

Appendix

Multiple Regression with Excel, JMP, Minitab, SPSS, CrunchIt!, R, and TI-83/-84 Calculators

Multiple regression in most cases uses the same regression dialogs as simple linear regression with additional predictor variables. See Chapters 2 and 12 for more details.



Examine the data for relationships. Excel cannot produce the scatterplot matrices that other technologies can. You will have to plot each pair of variables separately. If all the variables are in consecutive columns (the predictors will have to be for the regression), use **Data → Data Analysis → Correlation** to find the correlation for each pair.

Performing a multiple regression works just like simple regression; all the predictor variables must be in side-by-side columns. Specify the range of predictors as a block, for example, b1:h23.



1. To examine whether relationships are linear and whether predictor variables may be related to each other, click **Graph → Scatterplot Matrix**. Enter all variables into the “Y, Columns” role. Click **OK**.
2. Click **Analyze → Fit Model**.
3. Select and enter the response (Y) variable. Enter the predictor variables into the box labeled “Construct Model Effects.”
4. Create interaction variables by selecting two or more variables in the list and clicking “Cross.” Be careful here (and with transforms to create squared terms etc.). The JMP default is to “center” these by subtracting the mean. If you do not want that, click the red triangle next to “Model Specification” and uncheck “Center Polynomials.” If you’re using a nominal predictor variable, JMP will automatically add $k - 1$ indicator variables for you.
5. While JMP can create transforms of variables (exponentiating, squaring, etc.) in this dialog, results may not be what was intended. For that reason, we prefer computing a new variable using **Cols → Formula** before starting the fitting process.
6. Click **Run** to do the calculations after all variables have been entered and options checked.
7. If confidence intervals for parameters are desired, click the red triangle in the output next to Response and select **Regression Reports → Show all Confidence Intervals**.
8. Confidence intervals and prediction intervals for observed values are obtained by clicking the red triangle and selecting **Save Columns → Mean (or Indiv) Confidence Interval**.

For a video that shows how to use JMP here with an example, see the JMP Video Technology Manual, Multiple Regression: Fitting and Inference.



1. Examine relationships among the variables.
 - a. Plot them to look for linear (or curved) relationships. Click **Graph** → **Matrix Plot**. Select **Simple** (the default) and click **OK**. Select the variables of interest and click **OK**.
 - b. Examine pairs of variable for correlation (both among predictors—looking for multicollinearity) and with the response variable. Click **Stat** → **Basic Statistics** → **Correlation**. Click to enter all the variables of interest. Click **OK**.
2. Click **Stat** → **Regression** → **Regression** → **Fit Regression Model**.
3. Enter the list of predictors in the (Continuous) predictor box. If you have a categorical predictor with k categories, enter that in the “Categorical predictors” box. Minitab will create $k - 1$ indicator variables for you in fitting the model.
4. For interactions or powers of variables, click **Model** after entering the basic predictors. Selecting two variables in the dialog and clicking **Add** next to “Interactions Through Order 2” will add the basic interaction term. More explicitly, you can use **Calc** → **Calculator** to create the new variable as a function of the old ones before doing the regression. Clicking **Add** next to “Terms Through Order” (default is 2) would add the square of a predictor into a model.

Residuals plots are obtained just as they were in Chapter 2; prediction and confidence intervals for responses are still done through **Stat** → **Regression** → **Regression** → **Predict**.

For a video that shows how to use Minitab here with an example, see the JMP Video Technology Manual, Multiple Regression: Fitting and Inference.



1. Examine relationships among the variables:
 - a. Click **Graphs** → **Chart Builder** → **Scatter/Dot**. Select the plot type that looks like four blocks (the scatterplot matrix). Drag all the quantitative variables of interest into the Scattermatrix? area. Click **OK**.
 - b. Examine correlations between the variables. Click **Analyze** → **Correlate** → **Bivariate**. Click to enter all variables of interest. Click **OK**.
2. Click **Analyze** → **Regression** → **Linear** for a model that is linear in all predictors.
3. Click **Analyze** → **Regression** → **Curve Estimation** for a model that uses only one predictor variable but has powers of that variable (e.g., quadratic or cubic regression).
4. To add quadratic or cubic terms to a general linear model, use **Transform** → **Compute Variable** to create squares or cubes of individual predictors. To create indicator variables, use **Transform** → **Create Dummy Variables**. Click to enter the categorical variable into the “Create Dummy Variables for” box. Enter a category name into the “Root Names” box and click **OK**. SPSS will create one indicator variable for each value of the categorical variable. Do NOT use all of them in creating a model!

Residuals plots are still defined using **Plots** in the dialog and prediction and confidence intervals using the **Save** dialog option.

For a video that shows how to use SPSS with an example, see the SPSS Video Technology Manual, Multiple Regression: Fitting and Inference.

CRUNCHIT!

1. Examine relationships among the variables:
 - a. CrunchIt! cannot create scatterplot matrices. You will have to use **Graphics → Scatterplot** for each pair of variables of interest.
 - b. Examine the correlations between the variables. Click **Statistics → Correlation**. Click the box next to each variable of interest. Click **Calculate**.
2. Click **Statistics → Regression → Multiple Linear**.
3. Use the drop-down to select the response variable and check the boxes to select the predictor variables.
4. Use the second drop-down to select numeric results or one of two residuals plots.
5. Click **Calculate**.

If you want to add an interaction or power term to the model, use **Insert → Evaluate Formula** to create the new variable before trying the regression. If you want to create indicator variable(s) for a categorical variable, use an “If” statement in the formula. For example, `if([sex]==”M”,1,0)` will create a variable where Males are 1 and females are 0.

CrunchIt! cannot do confidence or prediction intervals for multiple regression responses.

For more information (and an example), see the CrunchIt! Help Video, Multiple Linear Regression.



TI-83/-84

These TI calculators cannot perform multiple regression.



There are many commands to create scatterplot matrices, but most depend on packages. The simplest that requires no package uses the “pairs” command as illustrated below.

```
> pairs(~y+x1+x2+x1:x2, data=dataframe)
```

To compute a correlation matrix for all numeric variables in a data frame, use a command like the following (note the use of square brackets):

```
> cor(dataframe[, unlist(lapply(dataframe, is.numeric))])
```

The basic fitting command is of the form

```
> model<-lm(y~x1+x2+x1:x2, data=dataframe)
```

The `x1:x2` portion of the statement is an interaction term. You can use `x1:x1` to create a quadratic term in your model.

For full information on the regression, use

```
> Summary(model)
```

For prediction and confidence intervals, use a command like

```
> predict(results, data.frame(x1=1000, x2=20000), interval="confidence")
```

For prediction intervals, replace “confidence” with “prediction” in the command.

For more information and an example, see the R Video Technology Manual videos, Multiple Regression: Fitting Models and Multiple Regression: Inference.