Wine Quality Analysis Report

# 1. Introduction

This report provides an analysis of the Wine Quality dataset, specifically focusing on white wines. The dataset comprises various physicochemical properties of the wines and their respective quality ratings.

# 2. Data Cleaning and Preparation

Upon inspection, the dataset was found to be complete without any missing values. Therefore, no further cleaning was required.

# 3. Data Exploration

The distribution of wine quality scores revealed that most wines are rated between 5 and 7, with 6 being the most common quality score.

# 4. Feature Analysis

The relationship between different features and wine quality was explored using boxplots. Key findings include:

- Wines with higher alcohol content tend to have higher quality scores.  
- Wines with lower density, fewer chlorides, and lower volatile acidity tend to be of higher quality.  
- The influence of other features like pH, sulphates, and residual sugar on wine quality is less pronounced.

# 5. Conclusion

Based on the dataset, the most influential factors affecting wine quality are alcohol content, density, chlorides, volatile acidity, and total sulfur dioxide. For a more nuanced understanding, advanced modeling techniques can be employed.