Multi-Cell Plotting

SP4R04d04.sas

1. Create a data table with the variable **Sex** that has groups *F* and *M*. For each group, generate 1000 observations with a seed of 123. If **Sex** is *F,* let **Height** be normally distributed with a mean of 66 and a standard deviation of 2. If **Sex** is *M*, let **Height** be normally distributed with a mean of 72 and a standard deviation of 2.

```
data sp4r.multi;
   call streaminit(123);
   do Sex='F', 'M';
      do j=1 to 1000;
         if sex='F' then height = rand('Normal',66,2);
         else height = rand('Normal',72,2);
         output;
      end;
   end;
run;
```

The DO loop can also iterate over a list of character variables. For each iteration, the variable is filled with the specified name. The DO loop does not require the TO syntax when a list is provided.

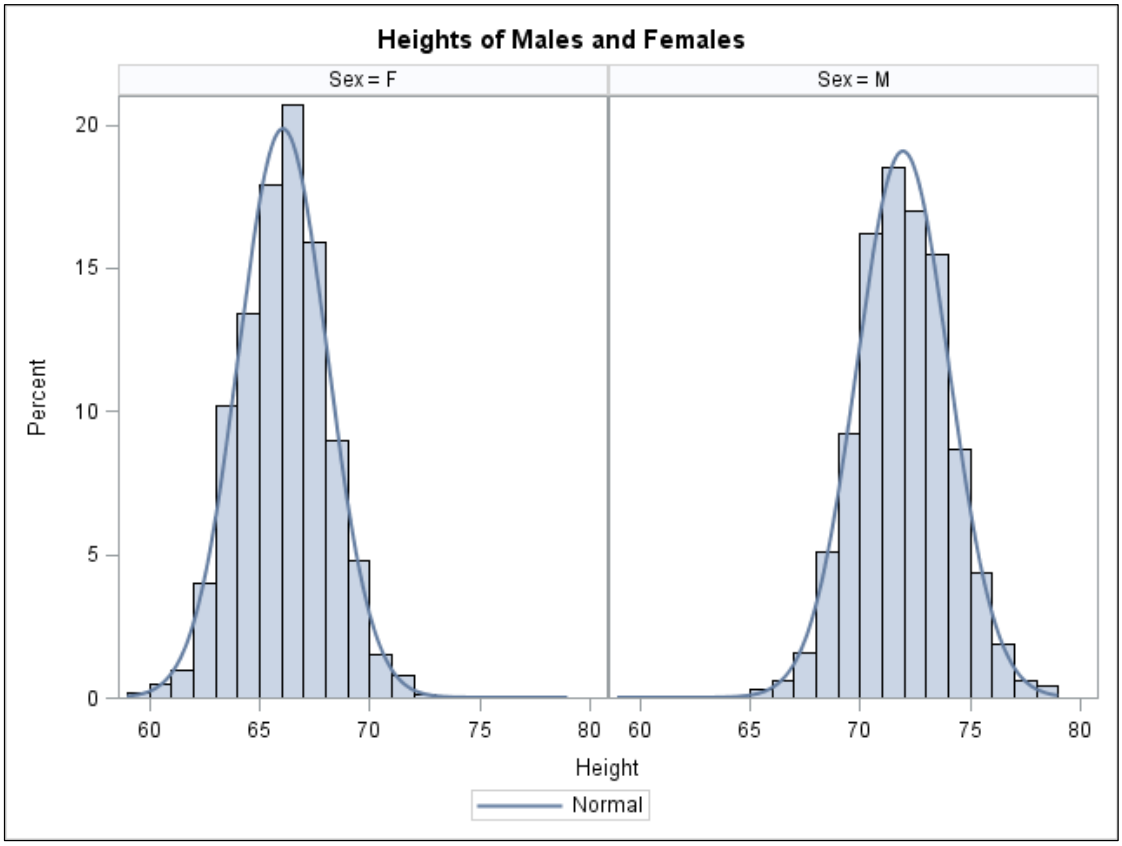
 A comma is needed for any list, character, or numeric values.

2. Use the SGPANEL procedure to plot histograms for each sex.
   a. Overlay a normal density estimate as well.
   b. Provide the title ‘Heights of Males and Females’ and the X-axis title ‘Height’.

```
proc sgpanel data=sp4r.multi;
   panelby sex;
   histogram height;
   density height / type=normal;
   title 'Heights of Males and Females';
   colaxis label='Height';
run;
```

Selected PROC SGPANEL statements:

PANELBY	creates a plot for each category of the proceeding variable and combines the plots in a single panel.
COLAXIS	allows for the specification of the Y-axis options. The YAXIS statement is not used in the SGPANEL procedure.



3. Using the same data table, create three different plots side by side. Use the ODS LAYOUT statement.
   a. Create a template with one row and three columns
   b. Specify a height and width of three inches for each plot.
   c. Use a WHERE statement to plot a histogram, a density estimate plot, and a box plot for Females only.
   d. Provide appropriate titles for each plot.

```
ods layout Start rows=1 columns=3 row_height=(1in) column_gutter=0;

ods region row=1 column=1;
proc sgplot data=sp4r.multi (where= (sex='F'));
   histogram height / binwidth=.5;
   title 'Histogram of Female Heights';
run;

ods region row=1 column=2;
proc sgplot data=sp4r.multi (where= (sex='F'));
   density height / type=kernel;
   title 'Density Estimate of Female Heights';
run;
```

```
ods region row=1 column=3;
proc sgplot data=sp4r.multi (where= (sex='F'));
   hbox height;
   title 'Boxplot of Female Heights';
run;

ods layout end;
```

The WHERE option is specified in parentheses inside the DATA statement. This is useful for plotting only a subset of data.