2. Collection of Data

Question 1. Frame at least four appropriate multiple-choice options for the following questions:

i. Which of the following is the most important when you buy a new dress?

ii. How often do you use computers?

iii. Which of the newspaper do you read regularly?

iv. Rise in the price of petrol is justified.

v. What is the monthly income of your family?

Answer:

i. Which of the following is the most important when you buy a new dress?
a. price
b. quality style
d. size

ii. How often do you use computers?

a. every day
b. 5-6 times a week 3-4 times a week
d. less than 3 times a week

iii. Which of the newspaper do you read regularly?

a. The Hindu
b. The Times of India
c. The Economics Times
d. Hindustan Times

iv. Rise in the price of petrol is justified

a. Because of increase in supply for petrol
b. Because of abundance of petrol
c. Because of hike in excise duty on petrol d. Because of appreciation of rupee

v. What is the monthly income of your family?

a. Less than Rs 30,000

b. Rs 30,000 to Rs 50,000

c. Rs 50,000 to Rs 1,00,000

d. More than Rs 1,00,000

Question 2. Frame five two-way questions (with 'Yes' or 'No').

Answer:

1. Do you have a car in your family?

2. Do you like economics?

3. Have you ever been out of your country?

4. Have you ever visited Taj Mahal?

5. Have to contribute towards Swaach Baharat Abhiyan?

Question 3. State whether the following statements are True or False.

(i) There are many sources of data.

(ii) Telephone survey is the most suitable method of collecting data, when the population is literate and spread over a large area.

(iii) Data collected by investigator is called the secondary data.

(iv)There is certain bias involved in the non- random selection of samples.

(v) Non-sampling errors can be minimized by taking large samples.

Answer:

(i) There are many sources of data. – True

Reason: There are many sources of data collection which can be categories under two heads Primary data sources and secondary data sources. Primary data sources consists of Direct personal interviews, Indirect oral investigations, Information through mailed Questionnaire or Information through questionnaire filled by enumerators. The sources of secondary data includes published sources like books, magazine, government survey reports etc. and unpublished sources like research papers, company documents, etc.

(ii) Telephone survey is the most suitable method of collecting data, when the population is literate and spread over a large area. - False

Reason: Telephone survey is most suitable method of collecting data when population is illiterate. When population is literate and spread over a large area the most suitable method is mailed questionnaire method.

(iii) Data collected by investigator is called the secondary data. - <u>False</u>

Reason: Data collected by the investigator himself is called primary data; the data collected by some other agency and which is already available for furthers studies are called secondary data.

(iv)There is certain bias involved in the non-random selection of samples. - <u>True</u>

Reason: Non-random sampling is done on the basis of judgement and convenience of the investigator hence it is not free of certain bias like personal prejudice.

(v) Non-sampling errors can be minimized by taking large samples - <u>False</u>

Reason: Non-sampling errors are more serious than sampling errors and it is difficult to minimize them, ever by taking a large sample. Some of the non-sampling errors are: errors in data Acquisition, non-response errors, sampling bias, processing errors, etc.

Question 4. What do you think about the following questions? Do you find any problems with these questions? Describe.

(i)   How far do you live from the closest market?

(ii)   If plastic bags are only 5 percent of our garbage, should it be banned?

(iii)   Wouldn't you be opposed to increase in price of petrol?

(iv)   Do you agree with the use of chemical fertilisers?

(v)   Do you use fetilisers in your fields?

(vi)   What is the yield per hectare in your field?

Answer:

(i) How far do you live from the closest market?

-   This question is quite ambiguous and respondent might find it difficult to answer it with certainty.

(ii) If plastic bags are only 5 percent of our garbage, should it be banned?

- This question is too long; it might discourage respondent to answer. It also gives a clue about how the respondent should answer which can influence the respondents answer, it should be avoided.

(iii) Wouldn't you be opposed to increase in price of petrol?

- This question contains double negatives. The questions starting with "wouldn't you" or "don't you" should be avoided as they may create confusion and lead to biased responses.

(iv)Do you agree with the use of chemical fertilisers? (v)Do you use fetilisers in your fields?

(vi)What is the yield per hectare in your field?

- The above three questions (iv, v and vi) are correct but the order in which they have been asked is not correct. One should always ask the general questions first and then should move towards more specific ones. This helps the respondents to feel comfortable.

- The correct order to ask these questions would have been.

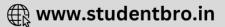i. What is the yield per hectare in your field?
ii. Do you use fetilisers in your fields?
iii. Do you agree with the use of chemical fertilisers?

Question 5. You want to do a research on the popularity of Vegetable Atta Noodles among children. Design a suitable questionnaire for collecting this information.

Answer:

- Name: ………………
- Age: …………………
- Gender: □ Male □ Female

1. Do you like eating noodles?

□ Yes □ No

2. How often do you eat noodles in a month?

□ Less than twice □ Less than three time

□ More than three times a month.

3. Have you eaten Vegetable Atta Noodles?

□ Yes □ No

4. Do you like Vegetable Atta Noodles?

□ Yes □ No

5. Do you prefer Vegetable Atta Noodles over other Masala noodles made of maida?

□ Yes □ No

6. In what meal time would you like to have Vegetable Atta Noodles?

□ Breakfast □ lunch

□ Evening snack □ dinner

7. Does your parents approve of you eating Vegetable Atta Noodles?

□ Yes □ No

Question 6. In a village of 200 farms, a study was conducted to find the cropping pattern. Out of the 50 farms surveyed, 50% grew only wheat. What is the population and the sample size?

Answer: Population or universe in statistics refers to the totality of the items under study. Since here the study was conducted to find out the cropping method of a village comprising 200 farms, the population size is 200 farms. Sample in statistics refers to a group or section of population from which the information is to be obtained or collected. Since here the information is collected from only 50 farms, the sample size for the study is 50 farms.

Question 7. Give two examples each of sample, population and variable.

Answer: Sample in statistics refers to a group or section of population from which the information is to be obtained or collected. Or that part of population or universe which is selected for the purpose of study as representative of the universe.

Population or universe in statistics refers to the totality of the items under study.

A variable refers to the characteristics of the sample or population under study that can be expressed in numbers and are generally represented by the letters X, Y, Z. Each value of a variable is an observation.

Example 1: To find out the average weight of students 10 students from a class of 50 students are selected.

Here,

- Sample is 10 students
- Population is 50 students
- Variable is height

Example 2: To find out the average income of workers in a factory comprising 250 workers, 50 workers are selected at random.

Here,

- Sample is 50 workers
- Population is 250 workers
- Variable is income

Question 8. Which of the following methods give better results and why?

(a) Census

(b) Sample

Answer: Census method refer to inclusion of all items in the field of statistical enquiry and sample methods refer to selection of few items as representatives of all the items.

Sample method gives better results than census method because of following reasons:

1. Economical method: Census method requires huge expenditure. As a large number of enumerators have to be employed. They have to be trained. Their work has to be coordinated and supervised. There will be expenditure on traveling, food etc in case of census method. The cost of the survey is much smaller if we a sample method because of less efforts involved in it.

2. Less time and efficiency consumption: The collection of data, the tabulation, and analysis take much less time in case of sampling method compared to census method. In fact the population census of India takes so much time that it takes place only once in ten years.

3. Accuracy: In sampling method it is possible to check the extent or errors and take corrective actions. While in census method, it is almost impossible to detect errors, owing to its large magnitude.

4. Less non-sampling errors: The magnitude of non- sampling errors is also much smaller in case of sampling method because of smaller size of data.

5. More reliable: In sampling method scientific methods and trained investigators are employed for the collection of data, which make it more reliable than census method.

Question 9. Which of the following errors is more serious and why?

(a) Sampling error

(b) Non- sampling error

Answer: Non-sampling errors are more serious than the sampling errors. Sampling errors arise due to drawing of inferences about the population on the basis of a few observations. For example the estimate of average income of people in a certain region on the basis of a small set of observations may not be equal to the actual average. The difference between the sample estimate and the true average income in the region is called the sampling error. Such kind of errors can be minimised by increasing the size of sample.

Non sample errors are those which may arise due to error in collection or measurement of data. The enumerators or respondents may misinterpret the questions, or some enumerators may not be very efficient. They may not record the data correctly. It is difficult to minimise such non-sampling errors even by increasing the size of sample.

Thus, non-sampling errors are more serious than sampling errors.

Question 10. Suppose there are 10 students in your class. You want to select three out of them. How many samples are possible?

Answer:

Population =10

Sample size = 3

Number of possible samples can be calculated using combination method, $^{10}C_3$:

Number of possible samples = 10X 9X 8 / 3 X 2 X 1 = 120 samples

Question 11. Discuss how you would use the lottery method to select 3 students out of 10 in your class.

Answer: To select 3 students out of 10 in a class by using lottery method we will follow the following 4 steps.

1st We will assign one number to each 10 student.

2nd We will right down all these numbers on the identical chits. (Make sure no two chits contain the same number).

3rd We will put all these chits in a box and mix them well.

4th Then we will pick out the 3 chits randomly out of the box one by one. (Without replacement)

The three students whose number is written on the drawn 3 chits will consider selected.   This way every student has an equal chance of getting selected.

Question 12. Does the lottery method always give you a random sample? Explain.

Answer: Yes, lottery method always gives a random sample. In random sampling, the items which get selected are beyond the control of the investigator, it depends entirely on chance. Each and every unit in the population has an equal chance of being selected. Similarly in lottery method, each individual unit is selected at random from the population and has equal opportunity of being selected.

For example if there are 50 students, out of which we have to select 5, we will put 50 slips containing the names of all the students in a bowl and mix them well. Then five slips are selected from the bowl one by one without replacement. The students corresponding to these 5 slips will constitute our sample.

Here each student has an equal opportunity of being selected same as in case of random selection. Hence, lottery method always gives a random sample.

Question 13. Explain the procedure for selecting a random sample of 3 students out of 10 in your class by using random number tables.

Answer: For selecting a simple random sample, random number tables can be used. Several standard tables of random numbers are available. For using these tables, the population units comprising N units should be numbered from 1 to N, which makes it possible to determine the range of numbers to be selected.

Since our population size in 10 students, two digit random tables are selected.

To select 3 students out of ten, we first identify 10 students with a number from 1 to 10, like 01, 02, 03, 04, 05, 06, 07, 08, 09, and 10.

Then we take any two digit random number table and choose any page from the table. Starting at any row or column, we select 3 two digit numbers one by one, discarding the numbers greater than ten.

Let us assume the first selected number is 04.

The two numbers successive to the selected number (04) either horizontally or vertically comprises the remaining two sample say 06 or 08.

Finally the students bearing these selected numbers will constitute our sample.

Question 14. Do samples provide better results than surveys? Give reasons for your answer.

Answer: Yes, samples do provide better results than surveys. The reasons behind the same are as follows:

1. Minimised costs- Expenditure is expected to be smaller if data collected is a small fraction of the total population using samples then if complete survey is done. In situation of limited financial resources at disposal, sample study is economical and viable choice.
2. Quick results –Data collection and summarisation is quicker in a sample method than in a survey method with complete count. It is a feasible choice when information is required on urgent basis.
3. Greater Scope- In certain circumstance or certain type of enquiry, to acquire and analyse data, highly trained personnel, specific equipments and expertise are required which are very limited in availability in such cases sampling methods are more practical.
4. Detailed enquiry and greater accuracy- Due to limited number of units under study in case of samples; it is possible to collect more detailed information by conducting intensive enquiries. Even highly trained and efficient people can be employed to achieve accurate results.
5. Highly convenient- The administration and organization of samples are easy in comparison to surveys. Samples can provide reasonably reliable and accurate information in shorter time and at a lower cost.