

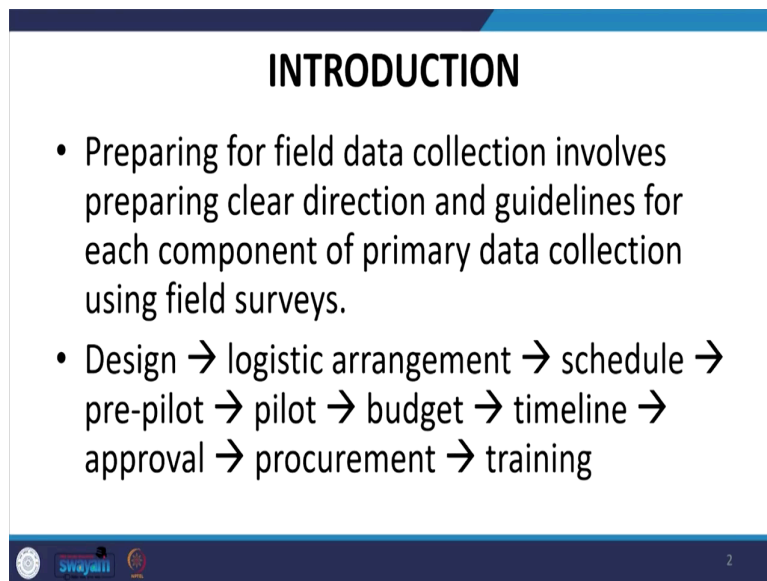
Exploring Survey Data on Health Care
Prof. Pratap C. Mohanty
Department of Humanities and Social Sciences
Indian Institute of Technology, Roorkee

Lecture - 06
Understanding Sampling Base and Design

Welcome friends once again to this NPTEL module on healthcare handling survey data on health care. In this lecture we are going to discuss about, how a surveyor can go to the field and can start for the field survey preparation. To prepare for the field survey on health care the first lecture I have decided to keep it on Understanding Sampling Base and Design.

We know that the sampling base and its design is very essential for research especially in health care research. Since health care is quite important not just its quantitative nature and responses, but also on its qualitative behavior or qualitative contents. So, how to decide our field, how we can design for the survey is important.

(Refer Slide Time: 01:35)



INTRODUCTION

- Preparing for field data collection involves preparing clear direction and guidelines for each component of primary data collection using field surveys.
- Design → logistic arrangement → schedule → pre-pilot → pilot → budget → timeline → approval → procurement → training

swajani 2

Let us move on and understand all those things. Starting with preparing for field data collection, it involves preparing a clear direction and guidelines for each component of the primary data collection from the field using the survey techniques.

So, in the introduction I will guide you that it requires number of stages starting with a clear cut design or the direction or the guidelines and then we will follow with some logistic

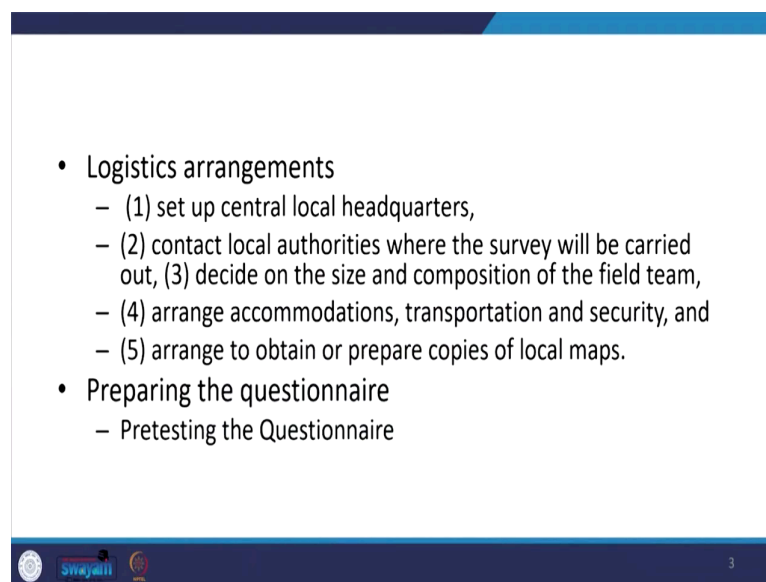
arrangements, how the arrangements are made so that the interviewer can be able to reach at the right place and can start following responses.

Then after the logistic arrangement we should also understand its schedule or the questions or the questionnaire, then there should be a pre-pilot survey and then pilot survey and pilot is a kind of pre-final survey.

Then we also need to understand the budget the extent of expenditure we can be able to bear for the survey and accordingly its timeline is set and due ethical approval should also be taken before the survey. And some sort of procurement and training about procurement of some equipment some instruments which are required for the survey should have been taken and training to the interviewer or enumerator are also required for the field survey.

I will discuss these things categorically, but let us discuss some difficulties or doubts on logistics arrangement it connects with setting of central or your local headquarters where your target is, second one is to contact the local authorities where the survey will be conducted.

(Refer Slide Time: 03:53)



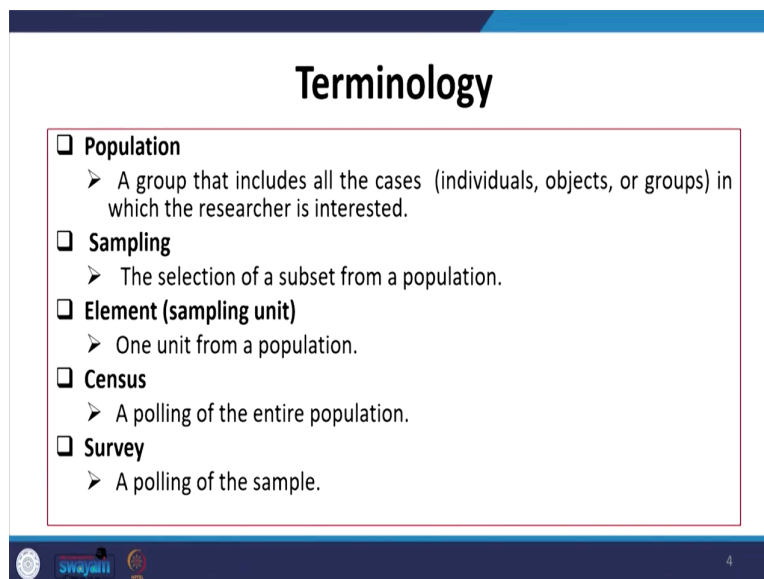
- Logistics arrangements
 - (1) set up central local headquarters,
 - (2) contact local authorities where the survey will be carried out, (3) decide on the size and composition of the field team,
 - (4) arrange accommodations, transportation and security, and
 - (5) arrange to obtain or prepare copies of local maps.
- Preparing the questionnaire
 - Pretesting the Questionnaire

And to decide on the size and composition of the field team and the fourth steps we need to talk about the accommodation where the responsible persons are supposed to stay. Their transportation, their security and also to arrange the copies for their local maps the road the exact location they are supposed to reach.

So, these logistic requirements are everywhere required in all the survey and design. Most importantly it starts with contacting the local resource to that of the map and in finally executing the survey. It involves with preparing the questionnaire, some pretesting of the questionnaire is required. On the next lecture we will discuss about what sort of pretesting are essential, how we can have pretesting of questionnaire for the survey.

I will also show you certain questionnaire or schedule files in the next lecture. In this lecture we are completely focusing on sample base and their design. There are certain terminologies used certain notations are used in every cellular questionnaire or in the sampling techniques. We should also understand those terminologies, one is population and as compared to sampling or element.

(Refer Slide Time: 05:39)



Terminology

- Population**
 - A group that includes all the cases (individuals, objects, or groups) in which the researcher is interested.
- Sampling**
 - The selection of a subset from a population.
- Element (sampling unit)**
 - One unit from a population.
- Census**
 - A polling of the entire population.
- Survey**
 - A polling of the sample.

swayam 4

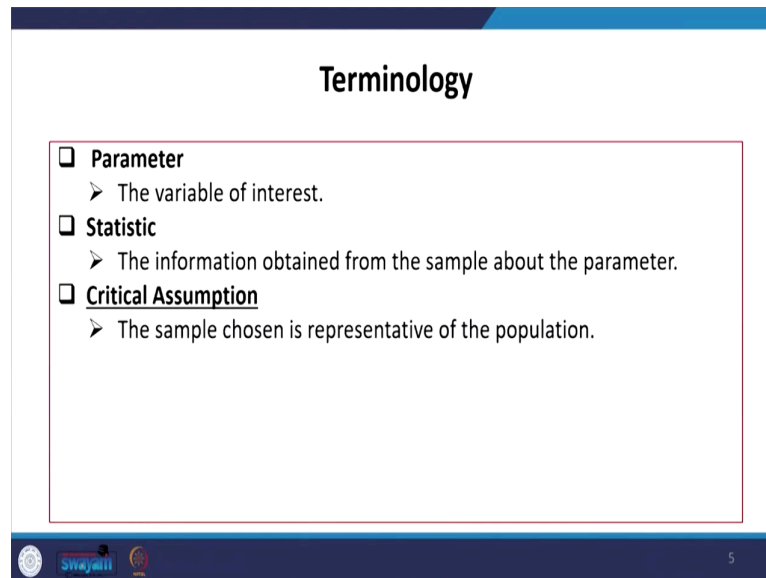
First when we discuss about population we should understand a group that includes all the cases within the study. All the individuals or the objects or the groups in which the researcher is interested. The entire genesis of your study is also required and within that we are supposed to take the sampling.

The sampling involves the selection of subset of that population unit and the population unit within the sampling unit we will take the particular unit from the population that particular unit is called sampling unit or is called element. And similarly we need to differentiate all those things very carefully one is population then sampling and the sampling unit or that is

called element. Try to understand these things through census, what happens in census we usually count the entire population.

So, each unit in the population is the sampling unit or the element. In case of survey we take certain pulling of the sample and all the elements are not included in the survey.

(Refer Slide Time: 07:06)



Terminology

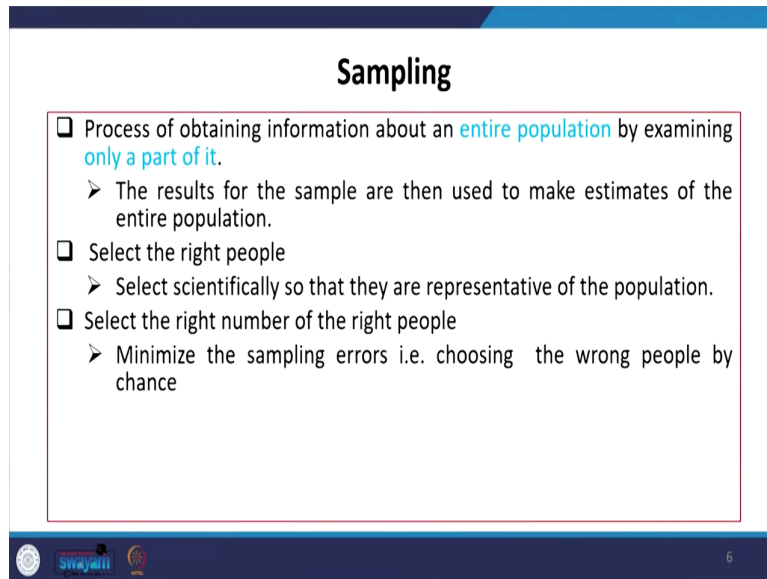
- Parameter**
 - The variable of interest.
- Statistic**
 - The information obtained from the sample about the parameter.
- Critical Assumption**
 - The sample chosen is representative of the population.

swjain 5

Some other terminologies are important like parameter that is the value of interest which is give certain coefficient for interpretation we call it a parameter, the statistics and; that means, the information obtained from the sample about the parameter. Parameter is the statistics of the population unit whereas, statistics is the information or the coefficient derived from the sample unit or from the sample.

So, statistics are the subset of the parameter likewise sampling population. There are some critical assumptions should also be understood like, first of all the most important assumption for sample and population is that your sample chosen should have been representative enough. If it is not representative to the population we are referring to your purpose of the sample or the sampling is not dealt correctly. So, let us discuss in detail about sampling.

(Refer Slide Time: 08:15)



Sampling

- ❑ Process of obtaining information about an **entire population** by examining **only a part of it**.
 - The results for the sample are then used to make estimates of the entire population.
- ❑ Select the right people
 - Select scientifically so that they are representative of the population.
- ❑ Select the right number of the right people
 - Minimize the sampling errors i.e. choosing the wrong people by chance

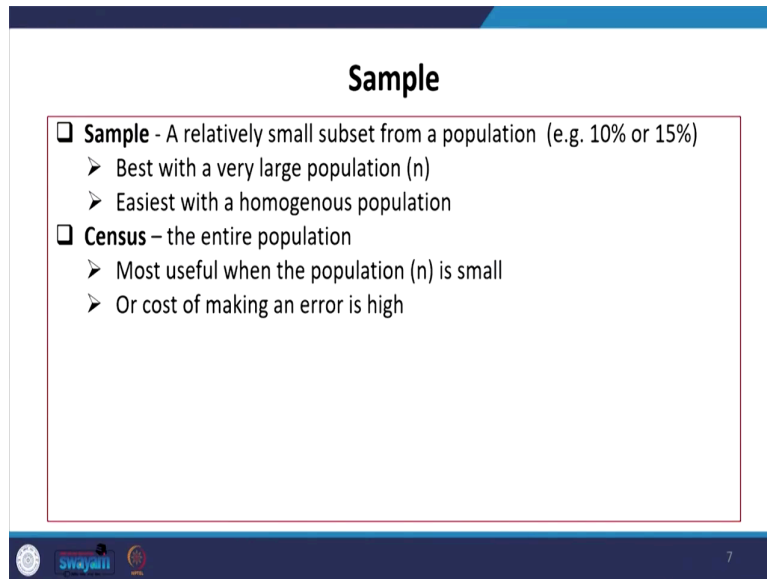
6

Sampling basically, is a process of obtaining information about the entire population by examining only a part of it. The results of the sample are then used to make estimates about the entire population we select the right people, that is, the techniques involved to select right people then select those people scientifically. So that, they can represent the population of our reference.

And then right number of people is also important, it is not just the right people but also the right number size i.e., sample size. So, that our errors should be minimum i.e., we can minimize our sampling error. So, we will discuss that while understanding the right sampling size. So, that our error is within the limit with its confidence limit.

So, now regarding sample once again that is basically the small subset of the population maybe 10 percent, maybe 15 percent depending upon the requirement, depending upon the confidence level confidence of your estimation.

(Refer Slide Time: 09:32)



Sample

- ❑ **Sample** - A relatively small subset from a population (e.g. 10% or 15%)
 - Best with a very large population (n)
 - Easiest with a homogenous population
- ❑ **Census** – the entire population
 - Most useful when the population (n) is small
 - Or cost of making an error is high

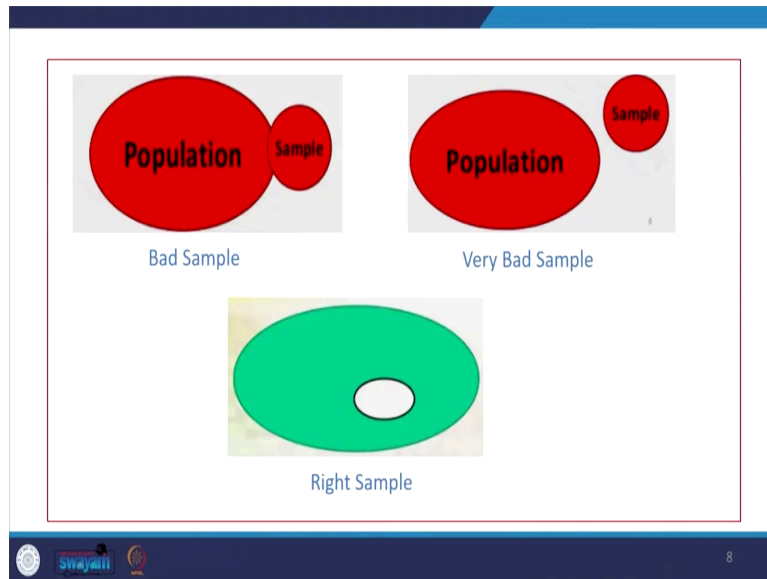
swayamii

When the confidence of estimation is expected to be higher we usually set relatively higher percentage of sample. Whereas, in case of census we are supposed to count the entire population. But in sample, though it is subset, sampling process becomes easiest when your population is homogeneous. When you have homogeneous population then your sample is expected to be better and going to represent your population.

Then usually sampling is going to be better for large population and large and homogeneous population and census as I already mentioned. So, we are supposed to count all the elements and in census we know that is usually referred to the accounting of population even for trees census are there, cattle census are there or domestic animal census are also there.

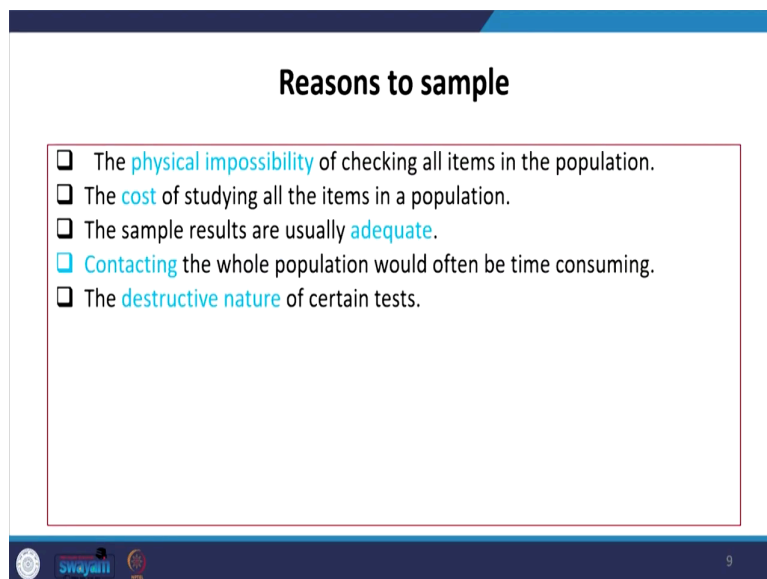
So, the cost of making everyone included in the sample that is called census is very high. So, cost is expensive so sampling techniques are important.

(Refer Slide Time: 10:58)



These 3 pictures show the quality of your sample, first one and the second one is highlighted in red color that shows bad sample because here the sample is not representing a population. On the first picture we say that your sample is only overlapping very less person, you have wrongly taken your sample from outside and the second one it is completely outside population, but in the third one your sample is right.

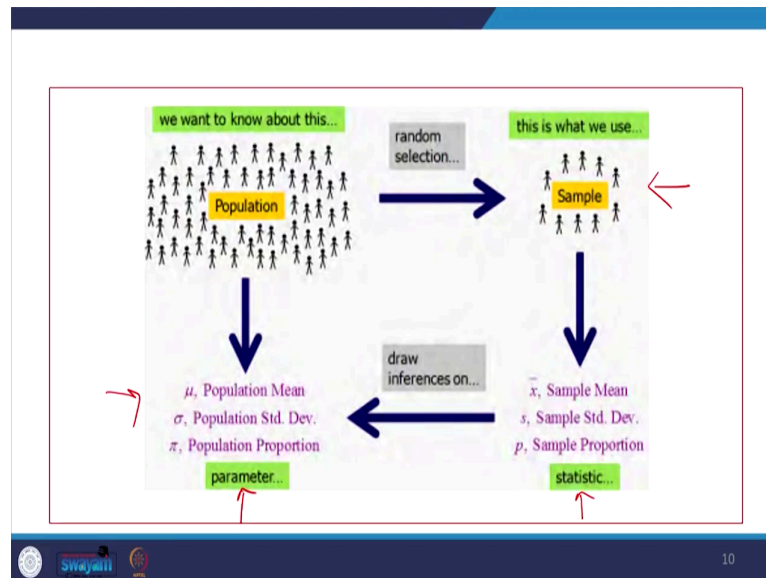
(Refer Slide Time: 11:27)



So, what are the reasons to sample as we already said that the physical impossibility of checking all items in the population and cost becomes very expensive.

So, better sampling is required contracting the whole population would often be time consuming, the destructive nature of certain test are also important while setting the sample. Here again we said that there are different ways of sampling.

(Refer Slide Time: 12:17)



Just for your example at this moment that we have discussed the set of population on the first part, from there we have randomly taken a sample like this. And when we go for estimation about the population those are called parameter, any sort of estimation like its standard deviation and population proportions. Those represents parameter when we estimate anything out of the sample those are called statistics.

And based on the statistic we can draw influences about our population. Like if you wanted to study child nutrition for example, or why am I referring because in the last two days newspaper the world hunger index has been published and this has shown that India's position has been again slide down to 101 ranks from 94 ranks last year. Now it has been again reduced to 101 out of 120 countries.

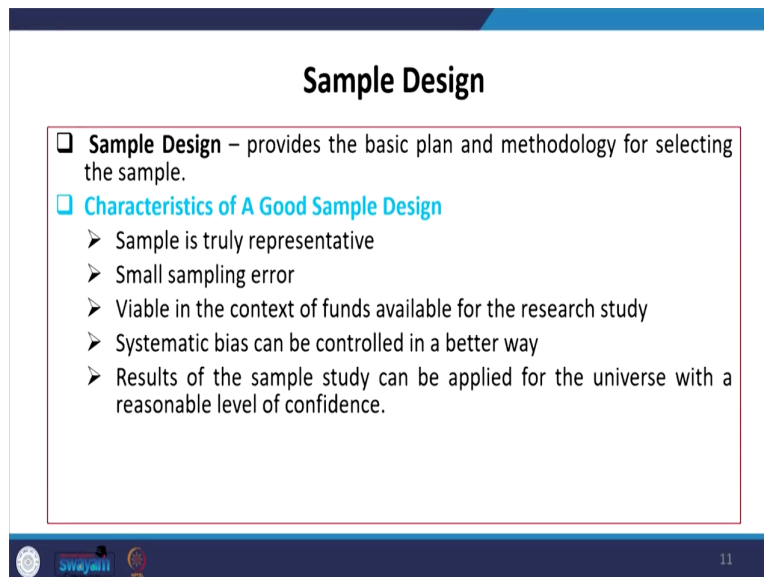
So, now India's child nutrition level is actually at a very low and it includes your nutrition indicators, your BMI indicators, wasting, stunting etc. If you wanted to study entire population and observing all the details is difficult usually we do it in Indian context a national family health survey is used to be the best one for calculation.

But if you wanted to do it in a small scale and for a state specific study then some estimation of the national level is there or in the state indicator already there they are considered to be your parameter. When you have taken sample and based on the sample you have collected sample means, sample standard deviation, sample proportion.

If now you are comparing these indicators with the population parameter and those parameters are usually understood based on literature, based on prior experiences, based on national reports or state reports. So, you can have certain comparison.

And accordingly you may identify whether your estimation is correct or not. Now, coming to sample design which is as per our title of this particular lecture. That you know sample design provides basic plan and methodology to select the sample. This gives a plan or a design including its methodology.

(Refer Slide Time: 15:27)



Sample Design

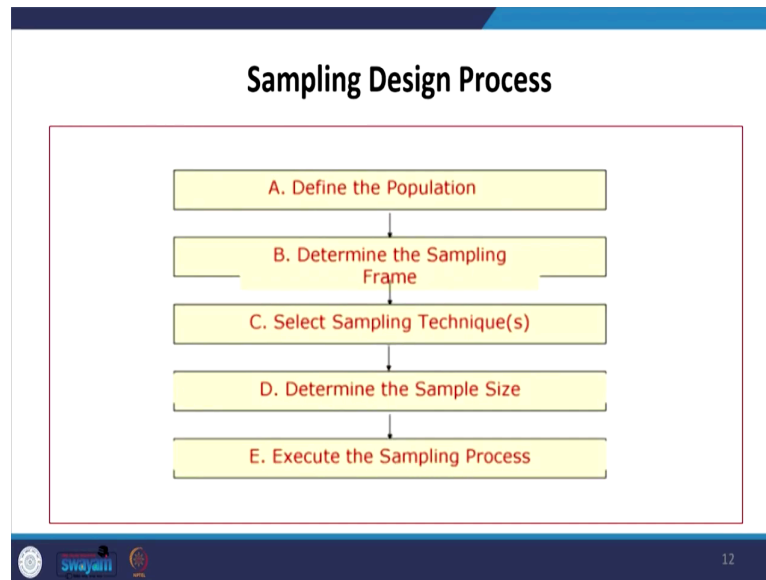
- ❑ **Sample Design** – provides the basic plan and methodology for selecting the sample.
- ❑ **Characteristics of A Good Sample Design**
 - Sample is truly representative
 - Small sampling error
 - Viable in the context of funds available for the research study
 - Systematic bias can be controlled in a better way
 - Results of the sample study can be applied for the universe with a reasonable level of confidence.

swjain 11

So, there are certain characteristics of good sample design those characteristics are like it should have small sampling error, it should be truly representative and should be viable to conduct the study; that means, it should have sufficient budget to carry out the study. And they are any sort of systematic bias that occurs should be properly controlled in a better way through the methodologies.

And the results of the sample study can be applied for the universe the reasonable level of confidence. Some with certain confidence level, your results must be applicable to a larger segment of the population, and then it is expected that you are having a good sample design.

(Refer Slide Time: 16:22)



Sampling design process involves understanding the population through the background or study. So, defining the population, then understanding sampling frame, then selecting the right sampling techniques, then to determine sample size and to execute sampling process. So, all these involves in the sampling design process. So, most importantly sample design then sampling frame, then its techniques of sampling right, sampling should have been taken, understanding the base of a sample and to execute is important.

Then coming to define the target population it addresses the question like, ideally who do you want to survey.

(Refer Slide Time: 17:21)

Define the Target Population

- ❑ It addresses the question “Ideally who do you want to survey ?”
 - i.e. those who have the information sought What are their characteristics. What should be excluded ?
- ❑ The target population can be **finite or infinite**
 - In finite universe the number of items is certain
e.g. population of city, workers in a hospital etc.
 - In infinite universe the number of item is infinite
e.g. listeners of a specific program, throwing of a dice etc.

Example - NSS 75th round on Social Consumption : Health
Target population – Whole of the Indian Union, except for those villages in Andaman and Nicobar Islands which are difficult to access.

swayam 13

And for this answer like we have the persons who have the information sought and what are their characteristics and what should be excluded. When exclusion and inclusion are well understood in your population then your study is well informed and I think your target population is well defined. The target population can be a finite or can be infinite.

So, it might be also finite as well infinite. The universe of the number is certain whereas, in other case like when your study is finite like your population of the city workers in a hospital etc. are usually referred. In infinite universe your number of items in that universe is not finite it is infinite like list notes of a specific program throwing off a dice what it is going to come how many options are going to come it's fine, but what is going to come it's very difficult.


So, there are some probabilities involved, in that case your universe is infinite. Similarly, in the example of 75th round of NSS national sample survey on social consumption that is particularly on health our target population is the whole India including its states and union territories.

(Refer Slide Time: 18:55)

Define the Target Population

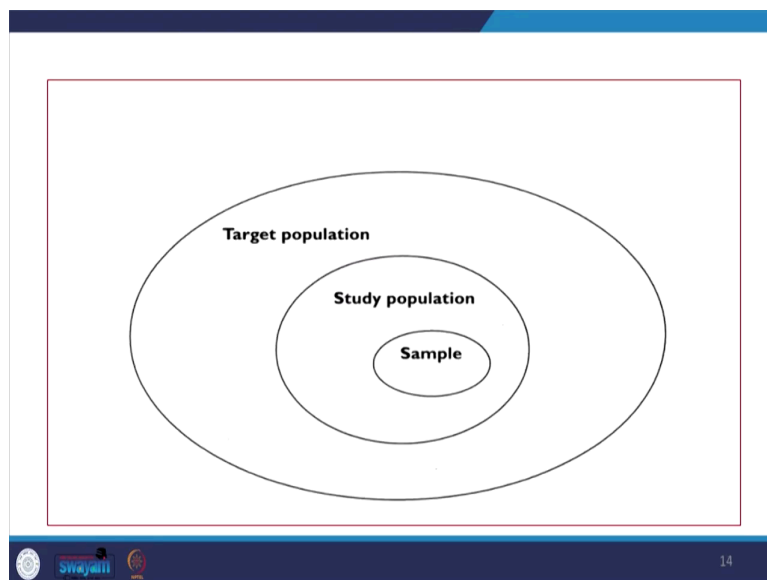
- ❑ It addresses the question “Ideally who do you want to survey ?”
 - i.e. those who have the information sought What are their characteristics. What should be excluded ?
- ❑ The target population can be **finite or infinite**
 - In finite universe the number of items is certain
e.g. population of city, workers in a hospital etc.
 - In infinite universe the number of item is infinite
e.g. listeners of a specific program, throwing of a dice etc.

Example - NSS 75th round on Social Consumption : Health
Target population – Whole of the Indian Union, except for those villages in Andaman and Nicobar Islands which are difficult to access.



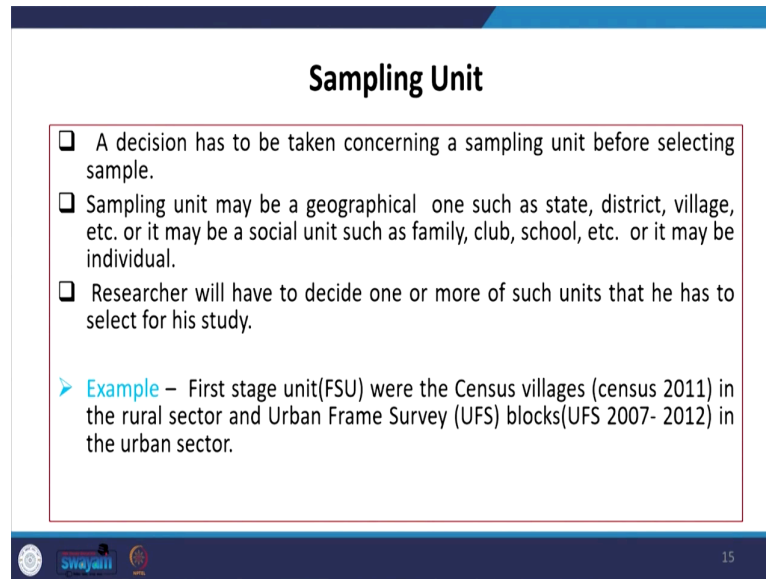
Except for those villages in Andaman and Nicobar island which are difficult to access. So, your target population is excepting these villages all other parts are included.

(Refer Slide Time: 19:08)



So, your sample is the subset of your target or study population and it should be inclusive.

(Refer Slide Time: 19:18)



Sampling Unit

- ❑ A decision has to be taken concerning a sampling unit before selecting sample.
- ❑ Sampling unit may be a geographical one such as state, district, village, etc. or it may be a social unit such as family, club, school, etc. or it may be individual.
- ❑ Researcher will have to decide one or more of such units that he has to select for his study.

➤ **Example** – First stage unit(FSU) were the Census villages (census 2011) in the rural sector and Urban Frame Survey (UFS) blocks(UFS 2007- 2012) in the urban sector.

15

What about sampling unit? Sampling unit is about a decision to be taken concerning a sampling unit before selecting the sample. Sampling unit may be geographically defined such as one state, district, village or it may be a social unit such as family club, school etc. it may be you individual as well.

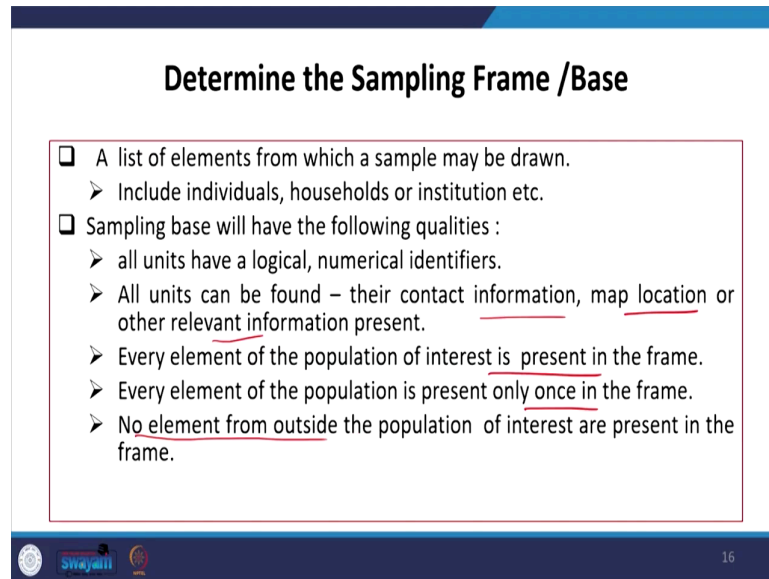
So, your sampling unit is referred to a particular unit. Then the researcher will have to decide one or more such units that he has to be selected for the study i.e., how many units to be taken? For example, in NSS it is clearly defined like FSU first stage unit and the census village where the first stage unit is defined.

In rural area first stage unit (FSU) is your census village whereas, in urban area it is called UFS (urban frame survey). And in NSS that I told you in the last class that our census first stage unit is your census village.

And census village this is as per the reference of census 2011 data whereas, the urban frame sample UFS is from the census enumeration block of UFS of 2007 and 2012 particularly in the urban sector so that is referred. Understanding the sampling frame or base this is basically a list of elements from which a sample may be drawn this includes individuals households or institutions etc.

Sampling base will have the following qualities. Those qualities are like all units have a logical and numerical identities or identifiers.

(Refer Slide Time: 21:34)



Determine the Sampling Frame /Base

- ❑ A list of elements from which a sample may be drawn.
 - Include individuals, households or institution etc.
- ❑ Sampling base will have the following qualities :
 - all units have a logical, numerical identifiers.
 - All units can be found – their contact information, map location or other relevant information present.
 - Every element of the population of interest is present in the frame.
 - Every element of the population is present only once in the frame.
 - No element from outside the population of interest are present in the frame.

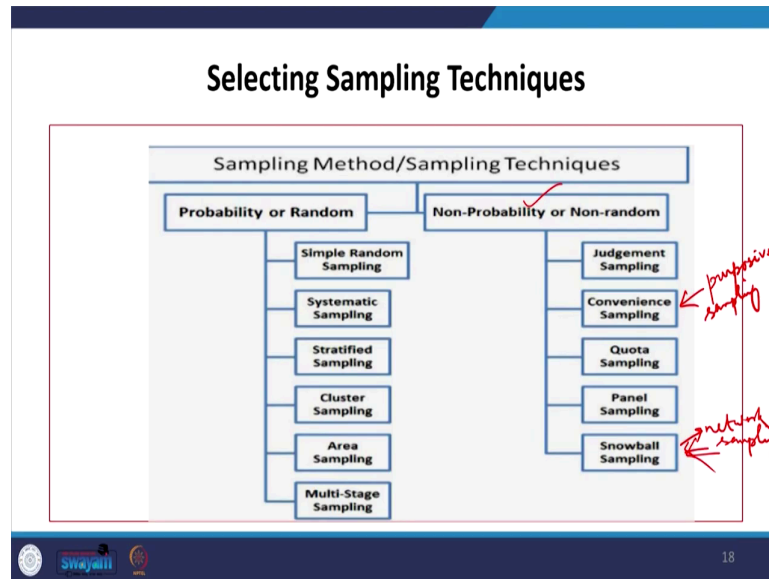
swayam 16

All units can be found their contact information, map location or other relevant information those are present. Every element of the population of interest is present in the frame and every element of the population is present only once in the frame that is also important.

So, these are the characteristics I have mentioned, no element from outside the population of interest are present in the frame outside the population that is going to lead wrong information for the statistics. Like in some examples I have cited here like sampling frame for the rural areas we have already said that there are some census villages of 2011.

Then UFS urban blocks is your urban frame survey and is from the UFS 2007-12 data. So, coming to the sampling techniques sampling base we have discussed, sampling frame we have discussed.

(Refer Slide Time: 22:56)



Then let us talk about sampling techniques. Sampling techniques or methods we have broadly two types: one is called probability sampling and another is called non- probability sampling. Probability sampling is also called random sampling design and non-probability sampling also called non random sampling.

Within the probability sampling we have different categories or the random sampling categories these are simple random sampling, systematic sampling, stratified random sampling, cluster sampling, area sampling and multi stage sampling. So, these three are mostly discussed and besides that we have non probability sampling that involves judgment sampling, then convenience sampling, then quota sampling, panel sampling and snowball sampling.

Nowadays people use snowball sampling. So, snowball sampling especially during the Covid period, it is very essential sometimes, it is also called network sampling. And most of the qualitative studies involve non-probability sampling and they go for their convenience based selection, therefore, they are also simply called purposive sampling.

(Refer Slide Time: 24:44)

- ❑ **Probability Sampling Design**
 - Each element has a known probability of being included in the sample.
- ❑ **Non – probability Sampling Design**
 - Each element in the population is not given equal chance of selection.

19

Then we will discuss all those things one by one, but not in detail we will address these things in the sampling lecture. The probability sampling design for the first one it involves that all have equal probabilities to be selected. Each element has an equal probability of being included in the sample.

In case of non probability sampling each element of the population is not given equal chance of selection. In probability sampling design, first one involves simple random sampling is also called SRS.

(Refer Slide Time: 25:21)

Probability Sampling Designs

- ❑ **Simple Random Sampling (SRS)**
 - Each and every element in the population is given equal chance of selection.
- ❑ **Systematic Sampling**
 - The selection of sample starts by picking some random point in the list and then every 'nth' element is selected until the desired number is secured

Simple random sample

Systematic sample

20

In short is written in different books SRS. In this case each and every element of the population has equal chance of being selected. Whereas, in case of systematic sampling we follow a systematic approach to make them involved in the sampling process the selection of sample starts by picking one random point in the start.

So, starting one is with first random point from the least and then we keep on adding the n^{th} item. So, that the next item or the next point (next unit) is selected. Like if your first one is selected in case of simple random sampling look at there is no ordering of selection like these are simply included randomly.

But in this case if the sampling interval is 3 every time you will be first is this is your first sample. If first is selected the second one is selected by this gap and the third one is selected again with the same gap fourth one is again selected after two person, again fifth one is selected after another two persons.

So, in this case we consider systematic sampling. Why systematic? You know this gives a better representation in the sample, but in case of simple random sampling if your population is very huge if you are randomly picking numbers there might be outliers in the data.

So, those outliers might mislead your interpretation whereas if you have kept all your observation (your population units) in a order and the first set you have taken from the first group as a random start then you keep on adding the interval. So, that from each block you are supposed to have certain representation. So, that is little better and we are going to give better chances of involvement of all the items.

(Refer Slide Time: 27:49)

Stratified Random Sampling

- The population is stratified into a number of non-overlapping subpopulation or strata and sample items are selected from each stratum.

Cluster Sampling

- It involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the same .

The slide contains two diagrams. The left diagram, labeled 'Stratified sample', shows a population of 10 people (5 green, 5 blue) divided into two strata. Red circles highlight a sample of 3 green people and 2 blue people. The right diagram, labeled 'Cluster sample', shows a population of 10 people (5 green, 5 blue) divided into two clusters of 5 people each. Red circles highlight one cluster of 3 green people and 2 blue people.

So, that is why it is called systematic random sampling. Another one is called stratified random sampling in that case we are supposed to divide or the entire population into different strata like in your population is divided into rural and urban. Whether rural or urban, there should be a clear differences. And the strata the population is stratified into number of non overlapping subpopulation or strata.

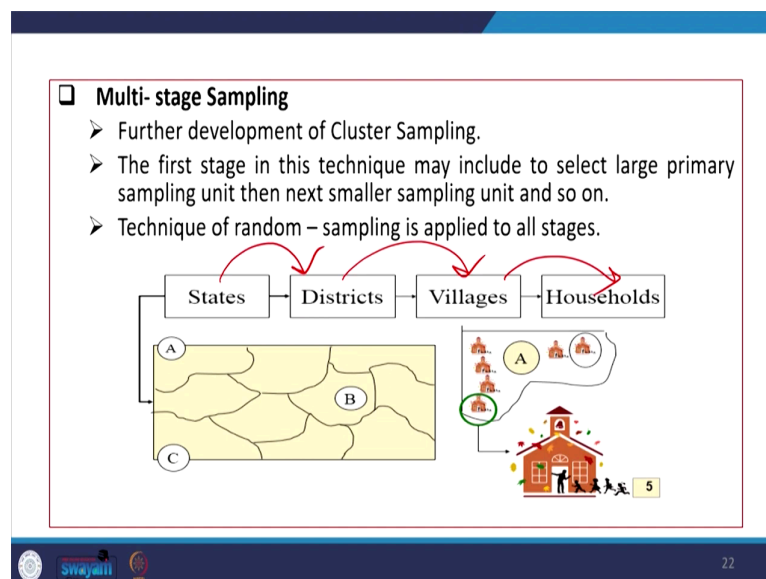
So, most importantly they are non overlapping and the sample items are selected from each of the stratum means your small unit where like rural is your stratum the urban is your another stratum. Then within that maybe rural you have divided into richer family and a poorer family those are also called stratum. Or we divided by religion within the community then they are also defined as stratum. Now about cluster sampling, this involves grouping of population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. So, we include clustering of sample. So, clustering of sample in case of stratified what happens your divisions are not overlapping and not comparable as well. Some features are going to be extracted within that particular category.

So, that category is of course, not comparable with another category that is urban, but in clusters they have each cluster have equal chances of being included in the population each clusters are more or less same. Like blocks are divided in the entire population we have different blocks in different part of our country.

Those blocks are divided in the principles of cluster sample, but the rural urban areas are defined not based on cluster they are different and they are not comparable. So, look at the features stratified sample. So, this is clearly different than previous one. There may be large numbers representing maybe rural of this one maybe urban area, but here almost same numbers are present not necessarily exactly but the features are more or less identical.

Now, to represent our population through the sample out of all those cluster if two are included that is going to represent the population. So, this is another way of selecting samples.

(Refer Slide Time: 30:53)



Now, multi-stage sampling after when we include all possible methods they are called multi stage sampling. In different stages we have different sampling steps involved. Like in multi-stage sampling it is more preferred we have some sampling design. This is developed form of cluster sampling the first stage in this technique may include to select large primary sampling unit then next smaller unit and so on.

For starting stage is your larger primary sample then it follows with sequential smaller units. Techniques of this sample are usually of random. Like in India there are different sampling techniques adopted by NSS, NFHS. So, starting with the large sampling base is with states, then we go to districts, then villages, then households. Then accordingly different techniques are adopted and that is why this is called multi-stage sampling.

(Refer Slide Time: 32:01)

Non-probability Sampling Designs

- ❑ **Convenience Sampling**
 - Elements for the sample are selected for the convenience of the researcher.
- ❑ **Judgement Sampling**
 - Element selected for the sample are chosen by the judgement of the researcher.

The diagram consists of two parts. On the left, a central grey figure is connected by lines to a group of 10 smaller human icons (5 blue, 5 green). A red arrow points down to the central figure, and a yellow box below is labeled 'CONVENIENCE'. On the right, a similar central grey figure is connected to a group of 10 smaller human icons. A red arrow points to one of the icons, and a yellow box below is labeled 'PURPOSIVE / JUDGMENTAL'. The slide footer includes logos for Swajathi and a page number '23'.

Another sampling method as against the probability sampling is called non-probability sampling design. We said that it involves convenience sampling, judgment sampling, snowball sampling etc. In case of convenience sampling people select based on their availability based on their convenience.

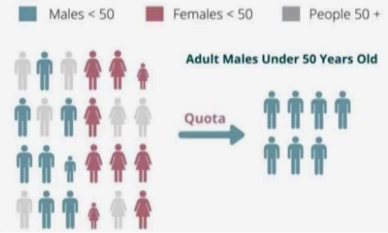
Judgment sampling involves the sample based on the judgment of the researcher, usually the researcher is expected to be more learned more experienced. In case of convenience the person selected here either surveyor or the enumerator selected as responses who are nearby based on his convenience.

Whereas in case of judgment, it is purposive. The person selected as per the requirements better representations are taken. Quota sampling: quota means certain quota certain reservations should be taken to include some of those units in the final sample.

(Refer Slide Time: 33:17)

Quota Sampling

- It appears to be similar to stratified random sampling.
- Instead of randomly sampling each stratum, the researcher uses a nonrandom sampling method to gather data from one stratum until the desired quota of samples is filled.

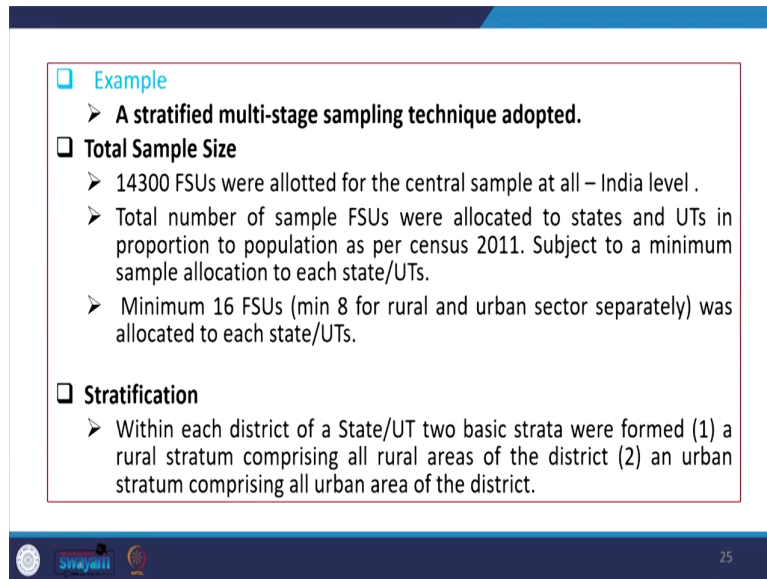


The diagram illustrates the process of quota sampling. It shows a population divided into three strata: Males < 50 (blue), Females < 50 (red), and People 50+ (grey). The population is represented by a grid of icons. A quota of 6 'Adult Males Under 50 Years Old' is indicated by a blue arrow pointing to a group of 6 blue icons. The text 'Quota' is written above the arrow.

It appears to be similar to stratified random sampling instead of randomly selecting each stratum the researcher uses a non-random sampling method. In this case, to gather data from one stratum until the desired quota samples is fulfilled. Like for example, male with less than 50 years of age, female with less than 50 years of age and people 50 years plus we have divided based on our requirement.

Now, in the quota adult males under 50 years of old should have been included for sure. So, the quota is that some numbers of adult females should have been there. So, it is taken those one by its quota method.

(Refer Slide Time: 34:04)



Example

- **A stratified multi-stage sampling technique adopted.**

Total Sample Size

- 14300 FSUs were allotted for the central sample at all – India level .
- Total number of sample FSUs were allocated to states and UTs in proportion to population as per census 2011. Subject to a minimum sample allocation to each state/UTs.
- Minimum 16 FSUs (min 8 for rural and urban sector separately) was allocated to each state/UTs.

Stratification

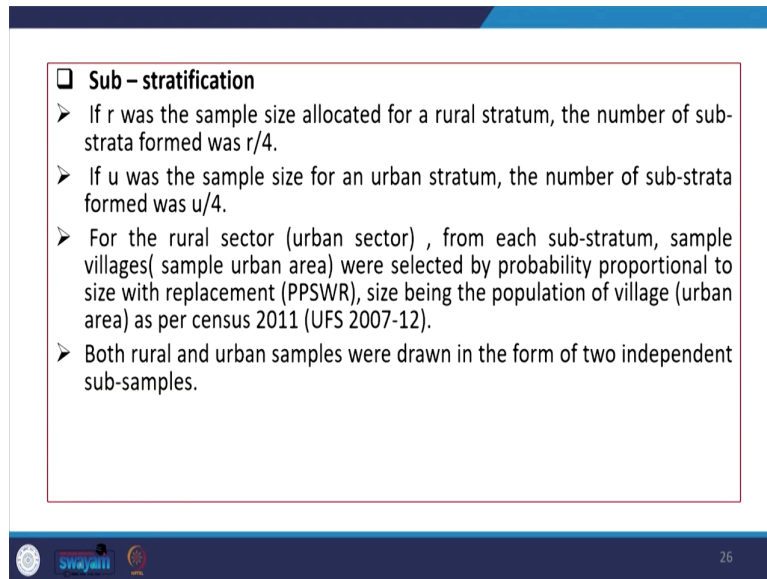
- Within each district of a State/UT two basic strata were formed (1) a rural stratum comprising all rural areas of the district (2) an urban stratum comprising all urban area of the district.

swayam 25

Then second for example, a stratified multistage sampling technique is adopted in most of the cases. Like in your sample design of NSS we have cited those FSU's how many they have adopted the stratified multi stage sampling technique in the NSS survey. They have been included 14300 FSUs from the urban areas and these FSUs are allocated to the central sample at all India level. Total number of sample FSUs were allocated to states and UT's in proportion to population that census 2011.

Minimum 16 FSU's for rural and minimum 8 for urban sector separately was allocated to each state in that survey. The stratification followed within each districts of the state or UTs is two basic strata were formed a rural stratum comprising of rural areas the district and urban comprising of its urban places of that district.

(Refer Slide Time: 35:18)



The slide contains a list of points under the heading 'Sub - stratification'. The points describe the sampling process for rural and urban strata, including the number of sub-strata formed and the selection method for sample villages.

- ❑ **Sub - stratification**
- If r was the sample size allocated for a rural stratum, the number of sub-strata formed was $r/4$.
- If u was the sample size for an urban stratum, the number of sub-strata formed was $u/4$.
- For the rural sector (urban sector) , from each sub-stratum, sample villages(sample urban area) were selected by probability proportional to size with replacement (PPSWR), size being the population of village (urban area) as per census 2011 (UFS 2007-12).
- Both rural and urban samples were drawn in the form of two independent sub-samples.

At the bottom of the slide, there are logos for 'swayamii' and '26'.

There are sub stratification made as well in the round if r was the sample size allocated for a rural stratum the number of sub strata informed is r by 4 because they have taken in 4 stratification so sub stratification is made. If u was the sample size for an urban stratum the number of sub strata accordingly is $u/4$.

For the rural sector from each substratum sample villages were selected by probability proportional to size with replacement and size being the population of village as per the census 2011. So, they followed the probability proportionate to size sampling technique we will also discuss that. In our module and size being in this case they have followed the UFS 2007 or the census 2011 for rural areas.

Both rural and urban samples were drawn in the form of two independent sub samples. And these are all details we have so far discussed and we have also mentioned about sample design, sample frame in the next lecture we will talk about how to go for the field, how to have a better questionnaire and schedule. I think I should close here and look forward your participation in next class.

Thank you.