Homework 2

Victoria Ignatova, BSC-163

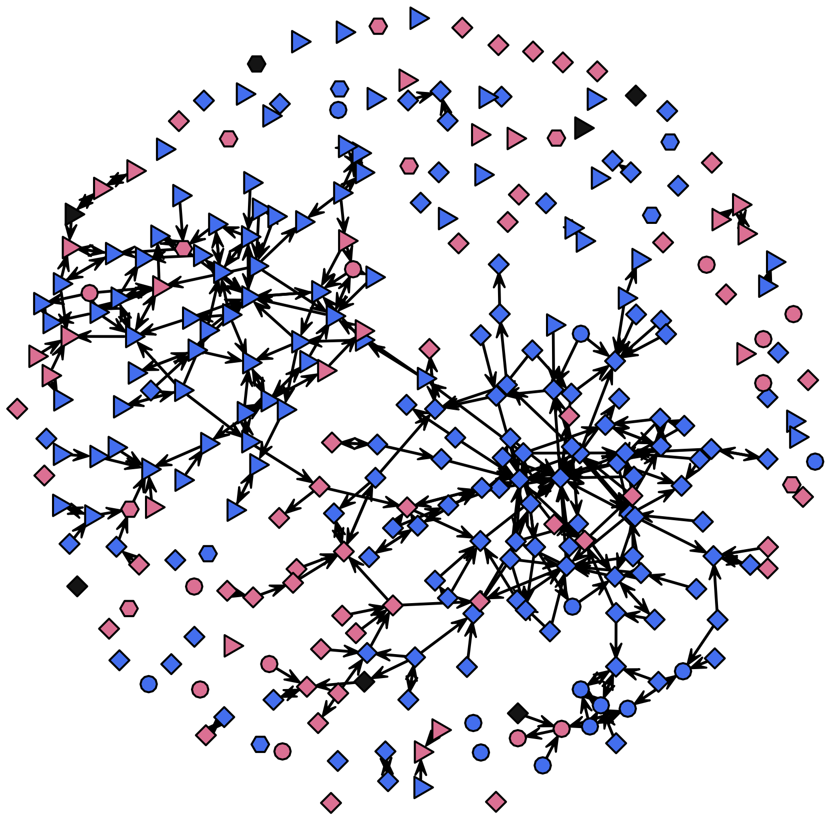
```
library(rmarkdown)
library(sna)
library(network)
```

```
# The drug network
drugpaj <- read.paj('drugnet2.paj')
drug <- drugpaj$networks[[1]]
#attributes
gender<-drugpaj$partitions[[1]]
ethnicity <- drugpaj$partitions[[2]]
```

1. We talked in class that for network analysis, the lines between descriptive and inferential analysis is blurred: most descriptive measures are automatically inferential. Examine the drug network we've drawn in this seminar. Does anything look odd to you?

As we see from the plot, there are two big gropus with network ties established tightly inside them and very few tie between them. There is a gender disbalance - the number of dots colored blue is much higher. There seem to be no gender differences between the two groups, althout it may be hard to make conclusions, as there too few nodes colored pink.

```
set.seed(8)
gender<-drugpaj$partitions[[1]]
ethnicity <- drugpaj$partitions[[2]]
sides<-ifelse(ethnicity==1,12,
              ifelse(ethnicity==2, 3, ifelse(ethnicity==3, 4, 6)))
colors<-ifelse(gender==2,"palevioletred",
               ifelse(gender==1,"royalblue2","gray8"))
par(mar=c(0,0,0,0))
plot(drug, vertex.col=colors,
     vertex.sides=sides, vertex.cex=1)
```



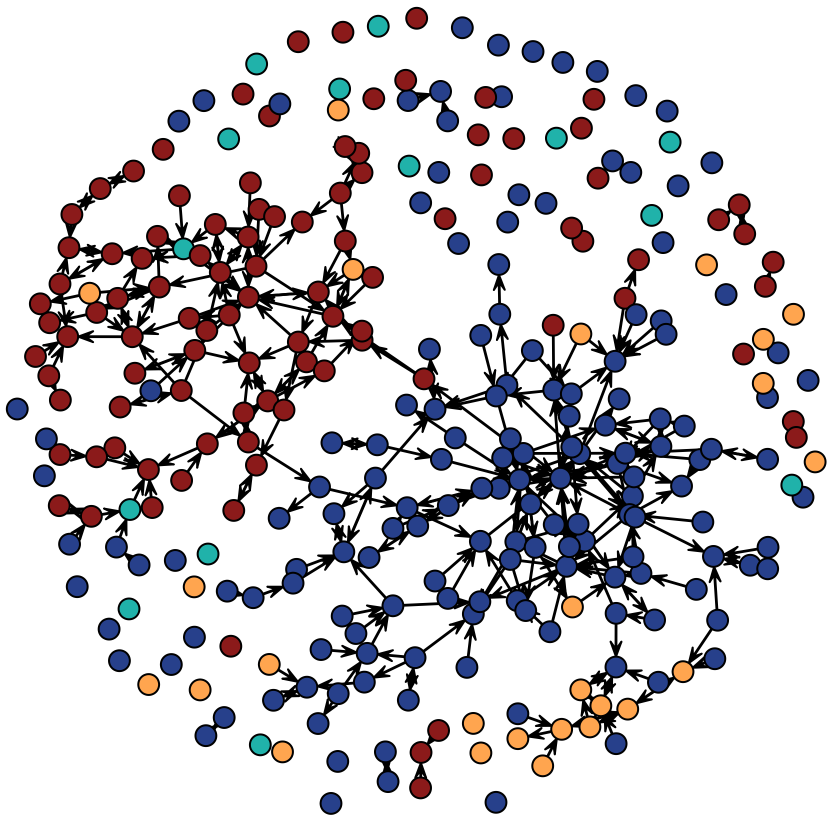
2. *For the drug network, draw the network with nodes colored by race. What do you conclude?*

*There are two obvious major groups of drug users and the race is seems to be a crucial factor of ties establishment.

*Drug users of the same race tend to establish ties within ther race and have or report very weak ties with drug users of another race.

*Drug user with the two other (not dominant) race denomination are weakly included into the network connections, especially the ones of the minority (turquoise color) race.

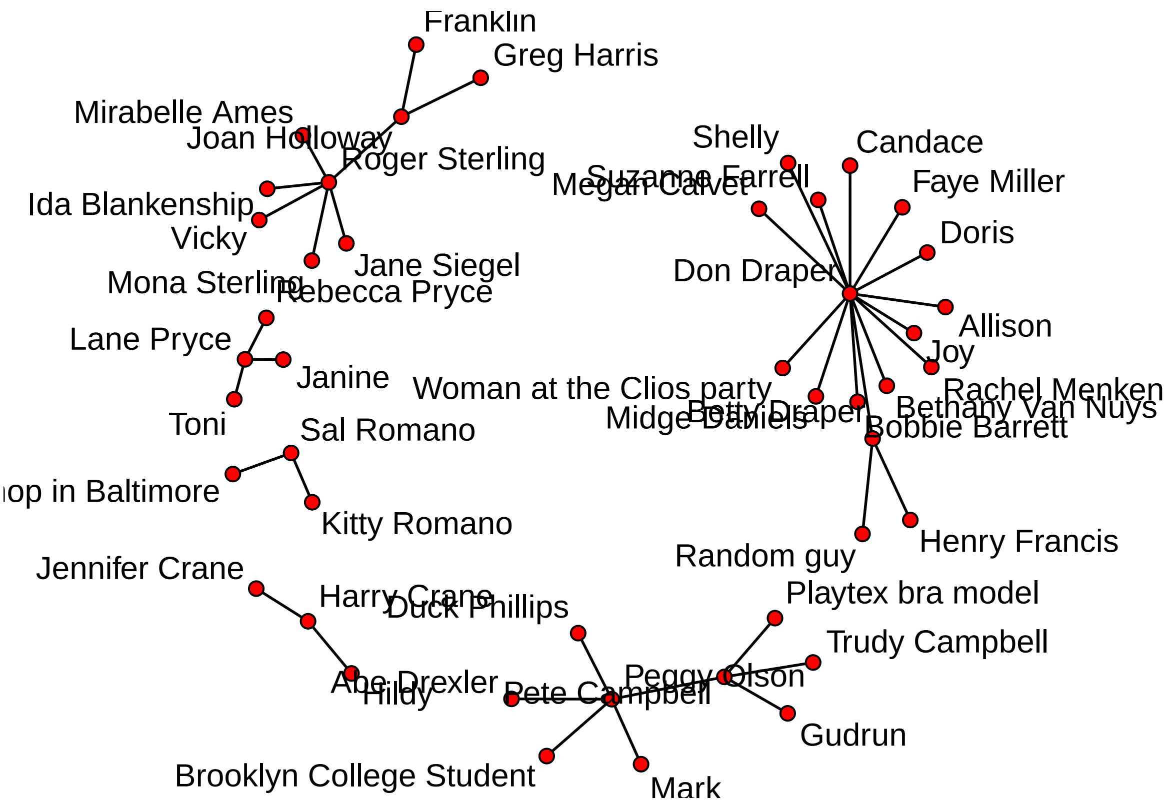
```
set.seed(8)
colors_ <-ifelse(ethnicity==1,"#FFA54F",
                 ifelse(ethnicity==2, "#8B1A1A",
                        ifelse(ethnicity==3, "#27408B", "#20B2AA")))
par(mar=c(0,0,0,0))
plot(drug, vertex.col=colors_, vertex.cex=1.3)
```



```
#Madmen network
load('madmen.Rdata')
mad.net <- as.network(mad.matrix,
                      directed=FALSE)
set.vertex.attribute(mad.net,
                     attrname='female',
                     value=mad.att[,2])
```

3. Please examine the options in the "network.layout" command and perform the following on the madmen data: * Create the madmen.net with labels.

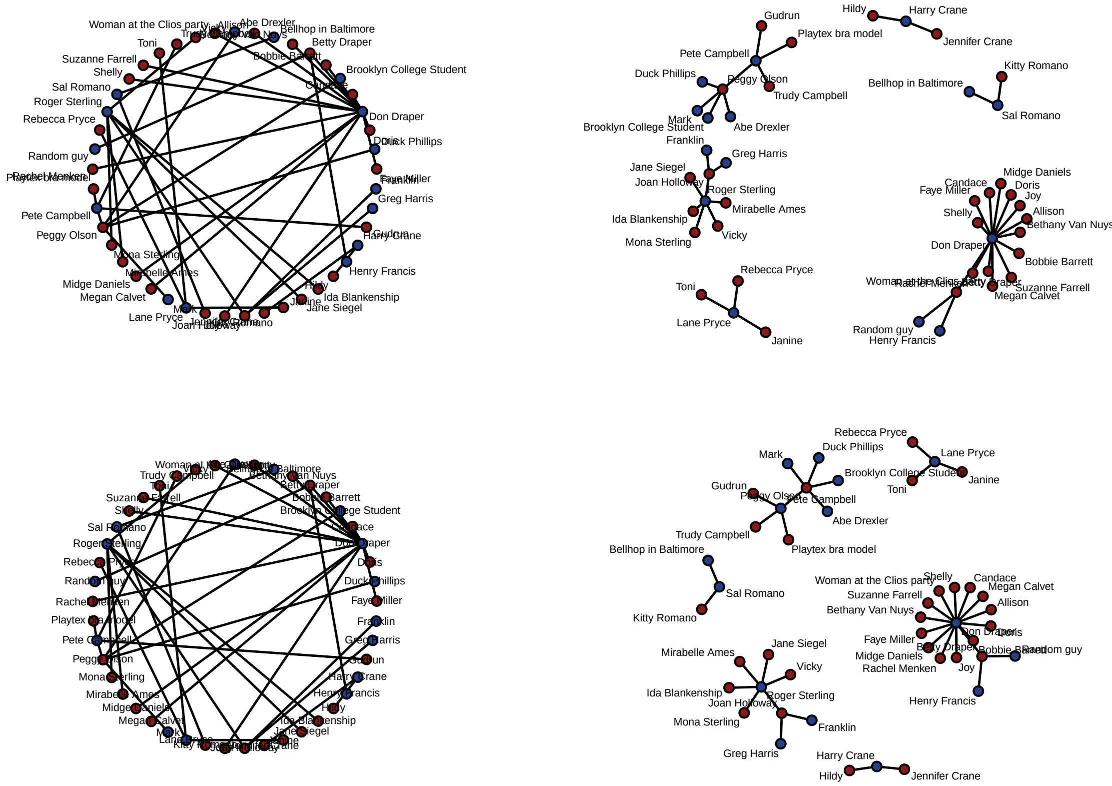
```
set.seed(8)
par(mar=c(1,1,1,1))
plot(mad.net, displaylabels = TRUE)
```



- Experiment with options by adding attributes, changing vertex or edge colors, finding the best position for labels. While this task may take a while, it will count as complete if you generate at least one graph that is different from the graphs I've shown you in this assignment

```
set.seed(10)
colors_gender <- ifelse(mad.att$Female==1, "#8B1A1A", "#27408B")
par(mar=c(1,1,1,1), mfrow=c(2,2))
plot(mad.net,
     displaylabels = TRUE,
     label.cex=0.4,
     label.pos=0,
     vertex.col = colors_gender,
     vertex.cex = 1.4,
     mode = 'circle')

plot(mad.net,
     displaylabels = TRUE,
     label.cex=0.4,
     label.pos=0,
     vertex.col = colors_gender,
     vertex.cex = 1.4,
     mode = 'fruchtermanreingold')
plot(mad.net,
```

```r
     displaylabels = TRUE,
     label.cex=0.4,
     label.pos=5,
     vertex.col = colors_gender,
     vertex.cex = 1.4,
     mode = 'circle')
plot(mad.net,
     displaylabels = TRUE,
     label.cex=0.4,
     label.pos=0,
     vertex.col = colors_gender,
     vertex.cex = 1.4,
     mode = 'kamadakawai')
```




```r
detach(package:sna)
detach(package:network)
library(igraph)
```

```
##
## Attaching package: 'igraph'

## The following objects are masked from 'package:stats':
##
##     decompose, spectrum

## The following object is masked from 'package:base':
```

```
##
##     union
```

```r
# Trade network
load('trade.Rdata')
trade.all<-as.matrix(trade.all)
trade.any <- ifelse(trade.all > 0, 1, 0)
trade.2 <- ifelse(trade.all > 1, 1, 0)
trade.max <- ifelse(trade.all == 5, 1, 0)

tradegraph.any <-graph.adjacency(trade.any, mode=c("directed"),
                                 weighted=NULL, diag=FALSE)
tradegraph.2 <-graph.adjacency(trade.2, mode=c("directed"),
                               weighted=NULL, diag=FALSE)
tradegraph.5 <-graph.adjacency(trade.max, mode=c("directed"),
                               weighted=NULL, diag=FALSE)
```

4. Think about the manipulations we are performing on the trade network. With respect to dichotomizing, please answer the following:

- How would you justify any of these choices? Please refer to specific social theories to make your answer more legitimate.
- What are the impirical implication of these choices?
- What differences do you observe between the graphs where the cutpoint is any tie, at least two ties, and all ties present?
- What information can you gather from these observed differences to help you expand on your earlier theoretical justification of ties? Alternatively, does your theoretical justification seem reasonable in light of new information obtained from these graphs?

Trade.any shows whether there are any ties or not (binary). This reduction can be used for geodesic distance measure counts, that can provide information about the importance of node for information transition over the network, it's remoteness from the power center and other.

Trade.2 can be used to exclude cases in which the presence of ties is reported, but the strength of the tie is really weak compared to other ties. To exclude ties, that are not distinguishable from null from the perspective of influence, but the respondents cannot report them as they remember they once had them.

Trade.max shows the strengthest ties. It keeps the strengthest directed ties we have in a trade network. While the visualizations for the first two cases looked like a lot of a) not NULL b)mutual ties are present, the trade.max reveals, that there are some countriesm that have strong mutual ties with only one partner and NULL with others (Madagascar, Liberia, Ecuador, Honduras). These cooountries are far from beeing the richest and we can assume. that they are really reliant on United states as their only strong-tie partne. Another conclusion we can make is that countries of Europe and America are much more included into the net than others.

5. There are several networks in the "trade.Rdata" file, described above. We have fully explored the "trade.all" network. Now, select one of the individual trade networks (manufacture, food, *crude, etc.) and show me everything you've learned in this class so far.*

```
library(sna)
```

```
## Loading required package: network

## network: Classes for Relational Data
## Version 1.16.0 created on 2019-11-30.
## copyright (c) 2005, Carter T. Butts, University of California-Irvine
##                     Mark S. Handcock, University of California -- Los Angeles
##                     David R. Hunter, Penn State University
##                     Martina Morris, University of Washington
##                     Skye Bender-deMoll, University of Washington
##  For citation information, type citation("network").
##  Type help("network-package") to get started.

##
## Attaching package: 'network'

## The following objects are masked from 'package:igraph':
##
##     %c%, %s%, add.edges, add.vertices, delete.edges,
##     delete.vertices, get.edge.attribute, get.edges,
##     get.vertex.attribute, is.bipartite, is.directed,
##     list.edge.attributes, list.vertex.attributes,
##     set.edge.attribute, set.vertex.attribute

## sna: Tools for Social Network Analysis
## Version 2.5 created on 2019-12-09.
## copyright (c) 2005, Carter T. Butts, University of California-Irvine
##  For citation information, type citation("sna").
##  Type help(package="sna") to get started.

##
## Attaching package: 'sna'

## The following objects are masked from 'package:igraph':
##
##     betweenness, bonpow, closeness, components, degree,
##     dyad.census, evcent, hierarchy, is.connected, neighborhood,
##     triad.census
```

```
library(network)
detach(package:igraph)
```

```
#food.data <- as.matrix(food)
crude.data <- as.matrix(crude)
crude.net <-as.network(crude.data, directed=TRUE)
```

Have 24 countries that will serve as nodes with ties base on crude trade realtions.

```r
dim(crude.data)
```

```
## [1] 24 24
```

We can see, that Algeria is not connected with Argentina, but Argentina has a tie with Algeria, so our network is directed.

```r
data.frame(crude.net[1:5, 1:5])
```

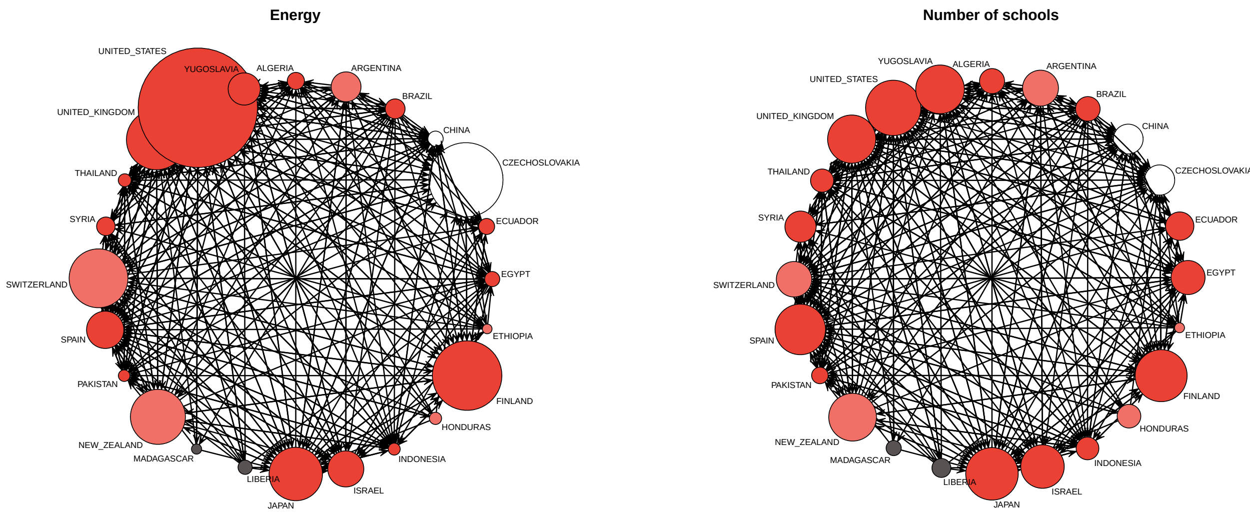
	ALGERIA	ARGENTINA	BRAZIL	CHINA	CZECHOSLOVAKIA
ALGERIA	0	0	0	1	1
ARGENTINA	1	0	1	1	1
BRAZIL	1	1	0	1	1
CHINA	0	1	0	0	1
CZECHOSLOVAKIA	0	0	0	0	0

```r
summary(trade.att)
```

```
##    POP_GROWTH         GNP             SCHOOLS          ENERGY
##  Min.    :0.100   Min.    :-1.900   Min.    : 1.00   Min.    :   24
##  1st Qu.:1.075   1st Qu.: 0.700   1st Qu.:29.75   1st Qu.:  469
##  Median :2.050   Median : 2.200   Median :45.00   Median : 1032
##  Mean    :1.954   Mean    : 2.473   Mean    :50.79   Mean    : 2572
##  3rd Qu.:2.700   3rd Qu.: 4.575   3rd Qu.:81.25   3rd Qu.: 4691
##  Max.    :3.700   Max.    : 5.900   Max.    :97.00   Max.    :11626
##                   NA's    :2
```

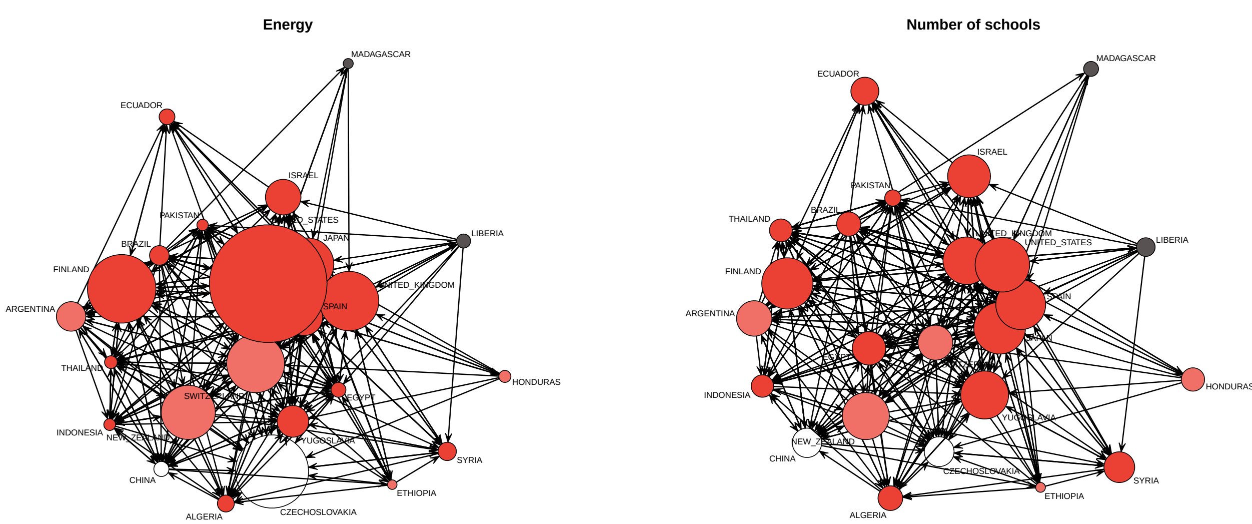
```r
# Extract attributes
pop_growth <- trade.att$POP_GROWTH
gnp <- trade.att$GNP
schools <- trade.att$SCHOOLS
energy <- trade.att$ENERGY

# GNP
gnp_col <- ifelse(gnp <0, '#595252', ifelse(gnp<=1, '#f07067', '#eb4034'))
#'#595252' less than 0
#'#f07067' 0 to 1
#'#eb4034' more than 1

set.seed(8)
par(mar = c(1,1,1,1), mfrow=c(1,2))
plot(crude.net, displaylabels = TRUE, mode = 'circle',
     label.cex=0.7, label.pos=0,
     vertex.cex = energy/1000 + 1, vertex.col = gnp_col)
title(main = 'Energy')
plot(crude.net, displaylabels = TRUE, mode = 'circle',
     label.cex=0.7, label.pos=0,
```

```
     vertex.cex = schools/20 + 1, vertex.col = gnp_col)
title(main = 'Number of schools')
```



And alternative version:

```
set.seed(8)
par(mar = c(1,1,1,1), mfrow=c(1,2))
plot(crude.net, displaylabels = TRUE,
     mode = 'fruchtermanreingold',
     label.cex=0.7, label.pos=0,
     vertex.cex = energy/1000 + 1, vertex.col = gnp_col)
title(main = 'Energy')
plot(crude.net, displaylabels = TRUE,
     mode = 'fruchtermanreingold',
     label.cex=0.7, label.pos=0,
     vertex.cex = schools/20 + 1, vertex.col = gnp_col)
title(main = 'Number of schools')
```



- Countries with highest GDP and Energy production have most of ties.

- GDP is correlated with Enegry (red bubbles are the biggest ones, grey are the smallest) and number of schools, that can also be the measure of development (assumed that the SCHOOL represents the amount of schools per 100 of population or measures like that, as the numbers given are too small to the raw values)
- Liberia and Honduras have very few mutual ties, while UK and USA, for exmple, have mutual connections with the vast majority of partners.