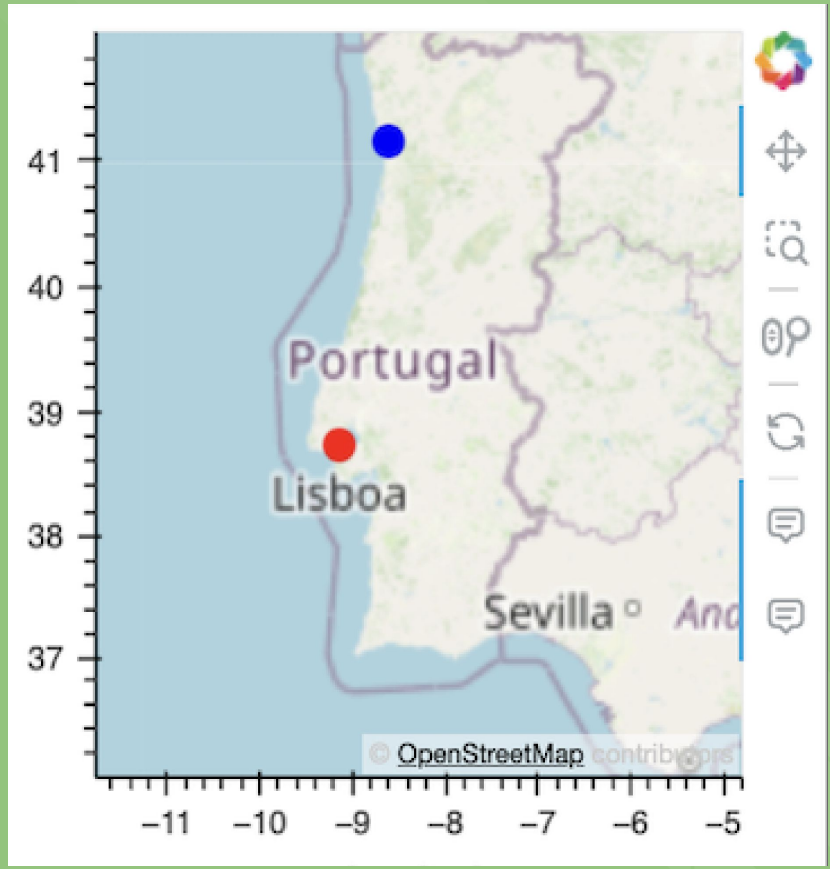
Wine Quality Analysis

Greg Labavitch, Varsha Sharma, Jonathan Gilbert, Brent Lang

Region of Northern Portugal

The region’s climate and grape varieties contribute to the unique taste of this young and slightly sparkling wine.



Overview

Wine Quality – Vinho Verde (Green Wine)

The Vinho Verde region is diverse and versatile in wine styles and profiles, known for producing not only light and fresh wines, but also complex, structured, and mineral wines.

The data that we are working with contains information collected from performing physicochemical tests on several thousand samples of wines from the Vinho Verde region of northern Portugal.

The data set consists exclusively of numerical data and is divided into two sets: red wine and white wine. The variables that we will be focusing on are fixed acidity, volatile acidity, and chloride levels.

.

Relevant Wine Definitions

01

Fixed Acidity

Fixed acidity is the naturally produced acids of the wine. The acids include tartaric, malic, citric, succinic, and lactic.

02

Volatile Acidity

Volatile acidity is the measurement of gaseous acids. Acetic acid is the primary volatile acid in wine and is associated with the smell and taste of vinegar.

03

Chlorides

Chloride indicates the level of sodium in the wine. The greater amount of sodium chloride is expressed through added saltiness. The chloride content is used to account for quality.

Project Objective

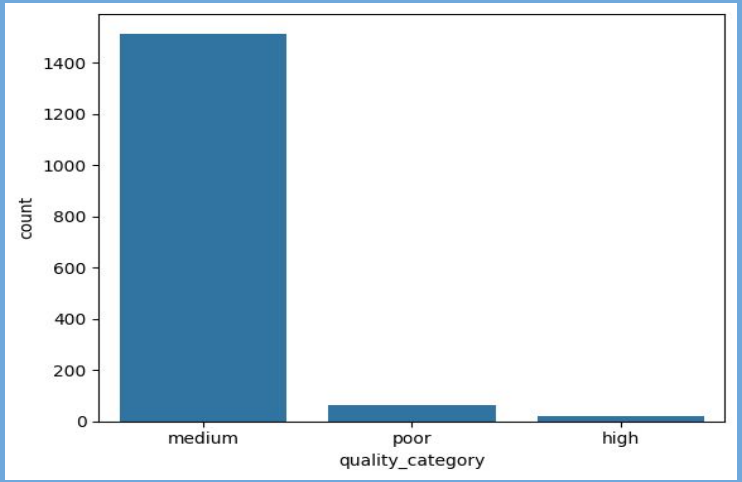
Analyze the quality of red and white wine

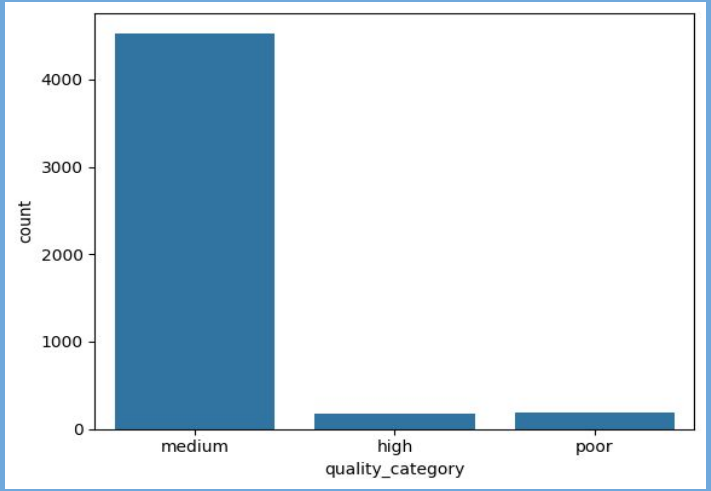
Compare overall quality

Analyze how wine quality is affected by fixed acidity, volatile acidity, and chloride levels

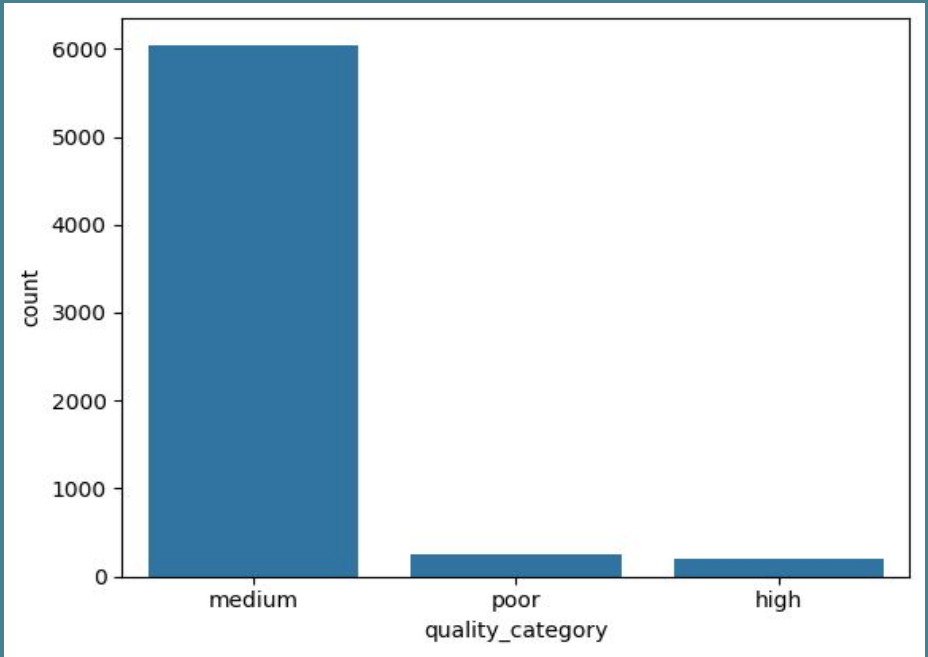
Bar Chart: Red (left) and White (right)

Quality category is based on quality score





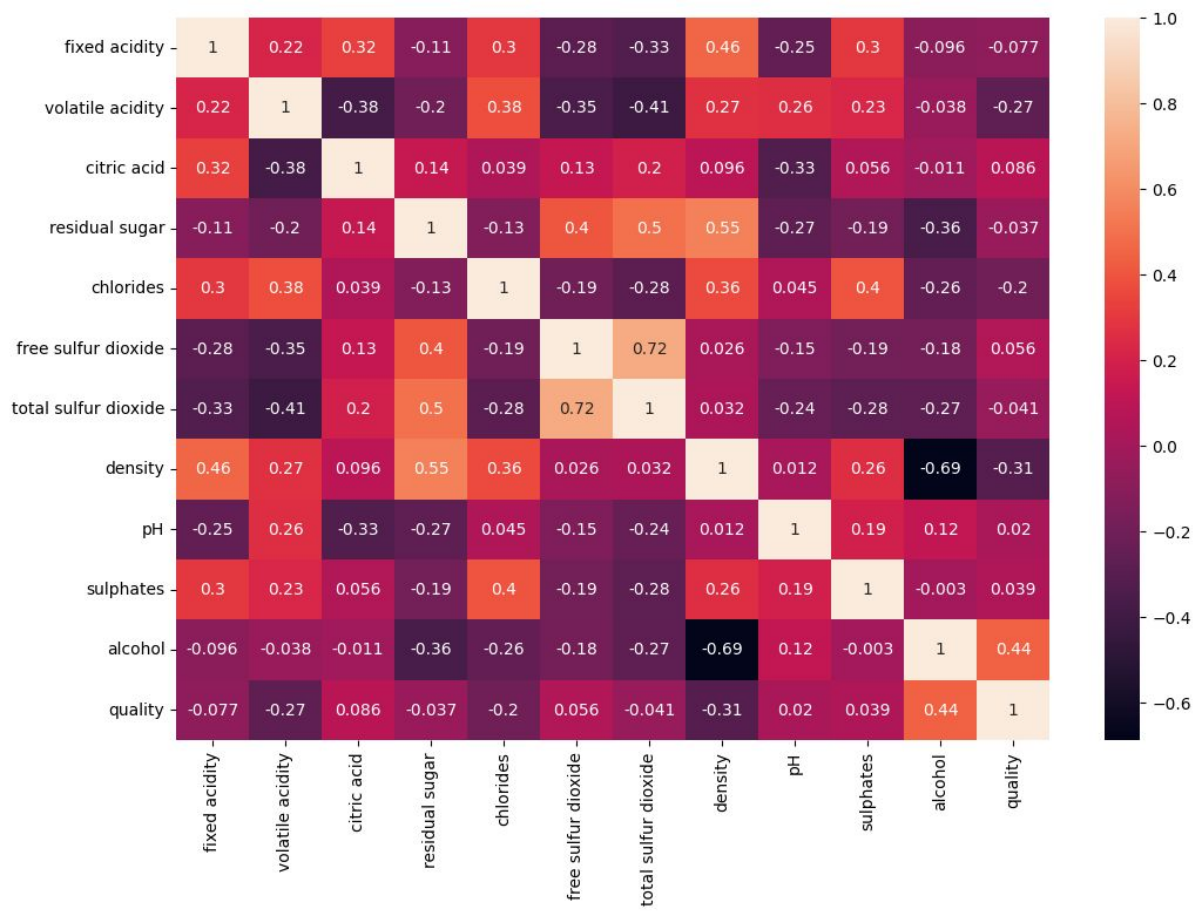
Bar Chart: Red and White Overall



Heat Map

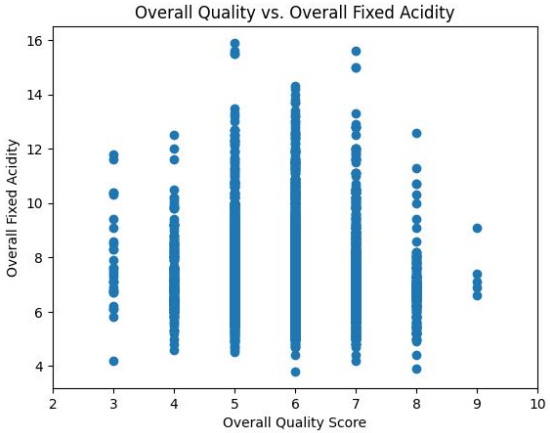
The heat map shows very low correlation between our chosen variables and quality. The overall quality across the dataset heavily favors medium quality. Because of the low numbers of poor and high quality wines, the charts and those that follow may be unreliable for determining correlation.

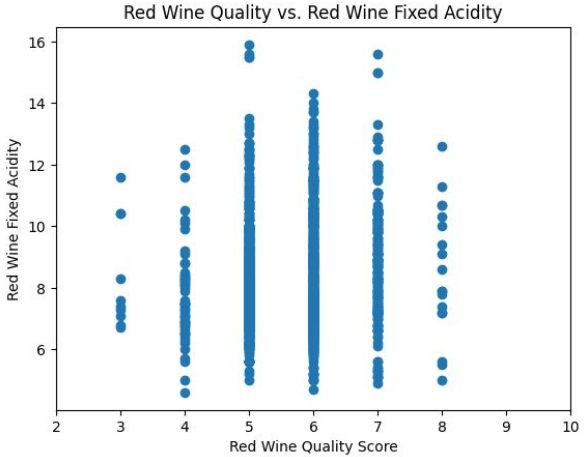
The correlation between alcohol content and quality is often observed, this two variables has good correlation.

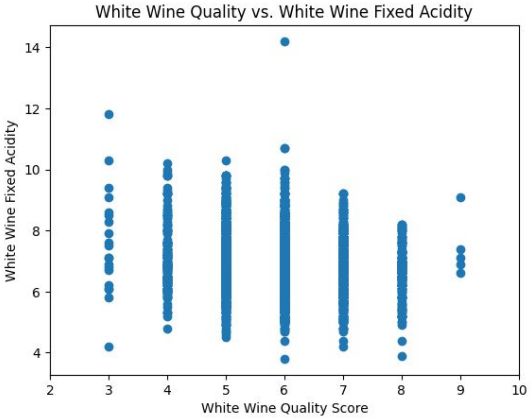


Fixed Acidity

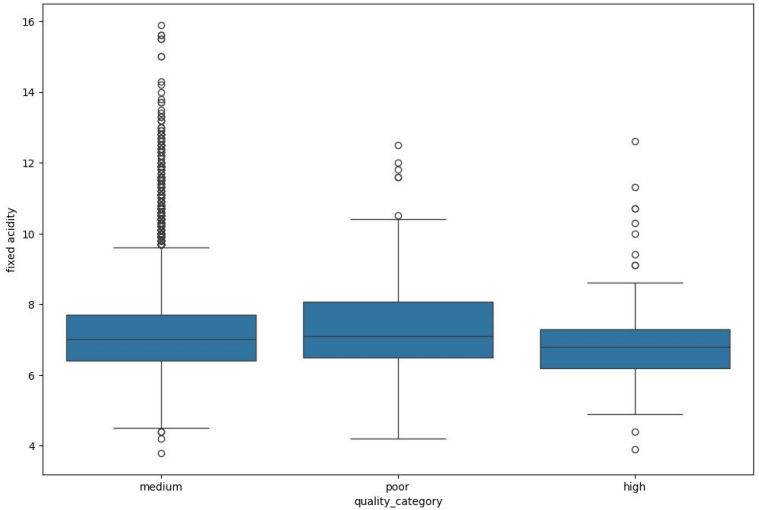
There does not appear to be much correlation between fixed acidity and quality, except maybe with white wine. This is so even if we ignore wines with scores of 3,4, 8, and 9.





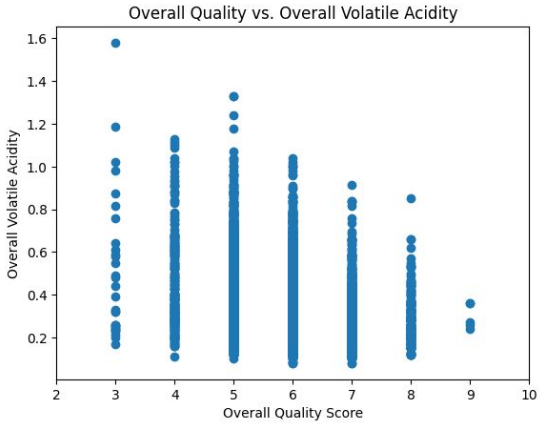


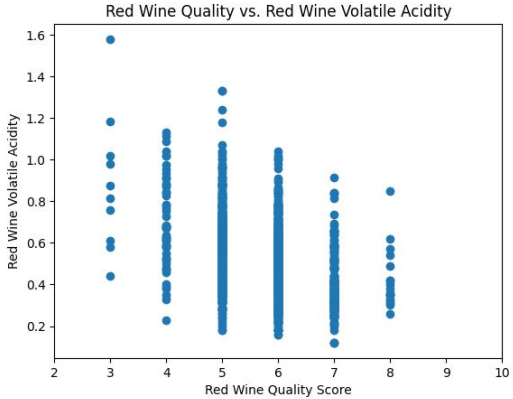
Box Plot for Fixed Acidity

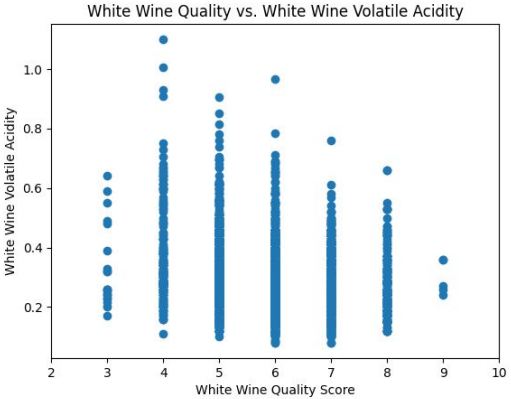


Volatile Acidity

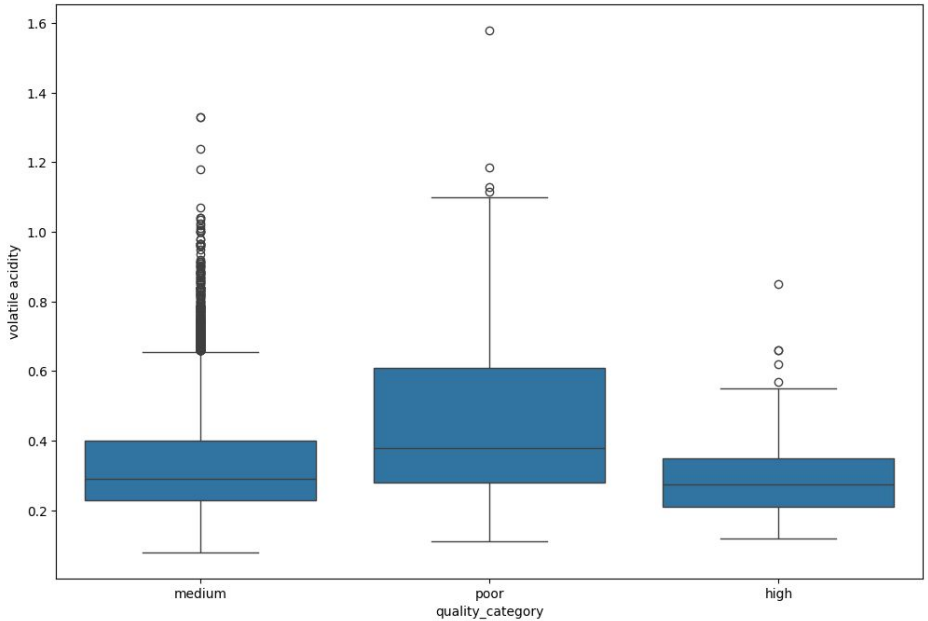
In general, there appears to be a negative correlation between volatile acidity and quality. This is especially noticeable if the wines with quality scores of 3, 4, 8, and 9 are ignored.





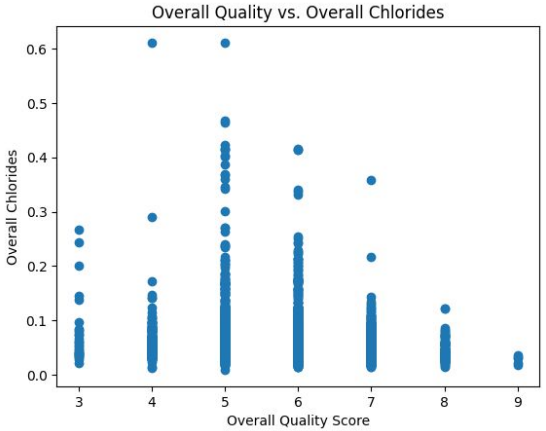


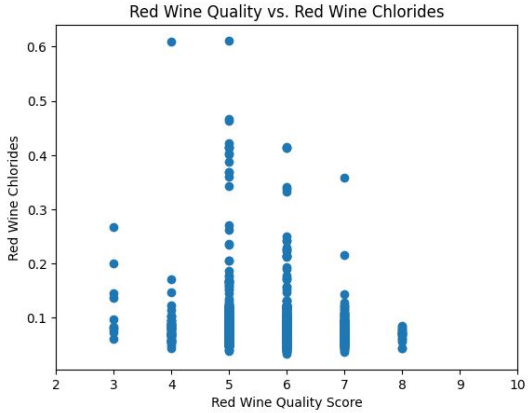
Box Plot for Volatile Acidity

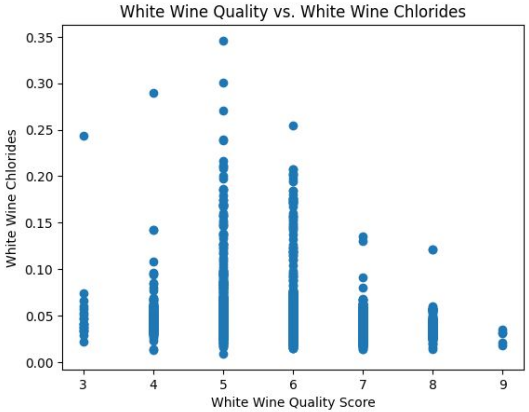


Chorides

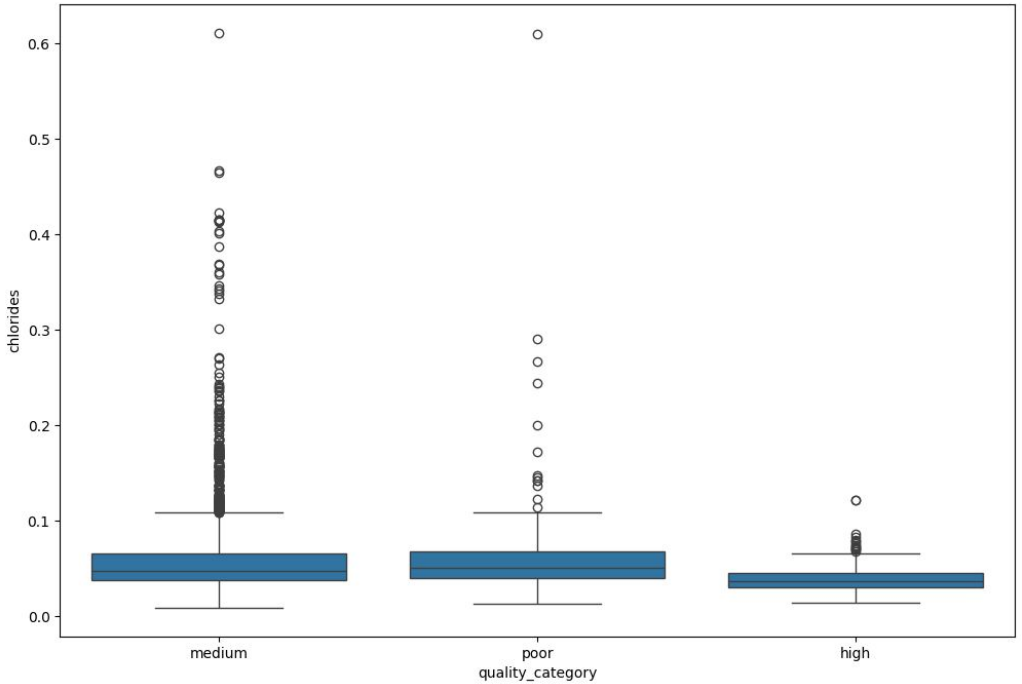
There does not appear to be much of a correlation between chloride levels and quality, unless wines with quality scores of 3, 4, 8, and 9 are ignored. Then there may be a negative correlation.







Box Plot for Chlorides



Digging Deeper

- Strip out the high and poor quality wines

- Drawing conclusions about how the quality of the wines is affected by our chosen variables may be easier when the statistics aren’t affected by a low number of high and low quality wines


Questions?

Sources

The source for our data is Kaggle under the title Wine Quality by Joakim Arvidsson. We also drew information from https://www.vinhoverde.pt/en/about-vinho-verde since this source was included with the page where we found the csv files we used.

This dataset is publicly available for research. The details are described in [Cortez et al., 2009].

Source definitions
Fixed Acidity; Agrovin Techniques for Correcting Wine Acidity
Volatile Acidity; Penn State Extension Volatile Acidity in Wine
Chloride; Mantech Inc. Application Note #105-Chloride in Wine Titration