



SWIRL. SMELL. SIP. REPEAT!

WINE ADVISER



A Journey Through the Grapevine in
My Capstone Project

BY MARY PETRICHENKO



Data

Wine Advisor

- Source of data - Kaggle dataset in CSV format
- ~130k wine reviews published on WineEnthusiast
- Dataset uploaded 3 years ago
- Data includes:

Wine title

\$ Price

Country

Review

Region, province

Points

Variety

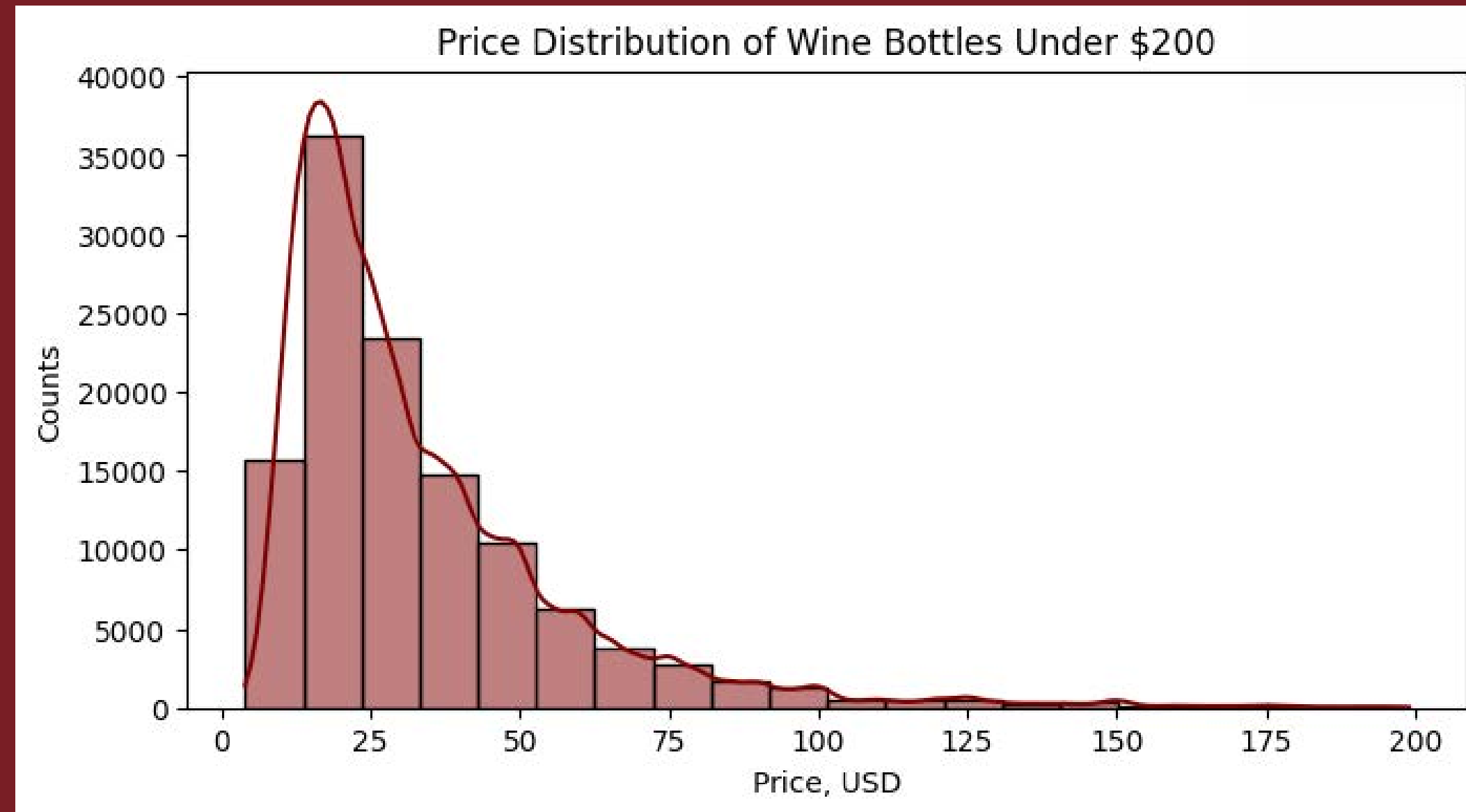
Taster name

Winery

Taster twitter account

Prices

Wine Advisor



The price range in the dataset is from \$4 to \$3.300 per bottle.

The majority of wines cost less than \$40.



Points

Wine Advisor

Classic - 98-100:

The pinnacle of quality.

Superb - 94-97:

A great achievement.

Excellent - 90-93:

Highly recommended.

Very Good - 87-89:

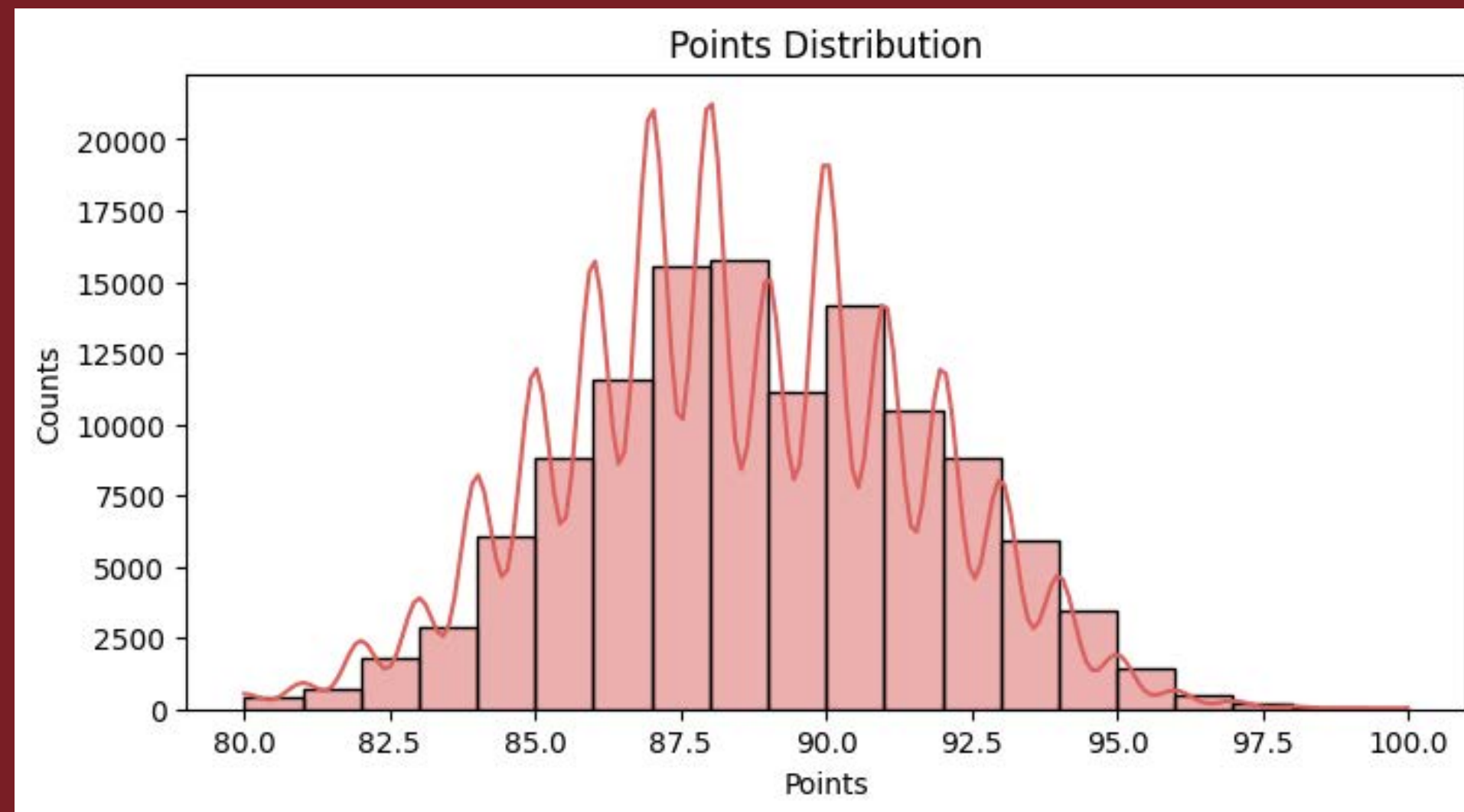
Often good value; well recommended.

Good - 83-86:

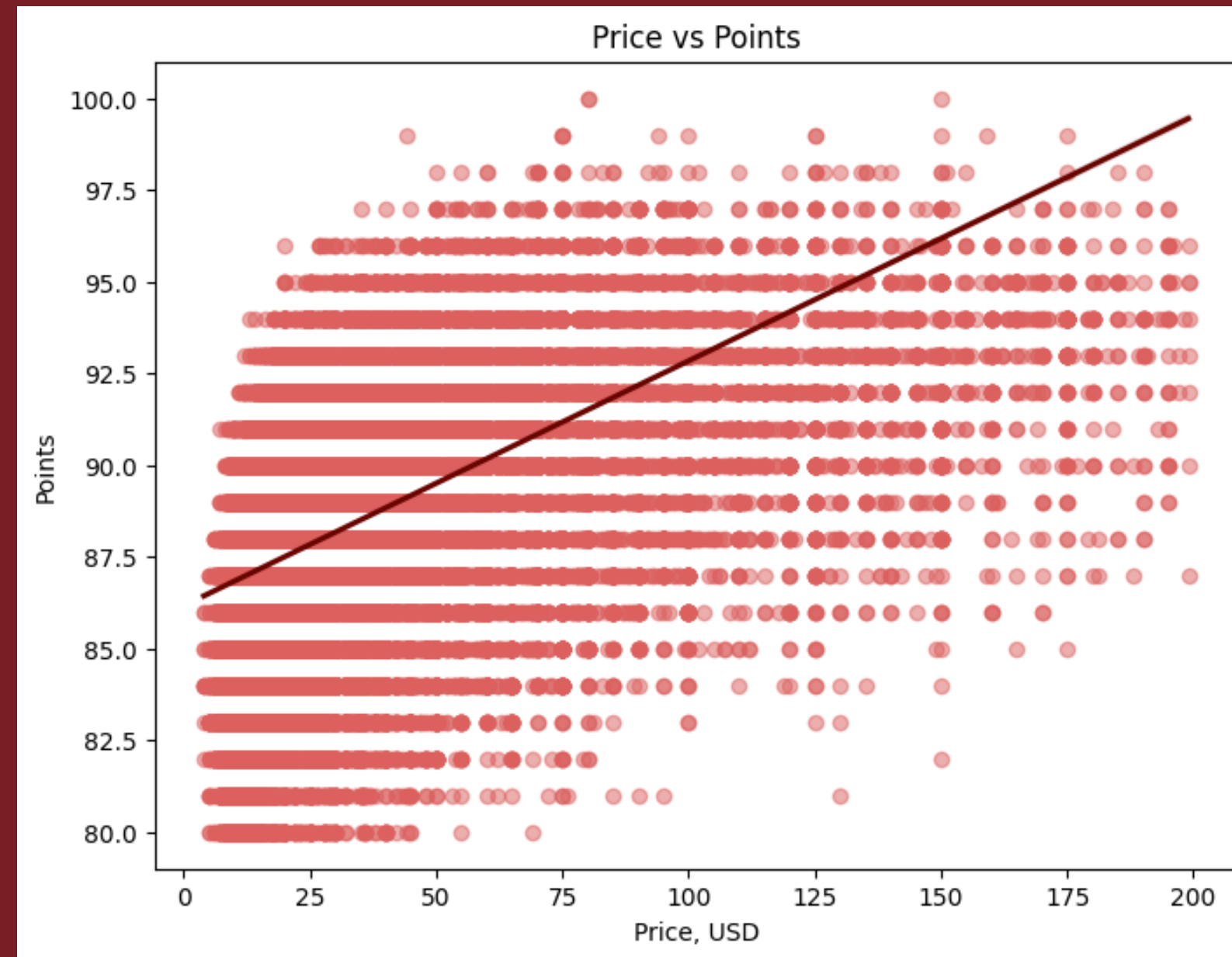
Suitable for everyday consumption; often good value.

Acceptable - 80-82:

Can be employed in casual, less-critical circumstances.



Points vs Price *Wine Advisor*



A positive correlation between wine quality, measured by points, and price can be seen. Generally, wines with higher prices tend to exhibit better quality, as reflected in their higher points.


Wines with points less than 83, which were defined as "acceptable", are more prevalent in the price range of less than \$75 per bottle.

Reviews

Wine Advisor

Wine reviews offer consumers valuable insights into a wine's characteristics, quality, and overall experience. Descriptions typically include details on aroma, taste, and mouthfeel, with descriptors like fruity, floral, spicy, oaky, or earthy notes.

Here is the example of the wine description from the dataset:



“Here's a bright, informal red that opens with aromas of candied berry, white pepper and savory herb that carry over to the palate. It's balanced with fresh acidity and soft tannins.”

white
pepper

savory herb

candied
berry

Reviews & types

Wine Advisor

Grouped
in types as:

- red wines
- white wines
- rose wines
- sparkling

Word Cloud for Red Wines



Word Cloud for White Wines



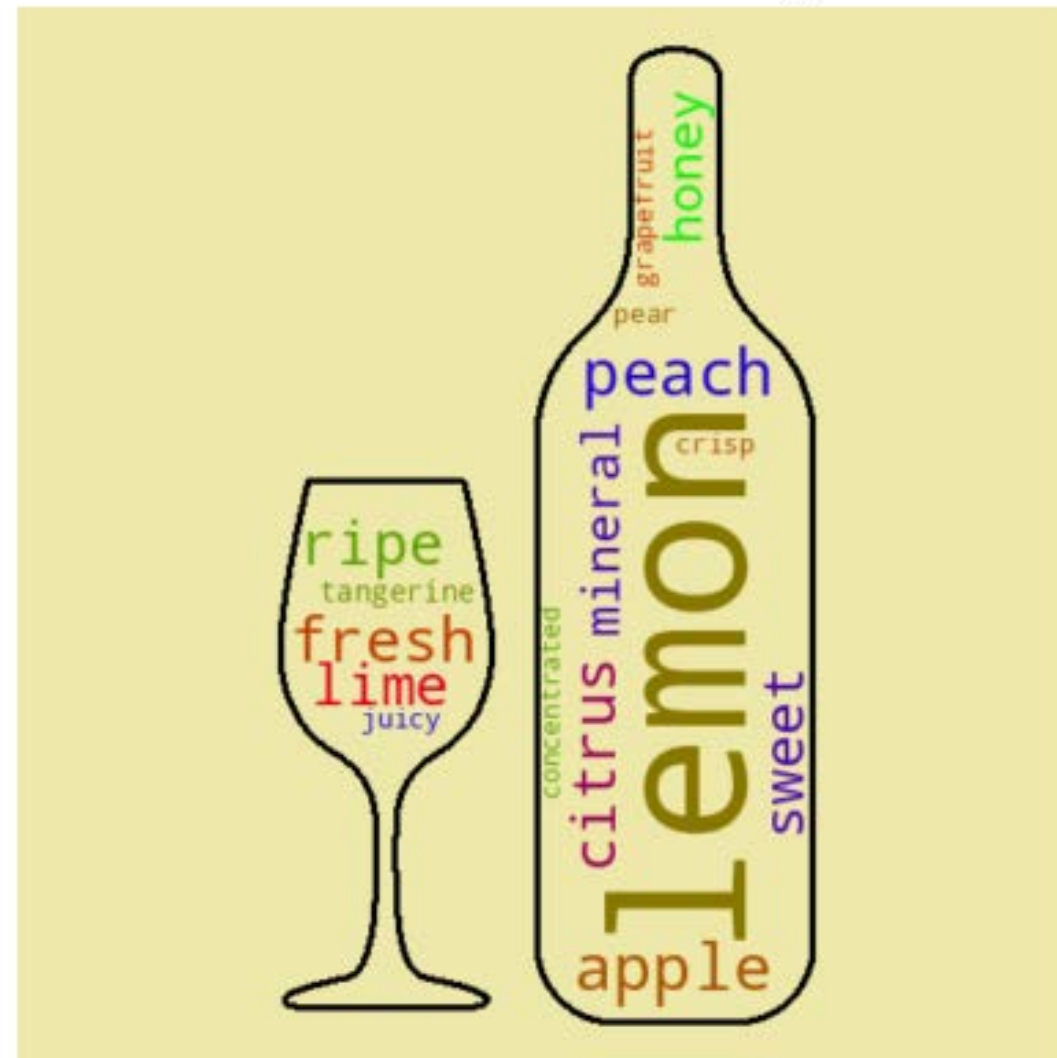


Reviews & varieties

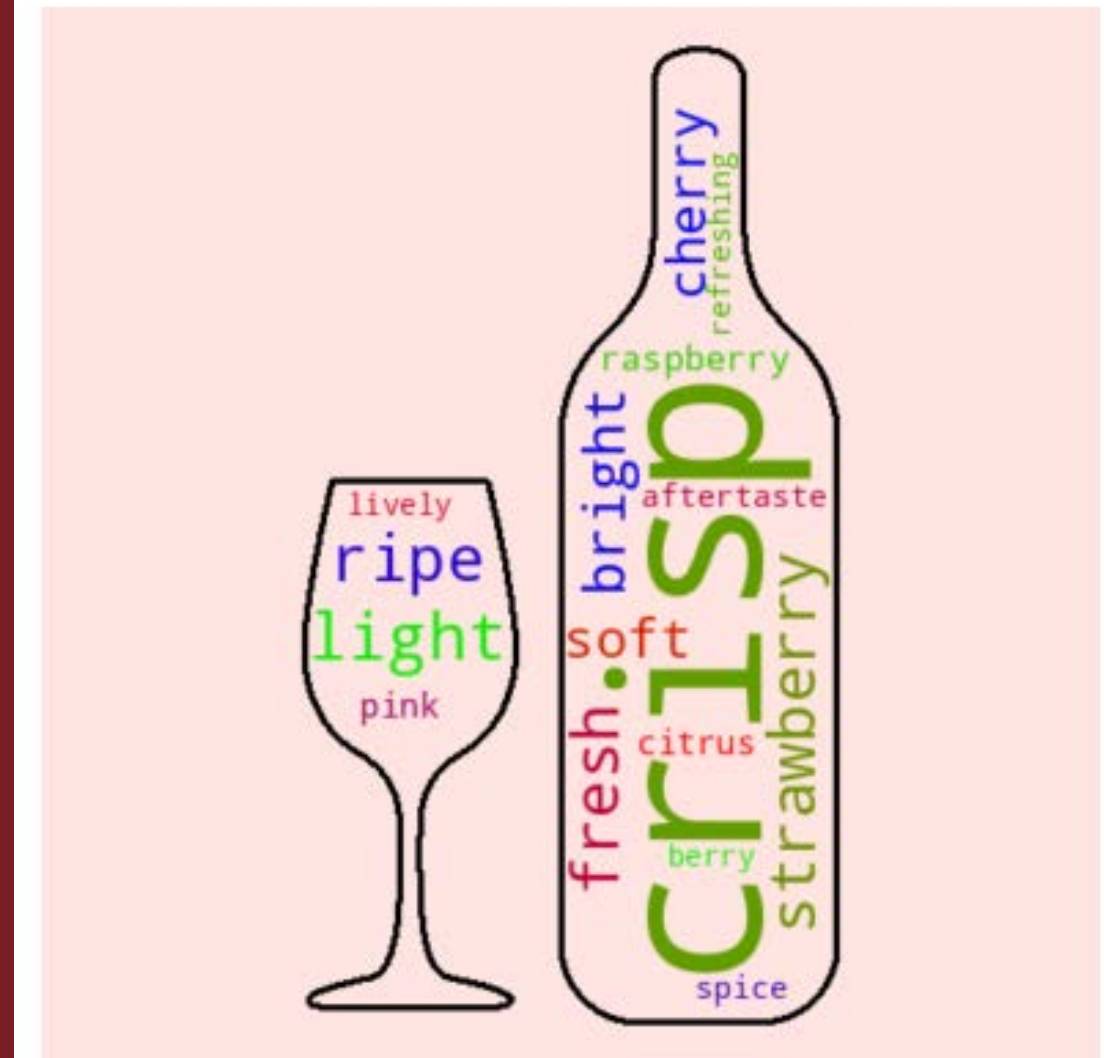
Wine Advisor

The dataset allows us to explore how each wine variety possesses its own unique set of frequently observed words in its description.

Word Cloud for Riesling



Word Cloud for Rosé

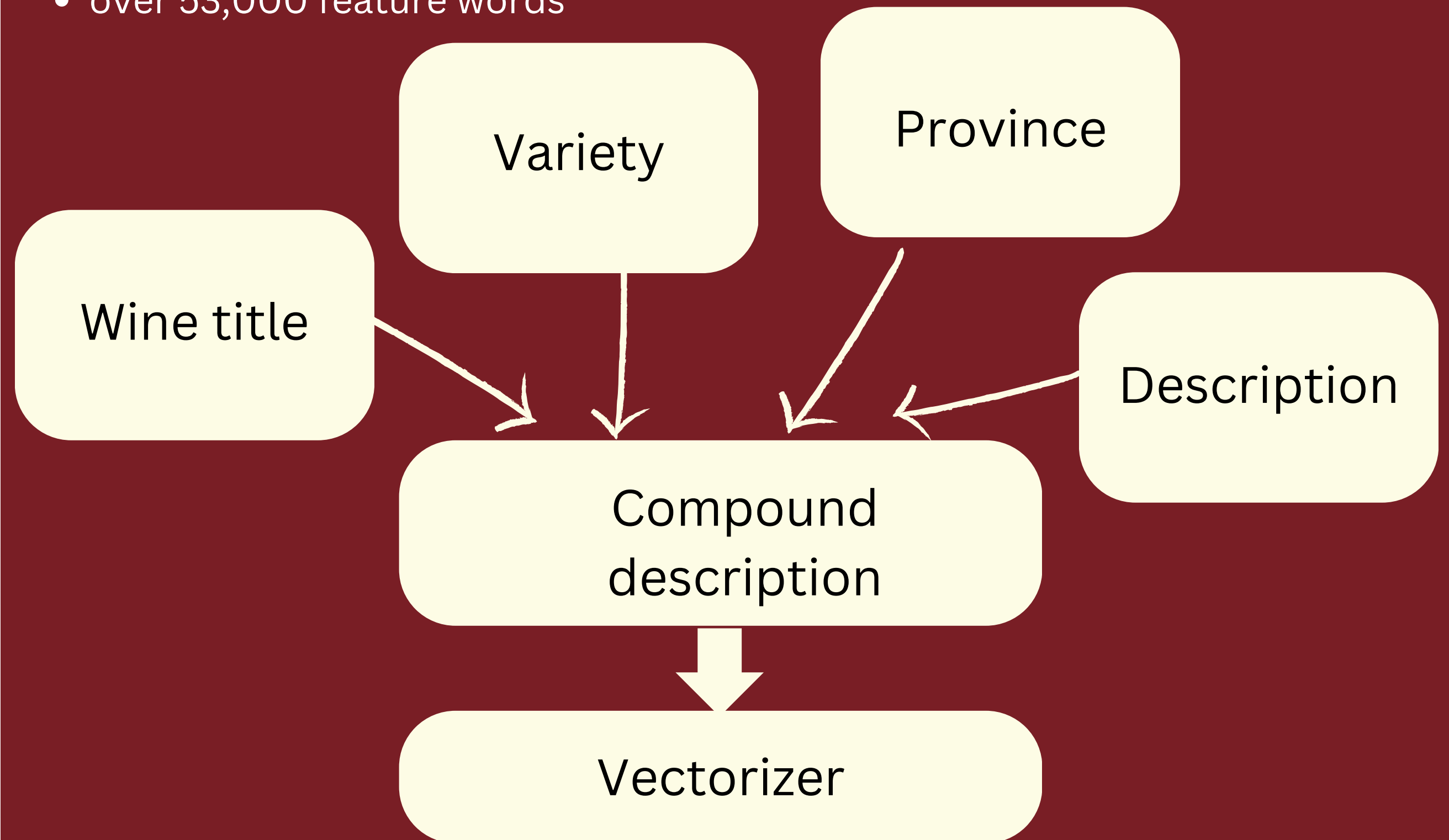




Vectorizer

Wine Advisor

- TF-IDF Vectorizer,
- analyzing by single words,
- fitted with the compound description of each wine
- over 53,000 feature words





Matrix & Linear kernel *Wine Advisor*

- Incorporating a feature for optional filtering by price and wine type
- Dividing the entire dataset into manageable chunks of less than 10,000 rows to tackle memory constraints
- Generating a matrix for each chunk of data: learning vocabulary and IDF, returning document-term matrix.
- Applying a linear kernel to each matrix
- Determining the best matching wine based on the user's prompt within each portion of the data
- Presenting the output of the best matching wines as a sorted dataframe

Streamlit

Wine Advisor

Please follow the link:

<https://wine-selector.streamlit.app>

Welcome to the Wine Advisor



Uncork the Perfect Wine Experience with Wine Advisor:
Your Palate's Personal Sommelier!

- ☐ White wines
- ☐ Red wines
- ☐ Rosé wines
- ☐ Let's go bubbles (Sparkling, Frizzante, Prosecco, Champagne...)

Minimal price of the wine

0 - +

Maximal price of the wine

100 - +

Price of the wine

0 100

To get the output you need to describe the taste of wine, you may specify the type of wine and a price range. After hitting the button the search begins.

This app utilizes data from Wine Enthusiasts.

I acknowledge and appreciate their contribution to enriching my app's content.



WineSearch

Wine Advisor

Prompt with the words: strawberry crisp fresh.
Price range - less than \$33.

The 3 top results were:

Miradou 2010 Rosé (Côtes de Provence)

Rosé (Rose)

Price: 12.0\$ (85 points)

Match: 34.9%

Very light in flavor and texture, the wine has hints of strawberry flavors with acidity and a crisp, clean character.

Vaza 2008 Crianza (Rioja)

Tempranillo (Red)

Price: 15.0\$ (87 points)

Match: 34.1%

This is roasted, savory and spicy up front, with solid berry aromas. It feels crisp and fresh, with a blend of cherry, strawberry and herb flavors. It's spicy, lively, fresh and clean on the finish.

Bott Frères NV Rosé Brut Sparkling (Crémant d'Alsace)

Sparkling Blend (Rose)

Price: 22.0\$ (86 points)

Match: 32.2%

Fresh and fruity, this is a refreshing, lively rosé. It is dry with strawberry flavors and a crisp acidity on the finish. Drink now.

Thank you for listening

