

CNT 5805 Project 1 Questions and Guidance

Please answer all of the questions that are bullets under each numbered paragraph below. Create corresponding numbered sections in your report.

1. There is a network extracted from the famous novel *Les Miserables* written in 1862 by the renowned French author Victor Hugo. Besides this novel, it has been performed as a musical and more recently as a movie. You are going to analyze this work using Gephi. You may download the file here [Gephi Github open source database](#). After downloading it, open it up with a text editor so you become familiar with the data.

Each node is a character and there is an edge between two characters if they appear in the same chapter. *Les Miserables* is written in over 300 short chapters, so two characters that appear in the same chapter are very likely to meet or talk in other chapters as the plot in the novel unfolds. The edges are weighted, and the weight is the number of chapters those characters appear together in.

- Explain your purpose (e.g., inform, persuade, educate, entertain, etc.) for analyzing this network?
2. Now open the file using Gephi.
 - What can you ascertain about the graph you see?
 3. Now run three to five Layout algorithms. Make sure to run the Yifan Hu algorithm as one of your choices.
 - Show a small screenshot of each one and briefly explain what changed and why?
 - At this point, which Layout seems most useful and why?
 4. Now add some emphasis to your diagram by sizing, color coding, and naming the nodes.
 - Show a screenshot of this action.
 - What have you learned from your new diagram?
 5. Now run some statistics about the network such as connectivity, centrality, clustering, etc. You decide which ones are most relevant.
 - Explain what statistical results you found. Please show graphs and/or numbers generated by Gephi in your explanation.
 - Are there any communities or giant components, please explain?
 - Do you see any homophily or density in your network, please explain?
 6. Final contemplation.
 - Can you think of any other relevant network data that would have enhanced or made the study more interesting? Please explain.
 - Summarize what you learned in this project (three paragraphs maximum).
 7. Submitting your project assignment: You must upload your paper into WebCourses in PDF format and your file name should follow this naming convention: *last name-first name-CNT5805-research project1.pdf*