Collection of Data

Multiple Choice Type Questions

1. Which of the following are known as the types of research data?
a. Organised data and unorganised data
b. Qualitative data and quantitative data
c. Processed data and unprocessed data
d. None of the above

2. Which of the following statements is true about the collection of data?
a. The data that is collected from the place of origin is known as primary data
b. The data that is collected from the place of origin is known as secondary data
c. The data that is collected from the place of origin is known as tertiary data
d. None of the above

3. Which of the following statements is true about the source of data?
a. The source of data that is collected and compiled by others is known as secondary data
b. The source of data that is collected and compiled by others is known as tertiary data
c. The source of data that is collected and compiled by others is known as primary data
d. None of the above

4. Which of the following statements is true about data in research?
a. The data used for research is quantitative
b. The data used for research can be qualitative but never quantitative
c. The data used for research can be both quantitative and qualitative
d. The data used for research can be quantitative but never qualitative

5. Which of the following statements is not true about the collection of data?
a. In the structured interview, the interviewer asks a set of pre-planned questions in a standard format
b. In the observation method, the researcher collects data with the help of their observational skills
c. In an indirect interview, the investigator directly asks the informants for the data
d. In a direct interview, the researcher meets the interviewees directly to collect the data

ANSWER

1. (b) 2. (a) 3. (a) 4. (c) 5. (c)

SHORT ANSWER QUESTIONS

1. Which of the following methods give better results and why?
(a) Census
(b) Sample
Answer: In terms of accuracy of results, census is better as it studies all the units of population but this method is very time consuming, expensive and sometimes not feasible to use. Hence, sampling is better due to following reasons

- Economical Sampling involves study, of a fraction of population and hence the cost involved In sampling is relatively low.
- Time Saving Huge amount of time is required to conduct a census survey while sample studies do not take that much time.

- Lesser Effort As only a part of the population is studied, it entails lesser effort on the part of the investigator than that required in census.
- Considerable Accuracy Results from sampling may not be as accurate as in case of sampling but the level of accuracy of these results can be established through statistical tests of significance and hence can be applied in general to the whole population if found significant.

2. Which of the following errors is more serious and why?
   (a) Sampling error
   (b) Non-sampling error
   Answer: Sampling error refers to the difference between the sample estimate and the actual value of a population characteristic. This type of error occurs when one makes an observation from the sample taken from the population. It is possible to reduce the magnitude of sampling error by taking a larger sample.

   Non-sampling errors are more serious than sampling errors because a sampling error can be minimised by taking a larger sample but it is difficult to minimise non-sampling error, even by taking a large sample. Even a Census can contain non-sampling errors. These include errors in data acquisition, non-response errors and sampling bias.

3. Does the lottery method always give you a random sample? Explain.
   Answer: Lottery method always gives a random sample if it is used in the proper manner without any bias. If the slips are prepared properly and drawn out one by one so that all the slips have equal chance of being selected in the sample, it will definitely give a random sample. But, if the slips are not mac . of identical size and identification is possible of the names or numbers on the slips, the selection will become biased.

   Similarly, if the same name or number is written on more than one slip and if some name or number is missed then also the chances of selection of different units of population in the sample will not be equal. In such cases even lottery method will not give random sample.

4. Explain the procedure of selecting a random sample of 3 students out of 10 in your class, by using random number tables.
   Answer: Random number tables have been devised to guarantee equal probability of selection of every individual unit in the population according to their listed serial number in the sampling frame. They are available either in a published form or can be generated by using appropriate software packages.
   The procedure of selecting a random sample of 3 students out of 10 in a class, by using random number tables is as follows

- Assign a specific number between 1 and 10 to all the 10 students.
- Here, the largest serial number is 10 which is a two digit number and therefore we consult two digit random numbers in sequence.
- We can start using the table from anywhere, i.e., from any page, column, row or point and select the first number randomly. We need to select a sample of 3 students out of 10 total students.
- We will select two more numbers from the table according to sequence. We will skip the random numbers greater than 10 since there is no student number greater than 10. Thus, the 3 selected students are with serial numbers.

5. Define secondary data.

   Answer: According to Wessel, “Data collected by another person is known as secondary data”. It is known as secondary data as it has already been collected by somebody else. These data are accessible in the form of a published and unpublished report.

LONG ANSWER QUESTION

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

   Answer:

   A survey, which includes every element of the population, is known as Census or the Method of Complete Enumeration. On the other hand, when a part of the population is studied and predictions are made about the population based on this part, it is called sampling.

   In terms of accuracy of results, census is better as it studies all the units of population but this method is very time consuming, expensive and sometimes not feasible to use. Hence, sampling is better due to following reasons

- Economical Sampling involves study of a fraction of population and hence the cost involved in sampling is relatively low. Census costs are high especially in case of large population with wide coverage in terms of area.
- Time Saving Huge amount of time is required to conduct a census survey if the population size is large or spread over a wide area while sample studies do not take that much time to be conducted.
- Lesser Effort As only a part of the population is studied, it entails lesser effort on the part of the investigator than that required in census.
- Inappropriateness of Census In certain case, when the population is infinite or exhaustible, census cannot be done and hence sampling is the only choice, e.g., one cannot burn all the units of coal available to know their calorific value; sample is the only means of testing it.
- Considerable Accuracy Results from sampling may not be as accurate as in case of sampling but the level of accuracy of these results can be established through statistical tests of significance and hence can be applied in general to the whole population if found significant.

2. What is Data Handling?

   Answer: Data handling can be defined as the method of performing statistical analysis on the given data. Now, you would think, what is Data? Why do we need data? Data can be defined as individual pieces of information, information about a particular system. The average human body temperature measures 37oC, which is data. Data can be made useful by data collection, data organization and this data can finally be processed through data analysis and finally it is depicted with the help of graphs or charts.

   For example if someone asks you how old you are, you reply to them with a number, correct. Do you know that you just handled data! Your age is a data, your height is a data and your roll numbers are all data! Now every single person in the world is constantly interacting with data and generating data; your messages, your mails all contain data in large quantities.

   What you handle day to day in our life is known as Raw Data and this kind of data by itself does not have any meaning until. Let’s take any test you may have recently had at the school. You have a fixed and you know the number of students in your class. The marks that you obtain are the data here. You can see that your marks by themselves don’t mean much to anyone but they do mean to you right? If I wanted to know how your class performed in the test then just your data is no good. It's only when the data is organized and processed that it is useful to us.

3. What is a Collection of Data?

Answer: Collection of data is more important than organizing data is. Suppose that you want to find out the marks your friend has obtained in a particular test. So you go to your teacher and ask to have a look at his paper but your teacher says that he cannot give you the answer sheet and she can only tell you the marks. So your teacher tells you the marks in the subject and you note these marks down and go to your friend. On your way you suddenly notice that your friend has scored exactly the same marks as you have scored. Is this a coincidence? What happened here was that the teacher told you your own marks assuming that you were there to ask your marks. So the entire data you have collected is wrong. In the concept of data handling, it is extremely important to know what the same data is being collected for before we actually collect it.

Organization of Data

We know that data needs to be always presented in a contextual and an organized manner! Here's an example, you must definitely have seen the old yellow pages. They were these huge books that used to have the name and number of every resident of a particular city. So that book contained the names and numbers of thousands of people. So if you go to a city you've never been to before and you want to meet any of your friends who you know live in that city, then all you need to do is flip through the yellow pages and you'll find his name and number there , this makes the task easy!

To make such a book, you need a lot of numbers from a lot of people .Receiving numbers is the easy bit. If you send a bunch of people around the town enquiring for people's phone numbers, soon enough you will have a considerable amount of data. For this data to be useful, the data will have to be arranged so that it helps in searching quickly. In the yellow pages, the names are arranged alphabetically which means you can easily look for it just like you would in a dictionary. That would help in tracking names down easily.

Types of Data-

Data handling methods can be performed based on the types of data. The data is generally classified into two types, such as:

Qualitative Data

Quantitative Data

Qualitative data can be defined as something that gives us descriptive information about something whereas quantitative data can be defined as something that gives numerical information about something. Here, the quantitative data can be further divided into two. They are discrete and continuous data. The discrete data can take only certain values, for example whole numbers. The continuous data can take a value within the provided range.

Qualitative data collection methods include observation method, interview method, questionnaire methods, schedules, etc.

- The observation method is used when the study is related to the behavioural sciences. It is a systematically planned method. The various kinds of observations are:
  a. Unstructured and structured observations
  b. Uncontrolled and controlled observations
  c. Non-participant, disguised and participant observations
- The interview method is used to collect data through verbal responses. The two ways in which this method is achieved are telephonic interviews and personal interviews.
- In the questionnaire method, a set of questions is provided to the respondent and he is asked to return it after filling in his replies. The questions should be printed in a definite order. The features of a good survey are:
  a. Simple and short
  b. Sequenced logically

c. Good appearance

d. Contains some space for answers

4. The scheduling method is similar to the interview method. It determines the objectives of the investigation. Here, enumerations are appointed to fill the schedules.

Answer: The data can be usually represented in any one of the following ways:

1. Bar Graph
2. Line Graphs
3. Pictographs
4. Histograms
5. Stem and Leaf Plot
6. Dot Plots
7. Frequency Distribution
8. Cumulative Tables and Graphs