Introduction

Wine history is extraordinarily long, and the oldest wine can be traced back to Neolithic or New Stone Age as the cavemen's tools were found to have traces of chemicals which were present in wine. The earliest wine industry started in Armenia around 4100 BCE, where the winemakers were mostly farmers who used simple clay pots and the process took a lot of time. An Armenian cave named Vayots Dzor is where archaeologists uncovered the remains of once a fully working winery along with equipment such as fermentation vats, grape press, range of glasses and cups. (Good Pair Days, 2024)

There are 9 primary wine styles: Sparking wines, Light-bodied white wines, Full-bodied white wines, Aromatic (sweet) white wines, Rose, Light-bodied red wines, Medium-bodied red wines, Full-bodied red wines, and Dessert wines (The Nine Primary Styles of Wine, 2020). Red and white wine are the two most common wines when it comes to discussion. The focus of the assessments would be on the white wine.

The white wine is made up of white, sometimes green grapes. The light skin of these grapes contributed to its juice in plain white colour. The acidity of white wine is higher and gives a refreshing quality due to early harvest (Good Pair Days, 2024). The pH level of white wine is between 3 to 3.4 (O'Donnell, 2017). Meanwhile, the alcohol content ranged from 5% to 14%, and an average of 10% alcohol by volume (ABV) (Weatherwax, 2024).

Factors that affect quality of wine included grape ripeness which affect the sweetness and acidity level, cold soaking that allows the skins to stay in contact for rich colour development and gives fruitiness taste in wine, temperature of fermentation to influence colour and flavour, aging container types where the oxygen exposure can be controlled, and capping method which possess different of risks in contaminating the scents of wine (o-a.com.au, 2018).

To study the wine quality, several statistical measures would be calculated and taken into consideration for estimation and hypothesis testing. The relationship between the variables would also be visualized via graphical methods to give a clearer access to their correlation towards to overall quality of wine. Assessment of the model will then be done by calculating the Coefficient of Determination (R-squared).

Dataset

The dataset of Wine Quality (UCI Machine Learning Repository, 2023) for white wine consists of 12 columns. The detail for each feature is shown in the table below:

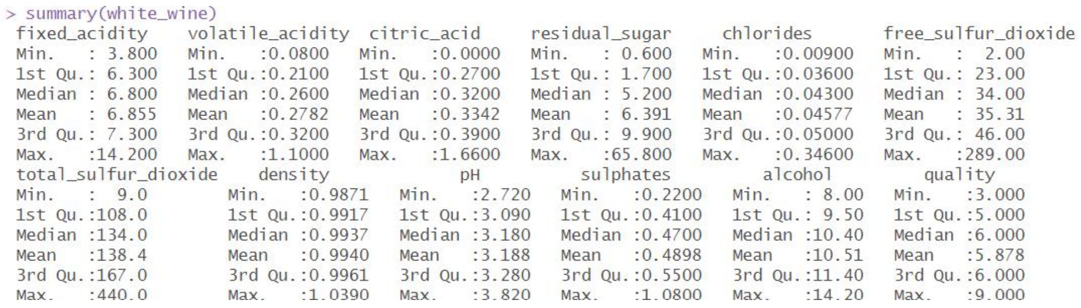
Column Name	Variable Name	Definition / Description
Fixed acidity	fixed_acidity	Wine's natural acids, grams of tartaric acid per liter. Acids that remain in the liquid when it is boiled. Unit: g/L **
Volatile acidity	volatile_acidity	The acids that are readily evaporate, the lower the better. Unit: g/L **
Citric acid	citric_acid	One of the fixed acids A type of acid present in grapes and citrus fruits. It gives a sour taste. Unit: g/L **
Residual sugar	residual_sugar	Sugar left in wine after the completion of alcoholic fermentation, and addition made when bottling the wine (Kaan, 2018) Unit: g/L **
Chlorides	chlorides	One of the contributors to production of TCA that causes cork taint (Purdue Extension Commercial Winemaking Production Series Chlorine Use in the Winery, n.d.). Gives the tongue a sensation of saltiness (Maltman, 2013). Unit: mg/L *
Free sulphur dioxide (FSO2)	free_sulphur_dioxide	Sulphur dioxides present in the wine that is not bound to other chemicals yet, able to protect the wine (Howes, 2017) as anti-microbial agent and

		antioxidant (Wine Education Topic: Sulfur Dioxide in Wine, 2024) Unit: mg/L *
Total sulphur dioxide (TSO2)	total_sulphur_dioxide	Total amount of Sulphur dioxide in the wine, including FSO2 and those that bound or react to other chemicals in the wine (https://www.facebook.com/midwestgrapeandwine, 2018). Unit: mg/L *
Density	density	Mass per unit volume of wine at 20°C. Range: 0.9912 to 1.0138 g/cm3 (Iwona Budziak-Wieczorek et al., 2023)
pH	pH	pH level, or acid concentration in wine. Range for white wine: 3.0 - 3.4 (O'Donnell, 2017)
Sulphates	sulphates	Mineral salts found in soil and some plants (What Are Wine Sulfites & Which Wines Are Low Sulfite Wines? - Bright Cellars, 2022). Product from redox reaction between Sulphur dioxide and oxygen to prevent other compounds from oxidation (OIV COLLECTIVE EXPERTISE DOCUMENT SO2 and WINE: A REVIEW, n.d.). Unit: mg/L
Alcohol	alcohol	Alcohol percentage of the wine, varies from 5% to 12%, average of 10% (Weatherwax, 2024)
Quality	quality	Quality of the wine, rate from 0 to 10

* mg/L = milligram per litre
** g/L = grams per litre

Descriptive analysis

Descriptive analysis is carried out for each of the features to find the respective values for mean, median, Q1, Q3, minimum and maximum



```
> summary(white_wine)
 fixed_acidity    volatile_acidity  citric_acid     residual_sugar     chlorides         free_sulfur_dioxide
 Min.   : 3.800   Min.   :0.0800   Min.   :0.0000   Min.   : 0.600   Min.   :0.00900   Min.   :  2.00    
 1st Qu.: 6.300   1st Qu.:0.2100   1st Qu.:0.2700   1st Qu.: 1.700   1st Qu.:0.03600   1st Qu.: 23.00    
 Median : 6.800   Median :0.2600   Median :0.3200   Median : 5.200   Median :0.04300   Median : 34.00    
 Mean   : 6.855   Mean   :0.2782   Mean   :0.3342   Mean   : 6.391   Mean   :0.04577   Mean   : 35.31    
 3rd Qu.: 7.300   3rd Qu.:0.3200   3rd Qu.:0.3900   3rd Qu.: 9.900   3rd Qu.:0.05000   3rd Qu.: 46.00    
 Max.   :14.200   Max.   :1.1000   Max.   :1.6600   Max.   :65.800   Max.   :0.34600   Max.   :289.00    
 total_sulfur_dioxide    density             pH          sulphates         alcohol          quality     
 Min.   :  9.0        Min.   :0.9871   Min.   :2.720   Min.   :0.2200   Min.   : 8.00   Min.   :3.000  
 1st Qu.:108.0        1st Qu.:0.9917   1st Qu.:3.090   1st Qu.:0.4100   1st Qu.: 9.50   1st Qu.:5.000  
 Median :134.0        Median :0.9937   Median :3.180   Median :0.4700   Median :10.40   Median :6.000  
 Mean   :138.4        Mean   :0.9940   Mean   :3.188   Mean   :0.4898   Mean   :10.51   Mean   :5.878  
 3rd Qu.:167.0        3rd Qu.:0.9961   3rd Qu.:3.280   3rd Qu.:0.5500   3rd Qu.:11.40   3rd Qu.:6.000  
 Max.   :440.0        Max.   :1.0390   Max.   :3.820   Max.   :1.0800   Max.   :14.20   Max.   :9.000  
```

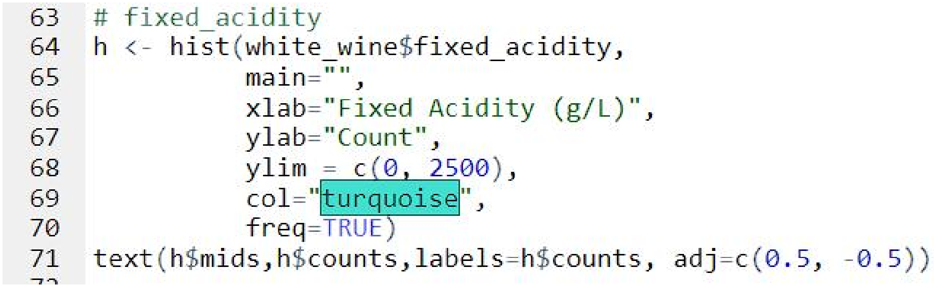
Graphs

To visualize the patterns and relationship between features, graphs are plotted along with the snippets shown

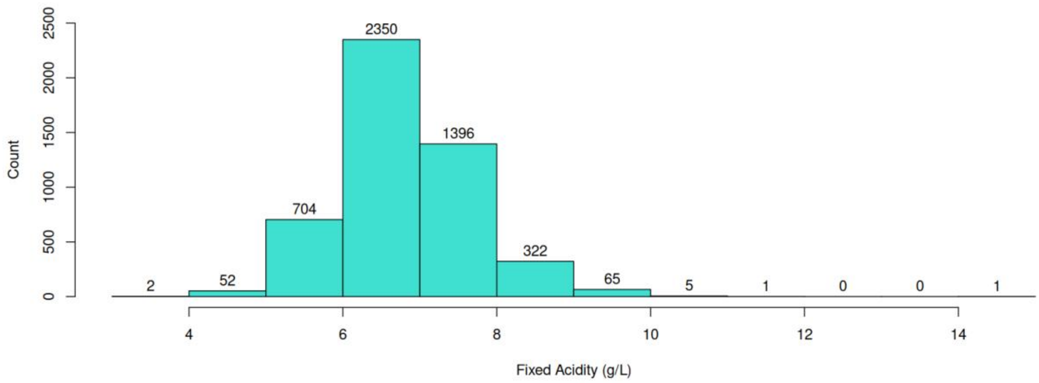
Bar chart and histogram

The graphs plotted below show a bell shape curve that follows normal distribution trend.

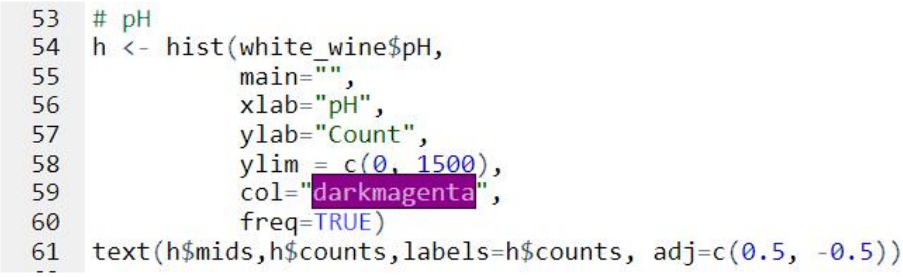
1. Fixed Acidity



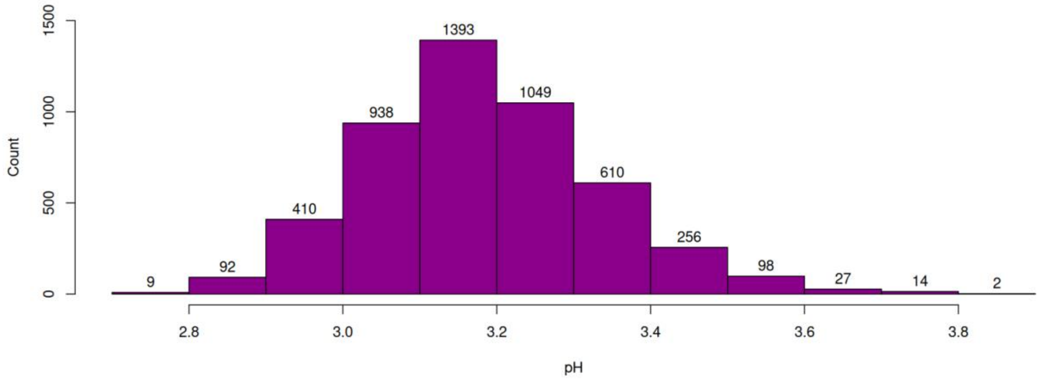
```
# fixed_acidity
h <- hist(white_wine$fixed_acidity,
          main="",
          xlab="Fixed Acidity (g/L)",
          ylab="Count",
          ylim = c(0, 2500),
          col="turquoise",
          freq=TRUE)
text(h$mids,h$counts,labels=h$counts, adj=c(0.5, -0.5))
```



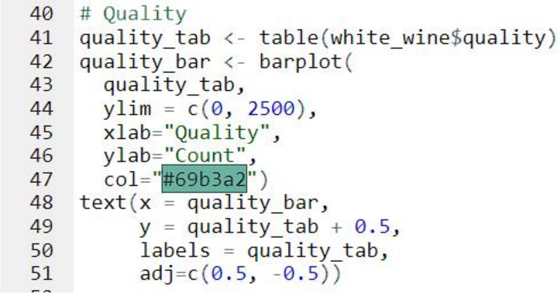
2. pH



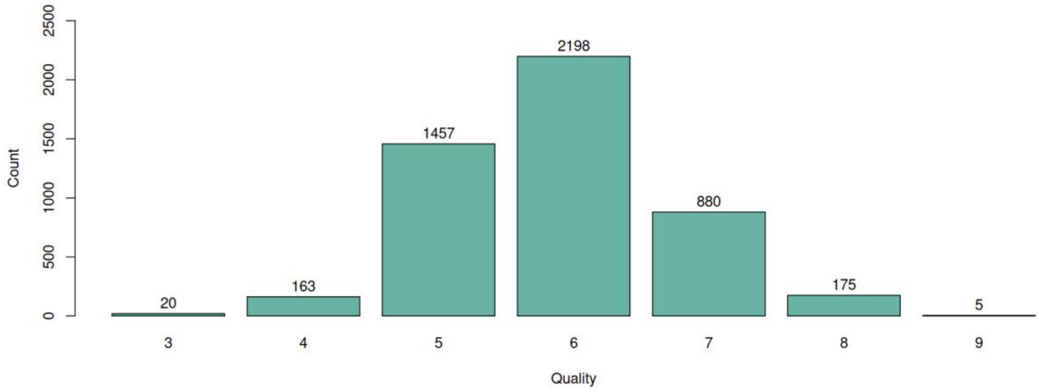
```r
# pH
h <- hist(white_wine$pH,
          main="",
          xlab="pH",
          ylab="Count",
          ylim = c(0, 1500),
          col="darkmagenta",
          freq=TRUE)
text(h$mids,h$counts,labels=h$counts, adj=c(0.5, -0.5))
```



3. Quality

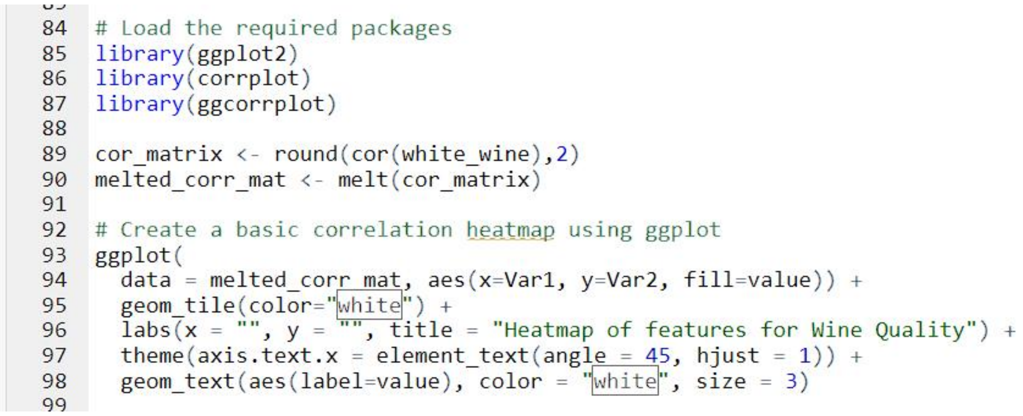


```r
# Quality
quality_tab <- table(white_wine$quality)
quality_bar <- barplot(
  quality_tab,
  ylim = c(0, 2500),
  xlab="Quality",
  ylab="Count",
  col="#69b3a2")
text(x = quality_bar,
     y = quality_tab + 0.5,
     labels = quality_tab,
     adj=c(0.5, -0.5))
```



Correlation Heatmap

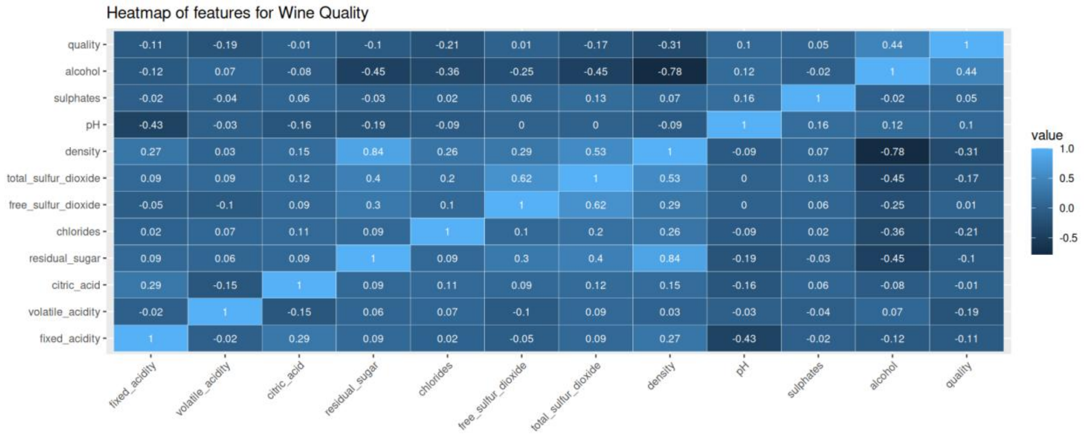
The heatmap shows the relationship between the features. From the colours and labels, Alcohol and density have strong negative correlation of -0.78, while Alcohol and Residual Sugar have strong positive correlation of 0.84.



```r
# Load the required packages
library(ggplot2)
library(corrplot)
library(ggcorrplot)

cor_matrix <- round(cor(white_wine),2)
melted_corr_mat <- melt(cor_matrix)

# Create a basic correlation heatmap using ggplot
ggplot(
  data = melted_corr_mat, aes(x=Var1, y=Var2, fill=value)) +
  geom_tile(color="white") +
  labs(x = "", y = "", title = "Heatmap of features for Wine Quality") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  geom_text(aes(label=value), color = "white", size = 3)
```

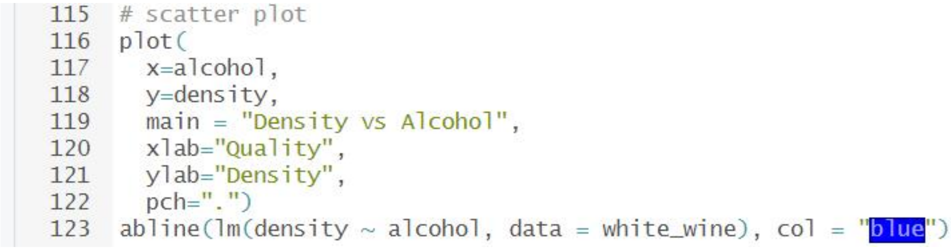


Heatmap of features for Wine Quality

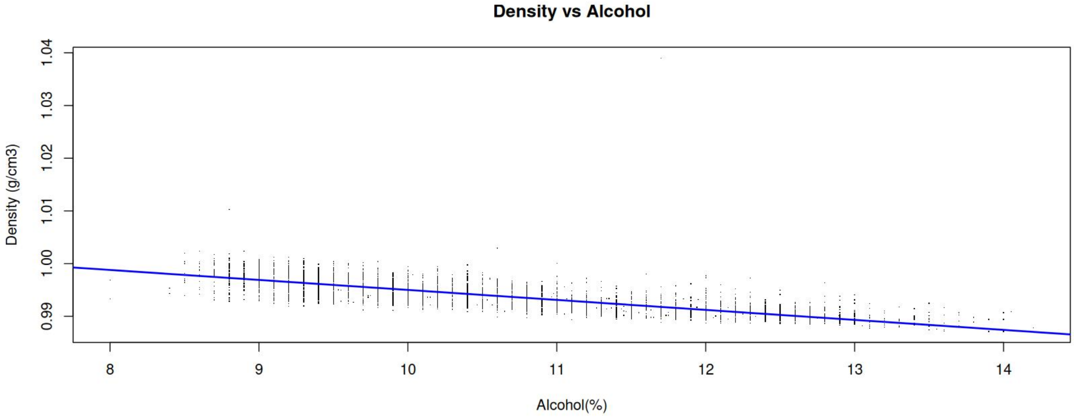
	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide	total_sulfur_dioxide	density	pH	sulphates	alcohol	quality
quality	-0.11	-0.19	-0.01	-0.1	-0.21	0.01	-0.17	-0.31	0.1	0.05	0.44	1
alcohol	-0.12	0.07	-0.08	-0.45	-0.36	-0.25	-0.45	-0.78	0.12	-0.02	1	0.44
sulphates	-0.02	-0.04	0.06	-0.03	0.02	0.06	0.13	0.07	0.16	1	-0.02	0.05
pH	-0.43	-0.03	-0.16	-0.19	-0.09	0	0	-0.09	1	0.16	0.12	0.1
density	0.27	0.03	0.15	0.84	0.26	0.29	0.53	1	-0.09	0.07	-0.78	-0.31
total_sulfur_dioxide	0.09	0.09	0.12	0.4	0.2	0.62	1	0.53	0	0.13	-0.45	-0.17
free_sulfur_dioxide	-0.05	-0.1	0.09	0.3	0.1	1	0.62	0.29	0	0.06	-0.25	0.01
chlorides	0.02	0.07	0.11	0.09	1	0.1	0.2	0.26	-0.09	0.02	-0.36	-0.21
residual_sugar	0.09	0.06	0.09	1	0.09	0.3	0.4	0.84	-0.19	-0.03	-0.45	-0.1
citric_acid	0.29	-0.15	1	0.09	0.11	0.09	0.12	0.15	-0.16	0.06	-0.08	-0.01
volatile_acidity	-0.02	1	-0.15	0.06	0.07	-0.1	0.09	0.03	-0.03	-0.04	0.07	-0.19
fixed_acidity	1	-0.02	0.29	0.09	0.02	-0.05	0.09	0.27	-0.43	-0.02	-0.12	-0.11

Scatter Plot

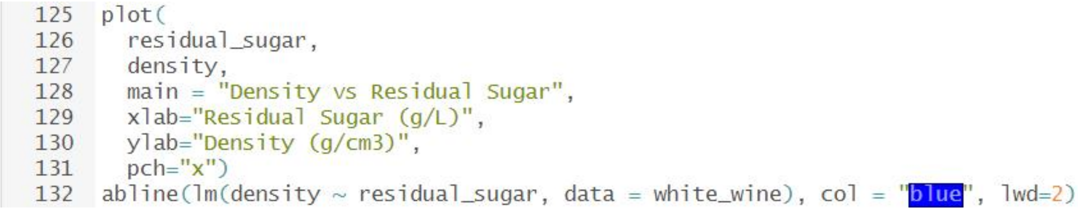
The scatter plot of Alcohol vs Density is plotted along with the line of best fit to show the strong negative correlation.



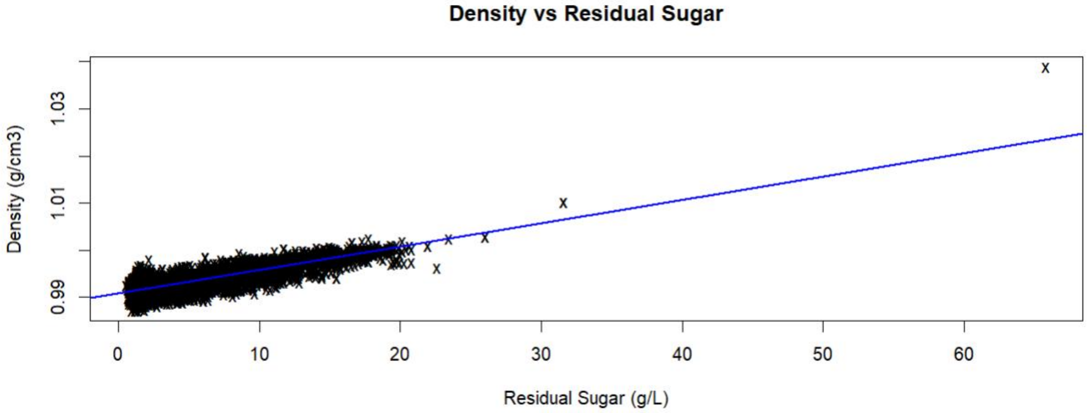
```
# scatter plot
plot(
  x=alcohol,
  y=density,
  main = "Density vs Alcohol",
  xlab="Quality",
  ylab="Density",
  pch=".")
abline(lm(density ~ alcohol, data = white_wine), col = "blue")
```



Residual Sugar and Density have strong positive correlation



```
plot(
  residual_sugar,
  density,
  main = "Density vs Residual Sugar",
  xlab="Residual Sugar (g/L)",
  ylab="Density (g/cm3)",
  pch="x")
abline(lm(density ~ residual_sugar, data = white_wine), col = "blue", lwd=2)
```



Objectives

1. Predict the quality of white wine based on alcohol percentage.
2. Find out the how Residual Sugar affect Density of wine.
3. Forecast the quality of white wine with alcohol, density, and chlorides.

Hypothesis

1. Alcohol vs Wine Quality

Null Hypothesis: Wine with high alcohol percentage of greater than or equals to 12% do not have higher mean quality than alcohol percentage lower than 12%.

Alternative Hypothesis: Wine with high alcohol percentage of greater than or equals to 12% have higher mean quality than alcohol percentage lower than 12%.

2. Residual Sugar and Density

Null Hypothesis: Residual sugar content is not significant to the density of wine.

Alternative Hypothesis: Residual sugar content is significant to density of wine.

3. Quality vs alcohol, density, and chlorides

Null Hypothesis: Alcohol, density, and chlorides are not sufficient predictors of white wine quality.

Alternative Hypothesis: Alcohol, density, and chlorides are sufficient predictors of white wine quality.

Method

Precondition:

- Set the confidence interval at 95%, and significance level at 5%
- Calculate summary statistics: mean, median, mode, standard deviation, variance, range, interquartile range.
- Display graphs related to the variables stated in objectives: Alcohol, Quality, Sulphates, Density, Chlorides

Hypothesis testing:

a) Two Sample z-test

- Used to test whether the quality of white wine is based on alcohol percentage
- Filter the data by the alcohol percentage into 2 subsets: -
    - low_pct: alcohol content less than 12%
    - High_pct: alcohol content greater than or equals to 12%
- Conduct a 2-sample z-test with calculated sample means and variance to get the p-value.

b) Correlation Analysis with Pearson correlation

- To find out the relationship between Residual Sugar and Density.
- Use `cor.test(sulphates, quality, method = "pearson")`
- Interpretation (Turney, 2022):
    - Strong correlation: ± 0.5 to ± 1
    - Moderate correlation: ± 0.3 to 0.5
    - No correlation: 0

c) Multiple linear regression model

- To determine the association between quality and alcohol, density, and chlorides.
- Use `lm(quality ~ alcohol + density + chlorides, data = white_wine)`, then summaries the result to gain the residuals, t value, coefficients, Multiple R-Squared, adjusted R-squared, and etc.
- Interpret:
    - Residuals: Measurement of vertical distance between a point and the regression line (Gohar, 2020).

  - Estimates: Effect estimated of each predictor variable on the white wine quality.
  - t value: Calculated by dividing the Estimate by the standard error, it represents the significance of coefficient of each predictor in influencing the response variable (Kumar, 2023). Higher t-value suggests that the coefficient is further away from 0 (Thieme, 2021), therefore has stronger predictive power, while lower value means the predictor is less significant to the response variable.
  - P-values, Pr(>|t|): Calculated from the t value earlier, it represents the significance of coefficients to the model. Lower p-value implies higher significance (Thieme, 2021), and vice versa.
  - Multiple R-Squared: Also known as coefficient of multiple determination, where the variance percentage in the response variable is explained by the predictors (Multiple Regression | Gunnison County, CO - Official Website, 2024).
  - Adjusted R-Squared: Shows variation percentage in response variable that can be explained by the predictors involved (Thieme, 2021). Hence, a higher number is preferred. The value is adjusted for the number of predictors in the model to give a value that is lower than Multiple R-squared. Therefore, it is not affected by predictors that do not improve the model.
  - F-statistic: The value is calculated by dividing Mean sum of squares regression (MSR) by Mean sum of squares error (MSE). Large value indicates the model is significant, and better than model without predictors when the value is higher than 1. High F-statistic comes with a small p-value (Kumar, 2023), and vice versa.

Result:

- If p-value is lesser than threshold, reject null-hypothesis and accept alternative hypothesis
- If p-value is greater than threshold, accept null hypothesis and reject alternative hypothesis.

Justification

Precondition:

Significance level, α, of 5% is accepted as threshold to differentiate significant to non-significant results for decades (Giovanni Di Leo & Sardanelli, 2020), while Confidence Interval (CI) is the complimentary of Significance Level, where CI = 1-α. Therefore, the CI is 95%.

Summary statistics gives a quick data summary, helpful in comparing projects (Summary Statistics, 2022). It also assists in understanding the data, especially the distribution and central tendency before conducting the hypothesis testing.

Graphs are the data visualization to show the trend, patterns, and distribution of the values. For instance, fixed acidity and quality shows a normal distribution curve. The relationship between 2 variables can also be displayed clearly in a graph. This gives a better understanding of the values calculated during hypothesis testing.

Hypothesis Testing:

a) Two Sample z-test

The hypothesis focused on 2 different samples of alcohol percentage: below and above 12%, where both samples have size larger than 30. Therefore, a z-test can be used as it is suitable for large sample sizes and known population variance (Yang, 2017).

b) Correlation Analysis with Pearson correlation

The conditions are met:

- Both predictor variable (Residual Sugar) and response variable (Density) are qualitative
- Variables are normally distributed.
- The data does not have outlier
- Have linear relationship (Turney, 2022).

c) Multiple linear regression model

There are multiple predictors involved, alcohol, density, and chlorides, to determine the quality of white wine. The relationship between quality and each predictor is linear, but not highly correlated. This assists in understanding the influence of predictors to the sole dependent variable and allows forecasts to be made from the predictors.

Result:

Three methods used above generate a p-value that contributes to the decision to reject or accept the null hypothesis.

Conclusion

The quality of white wine is determined by multiple factors, but not all of them carry the equal weight. Different proportions of features result in various quality rating, there is no one-size-fit-all-solution, due to the interaction between features. Thus, hypothesis testing would be carried out to investigate the relationship between 2 or more features.

1849 words

Reference

1. Cortez,Paulo, Cerdeira,A., Almeida,F., Matos,T., and Reis,J.. (2009). Wine Quality. UCI Machine Learning Repository. https://doi.org/10.24432/C56S3T
2. Good Pair Days. (2024). *Wine Subscriptions Paired To Your Tastes | Good Pair Days*. Good Pair Days. https://www.goodpairdays.com/guides/wine-101/article/history-of-wine/
3. Good Pair Days. (2024). Wine Subscriptions Paired To Your Tastes | Good Pair Days. Good Pair Days. https://www.goodpairdays.com/guides/wine-101/article/white-wine/?ref=gpd-guides.ghost.io
4. The nine primary styles of wine. (2020, September 16). Cult Wines. https://www.wineinvestment.com/learn/magazine/2020/09/the-nine-primary-styles-of-wine/
5. O'Donnell, D. (2017, December 6). Improving the Taste and Color of Wine with pH Control | Sensorex. Sensorex Liquid Analysis Technology. https://sensorex.com/ph-improve-taste-color-wine/#:~:text=Usually%2C%20a%20wine%20will%20fall,falling%20between%203.3%20and%203.6.

6. Weatherwax, J. (2024, June 5). Wine Alcohol Content Guide: How Much Alcohol Is In Wine? Binwise.com. https://home.binwise.com/blog/wine-alcohol-content#:~:text=The%20alcohol%20content%20in%20white,a%20lower%20rate%20as%20well.

7. o-a.com.au. (2018, March 19). Five Factors That Affect Wine Quality - Grapeworks - Tanium Machinery. Grapeworks - Tanium Machinery. https://grapeworks.com.au/news/winemaking/five-factors-that-affect-wine-quality/

8. Techniques for correcting wine acidity | Agrovin. (2021, May 10). Agrovin. https://agrovin.com/en/techniques-for-correcting-wine-acidity/#:~:text=Fixed%20acidity%20is%20the%20set,of%20tartaric%20acid%20per%20litre.

9. https://www.facebook.com/megmaker. (2023, February 25). Acid. Terroir Review. https://terroirreview.com/glossary/acid/#:~:text=Wine%20contains%20two%20types%20of,while%20volatile%20acids%20readily%20evaporate.

10. https://www.facebook.com/megmaker. (2023, February 23). Acidity In Wine – Terroir Review. Terroir Review. https://terroirreview.com/2023/02/23/acidity-in-wine/

11. Kaan, P. (2018, September 27). Residual Sugar. WINE DECODED. https://winedecoded.com.au/wine-words/residual-sugar/#:~:text=in%20new%20window)-,Residual%20Sugar%20refers%20to%20the%20amount%20of%20sugar%20left%20in,Typically%202g%2FL%20or%20less.

12. Purdue extension Commercial Winemaking Production Series Chlorine Use in the Winery. (n.d.). https://www.extension.purdue.edu/extmedia/FS/FS-50-W.pdf

13. Maltman, A. (2013). Minerality in wine: a geological perspective. Journal of Wine Research, 24(3), 169–181. https://doi.org/10.1080/09571264.2013.793176

14. https://www.facebook.com/midwestgrapeandwine. (2018, February 27). Total Sulfur Dioxide – Why it Matters, Too! | Midwest Grape and Wine Industry Institute. Midwest Grape and Wine Industry Institute. https://www.extension.iastate.edu/wine/total-sulfur-dioxide-why-it-matters-too/#:~:text=Simply%20put%2C%20Total%20Sulfur%20Dioxide,aldehydes%2C%20pigments%2C%20or%20sugars.

15. Wine Education Topic: Sulfur Dioxide in Wine. (2024). Aromadictionary.com. https://www.aromadictionary.com/articles/sulfurdioxide_article.html#:~:text=Sulfur%20dioxide%20plays%20two%20important,and%20protecting%20it%20against%20browning.
16. COMMENDIUM OF INTERNATIONAL METHODS OF ANALYSIS -OIV Density and Specific Gravity -Type IV method Density and Specific Gravity at 20 o C. (n.d.). https://www.oiv.int/public/medias/2468/oiv-ma-as2-01b.pdf
17. Iwona Budziak-Wieczorek, Vladimír Mašán, Klaudia Rząd, Bożena Gładyszewska, Karcz, D., Burg, P., Čížková, A., Mariusz Gagoś, & Arkadiusz Matwijczuk. (2023). Evaluation of the Quality of Selected White and Red Wines Produced from Moravia Region of Czech Republic Using Physicochemical Analysis, FTIR Infrared Spectroscopy and Chemometric Techniques. Molecules/Molecules Online/Molecules Annual, 28(17), 6326–6326. https://doi.org/10.3390/molecules28176326
18. Weatherwax, J. (2024, June 5). Wine Alcohol Content Guide: How Much Alcohol Is In Wine? Binwise.com. https://home.binwise.com/blog/wine-alcohol-content#:~:text=The%20alcohol%20content%20in%20white,keeps%20it%20light%20and%20refreshing.
19. OIV COLLECTIVE EXPERTISE DOCUMENT SO2 AND WINE: A REVIEW. (n.d.). https://www.oiv.int/public/medias/7840/oiv-collective-expertise-document-so2-and-wine-a-review.pdf
20. What Are Wine Sulfites & Which Wines Are Low Sulfite Wines? - Bright Cellars. (2022, January 27). Bright Cellars. https://www.brightcellars.com/blogs/learn/what-are-wine-sulfites#:~:text=That%20means%2C%20while%20sulfites%20and,released%20from%20mines%20or%20mills.
21. GeeksforGeeks. (2021, October 13). Rotating and spacing axis labels in ggplot2 in R. GeeksforGeeks; GeeksforGeeks. https://www.geeksforgeeks.org/rotating-and-spacing-axis-labels-in-ggplot2-in-r/
22. Muhammad Sopyan Yahya. (2023, July 9). Analyzing Red Wine Quality - Muhammad Sopyan Yahya - Medium. Medium; Medium. https://medium.com/@spynyahya/analyzing-red-wine-quality-69aadb08a303
23. Summary statistics. (2022). Better Evaluation. https://www.betterevaluation.org/methods-approaches/methods/summary-

statistics#:~:text=Summary%20statistics%20provide%20a%20quick,tendency%20and%20measures%20of%20dispersion.

24. Giovanni Di Leo, & Sardanelli, F. (2020). Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. European Radiology Experimental, 4(1). https://doi.org/10.1186/s41747-020-0145-y
25. Turney, S. (2022, May 13). Pearson Correlation Coefficient (r) | Guide & Examples. Scribbr. https://www.scribbr.com/statistics/pearson-correlation-coefficient/#:~:text=The%20Pearson%20correlation%20coefficient%20(r,the%20relationship%20between%20two%20variables.&text=When%20one%20variable%20changes%2C%20the,changes%20in%20the%20same%20direction.
26. Gohar, U. (2020, March 5). How to use Residual Plots for regression model validation? Medium; Towards Data Science. https://towardsdatascience.com/how-to-use-residual-plots-for-regression-model-validation-c3c70e8ab378
27. Multiple Regression | Gunnison County, CO - Official Website. (2024). Gunnisoncounty.org. https://www.gunnisoncounty.org/673/Multiple-Regression#:~:text=Multiple%20R%202%20is%20the,the%20coefficient%20of%20multiple%20determination.
28. Yang, X.-S. (2017). Regression and Curve Fitting. Elsevier EBooks, 215–228. https://doi.org/10.1016/b978-0-12-809730-4.00025-2
29. Kumar, A. (2023, December 11). F-test & F-statistics in Linear Regression: Formula, Examples. Analytics Yogi. https://vitalflux.com/interpreting-f-statistics-in-linear-regression-formula-examples/#:~:text=F%2Dstatistics%20is%20used%20to,known%20as%20the%20null%20model).
30. Thieme, C. (2021, March 12). Understanding Linear Regression Output in R - Towards Data Science. Medium; Towards Data Science. https://towardsdatascience.com/understanding-linear-regression-output-in-r-7a9cbda948b3