# Wine Quality Milestone Report

Introduction

I once had a sommelier tell me that the simplest way to select a good wine is by looking at the animal that was on the label. His theory was that the wine that had the deadliest animal on the label tasted the best. This is a fun anecdote, but how can we objectively quantify what makes a great wine?The quality of our favorite wines is determined by the chemical composition of its makeup. Factors such as pH, acidities, sulphur and alcohol contents greatly affect the flavor, aroma, and mouth feel of these wines. Comparing known components of a selection of wines and the quality scores that were given to these wines, we will find the common characteristics of a great wine and a poor wine.

Data learned from this project would be useful to wineries and producers looking to produce higher quality wines. The ability to fine tune which chemical compounds are used in producing the grapes allows the winemakers to tailor their wines for more enjoyability. Consumers could also use this data to select wines for purchase that will be in line with their tastes.

The dataset

The wine quality dataset comes from the UCI Machine Learning Repository. It consists of 2 datasets for the Portuguese Vinho Verde wine. First is the red wine data set that has 1599 wines with quality rankings from 3 to 8. The second set is the white wine dataset of 4898 wines with quality ranking of 3 to 9. Each of these datasets include the quantities of 12 variables : fixed acidity, volatile acidity,citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, alcohol, and quality. There is no data in regards to grape type, brand name, or price.
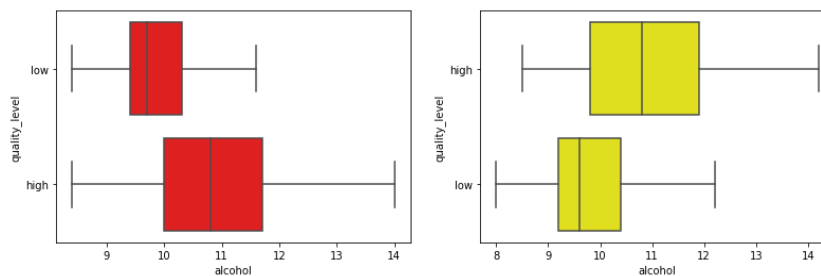
The data were clean and did not need any wrangling. There were no missing values for any of the variables. The distribution of the ratings is a normalized bell curve with many more 'normal' wines and few 'poor' or 'excellent' wines.
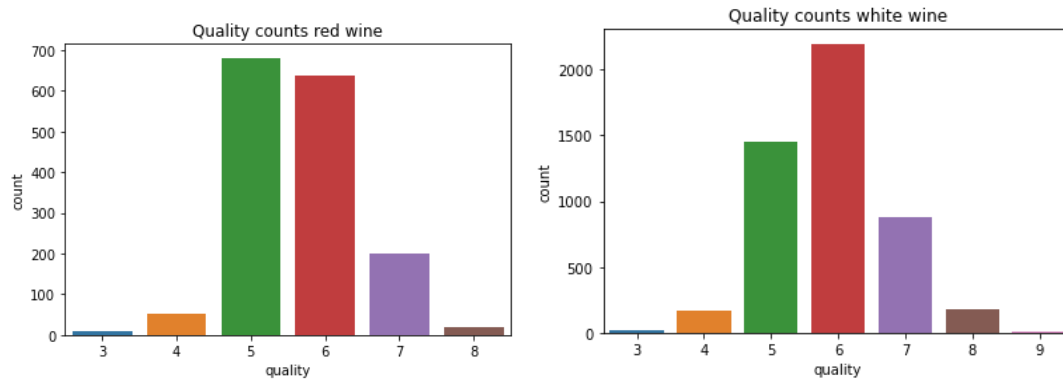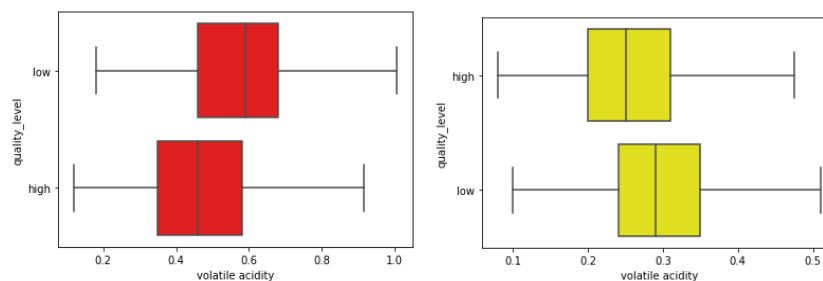
EDA

Initial mean, median, and range analysis produces the following hypothesis:
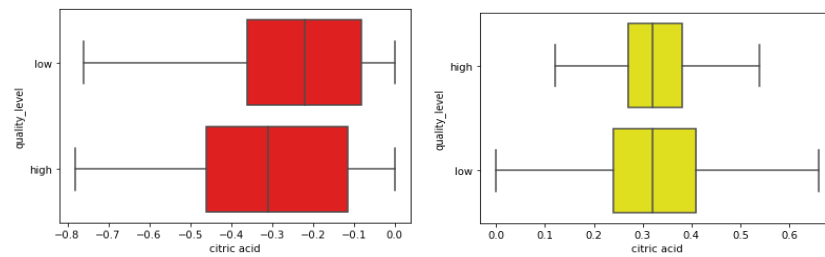
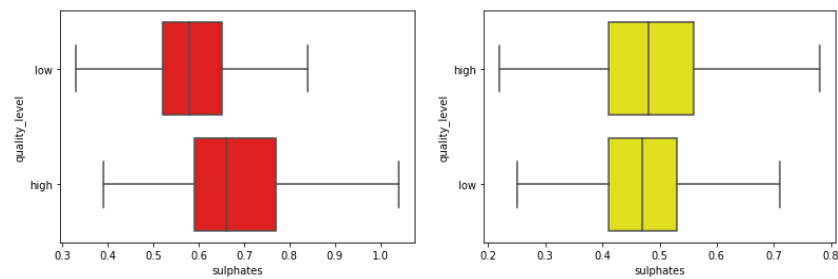- Higher alcohol content correlates to higher wine quality.

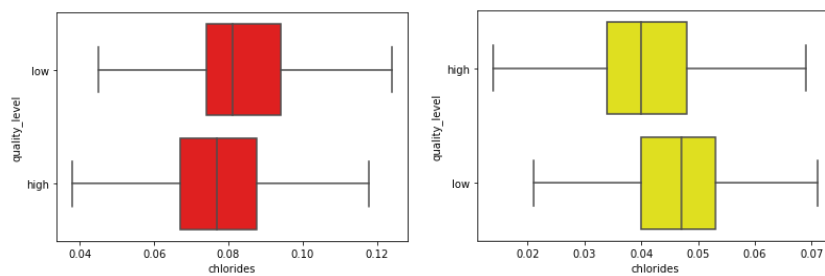- Higher volatIle acidity levels correlates to lower wine quality.

- Higher citric acid levels correlate to lower wine quality.

- Higher levels of sulfates correlates to higher wine quality.

- Lower chloride levels correlate to higher wine quality.

- Lower density correlates to higher wine quality.