

## Solution

SP4R04s06.sas

### 1. Using PROC SGSCATTER

- a. Create a data table with 300 observations and a seed of 123.
  - 1) Let **X** be the deviates from the standard normal distribution.
  - 2) Produce a variable **Y1**, which is **X** plus standard normal deviates.
  - 3) Produce another variable such that **Y2** is 5\***X** plus standard normal deviates.

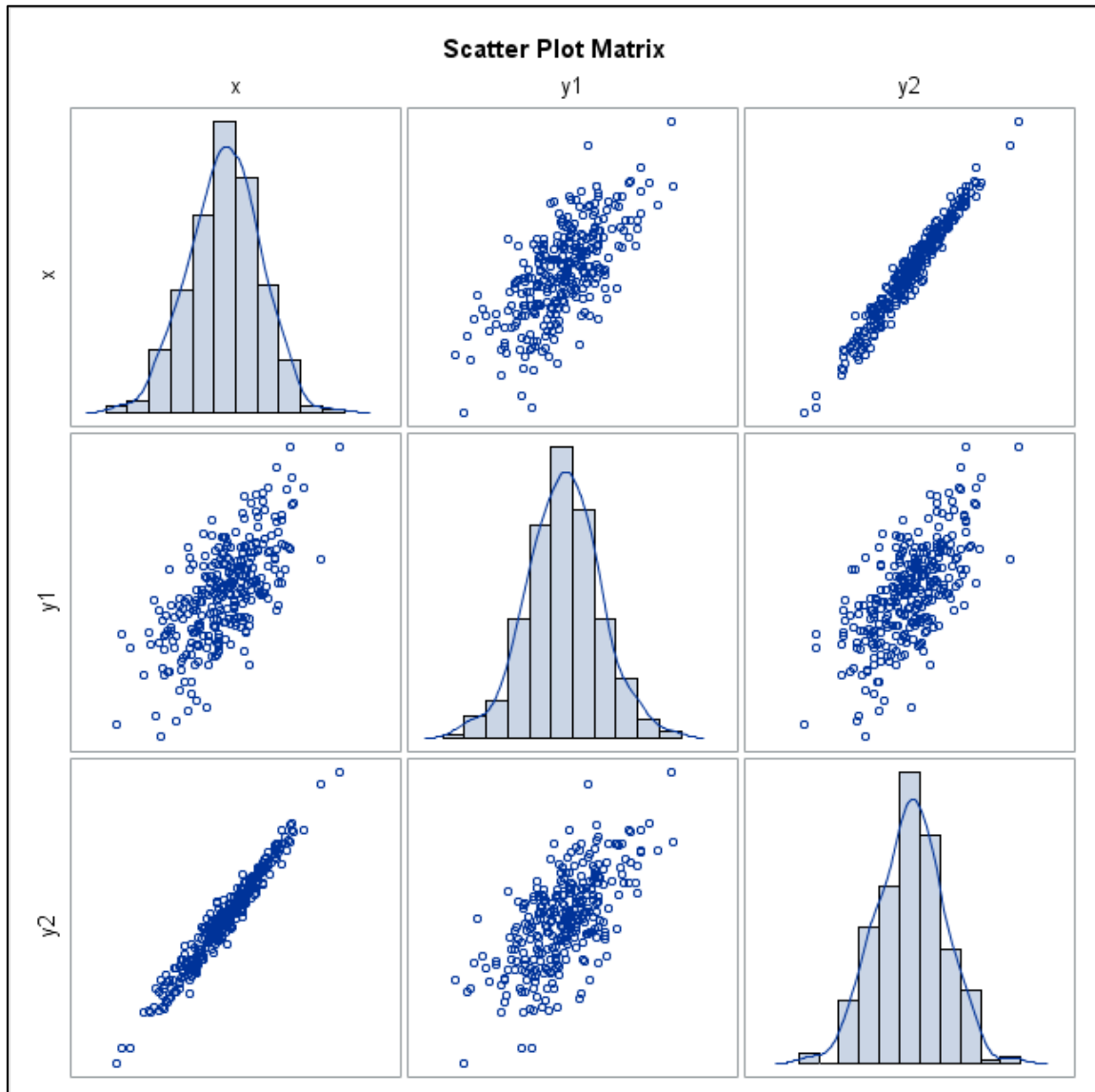
```
data sp4r.random;  
  call streaminit(123);  
  do i=1 to 300;  
    x = rand('Normal');  
    y1 = x + rand('Normal');  
    y2 = 5*x + rand('Normal');  
    output;  
  end;  
run;
```

- b. Use PROC SGSCATTER to create a scatter plot matrix of **X**, **Y1**, and **Y2**. Include histograms and kernel density estimates on the diagonal. (Hint: Look up the **DIAGONAL=** option in the **MATRIX** statement of the SGSCATTER procedure.)

```
proc sgscatter data=sp4r.random;  
  matrix x y1 y2 / diagonal=(histogram kernel);  
  title 'Scatter Plot Matrix';  
run;  
title;
```

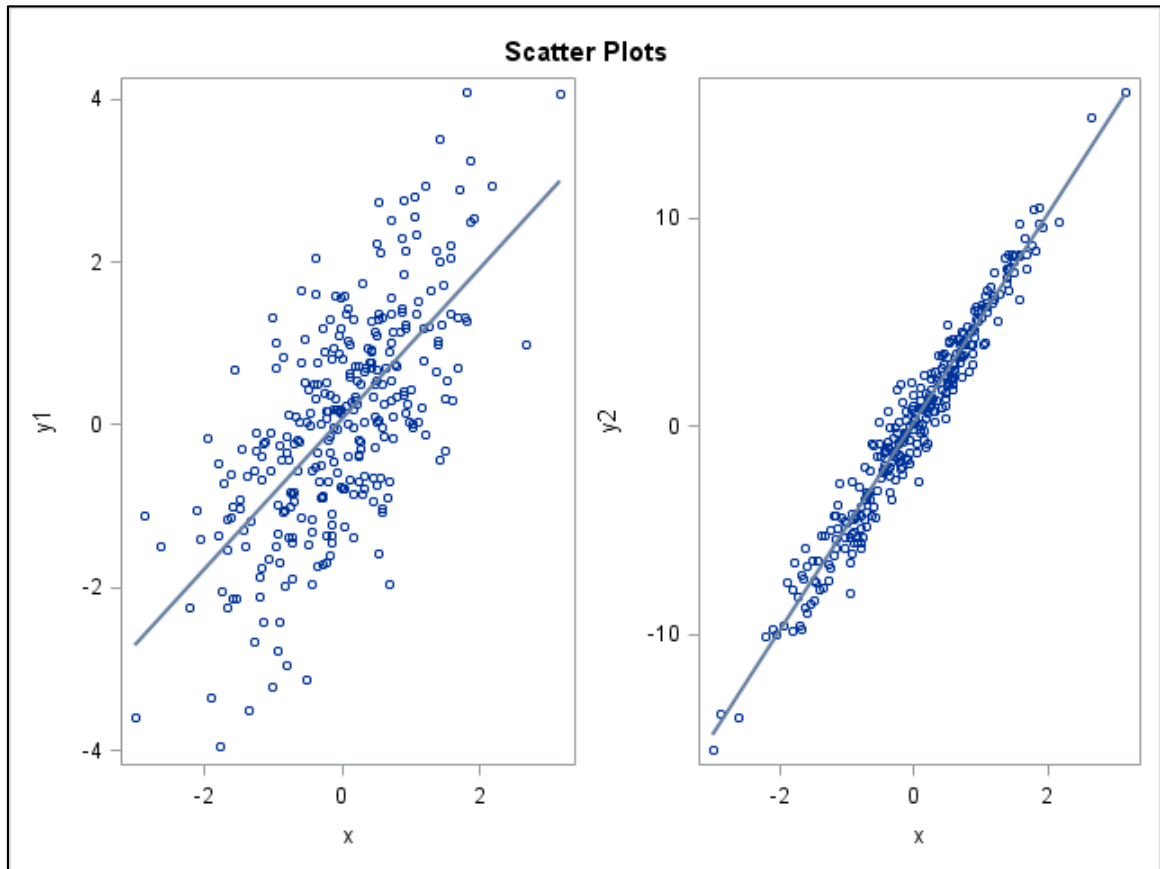
Selected PROC SGSCATTER statement and option:

**MATRIX** specifies the variables used to create a scatter plot matrix. Use the **DIAGONAL=** option to include a histogram, density estimates, or both as the diagonal elements of the scatter plot matrix.



- c. Use PROC SGSCATTER to create side-by-side scatter plots of **Y1** by **X** and **Y2** by **X** with the PLOT statement. Add the regression line to both plots with the REG option.

```
proc sgscatter data=sp4r.random;
  plot (y1 y2) * x / reg;
  title 'Scatter Plots';
run;
title;
```



- d. Use PROC SGSCATTER and the COMPARE statement to create the same scatter plot with shared axes.

```
proc sgscatter;  
  compare y=(y1 y2) x=x / reg;  
  title 'Scatter Plots';  
run;  
title;
```

