Good morning everyone welcome to the class of marketing research analysis in the last class we had discussed about case study and descriptive research we could not cover up the research error part which will be doing it in today session okay so today apart from the research errors which is always there in any research process as I had said in the last class we will also discuss about the data where I mean a secondary data that is very important in any research study first a fall let us understand what is a data.

They were there are subjects like data analysis, data analytics you know everywhere around and what is actually mean what is a data, so a data is something like you know you can understand a data is a piece of information.

(Refer Slide Time: 01:14)
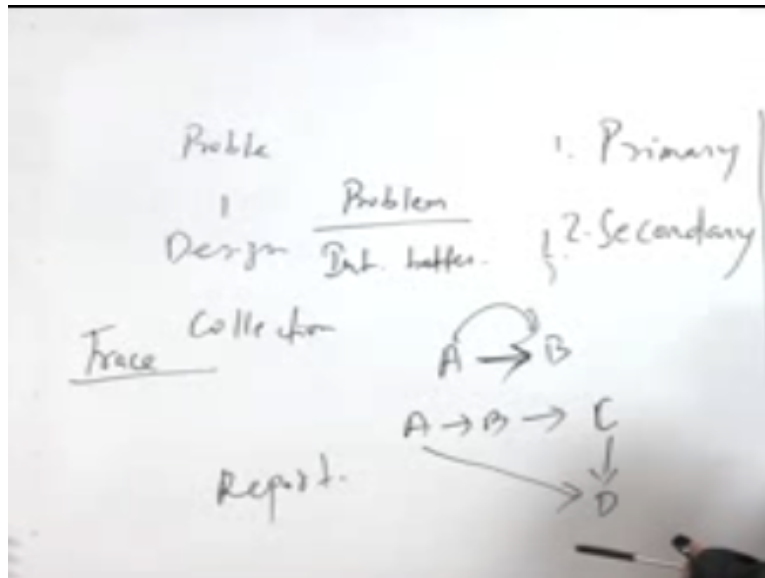
A piece or bit of information right bit of information so a these information's bits actually do not make a sense unless you can really bring it together and makes some real sense out of it right so how does one for collect data and or any even any researcher collecting data that is of very importance is the large importance you know in any study, so let us see what exactly it means so where does data comes from.

(Refer Slide Time: 01:49)
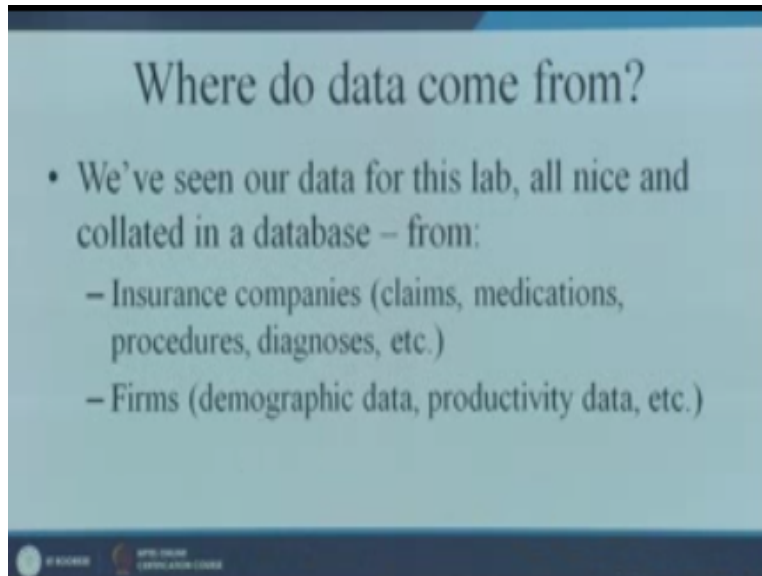
As it says data comes from everywhere in fact of you keep your eyes and ears open data is everywhere around you it comes from let say from fro example in the insurance companies claims you know medications the procedures diagnose is firms with demographic data a productivity data or let say taken a senses where you have the data of people of various religions different gender and their income group all these data right, so data is everywhere you are all flowed with data whether question is what kind of data do we require is of one prime question to any researcher.

So a primary data so two types of data basically that we discuss one is primary so and the second is called secondary, now both have their own importance right a primary and secondary so primary as the name suggests it is primary that means primary means basic so basic data so to understand these are originated by the researcher for the specific.

(Refer Slide Time: 02:59)
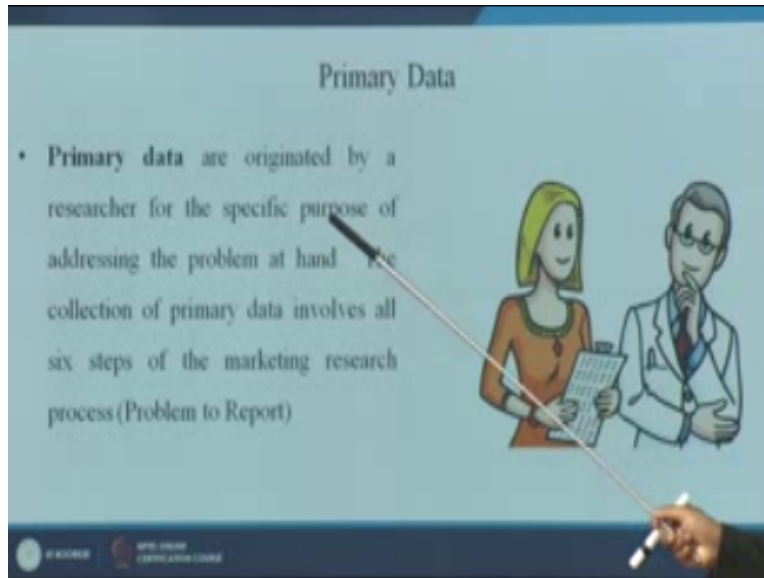
Purpose of addressing the problem in hand now suppose you have a problem as we you know our discussed earlier in the class every problem is different one might think keep this problem is very similar to another problem there might be similar but it is not same right so whenever the researcher is designing or working on a topic he has to understand that if he has a particular or a typical problem in hand then he is data requirement has to be separately or done in different manner or has to be collected separately okay.

So it dependents upon the specific purpose of the you know the purpose of the research right if you remember in one of the classes also I has said, without their problem definition or without addressing the problem properly one cannot a researcher cannot every conducted this is properly right, so the let us understand the collection of primary data involves all the six steps of the marketing research process.

Now six steps are we talking about the six steps that we are talking about are basically are basically of which starts with the you know problem right and it in between you have the design right then collection all this steps together till the report, report writing analysis and the finally the report okay, so let us take a this.

(Refer Slide Time: 04:30)

A Classification of Primary Data Source
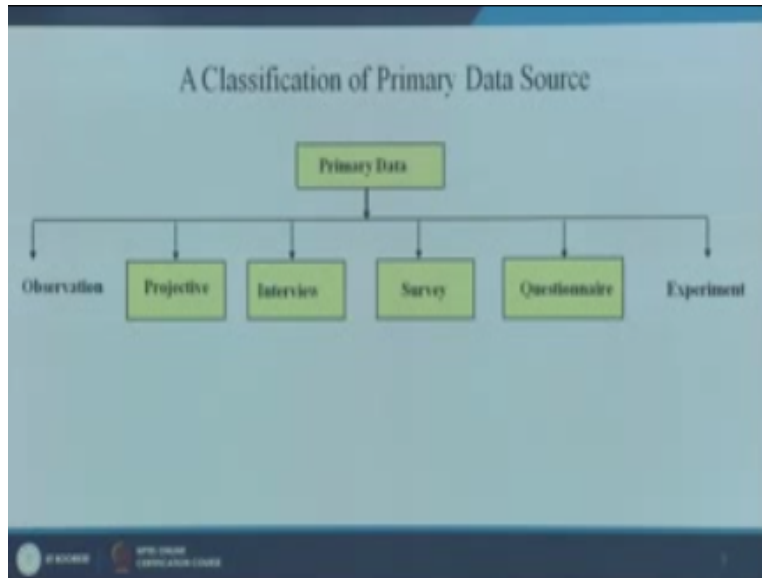
So the classification of primary source right so any data that is on you can understand another way also any data which is directly connected from field that means the area are the scope of research right or from the directly from the respondent directly it is the primary research, so what are these basically a observation so you observe the participant or respondent as you did in your exploratory research earlier projective techniques so you projected you understand somebody by asking like a third person right.

And understanding what he is trying to say about somebody else so that projects his own thinking process to for somebody else but that is one way of understanding how the respondent is thinking then is interview method we said in a depth interview in a special case where the you know the a interview is taken of a expert in a particular area so that we gain in knowledge about some specialized areas maybe of medical of defense.

It could be politics anything right survey is a very popular method survey in fact survey also sometimes you know the questionnaire goes into the survey so survey is very popular in the sense you do a you know try to collect data from a large pool of respondents, and through up maybe through a questionnaire and finally experiments in the labs for example or the experiment that you do on any particular field maybe it is or even sports right okay how does one particular medicine work in athletic.

Could be a kind of an experiment okay so let me start with this case this is very popular case in fact although it is many people might not be knowing it is origin religious is a chewing gum if you heared about it.

(Refer Slide Time: 06:31)
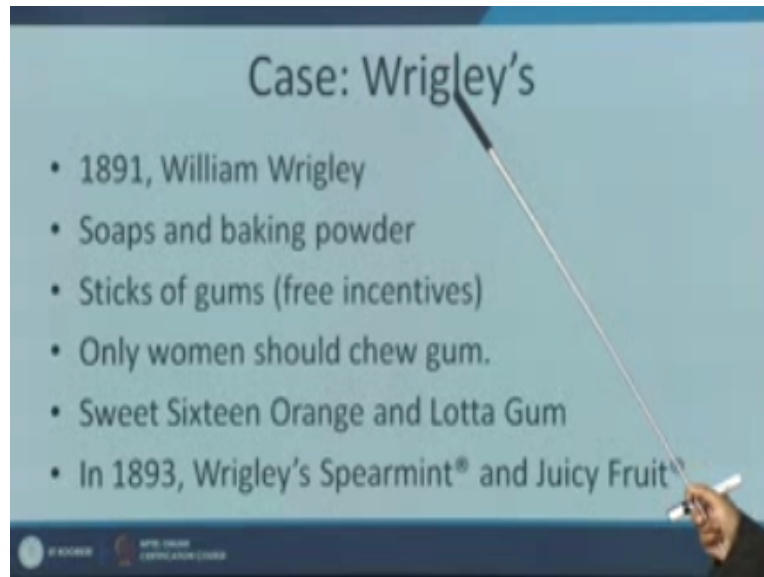
Wrigley's chewing gum it was in 1891 that William Wrigley started this you know what was it doing that time he was selling soaps and baking powder, and he was giving gum sticks of gums you know chewing gums like as a free incentives at that time the stereo type was that women would only use chewing gum right, but Wrigley William Wrigley observed very interesting trend he saw that his products that is soaps and baking powder where actually getting sold now because they were good.

But rather they were getting sold because of this free incentive that is the gums so he thought maybe this is product which is become more popular than the main product right, so what into he came out with two different brands one is the sweet 16 orange and the Lotta gum right so he observed it trend an observing the trend is thought there is a potential product you know in the which could be given to the consumers in 1893 Wrigley's came up with the to most iconic brands.

Wrigley's spearmint and juicy fruit right and everybody knows it became history of that right is a one of the classic cases of in marketing that we teach okay how a product became such a popular product and if a roper research would have been done through the primary data then maybe they

would never understood the potential of this product okay, now what are the advantages primary data as I said right it is on the filed the data collected from the field so the advantage as it is saying obviously you know.

The problem so you have define the problem so when you define the problem automatically you are interpretation, interpretation becomes better okay so you know the problem so you collect the data according to your objective so when you collected the data objective automatically our interpretation would be much superior right he shows of importance or the issues that are more specific to the case of research can be addressed in a better way right, third thing is efficient spending.

(Refer Slide Time: 08:57)
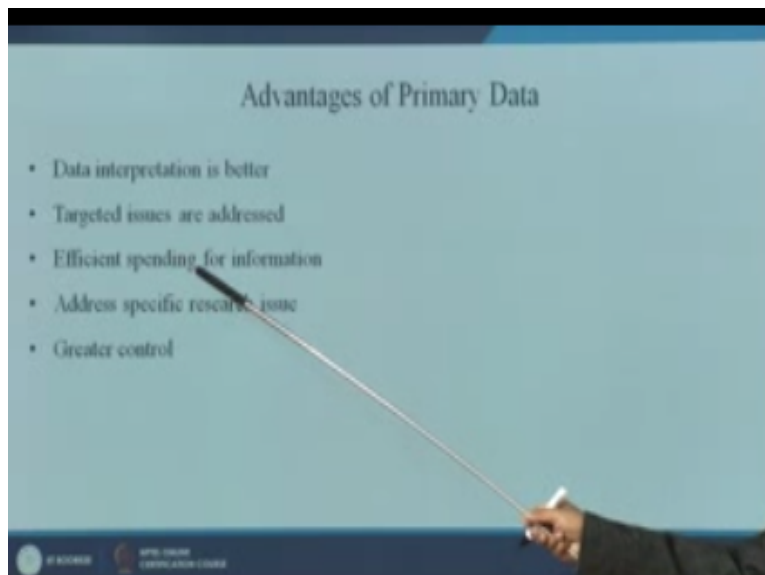
For information that means time and resort and money are two constraints for any research right, so these two concern these two resources have to be utilized carefully one cannot say that I am conducting a research for 6 years 7 years 10 years yeah that is possible in some pure research basic research but in applied research we always have short term you know time durations so researcher has to understand that time and money both are a constraint and has to spend wisely okay.

And even there is a problem if you take too much of a time maybe in applied research there is a problem of the trend completely changing, so maybe by the time your research is over already the trend has changed and the product or the idea which you had might not be any a longer valid

okay and addressing the specific research issue and you have a greater control why you have a greater control because you are you have designed the instrument you have you have understood the sample so you know whom to collect from.

And what to ask everything right so you have a better control similarly secondary data on the other hand is nothing but.

(Refer Slide Time: 10:06)
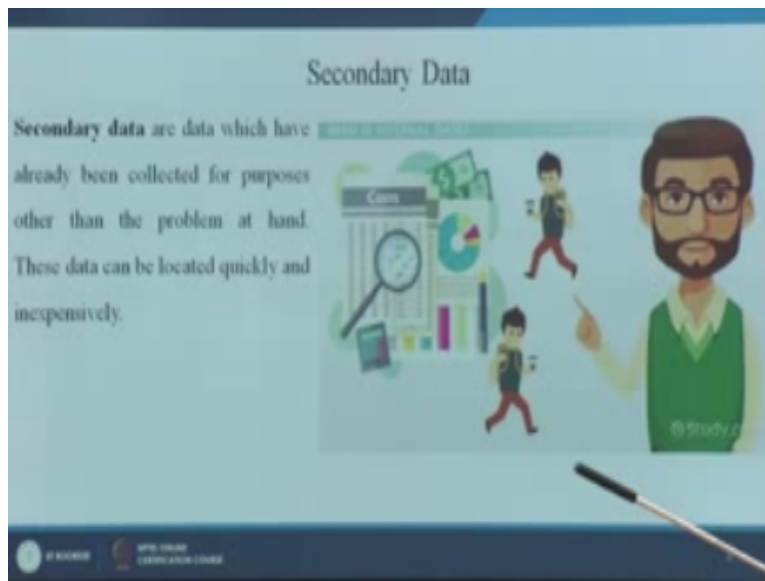
A primary data which has been done already earlier okay so secondary data which have already been collected for the purposes other than the problem at hand so as I said every research is different, so your problem is very specific or you know important to you but it might not be exact the same case but somebody else has done a similar situation or a under similar situation similarly looking study earlier right, so these data could be useful for you for your research in the future okay obviously this is quick because the data has been already collected and it is less extensive right.

(Refer Slide Time: 10:51)
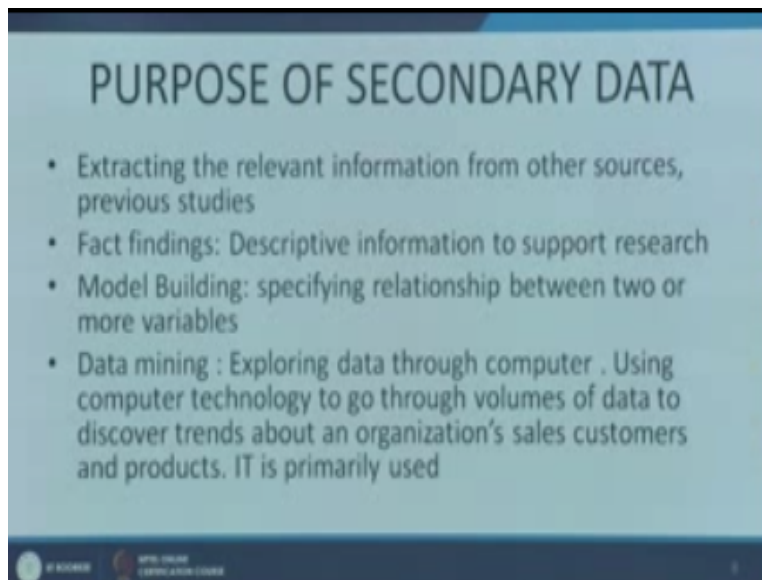
Purpose of the secondary data is to extract the relevant information from other sources like previous studies okay, so previous studies for example let say a company wants to know okay what would be the future of a particular let say brand of chocolates so he can he can look at the past records of other companies who have done some similar kind of experiments in the market had been successful had they have normal successful, if they had been successful or not successful then why it had happen.

So all what was the right channel of distribution what prices had they managed how did they promote that products so everything was in detail right, so fact findings descriptive information to support research is what secondary data helps you in third thing is it also helps you building of models basically models are nothing but specifying relationships now when I am saying for example let say life is never a linear relationship although we always say it is linearity to be checked normality linearity.

But the truth is it is sort the truth is okay it is not linear right you cannot say A affects B all the time in a linear fashion it could be a curl linear also right there might be a so what I am saying is they might not be one single relationship right like A affects B rather it could be A affects B, B affects C and then C affects D right A also affects D, so it could be a multitude of relationships or a multiple relationships which can be affecting the entire study or as I said it might not be linear but rather there might be an angular or it might be curl linear you know relationship so in such

events model building becomes comes to a very handy right finally data minding which is done through the experiment data through the computer.

So you have seen now a days the you know for example social sites for example facebook or twitter or for example all these things even you can see a Samsung data base or let say idem data base whatever you think of the sites right now there are n number of tweets n number of commands coming into it, through these comments through by understanding this comments can be understand the trend of the market for example if you have visited you must have for example any e commerce sight.

When you go into an e-commerce side by you buy something or you make a purchase or even you shown in interest you will see the next time when you go visit the site or you even checking some other site it comes that you have shown interest or you will be shown products which are something very close to you, you do not even realize okay why it is happening but the truth is okay somebody is already you know kind of a tracking you right, so that means I think in observation whenever talking about.

Observation methods one other things was trace analysis which I do not whether I have covered or not or I have missed so trace analysis was something where you try you every respondent use a trace like you know they say after we crime the suspect gives you know a trace right, so this is good enough to understand okay what are his motivations what are his underlying intensions what are his thought processes what does he like what it does not like everything okay, so it helps in discovering the trends right.

(Refer Slide Time: 14:33)

Secondary Data – Examples of Sources

- Vital Statistics – birth, death certificates
- Hospital, clinic, school nurse records
- Private and foundation databases
- City and county governments
- Surveillance data from state government programs
- National statistics - Census, weather etc.

So some of the secondary data I have given example country county governments surveillance data national studies like clean like school right.
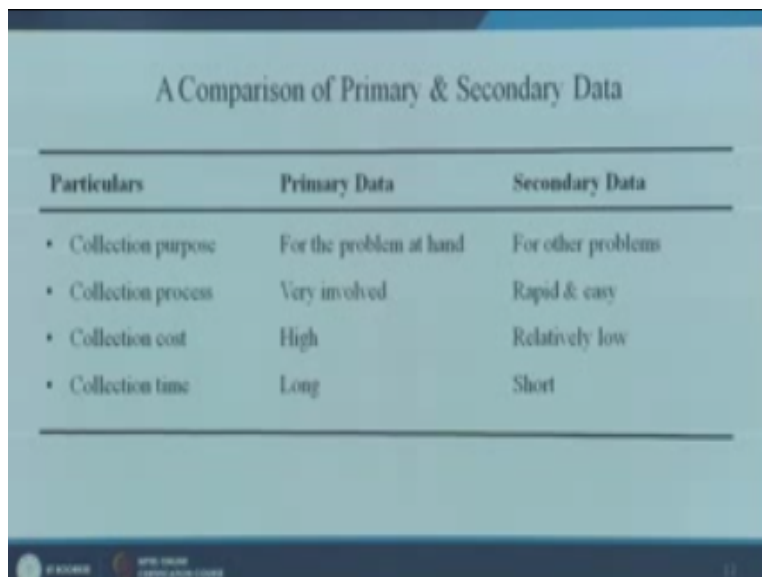
(Refer Slide Time: 14:41)

Case: High Touch Goes High Tech

- In many ageing nations - decline in the number of young (16-24 age) workers
- This potential shortage of young workers has caused many fast food restaurant to switch from a "high touch" to a "high tech" services orientation
- Many services formally rendered by workers are now performed by consumers by using high-tech equipment
- Touch screen kiosks is becoming a popular trend that provides a new avenue to cut labour cost and increase customer service.

Now this is a case very interesting now many other countries you will see that the aging population is becoming a higher and higher right, our Indian been exception where rather we have a large pool of youngsters but many of the countries in the world like Japan, USA and others they are facing a problem of age group higher age group right, so when a higher age group public is going to come it is not that in a day you will be getting ready for them right so it takes time to prepare the market.

To prepare the nation for it so in such a situation there is a potential shortage of young workers which has been seen that it would be a potential shortage of young workers which has been seen that it would be a potential problem, so companies have come up with solutions like understanding okay what would people require in the future, so to understand that fast food restaurants for example have move from a high touch that means a personal service to a high tech service right.

So now what is happening here touch screen is being popular trend that provides a new avenue to cut labor cost and increase in customer service, so when workers would not be aware automatically you have to develop you know not develop you have to dependent more on the machine the technology right maybe that is one reason you must have been seeing around the world this is a huge up search in research in robotics and use of robots as in even daily life okay.

(Refer Slide Time: 16:09)

A Comparison of Primary & Secondary Data

Particulars	Primary Data	Secondary Data
Collection purpose	For the problem at hand	For other problems
Collection process	Very involved	Rapid & easy
Collection cost	High	Relatively low
Collection time	Long	Short

A comparison the primary and the secondary data so as you had understood a the collection is for the problem at hand purpose and this is for other problems right, so this is very involved that means are part of the process it is easy because obviously it is there somewhere and you just have to get it, collection cost is obviously high you have to travel you have to go to the field or you have to pay the respondent this is comparatively low per unit cost right and this is long and this short.

But that does not make the primary data a weak case rather it is a stronger case right secondary data all certain could have some own limitations let see now first fall let start the advantages less times.

(Refer Slide Time: 16:53)
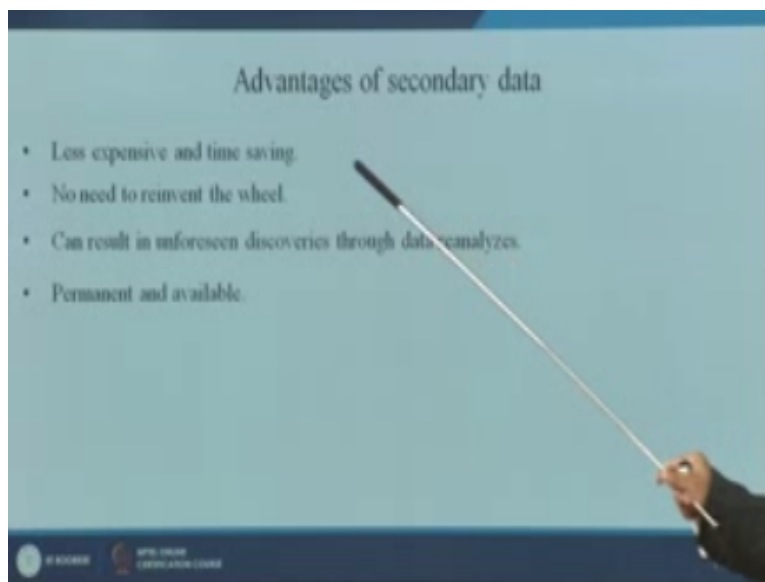
Expensive and time saving no need to reinvent the wheel because already it has been done by somebody so if it is close to your problem at hand or your rest topic so you can follow that you can use it you can result in unfortunate discoveries, now something that can you can realize whether even getting into the market by crunching those data or analyzing the trend for example is there a move in the market the people want to buy customize shoes now yes could be possible now it has been seen.

Day by day with obesity where with problems medical problems people want are not getting the right fit of shoes, now companies have realize okay this could be an new potential area there were they can find new customers even if can see the obtained industry for example the paint industry like Ancient paints and burges they do advertise that you can make your own color you can have your own color right so the given option and they ask you to make your own color and then you can you know you can say what you like so these are some of the advantages.

(Refer Slide Time: 18:04)
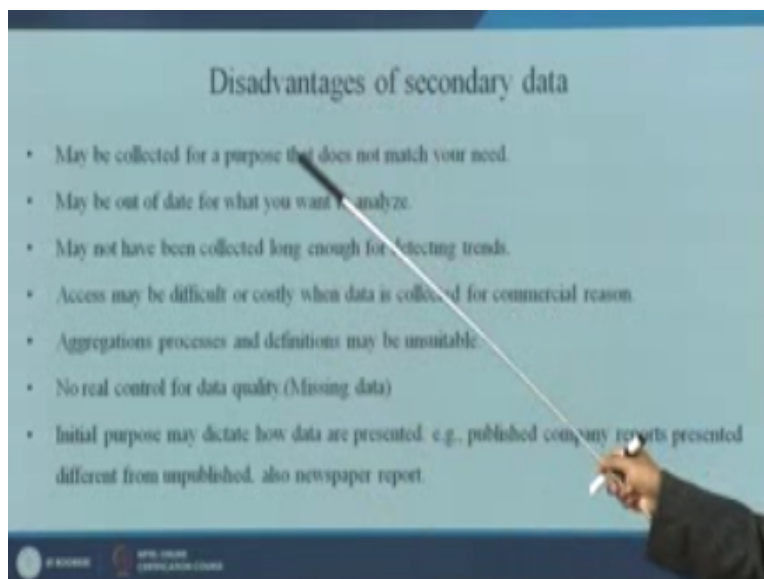
Now what are the disadvantages maybe it has been collected for a purpose that does not match your need this is important and very, very important why it is very important sometimes we do have an habit of adjusting and research is not about adjusting this is about knowing the right thing, so if you are trying to adjust okay my research is very similar to that I can adjust no, no research is not about adjustment please be very careful about it right maybe out of date for you what you want to analyze.

This may be was done in 9 2000 and let say five and in 2016 2017 already the trend has change right may not have been collected long enough for detect in trends, now this trends obviously this is connected to this so the trends have already changed access maybe difficult or costly when

data is collected for a commercial reason, so sometimes but still I would say that per unit cost remains less you know less because it is a large data base.

So that is a inducement might be have so sometimes the aggregation process the way has been aggregated the data might be unsuitable for your need right, this is very important thing again so they find lot of missing data's so a person has already you know made the research and if when you take this data, data base you find that there is a lot of missing data's now what will you do with this so these are some of the problems right, so some of the secondary data let say internal and external.
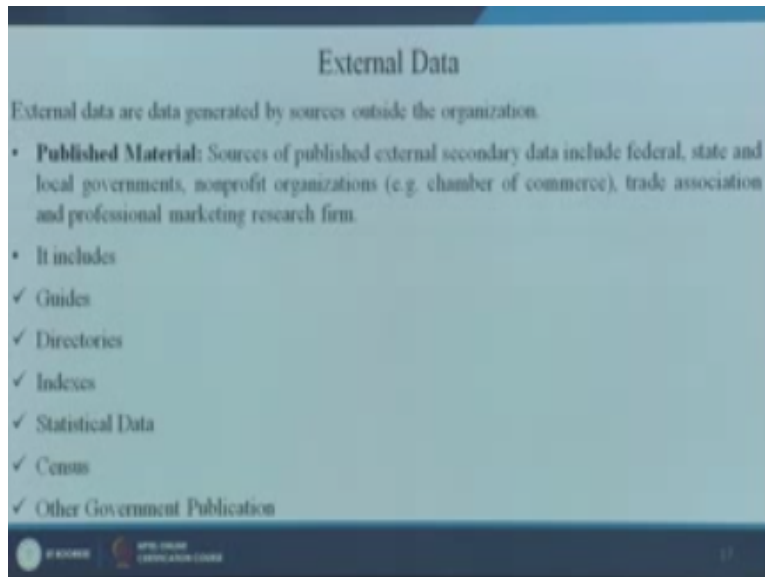
(Refer Slide Time: 19:41)
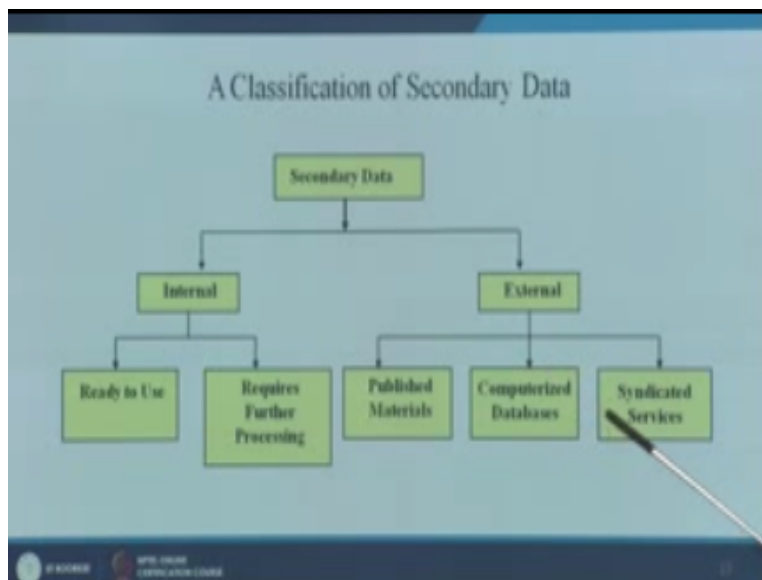
So as you understand the word internal case from inside external is from some external environment okay, now internal that might be you know something like your internally companies records something which is available ready to use and these are something like these required some processing some request for that processing external reports are like the publish materials in generals magazines new anywhere right material even newspapers computerized data which you can take from others.

What are the syndicators services syndicated services are basically those services are those services which are for example a done or conducted by a firm order in organization and they keep this data with them and all your request or according to your requirement they provide this data to you okay, so I just spoken about it.
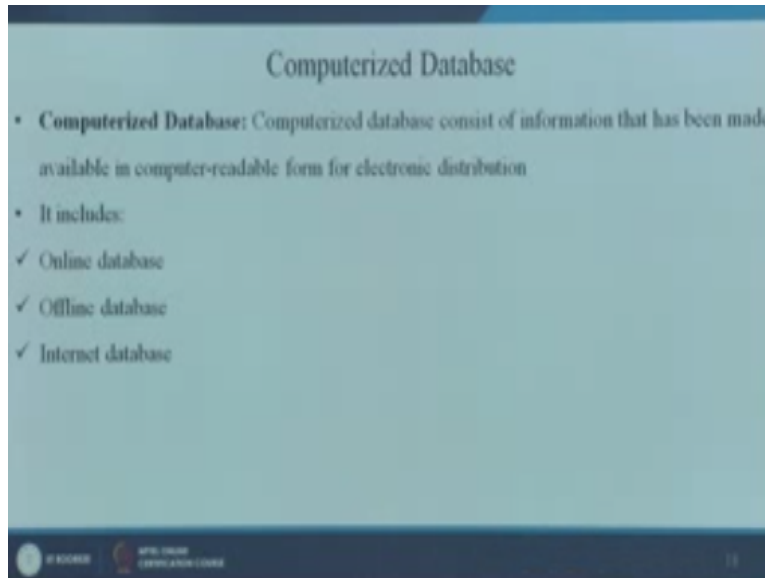
(Refer Slide Time: 20:38)

So I am not speaking much not spending time too much so for example this says the publish materials you can see some other publish materials you can see some of the publish materials directories, indexes, statistical data etc.

(Refer Slide Time: 20:48)

Computerized Database

- **Computerized Database:** Computerized database consist of information that has been made available in computer-readable form for electronic distribution
- It includes:
- ✓ Online database
- ✓ Offline database
- ✓ Internet database

Senses computerized data bases online data bases offline data base on the internet which I said social sides and all which is available now this is becoming a huge, hugely enter the data this companies are you know dying for and you know to understand those data and get is you know idea from it that is why there is a biggest reason why today there is a huge demand of data analysis and you know marketing research analyst to understand okay whether from this data can we understand the new trend.

Oh yes suppose a retail owner let say all mat comes to know that in a particular time of the day only people are coming right or in particular time of the day people are tending there is a tendency to bargain more and in a particular time of the day they are more relaxed so if I come to know this thing then maybe what I can do is I can how a dynamic pricing in the time in their more relaxed and I will have a more fixed time you know a kind of a system when they are more in a less flexible mood right so I said that also so these are the syndicators sources.
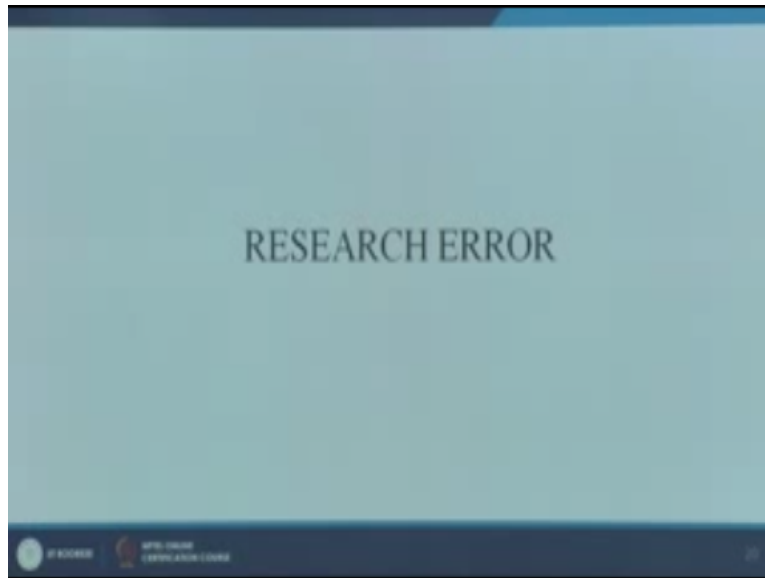
(Refer Slide Time: 21:57)

Common pools for a commercial value so you can buy from somebody write okay so this are already done so already the servant collected by somebody some syndicate some organization and you can buy it from them some price okay.

(Refer Slide Time: 22:13)

Now from here we will go into the research error as I have been saying always the researchers has to understand that research has it is own problems or scope of let say the problems or I can say where could be the errors line okay, where could the error be lying in during any research okay, so if you conduct a if you go with an error then there is a chance that you are entire outcome might be affected that buyers might creep in okay so what are the potentials sources let see this.

(Refer Slide Time: 22:53)

Potential Sources of Error in Research Designs
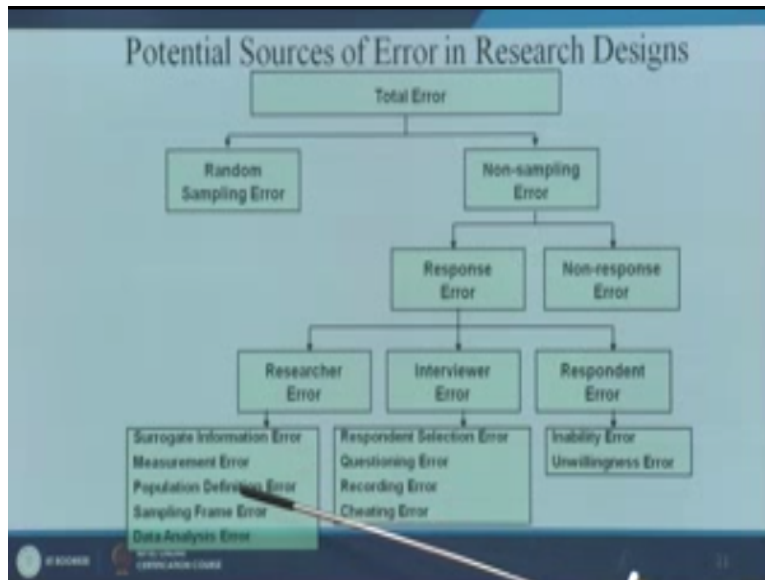
Total error can be broken up into two parts random and non random okay random and no random now let see this random sampling errors and non random sampling error so random sampling error is something like whenever you are correcting the data in a manner in a randomly they correct some selection choice sorry selection problem right, so this is called a buyers created in the process of selecting the respondent right then comes in non random or non sampling not exactly non random.

Sorry you can say non sampling error right so something that is that means in one way of you understand this is something connected to the sample this is not connected to the sampling non sampling okay I hope this is clear non sampling so in the non sampling there are several kinds of again buyers for example there is a respondents and they non respondents buyers non responds okay non responds buyers now let see this researcher error, interviewer error and respondent error now responds error is.

Linked with all the errors that are connected to the way the responds is getting collected okay now it could be from the researcher how as the researcher conducted error let see so there is a information share or get information error right there is a measurement error right there is a measurement error defining the population in a wrong way you have not even define the population what is your population be suppose the researcher is not very clear see it is looks very simple.

But when you will go into any live real case of research you will get into a complicacy of understanding the even the populations sometimes right what is the sampling frame first let us understand what is sampling frame the sampling frame is nothing but it is like a you can say the sampling frame is a frame or a.

(Refer Slide Time: 25:12)
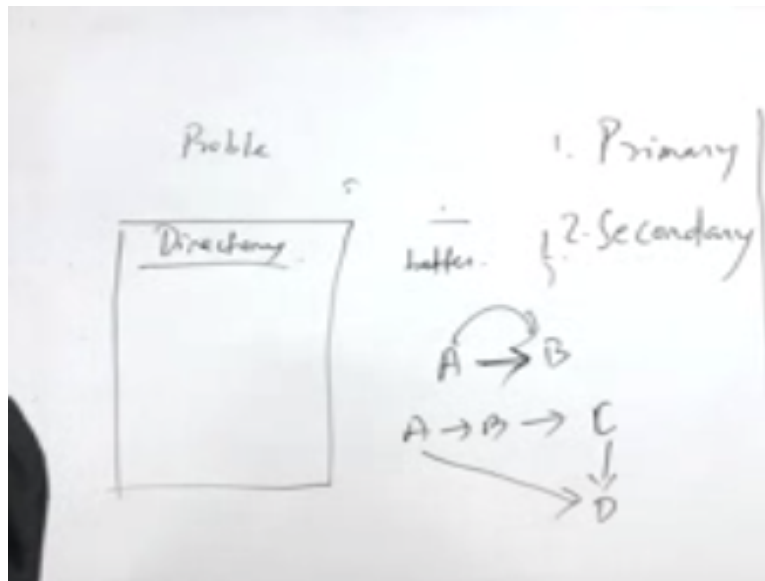
Is a source from where you collect the data right for example to collect the information about people and directory telephone directory could a source right it LOP it could be a source right so that is what is a sampling frame it is like a photo frame only right, finally the data analysis error now you most of the people are actually bothered about the data analysis error now this is only one out of it out of the many so there are other errors also one has to be very clear, so data analysis error for example.

As I had a said earlier what kind of what is a right research tools are statically tool that should be used that is suppose you have instead of doing a you know test, T- test you have instead of doing a an nova and so variance you are doing a T- test right, so instead of and now I am saying you are doing a T- test now that is wrong you should not be doing there is no reason although it can be done multiple T-test can be said as ANOVA it is a case of let say and GOVA right which is analysis of co variants.

Similarly we have suppose multiple independent a MONAVA right or similarly a MOVCOVA now what is your requirement accordingly you have to decide right selecting the respondent

while into taking an interview you may make is selecting a wrong respondent your respondent should have been suppose you are you have to collect the data for in the age group of let say up to 40 now and only let say mail but I am not understanding that clearly you have collected data of mail but in the age group if 55.

Or 60 so that is not fitting your data right is not your requirement so selecting the right respondent is again a critical challenge where you can commit mistakes which can be a fatal error how have you questioned now what you mean by how do you question first a fall is the question a valid question has then question been asked in the right frame right way has the scale that has been offered to the respondent good enough to explain the process how have you recorded the problems.

Now recording now suppose while conducting a research you have only heared it you have listen to it and you are not noted down and after that when you come to your place you have noted something which is exactly not where what the respondent I said so that is a big problem cheating error is something that I do not want to discuss because it is completely although it is a very large problem now a day's researchers in order to publish papers in order to prove you know effective to the bosses.

In a firm they try to give all data or all kind of inferences which are which would be liked right but that is not correct that might not be correct but it can be it may be light so that is a difference right, similarly I have seen scholars even for the data so data fudging is a crime data fudging is a crime so you should not be putting in the wrong data suppose you do not have an understanding of the data either go back to the field or there are this techniques through which you can correct that problem right.

For example we can handle missing data if there is a suppose re missing data or your data is not normal is not gaining the normal manner so you can you can convert into a normal distribution right so finally respondent error, now the respondent might not be in a position to answer you, you have asked about the very scientific question to a person who does not understand that so that is an inability or he is not willing to answer so you have to be very careful as a researcher okay so this is what we are said so the total error.

(Refer Slide Time: 29:04)
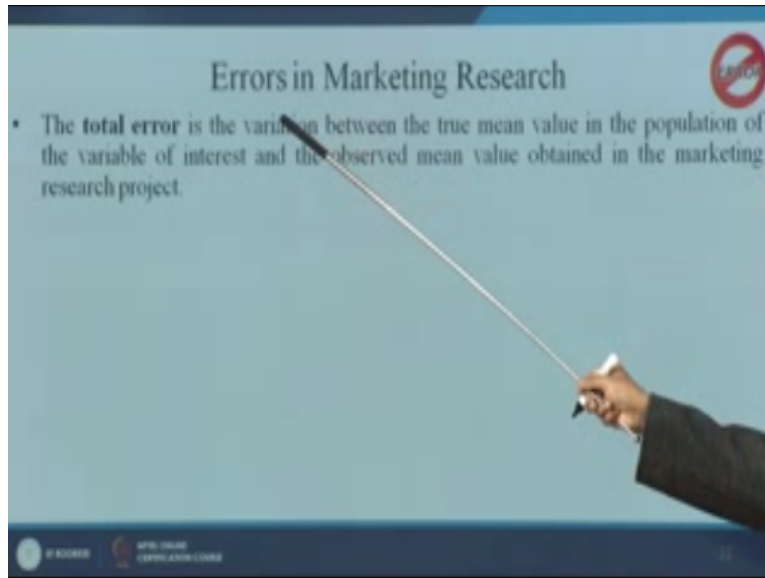
Is a variation between the true mean value of the population and the sample mean that is what I was I have said in the earlier classes also if you remember so the population mean this is mu given by the sample mean x bar so now x bar – mu should actually be tending towards 0 should be tending to 0 that should be approximate I think towards 0 but that means were they are similar but actually does not happen into never happen.
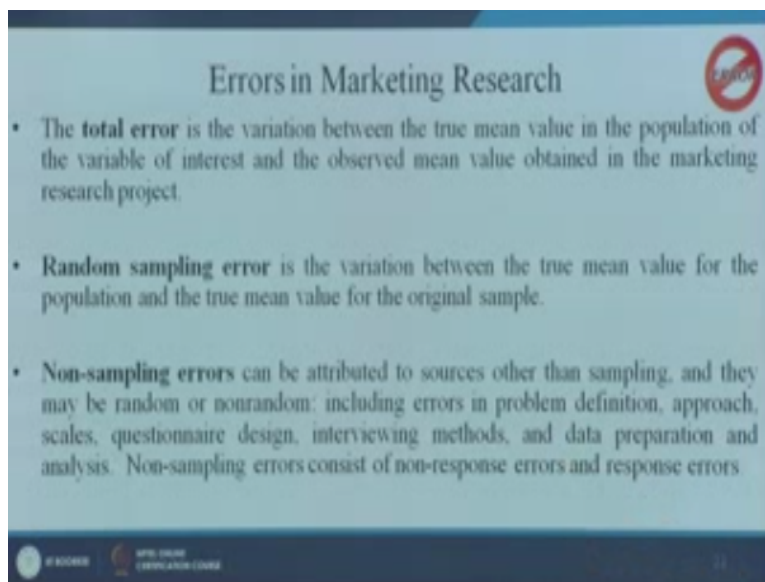
(Refer Slide Time: 29:56)
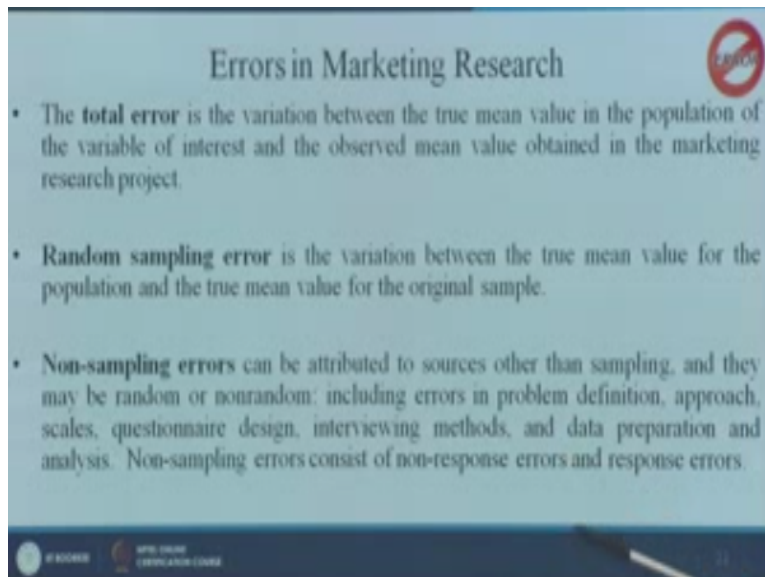
But you will always hope that this should be as closer possible right and of the if there is a difference then thus this difference whatever the difference whatever the difference lies will be

termed as the error so right now random sampling is that we have just said then the this are the other sources of sampling errors that is for example we are defining the problem and all this things right.

And now let see little bit the non respondents error included in the sample do not respond well okay and the response error arises when give you in accuracy answers right that is misrecorded or misanalyzed okay.

(Refer Slide Time: 30:29)
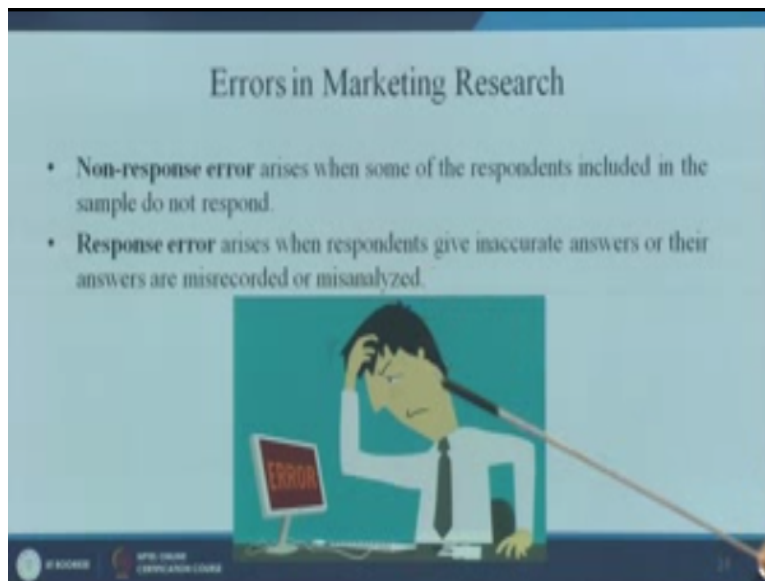
So response error could be surrogative information error as I said the variation between the information needed and the information sort by the researcher so the problem please understand once we had discussed in the class about symptoms and the problems so this is what I am talking about you have tail of elephant and you ar4e saying the such as a trying to say this is not an elephant or the elephant is like a rope no that is not it so you have to understand there is the difference between what is the research problem needs and what have you collected right so some many occurs because many times of the poor understanding of the researcher so that can only be collected through proper literature review.

The more you read the more you discuss with the experts in the area the better will be the clarity okay similarly measurement error how you measured the right the process define the population right at hand also this is the problem right.

As I said the case off male you know the age group higher age group so these are the some of the errors right so by interviewer so respondent selection error occurs when interviewers select respondents other than those specified the sampling design right questioning error which we also discussed how you have hearing interpreting and recording.

Errors in Marketing Research
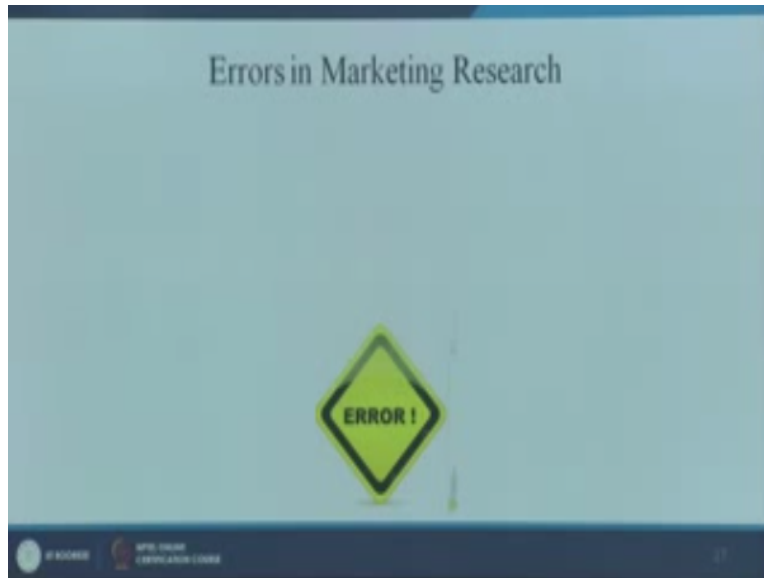
Cheating is what fabricate the answers if you want to cheat nobody can help you it is expected it is universally expected that researcher is honest will be the honest but somebody is not honest then it is a problem okay so this is where I will stop right so inability and unwillingness and this is all thank you for this session will meet you in next session thank you very much.

For Further Details Contact

Coordinator.Educational Technolgy Cell
Indian Institute of Technology Roorkee
Roorkee 247 667
E-Mail Etcellitrke@gmail.com etcell@itr.ernet.in
Website: www.itr.ac.in/centers/ETC. www.nptel.ac.in

Production Team
Sarath Koovery
Mohan Raj. S
Jithin. K
Pankaj saini
Graphics
Binoy. V.P
Camera
Arun. S

Online Editing
Arun.S
Video Editing
Arun.S

NPTEL Cooridinator
Prof B.K Gandhi

An Educational Technology Cell
IIT Roorkee Production