

Lecture # 9

Data Analysis – 1

Transformation of variables

- The process of changing data from original form to a form that is more suitable to perform data analysis.
 - Prior to any modification do two things:

1. Revisit the original variables

- Revisit the original variables, research questions and theoretical models you have developed.
- The data must be organized in such a way that these concepts/variables are appropriately measured for the final analysis.

2. Familiarize yourself with the data

- Obtain a frequency table for each variable.
- It ensures that all the codes are within the valid range.
- It provides a first look at some of the characteristics of the sample and the sorts of responses that have been given.
- Get an idea of the shape of the distributions on key variables and note: which variables have very little variation, which categories have almost no cases and so forth.
- Being familiar with the data is helpful in guiding the subsequent data preparation steps.

Collapsing the categories

- Give justification.
- Too many categories make the cross table cumbersome. Better collapse
- Very few cases in a category. Better combine the category with another suitable category. Very low frequencies can produce misleading tables and statistics.

Two approaches to collapsing categories

1. Substantive approach

- Combining categories that seem to fit together: the categories have something in common. Matter of commonsense. occupations, subjects.
- With ordinal and interval variables which have ranked categories, collapsing is mainly a matter of establishing cutting points along a continuum.
- Just divide scale into three groups so that approximately the same number of *codes* are collapsed into each category

2. The distributional approach

- This approach is restricted to receding variables where the categories or values have a natural order from low to high.

- Often the meaning of a particular response to a question is best interpreted in *relative* than in *absolute* terms.
- For example, how are we to regard the income level of a person who earns Rs. 30,000 a month: is this low, medium or high?
- It depends on the other incomes with which it is compared. If most people earn less than 30,000, then it is *relatively* high; if most earn more, then it may be *relatively* low.
- Classify a particular value of a variable as high or low depending on the values of other people in the sample.
- This approach to collapsing categories has the advantage of letting the data define what is low, medium or high rather than us imposing some external, unrealistic definition.

Receding according to the distribution of the variable

- Divide the sample up into roughly equal sized groups of *cases*.
- Selecting *cutting points* on the variable that achieve the sub-division into equal sized groups of cases.
- Dividing the sample into two equal sized groups is called *dichotomizing*. Dividing it into three equal sized groups is called *trichotomizing*.
- To divide the variable into a different number of groups simply select different cutting points using the cumulative per cent column.
- To dichotomize (two groups) simply select the code that is closest to the 50 cumulative percentage. To divide into four groups use the codes that correspond to the 25, 50 and 75 cumulative percentage figures.

Steps in trichotomising a variable using the distributional approach to recoding

1. Ensure that the variable is either ordinal or interval-level.
2. Decide on the number of categories into which the sample (and variable) is to be subdivided.
3. If three categories are required (e.g. low, medium and high) divide the sample into the third with the lowest scores, the middle third, and the top third with the highest scores.
4. Obtain a frequency table of the variable.
5. Examine the cumulative percentage column
6. Locate the cumulative percentage closest to 33 per cent (this gives the bottom third of cases)
7. Then look across to the code or value that corresponds to this cumulative percentage. All people who obtained this code or lower will be placed in one category of the recoded variable. This group will represent the 'low' group.
8. Locate the cumulative percentage closest to 67 per cent to select the middle third. The code or value that corresponds to the 67th per cent figure in the cumulative percentage column provides the cutting point between the 'middle' and 'high' groups.
9. The remaining cases (those above the 67 per cent mark) will constitute the top third of cases —the 'high' group.

Recode the variable

- If the codes for the variable are from 0 to 10 and the first one third corresponded with 0- 5, second third corresponded with 6 to 8, and third one corresponded to 9 and 10, then it can be recoded as:
- Recode: 0 to 5=low ■ 6 to 8=middle ■ 9, 10=high.
- Use a computer program to apply the recode.
- In SPSS the instruction would be:
- RECODE varname (0 thru 5=1)(6 thru 8=2)(9, 10=3).
- VALUE LABELS varname 1 'low'; 2 'Middle'; 3 'high'.
- Instead of low, middle, and high it could be poor, moderate, and good

Rearranging categories

- Rearrange categories into a more logical order. New order can have the effect of:
- ■ creating an order more appropriate to the focus of the analysis;
- ■ making tables easier to read;
- ■ changing the level of measurement of a variable and thus affecting the methods of analysis that can be applied to the variable.

Score index

- A score index is a composite measure of a concept that is created by examining a person's responses to a set of questions and then combining answers into a single composite measure of the underlying concept.
- Responses could be given a numeric score. Numeric scores may be the same as codes given to responses.

Reverse coding

- Sometimes even though the categories of a variable are rank ordered they are ordered in the *opposite* way to that which is required.
- A more common reason for reversing the order of categories is when constructing a score index.
- Where we want to combine items (questions) that are coded in different directions we need to reverse code of some items so that they are all coded in the same direction.
- It happens when some items (questions) have been asked in positive form and others in negative form. Negative response to a negative item question will give positive connotation. Therefore, its code has to be reversed.
- ■ select items where the coding is in the opposite direction.
- ■ reverse code of these items.

Creation of new variables

- New variables can be computed from existing ones by using information from a set of questions. In this way some very useful and sophisticated variables can be developed. Normally this is done in one of three ways:
- 1. Developing score indexes;
- 2. Using conditional transformations;
- 3. Using arithmetic transformations.

The problem of missing data

- Almost always survey data have missing values:
- 1. How can questions be improved to minimize missing values?

- 2. How are missing values coded?
- 3. Do missing values introduce a bias in the analysis?
- 4. How are missing values to be taken into account during analysis? Are there ways of substituting valid values for missing values?

Checking for missing data bias

- Are people for whom we have missing values on a variable different from those with valid values?
- For example, do those who refuse to answer questions about income tend to have other characteristics such as ethnic background or education level in common?
- If certain types of people produce missing data for particular variables then the results of the analysis could be biased because some types of people are underrepresented in the analysis of that variable.
- We need to assess whether missing data do introduce bias.
- The simplest way of doing this is to divide the sample into two groups: those with missing values and those with valid values on a particular

Missing values are a problem

- Because they reduce the number of cases available for analysis.
- In scale construction, bivariate and multivariate analysis, this can lead to an unacceptable loss of cases.
- Alternative ways of dealing with missing values.
- 1. Deleting either cases or variables from the analysis.
- 2. Substituting the missing value with a valid code. This substitution requires using a new, best guess value is called *imputation*.

Deleting missing cases

- Any case that has missing data on any one of the *set* of variables being analyzed is eliminated from further analysis.
- This method is called *listwise* deletion of missing data.
- Can lead to the loss of a lot of data: valid answers on many questions are disregarded simply because of a non-answer on one question. Can lead to an unacceptable reduction in sample size.
- Hertel (1976) advised against this method if it leads to a loss of more than 15 per cent of cases.
- If the missing data are clustered in a small number of cases, then you might wish to eliminate those cases as the quality of the data from those cases may be suspect.

Delete variables

- If a particular variable is responsible for a large number of the missing values, that variable can be dropped from the analysis.
- Advantage: we do not lose any cases and, at the same time, we eliminate an unreliable item.
- The advisability of this approach depends on how important that particular variable is for the analysis.

Sample mean approach

- If we do not know the value on a variable for any given person, then the best guess for that person is the same as the measure of central tendency for that variable.

- With interval-level variables we can replace missing values with the value of the mean of that variable for the sample.
- The problem with this approach is that it reduces the variability of the sample on the variable and thus reduces the correlation between this and other variables

The group mean approach

- One way of overcoming the problem of reducing variability on the variable is to use *group* means rather than the overall sample mean.
- Divide the sample into groups on a background variable (e.g. ethnicity, gender, education) that is well correlated with the variable for which we want to replace missing values.
- Obtain the mean for the 'missing data variable' within each category of the selected background variable.

Random assignment within groups

- This approach is similar to the group means approach.
- Divide the sample into subgroups on the basis of background variables that are likely to be correlated with the variable for which missing values are being imputed.
- Locate a case with missing data on a particular variable
- Simply look at the value on the same variable of the nearest preceding case with a valid code
- Substitute this value for the missing value.

Data preparation checklist

- 1. Are the categories of the variable the most meaningful for the problem being addressed? If not, recode to better categories.
- 2. Are there too many categories to allow data to be clearly presented in tables or graphs? If so, reduce the number of categories by recoding.
- 3. Are there categories with too few cases to allow meaningful analysis? If so, combine small categories with other appropriate categories.
- 4. Which method of collapsing categories makes most sense for the variable? Is the method selected suitable for the level of measurement of the variable?
- 5. If you select the substantive method of collapsing, how many categories will you collapse to? Have you put too many people in one category so that the variable has very little variation?
- 6. If selecting the distributional method of recoding is the variable ordinal or interval (it should be)? How many categories are you going to collapse to? Why?
- 7. Is the order of the categories of the variables suitable for the way in which the variable will be used? Is the order logical? Is it easy to read?
- 8. Can the level of measurement of the variable be increased by reordering some of the categories?
- 9. If the variable is ordinal or interval does the variable need to be reverse coded? If the variable is to be combined with other variables in a scale is it coded in the right direction?
- 10. Do you require variables not in the data set but that can be constructed from existing variables?
- 11. If so, what is the level of measurement of the variables you will use to build another variable? a. If at least one of the variables is nominal you will need to use conditional transformations of variable creation. b. If all the 'building block' variables are ordinal or interval

will you use conditional transformations or the arithmetic transformations to create the new variable?

- 12. Do you need to compare people on different variables?
- a Are these variables measured using different units of measurement?
- b Are the distributions of the variables to be compared quite different?
- c If so you will need a method of adjusting variables to convert them to a 'common currency'.
- 13 If the variables require converting to a common currency:
 - a. What is the level of measurement of the variables?
 - b. If ordinal, convert the variables using the 'ntiles approach' and rank cases. How many ntiles do you require?
 - c. If interval, will you use the 'ntiles' or 'z-score' method of adjusting the variables?
- 14 Do the variables have missing data?
- 15. Does the missing data create bias?
- 16. How will you handle missing data? Will you: a Ignore it?
- b Delete cases or variables that produce a lot of missing data? If you delete cases or variables will you use the listwise or pairwise method of deletion?
- c Impute valid values to substitute for missing data? If so which method of imputation will you use?

Analyzing Survey Data

Overview of survey data analysis

- Four broad factors which affect how data are analyzed:
 - 1. The number of variables being examined.
 - 2. The level of measurement of the variables.
 - 3. Methods of analysis (descriptive or inferential).
 - 4. Ethical responsibilities.

1. The number of variables

- How we analyze data depends on what we want to know.
- What research questions to be answered?
- If we simply wish to describe one characteristic of the sample at a time (e.g. sex, education, income level) we will use a *univariate* (one variable) method of analysis.
- If we are interested in two variables simultaneously we will use a *bivariate* (two variable) method.
- If our research question makes use of three or more variables we would use a *multivariate* technique.
- All depends upon the question to which answer is to be found.

2. The level of measurement

- Nominal
- Ordinal
- Interval
- Ratio

- Not only higher level of variables provide more information but they also open up a wider range of methods of statistical analysis.
- The higher the level of measurement the more powerful are the methods of analysis that can be used.

Which level of measurement to aim for?

- 1. A wider range of methods of analysis is appropriate as the level of measurement of variable increases.
- 2. More powerful and sophisticated techniques of analysis are only appropriate for interval-level variables.
- 3. Higher levels of measurement provide more information, if needed.
- 4. Questions that require a lot of precision and detail can be unreliable since people often do not have accurate, detailed information.
- 5. People may be reluctant to provide precise information but may provide it in more general terms (e.g. income bracket, age cohort).
- 6. Numerical data collected in grouped form (e.g. age, income categories) can be converted to interval data if we make particular assumptions.
- 7. If a variable is measured at an interval level it is simple to reduce it to ordinal or nominal levels. With the exception already noted, data collected at low levels of measurement cannot be converted to higher levels.

3. Methods of analysis

- The method of analysis adopted depends on the complexity of the research question.
- If the research question involves only one variable, select a method of analysis appropriate for univariate analysis.
- If the question involves two variables use a method designed for bivariate analysis and so on.
- Within each level of analysis (univariate, bivariate, multivariate) there is a range of methods of analysis.
- The choice between methods is determined in part by the level of measurement of the variables involved: some methods of analysis are appropriate only for variables measured at certain levels

Descriptive and inferential analysis

- Statistics are only a tool for analysis: choose the appropriate tool for the job in hand.
- Two basic types of statistics: descriptive, inferential

Descriptive statistics

- Summarizes patterns in the responses of cases in a sample.
- Provide information about, say, the 'average' income of respondents or indicate whether education level affects the consumer behavior of people in the sample.
- There are three broad ways in which descriptive analysis is conducted and presented:
 - Tabular (present results in tables → univariate, bivariate),
 - Graphical (bar chart, histogram, piechart), and
 - Statistical (summary measures of information contained in a set of cases – mean).

Use of descriptive statistics

- The use of descriptive statistics requires the most imagination and skill and is the most productive in terms of understanding any phenomenon.
- Once we have analyzed data using descriptive techniques we should use inferential statistics to see the likely match between the sample patterns and those in the population.
- The first and key task is to discover these patterns and processes. This is the task of descriptive statistics.

Inferential statistics

- Provide an idea about whether the patterns described in the sample are likely to apply in the population from which the sample is drawn.
- When a sample is obtained by probability sampling use inferential statistics for this purpose.

Choice of statistics

- The choice of statistics is determined by many previous decisions such as:
- the method of analysis (descriptive, inferential),
- level of measurement of the variables, and
- complexity of the research question (univariate, bivariate or multivariate).
- Work out the decision.

Some methods of survey analysis and statistics

Univariate methods	Bivariate methods	Multivariate methods
1 Frequency distributions	1 Crosstabulations	1 Conditional tables
2 Scattergrams	2 Partial rank-order correlation	
3 Regression	3 Multiple and partial correlation	
4 Rank-order correlation	4 Multiple and partial regression	
5 Comparison of means	5 Path analysis	

Lecture # 10

Univariate Analysis

- Descriptive analysis
 - Three main ways of doing descriptive analysis using:
 - 1. Tables,
 - 2. Graphs, and
 - 3. Summary statistics.

1. Frequency table

- Required information:
- Avoid unnecessary clutter. Normally the following information should be provided:
 - 1 Number the table and give title;
 - 2 Give labels for the categories of the variable;
 - 3 Give column headings to indicate what the numbers in the column represent;
 - 4 Give the total number on which the percentages are based;
 - 5 Give the number of missing cases, if any;
 - 6 Give the source of the data.

Let us take job satisfaction

- Assumption
- The variable was measured with multiple dimensions, elements, and statements (questions).
- It is an ordinal level variable.
- The data were converted into a score index which was further divided into three categories high job satisfaction, medium job satisfaction, and low job satisfaction.
- SPSS program has give a frequency table.
- **Table 1. Level of job satisfaction of the employees**

	Level of job satisfaction	Frequency	Percent	Valid percent	Cumulative %
	Delighted	153	9.0	10.8	10.8
	Very pleased	286	16.9	20.2	31.0

	Pleased	268	15.8	18.9	49.9
	Mostly satisfied	235	13.9	16.6	66.5
	Unhappy	238	14.0	16.8	83.3
	Terrible	237	13.9	16.7	100.0
	Total	1417	83.5	100.0	
Missing		279	16.5		
Total		1696	100.0		

Reading the frequency table

- Different ways:
 1. Examine the particular percentages in specific categories.
 2. Look at the table *as a whole*— examining the *shape* of the distribution.
- A scan of the valid percentages can quickly give a sense of the heterogeneity (diversity) and homogeneity (similarity) in the sample.
- Distribution of ordinal variables: Normal/skewed
- Make interpretations of the table.

Take another variable

- Level of motivation of the employees.

- Measured with the help of different dimensions, elements, and statements.
- Answers put on Likert scale.
- Data quantified, and score index constructed.
- Cases divided into three categories of high level of motivation, medium level of motivation, and low level of motivation.
- Accordingly variable recoded.
- SPSS gave this table.

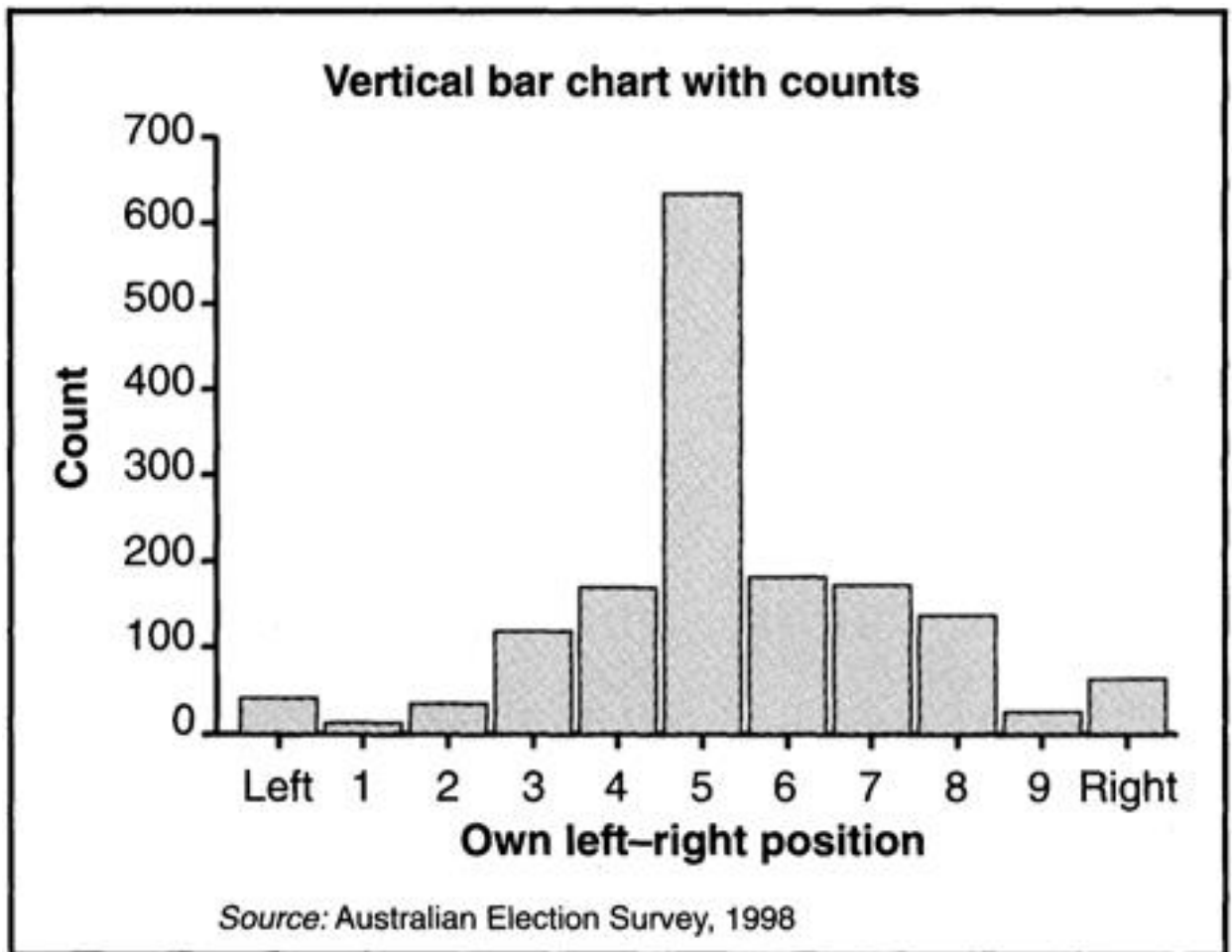
Table 2. Level of motivation of the employees

Level of motivation	Frequency	Percent	Valid percent	Cumulative %
Extremely motivated	100	5.9	5.9	5.9
Highly motivated	350	20.6	20.6	26.5
Mostly motivated	360	21.2	21.2	47.7
Somewhat motivated	330	19.5	19.5	67.2
Little motivated	380	22.4	22.4	89.6
Not motivated	176	10.4	10.4	100.0

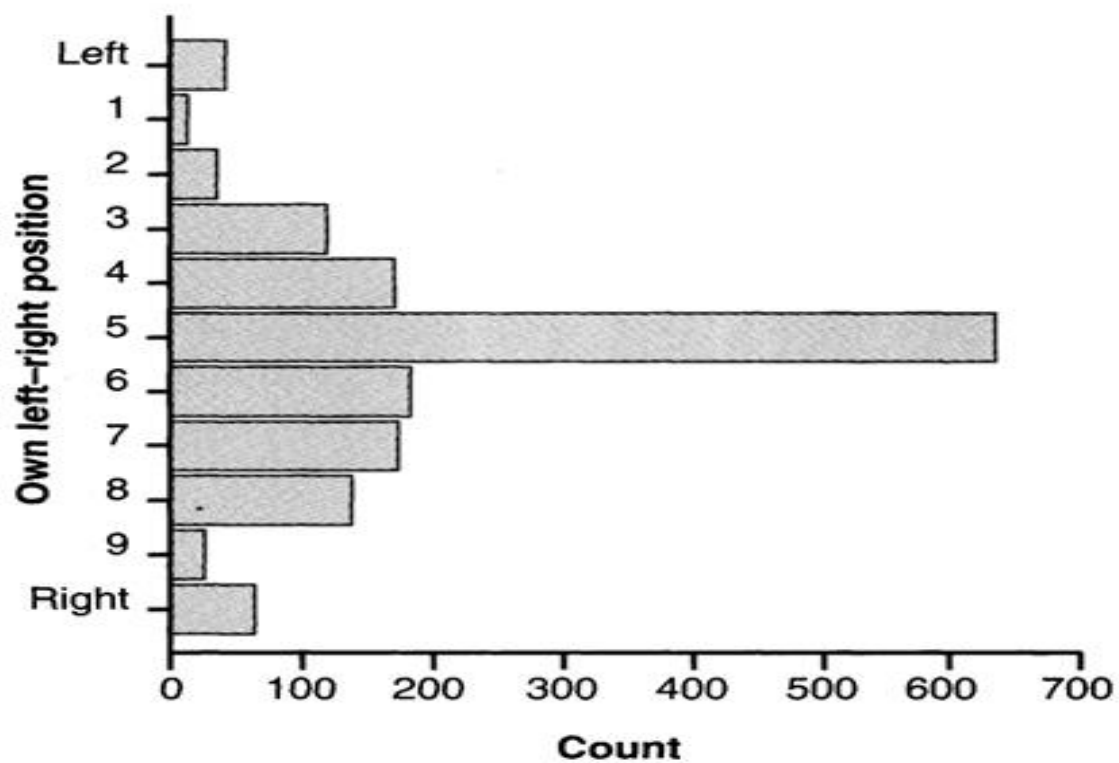
Total	1696	100.0	100.0	
-------	------	-------	-------	--

2. Graphic analysis

- Univariate distributions can often be displayed effectively with graphs.
- Range of graphs that can be used to display the distributions.
- Bar chart, line graphs, area graphs, histogram, pie chart.

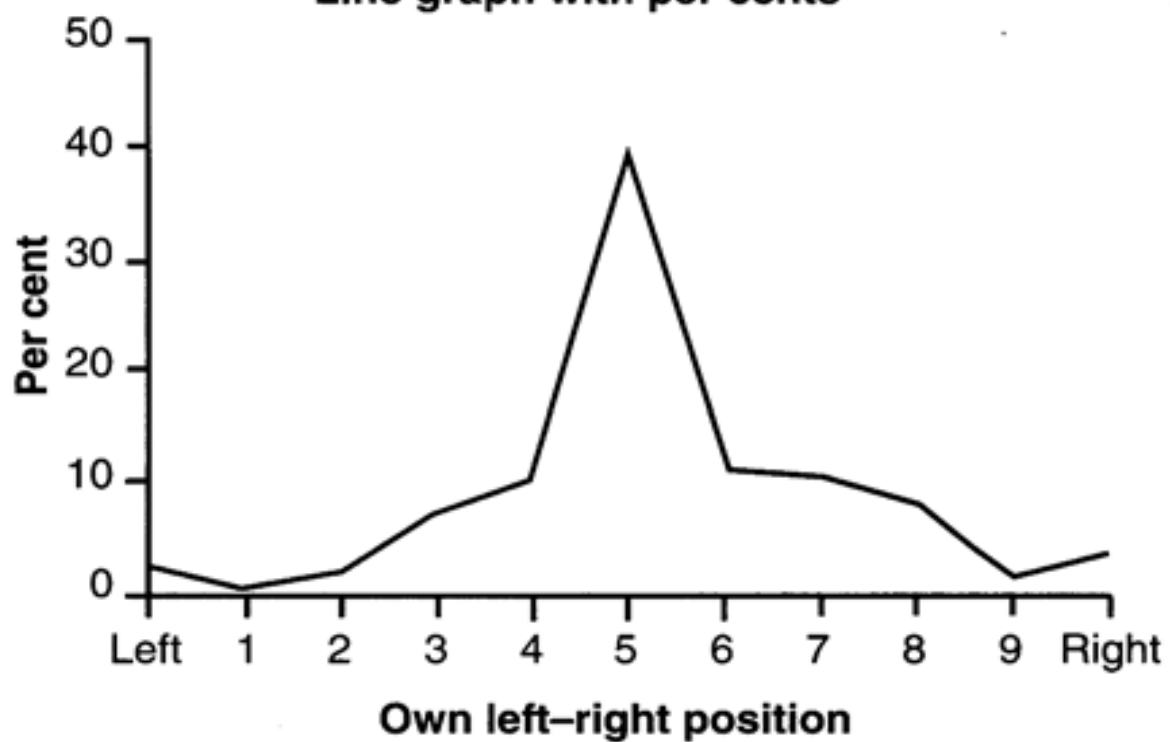


Horizontal bar chart with counts



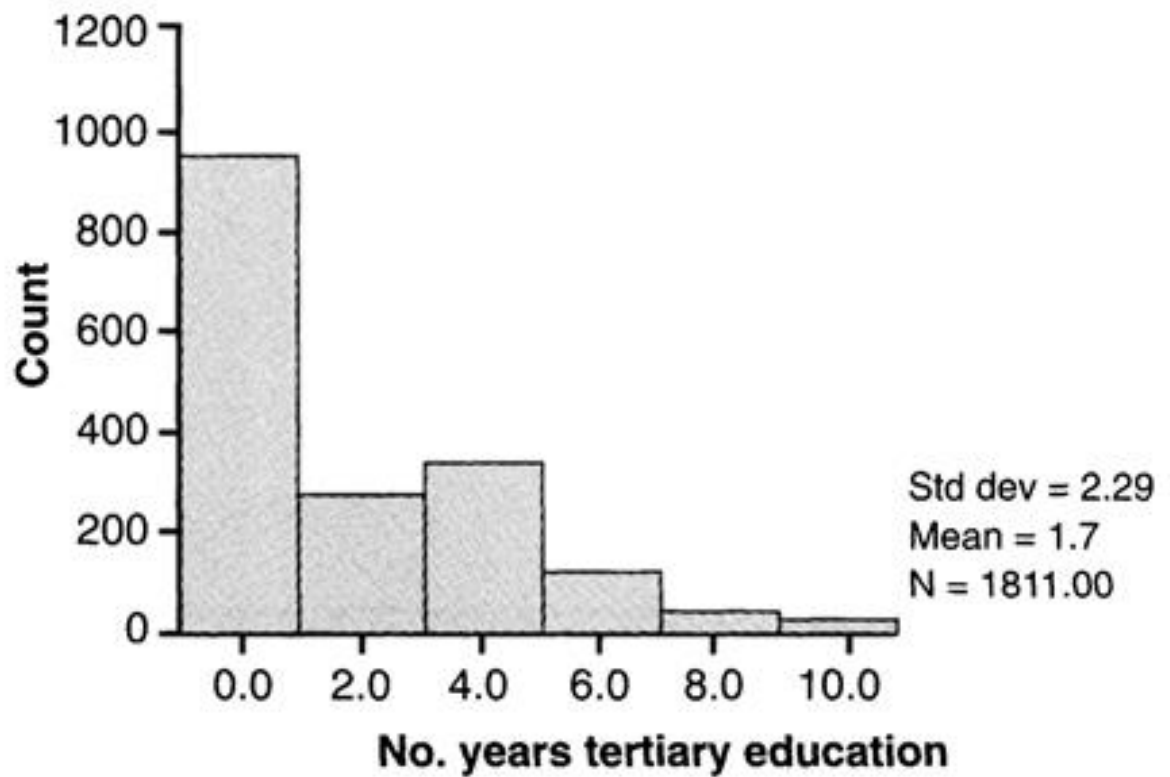
Source: Australian Election Survey, 1998

Line graph with per cents

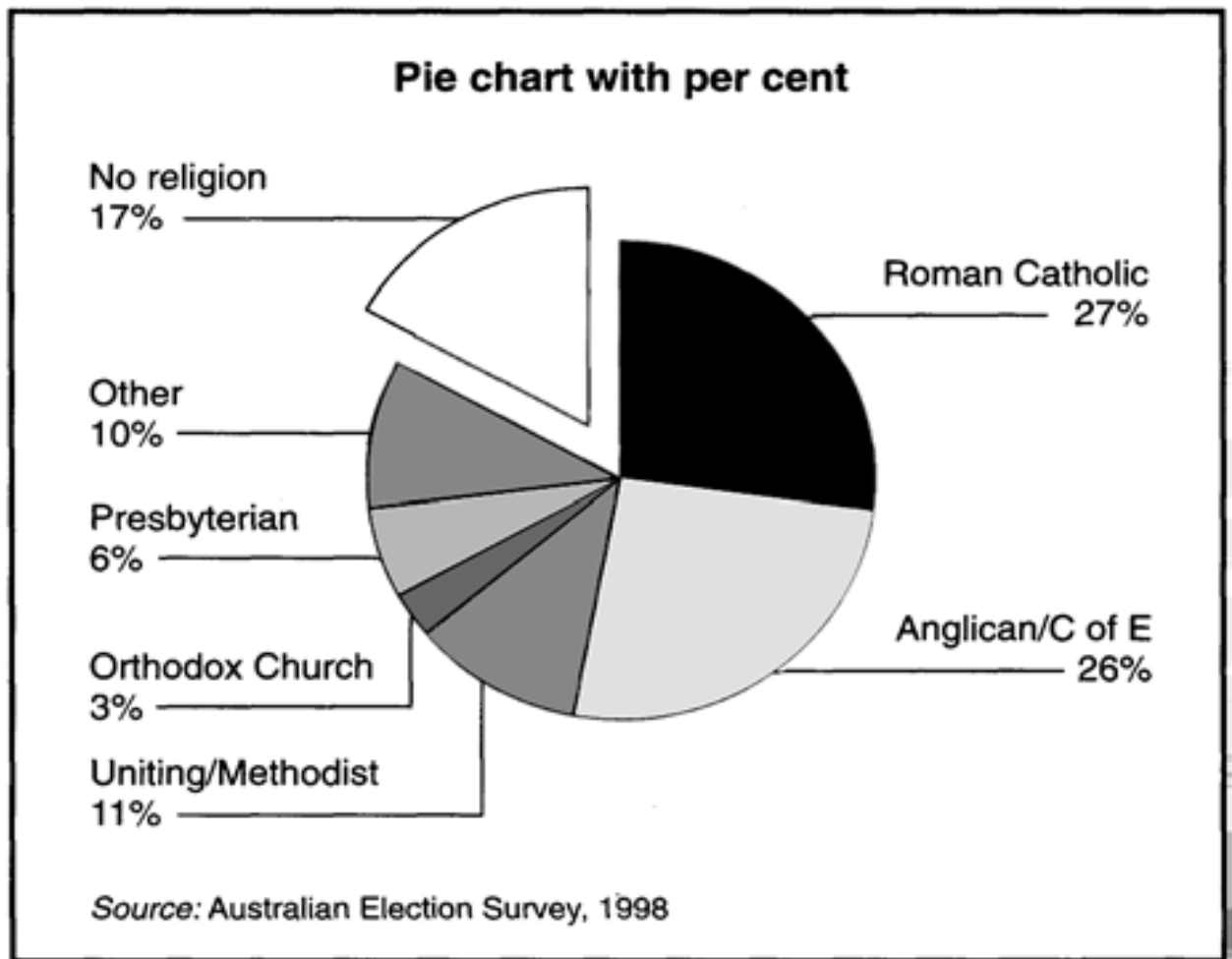


Source: Australian Election Survey, 1998

Histogram with counts



Source: Australian Election Survey, 1998.



3. Summary of descriptive statistics

- Characteristics can be summarized with simple and concise statistical measures
- Range of univariate statistics designed for this purpose.
- The choice of statistic depends on the level of measurement of the variable and the aspect of the distribution to be summarized.
- Central tendency: mean, median, mode.

Inferential Statistics

- Does the pattern in the sample reflect the pattern in the population from which the sample was drawn?
- Answering this type of question is the purpose of inferential analysis.
- There are two main approaches to inferential analysis: significance testing and making interval estimates.
- The particular methods of doing inferential analysis depends on the level of measurement of the variables.
- *Significance tests for nominal and ordinal variables*

Logic of significance testing

- It is standard to begin analysis by assuming a particular pattern in the population.

- For example, assume that the distribution of cases is even across the categories of the variable (e.g. 50 per cent in both categories of a two-category variable). This *assumption* about the population is called a *null hypothesis*.
- Examine the *actual* pattern in the *sample*. Is there the same percentage of cases in each category in the sample? It came out 56/44.
- It is unlikely that the distribution of cases in the sample will exactly match the assumption made about the population.
- The sample observation of 56/44 deviates from the 50/50 assumption for the population.
- There are two ways of interpreting the discrepancy between our *assumption* and the sample *observation*.
- 1 The sample is unrepresentative. Despite random sampling techniques we can still obtain poor samples. This is called **sampling error**.
- 2 The assumption of equal percentages in the population is incorrect. The difference between the pattern in the sample and the assumption for the population is much greater than can be accounted for by sampling error. If this is so then we would *reject the null hypothesis* of equal percentages in both categories in the population.

Tests of statistical significance

- When we have two alternative ways of interpreting results (sampling error vs. real differences) we have to have a way of working out which interpretation is correct. We do this with tests of statistical significance.
- *If there is no difference in the percentage of people in each category of the variable in the population, how likely is it that we would obtain a random sample in which sampling error produced a difference between categories as big as we have observed?*
- It is conventional to say that there is a chance that not more than five out of 100 samples would produce such differences due simply to sampling error.
- Our particular sample could have been one of those five.

Level of significance

- Level of confidence in the findings. 0.05 means 95% confidence.
- If we have a random sample then probability theory again provides the answer. If we took a large number of random samples most will come up with percentage estimates close to that which actually exists in the population. In only a few samples will the sample estimates be way off the mark. In fact the sample estimates would approximate a 'normal' distribution

One sample chi-square test

- Where the variable has *three or more categories* test whether the differences between the percentages across the categories is due to chance or is likely to reflect real percentage differences in the population.
- Assuming that the percentages in the population will be the same in all categories of the variable. Null hyp.
- Test to see if the sample fits this assumption.
- The one sample chi-square test is used to assess whether any misfit between the sample patterns and population assumptions is likely to be due to sampling error.

Bivariate analysis

Relationship between the variables

- Two variables are associated or related when the distribution of values on one variable differs for different values of the other.
- When subgroups (defined by belonging to one category or another of a given variable) differ systematically on another variable the variables are associated.
- Bivariate analysis provides a systematic way of measuring whether two variables are related and if so how strongly they are related.

Explanatory research

- Univariate analysis describes variation.
- Bivariate analysis provides explanation for the said variation in the variable.
- Variety of ways of establishing whether two variables are related.
- Which methods are used depends on:
 - the level of measurement of the variables,
 - the number of categories of each variable, and
 - the audience to which the analysis is directed.

Tabular method

- Cross-tabulations: displaying data for detecting an association between two variables.
- A set of frequency tables set side by side in one table.
- Let us take the two variables portrayed in the two univariate tables. Cross-tabulate.
- Six categories of each variable in the tables.
- Collapse the categories. Recode.

Elements of the crosstable

- **A cross-tabulation consists of:**
 - 1 *Labels and title*: the title indicates the two variables being cross-tabulated (dependent variable by independent variable), labels are provided for variables and for each category of both variables.
 - 2 *Rows and columns*: one *column* is allocated for each category of one variable and one row for each category of the 2nd variable.
 - 3 *Cells and cell contents*: cells represent cases who have *both* the characteristics indicated by the column *and* the characteristic indicated by the row. The contents of each cell may be the number of cases that have the two characteristics or the percentage with those characteristics.
 - 4. Marginals

Marginals

- *Marginals*: these represent the total number or percentage of cases in *a* particular category of a variable. These numbers will be very similar to the numbers in a frequency table for the same variable.
- Crosstabulation puts the data of two univariate (frequency) tables in one.
- One variable on one axis and the other on the other.
- Where XY axes meet the values of the XY variables are the lowest. Therefore the columns with lower levels of the two variables should be there.

Marginals

Table 3. Employees level of motivation by their level of job

	Level of job satisfaction			
Level of	Low	Medium	High	Total
Motivation	F		F	F
High				380
Medium				572
Low				465
Total	475	503	439	1417

Missing cases = 279

Source: Field data.

Complete this table

- Keeping in view the research hypothesis complete this in such a way that the hypothesis is validated.
- Hypothesis: There is a positive association between the employees level of motivation and their level of job satisfaction. Prove that:
 - Variables are associated.
 - The association is significant.
 - Association is in the proposed direction.
 - Is it linear?

Table 3. Employees level of motivation by their level of job satisfaction

	Level of job satisfaction			
Level of	Low	Medium	High	Total
Motivation	F		F	F
High	25		30	325 380
Medium		150		338 84 572
Low	300		135	30 465
Total	475	503	439	1417

Missing cases = 279

Source: Field data

Percentaging a cross-table

- Easier to interpret percentages than raw numbers when trying to detect association in a table.
- Convert cell frequencies into percentages.
- Can convert each cell frequency into three different percentages, each having an entirely different meaning.

Steps in detecting relationship

- 1 Determine which variable is to be treated as X.
- 2 Choose appropriate cell percentages:
- 3 Compare the percentages for each subgroup of the independent variable *within* one category of the dependent variable at a time.
- 4 If the independent variable is across the top, use column percentages and compare these across the table. Any difference between these reflects some association.
- 5 If the independent variable is on the side use row percentages and compare these down the table.

- Any difference between the percentages reflects some association. Amount of difference determines the level of significance.

Table 4. Employees level of motivation by their level of job satisfaction

Level of	Level of job satisfaction			
	Low	Medium	High	Total
Motivation	%	%	%	%
High	5.3	6.0	74.1	26.8
Medium		31.6	67.2	19.1
Low	63.1	26.8	6.8	32.8
Total	100.0	100.0	100.0	100.0
	(N= 475)		(N= 503)	(N= 439) (N=1417)

Missing cases = 279

Source: Field data.

The character of relationship

- Once the relationship is determined from a table, describe its character.
- There are three aspects to look at:
 - 1 strength
 - 2 direction
 - 3 nature.

Strength of relationship

- A strong relationship is one where the category of the independent variable to which a person belongs makes a very substantial difference to their characteristics on the dependent variable.
- If there are large differences between subgroups (as defined by the categories of the independent variable) there is a strong relationship.
- Can take arbitrary decision.

Direction of relationship

- Relationship can be positive or negative.
- A positive relationship is one in which people who score high on one variable are more likely than others to score high on the other variable; those who score low on one variable are more likely than others to score low on the other variable.* The
- Simplest way of detecting a positive relationship between variables is to examine the first row of the cross-tabulation (assuming the side variable is the dependent variable and is at least measured at the ordinal level). Simply compare the column percentages across this first row. If the percentages *become larger* as you move left to right across the row the relationship is probably positive. Also check the bottom row. If the percentages *decrease* as you move left to right across the bottom row the relationship probably is positive.

Nature of relationship

- Association of ordinal or interval variables can be linear, curvilinear or non-linear.
- A linear relationship means a 'straight line' relationship.

Table 4. Employees level of motivation by their level of job satisfaction

Level of	Level of job satisfaction			
	Low	Medium	High	Total

- | Motivation | % | % | % | % | |
|------------|-------|----------|----------|----------|----------|
| High | 5.3 | 6.0 | 74.1 | 26.8 | |
| Medium | | 31.6 | 67.2 | 19.1 | 40.4 |
| Low | 63.1 | 26.8 | 6.8 | 32.8 | |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | |
| | | (N= 475) | (N= 503) | (N= 439) | (N=1417) |
- Missing cases = 279
- Source: Field data.

When to use tables

- Although tables provide maximum information they are often inappropriate, especially when dealing with variables with a large number of categories.
- As a rule of thumb do not use tables when variables have more than six or seven categories and even then only do so with relatively large samples so that there are sufficient numbers in the categories.

Graphic presentation

- Outlines graphic methods and descriptive statistics suitable for relevant variables;

Using Summary Statistics

- While tables and graphs can provide detailed information about the way in which two variables are associated, summary statistics can provide a very concise index of the extent to which two variables are related.
- The main way in which to summarize the extent to which two variables are related is to use correlation coefficients (also called measures of association).
- Appropriate test.

Bivariate analysis for interval level variables

- Interval-level variables frequently have a large number of different values.
- Explores techniques suitable to analyzing interval variables *without* having to collapse these variables into a small number of categories;
- ■ Describes methods of analysis when the dependent variable is interval-level and the independent variable is categorical (comparison of means);
- ■ Describes the use of correlation coefficients when both the independent and dependent variables are interval-level (Pearson's correlation and rank—order correlation);
- ■ Introduces the analysis technique called regression analysis;
- ■ Describes the use of tests of significance and interval estimates suitable for interval-level variables.

Ethics and data analysis

- Data analysis is not just a technical matter. Social scientists have ethical responsibilities to analyze data properly and report it fairly.
- Selective reporting and selective, distorted analysis can readily paint a highly misleading picture.
- Huff's book *How to Lie with Statistics* (1954) provides plenty of lighthearted examples of how this can be done.
- Plenty of examples in scientific literature where people have either fabricated results entirely or changed figures to make them appear more impressive.

Unethical practices in data analysis

- Replication of experiments easier than sample surveys. [time and place become different]
- Results can be misrepresented without fabricating them. Inappropriate analysis of data.
- Inappropriate analysis may not be deliberate but may be due to lack of necessary skills to analyze data.
- Inappropriate analysis can be just as misleading as deliberate falsification of data. Unethical
- Instead of allowing the facts to speak for themselves, we make the facts speak for us.

- **Lecture # 11**
Data Analysis-2
(Qualitative)

Data management

Interviews, field notes, texts, visual data, transcripts.

Some similarities between quantitative and qualitative research

Differences between the two methods of data collection and analysis.

Critical of each other.

But there are points of similarity in data collection, management, and analysis.

1. Both are concerned with data reduction

- Both collect large amounts of data.
- Both distill and make it manageable.
- In quantitative research data reduction takes the form of statistical analysis. Frequency, percentage, mean.
- In qualitative research, develop concepts out of their often rich data.

2. Both are concerned with answering research questions

- Fundamentally concerned with answering research questions about the nature of social reality.
- More specific research questions in quantitative research.
- More open-ended research questions in qualitative research.

3. Both are concerned with relating data analysis to research literature

- Both relate findings to points thrown up by the literature relevant to the topic.
- Researcher's findings take significance when related to a body of literature.

4. Both are concerned with variation

- Seek to uncover variation.
- How phenomenon under study differs?
- What are the determinants of variation?

5. Both treat frequency as springboard for analysis

- Frequency with which the phenomenon under study occurs.
- In quantitative research it is in numbers (%).
- In qualitative research it is reported like "often", or "mostly".

6. Both ensure that deliberate distortion does not occur

- "Willful bias" or "consciously motivated misrepresentation" does not occur.
- No personal bias.

- Objectivity talked by the quantitative analysts. Reach the subjects' versions.
- Directly get the subject's interpretation of reality.

7. Both argue for the importance of transparency → Research design

- Both seek to be clear about their research procedures and how their findings were arrived at.
- Lay down clear research design.

8. Both have research methods appropriate to the research questions

- Both seek to ensure that, when they specify research questions, they select research methods appropriate to address those questions.
- Both at data collection and analysis stages.

Managing, analyzing and interpreting qualitative data

- Research design should include the appropriate plan.
- Some decision rules laid down.
- Mass of data. Like searching noodles from a soup.

Qualitative data management

- Like quantitative research there are no clear cut rules about how qualitative data analysis should be carried out.
- QR generates a large, cumbersome database because of its reliance on prose in the form of such media as field notes, interview transcripts, or documents.
- Qualitative data as "attractive nuisance", because of its richness but difficulty in finding analytic paths.

General strategy of data analysis

- Quantitative data analysis starts after data collection. Linear process
- Qualitative data analysis is iterative i.e. repetitive interplay between the collection and analysis of data. Interwoven procedure.
- Nevertheless, some strategy needs to be spelled out in the research proposal.

Analytic induction strategy:

- A question is a prerequisite for every research.
- Strategy begins with a rough definition of a research question (problem),
- proceeds to hypothetical explanation of that problem, and
- then continues on to the collection of data (cases). Cases that support the explanation.
- In case of encountering an inconsistent case, the researcher either redefines the hypothesis or excludes the deviant case. Process continues.
- Specifies the conditions that are sufficient for the occurrence of a phenomenon, but rarely specifies the necessary conditions.
- Tells why some people have adopted some behavior (become drug addicts) but does not tell why others have not done so.
- Does not tell how many cases to be studied for the confirmation of the validity of the hypothetical explanation. No guidelines provided.

Grounded theory

- Strauss – name associated with grounded theory.

- A theory that is derived from data, systematically gathered and analyzed through the research process.
- Data collection, analysis, and eventual theory stand in close relationship to one another.
- GT developed out of data and the approach is iterative –Data collection and analysis proceed in tandem, repeatedly referring back to each other.

Take grounded theory as a framework for analysis of data

- GT is not a theory but an approach to the generation of theory out of data. A framework for data analysis.
- In fact this approach generates concepts rather than a theory as such.
- Therefore:
- Grounded theory synonymous with inductive approach.
- It is a set of procedures (tools)

Procedures (tools) of grounded theory (4 Tools)

1. Theoretical sampling

- Theoretical sampling: process of data collection for generating theory whereby:
- the analyst jointly collects, codes, and analyzes his data and decides:
 - what data to collect next, and
 - where to find it, in order to develop his theory as it emerges. All based on researcher's theory (logic)
- The process of data collection is controlled by the emerging theory.
- On-going process
- Making comparisons to find out the variations

2. Theoretical coding

- Theoretical coding is the procedure for analyzing data, which have been collected in order to develop a grounded theory.
- Coding is the key process in grounded theory.
- Data are broken down into component parts which are given names.
- Code the emerging data as it is collected. Based on researcher's interpretation of data.
- Can be different levels of coding.

3. Theoretical saturation

- Saturation is a process that relates to two phases in grounded theory:
 - coding of data: you reach a point in reviewing your data to see how well they fit in with concepts or categories.
 - Collection of data: once a category has been developed, you may wish to continue collecting data to determine its nature and operation but then reach a point where new data are no longer illuminating the concept. Saturation point.

4. Constant comparison between data and conceptualization

- Maintaining a close connection between data and conceptualization, so that correspondence between concepts and categories with their indicators is not lost. Study of feminism.

- Compare phenomenon being coded under certain category so that theoretical elaboration of that category begins to emerge.
- Advice: write a memo on the category after a few phenomena had been coded.

Outcomes of grounded theory

- Concept(s): Labels given to discrete phenomena; concepts are referred to as building blocks of theory; concepts are produced through open coding.
- Category, categories: a concept that has been elaborated so that it is regarded as representing real world phenomena. Category may subsume 2+ concepts
- Categories are at higher level of abstraction around which other categories pivot.
- Properties: attributes or aspects of a category
- Hypotheses: initial hunches about relationships between concepts.
- Theory: A set of well developed categories ... that are systematically related through statements of relationship to form a theoretical framework that explains some relevant social ... or other phenomena.
- Substantive theory: empirical evidence of a substantive area (occupational socialization).
- Formal theory: Higher level of abstraction and wider range of applicability. Requires data collection in contrasting setting.

Memos

- Memos are notes that researchers might write for themselves and for those with whom they work.
- Serve as reminders about what is meant by the terms being used.
- Help in crystallizing ideas and not to lose track on various topics.

Limitations of grounded theory

1. Theory neutral observations are not possible

- Researchers are aware of the existing concepts and theories. Can they suspend their awareness until quite late in the process of analysis. No, because
- a)-- Conceptual armory of the discipline is already there. Researchers are aware and are sensitive to it.
- b)-- Observations are conditioned by many factors like what we already know about the social world.
- c)-- Researchers should be sensitive to the existing conceptualizations, and build on them.
- d)-- Researchers are required to spell out how the present study will contribute to the body of knowledge. They start with theories.

2. Doubtful whether the grounded theory really results in theory

- -- Provides a rigorous approach to the generation of concepts, but it is often difficult to see what theory, an explanation of something is being put forward.
- -- Most grounded theories are substantive in character i.e. specific phenomenon.

3. Grounded theory still vague on certain points

- What is the difference between concepts and categories?

4. GT is very much associated with an approach to data analysis

- Instead of being constructionist, grounded theory is mostly objectivist.

- Aims to uncover a reality that is external to social actors.
- Concepts and categories are labeled by the researcher using his own conceptual armory.
- Do not emerge out of the interaction of the researcher with the actors.

Lecture # 12

Data Analysis-2

(Qualitative)

Basic operations in qualitative data analysis

A continuum of analysis strategies

- Ideal types
- Prefigured technical (one extreme) → Objectivist end. Categories stipulated in advance.
- Emergent intuitive (other extreme) → Immersion/crystallization style. Templates. Developing codes/names. Depends upon researcher's intuitive capacities.
- Balance has to be struck. Problem with complexity of qualitative data.
- Nevertheless, analysis strategies have to be explained in the research proposal/R. design.

Generic Data Analysis Strategies

- Not a linear way. Not neat.
- Qualitative data analysis is a search for general statements about relationships and underlying themes; building of grounded theory.
- Generic term 'analysis' includes three activities:

Three major activities

- **1. Description**
- **2. Analysis**
- **3. Interpretation**
- The three activities are not mutually exclusive. Overlap. Each category shows varying emphasis.
- Typical researcher starts analyzing early. Needs to analyze as he goes along (to adjust his observation strategies). Nevertheless, some steps.

Analysis is data reduction

- Reams of collected data brought into manageable chunks, and interpretations made.
- Raw data have no inherent meaning; the interpretive act brings meaning to those data.
- Interpretation is the process of bringing meaning to raw, inexpressive data.
- Qualitative analysis transforms data into findings. No formula exists for that transformation. Only some guidance. No recipe.

ANALYTICAL PROCEDURES

- **Typical analytical procedures fall into seven phases:**
- 1. Organizing the data;
- 2. Immersion in the data;
- 3. Generating categories and themes;
- 4. Coding the data;
- 5. Offering interpretations;
- 6. Searching for alternative understandings; and

- 7. Writing the report.

1. Organizing the data

- Revisiting the “huge piles” of data.
- List on note cards the data that have been collected, perform the minor editing necessary to make field notes retrievable.
- Log the types of data according to dates, names, times, and places where, when, and with whom they were gathered.

Log of data gathering activities → example

Date	Place	Activity	Who	what
-----	XYZ	Focus group	6 teachers	Strategies for (names) doing research .
-----	XYZ	Observation	Aisha's	Seeing how is class room her teaching? .
-----	Aisha's	Interview	Aisha's	Challenges, spouse supports .

The researcher could also enter the data in computer software program

Call it: Data Preparation phase

- What data to be analyzed? Make a transcript of in-depth interviews or FGD
- Transcribing the data: translate from oral to written language. It is truth.
- Several key issues → how the data will be collected.

Key issues:

- Will you videotape or audiotape the session?
- Will you transcribe the entire data session? Summarize key passages.
- Will you transcribe all types of data you collect (laughter, pauses, emotions, non-verbal data – hand gestures)?
- Who will transcribe your data?
- What format? How will you represent the respondent's voice, nonverbal information, and so on?

2. Immersion in the data

- Reading, re-reading the data forces the researcher to be intimately familiar with people, events, and quotations.
- Description is of course part of it. Prepare appropriate schema for it (e.g. data recording charts). Help in streamlining data management, ensures reliability across several researchers.
- Guard against losing serendipitous findings.

3. Generating categories and themes

- This phase is the most difficult, complex, ambiguous, creative, and fun.
- Category generation involves noting patterns evident in the setting.

- Look at the meanings of categories → internally convergent (consistent) and externally divergent (distinct). In search of exhaustive and mutually exclusive similar to what is part of positivism.
- Categories become buckets or boxes into which segments of text is placed.

Since it is an inductive analysis, therefore:

- Discover patterns, themes, and categories in the data.
- Look for “indigenous typologies” created or expressed by the participants.
- Finally, these are analyst-constructed typologies grounded in the data (not explicitly used by people). Observer’s world rather than the world under study.
- Call it a terminology development process. Indexing
- Through logical reasoning the categories could be cross-classified. Matrix.
- **An empirical typology of teacher roles in dealing with high school dropouts**

- Teachers’ Behaviors towards dropouts.

beliefs about	Taking	Shifting
how to intervene	Responsibility	Responsibility
- Rehabilitation Counselor/friend: Referral agent;

	help kids directly	Refer them to other agents
--	--------------------	----------------------------
- Maintenance Traffic cop: Ostrich:

(caretaking)	Just keep them through the system	Ignore the situation and hope someone else does something
--------------	-----------------------------------	---
- Punishment Old fashioned Complainer:

	school master:	somebody should
	make them feel	remove the
	the consequences	problem kids

Source: Patton (1990) as quoted by Marshall and Rossman (2006, p. 160).

4. Coding the data

- The categories and themes have to be identified by some constructed names, signs, symbols, colors, numbers.
- The researcher marks the passages in the data using the codes.
- Hi-light passages with different colors. Each color can stand for a category/term.
- Software programs typically rely on abbreviations of key words, e.g.
- TCARE. LIS: Teacher’s caring as demonstrated through listening.

Coding process

- Data are broken down in its component parts.
- Process of giving a name, label, sign, symbol, a number, a color to a reality. Term that stands for this reality. Leads to developing a concept.
- Creating terminology relating to realities in text, observations. Also referred to as indexing.
- Can also be part of the vocabulary. Means to communication.

- Needs definition at some stage.
- Starting point in qualitative data analysis.

Bases for developing codes

- **Differentiating data and putting into different components (categories). What could be the bases?**
- 1. Of what general category this data item belong to? Men, women. Ethnicity
- 2. What does this item of data represent? A variable. Job satisfaction, OCB, brand culture.
- 3. What is this data item about? Element of a variable (brand culture).
- 4. Of what topic is this item of data an instance?
- 5. What question about a topic does this item of data suggest? What issue is the focus?
- 6. What sort of answer to a question about a topic does this item of data imply? Cause, effect
- 7. What is happening here? An event.
- 8. What are people doing? Playing, protesting.
- 9. What do people say they are doing? Local version
- 10. What kind of event is going on? Happy/sad, national sports, meeting.

Read through initial set of transcripts, field notes, documents

- While reading through and just at the end jot down a few general notes about what struck you i.e. important, interesting, or significant.
- These notes can contain terms which could be used as codes.

i) Open (ended) coding

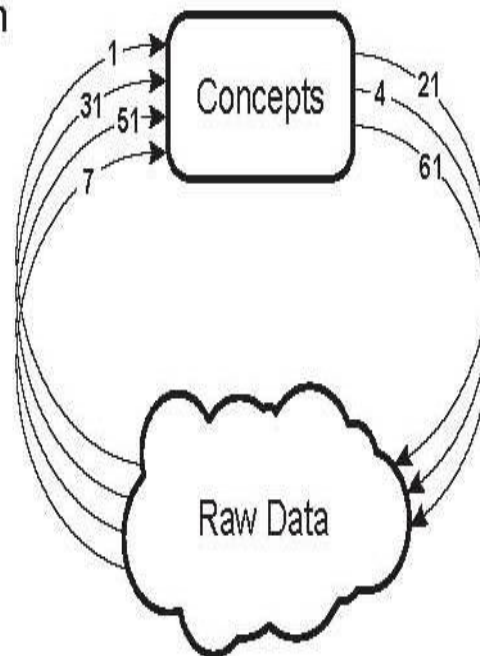
- First coding of data in which researcher examines the data to condense them into preliminary analytic categories, or codes.
- Process of breaking down, examining, comparing, and categorizing data. Open up the text and expose the thoughts, ideas, and meanings.
- Make notes as many as possible. Initially these will be very basic – perhaps key words used by respondents i.e. native or in vivo codes. Native words may be preserved for grounded theory.
- Generating an index of terms that will help to interpret and theorize in relation to data.

a) Generate as many codes as possible

- At least at the early stage of data analysis:
- Do line by line coding. No loss of data i.e. the perceptions and perspectives of those under study. Liberal coding.
- First the proliferation of codes. Do not worry about too many codes. Describe image of this reality (code). Name it i.e. concept. Leadership.

Codes to concepts

Successive Approximation

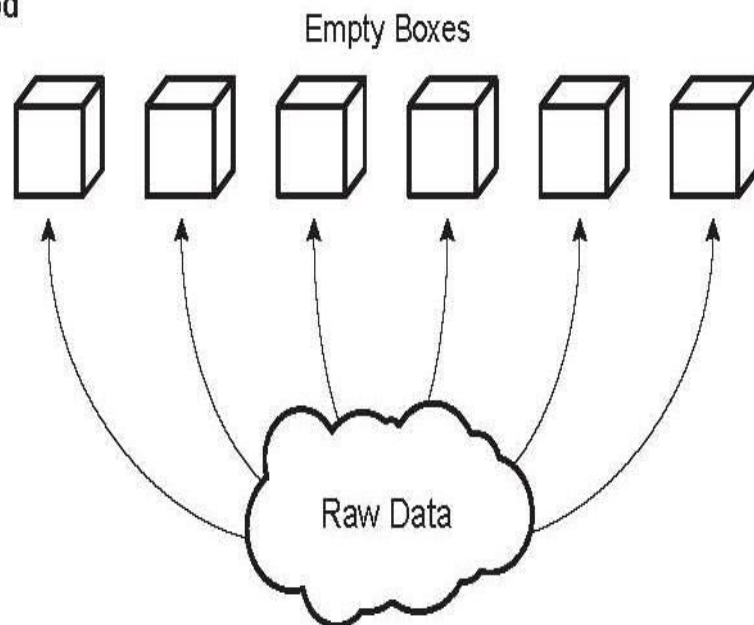


Concepts to categories

- Find out what is common in certain concepts
- Combine them into higher order and more abstract codes, called categories.
- These categories become buckets or boxes for pouring in the relevant data.
- Leadership to categories of leadership

Categories as empty boxes for sorting data

Illustrative Method



b) *Keep coding in perspective of the research question*

- Coding is not analysis. It is just a step.
- It is a means to data reduction. Locating abstract concepts in concrete data.
- Significance of the coded material for forging connections between codes in the light of research question.
- Interpret your findings.
- Reflective of the overall importance of the findings for research questions and the research literature.

c) *Review the codes*

- Review in relations to transcripts.
- Any overlapping. Two or more words for the same reality. Select the most appropriate, meaningful.
- Define the concept (keep it fluid)
- Do these words (codes) relate to concepts or categories available in the existing literature? Might be sensible to use the pre-existing terminology.
- Writing memos (basis of definition)

ii) *Axial coding*

- Axial coding is a set of procedures whereby data are put back together in new ways by making connections between categories. Linking codes to patterns, consequences. Focus more on the existing codes.
- Beginning to generate some general theoretical ideas about data.
- Relate the concepts, categories to the existing literature.
- Grouping answers to the research question. Code these connections.
- Create relational statements. Working hypotheses.
- Writing memos to the emerging theory.
- Go back to data for their confirmation.

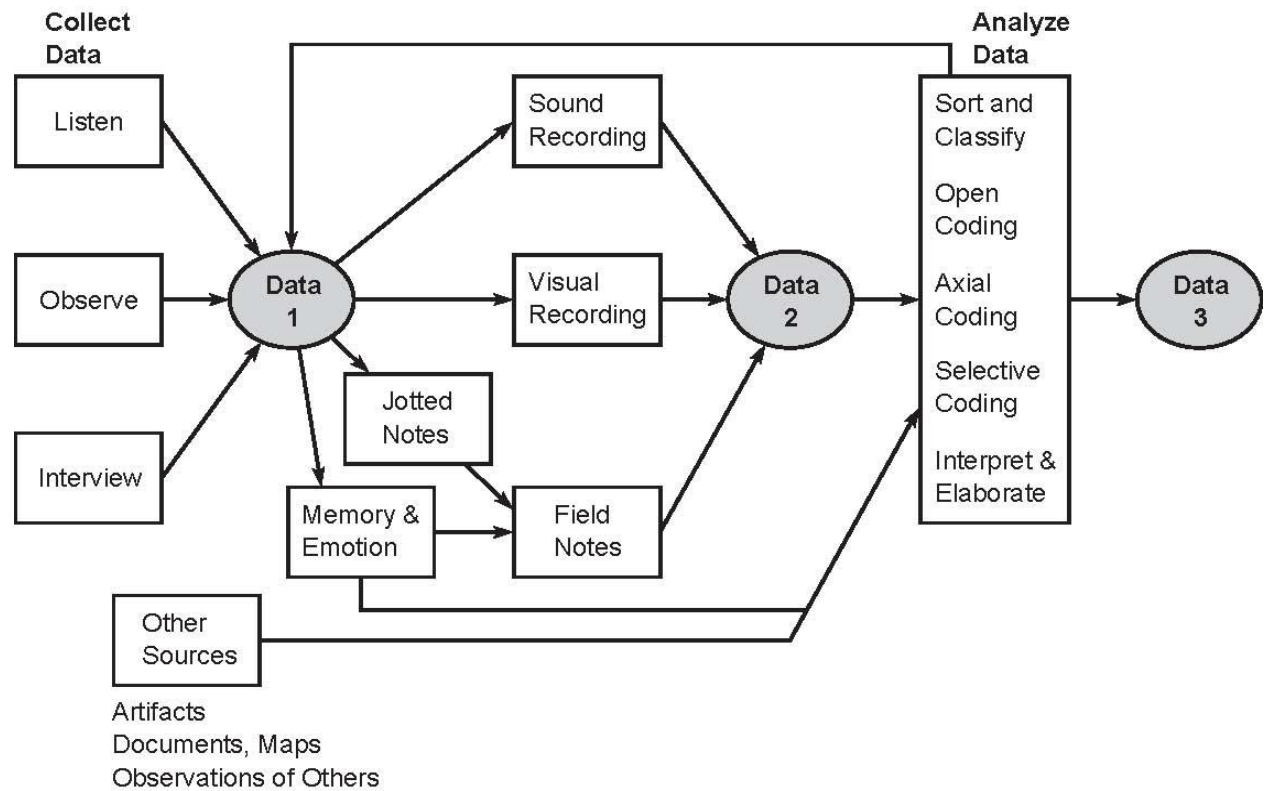
iii) *Selective coding*

- Also called focused coding.
- Emphasis on the most revealing codes. Core category.
- Core category is the central issue or focus around which all other categories are integrated. Storyline that frames the account. Gender relations. Possible situations: marriage, family, work, politics.
- Procedure of selecting a core category, systematically relating it to other categories, validating those relationships, and filling in those categories that need further refinement and development.
- Emerging theory is refined and elements integrated.

Some limitations of coding

- Problem of possible losing the context of what is said.
- Coding leads to fragmentation of data.
- The fragmented data may lose the flow of narrative.
- Some form of data may not be suitable for coding.

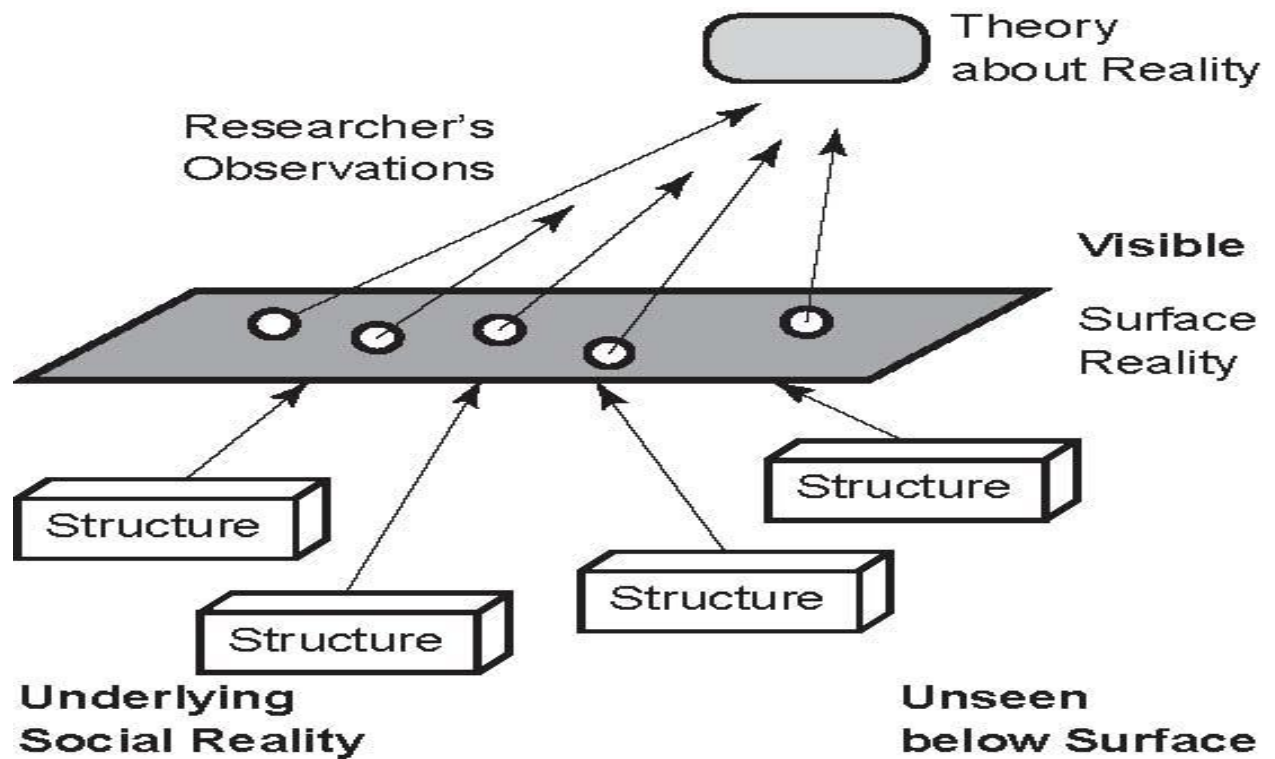
Analytic strategies



5. Offering interpretations

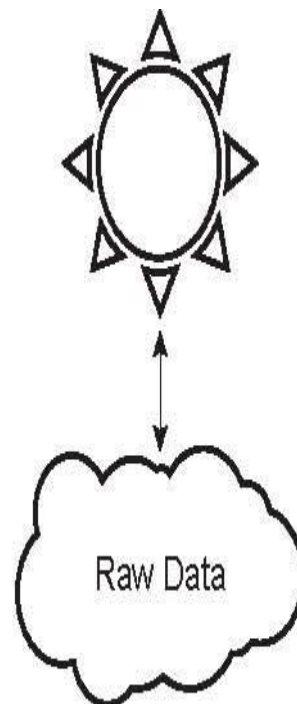
- Interpretation of the salient findings, making sense of the findings. Locating hidden realities.
- Findings as offering explanations, drawing conclusions, extrapolating lessons, and making inferences.
- Provide integrative interpretations of what has been learnt. "Telling the story." Story of OCB i.e. what has been found? Ideal type. Model of OCB. Compare and contrast with ideal type.
- Bring meaning and coherence to themes, patterns, and categories. Also develop linkages and story line that makes sense.

Interpreting the reality



Compare and contrast with ideal types

Ideal Type



6. Searching alternative understanding

- Search through the data for any alternatives to the apparent patterns that have been identified and codified.
- Look for anything hidden.

- Challenge the findings. Come up with alternative plausible explanations, and demonstrate that your explanation is the most plausible one.

7. Presenting the inquiry

- Written report the primary mode (dissertation divided into different chapters).
- Other modes could be theater skits, documentary, multimedia presentations.
- Different formats of presentation.
- Balancing the description, analysis, and interpretation.

Coding for thematic analysis

- Critical discourse analysis, qualitative content analysis, and narrative analysis.
- Search for themes – theme is same as code.
- Construct an index of recurring core themes and subthemes in the narrative.
- Are presented in a matrix. More like a variable indicated by its dimensions.

Thematic analysis

- Relevant to narrative analysis.
- Elucidation of life histories. Life events. Past, present, and its connection to future.
- *Thematic analysis*: an emphasis on what is said rather than how it is said. Contents of narration.
- *Structural analysis*: emphasis on the way story is related. It is not the contents of message rather how that message gets across. Persuasiveness.
- *Interactional analysis*: dialogue between story teller and listener. Co-construction of the meaning
- *Performative analysis*: Emphasis of performance – the use of words and gestures to get across a story.

• Lecture # 13

Discussion

Main purpose

- Explaining the meaning of results to the reader in the light of the objectives of the study.

Explaining the meaning of results

- You have carefully written the hypothesis.
- You have designed the study and collected the data.
- You have conducted the statistical analysis and grouped the summary results into tables and graphs. Given the interpretations.
- But what does all this mean?
- Explaining the meaning of the results to the reader is the purpose of the discussion section

Main functions of discussion

- To answer the questions posed in the Introduction,
- To explain how the results support the answers, and
- How the answers fit in with existing knowledge on the topic.
- Discussion is to provide a commentary and not a reiteration of the results.
- To help the reader to determine what can be positively learned and what is more speculative.
- The discussion has the potential to be the most satisfying, interesting, and meaningful aspect of the research.

- Discussion is the heart of the report

The Organization of the Discussion

- The organization of the Discussion is important.
- Try to develop an outline to organize your thoughts in a logical form.
- Here are some guidelines:
 - 1. Organize the Discussion from the specific to the general**
 - Organize the Discussion from the specific to the general: your findings to the literature, to theory, to practice.
 - Inductive approach
 - 2. Use the same key terms as used in research questions**
 - Use the same key terms, the same verb tense, and the same point of view that you used when posing the questions in the Introduction.
 - No contradictions. See what you said as part of your objective
 - 3. Begin by re-stating the research question**
 - Begin by re-stating the questions posed for answering or the hypothesis formulated for testing in the introduction.
 - More like restating the purpose of the study.
 - Take discussion as an independent unit.
 - If multiple research questions and hypotheses are posed, use the sub-headings: *"Research Question One," "Research Question Two," "Hypothesis One,"* and *"Hypothesis Two."*
 - 4. State the major findings of the study**
 - The discussion should begin with a statement of the major findings of the study. This should be the very first paragraph in the discussion.
 - It should be a direct, declarative, and succinct proclamation of the study results. However, it should not include data or reference to the study design.
 - Address all the results relating to the questions or hypothesis testing, regardless of whether or not the findings were statistically significant.
 - 5. Explain each finding in perspective**
 - Describe the patterns, principles, and relationships shown by each major finding/result and put them in perspective. The sequence is:
 - first state the answer to the research question,
 - then the relevant results, and
 - then cite the work of others.
 - No study is so novel and with such a restricted focus that it has no relation to other previously published papers.
 - Relate your study findings to those of other studies. Questions raised by previous studies may have served as the motivation for your study.
 - The findings of other studies may support your findings, which strengthens the importance of your study results.
 - It is also important to point out how your study differs from other similar studies.
 - What happened to the theoretical framework?

- Theoretical based on a borrowed theory or number of theories.
- It may have been one of the objectives to test the theory or part thereof.
- Ensuring that the review of literature is appropriately integrated into the discussion

6. Explain the meaning of the findings and why the findings are important

- No one has thought as long and as hard about your study as you have.
- You conceived, designed, and conducted the study, the meaning of the results and their importance seem obvious to you.
- Reader may not follow.
- Statistically testing the hypothesis is a finding.
- Explaining the process of tested relationship is lot more complex. Needs interpretation.
- Explanation if the relationship does not reach the significance level. Table/graphs show trends.
- Defend your emerging answers, if necessary, by explaining both why your answer is satisfactory and why others are not.
- Only by giving both sides to the argument can you make your explanation convincing.

Do not inflate the importance of the findings

- After all of the hard work that goes into a study, it is easy to attribute unwarranted importance to study findings.
- We all want our study to make an important contribution that will be cited for generations to come. However,
- Unwarranted inflation of the importance of the study results will disgust reviewers and readers.
- A measure of humility goes a long way.

7. Consider alternative explanations of the findings

- Despite efforts to remain objective, it is easy to consider only those explanations that fit your bias.
- The purpose of research is to *discover* and not to *prove*.
- Values in the selection of topic, perspective, research design and what not.
- Study designed to prove your bias rather than to discover the truth.
- When writing the discussion section, it is important to carefully consider all possible explanations for the study results, rather than just those that fit your biases.
- Discuss and evaluate conflicting explanations of the results. This is the sign of a good discussion.
- Do not use the discussion section as “bully pulpit” to criticize and attack other studies.
- Comparing your findings with alternative explanation should be done professionally.

8. Discuss any unexpected findings

- Might be additional findings.
- Discuss any unexpected findings. When discussing an unexpected finding, begin the paragraph with the finding and then describe it.
- May have implications for future research.
- It is possible that the discussion of the additional findings and implications are one of the most important outcomes of your research.

9. Acknowledge the study's limitations

- Even the best studies have limitations.
- Limitations of time, money, and other resources are external to the study. Every study has such bindings. These should not be mentioned.
- Limitations are the unanticipated constraints.
- Unanticipated problems associated with sampling, the measures, the treatment, the data analyses, and the execution of the procedures.
- Awareness of the limitations of the helps the reader to judge whether or not the methodology actually used provided an appropriate opportunity to answer the research questions or test the hypotheses.
- Others can learn lessons.
- Every research provides answers to certain question but opens doors for additional questions.
- New questions are not the limitations, could be the way forward.
- The limitations should not be fatal flaws that preclude the acceptance of the study.
- Better for you to identify and acknowledge your study's limitations than to have them pointed out by others. Don't over emphasize.
- Identify and comment on the relative importance of these to your interpretation of the results and how they may affect the validity of the findings.
- When identifying the limitations and weaknesses, avoid using an apologetic tone.
- Limitation of study could be used to make a suggestion for further research

10. Summarize concisely the principal implications of the findings

- The implications provides a series of recommendations based on the knowledge gained from the research.
- The implication of the study can be for:
- The advancement of knowledge (theory, research)
- The application of knowledge (practicing the findings, education and training, and public policy).
- Summarize concisely the principal implications of the findings, regardless of statistical significance.

11. Provide suggestions for further research

- Although a study may answer important questions, other questions related to the subject may remain unanswered.
- Some unanswered questions may become more focused because of your study.
- New questions may emerge.
- Provide recommendations (not too many) for further research.
- Do not offer suggestions which could have been easily addressed within the study, as this shows there has been inadequate examination and interpretation of the data.

12. Explain how the findings contribute to knowledge

- Explain how the results and conclusions of this study are important and how they influence our knowledge or understanding of the problem being examined.

- Strengthening the existing theory.
- Discovering new methodologies

13. Be precise and concise

- In writing of the Discussion, discuss everything, but be concise, brief, and specific.
- Care must be taken to provide a commentary and not a reiteration of the results. Side issues should not be included, as these tend to obscure the message.
- The key is to help the reader determine what can be positively learned and what is more speculative.
- No unwarranted speculation

What is the “take-home message”?

- Give the “Take-Home Message” in the Form of a Conclusion.
- What do you want the reader to remember from your study?
- The take-home message should be the first sentence of your conclusions

Qualitative research

- The structure of the discussion in a qualitative research can follow the same structure as in quantitative research reports.
- After a very short summary of your research question, you can repeat in one sentence the main result of your study.

Discuss how your findings relate to earlier research:

- Do they fill out the picture of what we already know, or possibly challenge or even contradict earlier findings?
- In what way has your study been important for the research community or for a larger audience?
- Can the results change the picture of similar phenomena in other cultures?
- Discuss the extent to which the findings with this data set may be relevant to the understanding of other situations.
- What are the concepts that can be transferred to other settings?
- Elaborate on the presented observations.
- Reflexively consider the research study and results.

Lesson 14

Thesis Writing

Every thesis is custom-made, yet some conventions of format

- Many universities have in-house, suggested formats or writing guides that researchers should be aware of.

Make an original contribution to knowledge

- The distinguishing mark of research at your level is *an original contribution to knowledge*.
- The thesis is a formal document whose sole purpose is to prove that you have made an original contribution to knowledge.
- Failure to prove that you have made such a contribution generally leads to failure.
- Therefore:

Thesis must show two important things:

- 1. you have identified a worthwhile problem or question which has not been previously answered, and
- 2. you have solved the problem or answered the question.
- Your contribution to knowledge generally lies in your solution or answer.

Make a clear statement of the question

- A very *clear* statement of the question is essential.
- To prove the originality and value of your contribution, you must present a *thorough* review of the existing literature on the subject, and on closely related subjects.
- Then, by making *direct* reference to your literature review, you must *demonstrate* that your question
 - (a) has not been previously answered, and
 - (b) is worth answering.
- Describing how you answered the question becomes easier.

A Generic Thesis Format:

- The general plan of organization for the parts.
- Tailoring the format to the research will help:
 - To obtain the proper level of formality, and
 - To decrease the complexity of the report.
- Formally a thesis/dissertation submitted to the university.
- Usually bound with a permanent cover.
- Consultants write long report for the organizations.

Nearly all theses/dissertations begin with four elements

- A title
- An abstract
- A table of contents
- Introduction
- These are all routine matters.
- The impression created at the start of thesis/dissertation is very important.
- Therefore the writing of first few pages should never be regarded as a triviality. Some suggestions:

The Title

- Give a short title of your research. Subject to change. Have an effective title (keep notes for possible changes).
- Titles should catch the readers' attention; also informing them about the main focus of study.
- Can also be a two part title. Breaking the title up into a title and subtitle when you have too many words A snappy main title and then a subtitle. *Policing the lying patient: Surveillance and self-regulation on consultations with adolescent diabetics*
- Have the most important words appear toward the beginning of the title.
- No use of ambiguous or confusing words,
- Include key words that will help researchers in the future to find your work.
- Follow a marketing approach.

- Should be short, catchy, meaningful. Do not use >15 words.

The Abstract

- Should cover the following:
- Your research problem.
- Why that problem is important and worth studying?
- Your data and methods.
- Your main findings.
- Their implications in the light of other research.
- Word limit usually of 100-150. Say as much as possible in as few words as possible.
- Make your abstract lively and informative.
- Emphasize your problem and content, not the fieldwork techniques.

The Table of Contents

- Not a trivial matter. A scrappy or uninformative table of contents will create terrible impression.
- To achieve 2 ends:
- 1. To demonstrate that you are a logical thinker, able to write a thesis/dissertation with a transparently clear organization.
- 2. To allow your readers to see this at once, to find their way easily between different parts of the thesis/dissertation and to pin point matters in which they have most interest.
- Use of double numbering system. ??

Table of Contents

Chapter	Title	Page .
I	Introduction	1
	Background	1
	Objectives	6
	Significance	9
II	Review of Literature	10
III	Theoretical Framework	18

When do you prepare the table of contents?

- It is outline of the structure of the thesis.
- Headings of chapters may remain the same. Subheadings likely to change as you proceed.
- Finalized towards the end especially with respect to pagination.

List of Tables

Number	Title	Page
1	Demographic Characteristics	35

1: Introduction

- Introduction is to answer the question: What is this thesis about? i.e.
- Why have you chosen this topic?
- Why this topic interests you?
- The kind of research approach or academic discipline you will utilize.
- State your research questions or problems.

- Role is to orient your readers. Do it clearly and succinctly. Do not over stretch. Also do not encroach upon other chapters (methodology).
- It is *not* just a description of the contents of each chapter.
- Chapter 1 to be last chapter, not in the literal sense. It should be rewritten and finalized at the end.

2: The Literature Review

- What should literature review contain:
- What do you already know about the topic?
- What do you have to say critically about what is already known?
- Has anyone else ever done anything exactly the same?
- Has anyone else done anything that is related?
- Where does your work fit in with what has gone before?
- Why is your research worth doing in the light of what has already been done?
- It displays your scholarly skills and credentials.
- You organize this section *by idea*, and not by author or by publication.

Do you need a literature review chapter?

- Know the relevant literature but just don't lump (dump) it into a chapter that remains unconnected to the rest of the study.
- Draw upon the literature selectively and appropriately as needed in telling the story of your research.
- So bring in appropriate literature as you need it, not in a separate chapter but in the course of your data analysis and/or any other discussion. Such a decision may be too radical. Write a conventional literature review chapter. Also cite literature in order to connect your narrow research topic to the directly relevant concerns of the broader research community.

3. Theoretical Framework

- Theory helps in understanding, explaining and prediction of the phenomenon.
- TF presents the theory which explains why the problem under study exists.
- Portrayal of how a particular theory provides explanation to the problem. Provides the structure that holds or supports the logic of research work
- This theory serves as a basis or foundation for conducting research on the issue.
- Need a strong foundation

4. Hypothesis and operationalization

- Hypothesis (es) or research questions.
- Show how the hypothesis (es) has (have) been drawn from the theoretical framework.
- Operationalization of the variables

5. Research Design

- Technical procedures must be explained.
- Supplement the material in this section with more details in the appendix. This part should address six topics:
- 1. *Purpose of study* → exploratory, descriptive, or explanatory. Why specific research design suited to the study?

- *2. Data collection methods.* Primary or secondary data used. How primary data were collected – survey, experiment, observation. Multiple techniques used – triangulation.
- *3. Sample design:* What was the target pop? Probability or non probability sample. Sampling frame. Type of sample. Selection process.
- *4. Instrument of data collection:* What instrument and why? Put a copy in appendix.
- *5. Fieldwork/Data collection:* how many, type of field workers used? Training/supervision. How was quality control assured?
- *6. Analysis strategy:* How was the analysis carried – score index applied, statistics used.

Limitations:

- No report is perfect, so indicate its limitations. For example problems with:
- Sampling procedures.
- Non response.
- Avoid over emphasizing the weaknesses.

Questions for a qualitative methods

- How did you go about your research?
- What overall strategy did you adopt and why?
- What design and techniques did you use?
- Why these and not others?

RD for Qualitative methods (Cont.)

To answer these questions describe the following:

- The data you have collected.
- How you obtained that data (e.g. issues of access and consent)
- What claims you are making about the data (representativeness of some pop./single case)
- The methods you have used to gather the data
- Why have you chosen these methods?
- How have you analyzed the data?
- The advantages and limitations of using your methods of data analysis.
- *Spell out your theoretical assumptions.*
- *Explain how you can generalize from your analysis.* [combining qualitative with quantitative, purposive sampling guided by time and resources, theoretical sampling]
- Avoid over-defensiveness. Also self-confidence should not mean lack of self-criticism. *Document* the rationale for your research design and data analysis. Ask colleagues' critique. Spencer et. al. gave the following guideline:
- Give an honest account of the conduct of the research.
- Provide full descriptions of what was actually done in regard to choosing your case (s) to study, choosing your method (s), collecting and analyzing data.
- Explain and justify each of your decisions.
- Discuss the strengths and weaknesses of what you did.
- Be open about what helped you and held you back.

6. Analysis of the data

- Present the findings in line with the objectives.

- Organize as a continuous narrative, designed to be convincing.
- Summary table and charts should be used
- Tables and charts may serve as points of reference to the data being discussed and free the prose from an excess figures.
- Detailed charts may be reserved for appendix.

7. Summary, Conclusions, and Recommendations

- Summary
- The Summary of contributions will be much sought and carefully read by the examiners.
- List the contributions of *new* knowledge that your thesis makes.
- The thesis itself must substantiate any claims made here. There is often some overlap with the Conclusions, but that's okay.
- Concise numbered paragraphs are again the best.
- Organize from *most* to *least* important.

Conclusions

- Conclusions are based on results.
- Conclusions are *not* a rambling summary of the thesis: they are *short, concise* statements of the inferences that you have made because of your work.
- Do not restate the research findings. Don't waste time of the reader. You got the results. So what?
- Organize conclusions as short numbered paragraphs, ordered from most to least important.
- All conclusions should be directly related to the research question.

Recommendations

- The biggest problem with this section is that the suggestions are often ones that could have been made prior to you conducting your research.
- Suggestions should emanate from experiences of conducting the research and the findings that have evolved.
- Make sure that your suggestions for further research serve to link your project with other projects in the future and provide a further opportunity for the reader to better understand what you have done.

References

- A bibliography is the listing of the works that are relevant to the topic of research interest arranged in alphabetical order of the last names of authors.
A reference list is a subset of the bibliography, which includes details of all the citations used in the literature survey and elsewhere in the report, arranged again, in the alphabetical order of the last names of authors.

Goals of referencing

- Crediting the author (s).
- Enabling the reader to find the works cited.

Give reference by following a style

- All references given *must* be referred to in the main body of the thesis.

- Organize the list of references alphabetically by author surname., depending upon the style you are using.

Different modes are followed

- APA, ASA, or any other.

Appendix:

- Appendix presents the “too ...” material.
- Any material that is too technical or too detailed should be in appendix. Material of interest only to some readers. Subsidiary materials.
- Any material which impedes the smooth development of presentation, but which is important to justify the results of a thesis.

Some tips

Review two or three well organized and presented dissertations

- Examine their use of headings, overall style, typeface and organization.
- Use them as a model for the preparation of your own thesis/dissertation.
- In this way you will have an idea at the beginning of your writing what your finished thesis/dissertation will look like.
- A most helpful perspective!

Don't have to proceed writing chapters sequentially

- The major myth in writing a thesis/dissertation is that you start writing at Chapter One and then proceed sequentially. This is seldom the case.
- The most productive approach in writing the thesis/dissertation is to begin writing those parts of the thesis/dissertation that you are most comfortable with.

Go with what interests you, start your writing there, and then keep building!

Always keep the reader's backgrounds in mind.

- Who is your audience? How much can you reasonably expect them to know about the subject before picking up your thesis?
- Usually your audience is pretty knowledgeable about the general problem, but they haven't been intimately involved with the details over the last couple of years like you have:
- Spell difficult new concepts out clearly.

Don't make the readers work too hard!

- The harder the examiners have to work to ferret out your problem, your defense of the problem, your answer to the problem, your conclusions and contributions, the worse mood they will be in, and the more likely that your thesis will need major revisions.
- Spell things out carefully, highlight important parts by appropriate titles etc.
- There's a huge amount of information in a thesis: make sure you direct the readers to the answers to the important questions.

Clear and unambiguous use of key concepts

- Prepare a list of key words that are important to your research and then your writing should use this set of key words throughout.
- There is nothing so frustrating to a reader as a manuscript that keeps using alternate words to mean the same thing.

- If you've decided that a key phrase for your research is "educational workshop", then do not try substituting other phrases like "in-service program", "learning workshop", "educational institute", or "educational program."

Table/figure presentation – a simple rule

- If you are presenting information in the form of a table or figure (graph, chart) make sure:
- You introduce the table or figure in your text.
- Insert the table/figure
- Following the insertion of the table/figure, make sure you discuss it.
- If there is nothing to discuss then you may want to question even inserting it.

Use the Table of Contents to help improve manuscript

- Use it to see:
- if you've left something out,
- if you are presenting your sections in the most logical order, or
- if you need to make your wording a bit more clear.
- See if the Table of Contents is clear and will make good sense to the reader. You will be amazed at how
- It is easy to see areas that may need some more attention.
- Don't wait until the end to do your Table of Contents.

Avoid

- Avoid using phrases like "Clearly, this is the case..." or "Obviously, it follows that ..."; these imply that, if the readers don't understand, then they must be stupid.
- Readers might not have understood because you explained it poorly.
- Avoid *red flags*, claims (like "Corporate social responsibility is the most important part of a management system"). Has to be demonstrated by evidence.

Lesson 15

Thesis Writing in Qualitative Research

No absolutes in this area.

- Some types of qualitative research will call for a different sort of report

Nearly all theses/dissertations begin with four elements

- A title
- An abstract
- A table of contents
- An Introduction
- These are all routine matters. But very important
- The impression created at the start of thesis/dissertation is very important.
- Therefore the writing of first few pages should never be regarded as a triviality. Some suggestions:

The title

- Give a short title of your research. Subject to change. Have an effective title (keep notes for possible changes).
- Titles should catch the readers' attention; also informing them about the main focus of study.

- Have the most important words appear toward the beginning of the title.
- No use of ambiguous or confusing words.
- Include key words that will help researchers in the future to find your work.
- Can also be a two part title. Breaking the title up into a title and subtitle when you have too many words A snappy main title and then a subtitle. *Building the brand by aligning employees: Contribution of Internal Branding in the shaping of Brand citizenship behavior*
- Follow a marketing approach.
- Should be short, catchy, meaningful. Do not use >15 words.

Abstract

- Should cover the following:
- Your research problem.
- Why that problem is important and worth studying?
- Your data and methods.
- Your main findings.
- Their implication in the light of other research.
- Word limit usually of 100 -150. Say as much as possible in as few words as possible.
- Make your abstract lively and informative.
- Emphasize your problem and content, not the fieldwork techniques.
- Written towards the end of the study.

Table of contents

- Not a trivial matter. A scrappy or uninformative table of contents (or none) will create terrible impression.
- Provides the macrostructure: logical progression.
- To achieve 2 ends:
- 1. To demonstrate that you are a logical thinker, able to write a thesis with a transparently clear organization. Not a confused thesis.
- 2. To allow your readers to see this at once, to find their way easily between different parts of the thesis and to pin point matters in which they have most interest.
- Use of double numbering system. ??

Table of contents

Chapter	Title	Page
I	Introduction:	
	aims of the study	1
II	Historical perspective	6
III	Review of Literature	10
IV	Research Design	15

Finalize towards the end of study

- It is an outline of the structure of the thesis.
- Headings of chapters may remain the same. Subheadings likely to change as you proceed.
- Finalized towards the end especially with respect to pagination.

1: Introduction

- Introduction is to answer the question: What is this thesis about? i.e.

- Why you have chosen this topic?
- Why this topic interests you?
- The kind of research approach or academic discipline you will utilize.
- Your research questions or problems.
- Role is to orient your readers. Do it clearly and succinctly. Do not over stretch. Also do not encroach upon other chapters (methodology).
- It is *not* just a description of the contents of each chapter.
- Introduction sets the scene and puts the research in context.
- If the research was about, for example, coping with stress by house doctors, the reader needs to know why the study was done and how it, broadly, relates to other researchs.
- It is useful to start with a sentence that describes exactly what this research is about.
- This is an account of a descriptive study of coping with stress in three groups of 10 house doctors working in three hospitals of Lahore.
- The study was a qualitative one involving interviews with a convenience sample of house doctors. Although there is a considerable amount of research carried out into whether or not internship is stressful, there is little known about the stress experienced and the coping strategies adopted by the house doctors.
- Here, the research question or the aim of the study is described.
- It can also be started directly like:
- The aim of this study was to address the question: 'What coping strategies for stress are adopted by the house doctors in their clinical and educational work settings'?
- At the end of the study, you are able to reflect back on the degree to which the aim was or was not achieved.

2: Review of Literature

- Misconceptions about the literature review chapter:
- It is done just to display knowledge that 'you know the area'.
- It is easier to do than your data analysis chapters.
- It is boring to read (and to write).
- It is best 'to get out of the way' at the start of research work.
- Begin by describing the literature that was searched.
- This involves describing the computer search engines used and the keywords entered into those engines.
- Needs thorough coverage and systematic review.
- Was the 'grey' literature reviewed?
- Grey literature is defined as:
- that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers.
- No formula for reviewing. Yet the reader needs to know who did the research and when. What was done and what was found? Thus an example of such reporting might be given:
- In a small scale study of 12 student nurses in a School of Nursing at the University of Health Sciences, Lahore, Saadia (2013) undertook two rounds of interviews to establish the factors that those students felt contributed to their ability to cope with stress. She found that most students

relied on family or friends for support. Some used stress reduction methods including breathing exercises, physical activities and diary keeping. Few expressed the view that they were unable to cope with stress. Ages and sex of the respondents were not quoted in the account of the study.

- Key research reports should be cited in this way. Make comparisons with others and give critical evaluation
- Others can be grouped together. For example, if a number of studies have been carried out using similar methods, with similar findings, these can be quoted thus:
- A number of studies, using the Pakistan Personal Stress Instrument (2001) – a free form reporting instrument – reported high levels of stress amongst younger students (Give multiple references).
- Finalize this chapter towards the end of your study, in order to cover the latest developments on the subject area.

3: Research Design

- David Silverman (2005) called this chapter as natural history of research conducted. All based on field notes, and diaries. Therefore:
- Be open and clearly state what actually happened during research.
- Demonstrate that you have the making of a competent researchers.
- The strategies, difficulties, and the way these were handled. Detective story.
- Document the rationale to back up your RD and data analysis.

(a) Questions to be looked at in RD

- Natural history of the research project should look at question like:
- How did you go about your research?
- What overall strategy did you adopt and why?
- What design and techniques did you use?
- Why these and not others?
- Context of the study.
- Why go for qualitative research?
- Why a particular population?

(b) Sample

- Sample: It is probably the case that convenience sampling is the most frequently used in qualitative studies. Applicable in case study of an organization.
- The reader needs to know the size and type of sample used in the reported study.
- If an unusual variant of sampling is used, it is useful to acknowledge the nature of it.
- Other comments about the sampling process may be helpful.
- A sample of 10 house doctors from each hospital, was invited to take part in the study. The sample was a convenience one and the snowball approach to sampling was adopted (Ref).
- Each house doctor was asked to recommend to the researcher another doctor who might be able to articulate views about his/her stress.
- There appears to be no general agreement about sample size in qualitative studies. Reports describe single-person studies (Refs). Other commentators suggest sample sizes ranging

from 6 (refs) to 30 (refs). It was felt that (10 X 3 = 30) respondents should be able to supply varied and detailed accounts for the purposes of this study.

(c) Data collection/field operations

- Describe what the researcher was aiming to find out,
- How the field operations were carried out?
- Problems and solutions (entry, rapport, logistics).
- Observations – participant, non participant.
- In-depth interviews (give all details: where, how long, one/two time, recording procedure, permission,)
- Focus- group discussions (same as in interview).
- Surveys.
- Use of documents.
- Triangulation
- Tell story, not critique – entry, rapport, ethical issues.

(d)Data processing method

- Part of RD or part of data analysis chapter?
- A variation is to be found in the amount of detail of reporting in this section.
- It is possible to describe, in full, how the researcher handled the data (transcription, transfer of data to computer, destroying the recordings on the completion of study).
- Or it is possible to write that 'The interviews were recorded and transcribed. The researcher then sorted those data into a range of categories and these are reported below'. Does not tell much.
- A comfortable compromise between these two extremes is probably achieved by reporting a little of what happened.
- Care should be taken with very general terms such as 'content analysis', when reporting data analysis. The term is probably so broad as to have little meaning. An example of how part of this section might be written is like this:
- All of the interview transcripts were read by the researcher and coded in the style of a grounded theory approach to data analysis (refs).
- Eight category headings were generated from the data and under these all of the data were accounted for.
- Two independent researchers were asked to verify the seeming accuracy of the category system and after discussion with them, minor modifications were made to it. In the grounded theory literature, a good category system is said to have 'emerged' from the data (refs).
- Other commentators have noted that, in the end, it is always the researcher who finds and generates that system (refs).

4.Data Analysis

- In empirical studies data analysis chapter is the key basis for the evaluation of your thesis.
- Data collection might have been highly laborious, even hazardous is neither here nor there. Final assessment is what do you do with your data. Actual writing up data analysis.

- Analysis may have begun earlier as part of data collection which helps you in theoretical sampling, yet there has to be a separate chapter in the thesis.
- You need to develop the skills to present your analysis clearly and cogently to your readers.
- A decision needs to be made, here, about whether or not (a) the researcher presents the findings on their own, without supporting discussion or (b) if he or she links the findings with the work of other researchers.

Structure the chapter

- Normally data analysis chapter will have three sections:
- An introduction: explain what you are going to do in advance.
- The main section: work through your data in terms of what you have already said.
- A conclusion: you summarize what you have shown and connect to next chapter.

(a) Introduction (to the chapter)

- At the outset, preface the chapter with:
- 1. Scene setting for the chapter, i.e. explaining the general area (s) that the chapter considers.
- 2. Locating the gap in knowledge which the chapter addresses.
- 3. Explaining how this chapter fills this gap.
- Providing a brief overview of what is in the chapter.

(b) Main section (of the chapter)

- Pull the areas outlined in the introduction and discuss each one separately.
- Guidelines:
- 1. Make one point at a time.
- 2. 'Top and tail' each data extract. Write a sentence or two before each extract to context it in the argument. Follow that up with detailed analysis.
- 3. Show that you know the limitations of both your data and your analysis of it.
- 4. Always number your extracts. Give your extracts two numbers, first will refer to chapter number and second to the order placed in the chapter (3.1). Line number should also be included for any extract over two lines.
- 5. Convince the reader. Convince of your interpretation.

(c) Conclusion (of the chapter)

- Tie the whole chapter together.
- Tell the reader what has been learnt in the preceding pages, but also to prepare for the chapter that follows. Therefore:
- 1. Explain what the chapter has done.
- 2. Describe the new questions the chapter has identified.
- 3. Explain where these questions will be addressed (next chapter).

5: Discussion and Conclusions

- The content of this section will be determined, to a considerable extent, on how the researcher has presented findings in the previous one.
- If links are made to previous research and some sort of critical debate is offered, it might be decided that a separate discussion section is not required.
- If the data were allowed, under the data analysis section, to stand on their own, then the discussion will enable the links to be made.

Suggested contents of the chapter

- 1. The relationship between the work done, the original research questions, and previous work. *Showing professionalism*. Discussion, theorizing.
- 2. Some answer to the classical examiner's question: if you were doing the study all over again is there anything you would do differently? Why so? Indicates the lessons learned from the study. *Limitations. Confessions and trumpets*.
- 3. Any implications for policy and practice. Applications should be realistic and no attempt should be made to extrapolate beyond the data.
- 4. Further research that might follow from your findings, methods or concepts used.

Abstract

- This is the piece of the work that will represent the researcher on bibliographic search engines and it may be all that many readers know of the work.
- A good abstract should contain details of the background to the study, the aim, the sample, the data collection and analysis methods and a summary of the findings.

Tips for tightening up

Make sure all parts are in place.

You are too close to your work to tell easily whether everything is properly in place.

Give yourself a critical distance.

1. Time. Put the thesis aside (back burner).
2. Give a talk on your research during the writing up stage. Listen to comments.
3. Have somebody to give a final reading.

Conclusions

- It is here that the researcher can both summarize his or her findings and suggest applications of those findings.
- Applications should be realistic and no attempt should be made to extrapolate beyond the data.
- Arguably, it is impossible to generalize from qualitative data (because of the sampling methods, the ways of collecting data and the methods of analysis) – and most would say that it is not the point of doing qualitative research to generalize in this way.
- However, it is sometimes a temptation for the researcher to project his or her findings into the future and to attempt to predict the implications of it through generalization out to a larger population.
- There is some debate amongst researchers about whether or not the researcher should offer an evaluation of the work, at the end of a paper. The researcher must decide whether or not to be critical of what he or she has done or to allow this function to rest with the reader. He or she should also be aware of the limitations of research itself (Shipman, 1997). Clearly, there are many questions that research cannot answer and the researcher needs to be open to the

possibility that his or her research does not answer all that many questions – and never answers any conclusively.

Abstract

- The final part of the writing process is writing an abstract for the paper. This is the piece of the work that will represent the researcher on bibliographic search engines and it may be all that many readers know of the work. A good abstract should contain details of the background to the study, the aim, the sample, the data collection and analysis methods and a summary of the findings. Good and bad examples of abstracts can be found within the pages of any international journals.
- **MA thesis = 10, 000 word**
- **Ph. D. = 70,000 to 100,000 words**

Thank you