Wine Recommendation System

By: Zachary Greenberg

Overview

- Business Problem
- Data Collection
- Exploratory Data Analysis
- Preprocessing
- Modeling
- Results
- Streamlit
- Next Steps

Business Problem / Proposal

- Tons of wines from around the world

- Find similar wines

- Educate consumers to make informed selections

- Wineries can recommend other bottles

Data Collection

- Scraped from Vivino.com
- 1800 wines

winery: str, the winery that makes that specific wine

wine_name: str, the name of the wine

wine_type: str, the classification of the wine (ie red, white..)

wine_country: str, the country in which the wine was made

average_rating: float, the average rating of the wine on Vivino

num_of_ratings: int, the number of ratings for that wine on Vivino

wine_price: float, the price per bottle

grapes: str, the grapes utilized in the wine (ie Chardonnay, Malbec..)

alcohol_content: the percent alcohol in the wines

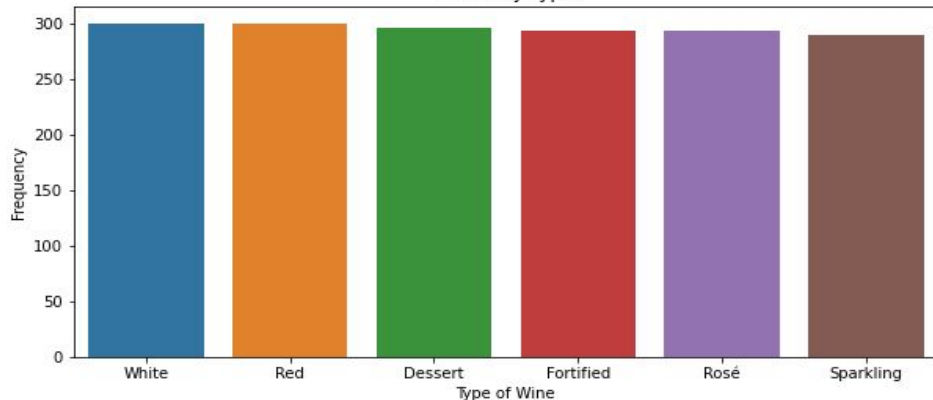**reviews**: str, descriptive text reviews of the wine
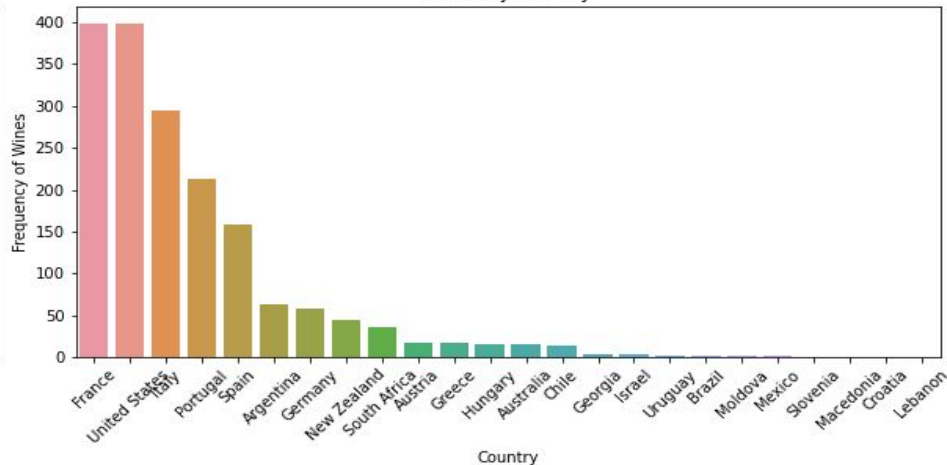
Exploratory Data Analysis

- Even distribution of wine by type
- Wines from countries around the world

Wines by Type

Wines by Country

Wines by Winery (Top 5 Representation)

| Winery	Number of Wines
Lustau	16
Quady	14
Sandeman	13
Taylor's	11
Warre's	10

- Multiple products from same winery

Preprocessing

Sample Text Review:

'Possibly my favorite SB. Enjoyed this at the winery. An oaked SB (18/82 new/neutral) that spends months on the lees. This was the first vintage where they used almost completely estate fruit. Delicious and fruity with vivid flavors of Meyer lemon, tangerine, peach, honeysuckle, and just a touch of grass. Really gorgeous.'
- Anonymous User

Natural Language Processing →
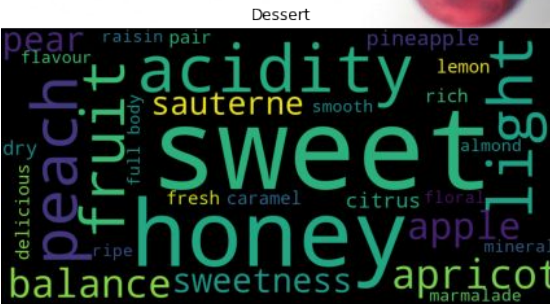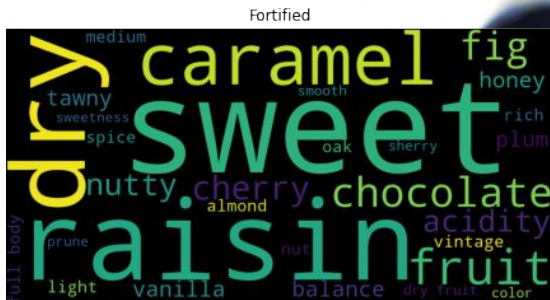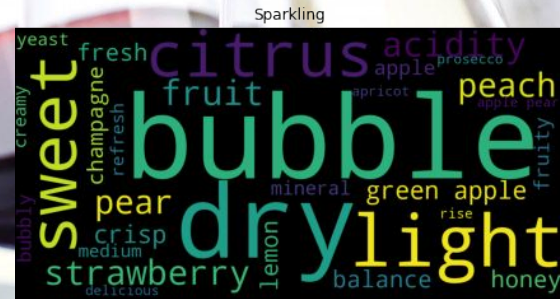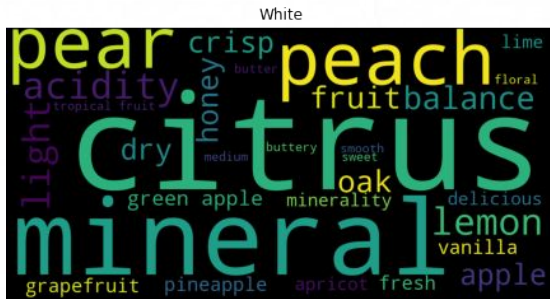
Modeling

- ## Predictive Modeling

Multinomial Naive Bayes
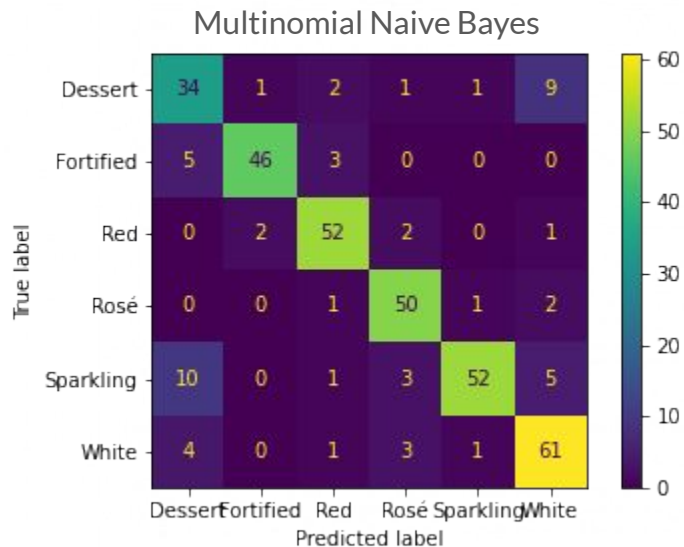
83% Accuracy

- ## Recommendation Systems

	1000 Stories Chardonnay 2018	Cosine Model	Euclidean Model
Average Rating	4.10	4.220000	4.160000
Wine Price	17.99	29.795587	25.755587

Results

- Rec System by Wine Name using Euclidean Distance

1000 Stories Chardonnay 2018 title	wine_type	wine_country	grapes	average_rating	wine_price	euclidean_distance_scores
Brown Estate House of Brown Chardonnay 2018	White	United States	Chardonnay	4.2	25.00000	14.060775
Miner Chardonnay 2018	White	United States	Chardonnay	4.2	26.99000	14.454876
The Prisoner The Snitch Chardonnay 2018	White	United States	Chardonnay	4.2	24.47000	15.849033
Ministry of The Vinterior Pinot Noir 2017	Red	United States	Pinot Noir	4.3	28.99000	16.072130
The Calling Dutton Ranch Chardonnay 2017	White	United States	Chardonnay	4.1	22.99000	16.322660
Le Mesnil Blanc de Blancs Brut Champagne Grand Cru	Sparkling	France	Chardonnay	4.0	35.19000	16.399854
Andresen 20 Year Old White Porto N.V.	Fortified	Portugal	Fortified Blend	4.2	27.59587	16.411672
Laguna Chardonnay 2017	White	United States	Chardonnay	4.2	18.35000	16.430939
Flor Brut Rosé N.V.	Sparkling	Italy	Pinot Noir	4.1	19.99000	16.435651
Domaine Carneros Brut 2016	Sparkling	United States	Sparkling Blend	4.1	27.99000	16.449608

1000 Stories
Chardonnay 2018

Streamlit App
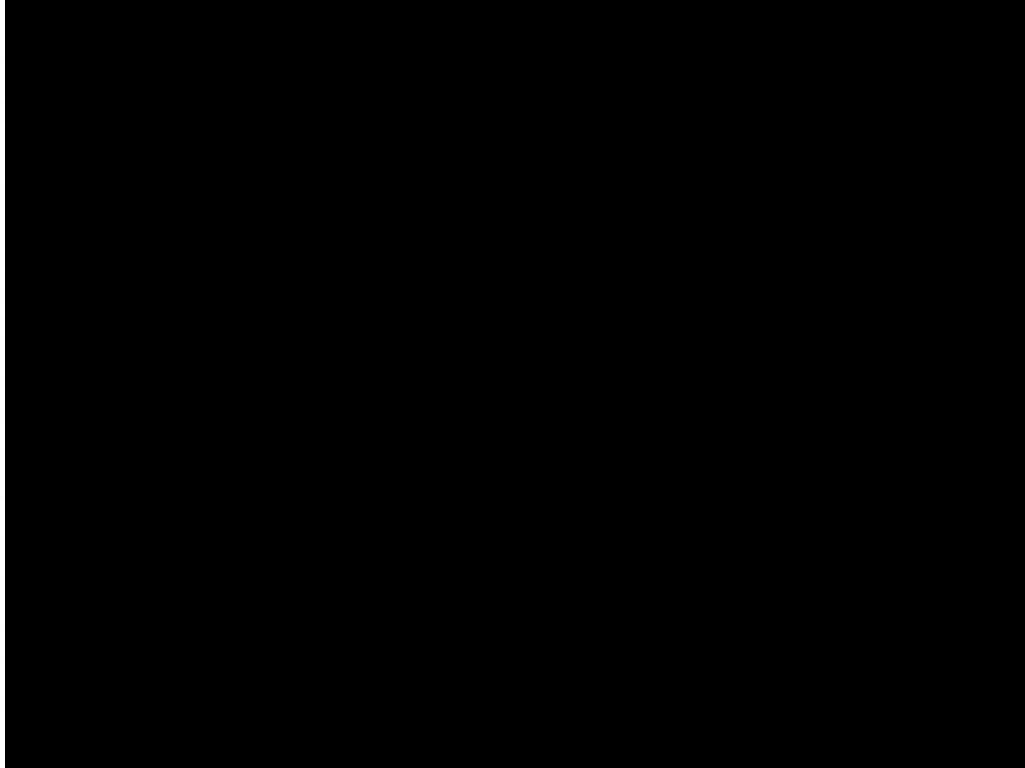
Next Steps

- Retrieve more wine data

- Use LDA for Topic Modeling

- Add more features to Streamlit App

Thank You!

Email: zach.a.greenberg@gmail.com
Github: @zachagreenberg
LinkedIn: https://www.linkedin.com/in/zachary-greenberg-496291208/