LUISS

DELIVERABLE 1
DATA VISUALIZATION PROJECT

Lucas Teixeira and Pedro Aiza

Version: 01 *Creation date: 31/03/23* *Version date: 06/04/23*

Description: this document includes the execution of tasks 3.1, 3.2 and 3.3 of the final Data Visualization project.

3.1. Selection of Dataset

The selection of the dataset was made according to the group's interest on the subject. Also, the opportunity to implement the knowledge learned in class about networks was strongly taken into consideration. Thus, the selected dataset is a real-world, undirected social network of frequent associations between 62 dolphins in a community. It can be found on this website as a *.gml* file.

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.

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 dolphins and their behavior. Furthermore, this information can be used to make a comprehensive approach to sociobiology, allowing one to compare dolphins and humans' communities, for example.

3.3. Visualization Design/Sketch

The main purpose of this project is to create two static and two interactive visualizations. The latter must be interrelated to include the following interactivity concepts: brushing and linking, and details-on-demand. This section includes further explanations on the visualizations the group plans to create, 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, the group will implement a **Kamada-Kawai** network visualization, 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.

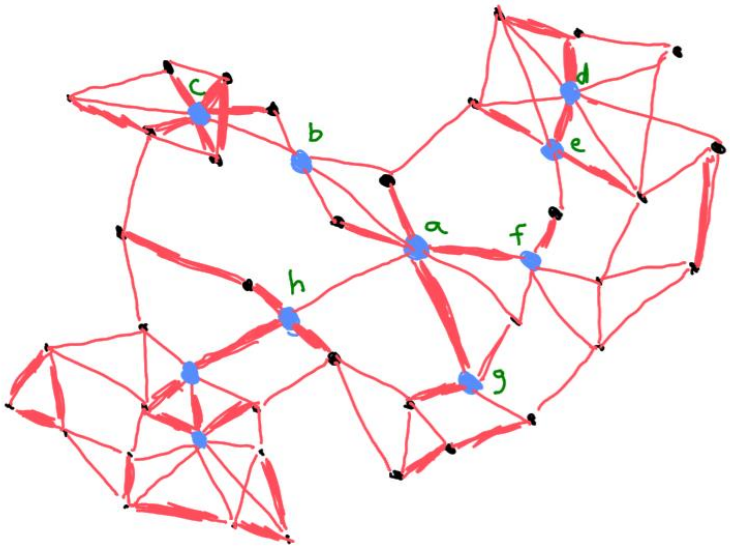

Figure 1: sketch of the Kamada-Kawai graph

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.

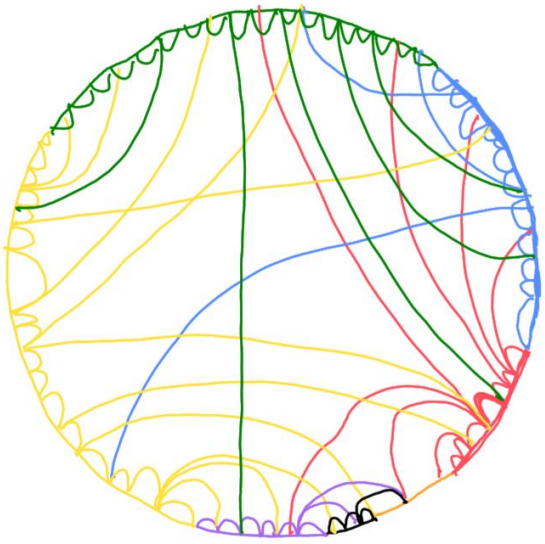

Figure 2: sketch of the chord diagram

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.

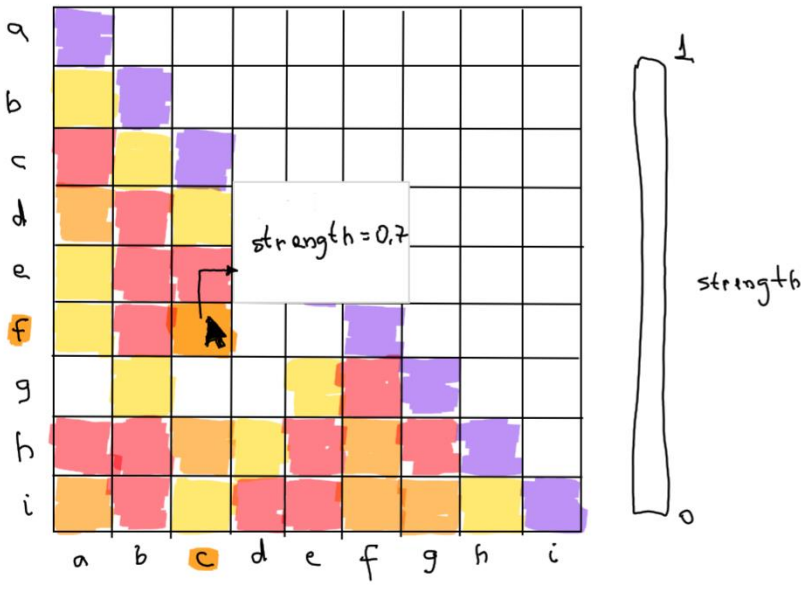

Figure 3: sketch of the heatmap

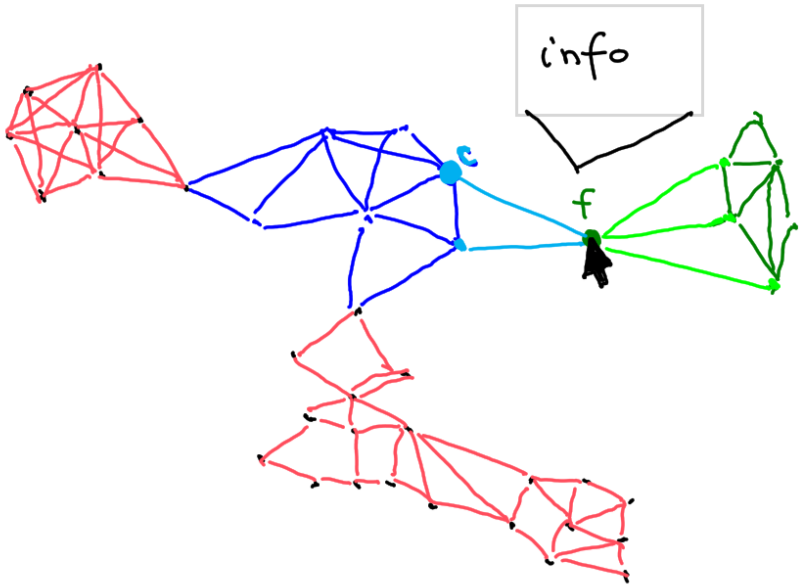

Figure 4: sketch of the force-directed graph