How to Analyze 'Network Data'

Visualization

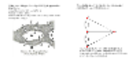
Characterization
- transitivity
- assortativity

Modeling, and more...
- Mathematical(Ch5)
- Statistical (Ch6)

Network Science

1930s-: Graph Theory

Recent 10-15yrs:
Sampling
Modeling
Inference of network topology
Network-indexed processes



Why do we study 'Network'?

Usefulness of a Network-based perspective

Chapter 2 Starts...! Codes are in BLUE

Chapter 2 Manipulating Network Data (Continued)

graph.formula

Structure function

Addition, Subtraction,
induced.subgraph(, 1:5)
graph.union(, ,)

```
get.edgelist( )
get.adjlist( )
get.adjacency( )

graph.edgelist( )
graph.adjlist( )
graph.adjacency( )
```

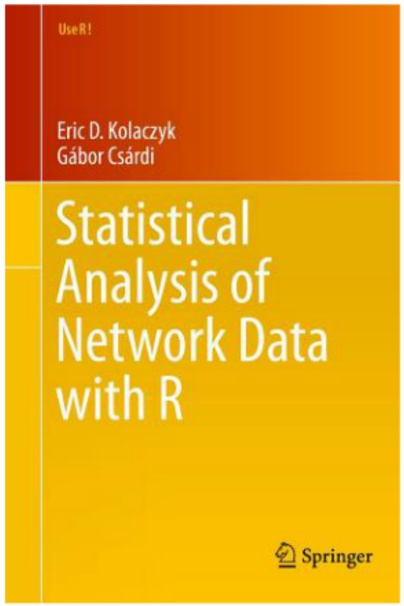
```
is.weighted( )
is.simple( )
```

```
<<< 2.20
library(sand)
g.lazega <- graph.data.frame(elist.lazega,
                             directed="FALSE",
                             vertices=v.attr.lazega)
g.lazega$name <- "Lazega Lawyers"
```

```
vcount( )
ecount( )

elist.lazega
v.attr.lazega
list.vertex.attributes(g.lazega)

neighbors( , 5)
degree( )  #number of neighbors
degree( , mode="in")
is.connected( )
is.connected( , mode= "weak")
```



Prezi

Data Visualization Basics
GeeS. Eun
Feb 13. 2015, Jamsil

Part 3-B. Statistical Analysis of Network Data with R

Ch1. Introduction
Ch2. Manipulating Network Data

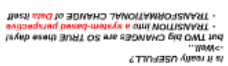
Prezi

Why do we study 'Network'?

Prezi

Network-indexed processes

Why do we study 'Network'?



Usefulness of a Network-based perspective

Prezi

Is it really USEFULL?

->Well...

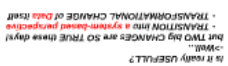
but TWO big CHANGEs are SO TRUE these days!

- ***TRANSITION into a system-based perspective***
- ***TRANSFORMATIONAL CHANGE of Data itself***

Prezi

Network-indexed processes

Why do we study 'Network'?



Usefulness of a Network-based perspective

Prezi

Network Science

1930s-: Graph Theory

http://mathworld.wolfram.com/Graph.html

Recent 10-15yrs:

Sampling

Modeling

Inference of network topology

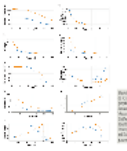
Network-indexed processes

Prezi

How to Analyze 'Network Data'

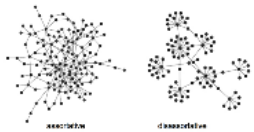
Visualization

Characterization



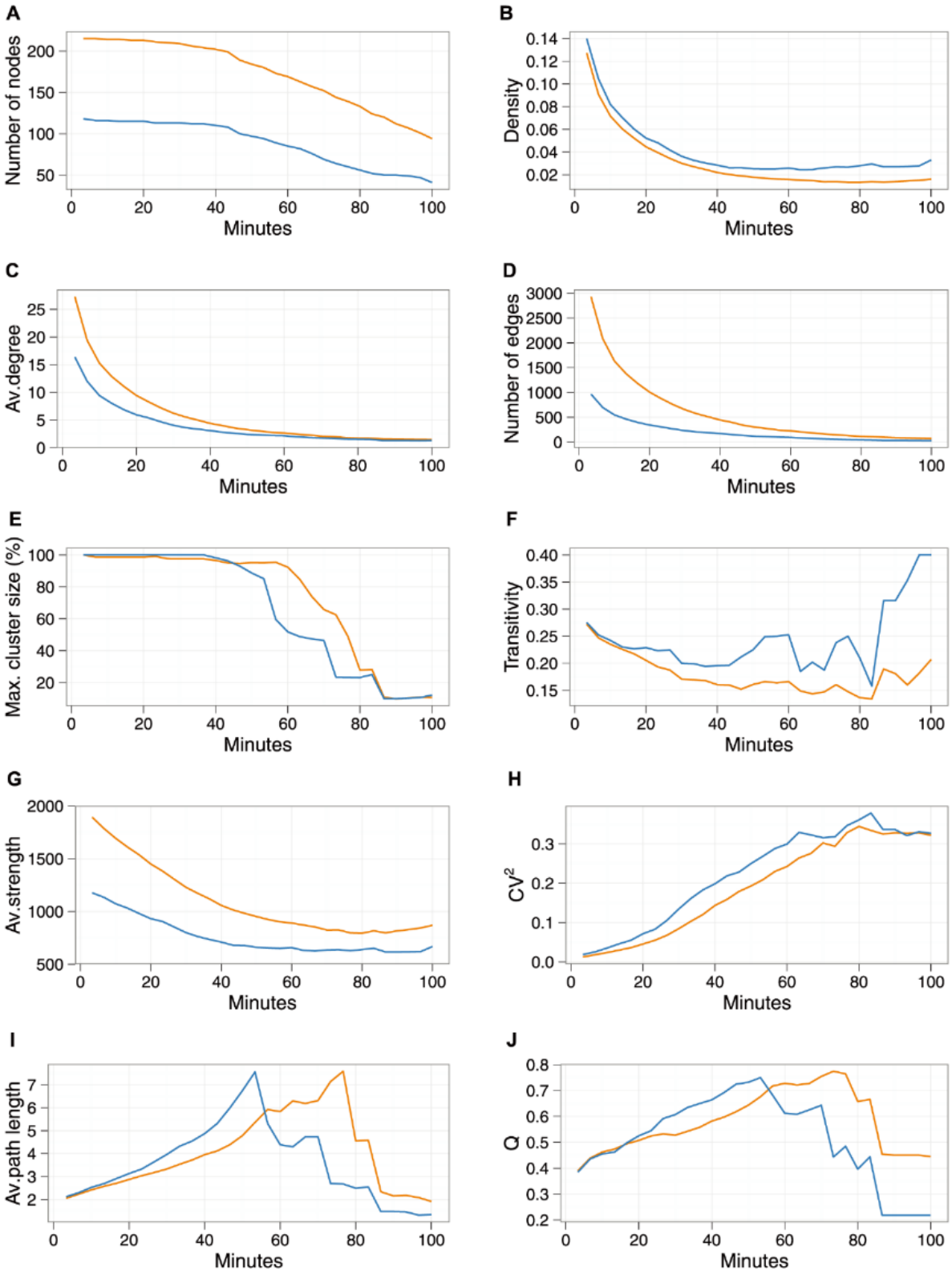
- transitivity
- assortativity

http://en.wikipedia.org/wiki/Transitivity
http://plus.maths.org/content/exploring-financial-ecosystem



Modeling, and more...

Prezi



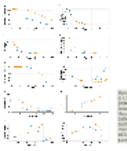
Barclay VC, Smi
J, Cao G, Rainey
(2014) Positive N
Assortativity of In
Vaccination at a
School: Implicati
Outbreak Risk a
Immunity. PLoS
e87042. doi:10.1
journal.pone.008

Prezi

How to Analyze 'Network Data'

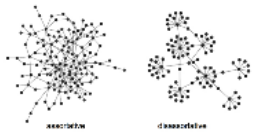
Visualization

Characterization



- transitivity
- assortativity

http://en.wikipedia.org/wiki/Transitivity
http://plus.maths.org/content/exploring-financial-ecosystem



Modeling, and more...

Prezi

assortative

disassortative

Prezi

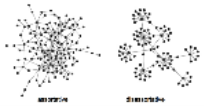
How to Analyze 'Network Data'

Visualization

Characterization

- transitivity
- assortativity

http://en.wikipedia.org/wiki/Transitivity
http://plus.maths.org/content/exploring-financial-ecosystem



Modeling, and more...

- **Mathematical(Ch5)**
- **Statistical (Ch6)**

Prezi

Data Visualization Basics
GeeS. Eun
Feb 13. 2015, Jamsil

Part 3-B. Statistical Analysis of Network Data with R

Ch1. Introduction
Ch2. Manipulating Network Data

Prezi

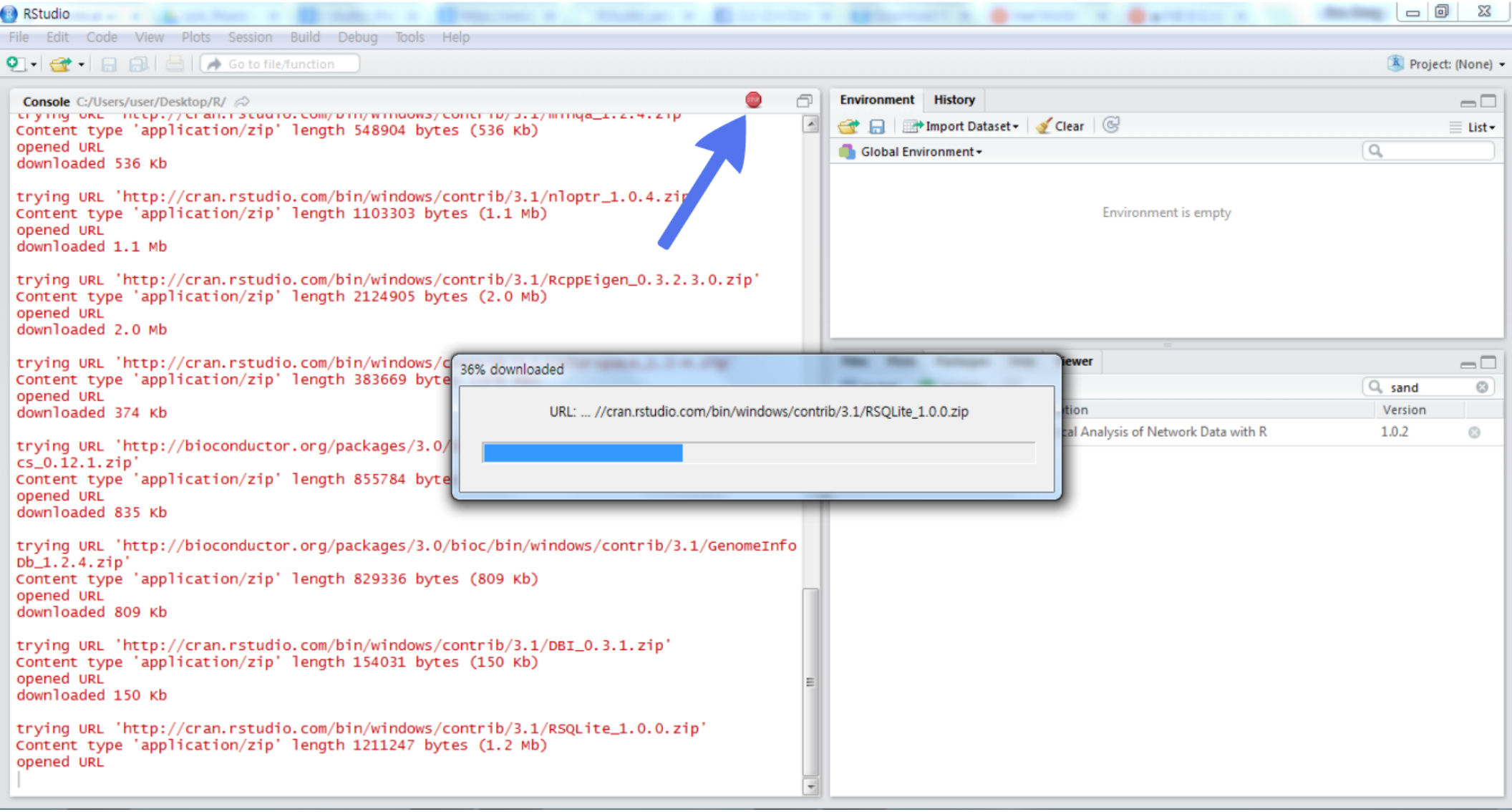
Chapter 2 Starts...! Codes are in BLUE

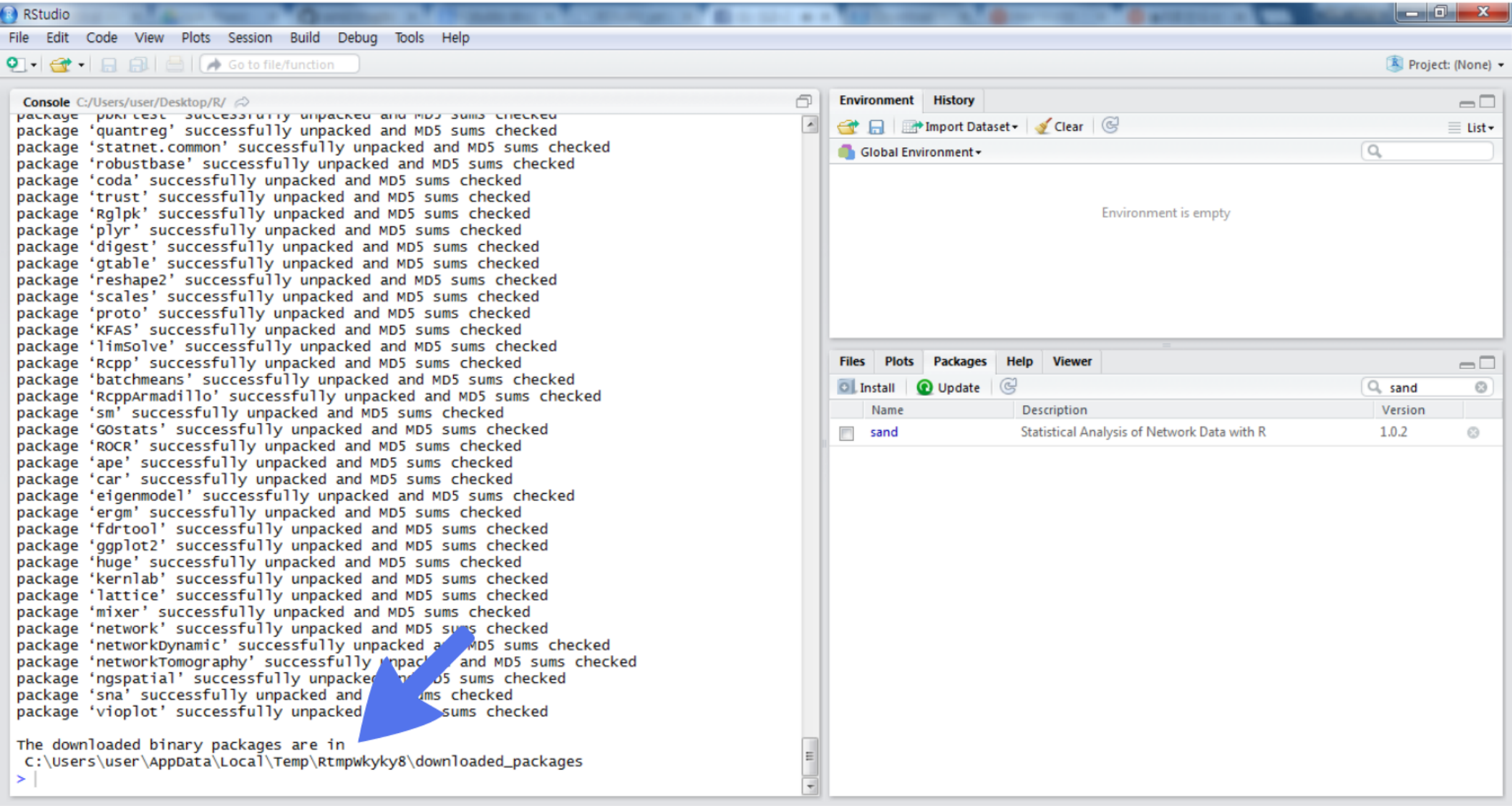
http://cran.nexr.com
http://www.rstudio.com
https://github.com/kolaczyk/sand

Before We Start:
Type in

```
install.packages("sand")
library(sand)
install_sand_packages()
```

Prezi


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Go to file/function
Project: (None)
Console C:/Users/user/Desktop/R/
Content type 'application/zip' length 548904 bytes (536 Kb)
opened URL
downloaded 536 Kb
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/nloptr_1.0.4.zip'
Content type 'application/zip' length 1103303 bytes (1.1 Mb)
opened URL
downloaded 1.1 Mb
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/RcppEigen_0.3.2.3.0.zip'
Content type 'application/zip' length 2124905 bytes (2.0 Mb)
opened URL
downloaded 2.0 Mb
trying URL 'http://cran.rstudio.com/bin/windows/
Content type 'application/zip' length 383669 byte
opened URL
downloaded 374 Kb
trying URL 'http://bioconductor.org/packages/3.0/
cs_0.12.1.zip'
Content type 'application/zip' length 855784 byte
opened URL
downloaded 835 Kb
trying URL 'http://bioconductor.org/packages/3.0/bioc/bin/windows/contrib/3.1/GenomeInfo
Db_1.2.4.zip'
Content type 'application/zip' length 829336 bytes (809 Kb)
opened URL
downloaded 809 Kb
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/DBI_0.3.1.zip'
Content type 'application/zip' length 154031 bytes (150 Kb)
opened URL
downloaded 150 Kb
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/RSQLite_1.0.0.zip'
Content type 'application/zip' length 1211247 bytes (1.2 Mb)
opened URL
Environment History
Import Dataset
Clear
List
Global Environment
Environment is empty
36% downloaded
URL: ... //cran.rstudio.com/bin/windows/contrib/3.1/RSQLite_1.0.0.zip
sand
Version
cal Analysis of Network Data with R
1.0.2


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Go to file/function
Project: (None)
Console C:/Users/user/Desktop/R/
package 'pbkrtest' successfully unpacked and MD5 sums checked
package 'quantreg' successfully unpacked and MD5 sums checked
package 'statnet.common' successfully unpacked and MD5 sums checked
package 'robustbase' successfully unpacked and MD5 sums checked
package 'coda' successfully unpacked and MD5 sums checked
package 'trust' successfully unpacked and MD5 sums checked
package 'Rglpk' successfully unpacked and MD5 sums checked
package 'plyr' successfully unpacked and MD5 sums checked
package 'digest' successfully unpacked and MD5 sums checked
package 'gtable' successfully unpacked and MD5 sums checked
package 'reshape2' successfully unpacked and MD5 sums checked
package 'scales' successfully unpacked and MD5 sums checked
package 'proto' successfully unpacked and MD5 sums checked
package 'KFAS' successfully unpacked and MD5 sums checked
package 'limSolve' successfully unpacked and MD5 sums checked
package 'Rcpp' successfully unpacked and MD5 sums checked
package 'batchmeans' successfully unpacked and MD5 sums checked
package 'RcppArmadillo' successfully unpacked and MD5 sums checked
package 'sm' successfully unpacked and MD5 sums checked
package 'GOstats' successfully unpacked and MD5 sums checked
package 'ROCR' successfully unpacked and MD5 sums checked
package 'ape' successfully unpacked and MD5 sums checked
package 'car' successfully unpacked and MD5 sums checked
package 'eigenmodel' successfully unpacked and MD5 sums checked
package 'ergm' successfully unpacked and MD5 sums checked
package 'fdrtool' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked
package 'huge' successfully unpacked and MD5 sums checked
package 'kernlab' successfully unpacked and MD5 sums checked
package 'lattice' successfully unpacked and MD5 sums checked
package 'mixer' successfully unpacked and MD5 sums checked
package 'network' successfully unpacked and MD5 sums checked
package 'networkDynamic' successfully unpacked and MD5 sums checked
package 'networkTomography' successfully unpacked and MD5 sums checked
package 'ngspatial' successfully unpacked and MD5 sums checked
package 'sna' successfully unpacked and MD5 sums checked
package 'vioplot' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:\Users\user\AppData\Local\Temp\RtmpWkyky8\downloaded_packages
>
Environment History
Import Dataset Clear List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Install Update
sand
Name Description Version
sand Statistical Analysis of Network Data with R 1.0.2

Prezi

Chapter 2 Manipulating Network Data (Continued)

After Installation,
Rstudio > Packages > sand (click it!)
Type in
C2
Type in
N
Type in
N
again... and again... ALL YOU NEED IS N

graph.formula

Structure function Str( )
Plot( )
V( )
E( )

Addition, Subtraction,
induced.subgraph(, 1:5)
graph.union(, ,)

Prezi

<- graph.formula(- , - , -)

cf. - , -+ , +- , ++

Prezi

Chapter 2 Manipulating Network Data (Continued)

After Installation,
Rstudio > Packages > sand (click it!)
Type in
C2
Type in
N
Type in
N
again... and again... ALL YOU NEED IS N

graph.formula

Structure function Str( )
Plot( )
V( )
E( )

Addition, Subtraction,
induced.subgraph(, 1:5)
graph.union(, ,)

Prezi

```
library(igraph)
g <- graph.formula(1-2, 1-3, 2-3, 2-4, 3-5, 4-5, 4-6,
                   4-7, 5-6, 6-7)

V(g)

Vertex sequence:
[1] "1" "2" "3" "4" "5" "6" "7"

E(g)

Edge sequence:

[1]  2 -- 1
[2]  3 -- 1
[3]  3 -- 2
[4]  4 -- 2
[5]  5 -- 3
[6]  5 -- 4
[7]  6 -- 4
[8]  7 -- 4
[9]  6 -- 5
[10] 7 -- 6

str(g)

IGRAPH UN-- 7 10 --
+ attr: name (v/c)
+ edges (vertex names):
1 -- 2, 3
2 -- 1, 3, 4
3 -- 1, 2, 5
4 -- 2, 5, 6, 7
5 -- 3, 4, 6
6 -- 4, 5, 7
7 -- 4, 6

plot(g)
```

Prezi

Chapter 2 Manipulating Network Data (Continued)

After Installation,
Rstudio > Packages > sand (click it!)
Type in
C2
Type in
N
Type in
N
again... and again... ALL YOU NEED IS N

graph.formula

Structure function Str( )
Plot( )
V( )
E( )

Addition, Subtraction,
induced.subgraph(, 1:5)
graph.union(, ,)

Prezi

graph.union(, ,)

get.edgelist( )
get.adjlist( )
get.adjacency( )

graph.edgelist( )
graph.adjlist( )
graph.adjacency( )

is.weighted( )
is.simple( )

chunk 2.26
E( ) $weight <- 1
E( ) $weight

<- simplify( )
E( ) $weight

Weighted True vs False
Multigraph vs simplegraph

```
<<< 2.20
library(sand)
g.lazega <- graph.data.frame(elist.lazega,
                             directed="FALSE",
                             vertices=v.attr.lazega)
g.lazega$name <- "Lazega Lawyers"
```

Prezi

chunk 2.26
E(●) $weight <- 1
E(●) $weight

●<- simplify(●)
E(●) $weight

Weighted True vs False
Multigraph vs simplegraph

Prezi

graph.union(, ,)

get.edgelist( )
get.adjlist( )
get.adjacency( )

graph.edgelist( )
graph.adjlist( )
graph.adjacency( )

is.weighted( )
is.simple( )

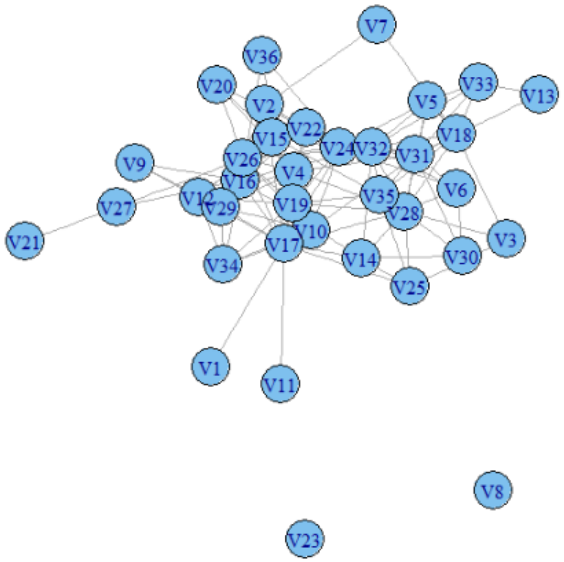
chunk 2.26
E( ) $weight <- 1
E( ) $weight

<- simplify( )
E( ) $weight

Weighted True vs False
Multigraph vs simplegraph

```
<<< 2.20
library(sand)
g.lazega <- graph.data.frame(elist.lazega,
                             directed="FALSE",
                             vertices=v.attr.lazega)
g.lazega$name <- "Lazega Lawyers"
```

Prezi



vcount( )
ecount( )

elist.lazega
v.attr.lazega
list.vertex.attributes(g.lazega)

neighbors(, 5)
degree( ) *=number of neighbors*
degree(, mode="in")
is.connected( )
is.connected(, mode= "weak")

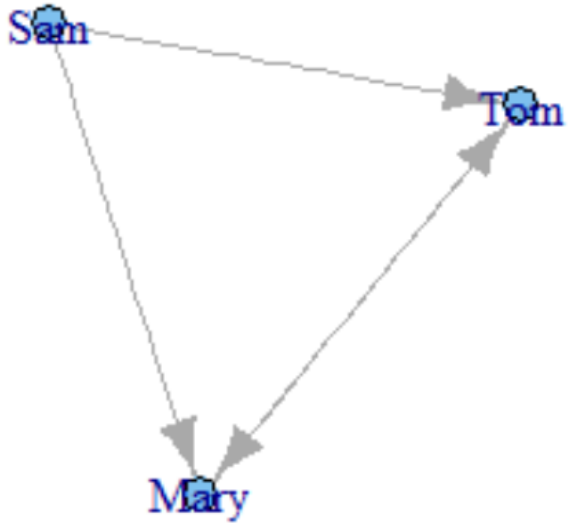
Prezi

#2.31
1 > degree(dg, mode="in")
2 >Sam Mary Tom
3 > 0   2   2

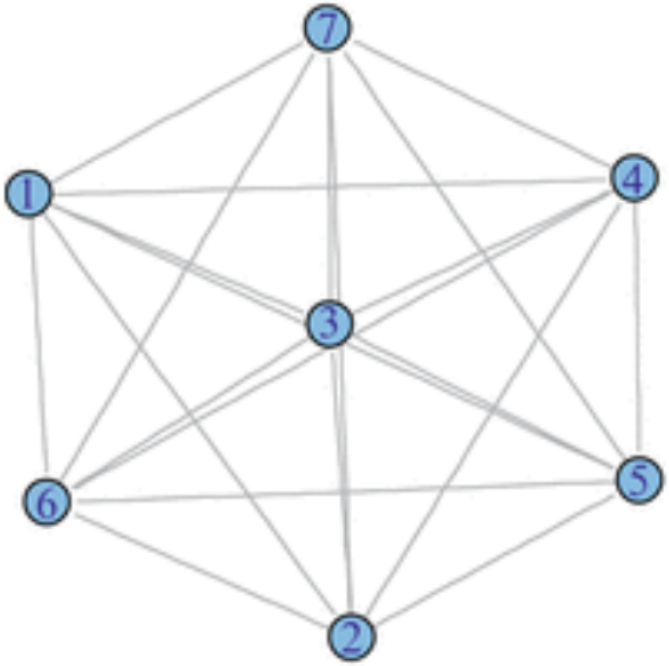
4 > degree(dg, mode="out")
5 >Sam Mary Tom
6 > 2   1   1

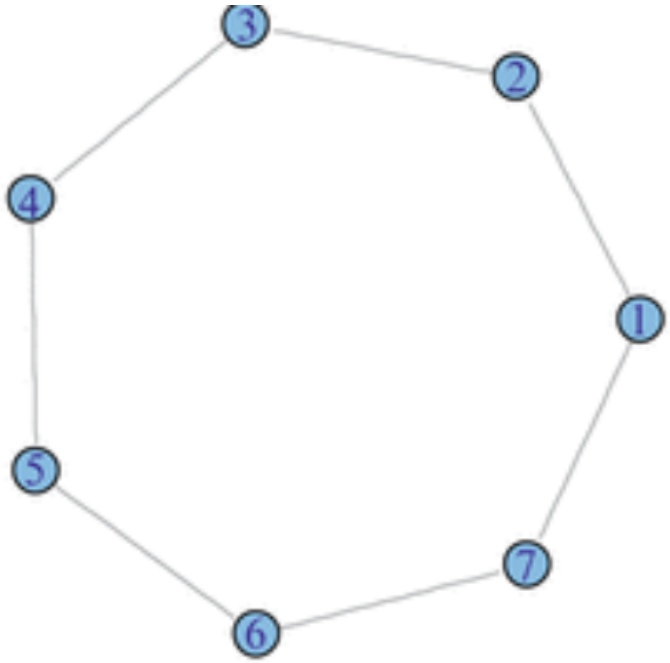
#2.34
1 > is.connected(dg, mode="weak")
2 [1] TRUE

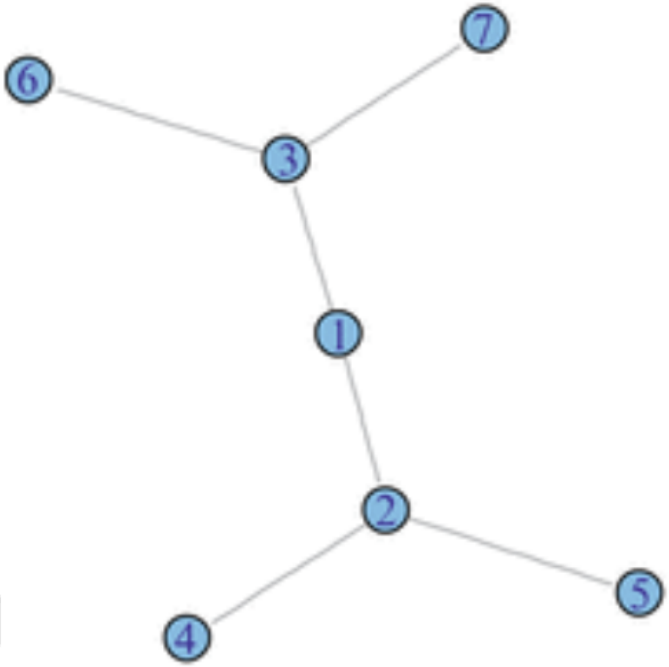
3 > is.connected(dg, mode="strong")
4 [1] FALSE

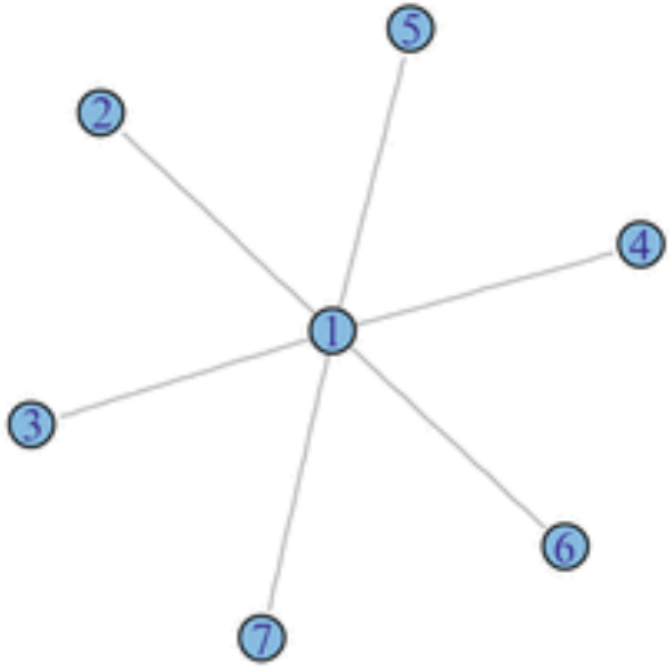


Prezi









Prezi

How to Analyze 'Network Data'

Visualization

Characterization

- transitivity
- assortativity

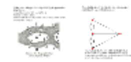
Modeling, and more...

- Mathematical(Ch5)
- Statistical (Ch6)

Network Science

1930s-: Graph Theory

Recent 10-15yrs:
Sampling
Modeling
Inference of network topology
Network-indexed processes



Why do we study 'Network'?

Usefulness of a Network-based perspective

Chapter 2 Starts...! Codes are in BLUE

Before We Start
Type in

```
install.packages("sand")
library(sand)
install_sand_packages()
```

Chapter 2 Manipulating Network Data (Continued)

graph.formula

Structure function
Plot()
V()
E()

Addition, Subtraction,
induced.subgraph(, 1:5)
graph.union(, ,)

```
get.edgelist( )
get.adjlist( )
get.adjacency( )

graph.edgelist( )
graph.adjlist( )
graph.adjacency( )
```

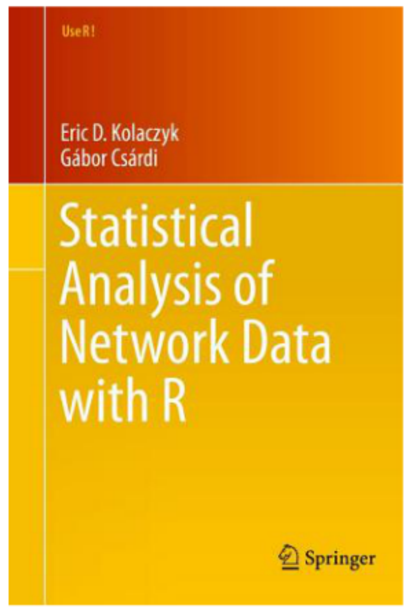
```
is.weighted( )
is.simple( )
```

```
<<< 2.20
library(sand)
g.lazega <- graph.data.frame(elist.lazega,
                             directed="FALSE",
                             vertices=v.attr.lazega)
g.lazega$name <- "Lazega Lawyers"
```

```
vcount( )
ecount( )

elist.lazega
v.attr.lazega
list.vertex.attributes(g.lazega)

neighbors( , 5)
degree( )  #number of neighbors
degree( , mode="in")
is.connected( )
is.connected( , mode= "weak")
```



Prezi