

Contents:

	Page
1. Index	3
2. Tools of Research	4
Measurements and statistical samples	4
Sampling in brief	6
Bias	7
3. Question design for SPSS	8
Ingredients of a good questionnaire	8
Types of question	8
4. Questionnaire design for SPSS	12
Question wording	12
Questionnaire structure	12
Questionnaire layout	13
5. Administering a survey	14
Pilot questionnaire	14
Coding the questionnaire	14
Collecting data	15
6. Using SPSS	16
Entering data via the keyboard	16
7. Other Useful SPSS Commands	24
Appendices	

1. INTRODUCTION:

'Research Powers the Modern World'

Research is about systematic investigation with the aim of developing knowledge and understanding. Research has one main goal, **discovery**. It begins with a question(s), to which the answer(s) is actively sought.

Data is collected in the form of facts and/or figures and is subjected to a form of scrutiny known as **analysis**, to help provide answers to the question(s) being asked. Facts must be judged logically and in an orderly fashion, and must be correctly **interpreted**.

Research is guided by a (**hypothesis**) reasonable guess as to the solution/answer. Findings of the research will either accept or reject the hypothesis. Research ends with a conclusion.

At the heart of all research is the need to tightly articulate the research problem, every thing else will flow from this.

Main Research Ingredients:

- Clear statement of the research problem
- Statement of hypothesis/hypotheses
- Recognition of relevant prior contributions- engage in the literature
- Data validly analysed / interpreted
- Interpretation given direction by the hypothesis
- Clearly stated conclusions
- Suggestion of further studies/research.

This brief handout aims to help the reader develop skills in one particular technique for gathering data – the survey- and one particular instrument for analysing the data, the software package SPSS. SPSS is a package that can be used to statistically analysis the data. It is normally used where the research is **quantitative** however it may also be useful when analysing **qualitative** surveys and interviews.

1 = 'best flavor', 2 = 'second best flavor', 3 = 'third best flavor', 4 = 'least best flavor'.

NOMINAL AND ORDINAL VARIABLES COLLECT WHAT IS KNOWN AS CATEGORICAL DATA, BY REQUIRING RESPONDENTS TO CHOOSE ANSWERS FROM INDEPENDENT AND MUTUALLY EXCLUSIVE CATEGORIES.

Categorical data are generated from questions which ask respondents to select one, or more categories which best describe their behaviour, attitudes, or characteristics.

Examples of categorical variables include gender, nationality, job description, eye colour etc. Individuals/objects can only be in one category.

Interval Scale- equal units of measurement.

This has all the qualities of an ordinal scale, but is characterised by equality of intervals. For example measuring time i.e. time could be measured in quarters of one year:

1 = 'January-March', 2 = 'April – June', 3 = 'July – September', 4 = 'October December'.

The interval between January and March is the same as the interval between October and December. Intervals are measurable.

A ratio scale is an interval scale that has an absolute zero.

An example of a ratio scale could be age:

1 = '0-19', 2 = '20-39', 3 = '40-59', 4 = '60-79' and so on.

Note:

You could ask individuals for their actual age and if required group the data later in your analysis.

Below is a sample of the popular attitude measurement techniques in social research. Are these scales interval or ordinal?

Likert Scale:

The workshop on SPSS has been excellent.

1 = 'strongly agree', 2 = 'agree', 3 = 'neither agree/disagree', 4 = 'disagree', 5 = 'strongly disagree'.

population made up of 100 females and 50 males. Do you sample randomly or do you select a sample that has 20 females and 10 males?

Good advice is to use probability sampling (random Sample) rather than non-probability sampling (influencing the sample composition). Sample error can be measured if you use probability sampling.

Other considerations:

When deciding on a sample you must try to avoid **bias**:

BIAS:

Any influence or condition that singly or together distort the data from what may have been obtained by pure chance, anything that disturbs the randomness of the data.

Example: if you select a sample from the telephone directory you are biased towards individuals who have telephones.

As a researcher you want to collect the **highest quality data**. Bias contaminates data, however don't panic, sometimes (more often than not in the social sciences), it is unavoidable, if this is the case, acknowledge it and suggest solutions.

3. Question design for SPSS

'To survey means 'to look' this can be extended to include questionnaires and interviews'

If you are using SPSS, one of the most important stages in the process of designing a survey is to design a questionnaire that will enable the required data to be gathered and prepared for input into the SPSS software.

Ingredients of a good questionnaire should include:

- Clear, unbiased questions using a variety of formats.
- Questions laid out in an appropriate order.
- An appreciation of the relationship between the format and the SPSS programme.

This workshop is based on a questionnaire designed to research characteristics of students as compared to lecturers in a local HE college.

Three main types of question:

Choice questions:

Used for behavioural questions, respondents are required to choose one or more answers from a list of two or more alternatives. Need some idea of the most likely response but it is normal to include an 'other' category or a 'don't know' category.

Dichotomous question:

What is your sex?

1= male, 0= female.

Multiple choice question:

How many units of alcohol did you consume in the last week?

1 = 0-4

2 = 5-9

3 = 10-14

4 = 15-19

5 = 20+

6= don't know

Multiple-response question:

What types of alcohol do you consume?

Wine	☐
Beer/Larger	☐
Spirits	☐
Alco-pop	☐
Other	☐ (please specify).....

Scaled questions:

Used to investigate attitudinal issues, as they are able to measure the intensity of responses.

Unipolar scales:

Measure the intensity of an attitude or opinion using a single descriptor.

How rewarding are your friends?

not at all rewarding	not very rewarding	rewarding	very rewarding	extremely rewarding
☐	☐	☐	☐	☐

4. Questionnaire design for SPSS

'A questionnaire that is difficult to understand is unlikely to be completed'

Question Wording:

Once you have decided on the type and style of question, the next step is to consider the clarity of the questions.

Good questions are:

- Relevant
- Answerable
- Impartial
- Unbiased -avoid emotive language
- Simple -avoid jargon and sophisticated descriptions
- Short – less than 20 words
- Unambiguous – difficult with attitude measurements.
- Specific – does over the last year mean since January or the last 12 months.
- Precise- tackle one issue at a time
- Positive- not hypothetical ask what have you done? not would you do?.
- Inoffensive
- Direct
- Structured

Questionnaire structure:

This is a craft that can be learnt and developed. There are three key issues to consider:

- Length-** there are no hard and fast rules, but the longer the questionnaire the more difficult it is to persuade respondents to complete it.
- Relevance-** make sure you only include relevant questions that will generate useful data. Review each question and ask yourself what data it is collecting.
- Sequence-** three key features
- coherence-* group related questions together.
 - sensitivity-* start with simple, non-emotive questions, use filter questions to ensure the respondent is a member of the population of interest. Questions requiring value judgements should come later.
 - influence awareness-* answers to questions can be influenced by the answers given to previous questions. Try to distance possible influences.

Poor structure
Alternative answers to closed questions.

Coding the questionnaire:

When you have piloted the questionnaire it is now essential to enter codes onto the final version, to give direction for entering the data into SPSS.

Closed questions are relatively simple to code.

Sex	Male	1
	Female	0

This workshop on SPSS has been excellent:

5 = 'strongly agree', 4 = 'agree', 3 = 'neither agree/disagree', 2 = 'disagree', 1 = 'strongly disagree'.

Open-ended questions can be coded by collapsing the responses into a number of categories. This can be done after the questionnaires have been completed.

Collecting data:

Methods:

- Face to face interviews
- Telephone interviews
- Mail surveys
- Self administered surveys
- E-mail

There are different issues to be addressed depending on the method chosen.

6. Using SPSS

Entering data via the keyboard:

We will now open the SPSS data file on characteristics of students and lecturers and enter a new question (students and lecturers 1)

The question has been devised using a Likert scale as follows:

Q.6 My relationship with my friends is rewarding

5 = 'strongly agree', 4 = 'agree', 3 = 'neither agree/disagree', 2 = 'disagree',
1 = 'strongly disagree'.

Step 1 Enter the question:

Step 5 Analysis of data:

Descriptive statistics:

The first step when carrying out analysis is to review the descriptive statistics:-

Frequencies: absolute frequencies – number of items of data for a particular question
relative frequencies - percentages.

- pull down analyze menu
- choose descriptive statistics
- choose frequencies
- highlight 'who are you?'
- click arrow-'who are you?' should now be in the variables box
- click Ok

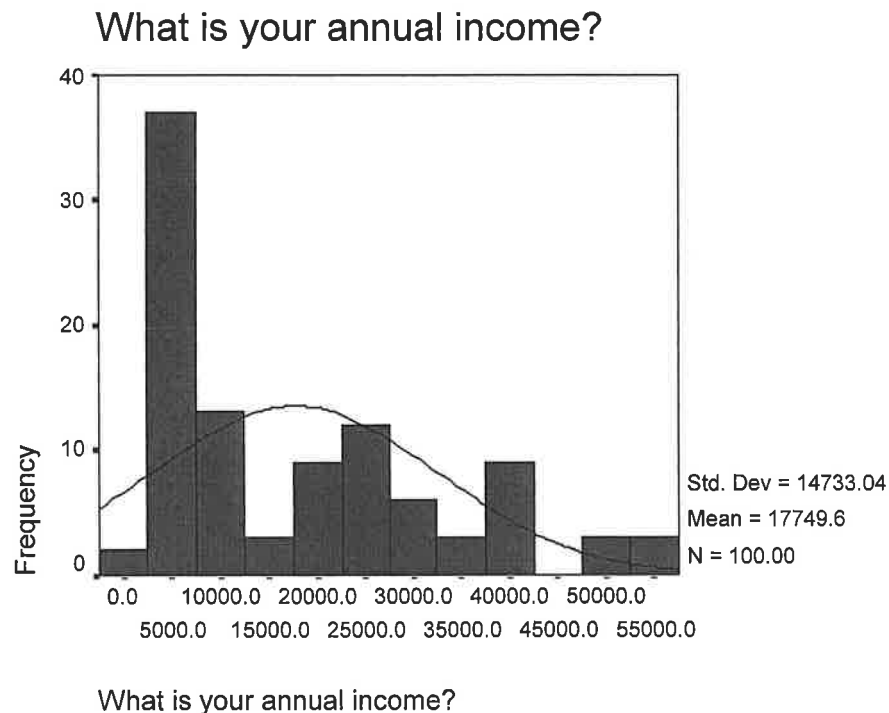
You will now move from the SPSS editor window to the SPSS output window.
Example:

Who are you?		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Lecturer	50	50.0	50.0	50.0
	student	50	50.0	50.0	100.0
	Total	100	100.0	100.0	

Descriptive statistics –mean, median, mode, standard deviation:

- pull down analyze menu
- choose descriptive statistics
- choose descriptives
- highlight annual income
- click arrow-annual income should now be in the variables box
- click options
- click mean-minimum-maximum-standard deviation-continue
- Click charts- bar chart or histogram with normal distribution curve.
- click Ok

HISTOGRAM WITH NORMAL DISTRIBUTION CURVE



Look for a normal distribution bell shaped curve. Check the **skewness and kurtosis** statistics. If you have a normal distribution in ideal circumstances the skewness and kurtosis statistics will be 0. The further from 0 the scores the more likely the data is not normally distributed.

Positive values for skewness indicate a pile up of scores on the left of the chart negative values a pile up on the right of the chart.

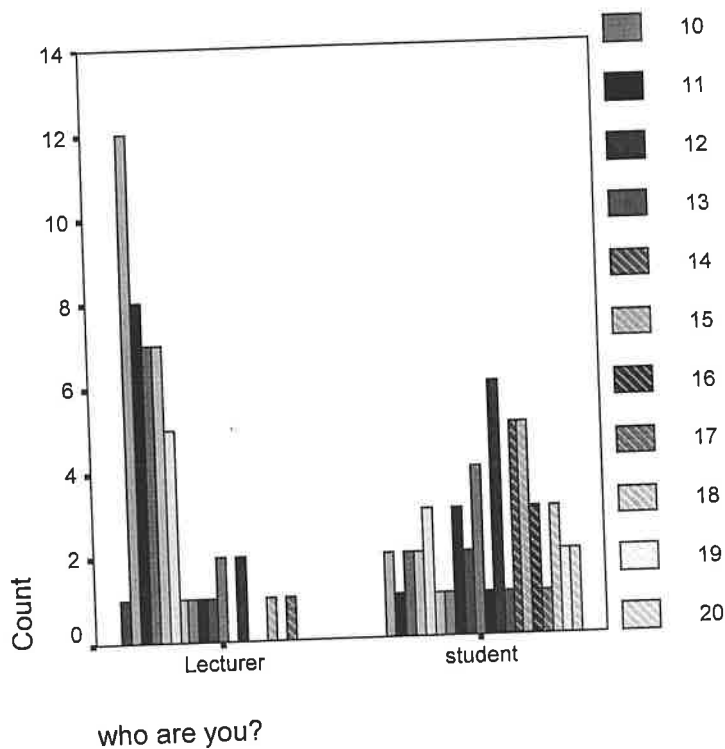
Positive values of kurtosis indicate a pointy distribution whereas negative values indicate a flat distribution.

If you decide to investigate further because you want to apply more sophisticated statistical techniques you should calculate z- scores for skewness and kurtosis or test for normality only relevant for parametric testing. (see page 21 or Field 2000 p 46).

In general you only need to consider the normality of the data if you intend to use parametric testing. In the social sciences students tend to collect nominal/categorical and ordinal data and therefore concentrate on non-parametric testing.

ASSUMPTIONS OF NON-PARAMETRIC TESTING:

- Random sample.
- Independent observations that can only be counted once, cases cannot influence each other.



To examine the relationship further we apply the **Pearson chi-square statistic**, this will tell us whether the relationship has arisen by chance (accept H0) or if there is a statistical relationship between the two variables, (accept H1).

Chi-Square Tests

Chi-Square Tests	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	43.310	20	.002
Likelihood Ratio	52.535	20	.000
Linear-by-Linear Association	37.704	1	.000
N of Valid Cases	100		

a 40 cells (95.2%) have expected count less than 5. The minimum expected count is .50.

The results from the chi-square are presented as above. Before we go any further you must:

Look at the last two line of the output, and check that the number of cells with expected frequency less than 5 is less than 20% and that minimum expected frequency is grater than 1. If these rules are not complied with the chi-square will be unreliable.

Clearly these conditions are not met. A solution if these rules are not met is to recode the variables reducing the number of categories by merging two or more categories.

The first column tells us the Chi-square value, the second column tells us the degrees of freedom the final column Asymp. Sig. (2 sided) tells you how confident you can be as whether or not there is a significant relationship between variables

.05 =	95% confidence
.01 =	99% confidence
.001 =	99.9% confidence
0.000 =	Greater than 99.9%

We can see that the results are very significant (greater than 99.9% chance) that the role of the participant will impact on the number of friends. We can therefore conclude that students have a larger friendship group than lecturers.

Pearson chi-square is a non-parametric test- it is less powerful than an equivalent parametric test but is easier to use as it has fewer requirements, particularly it lacks the requirement that the data is normally distributed.

NON-PARAMETRIC TESTING

If we are dealing with continuous variables (interval or ratio scales) we could use Pearson Correlation Coefficient- a test for a continuous linear relationship between variables.

Pearson Correlation Coefficient is a parametric test and should only be used on data that is normally distributed. You can test your data for normality by looking at a histogram of the data or by applying Kolmogorov-Smirnov test or the Shapiro and Wilk test. Alternatively you can use the parametric test anyway and hope that it does not invalidate your findings. Some statistics writers argue that most of the parametric approaches are fairly 'robust'; that is they will tolerate minor violations of assumptions. If you go ahead you will need to justify it look for other researchers who have done the same thing and quote them.

The nature of the data will affect the ways in which they can be measured and the types of statistical analysis that can be performed to generate useful information.

Continuous variables assume that any value within an interval is meaningful. Examples of continuous variables include, what is your age? Height? Weight? Income?

For example we could the hypothesis 'the older you are the more likely you are to fail your driving test' or 'the higher you income the longer your life expectancy'.

These two examples are of positive correlation (move in the same direction). There is also negative correlation (move in opposite direction), 'the more you revise the less likely you are make mistakes'

In our example we will hypothesis the greater the greater your income the more you drink.

- pull down analyze menu

7. Other useful SPSS commands

Merge Files

Select Data menu- merge file- add cases.

Select Lecturer and student data- open-OK.

Save to your existing file.

You should now find you have increased the number of cases.

Run the cross tab again to see what results you achieve, don't forget to recode the data as we did previously.

Select if:

Data – Select cases – and so on – use if you need to test accuracy of your entries or you need to condense data.

Computing a new variable:

Transform – compute – and so on- numerous applications if you want to manipulate data and create new variables.

Suggested formats for a dissertation:

Quantitative:

Title

List of contents

List of tables

Acknowledgements

Declaration

Abstract

Chapters

1. The research problem and the context.
2. Review of literature-prior contributions to the research problem
3. Methodology
4. Data findings and analysis
5. Discussion on findings
6. Conclusion and recommendations

Bibliography

Appendices

1. Definitions
2. Abbreviations
3. Other relevant appendices.

Qualitative

Title

List of contents

List of tables

Acknowledgements

Declaration

Abstract

Chapters

1. The research problem and the context.
2. Review of literature-prior contributions to the research problem
3. Methodology
4. Findings and discussion
5. Conclusion and recommendations

Bibliography

Appendices

1. Definitions
2. Abbreviations
3. Other relevant appendices.