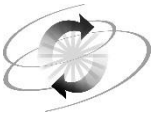


Solution

SP4R04s04.sas

1. Plotting Simple Linear Regression Data

a. Create a data table with $Y = \beta_0 + \beta_1 X + \varepsilon$ where ***X*** ranges from 1 to 30, $\beta_0 = 25$, $\beta_1 = 1$, and $\varepsilon \sim N(\mu = 0, \sigma = 5)$. Keep only the variables ***X*** and ***Y***.

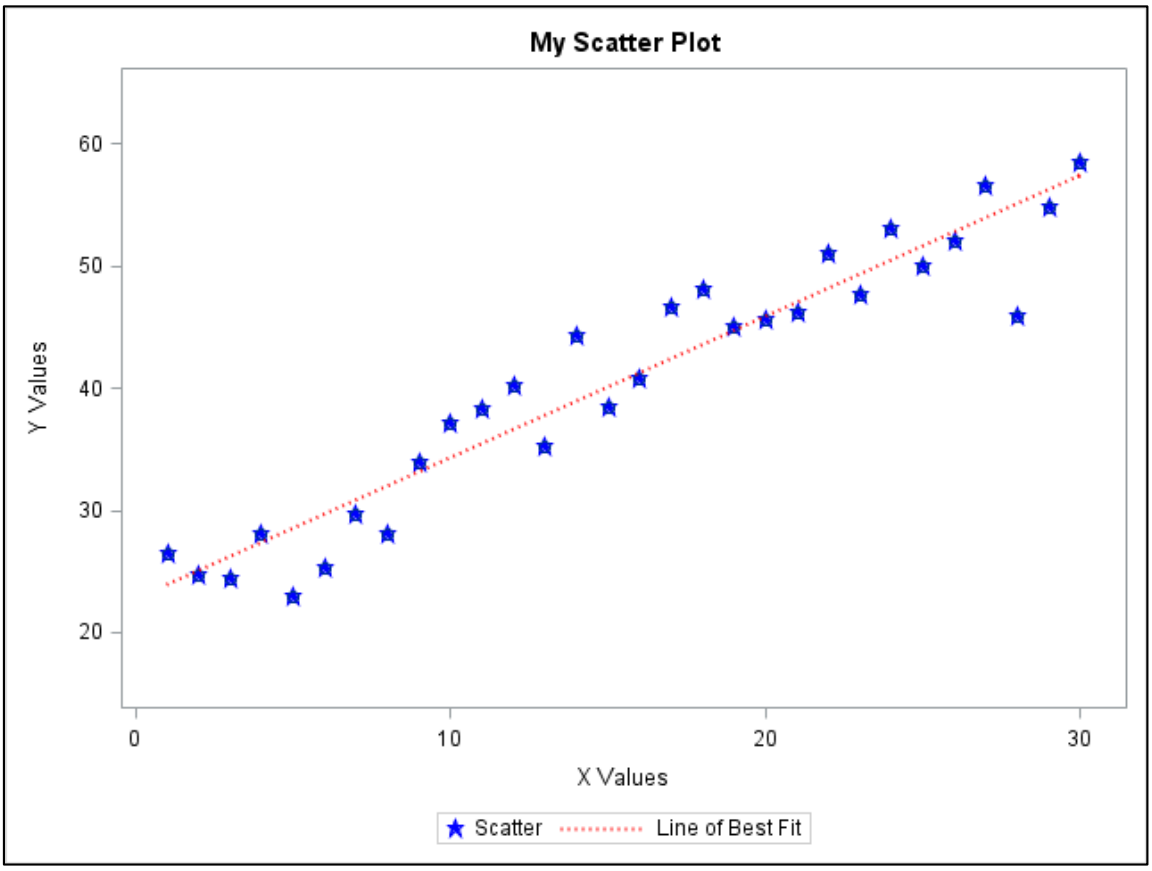
```
data sp4r.simple_lin (keep=x y);
   call streaminit(123);
   do x=1 to 30;
      beta01 = 25;
      beta11 = 1;
      y = beta01 + beta11*x + rand('Normal',0,5);
      output;
   end;
run;
```

b. Use PROC SGPLOT and the REG statement to plot the line of best fit for the data. Create a plot of the data. Use both the SCATTER and REG statement to plot the points and a line of best fit.

1) Enhance the plot by coloring the points blue and using the symbol STARFILLED
2) Color the regression line red and use the pattern DOT.
3) Add a title of your choosing to the X-axis, Y-axis, and the main title.
4) Use the X-axis limits from 0 to 31, and the Y-axis limits from 15 to 65.
5) Name the legend 'Scatter' and 'Line of Best Fit' for both plot types.

```
proc sgplot data=sp4r.simple_lin;
   scatter x=x y=y / legendlabel='Scatter' name='Scatter'
      markerattrs=(color=blue symbol=starfilled);
   reg x=x y=y / legendlabel='Line of Best Fit' name='Line'
      lineattrs=(color=red pattern=dot);

   title 'My Scatter Plot';
   xaxis label='X Values' min=0 max=31;
   yaxis label='Y Values' min=15 max=65;
   keylegend 'Scatter' 'Line';
run;
```



c. Alter the previous plot by changing the SCATTER statement to NEEDLE and the REG statement to PBSPLINE. (This demonstrates the ease in which plot types can be altered.)

```
proc sgplot data=sp4r.simple_lin;
   needle x=x y=y / legendlabel='Needle' name='Needle'
markerattrs=(color=blue symbol=starfilled);
   pbspline x=x y=y / legendlabel='Line of Best Fit' name='Line'
      lineattrs=(color=red pattern=dot);

   title 'My Needle Plot';
   xaxis label='X Values' min=0 max=31;
   yaxis label='Y Values' min=15 max=65;
   keylegend 'Needle' 'Line';
run;
```