Social Network Analysis
R Basics

Robin Burke

DePaul University

Chicago, IL

What I'm assuming

- Either
  - You can work with R
  - You can put in the effort to learn right now
- Also
  - Basic computing principles

"Work with R"

- You can
  - Use R Studio on your machine / in the lab
  - Load R libraries
  - Read and make sense of R documentation
  - Use R data types: factors, strings, vectors, lists, data frames
  - Create a data frame from a CSV file
  - Create a data frame from individual vectors
  - Extract subsets and projections of data frames
  - Compute summary statistics on data frames and vectors
  - Search on stackoverflow for answers ☺

What I don't expect

- Write your own functions / packages in R
- Know all of the packages available in R
  - Ha!
- Know about any of the packages we're learning in this class
  - sand, igraph, statnet, ggplot

Experimental attitude

○ If you don't understand something

 ● Create a toy example and test it

○ R is extremely interactive

How to actually learn any new programming concept

Essential

Changing Stuff and
Seeing What Happens

O RLY?

@ThePracticalDev

One Hour Rule

- If you are having trouble getting something to work
  - You MUST spend one hour trying to solve the problem
- If you spend an hour and the problem is not solved
  - You MUST ask for help
  - Use the slack forum ← Probably fastest
  - Send email to me ← Use if the question requires sharing your homework solution

a homework takes 3 hours to complete. If yo...
...omework after class (9 pm) on the day it is du...
...y times can you ask for help and still turn your homew...
in on time?

Zero

Once

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

e

R data types

○ Atomic values

- 1, 3.14159, "fred", fred

○ Collections

- Vectors
- Lists
- Data frames

String value

Factor value

Vectors and indexing
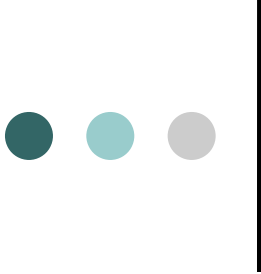
```
> v1 <- c(2,4,6,8,10)
> v1
[1]  2  4  6  8 10
> v1[-2]
[1]  2  6  8 10
> v1[0]
numeric(0)
```

English translation:
"This is not Python."

Element-wise and summary operations

```
> v1 * 2
[1]  4  8 12 16 20
> log(v1)
[1] 0.6931472 1.3862944 1.7917595 2.0794415 2.3025851
> v2 <- c(1,2,3,4,5)
> v1 * v2
[1]  2  8 18 32 50
> v1 + v2
[1]  3  6  9 12 15
> sum(v1)
[1] 30
> summary(v1 + v2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      3       6       9       9      12      15
```

Data frames and factors

```
df1 <- data.frame(col1=v1, col2=c("red", "red", "blue", "red", "black"))
> df1
   col1  col2
1    2   red
2    4   red
3    6   blue
4    8   red
5   10  black
> summary(df1)
     col1        col2
Min.   : 2    black:1
1st Qu.: 4    blue :1
Median : 6    red  :3
Mean   : 6
3rd Qu.: 8
Max.   :10
> df1$col2
[1] red    red    blue   red
blackLevels: black blue red
```

Data frame indexing

```
> df1[1,]
    col1 col2
1    2  red
df1[,1]
[1]  2  4  6  8 10
> df1[1,1]
[1] 2
> df1[-1,]
     col1  col2
2    4   red
3    6  blue
4    8   red
5   10 black
> df1[c(F, F, T, F, T),]
      col1  col2
3    6  blue
5   10 black
> df1[df1$col2!="red",]
    col1  col2
3    6  blue
5   10 black
```

...he data frame in the prior slides, what is the r...
...ng the following code: df1[df1$col2!="red",]$col...

Error

All of the values in col1
are set to -1

All of the values in col2
are set to -1

The values in col1 in rows
...nd 5 only are set to -1

Worked example
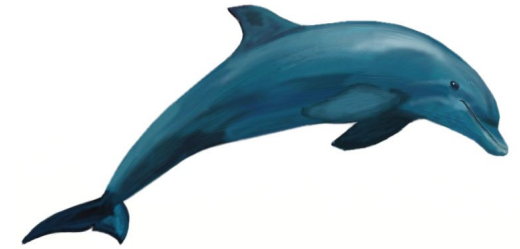
- Load a network into igraph
  - Dolphins!

- Examines its nodes and edges

- Visualize it

- Plot male / female distribution

- Plot its degree distribution

- Plot weighted degree

- Plot the degree boxplots for male and female dolphins

Homework 1

- Similar to the example
  - large network
- Turn in three files
  - HTML output
  - Rmd file that generated it
  - R file that supplies the code chunks
- I supply "skeleton" versions of R and Rmd files

Next week

- Homework 1 due
- Hahnemann and Riddle reading
- Write up the example from class for extra credit