Weekly Update 3

Hien An Ngo

February 23, 2026

Weekly Progress (5 points)

1. Downloaded contribution receipts for all candidates manually.

2. Spoke with OpenSecrets staff about issues with FEC data / deduplication.

3. Merged candidate receipts.

4. Continued to work on downloading candidate tweets.

5. Working on understandinghow to deduplicate data.

Challenges (3 points)

1. The FEC limits manual downloads on its website to max 500,000 records, which creates a challenge for candidates who have more than 500,000 itemized contributions between 2015-25. I had to download records in multiple different loads for certain candidates.

2. Because of the size of each file, I frequently hit rate limits on the FEC website and had to wait or try to clear my cookies.

3. Due to the size of the files, I wasn't able to push the candidate receipts to git.

Code Commits (2 points)

https://github.com/GW-datasci/26-spring-ngo-a/commits/main/

1. Tried to commit FEC data, but committed summary data for all of the files.

2. Committed script that merged files as well as script to summarize them.

References

[1] Mehran Shakarami. Twikit main.py. https://github.com/mehranshakarami/AI_Spectrum/blob/main/2024/Twikit/main.py, 2024. GitHub repository.

[2] Twikit Documentation. Twikit documentation. https://twikit.readthedocs.io/en/latest/twikit.html, 2024.