MBD, Impact Project

Demand Forecasting for ACME Textile Corp.

submitted to

Dr. Federico Castanedo
Data Scientist and Professor
Madrid, Spain

11th of July, 2023

Powered by

By
Daniel Sebastián de la Torre
Rita Otvalho
Laufey Sverrisdóttir
Lucas Trenzado Pérez
Vignesh Viswanathan Nambiar

"I hereby certify that this report and the accompanying presentation is my own original work in its entirety, unless where indicated and referenced."

Introduction & Problem Context

This capstone project aims to address the critical issue of demand forecasting inefficiency within ACME TEXTILE, a specialised company engaged in the production and sale of clothing for service professionals across various sectors. Currently, the demand forecasting process heavily relies on subjective assessments made by sales representatives, who predict the end-client product demand for the subsequent two months based on their experience and customer knowledge. However, the accuracy of these forecasts has notably declined due to the ever-changing market dynamics experienced throughout 2022.

The inadequacy of the existing forecasting method, which lacks adaptability to volatile market conditions, has resulted in significant error rates, subsequently leading to substantial financial inefficiencies. This is primarily manifested through overproduction or inadequate stock levels, causing undue strain on the company's production and inventory management processes. Consequently, ACME TEXTILE has suffered substantial losses and encountered cash flow challenges.

To mitigate these challenges and enable data-driven decision-making, a robust analytical model must be developed. The primary objective of this capstone project is to leverage historical sales data, including relevant data from the past years to enhance the accuracy of demand forecasts. The model will provide predictions for the subsequent two months, aiming to minimise the disparity between predicted and actual sales figures. By achieving greater precision in forecasting, the model will facilitate optimised production planning, efficient inventory management, and cost savings across the entire supply chain.

As the newly appointed team of data scientists, the responsibility lies in designing and implementing an advanced forecasting model that surpasses the existing average accuracy rate of 80% for all products. The successful execution of this model will empower ACME TEXTILE to make informed decisions, promptly adjust production and inventory levels in response to market demands, and ensure long-term financial stability. This capstone project carries substantial significance, as it holds the potential to reshape the future trajectory of the company by leveraging data-driven insights for enhanced operational efficiency and profitability.

The Data

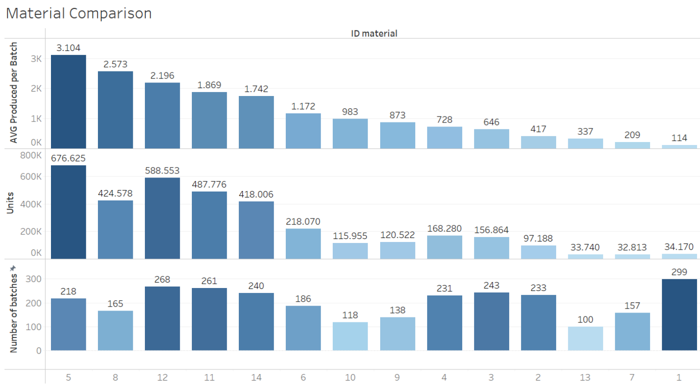
Time series data refers to a sequence of observations collected over a period of time. When handling such data the objective is to find patterns that offer an understanding of what the future state of the data could be. More specifically capture the temporal dependencies present in the data.

The provided time series dataset consists of three primary columns: “year-week”, “ID_material”, and “Units”. The first column, “year-week”, is a date type feature that ranges from April 2018 to October 2022 and it's formatted as YYYY/WW. As for “ID_material” this column features the numerical ID of 14 different products. To enhance the efficiency of data analysis, a numerical labelling system was implemented to represent the materials, replacing the use of their respective ID numbers. The assigned numerical labels are as follows: Material 1 (120004096), Material 2 (120009814), Material 3 (120009816), Material 4 (120010342), Material 5 (120010566), Material 6 (120010970), Material 7 (120011556), Material 8

(120011782), Material 9 (120012154), Material 10 (120012606), Material 11 (120014486), Material 12 (120014488), Material 13 (120015842), and Material 14 (120015996). Finally, “Units” is an integer data type column. Ranging from 0 to 8316 this feature represents the amount sold in a given week for each material.

As previously mentioned, this data can provide invaluable insight. ACME textiles should be able to forecast future values, and prepare their

Exploratory data analysis



To begin visualising the distribution of the data points the fluctuation of demand of each material was plotted over the given time period. This is done in order to better understand the fluctuations of demand. Discover patterns in the demand such as cyclical or seasonal, or if demand has an upward or a downward trend.

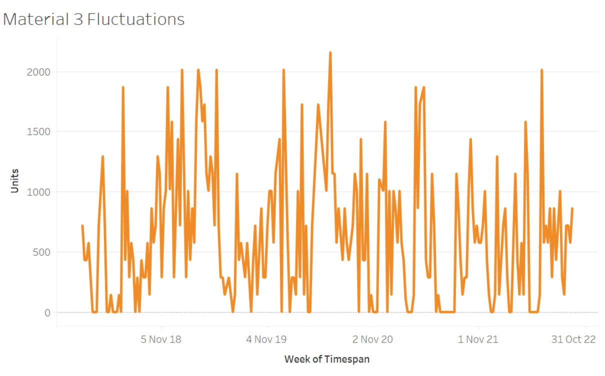
It’s important to note that the start dates of production for each material vary and so do the minimum and maximum units produced of each. Hence for each visualisation the scales change. To define outliers a box plot was created for each distribution. Following are short summaries for distributions of all materials, with some examples of time series are displayed.

Material 1
Sales of material 1 ran from the second week of 2017 to the end of the given time period or the 39th week of 2022, with a few weeks where units produced were equal to 0. The distribution of material one can be summarised in a mean 114, standard deviation of 88, a minimum of 0 and a maximum of 300. Material 1 is bought in 12 distinct batches that vary from 0 to 12. It seems that material one has a stable demand and has no outliers, one peak can be observed early in the year of 2021.

Material 2
Production of material 2 ran from the 16th week of 2018 to the 39 of 2022 with a few weeks where units produced were equal to 0. Material one has a mean of 417, a standard deviation of 313, a minimum of 0 and a maximum of 1404. Material 2 has one outlier which is the maximum. An increase in average sales can be observed in the year of 2021.

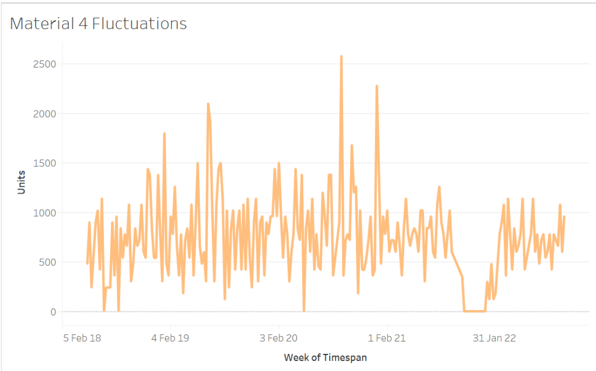
Material 3
Production of material 3 ran from the 6th week of 2018 to the end of the given time period, with a few weeks where units produced were equal to 0. Material 3 has a mean of 646 and a standard deviation of 552 it has a minimum of 0 and a maximum of 2160. Material 3 is sold at a greater scale than the previous two materials and fluctuates more extremely. It does not possess any outliers. On two occasions sales of material 3 dropped to zero for a prolonged

period of time, although no indication of decline in sales could be seen in the earlier weeks. This happens for 8 weeks from the 26th week of 2021 till the 33rd week and for 5 weeks from the 18th week of 2022 til the 22nd week that same year.



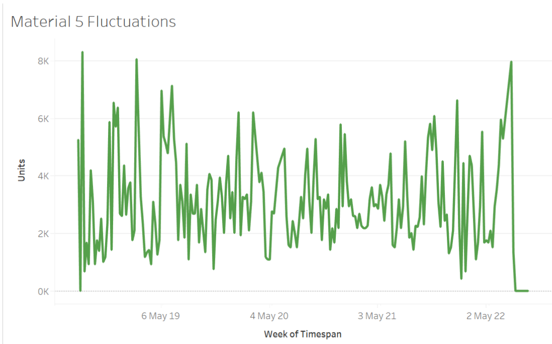
Material 4

Sales of material 4 ran from the 18th week of 2018 to the end of the given time period, with a few weeks where units produced were equal to 0. Data for unit sales of material 4 has a mean of 729 and a standard deviation of 408. It has a minimum point of 0 and a maximum of 2580. Material 4 has four outliers. Sales notably fall to zero for 11 weeks from the 43rd week of 2021 till the 1st week of 2022, this is the only instance where sales fall to 0 for a prolonged period of time for material 4 in the given time period. The maximum point of 2580 is quite a significant outlier.

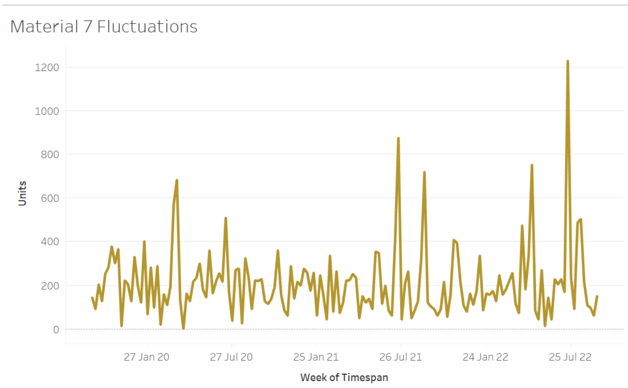


Material 5

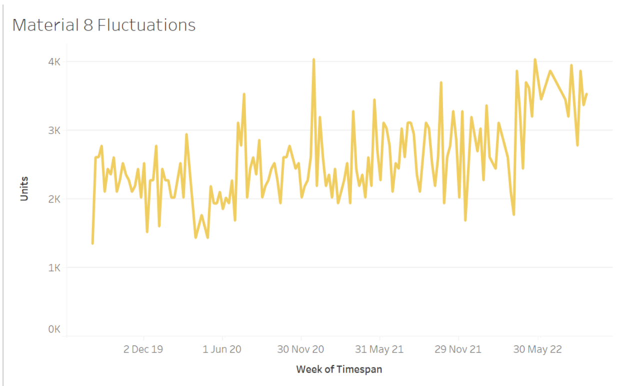
Sales of material 5 ran from the 31st week of 2018. At the 33rd week of 2022 production falls to zero and stays at zero till the end of the given time period. For the purpose of the prediction two possible scenarios can be considered. Either the assumption will be made that material 5 seized production and no prediction will be made. Or the last 8 weeks for which there is data available will be used for the prediction. These alternative scenarios will be evaluated at the model development stage.   h a few weeks where units produced were equal to 0.



Material 6
Sales of material 6 ran from the 11th week of 2019 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 6 strands at 1172 and the data has a standard deviation of 740. It has a minimum point of 0 and a maximum point of 3660. The distribution has four outliers, which can be observed in peaks which happen every summer in the given time period, similarly lows are observed around December every year. Sales of material 6 exhibit a cyclical pattern. Sales of material 6 never fall to zero for a prolonged period of time.
Material 7
Sales of material 7 ran from the 40th week of 2019 to the end of the given time period 2022 and had none of the weeks where units sold was 0. It can be seen towards 2021 and above the supply peaks are significantly higher than previous years, also it can be noted that this product follows an increasing average Units sold towards 2022. Peaks are observed in certain weeks of July every year.

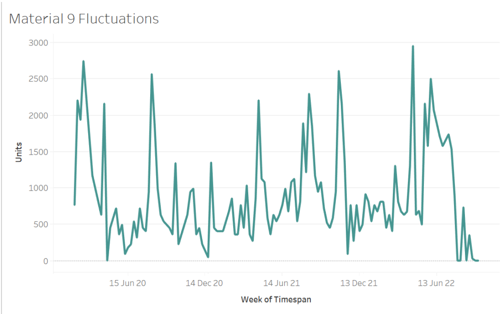


Material 8
Production of material 8 ran from the 32nd week of 2019 to the end of the given time period 2022 and had none of the weeks where units produced was 0. There is a gradual increasing variance with which we can say that this product is doing well. It can also be noted that there are no outliers in units sold which classifies this product into one of the best in terms of sales. The mean is around 2500 products, and the highest and lowest being 4000 and 1400 respectively



Material 9
Production of material 9 ran from the 7th week of 2020 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. This is one of the products with a very erratic pattern in sales and towards the last two weeks of our dataset it is seen that this product has been either discontinued or is under maintenance. We have around 7 outliers towards the maximum units sold, with the mean sales being around 650. product 9 may be a

covid product: last 2 months had 0 sales and before that there is a significant decrease in sales from 1000 to 23.

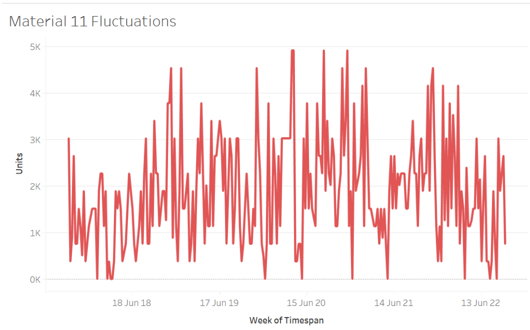


Material 10

Sales of material 10 ran from the 27th week of 2020 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 10 is around 1000. It has a minimum point of 0 and a maximum point of just under 4000. The distribution has two outliers, Sales of this product fluctuates massively by going from around 2K in a few weeks straight to near 0.

Material 11

Sales of material 11 ran from the 40th week of 2017 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 11 is around 2000. It has a minimum point of 0 and a maximum point of just under 5000. The distribution has no outliers, Sales of this product fluctuates by a lot but the mean remains fairly constant over a monthly or bi-monthly period.

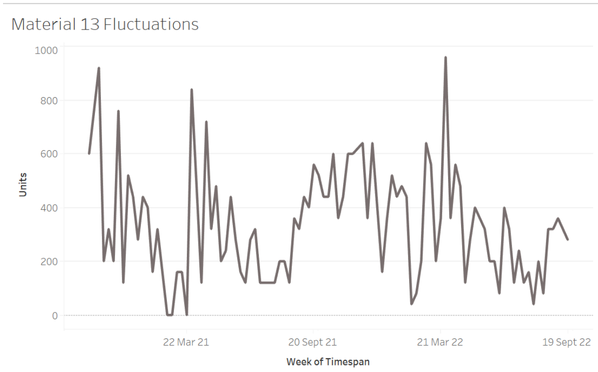


Material 12

Sales of material 12 ran from the 33rd week of 2017 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 12 is around 2000. It has a minimum point of 0 and a maximum point of just over 6000. The distribution has no outliers, **Sales of this product fluctuates by a lot but the mean remains fairly constant over a monthly or bi-monthly period**.

Material 13

Sales of material 13 ran from the 45rd week of 2020 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 13 is around 300. It has a minimum point of 0 and a maximum point of just over 900. The distribution has two outliers, **Sales of this product fluctuates by a lot but the mean remains fairly constant over a monthly or bi-monthly period**.



Material 14
Sales of material 14 ran from the 9th week of 2018 to the end of the given time period 2022 with a few weeks where units produced were equal to 0. The mean for material 13 is around 300. It has a minimum point of 0 and a maximum point of just over 900. The distribution has two outliers, **Sales of this product fluctuates by a lot but the mean remains fairly constant over a monthly or bi-monthly period**.

Preliminary data handling

No null values were identified within the dataset. However, it contained relevant zero values, which were considered significant within the problem context. Thus, these zero values were left unaltered. Furthermore, variations in start dates were observed across different material types. To address this variability, separate datasets were created for each material type, enabling the development of distinct models with unique starting points. Conversely, all data entries concluded simultaneously, sharing a common end date. Notably, Material 5 exhibited a lack of production activity in recent months. Decimal figures were discovered in the "Units" column, which deviate from the intended integer representation. To rectify this inconsistency, rounding procedures were applied to ensure that all values within the column were expressed as whole numbers.

Feature Engineering for ML Models

The data frame was filtered to find the decimal figures and manually. A year-month column was added by converting the year-week column into a datetime format. A feature of season was added by mapping seasons manually to every month. A package was imported called holidays and added a new column with the sum of holidays(Spain) in each month. After which months were changed to a numerical format. More features were added with GDP data and Unemployment data for Spain. All transformations can be found in DefinitiveDataGroupingMonth.ipynb

Statistical Testing for Models / Technical EDA

Most time series forecasting models assume stationarity, which means that the statistical properties of the data remain constant over time. In order to assure this, tests must be conducted, such as the Augmented Dickey-Fuller (ADF) test, to determine if the time series are stationary.

If the series are not stationary, techniques such as differencing or seasonal differencing should be applied to achieve stationarity. In order to start the ARIMA modelling that we will do later some of these procedures were applied. For each product a line graph was plotted. The ADF test was performed and stationarity evaluated with the help of the statsmodels library. Moreover OCSB test for Seasonal Unit Roots was performed to check seasonal differences.

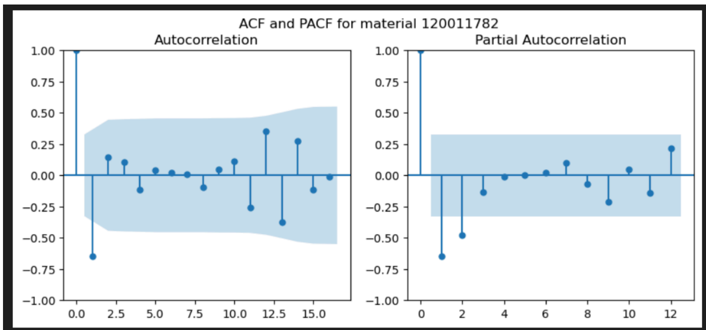
The tests indicated that only 2 materials (120004096 & 120011782) needed one regular difference to become stationary.

ADF output example:

```
ADF test for material 120004096
p-value: 0.1948228977087264
Regular differences: 1
Seasonal differences: 0
```

Autocorrelation Function (ACF) and **Partial Autocorrelation Function (PACF)** were plotted for those two materials. Plotting the ACF and PACF can help identify any significant lags or correlations in the data. In the case of white noise, there should be no significant lags, indicating no correlation. Moreover the **Ljung-Box Test** was performed. It is a statistical test used to check for autocorrelation in the residuals of a time series. The test evaluates whether a group of autocorrelations is significantly different from zero. In the case of white noise, the residuals should not exhibit significant autocorrelation. The data seems to be mostly white noise, with some exceptions.

ACF, PACF and Ljung-Box Test output example:




```
Ljung-Box test for material 120011782
p-value: 0.002366660240719144
(Our data DONT seems be White Noise)The residuals are likely not independently distributed (reject null hypothesis)
```

Normality Tests: Performing normality tests, such as the **Shapiro-Wilk test** or Anderson-Darling test, can assess if the data follows a normal distribution. White noise typically follows a normal distribution. The data was tested using a Shapiro-Wilk test, together with the previous tests this results:

Example output of our Shapiro-Wilk test

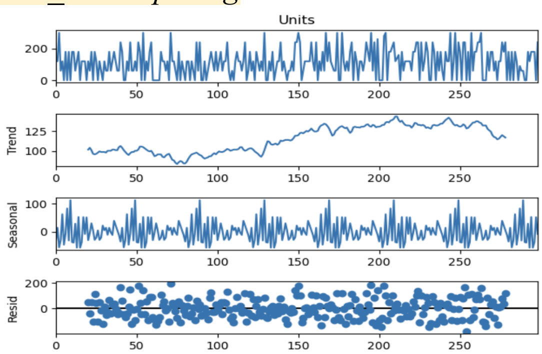
```
Shapiro-Wilk test for material 120014486
p-value: 0.13184821605682373
The data is likely normally distributed (fail to reject null hypothesis)
```

The results for each product on this tests can be checked on the annex folder inside the GitHub Repository: Appendix Number 1: *StatisticalTests_Results* and the ACF & PACF plots can be found at Appendix Number 2: *ACF&PACF_Results* the notebook of this tests and EDA can be found at *ACFPACF&TESTS.ipynb.*

Some questions can now arise, for example. Is there any correlation between different materials? This question is fairly simple to answer, through the composition of a correlation matrix. The notebook *CorrelationCalc.ipynb* will do exactly that and in the Appendix Number 3: *Correlation_Matrix* you can find the plot. This suggests that most of them are not at all highly correlated, except maybe one relationship (120012154 and 120010566). We will need further statistical analysis, for this we could perform a causality test, but because product 120012154 is not normally distributed we will be better performing a Granger Causality Test, that works with not normally distributed data which is a causality test that is specifically designed to assess whether one time series variable can predict another time series variable. It helps determine if the past values of one variable provide useful information for predicting the future values of another variable. This test is commonly used in econometrics and time series analysis to investigate causal relationships in data. The test can be checked in the notebook *GrangerCausalityTests.ipynb* the test is unable to reject the null hypothesis (one serie does not influence the other) even at the first lag, so we can confidently conclude that we should not worry about one product influencing the other if we do not want to overfit.

Data Decomposition, Seasonality and Trends

Identifying any seasonal patterns or trends present in the data. As previously mentioned Seasonality can be detected by plotting autocorrelation and partial autocorrelation functions (ACF and PACF) and examining the patterns. Trends can be assessed by smoothing techniques like moving averages or exponential smoothing. Perform decomposition of the time series to separate the components of trend, seasonality, and residual (error) using methods like additive or multiplicative decomposition. This can provide insights into the underlying patterns and help decide on appropriate modelling techniques. We performed decomposition for each time serie, using additive decomposition, this technique is the right one as you can clearly see that a multiplicative trend does not exist, the results it can be found at Appendix Number 4: *Seasonal_Decomposing*



Example of additive decomposition of a time serie of our dataset

Model Development

The goal of the modelling process in this project is to develop an advanced analytics model that can accurately forecast the demand for ACME TEXTILE CORP's products for the next two months. The aim is to improve the 80% average accuracy of the current model in use. The objective is to forecast as accurately as possible final sales, enabling the company to optimise response to demand. Thereby reducing inefficiencies and associated costs. The benefits of this will be further explained later. To achieve this goal, we will explore and compare different types of models, including traditional statistical approaches like ARIMA (AutoRegressive Integrated Moving Average) as well as more advanced deep learning / machine learning techniques. The different models will be assessed. The model which achieves the highest accuracy will be selected, evaluated and deployed. There are plenty of metrics, because the business case asked us to beat a 80% accuracy (Whatever that means for a regression problem), we decided to go with the most similar metric to accuracy but for a regression problem, MAPE. We were asked to predict a 2 point ahead forecast, but were open to whether to calculate the MAPE using 2 or 1 prediction, meaning 1 prediction for 1+2 or pred_1 and pred_2. We decided to go with the first approach as it would make sense production-wise, if the company wants to forecast the next 2 months from today it is probably because their production has to be arranged bi-monthly. For comparison purposes we will display the (1-MAPE) metric, where 100% means a perfect model. Keep in mind that different metrics and validations techniques have been used for each model, but with the purpose of comparing them and selecting the best ones in our final conclusions we will only look at the test set of the last 8 weeks, this will later be explained again.

Models

<u>Dummy/Lazy Models</u>

Several different models were tried, it is useful to categorise them in different sections,the first category are the so-called Dummy or Lazy Models, Including GlobalMeanGuessing, GlobalMedianGuessing, MonthSpecificMedian and RollingMedianWindow. The code for this models can be found at DummyLazyModels.ipynb

<u>GlobalMeanGuessing</u>

A basic approach used in business cases where a simple average or mean value of a variable is predicted. This model assumes that future values will be similar to the historical average, making it a straightforward benchmark for more complex models. We will divide the dataset by product and then into train and test, leaving the last 3 groups of 2-months together. We will first calculate the mean of all the train sets and see the bi-monthly MAPE for the first 2 months, then add this to training and predicting the next 2, repeating the process until we have tested the 3 test sets. This will result in a “robust” model, or at least tested to some extent in recent data. Because we have some outliers it will make sense to try out the median instead of the mean, we will try this approach in the future. We will not display the full results here but you can find them in the *Appendix Number 5: GlobalMeanGuessing_ResultsTable*. Here we will just mention that the Average (1-MAPE) across all products using this technique is 66.50%. With 7 out of 14 products above the 80% benchmark.

<u>GlobalMedianGuessing</u>

We will follow the same procedure as with the previous strategy, this time using the median, more robust to outliers. You can find the full results in the *Appendix Number*

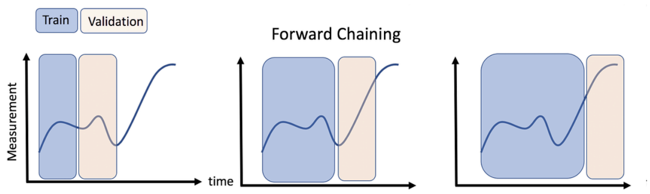
6:GlobalMedianGuessing_ResultsTable, (1-MAPE) is 66.79%. With 7 out of 14 products above the 80% benchmark. We can see it is just about the same as the Mean, slightly better.

MonthSpecificMedian

We will perform the same test as before, but using the historical mean of only that month, again we will use cross validation by using 3 validation/test sets ordered by time and retraining the model with the new data added. You can find the full results in the *Appendix Number 7: MonthSpecificMedian_ResultsTable*, (1-MAPE) is 62.03%. With 7 out of 14 products above the 80% benchmark. We can see it is just slightly worse than the previous strategies.

RollingMedianWindow

Our last dummy/lazy model will perform the same test as with the median, but using a window of months, this is, not taking the global median but only an X amount of previous months, in a maximum range of 3 years (34 months), for those products with less than 3 years of history maximum range possible was applied. We will try several windows. We discovered that our best window is 23 months, redoing the experiment with a window of 23 months. You can find the full results in the *Appendix Number 8:RollingMedianWindow_ResultsTable*, (1-MAPE) is 66.56%. With 7 out of 14 products above the 80% benchmark. We can see it is just about the same as the global median, slightly worse, this can be explained because our models do not have a significant trend.



Visual Representation of Forward Chaining Cross Validation

ARIMA , Exponential Smoothing & Prophet

We have also tried out some parametric / econometrics models. This ones can be found in the notebook ARIMAModels.ipynb

ARIMA-AIC, ARIMA-ManualFinding

(Autoregressive Integrated Moving Average): a powerful time series forecasting model commonly employed in business cases. It considers the autoregressive (AR), differencing (I), and moving average (MA) components of the data. ARIMA models are particularly useful for predicting trends, seasonal patterns, and irregular fluctuations in non-stationary time series data. They provide the capability to capture both short-term and long-term dependencies, making them valuable for forecasting tasks. ARIMA (Autoregressive Integrated Moving Average) models are widely used in time series analysis for forecasting. Unlike many other machine learning models, ARIMA models do not typically require cross-validation for model selection. This is because ARIMA models rely on historical data and aim to capture the underlying patterns and dynamics of the time series. Cross-validation, which involves dividing the data into training and validation sets, is often used to evaluate the generalisation performance of models when there is a concern about overfitting. However, in the case of ARIMA models, the model parameters are estimated using the entire available historical data, making cross-validation unnecessary. Instead of cross-validation, a common method for selecting the best ARIMA model is to use the Akaike Information Criterion (AIC). The AIC is a statistical measure that quantifies the quality of a model by balancing the trade-off between model complexity and goodness of fit. It penalises models with more parameters, favouring simpler models that adequately capture the patterns in the data. The AIC considers both the residual errors and the number of parameters in the model, making it a useful criterion for selecting the most appropriate ARIMA model. By minimising the AIC, one can identify the model that provides the best balance between accuracy and simplicity for a given time series data set. You can find the full results in the Annex: *Appendix Number 9.1: ARIMAAIC_ResultsTable*, (1-MAPE) is 49.65%. With 8 out of 14 products above the 80% benchmark. We can see that the AIC tends to play "very safe" to avoid overfitting and ends up being not that good. We will ditch the AIC-way of finding the best model and we will simply try to find the ARIMA with the best score on unseen data. You can find the full results in the *Appendix Number 9.2: ARIMAManual_ResultsTable*, (1-MAPE) is 74.33%. With 9 out of 14 products above the 80% benchmark. This has been our best try yet, and we will check that when adding manually our business rules, our score will be 86.79%

Prophet

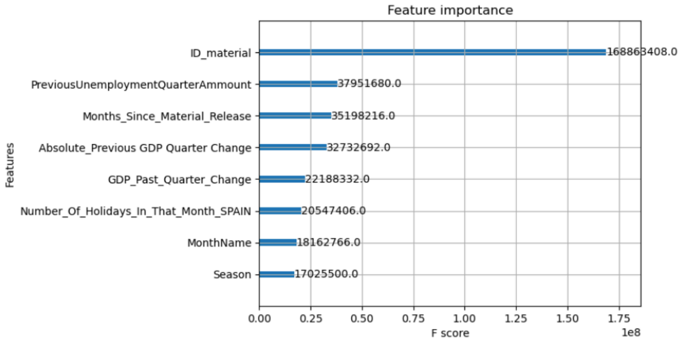
The results were similar to XGBoost but with some more error, this model is state of the art, but for its use we would need much more granular data, with a more clear seasonality and trends so we decided not to include the trials here.

Exponential Smoothing

Exponential smoothing is a popular forecasting method used for time series analysis. It is a simple yet effective approach that assigns exponentially decreasing weights to past observations, with more recent observations given higher weights. The basic idea behind exponential smoothing is to estimate the future values of a time series by giving more importance to recent observations while progressively reducing the influence of older observations. Exponential smoothing works well for time series for several reasons: Simplicity: Exponential smoothing is easy to understand and implement compared to more complex models like ARIMA, making it accessible to a wider audience. Flexibility: Exponential smoothing can be adapted to different types of time series data, including those with trends and seasonality. Variations such as double and triple smoothing can handle more complex patterns. Double and triple smoothing are extensions of exponential smoothing that incorporate additional components to handle more complex time series patterns. Double smoothing adds a trend component to the basic exponential smoothing equation.Triple smoothing extends double smoothing by incorporating a seasonal component in addition to the level and trend. Our code developed for testing our products with Exponential smoothing can be found as ExpoSmoothTrial.ipynb. Results were amazing, being able to predict with a 90% (1-MAPE) on non-white noise series with all of them below the 20% benchmark.

Machine Learning and Deep Learning Models

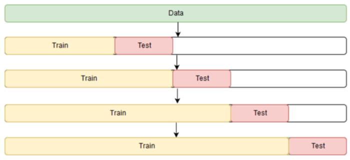
Machine learning models can be found on the notebook: MLModelCrossValidated.ipynb and DLModel.ipynb. Our first step was testing Machine Learning models, because of their ability to include external features on their training for forecasting we have added all our previously enriched columns to the training dataframe, although we tried many different ML models that usually work well with time-series, including LightGBM, KNN, SVR, LinearRegression, RandomForestRegressor and XGBoostRegressor, because their similarities in nature (they work very differently but all of them follow the same learning structure) and way of testing we will only explain the one that landed us the best results while being robust, XGBoost Our cross validation results (Cross-Validated 3 Bi-Month MAPE (%)) can be found in the annex *Appendix Number 10:CrossValidationML_ResultsTable.* XGBoost won in 8 of the 14 products, being the winner, behind SVR and LightGBM with 2 products where they are the best each and finally KNN and Linear Regression with 1 each. The main features of our best models for most the overall products are (based on the F-Score):



To the surprise of no-one, the most important feature overall is the ID_material, this makes sense as each product behaves differently and has different levels. Following this feature Unemployment is the second most important one, just around 4 times less important than material. This may just be a random finding and our model may be overfitting, but it is a fact that our model improves predictions on unseen data when this data is provided, following we find the Months_Since_Material_Release that is basically a point basis trend, that of course makes sense for those products that exhibit any trend, so it makes sense that this feature is important, then we find features about GDP, again we do not know if this affects our business because we are lacking context, but either way is important, then we find Holidays, Month and Season, that are typically important when forecasting sales.

XGBoost (Extreme Gradient Boosting)

An ensemble machine learning algorithm widely used in various business applications, including regression and classification tasks. By combining multiple weak prediction models, typically decision trees, XGBoost creates a strong predictive model. It excels at capturing complex interactions and nonlinear relationships in the data, outperforming other algorithms in many cases. This makes it highly effective for solving business problems like churn prediction, demand forecasting, and fraud detection. You can find the full results directly in the final results table, with a(1-MAPE) of 80.26%. Without products above the 80% benchmark. Because of the nature of MLModels we were able to develop a “Whole Company” model, this is, a single model able to predict all products of the company, it only needs to be trained once, (or retrained once) and deployed once, also its able to get more complex relationships between the products (probably just noise, but still some day maybe it finds some pattern that is not noise) Because here we are mixing products and we cannot aggregate bi-monthly we will measure the model on RMSE, so its more robust across products, with this approach we got a MAPE across all products, prediction sales of all products of 123519 when real sales across products were 126351, but it failed to grasp individual relationships. You can check both in the notebooks: XGBoostModelsGeneral&Individual.ipynb & XGBoostOnebyOneModels.ipynb. The model was saved as regressor_model.pickle



Visual Representation of Time Series K-Fold Cross Validation

RNN (Recurrent Neural Network)

A neural network architecture specifically designed to handle sequential dependencies in data. In business cases where temporal information is crucial, such as natural language processing, speech recognition, or time series forecasting, RNNs are often employed. Their unique feedback mechanism allows information to flow from previous steps to the current step, enabling the modelling of short-term and long-term dependencies. RNNs are particularly valuable when the order of data points matters in the analysis. We got a (1-MAPE) of 85.99%

The Metric to Evaluate Performance.

The metric used to assess the effectiveness of the various models is MAPE, which stands for Mean Absolute Percentage Error. MAPE measures the average percentage difference between the predicted values and the actual values, providing insights into the accuracy of the models. Specifically, we calculated the MAPE for the first month and the second month, and subsequently divided the sum of these two values by two to obtain the overall MAPE.

$$\text{Accuracy in percentage} = 100 - \text{MAPE}$$

MAPE 2 MONTHS TOGETHER	*Mean*	*Median*	*MonthMedian*	*RollingMedian*	*ARIMAAIC*	*ARIMAManual*	*ExpoSmo*	*ML (XGBoost)*	*DL (RNN)*	*Best Models*	*Best Models All Products*
Product1	18%	25%	11%	21%	9%	11%	2%	29%	20%	2%	**2%**
Product2 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**9%**
Product3 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**4%**
Product4	6%	8%	11%	7%	12%	1%	16%	10%	1%	1%	**1%**
Product5 (White Noise)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	**42%**
Product6 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**23%**
Product7 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**12%**
Product8	20%	20%	22%	19%	11%	1%	16%	8%	4%	1%	**1%**
Product9 (White Noise)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	(BR)	**11%**
Product10 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**15%**
Product11 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**44%**
Product12	10%	15%	15%	23%	0%	4%	0.2%	9%	1%	0.2%	**0.15%**
Product13	53%	60%	27%	60%	36%	17%	11%	27%	42%	11%	**11%**
Product14 (White Noise)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	(WN)	**56.00%**
MEAN ACROSS PRODUCTS	**22%**	**26%**	**17%**	**26%**	**14%**	**7%**	**9%**	**17%**	**14%**	**3%**	17%
(1-MAPE)	**78%**	**74%**	**83%**	**74%**	**86%**	**93%**	**91%**	**83%**	**86%**	**97%**	**83%**

After running all models it became clear that a combination of different models for each product would give the highest accuracy. Regarding feasibility of a deployment, a pipeline could be built with separate models per product. However this might be difficult to scale. If that is the case, using an ARIMA model for each product should be best for average overall predictions and accuracy. For 9 Products the best model was using the mean as we explained before, 2 of which also used business rules, then, Exponential Smoothing: 3, ManualARIMA: 2. Predictions can be checked in the folder Predictions_Graphs on our GitHub Repository

Recommendations & Next Steps

To improve forecasting results with only weekly sales data, ACME TEXTILE CORP can employ several strategies:

Data aggregation

The company can aggregate the weekly sales data into monthly totals. By summing up the sales data for each product over a month, the company can obtain a monthly sales series. This aggregation will provide a higher-level perspective of the demand pattern and eliminate some of the noise and fluctuations associated with weekly data. Monthly data will provide a smoother and more stable basis for forecasting models. If producing monthly/bi-monthly is necessary, collecting this data in this way would be preferred.

Additional Data

In addition to the aggregated monthly sales data, ACME TEXTILE CORP can incorporate additional features that may influence demand, such as promotional activities, product price, seasonality, holidays, or economic indicators. By including these relevant factors as input features in the forecasting model, the company can capture more nuanced patterns and improve the accuracy of predictions. For example, including the number of promotions or the consumer price index as features could help capture the impact of marketing efforts or changes in purchasing power on demand, as well as getting more granularity in the data by maybe measuring the demand by client, they could predict churn and improve forecasting. By implementing these strategies, ACME TEXTILE CORP can overcome the limitation of having only weekly sales data and improve the accuracy of their demand forecasting. The aggregated monthly data will provide a more stable foundation for modelling, while incorporating additional features will capture the various factors that influence demand. These enhancements will enable the company to make more accurate predictions and optimise production and stock management accordingly.

Next Steps

Deploying ARIMA and exponential smoothing models for sales forecasting on a cloud platform like MLOps can greatly benefit a business. MLOps, or Machine Learning Operations, provides a framework for efficiently deploying, managing, and monitoring machine learning models in production. By leveraging MLOps, the business can automate the entire lifecycle of the forecasting models. Firstly, the ARIMA and exponential smoothing models can be trained on historical sales data stored in a cloud database. The models can then be deployed on a cloud-based MLOps platform, which handles the infrastructure and scaling requirements. This allows the business to easily generate sales forecasts for any given time frame. The MLOps platform can automate the process of fetching new sales data, retraining the models periodically, and deploying the updated models. It also facilitates monitoring the performance of the models and alerts the business if any issues arise. With this deployment, the business can access accurate and up-to-date sales forecasts, enabling them to make informed decisions and optimise their operations effectively.

Business case

Sales forecasting allows businesses to estimate future sales volumes. Organisations can reduce costs and improve their overall profitability, by accurately predicting demand and allocating resources. As such, it is important to increase the forecasting accuracy as much as possible so that Inventory, Production, and Supplier Management are fully optimised.

Although ACME Textile already forecasts its demand with an 80% accuracy, it is important to note that there is still much to improve. As mentioned before, accuracy in predicting sales has a great impact on cost saving, as such having these many units poorly allocated not only has a great production, transportation, and storage cost but also an opportunity cost. This error severely impacts and endangers ACME Textiles' financial well-being and sustainability.

Having this in mind, as the newly appointed Data scientist, the responsibility lies in designing and implementing an advanced forecasting model that surpasses the existing average accuracy rate of 80% for all products. Consequently, an ARIMA model was developed that forecasts the demand with an almost 87% accuracy, surpassing the original model by 7 percentage points. This increase translates into a reduction in the overall costs and ultimately, a step forward into financial stability.

Furthermore, in order to translate these units into financial measurable values, there were two assumptions that needed to be taken into place 1) the revenue per unit 2) the profit margin. After some research on companies that are similar to ACME it was realised that the average revenue per unit was **359€** and the profit margin was around **20%**. Knowing these values, we can now measure the monetary impact that this **87%** accuracy ARIMA model can have on ACME. As previously mentioned, there should not only be taken into account transportation, fabrication and storage cost but also opportunity costs since the latter also severely impact the company. As such, being the basis of this analysis the percentage of over and under produced units calculated with our **87%** model which then was measured if the model had a **20%** MAPE, the overall impact was a decrease in **15 500** units that were wrongly allocated. This translates into a decrease of **13 000** over produced units and **2500** under produced units. Consequently, ACME was able to reduce its variable and fixed costs by **3 584 706€** and finally avoid a **175 804€** opportunity cost.

Given the assumptions listed above it is feasible to create alternative scenarios which ultimately support the effectiveness of the model and the use of models for demand prediction.

Scenarios 1 , Decreasing over production

In this scenario, ACME Textile focuses on reducing overproduction by implementing the model. By more accurately predicting demand, the company is able to optimise its production process and avoid manufacturing excessive units that may go unsold and result in both variable production cost incurred and additional overhead due to storage, labour cost and energy. With the ARIMA model in place, ACME Textile identifies that it has been overproduction of approximately 13,000 units. This reduction in costs amounts to a total savings of 3,584,706€.

This not only saves the company operational costs but also increases their leads to a positive impact on their sustainability efforts which could lead to a more positive brand image.

Additionally, this reduction in cost can be redistributed to coming financial periods or reinvested in PPE.

Business scenarios 2. Decreasing Underproduction

In this scenario, ACME Textile aims to address the issue of underproduction. Underproduction occurs when the company fails to meet the actual demand for its products, resulting in missed sales opportunities which translates to opportunity cost and potential customer dissatisfaction.

By implementing the ARIMA model, ACME Textile identifies that it has been underproducing by 2,500 units, which indicates a gap between the actual demand and the production output. This underproduction not only results in lost sales revenue but also has a negative impact on the company's market position and customer satisfaction.

By decreasing underproduction, ACME Textile ensures that it meets the actual demand for its products, thereby capturing additional sales revenue and enhancing customer satisfaction. The increased production of 2,500 units helps the company capitalise on missed sales opportunities and improve its market competitiveness.

From a financial perspective, decreasing underproduction has a positive impact on the company's revenue and profitability. With an average revenue per unit of 359€ and a profit margin of 20%, the additional 2,500 units produced due to improved forecasting accuracy translate into an increase in revenue of 897,500€ and a profit of 179,500€.

By implementing the advanced forecasting model and focusing on decreasing underproduction, ACME Textile not only improves its sales performance but also enhances its financial results. This scenario highlights the importance of accurate demand forecasting to ensure optimal production levels, meet customer demands, and drive business growth.

Conclusion

To conclude this report we believe that the implementation of an ARIMA model will significantly improve ACEMS textiles operational efficiency, profitability and sustainability. Creating an effectively on demand product line will lead to instrumental benefits not only to the business itself but additionally to the buyers and the planet's resources.

By improving the original models accuracy rate by around 7% ACME textiles will hopefully see more financial stability and a decrease in over or underproduction. ACME should continue the practice of demand prediction with appropriate model alterations, hyperparameter tuning, retraining and alert triggers. We recommend increased data governance and documentation as well as establishing a thought out ML pipeline.

As the lines of demand and supply converge, the market becomes more efficient, price becomes fair and resources will not be gluttonously utilised. The leverages used to stabilise the economy are becoming more and more sophisticated. The market finds balance, businesses make profits, consumers pay what they are willing to pay. The economy, like all finds balance with the help of data, and machine learning.

Acknowledgements

As a team we we would like to convey our deepest gratitude first and foremost to our mentor, Michelle Greaves who guided us through this turbulent process. Micelle, we admire your dedication to us, consistency and your unwavering diligence in getting us answers, time, aid and resources. We are entirely grateful to you.

We would also like to thank the admissions team at IE university for allowing us this opportunity in the first place and for their forever supportive presence around campus and in our daily lives.

We would also like to thank the KPMG team, for providing us with this opportunity to work on a real life problem, thereby enhancing our skills and making us more apt to face our soon to come professional challenges.

Lastly we would like to thank each other, for being the foundation needed to support each and everyone. For going further together, than we ever could alone.

Appendix

All extra Charts/Visualisations and Notebooks can be found at the GitHub Repository:

https://github.com/notvikke/KPMG_TimeSeries_Forcasting_Capstone