

---

Subject: Duplicate Molecules

Posted by [sbembenek18](#) on Sun, 01 Mar 2020 23:30:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Before loading an SDF file into DW, I remove duplicate molecules via canonical smiles. However, in DW, when I do: Data->Merge Equivalent Rows->Structure, DW still says there's duplicates -- I am unclear as to how this is possible.

---

---

Subject: Re: Duplicate Molecules

Posted by [thomas](#) on Mon, 02 Mar 2020 17:46:47 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

This would mean that two different canonical SMILES are considered the same structure by DataWarrior. I cannot perceive of a reason for that. If you delete all unique rows from your file, then you retain all these cases. If you can send me some of them, I may have a more detailed explanation.

---

---

Subject: Re: Duplicate Molecules

Posted by [sbembenek18](#) on Wed, 04 Mar 2020 20:18:28 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

I see what it is now. I have molecules where nitrogen might be considered a stereocenter. So, whereas the program that wrote the SDF considers that there are distinct stereoisomers, DW does not. I think this is consistent with the challenges of assessing this for these molecules.

---