

Stats218_HW1

Zihan Lin

2024-10-04

```
library(networkdata)
library(sna)
library(network)
library(tidyverse)
```

Question 1

```
data(butland_ppi)
data(addhealth9)
data(tribes)
```

a)

```
only_boy <- which(addhealth9$X[, "female"] == 0)
boys_match <- addhealth9$E[addhealth9$E[, 1] %in% only_boy & addhealth9$E[, 2] %in% only_boy,]
health_ntwk <- network(boys_match, directed = FALSE)
print(health_ntwk)
```

```
## Network attributes:
##   vertices = 254
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 273
##     missing edges= 0
##     non-missing edges= 273
##
## Vertex attribute names:
##   vertex.names
##
## No edge attributes
```

```
pos_tribe <- tribes[, "pos"]
tribe_ntwk <- network(pos_tribe, directed = FALSE)
print(tribe_ntwk)
```

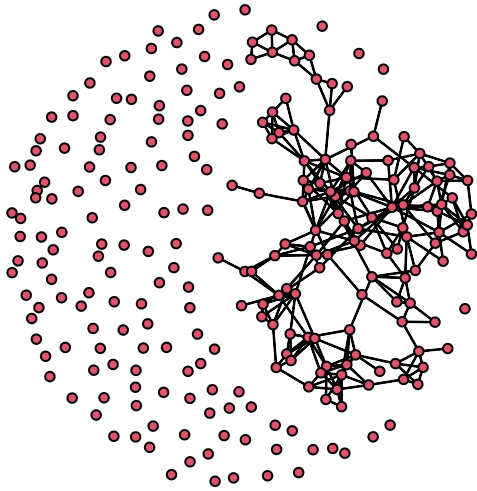
```
## Network attributes:
##   vertices = 16
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 29
##     missing edges= 0
##     non-missing edges= 29
##
## Vertex attribute names:
##   vertex.names
##
## No edge attributes
```

```
bl_ntwk <- network(butland_ppi, directed = FALSE)
print(bl_ntwk)
```

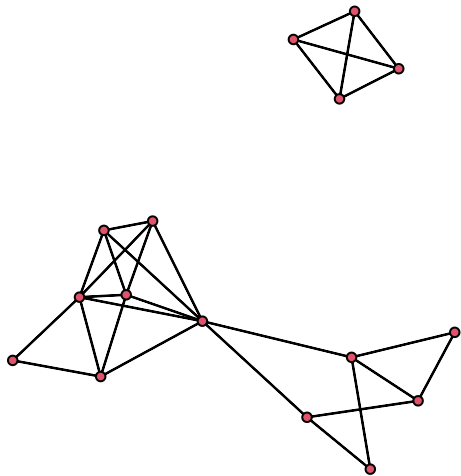
```
## Network attributes:
##   vertices = 270
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 716
##     missing edges= 0
##     non-missing edges= 716
##
## Vertex attribute names:
##   vertex.names
##
## No edge attributes
```

b)

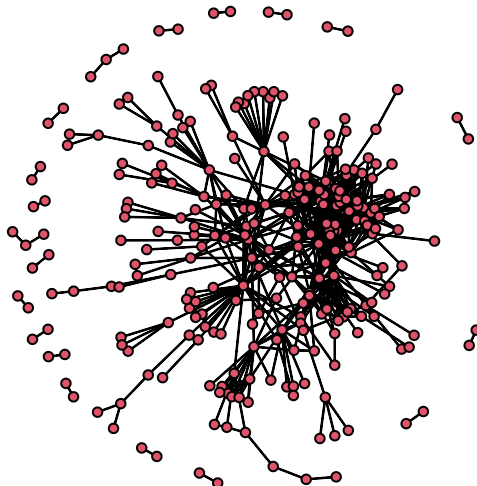
```
plot(health_ntwk)
```



```
plot(tribe_ntwk)
```



```
plot(bl_ntwk)
```



c)

```
health_degree <- degree(health_ntwk, gmode = "graph") %>% table()
tribe_degree <- degree(tribe_ntwk, gmode = "graph") %>% table()
bl_degree <- degree(bl_ntwk, gmode = "graph") %>% table()
health_degree
```

```
## .
##   0   1   2   3   4   5   6   7   8   9  10  11  13
## 136   8  10  16  29  22  10  11   8   1   1   1   1
```

```
tribe_degree
```

```
## .
## 2 3 4 5 6 7
## 3 6 4 1 1 1
```

```
bl_degree
```

```
## .
##   1   2   3   4   5   6   7   8  10  11  12  13  14  15  16  17  18  19  20  22
## 105  39  19  13  16  10  14   6   3   2   3   9   3   5   4   1   2   3   1   2
##  23  27  28  30  32  33  34  36
##   2   1   2   1   1   1   1   1
```

d)

```
# summary  
summary(health_degree)
```

```
## Number of cases in table: 254  
## Number of factors: 1
```

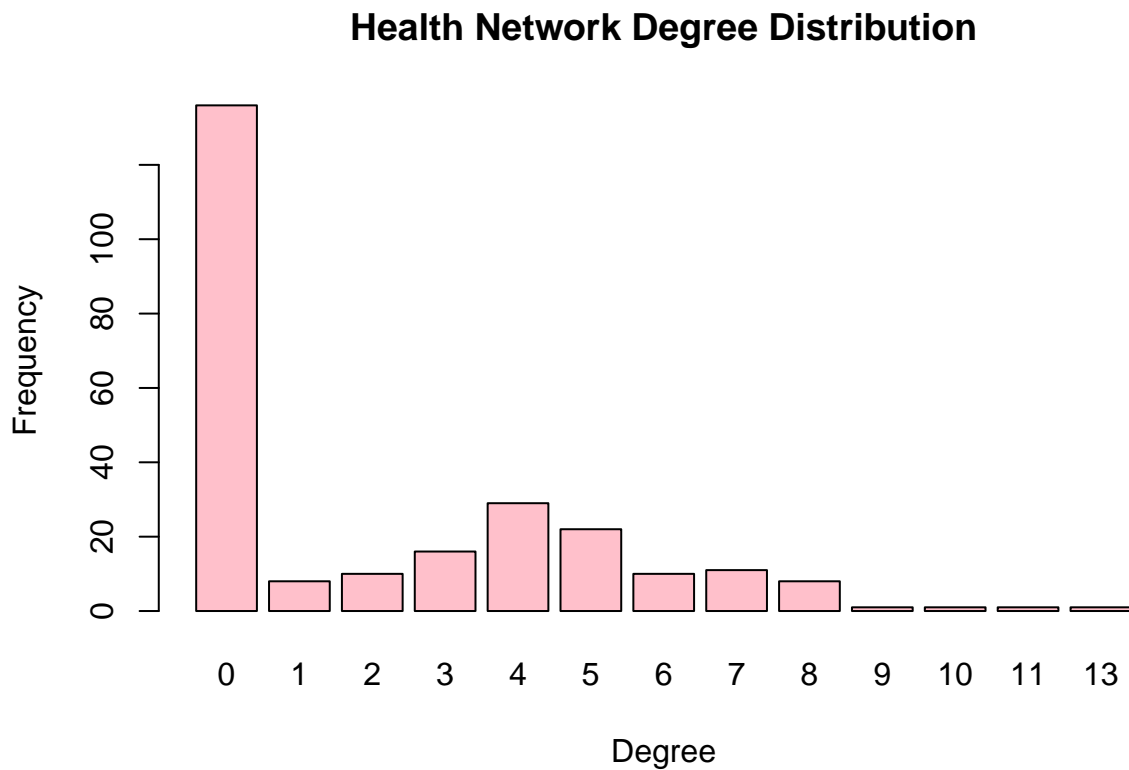
```
summary(tribe_degree)
```

```
## Number of cases in table: 16  
## Number of factors: 1
```

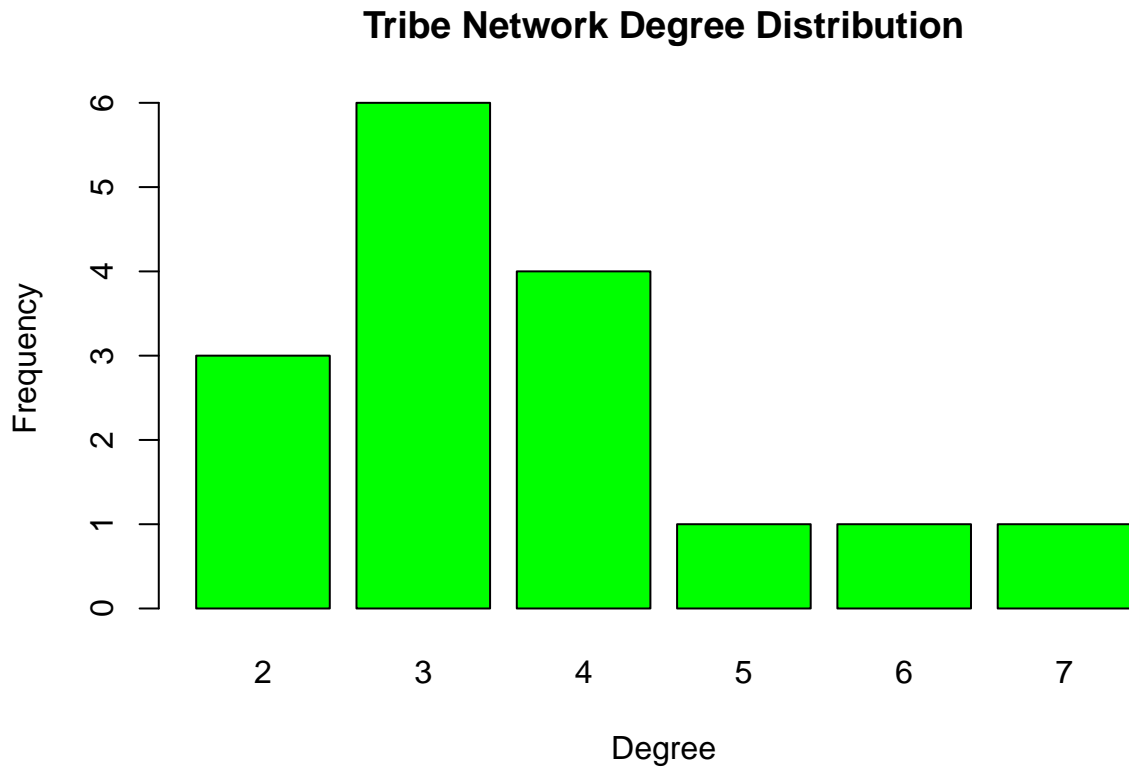
```
summary(bl_degree)
```

```
## Number of cases in table: 270  
## Number of factors: 1
```

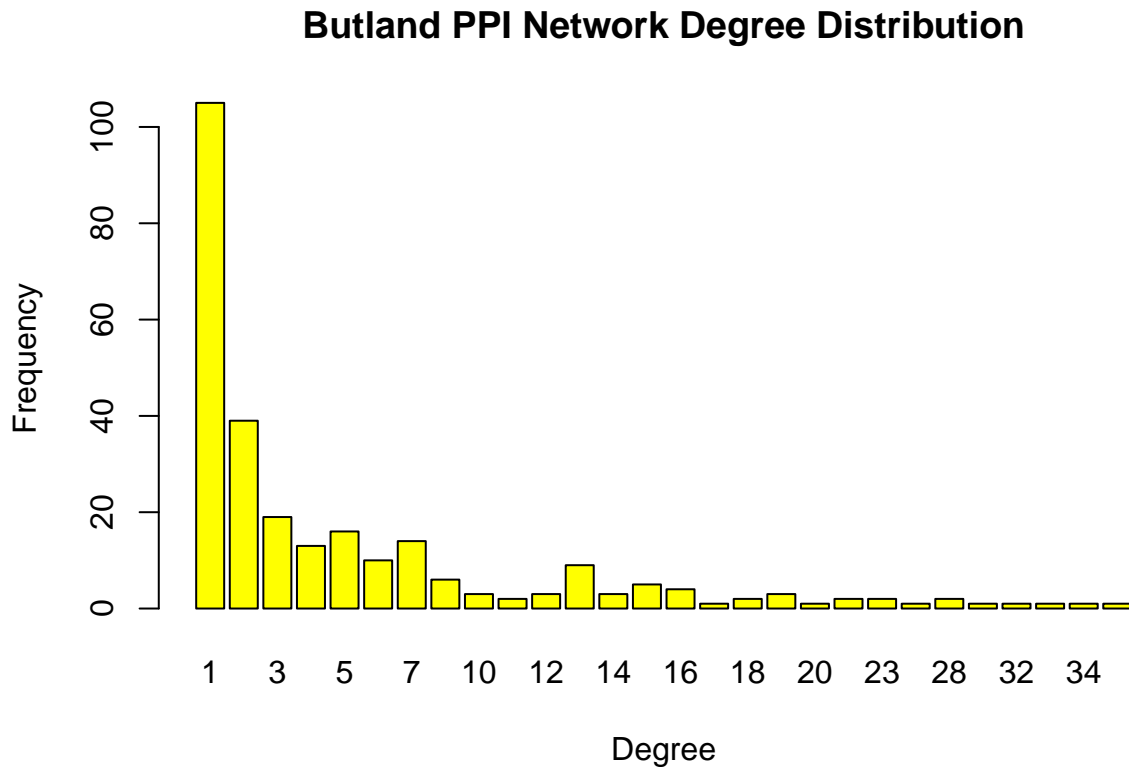
```
# barplot  
barplot(health_degree, main = "Health Network Degree Distribution",  
        xlab = "Degree", ylab = "Frequency", col = "pink")
```



```
barplot(tribe_degree, main = "Tribe Network Degree Distribution",
        xlab = "Degree", ylab = "Frequency", col = "green")
```



```
barplot(bl_degree, main = "Butland PPI Network Degree Distribution",
        xlab = "Degree", ylab = "Frequency", col = "yellow")
```

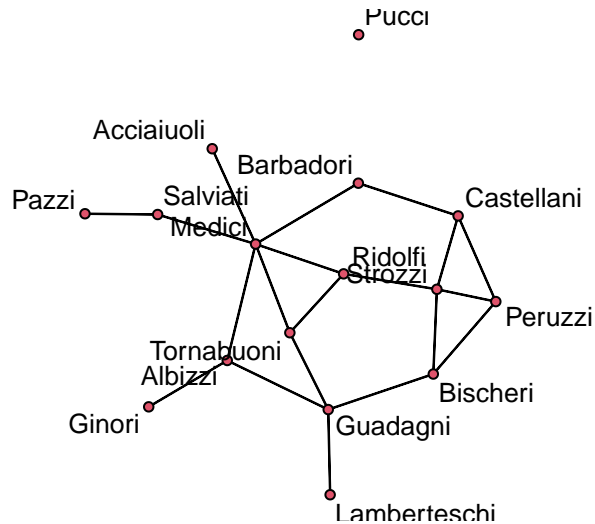


Question 2

```
data(florentine)
```

a)

```
network(flomarriage) %>%  
  plot(displaylabels = TRUE,  
        label = network.vertex.names(flomarriage),  
        label.cex = 0.8)
```



b)

```
fm_degree <- degree(network(flomarriage), gmode = "graph")  
  
cat("Familiy has the lowest degree: ",  
    network.vertex.names(flomarriage)[fm_degree == min(fm_degree)],  
    ". Its degree is ", min(fm_degree), ". \n")  
  
## Familiy has the lowest degree: Pucci . Its degree is 0 .  
  
cat("Familiy has the highest degree: ",  
    network.vertex.names(flomarriage)[fm_degree == max(fm_degree)],  
    ". Its degree is ", max(fm_degree), ". \n")  
  
## Familiy has the highest degree: Medici . Its degree is 6 .
```

c)

```
fm_degree
```

```
## [1] 1 3 2 3 3 1 4 1 6 1 3 0 3 2 4 3
```

d)

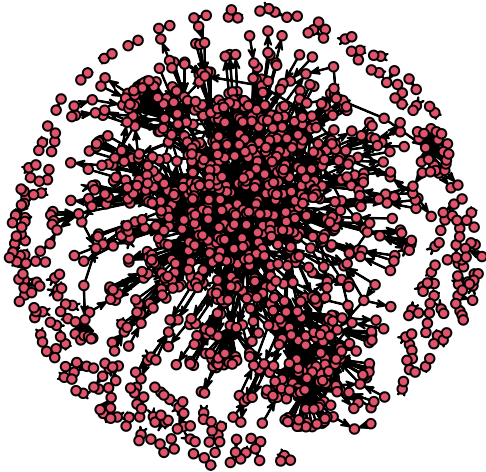
```
largest_fm <- which.max(component.dist(network(flomarriage))$csize)
largest_component <- component.dist(network(flomarriage))$membership == largest_fm
excluded <- network.vertex.names(flomarriage)[!largest_component]
cat("Family excluded from the largest component in the Florentine Marriage Network: ", excluded)
```

```
## Family excluded from the largest component in the Florentine Marriage Network: Pucci
```


Question 3

a)

```
ccsb <- read.table("CCSB-Y2H.txt")  
ccsb <- ccsb[ccsb[, 1] != ccsb[, 2], ]  
ccsb_network <- network(ccsb[, 1:2], directed = TRUE)  
plot(ccsb_network)
```



b)

```
out_degrees <- degree(ccsb_network, cmode = "outdegree")
in_degrees <- degree(ccsb_network, cmode = "indegree")

cat("Correlation between in-degree and out-degree:", cor(out_degrees, in_degrees), "\n")

## Correlation between in-degree and out-degree: -0.07393897
cat("Summary of total degrees:\n")

## Summary of total degrees:
summary(out_degrees + in_degrees)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000   1.000   1.000   2.662   3.000  85.000

barplot(table(out_degrees + in_degrees), main = "Protein Activity Network Degree Distribution",
        xlab = "Degree", ylab = "Frequency", col = "yellow")
```

