

2021 날씨 빅데이터 콘테스트



‘비건웨더’

날씨에 따른 비건 친환경 제품 추천 플랫폼 앱

구할 : 최수진 김수현 민지현 이정인

목차

공모 배경

데이터 정의

데이터 처리
방안

데이터 분석
기법

활용 방안
기대 효과



공모 배경

주제 선정 배경

#윤리적 소비 #가치 소비 # 착한 소비

화장품업계, 가치소비 확산 '비건' 트렌드 대세

■ MZ세대 뷰티 트렌드, 성분, 환경까지 고려 '비건' 제품 선호 2025년 시장가치 23조 전망

"착한 소비가 뜬다"...동물 실험 하지 않는 비건 화장품 인기



최근 기후 위기, 미세먼지, 코로나19 바이러스 등 급변하는 환경 속에서 윤리적 소비에 대한 관심과 마스크 착용에 따른 피부 자극에 대한 고민으로 '비건 화장품'에 대한 관심이 높아지고 있다.

주제 선정 배경



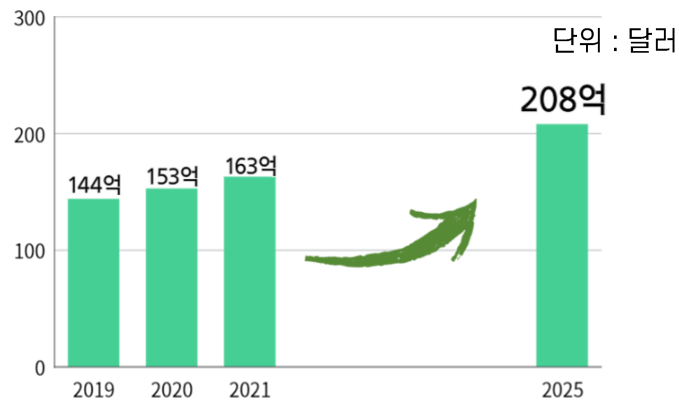
㉑ 비건 화장품이란?

비건(Vegan) 화장품이란 동물 유래 원재료를 이용하지 않고, 동물 실험을 실시하지 않으며, 생산 공정에서 비건이 아닌 제품에 의한 교차오염이 없는 화장품을 말한다.

㉒ 비건 화장품 시장 가치

세계 비건 화장품 시장은 2010년 중반 이후 연평균 6.3%씩 성장하고 있으며, 2025년 208억 달러 (한화 약 25조)에 이를 전망이다.

세계 비건 화장품 시장 성장 예측



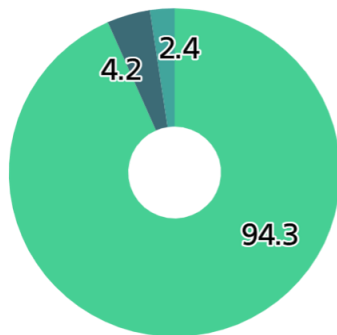
[출처 : 미국 그랜드 뷰 리서치]

주제 선정 배경



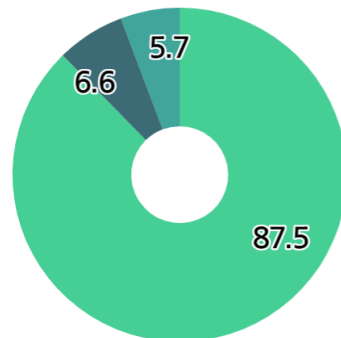
화장품 구매에 있어서 환경과 윤리적 행동을 고려하는 소비가 확대되고 있는 추세

같은 가격이라면 사회와 환경에 도움이 되는 화장품을 선택하겠다.



■ 그렇다 ■ 아니다 ■ 잘모름

화장품 브랜드의 가치관과 윤리적 행동에 대해 관심이 생겼다.

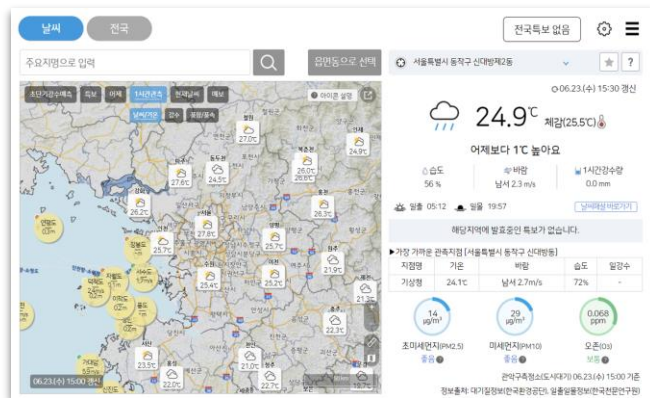


■ 그렇다 ■ 아니다 ■ 잘모름

[출처 : CJ올리브영 클린뷰티 설문조사]

환경을 생각하는 가치 소비의 시대, 비건 · 친환경 제품 추천 플랫폼을 만들자!

주제 선정 배경



실시간 날씨 데이터

W
'비건웨더'



블로그, 커뮤니티, 인스타그램
언급 건수 및 구매 건수 데이터

‘날씨 데이터와 SNS 언급 건수 및 구매 건수 예측에 기반하여
비건·친환경 화장품을 추천해주는 플랫폼 개발’을 주제로 선정





활용 데이터 정의

활용 데이터 정의



기상청

DATA 공공데이터포털
.GO.KR

V A I V

Corporation

출처

사용 데이터

기상청

2018-2019년(2년) 전국 일별 기상 관측 데이터
평균기온(°C), 평균일강수량(mm), 평균풍속(km/h),
평균습도(%rh), 일조합, 일사합, 체감온도(°C),
일 미세먼지 농도($\mu\text{g}/\text{m}^3$)

공공데이터 포털

공휴일 데이터

(주)엠코퍼레이션
(주)바이브컴퍼니

2018년~2019년(2년)
① 온라인 구매이력
② 소셜 데이터
일시, 성별, 연령대, 대분류명,
소분류명, 구매건수, 10만건 당 구매 건수

활용 데이터 정의



출처

한국비건인증원,
비건편의점,
기업 온라인 사이트

웨더아이

사용 데이터

비건 화장품 브랜드명, 상품명, 정가, 할인가 크롤링

실시간 서울 지역 기온(°C), 강수량(mm),
풍속(km/h), 미세먼지 농도($\mu\text{g}/\text{m}^3$) 크롤링

활용 데이터 정의

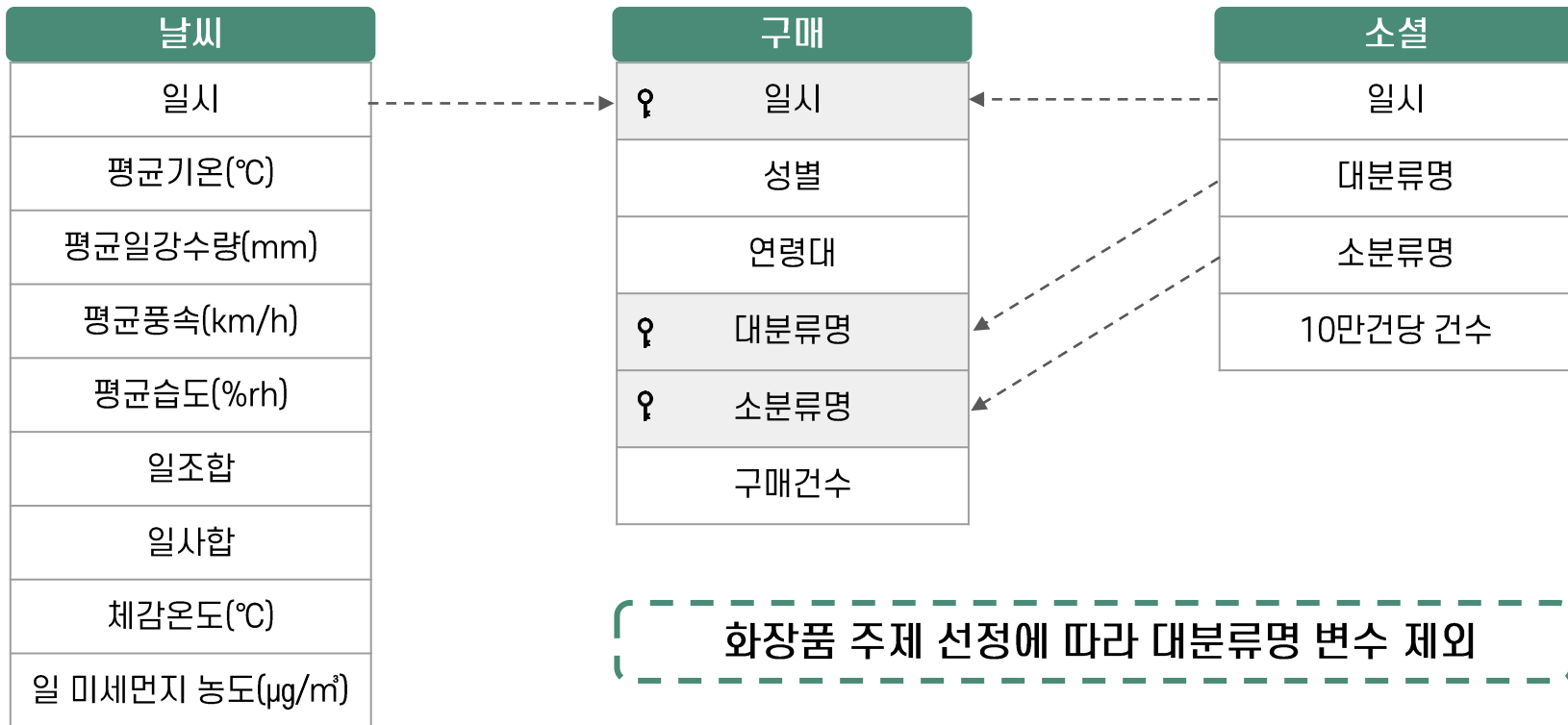
브랜드 종류 (총 37개)

1	Inglot	9	매닉패닉	17	러쉬	25	AHC	33	이즈그린
2	아렌시아	10	불독	18	아로마티카	26	수이스킨	34	쏘내추럴
3	아바마트	11	쿠모브러쉬	19	오엠	27	비브	35	오드리앤영
4	캐트리스	12	멜릭서	20	이든 미네랄	28	온도	36	발롱블랑
5	오셀로	13	보나쥬르	21	클레어스	29	에이트루	37	키피터치
6	더비건글로우	14	이솔	22	마녀공장	30	닥터올가		
7	더말로지카	15	허블룸	23	스킨푸드	31	베이지크		
8	라임크라임	16	닥터브로너스	24	카밀	32	싸이닉		

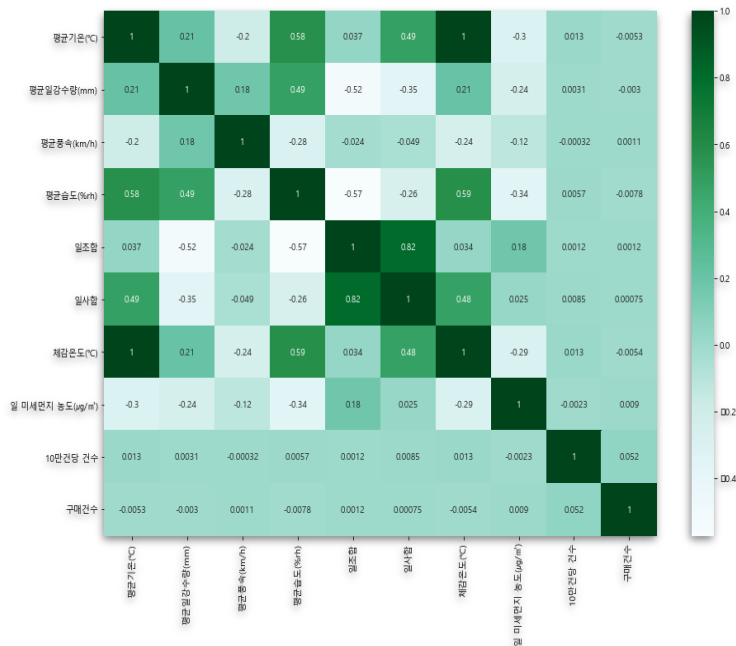


데이터 처리 방안

데이터 처리 방안 - 데이터 병합



데이터 처리 방안 - EDA



‘구매 건수’와의 상관 관계는 찾을 수 없었으나,
날씨 데이터 간에 뚜렷한 상관성이 나타나

다중공선성 확인

VIF

5356.72

51.17

11.62

변수

평균기온(°C)

일사합

평균습도(%rh)

VIF가 높은 평균기온, 일사합, 평균습도 변수 제외

데이터 처리 방안 - 파생변수 추가

기존 변수									
일시	성별	연령대	평균일강수량	평균풍속	체감온도	미세먼지농도	소분류명	10만건당건수	구매건수



파생 변수	
계절	봄: 3~5월, 여름: 6~8월, 가을: 9~11월, 겨울: 12~2월
월	-
일	-
요일	0: 월, 1: 화, 2: 수, 3: 목, 4: 금, 5: 토, 6: 일
공휴일	0: 공휴일X, 1: 공휴일O



파생 변수	
중분류명	기능성 화장품 외 19개

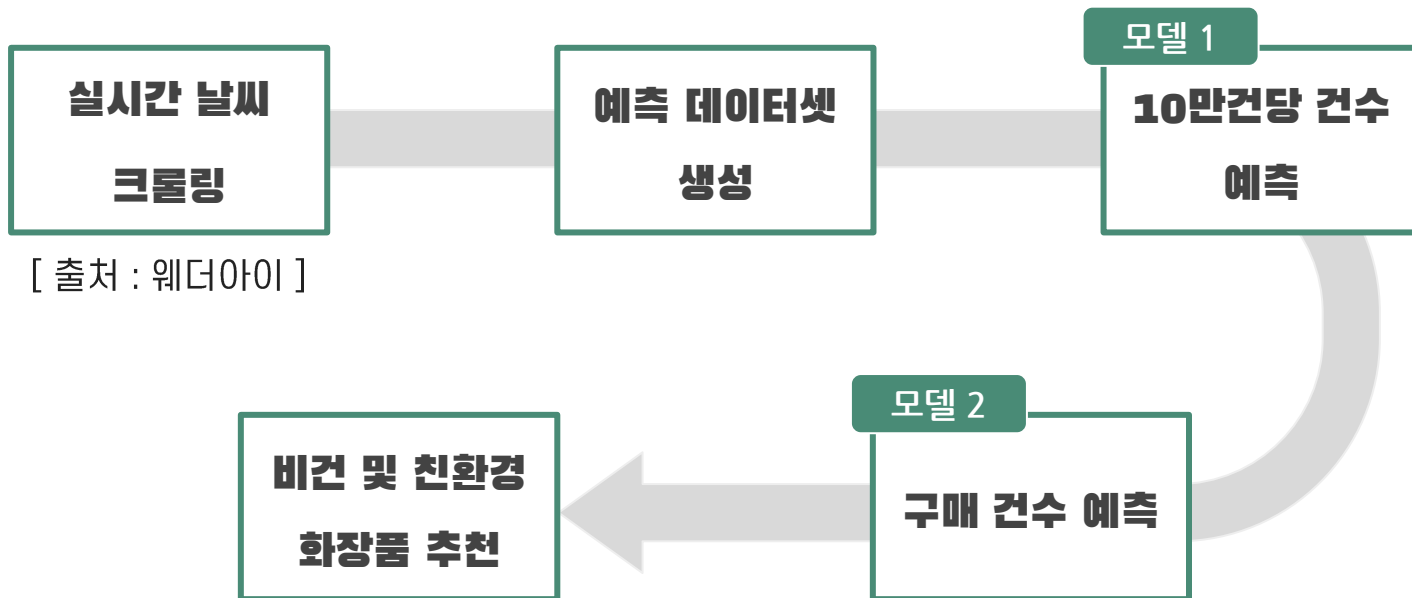
변수 선택
: 9개(기존) + 5개(파생) = 총 14개

- * 일시 변수는 파생 변수로 대체
- * 중분류명 변수는 추천 시스템에 활용



분석 기법

분석 기법



분석 기법

모델 1

**10만건당 건수
예측**

기상에 따른 '10만건당 건수' 예측 결과
'구매 건수' 예측에 사용

모델 2

구매 건수 예측

최종 예측인 '구매 건수'에 대한
설명 변수로 사용하기 위해
'10만건당 건수'에 대한 1차 예측 필요

분석 기법

사용 변수 및 전처리

사용 변수

계절, 월, 일, 요일, 공휴일, 성별, 연령대, 소분류명, 평균일강수량(mm),
평균풍속(km/h), 체감온도(°C), 일 미세먼지 농도($\mu\text{g}/\text{m}^3$), 10만건당 건수

전처리

- 결측치 제거
 - '일 미세먼지 농도($\mu\text{g}/\text{m}^3$)' 변수에 존재하는 1884건 제거
- 이상치 처리
 - 평균일강수량(mm), 평균풍속(km/h), 일 미세먼지 농도($\mu\text{g}/\text{m}^3$), 10만건당 건수에 대한 각 계절별 이상치를 평균으로 대체
- 라벨 인코딩
 - 계절(봄, 여름, 가을, 겨울), 성별 (F, M), 소분류명(131개)에 라벨 인코딩 적용
- '구매 건수' 로그 변환

분석 기법

모델 1 : 10만건당 건수 예측

사용 모델

Extra Trees Regressor, LightGBM

Training, Test data 8:2

Model	RMSE	RMSLE
Extra Trees Regressor	0.107	-
LightGBM	0.3946	0.1445

Extra Trees Regressor 옵션

- 결정트리 개수 (n_estimators) : 200
- max_features : 11
- min_samples_split : 2

모델 선택 : Extra Trees Regressor 모델을 이용하여 '10만건당 건수' 예측

분석 기법

모델 2 : 구매 건수 예측

사용 모델

RandomForest Regressor, XGBoost, LightGBM

Training, Test data 7:3

Model	RMSE	RMSLE
Random Forest Regressor	0.4371	0.1612
XGBoost	0.4662	0.1709
LightGBM	0.4702	0.1715

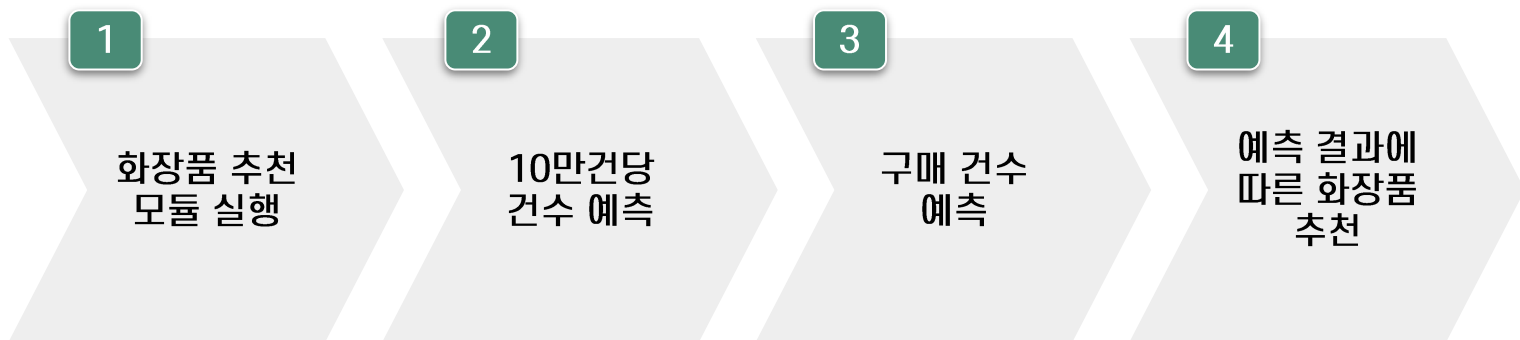
Random Forest Regressor 옵션

- 결정트리 개수 (n_estimators) : 100
min_samples_split : 2

모델 선택 : Random Forest Regressor 모델을 이용하여 '구매 건수' 예측

분석 결과

모델 1, 2를 기반으로 비건 · 친환경 화장품을 추천해주는 모듈 생성



분석 결과

```
import Vegan_Cosmetics
vc = Vegan_Cosmetics.Vegan_Cosmetics()
```

성별과 연령대를 입력하세요 : F 20

```
vc.extratrees_social() # 10만건당 건수 예측
vc.randomforest_buy() # 구매 건수 예측
```

```
df = vc.cosmetics_recomm()
df
```

린스 상품이 추천되었습니다.

	브랜드	중분류	소분류
0	더비건글로우	헤어 케어	린스
1	러쉬	헤어 케어	린스
2	아로마티카	헤어 케어	린스
3	아로마티카	헤어 케어	린스
4	아렌시아	헤어 케어	린스
5	아렌시아	헤어 케어	린스

```
df = vc.cosmetics_recomm()
df
```

메이크업 박스와(과) 관련한 메이크업 도구 상품이 추천되었습니다.

	브랜드	중분류	소분류	상품명	정가	할인가
0	쿠모 브러쉬	메이크업 도구	메이크업 브러쉬	스크류 브러쉬	8900	5900
1	쿠모 브러쉬	메이크업 도구	메이크업 브러쉬	블랙 핏 메이크업 스펀지 with 헤이다	11900	7900
2	Inglot	메이크업 도구	메이크업 브러쉬	브러쉬 클렌징 리퀴드	15000	
3	아바마트	메이크업 도구	메이크업 브러쉬	[RA007] 비건 순삭 아이라이너 브러쉬 2pi	15000	9000
4	쿠모 브러쉬	메이크업 도구	메이크업 브러쉬	앵글드 립 브러쉬	16000	12900
...

화장품 추천 모듈 실행

고객 정보를 입력받아
예측 데이터셋에 저장

예측 결과에 따른 화장품 추천

구매 건수가 가장 높게 예측된
소분류 상품 리스트 추천

비건 상품의 특성상 대응되는
화장품이 없을 경우
관련 중분류 상품 추천



활용 방안 및 기대 효과



‘비건웨더’

: 날씨에 따른 비건 친환경 제품 추천 플랫폼 앱

'비건웨더' 소개



1. 고객 맞춤 회원 서비스

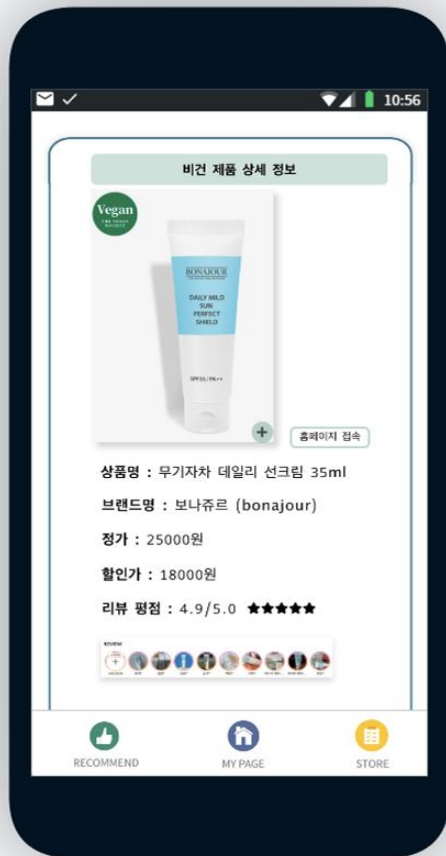
고객 정보 입력 : 연령대 & 성별
해당 정보 기반 추천 시스템 제공

2. 날씨 관련 비건 제품 추천 서비스

당일 및 주간 기상 정보 제공
추천하는 비건 제품 리스트 제공



‘비건웨더’ 소개

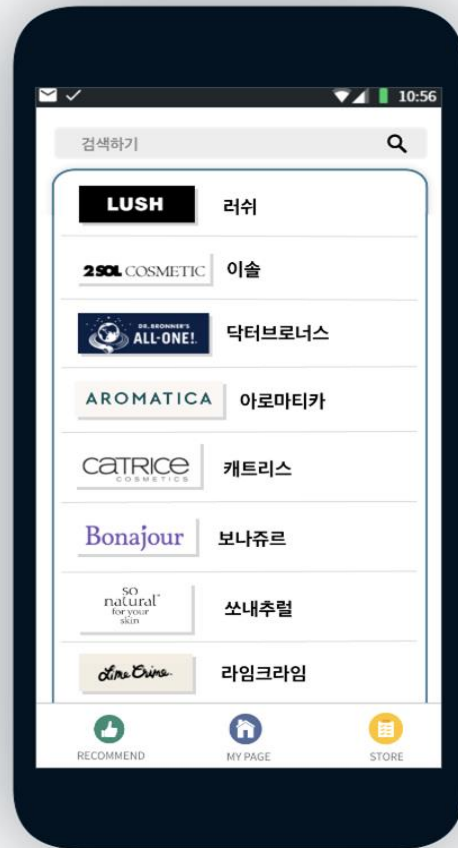


3. 상세 정보 제공 서비스

상품명, 브랜드명,
정가, 할인가, 리뷰평점 등
비건 제품 상세 정보 제공

4. 비건 스토어 리스트

비건 제품 관련
스토어 목록과 실시간 접속 가능



‘비건웨더’ 활용 방안 & 기대효과

❖ 차별점

- 여러 기상 정보를 일괄적으로 제공하는 편리성
- 기상 정보 & 개인 맞춤형 화장품 추천 서비스
- 환경을 생각하는 가치소비의 시대에 발맞춘 소비 플랫폼

❖ 비즈니스 효과

- 비건 기업과의 제휴 및 비건 제품의 생산 독려 가능
 - 개인 맞춤형 회원제 시스템, 쇼핑 이벤트, 적립금 시스템
- ☐ 일회성 구매가 아닌 장기적 활용으로 이어지도록 유도

‘비건웨더’ 기대효과



고객편리성 측면

고객편리성 측면

- ‘기상 맞춤형’ 화장품 추천이라는 차별성
- 기상 정보 제공의 편리성으로 이용자 확대 기대
- 비건 제품에 대한 접근성 상승 예상

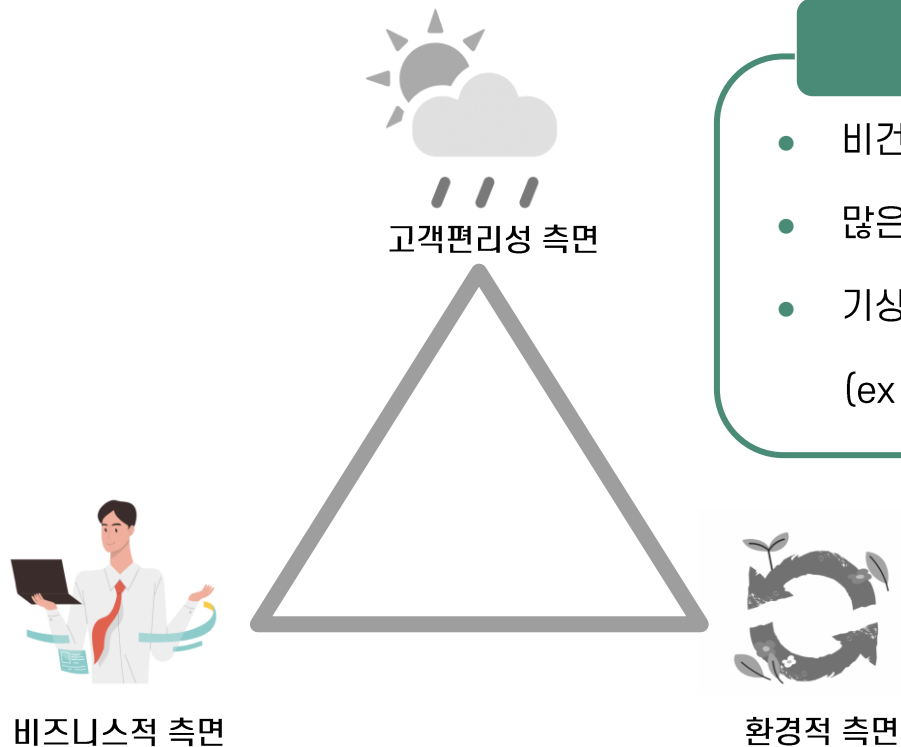


비즈니스적 측면



환경적 측면

‘비건웨더’ 기대효과



비즈니스적 측면

- 비건 기업과의 제휴
- 많은 기업이 비건/ 친환경 화장품 제조에 동참하도록 유도
- 기상 별 수요 예측 정보로 기업의 효율적 경영에 도움
(ex 계절 별 행사 아이템 선정, 시기 별 재고 관리에 참고)

‘비건웨더’ 기대효과



고객편리성 측면

환경적 측면

- 제품 성분부터 제조 공정까지
환경을 생각하는 기업 및 고객 증가
- 서비스 이용 기업과 고객이 증가하는 선순환에 따라
환경 보호, 기후 위기 해소에 이바지



비즈니스적 측면



환경적 측면



Thank you

