

1. Øvelse:

Intro til netværksanalyse i R

Alexander Gamerdinger

Department of Organization, Copenhagen Business School

E-mail: aga.ioa@cbs.dk

15 februar 2022

R set up

Installering af R og Rstudio

```
# funktion til at vise informationerne om R versionen + andre ting  
version <- R.Version()  
  
# kun R versionen  
version$version.string  
  
## [1] "R version 4.1.2 (2021-11-01)"
```

Sådan tjekker du din Rstudio version: Help -> About Rstudio

- **RStudio 2021.09.0+351**

Folder struktur

```
# funktionen til at lave en "working directory"
setwd("/Users/alexandergamerdinger/Desktop/PhD/teaching/virksomhedsstrategi_forår_2022")

# se hvilken "working directory" du har
getwd()

## [1]
"/Users/alexandergamerdinger/Desktop/PhD/teaching/virksomhedsstrategi_forår_2022"

# se filerne i din "working directory" - "." betyder at alt skal vises
list.files(path = ".")

## [1] "input"      "material"  "output"    "r"
```

Installer af pakker

```
# data manipulation  
install.packages('data.table')  
install.packages('tidyverse')
```

```
# data analysis & visualization  
install.packages('igraph')  
install.packages('ggraph')
```

```
# reading and writing data  
install.packages('readxl')  
install.packages('writexl')
```

```
# data manipulation  
library('data.table')  
library('tidyverse')
```

```
# data analysis & visualization  
library('igraph')  
library('ggraph')
```

```
# reading and writing data  
library('readxl')  
library('writexl')
```

Indlæsning & bearbejdning af data

Load data (data.table)

Først skal i downloade elite netværk datasæt fra canvas eller [her](#) og gem filen under **working_directory_folder/input**

```
#data.table
den <- fread("input/den17-no-nordic-letters.csv")
head(den)
```

##	name	affiliation	role	tags	position_id	id
## 1:	Aage Almtoft	Middelfart Sparekasse				
## 2:	Aage B. Andersen	Foreningen OEstifterne - Repraesentantskab (Medlemmer af delegeretforsamling)				
## 3:	Aage Christensen	AARHUS SOEMANDSHJEM				
## 4:	Aage Dam	Brancheforeningen automatik, tryk & transmission (bestyrelse)				
## 5:	Aage Dam	Dansk Erhverv (bestyrelse)				
## 6:	Aage Frandsen	Dommere valgt af Folketinget (Rigsretten)				
## 1:	Member	Corporation, FINA, Banks, Finance			1	95023

Load data (dplyr)

```
#dplyr
den <- read_csv("input/den17-no-nordic-letters.csv")

## Rows: 56849 Columns: 17

head(den)

## # A tibble: 6 × 17
##   name      affiliation role  tags  position_id      id sector type  description
##   <chr>      <chr>      <chr> <chr>      <dbl>  <dbl> <chr> <chr> <chr>
## 1 Aage Almt... Middelfart... Memb... Corp...         1  95023 Corpo... <NA> Automatisk...
## 2 Aage B. A... Foreningen... Memb... Char...         4  67511 NGO      Orga... Direktoer
## 3 Aage Chri... AARHUS SOE... Chai... Foun...         6 100903 Found... <NA> Automatisk...
## 4 Aage Dam    Branchefor... Chai... Busi...         8  69156 NGO      Orga... Formand, A...
## 5 Aage Dam    Dansk Erhv... Memb... Empl...         9  72204 NGO      Stat  Adm. dir. ...
## 6 Aage Fran... Dommere va... Memb... Judg...        15  73158 Parli... <NA> <NA>
## # ... with 8 more variables: created <dtm>, archived <dtm>,
## #   last_checked <dtm>, cvr_person <dbl>, cvr_affiliation <dbl>,
## #   person_id <dbl>, affiliation_id <dbl>, gender <chr>
```


Data bearbejdning (data.table)

#data.table select funktion

```
den[,.(name, gender)]
```

```
##              name gender
##      1:      Aage Almtoft   Men
##      2:      Aage B. Andersen   Men
##      3:      Aage Christensen   Men
##      4:      Aage Dam   Men
##      5:      Aage Dam   Men
##      ---
## 56845: Jacob Aarup-Andersen 195767   Men
## 56846:      Carsten Rasch Egeriis   Men
## 56847:      Marina Loenning   Women
## 56848:      Jaap-Jan Linze Postma
## 56849:      Andreas Albert Pfisterer   Men
```

#data.table count funktion

```
den[, .N, .(sector)]
```

```
##              sector      N
##      1: Corporations  7989
##      2:              NGO 17720
##      3: Foundations  6987
##      4: Parliament  1087
##      5:      Family   207
##      6:      State 13601
##      7:      Events  1948
##      8:              2349
##      9: VL_networks  3803
##     10: Municipal   320
##     11:      Politics   37
##     12: Organisation    6
##     13: Commissions  795
```

Data bearbejdning (dplyr)

```
#data.table select funktion  
den %>% select(name, gender)
```

```
## # A tibble: 56,849 × 2
```

```
##      name                gender  
##      <chr>              <chr>  
## 1 Aage Almtoft          Men  
## 2 Aage B. Andersen      Men  
## 3 Aage Christensen      Men  
## 4 Aage Dam              Men  
## 5 Aage Dam              Men  
## 6 Aage Frandsen         Men  
## 7 Aage Juhl Joergensen  Men  
## 8 Aage Krogsdam         Men  
## 9 Aage Larsen           Men  
## 10 Aage Lauridsen       Men  
## # ... with 56,839 more rows
```

```
#data.table count funktion  
den %>% group_by(sector) %>% summarize(N = n()) #  
Or simply den %>% count(sector)
```

```
## # A tibble: 13 × 2
```

```
##      sector              N  
##      <chr>             <int>  
## 1 Commissions          795  
## 2 Corporations         7989  
## 3 Events                1948  
## 4 Family                207  
## 5 Foundations          6987  
## 6 Municipal             320  
## 7 NGO                   17720  
## 8 Organisation           6  
## 9 Parliament           1087  
## 10 Politics              37  
## 11 State                13601  
## 12 VL_networks          3803  
## 13 <NA>                 2349
```

Netværk visualisering

Two-mode netværk

```
# subsetting Commissions
den1 <- den[sector == "Commissions"]
# dplyr way: den1 <- den %>% filter(sector == "Commissions")

# cross tabulation som giver en såkaldt "incidence matrix"
incidence_matrix <- xtabs(formula = ~ name + affiliation, sparse = T, data
= den1)

# load two-mode netværk
net1 <- graph_from_incidence_matrix(incidence_matrix, directed = FALSE)
```

One-mode netværk

Option 1

```
# split two-mode netværk i to one-  
mode netværk  
net2 <- bipartite.projection(net1)  
  
# one node netværk, individuals  
net3 <- net2$proj1  
  
# one node netværk, affiliation  
net4 <- net2$proj2
```

Option 2

```
# one node netværk, individuals  
adjacency <- incidence_matrix %*%  
Matrix::t(incidence_matrix)  
net5 <-  
graph_from_adjacency_matrix(adjacenc  
y, mode = "undirected")  
  
# one node netværk, affiliation  
adjacency <-  
Matrix::t(incidence_matrix) %*%  
incidence_matrix  
net6 <-  
graph_from_adjacency_matrix(adjacenc  
y, mode = "undirected")
```

Netværk visualisering med ggraph

```
# graph objekt
net1%>%
# graph layout
  ggraph(layout = "fr") +
# Tilføjer forbindelser mellem aktørene - alpha (fra 0-1 - 0 svag, 1 stærk)
  geom_edge_link0(color = "gray60", alpha = 0.8) +
# Tilføjer punkter eller aktørene
  geom_node_point(aes(color = type), size = 3) +
# Tilføjer Commissions label
  geom_node_text(aes(filter=type==TRUE, label = name), repel = TRUE, size = 4) +
# ændrer farver + labels
  scale_color_manual(values = c("lightblue", "darkred"), labels = c("Individuals",
"Commissions")) +
# graph tema som skal altid tilføjes
  theme_graph() +
# det gør at "color" titlen ikke bliver vist i forklaringerne
  labs(color = "")
```


Øvelse

Opgaver

1. Download tom r-fil: **lektion01-øvelse** [her](#)
2. Svar på følgende spørgsmål:
 - Hvilke styrelser (affiliation) har de fleste medlemmer?
 - Hvor mange kvinder findes der i datasæt?
 - Hvem sidder i de fleste styrelser?
3. Lav et nyt datasæt “den1” hvor I kigger kun på aktørene i sektor “Parliament”
 - Er der flere kvinder i dette dataseæt?
4. Lav et one-mode netværk af individer og visualisere dette.
5. Beskriv netværket