Chapter FIVE

Data Preparation and Analyses

Data Analysis

- After administering and scoring research tools scripts, data collected and organized.
- The collected data are known as 'raw data.'
- The raw data are **meaningless** unless certain **statistical** treatment is given to them.
- Analysis of data means:
  - To make the raw data meaningful; or
  - To draw some results from the data after the proper treatment.
- The 'null hypotheses' are tested with the help of **analysis data** so to obtain some significant results.

Main Functions of Data Analysis

1. To make the **raw data** meaningful

2. To test null hypothesis

3. To obtain the significant results

4. To draw some **inferences** or make generalization &

5. To **estimate** parameters

Steps of Data Analysis

1. Preparing the data for analysis,
2. Analyzing the data, and
3. Interpreting the data (i.e., testing the research hypotheses and drawing valid inferences).

- **Technically speaking,**
  - **Processing** implies editing, coding, classification and tabulation of collected data so that they are amenable to analysis.
  - **Analysis** refers to the computation of certain measures along with **searching** for patterns of relationship that exist among **data-groups**.

DATA PREPARATION

- All studies require some form of data **collection** and **entry** mechanisms
- Because of their critical nature, data should be treated with the **great** respect and care.
- In addition to ensuring the confidentiality and security of personal data, researchers should carefully plan the way how the data will be logged, entered, transformed (as necessary), and organized into a database that will facilitate accurate and efficient **statistical analysis**.
- We can use **computer applications** for logging, tracking & analyzing research data. e.g: Microsoft Access, Microsoft Excel, SPSS, STATA

Analyzing/Summarizing Data

- Data are a bunch of values of one or more variables
- A variable is something that has **different values**
- Values can be numbers or names, depending on the variable
  - **Numeric**, e.g. weight
  - **Counting**, e.g. number of injuries
  - **Ordinal**, e.g. competitive level (values are numbers/names)
  - **Nominal**, e.g. sex, marital status, religious affiliation and gender

Analyzing/Summarizing Data…

- When values are names, visualize the frequency of each value with a pie chart or a just a list of values and frequencies.
- When values are numbers, visualize the distribution of all values in stem and leaf plots or in a **frequency histogram**.
  - Can also use **normal probability plots** to visualize how well the values fit a normal distribution.

DATA ANALYSIS

- The methods you use to analyze your data will depend on whether you have chosen to conduct **qualitative** or **quantitative** research,
  - This choice will be influenced by personal and methodological preference and educational background.
- **Analysis**:- the computation of certain guides or measures along with searching for patterns of relationship that exist among the data groups.

- For **quantitative data analysis**, issues of validity and reliability are important.
  - Quantitative researchers attempt to show that their chosen methods succeed in measuring what they relevance to measure.
  - They want to make sure that their measurements are stable and consistent and that there are no errors or bias present, either from the respondents or from the researcher.
- Qualitative researchers, on the other hand, might acknowledge that participants are influenced by taking part in the research process.
- They might also acknowledge that researchers bring their own preferences and experience to the research.

- Qualitative data analysis is a very personal process.
- For example ask two researchers to analyze a transcript and they will probably come up with very different results.
  - This may be because they have studied different subjects, or because they come from different political or methodological standpoints.
  - It is for this reason that some researchers criticize qualitative methods as '***unscientific' or 'unreliable***'.
  - This is often because people who come from quantitative backgrounds try to ascribe their methods and processes to qualitative research.
  - The two approaches are very different and should be treated as such.

When to analyze data

- Quantitative and qualitative data are analyzed in different ways.
  - For **qualitative data**, the researcher might analyze as the research progresses, continually refining and reorganizing in light of the emerging results.
  - For **quantitative data**, the analysis can be left until the end of the data collection process, and if it is a **large survey, statistical software** is the easiest and most efficient method to use. For this type of analysis time has to be put aside for the data input process which can be long and laborious.
  - However, once this has been done the analysis is quick and efficient, with most software packages producing well presented graphs, pie charts and tables which can be used for the final report.

Analyzing Qualitative Data

- To help you with the analysis of qualitative data, it is useful to produce an interview summary form or a **focus group** summary form which you complete as soon as possible after each interview or focus group has taken place.
  - This includes practical details about the time and place, the participants, the duration of the interview or focus group, and details about the content and emerging themes.
  - It is useful to complete these forms as soon as possible after the interview and attach them to your transcripts.
  - The forms help to remind you about the contact and are useful when you come to analyze the data.

- There are many different types of qualitative data analysis.
  - The method you use will depend on your research topic, your personal preferences and the time, equipment and finances available to you.
  - Also, **qualitative data analysis** is a very personal process, with few rigid rules and procedures.
    - Formats for analysis
    - The qualitative continuum/Range
    - Thematic analysis
    - Comparative analysis
    - Content analysis
    - Discourse analysis

- Different processes involved in qualitative data analysis.
  - You need to **think** about the data from the moment you start to collect the information.
  - You need to **judge** the value of your data, especially that which may come from dubious sources.
  - As your research progresses you need to **interpret** the data so that you, and others, can gain an understanding of what is going on.
  - Finally, you need to undertake the **mechanical** process of analyzing the data.
- It is possible to undertake the **mechanical** process using computing software which can save you a lot of time, although it may stop you becoming really familiar with the data.

ANALYSING QUANTITATIVE DATA

- Computing software (data analysis software.)
- Statistical techniques: For those who do not have access to data analysis software, a basic knowledge of statistical techniques is needed to analyze your data.
- Statistics is the body of mathematical techniques or processes for gathering, describing organizing and interpreting numerical data.
  - For researches yields quantitative data, statistics (statistical analysis of data) is a basic tool of measurement and research.
- Analysis in quantitative approach may, therefore, be categorized as:
  - **Descriptive** Statistical analysis and
  - **Inferential** Statistical analysis (Inferential analysis is often known as statistical analysis).

Descriptive Statistical Analysis

- Concern the development of certain indices from the raw data.
- Descriptive statistical analysis is concerned with numerical description of a particular group observed and any similarity to those outside the group can not be taken for granted.
- Descriptive statistics allow the researcher to describe the data and examine relationships between variables,
- "Descriptive analysis is largely the study of distributions of one variable.
  - This study provides us with profiles of companies, work groups, persons and other subjects on any of a multiple of characteristics such as size, composition, efficiency, preferences, etc.".

- This sort of analysis may be:
  - In respect of one variable (described as **unidimensional analysis**), or
  - In respect of two variables (described as **bivariate analysis**) or
  - In respect of more than two variables (described as **multivariate analysis**).
- In this context we work out various measures that show the size and shape of a distribution(s) along with the study of measuring relationships between two or more variables.
- The principal objective of descriptive statistics is to accurately describe distributions of certain variables within a specific data set.

Univariate Analysis

- If your goal is to describe what you have found, all you need to do is count your responses and reproduce them.
- Simply a complete list of all possible values or scores for a particular variable, along with the number of times (frequency) that each value or score appears in the data set.
- Generally, this method concerned with calculations of several measures mostly concerning one variable.
- There are more precise ways of describing the shape of the distribution of values for a specific variable.
- These include measures of **central tendency** and **dispersion**.
- **Central Tendency :**
- The central tendency of a distribution is a number that represents the typical or most **representative value** in the distribution.

- Measures of central tendency provide researchers with a way of characterizing a data set with a single value.
- The most widely used measures of central tendency are the mean, median, and mode.
- **Dispersion**
- Measures of central tendency, like the mean, describe the most likely value, but they do not tell us anything about how the values vary.
- The most widely used measures of dispersion are range, variance, and standard deviation.

Bivariate analysis

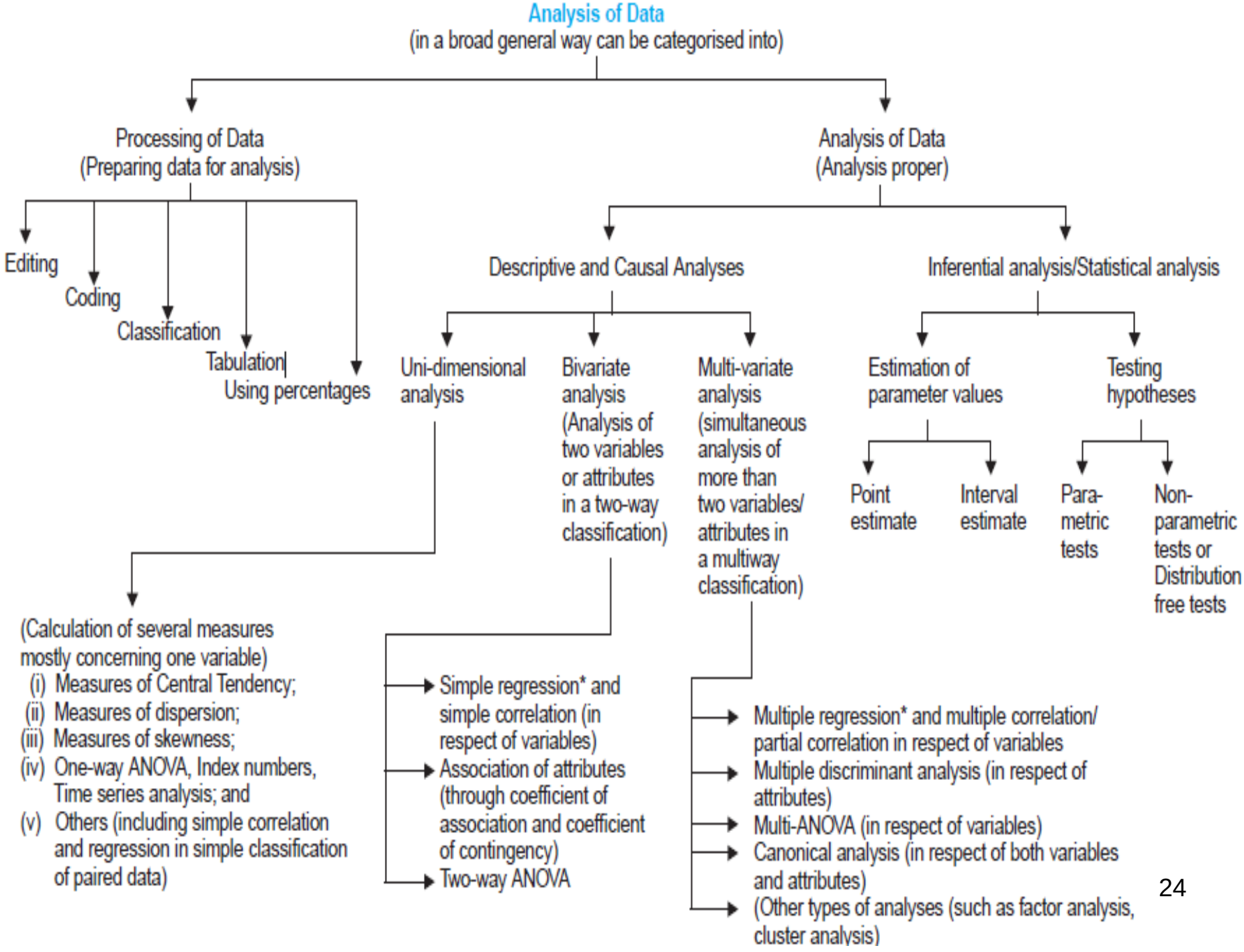
- Although frequency counts are a useful starting point in quantitative data analysis, you may find that you need to do more than merely describe your findings.
  - Often you will need to find out if there is a connection between one variable and a number of other variables.
  - Analysis of two variables/attributes.
  - For example, a researcher might want to find out whether there is a connection between watching violent films and aggressive behavior.

Multivariate Analysis

- The researcher is interested in exploring the connections among more than **two variables**.
  - For example, a researcher might be interested in finding out whether women aged 40–50, in professional occupations, are more likely to try complementary therapies than younger, non-professional women and men from all categories.
    - Professional women
    - Younger
    - Non professional women
    - Men

Inferential analysis

- In addition to describing and examining associations of variables within our data sets, we often conduct research to answer questions about the greater **population**.
  - Because it would not be feasible to collect data from the entire population, researchers conduct research with representative samples in an attempt to draw inferences about the populations from which the samples were drawn.
  - The analyses used to examine these inferences are appropriately referred to as *inferential statistics.*


Analysis of Data
(in a broad general way can be categorised into)
Processing of Data
(Preparing data for analysis)
Editing
Coding
Classification
Tabulation
Using percentages
Analysis of Data
(Analysis proper)
Descriptive and Causal Analyses
Inferential analysis/Statistical analysis
Uni-dimensional analysis
Bivariate analysis (Analysis of two variables or attributes in a two-way classification)
Multi-variate analysis (simultaneous analysis of more than two variables/ attributes in a multiway classification)
Estimation of parameter values
Testing hypotheses
Point estimate
Interval estimate
Para-metric tests
Non-parametric tests or Distribution free tests
(Calculation of several measures mostly concerning one variable)
(i) Measures of Central Tendency;
(ii) Measures of dispersion;
(iii) Measures of skewness;
(iv) One-way ANOVA, Index numbers, Time series analysis; and
(v) Others (including simple correlation and regression in simple classification of paired data)
Simple regression* and simple correlation (in respect of variables)
Association of attributes (through coefficient of association and coefficient of contingency)
Two-way ANOVA
Multiple regression* and multiple correlation/ partial correlation in respect of variables
Multiple discriminant analysis (in respect of attributes)
Multi-ANOVA (in respect of variables)
Canonical analysis (in respect of both variables and attributes)
(Other types of analyses (such as factor analysis, cluster analysis)

THANK YOU!!