CHAPTER 2 Appendix

Examining Relationships with Excel, JMP, Minitab, SPSS, CrunchIt!, R, and a TI-83/-84 Calculator

Scatter Plots

Excel

To create a scatterplot in Excel, the data for the two variables should be placed in two adjacent columns, with the column associated with the Y variable to the right of the column associated with the X variable. If the data file does not have the X variable first, copy and paste the data column(s) into another area of the spreadsheet. Click and drag the mouse to highlight the cells of the two columns of data.

1. Insert ➔ Scatter in the Charts group.
2. Choose the scatterplot option with no line connections.
3. If the gridlines are not desired, you can click on them and delete them by hitting the delete key or by right-clicking and selecting "Delete."
4. To add a grouping variable, watching the video referenced below is recommended. There is no automated procedure in Excel to do this.

Note: The layout of the scatterplot can also be manipulated by choosing among a variety of options offered within the Charts Layouts group found under the Design tab. To add the regression line, click "Add Chart Element," "Trend line," "Linear."

For videos to help with these topics, see the Excel Video Technology Manuals on Scatterplots and Scatterplot by Groups.

1. Analyze ➔ Fit Y by X
2. Select Y variable and click "Y, Response."
3. Select X variable and click "X, Factor."
4. Click "OK."

To add groups, use the Graph ➔ Graph Builder option. Click and drag the two variables to the X and Y axes; drag the grouping variable to the "Color" box.

For videos to help with these topics, see the JMP Video Technology Manuals on Scatterplots and Scatterplot by Groups.

Minitab

1. Graph ➔ Scatterplot
2. Select "Simple" for the type of scatterplot, then click "OK."
3. Click to highlight the name of the data column associated with the Y (response) variable, then select it into row 1 of the Y variables box. Repeat to select the X (explanatory) variable into row 1 of the X variables box. Click "OK."

To add groups, select the "With Groups" option instead of "Simple." Click and highlight the Y and X variables, then add the categorical grouping variable in the box below.

For videos to help with these topics, see the Minitab Video Technology Manuals on Scatterplots and Scatterplot by Groups.

SPSS

1. Graph ➔ Legacy Dialogs ➔ Scatter/Dot
2. Select "Simple Scatter" and click "Define."
3. Select the response variable and click the right arrow next to "Y Axis."
4. Select the predictor variable and click the right arrow next to "X Axis."
5. Click "OK."

To add a categorical grouping variable, select that variable and click to enter it into the box labeled "Set markers by."

For videos to help with these topics, see the SPSS Video Technology Manuals on Scatterplots and Scatterplot by Groups.

CRUNCHIT!

1. Graphics ➔ Scatterplot
2. Select X variable for "X" and Y variable for "Y."
3. In the parameters section you can specify additional options and labels.
4. Click "Calculate."

To add a categorical grouping variable, enter it in the box labeled "Group by."

For videos to help with these topics, see the Crunchit! Help Video on Scatterplots.

TEXAS INSTRUMENTS
TI-83/-84

1. Press "STAT" and select "Edit" to enter the list editor.
2. In L1, enter X values.
3. In L2, enter Y values.
4. Press STAT PLOT (2nd Y) and select the scatterplot.
5. Press 2nd 1 = [L1] for Xlist and 2nd 2 = [L2] for Ylist.
6. Press ZOOM 9.

To add a categorical grouping variable, you will need to have lists for each group. Define up to three scatterplots using different graphing symbols.

For videos to help with these topics, see the TI-83/-84 Video Technology Manuals on Scatterplots and Scatterplot by Groups.

R

1. The simplest command is

```
plot(x,y)
```

2. Add labels and titles and specify a data set by adding other parameters:

```
plot (x,y,main="title",xlab="X label",ylab="Y label",
data=dataset)
```

To add a grouping variable, modify the command to something like

```
plot (x,y,main="title",xlab="X label",ylab="Ylabel",
data=dataset,
bg=c("red","green","blue")[unclass(gpvar)])
```

For videos to help with these topics, see the R Video Technology Manuals on Scatterplots and Scatterplot by Groups.

Correlation

Excel

1. Select "Correlation" in the "Data Analysis" menu box and click "OK."
2. Enter the cell range of the data on the two variables (placed in adjacent columns) into the "Input Range" box. Click "OK."

The Excel Video Technology Manual on Correlation offers more help and an example.

JMP

1. Analyze ➔ Fit Y by X
2. Select Y variable and click "Y, Response."
3. Select X variable and click "X, Factor."
4. Click "OK."
5. Click the Red Triangle ➔ Density Ellipse ➔ 0.95
6. Expand the correlation section.

Alternatively:

1. Analyze ➔ Multivariate Methods ➔ Multivariate
2. Enter at least two variables into "Y, Column."
3. Click "OK."

The JMP Technology Video Manuals on Correlation offers more help and an example.

Minitab

1. Stat ➔ Basic Statistics ➔ Correlation
2. Highlight the variable names and select them into the "Variables" box.
3. Click "OK."

The Minitab Video Technology Manual on Correlation offers more help and an example.

SPSS

1. Analyze ➔ Correlate ➔ Bivariate
2. Select at least two variables and click the right arrow to enter variables.
3. Click "OK."

The SPSS Technology Video Manual on Correlation offers more help and an example.

CRUNCHIT!

1. Statistics ➔ Correlation
2. Check the boxes next to variables you wish to correlate (at least two).
3. Click "Calculate."

The Crunchit! Help Video on Correlation offers more help and an example.

TEXAS INSTRUMENTS
TI-83/-84

1. Press [STAT] and select "Edit" to enter the list editor.
2. In L1 enter X values.
3. In L2 enter Y values.
4. Press [▶] to Calc and select option 8: LinReg(a+bx). Press [ENTER].
5. Press [2nd] [1] [,] [2nd] [2] (for L1,L2 or the lists with the variables of interest) after LinReg(a+bx) in the command window, and press [ENTER].
6. The correlation coefficient is returned along with the linear regression line.

Note: You must first "turn on" the display of r and r^2. Do this by pressing [2nd] [0] (Catalog), [x^{-1}] (C in alpha mode), then arrow down to DiagnosticOn. Press [ENTER] twice.

The TI-83/-84 Technology Video Manual on Correlation offers more help and an example.

The correlation command is

```
cor (x,y)
```

You can specify the data set name with an optional "data=dataset" parameter.
The R Technology Video Manuals on Correlation offers more help and an example.

Least-Squares Regression

Excel

1. Select "Regression" in the "Data Analysis" menu box and click "OK."
2. Enter the cell range of the response data into the "Input Y Range" box and the cell range of the explanatory data into the "Input X Range" box.
3. If you want the residual values, then place a checkmark next to the "Residuals" option (this will produce a residual plot similar to Figure 2.17).
4. Click "OK" to obtain the regression output.
5. If you wish to produce a scatterplot superimposed with the least-squares regression line, then do not select the "Line Fit Plots." Instead, create a scatterplot as described earlier in this Appendix, click the "Design" tab, and select the "Layout 9" option found within the "Quick Layouts" group.

The Excel Technology Video Manuals on Correlation, Fitted Line Plot, Linear Regression, Prediction, and Residual Plots offer more help and examples.

1. Analyze ➔ Fit Y by X
2. Select Y variable and click "Y, Response."
3. Select X variable and click "X, Factor."
4. Click "OK."
5. Click the Red Triangle ➔ Fit Line

Note: To show diagnostic plots, click the red triangle next to "Linear Fit" and select "Plot Residuals." You can save residuals and predicted values from this menu.

The JMP Technology Video Manual on Linear Regression offers more help and examples.

Minitab

1. Stat ➔ Regression ➔ Regression ➔ Fit Regression Model
2. Click in the response variable data column into the "Response" box.
3. Click in the explanatory variable data column into the "Continuous Predictors" box.
4. If you want the residual values, click the "Storage button" and place a check-mark next to the "Residuals" option, then select the "Fits" option if you want the predicted values. Click "OK" to close this window.
5. Residual Plots can be created using the "Graphs" option in the regression menu box. The "4 in 1" plot does them all.
6. Click "OK" to obtain the regression output.

If you wish to produce a scatterplot superimposed with the least-squares regression line, do the following pull-down sequence:

1. Stat ➔ Regression ➔ Fitted Line Plot
2. Click in the response variable data column into the "Response (Y)" box and the explanatory variable data column into the "Predictor (X)" box.
3. Click "OK."

The Minitab Technology Video Manuals: Correlation, Fitted Line Plot, Linear Regression, Prediction, and Residual Plots offer more help and examples.

1. Analyze ➔ Regression ➔ Linear
2. Enter outcome variable in "Dependent Variable."
3. Enter predictor variable in "Independent(s)."
4. To save residuals, click "Save" and check the box next to the type of residual you wish to save (typically, Unstandardized).
5. Create a new scatterplot with these as the response (Y) variable.
6. Click "OK."

The SPSS Technology Video Manuals on Correlation, Fitted Line Plot, Linear Regression, Prediction, and Residual Plots offer more help and examples.

1. Statistics ➔ Regression ➔ Simple Linear
2. Enter outcome variable in "Dependent Variable."
3. Enter predictor variable in "Independent Variable."
4. Click "Calculate."

For residual plots or a fitted line plot, change the option in the "Display" drop-down.

The Crunchit! Help Video on Simple Linear Regression offers more help and examples.

TEXAS INSTRUMENTS
TI-83/-84

1. Press [STAT] and select "Edit" to enter the list editor.
2. In L1, enter X values.
3. In L2, enter Y values.
4. Press [STAT] and select "CALC" and select 8:LinReg(a+bx). Press [ENTER].
5. Enter L1,L2 after LinReg(a+bx) in the command window, and press [ENTER].
6. The correlation coefficient is returned along with the linear regression line if you have turned on Diagnostics (see the previous section, Correlation).

For a fitted line plot, save the regression equation into the Y1 variable by modifying the command by following adding the following keystrokes: [,]. [VARS], [▶] (to Y-Vars), [ENTER] (for Function) and [ENTER] (for Y1). The command should look like

```
LinReg(a+bx) L1,L2,Y1
```

Press [ZOOM] [9] to display the scatterplot again.

For residuals plots, find the list of residuals (RESID) in the [2nd], [STAT] = LIST menu.

The TI-83/-84 Technology Video Manuals on Correlation, Fitted Line Plot, Linear Regression, Prediction, and Residual Plots offer more help and examples.

1. The simplest regression command is

```
> lm(y~x)
```

2. To replicate output and get a complete table of results, use a command like

```
> model<-lm(y~x)
```

```
> summary(model)
```

3. To create diagnostic (residuals) plots, access them as resid(model) as in

```
> hist(resid(model))
```

The R Technology Video Manuals on Correlation, Fitted Line Plot, Linear Regression, Prediction, and Residual Plots offer more help and examples.