Section 1: Submission (10 points)

Please follow the link provided on Canvas and submit all your files on Gradescope. The interface is self-explanatory. Only one member of your group needs to submit. The other members can be added after submission. See the Gradescope video and documentation linked on Canvas for details.

A submission should contain three parts:

1. A report in PDF.

   The first paragraph must be a brief but specific statement on each member's contribution. A statement like "Rui worked on the project" is not acceptable. A report should have tables and figures and verbal descriptions and comments for each section and subsection of the project. The report could also include notes for Lisheng when running your code if necessary.

2. Code in Python notebook (.ipynb) Please put sufficient notes to indicate which cell does what.

3. Data files. Eight slots are preset. You submit as many as you have. You don't need to upload the data files given to you if you did not modify it for your code.

Section 2: Measuring the tone of FOMC statements (90 points)

You will follow the method in Tadle (2022). The paper is linked on Canvas.

1. **(10 points)** Scrape the text of the FOMC statements from January 2000 to the present. Use https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm for 2019-2024 and https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm for 2000-2018.

   Here are some hints provided by Joanna for scraping if scraping is new to you.

   - Use the **BeautifulSoup** package to extract the html from the website. Load this package by running the code: *from bs4 import BeautifulSoup.*
   - Use the command *soup.findall('a',href=True)* to get a list of all the URL links on the web page.
   - Inspect the URLs to figure out what the format of the links to statement files look like and narrow your list of links to only include these.
   - Iterate through the links to the statements and again use the **BeautifulSoup** package to extract the text.
   - Inspect the text to remove extraneous parts from the start and end.
   - Make sure to remove HTML tags.

- You will need the date of the statement to merge the text to the returns. It can be found in the URL. For example,
  https: //www.federalreserve.gov/newsevents/pressreleases/monetary20220126a.htm
  was released on Jan. 26, 2022, which is at the end of the URL.

Report the following after you extract the text: 1) Number of statements you obtain and the list of the dates of their release. The statements are released shortly after each FOMC meeting. 2) A summary statistic of the number of words in statements including mean, standard deviation, minimum, first quartile, median, third quartile, maximum.

2. **(30 points)** Follow Tadle (2022) section 3.1 (pages 4-8) to measure the tone of each FOMC statement. Here is a brief description of his method and some programming suggestions.

   - Split the statement into sentences using *sent_tokenize* from the **nltk** package in Python.
   - Remove punctuation and capitalization.
   - Remove all sentences that do not contain any keywords defined as those from the hawkish or dovish lists in the Excel file provided (Tadle_2022_keywords).
   - For each remaining sentence, count the number of positive and negative keywords, adjusting for negation terms. For example, "When a positive term is in the proximity (occurred after three words or less) of a negation term, then its effect is counted as negative." See Table 3 in Tadle (2022) for a list of examples.
   - Use equation (1) on page 7 of Tadle (2022) to assign sentiment scores to each sentence. Footnote 13 on page 7 tells you how to handle sentences that contain both hawkish and dovish keywords.
   - Aggregate across sentences to get the document-level score for each FOMC statement by following equation (2) on page 7 of Tadle (2022).

   Plot the computed sentiment of each FOMC statement since January 2000. Please plot the dots only; donot connect dots with lines.

3. **(10 points)** From FRED (Federal Reserve Economic Data, https://fred.stlouisfed.org/) find the following data series: 1) Federal Funds Target Rate (DFEDTAR) and 2) Federal Funds Target Range-Upper (DFEDTARU). You can search the "ticker symbol" (Capital letters in parenthesis) for each series given. FOMC used to announce a targeted value for the Federal Funds Rate (FFR) before Dec. 15th, 2008. Since that date, FOMC has announced a range. We will use the upper limit. Compute the change of the target on each of the FOMC statement release date. Plot these changes together with sentiment score you computed in the previous section. What do you find?

4. **(30 points)** Comment on the Tadle (2022) method. In light of your comments, describe and implement a different way to measure hawkish/dovish tone of FOMC statements. Be creative!

5. **(10 points)** Plot your sentiment score, together with Tadle's and changes in policy target. Compute the correlation matrix between the three. Comment on your results.

Section 3: Measuring the impact on the bond market (40 points)

We are going to treat the changes in the FFR targets and the calculated sentiments of FOMC statements a la Tadle (2022) as if they are surprises. See Aruoba and Drechsel (2022) to see why this is not a realistic assumption.

1. **(5 points)** From FRED find the following data series: 1) EFFR, 2) DGS1MO, and 3) DGS10. Make sure that you read their description. Compute the changes of values in these series on the date of FOMC meeting from the previous day. Provide summary statistics including, number of observations, mean, median, and standard deviation.

2. **(20 points)** Let y_{it} denote the change of value in series i in date t. Let $MPOL_t$ denote the change of FFR target and $HAWK_t$ sentiment of FOMC statement in date t. Run the following three specifications of regression:

$$\begin{aligned}
&\text{specification 1} \quad && y_{it} = \alpha_i + \beta_i^p MPOL_t + \epsilon_{it} \\
&\text{specification 2} \quad && y_{it} = \alpha_i + \beta_i^s HAWK_t + \epsilon_{it} \\
&\text{specification 3} \quad && y_{it} = \alpha_i + \beta_i^p MPOL_t + \beta_i^s HAWK_t + \epsilon_{it}
\end{aligned}$$

   Tabulate the estimated value, t-stats, and goodness-of-fit for 9 regressions.

3. **(5 points)** Comment on the difference between the three specifications. Pick one as your favorite and tell us why.

4. **(10 points)** Comment on the difference across the βs of three series using your favored regression specification. Do the regression results make economic sense? Why or why not?

Section 4: Measuring the impact on the stock market (60 points)

We will use the Tadle (2022) sentiments as monetary shocks. The steps below will guide you through the Fama-MacBeth (1973) procedure to estimate the monetary policy risk premium using industry returns. Industry return data are downloaded from the French Data Library, https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

1. **(25 points)** From the value-weighted returns in the daily industry return extract the values only for dates when FOMC met. For each of the 49 industries, regress returns

on the day of the FOMC meeting on monetary shocks. Let R_{it} denote the return of industry i on date t. The regression equation is the following:

$$R_{it} = \alpha_i + \beta_i HAWK_t + \epsilon_{it}$$

Generate a table with four columns: column 1 has the industry name, column 2 has the OLS regression coefficient of αs, column 3 has the OLS regression coefficient of βs, and column 4 has the p-value for βs. Sort the table from largest to smallest βs.

2. **(25 points)** Now turn to the monthly return data and still use value-weighted returns. Separately for each month from Jan. 2000 to Aug. 2024, regress returns on βs obtained in the previous step. Let $R_{i,T}$ denote the monthly return of industry i in month T. The regression equation is the following.

$$R_{i,T} = \gamma_T + \lambda_T \beta_i + \eta_{i,T}$$

Plot λ_T over time.

3. **(5 points)** What is the interpretation of λ_T? Comment on the plot obtained in the previous question.

4. **(5 points)** Plot the R-squared in part 2 over time. Comment on your plot.