MSc in Engineering | TSM Analysis of Sequential Data

Project on Time Series

Davide Nasi, Nicolò Rubattu

### **EXPLORATIVE ANALYSIS**

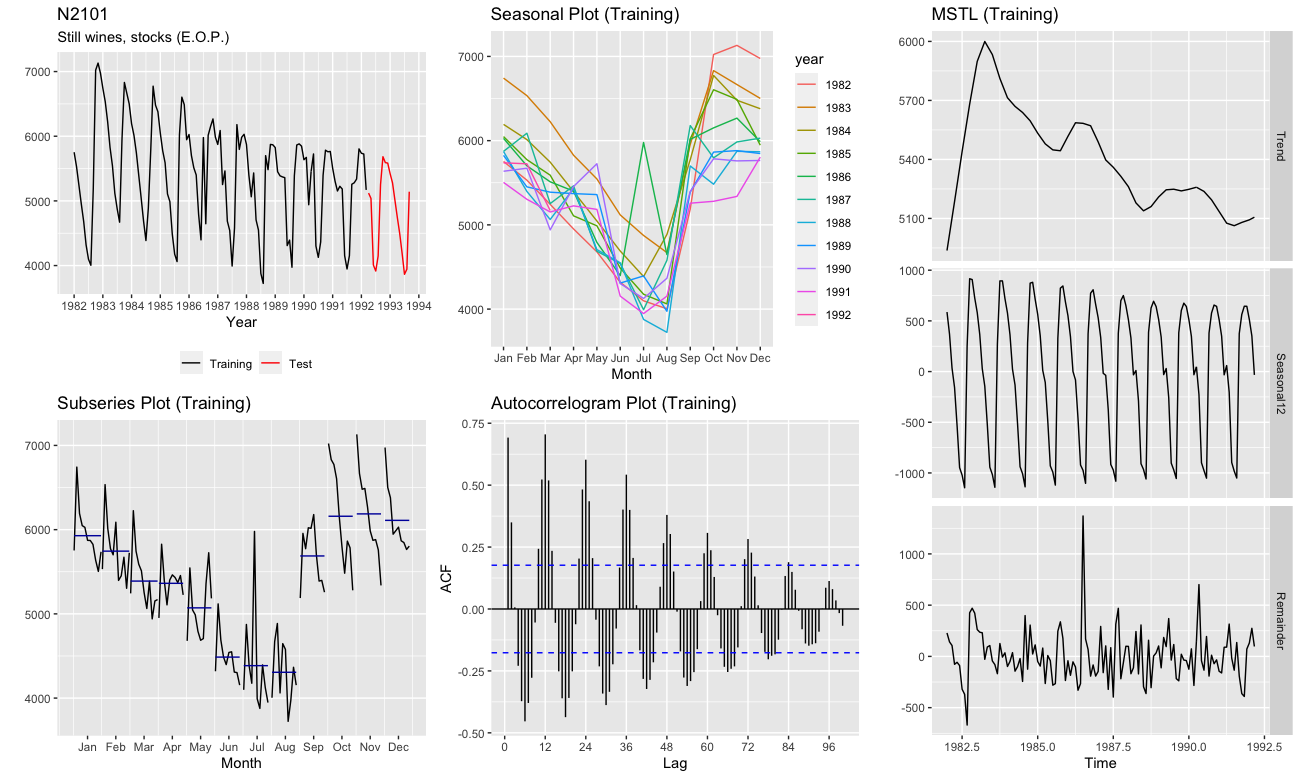
N2101 from the M3 competition is the selected time series for this study. It is a monthly (f=12) time series with 123 available observation for training and 18 for testing. It comes from industry sector and describe the stocks of still wines.  
  


Figure 1 - EDA

Plotting the time series a strong seasonal pattern is observable. Also a moderate negative trend is observable from 1983 to 1989, then the fluctuations stabilize around a constant value. 1982 was probably part of an uptrend which breaks in 1983 but a single year of observation is not enough to define it. With the seasonal plot we can observe peaks corresponding to winter seasons and lows to summer seasons. Two ‘outliers’ can be detected at July 1986 and May 1990. From the subseries plot we can observe that August is the global minimum (averaged over the year, i.e. horizontal blue line) and November the global maximum. Subseries plot also shows negative trends for almost all the subperiods. The autocorrelogram confirms previous statement of seasonality: a strong seasonal feature component is observable with statistically significant positive spikes at lags (n\*12) ±2 (i.e. multiples of seasonal frequency), alternated with negative spikes of same absolute amplitude. Up to lag 84 there are statistically significant correlations. The amplitudes decrease due to the trend. Seasonal and trend features were also observed individually using automatic time series decomposition.

### **INDICATORS OF FORECASTING PERFORMACE**

A scale independent metric is not relevant because we don’t have to calculate it across multiple time series. We opt for Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). MAE is popular as it is easy to both understand and compute. It measures the average magnitude of the errors, with equal weight, in units of the variable under study, with lower values that are better. RMSE, compared to MAE, has the benefit of penalizing large errors (high weight to large errors) that we don’t want. It is in units of the variable under study, with lower values that are better.

#TODO set targets

### **SIMPLE MODEL**

Among the baseline forecasting methods Seasonal Naïve is the most appropriate for our time series because of the seasonal component. Drift (random walk with drift) method could also be appropriate: it’s expected to perform better than Naïve (random walk) because it takes into account the negative trend, and better than Mean method too. Anyway, because the seasonal component is visually stronger respect to trend component, Seasonal Naïve is with few doubts the best choice.

|  |  |
| --- | --- |
|  |  |
|  |  |

Figure 2 - Simple models and Seasonal Naïve

**Analysis of residuals** From the autocorrelogram we can see that most of the peaks lie between the upper and lower significant limits. Exceptions are lag 1-4, 11-12, 36… where correlation is relevant. Formally, with the Ljung-Box test we *reject the white noise null hypothesis* (p-value=8.529e-08 < 0.05). Residuals distributions has mean=-35 (slightly overestimating overall). In 1983 we underestimate, from 1984 we overestimate due to the negative trend which Seasonal Naïve doesn’t model. From the histogram and the lineplot we can see the two outliers previously identified; since their sign is opposite with equal magnitude, their influence is balanced.

#TODO comments on MAE, RMSE and target

### **EXPONENTIAL SMOOTHING**

In the family of exponential smoothing models, we considered Holt-Winters’ method with seasonal component extension the one which is best appropriate for our time series because it models both seasonality and trend. We opt for the multiplicative variation of this model because seasonality has a slightly decreasing amplitude over time (Figure 1, seasonal component in the time series decomposition plot). Simple Exponential Smoothing (SES) and Holt’s methods didn’t model seasonality. For instance, fitting SES results with a model with α~1 (which is actually a RW), fitting Holt’s linear method results with a model with smoothing parameters α~1 and β\*~0 (slope doesn’t change over time, and actually a RW with a linear trend).

|  |  |  |
| --- | --- | --- |
|  |  | |
|  | |  |

Figure 3 - Holt-Winters seasonal multiplicative method

Estimated smoothing parameters are: α=0.2027, β\*=0.0073, γ=1e-04. α is small, so more weight is given to observations from the more distant past. The small value of β∗ means the slope component hardly changes over time. The small value of γ for the multiplicative model means that the seasonal component hardly changes over time. Indeed, the seasonal decreasing amplitude over time is little or irrelevant.

**Analysis of residuals** From the autocorrelogram we can see that most of the peaks lie between the upper and lower significant limits apart from two isolate lags. Formally, with the Ljung-Box test we again *reject the white noise null hypothesis* (p-value=5.874e-04 < 0.05), i.e. there is correlation. Distribution of residuals has mean=0 and is slightly asymmetrical (positive skewness).

With the performance (MAE and RMSE) on the test set a solid improvement is observable respect to Seasonal Naïve baseline model. Looking at the forecasting plot, it is nice to observe that the mean point forecasts of the model are visually close to actual testing values, more than in previous models. Confidence intervals are visually reliable too.

### **ETS**

|  |  |  |
| --- | --- | --- |
|  |  | |
|  | |  |

Figure 4 – ETS

The ETS function returns the best exponential smoothing model which is the one that minimize AIC. In our case it returns ETS(A,N,A): additive error, none trend and additive seasonality. Smoothing parameters are: α=0.3879, γ=1e-04.

#TODO comments…