DEN
Data Exchange Network

The "Camp '92" Dataset

Stephen P. Borgatti
University of Kentucky

H. Russell Bernard
University of Florida

Gery Ryan
Rand Corporation

Pertti Pelto
University of Conneticut

Rich DeJordy
Northeastern University

1. Overview

The Camp '92 dataset is a simple, small dataset collected in 1992 at the National Science Foundation's Summer Institute for Ethnographic Research Methods in Anthropology. This was a three-week workshop involving 14 participants (13 faculty and 1 post-doc) and 4 instructors (3 faculty and 1 PhD student). The participants lived together at a motel for the duration. The motel had a common room where group members hung out. Subsets of group members regularly ate together.

The data consist of one kind of network tie, collected at two points in time, along with a limited set of individual characteristics. It includes three relations and three attributes. In two of the three relations, individuals rank ordered each other based on how much time they spent together (once in the second week of the workshop, and once again in the third week). In the third relation, only the top three people each participant spent time with are included in the relation. In addition, gender, role (Student or Instructor), and position (Faculty, Post-Doc, and PhD Student) were also recorded.

2. Data Collection

Respondents were presented a deck of 18 randomly shuffled cards representing the 18 members of the group. Respondents were asked to find the card with their name on it, and put it face down on a table. They then found the card representing the person with whom they spent the most time during the last week, and placed that card on top of the last one. This was repeated for all members of the group, ending with the card corresponding with the person they interacted with the least. This data collection procedure was repeated at the end of each week of the 3-week workshop (but the first week's data were lost, so only the 2nd and 3rd week's data are available).

The data were entered as a rank order data matrix such that, for each respondent, the person they said they spent the most time with was coded as a 1, the next person as a 2, incrementing through the person they said they spent the least time with coded as a 17. A 0 was used for the diagonal.

3. Data Files and Formats

All the data are provided in one Excel Workbook, called **Camp92 Dataset.xlsx**, containing 4 worksheets (tabs).

Three worksheets contain the relational data. For the purposes of introducing various possible data formats, we provide each network file in a different format, though typically relational data will use only one of these formats.

- Week_2. Valued rank-order data are provided in a full matrix format. The value of the cell (i,j) indicates the rank that respondent i assigned to person j. A value of 1 indicates that i reported spending more time with person j than with any other person. A value of 17 indicates that i reported spending less time with person j than with any other person.

- Week_3. Valued rank-order data are provided in edgelist format. The edgelist format includes one line of three columns for each pair of nodes. The first column indicates the respondent, the second column indicates the person the respondent responded about, and the third column indicates the value the person in the first column assigned to the person in the second column for the relation. In these data, a row with the values i, j, and k means respondent i said s/he spent the kth most time with person j.

- Campnet. This is a dichotomized version of the Week_3 matrix, such that all values less than or equal to 3 were recoded to 1, and all other values were recoded to 0. Thus, this matrix gives the top 3 interaction partners for each respondent. These data are presented in nodelist format. This format, which can only represent dichotomous data, is a multicolumn format where the first column is the respondent and all subsequent columns are people nominated by the respondent in response to a dichotomous relationship (e.g., "Is this person one of your top three interaction partners?"). Thus, the row i, a, b, c means respondent i said people a, b, and c were his/her top 3 inter action partners, in no particular order.

- Campattr. As these are attribute data, they are also presented in full matrix format, where cell (i,j) is respondent i's value for attribute j. The following three attributes are provided, forming an 18x3 matrix:

  - *Gender*. The gender of each person, where 1 = female and 2 = male.

  - *Role*. The person's part in the workshop, where 1 = participant and 2 = instructor.

  - *Position*. The person's academic position, where 1 = faculty, 2 = post-doc and 3 = phd student.

4. Data Details

Response Rate	100%
Non-Respondent Bias	N/A
Theoretical Grouping	These data were collected in a methods workshop specifically to illustrate the methods. So, no real theoretical objectives.
Publications Using These Data	• Borgatti, SP, Everett, MG, Johnson, JC. 2013. Analyzing Network Data. Sage Publications. • Data are also distributed with UCINET software package as datasets camp92, campnet, and campattr.
Data Context	Research methods workshop/retreat
Respondents	Social science faculty and graduate students
Longitudinal	Yes, two points in time one week apart
Temporality	Low. Nothing about the data, collection, or context suggests the validity of the data will attenuate over time.
Analytical or Pedagogical Utility	• Demonstrating how valued graphs can be dichotomized • Illustrating homophily with categorical attributes • Finding cohesive subgroups • Contrasting with core/periphery network structures • Limitations of rank-ordering as a measurement method
Known Issues	None