**3.1. Selection of Dataset**

The selection of the dataset was made according to the interest on the subject and the opportunity to implement the knowledge learned in class about networks.

The data selected is divided into two datasets (BKFRAB and BKRFAC). It concerns interactions among students living in a fraternity at a West Virginia college. All subjects had been residents in the fraternity from three months to three years. BKFRAB records the number of times a pair of subjects were seen in conversation by an "unobtrusive" observer (who walked through the public areas of the building every fifteen minutes, 21 hours a day, for five days). BKFRAC contains rankings made by the subjects of how frequently they interacted with other subjects in the observation week. These data can be found on this [website](http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/ucidata.htm#bkfrat).

**3.2. Task Analysis**

A data visualization analysis of the selected datasets can help accomplish several tasks, such as identifying clusters and influential individuals, as well as inferring behaviors and personality traits (such as introversion and extroversion). These tasks can be done by analyzing some metrics, such as degree centrality and edge betweenness, and through clusters and cliques, for example. Also, a community algorithm, such as the Girvan-Newman method, can be used to efficiently identify the clusters within the network. Furthermore, a degree distribution analysis can offer insights into the social behavior and organization of society.

Also, the two datasets can be compared through the same visualization technique. This would allow to infer conclusions on how accurate the individuals recall their interactions, or whether the dataset collection’s method was accurate or not.

In this scenario, visualizing data is a useful tool to analyze patterns and trends on a social network, thus creating knowledge that allows a better understanding of the social dynamics of those monkeys and their behavior.

**3.3. Visualization Design/Sketch**

The main purpose of this project is to create a static and an interactive visualization. The latter must include the following interactivity concepts: brushing and linking, and details-on-demand. This section includes further explanations on the visualizations, as well as general sketches of how they should look. It is important to notice that the sketches are simplistic and do not represent the dataset and all the features the final project will present.

**3.3.1. Static Visualizations**

In order to visualize the general structure of the community, a **Kamada-Kawai** network visualization will be implemented, which may provide insights about groups inside the society, as well as individuals that play important roles. To make these points clearer, only the nodes with the highest degrees should have their labels displayed, as well with different colors. Also, the edges’ thickness should translate the frequency of interaction between dolphins. Figure 1 below represents a sketch of the Kamada-Kawai visualization.

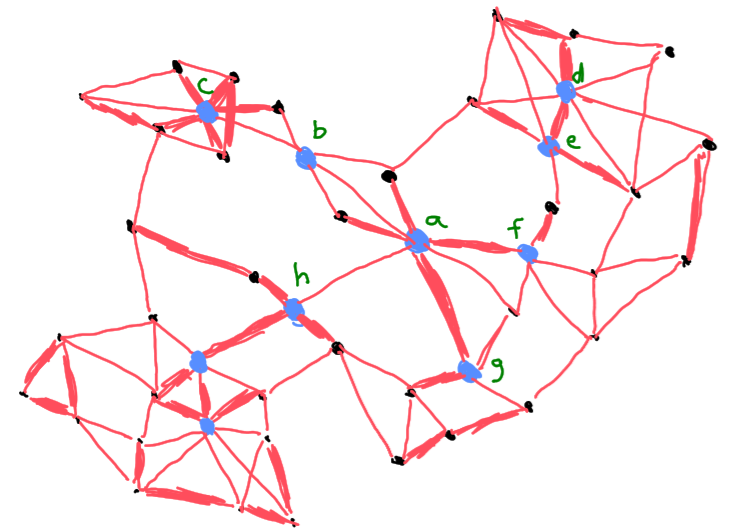


Figure 1: sketch of the Kamada-Kawai graph

As well as the Kamada-Kawai visualization, the group proposes a **chord diagram**. This graph should give more insights on clusters inside the community, as well as representing the strength of interaction between individuals through the ribbon’s thickness. Figure 2 below represents a sketch of the chord diagram visualization.

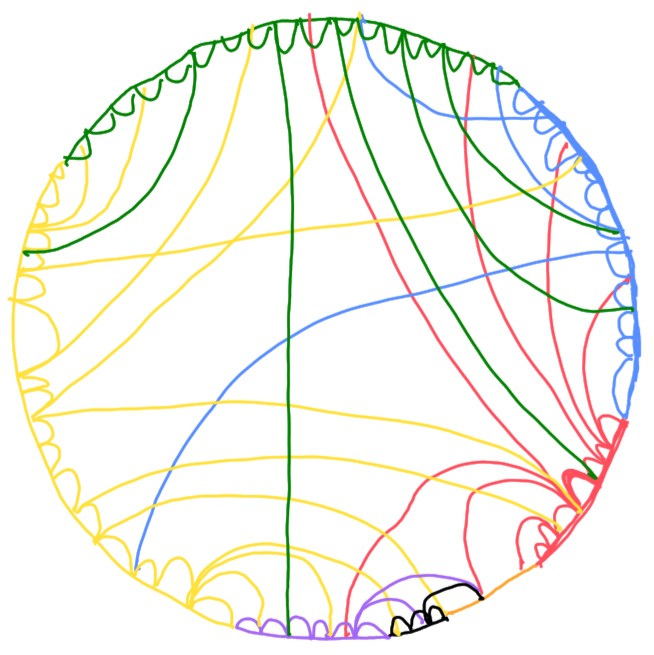


Figure 2: sketch of the chord diagram

**3.3.2. Interactive Visualizations**

An interactive visualization can be made relating a heatmap and a force-directed (spring) graph. The heatmap shows the frequency of associations between pairs of dolphins. It can be color-coded to highlight the strength of the association between each pair. The force-directed graph can have its nodes and edges change color when a subset from the heatmap is selected. Also, when the mouse hovers over a node, its label, cluster, and direct connections should be highlighted through the change of color and size. Similarly, if it hovers over an edge, new information, such as strength and nodes names, should pop up. Figures 3 and 4 below represent the sketches of the heatmap and the force-directed graph.

Chart, bar chart

Description automatically generated

Figure 3: sketch of the heatmap

Chart

Description automatically generated

Figure 4: sketch of the force-directed graph