since it is clear "what is to be observed". He may instruct his waiters to record this. This information is required to decide requirements of the chairs and tables and also the ambience.

Suppose, the manager wants to know how single customers and those with families behave and what their attitudes are like. This study is vague, and it needs a non-structured observation.

It is easier to record structured observation than non-structured observation.

Example 2:

To distinguish between structured and unstructured observations, consider a study, investigating the amount of search that goes into the purchase of a soap. On the one hand, the observers could be instructed to stand at one end of a supermarket and record each sample customer's search. This may be observed and recorded as follows: "The purchaser first paused after looking at HLL brand. He looked at the price on of the product, kept the product back on the shelf, then picked up a soap cake of HLL and glanced at the picture on the pack and its list of ingredients, and kept it back. He then checked the label and price for P&G product, kept that back down again, and after a slight pause, picked up a different flavour soap of M/s. Godrej Company and placed it in his trolley and moved down the aisle". On the other hand, observers might simply be told to record the "first soap cake examined", by checking the appropriate boxes in the observation form. The 'second situation' represents more structured than the first.

The observation method is the only method applicable to study the growth of plants and crops.

To use a more structured approach, it would be necessary to decide precisely what is to be observed and the specific categories and units that would be used to record the observations.

Disguised-Undisguised Observation

In disguised observation, the respondents do not know that they are being observed. In non-disguised observation, the respondents are well aware that they are being observed. In

disguised observation, observers often pose as shoppers. They are known as "mystery shoppers". They are paid by research organisations. The main strength of disguised observation is that it allows for registering the true of the individuals.

In the undisguised method, observations may be restrained due to induced error by the objects of observation. The ethical aspect of disguised observations is still open to question and debate.

Direct-Indirect Observation

In direct observation, the actual behaviour or phenomenon of interest is observed. In indirect observation, the results of the consequences of the phenomenon are observed. Suppose, a researcher is interested in knowing about the soft drinks consumption of a student in a hostel room. He may like to observe empty soft drink bottles dropped into the bin. Similarly, the observer may seek the permission of the hotel owner to visit the kitchen or stores. He may carry out a kitchen/stores audit, to find out the consumption of various brands of spice items being used by the hotel. It may be noted that the success of an indirect observation largely depends on "how best the observer is able to identify physical evidence of the problem under study".

Human-Mechanical Observation

Most of the studies in marketing research are based on human observation, wherein trained observers are required to observe and record their observation. In some cases, mechanical devices such as eye cameras are used for observation. One of the major advantages of electrical/mechanical devices is that their recordings are free from any subjective bias.

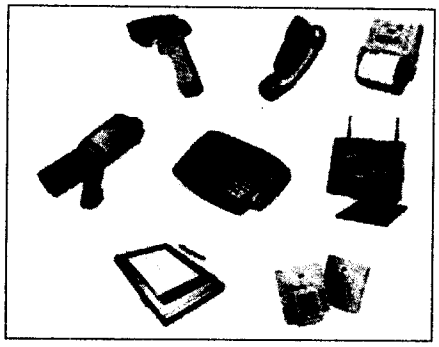
10.3.2 Advantages of Observation Method

1. The original data can be collected at the time of occurrence of the event.
2. Observation is done in natural surroundings. Therefore, the facts emerge more clearly, whereas in a questionnaire, experiments have environmental as well as time constraints.
3. Sometimes, the respondents may not like to part with some of the information. Such information can be obtained by the researcher through observation.
4. Observation can also be done on those who cannot articulate.
5. Any bias on the part of the researcher is greatly reduced in the observation method.

10.3.3 Limitations

1. The observer might wait for longer period at the point of observation. And yet the desired event may not take place. Observation is required over a long period of time and hence may not occur.

2. For observation, an extensive training of observers is required.
3. This is an expensive method.
4. External observation provides only superficial indications. To delve beneath the surface is very difficult. Only overt behaviour can be observed.
5. Two observers may observe the same event, but may draw different inferences.
6. It is very difficult to gather information on (1) Opinions (2) Intentions.



These are the various tools to collect primary data

10.4 DESIGNING THE QUESTIONNAIRE

Questionnaire: Its Importance and Characteristics.

Questionnaire: A questionnaire is a tool used to collect the data.

Importance of Questionnaire in MR: To study:

1. Behaviour, past and present.
2. Demographic characteristics such as age, sex, income, occupation.
3. Attitudes and opinions.
4. Level of knowledge.

Characteristics of Questionnaire

1. It must be simple. The respondents should be able to understand the questions.
2. It must generate replies that can be easily be recorded by the interviewer.
3. It should be specific, so as to allow the interviewer to keep the interview to the point.
4. It should be well arranged, to facilitate analysis and interpretation.
5. It must keep the respondent interested throughout.

10.5 DIFFERENT TYPES OF QUESTIONNAIRE

1. Structured and Non-disguised
2. Structured and Disguised
3. Non-structured and Disguised
4. Non-structured and Non-disguised

1. ***Structured and Non-disguised Questionnaire:*** Here, questions are structured so as to obtain the facts. The interviewer will ask the questions strictly in accordance with the pre-arranged order. For *example:* What are the strengths of soap A in comparison with soap B?

   - ❖ Cost is less
   - ❖ Lasts longer
   - ❖ Better fragrance
   - ❖ Produces more lather
   - ❖ Available in more convenient sizes

   Structured and non-disguised questionnaire is widely used in market research. Questions are presented with exactly the same wording and same order to all respondents. The reason for standardizing the question is to ensure that all respondents reply the same question. The purpose of the question is clear. The researcher wants the respondent to choose one of the five options given above. This type of questionnaire is easy to administer. The respondents have no difficulty in answering, because it is structured, the frame of reference is obvious.

   In a non-disguised type, the purpose of the questionnaire is known to the respondent.

   Example: "Subjects attitude towards Cyber laws and the need for government legislation to regulate it".

   Certainly, not needed at present

   Certainly not needed

   I can't say

   Very urgently needed

   Not urgently needed

2. ***Structured and disguised Questionnaire:*** This type of questionnaire is least used in marketing research. This type of questionnaire is used to know the peoples' attitude,

when a direct undisguised question produces a bias. In this type of questionnaire, what comes out is "what does the respondent know" rather than what he feels. Therefore, the endeavour in this method is to know the respondent's attitude.

Currently, the "Office of Profit" Bill is:

(a) In the Lok Sabha for approval.

(b) Approved by the Lok Sabha and pending in the Rajya Sabha.

(c) Passed by both the Houses, pending the presidential approval.

(d) The bill is being passed by the President.

Depending on which answer the respondent chooses, his knowledge on the subject is classified.

In a disguised type, the respondent is not informed of the purpose of the questionnaire. Here the purpose is to hide "what is expected from the respondent?" ***Example 1:*** "Tell me your opinion about Mr. Ben's healing effect show conducted at Bangalore?" ***Example 2:*** "What do you think about the Babri Masjid demolition?"

3. ***Non-Structured and Disguised Questionnaire:*** The main objective is to conceal the topic of enquiry by using a disguised stimulus. Though the stimulus is standardized by the researcher, the respondent is allowed to answer in an unstructured manner. The assumption made here is that individual's reaction is an indication of respondent's basic perception. Projective techniques are examples of non-structured disguised technique. The techniques involve the use of a vague stimulus, which an individual is asked to expand or describe or build a story, three common types under this category are (a) Word association (b) Sentence completion (c) Story telling.

4. ***Non structured and Non disguised Questionnaire:*** Here the purpose of the study is clear, but the responses to the question are open-ended. *Example:* "How do you feel about the Cyber law currently in practice and its need for further modification"? The initial part of the question is consistent. After presenting the initial question, the interview becomes very unstructured as the interviewer probes more deeply. Subsequent answers by the respondents determine the direction the interviewer takes next. The question asked by the interviewer varies from person to person. This method is called "the depth interview". The major advantage of this method is the freedom permitted to the interviewer. By not restricting the respondents to a set of replies, the experienced interviewers will be above to get the information from the respondent fairly and accurately. The main disadvantage of this method of interviewing is that it takes time, and the respondents may not co-operate. Another disadvantage is that coding of open-ended questions may pose a challenge. *For example:* When a researcher

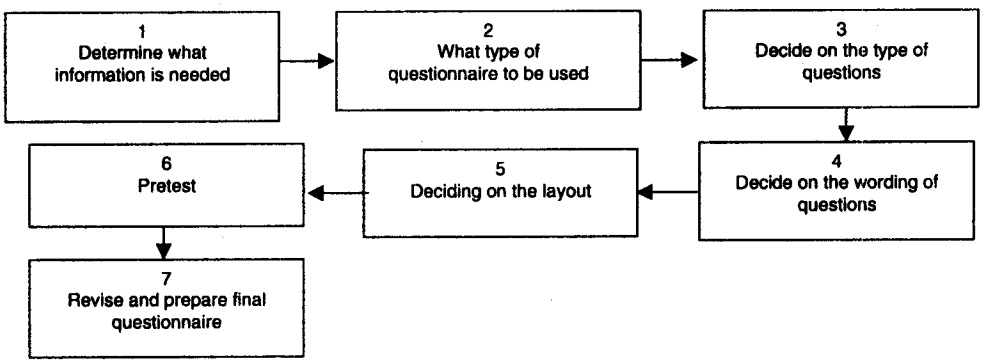
asks the respondent "Tell me something about your experience in this hospital". The answer may be "Well, the nurses are slow to attend and the doctor is rude. 'Slow' and 'rude' are different qualities needing separate coding. This type of interviewing is extremely helpful in exploratory studies.

Types of Questionnaires

Types	Characteristics
Structured – Disguised	• The same question is posed to each respondent. • Administering the questionnaire and post-administration work is simple i.e. coding tabulating, etc. is easy. • This type of questionnaire is least used in market research. • Respondents' bias is minimized.
Unstructured – Disguised	• This type of questionnaire is very commonly used for focus group discussions. • This is difficult to analyse, code etc, • No fixed set of questions. • The inner self (why) of an individual is researched. Eg: Motivation Research.
Unstructured – Undisguised	• No fixed questions. • Suitable for conducting depth interview. • Subject-matter can be questioned in great detail. • Coding, tabulating etc. are difficult not a very frequently used method.
Structured – Undisguised	• Fixed set of questions to every respondent. • Inappropriate when researcher wants to probe deeper. • Easy to administer coding, tabulating is easy. • Due to structuring and undisguised nature of the questionnaire, there is no possibility of the respondent misunderstanding the question. This is the most commonly used method.

10.6 PROCESS OF QUESTIONNAIRE DESIGNING

The following are the 7 steps:



Determine What Information is Required

The first question to be asked by the market researcher is "what type of information does he need from the survey?" This is valid because if he omits some information on relevant and vital aspects, his research is not likely to be successful. On the other hand, if he collects information which is not relevant, he is wasting his time and money.

At this stage, information required, and the scope of research should be clear. Therefore the steps to be followed at the planning stage are:

1. Decide on the topic for research.
2. Get additional information on the research issue, from secondary data and exploratory research. The exploratory research will suggest "what are the relevant variables?"
3. Gather what has been the experience with similar study.
4. The type of information required. There are several types of information such as (a) awareness, (b) facts, (c) opinions, (d) attitudes, (e) future plans, (f) reasons.

Facts are usually sought out in marketing research.

Example: Which television programme did you see last Saturday? This requires a reasonably good memory and the respondent may not remember. This is known as recall loss. Therefore questioning the distant past should be avoided. Memory of events depends on (1) Importance of the events and (2) Whether it is necessary for the respondent to remember. In the above case, both the factors are not fulfilled. Therefore, the respondent does not remember. On the contrary, a birthday or wedding anniversary of individuals is remembered without effort since the event is important. Therefore, the researcher should be careful while asking questions about the past. First, he must make sure that the respondent has the answer.

Example: Do you go to the club? He may answer 'yes', though it is untrue. This may be because the respondent wants to impress upon the interviewer that he belongs to a well-to-do family and can afford to spend money on clubs. To obtain facts, the respondents must be conditioned (by good support) to part with the correct facts.

Mode of Collecting the Data

The questionnaire can be used to collect information either through personal interview, mail or telephone. The method chosen depends on the information required and also the type of respondent. If the information is to be collected from illiterate individuals, a questionnaire would be the wrong choice.

Type of Questions

Open-ended Questions

These are questions where respondents are free to answer in their own words. *Example:* "What factor do you consider while buying a suit"? If multiple choices are given, it could

be colour, price, style, brand etc, but some respondents may mention attributes which may not occur to the researcher.

Therefore, open-ended questions are useful in exploratory research, where all possible alternatives are explored. The greatest disadvantage of open-ended questions is that the researcher has to note down the answer of the respondents verbatim. Therefore, there is a likelihood of the researcher failing to record some information.

Another problem with open-ended question is that the respondents may not use the same frame of reference.

Example: "What is the most important attribute in a job?"

Ans: Pay

The respondent may have meant "basic pay" but interviewer may think that the respondent is talking about "total pay including dearness allowance and incentive". Since both of them refer to pay, it is impossible to separate two different frames.

Dichotomous Question

These questions have only two answers, 'Yes' or 'no', 'true' or 'false' 'use' or 'don't use'.

Do you use toothpaste? Yes ..... No .....

There is no third answer. However sometimes, there can be a third answer: Example: "Do you like to watch movies?"

Ans: Neither like nor dislike

Dichotomous questions are most convenient and easy to answer.

Close-Ended Questions

There are two basic formats in this type:

- Make one or more choices among the alternatives
- Rate the alternatives

Choice Among Alternatives

Which of the following words or phrases best describes the kind of person you feel would be most likely to use this product, based on what you have seen in the commercial?

(a) Young ..... old .....

Single ..... Married .....

Modern ..... Old fashioned .....

(b) Rating Scale

(I) Please tell us your overall reaction to this commercial?

- A great commercial would like to see again.
- Just so-so, like other commercials.
- Another bad commercial.
- Pretty good commercial.

(II) Based on what you saw in the commercial, how interested do you feel, you would be buying the products?

- Definitely
- Probably I would buy
- I may or may not buy
- Probably I would not buy
- Definitely I would not buy.

Closed-ended questionnaires are easy to answer. It requires less effort on the part of the interviewer. Tabulation and analysis is easier. There are lesser errors, since the same questions are asked to everyone. The time taken to respond is lesser. We can compare the answer of one respondent to another respondent.

One basic criticism of closed-ended questionnaires is that middle alternatives are not included in this, such as "don't know". This will force the respondents to choose among the given alternative.

Wordings of Questions

Wordings of particular questions could have a large impact on how the respondent interprets them. Even a small shift in the wording could alter the respondent's answer.

Example 1: "Don't you think that Brazil played poorly in the FIFA cup?" The answer will be 'yes'. Many of them, who do not have any idea about the game, will also most likely say 'yes'. If the question is worded in a slightly different manner, the response will be different.

Example 2: "Do you think that, Brazil played poorly in the FIFA cup?" This is a straightforward question. The answer could be 'yes', 'no' or 'don't know' depending on the knowledge the respondents have about the game.

Example 3: "Do you think anything should be done to make it easier for people to pay their phone bill, electricity bill and water bill under one roof"?

Example 4: "Don't you think something might be done to make it easier for people to pay their phone bill, electricity bill, water bill under one roof"?

A change of just one word as above, can generate different responses by respondents.

Guidelines towards the use of correct wording:

Is the vocabulary simple and familiar to the respondents?

Example: Instead of using the word 'reasonably', 'usually', 'occasionally', 'generally', 'on the whole'.

Example: "How often do you go to a movie?" "Often, may be once a week, once a month, once in two months or even more."

Avoid Double-Barreled Questions

These are questions, in which the respondent can agree with one part of the question, but not agree with the other or cannot answer without making a particular assumption.

Example 1: "Do you feel that firms today are employee-oriented and customer-oriented?" There are two separate issues here – [yes] [no]

Example 2: "Are you happy with the price and quality of branded shampoo?" [yes] [no]

Avoid Leading and Loading Questions

Leading Questions

A leading question is one that suggests the answer to the respondent. The question itself will influence the answer, when respondents get an idea that the data is being collected by a company. The respondents have a tendency to respond positively. *Example 1:* "How do you like the programme on 'Radio Mirchy'? The answer is likely to be 'yes'. The unbiased way of asking is "which is your favorite F.M. Radio station? The answer could be any one of the four stations namely 1. Radio City 2. Mirchy 3. Rainbow 4. Radio-One.

Example 2: Do you think that offshore drilling for oil is environmentally unsound? The most probable response is 'yes'. The same question can be modified to eliminate the leading factor.

What is your feeling about the environmental impact of offshore drilling for oil? Give choices as follows:

1. Offshore drilling is environmentally sound.
2. Offshore drilling is environmentally unsound.
3. No opinion.

Loaded Questions

A leading question is also known as a loaded question. In a loaded question, special emphasis is given to a word or a phrase, which acts as a lead to respondent. *Example:* "Do you own a Kelvinator refrigerator." A better question would be "what brand of refrigerator do you own?" "Don't you think the civic body is 'incompetent'?" Here the word incompetent is 'loaded'.

Are the Questions Confusing?

If there is a question unclear or is confusing, then the respondent becomes more biased rather that getting enlightened. *Example:* "Do you think that the government publications are distributed effectively"? This is not the correct way, since respondent does not know what is the meaning of the word **effective distribution.** This is confusing. The correct way of asking questions is "Do you think that the government publications are readily available when you want to buy?" *Example:* "Do you think whether value price equation is attractive"? Here, respondents may not know the meaning of **value price** equation.

Applicability

"Is the question applicable to all respondents?" Respondents may try to answer a question even though they don't qualify to do so or may lack from any meaningful opinion. ***Examples:*** (1) "What is your present education level"? (2) "Where are you working" (assuming he is employed)? (3) "From which bank have you taken a housing loan" (assuming he has taken a loan)?

Avoid Implicit Assumptions

An implicit alternative is one that is not expressed in the options. Consider following two questions:

Would you like to have a job, if available?

Would you prefer to have a job, or do you prefer to do just domestic work?

Even though, we may say that these two questions look similar, they vary widely. The difference is that Q-2 makes explicit the alternative implied in Q-1.

Split Ballot Technique

This is a procedure used wherein (1) The question is split into two halves and (2) Different sequencing of questions is administered to each half. There are occasions when a single version of questions may not derive the correct answer and the choice is not obvious to the respondent.

Example: "Why do you use Ayurvedic soap"? One respondent might say "Ayurvedic soap is better for skin care". Another may say "Because the dermatologist has recommended".

A third might say "It is a soap used by my entire family for several years". The first respondent answers the reason for using it at present. The second respondent answers how he started using. The third respondent "the family tradition for using". As can be seen, different reference frames are used. The question may be balanced and rephrased.

Complex Questions?

In which of the following do you like to park your liquid funds?

- Debenture
- Preferential share
- Equity linked M.F
- I.P.O
- Fixed deposit

If this question is posed to the general public, they may not know the meaning of liquid fund. Most of the respondents will guess and tick one of them.

Are the Questions Too Long?

Generally as a thumb rule, it is advisable to keep the number of words in a question not exceeding 20. The question given below is too long for the respondent to comprehend, leave alone answer.

Example: Do you accept that the people whom you know, and associate yourself have been receiving ESI and P.F benefits from the government accept a reduction in those benefits, with a view to cut down government expenditure, to provide more resources for infrastructural development?

Yes _____ No_____ Can't say_____

Participation at the Expense of Accuracy

Sometimes the respondent may not have the information that is needed by the researcher.

Example 1: The husband is asked a question "How much does your family spend on groceries in a week"? Unless the respondent does the grocery shopping himself, he will not know how much has been spent. In a situation like this, it will be helpful to ask a 'filtered question'. An example of a filtered question can be, "Who buys the groceries in your family"?

Example 2: "Do you have the information of Mr. Ben's visit to Bangalore"? Not only should the individual have the information but also s(he) should remember the same. The inability to remember the information is known as "recall loss".

Sequence and Layout

Some guidelines for sequencing the questionnaire are as follows:

Divide the questionnaire into three parts:

(1) Basic information (2) Classification (3) Identification information. Items such as age, sex, income, education etc. are questioned in the classification section. The identification part involves body of the questionnaire. Always move from general to specific questions on the topic. This is known as funnel sequence. Sequencing of questions is illustrated below:

(1) Which TV shows do you watch?

Sports _____ News_____

(2) Which among the following are you most interested in?

Sports _____ News_____

Music _____ Cartoon_____

(3) Which show did you watch last week?

World Cup Football_____

Bournvita Quiz Contest_____

War News in the Middle East _____

Tom and Jerry cartoon show_____

The above three questions follow a funnel sequence. If we reverse the order of question and ask "which show was watched last week"?, the answer may be biased. This example shows the importance of sequencing.

Layout: How the questionnaire looks or appears.

Example: Clear instructions, gaps between questions, answers and spaces are part of layout. Two different layouts are shown below:

Layout – 1 **How old is your bike?**

_____Less than 1 year _____1 to 2 years _____2 to 4 years _____ more than 4 years.

Layout – 2 **How old is your bike?**

_____Less than 1 year

_____1 to 2 years.

_____2 to 4 years.

_____ More than 4 years.

From the above example, it is clear that layout – 2 is better. This is because likely respondent error due to confusion is minimised.

Therefore, while preparing a questionnaire start with a general question. This is followed by a direct and simple question. This is followed by more focused questions. This will elicit maximum information.

Forced and Unforced Scales

Suppose the questionnaire is not provided with 'don't know' or 'no option', then the respondent is forced to choose one side or the other. A 'don't know' is not a neutral response. This may be due to genuine lack of knowledge.

Balanced and Unbalanced Scales

In a balanced scale, the number of favourable responses are equal to the number of unfavourable responses. If the researcher knows that there is a possibility of a favourable response, it is best to use unbalanced scale.

Use Funnel Approach

Funnel sequencing gets the name from its shape, starting with broad questions and progressively narrowing down the scope. Move from general to specific examples.

1. How do you think this country is getting along in its relations with other countries?
2. How do you think we are doing in our relations with the US ?
3. Do you think we ought to be dealing with US ?
4. If yes, what should be done differently ?
5. Some say we are very weak on the nuclear deal with the US, while, some say we are OK. What do you feel ?

The first question introduces the general subject. In the next question, a specific country is mentioned. The third and fourth questions are asked to seek views. The fifth question is to seek a specific opinion.

Pre-testing of Questionnaire

Pre-testing of a questionnaire is done to detect any flaws that might be present. For *example,* the word used by researcher must convey the same meaning to the respondents. Are instructions clear skip questions clear? One of the prime conditions for pre-testing is that the sample chosen for pre-testing should be similar to the respondents who are ultimately going to participate. Just because a few chosen respondents fill in all the questions going does not mean that the questionnaire is sound.

How Many Questions to be Asked ?

The questionnaire should not be too long as the response will be poor. There is no rule to decide this. However, the researcher should consider that if he were the respondent, how would he react to a lengthy questionnaire. One way of deciding the length of the questionnaire is to calculate the time taken to complete the questionnaire. He can give the questionnaire to a few known people to seek their opinion.

10.7 MAIL QUESTIONNAIRE

10.7.1 Advantages

1. Easier to reach a larger number of respondents throughout the country.
2. Since the interviewer is not present face to face, the influence of interviewer on the respondent is eliminated.
3. Where the questions asked are such that they cannot be answered immediately, and needs some thinking on the part of the respondent, the respondent can think over leisurely and give the answer.
4. Saves cost (cheaper than interview).
5. No need to train interviewers.
6. Personal and sensitive questions are well answered.

10.7.2 Limitations

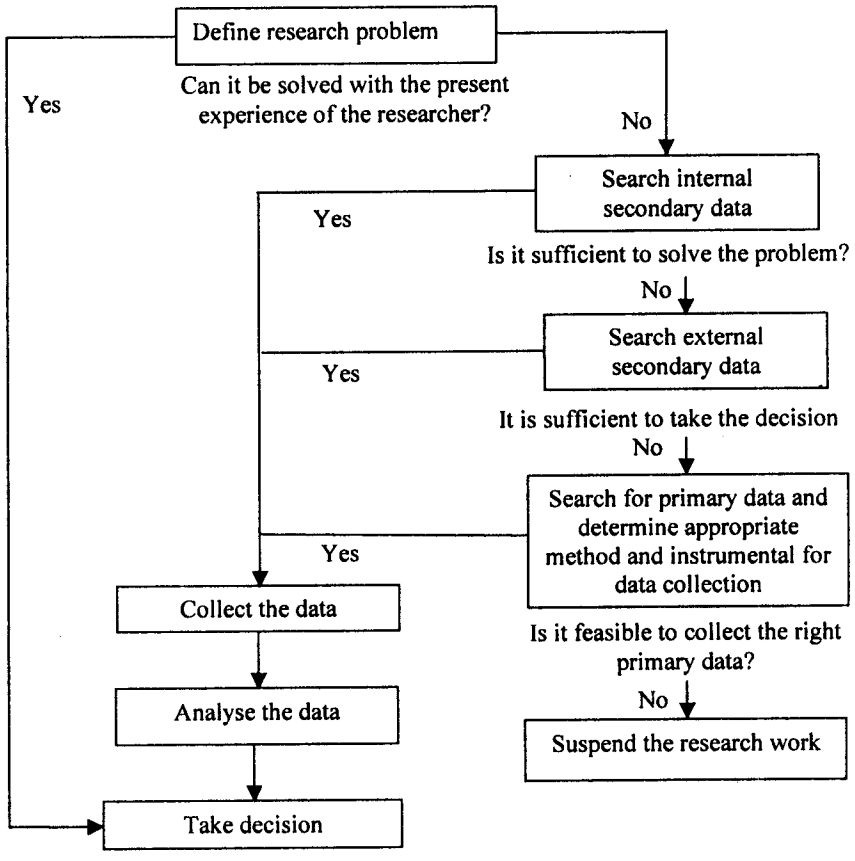
1. It is not suitable when questions are difficult and complicated. Example: "Do you believe in value price relationship"?
2. When the researcher is interested in a spontaneous response, this method is unsuitable. Because thinking time allowed to the respondent will influence the answer. *Example:* "Tell me spontaneously, what comes to your mind if I ask you about cigarette smoking".
3. In case of a mail questionnaire, it is not possible to verify whether the respondent himself/herself has filled the questionnaire. If the questionnaire is directed towards the housewife, say, to know her expenditure on kitchen items, she alone is supposed to answer it. Instead, if her husband answers the questionnaire, the answer may not be correct.
4. Any clarification required by the respondent regarding questions is not possible. *Example:* Prorated discount, product profile, marginal rate, etc., may not be understood by the respondents.

5. If the answers are not correct, the researcher cannot probe further.
6. Poor response (30%) – Not all reply.

10.7.3 Additional Consideration for the Preparation of Mail Questionnaire

1. It should be shorter than the questionnaire used for a personal interview.
2. The wording should be extremely simple.
3. If a lengthy questionnaire has to be made, first write a letter requesting the cooperation of the respondents.
4. Provide clear guidance, wherever necessary.
5. Send a pre-addressed and stamped envelope to receive the reply.

10.8 SYSTEM APPROACH TO DATA COLLECTION



10.9 SAMPLE QUESTIONNAIRES

10.9.1 A Study of Customer Retention as Adopted by Textile Retail Outlets

Note: Information gathered will be strictly confidential. We highly appreciate your cooperation in this regard.

1. Name of the outlet:
2. Address:
3. Do you have regular customers?

   Yes [ ] No [ ]
4. How often do your regular customers visit your outlet?

   Weekly [ ] Once in a month [ ] Twice in a month [ ]

   Once in 2 months [ ] 2 – 3 months [ ] Once in 6 months [ ]
5. Do you maintain any records of your regular customers?

   Yes [ ] No [ ]
6. What percentage of your customers are regular? % [ ]
7. Do you think that we can use the above as a retention strategy of customers for your outlets?

   Yes [ ] No [ ]
8. What are the different products that you handle in your outlets?

   Formals [ ] Casuals/Kids wear [ ] Ladies dress materials [ ]

   Sarees [ ] Others (Specify)
9. What type of customers (socio-economic) visit your outlets?

   Low income [ ] Middle income [ ] High income [ ]
10. Why do you think they come to your outlet?

    Product variety [ ] Price discount [ ] Easy gain to products [ ]

    Parking facility [ ] Store layout [ ] Quality [ ] Reasonable price [ ]

    Others (Specify)
11. Rank the factors that influence customers to visit your outlet:

    Credit facility [ ] Price discount [ ] Gifts [ ] Easy gain to products [ ]

Parking facility [ ] Store layout [ ] Product variety [ ]

Quality and reasonable price [ ] Others (Specify) ——————————

12. What do customers expect from the retail outlet?

Credit facility [ ] Gift coupon [ ] Price discount [ ]

Price reduction easy accessibility of product [ ] Quality and reasonable price [ ]

Others (Specify) ——————————

13. Do you have any retention strategy adopted to keep in touch with the customer?

Gifts on special occasion

(a) Birthday gift [ ] (b) Anniversary gift [ ]

(c) Festivals/Customer relationship [ ] Others (Specify) ——————————

14. Which one do you think is most effective? Please rank them?

(a) Birthday gift [ ] (b) Anniversary gift [ ]

(c) Festivals/Customer relationship [ ] Others (Specify) ——————————

Thanking You for Sparing Your Valuable Time

10.9.2 A Study on Customer Preferences of P.C.

Date: Place:

Form No: [ ] [ ] [ ] [ ] [ ]

1. Personal Profile

(a) Name: [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]

(b) Address: [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]

[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]

[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]

(c) Sex: Male [ ] Female [ ]

(d) Age: [ ] [ ] years

(e) Occupation: Self-employed [ ] Professional [ ] Service [ ] Housewife [ ]

2. Do you own a P.C.? Yes [ ] No [ ]

(a) If yes, whether: branded [ ] unbranded [ ]

(b) If no, do you plan to buy one? Near future [ ] Distant future [ ] Can't say [ ]

(Less than 6 months) (Less than a year)

If so, whether: branded [ ] unbranded [ ]

3. What is the utility of the PC to you?

   Education [ ] Business [ ] Infotainment [ ] Internet/Communication [ ]

4. What is the most important factor that matters while buying a PC?

   Quality [ ] Price [ ] Service [ ] Finance facility [ ]

5. Before deciding on the vendor, which factor goes into your consideration?

   Vendor's Reputation [ ] Technical Expertise [ ] Client Base [ ]

6. How did you come to know about the vendor?

   Friendly/Family [ ] Press Ads[ ] Direct Mailers [ ] Reference Website [ ]

7. Which configuration would you decide on while buying a PC?

   Standard [ ] Intermediate [ ] Latest/Advanced [ ]

8. In your PC, would you prefer? Conventional Design [ ] Innovative Design [ ]

   If new, why: New design distracts attention -

   New design means increased price -

   New design is hard to adapt -

   If Innovative, why? To create own identity

   Out of business need -

   Space management -

9. Rate the following four factors important for innovative design, starting with the most preferred:

   (A) Size (B) Shape (C) Colour/ordinary (D) Portability and Sturdiness

   1. _____ 3. _____

   2. _____ 4. _____

10. To what extent would the computer increase your efficiency?

    Negligible [ ] 20 – 40% [ ] 40 – 60% [ ] More [ ]

11. How many hours on an average per week would you use your PC?

    0 to 5 hours [ ] 6 to 12 hours [ ] 13 to 18 hours [ ] More [ ]

12. While using your PC, most of the time would be for:

    Education [ ] Accounting [ ] Net surfing [ ] Correspondence [ ]

13. Remarks ———————————————————————————

_____

_____

_____

Signature of the Respondent ———————————

SUMMARY

Sometimes, secondary data may not be able to solve the research problem. In that case researcher need to turn towards primary data. Primary data may pertain to life style, income, awareness or any other attribute of individuals or groups. There are 2 ways of collecting primary data namely. (a) Observation (b) By questioning the appropriate sample. Observation method has a limitation i.e, certain attitudes, knowledge, motivation etc cannot be measured by this method. For this reason, researcher needs to communicate.

Communication method is classified based on whether it is structured or disguised. Structured questionnaire is easy to administer. This type is most suited for descriptive research. If the researcher wants to do exploratory sturdy, unstructured method is better. In unstructured method questions will have to be framed based on the answer by the respondent. In disguised questionnaire, the purpose of research is not disclosed to respondents. This is done so that the respondents might speak the truth instead of giving some answer which satisfies the researcher.

Questionnaire can be administered either in person or online or Mail questionnaire. Each of these methods have advantages and disadvantages. Questions in a questionnaire may be classified into (a) Open question (b) Close ended questions (c) Dichotomous questions etc. While formulating questions, care has to be taken with respect to question wording, vocabulary, leading, loading and confusing questions should be avoided. Further it is desirable that questions should not be complex, nor too long. It is also implied that proper sequencing will enable the respondent to answer the question easily. The researcher must maintain a balanced scale and must use a funnel approach. Pretesting of the questionnaire is preferred before introducing to a large population. Personal interview to gather information is very costly. Therefore sometimes mail questionnaire is used by researcher to collect the data. However it has its own limitations.

KEYWORDS

Depth interview

Disguised

Undisguised

Mail questionnaire

Unstructured observation	Open ended questions
Closed ended questions	Dischotomous question
Double barrel question	Leading question
Split ballot technique	Pre test mail questionnaire

REVIEW QUESTIONS

1. What is primary data?
2. What are the various methods available for collecting primary data?
3. What are the advantages and disadvantages of a structured questionnaire?
4. What are the several methods used to collect data by observation method?
5. What are the advantages and limitations of collecting data by observation method?
6. What is a questionnaire? What are the different types?
7. What are the characteristics of a good questionnaire?
8. What are the limitations of a questionnaire?
9. Explain the steps involved in designing a questionnaire.
10. Explain Open ended & Closed ended questions in a questionnaire.
11. What is a split ballot method? When is it employed?
12. What is questionnaire pre-testing?
13. What is a dichotomous question? When is it most appropriate?
14. How does a questionnaire suffer compared to experimentation on account of validity & reliability?
15. What is meant by pre testing of questionnaire? Why is it required?
16. Distinguish qualitative and quantitative method of data collection.
17. What is mail questionnaire. Explain the advantages and limitations of the same?
18. What is meant by leading / loading question give example?
19. What is meant by double barreled questions?
20. Design a questionnaire to study brand preference for a consumer durable product?
21. One method of sequencing the question in a questionnaire is to proceed from general to specific. What is the logical reason behind this?

ASSIGNMENT-1

What observation technique would you use to gather the following information?

a. What kind of influence do children have on the purchase behaviour of their parents?

b. How do discounts influence the purchase behaviour of customers buying colour TV?

c. A study to find out the potential location for a snack bar in a city.

ASSIGNMENT-2

A nationalised bank wants to determine what is most effective way to increase responses to their mail questionnaire? Three possibilities are:

1. Issue a gift coupon for Rs. 25 so that the respondent can go to a specified store to avail the gift item.
2. Ask the respondent to note their name and address in the completed questionnaire. Thereafter they will be mailed a cheque for Rs. 50.
3. Along with a questionnaire, write a letter stating pen set as gifts would be sent to them, after receiving the completed questionnaire. Mail the questionnaire to 2,000 respondents chosen from four metros. Set up an experiment in which the above incentives can be tested and the most appropriate method identified.

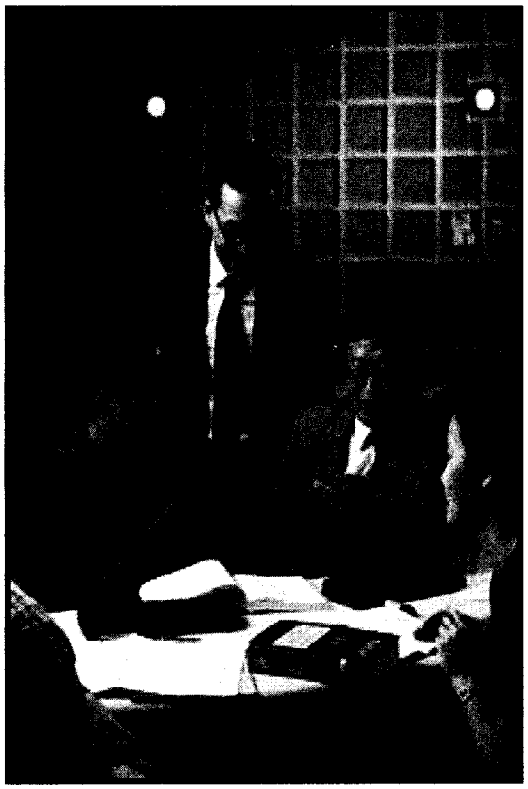
ASSIGNMENT-3

Give one example for each of the following type of the questions:

1. Leading question
2. Double-barreled question
3. Close-ended question
4. Fixed alternative question
5. Split-ballot question

CHAPTER eleven

Qualitative Techniques of Data Collection



In this chapter, the following questions are discussed:

- What are the advantages/ limitations of depth interview?
- What is a Delphi technique?
- What are focus group interviews?
- What are the types of projective techniques?
- What are the conditions for a successful interview?

11.1 QUALITATIVE TECHNIQUES OF DATA COLLECTION

Qualitative research is used to analyse those data which cannot be quantified. Qualitative research is used in exploratory research. The number of respondents covered in this type of research is small compared to quantitative research.

11.1.1 Qualitative v/s quantitative research

These two research methods vary in a number of ways. They vary in terms of

1. Measurability criteria
2. Features
3. Characteristics

Quantitative research

Measurability: quantitative data is measurable, for example; size of market, rate of product usage etc.

Features:

1. Data collected is numerical in nature.
2. Data collection methods are
   a) Mail questionnaire
   b) Personal interview
   c) Telephonic interview

Characteristics:

1. Sample size used is very large.
2. Structured questionnaire is used for data collection.

Qualitative research:

Measurability: Not possible or difficult to measure.

Features: It is a kind of exploratory research.

Characteristics:

a) Unstructured questionnaire is used.

b) Sample size is usually small

There are four major techniques in Qualitative Research. They are:

1. Depth Interview
2. Delphi Technique
3. Focus Group
4. Projective Technique

11.1.2 Depth Interview

Unstructured, direct interview is known as a depth interview. Here the interviewer will continue to ask probing questions of like, "What did you mean by that statement?", "Why did you feel this way?" and "What other reasons do you have?" etc., until he is satisfied

that he has obtained the information he wants. The unstructured interview is free from restrictions imposed by a formal list of questions. The interview may be conducted in a casual and informal manner in which the flow of the conversation determines what questions are to be asked and the order in which they should be asked.

Advantages

1. The primary advantage of the depth interview technique is its ability to discover motivations. Marketing decisions like the choice of product, methods of selling and advertising appeals etc., must be decided only after receiving feedback from consumer.
2. The second advantage of the depth interview procedure is that it encourages respondents to express any ideas they have.
3. The third advantage is that it provides a lot of flexibility to the interviewer. We have a two-way communication where both interviewer and the interviewee contribute to the knowledge gained.

Limitations

1. There are number of weaknesses in the depth interviewing approach. First of all, depth interview takes much longer than a typical structured questionnaire filling. It may lead to respondent fatigue and hence may lead to biased response.
2. The second weakness of the depth interview is the lack of systematic structure for interpretation of the information obtained. This requires a trained psycho-analyst. It is difficult to find the qualified and trained people for conducting depth interview.
3. Another difficulty is that no quantifiable data is obtained in the depth interviewing process. This means that human judgment is involved in summarising the findings. Different results will often be obtained by different people in the same situation. As a result, there is little or no opportunity for verification. Flexibility on the part of the interviewer is sometimes a major weakness.

11.1.3 Delphi Technique

This is a process where a group of experts in the field gather together. They may have to reach a consensus on forecasts. Sometimes, the judgment may be made by some group members who have strong personalities. In the Delphi approach, the group members are asked to make individual judgments about a particular subject, say 'sales forecast'. These judgments are compiled and returned to the group members, so that they can compare their previous judgment with those of others. Then they are given an opportunity to revise their judgments, especially if it differs from the others. They can say, why their judgment is accurate, even if it differs, from that of the other group members. After 5 to 6 rounds of interaction, the group members reach conclusion.

11.1.4 Focus Group Interview

They are the best known and most widely used type of indirect interviews. Here, a group of people jointly participate in an unstructured indirect interview conducted by a moderator. The group usually consists of six to ten people. In general, the selected persons have similar backgrounds. The moderator attempts to focus the discussion on the problem areas.

Focus groups are used primarily to provide background information and to generate hypothesis rather than to provide solution to problems. The areas of application include:

(1) Development of new product concept.

(2) The generation of ideas for improving established products.

(3) Development of creative concepts in advertising.

An example of the use of the focus group technique in the development of advertising may be looked at. Assume that company X wants to introduce electrical cars. Just prior to the introduction of the new car, the company conducts two focus group interviews to see "what is the dealers' perception about key benefits of the new type of car?" Assume that previous research indicated the customers would buy the new car, provided they were less expensive than the conventional cars. Since the new car was priced lower than price of a conventional car, the company expected no problems with the dealers accepting the new car. Instead, the focus group interviews found that the dealers were doubtful about the acceptance of electrical car in the Indian market, since it is new, despite the fact that it is cheaper than regular car. Customers were concerned about charging mode, facilities for doing so, battery life and above all, newness of the concept.

Advantages

This technique provides more sophisticated data because of the interaction among different members of the group.

It also offers other benefits of depth interviews and offers in addition the advantages of saving cost, time and resources during data collection stage.

Disadvantages

As the samples are small and invariably non-probabilistic, extrapolation of findings is not permitted.

Responses of individual members are not independent, and are influenced by what others have to say. Some respondents dominate the proceeding and try to force their opinion on others and some are very shy or nervous, and have very little or nothing to say, though they may feel strongly on the subject. The results of group interviews are difficult to quantify since they are unstructured.

The key component of a successful focus group interview is the skill of the group moderator (interviewer). By carefully guiding the discussion, avoiding dominance by a few group members and keeping the discussion focused on the topic of interest, a moderator can obtain valuable data from the participants.

Indirect Interviews

The direct questioning of respondents may not yield the desired results, due to the fact that the respondent is usually unable and unwilling to give accurate answers often to direct questions. To solve this, a number of techniques have been devised to obtain information by indirect means. Most of these interviewing techniques employ the principle of projection. That is, the respondent is given a non-personal, ambiguous situation and is required to respond. These techniques include the word association, sentence completion tests, interpretation of pictorial representation and other devices– that have been developed as means of inducing people to project their feelings. They have been most widely used for finding out attitude towards products, such as automobiles, soaps and detergents, cigarettes, food products and beverages. Indirect interviews are commonly referred to as projective techniques.

11.1.5 Projective Techniques

Projective techniques (Indirect method of gathering information/indirect interview) are unstructured and involve indirect form of questioning.

In projective techniques, respondents are asked to interpret the behaviour of users, rather than describe their own behaviour. In interpreting the behaviour of others, respondents indirectly project their own motivation and feelings into the situation. *Example:* Many a time, people do not want to reveal their true motive for fear of being branded 'old fashioned'. Questions such as "Do you do all household work yourself?" The answer may be 'no', though the truth is 'yes'. A 'yes' answer may not be given because it may suggest that the family is not financially sound and cannot afford a maid for help.

Two types of projective techniques are available:

1. An ambiguous stimulus is presented to respondents.
2. In reacting to the stimulus, the respondents will indirectly reveal their own feeling.

The general categories of projective techniques are:

1. Word association test
2. Completion technique
3. TAT and
4. Cartoon test

1. Word Association Test

This test consists of presenting a series of stimulus words to the respondent, who is asked to answer quickly with the first word that comes to his mind. The respondent, by answering quickly, gives the word that s(he) or she associates most closely with the stimulus word.

Example 1: What brand of detergent comes to your mind first, when I mention washing of an expensive cloth?

(1) Surf

(2) Tide

(3) Key

(4) Ariel

Example 2: Who drinks the milk most?

(1) Athletes

(2) Young boys

(3) Adults

(4) Children

Example 3: In a study of cigarettes, the respondent is asked to give the first word that comes to his mind.

(1) Injurious

(2) Style

(3) Strong

(4) Stimulus

(5) Bad manners

(6) Disease

(7) Pleasure

2. Completion Techniques

- Sentence completion
- Story completion

❖ *Sentence Completion*

Here the respondents have to finish a set of incomplete sentences. *Example:* Let us make a study dealing with people's inner feelings towards software professionals.

(a) Earnings of a software professional _____

(b) Being a software professional means _____

(c) Working hours for software professional are _____

(d) The personal life of a software professional is _____

(e) The social status of software professional is _____

Suppose you want to study the attitude towards a periodical:

(a) A person who reads *Women's Era periodical* is .....

(b) *Business World periodical* appeals to .....

(c) *Outlook periodical* is read by .....

(d) *Investor periodical* is mostly liked by .....

Suppose you want to provide a basis for developing advertising appeal for a brand of cooking oil, the following sentence may be used:

(a) People use cooking oil .....

(b) Most of the new cooking oil .....

(c) Costliest cooking oil .....

(d) The thing I enjoy about cooking oil used by my family.....

(e) One important feature to be highlighted in the advertisement about cooking oil is.....

❖ ***Story Completion:*** A situation is described to a respondent who is asked to complete the story based on his opinion and attitude. This technique will reveal the interest of the respondent, but it is difficult to interpret.

Example 1: Mr. X belongs to the upper-middle class. He received a telephone call, where the caller said that "I am from Globe Travels. Sir, I want to tell you about our recent offer, that is, if you travel to the US this summer, you will get two tickets free by the year end to fly to the Far East.

What was Mr. X's reaction? Why?

Example 2: Two children are quarreling at the breakfast table before going to school. The younger of the two, has spilled coffee on her brother's shirt which he was supposed to wear on the same day for attending annual sports event.

What will the mother do?

The story completion has numerous applications in solving marketing problem. The most important of which is to provide data to the seller, recognising the image and feelings people have about the company's products and services. This method is used before finalising an advertisement.

3. Thematic Apperception Test (TAT)

Definition: TAT is a projective technique. It is used to measure the attitude and perception of the individual. Some picture cards are shown to respondents. The respondent is required to tell the story by looking at the picture. When the subjects start telling the story, the researcher notices the respondents' expression, pauses and emotions to draw the inference.

Administration: TAT is administered to individuals in an atmosphere free from interruptions. The usual number of cards shown varies between 10 to 14. The original TAT developer Murry recommended the use of 20 cards. The original TAT consisted of 31 cards divided into 3 categories viz, for use with men or women only or for use with subjects of either sex. Of late, the use of separate set of cards for men and women has been discontinued.

The subject is then instructed to tell a story about the picture on each card with specific instructions to include a description of the event in the picture and the developments that led to the event, the thoughts and feelings of people in the picture, and the outcome of the story. The examiner keeps the cards in a pile, face down in front; gives the respondent one card at a time, and asks the respondents to place each card face down as the story is completed. Administering the TAT usually takes about an hour.

Recording: Murry's original practice was to take notes by hand on the subject's responses, including his (her) non-verbal behaviours. Research has indicated, however, that a great deal of significant material is lost when notes are recorded in this way. As a result, some examiners now use a tape recorder to record the respondent's answers. Another option involves asking the respondent to write down his (her) answers.

Interpretation: In interpreting responses to the TAT, examiners typically focus their attention on one of the three areas: the content of the stories that the respondent tells; the feeling or tone of the stories; or the respondents' behaviour apart from responses. These behaviours may include verbal remarks (for example, comments about feeling stressed by the situation or not being a good story teller) as well as non-verbal actions or signs (blushing, stammering, fidgeting in the chair, difficulties in making eye contact with the examiner, etc.). The story content usually reveals the respondents' attitudes, fantasies, wishes, inner conflicts, and views about the world outside. The story structure typically reflects the respondents' feelings, assumptions about the world, and an underlying attitude of optimism or pessimism.

The respondent is helped by asking questions such as:

(1) What is happening in the picture? (point of time)

(2) Why is it happening?

(3) What is your feeling about the character shown in the picture?

(4) Who is right?

(5) Who is the aggressor? Or who is right and who is wrong?

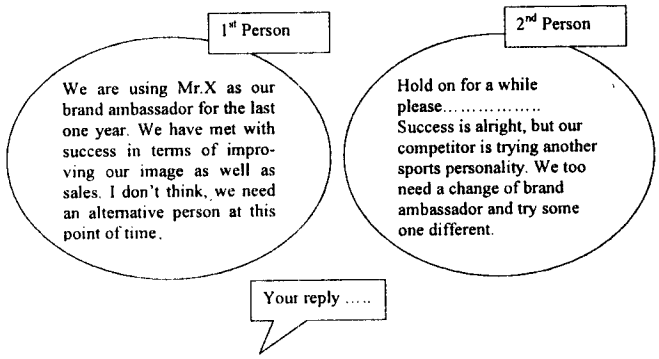
In the TAT, the test subject (the boy shown here) examines a set of cards that portray human figures in a variety of settings and situations, and is asked to tell a story about each card. The story includes the event shown in the picture, preceding events, emotions and thoughts of those portrayed, and the outcome of the event shown. The story content and structure are thought to reveal the subject's attitudes, inner conflicts, and views. Customer insights may be extracted by posing the questions given above to the respondents

***Source:** (http://www.minddisorders.com)*

Cartoon Test or Balloon Test

Here a cartoon is shown. The cartoon character belongs to a particular situation. One or more of 'balloons' include the conversation of the character, and is left open and the respondent is asked to fill in.

In comparing the cartoon technique with the direct question, let us take the example of "choosing a brand ambassador".



In the above case, with which person would you agree and why ?

11.2 CONDITIONS FOR A SUCCESSFUL INTERVIEW

An interview is an interaction between the researcher and the respondent. The purpose of the interview is to elicit information from the respondent. Interview is the most versatile and flexible method of communication. The questioning can be adapted to the situation.

planation, clarification can be requested if required. The drawback of the interview is that it is costly.

In order to complete an interview successfully, certain basic conditions should be satisfied. They are as follows:

1. Firstly, the interviewer should explain to the respondent what is expected from him. The respondent must know the subject-matter before he begins answering the researcher. If the respondent's answer is incomplete, the researcher should rephrase the question and explain the same to the respondent.

   Example: How do you rate the movie "The Devils Shock"? If the answer is 'somewhat OK', the researcher cannot reach a conclusion about the perception of the respondent over the movie, because the answer is incomplete. He should ask further questions like "Is it a horror movie? 'Was it shocking'?" etc.

2. Secondly, the researcher must make sure that respondent possess the information required at the time of conducting interview. The respondent may be facing many problems due to which he is unable to give the information.

   Example: Due to passage of time, he has forgotten, or may be the subject is sensitive and the respondent does not want to answer.

3. Thirdly, the respondents must be enthusiastic and co-operative. It is the job of the interviewer to create trust and maintain a good rapport. The interaction between the interviewer and respondent should not create a bias in the respondent's replies. Therefore, it is the job of the interviewer to motivate the respondent. If the above steps are not taken, there may be errors.

   The job of the interviewer can be classified as follows:

   i. Selecting the sample.
   ii. Fixing interviews.
   iii. Conduct interviews.
   iv. Record the answers.

Selecting the Sample

There are two elements involved in the selection of a sample:

(a) The originally-designated individuals from whom it is proposed to obtain information.

(b) Final determination of individuals from whom the information is actually obtained.

The toughest problem in fixing the sample is the difficulty in getting together the members of the universe. Unless obtained, it may not be possible to get a representative sample. For example, personal interviews made at the shopping centre, on different days of the week or different times of the day. Here, there is no control on the sample.

Fixing the Interview

After locating the sample, the next step is to organise the interview. A letter in advance is to be sent to the selected sample respondent. An ID cards have to be carried to identify him/ her- self, if the researcher is visiting an organisation. The respondent must be briefed about the subject-matter. He should be assured that confidentiality will be maintained and he is required to feel free in answering. The interviewer must put the respondent at ease and should not display any one-upmanship.

Conducting the Interview

At the commencement of the interview, start with simple and easy questions. Ask the questions in the same order as given in the questionnaire. The interviewer should not comment on the question's meaning or indicate in any way, what kind of answers are acceptable. Care must be taken with respect to, not influencing the respondent with his own ideas or thoughts or emphasizing on any particular aspect. Maintaining objectivity throughout the interview process is the key to a successful interview. Any clarification sought to various questions should be clarified by the interviewer.

Research studies say the response rate and accuracy of information is more in interview methods, when it is conducted at the shopping places.

Recording

The interviewer needs to record accurately, as answered by the respondent. If the recording is not done properly, analysis of data will be difficult. Also, the researcher cannot come again to conduct another interview with the same respondent. It is better to record the answers of the open-ended questions using a tape recorder instead of recording by hands. Quick recording can be done by using abbreviations such as A.O (any other reason) or A.E (anything else). At the end of the interview, the researcher should thank the respondent for spending his valuable time and co-operation extended.

Errors in Interviewing

i. Interview involves social relationship between two persons. Respondents adjust their conduct to what they consider to be appropriate to the situation. The interviewer should be able to establish a good rapport. If a good rapport is lacking, the respondent might answer half-heartedly. Errors can occur due to indifference of the respondent.

ii. If the interviewer is biased, i.e., he suggests the answer by stressing some part of the question or by his tone, errors can occur. The interview must be neutral and objective, if this error is to be avoided.

iii. Error can also occur due to inconsistencies in the reply of the respondent. The interviewer should bring this to the notice of the respondent. The interviewer can probe further to correct these inconsistencies.

iv. There can be errors in recording or interpreting the response. One type of error could be recording matter that were not said, as well as not recording matter that were said. Clerical mistakes, add to errors. The interviewer's experience, attitudes and opinions also affect the recorded answers.

v. Cheating by interviewers is a serious problem leading to serious errors. The most glaring example is the interviewer who fills the questions without making interviews. If respondents are difficult to locate and interview, cheating increases.

How to minimize the errors:

(a) Select and train the interviewers.

(b) Set up supervision to check the interviewers' work.

Selection of Interviewers

The following factors should be taken into account while selecting an interviewer:

- He should be a go-getter, social, freely mixing with people and should have no hesitation in contacting strangers.
- Should possess a pleasing personality.
- Must be able to speak the language of the respondent.
- Must possess good communication skills.
- Must have patience and perseverance.
- Must have flair to undertake field work.

Training of Interviewers

Interviewers may be provided with two types of training. It should be a) field b) classroom. Classroom training includes i) fixing interviews ii) asking questions iii) recording responses

iv) how to handle non-response, etc. For field training, the trainee, should accompany a senior and can be an observer initially. After sometime, he may be asked to conduct interviews. The superior has to point out his mistakes and correct the same. He must also suggest the ways and means of improving.

Setting up Supervision to Check Interviewers Work

Some of the methods adopted are:

- Check the forms filled in by the field workers.
- Make surprise visits to those places, where the field worker conducted interviews, to cross verify.
- Re-interview the respondent, but this may not be practicable.
- Send a self-addressed envelope to the respondent to verify whether the interview was held or not.
- In order to reduce cheating, the interviewer should be compensated reasonably.

SUMMARY

Qualitative techniques are used in exploratory research. Depth interview allows flexibility in gathering information, unstructured nature of the interview allows the respondent to tell whatever he wishes. The greatest advantages of depth interview is its ability to discover hidden motives. However its limitation is subjectivity. Focus group is another method used for gathering information. It is nothing but an indirect interview. However focus group suffers from certain disadvantages, such as influence of one member on others. Morever the success depends on the skill of the moderator.

Projective techniques are indirect from of questioning, various techniques are word association test, completion techniques, TAT, cartoon test etc. In word association consumer responds, with the first word that comes to his mind. In sentence competition test respondents are given a part of sentence. The remaining part has to be completed with first thoughts, that comes to mind. Necessary instructions are given to the respondent by researcher. In story telling, after reading the story, respondent is required to answer the questions given. This provides data image or feelings of the respondent. TAT is a projective technique used to measure perception of the individual. Picture cards are shown to individual and response sought.

For successful interview, interviewer need to select the sample, organize the interview and record the response correctively. However errors can occur in interviewing due to inconsistency in the reply of interviewer, training will help in reducing the error.

KEYWORDS

Depth interview
Focus group
Word association test
Disguised
Delphi technique
Projective technique
Sentence completion test
Cartoon test

REVIEW QUESTIONS

1. What is meant by qualitative techniques of data collection?
2. What are the advantages of the observation method?
3. Explain the different projective techniques in qualitative research.
4. Explain and give examples for each of the following
   a) Word association test
   b) Sentence completion
   c) Story completion
5. Distinguish between open and disguised observation.
6. What is Delphi technique
7. What is focus group interview? What are the advantages & limitations of this method?
8. What are the qualities required for a focus group moderator?
9. Why is training of interviewers important for field work?
10. What is TAT? When and why is it useful?
11. What are the conditions which are prerequisite for a successful interview?
12. What are the sources of error in interviewing?

ASSIGNMENT

Given below are some topics. In each case, indicate whether the research is qualitative or quantitative in nature. Also recommend specific techniques for each.

a) A company would like to come out with ideas to creatively communicate the benefits of a new detergent through a TV commercial.

b) Hospital authorities want to ascertain their patients' ratings of attributes like medical treatment, room service, emergency service, etc.

c) After discussing with several sales people, the sales manager suspects that the morale of the sales force is low, and wants to confirm this by using an employee morale questionnaire.

d) A firm marketing toffee has two alternative wrapper designs for the product and is wonders, which one will result in higher sales.

CHAPTER twelve

Sampling

In this chapter, the following questions are discussed:

- What is sampling?
- What are the steps involved in the sampling process?
- What are the types of sampling design?
- What are the differences between probability and non- probability sampling?
- What are the types of errors in sampling?
- How to determine the sampling size?

12.1 INTRODUCTION TO SAMPLING

12.1.1 Sample

A sample is a part of a target population, which is carefully selected to represent the population.

12.1.2 Sample Frame

Sampling frame is the list of elements from which the sample is actually drawn. Actually, sampling frame is nothing but the correct list of population. *Example:* Telephone directory, Product finder, Yellow pages.

12.1.3 Distinction between Census and Sampling

Census refers to complete inclusion of all elements in the population. A sample is a sub-group of the population.

12.1.4 When is a Census Appropriate?

1. A census is appropriate if the size of population is small. For *example:* A researcher may be interested in contacting firms in iron and steel or petroleum products industry. These industries are limited in number, so a census will be suitable.
2. Sometimes, the researcher is interested in gathering information from every individual. *Example:* Quality of food served in a mess.

12.1.5 When is Sample Appropriate?

1. When the size of population is large.
2. When time and cost are the main considerations in research.
3. If the population is homogeneous.
4. Also, there are circumstances when a census is not possible. *Example:* Reactions to global advertising by a company.

12.2 SAMPLING PROCESS

Sampling process consists of seven steps. They are:

1. Define the population.
2. Identify the sampling frame.
3. Specify the sampling unit.
4. Selection of sampling method.
5. Determination of sample size.

6. Specify sampling plan.
7. Selection of sample.

(1) ***Define the population:*** Population is defined in terms of:

(1) Elements

(2) Sampling units

(3) Extent

(4) Time.

Example: If we are monitoring the sale of a new product recently introduced by a company, say (shampoo sachet) the population will be:

1. Element - Company's product
2. Sampling unit - Retail outlet, super market
3. Extent - Hyderabad and Secunderabad
4. Time - April 10 to May 10, 2006

(2) ***Identify the sampling frame:*** Sampling frame could be (a) Telephone Directory (b) Localities of a city using the municipal corporation listing (c) Any other list consisting of all sampling units.

Example: You want to learn about scooter owners in a city. The RTO will be the frame, which provides you names, addresses and the types of vehicles possessed.

(3) ***Specify the sampling unit:*** Individuals who are to be contacted are the sampling units. If retailers are to be contacted in a locality, they are the sampling units.

Sampling unit may be husband or wife in a family. The selection of sampling unit is very important. If interviews are to be held during office timings, when the heads of families and other employed persons are away, interviewing would under-represent employed persons, and over-represent elderly persons, housewives and the unemployed.

(4) ***Selection of sampling method:*** This refers to whether (a) probability or (b) non-probability methods are used.

(5) ***Determine the sample size:*** This means we need to decide "how many elements of the target population are to be chosen?" The sample size depends upon the type of study that is being conducted. For *example:* If it is an exploratory research, the sample size will be generally small. For conclusive research, such as descriptive research, the sample size will be large.

The sample size also depends upon the resources available with the company. It depends on the accuracy required in the study and the permissible errors allowed.

(6) ***Specify the sampling plan:*** A sampling plan should clearly specify the target population. Improper defining would lead to wrong data collection.

Example: This means that, if a survey of a household is to be conducted, a sampling plan should define a "household" i.e., "Does the household consist of husband or wife or both", minors etc., "Who should be included or excluded". Instructions to the interviewer should include "How he should obtain a systematic sample of households, probability sampling non–probability sampling". Advise him on what he should do to the household, if no one is available.

(7) ***Select the sample:*** This is the final step in the sampling process.

12.3 TYPES OF SAMPLE DESIGN

Sampling is divided into two types:

Probability sampling: In a probability sample, every unit in the population has equal chances for being selected as a sample unit.

Non-probability sampling: In the non-probability sampling, the units in the population have unequal or negligible, almost no chances for being selected as a sample unit.

12.4 PROBABILITY SAMPLING TECHNIQUES

1. Random sampling.
2. Stratified random sampling.
3. Systematic sampling.
4. Cluster sampling.
5. Multi-stage sampling.

12.4.1 Random Sampling

Simple random sample is a process in which every item of the population has an equal probability of being chosen.

There are two methods used in the random sampling –

(1) Lottery method

(2) Using random number table.

(1) ***Lottery method:*** Take a population containing four departmental stores: A, B, C and D. Suppose we need to pick a sample of two stores from the population using a simple random procedure. We write down all possible samples of two. Six different combinations, each containing two stores from the population, are AB, AD, AC, BC, BD, CD. We can now write down six sample combination on six identical pieces of paper, fold the piece of paper so that they cannot be distinguished. Put them in a box. Mix them and pull one at random. This procedure is the lottery method of making a random selection.

(2) ***Using random number table:*** A random number table consists of a group of digits that are arranged in random order, i.e., any row, column, or diagonal in such a table contains digits that are not in any systematic order. There are three tables for random numbers (a) Tippet's table (b) Fisher and Yate's table (c) Kendall and Raington table.

The table for random number is as follows:

40743	39672
80833	18496
10743	39431
88103	23016
53946	43761
31230	41212
24323	18054

Example: Taking the earlier example of stores. We first number the stores.

1 A 2 B 3 C 4 D

The stores A, B, C and D have been numbered as 1, 2, 3 and 4.

We proceed as follows, in order to select two shops out of four randomly:

Suppose, we start with the second row in the first column of the table and decide to read diagonally. The starting digit is 8. There is no departmental stores with the number 8 in the population. There are only four stores. Move to the next digit on the diagonal, which is 0. Ignore it, since it does not correspond to any of the stores in the population. The next digit on the diagonal is 1 which corresponds to store A. Pick A and proceed until we get two samples. In this case, the two departmental stores are 1 and 4. The sample derived from this consists of departmental stores A and D.

In random sampling, there are two possibilities (1) Equal probability (2) Varying probability.

Equal Probability

This is also called as the random sampling with replacement.

Example: Put 100 chits in a box numbered 1 to 100. Pick one number at random. Now the population has 99 chits. Now, when a second number is being picked, there are 99 chits. In order to provide equal probability, the sample selected is being replaced in the population.

Varying probability

This is also called random sampling without replacement. Once a number is picked, it is not included again. Therefore, the probability of selecting a unit varies from the other. In our example, it is 1/100, 1/99, 1/98, 1/97 if we select four samples out of 100.

12.4.2 Systematic random sampling

There are three steps:

(1) Sampling interval K is determined by the following formula:

$$K = \frac{\text{No. of units in the population}}{\text{No. of units desired in the sample}}$$

(2) One unit between the first and K^{th} unit in the population list is randomly chosen.

(3) Add K^{th} unit to the randomly chosen number.

Example: Consider 1,000 households from which we want to select 50 units.

Calculate $K = \frac{1000}{50} = 20$

To select the first unit, we randomly pick one number between 1 to 20, say 17. So our sample begins with 17,37,57..... Please note that only the first item was randomly selected. The rest are systematically selected. This is a very popular method because we need only one random number.

12.4.3 Stratified Random Sampling

A probability sampling procedure in which simple random sub-samples are drawn from within different strata, that are, more or less equal on some characteristics. Stratified sampling are of two types:

1. ***Proportionate stratified sampling:*** The number of sampling units drawn from each stratum is in proportion to the population size of that stratum.
2. ***Disproportionate stratified sampling:*** The number of sampling units drawn from each stratum is based on the analytical consideration, but not in proportion to the size of the population of that stratum.

Sampling process is as follows:

1. The population to be sampled is divided into groups (stratified).
2. A simple random sample is chosen.

Reason for Stratified Sampling

Sometimes, marketing professionals want information about the component part of the population. Assume there are three stores. Each store forms a strata and the sampling from within each strata is being selected. The resultant might be used to plan different promotional activities for each store strata.

Suppose a researcher wishes to study the retail sales of products, such as tea in a universe of 1,000 grocery stores (Kirana shops included). The researcher can first divide this universe into three strata based on the size of the store. This benchmark for size could be only one of the following (a) floor space (b) volume of sales (c) variety displayed etc.

Size of stores	No. of stores	Percentage of stores
Large stores	2,000	20
Medium stores	3,000	30
Small stores	5,000	50
	10,000	**100**

Suppose we need 12 stores, then choose four from each strata, at random. If there was no stratification, simple random sampling from the population would be expected to choose two large stores (20% of 12) about four medium stores (30% of 12) and about six small stores (50% of 12).

As can be seen, each store can be studied separately using the stratified sample.

Selection by proportionate stratified sample

Assume that there are 60 students in a class of a management school, of this, 10 has to be selected to take part in a Business quiz competition. Assume that the class has students specializing in marketing, finance and HR stream.

The first step is to subdivide the students of the class into 3 homogeneous groups or stratify the student population, by the area in which they are specializing.

Marketing streaming			Finance Stream		HR Stream
1	32	8	11	33	34
2	36	12	13	35	37
3	40	15	17	38	39
4	43	18	20	41	42

Contd...

5	46	19	21	44	45
7	47	22	24	49	48
9	60	23	25	59	58
10	57	28	26	60	56
14	50	27	29	52	51
16	53	31	30	55	54

Second step is to calculate the sampling fraction f= n/N

n = Sample size required

N= Population size

Third step - Determine how many are to be selected from marketing stream(say n1)

n_1= 30 X 1/10=30 X 1/10

Sample to be selected from marketing strata n1= 30X1/10 = 3

Now we can select 3 numbers from among 30 numbers at random say 7,60,22

Similarly we can select n_2 n_3

n_2= 20X1/10 = 2

The 2 numbers selected at random from finance stream are 13,59

N_3= 10X1/10 =1

Stratified sampling can be carried out with:

1. Same proportion across the strata proportionate stratified sample.
2. Varying proportion across the strata disproportionate stratified sample.

Example

Size of stores	No. of stores (Population)	Sample Proportionate	Sample Disproportionate
Large	2,000	20	25
Medium	3,000	30	35
Small	5,000	50	40
	10,000	**100**	**100**

Estimation of universe mean with a stratified sample

Example

Size of stores	Sample Mean Sales per store	No. of stores	Percent of stores
Large	200	2000	20
Medium	80	3000	30
Small	40	5000	50
		10,000	**100**

The population mean of monthly sales is calculated by multiplying the sample mean by its relative weight.

$200 \times 0.2 + 80 \times 0.3 + 40 \times 0.5 = 84$

Sample Proportionate

If N is the size of the population.

n is the size of the sample.

i represents 1,2,3,.....k [number of strata in the population]

$\therefore$ Proportionate sampling

$$p = \frac{n_1}{N_1} = \frac{n_2}{N_2} = \ldots\ldots\ldots\ldots = \frac{n_k}{N_k} = \frac{n}{N}$$

$$\frac{n_1}{N_1} = \frac{n}{N} = n_1 = \frac{n}{N} \times n_1 \text{ and so on}$$

n_1 is the sample size to be drawn from stratum 1

$n_1, n_2, \ldots\ldots\ldots\ldots n_k = n$ [Total sample size of the all strata]

Illustration: A survey is planned to analyse the perception of people towards their own religious practices. The population consists of various religions, viz., Hindu, Muslim, Christian, Sikh, Jain, assuming a total of 10,000. Hindu, Muslim, Christian, Sikh and Jain consists of 6,000, 2,000, 1,000, 500 and 500 respectively. Determine the sample size of each stratum by applying proportionate stratified sampling, if the sample size required is 200.

Solution: Total population, N=10,000

Population in the strata of Hindus N_1=6,000

Population in the strata of Muslims N_2=2,000

Population in the strata of Christians N_3=1,000

Population in the strata of Sikhs N_4=500

Population in the strata of Jains N_5=500

Proportionate Stratified Sampling

$$p = \frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n_3}{N_3} = \frac{n_4}{N_4} = \frac{n_5}{N_5} = \frac{n}{N}$$

$\therefore$ Let us determine the sample size of strata N_1

$$\frac{n_1}{N_1} = \frac{n}{N} \times N_1 = \frac{200}{10,000} \times 6,000$$

$$= 20 \times 6 = 120$$

$$n_2 = \frac{n}{N} \times N_2 = \frac{200}{10,000} \times 2,000$$

$$= 40$$

$$n_3 = \frac{n}{N} \times N_3 = \frac{200}{10,000} \times 1,000$$

$$= 20$$

$$n_4 = \frac{n}{N} \times N_4 = \frac{200}{10,000} \times 500$$

$$= 10$$

$$n_5 = \frac{n}{N} \times N_5 = 10$$

$$n = n_1+n_2+n_3+n_4+n_5$$

$$= 120+40+20+10+10$$

$$= 200$$

Sample Disproportion

Let σ_i is the variance of the stratum i,

where i = 1,2,3.....k.

The formula to compute the sample size of the stratum i is the variance of the stratum i,

where size of stratum i

r_i = Sample size of stratum i

$$r_i = \frac{Ni}{N}$$

r_i = Ratio of the size of the stratum I with that of the population.

N_i = Population of stratum i

N = Total population.

Illustration: The Government of India wants to study the performance of women self help groups (WSHGs) in three regions viz. North, South and West. The total number of WSHGs is 1,500. The number of groups in North, South and West are 600, 500 and 400 respectively. The Government found more variation between WSHGs in the North, South and West regions. The variance of performance of WSHGs in these regions are 64, 25 and 16 respectively. If the disappropriate stratified sampling is to be used with the sample size of 100, determine the number of sampling units for each region.

Solutions:

Total Population N = 1,500

Size of the stratum 1, N_1 = 600

Size of the stratum 2, N_2 = 500

Size of the stratum 3, N_3 = 400

Variance of stratum 1, $\sigma=1^2$ = 64

Variance of stratum 2, $\sigma=2^2$ = 25

Variance of stratum 3, $\sigma=3^2$= 16

Sample size n = 100

Stratum Number	Size of the stratum N_i	$r_i = \frac{N_i}{N}$	σ_i	$r_i\sigma_{in}$	$r_i\sigma_{in} = \frac{r_i\sigma_{in}}{\sum_1^3 r_i\sigma_i}$
1	600	0.4	8	3.2	54
2	500	0.33	5	1.65	28
3	400	0.26	4	1.04	18
Total					100

12.4.4 Cluster sampling

The following steps are followed.

1. The population is divided into clusters.
2. A simple random sample of few clusters is selected.
3. All the units in the selected cluster are studied.

Step 1: The above mentioned cluster sampling is similar to the first step of stratified random sampling. But the two sampling methods are different. The key to cluster sampling is decided by how homogeneous or heterogeneous the clusters are.

A major advantage of simple cluster sampling is the case of sample selection. Suppose, we have a population of 20,000 units from which we wish to select 500 units. Choosing a sample of that size is a very time-consuming process, if we use Random Numbers table. Suppose, the entire population is divided into 80 clusters of 250 units each, we can choose two sample clusters (2 × 250=500) easily by using cluster sampling. The most difficult job is to form clusters. In marketing, the researcher forms clusters so that he can deal with each cluster differently.

Example

Assume there are 20 households in a locality.

Cross		Houses		
1	X_1	X_2	X_3	X_4
2	X_5	X_6	X_7	X_8
3	X_9	X_{10}	X_{11}	X_{12}
4	X_{13}	X_{14}	X_{15}	X_{16}

We need to select eight houses. We can choose eight houses at random. Alternatively, two clusters, each containing four houses can be chosen. In this method, every possible sample of eight houses would have a known probability of being chosen – i.e. chance of one in two. We must remember that in the cluster, each house has the same characteristics. With cluster sampling, it is impossible for certain random sample to be selected. For example, in the cluster sampling process described above, the following combination of houses could not occur: $X_1 X_2 X_5 X_6 X_9 X_{10} X_{13} X_{14}$. This is because the original universe of 16 houses have been redefined as a universe of four clusters. So only clusters can be chosen as a sample.

12.4.5 Multi-stage Sampling

The name implies that sampling is done in several stages. This is used with stratified/ cluster designs.

An illustration of double sampling is as follows.

The management of a newly-opened club is solicits new membership. During the first rounds, all corporates were sent details so that those who are interested may enroll. Having enrolled, the second round concentrates on how many are interested to enroll for various entertainment activities that club offers such as billiards, indoor sports, swimming, and gym etc. After obtaining this information, you might stratify the interested respondents.

This will also tell you the reaction of new members to various activities. This technique is considered to be scientific, since there is no possibility of ignoring the characteristics of the universe.

Advantage: May reduce cost, if first stage results are enough to stratify or cluster.

Disadvantage: Costs increase as more and more stages are included.

12.4.6 Area Sampling

This is a Probability sampling, a special form of cluster sampling.

Example: If someone wants to measure the sales of toffee in retail stores, one might choose a city locality and then audit toffee sales in retail outlets in those localities.

The main problem in area sampling is the non-availability of lists of shops selling toffee in a particular area. Therefore, it would be impossible to choose a probability sample from these outlets directly. Thus, the first job is to choose a geographical area and then list out outlets selling toffee. Then follows the probability sample for shops among the list prepared.

Example: You may like to choose shops which sell the brand—Cadbury dairy milk. The disadvantage of the area sampling is that it is expensive and time-consuming.

12.5 THE ADVANTAGES V/S DISADVANTAGES OF PROBABILITY SAMPLING?

Advantages:

- It is unbiased.
- Quantification is possible in probability sampling.
- Less knowledge of universe is sufficient.

Disadvantages:

- It takes time.
- It is costly.
- More resources are required to design and execute than in non-probability design.

In marketing research, non-probability sample is used due to time and budget constraints,.

12.6 NON-PROBABILITY SAMPLING TECHNIQUES

1. Deliberate sampling
2. Shopping Mall Intercept Sampling .

3. Sequential sampling
4. Quota sampling
5. Snowball sampling
6. Panel samples

12.6.1 Deliberate or Purposive Sampling

This is also known as the judgment sampling. The investigator uses his discretion in selecting sample observations from the universe. As a result, there is an element of bias in the selection. From the point of view of the investigator, the sample thus chosen may be a true representative of the universe. However, the units in the universe do not enjoy an equal chance of getting included in the sample. Therefore, it cannot be considered a probability sampling.

Example: Test market cities are being selected, based on the judgment sampling, because these cities are viewed as typical cities matching with certain demographical characteristics. Judgment sample is also frequently used to select stores for the purpose of introducing a new display.

12.6.2 Shopping Mall Intercept Sampling

This is a non-probability sampling method. In this method the respondents are recruited for individual interviews at fixed locations in shopping malls. (*Example:* Shopper's Shoppe, Food World, Sunday to Monday). This type of study would include several malls, each serving different socio-economic population.

Example: The researcher may wish to compare the responses of two or more TV commercials for two or more products. Mall samples can be informative for this kind of studies. Mall samples should not be used under following circumstances i.e., if the difference in effectiveness of two commercials varies with the frequency of mall shopping, change in the demographic characteristic of mall shoppers, or any other characteristic. The success of this method depends on "How well the sample is chosen".

Merits

1. It has a relatively small universe.
2. In most cases, it is expected to give quick results. The purpose of deliberate sampling has become a practical method in dealing with economic or practical problems.
3. In studies, where the level of accuracy can vary from the prescribed norms, this method can be used.

Demerits

1. Fundamentally, this is not considered a scientific approach, as it allows for bias.
2. The investigator may start with a preconceived idea and draw samples such that the units selected will be subjected to specific judgment of the enumerator.

12.6.3 Sequential Sampling

This is a method in which the sample is formed on the basis of a series of successive decisions. They aim at answering the research question on the basis of accumulated evidence. Sometimes, a researcher may want to take a modest sample and look at the results. Thereafter, s(he) will decide if more information is required for which larger samples are considered. If the evidence is not conclusive after a small sample, more samples are required. If the position is still inconclusive, still larger samples are taken. At each stage, a decision is made about whether more information should be collected or the evidence is now sufficient to permit a conclusion.

Example: Assume that a product needs to be evaluated.

A small probability sample is taken from among the current user. Suppose it is found that average annual usage is between 200 to 300 units. It is known that the product is economically viable only if the average consumption is 400 units. This information is sufficient to take a decision to drop the product. On the other hand, if the initial sample shows a consumption level of 450 to 600 units, additional samples are needed for further study.

12.6.4 Quota Sampling

Quota sampling is quite frequently used in marketing research. It involves the fixation of certain quotas, which are to be fulfilled by the interviewers.

Suppose, 2,00,000 students are appearing for a competitive examination. We need to select 1% of them based on quota sampling. The classification of quota may be as follows:

Example:

Classification of Samples

Category	Quota
General merit	1,000
Sport	600
NRI	100
SC/ST	300
TOTAL	2,000

Quota sampling involves the following steps:

1. The population is divided into segments on the basis of certain characteristics. Here, the segments are termed as cells.
2. A quota of unit is selected from each cell.

Advantages:

1. Quota sampling does not require prior knowledge about the cell to which each population unit belongs. Therefore, this sampling has a distinct advantage over stratified random sampling, where every population unit must be placed in the appropriate stratum before the actual sample selection.
2. It is simple to administer. Sampling can be done very quickly.
3. The necessity of the researcher going to various geographical locations is avoided and thus cost is reduced.

Limitations:

1. It may not be possible to get a "representative" sample within the quota as the selection depends entirely on the mood and convenience of the interviewer.
2. Since too much liberty is being allowed to the interviewer, the quality of work suffers if they are not competent.

12.6.5 Snowball Sampling

This is a non-probability sampling. In this method, the initial group of respondents are selected randomly. Subsequent respondents are being selected based on the opinion or referrals provided by the initial respondents. Further referrals will lead to more referrals, thus leading to a snowball sampling. The referrals will have demographic and psychographic characteristics that are relatively similar to the person referring them.

Example: College students bring in more students on the consumption of Pepsi. The major advantage of snowball sampling is that it monitors the desired characteristics in the population.

12.6.6 Panel Samples

Panel samples are frequently used in marketing research. To give an example, suppose that one is interested in knowing the change in the consumption pattern of households. A sample of households are drawn. These households are contacted to gather information on the pattern of consumption. Subsequently, say after a period of six months, the same households are approached once again and the necessary information on their consumption is collected.

12.7 THE DISTINCTION BETWEEN PROBABILITY SAMPLE AND NON-PROBABILITY SAMPLE

Probability Sample

1. Here, each member of a universe has a known chance of being selected and included in the sample.
2. Any personal bias is avoided. The researcher cannot exercise his discretion in the selection of sample items.

Examples: Random Sample, cluster sample.

Non-Probability Sample

In this case, the likelihood of choosing a particular universe element is unknown. The sample chosen in this method is based on aspects like convenience, quota etc.

Examples: Quota sampling, Judgment sampling.

12.8 ERRORS IN SAMPLING

12.8.1 Sampling Error

The only way to guarantee the minimization of sampling error is to choose the appropriate sample size. As the sample keeps on increasing, the sampling error decrease. Sampling error is the gap between the sample mean and population mean.

Example:

If a study is done amongst Maruti car-owners in a city to find the average monthly expenditure on the maintenance of car, it can be done by including all Maruti car-owners. It can also be done by choosing a sample without covering the entire population. There will be a difference between the two methods with regard to monthly expenditure.

12.8.2 Non-sampling Error

One way of distinguishing between the sampling and the non-sampling error is that, while sampling error relates to random variations which can be found out in the form of standard error, non-sampling error occurs in some systematic way which is difficult to estimate.

12.8.3 Sampling Frame Error

A sampling frame is a specific list of population units, from which the sample for a study being chosen.

Example 1

An MNC bank wants to pick up a sample among the credit card holders. They can readily get a complete list of credit card holders, which forms their data bank. From this frame, the desired individuals can be chosen. In this example, sample frame is identical to ideal population namely all credit card holders. There is no sampling error in this case.

Example 2

Assume that a bank wants to contact the people belonging to a particular profession over phone (doctors, lawyers) to market a home loan product. The sampling frame in this case is the telephone directory. This sampling frame may pose several problems: (1) People might have migrated. (2) Numbers have changed. (3) Many numbers were not yet listed. The question is "Are the residents who are included in the directory likely to differ from those who are not included"? The answer is yes. Thus in this case, there will be a sampling error.

12.8.4 Non-response Error

This occurs, because the planned sample and final sample vary significantly.

Example

Marketers want to know about the television viewing habits across the country. They choose 500 households and mail the questionnaire. Assume that only 200 respondents reply. This does not show a non-response error, which depends upon the discrepancy. If those 200 who replied did not differ from the chosen 500, there is no non-response error.

Consider an alternative. The people who responded are those who had plenty of leisure time. Therefore, it is implied that non-respondents do not have adequate leisure time. In this case, the final sample and the planned sample differ. If it was assumed that all the 500 chosen have leisure time, but in the final analysis only 200 have leisure time and not others. Therefore, a sample with respect to leisure time leads to response error.

Guidelines to increase the response rate

Every researcher likes to get maximum possible response from the respondents, and will be most delighted if cent percent respondent unfortunately, this does not happen. The non response error can be reduced by increasing the response rate. Higher the response rate, more accurate and reliable is the data. In order to achieve this, some useful hints could be as follows.

a) Intimate the respondents in advance through a letter. This will improve the preparedness.

b) Personalized Questionnaire should be accompanied by a covering letter.

c) Ensure/Assure that confidentiality will be maintained

d) Questionnaire length is to be restricted

e) Increase of personal interview, I.D card is essential to prove the bonafide.

f) Monetary incentives are gifts will act as motivator

g) Reminder / Revisits would help.

h) Send self addressed / stamped envelope to return the completed questionnaire.

12.8.5 Data Error

This occurs during the data collection, analysis of data or interpretation. Respondents sometimes give distorted answers unintentionally for questions which are difficult, or if the question is exceptionally long and the respondent may not have answer. Data errors "can also occur depending on the physical and social characteristics of the interviewer and the respondent. Things such as the tone and voice" can affect the responses. Therefore, we can say that the characteristics of the interviewer can also result in data error. Also, cheating on the part of the interviewer leads to data error. Data errors can also occur when answers to open-ended questions are being improperly recorded.

12.8.6 Failure of the Interviewer to Follow Instructions

The respondent must be briefed before beginning the interview, "What is expected"? "To what extent he should answer"? Also, the interviewer must make sure that respondent is familiar with the subject. If these are not made clear by the interviewer, errors will occur.

Editing mistakes made by the editors in transferring the data from questionnaire to computers are other causes for errors.

The respondent could terminate his/her participation in data gathering, because it may be felt that the questionnaire is too long and tedious.

12.9 HOW TO REDUCE NON-SAMPLING ERROR

1. For non-response – provide incentives such as a gift or cash. This enhances the possibility as well as incidence of response.
2. Data error: Don't ask question, which respondents cannot answer. Also, do not ask sensitive questions.
3. Train the interviewer to establish a good rapport with the respondents.

4. Avoid leading questions.
5. Pre-test the questionnaire.
6. Modify the sampling frame to make it a representative of the population.

12.10 SAMPLE SIZE DECISION

1. The first factor that must be considered in estimating sample size, is the error permissible.
2. Greater the desired precision, larger will be the sample size.
3. Higher the confidence level in the estimate, the larger the sample must be. There is a trade off between the degree of confidence and the degree of precision with a sample of fixed size.
4. The greater the number of sub-groups of interest within the sample, the greater its size must be.
5. Cost is a factor that determines the size of the sample.
6. The issue of response rate: The issue to be considered in deciding the necessary sample size is the actual number of questionnaires, that must be sent out. Calculation-wise, we may send questionnaires to the required number of people, but we may not receive the response. For *example,* we may like to obtain the family income level from a mail survey, but the researcher may not receive response from everyone. If the researcher feels the response rate is 40%, then he needs to despatch that many extra questionnaires. A low percentage of response can cause serious problems to the researcher. This is known as the non-response error.

Non-response error may be due to (1) failure to locate, (2) flat refusal

The failure to locate: People move to new destinations. However, if the sample frames used are of recent origin, this problem can be overcome.

Flat refusal: We do not know if those who did not respond hold different views or opinions from those who responded.

This implies that those who don't respond should be motivated. It can be done in any one of the following ways:

1. An advance letter informing the respondents that they will receive a questionnaire and requesting their cooperation. This will generally increase the rate of response.
2. Monetary incentive or gift given to respondents will yield a larger response rate.
3. Proper follow up is necessary after the potential respondent received the questionnaire.

Illustration: Determine the sample size if standard deviation of the population is 3.9, population mean is 36 and sample mean is 33 and the desired degree of precision is 99%.

Solution: Given $\sigma = 3.9$, $\mu = 36, \bar{x} = 33$ and z = 1% (99% precision implies 1% level of significance)

i.e. z_α = 2.576 (at 1% l.o.s)

(Table value)

We know that sample size n can be obtained using the relation

$$n = \left(\frac{z_\alpha \sigma}{d}\right)^2 \quad \text{where} \quad d = \mu - \bar{x}$$

$$n = \left(\frac{2.576 \times 3.9}{36 - 33}\right)^2 = 11.21 \simeq 11$$

Illustration: Determine the sample size if the standard deviation of population is 12 and the standard error (standard deviation of the sampling distribution) is 3.69.

Solution: Given the standard deviation of population

$$\sigma = 12$$

Standard error = $\sigma x = 3.69$

We know that

$$\sigma x = \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \sigma x^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow n = \frac{\sigma^2}{\sigma x^2} = \left(\frac{12}{3.69}\right)^2$$

$\Rightarrow$ n = 10.57 $\simeq$11

Illustration: Determine the sample size, if sample proportion p = 0.4 and standard error of proportion is 0.043.

Solution: Given that p = 0.4 $\Rightarrow$ q = 0.6 $\sigma_p = 0.043$

We know that $\sigma p = \sqrt{\frac{pq}{n}}$

$$\Rightarrow \sigma p^2 = \frac{pq}{n}$$

$$\Rightarrow n = \frac{pq}{\sigma p^2} = \frac{0.4 \times 0.6}{(0.043)^2}$$

$$= 129.79 \simeq 130$$

Illustration: Determine the sample size if the standard deviation of population is 8.66, sample mean is 45, population mean 43 and the desired degree of precision is 95%.

Solution:

Given that $\mu = 43$, $\overline{X} = 45$

$\sigma = 8.66$ z = 5% l.o.s

$\Rightarrow z_\alpha = 1.96$

We know that sample size n can be obtained using the relation

$$n = \left(\frac{z_\alpha \sigma}{d}\right)^2 \quad \text{where } d = \mu - \bar{x}$$

$$\simeq n = \left(\frac{1.96 \times 8.66}{43 - 45}\right)^2 = 72.03 \simeq 72.$$

SUMMARY

Sample is a representative of population. Census represents cent persent of population. The most important factors distinguishing whether to choose sample or census is cost and time. There are seven steps involved in selecting the sample. There are 2 types of sample (a) Probability sampling (b) Non probability sample. Probability sampling includes random sampling, stratified random sampling systematic sampling, cluster sampling, Multistage sampling. Random sampling can be chosen by Lottery method or using random number table. Samples can be chosen either with equal probability or varying probability. Random

sampling can be systematic or stratified. In systematic random sampling, only the first number is randomly selected. Then by adding a constant "K" remaining numbers are generated. In stratified sampling, random samples are drawn from several strata, which has more or less same characteristics. In multistage sampling, sampling is drawn in several stages.

KEYWORDS

Sample frame
Random sampling
Systematic sampling
Multistage sampling
Snow to all sampling
Panel sampling
Census
Stratified random sampling
Cluster sampling
Quota sampling
Deliberate sampling

REVIEW QUESTIONS

1. Distinguish census from sampling
2. What is sampling frame? Give an example.
3. What are the steps involved in the process of sampling?
4. What are the different types of sample designs?
5. What are the types of probability sampling techniques?
6. What is systemaic random sampling? How is the random start number and the sampling interval determined?
7. Explain the following
   a. Process of stratified sampling
   b. Proportionate stratified sampling
   c. Disproportionate stratified sampling
   d. Reasons for stratified sampling
8. What are the steps to be followed in the process of cluster sampling?
9. What are the advantages and disadvantages of multi stage sampling?
10. Discuss the advantages and disadvantages of probability sampling technique?

11. What is non-probability sampling technique?
12. What are the types of non-probability sampling techniques?
13. What is judgement sampling? Where is it applicable?
14. What are the merits and demerits of shopping mall intercept sampling?
15. What are the advantages and limitations of quota sampling?
16. How do you select a quota sampling?
17. Distinguish probability and non probability sampling.
18. What are the guidelines to determine the sample size of a population?
19. One mobile phone user is asked to recruit another mobile phone user. What sampling method is this known as and why?

ASSIGNMENT-1

Identify the appropriate target population and sampling frame for various situations listed below:

a. The regional marketing manager of a beverage company wants to test market three new flavours to gauge their acceptance.

b. A manufacturer wants to assess whether adequate inventories of spare parts are being maintained by the distributors to prevent shortages and loss of business.

c. A wholesaler dealing with audio/video equipments wants to evaluate the reaction of dealers towards a new promotion policy announced.

d. A TV channel wants to determine the viewing habits of housewives and their programme preferences.

e. A departmental chain such as Food World wants to determine the shopping behaviour of customers who use the credit cards.

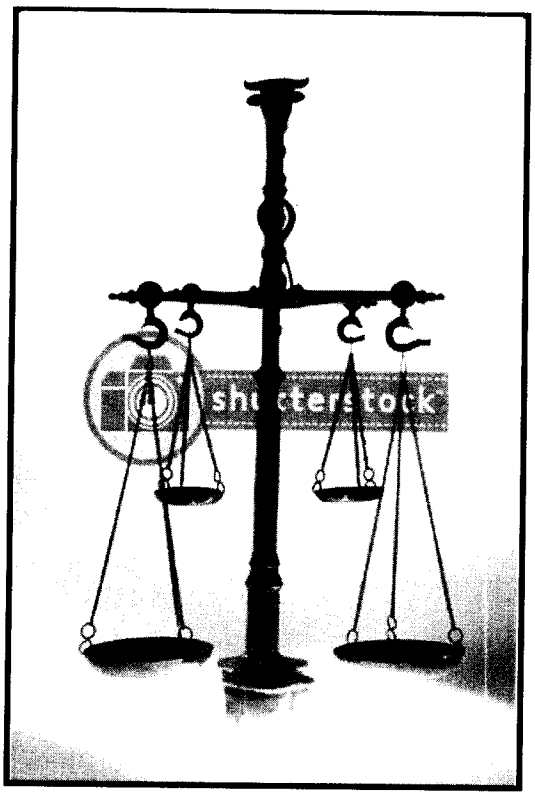
ASSIGNMENT-2

Prepare a sample plan including the sample size for Santoor soap, keeping in mind both the male and female customers. Use three economic strata, the educational level and the age group influencing the buyer behaviour. Prepare a sampling design for the following:

1. To measure the effectiveness for a TV Ad on soaps.
2. Consumer reaction to a new brand of coffee introduced.
3. To assess the market share of branded soap.

CHAPTER thirteen

Attitude Measurement and Scaling Techniques



In this chapter, the following questions are discussed:

- What are the components of attitude?
- What are the basic scaling techniques?
- What are the scales used in attitude measurement?
- What are multi-dimensional scaling techniques?
- What are the limitations of multi-dimensional scaling techniques?
- What is the criteria for a good test?

13.1 INTRODUCTION TO ATTITUDE

Meaning

Attitude is a degree of positive or negative effect associated with some psychological object. Attitudes are subjective and personal. Attitude influences the behaviour. Purchase decisions are based upon the attitudes. The attitudes can change over time.

Components of attitude: Attitude has three components, namely cognitive, affective and the behavioural.

Cognitive

This refers to the respondents' beliefs, knowledge or awareness about an event or an object. This is usually acquired from friends, periodicals etc. Sometimes, it is also known as the belief component. Statements like – (a) I am aware of the product 'X' (b) I have no idea about the product 'B' (c) That institute is excellent.

Cognitive component is very important in marketing in terms of creating awareness about the product, person etc.

Affective

This refers to the respondent's liking or preferences for an object. This is also known as the feeling component. (a) I like the product 'A' (b) Advertisement 'X' is poor. This component reveals the buyers' positive or negative attitude towards the product.

Behavioural

This refers to the respondent's intention to buy. This is a situation prior to the purchase. In marketing, the usage and buying pattern depends on this component. This is also known as action component.

13.2 DETERMINANTS OF ATTITUDE (WHAT ALTERS THE ATTITUDE?)

Attitudes are not static, but change continuously. Attitudes undergo change due to five factors:

- Information gathered in the past relating to the actual experience
- Individual perception and belief
- Exposure to new information
- Changes in the group membership
- Individual personality.

13.3 SCALING TECHNIQUES

These are of four kinds of scales, namely:

(a) Nominal scale

(b) Ordinal scale

(c) Interval scale

(d) Ratio scale

13.3.1 Nominal Scale

In this scale, numbers are used to identify the objects. For *example,* University Registration numbers assigned to students, numbers on their jerseys.

Example:

Have you ever visited Bangalore?

Yes-1

No-2

'Yes' is coded as 'One' and 'No' is coded as 'Two'. The numeric attached to the answers has no meaning, and is a mere identification. If numbers are interchanged as one for 'No' and two for 'Yes', it won't affect the answers given by respondents. The numbers used in nominal scales serve only the purpose of counting.

The telephone numbers are an example of nominal scale, where one number is assigned to one subscriber. The idea of using nominal scale is to make sure that no two persons or objects receive the same number. Similarly, bus route numbers are the example of nominal scale.

"How old are you"? This is an example of a nominal scale.

"What is your PAN Card number?"

Arranging the books in the library, subjectwise, authorwise – we use nominal scale.

Example: Physics-48, Chemistry-92 etc.

Limitations

(a) There is no rank ordering.

(b) No mathematical operation is possible.

(c) Statistical implication – Calculation of the standard deviation and the mean is not possible. It is possible to express the mode.

13.3.2 Ordinal scale (Ranking scale)

The Ordinal scale is used for ranking in most market research studies. Ordinal scales are used to ascertain the consumer perceptions, preferences etc. For *example,* The respondents may be given a list of brands which may be suitable and were asked to rank on the basis of ordinal scale.

- Lux
- Liril
- Cinthol

- Lifebuoy
- Hamam

Rank	Item	Number of respondents
I	Cinthol	150
II	Liril	300
III	Hamam	250
IV	Lux	200
V	Lifebuoy	100
Total		1,000

In the above example, II is mode and III is median.

Statistical implications: It is possible to calculate the mode and the median.

In market research, we often ask the respondents to rank the items, like for example, "A soft drink, based upon flavour or colour". In such a case, the ordinal scale is used. Ordinal scale is a ranking scale.

Rank the following attributes of 1–5 scale according to the importance in the microwave oven:

Attributes		Rank
A)	Company Image	5
B)	Functions	3
C)	Price	2
D)	Comfort	1
E)	Design	4

Difference between Nominal and Ordinal Scales

In nominal scale numbers can be interchanged, because it serves only for the purpose of counting. Numbers in Ordinal scale have meaning and it won't allow interchangeability.

13.3.3 Interval Scale

Interval scale is more powerful than the nominal and ordinal scales. The distance given on the scale represents equal distance on the property being measured. Interval scale may tell us "How far the objects are apart with respect to an attribute?" This means that the difference can be compared. The difference between "1" and "2" is equal to the difference between "2" and "3".

Example 1: Suppose we want to measure the rating of a refrigerator using interval scale. It will appear as follows:

1.	Brand name	Poor ———————	Good
2.	Price	High ———————	Low
3.	Service after-sales	Poor ———————	Good
4.	Utility	Poor ———————	Good

The researcher cannot conclude that the respondent who gives a rating of 6 is 3 times more favourable towards a product under study than another respondent who awards the rating of 2.

Example 2: How many hours you spend to do class assignment every day?

< 30 min.

30 min. to 1 hr.

1hr. to 1½ hrs.

> 1½ hrs.

Statistical implications: We can compute the range, mean, median etc.

Difference between Interval and Ordinal Scales

Ordinal scale gives only the ranking of the alternatives viz., one is greater than the other, but won't give the difference/distance between one and the other. Interval scales provide information about the difference between one and the other.

13.3.4 Ratio Scale

Ratio scale is a special kind of internal scale that has a meaningful zero point. With this scale, length, weight or distance can be measured. In this scale, it is possible to say, how many times greater or smaller one object is being compared to the other.

Example: Sales this year for product A are twice the sales of the same product last year.

Statistical implications: All statistical operations can be performed on this scale.

13.4 SCALES USED IN ATTITUDE MEASUREMENT

The following scales measure the attitude:

- Paired comparison
- Likert Scale

- Semantic Differential Scale
- Thurstone Scale

13.4.1 Paired Comparison

Example: Here a respondent is asked to show his preferences from among five brands of coffee – A, B, C, D and E with respect to flavours. He is required to indicate his preference in pairs. A number of pairs are calculated as follows. The brands to be rated are presented two at a time, so each brand in the category is compared once to every other brand. In each pair, the respondents were asked to divide 100 points on the basis of how much they liked one compared to the other. The score is totally for each brand.

$$\text{No. of pairs} = \frac{N(N-1)}{2}$$

In this case, it is $\frac{5(5-1)}{2}=2$

A&B	B&D
A&C	B&E
A&D	C&D
A&E	C&E
B&C	D&E

If there are 15 brands to be evaluated, then we have 105 paired comparison(s) and that is the limitation of this method.

13.4.2 Likert Scale

It is known as summated rating scale. This consists of a series of statements concerning an attitude object. Each statement has '5 points', *Agree* and *Disagree* on the scale. They are also called summated scales, because scores of individual items are summated to produce a total score for the respondent. The Likert Scale consists of two parts – item part and evaluation part. Item part is usually a statement about a certain product, event or attitude. Evaluation part is a list of responses like "strongly agree" to "strongly disagree". The five point-scale is used here. The numbers like +2, +1,0, -1,-2 are used. The *Likert Scale* must contain an equal number of favourable and unfavourable statements, Now, let us see with an example how the attitude of a customer is measured with respect to a shopping mall.

Evaluation of Globus—the Super Market by respondent

#	Likert scale items	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1.	Salesmen at the shopping mall are courteous	-	-	-	-	-
2.	Shopping mall does not have enough parking space	-	-	-	-	-
3.	Prices of items are reasonable.	-	-	-	-	-
4.	Mall has wide range of products to choose	-	-	-	-	-
5.	Mall operating hours are inconvenient	-	-	-	-	-
6.	The arrangement of items in the mall is confusing	-	-	-	-	-

The respondents' overall attitude is measured by summing up his (her) numerical rating on the statement making up the scale. Since some statements are favourable and others unfavourable, it is the one important task to be done before summing up the ratings. In other words, "strongly agree" category attached to favourable statement and "strongly disagree" category attached to unfavourable. The statement must always be assigned the same number, such as +2, or -2. The success of the *Likert Scale* depends on "How well the statements are generated?" The higher the respondent's score, the more favourable is the attitude. For *example,* if there are two shopping malls, ABC and XYZ and if the scores using the *Likert Scale* are 30 and 60 respectively, we can conclude that the customers' attitude towards XYZ is more favourable than ABC.

13.4.3 Semantic Differential Scale

This is very similar to the *Likert Scale.* It also consists of a number of items to be rated by the respondents. The essential difference between *Likert* and *Semantic Differential Scale* is as follows:

It uses "Bipolar" adjectives and phrases. There are no statements in the *Semantic Differential Scale.*

Each pair of adjective is separated by a seven point scale.

Some individuals have favourable descriptions on the right side, while some have on the left side. The reason for the reversal is to have a combination of both favourable and unfavourable statements.

Semantic Differential Scale Items

Please rate the five real estate developers mentioned below on the given scales for each of the five aspects. Developers are:

(1) Ansal (2) Raheja (3) Purvankara (4) Mantri (5) Salpuria

# Scale items	-3	-2	-1	0	+1	+2	+3	
1) Not reliable	_	_	_	_	_	_	_	Reliable
2) Expensive	_	_	_	_	_	_	_	Not expensive
3) Trustworthy	_	_	_	_	_	_	_	Not trustworthy
4) Untimely delivery	_	_	_	_	_	_	_	Timely delivery
5) Strong Brand Image	_	_	_	_	_	_	_	Poor brand image

The respondents were asked to tick one of the seven categories which describes their views on attitude. Computation is being done exactly the same way as in the *Likert Scale.* Suppose, we are trying to evaluate the packaging of a particular product. The seven point scale will be as follows:

"I feel .....

1. Delighted
2. Pleased
3. Mostly satisfied
4. Equally satisfied and dissatisfied
5. Mostly dissatisfied
6. Unhappy
7. Terrible."

13.4.4 Thurstone Scale

This is also known as an equal appearing interval scale. The following are the steps to construct a Thurstone Scale:

Step 1: To generate a large number of statements, relating to the attitude to be measured.

Step 2: These statements (75 to 100) are given to a group of judges, say 20 to 30, who were asked to classify them according to the degree of favourableness and unfavourableness.

Step 3: 11 piles are to be made by the judges. The piles vary from "most unfavourable" in pile 1 to neutral in pile 6 and most favourable statement in pile 11.

Step 4: Study the frequency distribution of ratings for each statement and eliminate those statements, which different judges have given widely scattered ratings.

Step 5: Select one or two statements from each of the 11 piles for the final scale. List the selected statements in random order to form the scale.

Step 6: The respondents whose attitudes are to be scaled were given the list of statements and asked to indicate their agreement or disagreement with each statement. Some may agree with one statement while some may agree with more than one statement.

Example 1: Crime and violence in movies:

1. All movies with crime and violence should be prohibited by law.
2. Watching crime and violence in movies is a waste of time.
3. Most movies with crime are bad and harmful.
4. The direction and theme in most crime movies are monotonous.
5. Watching a movie with crime and violence does not interfere with my routine life.
6. I have no opinion one way or the other, about watching movies with crime and violence.
7. I like to watch movies with crime and violence.
8. Most movies with crime and violence are interesting and absorbing.
9. Crime movies act as a knowledge bank gained by the audience.
10. People learn "how to be safe and protect oneself" by seeing a movie on crime.
11. Watching crime in a movie does not harm our lifestyle.

Conclusion: A respondent might agree with statements 8, 9 and 10. Such agreement represents a favourable attitude towards crime and violence. On the contrary, if items 1, 3, 4 are chosen by respondents, it shows that respondents are unfavourably disposed towards crime in movies. If the respondent chooses 1, 5 and 11, it could be interpreted to indicate that s(he) is not consistent in his(her) attitude about the subject.

Example 2: Suppose, we are interested in the attitude of certain socio-economic class of respondents towards savings and investments. The final list of statements would be as follows:

1. One should live for the present and not the future. So, savings are absolutely not required.
2. There are many attractions to spend the money saved.
3. It is better to spend savings than risk them in investments.
4. Investments are unsafe as the money is also blocked.
5. You earn to spend and not to invest.
6. It is not possible to save these days.

7. A certain amount of income should be saved and invested.
8. The future is uncertain and investments will protect us.
9. Some amount of savings and investments are a must for every individual.
10. One should try to save more so that most of it could be invested.
11. All savings should be invested for the future.

Conclusion: A respondent agreeing to statements 8, 9 and 11 would be considered having a favourable attitude towards savings and investments. The person agreeing with statements 2, 3 and 4 is an individual with an unfavourable attitude. Also, if a respondent chooses statements 1, 3, 7 or 9, his attitude is not considered consistent.

Merits of Thurstone Scale

1. Very reliable, if we are measuring a single attitude.
2. Used to find attitude towards issues like war, religion, language, culture, place of worship, etc.

Limitations

1. Limited use in marketing research, since it is time consuming.
2. Collecting a number of statements (100-200) makes it a very tedious job.
3. Bias on the part of the judges cannot be avoided.
4. It is an expensive method.

13.5 MULTI-DIMENSIONAL SCALING

This is used to study consumer attitudes, particularly with respect to perceptions and preferences. These techniques help identify the product attributes that are important to the customers and to measure their relative importance. Multi-Dimensional Scaling is useful in studying the following:

1. (a) What are the major attributes considered while choosing a product (soft drinks, modes of transportation)? (b) Which attributes do customers compare to evaluate different brands of the product? Is it price, quality, availability etc?
2. Which is the ideal combination of attributes according to the customer? (i.e., which two or more attributes consumer will consider before deciding to buy).
3. Which advertising messages is compatible with the consumer's brand perceptions?

This scaling is used to describe similarity and preference of brands. The respondents were asked to indicate their perception, or the similarity between various objects (products, brands, etc.) and preference among objects. This scaling is also known as perceptual mapping.

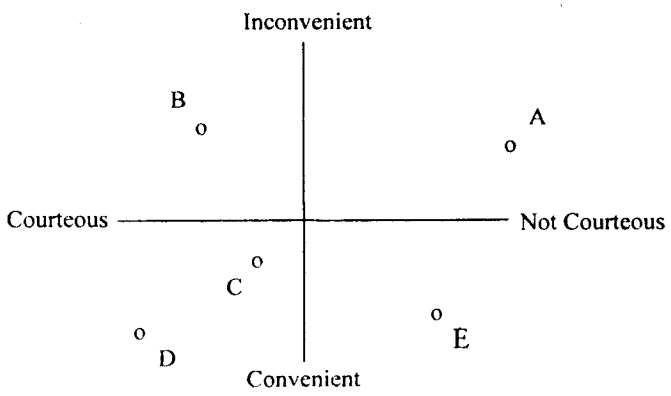
There are two ways of collecting the input data to plot perceptual mapping:

(1) Non-attribute method.

(2) Attribute method.

1. ***Non-attribute method:*** Here, the researcher asks the respondent to make a judgment about the objects directly. In this method, the criteria for comparing the objects is decided by the respondent himself.
2. ***Attribute method:*** In this method, instead of respondents selecting the criteria, they were asked to compare the objects based on the criteria specified by the researcher.

For example, to determine the perception of a consumer: Assume there are five insurance companies to be evaluated on two attributes namely (1) convenient locality (2) courteous personal service. Customers' perception regarding the five insurance companies are as follows:



A, B, C, D and E are five insurance companies.

According to the map, B & E are dissimilar insurance companies.

C is being located very conveniently.

A is a less convenient in location compared to E.

D is a less convenient in location than C.

E is a less convenient location compared to D.

RESEARCH INSIGHT

New Baby Care Product (Perceptual Mapping)

This method is particular about the steps adopted by searchers to assist a company in the new-born baby care market. The example cited is that of Marico. This map helps Marico

to identify the position of its competitors. Marico introduced a new brand baby oil named "Sparsh." This is an unorthodox entry. Marico was the first to rope in an ambassador actress in a market worth Rs. 300 crore. A second brand ambassador to speak in favour was Sonali Bendre, for both baby oil and a bathing bar. The reason for choosing a female ambassador was to lay emphasis on the concept of motherhood.

Marico is a leader in the world's largest coconut oil brand namely, Parachute. They are now switching over to health care products from hair oil and edible oil. Though adult health care products constitute a Rs. 1,500 crore market, baby care segment still continues to be a niche market. The following are some of the obstacles in developing loyalty towards the baby care products.

1. The family may use the same adult hair care product for children as well.
2. Customers repeating the product are hard to find due to the fact that women have fewer babies in present times.
3. Women stick to the product that their mothers recommend.
4. Herbal versions are still popular in urban, semi-urban and rural areas.
5. There are big players in the field of baby care products.

A few example are Johnson and Johnson, Dabur, Wipro etc. The market also has the Himalaya Drug Company which has established its own herbal baby care division.

The uniqueness of Sparsh lies in the fact that it meets consumer needs by using traditional ingredients in modern packing. Marico's main effort is to create brand differentiation.

Two parameters used by Sparsh of Marico are:

1. Price.
2. Value perception.

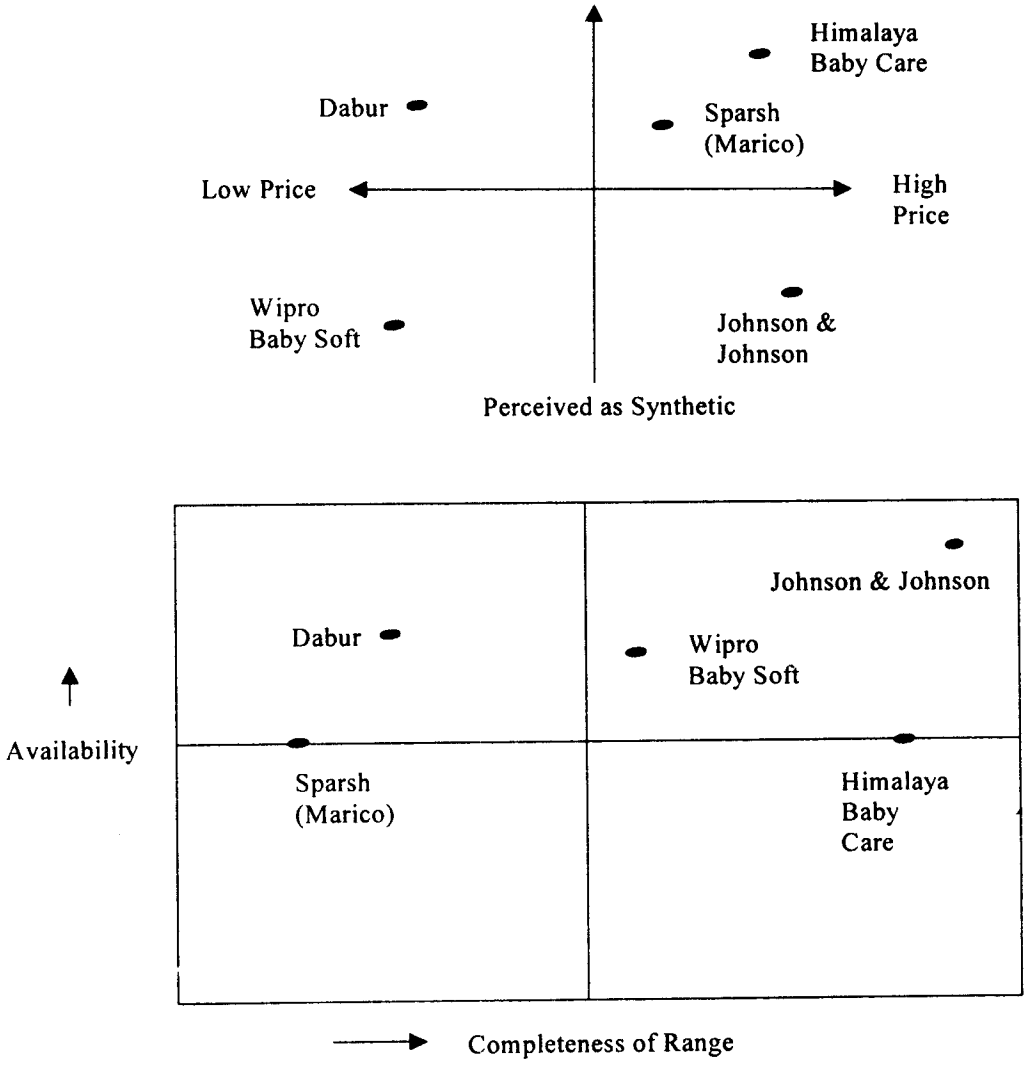
In a segment, where perceived quality governs decision-making, and value as a parameter is the choice, value proposition is central to Sparsh's marketing.

In a segment, where price is the consideration, the company has priced it on par with leaders.

In promotion, the company has used a two-pronged approach:

1. Build brand value.
2. Cut through the clutter.

Perceptual Mapping for Baby care Brands in India



Source: The Times of India

Use of Multi-Dimensional Scaling

1. To determine salient product attributes perceived by buyers in the market.
2. To know the combination of attributes buyers are likely to prefer.

3. To understand the products which are viewed as substitutes and those that are differentiated.
4. For segmenting the market.

What Tools are used in MDS?

Software such as SPSS, SAS and Excel are the packages used in MDS. Brand positioning research is one of SPSS's important features. SAS is a business intelligence software. Excel is also used to a certain extent.

Limitations of MDS

1. ***Conceptual problem:*** The criteria on which the similarities are gauged may vary during an interview with respondents. They vary depending on what the respondent thinks. A customer may buy something for himself or he may gift a product to others. In both cases, the criteria used for selection are different.
2. ***Preference:*** Keeps changing from time to time
3. Complicated computational problem.

13.6 CRITERIA FOR THE GOOD TEST

There are two criteria to decide whether the scale selected is good or not. They are:

(1) Reliability; and

(2) Validity

13.6.1 Reliability

Reliability means the extent to which the measurement process is free from errors. Reliability deals with accuracy and consistency. The scale is said to be reliable, if it yields the same results when repeated measurements are made under constant conditions. Example: Attitude towards a product or brand preference.

Reliability can be ensured by using the same scale on the same set of respondents, using the same method. However, in actual practice, this becomes difficult as:

(1) The same set of respondents may not be inclined to answer the same question again.

(2) Another point to note is that the same set of respondents may remember their previous answers to the question and give the same answers. This would erroneously indicate a greater reliability than what it is actually in reality.

(3) Another factor, that might distort the criteria of reliability is that the respondents may become extremely cautious and careful the second time, and may avoid giving accurate answers, or their attitudes might have changed during the interviewing period. In either of these cases, the use of reliability criteria will not be very helpful. If the time between test and re-test is a short one, then respondents will remember the answers and hence it will distort the reliability. On the contrary, if there is a long interval between test and re-test, there may be a change in their attitudes. Therefore, one has to choose the timing of the re-test in order to maintain reliability.

13.6.2 Validity

The paradigm of validity focused in the question "Are we measuring, what we think, we are measuring"? Success of the scale lies in measuring "What is intended to be measured?" Of the two attributes of scaling, validity is the most important.

There are several methods to check the validity of the scale used for measurement.

(i) ***Construct Validity:*** A sales manager believes that there is a clear relation between job satisfaction for a person and the degree to which a person is an extrovert and the work performance of his sales force. Therefore, those who enjoy high job satisfaction, and have extrovert personalities should exhibit high performance. If they do not, then we can question the construct validity of the measure.

(ii) ***Content Validity:*** A researcher should define the problem clearly. Identify the item to be measured. Evolve a suitable scale for this purpose. Despite these, the scale may be criticised for being lacking in content validity. Content validity is known as face validity. An example can be the introduction of new packaged food. When new packaged food is introduced, the product representing a major change in taste. Thousands of consumers may be asked to taste the new packaged food. Overwhelmingly, people may say that they liked the new flavour. With such a favourable reaction, the product when introduced on a commercial scale may still meet with failure. So, what is wrong? Perhaps a crucial question that was omitted. The people may be asked if liked the new packaged food, to which the majority might have "yes" but the same respondents were not asked, "Are you willing to give up the product which you are consuming currently?" In this case, the problem was not clearly identified and the item to be 'measured' was left out.

(iii) ***Predictive Validity:*** This pertains to "How best a researcher can guess the future performance from the knowledge of attitude score"?

Example: An opinion questionnaire, which is the basis for forecasting the demand for a product has predictive validity. The procedure for predictive validity is to first

measure the attitude and then predict the future behaviour. Finally, this is followed by the measurement of future behaviour at an appropriate time. Compare the two results (past and future). If the two scores are closely associated, then the scale is said to have predictive validity.

SUMMARY

The chapter deals with scales used to measure attitude. Measurement can be made using nominal, ordinal, interval or ratio scale. These scales show the extent of likes / dislikes, agreement / disagreement or belief towards an object. Each of the scale has certain statistical implications. There are four types of scales used in market research namely paired comparison, Likert, semantic differential and thurstone scale. Likert is a five point scale whereas semantic differential scale is a seven point scale. Bipolar adjectives are used in semantic differential scale. Thurstone scale is used to assess attitude of the respondents group regarding any issue of public interest MDS uses perceptional map to evaluate customers attitudes. The attribute or non attribute method could be used.

Last part of the chapter deals with criteria that is used to decide whether the scale chosen is good or not. Validity and reliability of the scale is verified before the scale is used for measurement. If repeated measurement gives the same result, then the scale said to be reliable. Validity refers to "Does the scale measure what it intends to measure". There are 3 methods to check the validity which type of validity is required depends on " What is being measured".

KEYWORDS

Nominal scale
Ordinal scale
Interval scale
Ratio scale
Paired comparison
Likert scale, Bipolar adjective
Thustone scale
Semantic differential
Non attribute method
Attribute method
Multi dimension scaling
Reliability
Perceptual map
Construct validity
Content validity
predictive validity
Internal validity

REVIEW QUESTIONS

1. What are the 4 types of scales used to measure attitude?
2. What is a paired comparison scale?
3. What are the statistical implication of various scales?
4. Identify the type of scale, you will use in each of the following (ordinal, nominal, internal, Ratio). Justify your answer.
5. Which one of the following courses, have you taken as specialization
   i. a. M.R c. Advertising Management
   ii. b. Sales Management d. Consumer Behaviour
6. On an average, how many cups of tea do you drink in a day?
   i. a. One cup b. 3 cups
   ii. c. 4 cups d. More than 4
7. Are you satisfied with outlook Magazine
   i. a. Very satisfied b. Dissatisfied
   ii. c. Somewhat satisfied d. Satisfied
8. What is your total house hold income?
9. Choose the soap of your choice and preference (from best to worst)
   i. a. Cintal b. Breeze
   ii. c. Hamam d. Lux
   iii. e. Pears f. Dove
10. Explain the construction of
    a. Likert scale
    b. Semantic differential scale
    c. Thurstone scale
11. What is forced and unforced scale?
12. What is attribute and non attribute method in scaling?
13. What is M.D.S? And what are the limitations?
14. What are the different types, sources and characteristics of hypothesis

ASSIGNMENT-1

1. Do you agree or disagree with the following?
   a. Validity is more crucial than reliability.
   b. Content validity is a more difficult type of validity to measure.
   c. A valid measurement is reliable, but a reliable measurement may not be valid.
2. What can be measured about the objects listed below?
   a. Packaged food..
   b. Output of a manufacturing unit.
   c. Job satisfaction.
   d. Subordinate staff in a college.

ASSIGNMENT-2

Suppose a cosmetic manufacturing company wants to ascertain the perception of its customers towards a product. Take the 7-item scale to measure the perceived perception of the product using Likert and Semantic Differential scales. The following are some of the likely adjectives which are used in Semantic Differential scale:

Unfavourable	X	Favourable
Soft	X	Hard
Organised	X	Disorganised
Quick	X	Slow
Formal	X	Informal
Pleasure	X	Displeasure
Complex	X	Simple
Cheap	X	Costly
Pleasant	X	Unpleasant
Fragrant	X	Less fragrant
Dominating	X	Submissive
Rational	X	Emotional

The normal categories for Semantic Differential will be as follows:

1. Most strongly agree
2. Strongly agree
3. Agree
4. Undecided
5. Disagree
6. Strongly disagree
7. Most strongly disagree.

ASSIGNMENT-3

A consumer durable company manufacturing TV sets is trying to measure consumer attitudes towards a product. For this purpose, the company wants the customer to complete a questionnaire which indicates several product attributes. It was decided by the company that only five attributes that affect the sale of the product would be considered for analysis. The attributes were appearance, quality of the picture, sound, after-sales service and price. The following scales are used to assess the product:

	Strongly disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly agree
1. Appearance is good	___	___	___	___	X
2. Price is reasonable	___	___	___	X	___
3. After-sales service is good	___	___	___	X	___
4. Sound quality is excellent	___	X	___	___	___
5. Picture is sharp	___	___	___	___	X
	1	2	3	4	5

Suppose the customer has inspected the product and the response is as shown in the table above:

1. What is the total score?
2. What according to you is the attitude of the customer? Is it favourable or unfavourable?

ASSIGNMENT-4

The above TV manufacturer is in the market for the past eight years. A survey conducted in the past by an MR agency produced the following score using *Likert Scale*. The data for various years is as below:

2000	-	18
2001	-	16
2002	-	17
2003	-	18
2004	-	20

1. What do you conclude about the customers' attitude? Is it favourable or unfavourable?

ASSIGNMENT-5

A manufacturer of packed bakery items wants to evaluate customer attitudes toward his product brand. 300 customers who buy this brand filled the questionnaire that was sent to them. The answers of this questionnaire were converted to scale and the results are as follows:

1. The average score from the above sample on a 10-item *Likert Scale* was 65.
2. Average score for a sample on 10-item *Semantic Differential Scale* was 60.

You are required to indicate whether these customers had a favourable or unfavourable attitude towards the product.

CHAPTER **fourteen**

Data Processing and Interpretation

In this chapter, the following questions are discussed:

- What are the steps involved in processing the data?
- How to edit and code the data collected?
- What is tabulation and what are the kinds of tabulation?
- How to summarise the data?
- How to make classification of data?
- How to apply central tendency techniques in MR?
- What is data interpretation?

14.1 DATA PROCESSING

Processing data is very important in market research. After collecting the data, the next task of the researcher is to analyse and interpret the data. The purpose of analysis is to draw conclusions. There are two parts in processing the data:

(1) Data analysis

(2) Interpretation of data

Analysis of the data involves organising the data in a particular manner. Interpretation of data is a method for deriving conclusions from the data analysed. Analysis of data is not complete, unless it is interpreted.

14.2 STEPS IN PROCESSING OF DATA

1. Preparing raw data
2. Editing
3. Coding
4. Tabulation
5. Summarising the data
6. Usage of statistical tools.

14.2.1 Preparing Raw Data

Data collection is a significant part of market research. Even more significant is to filter out the relevant data from the mass of data collected. Data continues to be in raw form, unless they are processed and analysed.

Primary data collected by surveys and observations by field investigations are hastily entered into questionnaires. Due to the pressure of interviewing, the researcher has to write down the responses immediately. Many times this may not be systematic. The information so collected by field staff is called raw data.

The information collected may be illegible, incomplete and inaccurate to a considerable extent. Also the information collected will be scattered in several data collection formats. The data lying in such a crude form are not ready for analysis. Keeping this in mind, the researcher must take some measures to organise the data so that it can be analysed.

The various steps which are required to be taken for this purpose are (a) editing and (b) coding and (c)tabulating.

14.2.2 Editing

The main purpose of editing is to eliminate errors and confusion. Editing involves inspection and correction of each questionnaire. The main role of editing is to identify commissions, ambiguities and errors in response.

Editing thus means the activity of inspecting, correcting and modifying the correct data. This can be done in two stages (a) Field editing (b) Office editing.

(a) ***Field editing:*** Objectives of field editing are – To make sure that proper procedure is followed in selecting the respondent, interview them and record their responses. In field editing, speed is the main criteria, since editing should be done when the study is still under progress. The main problems faced in field editing are:

(1) Inappropriate respondents

(2) Incomplete interviews

(3) Improper understanding

(4) Lack of consistency

(5) Legibility

(6) Fictitious interview

Example:

1. ***Inappropriate respondents:*** It is intended to include house owners in the sample for conducting the survey. If a tenant is interviewed, it would be wrong.

2. ***Incomplete interview:*** All questions are to be answered. There should not be any 'blanks'. Blanks can have different meanings, like (a) No answer (b) Refusal to answer (c) Question not applicable (d) Interviewer by oversight did not record. The reason for no answer could be that the respondent really does not know the answers. Sometimes, the respondent does not answer, may be because of the sensitive or emotional aspect of the question.

3. ***Lack of understanding:*** The interviewer, in a hurry, would have recorded some abbreviated answer. Later at the end of the day, s(he) cannot figure out what it meant.

4. ***Consistency:*** The earlier part of the questionnaire indicates that there are no children and in the later part the age of children is mentioned.

5. ***Legibility:*** If what is said is not clear, the interviewer must clarify the same on the spot.

6. ***Fictitious interview:*** This amounts to cheating by the interviewer. Here, the questionnaires are filled without conducting interviews. A surprise check by superiors is one way to minimise this.

(b) ***Office editing:*** Office editing is more thorough than field editing. The job of an office editor is more difficult than that of the field editor. In case of a mail questionnaire there are no other methods of cross-verification, except to conduct office audit.

Examples as below illustrate the kind of problems faced by office editors. Problems of consistency, rapport with respondents, etc., are some of the issues which get highlighted during office editing.

Example:

1. A respondent indicated that he doesn't drink coffee, but when questioned about his favourite brand, he replied "Bru".
2. A rating scale given to a respondent states Semantic Differential Scale with 10 items. The respondent has ticked "strongly agree" to the 10 items.
3. "What is the most expensive purchase you have made in the last one year?" is the question. Two respondents answer (1) LCD TV and (2) Trip to USA.

In example-1 above, there is inconsistency. There are two possibilities which an editor need to consider. (1) Was the respondent lying? (2) Did the interviewer record wrongly? The editor has to look into the answers to other questions on beverages, and interpret the right answer.

In example-2 above, it is to be remembered that Semantic Differential scale consists of items which has alternately positive and negative connotations. If a respondent has marked both positive and negative as 'agreed', the only conclusion the editor can draw is that the respondent filled the questionnaire without knowledge. The editor will have to discard this questionnaire, since there are no alternatives.

In example-3 above, both the respondents have answered correctly. The frame of reference is different. The main problem is, one of them is a product, whereas the other is a service. While coding the data, the two answers should be put under two different categories.

Answers to open-ended questions pose great difficulty in editing.

14.2.3 Coding

Coding refers to those activities which helps in transforming edited questionnaires into a form that is ready for analysis. Coding speeds up the tabulation while editing eliminates errors. Coding involves assigning numbers or other symbols to answers so that the responses can be grouped into limited number of classes or categories.

Examples: 1 is used for male and 2 for female.

Some guidelines to be followed in coding which is as follows:

1. Establishment of appropriate category.
2. Mutual exclusivity.
3. Single Dimension.

1. Establishment of appropriate category:

Example: Suppose the researcher is analysing the "inconvenience" that a car owner is facing with his present model. Therefore, the factor chosen for coding may be "inconvenience". Under this there could be 4 types (1) Inconvenience in entering the backseat (2) Inconvenience due to insufficient legroom (3) Inconvenience with respect to the interior (4) Inconvenience in door locking, and opening the dickey. Now the researcher may classify these four answers based on internal inconvenience and other inconveniences referring to the exterior. Each is assigned a different number for the purpose of codification.

2. Mutually exclusive:

This is important because the answer given by the respondent should be placed under one category. Example: Occupation of an individual may be responded to as (1) Professional (2) Sales (3) Executive (4) Manager etc.

Sometimes, respondents might think that they belong to more than one category. This is because a sales personnel may be doing a sales job and therefore should be placed under the sales category. Also, he may supervise the work of other sales executive(s). In this case, he is doing a managerial function. Viewed in this context, he should be placed under the managerial category, which has a different code. Therefore, he can only be put under one category, which is to be decided. One way of deciding this could be to analyse "in which of the two functions does he spend most time"?

Yet another scenario assumes that there is a salesman who is currently employed. Under the column of 'occupation', he will tick it as sales, while under the current employment column, he will mark unemployed. How does one codify this? Under which category should he be placed. One of the solutions is to have a classification, such as employed salesman, unemployed salesman to represent the two separate categories.

Questions	Answers	Codes
1. Do you own a vehicle	Yes	1
	No	2
2. What is your occupation	Salaried	S
	Business	B
	Retired	R
	Technical	T
	Consultant	C

14.2.4 Tabulation

Tabulation refers to counting the number of cases that fall into various categories. The results are summarized in the form of statistical tables. The raw data is divided into groups and sub-group(s). The counting and placing of data in a particular group and sub-group are done. The tabulation involves:

(1) Sorting and counting

(2) Summarising of data

Tabulation may be of two types (1) simple tabulation (2) cross tabulation. In simple tabulation, a single variable is counted. Cross-tabulation includes two or more variables, which are treated simultaneously. Tabulation can be done entirely by hand or by machine, or by both hand and machine.

The form in which tabulation is to be done is decided by taking into account (1) the purpose of study and (2) the use of statistical tools e.g. mean, mode, standard deviation etc. Improper tabulation may create difficulties in the use of these tools.

14.2.4.1 Sorting and Counting of Data

Sorting by manual method is as follows:

Sorting of data

Income (Rs.)	Tally Mark	Frequencies
1,000	𝍸	5
1,500	𝍸 𝍸	8
2,000	𝍸 𝍸 𝍸	12
2,500	𝍸 𝍸 𝍸	16

The above method is used commonly for sorting of data.

The tabulation may include table number, title, head note, stub, caption, sub-entries, body of the table, footnote and the source. The following example explains the component of a table.

Format of a Blank Table

TABLE No.

TITLE – Number of children per family

Head Note – Unit of measurement

Sub Heading	Caption	Total
	Body	
	Foot note	

The table must have a clear and brief title. The head note, usually the measurement unit, is placed at the top of the table in the right hand corner in a bracket.

Stub indicates the row title or the row headings and is placed in the left-hand column. Caption indicates what each column is meant for.

Sub-entries are the sub-group of the stub. The body of the table given full information of the frequency.

14.2.4.2 Kinds of Tabulation

1. Simple or one-way tabulation

The multiple choice questions which allow only one answer may use one-way tabulation or univariate. The questions are pre-determined and consist of counting the number of responses falling into a particular category and calculate the percentage. There may be two types of univariate tabulation:

(a) Question with only one response.

(b) Multiple responses to question

(a) ***Question with only one response:*** If the question has only one answer, the tabulation may be of the following type:

Table No. 1

Study of number of children in a family

No. of children	Family	Percentage
0	10	5
1	30	15
2	70	35
3	60	30
4	20	10
More than 4	10	5
	200	100

(b) ***Question with multiple response:*** Sometimes, respondents may give more than one answer to a given question. In this case, there will be an overlap, and responses when tabulated, need not add to 100 per cent.

Table No. 2
Choice of an automobile

What do you dislike about the car which you own at present?

Parameter	No. of respondents
Engine	10
Body	15
Mileage	15
Interior	06
Colour	18
Maintenance frequency	16
Inconvenience	20

There is duplication because respondents may be dissatisfied with the mileage given by vehicle and may dislike interior of the car. Here, there are more than one parameters to dislike the car by owner. Suppose we are tabulating the cause of inconvenience felt by the car owner, it can be classified as follows:

1. Cramped legroom.
2. Rear seat problem.
3. Difficulty in raising the window.
4. Difficult in locking the door.

Now, the tabulation of each of the specific factors would help to identify the real reason for dislike.

2. Cross tabulation or two-way tabulation

This is known as Bivariate Tabulation. The data may include two or more variables. Cross tabulation is very commonly used in market research.

Example: Popularity of a health drink among families having different incomes. Suppose 500 families are contacted and data collected is as follows:

Table No. 3
Use of health drink

Income per month	No. of children per family							No. of families
	0	1	2	3	4	5	More than 5	
<1000	5	0	8	9	11	15	25	73
1001-2000	10	5	8	10	13	18	27	91
2001-3000	20	10	12	14	20	22	32	130
3001-4000	12	3	6	7	13	20	30	91
4001-5000	6	2	6	5	10	15	20	64
> 5000	6	1	4	5	7	10	18	51
	59	21	44	50	74	100	152	500

Note: The above table shows that consumption of a health drink not only depends on income but also on the number of children per family.

Health drinks are also very popular among the family with no children. This shows that even adults consume this drink. It is obvious from the table that 59 out of 500 families consume health drinks even though they have no children. The table also shows that families in the income group of 2001 to 3000 consume health drinks the most.

14.2.5 Summarising the Data

Before taking up summarising, the data should be classified into (1) Relevant data, and (2) Irrelevant data. During the field study, the researcher collects lot of data which he may think would be of use. Summarizing the data includes:

(1) Classification of data

(2) Frequency distribution

(3) Use of appropriate statistical tool.

Classification of Data

(a) ***Number of groups:*** The number of groups should be sufficient to record all possible data. The classification should not be too narrow. If it is too narrow, there can be an overlap.

Example: If a researcher is conducting a survey on "Why does the current owner dislikes the car?" The car owner may indicate the following:

(1) Difficulty in seeking entry to the back seat

(2) Interior space

(3) Cramped leg room

(4) Mileage

(5) Rattling of the engine

(6) Dickey space

Now the above data can be classified into two or three categories such as (1) Discomfort (2) Expense (3) Pride (4) Safety (5) Design of the car.

(b) ***Width of the class interval:*** Class interval should be uniform and should be of equal width. This will provide consistency in the data distribution.

(c) ***Exclusive categories:*** The classification should be done in such a way that the response can be placed in only one category.

Example: Problem of leg room is the answer by respondent. This should be placed either under discomfort or design, but not both.

(d) ***Exhaustive categories:*** This should be made to include all responses including "Don't Know" answers. Sometimes this will influence the ultimate answer to the research problem.

(e) ***Avoid extremes:*** Avoid open-ended class interval.

14.2.6 Usage of Statistical Tools

Frequency Distribution

Frequency distribution simply reports the number of responses that each question receives. Frequency distribution organises the data into classes or groups. It shows the number of data that falls into particular class.

Example:

Income	No. of people
4000-6999	100
7000-9999	122
10000-12999	140

In marketing research, central value or **tendency plays a very important role.** The researcher may be interested in the average sales/shop, average consumption per month etc. The population parameters can be calculated with the help of simple average. The average of sample may be taken as population parameter. For *example,* if the average income of the population is to be computed, the researcher may select a sample, collect data on family income and calculate the relevant statistics which will be a representative of the population.

The total purchasing power of the community can be estimated on sample average. If the sample is stratified, the purchasing power of each income class may also be estimated. The median figure will reveal that half the population has more income than the median income, and the others half has less income than the median income. The mode will reveal the most common frequency. Based on this, shoppers can devise their strategy to sell the product.

The three most common ways to measure centrality or **central tendency are the mode, median and mean.**

Mode

The mode is the central value or item that occurs most often, when data is categorized in a frequency distribution, it is very easy to identify the mode, since the category in which the mode lies has the greatest number of observations.

Example: Data regarding household income of 300 people as tabulated by the researcher.

Income (Rs.)	Number (f)	Cumulative Frequency
upto 10000	30	30
10000-14999	125	155
20000-24999	50	205
25000-29999	30	235
30000-34999	33	268
35000-49999	20	288
above 35000	12	300

In the above table, 125 is the modal class.

Mode can be calculated using the formula:

$$M_0 = LM_0\left[\frac{D_1}{D_1+D_2}\right]\times i$$

LM_0 = Lower limit of modal class.

D_1 = Difference between the frequency of modal class and the class immediately preceding the modal class.

D_2 = Difference between the frequency of the modal class and the class immediately succeeding the modal class.

i = size of the modal class interval.

$$M_d = 10{,}000+\left(\frac{95}{95+75}\right)\times 5{,}000$$

substitute the values

$$= 10000 + \left(\frac{95}{170}\right)$$

= 5000 = 1000+2794 = 12794 Rs.

Conclusion: The majority have the income of Rs. 12,794. This is how statistical techniques are used in MR application.

Median

Median lies precisely halfway between the highest and lowest values. It is necessary to arrange the data into ascending or descending order before selecting the median value. For the ungrouped data with an odd number of observations, the median would be the middle value. For an even number of observations, the median value is half way between central value.

For a grouped data, the median is calculated using the formula:

$$M_d = LM_d \frac{\left(\frac{N}{2} - C.F\right)}{fM_d} \times i$$

M_d = Lower limit of median class.

CF = Cumulative frequency for the class just below the median class.

FM_d: Frequency of the median class.

i = Size of the class interval of median class.

In the table N = 300 N/2 = 150. The class containing the 150th person is the median class.

Substitute the value, we get median M_d = 21568.

Conclusion: Half of the population has income > Rs. 21,568' and half of the population has income < Rs. 21,568.

Mean:

In a grouped data, the midpoint of each category would be multiplied by the number of observation in that category. Sum up and the total to be divided by the total number of observation.

$$\text{Eqn., } X = \sum\left(\frac{fx}{\sum f}\right)$$

Example: Two students X, Y attend 3 class tests and the scores are as follows:

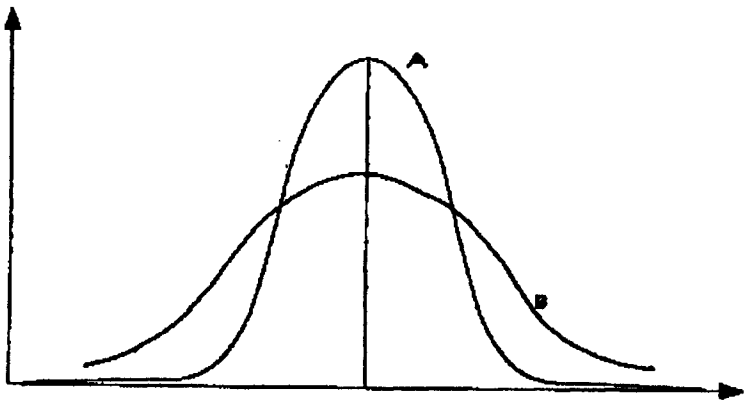
Marks		1st Test	2nd Test	3rd Test	Mean	
X		55%	60%	65%	60%	
Y		65%	60%	55%	60%	
Conclusion	X	- has improved				
	Y	- has deteriorated				

Though Mean is the same, X is better than Y.

14.2.7 Measures of Dispersion

Introduction

Dispersion is the spread of the data in a distribution. A measure of dispersion:



Mean

Indicates the degrees of the scatteredness of the observations. Let curves A and B represent two frequency distributions. Observe that A and B have the same mean. But curve A has less variability than B.

If we measure only the mean of these two distributions, we will miss an important difference between A and B. To increase our understanding of the pattern of the data, we must also measure its dispersion.

Measures of Dispersion

Range : It is the difference between the highest and lowest observed values.

i.e. range = H – L, H = Highest, L = Lowest.

Note 1 : Range is the crudest measure of dispersion.

2 : $\frac{H-L}{H+L}$ is called the coefficient of range.

Semi – Inter Quartile Range (Quartile deviation): Semi - Inter quartile range Q.

Q is given by $Q = \frac{Q_3 - Q_1}{2}$

Note 1: $\frac{Q_3 - Q_1}{Q_3 + Q_1}$ is called the **coefficient of quartile deviation.**

2: Quartile deviation is not a true measure of dispersion but only a distance of scale.

Mean Deviation (MD): If A is any average then mean deviation about A is given by:

$$MD(A) = \frac{\sum f_i |x_i - A|}{N}$$

Note 1: Mean deviation about mean $MD(\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{N}$

2: Of all the mean deviations taken about different averages mean derivation about the median is the least.

3: $\frac{MD(A)}{A}$ is called the coefficient of mean deviation.

Variance and Standard Deviation

Variance (σ^2): A measure of the average squared distance between the mean and each term in the population.

$$\sigma^2 = \frac{1}{N}\sum f_i(x_i - \bar{x})^2$$

Standard deviation (σ) is the positive square root of the variance:

$$\sigma = \sqrt{\frac{1}{N}\sum f_i(x_i - \bar{x})^2}$$

$$\sigma^2 = \frac{1}{N}\sum f_i(x_i^2 - (\bar{x})^2$$

Note: Combined variance of two sets of data of N_1 and N_2 items with means x_1 and x_2 and standard deviations σ_1 and σ_2 respectively is obtained by:

$$\sigma^2 = \frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_1d_1^2 + N_1d_2^2}{N_1 + N_2}$$

Where $d_1^2 = (x - x_1)^2$ $d_2^2 = (x - x_2)^2$

and $x = \frac{N_1\bar{x_1} + N_2\bar{x_2}}{N_1 + N_2}$

Sample variance (s^2) : Let $\bar{x}_1$, $\bar{x}_2$, x_3,..... x_n, represent a sample with mean x.

Then sample variance s^2 is given by:

$$s^2 = \frac{\sum(x - \bar{x})^2}{n-1}$$

$$= \frac{\sum x^2}{n-1} - \frac{n(\bar{x})^2}{n-1}$$

Note : $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{\sum x^2}{n-1} - \frac{n(\bar{x})^2}{n-1}}$ is called the sample standard deviation.

Coefficient of Variation (C.V)

It is a relative measure of dispersion that enables us to compare two distributions. It relates the standard deviation and the mean by expressing the standard deviation as a percentage of the mean.

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

Note: 1. Coefficient of variation is independent of the unit of the observation.

2. This measure cannot be used when x is zero or close to zero.

Illustration: For the data

103, 50, 68, 110, 105, 108, 174, 103, 150, 200, 225, 350, 103 find the Range, Coefficient of range and coefficient of quartile deviation.

Solution: Range = H – L = 350 – 50 = 300

$$\text{Coefficient of range} = \frac{H-L}{H+L} = \frac{300}{350+50} = 0.7$$

To find Q_1 and Q_3 we arrange the data in ascending order:

50, 68, 103, 103, 103, 103, 105, 108, 110, 150, 174, 200, 225, 350,

$$\frac{n+1}{4} = \frac{14}{4} = 3.5,$$

$$\frac{3(n+1)}{4} = 10.5$$

$$\therefore Q_1 = 103 + 0.5\,(103 - 103) = 103$$

$$Q_3 = 174 + 0.5\,(200 - 174) = 187$$

$$\text{Coefficient of QD} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$= \frac{84}{290} = 0.2896$$

Illustration: Calculate coefficient of mean deviation about:

(i) Median (ii) mean from the following data

X	14	16	18	20	22	24	26
f	2	4	5	3	2	1	4

X	F	cf	fx	$\lvert x - \bar{x} \rvert$	$\lvert x - M \rvert$	$f\lvert x - \bar{x} \rvert$	$f\lvert x - M \rvert$
14	2	2	28	5.71	4	11.42	8
16	4	6	64	3.71	2	14.84	8
18	5	11	90	1.71	0	8.55	0
20	3	14	60	0.29	2	0.87	6
22	2	16	44	2.29	4	4.58	8
24	1	17	24	4.29	6	4.29	6
26	4	21	104	6.29	8	25.16	32
	21		**414**		**69.71**		**68**

$$\bar{x} = \frac{\sum f_i x_i}{N} = \frac{414}{2i} = 19.71$$

$$\frac{N+1}{2} = \frac{22}{2} = 11 \text{ \Median } M = 18$$

$$\text{Now (i) M.D (x)} = \frac{\sum f_i \lvert x_i - \bar{x} \rvert}{N} = \frac{69.71}{21} = 3.32$$

$$\text{Coefficient of MD(x)} = \frac{MD(\bar{x})}{\bar{x}} = \frac{3.32}{19.71} = 0.16$$

$$\text{(ii) M.D (M)} = \frac{\sum f_i \lvert x_i - M \rvert}{N} = \frac{68}{21} = 3.24$$

$$\text{Coefficient of MD(M)} = \frac{MD(M)}{M} = \frac{3.24}{18} = 0.18$$

Illustration: A purchasing agent obtained a sample of incandescent lamps from two suppliers. He had the sample tested in his laboratory for length of life with the following results.

Length of light in hours	Sample A	Sample B
700 – 900	10	3
900 – 1100	16	42
1100 – 1300	26	12
1300 – 1500	8	3

Which company's lamps are more uniform?

Solution:

Table 1

Class interval	Sample A	Midpoint x	$u = \frac{x-1000}{200}$	fu	fu^2
700 – 900	10	800	- 1	- 10	10
900 – 1100	16	1000	0	0	0
1100 – 1300	26	1200	1	26	26
1300 – 1500	8	1400	2	16	32
	60			32	68

$$\bar{u}_A = \frac{32}{60} = 0.533$$

$$\bar{x}_A = 1000 + 200\,\bar{u}$$

$$\therefore \bar{x}_A = 1000 + 200\,(0.533) = 1106.67$$

$$\sigma_u^2 = \frac{1}{N}\sum fu^2 - (\bar{u})^2 = \frac{68}{60} - (0.533)^2$$

$$= 1.133 - 0.2809$$

$$\sigma_u^2 = 0.8524$$

$$\sigma_u = 0.9233$$

$$\sigma_x = 200 \times 0.9233 = 184.66$$

$$\therefore \text{ CV for sample A} = \frac{\sigma_A}{\bar{x}_A} \times 100$$

$$= \frac{184.66}{1106.67} \times 100 = 16.68\%$$

Table 2

Class interval	Sample B	Midpoint x	$u = \frac{x-1000}{200}$	fu	fu^2
700 – 900	3	800	- 1	- 3	3
900 – 1100	42	1000	0	0	0
1100 – 1300	12	1200	1	12	12
1300 – 1500	3	1400	2	6	12
	60			15	27

$$\bar{v} = \frac{15}{60} = 0.25$$

$$\bar{x}_B = 1000 + 200\ \bar{v}$$

$$= 1000 + 58$$

$$\therefore \bar{x}_B = 1058$$

$$\sigma_v^2 = \frac{1}{N}\sum fv^2 - (\bar{v})^2 = \frac{27}{60} - (0.25)^2$$

$$= 0.45 - 0.0625$$

$$\sigma_v^2 = 0.3875$$

$$\sigma_v = 0.6225$$

$$\sigma_B = 200\ \sigma_v$$

$$= 200 \times 0.6225$$

$$= 124.5$$

C.V for Sample B $= \frac{\sigma_B}{\bar{x}_B} \times 100$

$$= \frac{124.5}{1058} \times 100 = 11.77\%$$

Since C,V. for sample B is smaller, sample B lamps are more uniform.

14.3 STATISTICAL ESTIMATION

Estimation is needed in every walk of life. For eg. when you want to cross a road you estimate the speed of the approaching car. Based on this estimate you decide to cross the road by running or walking sales managers make estimates of future sales. Bank Managers make estimate of credit of take. College Hostel Warden might make an estimate of number of students who will have their meal in the hostel mess. Estimates helps in decision making. All statistical inferences are made based on estimation. All those who estimate use samples. Therefore anyone who is trying to draw conclusions something about a population will use information taken from a sample.