

# Social Network Analysis Assignment

Code ▾

*Raghu Sanugommula and Subhash Pemmaraju*

*27th March 2019*

## Introduction: Initial Summary

The purpose of this project is to analyze the data collected from a survey conducted by a professor in 512 class. This survey data talks about how students are connected for doing their assignments, are they pursuing careers in the same field, and how far the student is living from where the primary responder live, etc. The primary objective is to get students hands-on experience in network graphs visualization topics and gage the knowledge on how to design a compelling social networking graph, which became a dominant analytical approach for analyzing social networking data in today's world. By completing this assignment, we are confident that we would be able to solve real-world problems related to social networking data using data visualizations techniques

## Methodology:

igraph is a library collection for creating and manipulating graphs and analyzing networks. It is widely used in academic research in network science and related fields. It can be used for large social networking analysis efficiently. The data set for this assignment consists of 17 variables and 18 observations. We will be analyzing different categories of network connections between the students. We will use the network diagrams to show relationships visually. We will use the histogram to plot the degrees vs. frequency of the student connections. Besides, we will explore various factor and numeric variables using exploratory, and I graph visualization techniques.

Hide

```
par(mfrow=c(1,2))
set.seed(222)
## Network Diagram 1 ####

plot(network1,vertex.color="red",vertex.size=10,vertex.label.dist=0.1,edge.arrow.size=0.5,vertex.label.cex=1,main="WhoWorkedwithWho",layout=layout.fruchterman.reingold)

##Network Diagram2 ####

plot(network1,vertex.color=rainbow(52),vertex.size=V(network1)$degree*8,edge.arrow.size=1,layout=layout.fruchterman.reingold,main="WhoWorkedwithWho2")
```

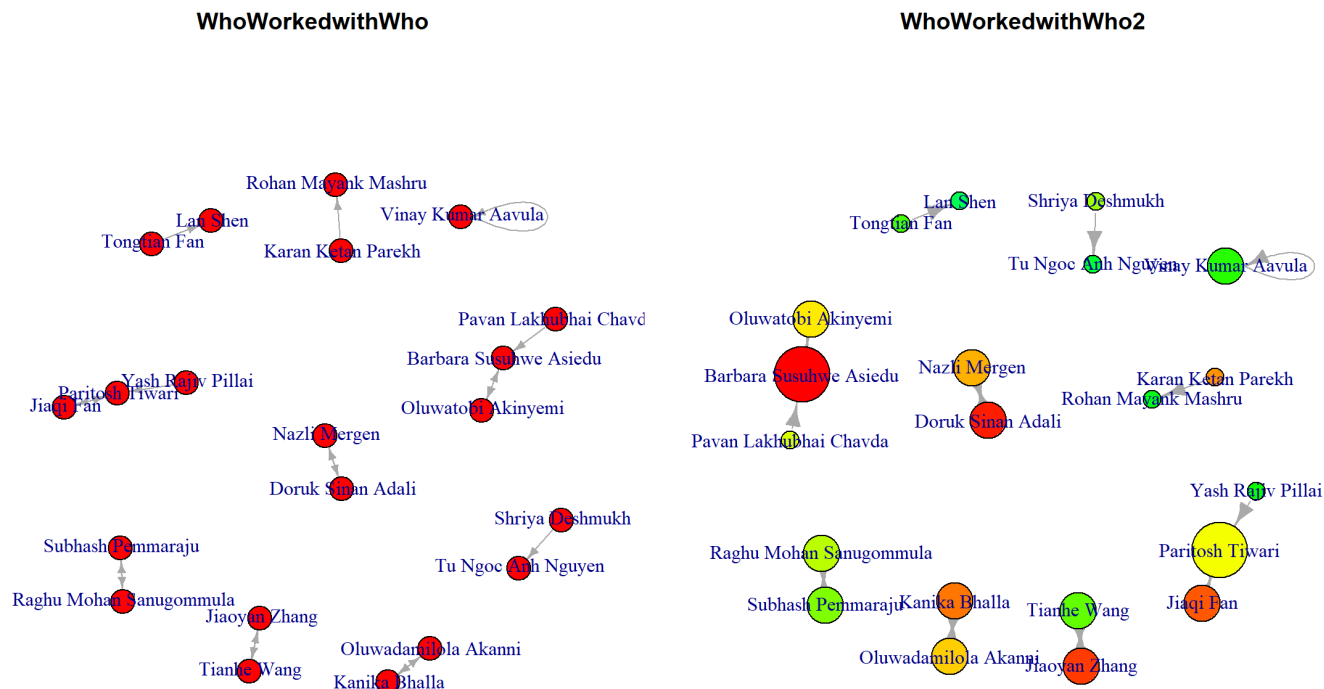


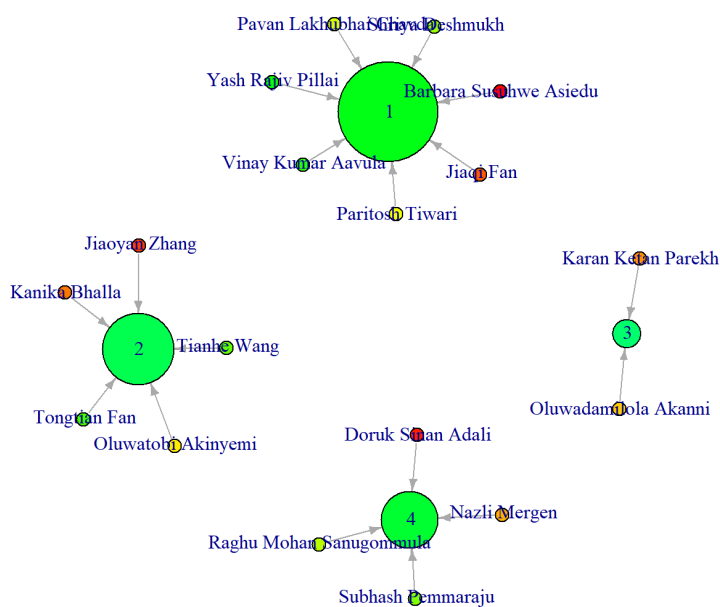
Figure 1 &amp; 2

Figure 1 and two talks relatively the same; however, it has differences in the visuals. We will talk about the differences in this section outlining the above graphs; Graph 1 is talking about who worked with who during their semester, and graph two talks about the same, but the vertices in graph two were derived from the degree of the two character variables. Besides, we tried a different color and basic format. Though there is no much to talk about these graphs since they are straightforward, however, the critical difference that should be observed in the graphs is how the data is being influencing the vertex, the more people have worked with the person, the more the size of the vertex. Also, we can observe how there are some outliers where people worked by themselves without having an assignment partner; in these cases, the arrows are pointing to themselves representing the same. That is another exciting learning from this exercise.

Hide

```
Networkgraph3<-data.frame(dataset$What.is.YOUR.name.,dataset$Student.1..Number.times.you.worked.
with.this.student.this.semester..enter.a.number.,dataset$Student.1..Is.this.student.male.or.female.)
network3<-graph.data.frame(Networkgraph3,directed=T)
V(network3)$degree<-degree(network3)
plot(network3, vertex.color=rainbow(52),vertex.size=V(network3)$degree*5,vertex.label.dist=0.1,edge.arrow.size=0.5,vertex.label.cex=1,layout=layout.fruchterman.reingold,main="How Many Times Student Worked with his partner")
```

### How Many Times Student Worked with his partner

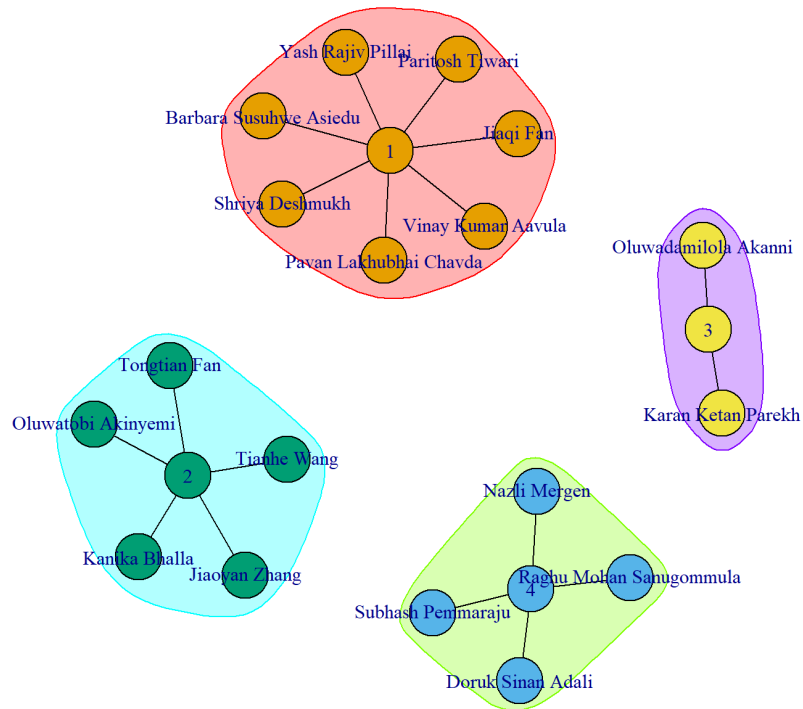


**Figure 3**

Hide

```
network3dense<-graph.data.frame(Networkgraph3,directed=F)
clusternest<-cluster_edge_betweenness(network3dense)
plot(clusternest,network3dense,main="ClusterAnalysis-How many students have chosen same number
of times worked with his partner",vertex.color=rainbow(52),layout=layout.fruchterman.reingold )
```

### ClusterAnalysis-How many students have chosen same number of times worked with his partner



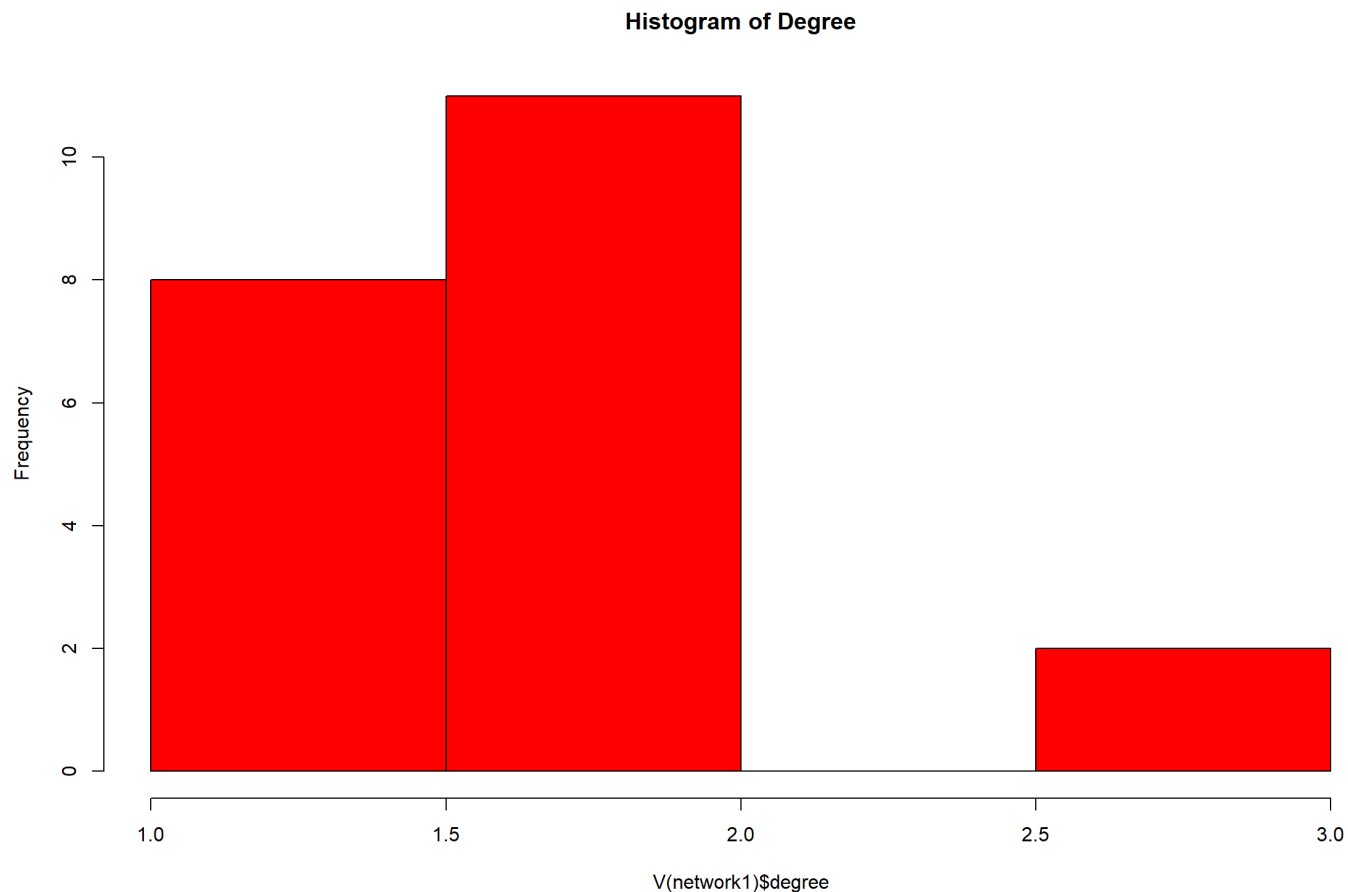
**Figure 4**

Figure 3 and 4 are an extension to the graph 1, it is talking about how many times each student had worked with his partner through out the semester. Also, it shows the clusters of numbers that student have worked with his partner. the larger the size of the edge, the most of the students have chosen that number.

## Histogram

Hide

```
hist(V(network1)$degree,col="red",main="Histogram of Degree",Xlab="Degree of Vertices",Ylab="Frequency")
```



**Figure 5-Histogram**

We have conducted more analysis by taking the degree and plot the values on the histogram. It seems like the degree of the connections fall between 1 and 3 as shown in the figure. By looking at the graph it is clear most of the students have opted between 1 and 2.0 so we should be considered that has mode for this analysis by using the histogram.

## Results and Conclusion

After conducting the minimal analysis using the survey dataset, We admit that it was a great practice exercise where we had an opportunity to use Igraph package and analyze how students are connected in the class and how reliable are the connections in terms of collaboration in the assignments and sharing similar academic and non-academic ideas throughout the semester. The critical observations are, it seems only two students had more connections compare to others and most of the connections have not interacted more than once throughout the semester. We conclude this research with excellent hands-on experience using Igraph package and will continue our research using more variables in the data set or any relevant dataset that could be found via the internet, which could potentially help us to attain knowledge in Igraph/network analysis.