Seminar 4

Lesnykh Kirill

February 9, 2019

```
wd<-getwd()
```

```
SWFile <- "/SouthernWomen/SouthernWomen.tsv"
SWFilePath <- paste(wd, SWFile, sep = "")
```

```
SWrawdata <- read.table(SWFilePath, sep = "\t", header = TRUE, row.names = 1)
SWrawdata
```

```
##           E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11 E12 E13 E14
## EVELYN     1  1  1  1  1  1  0  1  1   0   0   0   0   0
## LAURA      1  1  1  0  1  1  1  1  0   0   0   0   0   0
## THERESA    0  1  1  1  1  1  1  1  1   0   0   0   0   0
## BRENDA     1  0  1  1  1  1  1  1  0   0   0   0   0   0
## CHARLOTTE  0  0  1  1  1  0  1  0  0   0   0   0   0   0
## FRANCES    0  0  1  0  1  1  0  1  0   0   0   0   0   0
## ELEANOR    0  0  0  0  1  1  1  1  0   0   0   0   0   0
## PEARL      0  0  0  0  0  1  0  1  1   0   0   0   0   0
## RUTH       0  0  0  0  1  0  1  1  1   0   0   0   0   0
## VERNE      0  0  0  0  0  0  1  1  1   0   0   1   0   0
## MYRNA      0  0  0  0  0  0  0  1  1   1   0   1   0   0
## KATHERINE  0  0  0  0  0  0  0  1  1   1   0   1   1   1
## SYLVIA     0  0  0  0  0  0  1  1  1   1   0   1   1   1
## NORA       0  0  0  0  0  1  1  0  1   1   1   1   1   1
## HELEN      0  0  0  0  0  0  1  1  0   1   1   1   1   1
## DOROTHY    0  0  0  0  0  0  0  0  1   1   1   0   1   0   0
## OLIVIA     0  0  0  0  0  0  0  0  0   1   0   1   0   0   0
## FLORA      0  0  0  0  0  0  0  0  0   1   0   1   0   0   0
```

Assignment task. In your Seminar 4 folder, there is another file with data, HiTech. Create the code that will do the following: ??? Create a directory that will consist of a working directory name and the HiTech folder name, so that it opens the HiTech folder. ??? Create file paths to separate data files “GivesAdviceTo,” “IsFriendOf,” “ReportsTo,” and “HiTechAtt.”"

```
HiTechFolder <- paste(wd, "/HiTech/", sep = "")
HiTechFolder
```

```
## [1] "/Users/Lesnykh/Google \320\224\320\270\321\201\320\272/\320\241\320\265\321\202\320\265\
```

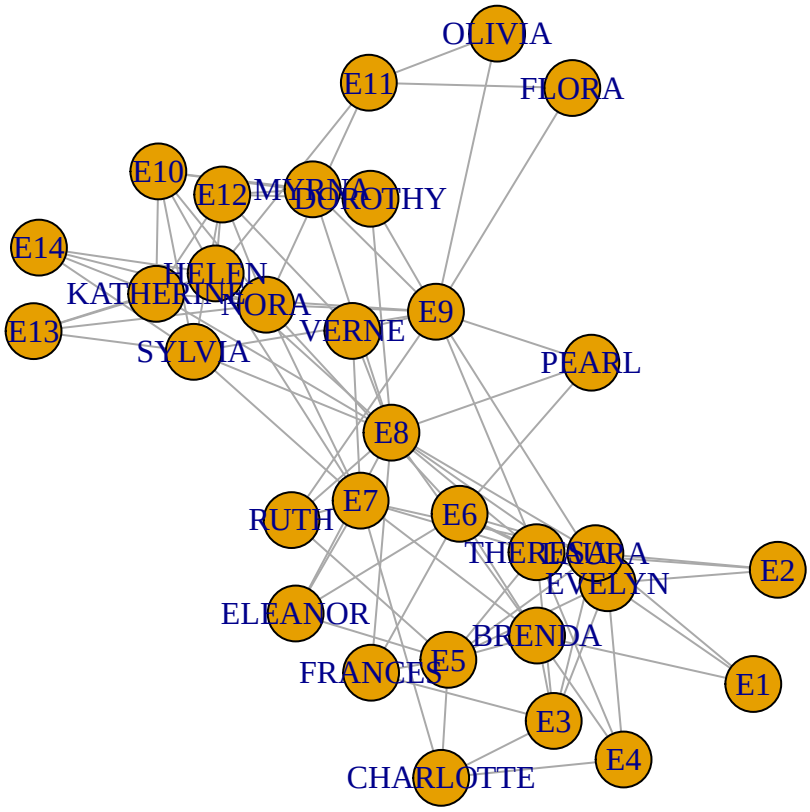
```
GivesAdviceToPath <- paste(HiTechFolder, "HiTech - GivesAdviceTo.tsv", sep="")
```

```
ReportsToPath <- paste(HiTechFolder, "HiTech - ReportsTo.tsv", sep="")

HiTechAttPath <- paste(HiTechFolder, "HiTech - HiTechAtt.tsv", sep="")

IsFriendsOfPath <- paste(HiTechFolder, "HiTech - IsFriendsOf.tsv", sep="")
```

```
library(igraph)
```

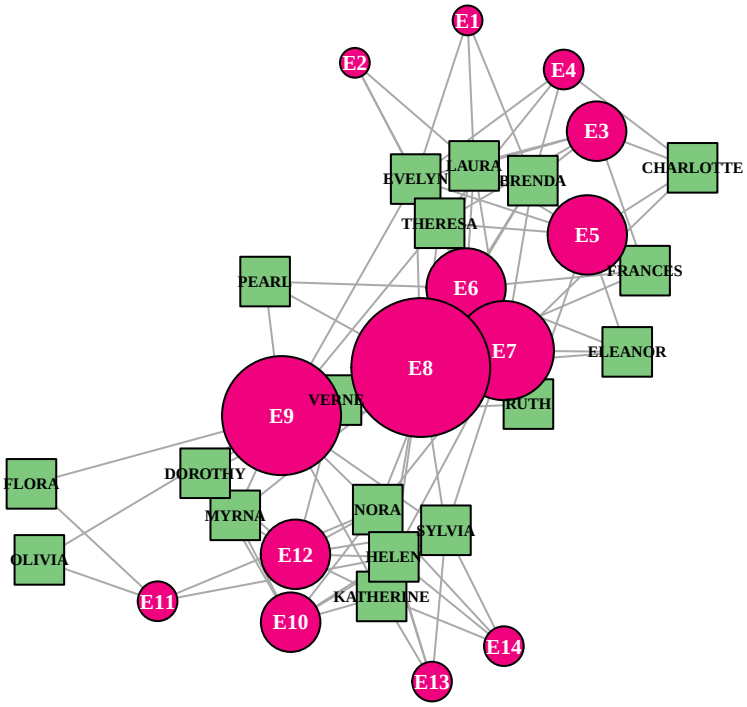
```
##
## Attaching package: 'igraph'
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
## The following object is masked from 'package:base':
##
##     union
```

```
SWnet<-graph_from_incidence_matrix(SWrawdata)

par(mar=c(0,0,0,0))
plot(SWnet)
```

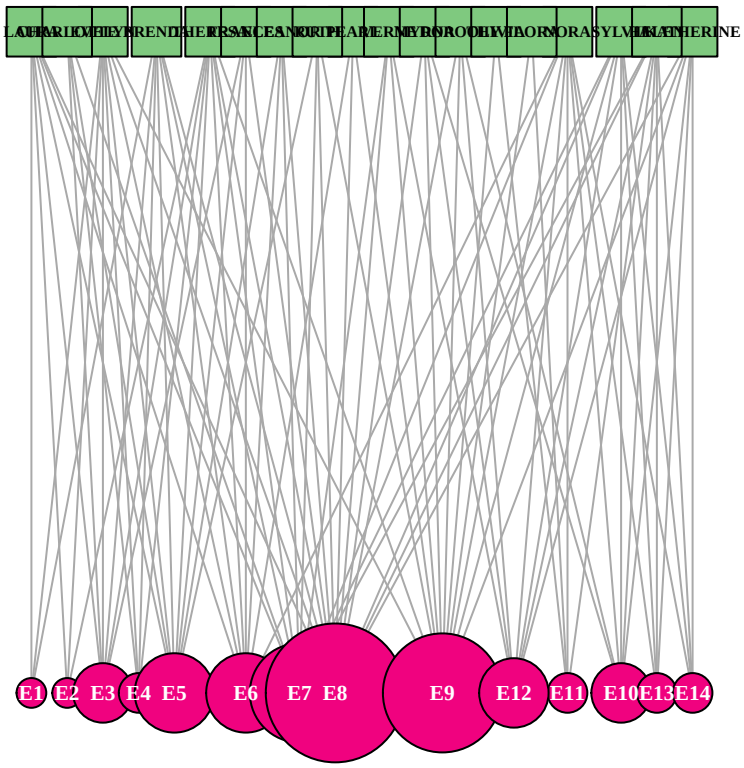


```
library(RColorBrewer)
colors<-brewer.pal(8, 'Accent')
```

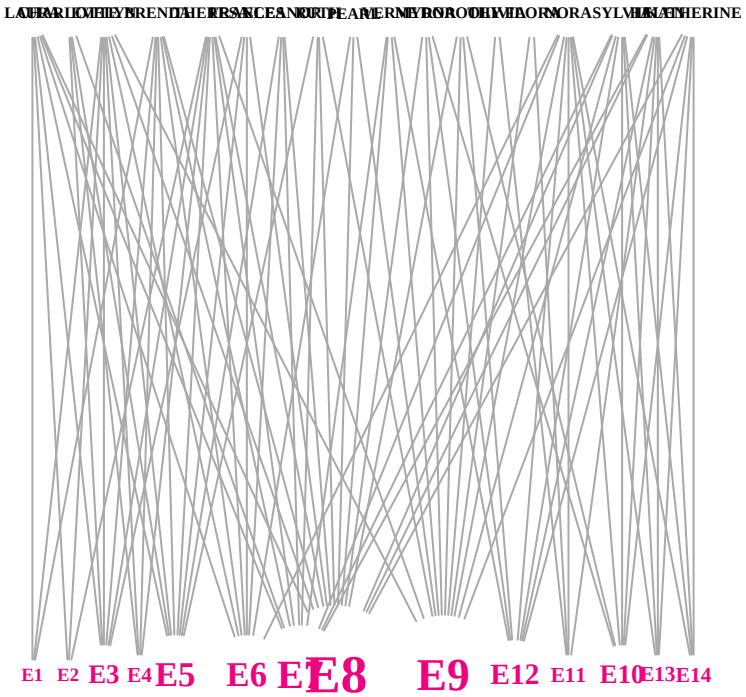
```
V(SWnet)$color <- c(colors[1],colors[6])[V(SWnet)$type+1]
V(SWnet)$shape <- c("square", "circle")[V(SWnet)$type+1]
V(SWnet)$label.color<-c("black","white")[V(SWnet)$type+1]
V(SWnet)$label.cex<-c(0.5, 0.7)[V(SWnet)$type+1]
V(SWnet)$label.font=2

V(SWnet)$indegree <- degree(SWnet, mode = "in")

V(SWnet)$size<-ifelse(V(SWnet)$type==TRUE,V(SWnet)$indegree*3,15)

par(mar=c(0,0,0,0))
plot(SWnet)
```




```
par(mar=c(0,0,0,0))
plot(SWnet, layout=layout.bipartite)
```

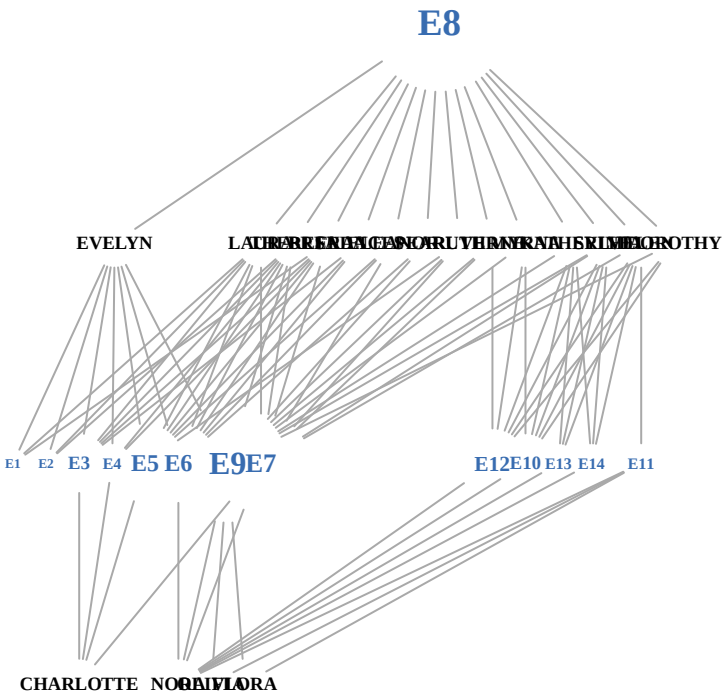



```
V(SWnet)$shape='none'
V(SWnet)$label.color<-c("black",colors[6])[V(SWnet)$type+1]
V(SWnet)$label.cex<-ifelse(V(SWnet)$type==TRUE,0.25+V(SWnet)$indegree/10,0.5)

par(mar=c(0,0,0,0))
plot(SWnet, layout=layout.bipartite)
```



Assignment task. Apparently, creating such graph is not as easy as it looks! You give it a try. Using the plotting function in igraph, improve my graph for the SouthernWomen data by changing at least three characteristics of the graph.

```
V(SWnet)$shape='none'
V(SWnet)$label.color<-c("black",colors[5])[V(SWnet)$type+1]
V(SWnet)$label.cex<-ifelse(V(SWnet)$type==TRUE,.2+V(SWnet)$indegree/15,0.6)

par(mar=c(0,0,0,0))
plot(SWnet, layout=layout_as_tree)
```



I have tried to depict the graph in some other and better way. I have changed the type of layout (the chosen type allows to cluster the data in some way), the color of labels and the size of it, so is is more readable. Unfortunatly, I was't able to change the labels of nodes for Women, so that they would be more visible.

```
Women.only<-as.matrix(SWrawdata)%*%t(as.matrix(SWrawdata)) # Matrix of women only
Events.only<-t(as.matrix(SWrawdata))%*%as.matrix(SWrawdata) # Matrix of events only

Women.only
```

```
##           EVELYN LAURA THERESA BRENDA CHARLOTTE FRANCES ELEANOR PEARL RUTH
## EVELYN         8     6       7      6         3       4       3     3    3
## LAURA          6     7       6      6         3       4       4     2    3
## THERESA        7     6       8      6         4       4       4     3    4
## BRENDA         6     6       6      7         4       4       4     2    3
## CHARLOTTE      3     3       4      4         4       2       2     0    2
## FRANCES        4     4       4      4         2       4       3     2    2
## ELEANOR        3     4       4      4         2       3       4     2    3
## PEARL          3     2       3      2         0       2       2     3    2
## RUTH           3     3       4      3         2       2       3     2    4
## VERNE          2     2       3      2         1       1       2     2    3
## MYRNA          2     1       2      1         0       1       1     2    2
## KATHERINE      2     1       2      1         0       1       1     2    2
## SYLVIA         2     2       3      2         1       1       2     2    3
## NORA           2     2       3      2         1       1       2     2    2
## HELEN          1     2       2      2         1       1       2     1    2
## DOROTHY        2     1       2      1         0       1       1     2    2
```

```
## OLIVIA      1     0      1     0       0      0      0     1     1
## FLORA       1     0      1     0       0      0      0     1     1
##       VERNE MYRNA KATHERINE SYLVIA NORA HELEN DOROTHY OLIVIA FLORA
## EVELYN      2     2      2     2     2     1      2      1     1
## LAURA       2     1      1     2     2     2      1      0     0
## THERESA     3     2      2     3     3     2      2      1     1
## BRENDA      2     1      1     2     2     2      1      0     0
## CHARLOTTE   1     0      0     1     1     1      0      0     0
## FRANCES     1     1      1     1     1     1      1      0     0
## ELEANOR     2     1      1     2     2     2      1      0     0
## PEARL       2     2      2     2     2     1      2      1     1
## RUTH        3     2      2     3     2     2      2      1     1
## VERNE       4     3      3     4     3     3      3      1     1
## MYRNA       3     4      4     4     3     3      4      1     1
## KATHERINE   3     4      6     6     5     5      4      1     1
## SYLVIA      4     4      6     7     6     6      4      1     1
## NORA        3     3      5     6     8     6      3      2     2
## HELEN       3     3      5     6     6     7      3      1     1
## DOROTHY     3     4      4     4     3     3      4      1     1
## OLIVIA      1     1      1     1     2     1      1      2     2
## FLORA       1     1      1     1     2     1      1      2     2
```

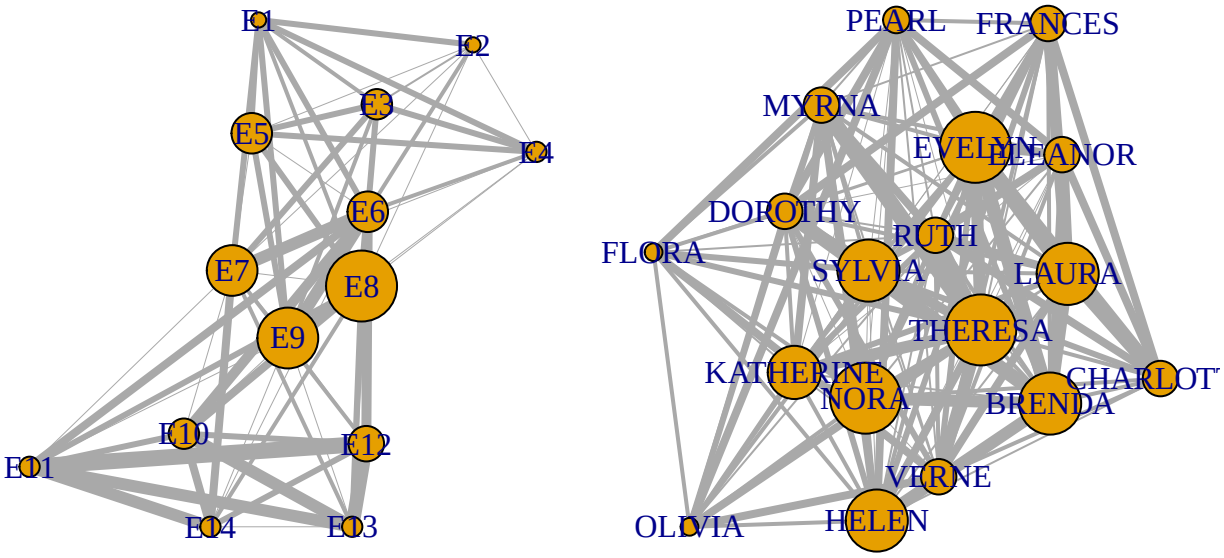
```
Events.only
```

```
##     E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11 E12 E13 E14
## E1   3  2  3  2  3  3  2  3  1   0   0   0   0   0
## E2   2  3  3  2  3  3  2  3  2   0   0   0   0   0
## E3   3  3  6  4  6  5  4  5  2   0   0   0   0   0
## E4   2  2  4  4  4  3  3  3  2   0   0   0   0   0
## E5   3  3  6  4  8  6  6  7  3   0   0   0   0   0
## E6   3  3  5  3  6  8  5  7  4   1   1   1   1   1
## E7   2  2  4  3  6  5 10  8  5   3   2   4   3   3
## E8   3  3  5  3  7  7  8 14  9   5   1   6   3   3
## E9   1  2  2  2  3  4  5  9 12   5   3   6   3   3
## E10  0  0  0  0  0  1  3  5  5   6   2   6   4   4
## E11  0  0  0  0  0  1  2  1  3   2   4   2   2   2
## E12  0  0  0  0  0  1  4  6  6   6   2   7   4   4
## E13  0  0  0  0  0  1  3  3  3   4   2   4   4   4
## E14  0  0  0  0  0  1  3  3  3   4   2   4   4   4
```

Assignment task. Display the contents of the Women.only and Events.only matrices we created above. ??? Explain what data in these matrices mean. ??? What are the benefits and problems with separating incidence matrices and creating adjacency matrices out of them?

```
Women.only1 <-graph.adjacency(Women.only,
mode=c("undirected"),
```

```
weighted=TRUE,
diag=FALSE)

Events.only1 <-graph.adjacency(Events.only,
mode=c("undirected"),
weighted=TRUE,
diag=FALSE)

par(mar=c(0,0,0,0), mfrow=c(1,2))
plot(Events.only1,
     edge.width=Events.only,
     vertex.size=diag(Events.only)*2)
plot(Women.only1,
     edge.width=Women.only,
     vertex.size=diag(Women.only)*3.5)
```



The data in those matrices indeicate the following: in Women.only - how many parties are have been visited by each woman (the size of the vertex on the graph Women.only1), and how many times women have visited the same party for each couple of women (the wildth of each edge in Women.only1).

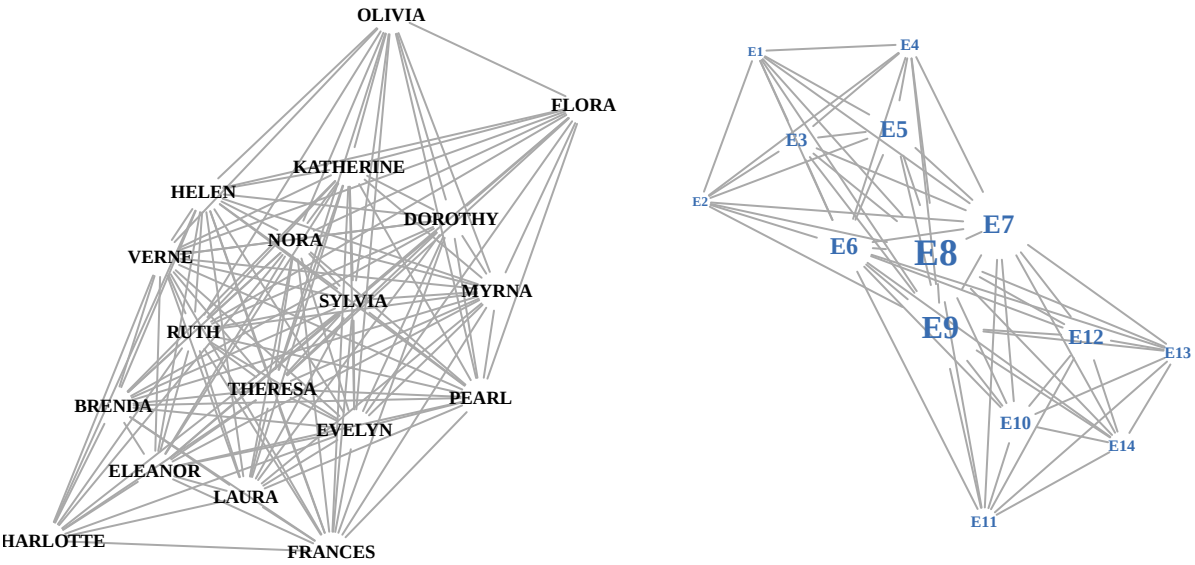
In Events.only - how many women from the poll have visited the party (the size of the vertex on the graph Events.only1), and how many the same women have visited the same party for each couple (the wildth of each edge in Events.only1).

The benefits with separating incidence matrices and creating adjacency matrices is that it is not a difficult way of researching the ties between the nodes of one type with relation to the other. For instance, I have accomplished to gather information on how many parties have been visited by each woman and how mant party they have visited together. So, I can tell which women attend parties more often and who they have met. The same is for the parties

also. Also, by creating adjacency matrices it possible to look on the data from another anlge and to conduct a deeper research.

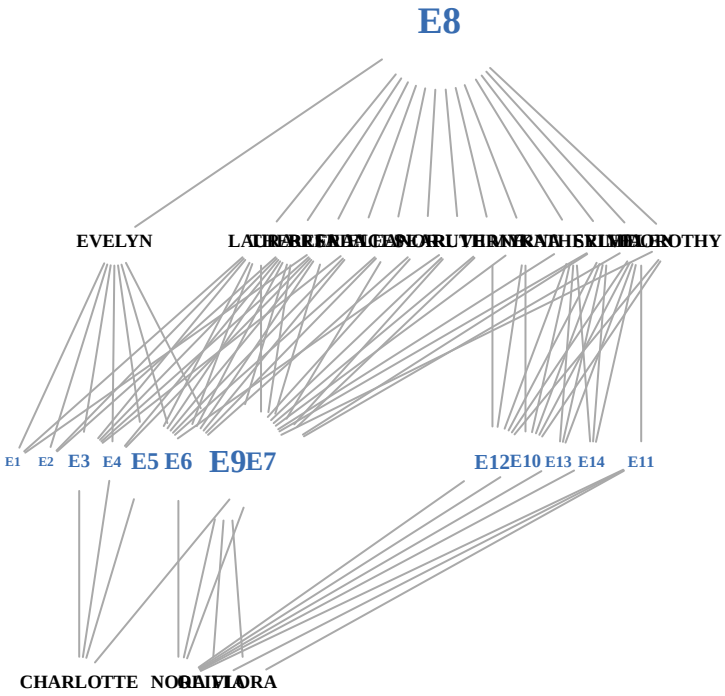
As for the problems of such action, it have to understood that such method can't always be applicable to the data. In the case with women and parties, it is possible to interpret and analyze the adjacency matrices we created. But it is not such an universal rule. Not all cases could be examined with that method. The other problem is that still it isn't allways possible to examine the data on a more individual level.

```
SWnet.sep<-bipartite.projection(SWnet)

par(mar=c(0,0,0,0), mfrow=c(1,2))
plot(SWnet.sep$proj1)
plot(SWnet.sep$proj2)
```



Assignment task. For the network SouthernWomen, please calculate the following network characteristics and briefl explain what they mean: ??? Indegree, outdegree, total degree; ??? Centrality: degree, betweenness, closeness, eigenvector, page rank; correlations between these measures. ??? Transitivity.

```
V(SWnet)$shape='none'
V(SWnet)$label.color<-c("black",colors[5])[V(SWnet)$type+1]
V(SWnet)$label.cex<-ifelse(V(SWnet)$type==TRUE,.2+V(SWnet)$indegree/15,0.6)

par(mar=c(0,0,0,0))
plot(SWnet, layout=layout_as_tree)
```




```
degree <- centralization.degree(SWnet, normalized = TRUE)

betweenness <- centralization.betweenness(SWnet)

closeness <- centralization.closeness(SWnet)

evcent <- centralization.evcent(SWnet)

PR <- page_rank(SWnet)
```

```
degree
```

```
## $res
##  [1]  8  7  8  7  4  4  4  3  4  4  4  6  7  8  7  4  2  2  3  3  6  4  8
## [24]  8 10 14 12  6  4  7  4  4
##
## $centralization
## [1] 0.2641129
##
## $theoretical_max
## [1] 992
```

```
betweenness
```

```
## $res
##  [1]  42.7599983  22.8565396  38.7392641  22.0119104   4.7279418
##  [6]   4.7516238   4.1356871   2.9762664   7.3608922   6.3676024
## [11]   5.9435242  16.2889210  25.2987307  43.9377835  30.7265229
```

```
## [16]   5.9435242   2.0866336   2.0866336   0.9737485   0.9440910
## [21]   8.1977508   3.4529893  16.9812110  28.0103112  58.0969236
## [26] 108.2616581  96.2294880   6.8185655   9.0194398  10.2353650
## [31]   1.8892292   1.8892292
##
## $centralization
## [1] 0.1959329
##
## $theoretical_max
## [1] 14415
```

closeness

```
## $res
##  [1] 0.5166667 0.4696970 0.5166667 0.4696970 0.3875000 0.4305556 0.4305556
##  [8] 0.4305556 0.4558824 0.4558824 0.4428571 0.4696970 0.5000000 0.5166667
## [15] 0.5000000 0.4428571 0.3780488 0.3780488 0.3690476 0.3690476 0.3974359
## [22] 0.3780488 0.4189189 0.4843750 0.5166667 0.5961538 0.5535714 0.3974359
## [29] 0.3780488 0.4078947 0.3780488 0.3780488
##
## $centralization
## [1] 0.3189281
##
## $theoretical_max
## [1] 15.2459
```

evcent

```
## $vector
##  [1] 0.6224572 0.5734522 0.6934195 0.5801231 0.3080944 0.3871692 0.4287008
##  [8] 0.3452842 0.4507862 0.4360231 0.3891268 0.4795512 0.5881534 0.5598766
## [15] 0.5057503 0.3891268 0.1393875 0.1393875 0.2585530 0.2750466 0.4607160
## [22] 0.3208698 0.5887508 0.6100461 0.7460019 1.0000000 0.7617536 0.4238656
## [29] 0.1957167 0.4873414 0.3105682 0.3105682
##
## $value
## [1] 6.869124
##
## $options
## $options$bmat
## [1] "I"
##
## $options$n
## [1] 32
##
## $options$which
```

```
## [1] "LA"
##
## $options$nev
## [1] 1
##
## $options$tol
## [1] 0
##
## $options$ncv
## [1] 0
##
## $options$ldv
## [1] 0
##
## $options$ishift
## [1] 1
##
## $options$maxiter
## [1] 1000
##
## $options$nb
## [1] 1
##
## $options$mode
## [1] 1
##
## $options$start
## [1] 1
##
## $options$sigma
## [1] 0
##
## $options$sigmai
## [1] 0
##
## $options$info
## [1] 0
##
## $options$iter
## [1] 2
##
## $options$nconv
## [1] 1
##
## $options$numop
```

```
## [1] 24
##
## $options$numopb
## [1] 0
##
## $options$numreo
## [1] 16
##
##
## $centralization
## [1] 0.5744777
##
## $theoretical_max
## [1] 30
```

```
PR
```

```
## $vector
##    EVELYN      LAURA    THERESA     BRENDA  CHARLOTTE    FRANCES
## 0.04153074 0.03649684 0.04070386 0.03624672 0.02261447 0.02210212
##   ELEANOR      PEARL       RUTH      VERNE      MYRNA  KATHERINE
## 0.02185696 0.01771749 0.02197103 0.02207701 0.02237325 0.03198349
##    SYLVIA       NORA      HELEN    DOROTHY     OLIVIA      FLORA
## 0.03623771 0.04198485 0.03732761 0.02237325 0.01473509 0.01473509
##         E1         E2         E3         E4         E5         E6
## 0.01793329 0.01785668 0.03176035 0.02263189 0.04107380 0.04108022
##         E7         E8         E9        E10        E11        E12
## 0.05004964 0.06895210 0.06323072 0.03212095 0.02620585 0.03681231
##        E13        E14
## 0.02261232 0.02261232
##
## $value
## [1] 1
##
## $options
## NULL
```

```
transitivity(SWnet)
```

```
## [1] 0
```

Based on the calculated above measures we can make the following conclusions: the gotten bipartiate network is not that tight and complex, it is far from all nodes being connected to all others. But the closeness and betweenness rates tells us, that nodes are not that close to each other based on geodesic distance.