

Wine Recommendation Engine

By: Zachary Greenberg



Overview

- Business Problem
- Data Collection
- Exploratory Data Analysis
- Modeling
- Results
- 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
- 1200 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

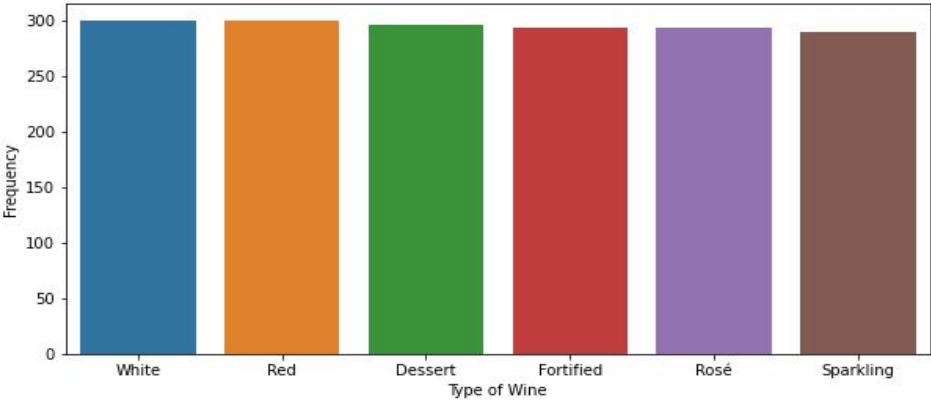


Exploratory Data Analysis

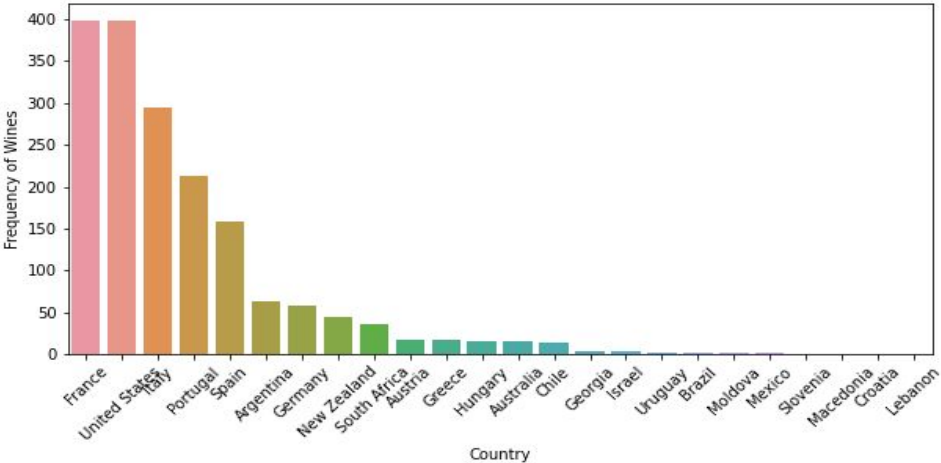
- Wines by Type/Country/Winery:



Wines by Type



Wines by Country



Wines by Winery

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

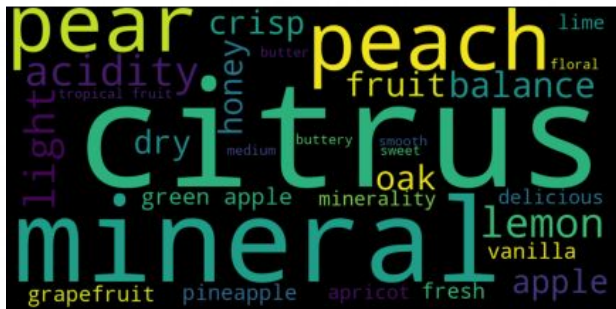


EDA continued....

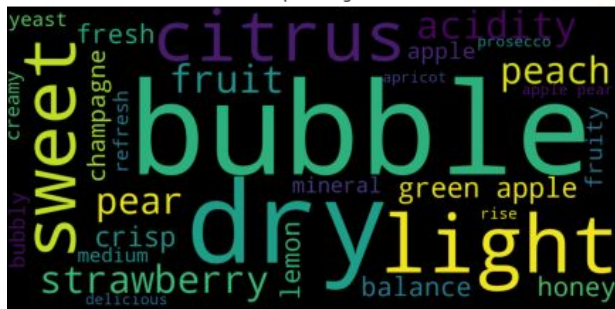
- Word Clouds of Text Reviews



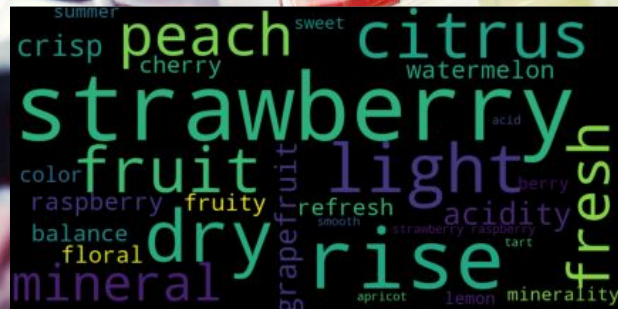
White



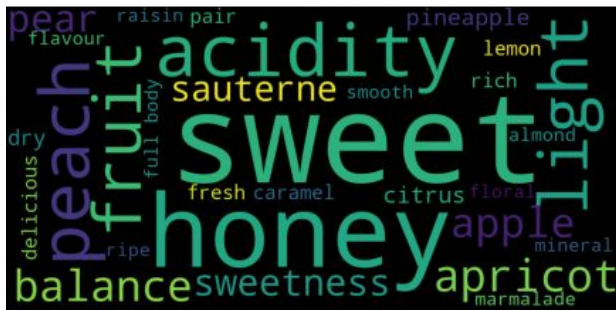
Sparkling



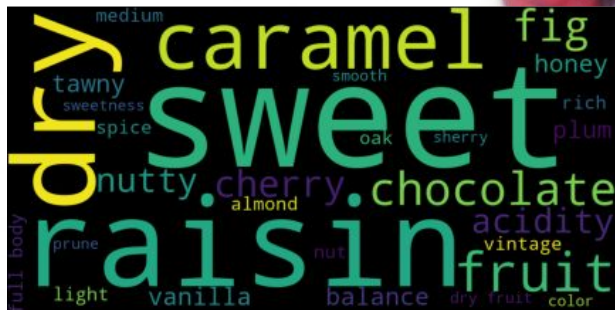
Rosé



Dessert



Fortified

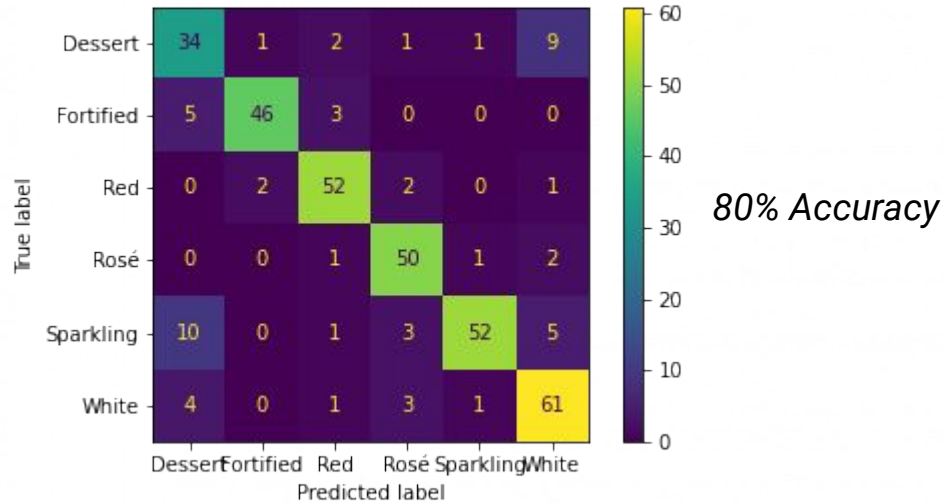


Red



Modeling

- MultiNomial NB Classification Algorithm



- Recommender Systems

Cosine Similarity & Euclidean Distance



Results

- Rec System by Wine Name using Cosine Similarity



Lloyd Chardonnay
2016

Lloyd Chardonnay 2019

	wine_type	wine_country	grapes	average_rating	wine_price	similarity_scores
title	Cosine Similarity					
Tyler Chardonnay 2018	White	United States	Chardonnay	4.1	27.59587	0.799311
Matthiasson Linda Vista Vineyard Chardonnay 2018	White	United States	Chardonnay	4.1	34.99000	0.748776
Walter Hansel The Estate Vineyard Chardonnay 2018	White	United States	Chardonnay	4.2	29.99000	0.746672
Mer Soleil Reserve Chardonnay 2012	White	United States	Chardonnay	4.1	27.99000	0.740810
Brioche Chardonnay 2017	White	United States	Chardonnay	4.2	21.99000	0.740417
Brick House Chardonnay 2015	White	United States	Chardonnay	4.2	30.59000	0.730619
Jordan Chardonnay 2017	White	United States	Chardonnay	4.1	36.99000	0.728534
Gary Farrell Russian River Selection Chardonnay 2017	White	United States	Chardonnay	4.2	32.99000	0.725776
Alexana Terroir Series (Terroir Selection) Chardonnay 2016	White	United States	Chardonnay	4.2	29.99000	0.724258
Tyler Chardonnay 2016	White	United States	Chardonnay	4.1	29.99000	0.717020

	Lloyd Chardonnay 2019	Cosine Model	Euclidean Model
Average Rating	4.30	4.150000	4.140000
Wine Price	37.99	30.310587	29.910587



Next Steps

- Retrieve more wine data
- Use LDA for Topic Modeling
- Create interface for user interaction



Thank You!

Email: zach.a.greenberg@gmail.com

Github: @zachagreenberg

LinkedIn: <https://www.linkedin.com/in/zachary-greenberg-496291208/>

