

Subject: .csv data input including SMILES  
Posted by [nbehrnd](#) on Thu, 13 Jun 2019 13:40:41 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Dear Thomas,

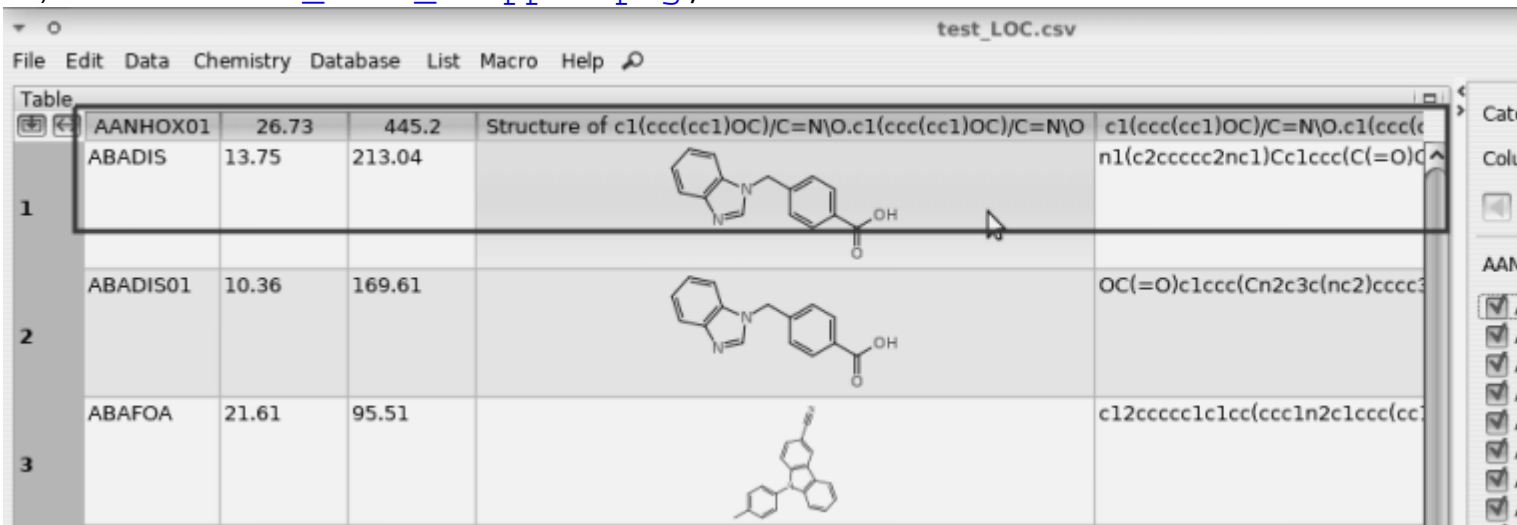
while importing comma-separated tables including a column containing SMILES, I noticed DataWarrior (version 5.0.0; running in Linux) appears to work well starting with the second line of the table only. It is not an issue related to file endings which could be resolved by the user by unix2dos. Apparently, DataWarrior expects the first line to be a header; which currently is not described-as-such in the online manual pages "Loading Data into DataWarrior".

Example of hampered reading of a .csv file lacking an explicit header:

Example of reading a .csv file containing explicit header:

## File Attachments

- 1) [test\\_data\\_with\\_header.csv](#), downloaded 803 times
- 2) [test\\_LOC.csv](#), downloaded 774 times
- 3) [headerless\\_data\\_cropped.png](#), downloaded 1479 times



The screenshot shows the DataWarrior interface with a table titled "test\_LOC.csv". The table has four columns: ID, two numerical columns, and a SMILES column. The first row is highlighted and contains the following data:

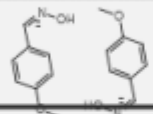
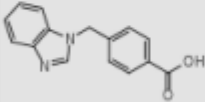
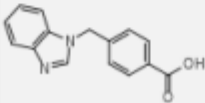
ID	Col 2	Col 3	SMILES
1	AANH0X01	26.73	445.2
2	ABADIS	13.75	213.04
3	ABADIS01	10.36	169.61
4	ABAFOA	21.61	95.51

The SMILES column contains chemical structures. The first structure is a benzimidazole derivative with a carboxylic acid group. The second structure is a similar benzimidazole derivative. The third structure is a more complex polycyclic aromatic hydrocarbon.

- 4) [data\\_with\\_header\\_cropped.png](#), downloaded 1366 times

Marcus\_parameters\_and\_smiles.csv

File Edit Data Chemistry Database List Macro Help

Table	CCDC code	Hab	lambda	Structure of Smiles	Smiles	CODE check
1	AANH0X01	26.73	445.20		<chem>c1(ccc(cc1)OC)OC</chem>	AANH0X01
2	ABADIS	13.75	213.04		<chem>n1(c2ccccc2n1)Cc3ccc(C(=O)O)cc3</chem>	ABADIS
3	ABADIS01	10.36	169.61		<chem>OC(=O)c1ccc2n(c1)c3ccccc3n2</chem>	ABADIS01

---

Subject: Re: .csv data input including SMILES

Posted by [thomas](#) on Thu, 27 Jun 2019 06:12:55 GMT

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

---

Dear Norwid, thank you for the hint. DataWarrior does a minimum data analysis trying to figure out whether there is a header line or not. I will improve this check and also update the manual to explain.

---