Assignment 1

Network Analysis

Please submit your answers in a `.pdf` file containing your report and a `.R` file containing your codes (use comments to indicate where you are answering each question) to Blackboard *before* 15.00 Thursday November 10.

Please work through the (very) short introduction to R (on blackboard and available at `https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf`) *before* starting this assignment. You can skip section 11 on programming.

Practical questions

Question 1:

```
a <- c(1,6,NA,8,9,10)
mean(a)

## [1] NA
```

a. This gives `NA`, why? Compute the mean using an argument of the `mean()` function.

b. Now, remove the `NA` from the vector instead and compute the mean.

Question 2:

a. Create the following matrix in `R` and store the result in an object

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

b. Install and load the qgraph package and run the function `qgraph(input)` with the object containing your matrix as input. Copy the resulting plot in your report.

Question 3:

a. Create a *list* in `R` containing as first element the vector $7, 8, 2, 5$, as second element the vector $2, 8, 1, 7$ and as third element the vector $7, 6, 10, 4$. Call the list `mylist`

b. Run the commands `sapply(mylist, mean)` and `sapply(mylist, sd)`. Can you explain what happened? What do the `sapply`, `mean` and `sd` functions do?

c. What do the following two commands to? How do they relate to one-another?

```
cbind(mylist[[1]], mylist[[2]], mylist[[3]])
do.call(cbind, mylist)
```

Question 4:

a. Install and load the R package `psych` into R. Once `psych` is loaded, load the 'bfi' dataset:

```
# Load bfi dataset:
data(bfi)

# Look at data:
View(bfi)
```

Where can you find more information about this dataset and what the variables represent?

b. How many men are in the sample? And how many women?

c. Create a separate dataset containing only the first 25 columns (the items). Use the psych package to perform an "Exploratory Factor Analysis" (EFA) on this dataset with 5 factors, using "promax" rotation. Where did you find the information needed to perform a EFA using the psych package?

d. Make a plot of the estimated factor loadings using the `qgraph.loadings` function of the *qgraph* package. Tip, you might need the `loadings` function.

Question 5

Run the following codes in R:

```
set.seed(112)
Nperson <- 1000
Ability <- rnorm(Nperson)
Item1 <- 1*(exp(Ability)/(1+exp(Ability)) > runif(Nperson))
Item2 <- 1*(exp(Ability)/(1+exp(Ability)) > runif(Nperson))
```

This simulates three objects: `Ability` denotes the mathematical ability of 1000 subjects, and `Item1` and `Item2` represent scores (0 is incorrect, 1 is correct) on two mathematical questions. Those of you with a background in methods will recognize this as the Rasch model. In this model, people with a higher (lower) level of ability have a higher chance of making both items correct (incorrect).

a. Investigate the correlation between the two items. Is the correlation significant? Report both the correlation and the p-value.

b. Find a way to investigate the *partial correlation* between item 1 and item 2, after controlling for ability. Again, report both the correlation and the p-value.

c. What happened? Can you explain why this happened?

Question 6

A company wants to know how job performance relates to IQ, motivation and social support. They collect data on 60 employees, resulting in the file `job_performance.sav` on Blackboard.[1] The company made the crucial mistake of asking academics to analyze the data, who cannot afford a copy of SPSS. Hence, we need to analyze the data using R.

a. To load the data, we need to set the working directory. What is a working directory? And how do we set it?

b. Load the data into R as a data frame. Tip, the foreign package contains a function for this, but it requires an argument `to.data.frame=TRUE`. There are also other packages you can use.

c. Compute and report the correlation matrix of the four numeric items

d. Run the following command:

```
fit <- lm(perf ~ iq + mot + soc, data = mydata)
```

Where `mydata` is the object containing your dataset. What analysis did you perform?

d. Investigate the object `fit` to see if IQ is a significant predictor of performance.

[1] I have taken this example from https://www.spss-tutorials.com/linear-regression-in-spss-example/