
Subject: Re: data with number of columns variables

Posted by on Thu, 05 Mar 2015 18:55:08 GMT

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

Den torsdag 5 mars 2015 kl. 19:12:00 UTC+1 skrev luc...@gmail.com:

> Hello
> I have a set of data like this
> The rows can have 3, 7 8, or 9 columns. and they don't repeat in any particular order.
>
> I want to obtain different arrays depending on the first element of the row
> any ideas?
>
> header
> header
> 1500.0 1.0 2.0
> 3000.0 1.0 40.0 3.0 212.0 0.0 0.0
> 10.3 7.9 -2.4 -0.6 -0.8 0.1 10.0 1.0 0.0
> 3000.0 2.0 48.0 139.0 3.0 122.0 6.0 0.0
> 5.7 2.5 -1.6 -0.6 -0.8 0.1 10.0 1.0 0.0
> 5000.0 6.0 78000.0 1.0 3.0 115.0 5.0 1.0
> 3.9 0.3 -0.5 -0.5 -0.2 0.8 1.2 1.0 0.0
> 2011.0 6.0 2.0 21.0 3.0 115.0 5.0 1034.0
> 3.9 0.3 -0.5 -0.8 -0.2 -0.5 0.0 1.0 0.0
> 5000.0 17.0 78000.0 3.0 3.0 123.0 5.0 4.0
> 5.0 0.5 1.3 -1.0 -0.1 0.3 0.3 1.0 0.0
> 2011.0 17.0 2.0 23.0 3.0 123.0 5.0 1005.0
> 5.0 0.5 1.3 -1.0 0.2 -0.1 0.0 1.0 0.0
> 5000.0 19.0 78000.0 3.0 3.0 123.0 5.0 2.0
> 5.0 0.5 1.3 0.3 0.0 -1.0 0.0 1.0 0.0
> 2014.0 19.0 2.0 25.0 3.0 123.0 5.0 485.0
> 5.0 0.5 1.3 0.5 -0.6 0.6 0.0 1.0 0.0
> 1635.0 1.0 3.0
> 3000.0 1.0 40.0 3.0 212.0 0.0 0.0
> 7.3 9.8 2.5 -0.2 -0.9 -0.4 10.0 1.0 0.0
> 3000.0 2.0 48.0 155.0 3.0 122.0 6.0 0.0
> 5.6 2.5 -0.4 -0.2 -0.9 -0.4 10.0 1.0 0.0
> 3000.0 3.0 47.0 168.0 3.0 211.0 0.0 5.0
> 5.6 2.5 -0.4 0.1 -1.0 -0.2 6.9 1.0 0.0
> 1810.0 1000.0 4.0
> 3000.0 1.0 40.0 3.0 212.0 0.0 0.0
> -3.0 -4.1 -1.3 0.7 0.7 0.3 10.0 1.0 0.0
> 3000.0 2.0 1.0 47.0 3.0 201.0 0.0 0.0
> 0.0 -1.1 -0.2 0.7 0.7 0.3 10.0 1.0 0.0
> 5000.0 3.0 115.0 44.0 3.0 108.0 5.0 0.0
> 1.7 0.6 0.4 0.7 0.7 0.3 10.0 1.0 0.0
> 2012.0 3.0 2.0 1.0 3.0 108.0 5.0 1605.0
> 1.7 0.6 0.6 -0.7 -0.6 -0.4 0.0 1.0 0.0
> 5000.0 11.0 29000.0 1.0 3.0 111.0 7.0 1.0

```
> 3.5 -0.2 -0.5 0.6 -0.8 -0.1 0.3 1.0 0.0
> 2011.0 11.0 2.0 3.0 3.0 111.0 7.0 532.0
> 3.5 -0.2 -0.5 -0.8 0.6 0.1 0.0 1.0 0.0
> 5000.0 5.0 78000.0 3.0 3.0 108.0 5.0 1.0
> 1.7 0.7 0.6 -0.8 0.0 -0.6 0.1 1.0 0.0
> 2011.0 5.0 2.0 5.0 3.0 108.0 5.0 736.0
> 1.7 0.7 0.6 -1.0 -0.1 -0.1 0.0 1.0 0.0
> 5000.0 5.0 78000.0 3.0 3.0 108.0 5.0 1.0
> 1.7 0.7 0.6 0.3 -0.9 0.2 0.0 1.0 0.0
> 2011.0 5.0 2.0 7.0 3.0 108.0 5.0 144.0
> 1.7 0.7 0.6 0.0 -1.0 0.1 0.0 1.0 0.0
> 5000.0 5.0 78000.0 3.0 3.0 108.0 5.0 1.0
> 1.7 0.7 0.5 -0.1 0.6 0.8 0.2 1.0 0.0
> 2011.0 5.0 2.0 9.0 3.0 108.0 5.0 909.0
> 1.7 0.7 0.5 -0.1 -0.7 0.7 0.0 1.0 0.0
> 5000.0 5.0 78000.0 3.0 3.0 108.0 5.0 1.0
> 1.7 0.7 0.5 0.5 -0.6 0.6 0.0 1.0 0.0
> 2011.0 5.0 2.0 11.0 3.0 108.0 5.0 648.0
> 1.7 0.7 0.5 0.4 -0.7 0.6 0.0 1.0 0.0
```

* Make a template with `ascii_template` based on the rows with the largest number of columns, specify NaNs for missing data.

* Read the file with `read_ascii` using this template

* The three-column rows are the ones with NaN in the fourth column.

* The seven-column rows are the ones with NaN in the eighth column but not in the fourth.

* Similar tests for eight-column and nine-column rows.
