Collection of data

Primary and Secondary Data:
Primary data are those which are collected for the first time and are always given in the form of raw materials and originals in character. These types of data need the application of statistics methods for the purpose of analysis and interpretation. While secondary data are those which have already been collected by someone and have gone thought the statistical machines. They are usually refined of the raw materials .when statistical methods are applied on primary their shape and become secondary data.

Methods of Collection of Primary Data:
The primary data are collected by the following methods.

1. Direct personal investigation.
2. indirect personal investigation
3. Investigation thought questionnaire.
4. investigation through questionnaire in the charge if enumerator
5. Investigation through local's reports.

1. **Direct Personal Investigation:** According to this methods the investigator has to collect his information himself personally form the source concerned. It means the investigator should be are the spot where the enquiry concerned. It means the investigator should be at the spot where the enquiry is being conducted, it is also expected that the investigator should be very polite and courteous. Further he should acquaint himself with the surrounding situation and must know their local customs and tradition.
   Advantages:
   1. the information collected by this methods is reliable and accurate
   2. it is a good methods for intensive investigation
   3. This method gives a satisfactory result provided the scope of inquiry is narrow.

   Disadvantages:
   1. this methods is not suitable for extensive inquiry
   2. its required a lot of expenses and time
   3. the bias on the part of investigator can damage the whole inquiry
   4. sometimes the informant may be reluctant to answered the question
2. **Indirect Personal Investigation:** This method is used when the informants are reluctant to give the definite information. e.g., if a government servant is asked to give the information regarding his income. He will not be willing to give the information for the additional income which he earned by doing part time worked. In such cases what is done? The investigators puts the informant some suitable indirect question which provides him some suitable information. Thus the only difference between the first and the second methods is that in the first methods he investigator puts directs question and collect the information while in the second methods no direct question is put to the informant but only indirect questions are asked. Even then, if it is not possible for the investigator to collect the information by the above methods then the information is collected through indirect sources, i.e. from the persons who have full knowledge of the problem under study. The persons from whom the desired information is collected are known as witnesses. Usually a list of question is prepared which is put before the

collected by this method largely depends upon the persons who are selected to give information. Hence it is necessary to take the following precautions for the selection of the informant.

3. **Investigation through questionnaire:** According to this method a standard list of questions relating to the particular investigation is prepared. This list of questions is called a questionnaire. The data are collected "By sending the questionnaire to the informants and requesting them to return the questionnaire after answering the questions. " This method is an important one and is usually used by research workers, non-official bodies and private individuals.

   Choice of Questionnaire: The success of the investigation largely depends upon the proper choice the questions to be put to the informants. While preparing a questionnaire the following points should be kept in mind.

   i. **Short and clear: -** The questions should be short and clear so as to be easily intelligible to every man. There should be no ambiguity in the questions. If some technical terms are used in the questionnaire, their definitions should be given.

   ii. **Few in number and easy:** The questions should be few in number. A large number of the questions would harm the informants because they take much time to answer, with the result they would not pay much attention to ever question and would try to save their akin by giving vague answers. Moreover the questions should be easy to answer.

   a. Definiteness: The questions should be such the answers of which are definite and exact. Preferably the questions should be such the replies of which are in the form of "Yes" or "No" Such questions should not be framed the replies of which are vague in nature because such replies are of no use to a statistician.

   b. Corroborating in nature: The questions should be such that their replies check the value replies and truth can be easily verified from them.

   c. Non-confidential information: The questions framed should not be such which call the confidential information of the informants. This will injure the feelings with the result that they would not give proper answer.

   d. Logical sequence: The questions framed should be put in some logical order; their replies should also be put in the same order because this would facilities the work.

4. **Investigation through questionnaire in charge of enumerators:** According into this method enumerators are appointed who go to the informants with the questionnaire and help them in recording the answer. Here the enumerators explain the background, aim and object of the problem under investigation and emphasize the necessity of giving correct answer. They also help the informants in understanding some t4echnical terms of question the concept of which is not clear to the informants. Thus the questionnaire is filled by the informants in the presence and help of the enumerators.

5. **Investigation through local reports:** According to this method the collection of data is neither through the questionnaire nor through the enumerators but through local correspondents. This method of collecting the data is not reliable and it should be used only at those places where the purpose the investigation is served by rough estimates.

COLLECTION OF SECONDARY DATA

The secondary data are those which have already been collected by someone other than the investigator himself, and as such the problems associated with the original collection of data do not arise here. The secondary data can be collected directly either form published or unpublished sources. The following are the sources of published at from which secondary data can be collected.

1. Official publications, i.e. the publication of the central statistical office, Karachi , Ministry of Finance , Ministry of Food, Agriculture, Lahore, Industry, etc… the provincial statistical Bureau, etc.
2. Semi-Official publications , etc., the publication issued by the state Bank of Pakistan Railway Board , Board of Economic Enquiry , District councils, Municipalities, Central Cotton Committee, etc
3. Publication of trade-association, chambers of commerce, co-operative societies, and unions.
4. Research publication, submitted by research workers, economists, University bureaus, and other institutions.
5. Technical or trade journals.

Sources of Unpublished Data: The secondary data are also available from the unpublished data. Type of material can be obtained from the chamber of commerce, trade associations, labor bureaus and research workers.

Scrutiny of Secondary Data: In the words of Bowley, " It is never safe to take published statistics at their face value without knowing their meaning and limitations and it is always necessary to criticize arguments that can be based on them , " Thus the data collected by some other person should not be fully depended as they might have pitfalls. Thus it becomes necessary to find out the inconsistencies probable errors and omissions in the data. This necessities the scrutiny of secondary data because it is just possible that the data might be inaccurate, inadequate or even unsuitable for the purposes of investigation. Hence the secondary data should possess the following qualities:

1. Reliability
2. Suitability
3. Adequacy

1. **RELIABILITY:** In order to test the reliability of the data following points should be considered:

i. Who collected the data?
ii. ii. The source of collection of the data

i. Is the reliability of the compiler dependable?
ii. Is the source of the collection of the data dependable?
iii. What was the scope and object of the investigation?
iv. Were the data collected by the use of proper methods?
v. Were the statistical units defined in which the compiler collected the data?

vi. What was the period of the collection of data?
vii. What was the type of inquiry? Was it census or sample?
viii. What was the degree of accuracy desired and achieved?
ix. Were the data in comparable form?

2. **SUITABILITY:** If the data are reliable it does not mean that they are suitable for every investigation. Data which are found suitable for one inquiry might not be suitable for another one. These necessities that the suitability of the data for the inquiry under investigation is very essential.