

4

DATA PRESENTATION**Learning Objectives:***To Understand:*

- Data presentation and Analysis*
- Editing*
- Classification of Data*
- Data Entry*
- Tabulation*
- Types of Frequency Distributions*
- Introduction*
- Significance of Measuring Variation*
- Range*
- Quartile Deviation*
- Mean Deviation (m.d.)*
- Standard Deviation*
- Coefficient of Variation*
- Skewness and Kurtosis*

4.1 DATA PRESENTATION AND ANALYSIS

All business decisions are base upon evaluation of some data. Availability of right information is very important for making right decision. Once the data begin to flow in attention turns to data analysis.

It is almost impossible for the management to deal with all this data in the raw form. Such data must be presented in a suitable and summarised form without any loss of relevant information so that it can be efficiently used for decision making.

Whenever there is a need to present numerous figures or to describe a technical process or procedures, graphic aids can help communicate this information to your audience more quickly. The two graphic aids most used in research reports are tables and graphs. Besides making the report easier to read and understand, graphic aids improve its physical appearance.

4.2 EDITING

The first step in analysis is to edit the raw data. Editing detects errors and omissions, correct then when possible.

4.2.1 Purpose of Editing

1. Accurate,
2. Consistent with other information,
3. Uniformly entered,

4. Complete, and
5. Arranged to simplify coding and tabulation.

The editor's responsibility is to decide which of the responses is consistent with the intent of the question or other information in the survey and is most accurate for this individual respondent.

4.3 CLASSIFICATION OF DATA

When the raw data has been collected and edited it should be put into an ordered form (ascending or descending order) so that it can be looked at more objectively. The next important step towards processing the data is classification. Classification means separating items according to similar characteristics and grouping them into various classes. The items in different classes will differ from each other on the basis of some characteristics or attributes.

The classification of data into limited categories sacrifices some data detail but is necessary for efficient analysis.

4.3.1 Coding

Coding is the process whereby the data are assigned a numerical code a value so that the data can be more easily fitted into the appropriate categories. Coding is necessary where machine tabulation is used.

Coding Rules: Four rules guide for the categories:

1. Appropriate to the research problem and purpose
2. Exhaustive
3. Mutually exclusive
4. Derived from one classification principle.

4.4 DATA ENTRY

Data entry convert information gathered by a researcher through a primary and/or secondary method to a medium for viewer & manipulation.

4.5 TABULATION

Tabulation consists of counting the number of cases, which fall into the established categories. Before tabulation can take place the plan of analysis must be thought through. A tabulation plan which specified the precise counts to be obtained needs to be prepared. Usually the plan consists of setting up dummy tables completes with headlines and stubs plus a description of the data to be included. Tabulation may be accomplished using manual or machine methods. Both processes possess unique advantages and disadvantages.

Machine tabulation is preferable where a large number of questionnaires is involved and where cross tabulation are planned.

After the tabulation step has been completed it is necessary to determine whether any differences exist and whether these differences are statistically significant. If the results from a marketing study are to be useful they must be organised and summarised in such a way as to enable the objectives of the study to be achieved. This requires the use of those summary statistical measures, which typify the group, such as the arithmetic mean and percentages, and the use of those measures which indicate

dispersion, such as the standard deviation. All of these measures have their advantages and disadvantages, and the analyst must be cognisant of these in selection the most appropriate measure to use.

Further, since most marketing data are collected using a sample, the reliability of the summary estimates must be determined. The standard error and confidence limits are two terms and concepts, which are fundamental to the general subject of reliability.

To determine the significance of a difference existing between a sample estimate and a similar estimate from another from another sample requires the use of tests of significance. Such tests indicate the probability of the difference having a cause due to random sampling variations. All these tests make use of the standard error and confidence limit concepts.

The next step in the analysis procedure consists in trying to explain 'why' differences exist. This is accomplished by attempting to identify a causal relationship between two or more variables. To accomplish this the survey data are arranged to approximate to a controlled experiment. Causal analysis requires cross tabulations. This procedure introduces one or more additional variables into the analysis by dividing the sample into subgroups. This helps to confirm or reject earlier relationship. In introducing additional variables the analyst must consider the size of the sample in the various subgroups and also be on guard against spurious correlation.

Example 1: Construction of a frequency table for ungrouped data. Given below is the number of children in 30 families.

2, 1, 3, 1, 5, 2, 3, 4, 2, 1, 4, 3, 2, 2, 1, 3, 3, 2, 2, 1, 5, 4, 3, 3, 2, 1, 1, 3, 2

Solution: # 1: Arranging the data in Ascending order of magnitude: (An Array):

1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5

2: Frequency table may now be presented as under, using tally mark:

No. of Children	Tally Marks	Frequency (No. of Families)
1		7
2		9
3		9
4		3
5		2

Example 2: Present the given below data of marks obtained by 50 students out of 100 marks in an examination of statistics paper.

70, 65, 45, 10, 40, 50, 60, 55, 65, 25, 45, 85, 91, 73, 12, 23, 57, 61, 90, 72, 27, 17, 33, 93, 82, 38, 62, 47, 98, 37, 90, 99, 38, 98, 68, 58, 66, 38, 81, 62, 60, 45, 55, 65, 70, 42, 61, 39, 07, 87.

Solution #: Classify the marks into groups or classes as shown in the table given below: