
Subject: Re: Box and Whisker Plotting

Posted by [David Fanning](#) on Tue, 21 Jan 2014 13:26:32 GMT

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

Madhavan Bomidi writes:

> I am having a dataset with 100 columns (correspond to variable values) and 86400 rows (correspond to seconds or time). In order to identify the outlier data, I am attempting to make a Box and Whisker plot using Coyote's graphic function - cgBoxPlot.pro to ascertain the outliers in my data at every time stamp. I know that cgBoxPlot requires data in columns and so I will apply transpose before I adopt the graphics function. My queries are as below:

>

> 1. Are the number of columns of data restricted to only 28 columns if I want to use cgBoxPlot function?

IDL has a limit on the number of labels you can use, which if I remember correctly is just under 50. Since each "box" uses a label, this imposes a limit on the number of boxes or columns you can specify.

> 2. I also have lot of !Values.F_NaN values, which I will define with MISSING_DATA_VALUE=!Values.F_NaN. If all the variable values (or column data) are !Values.F_NaN, then will this cgBoxPlot ignores this and then will the function move to the next column of data automatically?

I don't know the answer to this. I can't imagine the author of the program was thinking about someone doing a box plot that consisted entirely of missing data. The code is available, of course, so you can modify it however you like.

> 3. Supposing that I don't want to reduce my time stamps, if there is a limitation for cgBoxPlot then what is the best possible solution to my problem?

I don't know the best solution. You seem to be exploring all of the possibilities. It may be you have to write your own display program if nothing "off the shelf" will work.

Cheers,

David

--

David Fanning, Ph.D.

Fanning Software Consulting, Inc.

Coyote's Guide to IDL Programming: <http://www.idlcoyote.com/>

Sepore ma de ni thue. ("Perhaps thou speakest truth.")
