

Quantitative data analysis

Business



**ASSIGN
BUSTER**

Practitioner Research Techniques Assessment.

Part 2; Quantitative Analysis By Stephenson J Sylvester, University of Huddersfield. July 2006 This paper is a discussion in support of the Doctorate of Education Programme Unit 3 Assignment part 2 on Quantitative Analysis. The discussion looks at the usage and value of quantitative data analysis packages like SPSS. A brief comparison is drawn between other methods of data analysis and a summary conclusion is made. A step by step procedure is adopted in answering some key points on the use of SPSS for Windows.

Where indicated, the data and information from SPSS version 12.

0. 1 has been inserted directly into the discussion. One of the first questions that surface is “ What does SPSS stand for? ” Quite simply, it means “ Statistical Product and Service Solutions. ” It is an amalgam of several software packages designed and developed over many years to analyse statistical data and make predictions based on probability. These packages have been around for about 38 years and it was formerly called “ Statistical Package for the Social Sciences”. An in-depth history and makeup of SPSS can be found online at spss.

om. 1 It is very useful in social and/or educational research, where there is a need to handle a large amount of statistical data generated in the investigation through the use of quantitative analysis and questionnaires. Indeed, where data is gathered from a very large number of respondents by the use of questionnaires, it can be very difficult to arrange the presentation of the data for the readers benefit. It can also be difficult to present

supporting analysis in the form of evidence for any conclusions the researcher may draw or infer from the statistics.

SPSS for Windows is a tool, one of many, which aid this process.

There are other software packages like NCSS, WINKS, Statsoft, Stata, etc, but by far the most popular and widely used by social researchers in quantitative analysis is SPSS. To set the scene, we have a research project to determine the relationship between sixth form students and their desire to go to university. The comparison is made across both genders and an assumption is made that there is no difference between the genders. This is called the “null hypothesis”.

In its simplest form we measure the difference between what we expect the respondents to say and what the respondents actually say.

If the difference is small, we put this down to chance and accept the basic assumption of there being no difference. If the difference large we reject the null hypothesis and 1 SPSS Inc, (2006), About SPSS Inc, [online], Available at:

. Accessed – 3rd July 2006 1 Chi-Square Analysis: Sample Questionnaire

Items Please place a tick in the appropriate box to indicate your response. 1.

What is your gender? Male ? [1] [2] Female ? .

Do you wish to go to university after the Sixth Form? Yes No ? ? [1] [2]

Picture i. Picture i. shows a sample questionnaire used during this assignment. There are two questions which each have two possible answers.

The questions are numbered 1 and 2 respectively. In the first question, the subject of gender is addressed.

The respondent has the option of answering male or female to the following question. “ What is your gender? ” The subject of desire is addressed using the following question. “ Do you wish to go to university after the Sixth Form? Here the respondent has the option of “ yes” or “ no” to indicate their desires on university study after Sixth Form College. Each question has been given an identification that will be used in SPSS later. For question 1 the identification is “ gend”.

For question 2 the identification is “ uni”. In each question, the respondent has an option of two answers. For the purpose of this exercise, these are coded with two unique numbers. In the first question “ 1” is used for 2 Male and “ 2” is used for Female. In the second question “ 1” is used for a “ Yes” response and “ 2” is used for a “ No” response.

These codes are to be used as data references in SPSS so the actual numbers used are insignificant. There is a need for consistency throughout the exercise so the numbers used throughout the analysis must be consistent. The responses to the questionnaire are listed in the table below.

Data Input Printout gend 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
45 1. 00 1. 00 2.

00 1. 00 2. 00 1. 00 2. 00 2.

00 2. 00 1. 00 1. 00 2. 00 2. 00 2.

00 1. 00 2. 00 1. 00 1. 00 1.

00 2. 00 1. 00 1. 00 1. 00 2.

00 2. 00 1. 0 2. 00 1. 00 2. 00 2.

00 1. 00 1. 00 2. 00 1. 00 2. 00 1.

00 2. 00 2. 00 1. 00 1. 00 2.

00 2. 00 1. 00 1. 00 1. 00 uni 1.

00 2. 00 1. 00 2. 00 2. 00 1. 00 1.

00 1. 00 1. 00 1. 00 1. 00 1. 00 2.

00 1. 00 2. 00 1. 00 1. 00 2.

00 1. 00 1. 00 2. 00 1. 00 1.

00 2. 00 1. 00 2. 00 1. 00 1.

00 1. 00 1. 00 1. 00 2. 00 2. 00 2.

00 1. 00 2. 00 1. 00 2. 00 2.

00 2. 00 1. 00 1. 00 2. 00 2. 00 2.

00 3 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 2. 00 2. 00 1. 00 2. 00 1.

00 1. 00 1. 00 1. 00 1. 00 2. 00 2.

00 2. 00 2. 00 2. 00 2. 00 1.

00 1. 00 1. 00 1. 00 1. 00 1.

00 1. 00 1. 00 2. 00 1. 00 1. 0 1.

00 1. 00 1. 00 1. 00 Table i. In table i, we have the resulting response from 60 respondents tabulated for processing in SPSS. The table has been exported directly from SPSS into a word document.

The first column which has no header shows a list of 1 to 60. These represent 60 different responses from 60 different respondents to each question.

Column 2 has a header “gend”. This represents answers to the first question. Finally, column 3 had a header of “uni”. This represents answers to the second question.

Foster (2002)² stated that this type of measurement is on a Nominal Scale.

There are 3 other types of scales, Interval, Ratio and Ordinal (or Rank) scales. These are discussed later. Each row represents the response received from a given respondent to the questions. So, if we looked at row number 10 for example we see that the respondent answered male to the first question and yes to the second question as indicated by “1” and “1” in the following cells on row 10. The codes assigned have no significance to the respondent.

In many cases, the codes can be added after the respondent has completed the form as the extra data may confuse the respondent.

It is possible that the respondent may feel the higher or lower value represented by the code might indicate a correct or incorrect answer. Gillham (2004)³ argues that the inclusion of coding data can add confusion to the respondent and will inevitably take up valuable space on the questionnaire. With Ordinal scales it is not possible to quantify the actual difference between the scales but the information is conveyed as to the

position of a given individual data in relation to the others of that scale. A good example is that of a Grand Prix Formula 1 motor race result.

At the British round in Silverstone (ITV-F1, 2006)⁴ the racing driver Alonso won by some 14 seconds.

In second place was racing driver Schumacher. Racing driver Raikkonen was some 5 seconds behind Schumacher in third place. The in the finish ranking scale the difference in relation to each other's positions is 1, 2 and 3 respectively. The ranking or ordinal scale gives no indication of the magnitude of the differences between each value. For each of the four cells: Square the discrepancy and divide this by the expected frequency for that cell. 2.

3. Add together the results for the four cells. The resulting number is the Chi-Square statistic.