Physicochemical Properties Related to the Quality of Red Wines

: Analyzing Numerical properties that affect the quality of red wines using ordered probit regression models

Name: Jeesoo, Kim (Business/Computer Science)
Student ID: 20190741

Time Series Data Analysis and Forecasting (Capstone Design)
2024 Spring Semester
Professor Myungsuk, Kim

Ⅰ. Problem Motivation

The demand of imported wines have been increased in korea. The imported values of wines grew from 147.26 million dollars in 2012 to 192.45 million dollars in 2016.[1] Moreover, after the COVID-19 pandemic, imported wine consumption had been increased 69.6% in 2021 and it will increase 6.22% by value annually according to the article.[2]

Unlike in countries where high-end alcoholic beverages such as wine and whiskey are active, where most of them choose a tax method that levies taxes according to the volume of alcohol, Korea chooses a tax method that levies taxes on the price of alcohol to favor the price of distilled soju in its own country.[3] In addition, since Korea imposes tariffs, education taxes, liquor taxes, and value-added taxes on imported alcoholic beverages, the price of the same product is much higher than other OECD countries. The reality is that it is not easy to revise such laws because switching to a specific duty can lead to various complaints as the price of distilled soju, the most sold alcohol in Korea, becomes higher while the price of imported alcoholic beverages decreases.

For these social reasons, consumers who enjoy imported liquor such as wine and whiskey in Korea are bound to be particularly sensitive to price. There is a direct example about this. Wine-Searcher is a site that tells the average overseas price of the wine sold. Based on this site, the Chateau Mouton Rothschild 2016 is $923 on average. However, if you consider purchasing it in Korea, it varies from $1200 to $1900 depending on the seller as the prices of wines differ between the wine discounting shops and department stores.[4] The price range is much higher than the worldwide average.

In general, wine-buying consumers refer to the scores given by famous wine critics such as Robert Parker and James Suckling. If these famous critics give the wine a high score, it becomes famous in a short period of time with marketing phrases such as RP100(Robert Parker 100 points), resulting in a price premium. As a result the price will immediately rise due to the principle of supply and demand. For example, Chateau Mouton Rothschild in 2013, RP93 points, averages $564, but the average price of the same wine in 2016, RP100 points, averages $923.

Meanwhile, although the quality of wine is obviously good there are some wines that are sold at much lower prices because they did not have gained popularity yet. They are undervalued for some reasons such as not being a famous brand or not being evaluated by critics. These kinds of wines are not particularly expensive in Korea compared to the world average price as the importers make lower margins to sell this products.

In case of wine, since wine producers upload various numerical data along with the acidity and alcohol content of the wine on their site, it will be helpful for purchasing wines if we can know which factors affect the quality of wines. Therefore, in this study, for domestic consumers who are sensitive to price of wines (including the author of this paper), we would

1) 최민영, "레드와인 평균 판매가격, 수입가격의 11배", 경향신문(2018)
2) Jihyung Kim. "A Study on Demand Forecasting Change of Korea's Imported Wine Market after COVID-19 Pandemic." The Korean Journal of BigData 8, 2 (2023): 189-200.
3) 황혜빈, "위스키 가격, 한국이 해외보다 3배 비싼 이유... 세금 부과 방식 탓", IT조선(2023)
4) 김도균, "연말 와인값 천차만별... 백화점-점문점 26만원 차이", 중부일보(2022)

like to identify the attributes of wine that generally affect the quality of wines. And through this, we would like to suggest a way to find wines of good quality, although it has not gained popularity yet. Additionally, researching this topic will help wine producers make better wines. In fact, since champagne or red wine producers control sugar supplementation, acidity, and alcohol levels in their processes of producing wine, the following research may reduce their trial and error.

Ⅱ. Literature Review

In domestic research, studies have primarily focused on the influence of branding and label design on consumer psychology rather than predicting the quality of wine using quantitative data. Since wine is an unfamiliar product in contrast to the popular drink such as soju in Korea, it is natural that analysis related to the consumer behavior has been conducted in the past. This is also because it would be effective to derive strategies to promote sales by analyzing factors that influence purchases in the mid to low-priced wine segment, where the majority of consumption takes place. Apart from this, there have been attempts to predict the demand for wine using time series data analysis such as a study on demand forecasting change of Korea's imported wine market after COVID-19 pandemic.

On the other hand, in countries where wine is popularly consumed and produced, such as the United States and France, research has been conducted not only on marketing factors between wine and consumers, but also on the wine brewing research and the relationship between the numerical characteristics and quality of wine. According to the paper modeling wine preferences by data mining from physicochemical properties[5] a large dataset is considered when compared to other studies in this data mining approach to predict human wine taste preferences. The orginal article used SVD(support vector machine) to examine if there is any relationship between the physicochemical properties of wine and its quality. On the other hand, in this paper, the ordered logistic regression model was used to analyze this relationship in line with the project requirements. Using the logistic model could directly explain the impact of independent variables toward the quality of wine while SVD focuses on the classification.

Ⅲ. Statement of Research Objective

In this analysis, we tried to analyze what numerical factors affect the quality of red wine. In detail, the red wine data registered in the UCI repository used in this analysis include measurable numerical independent variables such as fixed acidity, volatile acid, citric acid, residual sugar, chloride, free sulfur dioxide, density, pH, sulfates, and alcohol. Among these variables, which variables affect the quality of wine and which variables have a greater influence were analyzed in this report. If this approach proves successfully, it may encourage wine enthusiasts, who previously relied on critic scores for wine quality, to attempt identifying

5) Cortez, P., Antonio Luíz Cerdeira, Fernando Almeida, Telmo Matos and José Reis. "Modeling wine preferences by data mining from physicochemical properties." Decis. Support Syst. 47 (2009): 547-553.

high-quality wines based on the physicochemical characteristics which are partially disclosed by wine makers. In short, the main research objective is to explain the impact of the physicochemical properties on the quality of wine so that customers, marketers, and even wine makers could reference the result to get better results on their purposes. The limitations of this analysis were discussed in the final part of the report.

Ⅳ. Description of Data & Applied Methodologies

This red wine dataset is related to 1599 red variants of the Portuguese "Vinho Verde" win e[6] and is downloaded from the UCI machine learning repository. Wine grade and physicochemical data such as pH, acidity, sugar, and alcohol content for 1599 red wine samples are summarized, and the following is a description of each survey item variable.

Dependent variable and independent variables used in this paper are explained in detail. Wine quality is the dependent variable that we want to analyze in order to find which independent variables influences the quality of wine. Independent variables are physicochemical properties that are related to wine. These variables are selected as wine certification is generally assessed by sensory tests and physicochemical tests.[7] Physicochemical tests are routinely used to characterize wines including pH values, alcohol, density. However in reality, wine qualities are measured by human expert's sensory tasting which can be stressed. Moreover, according to the previous research the relationships between the sensory analysis and physicochemical are very complex and still not fully understood.[8] Therefore, analysis on the physicochemical variables are taken in this paper and their characteristics are shown as below table.

Name	Units	Type	Description
Fixed acidity	g/dm^3	Continuous	Most acids involved with wine
Volatile acidity	g/dm^3	Continuous	Amount of acetic acid in wine, which at too high can lead to an unpleasant, vinegar taste
Citric acid	g/dm^3	Continuous	Found in small quantities. Citric acid can add freshness and flavor to wines
Residual sugar	g/dm^3	Continuous	Amount of sugar remaining after fermentation
Chlorides	g/dm^3	Continuous	Amount of salt in the wine
Free sulfur dioxide	mg/dm^3	Continuous	Free form of SO2 exists in equilibrium between molecular SO2 and bisulfite ion; prevents microbial growth and the oxidation of wine

6) Cortez,Paulo, Cerdeira,A., Almeida,F., Matos,T., and Reis,J.. (2009). Wine Quality. UCI Machine Learning Repository. https://doi.org/10.24432/C56S3T.
7) S. Ebeler, Flavor Chemistry — Thirty Years of Progress, Kluwer Academic Publishers, 1999, pp. 409-422, chapter Linking flavour chemistry to sensory analysis of wine.
8) A. Legin, A. Rudnitskaya, L. Luvova, Y. Vlasov, C. Natale, A. D'Amico, Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception, Analytica Chimica Acta 484 (1) (2003) 33-34.

Total sulfur dioxide	mg/dm^3	Continuous	Amount of free and bound forms of S02; in low concentrations, SO2 is mostly undetectable in wine, but at free SO2 concentrations over 50 ppm, SO2 becomes evident in the nose and taste of wine
Density	g/cm^3	Continuous	Density of water is close to that of water depending on the percent alcohol and sugar content
pH	-	Continuous	Describes how acidic or basic a wine is on a scale from 0 (very acidic) to 14 (very basic); most wines are between 3-4 on the pH scale
Sulphates(SO_4^{2-})	g/dm^3	Continuous	A wine additive which can contribute to sulfur dioxide gas (S02) levels, which acts as an antimicrobial and antioxidant
Alcohol	vol.%	Continuous	the percent alcohol content of the wine
Quality	score	Discrete	Output variable. Integer score between 0 and 10.

<Table 1: Variables information>

Summary statistics are given as the below table. In this part, some important points were mentioned among the data in the table below. First, the quality variable, which is a dependent variable, is not a continuous variable. Instead, it is a discrete variable about wine score which could be integer values between 0 and 10. That is, there is no data value between any two consecutive integers. Note that in this data, the minimum value of wine quality score is 3 and maximum is 8. Therefore, there are six possible wine score values: 3, 4, 5, 6, 7, and 8. As a result, in this paper, we analyzed this quality data using the ordered probit regression(oprobit for short) method. According to encyclopedia, ordered probit is a generalization of the widely used probit analysis to the case of more than two outcomes of an ordinal dependant variable.[9] For example, a dependant variable for which the possible values have a natural ordering, as in poor, fair, good, very good, etc. Moreover the model is selected as it provides an appropriate fit to these data, while preserving the ordering of response options making no assumptions of the interval distances between options.[10]

Next, independent variables of the below table are physicochemical values that are recorded as continuous numerical values. It can be seen in the table below that the density variable have a very small deviation between the values, while that of other dependent variables are quite large.

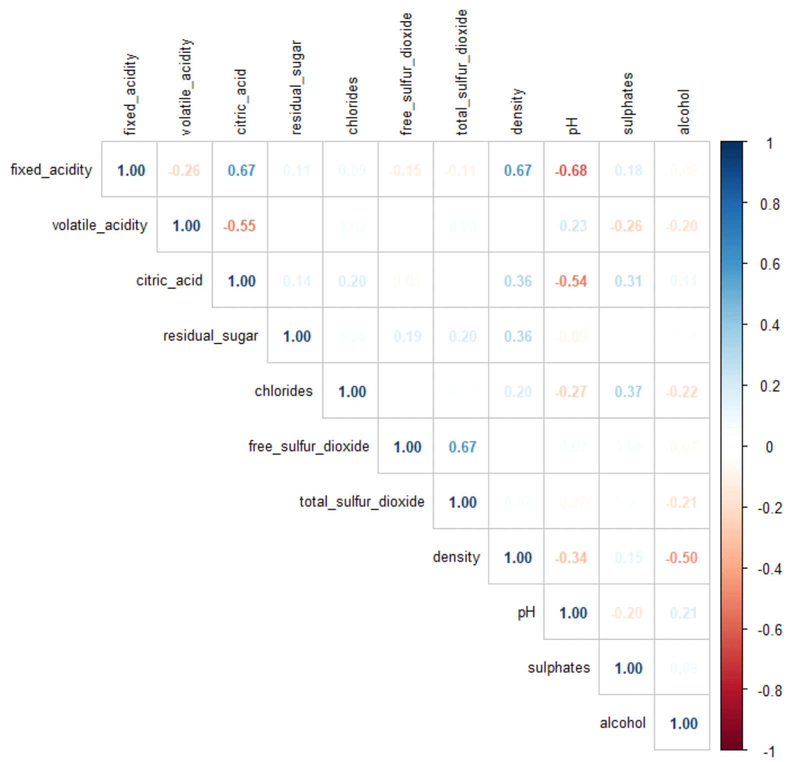
9) Ordered probit. (n.d.). In Wikipedia. Retrieved April 6, 2024, from https://en.wikipedia.org/wiki/Ordered_probit

10) Liddell, T; Kruschke, J (2018). "Analyzing ordinal data with metric models: What could possibly go wrong?". Journal of Experimental Social Psychology. 79: 328-348.

Name	fixed_acidity	volatile_acidity	citric_acid	residual_sugar
Min	4.60	0.1200	0.000	0.900
1st Quarter	7.10	0.3900	0.090	1.900
Median	7.90	0.5200	0.260	2.200
Mean	8.32	0.5278	0.271	2.539
3rd Quarter	9.20	0.6400	0.420	2.600
Max	15.90	1.5800	1.000	15.500
Name	**chlorides**	**free_sulfur_dioxide**	**total_sulfur_dioxide**	**density**
Min	0.01200	1.00	6.00	0.9901
1st Quarter	0.07000	7.00	22.00	0.9956
Median	0.07900	14.00	38.00	0.9968
Mean	0.08747	15.87	46.47	0.9967
3rd Quarter	0.09000	21.00	62.00	0.9978
Max	0.61100	72.00	289.00	1.0037
Name	**pH**	**sulphates**	**alcohol**	**quality**
Min	2.740	0.3300	8.40	3.000
1st Quarter	3.210	0.5500	9.50	5.000
Median	3.310	0.6200	10.20	6.000
Mean	3.311	0.6581	10.42	5.636
3rd Quarter	3.400	0.7300	11.10	6.000
Max	4.010	2.0000	14.90	8.000

<Table 2: Summary statistics on each attributes>

Meanwhile, prior to the analysis, the below diagram was created to examine whether there is any significant correlations between each independent variable. In the diagram below, correlations between independent variables are indicated by numbers. Note that darker the color, the more significant the corresponding value is. (Fixed_acidity, Citric_acid), (Fixed_acidity, density), (Fixed_acidity, pH), (Volatile_acidity, Citric_acid), (Citric_acid, pH) have correlation values bigger than 0.5. It is speculated that this is probably because the independent variables are different variables that exhibit similar characteristics of wine acidity. Therefore, it is suggested that the variable selection of the final model should be carefully taken considering the correlations between independent variables.



<Figure 1: Correlation matrix between independent variables>

Ⅴ. Analysis result

In this part, the analysis results and additional tests were executed using the regression model selected based on the evidence explained before.

The ordinal probit regression analysis was done by using polr() with probit method in R language. The model was first built with the full model with 11 independent variables. Then, we select the variables that is appropriate for the predicting wine quality scores using stepwise variable selection method. The used direction for stepwise variable selection was both. Finally, we analyzed the impact of each independent variable on the wine quality score with analysis of deviance table(type 2 tests) to carefully select the physicochemical independent variable that is likely to influence the quality of wine via p-value approach. As a result this paper select 4 independent variables which could not be removed from the model to predict a wine quality score and explain the impact of those variables. The analysis results of the ordinal probit model using stepwise variable selection method is given on the below table.

```
> summary(ordinal_model)
Call:
polr(formula = quality ~ volatile_acidity + citric_acid + chlorides +
    free_sulfur_dioxide + total_sulfur_dioxide + pH + sulphates +
    alcohol, data = train_set, Hess = TRUE, method = "probit")

Coefficients:
                         Value Std. Error t value
volatile_acidity     -1.907193   0.236181  -8.075
citric_acid          -0.364365   0.243650  -1.495
chlorides            -3.157698   0.786822  -4.013
free_sulfur_dioxide   0.007553   0.004527   1.668
total_sulfur_dioxide -0.007229   0.001404  -5.149
pH                   -0.876357   0.267399  -3.277
sulphates             1.265030   0.214074   5.909
alcohol               0.515123   0.035483  14.518

Intercepts:
    Value   Std. Error t value
3|4 -1.4718  0.9578    -1.5366
4|5 -0.5977  0.9416    -0.6348
5|6  1.5215  0.9360     1.6256
6|7  3.0838  0.9408     3.2778
7|8  4.7650  0.9534     4.9976

Residual Deviance: 2291.159
AIC: 2317.159

> Anova(ordinal_model, type = "II")
Analysis of Deviance Table (Type II tests)

Response: quality
```

```
                       LR Chisq Df Pr(>Chisq)
volatile_acidity         49.326  1  2.168e-12 ***
citric_acid              -8.651  1     1.0000
chlorides                 2.589  1     0.1076
free_sulfur_dioxide      -6.938  1     1.0000
total_sulfur_dioxide     20.081  1  7.422e-06 ***
pH                       -2.926  1     1.0000
sulphates                27.980  1  1.225e-07 ***
alcohol                 202.595  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

[1] "Ordinal Model Accuracy: 0.61"
```

<Table 3: Code result of ordinal model>

Notice the model uses 8 independent variables to predict the quality of red wines and the accuracy of the prediction was 0.61. However, by checking the analysis of deviance table, only 4 variables was statistically significant variable to predict the quality of red wines. That means it is enough for predicting the quality of wines using only 4 variables: volatile_acidity, total_sulfur_dioxide, sulphates, alcohol. Therefore this paper selected those 4 independent variables for building the model and the following result is as follows.

```
> summary(ordinal_model_revised)
Call:
polr(formula = quality ~ volatile_acidity + total_sulfur_dioxide +
    sulphates + alcohol, data = train_set, Hess = TRUE, method = "probit")

Coefficients:
                          Value Std. Error t value
volatile_acidity      -1.975186   0.198315  -9.960
total_sulfur_dioxide -0.005427   0.001003  -5.412
sulphates              0.964964   0.191601   5.036
alcohol                0.514164   0.033313  15.434

Intercepts:
    Value   Std. Error t value
3|4  1.5789  0.4288     3.6820
4|5  2.4379  0.4021     6.0625
5|6  4.5282  0.4006    11.3022
6|7  6.0654  0.4181    14.5059
7|8  7.7308  0.4504    17.1638

Residual Deviance: 2318.809
AIC: 2336.809

> Anova(ordinal_model_revised, type = "II")
Analysis of Deviance Table (Type II tests)

Response: quality
                       LR Chisq Df Pr(>Chisq)
volatile_acidity         74.697  1  < 2.2e-16 ***
total_sulfur_dioxide     19.411  1  1.054e-05 ***
```

```
sulphates                15.059  1  0.0001042 ***
alcohol                 230.217  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

[1] "Revised Ordinal Model Accuracy: 0.61"
```

The result is that even though we removed 4 variables from the original model, the accuracy results of the prediction were same. Therefore it is obvious to choose the reduced model for predicting the quality of red wines. In the rest of this section, interpretation of the model and additional verification regarding the model is explained in detail.

Prior to explain the result, correlations between the independent variables that we calculated in the earlier section is revisited. That is checking whether or not there is a pair of independent variables that has high relationship between each other. As the model selected volatile acidity, total sulfur dioxide, sulphate, and alcohol there is no such pair in the model.

Also, regarding the VIF of the model, there is no such problem found as the value stays very small as seen in below table.

volatile_acidity	total_sulfur_dioxide	sulphates	alcohol
1.045860	1.020877	1.048296	1.017230

The actual impact of each independent variable can be explained using the above coefficients result of the revised ordinal model. In this part, the impacts are explained in probability wise using the calculated results. The possible impacts on the wine customers, importers, and wine makers that could use these numerical data are explained at the next section.

Volatile acidity	As volatile acidity increases, the wine quality rating tends to decrease. In terms of odds, for each one-unit increase in volatile acidity while others unchanged, the odds of belonging to the respective grade decrease by approximately 0.139 times(e^-1.975).
Total sulfur dioxide	An increase in total sulfur dioxide is associated with a decrease in wine quality rating. In terms of odds, for each one-unit increase in total sulfur dioxide while others unchanged, the odds of belonging to the respective grade decrease by approximately 0.995 times (e^-0.005).
Sulphate	Higher levels of sulphates are associated with higher wine quality ratings. In terms of odds, for each one-unit increase in sulphate while others unchanged, the odds of belonging to the respective grade increase by approximately 2.625 times (e^0.965).
Alcohol	Increasing alcohol content is associated with higher wine quality ratings. In terms of odds, for each one-unit increase in alcohol content while others unchanged, the odds of belonging to the respective grade increase by approximately 1.673 times (e^0.514).

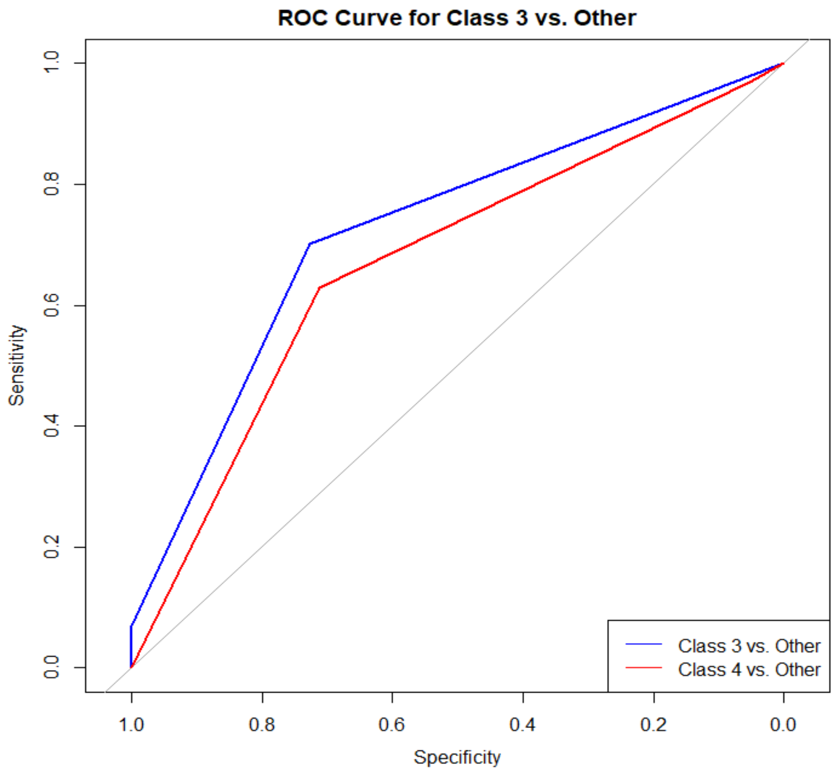
The correlation between total sulfur dioxide and sulphate is 0.04. That means those are different types that are related to So2 but there is no linear relationship between them. Note that a wine addictive which can contribute to So2 levels that are informed by the wine label(So2 warning) is total sulfur dioxide(mg/dm^3). To examine whether the model result could change if we only use one of two variables between total sulfur dioxide and sulphate, the coefficient result was almost same. However as the model accuracy downs a bit, the model with above four variables were used in this research. In the interpretation, the impact of sulphate is much smaller than the total sulfur dioxide as the Anova test(type 2) tells. More specifically, the range of total sulfur dioxide of the data 260 while with higher variance that those of sulphate. Using standard deviation, total sulfur dioxide is 32.9 and sulphate is 0.17 in value. Meanwhile in average, the difference in total sulfur dioxide influences the wine rating for about $0.995^{33}=0.86$times and sulphate influences the wine rating for about $2.625^{0.17}=1.17$times. Such results could make some confusion in interpretation, but consumers could think it is better to be low in total sulfur dioxide and high in sulphate which do not has correlation but affects the So2. Actually data that have low total sulfur dioxide and high sulphate averages of 6.23 wine quality ratings where maximum quality is 8 and total average rating is 5.63. As a result, rather than considering just either one variable, considering both could be a better choice as considering either one had resulting average ratings of 5.76 and 5.64 each.

It is obvious that volatile acidity impacts negatively toward the quality of red wines as volatile acidity is amount of acetic acid in wine which can lead to an unpleasant, vinegar taste by definition. Finally, most importantly for wine consumers, the alcohol content is associated with higher quality of wine when it is high. All others equal, the higher alcohol(>3rd Quarter) resulted in wine quality ratings of average 6.23 while lower alcohol(<1st Quarter) resulted in that of average 5.23 which is even much lower than the total average of 5.64.

Finally, the confusion matrix and the area under curve was analyzed in this section. The confusion matrix is as follows.

	Predicted 1	Predicted 2	Predicted 3	Predicted 4	Predicted 5	Predicted 6
Actual 1	0	0	5	0	0	0
Actual 2	0	0	11	7	0	0
Actual 3	0	0	123	46	0	0
Actual 4	0	0	50	115	9	0
Actual 5	0	0	3	20	6	0
Actual 6	0	0	0	4	1	0

Calculated accuracy using the confusion matrix is 0.61 which is same as the result in table 3. Multi-class AUC is calculated using multiclass.roc method in R included in the pROC library. The calculated result is 0.748 which is quite good considering that the original model used for this problem is support vector machine model which focuses on classification while the approach used in this paper could focus on the interpretation of each variable as well. AUC visualization is executed for class 3 versus other and for class 4 versus other as there are so many lines that are shown as the result of multi-class AUC. The selected are the most frequent variables as seen in the above table. The visualization results are shown below. Each calculated result is 0.724 and 0.669.



Ⅴ. Expected Original Contribution

The original expected result was there is some physicochemical properties that are related to the quality of red wines. From the domain knowledge, the author expected that some type of acidity(Fixed acidity, Volatile acidity, citric acid, pH), residual sugar, alcohol, and sulphate could affect the quality of red wines. First is acidity as many wine experts evaluate the potential of the red wine with its acidity. The higher acidity wine could be aged for a long time which means it can develops over time. Second, there exists natural wines which do not add sulphate/sulfur dioxide to the wine for the preserving purposes while there is some controversies regarding the quality of the natural wine. So, the author expected that sulphate could affect the quality of red wines. Finally, just as the champagne is categorized based on the amount of added sugar, although it is not red wine, it was thought that quality of red wines may also be related to residual sugar.

As a result of the section before, four physicochemical variables are related to the quality of red wines. Those are volatile acidity, total sulfur dioxide, sulphate, and alcohol. The impact in terms of odds ratio was described in detail in the previous section.

This results could have some contribution to the wine business. In regard to the sulphate added for the preserving wines, it has partial effects on the quality of red wines. That is, it is useful when both considering total sulfur dioxide and sulphate that are added in red wines. So, for this reason, consumers won't have to overpay for natural wines, which have become hugely popular for not adding So2 regarding elements in recent years as adding total sulfur dioxide itself has extremely small impacts on the quality of red wines. Also, importers that are looking for the new wines and wine makers that are recently considering starting the business would rather consider the fact that adding sulphate is even good sometimes for preserving wines as well as its quality when total sulfur dioxide is low and sulphate is high. In that case importer could import better quality wines as natural wines have higher price as they need more money to produce wines with bio-dynamic techniques that do not use sulphate.

Next, in regard to the alcohol volume of the wine, it has positive effects on the quality of red wines as alcohol increases. Therefore, consumers who hesitate to purchase wines with higher alcohol volume than those they usually drink might actually find it to be an opportunity to encounter better wines. By trying wines with higher alcohol content, they may discover wines of better quality than what they are accustomed to. This applies to the wine importers and wine makers as well.

Finally, the volatile acidity can directly affect the quality of red wines negatively. This information could have some implication to the wine makers and importers as they can directly evaluate this type of physicochemical values. Volatile acidity is amount of acetic acid in wine, which at too high can lead to an unpleasant, vinegar taste. So, the importers can evaluate this property through the samples they have before making the decision.

The importance of interpreting the approximate effects of each variable on wine quality lies in the fact that, in reality, useful numerical data that is available for consumers are limited, making it impossible to input all variables directly into classification models. Therefore, providing rough interpretations of variables such as alcohol content, acidity, and sulfur dioxide addition, when everything else is held constant, can serve as an effective indicator for consumers who are looking for some good value wines.

Ⅵ. Limitation of this Paper

Although the results were clear, there are some limitations of this paper. First, wine is not all about numerical data. Sometimes the mood and environment the person enjoys that wine could affect more to the recognized quality of wine than the actual quality of wine. So, there are many researches about the relationships between the brand, price, mood, and other situational variables of wine and the customer satisfaction. Even there is a word that who we enjoys with is more important that than the actual quality of wines as the perceived responses are subjective for each person.

Considering the domain knowledge, this result could be very limited as the sample data is 1599 red variants of the Portuguese “Vinho Verde” wine because there are some differences in characteristics of wine between the country of origin, grape type(Cabernet sauvignon, Pinot noir, Cabernet France, etc.), specific region as experts refer to the characteristic differences in regard to the region as terroir. The difference caused by terroir can be said to be the difference induced from the characteristics of the soil that is differ from each region. Thus, as a result, the data only from the Portuguese “Vinho Verde” region could be difficult to generalize to all the red wines. As a result, later studies could be executed targeting the popular country and popular grape type such as Cabernet Sauvignon red wines in France. That could be practical as many volume of wines that are consumed are cabernet sauvignon(grape type) wines in France regardless of the price. This study could give an alternative evaluation criteria that might be more useful than Grand cru classification which was established in 1855 for bordeaux red wines in France.

Ⅶ. References

1. 최민영, “레드와인 평균 판매가격, 수입가격의 11배”, 경향신문(2018)

2. Jihyung Kim. "A Study on Demand Forecasting Change of Korea’s Imported Wine Market after COVID-19 Pandemic." The Korean Journal of BigData 8, 2 (2023): 189-200.

3. 황혜빈, “위스키 가격, 한국이 해외보다 3배 비싼 이유... 세금 부과 방식 탓”, IT조선(2023)

4. 김도균, “연말 와인값 천차만별... 백화점-점문점 26만원 차이”, 중부일보(2022)

5. Cortez, P., Antonio Luíz Cerdeira, Fernando Almeida, Telmo Matos and José Reis. “Modeling wine preferences by data mining from physicochemical properties.” Decis. Support Syst. 47 (2009): 547-553.

6. S. Ebeler, Flavor Chemistry — Thirty Years of Progress, Kluwer Academic Publishers, 1999, pp. 409-422, chapter Linking flavour chemistry to sensory analysis of wine.

7. A. Legin, A. Rudnitskaya, L. Luvova, Y. Vlasov, C. Natale, A. D'Amico, Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception, Analytica Chimica Acta 484 (1) (2003) 33-34.

8. Ordered probit. (n.d.). In Wikipedia. Retrieved April 6, 2024, from https://en.wikipedia.org/wiki/Ordered_probit

9. Liddell, T; Kruschke, J (2018). "Analyzing ordinal data with metric models: What could possibly go wrong?". Journal of Experimental Social Psychology. 79: 328-348.