Adding Statistics and Text to the Panel Graphs using INSET option in PROC SGPANEL

Ajay Gupta, PPD, Morrisville, NC

ABSTRACT

The SGPANEL procedure creates a panel of graph cells for the values of one or more classification variables. Business users often request to add a text or statistics to each panel cell of the graph. Normally, this task can be accomplished using Graph Template Language (GTL). But, the GTL approach requires a good amount of programming time and can be tedious. In SAS v9.4, the INSET statement is added in PROC SGPANEL. Using the INSET statement, text and statistics can be added in the panel graph very easily. This paper will introduce the INSET statement in the SGPANEL procedure and provide a brief overview of all options (with examples) which can be used in the INSET statement to enhance the appearance of text and statistics in the panel graph.

INTRODUCTION

The SGPANEL procedure creates a panel of graph cells for the values of one or more classification variables. For example, if a data set contains three variables (A, B and C) and you want to compare the scatter plots of B*C for each value of A, then you can use the SGPANEL procedure to create this panel. The SGPANEL procedure creates a layout for you automatically and splits the panel into multiple graphs if necessary. The SGPANEL procedure can create a wide variety of plot types, and overlay multiple plots together in each graph cell in the panel. It can also produce several types of layouts. For e.g. box plot, histogram, bar chart, and scatter plot.

In SAS v9.4, the INSET statement is added in PROC SGPANEL. Using the INSET statement, descriptive text and statistics can be added in any panel graph very easily. This paper will introduce the INSET statement in the SGPANEL procedure and provide a brief overview of all options (with examples) which can be used in the INSET statement to enhance the appearance of text and statistics in the panel graph.

On another note, only one INSET statement can be specified in the PROC SGPANEL step. If more than one is specified, the first inset is generated and a message is logged for the additional statements. Also, unlike the SGPLOT procedure, panel insets do not accept text strings as arguments. The INSET statement in the SGPANEL procedure generates data-driven text from one or more variables. Typically, the variable is a computed numeric value, such as a mean or a sum. For non-computed variables, the statement displays the value of the first observation for each classification. The inset labels are derived from the variable labels, or variable names if the labels are not present. (You can suppress the labels using a NOLABEL option.) The inset values come from the variable data.

SYNTAX FOR INSET STATEMENT IN PROC SGPANEL:

INSET variable <...variable-n> </option(s)>;

Please see table below for detail list of options.

Options	Detail Summary
BACKCOLOR=color	Specifies the background color of the inset. You can specify colors using a number of different color-naming schemes. For more information, see Color-Naming Schemes. By default, the background is transparent.
BORDER | NOBORDER	Specifies whether to display a border around the text box. The BORDER option displays the border. The NOBORDER option hides the border.
NOLABEL	Suppresses the label.
OPAQUE	Forces the inset background to be opaque rather than transparent. This option is useful when a transparent background makes the text in the inset difficult to see. For example, if the inset is positioned on top of a histogram, specifying OPAQUE sets the inset off from the histogram. The background is transparent unless BACKCOLOR= is also specified.
POSITION=position-value	Specifies the position of the text box within the plot. If you do not

	specify a position, then a position is determined automatically. The position values are as follows: BOTTOM, BOTTOMLEFT, BOTTOMRIGHT, TOPLEFT, TOPRIGHT, LEFT, and RIGHT.
SEPARATOR="string"	Specifies one or more characters to place between the data label and the data value. With fonts that support Unicode, you can produce specific characters by specifying a hexadecimal value. A trailing x identifies a string as a hexadecimal value. You must also enclose the character specification in a special ODS handler string, in the format (*ESC*){Unicode 'hexadecimal-value'x}. For example: separator="(*ESC*){unicode '03B1'x}"; This option produces the lowercase Greek letter alpha for the separator.
TEXTATTRS=style-element <(options)> | (options)	Specifies the appearance of the text in the text box. Example: TEXTATTRS=(Color=Green Family=Arial Size=8 Style=Italic Weight=Bold)
TITLE="text-string"	Specifies a title for the text box. The title text is always center-aligned.
TITLEATTRS=style-element <(options)> | (options)	Specifies the appearance of the title. Example: TITLEATTRS=(Color=Green Family=Arial Size=8 Style=Italic Weight=Bold)

Table 1. List of options use in INSET statement

TECHNIQUE & MECHANISM

The general process of adding the statistics to the panel graph is as follows:

1. Create all the required statistics using a SAS procedure. For e.g. PROC MEANS to get summary statistics.
2. Merge the statistics to the initial graph dataset.
3. Execute the PROC SGPANEL procedure with the INSET statement.
4. Use the INSET options to enhance the appearance and adjust the location of the statistical results.

For demonstration purpose, we will use SAS data set CLASS from library SASHELP. Please see Screenshot below for data set SASHELP.CLASS.

	Name	Sex	Age	Height	Weight
1	Alice	F	13	56.5	84
2	Barbara	F	13	65.3	98
3	Carol	F	14	62.8	102.5
4	Jane	F	12	59.8	84.5
5	Janet	F	15	62.5	112.5
6	Joyce	F	11	51.3	50.5
7	Judy	F	14	64.3	90
8	Louise	F	12	56.3	77
9	Mary	F	15	66.5	112
10	Alfred	M	14	69	112.5
11	Henry	M	14	63.5	102.5
12	James	M	12	57.3	83
13	Jeffrey	M	13	62.5	84
14	John	M	12	59	99.5
15	Philip	M	16	72	150
16	Robert	M	12	64.8	128
17	Ronald	M	15	67	133
18	Thomas	M	11	57.5	85
19	William	M	15	66.5	112

Output 1. Dataset SASSHELP.CLASS.

Example 1: In the example below (taken from SAS knowledge base) summary statistics are added to the panel histogram using the INSET statement. Please see SAS code comments for details.

```
/* Calculate the mean height and weight for the inset */

    PROC MEANS DATA=sashelp.class nway;
        CLASS sex;
        VAR weight height;
            OUTPUT OUT=stats mean(weight)=mean_weight
            mean(height)=mean_height n(height)=n_height
            n(weight)=n_weight std(height)=sd_height
            std(weight)=sd_weight;

    RUN;

/* Sort the data */
    PROC SORT DATA=sashelp.class OUT=class;
        BY sex;
    RUN;

/* Match-merge the inset data with the original data and label*/

    DATA merged;
        MERGE class stats;
        BY sex;
        LABEL mean_height = "mean(height)";
        LABEL mean_weight = "mean(weight)";
        LABEL n_height = "n(height)";
        LABEL n_weight = "n(weight)";
        LABEL sd_height = "sd(height)";
        LABEL sd_weight = "sd(weight)";
    RUN;
```
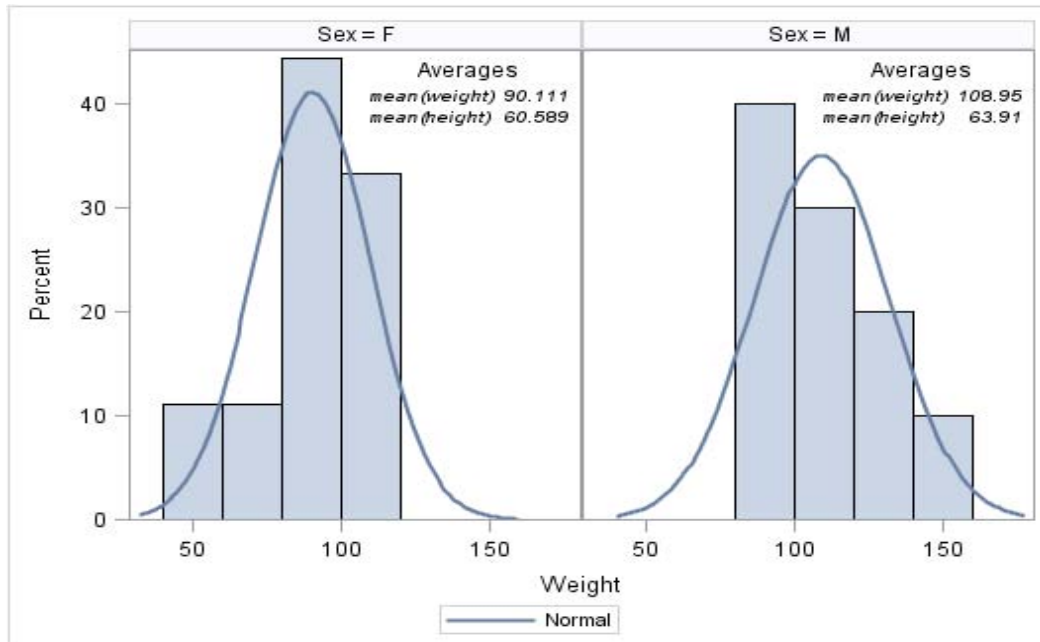
```
/*Create Panel Graph using INSET statement*/
    PROC SGPANEL DATA=merged;
        PANELBY sex;
        HISTOGRAM weight;
        DENSITY weight;
        INSET mean_weight mean_height /
                POSITION=topright TEXTATTRS=(style=italic)
                TITLE="Averages";
    RUN;
```

Averages are added in the Panel Histogram below:

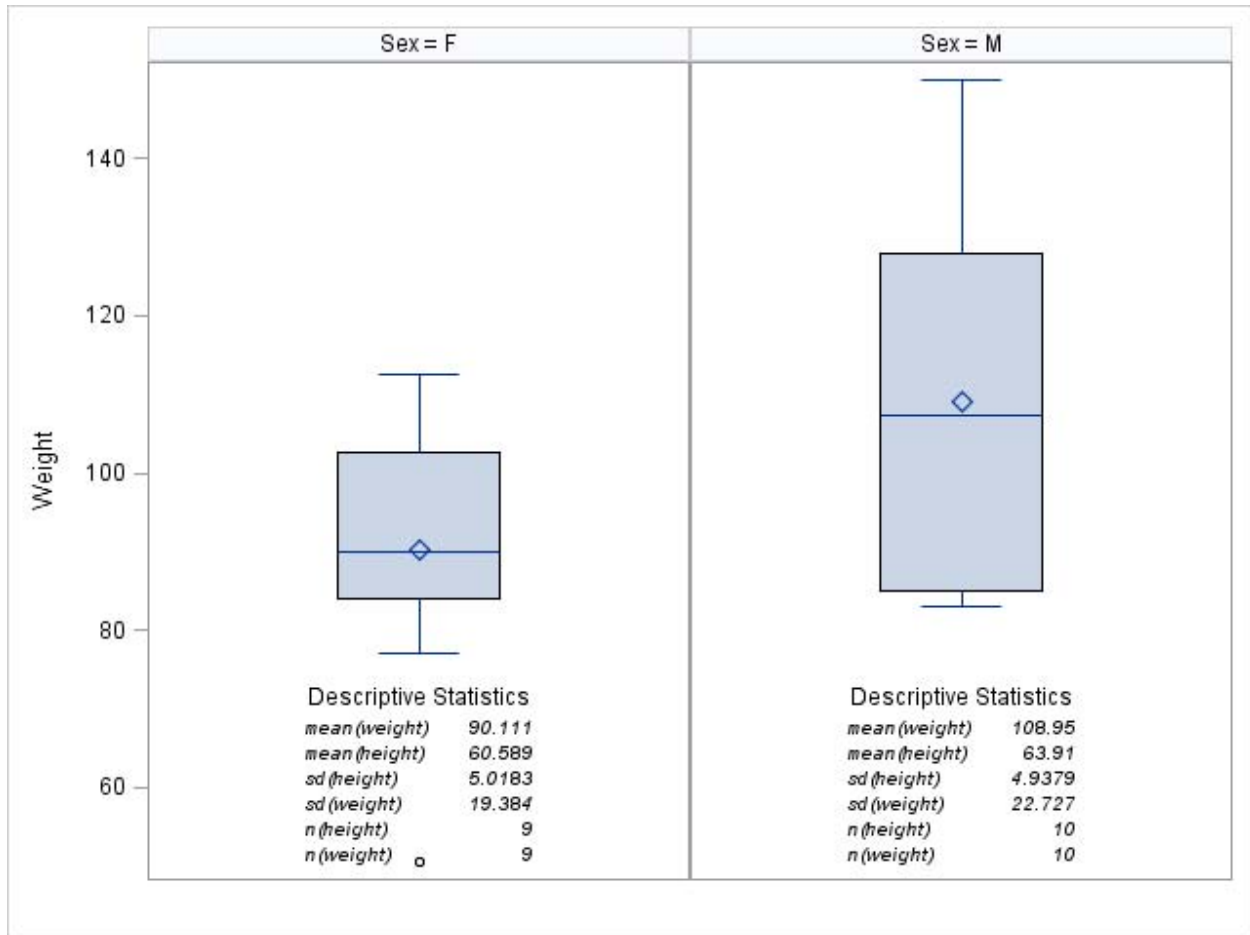**Figure 1: Panel Histogram with descriptive stats added using INSET statement**

Example 2: In the example below the summary statistics are added to the panel box plot using the INSET statement. Please see SAS code comments below for details.

```
/*Create Panel BOXPLOT using INSET statement*/

PROC SGPANEL DATA=merged;
      PANELBY sex;
      VBOX weight;
      INSET mean_weight mean_height sd_height sd_weight n_height n_weight /
      POSITION=bottom TEXTATTRS=(style=italic)
      TITLE="Descriptive Statistics";
RUN;
```

Averages are added in the Panel BOXPLOT below:

Figure 2: Panel BOXPLOT with Descriptive Stats added using INSET statement

Note, in above example only few INSET statement options are used for e.g. POSITION, TEXTATTRS, and TITLE. But, as needed other options can be easily added in above example.

CONCLUSION

Using the INSET statement, descriptive statistics can be added in the SGPANEL graph very easily. This option will bypass the complex GTL programming options to add descriptive statistics and will save good amount of programming time.

REFERENCES

http://support.sas.com

http://support.sas.com/documentation/cdl/en/grstatproc/67909/HTML/default/viewer.htm#n1xboke5qwjvh3n1w00ckrk elyxp.htm

http://support.sas.com/documentation/cdl/en/grstatproc/67909/HTML/default/viewer.htm#p0v3zd08d04up6n1uu8kzkz j4bkf.htm

http://support.sas.com/documentation/cdl/en/grstatproc/67909/HTML/default/viewer.htm#n0h71t919qh9g2n1rhvrpq3e 2cgu.htm

http://support.sas.com/documentation/cdl/en/grstatproc/67909/HTML/default/viewer.htm#n19gxtzyuf79t3n16g5v26b7 3ckv.htm#n1e9nkst4ozgbtn1myo0g4do40ny

Matange Sanjay, 2014. Annotate your SGPLOT Graphs. Proceedings of the PharmaSUG 2014 Conference, paper CC01.

ACKNOWLEDGMENTS

DISCLAIMER

The content of this paper are the works of the authors and do not necessarily represent the opinions, recommendations, or practices of PPD.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Ajay Gupta, M.S.
Enterprise: PPD
Address: 3900 Paramount Parkway
City, State ZIP: Morrisville, NC-27560
Work Phone: (919)-456-6461
Fax: (919)-654-9990
E-mail: Ajay.Gupta@ppdi.com,
      Ajaykailasgupta@aol.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.