Predict wine quality rating

results and conclusions

github

Elena Kochetova
linkedin.com/in/elena-kochetova

Description

Predict wine quality rating

Sarmo Labs Internship
Sep - Oct 2023

Objective:

Predict wine quality rating and identify the key factors that influence it.

Dataset Details:

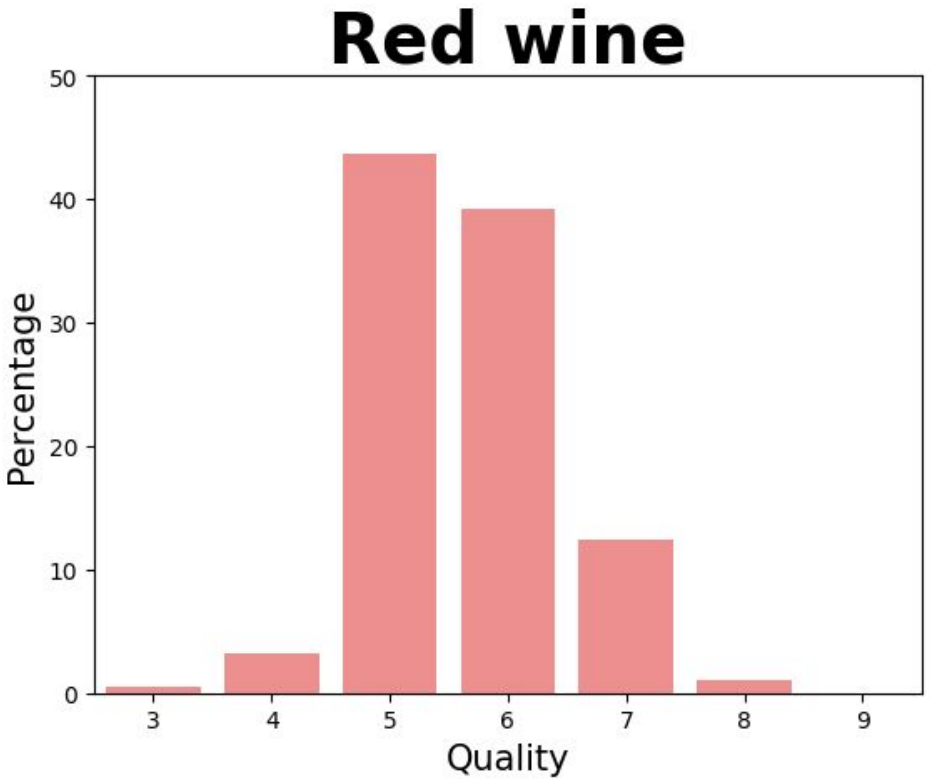
The datasets are related to the red and white variants of the Portuguese Vinho Verde wine. The goal is to model wine quality based on physicochemical tests.

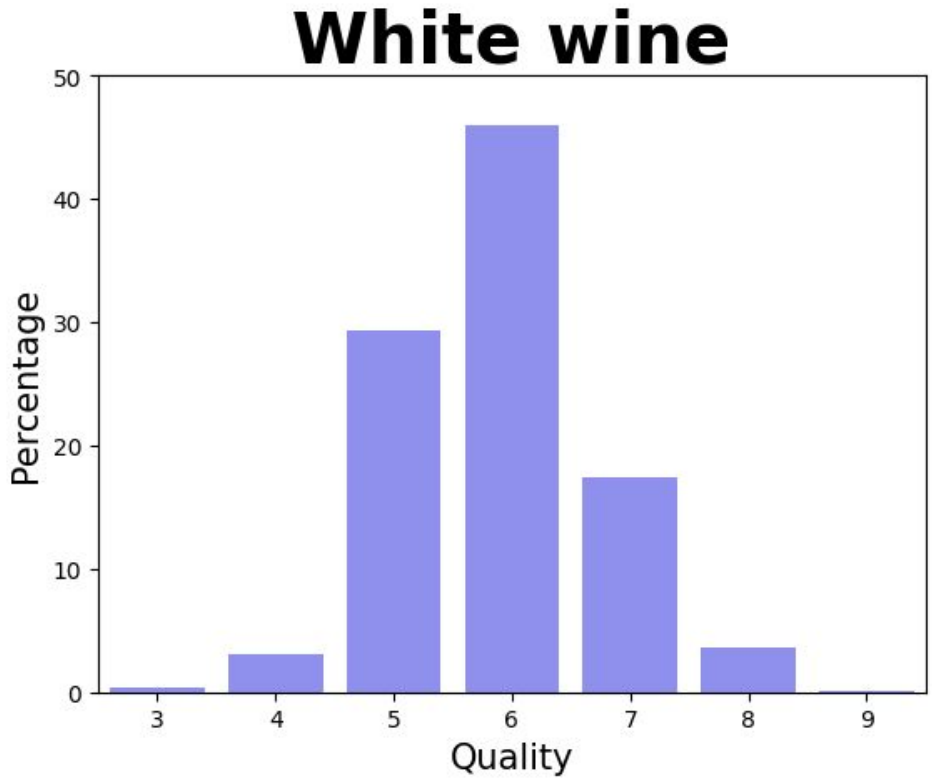
Business Insights and Recommendations:

- Rank features by their influence on wine quality.
- Provide insights on which physicochemical properties should be prioritized to produce high-quality wine.
- Offer any other relevant recommendations or insights derived from the analysis.

Results & Business Insights

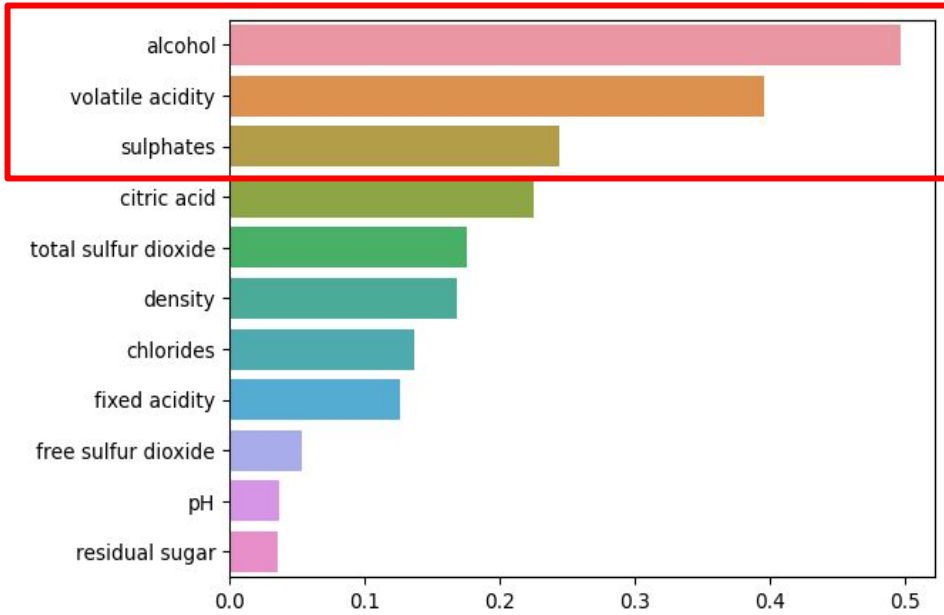
Wine type - Quality



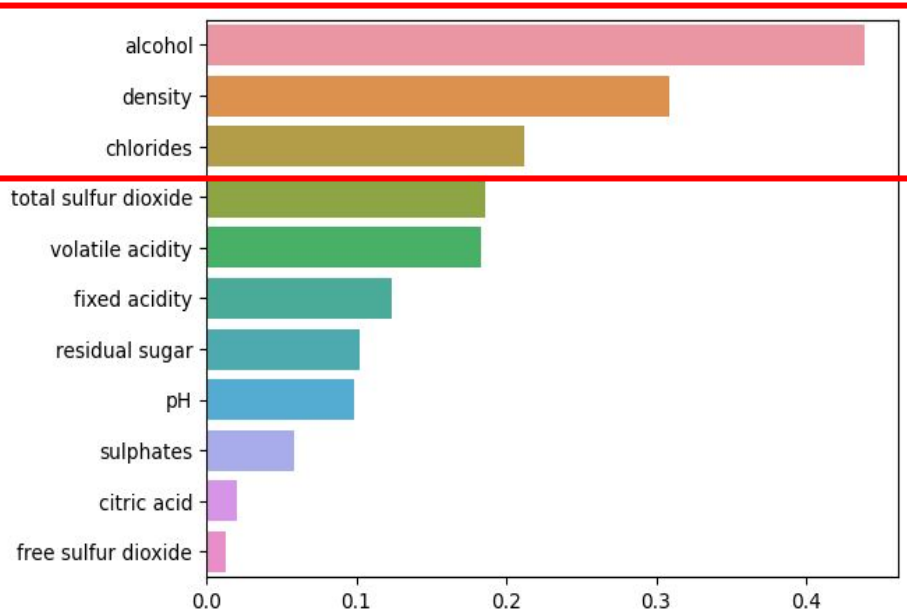


Wine type

Red wine



White wine

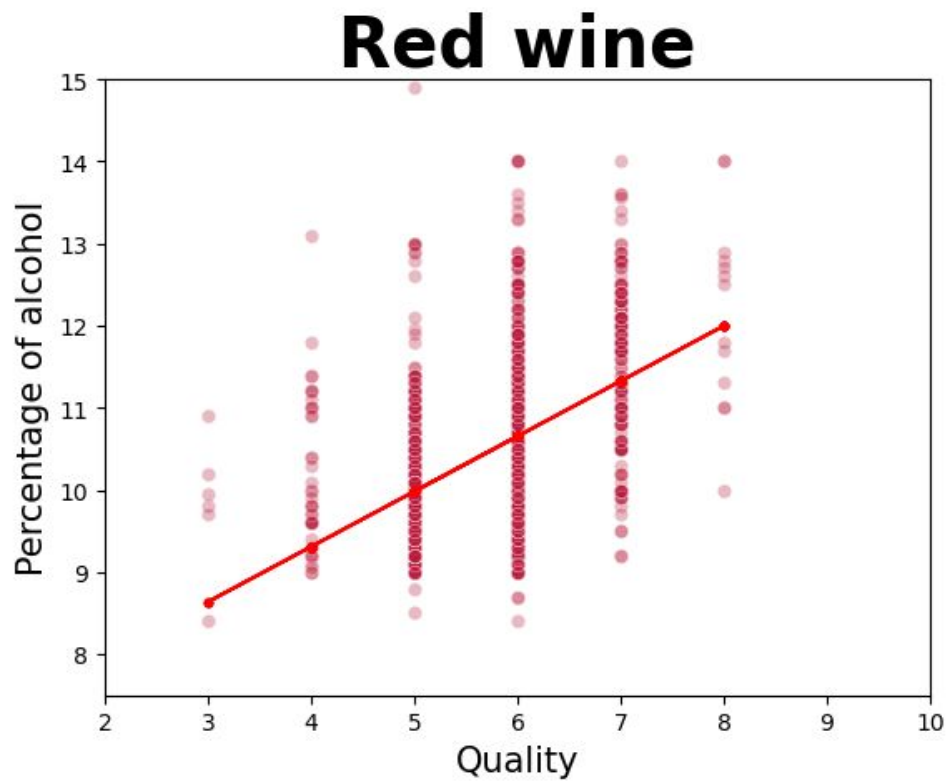


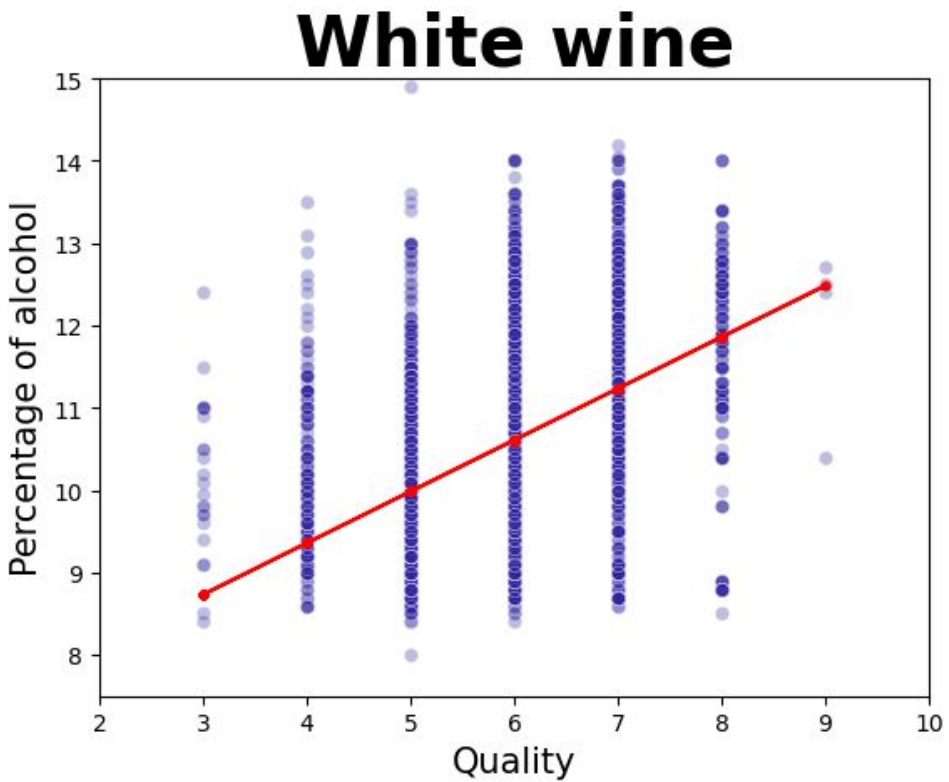
Recommendations:

Consideration of Key Parameters:

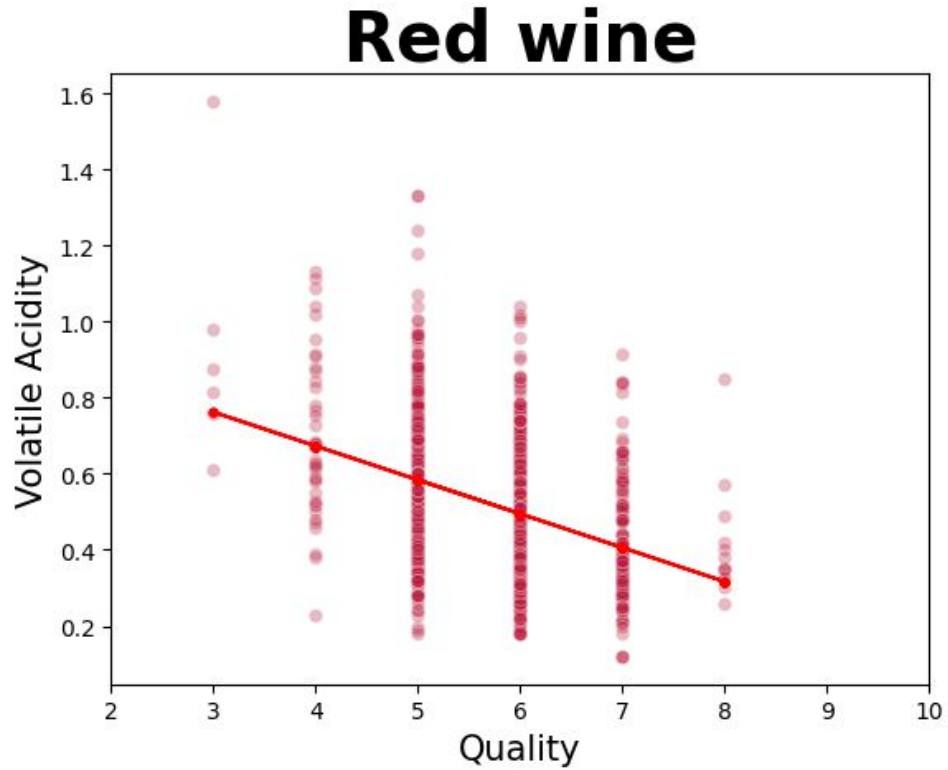
Red wine	**White wine**
Alcohol	**Alcohol**
Volatile Acidity	Density
Sulphates	Chlorides

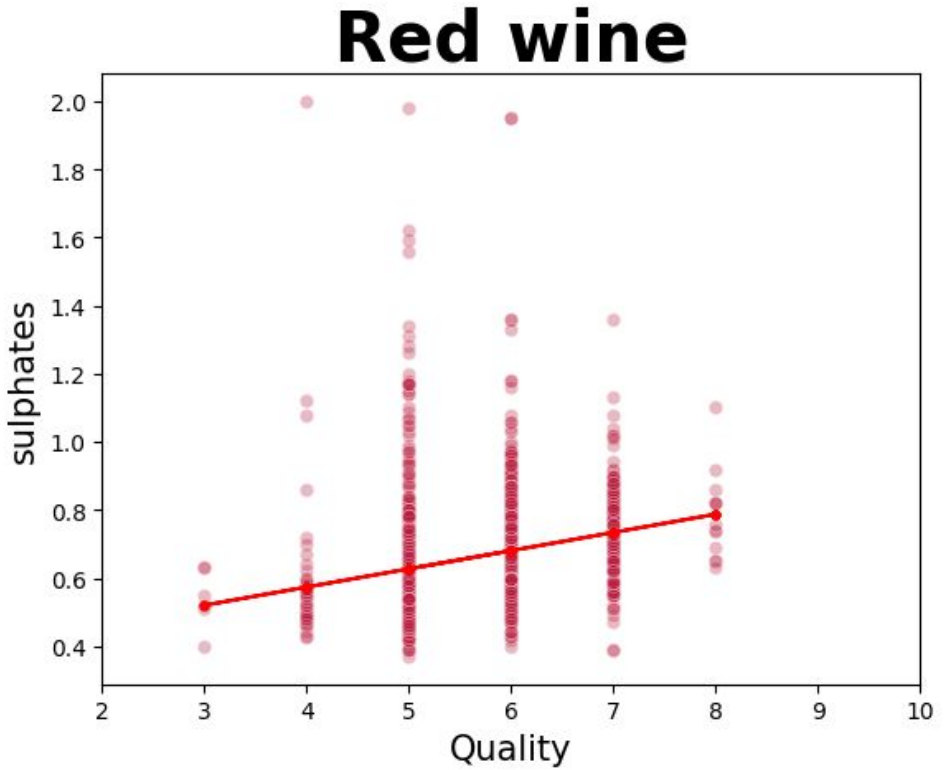
Alcohol content for Red & White wine



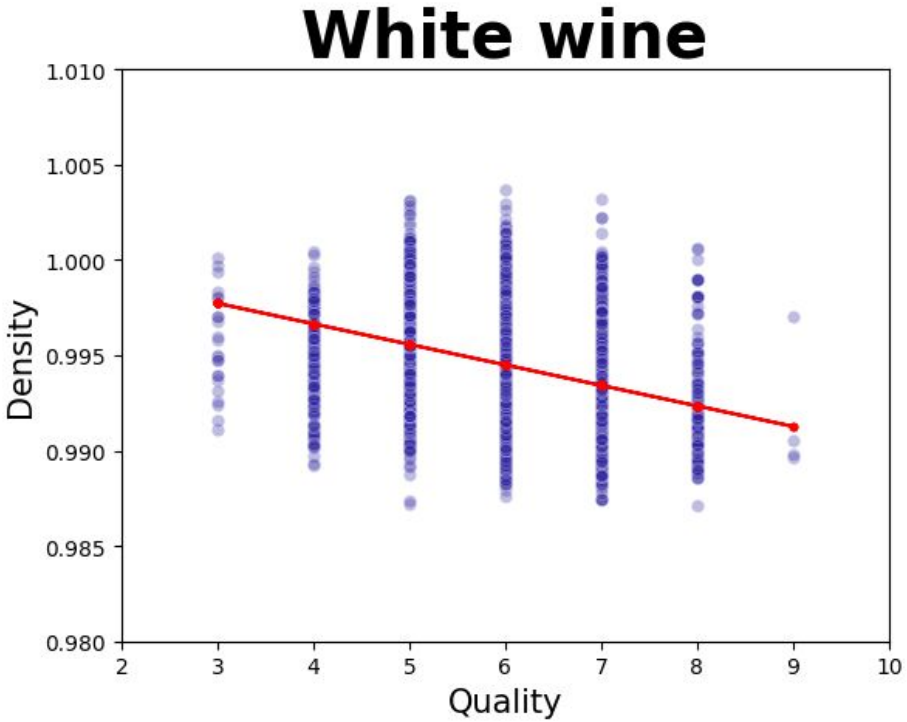


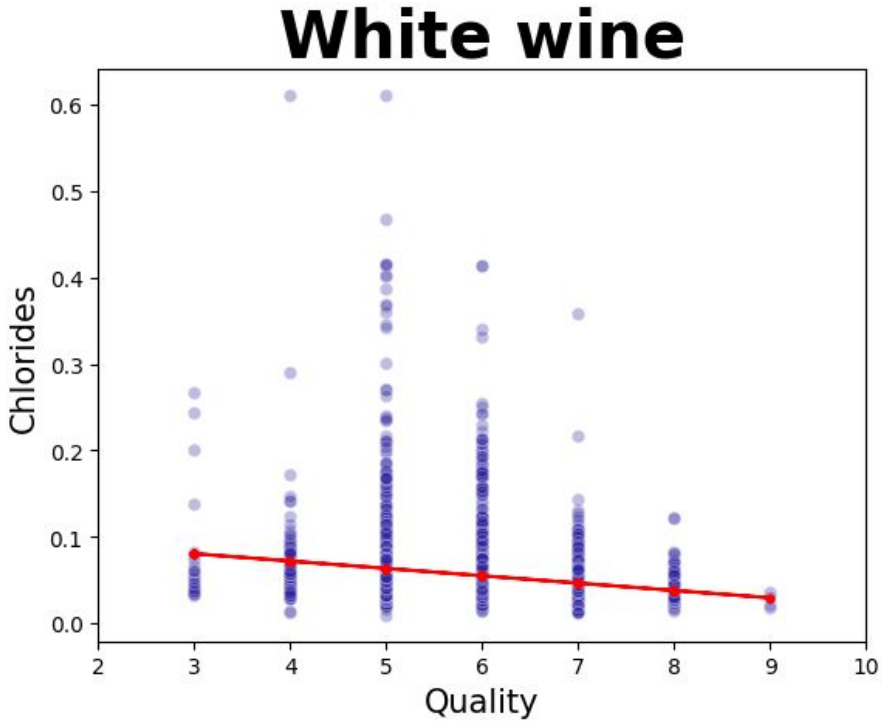
Volatile Acidity & Sulphates for Red wine





Density & Chlorides for White wine





Recommendations:

Consideration of Key Parameters:

Red wine	White wine
Alcohol	Alcohol
Volatile Acidity	Density
Sulphates	Chlorides

To enhance the quality of wine, it is essential to focus on controlling key parameters that have a significant impact on the wine's quality assessment

Recommendations:

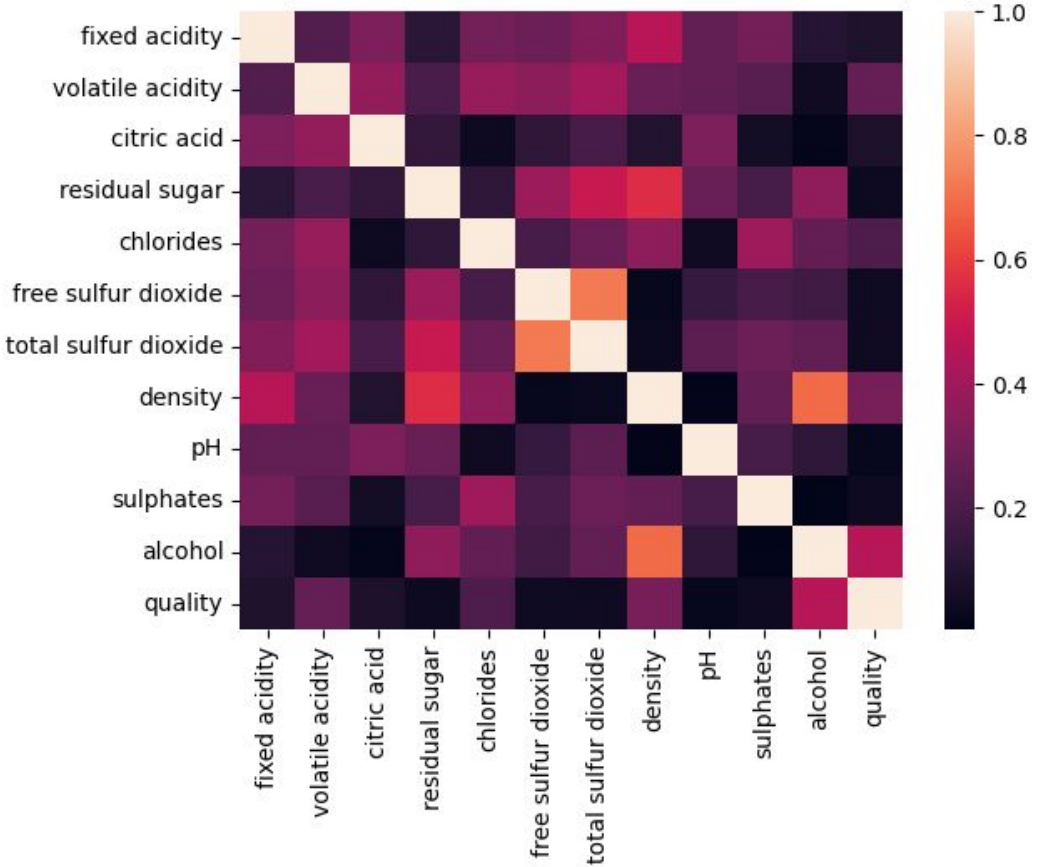
Discarding Insignificant Parameters:

Red wine	White wine
Citric Acid Total Sulfur Dioxide Density Chlorides Fixed Acidity Free Sulfur Dioxide pH Residual Sugar	Total Sulfur Dioxide Volatile Acidity Fixed Acidity Residual Sugar pH Sulphates Citric Acid Free Sulfur Dioxide

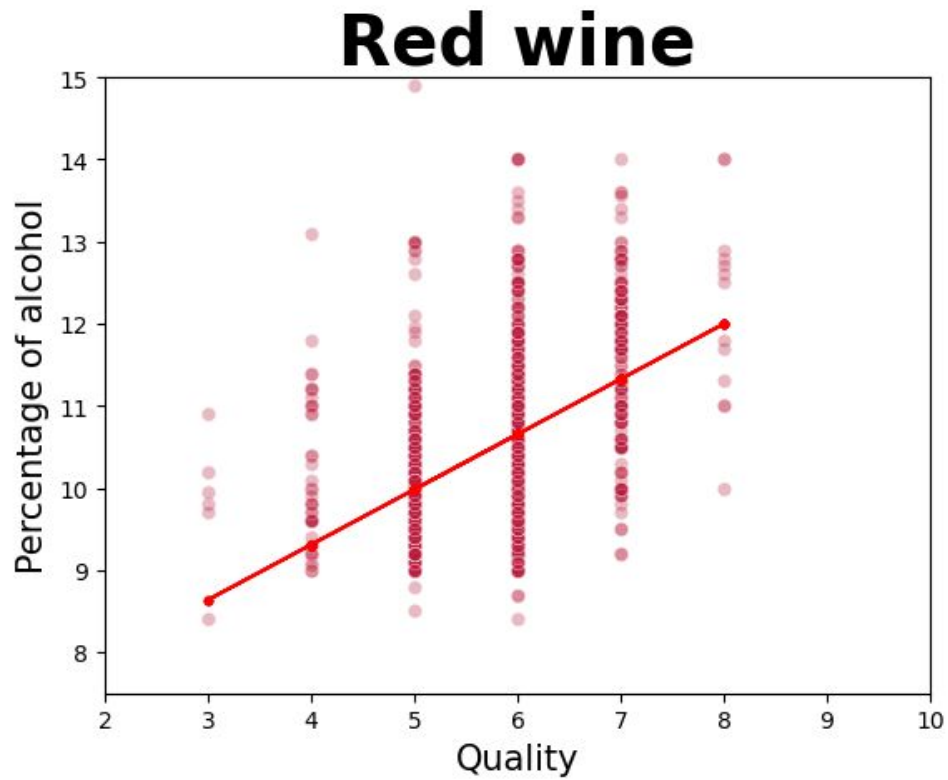
This will help reduce production costs and resources without compromising quality.

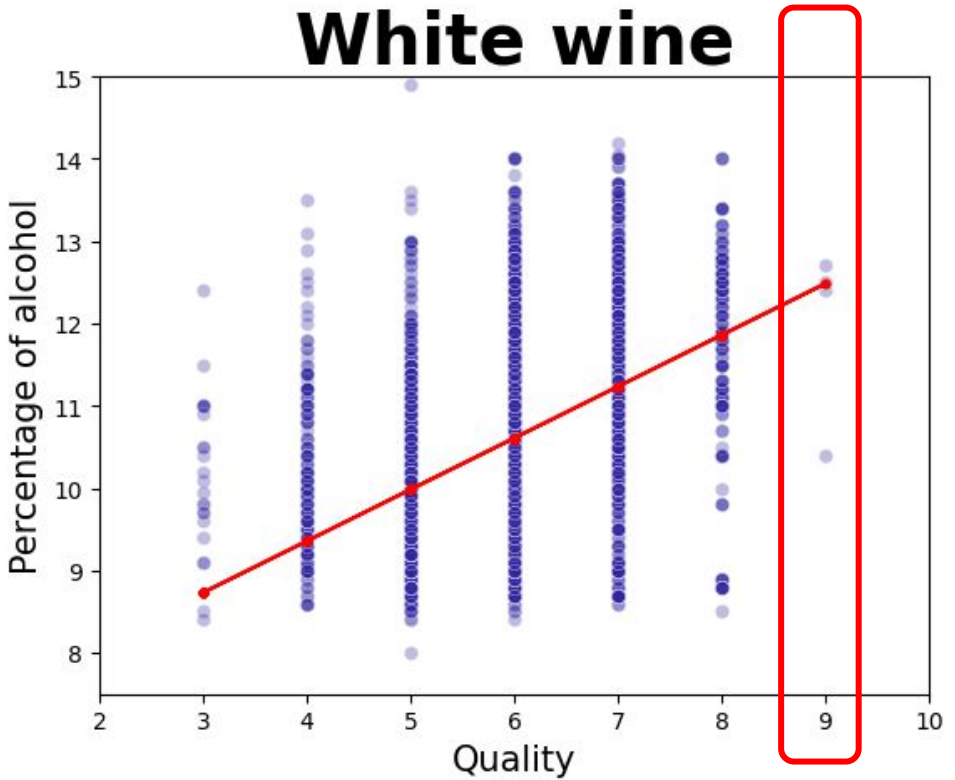
Parameters Correlation Heatmap:

A correlation heatmap of parameters provides valuable insights for optimizing wine production and enhancing its quality.



Alcohol content – Quality





Recommendations:

Creating a Segment of High-Quality Wine:

Define a threshold value for wine quality that will be considered high.

Allocate production and marketing efforts for wines with a quality rating above the established threshold.

This will allow you to create an attractive segment for wine connoisseurs and promote them more effectively in the market.

Maintain strict quality control for wines included in this segment to ensure consistent quality.

Additional Data Analysis:

Utilize machine learning models to predict which wines will belong to the high-quality segment based on their current characteristics.

Prediction Model:

Developed a ML model for predicting wine quality based on physicochemical parameters.

Achieved the following RMSE scores on the test data:

- DecisionTreeRegressor: RMSE = 0.81
- Linear Regression: RMSE = 0.57
- RandomForestRegressor: RMSE = 0.39

The **RandomForestRegressor** model demonstrated the best performance, making it the most suitable for predicting wine quality.

Thank you for your attention!

Elena Kochetova

https://www.linkedin.com/in/elena-kochetova/