Market Models: Examining Market Dynamics in Agriculture

PURDUE UNIVERSITY | The Data Mine

BASF We create chemistry

Nayeli Gurrola, Jalen Hardiman, Ehi Oboh, Margaret Harris, and Dawson Roach

Introduction

During the time with BASF, we evaluated different market segments within different geographical regions to determine how variables such as weather impact the transactional data. Our approach to the project was to find trends and anomalies and put them into the real-world context so that the transactional data can be predicted in a more optimal way.

Central Questions

- What are the key features that drive an agricultural market segment?
- Which features have the largest impact on the market segment?
- Can we improve existing models with the incorporation of new features?

Future Goals

- For future goals we would like to research and explore additional new features (other than weather) that have an impact on the transactional data.
- By adding features, we will be able to get a better accuracy when predicting/forecasting the transactional data.
- Additionally, we want to investigate whether there is a correlation or causation between the two variables and whether both the markets have an impact on each other.

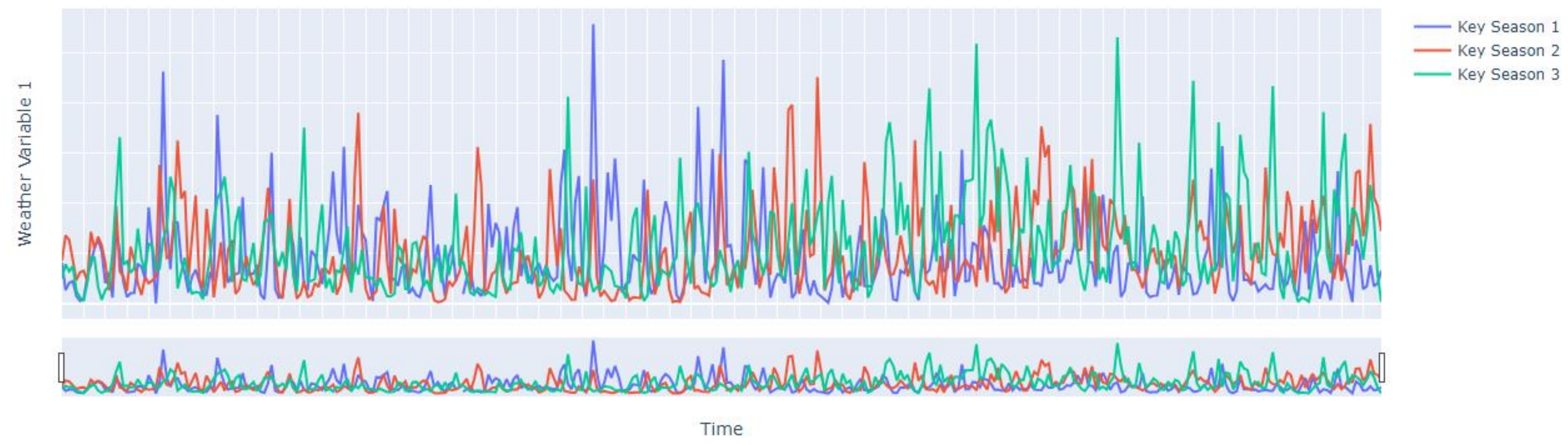
Exploratory Data Analysis – Market Segments 1 and 2

Data Collection : We collected weather and transactional data for specific regions to determine whether certain weather conditions have an impact on market trends.

Data Preprocessing : Data was aggregated for proper time series analyses. Subsequently, there was a visual inspection of the data and statistical tests to support the identification and confirmation of any anomalies or trends.

Model Selection and Evaluation : After selecting an appropriate time series model based on the data characteristics, we proceed to evaluate the model's performance on the testing set using PyCaret.

Market Segment 1



The trends for Weather Variable 1 (Fig. 1) and Weather Variable 2 (Fig. 2) in Geographic Regions 1, 2, and 3 during seasons..

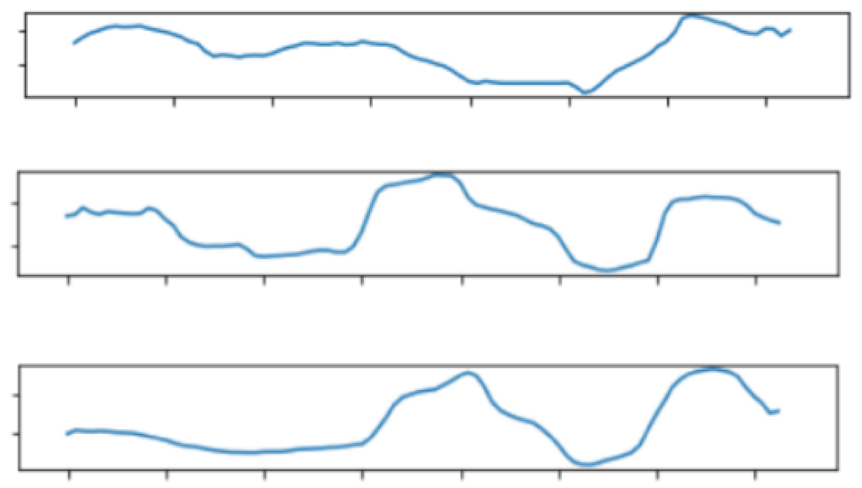


Fig. 1

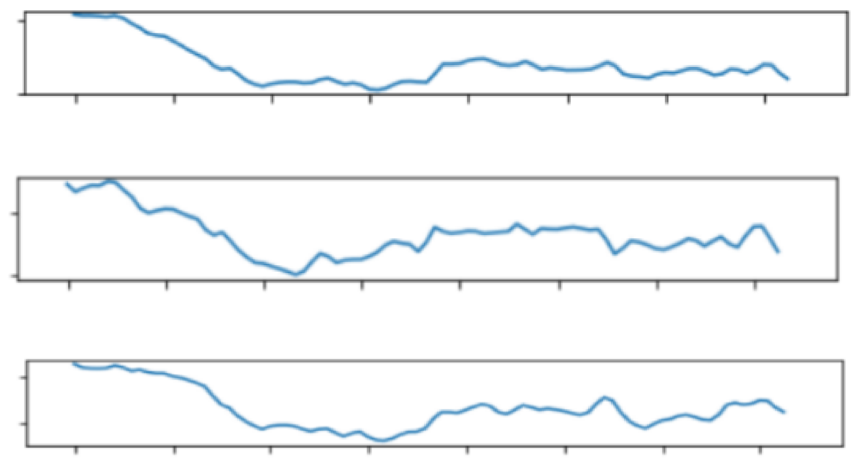


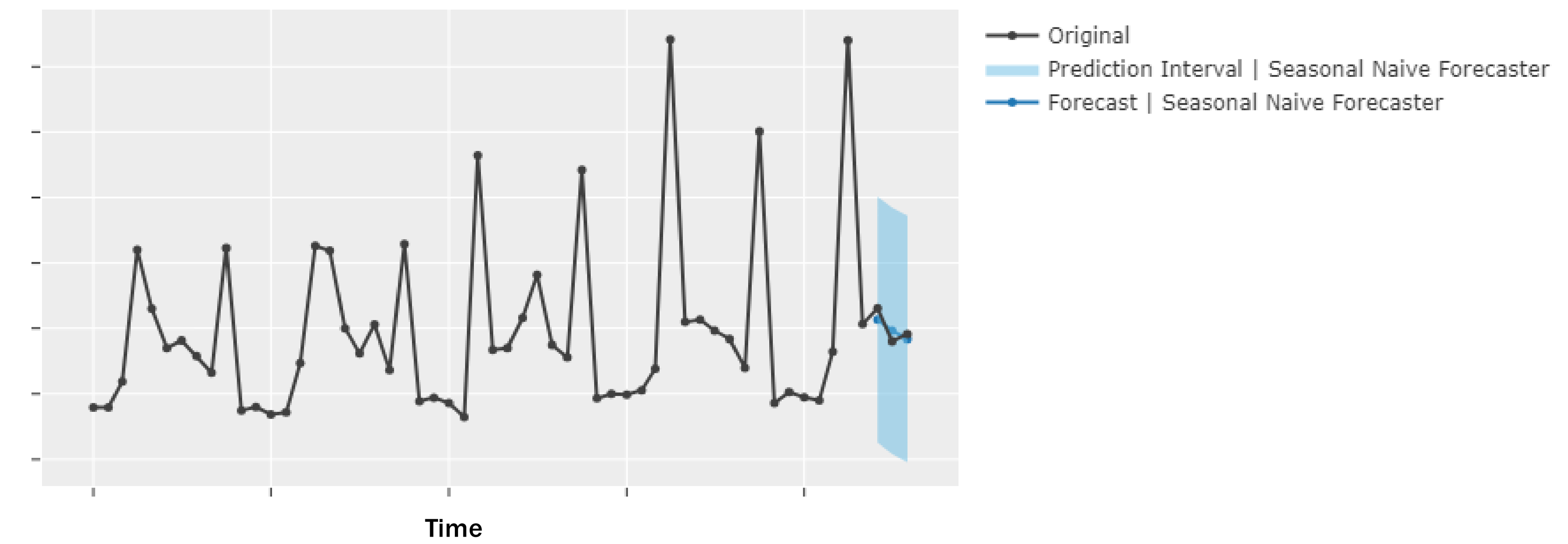
Fig. 2

Results and Conclusion

- Weather Variable 1 and Weather Variable 2 stood out for their impact on agriculture. We noted variations in Weather Variable 2 trends from January to March and their correlation with other variables.
- Focusing on Geographic Regions 1, 2, and 3, we found inflection points in Weather Variable 1 for Geographic Region 2 and Geographic Region 3 in Key Years 2 and 3, and for Geographic Region 1 in a Non-Key Year. Weather Variable 2 trends were similar across all three regions.
- **Discovered that El Niño and La Niña weather patterns significantly affect Market Segment 1's climate, and El Niño leads to warmer temperatures and less rainfall, while La Niña causes colder winters and increased precipitation.**
- **After discovering the effects of El Niño and La Niña, we analyzed the period of high temperatures in 2016. Upon further research, we discovered this period of high temperatures was due to an El Niño. This led to temperatures across Market Segment 1 being much higher than normal for the year.**
- **By incorporating our findings, there was a 16% decrease in Root Mean Squared Error (RMSE) and a 32% increase on R-squared when applied to the internal models of our studied market segments.**

Market Segment 2

This is a univariate time series model for Market Segment 2. The data was aggregated to the appropriate time series and split into geographical regions. We created a univariate time series to understand the transactional data and then added weather layers to understand the impact of weather.



Results and Conclusion

According to the forecast error, this combination of variables yielded the lowest error rate, showing its effectiveness in predicting transactional data for Market Segment 2. As we add weather layers to the existing transactional data, these results inform our approach.

Region	Geographical 1	Best Model	MASE	RMSSE	MAE	RMSE	MAPE	SAMPLE	R2
Approach	external features	Seasonal Naïve							
Univariate - Sales	N/A								
Univariate	Weather Variable 4	Seasonal Arima	0.359	0.4777	207352.2	384110.5	0.4541	0.4971	0.7424
Region	**Geographical 2**								
Multivariate	Weather Variable - 1, 2, 3	Seasonal Arima	0.831	0.5824	360652	406656.3	0.8502	0.8029	0.6824
Multivariate	Weather Variable - 2, 3, 4	Seasonal Arima	0.83	0.5822	360247.8	406455.9	0.8484	0.8018	0.6827
Multivariate	Weather Variable - 1, 2, 3, 4	Seasonal Arima	0.809	0.5805	351354.6	405285.3	0.8003	0.7599	0.6845
Region	**Geographical 3**								
Multivariate	Weather Variable - 1, 3	CatBoostRegressor	0.665	0.5915	242807.2	304551.4	0.6002	0.4466	0.7661
Multivariate	Weather Variable - 1, 2, 3, 4	Seasonal Arima	0.534	0.4225	195055.6	217505.5	0.6659	0.6866	0.8807
Multivariate	Weather Variable - 2, 3, 4	Seasonal Arima	0.527	0.4181	192523.9	215263	0.6577	0.6752	0.8831

Acknowledgements

Katelyn Brandt – BASF Mentor, Melissa Barona – BASF Mentor, Emily Hoeing - [Data Mine Liaison], Cai Chen - [Data Mine Liaison], Maha Bayana – Purdue TA

The Data Mine Corporate Partners Symposium 2024