**Data Analytics Capstone Topic Approval Form**

**Student Name:** Matt Boruff

**Student ID:** 000406586

**Capstone Project Name:** Corn Futures for a farmers needs

**Project Topic**: Time Series on corn prices to project future planting needs

**This project does not involve human subjects research and is exempt from WGU IRB review.**

**Research Question:** Can we use a weekly corn price index from past years to predict future corn prices?

**Hypothesis**: Weekly prices will need to be predicted to at least 90% accuracy through using ARIMA and MSE checking in order for the hypothesis to be true.

**Context:** Farmers would benefit in several areas, such as improved planning for their field sizes. With reliable future prices farmers can make a more informed decision on what to plant and harvest and best times to enter the market with that crop for pricing. Risk Management is another area that they can benefit, maximizing their yields, profits and minimize possible losses to price drops. Accurate price predictions enable farmer to allocate their resources more effectively. Adjusting their purchase of seed, fuel, fertilizer and equipment to fit the needs of the coming harvest. As well they can develop a marketing plan to maximize their profit by capitalizing on the anticipated prices. Beyond the immediate planning, access to a reliable forecast allows farmers to plan their field rotations, fertilization and infrastructure upgrades(drainage) to land knowing what funds should be available each season. In summary this study could provide farmers with valuable information and tools to navigate the fluctuating agriculture market more effectively.

**Data:**Using a data set that has weekly prices at market closings allowing us to build a time series visualization and help predict (if possible) future prices.

*If an existing data set will be used, describe the data set.* [*https://www.kaggle.com/datasets/nickwong64/corn2015-2017*](https://www.kaggle.com/datasets/nickwong64/corn2015-2017)This is the data set I am looking to start with and use initially. If further data is needed I will update as I go.

*Explain who owns the data and why you are allowed to use this data for your capstone project.* <https://www.kaggle.com/datasets/nickwong64/corn2015-2017>

**Data Gathering:** I will be gathering a data set from Kaggle on the corn price index from 2015 to 2017

**Data Analytics Tools and Techniques**: I will be using Python to perform an Arima analysis with a time series visualization to help predict corn prices.

**Justification of Tools/Techniques:** ARIMA (AutoRegressive Integrated Moving Average) models are well-established and widely used for time series forecasting. They are particularly suitable for capturing temporal dependencies and patterns present in sequential data, making them a natural choice for predicting corn prices, which exhibit time-dependent behaviors influenced by various factors such as seasonality, trends, and cyclical patterns. Justifying the use of ARIMA models involves highlighting their proven effectiveness in capturing these complex dynamics and providing accurate forecasts.

**Project Outcomes**: The project on time series prediction of corn prices using Python, ARIMA modeling, and time series visualizations promises several valuable outcomes and deliverables, each contributing to a comprehensive understanding of corn price dynamics and facilitating informed decision-making for stakeholders in the agricultural industry.

1. **Predictive Models**: One of the primary outcomes of the project will be the development of predictive models capable of forecasting future corn prices. Utilizing historical corn price data, ARIMA modeling techniques implemented in Python will be employed to build robust forecasting models. These models will capture the underlying patterns, trends, and seasonality present in the data, enabling accurate predictions of corn prices over various time horizons.
2. **Forecasting Accuracy Evaluation**: A critical deliverable of the project will involve evaluating the forecasting accuracy of the developed models. This evaluation will entail comparing the predicted corn prices against actual observed prices over a specific validation period. Various metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) will be calculated to assess the performance of the models. This evaluation will provide insights into the reliability and effectiveness of the forecasting models in real-world scenarios.
3. **Visualization Tools**: The project will also yield a set of time series visualizations aimed at enhancing understanding and interpretation of corn price dynamics. Using Python libraries such as Matplotlib and Seaborn, these visualizations will illustrate important features of the data, including trends, seasonality, and irregular patterns. Line plots, scatter plots, and decomposition plots will be employed to visualize historical price data, model predictions, and forecast intervals, aiding in the identification of relevant patterns and informing decision-making processes.
4. **Documentation and Reports**: A comprehensive documentation package will be prepared as a key deliverable of the project. This documentation will include detailed descriptions of data preprocessing steps, model development methodologies, parameter tuning processes, and evaluation results. Additionally, reports summarizing the findings, insights, and implications of the study will be generated. These reports will provide stakeholders with actionable information and recommendations based on the analysis of corn price trends and forecasts.
5. **User-Friendly Interface**: To facilitate accessibility and usability, a user-friendly interface may be developed as an optional deliverable. This interface could allow stakeholders, such as farmers, traders, and policymakers, to interact with the predictive models and visualizations in a intuitive manner. Through the interface, users can input parameters, generate customized forecasts, and explore different scenarios, empowering them to make informed decisions based on the latest corn price predictions.

In conclusion, the project on time series prediction of corn prices offers a range of outcomes and deliverables aimed at providing valuable insights, tools, and resources to stakeholders in the agricultural industry. From accurate forecasting models to informative visualizations and user-friendly interfaces, these deliverables have the potential to enhance decision-making, improve risk management, and optimize resource allocation in corn production and trading activities.

**Projected Project End Date**: 3/15/2024

**Sources**: https://www.kaggle.com/datasets/nickwong64/corn2015-2017,

DataCamp,

python.org,

<https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>

**Course Instructor Signature/Date:**

The research is exempt from an IRB Review.

An IRB approval is in place (provide proof in appendix B).

Course Instructor’s Approval Status: Approved

Date: Click here to enter a date.

Reviewed by:

Comments: Click here to enter text.