

**FROM**

---

**FOOD TO**

---

**WINE**

**Justine Meier**





# Introduction

- 음식과 와인의 페어링은 오래전부터 행해졌던 기술의 일종
- 페어링 과정에서 엄격한 룰 적용
- 개인차로 인한 페어링 복잡성 증대





# Overall Goal

- Naïve Bayes와 K-Means 를 통해 와인의 색(레드, 화이트) 선택  
또는 음식과 알맞은 특정 와인 선택의 **단순화**



# Dataset

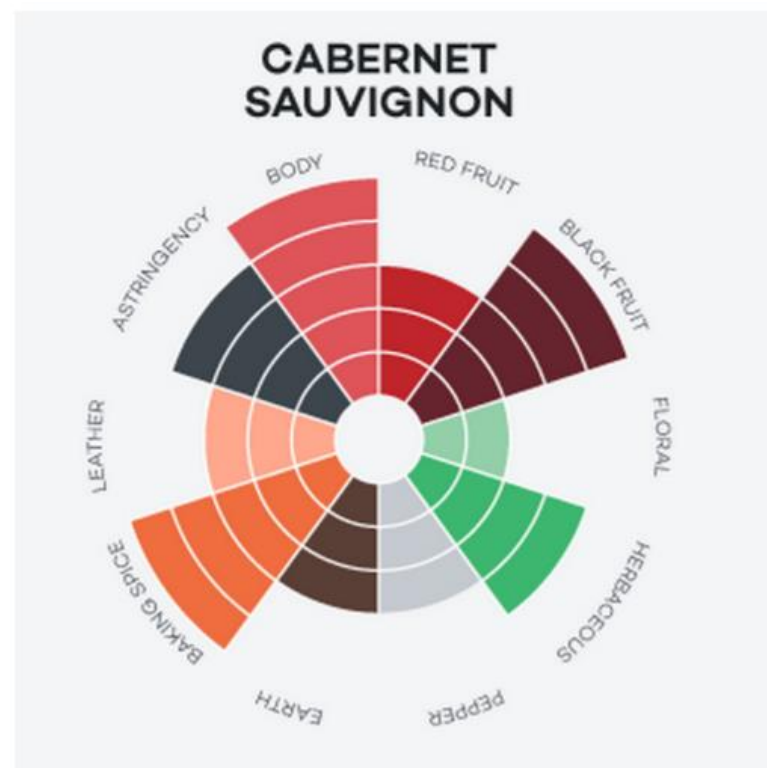
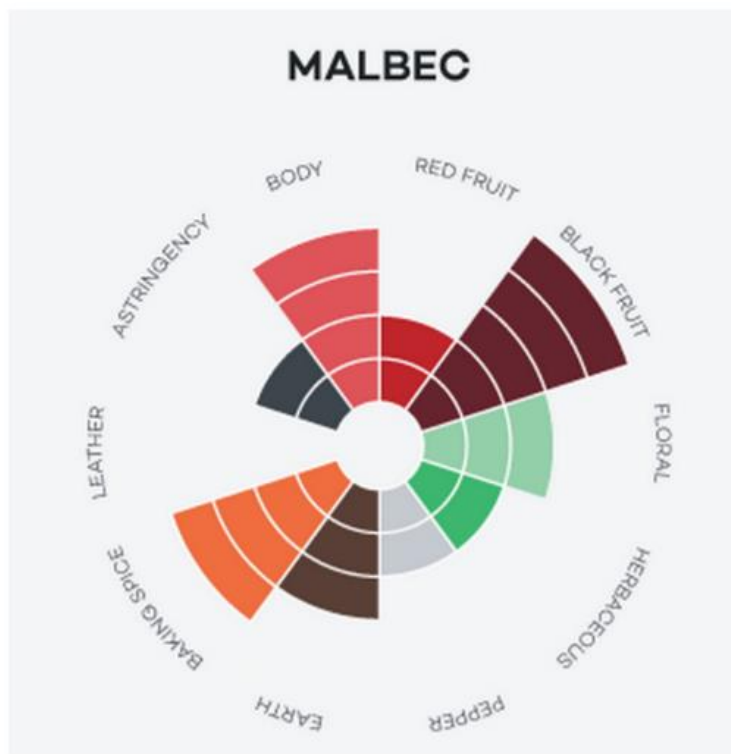
- 기존의 음식과 와인의 페어링에 대한 평가 등급
- 1~5 등급

Food ▲ ▼		Type ▲ ▼	Varietal ▲ ▼	Rating ▲ ▼
▶ Tomatoes Provencale	Paired with:	Rosé	Côtes de Provence, Rosé AOC	★★★★★
Fried Green Tomatoes	Paired with:	White Wine	Sauvignon Blanc	★★★★★
Fried Green Tomatoes	Paired with:	White Wine	Pinot Grigio	★★★★★
Tomatoes	Paired with:	Red Wine	Barbera DOC	★★★★★
Tomatoes	Paired with:	Red Wine	Chianti DOCG	★★★★★
Tomatoes	Paired with:	Red Wine	Sangiovese	★★★★★



# Dataset

- 와인의 맛에 대한 평가 그래프
- 10가지 flavor, 등급은 0~5



# Dataset

- 와인 페어링 관련 회사들의 자료
- 테스트 데이터

## Steamed Clams and Mussels with Grilled Bread

### **Wine:**

**Color:** White, **Varietal:** Sauvignon Blanc, **Vineyard:** Beringer,  
**Year:** 2006, **Region:** Napa Valley, California

Seafood - Shell Fish

**Users rating:**



[view details...→](#)





# Features & Preprocessing

$$FlavProf(food) = \frac{1}{RateSum(food)} \sum flavorProf(wine) * rating$$

- 각 재료에 대한 와인의 flavor 평균
- 높은 순위의 와인일 수록 재료와의 pairing 결정 시 선택 될 확률 증가

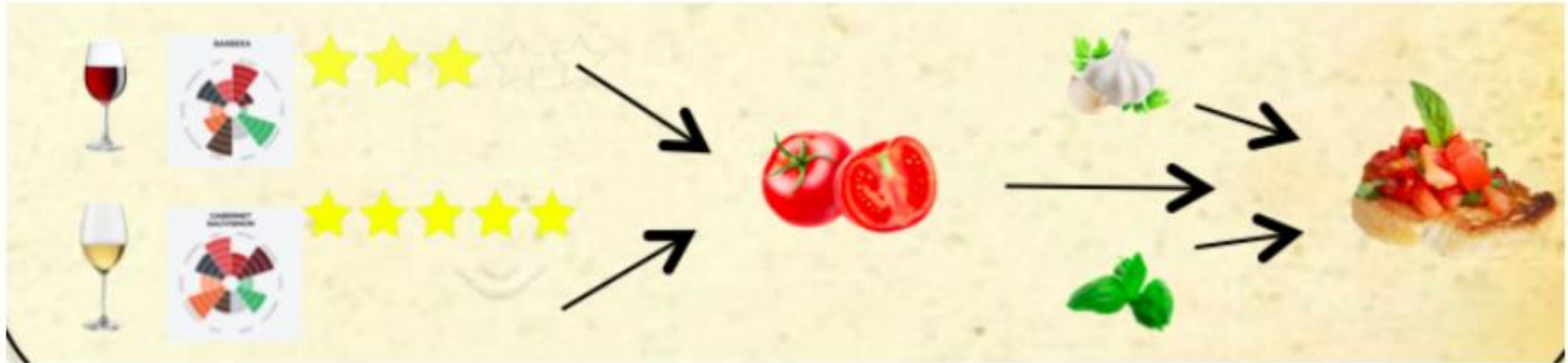
$$FlavProf(dinner) = \frac{1}{numOfIngredients} \sum flavorProf(food)$$

- 음식 flavor의 평균
- 각 재료 flavor의 평균과 와인의 flavor의 평균에 영향 받음





# Features & Preprocessing

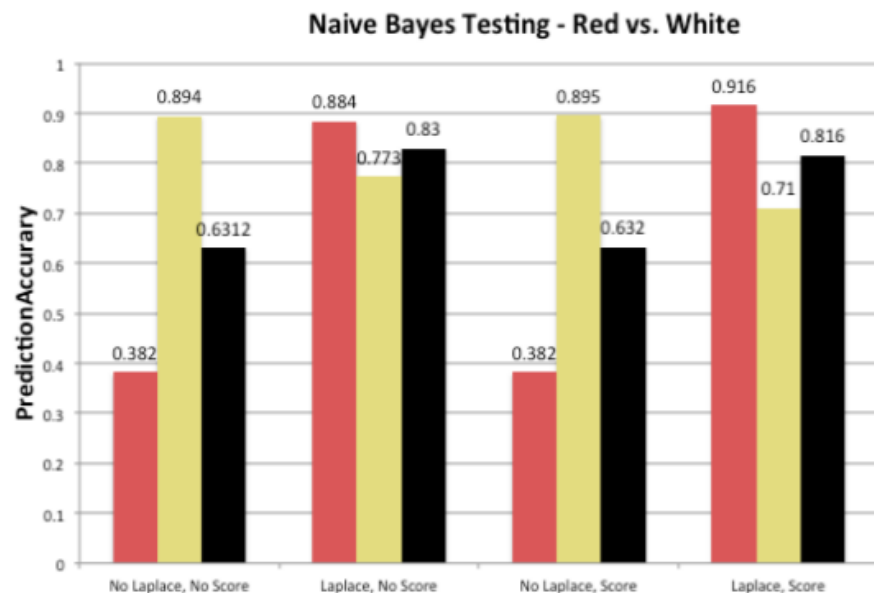






# Models : Naïve Bayes - Color Select

- 순위를 고려했을 때
  - 더 높은 순위의 와인이 선택 예측율이 더 높게 나옴
  - Laplace smoothing이 정상작동 X, 모든 등급이 평균으로 주어짐
- 순위를 고려 하지 않을 때
  - Laplace 정상작용



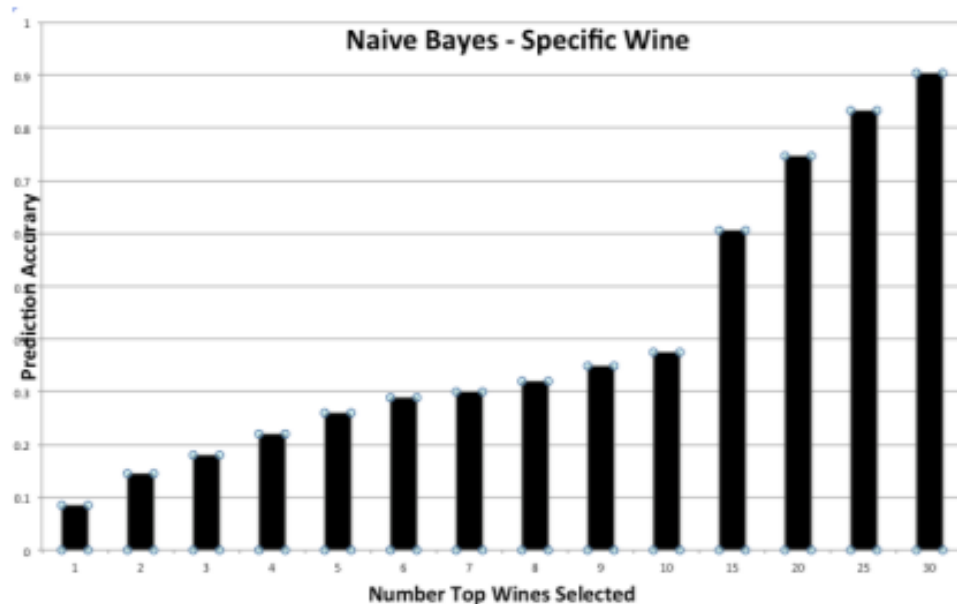
Red vs. White			
	Red Pred. Accuracy	White Pred. Accuracy	Both Pred. Accuracy
No Laplace No Score	38.2%	89.4%	63.1%
Laplace No Score	88.4%	77.3%	83.0%
No Laplace Score	38.2%	89.5%	63.2%
Laplace	91.6%	73.0%	81.6%





# Models : Naïve Bayes - Specific

- Color Prediction에서 전체 정확도가 높은 모델 사용
  - Laplace, No Score 모델
- N이 클 수록 예측 정확성 높아짐



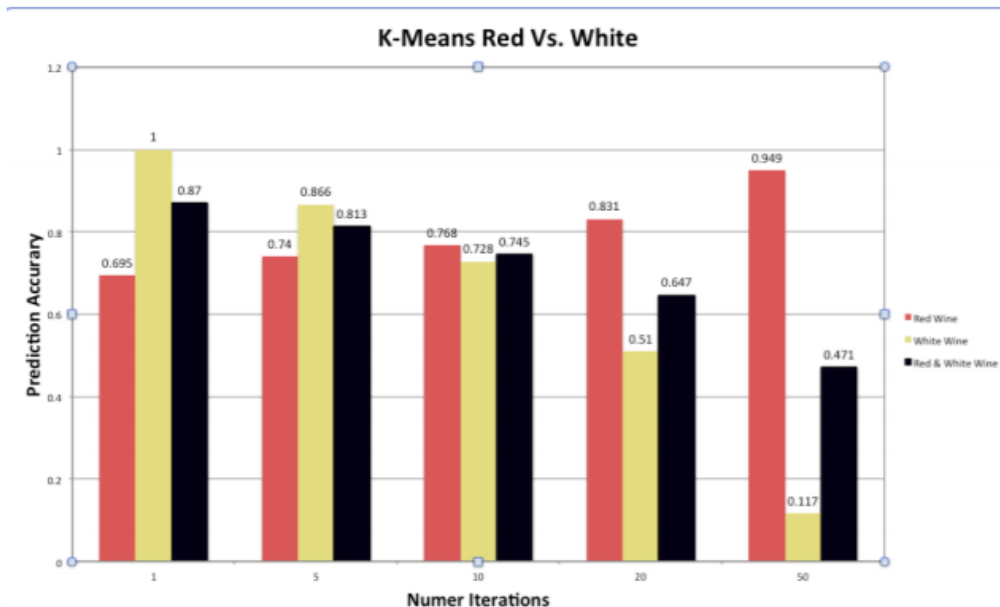
Specific Wine (with top N wines)	
N = 1	8.5%
N = 3	18.0%
N = 5	26.0%
N = 10	37.5%
N = 20	74.7%
N = 30	90.4%





# Models : K-Means

- Iteration이 증가할 수록 화이트 와인 선택 예측 확률이 낮아짐



Red vs. White from Specific Wine			
	Red Pred. Accuracy	White Pred. Accuracy	Both Pred. Accuracy
Iters = 1	69.5%	100%	87.0%
Iters = 5	74.0%	86.6%	81.3%
Iters = 10	76.8%	72.8%	74.5%
Iters = 20	83.1%	51.0%	64.7%
Iters = 50	94.9%	11.7%	47.1%



# Models : K-Means



- 육류는 Red Wine과 Pairing 등급 높음
- 채소류는 White Wine과 Pairing 등급 높음

→ White와인과 잘 Pairing되는

스테이크여도 Iteration이 증가 할 수록

육류와 Pairing되는 Red와인의 Weight가

증가해서 예측 실패

