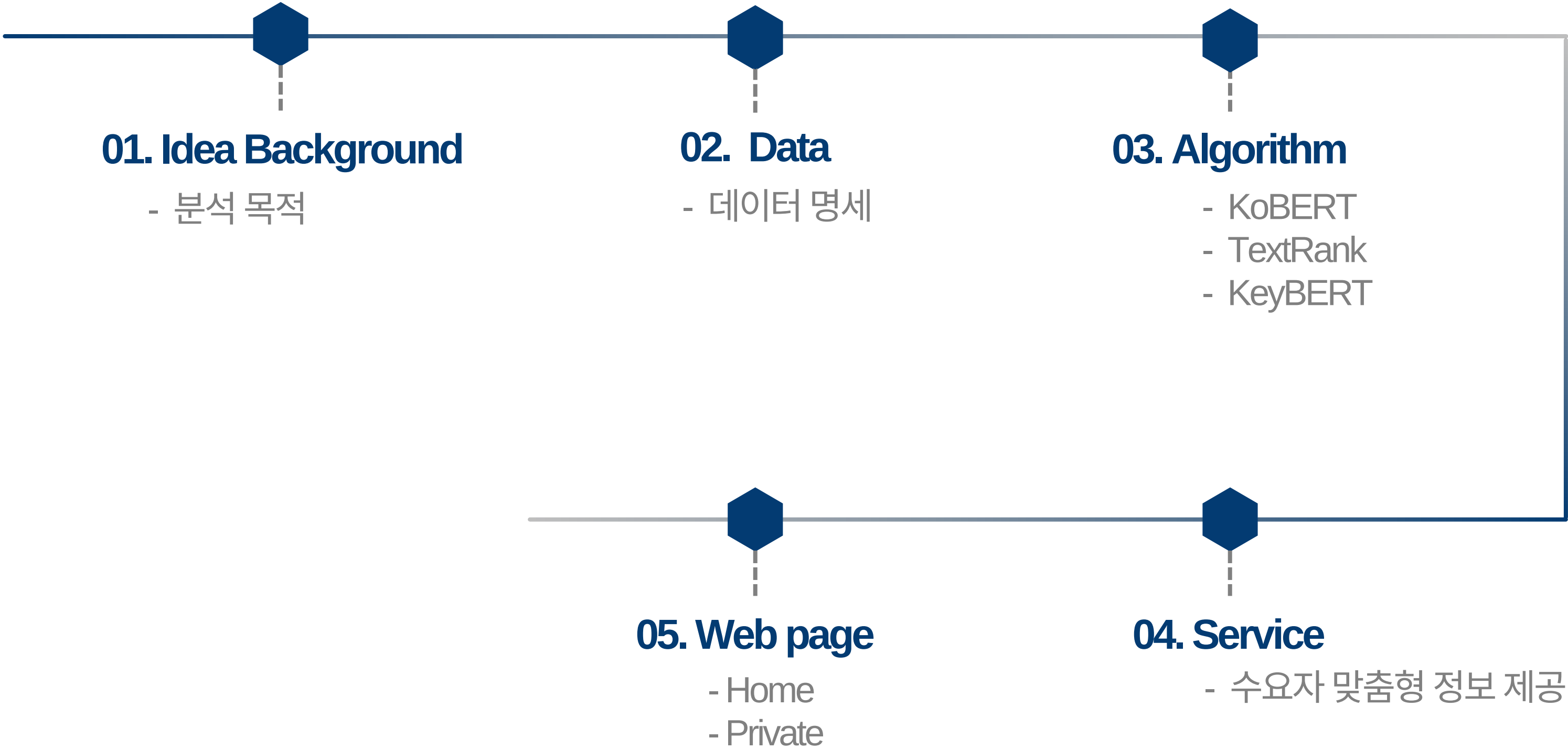


| 수요자 맞춤형 한국소비자원 웹사이트 구현

권동구 김기남 유태권





한국 소비자원 실증분석 목적

고객 결제 패턴 분석을 통한 소비자 맞춤형 뉴스 추천서비스

제5회 대구 빅데이터 분석 경진대회

소비자 부문 추가 모집

2023 6.7.(수) - 6.25.(일)

신청 방법
■ bigdata01@dip.or.kr (이메일 접수)

참가 대상
■ 데이터 분석 및 활용 분야
일반인·대학생 참여 (개인 또는 3인 이내 팀)
■ 데이터 시각화 분야
중·고등학생 참여 (지도교사 포함 2인 또는 4인 이내 팀)

공모 과제
■ 데이터 분석 및 활용 분야
금융 부문 | 빅데이터를 활용한 금융 등 자유주제
금융 부문 | 대구·경북 지역 내 소상공인 지원
ex1) 소상공인 지원 정책 효과 분석
ex2) 소상공인 지원 정책 효과 분석

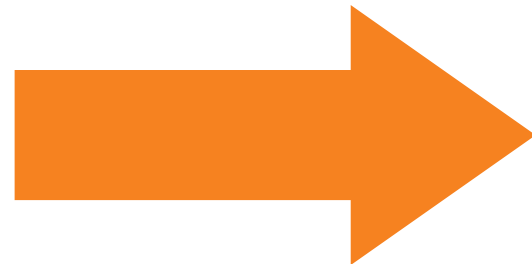
소비자 부문 | 소비자 관련 데이터를 활용한 자유주제
ex1) 디지털 헬스케어 제품 등 신산업 분야 및 온라인몰 등을 시장 내 소비자 권리 보호
ex2) 소비자 이슈 탐지 모델 개발

데이터 시각화 분야
"일기 좋은 대구"를 주제로
- 빅데이터를 활용한 대구의
시각화 결과를 소개

시상내역
■ 19명(팀) 선정, 총 상금 3,450만원

구분	상 등	시 상 수	상 금
금융 부문	최우수상 (한국경제신문 후원)	1명	300만원
	우수상 (한국경제신문 후원)	1명	200만원
	장려상 (한국경제신문 후원)	1명	100만원
금융 부문	최우수상 (대구경제신문 후원)	1명	300만원
	우수상 (대구경제신문 후원)	1명	200만원
	장려상 (대구경제신문 후원)	1명	100만원
데이터 분석 및 활용 분야	최우수상 (한국소비자원 후원)	1명	300만원
	우수상 (한국소비자원 후원)	1명	200만원
	장려상 (한국소비자원 후원)	1명	100만원
데이터 시각화 분야	최우수상 (한국소비자원 후원)	1명	300만원
	우수상 (한국소비자원 후원)	1명	200만원
	장려상 (한국소비자원 후원)	1명	100만원

문의처
■ (재)대구디지털혁신진흥원 경진대회 담당자
T. 053-951-8977 E. bigdata01@dip.or.kr



수요자 맞춤형 한국소비자원 웹사이트 구현

HOME PRIVATE

한국소비자원
Korea Consumer Agency

소비생활에 가치와 신뢰를 더하여
국민의 삶의 질 향상에 기여합니다.

보도자료
"편의점 도시락, 고가·반찬 양 제품별로 차이 커"
: 나트륨은 1일 영양성분 기준 대비 55~86%로 한 끼로는 많은 편
더 많은 정보를 알아보세요.

피해예방주의보
"휴대용 빙프로젝터, 영상·음향품질 등 주요 성능 차이 있어"
: 발기·배터리 사용시간은 표시개선 필요
더 많은 정보를 알아보세요.

소비자안전주의보
"알상 회복에 따라 어린이 안전사고, 전년 대비 36.4% 증가"
: 가정 밖 안전사고 높고 영아기엔 '추락', 걸음마기·유아기는 '이물 삼킴' 사고 많아
더 많은 정보를 알아보세요.

상담다발품목

품목	피해유형	거래유형
대형승용차 동차	품질(물품/용역)	일반판매
악곡	계약해제·해지/위약금	일반판매
신유행상품권	약관	기타
헬스장	계약해제·해지/위약금	일반판매

국내리콜정보

제품명	사업자명	공표일
하루농	(주) 씨앗푸드	~24.11.22

소비자24로 바로가기

더 많은 정보를 알아보세요.

소비자 단어
"체리 피커 (cherry picker)"
: 기업의 상품이나 서비스를 구매하지 않으면서 자신의 실속을 자라기에만 관심을 두고 있는 소비자
네이버 지식백과로 바로가기

한국 소비자원이 생산하는 소비자 정보를 수요자 맞춤으로 제공하는 웹사이트 구현

데이터 명세

대구은행 BC카드 결제 데이터

◎ 주요 자료항목
[데이터 경진대회]

컬럼ID	컬럼명	타입	비고
승인월	승인월	VARCHAR	년월
고객번호	고객번호	VARCHAR	고객번호
가맹점업종	가맹점업종	VARCHAR	업종코드
가맹점명	가맹점명	VARCHAR	가맹점명
가맹점 시도	가맹점 시도	VARCHAR	시도명
승인금액	승인금액	NUMBER	금액
승인건수	승인건수	NUMBER	건수

데이터 명세

데이터명	형식	대상기간	사용변수	출처	비고
대구은행 BC카드 결제 데이터	CSV	2022.12.01 ~ 2022.12.31	고객ID, 가맹점업종코드, 가맹점업종명, 승인금액, 승인건수	대구은행	-
BC카드 가맹점 업종코드	CSV	-	가맹점업종, 가맹점분류1, 가맹점분류2, 가맹점업종명	대구은행	-
네이버뉴스 크롤링 학습데이터	CSV	2018.06 ~ 2023.05	url, title, content, label	네이버 뉴스	https://n.news.naver.com/
한국 소비자원 보도자료	CSV	2022.07 ~ 2023.11	title, content	한국 소비자원	https://www.kca.go.kr/home/main.do

데이터 명세

한국소비자원 소비자 뉴스

소비자뉴스	
보도자료	—
소비자원동향	+
공지사항	+
카드뉴스	+
피해예방주의보	+
소비자안전주의보	+

소비자24 국내 리콜

소비자24
24시간 열린 소비자포털

상품·안전정보

☞ > 상품·안전정보 > 국내리콜

국내리콜

전체


공산품

생활방사선제품

식품

축산물

의약품




11007. 큰맘해장국

사업자명 | (주) 포두미트

리콜공표일 | ~ 2024.11.12

출처 | 식품의약품안전처



11006. 냉동홍고추

사업자명 | 한성 글로벌

리콜공표일 | ~ 2024.12.14

출처 | 식품의약품안전처

소비자 상담 데이터

E	F	G	H	I	J	K
SEX_CD	SEX	AGE_CD	AGE	AREA_CD	AREA	TOP_CLS
1	남성	06	51	1403	경상북도구미시	스포츠·레저·취미용품
2	여성	04	33	1410	경상북도포항시	문화·오락서비스
1	남성	03	28	1410	경상북도포항시	교육서비스
1	남성	04	33	0302	대구광역시달서구	스포츠·레저·취미용품
1	남성	03	1	1403	경상북도구미시	운수·보관·관리서비스
1	남성	04	38	0308	대구광역시달성군	문화·오락서비스

- 성별
- 연령
- 상담 품목

알고리즘

KoBERT → 뉴스 분류 모델

뉴스 Feature Vector

Fine-Tuning

Task에 적합한 다중 분류 모델을 얻기 위해 Fine-Tuning

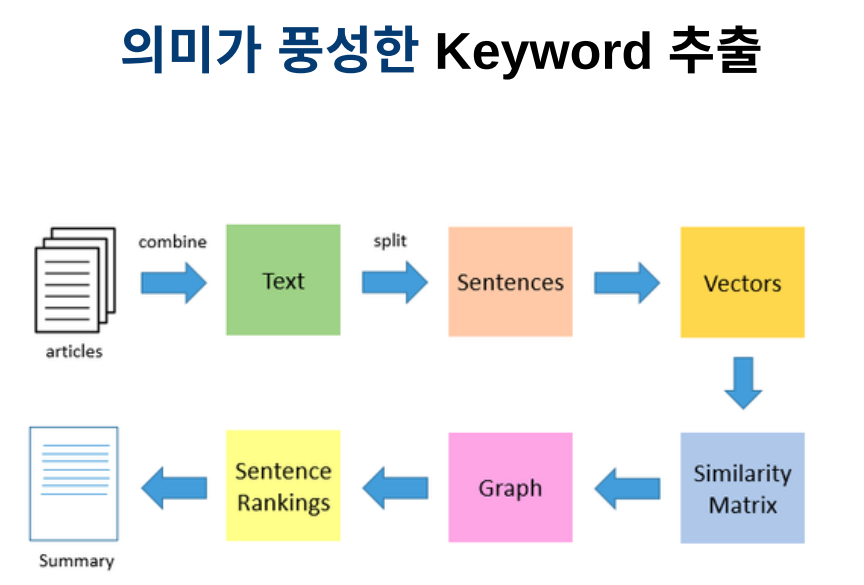
뉴스 Feature Vector

TextRank → 요약

추출적 요약
(Extractive Summarization)

사실을 정확하고 공정하게 전달해야 하는 뉴스의 특성 고려

KeyBERT → 키워드 추출



코사인 유사도 → 추천 알고리즘

코사인 유사도
(Cosine Similarity)

벡터의 크기가 아닌 방향의 유사도를 측정하는 방법

$$similarity = \frac{A \cdot B}{||A|| ||B||}$$
$$= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

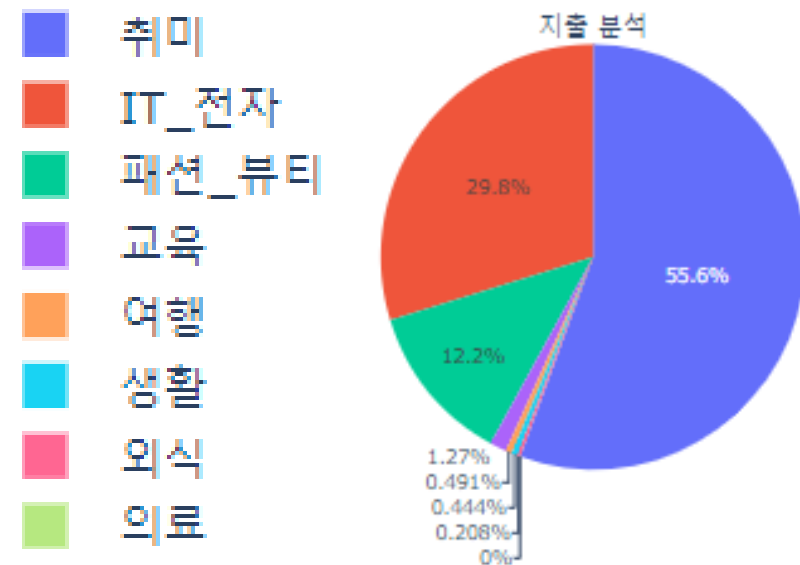
수요자 맞춤형 정보 제공

수요자 소비 패턴 기반 페르소나

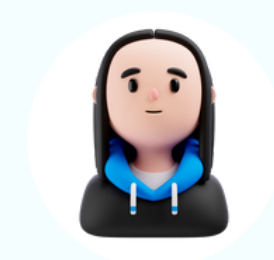
취미 55.6%, IT_전자 29.8%, 패션_뷰티 12% ..



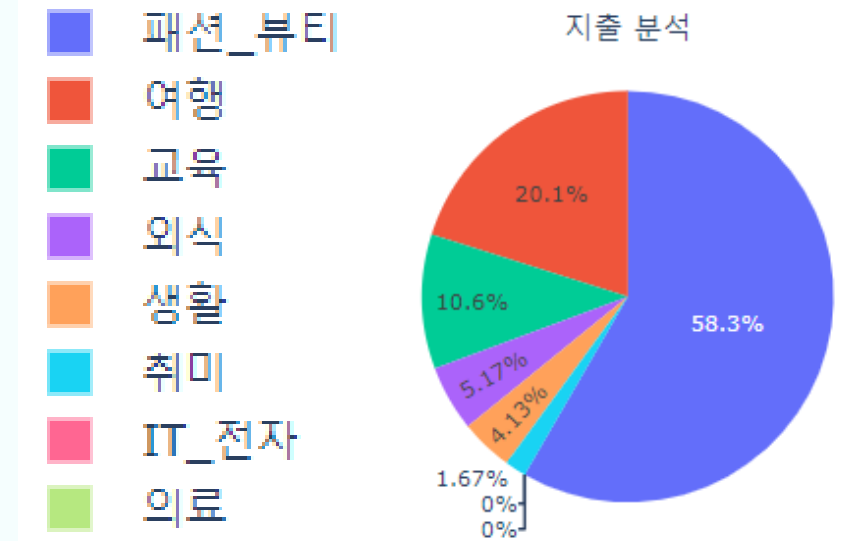
[20대 남]



패션/뷰티 58.3%, 여행 20.1%, 교육 10.6% 1% ..



[20대 여]



‘맞춤형 소비자 자원 사이트’를 고안

장점1

소비 패턴의 비율을 고려한 맞춤형 정보 추천

소비패턴의 비율을 고려하여
유사도가 가장 높은 보도자료를 추천
성별 연령별 상담다발품목 제공

장점2


한눈에 파악 가능한 핵심 요약

키워드와 핵심 문장 요약을 통해 한눈에 뉴스의
본질을 파악할 수 있도록 서비스 구현

웹페이지 구현 - HOME

HOME

PRIVATE




한국소비자원
Korea Consumer Agency



보도자료


"공유 전기자전거 이용자 99% 안전모 미착용, 공용 안전모 제공 필요"
: 올바른 주차구역에 대한 정보제공 강화 필요

 더 많은 정보를 알아보세요.



피해예방주의보


"프라이팬의 핵심 성능인 코팅 내구성, 제품 간 차이 있어"
: 손잡이 강도·기름 누설성·유해물질 등의 안전성은 모든 제품이 이상 없어

 더 많은 정보를 알아보세요.



소비자안전주의보


"화장품 용기, 재활용이 어려운 경우 많아"
: 친환경 소비를 위해 화장품 용기의 재활용 등급 표시 확대 필요


 더 많은 정보를 알아보세요.



상담다발품목

품목	피해유형	거래유형
외과	단순문의·상담	기타
스마트폰	AS불만	일반판매
기타가방	계약불이행(불완전이행)	국내온라인거래
전화이용서비스	약관	국내온라인거래


 더 많은 정보를 알아보세요.



소비자 단어


"바이럴 마케팅 (viral marketing)"
: 네티즌들이 이메일이나 다른 전파 가능한 매체를 통해 자발적으로 어떤 기업이나 기업의 제품을 홍보할 수 있도록 제작하여 널리 퍼지는 마케팅 기법

 네이버 지식백과로 바로가기



국내리콜정보

제품명	사업자명	공표일
빠투능	(주) 씨앰루드	~24.11.22

 소비자24로 바로가기

소비자뉴스

보도자료

소비자원동향

공지사항

카드뉴스

피해예방주의보

소비자안전주의보

소비자24 국내 리콜

상품·안전정보

국내리콜

11007. 콘암해강국

사업자명 (주) 포듀메트
리콜공표일 ~ 2024.11.12
출처 식품의약품안전처


11006. 냉동홍고추

사업자명 한성 글로벌
리콜공표일 ~ 2024.12.14
출처 식품의약품안전처

웹페이지 구현 - PRIVATE

HOME

PRIVATE

 한국소비자원
Korea Consumer Agency

고객 ID를 입력해주세요. 고객 ID의 앞자리는 연령대(1~8), 뒷자리는 성별(0,1)을 의미합니다.

고객 ID 예시) 20 : 20대 남자, 41 : 40대 여자

20

고객 ID

추천 받기

안녕하세요 😊 소비 패턴 분석을 통해 3개의 보도자료가 추천되었습니다.

1st

제목

"공유 전기자전거 이용자 99% 안전모 미착용, 공용 안전모 제공 필요"


요약

공유 전기자전거 이용자의 99%가 안전모 미착용, 안전모 제공 사업자는 없어
통행·소방시설 이용에 방해되는 위치에 전기자전거 주차 사례 빈번

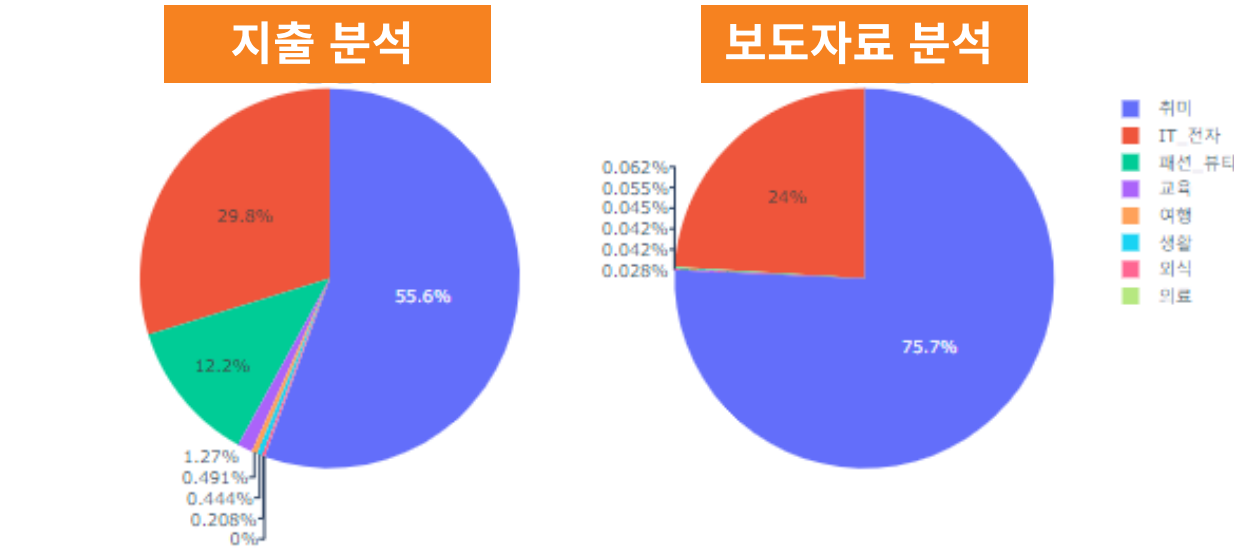
키워드

#지하철역 #이용자 #안전사고

시각화



🔥 분석결과 ➡ 96.5%



🗑️ 상담다발품목

