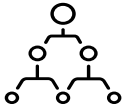
Ribbed Smoked Sheet Forecasting

Suwicha saeling

Table of Contents

01 PROBLEM STATEMENT

02 EDA

03 MODELING and EVALUATION 

04 CONCLUSION and RECOMMENDATION

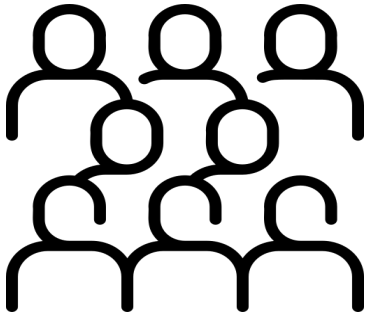
Who Am I

Suwicha Saeling
DSI STUDENT
Son of Rubber Shop

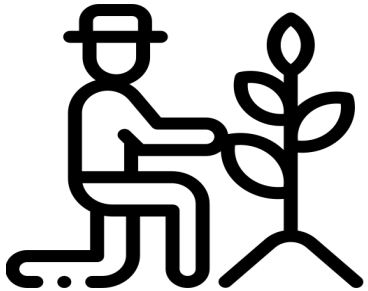
Junior Data Scientist at
Rubber cooperative

Part 1:
Problem Statement

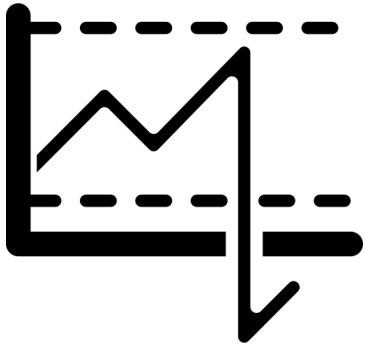
Why I chose Rubber



Households
2.9M



Rubber Plantation owners
1M



Facing Rubber Price Fluctuations

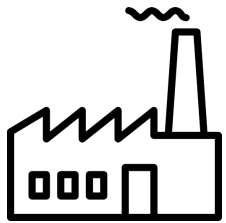
Why I chose Rubber

The Rubber price fluctuated due to various world events such as:

1. Downward trending demand for medical gloves decreased during the pandemic.
2. Semiconductor chips in short supply which dramatically curtailed production in the auto industry. (source)

Rubber Cooperative Cycle

The problem in Rubber cooperative is using **Instinct-Decision** based from the experience.



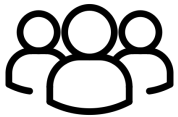
Manufacturer

Loss

Sold

Rubber cooperative

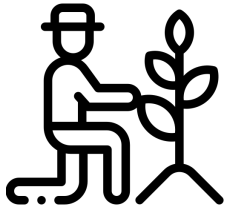
Dividend



Member of Rubber cooperative

Lower price

Sold to cooperative



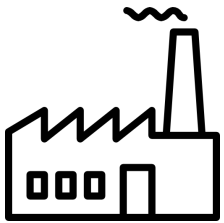
Rubber Plantation Owner

Problem Statement

"Rubber cooperative wants a model that forecasts **Ribbed smoked sheet** prices to estimate trend 1 month out to increase chance of success speculating in the Ribbed smoked sheet market"

Rubber Cooperative Cycle

If we can change to **Data-Driven-Decision.**



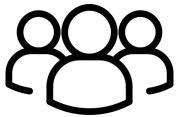
Manufacturer

Profit

Sold

Rubber cooperative

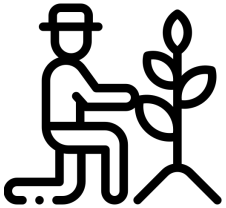
Dividend



Member of Rubber cooperative

Higher price

Sold to cooperative



Rubber Plantation Owner

Direct Variables

Latex

Rubber Sheet

Cup lump

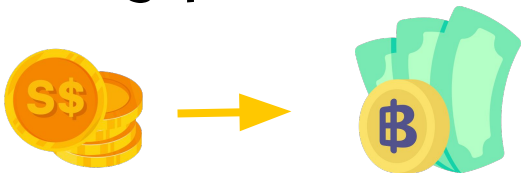
Ribbed Smoked Sheet

- Tocom
- FOB
- Sicom

Indirect Variables

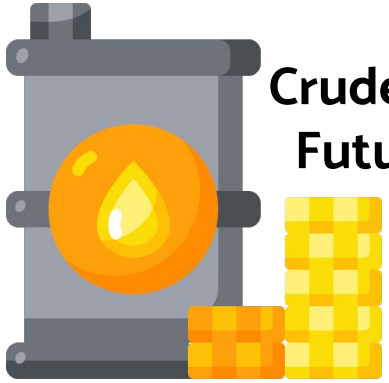
Currency **Appreciation** & **Depreciation** also impact the Rubber price because they are related to **imports** and **exports**

Singaporean Dollar



U.S. Dollar

Yen

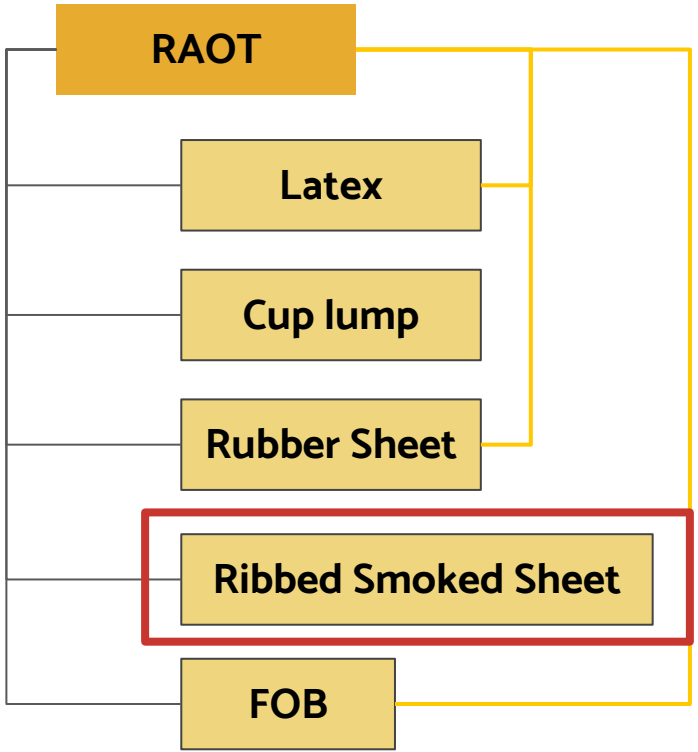


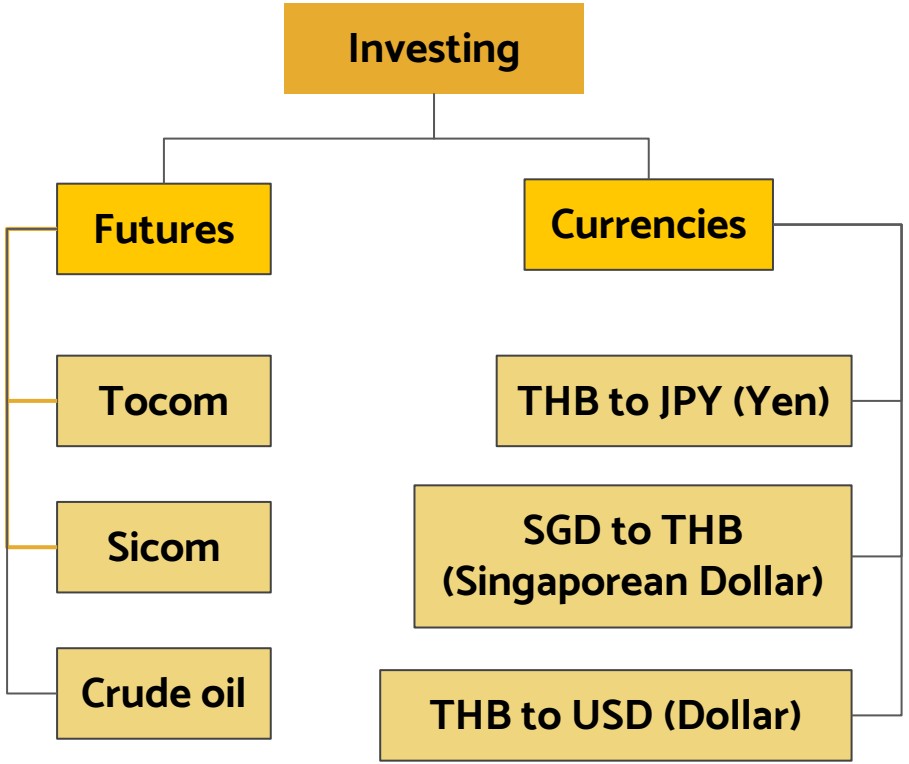
Crude oil Future

Crude oil can make **Synthetic Rubber**

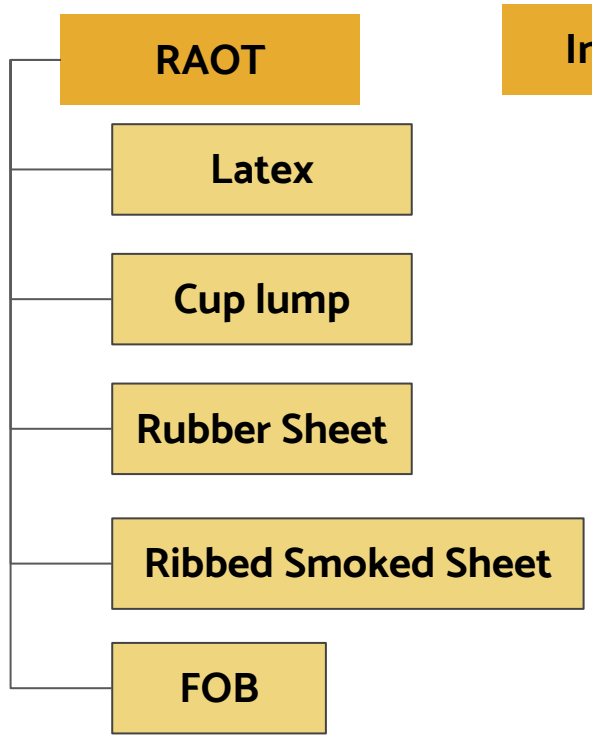
Data Gathering

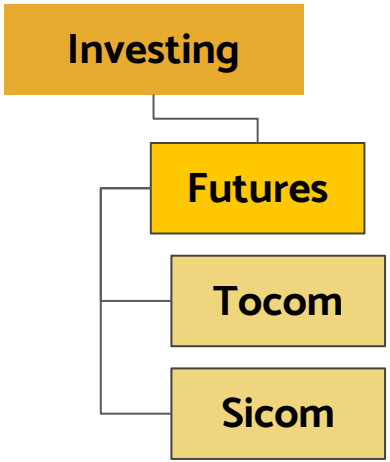
Year ranges in data
- 2017-2022
- 2010-2022



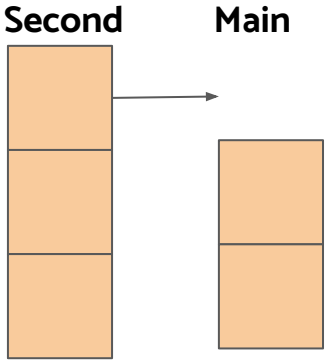


Data Cleaning - Missing Values





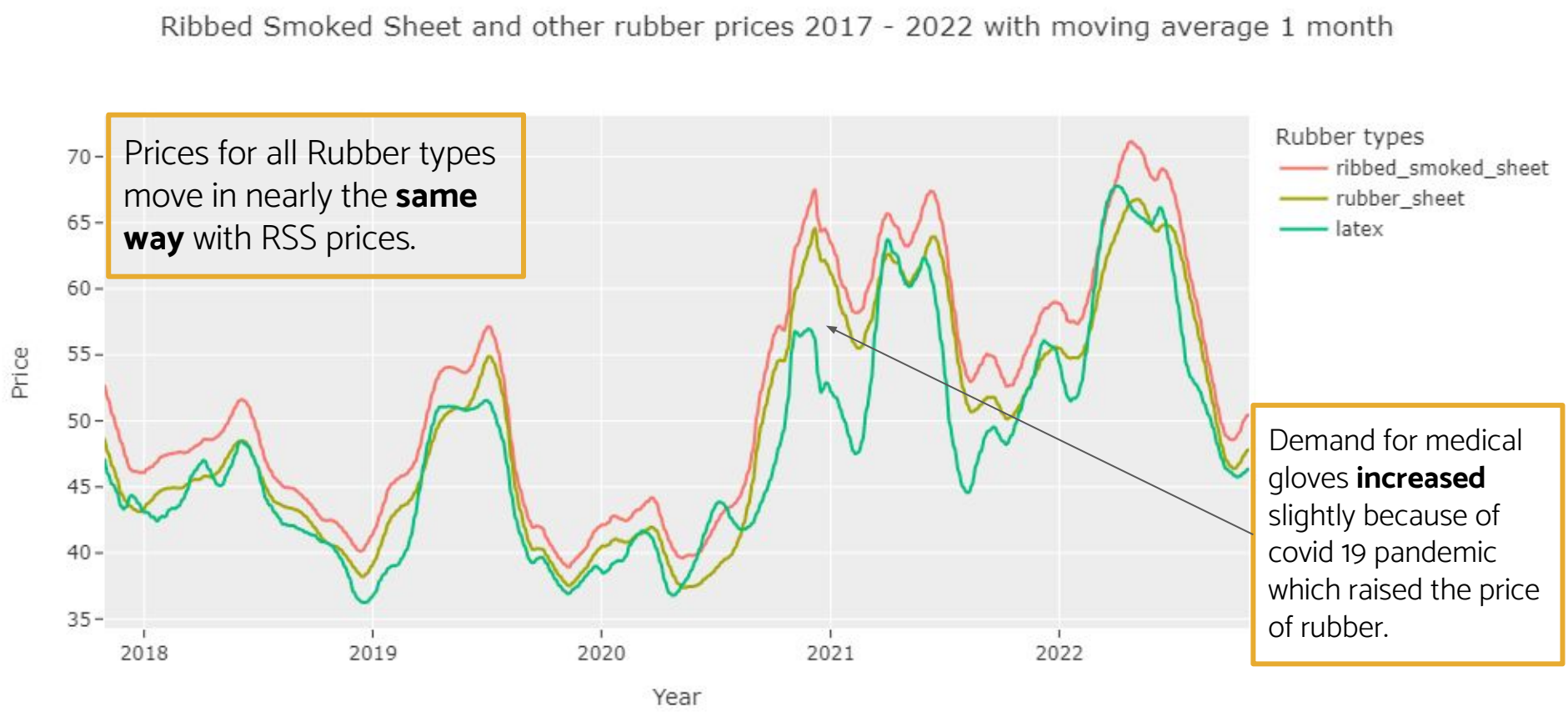
Fill missing values with last observed value because it is a holiday.



Main Data in year 2017 and 2018 half year of Latex are **missing**. Fill missing values from my second data which have data from year 2010 to 2022 and all of data nearly matches.

Part 2:
Exploratory Data Analysis

RSS and Other Rubber Type Prices

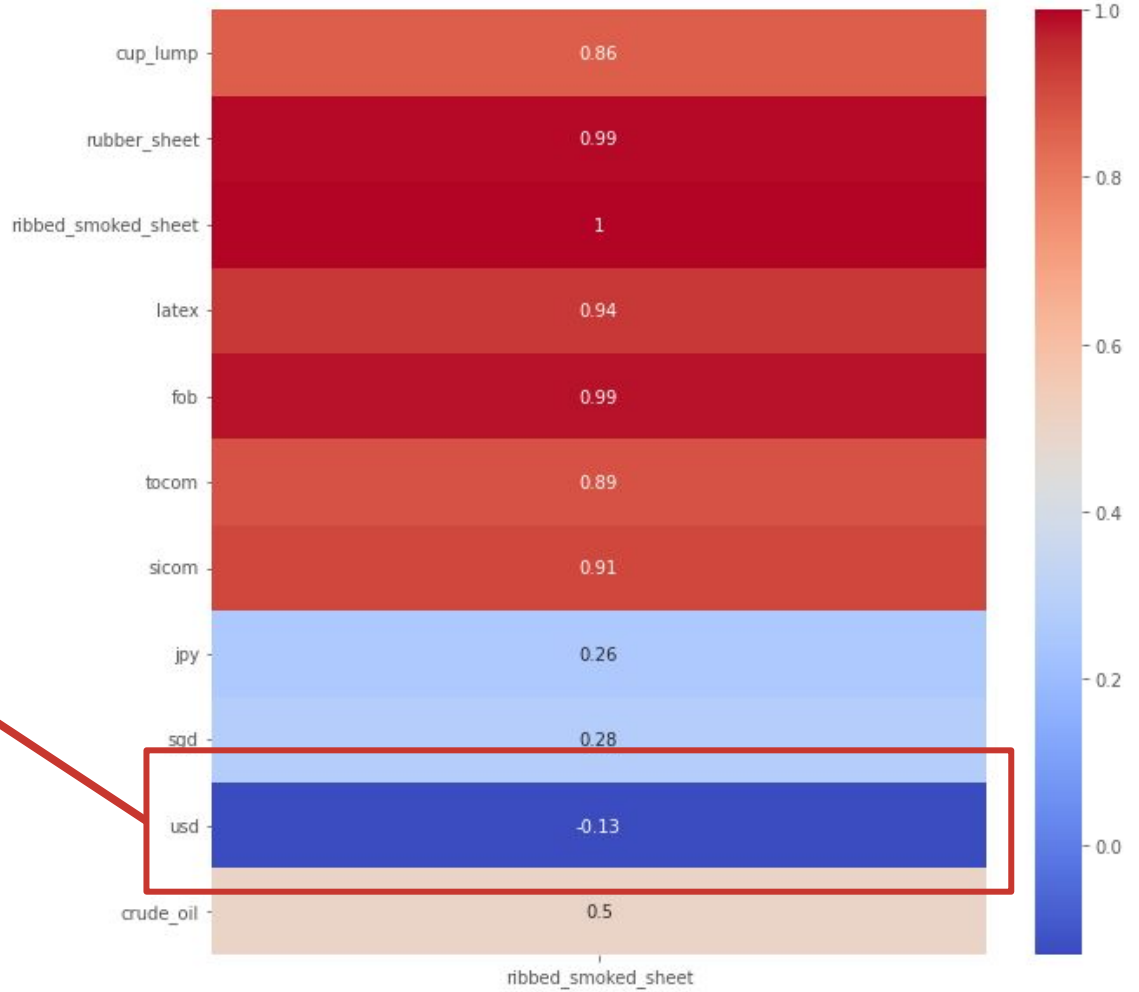


RSS and Futures Prices

Feature Selection

My criteria is features have correlation more than or equal to **moderate** correlation.

Delete usd variable because correlation does not meet my threshold of moderate correlation.



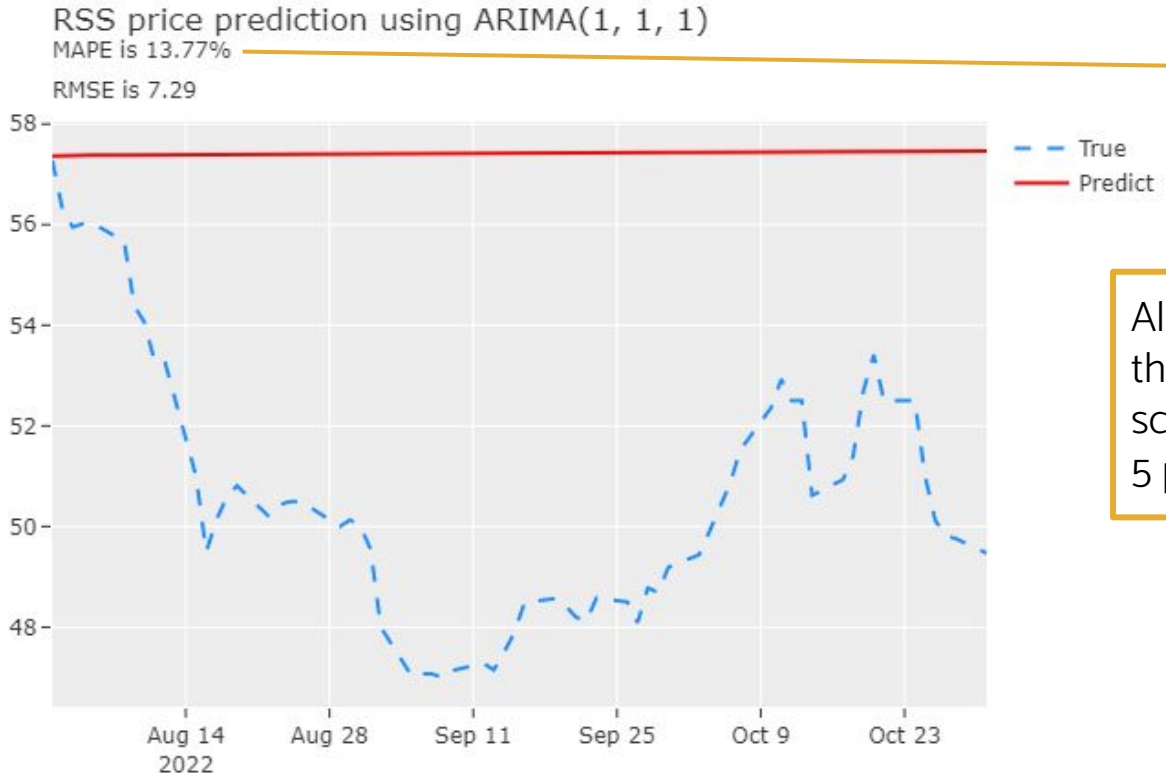
Part 3:
Modeling & Evaluation

Goal for Developing Model

"Build model to **forecast Ribbed smoked sheet** prices 1 month out to estimate trend of RSS prices for Rubber cooperative. Model performance will evaluated by MAPE. Model will be considered for production if MAPE score is **better** than baseline by 5% or more."

It is a convention that we use MAPE for measure of prediction accuracy of a forecasting method

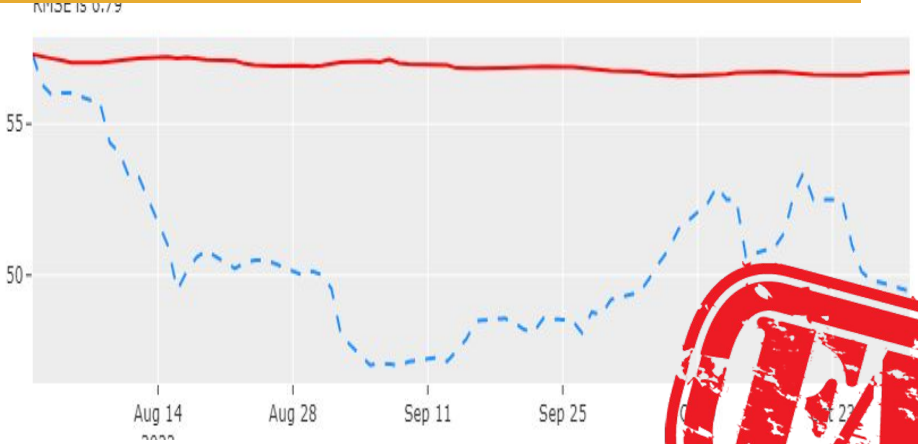
Baseline Model



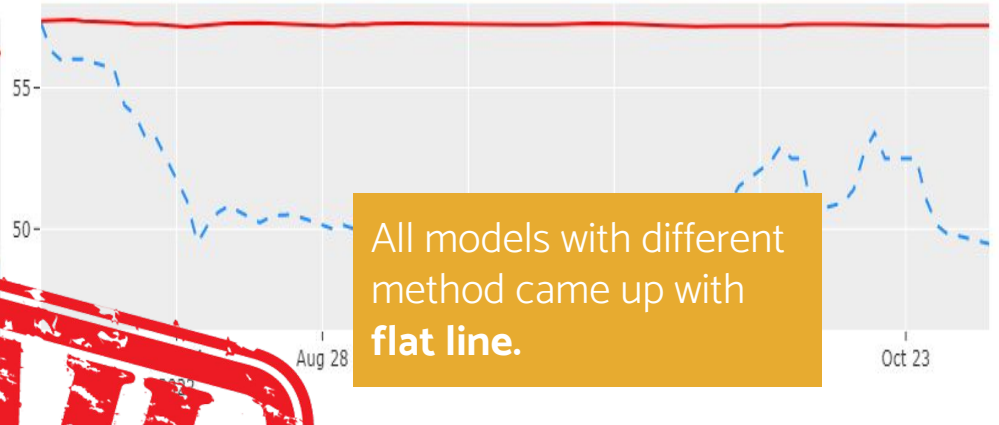
MAPE is 13.77%

All models will be **compared** with the score above. If a model's **MAPE** score is lower than baseline score by 5 points or more, it will **pass**.

Models with features



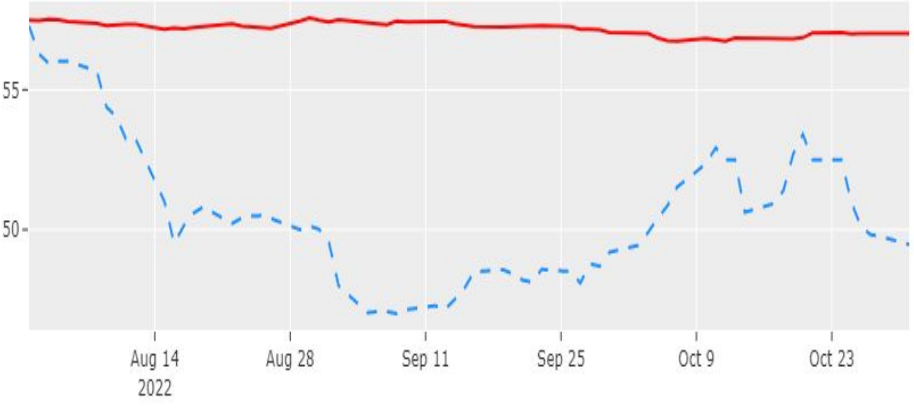
RSS price prediction with futures price using AutoARIMA
MAPE is 13.43%
RMSE is 7.12



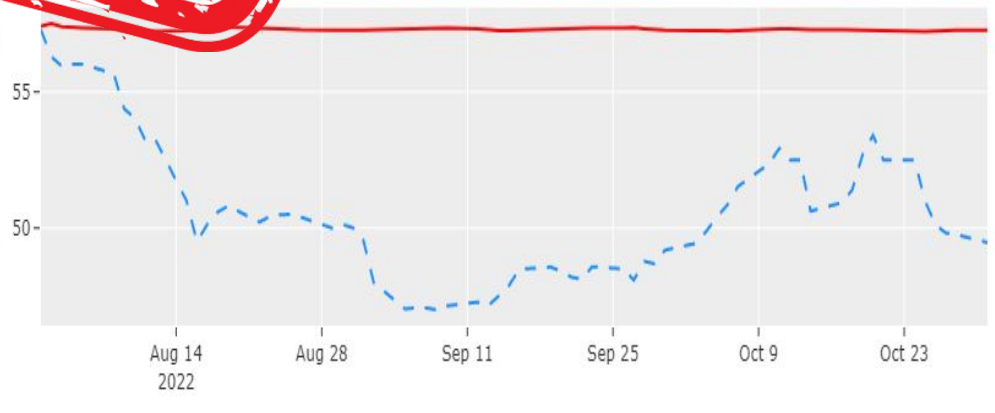
All models with different method came up with **flat line.**

FAIL

RSS price prediction with currencies using AutoARIMA
MAPE is 13.32%
RMSE is 7.09

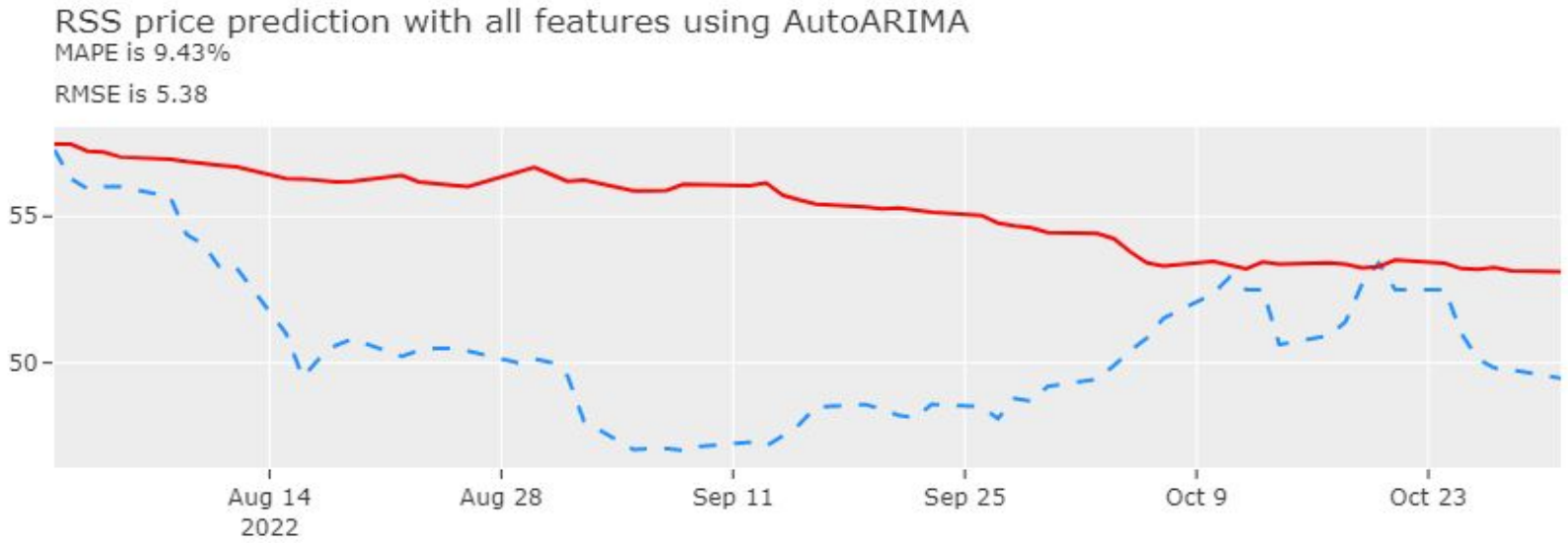


Crude oil using AutoARIMA

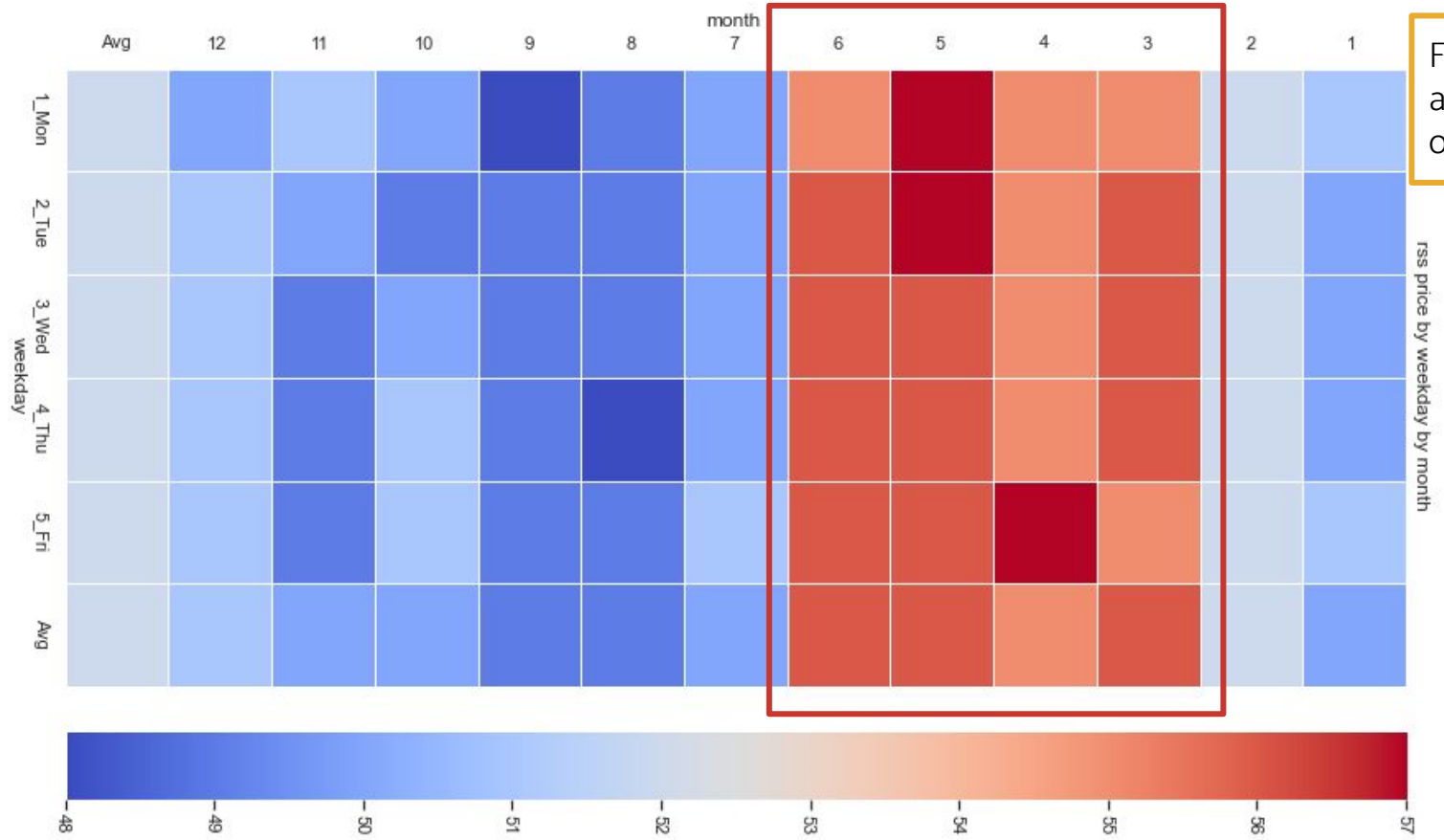


Model with all features

This model is fed **all features**, and predicted values have **downward trend** that fails to capture the pattern of the data.



Feature Engineer - Date



Found months 3 - 6 have average price **more** than other month.

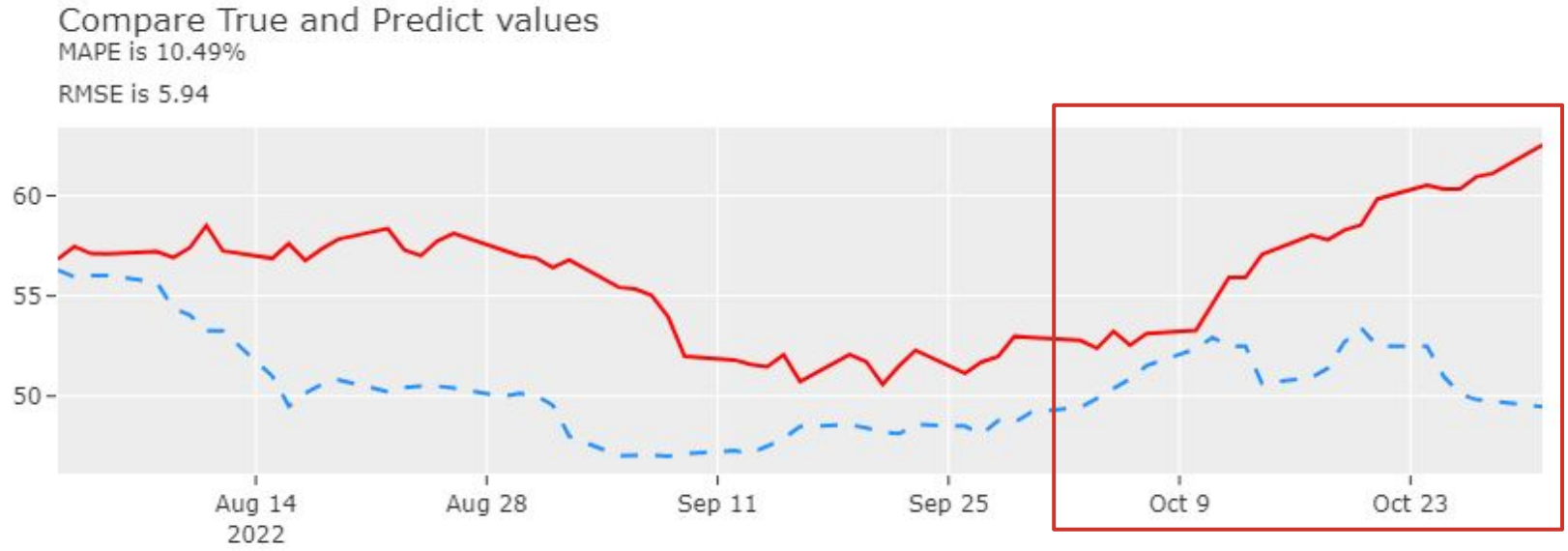
Features Engineer

Quarter

Month

Modeling

The model perform better in capture trends but still make a lot of mistake and if looking carefully in **October** that model went too much **wrong**.

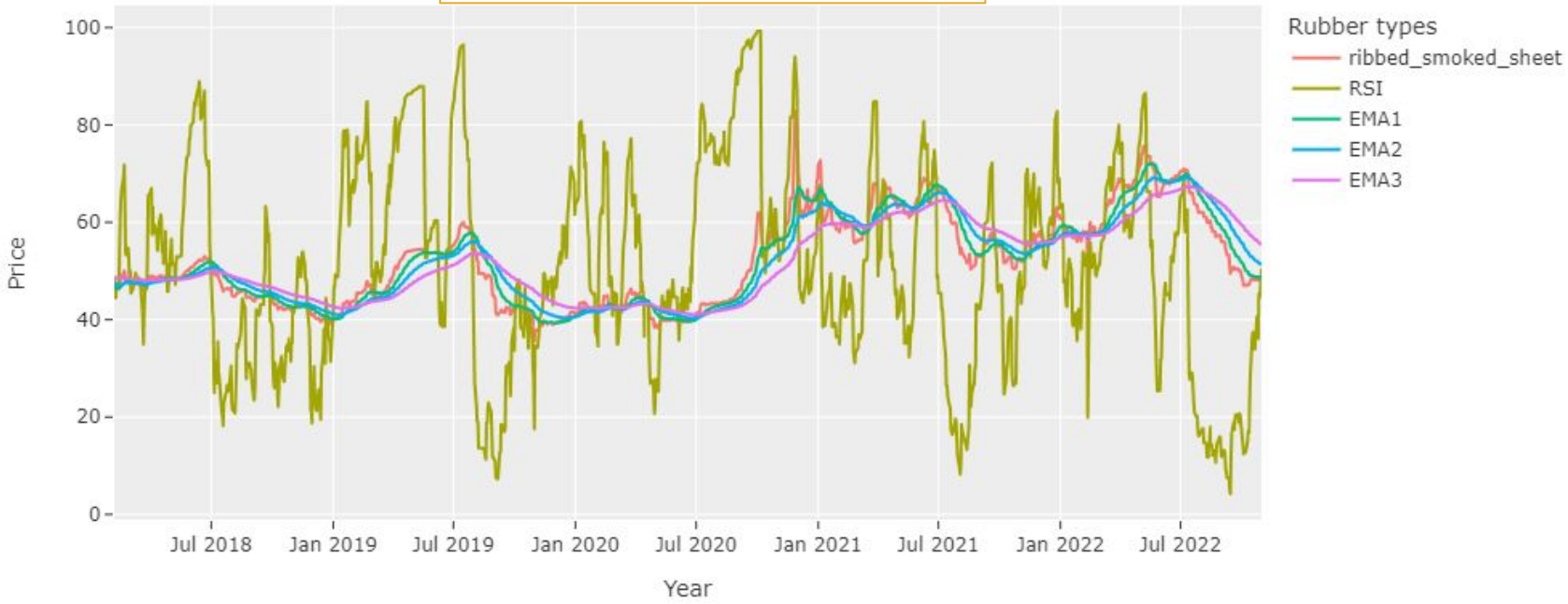


**Move up to more complexity model

Features engineer - Adding Indicators

EMA
tells the model whether there is an **up or down trend** in the data..

RSI
tells the price **momentum** to the model.

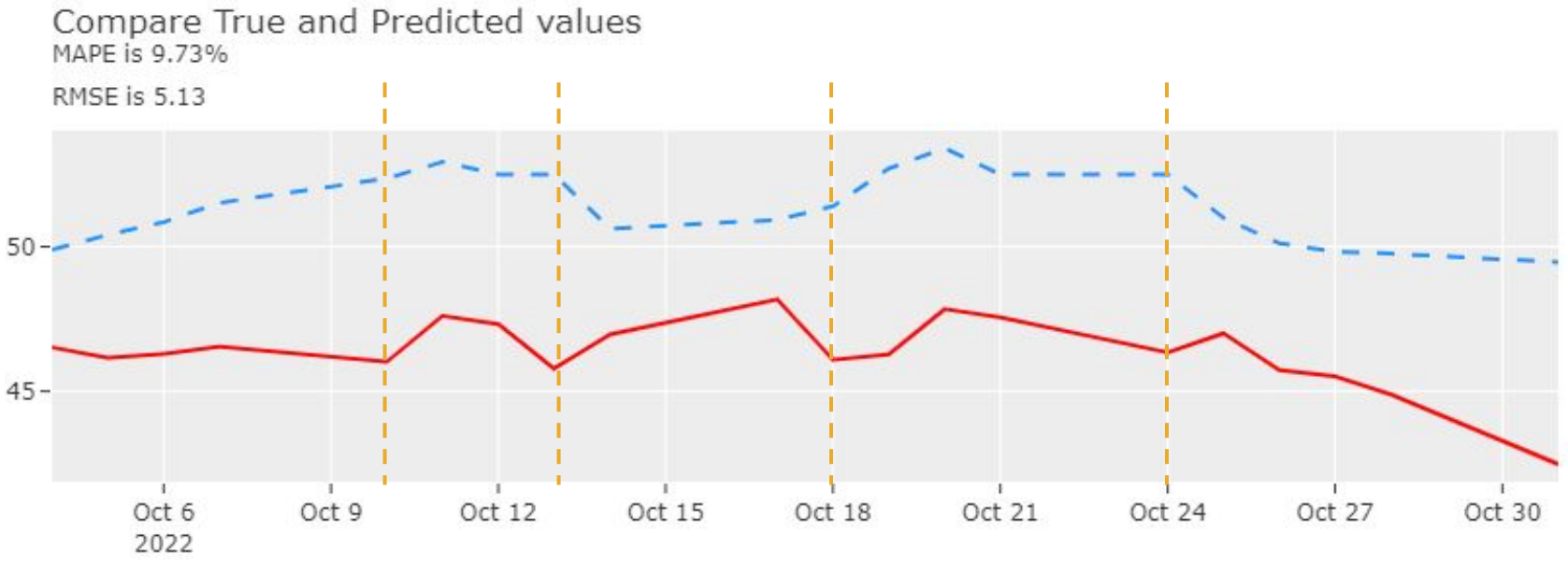


**Gave all data to predict october

Final model

Criteria for pass | **13.78% - model score <= 5%** |

Model perform a lot better in catching trend, but still have too many **errors** and the **score doesn't meet** criteria.



Part 4:
Conclusion & Recommendation

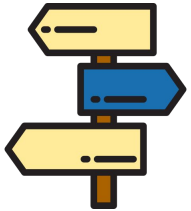
Conclusion

"Build model to forecast Ribbed smoked sheet prices next 1 month to estimate trend of RSS prices for Rubber cooperative. Model performance will evaluate by MAPE, Will working in production part if **score better than baseline 5%**"

To conclude, this project **fails** because the data has complex trend and confounding variables. Final model captures wrong trend though it improved with more feature engineering. This project can be improved further if I can find **additional features and add more Indicators**.

Recommendation

1. RAOT has high quantities of rubber stock in warehouses which if brought to market could decrease RSS price in Thailand and globally. Thus, what needs to be included is **RAOT rubber stock quantity** as it might help model perform more accurately. (source)



2. Find the **appropriate Indicators** to help forecast rubber prices, as in the final model, the indicators helped to model data trends.

THANK YOU

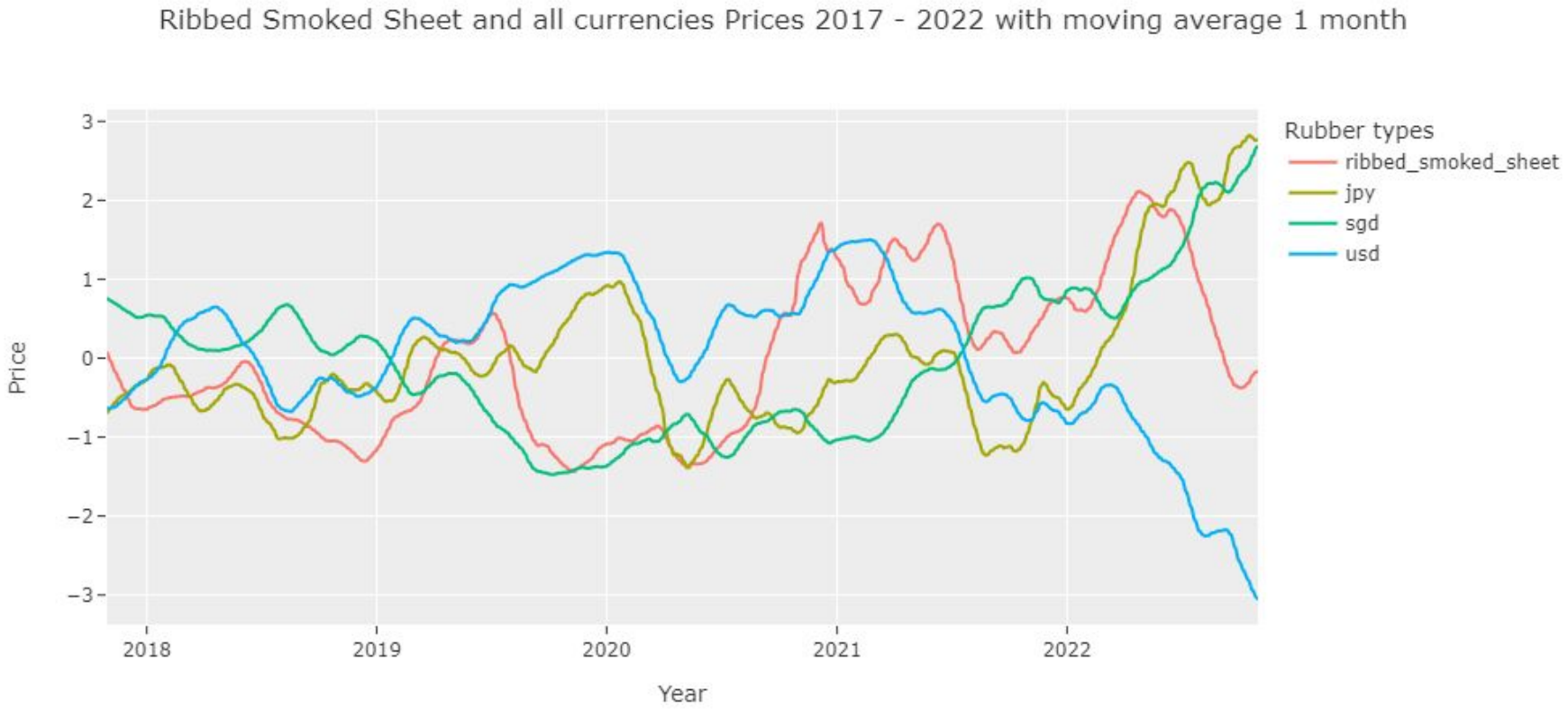
- My **Father** and **Mother** who always give me the important information.
- **James**, **Nozomi,** and **Mood**, my instructional team. They gave me guidance in this capstone project.
- My **classmates** for giving great ideas.
- **Everyone** in this room for paying attention to my presentation.

Presentation template by Slidesgo Icons by Flaticon Infographics by Freepik Images created by Freepik

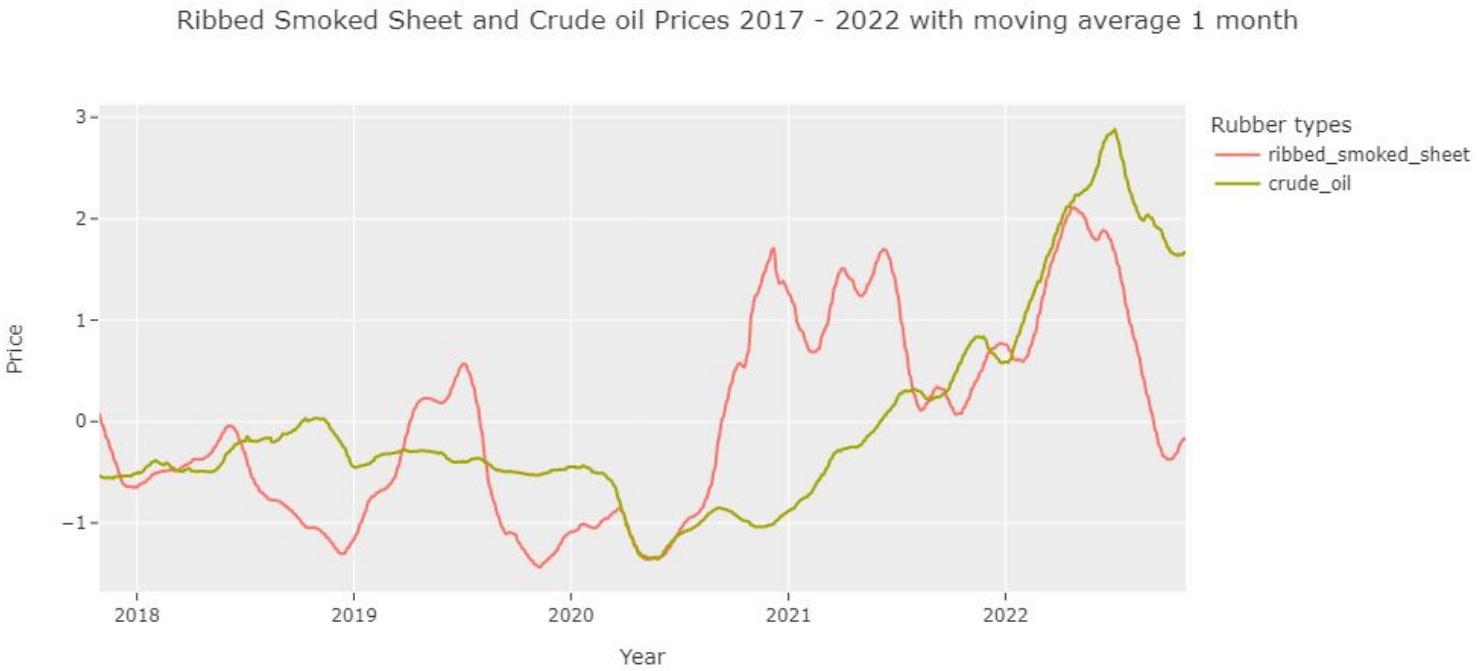
Does anyone have any questions?

Additional

RSS and Currencies Prices

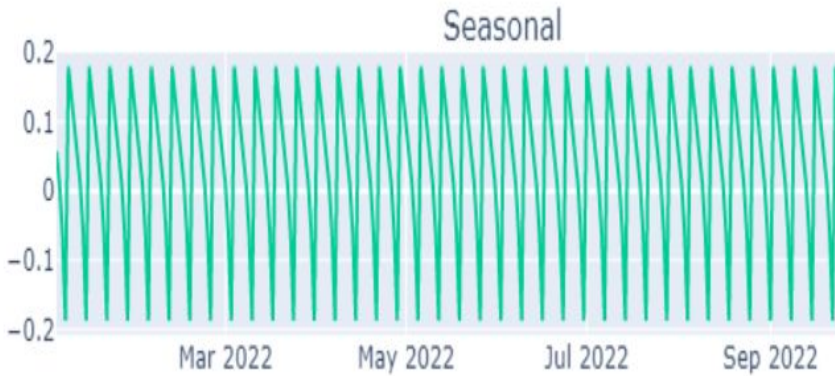


RSS and Crude oil Prices

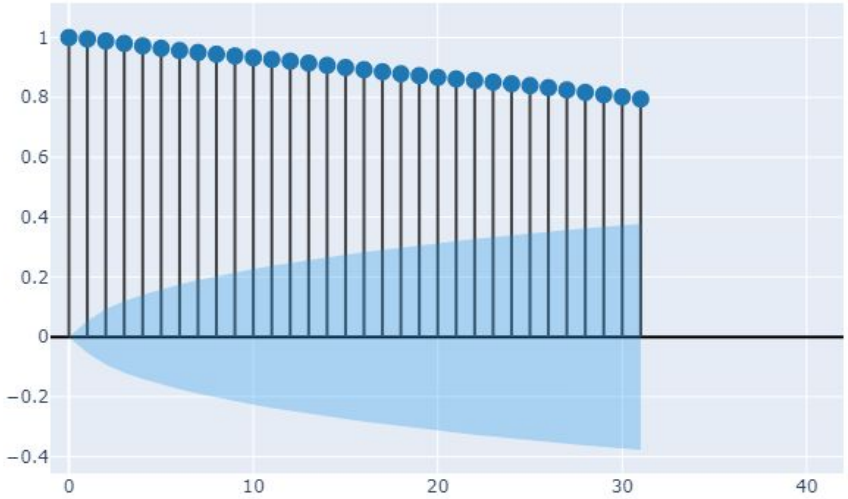


Seasonal Decompose

There is no **sign** of seasonality of RSS price.



Ribbed Smoked Sheet Autocorrelation (ACF)



Stationary Test

Used **Augmented Dickey-Fuller** test for check **stationary** if P-value < 0.05 that dataset are stationary.

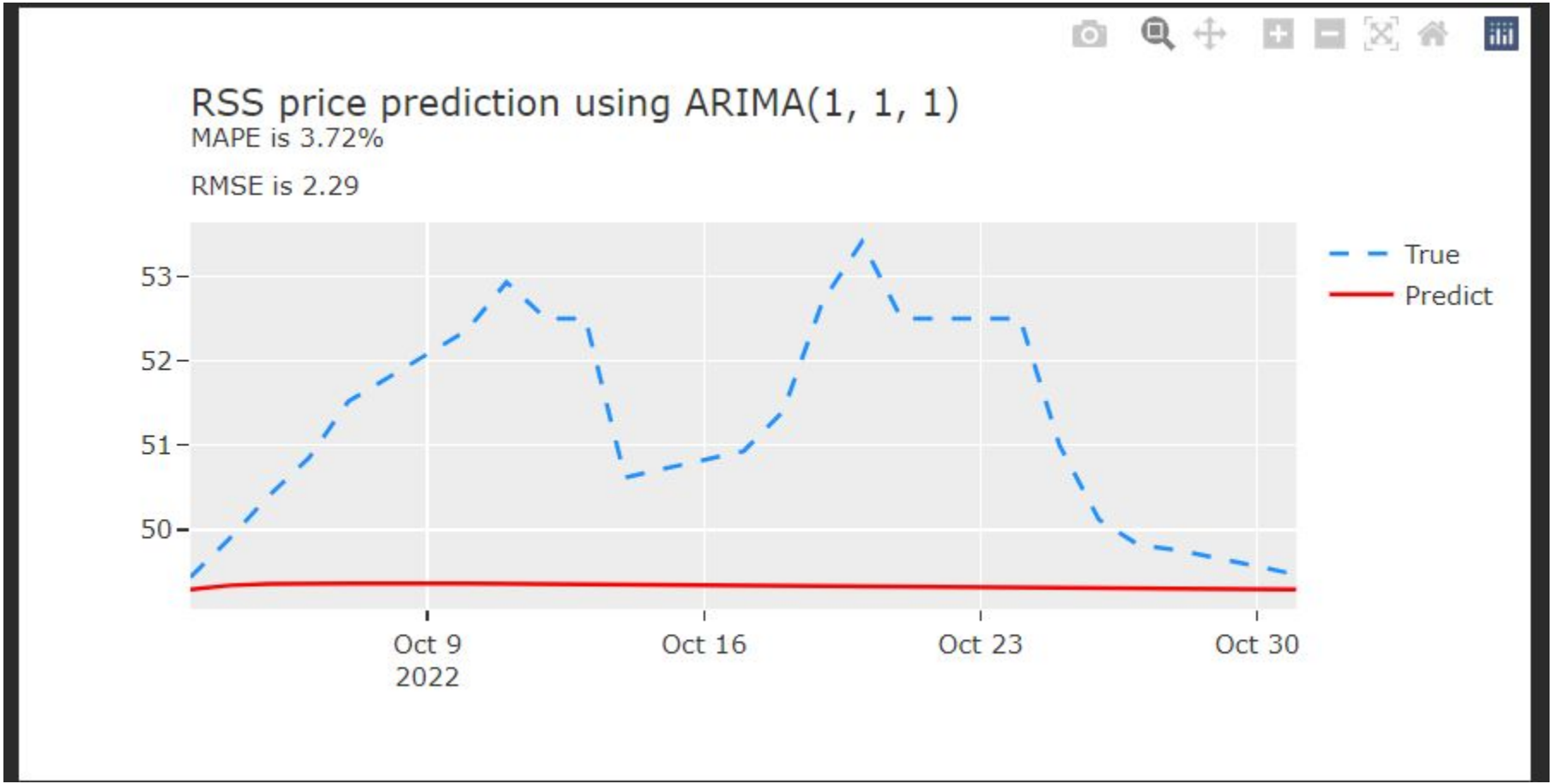
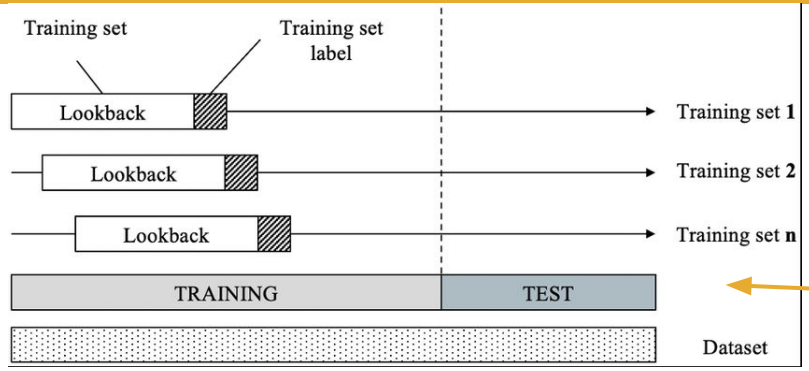
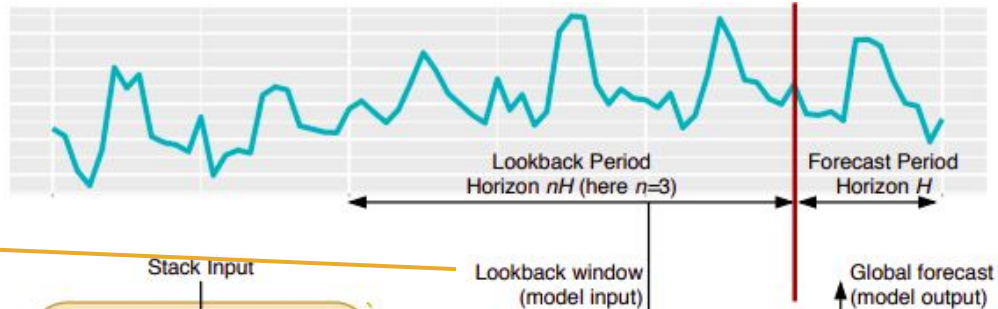
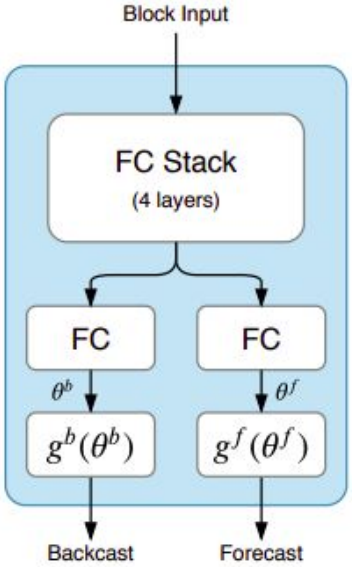
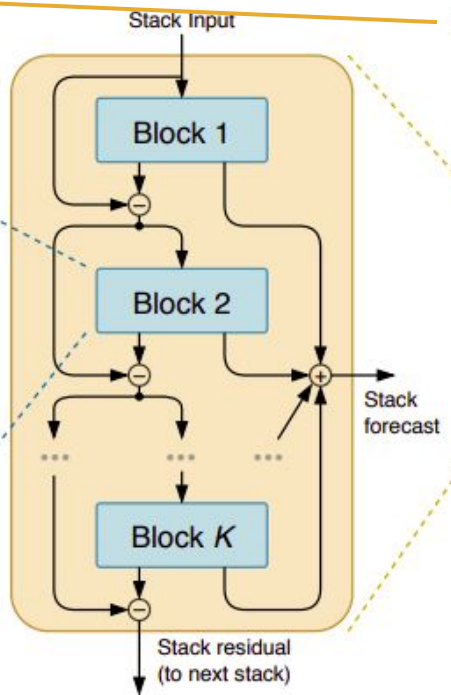
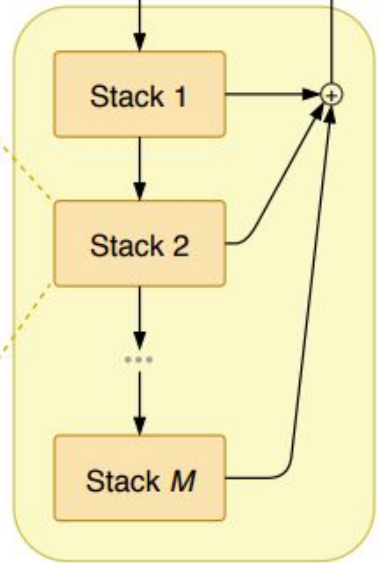
RSS prices after differencing one time

P-value

0.00...

ARIMA

Baseline model only october



N-BEATS - Architecture