California State University, Fullerton
ISDS 526
Forecasting for Analytical Decision Making

Project Report 2
Spring 2017

Northern Napa Valley Winery Inc.

Prof: Dawit Zerom

By

Amogh Newgi (802787945)
Jitesh Punjabi(803181882)
Santosh Konchada (803010297)
Yashwanth Vattikuti (89355377)

Table of Contents

Content	Page Number
Executive Summary	3
1. The Forecasting Problem Set-up	4
2. Examination of Data Patterns	5
3. Model Selection	8
4. The Forecast	12
5. Conclusion and Recommendation	14

Executive Summary

Ms. Quintana CEO of Northern Napa Valley Winery Inc. was considering collaboration with TransContinental stores to sell excess grapes from the 1996 harvest to increase profits. Before taking a decision she must determine the quantity of the harvest that must be retained for the production of Northern Napa's own red table wine. She realized that the sales are related to the amount of red wine produced. She used the sales data of last 8 years to forecast the sales for next one year. The issue she faced was the seasonality of wine sales as December is the peak time for wine sales. We recommend implementing a variable production strategy for months with higher sales as compared to the rest of the year. This strategy would ensure that the cost of lost sales during months of October to December would be kept low. If any wine is left over from this period it can be sold using promotions in the following months and revenues can also be brought in by selling the excessive grapes to TransContinental. We have made this recommendation based on the forecast we created. The time series of the last 8 years of data was analyzed to check for patterns such as trend and seasonality. We then used several forecasting models and compared them based on the accuracy measures each generated. On analysis, Multiplicative Winter's Exponential Smoothing model produced the best results. The fit and accuracy for the model was good and provided us with confidence to forecast the sales for the period October 1996 to September 1997. The following report will explain in detail the steps that we have taken to make this recommendation.

1. Problem Set-Up

Ms. Quintana, major shareholder and CEO of Northern Napa Valley Winery Inc., was proposed an offer by TransContinental Stores for purchasing the excess grapes of 1996 harvest. The framework below in Figure 1 shows the process of the issue at hand. The Goal is to determine the amount of excessive grapes to sell. The issue for Ms. Quintana is to forecast the amount of grapes required for their own production of wines before deciding to sale the excess to TransContinental. The forecast also needs to account for the fluctuations in the demand for wine observed annually. An important aspect mentioned in the case is that wine does not improve with time, so it would be a good option to sale the excess grapes production. Since, there are fluctuations in the demand for wine, with peaks in December; trend, seasonality or cyclicality factors needs to be taken into consideration for accurate forecasts. They should ensure that they the excess grapes they sell to TransContinental does not exceed their requirements for the year and to minimize the cost of lost sales or of overproduction.

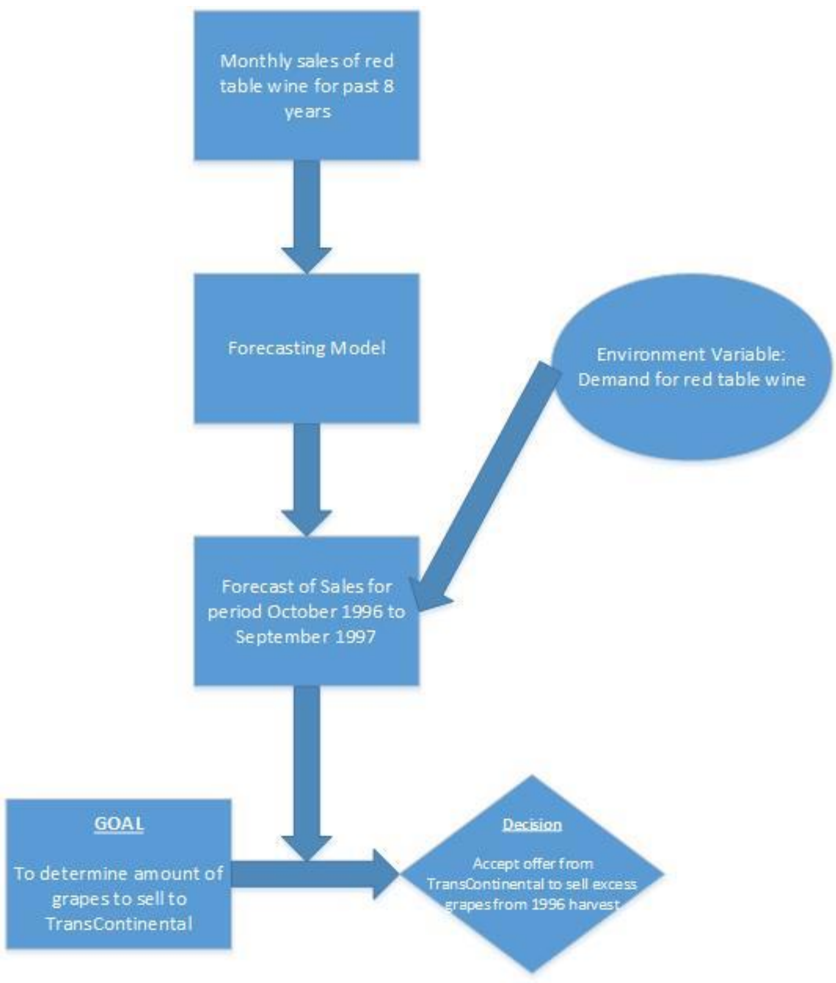

Figure 1: Graphical representation of Decision Making

In the next section, we examine the time series to analyze any patterns such as trend and seasonality that might exist. This is done by using graphical representations and the autocorrelation function of Forecast Pro.

2. Examination of Data Patterns

2.1 Objective

In this section the aim is to produce find patterns within the time series using Forecast Pro. We use graphical visuals and autocorrelation analysis to achieve this aim. To further analyze the time series for existence of trend or seasonality, we will use simple and seasonal differencing.

2.2 Analysis of Time Series

The data available to use consisted of monthly wine sales for Northern Napa Valley Winery from January 1998 through July 1996. For examining the data patterns, we have used MS Excel to showcase the historical trend in the data. The graph shown in Figure 2 exhibits the sales for red table wine.

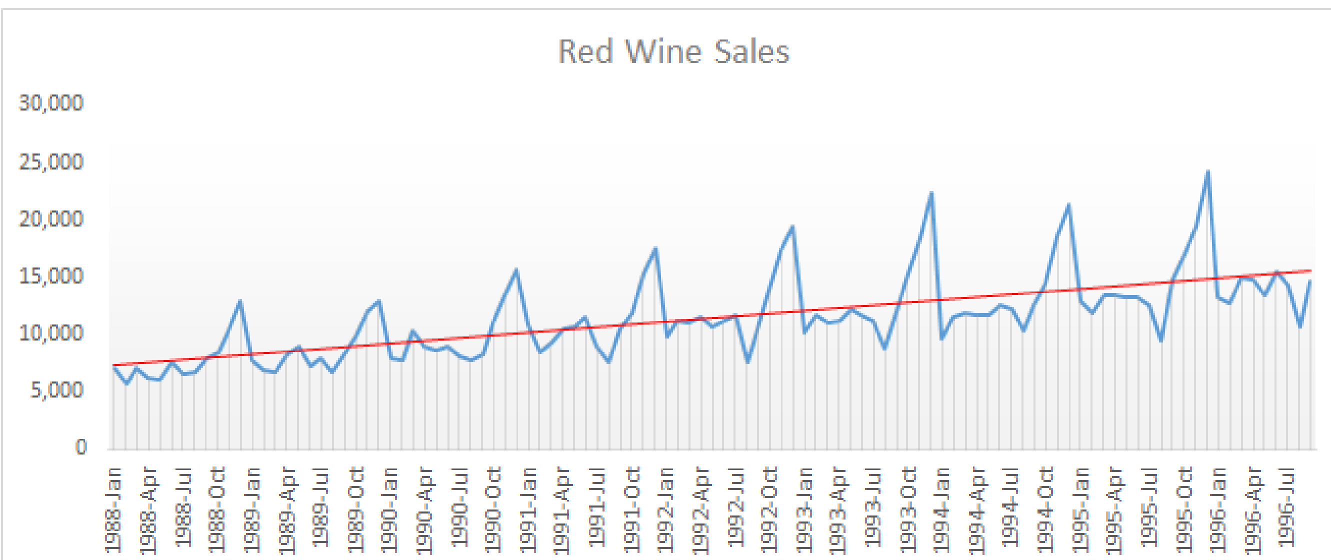

Figure 2: Historical trend of red wine sales

As seen from Figure 2, we can clearly see that there is an increasing trend in the sales of wine for the period ranging from 1988 to 1996. A time series is said to be exhibiting a trend when it shows a long-term increase or decrease in the data. From this we can infer that, since the time series has trend, it cannot be stationary. Also, the trend line indicates that there is a positive trend in the sales of red wine because the average is increasing over the period.

We also observe that the time series has peaks during the quarter of October to January of every year, which is an indication of seasonality. A time series that shows periodic increases or decreases that may be weekly, monthly or yearly is said to be seasonal.

However, a graphical representation is not always conclusive so we perform Autocorrelation analysis to understand and conclude on the aspects of trend and seasonality.

2.3 Autocorrelation Analysis

Autocorrelation is the relationship between a variable and delayed copy of the same variable. To further analyze the time series, we use the autocorrelation function in Forecast Pro to generate correlogram.

Figure 3 below shows the correlogram of the time series as generate by Forecast Pro. If a time series has seasonality, time lags at consistent intervals such seasonal time lag or multiples of seasonal lag need to be analyzed. As seen from the correlogram in Figure 3, we see that values at time lags 1-12-24-36-48 are highly significant when compared to other value. This pattern in the time series implies seasonality. However, ACF is not a clear indicator of time series patterns and requires further analysis.

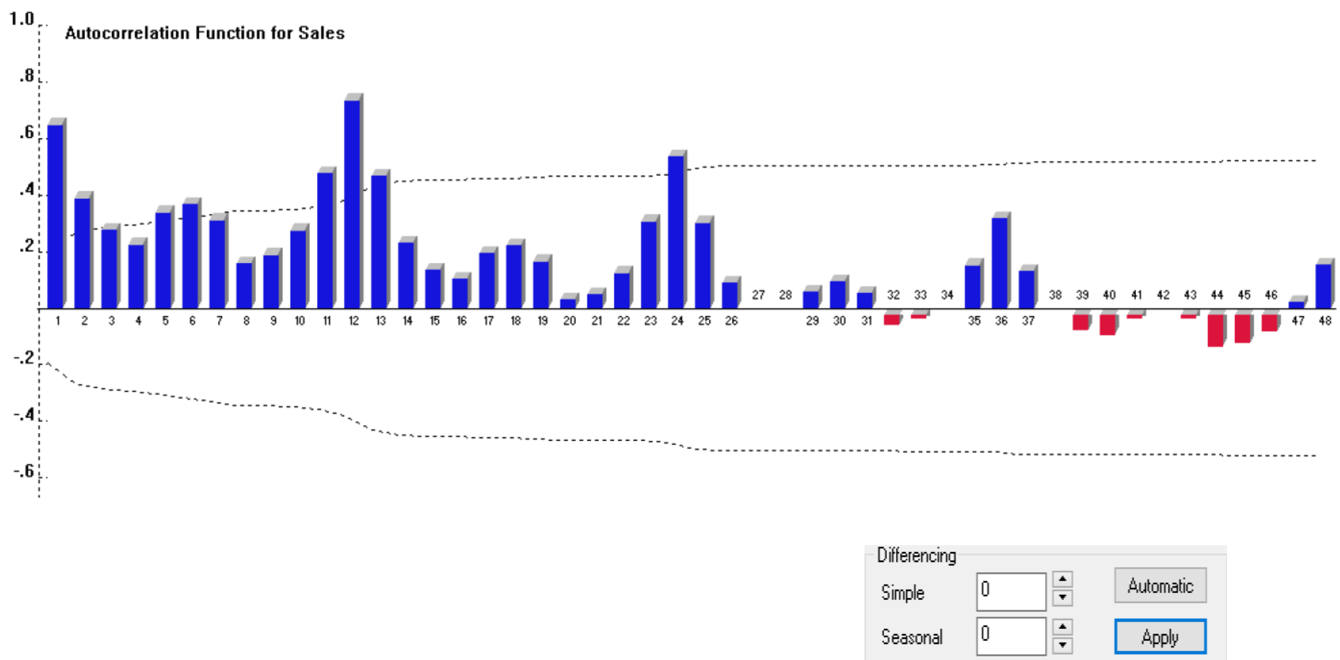

Figure 3: ACF without any differencing

To analyze only the seasonal component in the time series we can make the trend component zero by applying simple differencing (simple = 1) to time series. Figure 4, as shown below, is a graphical representation of first order simple differencing as generated by Forecast Pro. We can clearly see that there are spikes in the time series which are 12 months apart. Specifically, there are spikes in the time series at lags 12, 24, 36 etc. which confirms our earlier analysis regarding the presence of seasonality.

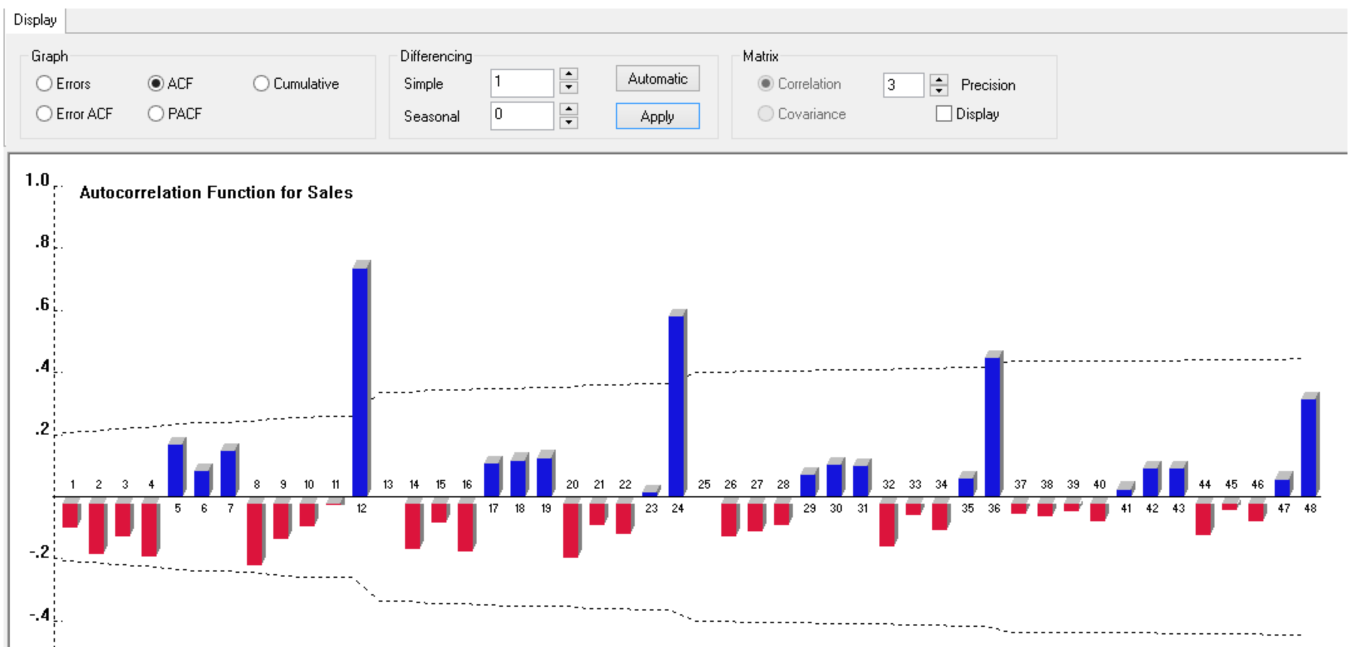

Figure 4: ACF with simple differencing of time series

Furthermore, from figure 3, we can see that successive values of time series for the sales of red wine are correlated to each other. Also, we can see that the values of the time series are gradually declining to zero. This pattern in the time series indicates the presence of trend. However, as stated earlier, ACF is not always conclusive, hence further analysis is required. To check for trends in a time series, we apply seasonal differencing to remove seasonality.

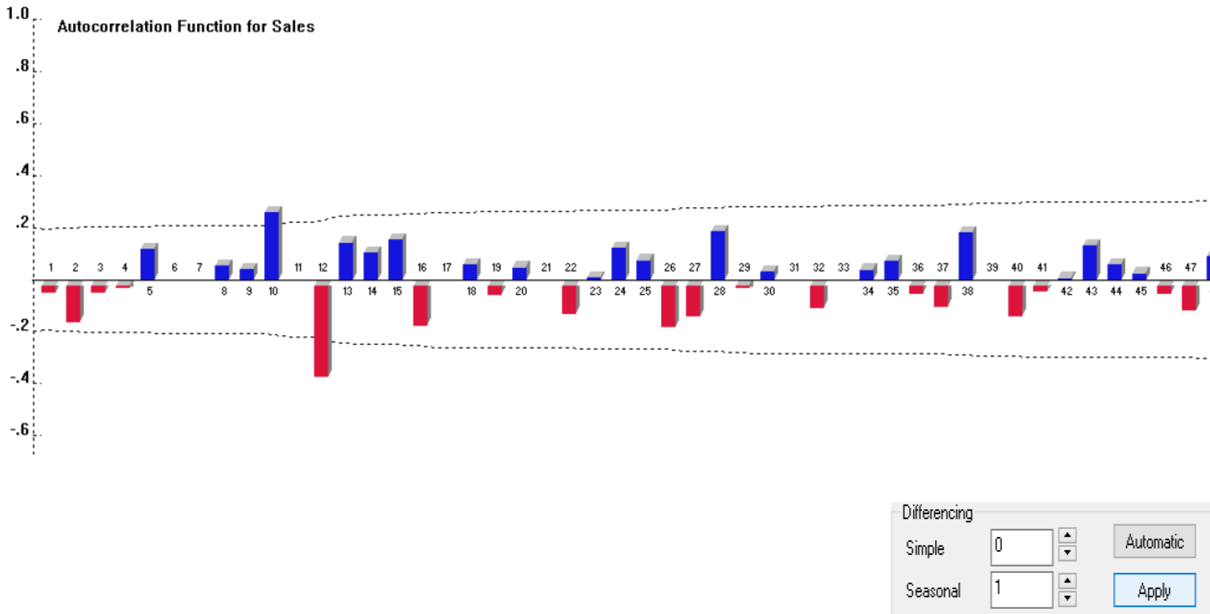

Figure 5: ACF with seasonal differencing of time series

The correlogram in Figure 5 as shown above, is a result of seasonal differencing on the time series. When a trend exists in a time series, the initial autocorrelation coefficients tend to be similar and different from zero but eventually drop to zero. Analyzing these characteristics, we can see from the correlogram that there is not clear indication of trend in the time series.

In the next section, we compare various models and select the model with the best performance.

3. Model Selection

3.1 Objective

In this section, we aim to apply and compare models from the exponential smoothing family. The comparison between the models is based on the MAPE value that each model generates which is used to measure accuracy. Furthermore, we discuss the seasonal indexes generated by the model we select to analyze the variations in demand.

3.2 Exponential Smoothing Family

To study in more detail about the wine sales we will be using exponential smoothing models. The reason for using these models is that they are specifically build to capture features of trend and seasonality.

Any exponential smoothing model is made up of at least level, random events and noise. Level describes the smooth and slow changing non-seasonal part of the time series, random effects are the values that change unpredictably. The other components trend, seasonal indexes and event indexes are optional which capture the presence of trend, seasonality and events that might happen over the course of the time series.

Further, trend within a time series can be classified into 4 types, namely, un trended where forecasts are flat, linear trend where forecasts are extrapolated from the last estimate of the trend, damped trend where trend shows linearity in the early stages but dies off exponentially, and exponential where trends begins linearly but increases as a percentage of itself. Moreover, seasonality in a time series is captured in three ways: none, multiplicative and additive. If indexes are multiplicative, the seasonal adjustments are made as a product of indexes and the de-seasonalized values. Whereas, for additive, the adjustment is made by adding the index into the de-seasonalized values.

All the models which contain the components discussed above come under exponential smoothing family. To compare and choose the best model we consider Mean Absolute Percentage Error (MAPE). MAPE captures the size of the forecast error in terms of percentage. On comparing all the models, we select the model that generates the lowest value of MAPE.

Using Forecast Pro, we ran all the different models included in the exponential smoothing family that are built with different types of trend and seasonality. The MAPE value generated by Forecast Pro was recorded for each of the model as shown in the table in Figure 5.

We can observe at the beginning that MAPE values are very high for both Simple exponential smoothing and Holt’s model. The model which contains only the three basic components is called Simple exponential smoothing (SES). We will not be using this model as the MAPE value is relatively higher. We also eliminate using Holt’s smoothing model because of its high error rate and this model is best suited for data which has trend and is non-seasonal. Holt’s model is slightly better than Simple exponential because it considers trend.

<table>
<tr><th colspan="3">Model</th><th>MAPE</th></tr>
<tr><td rowspan="8">Custom Selection of Exponential Smoothing Models</td><td rowspan="4">Additive Seasonality</td><td>No Trend</td><td>7.12%</td></tr>
<tr><td>Linear Trend</td><td>4.52%</td></tr>
<tr><td>Dampened Trend</td><td>4.55%</td></tr>
<tr><td>Exponential Trend</td><td>4.54%</td></tr>
<tr><td rowspan="4">Multiplicative Seasonality</td><td>No Trend</td><td>6.09%</td></tr>
<tr><td>Linear Trend</td><td>4.51%</td></tr>
<tr><td>Dampened Trend</td><td>4.56%</td></tr>
<tr><td>Exponential Trend</td><td>4.90%</td></tr>
<tr><td colspan="3">Simple Exponential Smoothing</td><td>15.26%</td></tr>
<tr><td colspan="3">Holt's Exponential Smoothing</td><td>13.85%</td></tr>
<tr><td colspan="3">Winter's Exponential Smoothing</td><td>4.51%</td></tr>
</table>

Figure 5: MAPE for all models

Winters exponential smoothing is one of the best suited model for our analysis, as this model is considered for data which have trend and seasonality. On running Winter's exponential model, we can observe that MAPE is lowest at 4.51 as it considers both trend and seasonality. The model forecast pro selected for this is linear trend and multiplicative seasonality. From the rest of the custom models, we can confirm that the lowest MAPE is for Linear trend and Multiplicative seasonality which is the same result that we obtained for winters exponential smoothing. Therefore, after analyzing all the models we decided that Multiplicative Winters exponential smoothing model is best to forecast the wine sales of the next 12 months.

3.3 Seasonal index

From the analysis done so far, we can observe that there exists seasonality in wine sales. To measure these seasonal changes in demand we will be analyzing seasonal indexes which is a measure of seasonal variations. Below table shows us the multiplicative seasonal indexes for every month. These values are extracted from results of the winters exponential smoothing model.

Month	Seasonal Indexes
January	0.8923
February	0.8884
March	0.9428
April	0.9478
May	0.9407
June	0.9554
July	0.9055
August	0.7243
September	0.971
October	1.118
November	1.387
December	1.591

Figure 6: Seasonal Indexes Jan-Dec

Multiplicative seasonal indexes are measured around the base number 1. When the seasonal index is below 1, then sales should be adjusted downwards. Whereas, if the seasonal index is above 1 then sales should be adjusted upwards. We can see that the sales from October to December are more than 1. For example, the seasonal index for December 1.591, which is the highest, which means that the sales must be adjusted upward by 59%. The seasonal index for the month August 0.7243, which is lowest, meaning that the sales should be adjusted downwards by 28%. From the indexes, we observe that wine sales are highest in the months of October, November and December. This can be attributed to the start of holiday season.

The next section we apply the Winter's Exponential Smoothing model to the data to determine the accuracy of the model as well as any production strategies that the company should consider

4. The Forecast

4.1 Objective

In this section, the objective is to produce forecasts from October 1996 to September 1997 based on the past 7 years and 9 month's data of Northern Napa Valley wine sales using Forecast Pro. We will then use different confidence levels to understand the range of sales.

4.2 Forecast Using Winter's Exponential Smoothing Model

After confirming the best model from the previous section as Multiplicative Winter's exponential smoothing model, we then forecasted the red table wine sales from October 1996 to September 1997. Figure 7 as shown below is a graphical representation of forecast sales generated on Forecast Pro. The historical data is represented in green, forecasted values in red for both in-sample and out-of-sample periods while the blue represents the upper and lower confidence levels. As can be seen graphically, Winter's exponential smoothing model has a good fit on the in-sample data with small deviations.

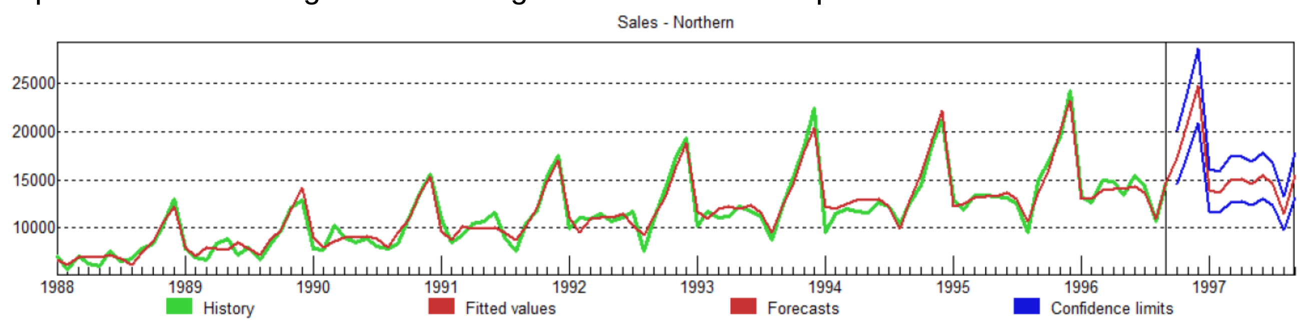

Figure 7: Forecast Using Winter's Exponential Smoothing Method

Further, we analyze the forecast of sales for the out-of-sample time period of October 1996 to September 1997 using a confidence level of 95%. Forecast Pro uses confidence intervals of 97.5 as upper limit and 2.5 as lower limit. This is also known as 95%(97.5-2.5) Confidence Interval. It means that the probability of future sales falling between the upper and lower limit is 0.95. These values as shown below in Figure 8 will be a good indicator of how much grapes the company should consider selling to TransContinental.

Date	Forecast Sales	95% Confidence Interval		97.5% Confidence Interval		99% Confidence Interval	
		Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit
1996-Oct	17,152	14,479	20,318.54	14,131	20,819	13,728	21,430
1996-Nov	21,281	17,956	25,221.39	17,523	25,844	17,023	26,604
1996-Dec	25,010	21,093	29,655.25	20,583	30,390	19,994	31,286
1997-Jan	14,673	12,369	17,406.14	12,070	17,838	11,723	18,366
1997-Feb	13,656	11,506	16,206.63	11,227	16,610	10,904	17,103
1997-Mar	15,642	13,174	18,572.23	12,853	19,036	12,482	19,602
1997-Apr	15,584	13,120	18,512.24	12,799	18,976	12,429	19,541
1997-May	14,893	12,532	17,698.87	12,225	18,143	11,870	18,685
1997-Jun	15,650	13,163	18,606.18	12,840	19,074	12,466	19,646
1997-Jul	14,729	12,383	17,519.55	12,078	17,961	11,726	18,501
1997-Aug	11,121	9,345	13,233.22	9,115	13,568	8,848	13,977
1997-Sep	16,516	13,873	19,661.72	13,530	20,160	13,133	20,769
Total	**195,907**	**164,994**	**232,612**	**160,975**	**238,419**	**156,326**	**245,509**

Figure 8: Comparison of Forecast Values of Sales with Different Confidence Intervals

The table above shows the forecasted values for the period October 1996 to September 1997 are 195,907 cases. The variation in the sales for each month is quite considerable and can be attributed to the seasonal nature and high sales in the holiday period starting October to December. The upper and lower limit are represented by a 95% confidence interval for each forecast. This means that if the company is producing 232,612 cases(upper confidence limit) in total, it means that the company will not be going out of stock 97.5 percent of the time and vice versa for lower limit. These upper and lower limits are crucial as it will help determine the strategy that will be implemented and how many grapes should be sold to TransContinental.

Further, to explore the range of sales for next 1 year, the forecasting was carried out with different confidence levels of 97.5% and 99% as shown in Figure 8 above. For 97.5% confidence level the upper limit was set to 98.75% and lower limit to 1.25%. Similarly, for 99% confidence level the upper limit was set to 99.5% and the lower limit to 0.5%. It can be observed that with increase in confidence levels the values of upper limit also increases and the values of lower limit decreases. This means that by increasing the confidence level, the variation in the forecast also increases by a considerable amount which can create production challenges. However, with a MAPE of 4.51 for the holdout period as generated by Winter's exponential smoothing model, we can say that on average the sales absolute forecasted error were 4.51% of the actual sales data. This is good indicator of the accuracy of the forecast model which was also confirmed above visually in Figure 7.

We also see from the Figure 9, as shown below, the smoothing parameters as generated by Forecast Pro. A seasonal weight of 0.2800 is an average indicator as to how seasonality is affecting the time series. However, the trend smoothing parameter of 0.15 indicates that trend does not have a major influence on the time series and hence the lower value of Level.

User Defined \WINTERS			
Multiplicative Winters: Linear trend, Multiplicative seasonality			
LM(0.030, 0.159, 0.280)			
Confidence limits proportional to indexes			

Component	Smoothing Wgt		Final Value
Level	0.03034		14,158
Trend	0.1589		57.59
Seasonal	0.2800		
Seasonal Indexes			
Jan - Mar	0.8923	0.8884	0.9428
Apr - Jun	0.9478	0.9407	0.9554
Jul - Sep	0.9055	0.7243	0.9710
Oct - Dec	1.118	1.387	1.591

Figure 9: Smoothing parameters

5. Conclusion and Recommendations

Patricia Quintana, the CEO, and shareholder at Northern Napa Valley Winery have an important and difficult decision to make. To consider an offer from TransContinental to buy the leftover grapes from their harvest, Ms. Quintana must understand and consider the sales they will make for their table wine. Our forecast shows that the average sales for the period October 1996 to Sept 1997 will be 195,907 using a 95% confidence level. Selling grapes to TransContinental highly depends on the strategy the CEO and shareholders want to adopt at Northern Napa Valley Winery. We would like to recommend either one of the two strategies as mentioned next. One strategy would be to produce at the lower limit of the forecast at 95% confidence level throughout the year with total forecasted sales of 164,994. Adopting this strategy would mean there is only a 2.5% chance that the total sales will be below that value and costs will be saved on wastage. Also, the company can sell the excessive grapes to TransContinental rather than overproducing. However, an important thing to consider here is the cost of lost sales by producing at the lower limit. A cost analysis should be conducted to see if more revenues can be brought in by producing more wine or by selling more grapes to TransContinental and avoiding wastage. The second strategy we recommend is to produce variably. In this strategy, the company would be producing at the forecasted sales for the months with high sales such as October, November and December. And then produce at the lower limits for the rest of the year. This would allow company to push the sales in the holiday period of October to December. Any leftover wine from this period can be sold in the following months using promotions. Moreover, the company can also make money selling the leftover grapes to TransContinental. A cost analysis of this strategy would provide a clearer picture and help grow revenues.