Subject: How to customize Merge Equivalent Rows
Posted by yuan on Thu, 23 May 2024 04:44:03 GMT
View Forum Message <> Reply to Message

Hi Thomas,

Datawarrior is a very useful open-source cheminformatics tools. Thanks for sharing it.

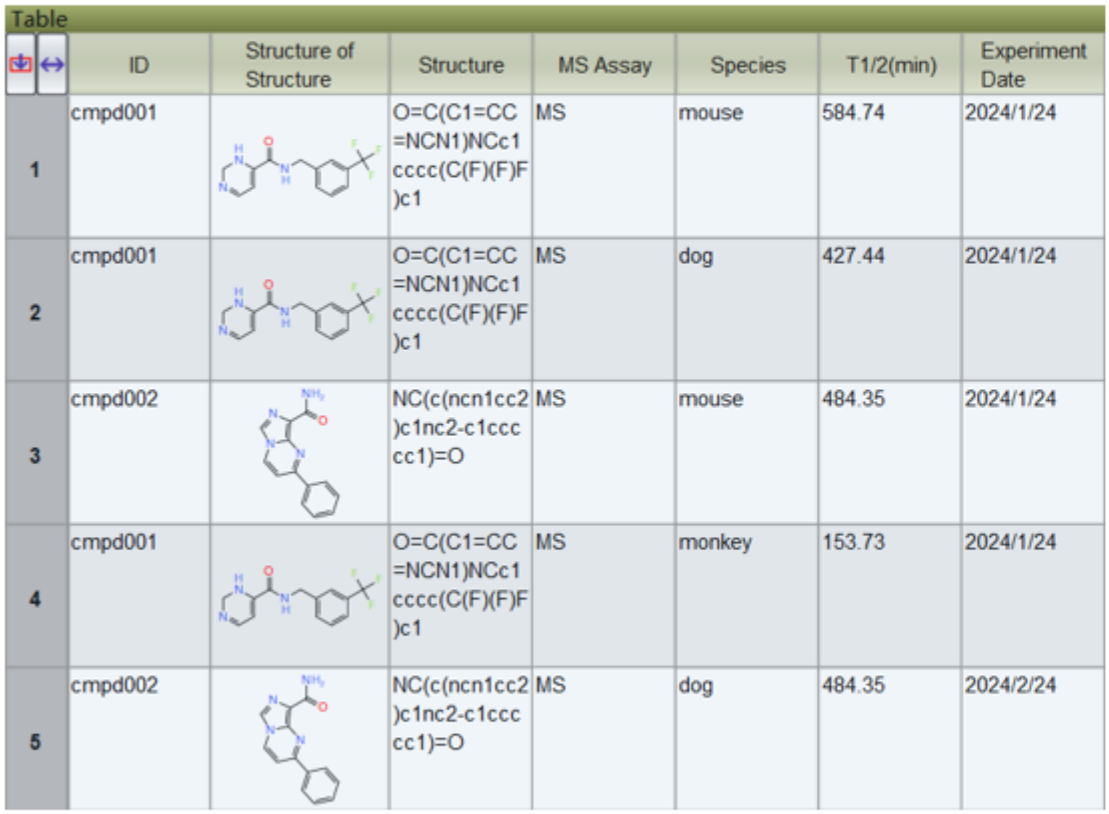
We wanted to use it organize experiment data by merging different assay results. When we used merge equivalent rows, we found that except for selected column(s) for equivalent check, other columns with the same values were also merged as a single value. Can we keep them individual or modify this function to let user determine whether merging equivalent values?

Many thanks!

Yuan

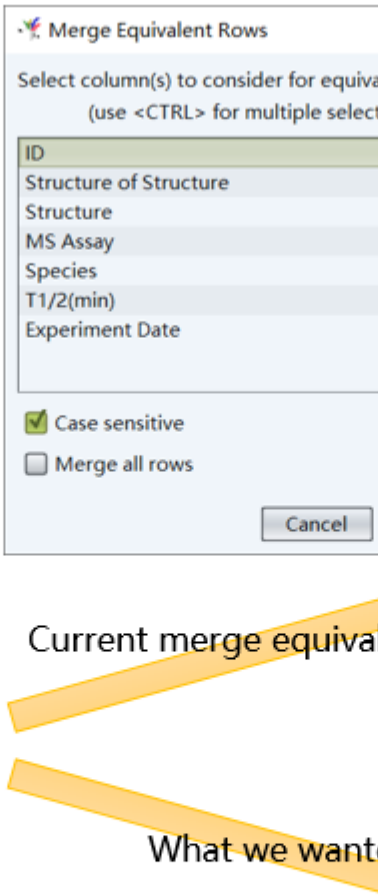
File Attachments

1) merge equivalent rows 1.png, downloaded 895 times



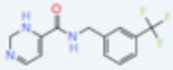
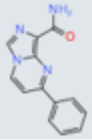
Table

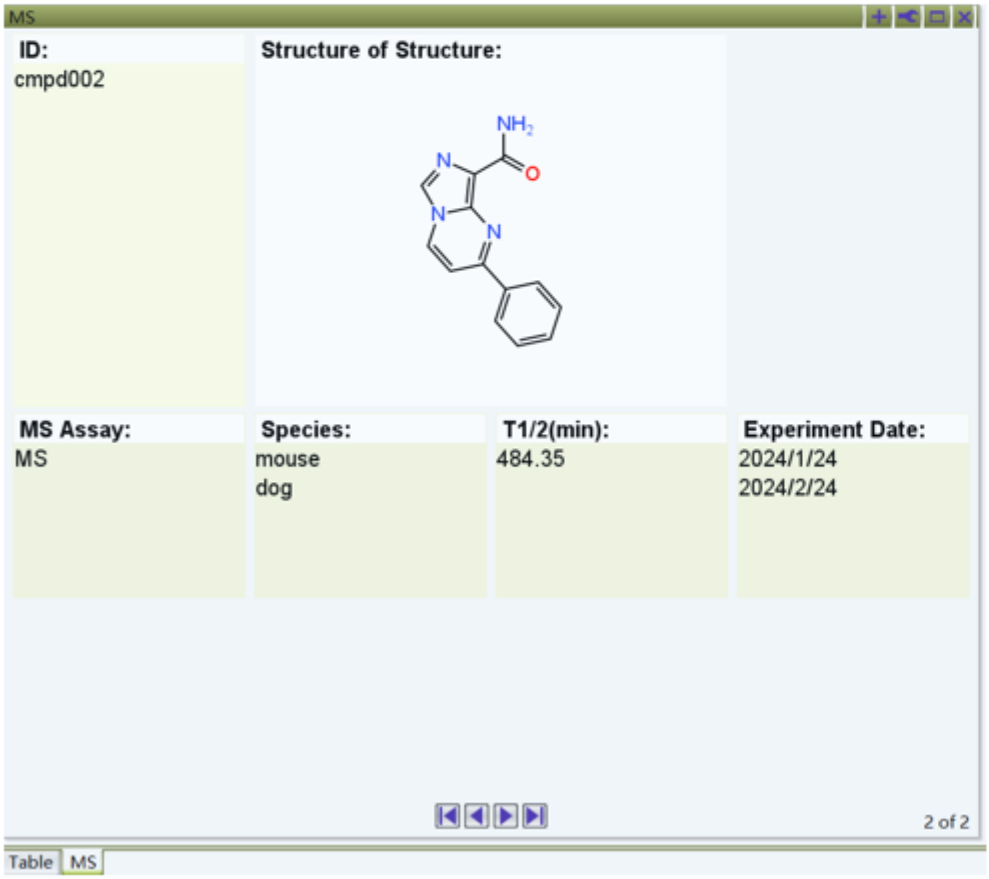
	ID	Structure of Structure	Structure	MS Assay	Species	T1/2(min)	Experiment Date
1	cmpd001		O=C(C1=CC=NCN1)NCc1cccc(C(F)(F)F)c1	MS	mouse	584.74	2024/1/24
2	cmpd001		O=C(C1=CC=NCN1)NCc1cccc(C(F)(F)F)c1	MS	dog	427.44	2024/1/24
3	cmpd002		NC(c(ncn1cc2)c1nc2-c1ccccc1)=O	MS	mouse	484.35	2024/1/24
4	cmpd001		O=C(C1=CC=NCN1)NCc1cccc(C(F)(F)F)c1	MS	monkey	153.73	2024/1/24
5	cmpd002		NC(c(ncn1cc2)c1nc2-c1ccccc1)=O	MS	dog	484.35	2024/2/24



2) merge equivalent rows 2.png, downloaded 800 times

Table

	ID	Structure of Structure	Structure	MS Assay	Species	T1/2(min)	Experiment Date
1	cmpd001		O=C(C1=CC=NCN1)NCc1cccc(C(F)(F)F)c1	MS	mouse dog monkey	584.74 427.44 153.73	2024/1/24
2	cmpd002		NC(c(ncn1cc2)c1nc2-c1ccccc1)=O	MS	mouse dog	484.35	2024/1/24 2024/2/24



What we wanted

3) test data.xlsx, downloaded 548 times
4) Test-keep individual.dwar, downloaded 595 times

Subject: Re: How to customize Merge Equivalent Rows
Posted by thomas on Thu, 23 May 2024 20:12:50 GMT
View Forum Message <> Reply to Message

Hi Yuan,

thanks you for the hint. Of course, your suggestion makes a lot of sense. I have update the source code now to offer three options:

- 'remove duplicates', which is similar to the current handling: if the cells to be merged all contain the same value,
  then this value is placed into the merged cell only once. Otherwise we see all entries in original order
- 'keep duplicates', which is your suggested handling: we always see all entries in original order
- 'depend on columns groups', which assumes that related columns, e.g. result columns of one assay, are grouped using
  the column grouping feature. Then it checks, whether all cells of the same column group are equal and only then it
  removes for all these columns the same entry.

In the next version this will go live...

Thomas

Subject: Re: How to customize Merge Equivalent Rows
Posted by yuan on Fri, 24 May 2024 05:11:29 GMT
View Forum Message <> Reply to Message

Hi Thomas,

Thanks for your quickly reply! I'm glad to hear that my suggestion was helpful. The updates you've made to the source code sound fantastic, and I'm sure they'll improve the user experience. Offering three different options for handling duplicates will give users more control and flexibility when merging cells.

Great job on implementing these changes!

By the way, can we take a look at the updated version?

Yuan

Subject: Re: How to customize Merge Equivalent Rows
Posted by thomas on Fri, 24 May 2024 13:50:10 GMT
View Forum Message <> Reply to Message

I prepared a test file for you. You can download it from https://openmolecules.org/datawarrior/datawarrior_v06.02.02a lpha.jar Then copy it into your datawarrior update folder. On Windows this is C:\\ProgramData\DataWarrior\

The standard procedure to test any non-yet-released functionality would be:
On a Linux or MacOS computer with JDK21/JFX (I suggest Liberica full JDK 21) installed clone the DataWarrior project from Github and build the jar file:
> git clone https://github.com/thsa/datawarrior
> buildDataWarrior

Then rename the built 'datawarrior_all.jar' file to any version newer than the current one, e.g. 'datawarrior_v06.02.02.jar' and put it into your datawarrior update folder, where the other automatic updates are located. This works on any platform (Linux, Mac, or Windows). If you then launch DataWarrior, it will use this version. Don't forget to remove the file after testing it.

Subject: Re: How to customize Merge Equivalent Rows
Posted by yuan on Tue, 28 May 2024 01:40:52 GMT
View Forum Message <> Reply to Message

Hi Thomas,

Thank you for your reply. I downloaded the updated .jar file and tested it. Really fantastic!
I believe more users will benefit from it.

Yuan

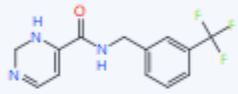
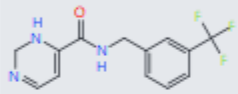
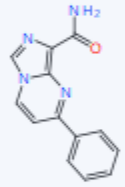
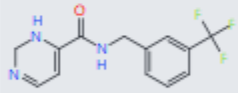
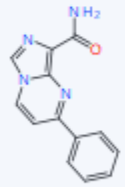
File Attachments

1) Merge_results.png, downloaded 808 times

Test.dwar

File Edit Data Chemistry Database List Macro Help

Table

	ID	Structure of Structure	Structure	MS Assay	Species	T1/2(min)	Experimen Date
1	cmpd001		O=C(C1=CC=	MS	mouse	584.74	2024/1/24
2	cmpd001		O=C(C1=CC=	MS	dog	427.44	2024/1/24
3	cmpd002		NC(c(ncn1cc2	MS	mouse	484.35	2024/1/24
4	cmpd001		O=C(C1=CC=	MS	monkey	153.73	2024/1/24
5	cmpd002		NC(c(ncn1cc2	MS	dog	484.35	2024/2/24

Subject: Re: How to customize Merge Equivalent Rows
Posted by thomas on Tue, 11 Jun 2024 12:25:56 GMT
View Forum Message <> Reply to Message

Hi Yuan, your example image is a perfect illustration for this feature. May I use it for the manual?