FRE-GY-6971, Homework #2, due 4/11/2023, 9pm

1. Definitions:
   1. Sample1: 09/30/2012 to 09/30/2016
   2. WFLY: 3Y\*w1 – 5Y + 7Y\*w2, weights = (w1, -1, w2)
2. Build a Jupyter Notebook to do the following:
   1. Download a panel of CMT rates into pandas data frame & remove ‘1M column from the dataset
   2. Perform PCA on the dataset using Sample1
   3. Use this PCA model to analyze the CMT curve move on the 2016 Election Day: 11/8/2016 to 11/9/2016
      1. Plot CMT curve move vs the move explained by the first PCA factor, first 2 PCA factors, first 3 PCA factors
      2. Explain your calculations and results
   4. Compute weights of the WFLY to make sure that WFLY does not have PCA1,2 risk exposure in Sample1. Let’s call this combination WFLY1
   5. Choose weights of the WFLY from cointegration analysis (weights correspond to the best cointegrated vector). Let’s call this combination WFLY2
      1. Use Box-Tiao estimation procedure
      2. If cointegration estimation fails for you – use a linear regression of levels instead (regressing 5Y rate on [3Y,7Y] rate)
3. Compute Half-Life & ADF statistic for WFLY2 using Sample1
   1. Note that you are using time series of levels, not daily differences
4. Repeat Step #3 out-of-sample: using 3m, 6m out of sample periods
   1. How do out-of-sample results compare across periods and combinations?

**Independent reading assignment**:

1. What is CCA (Canonical Correlation Analysis)? Find 2 -3 examples independently.