

KU LEUVEN  
NETWORK ANALYSIS  
PROF. DR. ANDRÁS VÖRÖS

---

## Course Paper

---

Halyna Aleksandrovyh  
(r0648181)



April 16, 2018

# Contents

<b>1. Introduction</b>	<b>1</b>
1.1 Dataset . . . . .	1
1.2 Descriptive statistics . . . . .	1
1.3 Missing data . . . . .	4
<b>2. Analysis</b>	<b>4</b>
2.1 Cluster analysis . . . . .	4
2.2 Assortativity . . . . .	5
2.3 Random graphs . . . . .	6
2.4 MRQAP . . . . .	7
2.5 SIENA . . . . .	8
<b>3. Conclusion</b>	<b>9</b>
<b>4. References</b>	<b>10</b>
<b>Appendix A: R Code</b>	<b>10</b>

# **1. Introduction**

## **1.1 Dataset**

The goal of this paper was to study the network dynamics of a classroom of students. The dataset contained information on 34 students from class number 3200 on the variables of friendship, gender, drinking habits, and trust. Drinking habits for each student were measured in two waves. The same was done for measuring friendship and trust within the class.

Previous research in this field has struggled to separate the concepts of selection and influence (Veenstra & Dijkstra, 2011) when choosing and affecting one's friends - do similar people become friends or do people become more similar due to their friendship and not other outside factors? Additional research has suggested an interaction effect of clique gender and alcohol consumption (Jacobs et al, 2016), suggesting that more gender homogeneous cliques might be associated with less drinking as compared to more gender mixed groups.

Based on this, the goal of this paper will be to examine which role, if any, does gender play in the structure of a classroom social network.

## **1.2 Descriptive statistics**

First, the classroom data was visualized for both time points. The plots below include information on gender, plotting the 26 girls as squares and the 8 boys as circles. Drinking information was color-coded in blue, and lastly, affective attitudes were captured by the tie color.

Visually inspecting the two plots, several changes can be observed. Firstly, there seems to be an increase in negative ties at the second time point, though they are still few. Secondly, most students seem to have either remained at the same level or increased their drinking habits. The classroom also seems overall tightly-knit, with just a few outliers such as students #1 and #8 and there seems to be some gender effect with the boys being fairly close together. Lastly, important to mention, three students (#1, #8, and #21 as labeled on the plots) were absent during the second measurement and thus their out-going ties as well as drinking levels are not known for the second wave. The reason for their non-participation in the second round is not known, but perhaps because multiple negative ties can be seen from other classmates towards students #1 and #8, negative interactions have affected their school attendance.

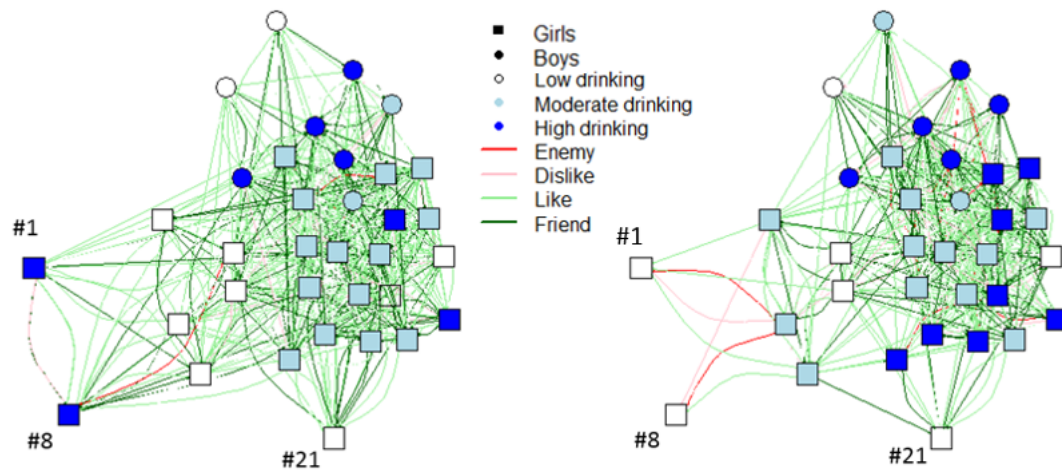


Figure 1: Affective graphs of wave 1 (left) and wave 2 (right)

Looking at a frequency table of tie types, it is confirmed that negatives ties increased over time while neutral and positive ties decreased. This could be due to a natural process of initially being polite and over time establishing stronger bonds and opinions of one's classmates.

Table 1: Affective descriptive statistics

	Enemy	Dislike	Neutral	Like	Friend
Wave 1	3	28	403	459	229
Wave 2	13	54	334	396	205

Because there are too few negatives ties to properly study, it was decided to focus on the friendships in this group, and thus the network was recoded to a friendship network, setting all ties equal zero except for "friend" ties:

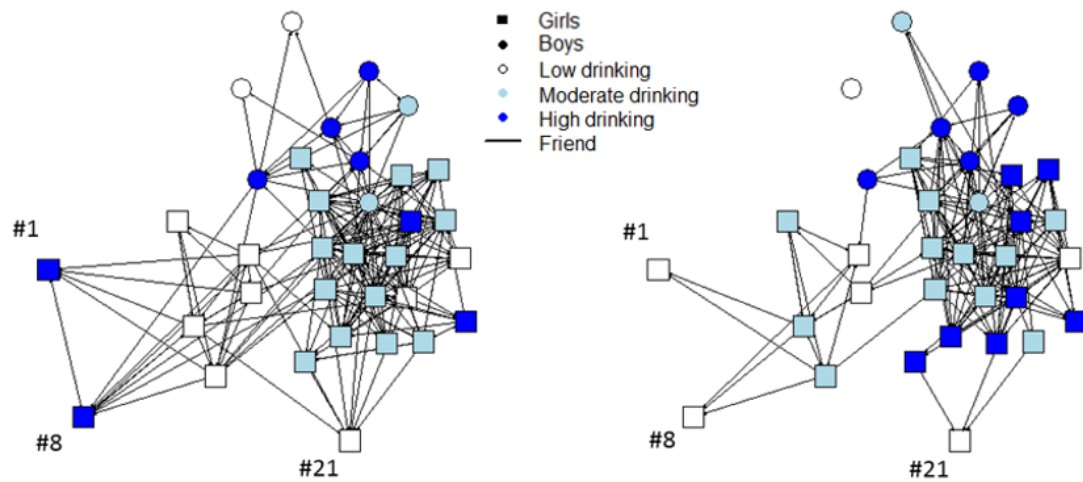


Figure 2: Friendship graphs of wave 1 (left) and wave 2 (right)

Fewer ties can be observed at the second time point and a much more defined cluster at the right with the remaining students having fewer ties. Looking at some descriptive statistics in 2, we can see that friendship ties indeed decreased slightly over time, and that reciprocity jumped from 0.355 to 0.485 while density remained largely the same. This is possibly due to the fact that as more time passed, pairs of students became more clear in their friendships than at the beginning of the study.

Table 2: Friendship descriptive statistics

	Ties	Reciprocity	Density
Wave 1	229	0.355	0.204
Wave 2	205	0.485	0.205

In- and out-degree distributions were also plotted to note any changes over time. At both time points, most students sent out up to 12 out-degrees, though out-degrees overall had a wider range of values than in-degrees which ranged from zero to 13. This suggests that even in the presence of some students which sent out up to 20 friendship ties, there was no clear single popular classmate receiving all of them, as in some other scenarios. 3

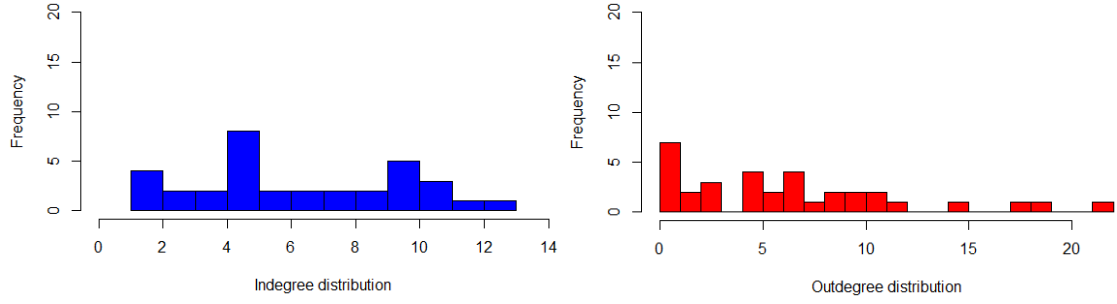


Figure 3: In- and out- degree distribution of friendship wave 1

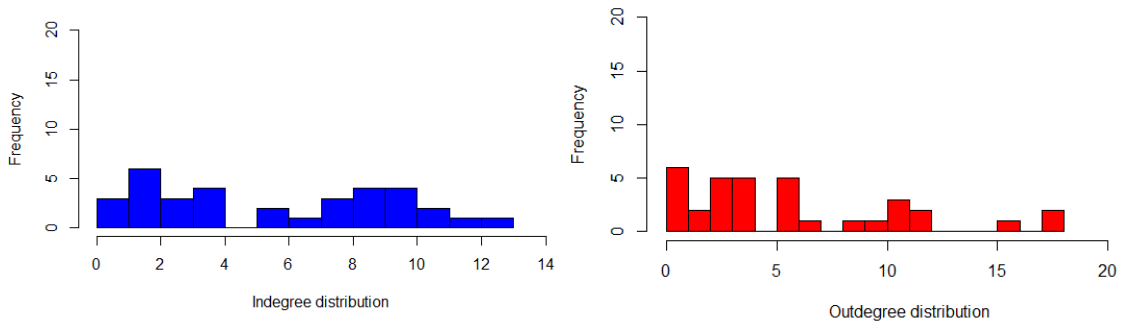


Figure 4: In- and out- degree distribution of friendship wave 2

### 1.3 Missing data

Lastly, as was already mentioned, the data contained some missing values. Specifically, wave 2 information across all variables was missing for students #1, #8, and #21. In addition to this, there were 20 dispersed missing values in friendship wave 2, and one missing values in drinking wave 1. The reasons for the missing data are unclear, as well as whether these reasons are related to this research. This is particularly the case for the three absent students, whose absence might be related to the network dynamics. However, in order to determine anything with greater certainty, more information on the data collection process would be needed.

## 2. Analysis

### 2.1 Cluster analysis

A visual analysis of friendship graphs suggested the presence of one main cluster with some lesser-connected classmates. Multiple cluster analysis approaches were used to test for this, including blockmodeling, k-cores, and fast-and-greedy algorithm.

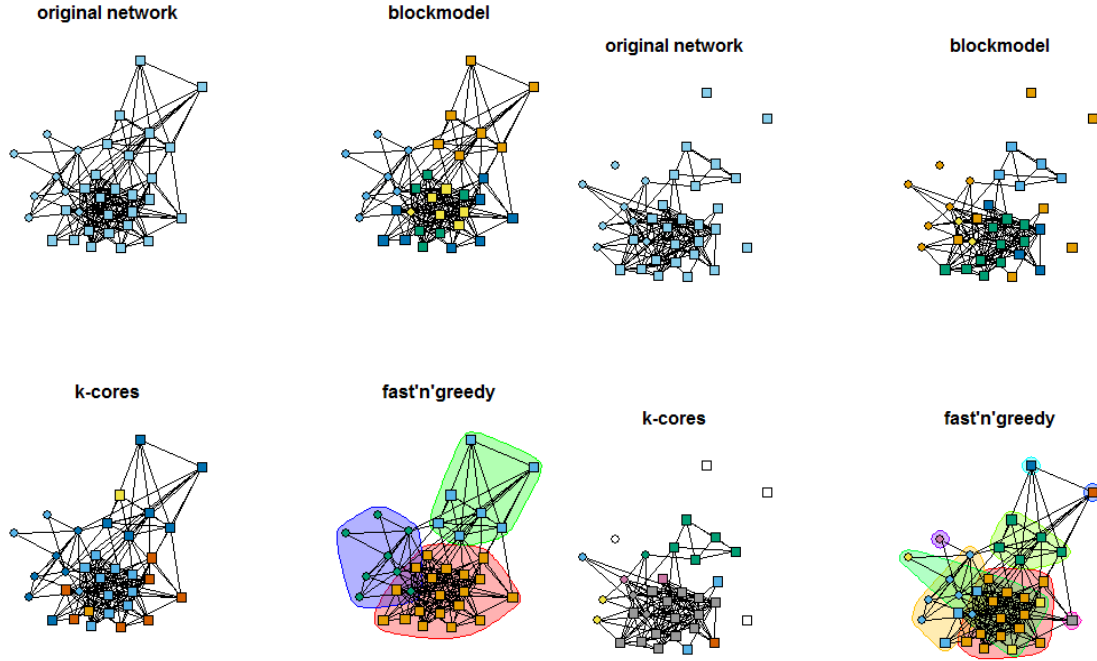


Figure 5: Community detection in friendship wave 1 (left) and wave 2 (right)

All methods show some polarization in the second wave - a cluster that is even tighter-knit, while many students that initially had few ties, having even fewer and being put into small cliques or into cliques of their own. All methods identify different sizes of the same core clique that is the most interconnected and then more loosely-connected periphery communities. Lastly, gender-based cliques are most evident in the first wave where the boys of the class are nearly entirely grouped together by every algorithm. Next, assortativity was looked at, in order to better understand these potential homophily patterns.

## 2.2 Assortativity

First, an assortativity, or homophily, analysis was performed on the friendship network to examine whether there is a higher likelihood of of friendship ties between students that are similar and if so, similar in which ways. For this, the variables of gender, drinking, as well as the number of ties an individual sends and receives were looked at:

Table 3: Assortativity measures of friendship

	Gender	Drinking	In-degree	Out-degree
Wave 1	0.498	0.161	0.420	0.043
Wave 2	0.354	-0.020	0.152	0.089

All assortativity values, except for out-degree decreased over time, suggesting that friend-

ships at the second time point were between less similar students than at the beginning, in terms of these variables. From this table alone though, it is not clear if this was the case of students changing within a friendship, or becoming friends with others.

Nevertheless, some conclusions could be made. The values for gender were fairly high in both friendship waves, suggesting that students were more likely to be friends with those of the same sex. The number of boys and girls in this class though is very unequal (26 girls and 8 boys), and thus it cannot be ruled out that inequality might further encourage such behavior. Gender-based friendship selection tables confirmed this homophily, with the odds ratios in the first friendship wave for a girl to send a tie to a girl rather than a boy being 15.56, and for a boy to send a tie to a boy rather than a girl being 10.32. In the second wave, these numbers were 6.80 for girls and 9.73 for boys. This reflects both that gender homophily decreased slightly for the second wave, as well that gender homophily became for important for boys during the second wave, perhaps because they were a minority.

Drinking showed a slight positive relationship during the first measurement and a very small negative relationship during the second, suggesting that drinking played less of a role in determining friendship as time went on. Lastly, the number of degrees one receives seemed to be more strongly associated with friendship than sending out a similar number of ties.

### **2.3 Random graphs**

Before comparing the friendship network to simulated random graphs, the structure of the two waves was compared using the Jaccard index. It was chosen because of the small sample size as well as high number of non-existent ties, as it disregards them. The Jaccard index was 0.43, meaning that of the ties that existed at any measurement point (either during both or one of the waves, in this case), 43% remained the same. In other words, 43% of the ties existed during both waves while 57% existed only during one wave. This can be interpreted as a measure of stability of friendship, with a little under half of friendships formed during wave 1 surviving until the second measurement period.

Next, the friendship network was compared to a generated set of random graphs in order to examine whether the network structure is random or whether it reflects certain variables, such as forming friendships based on gender or drinking habits. For both waves, 100 random graphs of the same size and density were generated and then the number of dyads and triads compared using violin plots. Prior to those though, frequency plots of dyads and triads were created for the friendship network, showing that nearly all types of tie configurations decreased in the second wave:



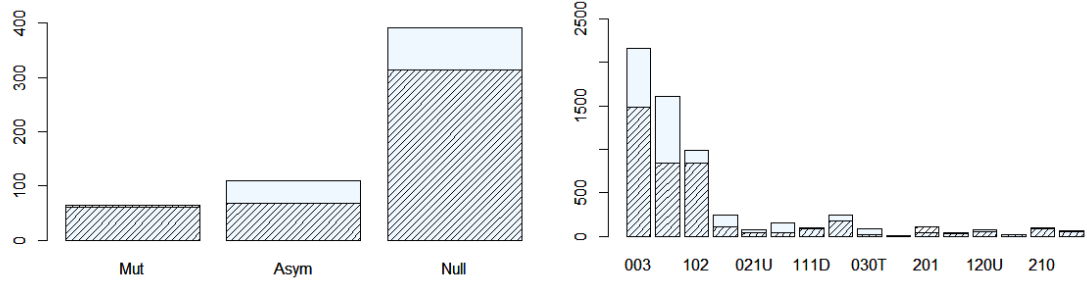


Figure 6: Barplots of dyad (left) and triad (right) types for friendship waves 1 (light blue) and 2 (cross-hatched)

The first two types of triads (003 and 012) were excluded from the triad violin plots to improve visualization but in both waves, the violin plots showed a poor fit with the randomly generated graphs. This suggested that neither of the friendship waves tie structures was random, but that underlying processes affected the likelihood of a tie existing, supporting a further analysis of node characteristic variables and their influence of tie formation.

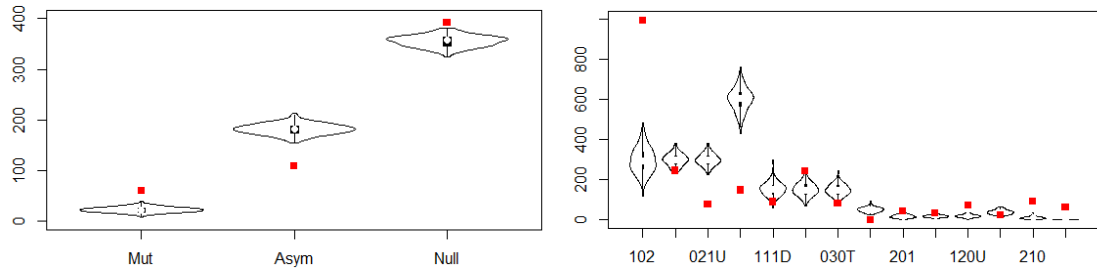


Figure 7: Violin plots of dyads and triads of friendship wave 1

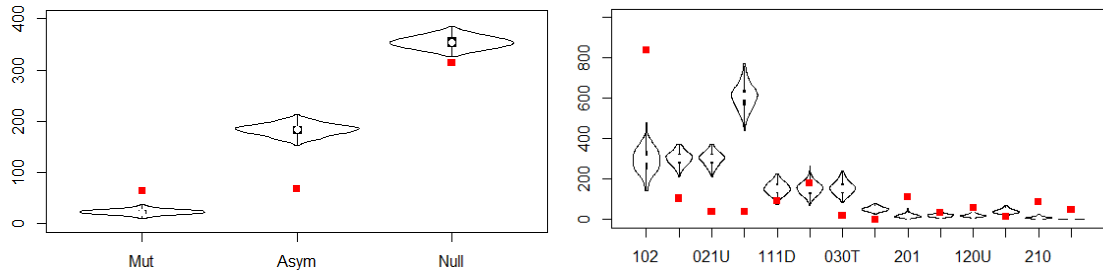


Figure 8: Violin plots of dyads and triads of friendship wave 2

## 2.4 MRQAP

A multiple regression quadratic assignment procedure (MRQAP) was performed to establish whether the correlation between two social network matrices can be explained by a set

of variables, in this case meaning whether the friendship network at time point two can be explained through prior friendships and/or characteristic variables. Though it is similar in intuition to classic linear regression, the latter cannot be used on network data as it does not take into account the dependency structure of the data, making standard errors unreliable. The MRQAP also builds upon the simpler QAP which evaluates the correlation between two matrices against a baseline correlation derived from permuting the network while preserving its structure. The MRQAP additionally allows to include variables which could be used to explain the above correlation.

The MRQAP summary can be seen below:

Table 4: MRQAP summary

	Estimate	Exp(B)	p-value
Intercept	-2.533	0.079	0.000
Wave 1	2.401	11.032	0.000
Gender	0.528	1.695	0.08

The results show a significant effect of wave one friendships on wave two friendships, but not a significant effect of gender (specifically, same-gender choices) on friendship during the second wave. These results ought to be taken with a grain of salt however, given the sample size at hand, as well as the unbalanced gender groups. With a larger and more balanced group, it is possible that the effect of gender does indeed exist as some of the descriptive measures have suggested.

## 2.5 SIENA

Lastly, the SIENA model was fitted. The SIENA model, unlike ERGMs is actor-oriented, rather than tie-oriented and focuses on which potential tie of a given actor could change during the next time period, taking a longitudinal approach. In addition to structural variables, the same covariates were used as in the MRQAP model - friendship during wave one and gender:

Table 5: SIENA summary

	Estimate	St. error	Convergence ratio
outdegree (density)	-2.312	0.265	0.043
reciprocity	1.917	0.263	0.023
transitive triplets	0.268	0.039	0.044
3-cycles	-0.379	0.074	0.053
indegree (popularity)	0.019	0.024	0.039
gender.coCovar alter	-0.536	0.224	-0.009
gender.coCovar ego	0.128	0.216	-0.008
same gender.coCovar	0.486	0.193	0.023

Both certain structural variables and covariates were found to have an effect in the model, including out-degree, reciprocity, transitive triplets, 3-cycles, gender alter (or receiver), as well as same-gender effect. This means that at least in the context of this specific classroom data, these variables were associated with the likelihood of an actor's potential tie changing, for example increased likelihood of a tie appearing if a potential receiver-node is of the same gender as the sender. Once again, caution should be used in extrapolating these results as the sample size was small.

### **3. Conclusion**

The goal of this paper was to examine the role of gender in the network dynamics of a classroom of 34 students. Visual and cluster analysis as well as assortativity measures all pointed towards an association of gender and likelihood of friendship ties. However, this gender effect was not found to be significant in the MRQAP analysis, but was in the SIENA model.

These analyses had many limitations. Firstly, the sample size was small, and gender-wise unbalanced. However, given the often complex dependency structures within social networks, it also cannot be assumed that having a larger social network could yield results representative to that of a single classroom. This shows a need for on one hand weighing statistical advantages and on the other, understanding the changing dynamics between smaller and larger groups. Nonetheless, in order to better study and control for, the effect of gender, a more balanced dataset could have been useful. As mentioned earlier, it is also possible that certain gender dynamics activate when there is a large inequality in size - perhaps two boys might otherwise overlook each other if they were not so keenly aware that there are so few of them in this classroom.

Another limitation were the three missing students during wave two and additional missing values, but particularly the former as their absence could have been the result of social dynamics and thus informational rather than random. Thus, their evolution over time in terms of affective ties and drinking was not captured and hence not included in the analysis.

Nevertheless, there were some clear indications homophily based on gender, which could be further studied in either more balanced classrooms, as well as across more classroom or similar groups, in order to study these gender dynamics on a smaller group level, but have greater data on which to base conclusions.

## 4. References

Wura Jacobs, Patricia Goodson, Adam E. Barry, Kenneth R. McLeroy, E. Lisako J. McKyer & Thomas W. Valente (2016) Adolescent Social Networks and Alcohol Use: Variability by Gender and Type, Substance Use & Misuse, 52:4, 477-487, DOI: 10.1080/10826084.2016.1245333

Veenstra, R., & Dijkstra, J. K. (2011). Transformations in adolescent peer networks. In B. Laursen & W. A. Collins (Eds.), Relationship pathways: From adolescence to young adulthood (pp. 135–154). Los Angeles: Sage.

## Appendix A: R Code

```
setwd('C:\\Users\\Halyna\\Google Drive\\KU Leuven\\Network Analysis\\Week 1')
library(sna)

#load data
affective_w1 <- as.matrix(read.csv("RECENS_data/3200_affective_w1.csv",
                                   header=TRUE, row.names=1, sep=","))
affective_w2 <- as.matrix(read.csv("RECENS_data/3200_affective_w2.csv",
                                   header=TRUE, row.names=1, sep=","))
trust_w1 <- as.matrix(read.csv("RECENS_data/3200_trust_w1.csv",
                               header=TRUE, row.names=1, sep=","))
trust_w2 <- as.matrix(read.csv("RECENS_data/3200_trust_w2.csv",
                               header=TRUE, row.names=1, sep=","))
sex <- as.matrix(read.csv("RECENS_data/3200_sex.csv",
                          header=TRUE, row.names=1, sep=","))
drink <- as.matrix(read.csv("RECENS_data/3200_drink.csv",
                            header=TRUE, row.names=1, sep=","))

affective_w1 # -2 to 2
tail(affective_w2)
sex
drink
trust_w1 # binary
trust_w2

#what is the number of ties of different affective types?
enemy_w1 <- length(which(affective_w1 == -2))
enemy_w2 <- length(which(affective_w2 == -2))
dislike_w1 <- length(which(affective_w1 == -1))
dislike_w2 <- length(which(affective_w2 == -1))
neutral_w1 <- length(which(affective_w1 == 0))
neutral_w2 <- length(which(affective_w2 == 0))
like_w1 <- length(which(affective_w1 == 1))
like_w2 <- length(which(affective_w2 == 1))
friend_w1 <- length(which(affective_w1 == 2))
friend_w2 <- length(which(affective_w2 == 2))
```

```

enemy_w1
enemy_w2
dislike_w1
dislike_w2
neutral_w1
neutral_w2
like_w1
like_w2
friend_w1
friend_w2 #negatives increase, positives and neutrals decrease

#descriptives
class_size <- nrow(affective_w1)
class_size

gender_vector<- c("boys", "girls", sum(sex==1), sum(sex==2))
gender_comp <- matrix(gender_vector, 2,2, byrow=TRUE)
gender_comp

#visualization
library(igraph)
graph.1 <- graph.adjacency(affective_w1)
graph.2 <- graph.adjacency(affective_w2)
graph.12 <- graph.adjacency(affective_w1 + affective_w2)
myLayout <- layout.fruchterman.reingold(graph.12)

#reformat drinking
drinking.w1<- drink[,1]
drinking.w2<- drink[,2]
head(drink)
head(drinking.w1)
head(drinking.w2)

par(mfrow=c(1,2))
plot(graph.1,
      vertex.color = ifelse(drinking.w1 == 1, "white", ifelse(drinking.w1 == 2, "
      lightblue", "blue")),
      vertex.shape = ifelse(sex == 2, "square", "circle"),
      edge.color = ifelse(affective_w1== -2, "red", ifelse(affective_w1==-1, "pink",
      ifelse(affective_w1==0, "white", ifelse(affective_w1==1, "lightgreen", "
      darkgreen")))),
      edge.width = 0.2,
      edge.arrow.size = 0.05,
      vertex.size = 10,
      vertex.label = "",
      layout = myLayout,
      main = "Affective network - wave 1")
plot(graph.2,
      vertex.color = ifelse(drinking.w2 == 1, "white", ifelse(drinking.w2 == 2, "
      lightblue", "blue")),
      vertex.shape = ifelse(sex == 2, "square", "circle"),
      edge.color = ifelse(affective_w2== -2, "red", ifelse(affective_w2==-1, "pink",
      ifelse(affective_w2==0, "white", ifelse(affective_w2==1, "lightgreen", "

```

```

        darkgreen")))),
    edge.width = 0.2,
    edge.arrow.size = 0.05,
    vertex.size = 10,
    vertex.label = "",
    layout = myLayout,
    main = "Affective network - wave 2")
legend('topleft', legend=c("Boys", "Girls", "Low drinking", "Moderate drinking", "
    High drinking",
    "Enemy", "Dislike", "Like", "Friend"), cex=1, pch=c(15,16,1,16,16,NA,NA,NA,NA),
    col=c('black', 'black', "black", "lightblue", "blue", "red", "pink", "lightgreen",
    "darkgreen"),
    lty=c(NA,NA,NA,NA,NA,1,1,1,1), lwd=c(NA,NA,NA,NA,NA,2,2,2,2))

par(mfrow=c(1,1))

#missing data
affective_miss_w1 <- sum(is.na(affective_w1))
affective_miss_w1 #diagonal which is okay
affective_miss_w2 <- sum(is.na(affective_w2))
affective_miss_w2 #students 1,8 & 21 missing plus some random
drink_miss <- sum(is.na(drink))
drink_miss
sex_miss <- sum(is.na(sex))
sex_miss
trust_miss_w1 <- sum(is.na(trust_w1))
trust_miss_w1 #just diagonal
trust_miss_w2 <- sum(is.na(trust_w2))
trust_miss_w2 #students 1,8, & 21, no random

#subtract students 1,8, and 21
affective_w2_full <- affective_w2[-c(1,8,21),]
affective_w2_full
affective_w2_full_miss <- sum(is.na(affective_w2_full))
affective_w2_full_miss

trust_w2_full <- trust_w2[-c(1,8,21),]
trust_w2_full
trust_w2_full_miss <- sum(is.na(trust_w2_full))
trust_w2_full_miss #just diagonal NAs

#check proportion of missing ties
affective_miss_w1 <- affective_miss_w1 / ( nrow(affective_w1) * (ncol(affective_w1)
- 1) )
affective_miss_w1
affective_miss_w2 <- affective_miss_w2 / ( nrow(affective_w2) * (ncol(affective_w2)
- 1) )
affective_miss_w2
trust_miss_w1 <- trust_miss_w1 / ( nrow(trust_w1) * (ncol(trust_w1) - 1))
trust_miss_w1
trust_miss_w2 <- trust_miss_w2 / ( nrow(trust_w2) * (ncol(trust_w2) - 1))
trust_miss_w2

```

```

#FRIENDSHIP ONLY
#affective => friendship
friendship_w1 <- affective_w1
friendship_w1[friendship_w1== -2] <- 0
friendship_w1[friendship_w1== -1] <- 0
friendship_w1[friendship_w1== 0] <- 0
friendship_w1[friendship_w1== 1] <- 0
friendship_w1[friendship_w1== 2] <- 1
table(friendship_w1)

friendship_w2 <- affective_w2
friendship_w2[friendship_w2== -2] <- 0
friendship_w2[friendship_w2== -1] <- 0
friendship_w2[friendship_w2== 0] <- 0
friendship_w2[friendship_w2== 1] <- 0
friendship_w2[friendship_w2== 2] <- 1
table(friendship_w2)

#plots
library(igraph)
graph.3 <- graph.adjacency(friendship_w1)
graph.4 <- graph.adjacency(friendship_w2)
graph.34 <- graph.adjacency(friendship_w1 + friendship_w2)
myLayout2 <- layout.fruchterman.reingold(graph.34)

par(mfrow=c(1,2))
plot(graph.3,
      vertex.color = ifelse(drinking.w1 == 1, "white", ifelse(drinking.w1 == 2, "
        lightblue", "blue")),
      vertex.shape = ifelse(sex == 2, "square", "circle"),
      edge.color = "black",
      edge.width = 0.2,
      edge.arrow.size = 0.05,
      vertex.size = 10,
      vertex.label = "",
      layout = myLayout,
      main = "Friendship network - wave 1")
plot(graph.4,
      vertex.color = ifelse(drinking.w2 == 1, "white", ifelse(drinking.w2 == 2, "
        lightblue", "blue")),
      vertex.shape = ifelse(sex == 2, "square", "circle"),
      edge.color = "black",
      edge.width = 0.2,
      edge.arrow.size = 0.05,
      vertex.size = 10,
      vertex.label = "",
      layout = myLayout,
      main = "Friendship network - wave 2")
legend('topleft', legend=c("Boys", "Girls", "Low drinking", "Moderate drinking", "
  High drinking",
    "Friend"), cex=1, pch=c(15,16,1,16,16,NA),
      col=c('black', 'black', "black", "lightblue", "blue", "black"),

```

```

lty=c(NA,NA,NA,NA,NA,1) , lwd=c(NA,NA,NA,NA,NA,2))
par(mfrow=c(1,1))

#ties
table(friendship_w1)
table(friendship_w2)
#reciprocity
friend_rec_w1 <- grecip(friendship_w1 , measure="dyadic.nonnull")
friend_rec_w2 <- grecip(friendship_w2 , measure="dyadic.nonnull")
friend_rec_w1
friend_rec_w2
#density
friend_dens_w1 <- gden(friendship_w1)
friend_dens_w2 <- gden(friendship_w2)
friend_dens_w1
friend_dens_w2
#in- out- degree histograms
detach("package:igraph")
friend_ind_w1 <- degree(friendship_w1 , cmode = "indegree")
friend_outd_w1 <- degree(friendship_w1 , cmode = "outdegree")
friend_ind_w2 <- degree(friendship_w2 , cmode = "indegree")
friend_outd_w2 <- degree(friendship_w2 , cmode = "outdegree")
friend_ind_w1
friend_outd_w1
friend_ind_w2
friend_outd_w2
#12, 22, 13 18
hist(friend_ind_w1 , xlim=c(0,14), ylim=c(0,20), breaks=13, main ="Histogram of
indegrees", xlab="Indegree distribution", col="blue")
hist(friend_outd_w1 , xlim=c(0,22), ylim=c(0,20), breaks=21, main="Histogram of
outdegrees", xlab="Outdegree distribution", col="red")
hist(friend_ind_w2 , xlim=c(0,14), ylim=c(0,20), breaks=13, main ="Histogram of
indegrees", xlab="Indegree distribution", col="blue")
hist(friend_outd_w2 , xlim=c(0,22), ylim=c(0,20), breaks=21, main="Histogram of
outdegrees", xlab="Outdegree distribution", col="red")

#assortativity
library(igraph)
assortativity.nominal(graph.3, sex)
assortativity.nominal(graph.4, sex)
drinking.w1[is.na(drinking.w1)] <- median(drinking.w1, na.rm=TRUE)
drinking.w2[is.na(drinking.w2)] <- median(drinking.w2, na.rm=TRUE)
assortativity(graph.3, drinking.w1)
assortativity(graph.4, drinking.w2)
assortativity(graph.3, friend_ind_w1)
assortativity(graph.3, friend_outd_w1)
assortativity(graph.4, friend_ind_w2)
assortativity(graph.4, friend_outd_w2)

assortativity.degree(graph.3, directed=TRUE)
assortativity.degree(graph.4, directed=TRUE)

#friendship selection tables

```



```

gg.1 <- friendship_w1[sex==2, sex==2]
gb.1 <- friendship_w1[sex==2, sex==1]
bb.1 <- friendship_w1[sex==1, sex==1]
bg.1 <- friendship_w1[sex==1, sex==2]

friend.selection <- matrix(NA, 2, 2)
rownames(friend.selection) <- c("girl", "boy")
colnames(friend.selection) <- c("girl", "boy")
detach(package:igraph)
library(sna)
friend.selection[1,1] <- gden(gg.1, diag=FALSE)
friend.selection[1,2] <- gden(gb.1, diag=TRUE)
friend.selection[2,2] <- gden(bb.1, diag=FALSE)
friend.selection[2,1] <- gden(bg.1, diag=TRUE)

(friend.selection.norm <- friend.selection / gden(friendship_w1))
friend.selection[1,1]/friend.selection[1,2]
friend.selection[2,2]/friend.selection[2,1]

gg.2 <- friendship_w2[sex==2, sex==2]
gb.2 <- friendship_w2[sex==2, sex==1]
bb.2 <- friendship_w2[sex==1, sex==1]
bg.2 <- friendship_w2[sex==1, sex==2]

friend.selection2 <- matrix(NA, 2, 2)
rownames(friend.selection2) <- c("girl", "boy")
colnames(friend.selection2) <- c("girl", "boy")
detach(package:igraph)
library(sna)
friend.selection2[1,1] <- gden(gg.2, diag=FALSE)
friend.selection2[1,2] <- gden(gb.2, diag=TRUE)
friend.selection2[2,2] <- gden(bb.2, diag=FALSE)
friend.selection2[2,1] <- gden(bg.2, diag=TRUE)

(friend.selection.norm2 <- friend.selection2 / gden(friendship_w2))
friend.selection2[1,1]/friend.selection2[1,2]
friend.selection2[2,2]/friend.selection2[2,1]

#comparing the two wave networks
(hamming <- hdist(friendship_w1, friendship_w2))
(hamming.prop <- hamming/nties(friendship_w1))
(matching <- 1 - hamming.prop)
#shows a stable network, but
#because most ties are absent and hamming sensitive to that
#Jaccard index - disregards absent ties
A <- sum((friendship_w1 * friendship_w2)==1, na.rm=TRUE)
BplusC <- sum((friendship_w1 + friendship_w2)==1, na.rm=TRUE)
(jaccard <- A/(A+BplusC))

#dyads and triads
dyad.count <- dyad.census(friendship_w1)
triad.count <- triad.census(friendship_w1)
dyad.count2 <- dyad.census(friendship_w2)

```

```

triad.count2 <- triad.census(friendship_w2)
dyad.count
triad.count
dyad.count2
triad.count2
barplot(dyad.count, col="aliceblue", border=TRUE, ylim = c(0,400))
barplot(dyad.count2, add=TRUE, col="black", density=20, border = TRUE)
barplot(triad.count, col="aliceblue", border=TRUE, ylim = c(0,2500),
names.arg=c("003", "012", "102", "021D", "021U", "021C", "111D", "111U", "030T",
"030C", "201", "120D", "120U", "120C", "210", "300"))
barplot(triad.count2, add=TRUE, col='black', density=20)

#random graphs
#testing dyad/triad census against random graphs
net.size <- nrow(friendship_w1)
net.dens <- gden(friendship_w1)
random.nets <- rgraph(net.size, 200, net.dens)
dim(random.nets)
random.dens <- gden(random.nets)
hist(random.dens)
mean(random.dens)
net.dens #super similar
random.triad <- triad.census(random.nets)
random.dyad <- dyad.census(random.nets)
library(vioplot)
vioplot(random.triad[,3],random.triad[,4],random.triad[,5],
        random.triad[,6],random.triad[,7],random.triad[,8],random.triad[,9],random.
        triad[,10],
        random.triad[,11], random.triad[,12], random.triad[,13], random.triad[,14],
        random.triad[,15],
        random.triad[,16],
        names=colnames(random.triad)[3:16],
        col="transparent",
        ylim=c(0, 1000))
points(1:14, triad.count[c(3,4,5,6,7,8,9,10,11,12,13,14,15,16)],
        col="red",
        pch=15) #1 003 above, 2 012 below
vioplot(random.dyad[,1], random.dyad[,2], random.dyad[,3],
        names=colnames(random.dyad), col = "transparent",
        ylim=c(0,400))
points(1:3, dyad.count[c(1,2,3)],
        col="red",
        pch=15)

net.size2 <- nrow(friendship_w2)
net.dens2 <- gden(friendship_w2)
random.nets2 <- rgraph(net.size2, 200, net.dens2)
dim(random.nets2)
random.dens2 <- gden(random.nets2)
hist(random.dens2)
mean(random.dens2)
net.dens2 #super similar
random.triad2 <- triad.census(random.nets2)

```

```

random.dyad2 <- dyad.census(random.nets2)
library(vioplplot)
vioplplot(random.triad2[,3],random.triad2[,4],random.triad2[,5],
          random.triad2[,6],random.triad2[,7],random.triad2[,8],random.triad2[,9],
          random.triad2[,10],
          random.triad2[,11], random.triad2[,12], random.triad2[,13], random.triad2
            [,14], random.triad2[,15],
          random.triad2[,16],
          names=colnames(random.triad2)[3:16],
          col="transparent",
          ylim=c(0, 1000))
points(1:14, triad.count2[c(3,4,5,6,7,8,9,10,11,12,13,14,15,16)] ,
       col="red",
       pch=15) #1 003 fits , 2 012 way below
vioplplot(random.dyad2[,1], random.dyad2[,2], random.dyad2[,3],
          names=colnames(random.dyad2), col = "transparent",
          ylim=c(0,400))
points(1:3, dyad.count2[c(1,2,3)] ,
       col="red",
       pch=15)

#cluster detection
friend_w1 <- friendship_w1 + t(friendship_w1)
friend_w1[friend_w1==2] <- 1
detach(package:sna)
library(igraph)

friend1 <- graph.adjacency(friend_w1)
friend1 <- as.undirected(friend1)
cliques <- cliques(friend1)
length(cliques)
table(sapply(maximal.cliques(friend1), length))
cores <- graph.coreness(friend1)

detach(package:igraph)
library(sna)
gplot(friend_w1)
equiv.w1 <- equiv.clust(friend_w1, cluster.method="ward.D2", method="hamming")
plot(equiv.w1)
bm.w1 <- blockmodel(friend_w1, equiv.w1, k=5)
plot.sociomatrix(friend_w1, diaglab = FALSE)
plot.sociomatrix(bm.w1$blocked.data, diaglab=FALSE)
block.members <- bm.w1$block.membership[order(bm.w1$order.vector)]
gplot(friend_w1, vertex.col=block.members)

communities <- fastgreedy.community(friend1)
length(communities)
sizes(communities)
membership(communities)

par(mfrow=c(2,2))
# the original friendship network
plot(friend1,

```

```

    edge.color = "black",
    edge.width = 1.5,
    edge.arrow.size = 0.25,
    vertex.size = 10,
    vertex.label = "",
    vertex.color="skyblue",
    vertex.shape = ifelse(sex == 2, "square", "circle"),
    layout=myLayout,
    main="original network")
# the blockmodel (from the previous script)
plot(friend1,
     edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.color= block.members,
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="blockmodel")
# the k-cores
plot(friend1,
     edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.color= cores,
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="k-cores")
# the fast-greedy communities
plot(communities, friend1, edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="fast 'n' greedy")
par(mfrow=c(1,1))

friend_w2 <- friendship_w2 + t(friendship_w2)
friend_w2[friend_w1==2] <- 1
detach(package:sna)
library(igraph)

friend2 <- graph.adjacency(friend_w2)
friend2 <- as.undirected(friend2)
cliques2 <- cliques(friend2)
length(cliques2)
table(sapply(maximal.cliques2(friend2), length))
cores2 <- graph.coreness(friend2)

```

```

detach(package:igraph)
library(sna)
gplot(friend_w2)
equiv.w2 <- equiv.clust(friend_w2, cluster.method="ward.D2", method="hamming")
plot(equiv.w2)
bm.w2 <- blockmodel(friend_w2, equiv.w2, k=5)
plot.sociomatrix(friend_w2, diaglab = FALSE)
plot.sociomatrix(bm.w2$blocked.data, diaglab=FALSE)
block.members2 <- bm.w2$block.membership[order(bm.w2$order.vector)]
gplot(friend_w2, vertex.col=block.members2)
library(igraph)
communities2 <- fastgreedy.community(friend2)
length(communities2)
sizes(communities2)
membership(communities2)

par(mfrow=c(2,2))
# the original friendship network
plot(friend2,
     edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.color="skyblue",
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="original network")
# the blockmodel (from the previous script)
plot(friend2,
     edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.color= block.members2,
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="blockmodel")
# the k-cores
plot(friend2,
     edge.color = "black",
     edge.width = 1.5,
     edge.arrow.size = 0.25,
     vertex.size = 10,
     vertex.label = "",
     vertex.color= cores2,
     vertex.shape = ifelse(sex == 2, "square", "circle"),
     layout=myLayout,
     main="k-cores")
# the fast-greedy communities
plot(communities2, friend1, edge.color = "black",

```

```

    edge.width = 1.5,
    edge.arrow.size = 0.25,
    vertex.size = 10,
    vertex.label = "",
    layout=myLayout,
    vertex.shape = ifelse(sex == 2, "square", "circle"),
    main="fast 'n' greedy")
par(mfrow=c(1,1))

#MRQAP
qap1 <- netlogit(friendship_w2, friendship_w1, nullhyp="qap", reps=100)
qap1

same.sex <- sex %>% t(sex)
same.sex[same.sex==2] <- 0
same.sex[same.sex==4] <- 1

qap2<- netlogit(friendship_w2, list(friendship_w1, same.sex), nullhyp='qap', reps
=100)
qap2
summary(qap2)

#SIENA
library(RSiena)
library(igraph)

nActors <- dim(friendship_w1)[1]
friendship.dependent <- sienaDependent(array(c(friendship_w1, friendship_w2),
dim=c(nActors, nActors, 2)))
gender<- unlist(read.csv("RECENS_data/3200_sex.csv",
header=TRUE, row.names=1, sep=","))
gender.coCovar <- coCovar(gender)
mySienaData <- sienaDataCreate(friendship.dependent,
gender.coCovar)
print01Report(mySienaData,
modelname="friendship siena")

mySienaEffects <- getEffects(mySienaData)
mySienaEffects <- includeEffects(mySienaEffects, transTrip, cycle3)
mySienaEffects <- includeEffects(mySienaEffects, inPop)
mySienaEffects <- includeEffects(mySienaEffects, egoX, altX, sameX, interaction1="
gender.coCovar")
mySienaEffects

mySienaAlgorithm <- sienaAlgorithmCreate(projname="friendship siena",
MaxDegree=c(friendship.dependent=50))
result <- siena07(mySienaAlgorithm,
data=mySienaData,
effects=mySienaEffects)

result

siena.table(result, type="html", file="results.html")

```