Subject: Re: Box-Whisker plots in IDL
Posted by jschwab@gmail.com on Tue, 21 Aug 2007 04:46:51 GMT
View Forum Message <> Reply to Message

On Aug 21, 12:01 am, David Fanning <n...@dfanning.com> wrote:
> jsch...@gmail.com writes:
>> This looks good to me. I wrote a routine to find the energy quartiles
>> of some x-ray data a few weeks back and that was the way I ended up
>> doing it. Not that speed is an issue, but I'd be curious to see how
>> this method compares with a SORT, or some other (yet undiscussed)
>> method. Maybe I'll play around with that tonight if I have some extra
>> time.
>
> I wrote a short article to illustrate how I would go
> about creating a box and whisker plot in IDL:
>
> http://www.dfanning.com/graphics_tips/box&whisker.html
>
> Cheers,
>
> David
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming:http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Well I see David beat me to it, but I was also playing around writing
some box-and-whisker plotting code. I think David's is nicer than mine
(big surprise), but I'll post what I wrote so anyone else can play
with it if they like.

Cheers,
Josiah

--
```
;+
; NAME: BWPLOT
;
; PURPOSE: Draw a box-and-whisker plot
;
; CATEGORY: Plotting, Graphics
;
; CALLING SEQUENCE: BWPLOT, data
;
; INPUTS: data = data to be plotted
;
; OPTIONAL INPUTS: None
```

```
;
; KEYWORD PARAMETERS:
;
;    OUTLIERS - if set, plots outliers (points which are > 1.5x
;             the interquartile range away from Q25 or Q75
;
;    BOXWIDTH - height of the box as a percentage of the screen
;             height; defaults to 10%
;
;    WHISKWIDTH - height of the whiskers as a percentage of the
;              screen height; defaults to half of the box width
;
;    BOXCOLOR - color to make the box portion of the plot; specify in
;             the same manner that one would set COLOR when using
;             PLOT
;
;    WHISKCOLOR - color to make the whisker portion of the plot;
;              specify in the same manner that one would set
;              COLOR when using PLOT
;
;    OUTSYM - plot symbol to use for outliers; only relavent when the
;            OUTLIERS keyword is set; specify in the same manner
;            that one would set PSYM when using PLOT
;
;    OUTCOLOR - color to make the outliers in the plot; only relavent
;             when the OUTLIERS keyword is set; specify in the
;             same manner that one would set COLOR when using PLOT
;
;    QUARTILES - variable to contain the 5 values used to contruct
;              the plot; a 5 element array
;              [min, q25, median, q75, max]
;
;    IQR - variable to contain the value of the interquartile range
;
; OUTPUTS: None (see keywords QUARTILES and IQR)
;
; OPTIONAL OUTPUTS: None
;
; COMMON BLOCKS: None
;
; SIDE EFFECTS: None
;
; RESTRICTIONS:
;      Does not produce vertical plots
;      Does not produce multiple plots
;      Does not explictly label quartiles
;
; PROCEDURE: Straightforward
```

```
;
; EXAMPLE:
;    Make a box-and-whisker plot of some random data
;
;    random_data = randomu(seed, 1000) * 100.
;    bwplot, random_data
;
; MODIFICATION HISTORY:
;
;   Created:
;       Mon Aug 20, Josiah Schwab
;
; LICENSE
;    Copyright (c) 2007 Josiah Schwab
;
;Permission is hereby granted, free of charge, to any person obtaining
a copy
;of this software and associated documentation files (the "Software"),
to deal
;in the Software without restriction, including without limitation the
rights
;to use, copy, modify, merge, publish, distribute, sublicense, and/or
sell
;copies of the Software, and to permit persons to whom the Software is
;furnished to do so, subject to the following conditions:
;
;The above copyright notice and this permission notice shall be
included in
;all copies or substantial portions of the Software.
;
;THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR
;IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF
MERCHANTABILITY,
;FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT
SHALL THE
;AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR
OTHER
;LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE,
ARISING FROM,
;OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER
DEALINGS IN
;THE SOFTWARE.
;
;-

PRO bwplot, data, $
          OUTLIERS = OUTLIERS, $
```

```
        BOXWIDTH = BOXWIDTH, WHISKWIDTH = WHISKWIDTH, $
        BOXCOLOR = BOXCOLOR, WHISKCOLOR = WHISKCOLOR, $
        QUARTILES = QUARTILES, IQR = IQR, $
        OUTSYM = OUTSYM, OUTCOLOR = OUTCOLOR

COMPILE_OPT IDL2
ON_ERROR, 2

;test for at least 5 pts
if n_elements(data) lt 5 then message, "Must have at least 5 points"

;; set keywords
if not keyword_set(outliers) then outliers = 0
if not keyword_set(boxwidth) then boxwidth = 0.1
if not keyword_set(whiskwidth) then whiskwidth = boxwidth / 2.
if not keyword_set(boxcolor) then boxcolor = !P.color
if not keyword_set(whiskcolor) then whiskcolor = !P.color
if not keyword_set(outsym) then outsym = 1
if not keyword_set(outcolor) then outcolor = !P.color

;; calculate quartiles
;; they are returned in variable "Quartiles"
minVal = min(data, max = maxVal)
medVal = median(data,/EVEN)
q25Val = median(data[where(data LE medVal)], /even)
q75Val = median(data[where(data GT medVal)], /even)

quartiles = [minVal, q25Val, medVal, q75Val, maxVal]

;; calculate interquartile range
IQR = q75Val - q25Val

;; set up plot
left = floor(minVal - 0.1 * IQR)
right = ceil(maxVal + 0.1 * IQR)
plot, data, data, /nodata, $
     xrange = [left, right], yrange = [-1, 1], $
     xstyle = 1, ystyle = 1, $
     yticks = 1, ytickname = [' ', ' ']

;; OUTLIERS == 1 --> PLOT OUTLIERS SEPERATELY

if outliers eq 1 then begin

   low = q25Val - 1.5 * IQR
   high = q75Val + 1.5 * IQR

   out = where( (data LT low) OR (data GT high) , out_count, $
```

```
                complement = not_out )
  if out_count gt 0 then $
    oplot, data[out], rebin([0], out_count), $
          psym = outsym, color = outcolor
  whisk_min = min(data[not_out], max = whisk_max)

endif else begin

  whisk_min = minVal
  whisk_max = maxVal

endelse

;draw box
plots, [q25Val, q75Val, q75Val, q25Val, q25val], $
      boxwidth * [1, 1, -1, -1, 1], $
      color = boxcolor
plots, [medVal, medVal], boxwidth * [1, -1], color = boxcolor

;draw whiskers
plots, [q75val, whisk_max], [0, 0], color = whiskcolor
plots, [q25val, whisk_min], [0, 0], color = whiskcolor
plots, [whisk_max, whisk_max], whiskwidth * [-1, 1], color =
whiskcolor
plots, [whisk_min, whisk_min], whiskwidth * [-1, 1], color =
whiskcolor


end
```