가격 만족도	배송/포장 만족도
컨디션	(3) 32***	(3) 10***	(3) 9***	(5) -59***
레디큐	(4) 17***	(1) 53***	(-) 4	(1) -23***
여명808	(2) 39***	(4) 10**	(2) 12***	(4) -40***
깨수깡	(5) 14**	(2) 39***	(4) 11*	(3) -38***
상쾌환	(1) 56***	(-) 1	(1) 17***	(2) -32***