Traceis™ Data Exploration Studio
Clustering

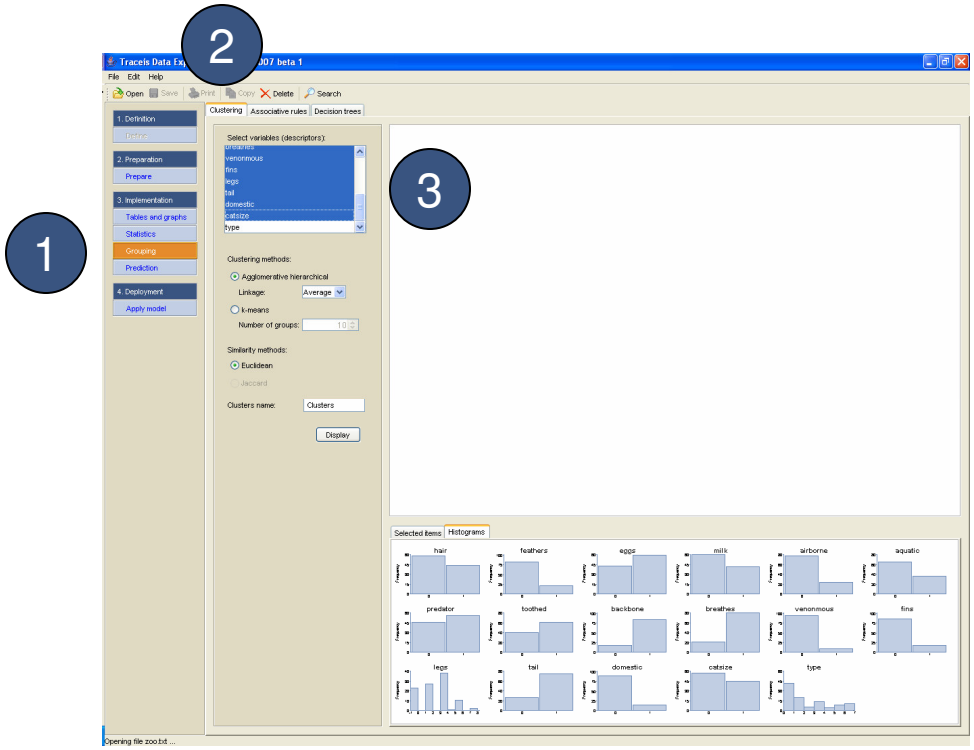
1 Select the grouping step

In this example, the zoo dataset was used. First, click on the grouping step.

2 Select the clustering tab

3 Select the descriptors

Select all variables to use as descriptors. In this example, all variables except type were selected. To select multiple non-contiguous variables, use ctrl-click and for contiguous variables use the shift-click.



1 Select clustering type and options

In this example, the agglomerative hierarchical clustering method was selected with the average linkage joining option.

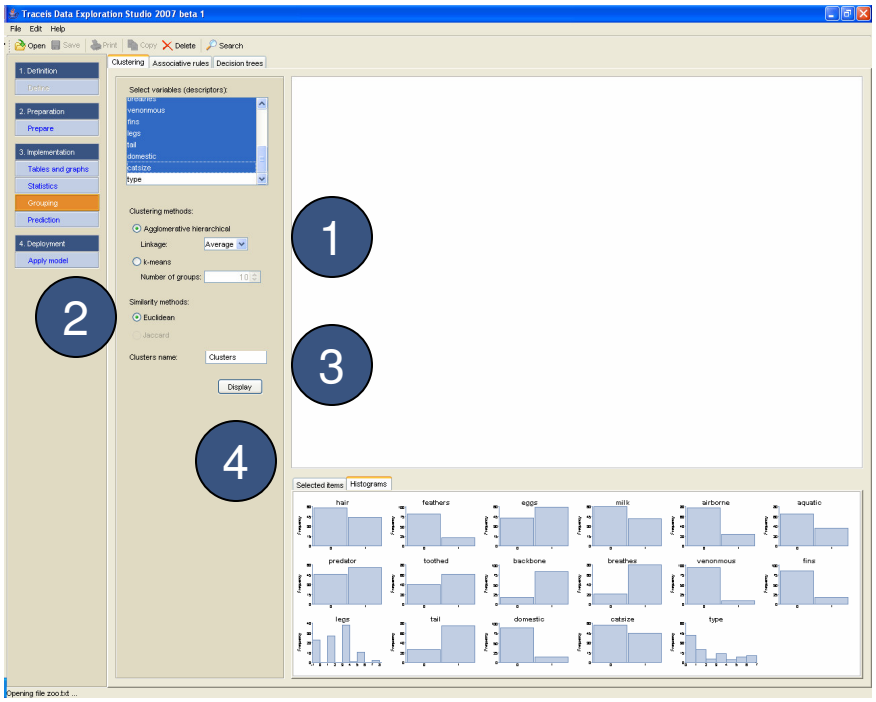
2 Select similarity method

The euclidean distance method is currently available.

3 Name the cluster variable

A variable will be generated for the selected cluster, where each observation is assigned to a cluster. The name of the cluster variable can be set here.

4 Display the clusters

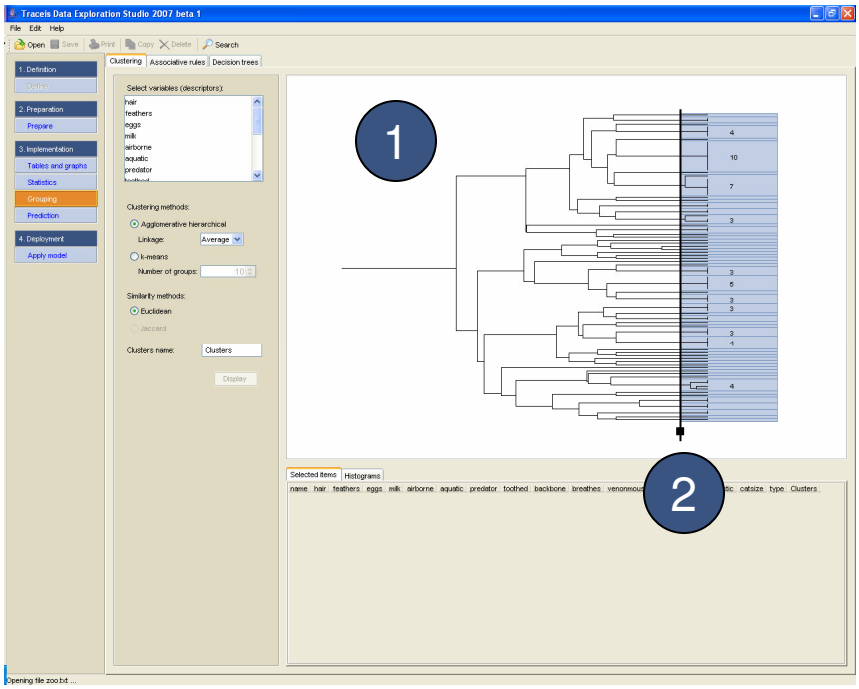


1 View the clusters

The results are presented in the form of a clustering dendrogram. The individual cluster are presented as rectangles to the right of the dendrogram.

2 Adjust the cut-off value

To change the distance at which the clusters are generated, click on the black square and move it to the left or right.

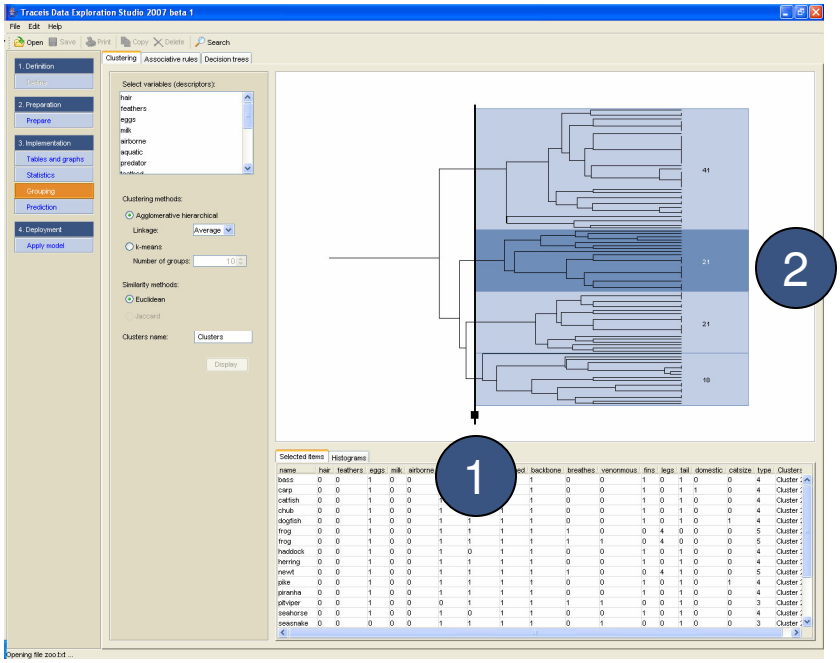


1 Adjust the cut-off value

Moving the cut-off value to the left will results in fewer clusters.

2 View cluster observations

To view the observations within a cluster, click within the blue cluster rectangle to the right of the cut-off line.

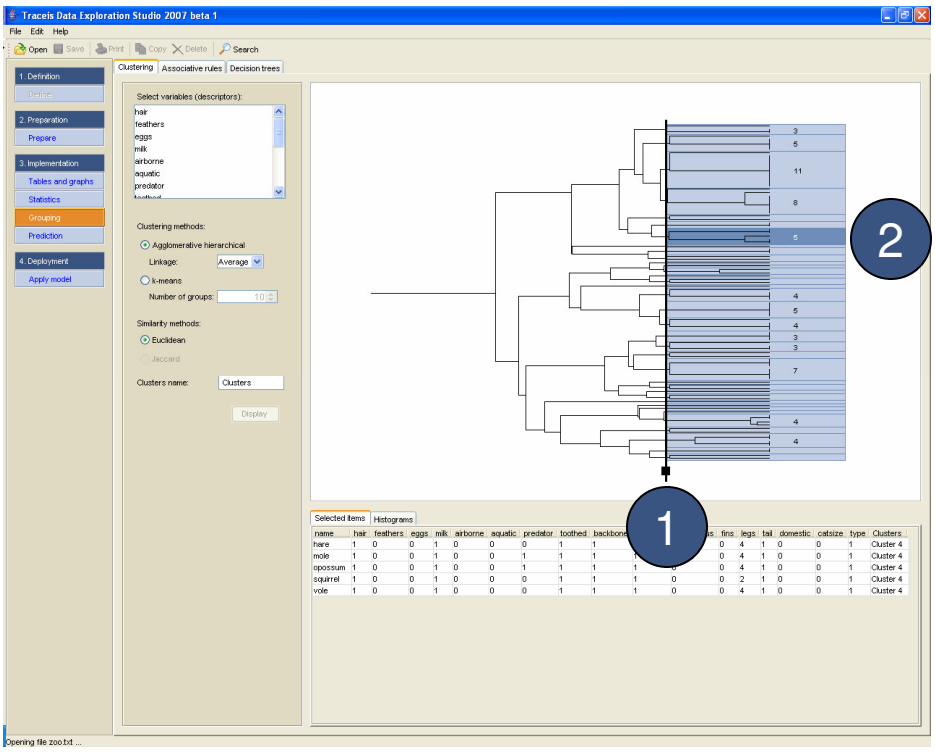


1 Adjust the cut-off value

Moving the cut-off value to the right will results in more clusters.

2 View cluster observations

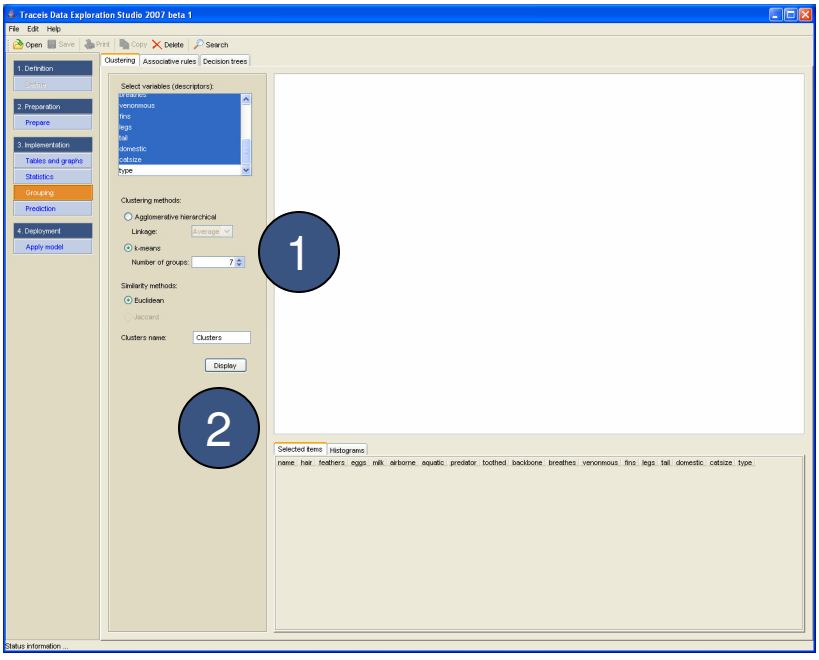
To view the observations within a cluster, click within the blue cluster rectangle to the right of the cut-off line.



1 Select clustering type and options

In this example, the k-means clustering method was selected to generate 7 clusters.

2 Display the clusters



1 View the clusters

The results are presented in the form of a series of rectangles. The vertical length of the rectangles is proportionate to the number of observations in each cluster.

2 View cluster observations

To view the observations within a cluster, click within the blue cluster rectangle.