**Ana Krishnan**

**Project Overview** *[Maximum 100 words]*

I used the Federalist Papers from the Gutenberg Project. I saved it as a text file, and then took the first 10 papers and found the most used words in them. I also ran a sentiment analysis that took the average sentiment of each of the three authors (Hamilton, Jay, and Madison). I hoped to create something that would clearly show the difference in tones between the different authors and would show the most used words in the first 10 papers.

**Implementation** *[~2-3 paragraphs]*

First, I began with the Federalist Papers. I saved the Papers to a .txt file, and then found the lines where each one began and ended. I did this for the first 10 papers, and then saved each group of lines into its own text file by joining them to form one line and turning them into strings. I then tokenized each one of the papers and removed punctuation, spaces, and left the option of getting rid of common words. Then I rewrote these lists in a string to the files. I used these strings to calculate the frequency of each word. In addition to this, I ran sentiment analysis on the previously selected lines of the entirety of the Federalist Papers, and got an average sentiment for Hamilton and Jay (Madison only wrote one paper in the first 10). I chose to save the file as a .txt instead of pickling it because I was pretty confused as to what pickling did. Just using a .txt file taught me how to modify a .txt file in a Python script and how to utilize .txt in Python.

The input for all of the functions is the list "papers" which contains the first ten papers in .txt files. Each function returns a different thing which is plugged into the next function.

**Results** *[~2-3 paragraphs + figures/examples]*

[('the', 134), ('of', 106), ('to', 72), ('and', 40), ('be', 34), ('that', 28), ('in', 27), ('a', 25), ('will', 25), ('it', 20), ('which', 18), ('not', 15), ('this', 14), ('i', 14), ('by', 14), ('their', 14), ('is', 13), ('for', 13), ('are', 12), ('than', 11), ('from', 11), ('may', 11), ('an', 11), ('its', 10), ('as', 10), ('have', 10), ('your', 10), ('on', 9), ('government', 9), ('those', 9), ('all', 9), ('constitution', 8), ('we', 8), ('who', 8), ('at', 8), ('my', 8), ('men', 8), ('shall', 7), ('more', 7), ('union', 7), ('new', 7), ('you', 7), ('any', 6), ('they', 6), ('people', 6), ('state', 6), ('good', 6), ('or', 6), ('upon', 6), ('with', 6), ('has', 6), ('appearance', 5), ('liberty', 5), ('many', 5), ('truth', 5), ('general', 4), ('own', 4), ('subject', 4), ('if', 4), ('one', 4), ('can', 4), ('been', 3), ('arguments', 3), ('am', 3), ('discussion', 3), ('well', 3), ('our', 3), ('interest', 3), ('opposition', 3), ('these', 3), ('must', 3), ('equally', 3), ('political', 3), ('certain', 3), ('zeal', 3), ('too', 3), ('nothing', 3), ('great', 3), ('however', 3), ('every', 3), ('much', 3), ('principles', 3), ('so', 3), ('ambition', 3), ('made', 3), ('other', 3), ('right', 3), ('no', 3), ('spirit', 3), ('under', 3), ('least', 3), ('judgment', 2), ('reflection', 2), ('either', 2), ('former', 2), ('proposed', 2), ('passions', 2), ('public', 2), ('open', 2), ('offered', 2), ('rights', 2), ('even', 2), ('necessity', 2), ('parties', 2), ('idea', 2), ('feel', 2), ('consequences', 2), ('view', 2), ('after', 2), ('most', 2), ('question', 2), ('led', 2), ('into', 2), ('adoption', 2), ('there', 2), ('might', 2), ('empire', 2), ('choice', 2), ('power', 2), ('confederacies', 2), ('several', 2), ('decision', 2), ('interests', 2), ('wrong', 2), ('course', 2), ('inducements', 2), ('would', 2), ('motives', 2), ('views', 2), ('introduction', 2), ('doubt', 2), ('states', 2), ('apt', 2), ('forgotten', 2), ('number', 2), ('hand', 2), ('efficiency', 2), ('hope', 2), ('security', 2), ('conduct', 2), ('them', 2), ('true', 2), ('same', 2),

('already', 2), ('intentions', 2), ('class', 2), ('themselves', 2), ('against', 2), ('interesting', 2), ('whether', 2), ('head', 2), ('country', 2), ('utility', 2), ('welfare', 2), ('us', 2), ('side', 2), ('observations', 2), ('give', 2), ('oppose', 2), ('but', 2), ('moderation', 2), ('jealousy', 2)]

Alexander Hamilton wrote Federalist Papers 1 and 6-9. This the average sentiment of those papers: {'neu': 0.7859999999999999, 'neg': 0.0854, 'pos': 0.12840000000000001, 'compound': 0.9686400000000001}

John Jay wrote Federalist Papers 2-5. This the average sentiment of Papers 2-5: {'neg': 0.08150000000000002, 'pos': 0.167, 'compound': 0.99885, 'neu': 0.7517499999999999}

James Madison wrote Federalist Paper 10. This the average sentiment of Paper 10: {'neg': 0.075, 'neu': 0.723, 'compound': 0.9999, 'pos': 0.202}

My code analyzes first 10 Papers in the Federalist Papers. I first saved the Papers to a text file and then found the frequencies of the words used and also ran a sentiment analysis on it. "The" is the most used word in the first 10 Federalist Papers, used a whopping 134 times. The most used word that isn't along the lines of "the", "an", "and", etc. are "government" and "constitution", which are used 9 and 8 times respectively. The sentiment analysis showed that Hamilton was the most neutral and negative author, while Madison was the most positive (but this may have been skewed in his favor because he only wrote one paper in the section that I was testing). Overall, the papers all leaned positive.

**Reflection** *[~1 paragraph]*

From a process point of view, what went well? What could you improve? Other possible reflection topics: Was your project appropriately scoped? Did you have a good plan for unit testing? How will you use what you learned going forward? What do you wish you knew before you started that would have helped you succeed?

Actually having a working final product was really exciting for me. I worked on and wrestled with this for a long time, but I couldn't get some parts to work. I think I should have done something a bit easier, because I have virtually no coding experience and had a hard time figuring out what I should do to my code to make it work. I'm really excited that I figured out how to do everything by myself. I originally wanted to create something that would take all of the Federalist Papers and output who wrote which one, but I struggled a lot with the code. Going forward, I will be sure to look out for small syntax errors, and I have learned how to Google for help in a much more concise way. I wasn't sure about how to go about unit testing my data. In the future, I would like to analyze all of the Federalist Papers, instead of just the first 10.