

Become A Sommelier

Using Topic Modeling to Investigate Wine Style and Structure



Goals

Use Topic Modeling(Specifically NMF) to investigate the style and structure of Wines

Use the results of my topic model as inputs for a content based recommendation system

Data

The data was downloaded from kaggle in csv form. But originated from winemag.com and was kindly scraped by Zack Thoutt.

There are 280K rows of unique wines and 10 feature columns. Some of the columns are Country , Description, Price, Points, Province, Region, and Wine Name.

I scraped an additional 300 rows of Alcohol By Volume Percentages from winemag.com corresponding to the Wine Names of the first 300 rows in the original data set.

Topic Investigation



Topic 2

White Wine



White Peach Flavor



Topic 1

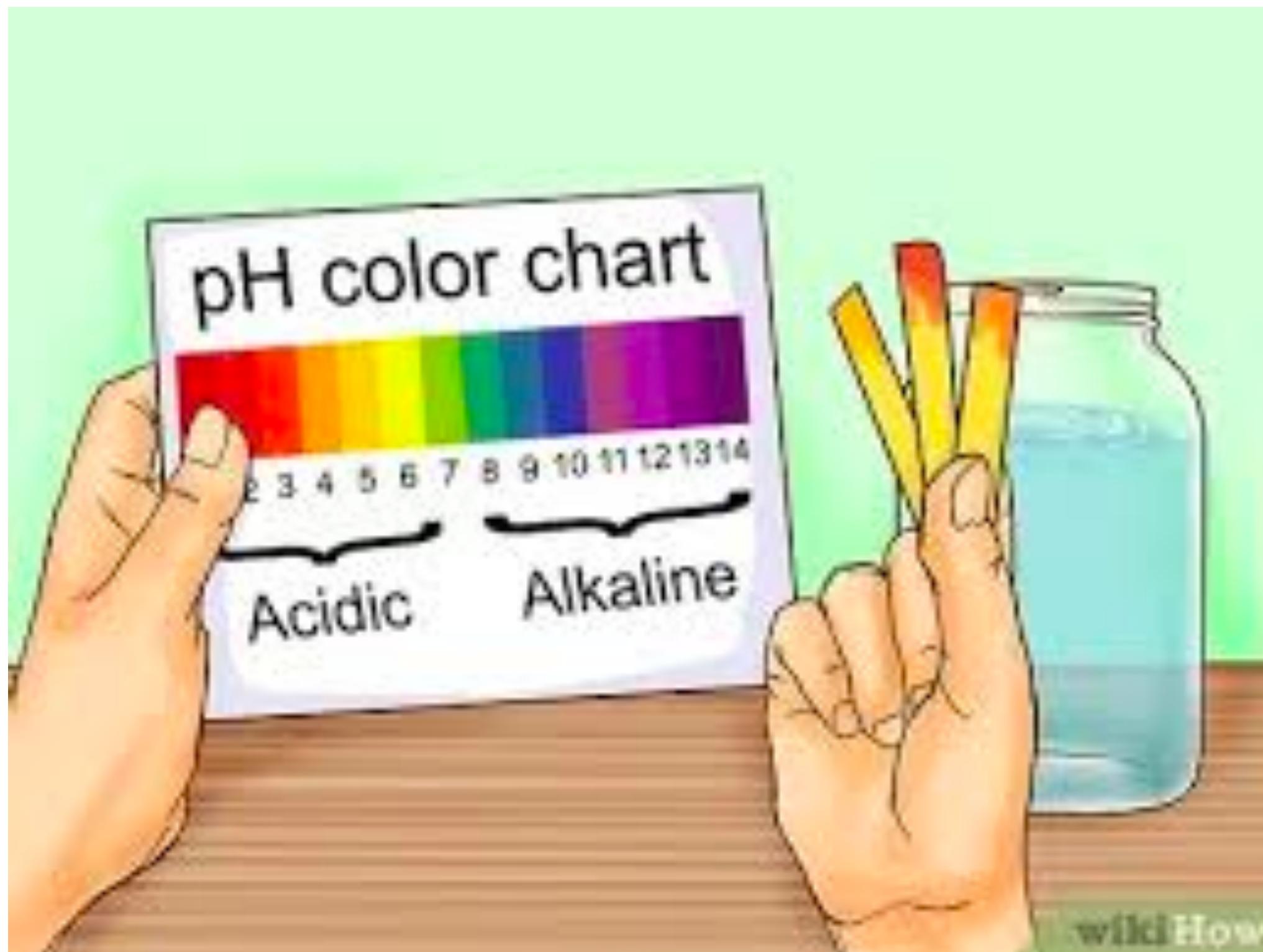
(Sample Description)

Floral aromas of **white** spring flowers partner hints of pear and stone fruit. The juicy palate offers **white peach**, apple and green melon alongside a mineral note and fresh acidity.

White Blend

Topic 5

Acidity



Light(body) Reds



Topic 5 (Sample Description)

This is a **soft** fruity wine. It has attractive red-fruit flavors, very ripe **fruit** and **light acidity** at the end. The wine is ready to drink.

Portuguese Red

Topic 4

Oak Barrel, Oaky Flavor



Red Blend



Topic 4 (Sample Description)

Here's a basic red wine, full-bodied and dry, with ripe **blackberry**, cherry, currant and **oak flavors**. The blend is Merlot-dominated, with Cabernet Sauvignon and Petit Verdot.

Bordeaux-style **Red Blend**

Wine Comparison (Stock Descriptions)



This is a highly **herbal** Cab with strong aromas of juniper and pyrazines. The palate feels full but tannic, while leafy, dry, roasted

berry flavors finish with baked

notes and a lasting **herbal** flavor.

Price: 15.0

Wine Name: Santa Alicia 2011 Gran Reserva de Los Andes Estate Bottled Cabernet Sauvignon (Maipo Valley)

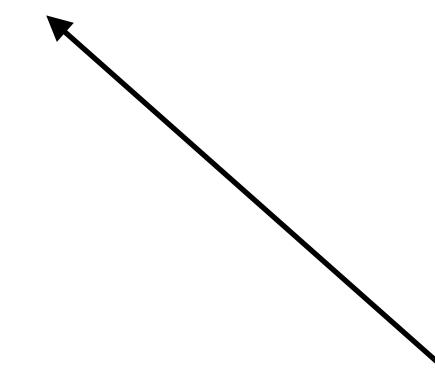


Grassy red **berry** and plum aromas lead to a lean, somewhat high-acid palate with

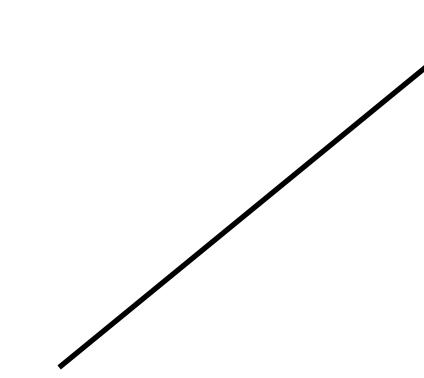
herbal, oaky flavors of raspberry and plum that don't last that long. Fresh and healthy feeling on the finish, but tasting more of oak and spice than fruit.

Price: 13.0

Wine Name: Maipe 2010 Andean Culture Malbec (Mendoza)



Wine of Comparison



Most Similar Wine

Wine Comparison (Stock Descriptions)

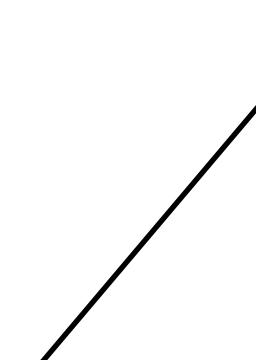
This wine demonstrates Viognier's exotic side, with flavors of tropical fruit, tangerine, **white** flower, vanilla, honey and spice.

Despite the richness, it's **dry**

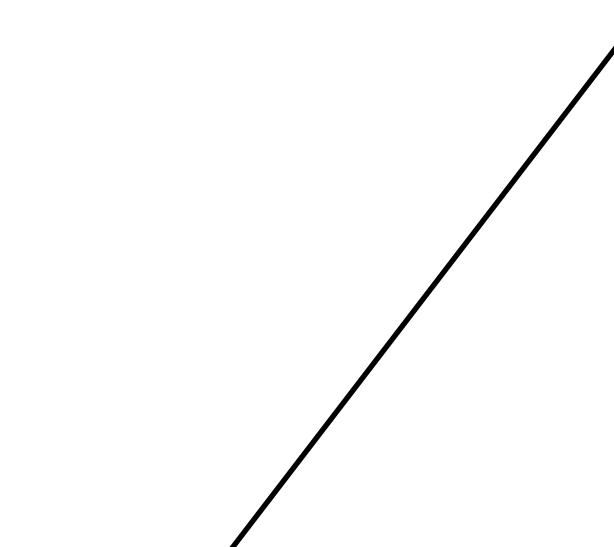
and the **acidity** is mouthwatering. Don't serve this too cold, and pair it with Asian fare that's slightly sweet.

Variety: Viognier(White)

Wine of Comparison
ABV: 14.5%

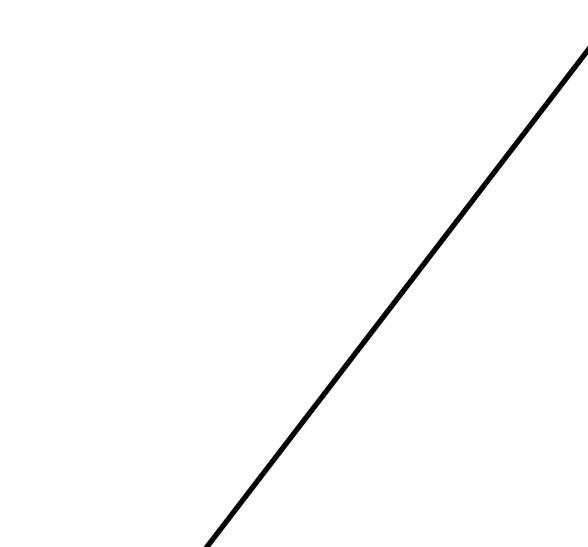


Similar Wine A
ABV 12%

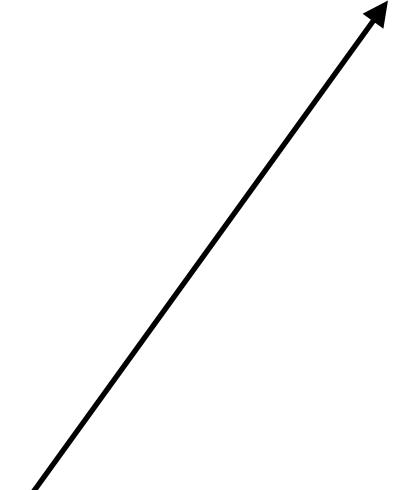


Plump apricot, melon and **white** peach abound are luscious. Yet it's

dry in style and bracing too, edged by taut **acidity** and a steely, mineral edge. Accessible and elegantly balanced, its a four-square wine with excellent value.
Variety: Riesling(White)



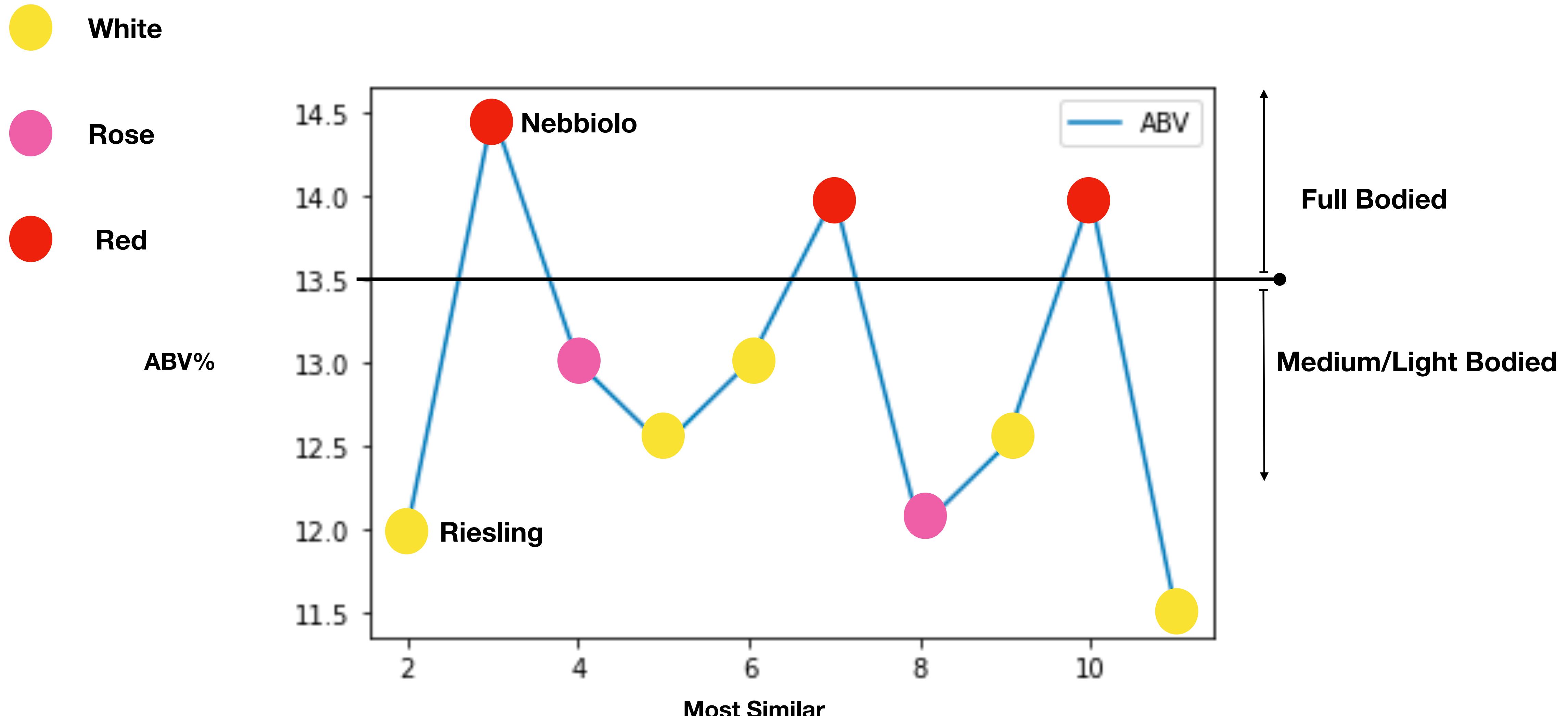
Similar Wine B
ABV 14.5%



This intensely fragrant wine opens with violet, woodland berry and star anise aromas. The racy palate offers sour cherry, cranberry and ground clove flavors framed by taut tannins and brisk

acidity, feeling just underripe.
Variety: Nebbiolo(Red)

ABV% Analysis



Wine Comparison (Augmented Descriptions)



Earthy in character, this is a rustic wine. It does have attractive red **fruits**, and balancing **acidity** but these elements come up against a barnyard character and firm **tannins**. This Wine is from

France in Southwest France and from the Cahors Region and is a Malbec Price: 13.0 Wine Name: Château Nozières 2011 Malbec (Cahors)



Balanced and fruity, this is a crisp wine that has soft **tannins**. With its bright black-currant **fruits**, **acidity**, and touch of spice, it is ready to drink. This Wine is from

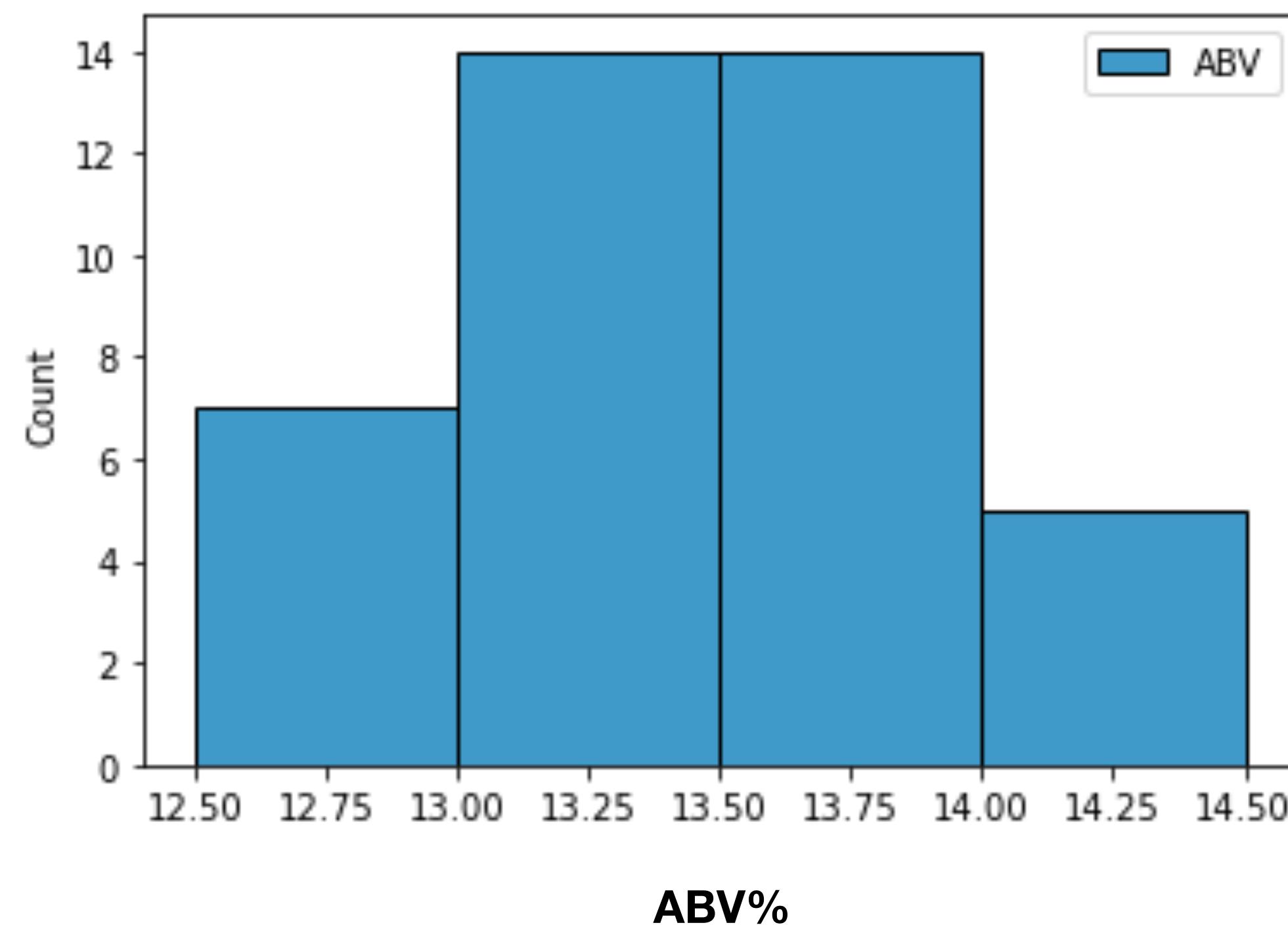
France in Bordeaux and from the Bordeaux Region and is a Bordeaux-style Red Blend Price: 13.0 Château Haut Mouleyre 2015 Bordeaux

Wine of Comparison
(14.5% ABV)

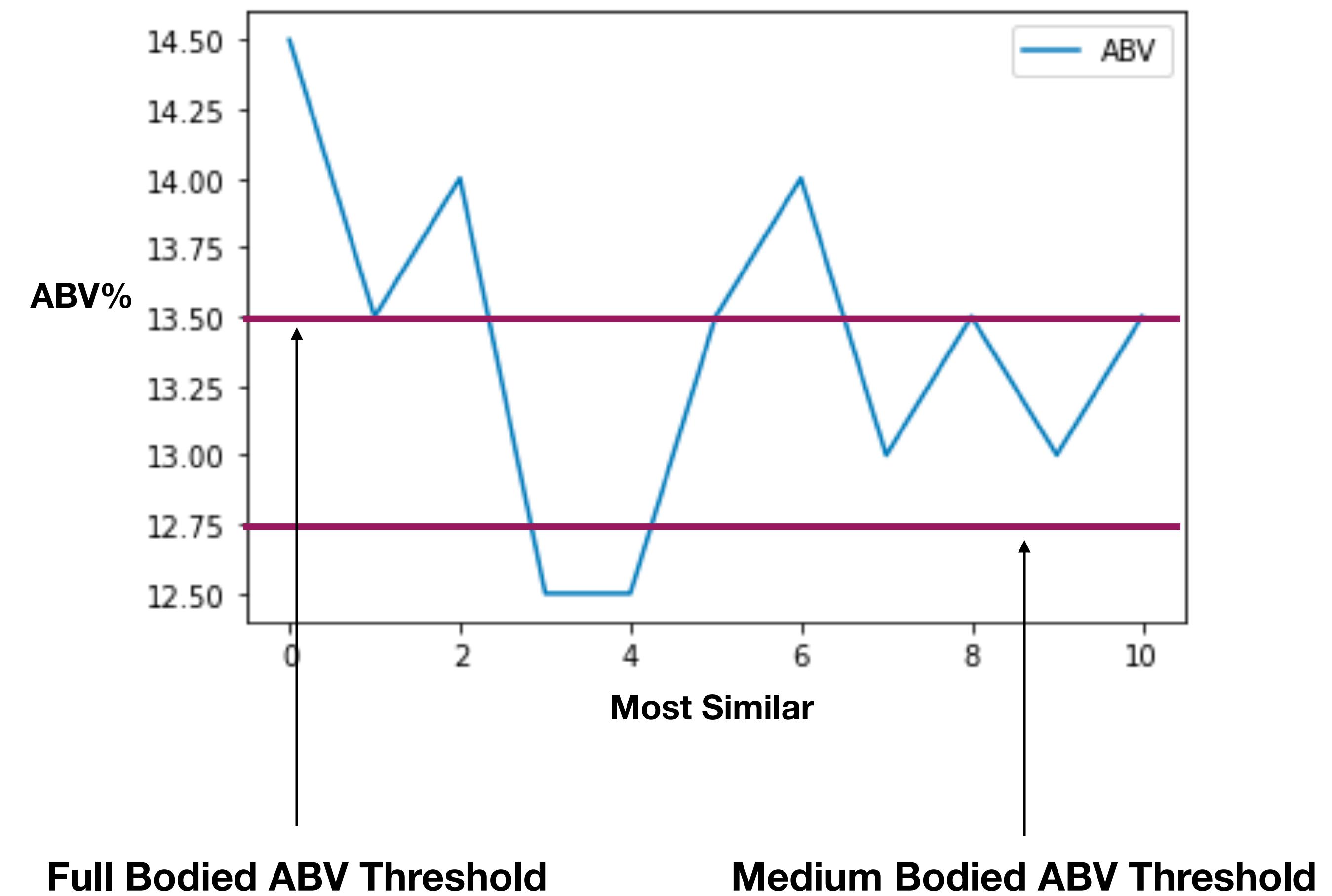
Similar Wine
(14% ABV)

ABV% Analysis

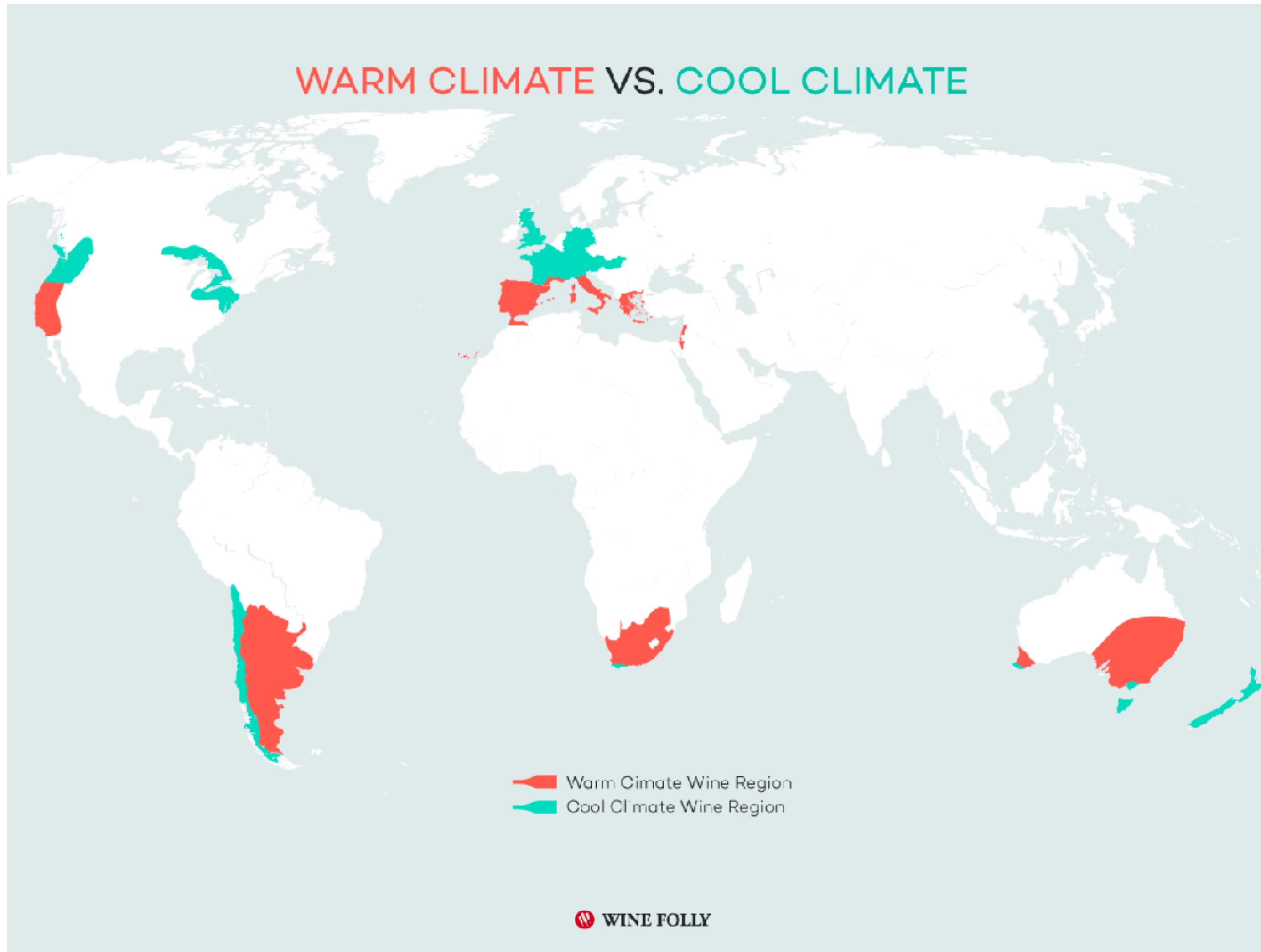
ABV% DIST TOP 40



ABV% TOP 10 Wines



Climate Impact on Wine Body



APPENDIX

- 1. SUBJECTIVE INTERPRETATION OF WINE DESCRIPTIONS**
- 2. TRY DO DRAW SIMILARITY BASED ON INSTITUTIONAL KNOWLEDGE**

Stock Descriptions

(Only Includes Wine Style in Corpus)

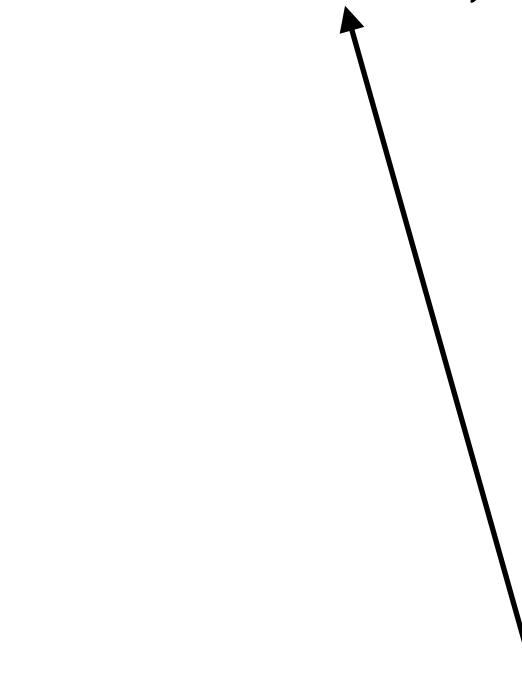


A blend of 85% Nebbiolo and 15% Uva Rara, this fragrant red boasts alluring scents of pressed rose, moist earth, **aromatic herb** and whiff of exotic spice. The racy straightforward palate doles out tart **red cherry**, **star anise** and a **hint of clove** alongside fine-grained tannins and firm acidity.

Price: 31.0

Points: 88

Wine Name: Podere ai Valloni 2014 Gratus Bio Red
(Colline Novaresi)



Wine of Comparison

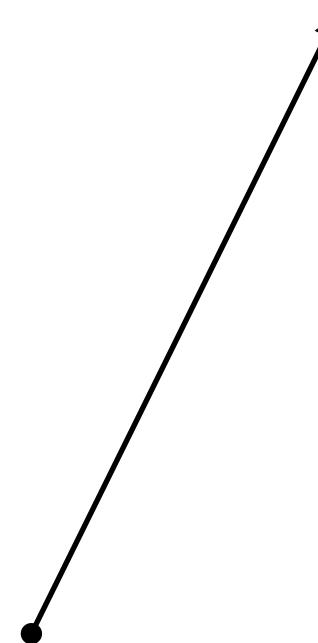


Spicy and dry on the nose, with red cherry, plum and candied sweetness. The palate has good body and feel, **with medicinal plum and berry** flavors along with an olive accent. Juicy and forward, with a **lightly herbal** finish.

Price: 10.0

Points: 86

Wine Name: Vistamar 2009 Sepia Reserva
Cabernet Sauvignon (Maipo Valley)

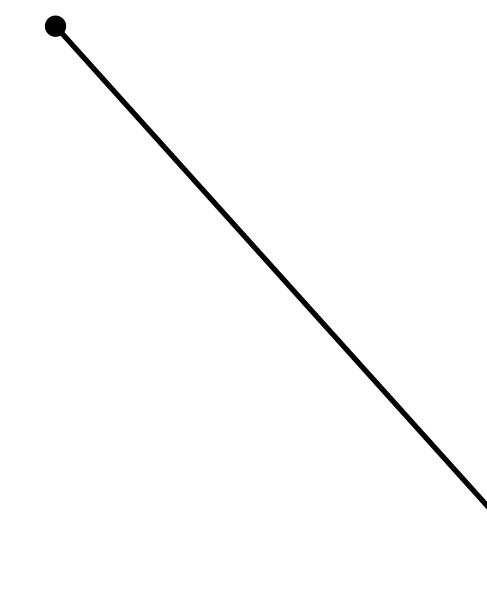


Fourth Most Similar

Taste Profile Similarities

Pancakes made with Star Anise

Star Anise



Plum



Is very sweet and licorice-like....

Medicinal Plum Flavor

Taste Profiles

Pinot Grape



Pinots are traditionally medium bodied wines
With lower alcohol content

Cabernet Sauv Grape



Cabernet Sauvignons are traditionally full bodied wines
With higher alcohol content

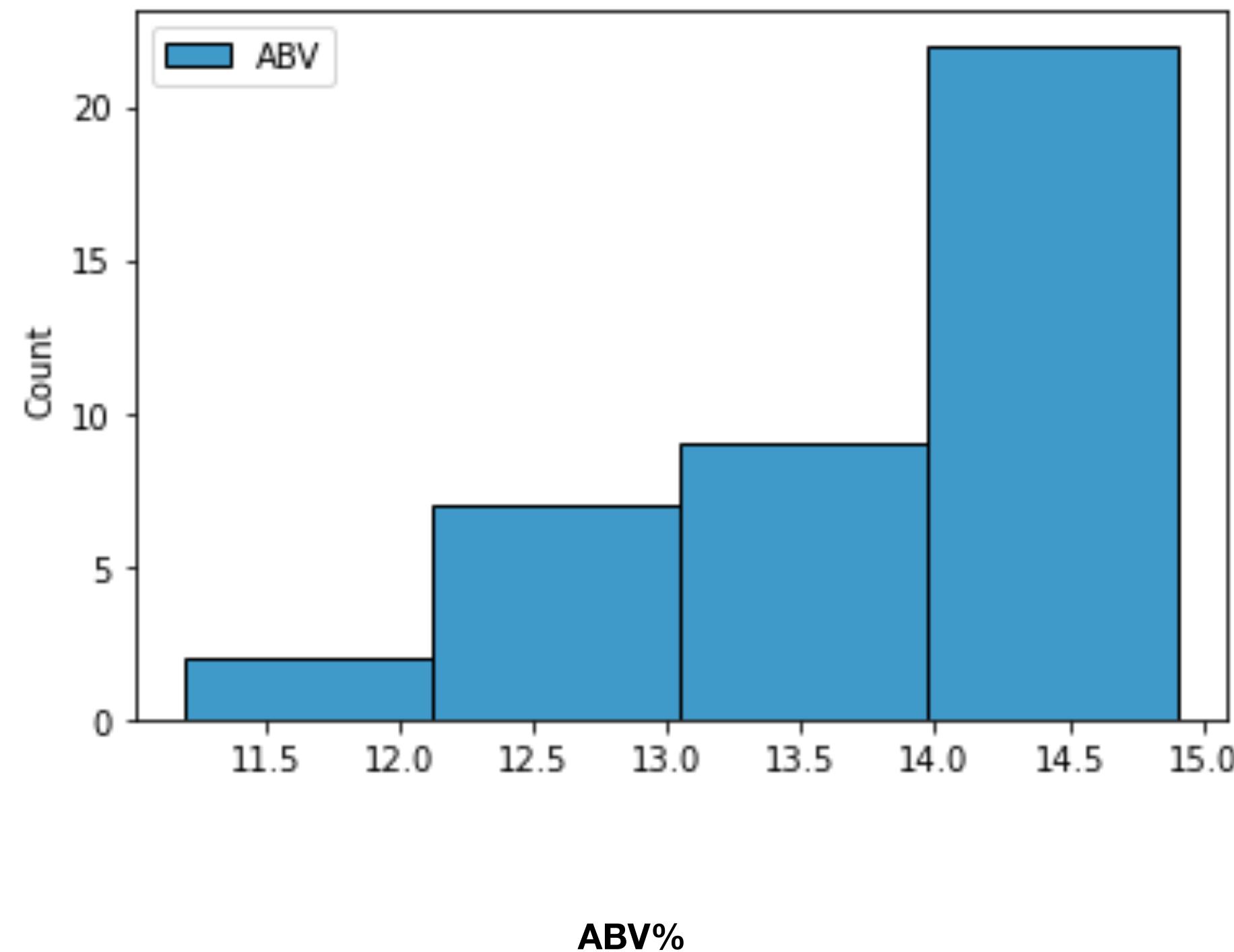
Very Similar... But why ?

AVB: 14.3%

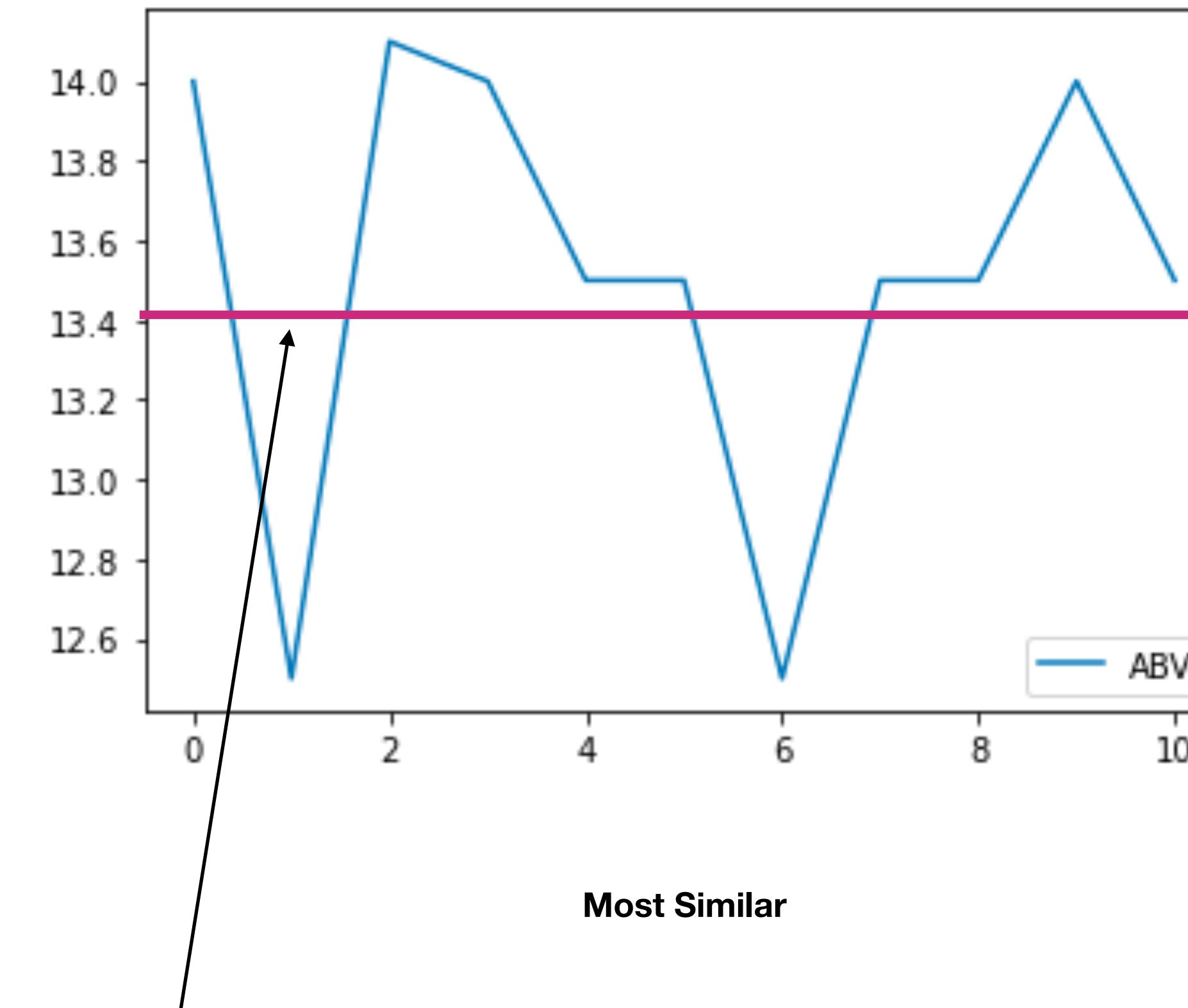


AVB: 14.58%

ABV% DIST TOP 40



ABV% TOP 10 Wines



Full Bodied Threshold