**AGR 333 Data Science for Agriculture**

**Agricultural Economics Assignment**

For this assignment, you will draw on the exercises you performed in Labs 14 and 15. While most of the information you need can be found in the labs, some steps may require additional research. You will be asked to provide visuals (graphs and tables) and write a brief response about the underlying economics. **Upload your Lab 14 and Lab 15 code to GitHub and include the repository link in your submission.**

**Lab 14 Questions** (WASDE.csv)

**GitHub repository link to the code for Lab 14 (**2.50 points**):**

**0.50 points each question**

1. Make a scatter plot of corn prices on the y-axis and the stock-to-use ratio (SUR) on the x-axis with a linear regression line. What is the general relationship between corn prices and the SUR? Is a linear regression line a good representation of this relationship?
2. Run the linear regression: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. How do you interpret the coefficient on SUR? Is this consistent with the theory of supply and demand? How do you interpret the R-squared value? Describe the overall fit of this model.
3. Run the non-linear regression: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. From this model, what is the marginal effect of SUR on the price of corn? Hint: if then .
4. Create a scatter plot of corn prices on SUR with regression lines for two distinct periods: 1973-2005 vs. 2006-2019. How did the relationship change between the two time periods? What could explain this?
5. Run the linear regression with a dummy variable for the post 2006 period and an interaction between the post-2006 dummy and SUR: and show the regression output in a table. Include the regression coefficients, their standard errors, and the R-squared value. From this model, what is the marginal effect of SUR on the price of corn in the 1973-2005 period? What is the effect of SUR in the 2006-2019 period? How do you interpret the coefficient on the P2006 dummy variable?

**Lab 15 Questions** (soybeans-prices-monthly.csv, CPI.csv)

**GitHub repository link to the code for Lab 14 (**2.50 points**).**

**0.42 points each question**

1. Graph the trends in the nominal (not inflation adjusted) and real (inflation adjusted) price of soybeans. These can be two separate graphs or one plot with both series. What explains the difference between the two trends? Are these series stationary? Why or why not?
2. Decompose the nominal soybeans prices time-series and show the seasonal effects in a table by month. Which months have the largest seasonal effects?
3. Graph the simple moving average (SMA) of nominal prices for 3, 12, and 48 months. Describe what happens as you increase the “order” (number of months) included in the SMA.
4. Graph the log differenced time-series of prices (i.e. percent returns). Does this series appear stationary?
5. Run an auto-regressive model on log-differenced prices and report the coefficients associated with each lag. The number of coefficients you report will depend on the number of lags you choose to put in your model. I.e., if you run an AR(5), you will report five coefficients. You do not need to report standard errors or R-squared, just the coefficients.
6. Using your AR model, forecast future percent returns for the six months after the end of the time series (i.e. Aug, 2020-Jan, 2021). Graph your forecast along with the most recent 36 months. How confident are you in your forecast? How close was your forecast to the actual percent returns in soybeans prices for July-Oct of 2020?