

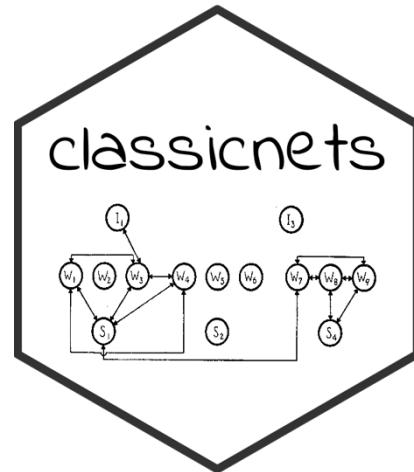
Title:

Using R and GitHub to compile/conserve classic books:  
the case of a repository with social networks data

Authors:

Francisca Ortiz Ruiz, Centre for Economics and Social Policy,  
and School of Government and Public Administration,  
Universidad Mayor, Chile. Millennium Institute for Care  
Research (MICARE), Santiago, Chile.

Alejandro Espinosa-Rada, Instituto de Sociología, P. Universidad  
Católica de Chile.



Abstract:

In this presentation, we will share some of the learnings and challenges of using R, complemented by GitHub, to build a database repository. Within the open science framework, we used a specific case of study called the “classinets” R package (<https://github.com/anespinoza/classinets>). Its name is translated as “Classic Data of Social Networks”. Five main learnings and challenges during the process of construction were identified. In the first place, the access to the books (or primary source) in which the network information is available. Once we could access those books, there was a second challenge: finding the information necessarily for the documentation of the data sets. Third, achieving a standard representation for uploading the networks. As networks can be represented using different data structures such as adjacency lists, edgelists, matrices, to name some of the most common representations. Fourth, managing to find a way to present clear, organised and well-informed data collected from a book into a GitHub repository required some creativity. Finally, disseminating the R package, as it is trying to recover “old” datasets and contents of the network perspective, has been challenging to engage with the audience.

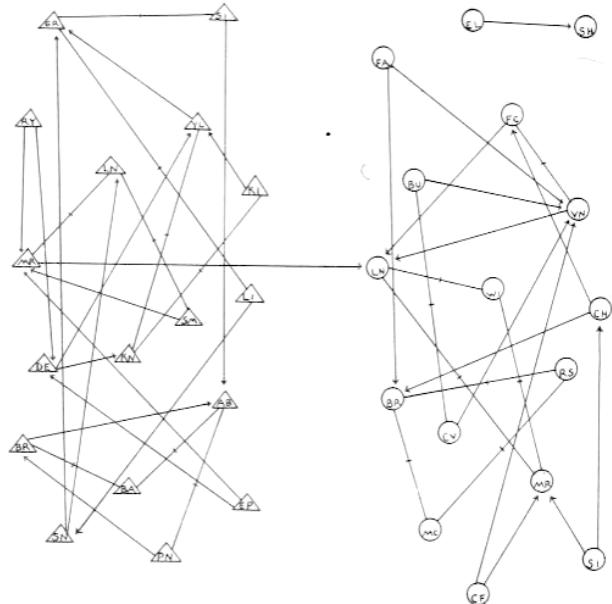
In addition, we reflect on the benefits of open access in research for social network analysis. These include the following three main: First, open-access research increases visibility so that more people can use this database and, eventually, have data available from “classic books” selected in the repository, which otherwise might not be accessible for broader audiences. Second, more collaboration between researchers should be promoted through the use of completely available databases. Third, transparency of this data type for social network research may be helpful for teaching purposes. We illustrate the case of the Zachary karate dataset often used to optimise community detection algorithms.

The goal of “classinets” is to make available, compile, and conserve classic data of social networks. We follow the history of this perspective based on Linton Freeman, John Scott, and reviews from Ibero-America. The data are in the classic matrix format, which is one of the most common formats used in the social networks perspective.

During the presentation, we will show some classic examples of the network's perspectives, incorporating some information about the theory and code behind it. Among others:

“Friendship choices among pupils” by Jacob Moreno (1934)  
data("pupils\_classroom")

### EVOLUTION OF GROUPS



### CLASS STRUCTURE, 4TH GRADE

17 boys and 16 girls. *Unchosen*, 6, EP, RY, EL, FA, SI, CF; *Pairs*, 17, GR-SI, GR-LI, MR-LN, LN-SM, YL-KN, AB-BA, BA-BR, KI-KN, AB-PN, FC-VN, BU-CV, LN-WI, LN-MR, BR-MC, BR-RS, WI-MR, MC-RS; *Stars*, 2, LN, VN; *Chains*, 0; *Triangles*, 2, BR-RS-MC; LN-WI-MR; *Inter-sexual Attraction*s, 1.

“Western Electric Company” by Roethlisberger, Fritz Jules and Dickson, William John (1939)  
data("bank\_room")

```
matrix_to_edgelist(bank_room$friendship, digraph = TRUE)
```

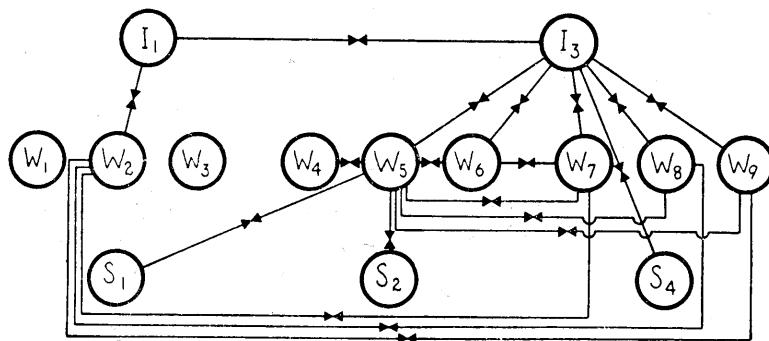


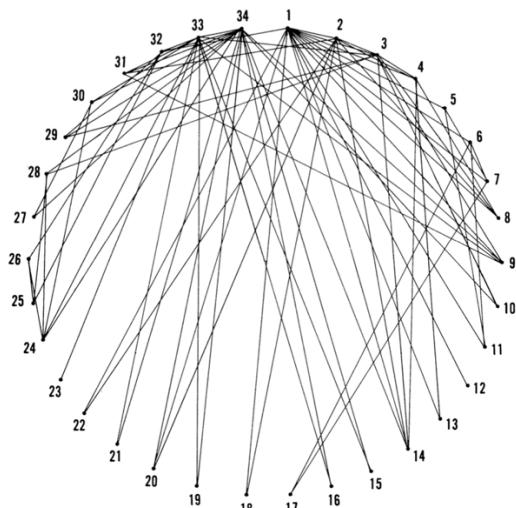
FIGURE 44  
ANTAGONISMS

BANK WIRING OBSERVATION ROOM

“Southern Women” by Davis, Allison; Gardner, Burleigh B. and Mary. R. Gardner (1941).  
data("southern\_women")

NAME OF PARTICIPANTS OR GROUP I	CODE NUMBERS AND DATES OF SOCIAL EVENTS REPORTED IN <i>Old City Herald</i>													
	(1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/15	(9) 4/8	(10) 6/10	(11) 2/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson.....	X	X	X	X	X	X	...	X	X					
2. Miss Laura Mandeville.....	X	X	X	...	X	X	X	X	...					
3. Miss Theresa Anderson.....	...	X	X	X	X	X	X	X	X	...				
4. Miss Brenda Rogers.....	X	...	X	X	X	X	X	X	...					
5. Miss Charlotte McDowd.....	...	X	X	X	X	...	X	...						
6. Miss Frances Anderson.....	...	X	...	X	X	...	X	...						
7. Miss Eleanor Nye.....	...	...	...	X	X	X	X	...						
8. Miss Pearl Oglethorpe.....	...	...	...	...	X	...	X	...						
9. Miss Ruth DeSand.....	...	...	...	...	X	...	X	X	X	...				
10. Miss Verne Sanderson.....	...	...	...	...	...	X	...	X	X	X	...	X	...	
11. Miss Myra Liddell.....	...	...	...	...	...	...	...	X	X	X	X	...	...	
12. Miss Katherine Rogers.....	...	...	...	...	...	...	X	X	X	X	X	...	X	X
13. Mrs. Sylvia Avondale.....	...	...	...	...	...	...	X	X	X	X	X	...	X	X
14. Mrs. Nora Fayette.....	...	...	...	...	...	X	X	...	X	X	X	...	X	X
15. Mrs. Helen Lloyd.....	...	...	...	...	...	X	X	...	X	X	X	...	X	X
16. Mrs. Dorothy Murchison.....	...	...	...	...	...	...	X	X	...	X	X	...		
17. Mrs. Olivia Carleton.....	...	...	...	...	...	...	...	X	...	X	...	...		
18. Mrs. Flora Price.....	...	...	...	...	...	...	...	...	X	...	X	...		

Zachary, Wayne W. (1977).  
 data("karate\_club")



#### Bibliographic references:

- Davis, Allison; Gardner, Burleigh B. and Mary. R. Gardner (1941). Deep South: A Social Anthropological Study of Caste and Class. The University of Chicago Press, Chicago.
- Moreno, Jacob L. (1934). Who Shall Survive? Washington, DC: Nervous and Mental Disease Publishing Company.
- Roethlisberger, Fritz Jules and Dickson, William John (1939). Management and the worker. An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago. Harvard University Press: Cambridge, Massachusetts.
- Zachary, Wayne W. (1977). An information flow model for conflict and fission in small groups. Journal of Anthropological Research, 33(4), 452-473.