SRIMAAN COACHING CENTRE-TRICHY-PG-TRB-ZOOLOGY-UNIT-III-COLLECTION OF DATA
TO CONTACT: 8072230063.

2021-22
SRIMAAN

SRIMAAN

PG-TRB

ZOOLOGY

UNIT-III-(ENGLISH MEDIUM)
STUDY MATERIAL

SRIMAAN COACHING CENTRE-TRICHY.

TO CONTACT:8072230063.

PG-TRB:ZOOLOGY

UNIT-III

COLLECTION OF DATA, CLASSIFICATION AND TABULATION

Introduction:

Everybody collects, interprets and uses information, much of it in a numerical or statistical forms in day-to-day life. It is a common practice that people receive large quantities of information everyday through conversations, televisions, computers, the radios, newspapers, posters, notices and instructions. It is just because there is so much information available that people need to be able to absorb, select and reject it. In everyday life, in business and industry, certain statistical information is necessary and it is independent to know where to find it how to collect it. As consequences, everybody has to compare prices and quality before making any decision about what goods to buy. As employees of any firm, people want to compare their salaries and working conditions, promotion opportunities and so on. In time the firms on their part want to controlcosts and expand their profits.

One of the main functions of statistics is to provide information which will help on making decisions. Statistics provides the type of information by providing a description of the present, a profile of the past and an estimate of the future. The following are some of the objectives of collecting statistical information.

1. To describe the methods of collecting primary statistical information.
2. To consider the status involved in carrying out a survey.
3. To analyse the process involved in observation and interpreting.
4. To define and describe sampling.
5. To analyse the basis of sampling.
6. To describe a variety of sampling methods.

Statistical investigation is a comprehensive and requires systematic collection of data about some group of people or objects, describing and organizing the data, analyzing the data with the help of different statistical method, summarizing the analysis and using these results for making judgements, decisions and predictions. The validity and accuracy of final judgement is most crucial and depends heavily on how well the data was collected in the first place. The quality of data will greatly affect the conditions and hence at most importance must be given to this process and every possible precautions should be taken to ensure accuracy while collecting the data.

Nature of data:

It may be noted that different types of data can be collected for different purposes. The data can be collected in connection with time or geographical location or in connection with time and location. The following are the three types of data:

1. Time series data.
2. Spatial data
3. Spacio-temporal data.

Time series data:

It is a collection of a set of numerical values, collected over a period of time. The data might have been collected either at regular intervals of time or irregular intervals of time.

Example 1:

The following is the data for the three types of expenditures in rupees for a family for the four years 2001,2002,2003,2004.

Year	Food	Education	Others	Total
2001	3000	2000	3000	8000
2002	3500	3000	4000	10500
2003	4000	3500	5000	12500
2004	5000	5000	6000	16000

Spatial Data:

If the data collected is connected with that of a place, then it is termed as spatial data. For example, the data may be

1. Number of runs scored by a batsman in different test matches in a test series at different places
2. District wise rainfall in Tamilnadu
3. Prices of silver in four metropolitan cities

Example 2:

The population of the southern states of India in 1991.

State	Population
Tamilnadu	5,56,38,318
Andhra Pradesh	6,63,04,854
Karnataka	4,48,17,398
Kerala	2,90,11,237
Pondicherry	7,89,416

Spacio Temporal Data:

If the data collected is connected to the time as well as place then it is known as spacio temporal data.

Example 3:

State	Population	
	1981	1991
Tamil Nadu	4,82,97,456	5,56,38,318
Andhra Pradesh	5,34,03,619	6,63,04,854
Karnataka	3,70,43,451	4,48,17,398
Kerala	2,54,03,217	2,90,11,237
Pondicherry	6,04,136	7,89,416

Categories of data:

Any statistical data can be classified under two categories depending upon the sources utilized.

These categories are,

1. Primary data   2. Secondary data

Primary data:

Primary data is the one, which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by survey conducted by individuals or research institution or any organisation.

Example 4:

If a researcher is interested to know the impact of noon-meal scheme for the school children, he has to undertake a survey and collect data on the opinion of parents and children by asking relevant questions. Such a data collected for the purpose is called primary data.

The primary data can be collected by the following five methods.

1. Direct personal interviews.
2. Indirect Oral interviews.
3. Information from correspondents.
4. Mailed questionnaire method.
5. Schedules sent through enumerators.

1. Direct personal interviews:

The persons from whom informations are collected are known as informants. The investigator personally meets them and asks questions to gather the necessary informations. It is the suitable method for intensive rather than extensive field surveys. It suits best for intensive study of the limited field.

Merits:

1. People willingly supply informations because they are approached personally. Hence, more response noticed in this method than in any other method.
2. The collected informations are likely to be uniform and accurate. The investigator is there to clear the doubts of the informants.
3. Supplementary informations on informant's personal aspects can be noted. Informations on character and environment may help later to interpret some of the results.
4. Answers for questions about which the informant is likely to be sensitive can be gathered by this method.
5. The wordings in one or more questions can be altered to suit any informant. Explanations may be given in other languages also. Inconvenience and misinterpretations are thereby avoided.

Limitations:

1. It is very costly and time consuming.
2. It is very difficult, when the number of persons to be interviewed is large and the persons are spread over a wide area.
3. Personal prejudice and bias are greater under this method.

2. Indirect Oral Interviews:

Under this method the investigator contacts witnesses or neighbours or friends or some other third parties who are capable of supplying the necessary information. This method is preferred if the required information is on addiction or cause of fire or theft or murder etc., If a fire has broken out a certain place, the persons living in neighbourhood and witnesses are likely to give information on the cause of fire. In some cases, police interrogated third parties who are supposed to have knowledge of a theft or a murder and get some clues. Enquiry committees appointed by governments generally adopt this method and get people' s views and all possible details of facts relating to the enquiry. This method is suitable whenever direct sources do not exists or cannot be relied upon or would be unwilling to part with the information.

The validity of the results depends upon a few factors, such as the nature of the person whose evidence is being recorded, the ability of the interviewer to draw out information from the third parties by means of appropriate questions and cross examinations, and the number of persons interviewed. For the success of this method one person or one group alone should not be relied upon.

3. Information from correspondents:

The investigator appoints local agents or correspondents in different places and compiles the information sent by them. Informations to Newspapers and some departments of Government come by this method. The advantage of this method is that it is cheap and appropriate for extensive investigations. But it may not ensure accurate results because the correspondents are likely to be negligent, prejudiced and biased. This method is adopted in those cases where informations are to be collected periodically from a wide area for a long time.

4. Mailed questionnaire method:

Under this method a list of questions is prepared and is sent to all the informants by post. The list of questions is technically called questionnaire. A covering letter accompanying the questionnaire explains the purpose of the investigation and the importance of correct informations and request the informants to fill in the blank spaces provided and to return the form within a specified time. This method is appropriate in those cases where the informants are literates and are spread over a wide area.

Merits:

1. It is relatively cheap.
2. It is preferable when the informants are spread over the wide area.

Limitations:

1. The greatest limitation is that the informants should be literates who are able to understand and reply the questions.
2. It is possible that some of the persons who receive the questionnaires do not return them.
3. It is difficult to verify the correctness of the informations furnished by the respondents.

With the view of minimizing non-respondents and collecting correct information, the questionnaire should be carefully drafted. There is no hard and fast rule. But the following general principles may be helpful in framing the questionnaire. A covering letter and a self addressed and stamped envelope should accompany the questionnaire. The covering letter should politely point out the purpose of the survey and privilege of the respondent who is one among the few associated with the investigation. It should assure that the informations would be kept confidential and would never be misused. It may promise a copy of the findings or free gifts or concessions etc.,

Characteristics of a good questionnaire:

1. Number of questions should be minimum.
2. Questions should be in logical orders, moving from easy to more difficult questions.

3. Questions should be short and simple. Technical terms and vague expressions capable of different interpretations should be avoided.
4. Questions fetching YES or NO answers are preferable. There may be some multiple choice questions requiring lengthy answers are to be avoided.
5. Personal questions and questions which require memory power and calculations should also be avoided.
6. Question should enable cross check. Deliberate or unconscious mistakes can be detected to an extent.
7. Questions should be carefully framed so as to cover the entire scope of the survey.
8. The wording of the questions should be proper without hurting the feelings or arousing resentment.
9. As far as possible confidential informations should not be sought.
10. Physical appearance should be attractive, sufficient space should be provided for answering each questions.

5. Schedules sent through Enumerators:

Under this method enumerators or interviewers take the schedules, meet the informants and filling their replies. Often distinction is made between the schedule and a questionnaire. A schedule is filled by the interviewers in a face-to-face situation with the informant. A questionnaire is filled by the informant which he receives and returns by post. It is suitable for extensive surveys.

Merits:

1. It can be adopted even if the informants are illiterates.
2. Answers for questions of personal and pecuniary nature can be collected.
3. Non-response is minimum as enumerators go personally and contact the informants.
4. The informations collected are reliable. The enumerators can be properly trained for the same.
5. It is most popular methods. **Limitations:**
1. It is the costliest method.

2. Extensive training is to be given to the enumerators for collecting correct and uniform informations.
3. Interviewing requires experience. Unskilled investigators are likely to fail in their work.

Before the actual survey, a pilot survey is conducted. The questionnaire/Schedule is pre-tested in a pilot survey. A few among the people from whom actual information is needed are asked to reply. If they misunderstand a question or find it difficult to answer or do not like its wordings etc., it is to be altered. Further it is to be ensured that every questions fetches the desired answer.

Merits and Demerits of primary data:

1. The collection of data by the method of personal survey is possible only if the area covered by the investigator is small. Collection of data by sending the enumerator is bound to be expensive. Care should be taken twice that the enumerator record correct information provided by the informants.
2. Collection of primary data by framing a schedules or distributing and collecting questionnaires by post is less expensive and can be completed in shorter time.
3. Suppose the questions are embarrassing or of complicated nature or the questions probe into personnel affairs of individuals, then the schedules may not be filled with accurate and correct information and hence this method is unsuitable.
4. The information collected for primary data is mere reliable than those collected from the secondary data.

Secondary Data:

Secondary data are those data which have been already collected and analysed by some earlier agency for its own use; and later the same data are used by a different agency. According to W.A.Neiswanger, ' A primary source is a publication in which the data are published by the same authority which gathered and analysed them. A secondary source is a publication, reporting the data which have been gathered by other authorities and for which others are responsible' .

Sources of Secondary data:

In most of the studies the investigator finds it impracticable to collect first-hand information on all related issues and as such he makes use of the data collected by others. There is a vast amount of published information from which statistical studies may be made and fresh statistics are constantly in a state of production. The sources of secondary data can broadly be classified under two heads:

1. Published sources, and
2. Unpublished sources.

1. Published Sources:

The various sources of published data are:

1. Reports and official publications of
   (i) International bodies such as the International Monetary Fund, International Finance Corporation and United Nations Organisation.
   (ii) Central and State Governments such as the Report of the Tandon Committee and Pay Commission.
2. Semi-official publication of various local bodies such as Municipal Corporations and District Boards.
3. Private publications-such as the publications of –
   (i) Trade and professional bodies such as the Federation of Indian Chambers of Commerce and Institute of Chartered Accountants.
   (ii) Financial and economic journals such as ‘ Commerce’ , ‘ Capital’ and ‘ Indian Finance’ .
   (iii) Annual reports of joint stock companies.
   (iv) Publications brought out by research agencies, research scholars, etc.

It should be noted that the publications mentioned above vary with regard to the periodically of publication. Some are published at regular intervals (yearly, monthly, weekly etc.,) whereas others are ad hoc publications, i.e., with no regularity about periodicity of publications.

Note: A lot of secondary data is available in the internet. We can access it at any time for the further studies.

2. Unpublished Sources

All statistical material is not always published. There are various sources of unpublished data such as records maintained by various Government and private offices, studies made by research institutions, scholars, etc. Such sources can also be used where necessary

Precautions in the use of Secondary data

The following are some of the points that are to be considered in the use of secondary data

1. How the data has been collected and processed
2. The accuracy of the data
3. How far the data has been summarized
4. How comparable the data is with other tabulations
5. How to interpret the data, especially when figures collected for one purpose is used for another

Generally speaking, with secondary data, people have to compromise between what they want and what they are able to find.

Merits and Demerits of Secondary Data:

1. Secondary data is cheap to obtain. Many government publications are relatively cheap and libraries stock quantities of secondary data produced by the government, by companies and other organisations.
2. Large quantities of secondary data can be got through internet.
3. Much of the secondary data available has been collected for many years and therefore it can be used to plot trends.
4. Secondary data is of value to:
   - The government – help in making decisions and planning future policy.
   - Business and industry – in areas such as marketing, and sales in order to appreciate the general economic and social conditions and to provide information on competitors.
   - Research organisations – by providing social, economical and industrial information.

Classification:

The collected data, also known as raw data or ungrouped data are always in an un organised form and need to be organised and presented in meaningful and readily comprehensible form in order to facilitate further statistical analysis. It is, therefore, essential for an investigator to condense a mass of data into more and more comprehensible and assimilable form. The process of grouping into different classes or sub classes according to some characteristics is known as classification, tabulation is concerned with the systematic arrangement and presentation of classified data. Thus classification is the first step in tabulation.

For Example, letters in the post office are classified according to their destinations viz., Delhi, Madurai, Bangalore, Mumbai etc.,

Objects of Classification:

The following are main objectives of classifying the data:

1. It condenses the mass of data in an easily assimilable form.
2. It eliminates unnecessary details.
3. It facilitates comparison and highlights the significant aspect of data.
4. It enables one to get a mental picture of the information and helps in drawing inferences.
5. It helps in the statistical treatment of the information collected.

Types of classification:

Statistical data are classified in respect of their characteristics. Broadly there are four basic types of classification namely

a) Chronological classification
b) Geographical classification
c) Qualitative classification
d) Quantitative classification

a) Chronological classification:

In chronological classification the collected data are arranged according to the order of time expressed in years, months, weeks, etc., The data is generally classified in ascending order of time. For example, the data related with population, sales of a firm, imports and exports of a country are always subjected to chronological classification.

Example 5:

The estimates of birth rates in India during 1970 – 76 are

Year	1970	1971	1972	1973	1974	1975	1976
Birth Rate	36.8	36.9	36.6	34.6	34.5	35.2	34.2

b) Geographical classification:

In this type of classification the data are classified according to geographical region or place. For instance, the production of paddy in different states in India, production of wheat in different countries etc.,

Example 6:

Country	America	China	Denmark	France	India
Yield of wheat in (kg/acre)	1925	893	225	439	862

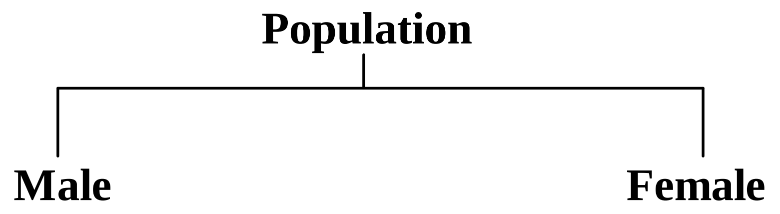
c) Qualitative classification:

In this type of classification data are classified on the basis of same attributes or quality like sex, literacy, religion, employment etc., Such attributes cannot be measured along with a scale.

For example, if the population to be classified in respect to one attribute, say sex, then we can classify them into two namely that of males and females. Similarly, they can also be classified into 'employed' or 'unemployed' on the basis of another attribute 'employment'.

Thus when the classification is done with respect to one attribute, which is dichotomous in nature, two classes are formed, one possessing the attribute and the other not possessing the attribute. This type of classification is called simple or dichotomous classification.

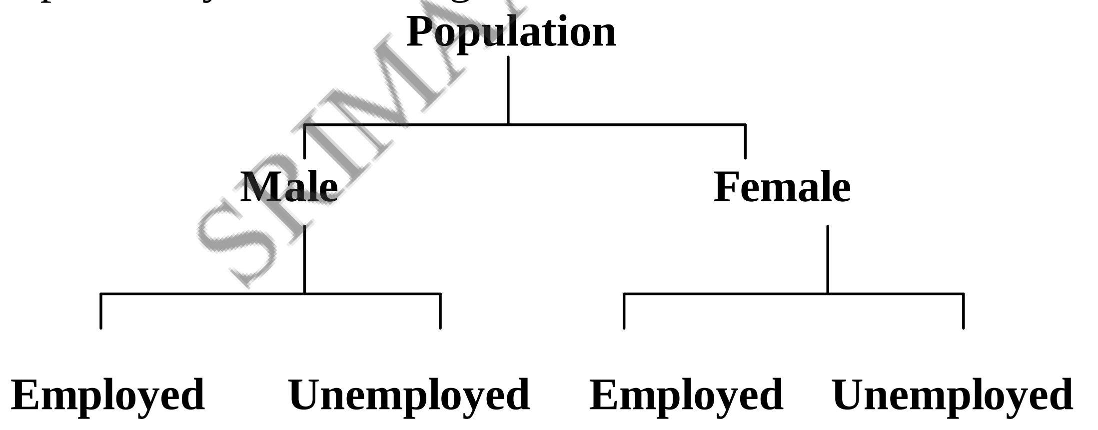
A simple classification may be shown as under



The classification, where two or more attributes are considered and several classes are formed, is called a manifold classification. For example, if we classify population simultaneously with respect to two attributes, e.g sex and employment, then population are first classified with respect to ‘sex’ into ‘males’ and ‘females’. Each of these classes may then be further classified into ‘employment’ and ‘unemployment’ on the basis of attribute ‘employment’ and as such Population are classified into four classes namely.

(i) Male employed
(ii) Male unemployed
(iii) Female employed
(iv) Female unemployed

Still the classification may be further extended by considering other attributes like marital status etc. This can be explained by the following chart



d) Quantitative classification:

Quantitative classification refers to the classification of data according to some characteristics that can be measured such as height, weight, etc., For example the students of a college may be classified according to weight as given below.

Weight (in lbs)	No of Students
90-100	50
100-110	200
110-120	260
120-130	360
130-140	90
140-150	40
Total	1000

In this type of classification there are two elements, namely (i) the variable (i.e) the weight in the above example, and (ii) the frequency in the number of students in each class. There are 50 students having weights ranging from 90 to 100 lb, 200 students having weight ranging between 100 to 110 lb and so on.

Tabulation:

Tabulation is the process of summarizing classified or grouped data in the form of a table so that it is easily understood and an investigator is quickly able to locate the desired information. A table is a systematic arrangement of classified data in columns and rows. Thus, a statistical table makes it possible for the investigator to present a huge mass of data in a detailed and orderly form. It facilitates comparison and often reveals certain patterns in data which are otherwise not obvious.Classification and ' Tabulation' , as a matter of fact, are not two distinct processes. Actually they go together. Before tabulation data are classified and then displayed under different columns and rows of a table.

Advantages of Tabulation:

Statistical data arranged in a tabular form serve following objectives:

1. It simplifies complex data and the data presented are easily understood.
2. It facilitates comparison of related facts.
3. It facilitates computation of various statistical measures like averages, dispersion, correlation etc.

4. It presents facts in minimum possible space and unnecessary repetitions and explanations are avoided. Moreover, the needed information can be easily located.
5. Tabulated data are good for references and they make it easier to present the information in the form of graphs and diagrams.

Preparing a Table:

The making of a compact table itself an art. This should contain all the information needed within the smallest possible space. What the purpose of tabulation is and how the tabulated information is to be used are the main points to be kept in mind while preparing for a statistical table. An ideal table should consist of the following main parts:

1. Table number
2. Title of the table
3. Captions or column headings
4. Stubs or row designation
5. Body of the table
6. Footnotes
7. Sources of data

Table Number:

A table should be numbered for easy reference and identification. This number, if possible, should be written in the centre at the top of the table. Sometimes it is also written just before the title of the table.

Title:

A good table should have a clearly worded, brief but unambiguous title explaining the nature of data contained in the table. It should also state arrangement of data and the period covered. The title should be placed centrally on the top of a table just below the table number (or just after table number in the same line).

Captions or column Headings:

Captions in a table stands for brief and self explanatory headings of vertical columns. Captions may involve headings and sub-headings as well. The unit of data contained should also be given for each column. Usually, a relatively less important and shorter classification should be tabulated in the columns.

Stubs or Row Designations:

Stubs stands for brief and self explanatory headings of horizontal rows. Normally, a relatively more important classification is given in rows. Also a variable with a large number of classes is usually represented in rows. For example, rows may stand for score of classes and columns for data related to sex of students. In the process, there will be many rows for scores classes but only two columns for male and female students.

A model structure of a table is given below:

Table Number **Title of the Table**

Sub Heading	Caption Headings	Total
	Caption Sub-Headings	
Stub Sub- Headings	Body	
Total		

Foot notes:
Sources Note:

Body:

The body of the table contains the numerical information of frequency of observations in the different cells. This arrangement of data is according to the discription of captions and stubs.

Footnotes:

Footnotes are given at the foot of the table for explanation of any fact or information included in the table which needs some explanation. Thus, they are meant for explaining or providing further details about the data, that have not been covered in title, captions and stubs.

Sources of data:

Lastly one should also mention the source of information from which data are taken. This may preferably include the name of the author, volume, page and the year of publication. This should also state whether the data contained in the table is of ‘ primary or secondary’ nature.

Requirements of a Good Table:

A good statistical table is not merely a careless grouping of columns and rows but should be such that it summarizes the total information in an easily accessible form in minimum possible space. Thus while preparing a table, one must have a clear idea of the information to be presented, the facts to be compared and he points to be stressed.

Though, there is no hard and fast rule for forming a table yet a few general point should be kept in mind:

1. A table should be formed in keeping with the objects of statistical enquiry.
2. A table should be carefully prepared so that it is easily understandable.
3. A table should be formed so as to suit the size of the paper. But such an adjustment should not be at the cost of legibility.
4. If the figures in the table are large, they should be suitably rounded or approximated. The method of approximation and units of measurements too should be specified.

5. Rows and columns in a table should be numbered and certain figures to be stressed may be put in ' box' or ' circle' or in bold letters.
6. The arrangements of rows and columns should be in a logical and systematic order. This arrangement may be alphabetical, chronological or according to size.
7. The rows and columns are separated by single, double or thick lines to represent various classes and sub-classes used. The corresponding proportions or percentages should be given in adjoining rows and columns to enable comparison. A vertical expansion of the table is generally more convenient than the horizontal one.
8. The averages or totals of different rows should be given at the right of the table and that of columns at the bottom of the table. Totals for every sub-class too should be mentioned.
9. In case it is not possible to accommodate all the information in a single table, it is better to have two or more related tables.

Type of Tables:

Tables can be classified according to their purpose, stage of enquiry, nature of data or number of characteristics used. On the basis of the number of characteristics, tables may be classified as follows:

1. Simple or one-way table
2. Two way table
3. Manifold table

Simple or one-way Table:

A simple or one-way table is the simplest table which contains data of one characteristic only. A simple table is easy to construct and simple to follow. For example, the blank table given below may be used to show the number of adults in different occupations in a locality.

The number of adults in different occupations in a locality

Occupations	No. Of Adults
Total	

Two-way Table:

A table, which contains data on two characteristics, is called a two-way table. In such case, therefore, either stub or caption is divided into two co-ordinate parts. In the given table, as an example the caption may be further divided in respect of ‘ sex’ . This subdivision is shown in two-way table, which now contains two characteristics namely, occupation and sex.

The umber of adults in a locality in respect of occupation and sex

Occupation	**No. of Adults**		**Total**
	Male	**Female**	
Total			

Manifold Table:

Thus, more and more complex tables can be formed by including other characteristics. For example, we may further classify the caption sub-headings in the above table in respect of “marital status”, “ religion” and “socio-economic status” etc. A table ,which has more than two characteristics of data is considered as a manifold table. For instance , table shown below shows three characteristics namely, occupation, sex and marital status.

Occupation	No. of Adults						Total
	Male			Female			
	M	**U**	**Total**	**M**	**U**	**Total**	
Total							

Foot note: M Stands for Married and U stands for unmarried.

Manifold tables, though complex are good in practice as these enable full information to be incorporated and facilitate analysis of all related facts. Still, as a normal practice, not more than four characteristics should be represented in one table to avoid confusion. Other related tables may be formed to show the remaining characteristics

TO BE CONTINUED....

PG-TRB MATERIALS

- PG TRB: ENGLISH MATERIAL WITH QUESTION BANK.
- PG TRB: TAMIL STUDY MATERIAL +QUESTION BANK (T/M)
- MATHEMATICS MATERIAL WITH QUESTION BANK (E/M)
- PG TRB: PHYSICS MATERIAL WITH QUESTION BANK (E/M)
- PG TRB: CHEMISTRY MATERIAL + QUESTION BANK (E/M)
- PG TRB: COMMERCE MATERIAL WITH Q.BANK (T/M)&(E/M)
- PG TRB:ECONOMICS MATERIAL+Q.BANK (T/M & E/M)
- PG TRB: HISTORY MATERIAL + Q.BANK (T/M & E/M)
- PG TRB: ZOOLOGY MATERIAL + QUESTION BANK (E/M)
- PG TRB: BOTANY MATERIAL +QUESTION BANK (T/M& E/M)
- PG TRB: GEOGRAPHY STUDY MATERIAL (E/M)
- TRB-BEO STUDY MATERIAL (T/M & E/M) AVAILABLE

PG-TRB COMPUTER INSTRUCTOR-GRADE-I (NEW SYLLABUS)-2021-2022