Assignment #7

This homework asks you to tell a story about a subject of much debate in the social sciences – what is culture? To answer this question, we will look at how culture has been used in sociology, specifically the journals *American Journal of Sociology*, *American Sociological Review*, & *Social Forces*, and political science, specifically the journals *American Journal of Political Science*, *American Political Science Review*, and the *Journal of Politics*. Specifically, we will analyze articles that include the term "culture" in keywords, titles, and abstracts in the Web of Science database. (searched: "cultur*").

The corpus has already be stored as a data frame in the data folder for the homework. Several fields of interest include: AU = authors; CR = cited references (e.g., the references cited in the article); DE = author provided keywords; TI = title; AB = abstract; PY= publication year; SO = journal title); WC = Discipline (e.g., Political Science or Sociology); text = a combined column of the title, abstract, and keywords and lower case. When relevant, like authors, the field is separated by semi-colons.

1. **Introductory Inspection:** Provide an overview of the culture in social science data set.
   a. How many articles are in the data set?
   b. How many unique authors are in the data set?
   c. Plot how the corpus changed over time in terms of volume?
   d. Bonus: How do political science and sociology differ in terms of shifts over time?

2. **Culture in the Social Sciences**: Use the strategy of your choice to build a two-mode network or one-mode projection of a two-mode network for analyzing the use of culture in the social sciences. For example, you may build a coauthorship network, a bibliometric coupling network (e.g, how the articles connect via overlapping citations), a text network, or keyword network. Without consideration of discipline, what does this network tell us about culture in the social sciences.
   a. How many nodes and edges are in the network?
   b. Are there numerous components?
   c. What is the centralization?
   d. Average density?
   e. Interpret the graph has a whole.

3. **Decomposing the Network:**) Examining the largest component(s).
   a. Is the component dominated by a single discipline?
   b. Does the network have a community structure and does this structure vary by discipline?