Assignment 1

by Olanrewaju Labulo (19-950-179) Ximena Sutter (15-809-783) Pavel Senchanka (19-950-997)

Loading Libraries:

1a)

```
affective_network <- read.csv('2400_affective_w1.csv')
gender <- read.csv('2400_sex.csv')

##Dropiing IDs

affective_network = affective_network[,-1]
colnames(affective_network) <- gender$student.ID
```

1b)

code:

affective_network[affective_network < 2] = 0

affective_network[affective_network == 2] = 1

affective_network[is.na(affective_network)] = 0

affective_network_mat = data.matrix(affective_network)

1c)

```
friendship.igraph <- graph_from_adjacency_matrix(affective_network_mat,
                                                 mode = "directed",
                                                 diag = FALSE
)

num_students = dim(affective_network_mat)[1]

paste("Network Size: ", num_students)
```

```
## [1] "Network Size:  27"
```

```
paste("Density: ", sum(affective_network_mat) / length(affective_network_mat))
```

```
## [1] "Density:  0.181069958847737"
```

```
paste("Average Degree: ", sum(affective_network_mat) / num_students)
```

```
## [1] "Average Degree:  4.88888888888889"
```

```
paste("Reciprocity: ", grecip(affective_network_mat, measure = "dyadic.nonnull"))
```

```
## [1] "Reciprocity:  0.360824742268041"
```

```
paste("Number of Males: ", sum(gender$sex == 1))
```

```
## [1] "Number of Males:  11"
```

```
paste("Number of Females:", sum(gender$sex == 2))
```

```
## [1] "Number of Females: 16"
```

```
paste("Female to Female connections: ", sum(affective_network[gender$sex == 2,gender$sex == 2]))
```

```
## [1] "Female to Female connections:  72"
```

```
paste("Male to Male connections", sum(affective_network[gender$sex == 1,gender$sex == 1]))
```

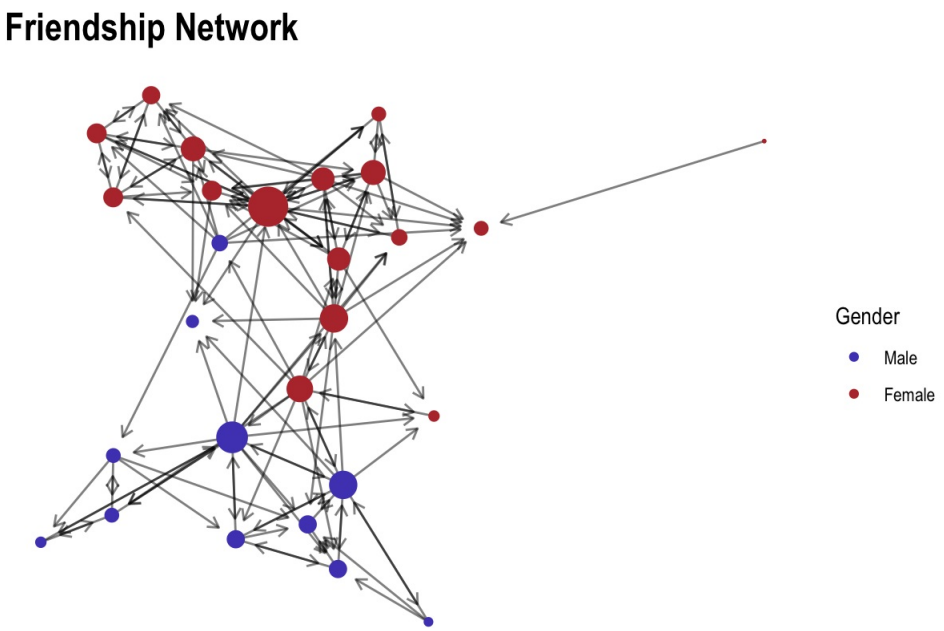
```
## [1] "Male to Male connections 36"
```

```
paste("Max Betweeness: ", max(igraph::betweenness(friendship.igraph)))
```

```
## [1] "Max Betweeness:  144.975"
```

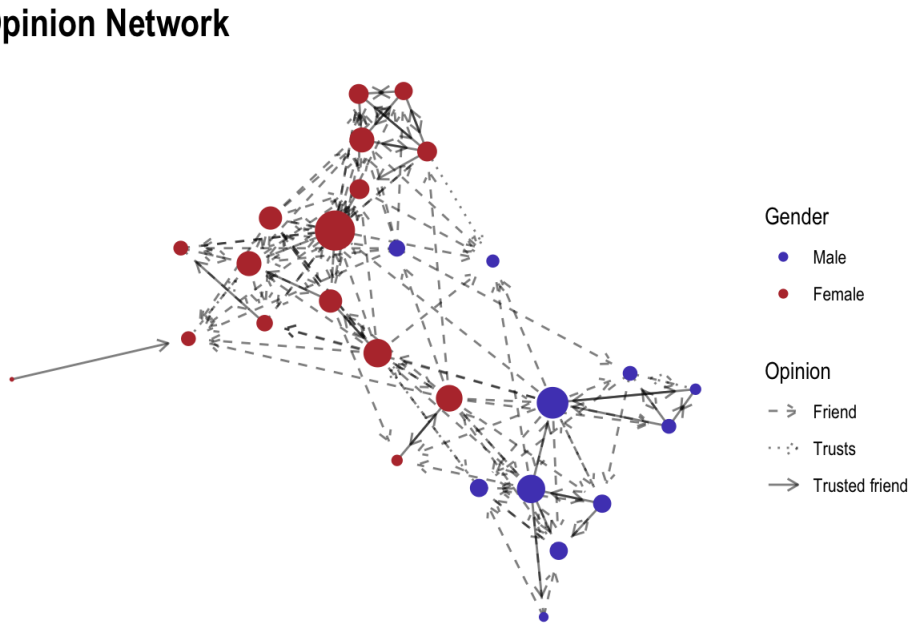
1d)

```
node_gender_colour <- function() {
  return(
    scale_colour_manual(
      breaks=1:2,
      values=c('#5147bf', '#b83739'),
      labels=c('Male', 'Female')))
}

students = gender
students$degree <- sna::degree(affective_network_mat, cmode = 'freeman')

my.graph <- create_layout(friendship.igraph,
                          layout = 'fr')
set.seed(52)
ggraph(my.graph)+
  geom_edge_link(alpha = .5,
                 arrow = arrow(length = unit(2.0, 'mm')),
                 end_cap = circle(3.4, 'mm'))+
  geom_node_point(
    aes(colour = as.factor(gender$sex)),
    size = (students$degree/2.7))+
  scale_size_continuous(range = c(2,4)) +
  node_gender_colour()+
  labs(colour = "Gender")+
  theme_graph()+
  ggtitle("Friendship Network")
```

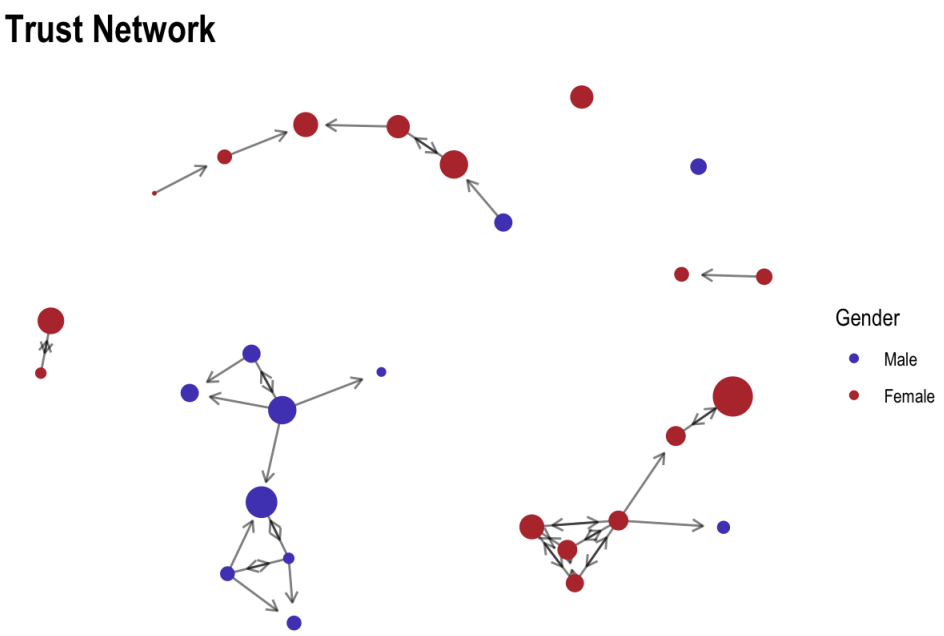


1e)

```
trust <- read.csv('2400_trust_w1.csv')
trust <- trust[,-1]
trust[is.na(trust)] = 0

trust_mat <- data.matrix(trust)

combined_friend_trust_mat <- affective_network_mat + 2*trust_mat

interpret_weight <- function(weight) {
  result <- rep('only_friend', length(weight))
  result[weight == 2] <- 'only_trust'
  result[weight == 3] <- 'trusted_friend'
  return(result)
}

set.seed(52)
ggraph(graph.adjacency(combined_friend_trust_mat, weighted = T),
       layout='fr')+
  geom_edge_link(alpha = .5,
                 aes(
                   lty=interpret_weight(weight),
                 ),
                 arrow = arrow(length = unit(2.0, 'mm')),
                 end_cap = circle(3.4, 'mm'))+
  scale_edge_linetype_manual(
    breaks=c('only_friend', 'only_trust', 'trusted_friend'),
    values=c('dashed', 'dotted', 'solid'),
    labels=c('Friend', 'Trusts', 'Trusted friend'))+
  geom_node_point(
    aes(colour = as.factor(gender$sex)),
    size = (students$degree/2.7) )+
  scale_size_continuous(range = c(2,4))+
  node_gender_colour()+
  labs(colour = "Gender",
       edge_linetype = "Opinion")+
  theme_graph()+
  ggtitle("Opinion Network")
```



```
# Part E (bonus): trust network only

set.seed(52)
ggraph(graph.adjacency(trust_mat),
       layout='fr')+
  geom_edge_link(alpha = .5,
                 arrow = arrow(length = unit(2.0, 'mm')),
                 end_cap = circle(3.4, 'mm'))+
  geom_node_point(
    aes(colour = as.factor(gender$sex)),
    size = (students$degree/2.7))+
  scale_size_continuous(range = c(2,4)) +
  node_gender_colour()+
  labs(colour = "Gender")+
  theme_graph()+
  ggtitle("Trust Network")
```



1f)

```
num_only_friend = sum(combined_friend_trust_mat == 1)
num_only_trust = sum(combined_friend_trust_mat == 2)
num_trusted_friend = sum(combined_friend_trust_mat == 3)

paste('Friend, no trust: ', num_only_friend)
```

```
## [1] "Friend, no trust:  98"
```

```
paste('Trust, not a friend: ', num_only_trust)
```

```
## [1] "Trust, not a friend:  4"
```

```
paste('Trusted friend: ', num_trusted_friend)
```

```
## [1] "Trusted friend:  34"
```

1g)

```
paste('Trust reciprocity: ', grecip(trust_mat, measure="dyadic.nonnull"))
```

```
## [1] "Trust reciprocity:  0.461538461538462"
```

It is quite uncommon to see trust without high level of friendship (only 4 occurences out of 38 trust edges). Friendship, on the other hand, does not necessarily imply trust (only 34 of the 132 friendship edges are also trusted). Trust between genders is basically non-existent (only two instances). In general, the network is well clustered by gender. There is a visible center node in the female part, the male part is more decentralized. There are, interestingly, two nodes where in-degree is high, but out-degree is near zero. None of the connections to these nodes are trusted friendships though, whichindicates a certain level of isolation. Both friendship and trust reciprocity is not high at 36% and 46% respectively.

Task 2: Empirical Research Cycle

(a) You want to examine the social factors contributing to why some students of a university class drink more alcohol than others. How would you go about? What hypothesis would you test? Which kind of data would you collect? Structure your response using the steps of the empirical research cycle. You do not need to analyze any data { just indicate what you would do in each step of the cycle.

Theory/Observations

A possibility why some students drink more often alcohol could be, that they encounter more situations in which alcohol is been drunk (Thursday and Friday beer with friends) or that they drink less often but in a large amount like in a party. Another observation is that especially women do not like to drink alcohol alone. Therefore, we could conclude that people who like to drink, do not drink alone and if they have more occasions, they drink more often and consequently a larger amount.

Induction

Hypothesis: People who have more connections to people who drink, drink more.

Deduction

The prediction is that people who drink more alcohol have a higher degree of connections to people who drink than people who do not drink much. This will be evaluated with the degree centrality value.

Testing on Data

Data collection : Computer survey.
Questions:

- With whom do you interact at least once a week beside in the lectures? (to have a list of all students enrolled in the university class) How good are you friends with them? (rate from 1-5) Do you drink with them? (no:0, yes:1) -> Frequency of contact
- Do you drink? (No/ less than once a month:0) If yes how often? (Yes: once a month - 2 beers a week: 1; more than two beers a week 2)
- Gender?

Research design: Cross sectional study. Select a university class with mixed gender. Major of the university class should not be related to health science, like for example health science and technology. The knowledge about the health damage of alcohol could be a possible bias. Clear boundary because only students in the class will be in the network.

Data management: Excel -> RStudio. In Excel the answers are defined in numbers so we can transfer it easier to RStudio and process it.

Data Visualization: network: include gender (shape), How often they drink (3 colours, 1st for no drinker, 2nd for little amount, 3rd for much), connections of the nodes are made with

whom the students interact on a regular basis (at least once a week) beside the lectures. No names will be presented in the network -> anonymous. To link the connections, we need the names in the survey, because only linkages are made if both students name each other.

Evaluation

Calculate the degree centrality of the nodes. Comparison between nodes who drink a lot and nodes who drink less. The degree of centrality can be assessed in two ways. First, we only look at the connections with people who drink a lot and then in a second step we calculate the degree considering all connections. We take the ratio of both values to receive a fraction of how many people we hang out with, drink a lot of alcohol.
Further, we can compare the values between men and women.

(b) Chose one of the introduced network theories, derive a hypothesis and design your own social network study. Structure your response using the steps of the empirical research cycle. Again: You do not need to analyze any data { just indicate what you would do in each step of the cycle.

Theory/Observations

People who are known to be extroverted and sociable tend to know a lot of people from different places and background. They easily form friendships and stand up for themselves. Therefore, they do not need the feeling of protection of a clique.

Induction

Hypothesis: Extroverted people who easily form friendships are in a structural hole.

Deduction

The extroverted people are the connection between several cliques. This should be seen in high value of betweenness centrality. They also have a high degree of centrality, because they easily form friendships.

Testing on Data

Data collection: Interview and survey. Survey should reflect how people perceive themselves. Interview is more as a control if the person is classified as being extroverted or not. We should keep in mind, that some people are something in between and having both evaluations hopefully make the classification easier.

Questions:

- Are you an active member in the association? If yes, which function do you have in the association?
- Can you make easily friends?
- With whom in the association are you friends with?
- How extroverted are you? (have a scale)

Research Design: Cross sectional study. Select a student association with about 200 members and has male and female members. Clear boundary because we look at a student association.

Data Management: Excel -> RStudio. In Excel the answers are defined in numbers so we can transfer it easier to RStudio and process it.

Data Visualization: network: include gender (shape), extroverted or not (colour), Arrows to nodes participant consider to be friends with.

Evaluation

Calculate betweenness centrality and degree centrality. With the betweenness centrality we can assess the brokers. Degree centrality should represent how many friends we have. Investigate also the correlation between being extroverted and the function in the association.