

Reason of Intent – ERGM Project

My analysis of the social networks in the show BoJack Horseman has allowed me to learn and advance my data analysis skills in several ways. Firstly, for this project, I had to gather my data based on my opinions of the social networks in BoJack Horseman. I learned how to organize data appropriately for social network analysis, such as a collecting list of nodes (characters) and edges (relationships). I also had to verbalize how I collected the data and why I chose the relationships I did. I collected the data by finding a list of characters online and choosing only those considered "main" or "reoccurring" and included them in my list through my own opinion about the relationships from the show. With this, I learned the importance of communicating how and why data is collected in a certain way, which makes it simpler for someone else to understand and reproduce the work if wanted.

The project also taught me how to use Ideanet, a newer package in R created by Tom Wolff, to analyze the social network data I had collected. With Ideanet, I was able to load my data into R in the correct form (edges and nodes) in an easier, more efficient way. This is especially useful if I ever intend to go further into social network analysis as I have a deep understanding of the form the data must be in as well as how to get that data into R for analysis.

Once my data was in R, I was able to learn a variety of techniques for visualization and analysis of social networks. For visualizations, social networks are plotted in specific graphs that allow you to see the nodes, edges, and connectedness of the characters in the network. I learned how to use the Igraph package in R to plot my networks from the show, allowing me to begin to draw conclusions about the networks in BoJack Horseman. Furthermore, I was able to use tools in R to create visualizations of the network categorized by differing traits and with nodes sized by character prominence. These visualizations were combined with statistical analysis to provide further evidence of any patterns I found in the network. For example, I learned how to analyze the prominence of characters in the network by looking at the centrality of the network. Specifically, I analyzed degree, closeness, and betweenness centrality. This analysis and the visualizations allowed me to conclude who are the most prevalent actors on the show.

I was further able to analyze the networks in the show with the help of ERGMs or exponentiated random graph models. I used ERGMs to understand the impact of sociality and

selective mixing in my networks. Again, I used R to create and analyze my network in this way and then created a table to look at the results for my networks.

After the analysis, I used RMarkdown as a way of communicating my findings. I was initially interested in studying the social networks in BoJack Horseman because the main character has a mental illness, and I was interested in seeing if this impacted the relationships on the show. The show also takes place in Hollywood and is centered around the main character, BoJack, a washed-up actor. I was interested in understanding how occupation may impact the social networks of the show, as famous people tend to only associate with other famous people. With the use of visualizations and analysis that I described earlier, I was able to understand a lot more about the show. I developed a clearer understanding of the friendship and economic networks in the show and which characters were essential to those networks – which often were the main characters in the show. I also found that, when it comes to friendship, characters that are in the same occupation are much more likely to be friends, which is what I suspected. All in all, these are just some interesting insights I was able to gain from analyzing the social networks in the show.

However, the biggest takeaway from this project was that I learned the necessary skills needed to begin to analyze social networks. I have a deeper understanding of how data needs to be organized for social network analysis, how to use R for visualizations and statistical analysis, and how to interpret my findings. With this knowledge, I feel prepared to analyze other networks in the future!