

AMORE PACIFIC

연세대학교 산학 협력 프로젝트

Taeyoung Park, Professor of Applied Statistics

Seunghan Lee | Kyoungmin Cho

Jungkyu Kim | Seseung Park

Yonsei University

분석 목적 및 목차

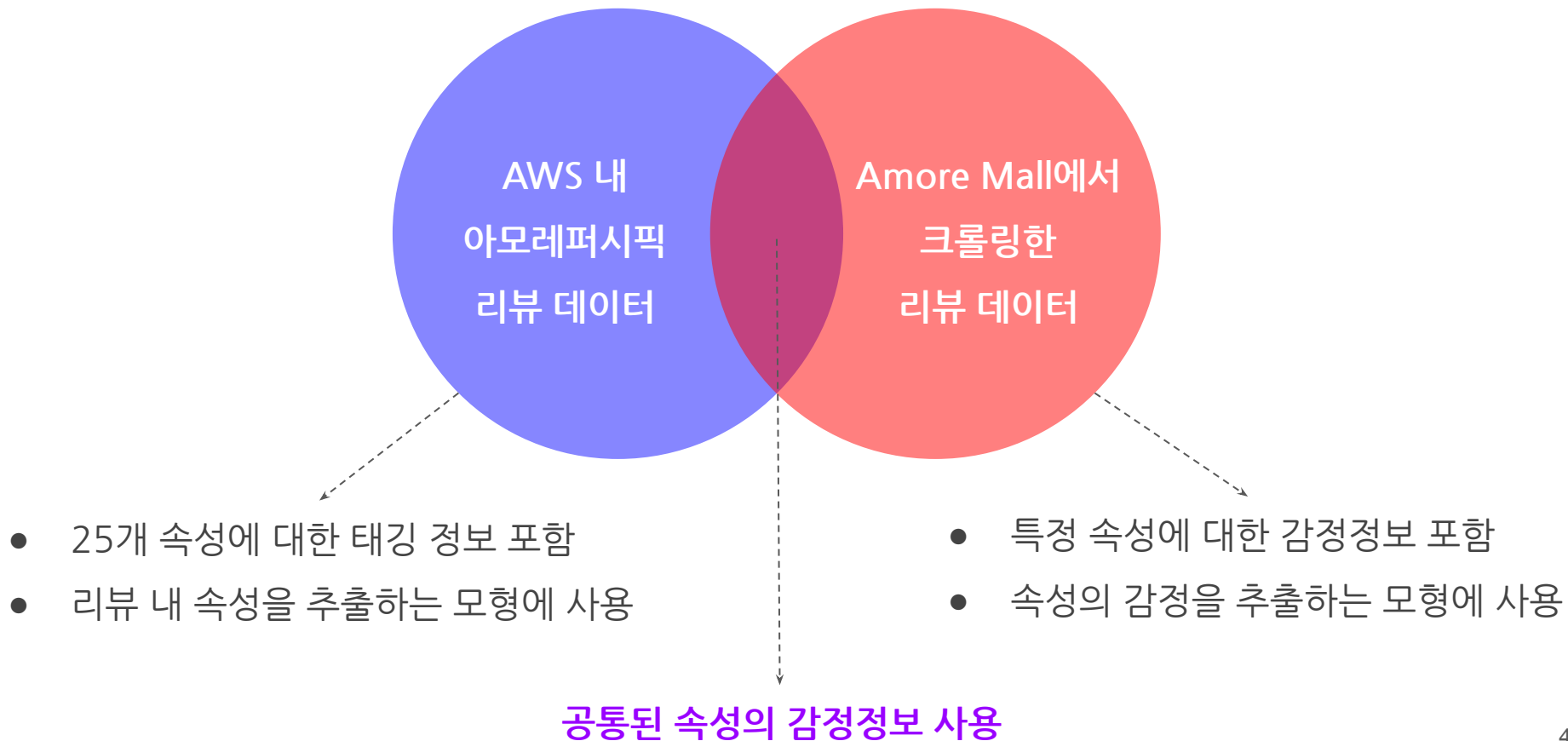
- 리뷰 및 SNS 데이터를 이용해 **브랜드 중요성**을 나타낼 수 있는 지표 개발
- 브랜드 중요성의 **모니터링**을 통해 브랜드 이미지의 전략적 관리
- 브랜드 가치에 대한 지표를 소비자 반응 및 수요 예측 등의 **부가적인 목적에 활용**

- 사용한 데이터 소개 및 설명
- Semantic Brand Score (SBS)
 - 확산성, 다양성, 연결성의 3차원 요소로 구성
- Contextual Brand Score (CBS) — **리뷰 및 속성의 감정을 고려**
 - 인상, 다각성, 연관성의 3차원 요소로 구성
- Contextual Brand Score Extended (CBSx) — **추가적으로 리뷰의 품질을 고려**
 - 인상, 다각성, 연관성, 품질의 4차원 요소로 구성

01

Data Preparation

사용한 데이터 설명



웹크롤링한 리뷰 데이터

- 총 8개의 **대표적인 브랜드**에 대해 크롤링 진행: 바이탈뷰티는 건강식품 위주이므로 제외

BEST BRAND

Sulwhasoo

H E R A

primera

AMORE PACIFIC



IOPE

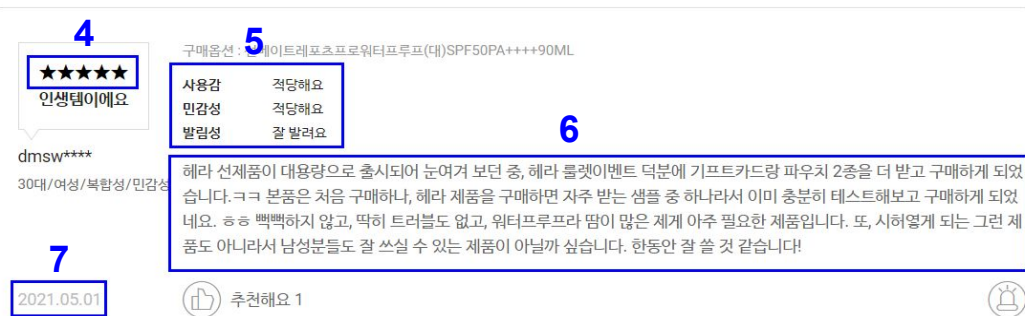
LANEIGE

Mamonde

한울

- 웹크롤링한 데이터 목록

- 1) 브랜드명
- 2) 제품명
- 3) 제품군
- 4) 전체평점
- 5) 속성 종류 및 감정
- 6) 리뷰 텍스트
- 7) 날짜



최종 데이터

1		7		2		4		6		5													
brand		date		prod_id	prod_nm	rating	review		A 용 량	A 민 감 성	A 향 기	A 커 버 력	...	A 사 용 감	A 발 림 성	A 세 정 력	A 촉 촉 함	A 유 분 기	A 발 색 감	A 제 형	A 패 키 지	A 실 용 성	A 보 습 력
0	sulhwasoo	2021.07.11	21802	핸드워시 백은향 300ml	5	평소에 너무 써 보고 싶었던 제 품이에요~ 쿠폰 적용해서 엄청 저렴하게 구매했 어요!! ...		적 당 해 요	적 당 해 요	적 당 해 요	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	sulhwasoo	2021.02.25	21802	핸드워시 백은향 300ml	5	바디워시 쓰고 반해서 핸드워시 도 사게 됐어요 바디워시와 같은 향이라 그런지 좋네요 건...		적 당 해 요	적 당 해 요	향 이 줄 아 요	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

3. 카테고리(대분류) : 9개의 제품군 (클렌징, 모이스처, 선케어, 페이스, 립, 아이, 색조, 바디, 헤어)

데이터 전처리

- 2020년 8월 1일 ~ 2021년 7월 31일까지 8개 브랜드의 총 50,183개 리뷰 데이터 사용
- 전체 평점은 클래스 비율을 고려하여 “1~3점 / 4점 / 5점”을 각각 부정(-1), 중립(0.5), 긍정(1)으로 분류



- 클래스 비율: 4%(1~3점) / 19%(4점) / 77%(5점) ➡ 클래스의 불균형이 심함
- 속성 표현(ex. 부드럽다, 건조하다 등)을 감정으로 변환: 부정(-1), 중립(0), 긍정(1)
- 제품군 분류는 아래와 같음

제품군 (대분류)	제품군 (소분류)	제품군 (대분류)	제품군 (소분류)
catA (스킨케어)	스킨케어_클렌징	catD (눈M)	메이크업_아이
	스킨케어_모이스처라이징		메이크업_기타색조(치크)
	스킨케어_선케어	catE (사용X)	향수
catB (얼굴M)	메이크업_페이스	catF (생활용품)	생활용품_바디용품
catC (입술M)	메이크업_립		생활용품_헤어용품

02

Semantic Brand Score (SBS)

Semantic Brand Score (SBS)

1) Prevalence (확산성)

- 브랜드가 얼마나 많이 언급되는지
- SNS에서 브랜드가 언급되는 빈도 계산

2) Diversity (다양성)

- 얼마나 다양한 속성과 함께 언급되는지
- 브랜드와 함께 언급되는 단어의 개수 계산

3) Connectivity (연결성)

- 단어 간 관계에서 얼마나 많이 나타나는지
- 브랜드가 단어들을 바로 연결하는 빈도 계산



리뷰 데이터에서 SBS의 계산

- **SNS 데이터**는 불특정 다수가 브랜드에 대해 언급한 데이터
- **리뷰 데이터**는 제품을 구매한 고객이 제품을 평가하고 브랜드에 대한 직접적인 언급이 없음
 - 제품의 리뷰를 브랜드에 대한 리뷰로 볼 수 있음
 - 제품이 많이 팔릴수록 리뷰의 개수가 많아짐
 - 리뷰의 길이가 다양함



보습감
향
민감성

촉촉해요
향이 좋아요
은해서 좋아요

저 3년만에 인생화장을 만났습니다 제 피부타입을 말하자면 국메인피부예요 겉으로 돌아나는 트러블은 조금밖에 안나서 다행이지만 저나 아주 가까이에서 본 분들은 아는 그런 트러블이요 어릴때부터 화장품만 쓰면 트러블나서 베이비로션 천연화장품들도 써왔지만 저의 예민한 피부는 그들도 이기더라고요 아이러니하게 피부에 아무것도 안발랐을 때가 제일 피부가 좋았어요 성인되고 안바르기에는 이제 힘들어서 겨우 찾은게 크리크 분홍크림 켈 타임 이에요 제가 크림의 무게는 느낌 변들거리는 느낌이 너무 싫어서 트러블까지 고려해서 선택한게 분홍크림인데 몇년 그럭저럭 썼는데 30대가 넘어가니 켈타임은 아무래도 너무 가벼워서 또 화장품 유무한 생활을 했어요 그러다 지인에게 추천받은 한울 이런쪽 진정크림 !! 리뷰까지 꼼꼼히 읽고 마침 행사중이라 또 속아보자 하는 마음으로 구매했어요 당시 제 피부 상태는 화장술 사용으로 한겨울에도 안그릴만큼 기절기질 사포갈고 얼굴에 붉은기가 장난아니였거든요 도약해서 발라볼마다 제일 눈에 보이는 트러블 다행히도 안났습니다 발림성도 크림인데도 놀랍게도 무겁지 않고 가벼우면서 촉촉해요 변들거리는거 전혀 없구요 크림 제품중에는 얼굴에 땀가 빠진기 처럼 답답하게 만드는 것도 있잖아요 한울크림은 전혀 안그래요 발림성이 너무 좋아서 제발 저에게 맞이라 맞이라 했는데 일주일정도 써본 결과 트러블 전혀 없구요 사포갈이 거친 피부도 다시 보습보습으로 돌아왔어요 얼굴이 부들부들해져서 너무 기분 좋고 얼굴빛도 조금은 밝아진거 같은 착각 아닌 착각도 드네요 붉은기도 물론 박 없어졌습니다 한울과 추천해주 지인에게 절하고 싶어요 ㅠㅠ강추

VS.

보습감
향
민감성

적당해요
향이 별로예요
적당해요

얼마가 흘러구 싶는데 향 별로예요 ㅠㅠ

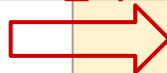
- **리뷰 데이터**로 SBS를 계산 시, 리뷰의 개수 및 길이에 의해 편향 발생 가능성
 - **확산성**: 리뷰의 개수가 많을수록 유리함
 - **다양성**: 리뷰의 개수가 많고 리뷰의 길이가 길수록 유리하고, 긴 리뷰를 남긴 고객에 의존함
 - **연결성**: 리뷰의 길이가 길수록 유리함

리뷰 데이터에서 SBS의 계산

- **확산성**: 리뷰를 브랜드에 대한 언급으로 취급하고 전체 리뷰 중 해당 브랜드 리뷰의 빈도 계산

제품명	리뷰 개수	확산성
자연을 닮은 시트마스크	1948	7.732658
오가니언스 워터	1069	3.998003
달빛유자 수면팩	957	3.522142
...
로즈 블레미쉬 클리어링 세럼	1	-0.539668

표준화

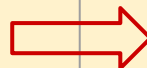


리뷰 데이터에서 SBS의 계산

- 다양성**: 리뷰 내 등장하는 단어들의 언급 개수로 표현 다양성에 대한 계산

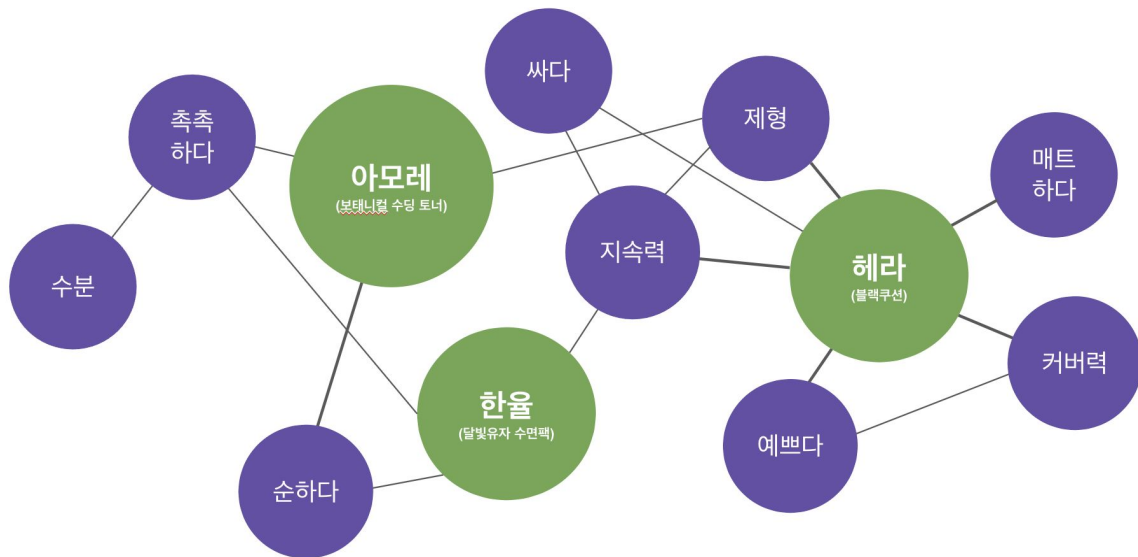
제품명	리뷰	명/동/형	고유한 명/동/형	다양성
진설크림	저희 어머니께서는 이제품을 계속 사용하고 계십니다. 제가 매번 주문해드려요. 저고 몇번사용해보니 어마니가 왜 계속 사용하는지 알거 같더라고요. 기름지지도 않고 적당히 촉촉해요. 가격이 많이 사악한편이지만 믿고 구매하는 제품입니다.	'기름', '지지도', ... '촉촉하다', '저', '몇번', '사용', ... '왜', '계속', '사용', '하다', ... '가격', '사악', '한편', '믿다', '구매', '하다', '제품', '이다', '저희', '어머니', '제품', ...	30	0.712
...		
자연을 닮은 시트마스크	빨간쌀 스킨제품 사용하고 있었는데 피부에도 맞고 좋은것 같아 마스크 제품도 구매했어요. 빨간쌀의 보습효과를 느낄수 있어요. 마스크 하고나면 피부가 촉촉한게 느껴지네요. 피부상태에 따라 산들박하와 빨간쌀 마스크를 번갈아 사용하려구요	'피부', '상태', '따르다', '산들', '박하', '빨강다', '쌀', '마스크', '번갈아', '사용', '하다', '마스크', '하다', '피부', '촉촉하다', '느껴지다', '빨강다', '쌀', '보습', '효과', ...	25	0.538

표준화

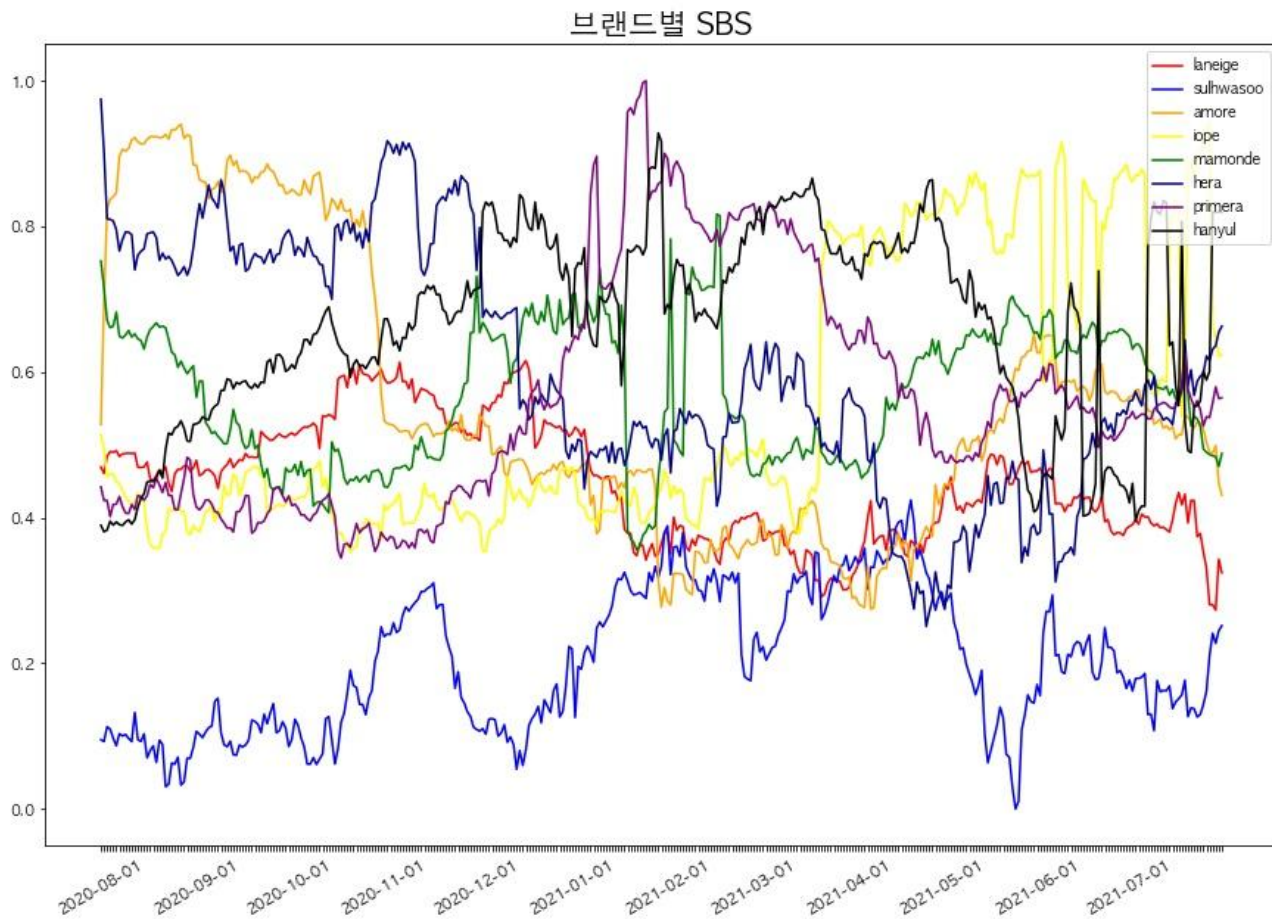


리뷰 데이터에서 SBS의 계산

- **연결성**: 브랜드가 단어들을 바로 연결하는 빈도에 대한 계산
 - 리뷰 자체를 “제품명”의 노드(Node)로 사용
 - 네트워크 연산량을 고려하여 최빈 단어(명/동/형) 상위 K개를 노드로 사용
 - 모든 “제품명”과 최빈 단어 상위 K개의 네트워크 구축 후, **매개 중심성** 계산



리뷰 데이터의 SBS 시각화



03

Contextual Brand Score (CBS)

Contextual Brand Score (CBS)

- SBS는 브랜드에 대한 **감정**을 담고 있지 않아 브랜드의 가치를 제대로 전달하지 못할 수 있음
 - 긍정적이든 부정적이든 브랜드에 대한 언급이 많으면 SBS가 높게 나오는 경향
 - 긍정적이든 부정적이든 브랜드와 관련된 속성이 많으면 SBS가 높게 나오는 경향
 - 리뷰 이벤트나 프로모션의 효과에 지나치게 민감한 경향
- CBS(**문맥적 브랜드 점수**)는 브랜드 및 브랜드와 관련된 속성의 감정을 함께 고려하고자 함
- CBS의 특징
 - 브랜드에서 직관적으로 받게 되는 **전반적인 인상**을 설명
 - 브랜드와 구체적으로 연관된 **속성들의 감정을 다양한 각도에서 설명**
 - 속성이 가진 **긍/부정적 이미지의 연관성**을 종합적으로 반영

Contextual Brand Score (CBS)

1) Impression (인상)

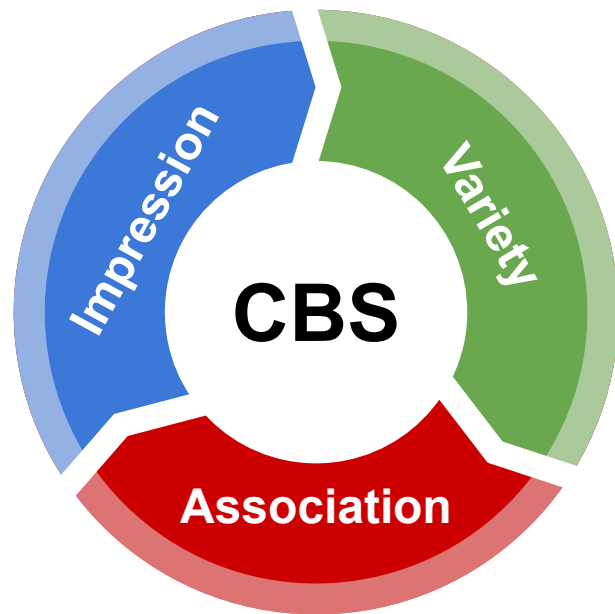
- 브랜드에 대해 전반적으로 느끼는 긍정(+)/부정(-) 감정
- 브랜드에 대한 **긍·부정값**의 합산

2) Variety (다각성)

- 브랜드와 같이 등장하는 다양한 긍정(+)/부정(-)적 속성
- 브랜드와 **함께 언급되는 속성에 대한 긍·부정값**의 합산

3) Association (연관성)

- 브랜드를 통해 긍정(+)/부정(-)적 속성이 연결되는 양상
- 직접 연결되지 않은 **두 속성이 브랜드를 통해 연결되는 긍·부정 연결강도**의 합산



리뷰 데이터에서 CBS의 계산

1) Impression (인상)

- 의미: 브랜드에 대해 전반적으로 느끼는 **긍·부정 감정**
- 방법: 리뷰에 대한 **전체적인 감정(예측된 평점 기준)**을 종합하여 계산
- 실행: “**속성 탐지 + 다중 작업 학습**”을 통해 나온 **전체 평점값** 활용

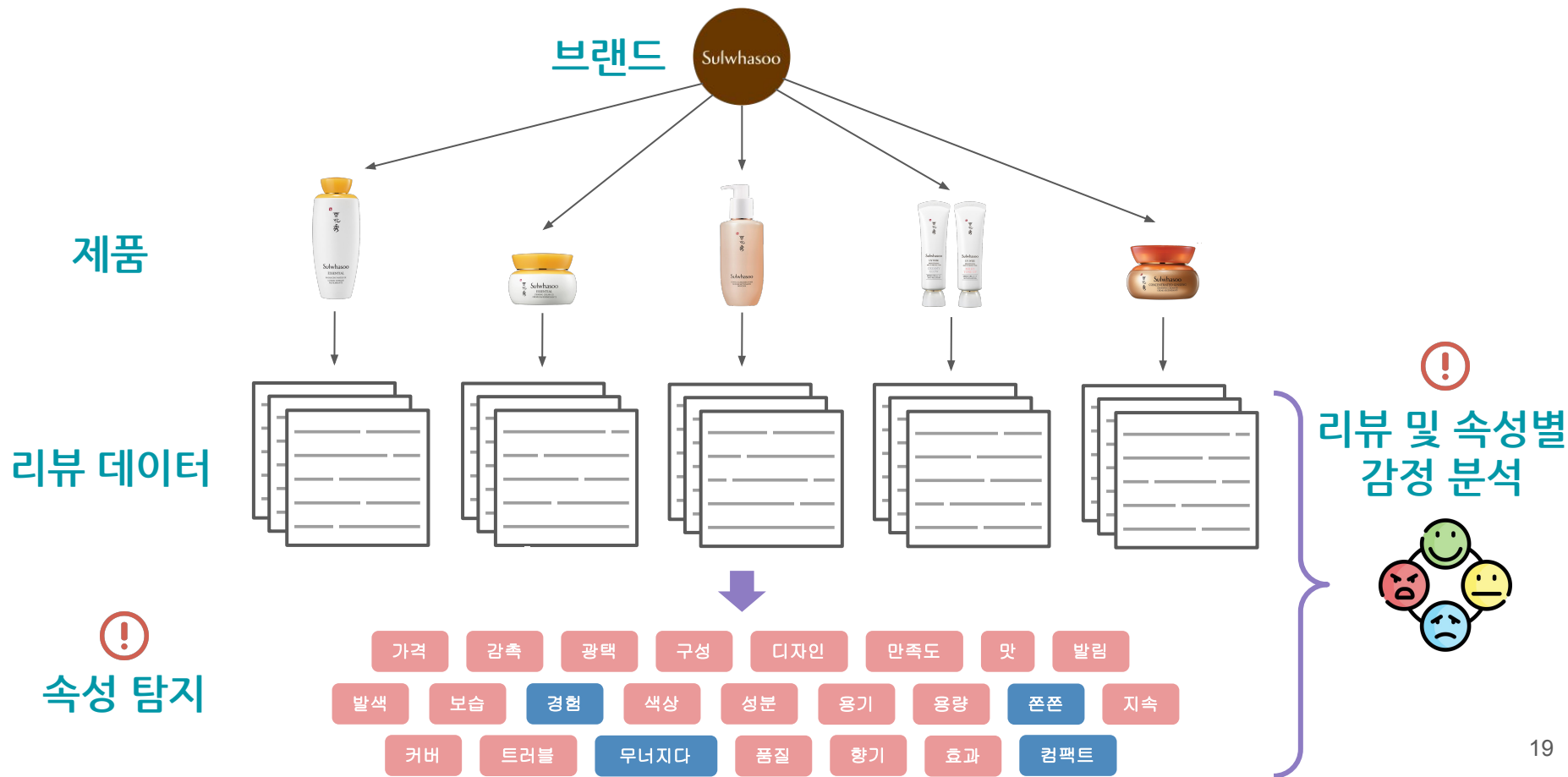
2) Variety (다각성)

- 의미: 브랜드와 **함께 언급되는 다양한 긍·부정 속성**
- 방법: 제품에 대한 **속성의 감정(예측된 감정 기준)**을 종합하여 계산
- 실행: “**속성 탐지 + 다중 작업 학습**”을 통해 나온 **속성 감정값** 활용

3) Association (연관성)

- 의미: 브랜드를 통해 **긍·부정 속성이 연결되는 양상**
- 방법: 제품에 대한 속성 네트워크에서 **긍·부정 연결강도를 고려한 매개 중심성** 계산
- 실행: “**속성 탐지 + 다중 작업 학습**”을 통해 나온 **속성 감정값** 활용

리뷰 데이터에서 CBS의 계산



리뷰 데이터에서 CBS의 계산

- **속성 탐지**: 속성 탐지 (Aspect Detection, AD) 모형 적용
- **감정 분석**: 어텐션 기반 다중 작업 학습 (Attention-based Multi Task Learning, AMTL) 모형 적용
 - 리뷰에 대한 **전반적인 감정** 추출
 - **리뷰의 감정 분석**: 고객이 남긴 평점을 기반으로 감정 학습
 - 리뷰가 포함하고 있는 **속성에 대한 감정** 추출
 - **속성기반 감정 분석**: 고객이 남긴 속성에 대한 평점을 기반으로 감정 학습
- 고객이 남긴 평점을 “**그대로**” 사용하지 않고 “**예측값**”을 사용
 - 사람마다 판단 기준점이 다르기 때문에 계단식 평점 대신 **연속형 평점**을 사용
 - 각 평점에 속할 확률을 기반으로 가중평균 계산
 - 평점에 대한 **보정 효과** 기대 가능

1. 속성 탐지

- 속성 탐지 모형:

- 리뷰 데이터가 주어졌을 때, 속성의 탐지를 자동화
- 총 25개의 속성에 대해 태깅이 되어 있는 AWS 내 리뷰 데이터로 모형을 훈련

리뷰	보습	가격	커버력	트러블	...	발색
버터가 촉촉하니 피부가 당기지 않네요 보습력이 좋아요~ 트러블도 없고 가격도 만족해요	1	1	0	1	...	0
바르자마자 선명하게 색감이 올라오는게 너무 좋았어요! 착색도 원래 색 그대로였습니다. 가성비 좋아요!	0	1	0	0	...	1
...
건성 피부엔 괜찮을듯 하고 요건 커버력은 별로 없어요 ㅠㅠ	1	0	1	0	...	0

- Amore Mall에서 크롤링한 데이터를 가지고 속성을 탐지

1. 속성 탐지

(1,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0)

출력

Fully Connected Layer (25차원)

3) 추출된 특성을 바탕으로 속성 유무를 판단하는 다중 분류 수행

Bidirectional LSTM

2) 양방향 LSTM을 사용해 문맥을 압축한 특성 추출

Word2Vec 임베딩 (128차원)

1) Word2Vec을 사용하여 리뷰 텍스트 임베딩

리뷰 #1

리뷰 #2

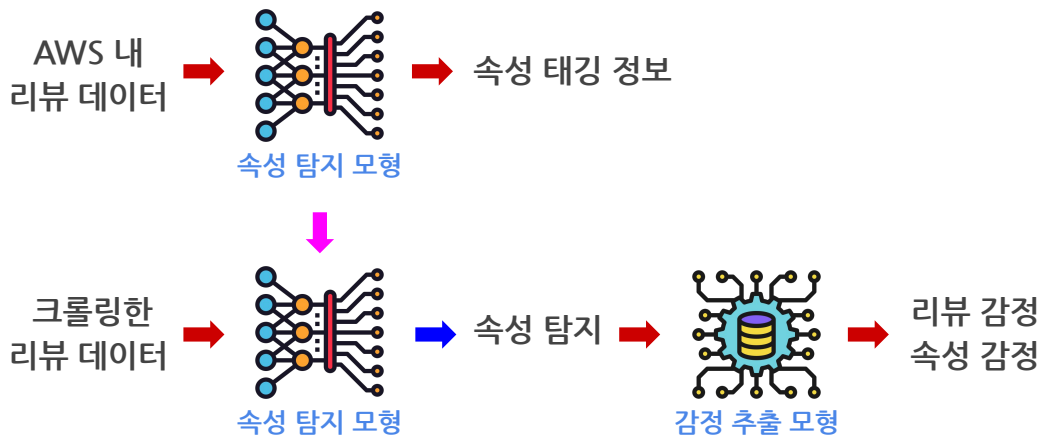
...

리뷰 #N

입력

1. 속성 탐지

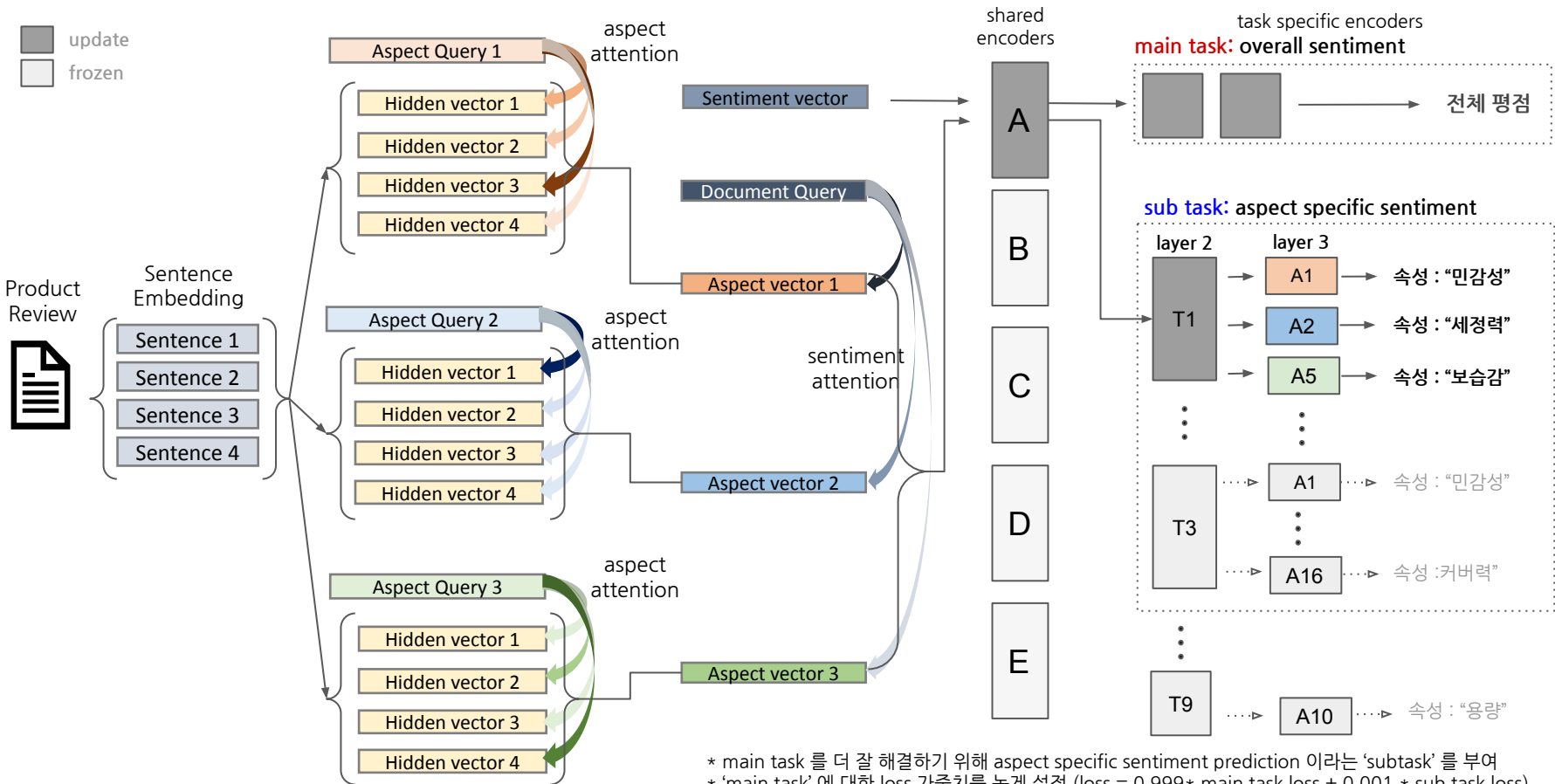
- 약 65만 개의 **AWS 내 리뷰 데이터**를 정제한 뒤 **속성 탐지 모형** 학습에 이용
- Amore Mall에서 **크롤링한 리뷰 데이터**를, 훈련된 모형의 입력값으로 사용해 속성 탐지
 - 예측 성능 및 결과 : **정확도: 0.93 / 정밀도: 0.8 / 재현율: 0.71 / F1 점수: 0.75**
 - 예측된 결과와 리뷰 텍스트 원본을 눈으로 직접 비교해도 합리적인 결과 도출
- 추출된 속성을 대상으로 **감정 추출 모형**을 학습



2. 감정 추출

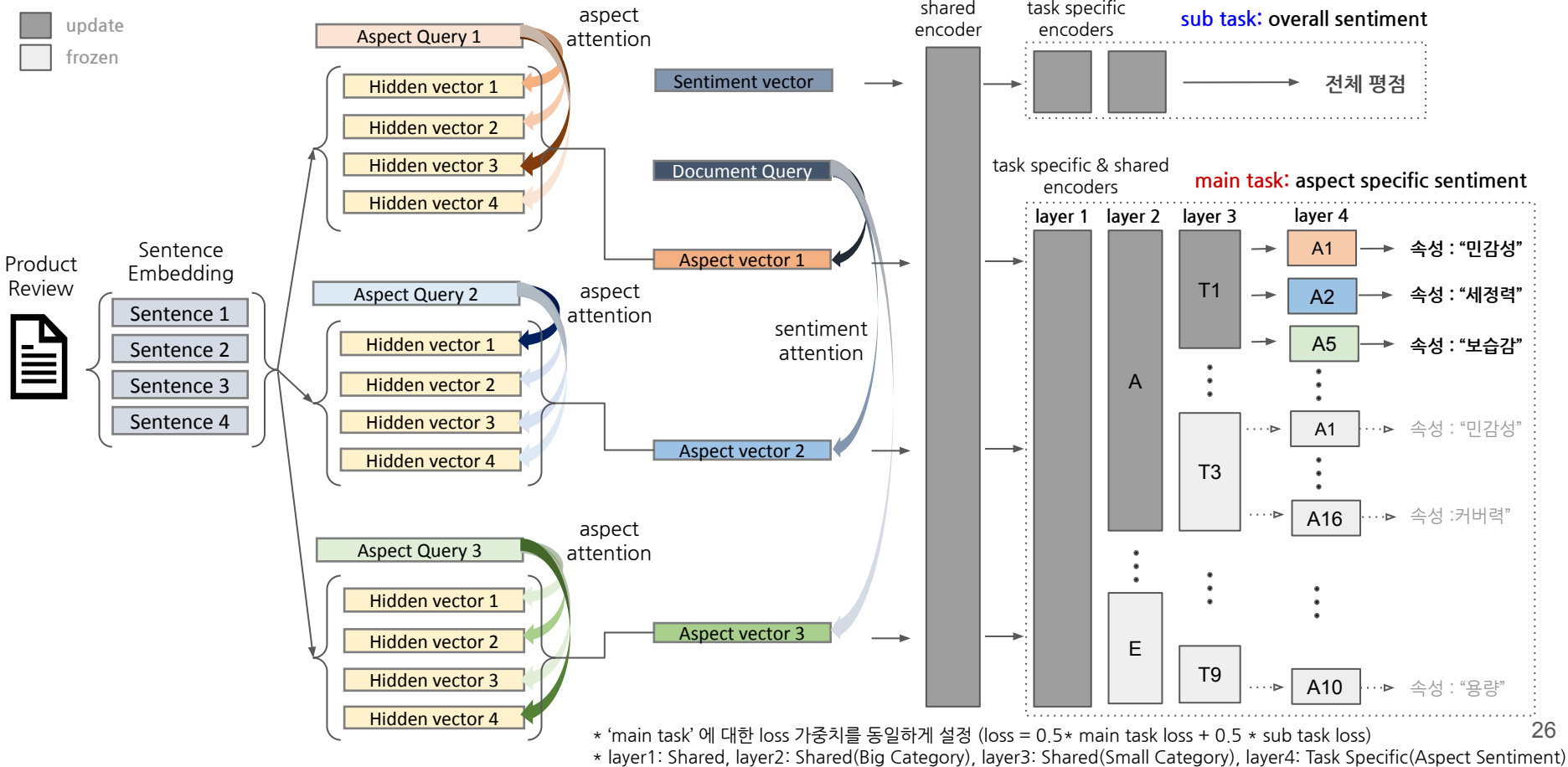
- **AMTL**(Attention-based Multi-Task-Learning; **어텐션 기반 다중 작업 학습**) 모형 구축
 - 작업#1 (General): 리뷰의 **전반적인 감정** (부정/중립/긍정) 예측
 - 작업#2 (Specific): 제품군의 여러 **속성에 관한 감정** (부정/중립/긍정) 추론
 - 예) 제품군 = **스킨케어_클렌징**, 속성 = **민감성, 세정력, 촉촉함**
 - 각 리뷰가 속한 총 9개의 제품군에 대해 작업#1과 작업#2를 **동시에** 추론
 - **AMTL-1** (전반적인 감정 추출)
 - **전반적인 감정**을 중심(main)으로, **속성에 관한 감정**을 부가적(sub)으로 해결, **정확도는 약 73%**
 - **AMTL-2** (속성에 관한 감정 추출)
 - **속성에 관한 감정**을 중심(main)으로, **전반적인 감정**을 부가적(sub)으로 해결, **정확도는 약 70%**
- **어텐션 기반 다중 작업 학습**의 장점
 - **자원 효율화**: 제품군별로 여러개의 개별 모델이 아닌, 제품군을 통합한 한 개의 모델만을 학습
 - **추론 능력 향상**: 제품군 간 공유되는 **전역적인** 특징과 제품군 내의 고유한 **지역적인** 특징을 동시에 고려
 - **데이터 불균형 해결**: 데이터가 불충분한 특정 제품군의 경우, 다른 제품군의 추론 정보를 활용

2. AMTL-1 (전반적인 감정 추출)



* main task 를 더 잘 해결하기 위해 aspect specific sentiment prediction 이라는 'subtask' 를 부여
 * 'main task' 에 대한 loss 가중치를 높게 설정 ($\text{loss} = 0.999 * \text{main task loss} + 0.001 * \text{sub task loss}$)
 * layer1: Shared(Big Category), layer2: Shared(Small Category), layer3: Task Specific (Aspect Sentiment)

2. AMTL-2 (속성에 관한 감정 추출)



CBS - 인상

- 인상(Impression) 점수

- “속성 탐지 + 감정 추출”을 거쳐서 나오게된 특정 리뷰에 대한 종합 감정 점수를 활용
- 학습된 모델을 통해 리뷰별로 특정 클래스에 속할 확률을 계산
- 평점 1~3점/4점/5점을 각각 -1(부정), 0.5(중립), 1(긍정)로 두고 확률을 가중치로 하여 가중 평균

평점	부여 감정점수	리뷰 텍스트	클래스에 속할 확률	고객이 남긴 평점
1~3 점	-1	항상 마스크를 쓰고다니다 보니 자연스레 화장을 안하게 되는데 마침 좋은기회가 왔네요~~. 새롭게 리뉴얼 되서 나왔는데 커버력과 지속력이 좋으며 얇게 발려서 밀착력도 좋네요~~퍼프 또한 물방울 모양이라 마음에 들어요~	0.09	4점
4점	0.5		0.31	
5점	1		0.6	



인상 점수 : $(-1 \times 0.09) + (0.5 \times 0.31) + (1 \times 0.6) = 0.665$

CBS - 인상: 평점 보정효과

제품명	실제 고객 평점 (1~5점)	실제 리뷰	모델 예측 평점 (-1~1점)
더마트리블 에멀전	5점	아이크림 발림성 좋아요. 적극 추천흡수 잘되구요향도 좋아요. 샘플도 많이 주시네요. 강추입니다!!!	0.81 (= (-1)x(0.07) + (0.5)x(0.11) + (1)x(0.82))
스텝 III 앰플	4점	평소에 촉촉한 립이나 글로스한 립을 선호하는 편이라 이 제품도 잘 맞을 것 같아서 구매했어요. 색상 정말 이쁘면서 뽕얼에도 잘 어울려요!! 그리고 향이 조금 있는데 저는 괜찮은 것 같아요. 근데 정말 지속력은 제로입니다... 아예 제로인 것 같아요. 편하게 쓱바르기 좋으면서 자주 덧발라야 하는..?	-0.03 (= (-1)x(0.44) + (0.5)x(0.30) + (1)x(0.26))

CBS - 다각성

● 다각성(Variety) 점수

- “속성 탐지 + 감정 추출”을 거쳐서 나오게된 **특정 리뷰에 대한 속성별 감정 점수**를 활용

	[속성] 보습력				[속성] 향기				...
리뷰	att	-1(부)	0(중)	1(긍)	att	-1(부)	0(중)	1(긍)	...
#1	0.3	0.2	0.4	0.4	0.7	0.15	0.35	0.5	...
#2	언급 안함				0.4	0.5	0.3	0.2	...
...									...
#98	0.5	0.5	0.4	0.1	언급 안함				...
#99	0.3	0.15	0.2	0.65	0.4	0.2	0.4	0.4	
#100	언급 안함				0.6	0.1	0.1	0.8	

리뷰 전체 평점에 대한 각 속성의 **기여도**



해당 리뷰 내 **“언급된”** 속성들이 리뷰 평점에 어느정도 기여했는지 표현 (언급되지 않은 속성은 att=0)

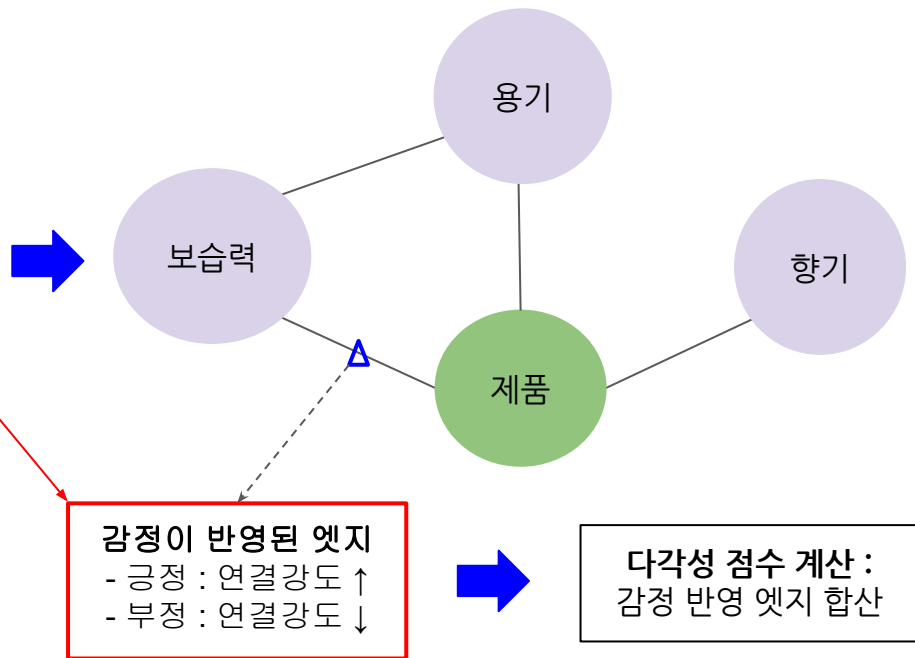
해당 속성이 각 감정 클래스에 속할 **확률**

CBS - 다각성

- 다각성(Variety) 점수

- 속성별 감정 점수를 바탕으로 속성 네트워크 형성하여 엣지값 합산

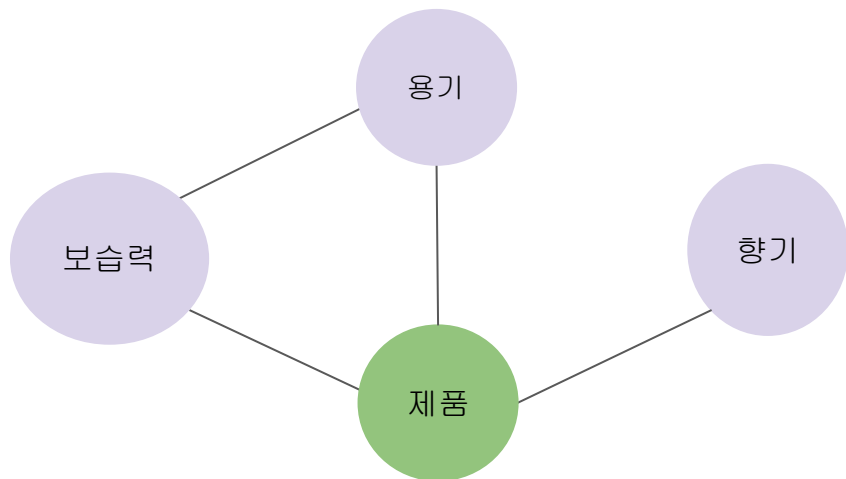
	[속성] 보습력				...
리뷰	att	-1(부)	0(중)	1(금)	...
#1	0.3 x (-0.2 + 0 + 0.4)				...
#2	언급 안함				...
...					...
#98	0.5 x (-0.5 + 0 + 0.1)				...
#99	0.3 x (-0.15 + 0 + 0.65)				
#100	언급 안함				



CBS - 연관성

- 연관성(Association) 점수

- 속성별 감정 점수를 바탕으로 속성 네트워크 형성하여 매개 중심성 계산

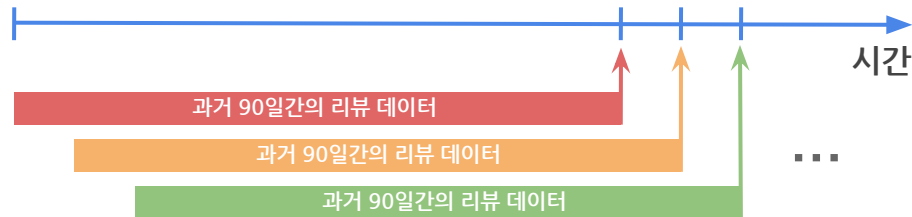


리뷰	속성
안티에이징 제품이라 기름지지 않을까 했지만 전혀 그렇지 않고, 촉촉 합니다. 속부터 채워준다는 느낌을 많이 받았어요.	보습력
개인적으로 화장품 향기 가 좀 나는 걸 좋아하는데 이건 약간 연한 두유 냄새 라고해야하나 공말린거 냄새 라해야하나 향기 는 제 취향이 아니었어요.	향기
용기 가 펌핑이라 공기유입이 덜 되느니라 세로로 추천해요. 두 방울 발리 가만히 퍼바르고선 지켜봤는데 어직까지도 촉촉함 과 시원함이 달라요.	용기, 보습력

리뷰 데이터를 이용한 CBS 계산

● CBS 계산 구조

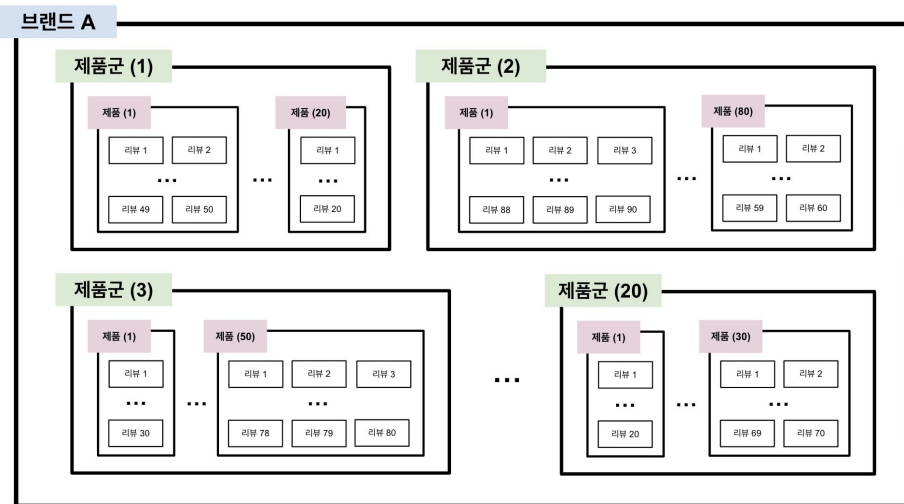
- 1) 해당일로부터 과거 90일간의 리뷰 데이터를 이용해 CBS의 3가지 요소(인상/다각성/연관성)를 계산
- 2) 제품별로 3가지 요소를 합산
- 3) 제품군별로 3가지 요소를 합산
- 4) 브랜드에 대해 3가지 요소를 합산



● CBS 특징

- 각 요소의 최대값은 해당 브랜드의 리뷰 개수가 됨
- 브랜드 점수는 제품군, 제품으로 **분해**될 수 있음
- 제품 및 속성에 대한 **감정이 반영**됨
- 과거 90일간의 리뷰 데이터를 기반으로 하여

이벤트나 프로모션의 영향력을 최소화



04

Contextual Brand Score Extended (CBSx)

Contextual Brand Score Extended (CBSx)

1) Impression (인상)

- 브랜드에 대해 전반적으로 느끼는 긍정(+)/부정(-) 감성
- 브랜드에 대한 **긍·부정값**의 합산

2) Variety (다각성)

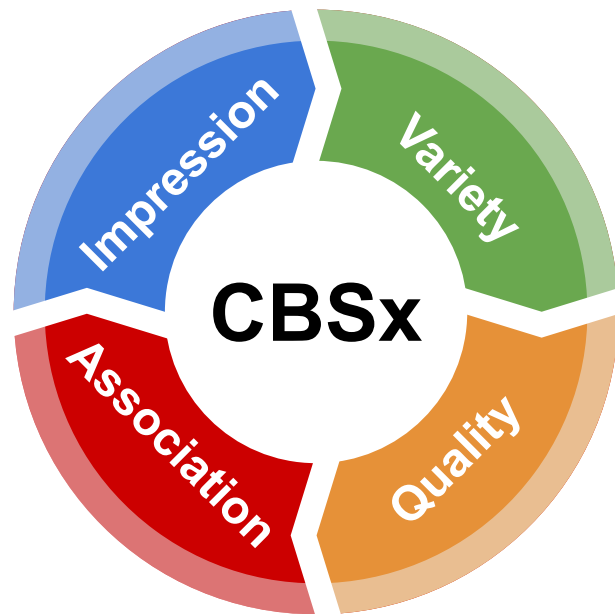
- 브랜드와 같이 등장하는 다양한 긍정(+)/부정(-)적 속성
- 브랜드와 **함께 언급되는 속성에 대한 긍·부정값**의 합산

3) Association (연관성)

- 브랜드를 통해 긍정(+)/부정(-)적 속성이 연결되는 양상
- 직접 연결되지 않은 **두 속성이 브랜드를 통해 연결되는 긍·부정 연결강도**의 합산

4) Quality (품질)

- 리뷰 데이터의 품질을 나타내는 추가적인 요소
- **정성도**(리뷰의 길이), **신뢰도**(고유한 문장의 비율), **표현력**(어휘의 다양성)



CBSx - 품질

- 품질(Quality) 점수

- 리뷰의 **정성도** : 리뷰의 길이를 (0,1)의 범위로 변환 (최소 리뷰 길이(20자)를 고려한 뒤, sigmoid-log-log 변환)
- 리뷰의 **신뢰도** : 고유한 문장의 비율 (중복 문장 제거)
- 리뷰의 **표현력** : 해당 리뷰에서 사용한 어휘의 다양성 (= 고유한 명.동.형 개수 / 전체 명.동.형 개수)

브랜드 A 내 제품(1)의 리뷰				
	리뷰의 정성도	리뷰의 신뢰도	리뷰의 표현력	품질 점수
리뷰 1	53	1	0.857	= 0.685
리뷰 2	116	1	0.735	= 0.607
리뷰 3	117	0.6	0.848	= 0.421

sigmoid-log-log 변환

CBSx - 품질: 보정효과

제품명	실제 리뷰	정성도	신뢰도	표현력	품질 점수
어린쑥 수분진정 크림	순하고 촉촉하네요. 배송도 빨랐어요. 잘 쓰겠습니다. 순하고 촉촉하네요. 배송도 빨랐어요. 잘 쓰겠습니다. 순하고 촉촉하네요. 배송도 빨랐어요. 잘 쓰겠습니다. 순하고 촉촉하네요. 배송도 빨랐어요. 잘 쓰겠습니다!!!	137	0.33	0.25	0.069
자연을 닮은 시트마스크	한올 시트팩만큼 밀착력 좋은제품 아직 못만난거같아요 전에 스페셜기프트로 받고 써봤는데 너무 좋아서 입소문내는중입자 최고예요	54	1.0	0.947	0.757

CBSx - 품질

- CBSx 계산 구조

- 1) 해당일로부터 과거 90일간의 리뷰 데이터를 이용해 CBS의 3가지 요소(인상/다각성/연관성)를 계산
- 2) **리뷰별 품질 점수**를 계산하여 CBS에 품질 점수를 곱함
- 3) 제품별로 3가지 요소를 합산
- 4) 제품군별로 3가지 요소를 합산
- 5) 브랜드에 대해 3가지 요소를 합산

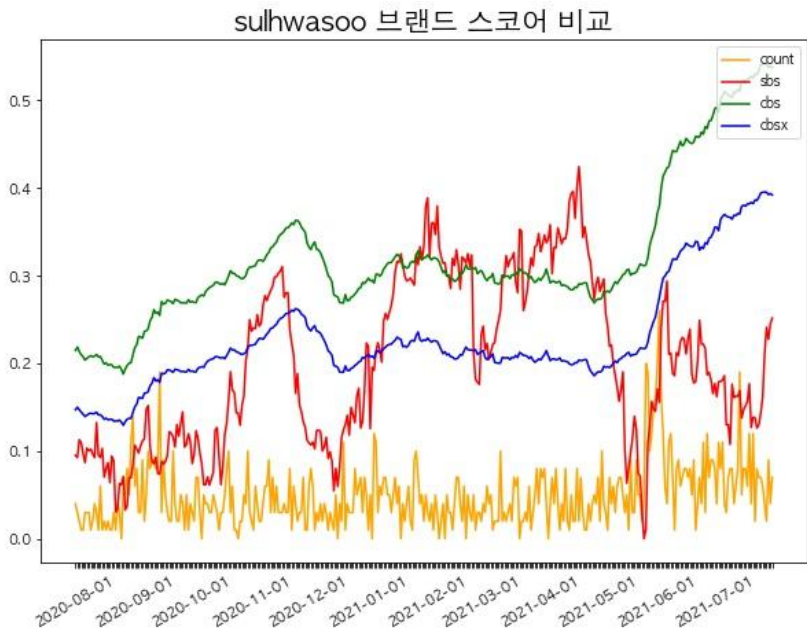
- CBSx 특징

- CBSx는 **리뷰의 품질**을 반영하여 CBS에 비해 낮아질 수 있음
- 리뷰의 품질은 **리뷰의 길이, 단순 반복 문장, 어휘의 다양성**을 종합하여 판단

05

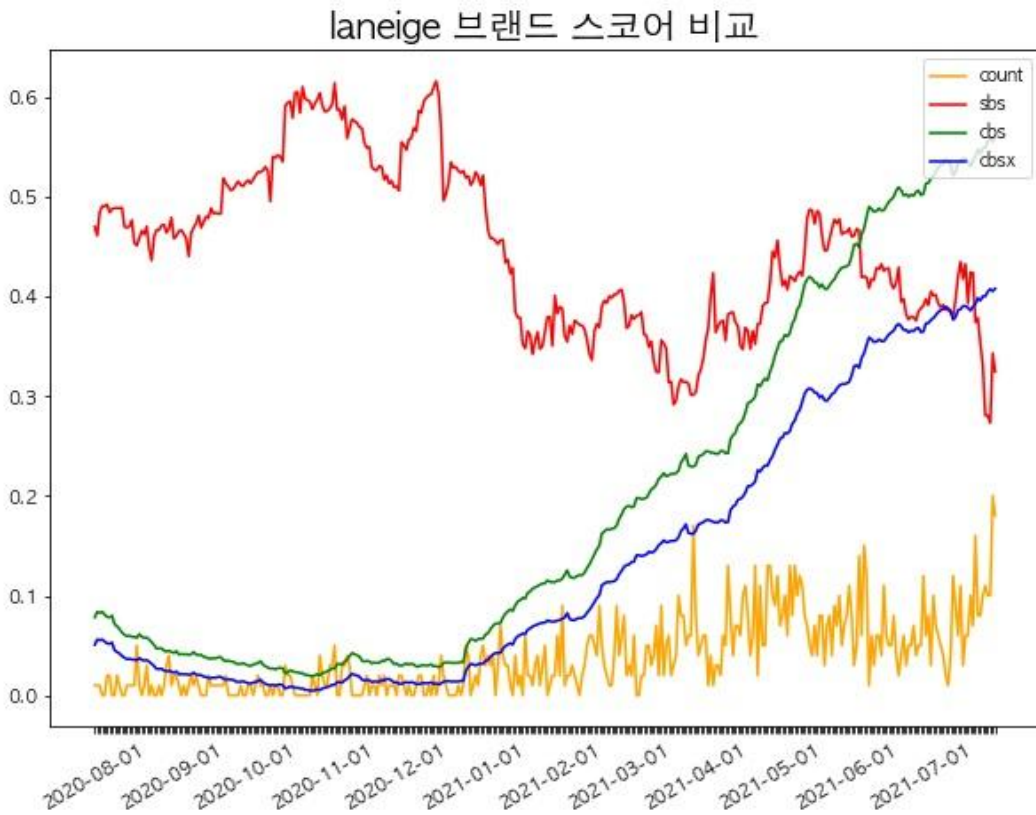
Score Visualization

브랜드 점수 비교 (설화수)

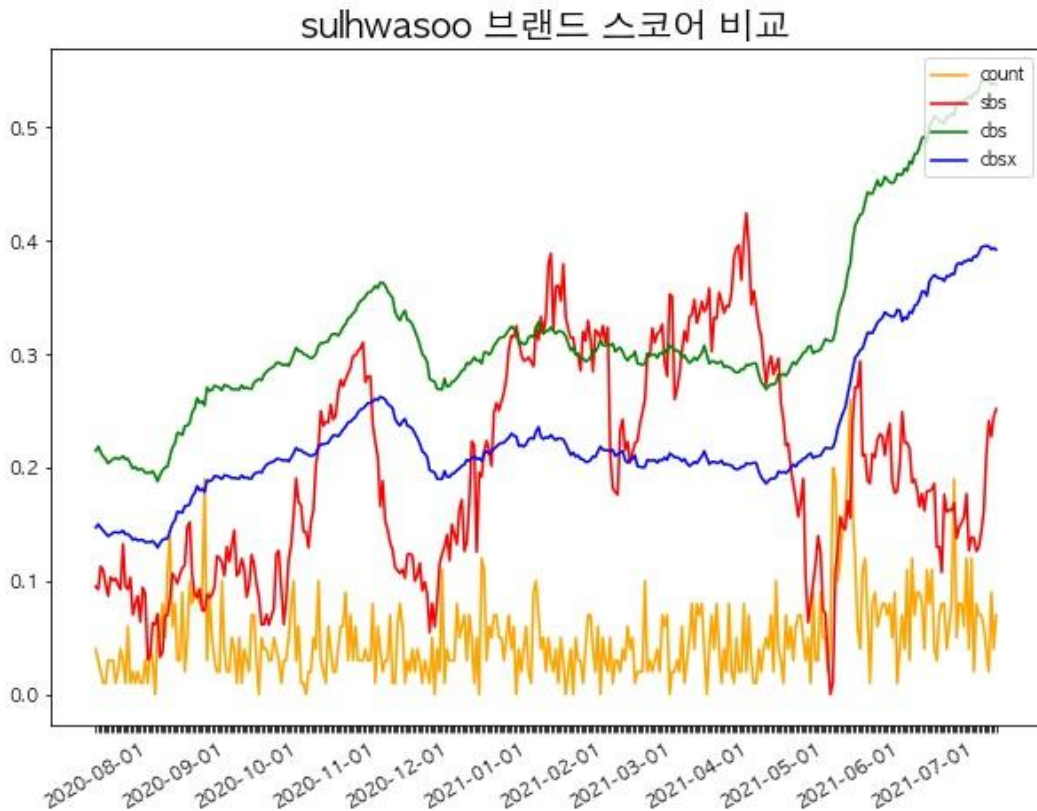


- SBS, CBS, CBSx를 각각 최소값과 최대값을 사용해 0~1 사이의 범위로 변환한 뒤, 한번에 비교
- SBS, CBS, CBSx는 모두 해당일로부터 **과거 90일간의 리뷰 데이터를 사용**하여 특정일의 리뷰 수와 반드시 연관성이 높지는 않음
- SBS는 **일별 변동성이 다소 높음**
- CBS 및 CBSx는 **감정 점수를 반영**한 추세를 보여줌
- CBSx는 **리뷰의 품질을 반영**하여 CBS보다 낮을 수 있음

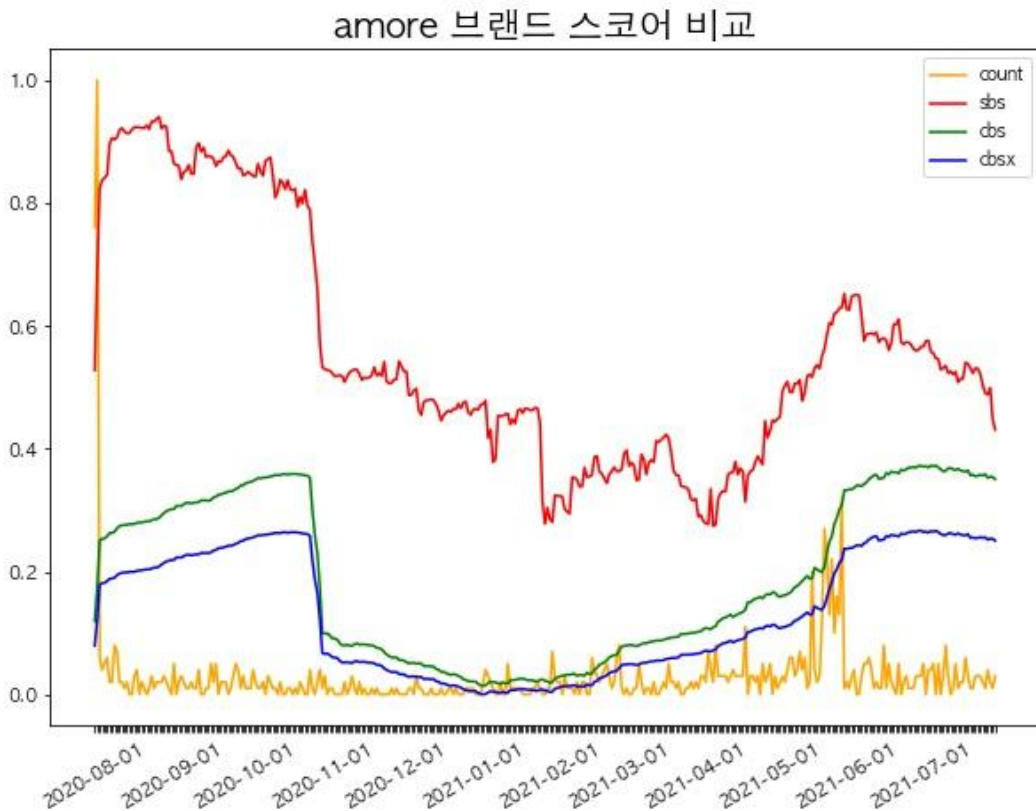
브랜드 점수 비교 (라네즈)



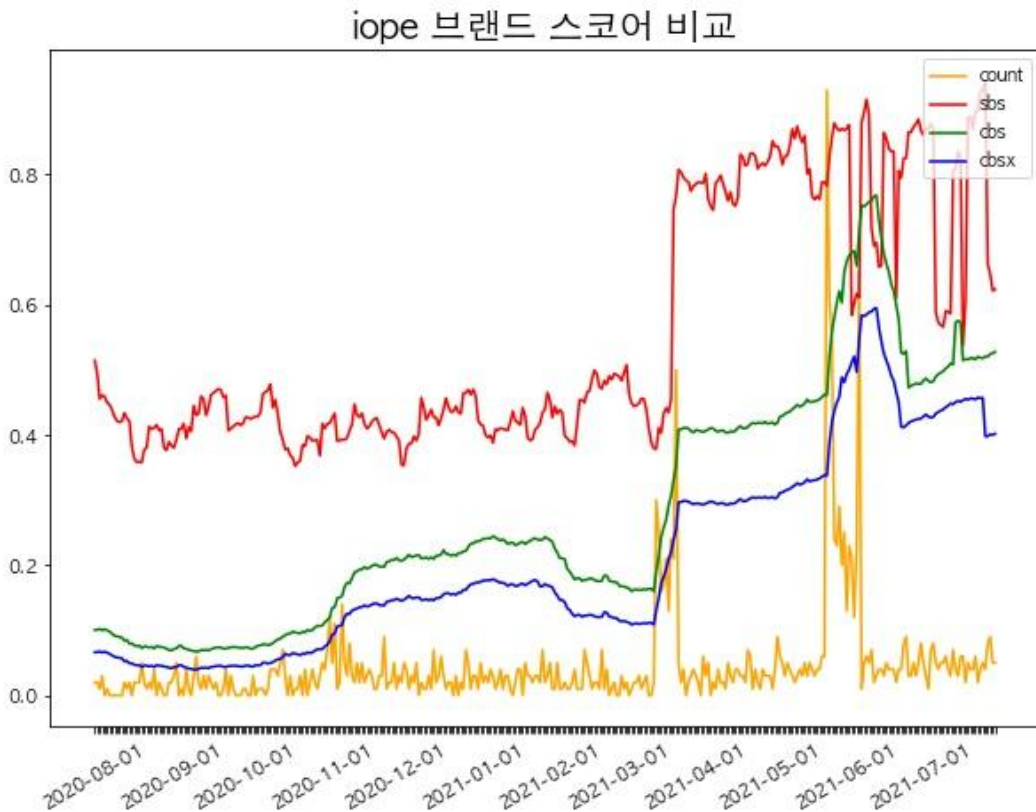
브랜드 점수 비교 (설화수)



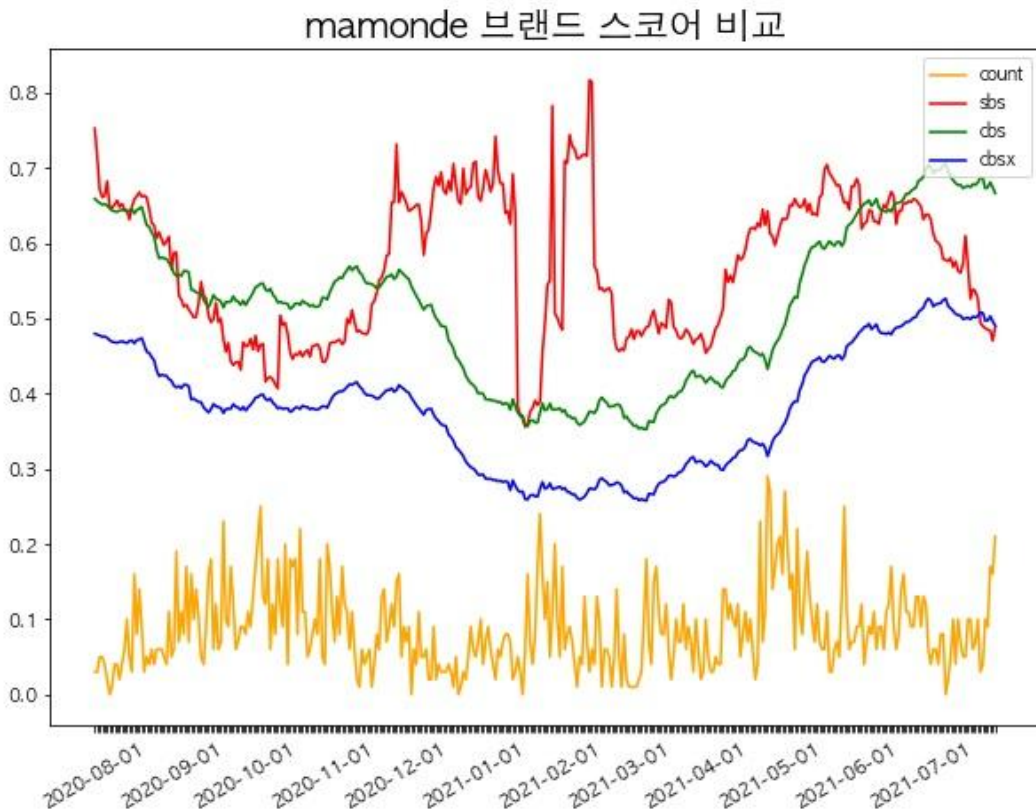
브랜드 점수 비교 (아모레)



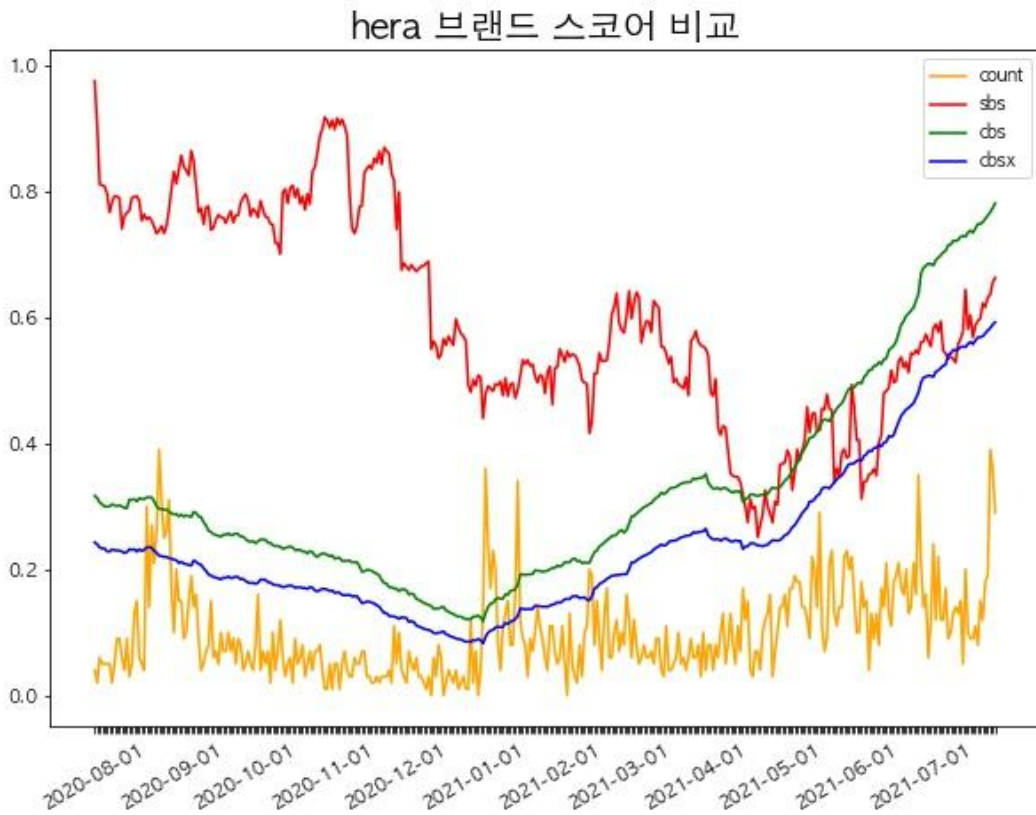
브랜드 점수 비교 (아이오펜)



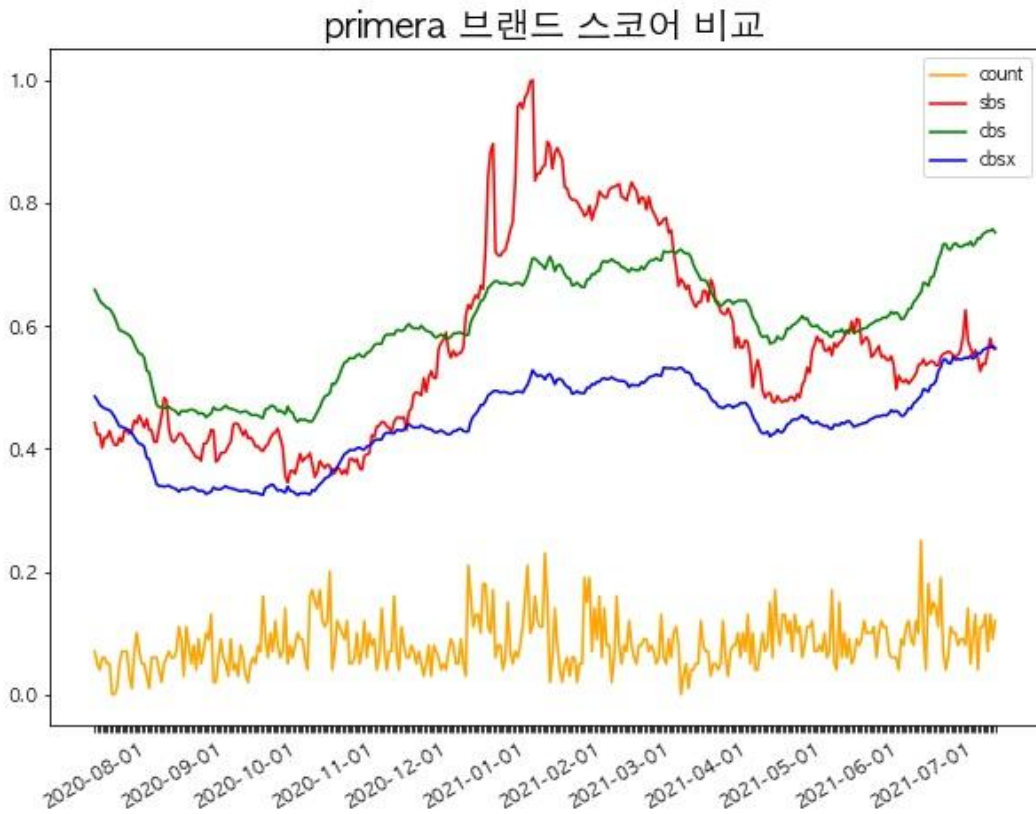
브랜드 점수 비교 (마몽드)



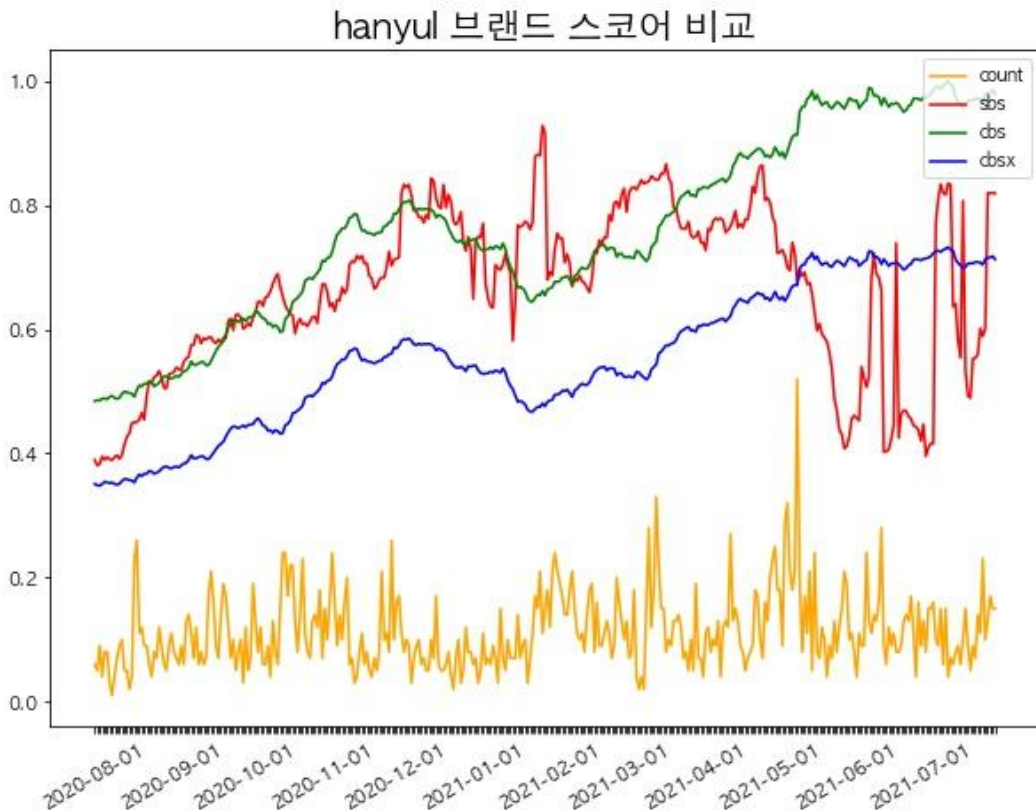
브랜드 점수 비교 (헤라)



브랜드 점수 비교 (프리메라)

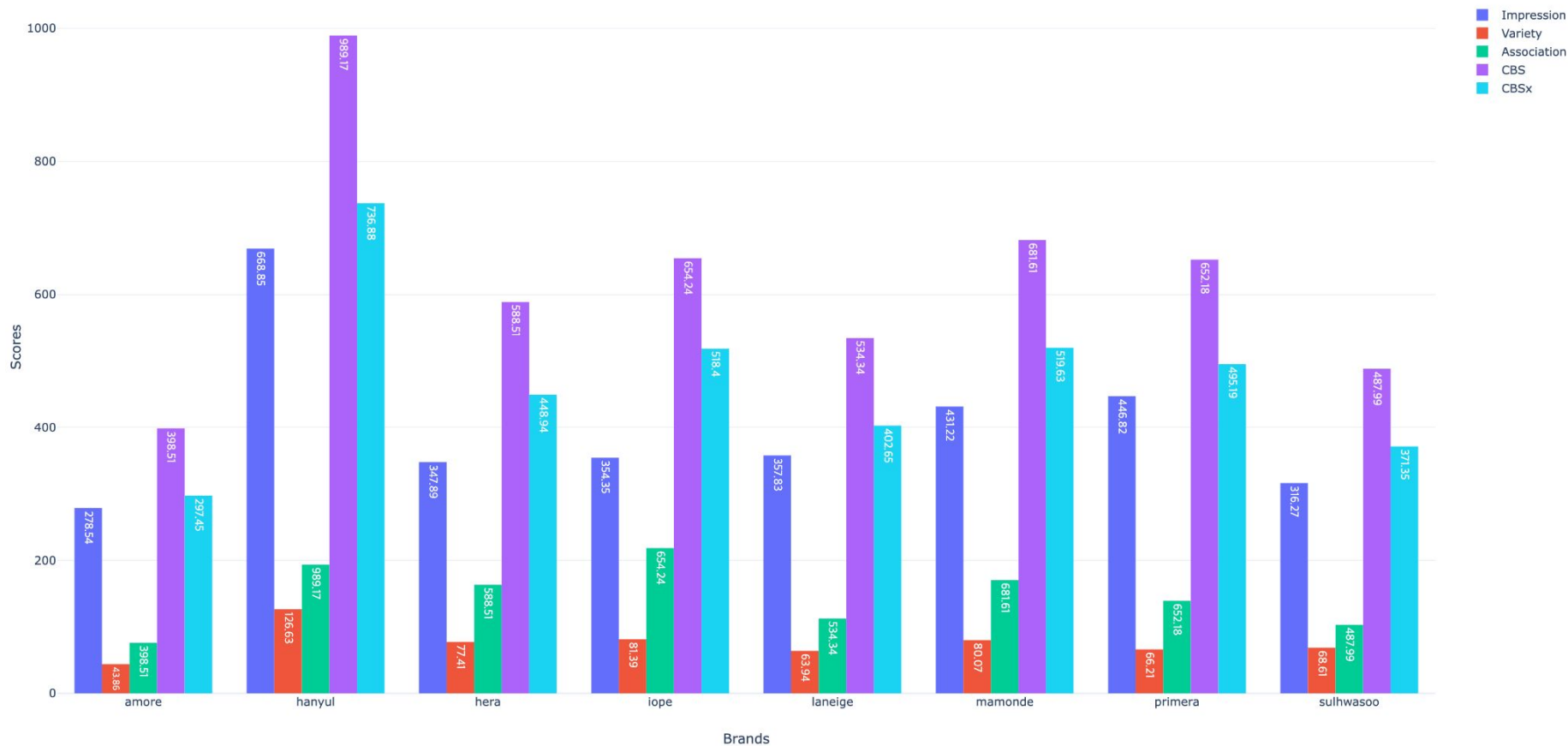


브랜드 점수 비교 (한울)

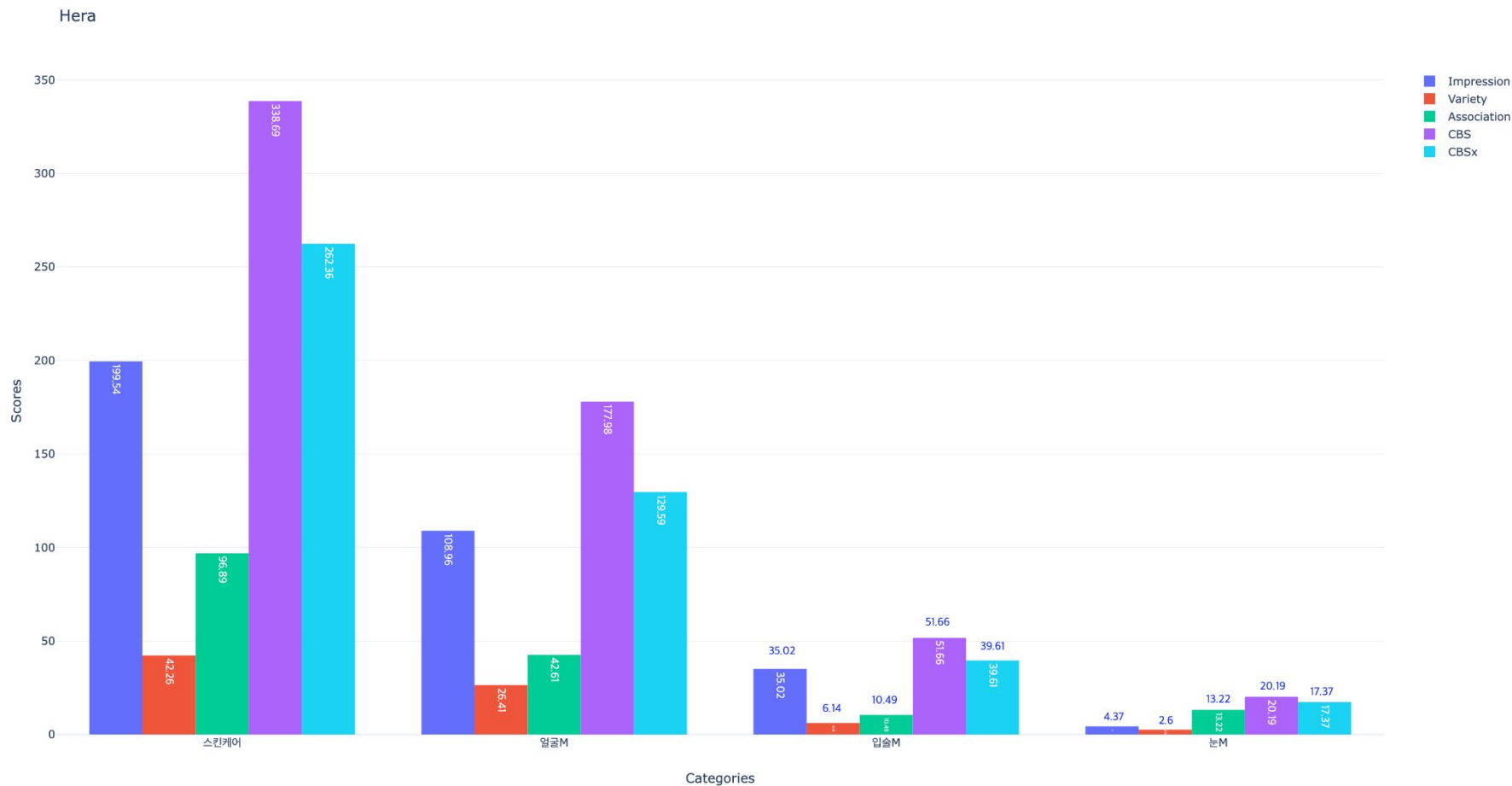


CBS & CBSx - 막대 그림 (모든 브랜드)

All Brands



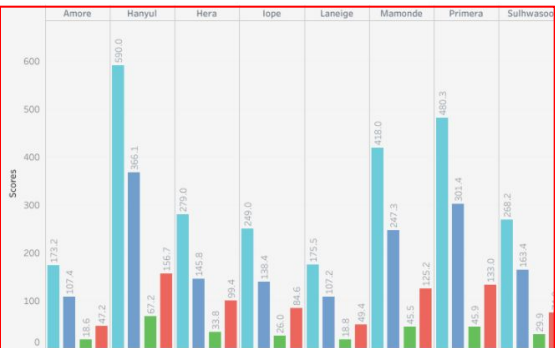
CBS & CBSx - 막대 그림 (헤라)



CBSx - 대시보드

Contextual Brand Score Extended (CBSx) Dashboard

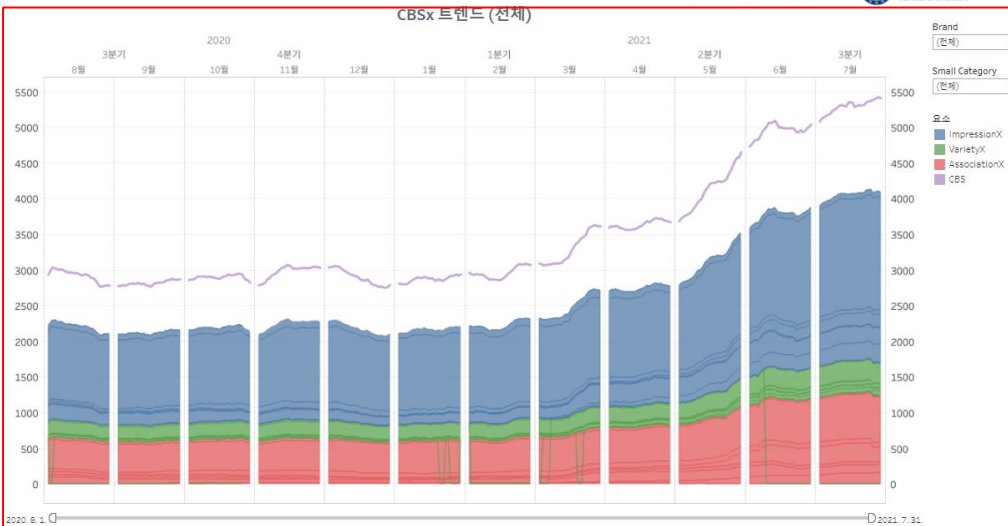
① 브랜드별 CBSx (Side-by-Side Bar Plot)



② 제품별 CBSx (Side-by-Side Bar Plot)



③ CBSx 트렌드



- * 공통사항: 관심기간 및 관심 브랜드를 선택
- * 마우스오버시 세부 사항 확인 가능
- * 각 측정시점의 값은 측정시점으로부터 과거 90일 동안의 리뷰 데이터를 사용해 산출

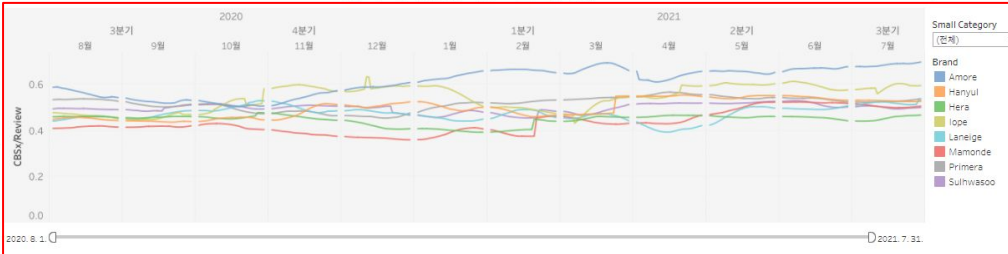
① 브랜드간 일평균 CBSx 값 및 각 요소의 절대적 크기를 비교

② 소제품군간 일평균 CBSx 절대적 크기를 비교

③ 일평균 CBSx의 요소별/소제품군별 트렌드를 표현함. 좌측 y축은 CBSx 값, 우측 y축은 리뷰수 대비 CBSx 비율

④ 리뷰 수 대비 CBSx 비율의 트렌드를 표현함. (30일 window에 대해 kernel smoothing 적용)

④ 리뷰 수 대비 CBSx 비율 트렌드



감사합니다!

AMORE PACIFIC
CORPORATION



연세대학교
YONSEI UNIVERSITY