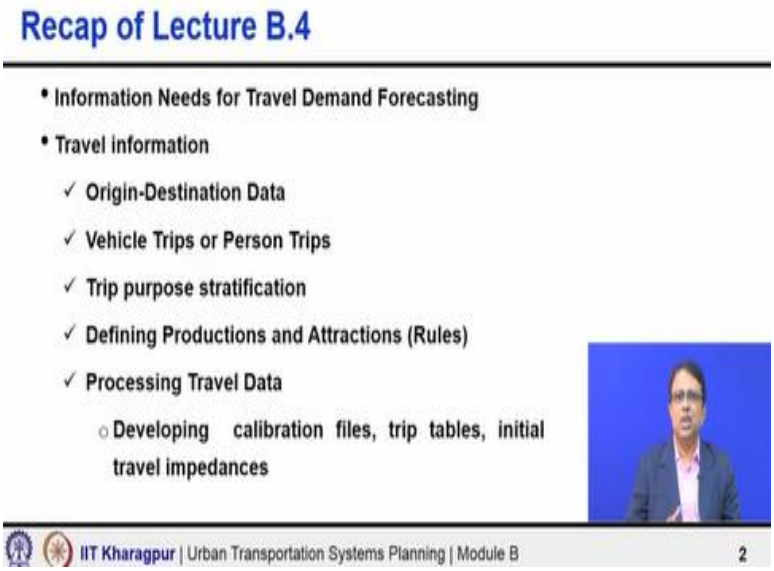
Urban Transportation Systems Planning
Prof. Bhargab Maitra
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture - 10
Data Collection and Techniques

Welcome to module B lecture 5, this is the last lecture of this module.

(Refer Slide Time: 00:19)



In lecture 4 we discussed about the information needs for travel demand forecasting and in specific we discussed about the travel related information, specifically the importance of the origin destination data. Then the issue related to considering vehicle trips or person trips. We said that it is context specific and also if you are using public transport related to work then the person trips data is very, very important.

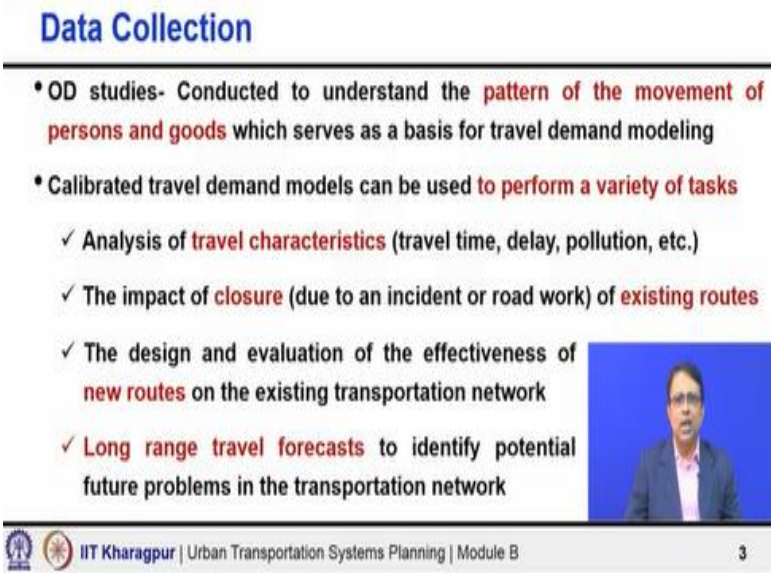
Then also we said the kind of trip purpose stratification we should use depending upon the context, we give lot of examples as well. Then the confusion related to defining production and attractions taking the origin and destination that discussion we had. We said that if home is there either as origin or as destination then trip is always produced at home and get attracted to the other end. Other end is always the attraction.

And, if neither end is the home then the origin is the point of production and the trip is getting attracted and destination is the attraction. Then we talked about the processing of travel data, how to develop calibration files, what it means then how the trip tables are developed based on the calibration files and then finally how to develop initial travel impedance because that is very important.

When we start the fore stretch planning particularly in set of distribution or mode choice we need to use the travel time data. Now, of course the actual loading is not known at the time but you start with something, you need to start with something. So that is the initial values that we can get from the processing of travel data because you have a lot of data. We know that people are travelling from one point to another point, one zone to another zone using different modes.

All those data are available, so the initial travel impedances may be developed from this data which subsequently once you do the actual modeling then you know the actual loading and then may be the travel time and the travel cost may be revised. So according to it may be obtained but initial travel impedances may be developed based on that.

(Refer Slide Time: 02:54)



With this background in this lecture we are going to mainly focus on the data collection. Although we use various types of data but in the forestage transportation planning context it is very important to have the OD data. So we are now specifically going to talk about the OD

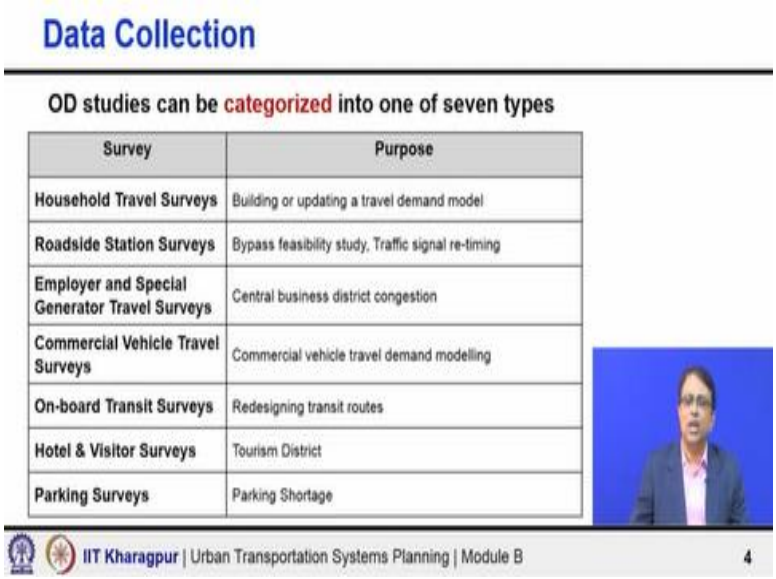
studies which are conducted to understand the pattern of movement of persons and goods which serve as a basis for travel demand modeling.

You want to develop the trip distribution; you need it sometimes even the mode wise OD data. It is also used, right for other works related to forestage transportation planning and why it is very, very important because OD data is used basically for calibration of travel demand model and once you have a travel demand model you can do so many important and very practical things for example, analysis of travel characteristics.

What would be the travel time, what would be the delay and if there is a travel scenario then what kind of pollution or the vehicular emission that is expected. Then what would be the impact of closure of existing routes maybe if due to certain incidents or road work, some lanes or some segments are to be closed then what would be the impact? We can assess it. We can save the design and evaluation of the effectiveness of new routes introduced in a transportation network.

Also we can do the long range travel demand forecasts to identify potential future problems that if it happens then down the line then what kind of transportation problems that we are going to encounter and all such kinds of extremely valuable things can be used using a calibrated travel demand model. For doing this calibration you need the OD data. So the OD studies are really very important.

(Refer Slide Time: 04:56)



Here we are telling that the OD studies can be categorized into one of seven types. We will say what are those seven types? a detailed discussion will be made subsequently in this lecture only. First is the household travel survey here, basically you know we go to the households and then collect the travel data. How and what further details we will discuss later but these are typically done for building or updating a travel demand model.

Next is the roadside station surveys, in a variety of ways we can collect the OD data from the roadside station surveys. Typically may be for bypass feasibility study, or even for traffic signal re-timings we need to do this kind of data may be very useful. This kind of data collection may be very useful. Then the employer and special generator travel surveys also may be done as a part of OD studies.

What we do typically when the central business district, the CBD congestion is really the keys. When we are thinking what we can do, how we can avoid bringing so many people at a given time at this location. So we know we do some OD studies covering employers and special generator travel surveys that is the thing. Then commercial vehicle travel survey, this is also one way of doing the OD studies.

Particularly we do it when we want to do commercial vehicle travel demand modeling. So we need to identify and get the commercial OD data, so we can get it on board transit survey. This is

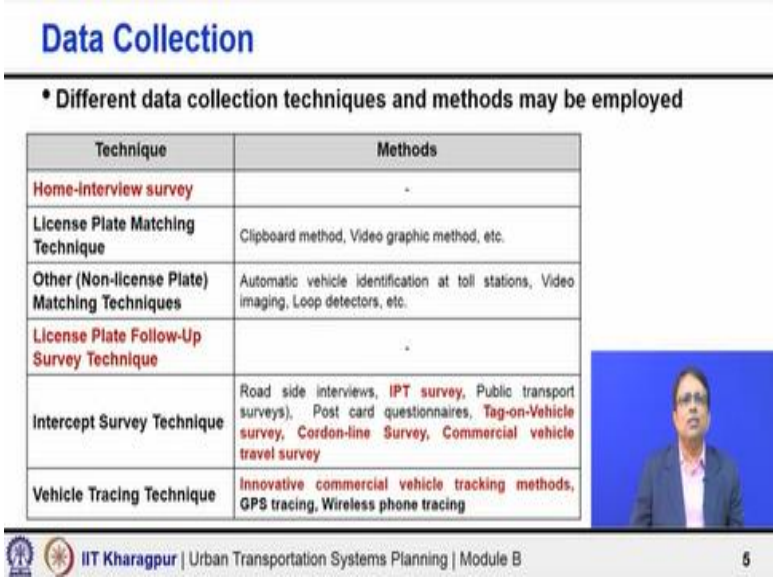
again another form typically while we do we want to do work related to transit. May be redesign of transit routes or you know further improvement of the service characteristics of the transit. All this can be done then hotel and visitor surveys are also done as a part of OD studies.

Typically where we do where there is a tourism district or there are certain routes in the study area where a lot of tourists come from outside. So they come, they stay. You know, always in urban areas where tourism is really affected an important consideration. There are certain catchments where the tourists come and stay. In the city anybody you ask will know that where the hotels are located and typically tourists come and stay there.

So there we do again as a part of OD studies, we do hotel and visitor surveys. Because we interact with them from where they are coming, what are the places they want to travel and what are their expectations and so you know extract the relevant data which are important for the forestage demand modeling. Also OD studies may be done as part of the parking surveys because you know urban areas parking is a major problem.

There are quite often imbalance between the demand and supply of parking in civil areas. So you need to understand what kind of parking shortage is there or even what king of parking management or even what other improvements are possible. So as a part of parking surveys also in the form of parking surveys we can also still do the OD studies. So that is the different categories that is the different ways OD studies can be categorized.

(Refer Slide Time: 08:41)



There are different data collection techniques and methods which may be implied for this OD data collection. One is basically the home interview survey that simply you go to different households and you do home interview data home interview survey and collect the data. Then license plate matching technique; this one technique which is used extensively again, so at various locations we try to match the license plate.

Then automatically you know how the vehicle is moving. So may be two external card and plate match the license plate. I find the same license plate. So I know that basically external to external trips. Just one instant example, there could be many other many other applications. So that kind of data you can use it may be a clipboard method or maybe you can record things manually in a piece of paper or using some other means or you can do a video graphic method.

Video graphy you say, from the video graphic data you collect that one. There are other matching techniques not using license plate. There are other matching techniques, a lot of other matching techniques are available may be automatic vehicle identification at toll locations that is there you know in different toll plazas located in different place vehicles things and there. So toll stations one can use video imaging even in many cases if you are using loop detectors are used for data collection.

So they can also be matched. So even other non license plate matching techniques, where we do not really match the license plate, we do not record. But some other ways we will try to understand the origin destination. Then license plate follow up survey, when we get the license plate and we know that the vehicle is at a given location at a given time and then subsequently we can communicate through telephonic interview or through post card or through email, so many ways communication is possible.

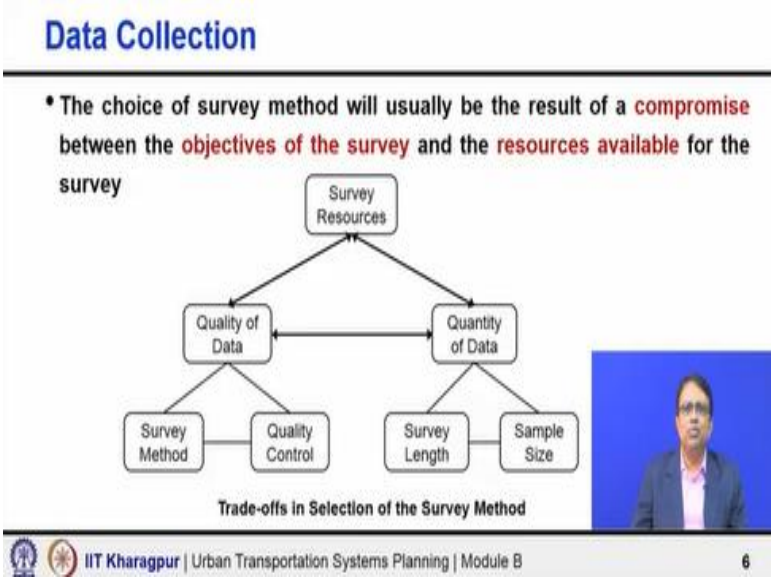
Then intercept survey techniques. This is generally you can do it for any kind of roads even for pedestrians also one can do. For passengers you know different you know passengers of public transport, intermediate public transport. In all cases you can do. So roadside interviews you can do it for we can do it, or we can do it in IPT survey, may be at the stations, or the stand, public transport also at the bus stops or the stations or even inside vehicles.

So basically we are intercept survey techniques we say intercept users or transport users are taking. You can also use a postcard questionnaire, you can use tag-on vehicle survey that these are again other intercepting vehicle not intercepting passenger. We can intercept vehicles, intercept passengers so many ways. Generally we say it as intercept survey technique. So tags on vehicles are civil intercepting vehicles.

So in a cordon-line survey, commercial vehicle may also be surveyed by intercepting. So there are so many most of techniques are there which may be or approaches are there which may be used to do the, to intercept either vehicle or passenger or so, passenger vehicle, goods vehicle or individual passenger to and then collect the required data. We can also use vehicle tracing techniques.

Which may include innovative commercial tracking methods, GPS tracking, mobile data, you know by use of mobile data, wireless phone tracking. There are a number of things can be done here. So these are all different techniques.

(Refer Slide Time: 13:04)



Now, when we are doing the data collection the choice of survey method is obviously a compromise between what is the objective of the survey and what kind of resource is available because these are the two major determinants. What is the objective, what will I do with this data and for I am collecting this data and second, how much resource is available to me. So typically I would say that I have shown it like this, you can see, one is the survey resource.

How much time it is available then what is or what is the expectation in terms of the quality of the data and what is the expectations in terms of the quantity of data. So it is quality of data, quantity of data and the available resources. When I am saying quality of data then two aspects are again very important. The quality expectation of about what level of quality, what quality data is requiring that will say what kind of survey method we would deploy?

And also, what kind of quality control I need to adopt. So what survey method and what type of quality control is needed that covers the quality of the data. When quality of data means quality aspect of the data, I should not say the quality of data but the quality aspects of the data. So the quality aspect of the data is important to judiciously select the survey method and formulate the quality control aspect of the data.

You send the data collected when you are doing a personal interview, so then you know you tell the enumerators you do some pilot team you know you just ensure that they are able to speak

clearly explain the things clearly in a face to face interview, so all this are there. So the quality of data means the quality aspects of the data and for that which survey method that is very important and what kind of quality control.

When you see the quantity aspects of the data then that means how much data I require basically that will basically relate to the survey length. For how much duration I should do the survey, one day, two day, or only during the peak hour, full day, multiple days in a week, weekdays and weekends and holidays separately and also it will relate to the sample size. The sample size is very, very important.

Because we are always use all the data we are collecting we use sample data. So the sample has to be representative of the population. There are different sampling techniques and you know that itself is a subject area. So you really need to understand different sampling techniques and then what kind of sampling techniques would be appropriate for a given context, so all these considerations are there.

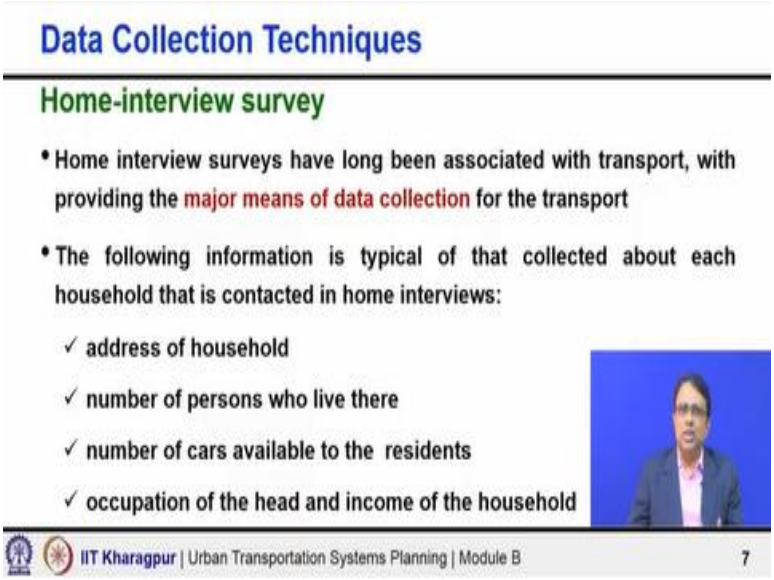
I am not going into details because detailed discussion about sample size and sampling strategy this is actually beyond the context. Sampling strategy is again related to the quality of the data, sampling strategy because if your sampling strategy is not there then you will not get the good quality. So quality may include enumerated training, quality may include also the type of sampling what you are using for a given context or what you are adopting for a given context.

So you know given a resource how best we can handle the quality aspect and how best we can handle the quantity aspect and in all together the survey methods, quality control aspect of the data. Quality control as I said in terms of enumerator, training, the correctness if you are using automated image in data collection then what is the reliability or how good you can extract the data, the correctness of that all these are important and even the sampling strategy all will come.

And in the quantity of data in that perspective maybe what should be my survey length and how much sample I require there are different sample size calculated for different surveys. Well

establish sample size calculators there depending on certain parameters. We can understand that what would be the minimum sample size requirement. These are the very, very important aspect.

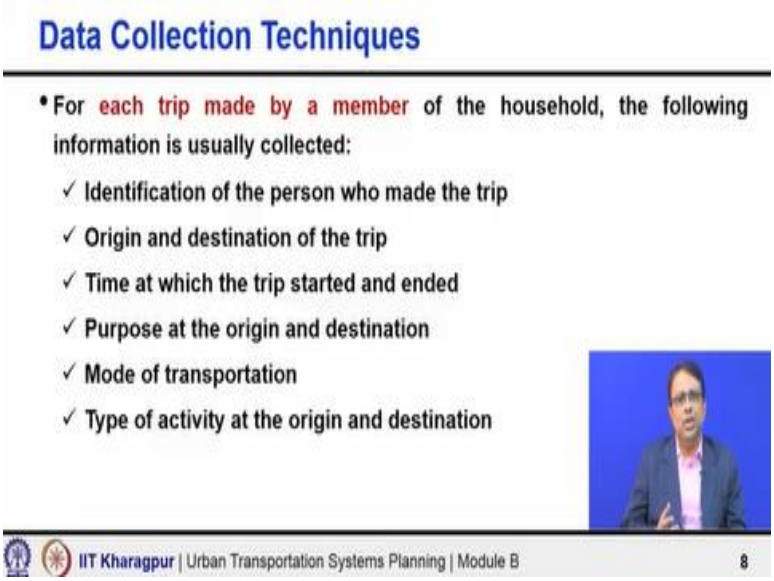
(Refer Slide Time: 18:19)



Then we now briefly discuss about the various techniques as I say in home interview techniques it has long been associated with the transport. Lot of people have been using it which is providing the major means of data collection for the transport. Typically what we collect is like these, address of the household, number of persons live there, number of cars which are available to the residents.

What is the occupation of the head of the household and then what is the income of the household? so all these kinds of data are collected.

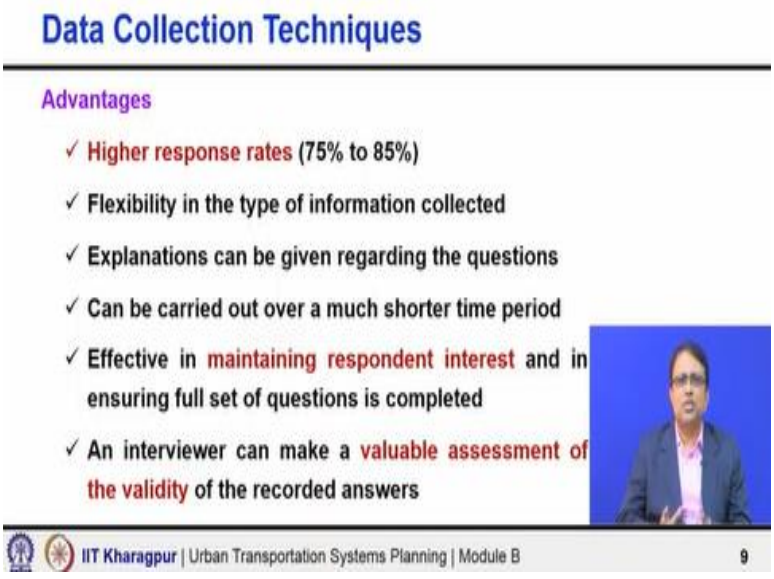
(Refer Slide Time: 18:54)



Along with these for each trip the trip details. For example, the identification of the persons who made the trip in a family there may be 4 persons who are making trips. So each of them each person each trip details are recorded. So we need to identify the persons who make the trip, we need to record the origin and destination of the trip, time at which the trip started, time at which it ended so you know how much it took for the trip to complete.

Then purpose at the origin and the purpose at the destination, what is the mode of transport that is used, what is the type of activity at the original destination. All these data's are collected from typical home interview survey.

(Refer Slide Time: 19:45)



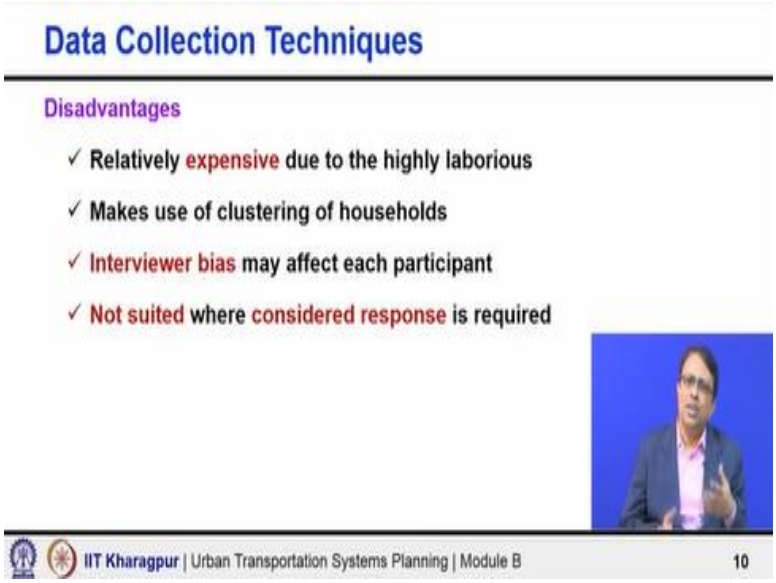
It has got certain advantages, for example, you can get very high generally high response rates if you are personally doing a home interview survey. Just normally from 75 to 85 % you know the response rate success rate you can get and there is a lot of flexibility in terms of the type of information one wants to collect. Explanation can be given regarding the questionnaire because face to face you are going in the household level and collecting the data, talking to them and normally people hopefully listen and spend little bit of time.

If you catch somebody when he is going to the office he will be very busy and may not be interested but if somebody is willing to talk to you at home and they can ask you explanation and you can get a chance to explain also or clarify some doubts. Then it is effective in maintaining respondent interest and in ensuring that full set is completed. You are not like the postcard survey or something you leave it to them.

And, you may not know whether they may return or they may complete certain questionnaire they may not complete some questionnaire but here we have better control. An interviewer can make a valuable assessment of the validity of the recorded answer. So if somebody is going and doing a home interview survey then the respondent is actually responding. So from the response one can judge about the quality of data.

So if you know that no, he has just had to give some reply so he just gave some reply and completed it you may not use the data. You can apply that kind of judgment that is really an important part in this context.

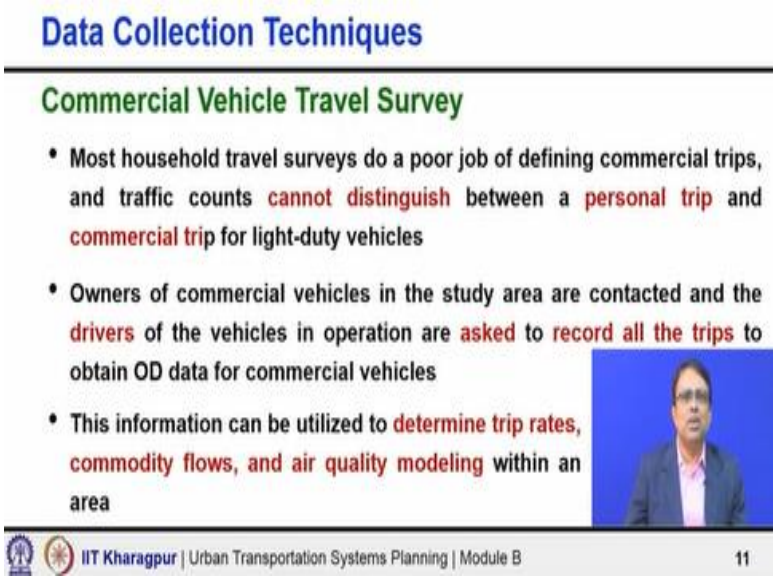
(Refer Slide Time: 21:32)



But it has got as usual some disadvantage because it is expensive, I will say not relatively expensive but it is really expensive in big cities, for say Calcutta or Bombay or Chennai or Bangalore, if you want to and if you really want to do correct requirements in symbol. It is enormously expensive. It is valuable but it is really, really valuable but expensive level. Another disadvantage is it is sometimes you know interviewer bias they are may affect the participant because ultimately you are employing large number of enumerators.

Because it is not a job of 1 or 2% you need a much bigger team to do it. So although you do such quality control training and everything, still there may be a problem it is not suited where considered response is required that is another part.

(Refer Slide Time: 22:34)

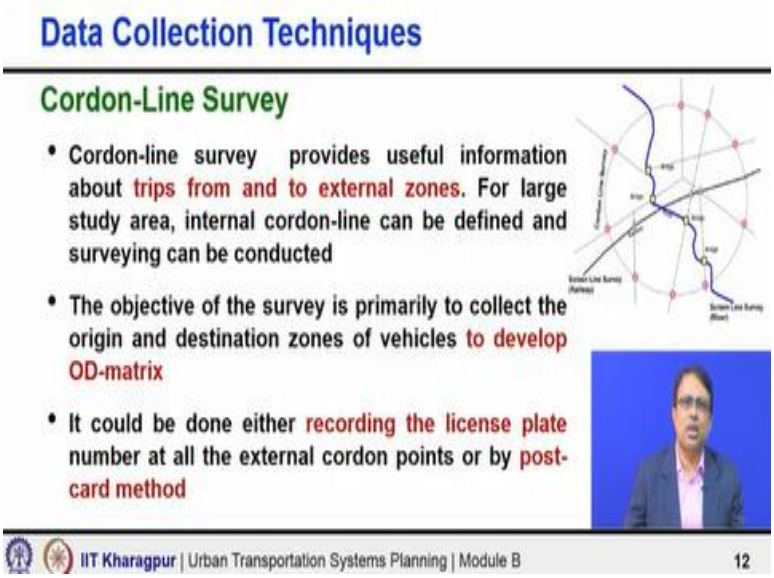


Then the next method is the commercial vehicle travel survey. Next part is commercial vehicle travel survey this is most household travel surveys do a poor job. We have seen that in terms of commercial trips they do not come, they do not come. Personal trip and commercial trip come with different speeds but not the commercial trip. So you need to do something else something extra that is what this one.

Owners of the commercial vehicles in the study area are contacted and the drivers of the vehicles in operation are asked to record all the trips to obtain OD data for commercial vehicles. So you can contact the operator contact the driver wherever applicable and which one is suitable for the given context which is appropriate for the given context and then get all the data. What commodity they are studying where they are going.

So especially I mean normally the personal trip data collection and the commercial vehicle data collection is not usually the same. And I would say that getting the commercial vehicle data or the goods movement data is much more challenging. So we need an extra effort. This information can be utilized to determine trip rates, commodity wise flow and also indirectly related to the air quality modeling within an area, so that is the thing.

(Refer Slide Time: 24:11)

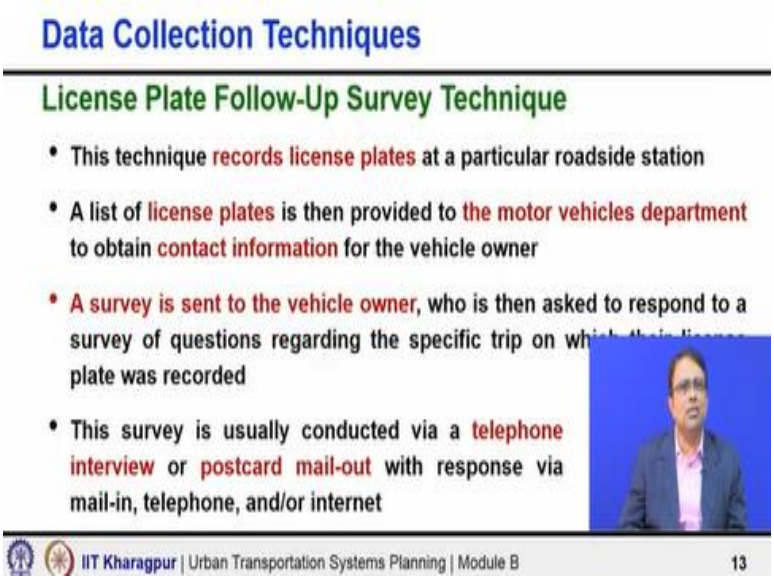


Then the cordon-line survey; we say what is a cordon-line you know that cordon-line is the external to the internal, internal to the external trips are getting interchanged, you know that is the imaginary line. So we have to do a cordon-line survey because it provides useful information about the trip from and to external zones because it is the demarcation of the study area boundary.

If you do not remember what is cordon-line it is the imaginary line demarcating the study area. Study area boundary is delineated by this cordon-line. So obviously many roads will intersect and on all those locations we need to do the traffic count and OD and understand how this movement is taken. So the objective of this study survey is to primarily collect origin and destination zones of vehicles to develop OD-matrix.

And it could be done either by recording the license plate number at all the external cordon points or by postcard method or even a variety of methods. You can simply note down or you can using the camera to record it. You can stop at the cordon-line location. Several techniques can be deployed depending on the context, not just one but cordon line survey means at the cordon location we are intercepting vehicles, collecting the information and trying to establish that OD, that is the basic objective.

(Refer Slide Time: 25:55)

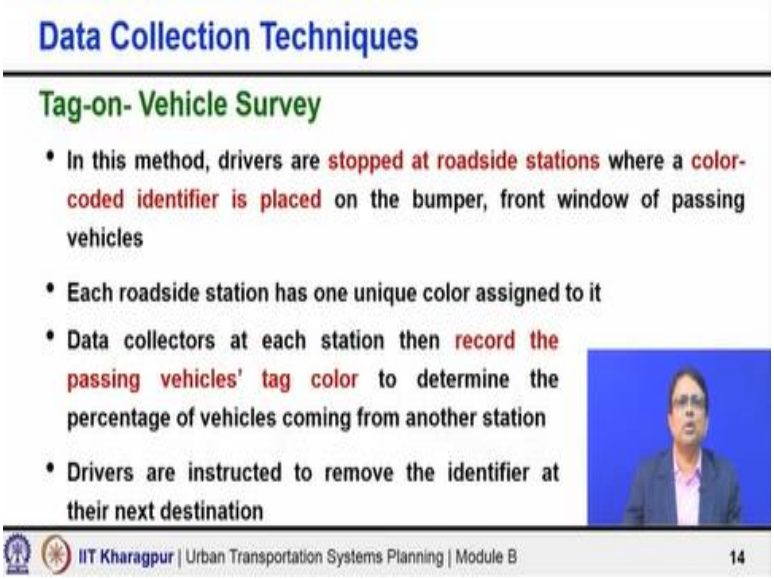


Then license plate follow up survey technique also may be used. This technique records license plates at a particular roadside station. So basically just to record this vehicle was there at this location at this time. So then what we do a list of license plate is provided to the motor vehicle department and we get the contact information of the vehicle owner may or may not be applicable in all situations.

You may not be able to do this everywhere but yes these are the methods which are available. Then a survey is sent to the vehicle owner who is asked to respond to a survey of questions regarding the specific trip on which their license plate was recorded. So it is something like this vehicle was located in this area at this time at this thing so what for the travel was made, where from the travel was made.

What was the purpose of the travel and so on whatever relevant data, this survey is usually conducted via a telephone interview or postcard mail-out with response via mailing, telephone, and or internet. So many ways we can collect the data whatever is suitable.

(Refer Slide Time: 27:17)

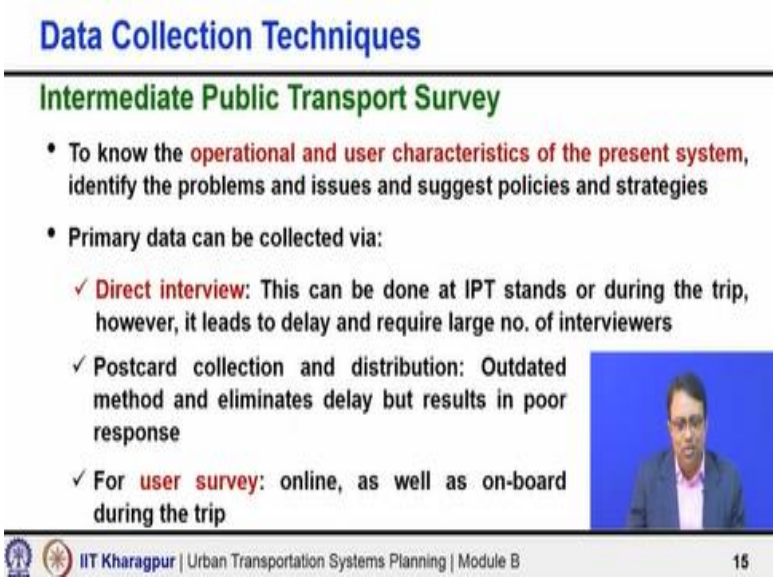


Then tag on vehicle survey; in this method, drivers are stopped at roadside stations and what a color-coded identifier is placed on the bumper, front window of passing vehicles. So maybe you will do it at different strategic locations may be even for external vehicles also one can do it. So every location has got a color tag. So when a vehicle is entering from that the tag is given. So wherever at other point the vehicle will be noted they know that the vehicle which has entered from that point actually has come to this area.

So that tells you the movement, the pattern of movement that is happening. So each roadside station has unique color assigned to it and data collectors at each station then record the passing vehicles' tag color coming from another station. So I have observed 10 vehicles out of which 2 are red out of 10, I noted that wherever the red tag has been used whatever I am getting is a only 20% of the vehicle is actually coming from that location or through that point actually.

So drivers are instructed to remove any identifier at their next destination. So once it is done you remove it because you do not want to create further confusion.

(Refer Slide Time: 28:56)



Intermediate public transport survey as I said to know the operational user characteristics of the present system. It is not only the intermediate public transport survey but it is also for public transport survey for all IPT, UPT modes. Then again so many ways primary data can be collected, say by direct interview this can be done at IPT or UPT public transport or intermediate public transport station or even during the trip.

So you can do it at the stop when people are waiting for the bus, or waiting for a shared auto or shared taxi to come you can talk to them and get it or for bus particularly during the trip also. While you are travelling then you can interact with them in the vehicle. So, you know there are obvious advantage in terms of how many locations you can cover or how many samples you can get all those.

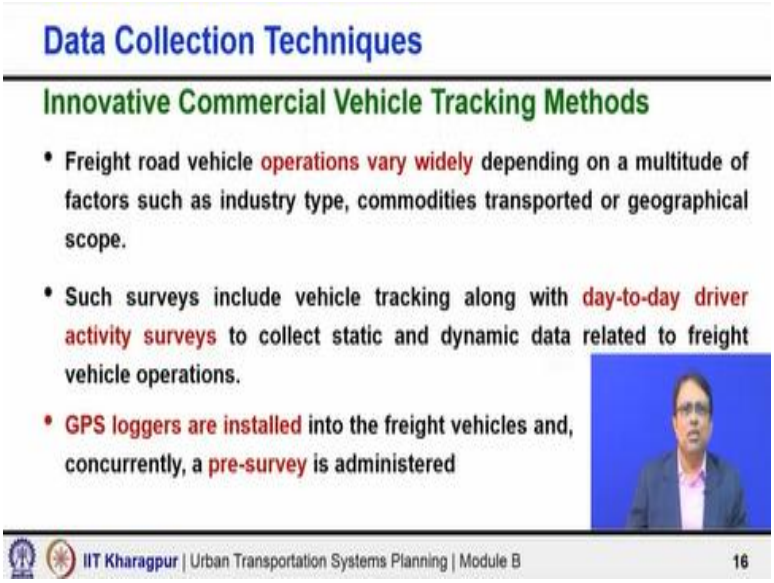
Then postcard collection and distribution also is there but I will say it is basically outdated. Nowadays I do not think it is very appropriate and response rate is also not good. So the face to face interview at least in the Indian context is works much better. For user survey you know it can be online as well as on board during the trip. So the first one was direct interview and the second one is user survey.

So online we have done lots of services regarding the IPT, NPT usage. So, distributed in central opinion; social media using Facebook, Twitter or even Whatsapp. So to you know larger

audience and request them to participate sometimes by email. So all these platforms can be used or user survey can be also done on board during the trip, I mean simply put the person, enumerator we call who collects the data.

Enumerator is also travelling in the bus or train and can interact with people and get all the necessary information. So these are all different methods.

(Refer Slide Time: 31:22)

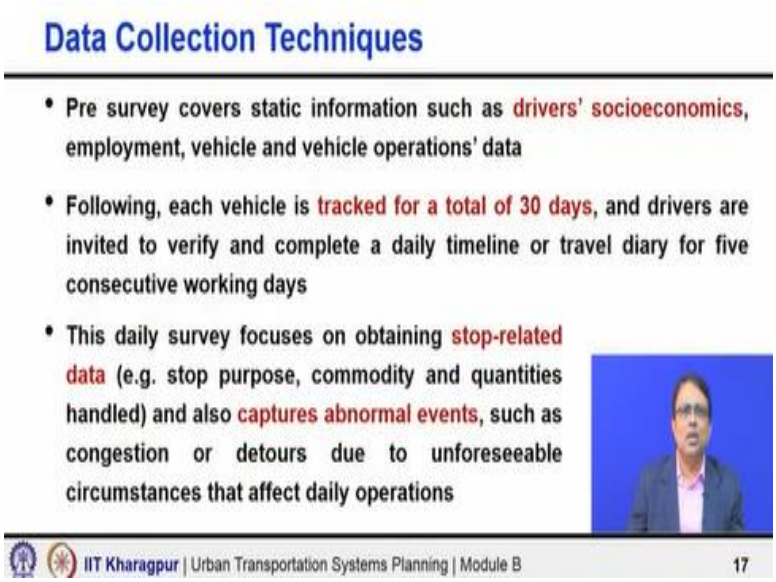


Then one can also use innovative commercial vehicle tracking methods. This basically freight road vehicle operations vary widely depending on many factors, say the type of industry, commodities transported and geographical scope lot of things. You know every commodity has got different factors, you know what kind of commercial vehicle movement take place that they also depend on what commodity the vehicle is carrying.

It is very much requirement very much commodity specific and it is a different world all together. Goods vehicle movement is a different world and sometimes much more complex. So this kind of surveys include vehicle tracking along with day to day driver activity surveys to collect static and dynamic data related to freight vehicle operations. So in many vehicles are going with GPS loggers are involved for the commercial vehicle.

So you know the complete tracking is done and from that where the vehicle went and where is the destination and everything. A lot of information would come out of that a pre-survey is also administered.

(Refer Slide Time: 32:49)

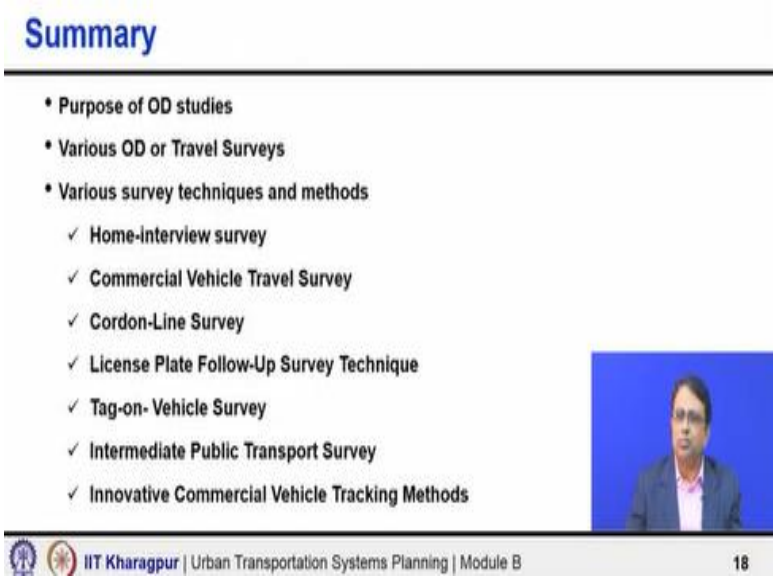


And the pre survey covers static information such as drivers' socioeconomics, employment, vehicle and vehicle operations data all these are collected and then GPS data loggers and GPS data is also taken. Following each vehicle is tracked actually for a total of 30 days after doing the pre survey and the drivers are invited to verify the data that have come out just to do as a check.

So if there is any mismatch or anything and complete a daily timeline or travel diary for five consecutive working days what they did or not and what all this daily survey focuses on obtaining stop-related details because obviously the commercial vehicles will stop for parking, for food, rest and all other things. So you can say stop purpose, commodity they were carrying, how much quantity they were handling all site related information.

And also captures abnormal events, such as where they have faced congestion and where they had to do detouring due to unforeseeable circumstances that affect daily operations. All those kinds of things, you know lot of lot of data will come out of these innovative techniques.

(Refer Slide Time: 34:18)



So in summary, we discussed in this lecture the purpose of OD studies, various OD or travel surveys and various survey techniques like home-interview survey, commercial vehicle travel survey, cordon-line survey, license plate you know follow-up survey technique, tag-on-vehicle survey. Then IPT, NPT surveys, innovative commercial vehicle tracking. So these are the different methods which help us.

And different techniques which can be deployed or used judiciously to get the required data for the travel data for the travel demand forecasting. So with these we complete this module B that is on overview of the forestage transport planning process. Thank you so much.