**Sentiment analysis of Federal Open Market Committee (FOMC) minutes: are they leading indicators of a recession similar to traditional Metrics?**

**Literature Review**

1. The first article uses 2017 FOMC minutes and focuses on sentiment analysis of each member of the board’s statements inside these minutes [1]. These data represent a small subset of the minutes I use. The use of such a small subset is likely due to pre-processing complexities of attributing statements to the individual board members. Ultimately each member of the board is ranked in a correlation matrix to the sentiment of the board. That is, which members most closely represent the median view of the board in a ranking of 1 to 10.
2. The second article uses FOMC statements, which are more timely and abbreviated summaries of the meetings than the FOMC minutes [2]. This study examines the semantic similarity from meeting to meeting over time. The cosine similarity of word persistence does appear to trend with the business cycle of the economy [2]. The authors use a moving average for smoothing out a lot of the noise in the cosine similarity charts which help us identify the trend.

**Pre-Processing of Data**

The minutes are available in .htm files through the Federal Reserve Public Website. Using the Beautiful Soup python library, I collected links to minutes from all federal reserve meetings between 1996 and 2019. I then looped through each one of these URLs collecting the text from each page and storing this in a data frame along with the meeting date, the year and the length on the meeting minutes text. I saved the final data frame as a .json in order to avoid re-running this step. I visually spot checked several rows to verify that the text matched the date it said on the public website.

**Exploratory Data Analysis**

Once the data were processed, the Vader Sentiment analysis python library was used to determine neutral, positive and negative scores of each of the minutes. I created a table of summary statistics based on the sentiment scores of the minutes. I also plotted the scores over time using matplotlib, and also did the same with length of the meeting minutes. And finally I measured the correlation between the scores and employment level, the job openings rate and the hires rate as these metrics tend to trend with the business cycle and in the case of the employment level are used to determine recession dates [3].

One of the three principal reasons to explore developing a model using FOMC minutes and assessing if such a model could serve as a leading indicator of a recession. Moving forward, using moving averages smooth out some of the seasonal volatility in the meeting minutes sentiment [2].

**Reference List:**

[1] H. Ramachandran and D. DeRose Jr., “A Text Analysis of Federal Reserve meeting minutes” arXiv preprint arXiv:1805.07851, 2018, Available: <https://arxiv.org/ftp/arxiv/papers/1805/1805.07851.pdf>. [Accessed February 12, 2020].

[2] M. Acosta and E. Meade, “Hanging on every word: Semantic analysis of the FOMC's postmeeting statemen” FEDS Notes, September 30, 2015. Available: <https://www.federalreserve.gov/econresdata/notes/feds-notes/2015/semantic-analysis-of-the-FOMCs-postmeeting-statement-20150930.html>. [Accessed February 12, 2020].

[3] S. Ng and J. Wright, “Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling” Journal of Economic Literature 2013, 51(4), 1120–1154 <http://dx.doi.org/10.1257/jel.51.4.1120>. [Accessed February 12, 2020].