**SPSS: Statistical Package for the Social Sciences**

**Development of SPSS**

This chapter aims to outline some ways in which one of the most widely used statistical packages, the Statistical Package for the Social Sciences (SPSS), can be used to input, manipulate, and statistically analyse data in the behavioural and social sciences. SPSS was initially developed by Norman Nie, Dale Bent and Hadlai Hull in 1965 in the Political Sciences Department of Stanford University. In 1968 it became a commercial organisation when the software was first publicly released. Its development continued in 1970 at the National Opinion Research Center (NORC) of the University of Chicago. In 2006 Nie was awarded Exceptionally Distinguished Achievement from the American Association for Public Opinion Research for his substantive as well as methodological contribution to public opinion research.

New SPSS versions have regularly appeared since 1968. Apart from SPSSX (or SPSS-X) which came out in 1983, these releases were given Arabic numerals. The Roman X in this case referred to version 10. At that time, SPSS was usually referred to as SPSSX1. After X, the numbering restarted at 1. The most recent version at the time of writing was 29, which was released on 12 September 2022 (https://community.ibm.com/community/user/ai-datascience/blogs/kennia-garcia/2022/08/08/whats-new-in-spss-statistics-29). Table 1 shows the year of some developments in SPSS together with that of some of its major alternatives.

-------------------------

Table 1 about here

-------------------------

While some of these SPSS releases are clearly major modifications, such as the introduction of the Windows version in 1993, most have involved what appears to be minor changes for many users. In 2009, SPSS was bought by the International Business Machines (IBM) Corporation. Because of a dispute about the ownership of its name, the package was called the *P*redictive *A*nalytic *S*oft*w*are (PASW) in 2009-2010. Since then, the software has officially been referred to as IBM SPSS Statistics although many people still simply call it SPSS.

The first book on SPSS was published in 1970 by Nie et al., followed by a second edition in 1975. At that time, computer instructions and data were punched onto 80-column 12-row computer cards. These were fed in batches into a card reader linked to a large, mainframe computer. Subsequently, running SPSS was made easier by typing computer instructions (as commands or syntax) directly into the computer. Ease was further enhanced by the development of small personal computers (PCs) using a graphical user interface (GUI). This enabled options to be selected from drop-down menus and dialog boxes (Norušis, 1986).

Commands can still usefully be written in a Syntax window, as will be described later.

**Comparing Statistical Packages**

As indicated in Table 1, several other statistical packages were produced around 1965 and later. Among them were:

* Bio-Medical Data (BMD) Program (BMDP)2 by Wilfrid Dixon in 1961 at the University of California, Los Angeles (UCLA) (Dixon, 1975; Flournoy, 2010);
* Statistical Analysis System (SAS) at North Carolina State University in 1966-1976 (Helwig & Council, 1979);
* Minitab in 1972 at Penn State University (Ryan et al., 1976); and
* SYSTAT in 1980-1983 by Leland Wilkinson (1984) at the University of Illinois in Chicago.

A few authors have compared some of these statistical packages, either directly or indirectly. An initial direct comparison was that of Tabachnik and Fidell (1982) who looked at SPSS, SAS and SYSTAT in their handling of multivariate statistics. They did not evaluate these packages. Their (and Ullman’s) seventh and most recent edition of their book was published in 2021. Some authors have implicitly compared packages by essentially writing the same book but using different statistical software such as SPSS and Minitab. I did this with Bryman (1990, 1996) and on my own (1994a, 1997, 1998). There was also no attempt to appraise the software. As the sales and citations for the SPSS books were much higher than those for Minitab, no further editions of the Minitab books were produced. Field et al. (2012) has written an R version3 of his 2009 SPSS book, enabling these two packages to be indirectly compared.

I used a small set of data of fewer than four variables to directly check the statistical algorithms or formulae used by SPSS (1983) in essentially what was the same book (1994a, 1998). The results were basically the same. When realising that my hand calculation of the median was not the same as that of SPSS in my 1998 book, I used the SPSS algorithm so I could obtain the same value (pp. 21-23). It was also through hand calculations (pp. 30-31) that I realised that SPSS produced the sample or estimated population, and not the population, variance and standard deviation (p. 56).4 My colleague Dennis Howitt and myself have compared the results of SPSS and hand computations less directly in the various editions of one book using manual calculations (e.g. 1997a, 2024) and editions of a companion book using SPSS (e.g. 1997b, 2017). We have also done this in a shorter, combined edition directed at social science, rather than psychology, students (Norris et al., 2012). We have computed these calculations with no more than three variables since the calculations with more variables typically involve matrix algebra. This was not done as it would have required greater abstraction and complexity as well as a lengthier exposition.

Among later papers comparing the results of SPSS with some other statistical software on particular statistical aspects is that of Peng and So (2002). They looked at all the packages listed in Table 1, apart from R, on a logistic regression of 752 married, white, working women from the US 1975 Panel Study of Income Dynamics. They recommended Minitab and Stata for beginning researchers and BMDP and SAS for experienced ones. Their objection to SPSS and SAS was that their goodness-of-fit and diagnostic statistics were calculated from individual observations and so should not be interpreted as chi-square values.

In another comparison, Keeling and Pavur (2007) examined five of the packages shown in Table 1, including R, on the accuracy of the number of correct digits provided by the log relative error for various statistics including univariate summary statistics, one-way analysis of variance and linear regression. They did this on *St*atistical *R*eference *D*ata (StRD) sets available at the American *N*ational *I*nstitute of *S*tandards and *T*echnology (<https://www.itl.nist.gov/div898/strd/>). The very large number of results presented in this study are difficult to summarise and the authors did not attempt to do so. Improvements to the packages are continually being made and so may not apply to their current versions. Furthermore, in view of the relatively crude nature of many social and psychological measures, it is not immediately obvious what degree of accuracy is needed for most behavioural statistical analyses.

The latest comparison study seems to be that of Hodges et al. (2022) who investigated very recent versions of four of the packages presented in Table 1 on various nonparametric tests. Using the 2018 Race Implicit Association Test data as well as simulated data, they found that the most inconsistent results were due to algorithmic variations in calculating Pearson’s chi-square text for 2 x 2 tables. The two-tailed *p*-values provided were the same for SPSS and SAS across large, medium, and small samples. Algorithmic differences generally resulted in small test statistic or *p*-value discrepancies and so may not substantially impact statistical conclusions. They recommended that at least two statistical packages should be used for analyses and that these should be named.

Muenchen (2013) has periodically tried to gauge the popularity of up to 12 statistical packages in various ways several times a year on his website (http://r4stats.com), the last time being in 2022. While popularity of a package may not be related to its soundness, this information may be useful in highlighting which packages require further consideration. The ideal way of estimating popularity would be a representative survey. The best method that Muenchen used seems to have been the number of research papers specifying the statistical package used. He searched the literature with Google Scholar, a search engine which itself is not without its difficulties (Jacsó, 2008). In terms of this index, the most popular package was SPSS Statistics followed by R (https://r4stats.com/articles/popularity/#google\_vignette).

In an earlier questionnaire survey of the Chairs of 51 Canadian University Psychology Departments, Alder and Vollick (2000) noted that, of the 59 per cent responding, 97 per cent of the departments reported having a mandatory statistics course. Statistical software was included in 79 per cent of the mandatory courses and 95 per cent of the non-mandatory courses. Over 70 percent of these courses taught SPSS, followed by SYSTAT in 13-22 per cent of them.

Enlyft (2023) provided information about the sales and market share of companies using Analytics software such as SPSS (<https://enlyft.com/tech/analytics>). Although Matlab and Mathworks are not statistical packages, they had the largest share of the market at 23 and 22 per cent respectively, followed by SPSS at 19 per cent. The number of companies which had the SPSS base edition installed was 54,070. The next most frequently installed statistical packages included SAS/STAT (4,714 or .017 per cent) and Stata (1,016 or .004 per cent). The three most common sectors having an SPSS base edition installed were Higher Education (4,342 or 8 per cent), Information Technology and Services (4,123 or 8 per cent) and Hospital and Health Care (2,990 or 6 per cent). Most of these companies were in the US (27,437 or 51 per cent), followed by the UK (3,717 or 7 per cent) and India (3,068 or 6 per cent).

The SPSS base edition includes many of the most widely used statistical procedures such as Compare Means, Correlate, Regression, Data Reduction and Non-Parametric Tests (<https://www.sv-europe.com/ibm-spss-statistics/spss-statistics-base/>). For comparison, the Advanced Statistics or Standard edition includes such statistics as General Linear Modelling Multivariate Analysis and Repeated Analysis, General Loglinear Analysis and Logit Loglinear Analysis (https://www.ibm.com/docs/en/spss-statistics/saas?topic=statistics-advanced). Separate SPSS manuals have been published to support these modules (e.g. Norušis, 1993, 1994a, 1994b). Many higher education institutions subscribe to the full SPSS package which may also incorporate the structural equation modelling software AMOS (*A*nalysis of *Mo*ment *S*tructures). AMOS was acquired by IBM in what seems to be around 1997 (Arbuckle, 1997; Bacon et al., 1997). AMOS is now known officially as IBM SPSS Amos (Arbuckle, 2016).

**Entering or Inputting Raw Data**

Ideally, data to be analysed should be entered directly by a participant into an SPSS data file, or a txt or Unicode file. The latter two file types can be readily converted into an SPSS data file. Participants answering questions online should stop them making such mistakes as inadvertently choosing two answers to the same question while not seeming to answer an adjoining question. For the researcher, this input method saves time and mistakes if the data is to be entered manually. Researchers, with little experience of entering data or using SPSS, should preferably familiarise themselves with SPSS before collecting the kind of data they expect to obtain. The potential size of an SPSS data file is very large with an ASCII (*A*merican *S*tandard *C*ode for *I*nformation *I*nterchange) record length of up to 32,767 characters which can be further extended (https://www.ibm.com/support/pages/maximum-record-length-read-ibm-spss-statistics). The number of cases will be limited by the size of the computer rather than the capacity of the software (https://www.ibm.com/support/pages/maximum-number-cases-spss-dataset).

It is advisable to give the SPSS data file a unique name before constructing it in case naming it is inadvertently not done, and the process needs to be restarted. The data file name can have up to 64 characters, ending with the suffix “.sav”. The data file itself has a very large number of rows and columns. It can be viewed in terms of the nature of the variables (“Variable View” as shown in Figure1) or the values given to the data (“Data View” as displayed in Figure 2). Which view is being displayed is indicated, and can be selected, by the two options on the bottom left of the data window.

These two views are illustrated with part of a data set, which among other variables, assessed two of the relationship attitudes that Rogers (1959) proposed led to relationship satisfaction (Cramer, 2003a). These two attitudes were assessed by the six highest loading regard and empathy items from a factor analysis (Cramer, 1986) of the 64-item revised Barrett-Lennard (1964) Relationship Inventory completed by 150 participants describing their closest romantic relationship. These items, presented in Table 2, were slightly modified so that three were worded negatively and three positively. They were answered on a 6-point scale from “totally untrue” (1) to “totally true” (6). Higher scores indicated higher regard and empathy.

------------------------------------------------

Table 2, Figures 1 and 2 about here

------------------------------------------------

In the SPSS data file, variables are best given unique short “Names” of up to eight characters. These names should provide as close an idea as possible as to what the variable is. There is also an opportunity in the fifth column of the “Variable View” window to give a longer label of up to 256 characters. This may include a partial or full wording of variables such as survey questions. Including more detailed labels is useful when interpreting the factor structure and alpha reliability of variables such as questionnaire items.

Data may be missing for various reasons (e.g. Bodner, 2006; Rousseau et al., 2012). Specifying these reasons may be informative. Giving a specific value for missing data is preferable to leaving the data cell blank which SPSS indicates with a full stop. The reason for specifying a missing value is that an empty cell may result from a valid value being inadvertently skipped by the person entering the data. Traditionally, SPSS has reserved three numerical values to indicate missing data. As SPSS was originally designed to analyse survey data, these three values were often illustrated as 8, 9 or 97 for “don’t know”, 98 for “refusing to answer”, and -1, 0, 9 or 99 for “not applicable”. It is important that whatever values are chosen to represent missing value cannot also reflect a valid value.5 For the data in Figure 2, zero denotes a missing value. The “1” in the “filter\_$” column indicates which cases have been selected. These are the participants who have reported they are describing their closest romantic relationship.

Variables can be thought of as comprising two main types. One type is category or nominal variables with more than two categories. Numbers here are primarily used to denote different categories such as country or nationality. The other type is continuous scores. These scores are arranged in numerical order from lower to higher values for such variables as age, reaction time or multiple point rating scales. A binomial or dichotomous variable with only two categories can be treated as either a category or a score with two values. For example, relationship type may be coded as 1 for a romantic one and 2 as a friendship. A higher score on this variable would indicate greater friendship and a lower score greater romantic involvement.

**Double Data Entry**

Entering a large set of data can be time-consuming and error-prone (e.g. Barchard & Pace, 2011; Barchard & Verenikina, 2013). In the 1950s, IBM developed an input system where one operator would enter the data on an IBM (1964) 26 printing card punch machine and another operator would re-enter the same data on an IBM (1962) 56 Card Verifier machine. Ideally, having two such “inputters” is better as they may be less likely to make the same mistakes. If a second inputter cannot be recruited, then the same person should input the data twice, preferably on two occasions separated by some time such as a week.

In 2012, SPSS introduced the Compare Datasets procedure to inspect values for the same variables and cases in two separate files. To demonstrate this procedure, I have called the files “Input1” and “Input2”. For simplicity, all six values in “Input1” are correct in my example. There are three mistakes in “Input2”. However, errors are likely to occur in both files and these need to be corrected in both files. When there are no more mistakes, one of the files can be used as the one to analyse.

This procedure would usually be run using menus and dialog boxes. So “Compare Datasets” would be selected from the “Data” dropdown window as displayed in Figure 3.

---------------------------

Figure 3 about here

--------------------------

This option opens up the “Compare to” dialog box as presented in Figure 4.

---------------------------

Figure 4 about here

--------------------------

The “Browse…” option in the dialog box is used to choose the other file to be compared (i.e. “Input2”) as displayed in Figure 5.

---------------------------

Figure 5 about here

--------------------------

“Open” is selected to choose the variables to be matched as depicted in Figure 6.

---------------------------

Figure 6 about here

--------------------------

The last table of output in the Output window shows the discrepant values and their row (and not id) numbers as displayed in Figure 7.

---------------------------

Figure 7 about here

--------------------------

Note that for “s3” in row 1, there is no discrepant value. For the first participant, 40, there are two mistakes. The first error is to put two values (54) in the second column and, consequently, none in the third column which is the second mistake. The third mistake is for the second participant, 54, where the value for “s1” is wrongly inputted as “3” rather than “1”.

These steps may be summarised with the following notation where ➔ denotes select (Bryman & Cramer, 2011; Cramer, 1998):

➔**Data** ➔**Data** **Compare Datasets**… [as shown in Figure 3 and which opens

up **Compare to** dialog box as depicted in Figure 4]

**➔Browse…** [opens up **Compare Datasets: Read File** dialog box as

displayed in Figure 5]

**➔**comparison file [e.g. **Input2.sav**]

**➔Open** [returns to **Compare Datasets: Read File** dialog box as shown in Figure 5]

➔**Continue** [which closes **Compare Datasets: Read File** dialog box and opens up **Compare Datasets** as presented in Figure 6]

**➜comparison** variable/s [e.g**. s1** to **s3**] in **Matched fields:** box ➔➥ button ➔**Fields to Compare:** box

**➜**identifying case variable/s [e.g. **id**] in **Matched fields:** box ➔➥button ➔ **Case IDs:** box

**➜OK** [the last table of output is displayed in Figure 7]

This procedure may also be conducted with syntax files. Syntax files have the advantage of keeping a concise record of what was done and of being exactly repeatable if needed.

DATA LIST LIST

/ID S1 S2 S3.

BEGIN DATA.

40 5 4 5

54 1 1 3

73 4 5 4

END DATA.

DATASET NAME INPUT1.

DATA LIST LIST

/ID S1 S2 S3.

BEGIN DATA.

40 54 5

54 3 1 3

73 4 5 4

END DATA.

DATASET NAME INPUT2.

DATASET ACTIVATE INPUT1.

COMPARE DATASETS

/COMPDATASET INPUT2

/VARIABLES ALL.

Note that each complete command ends in a full stop. A continuation of a command on two or more lines is indicated by a forward slash. These commands are then carried out using the “Run” option.

Entering the data for a large file within the syntax file would be unwieldly. In such situations, it is better to use a “get file” command as shown below.

GET

FILE='C:\USERS\DUNCAN CRAMER\DOCUMENTS\BLRI\2000\INPUT1.SAV'.

DATASET NAME INPUT1 WINDOW=FRONT.

GET

FILE='C:\USERS\DUNCAN CRAMER\DOCUMENTS\BLRI\2000\INPUT2.SAV'.

DATASET NAME INPUT2 WINDOW=FRONT.

DATASET ACTIVATE INPUT1.

COMPARE DATASETS

/COMPDATASET INPUT2

/VARIABLES ALL

The SPSS keyword “FRONT” is used to keep that file active. The output for these commands is the same as that for the previous syntax.

Although prior to 2012, SPSS had software for comparing two data files (*SPSS Data Collection Data Entry*), it was expensive and thus not readily available. So, for myself and my students, I put together some SPSS commands for making this comparison and included a description of the procedure in the fifth and sixth edition of the SPSS book (Howitt & Cramer, 2011, 2017). Although this procedure may make the underlying process more explicit, it is easier to use the above syntax.

**Entering or Inputting Aggregated Data**

Conducting some SPSS statistical analyses is possible on data which has already been aggregated or summarised in the form of frequency tables, means, standard deviations and correlations. Frequency tables can be created by weighting categories for the following statistics as shown, for example, in Howitt and Cramer (2017): pie diagrams, bar charts and histograms (Chapter 4); compound bar charts and histograms (Chapter 9); chi-square, Fisher’s exact test and one-sample chi-square (Chapter 16); McNemar’s test (Chapter 17); and log-linear analysis (Chapter 37). Means, standard deviations and a correlation matrix can be used for multiple regression (Chapter 48). SPSS syntax files using “compute” statements can also be written to carry out various statistics on summarised data. For example, because I needed to compare a series of correlations with *z*, *T2* and *Z2\** tests, I wrote SPSS syntax files for conducting these (Cramer, 1994a, 1998).

**Grouping Related Variables**

A correlation matrix, if available, can also inputted to conduct a factor analysis as illustrated with the following syntax file.

MATRIX DATA VARIABLES R1 E1 R2 E2 R3 E3 R4 E4 R5 E5 R6 E6

/N =150

/CONTENTS = CORR.

BEGIN DATA.

1

.463 1

-.575 -.444 1

.331 .665 -.394 1

-.448 -.355 .541 -.223 1

-.347 -.537 .314 -.410 .341 1

**.673** .492 -.625 .453 -.568 -.387 1

-.435 -.537 .449 -.560 .408 .563 -.565 1

.496 .280 -.523 .275 -.476 -.433 .556 -.382 1

-.179 -.406 .183 -.439 **.024** .220 -.145 .397 -.083 1

-.451 -.371 .453 -.353 .455 .357 -.457 .491 -.337 .415 1

.243 .518 -.308 .602 -.152 -.395 .300 -.471 .200 -.507 -.396 1

END DATA.

FACTOR matrix = in(COR=\*)

/MISSING LISTWISE

/ANALYSIS r1 e1 r2 e2 r3 e3 r4 e4 r5 e5 r6 e6

/PRINT UNIVARIATE INITIAL ROTATION

/FORMAT SORT

/CRITERIA MINEIGEN(1) ITERATE(250)

/EXTRACTION PC

/CRITERIA ITERATE(250)

/ROTATION VARIMAX

/METHOD=CORRELATION.

The data was taken from Cramer (2003a).

The Windows procedure for this factor analysis may be notated as follows:

➔**Analyze** ➔**Dimension Reduction** ➔**Factor...** [opens **Factor Analysis**

dialog box shown in Figure 8]

➔**r1 e1 r2 e2 r3 e3 r4 e4 r5 e5 r6 e6** ➔➥ button [puts **r1** to **e6** under

**Variables:**]

➔**Descriptives...** [opens **Factor Analysis: Descriptives** dialog box]

➔**Univariate statistics** [for the *N*s and standard deviations]

➔**Initial solution**

➔**Coefficients** [for correlation coefficients]

➔**Continue** [closes **Factor Analysis: Descriptives** dialog box]

➔**Extraction...** [opens **Factor Analysis: Extraction** dialog box]

**Principal components** is the default option

➔**Scree plot** ➔**Number of factors:** and type number of factors if

different from default option of Kaiser’s criterion greater than 1

➔**Maximum Iterations for Convergence:** and type number if default of **25**

not wanted [e.g. **250**] ➔**Continue** [closes **Factor Analysis: Extraction** dialog

box]

➔**Rotation...** [opens **Factor Analysis: Rotation** dialog box] ➔**Varimax**

➔**Continue** [closes **Factor Analysis: Rotation** dialog box]

➔**Options...** [opens **Factor Analysis: Options** dialog box

➔**Sorted by size** [highest loadings shown first]

➔**Continue** [closes **Factor Analysis: Options** dialog box]

➔**OK**

--------------------------

Figure 8 about here

---------------------------

Although it is not easy to discern by visual inspection how a large number of variables may group together as being related, the highest and lowest correlated variables can be easily picked out. The highest correlation (.673) is between the two regard items of “R4 They really value me” and “R1 They are truly interested in me”. The lowest correlation (.024) is between the regard item “R3 They don't care for me” and the empathy item “E5 They don't always appreciate exactly how the things I experience feel to me”. The way variables are grouped together is conducted statistically by a “dimension reduction” procedure such as principal components analysis.

This analysis revealed two components which each had a greater variance than a single item (i.e. an eigenvalue of more than 1). The 12 items are ordered in terms of the size of their larger loadings on the two factors in columns 2 and 3 of Table 2. Five regard items loaded or correlated substantially higher on the first than the second factor. Five empathy items did this on the second factor. The analysis produces a correlation matrix for the 12 items, initially shown to three decimal places although six may be obtained by double clicking on them. Very similar results emerged from a principal components analysis of this correlation matrix as displayed in columns 4 and 5 of Table 2. Correlations have traditionally been reported to two decimal places in research publications. For two decimal places, the findings are essentially the same as the two previous ones, as presented in columns 6, 7 and 8 of Table 2.

**Recoding Variables**

As in the case of social measures such as regard and empathy, several items are normally used to capture the different ways in which these constructs can be expressed. Factor analysis is initially applied to determine how similar the variables are. Where possible, it is usually recommended that half the items on a scale should be reversed to control for response set, such as the tendency to agree or disagree with items regardless of their content. If the reverse items work as intended, their scoring needs to be reversed before being combined with the other items to make up the total score for the multi-item measure. Ideally, the direction of scoring should reflect the name of the measure. So, it appears more logical for higher, rather than lower, scores to represent a variable called “regard”.

The syntax for the scoring of the “regard” items to be reversed is:

DATASET ACTIVATE DataSet10.

RECODE r2 r3 r6 (1=6) (2=5) (3=4) (4=3) (5=2) (6=1) (MISSING=SYSMIS) INTO r2n r3n n r6n.

EXECUTE.

The Windows procedure for recoding values may be described as follows:

➔**Transform** ➔**Recode into Different Variables…** [opens **Recode into**

**Different Variables** dialog box shown in Figure 9]

➔**r2** ➔➥ button [puts **r2 i**n **Numeric Variable -> Output Variable:** box]

➔Type **r2n** in **Name:** box ➔**Change** [**r2 -> ?** in **Numeric Variable -> Output**

**Variable:** box changes to **r2 -> r2n**]

➔Repeat for **r3** and **r6**

➔**Old and New Values…** [opens **Recode into Different Variables: Old and**

**New Values** dialog box]

➔Type **1** in **Value:** box under **Old Value** and **6** in **Value:** box under

**New Value**

➔**Add** [puts **1 -> 6** in **Old -> New:** box]

➔Change **2 -> 5, 3 -> 4, 4 -> 3, 5 -> 2** and **6 -> 1**

➔**System- or user-missing** under **Old Value** and **System-missing** under

**New Value**

➔**Continue** ➔**OK**

---------------------------

Figure 9 about here

---------------------------

This recoding can be checked with the following syntax on, say, the first six participants:

SUMMARIZE

/TABLES=r2 r2n r3 r3n r6 r6n

/FORMAT=VALIDLIST NOCASENUM TOTAL LIMIT=6

/TITLE='Case Summaries'

/MISSING=VARIABLE

/CELLS=COUNT.

The Windows procedure for checking the accuracy of the data manipulation of recoding values may be summarised as follows:

➔**Analyze** ➔**Reports** ➔**Case Summaries...** [opens **Summarize Cases**

dialog box as shown in Figure 10]

➔select variables [e.g. **r2**] ➔➥ button ➔**Limit cases to first** and type

number [e.g. **6**]

➔**OK**

----------------------------

Figure 10 about here

----------------------------

Figure 11 shows the output table containing the values of the original and recoded values to enable to check whether they have been appropriately recoded.

----------------------------

Figure 11 about here

----------------------------

**Alpha Reliability Measurement**

The variables making up a measure should be internally consistent which means they should be highly related to one another. The most common statistic of internal consistency is Cronbach’s (1960) alpha reliability. This is usually described as the mean correlation of all possible split-half reliabilities of a measure. The way it can be calculated is as a single factor repeated measures analysis of variance (ANOVA) in which the error variance is subtracted from the between-subjects variance and the result divided by the between-subjects variance. In other words, it is the proportion of error variance to 1 minus the between-subjects variance. The bigger the error variance, the smaller the alpha or proportion is (Cramer, 1994a, 1998).

As with the factor analysis, alpha is computed for the raw data on the six-item regard scale, followed by the syntax commands using the correlation matrix, means, standard deviations and *N*s. Alpha is .85 as shown in the output table in Figure 12.

----------------------------

Figure 12 about here

---------------------------

Alpha can be calculated from the ANOVA results in Figure 13 as 1 – 0.488/3.305.

----------------------------

Figure 13 about here

---------------------------

Nunnally and Bernstein (1994) have suggested a research measure should have a minimum reliability of .70 which this scale meets.

Where the reliability is below .70, this level or higher may be achieved by dropping one or more items. This information is obtained from the Item-Total Statistics table depicted in Figure 14. In this case, alpha is not substantially increased if the last item r6n is deleted so using all six items is preferable.

----------------------------

Figure 14 about here

----------------------------

MATRIX DATA VARIABLES=ROWTYPE\_ r1 r2n r3n r4 r5 r6n.

BEGIN DATA.

MEAN 5.221 5.214 5.552 5.026 5.195 4.968

STD 0.786 1.022 0.908 0.921 0.950 1.228

N 205 205 205 205 205 205

CORR 1

CORR .575 1

CORR .451 .541 1

CORR .669 .625 .569 1

CORR .493 .522 .481 .562 1

CORR .448 .453 .456 .457 .341 1

END DATA.

RELIABILITY VARIABLES = r1 r2n r3n r4 r5 r6n

/MATRIX = IN(\*)

/MODEL=ALPHA

/SCALE('ALL VARIABLES') ALL

/STATISTICS ANOVA

/SUMMARY=TOTAL.

The Windows procedure for computing the alpha reliability of a measure may be encapsulated as follows:

➔**Analyze** ➔**Scale** ➔**Reliability Analysis...** [opens **Reliability Analysis**

dialog box shown in Figure 15]

➔**r1**, **r2n**, r**3n**, **r4**, **r5** and **r6n** ➔➥ button [puts **r1**, **r2n**, r**3n**, **r4**, **r5** and **r6n**

in the **Items:** box]

➔**Model:** ➔**Alpha i**s the default option

➔**Statistics** ➔**Scale if item deleted** ➔**F test**

➔**OK**

----------------------------

Figure 15 about here

----------------------------

**Combining Variables**

The scores of individual items making up a scale such as “regard” need to be combined to form an overall scale score. The factor structure and alpha reliability of potential items a sample should always be checked as these can vary depending on the composition of the sample. The overall score could be a mean or total one, depending on how such variables are usually combined. In addition, it is necessary to decide how to deal with participants who have a relatively small amount of missing data. These kinds of issues are often not described in research papers. For a six-item scale such as “regard” it may be reasonable to permit the score of one item to be missing. If this is to be done for computing a total score, the mean score has to be calculated first before multiplying by the number of valid scores. This is to ensure that all participants can attain the same maximum score (e.g. Bryman & Cramer, 2011, pp. 60-61; Howitt & Cramer, 2017, Chapter 37). The syntax for computing a mean score based on 5 or 6 items is:

DATASET ACTIVATE DataSet7.

COMPUTE regard5.6=mean.5(r1, r2n, r3n, r4, r5, r6n).

EXECUTE.

The Windows procedure for conducting this computation may be reported as follows:

➔**Transform** ➔**Compute Variable…** [opens up **Compute Variable** dialog

box shown in Figure 16]

➔Type name of new variable (e.g. **regard5.6**) in box under **Target Variable:**

➔Type **MEAN.5()** in box under **Numeric Expression:** or ➔**All** in box under

**Function group:** ➔**Mean** in box under **Functions and Special Variables:**

➔⮭ button [which puts **MEAN(?,?)** in box under **Numeric Expression:]**

➔**r1** ➔➥ button to put **r1** after **(** in box under **Numeric Expression:]**

➔for legibility, type **comma** and **space** and repeat for remaining variables

[e.g. **r2n, r3n, r4, r5, r6n**]

➔**OK**

----------------------------

Figure 16 about here

----------------------------

This computation can be checked with the following syntax on, say, the first four participants.

SUMMARIZE

/TABLES=r1 r2n r3n r4 r5 r6n regard5.6

/FORMAT=VALIDLIST NOCASENUM TOTAL LIMIT=4

/TITLE='Case Summaries'

/MISSING=VARIABLE

/CELLS=COUNT.

Figure 17 shows the values of the six items to calculate the mean score based on more than four valid values.

----------------------------

Figure 17 about here

----------------------------

**Distribution Shape of Variables**

It is useful to know and report the shape of the distribution such as its number of peaks, its skewness or asymmetry and its kurtosis or tailedness (Cramer, 1994a, 1998). For example, a bi-modal rather than a uni-modal distribution of relationship satisfaction implies that this measure tends to categorise satisfaction into two groups. Number of peaks may be broadly classified as uni-, bi- or multi-modal. This judgement may be subjective as illustrated in Figure 20. Are there three peaks or one in this distribution? The apparent number may vary with the size of the categories into which the scores are grouped. Skewness can be positive, symmetrical, or negative. As skewness is calculated by subtracting the individual scores from the mean, positive skewness is indicated by a predominance of higher scores on the left of the distribution. Negative skewness is reflected by a majority of higher scores being on the right of the distribution as shown in Figure 20. Kurtosis can be positive or leptokurtic “thin curve”). It can also be or negative or platykurtic (“flat curve”). Between these two shapes is a third form called mesokurtic (“middle curve”) which has values close to zero and is characteristic of a normal distribution. These distribution statistics can only be computed on raw data. This analysis can be done with the following syntax for regard5.6.

FREQUENCIES VARIABLES=regard5.6

/STATISTICS=SKEWNESS SESKEW KURTOSIS SEKURT

/HISTOGRAM NORMAL

/ORDER=ANALYSIS.

The Windows procedure may be written as follows:

➔**Analyze** ➔**Descriptive Statistics** ➔**Frequencies...** [opens **Frequencies**

dialog box shown in Figure 18]

➔**regard5.6** ➔➥ button [puts in box under **Variable(s):**]

➔**Statistics…** [opens **Frequencies: Statistics** dialog box]

➔**Skewness** and **Kurtosis** below **Distribution**

➔**Continue** [closes **Frequencies: Statistics** dialog box]

➔**Charts…** [opens **Frequencies: Charts** dialog box]

➔**Histograms** ➔**Show normal curve on histogram**

➔**Continue** [closes **Frequencies: Charts** dialog box]

➔**OK**

----------------------------

Figure 18 about here

----------------------------

Skewness and kurtosis together with their standard errors (Std. Error) are displayed in Figure 19 on the full sample of 256 participants. SPSS does not report the statistical significance of these statistics. This significance is calculated by dividing the statistic by its standard error which gives a value comparable to *z.* *Z* is 6.01 (-.913/.152) for skewness and 3.26 (.988/.303) for kurtosis. These *z* values can be looked up in a standard normal distribution table or computed with a web app (e.g. https://www.socscistatistics.com/pvalues/normaldistribution.aspx). A *z* value of 6.01 or more is statistically significant at the two-tailed .00001 or less probability level. One of 3.26 is statistically significant at the two-tailed .002 or less probability level. Regard scores are significantly negatively skewed with most scores to the right of the distribution. The mean empathy scores are also significantly negatively skewed (2.61 = -.397/.152) at the two-tailed .005 or less probability level.

----------------------------------------

Figures 19 and 20 about here

----------------------------------------

**Normalising Asymmetrical Scores and Correlations**

It has been argued that the statistical significance of the more commonly used parametric statistics, such as Pearson’s product-moment correlation, are based on the scores being normally distributed (Siegel, 1957). It is possible to transform scores so that they are more normally distributed. Where scores are not normally distributed such as in the case of the regard and empathy scores, it may be worth checking whether their normalisation affects their size and/or their statistical significance. Tabachnik and Fidell (2019) suggested that negatively skewed scores can be made more normal by taking the square root of subtracting each score from 1 added to the highest score. The highest score for both regard and empathy was 6.00. This computation can be conducted with a “Compute Variable” SPSS procedure. Doing this makes skewness for regard less statistically significant (2.96 = .450/.152, *p* = .002). Kurtosis becomes non-significant for regard (0.71 = -.214/.303, *p* = .239) and empathy (0.62 = -.187/.303, *p* = .239). Figure 21 shows an SPSS histogram for the transformed regard scores.

----------------------------

Figure 21 about here

----------------------------

These transformations hardly affected the size or the statistical significance of the Pearson’s correlation between the two variables. This was .549 (*p* < .001) for the untransformed variables and .551 (*p* < .001) for the transformed ones. The non-parametric Spearman correlation between the two untransformed variables was similar at .546 (*p* < .001).

The syntax commands for these correlations are:

DATASET ACTIVATE DataSet1.

CORRELATIONS

/VARIABLES=regard5.6 empathy5.6 regard5.6sqrt1 empathy5.6sqrt1

/PRINT=TWOTAIL NOSIG LOWER

/MISSING=PAIRWISE.

NONPAR CORR

/VARIABLES=regard5.6 empathy5.6 regard5.6sqrt1 empathy5.6sqrt1

/PRINT=SPEARMAN TWOTAIL NOSIG LOWER

/MISSING=PAIRWISE.

The names “regard5.6sqrt1” and “empathy5.6sqrt1” refer to the transformed variables where the square root (sqrt) of each score is subtracted from 1 added to the highest score.

The Windows procedure is:

➔**Analyze** ➔**Correlate** ➔**Bivariate...** [opens **Bivariate Correlations**

dialog box shown in Figure 22]

➔**regard5.6** ➔➥button ➔**empathy5.6** ➔➥ button ➔➥**regard5.6sqrt1**

➔➥ button **empathy5.6sqrt1** [puts **regard5.6**, **empathy5.6**, **regard5.6sqrt1**

and **empathy5.6sqrt1** in **Variables:** box] ➔**Spearman** ➔**Show only the**

**lower triangle**

➔**OK**

----------------------------

Figure 22 about here

----------------------------

Figure 23 shows the Pearson’s correlation table first followed by the Spearman’s correlations.

----------------------------

Figure 23 about here

----------------------------

**Linearity and Homoscedasticity Assumptions**

Two further assumptions of tests of correlation and regression are that the relationship between variables is linear and that the scatter of the joint values on two variables is evenly distributed around the straight line of best fit drawn between them. Even distribution is known as homoscedasticity (“same dispersion”) and uneven distribution (“different dispersion”) as heteroscedasticity. Determining these two assumptions may be done visually by drawing a scatterplot or scatter diagram. This can be conducted with the following syntax command for the two variables of “regard5.6” and “empathy5.6”:

DATASET ACTIVATE DataSet1.

GRAPH

/SCATTERPLOT(BIVAR)=regard5.6 WITH empathy5.6

/MISSING=LISTWISE.

The resulting graph can be edited to fit a straight line that runs closest to all the points on the plot as shown in Figure 24 and as outlined below in the description of the Windows procedure.

----------------------------

Figure 24 about here

----------------------------

The Windows procedure for producing and editing this chart is as follows:

➔**Graphs** ➔**Legacy Dialogs** ➔**Scatter/Dot…** ➔**Simple Scatter** [opens **Scatter/Dot** dialog box] ➔**Define** [closes **Scatter/Dot** dialog box and opens **Simple Scatterplot** dialog box shown in Figure 25]

➔one variable (e.g. **empathy5.6**)➔➥ button [to put in box under **Y Axis:**]

➔other variable (e.g. **regard5.6**)➔➥ button [to put in box under **X Axis:**] ➔**OK**

*double-click* anywhere in the scatterplot [opens **Chart Editor**]

➔**Elements** ➔**Fit Line at Total** [opens **Properties** dialogbox] ➔**x** at top right

corner of **Properties** dialog boxor **Close** to close]

----------------------------

Figure 25 about here

----------------------------

As can be seen in Figure 25, the line runs from the lower left to the upper right indicating a positive linear relationship with higher scores on “regard5.6” going with higher scores on “empathy5.6”. The scatter does not appear even with scores on the left showing a degree of heteroscedasticity. This would be reduced by removing these values. Deleting, or preferably selecting out the more extreme “regard5.6” scores using “Select Cases…” under “Data”, reduces the Pearson’s correlation. For example, removing the 10 cases with scores of 3.50 or less lowers it from .549 to .504. Omitting the 35 cases with the lowest value of 4.25 or less further decreases it to .395. It is unclear to what extent this smaller correlation is due to reduced variance rather than to lower heteroscedasticity.

There are various statistical tests for assessing heteroscedasticity in correlation and regression (e.g. Lyon & Tsai, 1996). These include Breusch and Pagan’s (1979) and Cook and Weisberg’s (1983) score test, White’s (1980) test and Wooldridge’s (2009) modification, Koenker’s (1981) test, and Spearman’s rank correlation test (Yin & Carroll, 1990). The last statistic correlates the standardised predicted residuals with the standardised absolute residuals. These values can be obtained from the “Save” option (Figure 26) in the “Linear Regression” dialog box produced by selecting “Analyze”, “Regression” and then “Linear…” on the drop-down menu. The rank correlation is not significant for either the full sample (σ = .007, *p* = .916, *n* = 256) or the reduced sample (σ = -.255, *p* < .001, *n* = 221), indicating that the homoscedasticity assumption was met in the full sample. The first three tests can be selected in the “Univariate: Options” dialog box (Figure 27) obtained by choosing “Analyze”, “General Linear Model”, and “Univariate…” on the drop-down menu and then “Options…” in the “Univariate” dialog box.

**Structural Equation Modelling**

Structural equation modelling (SEM) distinguishes latent variables (or factors) from manifest ones (or indicators). For example, the indicators of “regard” are the individual items that measure it while the factor is the unobserved variable that underlies these indicators. A variable in a structural model can be a total score such as “regard6.5”, the total score and its alpha reliability, or effectively its factor score (Cramer, 2003b). The structural model is the relationship between the main variables which should be more than two. Structural equation modelling has the advantages of testing more than one criterion or outcome variable, controlling for measurement error, and comparing the statistical fit of models to each other and to the original data. The absence of a statistical difference between two models means that they are equivalent.

These models may be referred to and treated as if they are “causal” models, even if the data is cross-sectional rather than longitudinal. The minimum requirement of a causal model is that the cause should temporally precede an effect so the design should be prospective. I have published several two-variable two-and more wave models to determine the temporal relationship between the two variables (Cramer, 1990, 1994b, 1995,1996, 1996; Cramer & Takens, 1992; Cramer et al., 1996). A potential problem with this type of analysis is that the sign of cross-lagged correlations may be reversed and its size larger than one because of high synchronous and test-retest correlations (Cramer, 2003c). The maximum size of a correlation is set at +1.

Joreskog and van Thillo (1972) developed the first software to conduct SEM. They called this software LISREL, which stands for *L*inear *S*tructural *Rel*ationship Analysis. This software enables latent variables based on alpha reliability to be tested. In 1997 IBM bought an alternative SEM package called AMOS and incorporated it in SPSS.

Byrne (2001) has compared these two programs, together with EQS6, in her 1993 study of the validity of the Maslach Burnout Inventory in 580 teachers. She did not say which package she would advocate or for what kind of analysis. She has also written similar books on LISREL (1998) and AMOS (2016) which used data from this study published in 1993 and 1994 on burnout in teachers. Narayanan (2012) compared eight SEM packages which included LISREL, AMOS and EQS. She listed the particular strengths of the software but did not further elaborate their advantages. LISREL, AMOS and EQS were reported as all having a graphical interface. However, R had introduced such an interface before the publication of her paper in 2012 despite her referencing the Fox paper which discussed this interface (Fox, 2005; https://semlj.github.io/gui.html).

**Summary**

The statistical package SPSS was first released by Nie et al. at Stanford University in 1968. Since then, new versions have been regularly issued, with the latest, at the time of writing, released in 2022 and numbered 29. While some of these updates have been major modifications, users may have hardly noticed the differences in their use of many of them. SPSS has grown to become by far the major dominant statistical software in academia. The accuracy of its data handling and statistical procedures on small, limited data sets has been repeatedly confirmed. Double data entry by preferably two independent people is important to ensure that data entry is accurate. The SPSS “Compare Datasets” procedure has greatly facilitated this process.

SPSS can analyse some forms of aggregated data such as frequency tables and correlation matrices. SPSS “compute” commands may be used on other summarised statistics to conduct further tests unavailable on SPSS, such as *z*, *T2* and *Z2\** tests, or for aggregated data, such as related and unrelated *t*-tests. Syntax instructions provide a concise way of recording what procedures have been conducted and enable those procedures to be exactly replicated if necessary. In the late 1990s SPSS offered the structural equation modelling software AMOS. This provides further important statistical procedures. These statistics are being increasingly used in published research and deserve to be better known to students and researchers, despite their apparent complexity.

**References**

AAPOR Award (2006). Public Opinion Quarterly, *70*(3), 423.

Alder, A. G., & Vollick, D. (2000). Undergraduate statistics in psychology: A survey of Canadian institutions. *Canadian Psychology/Psychologie canadienne*, *41*(3), 149-151.

Arbuckle, J. L. (1989). AMOS: Analysis of moment structures. *American Statistician*, *43*, 66-67.

Arbuckle, J. (1997). *Amos users’ guide Version 3.6.* Smallwaters Corporation.

Arbuckle, J. L. (2016). *IBM SPSS Amos user’s guide*. IBM Corp.

Bacon, L. D. et al. (1997). Using Amos for structural equation modelling in market research. SPSS White Paper. https://www.bauer.uh.edu/jhess/documents/3.pdf

Barchard, K. A., & Pace, L. A. (2011). Preventing human error: The impact of data entry methods on data accuracy and statistical results. *Computers in Human Behavior*, *27*(5), 1834-1839.

Barchard, K. A., & Verenikina, Y. (2013). Improving data accuracy: Selecting the best data checking technique. *Computers in Human Behavior*, *29*(5), 1917-1922.

Barrett-Lennard, G. T. (1964). *The Relationship Inventory: Forms OS-M-64, OS-F-64 and MO-M-64 plus MO-F-64*. Unpublished manuscript, University of New England, Australia.

Bodner, T. E. (2006). Missing data: Prevalence and reporting practices. Psychological Reports, 99(3), 675–680.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, *47*(5),1287-1294.

Bryman, A., & Cramer, D. (1990). *Quantitative data analysis for social scientists*. Routledge.

Bryman, A., & Cramer, D. (1996). *Quantitative data analysis with Minitab: A guide for social scientists*. Routledge.

Bryman, A. & Cramer, D. (2011). *Quantitative data analysis with IBM SPSS Statistics 17, 18 and 19: A guide for social scientists*. Routledge.

Byrne, B. M. (1993). The Maslach Burnout Inventory: Testing for factorial validity and invariance across elementary, intermediate and secondary teachers. *Journal of Occupational and Organizational Psychology*, *66*(3), 197-212.

Byrne, B. M. (1994). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Educational Research Journal*, *31*(3), 645-673.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Lawrence Erlbaum Associates.

Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing*, *1*(1), 55-86.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed). Lawrence Erlbaum Associates.

Cook, R. D., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, *70*(1), 1-10.

Cramer, D. (1986). An item factor analysis of the revised Barrett-Lennard Relationship Inventory. *British Journal of Guidance and Counselling*, *14*(3), 314-325.

Cramer, D. (1990). Self-esteem and close relationships: A statistical refinement. *British Journal of Social Psychology*, *29*(2), 189-191.

Cramer, D. (1994a). *Introducing statistics for social research: Step-by-step calculations and computer techniques using SPSS*. Routledge.

Cramer, D. (1994b). Self-esteem and Rogers' core conditions in close friends: A latent variable path analysis of panel data. *Counselling Psychology Quarterly*, *7*(3), 327-337.

Cramer, D. (1995). Life and job satisfaction: A two-wave panel study. *Journal of Psychology*, *129*(3), 261-267.

Cramer, D. (1996). Job satisfaction and organizational continuance commitment: A two-wave panel study. *Journal of Organizational Behavior*, *17*(4), 389-400.

Cramer, D., Henderson, S., & Scott, R. (1996). Mental health and adequacy of social support: A four-wave panel study. *British Journal of Social Psychology*, *35*(2), 285-295.

Cramer, D. (1997). *Basic statistics for social research: Step-by-step calculations and computer techniques using Minitab*. Routledge.

Cramer, D. (1998). *Fundamental statistics for social research: Step-by-step calculations and computer techniques using SPSS for Windows*. London: Routledge.

Cramer, D. (2003a). Facilitativeness, conflict, demand for approval, self-esteem and satisfaction with romantic relationships. *Journal of Psychology*, *137*(1), 85-98.

Cramer, D. (2003b). *Advanced quantitative data analysis*. Open University Press.

Cramer, D. (2003c)**.** A cautionary tale of two statistics: Partial correlation and standardised partial regression. *Journal of Psychology, 137*(5), 507-511.

Cramer, D., & Takens, R. J. (1992). Therapeutic relationship and progress in the first six sessions of individual psychotherapy: A panel study. *Counselling Psychology Quarterly*, *5*(1), 25-36.

Cronbach, L. J. (1960). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Dixon, W. J. (1975). *BMDP: Biomedical computer programs*. University of California Press.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Flournoy, N. (2010). Wilfrid Joseph Dixon, 1915-2008. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *173*(2), 455-457.

Fox, J. (2005). The R commander: a basic-statistics graphical user interface to R. *Journal of Statistical Software*, *14*(9), 1-42.

Helwig, J. T., & Council, K. A. (Eds.) (1979). *SAS user’s guide*. SAS Institute Inc.

Hodges, C. B., Stone, B. M., Johnson, P. K., Carter III, J. H., Sawyers, C. K., Roby, P. R., & Lindsey, H. M. (2022). Researcher degrees of freedom in statistical software contribute to unreliable results: A comparison of nonparametric analyses conducted in SPSS, SAS, Stata, and R. *Behavior Research Methods*, 1-25.

Howitt, D., & Cramer, D. (1997a). *An introduction to statistics in psychology: A complete guide for students* (1st ed.). Hemel Hempstead: Prentice Hall/Harvester Wheatsheaf.

Howitt, D., & Cramer, D. (1997b). *A guide to computing statistics with SPSS for Windows* (1st ed.). Prentice Hall/Harvester Wheatsheaf.

Howitt, D. & Cramer, D. (2011). *Introduction to SPSS Statistics in psychology: For version 19 and earlier* (5th ed.). Harlow: Prentice Hall.

Howitt, D. & Cramer, D. (2014). *Introduction to SPSS in psychology: For version 22 and earlier* (6th ed.). Harlow: Prentice Hall.

Howitt, D., & Cramer, D. (2017). *Introduction to* *SPSS in psychology* (7th ed.). Prentice Hall.

Howitt, D., & Cramer, D. (2024). *Understanding statistics in psychology* *with SPSS* (9th ed.). Prentice Hall.

IBM (1962). *Reference Manual IBM 56 Card Verifier*. IBM Corp.

IBM (1964). *Reference Manual IBM 24 Card Punch, IBM 26 Printing Card Punch*. IBM Corp.

IBM Corp. (2022). *IBM SPSS Statistics for Windows, Version 29.0*. IBM Corp.

Jacsó, P. (2008). Google Scholar revisited. *Online Information Review*, *32*(1), 102-114.

Joreskog, K. G., & Van Thillo, M. (1972). LISREL: A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. ETS.

Lyon, J. D., & Tsai, C. L. (1996). A comparison of tests for heteroscedasticity. *Journal of the Royal Statistical Society: Series D: The Statistician*, *45*(3), 337-349.

Hull, C. H., & Nie, N. H. (1981). *SPSS updates 7-9: New procedures and facilities for releases 7-9*. McGraw-Hill.

Keeling, K. B., & Pavur, R. J. (2007). A comparative study of the reliability of nine statistical software packages. *Computational Statistics & Data Analysis*, *51*(8), 3811-3831.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics*, *17*(1), 107-112.

Muenchen, R. A. (2013). *The popularity of data analysis software*.URL http://r4stats. com/popularity.

Narayanan, A. (2012). A review of eight software packages for structural equation modeling. *American Statistician*, *66*(2), 129-138.

Nie, N. H. (1983). *SPSS-X user’s guide* (1st ed.). McGraw-Hill.

Nie, N. H., Bent, D. H., & Hull, C. H. (1970). *SPSS: Statistical package for the social sciences* (1st ed.). McGraw-Hill.

Nie, N. H., Hull, C. H., Jenkins, J. G. Steinbrenner, K., & Bent, D. H. (1975). *Statistical package for the social sciences* (2nd ed.). McGraw-Hill.

Norris, G., Qureshi, F., Howitt, D., & Cramer, D. (2012). *Introduction to* *statistics with SPSS for social scientists*. Pearson.

Norušis, M. J. (1986). *SPSS/PC+ for the IBM PC/XT/AT*. SPSS Inc.

Norušis, M. J. (1988). *SPSS/PC+V2.0 base manual*. SPSS Inc.

Norušis, M. J. (1993). *SPSS: SPSS for Windows base system user’s guide Release 6.0*. SPSS Inc.

Norušis, M. J. (1994a). *SPSS advanced statistics 6.1*. SPSS Inc.

Norušis, M. J. (1994b). *SPSS professional statistics 6.1*. SPSS Inc.

Nunnally, J. M., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Peng, C. Y. J., & So, T. S. H. (2002). Modeling strategies in logistic regression with SAS, SPSS, Systat, BMDP, Minitab, and STATA. *Journal of Modern Applied Statistical Methods*, *1*(1), 147-156.

Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 3, pp. 184-256). McGraw-Hill.

Rousseau, M., Simon, M., Bertrand, R., & Hachey, K. (2012). Reporting missing data: A study of selected articles published from 2003–2007. Quality & Quantity: International Journal of Methodology, 46(5), 1393–1406.

Ryan, T. A., Joiner, B. L., & Ryan. B. F. (1976). *Minitab student handbook*. Duxbury Press.

Sachdev, A. (2009, September 27). What’s in a name? Recognition, plenty of money. *Chicago Tribune*. <https://www.chicagotribune.com/news/ct-xpm-2009-09-27-0909250481-story.html>

Siegel, S. (1957). Nonparametric statistics. *American Statistician*, *11*(3), 13-19.

SPSS Inc. (1983). *SPSSX statistical algorithms*. SPSS Inc.

Tabachnik, B. G., & Fidell, L. S. (1982). *Using multivariate statistics* (1st ed.). Longmans Higher Education.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2021). *Using multivariate statistics* (7th ed.). Pearson.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, *48*(4), 817-838.

Wilkinson, L. (1984). *SYSTAT, Version 1*. SYSTAT Inc.

Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach* (4th ed.). Cengage Learning.

Yin, Y., & Carroll, R. J. (1990). A diagnostic for heteroscedasticity based on the Spearman rank correlation. *Statistics & Probability Letters*, *10*(1), 69-76.

**Footnotes**

1I began using SPSS-X in 1983 and started working on a book suggested by and with Bryman in 1988 when there seemed to be few user-friendly SPSS texts.

2I used a punch card version of BMDP for my PhD in the early 1970s at the Institute of Psychiatry in London.

3R seems to refer to the initial letter of the first name of Ross Ihaka and Richard Gentleman who developed it based on a computer language simply called S.

4 Nie et al. (1975) did state that the sample variance was being calculated (p. 184). Of course, it is simple to convert one into the other by hand. Moreover, with larger samples, the difference between these two statistics becomes small (Cramer, 1998, p. 28).

5Ironically the US General Social Survey, produced by the organisation NORC that further developed SPSS initially, is not going to produce an SPSS data file of the survey from 2022 onwards because it only has three missing values options (https://gss.norc.org/get-the-data/spss#:~:text=Starting%20with%20the%202020%20GSS,to%20implement%20consistent%20missing%20values.).

6EQS most probably stands for *Eq*uation*s.*

Table 1

*Timeline comparison of some major developments in SPSS and its alternatives*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Year | SPSS | BMDP | SAS | Minitab | SYSTAT | Stata | R |
| 1961 |  | Began in Preventive Medicine Department, UCLA |  |  |  |  |  |
|  |  | First manual |  |  |  |  |  |
| 1965 | Began in Political Sciences Department,  Stanford University |  |  |  |  |  |  |
| 1966 |  |  | Began in Agricultural Department, North Carolina State University |  |  |  |  |
| 1970 | Moved to NORC, University of Chicago |  |  |  |  |  |  |
|  | First manual |  |  |  |  |  |  |
| 1972 |  |  |  | Began in Statistics Department, Penn State University |  |  | R Commander |
| 1973 |  | First manual |  |  |  |  |  |
| 1980 |  |  |  |  | Began in Psychology Department, University of Illinois in Chicago |  |  |
| 1982 | Tabachnik & Fidell multivariate analysis procedures |  | Tabachnik & Fidell multivariate analysis procedures |  | Tabachnik & Fidell multivariate analysis procedures |  |  |
| 1983 | SPSS-X (10) |  |  |  |  |  |  |
| 1984 |  |  |  |  |  | Began in Computer Resource Centre, Santa Monica, CA |  |
| 1985 |  |  |  |  |  | Stata 1.0 |  |
| 1986 | SPSS/PC+V1.0 |  |  |  |  |  |  |
| 1990 |  |  |  |  |  | Statistics with Stata |  |
| 1993 | SPSS Windows 6.0 |  |  |  |  | Moved to College Station, Texas  StataCorp |  |
| 1994 |  | BMDP 1.0 for Windows |  |  |  |  |  |
| 1996 |  |  |  |  |  |  | Began Statistics Department, Auckland University |
| 1997 | SPSS Amos 3.6 |  |  |  |  |  |  |
| 2009 | Sold to IBM |  |  |  |  |  | R Commander 2.9.0 |
| 2012 | Compare Datasets |  |  |  |  |  |  |
| 2017 |  |  |  |  | SYSTAT 13.2 |  |  |
| 2020 |  |  | SAS 9.4/ STAT 15.2 |  |  |  |  |
| 2022 | SPSS 29.0 |  |  |  |  |  |  |
| 2023 |  |  |  | Minitab 21.4 |  | Stata 18.0 | R-4.3.1 for Windows |

Table 2

*Two Varimax rotated factors of regard and empathy 6-item subscales from raw data, and three and two decimal correlation matrices*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Items from the Cramer Shortened Barrett-Lennard Relationship Inventory | Raw data | | Correlation matrix | | | | |
|  |  |  | 3 decimals | | 2 decimals | | |
|  | F1 | F2 | F1 | F2 | Items | F1 | F2 |
| *R4* They really value me | .812 | .264 | .812 | .263 | *R4* | .814 | .265 |
| *R3* They don't care for me | -.783 | -.049 | -.783 | -.049 | *R3* | -.783 | -.050 |
| *R2* They don't feel deep affection for me | -.755 | -.241 | -.755 | -.240 | *R2* | -.756 | -.237 |
| R5 They feel a true liking for me | .751 | .088 | .751 | .089 | R1 | .752 | .216 |
| R1 They are truly interested in me | .750 | .217 | .750 | .217 | R5 | .749 | .091 |
| R6 I do not feel appreciated by them | -.497 | -.458 | -.497 | -.458 | R5 | -.498 | -.459 |
| *E6* They usually sense or realise what I am feeling | .104 | .811 | .104 | .811 | *E6* | .103 | .811 |
| E2 They usually understand the whole of what I mean | .254 | .776 | .254 | .776 | E5 | .067 | -.776 |
| E5 They don't always appreciate exactly how the things I experience feel to me | .066 | -.774 | .066 | -.774 | E2 | .253 | .775 |
| *E1* They nearly always know exactly what I mean | .386 | .700 | .386 | .699 | *E1* | .384 | .703 |
| *E4* They don't understand me | -.489 | -.621 | -.489 | -.621 | *E4* | -.491 | -.620 |
| *E3* They don't realise what I mean, particularly when I have difficulty in saying it | -.435 | -.494 | -.435 | -.495 | *E3* | -.432 | -.498 |

Figure 1

*SPSS “Variable View” window”*

A screenshot of a computer

Description automatically generated

Figure 2

*SPSS “Data View” window”*

A screenshot of a computer

Description automatically generated

Figure 3

*SPSS “Data” drop-down menu showing “Compare Datasets…”*

A screenshot of a computer

Description automatically generated

Figure 4

*SPSS “Compare to” dialog box*

A screenshot of a computer

Description automatically generated

Figure 5

*Select file to be compared (“Input2”) into the SPSS “Compare Datasets” dialog box*

A screenshot of a computer

Description automatically generated

Figure 6

*“Compare Datasets” dialog box with variables to be matched, selected*

A screenshot of a computer

Description automatically generated

Figure 7

*Last table of output for “Compare Datasets” procedure showing discrepant values and their row number*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Case By Case Comparison** | | | | | |
| Row | | id | s1 | s2 | s3 |
| Active | Compare |
| 61 | 1 |  | (1) 5.00  (2) 54.00 | (1) 4.00  (2) 5.00 | (1) 5.00  (2) . |
| 2 | 2 |  | (1) 1.00  (2) 3.00 |  |  |
| (1) is the Active Dataset and (2) is the Comparison Dataset | | | | | |

Figure 8

*“Factor Analysis” initial dialog box*

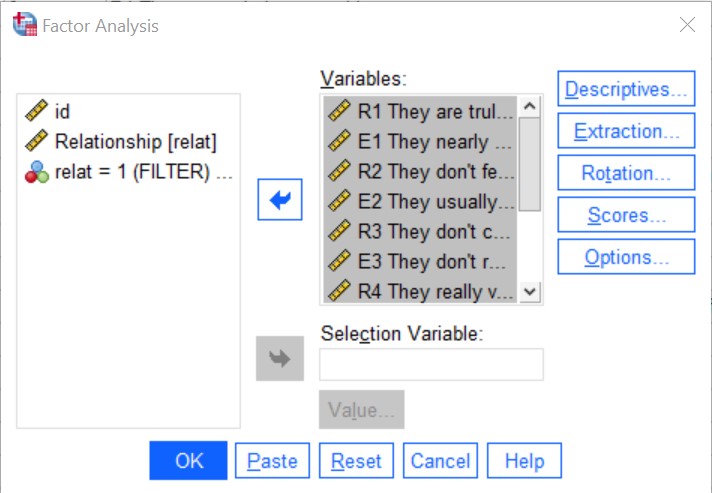


Figure 9

*“Recode into Different Variables” initial dialog box*

A screenshot of a computer

Description automatically generated

Figure 10

*“Summarize Cases” dialog box*

A screenshot of a computer

Description automatically generated

Figure 11

*Output table of “Case Summaries” showing original and recoded values of three variables*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Case Summariesa** | | | | | | | |
|  | | R2 They don't feel deep affection for me | r2n | R3 They don't care for me | r3n | R6 I do not feel appreciated by them | r6n |
| 1 | | 2 | 5.00 | 1 | 6.00 | 4 | 3.00 |
| 2 | | 1 | 6.00 | 1 | 6.00 | 4 | 3.00 |
| 3 | | 1 | 6.00 | 0 | . | 0 | . |
| 4 | | 1 | 6.00 | 1 | 6.00 | 1 | 6.00 |
| 5 | | 1 | 6.00 | 1 | 6.00 | 1 | 6.00 |
| 6 | | 2 | 5.00 | 2 | 5.00 | 2 | 5.00 |
| Total | N | 6 | 6 | 5 | 5 | 5 | 5 |
| a. Limited to first 6 cases. | | | | | | | |

Figure 12

*Output table showing Cronbach’s alpha*

|  |  |  |
| --- | --- | --- |
| **Reliability Statistics** | | |
| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
| .852 | .862 | 6 |

Figure 13

*Output showing ANOVA results for calculating alpha*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVA** | | | | | | |
|  | | Sum of Squares | df | Mean Square | F | Sig |
| Between People | | 505.711 | 153 | 3.305 |  |  |
| Within People | Between Items | 32.148 | 5 | 6.430 | 13.163 | <.001 |
| Residual | 373.685 | 765 | .488 |  |  |
| Total | 405.833 | 770 | .527 |  |  |
| Total | | 911.544 | 923 | .988 |  |  |
| Grand Mean = 5.1959 | | | | | | |

Figure 14

*“Item-Total Statistics” table showing alpha reliability with one item removed at a time*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item-Total Statistics** | | | | |
|  | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted |
| R1 They are truly interested in me | 25.9545 | 15.076 | .678 | .825 |
| R4 They really value me | 26.1494 | 13.866 | .746 | .808 |
| R5 They feel a true liking for me | 25.9805 | 14.568 | .601 | .834 |
| r2n | 25.9610 | 13.528 | .699 | .815 |
| r3n | 25.6234 | 14.563 | .641 | .827 |
| r6n | 26.2078 | 13.473 | .538 | .857 |

Figure 15

*“Reliability Analysis” dialog box*

A screenshot of a computer

Description automatically generated

Figure 16

*“Compute Variable” dialog box*

A screenshot of a computer

Description automatically generated

Figure 17

*“Case Summaries” output showing the values of the six items making up the overall mean scale score*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Case Summariesa** | | | | | | | | |
|  | | R1 They are truly interested in me | r2n | r3n | R4 They really value me | R5 They feel a true liking for me | r6n | regard5.6 |
| 1 | | 4 | 5.00 | 6.00 | 4 | 4 | 3.00 | 4.33 |
| 2 | | 5 | 6.00 | 6.00 | 5 | 6 | 3.00 | 5.17 |
| 3 | | 6 | 6.00 | . | 0 | 0 | . | . |
| 4 | | 6 | 6.00 | 6.00 | 5 | 6 | 6.00 | 5.83 |
| Total | N | 4 | 4 | 3 | 3 | 3 | 3 | 3 |
| a. Limited to first 4 cases. | | | | | | | | |

Figure 18

*“Frequencies” dialog box*

A screenshot of a computer

Description automatically generated

Figure 19

*SPSS skewness and kurtosis statistics output table*

|  |  |  |
| --- | --- | --- |
| **Statistics** | | |
| regard5.6 | | |
| N | Valid | 256 |
| Missing | 5 |
| Skewness | | -.913 |
| Std. Error of Skewness | | .152 |
| Kurtosis | | .988 |
| Std. Error of Kurtosis | | .303 |

Figure 20

*SPSS histogram for regard scores with superimposed normal curve*

A graph with blue lines and black lines

Description automatically generated

Figure 21

*SPSS histogram for transformed regard scores with superimposed normal curve*

A graph with a line

Description automatically generated

Figure 22

*“Bivariate Correlations” dialog box*

A screenshot of a computer

Description automatically generated

Figure 23

*Pearson and Spearman correlation tables*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
|  | | regard5.6 | empathy5.6 | regard5.6sqrt1 | empathy5.6sqrt1 |
| regard5.6 | Pearson Correlation | -- |  |  |  |
| N | 256 |  |  |  |
| empathy5.6 | Pearson Correlation | .549\*\* | -- |  |  |
| Sig. (2-tailed) | <.001 |  |  |  |
| N | 256 | 256 |  |  |
| regard5.6sqrt1 | Pearson Correlation | -.993\*\* | -.548\*\* | -- |  |
| Sig. (2-tailed) | <.001 | <.001 |  |  |
| N | 256 | 256 | 256 |  |
| empathy5.6sqrt1 | Pearson Correlation | -.548\*\* | -.994\*\* | .551\*\* | -- |
| Sig. (2-tailed) | <.001 | <.001 | <.001 |  |
| N | 256 | 256 | 256 | 256 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | | |
|  | | | regard5.6 | empathy5.6 | regard5.6sqrt1 | empathy5.6sqrt1 |
| Spearman's rho | regard5.6 | Correlation Coefficient | -- |  |  |  |
| Sig. (2-tailed) | . |  |  |  |
| N | 256 |  |  |  |
| empathy5.6 | Correlation Coefficient | .546\*\* | -- |  |  |
| Sig. (2-tailed) | <.001 | . |  |  |
| N | 256 | 256 |  |  |
| regard5.6sqrt1 | Correlation Coefficient | -1.000\*\* | -.546\*\* | -- |  |
| Sig. (2-tailed) | .000 | <.001 | . |  |
| N | 256 | 256 | 256 |  |
| empathy5.6sqrt1 | Correlation Coefficient | -.546\*\* | -1.000\*\* | .546\*\* | -- |
| Sig. (2-tailed) | <.001 | .000 | <.001 | . |
| N | 256 | 256 | 256 | 256 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | | |

Figure 24

*Edited scatterplot of “regard5.6” and “empathy”*

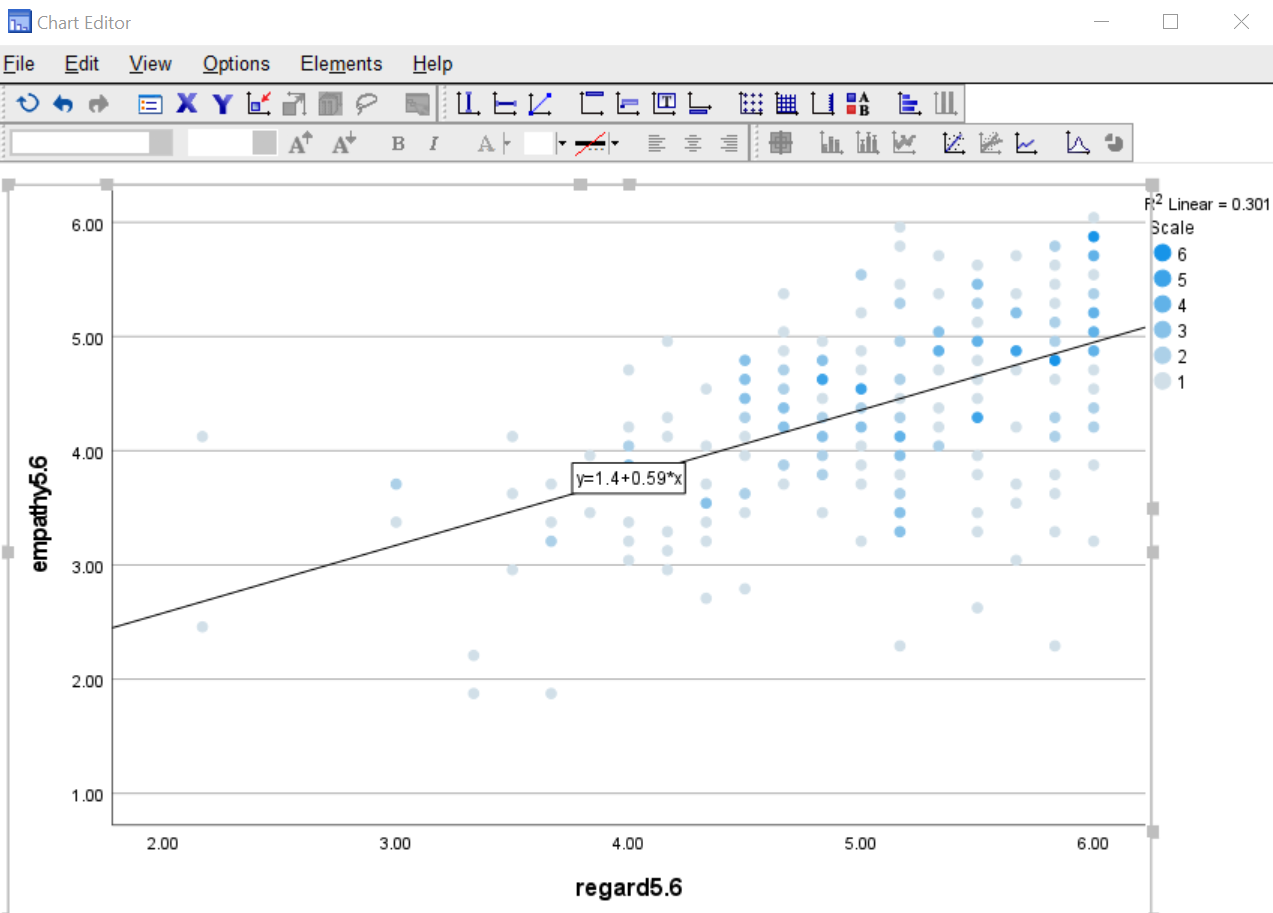
**

Figure 25

*“Simple Scatterplot” dialog box*

*A screenshot of a computer

Description automatically generated*

Figure 26

*“Linear Regression: Save” dialog box*

A screenshot of a computer

Description automatically generated

Figure 27

*“Univariate: Options” dialog box*

A screenshot of a computer

Description automatically generated