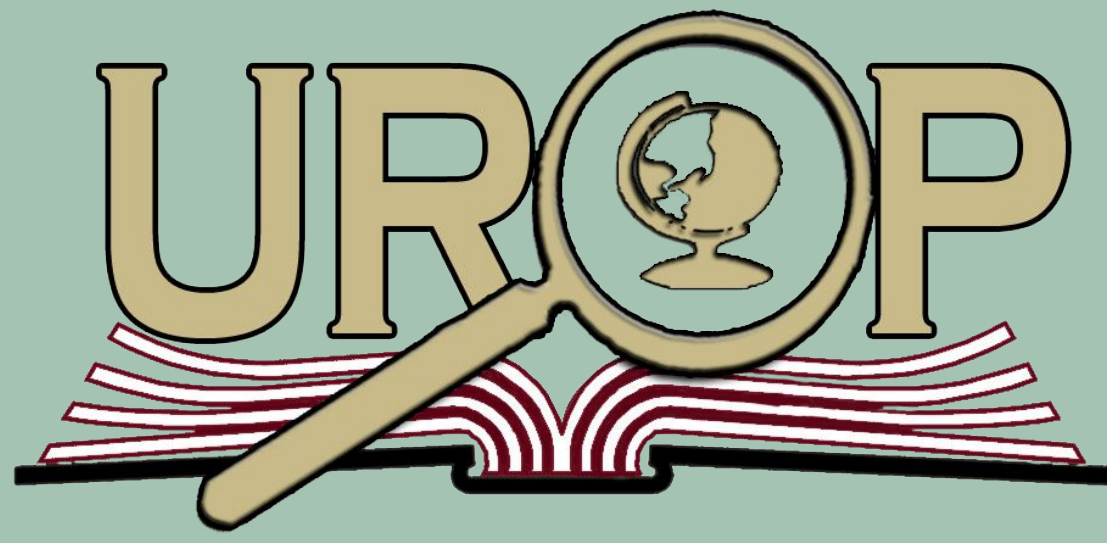# Data Management in Respect to Social Network Analysis

Kelli Cardenas, Dr. Shannon Montgomery and Dr. Joseph Grzywacz

Department of Human Development and Family Sciences

Introduction

Purpose: This study aims to document the steps required to prepare, clean, manage, and analyze social network data

Hypothesis:
When data is managed and analyzed properly it can be used for social network analysis to better understand how social networks can be used for social network interventions

Background

Social Network Analysis
- A research method used to analyze relationships within a group of individuals, primarily focusing on patterns of relationships (O'Malley 2008).
- Important for understanding the social structure and provides a visualization of the flow of relationships.
- Has a wide variety of uses, this can be applied to anything from the medical field to the business industry, more specifically can be used for health behavior interventions when done properly (Hurtado 2018).

Data Management
- Data must be applied correctly to be successful when analyzing social networks.
- One must be able to manage, analyze, and visualize the data.
- When working with social network data the data is clustered, therefore making it violate the typical independence often seen in analysis.

Methods

- **Participants**: 3 mock farmworker crews (each made of 15 individuals)
- **Process**: the data was put into excel manually, then anonymized, and finally transformed from a wide data set to a long data format.
  - The edgelist was uploaded into the program R to be able to perform social network analysis.
  - Packages readxl, tidyverse, network, dplyr, igraph, rlang, and pillar were downloaded to implement in R.
  - R Script was then used to code how to plot the network and make it more visually appealing.

Results

- The data collected shows clustered networks, but when we manipulate the social network itself, we can get a clearer visual of what these networks actually tell us.
- The individual and bridge interventions can be used due to the high centrality of certain nodes and the bridge node shown within the social network (Valente 2012).
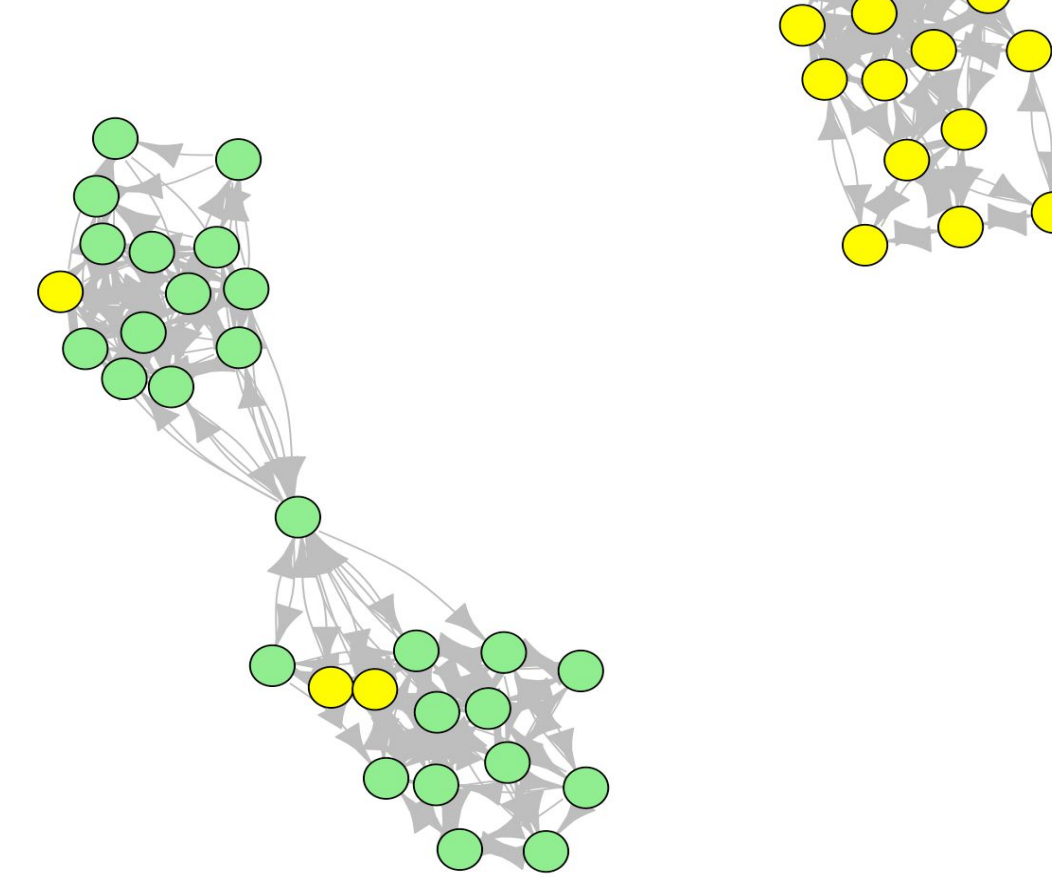
Social Network A
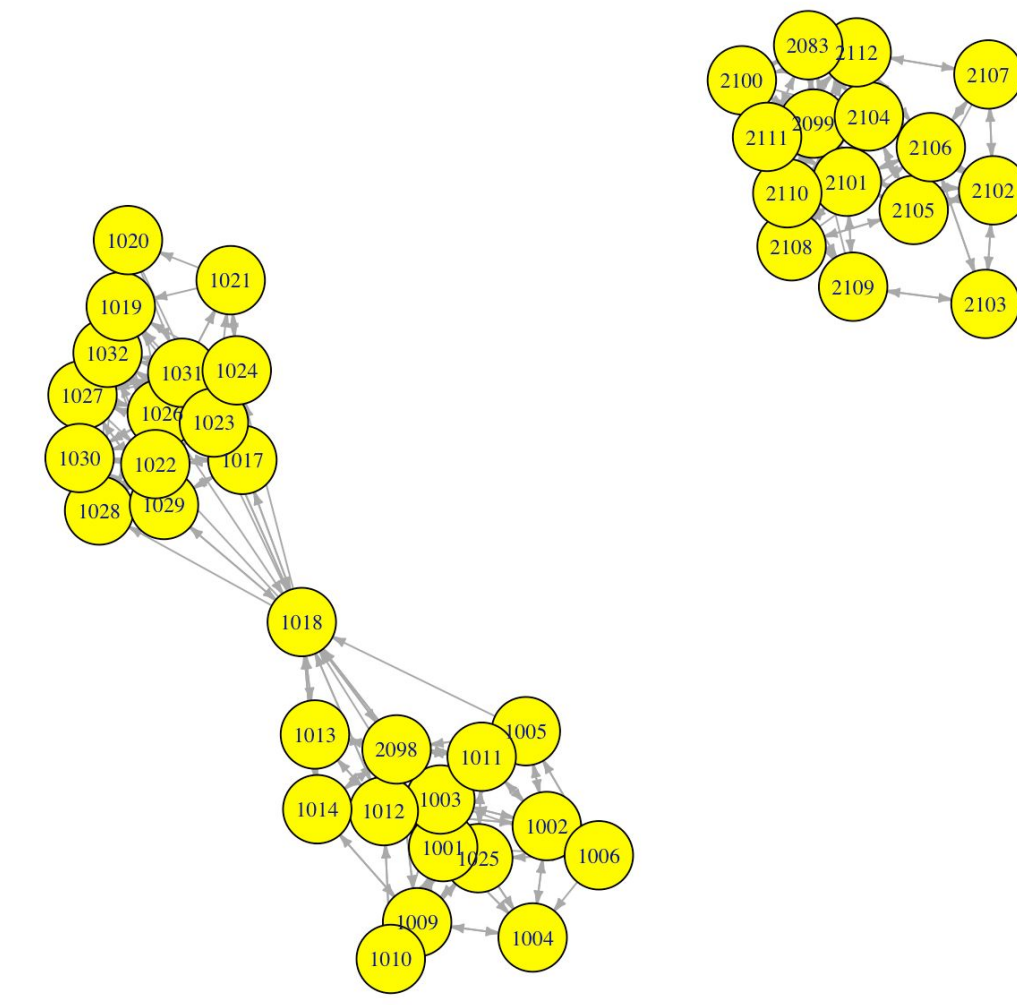
Social network A resembles what it would look like if a network was plotted with no manipulation. As shown, the network is clustered and is not easy to interpret.
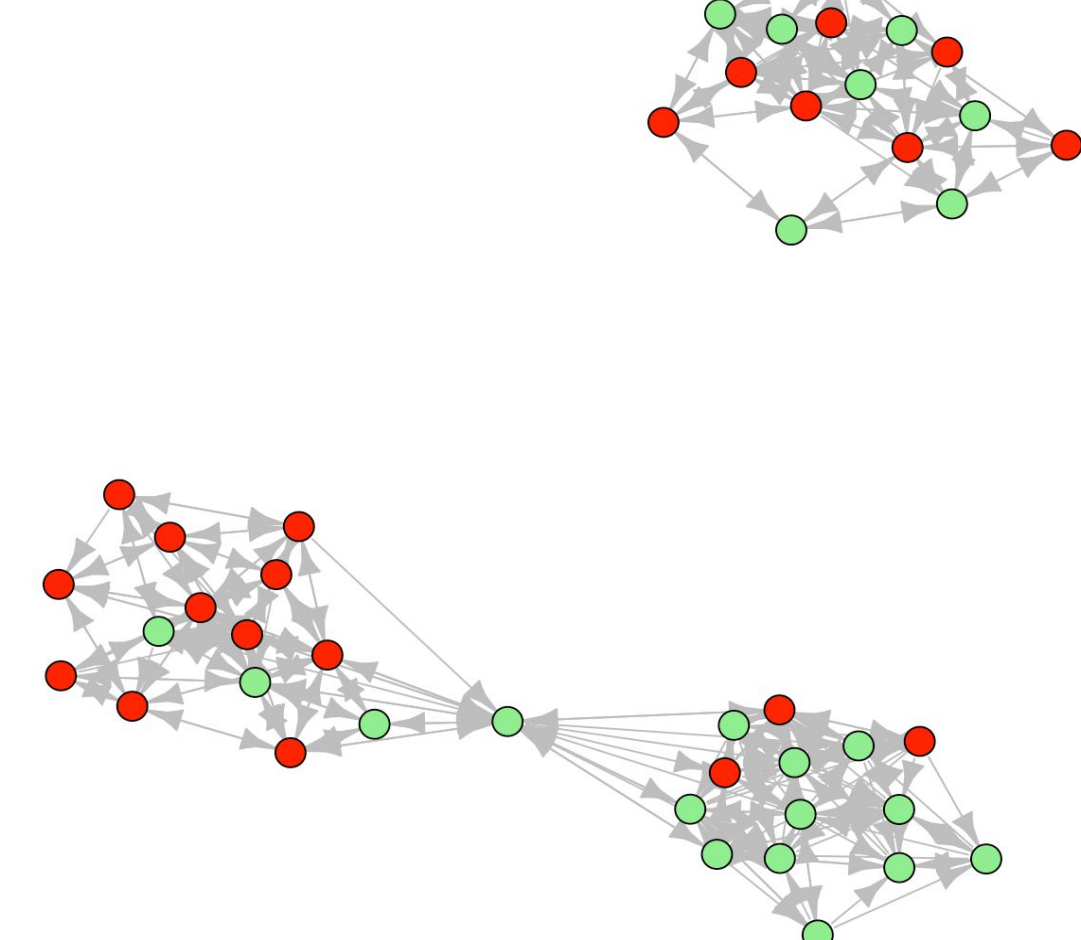
Social Network B
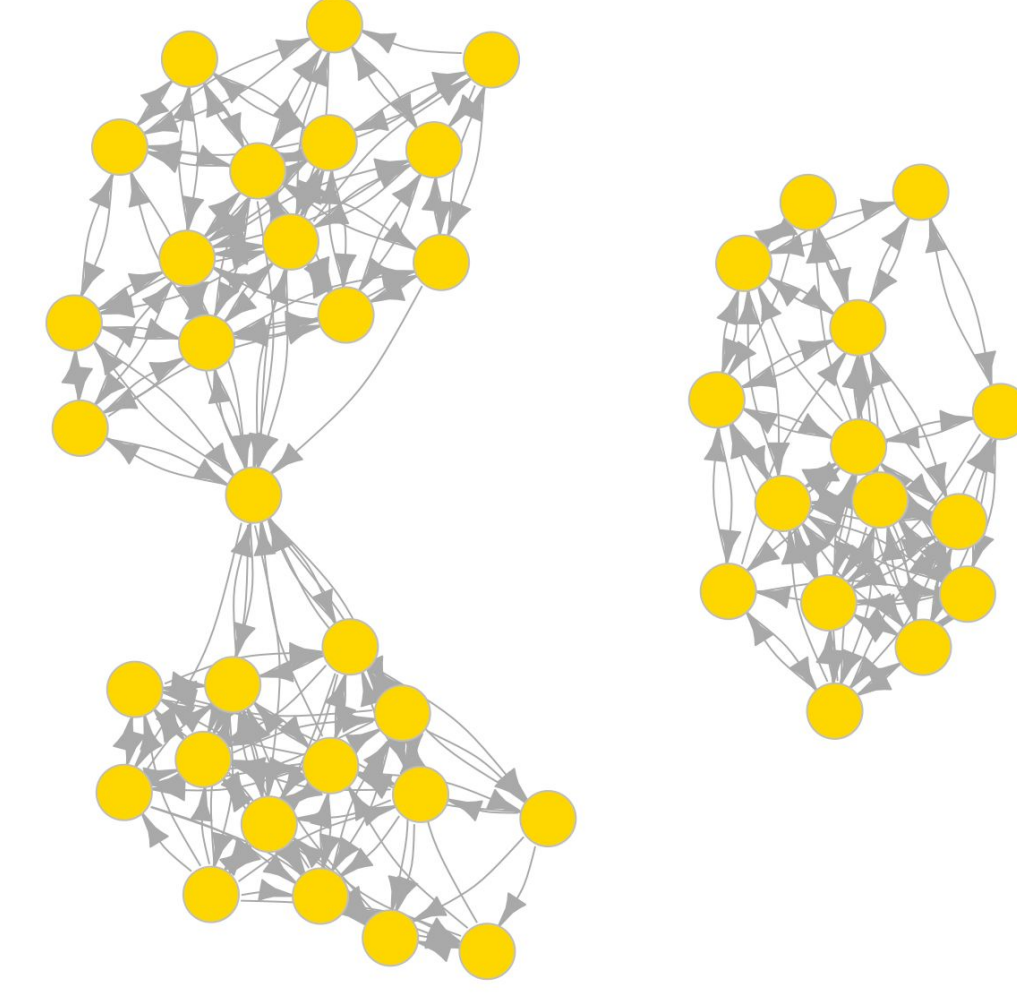
Social network B is a lot easier to interpret than social network A. The nodes have been spread out and the ID numbers are no longer in the plot.

Social Network C

Social network C plots the gender of the participants. The yellow indicates a woman and the green indicates a man. This is an example of what attributes can be shown for further analysis.

Social Network D

Social network D plots whether or not a participant has a high grade in understanding heat related risks. A grade of C or above is indicated with a green node and any grade lower is indicated with red.

Discussion

- The results are important because they show the network's different characteristics, thus allowing us to determine the right type of social network intervention to be used.
- When the right kind of intervention is used, we can then implement different behaviors within a network using the intervention chosen (Valente 2012).
  - For example, within the field of health we can implement behaviors that will encourage better hygiene to help protect the health of farm workers.

Conclusions

- These results help support the hypothesis that when data is managed and analyzed properly it can be used for social network analysis to better understand how they can be used for health behavior interventions
- Social Networks can also be utilized to show the different attributes within the data. This can aid in the selection of which interventions should be implemented (O'Malley 2008).
- The results do not show the issues that may be encountered with data management and analysis.
  - For example, there may be issues with unidentifiable subjects which must be removed or the steps that must be taken to adjust the clustering of networks.

References

Hurtado, David A., et al. "Identifying Safety Peer Leaders with Social Network Analysis." *Occupational Health Science*, vol. 2, no. 4, 2018, pp. 437–450., https://doi.org/10.1007/s41542-018-0026-4.

O'Malley, A. James, and Peter V. Marsden. "The Analysis of Social Networks." *Health Services and Outcomes Research Methodology*, vol. 8, no. 4, 1 Dec. 2008, pp. 222–269., https://doi.org/10.1007/s10742-008-0041-z.

Valente, Thomas W. "Network Interventions." *Science*, vol. 337, no. 6090, 2012, pp. 49–53., https://doi.org/10.1126/science.1217330.