DS8006: Lab 4 "Introductory Social Network Analysis with Netlytic"
Student's name: NAJLIS, BERNARDO (#500744793)

1. Briefly explain what topics you chose for this lab and what were your search keywords/hashtags used for data collection.
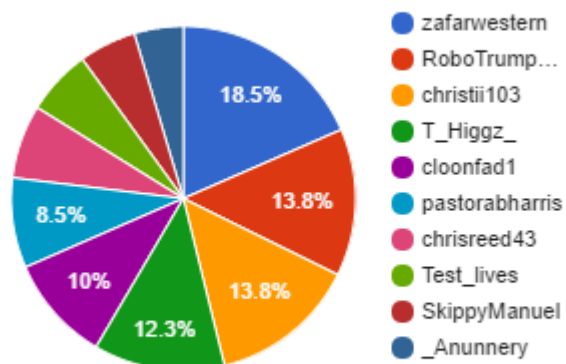
The topic chosen was the President of the United States Address to the Congress, and the hashtag used in the search was #JointAddress (that was also trending topic on that same day). The address was done on February 28, the data was collected the next day on March 1st. and an extra load on March 6th.

2. Include a brief description of the collected dataset: # of posts, # of users, etc., and any other relevant information that you think would be useful to know about this dataset.

The dataset includes 2,000 tweets from 1,639 unique twitter users / posters. Doing the basic text analytics that Netlytic allows, the top 10 more frequently used words are:
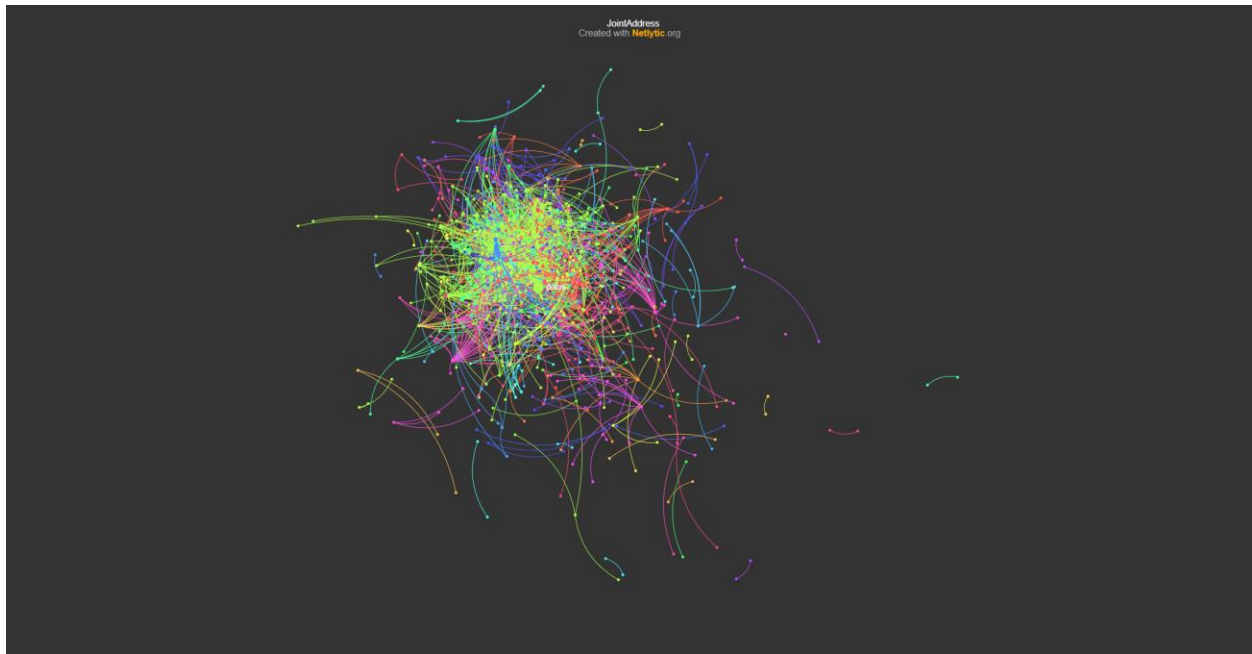
1. *#jointaddress*
2. *Trump*
3. *@potus*
4. 😺😺😺
5. *Rejoicing*
6. *Demoncrats*
7. *Freed*
8. *@tponews*
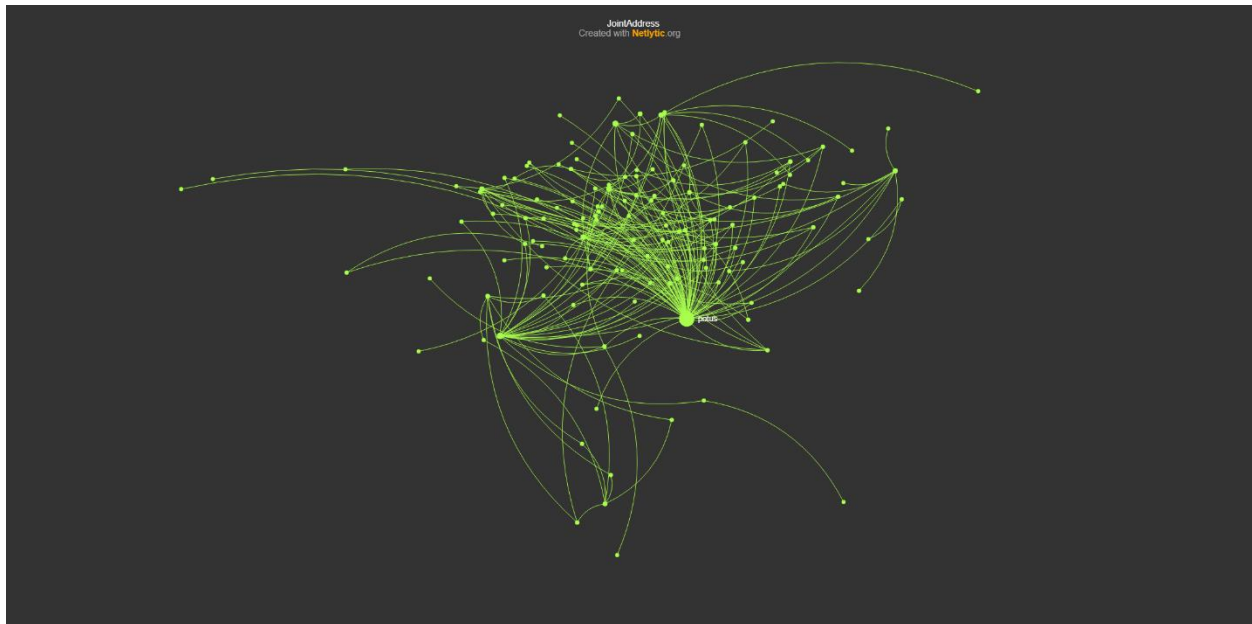9. *https://t.co/gudlfogknb*
10. *#jointsession*
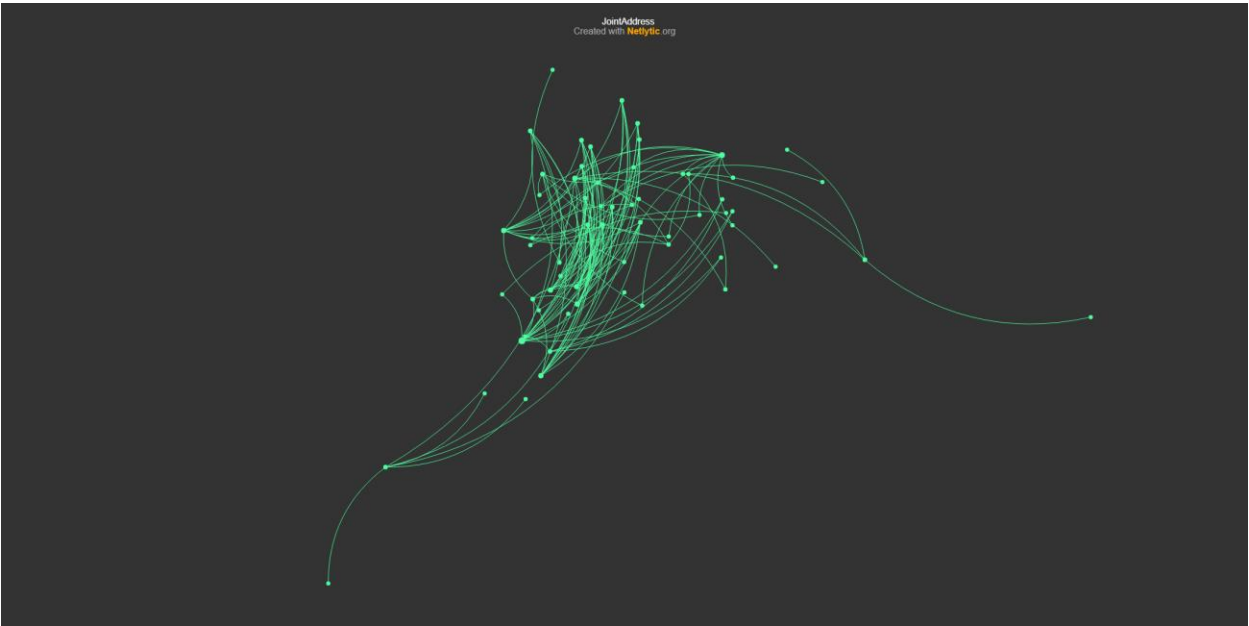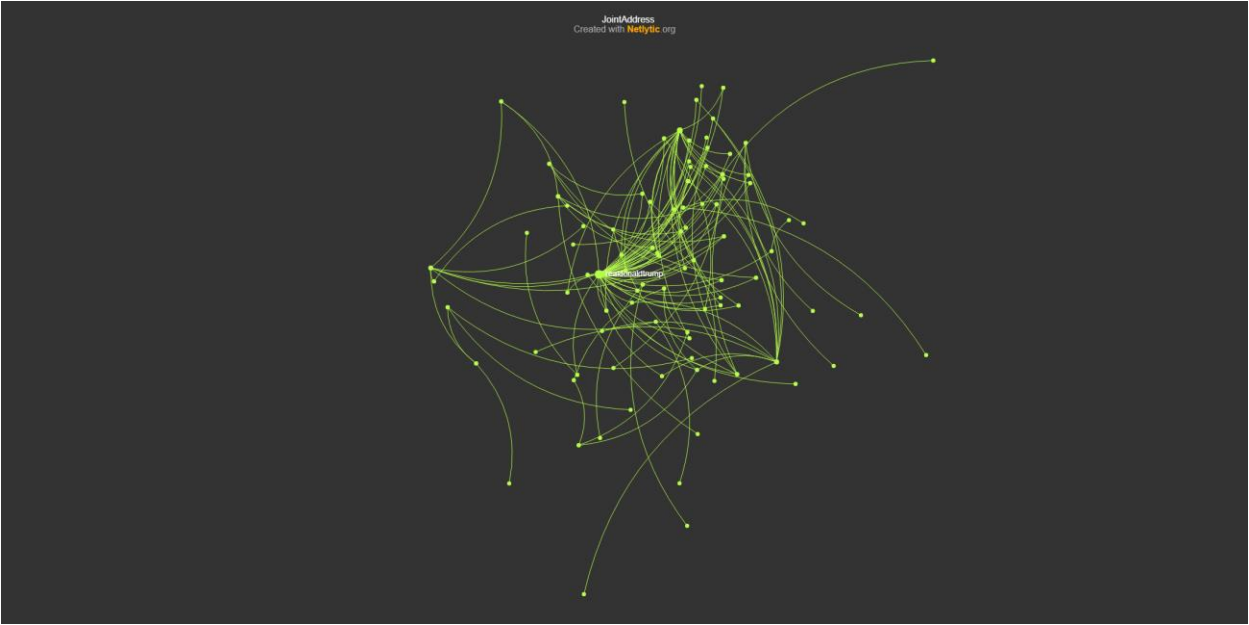
These are the top 10 posters:

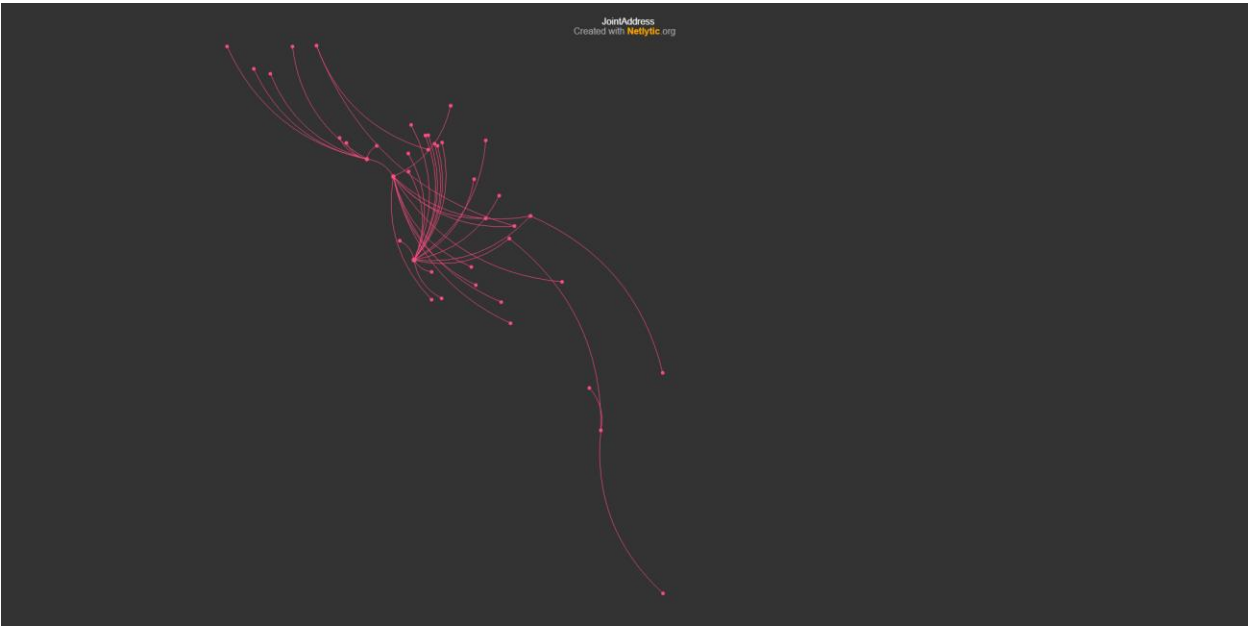**3. Include a screenshot(s) of the resulting network visualization.**

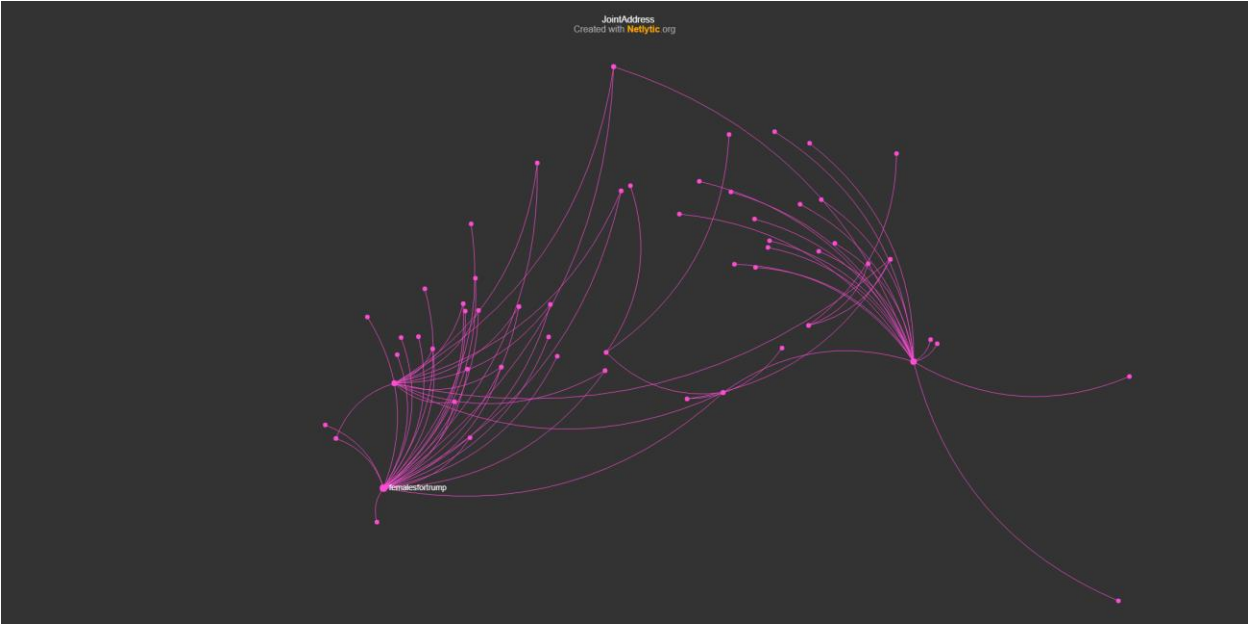*This is the main network visualization generated by Netlytic*

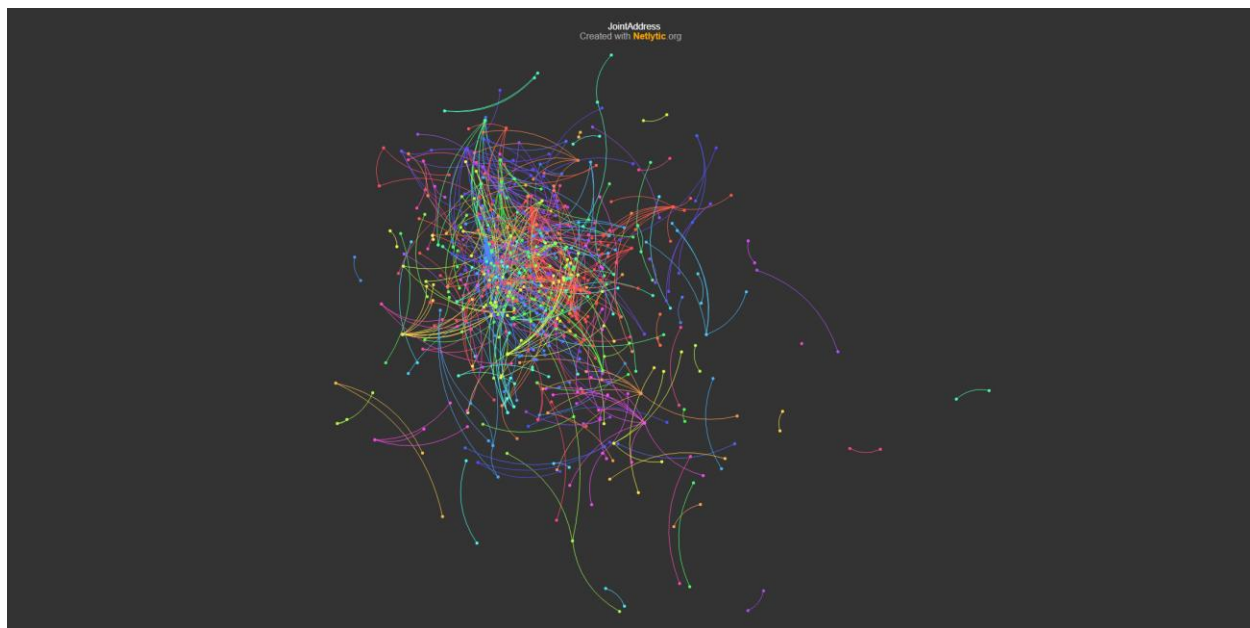*The following are independent visualizations for each one of the clusters detected.*

JointAddress
Created with Netlytic.org

JointAddress
Created with Netlytic.org

JointAddress
Created with Netlytic.org

JointAddress
Created with Netlytic.org

JointAddress
Created with Netlytic.org

4. Use the resulting network visualization to identify the top 10 key influencers (both individual and organizational accounts), and discuss why the are influential in this network. You can use the following table structure. Also <u>discuss what criteria were used to identify the 10 key influencers</u>.

These are the top 10 influencers and the corresponding criteria used for each:

#	Influencer @handle	Criteria used for selection	Comments / Why are they influential ?
1	@potus	Main node on Cluster #1	Official account for the President of the United States.
2	@realdonaldtrump	Main node on Cluster #2	Personal account for the President of the United States, Donald J. Trump.
3	@cnn	Main node on Cluster #3	Official account of CNN, a major news media outlet that covered the event and generally opposes Donald Trump.
4	@femalesfortrump	Main node on Cluster #4	Account for a political group that supports the current President of the United States.
5	@barryobozo	Main node on Cluster #5	Probably a bot account, as of March 6th the account doesn't exist anymore.
6	@whitehouse	Secondary node on Cluster #1	Official account for the White House, residence of the President of the United States.
7	@votetrumppics	Secondary node on Cluster #2	Account managed by Donald Trump's supporters with political message on images.
8	@foxnews	Secondary node on Cluster #3	Official account of Fox News, a major news media outlet that covered the event and generally supports Donald Trump.
9	@drmartyfox	Secondary node on Cluster #4	Personal account for an individual who expresses as Donald Trump supporter.

10	@realjack	Secondary node on Cluster #5	Personal account for an individual who expresses as Donald Trump supporter.

5. What was the most challenging part of this lab?

The most challenging part was finding a topic interesting enough to generate a large dataset for the lab. After assessing options that were more interesting from a personal perspective (as part of a non-profit organization I collaborate with) I had to switch to other topic that was more mainstream in order to collect a larger dataset.