

POLITICAL POLARIZATION ON R/POLITICS AND R/THE_DONALD

Sarah Bolden | IST 719 | School of Information Studies | Syracuse University | April 2019

CONTEXT & MOTIVATION

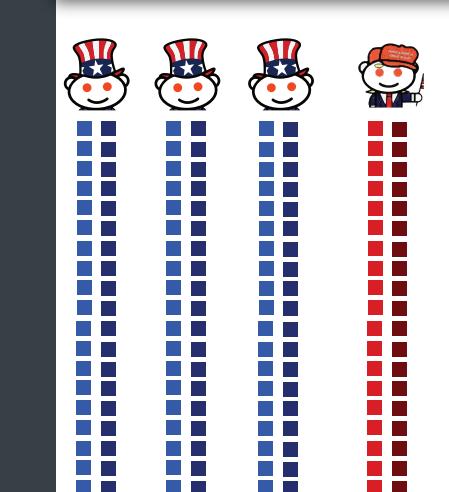
This project examines how partisan Reddit networks differed in their reactions to Dr. Christine Blasey Ford's and Judge Brett Kavanaugh's testimonies to the United States Senate Judiciary Committee on 9/27/18. Looking at live megathreads on r/Politics (progressive) and r/The_Donald (conservative), this project evaluates whether the subreddit(s) and conversation(s) that users participate in reveal evidence of polarization.



QUESTIONS

- Are there meaningful differences in the network structure of each subreddit and thread? Does network centrality correlate with other measures (such as comment score)?
- To what extent – if any – does the quantity of, and reception to, content vary across each subreddit and thread?
- What topics surface across each subreddit and thread? Are there noteworthy differences across each subreddit and thread?

DATA



COLOR LEGEND

r/The_Donald: controversial	r/The_Donald: top
r/Politics: controversial	r/Politics: top

From each megathread I scraped the top 20 comment threads sorted by "top," and the top 20 sorted by "controversial." Each comment thread (consisting of various numbers of comments) is represented on the left as a square.

The final data frame, after adding columns allowing me to subset by megathread, comment thread type, and individual comment threads, had 3240 rows and 21 columns. Metadata I collected includes: comment author; comment score; date posted; number of upvotes; controversiality; and network structure (for each comment thread).

The only cleaning I performed was for text analysis: I removed stopwords; special characters; white space; and converted comments to lowercase.

SOURCES

IMAGES

Christine Blasey Ford: <https://www.vox.com/photos-christine-blasey-ford-hearing-brett-kavanaugh-709b170-cf4a-4297-9431-12e3a620526b.html>

Brett Kavanaugh: <https://www.vox.com/policy-and-politics/2018/9/27/1791652/brett-kavanaugh-senate-testimony-christine-blasey-ford>

Reddit logo: <https://www.competethemes.com/blog/create-site-like-reddit/>

r/Politics logo: <https://rslashpolitics.com/rslashpolitics>

r/The_Donald logo: https://en.wikipedia.org/wiki/r/The_Donald

R CODE

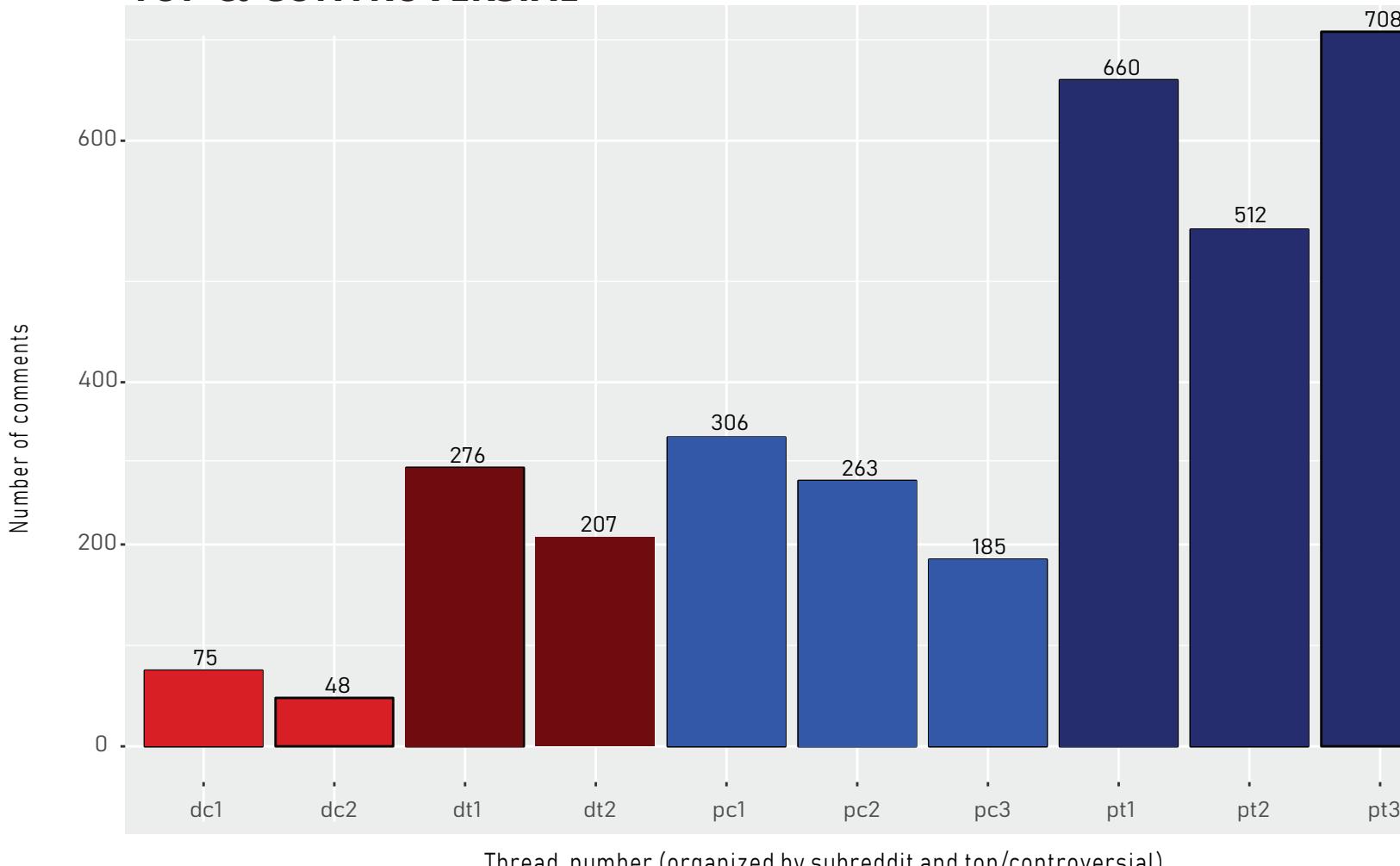
TF-IDF: <https://www.tidytextmining.com/tfidf.html>

Correlations: <https://www.tidytextmining.com/ngrams.html>

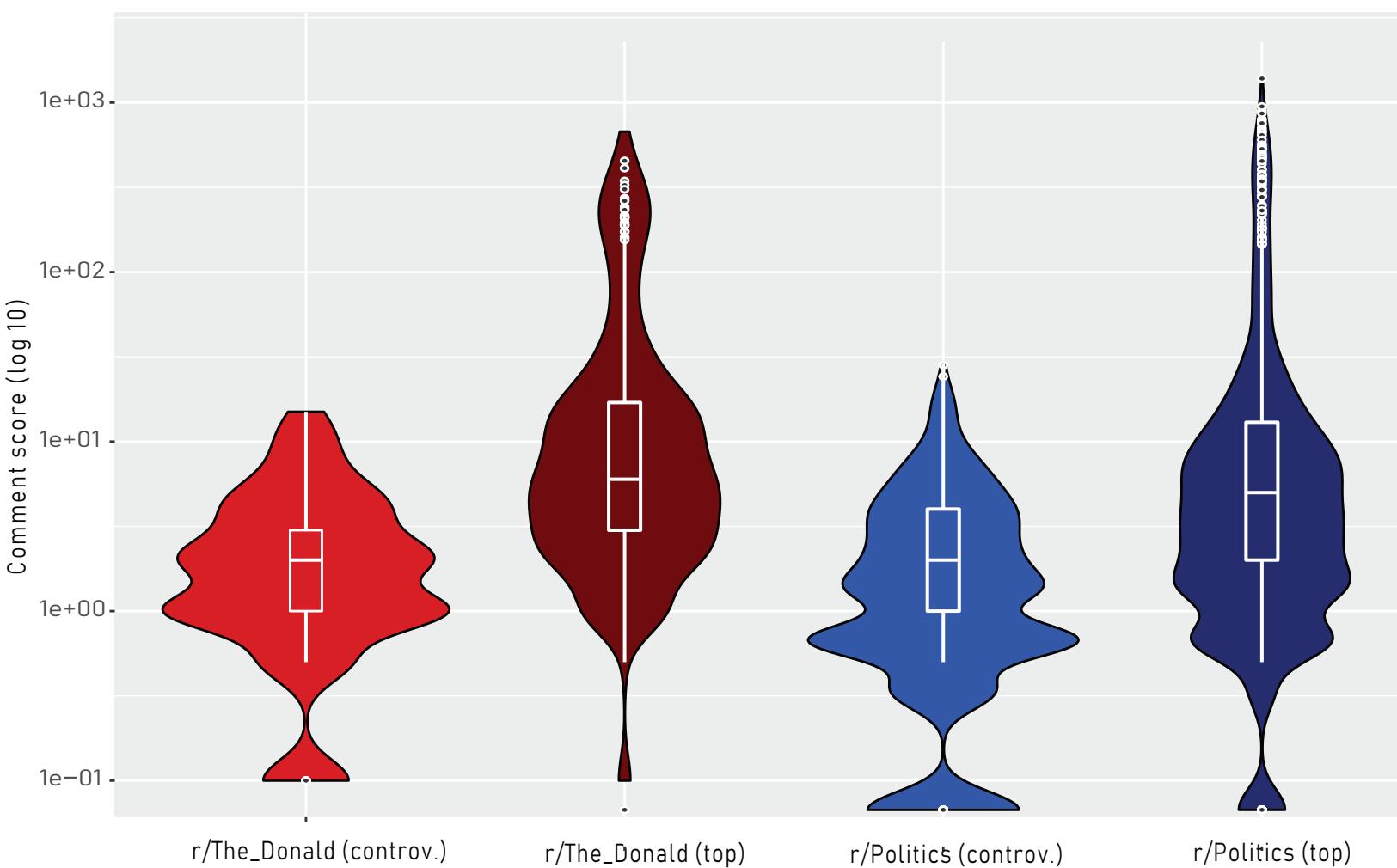
Chord diagrams: https://jokergoo.github.io/circulize_book/book/the-chorddiagram-function.html

DISTRIBUTIONS

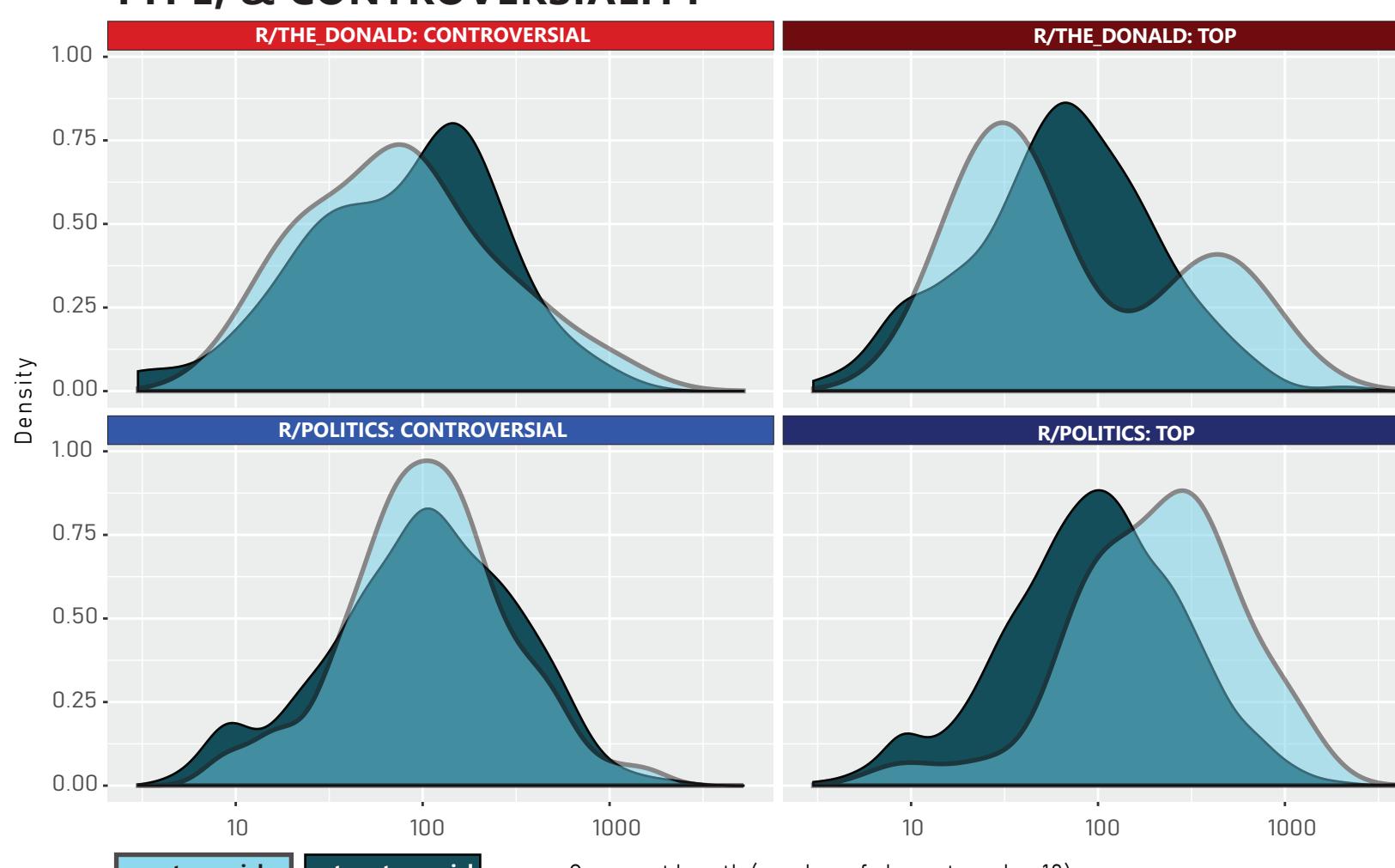
TOTAL NUMBER OF COMMENTS PER MEGATHREAD, ORGANIZED BY TOP & CONTROVERSIAL



DISTRIBUTION OF COMMENT SCORES BY SUBREDDIT & THREAD TYPE



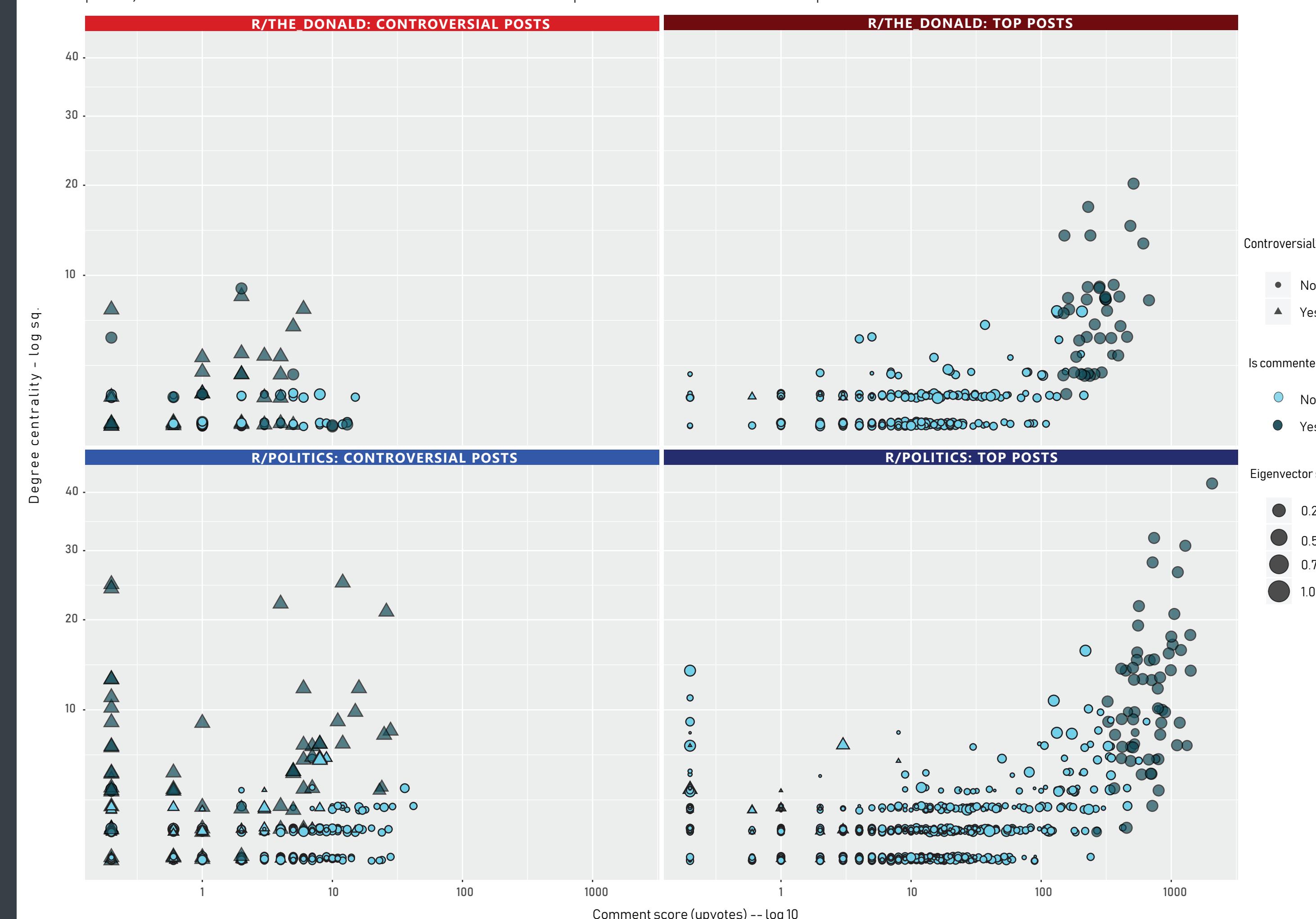
DISTRIBUTION OF COMMENT LENGTH BY SUBREDDIT, THREAD TYPE, & CONTROVERSIALITY



HOW DO NETWORKS VARY IN STRUCTURE/RECEPTION?

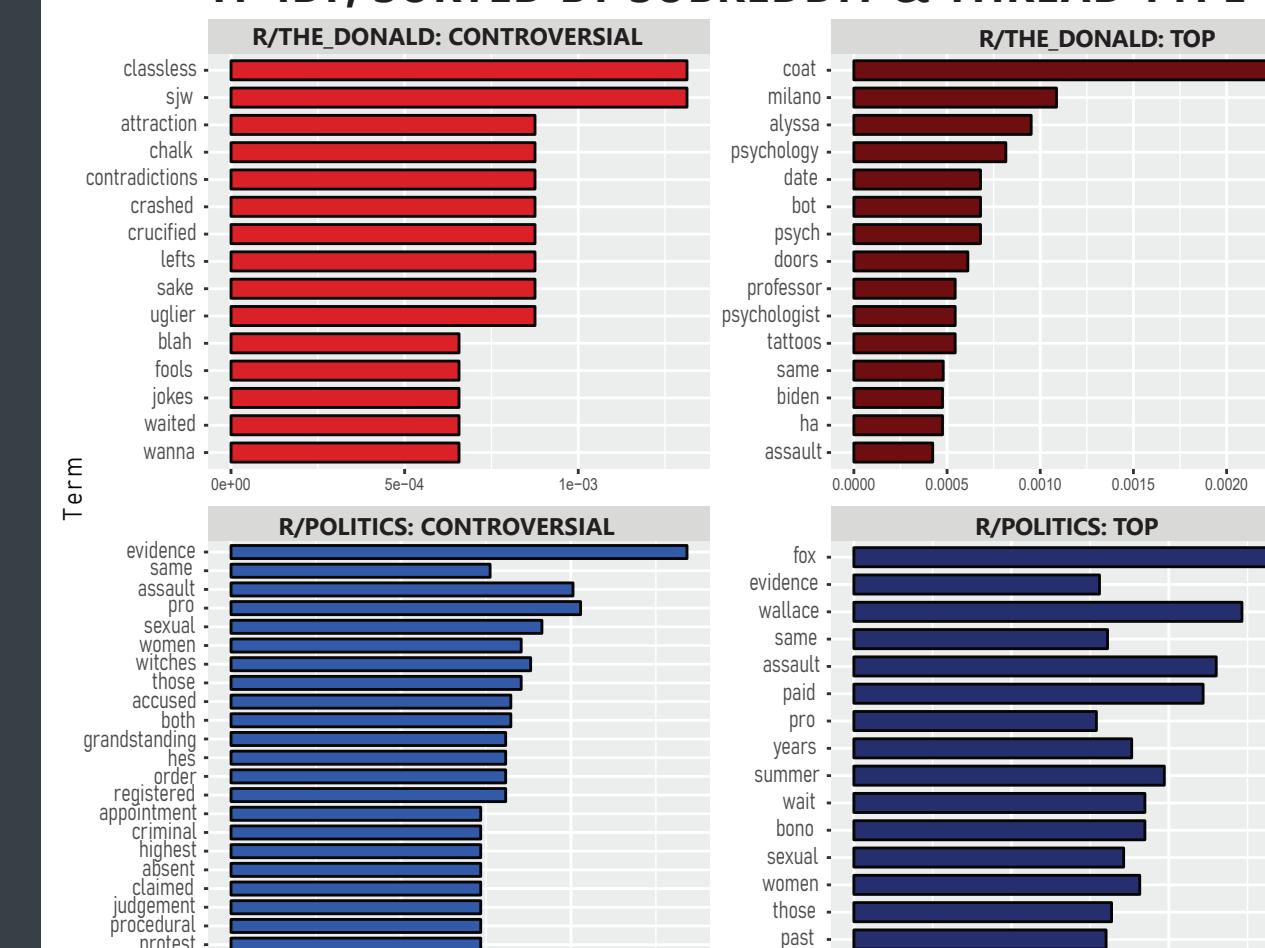
COMMENT SCORE BY DEGREE CENTRALITY, ORGANIZED BY THREAD TYPE

Comments in controversial threads have a wide range of degree and eigenvector centrality scores, despite having lower comment scores on average. Controversial comments from r/Politics have higher a larger range of degree centrality scores than controversial comments from r/The_Donald. OP (original poster) comments from controversial threads have less clustered upvotes than OP comments from top threads.

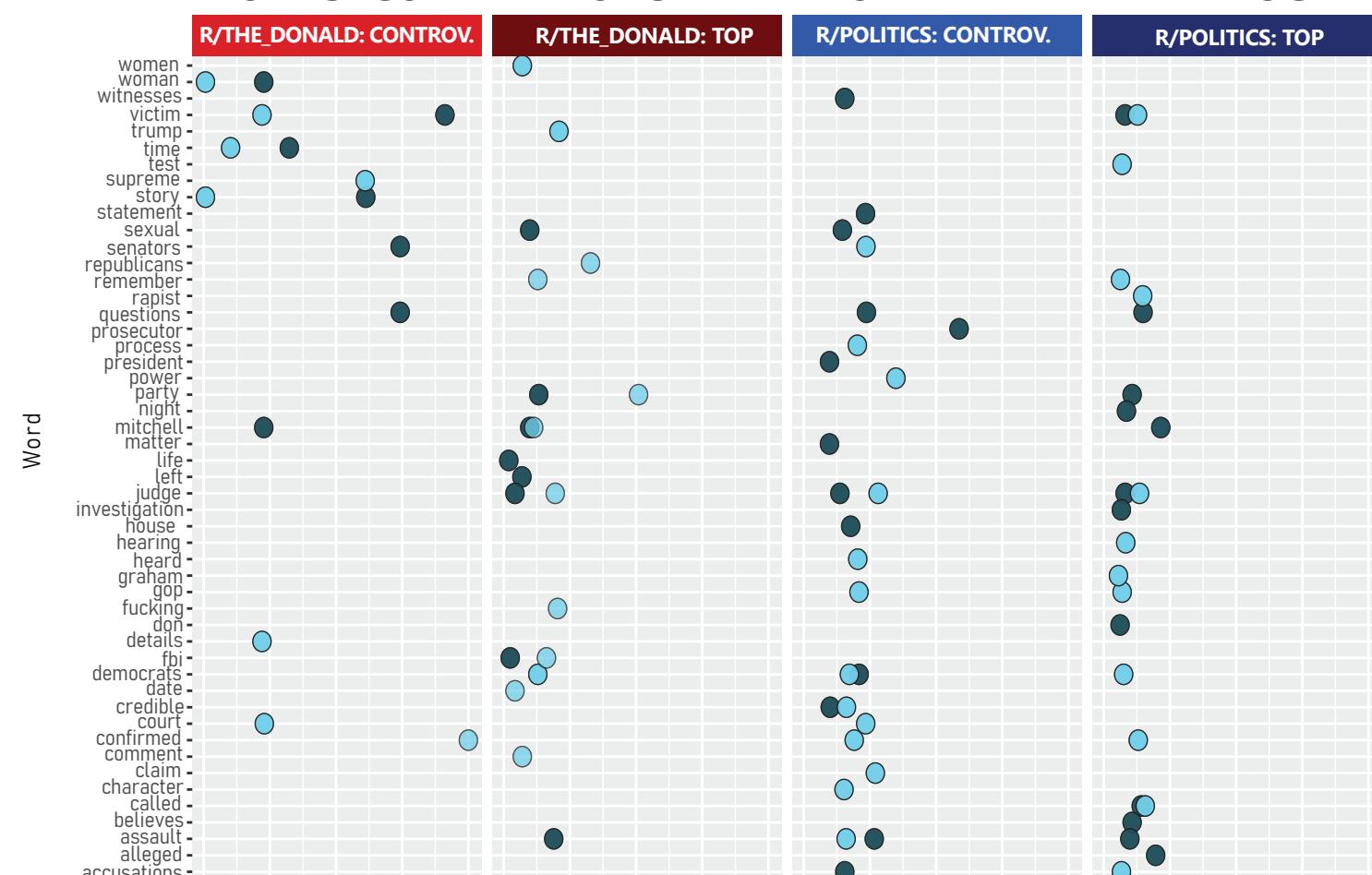


TEXT

TF-IDF, SORTED BY SUBREDDIT & THREAD TYPE

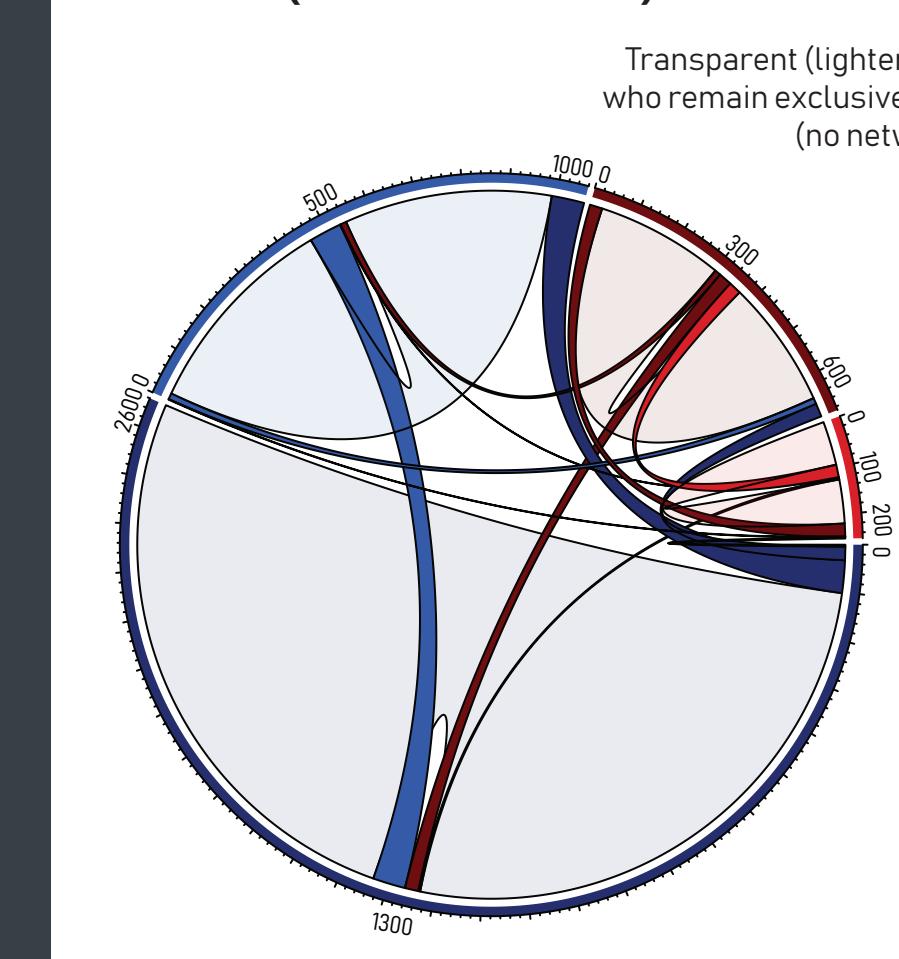


WORDS' CORRELATIONS WITH "FORD" AND "KAVANAUGH"

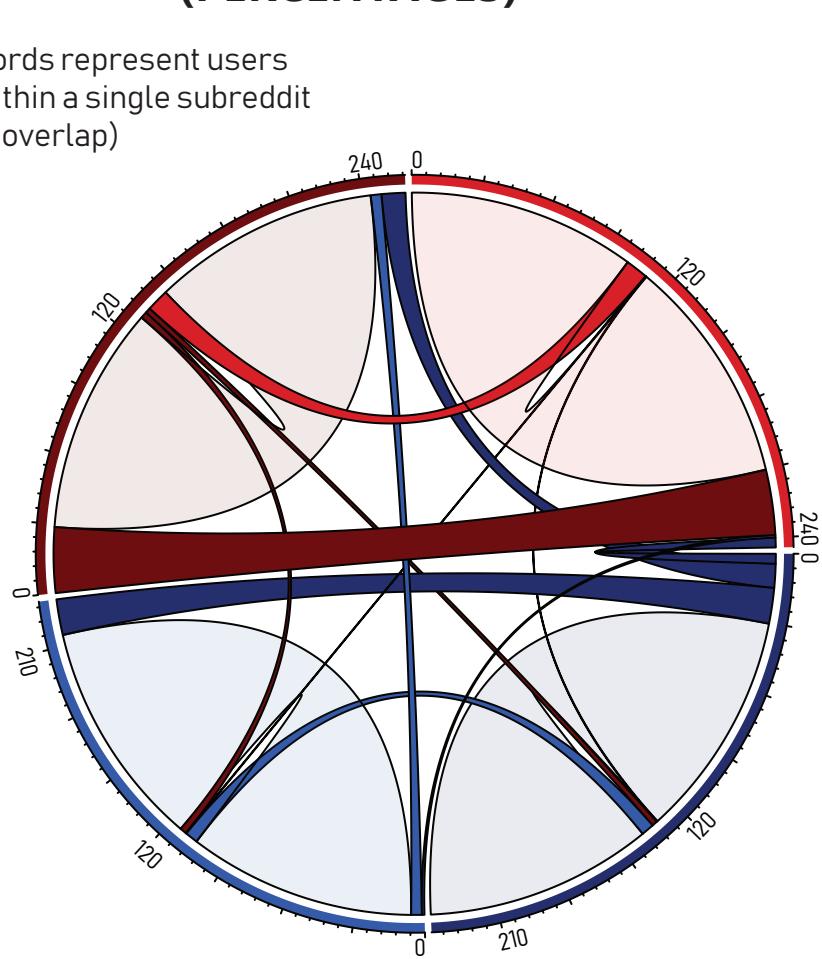


NETWORK

UNIQUE NETWORK OVERLAP (RAW NUMBERS)

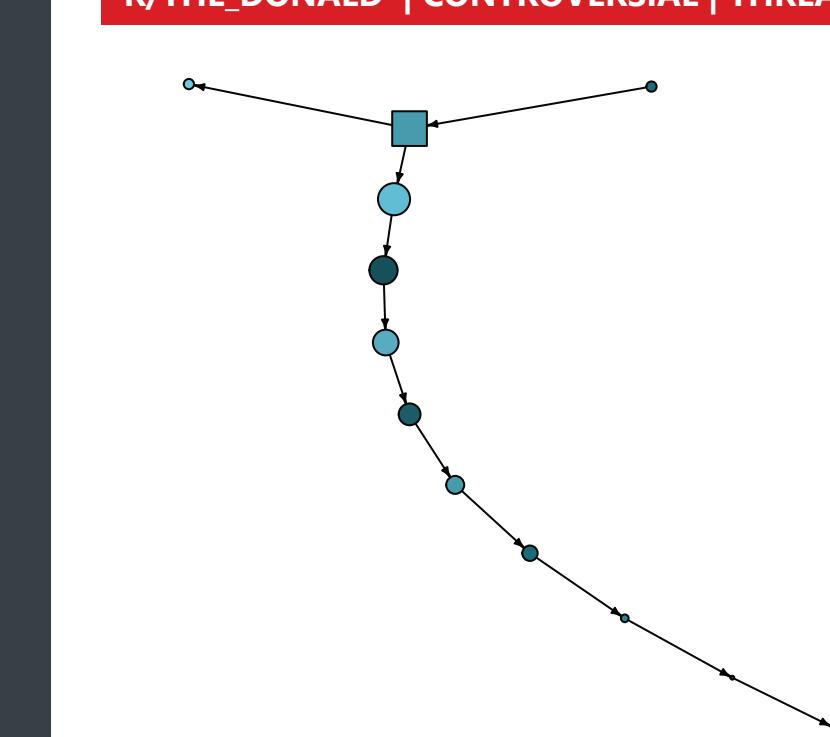


UNIQUE NETWORK OVERLAP (PERCENTAGES)

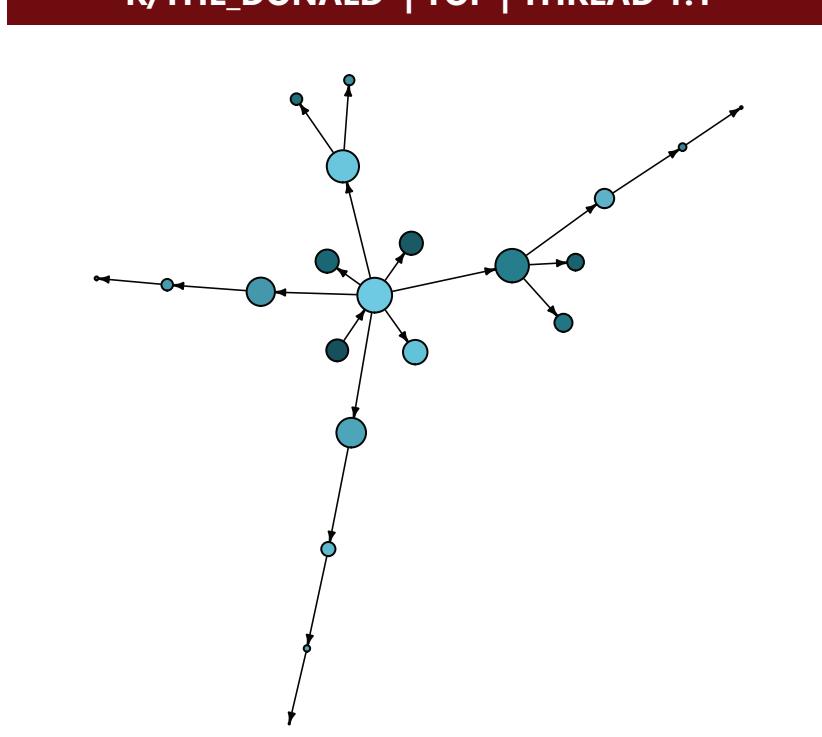


EXAMPLE NETWORKS AND COMMENTS

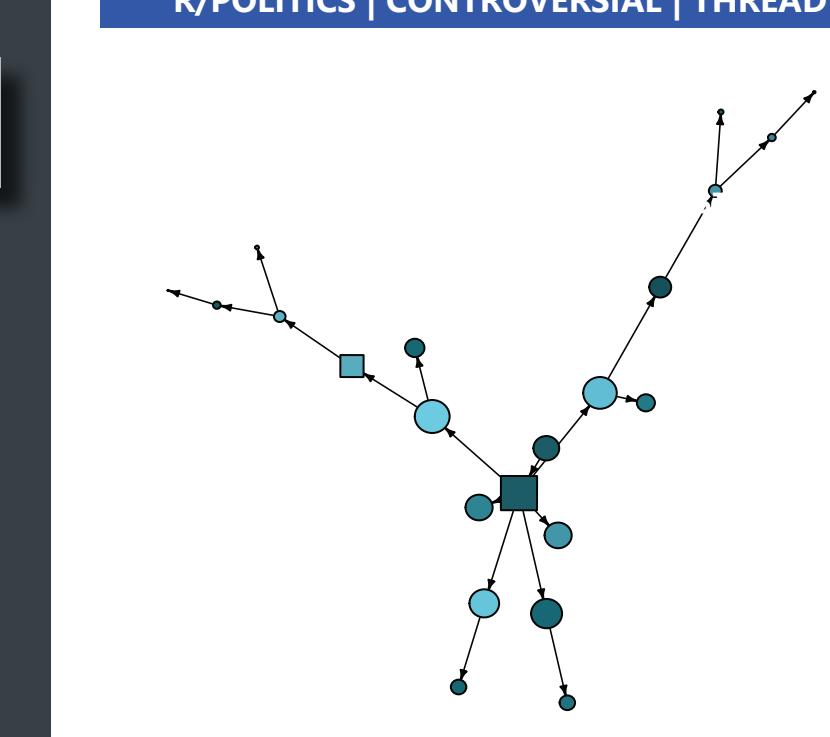
R/THE_DONALD | CONTROVERSIAL | THREAD 1.1



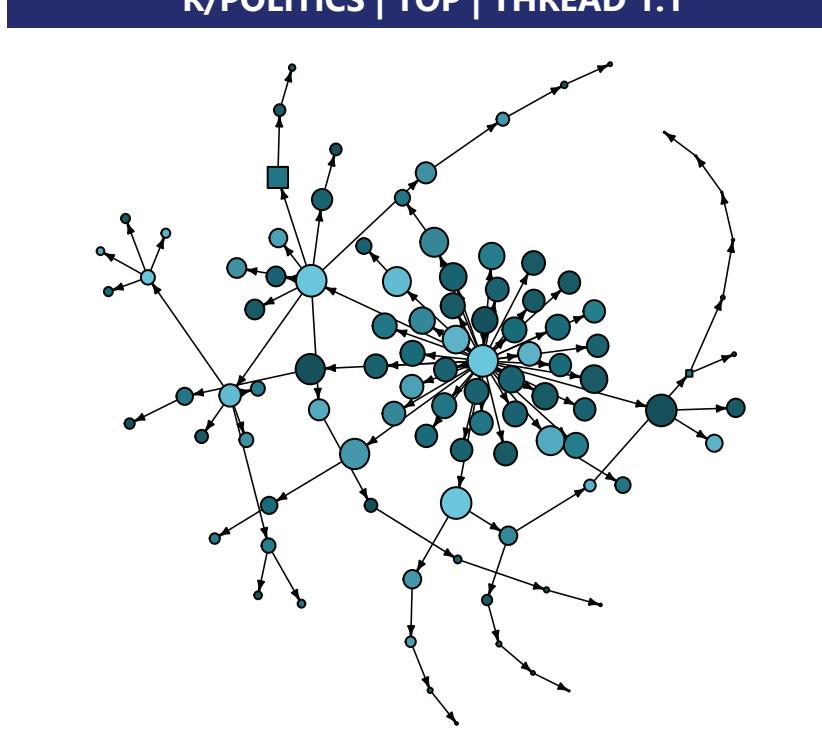
R/THE_DONALD | TOP | THREAD 1.1



R/POLITICS | CONTROVERSIAL | THREAD 1.1



R/POLITICS | TOP | THREAD 1.1



NETWORKS LEGEND

COLOR: COMMENT SCORES, LOW (DARK) TO HIGH (LIGHT)
SIZE: EIGENVECTOR SCORES, SMALL TO LARGE
SHAPE: CIRCLES (NOT CONTROVERSIAL); SQUARES (CONTROVERSIAL)