Josiah Thulin
Winter 2020
SPA-DRP Project Write-Up
Networks

My project for Winter 2020 SPA-DRP is about networks. Sarah Teichman, my graduate student mentor, and I worked through chapters in two books about networks: *Statistical Analysis of Network Data* by Eric D. Kolaczyk and *Statistical Analysis of Network Data with R* by Eric D. Kolaczyk and Gábor Csárdi. Over the course of the program, I learned about graph theory that underlies the study of networks, various measures and statistics used to summarize or analyze different characteristics of networks, ways to visualize networks to best communicate information about the system being represented, and methods for generating random networks. We also discussed ways in which the use and analysis of networks can be applied to several real-world applications in a variety of disciplines, including sociology, computer science, medicine, and economics.

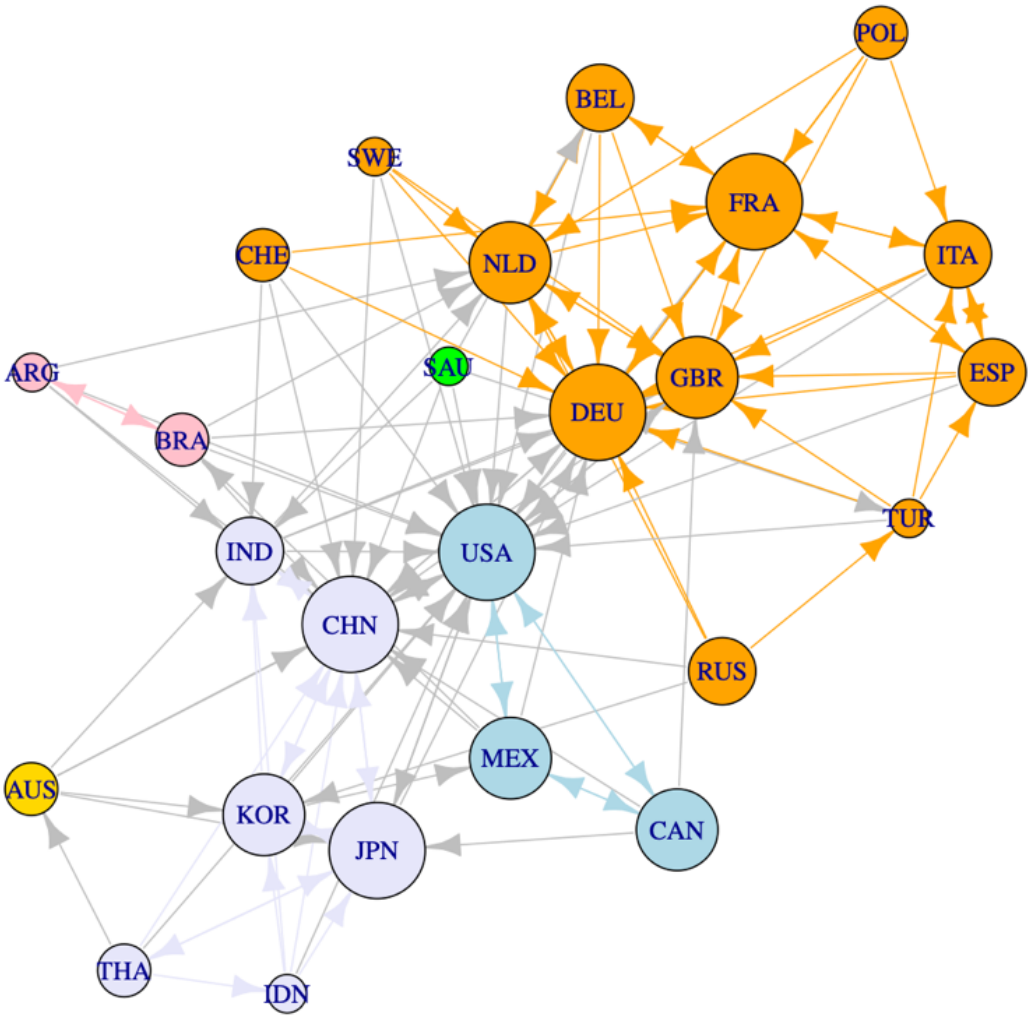
I am particularly interested in the ways that the use of networks can be applied to the field of economics. This motivated my project, in which I chose to investigate how international trade relationships could be modeled and analyzed using networks. Using information from the World Bank[1] and trade data from the UN Comtrade Database[2], I gathered trade data for the top 25 countries by nominal GDP in 2018 to create the network. To simplify my analysis, I decided to focus on the export data for each of the 25 countries, specifically looking at the five largest trading partners for each country in terms of exports. The resulting network had 25 nodes (one representing each country) and 125 edges (with five edges originating from each node representing the top five trade partners of each country with regard to that country's exports). Something I noticed within this network was that the nodes representing the United States, China, and Germany all have a large number of edges directed towards them; the in-degree of these three nodes (22, 17, and 16, respectively) seemed to be much greater than that of other nodes in the network. To see if this characteristic of the export network is significant among networks with the same number of vertices and edges, I conducted an informal hypothesis test for the parameter d_{max}^{in} (maximum in-degree for directed graphs with 25 nodes and 125 edges). For the test, I used R to randomly generate 100 directed networks with the same number of nodes (25) and edges (125) as the export network using the G(n, M) Erdős-Réyni model with n = 25

and M = 125. Once the random networks were simulated, I observed the degree distributions of each generated network to see how many (if any) had a value of d_{max}^{in} as large or larger than that of Germany, China, or the United States in the export network (indicating that such high in-degree values might appear simply due to the size of the network). The largest values of d_{max}^{in} that appeared were 11 (appearing six times) and 12 (appearing once). The mean value of d_{max}^{in} among the 100 simulations was 9.11. While not a strict hypothesis test, I felt that the results of the simulation gave fairly strong evidence that the in-degree values for Germany (16), China (17), and the United States (22) were significantly large. One thing to note while considering these results is the degree to which the G(25, 125) Erdős-Réyni model captures the structure of the export network I was trying to model. While this model may not most accurately model the export network's structure, it was still able to provide useful information for the purposes of my informal hypothesis test. Further analysis of the export network would benefit from the use of a different model that takes other characteristics of the export network into account (such as its highly skewed degree distribution).

In a broader context, this evidence indicates that some countries play a significant role in global trade. The high in-degrees of the United States, China, and Germany show that they are disproportionately represented among the top five trading partners for the world's largest economies. This conclusion may not seem particularly insightful on its own, but the greater takeaway for me was the example it provided of a way in which network analysis could be applied to economic ideas like global trade relationships. Another idea I could explore with further analysis of this export network is the role that geography plays in determining the top trade partners for each country (looking at the balance between geographic neighbors and the world's economic giants within a country's top five export partners). Overall, this project has been an incredible chance for me to learn about networks and discover ways in which statistics can be applied to the real world.

Josiah Thulin
Winter 2020
SPA-DRP Project Write-Up
Networks

Network Graph of Export Data



Key

ARG	Argentina	**DEU**	Germany	**JPN**	Japan
AUS	Australia	**ESP**	Spain	**KOR**	South Korea
BEL	Belgium	**FRA**	France	**MEX**	Mexico
BRA	Brazil	**GBR**	United Kingdom	**NLD**	Netherlands
CAN	Canada	**IDN**	Indonesia	**POL**	Poland
CHE	Switzerland	**IND**	India	**RUS**	Russia
CHN	China	**ITA**	Italy	**SAU**	Saudi Arabia

Josiah Thulin
Winter 2020
SPA-DRP Project Write-Up
Networks

SWE	Sweden	**TUR**	Turkey
THA	Thailand	**USA**	United States

Data Sources

1. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?most_recent_value_desc=true&year_high_desc=true
2. https://comtrade.un.org/data/