# Introduction

## 1:1 Problem statement and research motivation

The ever-growing wine industry has seen a decline in both consumption and production recently. The growth of the industry relies on the satisfaction of consumers. Thus, the relationship between alcohol content and wine quality is a critical topic in the wine industry, as prior research suggests that alcohol content can significantly impact the sensory characteristics of a wine (Teng-Zhen Ma , et al., 2021). Ensuring the wine industry relishes its glory days again with the help of our research provided us the motivation to figure out whether wine with higher alcohol content has a significantly higher quality rating than then the wines with lower alcohol content.

## 1:2 The data set

Our dataset “wine quality-white” was derived from Kaggle containing a total of 4898 samples of wine along with 12 features next to them. Although our research primarily prioritises on the feature of Alcohol (percentage of Alcohol) and wine quality (rating from 1-10), other features such as density, PH, residual sugar, citric acid etc are also key components of a wine sample which are present in the dataset that can lead many more compelling and insightful research. The dataset is ideal for studying the impact of alcohol content on wine quality.

## 1:3 Research question

To ascertain whether the proportion of alcohol quantity connects to alcohol quality we needed to craft an excellent research question. Therefore, it lead to the research question of –

Is there a difference in the mean quality rating between wines with low alcohol content and wines with high alcohol content?

We plan to answer our research questions through statistical testing revolving around our hypothesis. We want use visualizations such as histogram to test the normality of the data, boxplots to illustrate differences and statistical test such as T-test, Wilcoxon test to evaluate significance. We intend to use the language R through R Studio, utilizing the dataset to explore the effects of alcohol content on wine quality.

## 1:4 Null hypothesis and alternative hypothesis (H0/H1)

Over the course of this assessment, we want to investigate the effect of alcohol content on the quality ratings of white wine. To establish this, we approach it scientifically, establishing two competing hypotheses-

**Null Hypothesis**: There is no difference in the mean quality ratings between the wines with low alcohol content and wines with high alcohol content.

**Alternative Hypothesis**: There is a difference in the mean quality ratings between wines with low alcohol content and wines with high alcohol content.

Though our statistical testing we will be able to evaluate the hypothesis under defined significance level leading us to either reject or accept the null hypothesis.

# Bibliography

T.-Z. M.et al., 2021. Techniques for Dealcoholization of Wines: Their Impact on Wine Phenolic Composition, Volatile Composition, and Sensory Characteristics.