INDUSTRIAL & SYSTEMS ENGINEERING
TEXAS A&M UNIVERSITY

ISEN 615: Final project report

Date – 12/3/2018

ARCONIC

VT INDUSTRIES

Team members:

1. Aparajit Koshal (326004962)
2. Harish Chellappa (826006615)
3. Himanshu Gupta (426007282)
4. Madan Taldevkar (523007674)
5. Rashi Sinha (927000967)
6. Siddharth Ajit (227001842)

Contents

Arconic demand forecasting

To begin with, we performed some exploratory data analysis on Arconic demand. We used Market A data for basic exploration and understanding. We investigated the significance of each column and how can we use them in forecasting. We plotted demand by months (After rolling up at month level) at product family level and item level using excel. It gave us an idea that lot of product families/items are not having orders in all of the months. Also, it gave us an idea what should be the appropriate level of aggregation to forecast demand. We chose important columns based on the analysis and performed some pre-processing on data before applying any forecasting technique. Below are some observations and decisions based on exploratory analysis –

1. For Market A, there was 41227 rows of data in historical data sheet and 6850 rows in open data sheet. As historical and open data both were having maximum order date of September'2018, we decided to combine them. This gave us the actual sense of demand in the June 2014 – September 2018 period that we are developing our forecast on.
2. After looking at the data carefully, we aggregated demand at Product family level and item level. We decided that we would go ahead with forecast on Product family-month level and item-quarter level, as they seemed appropriate looking at the demand numbers and patterns.
3. We extracted mainly 4 columns from Arconic data – Product family, Item no., order date (to extract months and quarters), and Demand quantity.
4. We cleaned data and removed observations wherever quantity < 0 (about 1.3% of data).

We have forecasted on two levels of demand aggregation – monthly aggregation for product families & quarterly aggregation at SKU (item#) level. According to the textbook, monthly forecast at family level is a short-term forecasting technique. We have used 6 techniques namely- Moving average, Exponential smoothing, Seasonality, Regression method, Holt's method and ARIMA. Quarterly forecasts at item level is a long-term forecasting technique. Holt's method has the parameter Tau that is used for multiple step ahead forecasts for future months.
In addition to all these models, we tried to implement Holt Winter method, but we faced some problems due to missing demand. The averaged seasonal factors were coming out to be 0 for some missing demand months which is used as a denominator in calculation St (level) at each step of series. So, we did not go ahead with this approach.

As of now, our models use Market A data for forecasting purpose but our analysis can easily be replicated to Market B and C datasets as well.

Product family and Monthly level - short term forecast

Data Preprocessing

As we wanted to get our data ready for forecasting, we applied some pre-processing touches to the dataset. We started off with the actual data format as per the dataset. Daily and weekly aggregation of demand did not yield good results. We found sufficient data at month level aggregation for each product family.

- We used three columns from main dataset – Product family, Order Month (**yyyymm** format), demand quantity and aggregated data at product family-month level

- we found some months with missing demand. So, we added missing months demand as 0 to resolve inconsistency in data. This increased number of records increased to 62618.
- The aggregated version had 4680 records. There were 90 product families to be analyzed. Each product family had 52 months of data from 201406-201809 (removed one 201404 observation).

1. Moving average model

We used three types of averages - MA (3), MA (6) and MA (9) to forecast demand and calculated absolute error at each product family-month level for all 90 families on 52 months data. After that, we calculated Mean Absolute percentage error for each family at all three models and rounded it off to nearest decile to compared them. The one which gave the minimum MAPE value, was considered for that family. So, in summary, some product families got MA (3) as best model while some other got MA (6) or MA (9).

Below is the chart of least rounded off MAPE vs count of families. We found out that 42 out of 90 product families had <= 50% MAPE.

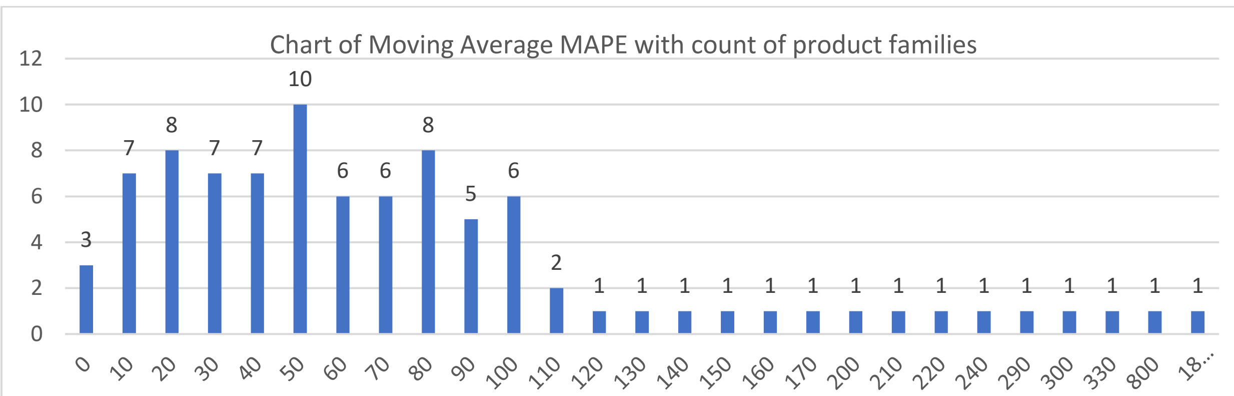


Fig1: Moving average MAPE chart

2. Exponential Smoothing

We tried similar approach for exponential smoothing using 4 different values of Alpha as 0.1, 0.2, 0.5 and 0.8. We chose these 4 α values considering the varying weights placed on previous demand. Around 29 families out of 90 gave <= 50% MAPE.

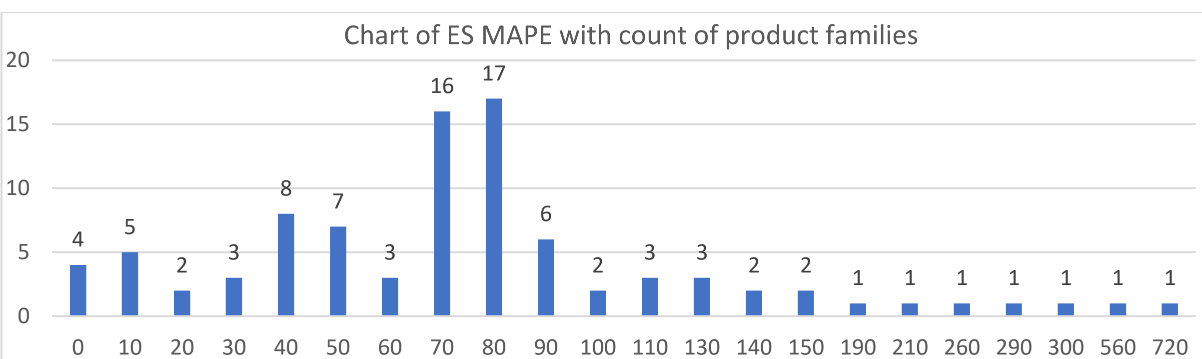


Fig2: Exponential smoothing MAPE chart

Also, table (click here) explains How many families had which Alpha value that gave minimum error.

3. Seasonality approach

We used programming tool R for implementing this method to all the families. First, we calculated the average of demand for each family across all 52 months and calculated seasonal factors for each month by dividing monthly demand by average demand. After that Monthly seasonal factors were calculated by taking an average of all seasonal factors for each month. Thus, we had 12 seasonal factors, one for each month in the year. We multiplied the seasonal factors with average demand of the product family to forecast demand for all 52 months. Around 31 families out of 90 gave <= 50% MAPE.

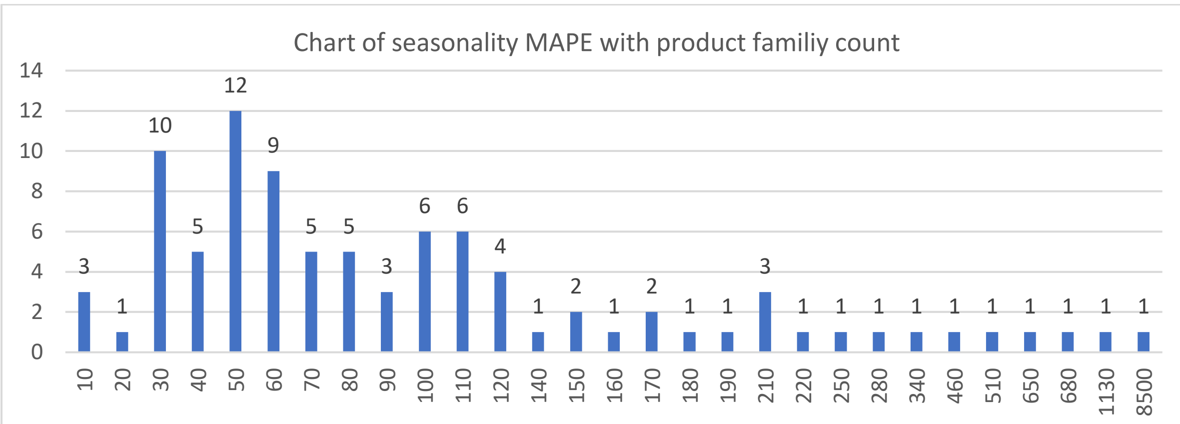


Fig3: Seasonality MAPE chart

4. Regression modelling approach

We fitted a regression model on each product family using R. We regressed monthly demand with month no. (1 to 52 from their sequence) and predicted on same data using the fitted model. Thus, we used this predicted demand as forecast to calculate MAPE at family level. Click here to access its MAPE chart. Regression did not perform that well compared to other models.

5. Holt's method approach

We used double exponential smoothing method - Holt's method, to capture trend and series in the data. We used alpha = 0.2 and beta = 0.2 for this method and calculated values of trend and level at each step using R (as we wanted to replicated Holt's logic to all families which could be done using a programming language). We chose initial values of level and trend using mean and slope of each family demand. As it can be seen in chart (click here), Holt's method forecasted 6 families <= 50% MAPE.

6. ARIMA model

Non Seasonal Arima models are defined as ARIMA (p,d,q) where p stands for the lag order, d for the differencing order to make the series stationary, q for the order of moving average. To estimate the parameters of the ARIMA model, *forecast* package in R environment was used. Each of the family demand for the first 36 months was fit using auto.arima function. The ARIMA model parameters with lowest value of AIC was selected to further predict the demand for the next 16 months. Finally, the MAPE for the predicted periods were calculated to compare the model performance. 40% of the family gives us MAPE value less than 50%

Comparing all these models, a dashboard has been prepared which choses best model for each family automatically and provides method name as well as % MAPE.

Summary table of outputs

Model	% of Product families that had <=50% MAPE (rounded)
Moving average	47%
ARIMA	40%
Seasonality	34%
Exponential Smoothing	32%
Holt's Method	7%
Regression	4%

Table 1: MAPE summary of models at product family month level

As it can be seen in the table above, moving average proved to be the best model for short term forecast.

Item# (SKU) and quarterly level - long term forecasts

As the short-term forecasts at product family level proved to be reasonably accurate, it made sense to implement long term forecasts at a more granular level. Hence, we started our exploratory analysis at item level.

Data Preprocessing

We extracted product family, item no, order date and demand columns from Market A dataset We started by looking at monthly data but that was highly sparse, so we tried quarter and half yearly level demand aggregations at item level. Below are some observations –At Quarter level -

- The data starts from Q2-2014 and ranges to Q3-2018, giving us 18 quarters of demand with 1180 unique items
- 774 items of these 1180 do not have any demand in at least 15 of the 18 quarters which means they would not give good results for any forecasting technique, including seasonality. So, we decided to drop them

At Half-yearly level -

- Data was divided into 8 halves for each item from H2-2014 till H1-2018
- 690 out of 1180 SKUs have zero demands in 7 out of 8 halves. 784 out of 1180 have zero demands in 6 / 8 halves, and 829 out of 1180 have zero demands in 5 / 8 halves

As both the levels of aggregations generate huge sparsity in demand at their level, we can conclude we don't have sufficient data to provide conclusive results at item level granularity.
Hence, after team discussions, we decided to provide forecast for only those Items that had at least 4 quarters of data.

- We came up with (1180-774 = 406) distinct Items no. that had atleast 4 quarters of data
- We removed Q2-2014 data because that had only 1 item ordered

- For missing quarter demand, we imputed 0 demand. Now, each Item had 17 quarters of data
- We created quarters in YYYY_Q format where Q was no. of quarter (Jan-Mar months is 1st quarter and so on)
- The aggregated dataset had 6902 rows

We used all 5 techniques that were used for product family-month level forecast and replicated the logic in excel/R on Product Family-Item-quarter level.
Here are the charts of each approach –

1. Moving average

MA (3), MA (6), and MA (9) were used for forecast and model which gave least MAPE chosen for that item-quarter level forecast. As it can be seen from chart below, 161 items out of 406 had <= 50% error.

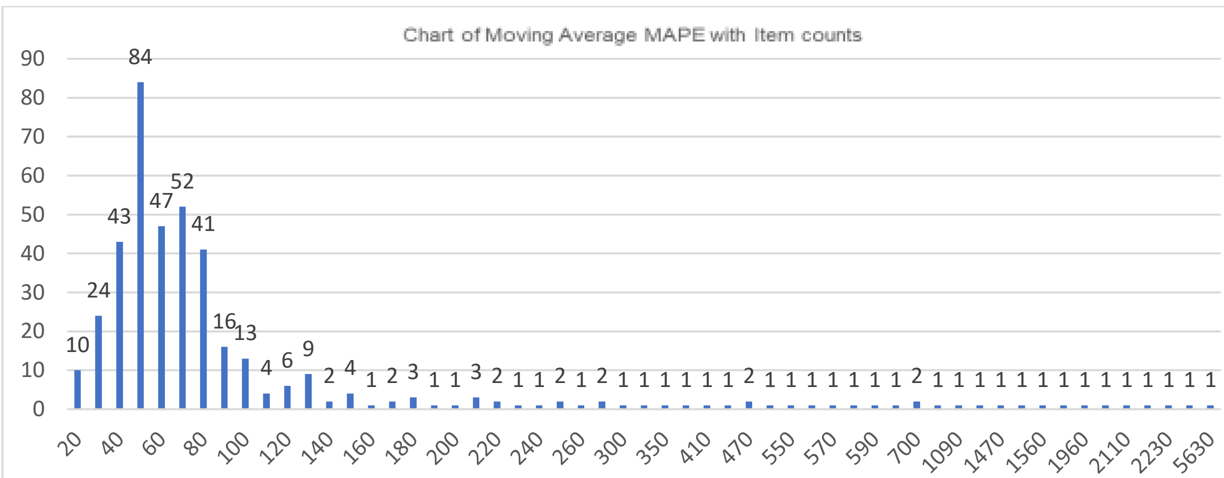


Fig4: Moving average MAPE chart

2. Exponential Smoothing

Alpha values 0.1, 0.2, 0.5 and 0.8 were chosen to forecast. Around 121 items out of 406 items gave <= 50% MAPE. Click here to see the chart.

3. Seasonality

Here, 81 items out of 406 Items gave <= 50% MAPE. Click here to see the chart.

4. Regression

Only 59 out of 406 items gave <=50% MAPE. Click here to see the chart.

5. Holt's approach

Only 46 items gave <= 50% MAPE. Click here to see the chart.

Summary table of outputs

Model	% of items that gave <=50% MAPE (rounded)
Moving average	40%
Seasonality	30%
Exponential Smoothing	20%
Regression	15%
Holt's Method	11%

Table2: MAPE summary of models at Item-quarter level

On item level too, moving average came out to be the best forecasting approach.

The Dashboard that gives you the best model at one-click

Based on the above results, we have 2 dashboards - 1 each for short term forecasts and long term forecasts. In each of the dashboard, the underlying data is being referenced from the models mentioned above, and its results. We have a drop-down feature to select a particular product-family / item# which initially plots the demand pattern. It also gives the best model which has the least Mean Absolute Percentage Error (MAPE). This technique can be adopted for that particular family / item for future predictions.

For this purpose, we have made a separate interactive dashboard which asks the user to enter demands for 12 time periods, and predicts the demand for 3 methods for any user-specified multiple step ahead forecast.

VTI

About VTI – Exploratory Data Analysis (Accomplishments)

- There are 3 product families (Particle board, MDF, and plywood) comprising 27 different items varying in material, density and sizes.
- These 27 items are supplied from 4 major suppliers namely:
    a. Georgia Pacific (Atlanta, GA)
    b. Roseburg (Springfield, OR)
    c. Arauco (Haywood, NC)
    d. Pacific Wood (Diboll, TX)
- Some of the items are high volume items namely: PB 8, PB10 and PB12. They are ordered in full truckloads. Other 24 items are ordered as shared truckload. Depending on the supplier products are combined in the following manner:

Supplier	Items
Georgia Pacific	Wide PB 8, Wide PB10 and Wide PB12
Arauco	MDF8, MDF10, MDF12, WideMDF8, WideMDF10, WideMDF12
Pacific Wood	Plywood8, Plywood10, Plywood12, Wide ply 8, Wide ply 10, Wide ply 12

Roseburg	FSC 8, FSC 10, FSC 12, Wide FSC 8, Wide FSC 10, Wide FSC 12, Cove 8, Cove 10, Cove 12

Table3: Supplier item

- With the demand statistics provided by VTI, coefficient of variation (CoV) was calculated. As can be seen from the graph most of the item's category (67%) has stable demand. 4x12'Plywood has the highest variability in the demand with CoV=25, followed by 4x8'SkyblendFSC 4x10'SkyblendFSC. This data will help in developing recommendations as product with high variability

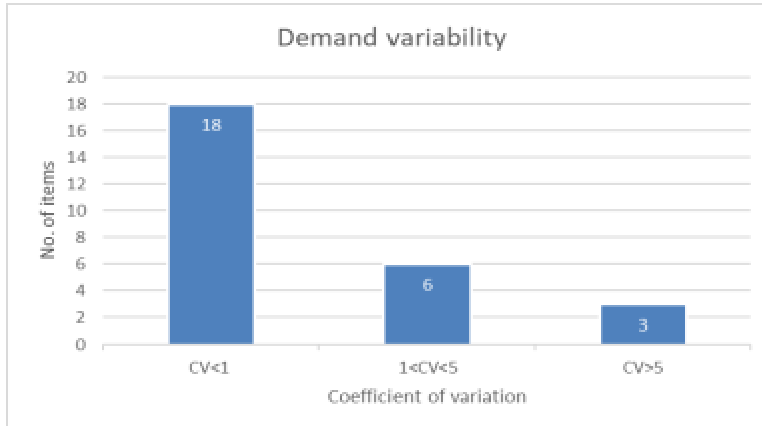


Fig5: Demand variability

- VTI currently uses an excel tool to monitor safety stock, on hand inventory and reorder quantity. This sheet serves as an ERP model as it stores weekly data for previous orders. QR strategy has been used to determine maximum inventory level (up to level) and reorder point in accordance with service level (95%)

Analysis of the existing Excel tool

The current excel tool helps the material management team to place purchase orders in terms of truckloads. To calculate numbers of truckloads to order is made with the help of following metrics.

1. Reorder point, maximum inventory:

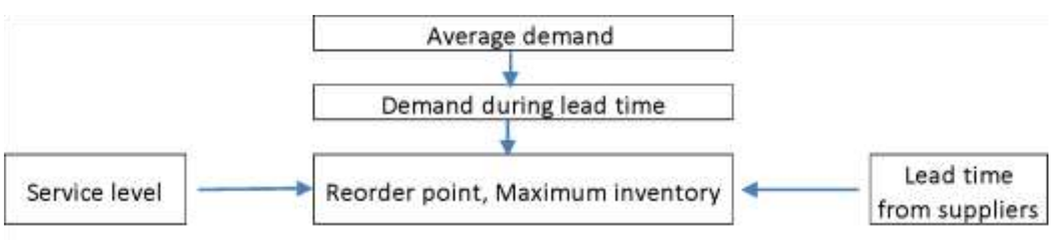


Chart1: Reorder point, maximum inventory

2. With the help of ROP and maximum inventory = evaluation of number of trucks

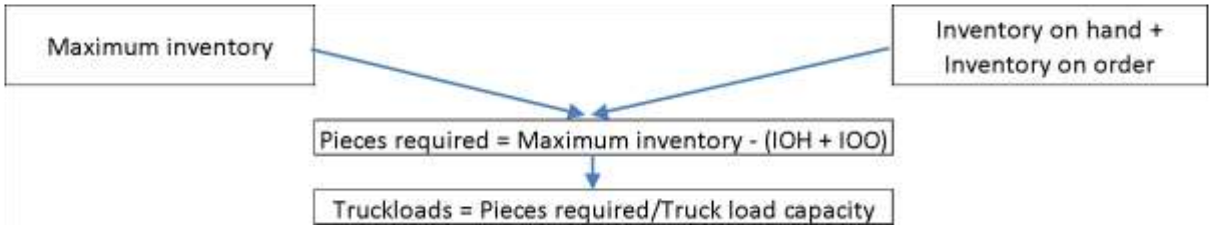


Chart2: Number of trucks

3. Visual inventory:
   With the help of conditional formatting, maximum inventory and reorder window has been shown for every product in the visual inspection tab. Color coding has been performed in such a manner which helps in distinguishing the level of inventory. It also takes into account the lead time for each product-supplier combination which enhances the visualization.

What changed?

Initial plan	Updated plan	Changed. Why?
To develop a mathematical model to derive an optimal purchasing strategy (MILP model)	To revise the strategic tool that VTI already has in place which already incorporates a heuristic approach to calculate the number of trucks	A mathematical model will be hard to explain as well as it is only justified when the problem at hand is very large in scale. Therefore, the existing heuristics works better
No dashboard was included in the initial plan	Interactive dashboard will help better understand the tool VTI has in place. As of now the tool is incomprehensible	Realized the need of VTI. Updating their tool which makes sense.

Table4: changes

Assumptions:

1. Apart from the aforementioned suppliers, there are no other alternative suppliers for 27 products.
2. We have only considered one plant of VTI for the analysis, located in San Antonio.
3. Distributions with a coefficient of variation to be less than 1 are low-variance, whereas those with a CV higher than 1 are high variance.

Data Visualization

Current tool/dashboard:

VTI currently has six sheets which it uses to display all the necessary parameters related to inventory, namely:

- How much to order
- When to order
- Safety stock
- Reorder range
- Maximum inventory possible to hold

Apart from this information, VTI uses a visual inspection to provide details about their current inventory position by considering how much material already has been ordered as well as how much inventory VTI currently holds.

Based on the truck capacity, inventory position VTI also evaluates recommended trucks needed to satisfy the current demand for 27 products.

PB 8	PB 10	PB 12
40202	40201	40200

Apart from PB 8, PB 10, PB 12 supplied by Georgia-Pacific, no other product is purchased as FTL (full truck load) from the vendors. Hence for other 24 products, generally the items are purchased in mixed

purchase order format. This is due to comparatively less demand for 24 products as compared to PB 8, PB 10, PB 12.

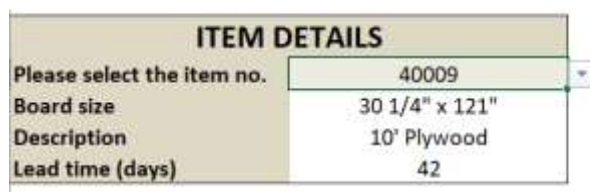
The current tool VTI has in place makes it difficult to understand all the information relating to inventory management. Although the tool is very useful, our objective has been to make it more user friendly. In order to enable that we did complete overhauling of the dashboard and developed a new one.

New dashboard details and instructions:

Due to space constraints, Dashboard screenshot image is put up in appendix section. Please click here to see it

In this dashboard, we have tried to keep all the necessary information in one single sheet. There are 4 different components of this sheet.

1. ITEM DETAILS: In this section as the name suggest once an item number is selected, item description including board size, name, lead time gets displayed. The provision to select a product has been enabled with the help of a drop-down list.



ITEM DETAILS	
Please select the item no.	40009
Board size	30 1/4" x 121"
Description	10' Plywood
Lead time (days)	42

Image1: Item details

Conclusion

We understood the importance of forecast accuracies in other related departments like production planning. For routine time series models, using simple methods such as moving average or exponential smoothing made a great deal of sense. Rightly, the textbook quotes "At the individual item level, short-term forecasts for a large number of items are required, and monitoring the forecast for each item is impractical at best. The risk of severe errors is minimized if simple methods are used." This is now seen as proven result from this project.

2. INVENTORY DETAILS: Once the product is selected. This section enables the user to enter the information like how much inventory is present in the plant (highlighted in yellow). Taking previous orders into account, this tool automatically evaluates how much inventory is on order. Based on this dashboard provides the recommended truckload utilization. In the end user can select how many trucks to order.

INVENTORY DETAILS (lbs)	
Inventory on hand	70
Inventory on order	0
Maximum capacity	400
Top of reorder window	240
Bottom of reorder window	159
Recommended order	330
Truck capacity	789
Recommended trucks	0.42
Ordered	0

Image2: Inventory details

Supplier
Pacific Wood

Product mix
12' Plywood
10' Plywood
8' Plywood
4x12' Plywood
4x10' Plywood
4x8' Plywood

Image3

3. SUPPLIER DETAILS & PRODUCT MIX: Since most products are ordered in combination of another, this section will help the material management team to understand what all products can be combined in one order and who is the designated vendor for that product.

4. VISUAL INSPECTION: The visual inspection chart enables user to monitor inventory levels of desired item. Depending on reorder point and demand during lead time, Inventory is divided in three zones, namely safe zone, reorder window and critical zone.

5. DATA LOGGING: Once recommended order quantity is confirmed by user, order details gest stored in log sheet and order quantity is added to inventory level

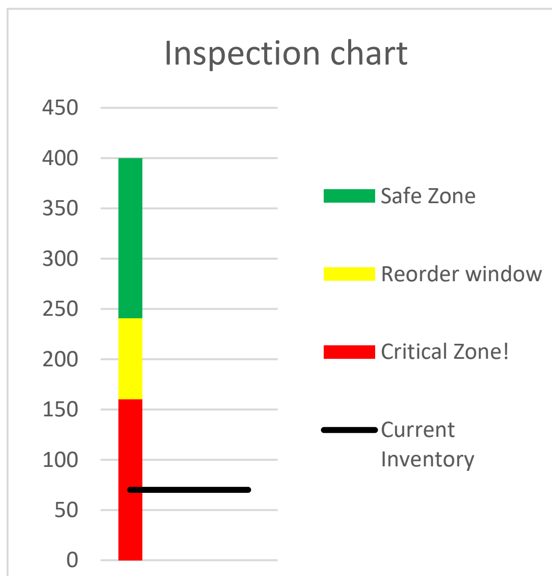


Image4: Inspection chart

Appendix

Product family – Month level forecasting charts

α value	Count of Product families for which this value of α has minimum % Error
0.1	43
0.2	17
0.5	16

0.8	23

Table5: exponential smoothing Alpha values with # of product families

Back to top

1. Regression forecasting chart

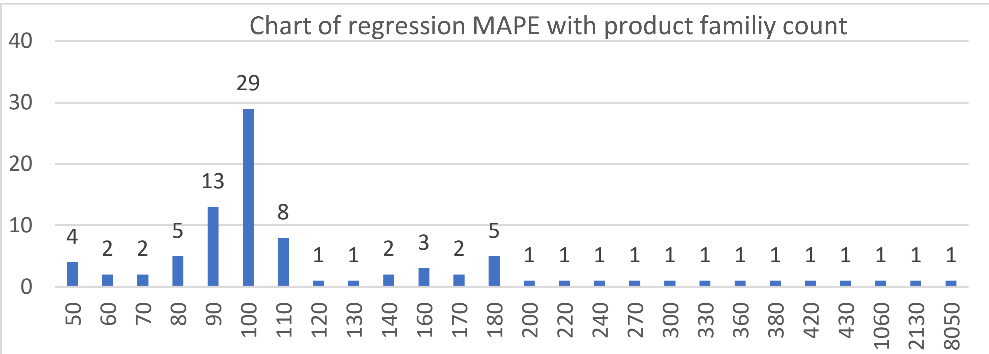


Fig6: Regression MAPE

Back to top

2. Holt's forecasting chart –

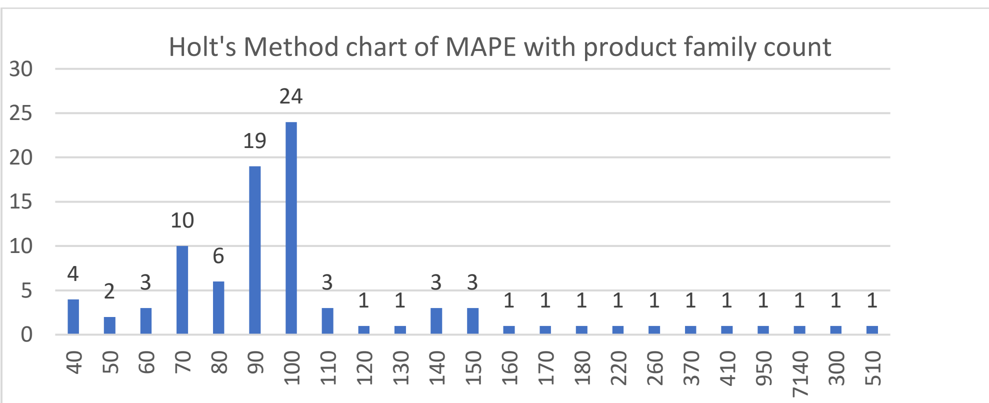


Fig7: Holt's MAPE

Back to top

Item No.– Quarter level forecasting charts

1. Exponential smoothing forecasting chart –

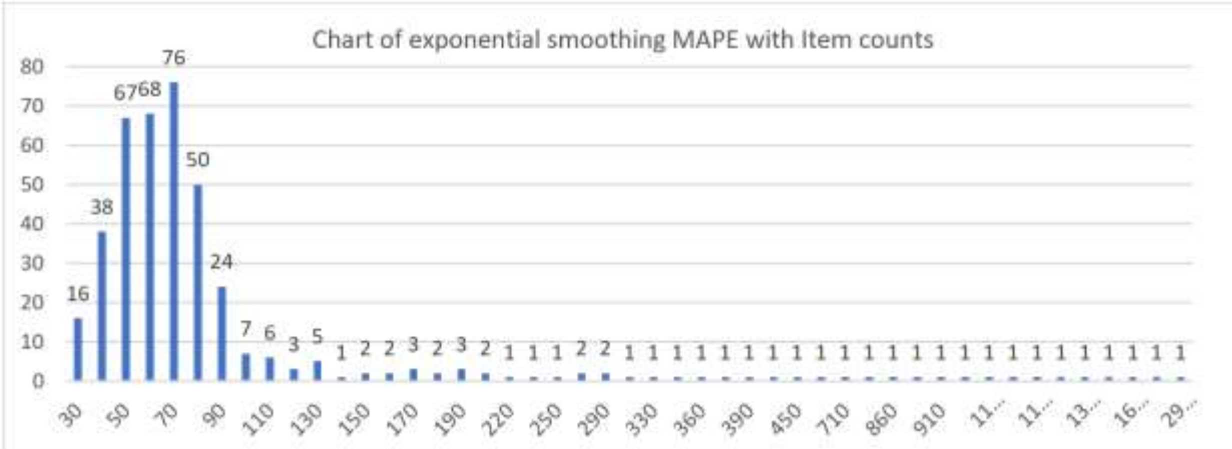


Fig7: ES chart

Back to top

2. Seasonality forecasting chart –

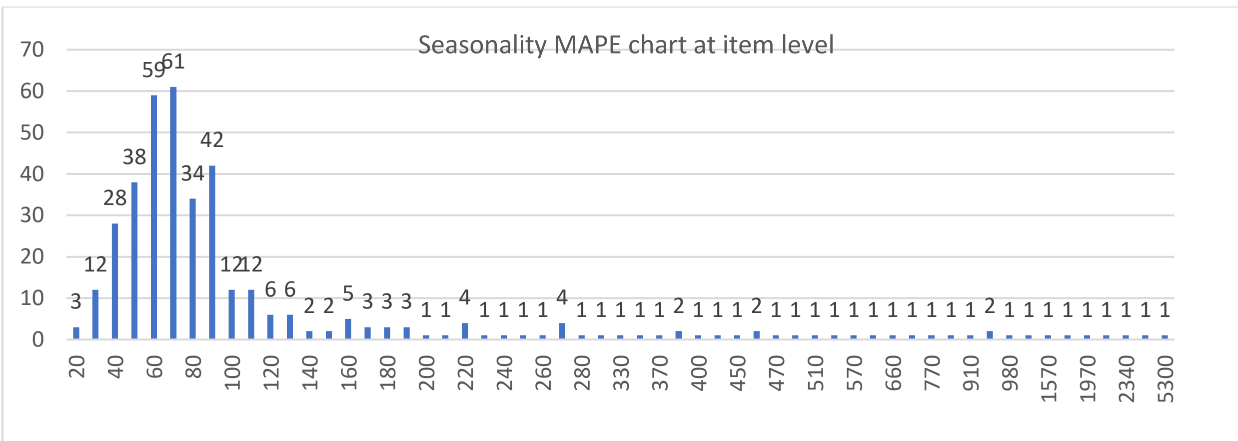


Fig8: Seasonality MAPE

Back to top

3. Regression forecasting chart –

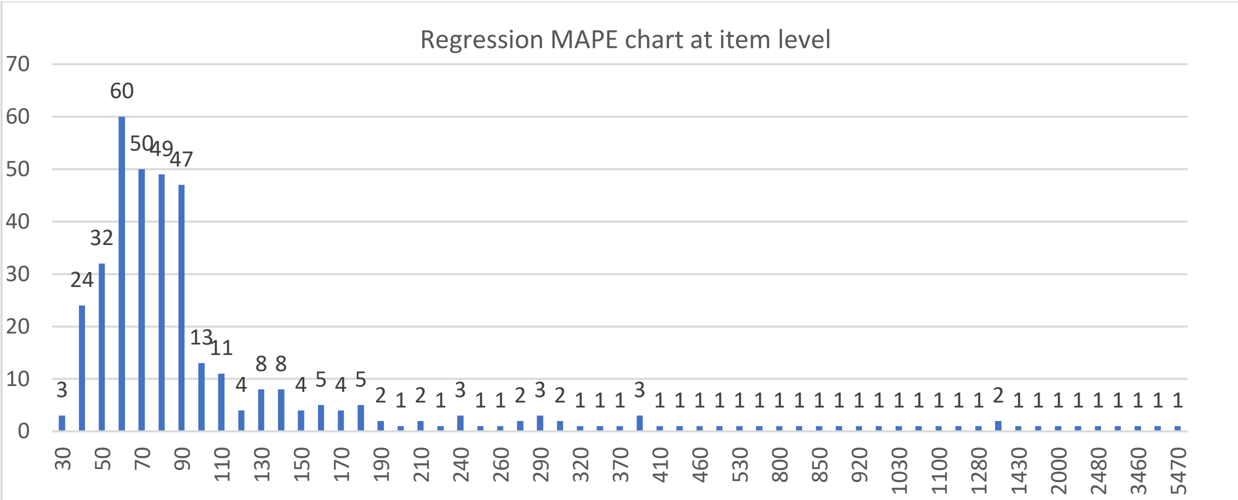


Fig9: Regression MAPE

Back to top

4. Holt's forecasting chart -

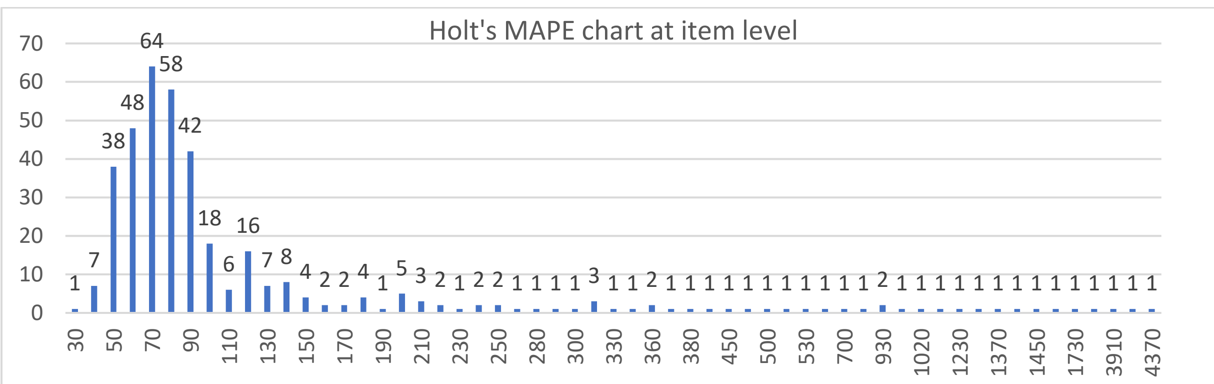


Fig10: Holt's MAPE

Back to top

VTI Dashboard –

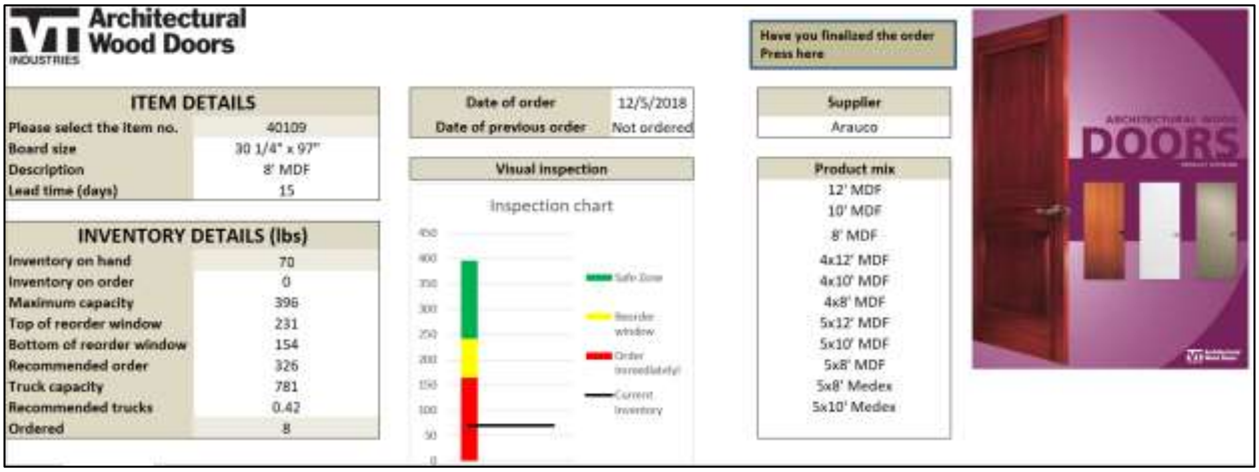


Image5: VTI Dashboard

[Back to top](#)