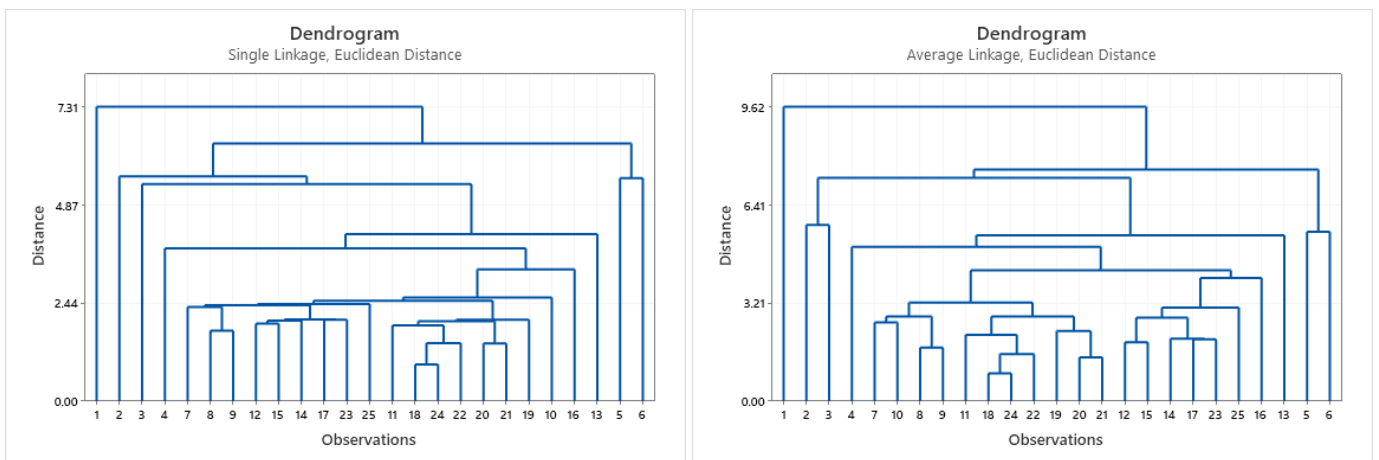


## Additional Multivariate Exercise 6

*Brief solution, for Steneryd data only*

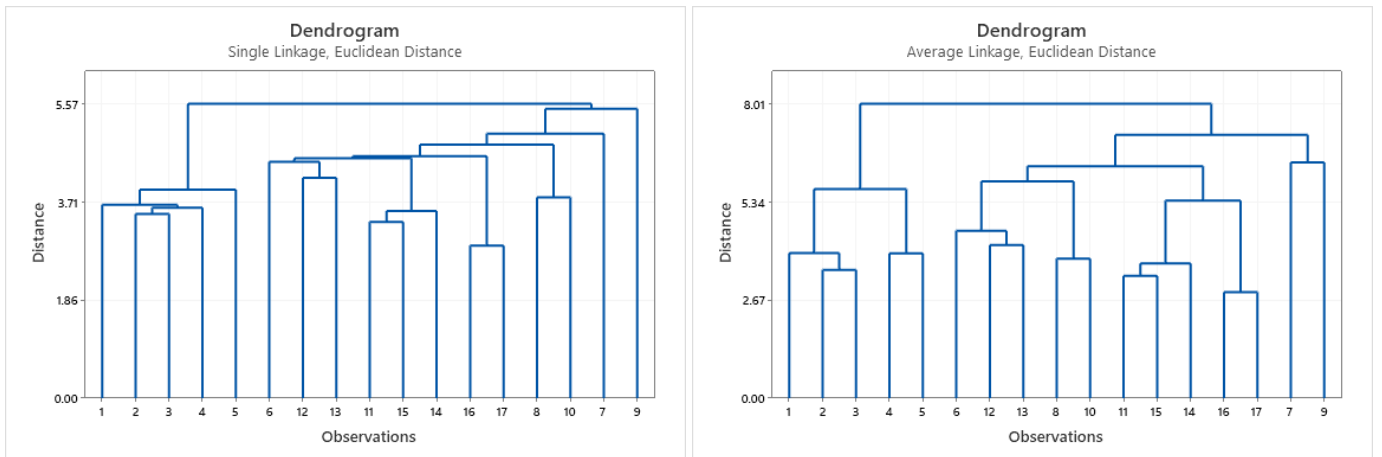
Data: The Steneryd data comprises indices for the 25 most abundant species observed in 1965 on 17 plots at a grazed meadow in the Steneryd Nature Reserve in Sweden (data source: Persson (1981), *Journal of Ecology* **69**, 71–84). The values represents the sum across 9 nine sample quadrants each scored by cover indices ranging 0–5 (where 5 ~ complete cover by the species in a quadrant). Note that the species are ordered by abundance and the plots by some of the environmental variables.

Hierarchical cluster analysis (by species, labelled 1–25): The dendrograms from two analyses, based on standardized variables and Euclidean distances, with single and average linkages performed in Minitab are shown below.

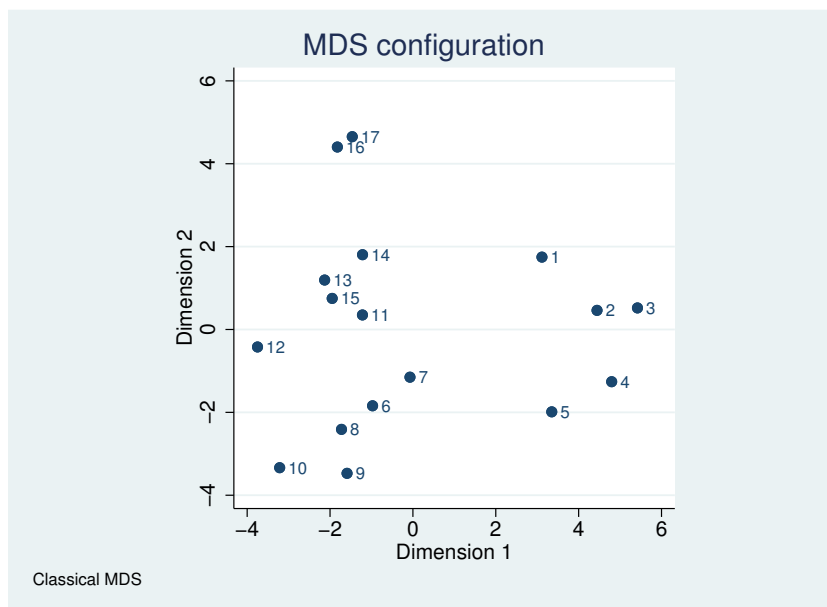


With both linkages, the most similar species are (20, 21) and the triplet (18, 22, 24). Both dendrograms also show species 1 being separated out as the first cluster (from the top); this seems very plausible based on the raw data where this species has the by far largest values for plots 1–7 and very low counts for the remaining plots — clearly contrasting the patterns amount the other species. Also the species 2–6 are separated out as individual or paired clusters in the first few steps of the dendrogram. This suggests that although the distances were based on standardized plot indices, the method gives much stronger weights to the frequently occurring species. In effect, this may be desirable, because it would seem more questionable to base clusters on very low indices (representing scarce occurrences). It might be of interest to explore how (non-linear) scale changes affect the patterns obtained among the species, but we stop here.

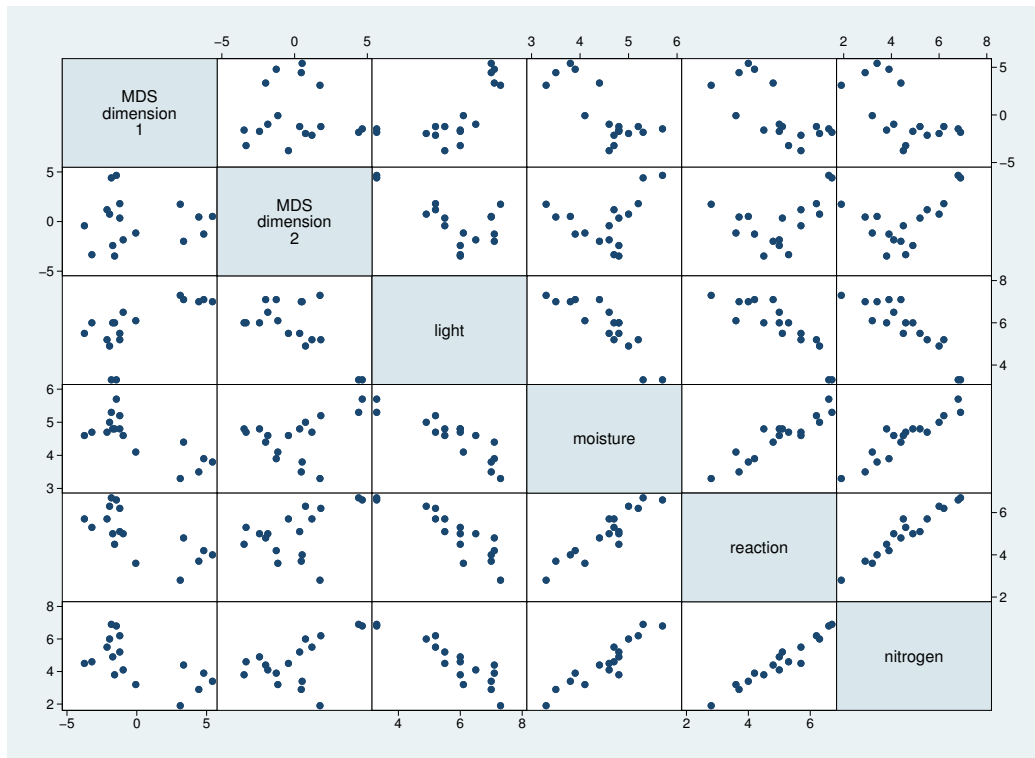
Hierarchical cluster analysis (by plots, labelled 1–17): Again we show, on the next page, the dendrograms from analyses with single and average linkages performed in Minitab. Despite somewhat different appearances of the two dendrograms, one major characteristic is common to both of them, namely that the plots are kept quite closely in their original order in the dendrogram. More precisely, plots close in the dendrogram are typically also close in the original numbering (which did reflect some environmental characteristics). It is however not obvious how to utilize the values of the environmental variables in the analysis, or how to incorporate them in the display.



Multidimensional scaling (by plots): This technique offers additional options to explore the relationships with the environmental variables. We illustrate a classical MDS with two components, obtained from Stata. Our results are equivalent to those shown in Table 12.1 of Manly for a principal components analysis.



The plot shows similar clustering of plots by their numbering. However, now we have the coordinates of the first two dimensions available, and we can explore their relationships with the environmental variables.



The scatterplot matrix shows that the first MDS component is positively correlated with light and negatively correlated with the three other environmental variables. The second MDS component seems negatively correlated with light, whereas the other associations are less strong. For these interpretations it is also helpful to compute the pairwise correlations between the variables. Note that the two MDS components are uncorrelated by their construction.

```
. pwcorr dim1 dim2 light moisture reaction nitrogen
```

	dim1	dim2	light	moisture	reaction	nitrogen
dim1	1.0000					
dim2	-0.0000	1.0000				
light	0.6719	-0.6339	1.0000			
moisture	-0.7569	0.2847	-0.8648	1.0000		
reaction	-0.6830	0.4116	-0.8591	0.9204	1.0000	
nitrogen	-0.5991	0.4826	-0.8820	0.9246	0.9673	1.0000

It is seen that the environmental variables are quite strongly correlated themselves (which is also apparent from the scatterplot matrix). Further refinements of the analysis could consist in looking at additional MDS components or trying modern MDS algorithms. For this solution we stop here.