Louvain School of Management

Exploring the influence of environmental variables on the quality ratings of four Californian wine varieties

Author: Madeleine Hayez
Supervisor: Professor L. Iania
Academic year 2022-2023
Dissertation for the master of Business engineering
Master subject and focus: Financial engineering
Daytime schedule

I would like to thank my research advisor, Leonardo Iania, for his guidance and support. His expertise, commitment, and his dedicated time have been incredibly valuable throughout my analyses and the writing of this paper.

Furthermore, I would like to thank my friends and family for the continuous support they gave me. They encouraged me during all my studies and particularly during the writing of this paper. They gave me good advice when needed, which helped me push further my analyses and take a critical look at my work to achieve this final version, which I'm proud of.

This paper examines the relationship between weather variables and the quality ratings of four wine varieties: The Californian North Coast Cabernet Sauvignon, Chardonnay, Zinfandel and Pinot Noir. The aim of this research is to identify weather variables that can explain a great proportion of the variance in the quality ratings, despite a limited impact on wines from regions with stable climates, such as California. To achieve this objective, rating qualities (Robert Parker Wine Advocate, 2023b) and weather variables (NCEI National Centers for Environmental Information, 2023d) were downloaded. The quality ratings were transformed to account for the fact that they are bounded between 50 and 100, and the weather variables were averaged over two counties (Sonoma and Napa) for greater precision. The weather variables used were the average temperature, the average temperature squared, the maximum and minimum temperature, the difference between the maximum and minimum temperature, the precipitation, and the Southern Oscillation Index (SOI). In this paper, a yearly and seasonal analysis is done to test whether by splitting the variables by season, the data fit the regression models better. The main finding of this research is that these weather variables do play a role in explaining the variations in wine quality. However, their role is quite limited. The adjusted R-squared of the models are highly influenced by a positive trend, that accounts for the improvement of qualities, not related to the climate variables. The difference between the maximum and minimum temperature and the SOI are the variables that are the most often statistically significant, apart from the trend, indicating they probably have a significant influence on the quality of the wine. Climate change will likely impact wine quality, and therefore vineyard production and profitability. It is thus crucial to assess the influence of weather on the quality of the different wine varieties, allowing winemakers to adapt accordingly.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm

Accronyms

CNC = Californian North Coast

RPWA = Robert Parker Wine Advocate

NOAA NCEI = National Oceanic and Atmospheric Administration: National Centers for Environmental Information

ENSO = El Ninõ-Southern Oscillation

SOI = Southern Oscillation Index

CS = Cabernet Sauvignon

CH = Chardonnay

Zi = Zinfandel

PN = Pinot Noir

yAT = Yearly Average Temperature

yPrec= Yearly Precipitation

yMinT = yearly Minimum Temperature

yMaxT= yearly Maximum Temperature

yDiff = Difference of yearly maximum and minimum temperature (= yMaxT - yMinT)

W = Winter season (from November until February)

EG = Early-Growing season (Mars and April)

G = Growing season (from May until August)

H = Harvest season (September and October)

AT_W = Average temperature during the winter season

$Prec_W$ = Precipitation during the winter season

$MinT_W$ = Minimum temperature of the winter season

$MaxT_W$ = Maximum temperature of the winter season

$Diff_W$ = Difference between the maximum and minimum temperature during the winter season ($= MaxT_W - MinT_W$)

AT_{EG} = Average temperature during the early-growing season

$Prec_{EG}$ = Precipitation during the early-growing season

$MinT_{EG}$ = Minimum temperature of the early-growing season

$MaxT_{EG}$ = Maximum temperature of the early-growing season

$Diff_{EG}$ = Difference between the maximum and minimum temperature during the early-growing season ($= MaxT_{EG} - MinT_{EG}$)

AT_G = Average temperature during the growing season

$Prec_G$ = Precipitation during the growing season

$MinT_G$ = Minimum temperature of the growing season

$MaxT_G$ = Maximum temperature of the growing season

$Diff_G$ = Difference between the maximum and minimum temperature during the growing season ($= MaxT_G - MinT_G$)

AT_H = Average temperature during the winter season

$Prec_H$ = Precipitation during the winter season

$MinT_H$ = Minimum temperature of the winter season

$MaxT_H$ = Maximum temperature of the winter season

$Diff_H$ = Difference between the maximum and minimum temperature during the harvest season ($= MaxT_H - MinT_H$)

CV = Coefficient of variation

SD = Standard deviation

VIF= Variance Inflation Factor

Int. = Intercept

JB test = Jarque-Bera test

BP test = Breusch-Pagan test

Adj. R^2 = Adjusted R-squared

Contents

1 Introduction

The quality of a wine depends on a various number of factors, including wine-making techniques, the soil composition, grapevine characteristics, the region and the climatic conditions of the year (Ollat et al., 2016). Among these factors, the climate has been identified as particularly affecting the quality of the wine (Ollat et al., 2016; Ramirez, 2008; van Leeuwen et al., 2004). The growth cycle of the grapes is influenced by factors such as rainfall, temperature and other weather variables, making it interesting to analyze the specific influence of the weather on the quality of the wine.

In parallel with that, the effects of climate change are becoming more evident worldwide. It's an important and urgent issue for the wine market because climate change can directly influence the characteristics of a wine, its quality, the overall quantities produced and thus also the prices of wines, thereby influencing the profitability of the vineyards (Ashenfelter, 2010; Ashenfelter & Storchmann, 2016; Oczkowski, 2016; Ollat et al., 2016).

Exploring how the weather can impact the quality of wine, can give valuable insights into how weather affects the taste and value of wines. This knowledge is essential for wine producers, consumers, and the industry as a whole, as it helps us understand the challenges and opportunities that come with changing climate conditions (Oczkowski, 2016; Ollat et al., 2016). It can also give an insight into the suited adaptation strategy, such as technological innovations or relocating the vineyards into more suited regions. Another aspect that makes wine an interesting product to analyze is the fact that wine can be easily stored for long periods of time, and in some cases, it appreciates over time. This makes wine an attractive investment option and justifies research dedicated to the influence of climate on wine quality (Ashenfelter & Storchmann, 2016).

This research analyses the relationship between several climate variables and specific quality ratings that were given to some wines. Specifically, it focuses on four different grape varieties produced in the Californian North Coast (CNC): The *CNC Cabernet Sauvignon, CNC Chardonnay, CNC Zinfandel*, and *CNC Pinot Noir*.

To answer the research question of *"How do the weather variables, such as rainfall, average temperature, maximum temperature, minimum temperature, the difference between maximum temperature and minimum temperature, and the Southern Oscillation Index, influence the quality ratings of the four North Coast Californian wine varieties?"*, a quantitative methodological approach is used, necessitating the collection of data. The quality of the wines is approximated by quality ratings. These quality ratings were collected from the website "Robert Parker Wine Advocate". This is a reputable website, that is known for its wine expertise. The quality ratings are computed as an average of 12,989, 6,268, 6,124, and 3,599 observations between 1978 and 2018 for the Cabernet Sauvignon, the Chardonnay, the Zinfandel and the Pinot Noir respectively. The rating quality underwent two transformations to account for the fact that they are bounded between 50 and 100: a linear transformation to fit a logistic regression and a transformation to estimate the dependent variables as multiple linear functions. Additionally, climate data was collected from the recognized website of the "NOAA National Centers for Environmental Information (NCEI). The average precipitation, average temperature, maximum temperature, and minimum temperature were collected on a monthly basis, going from November 1977 until October 2018. These weather variables were collected for the Napa and Sonoma county, which is where the great majority of the wines included in the quality ratings, comes from. A weighted average of the weather variables from both regions was computed, taking into account the numbers of wines originating from the Sonoma County and the Napa County, for each wine variety. This permits to have a precise database for each wine variety. The Southern Oscillation Index was also collected on a monthly basis, going from June 1977 until May 2018, to be in line with the life cycle of the SOI. Lastly, a trend was included in the models to account for improvements in quality ratings, not related to the weather. Both websites were meticulously chosen, based on the quality of their available data, which is what ensures the reliability and validity of the analysis conducted in this research.

Furthermore, as models described in previous literature performed badly, others models were tested in order to assess the ability of some models to better describe variances in the quality ratings of the four wine varieties. In order to achieve this, a yearly and seasonal analysis was performed and the results are given in this paper.

First, the literature review will be described, which aims at providing a critical overview of the existing literature that serves as the foundation for this research. Following that, the methodology employed and the collected data are described. This includes the research objective, the data collection, the tested hypotheses and the statistical models that are tested in this research. In the next section, the detailed statistical analysis is presented. Lastly, the limits and further improvement points are given, followed by the conclusion of this research.

2 Literature review

2.1 Theoretical Framework and Existing Literature

Many studies have analyzed the influence of factors such as climate, soil, terroir, region of origin, appellation and grape variety on wine prices (Outreville & Fur, 2020). As demonstrated by numerous papers, weather conditions have been found to affect the grape quality and, consequently, wine quality (Ashenfelter, 2010; Charlin & Cifuentes, 2023; Oczkowski, 2016; Ollat et al., 2016; Ramirez, 2008). This paragraph aims at providing more insight into the existing research on this subject.

There is a substantial amount of literature available that focuses on the hedonic price function of wines. This approach assumes a direct association between the price of a wine and some characteristics of this wine. The theoretical framework proposed by Rosen (1974) is widely referenced in the literature on the hedonic price functions of wine. In most articles, the wine prices are estimated as a log-linear function. Several articles analyze the impact of the weather, and often also of other variables such as the quality or the age of the wine, on the wine prices (Ashenfelter, 2010; Haeger & Storchmann, 2006; Niklas & Rinke, 2020; Oczkowski, 2016; Ramirez, 2008).

This research will not focus on the hedonic price function, but it will analyze the influence of the climate variables on wine quality. The quality is approximated by rating scores given by the website Robert Parker Wine Advocate. According to Haeger & Storchmann (2006), the wine critics of *Robert M. Parker, Jr.* and the *Wine Spectator magazine* are extensively referenced and the most famous. More information about this website is given in paragraph 2.3. This research will thus use a wine rating as a proxy for the quality of wine, as done in several other papers (Charlin & Cifuentes, 2022, 2023; Jones et al., 2005; Niklas & Rinke, 2020; Oczkowski, 2016; Ramirez, 2008).

It is important to note that, given the extensive body of literature on this subject, it is impossible to provide a comprehensive list of all relevant studies in this literature review. However, the following paragraphs will highlight three notable articles that significantly influenced this research. The choice of these articles is based on the fact that they are widely cited in other papers on this subject, especially the ones of Oczkowski (2016) and Ramirez (2008). Furthermore, most of the other papers analyze the influence of the weather on wine prices, while the three articles described below analyze the impact of weather on wine quality, similarly to what is done in this research.

2.1.1 Wine quality and weather

Among the existing research, several different techniques have been deployed. Some articles use a cross-sectional analysis of wines to examine the impact of weather on wine in one specific year. Oczkowski (2016) for example did a cross-sectional analysis of weather factors on a wine-quality measure and wine prices of several Australian wine varieties, produced in 2014. In this article, the following model was used for each wine variety separately, as it is considered to have the most statistically significant variables and the biggest explanatory power. It is used employing robust standard errors :

$$(Rating)_i = \alpha_0 + \alpha_1(Rain)_i + \alpha_2(Diff)_i + \alpha_3(Temp)_i + \alpha_4(Temp)_i^2 + \epsilon_i \tag{1}$$

The selection of weather variables, used by Oczkowski (2016), is based on those commonly used in other articles on the same topic, and the rating variable is obtained from Halliday (2014). ϵ represents the error term and *Rain* represents the average monthly rain, in inches, during the Australian harvest period (January to March). *Diff* represents the monthly average variation between the highest and lowest temperatures (in degrees Celsius), throughout the Australian growing season (October to March). *Temp*, on the other hand, represents the average temperature during the growing season, also based on monthly averages. Lastly, a quadratic form is used on the variable *Temp*. It is expected that the coefficient α_3 will have a positive sign, while α_4 will have a negative sign (Oczkowski, 2016). Indeed, high temperatures during the growing season tend to be optimal for the vineyards, up to a certain threshold. Excessively high temperatures can significantly harm the quality of grapes, which would negatively impact the wine quality (Ashenfelter & Storchmann, 2016; Charlin & Cifuentes, 2023; Haeger & Storchmann, 2006; Jones et al., 2005; Niklas & Rinke, 2020; Oczkowski, 2016; Ollat et al., 2016; Ramirez, 2008). By adding the quadratic form of *Temp*, it is possible to test whether this theory can be statistically validated.

The paper of Oczkowski (2016) concludes that variation in wine quality can be explained by weather differences between the regions, for each wine variety separately. Furthermore, the variables *Temp* and $Temp^2$ were the most statistically significant, with α_3 positive and α_4 negative for all varieties except one. This validates that too high temperatures during the growing season tend to have a negative impact on the wine quality. The variable *Diff* was mostly not statistically significant while *Rain* was statistically significant for 5 varieties out of the 8, with α_1 negative for all varieties except one. This indicates that rain during the harvest period negatively influences wine quality. Nevertheless, the explanatory power of the models is quite low, ranging from 3.5% to 17.5%. This suggests that weather plays a relatively low role in explaining variations in the quality of wine in Australia (Oczkowski, 2016) .

Another article that is worth mentioning is the one written by Ramirez (2008). Instead of focusing on one specific year, such as Oczkowski (2016), a cross-sectional analysis is used, to test the influence of the weather on one specific grape variety of a single region, over various years. In this paper, an analysis of the impact of some weather factors on all Cabernet Sauvignon wines, rated by the Wine Spectator, and located in the Napa Valley region of California, from 1970 to 2004, is done. For this analysis, a linear-, quadratic-, and log-linear specification is described, in function of the following climate variables:

$$R_{i,t} = f(trend, temperature, precipitation) \tag{2}$$

The models are used with robust standard errors. The climate variables used are quite similar to model 1 (Oczkowski, 2016), with some differences. Similarly to model 1, Ramirez (2008) also used the precipitation and average temperature variables. However, he also added a trend in the model. Furthermore, he split the temperature and precipitation into different periods to permit more flexibility and a nonlinear effect throughout the seasons. For the average temperature, the temperature of the growing season was employed, and divided into 3 sub-periods, namely the average temperature of April-May, June-July, and August-September. Concerning the precipitation, the variable was split by taking the average precipitation of the winter and growing season, divided into 4 different sub-periods, being January-February, April-May, June-July, and August-September (Ramirez, 2008).

The conclusion was made that both precipitation and temperature variations have an impact on wine quality and wine prices. However, the impact on wine quality was much smaller. For the linear model, the weather variables contribute only approximately for 3% of the model variation. This was computed by removing the influence of the constant on the model variation. The model variation, including the influence of the constant, was in fact 29.12%. The adjusted R-squared, which represents the model variation, of the quadratic-variation model and log-linear model were respectively 30.68% and 28.19% (Ramirez, 2008). This aligns with the results of other studies that have indicated that in regions characterized by consistent weather patterns, which is the case for California, the influence of weather on wine quality is less significant than in regions with highly fluctuating weather (Charlin & Cifuentes, 2023; Haeger & Storchmann, 2006; Jones et al., 2005; Oczkowski, 2016). Despite that, the quadratic terms added in the quadratic model were mostly statistically significant. Despite the presence of multicollinearity, the fact they are statistically significant could suggest a non-linear impact of the weather on the wine quality and prices, as mentioned above. Lastly, the results obtained from the log-linear model were highly comparable to those of the linear model, with as main difference that in the log-linear model, the variable *Trend* was statistically significant. This could indicate that it is an important variable to add in future models (Ramirez, 2008). The inclusion of a trend variable has, however, been criticized by Charlin & Cifuentes (2023), suggesting that it may lead to issues of multicollinearity and provide little improvement to the explained variation of the model.

A third article important to mention is therefore the article written by Charlin & Cifuentes (2023). By using the Wine Spectator ratings as a proxy for the quality of wines, they analyze the impact of weather variables on the wine quality of Malbec wines in the Mendoza region, from 1995 to 2020. As did Ramirez (2008) and Oczkowski (2016), Charlin & Cifuentes (2023) used the average temperature (*Temp*) and the precipitation (*Rain*) in their analysis, in function of different sub-periods (being here the winter- (W), harvest- (H), first part of growing- (*G1*) and second part of growing season (*G2*). However, their analysis also includes additional variables, such as the average of the minimum and maximum daily temperature (T_{max} & T_{min}), the most extreme values of the season ($Max(T_{max})$ & $Min(T_{min})$), the humidity, the cloudiness (*Cloud*) and the number of days with a temperature above 32°C. Lastly, they also use the wineries coefficient linked with a dummy variable (0,1) to associate each observation with its corresponding winery or group of wineries. They describe three final models in their article.

The first model they examined, represented by model 3 in this research, is very similar to the ones cited above, as it only includes the rainfall and average temperature during the two growing seasons. It can be described as follows:

$$(Rating)_i = \alpha_0 + \alpha_1(Rain_{G1})_i + \alpha_2(Rain_{G2})_i + \alpha_3(Temp_{G1})_i + \alpha_4(Temp_{G2})_i + \epsilon_i \quad (3)$$

The second model, represented by model 4, includes more variables. It was chosen by keeping the variables with the greatest explanatory power and considering multicollinearity problems. It can be described in the following way:

$$\begin{aligned}(Rating)_i = \alpha_0 &+ \alpha_1((T_{max})_W)_i + \alpha_2(Cloud_W)_i + \alpha_3((Max(T_{max}))_{G1})_i + \alpha_4(Cloud_{G1})_i \\ &+ \alpha_5(Temp_{G2})_i + \alpha_6(Temp^2_{G2})_i + \alpha_7((Max(T_{max}))_{G2})_i + \alpha_8(Cloud_{G2})_i \\ &+ \alpha_9(Temp^2_H)_i + \alpha_{10}((Min(T_{min}))_H)_i + \alpha_{11}(Cloud_H)_i + \epsilon_i\end{aligned} \quad (4)$$

The reason the variables describing the rainfall and the humidity were not used is because of their high correlation with the variable *Cloud*, which had a higher explanatory power. Moreover, the reason the variable $Temp^2$ was used is the same as explained for the model described by Oczkowski (2016). Lastly, the absence of the variable describing the number of days with a temperature above 32°C is due to its high correlation with $Max(T_{max})$.

The third and last model described, uses the same variable as model 4, with as only additional variable an index to identify the different wineries.

The models described in the paper of Charlin & Cifuentes (2023) showed various results. A first important aspect to highlight is the fact that all the variables of each model are statistically significant at a level of 1%, except $Rain_{G2}$ of their first model which is not statistically significant, and $Cloud_H$ that is at a level of 10%. Secondly, the adjusted R^2 of the third model was a lot higher than their first and second model, being respectively 24.9%, 3.2%, and 7.3%. This indicates that the wine quality of the Mendoza's Malbec is more influenced by the winery effects than by weather effects. This can be attributed to the fact that this winery effect captures several aspects that have an influence on the wine quality, such as the soil composition and the altitude at which the vines are planted, the expertise and wine-making processes of each winery or the technological tools the winery uses (Charlin & Cifuentes, 2023). Lastly, the adjusted R^2 of the first two models indicate that the weather variables of these models explain respectively only 3.2% and 7.3% of the model variation. The fact they are small can be explained by the fact that the weather in the Mendoza region is quite stable. As seen in the works of Oczkowski (2016) and Ramirez (2008), the wine quality tends to be only moderately explained by the weather variables in regions with stable weather conditions. However, the fact that their second model (represented by model 4), has a slightly higher explanatory power compared to their first model (represented by model 3), could indicate the importance of including variables of the four seasons and including variables that capture extreme conditions (Charlin & Cifuentes, 2023). This aligns with what is described in another of their articles, written in 2022. In this article, the authors analyze the impact of the weather on the wine quality of the Chilean Cabernet Sauvignon, which is located in a region with even more stable weather conditions than in California or Australia, according to their calculations. By incorporating 27 weather variables and one constant in their model, they achieve an adjusted R-squared of 31% (Charlin & Cifuentes, 2022). Although there is probably a problem of multicollinearity in their model which makes the coefficients and adjusted R-squared of this model difficult to interpret, it could indicate that adding more weather variables can improve the explanatory power of a model. This multicollinearity problem is the reason this article is not explained more profoundly in this paragraph.

In conclusion, these three articles highlight that the weather plays a small role in explaining the variation in wine quality of regions with consistent climate conditions. Additionally, splitting the weather factors into different periods enables a more detailed and accurate analysis of this subject. This is crucial as the grapes require varying climate conditions during their whole life cycle. Lastly, although including a quadratic term for the average temperature during the growing season may bring multicollinearity concerns, this variable tends to be statistically significant, which confirms that temperatures positively affect wine quality, but at a decreasing rate.

The next paragraph aims at giving more information about the theoretical optimal climate conditions for the various grape varieties that will be analyzed in this research.

2.2 Wine varieties and their optimal climate conditions

Several analyses look at the impact of weather variables on wine quality or wine prices, by splitting the weather variables according to several seasons. Most often, the variables are split into the winter season, the harvest season, and the growing season and the variables of one or more seasons are used in the models (Ashenfelter, 2010; Ashenfelter & Storchmann, 2016; Charlin & Cifuentes, 2022, 2023; Haeger & Storchmann, 2006; Jones et al., 2005; Niklas & Rinke, 2020; Oczkowski, 2016; Ramirez, 2008).

The reason for this is that the grapes don't need the same climatic conditions during the whole year. There are periods where they need more water or hotter temperatures than others, to produce optimal grapes. Splitting the weather variables into the different seasons can therefore lead to more precise and optimal analyses. For example, freezing temperatures during the growing season can be destructive for the grapes (Ashenfelter & Storchmann, 2016), while it can reduce the risk of vine damage when they occur during the winter season (Jones et al., 2005). Furthermore, rainfall during the harvest period tends to have a negative effect on wine quality as it increases the chances of diseases (Ashenfelter & Storchmann, 2016; Charlin & Cifuentes, 2023; Haeger & Storchmann, 2006; Oczkowski, 2016; Ramirez, 2008), while rainfall during the winter season tends to improve the wine quality (Ashenfelter & Storchmann, 2016; Ollat et al., 2016; Ramirez, 2008). Moreover, rather cool and wet winters and relatively warm and dry summers are considered to have a positive impact on the wine (Ashenfelter, 2010; Ramirez, 2008). Lastly, extremely hot temperatures during the growing season tend to negatively impact the wine quality (Ashenfelter & Storchmann, 2016; Charlin & Cifuentes, 2022, 2023; Haeger & Storchmann, 2006; Jones et al., 2005; Niklas & Rinke, 2020; Oczkowski, 2016; Ollat et al., 2016; Ramirez, 2008).

Another important aspect to note is the fact that not all grape varieties need the same type of climatic conditions, and it is therefore important to analyze them separately (Oczkowski, 2016). Indeed, the Pinot Noir for example is known to be more resistant to extreme temperatures than other varieties. Cooler temperatures are often appropriate for this variety (Haeger & Storchmann, 2006). The Zinfandel needs hot summer and rainy winter season, which makes California an optimal place to grow Zinfandel. However, all the grapes of the Zinfandel don't ripen at the same time, which can influence its quality (Jancis Robinson, 2023). This underscores that each grape variety should be analyzed separately.

2.3 Robert Parker or "The Wine Advocate"

2.3.1 About Robert Parker

This research will analyze the quality ratings of the wines given by the website "Robert Parker Wine Advocate (RPWA)". Originally, a man named Robert Parker created a wine guide, titled "The Baltimore-Washington Wine Advocate", which became later "The Wine Advocate". With this magazine, Robert Parker had as goal the creation of a literature that would give unbiased information about wine, without being linked to financial terms. By doing so, he enabled to make the information about wine accessible to every interested wine consumer. Over the years, Parker managed to prove his expertise and created a strong positive reputation for his magazine "The wine advocates" (Robert Parker Wine Advocate, 2023a).

At first, "The Wine Advocate" was available only on paper, but in 2002, the website "RobertParker.com" was created. The website "Robert Parker Wine Advocate" counts nowadays subscribers in more than 40 countries and in every state of the United States. In 2019, RPWA even has been acquired by the Michelin Group. Furthermore, RPWA is now considered by every knowledgeable observer as the literature that has the greatest impact on wine buying habits and trends throughout most wine markets in the world, such as in the United States, France, the United Kingdom, Switzerland, Japan, Singapore, Russia, Mexico, Brazil, and China. RPWA publishes essentially reviews of newly released wine, but it also creates vintages retrospectives as well as occasional profiles of individual wine producers and detailed examinations of specific wine labels. Each month, "Robert Parker" published between 2,000 and 6,000 wine reviews, with the goal of reviewing most of the major wine regions on an annual basis. Some smaller regions are reviewed every two to three years (Robert Parker Wine Advocate, 2023a).

2.3.2 The Wine Advocate's wine rating

"The Wine Advocates" gives almost every wine a rating. This rating represents the relative quality of the wine, with respect to its style, region, and grape variety, at the time it was tasted. The rating ranges go from 50 to 100 points and tries to be the most precise possible. "Robert Parker Wine Advocates" consider their ratings to be strict and even mention they rather underestimate the quality of wine than overestimate it. In addition to the rating, the site also assigns a letter for each wine, which describes the maturity of the wine (Robert Parker Wine Advocate, 2023a). The precise meaning of the ratings is described in paragraph 3.2.1.

"Robert Parker Wine Advocate" publishes, what they call, a Vintage Chart. This chart shows the rating they give for several grape varieties, in function of the year the wine was produced, in several countries and regions. This research will focus on the rating of four North Coast Californian grape varieties of this Vintage Chart. RPWA is therefore primordial for this research (Robert Parker Wine Advocate, 2023b).

2.4 NOAA National Centers for Environmental Information

The weather variables described in this research all come from the website of the National Centers for Environmental Information (NCEI). NCEI is the *"Nation's leading authority for environmental data, and manages one of the largest archives of atmospheric, coastal, geophysical, and oceanic research in the world"* (NOAA National Centers for Environmental Information, 2023a). Among other things, they use objective methods to collect climate information. They do this by looking at key climate indicators which permits them to provide data about temperature and precipitation, snow and ice, drought and wildfire, storms and wind, and various weather patterns (NOAA National Centers for Environmental Information, 2023a).

In this research, the influence of the climate indicators on the wine ratings, given by Robert Parker, will be analyzed. All the weather variables mentioned in this research come from the NOAA NCEI as it is a reliable and notable website.

3 Methodology and Data

3.1 Research objective

As mentioned above, the main objective of this research is to analyze the impact of some climate factors on the quality ratings of four different grapes varieties of wine, produced in the Californian North Coast. The goal is to see if the wine of each grape variety reacts differently to the weather and how the weather impacts their quality. This research gives an answer to the research question: *"How do the weather variables, such as rainfall, average temperature, maximum temperature, minimum temperature, the difference between maximum temperature and minimum temperature, and the Southern Oscillation Index, influence the quality ratings of the four North Coast Californian wine varieties? "*.

Because the models described in the other researches performed badly, with low adjusted R-squared, alternative models were tested in this research to try to find a model for each wine variety that is able to explain the variation in the quality ratings better.

To achieve this objective, a quantitative methodology is employed. This means that a lot of data was collected to be able to analyze patterns, test hypotheses and create statistical models. (Grad Coach, 2021). To achieve reliable statistical analyses, it is primordial to use data of good quality. This encouraged the choices of the websites from which the data was collected, as they both provide accurate and high-quality data. By selecting notable and qualitative sources, potential inaccuracies in the collected data are minimized, which ensures better statistical results.

3.2 Data collection

Data about the quality ratings and the weather was collected. The weather variables were averaged into yearly variables and also split into four different seasons; the winter season, the early-growing season, the growing season and the harvest season. This was done in accordance with previous literature on this subject. The winter season considered in this research goes from the month of November until February. This is the season when wines are dormant. The early-growing season goes from March until April, which is the period when the wines are the most vulnerable. The growing season goes from May until August. Lastly, the harvest season goes from September until October. The yearly growing cycle of the grapes starts thus in November of the previous year and ends in October of the following year, which is why the data starts in November 1977 until October 2018 (Calwineries, 2023; Sonoma Valley Visitors Bureau, 2022).

3.2.1 Quality Ratings of Robert Parker Wine Advocate

This section aims to provide a better understanding of the quality ratings given by RPWA. The quality ratings, obtained from the website "Robert Parker Wine Advocate", are the ones of four different grapes varieties, being *CNC Cabernet Sauvignon, CNC Chardonnay, CNC Zinfandel*, and *CNC Pinot Noir*. The data was collected for the period going from 1978 to 2018 from Robert Parker Wine Advocate (2023b). For every of these four grape varieties, a rating is given every year. These yearly ratings are thus an average of all the quality ratings given to the corresponding wine variety, produced in the North Coast of California, in a specific year. The numbers of wines taken into account for each variety are not the same. The ratings given between 1978 and 2018 are based on 12,989 observations for the Cabernet Sauvignon, 6,268 for the Chardonnay, 6,124 for the Pinot Noir and 3,599 for the Zinfandel (Robert Parker Wine Advocate, 2023b).

Due to the unavailability of automated data download options, the collection of the data was done manually. This was done by recording the ratings in an Excel file so that it could be used later for further analysis. Because the quality ratings are divided into two parts, namely the *dating ranges* and the *maturity*, they were split in the Excel file in order to form two databases that can be analyzed independently. The rating ranges are points, ranging from 50 to 100, and they are described as follows:

- Wines with a rating between 96 and 100 are considered extraordinary, impressively complex and profound.
- Ratings between 90 and 95 are given to outstanding wines, that are really complex and of incredible quality.
- Wines with a rating between 80 and 89 are above average to very good, with "*varying levels of finesse, spice and character*" (Robert Parker Wine Advocate, 2023a)
- Wines that are considered average receive a rating between 70 and 79.

- Wines described as below average, that contain some perceptible flaws, receive a rating ranging from 60 to 69.
- Wines with a rating between 50 and 59 are considered unacceptable.

Regarding the maturity, RPWA utilizes seven different maturity ratings, indicated by letters, to provide additional insights:

- The letter "C" indicates the wine may be too old.
- An "E" indicates an early maturing and accessible wine.
- A wine whose vintage is not declared, is indicated by the letters "NV".
- An irregular wine is indicated by the letter "I".
- The letters "NT" are used when the wine has not been sufficiently tasted to assign a rating.
- The letter "R" indicates that the wine is ready to drink;
- The letter "T" signifies that the wine is still tannic, youthful, or slow to mature

(Robert Parker Wine Advocate, 2023a)

The collected data did not contain any observations with a maturity rating "NV". Moreover, the ratings after 2018 were assigned by the letters "NT". These ratings are considered unavailable data and were thus removed from the database. This is why this research will only analyze the Californian North Coast grape varieties, produced between 1978 and 2018.

An important distinction to make is that the rating ranges and the maturity assessments represent different types of variables. The dating ranges are discrete quantitative variables as they consist of non-continuous, individual values that represent real amounts (Bevans, 2022). However, they are bounded between 50 and 100. On the other hand, the maturity assessments are nominal categorical variables that describe the wine's stage of development and drinking readiness. These assessments use letters to indicate specific characteristics related to the wine's maturity. The wines are thus grouped by category, without really having a clear ranking between the different categories. All this indicates the fact they are nominal categorical variables (Bevans, 2022).

By recognizing the fact that the independent variable of our model is thus composed of one discrete quantitative variable (rating ranges), that provide a numerical representation of the wine's quality, and one categorical variable (maturity assessments), that offers qualitative information about its stage of development, we can make appropriate statistical analyses. Indeed, this distinction is needed to interpret and analyze the data in a way that aligns with the nature of these variables.

This research will focus on the rating ranges, as this is what really describes the quality of the wines. The maturity assessment is mainly linked to the age of the wine, and analyzing this part is thus not the main goal of this research. For simplicity reasons, the term *quality ratings* is also used in this research when referring to the *rating ranges*.

3.2.2 Transformation of the Quality Ratings

The quality ratings are bounded between 50 and 100, meaning they can never be smaller than 50 or higher than 100. This is an important aspect to highlight as this means that estimating the quality ratings with multiple linear models, as observed in the literature review articles, may not be appropriate. Multiple linear regressions can have explanatory variables of any value, which means the dependent variable could go from negative to positive infinity. It could therefore estimate ratings that are not between 50 and 100, which should not be possible. To solve this issue, the quality ratings of the four varieties were transformed before estimating them.

First, fifty was subtracted from the quality ratings, and then they were divided by fifty. This linear transformation permits to obtain all quality ratings ranging between 0 and 1. The quality ratings can then be compared to probabilities. Following this, the logistic response function can be used to constrain the outcomes of the regression. Unlike the linear regression, the logistic regression models the probability of an event happening, based on its independent variables. As a result, the dependent variable is constrained between 0 and 1. By doing the following transformation on the logistic regression, the dependent variables can be estimated with a multiple linear function (Kantar, 2021).

$$P(X)_i = Y_i = \frac{1}{1+e^{-(\beta_0+\beta_1 X_{1,i}+\cdots+\beta_n X_{n,i})}} \tag{5a}$$

$$\Rightarrow \frac{Y_i}{1-Y_i} = \exp(\beta_0+\beta_1 X_{1,i}+\cdots+\beta_n X_{n,i}) \tag{5b}$$

$$\Rightarrow \ln(\frac{Y_i}{1-Y_i}) = \beta_0+\beta_1 X_{1,i}+\cdots+\beta_n X_{n,i} \tag{5c}$$

(Kanade, 2022; Kantar, 2021)

For this reason, after performing the linear transformation of the quality ratings to constrain them between zero and one, they underwent the transformation of equation 5b and 5c. In the models, the betas represent thus changes in the log-odds scale, which can be described as the logarithm of the ratio of success to failure. In this case, as the dependent variables are not probabilities, the beta's will need to be transformed again to understand the impact more easily. However, this analysis is mostly interested in the signs of the betas. Therefore, understanding the models will not be too constrained by this transformation. Moreover, the regressions were more statistically significant after this logistic transformation. The summary statistics of the transformed variables can be seen in table A3 of the Appendix, section A.2. Furthermore, the evolution of the quality ratings before and after transformation is also given, in figure A2. By looking at this figure and comparing the summary statistics of the ratings before and after transformation, it is clear that the ratings still evolve with similar patterns and have, proportionally, similar summary statistics. This indicates that the transformation does not change the proportional relationship of the quality ratings. Lastly, the results of all the regressions used in this thesis without performing the transformation are given in the appendix, section A.3. The results indicate that performing the transformation indeed permitted to have models with higher adjusted R-squared and more statistically significant variables, except for the Pinot Noir.

3.2.3 Climate variables of NOAA NCEI

To examine the impact of climate factors on the wine ratings, relevant climate data of the state of California was collected from the website *NOAA National Centers for Environmental Information*, known for giving accessibility to environmental data. These variables represent the dependent variables of the statistical models. The climate variables used in this analysis are collected on a monthly basis, going from November 1977 to October 2018, to correspond to the years of the quality ratings. The grape harvest for wine production occurs every year and these grapes, which will influence the wine quality, are thus affected by the climatic conditions of the corresponding year. These monthly variables were saved in Excel, and after rearranging them by hand into 12 columns, representing the months, and 41 lines, representing the years, the seasonal averages for each variable were computed. It is important to note that, when computing the average values for the winter season, the data for November and December was taken from the preceding year, which is the reason why the collected weather data starts in November 1977. Indeed, as mentioned above, the growing cycle of the wine starts in November of the previous year and ends when the grapes are being harvested, in September or October of the following year.

Moreover, the goal is to download the data the most precisely possible. Indeed, the wine is influenced by the weather of the region where it grows. Collecting data about the weather that is the closest to the place the wine grows is what makes this analysis more reliable. In order to collect the right data about weather, the appellations of the different wine varieties were analyzed. From the 28,980 different wines of which the ratings are computed, 14,740 wines come from wineries located in the Napa Valley and 13,228 are from wineries located in the Sonoma County (Robert Parker Wine Advocate, 2023b). The rest of the wines are originated from 5 additional regions, that have quite similar climate conditions. To facilitate the computation of the weather variables for the different wine varieties, only these two regions are considered, because 96.51% of the wines that are analyzed come from there. The website NOAA NCEI provides monthly weather data about the Average Temperature, Precipitation, Maximum Temperature and Minimum Temperature for each county (NOAA National Centers for Environmental Information, 2023d). Therefore, the weather of the Napa Valley is approximated by considering the weather of the whole Napa County. The location of the Napa County and the Sonoma County is given in figure A1 (Appendix section A.1).

- **Average Temperature, Precipitation, Maximum Temperature and Minimum Temperature**

The Average Temperature (AT), Precipitation (Prec), Maximum Temperature (MaxT) and Minimum Temperature (MinT) of each grape variety were computed as a specific weighted average of the observations from the Napa County and the Sonoma County. The data was downloaded on a monthly basis. The Average Temperature, Maximum Temperature and Minimum Temperature are given in degrees Fahrenheit, while the Precipitation is computed in Inches (NOAA National Centers for Environmental Information, 2023d). The weighted averages were computed as follows:

$$Observation/_{variety} = W_1 * Observation_{Sonoma} + W_2 * Observation_{Napa} \tag{6}$$

With W_1 and W_2 different for every variety. The calculations to determine the weights for each variety rely on the number of wines originating from the Sonoma County and the Napa Valley (approximated by the Napa County). It is important to note these calculations concern the Average Temperature, Precipitation, Maximum Temperature and Minimum Temperature, as these are the weather variables available per county.

Out of the 12,989 Cabernet Sauvignon wines, 1,744 of them come from the Sonoma County and 11,145 from the Napa Valley (Robert Parker Wine Advocate, 2023b). Therefore, the weather variables for the Cabernet Sauvignon are computed with $W_1 = 0.134(= 1.7/12.7)$ and $W_2 = 0.866(= 11/12.7)$. Concerning the Chardonnay, there are 4,061 wines originating from the Sonoma County and 2,021 from the Napa Valley (Robert Parker Wine Advocate, 2023b). Hence, the weather variables are computed with $W_1 = 0.667$ and $W_2 = 0.333$. As for the Zinfandel, there are 2,406 wines originating from the Sonoma

County and 979 from the Napa Valley (Robert Parker Wine Advocate, 2023b). Consequently, the values for W_1 and W_2 are respectively 0.71 and 0.29. Finally, out of the 6,124 Pinot Noir wines, 5,017 are produced in the Sonoma County and 595 in the Nappa Valley (Robert Parker Wine Advocate, 2023b). Therefore, $W_1 = 0.893$ and $W_2 = 0.107$. The table below gives a summary of the values for W_1 and W_2. Cabernet Sauvignon is indicated by CS, Chardonnay by CH, Zinfandel by Zi and Pinot Noir by PN.

Table 1: Summary of the values of W_1 and W_2

	CS	CH	Zi	PN
W_1	0.134	0.667	0.71	0.893
W_2	0.866	0.333	0.29	0.107

By doing these computations, the goal is to approximate the impact of the weather on each grape variety the most precisely and correctly possible. However, to extend the scope of the analysis done in this research, the difference between the maximum and minimum temperature was computed and data about the Southern Oscillation Index (SOI) was also downloaded. The variable *Diff* is thus not directly downloaded from the NOAA NCEI website but is calculated based on *MaxT* and *MinT*. The SOI is a global variable and is explained in the next paragraph.

- **The Southern Oscillation Index (SOI)**

The Southern Oscillation Index is a standardized index that measures the air pressure fluctuations between the western and eastern tropical Pacific, by observing the difference in sea level pressure between Tahiti and Darwin, in Australia. It describes the part concerning the atmospheric changes of the El Ninõ-Southern Oscillation (ENSO) cycle. The ENSO is a climate phenomenon that happens irregularly every two to seven years, where the ocean near the Eastern tropical Pacific coast becomes warmer (El Ninõ), and the atmospheric pressures at the ocean surface of the Western tropical Pacific decrease (Southern Oscillation). La Ninã refers to the opposite weather pattern. The SOI time series corresponds thus to the changes in ocean temperatures. Consequently, negative values of the SOI indicate below-normal air pressure at Tahiti and above-normal air pressure at Darwin. When negative (positive) SOI values are observed for an extended period of time, it means the ocean near the Eastern tropical Pacific coast is abnormally warm (cold), which means the world enters an El Ninõ (La Ninã) phase. The neutral phase is indicated with values close/equal to zero.

This world phenomenon has an impact on the ocean temperature, strength of ocean currents, and the fish in the ocean, but also on the global atmospheric circulation, which has an impact on the global temperatures, wind, precipitations and extreme weather events. All of this makes the ENSO one of the most crucial phenomena on Earth (Michelle l'Heureux, 2014; National Geographic, 2022; NOAA National Centers for Environmental Information, 2023b). Given its importance, the impact of this variable on wine quality was analyzed.

The El Ninõ en La Ninã cycles tends to last 9 to 12 months, developing during the spring and weakening during the early summer of the following year. Their influence on the U.S. seasonal climate tends to peak during the winter. During La Ninã, the winter seasons in California tend to be drier and warmer than average (Michelle l'Heureux, 2014; William Roberts and Jayasankar Pillai, 2023). The SOI is therefore computed on a yearly basis, by taking the average of the monthly SOI, starting in June and ending in May of the following year. As a consequence, the data for the SOI was downloaded from June 1977 until May 2018.

A last thing that is worth mentioning is that, although snow could be an important factor influencing wine quality, it was not taken into account in this research. The reason is that it almost never snows in California. According to NOAA NCEI (2023c), the Napa and Sonoma Counties only experienced one day of heavy snow between 1977 and 2018, which was in January 2011. Furthermore, the average annual snowfall between 1981 and 2010 is equal to zero inches in California according to Current Results (2023). Although these are not the exact years of this analysis, it is still a good indicator that this analysis can ignore the level of snow.

3.2.4 Summary statistics

The summary statistics of all the variables are given in table 2. Furthermore, the summary statistics of the weather variables for each season separately are shown in table A1 and A2 (Appendix). "N" represents the number of observations for each variable. N is equal to 492 for the monthly variables as there is one observation every month for 41 years. For the seasonal variables, N is equal to 164 or 82, depending on the number of months included in the corresponding season.

In these tables, the mean, standard deviation, coefficient of variation, minimum value and maximum value are given. The Coefficient of variation is calculated by dividing the standard deviation by the mean. This measure enables to compare the degree of variation of the different variables, although they are measured in different units. However, it should not be used when the mean is negative and/or close to zero, which is why it is not computed for the SOI (Adam Hayes, 2023).

Table 2:
Summary Statistics of all the monthly variables

Variables	Mean (μ)	SD (σ)	CV (σ/μ)	Min	Max
Cabernet Sauvignon (N= 492)	90.63	5.19	0.06	76	98
Monthly AT (N= 492)	59.10	9.37	0.16	41.57	75.53
Monthly Prec (N= 492)	2.81	3.96	1.41	0.00	23.01
Monthly MaxT (N= 492)	71.55	12.06	0.17	50.95	92.46
Monthly MinT (N= 492)	46.65	6.88	0.15	29.48	58.89
Monthly Diff (N= 492)	24.90	5.85	0.23	9.19	35.44
Chardonnay (N= 492)	88.98	4.42	0.05	75	96
Monthly AT (N= 492)	57.90	8.07	0.14	41.83	72.07
Monthly Prec (N= 492)	3.26	4.53	1.39	0.00	26.33
Monthly MaxT (N= 492)	69.59	10.39	0.15	51.53	87.93
Monthly MinT (N= 492)	46.20	6.00	0.13	29.80	56.87
Monthly Diff (N=492)	23.39	5.24	0.22	8.77	33.20
Zinfandel (N= 492)	86.56	5.11	0.06	74	95
Monthly AT (N= 492)	57.81	7.97	0.14	41.86	71.78
Monthly Prec (N= 492)	3.30	4.57	1.38	0.00	26.60
Monthly MaxT (N= 492)	69.44	10.25	0.15	51.58	87.56
Monthly MinT (N= 492)	46.17	5.93	0.13	29.83	56.70
Monthly Diff (N= 492)	23.27	5.20	0.22	8.73	33.02
Pinot Noir (N= 492)	87.98	3.20	0.04	80	94
Monthly AT (N= 492)	57.39	7.53	0.13	41.95	70.60
Monthly Prec (N= 492)	3.46	4.77	1.38	0.00	27.74
Monthly MaxT (N= 492)	68.76	9.69	0.14	51.78	86.01
Monthly MinT (N= 492)	46.01	5.65	0.12	29.94	56.08
Monthly Diff (N=492)	22.75	5.01	0.22	8.59	32.25
Monthly SOI (N= 492)	0.03	0.97	/ [1]	-3.6	2.9

Source: Table prepared by the author. *AT*, *MaxT*, *MinT* and *Diff* are computed in degrees Fahrenheit. *Prec* is computed in Inches.

[1] The Coefficient of variation can not be computed for the SOI because the mean is close to zero (Adam Hayes, 2023)

On average, we can see that the CNC Cabernet Sauvignon tends to be of better quality than the Chardonnay, Zinfandel and Pinot Noir, with a maximum quality rating of 98. However, it has the second-highest coefficient of variation, which means the spread in rating qualities is quite high. The quality of the Cabernet Sauvignon is thus less stable than the Chardonnay and the Pinot Noir. In comparison to that, the Pinot Noir has clearly the most stable quality ratings: it has the lowest SD, the lowest CV and the highest minimum rating. All this indicates this variety of wine has a very stable rating quality. The Zinfandel has the lowest mean quality ratings between the four varieties and the highest CV. The rating qualities are thus less stable and lower on average than the rest. Despite this, the four grape varieties have average ratings that are attributed to wines above average, which indicates that the North Coast of California is a rather good place to produce wine, or at least, was, between 1978 and 2018.

Concerning the weather variables, the coefficient of variation (CV) of the variables describing the average-, minimum-, and maximum temperature tend to be very small, ranging from 0.12 to 0.17. This confirms thus that the temperature in the Napa and Sonama county are very stable, with a mean temperature ranging around 57-59 degrees Fahrenheit. In comparison, the CV of the precipitation is much higher, which indicates that the monthly amount of rain is less stable than the temperatures. The coefficients of variation of the variables *Diff* range in between. In addition to that, the CVs of the weather variables concerning the Pinot Noir are lower than the others, which means the climate variables influencing the Pinot Noir are more stable than the others. The Pinot Noir has thus the most stable climate and the most stable quality ratings. In addition, the CVs of the weather variables concerning the Cabernet Sauvignon are the highest, which is again in line with the fact that this variety has the highest variation in quality rating.

The tables A1 and A2 in the Appendix provide a more detailed overview of the weather variables, by providing the seasonal summary statistics. The first thing that can be noticed by looking at the CVs is that, by splitting the weather variables per season, they became even more stable, confirming that the weather in California is very stable and shows the same weather patterns every season. This is in line with what is said in the Literature review. Although the differences are very small, the AT, MinT and MaxT vary the most during the early-growing season, followed by the winter season, and the harvest season. The highest consistency is observed during the growing season. However, the variable Diff does not completely follow the same pattern as it is more stable during the early-growing season than the winter season. During the winter, the difference between the highest and the lowest temperature is the most fluctuating. Lastly, the minimum temperature during the early-growing season is never below 32 degrees Fahrenheit, indicating that it is on average never freezing during the early-growing season. This is very positive for the wine quality.

In comparison to variables in relation to the temperature, the variability of the precipitation is a lot higher. The average precipitation during the winter is much higher than during the growing season, whereas the CV of the precipitation is the smallest during the winter and the highest during the growing season. Basically, it rains a lot every year during the winter season, while the rain during the growing season is more scarce and more fluctuating. The season with the second most stable precipitation level is the early-growing season, followed by the harvest season.

The variations in SOI are more difficult to analyze as their CV could not be computed. However, El Ninõ phases tend to be more extreme as the minimum value goes until -3.6, while El Ninã phases are less extreme. Moreover, the fact that the SOI has a mean close to zero confirms that it describes a cycle of positive and negative values, that fluctuates around zero.

To conclude, the four wine varieties are influenced by weather variables that have similar patterns. The weather of the Napa and Sonoma counties can be described as rainy and cold during the winter and early-growing seasons and dry and hot during the growing and harvest seasons. It has thus an optimal climate to grow wine.

3.3 Tested Hypotheses

Based on the literature review and wine knowledge, the hypotheses cited below will be tested in this research:

- H1(a): Higher temperatures during the growing season will positively impact the wine ratings for CNC Cabernet Sauvignon, CNC Chardonnay, CNC Zinfandel, and CNC Pinot Noir.
- H1(b): The quadratic temperature during the growing season will negatively impact the wine ratings, as too high temperatures are not optimal for the grapes.
- H1(c): Lower minimum temperatures during the early-growing season will have a negative influence on the wine ratings for all four grape varieties.
- H2(a): Increased drought during the harvest season has a positive influence on the wine ratings for all four grape varieties.
- H2(b): Increased drought during the winter season has a negative influence on the wine ratings for all four grape varieties.

- H3: The Southern Oscillation Index (SOI) is negatively related to wine ratings, indicating that La Niña events contribute to lower quality ratings.

3.4 Statistical Models

For every wine variety, several models were analyzed. The regression below was used, with $Y_i = \frac{Rating_i - 50}{50}$, and $Z_i = ln(\frac{Y_i}{1-Y_i})$, following the transformation shown in equation 5:

$$Z_{i,t} = f(trend_i, precipitation_i, AT_i, AT_i^2, MaxT_i, MinT_i, Diff_i, SOI) \quad (7)$$

Note that every wine variety is analyzed separately, with its separate weather database, described in section 3.2.4. This is represented by the letter i in the equation above. Because SOI is a global variable, it is not dependent on the different wine varieties. AT represents the average temperature, AT^2 is the average temperature squared, MaxT is the maximum temperature, and MinT is the minimum temperature. Diff represents the difference between the maximum and minimum temperature and SOI is the Southern Oscillation Index.

A trend variable was included in all the models in order to take into account improvements of quality that are not related to the climate variables, as done in the works of Ashenfelter and Storchmann (2016), Jones et al. (2005), Niklas and Rinke (2020) and Ramirez (2008). The reason for doing this is that, for every wine variety, the quality ratings experienced a clear positive trend, which is shown in figure A3 of the Appendix, section A.1. The Chardonnay and the Pinot Noir are the varieties that have the most statistically significant trend. This trend could be attributed to technological improvements, more experience and wine knowledge, and/or a time-dependent bias of wine critics known as "score inflation" (Jones et al., 2005). Moreover, according to Ashenfelter and Storchmann (2016), the higher sugar level of wines, associated with higher quality, is mostly due to better vineyard management practices rather than increasing temperatures. California is known to use developed irrigation techniques to control the water supply and wind turbines or braziers to have some control over the temperatures (Dominé, 2001). This is a concrete example of the impact of technology that can lead to better wines. Lastly, contrary to what is mentioned by Charlin & Cifuentes (2023), the trend does not pose a multicollinearity problem in this research, being only weakly correlated with the other variables. Adding a trend to the regression allows them to perform better.

Furthermore, for every model, the correlation between the variables and the Variance Inflation Factor (VIF) was computed to check the presence of multicollinearity. The VIF factor computes how much the variance of a regression coefficient increases due to the presence of collinearity between the explanatory variables. A VIF equal to one indicates that the variables are not correlated, while a VIF higher than 5 indicates a multicollinearity problem (The Investopedia Team, 2023). Some authors use other thresholds, such as 2.5 or 10, however, a VIF higher than five is also the threshold used by Haeger & Storchmann (2006). A distinction to make is that, if the VIF is high due to the presence of a variable and its quadratic term, the multicollinearity is not considered problematic (CFI Team, 2020). All the models described in this research have all VIF values below 2.5, except the ones with a quadratic term, due to the high correlation between the variable and its quadratic term.

In addition to this, for every model, the Breusch-Pagan Test (BP test) was performed on every regression and the Jarque Bera test (JB test) on the residuals of every regression, to test the presence of respectively, heteroscedasticity and the normality of the residuals. If the null hypothesis of homoscedasticity and/or the normality of the residuals is rejected at a 10% level, robust standard errors are applied on the models to correct for this (indicate with the standard errors being between brackets). If not, the Ordinary Least Square (OLS) estimates and standard errors are used (indicated with the standard errors being between parentheses). The robust standard errors are thus not always applied to avoid having less precise estimates (Mansournia et al., 2020).

Lastly, the first model that was tested for each wine variety is based on the tested hypotheses, presented in paragraph 3.3. The first model tests thus the influence of the temperature during the early-growing and growing period, and the rain during the harvest and the winter period. For every regression, the winter, early-growing, growing, and harvest seasons are indicated by adding respectively a "W", "EG", "G", or "H" after the abbreviation of the weather variable. It is good to know that because AT_{EG} is highly negatively correlated with $Prec_{EG}$, the level of precipitation during the early-growing season is also captured a bit by this model:

$$Z_{i,t} = \beta_0 + \beta_1 AT_{G,i} + \beta_2 AT_{G,i}^2 + \beta_3 AT_{EG,i} + \beta_4 Prec_{H,i} + \beta_5 Prec_{W,i} + SOI + trend_i \quad (8)$$

However, because this model performs badly for the four wine varieties, alternative models were tested to try to find a model for each wine variety that can describe the variation in the quality ratings better. The goal is to see if, by including other weather variables, it is possible to describe the wine quality better. It is also important to note that the models described in the literature review also perform badly, which is why they are not shown in this research.

To achieve better models, a yearly analysis was computed before splitting the weather variables into seasonal periods. The goal is to analyze if yearly weather variables can explain the variation in the quality ratings of the wine variety. Furthermore, doing a yearly analysis could allow to see which variables potentially impact the quality of each wine variety the most. Yearly weather variables are denoted by adding the letter "y" before the abbreviation of the weather variable. For example, yAT stands for yearly temperature. For each variety, several models were analyzed. The models with the highest adjusted R-squared and the most statistically significant variables were kept. It is important to note that, due to a high correlation between the weather variables, the variables could not all be put together in one model, as it would induce a multicollinearity problem. The correlations of the yearly weather variables for the four wine varieties are given in figure A4, A5, A6 and A7 in the section A.1 of the Appendix. Although they differ slightly, the correlation plots are very similar between the different varieties, which is foreseeable as the climate is very stable in California.

The correlation plots indicate that the yearly average temperature, the yearly maximum temperature, and the yearly minimum temperature are highly correlated, which makes sense. Furthermore, the variable $yDiff(= yMaxT - yMinT)$ is highly negatively correlated with the yearly precipitation and positively correlated with $yMaxT$, indicating that when the difference between the maximum temperature and the minimum temperature is high, it tends to rain less. This can be explained by the fact that *yDiff* is mostly dependent on *yMaxT*, and that the precipitation is negatively correlated with the maximum temperatures. Basically, it tends to rain more in California when it is colder. This makes sense as in California, it is mostly raining during the winter season, while the growing seasons are very dry. This is confirmed by the seasonal summary statistics (table A1 and A2). In addition to that, the trend is a bit correlated with yAT and *yMaxT* because they also experience a positive trend, that can be linked with climate change. Lastly, the SOI is correlated with no variable, which can potentially make it interesting, as it adds new information.

To conclude, because the climate in California is very similar every year, the weather variables tend to be correlated with each other. It is therefore important to take this into account when searching for the best model describing the wine quality.

4 Statistical results

4.1 Results of the first model tested for the four wine varieties

The results of the regression, described by model 8, are given in table 3. The first thing that stands out is the fact that almost no variable is statistically significant, even at a level of 10%. Apart from the trend, $Prec_w$ is the most frequently statistically significant weather variable, being significant for three of the four varieties. However, while it was seen in the literature review that rain in the winter tends to positively impact the quality of the wine, here it seems that it is the opposite: more rain during the winter seems to negatively impact the Cabernet Sauvignon, the Zinfandel and the Pinot Noir. Another counterintuitive observation is that, for the Chardonnay, the Zinfandel and the Pinot Noir, although they are not statistically significant at a level of 10%, the coefficients of the average temperature during the growing season are negative, while they are positive for the quadratic terms. This could indicate that, because the temperatures during the growing season in the Sonoma and Napa Counties are warm on average, higher temperatures could have a negative impact on wine quality. It could be possible that the temperatures already exceed the optimum temperature to grow wine in these regions, as mentioned by Haeger and Storchmann (2006). However, as the variables are not statistically significant, this information is not reliable, which is confirmed in table 7, which tested another model for the Zinfandel and shows opposite results with the variables being statistically significant. However, this table will be described later in this analysis (section 4.3).

In spite of this, there are still observations that coincide with what was expected, notably the coefficient of average temperature in the early-growing season that is slightly positive for the four wine varieties, indicating that higher temperatures during the early-growing season can lead to better wine. Moreover, the coefficient for rain during the harvest season is negative for the four wines. Unfortunately, it is difficult to draw any conclusions, as these two observations are only statistically significant at a level of 10% for one of the four wine varieties, being the Chardonnay and Pinot Noir respectively. Lastly, the SOI seems to have a negative impact on wine quality, meaning La Ninã event negatively impacts the wine quality. However, the SOI coefficient is not statistically significant for the Cabernet Sauvignon and the Pinot Noir.

To conclude, the influence of the climate on the wine quality seems to be very limited. Although every adjusted R-squared is quite good, the trend is the most statistically significant variable, which explains most of the R-squared of the models. Indeed, the models with the highest adjusted R-squared are the ones of the Chardonnay and the Pinot Noir, which are also the wine varieties with the most statistically significant trend (given as

Table 3:
Results of the four regressions described by model 8

Model 8	CS	CH	Zi	PN
(Intercept)	-131.500	141.293	183.811	2.408
	(238.251)	[124.659]	[143.238]	[14.935]
AT_G	3.429	-4.733	-5.981	-0.343
	(7.011)	[3.828]	[4.405]	[0.444]
AT_G^2	-0.025	0.036	0.046	0.003
	(0.052)	[0.029]	[0.033]	[0.003]
AT_{EG}	0.054	0.058 .	0.044	0.0007
	(0.052)	[0.030]	[0.044]	[0.028]
$Prec_H$	-0.099	-0.063	-0.095	-0.061 .
	(0.120)	[0.049]	[0.082]	[0.035]
$Prec_W$	-0.085 .	0.018	-0.083**	-0.025 .
	(0.044)	[0.026]	[0.030]	[0.015]
SOI	-0.206	-0.172*	-0.309*	-0.103
	(0.138)	[0.084]	[0.115]	[0.070]
Trend	0.142**	0.121***	0.107.	0.113***
	(0.040)	[0.018]	[0.053]	[0.018]
Adjusted R^2	0.367	0.591	0.293	0.499
p-value JB test	0.356	1.094e-05	0.093	0.641
p-value BP test	0.666	0.686	0.444	0.066

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

footnote of figure A3). Furthermore, table A10 in the Appendix shows the results of the regressions with the trend as the only explanatory variable for every wine variety. The adjusted R-squared of the regressions are 29.5%, 49.1%, 10%, and 50.4% for the Cabernet Sauvignon, the Chardonnay, the Zinfandel, and the Pinot Noir respectively. By adding the climate variables shown in table 3 to the regression, the adjusted R-squared of the regressions only slightly increased, and even decreased for the Pinot Noir (36.7%, 59.1%, 29.3% and 49.9% respectively). This means that the model with only the trend as explanatory variable explains the variation in the rating quality of the Pinot Noir better. The models were also tested without the trend. The results are given in table A4 of the Appendix. By removing the trend, the sign of $Prec_W$ for the Chardonnay, and AT_G and AT_G^2 for the Pinot Noir changed. This confirms that, when the variables are not statistically significant, they are very unstable and it is not reliable to take conclusions from them. Lastly, the adjusted R^2 without the trend are much lower, going from 3.8% for the Pinot Noir, to 22.8% for the Zinfandel. The Zinfandel is the less affected by the removal of the trend, while the Pinot Noir is the most affected. The Pinot Noir is a wine variety that is known to be more resistant to extreme temperatures (Haeger & Storchmann, 2006), which could explain that, when growing in very stable regions, such as California, the climate doesn't influence the rating quality of this wine.

Although these results indicate that the climate probably doesn't influence the rating qualities of the North Coast Californian wines a lot, the following paragraphs will describe some models that gave better results for each wine variety in order to assess which model is more able to describe the rating quality of the different wine varieties. The goal is to assess if the varieties are influenced by different climate variables and how their quality can be best modelled.

4.2 The Yearly analysis

As mentioned above, for every wine variety, a regression with the yearly weather variables was first tested. Table 4 shows the results of the best yearly regression for the four wine varieties found for this research. The model with the most statistically significant variables and the highest adjusted R-squared for the Cabernet Sauvignon, Zinfandel and Pinot Noir is obtained by adding *yDiff, SOI* and their trend to the model. One model that includes *yPrec* instead of *yDiff* is shown for the Cabernet Sauvignon. Although the models are very similar, which is expected due to the strong negative correlation between *yPrec* and *yDiff* (see figure A4), both models are shown to reinforce the reliability of the models. The fact that *yPrec* and *yDiff* have opposite signs support the validity of the models. Similar results were obtained for the Zinfandel and the Pinot Noir, but the results are not shown to avoid redundancy. In addition to that, $yMaxT$ also tended to give similar results for the CS and the Zi, as it is positively correlated with *yDiff*, but at a lower significance level compared to *yPrec*. $yMaxT$ was not statistically significant for the PN wine.

By looking at the results, it seems the Cabernet Sauvignon (CS) and the Zinfandel (Zi) are similarly impacted by the yearly weather variables. The variables yAT and $yMinT$ were never statistically significant for these wines (which is why they are not included in the models). Furthermore, how bigger the difference between the maximum and the minimum temperature, how better the rating quality of these wines. Due to their high correlation, it could indicate that yearly precipitation negatively impacts these wine ratings, while $yMaxT$ positively impacts the rating, meaning that higher temperatures can give better wines. However, these variables were less statistically significant, which is why *yDiff* is used in the models. The Pinot Noir is the variety whose quality ratings are less influenced by the weather variables; in addition to its high adjusted R^2 linked to the trend alone, it is the only variety for which SOI is not statistically significant. This variety is less affected by extreme weather, which can explain these results (Haeger & Storchmann, 2006).

Table 4:
Results of the yearly regressions

Variable	CS	Var.	CS	Zi	PN	Var.	CH
(Int.)	-11.937 **	(Int.)	-16.644 ***	-12.035**	-9.871***	(Int.)	-15.708***
	[3.377]		[3.562]	(4.072)	[1.622]		[2.425]
yPrec	-0.286 *	**yDiff**	0.233 **	0.216**	0.074 .	**yAT**	0.114*
	[0.110]		[0.080]	(0.061)	[0.041]		[0.046]
SOI	-0.288 .	**SOI**	-0.302 *	-0.297**	-0.108	**SOI**	-0.169*
	[0.155]		[0.148]	(0.099)	[0.066]		[0.074]
Trend	0.158 ***	**Trend**	0.137 **	0.093 .	0.106***	**Trend**	0.118***
	[0.037]		[0.041]	(0.050)	[0.019]		[0.018]
Adjusted R^2	**0.436**		**0.450**	**0.351**	**0.556**		**0.595**
p-value JB test	0.546		0.616	0.101	0.722		1.532e-06
p-value BP test	0.095		0.063	0.393	0.0138		0.626

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

The Chardonnay stands out from this yearly analysis. It is the only variety that is not affected by *yDiff, yPrec* nor *yMaxT*, but that gives better results by including the average temperature in the regression. This could indicate that the Chardonnay is less sensitive to rain and more sensitive to temperature. The sign of the coefficient of the temperature is in line with what has been seen in the literature review.

The highest adjusted R-squared is obtained by the regression on the Chardonnay, followed by the Pinot Noir, Cabernet Sauvignon and Zinfandel. However, because this is probably due to the trend, all the models were analyzed without their trend. The results are given in table A5 of the Appendix. By removing the trend, the adjusted R-squared of all the models, except one, dropped consequently. Only the model of the Zinfandel is not highly affected by its trend: all the variables are still statistically significant and the adjusted R-squared only decreased by 4.3%. The quality ratings of the Zinfandel seem to be more affected by the weather than the three other varieties. In contrast to the Zinfandel, the Pinot Noir shows completely different results. By removing the trend, its adjusted R^2 drops sharply, down to 16.1%. The decreases in the adjusted R^2 of the Cabernet Sauvignon and the Chardonnay are smaller than for the Pinot Noir, but still higher than for the Zinfandel.

A last thing that can be observed is that, by removing the trend, the statistical significance of the variables *yPrec, yDiff* and yAT increased, becoming statistically significant at a level of 0.1% and 1%. This could be linked to the fact that, for all varieties, it could not be rejected at a level of 5% that yAT and *yDiff* experience a positive trend. They take thus the explanatory part of the trend partially over.

The yearly regressions indicate that the weather probably has an impact on the quality ratings of the wines. However, all the varieties don't react the same way and the impact seems to be limited. The yearly regressions performed better than the regressions described by model 8, which suggests that it is probably possible to find better models by taking the right seasonal weather variables into account. Therefore, for every wine variety, various models were tested to try to find the one that best describes the variation in wine quality ratings.

4.3 Seasonal analysis

This paragraph will give the results of the seasonal regressions that had the most statistically significant variables and the highest adjusted R^2 found in this research. Again, to not induce a multicollinearity problem, not all variables could be put together in a regression. The main correlation patterns are that the average temperature of every season is positively correlated with the maximum and minimum temperature and the difference between both (=*Diff*) of the same season. Moreover, the precipitation of each season was again negatively correlated with the variable "*Diff*" of the same season. SOI and trend have no high correlation with the variables and they could thus always be added to the regressions. Lastly, for every wine variety, other variables are used in the models. The reason for this is that, when the variables used for the regressions on the rating qualities of one variety were tested on another variety, the models performed badly. This can be a sign that the wine varieties don't react exactly the same way to the weather variables.

The first results that are analyzed are the one of the Cabernet Sauvignon. The results are given in table 5. Not surprisingly, the trend is statistically significant at a level of 0.1%. The variable that seems to have the highest influence on the quality rating of the CS, after the trend, is the SOI, being again negative. Following this, the minimum temperatures during the harvest season were often statistically significant. Although it could seem a bit counterintuitive that the coefficient of $MinT_H$ is negative, indicating that if the minimum temperatures during the harvest season increase, it will negatively impact the wine rating, this can be explained. Because it is, on average, hot in California, this could indicate that having some colder temperatures will positively impact the quality of the wine. When the temperatures are too high, the grapes tend to ripen too fast. This can lead to dehydrated fruits and unbalanced wines, with a high alcohol level and low acidity. Having periods of colder temperatures (during the night for example) can help offset this effect (Dominé, 2001), which could explain that $MinT_H$ has a negative impact on the wine quality, and not AT_H (Jones et al., 2005; Ollat et al., 2016). This also confirms that higher temperatures are not always better when it comes to wine, as was mentioned in this literature review. In addition, model A indicates that the precipitation seems to influence the wine quality negatively, again. More rain leads to lower quality of the Cabernet Sauvignon. Model B was tested with, $Diff_{EG}$ and $Diff_G$. They are positively correlated with the variable $MaxT$, but not with the variable $MinT$ of the same period. A positive sign indicated thus that, although too high temperatures during the harvest period negatively impact wine quality, having hotter temperatures during the growing and early-growing season will increase the wine quality for the Cabernet Sauvignon. The results of the same regression, without including the trend variable are given in the Appendix, table A6. Although model A, including $Prec_W$ and $Prec_G$ and the trend, performed better than model B, when removing the trend in both models, the adjusted R^2 of model A dropped more sharply than for model B. This is why model B is also given here. Without trend, $Diff_{EG}$ is statistically significant at a level of 1%. This reinforces the importance of including a trend variable if the goal is to describe the wine quality more accurately. Despite the sharp drop in the adjusted R^2, the signs of the weather coefficient all remain the same when removing the trend, which strengthens the validity of the model.

Table 5:
Results of the seasonal regression of the Cabernet Sauvignon

Model A	Estimate	Std. Error	**Model B**	Estimate	Std. Error
(Intercept)	-3.486	(4.092)	(Intercept)	-11.990 *	(5.282)
$Prec_W$	-0.078 *	(0.037)	$Diff_{EG}$	0.062 .	(0.036)
$Prec_G$	-0.479 .	(0.259)	$Diff_G$	0.134 .	(0.079)
$MinT_H$	-0.230 **	(0.068)	$MinT_H$	-0.140 .	(0.073)
SOI	-0.375 **	(0.116)	SOI	-0.323 **	(0.116)
trend	0.193 ***	(0.033)	trend	0.170 ***	(0.036)
Adjusted R^2		0.524	Adjusted R^2		0.506
p-value JB test		0.999	p-value JB test		0.479
p-value BP test		0.154	p-value BP test		0.644

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

To conclude, the rain seems to have a negative impact on the Cabernet Sauvignon, whereas warmer temperatures during the growing and early-growing season, combined with some colder temperatures during the harvest season (to avoid the grapes ripening too fast), have a positive impact on the quality of the Cabernet Sauvignon. However, the statistical significances of the variables are quite unstable. This can clearly be observed with the results of the same regressions without trend, which show almost no variables being statistically significant. Furthermore, when testing $Prec_W$ and $Diff_{EG}$ together in a model (the two only variables that are statistically significant without trend), $Prec_W$ was not statistically significant anymore. This confirms that the models are quite unstable. Despite that, $Diff_{EG}$ seems to really impact the Cabernet Sauvignon because a model with only $Diff_{EG}$ as an explanatory variable has an adjusted R^2 of 19.79%. Nevertheless, it is important to nuance this information with the results of the Mann-Kendall test of $Diff_{EG}$, which rejected the assumption of no monotonic trend for this variable at a significance level of 7% (2sided p-value = 0.062). It could thus be that the increased statistical significance of $Diff_{EG}$ is due to its ability to capture a portion of the improving wine qualities, associated with better technology and vineyard management. Lastly, including the trend probably allows the weather variable to explain the small variation in quality not explained by the trend, which is why they are not significant anymore without the trend.

Concerning the Chardonnay, the same kind of results are obtained (see table 6). The fact that $MinT_{EG}$ and AT_G are statistically significant at a level of 5% indicates that the quality of the Chardonnay is somewhat positively impacted by the temperature, as was the case in the yearly analysis. Colder minimum temperatures during the early-growing season have a negative impact on the Chardonnay, which is in line with what is expected: freezing temperatures during the early-growing season can be disastrous for the grapes (Ashenfelter & Storchmann, 2016). The coefficient of $Diff_H$ has a positive sign and is somewhat statistically significant, indicating that higher temperatures during harvest are positive for the wine quality. Due to its negative correlation with $Prec_H$, this can somewhat indicate that rain during harvest has a negative impact on quality ratings, which is expected. SOI is again negative, confirming the negative impact of La Ninã episodes on the wine quality. When removing the trend, model B performed better than model A, with an adjusted R^2 of 20.2%, compared to 8.4% for model A. The results are given in table A7 of the Appendix. The fact that all the signs of the weather coefficient remain the same with or without a trend, supports the credibility of the model. When the trend is not included, AT_G seems to have the highest influence on the rating quality of the Chardonnay. However, this information is to be treated with caution because the null hypothesis of the Mann-Kendall test (that there is no monotonic trend in the series) was rejected for the variable AT_G at a level of 6% (2-sided p-value = 0.053). Similarly to what is observed for the Cabernet Sauvignon, it is difficult to know if the higher adjusted R^2 of model B without trend is because AT_G has a higher influence on the wine quality of the Chardonnay, or if it is because AT_G captures a part of the higher ratings due to better technology and vineyard management.

Table 6:
Results of the seasonal regression of the Chardonnay

Model A	Estimate	Std. Error	**Model B**	Estimate	Std. Error
(Intercept)	-14.822 ***	[2.063]	(Intercept)	-14.753 ***	[3.23]
$MinT_{EG}$	0.079 *	[0.035]	AT_G	0.106 *	[0.052]
$Diff_H$	0.052 .	[0.028]	$Diff_G$	-0.080	[0.048]
SOI	-0.166 .	[0.094]	$Diff_H$	0.033	[0.031]
Trend	0.128 ***	[0.017]	SOI	-0.155 .	[0.079]
			Trend	0.117***	[0.016]
Adjusted R^2		0.600	Adjusted R^2		0.577
p-value JB test		3.399e-05	p-value JB test		3.344e-05
p-value BP test		0.497	p-value BP test		0.832

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

To conclude, similarly to the yearly regression, the variables related to temperature are more statistically significant than the precipitation, concerning the quality ratings of the Chardonnay. By including the trend, model A has the highest adjusted R-squared of all the models that were tested for this research. This model is able to explain 60% of the variation in the quality ratings of the Chardonnay wine, which is quite good. Despite this, it is important to keep in mind that the regression including only the trend has an adjusted R^2 of 49.1%. Including the weather variables improve the model only slightly.

The results of the seasonal regression for the Zinfandel are given in table 7. As was mentioned in the yearly analysis, the Zinfandel variety has the lowest two-sided p-value in the Mann-Kendall test and is thus the wine variety where the hypothesis of no monotonic trend in the series can be the least rejected. This can partially explain the lower adjusted R-squared obtained for both seasonal regressions. Except for the trend and the SOI (which show the same results as expected, similar to the other varieties), the variable that is the most often statistically significant is $Prec_W$ and $Diff_W$. Because better statistical results are obtained by including $Diff_W$, this variable is kept in the model shown in this paper (both cannot be together in a model due to their high negative correlation). The fact that the coefficient of $Diff_W$ is positive is in line with the results of table 3, but is counterintuitive. Higher maximum temperatures during the winter, and thus less rain (given their high negative correlation), tend to positively impact the quality of the Zinfandel. This is counter-intuitive as other results seen in the literature review indicated that having more rain during the winter improved the wine quality. Another observation that can be made is that, by adding the variables $Prec_G$, AT_G, and AT_{EG}^2 to model B, the adjusted R-squared increases by 8.3%. However, these variables are not statistically significant anymore after removing the trend from the model (see table A8 in the Appendix). The fact that the coefficient of rain during the growing period is negative, implies that rain during the growing period has also a negative impact on the Zinfandel's quality. The results of this regression indicate that the Zinfandel needs very dry summers to give qualitative wines. Here, the signs of AT_G and AT_{EG}^2 are the ones expected, indicating that warmer growing seasons positively impact wine quality, up to a certain point. The signs of all the coefficients remain the same after removing the trend from the model, which reinforces the validity of the models.

Table 7:
Results of the seasonal regression of the Zinfandel

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-117.740 *	(45.654)	(Intercept)	-9.925 *	(4.072)
$Diff_W$	0.099 **	(0.029)	$Diff_W$	0.114 ***	(0.031)
$Prec_G$	-0.336 .	(0.179)	SOI	-0.276 **	(0.096)
AT_{EG}	3.853 *	(1.633)	Trend	0.103 *	(0.048)
AT_{EG}^2	-0.035 *	(0.015)			
SOI	-0.294 **	(0.093)			
Trend	0.151 **	(0.051)			
Adjusted R^2		0.445	Adjusted R^2		0.362
p-value JB test		0.594	p-value JB test		0.368
p-value BP test		0.384	p-value BP test		0.188

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

The Zinfandel is the variety that obtains the highest adjusted R^2 without the trend included in the regression. It is also the only one for which the SOI was statistically significant without the trend. The fact that the trend is less statistically significant could be due to several factors. It could be that the weather has a bigger influence on this wine variety, which makes it less stable. However, as the climate is really stable in California, and that the adjusted R^2 of the regressions are relatively low, this is probably not the only reason. What could explain this higher volatility in rating quality is the fact that all grapes of the Zinfandel don't ripen at the same speed. This means that the winemakers often let the grapes ripen longer, and the grapes can therefore rot more easily (Jancis

Robinson, 2023). This could explain why AT_{EG} and AT_{EG}^2 are statistically significant for this wine. If temperatures are too hot, the uneven ripening process could be even stronger, as some grapes will then ripen too fast, leading to less good wines (Ewing-Mulligan & McCarthy, 2016; Jancis Robinson, 2023). Furthermore, this variety tends to produce an excessive amount of grapes, resulting in a wine that lacks concentration. The Zinfandel requires thus more attention than other wine varieties (Jancis Robinson, 2023), which could explain why the ratings of this wine variety are more volatile.

Lastly, the results for the seasonal analysis of the Pinot Noir are given in table 8. It is the only model for which no model was found for this research with weather variables statistically significant at a level of 5%, or better. It is striking that when the trend (see results table A8) is removed, the adjusted R-squared becomes extremely small (0.0989%), and none of the variables are statistically significant. This confirms once again that Pinot Noir does not seem to be affected by climate in the Sonoma and Napa Valley and that the trend has the highest explanatory power.

Table 8:
Results of the seasonal regression of the Pinot Noir

Variable	Estimate	Std. Error
(Intercept)	-11.281***	[1.889]
Diff_G	0.074 .	[0.040]
SOI	-0.112 .	[0.064]
Trend	0.119***	[0.018]
Ajd. R^2		0.560
p-value JB test		0.925
p-value BP test		0.018

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

4.4 Summary of the results

While a lot of models were analyzed, some conclusions can be made. The explanatory power of the trend is the most striking. The quality ratings of the wine varieties experience a clear upward trend. This indicates that factors such as more experience, better technology, and/or the influence of score inflation indeed played a role in the evolution of wine quality of the four North Coast Californian wines. It is possible that, by including a trend, it allows the weather variables to explain the small variation in wine quality, not linked to the trend. However, as the trend plays a crucial role in capturing the variation in wine quality, the weather variables cannot explain independently the variation in the wine quality. Their predictive power is quite weak.

However, the variation in the quality ratings of the Cabernet Sauvignon and Zinfandel are less explained by the trend than the other variables. They are also the wine varieties that have the quality ratings with the highest standard deviation and CV, meaning their quality ratings fluctuate more than the others. This could be the first indication that these varieties of wine are more influenced by the weather than the two other varieties. The Cabernet Sauvignon, however, is still more impacted by the trend than the Zinfandel, which can explain the fact that its models obtain higher adjusted R^2. As mentioned before, the Zinfandel is known to demand higher care due to its uneven ripening of the grapes and its excessive production of amounts of grapes per plant (Jancis Robinson, 2023), which could explain a part of its higher volatility in quality ratings. This could explain why the trend is less statistically significant.

The Pinot Noir and the Chardonnay, on the other hand, have a really clear influence of a trend. The results of the Mann-Kendall test (given as footnote of figure A3) and their high adjusted R^2 linked to their trend alone (see table A10) show this. However, the Chardonnay seems to be more influenced by the weather than the Pinot Noir as it shows a higher adjusted R-squared than the Pinot Noir, with and without trend. The Pinot Noir is the wine variety that improved the least by adding weather variables in the models, which indicates that its quality is almost not influenced by the weather.

Another striking finding is the consistent statistical significance of the Southern Oscillation index at a level of 10% or better. However, when the trend is removed, the significance of this variable decreases and is only statistically significant for the Zinfandel. The SOI is thus a good variable to add when analyzing the wine quality. Based on the results, it can be concluded that La Ninã episodes negatively impact wine quality, while El Ninõ can have a positive influence on wine quality.

Furthermore, the rain seems to have negative coefficients, indicating that more rain tends to lower the quality of the wines. The variables $Diff$ are often statistically significant, indicating that taking extreme weather variables into account, instead of only looking at the average temperatures, can help to improve the models.

Although for every wine variety, a performing model was found, the impact of the climate on the quality ratings stays very limited. The yearly regression performed almost as well as the seasonal regression. The reason for that could be that the climate in the Napa and Sonoma county region is consistent throughout the year, with distinct patterns such as warm and dry summers and rainy winters. As a result, the yearly weather variables capture enough information. The stability of the weather variables is evident when considering the coefficient of variation (CV) of the seasonal weather variables, which ranges from 0.2 to 1.09. The weather is very stable and splitting the weather variables by season seems to only improve the models slightly. The fact they are so stable is probably the reason why they cannot explain the variation in the quality rating of the different wine varieties.

To conclude, by providing a more profound analysis of each wine variety, the goal was to indicate that, despite the very stable climate and the improvement in wine knowledge and technology, the weather variables still have some degree of influence on the variations in wine quality. For the four wine varieties, the variable trend, which includes non-weather related variables, such as technological improvements, better experience or increase in wine knowledge, and/or a time-dependent bias in wine critics, has the most explanatory power for wine quality. This explains the limited explanatory power of the weather variables. However, this research showed that the weather variables cannot be completely ignored when analyzing the quality of North Coast Californian wines. Including them in the models results in an improvement of the adjusted R^2 of the models, which is modest, but not negligible.

5 Discussion and further improvement points

As mentioned above, despite the very stable climate in the Napa and Sonoma county and the high explanatory power of non-weather variables (trend), this research could demonstrate that the weather variables have some degree of influence on the variations in wine quality. However, this research has some limitations. Firstly, the results of this research are only applicable to the Californian state. Specifically, it is only applicable on the *CNC Cabernet Sauvignon, CNC Chardonnay, CNC Zinfandel*, and *CNC Pinot Noir* from 1978 to 2018. It is therefore not possible to generalize the findings of the statistical analyses done in this research to other grapes varieties, regions or years.

Another limitation is that the analysis done in this research is based on the quality ratings provided by RPWA. It is important to note that RPWA's ratings are very specific and tailored to their criteria. Secondly, while RPWA claims to have a rigorous selection process for their reviewers, including individuals with extensive expertise and knowledge in specific wine regions, there is still the possibility of subjective elements influencing the ratings. Indeed, although the reviewer team adheres to a strict code of ethics to maintain impartiality, the subjectivity of wine tasting cannot be completely eliminated. Furthermore, RPWA itself acknowledges that the written commentary, or tasting notes, accompanying the ratings offer more comprehensive insights into a wine style, personality, value, aging potential, and quality in comparison to peers than the scores alone (Robert Parker Wine Advocate, 2023a). Therefore, it is important to consider these limitations when interpreting and drawing conclusions from the quality ratings used in this research.

Moreover, the quality ratings analyzed in this research are calculated as a mean of several individual ratings, given to wines of a specific variety, in a specific year. The fact that it is the mean of several different ratings may have reduced the volatility of the quality ratings. This could potentially "hide" a significant part of the volatility in the ratings that could have been attributed to weather factors. Analyzing wines more precisely, by taking a lower number of different wines per grape variety into account could help find a higher degree of influence of the climate on the wine quality.

In addition to that, the weather variables considered in this research are also averaged values. Including only average values of the weather variables in the multiple linear regressions may neglect the impact of extreme weather conditions on wine quality. Efforts were made to reduce this effect, by taking $MinT$, $MaxT$ and *Diff* into account. Despite that, these variables are still computed as an average of the minimum and maximum temperatures, and the difference between both respectively and are thus not completely representative of extreme weather patterns. Including $Max(MaxT)$ and $Min(MinT)$ in the models, as done in the work of Charlin & Cifuentes (2023), or other extreme weather patterns, could have potentially improved the explanatory power of the models.

Furthermore, as well as taking the average values into account, the variables are also computed as average values of the climate throughout the whole of Sonoma and Nappa Counties. Although this effect was a bit reduced by taking a weighted average, specific to every wine variety, into account, taking the climate of the whole county probably makes the weather variables more stable than they really are, which could also have an impact on the outcome of this research. In these counties, the different varieties grow in very specific regions, that are located at different altitudes and are thus influenced by slightly different climates, which is not represented in this research.

Following this, although a trend is included in the models, which takes some effects not related to climate into account, the effects of the different vineyard characteristics, such as soil, terroir, and altitude are not taken into account as the quality ratings are averaged values. It was thus not possible to take these into account for this research. It could therefore be interesting to add these variables in future models by computing the same kind of analysis, while focusing on specific wines, for which it is possible to add variables related to their specific location, in the models.

Another limitation of this research is that, although the statistical analyses done in this research are based on several climate factors, they certainly do not encompass all potential weather variables that impact the quality of the wine. Cloudiness, humidity and sun hours could be examples of interesting variables to analyze (Charlin & Cifuentes, 2022, 2023)). However, these variables are difficult to collect and thus may be less reliable.

An additional aspect to mention is that the yearly average temperatures described in this research experience a positive trend over time. It is therefore difficult to assess if a part of the increasing quality ratings is in fact due to these increasing temperatures, and if the trend added to the models does take into account the effect of increased temperatures over time, or not. The trend seems to improve the models, which is why it was used.

Another aspect to consider is that, because the climate is so stable in these regions, the weather variables are a lot correlated between them. The reason why the yearly precipitation is negatively correlated with the yearly average temperature is that it tends to rain a lot during the winter and not in the summer. This impact is reduced by taking seasonal variables. Despite this, the correlation between the seasonal weather variables didn't completely disappear and some of them stay correlated. This makes the models somewhat unstable. For example, $Diff$ and $Prec$ were often statistically significant, but $Diff$ gave better results. It is thus difficult to assess to what extent the precipitation has an influence on the wine quality ratings, or if the temperature indeed has more impact on the wine quality. Moreover, some variables are statistically significant, and by adding other variables, become not statistically significant anymore. Although this research has tried to present the best-performing models for each wine variety (which are models that have the highest adjusted R^2 and the most statistically significant variables), it is difficult to be certain which weather variable actually has the greatest impact on each wine variety. A good example is the case of AT_{EG}, which is not statistically significant for the first model analyzed for the Zinfandel (see table 3), but is well in the seasonal analysis of the Zinfandel (see table 7). This makes the impact of the average temperature on the Zinfandel during the early-growing season therefore difficult to assess.

To conclude, the models presented in this research are a good illustration that the weather variables play a (modest) role in the quality of the various wines. However, the exact impact is difficult to conclude based on these models, as other models were possible, with other variables being statistically significant. Despite this, the models presented in this research are the ones that have the highest adjusted R-squared, found during this research. This could thus be a first indication that the weather variables presented in the models are the ones that impact the quality ratings of the four wine varieties the most, although additional analyses are needed to further confirm this.

6 Conclusion

This research examines the relationship between weather conditions and the quality of four distinct wine varieties: the Californian North Coast Cabernet Sauvignon, Chardonnay, Zinfandel and Pinot Noir. This paper relies on a data set of respectively 12,989, 6,268, 6,124, and 3,599 quality ratings, along with monthly weather variables, collected over a period of 41 years. The objective of this analysis is to determine whether including specific weather factors in the models can permit to obtain performing models, capable of describing a greater variation in the quality ratings of the four wine varieties. The quality ratings were transformed to account for the fact that they are bounded between 50 and 100, and the weather variables were averaged over two counties for greater precision.

The main conclusion of this research is that weather does play a role in explaining the variations in wine quality. However, this role is quite limited. Winemakers are able to limit the impact of the weather by employing new technologies and knowing which grape varieties grow best in which specific place. This knowledge helps them achieve better results and mitigate the impact of weather on their wine production. The statistical significance of the trends for each wine variety confirms an overall improvement in wine quality over time. Without considering the trend, the explanatory power of the models is reduced, indicating that the impact of the weather on the wine quality is indeed limited. The consistent weather of California is probably also reducing this impact, as the weather there is optimal to produce wine.

Despite this limited impact, the results of this research reveal that the wine varieties don't react the same way to the weather variables. The models for the Zinfandel have the smallest adjusted R^2, and it is also the variety that is the least influenced by the trend. The quality ratings of this variety are also the most volatile. This can be explained by the fact that the Zinfandel demands higher care due to the uneven ripening process of its grapes and because it tends to give too many grapes, which results in a wine that lacks concentration (Jancis Robinson, 2023). The Cabernet Sauvignon also seems to be quite influenced by the weather. By adding weather variables to the models, the adjusted R^2 of the models improved consequently (52.4% and 50.6%), compared to the model with the trend as the only explanatory variable (29.5%). In addition to this, the adjusted R^2 of the Chardonnay's models also improves a bit by adding weather variables to the models, but it improves less than it does for the CS. The influence of the weather is limited. Despite the limited impact of the weather on the quality ratings, it is the variety that obtained the model with the highest adjusted R^2. It is thus the model for which the highest proportion of variance in the quality ratings is described by the dependent variables. Lastly, the Pinot Noir is the variety that seems the least influenced by the weather variables. The adjusted R-squared of the models only slightly increased, and no model was found with weather variables statistically significant at a level of 5%, or better. Furthermore, without the trend, the adjusted R^2 of its model becomes ridiculously small (going from 56.0% with a trend to 0.0989% without a trend). This is in line with the research of Haeger & Storchmann (2006) that states that the Pinor Noir is less affected by extreme weather.

All this highlights the importance of analyzing wine varieties separately in future research. Nonetheless, the precipitation and the Southern Oscillation Index seem to have a negative impact on the quality of the four wine varieties. The variable *Diff* is often statistically significant, and often has a positive coefficient, indicating that a greater difference between the maximum and minimum temperatures can have a positive impact on the wine quality. The fact it is often statistically significant highlights the importance of taking extreme weather factors into account when analyzing the influence of the weather on the wine quality (or wine prices). The average temperature, on the other hand, does not appear to be the factor that impacts the wine quality the most. When it does, it is positive, indicating that warmer temperatures positively affect wine quality. Although the increasing temperatures have been shown to be mostly beneficial for wine until now, "the warmer the better" has been shown to be incorrect for several wine varieties, in several studies (Ashenfelter & Storchmann, 2016; Charlin & Cifuentes, 2023; Haeger & Storchmann, 2006; Jones et al., 2005; Niklas & Rinke, 2020; Oczkowski, 2016; Ollat et al., 2016; Ramirez, 2008). The results obtained for the seasonal analysis of the Zinfandel confirm that excessively high temperatures can have a negative impact on the wine quality.

As climate change continues to be a concern worldwide, it is crucial to acknowledge that its influence on wine quality, and thus wine production and prices, will continue, and could become more important around the world. The impact of climate change will probably differ in function of the various regions and wine varieties. Furthermore, climate phenomena, such as El Ninõ and La Ninã episodes could also be influenced by climate change, increasing their impact on wine quality. Assessing which weather variables has a greater influence on the wine quality could therefore help the winemakers to reduce the impact of the weather on the wine quality. This could have an impact on wine prices and therefore help winegrowers to better manage their vineyards and remain profitable.

Although this research shows the impact of the weather on the wine quality is limited, it still shows that they have some influence on the wine quality, that cannot be neglected. This research helped to develop a better understanding of which variables impact the wine quality of the different varieties. This paper confirms that grape varieties need to be analysed separately and that their quality reacts differently to the weather. Moreover, the impact of the Southern Oscillation Index is shown and it could therefore be an important variable to include in further research. This paper also shows the positive impact of performing a logistic transformation on the quality ratings, in order to obtain data that better fit the regression models.

It is thus primordial to keep an eye on the changing weather conditions, to give an insight on which future adaptation strategies winemakers could have to make, such as new technological innovation or relocating their vineyards to more suited regions.

References

Adam Hayes. (2023). *Co-efficient of variation meaning and how to use it*. https://www.investopedia.com/terms/c/coefficientofvariation.asp

Ashenfelter, O. (2010). Predicting the Quality and Prices of Bordeaux Wine. *Journal of Wine Economics*, *5*(1), 40–52. https://doi.org/10.1017/S193143610000136X

Ashenfelter, O., & Storchmann, K. (2016). Climate Change and Wine: A Review of the Economic Implications. *Journal of Wine Economics*, *11*(1), 105–138. https://doi.org/10.1017/jwe.2016.5

Bevans, R. (2022). *Types of Variables in Research & Statistics*. Scribbr. https://www.scribbr.com/methodology/types-of-variables/

Calwineries. (2023). *Annual Life Cycle of the Grape Vine*. https://www.calwineries.com/learn/grape-growing/annual-life-cycle-of-the-vine

CFI Team. (2020). *Variance Inflation Factor (VIF)*. https://corporatefinanceinstitute.com/resources/data-science/variance-inflation-factor-vif/

Charlin, V., & Cifuentes, A. (2022). On Weather and Wine Quality: The Case of the Chilean (Maipo Valley) Cabernet-Sauvignon. https://doi.org/10.13140/RG.2.2.27394.15041

Charlin, V., & Cifuentes, A. (2023). The quality of the Argentinean Malbec and the weather in the Mendoza region. *International Journal of Wine Business Research*, Advance online publication. https://doi.org/10.1108/IJWBR-10-2022-0036

Current Results. (2023). *Average Snowfall by USA State - Current Results*. https://www.currentresults.com/Weather/US/average-snowfall-by-state.php

Dominé, A. (2001). *Le vin*. Könemann.

Ewing-Mulligan, M., & McCarthy, E. (2016). *Le vin pour les nuls* (9th ed.).

Grad Coach. (2021). Qualitative vs Quantitative vs Mixed Methods Research: How to Choose Research Methodology [Video file]. https://www.youtube.com/watch?v=hECPeKv5tPM

Haeger, J. W., & Storchmann, K. (2006). Prices of American Pinot Noir wines: climate, craftsmanship, critics. *Agricultural Economics*, *35*(1), 67–78. https://doi.org/10.1111/j.1574-0862.2006.00140.x

Halliday, J. (2014). *Australian Wine Companion, 2015 edition*. Hardie Grant Books.

Jancis Robinson. (2023). *Zinfandel*. https://www.jancisrobinson.com/learn/grape-varieties/red/zinfandel

Jones, G. V., White, M. A., Cooper, O. R., & Storchmann, K. (2005). Climate change and global wine quality. *Climatic Change*, *73*(3), 319–343. https://doi.org/10.1007/s10584-005-4704-2

Kanade, V. (2022). *What Is Logistic Regression? Equation, Assumptions, Types, and Best Practices*. https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-logistic-regression/

Kantar, L. (2021). Limited Dependent Variables (Logit and Probit Models) and an Application on BIST-100: Logit and Probit Models. In B. Adıgüzel Mercangöz (Ed.), *Handbook of Research on Emerging Theories, Models, and Applications of Financial Econometrics* (pp. 167–195). Springer, Cham. https://doi.org/10.1007/978-3-030-54108-8_7

Mansournia, M. A., Nazemipour, M., Naimi, A. I., Collins, G. S., & Campbell, M. J. (2020). Reflection on modern methods: demystifying robust standard errors for epidemiologists. *International Journal of Epidemiology*, *50*(1), 346–351. https://doi.org/10.1093/ije/dyaa260

Michelle l'Heureux. (2014). *What is the El Niño–Southern Oscillation (ENSO) in a nutshell?* https://www.climate.gov/news-features/blogs/enso/what-el-ni%C3%B1o%E2%80%93southern-oscillation-enso-nutshell

National Geographic. (2022). *El niño*. https://education.nationalgeographic.org/resource/el-nino/

Niklas, B., & Rinke, W. (2020). Pricing Models for German Wine: Hedonic Regression vs. Machine Learning. *Journal of Wine Economics*, *15*(3), 284–311. https://doi.org/10.1017/jwe.2020.16

NOAA National Centers for Environmental Information. (2023a). *About us*. https://www.ncei.noaa.gov/about-us

NOAA National Centers for Environmental Information. (2023b). *El Niño/Southern Oscillation (ENSO): Technical Discussion*. https://www.ncei.noaa.gov/access/monitoring/enso/technical-discussion

NOAA National Centers for Environmental Information. (2023c). *Storm events database - search results for napa and sonoma counties, california*. https://www.ncdc.noaa.gov/stormevents/listevents.jsp?eventType=%28Z%29+Heavy+Snow&eventType=%28Z%29+Ice+Storm&beginDate_mm=11&beginDate_dd=01&beginDate_yyyy=1977&endDate_mm=10&endDate_dd=31&endDate_yyyy=2018&county=NAPA%3A55&county=SONOMA%3A97&hailfilter=0.00&tornfilter=0&windfilter=000&sort=DT&submitbutton=Search&statefips=6%2CCALIFORNIA

NOAA National Centers for Environmental Information. (2023d). *Climate at a Glance: County Time Series*. Retrieved May 20, 2023, from https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/time-series

Oczkowski, E. (2016). The Effect of Weather on Wine Quality and Prices: An Australian Spatial Analysis. *Journal of Wine Economics*, *11*(1), 48–65. https://doi.org/10.1017/jwe.2015.14

Ollat, N., Touzard, J.-M., & van Leeuwen, C. (2016). Climate Change Impacts and Adaptations: New Challenges for the Wine Industry. *Journal of Wine Economics*, *11*(1), 139–149. https://doi.org/10.1017/jwe.2016.3

Outreville, J.-F., & Fur, E. L. (2020). Hedonic Price Functions and Wine Price Determinants: A Review of Empirical Research. *Journal of Agricultural & Food Industrial Organization*, *18*(2). https://doi.org/10.1515/jafio-2019-0028

Ramirez, C. D. (2008). Wine Quality, Wine Prices, and the Weather: Is Napa "Different"? *Journal of Wine Economics*, *3*(2), 114–131. https://doi.org/10.1017/S1931436100001164

Robert Parker Wine Advocate. (2023a). *About The Wine Advocate*. https://www.robertparker.com/about

Robert Parker Wine Advocate. (2023b). *The Wine Advocate Vintage Guide*. Retrieved May 18, 2023, from https://www.robertparker.com/resources/vintage-chart

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Sonoma Valley Visitors Bureau. (2022). *Understanding the Wine Growing Season*. https://www.sonomavalley.com/blog/stories/post/understanding-the-growing-season/

The Investopedia Team. (2023). *Variance Inflation Factor (VIF)*. https://www.investopedia.com/terms/v/variance-inflation-factor.asp

van Leeuwen, C., Friant, P., Choné, X., Tregoat, O., Koundouras, S., & Dubourdieu, D. (2004). Influence of Climate, Soil, and Cultivar on Terroir. *American Journal of Enology and Viticultre*, *55*(3), 207–217. https://doi.org/10.5344/ajev.2004.55.3.207

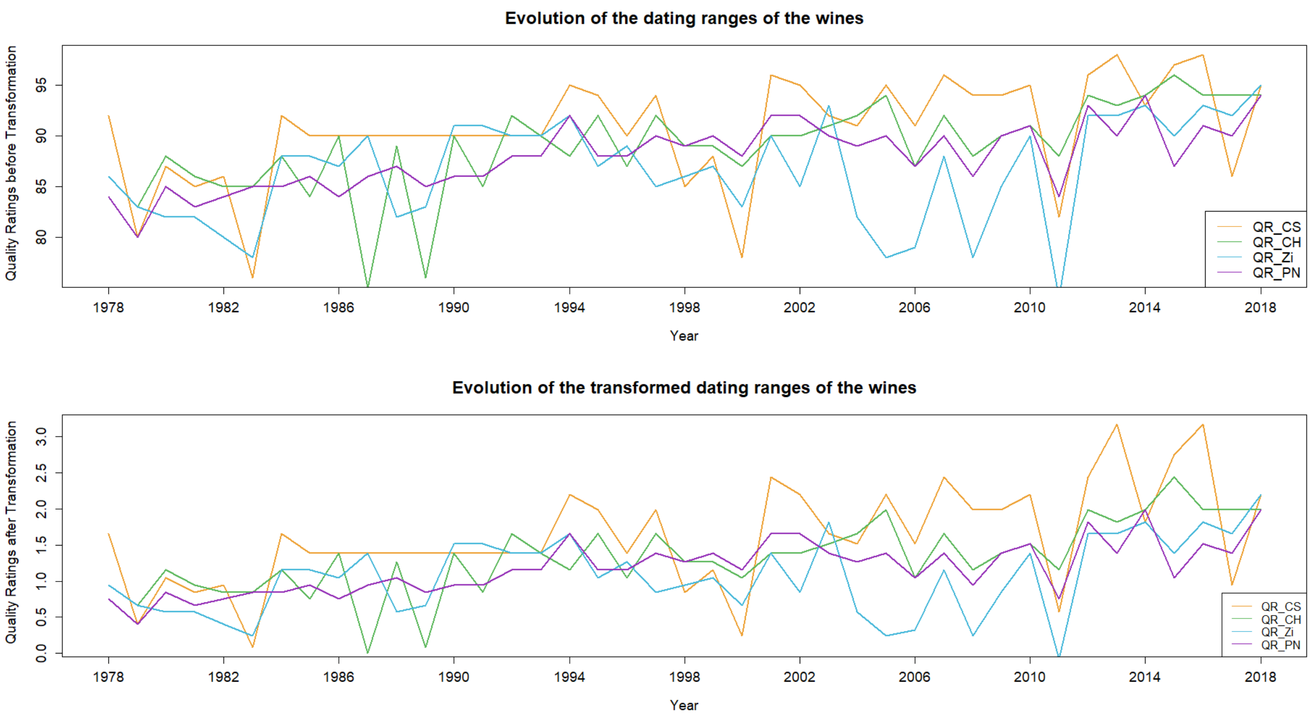
William Roberts and Jayasankar Pillai. (2023). *La niña is finishing an extremely unusual three-year cycle. how it affected weather around the world*. https://theconversation.com/la-nina-is-finishing-an-extremely-unusual-three-year-cycle-heres-how-it-affected-weather-around-the-world-196561

A Appendix

A.1 Figures

Figure A1: **Californian Counties**[1]

Figure A2: **Evolution of the dating ranges before and after transformation**



[1](NOAA National Centers for Environmental Information, 2023d)

Figure A3:
Evolution of the dating ranges and their trend [2]

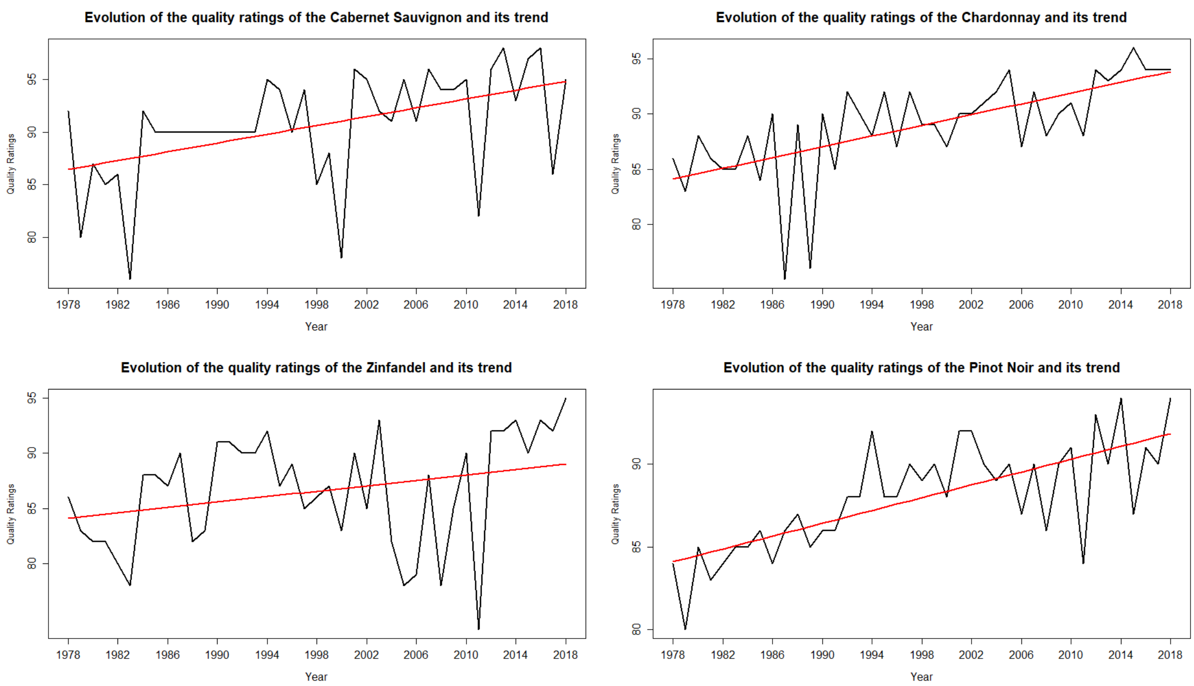


Figure A4: Correlation between the yearly variables for the Cabernet Sauvignon

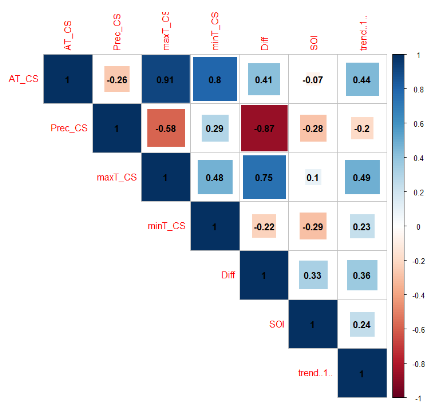
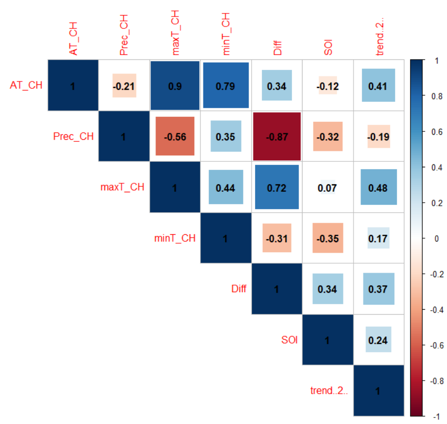


Figure A5: Correlation between the yearly variables for the Chardonnay



[2]The 2-sided p-value of the Mann Kendall test, that has as null hypothesis that there is no monotic trend in the series, were respectively 0.00012517, 4.7684e-07, 0.02405 and 1.1921e-07.

Figure A6: Correlation between the yearly variables for the Zinfandel

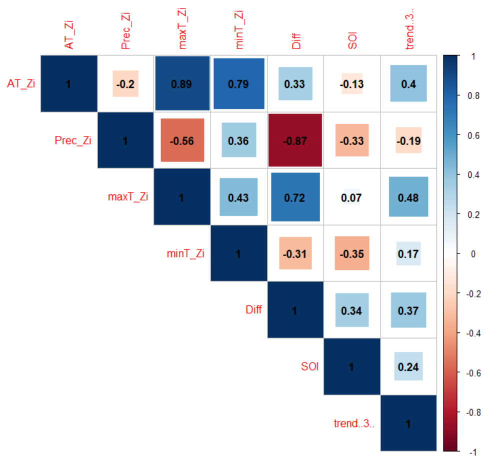
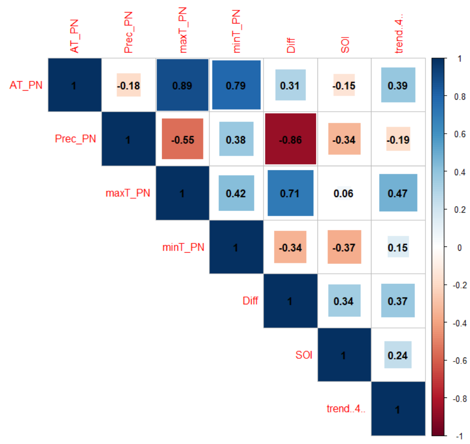


Figure A7: Correlation between the yearly variables for the Pinot Noir



A.2 Tables

Table A1: **Summary Statistics of the yearly SOI and the weather variables per county: Winter and Early-Growing season**

Weather Variables	Mean (μ)	SD (σ)	CV (σ/μ)	Min	Max
Yearly SOI (N=492)	0.02	0.74	/ [1]	-1.85	1.82
WINTER (Nov-Feb)					
Cabernet Sauvignon					
AT (N= 164)	48.98	1.46	0.03	45.96	53.25
Prec (N=164)	5.88	2.28	0.39	1.62	10.41
MaxT (N=164)	58.42	1.99	0.03	54.95	63.85
MinT (N=164)	39.53	1.71	0.04	36.55	43.59
Diff (N=164)	18.89	2.28	0.12	15.05	24.46
Chardonnay					
AT (N= 164)	49.32	1.41	0.03	46.40	53.67
Prec (N=164)	6.81	2.59	0.38	1.91	12.78
MaxT (N=164)	58.44	1.90	0.03	55.26	63.77
MinT (N=164)	40.19	1.77	0.04	36.82	44.37
Diff (N=164)	18.26	2.35	0.13	2.10	14.17
Zinfandel					
AT (N= 164)	49.35	1.41	0.03	46.44	53.7
Prec (N= 164)	6.89	2.62	0.38	1.93	12.97
MaxT (N= 164)	58.44	1.89	0.03	55.28	63.77
MinT (N= 164)	40.24	1.78	0.04	36.84	44.43
Diff (N=164)	18.20	2.35	0.13	14.10	24.07
Pinot Noir					
AT (N= 164)	49.46	1.39	0.03	46.59	53.84
Prec (N=164)	7.21	2.73	0.38	2.03	13.79
MaxT (N=164)	58.45	1.86	0.03	55.39	63.74
MinT (N=164)	40.47	1.8	0.04	36.92	44.69
Diff(N=164)	17.99	2.38	0.13	13.80	23.95
EARLY GROWING (Mar-Apr)					
Cabernet Sauvignon					
AT (N= 82)	54.34	2.06	0.04	50.30	58.76
Prec (N=82)	3.53	2.36	0.67	0.20	9.82
MaxT (N=82)	65.83	3.06	0.05	59.16	72.26
MinT (N=82)	42.85	1.54	0.04	39.77	45.47
Diff (N=82)	22.98	2.52	0.11	16.71	27.67
Chardonnay					
AT (N= 82)	53.58	1.94	0.04	49.72	57.58
Prec (N=82)	4.11	2.73	0.66	0.29	12.03
MaxT (N=82)	64.41	2.89	0.04	58.02	70.50
MinT (N=82)	42.74	1.51	0.04	39.45	45.33
Diff (N=82)	21.67	2.48	0.11	15.57	26.33
Zinfandel					
AT (N= 82)	53.52	1.93	0.04	49.67	57.49
Prec (N= 82)	4.16	2.76	0.66	0.29	12.21
MaxT (N= 82)	64.29	2.88	0.04	57.92	70.36
MinT (N= 82)	42.73	1.51	0.04	39.42	45.32
Diff (N= 82)	21.57	2.48	0.11	15.47	26.22
Pinot Noir					
AT (N= 82)	53.26	1.89	0.04	49.47	57.09
Prec (N= 82)	4.36	2.9	0.67	0.32	12.97
MaxT (N= 82)	63.81	2.83	0.04	57.53	69.75
MinT (N= 82)	42.69	1.5	0.04	39.31	45.28
Diff (N= 82)	21.12	2.48	0.12	15.08	25.77

Source: Table prepared by author. *AT*, *MaxT* and *MinT* are computed in degrees Fahrenheit, *Prec* is computed in Inches.

[1] The Coefficient of variation can not be computed for the SOI because the mean is close to zero (Adam Hayes, 2023)

Table A2: **Summary Statistics of the weather variables per county: Growing and Harvest Season**

Weather Variables	Mean (μ)	SD (σ)	CV (σ/μ)	Min	Max
GROWING (May-Aug)					
Cabernet Sauvignon					
AT (N= 164)	68.29	1.30	0.02	65.20	70.47
Prec (N= 164)	0.31	0.32	1.03	0.02	1.20
MaxT (N= 164)	83.29	1.72	0.02	79.43	86.29
MinT (N= 164)	53.27	1.05	0.02	50.44	55.24
Diff (N= 164)	30.01	1.17	0.04	26.88	32.07
Chardonnay					
AT (N= 164)	65.55	1.19	0.02	62.83	67.57
Prec (N= 164)	0.35	0.38	1.09	0.02	1.47
MaxT (N= 164)	79.34	1.61	0.02	75.78	82.27
MinT (N= 164)	51.74	0.98	0.02	49.38	53.41
Diff (N= 164)	27.60	1.20	0.04	24.52	29.88
Zinfandel					
AT (N= 164)	65.32	1.18	0.02	62.64	67.33
Prec (N= 164)	0.36	0.38	1.06	0.02	1.49
MaxT (N= 164)	79.02	1.6	0.02	75.49	81.94
MinT (N= 164)	51.62	0.98	0.02	49.3	53.31
Diff (N= 164)	27.40	1.21	0.04	24.33	29.70
Pinot Noir					
AT (N= 164)	64.38	1.15	0.02	61.83	66.33
Prec (N= 164)	0.37	0.4	1.08	0.01	1.58
MaxT (N= 164)	77.66	1.57	0.02	74.23	80.56
MinT (N= 164)	51.09	0.97	0.02	48.94	52.88
Diff (N= 164)	26.57	1.23	0.05	23.52	28.94
HARVEST (Sept-Oct)					
Cabernet Sauvignon					
AT (N= 82)	65.74	1.56	0.02	62.70	69.16
Prec(N=82)	0.96	0.80	0.83	0.00	2.88
MaxT (N=82)	80.06	2.17	0.03	75.36	84.32
MinT (N=82)	51.43	1.20	0.02	49.53	54.21
Diff (N= 82)	28.63	1.61	0.06	25.37	31.79
Chardonnay					
AT (N= 82)	64.11	1.50	0.02	61.32	67.58
Prec (N=82)	1.13	0.91	0.8	0.00	3.61
MaxT (N=82)	77.60	2.08	0.03	73.02	81.95
MinT (N=82)	50.62	1.21	0.02	48.83	53.25
Diff (N= 82)	26.98	1.61	0.06	23.48	30.13
Zinfandel					
AT (N= 82)	63.98	1.5	0.02	61.2	67.46
Prec (N= 82)	1.15	0.92	0.80	0.00	3.67
MaxT (N= 82)	77.4	2.08	0.03	72.83	81.77
MinT (N= 82)	50.56	1.21	0.02	48.78	53.17
Diff (N= 82)	26.84	1.62	0.06	23.33	30.00
Pinot Noir					
AT (N= 82)	63.42	1.49	0.02	60.73	66.92
Prec (N= 82)	1.21	0.96	0.79	0.00	3.92
MaxT (N= 82)	76.55	2.06	0.03	72.02	81.03
MinT (N= 82)	50.28	1.23	0.02	48.46	52.98
Diff (N= 82)	26.28	1.64	0.06	22.68	29.43

Source: Table prepared by author. *AT*, *MaxT* and *MinT* are computed in degrees Fahrenheit, *Prec* is computed in Inches.

Table A3:
Summary Statistics of the transformed quality rating [1]

Variables	Mean (μ)	SD (σ)	CV (σ/μ)	Min	Max
Cabernet Sauvignon (N= 41)	1.61	0.72	0.45	0.08	3.18
Chardonnay (N= 41)	1.33	0.51	0.38	0.00	2.44
Zinfandel (N= 41)	1.06	0.53	0.5	-0.08	2.20
Pinot Noir (N= 41)	1.18	0.37	0.31	0.41	1.99

Source: Table prepared by authors.
[1] Here, N= 41 as the transformation is done on the yearly quality ratings.

Table A4:
Results of the four regressions described by model 8 without trend

Model 8	CS	CH	Zi	PN
(Intercept)	-19.451	320.282	172.593	-13.027
	[273.703]	(214.903)	(149.583)	[22.485]
AT_G	0.423	-9.984	-5.392	0.373
	[8.066]	(6.593)	(4.594)	[0.683]
AT_G^2	-0.002	0.077	0.042	-0.002
	[0.059]	(0.050)	(0.035)	[0.005]
AT_{EG}	0.082	0.079 .	0.050	0.005
	[0.060]	(0.045)	(0.046)	[0.032]
$Prec_H$	-0.174	-0.117	-0.120	-0.088
	[0.137]	(0.085)	(0.085)	[0.077]
$Prec_W$	-0.110*	-0.004	-0.091**	-0.038 .
	[0.050]	(0.031)	(0.031)	[0.022]
SOI	-0.114	-0.094	-0.254*	-0.003
	[0.157]	(0.114)	(0.117)	[0.082]
Adjusted R^2	0.151	0.146	0.228	0.038
p-value JB test	0.871	0.131	0.661	0.462
p-value BP test	0.099	0.226	0.417	0.085

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A5:
Results of the yearly regressions without trend

Variable	CS	Var.	CS	Zi	PN	Var.	CH
(Int.)	2.593 ***	(Int.)	-6.348 **	-4.792 **	-1.982 .	(Int.)	-14.256 ***
	(0.343)		[1.868]	(1.382)	(1.023)		[3.429]
yPrec	- 0.347 **	**yDiff**	0.320 ***	0.252 ***	0.139 **	**yAT**	0.269 ***
	(0.116)		[0.075]	(0.059)	(0.045)		[0.059]
SOI	-0.180	**SOI**	-0.239	-0.273 *	-0.066	**SOI**	-0.032
	(0.147)		[0.147]	(0.101)	(0.076)		[0.097]
Adjusted R^2	**0.150**		**0.257**	**0.308**	**0.161**		**0.242**
p-value JB test	0.898		0.788	0.623	0.598		0.014
p-value BP test	0.266		0.063	0.612	0.205		0.886

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A6:
Results of the seasonal regression of the Cabernet Sauvignon without trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	9.330.	(4.741)	(Intercept)	-2.299	(6.169)
Prec_W	-0.110*	(0.050)	Diff_EG	0.118**	(0.043)
Prec_G	-0.288	(0.354)	Diff_G	0.125	(0.099)
MinT_H	-0.136	(0.092)	MinT_H	-0.050	(0.090)
SOI	-0.206	(0.155)	SOI	-0.189	(0.143)
Adjusted R^2		0.094	Adjusted R^2		0.208
p-value JB test		0.917	p-value JB test		0.805
p-value BP test		0.306	p-value BP test		0.772

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A7:
Results of the seasonal regression of the Chardonnay without trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-5.999 .	(3.082)	(Intercept)	-11.719**	4.129
MinT_EG	0.104 .	(0.056)	AT_G	0.241**	(0.075)
Diff_H	0.107 *	(0.050)	Diff_G	-0.174 *	(0.078)
SOI	-0.043	(0.108)	Diff_H	0.076 .	(0.045)
			SOI	-0.006	(0.104)
Adjused R^2		0.084	Adjused R^2		0.202
p-value JB test		0.236	p-value JB test		0.316
p-value BP test		0.544	p-value BP test		0.685

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A8:
Results of the seasonal regression of the Zinfandel without trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-51.794	(44.028)	(Intercept)	-1.311*	(0.571)
Diff_W	0.122***	(0.031)	Diff_W	0.131 ***	(0.031)
Prec_G	-0.204	(0.192)	SOI	-0.241 *	(0.099)
AT_EG	1.841	(1.642)			
AT_{EG}^2	-0.016	(0.015)			
SOI	-0.229**	(0.010)			
Adjused R^2		0.322	Adjusted R^2		0.303
p-value JB test		0.717	p-value JB test		0.852
p-value BP test		0.534	p-value BP test		0.777

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A9:
Results of the seasonal regression of the Pinot Noir

Variables	Estimate	Std. Error
(Intercept)	-0.672	(1.313)
Diff_G	0.070	(0.049)
SOI	-0.019	(0.082)
Ajd. R^2		0.000989
p-value JB test		0.858
p-value BP test		0.339

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A10:
Results of the regression of the four varieties with only their trend

Model: only trend	Cabernet Sauvignon	Chardonnay	Zinfandel	Pinot Noir
(Intercept)	-12.825***	-9.603 ***	-9.814*	-8.771***
	(3.432)	[1.603]	[4.55]	[1.642]
Trend	0.159***	0.123***	0.125*	0.113***
	(0.038)	[0.018]	[0.053]	[0.019]
Ajd. R^2	0.295	0.491	0.100	0.504
p-value JB test	0.110	0.020	0.218	0.397
p-value BP test	0.0701	0.490	0.039	0.020

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

A.3 Results of the models without the logistic transformation

Table A11:
Results of the four regressions described by model 8 w/o transformation

Model 8	CS	CH	Zi	PN
(Intercept)	-655.341	1379.437	1941.721	2.542
	[1995.752]	[1157.237]	[1727.623]	(1093.000)
AT_G	18.980	-43.107	-59.922	-0.067
	[59.036]	[35.623]	[53.035]	(34.040)
AT_G^2	-0.137	0.331	0.464	6.105e-04
	[0.434]	[0.274]	[0.407]	(0.265)
AT_{EG}	0.391	0.392	0.486	0.031
	[0.409]	[0.275]	[0.500]	(0.222)
$Prec_H$	-1.231	-0.591	-0.742	-0.540
	[0.928]	[0.473]	[0.635]	(0.388)
$Prec_W$	-0.650 *	0.250	-0.801 **	-0.185
	[0.315]	[0.264]	[0.227]	(0.144)
SOI	-0.992	-1.339	-2.632 .	-0.868
	[1.570]	[0.859]	[1.310]	(0.554)
Trend	0.800 *	1.001 ***	0.684	0.995 ***
	[0.301]	[0.174]	[0.506]	(0.172)
Adjusted R^2	0.267	0.500	0.230	0.512
p-value JB test	0.003	2.442e-15	0.055	0.213
p-value BP test	0.576	0.616	0.443	0.257

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A12:
Results of the four regressions described by model 8 w/o transformation and trend

Model 8	CS	CH	Zi	PN
(Intercept)	-25.033	2854.047	2500.397	1384.999
	(1990.528)	[1716.074]	(2094.389)	(1491.659)
AT_G	2.074	-86.373	-75.564	-41.119
	(58.661)	[52.663]	(64.470)	(46.573)
AT_G^2	-0.010	0.667	0.586	0.325
	(0.431)	[0.403]	(0.495)	(0.363)
AT_{EG}	0.551	0.562 .	0.543	0.138
	(0.435)	[0.317]	(0.439)	(0.309)
$Prec_H$	-1.658	-1.038	-0.893	-0.842
	(0.993)	[0.852]	(0.814)	(0.538)
$Prec_W$	-0.789 *	0.069	-0.864 **	-0.322
	(0.363)	[0.288]	(0.301)	(0.199)
SOI	-0.471	-0.696	-2.411 *	-0.392
	(1.142)	[1.193]	(1.103)	(0.766)
Adjusted R^2	0.147	0.108	0.216	0.045
p-value JB test	0.130	1.386e-05	0.220	0.460
p-value BP test	0.169	0.275	0.539	0.145

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A13:
Results of the yearly regressions w/o transformation

Variable	**CS**	Var.	**CS**	**Zi**	**PN**	Var.	**CH**
(Int.)	10.130	(Int.)	-27.132	-21.718	-10.076	(Int.)	-41.620 .
	[23.762]		[25.695]	[35.135]	(13.130)		[21.337]
yPrec	-2.170 *	**yDiff**	1.914 **	2.082 **	0.601 *	**yAT**	0.741 .
	[0.874]		[0.624]	[0.648]	(0.296)		[0.389]
SOI	-1.654	**SOI**	-1.821	-2.897 *	-1.053 *	**SOI**	- 1.435 .
	[1.313]		[1.236]	[1.089]	(0.481)		[0.754]
Trend	0.956***	**Trend**	0.774 **	0.692	0.960 ***	**Trend**	0.986 ***
	[0.257]		[0.275]	[0.442]	(0.155)		[0.176]
Adjusted R^2	**0.340**		**0.384**	**0.318**	**0.571**		**0.493**
p-value JB test	0.020		0.067	0.027	0.435		< 2.2e-16
p-value BP test	0.055		0.035	0.654	0.129		0.513

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A14:
Results of the yearly regressions w/o transformation and trend

Variable	CS	Var.	CS	Zi	PN	Var.	CH
(Int.)	97.787 ***	(Int.)	30.864 *	31.946 *	61.041 ***	(Int.)	-29.442
	[2.348]		[14.457]	(13.288)	(8.997)		[28.184]
yPrec	-2.536 **	**yDiff**	2.402 ***	2.350 ***	1.184 **	**yAT**	2.045 ***
	[0.881]		[0.580]	(0.571)	(0.395)		[0.483]
SOI	-1.003	**SOI**	-1.465	-2.717 *	-0.666	**SOI**	-0.283
	[1.361]		[1.003]	(0.973)	(0.673)		[0.937]
Adjusted R^2	**0.150**		**0.275**	**0.300**	**0.149**		**0.171**
p-value JB test	0.061		0.228	0.170	0.491		2.307e-09
p-value BP test	0.340		0.036	0.947	0.735		0.817

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A15:
Results of the seasonal regression of the Cabernet Sauvignon w/o transformation

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	73.790 *	[31.600]	(Intercept)	11.149	[40.420]
Prec_W	-0.561 *	[0.255]	Diff_EG	0.525	[0.314]
Prec_G	-3.030 .	[1.761]	Diff_G	0.985	[0.664]
MinT_H	-1.737 **	[0.566]	MinT_H	-1.069 .	[0.557]
SOI	-2.239 .	[1.174]	SOI	-1.926 .	[1.129]
Trend	1.219 ***	[0.241]	Trend	1.025 ***	[0.223]
Adjusted R^2		0.408	Adjusted R^2		0.429
p-value JB test		0.036	p-value JB test		0.001
p-value BP test		0.176	p-value BP test		0.620

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A16:
Results of the seasonal regression of the Cabernet Sauvignon w/o transformation and trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	154.541 ***	[39.602]	(Intercept)	69.678	(44.164)
Prec_W	-0.761**	[0.264]	Diff_EG	0.865 **	(0.305)
Prec_G	-1.827	[2.004]	Diff_G	0.934	(0.713)
MinT_H	-1.144	[0.766]	MinT_H	-0.524	(0.643)
SOI	-1.173	[1.440]	SOI	-1.118	(1.022)
Adjusted R^2		0.092	Adjusted R^2		0.229
p-value JB test		0.076	p-value JB test		0.147
p-value BP test		0.382	p-value BP test		0.560

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A17:
Results of the seasonal regression of the Chardonnay w/o transformation

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-37.433 *	[18.060]	(Intercept)	-41.056	[29.236]
MinT_EG	0.501	[0.348]	AT_G	0.788 .	[0.434]
Diff_H	0.468 .	[0.266]	Diff_G	-0.583	[0.361]
SOI	-1.437	[0.998]	Diff_H	0.354	[0.310]
Trend	1.039 ***	[0.158]	SOI	-1.315	[0.786]
			Trend	0.955 ***	[0.150]
Adjusted R-squared		0.496	Adjusted R-squared		0.482
p-value JB test		<2.2e-16	p-value JB test		<2.2e-16
p-value BP test		0.497	p-value BP test		0.750

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A18:
Results of the seasonal regression of the Chardonnay w/o transformation and trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	34.099	[23.760]	(Intercept)	-16.354	[29.715]
MinT_EG	0.704	[0.530]	AT_G	1.884 ***	[0.523]
Diff_H	0.919 *	[0.407]	Diff_G	-1.346 *	[0.496]
SOI	-0.439	[1.275]	Diff_H	0.705 .	[0.399]
			SOI	-0.102	[0.970]
Adjusted R-squared		0.056	Adjusted R-squared		0.161
p-value JB test		1.437e-05	p-value JB test		4.78e-06
p-value BP test		0.603	p-value BP test		0.654

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A19:
Results of the seasonal regression of the Zinfandel w/o transformation

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-1015.708 *	(485.089)	(Intercept)	-2.205	(0.551)
Diff_W	0.904 **	(0.296)	Diff_W	1.046 **	(0.307)
Prec_G	-3.129 .	(1.815)	SOI	-2.656 **	(0.960)
AT_EG	36.762 *	(17.636)	Trend	0.806	(0.481)
AT^2_{EG}	-0.340 *	(0.165)			
SOI	-2.940 **	(0.947)			
Trend	1.090 *	(0.492)			
Adjused R^2		0.387	Adjusted R^2		0.308
p-value JB test		0.289	p-value JB test		0.103
p-value BP test		0.251	p-value BP test		0.207

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A20:
Results of the seasonal regression of the Zinfandel w/o transformation and trend

Model A	Estimate	Std. Error	Model B	Estimate	Std. Error
(Intercept)	-588.933	(469.344)	(Intercept)	65.136 ***	(5.564)
Diff_W	1.074 **	(0.302)	Diff_W	1.180 ***	(0.304)
Prec_G	-2.206	(1.862)	SOI	-2.382 *	(0.968)
AT_EG	24.028	(17.578)			
AT^2_{EG}	-0.219	(0.164)			
SOI	-2.393 *	(0.964)			
Adjused R^2		0.318	Adjusted R^2		0.276
p-value JB test		0.634	p-value JB test		0.399
p-value BP test		0.283	p-value BP test		0.392

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1%, 1%, 5% and 10% level.

Table A21:
Results of the seasonal regression of the Pinot Noir w/o transformation

Variables	Estimate	Std. Error
(Intercept)	-21.741	(14.950)
Diff_G	0.605 *	(0.281)
SOI	-1.085 *	(0.481)
Trend	1.065 ***	(0.146)
Ajd. R^2		0.576
p-value JB test		0.569
p-value BP test		0.160

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A22:
Results of the seasonal regression of the Pinot Noir w/o transformation and trend

Variables	Estimate	Std. Error
(Intercept)	72.809 ***	(11.512)
Diff_G	0.571	(0.433)
SOI	-0.261	(0.702)
Ajd. R^2		-0.0065
p-value JB test		0.748
p-value BP test		0.467

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.

Table A23:
Results of the regression of the four varieties with only their trend w/o transformation

Model: only trend	Cabernet Sauvignon	Chardonnay	Zinfandel	Pinot Noir
(Intercept)	-1.012e-12	-2.308e-13	1.349e-12	2.841e-13
	[28.176]	[14.478]	[43.920]	[14.039]
Trend	1.000 **	1.000 ***	1.000 .	1.000 ***
	[0.311]	[0.159]	[0.511]	[0.161]
Ajd. R^2	0.214	0.422	0.062	0.517
p-value JB test	8.862e-05	2.686e-10	0.041	0.089
p-value BP test	0.891	0.175	0.099	0.133

Note: t-statistics are computed in parentheses, heteroscedasticity-consistent t-statistics are computed in brackets. ***, **, * and . indicate significance at the 0.1% , 1% , 5% and 10% level.