



PhUSE US Connect 2018
AB06

Shaken but Not Stirred: an Example of Subject Classification using Multidimensional Scaling

////////

05Jun2018 / Manuel Sandoval / V1.0





Agenda

// **Multidimensional Scaling**

// Definition

// Dissimilarity Functions

// SAS code & output

// First example

// Second example

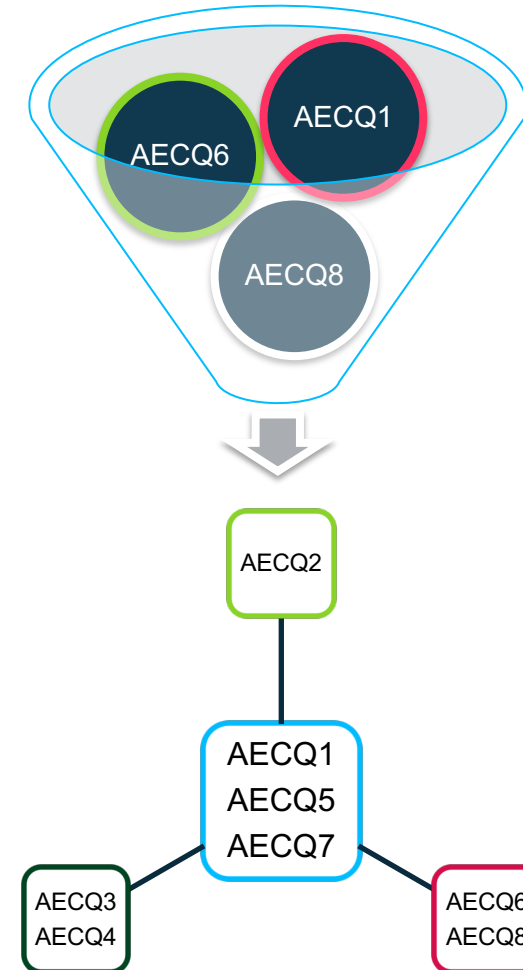


First example - Rationale

Safety – AE Classification

// We have a clinical trial with 8 Adverse events groups of special interest

// We want to look at similarities between them, to explore different groups



Multidimensional Scaling

Definitions

“the search for a low dimensional space, usually Euclidean, in which points in the space represent the objects [...], one point representing one object, and such that the distances between the points in the space match, as well as possible, the original dissimilarities”. [Cox and Cox, *Multidimensional Scaling*]



“a low-dimensional representation of the differences in a set of points.”





Multidimensional Scaling

Elements

// What do we need?

// A set of points to be classified

// A dissimilarity (or distance) between any two points to compare

// What do we want to do?

// A graphical representation in two dimensions of the dissimilarity.



Multidimensional scaling – Dissimilarity function

Any function that satisfies the following four properties can be considered a dissimilarity (or distance) function:

1. For any given point, the distance to itself is 0.
2. Given any two points, a and b , the distance from a to b is greater or equal to 0.
3. Given any two points, a and b , the distance from a to b is equal to the distance from b to a .
4. Given three points a , b and c , the distance from a to c is less than or equal to the sum of the distances from a to b and b to c .



Multidimensional Scaling

SAS Code

```
ods graphics on;
```

```
proc mds data = <datain> out = <dataout> pdata pfit similar;
```

```
by <byvar>;
```

```
var <matrix_vars>;
```

```
RUN;
```

```
ods graphics off;
```



Multidimensional Scaling

SAS Output

Multidimensional Scaling: Data=WORK.ADAE_CORR_USUBJID.DATA
Shape=TRIANGLE Condition=MATRIX Level=ORDINAL
Coef=IDENTITY Dimension=2 Formula=1 Fit=1

Mconverge=0.01 Gconverge=0.01 Maxiter=100 Over=2 Ridge=0.0001

Iteration	Type	Badness-of-Fit Criterion	Convergence Measures		
			Change in Criterion	Monotone	Gradient
0	Initial	0.3017	.	.	.
1	Monotone	0.1934	0.1083	0.2240	0.7793
2	Gau-New	0.1111	0.0823	.	.
3	Monotone	0.0956	0.0155	0.0454	0.4318
4	Gau-New	0.0920	0.003565	.	.
5	Monotone	0.0788	0.0133	0.0376	0.2792
6	Gau-New	0.0781	0.000621	.	.
7	Monotone	0.0769	0.001234	0.0130	0.2143
8	Gau-New	0.0764	0.000544	.	.
9	Monotone	0.0763	0.000107	0.003935	0.1749
10	Gau-New	0.0751	0.001105	.	0.0152
11	Gau-New	0.0751	7.909E-6	.	0.001728

MATRIX	Number of Nonmissing Data	Weight	Badness-of-Fit Criterion	Distance Correlation	Uncorrected Distance Correlation
1	28	1.00	0.08	0.97	1.00

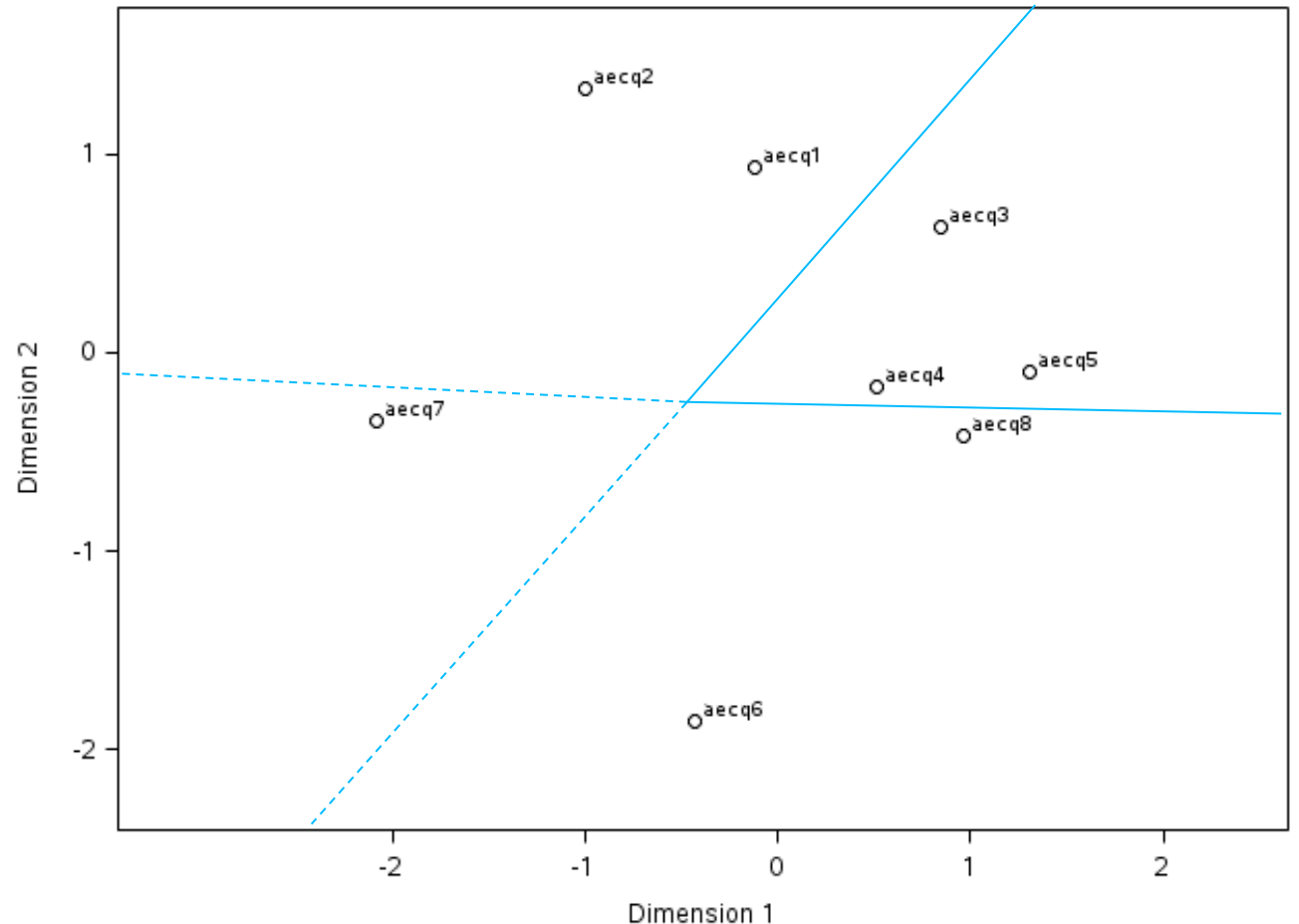
Convergence criteria are satisfied.



First example: Safety – AE Classification

Multidimensional scaling results, AEs summarized by subject

- // Summarized number of Adverse Events by subject
- // Dissimilarity function used: squared correlation coefficients
- // PROC MDS SIMILAR option used
- // LEVEL = ORDINAL option used
- // Groups 1 and 2 related to condition
- // Groups 3, 4 and 5 related to treatment
- // Groups 6, 7 and 8; expected conditions that required special care

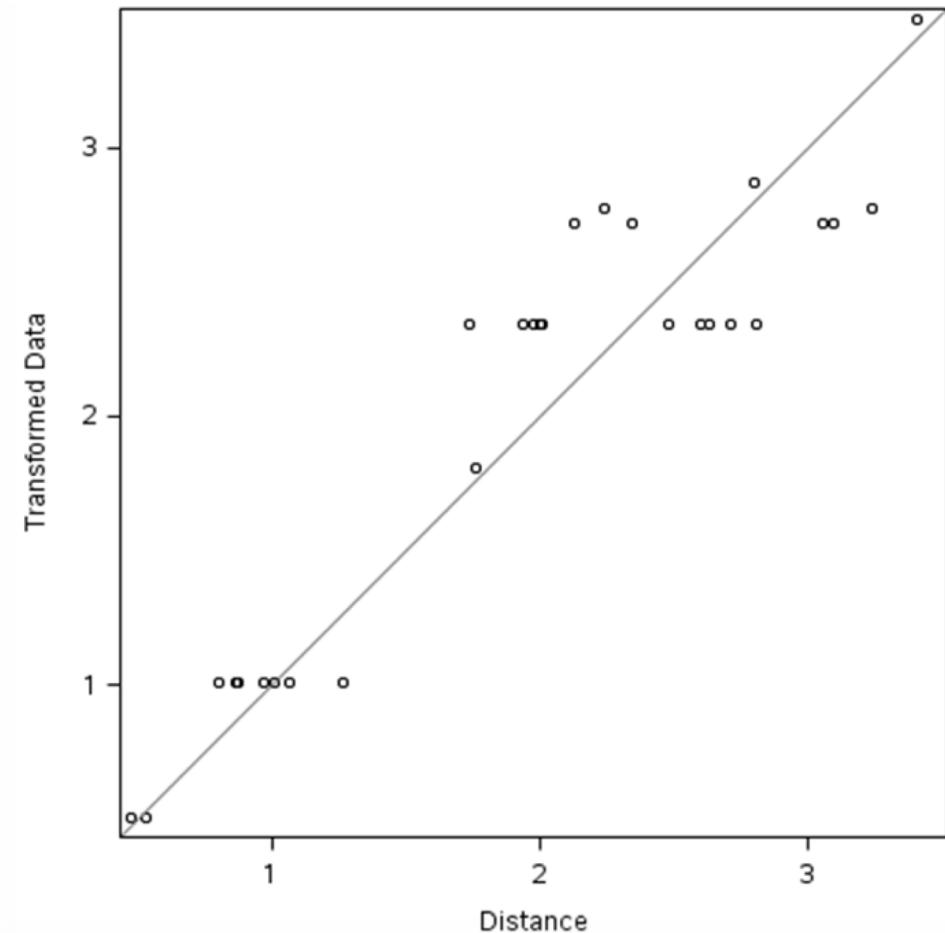




First example: Safety – AE Classification

Evaluation of results.

- // Graph included in ODS GRAPHICS SAS output
- // Actual distance vs. transformed data for 2D plot
- // Badness-of-fit statistic: 0.15
- // According to one criteria reviewed, it is a borderline good fit.

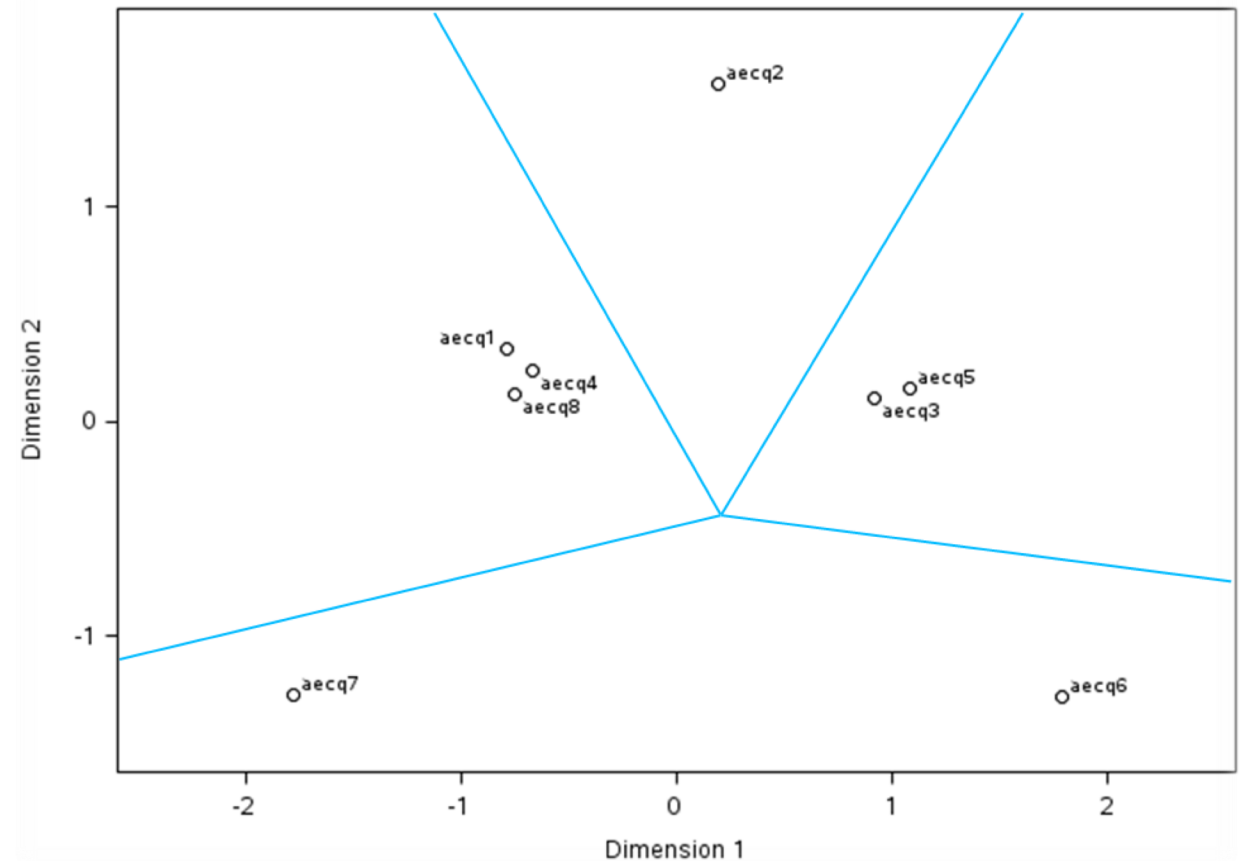




First example: Safety – AE Classification

Multidimensional scaling results, AEs summarized by condition severity

- // Summarized number of Adverse Events by condition severity at baseline
- // Partitions (unfolding) set arbitrarily
- // Several questions for further research
- // Badness-of-fit statistic: 0.04

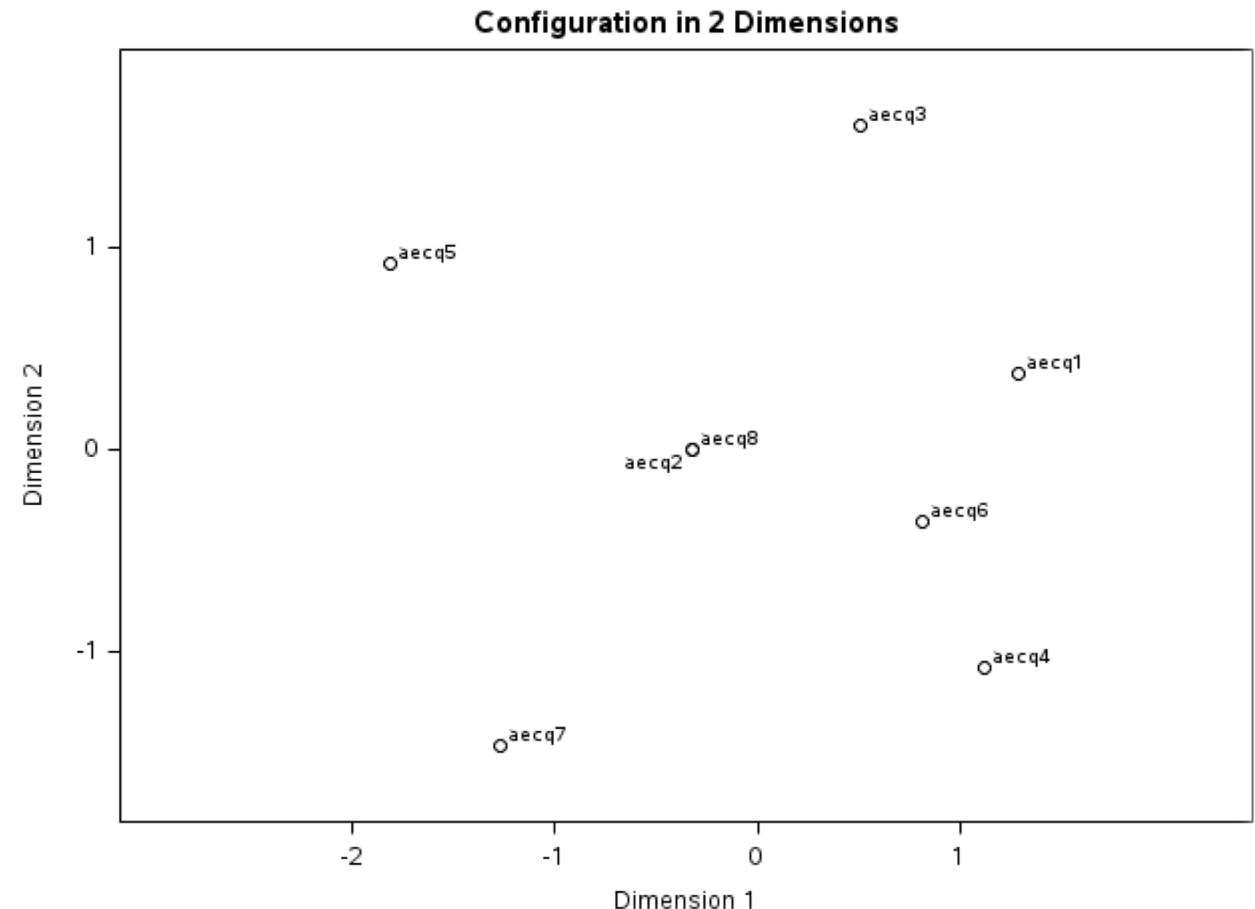




First example: Safety – AE Classification

Multidimensional scaling results, AEs summarized by condition severity

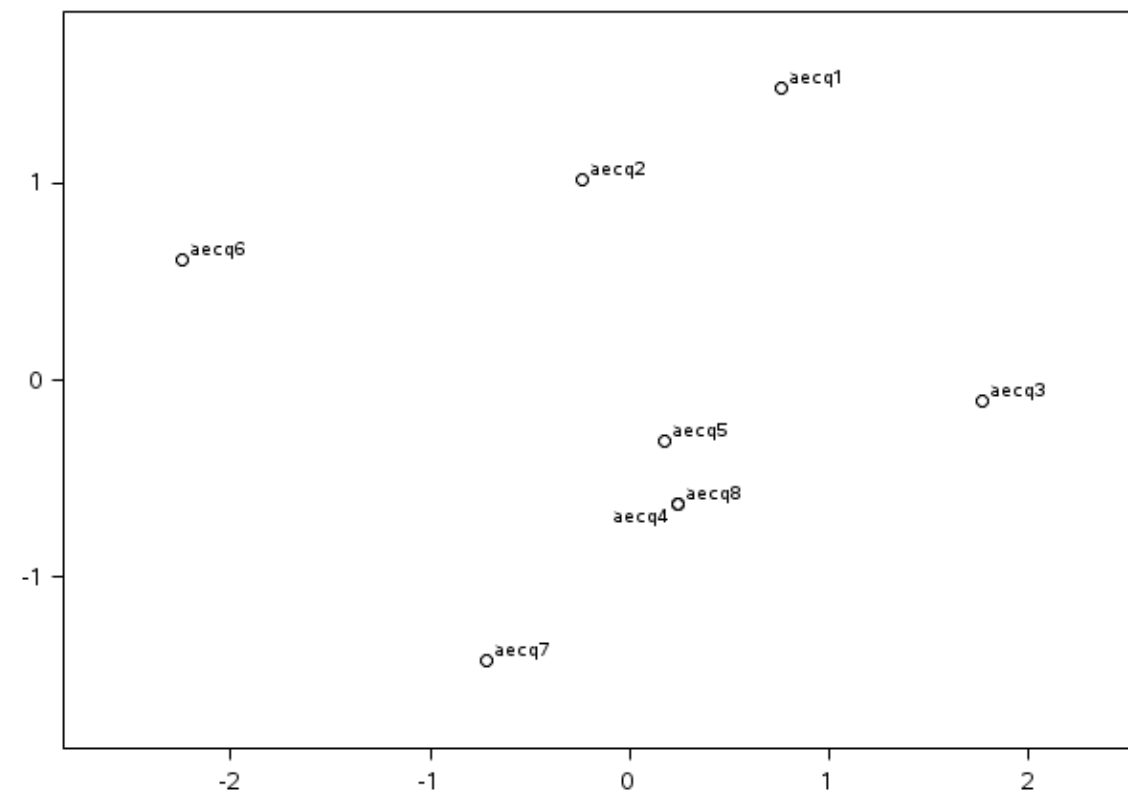
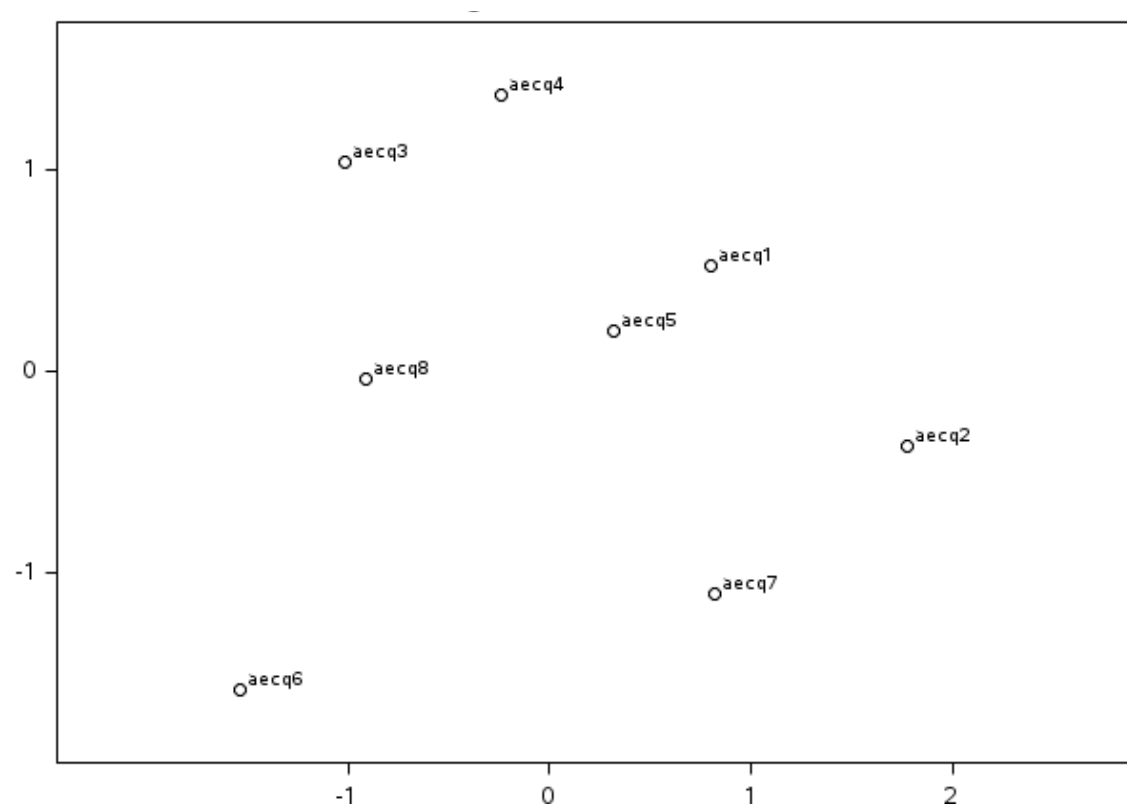
- // Summarized number of Adverse Events by treatment
- // Results between treatments were too similar, so most correlations were close to 1.
- // Additional analysis created, summarizing each treatment separately, by subject (next slide).
- // Two different graphs were needed because there was no dissimilarity measure between AEs from subjects with different treatments.





First example: Safety – AE Classification

Graphs by treatment





Second example: Efficacy – Endpoint result

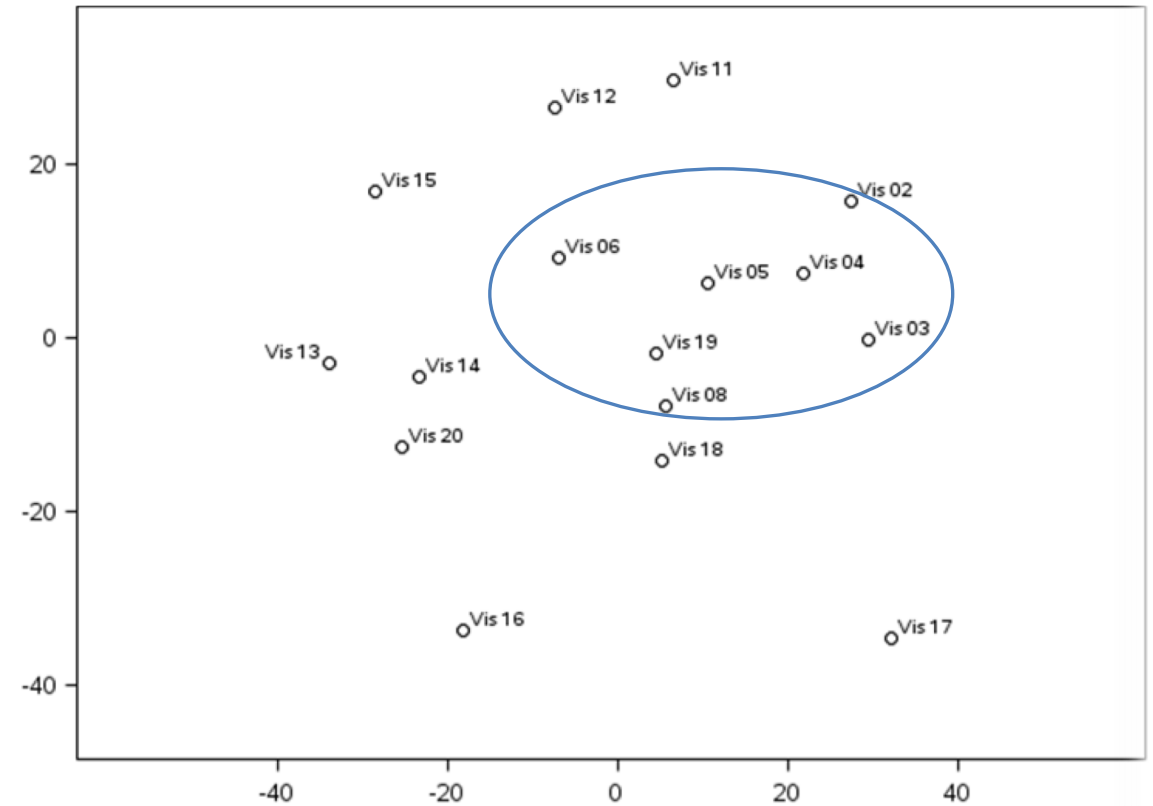
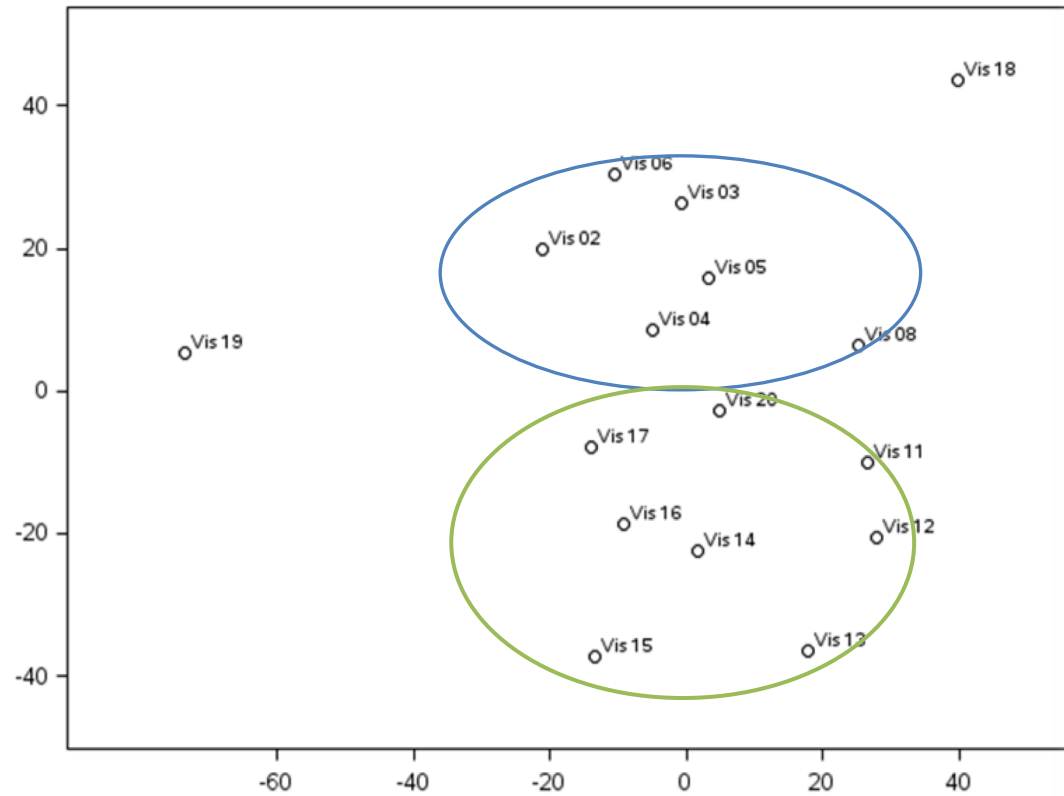
Visits Classification

- // We have a clinical trial with 20 visits, and a particular endpoint that we measure on all visits except 1, 9 and 10.
- // The first 10 visits correspond to a first phase of the study; the following 10 to a second, long-term extension phase.
- // Multidimensional scaling was used to visualize the similarities of improvement between each visit.
- // Dissimilarity function, absolute difference between the value for each couple of visits
- // Option LEVEL=ABSOLUTE in PROC MDS statement
- // Results separated by treatment



Second example: Efficacy – Endpoint result

Multidimensional Scaling Results - Visits

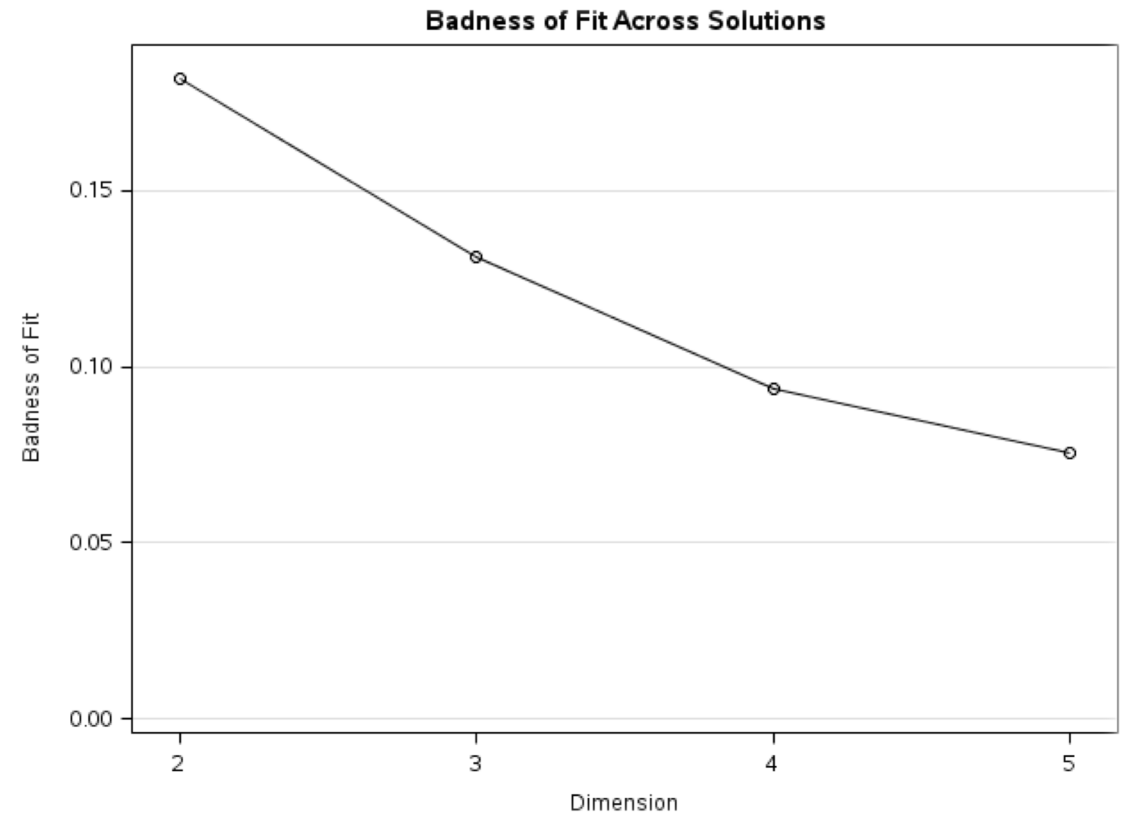




Second example: Efficacy – Endpoint result

Multiple dimensions

- // In the previous analysis, badness-of-fit is 0.14 (treatment 1) and 0.18 (treatment 2)
- // Increasing the dimensions decreases the badness-of-fit statistic
- // Using the option DIM = 2 to 5 BY = 1 in the PROC MDS statement, SAS creates the scaling for all dimensions and the figure to the right.

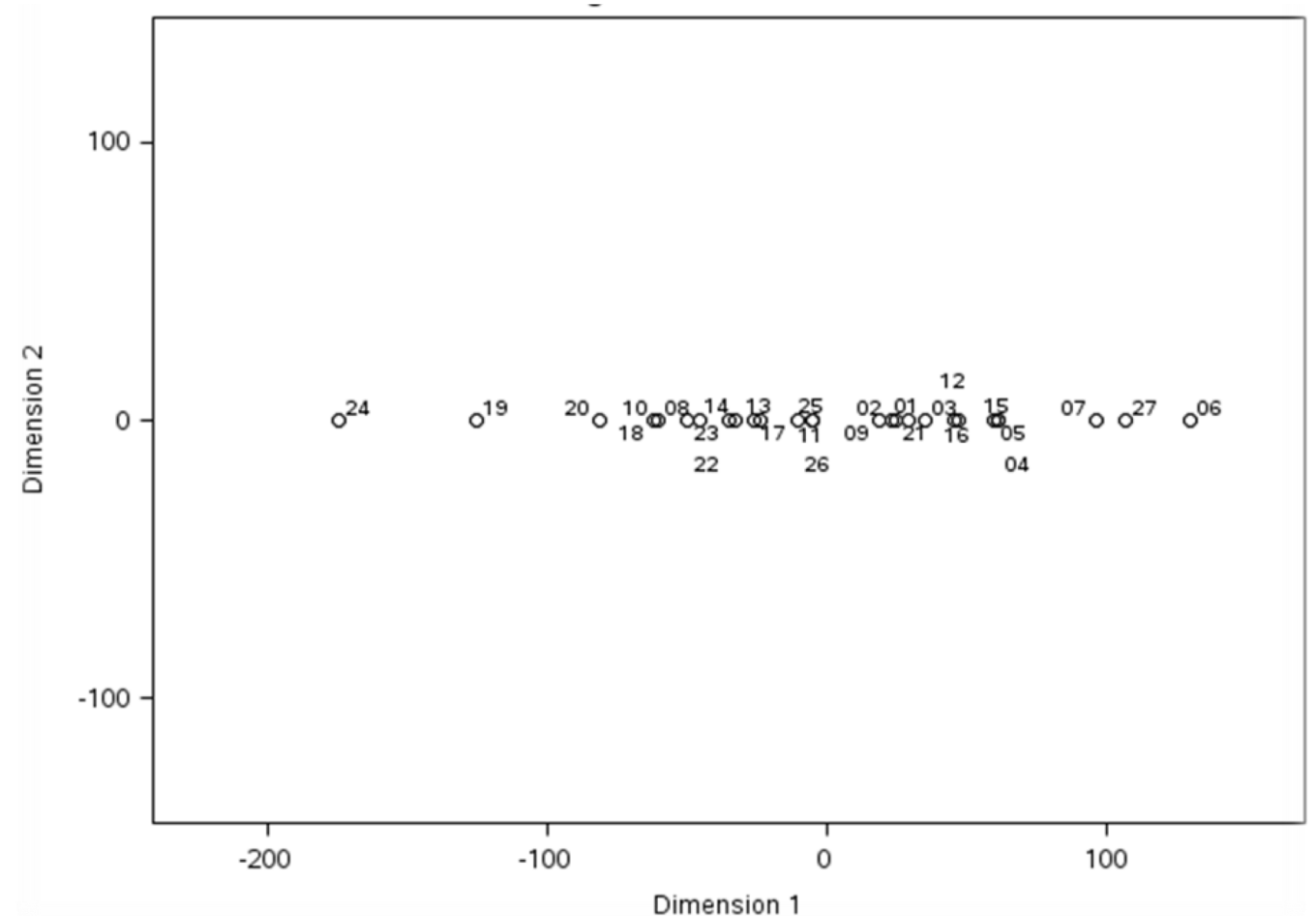




Second example: Efficacy – Endpoint result

Subject classification

- // Multidimensional scaling analysis on a subset of all study subjects, using their maximum endpoint result
- // Dissimilarity function: Absolute difference
- // To the right: SAS generated graph

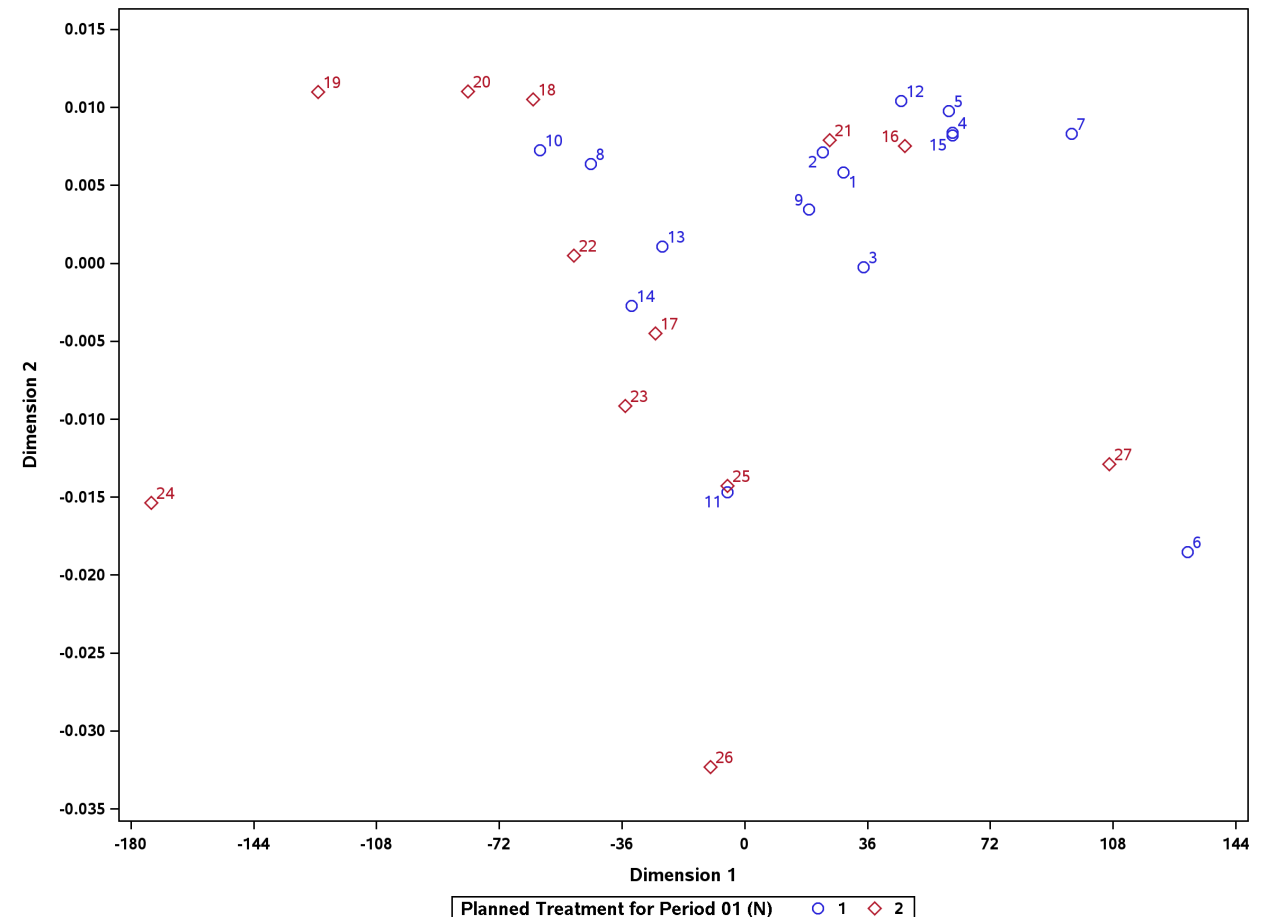




Second example: Efficacy – Endpoint result

Subject classification – Custom graph

- // Custom graph based on coordinates given by SAS output dataset
- // Treatments can be differentiated
- // Relative position shows difference between treatments, but we cannot say by this analysis which one is better: a mirror graph is equivalent





Final Thoughts

- // Multidimensional scaling is a powerful tool for screening differences.
- // To generate meaningful results, subject matter experts are needed, as usual.
- // This is only an introduction, there is a lot more there
- // Using Euclidean distance, Multidimensional scaling is equivalent to Principal components analysis



Thank you!

