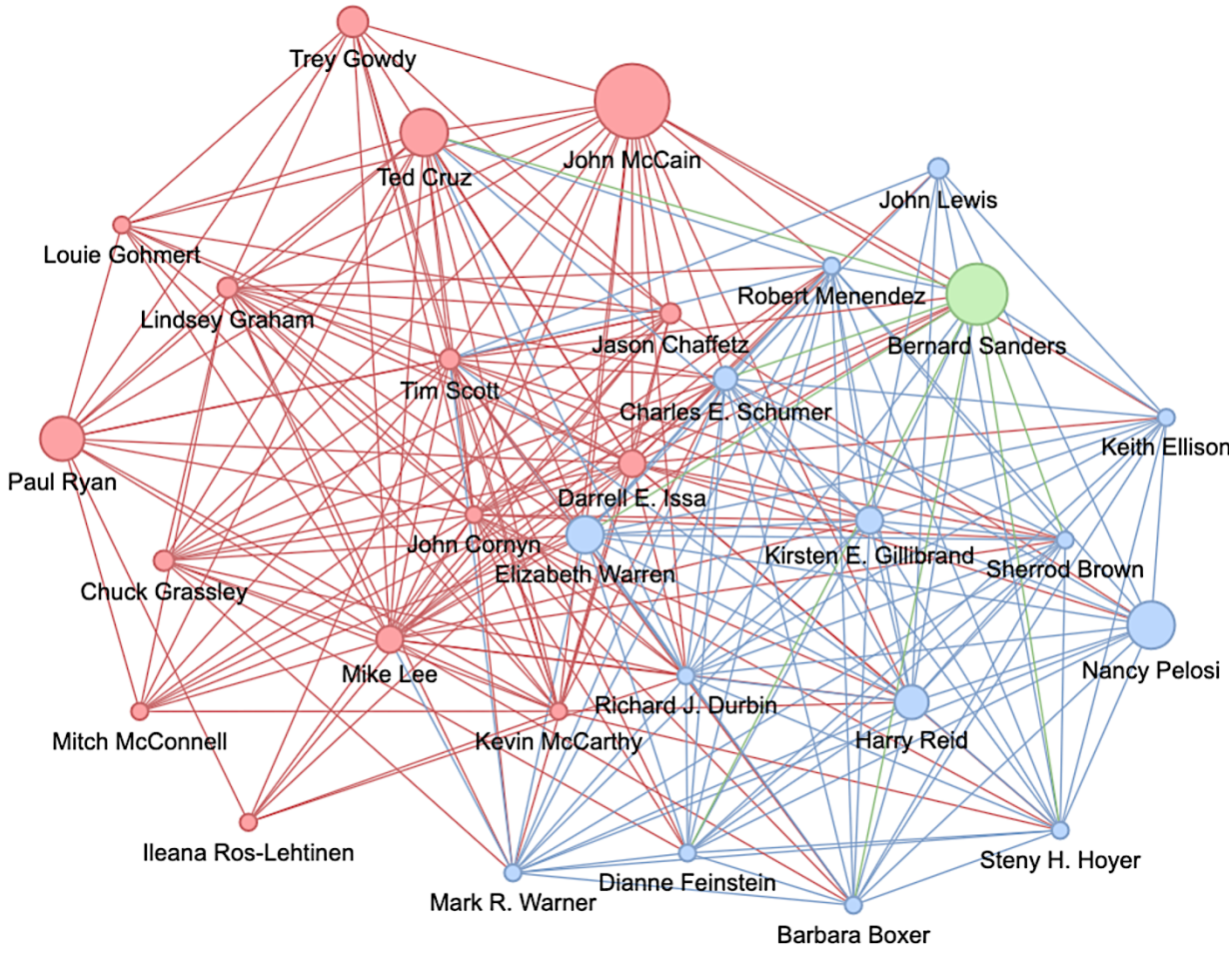
Data 22700
Data Visualization and Communication
Trimble
Ashley Hitchings

Homework 5: Network Visualization

Question: Who are the most-followed Congresspeople on Twitter, to what parties do they belong, and to what extent do politicians interact within their party versus across the aisle on social media platforms?

Twitter Interactions Between the Most-Followed U.S. Congresspeople (2015)



Caption: Among the thirty most-followed Congresspeople on Twitter in 2015, fifteen belonged to the Republican party, fourteen to the Democratic party, and one to the Independent party, representing a balanced composition across party lines. Senator John McCain had the largest following at nearly 2m followers and Sen. Bernie Sanders trailed at 1.2m, followed successively by House Minority Leader Nancy Pelosi, Sen. Ted Cruz, Speaker of the House Paul Ryan, and Senator Elizabeth Warren. For both Democrats and Republicans, most Congresspeople primarily interacted with members of their own party on Twitter, a trend especially prominent for party leaders Pelosi and Ryan, who appear at opposite sides of the

network. Nevertheless, there exists a large amount of cross-party digital interaction senators and representatives at the lower end of the top 30, including Sen. Tim Scott, Sen. Chuck Schumer, and Sen. Dick Durbin. Additionally, a subset of politicians who hold relatively radical positions within their party, including Senators Warren and Sanders, appear to display high cross-party Twitter engagement, though further research is necessary to substantiate this finding. Data Source: Pablo Barbera (GitHub: pablobarbera), "Data Science Workshop - Congress Twitter Network," https://github.com/pablobarbera/data-science-workshop/tree/master/sna/data. Note: though the source does not specify when the data was pulled from Twitter, the files were uploaded to GitHub in 2015, which is assumed to be the year of data collection.

For my network visualization of Congressional Twitter interactions, I initially began with a dataset containing 517 congresspeople; I chose to cut the data to only include the top 30 most-followed accounts to prevent the network graph from becoming overly cluttered, which would render it more difficult for viewers to identify, reference, and draw inferences about individual nodes. Using NetworkX, I shaded nodes according to party to make observing high-level patterns in intra/inter-party Twitter engagement clearer, with red nodes representing Republicans, blue nodes representing Democrats, and green nodes representing Independents (in this cut of the data, only those three parties were represented). After my first iteration of this graphic, I re-shaded the nodes with pastel versions of red, green, and blue, as the colored edges were initially too dark and made it difficult to read the node labels. Each node is labeled with the name of the member of Congress represented, and node areas correspond to the size of a Congressperson's following (larger nodes indicate more Twitter followers). The weights of the line have no significance in this graphic, as they are all a uniform thickness; while more data on the number of Twitter interactions between each pair of nodes would have been interesting to analyze, it was not provided in the dataset. Finally, I toggled with the physics pre-sets in Pyvis, increasing spring length, decreasing overlap, and making adjustments to central gravity and other parameters to improve viewability and prevent graphical mishaps like overlapping labels. On the interactive html file attached with my submission, viewers can also click on or hover over nodes to access additional information about that Congressperson, including their specific number of Twitter followers, as well as a metric labeled "Degree," which represents the number of other Congresspeople they are connected to in the network.