www.prepnext.com

Notes Prepared By:

RAKESH AGARWAL

M.Com, MBA, FIII
E-mail: messagerakesh@gmail.com
WhatsApp No: 8486118428

Coaching Available for: *Bank P.O./ Clerk, SSC, Railway, MAT, CA, CS, B.Com and M.Com. For details, call 8486118428 or email at info@prepnext.com*

Basics of Academic Project Preparation

Miscellaneous - 2

Q: What is data processing? Evaluate the steps involved in processing data. *(www.prepnext.com)*

Ans:

Once the data are collected for any research project/research work, the next step is to process the date according to research plan. This is an important stage that acts as an intermediary between data collection and their analysis and interpretation. Data processing converts the raw data to the form that is suitable for the required analysis. It involves the following steps:

Please WhatsApp your suggestions/ feedback at: 8486118428

1. Questionnaire Checking
2. Editing
3. Coding
4. Classification
5. Tabulation
6. Graphical Representation
7. Data Cleaning

1. Questionnaire Checking:
When the data is collected through questionnaires, the first step of data processing process is to check the questionnaires if they are acceptable or not. This involves the examination of all questionnaires for their completeness and interviewing quality. A questionnaire may not be acceptable if:

(i)   It is incomplete partially or fully.
(ii) It is answered by a person who has inadequate knowledge or does not qualify for the participation.
(iii) It is answered in such a way which gives the impression that the respondent could not understand the questions.

2. Editing:
Editing of data is a process of examining the collected raw data (specially in surveys) 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 and/or schedules. Editing is done to assure that the data are accurate, consistent with other facts gathered, uniformly entered, as completed as possible and have been well arranged to facilitate coding and tabulation.

Editing, basically, can be done in two ways: **field editing** and **central editing.**
Field editing consists of the review of the reporting forms by the investigator for completing (translating or rewriting) what the latter has written in abbreviated and/or in illegible form at the

time of recording the respondents' responses. While doing field editing, the investigator must restrain himself and must not correct errors of omission by simply guessing what the informant would have said if the question had been asked.

Central editing should take place when all forms or schedules have been completed and returned to the office. This type of editing implies that all forms should get a thorough editing by a single editor in a small study and by a team of editors in case of a large inquiry. Editor(s) may correct the obvious errors such as an entry in the wrong place, entry recorded in months when it should have been recorded in weeks, and the like. In case of inappropriate or missing replies, the editor can sometimes determine the proper answer by reviewing the other information in the schedule. At times, the respondent can be contacted for clarification. The editor must strike out the answer if the same is inappropriate and he has no basis for determining the correct answer or the response. In such a case an editing entry of 'no answer' is called for. All the wrong replies, which are quite obvious, must be dropped from the final results.

3. **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 characteristic of exhaustiveness (i.e., there must be a class for every data item) and also that of mutual exclusively 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 unidimensionality by which is meant that every class is defined in terms of only one concept. Coding is necessary for efficient analysis and through it the several replies may be reduced to a small number of classes which contain the critical information required for analysis.

**Please WhatsApp your suggestions/ feedback at*: 8486118428*

4. Classification:

Most research studies result in a large volume of raw data which must be reduced into homogeneous groups if we are to get meaningful relationships. This fact necessitates classification of data which happens to be the process of arranging data in groups or classes on the basis of common characteristics. Data having a common characteristic are placed in one class and in this way the entire data get divided into a number of groups or classes.

Classification can be one of the following two types, depending upon the nature of the phenomenon involved:

(a) Classification according to attributes: Descriptive characteristics such as literacy, sex, honesty, etc. cannot be measured quantitatively; only their presence or absence in an individual item can be noticed. Data obtained this way on the basis of certain attributes are known as **statistics of attributes** and their classification is said to be classification according to attributes.

Such classification can be **simple classification** or **manifold classification**. In **simple classification** we consider only one attribute and divide the universe into two classes – one class consisting of items possessing the given attribute and the other class consisting of items which do not possess the given attribute. But in **manifold classification** we consider two or more attributes simultaneously, and divide that data into a number of classes.

(b) Classification according to class-intervals: Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomenon which can be measured through some statistical units. Data relating 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. For instance, persons whose incomes, say, are within Rs.5001 to

Rs.10000 can form one group, those whose incomes are within Rs.10001 to Rs.20000 can form another group and so on. In this way the entire data may be divided into a number of groups or classes or what are usually called, 'class-intervals.'

5. 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 procedure is referred to as tabulation. Thus, tabulation is the process of summarising raw data and displaying the same in compact form (i.e., in the form of statistical tables) for further analysis. In a broader sense, tabulation is an orderly arrangement of data in columns and rows.

Tabulation is essential because of the following reasons.
1. It conserves space and reduces explanatory and descriptive statement to a minimum.
2. It facilitates the process of comparison.
3. It facilitates the summation of items and the detection of errors and omissions.
4. It provides a basis for various statistical computations.

Generally Accepted Principles of Tabulation:
Such principles of tabulation can be briefly stated as follows:
1. Every table should have a clear, concise and adequate title so as to make the table intelligible without reference to the text and this title should always be placed just above the body of the table.
2. Every table should be given a distinct number to facilitate easy reference.
3. The column headings (captions) and the row headings (stubs) of the table should be clear and brief.
4. The units of measurement under each heading or sub-heading must always be indicated.

5. Explanatory footnotes, if any, concerning the table should be placed directly beneath the table, along with the reference symbols used in the table.

6. Source or sources from where the data in the table have been obtained must be indicated just below the table.

7. Usually the columns are separated from one another by lines which make the table more readable and attractive.

8. Lines are always drawn at the top and bottom of the table and below the captions.

9. There should be thick lines to separate the data under one class from the data under another class and the lines separating the sub-divisions of the classes should be comparatively thin lines.

10. The columns may be numbered to facilitate reference.

11. Those columns whose data are to be compared should be kept side by side. Similarly, percentages and/or averages must also be kept close to the data.

12. It is generally considered better to approximate figures before tabulation as the same would reduce unnecessary details in the table itself.

13. In order to emphasis the relative significance of certain categories, different kinds of type, spacing and indentations may be used.

14. It is important that all column figures be properly aligned. Decimal points and (+) or (-) signs should be in perfect alignment

15. Abbreviations should be avoided to the extent possible and ditto marks should not be used in the table.

16. Miscellaneous and exceptional items, if any, should be usually placed in the last row of the table.

17. Table should be made as logical, clear, accurate and simple as possible. If the data happen to be very large, they should not be crowded in a single table for that would make the table unwieldy and inconvenient

**Please WhatsApp your suggestions/ feedback at*: 8486118428*

18. Total of rows should normally be placed in the extreme right column and that of columns should be placed at the bottom.

19. The arrangement of the categories in a table may be chronological, geographical, alphabetical or according to magnitude to facilitate comparison. Above all, the table must suit the needs and requirements of an investigation.

6. Graphical Representation:

Graphs help to understand the data easily. Most common graphs are bar charts, histograms and pie charts.

7. Data Cleaning:

This includes checking the data for consistency and treatment for missing value. Consistency checks look for the data which are not consistent or outlines. Such data may either be discarded or replaced by the mean value. However, the researcher should be careful while doing this. Extreme values or outlines are not always erroneous.

Missing values are the values which are unknown or not answered by the respondent. In place of such missing values, some neutral value may be used. This neutral value may be the mean of available values. The other option could be to use the pattern of responses to other questions to calculate a suitable substitute to the missing values.

Please WhatsApp your suggestions/ feedback at*: 8486118428*

DATA INTERPRETATION

It refers to drawing inferences from the collected facts after an analytical or experimental study. It is only through interpretation that a researcher is able to explain the relationship and processes underlying his findings. In simple words, an interpretation is an explanation. It is a means to present information in terms that are understandable.

FOCUS OF INTERPRETATION:

1) the researcher should try to establish continuity in his research through linking the results of his study with those of the others.
2) He should try to establish some explanatory concepts.

IMPORTANCE OF INTERPRETATION:

Following functions are performed by interpretation:

1) It enables a researcher in understanding the abstract principles, working beneath his findings.

2) It enables the researcher to link his findings with the findings of other studies that are having the same abstract principles.

3) Interpreting the findings of explanatory research study many a times results into hypotheses for experimental research

4) Interpretation results into establishment of explanatory concepts. These concepts often serve as a guide for future studies.

5) Interpretation opens up new avenues of intellectual adventure and journey.

6) It motivates and encourages the researcher, and stimulates his quest for more knowledge.

7) It enables the researcher to understand his research findings.

8) It can make others understand the significance of the research findings.

REPORT WRITING

Reporting refers to the communication function of the findings of the research to the target audience following the standards of communication practices specially designed for.

Writing report is the last step in a research study. It consists of the following steps:

STEPS:

1) Logical Analysis of the subject matter: It is related with the development of the subject. For this, there are two different routes:
(a) Logical Development : It is made on the basis of mental associations and connections between one thing and another by way of analysis. It moves from simple to complex structure.
(b) Chronological Development: It is made on the basis of time based connections or sequences in time or occurrences.

2) Preparation of Final Outline: Outline refers to the framework on which long written works are built or constructed. This provides the reader logical connections and reminder of points not to be overlooked.

3) Preparation of Rough Draft: Here the researcher writes as it comes to him in the context of the research study.

4) Rewriting and Polishing: Here the focus is on correcting and improving the weakness in logical developments and presentations. This step is more time consuming than preparation of rough draft.

5) Preparation of Final Bibliography: The researcher should arrange the Bibliography in an alphabetical order. Bibliography is basically divided into two parts:

Please WhatsApp your suggestions/ feedback at*: 8486118428*

(i) Names of books and pamphlets. (It should be mentioned in the following order)

    (a) Name of the author, last name first

    (b) Title

    (c) Place, Publisher and date of Publication

    (d) Number of Volume

(ii) Magazines and newspaper articles (It should be mentioned in the following order)

    (a) Name of the author, last name first

    (b) Title

    (c) Name of the periodical

    (d) Volume and Number

    (e) The date of Issue

    (f) The Pagination

6) Writing the Final Draft:

The final draft should be written in concise and objective style. Simple sentences should be used. Use of abstract terminology, technical words and jargons, vague expressions such as 'it appears', "there may be', etc must be avoided. In short, the report should be drafted in such a manner that it adds to existing knowledge.

CONTENT OR LAYOUT OF THE RESEARCH REPORT

There are well – designed rules and guidelines for the content or layout of the research report. A research report includes:

I) Preliminary pages:

Preliminary pages include a title and date, acknowledgement in the form of foreword or preface, table of contents, list of tables and illustrations.

Please WhatsApp your suggestions/ feedback at*: 8486118428*

II) Main Text:

It includes the following sections:

(a) Introduction

(b) Statement of findings and recommendations

(c) Result

(d) Implications drawn from the results

(e) Summary

III) THE END MATTER:

It includes the appendices related to questionnaire, bibliography, sample information mathematical derivations, etc.

**Please WhatsApp your suggestions/ feedback at*: 8486118428*