



# Data Analysis and Report Writing

(BA (JMC) 303)

Unit – 4

by

**Dr. Sheel Nidhi Tripathi**  
(Associate Professor, BVICAM, New Delhi)

2023

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---

---

---

---

---

---

## LESSON 1 Processing of Data: Editing, Coding, Classification, Tabulation

- After collection, the data has to be analyzed in accordance with the outline laid down for the purpose at the time of developing the research plan.
- This is essential for a scientific study and for ensuring that the data collected is relevant for making contemplated comparisons and analysis. Processing of data refers to editing, coding, classification and tabulation of collected data so that they are amenable for analysis.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---

---

---

---

---

---

## Editing

- It is a process of examining the collected data to detect errors and omissions and to correct these when possible.
- As a matter of fact, editing involves a careful scrutiny of the completed questionnaires or schedules.
- Editing is done to assure that the data are accurate, consistent with other facts gathered, uniformly entered, as much completed as possible and have been well arranged to facilitate coding and tabulation.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---


---

---

---

---

---

 **Points to be remembered by the Editor**

- Should be familiar with instructions give to the interviewers and coders as well as with the editing instructions supplied to them for the purpose.
- While crossing out an original entry for one reason or another, they should just draw a single line on it so that the same may remain legible.
- They must make entries (if any) on the form in some distinctive colour and that too in a standardized form.
- They should initial all answers which they change or supply.
- Editor"s initials and the date of editing should be placed on each completed form.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---


---

---

---

---

---

 **Coding**

- Coding refers to the process of assigning numerals or other symbols to answers so that responses can be put into a limited number of categories or classes.
- Such classes should be appropriate to the research problem under consideration. They must also possess the characteristics of exhaustiveness ( i.e. there must be a class for every data item) and also that of mutual exclusiveness which means that a specific answer can be placed in one and only one cell in a given category set.
- Another rule to be observed is that of uni-dimensionality by which is meant that every class is defined in terms of only one concept.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---


---

---

---

---

---

 **Classification**

- Classification means to arrange data in groups or classes on the basis of common characteristics. Data having common characteristics are placed in one class and in this way the entire data gets divided into a number of groups or classes.

**Classification can be one of the following two types:**

- **Classification according to Attributes** Data can be classified on the basis of common characteristics which can either be descriptive (such as literacy, sex, honesty etc) or numerical (such as weight, height, income etc). Descriptive characteristics refer to qualitative phenomenon which cannot be measured quantitatively only their presence or absence in an individual item can be noticed. Data obtained on the basis of certain attributes are known as statistics of attributes.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---


---

---

---

---

---

 **Classification can be one of the following two types:**

- **Classification according to Class** – intervals Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomenon which can be measured through some statistical units. Data related to income, production, age, weight etc come under this category. Such data are known as statistics of variables and are classified on the basis of class intervals.

---

---

---

---

---

---

---

---

---

---

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

 **Tabulation**

- When a mass of data has been assembled, it becomes necessary for the researcher to arrange the same in some kind of concise and logical order. This process is called tabulation. Tabulation is the process of summarizing raw data and displaying the same in compact form for further analysis. Tabulation is an orderly arrangement of data in columns and rows. Tabulation is essential because of the following reasons:
- It conserves space and reduces explanatory and descriptive statement to a minimum. It facilitates the process of comparison.
- It facilitates the summation of items and the detection of error and omissions.

---

---

---

---

---

---

---

---

---

---

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

 **Tabulation**

- It provides a basis for various statistical comparisons.
- Tabulation can be done manually or mechanically. The choice depends on the size and type of study, cost considerations, time pressures and the availability of tabulating machines or computers.
- Tabulation may also be classified as simple and complex tabulation. Simple tabulation gives information about one or more groups of independent questions, whereas the complex tabulation shows the divisions of data in 2 or more sets of inter – related questions.

---

---

---

---

---

---

---

---

---

---

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

## Sample of Tabulation

- The sample explains how tabulation is done for every coded & classified question. The variable used is age.
- The researcher has to first represent the number of respondents who agree to have bought or downloaded pirated content. In the statistical representation the researcher must denote each 'yes' in a line and count five in one set.
- All the sets are then counted to generate the total number. In case of even number, like 14, the stand alone number must also be counted. After the statistical representation, the researcher then gives the numerical value. This data is then represented through graphs and charts.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---

---

---

---

---

---

---

---

## Sample of Tabulation

S. No.	Age	Statistical Representation	Numerical	
			Yes	No
1.	18-25		20	0
2.	26-50		14	2

**Code A: Table 1: No. of respondents who have bought or downloaded pirated content**

The sample explains how tabulation is done for every coded & classified question. The variable used is age. The researcher has to first represent the number of respondents who agree to have bought or downloaded pirated content. In the statistical representation the researcher must denote each 'yes' in a line and count five in one set. All the sets are then counted to generate the total number. In case of even number, like 14, the stand alone number must also be counted. After the statistical representation, the researcher then gives the numerical value. This data is then represented through graphs and charts.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---

---

---

---

---

---

---

---

## Problem in Data Processing

- The problem concerning **DK Response "don't know" responses**: While processing the data, the researcher often comes across some responses that are difficult to handle. One category of such responses may be "don't know response" or simply DK Response. When the DK Response group is small, it is of little significance. But when it is relatively big, it becomes a matter of major concern in which case the question arises: is the question which elicited DK Response useless? The answer depends upon two points i.e. the respondent actually may not know the answer or the researcher may fail in obtaining the appropriate information. In the 1st case, the concerned question is said to be alright and DK Response is taken as legitimate DK Response. But in the second case, DK Response is more likely to be a failure of the questioning process.

© Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi-63, by Dr. Sheel Nidhi Tripathi

---

---

---

---

---

---

---

---

---

---

---

---



## Problem in Data Processing

- The researcher must design better type of questions to deal with DK Responses. Good rapport of the interviewer with the respondents will result in minimizing DK responses. To tackle DK Responses which have already taken place is to estimate the allocation of DK answers from other data in the questionnaire. The other way is to keep DK Responses as a separate category in tabulation where we can consider it as a separate reply category if DK Responses happen to be legitimate, otherwise we should let the reader make his own decision.

---

---

---

---

---

---

---

---