



Optimizing Wine Quality: Insights and Recommendations for

BlueBerry Winery

Leveraging Data Analytics to Enhance Product Quality

Fehmi Katar – 29.08.2023

Part 1: Wine Analysis and Feature Selection

- Introduction to Wine Dataset
- Comparative Analysis: White Wine vs. Red Wine
- Feature Importance Analysis

Part 2: Modeling and Business Recommendations

- Model Selection and Sampling Techniques
- Modeling Results
- Business Recommendations



BlueBerry Winery's Goal



To set **the right price** for their wines based on how good they are.

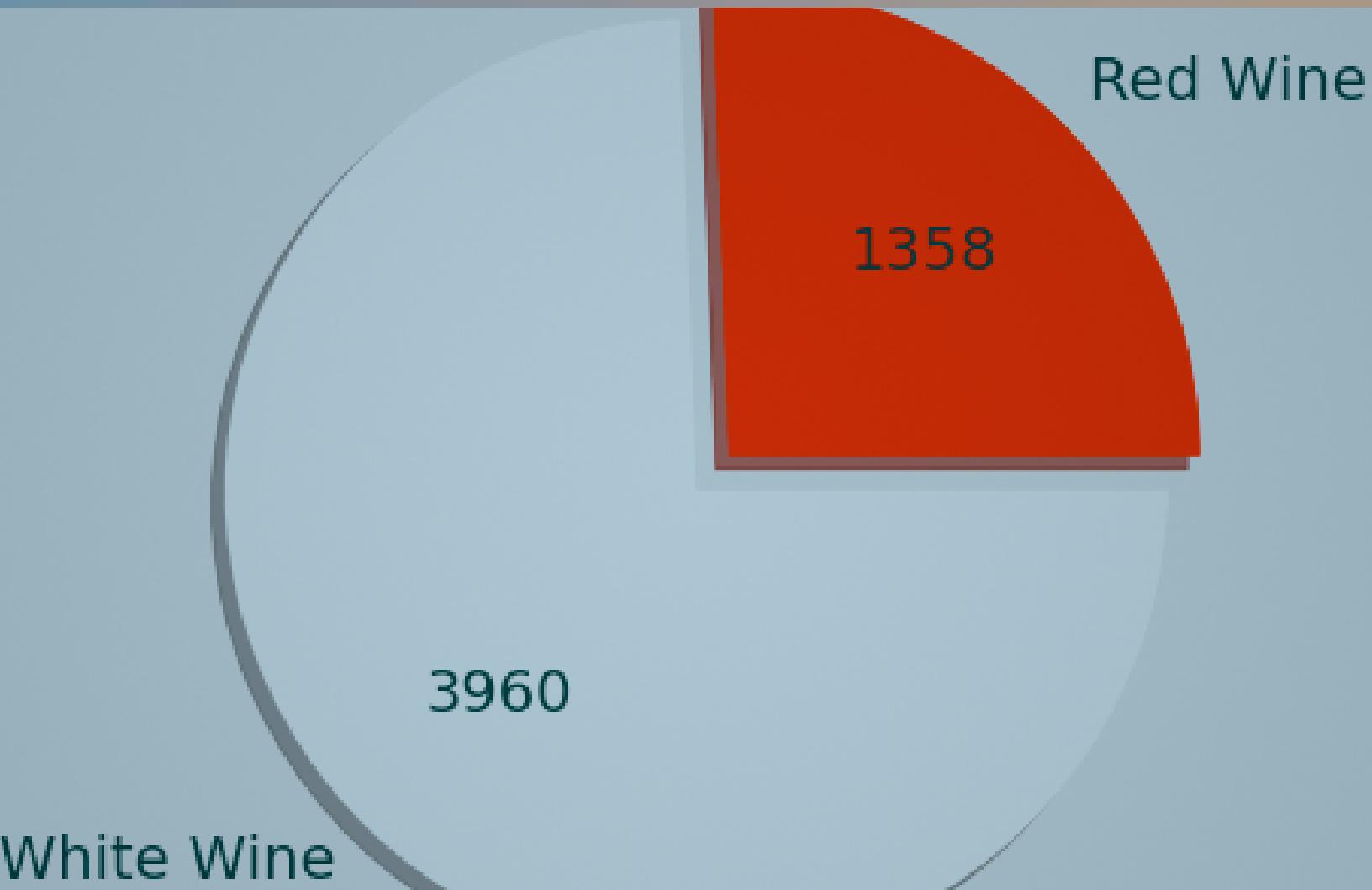


Data Analytic's Goal



To create a **system that checks the quality** of the wines by analyzing the ingredients, enabling **better pricing decisions**.

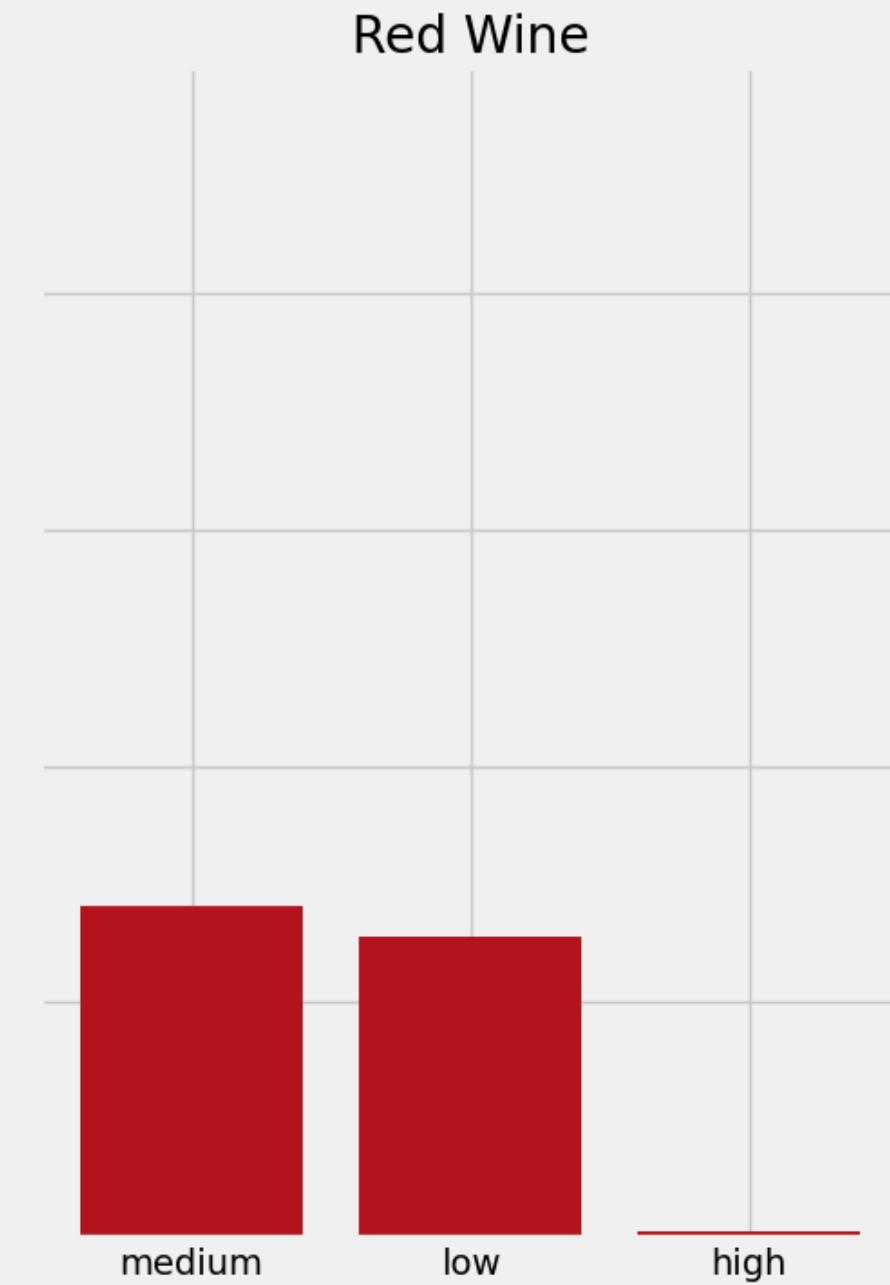
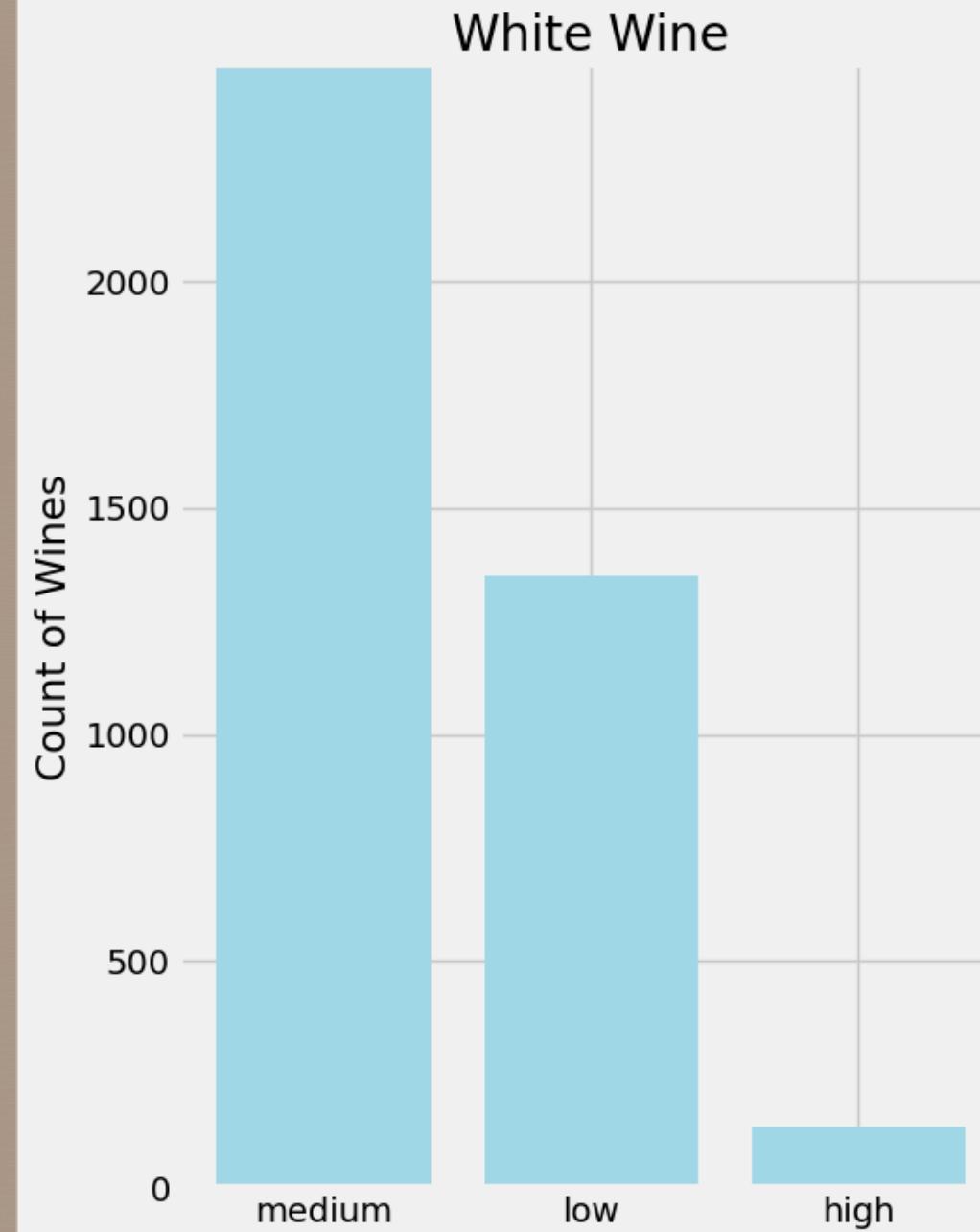
Total count of white and red
wines: 5318



Ingredients / Features

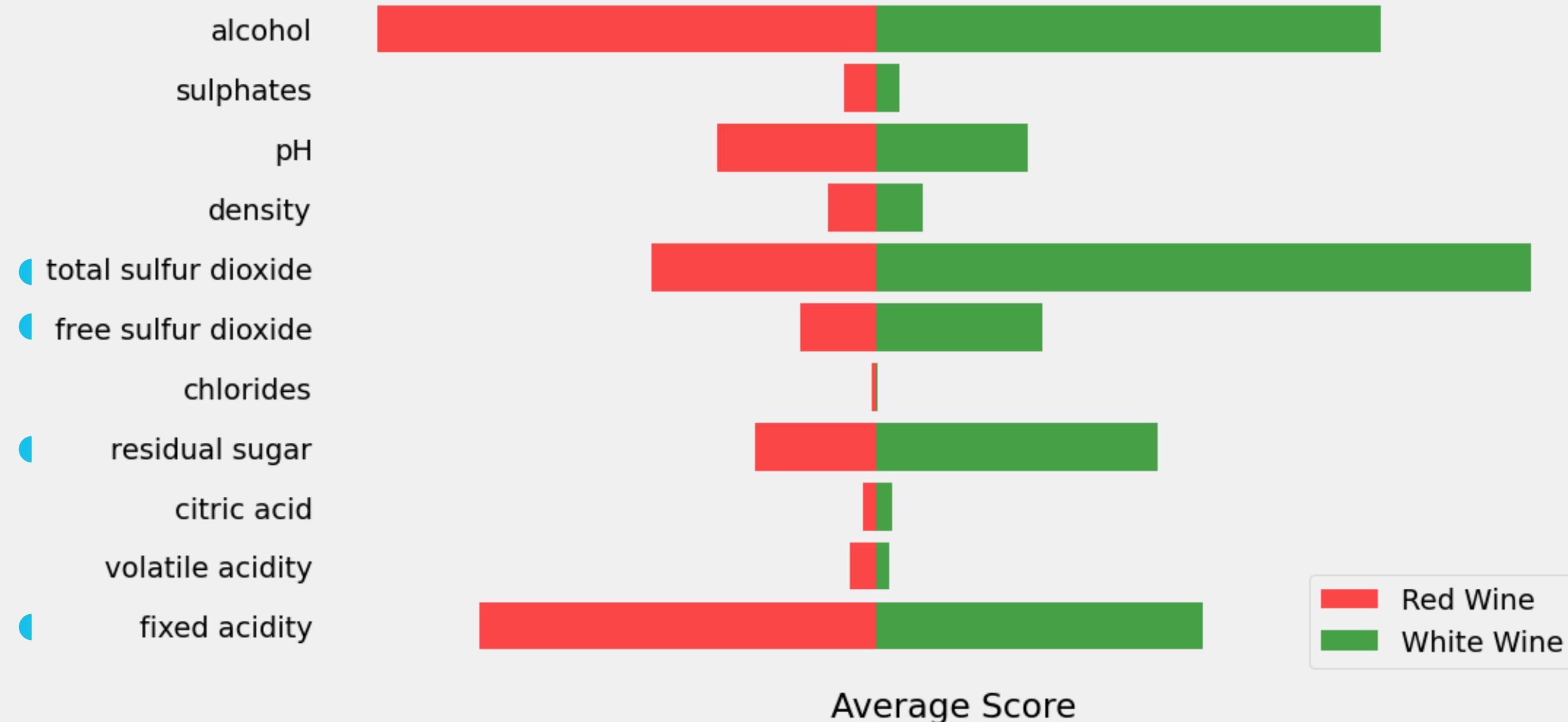
- Fixed Acidity
- Volatile Acidity
- Citric Acid
- Residual Sugar
- Chlorides
- Free Sulfur Dioxide
- Total Sulfur Dioxide
- Density
- pH
- Sulphates
- Alcohol
- Quality

Quality future in Dataset



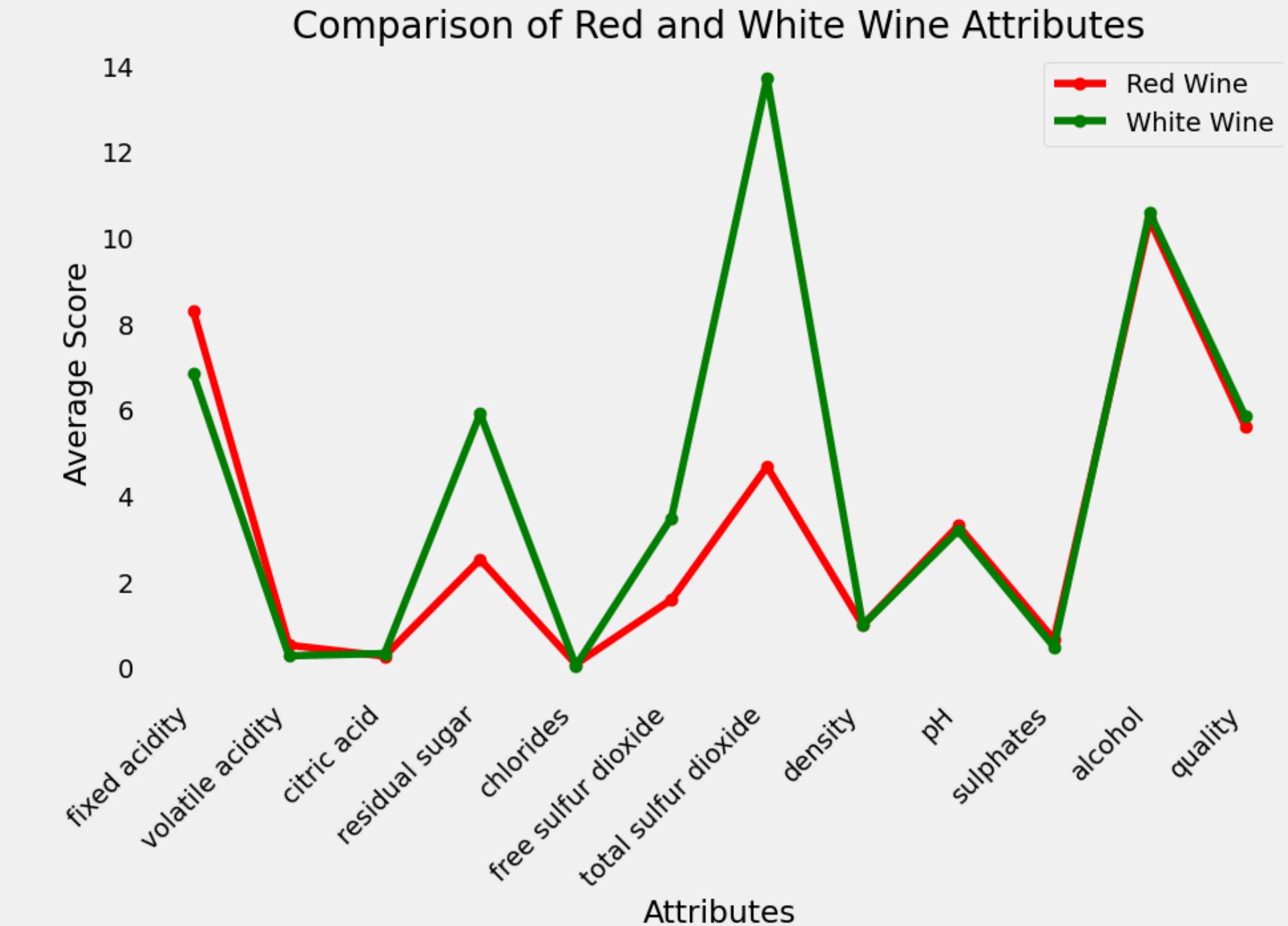


Comparison of Red and White Wine Attributes



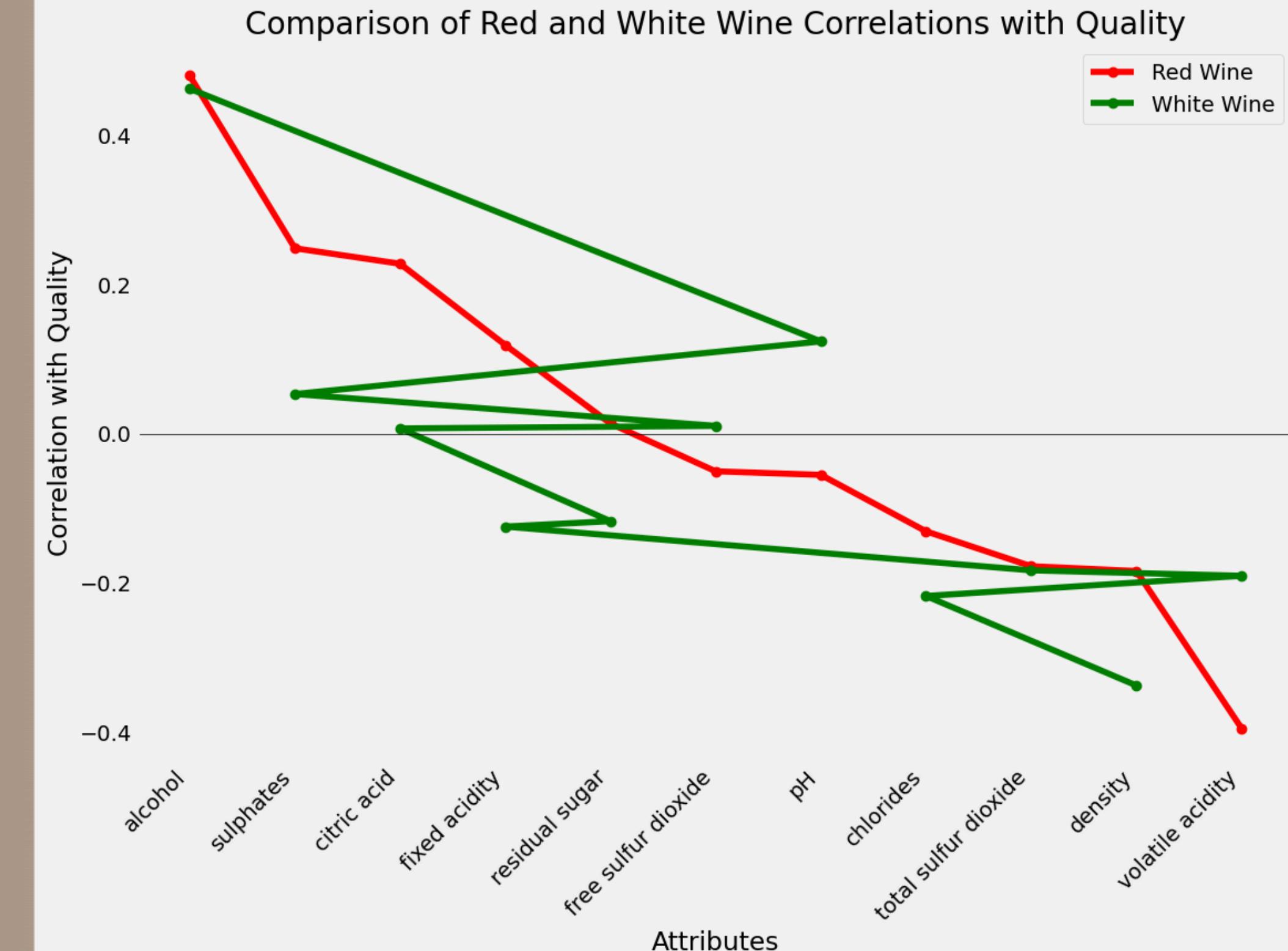
Overall, we can see differences mainly in 4 attributes:

- Fixed Acidity
- Residual Sugar
- Free Sulfur Dioxide
- Total Sulfur Dioxide



Key Attribute Correlations with Quality: Red vs White Wines

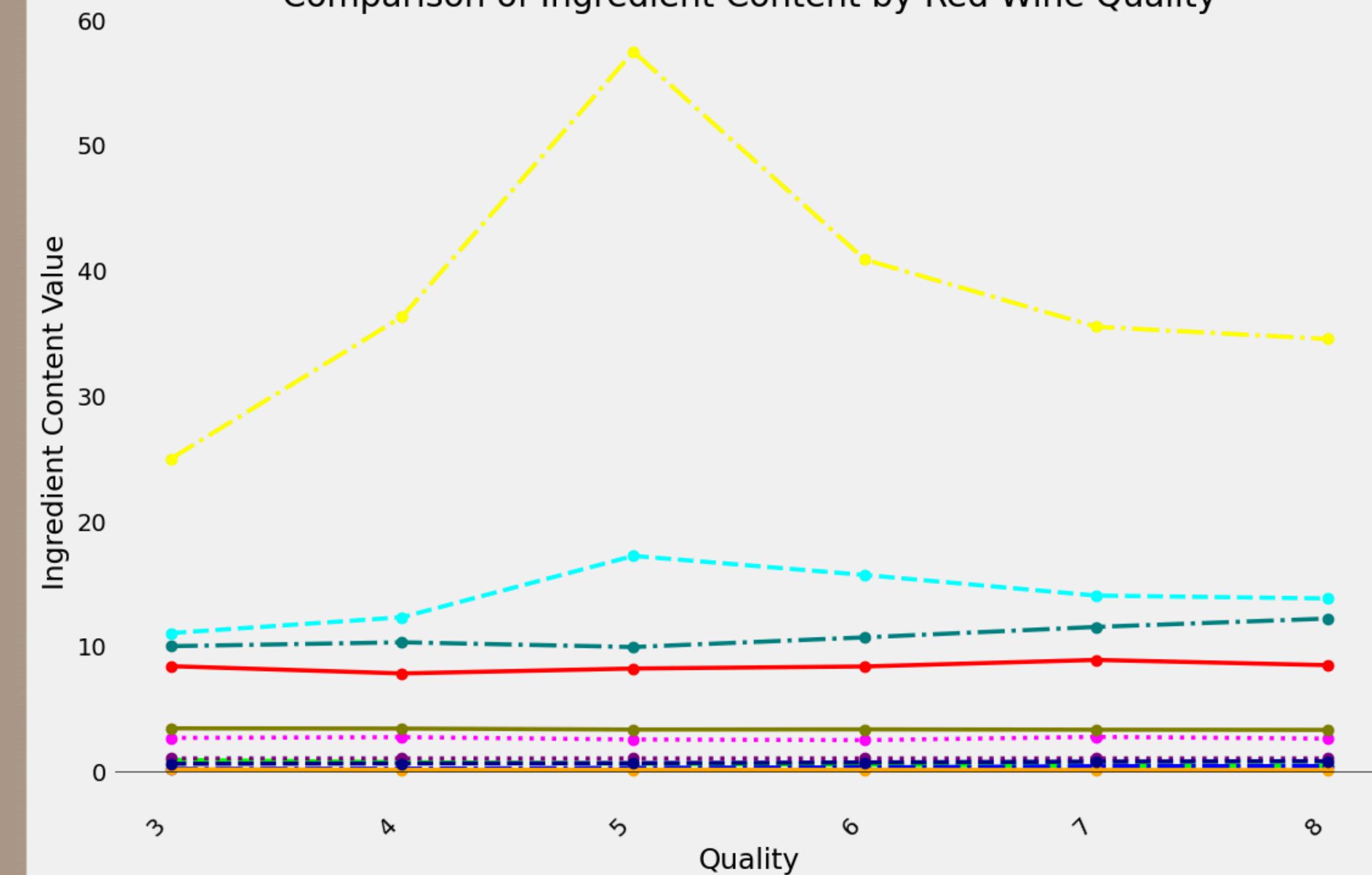
- Alcohol: Stronger impact on red wine quality.
- Sulphates: Enhance red wine more than white.
- Volatile Acidity: More harmful to red wine quality.
- Citric Acid: Benefits red wine quality more.
- Other Attributes: Minimal impact across both wines.



Ingredient Stability Across Red Wine Quality Levels

- Minimal Variation in Red Wine Ingredients
- Stable Ingredients Below Red Line; Variations Above
- Consistent Ingredient Content with Few Changes

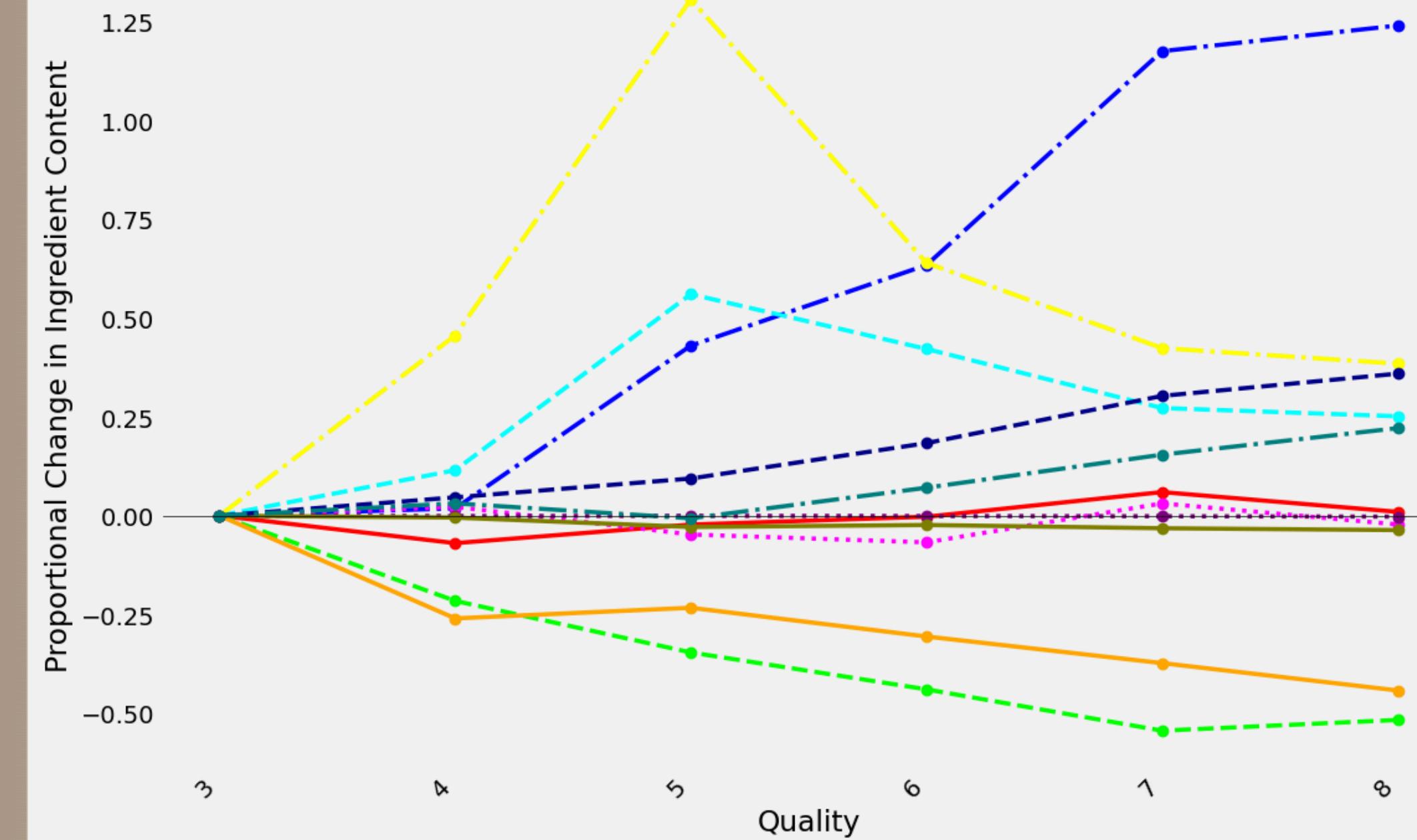
Comparison of Ingredient Content by Red Wine Quality



Proportional Changes in Red Wine Ingredients by Quality

- Most ingredients show >25 % proportional changes.
- Four ingredients stay within 10 % change.

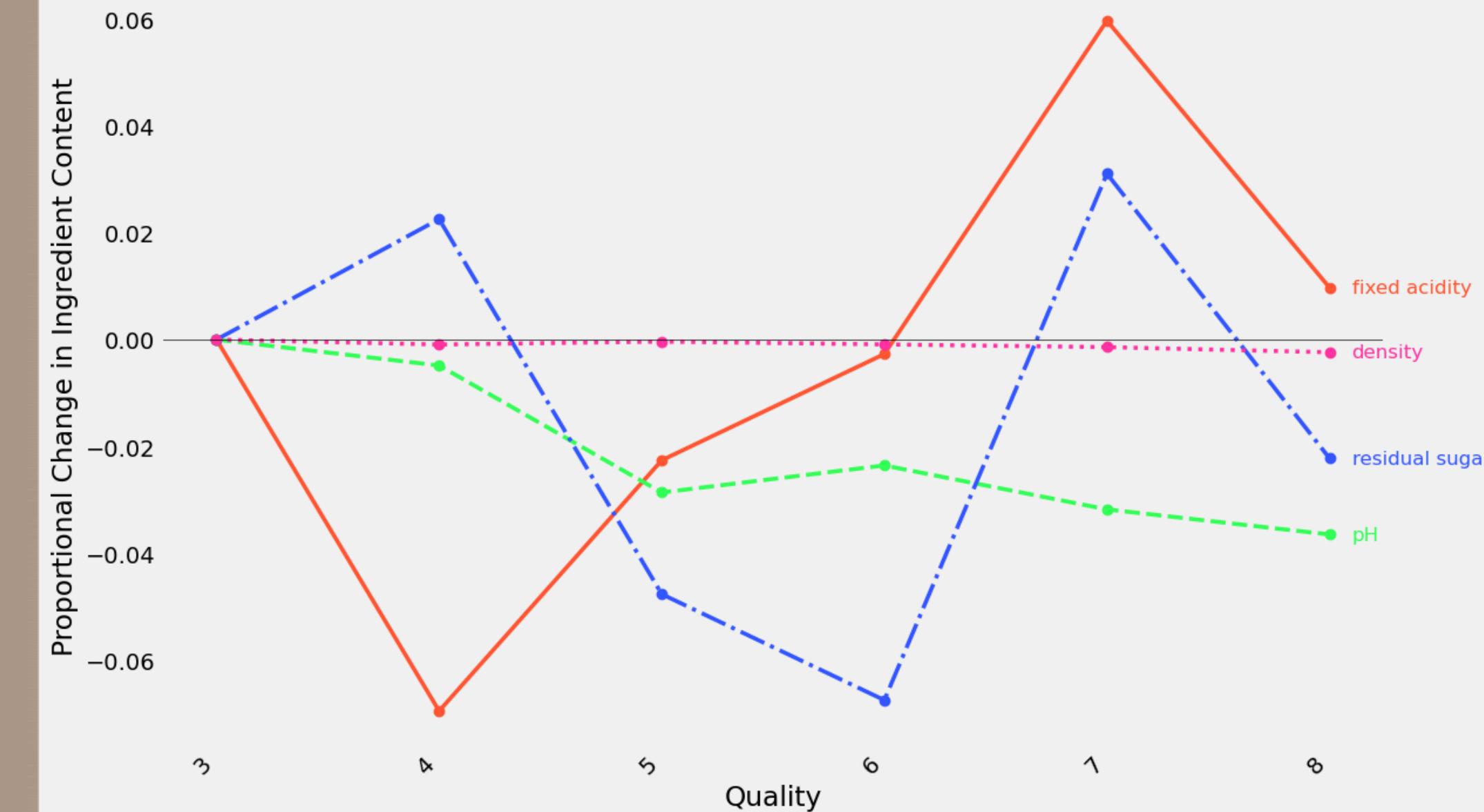
Proportional Changes in Ingredient Content by Red Wine Quality



Proportional Changes in Selected Ingredients

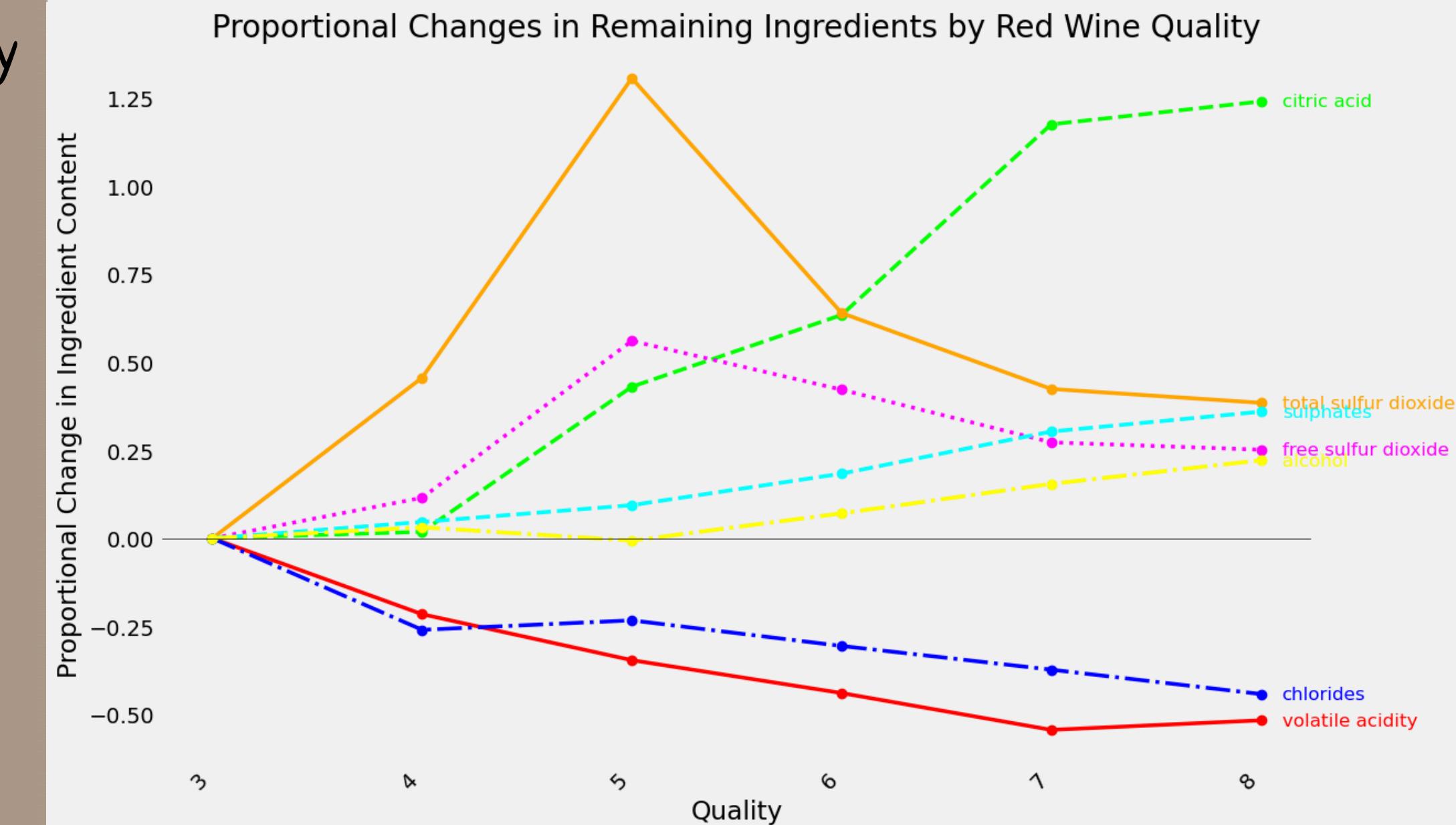
- Changes are within a 6% range for all selected ingredients.
- Density and pH show minimal fluctuations.
- Residual sugar and fixed acidity display more noticeable changes.

Proportional Changes in Selected Ingredients by Wine Quality



Key Ingredients Driving Red Wine Quality

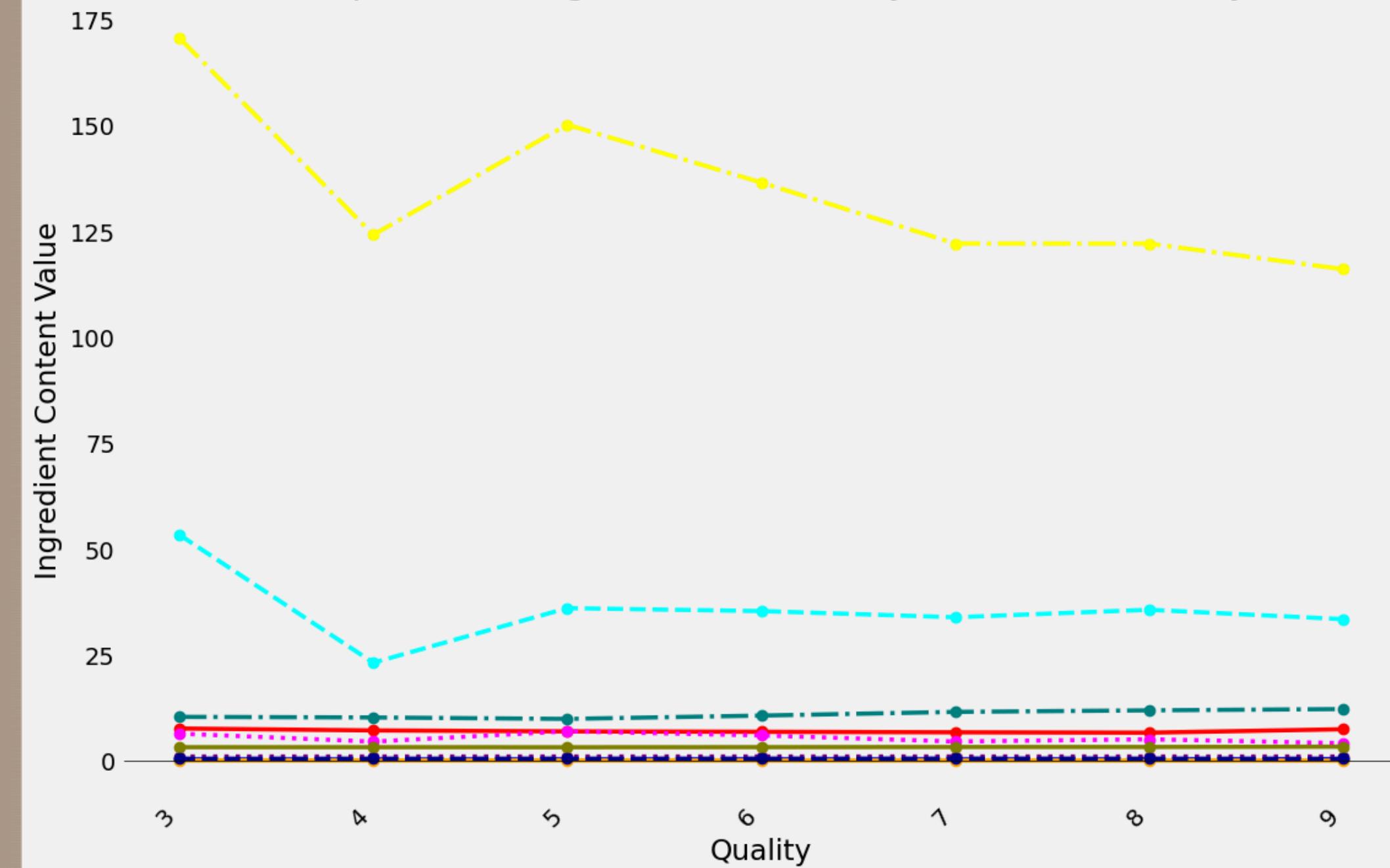
- Citric acid, sulphates, and alcohol consistently increase with quality.
- Chlorides and volatile acidity decrease, potentially indicating better quality.
- Sulfur dioxide levels show mixed changes, suggesting a nuanced role.



Ingredient Consistency in White Wine Quality

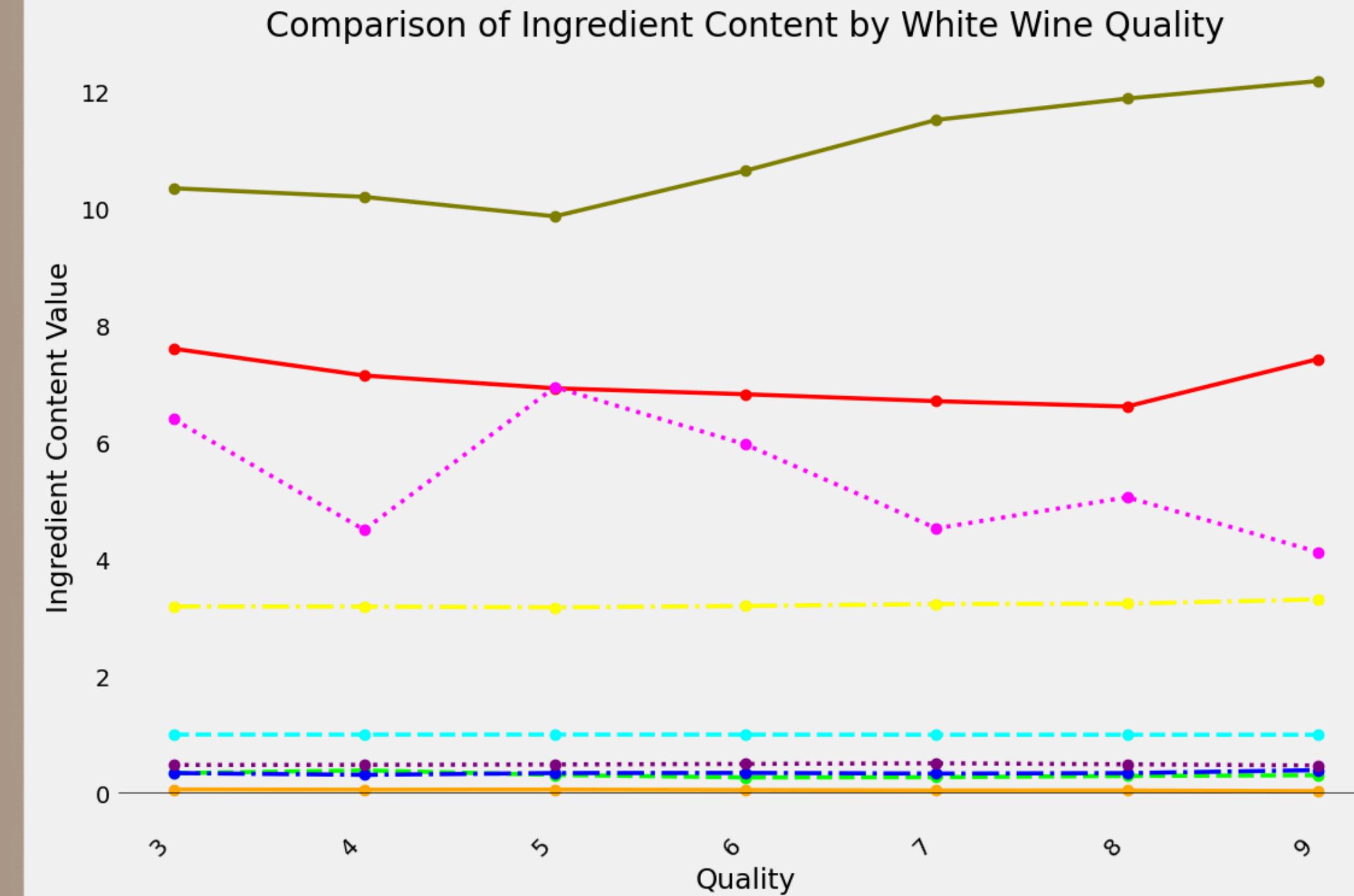
- Stable content across quality levels.
- Sulfur dioxide dominates due to unit difference.

Comparison of Ingredient Content by White Wine Quality



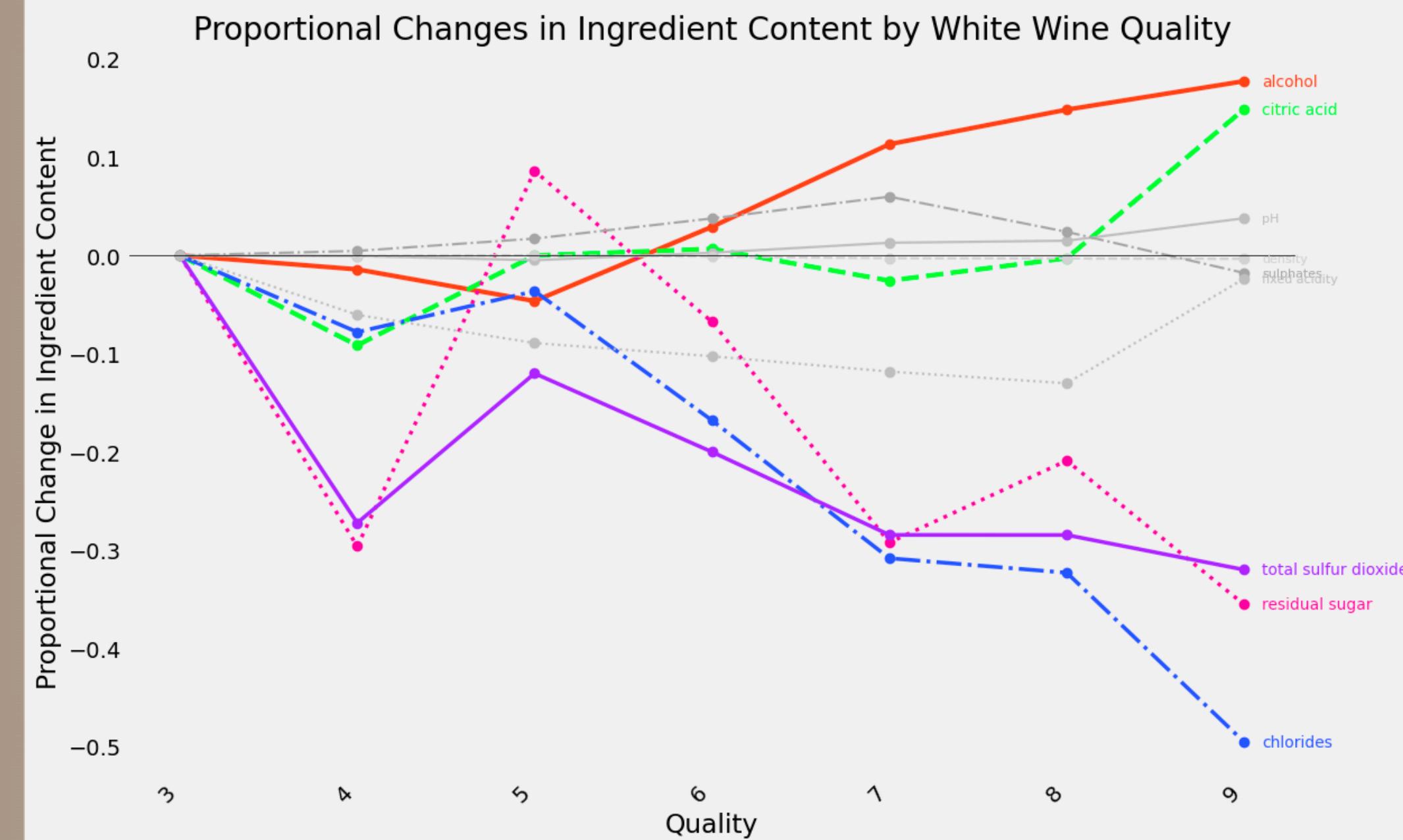
Visible Ingredient Variations Across White Wine Quality

- Clear changes in ingredients above yellow line.
- Minimal change for ingredients below yellow line.
- Proportional analysis needed for better comparison.



Proportional Ingredient Changes in White Wine Quality

- Ein Minimal changes for pH, density, sulphates, and fixed acidity.
- Alcohol and citric acid show increasing trends.
- Volatile acidity, sulfur dioxide, residual sugar, and chlorides decrease.
- Lower quality wines (3-5) exhibit unbalanced ingredient changes.



Key Takeaways on Wine Quality

(Red and White)

- Stable Ingredients: High-quality wines show consistent ingredient levels.
- Key Indicators: Quality linked to higher citric acid, alcohol; lower volatile acidity, chlorides.
- Red vs. White: Red wines are more stable; white wines vary more, especially at lower quality. Sulfur dioxide is positively correlated with quality in red wine and negatively correlated in white wine.
- Balanced Composition: High-quality wines have balanced, moderate ingredient profiles.

Selection of Features as Conclusion

- Selected features show significant changes at different quality levels.
- Features either increase or decrease noticeably.
- Valuable for accurate wine quality prediction

White Wine Model Features:

- Alcohol
- Chlorides
- Total Sulfur Dioxide
- Residual Sugar
- Sulphates

Red Wine Model Features:

- Volatile Acidity
- Chlorides
- Citric Acid
- Sulphates
- Alcohol
- Total Sulfur Dioxide

Modeling Approach

- Logistic Regression
- Random Forest
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)

Optimization Techniques

- Fine-Tuning
 - Adjusting settings for best accuracy. (Number of trees in Random Forest)
- Sampling Techniques
 - SMOTE: Adds more examples.
 - Under-Sampling: Reduces larger groups.
 - Over-Sampling:Duplicates examples

Best model for red wine: Random Forest

Random Forest

- Consistently High Accuracy
- Strong Performance (0.84)

K-Nearest Neighbors (KNN)

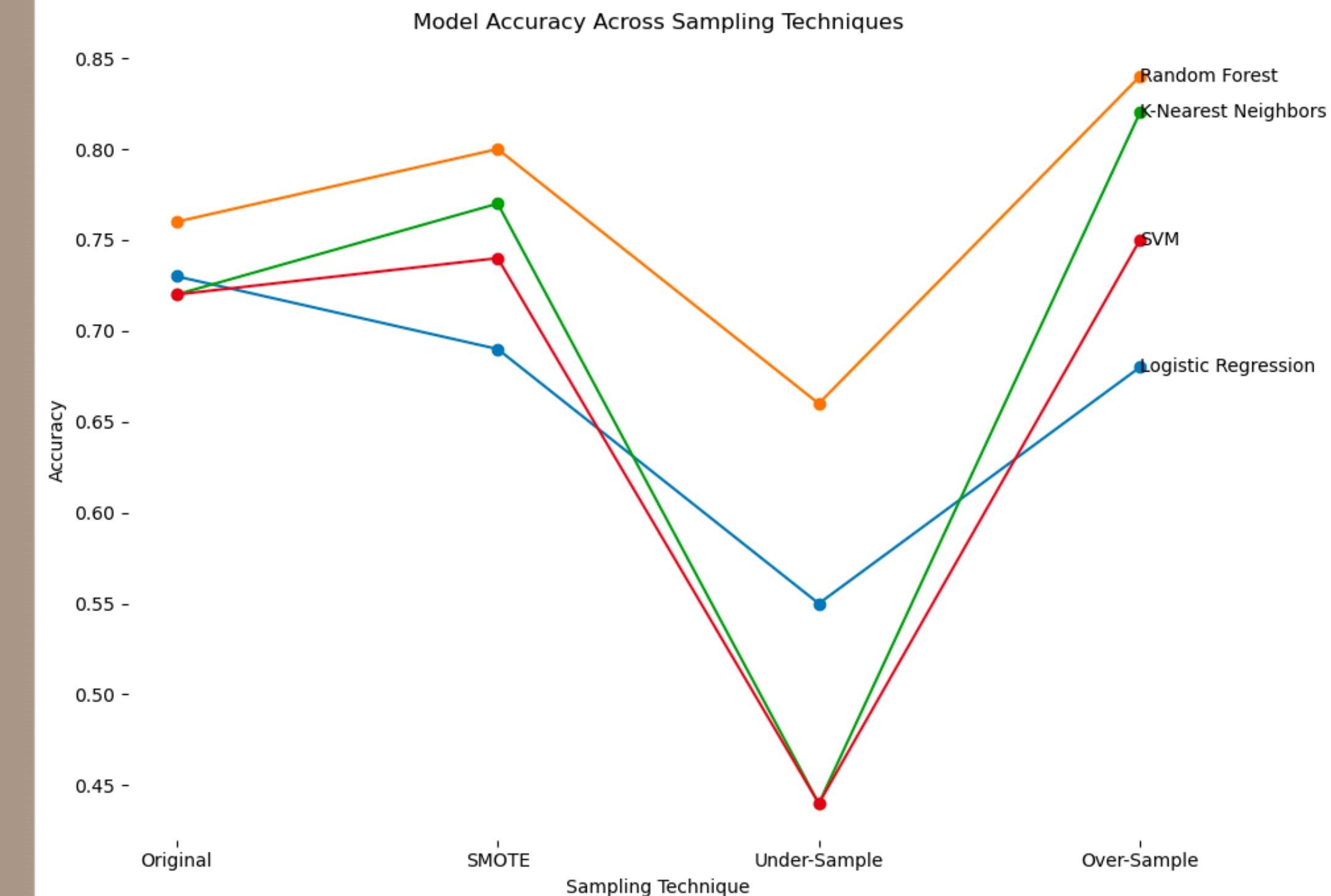
- High Performance: Over-Sampling (0.82)

SVM

- Stable but Moderate Performance
- Over-Sampling (0.75)

Logistic Regression

- Worst Among Models



Best model for white wine: Random Forest

Random Forest:

- Consistently High Accuracy
- Strong Performance (0.86)

K-Nearest Neighbors (KNN):

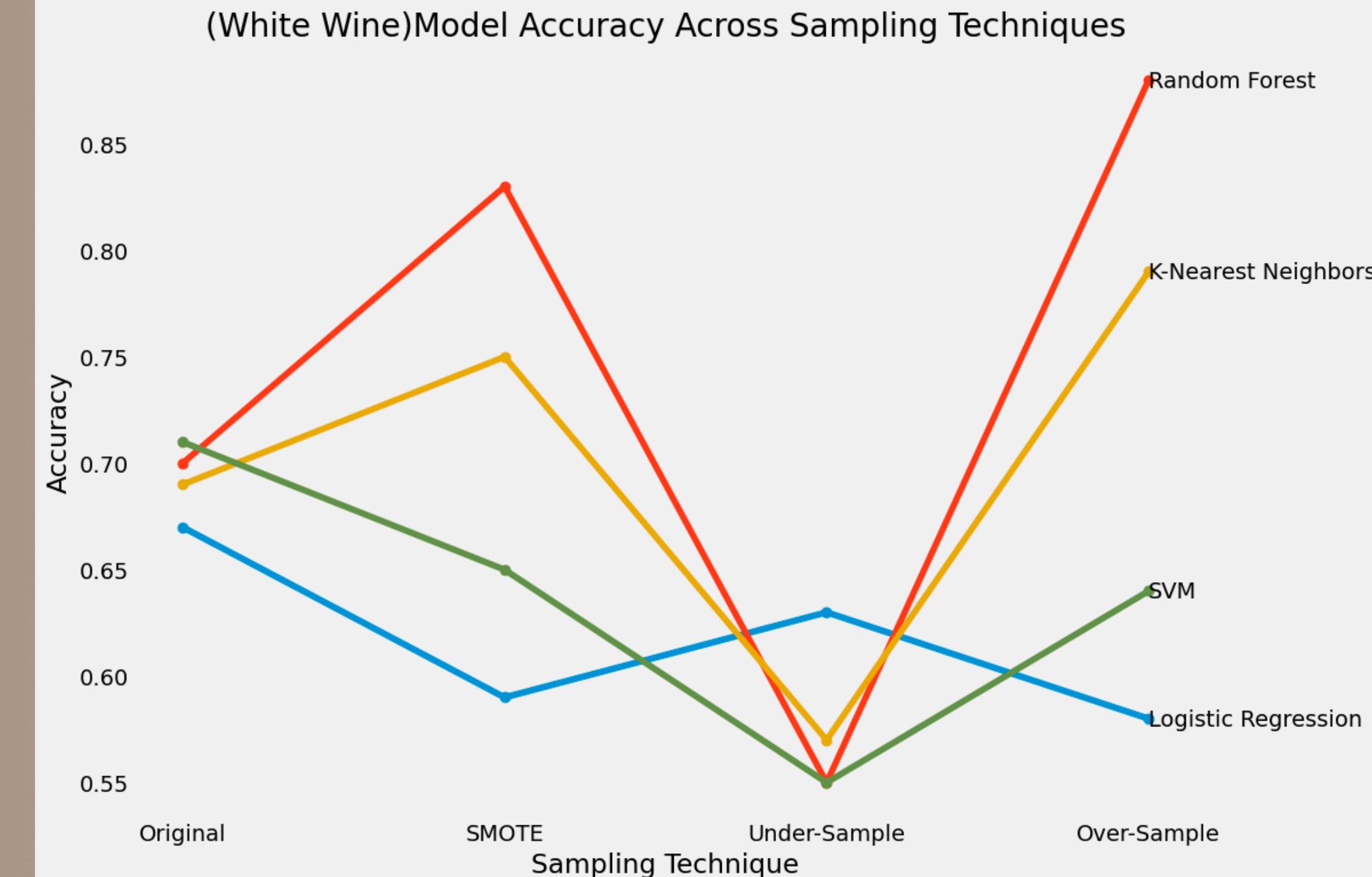
- High Performance: Over-Sampling (0.83)

SVM:

- Stable but Moderate Performance
- Over-Sampling (0.76)

Logistic Regression:

- Lowest Performance (~0.66)



Red Wine vs White Wine

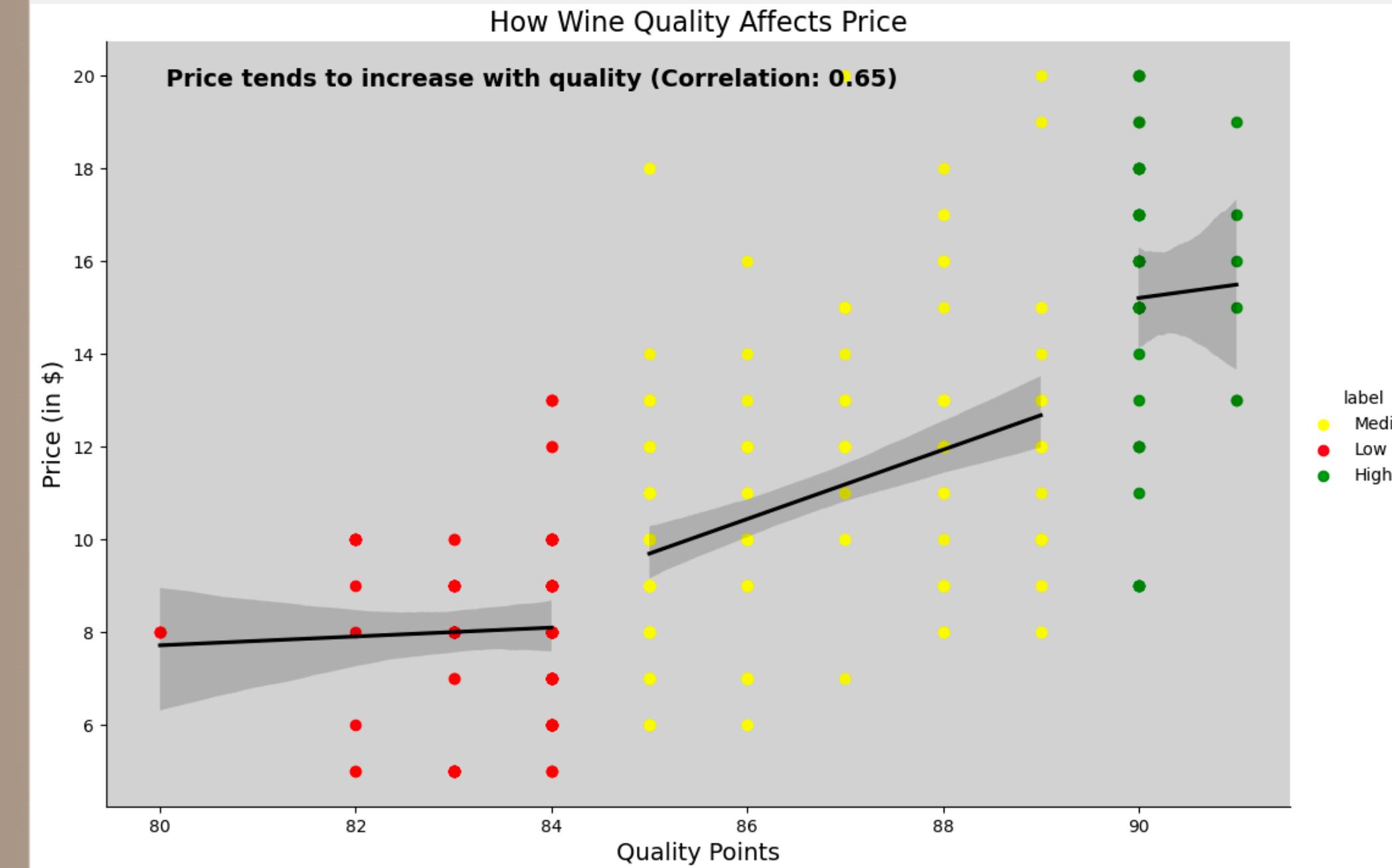
Higher Quality in White Wine: White wine has more high-quality samples, leading to better model performance.

Effect of Sampling:

- Red Wine: Higher accuracy without sampling, but models struggle with sampling techniques due to data imbalance.
- White Wine: More consistent accuracy with or without sampling, showing less impact from data imbalance.

Price Increases with Quality

- Correlation: 0.65
- Medium-Quality Wines: Responsive to quality improvements
- Low-Quality Wines: Stable pricing, less variation
- High-Quality Wines: Slight price change with quality



Business Recommendations for BlueBerry Winery

Improve Medium-Quality Wines:

- Slightly improve quality to raise prices and boost profits.



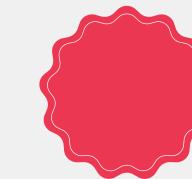
Keep Low-Quality Wines Affordable:

- Maintain lower prices to attract budget-conscious customers.



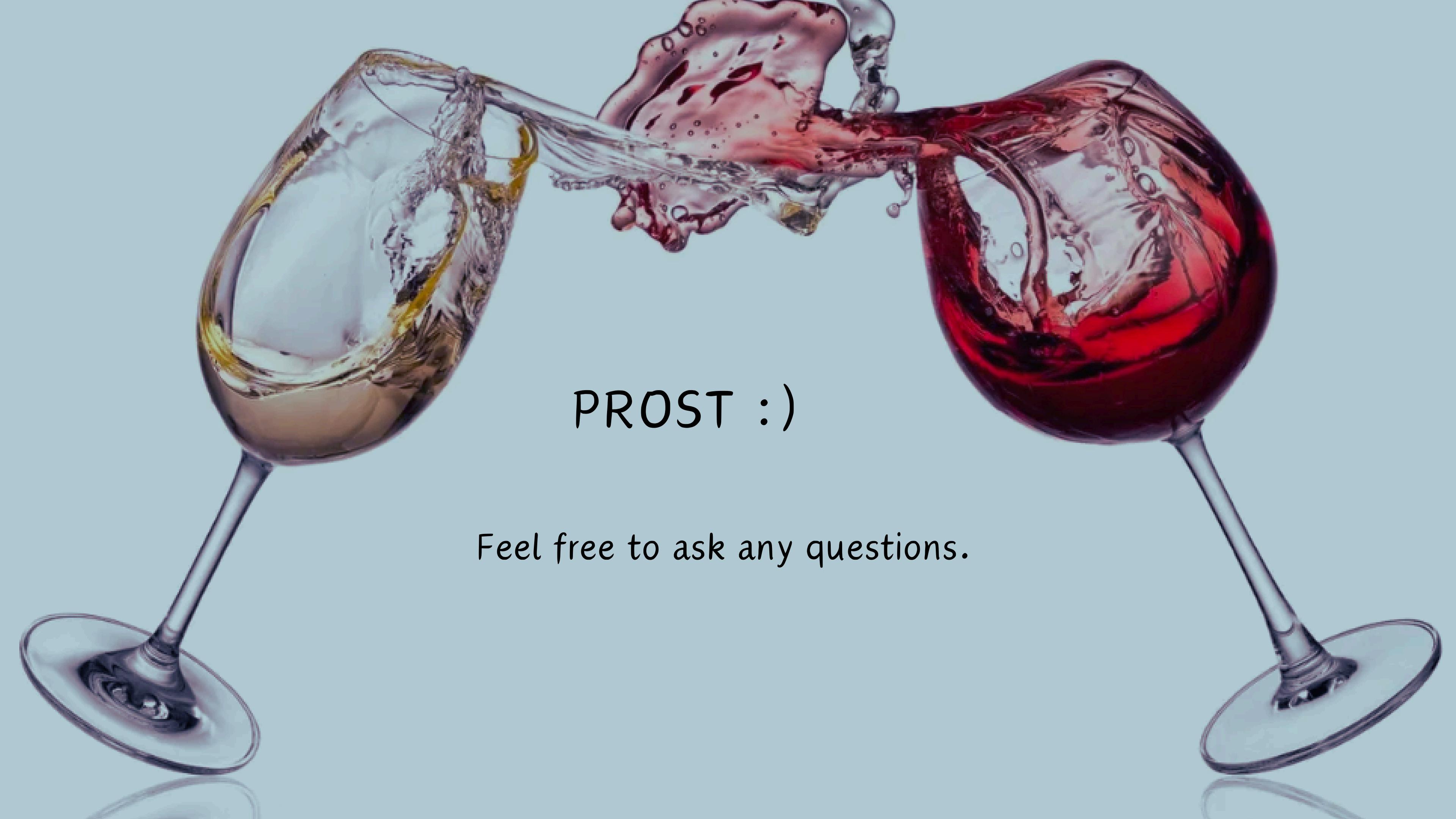
Differentiate High-Quality Wines:

- Focus on unique features to justify premium pricing.



Optimize Quality and Pricing with Data:

- Use the predictive model to produce consistent quality and set prices that reflect the true value.



PROST :)

Feel free to ask any questions.

