FRE 501
Module 1 Assignment
Due November 26, 2021
(but strongly recommend to complete this question and submit it before the midterm exam).

Question 1 (Estimating Seasonality with R)

During the October 14th lecture most of you used Excel (rather than R) to estimate an enhanced dummy variable model with the goal of isolating long-term average quarterly prices, both in normal years and in high stock years. This first assignment question requires you to estimate the same model in R. If you get the same answer as what you obtained with Excel, then you very likely have the correct answer.

Below is a possible work flow along with some coding hints. Begin by opening the file titled "seasonality_exercise_student.R" which is in the Code folder of the October 14th zip file titled "stock-class-exercise.zip". Run the project folder and then run the short bit of code in "seasonality_exercise_student.R". The code will load the necessary packages and read in the two data files: "prices.csv" and "stocks.csv".

Krisha will provide you with separate instructions regarding the format of your submitted answer (e.g., Markdown file). She will also provide the grading rubric. The full assignment for Module 2 consists of four equally weighted questions, where each question is worth 2.5 points toward your final grade.

Be sure to carefully read "dummies_difference.pdf" (posted in Canvas) before starting this assignment.

Here is what you need to do to complete the assignment.

1. Use R to estimate the dummy variable model in first difference format and with interaction terms. Specifically, estimate the following equation:

$$P_{j,t} - P_{j,t-1} = \beta_1(D_{1,t} - D_{1,t-1}) + \beta_2(D_{2,t} - D_{2,t-1}) + \beta_3(D_{3,t} - D_{3,t-1}) + \gamma_0(\Phi_{j,t} - \Phi_{j,t-1}) + \gamma_1(D_{1,t}\Phi_{j,t} - D_{1,t-1}\Phi_{j,t-1}) + \gamma_2(D_{2,t}\Phi_{j,t} - D_{2,t-1}\Phi_{j,t-1}) + \gamma_3(D_{3,t}\Phi_{j,t} - D_{3,t-1}\Phi_{j,t-1}) + e_{j,t} - e_{j,t-1}$$

2. Following Krisha's instructions, show all of your code and the results from running the code. Use the head() function to show the first six rows of the most important data frames (e.g., after merging the annual and quarterly data). Be sure to show the set of estimated coefficients and p values, as reported by R. Your results should match the results in "Seasonality Estimation Solution.xlsx" (posted in Canvas).
3. Following Krisha's instructions, briefly describe your code, moving from importing the data to generating the final regression results.
4. Answer the following questions (please provide specific answers in terms of differences in long term average prices):
    a. What is your estimate of β_3 measuring?
    b. What is your estimate of $\gamma_0 + \beta_2$ measuring? Use "-" instead of "+"
    c. What is your estimate of γ_4 measuring? Subscript "3" instead of "4"

What follows is only a suggestion for how to structure your code. You are welcome to develop your code with a partner but you must submit individual work and indicate who you worked with. Your explanations of the code and the answers to the questions must be completed individually. You must not copy large chunks of other people's code – that is plagiarism. You are welcome to use small chunks of someone else's code (e.g., a suggestion from Piazza). Please do not post your full code on Piazza. Use this forum to post specific questions and receive answers about specific chunks of your code (e.g., problems with merging the annual and quarterly data).

Limit the amount of time you spend on this question (e.g., 4 – 5 hours). If you run into a coding problem don't become stubborn and fight with it. Simply post your specific code chunk on Piazza and ask for help. Each Piazza post should have just one specific question, similar to how things work on sites such as Stack Overflow.

Your code should make extensive use of the *mutate()* function which Krisha has shown you on several occasions.

1. Add an indicator variable to the stocks data frame: 1 if annual stocks > **108** and 0 otherwise.
    a. Try using the *mutate()* and *elseif()* functions together so that you can create the variable and add it to your stocks data frame in one step.
    b. See https://rstudio-pubs-static.s3.amazonaws.com/116317_e6922e81e72e4e3f83995485ce686c14.html#/5

```
> head(stock_data)
  year stocks stckDum
1    1  119.6       1
2    2   96.1       0
3    3  103.5       0
4    4   91.7       0
5    5  111.9       1
6    6  107.5       0
```

2. Merge your annual stock data and quarterly price data using an appropriate method. Consider using the *left_join()* function which Krisha showed you.
    a. Check out the Left Join section of https://hollyemblem.medium.com/joining-data-with-dplyr-in-r-874698eb8898
    b. leftJoinDf <- left_join(tableA,tableB,by="Customer.ID")

3. Use the same combination of *mutate()* and *elseif()* to create a set of quarterly dummies within your merged quarterly price data frame.
4. Use the *mutate()* function to create the interaction variables within the merged quarterly price data frame.
5. Use the *mutate()* and *lag()* function to add the first differences of price, dummies and interaction variables to your merged quarterly price data frame (e.g., d1Diff = d1 - lag(d1)). Be sure to include *slice(-1)* before completing the *mutate()* function – this is used to remove the first row of the data frame.
6. You are now set to use R's regression function to estimate the dummy variable model.