Managing stockouts in a health care manufacturing factory

Name - Utkarsh Balooni

Roll No. - 22ds1000021

1. Executive Summary:

The aim of this report is to provide feasible strategies to minimize stockouts and optimize inventory levels for a pharma manufacturing factory. The techniques used for this purpose are demand forecasting and safety stock analysis.

This final submission report contains a detailed explanation of the analysis carried out, the results and findings that follow and the conclusions drawn from it.

A combination of ABC-VED analysis is carried out on the data to shortlist the most profitable products. Basic pre-processing is carried out on the dataset and lead time for the selected products is calculated. Time series components for the data like trend, seasonality and noise are analysed and inferences are drawn. Demand for the products is forecasted using SARIMA, LSTM and Prophet model. Mean absolute error for all the models is analysed and future predictions for demand are made using the best model. Safety stock and reorder point for the products is calculated using appropriate formulae. All the analysis is accompanied by relevant graphical representations.

Finally, results are interpreted and practical recommendations are listed. September and October are found to be the prime production months in the factory.

The shortcomings of the analysis are also addressed and scope for improvement is provided.

2. Detailed explanation of analysis:

ABC-VED analysis – The most profitable products are shortlisted using a combination of ABC-VED analysis (explained in detail in the mid-term submission). Products which belong to both category A and category V are chosen for further analysis. The chosen products account for 86% of the total issue amount and 40% of the total issue frequency. 6 raw materials were shortlisted using ABC-VED analysis.

Lead Time calculation – Lead time is calculated as the number of days between procurement of purchase order (date of placing order) and receiving of raw materials. The average of lead time taken over all POs and converted to weekly units is used for calculations.

Demand Forecasting – The issue quantity of the raw materials is used for forecasting. Basic pre-processing is carried out on the dataset and only the relevant columns and items are selected. The data is converted into a format suitable for forecasting and is grouped by months. Here is a snapshot of the processed data for 3 raw materials.

`atenolol`

date	issue_qty
2021-02-28	600.32
2021-03-31	950.00
2021-04-30	200.00
2021-05-31	50.00
2021-06-30	150.00
2021-07-31	0.00
2021-08-31	200.00
2021-09-30	200.00
2021-10-31	250.00
2021-11-30	200.00
2021-12-31	0.00

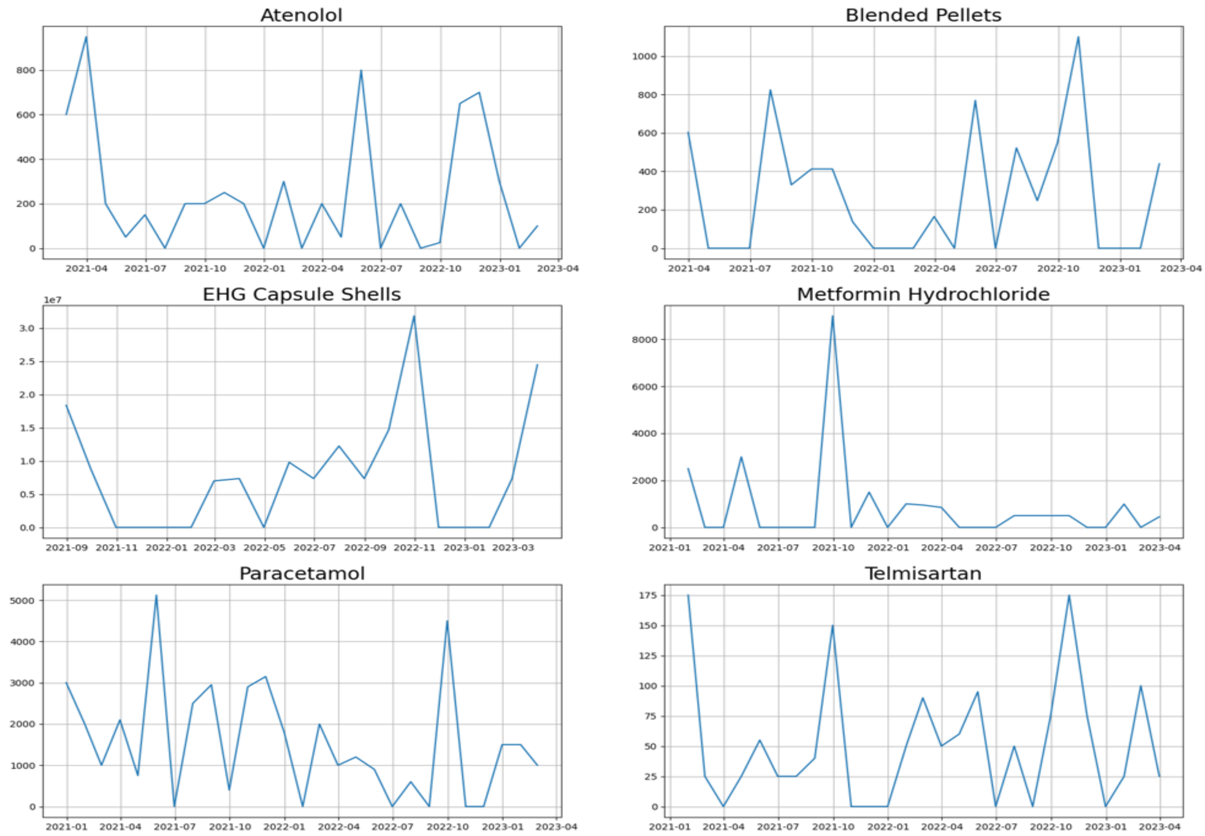
`paracetamol`

date	issue_qty
2020-12-31	3000.0
2021-01-31	2000.0
2021-02-28	1000.0
2021-03-31	2100.0
2021-04-30	750.0
2021-05-31	5125.0
2021-06-30	0.0
2021-07-31	2500.0
2021-08-31	2950.0
2021-09-30	400.0
2021-10-31	2900.0

`capsule`

date	issue_qty
2021-08-31	18375000.0
2021-09-30	8820000.0
2021-10-31	0.0
2021-11-30	0.0
2021-12-31	0.0
2022-01-31	0.0
2022-02-28	7000000.0
2022-03-31	7350000.0
2022-04-30	0.0
2022-05-31	9800000.0
2022-06-30	7350000.0

Line plot for the issue_qty of all shortlisted products is shown below:



To check for trend, seasonality and outliers in the data, we use seasonal_decompose package from statsmodels API of python. Trend, seasonality and residuals for 5 raw materials is plotted.

Seasonal decomposition for "BLENDED PELLETS OF RABEPRAZOLE EC 20MG & DOMPERIDONE SR 30MG" could not be carried out since the number of data points available were not enough.

```
ValueError: x must have 2 complete cycles requires 24 observations. x only has 20 observation(s)
```

We will be comparing 3 ML models based on (normalized) mean absolute error on training data and select the best one for making predictions:

- SARIMA (Seasonal Autoregressive Integrated Moving Average),

- LSTM (Long Short-Term Memory), and
- Facebook Prophet model.

SARIMA: In order to fit SARIMA model, we need to ensure that the data is stationary. To check if the data is stationary, we use the ADF test (p-value<=0.05 cut-off). Data for all raw materials except "TELMISARTAN" was stationary.

```
adfuller_test(tel['issue_qty'])
```

```
ADF Test Statistic : 0.85674015973261
p-value : 0.99249898957295
#Lags Used : 9
Number of Observations Used : 17
weak evidence against null hypothesis, time series has a unit root, indicating it is non-stationary
```

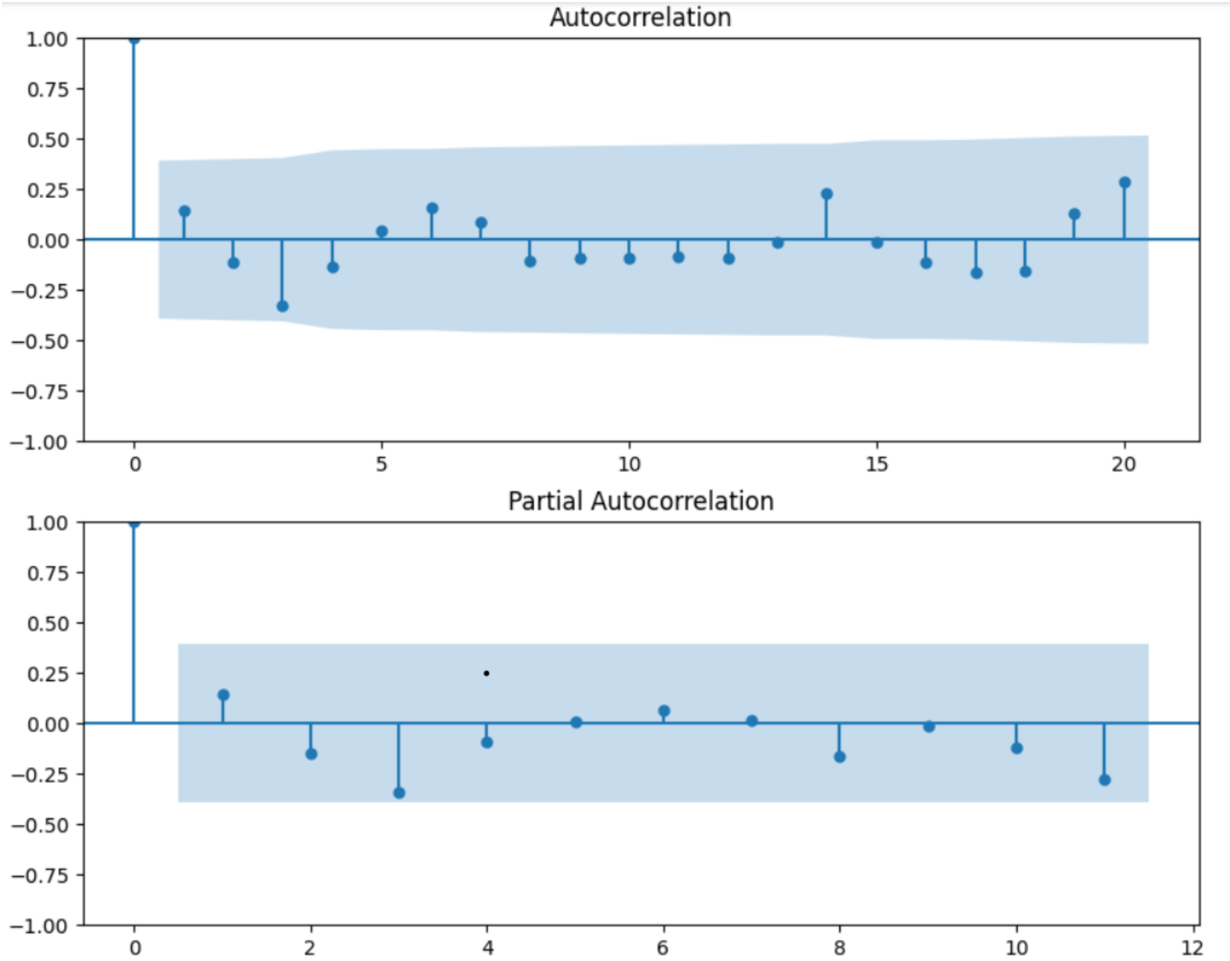
We use 2^{nd} order differencing to make the data for Telmisartan stationary.

```
tel_diff=np.abs(tel.diff(periods=2).dropna())
adfuller_test(tel_diff['issue_qty'])
```

```
ADF Test Statistic : -3.849277541893278
p-value : 0.0024415236580231756
#Lags Used : 9
Number of Observations Used : 15
strong evidence against the null hypothesis(Ho), reject the null hypothesis. Data has no unit root and is stationary
```

Parameters of the SARIMA model are estimated for each dataset using the autocorrelation and partial autocorrelation plots.

Plot of ACF and PACF for "ATENOLOL" is shown below:



Finally, we fit the SARIMA model on all datasets and calculate the mean absolute error for each raw material.

LSTM: In order to fit LSTM, we first scale the data using MinMaxScaler from the sklearn library. Then we convert the scaled data frame to a Time series Generator object (window size = 3). The time series generator converts the data points into a format like this:

```
[[[1st pt], [2nd pt], [3rd pt]], [4th pt]]
[[[2nd pt], [3rd pt], [4th pt]], [5th pt]]
[[[3rd pt], [4th pt], [5th pt]], [6th pt]]
......
```

```
generator[0]
```

```
(array([[[0.63191579],
         [1.        ],
         [0.21052632]]]),
 array([0.05263158]))
```

Next, we compile and fit a neural network consisting of 1 Input Layer, 1 LSTM layer, 1 Dense layer and 1 output layer on the generator object. A small value of learning rate is used to ensure less fluctuations in the loss.

```
#define the model
model = Sequential()
model.add(InputLayer((3, 1)))
model.add(LSTM(64))
model.add(Dense(8, 'relu'))
model.add(Dense(1, 'linear'))
model.compile(loss=MeanSquaredError(), optimizer=Adam(learning_rate=0.0001), metrics=[RootMeanSquaredError()])

#fit the model
model.fit(generator, epochs=40, verbose=0)
```

Mean absolute error of the model is calculated and reported for all raw materials.

Prophet model: The columns of the dataset are renamed 'ds' (date) and 'y' (issue_qty) before feeding the dataset to the prophet model.

	ds	y
0	2021-02-28	600.32
1	2021-03-31	950.00
2	2021-04-30	200.00
3	2021-05-31	50.00
4	2021-06-30	150.00
5	2021-07-31	0.00
6	2021-08-31	200.00

The model is fit and mean absolute error is calculated.

Finally, we compare the MAE of the 3 models and report the predictions of the best model.

Safety stock and Reorder point calculation:

Safety stock and reorder point is calculated using actual demand data for the products. Lead time is converted into units of weeks. Average weekly demand and standard deviation of demand during lead time (in weeks) is calculated for all shortlisted products.

The formula used for safety stock calculation is:

$$\text{Safety Stock} = Z \times \sigma_{demand} \times \sqrt{L}$$

where,

Z – number of standard deviations required to achieve the desired service level. Can be found in a Standard Normal distribution table.

σ_{demand} – Standard Deviation of demand during lead time.

L – Lead time in weeks.

We will calculate safety stock for a desired service level of 90%. Using a standard normal distribution table, it can be obtained that <u>Z = 1.28 for a 90% service level</u>.

Formula used for reorder point calculation is:

Reorder Point = (Average weekly demand × Lead time) + (safety stock)

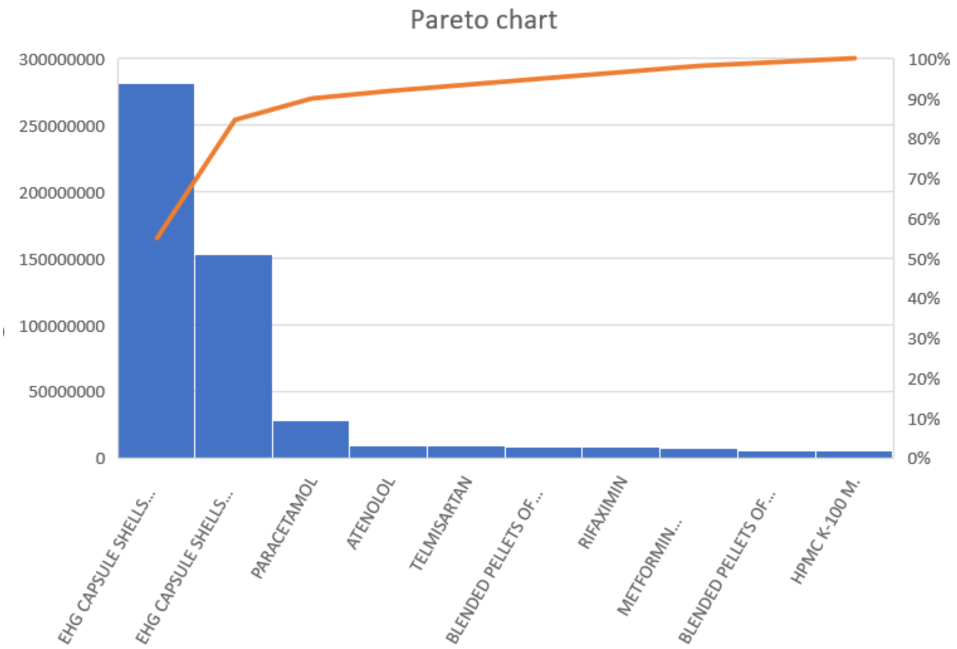
<u>Note</u> – The unit of lead time used in mid-term submission was months. However, this had to be changed to weeks since most of the products had <u>shorter lead times (≈ 0 months).</u> Taking lead time ≈ 0 was interfering with the standard deviation of demand during lead time calculation.

3. Results and Findings:

ABC-VED analysis: As a result of ABC-VED analysis 6 raw materials were shortlisted.

	Shortlisted Products
1	ATENOLOL
2	BLENDED PELLETS OF RABEPRAZOLE EC 20MG & DOMPERIDONE SR 30MG
3	EHG CAPSULE SHELLS (PINK/CT #2)
4	METFORMIN HYDROCHLORIDE
5	PARACETAMOL
6	TELMISARTAN

Pareto chart using total issue amount as values is plotted and it is evident from the chart that the data follows the pareto (80-20) principle.

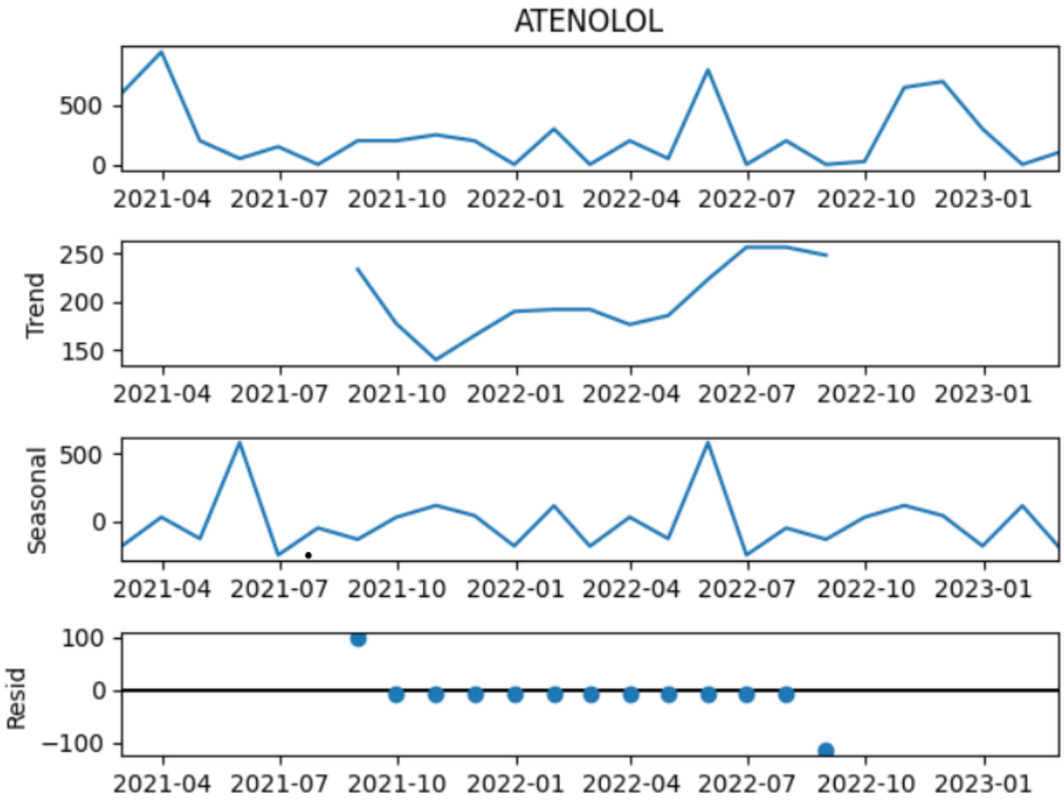


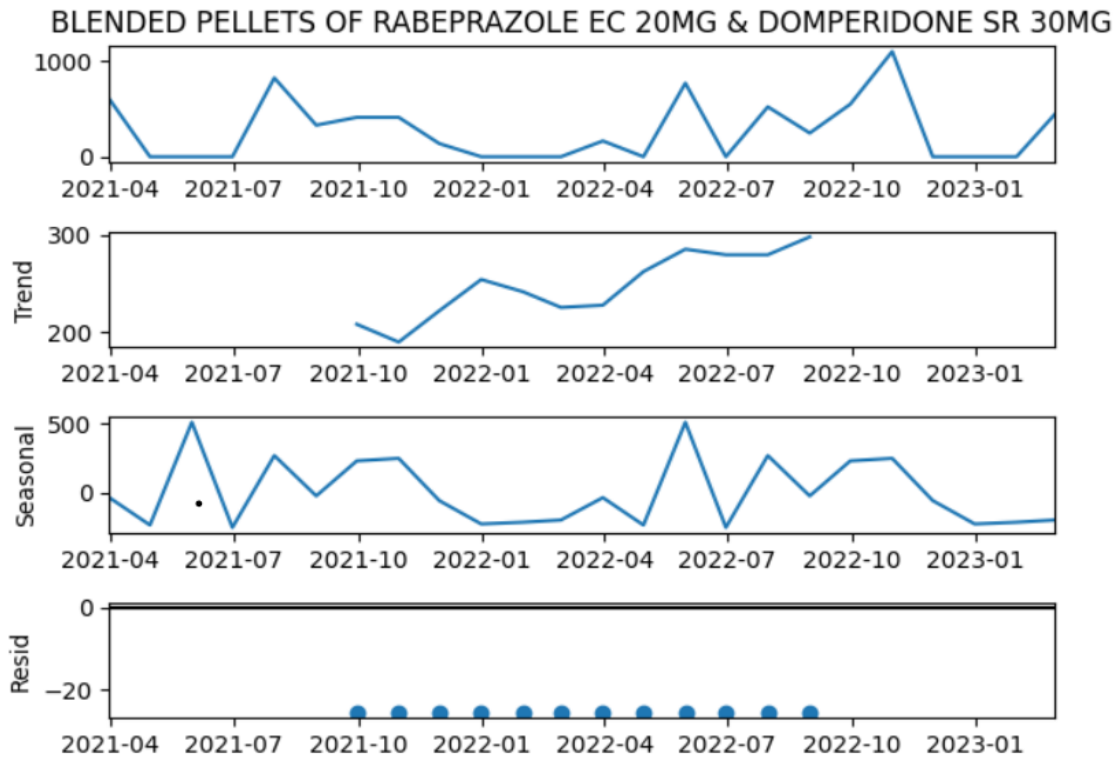
Lead time: Average lead time in weeks for the shortlisted products is shown below:

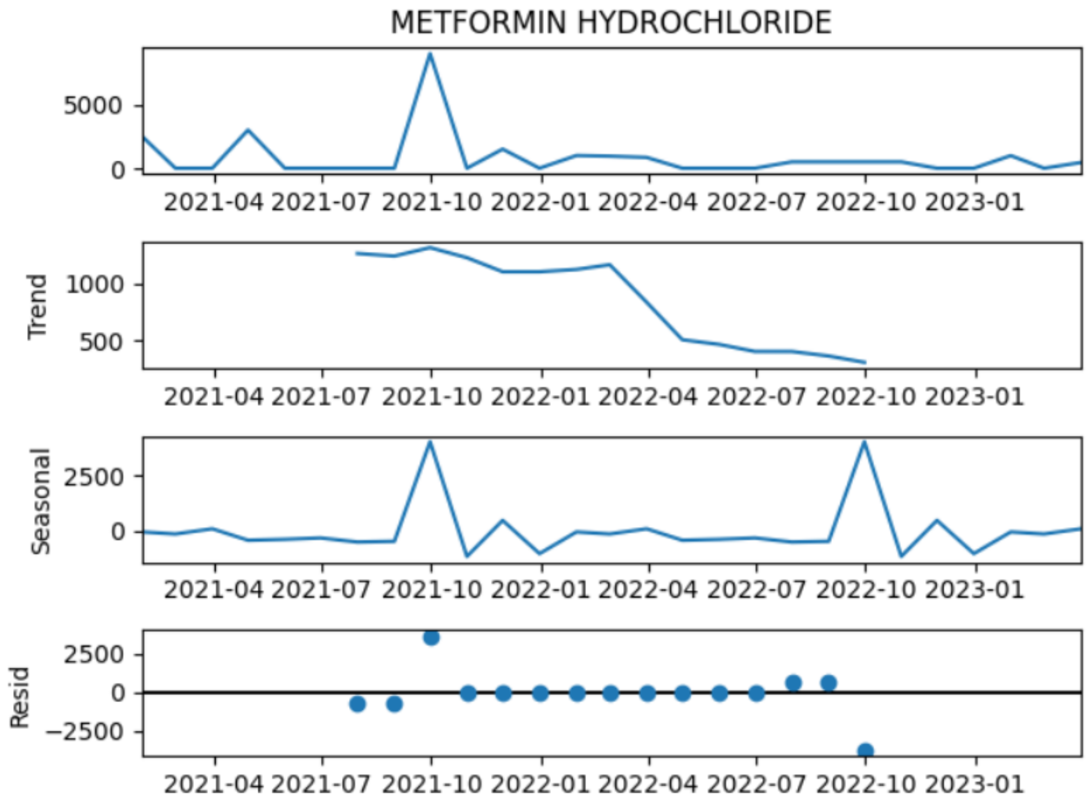
Item Name	Average lead time (weeks)
ATENOLOL	5.35
EHG CAPSULE SHELLS (PINK/CT #2)	2.83
TELMISARTAN	2.62
METFORMIN HYDROCHLORIDE	2.32
BLENDED PELLETS OF RABEPRAZOLE EC 20MG & DOMPERIDONE SR 30MG	2.10
PARACETAMOL	1.92

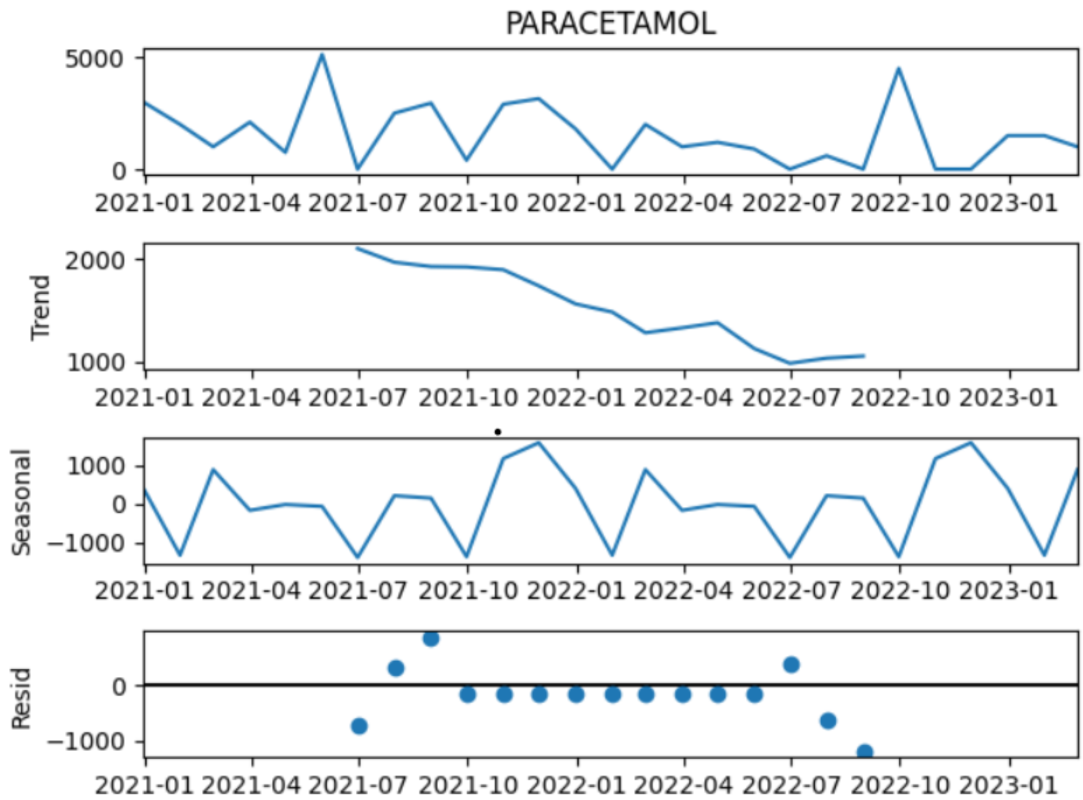
Demand Forecasting:

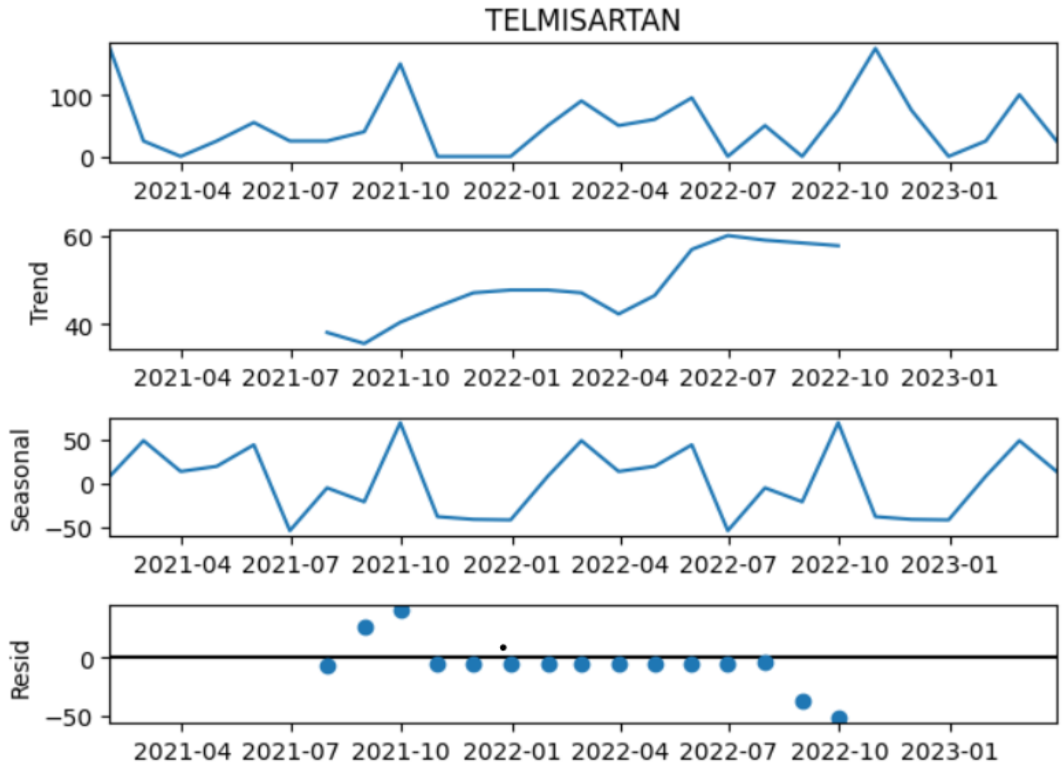
Seasonal decomposition of the shortlisted products is shown below:





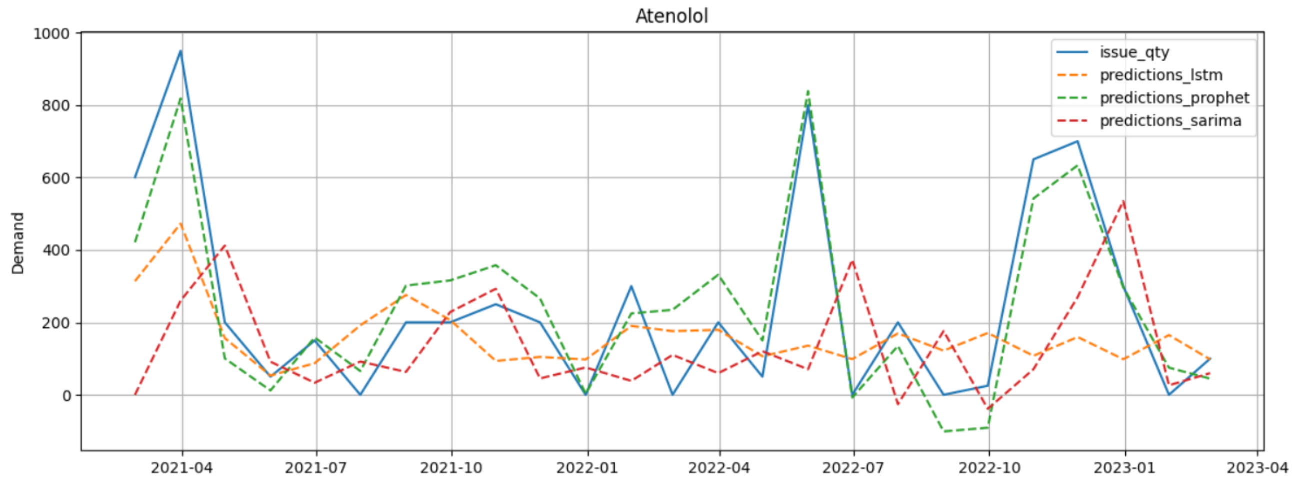

METFORMIN HYDROCHLORIDE
Trend
Seasonal
Resid


PARACETAMOL
Trend
Seasonal
Resid

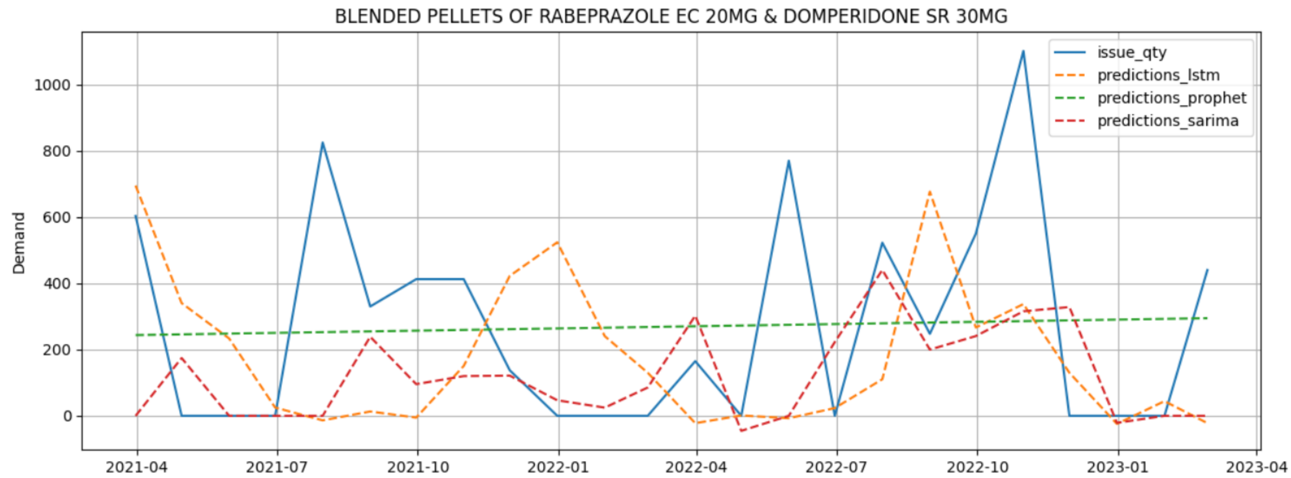


- As seen from the trend charts, Atenolol, Blended pellets, and Telmisartan depict an upward trend in demand, while Metformin Hydrochloride, and Paracetamol show a downward trend.
- The seasonal charts show that all the products exhibit a seasonality pattern of approximately 12 months. (yearly).
- It is clear from the residue plots that the data for all products contains noise.

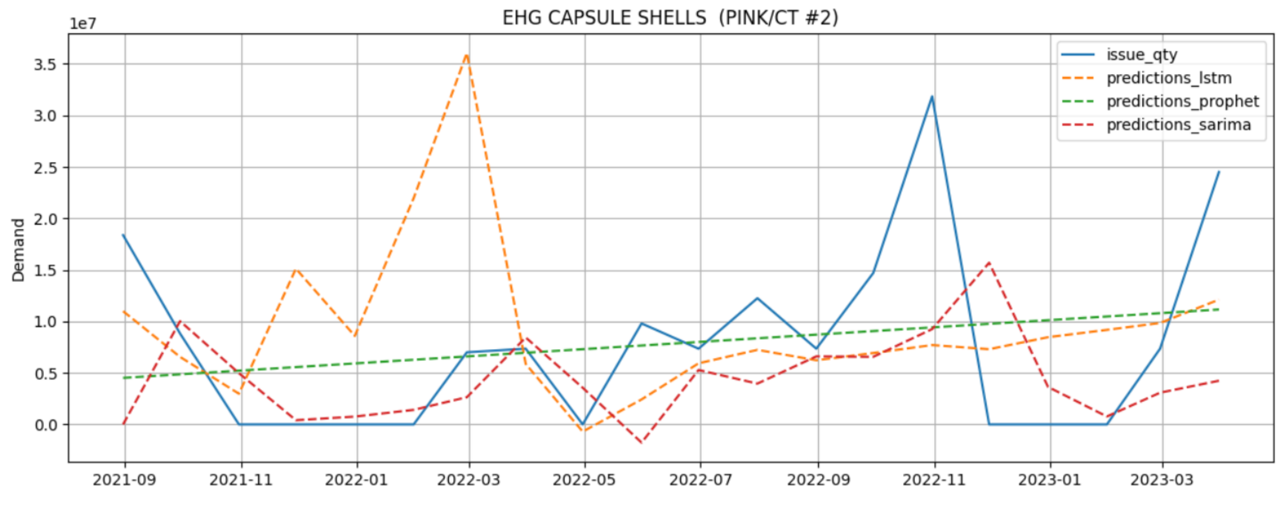
The predictions on training data for all the models is shown below (for all RMs):



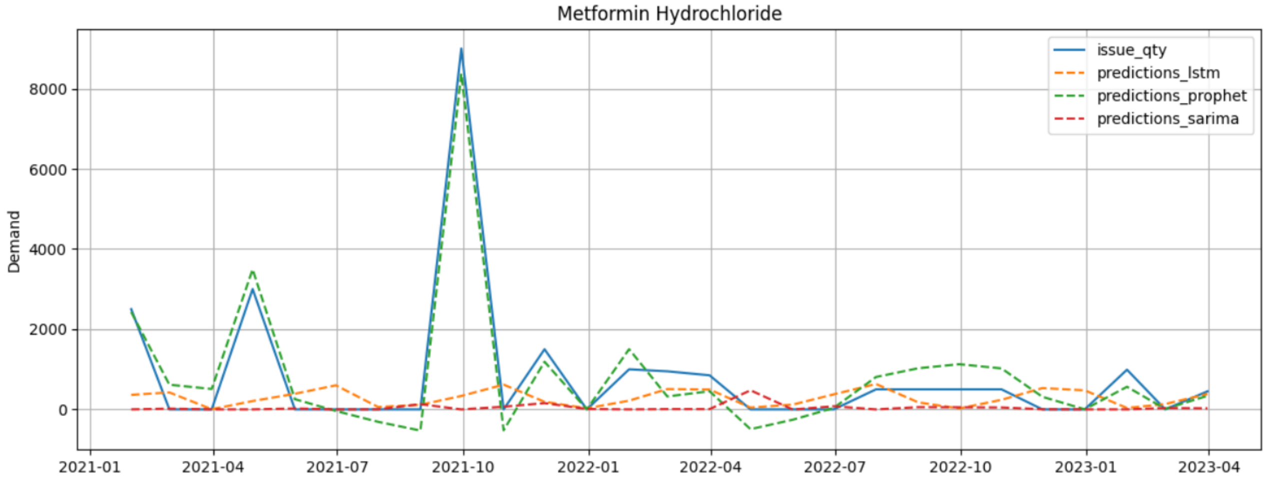
Atenolol
issue_qty
predictions_lstm
predictions_prophet
predictions_sarima
Demand
1000
800
600
400
200
0
2021-04
2021-07
2021-10
2022-01
2022-04
2022-07
2022-10
2023-01
2023-04

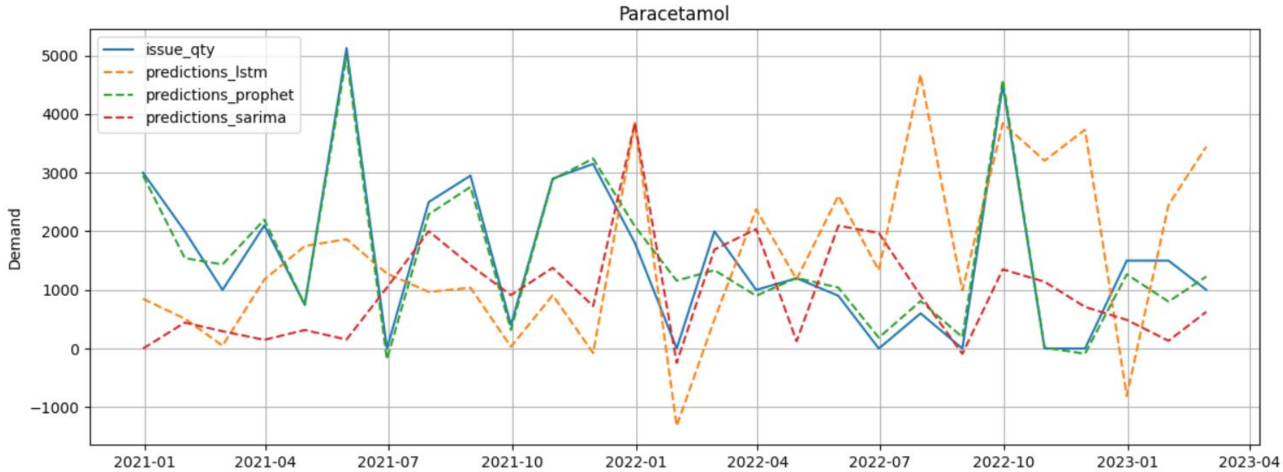


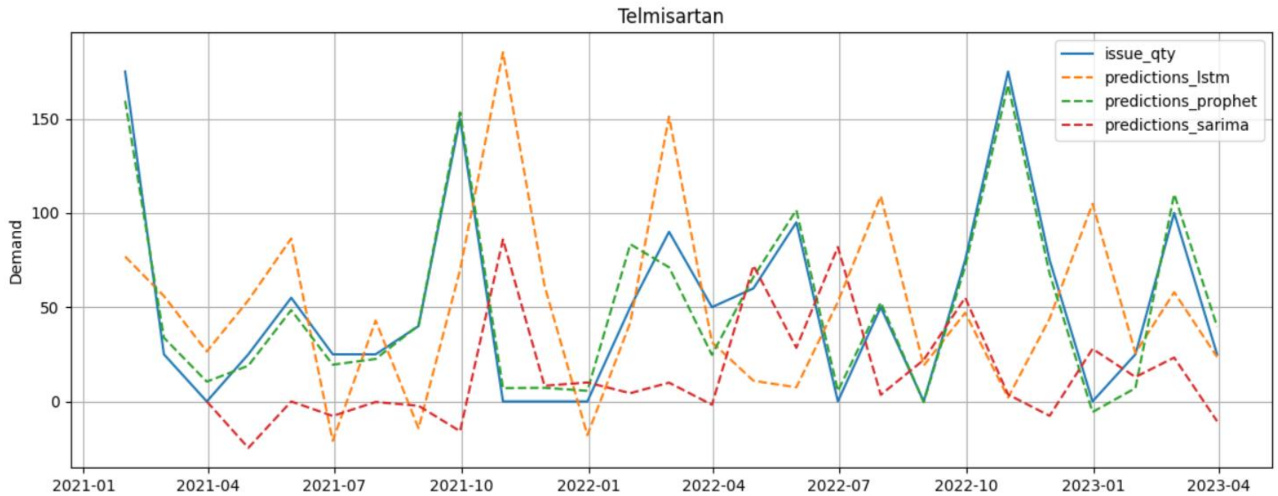
BLENDED PELLETS OF RABEPRAZOLE EC 20MG & DOMPERIDONE SR 30MG
issue_qty
predictions_lstm
predictions_prophet
predictions_sarima
Demand
1000
800
600
400
200
0
2021-04
2021-07
2021-10
2022-01
2022-04
2022-07
2022-10
2023-01
2023-04



EHG CAPSULE SHELLS (PINK/CT #2)
1e7
issue_qty
predictions_lstm
predictions_prophet
predictions_sarima
Demand
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2021-09
2021-11
2022-01
2022-03
2022-05
2022-07
2022-09
2022-11
2023-01
2023-03







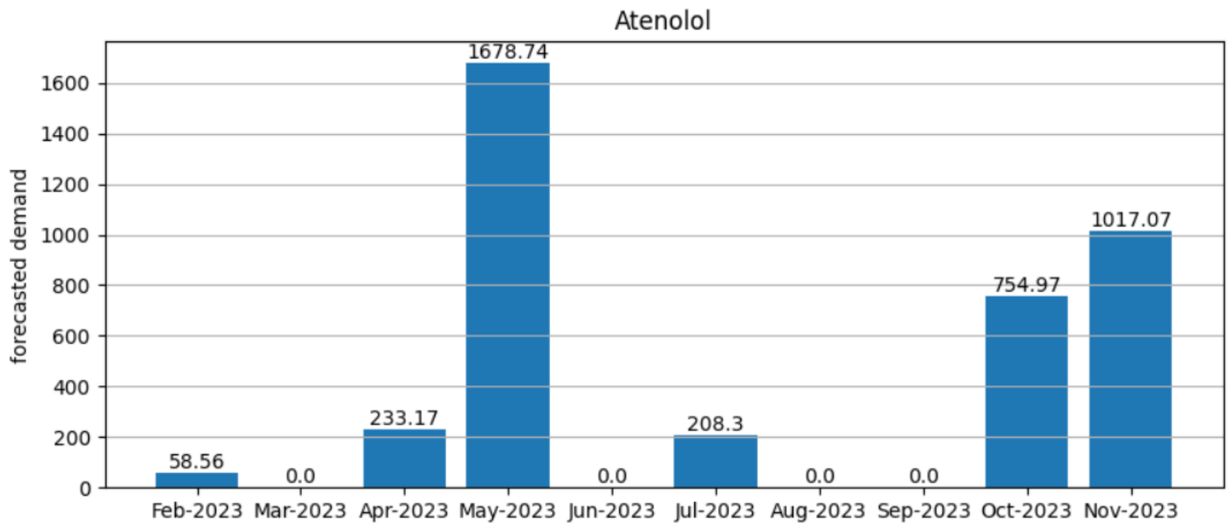
Mean absolute errors for all RMs are normalized by dividing with the mean of the issue quantity. This is done as a sanity check and also to ensure that all errors are on the same scale.

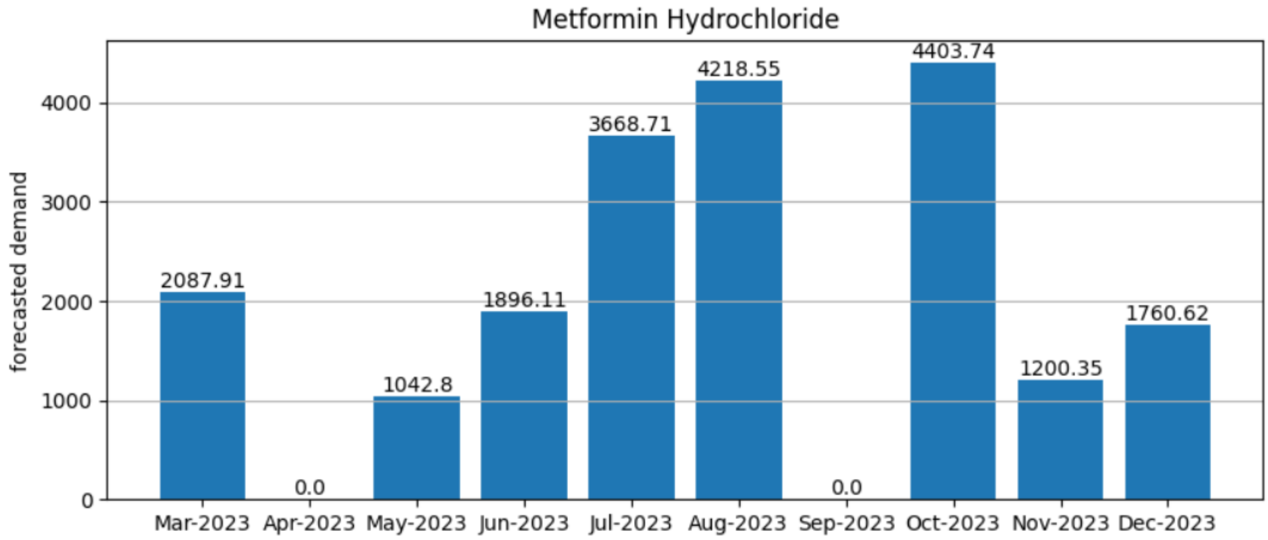
Table for MAE is shown below:

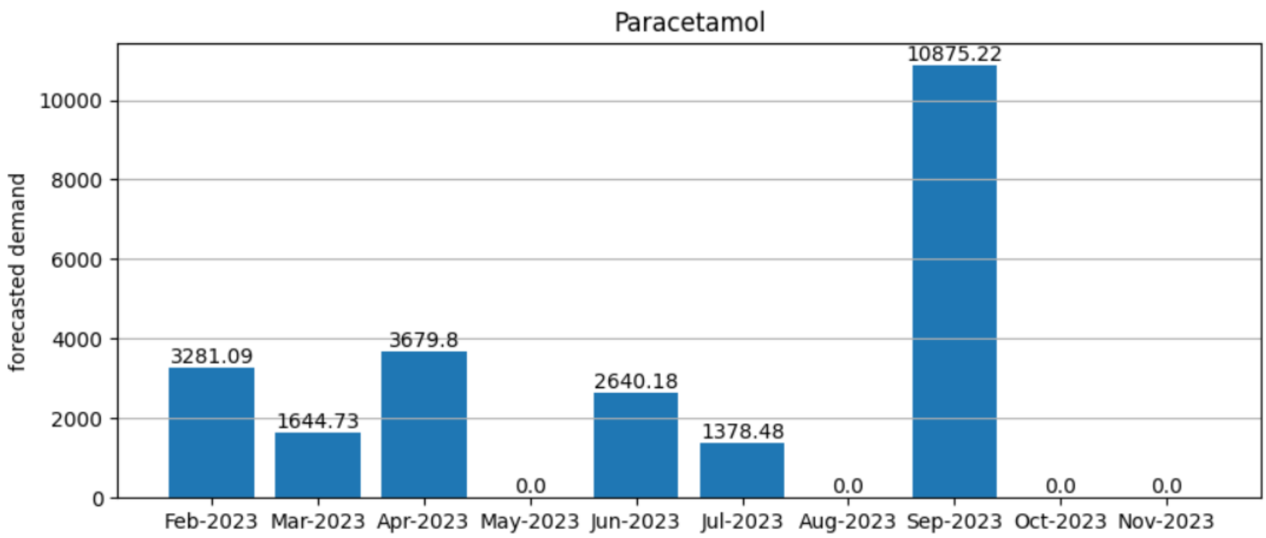
Item Name	Model		
	LSTM	SARIMA	Prophet
Atenolol	0.785	0.922	0.342
Blended Pellets	0.995	0.981	0.998
EHG Capsule shells	0.817	0.856	0.843
Metformin Hydorchloride	1.186	0.977	0.425
Paracetamol	0.705	0.819	0.149
Telmisartan	0.76	0.912	0.175

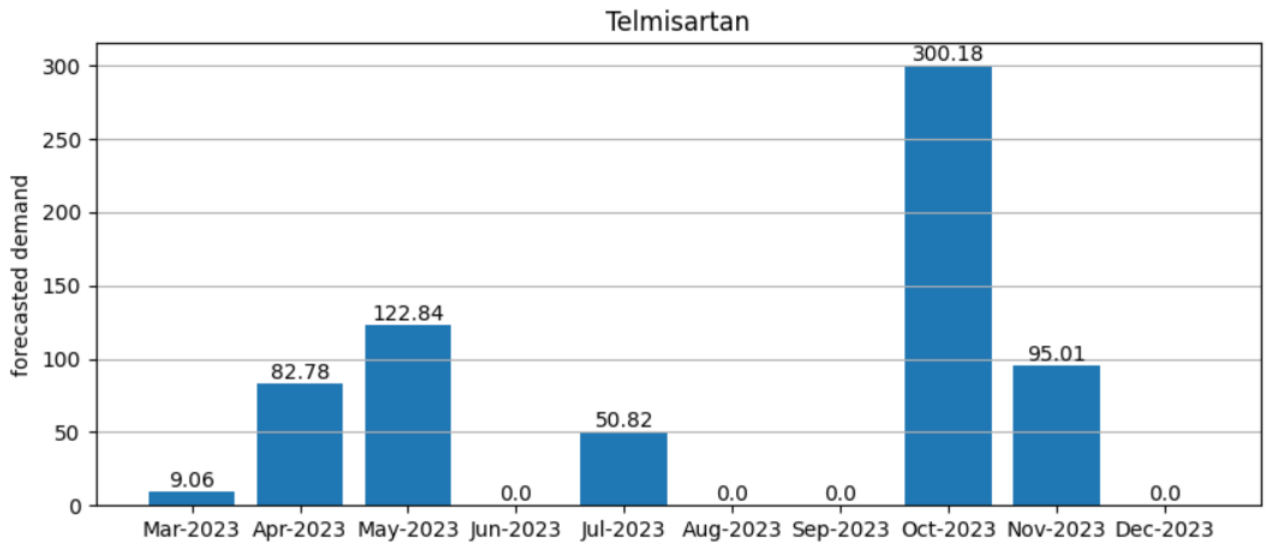
It is clear from the table that the prophet model performs significantly better than LSTM and SARIMA.

Finally, we select the model with the least MAE and report our predictions using it. We will not be forecasting demand for Blended pellets and EHG Capsule Shells as the MAE for them is quite high.




Metformin Hydrochloride
forecasted demand
0
1000
2000
3000
4000
2087.91
0.0
1042.8
1896.11
3668.71
4218.55
0.0
4403.74
1200.35
1760.62
Mar-2023
Apr-2023
May-2023
Jun-2023
Jul-2023
Aug-2023
Sep-2023
Oct-2023
Nov-2023
Dec-2023

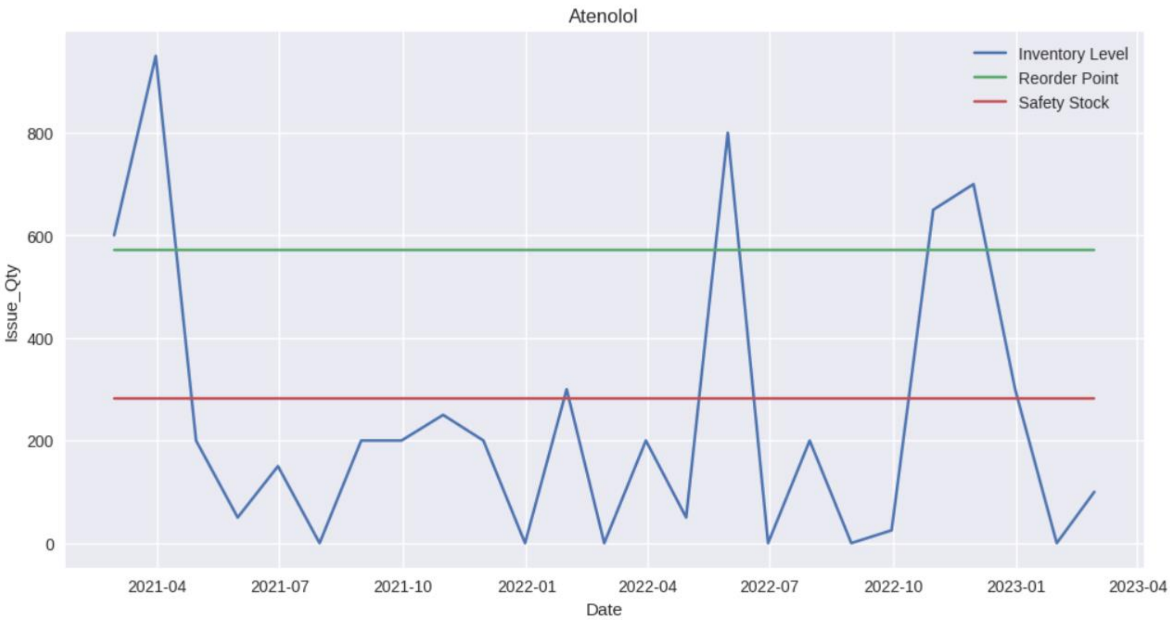

Paracetamol
forecasted demand
0
2000
4000
6000
8000
10000
3281.09
1644.73
3679.8
0.0
2640.18
1378.48
0.0
10875.22
0.0
0.0
Feb-2023
Mar-2023
Apr-2023
May-2023
Jun-2023
Jul-2023
Aug-2023
Sep-2023
Oct-2023
Nov-2023


Telmisartan
forecasted demand
0
50
100
150
200
250
300
9.06
82.78
122.84
0.0
50.82
0.0
0.0
300.18
95.01
0.0
Mar-2023
Apr-2023
May-2023
Jun-2023
Jul-2023
Aug-2023
Sep-2023
Oct-2023
Nov-2023
Dec-2023

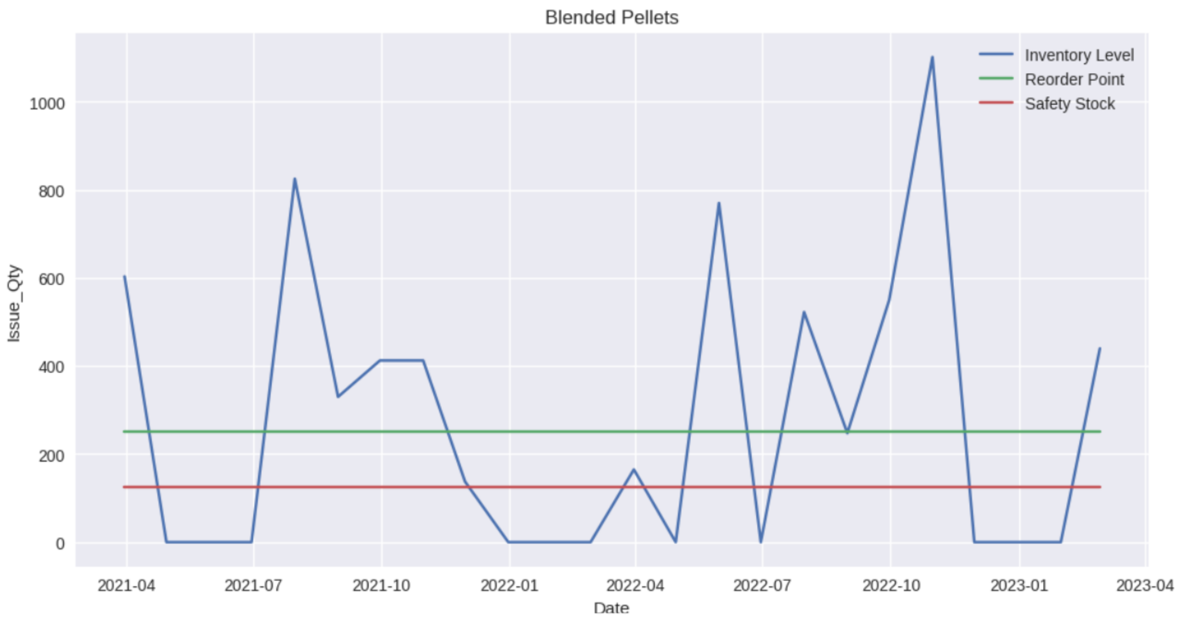
Safety Stock and Reorder Point:

Safety stock and reorder point for all RMs are as shown:

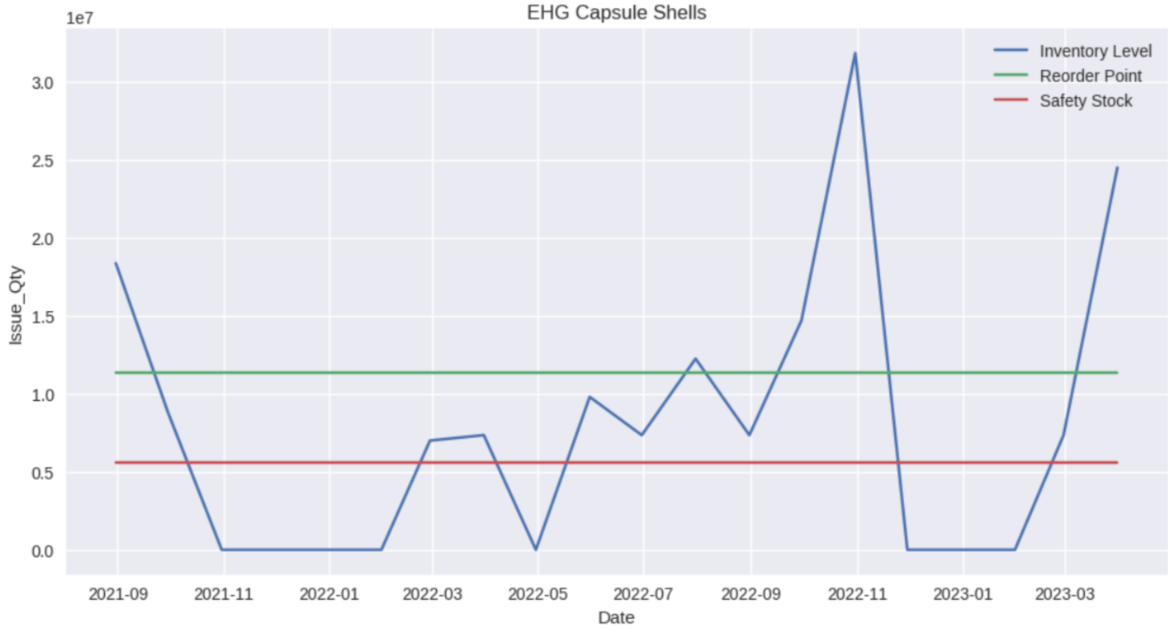
```
Atenolol
Average weekly demand: 57.79 units/week
Lead Time: 5 weeks
Standard deviation of demand during lead time: 98.49 units
Safety stock at 90% service level: 281.9 units
Reorder point: 570.83 units
```



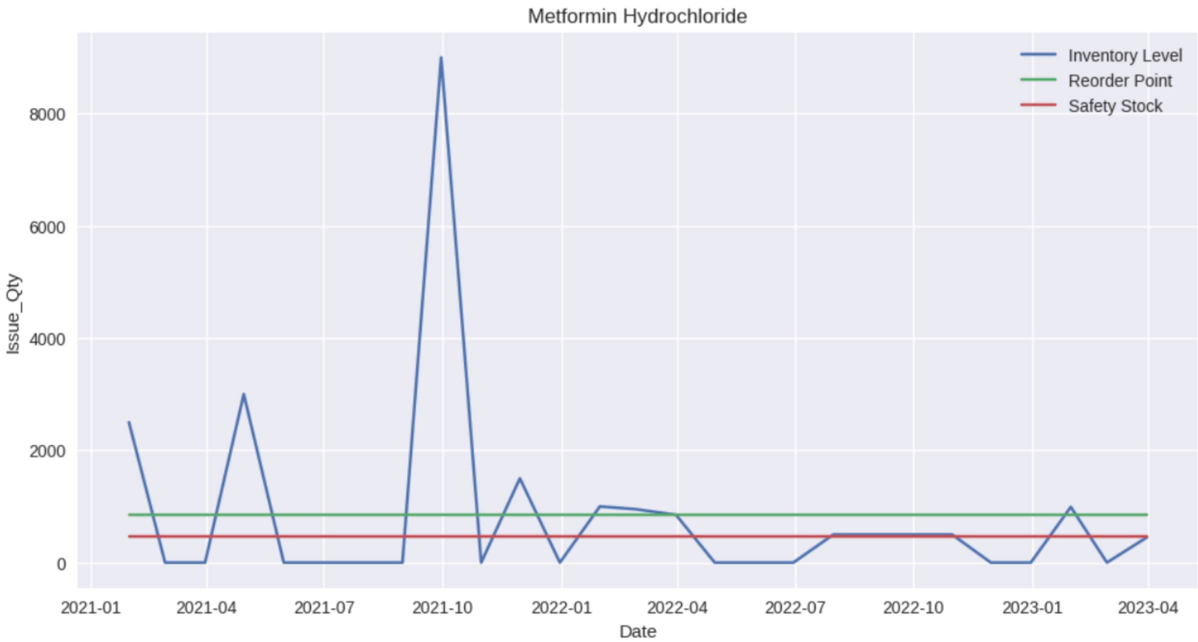

```
Blended Pellets
Average weekly demand: 63.27 units/week
Lead Time: 2 weeks
Standard deviation of demand during lead time: 68.82 units
Safety stock at 90% service level: 124.58 units
Reorder point: 251.13 units
```



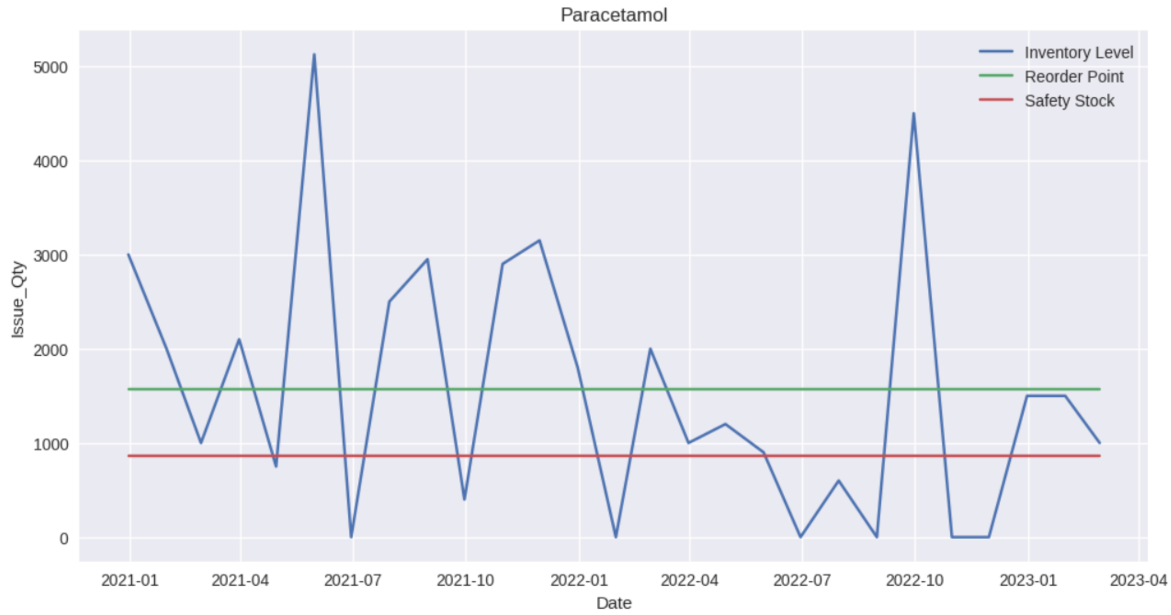
```
EHG Capsule Shells
Average weekly demand: 1910914.63 units/week
Lead Time: 3 weeks
Standard deviation of demand during lead time: 2521842.61 units
Safety stock at 90% service level: 5590988.19 units
Reorder point: 11323732.09 units
```



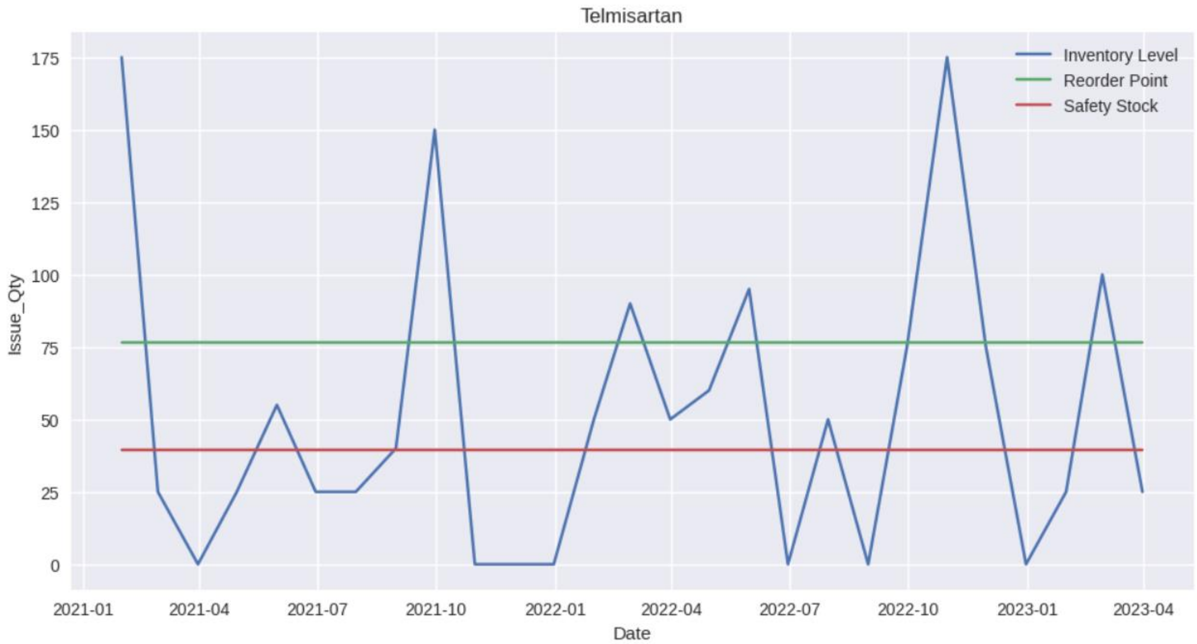

```
Metformin Hydrochloride
Average weekly demand: 195.09 units/week
Lead Time: 2 weeks
Standard deviation of demand during lead time: 253.62 units
Safety stock at 90% service level: 459.1 units
Reorder point: 849.28 units
```



```
Paracetamol
Average weekly demand: 354.87 units/week
Lead Time: 2 weeks
Standard deviation of demand during lead time: 473.52 units
Safety stock at 90% service level: 857.16 units
Reorder point: 1566.91 units
```




```
Telmisartan
Average weekly demand: 12.3 units/week
Lead Time: 3 weeks
Standard deviation of demand during lead time: 17.84 units
Safety stock at 90% service level: 39.55 units
Reorder point: 76.45 units
```



4. Interpretation of Results:

- The ABC-VED analysis reveals that out of the 183 raw materials only 6 have a significant contribution towards the profitability of the factory.
- Issue quantity for "Metformin Hydrochloride" and "Paracetamol" is declining as seen from the trend charts.
- Demand for the products peaks around the months September-October. (May for Atenolol)
- Demand for "Metformin Hydrochloride" is expected to rise and follow an upward trend in the upcoming period.
- Values of safety stock and reorder point for the raw materials are self-evident.

5. Recommendations:

- The factory must harness its resources to manage the most profitable products and monitor their sales routinely.
- Use the forecasted demand to plan production more efficiently.
- Increase manpower and resource allocation during the months of September-October.
- Keep a close eye on the product "Metformin Hydrochloride".
- Analyse the values of safety stock provided and consider keeping a buffer stock to prevent stockouts.
- Place future orders using the value of reorder point provided.

6. Limitations of the Analysis:

Although the data was collected for 2.5 years, the data available for issue quantity of individual raw materials was still not enough for accurate demand forecasting. The data also contained a lot of noise which would have impacted the forecast.

The value of safety stock provided needs further analysis before implementation.