

Student Number: **1007045118**

Assignment Title: Social Network Analysis

Course Code: RSM8431

Course Title: Colloquium

Instructor Name: Kanchana Padmanabhan

In submitting this work for grading, I confirm:

- The work is original, and due credit is given to others where appropriate
- Acceptance and acknowledgement that assignments found to be plagiarized in any way will be subject to sanctions under the University's Code of Behaviour on Academic Matters.

Assignments are to be submitted using student ID numbers only; do not include your name.

Assignments that include names OR that do not have the box below checked **will not be graded.**

Please check the box and record your Student Number below to indicate that you have read and abide by the statements above.

☐ **1007045118**

Introduction

The aim of this project is to utilize a Game of Thrones network dataset and apply different social network analysis techniques to derive insights on the Game of Thrones tv series. Throughout our analysis, we hope to understand the complex system of interactions that occurred during the show and understand who truly mattered in the network. Additionally, we would like to determine the different factions or communities present in the show and determine who has the most important interactions in the show. The scope of this report will be on season 1 of the series and the whole show. We choose to focus on season 1 only to grasp how much the show has changed from its beginning compared to what it became.

Game of Thrones was an American fantasy drama airing on HBO based on the popular series A Song of Ice and Fire by George R. R. Martin. It follows the story of several factions and their ultimate goal to sit on the iron throne of Westeros, hence the title. The show lasted for 7 seasons and provided one of the grandest stories told in the television medium. The very nature of the TV show lends itself well for network analysis as there are several branching plot points simultaneously running throughout the show, with many characters never interacting with each other until the final season.

Data Description

The datasets provided are character nodes, character interactions and the weights for each interaction in Game of Thrones. The weights were additionally min-max scaled for comparability. Data is provided for each season of Game of Thrones and is analyzed separately and together.

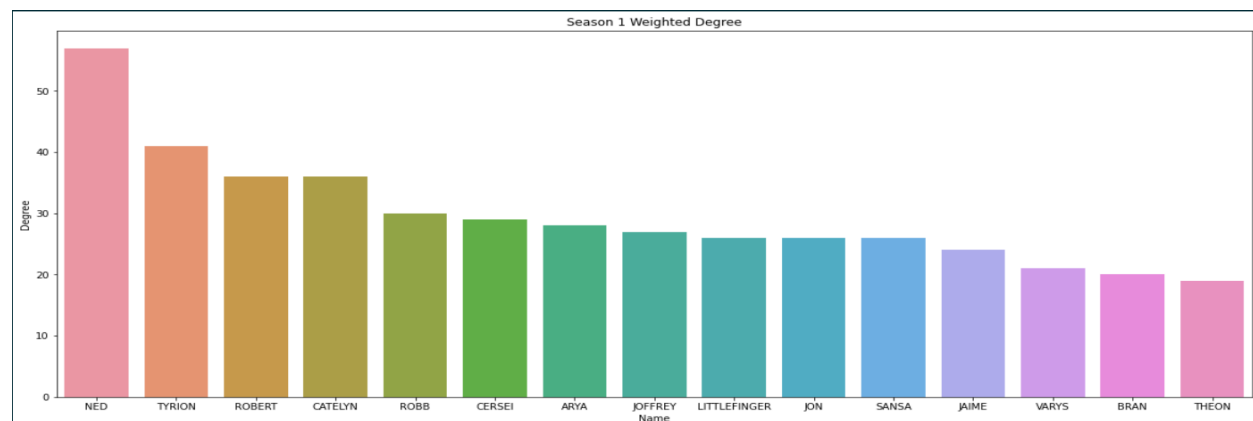
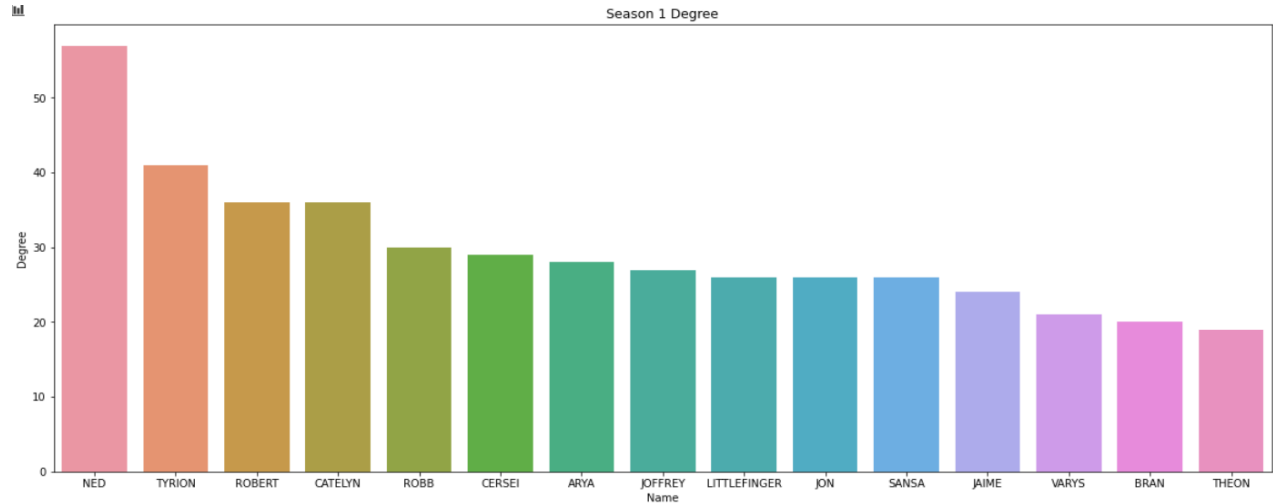
Season 1

Network Description

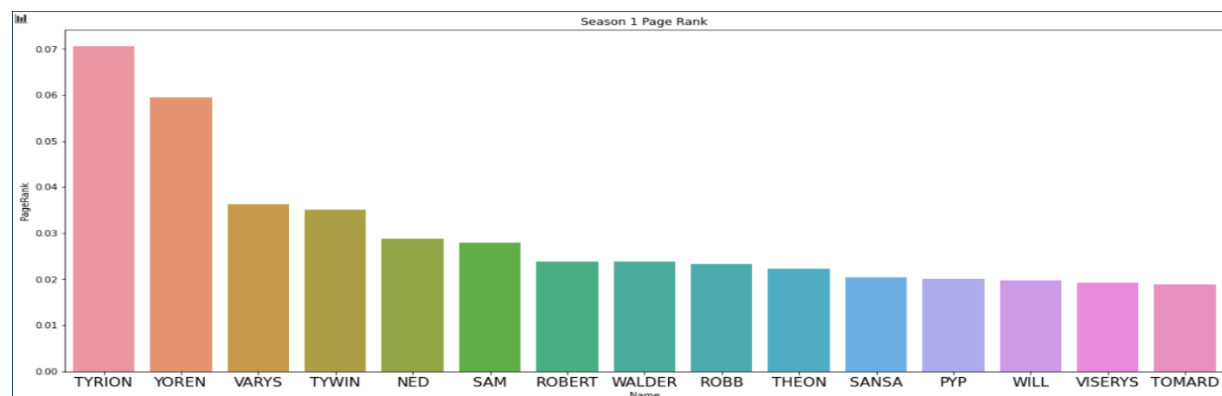
The first season contains 126 characters or nodes and 549 edges or interactions.

Node Importance

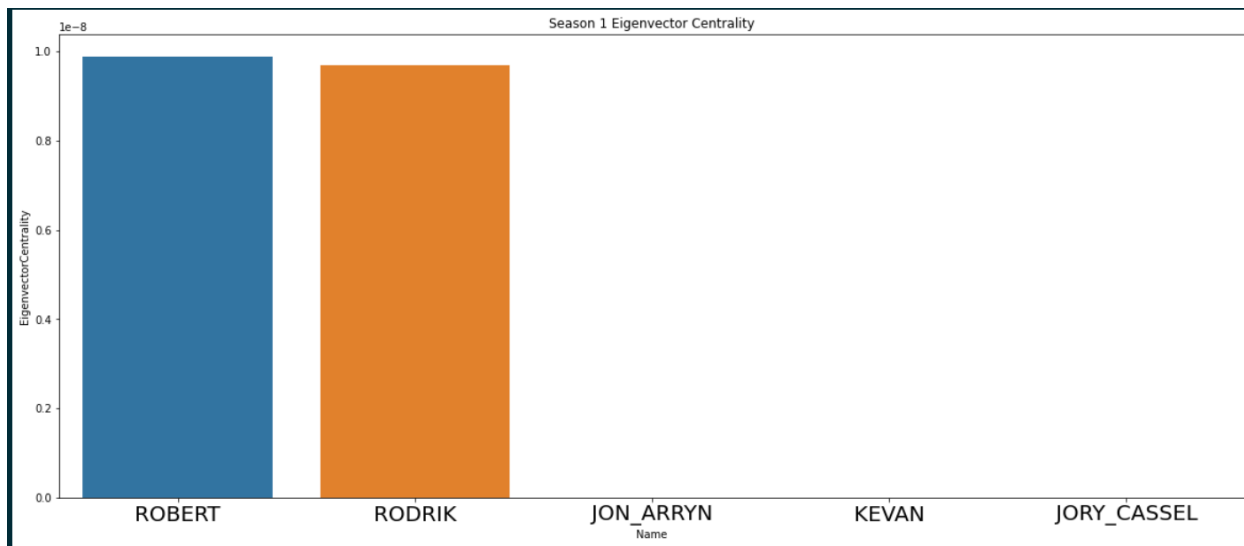
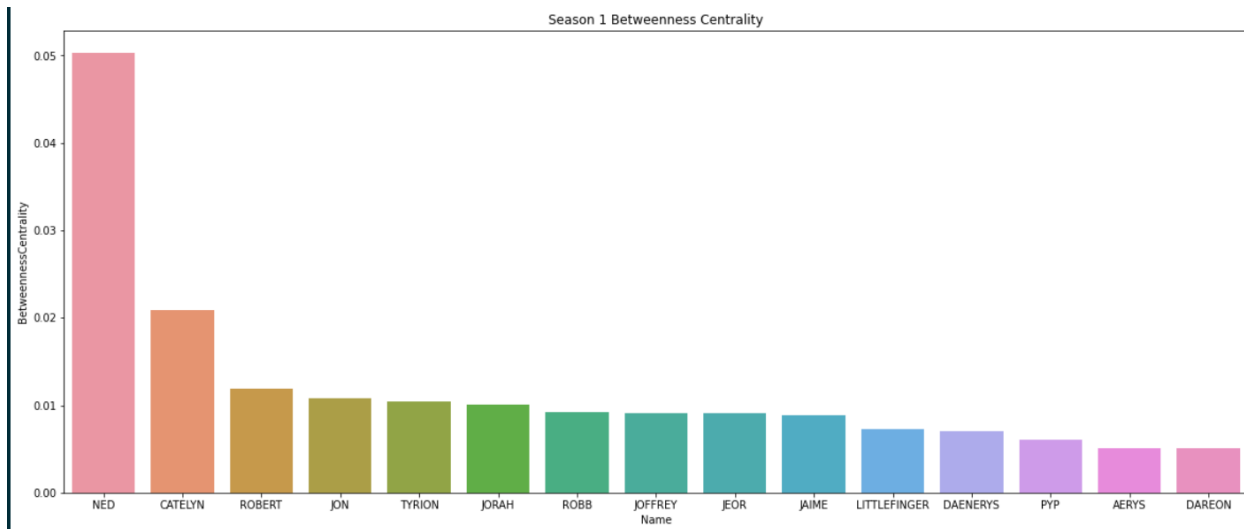
To start of our analysis, we begin by measuring the node importance or character importance of every character in Game of Thrones. The measures being utilized are degree centrality, weighted degree centrality, betweenness centrality, eigenvector centrality and page rank. The top 15 characters for these rankings are shown in the bar charts below.



Degree and Weighted Degree Centrality measures the number of neighbors of a node in an undirected graph and provides us with an idea of who is most influential in the show. As more people are connected to you the more influence you have. In this case, for season 1, the most influential characters according to degree centrality is Ned, Tyrion, and Robert.



Page Rank is another measure of node importance that works by counting the number and quality of links to a node to determine how important a node is. It can be used to truly determine a character's influence. In this case, Tyrion, Yoren, and Varys seem to have the highest page rank indicating that they were crucial characters for season 1. These are the characters that are most likely to get things done in the show due to the quality and quantity of their influence. This makes sense as Varys was the Master of Whisperers on the show, Tyrion is from the Lannister family, and Yoren was a recruiter at the Night's Watch.



Finally, in terms of Betweenness and Eigenvector Centrality, the most important characters are Ned, Catelyn, and Robert for Betweenness with Robert and Rodrik being the most important for Eigenvector Centrality. A character with a higher betweenness will have greater influence on the information flow in a network as

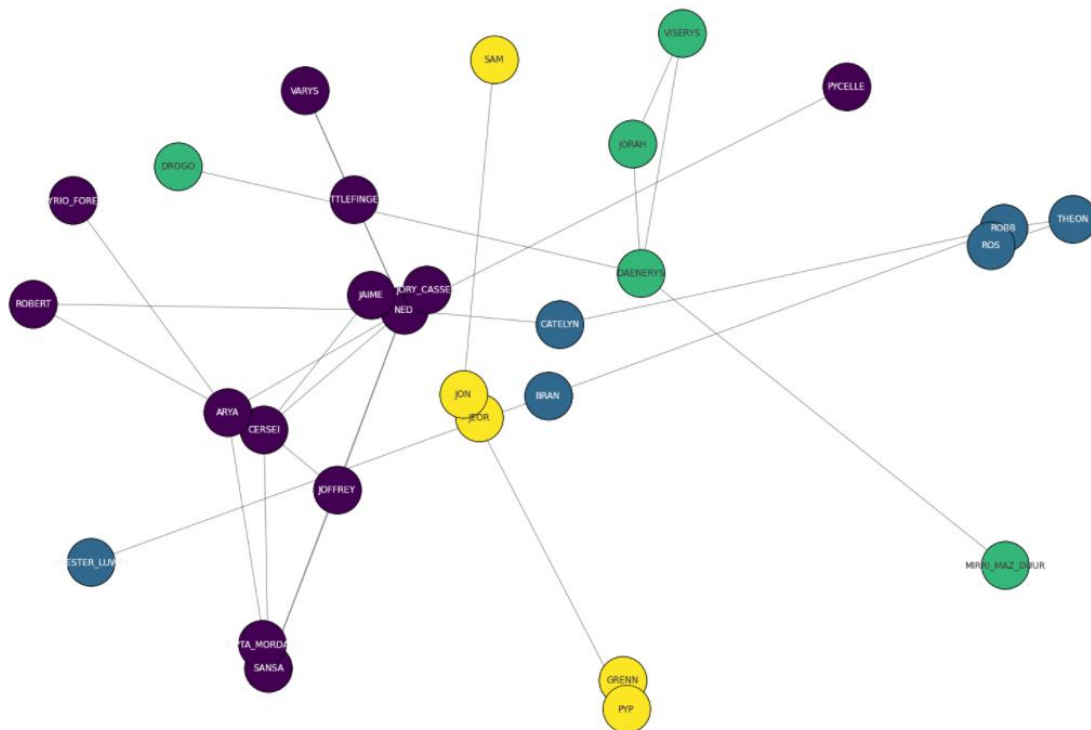
they exist between the most individuals in the set of nodes. In terms of eigenvector centrality, the higher the score the more important individuals a character knows.

Community Detection

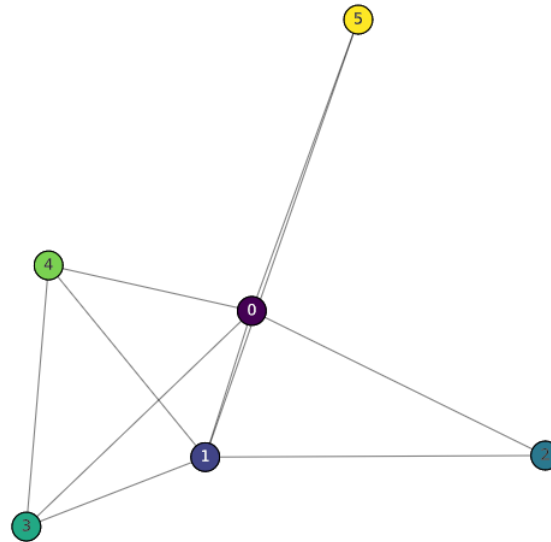
In terms of community detection, we first look at the first season of Game of Thrones to figure out how the narrative began. Only the top 40-character interactions in terms of weight were utilized to create readable graphs and focus on insights. From the network cluster graph there seems to be 4 main clusters or communities that matter for the show. This makes sense as the first season focused on the events that occurred in Westeros with Cersei, the Wall with Jon Snow, the Stark revolution with Robb, and finally the revolution of Daenerys. The network cluster graph is seen below.

III

Season1



Note that when all the data is utilized, other communities appear to be created as well, however they are not as highly weighted or important to the story. This can be demonstrated in the network community graph below.



As you can see the community graph above there are 6 distinct communities, however two of them appear to be farther away and less interconnected than the other communities. This falls in line with the show having four main storylines with some additional side stories.

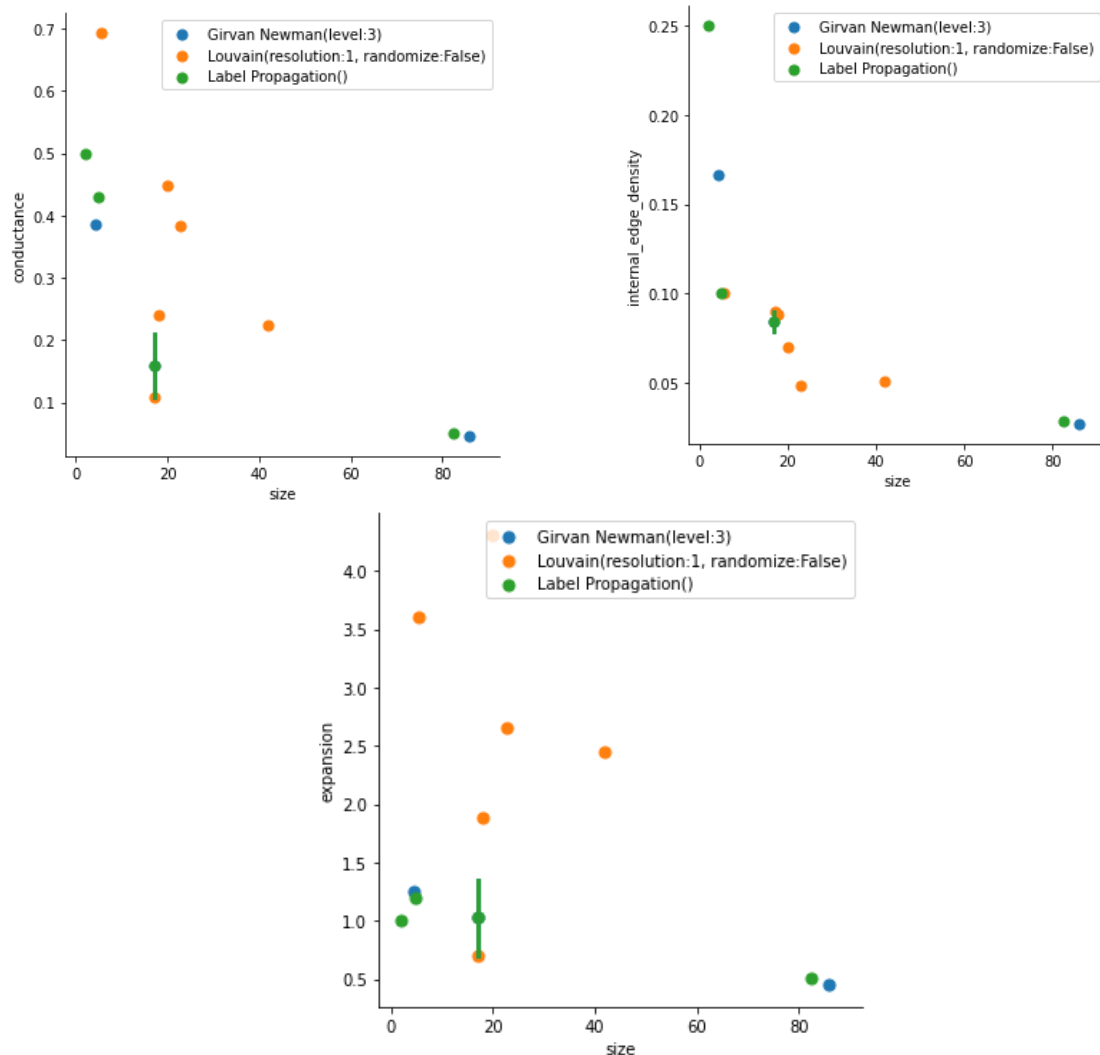
In terms of quality metrics, multiple different algorithms were utilized and summarized in the table below.

Metrics	Girvan Newman	Louvain	Label Propagation
Conductance	0.188 +/- 0.128	0.274 +/- 0.124	0.259 +/- 0.176
Expansion	0.941 +/- 0.373	2.229 +/- 1.04	0.953 +/- 0.311
Internal Density	0.0907 +/- 0.0499	0.0665 +/- 0.0172	0.109 +/- 0.074

Girvan Newman is an algorithm that works by detecting communities by progressively removing edges from the original network. The remaining connected components are the communities.

Louvain is another algorithm like Girvan Newman that detects communities, however it does so by optimizing a modularity metric, which measures the strength of division of a network into modules or communities.

Label Propagation works by detecting communities in a graph using network structure only as a guide. It does not use an objective function or any prior information on the communities. The algorithm propagates labels throughout the network and forms communities.



Conductance measures how well-knit a graph is. Expansion measures the number of edges with nodes outside of the community. Internal Density measures the number of edges within a community. When it comes to finding the optimal algorithm, it should maximize Density, minimize Expansion. The Conductance metric is a ratio of Expansion to Expansion plus Internal Density. As you can see from the table above it was found that Girvan Newman provided the best algorithm for community detection in terms of Conductance.

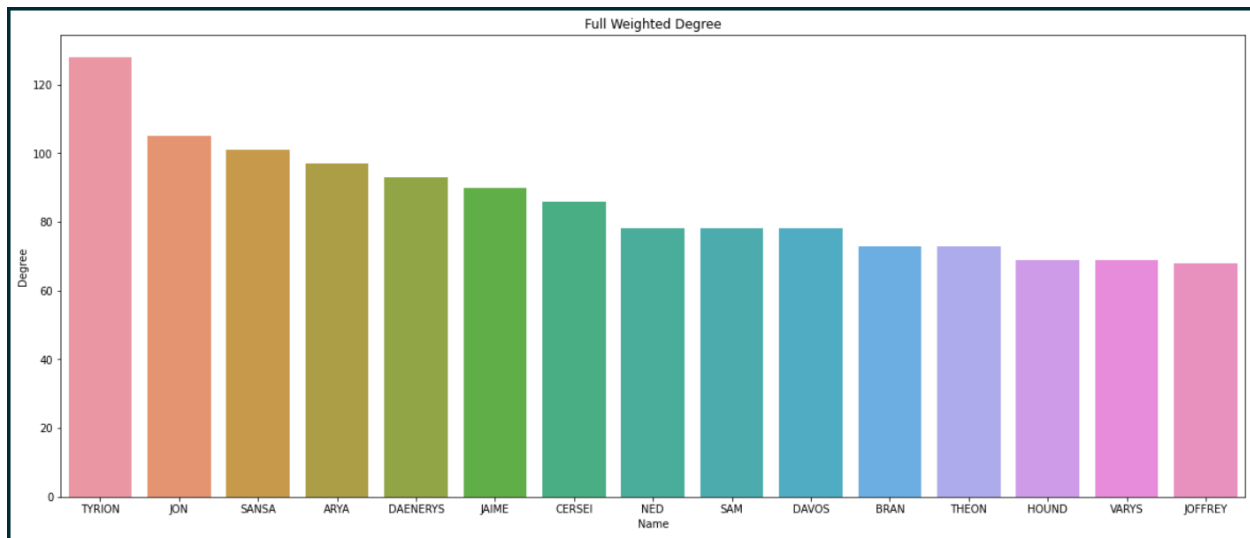
Full Season

Network Description

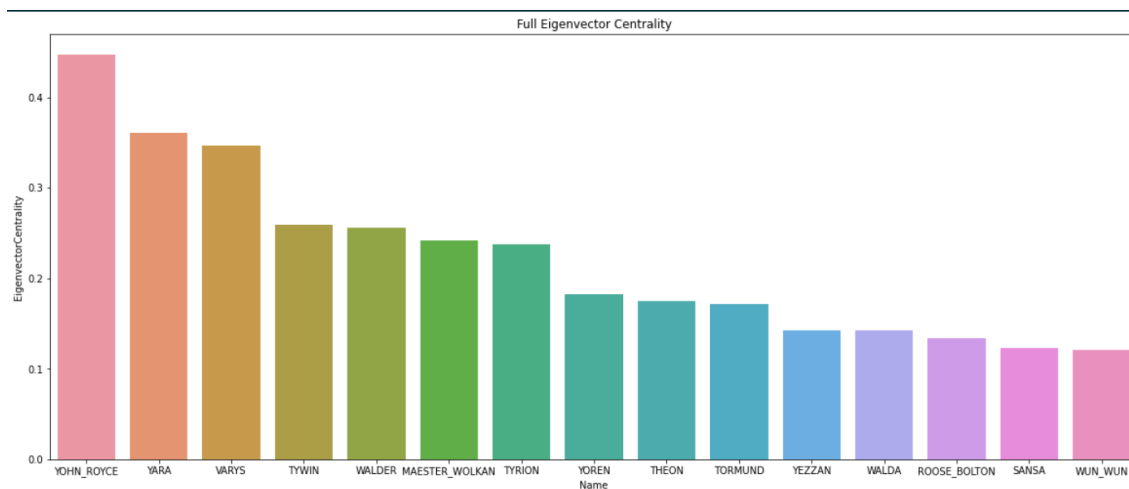
The full show dataset includes 406 nodes or characters and 2639 interactions or edges.

Node Importance

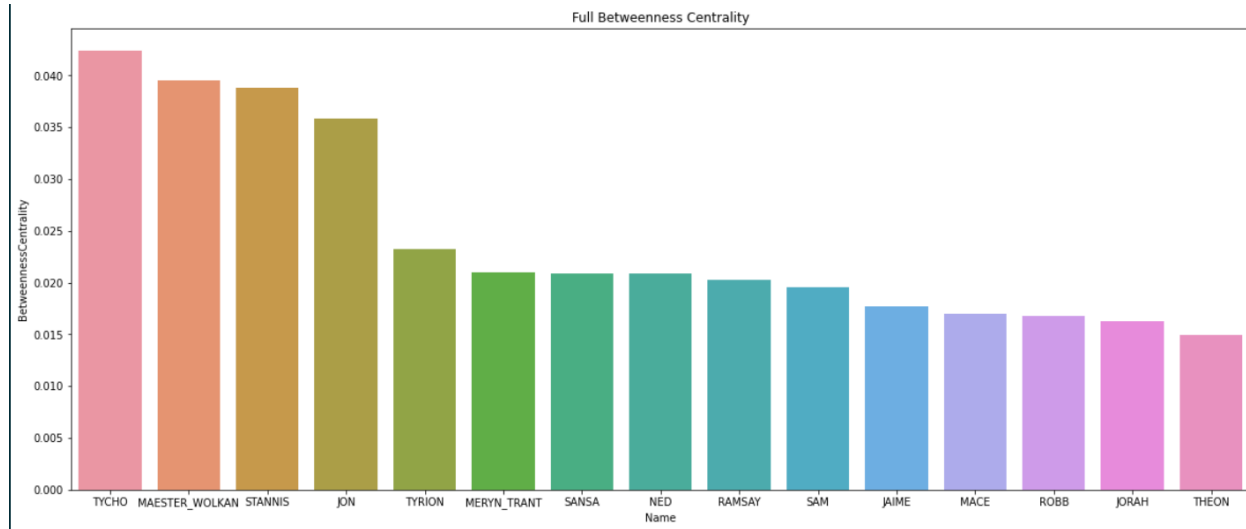
The node importance metrics for season 1 were utilized again to compare against the entire show, however Degree was replaced with Weighted Degree only.



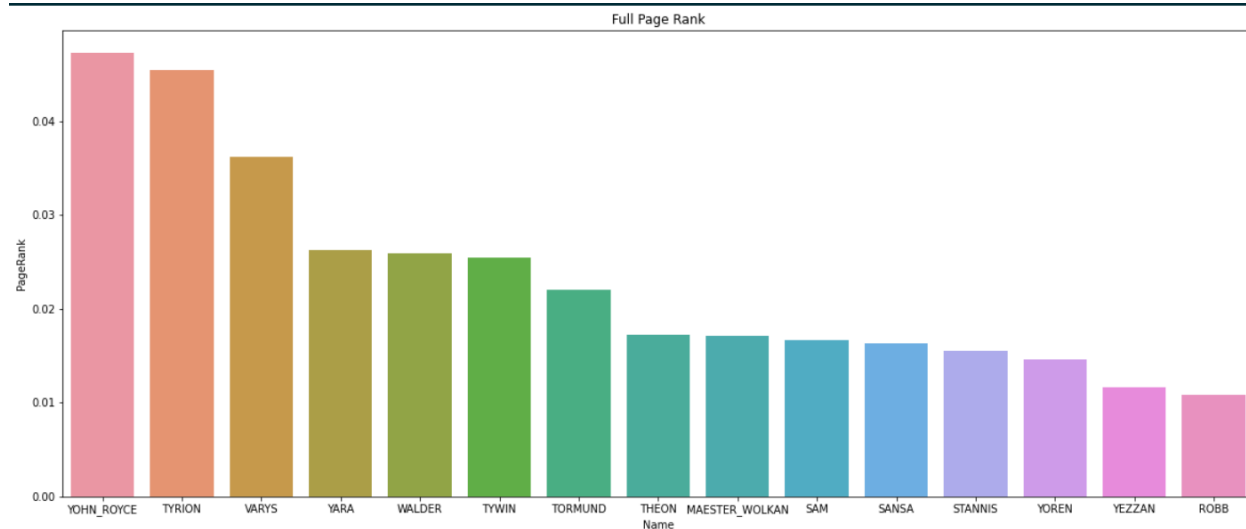
In terms of Weighted Degree, Tyrion Lannister overtakes Ned Stark for the character with the most neighbors. This makes sense as in subsequent seasons after season 1, Tyrion started becoming more prevalent and popular to the audience. Additionally, you can notice characters such as Sansa, Jon, Daenerys, and Cersei take the lead from Ned as they become more central characters in the story. Interestingly, Ned Stark is still one of the most influential characters in the show, even though he only was around for a single season.



For Eigenvector Centrality, surprisingly characters such as Yohn Royce, Yara, and Varys seem to have the highest values. This makes sense for Varys as he is known for his connections, however the other two is unexpected. Tyrion Lannister also appears on this list.



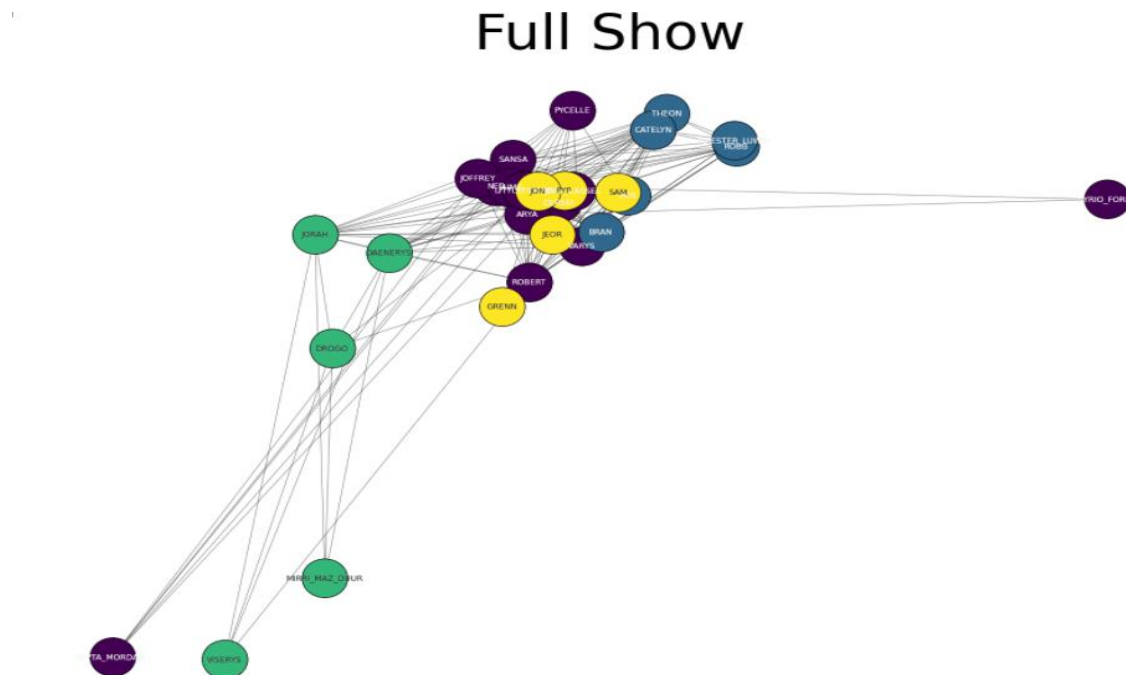
In terms of Betweenness Centrality, unsurprisingly characters like Jon Snow, Tyrion, and Sansa demonstrate the highest betweenness. This makes sense as they are central characters and have the greatest influence in spreading information across the network.



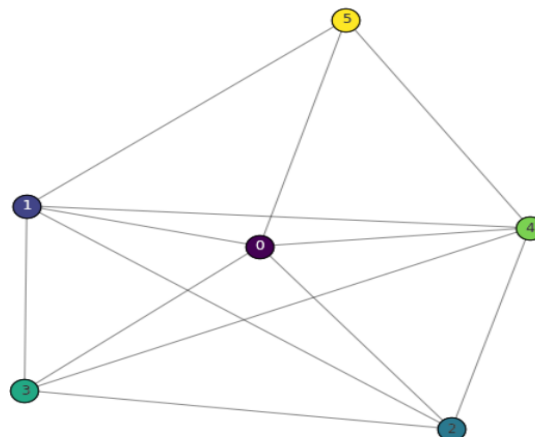
Finally, for Page Rank, the most important characters are Tyrion, Varys, and Yohn Royce. This is aligned with the results of season 1 to a degree and demonstrates that these characters have never lost their influence on the story.

Community Detection

Once again, only the top 40-character interactions in terms of weight were utilized to create readable graphs and focus on insights. From the network cluster graph there still seems to be 4 main clusters or communities that matter for the show. However, in this case the factions seem to be closer together than in season 1. This makes sense as the final enemy of Game of Thrones was the Night King, where every faction had to come together to kill it. This would naturally lead to storylines intersecting over the course of the show.

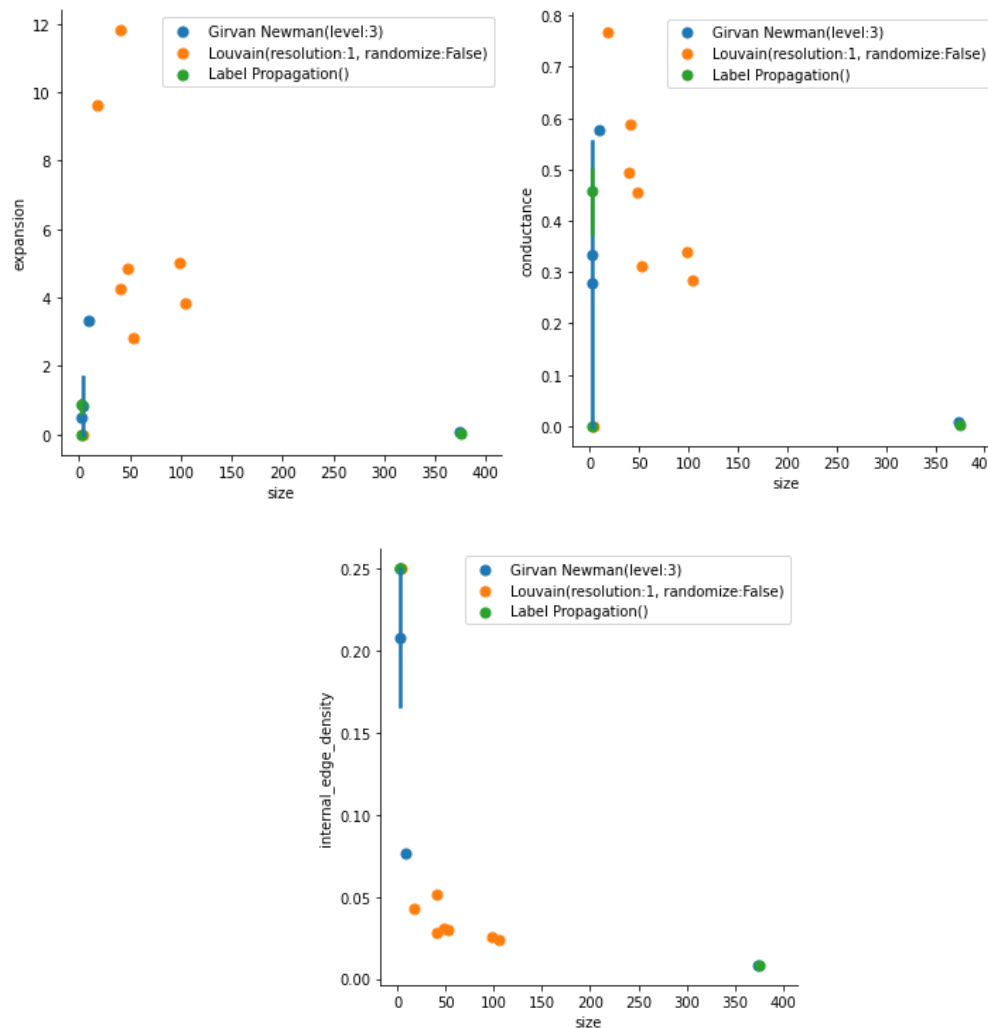


This idea is even better demonstrated in the network community graph below. As you can see, the six communities that previously existed in the show are much more connected now than they were before.



Multiple different algorithms were utilized again and summarized in the table below.

Metrics	Girvan Newman	Louvain	Label Propagation
Conductance	0.294 +/- 0.252	0.405 +/- 0.214	0.306 +/- 0.224
Expansion	1.118 +/- 1.256	6.28 +/- 3.51	0.587 +/- 0.445
Internal Density	0.150 +/- 0.0956	0.0602 +/- 0.0722	0.209 +/- 0.09

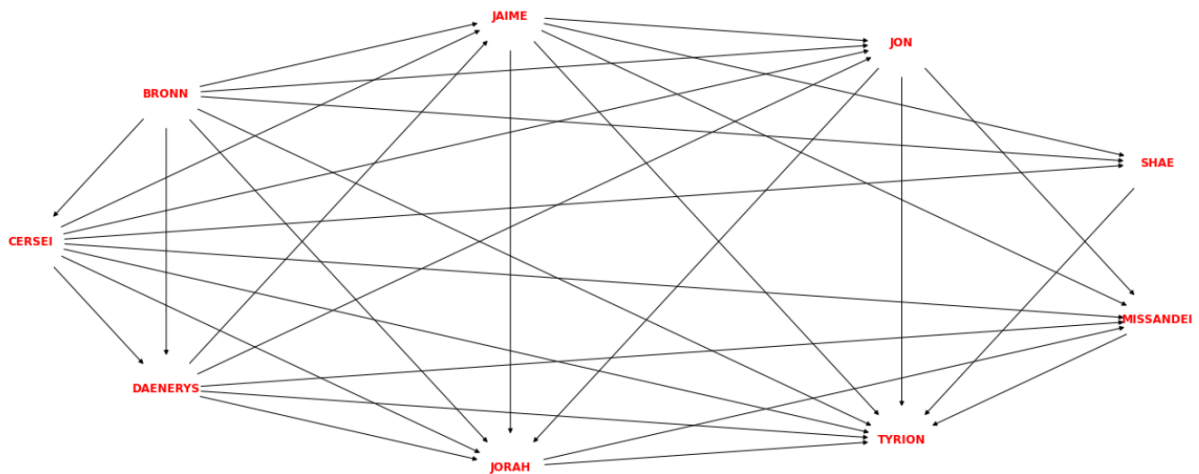


Using the same metrics as season 1, it was once again found that Girvan Newman provided the best algorithm for community detection in terms of Conductance.

Most Influential Character

After analyzing the first season and the entire show, one of if not the most influential character on the show is, Tyrion Lannister. Narratively this makes sense due to the various high positions he has held and the connections he has with the central characters of every faction. The network demonstrates the influence he has on the show and how paramount he is in the information flow in Game of Thrones. This can be seen in the graph below as well.

Tyrion Lannister Top 15 Interactions



The graph above describes the top 15 interactions in terms of weight Tyrion Lannister has had in the show. Just from this list, you can notice how strong his influence is with his connections to the shows central characters (Sansa, Jon, Cersei etc.). Tyrion Lannister's network is further summarized in the table below. Ned Stark was utilized as well as he was the most important character in season 1.

Metrics	Tyrion	Ned
In Degree	118	46
Out Degree	10	32

Shortest Path for interaction between Tyrion and Ned:

Tyrion=> Maester Wolkan => Roose Bolton => Stannis => Tycho => Meryn Trant
=> Ned

Conclusion

From our analysis of season 1 and the whole show, we notice some general trends. Firstly, Ned Stark loses importance significantly after season 1 because he dies during the season 1 finale, however his influence is still relevant seven seasons after his departure. Additionally, characters such as Jon Snow, Cersei, and Sansa are shown to gain much more influence over the course of the show. This is seen by the centrality measurements and the quality and quantity of influence these characters gain throughout the show. At the end, the communities stayed the same, however they grew more interconnected. Finally, the most important character according to this network analysis is Tyrion Lannister.

Through our network analysis, we were able to conclude who the most important character was, how many factions or storylines existed in the show, and how the show evolved over time. Although Game of Thrones had a controversial last season, the seeds that led the various factions to combine were setup since the beginning.