

Market Indicators for Trend Identification

🗆 • BASF We create chemistry

Ishana Didwania, Aryan Bakshi, Augusto Butkewitsch, Maxsen Anderson, Rithvik Krishnan, Alice Tang, TA: Suhani Dadhich

1. Introduction

BASF creates chemistry for a sustainable future. We want to help.

As statisticians in Purdue's Data Mine, our goal is to aid BASF in this mission by accurately predicting corn and sovbean prices for the next three years.

2. Objective

Use statistics and machine learning to identify trends and market indicators in agricultural data to predict future prices for corn and soybean crops.

3. Methodology

Our quantitative research methodology begins with data scraping from USDA's WASDE (World Agricultural Supply and Demand Estimates) report, utilizing yearly estimates since 1995.

This data was cleaned by removing a few missing values, analyzed with seaborn and matplotlib graphs of various variables over time, and parsed into several predictive models, namely ARIMA, LSTMs, and Gradient Boosting models, to identify patterns and draw meaningful, data-driven conclusions that answer our overarching objective.



5. Modeling & Data Visualization

Visualization

vg.

Below are three models with data spanning 1993-2022, along with predictions for the next few years. ARIMA Model





Gradient Boosting Model



- The ARIMA models predict both soybean and corn prices will fall to a steady state in the coming years.
- Predicted soybean prices will be between \$13-\$14 through 2023-2025
- Predicted Corn prices will fall to around \$5.25 through 2023-2025
- The LSTM Models predict both corn and sovbean prices will remain level with slight upward and downward within a dollar. Predicted sovbean prices will be between \$13-\$14 through 2023-2025 Predicted corn prices will
- be between \$4-\$5 through 2023-2025

The Gradient Boosting models predict both soybean and corn prices will fall within the next three years.

- Predicted soybean prices will be between \$13-\$14 through 2023-
- 2025 Predicted Corn prices will be between \$4,75-\$5,75 through 2023-2025

6. Model Evaluation

- The ARIMA model produced a root mean squared error of 1.22 for corn, and 2.45 for soybean. This is the average deviation of the forecasted prices from the actual prices in the test data set.
- The LSTM model produced an RMSE of 0.953 on the corn testing model, and a 1.58 for the soybean testing model.
- The Gradient Boosting model had an RMSE of 0.68 and 0.5 for corn and soybean predictions, respectively, over the test interval from 2017-2022.

7. Future Goals

- In the future, we will look further into machine learning models like exponential smoothing or random forest regressors.
- · We will create a more interactive dashboard interface for visualizing these data and results.

8. Conclusion

- Linear regression is not suitable for our data.
- Our data is suitable for the ARIMA, LSTMs and Gradient Boosting models, which were used to predict prices for the next 2-3 years.
- Our models indicate that the prices of soybean are predicted to fall to between \$13 to \$14, and that corn will fall to between \$4.50 to \$6.50.

9. Acknowledgements

Our project would not have been completed without:

- The Data Mine Data Science Team
- Cai Chen, Corporate Partners advisor
- Our BASF team mentors, Katelyn Brandt & Audrey Gould
- Ashley Arroyo for technical & statistical help