Contents

Preface  xiii

CHAPTER 1  INTRODUCTION  1

1.1  AIMS AND OBJECTIVES  1

1.2  MEASUREMENTS AND DATA  2
1.2.1  Variables: quantitative and qualitative  2
1.2.2  Levels of measurement: scale, ordinal and nominal data  2
1.2.3  Univariate, bivariate and multivariate data sets  3

1.3  DISTRIBUTIONS  4
1.3.1  The three most important properties of a distribution  5
1.3.2  Some common distribution shapes  6

1.4  EXPERIMENTAL VERSUS CORRELATIONAL RESEARCH  7
1.4.1  A simple experiment  8
1.4.2  A more complex experiment  9
1.4.3  A correlational study  10
1.4.4  Quasi-experiments  13

1.5  CHOOSING A STATISTICAL TEST  13
1.5.1  Considerations in choosing a statistical test  13
1.5.2  Five common research situations  14

1.6  A SIGNIFICANT DIFFERENCE?  15
1.6.1  Independent or related samples?  15
1.6.2  Flow chart  16

1.7  ARE TWO VARIABLES ASSOCIATED?  17
1.7.1  Flow chart  18
1.7.2  Measuring association in ordinal data  18
1.7.3  Measuring association in nominal data: contingency tables  19
1.7.4  Multi-way contingency tables  20

1.8  MAKING PREDICTIONS  20
1.8.1  Flow chart  20
1.8.2  Simple regression  21
1.8.3  Multiple regression  21
1.8.4  Predicting category membership  22

1.9  FROM A SINGLE SAMPLE TO THE POPULATION  22
1.9.1  Flow chart  22
1.9.2  Goodness-of-fit: nominal data  23
1.9.3  Inferences about the mean of a single population  23

1.10  THE SEARCH FOR LATENT VARIABLES  24

1.11  MULTIVARIATE STATISTICS  24

http://www.psypress.com/spss-made-simple/
1.12 A FINAL WORD 25

CHAPTER 2 Getting started with IBM SPSS Statistics 19 26

2.1 INTRODUCTION 26

2.2 DESCRIBING THE DATA FROM THE CAFFEINE EXPERIMENT 29

2.2.1 Opening SPSS 29
2.2.2 The SPSS Statistics Data Editor 30
2.2.3 Working in Variable View 30
2.2.4 Working in Data View 35
2.2.5 Computing the group means 39
2.2.6 The SPSS Statistics Viewer 42
2.2.7 The output from the Means procedure 46
2.2.8 Histograms 47
2.2.9 Editing items in the Viewer 52
2.2.10 Ending the session 53
2.2.11 Resuming work on a saved data set 53

2.3 THE FOUR DRUGS EXPERIMENT 53

2.3.1 In Variable View 54
2.3.2 In Data View 56
2.3.3 Using the Means procedure 57
2.3.4 The histograms 59

2.4 PRINTING FROM THE STATISTICS VIEWER 60

2.5 USING SPSS SYNTAX 64

2.5.1 The Syntax Editor 65
2.5.2 Running the Means procedure with Syntax 66
2.5.3 Looping functions in Syntax 69

2.6 A FINAL WORD 69

CHAPTER 3 Editing data sets 70

3.1 MORE ON THE DATA EDITOR 70

3.1.1 A preliminary check on the default settings 70
3.1.2 Inserting new variables 71
3.1.3 Rearranging the order of variables in the Data Editor 72
3.1.4 Inserting case numbers 73
3.1.5 Using Syntax to insert case numbers into a data set 76
3.1.6 Inserting case numbers into an empty Data Editor 76
3.1.7 Using Syntax to insert case numbers into an empty Data Editor 77
3.1.8 Changing the default settings for Width and Decimals 77
3.1.9 String variables 78
3.1.10 Specifying missing values 80
3.1.11 Changing the Alignment settings 82
3.1.12 Opening an SPSS file 82
3.1.13 Entering data from other applications 83
3.1.14 Creating new variables while in Data View 84
3.1.15 Adding new cases while in Data View 85

3.2 VALIDATION OF THE DATA 85

http://www.psypress.com/spss-made-simple/
5.2.5 Editing a bar chart 167

5.3 **ERROR BAR CHARTS** 171

5.4 **BOXPLOTS** 173

5.5 **PIE CHARTS** 173

5.6 **LINE GRAPHS** 174
- 5.6.1 The Visual Binning procedure 175
- 5.6.2 Plotting line graphs 177

5.7 **USING CHART TEMPLATES** 179

5.8 A FINAL WORD 183

---

**CHAPTER 6** Comparing averages: Two-sample and one-sample tests 184

6.1 **OVERVIEW** 184

6.2 **THE INDEPENDENT-SAMPLES T TEST WITH SPSS** 185
- 6.2.1 Running the $t$ test 185
- 6.2.2 Interpreting the output 188
- 6.2.3 Two-tailed and one-tailed $p$-values 190
- 6.2.4 Measuring effect size 191
- 6.2.5 Reporting the results of a statistical test 193

6.3 **THE RELATED-SAMPLES (OR PAIRED-SAMPLES) T TEST** 194
- 6.3.1 Preparing the data file 195
- 6.3.2 Exploring the data 195
- 6.3.3 Running the $t$ test 196
- 6.3.4 Interpreting the output 197
- 6.3.5 Measuring effect size 198
- 6.3.6 Reporting the results of the test 199
- 6.3.7 A one-sample test 199

6.4 **NONPARAMETRIC TESTS** 199
- 6.4.1 Nonparametric tests in SPSS 200
- 6.4.2 Independent samples: The Mann-Whitney $U$ test 201
- 6.4.3 Output from the Mann-Whitney $U$ test 203
- 6.4.4 Exact tests 206
- 6.4.5 Effect size 207
- 6.4.6 The report 208

6.5 **THE WILCOXON MATCHED-PAIRS TEST** 209
- 6.5.1 The Wilcoxon matched-pairs tests in SPSS 209
- 6.5.2 The output 210
- 6.5.3 Effect size 214
- 6.5.4 The report 215

6.6 **THE SIGN AND BINOMIAL TESTS** 215
- 6.6.1 The sign test in SPSS 216
- 6.6.2 Bernoulli trials: the binomial test 219

6.7 **EFFECT SIZE, POWER AND NUMBER OF PARTICIPANTS** 223

6.8 A FINAL WORD 225
## Chapter 7  The one-way ANOVA 227

7.1  Introduction 227  
7.2  Entering the data 233  
7.3  Running the one-way ANOVA on GLM 236  
   7.3.1 Finding the GLM menu 236  
   7.3.2 Descriptives and the ANOVA summary table 241  
   7.3.3 The profile plot 243  
   7.3.4 Measuring effect size 244  
   7.3.5 Report of the analysis 246  
7.4  Making comparisons among the treatment means 247  
   7.4.1 Planned and unplanned comparisons 247  
   7.4.2 Linear contrasts 251  
   7.4.3 Helmert contrasts 252  
7.5  Power and effect size in the one-way ANOVA 258  
7.6  Alternatives to the one-way ANOVA 260  
7.7  A final word 266  

## Chapter 8  Between subjects factorial experiments 268

8.1  Introduction 268  
   8.1.1 An experiment with two treatment factors 268  
   8.1.2 Main effects and interactions 270  
   8.1.3 Profile plots 271  
8.2  How the two-way ANOVA works 272  
   8.2.1 The two-way ANOVA 272  
   8.2.2 Degrees of freedom 273  
   8.2.3 The two-way ANOVA summary table 274  
8.3  The two-way ANOVA with SPSS 275  
   8.3.1 Entering the data 276  
   8.3.2 Exploring the data: boxplots 277  
   8.3.3 Choosing a factorial ANOVA 277  
   8.3.4 Output for a factorial ANOVA 280  
   8.3.5 Measuring effect size in the two-way ANOVA 282  
   8.3.6 Reporting the results of the two-way ANOVA 285  
8.4  Further analysis 286  
   8.4.1 A problem with multiple comparisons 286  
   8.4.2 Unpacking significant main effects: post hoc tests 286  
   8.4.3 The analysis of interactions 287  
8.5  Testing for simple main effects with syntax 289  
   8.5.1 Using the MANOVA command to run the univariate ANOVA 289  
   8.5.2 Including simple effects in a MANOVA command 291  
8.6  How many participants shall I need? 297  
8.7  More complex experiments 298  
   8.7.1 Three-way interactions 298  
   8.7.2 The three-way ANOVA 299  
   8.7.3 How the three-way ANOVA works 301  
   8.7.4 The three-way ANOVA with SPSS 302
8.7.5 Follow-up analysis following a significant three-way interaction 304
8.7.6 Testing for simple interactions and simple, simple main effects 305
8.7.7 Unplanned multiple comparisons 308

8.8 A FINAL WORD 309

CHAPTER 9  Within subjects experiments 311

9.1 INTRODUCTION 311
9.1.1 Rationale of a within subjects experiment 311
9.1.2 How the within subjects ANOVA works 312
9.1.3 A within subjects experiment on the effect of target shape on shooting accuracy 315
9.1.4 Order effects: counterbalancing 316
9.1.5 Assumptions underlying the within subjects ANOVA: homogeneity of covariance 317

9.2 A ONE-FACTOR WITHIN SUBJECTS ANOVA WITH SPSS 318
9.2.1 Entering the data 319
9.2.2 Running the one-factor within subjects ANOVA 319
9.2.3 Output for a one-factor within subjects ANOVA 322
9.2.4 Effect size in the within subjects ANOVA 327

9.3 HOW MANY PARTICIPANTS SHALL I NEED? 328

9.4 NONPARAMETRIC EQUIVALENTS OF THE WITHIN SUBJECTS ANOVA 329
9.4.1 The Friedman test for ordinal data 329
9.4.2 Cochran’s Q test for nominal data 333

9.5 THE TWO-FACTOR WITHIN SUBJECTS ANOVA 336
9.5.1 Preparing the data set 337
9.5.2 Running the two-factor within subjects ANOVA 338
9.5.3 Output for a two-factor within subjects ANOVA 341
9.5.4 Unpacking a significant interaction with multiple comparisons 344

9.6 A FINAL WORD 346

CHAPTER 10  Mixed factorial experiments 349

10.1 INTRODUCTION 349
10.1.1 A mixed factorial experiment 349
10.1.2 Classifying mixed factorial designs 351
10.1.3 Rationale of the mixed ANOVA 351

10.2 THE TWO-FACTOR MIXED FACTORIAL ANOVA WITH SPSS 354
10.2.1 Preparing the SPSS data set 354
10.2.2 Running the ANOVA 355
10.2.3 Output for the two-factor mixed ANOVA 357
10.2.4 Simple effects analysis with syntax 364

10.3 THE THREE-FACTOR MIXED ANOVA 369
10.3.1 The two three-factor mixed designs 369
10.3.2 Two within subjects factors 370
10.3.3 Using syntax to test for simple effects 375
10.3.4 One within subjects factor and two between subjects factors: the A×B×(C) mixed factorial design 380

http://www.psypress.com/spss-made-simple/
### Chapter 10: The Multivariate Analysis of Variance (MANOVA) 387

- 10.4.1 What the MANOVA does 387
- 10.4.2 How the MANOVA works 389
- 10.4.3 Assumptions of the MANOVA 392
- 10.4.4 Application of the MANOVA to the shape recognition experiment 393

### Chapter 11: Measuring Statistical Association 399

- 11.1.1 Introduction 399
- 11.2.1 Formula for the Pearson correlation 402
- 11.2.2 The range of values of the Pearson correlation 403
- 11.2.3 The sign of a correlation 404
- 11.4.1 Spearman’s rank correlation 409
- 11.4.2 Kendall’s tau statistics 410

### Chapter 12: Regression 448

- 12.1.1 Simple, two-variable regression 449
- 12.1.2 Residuals 450
12.1.3 The least squares criterion for ‘the best-fitting line’ 451
12.1.4 Regression and correlation 452
12.1.5 The coefficient of determination revisited 453
12.1.6 Shrinkage with resampling: cross-validation 455
12.1.7 Beta coefficients 456
12.1.8 Effects of linear transformations on correlation and regression coefficients 458

12.2 SIMPLE REGRESSION WITH SPSS 458
12.2.1 Drawing scatterplots with regression lines 458
12.2.2 The output for simple regression 464

12.3 MULTIPLE REGRESSION 469
12.3.1 The multiple regression equation 469
12.3.2 Partial and semipartial (part) correlations 471
12.3.3 Measuring the importance of an IV in multiple regression 473
12.3.4 Strategies in multiple regression 474

12.4 MULTIPLE REGRESSION WITH SPSS 475
12.4.1 Running a simultaneous multiple regression with SPSS 475
12.4.2 The output from simultaneous multiple regression 478
12.4.3 An hierarchical multiple regression 482

12.5 RUNNING THE ANOVA AS A MULTIPLE REGRESSION 486
12.5.1 The two-group case 486
12.5.2 The k-group case 492
12.5.3 Other systems of coding: contrast coding and effects coding 498
12.5.4 The factorial ANOVA as a regression 499

12.6 MULTILEVEL REGRESSION MODELS 503

12.7 A FINAL WORD 504

CHAPTER 13 The analysis of covariance (ANCOVA) 507

13.1 INTRODUCTION 505
13.2 THE ANCOVA WITH SPSS 508
13.2.1 Preliminary analysis 508
13.2.2 The five within groups regression lines 509
13.2.3 The ANCOVA 511
13.2.4 Further analysis 513

13.3 THE ANCOVA AS AN HIERARCHICAL MULTIPLE REGRESSION 518
13.3.1 Setting up the regression 519
13.3.2 Running the regression 520
13.3.3 Interpreting the output 522

13.4 A FACTORIAL ANCOVA 525

13.5 A FINAL WORD 529

CHAPTER 14 Analyses of multiway frequency tables 531

14.1 INTRODUCTION 531
14.2 SOME BASICS OF LOGLINEAR MODELLING 532
14.2.1 Loglinear models and ANOVA models 532
14.2.2 Model-building and the hierarchical principle 534

http://www.psypress.com/spss-made-simple/
14.2.3 The main-effects-only loglinear model and the traditional chi-square test for
association 536
14.2.4 Analysis of the residuals 537
14.3 MODELLING A TWO-WAY CONTINGENCY TABLE 538
14.3.1 SPSS procedures for loglinear analysis 538
14.3.2 Fitting an unsaturated model 543
14.3.3 Summary 548
14.4 MODELLING A THREE-WAY FREQUENCY TABLE 548
14.4.1 Exploring the data 549
14.4.2 Loglinear analysis of the data on gender and helpfulness 550
14.4.3 The main-effects-only model and the traditional chi-square test 555
14.4.4 Collapsing a multi-way table: the requirement of conditional independence 557
14.4.5 An alternative data set for the gender and helpfulness experiment 558
14.4.6 Reporting the results of a loglinear analysis 561
14.5 A FINAL WORD 562

CHAPTER 15 Predicting category membership: logistic regression 564
15.1 INTRODUCTION 564
15.1.1 Logistic regression 565
15.1.2 Binary and multinomial logistic regression 565
15.2 BINARY LOGISTIC REGRESSION 566
15.2.1 How logistic regression works 568
15.2.2 A binary logistic regression with quantitative variables 574
15.2.3 A binary logistic regression with categorical independent variables 584
15.3 MULTINOMIAL LOGISTIC REGRESSION 588
15.3.1 Accessing the data set 589
15.3.2 Running multinomial logistic regression 590
15.5 A FINAL WORD 599

CHAPTER 16 The search for latent variables: factor analysis 601
16.1 INTRODUCTION 601
16.1.1 Stages in an exploratory factor analysis 603
16.1.2 The extraction of factors 604
16.1.3 The rationale of rotation 604
16.1.4 Some issues in factor analysis 604
16.1.5 Some key technical terms 605
16.1.6 Preliminaries 606
16.2 AN EXPLORATORY FACTOR ANALYSIS 607
16.2.1 Entering the data for a factor analysis 607
16.2.2 Running a factor analysis on SPSS 608
16.2.3 Output for factor analysis 610
16.3 USING SPSS SYNTAX TO RUN AN EXPLORATORY FACTOR ANALYSIS 619
16.3.1 Procedure with the raw data as input 619
16.3.2 Procedure with a correlation matrix as input 620
CHAPTER 5

More on graphs and charts

5.1 Introduction
5.2 Bar charts
5.3 Error bar charts
5.4 Boxplots
5.5 Pie charts
5.6 Line graphs
5.7 Using chart templates
5.8 A final word.

5.1 INTRODUCTION

SPSS offers a wide range of graphs and charts, some of which we have already made use of in Chapters 2, 3 and 4. In this chapter, we shall look more closely at some aspects of graph-drawing in SPSS.

For monochrome printing, the Chart Editor can be used once a graph or chart has appeared in the Viewer to remove the colours and replace them with patterns. It is better, however, to change the default settings beforehand as follows:

- Choose Edit➜Options… and select the Charts tab in the Options dialog box.
- Within the Style Cycle Preference selection panel, select Cycle through patterns only.
- Click Fills… . Select the pattern you want for Simple Charts and delete the empty pattern box in Grouped Charts by clicking the radio button for Grouped Charts, selecting the empty box pattern and clicking Remove. Click Continue.
- Click Apply, then OK.
5.2 BAR CHARTS

This section describes the production of simple bar charts, clustered bar charts, and panelled bar charts.

5.2.1 Simple bar charts

A simple bar chart summarizes the distributions of a scale or continuous variable at different levels of one categorical variable only, such as the experimental condition under which the participants in a study performed. We shall use the data from the caffeine experiment in Ch5 Caffeine and Gender.sav at http://www.psypress.com/spss-made-simple.

A simple bar chart for comparing the means scores of groups of participants such as those in the caffeine experiment (Chapter 2) can be obtained as follows:

- Choose Graphs ➔ Chart Builder…

- A warning box will appear asking you to ensure that each variable has been defined in the Measure column of Variable View as Scale, Ordinal or Nominal, and that the values of categorical variables have been labelled. Should you have forgotten to do either of these things, you can enter the information at this point by completing the warning dialog. Otherwise, click OK to continue.

- In the Choose From list, highlight Bar to display, in the Gallery, pictures of the different kinds of bar chart. Click the first (top left) picture of simple bars to highlight it and then drag the template to the Chart preview in the panel above (Figure 1).

- In the Element Properties dialog box, check Display error bars and Apply to return to the Chart Builder and see the error bars included in the template.

- In the Variables list, click Score and drag it to the Y-Axis drop zone in the Chart preview. Move Experimental Condition to the X-Axis drop zone.

- To add a title, click Titles/Footnotes tab and then, in the list of check-boxes that will appear in place of the gallery of graphics choices, click Title 1. A panel will appear in the Element Properties dialog box, where a title such as ‘Means and 95% Confidence Intervals’ can be typed in. Click Apply, followed by Close. The marker T1 will appear at the top of the preview if a title has been requested.

- Finally, back in the Chart Builder, click OK to create the chart (Output 1).
5.2.2 Clustered bar charts

A **clustered bar chart** is a graph in which, instead of only a single bar over each of the categories on the horizontal axis, there is a **cluster** of bars, each bar in the cluster representing a category in a second categorical variable. Suppose that, in addition to the Experimental Condition variable, the data set also contained the Gender of the participants. A clustered bar chart could then be plotted with Experimental Condition as the first categorical variable and Gender as the clustering variable.
- Select Bar from the **Choose from** list. The appropriate array of choices will appear in the **Gallery** (Figure 2).

- Drag the **Clustered Bar** template into the **Chart Preview** area.

- From the **Variables** list, drag the three variables into their appropriate drop zones (Figure 3).

- In **Element Properties**, check **Display error bars** and click **Apply** to see the error bars added to the outline in the Chart preview (Figure 4).

---

**Figure 2.** The *bar chart gallery*

**Figure 3.** The three drop zones

http://www.psypress.com/spss-made-simple/
Figure 4. Chart preview after specification of the cluster variable and error bars

- Click **OK** to produce the clustered bar chart (Output 2).
5.2.3 Panelled bar charts

In Chapter 2, we showed how to display histograms for different groups either in rows or in columns. This operation is known as **panelling**. Bar charts can be panelled in exactly the same way as can histograms. Suppose that we want to display simple bar charts of the scores in the Caffeine experiment for the males and females in a side-by-side panelled display. Proceed as follows.

- Select **Bar** from the **Choose from** list. The appropriate array of choices will appear in the **Gallery**.
- Choose **Simple Bar** from the **Gallery** and drag the template into the **Chart Preview** area.
- From the **Variables** list, click and drag **Score** into **Y-Axis** drop zone in the **Chart Preview** and similarly transfer the grouping variable to the **X-Axis** drop zone.
- Click on **Element Properties**, order error bars and click **Apply** to return to the **Chart Builder** to see the error bars in the Chart preview.
- In the **Chart Builder**, click the **Groups/Point ID** tab and check **Columns panel variable**. (**Columns panelling** displays the figures in a row; **Rows panelling** displays them in a column.)
- Drag the grouping variable into the drop zone labelled **Panel?** (Figure 5)
- Click **OK** to produce the panelled bar charts, which are shown in Output 3.

![Figure 5. Drop zone for the panelling variable](http://www.psypress.com/spss-made-simple/)
Output 3. Panelled bar charts

5.2.4 Editing a bar chart

The figure in Output 3 has been slightly edited: the thickness of the bars has been adjusted; and shading has been introduced to distinguish more sharply between the Caffeine and Placebo groups. The original version in the Viewer had thicker bars and no shading.

- Double-click the chart to open the Chart Editor (Figure 6). In the Editor, you will see the unedited figure that first appeared in the Viewer.
- Double-click on the bars to open the Properties dialog box. If necessary, click the Fill & Border tab.
- You will notice that all four bars are surrounded by a faint yellow border. We want to change the fill of the Caffeine bars from a self-colour to a pattern. Single-click, say, the Caffeine bar in the Male part of the figure. You will see that this bar retains its yellow border; whereas the remaining bars lose theirs. This means that any editing requested will apply to that bar only.

TIP
Single-click one of the selected bars to select that bar alone

http://www.psypress.com/spss-made-simple/
In **Properties**, click the drop-down menu labelled **Pattern** and make a selection from the array (Figure 7). Click **Apply** to produce the pattern in the selected bar only.

Single-click the other Caffeine bar to select that bar only and repeat the operation. Figure 7 shows the point where the selection has been made, but the **Apply** button has yet to be clicked and the target bar remains self-coloured.

Click on the tab labelled **Bar Options** and adjust the width of the bars as shown in Figure 8.

Click **Apply** to implement the change and **Close** to leave **Properties**.

Back in the Chart Editor, choose **File** ➔ **Close** to leave the Chart Editor. The figure should now appear as in Output 3.

---

**Figure 6. The Chart Editor**

http://www.psypress.com/spss-made-simple/
Figure 7. Selecting a filler pattern for the Caffeine bar in the Female column of the panel.

1. Click to choose a pattern for the bar.
2. Click to run.

Pattern inserted by previous operation.

Chosen pattern will appear in selected bar.

Click to leave properties.
Figure 8. Adjusting the widths of the bars

**Editing captions and titles on figures**

When there is a figure in the Chart Editor, it is possible to edit titles, subtitles footnotes or captions by double-clicking on the area of the figure concerned to produce a selection frame, then single-clicking to produce a red cursor. This is crucial – double-clicking will not produce the cursor. See Figure 9.
Changing the orientation of a label on the vertical axis of a graph

The initial orientation of the caption on the vertical axis of the figure was vertical: the text read in a bottom-up direction (Figure 9). To change the orientation of the caption, so that the text will read from right to left, proceed as shown in Figure 10.

5.3 ERROR BAR CHARTS

An alternative to a bar graph is an Error Bar chart, in which the mean of the scores in a particular category is represented by a single point. The spread is represented by a vertical line (T-bar or whiskers) passing through the point. The user can choose, as a measure of spread, the confidence interval on the mean, multiples of the standard deviation or multiples of the standard error of the mean. Output 4 is a clustered error bar chart summarising the results of the drug experiment.

http://www.psypress.com/spss-made-simple/
The production of an error bar chart with the Chart Builder is analogous to the production of bar graphs and raises no new issues.

Output 4. A clustered error bar chart with Experimental Condition as the category variable and Gender as the cluster variable

The symbols used for the means and the form of the lines used for the error bars can be changed by double-clicking anywhere within the graphic to open the Chart Editor. Double-clicking on the appropriate symbol or line in the Gender key will open the corresponding Properties dialog box where changes can be made.

You will notice that in Output 4, there are no lines linking the error bars. This is entirely appropriate, since the bars represent qualitatively distinct categories. In other circumstances, however, as when the categories are ordered, it may be desirable to join up the points (when there are more than two) with interpolation lines. This is easily achieved in the Chart Editor by clicking the means to highlight them, selecting the Elements drop-down menu and clicking Interpolation line (or alternatively clicking the icon at right).
5.4 BOXPLOTS

Three types of boxplots are available in Chart Builder, the single boxplot (called 1-D Boxplot in the gallery), the simple boxplot for plotting the boxplots across categories of a grouping variable and the clustered boxplot for plotting boxplots across categories of two grouping variables.

Output 5 shows the boxplots of Score for the Placebo and Caffeine groups, clustered by Sex. Notice in the output that there is one case identified as an outlier with ‘o’. Any extreme case would have been identified with an asterisk (*).

The production of boxplots with the Chart Builder proceeds in a manner analogous to the drawing of histograms and bar charts and presents no new issues.

Output 5. Boxplots of scores of the Placebo and Caffeine groups clustered by Sex

5.5 PIE CHARTS

The pie chart is an alternative to a bar graph, which provides a picturesque display of the frequency distribution of a qualitative variable. It is useful for displaying the relative frequencies of observations in the same set of categories over time or for bringing out the varying compositions of two things, such as conservative versus risky investment portfolios.
Pie charts can be panelled in a similar way to bar charts and histograms. Output 6 is a pie chart showing the percentages of the different blood groups in the sample we studied in Chapter 4. (The data are in Ch4 Height, Weight, Sex & Blood group.sav at http://www.psypress.com/spss-made-simple.)

Output 6. A 3-D Pie Chart showing the distribution of Blood Group with the Group B sector 'exploded' for greater salience

The pie chart in Output 6 has been edited in the Chart Editor to impart a three-dimensional appearance and 'explode' the Group B sector.

5.6 LINE GRAPHS

Like a scatterplot, a line graph depicts the relationship between two continuous or scale variables, such as weight and height. In a line graph, as in a histogram, the entire range of a one variable (say height) is stepped out in equal intervals along the horizontal axis. Above the midpoint of each interval, in the body of the graph, a point is placed with height on the y-axis proportional to the mean weight of all cases with heights falling within the interval on the horizontal axis. Finally, adjacent points are joined by straight lines.

Line graphs can be drawn with just one line or more than one line in the graph; and like bar charts and pie charts, they can also be panelled.

In this section, we shall use the Chart Builder to draw line graphs depicting the relationship between weight and height in the men and women in one of the data sets we explored in Chapter 4. To do this, we must first divide the total range of height into equal intervals, a task for which we need a special procedure known as visual binning.
5.6.1 The Visual Binning procedure

In order to draw the line graphs, we must first divide the entire range of the variable that is going to be on the horizontal axis of the graph (height) into equal intervals and specify a representative value for each interval. This is done automatically in SPSS’s Histogram procedure. In the histogram, the intervals are known as class intervals. Elsewhere, however, class intervals are known as bins and we shall need to use a special procedure known as binning to divide the total range of height into bins, with fixed bin width.

We shall use SPSS’s Visual Binning procedure to divide the total range of the men’s heights into intervals or bins, the largest and smallest of which are open-ended, so that all scores are included. Proceed as follows.

- Choose Transform ➔ Visual Binning to open the initial Visual Binning dialog box and transfer the variable Height to the Variables to bin box on the right. (See Figure 11.)

![Visual Binning dialog box](http://www.psypress.com/spss-made-simple/)

**Figure 11. The initial Visual Binning dialog box with the target variable transferred**

Click the Continue button to enter the main Visual Binning dialog box, which shows a histogram of the distribution of Height, tells us that 2000 cases have been scanned and gives the minimum and maximum values of Height in our data set as 156.28cm and 205.09cm, respectively (See Figure 12).

We shall want about 12 class intervals or bins. The cutpoints will appear in the histogram when the bins are specified. We are going to enter, as Values in the Value column, the upper limits of the intervals. If we want 12 bins, we shall need to specify only 11 cutpoints.
Chapter 5

Figure 12. The main Visual Binning dialog box

Click the Make Cutpoints button. The dialog is shown in Figure 13.

Figure 13. The Make Cutpoints dialog. The third slot must be clicked before the Apply button will become active.

http://www.psyspress.com/spss-made-simple/
On returning to the main Visual Binning dialog box, click the Make Labels Button. A column of labels now appears in the spaces under Label. Double-click on these to shorten them (Figure 18). Click the OK button to run the procedure.

![Visual Binning dialog box](http://www.psypress.com/spss-made-simple/)  

**Figure 14.** The main Visual Binning dialog box again, after clicking Make Labels and editing the labels

By choosing View ➔ Value labels, you can see, in Data View, the bin intervals shown in the Label column of the Visual Binning dialog. These intervals also appear in the Values column of Variable View for the variable Hght. With Value labels inactive, you would see only the ordinal numbers of the intervals. If, while in Variable View, you look in the Measure column, you will see that the binned version of Height has been automatically entered as an ordinal variable.

### 5.6.2 Plotting line graphs

We have used the Visual Binning procedure to divide the total range of heights of the participants into 11 bins. This binned version of the Height variable has been stored as the ordinal variable Hght. (Hght is fine for a variable name, provided the technical term bin does not appear in the variable label. You will need to edit the variable label in Variable View: a label such as ‘Height in Centimeters (binned)’ should be avoided.) We are going to plot line graphs on this new variable.
graphs of weight against height for the males and females and present them in a panelled display for comparison. Proceed as follows:

- Open the **Chart Builder**, select **Line** from the **Choose from** list, click the **Simple Line** picture in the gallery and drag it into the **Chart preview** box.
- Click the **Groups/Point ID** tab and choose **Columns panel variable**.
- Transfer the three variables to their correct drop zones, as shown in Figure 15.

![Figure 15. Chart preview with the variables and error bars specified](http://www.psypress.com/spss-made-simple/)

- In **Element Properties**, specify that the statistic for the Y-axis is the mean. Order error bars as well.
- Click **OK** to run the procedure.

The panelled line graph is shown in **Output 7**. **Output 7** has been edited by double-clicking to enter the Chart Editor and specify markers.

http://www.psypress.com/spss-made-simple/
5.7 USING CHART TEMPLATES

Many or all of the attributes of an edited chart can be saved for future use, in what is known as a chart template. Should it be necessary to produce a similar chart on future occasions, time can be saved by opening the template, which will automatically incorporate the final attributes of the edited chart in the new chart that appears in the Viewer. Such attributes include colour and shading and width and spacing of the bars, as well as headings, subheadings and footnotes.

Figure 16 shows the edited version of a chart in the Chart Editor. (The right-hand margin has been dragged to the right to gain control of the position of the original footnote about the error bars.) Realising that we may need to produce a similar chart in the future, we can store the edited version in a chart template as follows.

- Choose File→Save Chart Template… to open the Save Chart Template dialog box (Figure 17).

- Complete the Save Chart Template dialog as shown. On clicking Continue, you enter the Save Template dialog box. Save the template somewhere in your own file space as shown in Figure 18. Leave the Chart Editor.
Open the **Chart Builder** and order a clustered bar chart as before.

Click the **Options** button to enter the **Options** dialog box, part of which is shown in Figure 19.

Click the **Add** button, locate the folder in which you saved the file and open the file to see the file name appear in the **Template Files** box of the **Options** dialog box, below the **Add** button.

Click **OK** to return to the **Chart Builder**. *At this stage you will see no change in the Chart preview.*

In the **Chart Builder**, click **OK** to produce the chart  (Output 8).

It is clear from Output 8, which appeared immediately in the Viewer, that the chart template has preserved all the features of the final edited version of the chart.
More on graphs and charts

1. Check All settings.

2. Make a brief note.

3. Click.

Figure 17. Completing the **Save Chart Template** dialog box

Figure 18. The **Save Template** dialog box
Chart templates are obviously very useful tools. There are, however, some editorial changes that can be made in the Chart Editor, but which would not be preserved in a chart template. In a simple bar chart, for example, you might want to use different shades or patterns for different bars. Such a change would not be preserved in a template. A similar problem would be
encountered should you have used different patterns for selected categories of the x-axis variable separated by panel variable category. It can be done in the Chart Editor, but the change would not be preserved in the chart template. In this context, a clustered bar chart should be seen as two or more simple bar charts, one imposed upon the other: the bars in a cluster can be shaded differently, but are homogeneous within each of the simple component charts. Clustered bar charts, therefore, even when edited, are suitable for storage in a template. This requirement of pattern homogeneity within a simple chart unit, however, is not met when the categories of an X-variable have different patterns or colours across categories of a panel variable.

5.8 A FINAL WORD

In this chapter, we have tried to focus on some aspects of the editing of charts with which users often have difficulty. We have also looked more closely at the Chart Builder and drawn some new displays, such as line graphs.

When you are producing, say, a bar chart, you have the option of creating special effects, such as giving the graph a three-dimensional appearance. The fanciest graphs, however, do not necessarily present the clearest picture of the results of an investigation. Three-dimensional effects, for example, though aesthetically attractive to some, require careful handling; otherwise, they may actually obscure your presentation. The addition of error bars to such a figure creates the impression of tower blocks with a forest of radio masts on their roofs. It is sometimes difficult to discern, in such a cluttered display, the most important features of your results.

It is possible, having spent time editing a chart, to save many or all of its features in a chart template, which can be used to reproduce the final version instantly, whenever this is required.

Exercises

Exercise 6 Charts and graphs, and Exercise 7 Recoding data; selecting cases; line graph are available in www.psypress.com/spss-made-simple and click on Exercises.