



# Perfume Recommendation System



시계열자료분석팀    오정민 강현주 김민정 배지현



# Contents

---

향수 추천시스템

01

주제 선정

02

데이터

03

EDA

04

4주차 예고



주제 선정

데이터

EDA

4주차 예고

# 01. 주제 선정



기존 향수 추천 시스템



파페, 퍼퐁그래피 등

추천시스템을 가진  
회사에서 만든 제품을  
홍보하고 판매하기 위한 목적!



주제 선정

데이터

EDA

4주차 예고

## 01. 주제 선정

어떤 방식의 추천 시스템이 필요?

// 향수 특성을 기반으로 한 //

### 소비자의 취향

향기, 지속성 등

- ✓ 향수의 추상적 특성 고려하여 추천
- ✓ 향수의 구체적 특성을 잘 모르는 사람도 쉽게 추천 받을 수 있도록



주제 선정

데이터

EDA

4주차 예고

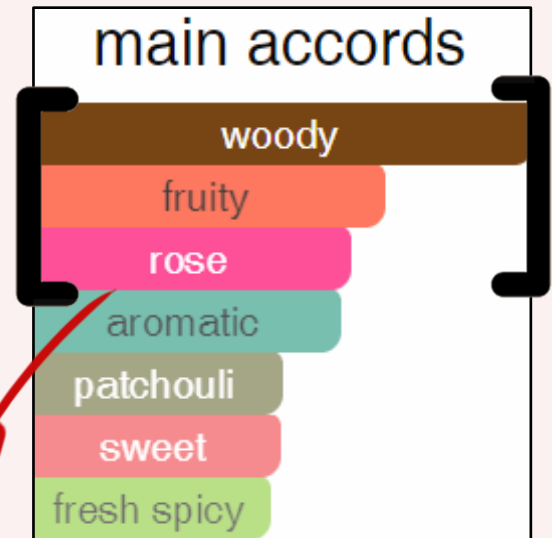
## 02. 데이터

데이터 수집 변수설명 전처리



크롤링 데이터 - main accords

Name	Brand	Main1	Main2	Main3
Gucci Guilty Absolute pour Femme	Gucci	Woody	Fruity	Rose
Coromandel Parfum	Chanel	Amber	Citrus	Warm Spicy
Sauvage Eau de Parfum	Christian Dior	Aromatic	Fresh spicy	Lavender
...	...	...	...	...



Main accords 상위 3개 추출!



주제 선정

데이터

EDA

4주차 예고

## 02. 데이터

데이터 수집 변수설명 전처리



크롤링 데이터 - 계절/ 낮/ 밤

Name	Brand	Cold	Spring	Hot	Fall	Day	Night
Gucci Guilty Absolute pour Femme	Gucci	71.4674	65.2174	28.2609	100	81.25	78.5326
Coromandel Parfum	Chanel	63.6364	63.6364	45.4545	81.8182	77.2727	100
Sauvage Eau de Parfum	Christian Dior	71.4674	65.2174	28.2609	100	96.4286	94.335
...	...	...	...	...	...	...	...

각 계절에 얼마나 어울리는지를 수치로 표현!



주제 선정

데이터

EDA

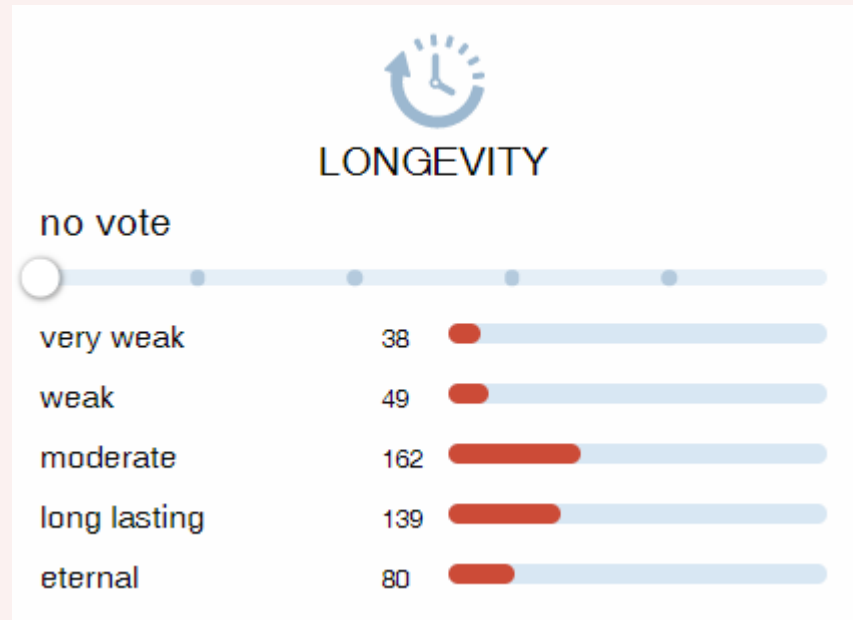
4주차 예고

## 02. 데이터

데이터 수집 변수설명 전처리



전처리 - Longevity



투표수가 max인 범주를 **ordinal variable** 형식으로 가져옴



주제 선정

데이터

EDA

4주차 예고

## 02. 데이터

데이터 수집 변수설명 전처리



크롤링 데이터 - review

상위 5개 **긍정 리뷰**

상위 5개 **부정 리뷰**

Name	Brand	pos_review	neg_review
Mon Guerlain Eau de Parfum Intense	Guerlain	["While I may not have been impressed with the first Mon Guerlain and its	['Alberto Morillas increased my hopes for this fragrance. Old Gucci fragrances are the best, but the new ones are the worst ...']
Coco Mademoiselle Intense	Chanel	['I asked my 5 year old son to do this review. Here are his thoughts...Wn"Smells like roses! ...	['Unlike anything I own, so I figured Iw'd try to leave my comfort zone and ...
Chance Eau Tendre Eau de Parfum	Chanel	['Success !!! Making this my signature again, edt was my signature 6 years ago ...	["In Chance eau Tendre EDT the iris was strong on my skin, changing the fresh fragrance into the creamy soap. ...
...	...	...	...





주제 선정

데이터

EDA

4주차 예고

## 02. 데이터

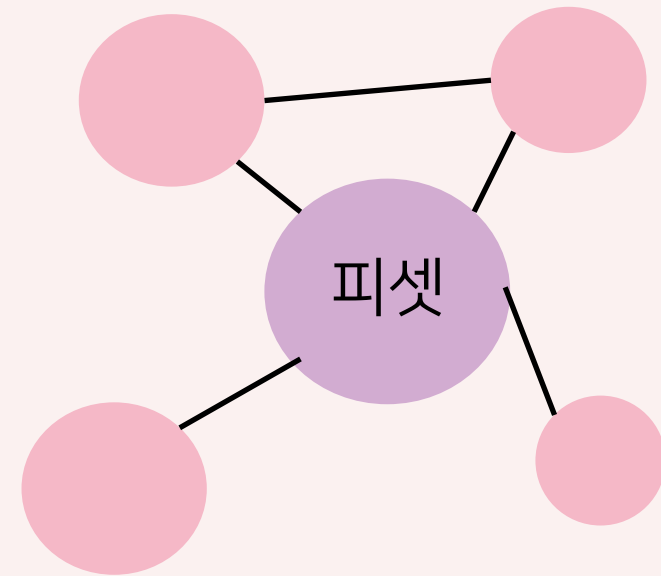
데이터 수집 변수설명 전처리

- 리뷰 - 핵심 키워드 추출하기

### TextRank

: PageRank에 기반한 Graph Ranking 알고리즘

나는 피셋이 좋아요.  
피셋에 뼈를 묻자.  
주제분석 너무 좋아.



“단어 그래프를 대표하는 **핵심 단어**를 선택하는 문제!”



주제 선정

데이터

EDA

4주차 예고

## 02. 데이터

데이터 수집 변수설명 전처리

- 리뷰 - 핵심 키워드 추출하기

### TextRank

Name	Brand	pos_review	neg_review
Mon Guerlain Eau de Parfum Intense	Guerlain	["While I may not have been impressed with the first Mon Guerlain and its	['Alberto Morillas increased my hopes for this fragrance. Old Gucci fragrances are the best, but the new ones are the worst ...']



vanilla

original

mon

guerlain

intense

TR값이 높은 5개의 값을  
핵심 키워드로 추출!



주제 선정

데이터

EDA

4주차 예고

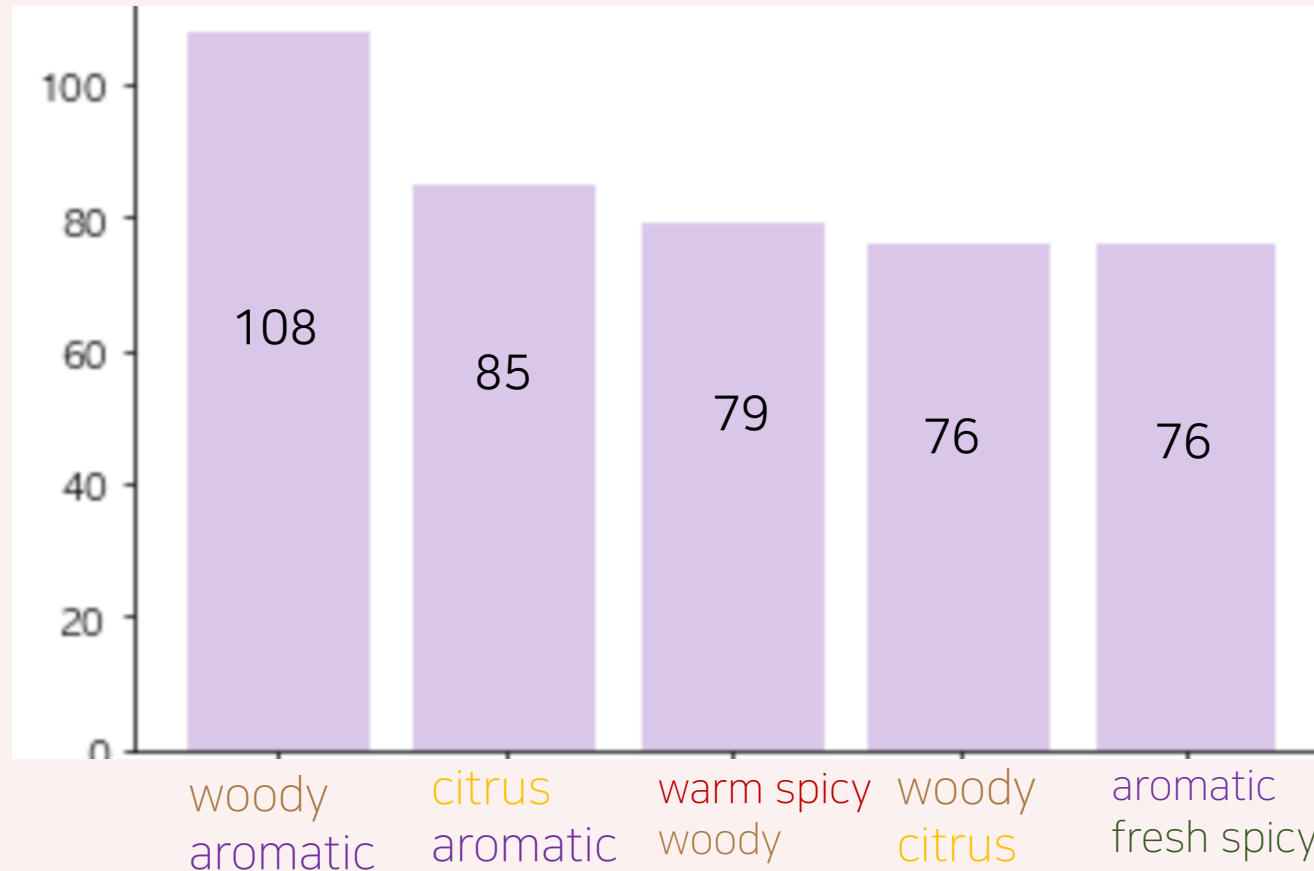
## 03.EDA

빈도 순위

변수 간 관계

리뷰/태그

### <가장 자주 쓰인 *main accord* 조합 Top5>



Top5 순위권 안의  
main accord들로  
이루어진 조합!



## 03.EDA

빈도 순위

변수 간 관계

리뷰/태그

주제 선정

데이터

EDA

4주차 예고

### <성별과 Middle Notes>



Female과  
Unisex에서

Jasmine, rose가

자주쓰임

Male은 조금 다르게

나타남



주제 선정

데이터

EDA

4주차 예고

## 03. EDA

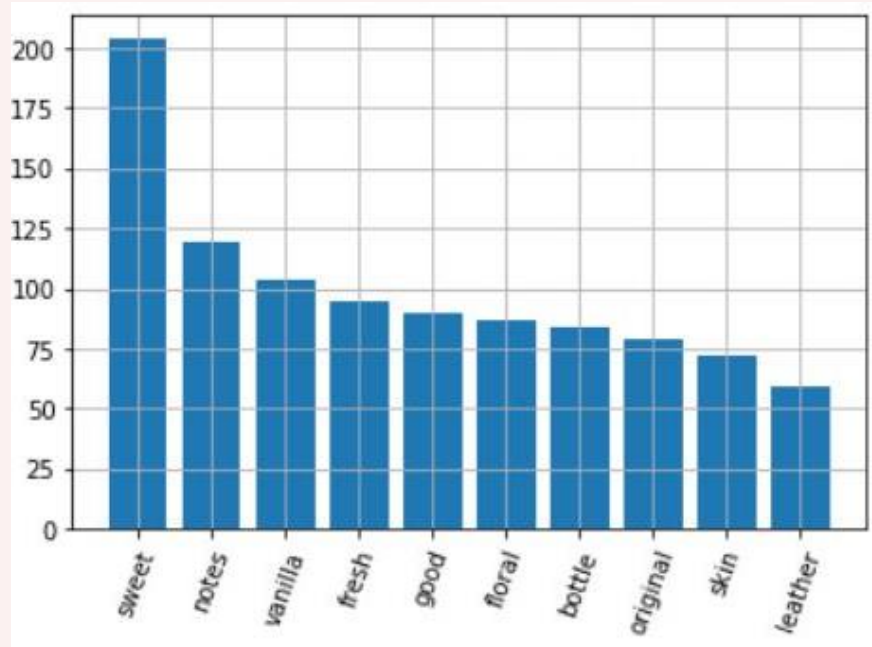
빈도 순위

변수 간 관계

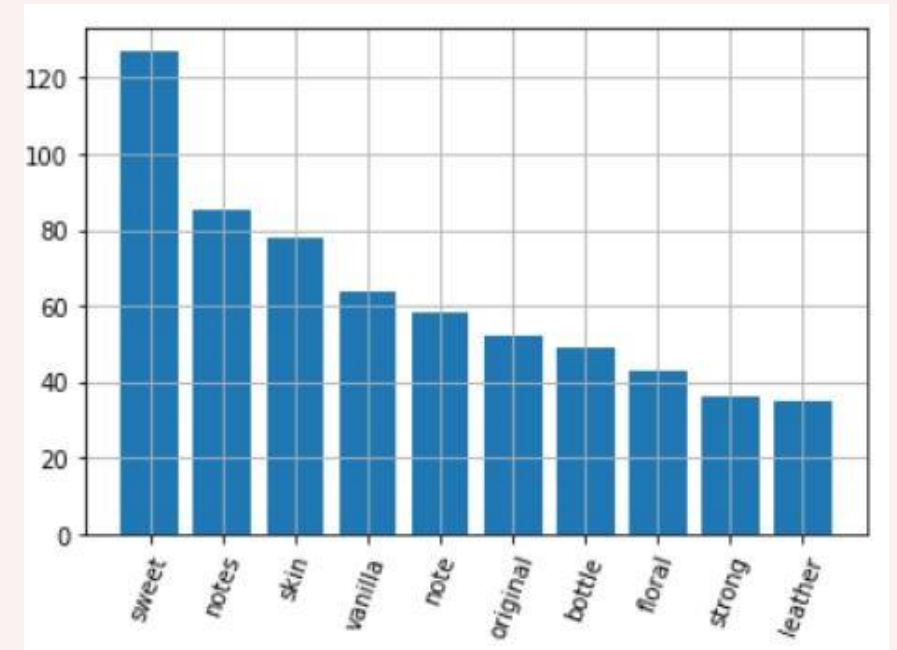
리뷰/태그

- 태그 데이터

긍정리뷰 태그 빈도수



부정리뷰 데이터 빈도수



공통적으로 나온 단어: sweet, notes, vanilla, floral, bottle, skin, leather, original

“빈도수가 높은 상위 10개의 단어가 거의 같다”



주제 선정

데이터

EDA

4주차 예고

## 03. EDA

빈도 순위

변수 간 관계

리뷰/태그

### • 태그 데이터

예시 중 하나만 가져왔어요

To me this is a bottle of Olympea and Ellie Saab Le Parfum combined and stirred together. It smells like a combination of both at the same time. I'm a little let down by that because it lacks originality. On the upside, you don't need any of the above mentioned because you would already have them in this perfume.

**"향수의 독창성이 떨어지기 때문에 조금 실망했다"**

소비자는 자신의 개성을 표현할 향수를 원한다!

독창성에 따라 향수 평가 달라질 수 있다



공통적으로 나온 단어: sweet, notes, vanilla, floral, bottle, skin, leather, original

**"빈도수가 높은 상위 10개의 단어가 거의 같다"**