Interrelationship Digraphs and Sociograms Using SAS®

Melvin Alexander – Social Security Administration

ABSTRACT

This presentation is about two important visualizations: one used in quality management and planning (MP), the other for studying patterns of interactions between people in groups. The two tools are the Interrelationship Digraphs and Sociograms.

The Interrelationship Digraph (ID) or Relations Diagram is designed to clarify intertwined, causal relationships among a group of items, issues, problems, or opportunities. This tool helps analysts gain insights into potential complex relationships of root causes that may underlie recurring problems despite efforts to resolve them.

The Sociogram is a graphic representation of the connections or links that a person has with other persons in a social network. Sociograms are useful in analyzing interrelationships between people such as: determining which individuals direct most of their comments to others or the group, who interrupts others, which person is the center of attention, how members of the group are connected, and other observable, interpersonal relationships.

IDs and Sociograms are specialized applications of Graph Theory, and Social Network Analysis where connections and interactions between objects and systems are made.

This presentation will give some background about IDs; describe how to construct sociograms and the two standard forms of IDs using base SAS®, SAS/IML®, the %DS2CONST macro; and conclude with final notes and thoughts.

INTRODUCTION

This presentation is about two important visualizations: one used in quality management and planning (MP), the other for studying patterns of interactions between people in groups. The two tools are the Interrelationship Digraphs and Sociograms.

The Interrelationship Digraph (ID) or Relations Diagram tool. The ID is an important quality management and planning (MP) tool designed to clarify intertwined, causal relationships among a group of items, issues, problems, or opportunities. This tool helps analysts gain insights into potential complex relationships of root causes that may underlie recurring problems despite efforts to resolve them.

The ID is a special network visualization that consists of a set of nodes (or vertices) connected by arrows. Arrows show directional relationships between “source” (sender) nodes into “target” (receiver) nodes. This representation turns the ID into a Sociogram [1], i.e., a form of Social Network Analysis that shows the connections and interactions between items, objects, and systems.

This paper includes background information about IDs; a description of methods used in constructing two standard forms of IDs and Sociograms using SAS/IML and the %DS2CONST macro of base SAS; and conclusions.

BACKGROUND

The ID is one of the seven Management & Planning Tools (7-MP) originally developed in 1976 by the collaboration between the Union of Japanese Scientists and Engineers (JUSE) and the Society of Quality Control Technique Development. Shigeru Mizuno [2] published information about IDs and six other tools in 1988 as a collection of methods for engineering, economic planning, and management. The other New

QC tools are: Affinity Diagrams (a.k.a. KJ Method), Systematic (tree) diagrams, Matrix diagrams, Matrix Data Analysis, Process Decision Program Charts (PDPC), and Arrow diagrams.

IDs belong to a class of knowledge-based quality tools used in identifying and processing ideas that lead to a better understanding of causal relationships within the “Define” phase of the Six Sigma Define, Measure, Analyze, Improve, Control (DMAIC) framework, see Hoerl and Snee [4].

IDs outdo the capability of Ishikawa (Fishbone, Cause-and-Effect) Diagrams. Ishikawa Diagrams assume that causal relationships between items are totally independent of each other, which usually is not the case. Analysts use IDs to determine which set of causes has the most influence on effects and how causes interact (or interrelate) with each other.

Since IDs are knowledge-based, they provide a preliminary step before data collection, experimentation, and root-cause verification process occurs in later MAIC Six Sigma stages
IDs graphically show the logical, causal relationships between several factors, issues, or ideas. They are useful in prioritizing choices when executives and organizations have difficulty reaching consensus, especially when credible data are not available.
The steps in creating IDs are:
(1) Knowledgeable experts from cross-functional departments agree on the issue to be resolved or question to be answered.
(2) Each member writes down the choices, factors, ideas, or items on index cards or sticky notes that affect the issue or question.
(3) The team arranges the items from step (2) in an octagonal or circular shape.
(4) Members pit each choice, or item against each other, individually or collectively. Next, they consider all combinations of the choices or items.
(5) For each pair of connected items, each member is asked which item influences (is stronger or causes) the other item and by how much on a degree of strength scale of 1 for weak or equal strength; 3 if the relationship has medium strength; or 9 if the relationship is strong or very strong.
(6) Members draw outbound one-headed arrows from "source" items to "target" items where a relationship between them exists or where one item influences the other item.
(7) Finally, the team adds the number of items that have the most inbound arrows (as key outcomes or results) and most outgoing arrows (as root causes or drivers).

IDs are presented in two basic formats: the Traditional Directed Graph (*Digraph*) and the Matrix format.

- The Traditional Directed Graph format shows how each item is connected to every item by arrows, either going out to (or coming in from) other items. The Traditional graphic format is visually more familiar to most people. The Traditional (Graphical) Format of the Interrelationship Digraph was produced using SAS Data steps and the %DS2CONST Constellation macro [5].
- The tabular Matrix Format ID (A*djacency matrix*) expresses connecting arrows between nodes of the Traditional ID elements by using the values “-1", "0", or “+1" to relate the adjacent row i and corresponding column j cells. Note that Bi-directional, two-way arrow connections between adjacent nodes could use the fourth value “±1” (or “0.5”), where arrows *go to* and *come from* row and columns items i and j. Although bi-directional connections are possible (e.g., Sociograms [1]), only one-way relationship arrows in the direction of the stronger cause or influence constructs the tabular IDs. The Matrix ID format tends to be more systematic and more manageable, especially when there are many other issues to consider.

An IML module, ID(K,W), expresses the relationships of issues in a matrix format using values of -1, 0, or +1 that denote "in" arrows (effects coming from column element j into row element i); “no relationship” between elements i, j; or "out" arrows (causes or drivers out of row element i into column element j) for the comparison pair, C_{ij}, consisting of elements i and j, respectively. Zero (0) is assigned to the C_{ij} pair when i=j (i.e., $C_{ii} = C_{jj} = 0$), where no connection between row element i and column element j exists, or from row item j into column item i.

ID(K,W) performs the ID calculations and returns an idmatrix output matrix.

PROC IML's matrix operators and functions produces the tabular matrix ID format.

The strength vector, c, is determined by computing the diagonal elements of the product of two matrices, absolute value of matrix k and the strength of relationship matrix w.

The Vecdiag function returns the main diagonal of the weighted sum of the products of the strength and adjacency matrices. If the values of the strength of relationships matrix are equal, then the strength matrix (W) has values of one (1), as expressed using the J function as:

W = J(nrow(K),ncol(K),1);

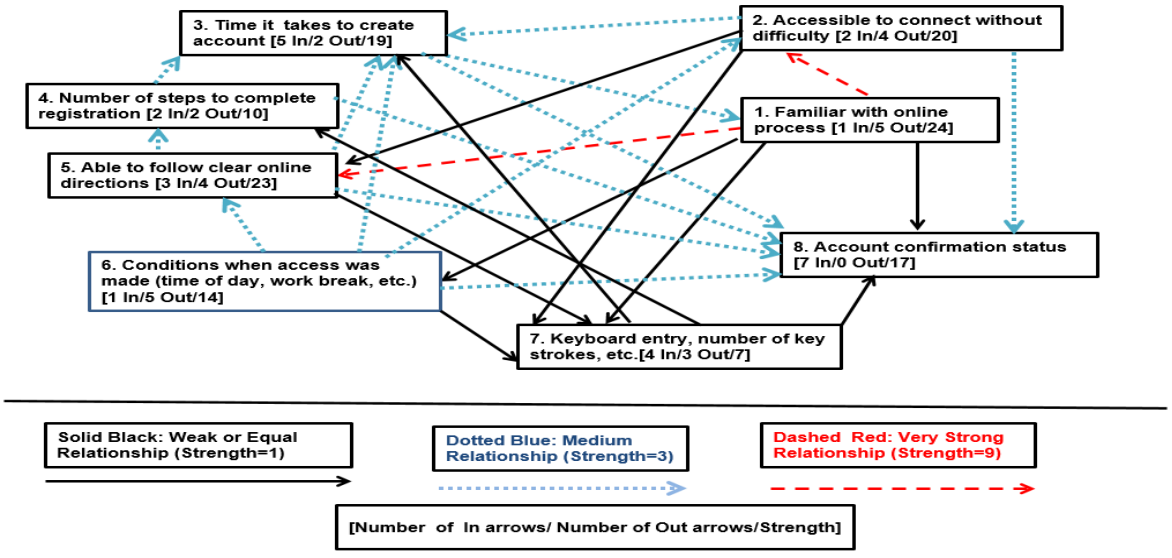
This results in the strength vector being the sum of "Total In(-1)" and "Total Out(+1)" columns in the output idmatrix. More information about the Relations Diagram and other examples can be found in the references.

EXAMPLES

Example 1: The Interrelationship Digraph

Example 1 from Alexander [6]: An ID was applied to the situation where a public service organization wanted customers to minimize office visits by getting them to set up and use online services. A Quality improvement team produced an Interrelationship Digraph to identify and understand the factors thought to be important in the successful creation of online accounts. Figure 1 shows the Traditional Format ID the team produced.

Figure 1: Interrelationship Digraph of Factors Affecting the Successful Creation of Online Accounts



The team listed eight factors as important for successful online account creation. The factors were:

- Customers being familiar with the online process (Factor 1),
- Customers Access to connect online without difficulty (Factor 2),
- The time it takes to create online accounts (Factor 3),

- Number of steps to complete registration (Factor 4),
- Customer's ability to follow clear online directions (Factor 5),
- Conditions when access was made, e.g., during work break, time of day, etc. (Factor 6),
- Keyboard entry, number of keystrokes, etc. (Factor 7), and
- Getting Account confirmation status (Factor 8).

The team determined that Familiarity with the online process (Factor 1) had a very strong, causal influence on accessibility to connect without difficulty (Factor 2) and the ability to follow clear online directions (Factor 5) – indicated by the dashed red arrows. Thirteen medium strength relationships (dotted blue arrows) were established between factors. Ten equal strength relationships (solid black arrows) were displayed.

Table 1 shows the Relations diagram in a tabular matrix format. Incoming and outgoing arrows are represented as "-1" and "+1" in the adjacent node pairs (connected factors) for row-item i and column-item j. No-connection-between-nodes is represented as "0". Factors with the sum of the incoming arrows and outgoing arrows are listed in the "Total In(-1)" and "Total Out(+1)" columns, respectively. The "Strength" column vector was computed using PROC IML's matrix product operations of the interrelationship and strength-of-relationship matrices.

Table 1: Tabular Matrix Format

Online Account Interrelations Digraph: Factors Affecting Successful Online Account Creation

Factor	1.Familiar with online process	2.Accessible to connect without difficulty	3.Time it takes to create account	4. Number of steps to complete registration	5.Able to follow clear online directions	6.Conditions when access was made(workbreak, time of day, etc.)	7.Keyboard entry, number of keystrokes, etc.	8.Account confirmation status	Total In (-1)	Total Out (+1)	Strength
1.Familiar with online process	0	9 1	3 -1	0	9 1	1 1	1 1	1 1	1	5	24
2.Accessible to connect without difficulty	9 -1	0	3 1	0	1 1	3 -1	1 1	3 1	2	4	20
3.Time it takes to create account	3 1	3 -1	0	3 -1	3 -1	3 -1	1 -1	3 1	5	2	19
4. Number of steps to complete registration	0	0	3 1	0	3 -1	0	1 -1	3 1	2	2	10
5.Able to follow clear online directions	9 -1	1 -1	3 1	3 1	0	3 -1	1 1	3 1	3	4	23
6.Conditions when access was made (workbreak, time of day, etc.)	1 -1	3 1	3 1	0	3 1	0	1 1	3 1	1	5	14
7.Keyboard entry, number of keystrokes, etc.	1 -1	1 -1	1 1	1 1	1 -1	1 -1	0	1 1	4	3	7
8.Account confirmation status	1 -1	3 -1	3 -1	3 -1	3 -1	3 -1	1 -1	0	7	0	17

Relationship Strength Legend for Symbols: 9 = 9 = Strong relationship; 3 = 3 = Medium Relationship; 1 = 1 = Weak or Equal Relationship;
"-1" means *incoming* arrow into row i from column j ; "0" means no connection between row i and column j;
"+1" means *outgoing* arrow from row i to column j; Strength value of 24 = 2 x 9 + 1 x 3 + 3 x 1.

Note: **Familiarity with the online process (Factor 1) was the key driver** because it had a maximum strength value of 24 and the most outgoing arrows ("Total Out(+1)" = 5). **The key indicator or effect was Account confirmation status (Factor 8)** having the most incoming arrows("Total In(-1)" = 7).

The matrix format of Table 1 and graph in Figure 1 show that Familiarity with the online process (Factor 1) was a very strong driver to Accessibility to connect without difficulty (Factor 2) and being Able to follow clear directions (Factor 5). Familiarity with the online process and the Conditions when access was made

(Factor 6) had the most outgoing arrows ("Total Out(+1)" = 5). The key indicator or effect was Account confirmation status (Factor 8) having the most incoming arrows ("Total In(-1)" = 7). The "key driver" among all factors was Familiarity with the online process (Factor 1) since it had the most outgoing arrows ("Total Out(+1)" = 5) and the maximum "Strength" value of 24. Pilot study results showed that emphasis on customer's familiarity with the online process, on-line access conditions, and ability to follow clear instructions affected other factors. Both factors served as sources of quality or performance metrics. Account Confirmation status had seven input arrows. Confirming customer account status indicated the magnitude of online service concerns that helped determine satisfaction levels of online services. In addition, item nodes can be grouped with colors that indicate three different group memberships. For example, Factors 1, 2, and 3 could represent an "Account Access" Group (in **Orange**). Factors 4, 5, and 7 could refer to an "Account Set-up" Group (in **Green**). Factors 6 and 8 could be classified as the "Account Conditions and Status" Group (in **Violet**).

The interrelations_digraph (ID) expression takes the Adjacency matrix, k=[C(i,j)], and strength of relationship matrix, w, to form a full matrix (fullmatrx). Fullmatrx concatenates matrix k with the sum of incoming arrows (sumin), outgoing arrows (sumout), and the strength vector (c). The strength vector, c, is determined by computing the diagonal elements of the product of two matrices, absolute value of matrix k and the strength of relationship matrix w using the IML Vecdiag () function.

Note: The matrix k has similarities to directed path effects of structural equation models (SEMs). The objective of SEM is to minimize differences between the observed covariance matrix and the implied structure of the model where the covariance of the data represents how connections between two or more nodes are related. The strength relationship matrix w would represent latent (interaction) effects variables in SEMs. IDs have graphical similarities to SEMs used in Path Analysis and Directed Acyclic Graphs (DAGs) of Social Network Analysis.

Figure 2 shows the Traditional ID using SAS' %DS2CONST macro.

Figure 2: Interrelationship Digraph – Traditional Format (using SAS® %DS2CONST macro) Red: Very Strong Relationship Strength=9, Blue: Medium Relationship (Strength=3), Black: Weak or Equal Relationship (Strength = 1)

Color Key for nodes: **Orange referred to Account Access**; **Green represented Account Conditions and status**; **Violet denoted Account Set Up.**

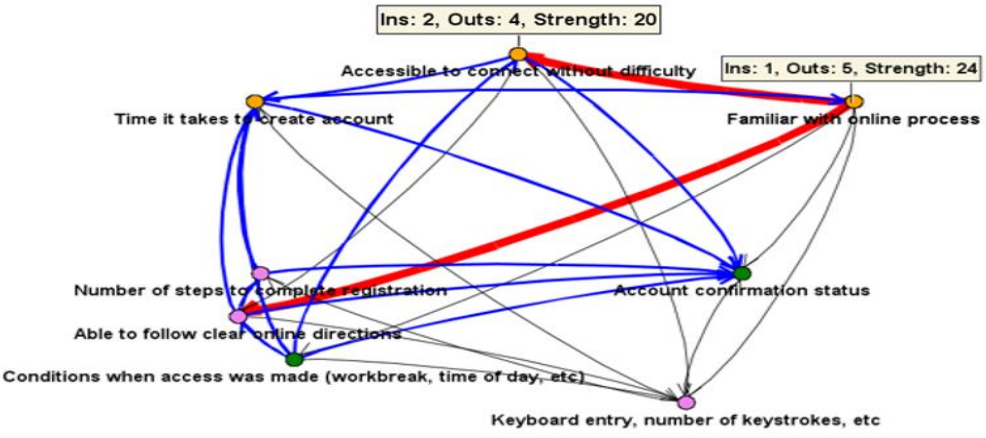


Table 3: NODEDATA Dataset for Example 2

NODEDATA

Filter and Sort | Query Builder | Data | Describe | Graph | Analyze | Export | Send To

	nshape	tname	fnwt	ncolor	nodeID	nlabel	value	tip	xlo	yloc	nsiz
1	circle	sansserif	bold	Orange	1	Familiar with online process	24	Ins: 1, Outs: 5, Strength: 24	600	100	10
2	circle	sansserif	bold	Orange	2	Accessible to connect without difficulty	20	Ins: 2, Outs: 4, Strength: 20	300	45	10
3	circle	sansserif	bold	Orange	3	Time it takes to create account	19	Ins: 5, Outs: 2, Strength: 19	65	100	10
4	circle	sansserif	bold	Violet	4	Number of steps to complete registration	10	Ins: 2, Outs: 2, Strength: 10	70	300	10
5	circle	sansserif	bold	Violet	5	Able to follow clear online directions	23	Ins: 3, Outs: 4, Strength: 23	50	350	10
6	circle	sansserif	bold	Green	6	Conditions when access was made (workbreak, time of day, etc)	14	Ins: 1, Outs: 5, Strength: 14	100	400	10
7	circle	sansserif	bold	Violet	7	Keyboard entry, number of keystrokes, etc	7	Ins: 4, Outs: 3, Strength: 7	450	450	10
8	circle	sansserif	bold	Green	8	Account confirmation status	17	Ins: 7, Outs: 0, Strength: 17	500	300	10

Table 4: LINKDATA Dataset for Example 2

LINKDATA

Filter and Sort | Query Builder | Data | Describe | Graph | Analyze | Export

	ltip	from	to	lcolor	value
1	Red: Very Strong Relationship (Strength=9)	1	2	Red	4
2	Red: Very Strong Relationship (Strength=9)	1	5	Red	4
3	Black: Weak or Equal Relationship (Strength = 1)	1	6	Black	1
4	Black: Weak or Equal Relationship (Strength = 1)	1	7	Black	1
5	Black: Weak or Equal Relationship (Strength = 1)	1	8	Black	1
6	Blue: Medium Relationship (Strength=3)	2	3	Blue	2
7	Black: Weak or Equal Relationship (Strength = 1)	2	5	Black	1
8	Black: Weak or Equal Relationship (Strength = 1)	2	7	Black	1
9	Blue: Medium Relationship (Strength=3)	2	8	Blue	2
10	Blue: Medium Relationship (Strength=3)	3	1	Blue	2
11	Blue: Medium Relationship (Strength=3)	3	8	Blue	2
12	Blue: Medium Relationship (Strength=3)	4	3	Blue	2
13	Blue: Medium Relationship (Strength=3)	4	8	Blue	2
14	Blue: Medium Relationship (Strength=3)	5	3	Blue	2
15	Blue: Medium Relationship (Strength=3)	5	4	Blue	2
16	Black: Weak or Equal Relationship (Strength = 1)	5	7	Black	1
17	Blue: Medium Relationship (Strength=3)	5	8	Blue	2
18	Blue: Medium Relationship (Strength=3)	6	2	Blue	2
19	Blue: Medium Relationship (Strength=3)	6	3	Blue	2
20	Blue: Medium Relationship (Strength=3)	6	5	Blue	2
21	Black: Weak or Equal Relationship (Strength = 1)	6	7	Black	1
22	Blue: Medium Relationship (Strength=3)	6	8	Blue	2
23	Black: Weak or Equal Relationship (Strength = 1)	7	3	Black	1
24	Black: Weak or Equal Relationship (Strength = 1)	7	4	Black	1
25	Black: Weak or Equal Relationship (Strength = 1)	7	8	Black	1

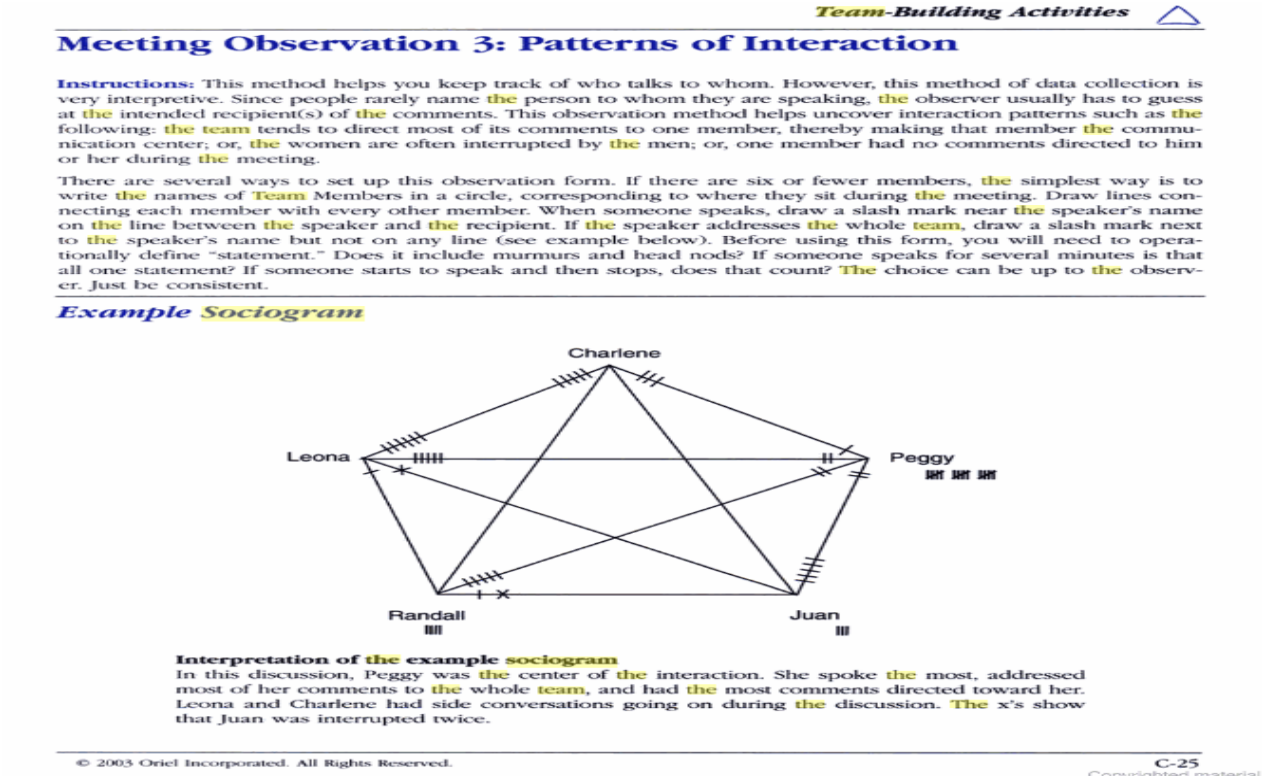
Example 2: The Sociogram

The Sociogram is a special type of interrelationship digraph that reveals relationships among people within a group that allows two-directional links between nodes. Jacob L. Moreno [7] developed sociograms as a way of showing people's connections and interactions with each other. His approaches evolved into the modern data and social science disciplines known as network analysis (computer science) and graph theory (mathematics). Sociograms, like most network graphs, are defined as: G=*{N,E,W}*, where *N* denotes a set of **nodes** (objects or vertices such as people, organizations, transactions, etc.). *E* represents **edges** (links or arcs) that connect between nodes. *W* represents weights that measure the level of strength that edges have between connected nodes.

Sociograms are typically used to identify and understand the dynamics and interactions between different nodes or other entities such as: (1) identifying the most vocal or influential individuals in team or group meetings; or 2) uncovering the emotional and behavioral dynamics between students in classrooms.

The example in Figure 3 came from Scholtes, Joiner, and Streibel [8]. Tally marks next to the speaker's name, away from the connecting lines (arc links), represent the number of times a speaker addressed the group. Slash marks on the arc links near the speaker's name shows the number of times the speaker addressed the recipient. The x's show that the speaker interrupted the recipient. Nodes that have the largest frequency of slash marks on the network indicates the individuals who have dominant, stronger, influence or weight over other group members they interact. The example shows that Peggy spoke more times (22) than any other group member, followed by Leona who spoke 13 times to Charlene, Randall, Peggy, and Juan.

Figure 3: Sociogram Example from Scholtes, Joiner,and Streibel, [8]
(Reprinted with permission from the Publisher's Owner Oriel STAT A MATRIX)

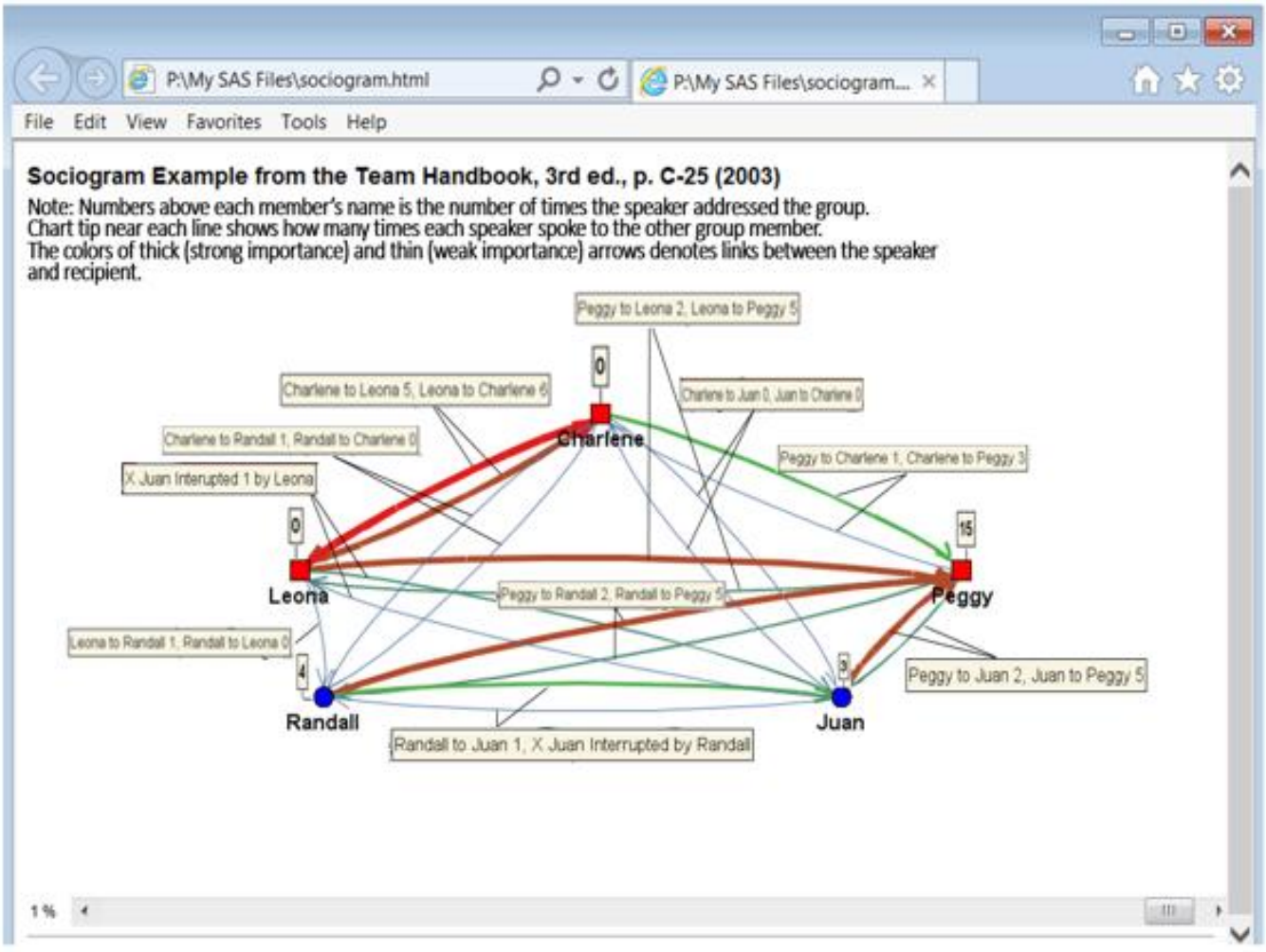


One objective of this example is to identify who are the most vocal and influential members of a group. The sociogram analyzed how group-member dynamics and interactions are studied analyzed and understood. My approach conducts this analysis using SAS/IML and %DS2Const macro. Sociograms are presented both graphically and in the matrix formats. Additional key social network metrics can be estimated.

Appendix 2 contained the SAS code that produced Figure 4. The nodedata datastep defined the variables about the meeting participants such as the individual's names, symbols distinguishing gender, the number of time each member spoke to the group, and the x-y coordinates for plotting the graph. The linkdata datastep created the weighted adjacency matrix. The weighted adjacency matrix is a square matrix in which cell values (linkvalue) of 0 indicated no connection between nodes (members). Linkvalues greater than zero represented the frequency of times each member (node) spoke to another member in the group. The chart tips, address_tip from nodedata, and tip from linkdata, displays the values whenever the pointer is placed over the node or edge of the graph.

Figure 4 shows the sociogram produced from the %DS2CONST macro.

Figure 4: Sociogram Produced by the %DS2CONST macro



CONCLUSIONS

IDs are special cases of Directed Acyclic Graphs (DAGs) used in Graph Theory and Social Network Analysis.

IDs are powerful tools that: (1) are useful in analyzing qualitative, "idea" data (i.e., data collected from experiences, opinions, creative thoughts of a team working towards solving a problem, answering a question, or developing a goal or mission); (2) transcend cause and effect diagrams by enabling teams to systematically depict complex relationships between different issues; and (3) force teams to establish logical, sequential connections between products, processes, and related ideas.

The PROC IML ID(K,W) module and %DS2CONST macro allows teams to produce combined Traditional Graphic and Matrix IDs in a more structured and organized manner that interrelate different ideas teams may likely face during six-sigma project planning.

The resulting report windows help teams focus on the more important causes and effects, enabling them to explore future quality improvement opportunities. Key drawbacks of IDs are that identified items may not be complete and are dependent on the knowledge of team members creating the ID. Therefore, having team members who are most knowledgeable; representing the diverse perspectives, viewpoints, and functions of the organization; and are familiar with the needs of stakeholders and customers will help minimize the shortcomings in constructing IDs.
IDs developed in the manner described above set the stage for forming "useful" models that establish functional relationships between process input causal factors and process effects (outcomes) that can be studied in later six sigma measure, analyze, improve, and control (MAIC) stages.

IDs lead into tracing the known reasons of causal relationships using designed experiments and other Lean Six Sigma (LSS) methods like Failure-Mode-and-Effects Analysis (FMEA), Fault Tree Analysis (FTA), Reliability Block Diagrams (RBD), 5-Whys, Root Cause Analysis (RCA), and Is-Is Not Analysis. For situations when the reasons for root causes are unknown, IDs are ideal for finding root causes and effects. Additional Network-relational and Statistical models relate explanatory variables to outcome measures. O'Malley and Marsden [9] discuss these models in more detail.

Other SAS tools that can be used to create Interrelations Digraphs include: SAS/GRAPH® Network Visualization Workshop and SGPLOT procedure (see Matange, [10]), SAS® OPTGRAPH and SAS/OR® OPTNET procedures.

The ID is a great tool to use in the planning stages of basic problem solving. They help teams focus on the key issues and confirm root causes before implementing solutions or rolling out new services.

Sociograms are useful tools that help teams understand group dynamics and interactions of meeting participants. Sociograms from different meetings help meeting facilitators identify the key influencers, "isolates", and "interrupters" so that meeting participants and focus groups can be more productive.

The variety of SAS software tools empowers users to build IDs and Sociograms with relative ease as part of a Social Network Analysis program.

REFERENCES:

1. Wikipedia, "Sociogram", https://en.wikipedia.org/wiki/Sociogram.

2. Mizuno, S. (1988), *Management for Quality Improvement: The 7 New QC Tools*, Cambridge, MA, Productivity Press, Inc.

3. Michalski, W.J. and D.G. King, D.G., (2005), *Six Sigma Tool Navigator ©: The Master Guide for Teams*, Tool 141, http://flylib.com/books/en/2.890.1.227/1/ .

4. Hoerl, R. and Snee, R. (2012), *Statistical Thinking: Improving Business Performance, 2nd ed.,* John Wiley & Sons: Hoboken, NJ, pp. 128 -138, 185-188.

5. SAS Institute, Inc., *Creating Constellation Diagrams*, http://documentation.sas.com/?docsetId=grapaxug&docsetTarget=p0bdoumann9sudn1ls65pzf4yrgf.htm&docsetVersion=9.4&locale=en .

6. Alexander, M. (2013), "When Little Objective Data Are Available, Find Root Causes and Effects with Interrelationship Digraphs and JMP®", *SESUG 2013: The Proceedings of the SouthEast SAS® Users Group*, St. Pete Beach, FL, 2013, http://analytics.ncsu.edu/sesug/2013/JMP-01.pdf.

7. Moreno, J.L. (1934), *Who Shall Survive? A new Approach to the Problem of Human Interrelations*, New York, NY, Beacon House.

8. Scholtes, Joiner, B.L., and B.J., Streibel, B.J. (2003), *The Team Handbook, 3rd edition*, Union, NJ, Oriel STAT A MATRIX, p. C-25.

9. O'Malley, A.J. and Marsden, P.V., (2008), "The Analysis of Social Networks", *Health Services Outcomes & Research Methodology*, 8(4), 222–269, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2799303/pdf/nihms95627.pdf.

10. Matange, S. (2016), "Outside-the-box: Directed circle link graphs", https://blogs.sas.com/content/graphicallyspeaking/2016/11/01/outside-box-directed-circle-link-graphs.

11. Darom, E. (1997), "Social Network Analysis with SAS/IML", http://www.sascommunity.org/seugi/SEUGI1997/DAROM_DATAMIN.PDF

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Melvin Alexander
Social Security Administration
Office of Analytics and Improvements - Division of Modeling
6401 Security Blvd.; East High Rise Building (4-A-5, Cube 4019)
Baltimore, MD 21235
Phone: (410) 966-2155
Fax: (410) 966-4337
E-mail: Melvin.Alexander@ssa.gov

DISCLAIMER

The views expressed in this presentation are the author's and do not represent the views of the Social Security Administration or SAS Institute, Inc.

ACKNOWLEDGMENTS

I thank Lucia Ward-Alexander for review and editorial assistance.

APPENDIX 1: Code Listing of PROC IML ID(K,D) module and %DS2CONST Macro to produce the Interrelationship Digraph

```
/************************************************************
  Interrelations digraph (ID) using proc iml

  Interrelations digraphs (a.k.a Relations diagrams or ID)
  show how different issues relate to one another based
  on cause and effect. IDs map out the ways in which each
  part of a system affects other system parts. IDs
  identify which parts are root causes or effects of the
  other parts.

  The ID also shows the strength of each influence.

  The IML procedure below expresses the relationships
  of issues in a matrix format using values of -1, 0, or +1
  to denote "in" arrows (effects from element j into element i);
  no relationship between elements i, j; or "out" arrows (causes
  or drivers from element i out to element j) of the comparison pair,
  Cij, of elements i and j. Zero (0) is assigned to the Cij pair
  when i=j (i.e., Cii = Cjj = 0).

  Author: Mel Alexander

  INPUT  K: Interrelationship matrix of Cij pairs
         W: Strength of relationships between Cij pairs
            Relationship strength has values
            Significant/Strong = 9
            Medium = 3
            Weak or equal = 1.

            If all comparison pair are equal, strength
            have values of 1s which expressed using the
            J function as: W = J(nrow(K),ncol(K));

  OUTPUT idmatrix: Concatenated matrix of the Cij pairs,
                   Ins, Outs, and Strength (Total) vectors.

 The ID(K,W) module performs the ID calculations and returns
 the idmatrix output matrix.

************************************************************/
proc iml;
start id(k,w) ;
/* create matrix of into (-1) arrows, out arrows (+1), and no
   arrows (0) of the comparison pair, Cij, of connected entities i and j
```

```
*/
k = k ;
w = w ;
/*  create matrix of strength of relationships, w, (where 0= no relationship,
      1 = weak or equal relationship, 3 = medium relationship,
      9 = strong relationship)
*/

p =abs(k) ;
weight = p * w ;
c= vecdiag(weight)  ; * the diagonals represent the strength of relationship
vector ;

sumin = J(1, nrow(k), 0 ) ;  * initialize row vector with zeroes  ;
sumout = J(1, nrow(k),0 ) ;  * initialize row vector with zeroes  ;
sumi = 0; sumo = 0 ; * initialize sum in(sumi) and sum out (sumo) to zero  ;
sumin = sumi + (k < 0 ) ; * sum all the ins (-1) ;
sumout = sumo + (k > 0 ) ; * sum all the outs (+1) ;
totalin = sumin[,+] ;   * sum all the ins vector ;
totalout = sumout[,+] ; * sum all the outs vector;

idmatrix = k||totalin||totalout||c ; *concatenate the ins, out, and strength
vectors ; ;

*print c[rowname=rnames colname=cnames label="Combination ID/Matrix Method"]
;
*print d[rowname=rnames colname=cnames label="ID Matrix"] ;
*print totalin[label="Ins"] ;
*print totalout[label="Outs"] ;


return (idmatrix);
finish id ;

/* Example A Public service organization wants customers to minimize
   office visits by getting them to set up online accounts. A
   Quality Improvement team brainstormed factors thought to
   affect users ability to successfully create online accounts.

   An Online Account Interrelationship Digraph was produced that
   showed cause-and-effect relationships between factors.
*/
k = {0  1 -1  0  1  1  1  1,  -1  0  1  0  1 -1  1  1,
     1 -1  0 -1 -1 -1 -1  1,   0  0  1  0 -1  0 -1  1,
    -1 -1  1  1  0 -1  1  1,  -1  1  1  0  1  0  1  1,
    -1 -1  1  1 -1 -1  0  1,  -1 -1 -1 -1 -1 -1 -1  0 };
w = {0 9 3 0 9 1 1 1,  9 0 3 0 1 3 1 3, 3 3 0 3 3 3 1 3,
     0 0 3 0 3 0 1 3,  9 1 3 3 0 3 1 3, 1 3 3 0 3 0 1 3,
     1 1 1 1 1 1 0 1,  1 3 3 3 3 3 1 0 };

idmatrix = id(k,w);
rnames={"1.Familiar with online process",
        "2.Accessible to connect without difficulty",
        "3.Time it takes to create account",
        "4. Number of steps to complete registration",
```

```
        "5.Able to follow clear online directions",
        "6.Conditions when access was made(workbreak, time of day, etc.)",
        "7.Keyboard entry, number of keystrokes, etc.",
        "8.Account confirmation status"} ;
combid ={"1.Familiar with online process",
         "2.Accessible to connect without difficulty",
         "3.Time it takes to create account",
         "4. Number of steps to complete registration",
         "5.Able to follow clear online directions",
         "6.Conditions when access was made(workbreak, time of day, etc.)",
         "7.Keyboard entry, number of keystrokes, etc.",
         "8.Account confirmation status",
         "Total In(-1)", "Total Out(+1)", "Strength"} ;
dsnames = {"Row","1.Familiar with online process",
         "2.Accessible to connect without difficulty",
         "3.Time it takes to create account",
         "4. Number of steps to complete registration",
         "5.Able to follow clear online directions",
         "6.Conditions when access was made(workbreak, time of day, etc.)",
         "7.Keyboard entry, number of keystrokes, etc.",
         "8.Account confirmation status",
         "Total In(-1)", "Total Out(+1)", "Strength"}
             ;
row = {1,2,3,4,5,6,7,8} ;
tradid = row || idmatrix ;
mattrib idmatrix rowname=rnames
          colname=combid
          label={"Online Account Interrelations Digraph: Factors Affecting
Successful Online Account Creation"}
          ;
print  idmatrix;

/*create trad2 dataset containing the ID matrix format variables  */
create trad2 from tradid [colname = dsnames]  ;
append from tradid;
close trad2 ;
quit;

/* use rnames as formats for column row */
proc format ;   value rowname
1 = "1.Familiar with online process"        2 = "2.Accessible to connect
without difficulty"
3 = "3.Time it takes to create account"   4 = "4. Number of steps to complete
registration"
5 = "5.Able to follow clear online directions" 6 = "6.Conditions when access
was made(workbreak, time of day, etc.)"    7 = "7.Keyboard entry, number of
keystrokes, etc."
8 = "8.Account confirmation status" ;  run;
data trad2; set trad2 ;  format row rowname. ; run;
title "Online Account Interrelations Digraph: Factors Affecting
Successful Online Account Creation" ;
proc print data= trad2 noobs split= '*' style(data) =
background=backcolor. foreground=forecolor. font_weight=bold}
```

```
style(header obs obsheader) = {backgroundcolor=light grey
color=black};
label Row = "Factor" '2.Accessible to connect without'n =
"2.Accessible to connect without difficulty"
'3.Time it takes to create accoun'n = "3.Time it takes to create
account" '4. Number of steps to complete r'n = "4. Number of steps to
complete registration" '5.Able to follow clear online di'n = "5.Able
to follow clear online directions" 6.Conditions when access was mad'n
= "6.Conditions when access was made(workbreak, time of day, etc.)"
'7.Keyboard entry, number of keys'n = "7.Keyboard entry, number of
keystrokes, etc." ;
run;

/*************************************************************
 Example 2: Constellation Chart with DATATYPE=ARCS using
            the %DS2CONST macro
*************************************************************/

/* Define the path to the JAR files and the HTML output file name */
%let jarfiles = file:///C:\Program Files\SASHome\SASGraphJavaApplets\9.4 ;
*%let jarfiles = path-to-archive-files;
*%let htmlfile = html-filename.htm;
*%let htmlfile = P:\My SAS Files\arctype1.html;
%let htmlfile = P:\My SAS Files\inter.html;
/* Define a nodes data set of the digraph */
data nodedata;
length nshape $8. tname $10. fnwt $4. ncolor $7. ;
input nodeID @3 nlabel $char62. @64 value  @67 tip $32. @100 xloc @103 yloc;
nshape='circle '; * node shape ;
nsize=10;  * node size ;
* Assign node color ;
if nodeID in (1,2,3) then ncolor = 'Orange' ;
else if nodeID in (6,8) then ncolor = 'Green' ;
else ncolor = 'Violet' ;
tname='sansserif'; * text font ;
fnwt='bold'; * font weight ;
*----+----1----+----2----+----3----+----4----+----5----+----6----+----7----+-
---8----+----9----+----c----+----0;
datalines;
1 Familiar with online process                                  24  Ins: 1,
Outs: 5, Strength: 24  600 100
2 Accessible to connect without difficulty                      20  Ins: 2,
Outs: 4, Strength: 20  300  45
3 Time it takes to create account                               19  Ins: 5,
Outs: 2, Strength: 19   65 100
4 Number of steps to complete registration                      10  Ins: 2,
Outs: 2, Strength: 10   70 300
5 Able to follow clear online directions                        23  Ins: 3,
Outs: 4, Strength: 23   50 350
6 Conditions when access was made (workbreak, time of day, etc) 14  Ins: 1,
Outs: 5, Strength: 14  100 400
7 Keyboard entry, number of keystrokes, etc                      7  Ins: 4,
Outs: 3, Strength:  7  450 450
```

```
8 Account confirmation status                                    17  Ins: 7,
Outs: 0, Strength: 17  500 300
;
run;

/* Define a links data set */
data linkdata;
length ltip $49. ;
input from to @6 lcolor $6. ;
/* define value based on the strength of relationships between source node
(from)
   and target node (to) where value = 1 denotes "weak"; value = 2 denotes
"medium";
   and value = 4 denotes "very strong"
*/
if (from = 1 and to = 2) or (from = 1 and to = 5) then value = 9 ;

if ((from = 1 and to = 6) or
    (from = 1 and to = 7) or
    (from = 1 and to = 8) or
    (from = 2 and to = 5) or
      (from = 2 and to = 7) or
      (from = 5 and to = 7) or
      (from = 6 and to = 7) or
      (from = 7 and to = 3) or
      (from = 7 and to = 4) or
      (from = 7 and to = 8))
  then value = 1 ;
if ((from = 2 and to = 3) or
    (from = 2 and to = 8) or
    (from = 3 and to = 1) or
      (from = 3 and to = 8) or
      (from = 4 and to = 3) or
      (from = 4 and to = 8) or
      (from = 5 and to = 3) or
      (from = 5 and to = 4) or
      (from = 5 and to = 8) or
      (from = 6 and to = 2) or
      (from = 6 and to = 3) or
      (from = 6 and to = 5) or
      (from = 6 and to = 8)
    ) then value = 3;
/* line color (lcolor) assigns color based on strength of relationships
   connecting source node (from) and target node (to) that associates
   the arrow tip (ltip) when the pointer hovers over the arrow.
*/
if lcolor = "Black" then ltip = "Black: Weak or Equal Relationship (Strength
= 1)" ;
if lcolor = "Blue" then ltip  = "Blue: Medium Relationship (Strength=3)" ;
if lcolor = "Red"  then ltip  = "Red: Very Strong Relationship (Strength=9)"
;
cards;
1  2  Red
1  5  Red
1  6  Black
```

```
1  7  Black
1  8  Black
2  3  Blue
2  5  Black
2  7  Black
2  8  Blue
3  1  Blue
3  8  Blue
4  3  Blue
4  8  Blue
5  3  Blue
5  4  Blue
5  7  Black
5  8  Blue
6  2  Blue
6  3  Blue
6  5  Blue
6  7  Black
6  8  Blue
7  3  Black
7  4  Black
7  8  Black
;
run;

goptions reset=all;

/* Close ODS HTML and open ODS LISTING */
ods html close;
ods listing;

/* Run the DS2CONST macro */


title1 "Interrelationship Digraph - Traditional Format (using DS2CONST macro)
";
*title2 "Red: Very Strong Relationship (Strength=9)" color=red ;
*title3 "Medium Strength Relationship (Strength=3)"  color=blue ;
*title4 "Black: Weak or Equal Relationship (Strength=1)" color=black ;


%ds2const(ndata=nodedata,
ldata=linkdata, datatype=assoc, codebase=&jarfiles,
htmlfile=&htmlfile, colormap=y, nid=nodeid, layout=user,
nlabel=nlabel, nx=xLoc, nfntname=tname,nfntstyl=fnwt,
ny=yLoc, nsize=nsize, ntip=tip, nshape=nshape,
lfrom=from, lto=to, lvalue=value, lwidth=value,
lcolor=lcolor, minlnkwt=1.6, ltip=ltip, linktype=arrow,
fntsize=17, height=540, width=960, cback=white,
ncolor=ncolor, cselect=red, tface=arial, fface=arial,
tcolor=black, fcolor=black, tsize=4, fsize=3, bg=white,
bgtype=color
);
run;
ods _all_ close;
```

```
run;


/* Close ODS LISTING and open ODS HTML */
ods listing close;
ods html;
```

APPENDIX 2: Code Listing of Base SAS and %DS2CONST Macro to produce the Sociogram

```
/* Define the path to the JAR files and the HTML output file name */
%let jarfiles =file:///C:\Program Files\SASHome\SASGraphJavaApplets\9.4;
%let htmlfile = P:\My SAS Files\sociogram.html; * use another output
destination location ;
/* Define nodes data set of labels, arrow in-out summaries, x-y  display
positions  */
data nodedata;
length nshape $8. tname $10. fnwt $4.;
*input nodeID @3 nshape $char62. @64 value  @67 tip $32. @100 xloc @103 yloc;
input nodeID @3 nodeLabel $char12. @15 nshape  $8. @24 ncolor $4. @29
address_team_tip $2. @33 xloc @37 yloc;
/* address_team_tip is the number of times the person (nodeID) spoke to the
group */
/* add node shapes, color, and font weight */
*nshape='circle ';
nsize=10;
*ncolor = 'Yellow' ;
tname='sansserif';
fnwt='bold';
datalines;  *n
1 Peggy    square   red   15  600  200
2 Charlene square   red    0  300   75
3 Leona    square   red    0   50  200
4 Randall  circle   blue   4   70  300
5 Juan     circle   blue   3   500 300
;
/* Define the link data. */
data linkdata;
length from to linkvalue 8 tip $60;
input from to linkvalue @10 tip $char60.;
cards;
1  2  1  Peggy to Charlene 1, Charlene to Peggy 3
1  3  2  Peggy to Leona 2, Leona to Peggy 5
1  4  2  Peggy to Randall 2, Randall to Peggy 5
1  5  2  Peggy to Juan 2, Juan to Peggy 5
2  1  3  Peggy to Charlene 1, Charlene to Peggy 3
2  3  5  Charlene to Leona 5, Leona to Charlene 6
2  4  0  Charlene to Randall 0, Randall to Charlene 0
2  5  0  Charlene to Juan 0, Juan to Charlene 0
3  1  5  Peggy to Leona 2, Leona to Peggy 5
3  2  6  Charlene to Leona 5, Leona to Charlene 6
3  4  1  Leona to Randall 1, Randall to Leona 0
3  5  1  X Juan Interrupted twice by Leona
```

```
4  1  5  Peggy to Randall 2, Randall to Peggy 5
4  2  0  Charlene to Randall 0, Randall to Charlene 0
4  3  1  Leona to Randall 1, Randall to Leona 0
4  5  2  Randall to Juan 1, X Juan Interrupted by Randall
5  1  5  Peggy to Juan 2, Juan to Peggy 5
5  2  0  Charlene to Juan 0, Juan to Charlene 0
5  3  1  X Juan Interrupted 1 by Leona
5  4  1  Randall to Juan 1, X Juan Interrupted by Randall
 ;
run;
ods listing;
Title1
 "Sociogram Example from the Team Handbook, 3rd ed., p. C-25 (2003) ";
%ds2const(ndata=nodedata,
ldata=linkdata, datatype=assoc, codebase=&jarfiles,
htmlfile=&htmlfile, colormap=y, nid=nodeid,
layout=user, nlabel=nodelabel, nx=xLoc,
nfntname=tname, nfntstyl=fnwt, ny=yLoc,
nsize=nsize, ntip=address_team_tip, nshape=nshape,
lfrom=from, lto=to, lvalue=linkvalue, lwidth=linkvalue,
minlnkwt=1.6, ltip=tip, linktype=arrow, fntsize=17,
height=540, width=960, cback=white, ncolor=ncolor,
cselect=red, tface=arial, fface=arial, tcolor=black,
fcolor=black, tsize=4, fsize=3, bg=white, bgtype=color
);
run;
ods _all_ close;
run;
```