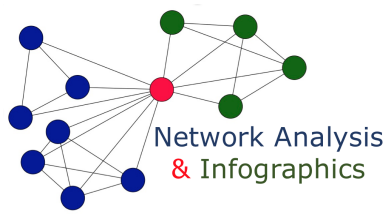
Network Definition
–Seminar–

Yasemin Aslan

SPRU (Science Policy Research Unit)
Business School
University of Sussex



Week 2: 3 February 2022

Overview

1. R objects
2. R operators
3. Importing/exporting data in R

Ready ...?

Ready ...?

R and RStudio

On campus

1. Go to http://rstudio.uscs.susx.ac.uk/
2. Access the website by using your University of Sussex account

On your personal computer

1. Install R
2. Install RStudio

RStudio cloud

1. Go to `https://rstudio.cloud`
2. Register and sign in

Ready ...?
Installing igraph

Through the RStudio interface:

1. Tools
2. Install Packages
3. Search for igraph
4. Tick the box "Install Dependencies"
5. Install

or

Through the RStudio console

```
1 install.packages("igraph")
2 library("igraph")
```

R objects

R objects

- R is an object-oriented language
- R creates and manipulate objects
- Different types or classes of objects
- The most used objects are
  - Data objects
  - Function objects
- Objects have attributes
- Objects are stored in a workspace called *environment* in RStudio

Data **type**

Data **type**

- numeric
  (*1.5*, *-102*, *0.001*, *...*)

Data **type**

- numeric
  (*1.5, -102, 0.001, ...*)
- integer
  (*1, -4, 4, ...*)

R objects
Data objects

Data **type**

- numeric
  (*1.5, -102, 0.001, ...*)
- integer
  (*1, -4, 4, ...*)
- character
  (*“a”, “weather”, “you are”, ...*)

R objects
Data objects

Data **type**

- numeric
  (*1.5, -102, 0.001, ...*)
- integer
  (*1, -4, 4, ...*)
- character
  (*"a", "weather", "you are", ...*)
- logical
  (*TRUE, FALSE* or *T, F*)

R objects
Data objects

Data **type**

- numeric
  (*1.5, -102, 0.001, ...*)
- integer
  (*1, -4, 4, ...*)
- character
  (*"a", "weather", "you are", ...*)
- logical
  (*TRUE, FALSE* or *T, F*)
- ...

Data **structure**

Data **structure**

- vector: an ordered collection of values

$$\mathbf{a} = \begin{pmatrix} 0 \\ 20 \\ 30 \\ -4 \\ 12 \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} \textit{Peter} \\ \textit{Sara} \\ \textit{Andrew} \\ \textit{Charlotte} \\ \textit{Rachel} \end{pmatrix}$$

Data **structure**

- vector: an ordered collection of values
- matrix: a 2-dimensional vector
  (a vector with > 2 dimensions is called array)

$$\mathbf{A} = \begin{pmatrix} 0 & 10 & 20 \\ 20 & -2 & 10 \\ 30 & -10 & 5 \\ -4 & 0 & 0 \\ 12 & -23 & 2 \end{pmatrix}$$

Data **structure**

- vector: an ordered collection of values
- matrix: a 2-dimensional vector
  (a vector with > 2 dimensions is called array)
- data frame: variables and observations

Table: Students in 2021/22

Student	Course
Student 1	SD
Student 2	STP
Student 3	SD
Student 4	SIM
Student 5	SD
Student 6	STP
Student 7	SD
Student 8	SD
Student 9	SIM
Student 10	SIM
Student 11	SD
Student 12	STP
Student 13	SD
Student 14	STP
Student 15	SD
Student 16	SIM
Student 17	SIM
Student 18	SD
Student 19	SIM
Student 20	SIM

R objects
Data objects

Data **structure**

- vector: an ordered collection of values
- matrix: a 2-dimensional vector
  (a vector with > 2 dimensions is called array)
- data frame: variables and observations
- list: an ordered sequences of objects

$$\mathbf{a} = \begin{pmatrix} 0 \\ 20 \\ 30 \\ -4 \\ 12 \end{pmatrix}$$

$$\mathbf{b} = \begin{pmatrix} red \\ red \\ green \\ yellow \\ yellow \end{pmatrix}$$

$$\mathbf{A} = \begin{pmatrix} 0 & 10 & 20 \\ 20 & -2 & 10 \\ 30 & -10 & 5 \\ -4 & 0 & 0 \\ 12 & -23 & 2 \end{pmatrix}$$

Data **structure**

- vector: an ordered collection of values
- matrix: a 2-dimensional vector
  (a vector with > 2 dimensions is called array)
- data frame: variables and observations
- list: an ordered sequences of objects
- factor: categorical data (e.g. "male", "female")

$$\mathbf{a} = \begin{pmatrix} red \\ red \\ green \\ yellow \\ yellow \end{pmatrix}$$

factors = green, red, yellow

R objects
Functions

- A function is a command in R that returns a given outcome
- Functions can read, manipulate and analyse data
- Certain base functions are already integrated in the basic R
- Packages provide users with additional functions (e.g. igraph)

plot()

mean()

read_csv(...)

write_csv(...)

...

R objects
Attributes

- name: the name of the object
- mode: the type of data
- length: number of elements in the object
- ..

R operators

R operators

Arithmetic operators

Operator	Description
+	addition
−	subtraction
*	multiplication
/	division
^	exponentiation

R operators

Logical operators

Operator	Description
$<$	less than
$<=$	less than or equal to
$>$	greater than
$>=$	greater than or equal to
$==$	equal to
$!=$	not equal to
$!x$	Not x
$x \vert y$	x OR y
$x \& y$	x AND y

R operators

Let's explore objects and operators in RStudio

Importing/exporting data in R

Importing/exporting data in R

- R can deal with many different data formats

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:
  - "readr" (csv, txt, delimited)

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:
  - "readr" (csv, txt, delimited)
  - "readxl" (Excel)

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:
  - "readr" (csv, txt, delimited)
  - "readxl" (Excel)
  - "haven" (STATA, SPSS, SAS)

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:
  - "readr" (csv, txt, delimited)
  - "readxl" (Excel)
  - "haven" (STATA, SPSS, SAS)
  - "tidyverse"

Importing/exporting data in R

- R can deal with many different data formats
- Which data formats do you know?
- Most commonly used data formats
  - txt: fields are usually separated by *comma*, *tab*, *colon*
  - csv: fields are separated by *comma*
  - dta: format used by STATA to store data
  - xls/xlsx: format used by Excel to store data
  - ...
- Text editors: Notepad++ (Windows), TextWrangler (Mac)
- R packages to read and transform data:
  - "readr" (csv, txt, delimited)
  - "readxl" (Excel)
  - "haven" (STATA, SPSS, SAS)
  - "tidyverse"
- Install these packages in R (you should now know how to do it)

Importing/exporting data in R

Let's import/export data in RStudio

Next time ...

Next time ...

- **Lecture: Network data collection**
  - Main approaches to collect and sample network data
  - Network boundary specification problem

- **Seminar: Network data collection**
  - Network file formats
  - How to create/import and manipulate network data in igraph

Questions