

**Population and Society**  
**Prof. A. K. Sharma**  
**Department of Humanities and Social Sciences**  
**Indian Institute of Technology, Kanpur**

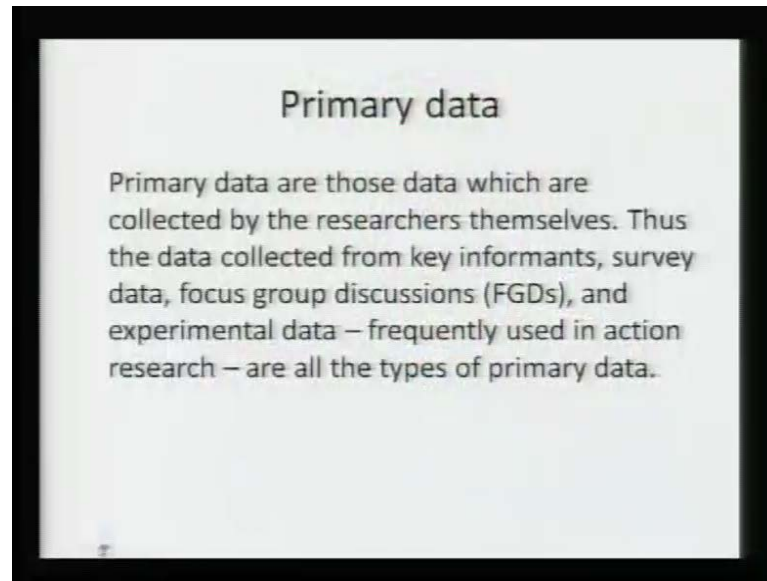
**Lecture No. # 08**

**Primary and Secondary Data and Related Issues.**

Well friends, this is the second lecture of the module on methodology. In the last lecture, I made a distinction between quantitative and qualitative methods. I said that in sociology in general and in population studies, we use quantitative methods and qualitative methods both. And sometime, in particular these days in operations research scientists tend to favor a combination of both quantitative and qualitative methods. Quantitative methods are particularly good when the purpose is to test a hypothesis or to generalize results for larger sample or larger population. And qualitative methods are particularly good when the purpose is to obtain insightful information, rich data, holistic understanding of the subject or subjective meanings.

Now, in practical research we call operations research, in which the purpose is to test hypotheses for program interventions donors, these days most of the time expect that we combine the methods. So, the old debate old sociological debate, but when a quantitative and qualitative method is not so important now. We are trying to combine methods so that we can benefit from the good points or stands of merits of both the methods. Today we will talk about various types of data and tools and techniques that are used in collection of data. This will cover both quantitative and qualitative, when it comes to collection of real data sociologists often make a distinction between primary and secondary data and we will talk about them and related issues.

(Refer Slide Time: 02:22)



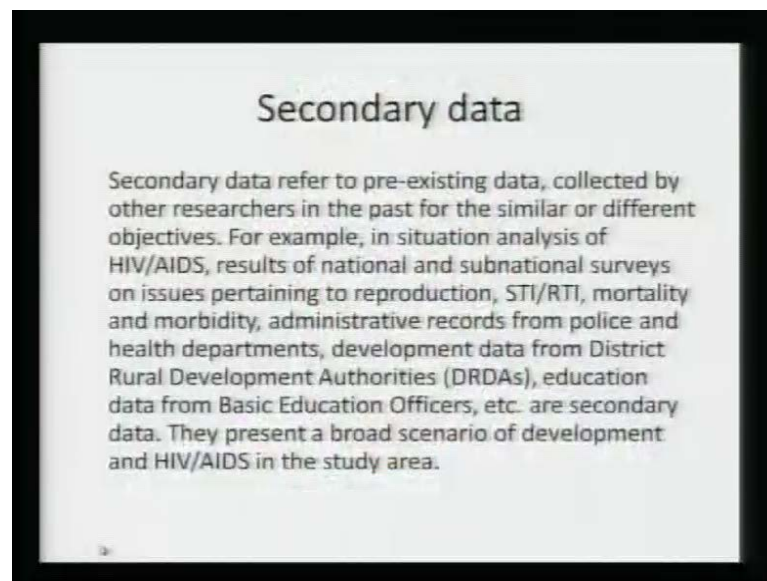
Primary data are those data which are collected by the researchers themselves. You know not always all researchers work with their own data, most sociologists have worked with data collected by others. Even important people Sorokin, Kingsley Davis, Emile Durkheim, Karl Marx, Boudreaux very few important sociologists have collected their own data; they have worked with others data but these days the field of practical research or operations research or action research we are also expected to collect our own data. Because the data on the subject on which we want to work may be lacking. And to both these types of data collected by others and data collected by us pose different kinds of difficulties. They have their own merits and demerits and therefore, sociologists often make a distinction between primary data and secondary data.

Primary data as I said are those which are collected by researchers themselves. So, If am interested and say international migration and I collect my own data from Kanpur city regarding from a selected area or from a selected sample, how many people have ever gone to other countries means abroad? For what purpose? What were their demographics, age composition, sex composition, marital status, education? What has been their experience of going abroad? Did they send remittance or how did they spend their money which they earned in abroad? And did they continue to stay at the place of destination or have they come back? If I do a survey of this kind I am collecting my own data this is called primary data.

But I may also write a good research paper or write a thesis or develop a theory on the basis of secondary data means, data on international migration already existing collected by some other researcher or by some other agency dealing with migration statistics. It may be a passport office or it may be data collected compiled or provided by United Nations organizations dealing with population trends. They are lots of agencies to collect to compile and to disseminate information about migration. International migration is anyway an important issue. If I am using data collected compiled disseminated or even analyzed by others. I am making use of secondary data.

So, that way primary is my own and what comes under primary data. We will talk about all these methods tools a little bit today. Data collected from key informant's survey data, focus group discussions and experimental data, frequently used in action research are all the types of primary data. They are data collected by the researchers and researchers or groups of researchers collect their own data using interviews of key informants or sample survey or focus group discussion or by conducting experiments.

(Refer Slide Time: 06:21)



Then the secondary data as I said that, secondary data refer to preexisting data. Data that already exist, collected by other researchers in the past for the similar purpose of a different objectives. Sometimes data may have been collected for some other purpose, but the nature of data available is such that it can be used for our own purpose also. You know like routine, administrative statistics collected by different departments of

government states, central can be used for research purposes. For an example in situation analysis of HIV aids, results of national and sub national surveys. National surveys are conducted by population council, international council for women's research, international institute of population sciences, and ministry of health and family welfare. Private research organization like ORG operations research group, mod mar and so on.

They are national surveys sometime the same agencies also conduct in-depth studies in specific states or specific districts. They may be called sub national surveys like cheer cop project, which was taken up to study the impact of interventions in HIV aids. This project was implemented only in its limited number of only in certain districts of certain states. And then base line and end line surveys were conducted they can be called sub national. So, surveys which are conducted at the state level district level or lower level, but which would provide insightful information for our purpose. They are sub national surveys and is STI, RTI these are the topics on which surveys national and sub national have been conducted STI sexually transmitted infection, RTI reproductive tract from infection, mortality, morbidity patterns or diseases, administrative records from police and health department.

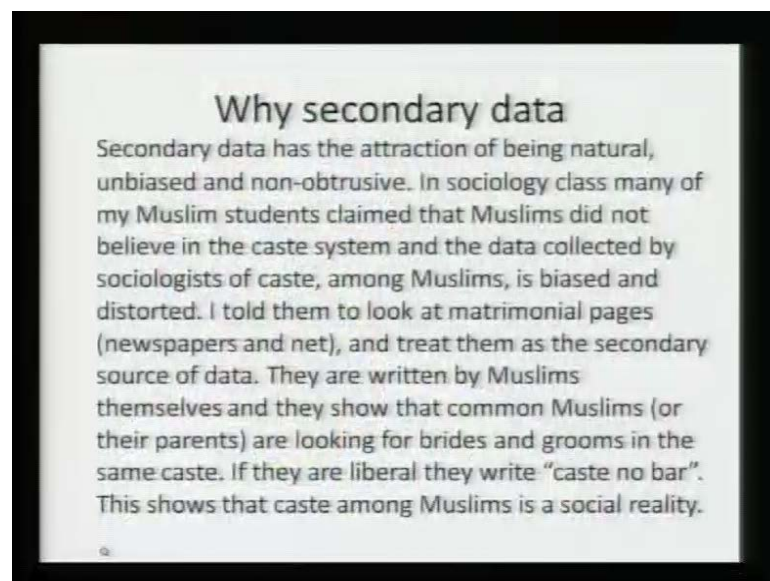
Development data from district rural development authorities, many types of data are collected by district rural development authorities not for the purpose of research. But for the purpose of monitoring of the program, or as part of routine collection of statistics or data pertaining to development program. So, DRDA can provide you data, on how many groups were formed in which year, whether they are active or inactive, who finance them what about the size and what kind of activities were taken up by different self help groups. How much money was in all spent on different types of rural development activities in different years in different financial years for last 5 or 10 years.

Similarly, we have education data, education data routinely collected by basic education of research, if it pertains to basic or elementary education and by district inspectors of schools, if it pertains to say high school and you also have data on higher education with UGC or state departments of higher education. And they can be used they can they are called secondary data, suppose somebody wants to work on crime in India then certain departments dealing with justice, crime, punishment you know police department for example, they are concerned with crime situation and they routinely collect data or they maintain data on various types of crimes in the country.

So, from those data you can know, what has been the crime rate, what types of crimes were involved, and are there regional differences urban rural differences, it depends on what kind of data are collected. If the agencies for routine administrative purposes collect or classify crimes according to urban and rural areas then you can find out differences in crime rate between urban and rural areas. If they maintain record state wise then you can find differences in rates in state wise, if they also maintain records according to age of the victim or off ender or gender then you can have data according to that.

They represent a broad scenario of development in HIV aids in the study area. You know the things which I mentioned RTI, STI reproduction, mortality, morbidity and so on. These are these are the data which can be obtained from ministry of these days many ministries have provided lot of data on their website. Recently I was using the data of Mahatma Gandhi national rural employment scheme; ministry of rural development provides lot of data regarding Noriega. Implementation of Noriega the country including number of beneficiary's categories of beneficiaries and many other related issues including issues connected with social audit. These data can be obtained from websites total sanitation program Indira abasyojana there are many government schemes on which secondary data are available and one can use this secondary data.

(Refer Slide Time: 12:40)



Why secondary data? Secondary data has the attraction of being natural unbiased and non-obtrusive is non-obtrusive this term requires some clarification. In sociology we often say that the method of data collection affects the quality and nature of data, means methodology affects the outcome of the study. Yes sometime methodology may affect the outcome of the study. Your respondents may not reveal the truth or they may underestimate or overestimate certain things or there may be non response or by the very fact that somebody is observing them, but the very fact that an outsider is present amongst them and observing their behavior can change their behavior.

You want to study say communal relations in a village and you are present in the village and you are and the people know that you have come from outside. They may have their own interpretation of yours whether you are from university, from government, from police, from an outside agency. And they will have their own assessment of your motive also. Their assessment of your background and their assessment of your motive can certainly influence the relationship between different communities in the village. One example which I always give in the class, the problem of surveys right from the beginning since 1950 when surveys of say ideal family size are conducted and we ask what is the ideal number of children or how for a family of your type for a family in your social economic circumstances. What is the ideal number of sons? And what is the ideal number of daughter?

They know that you have come either from university or from government and the standard answer, they will give for which we have a technical term social desirability, always our surveys have yielded the same result two sons and one daughter. When fertility was seven or eight total fertility rate was seven or eight our ideal family size was around three. And today when fertility has declined to around three even then our ideal family size is same. Why does this happen? Because our methodology of data collection affects the outcome of research, but if people are unaware that somebody is watching them, somebody is observing them and the data are collected in a natural setup then such data can be more reliable.

So, for survey kind of research or for our own focus group discussion sometime we say it is obtrusive. And for data already existing in some form and may be which was collected or which was presented by other parties for some other purpose those data are more reliable. In sociology class many of my Muslim students claim that Muslims do not

believe in the caste system. So, when I say that in India is universal and caste is found in all religions, caste is found among Hindus, Muslims, Buddhist, Janis, Sikhs Parsifal community all religious communities have caste.

Partly because today people belonging to different communities other than Hindus a large number of them are converts. Some 70 these are true estimates, but one can say that say around seventy percent of Muslims are converts from some other caste. Similarly Sikhs are converts Buddhists are converts and therefore, caste system is everywhere. But due to religious or ideological bias political ideology or religious ideology students tend to say that no Muslims do not believe in the caste system. If I decide to collect primary data on caste among Muslims mean I go to the field I talk to people conduct a survey or focus group or interview people Muslim people and ask. Do you believe in caste? What are different castes among Muslims? Their responses may be bias and there is a tendency among Muslims to say that no they do not believe in the caste, in the institutions of caste because such institutions are against the basic spirit of Islam.

So, I tell my students that look at the matrimonial pages of news papers and net and treat them as secondary data. Now this matrimonial are not sent by government, if you think that state has invested interest in maintaining caste among Muslims. The matrimonial are not sent by state, matrimonial are not sent by intellectuals or social matrimonial are sent by people, who need them Muslim themselves. And then I tell students note down how many matrimonial from Muslims mention their caste. And you know in our country and anybody who sends matrimonial knows that for writing caste no bar, there is some commission, some concession, and some twenty percent it depends on newspaper twenty percent, thirty percent. If you write caste no bar then you will get thirty percent concession.

And therefore, it is quite rational for writing it is quite rational to write caste no bar. So, that you can get the concession, but still people write caste that is because they are looking for a suitable match in their own caste. And if you analyze the data on data from matrimonial form Muslims you find that most of the time caste is mentioned they write (( )) this or that. And sometime they also write that match is invited only from such caste and this is, this practice is so common that not only upper cast. But even lower caste maintain the caste distinction in matrimonial we show the caste is universal secondary data.

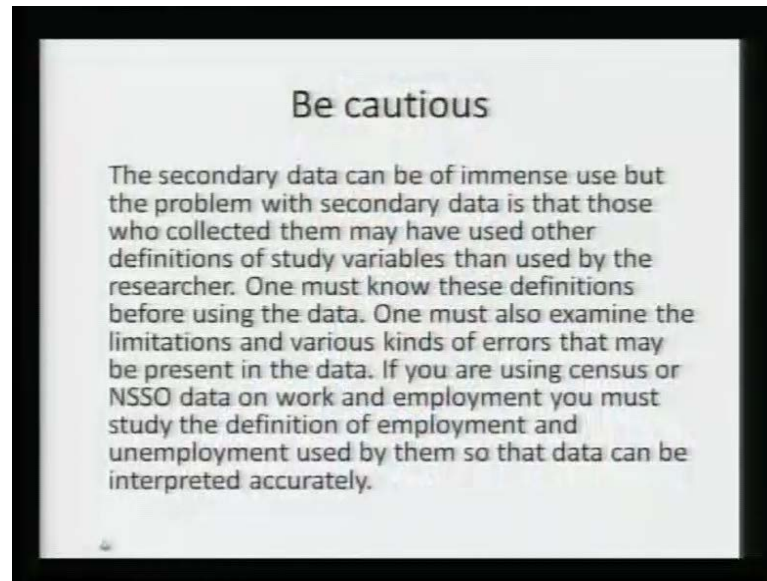
The other day I got some matrimonials to the class and I showed that one matrimonial sent by a Gujarat. Gujarat you know that Gujjars are fighting for getting the status of schedule tribe in Rajasthan. Right now they are under OBC and they want the status of schedule tribe. They think that now in the category of OBC. They cannot compete with other OBC like jatts, but if they are put in the category of schedule tribe then they will benefit.

So, on the one hand there is a claim that they belong to a lower community or a lower caste or they have most disadvantages and all. And at the same time they maintain the caste side, even the lowest of the low maintain caste identity. So, the matrimonial I was referring to from some Gujarat that say that they belong to Gujarat community and they are looking for this kind of match and certain categories of Gujjars they mentioned and Bhatt's that they should not apply. So, this is secondary data secondary we are not collecting data, data already exist. When we collect data from basic education officer's office and suppose that shows, that there is there is a great gap number of enrolments between males and females urban and rural that is real.

In such cases we can expect that there is an over estimate sometime over estimate of enrollment because the junior officers have to send these data to senior officers for and that may somehow be connected with their evaluation. So, there may be a tendency to over report, but if you still find that there are gaps; there are differences in enrolment rates according to gender. Or if you find that by the time students should reach class ten seventy percent of them dropout. This can be taken as secondary data to support the claim that there is lot of school dropout in certain regions in the country and that there are differences in enrolment rates according to gender or there are differences in dropout rates at various levels, primary level, middle level, high school level intermediate or higher levels.



(Refer Slide Time: 21:56)



However one has to exercise caution when using secondary data, because secondary data can be collected with a different purpose and with different definitions with different motives. So, the secondary data can be of human's use, but the problem with secondary data is that those who collected them may have used other definitions of study variables than your definitions. One must know these definitions before using the data, one must also examine the limitations and various kinds of errors that may be present in the data. If you are using census or NSSO data on work and employment you must study the definition of employment and unemployment used by them so that data can be interpreted accurately.

This is a very common problem in studies of employment rates; everybody defines employment in a different way or studies of poverty. So, many agencies have collected data on poverty. So, many agencies have collected data on employment your planning commission estimates, you have NSSO estimates, you have estimates from say national applied council of economic research sorry national council of applied economic research and there are other scholars, individual scholars very important reputed persons well estimated poverty. They many times the differences in their findings arise from the fact that the surveys conducted by them use different definitions of employment. Sometime in the same source also like census. I told once that census started in India around 1872 and after that we have had decadal censuses conducted in uninterrupted manner every ten years. And many definitions have been changing definition of, what is

urban has been changing definition of what is work? Work has been defined in different censuses in different ways. So, if you compute level of urbanization or if you compute work participation rate and compare them over the years over census decades. Then there is a difficulty of correct interpretation of changes in work participation rates or urbanization. Because in those years in which definitions change you cannot compare two census figures, census figures from two consecutive censuses.

When the definition have changed are exactly cannot be compared. So, sometimes the change of definition can lead to over estimation of urbanization or sometimes it can lead to under of urbanization. One example of this is literacy up to a certain point, literacy rate in India was computed by taking all the number of literate people in the numerator and the whole population in the denominator. But in one census here it was decided that in the denominator we will take population of seven plus only. Now, once you reduce your denominator from total population to seven plus, immediately your literacy rate goes up.

So, some people may be some suspicion or cynicism may also say that the purpose was to show that the literacy rate has gone beyond 50 years 50 percent. That may not be the motive and this may be quite to have in denominator seven plus, because we do not expect zero to six a small children to be literate. So, seven plus make sense lot of sense, but the point is that if the definition of literacy or definition of literacy rate is changing then the data on literacy rate obtain from consecutive censuses in which different definitions were used are not comparable.

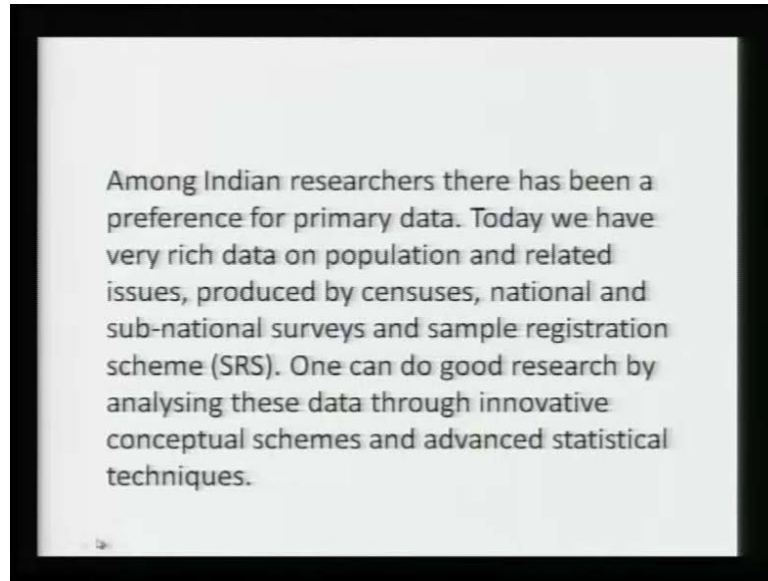
So, when you are using secondary data you must be very careful about, what definitions were used? What wordings were used? How question or how items or how scales were standardize? What were the reference periods? In the study of poverty reference period creates special problems sometimes in consumption surveys reference period is one week sometime one month for some purposes it may be one year. So, it depends in lot if you are using secondary data then you must be thorough about these things. Sometimes secondary data may also contain other types of errors. For example, age data of census, age data of census contains lot of errors there is something called age hiping, students of population know that ah our respondents have a tendency to report ages ending in certain digits like 0 and 5 and a tendency to avoid digits like nine and three.

So, if you look at age distribution of a population in single years, you find lots of ups and downs which are absurd and this absurdity arises only from digit preference in reporting age data. So, you must be aware of that and ways must be found to identify the kind of errors you have in the data. And to modify data or to rectify those errors in the data before the data are used for planning or forecasting for analytical purposes. Now, you have why secondary data is very useful, sometime more useful and cheaper also, you know you want to study seasonal migration from some region. Now you can very well conduct a survey, but conducting a survey would; obviously, require lot of resources in terms of man power time money.

If you have secondary data on how many tickets are sold for which destination from the local railway station that is a very good source of data on migration. So, you'll find that in certain month from some eastern U P district more tickets were sold for Mumbai, in some other seasons less ticket were sold. Similarly, more people one can observe by sitting at the railway station that in some months more people arrive from Mumbai or other destinations in some months and less in others. This data the data secondary data already existing data on number of ticket sold in different months for different destination are very useful data.

And you can combine two types of data, you can use secondary data for outmigration for seasonal outmigration. You can use observations on primary data for seasonal migration to the to the place where you are studying means in migration. For outmigration you can see data on tickets in migration .You can go for observations or you can go for seasonal in migration to railway stations of destinations Mumbai or Calcutta or whichever other major destinations and obtain data on ticket sold to different places from there.

(Refer Slide Time: 29:46)



There among Indian researchers there has been a preference for primary data; however, we have very rich data on population and related issues produce by censuses, national and sub national surveys and sample registration scheme. One can do very good research by analyzing this data, through innovative conceptual schemes and advance statistical techniques. You can produce very good papers. Our PHD students generally M Phil, PHD students have a tendency to collect their own data and they feel very happy by saying that they have their own primary data to support their hypotheses.

Actually there is so much of data available from secondary sources from ministry websites, from IIPS, from research organizations and from government departments, from statistics collected routinely for administrative purposes. That you can do very good research, extensive research, in-depth research, just by analyzing those data. Unfortunately lots of data are collected in census and in SSO and our academicians have not done justice to those data. Professor Bose in one of his essay in EPW said that all research organizations in the country are more busy analyzing data from NFH SRCA national familiar survey, reproductive and child health survey etcetera.

And so much of data is wasted in censuses and surveys which are of very god quality. Census data, NSSO data, very good quality of data different dimensions on complex issues of socioeconomic phenomena and there is a need to do research on that. So, somebody you need not always collect your own data of course, if you are working on

something new, like somebody is working on the issue of palliative health care. There is no research on palliative health care even the idea of palliative health care is new. So, you have to collect your own data on palliative health care, to study what are the needs of the clients, to study what are the problems of care givers, to study what are the motivations of volunteers, to study in what kind of setup palliative health care is more successful.

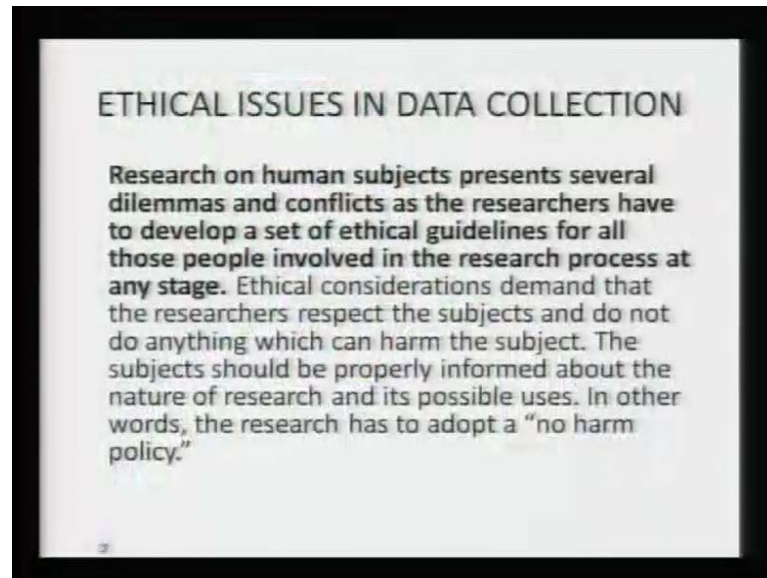
These days we UN agencies are also talking of Kerala model of palliative care. So, some to study Kerala model of palliative care, you have to collect primary data. And we have seen that in Kerala also there are diverse approaches to palliative health care not in all the districts you have the same model. In some district you have community based model, in some district you have NGO based model, hospital based model and there are differences. There are differences in successes, failures and one needs to collect primary data on needs of client families on care givers, or those specific members of family, who provide palliative health care in the family setup or volunteers, or doctors or professionals, or community members, or donors.

Sometime religious organization sometime NGOS sometime government departments are providing funds for running palliative care units. So, you require primary data there see it depends, but while discussing primary and secondary data among students like students of sociology and population. I would say that in our country today there is lot of good quality secondary data of collected by decadal censuses and NSSO and organizations. Like if you approach organizations like applied economic research they will be happy to give you data these days research organizations are ready to provide their data. So, national familiar survey RCH these data you want their data, you write to them to their director whoever is in charge of data collection unit and they will give you data sometime they charge some money for that.

And these data can be the advantage of using secondary data would be that these people who have these organizations who have collected secondary data are specialized organizations. And they have spent lot of funds. That they have ensured good that the data of good quality are collected by training, by spending more money, by designing tools and techniques, more properly, academically, by consulting national and international experts in the field. So, you have good quality data and you in your own PHD or M Phil research if you go for primary data. May be you can collect data from

hundred people two hundred three hundred here in secondary data you have as large sample as 50000. So, why not use secondary data I would say that you should use secondary data as often as possible.

(Refer Slide Time: 35:13)



There are also some ethical issues in data collection, the term ethics means moral, social. So, far unfortunately in our country though the total size of the quantum of social science research was very high. Generally we did not bother about ethical issues very few students of social sciences have even heard this term ethical issues in data collection. Whatever guidelines for ethics that develop they were developed in the context of medical or biological researches. Social scientist who are completely free to take up any kind of problem use any method, it was on their own fancy and whims to decide what kind of data and how will be collected. Ethical concerns are new concerns let me tell you that if you are going to publish your paper in international journal today in reputed journal, they will certainly ask you whether your research was permitted approved by an ethical committee of their institute.

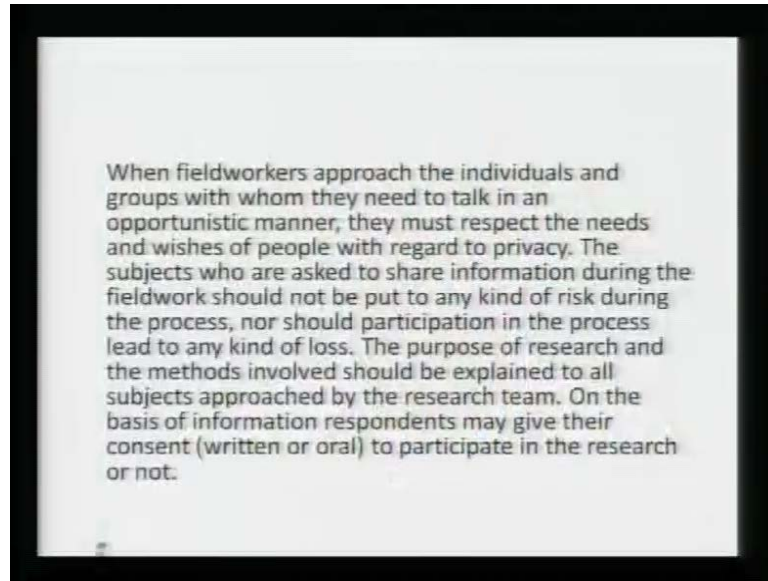
In if you are working for a donor organization like path finder's international or population council they will ask whether you have followed the ethical guidelines. And whether your tools and techniques, your survey methodology, your items and your whole approach, you know they were approved by the ethical committee of your institute. Ethic means that behave in a moral manner, behave in a socially responsible manner. What is

to behave in a morally responsible or socially responsible manner in research? It is like this, Research on human subjects presents several dilemmas and conflicts as the researchers have to develop a set of ethical guidelines, for all those people who are involved in the research process at any stage.

You know they are your senior researchers, supervisors, field workers they are your subjects they are your coder's data analyst statisticians donors even so anybody who is involved in the research has to follow certain ethical guideline. Ethical considerations demand that the researchers respect the subjects and do not do anything which can harm the subject. The subject should be properly informed about the nature of research and its possible uses, in other words the research has to adopt a no harm policy. Now I cannot go to people to say bureaucrats or middle ranking officers and tell them that I am collecting data on corruption on bribe, please provide me data and suppose in good faith thinking that I am only a social scientist and I will only publish papers. I collect they give me lots of confidential or personal information about them about their malpractices corruption bribe with their own justifications and or may be if they have hidden some income from income tax returns.

Now, I am not supposed to use that information. I am not supposed to write to commissioner of income tax that such and such person has falsely reported his or her income on the income tax return. I am not supposed to diver information which is of confidential nature to the community. Normally in surveys I am suppose to present only the findings of the study and information of confidential nature about a specific individuals is not to be diverge that will be unethical. Anything which can harm somebody which can harm my researchers or which can harm a particular community. Normally such types of data should not be diverge or such type of information about specific individuals should not be diverge.

(Refer Slide Time: 39:51)

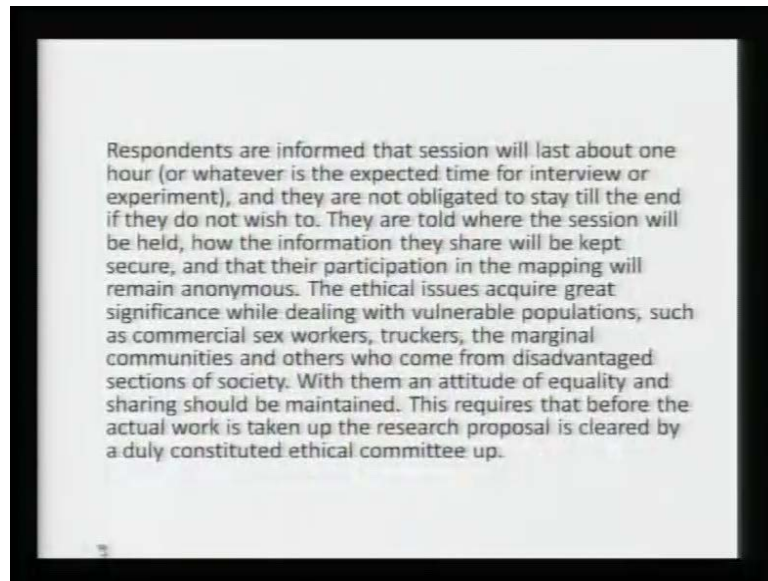


When field workers approach the individuals and groups within with whom they need to talk in an opportunistic manner. They must respect the needs and wishes of people with regard to privacy in this privacy, confidentiality, privacy respecting your subject. These are other issues the subjects who are asked to share information during the field work should not be put to any kind of risk, during the process nor should participation in the process lead to any kind of loss may be loss of wages or loss of prestige. The purpose of research and the methods involve should be explained to all the subjects approached by the researchers team.

On the basis of information respondents may give their consent to participate in the research or not. So, written or verbal these days it has become an important piece of information in research to right whether written or oral consent of the participants was taken.

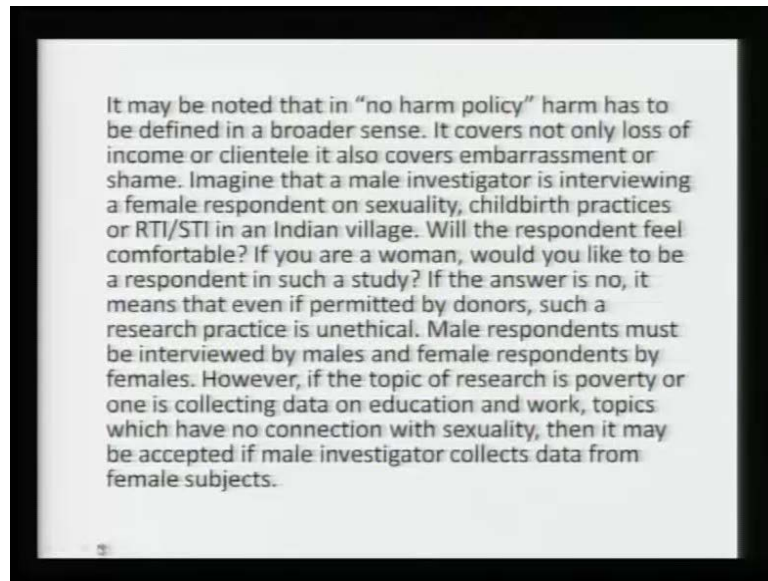


(Refer Slide Time: 40:56)



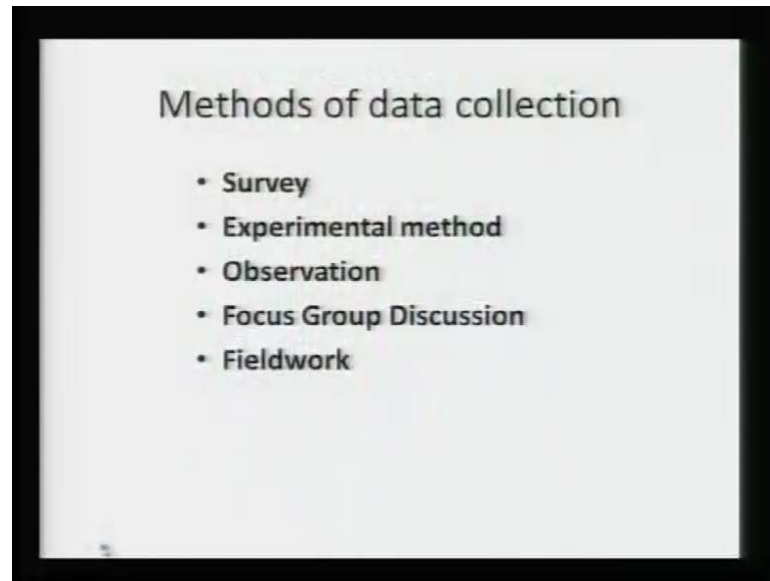
Respondents are informed the sessions will how long your sessions will last? What kind of information will be collected? How you are going to make use of that information for whose purpose this research is being conducted? The ethical issues acquire great significance while dealing with vulnerable population such as commercial sex workers, suppose in some group in confident they tell me that yes there are sex workers, number and distribution. I am not suppose to provide this research information to local police or local administration, because then it can be use to exploit the sex worker truckers marginal communities and others who come from disadvantaged sections society. While working with them an attitude of equality and sharing should be maintained. This require that before the actual work is taken up the research proposal is cleared by a duly constituted ethical committee.

(Refer Slide Time: 42:02)



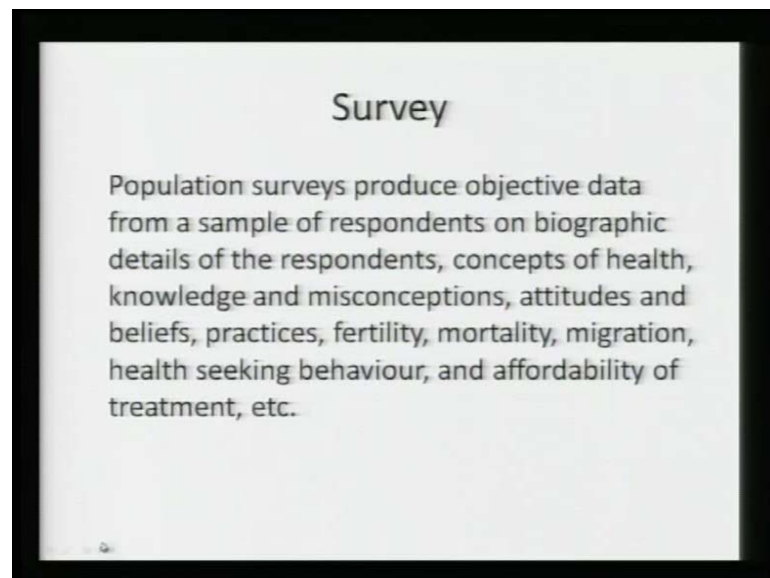
And this no harm policy is not simply loss of wages. I just give an example like you are collecting data on R T I, S T I from women or P P H postpartum hemorrhage from women. It is not ethical or rightful by sending it is not ethical or rightful to send men investigators for collecting data from women. In such sensitive matters male respondents must be interviewed by males and female respondents by female. It will be very embarrassing for women and good quality data will also not come if you are sending men to interview women. However if the topic of research is poverty or one is collecting data on education which are sex neutral topics then it may be accepted if male investigator collects data from female subjects.

(Refer Slide Time: 43:07)



The methods of data collection primary data are survey, experimental method, observation, focus group discussion and fieldwork.

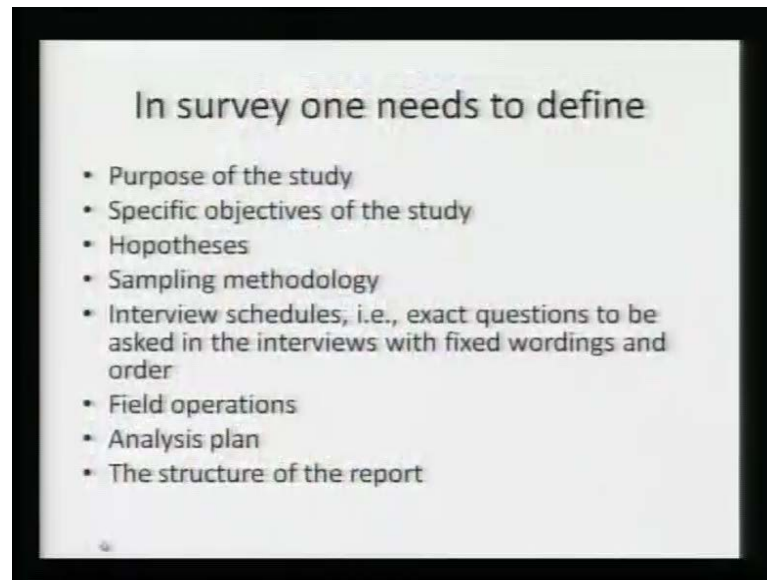
(Refer Slide Time: 43:19)



Survey, population surveys produce objective data from a sample of respondents on biographic details of the respondent's concepts of health, knowledge, misconception, attitudes, beliefs, practices, fertility, mortality, migration, health seeking behavior and affordability of treatment. Sometime in operation research quality of service, quality of services or counseling these have become some of the issues. Some of the earliest

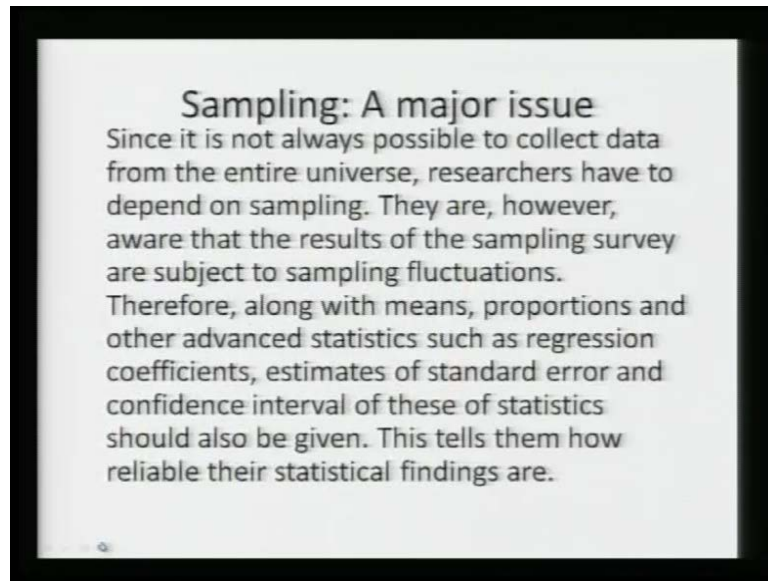
surveys in population studies were KAP type knowledge attitude and practice regarding fertility and family planning. These days more surveys of abuse, drug abuse, risk behavior, health, accessibility to health services, belief system, social representation these kinds of studies have become more common.

(Refer Slide Time: 44:17)



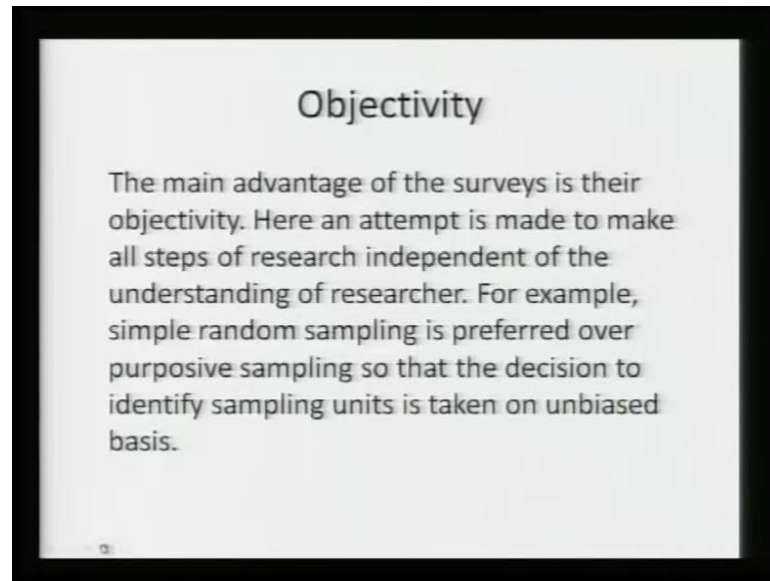
To conduct a survey one needs to define purpose of the study very clearly. A specific objective of the study hypotheses. If there are some hypotheses, sampling methodology or sample design, interview schedules that are exact questions to be asked in the interviews with fixed wordings and order. Surveys are believed to be part of scientific methodology. So, as far as possible our survey methodology should be free from any kind of bias, interviewers bias, researchers bias in selection of sample in wordings, in order of questions, in data analysis at all stages care must be exercise to make data free from biases of researchers or all those involved in the research. Field operations then analysis plan and the structure of the report.

(Refer Slide Time: 45:21)



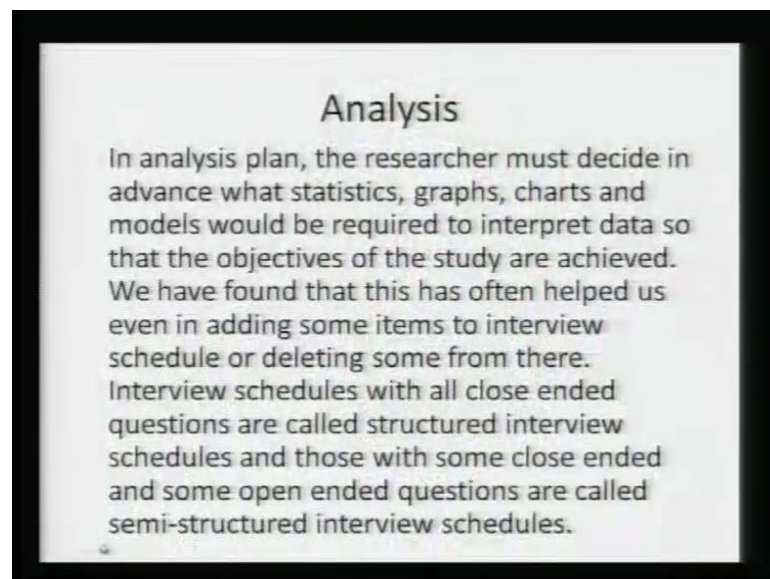
These are the things sampling. Sampling is a major issue; since it is not always possible to collect data from the entire universe researchers have to depend on sampling. They are; however, aware that the results of the sampling are subject to certain what we call sample fluctuations. You are not studying the whole population; you are studying only a party of that. Therefore, along with means, proportions and others advance statistics such as regression coefficients, path coefficient, and estimates of standard error. Even if you give say, (( )) of your data, courtesies of your data, fourth moment third moment regression coefficient of beta, everything that comes from sample study as sampling fluctuation. A subject to sampling fluctuations and therefore, has a standard error associated with it.

(Refer Slide Time: 46:27)



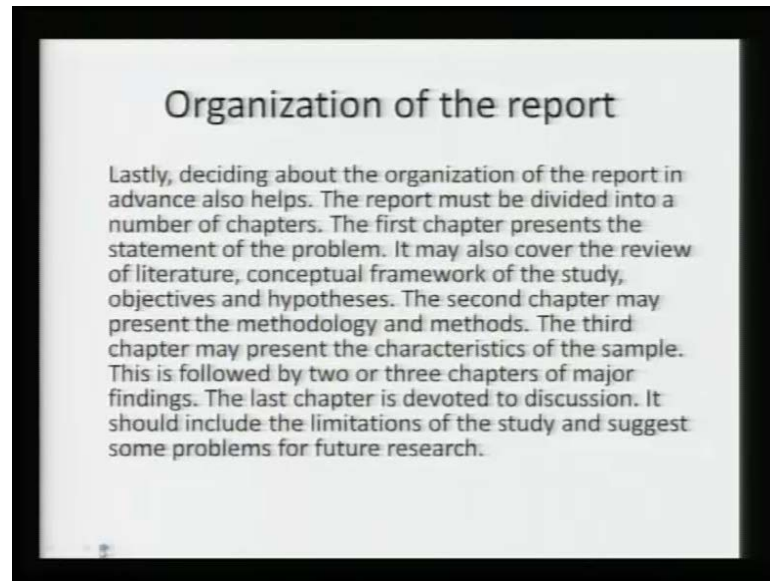
So, we must provide standard errors and these standard errors depend on sample size and therefore, it is important to think about the sample size. The main advantage their objective status. I have already said that the surveys are known for your objectivity and therefore, they are considered to be a part of scientific method and this also means that in place of going for purpose sampling where sample is decided by the researcher.

(Refer Slide Time: 46:53)



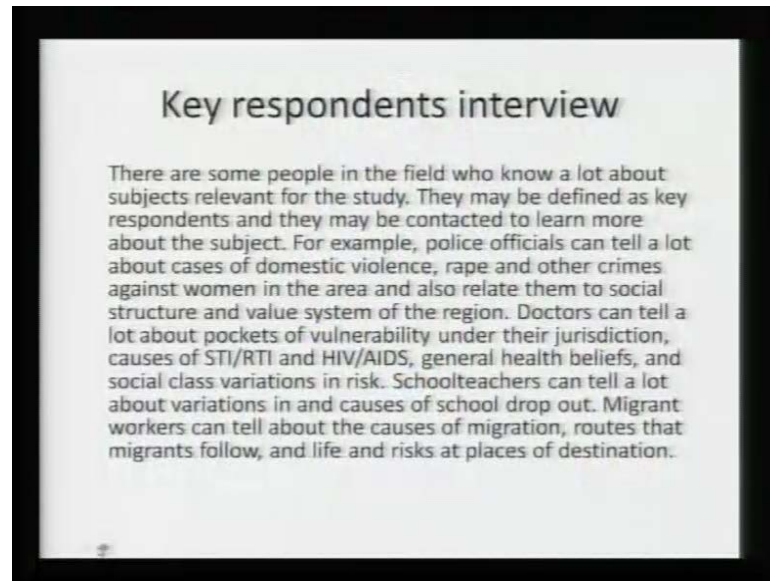
It is better to go for probability sampling then analysis some time later we will talk of about analysis more.

(Refer Slide Time: 46:57)



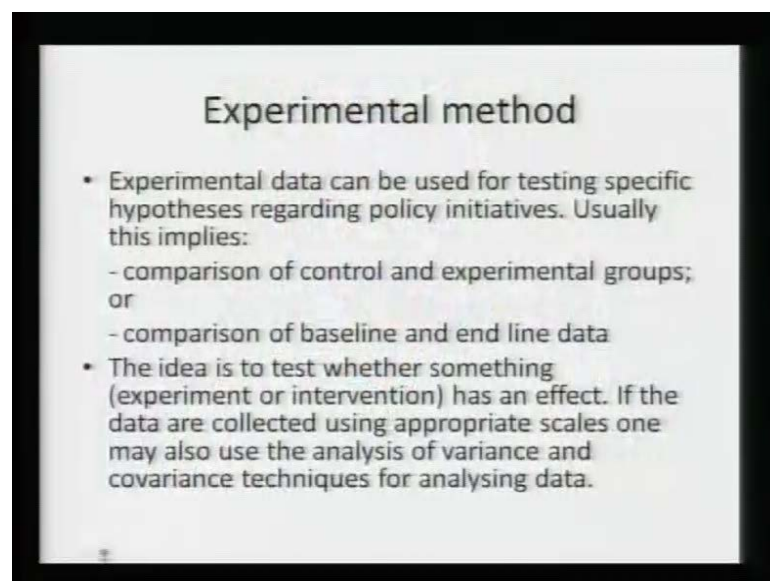
And then organization of the report normally organization of the report should be like this, it must be divided into a number of chapters. First chapter about the statement of the problem, review of literature, conceptual framework of the study, hypotheses and objective. These kinds of things must be stated about the research means everything about the research. Then the second chapter what methodology and methods you have followed quantitative, qualitative, combination, triangulation what. Then the third chapter characteristics of sample, because of finding depend heavily on whom your respondents for and then one or two or three chapters on the findings main findings of the study and last chapter is devoted to discussion.

(Refer Slide Time: 47:45)



Another method of key respondents interview. People who are knowledgeable about the field are called key respondents like doctors know about the nature of epidemics or mortality, police superintendent or senior police officers may know about the crime, village heads may know about caste relations, community relations or development activities or some rickshaw pullers in some high risk area may know about nature of sexual locations size and so on. They are key respondents and key respondents interviews can reveal a lot.

(Refer Slide Time: 48:29)

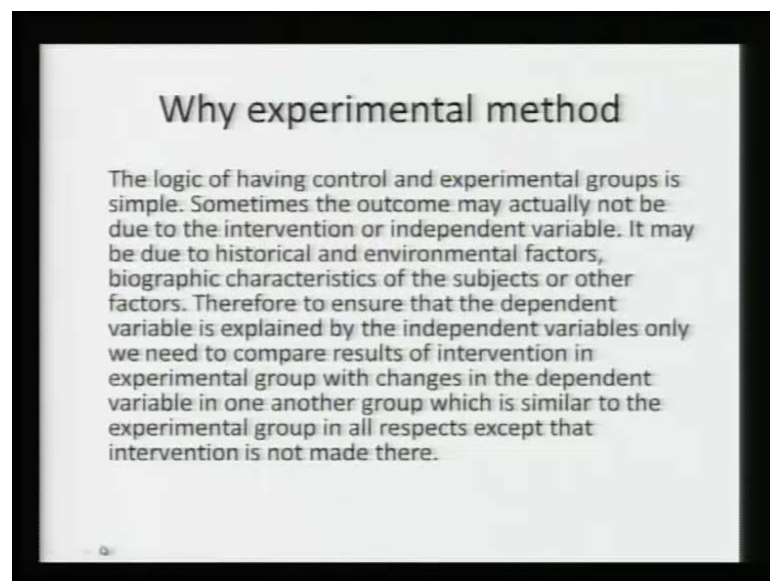




Experimental method wants in passing. I mentioned about experimental method. Experimental method can be used for testing specific hypotheses, usually this implies comparison of control and experimental groups or comparison of base line and end line data. Before base line means before an intervention is made, you collect social economic data and data pertaining to your study variables and then after the intervention has been made. Again you collect the same socioeconomic data and data pertaining to your study variables compare them. This is based line end line sometime you require only two data point sometime you may require collecting data at three four or more times.

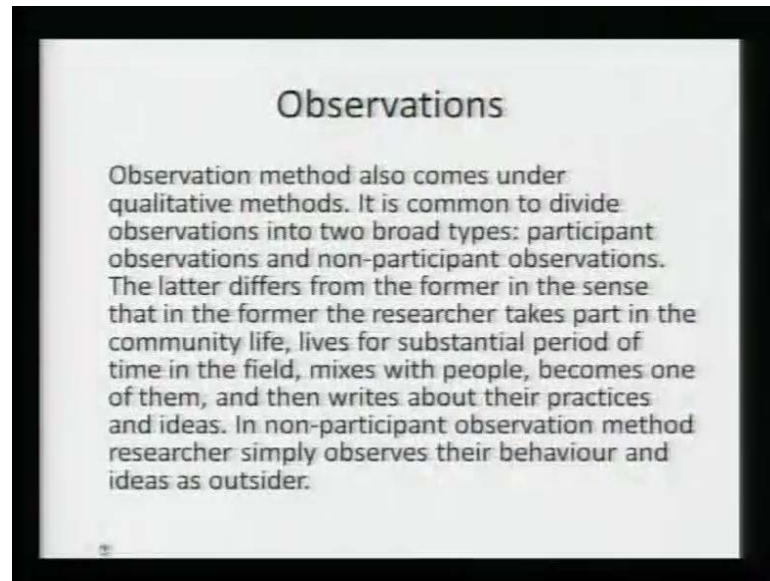
The idea is to test whether something has an effect and if the data are collected using appropriate skills, one may ask with the analysis of variants and co-variants techniques for analyzing data.

(Refer Slide Time: 49:24)



Experimental method is important because you know the findings may depend not only on intervention, but also on many other factors socioeconomic, biographic, demographic and by using the idea of control and experimental we try to ensure that the findings are the findings about effectiveness of intervention only and not about anything else.

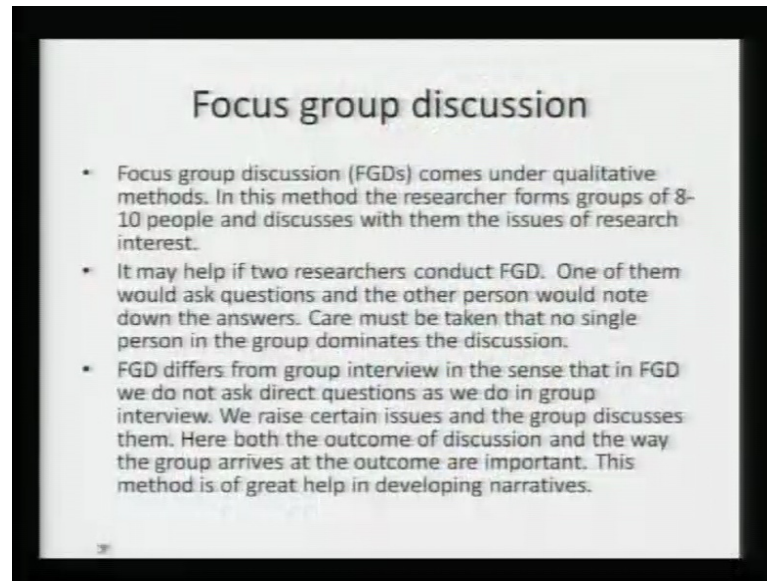
(Refer Slide Time: 49:53)



Then observations, observation method means observing you go to the field and observe. It is common to divide observations into two broad types, participant observations and non-participant observation. The later differs from the former means non-participant observation method differs from the participant observation method. In the sense that in the in the former, the researcher takes apart in the community life. A life for a substantial part of his life or for a substantial period of time in the field mixed with people becomes one of them and then writes about their practices and ideas. Anthropologists working in preliterate and small societies in alien cultures have generally used this kind of participant observation method.

Some went to the extent that they settled down in the community, that was to be observed even married their and spend entire life in the cultures that they wanted to study. It is not always possible for social scientist to become part of the reality and spend a large amount of time in the community and therefore, non-participant observation method, here researcher simply observes their behavior and ideas as outsider.

(Refer Slide Time: 51:18)



Then you also have focus group discussion. Focus group discussion comes under qualitative method, but it is a part of primary data; in this method the researcher forms groups of eight to ten people usually homogenous groups should be formed. It should not happen that all kinds of people are brought together and you are discussing something as far as like in a village situation in the study of population. It will be better if you have groups of men, women, young, old, unmarried, married recently we conducted some district situation analysis you wanted to know whether people know about condoms and whether they know about HIV aids.

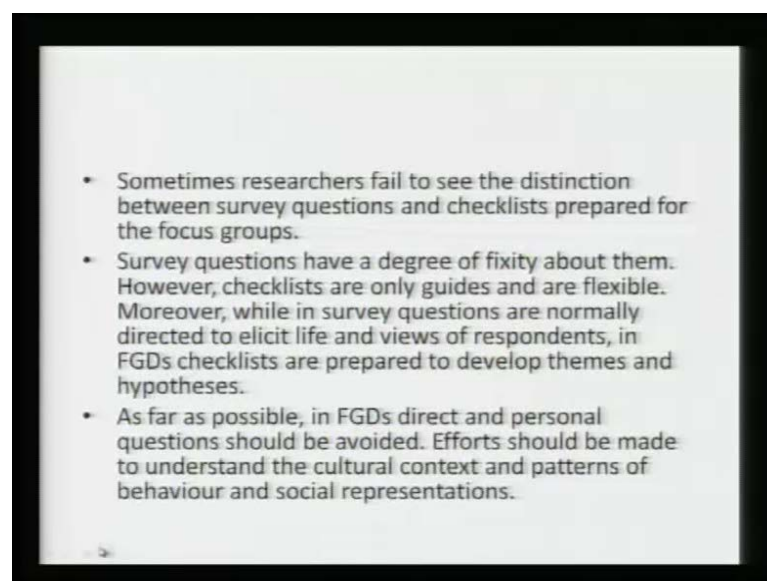
So, we form separate group we conducted focus group discussions separately among men and women and men and women were divided into two groups unmarried and married. It may help if two researchers conduct FGD. One of them would ask questions and the other person would take notes, here must be taken that no single person in the group dominate the discussion which is usually the tendency when you talk to people then and you have group of six to eight or ten people then not everybody talks.

So, if only one person or two person are talking then it is not focus group discussion. You are only reporting the views and attitudes express by this one or two persons. You must involve everybody because FGD differs from group interview or interviews in the sense that if FGD we do not ask direct questions. As we do in group interview in FGD

we will not ask a direct question like what is this or do you know this or what is your educational qualification.

We raise certain issues and the group discusses them we provide an opportunity to the group to discuss our issues here both the outcome of the discussion and the way the group arrives at the outcome are important. Both of both the things are important the outcome of the discussion and the way the groups arrives at the outcome or decision this method is of great help in developing narratives.

(Refer Slide Time: 53:46)

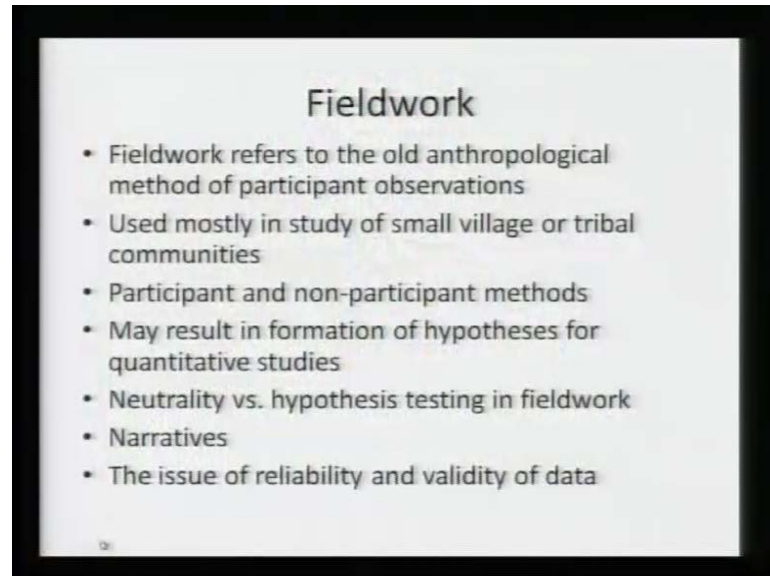


Sometime researchers fail to see the distinction between survey questions and checklist in focus for focus group we prepare checklist. Checklist are more flexible lose they can be modified in the field depending on the needs or depending on what direction the focus group takes survey questions. Survey items, wordings, orders they are fixed in that sense survey questions and checklist are different. For focus group discussion we do not make schedules or questionnaire; we make checklist survey questions have a degree of fixity about them; however, checklists are only guide and flexible therefore, while in survey questions are normally directed to elicit life and views of respondents in FGDS checklists are prepared to develop themes and hypotheses only.

As far as possible in FGDS direct and personal questions should be avoided. They can be embarrassing, unethical, immoral and personal private confidential things should be

avoided FGDS should be used to study culture or communal community's ideas, community's beliefs community's practices rather than individual.

(Refer Slide Time: 55:15)



Field work is the term that is used for old anthropological methods. Field work refers to usually method of participant observations and use mostly in the study of small village or tribal communities. It can be participant non-participant and may result in formation of usually we say that qualitative studies based on field work must provide hypotheses which are later on tested in quantitative studies conducted through surveys or experimental research.

Field work also generates stories or narratives and it gives us ideas or insights. That is why say that qualitative data rich, holistic, insightful. It the complexity of the social phenomenon. Usually the term reliability and the term validity are used for survey data or experimental data, but these days people are also asking questions about quality of data in qualitative research and issues regarding reliability and validity of qualitative methods come.

In qualitative method I would say that while for surveys you have back alpha and spilt half reliability and those things in qualitative research reliability and validity will mean whether you have exactly codified the procedures that you have used in data collection. So, the idea that if your reader not only knows the findings of your research, but also knows exactly what you have done, exactly what your transcripts were, exactly what did

your respondents say, exactly how did you analyze those data. Content analysis manual analysis or using statistical packages then it should be possible to verify them. This degree to which you can communicate your procedures that you use in collecting qualitative data show the degree of validity of your findings. This is all that I wanted to discuss today in the next class then we will talk about various sources of demographic data in the country. Thank You.