Chapter 19 - Cluster analysis

Politics.csv contains the views of 150 football supporters on political issues, in 5 different age groups ranging from 61+ to 18-30. We are interested in differences between age groups. Produce a dendrogram and provide a (tentative) opinion on what it may represent. It is recommended that you reduce the sample size for clarity.

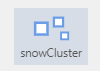


First, you load the snowCluster module from the Jamovi library (press the Modules tab at the top right, then press ‘jamovi library’, then install the snowCluster module from the list).

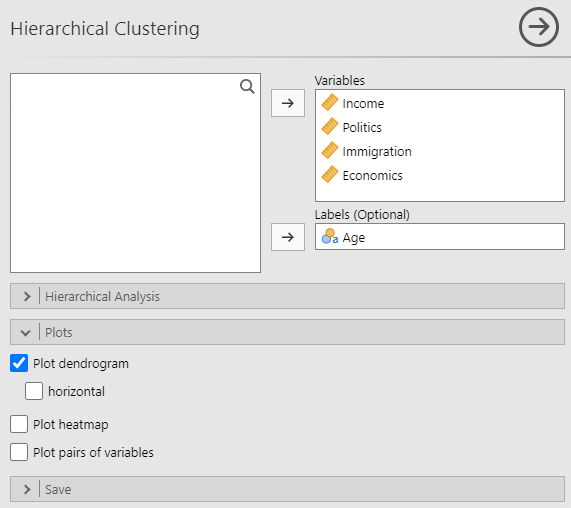
I created a sample using a random 6 rows from each of 5 age groupings:



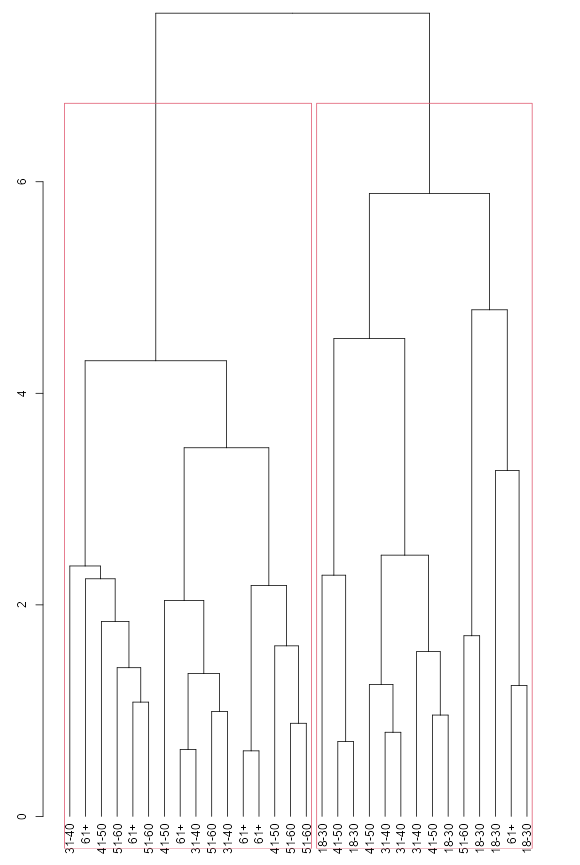
Press on the snowCluster tab:



Select **Hierarchical Clustering**.



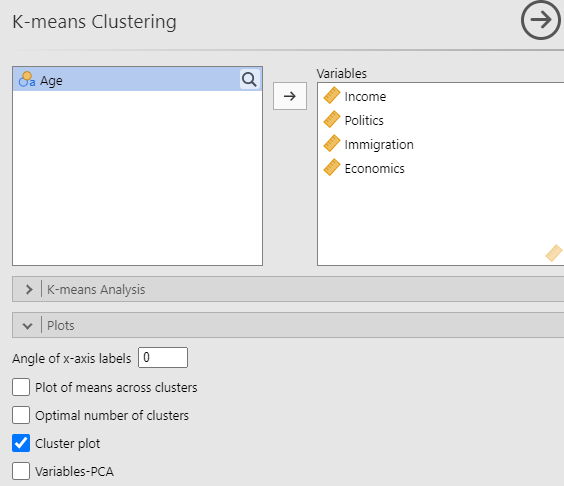
Here, I have left the Hierarchical Analysis panel alone, on default, and have selected Plot dendrogram from the Plots panel.



The red clustering boxes contain rather different age groupings. People in their 50s and 60s tend to occupy one uber-cluster, here shown on the left. Younger people tend to appear on the right. I have the impression that older respondents tend to have different points of view on the issues than younger ones. (Not that I think that Brexit was that clear-cut!)

Do note that your own subsamples won’t necessarily look like this, and you can of course change the size of the sub-sample, and maybe try out different cluster numbers (the default is two), and different distance and cluster options.

You may also want to try K-means Clustering.



I would suggest that the Cluster plot option would provide you with a way of identifying individual cases within clusters, and thus forming your own idea of what clusters mean.

