Development Research Methods
Dr. Rajshree Bedamatta
Department of Humanities & Social Sciences
Indian Institute of Technology-Guwahati

Lecture - 16
Field Surveys and Inventories

Hello and welcome to week 5 and lesson 3 of the NPTEL MOOC's course on development research methods. In today's lesson, we will study about survey research methods. I have titled it as field surveys or inventories.

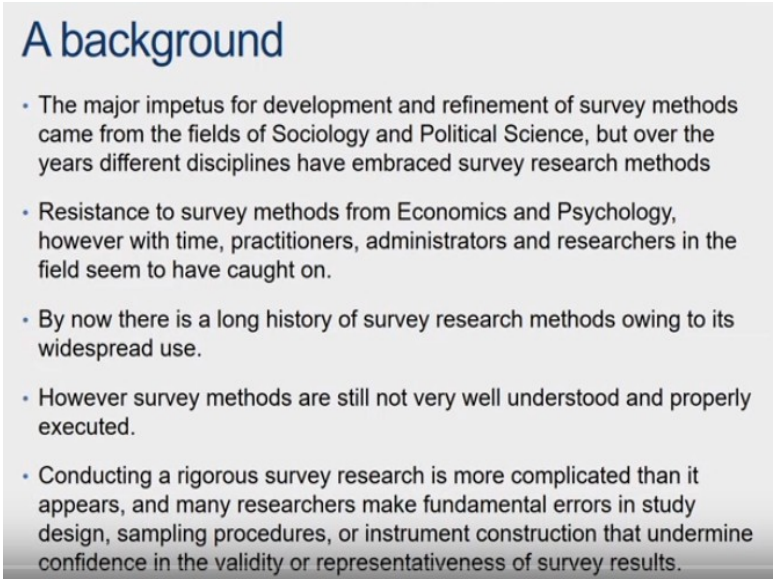
Now surveys are one of the most important methods of inquiry as far as development research is concerned. If we look at the definitive ways in which survey research methods are studied in general in the field of research methods, it is very extensive and it covers a large number of materials. However, for the purpose of today's lesson, particularly suiting to the needs of development researches and development research methods, there are three important things that I would like to cover, which is very frequently used in development research.

(Refer Slide Time: 01:30)

So, what are these three things? First, we will look at what is survey research, what is survey design, what do we mean when we are saying that a survey needs to be designed and what are the few things that needs to be kept in mind; very general points with regard to the design aspect of it. We will also cover about sampling. The

most important methods of sampling that are executed as far as development research methods is concerned. Finally, I will end today's lesson with field surveys and inventories particularly looking at the characteristic features of field mapping and field inventories. Looking at some of the examples and how those examples can be best suited to the needs of the research investigation and development research methods. So, let me begin with what is survey research.

(Refer Slide Time: 02:27)



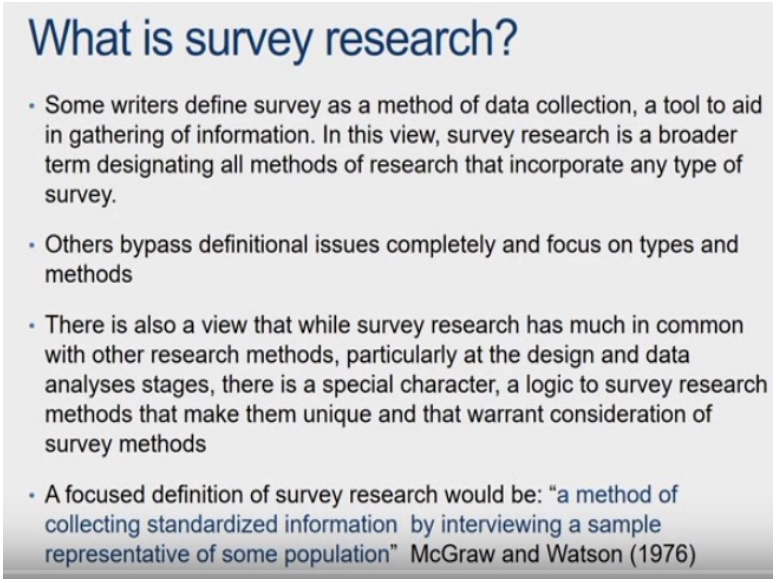
Now if we look at a brief background of which are those social sciences that have extensively used survey research, we would see that the major impetus for development and refinement of survey methods came from the fields of sociology and political science. But over time, various other disciplines within the social sciences have also embraced survey research methods as their own. For example, anthropology and economics are two important social sciences which make extensive use of survey research methods. But if go to the roots of when survey research methods began, we see a lot of its use within the disciplines of sociology and political science.

Actually, resistance to survey methods came from the disciplines of economics and psychology. However, with time practitioners, administrators and researchers in the field seem to have caught on with the idea of survey research. And now, there is a long history of survey research methods owing to its widespread use.

Whenever we talk about development research, particularly with regard to project evaluations or evaluation studies, survey is one of the most important techniques or methods of inquiry that is carried out. However, survey methods are still not very well understood and properly executed, and that is primarily because of not understanding the technicalities with regard to survey design and sampling methods, and that is something that we will touch upon in our lesson today.

Now conducting a rigorous survey research is more complicated than it appears and many researchers actually make very fundamental errors in study design and sampling procedures. And this adversely affects the instrument that is being prepared to be able to carry out the survey, particularly with regard to questionnaire design. Although questionnaire design and the important aspects to keep in mind for questionnaire design is not being taken as a part of today's lesson. However, some of the important points with regard to the designing aspect of it can go on to inform better designing with respect to survey instruments or questionnaires.

(Refer Slide Time: 04:41)



Now what is survey research? Some writers define surveys as a method of data collection or tool to aid in gathering information. And in this view survey research is a broader term designating all methods of research that incorporate any type of survey and with time and with the different methods of surveys being carried out online as well as offline, face to face interviews, different methods of being able to carry out

the survey has also come up like we have telephone interviews, we have web surveys and so on and so forth. Therefore, the nature of survey research has become very varied. However, the very basic considerations with regard to the design and sampling procedures have not changed much.

Now others bypass definitional issues completely and focused on types and methods. And there is also a view that while survey research has much in common with other research methods, particularly at the design and data analysis stages, there is a special character or logic to survey research methods that make them unique and that warrant consideration of survey methods. McGraw and Watson 1976, which is a very acceptable definition of survey research, they define it as 'a method of collecting standardized information by interviewing a sample representative of some population.' And notice the terms here 'interviewing a sample, which is representative of some population'. This is precisely one of the main reasons why survey research methods are carried out in development research. The idea being to support representativeness of the data that we are collecting, representativeness of the population that we are doing our research on.

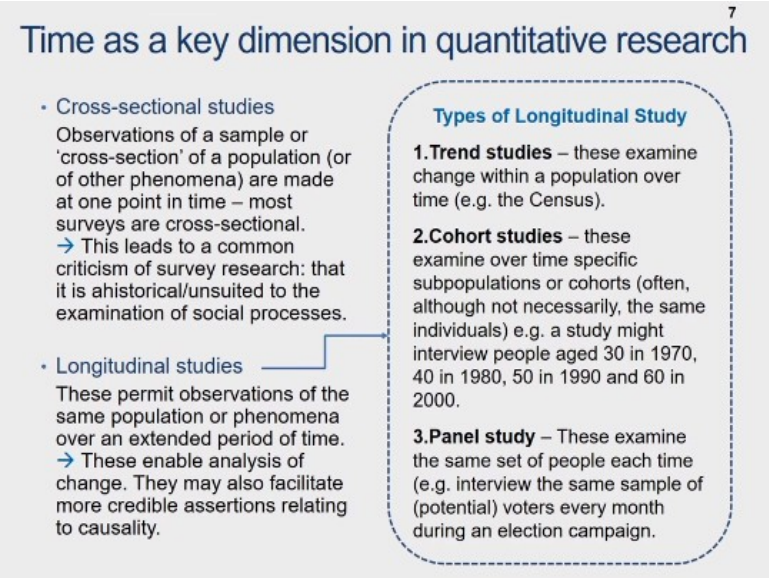
Now given the very basic nature of how survey research design takes place, it will not be wrong to say that survey research or survey techniques follow the overall paradigm of positive methods or experimental methods as we now call them. However, understand that the mixed methods paradigm or the pragmatic paradigm also consider survey research methods is one of the important methods of inquiry.

(Refer Slide Time: 07:05)

A useful way to begin looking at survey research method would be this. First, we need to have a choice of a survey design. Next, we move on to the selection of a representative sample and choosing a sample is of utmost importance because the question here is whether we want the sample to be representative of the population or not. And often when we are carrying out surveys, representativeness is possibly the most important thing that we are trying to focus on. Then of course, interviews, questionnaire design and administration or administering of the questionnaires, how we want to administer them, whether we want to do it physically by caring out face to face interviews, or we want to carry out telephone interviews or online surveys and so on.

Now the major disciplines today that make frequent use of survey research methods are the following; sociology, political science, psychology, economics, social anthropology, education, social work, public health, and medicine. Also, in the fields of evaluation studies falling under each of these themes that I have just cited, survey research methods are used quite frequently.

(Refer Slide Time: 06:16)



Now, time is a very important dimension with regard to quantitative research and that helps us to identify, which are the kinds of method of inquiry that should be carried out and usually survey research methods or surveys are carried out with respect to cross section studies or longitudinal studies. So, we can broadly divide survey research methods' application to studies in the form of cross-sectional studies and longitudinal studies. So cross-sectional studies are basically a point of time study. You are carrying out a study at a point in time, whereas longitudinal studies take place over a period of time.

So, under cross-sectional studies, what we do is observations of a sample or cross section of a population are made at one point in time and most surveys usually use cross-section data. And some of the limitations with respect to survey data making use of cross-section data also comes from the fact that the social processes social phenomenon is not best explained with the help of cross-section data. So, these limitations are pointed out. As a result of which most evaluation studies make use of cross-section data quite often. So, this leads to a common criticism of survey research that it is ahistorical and unsuited to the examination of social processes.

Second is longitudinal studies and these studies permit observations of the same population or phenomena over an extended period of time and therefore, it is said that longitudinal studies enable analysis of change. That they may also facilitate more credible assertions relating to causality. What causes, which are those variables that cause a certain change in the model that we are trying to investigate. So, it is said that,

the survey research methods can bring out best results based upon the longitudinal studies.

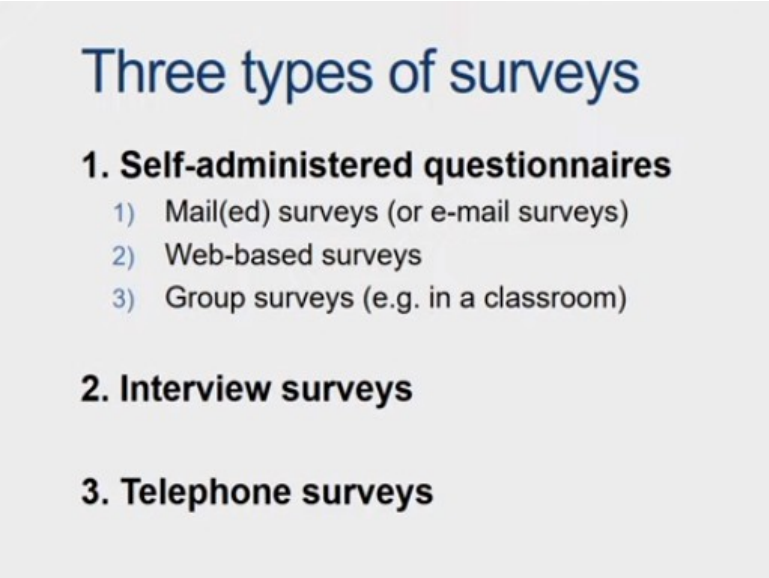
Now there are three very popular longitudinal studies that make use of survey research. One is the trend studies, cohort studies, and panel studies. Trend studies basically examine change within a population over time. For example, we have the census of India surveys carried out decennially or we also have the National Sample Survey Organization surveys which are carried out every 5 years or quinquennially. And similarly, we have various other surveys for example, agricultural census of India or various other labor bureau studies that are carried out, education surveys that are carried out that take place over a period of time. So, in this case the sample that may be studied is not the same over a period of time. However, that gives us a sense of a pattern change that is happening over a period of time.

So, there are trend studies there are also cohort studies. And these cohort studies examine over time specific subpopulations or cohorts. They may not necessarily be the same, the individuals may be different. For example, we may study or interview people aged 30 years in 1970 and then 40 years in 1980, 50 years in 1990, 60 in 2000 and so on and so forth. So, we are studying cohorts at different points in time. And that also gives us some kind of longitudinal data. However, the technique or the method of inquiry that is being carried out is survey research.

A third kind of study with respect to longitudinal research studies are panel studies. These basically examine the same set of people each time. Example, when we are interviewing the same sample of potential voters every month during an election campaign. So, the cohort studies and panel studies differ with respect to the population that we are interviewing. The cohort studies may interview similar groups of population, but the individuals may not be the same. Whereas in panel study, we are investigating the same set of population over a period of time. So, and the method of inquiry is again the survey research methods.

So, these are the few aspects to keep in mind with respect to the time dimension of sampling consequent upon the survey design that we want to carry out.

(Refer Slide Time: 13:00)



Based upon the discussion that I have had so far; we can have three types of surveys. One is self-administered questionnaires, self-administered surveys wherein the interviewer approaches the interviewee personally. There is a physical interview. It can be either physical interview or we are self-administering the questionnaires through a certain medium. It may be email surveys, web-based surveys or group surveys in a classroom.

Now there is a slight bit of a difference made between questionnaires and schedules. Usually schedules are questionnaires made where the interviewer himself or herself physically approaches the interviewees for writing down the details in the schedule. Whereas questionnaires are usually close ended and does not leave much room for interpretation by the interviewees, and it is self-administered, but it can be sent across with instructions to the interviewees to be filled up by the interviewees themselves. And then we can also have telephone service.

(Refer Slide Time: 14:15)

9

Method	Advantages	Disadvantages	Tips to Remember
Self-completion	•Cheap •Cover wide area •Anonymity protected •No Interviewer bias •People can take their time	•Low response rate (and possible bias from this) •Questions need to be simple •No control over interpretation •No control over who fills it in •Slow	•Simplify questions •Include covering letter •Include stamped addressed response envelope •Send a reminder
Telephone survey	•Can do it all from one place •Can clarify answers •People may be relatively happy to talk on the phone •Relatively cheap •Quick	•People may not have home phones/be ex-directory •You may get wrong person or call at wrong time •May be a bias from whose name is listed/who's at home •Easy for people to break off •No context to interview	•Because you rely totally on verbal communication – questions must be short and words easy to pronounce •Minimize number of response categories (so people can remember them)
Face-to-face interview	•High response rate •High control of the interview situation •Ability to clarify responses	•Slow •Expensive •Interviewer presence may influence way questions are answered •If there is more than one interviewer, they may have different effects	•Interviewer should be non-threatening •Interviewer can clarify questions, but should be wary of elaborations that affect the content •Ask questions in a clear, standardized way •If the list of possible responses is long, show them to the respondent for them to read while the question is read out

So, let us look at some of the advantages and disadvantages with regard to these different kinds of surveys. Self-completion, telephone survey, and face-to-face interview. Now the advantages of self-administered surveys are one is that it is cheap. The interviewer does not need to be present. We can cover wide areas. The anonymity of the interviewee is protected. And there are no interviewer bias and people can take their own time for completing the survey questionnaires.

The disadvantages are that because the interviewer is not present for a face-to-face interview, and the questionnaires are being sent across either through emails or through group-based surveys, the response rate may be very low. We will come to how response rate is calculated in some time. And possibly there may also be a bias with respect to the responses that we are receiving from the interviewees. The disadvantages, questions need to be very simple because there is no scope for the interviewer to simplify the questions for the interviewees when they are filling up the questionnaires. There is no control over interpretation and there is no control over who fills it. And of course, with respect to self-administered questionnaires or the right term to use is actually self-completion of the questionnaires. Because questionnaires are usually administered by the investigator or the interviewer, and it is completed by the interviewees. Usually, self-completed questionnaires need a lot of reminders from the investigators to the interviewees to be sent back to them.

So, some of the tips that one should remember when one is focusing on self-completed questionnaires is that the questions must be very simple or it must be

simplified and not leave much room for interpretation. There must be a cover letter explaining the intent and motivation of carrying out the survey and possibly also giving instructions with regard to the surveys that we are conducting. Include stamped addressed response envelope and also keep sending reminders.

Telephone survey's advantages are that we can do it all from one place. We can clarify answers because we are speaking to the interviewer ourselves. People may be relatively happy to talk on the phone because it also enables some kind of anonymity on the part of the interviewee. It is also relatively cheap and quick. And the disadvantages are that people may not have phones. You may get a wrong person or call at a wrong time. There may be a bias from whose name is listed or who is at home, and it is easy for people to break off and there is no context to interview.

So, what are the tips to remember with regard to telephone surveys? You rely totally on verbal communication. Therefore, questions must be short and words easy to pronounce, and to minimize number of response category- so people can remember them because you are asking them. Usually telephone surveys options could be either in the form of a yes or a no. But when we have more options or more categories to choose from, it becomes difficult for people to remember.

Face-to-face interviews are the methods of inquiry that are most often used in development research methods, because when we are moving out to the field to carry out the kind of research questions that I have discussed over the period of last few weeks, these are the questions that are best investigated when the investigator is present on the field and therefore face-to-face interviews is something which is very extensively used in development research methods. And there are some very robust advantages with regard to these interviews. First is that you have a very high response rate and you are in control of the interview situation and you are also able to clarify responses. So, the respondent bias is reduced as far as face-to-face interviews are concerned.

Disadvantage is it may be very expensive. It may be very time taking. It is a resource intensive exercise. There is a lot of time and money spent on being able to carry out face-to-face interviews. The interviewer's presence may also influence a lot of questions that are to be answered, which is also called the investigators' bias. So, both

the respondent bias and investigator bias are frequent. Or the incidence of this is very high when we have face-to-face interviews. And if there is more than one interviewer, they may have different effects. Suppose usually when we carry out surveys in rural areas or urban areas, we carry out household surveys, door to door surveys and we have different teams of investigator approaching these households with the same survey instrument, the same questionnaire. The responses from the interviewers may be varied based upon the kind of questions that are being put forth by the investigator. So that also affects the quality of the responses that is been received.

Some of the tips to remember with regard to face to face interviews are that the interviewer should be non-threatening and therefore, usually when we go for development surveys in different locations, we always give an introduction of ourselves pointing out to the intent of the research, where it is being used, whether we are trying to publish it, who are we collaborating with, with respect to the dissemination results and so on. So, all of this information is made very clear to the interviewee to begin with. Therefore, the interviewer should be non-threatening.

The interviewer can clarify questions, but should be wary of elaborations that affect the content. Usually when we explain more than required to the interviewees the kind of answers responses that we are getting also gets affected. So, we have to apply some amount of judgment with regard to how much should be discussed with the interviewees. We need to ask questions in a very clear and standardized way and if the list of possible responses is long, we need to show them to the respondent for them to read while the question is being read out.

(Refer Slide Time: 20:45)

So, these are some 10 steps to a good survey design. First, we need to identify what is it that we want to cover in a survey. Suppose we are carrying out a study on complimentary feeding practices. This is the example which I have been taking in the last few weeks. And we want to investigate the particular questions with regard to food diversity, food composition, the frequency in which food is being fed, and the various socioeconomic correlates that go on to affect these outcomes, complimentary feeding outcomes. However, along with asking questions and complimentary feeding, if we go on to ask remotely related questions, let us say on different other kinds of nutrition outcomes with regard to the households, then it may go on to jeopardize the research investigation that we are carrying out in the first place. So therefore, we need to make it very clear with regard to what is it that we want to cover in our survey that we have chosen to do.

The next step is to put questions that are necessary. The first is this identify. Second is we need to put questions that are necessary. Unnecessary questions should be removed because that affects the length of the questionnaire. And when the length of the questionnaire gets affected, it is very difficult to make the respondent sit down for a very long time and answer the question. So, the respondents may clearly be very bored with the questionnaire the length of the questionnaire that we have and we may not get the suitable answers. So therefore, there is a need to keep it short and simple.

We ask one question at a time, avoid using jargons. Often when there are technical terms that we put in the questionnaires it also needs to be seen how best to deconstruct

those technical jargons and ask the respondents about them. And these issues can be dealt with only when we have face-to-face interviews. And this is also one of the advantages of face-to-face interviews. Suppose we are carrying out a study on nutrition, a nutrition survey is being carried out and we do know that there are certain technical terms that are used with regard to anthropometric measurements let us say for example. We are taking estimates on mid upper arm circumference or we are taking measurements on the body mass index, taking measurements that can inform us about stunting and wasting and underweight. Now if we start using these terminologies on the field, it is but obvious that we may not get the required answers because the respondents may not be aware of these technical terms and therefore, it is important for the investigators to simplify these terms or not use technical jargons at all.

Whether to use open ended questions or closed ended question is something which needs to be deliberated upon by the investigators. Because face-to-face interviewing gives us a lot of scope to carry out investigation with the help of both open ended and close ended questions, and well-designed surveys usually contain both of them, close ended questions as well as few open-ended questions. We need to spend time to design our survey, analyze the responses after collecting, put a summary report together and conclude our plan of action. These are the 10 steps to a good survey design that needs to be kept in mind.

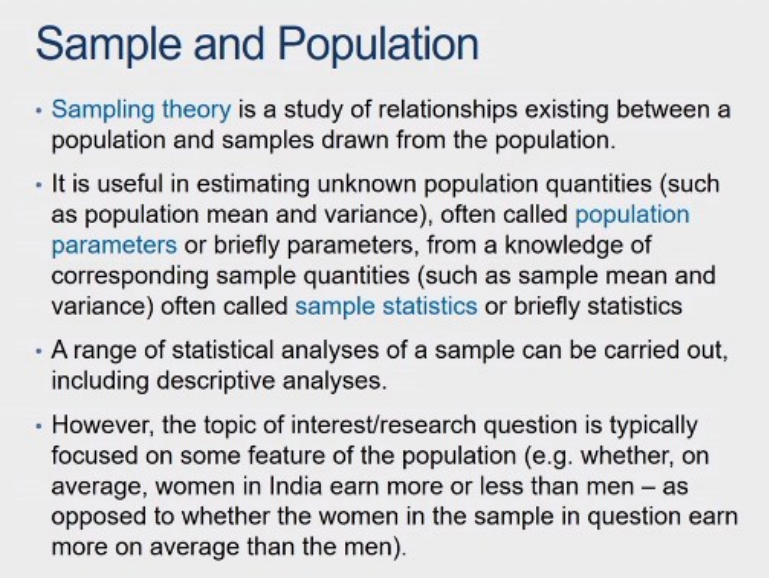
Those who are interested in looking at more aspects of a good survey design, I would urge you to look up this website, questionpro.com from where we have retrieved this figure on 10 steps to a good survey design.

(Refer Slide Time: 24:45)

Now having discussed some of the very basic issues with regard to survey design or what is survey research and different kinds of surveys that are carried out. This is a very elementary level of looking at the sampling procedures. However, this has been kept at an elementary level so that development researchers who want to make use of quantitative data can begin to think about why and where which procedure of sampling is to be used.

So, before we introduce some of the very common forms of sampling that are carried out in development research, let us also get clarified certain terminologies with regard to sampling procedures.

(Refer Slide Time: 25:35)



First is the difference between a sample and a population. So usually when we say that we are choosing a sample that is representative of a population, what is it that we mean? Because the concept sample and the population are used in a certain manner, when we are looking at sampling techniques. So, before that what is sampling theory? Sampling theory is basically a study of relationships between a population and samples drawn from a population. So, for example, you have a class of hundred students, and you want to look at the weight and height of children in a class. And if you are looking at hundred students in a class and the total number of students in the class is hundred, then the 100 N is equal to 100 is your population. Whereas, if you want to draw sample from the class, let us say, a 30% sample from a class of hundred students to look at the various other socioeconomic characteristics of those children then what you are essentially doing is you are drawing a sample from the population.

So and why do we draw a sample from this population, because if we do not have the time and resources to study all the students in a population, and we want to carry out some procedure of sampling to be able to draw a sample so that, that sample can tell us a lot about the larger population from which the sample is being drawn. So, in that sense a sampling theory is a study of relationships existing between population and samples drawn from the population.

Now it is useful in estimating unknown population quantities such as population mean and variants often called population parameters. When we are trying to calculate the average weight and height of a class of hundred children, let us say, based upon the population and we are carrying out some averages there, what we are essentially doing is we are estimating population parameters and these are referred to as parameters. So, when we are calculating the average age of children in a class of hundred, so here age becomes a parameter. If we are calculating height, height becomes a parameter and so on from a knowledge of corresponding sample quantity such as sample mean and variance. And this is also referred to as sample statistics. When we are carrying out these estimations based upon the sample and not the entire population then essentially what we are expressing is sample statistics. So, population parameters, we want to estimate a population parameter let us say age or height based upon the sample statistics from the sample that we have drawn from the population of hundred children.

Imagine increasing this number of 100 to 1000 or 10,000 and then drawing a sample from a population of 1000 would then start making sense. For example, based upon the census data, when we calculate the average literacy rate of the country or we calculate the average life expectancy of adults in a country, so essentially what we are doing is we are giving you population parameters here. In the case of National Samples Survey Organization when we are carrying out employment statistics, we are saying what is the employment rate in the country, what is the usual status unemployment rate, what is the daily status unemployment rate. What we are essentially doing is we are carrying out samples' statistics. We have drawn a sample based upon a sampling procedure and then we are carrying out, we are estimating sample statistics based upon the sample that we have drawn. So that is about population parameters and sample statistics.

Now a range of statistical analysis of a sample can be carried out including descriptive statistics and that comes under the larger domain of development statistics. For those who are interested can look up various books related to development statistics. One of them being Chandan Mukherjee's Econometrics in Underdeveloped Countries. This is a highly recommended book for those interested in studying development statistics that are applicable to development research methods.

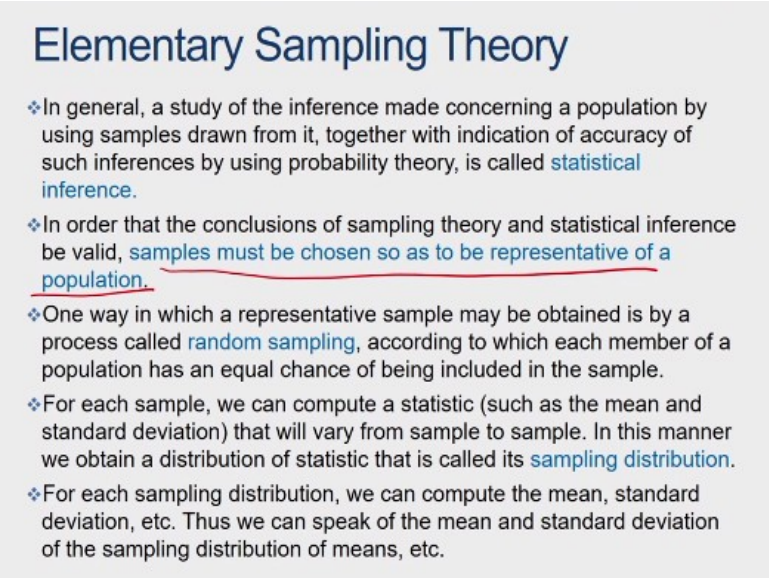
Now the thing to keep in mind is the topic of interest or research question is typically focused on some feature of the population. Example whether on average women in India earn more or less than men as opposed to whether the women in the sample in question earn more on average than the men. So, this is a point that needs to be remembered, what is the motivation for carrying out a survey? What is the motivation for drawing a sample from the population that we want to study? So, when we are posing a question with regard to let us say complimentary feeding and we are saying that, what are the complimentary feeding practices that are followed in a certain community, in a certain location, in a certain state in India let us say. So, if we want to study the different kinds of feeding practices in a certain community and the community has large numbers of people, then obviously we will be drawing the sample from them, from that population to check if we want to adopt the method of survey to do an inquiry with regard to their complimentary feeding practices.

So usually when we are posing the research question for which survey inquiry is to be carried out, research questions are never posed for the sample that we are studying, the research questions are posed for the population that we are studying and to be able to answer those questions we draw a sample from the population and then say that okay, since we have followed the rules of sampling for drawing a sample, so the estimates that are coming from that sample are representative of the entire population that we are studying.

So just to summarize, this discussion that we have had, I will just read out the whole slide for you. Sampling theory is a study of relationships existing between a population and samples drawn from the population, number one. Number two, what are population parameters? Population parameters, it is useful in estimating unknown population quantity such as population mean and variances. These are often called population parameters or briefly the parameters from a knowledge of corresponding sample quantities. And when we are estimating sample mean and sample variance it is referred to a sample statistics or briefly statistics. And there is a range of statistical analysis of a sample that can be carried out including descriptive statistics. Descriptive statistics is usually the very basic estimations that we carry out with regard to the measures of central tendency let us say carrying out the arithmetic average- the arithmetic mean, the media, and the mode and the skewness in the data, the variance in the data and so on and so forth, the standard deviation in the distribution that we are studying and so on.

Lastly the topic of interest or research question is typically focused on some feature of the population example we can ask the question whether women in India on an average earning more or less than the men and we hardly ever ask the question whether women in the sample in question are earning more or less on an average than men.

(Refer Slide Time: 32:50)



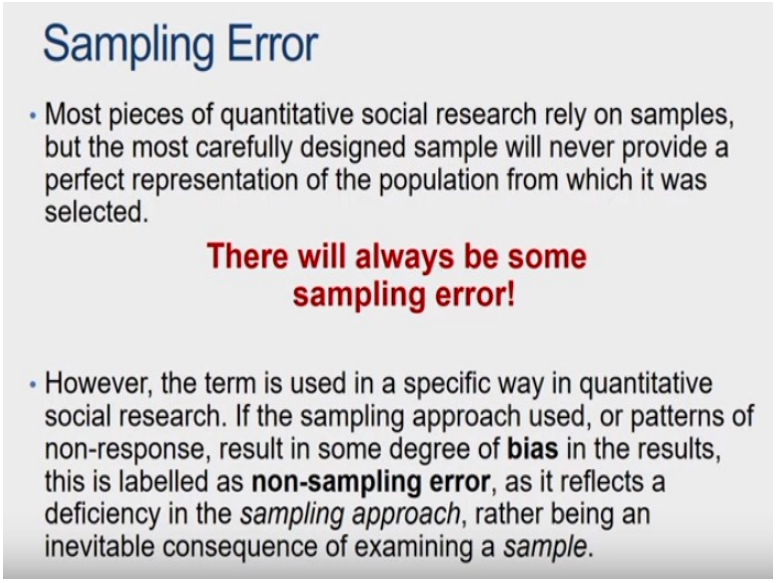
Now in general, study of the inference is made concerning a population by using samples drawn from it and together with indication of accuracy of such inferences we are using probability theory and this is called statistical inference. When we are trying to use the methods of probability sampling, we are following the rules of probability to be able to make certain inferences from the sample for the population.

We say that we are making use of statistical inference and in order that the conclusions of sampling theory and statistical inference be valid, samples must be chosen so as to be representative of a population. And this is an issue that are generally faced in different kinds of methods of inquiry by development researchers. Often in this field we make a distinction between probability sampling and non-probability sampling. And it is only probability sampling that makes use of methods wherein we are drawing a sample and we can very conclusively say that these estimates are representative of the entire population. Non-probability sampling usually is not representative. However, it is carried out and publications also make references to non-probability sampling. It has in a way become practice to make use of non-probability samples, particularly with regard to social studies or development research methods.

Now one way in which a representative sample may be obtained is via process called random sampling. According to which each member of a population has an equal chance of being included in the sample. I will elaborate on random sample in some time. So, we will keep it to later. For each sample, we can compute a statistic such as

the mean and standard deviation that will vary from sample to sample. And in this manner, we obtain a distribution of statistics that is called its sampling distribution. So, when we have a certain distribution and when we are computing certain statistics based upon the distribution that we have, and suppose we have different samples and we are computing these statistics for these different samples. And in this manner when we obtain a distribution of statistic, we call it sampling distribution. And for each sampling distribution we can compute mean standard deviation in all of these. So, we can speak of mean and standard deviation of the sampling distribution of means and so on.

(Refer Slide Time: 35:09)

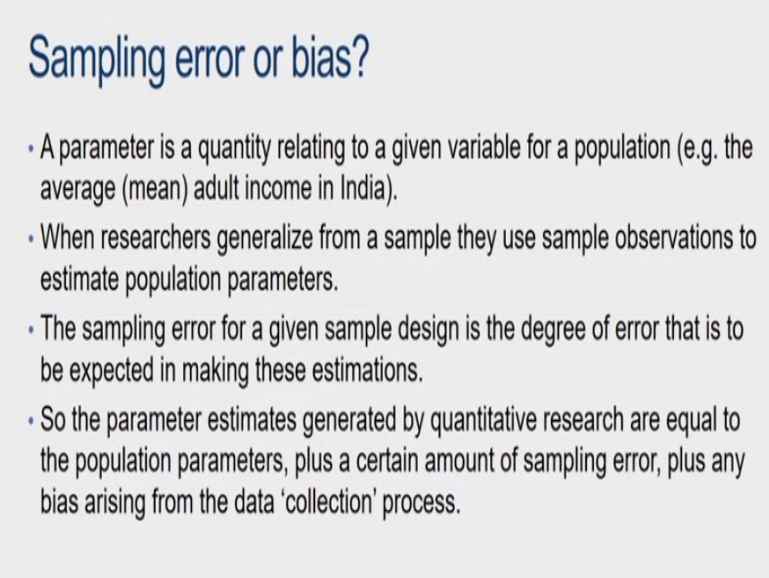


Before moving on to random samples, let us also spend a little time on sampling error. Often, when we are presenting our results from the sample based upon some sampling that we have carried out and we are discussing our sample statistics there is also mention made of sampling error. So, what is sampling error and why is reporting sampling error also important.

So, most pieces of quantitative social research, as we now know rely on samples, but the most carefully designed sample will never provide a perfect representation of the population from which it is selected, there will be some amount of error. So, the term is used in a specific way in quantitative social research. If the sampling approach used, or patterns of non-response result in some degree of bias, remember we spoke

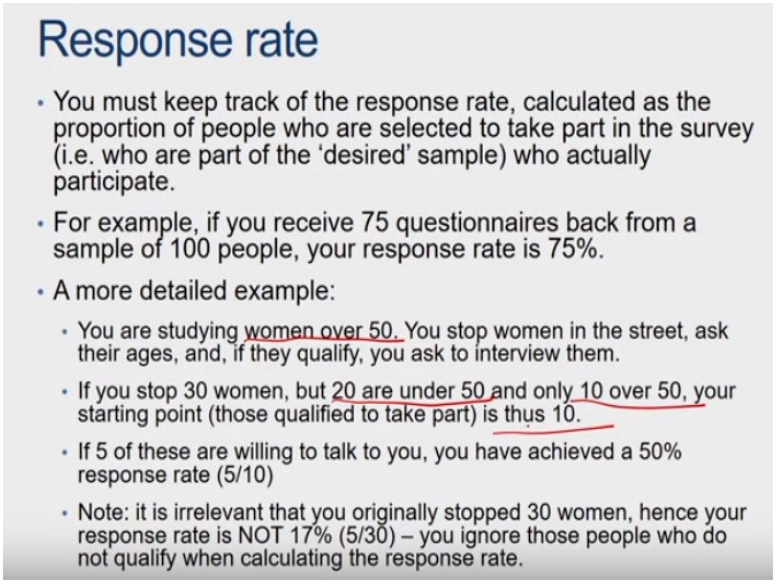
about investigator bias, respondent bias when we were talking about survey design and therefore when we are designing the questionnaires we have to keep in mind regarding these biases. It is these biases that enter into error. And these are referred to as non-sampling error as it reflects the deficiency in the sampling approach, the way we have carried out the sampling rather being an inevitable consequence of examining a sample. So, should we call it a sampling error or bias? Now, a parameter is a quantity relating to a given variable for a population. Example, the average adult income in India.

(Refer Slide Time: 36:46)



Now when researchers generalize from a sample, they use sample observations to estimate population parameters and sampling error for a given sample design is the degree of error that is to be expected in making these estimations. So, you have non-sampling error and you have a sampling error. Are all those biases enter into non-sampling error but sampling error is a degree of error that is expected in when we are making these estimations from sample to population parameters. So, the parameter estimates generated by quantitative research are equal to the population parameters plus a certain amount of sampling error plus any bias arising from the data collection process. Because of investigator and respondent bias and the methods of sampling that we are employing to be able to draw a sample from the population, all of these taken together can result in some amount of errors and that is what is referred to as sample error.

(Refer Slide Time: 37:44)



Let us spend a little time on what is response rate. Usually when we carry out survey research, one of the first things that we report based upon our survey data is what is the response rate of the sample that we have studied. We will discuss it with the help of an example. It is important to keep track of response rate calculated as the proportion of people who are selected to take part in the survey. That is who are part of the desired sample and who actually participate. For example, if you receive 75 questionnaires back from a sample of 100 people your response rate is 75%.

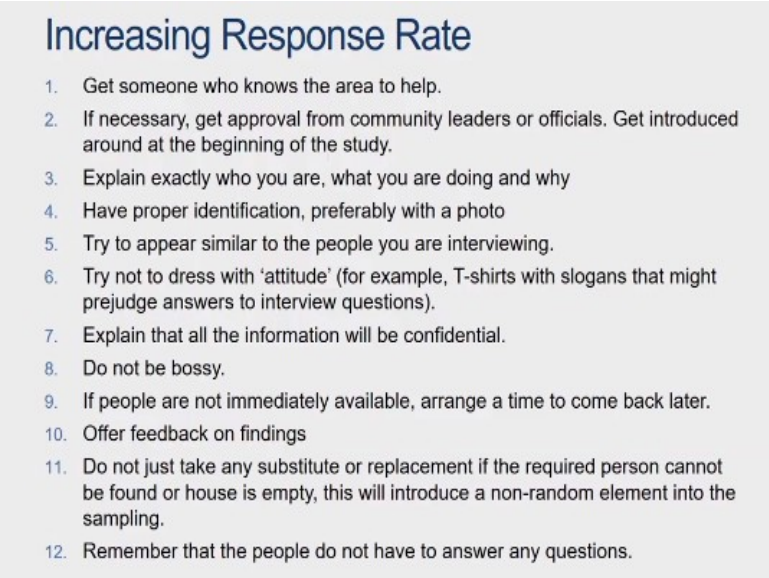
Suppose, you are studying for your research, you are studying women over 50 and you stop women in the street you ask their ages and if they qualify, you ask to interview them. Suppose you have stopped 30 women but 20 are remember that you are studying women over 50, you stop these 30 women and of them 20 or under 50 and only 10 are over 50. So, your starting point that is those qualified to take part is thus 10. Now of these 20 who are under 50 but who you stopped to ask questions and they may have answered you do not include them as a part of the responses that you have received because they do not qualify to be answering your questions in the first place. So, your starting point is 10. And if of these 10, five are willing to talk to you, then you have achieved a 50% response rate, which is five 5 on 10. Because 10 are the people who are the women who have qualified to participate in your survey in the first place. So, it is important to know that it is irrelevant that you originally stopped

30 women, hence your response rate is not 17%, which is 5 on 30. But you ignore those people who do not qualify when calculating the response rate.

However, there is a word of caution. It needs to be borne in mind regarding this number 10. Now what you are essentially doing here is you are sampling a number of women for your study. Whether you want to stick to only the number 10 or you need to increase the size of your sample is something which depends upon the study that you are carrying out. And there are usually certain benchmarks with regard to what is the sample that should be drawn from a population. When we are carrying out survey research methods, when we are carrying out survey research for development studies, let us say we are doing a village survey, where we want to study report findings from households, let us say agriculture related surveys and so on and so forth benchmarks are usually in the range of 15% to 30%. So, we usually draw a sample of 15% from the population such that, that can be representative of the entire population.

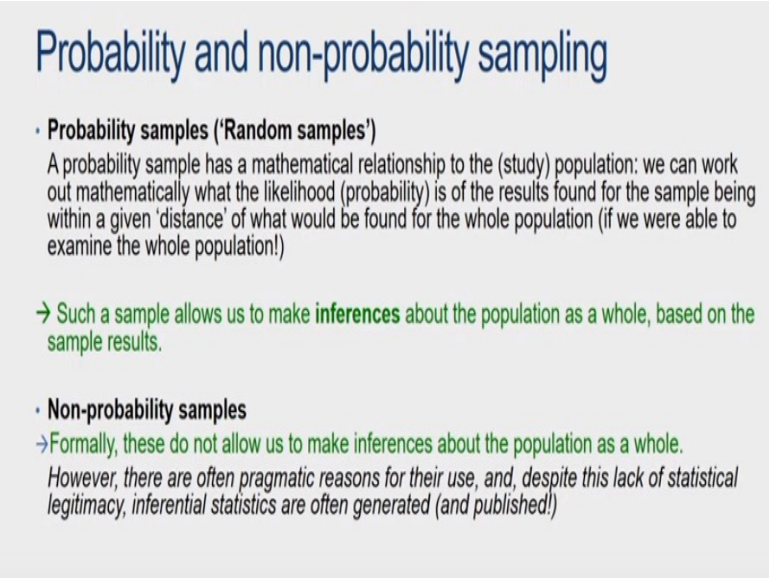
However, when we carry out large samples, large sample surveys when we are doing project evaluations and so on the benchmarks can also come down to 5%. Suppose we are covering 100 districts or we are covering let us say 20 states and so on and we are carrying out let us say 500 blocks in a country for reporting our results with regard to farmers incomes, let us say, then the benchmark can come down to 5%. We can choose 5% of farmers from each block of the country and that adds up to a certain number. So, these benchmarks have to be kept in mind, borne in mind with respect to the context of the research question that we have chosen.

(Refer Slide Time: 41:29)



So these are some of the tips to increase response rate. You get someone who knows the area to help you. If necessary, get approval from community leaders or officials. Get introduced around at the beginning of the study. You explain exactly who you are, what you are doing and why. Have proper identification, preferably with a photograph. You try to appear similar to the people you are interviewing. Your attitude and dressing sense also add on a lot to what you were trying to bring to the survey research. You explain that all the information will be confidential. You do not try to be very bossy or arrogant about the research that you are carrying out. If people are not immediately available you arrange a time to come back later. So therefore, patience is to be exercised when we are carrying out survey research. And offer feedbacks on findings and we do not just take any substitute or replacement if the required person cannot be found or house is empty. This will introduce a non random element into the sampling. We must remember that the people do not have to answer any questions and therefore a lot of judgment and patience needs to come in with regard to survey research.

(Refer Slide Time: 42:36)

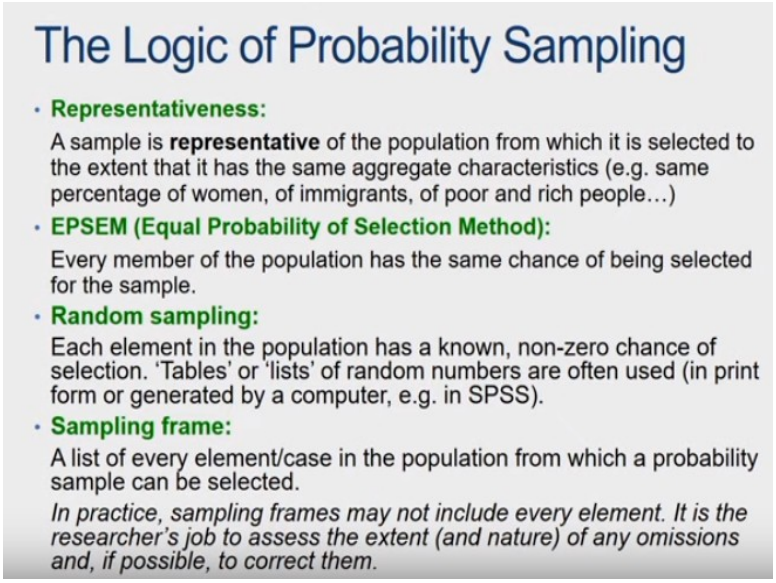


Now let us come to some of the very important probability and non-probability sampling methods that we carry out. Usually, we make a distinction between probability samples and non-probability samples. And when we say random samples, we are actually talking about probability samples. There is a lot of misinformation with regard to random samples and how random sampling takes place. Random sampling is a purely mathematical method of drawing a sample from the population and also one of the robust methods of being able to conclude about the population that you are studying. So, it is one of the purest forms of probability samples where it shows a mathematical relationship to the study population. And we can work out mathematically what is the probability of the results found for the sample being within a given distance of what would be found for the whole population. So, such a sample allows us to make inferences about the population as a whole based on the sample results. So, if we want representativeness to be the hallmark of our study, it is usually advisable to draw a random sample based upon the techniques of sampling.

Non-probability samples as I have briefly pointed out earlier, these do not allow us to make inferences about the population as a whole and social science research since we are investigating human beings, we are investigating the context within which human beings survive often make use of non-probability samples. And it is important to bear in mind that sampling methods which does not make use of probability sample does not qualify to represent the entire population. So, we should exercise caution when we are coming up with our conclusions with regard to the population when we are following non-probability sample.

So, there are often pragmatic reasons for using non-probability sample, but it lacks statistical legitimacy in the sense that we cannot represent much about the population. Whether you are carrying out a one point study, and if you are carrying out a one point study, doing survey based upon the sample that you have chosen, at best we can only describe the sample that we are studying without being able to draw any inferences out of them for the entire population. And as I just mentioned, inferential statistics are often generated and published based upon non-probability sampling. However, statisticians will not be very happy to look at non-probability sampling coming up with conclusions that claim to represent the entire population.

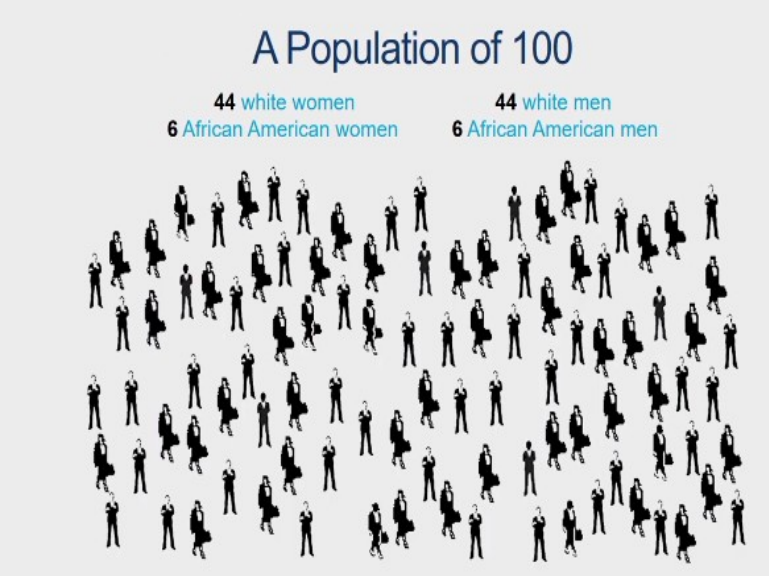
(Refer Slide Time: 45:25)



So, what is the logic of probability sampling? We can group them as following. First is representativeness. Sample is representative of the population from which it is selected to the extent that it has the same aggregate characteristics. Example, same percentage of women, immigrants, poor and rich people. Equal probability of selection. So, when we are drawing a random sample, every member of the population has the same chance of being included in the sample, so random sample. Each element in the population has a known or non-zero chance of selection. So, tables or lists of random numbers are often used either in print form or generated from the computer, from the excel sheets and so on.

Fourth thing to remember is sampling frame. The sampling frame is very important. What is the sample frame? From where we are drawing the sample? So, let us say we are studying households in a certain location. Now how are we going to sample the households in that location? Whether we are doing a district level study, we are doing a block level study, we are doing a village level study, how are we to draw the sample? What is the sample frame? What is the list of households from which we are to draw the sample and that is of paramount importance with regard to sampling frame. So, it is a list of every element or case in the population from which a probability sample can be selected. And in practice sampling frames may not include every element. It is the researcher's job to assess the extent and nature of any omissions and if possible, to correct them.

(Refer Slide Time: 46:58)

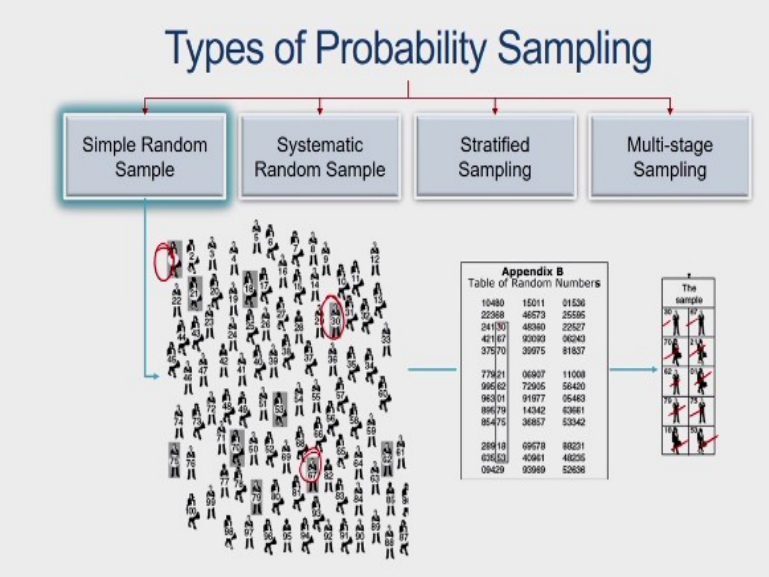


Let us look at this population. Here we have a population of 100 constituting of 44 white women and white men respectively and 6 African American women and men respectively. Now this is a population of 100 people and we want to draw a sample from this population. Now if our research question has nothing to do with the ethnicity of the people that we are studying, then we might consider our sample frame to be constituting of 100 persons here. And suppose we are trying to estimate average incomes by drawing a sample from these 100 peesons that is just showing on your slide here then essentially what we are doing is we are not laying much emphasis with regard to the ethnicity of these persons that are being shown in this population. We

are only concerned about drawing let us say 15% sample which will go on to give us some sample estimates, so that we can draw some inferences about the population.

But the situation changes if we are concerned about the ethnicity backgrounds of the population also. Let us say we are looking only at the average incomes of white men and women. And our sample frame then becomes only white men and women. That is 88 persons in the total population and we draw a sample from the 88. Now further, if we are looking only at the average incomes of white women vis-a-vis all the others then our sample frame is 44 and not 100 as was earlier pointed out. So, it is very important, depending upon the researcher's questions, it is very important to come up with a conclusion with regard to what is a sample frame that we are finalizing and from that sample frame to draw a sample.

(Refer Slide Time: 49:04)

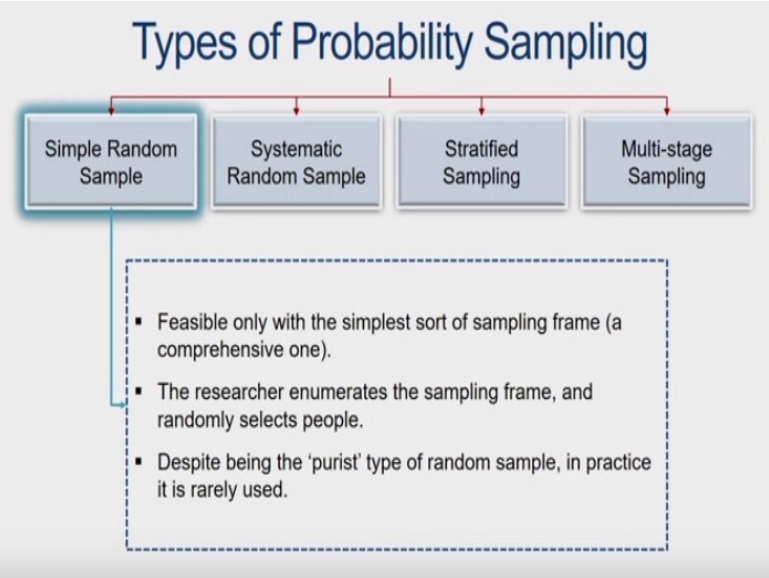


So, these are the different types of probability sampling that is usually used. First is a simple random sample, systematic random sample, stratified sampling, and multistage sampling. As I pointed out, simple random sample is one of the most widely used. So, what we essentially do, suppose this is a sample frame and we have numbers of people from 1 to 100 in this. So, we have a population of 1 to 100. And let us say we want to choose 10 people; we want to draw a sample of 10 people then how do we draw a sample of 10 people. If you look at the back of the statistics textbooks, you will see random number tables. So, you pick up the random number tables and point

out one random number and you start from there, and you generate 10 random numbers from the random number table. And depending upon the number that has come up on your table, you choose only those numbers from the population here. So, let us say you have 100 numbers here and the random numbers that you have chosen is 30, 67, 70, 21, 62, 1, 79, 75, 18, 53. So there are 10 numbers here. 30, so your first sample is 30. 67, so you choose 67 from here as one of your sample. You have a sample 1, so 1. So when we are choosing 10 persons from the sample of 100 population, given the random numbers that are showing up on your computer or that you have chosen from your random number table here, you look up these numbers and you select these numbers and you make them a part of your sample. So that is how you draw your sample.

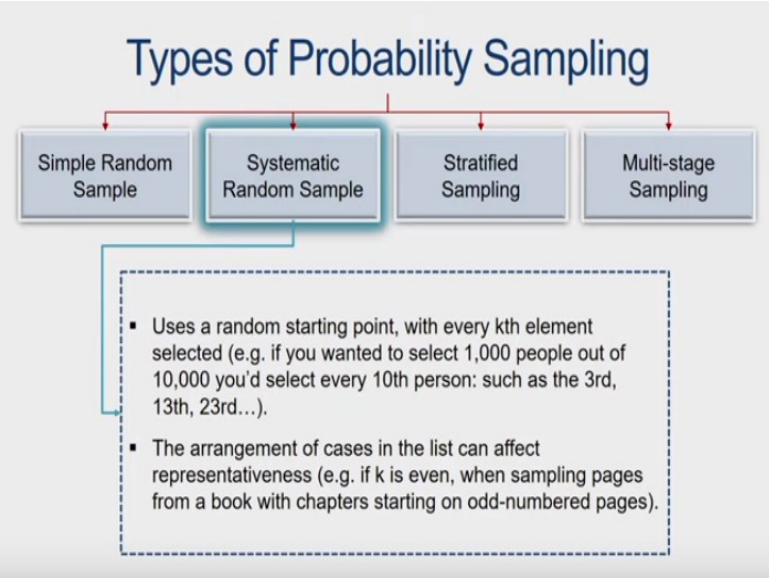
It is very akin to the lottery method. So when you close your eyes and draw a lottery, so suppose you have 100 persons again and you want to choose randomly 10 persons from the 100 which is a sampling frame, you write a chit on which you number them as 1 to 100 and put it in a bowl and close your eyes and draw a lottery of 10 people. And that is a finest example of how random samples are actually chosen. Because there is no bias involved here. Your eyes are closed. There is no judgment with regard to what is it that you are trying to choose here. So that is the best example of a random method. But of course, the lottery method fails when the number increases, when your population is large, when your sampling frame is large you cannot make use of lottery methods. And therefore, you make use of random number tables or you generate random numbers on excel sheets or your spreadsheets and then choose the designated numbers.

(Refer Slide Time: 52:11)



So simple random sample is feasible only with the simplest sort of sampling frame but it is a comprehensive one and the researcher enumerates the sampling frame and randomly selects people. Despite being the purest type of random sample, in practice it is rarely used. However, most village surveys, evaluation studies, which has door to door surveys, we make use of simple random sampling methods.

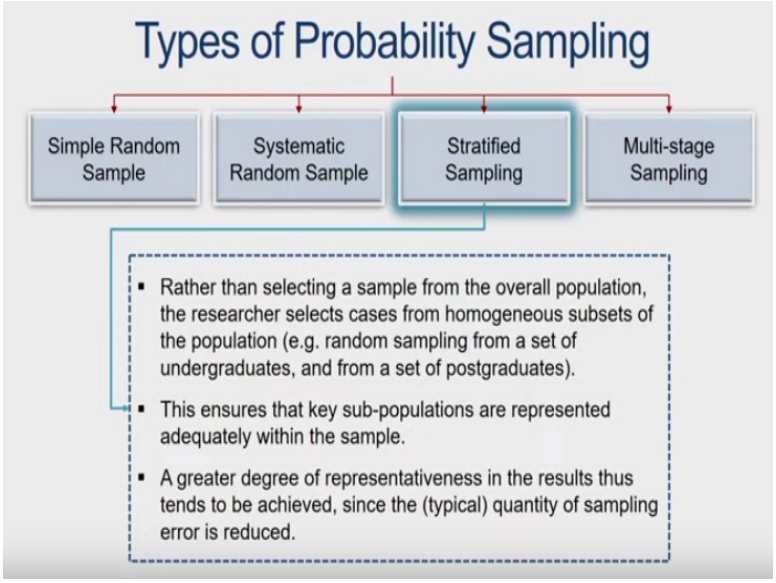
(Refer Slide Time: 52:41)



Now systematic random sample is another form of probability sampling. Here it uses a random starting point, let us say every kth element selected. For example, if you want to select 1000 people out of 10,000 you select every 10th person such as the 3rd, 13th, 23rd, 33rd, 43rd, 53rd and so on get selected. So, this is what is referred to as

systematic random sample. And the arrangement of cases in the list can affect representativeness. It is very important how the cases are arranged in the list. If we follow a certain strategy for arranging people in the list such that we know who is the person that is getting selected in the 13th and the 23rd in the 33rd and so on, then there is a lot of bias. So, these are things to be kept in mind when we are following systematic random sample. However, nevertheless systematic random sample is also a very robust form of probability sampling methods.

(Refer Slide Time: 53:38)



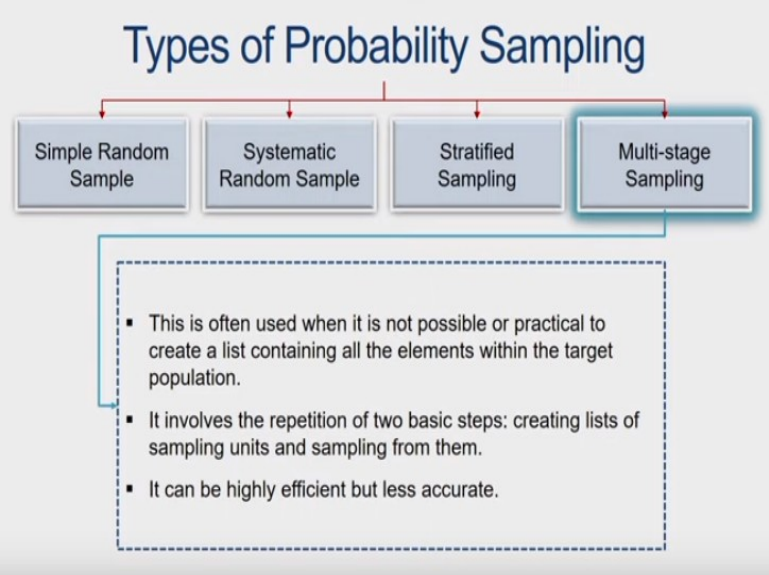
Stratified sampling is another form of probability sampling. Here, rather than selecting a sample from the overall population, the researchers selects cases from homogeneous subsets of population. Example random sampling from a set of undergraduates and from a set of postgraduates. So, you are stratifying the population to begin with and then you are drawing a random sample from each of the strata.

So, let us say you want to carry out a survey of farm incomes and you approach a certain location and you see that there are different size classes of farmers holding different size classes of farms. So instead of straightaway carrying out a random sample, the best thing to do would be to stratify the farmers into different classes. Let us say you have large farmers, medium farmers, small farmers, marginal farmers, and so on. Once that first stage of stratification is complete, you can then draw a random sample from the large farmers, the medium farmers, or small farmers, such that each

of the farm size classes have an equal chance of being represented in the sample. So that is what we mean by stratified sampling.

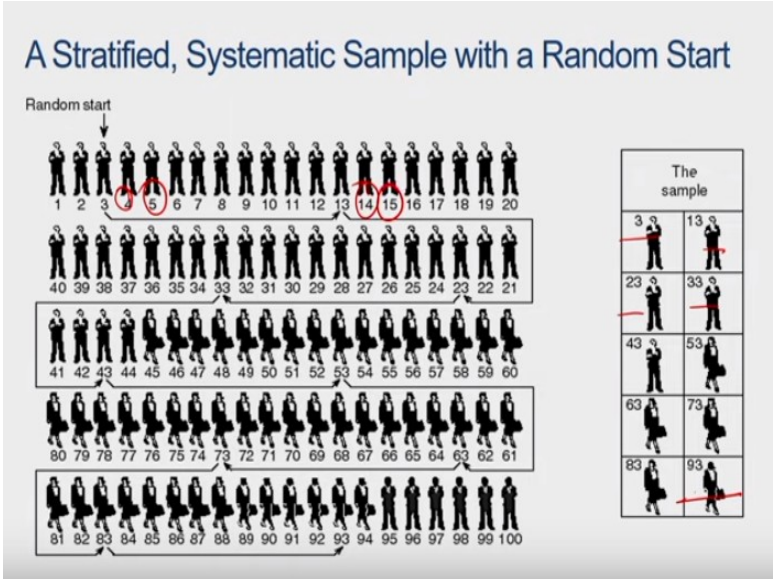
And this ensures that key sub populations are represented adequately within the sample. This is also true of a multi ethnic society, of a multicast society, class society where there are different numbers of ethnic groups of populations and population and we want each of these strata to be represented in the sample in a study that we are conducting and that is when stratified sampling is of paramount importance. A greater degree of representativeness in the results tend to be achieved since the quantity of sampling error is reduced.

(Refer Slide Time: 55:23)



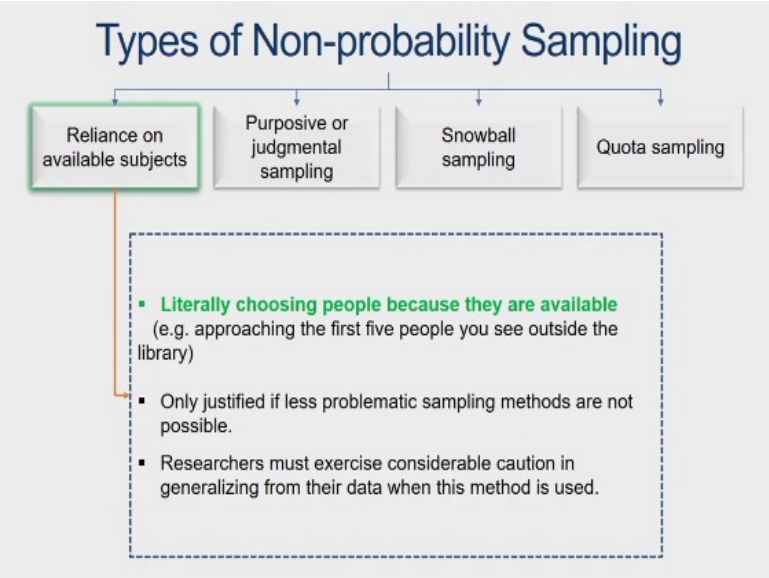
Multi-stage sampling. This is often used when it is not possible or practical to create a list containing all the elements within the target population. And it involves repetition of two basic steps or many steps creating lists of sampling units and sampling from them. It can be highly efficient, but less accurate with regard to the estimates that we are calculating.

(Refer Slide Time: 55:43)



So, this is an example of a stratified systematic sample with a random start. So here, if you look at the pictures here, you would see that the sample have been stratified differently they are different looking persons and individuals in this picture. However, we are carrying out a systematic sample with the random start. Here we start with the third person, move to the 13th, 23, 33 and this way we go on to the 93rd person. So, we have a random start. And one of the ways to do is to just close your eyes and put a finger somewhere. If you are choosing the fourth person, then you can go to the 14th. If you are choosing the 5th person, then you can go to the 15th for the next sample and so on, 15th and then 25th and so on. So, you have a random start. You stratify the population, but then carry out systematic random sampling.

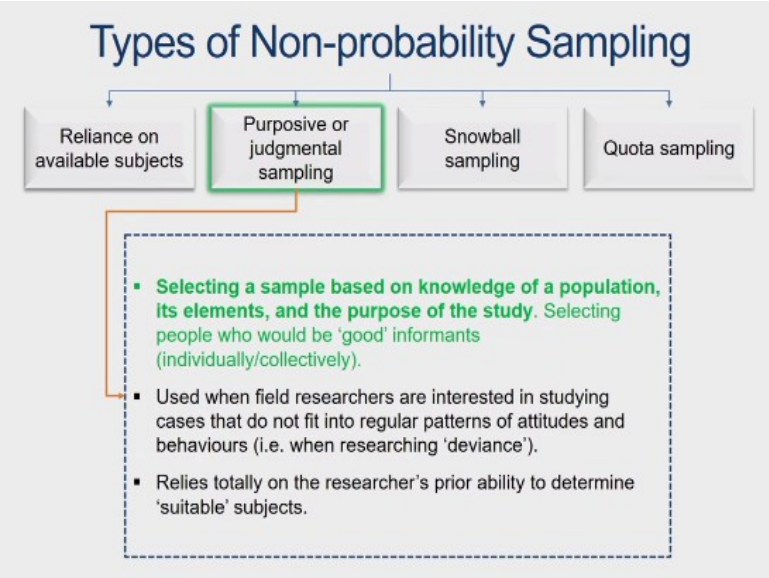
(Refer Slide Time: 56:43)



These are some of the different types of non-probability sampling. One is, there has to be reliance on the available subjects. And usually when we are carrying out mixed methods study, with a very large qualitative component when the focus is on the qualitative component and often there are different characteristic features that needs to be kept in mind for deciding on the sample, then probability samples may be very hard to get. And that is where you will rely on available subjects. You are literally choosing people because they are available. Example, approaching the first five people you see sitting outside the library or in cases of migration let us say when you are carrying out a study on people who are migrating out during a certain season and you have come back after the migration season and it happens that the people have not yet returned, then you may just want to investigate or ask questions to the first five people that you see sitting in the location that you are studying.

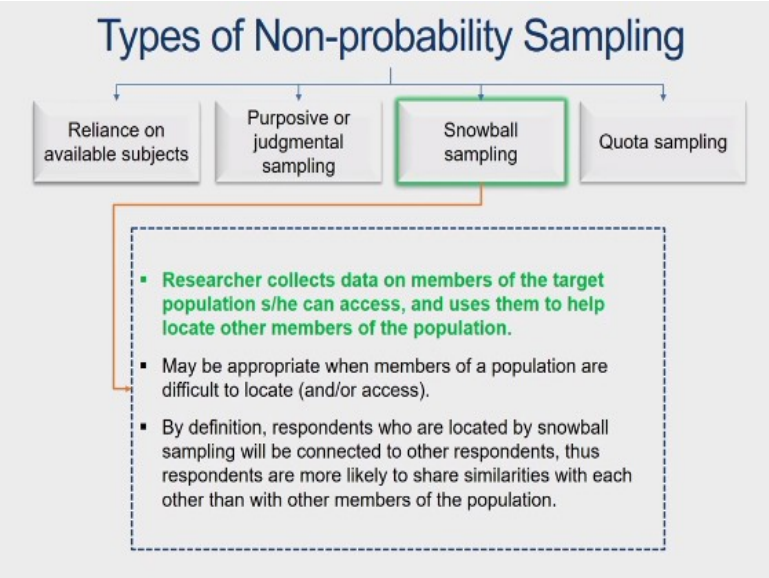
It is only justified if less problematic sampling methods are not possible and researchers must exercise considerable caution in generalizing from their data when this method is used.

(Refer Slide Time: 58:02)



Purposive or judgmental sampling is something which is very frequently used in qualitative studies. And this is basically when you are selecting a sample based on knowledge of a population, its elements and the purposes of the study, you know what is the population like and then you are selecting, you are choosing a sample from them. So, you are selecting people who would be good informants. You know that there are certain sample, certain cases, or certain people who can provide you better information and then you choose those informants individually or collectively. So, it is used when field researchers are interested in studying cases that do not fit into regular patterns of attitudes and behaviors. And relies totally on the researcher's prior ability to determine suitable subjects and equally in the case of purposive sample also we need to exercise caution with regard to drawing inferences for the population as a whole.

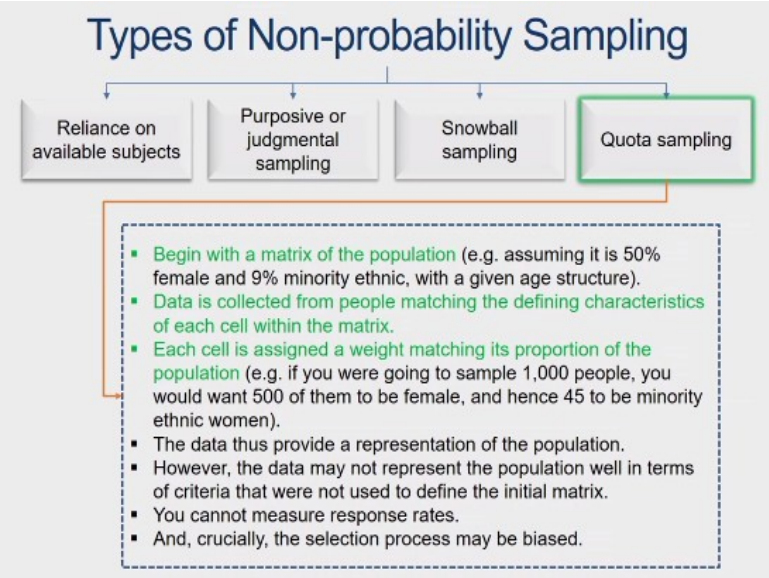
(Refer Slide Time: 58:53)



Snowball sampling is something which is very frequently used again in development research. Here, the researchers collect data on members of the target population that they can access and they use them to help locate other members of the population. Like for example, I was taking the example of migration. You enter a village; you see that there is a certain person who has returned from seasonal migration. And you are discussing with this informant with regard to the living conditions of how it was in the place that they had migrated to. And you get more information from this person regarding someone else who might have also migrated during the season and then based upon the information received from the first person you approach the second person for receiving more information. So that is how it takes a snowballing format.

Now these methods, snowball sampling may be appropriate when members of a population are difficult to locate or access. And by definition, respondents who are located by snowball sampling will be connected to other respondents, they know them. So, respondents are more likely to share similarities with each other and with other members of the population. And therefore, again, from snowball sampling we cannot draw inferences for the entire population.

(Refer Slide Time: 1:00:12)

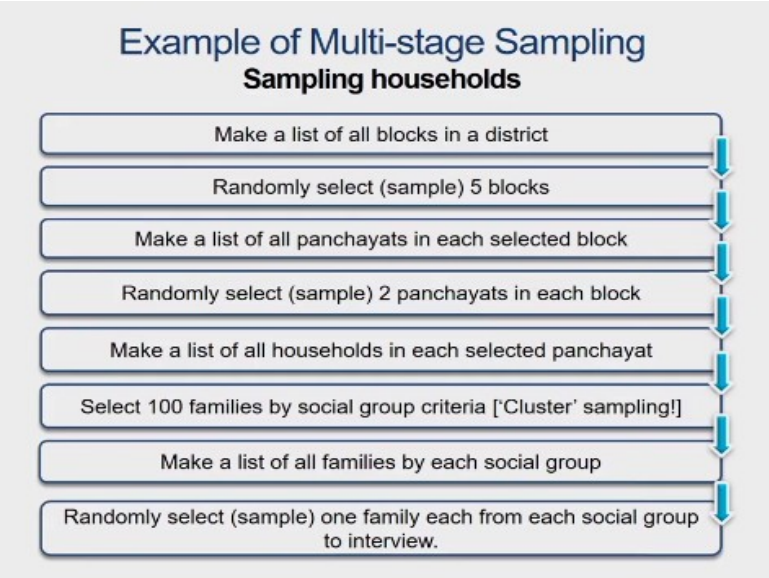


Lastly, quota sampling, this is another form of non-probability sample where here you begin with a matrix of population. Let us say you are assuming it is 50%, female and 9% minority ethnic with a given age structure. So, data is collected from people matching defining characteristics of each cell within a matrix. And each cell is assigned a weight matching its proportion of the population. So, example if you are going to sample 1000 people, you would want 500 to be female, because you are talking about 50% female here, you will have 45 minority ethnic women because you have already identified 9% minority ethnic.

So, what happens here the data does provide a representation of the population because you know that in this population 50% are women, 9% are minority ethnic. So, you are trying to represent the entire population in the sample. However, the data may not represent the population well in terms of criteria that were not used to define the initial matrix. Like for example, there may be various other characteristic features, identity related characteristic features which we might not have considered in the matrix and that might result in some kind of a bias in the study that we are carrying out.

Also, you cannot measure response rates and crucially the selection process may be extremely biased as far as the sample drawing is concerned.

(Refer Slide Time: 1:00:34)



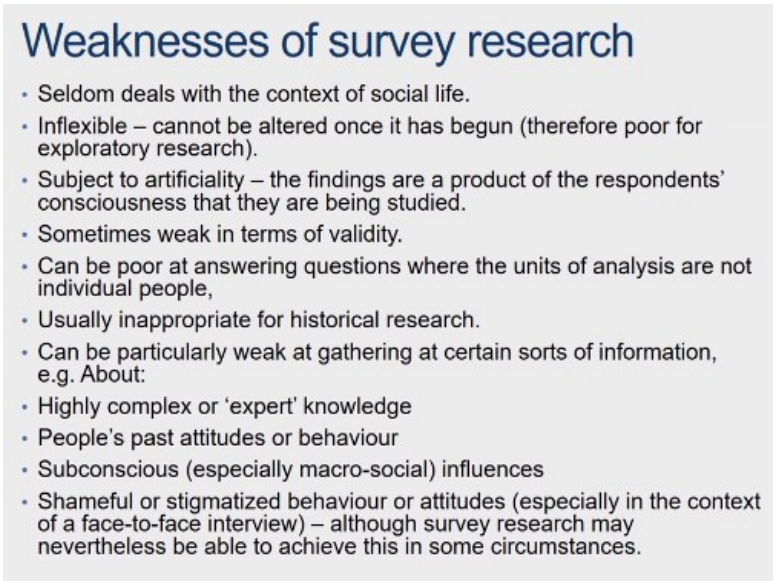
These are some examples of the multi-stage sampling method. So, this is with regard to the multi-stage sampling that I was talking about in probability samples. So, when we are sampling households, this is just an example of how to identify the different stages. So maybe you first make a list of all blocks in a district. You randomly select let us say five blocks. Then you make a list of all panchayats in a selected block. You randomly select let us say two panchayats in each block. Make a list of all households in each selected panchayat and select hundred families by social group criteria and make a list of all families by each social group. And then you randomly select one family each from each social group to interview.

So here essentially in the cases of multi-stage sampling, what you are essentially doing is you are drawing a random sample at each stage. But you are defining each stage purposively as to which stage should be identified for giving you a sample frame.

(Refer Slide Time: 1:02:41)

Now let us conclude this part of the discussion with focus on strengths of survey research and disadvantages. Some of the strengths of survey research are the following. It is useful for describing characteristics of a large population. It makes large samples feasible. It is flexible because you can ask many questions on a given topic. And it has a high degree of reliability and replicability. When we want to replicate our results, we make use of survey research methods and is a relatively transparent process as well.

(Refer Slide Time: 1:03:13)



Weaknesses are enormous. However, given the robustness of replicability and representativeness, survey research is very frequently used. Some of the important

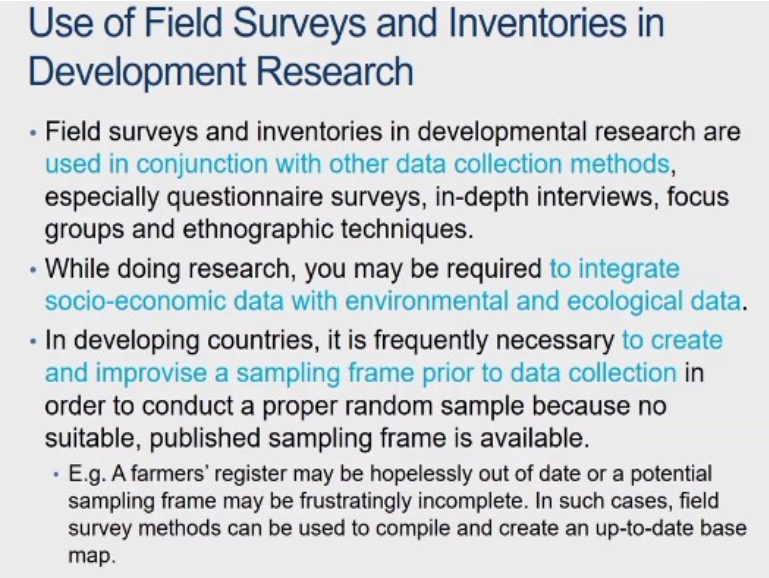
weaknesses are it seldom deals with the context of social life. It is inflexible because it cannot be altered once it has begun. It cannot be altered, because it is extremely resource intensive. It is a drain on the energy and financial resources of the survey team. Therefore, once you have begun the survey process, it cannot be altered. And therefore, it is said that it is poor for exploratory research. It is subject to artificiality. The findings are a product of the respondent's consciousness that they are being studied. Sometimes it is weak in terms of validity. It can be poor at answering questions where unit of analysis are not individual people.

Usually inappropriate for historical research and it can be particularly weak at gathering certain sorts of information. Example about highly complex or expert knowledge, people's past attitudes or behavior, subconscious influences, stigmatized behavior or attitudes and so on. Also, although survey research may nevertheless be able to achieve this in some circumstances.

(Refer Slide Time: 1:04:30)

Now, let us come to the final part of our lesson on field surveys and inventories. In this part I will make a very clear distinction between field mapping and field inventories and then come up with a few examples of them and that is how we will end this lesson.

(Refer Slide Time: 1:04:43)

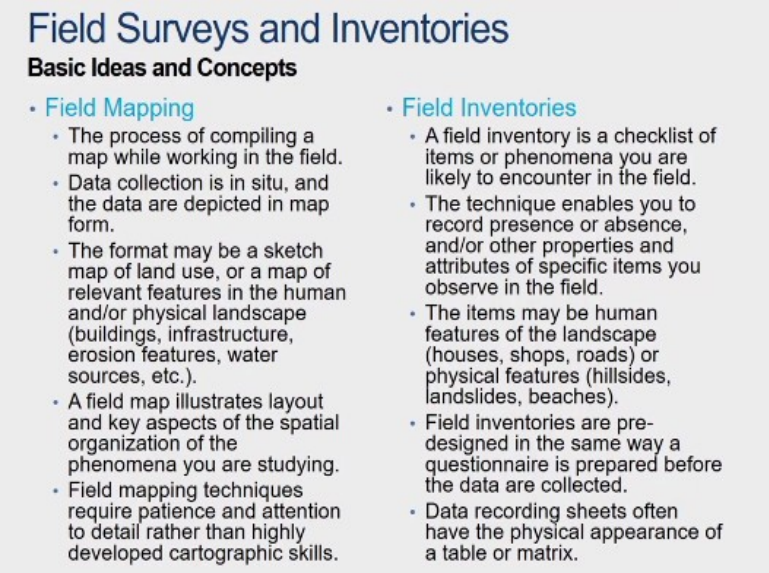


So, what is the use of field surveys and inventories in development research? Now field surveys and inventories are used in conjunction with other data collection methods especially questionnaire surveys and in-depth interviews, focus groups and ethnographic techniques. And while doing research we may be required to integrate socioeconomic data with environmental and ecological data.

In developing countries, it is frequently necessary to create an improviser sampling frame prior to data collection in order to conduct a proper random sample because no suitable published sampling frame is available usually. Now usually because in developing countries we do not have much computerized data of different locations that we are studying, it becomes important for the researcher, the investigator to design the sampling frame in the first place. But designing a sampling frame is a task in itself because you do not have pre-ordered numbers. You are creating those numbers keeping certain characteristics in mind. And that is where field mapping and field inventories bear a lot of importance.

Example- a farmer's register may be hopelessly out of date or a potential sampling frame maybe frustratingly incomplete. Suppose we are studying land reforms in a location and the farmers do not have pattas to be able to tell you who holds what amount of land. So, any amount of further inquiry with regard to land reforms will break down. So that is where you need to do some kind of field mapping in the first place to be able to carry out your research investigation. So, field survey methods can be used to compile and create an up-to-date base map.

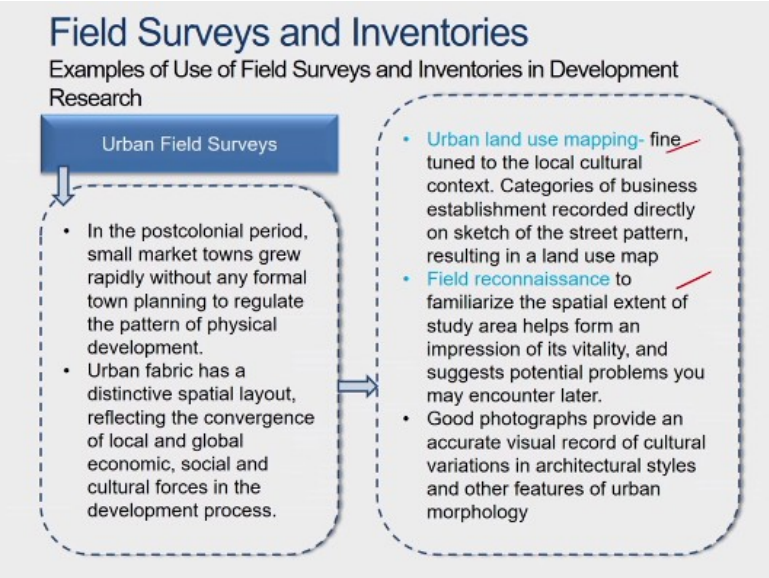
(Refer Slide Time: 1:06:28)



Now the basic idea and concept, difference between field mapping and field inventories are as follows. Field mapping is a process of compiling a map while working in the field. Data collection is in situ, it is collected there, and the data depicted in a map form and the format maybe a sketch map of land use or a map of relevant features in the human or physical landscape. For example, buildings infrastructure, water bodies, irrigation facilities, and so on. A field map illustrates layout and key aspects of the spatial organization or the phenomena you are studying. And field mapping techniques requires patience and attention to detail rather than highly developed cartographic skills. Because these maps are to be practically used for furthering the process of investigation.

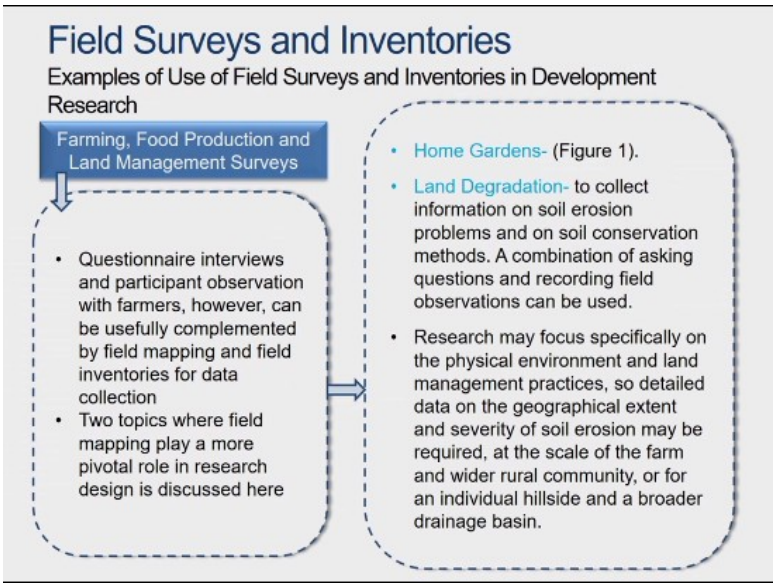
Field inventories on the other hand is a checklist of items or phenomena you are likely to encounter in the field. The technique enables you to record presence or absence or other properties and attributes of specific items you observe in the field. And the items may be human features of landscape or physical features. Field inventories are pre-designed in the same way a questionnaire is prepared before the data are collected. And data recording sheets often have the physical appearance of a table or matrix.

(Refer Slide Time: 1:07:41)



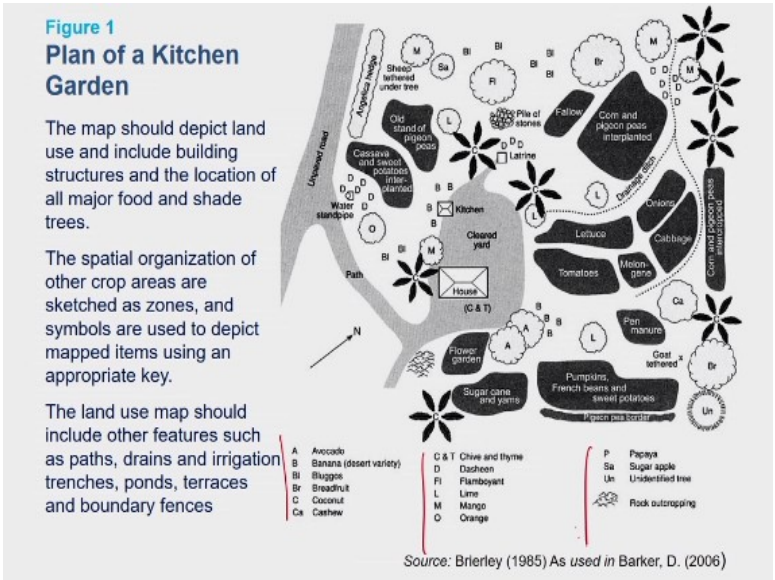
Now example of use of field service and inventories in development research. I will end the lesson with just three, four examples. One is urban field service. In the post colonial period, small market towns grew rapidly without any formal town planning to regulate the pattern of physical development. So urban fabric has a distinctive spatial layout reflecting the convergence of local and global economic, social and cultural forces in the development process. So, you can have urban land mapping field reconnaissance. Urban land use mapping, fine tune to the local cultural context. Categories of business establishment recorded directly on sketch of the street pattern resulting in a land use map. Field reconnaissance to familiarize the spatial extent of study area helps form an impression of its vitality and suggests potential problems you may encounter later. Good photographs also provide an accurate visual record of cultural variations in architectural styles and other features of urban morphology.

(Refer Slide Time: 1:08:41)



Another example of use of field surveys and inventories is farming, food production and land management surveys. Now questionnaire interviews and participant observation with farmers can be usefully complemented by field mapping and field inventories. Two topics where field mapping play a more pivotal role in research design is discussed here. One is home gardens, land degradation.

(Refer Slide Time: 1:09:04)

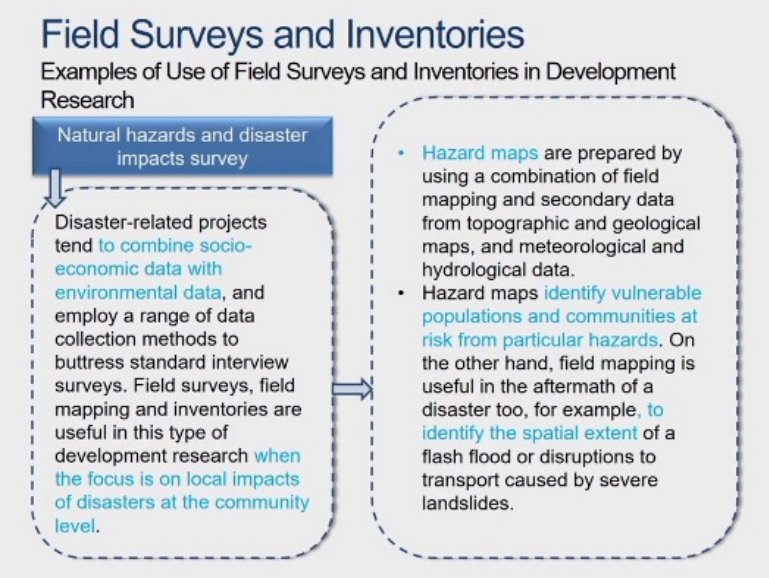


Home gardens if you look at the figure here this is a plan of a kitchen garden. This can help us explore this matter further. This map depicts land use and include building structures and location of all major food and shade trees. Spatial organization of other crop areas are sketched as zones and symbols are used to depict mapped items using

an appropriate key. And the land use map should include other features such as path, drains and irrigation trenches, ponds, terraces, and boundary fences. So, you have the legends here, the keys that help you explain the kitchen garden map.

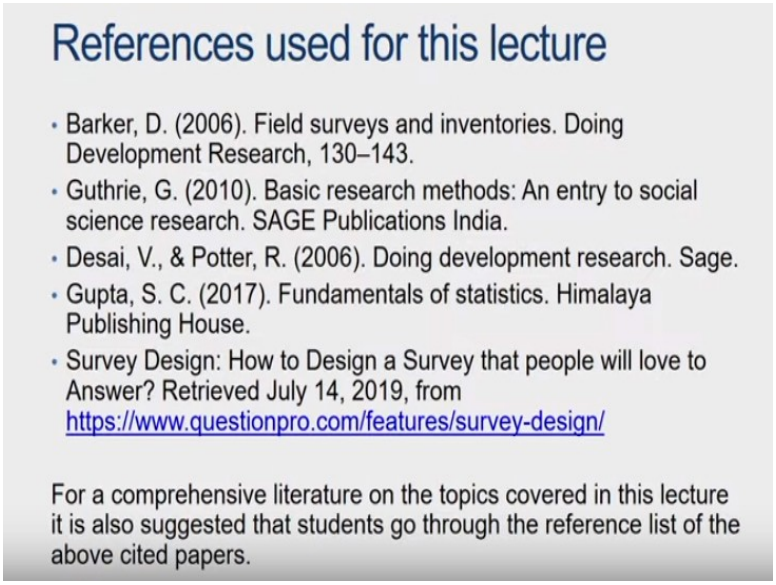
Land degradation to collect information and soil erosion problems and on soil conservation methods. A combination of asking questions and recording field observations can be used. And this research may focus specifically on physical environment and land management practices and so on.

(Refer Slide Time: 1:09:55)



Natural hazards and disaster impact survey, this is another example. Disaster related projects tend to combine socio-economic data with environmental data and employ a range of data collection methods to buttress standard interview surveys. So here also it is to look at the local impacts of disaster at the community level you can create hazard maps. They identify vulnerable population and communities at risk from particular hazards.

(Refer Slide Time: 1:10:23)



These are the references that I have used for this lesson. For those interested in exploring the details of survey research methods are requested to go through these references and also for a comprehensive literature on the topics that we have covered, I suggest that students go through the reference list of the above cited papers in the books.

I will see you in the next class. Thank you.