

## Transforming data with a statistical procedure

This is an example of using a statistical procedure to transform a set of data. The details of the algorithm are (sometimes) documented in text, perhaps with mathematical formulas, but may be quite complex and the actual code procedure code may be proprietary.

The original data for this example are a set of self-reported distances, where each person is asked to rate how “far” they felt they were from each of the other people (0 from yourself). The original data might appear in two forms:

### A pairwise list:

Tom Dick 3  
Tom Sally 3  
Tom Harry 5  
Sally Dick 3  
Sally Harry 4  
Dick Harry 5

### A Matrix Representation:

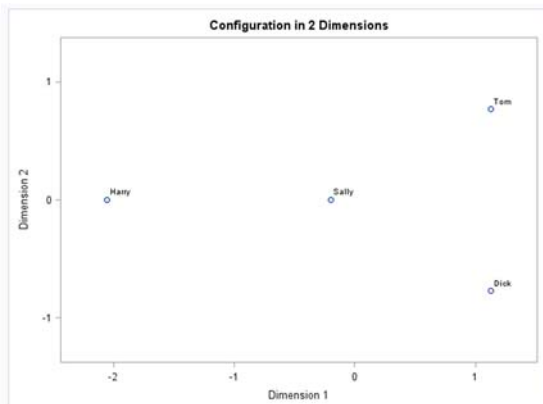
ID	Name	Tom	Dick	Harry	Sally
1	Tom	0	3	5	3
2	Dick	3	0	5	3
3	Harry	5	5	0	4
4	Sally	3	3	4	0

A common transformation for this type of network data is to generate some spatial representation. One approach is to use a Multidimensional Scaling procedure to reduce the dimensionality of the matrix representation. Applying the SAS MDS procedure requesting two resulting dimensions yields:

### Spatial Representation:

_NAME_	Dim1	Dim2
Tom	1.127904	0.771126
Dick	1.127915	-0.77112
Harry	-2.05582	3.07E-06
Sally	-0.2	-1.4E-05

### Plotted Data



## SAS Code for this Example:

```
/* initial data is self rated distances from each other on a 0 to 5 scale */
libname Mydata
'C:\DDRIVE\projects\various\DDI\DDI4\SimpleDataDescription\Transformations';

data Mydata.Distances;
input ID Name $ Tom Dick Harry Sally;
datalines;
1 Tom 0 3 5 3
2 Dick 3 0 5 3
3 Harry 5 5 0 4
4 Sally 3 3 4 0
;
run;

/* Use a multidimensional scaling to
transform the data into a 2 dimensional spatial representation
Use the default parameters other than the number of dimensions */

ods graphics on;
PROC MDS data=Mydata.Distances
out=Mydata.Spatial
DIMENSION=2
PLOTS;
var Tom Dick Harry Sally;
run;

Proc print data=Mydata.Spatial;
where _type_="CONFIG";
run;
```