Please create a corpus of the FOMC Meeting Minutes, Fed speeches and Press Conference transcripts and evaluate the level of hawkishness and dovishness of those statements using an appropriate polarity score for each document. Recall that on **January 25, 2012 (PDF):** Participants' assumptions about the appropriate level of the federal funds rate at year-ends and in the longer run are added to the SEP. Moreover, the SEP summary accompanying the post meeting statement now includes a depiction of individual assumptions for the federal funds rate; this figure is informally known as the "dot plot." These changes coincide with the first annual release of the FOMC's Statement on Longer-Run Goals and Monetary Policy Strategy and represent a new era of FOMC communications with market participants.

Evaluate how the bond market yields, the Treasury Yield Curve spread, and other relevant financial indicators respond to the level of **Hawkishness vs. Dovishness** in Meeting Minutes, Press Conferences, and Fed speeches. Employ word dictionary and factor similarity analysis.

Write a two-page summary of your findings and submit your corresponding Python code.

Best regards,

Prof. Rodriguez

**Project Grading Rubric Below**

| **Criteria** | **Evaluation** |
| --- | --- |
| 1. Correctly differentiated between hawkish and dovish sentiment across FOMC communications using an appropriate dictionary and factor similarity analysis. (25 pts) | You've implemented the **TF-IDF** calculations correctly in order to properly compute the level of hawkishness or dovishness for each FOMC communication. |
| 2. Accurate Calculation of Statistically Significant Market Reactions to very hawkish versus very dovish speeches/communications. (15 points) | You've correctly calculated statistical significance on the days mapped to specific FOMC communications. |
| 3. Interpretation of Results and Comparison (20 points) | You have compared the results between using a word list/dictionary and factor similarity and determined which method is preferable. You have provided an insightful summary of your results. |
| 4. Code Quality and Documentation (20 points) | The code is structured well and is modular and you have included detailed comments to explain the logic for your code. |
| 5. Data Handling and Integrity (10 points) | Your data handling is good. The alignment between the communication dates, sentiment data, and market reactions are accurate and valid. |