CentraleSupélec

Inventory and Demand Planning

Analysis report for the Duvel brewery

[redacted] in Belgium

January 2025

I - Introduction

As part of the inventory management course, Duvel asked us to look at their demand forecasting and stock management methods.
Duvel, a major player in the Belgian beer market, is faced with a dilemma concerning the effectiveness of its statistical forecasting engine compared with the forecasts made by its planners. The aim of this report is to analyze the impact of these two approaches on stock management for [redacted] on the Belgian market.

The study focuses on the period from week 26 of 2023 to week 52 of 2024. This period includes weeks 36 to 38 of 2023, when data logging problems were identified, making the data unavailable for analysis. The main objective is to determine which forecasting method would make it possible to optimize stock levels, minimizing both overstocking costs and the risk of stock-outs.
Based on quantitative and qualitative analyses, this report will provide recommendations on the best stock policy for Duvel, considering the specificities of each forecasting method and their potential impact on the company's performance.

II - Qualitative data analysis

Figure 1 illustrates the discrepancies between forecasts and actual sales of [redacted] in Belgium. Historical sales (blue) show high volatility, with marked oscillations between periods, including a major peak around week 2024-W16. In contrast, statistical forecasts (orange), although showing a smoother trend, do not incorporate abrupt variations, suggesting an underestimation of fluctuations. The consensus forecasts (green) more closely follow the general trends in actual sales, particularly after the week 2024-W16, although they also struggle to capture the extremes. Weeks 36 to 38 of 2023 show anomalies, probably due to missing data. The data range is too short to identify any seasonality. This discrepancy between forecasts and actual sales highlights the need for a thorough evaluation to determine the most effective approach to stock management.

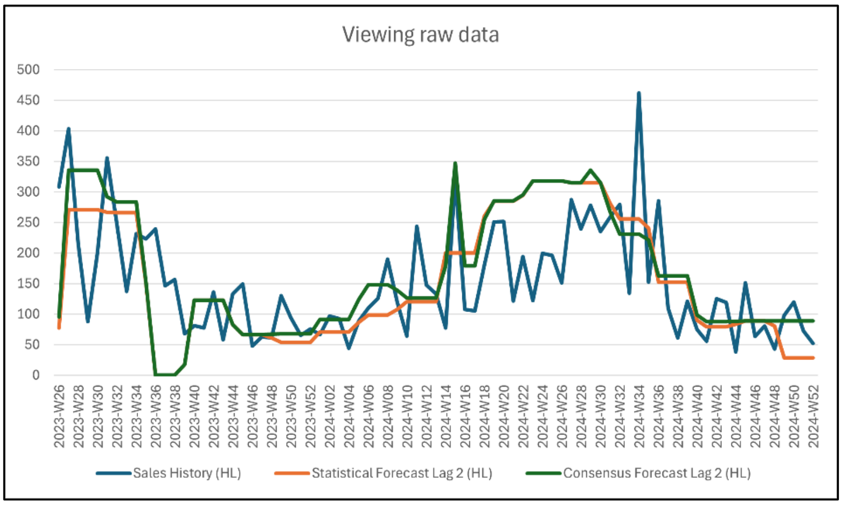

Figure 1 : Graphical representation of sales history, statistical forecasts and consensus forecasts over the analysis period

III – Choice of KPI's

In the rest of our analysis, we will focus on choosing the best method for 'completing' and processing the missing data for weeks W36, W37 and W38. In order to be able to compare forecasts quantitatively between methods, we need to introduce performance metrics (KPI's). For this study, I consider the following KPI's:

- Bias: It measures whether our model tends to systematically overestimate or underestimate demand, which has a direct impact on stock decisions.

- MAE: It measures the average of the absolute differences between forecasts and actual sales. It is a good, simple indicator for quantifying the overall scale of errors, regardless of their direction (over- or under-estimation).

- RMSE: It measures the overall accuracy of our model by quantifying the average squared deviation between forecasts and actual values. In our case, for heretical demand, it is essential to have a metric that strongly penalizes major deviations, which can have a significant impact on stock management.

We will therefore define our performance indicator by the following score:

$$\text{Score} = |\text{Bias}| + \text{MAE} + \text{RMSE}$$

IV – Treatment of missing data and quantitative analysis of the two forecasting models

a) Treatment of missing data

Our client Duvel has informed us of a problem with the generation of data for weeks W36, W37 and W38 in 2023. To consolidate the forecasts, we will test several correction methods, which will be compared using the performance indicators defined previously. The methods used are explained below :

1- W36, W37, W38 forecasted at 0
2- W36, W37, W38 excluded
3- 3-month rolling average
4- 5-month rolling average
5- Weighting of the last 3 months: $0{,}5 \times W35 + 0{,}3 \times W34 + 0{,}2 \times W32$
6- Weighting of the last 5 months: $0{,}4 \times W35 + 0{,}2 \times W34 + 0{,}2 \times W33 + 0{,}1 \times W32 + 0{,}1 \times W31$

The results are shown in Figure 2 below:

	Statistical Forecast			Consensus Forecast		
Correction method	**Bias**	**MAE**	**RMSE**	**Bias**	**MAE**	**RMSE**
W36, W37, W38 forecasted at 0 (baseline)	1,8%	42,0%	56,4%	10,7%	40,7%	56,2%
W36, W37, W38 excluded	6,6%	42,8%	54,9%	15,8%	41,3%	54,6%
3-month rolling average	7,2%	38,7%	51,9%	16,3%	37,5%	51,8%
5-month rolling average	7,8%	39,1%	52,3%	17,0%	38,2%	52,3%
Weighting of the last 3 months	6,9%	38,8%	51,9%	16,0%	37,5%	51,7%
Weighting of the last 5 months	7,2%	39,0%	52,1%	16,3%	37,8%	51,9%

Figure 2: Kpi's for each forecast according to the method used to correct for missing values.

Interpreting the results:

1- W36, W37, W38 forecasted at 0: This method has a low bias (1.8% for the Statistical Forecast, 10.7% for the Consensus Forecast), indicating a balanced forecast. However, the large errors (MAE of 42.0% and RMSE of 56.4% for the Statistical Forecast, MAE of 40.7% and RMSE of 56.2% for the Consensus Forecast) reflect low accuracy.

2- Exclusion of weeks W36, W37 and W38: This slightly improves the RMSE (54.9% for the Statistical Forecast, 54.6% for the Consensus Forecast) compared with the forecast at 0, but the bias increases significantly (6.6% and 15.8% respectively). The MAE remains high, which limits the effectiveness of this approach.

3- 3- month rolling average: This offers a significant improvement in forecast accuracy. With an MAE of 38.7% and an RMSE of 51.9% for the Statistical Forecast, this method reduces errors compared with previous approaches. The results for the Consensus Forecast also perform well, with an MAE of 37.5% and an RMSE of 51.8%. Nevertheless, the bias is moderate (7.2% for the Statistical Forecast and 16.3% for the Consensus Forecast), indicating a slight tendency to underestimate values. Overall, this method offers a good balance between simplicity and accuracy.

4- 5-month rolling average: The results are comparable to those of the 3-month moving average (MAE of 39.1% and RMSE of 52.3% for the Statistical Forecast, MAE of 38.2% and identical RMSE for the Consensus Forecast). However, the bias is slightly higher (7.8% and 17.0%), which makes it less attractive than the 3-month moving average.

5- Weighting of the last 3 months: This method performed very well overall, with an MAE of 38.8% and an RMSE of 51.9% for the Statistical Forecast. For the Consensus Forecast, the MAE is 37.5%, and the RMSE is reduced to 51.7%, the lowest of all the approaches evaluated. The bias remains moderate, at 6.9% for the Statistical Forecast and 16.0% for the Consensus Forecast. This method shows better management of large errors, making it a relevant alternative to the 3-month moving average.

6- Weighting of the last 5 months: Performance is like that of the 3-month weighting (MAE of 39.0% and RMSE of 52.1% for the Statistical Forecast, MAE of 37.8% and RMSE of 51.9% for the Consensus Forecast). The slightly higher bias (7.2% and 16.3%) makes it less competitive.

We will therefore use the ***Weighting of the last 3 months method*** to correct the values for weeks W36, W37 and W38.

b) Quantitative analysis of the two forecast models

	Statistical Forecast			Consensus Forecast		
Correction method	**Bias**	**MAE**	**RMSE**	**Bias**	**MAE**	**RMSE**
Weighting of the last 3 months	6,9%	38,8%	51,9%	16,0%	37,5%	51,7%

Figure 3: Kpi's for each forecast with the correction method used

In Figure 3, the statistical forecast shows a bias of 6.9%, reflecting greater neutrality in the forecasts. In comparison, the consensus forecast has a significantly higher bias of 16.0%, indicating a tendency to over forecast. This high bias can lead to inappropriate decisions and lead to high stock costs. The MAE is slightly lower for the consensus forecast (37.5%) than for the statistical forecast (38.8%). This shows that the consensus, by incorporating human adjustments, slightly reduces the average absolute differences between forecasts and actual values. This improvement is marginal but may be useful in contexts requiring greater day-to-day accuracy. The consensus forecast also has a lower RMSE (51.7%) than the statistical forecast (51.9%). Although this difference is small, it indicates that the consensus forecast is better at handling large or extreme errors, which can limit the risks associated with unforeseen variations in demand.

The choice of forecast therefore depends on Duvel's priority:

- The statistical forecast is preferable if the objective is to guarantee neutral and automated forecasts, as it minimizes the bias to 6.9%.

- The consensus forecast is recommended if the objective is to reduce absolute deviations and significant errors (MAE and RMSE), particularly in contexts where human expertise provides real added value.

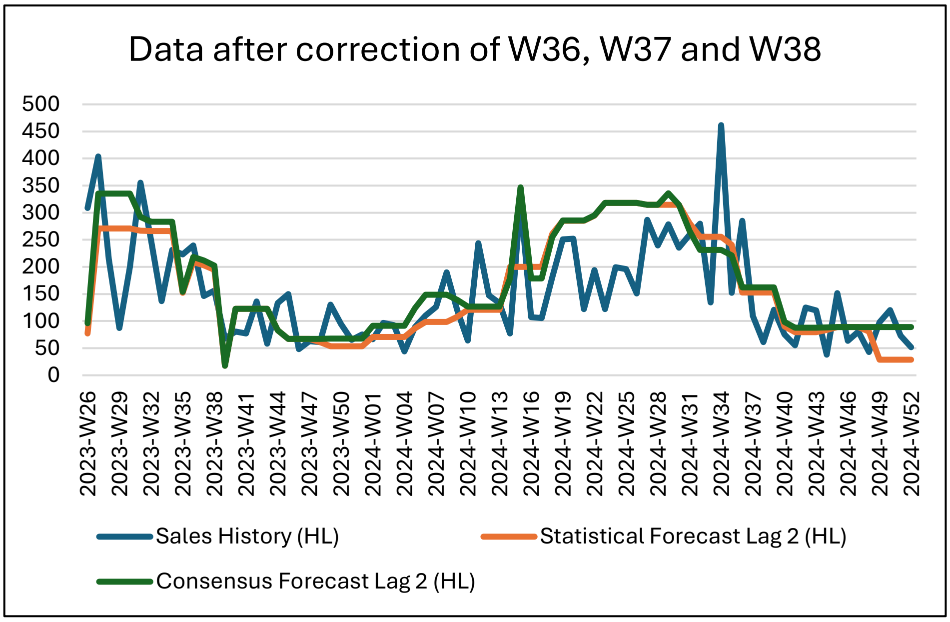

Figure 4 :Data after correction of W36, W37 and W38

V - Identification, analysis and selection of the appropriate inventory management model

We concluded in the previous part of the analysis that the use of each model depends on Duvel's priorities. Thus, in this section we will determine the ideal inventory management policy for each of the two forecasts (statistical and consensus).
However, Duvel does not provide us with any information concerning shortage costs, holding costs or any other information specific to the company. It is therefore necessary to introduce hypothetical values for the holding cost per unit and the shortage cost per unit. These values, although theoretical, are defined based on realistic assumptions adapted to the brewing sector. The holding cost is set at *0.20€ per pack,* reflecting the costs associated with storage, refrigeration and obsolescence. The lost sales cost is estimated at *50.00€ per pack*, including lost margin, customer dissatisfaction and missed opportunities.

We assume that we will start with *an initial stock of 350 packs*.
We also notice that the lead time (L) is 3 period, and the review period (R) is equal to 1. So, ***the risk horizon*** is equal to 4 period. We will consider the lead time in the calculation of the level of the ***Up to Level value*** in the 3 different inventory management models.

For each of the two models we will test the following inventory management models:

1- Fixed safety stock policy.
2- Dynamic safety stock policy (Up to level: $F_T + F_{T+1} + F_{T+2} + F_{T+3} + Safety\ Stock$).
3- Policy with the usual safety stock formula. We use the Cycle Service Level and the risk horizon in the formula: $\alpha = z \times \sigma \times \sqrt{L + R}$

In the 3 stock policies we take backlog into account. So, to each forecast for period ***T,*** we add backlog from period ***T-1***. (This represents the additional deliveries to be made in period T to rectify the shortage in period T-1.)

a) Choice of service level metrics

Initially, we will consider 3 different service level indicators. We will then analyze their relevance to our study.
Consider the following 3 indicators:

- **Period Service Level:** Fraction of periods (a day, a week, a month) without shortages (= inventory on-hand > 0). (Often called on-shelf availability)
- **Cycle Service Level:** The probability of not stocking out during an order replenishment cycle.
- **Fill Rate:** Fraction of items directly supplied from on hand inventory.

The Period Service Level offers a limited analysis here. Although it measures the proportion of periods without interruptions, it does not distinguish between minor interruptions during periods of low demand (W48) and major interruptions during peaks. This simplification minimizes the impact of critical periods, which reduces its relevance in such a variable context.

Cycle Service Level is of little relevance in the face of highly fluctuating demand, with peaks more than 300 HL (W28-W35) and troughs (<100 HL). It is based on a binary analysis (success or failure per cycle) and does not measure the extent of breakages. As a result, it underestimates the impact of periods of high demand, such as W16, when breakages can be critical.

The Fill Rate is best suited to this demand profile. It accurately quantifies the percentage of total demand satisfied, regardless of fluctuations. For example, it assesses the impact of a break during a peak (W16 to 400 HL) while considering calmer periods (W46-W50). This granularity enables better performance management in volatile environments.
We will therefore only consider the **fill rate** as an indicator of service level.

We will identify the parameters of the models with the following abbreviations:

- Safety Stock Coverage: SSC
- Fill Rate: β
- Total Cost (Holding Cost + Shortage cost): TC

a) Statistical forecast

The optimum parameters were optimized using Excel solver software with the aim of minimizing total costs.

Inventory Policy	Optimal parameters	Values
Fixed safety stock policy	Safety Stock = 77,9	β = 95,2 % TC = 35 939 €
Dynamic safety stock policy	SSC = 1,1	β = 95,2 % TC = 37 970 €
Policy with the usual safety stock formula	z = 2,24	β = 95,2 % TC = 40 160 €

The most effective inventory management model, combined with statistical forecasting, is based on a ***fixed inventory management policy*** (with a Safety Stock equals to 77,9). This policy enables us to obtain the lowest total cost for an equal level of service

b) Consensus forecast

The optimum parameters were optimized using Excel solver software with the aim of minimizing total costs.

Inventory Policy	Optimal parameters	Values
Fixed safety stock policy	Safety Stock = 86,3	β = 95,2 % TC = 37 325 €
Dynamic safety stock policy	SSC = 0,19	β = 95,2 % TC = 36 686 €
Policy with the usual safety stock formula	z = 1,89	β = 95,2 % TC = 41 435 €

The most effective inventory management model, combined with consensus forecasting, is based on a ***Dynamic safety stock policy*** (with an SSC equals to 0,19). This policy enables us to obtain the lowest total cost for an equal level of service.

VI - Conclusion

We have identified the optimal inventory management policy for each of the two forecasts. In this way, depending on the operational constraints it wishes to prioritize, Duvel will be able to select the most appropriate forecast type and inventory management policy:

- If the priority is to guarantee a neutral forecast, with a low bias, the analysis shows that Duvel should opt for the statistical forecast model coupled with a fixed safety stock policy.

- If the priority is to reduce significant errors, the analysis shows that Duvel should move towards the consensus forecast model coupled with a Dynamic safety stock policy.