Multivariate Statistics

SPSS Commands

Test for Normality

1. ***Analyze -> Descriptive Statistics -> Explore***
2. Click the variable you want to test for normality in the box on the left side of the window, and then drag it into the ***Dependent List*** box on the right side.
3. Click the ***Statistics*** button at the upper right corner of the window to open the Statistics pop-up window. Don't make any changes to the options displayed and click the ***Continue*** button.
4. Click the ***Plots*** button on the ***Explore*** window. The Plots window opens. Select ***None*** in the ***Boxplots*** section and de-select all options in the ***Descriptive*** section.
5. Select the ***Normality plots with tests*** check box, and then click Continue.
6. ***OK***
7. Review the results in the ***Tests of Normality*** box to determine if the data is normally distributed. The box displays the results from two tests; the Kolmogorov-Smirnov test and the Shapiro-Wilk test. The Kolmogorov-Smirnov test is used to test large data sets while the Shapiro-Wilk test is more appropriate for a smaller sample, such as 50 numbers or less. If the "Sig" column of either test is above 0.05, your data is normally distributed.

Bivariate Correlation and Regression

Requesting a Scatterplot

1. Select

   Graphs -> Legacy Dialogs -> Scatter/Dot...

2. Select ***Simple Scatter***, and then click the ***Define*** button.
3. Move the independent variable to ***Y Axis*** and the independent variable to ***X Axis***
4. Click ***Titles***. Choose a title.
5. ***Continue.***
6. ***OK.***

Optional Features

1. Double-click on the scatterplot to edit it.
2. In the ***Chart Editor*** that opens, select ***Fit Line at Total*** from the ***Elements*** menu.
3. In the ***Properties*** window that opens, select ***Linear***.
4. Click ***Close*** to proceed.

Another useful option for a scatterplot is a label for each data point. The label can be a case number or a value obtained from a variable in the ***SPSA Statistics Data Editor***:

1. In the ***Simple Scatterplot*** filed, click ***Options***.
2. Select the option ***Display chart with case labels***.
3. ***Continue***.
4. ***OK***.

Correlation Analysis

1. Select

   Analyze -> Correlate -> Bivariate
2. Select the two variables.
3. Select ***Pearson*** as the ***Correlation Coefficients***.
4. ***OK***.

Requesting a Regression Analysis

1. From the ***Analyze*** menu select ***Regression*** then ***Linear***.
2. Click on the name of the *independent variable* to highlight it, then transfer it to the ***Independent(s)*** field.
3. Click on the name of the dependent variable to highlight it and then transfer it to the ***Dependent*** field.
4. Click on ***Statistics*** and select the following statistics:

   Estimates, Model fit, Descriptives, Casewise diagnostics, Outliers outside
5. ***Continue***.
6. Click on ***Plots...***
7. Select ***Normal probability plot***.
8. ***Continue.***
9. ***Save...***

10. Select

    - ***Unstandardised*** in ***Predicted Values.***
    - ***Unstandardised*** and ***Studentized*** in ***Residuals***.
    - ***Mean*** (to obtain a confidence interval output in the ***Data*** window and ***Individual*** in ***Prediction Intervals*** to obtain a prediction interval output in the **Data** window at the 95% level (or whatever level the problem requires).

11. ***Continue***.

12. ***OK***

Multiple Linear Regression

1. Select
   Analyze -> Regression -> Linear...
2. Move the dependent variable to the ***Dependent*** window and the independent variable to the ***Independent(s)*** window**.**
3. Select the ***Backward*** method.
4. Click ***Statistics...***
5. Select ***Estimates*** and ***Confidence intervals*** for the ***Regression*** *Coefficients.*
6. Select ***Model fit*** and ***Collinearity diagnostics***.
7. ***Continue***.
8. Click ***Plots...***
9. Select "***Normal Probability Plot***", "***Produce all partial plots***".
10. ***Continue***.
11. Click ***Save…***

12. Select in ***Predict Values*** part ***Unstandardised*** and ***S.E. of mean predictions***.

13. Select in ***Residuals*** part ***Unstandardised*** and ***Studentized***.

14. Select in ***Prediction Intervals*** part ***Mean*** and ***Individual*** at for example 95% level.

15. ***Continue***

16. ***OK***

17. Select

    Graphs -> Legacy Dialogs -> Scatter Dot

18. Select

    Simple Scatter -> Define

19. Move the residual ***RES_1*** to ***Y Axis*** and ***PRE_1*** to ***X Axis***.

20. Click ***Titles*** to enter ***Studentized Regional Plot*** as the title for your graph.

21. ***Continue***

22. ***OK***

23. Double-click the resulting graph in the ***Output*** window.

24. Select

    Options -> Y Axis Reference Line

25. Select the ***Reference Line*** tab in the ***Properties*** window, add ***Position*** of line 0, and click

    Apply

26. Click the ***Close*** box to exit the ***Charter Editor***

One-Way ANOVA

1. Select

   Analyze -> Compare Means -> One-Way ANOVA...

2. Place the dependent variable in the ***Dependent List*** box and the factor in the ***Factor*** box.

3. Select ***Options...***

4. Choose ***Descriptive, Homogeneity of variance test, Means plot***.

5. ***Continue***

6. Select ***Post Hoc...***

7. Choose ***LSD*** (***Tukey, Scheffe***)

8. ***Continue***

9. ***OK***

10. Select

    Analyze -> Descriptive Statistics -> Explore...

11. Click the ***Plots*** button. Click the ***Plots***

12. Move the dependent variable to ***Variable*** and the factor to ***Category Axis***

13. Choose ***Normality plots with tests, Untransformed***

14. ***Continue***

15. ***OK***

16. Select

    Graphs -> Legacy Dialogs...

17. Choose ***Error Bar...***

18. Choose ***Simple***, ***Summaries for groups of cases***

19. ***Define***

20. Move the dependent variable to ***Variable*** and the factor to ***Category Axis***

21. Set the ***Bars represent*** Option to be ***Standard error of mean***

22. Use a ***Multiplier 1***

23. ***OK***

Two-Way ANOVA

1. From the menu at the top of the scree, click on ***Analyze***, then click on ***General Linear Model***, then on ***Univariate***.

2. Click on the dependent and move it into the box labeled ***Dependent variable***.

3. Click on the two independent variables and move them into the box labeled ***Fixed Factors***.

4. Click on the ***Options*** button.

    - Click on ***Descriptive Statistics***, ***Estimates of Effect Size*** and ***Homogeneity tests***.
    - Click on ***Continue***.

5. Click on the ***Post Hoc*** button.

    - From the ***Factors*** listed on the left-hand side, choose the independent variables.
    - Click on the arrow button to move it into the ***Post Hoc Tests for*** section.
    - Choose the test you wish to use (for instance, ***Tukey***).
    - Click on Continue.

6. Click on the ***Plots*** button.

    - In the ***Horizontal*** box, put the variable that has the most groups.
    - In the box labeled ***Separate Lines***, put the other independent variable
    - Click on ***Add***.
    - In the section ***Plots***, you should now see your two variables listed.

7. Click on ***Continue*** and then ***OK***.

Cluster Analysis (K-Means Cluster)

1. Select

    Analyze -> Classify -> K-Means Cluster...

- Choose the relevant variables.

- Move (optionally) the case variable to ***Label Cases by***.

- Check ***Iterate and classify***.

- Choose the ***Number of Clusters***.

2. Select ***Options...***

- Choose ***Cluster information for each case***

- ***Continue***

3. Select ***Save...***

- Choose ***Cluster membership***

- ***Continue***

- ***OK.***

(Last revised: 5.10.2016)