Subject: Suggest: select within cluster by property
Posted by DrCJM on Fri, 29 May 2020 02:06:26 GMT
View Forum Message <> Reply to Message

Something I'd find really useful, that I can't see how to do at present, is to select compounds after clustering by a particular property. The "representative" compound in a cluster is identified, I assume, as the most "central" compound in the cluster, which is fine. But sometimes I want to select the most active compound in the cluster, or the smallest MW in the cluster etc. etc.

If that's already possible, advice on how would be great!

Otherwise, a method of saying "Select from each cluster by value of property X" is something I'd use a lot.

Cheers,

Craig.

Subject: Re: Suggest: select within cluster by property
Posted by thomas on Sat, 30 May 2020 15:19:04 GMT
View Forum Message <> Reply to Message

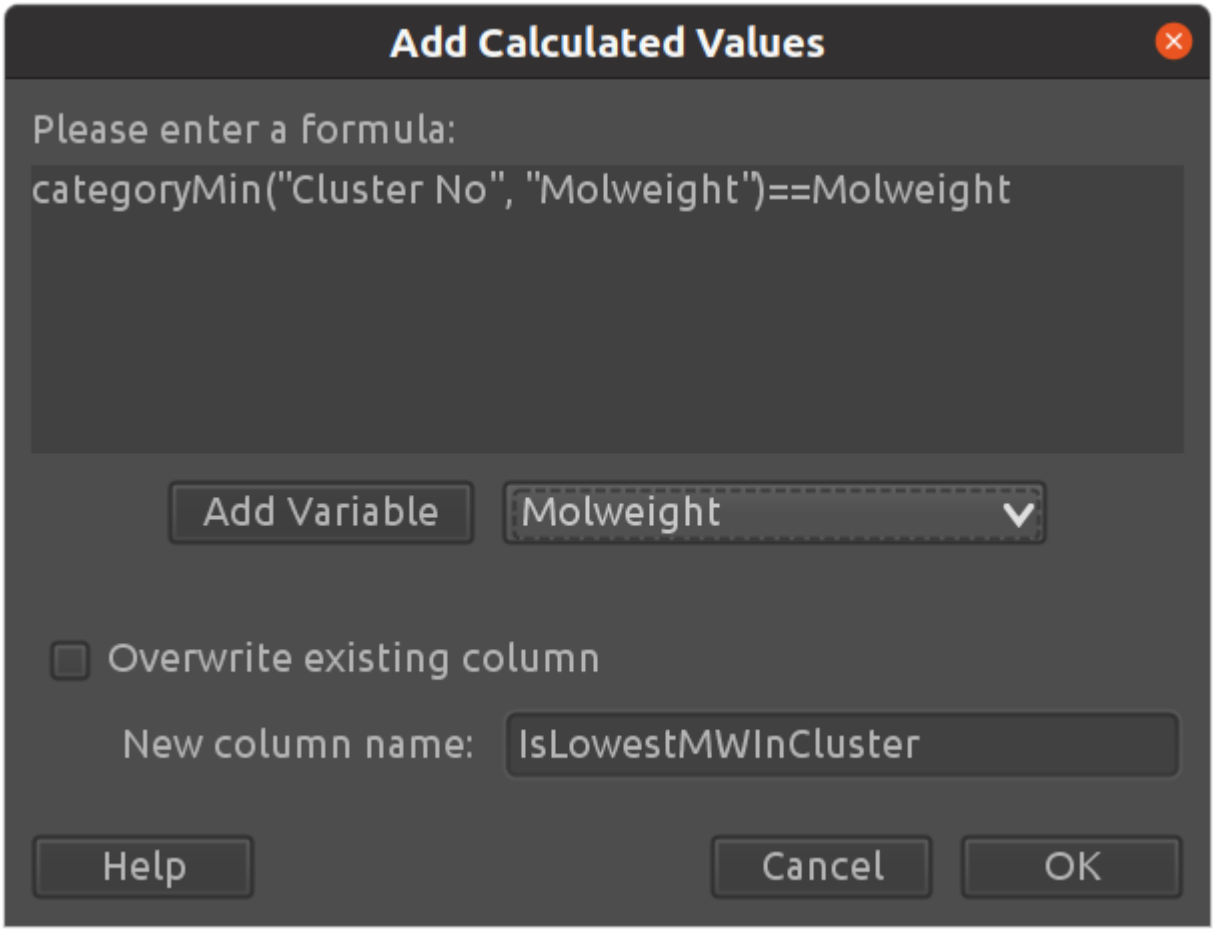
Assuming you have cluster numbers in a column called 'Cluster No' and molecular weights in a column called 'Molweight', then you could use "Data->Add Calculated Values...", configure the dialog as follows:

This creates a new column with a '1' value for all cluster members that have the lowest molecular weight within their own cluster. All other rows have a '0' in that column. Note that the Molweight column should not be set to use logarithmic behaviour. I need to fix an issue there.

Thomas

File Attachments

1) smallestInCategory.png, downloaded 1458 times



Subject: Re: Suggest: select within cluster by property
Posted by DrCJM on Mon, 01 Jun 2020 01:35:54 GMT
View Forum Message <> Reply to Message

Hi Thomas,

when I try that I get an error: Function "categoryMin" requires 3 parameters

I'll play around and see if I can figure out how to get it working, but any further advice would be great.

EDIT

OK, got it to work, it needed the category value defined, so format is:

categoryMin(ClusterNo,"Cluster No","Molweight")

which generates a new column with the lowest Molweight for the cluster added to every row (same value for all members of the cluster). With that, and my now finding the appropriate

functions in the Help page I think I can do what I need to do.

Cheers,

Craig.

Subject: Re: Suggest: select within cluster by property
Posted by thomas on Fri, 05 Jun 2020 14:34:39 GMT
View Forum Message <> Reply to Message

Hi Craig,

sorry this slipped my attention. The category functions were simplified during recent months to only use two parameters. The official version 5.2.1 still uses the 3 parameter form. Under this link you always find about weekly releases of replacement files of the most recent development cycle: openmolecules/datawarrior/dw521win.zip (Windows) and dw521x.zip (Mac/Linux). The have many fixes, new features, but often not up-to-date manual pages. And new features are sometimes in experimental state.

Best wishes, Thomas
