



실제 라이브쇼핑 **댓글** 분석 예시를 통한 프로젝트 진행 **과정** 및 **견적**

2023 / 09

본 자료는 **엠포스 데이터전략실**에서 진행하는
<쇼핑라이브 댓글 분석 서비스>를 소개하기 위해 제작된 것으로,
실제 사례 분석을 통해, 제공 가능한 **분석 결과물**을 명시하였으며
한 달 기준 **진행 비용**도 기재되어 있으니, 편하게 문의 부탁드립니다.

CHAPTER.1

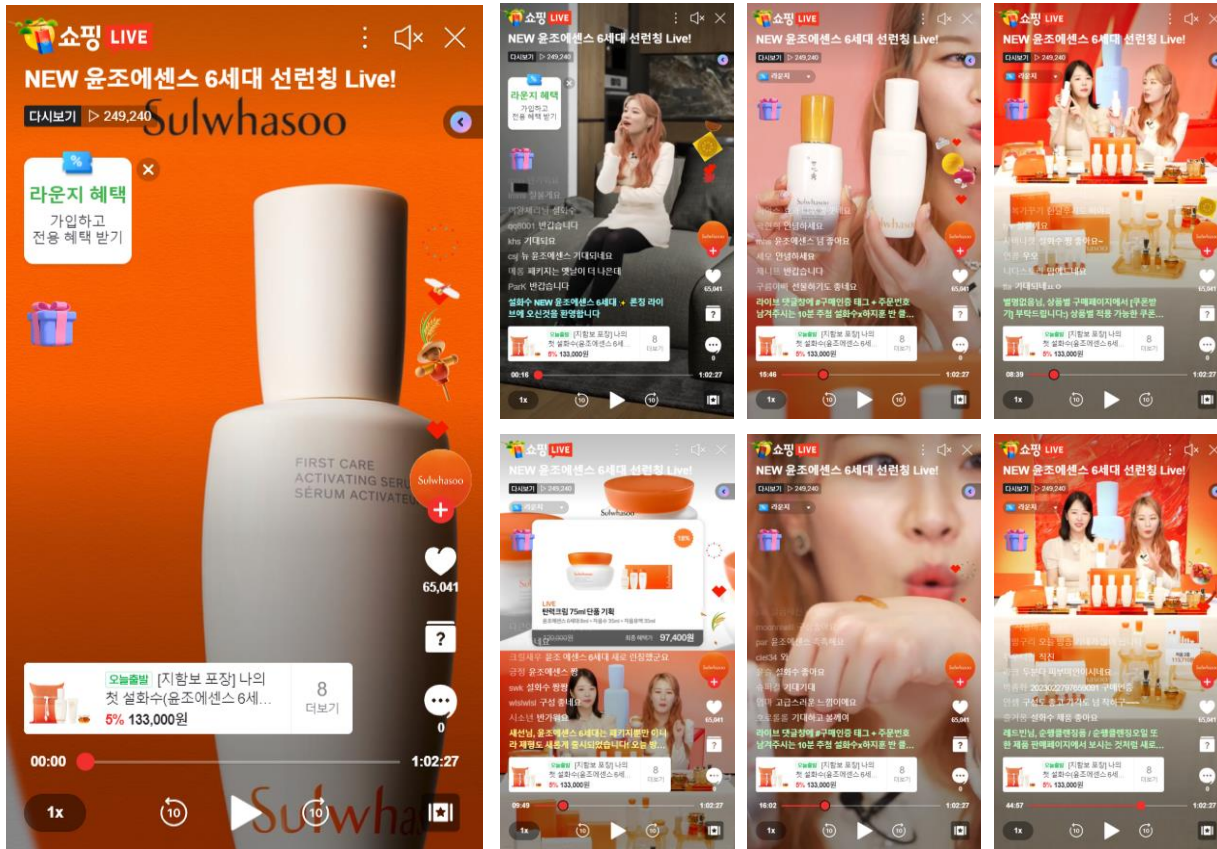
데이터 설계 기준 및 분석/분류 방법

댓글 분류 기준은 고객 KPI, 댓글 내용 등에 따라 상이하게 적용

데이터 설계

설화수 브랜드에서 진행한 라이브쇼핑 댓글 추출

- 파이썬 크롤링 코드를 활용하여 방송 중 생성된 댓글 **13,475개**를 수집
- 라이브방송 시청 인원은 **5,885명**이며, 『다시보기 수』는 **약 25만 회**로 집계



항목	내용
방송명	NEW 윤조에센스 6세대 선런칭 Live!
댓글 수	13,475 개
시청자 수	5,885 명 (Unique_ID 기준)
다시보기 수	249,241 회 (9/22 조회기준)
수집 방법	자체 개발한 파이썬 코드 활용

```

1 def collect_comments_realtime(self,
2   full_url = f"{self.URL_PREFIX}{self.URL_LAT}"
3   self.driver.get(full_url)
4
5   comments = [] # 실시간 댓글을 저장할 리스트

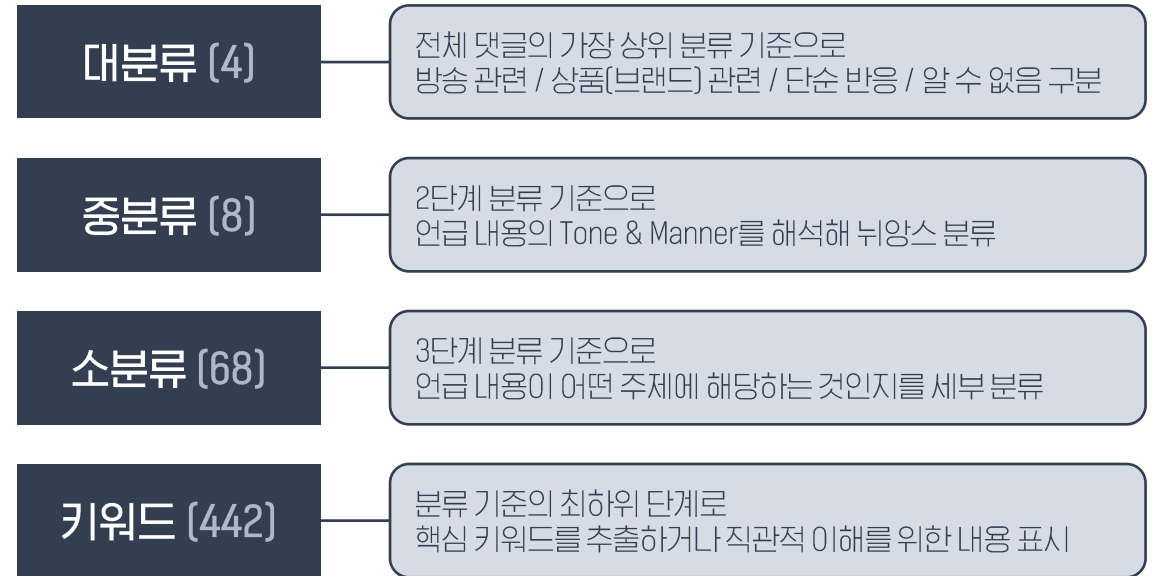
```

EXAMPLE

댓글 분류 기준

- 라이브쇼핑 댓글 분석은, **최대한 내용을 세분화하고 방영 의도나 산업 특성에 따라 정의하는 것이 중요!**
- 설화수 브랜드의 경우 단순 구매 인증 외에도 **제품에 대한 상세한 평가나, 선물 용도 등 구입 목적, 타겟 이미지, 심지어 진행자에 대한 호감** 등이 방송 호감도에 복합적으로 영향을 미친다고 판단하여, 총 4단계로 세밀하게 구분

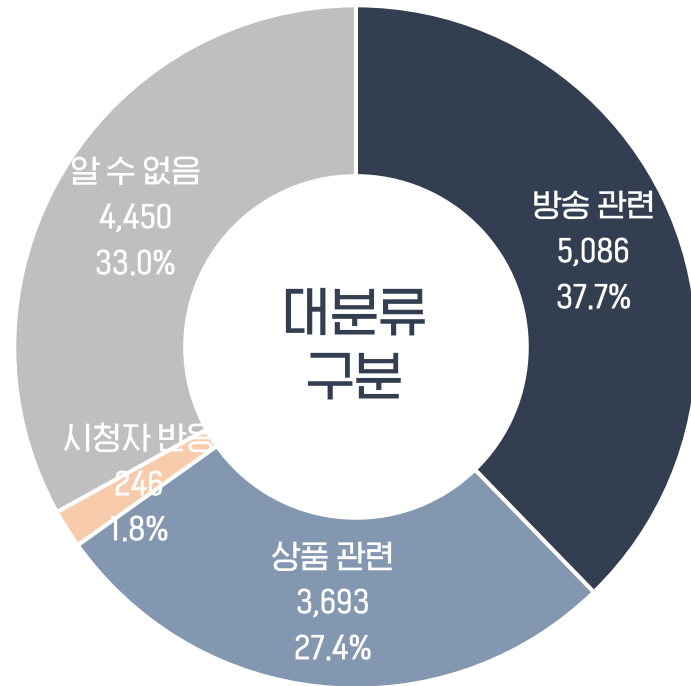
대분류 [4개]	중분류 [8개]	소분류 [68개]	키워드 [442개]
방송 관련	호감표시	단순 호감	방송
상품 관련	인사	참가인사	좋아요
시청자 반응	해석불가	상품 호감	기대
알 수 없음	단순언급	브랜드 호감	감탄
	질문	해석불가	해석불가
	불만	혜택 관련	구성
	요청	상품 구성	설화수 호감
	답변	구매인증	구매인증
		진행자 호감	혜택 좋아요
		가격 호감	가격 좋아요
		디자인	선물용
		단순 언급	에센스
		선물	도배
		방송 호감	설화수 언급
		제형	윤조에센스
		상품 특성	사고싶다
	



**데이터 분류 기준 및 체계는
상품 카테고리(산업)이나 언급 내용 및
뉘앙스 등에 따라 다양한 형태로 설계**

댓글 분류 기준 예시 : 대분류

- 통상 라이브쇼핑 댓글은 크게 '방송'에 대한 부분과 '상품'에 대한 부분으로 나뉘며 그 외 단순 리액션에 해당하는 부분이나 의미 해석이 어려운 부분도 존재



■ 방송 관련 ■ 상품 관련 ■ 시청자 반응 ■ 알 수 없음

no	대분류	설명	언급 예시
01	방송 관련	라이브 방송과 관련된 내용	“기다리던 방송이에요” / “라이브 기대됩니다.”
02	상품 관련	상품과 관련된 다양한 정보	“탄력크림 가격 좋네요” / “설화수 에센스 최고”
03	시청자 반응	소통과 관련된 내용	“오아 완전 예뻐요!” / “구매 각입니다.”
04	알 수 없음	도배, 이모티콘, 해석불가내용	“♡♡♡♡♡” / “ㅎㅎㅎ”

▶
현재는 이모티콘 등을 “알 수 없음”, “판단 불가”로 표현했지만 댓글을 통한 호응도를 판별하는 디지털 환경에서는 이모티콘이나 “ㅎㅎㅎ” 등 언급 내용 역시 중요 변수로 작용 가능하여 Data Tracking 과정에서 일부 조정이 필요할 수도 있을 것으로 판단

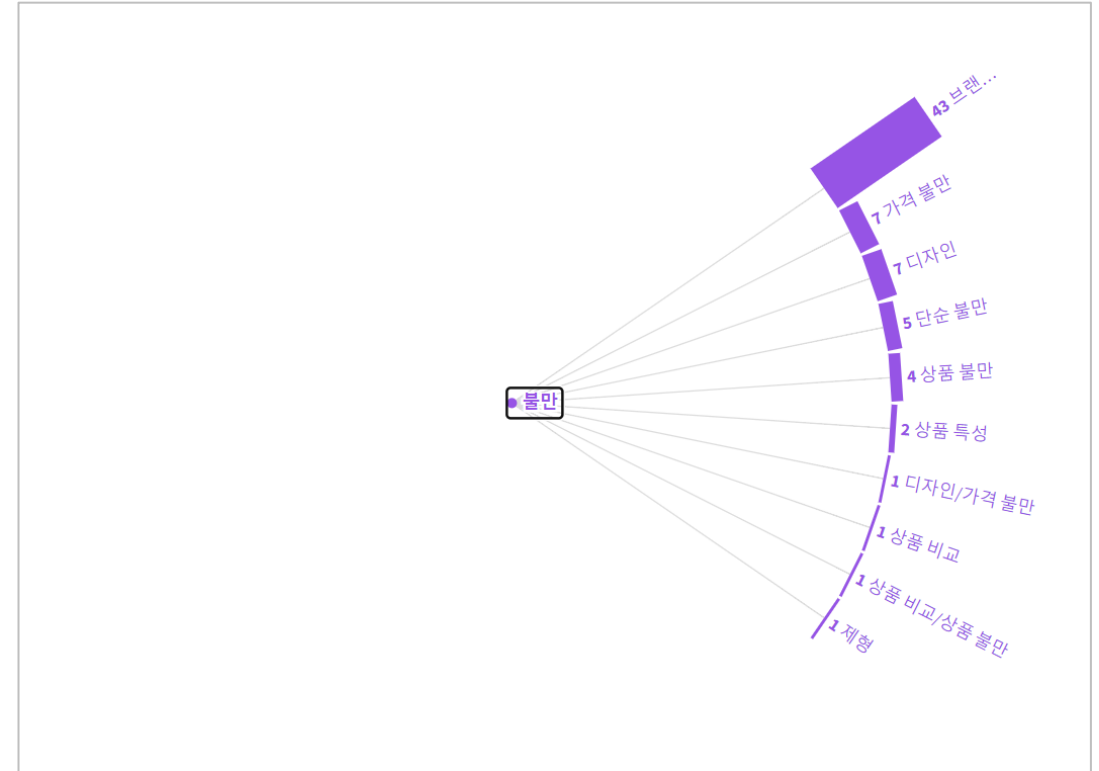
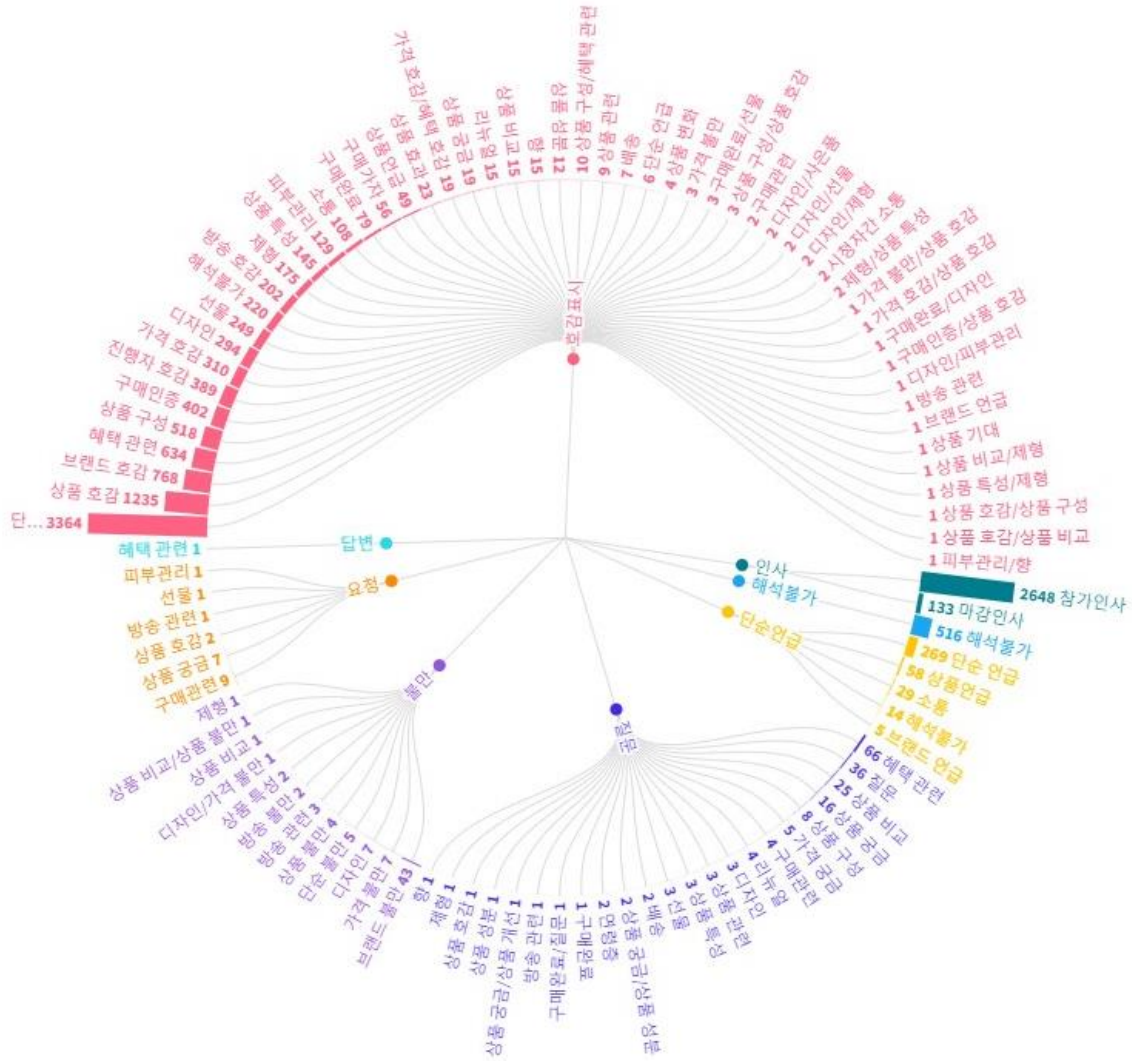
CHAPTER.2

댓글 분류 결과 시각화 예시

시각화는 예시로 제안된 방법 외에도 데이터 특성 등에 따라 다양하게 구현될 예정

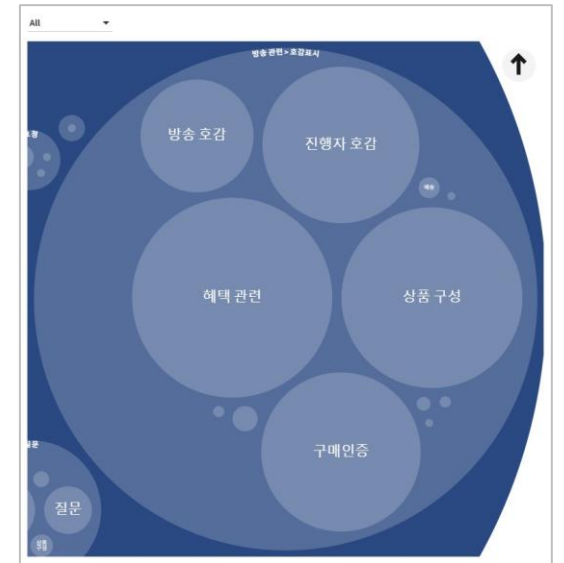
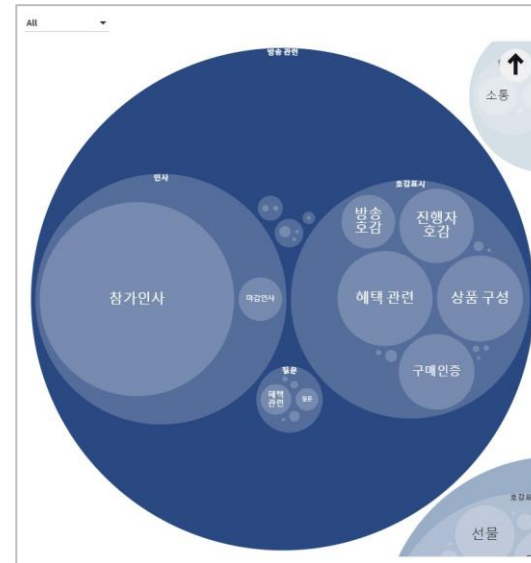
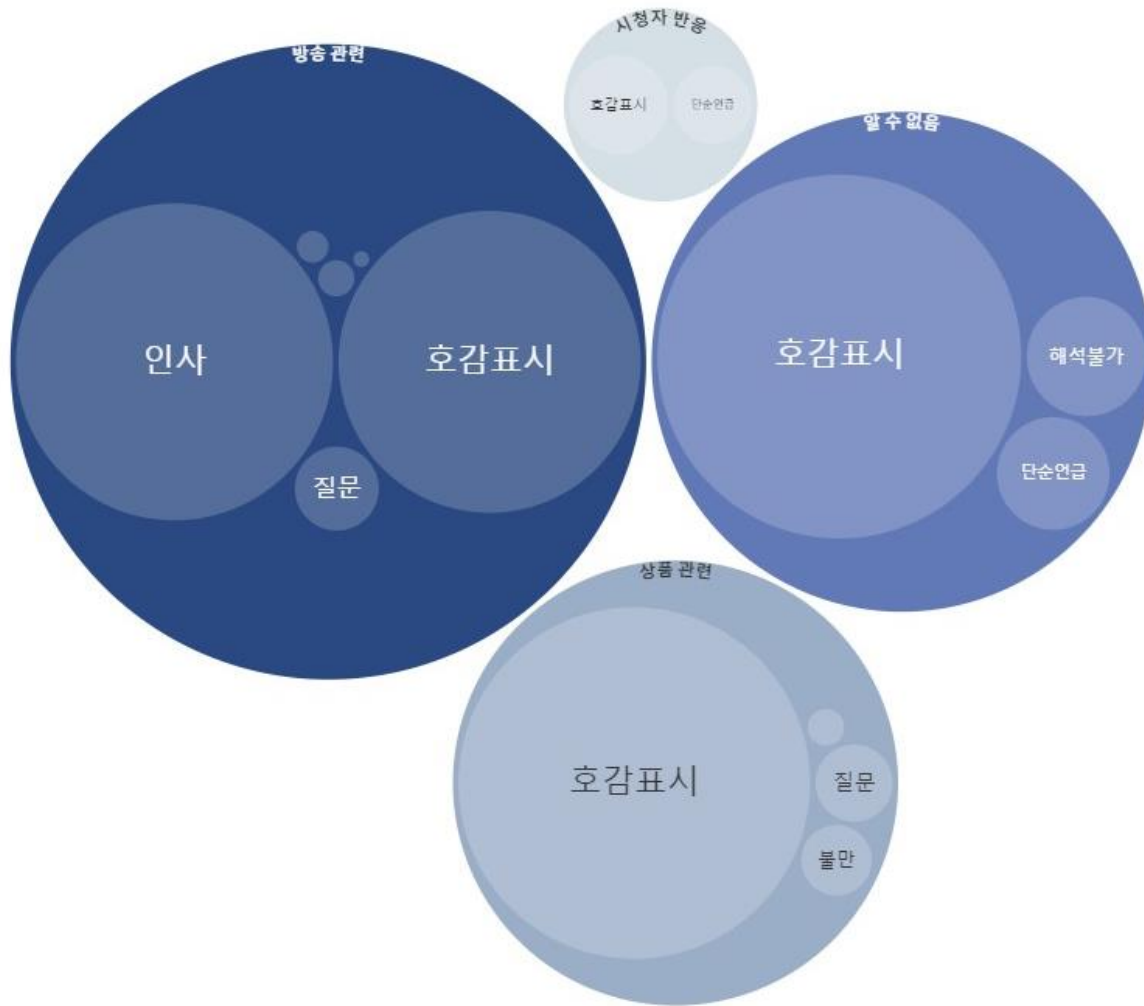
시각화 예시 01. Radial Tree

- 본 시각화에서는 중분류와 소분류 데이터만 활용한 결과
- 실제 구현된 모습은 → [이 부분을 클릭하여 이동](#) (마우스로 세부 영역을 클릭하면 상세하게 확인 가능)



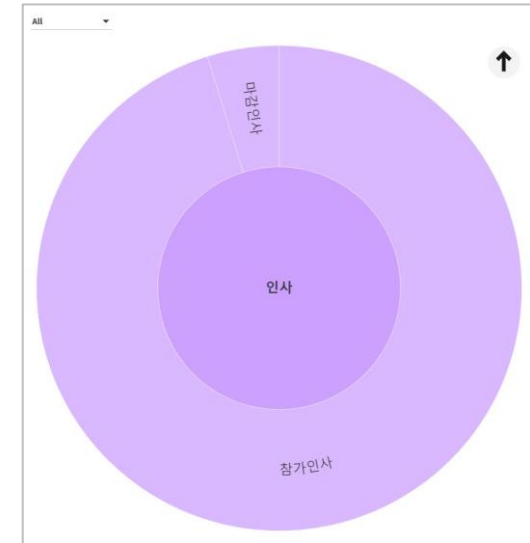
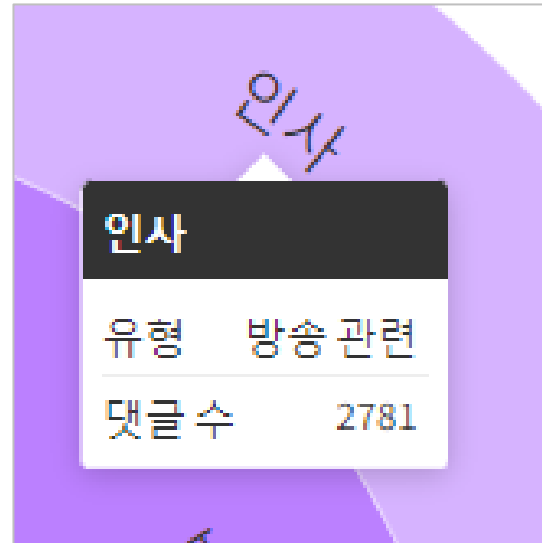
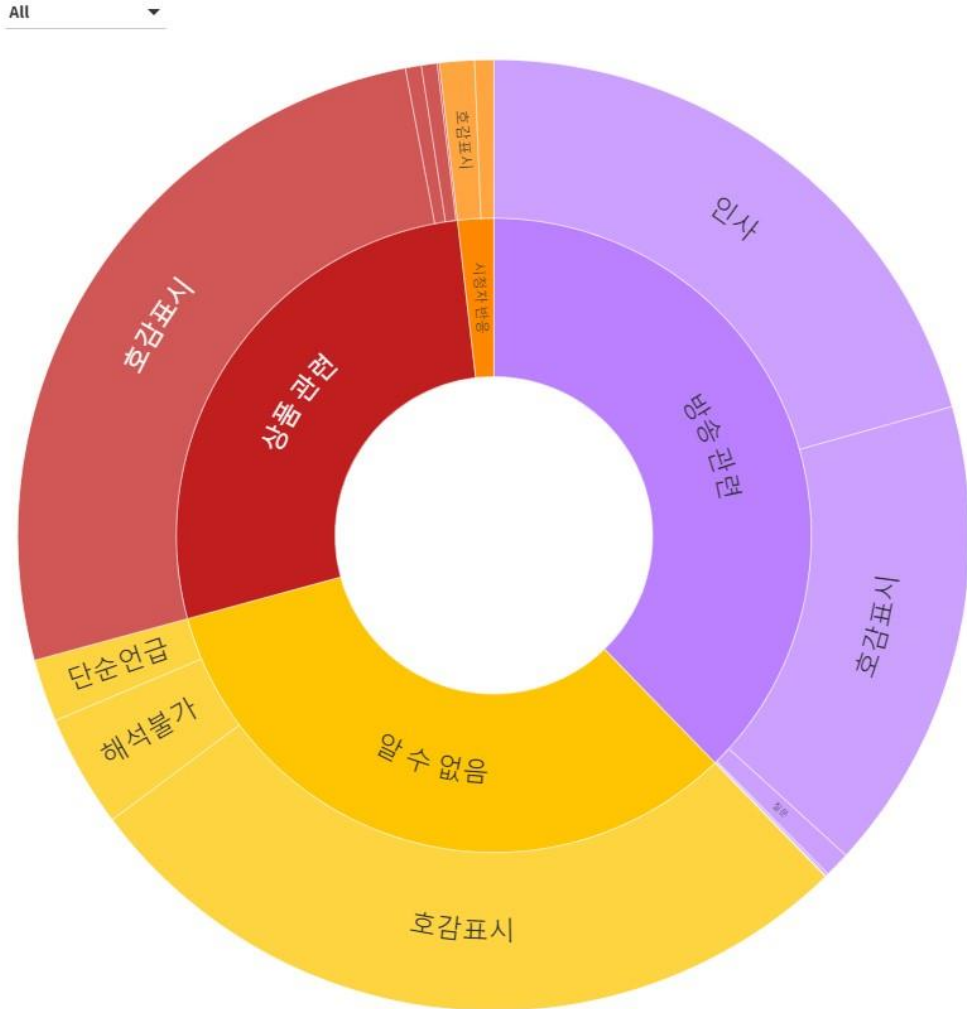
시각화 예시 02. Packed circles

- 본 시각화에서는 대분류를 포함하여 시각화
- 실제 구현된 모습은 → [이 부분을 클릭하여 이동](#) (마우스로 세부 영역을 클릭하면 상세하게 확인 가능)



시각화 예시 03. Sunburst

- 본 시각화에서는 대분류를 포함하여 시각화
- 실제 구현된 모습은 → [이 부분을 클릭하여 이동](#) (마우스로 세부 영역을 클릭하면 상세하게 확인 가능)



CHAPTER.3

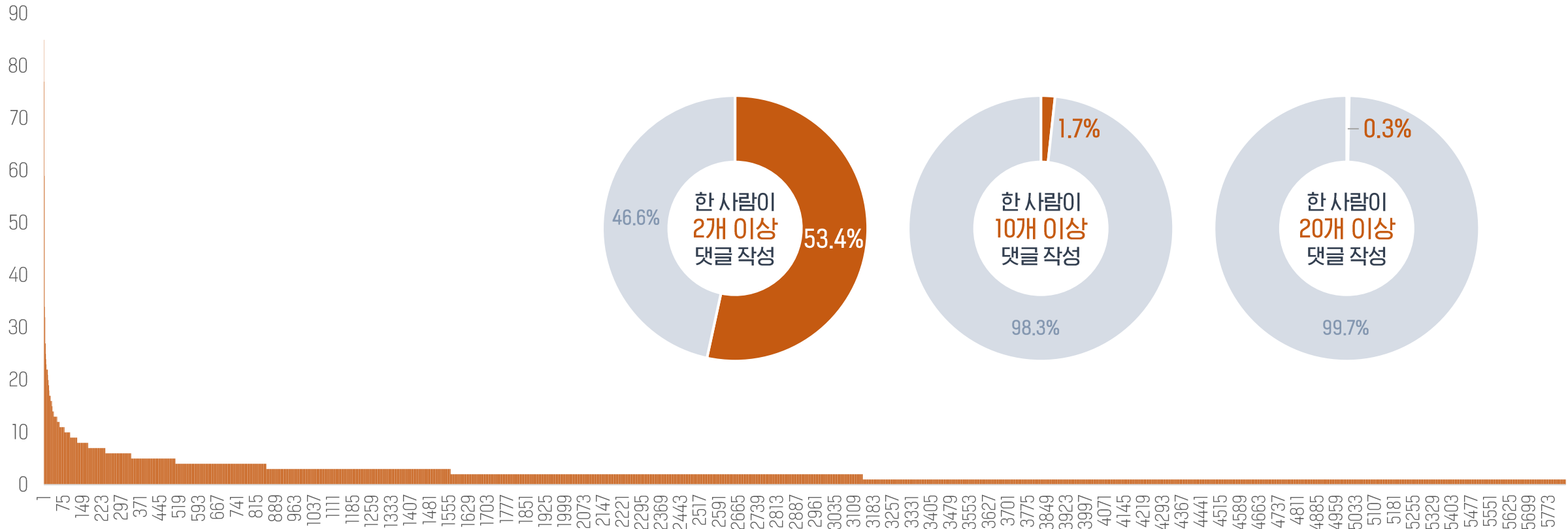
기타 분석 내용

댓글 내용 중 눈에 띄는 내용 추가 정리

댓글 작성 구조 파악

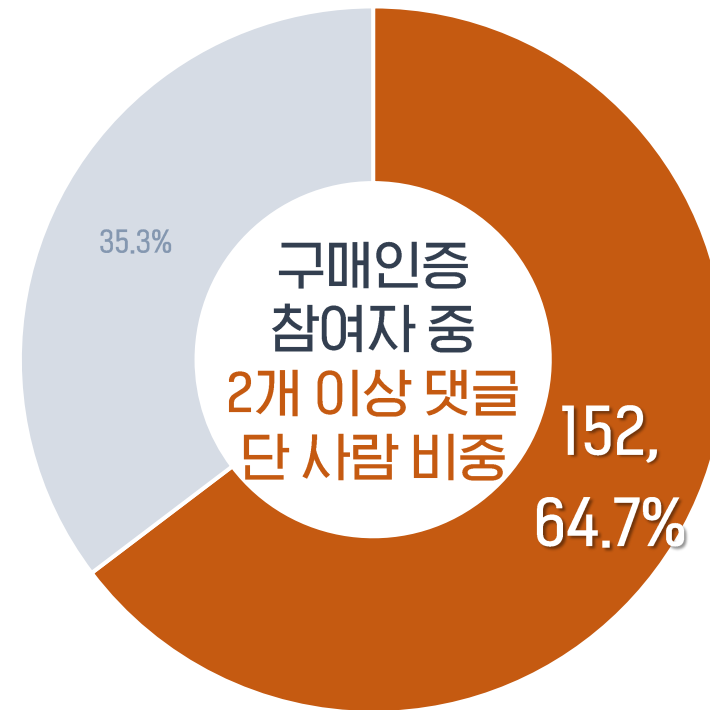
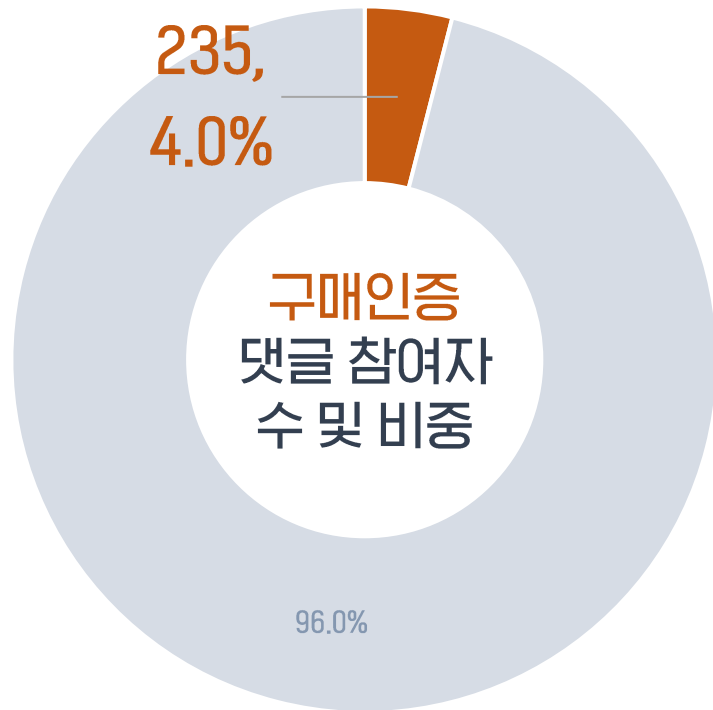
- 전체 참여자 5,885명 중, 인당 최대 댓글 수는 85개, 즉 가장 많이 댓글을 작성한 사람이 무려 85개를 달았다는 것이며 → 전체 기준, 인당 평균 2.30개 댓글 작성
- 전체 중 댓글을 2개 이상 작성한 참여자 비중은 53.6%였으며 10개 이상 작성한 참여자 비중은 1.7%, 20개 이상 작성한 참여자 비중은 0.3%로 집계
- 10개 이상의 댓글을 작성한 참여자 비중이 1.7%로 매우 적어 보이지만 인원 규모로 보면 100명 정도로, 라이브 방송 중 활발한 댓글 조장을 위한 인원으로 생각하면 적지 않다고 판단

ID별 작성 댓글 수 분포



구매인증 댓글 참여자 비중

- 구매인증을 한 참여자는 235명으로 전체 중 4%에 해당하며 이들 중 2개 이상의 댓글을 작성한 참여자의 비중은 64.7%(152명)에 해당
- 해당 데이터는 라이브방송 시 구매 건수 데이터와 함께 비교해 보는 것이 필요



구매인증을 한 시청자 중 2개 이상의 댓글을 작성한 152명에 대한 상세 키워드 분석

- 대분류, 중분류의 경우 상위 5개의 분류만 기입
구매인증 댓글이 주를 이루기 때문에 상위 5개의 키워드 제외
- 혜택 관련 질문과 엄마를 위한 선물 주문으로 꼼꼼한 확인을 기반한 질문이 주를 이룸



- ✓ 언급 내용 기반 상세 추출 키워드를 언급 빈도에 따라 WordCloud 형태로 시각화
- ✓ 설화수, 윤조, 구매, 인증, 에센스 등 상품 및 브랜드, 이벤트 참여 등의 키워드는 제외

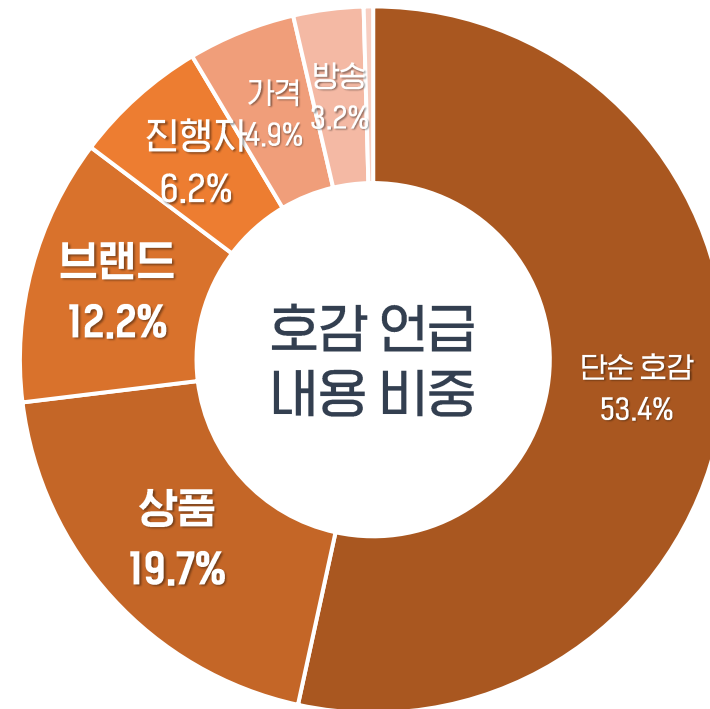
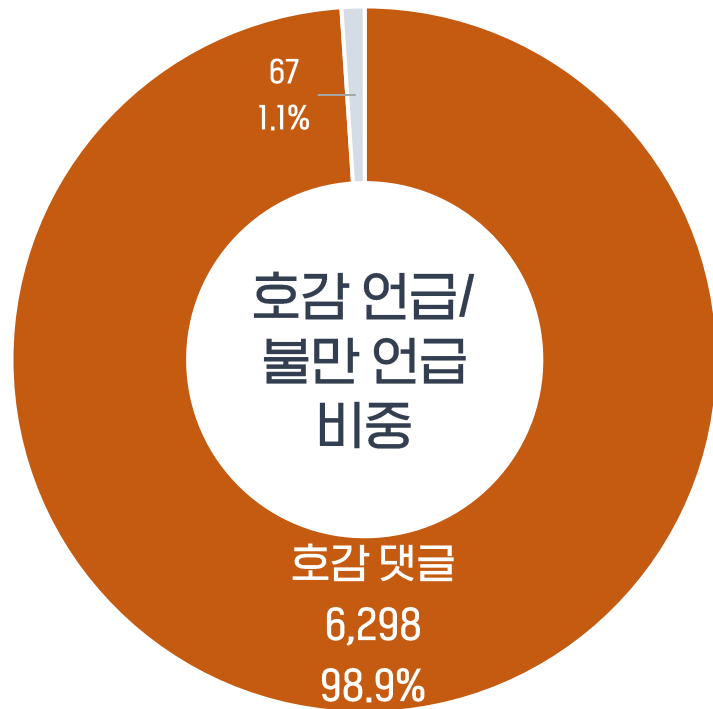
구매인증을 하지 않고 댓글을 작성한 참여자에 대한 상세 키워드 분석

- 상품에 대한 단순 언급이나 인사, 도배글을 제외하고 확인해보면
방송을 시청함으로써 얻을 수 있는 혜택, 가격 구성에 대해 높은 관심 표출



구매인증 댓글 참여자 비중

- 호감을 포함한 댓글 수 : 6,298개 / 불만을 포함한 댓글 수 : 67개 를 중심으로 비중 산정
- 상세 호감 내용을 보면 “좋아요~” 등 단순 호감을 표현한 비중이 53.4%로 가장 많지만 그 외 상품과 브랜드에 대한 호감도를 합치면 31.9%로 적지 않은 수준
- 또한 화장품 특성 상 진행자에 대한 호감이 적지 않게 집계된 것이 눈에 띄며 실제 내용을 보면 진행자에 대한 호감으로 방송을 챙겨본다는 의견도 적지 않게 형성



라이브방송에 달리는 댓글은 대부분 단순언급이 많다.
안녕하세요, 좋아요, 구매인증, ㅎㅎㅎㅎ, 등등

그래서 겉으로 보면, 대부분 이벤트 당첨만을 위한 행동 같은데
내용을 쪼개다 보면 상품에 대한 평가, 브랜드에 대한 인식도 적지 않다.

물론, 겨우 댓글로 달린 몇몇 내용을 가지고
시장을 진단하고 위상을 평가하는 건 부적절하다. 그건 과하다.

다만 경험상, 자발적으로 기재된 소비자의 언어에는 언제나 **함의(含意)**가 있다.

텍스트 같은 비정형 데이터일수록 거시적인 표면의 Wants보다
미래 가치가 있는 **미시적인 내면의 Needs**에 주목할 필요가 있다.

쇼핑라이브 댓글을 분석해야 하는 이유는 간단하다.

1. 우리는 원래 우리가 가진 **모든 데이터를 분석해야** 한다.
2. 방송의 성과를 평가하기 위한 **KPI 자료로 활용 가능**하다.
3. 소비자 반응을 모니터링하면 **개선 아이디어**를 얻을 수 있다.

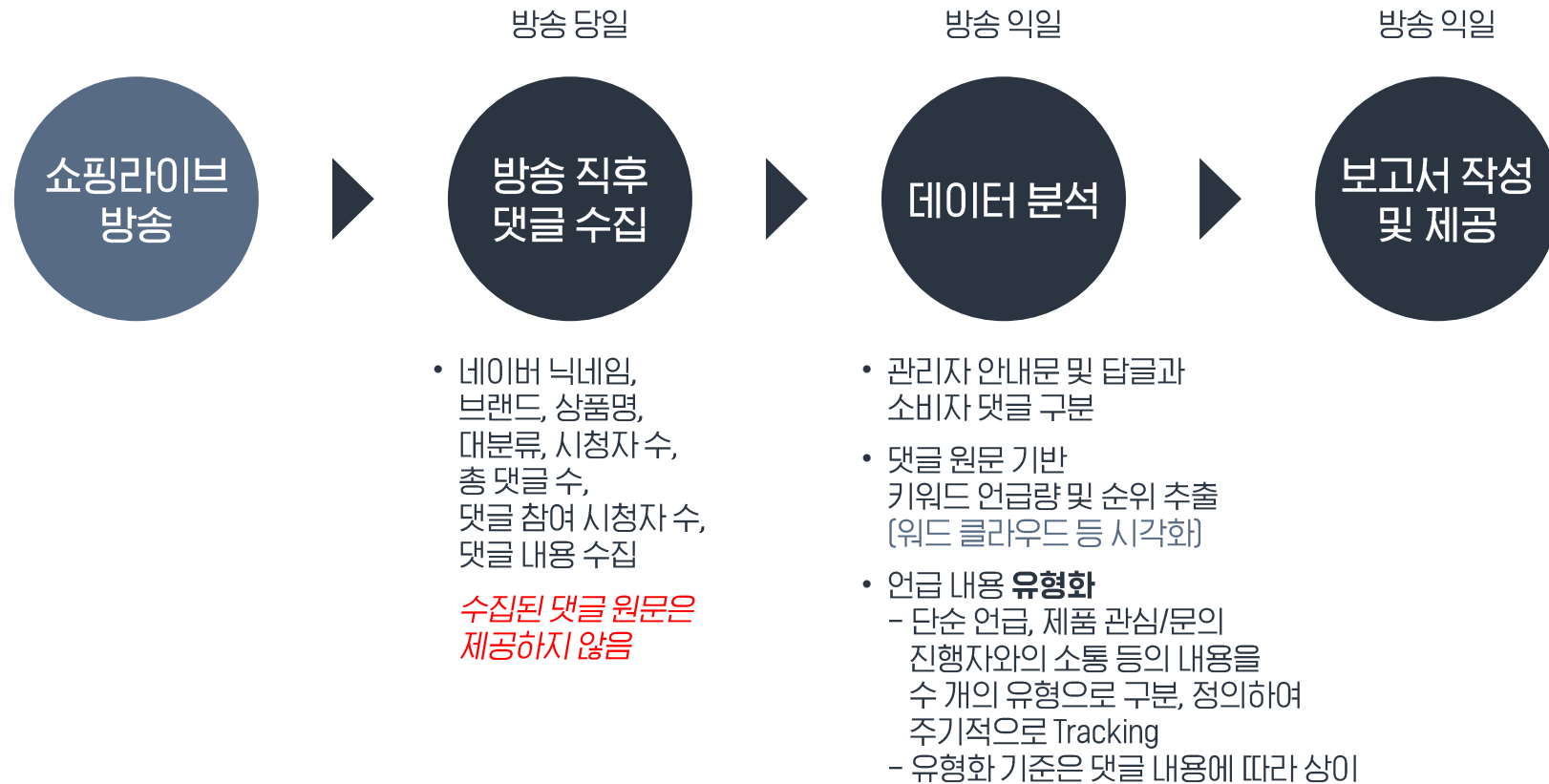
CHAPTER.4

댓글 분석 의뢰 비용 및 조건

일별 데이터 분석 및 보고서 제공 가능

쇼핑라이브 댓글 분석, 업무 진행 과정

- 라이브 방송 직후 1~2일만에 댓글 내용 분석 리포트 제공 (단, 초기 분류 기준 세팅시 2~3일 소요)
- 수집된 댓글 원문 파일은 제공하지 않으며, 댓글 분류 기준과 리포트 제공 양식, 전달 시기 등은 의뢰 기업과 협의하여 결정



쇼핑라이브 댓글 분석, 제안 견적

- 1개월 기준 견적, **일백만 원(1,000,000 원 / VAT별도)**로 제안
 ※ 본 견적은 데이터 연구를 위한 협업 차원에서 설정된 초기 견적으로,
 프로젝트 의뢰 기업이 일정 수준 이상 많아지면 마감될 수 있습니다.
- 문의 : 엠포스 데이터전략실, 박경하 실장 [02-6177-1871] / khbak@emforce.co.kr

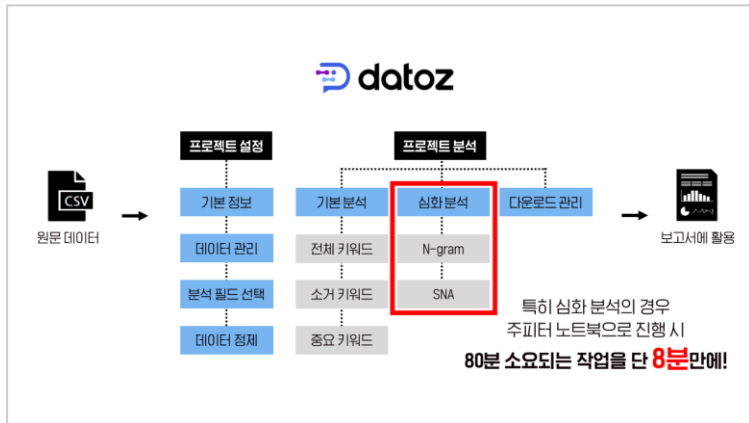
✓ **개인사업자, SOHO**를 운영자 및 **단건 분석**은 별도 문의 바랍니다.

업무 범위 및 견적 산정 조건	1개월 기준 견적
<ul style="list-style-type: none"> • 업무 범위 : 1일 1건 방송되는 콘텐츠에 대한 댓글 수집과 분석, 보고서 작성 <ul style="list-style-type: none"> - 한 달 평균 라이브방송 횟수 20개 기준 [분석 대상 방송 추가는, 별도 협의] - 데이터 분석 및 보고서 작성은 평균 2일 내외 소요 [초기 세팅 2~3일 예상] • 1건 당 5만원 X 20개 콘텐츠 분량 고려하여 견적 산출 • 1개월 단위 계약 가능 	<p>일금 일백만 원 (1,000,000 원 / VAT별도)</p>

추가 제공 가능 서비스

- 방송 직후부터 분석 보고서가 제공되는 1~2일차 기간 사이에 미리 기본 분석 결과를 확인할 수 있도록 수집된 원문을 '자체 개발 솔루션'에 탑재해 제공 가능
- 활용 솔루션 : 텍스트 데이터 분석을 위해 엠포스에서 자체 개발한 『데이토즈 (Datoz)』, (www.datoz.net)
- 확인 가능 내용 : 언급 키워드 종류, 키워드 언급량 기준 순위, 키워드 관계 등

데이터의 인사이트를 알려줄께! 데이토즈

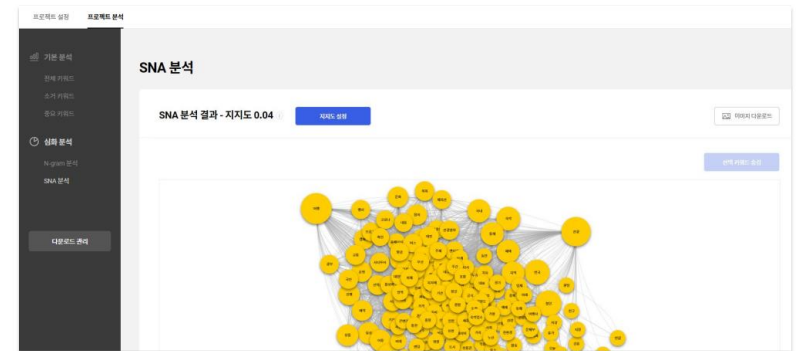


04 기본 분석하기 > 전체 키워드



연관어 맵에서는 빈도 수 상위 50개의 상단의 컨트롤 바를 통해 워드 클라우드 키워드를 클릭하면 해당 키워드를 클릭하면 해당 키워드를 클릭하면 해당 키워드를 클릭하면 해당 키워드를 클릭하면

05 심화 분석하기 > SNA



<SNA 분석>은 빈번히 언급된 키워드 간의 관계를 Line으로 표현한 것으로, 키워드 노드를 클릭하여 관계망을 확인하거나 [선택 키워드 숨김] 처리할 수 있습니다. 또한 키워드를 드래그하며 위치를 바꾸어 그래프 커스터마이징을 할 수 있습니다.

End Of Document

Contact Us

Website URL <http://bigdata.emforce.co.kr>

T 02. 6177. 1871 eMAIL khbak@emforce.co.kr