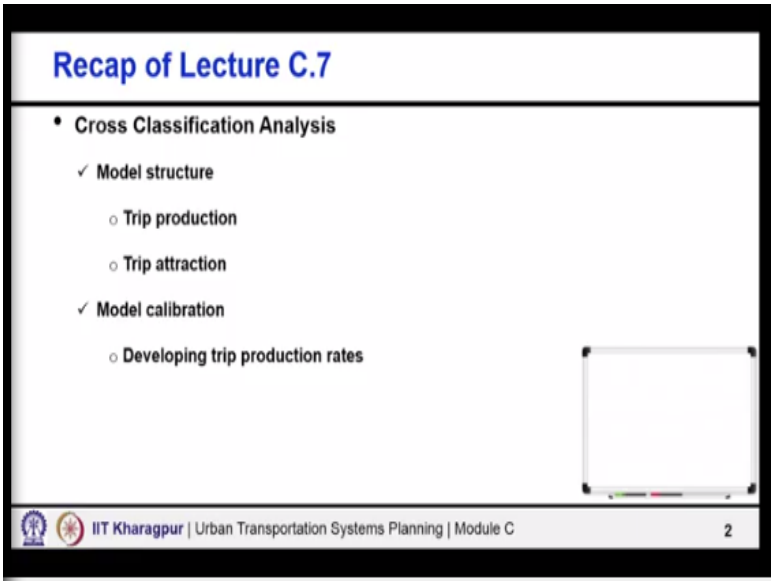
Urban Transportation Systems Planning
Prof. Bhargab Maitra
Department of Civil Engineering
Indian Institute of Technology-Kharagpur

Lecture-18
Cross Classification Analysis: Model Calibration

Welcome to module C, lecture 8.

(Refer Slide Time: 00:21)



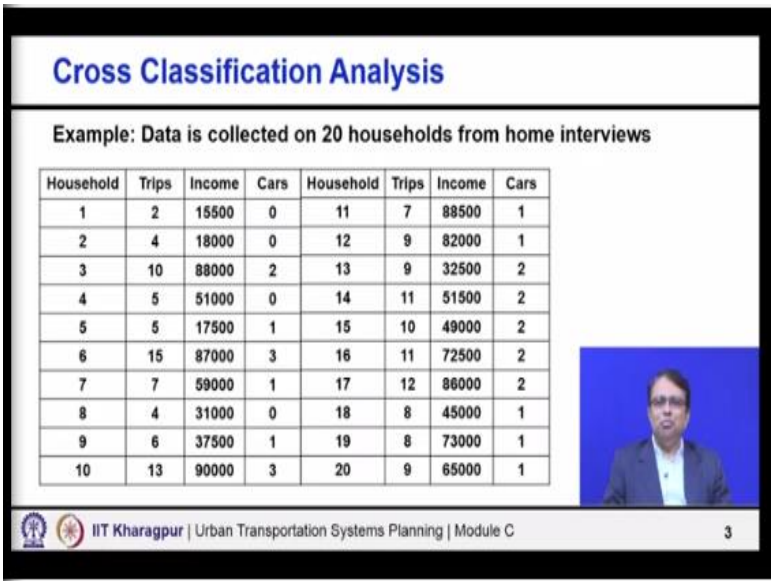
We were discussing about the cross classification techniques, what we do in the cross classification, we stratify households based on certain household characteristics, create certain categories and then calculate the trip rates for each category. Similar thing we do it for trip attractions as well using cross classification techniques. The basic assumption of cross classification technique the trip making behavior of households of certain particular stratification remain stable over time, and that is a very, very vital assumption.

I shall repeat it probably in every lecture, I will be talking about cross classification or category analysis. We discussed about the model structure, the very first step to go ahead do with cross classification analysis is to develop the model structure. We discussed how we structured the trip production model, how we structured the trip attraction models, logically identifying the variables which influenced the trip rate.

And using those variables we create suitable number of groups. And then create the number of categories or classes using this cross classification technique. Once the model structure is defined, we know exactly how many categories we are creating or we have created, and what all variables we have considered, how exactly the groups have been created? So, you have seen that sometimes using income we have created 5 groups, sometimes using a car ownership, we considered 3 groups.

So, 5 into 3 in that case, there are 15 groups. Similarly the attractions also 3 types of employment we considered in that example, and 3 types of trip home based work, home based other and non home based trips, so 3 into 3, 9 categories we have created. So, first case maybe 15 categories and the second case 9 categories. So, similarly once we have created these categories, the next step calibration is basically establishing the trip rate for each such categories we have created. So, what we did?

(Refer Slide Time: 03:15)

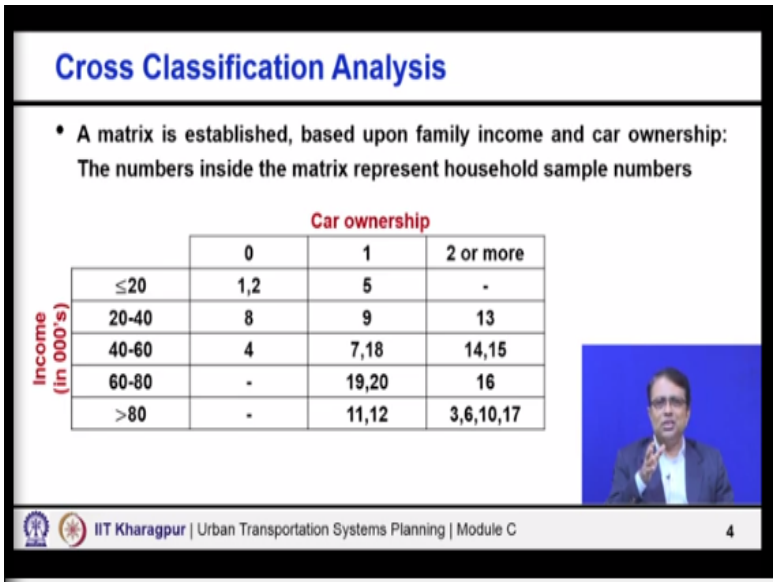


Household	Trips	Income	Cars	Household	Trips	Income	Cars
1	2	15500	0	11	7	88500	1
2	4	18000	0	12	9	82000	1
3	10	88000	2	13	9	32500	2
4	5	51000	0	14	11	51500	2
5	5	17500	1	15	10	49000	2
6	15	87000	3	16	11	72500	2
7	7	59000	1	17	12	86000	2
8	4	31000	0	18	8	45000	1
9	6	37500	1	19	8	73000	1
10	13	90000	3	20	9	65000	1

We took an example just to say that, how the household trip rates can be established. So, we consider an example where 20 household data is available with us. Each case we know for each household, how many trips are being made and what is the income? I would like to remind you once again that in no practical example you will be able to develop trip rates for so many categories using 20 household data, it will simply not be acceptable, please understand that.

I have taken 20 because I want you to explain the procedure. So, to explain the procedure, I did not really consider some several 1000 household and their characteristics, but I took only 20 households, the maximum data probably I can show in one slide. So, with these now how we develop the trip rates.

(Refer Slide Time: 04:29)



So, the next step is, I know that we have 5 income categories, let us say less than 20,000 rupees income per household 20 to 40,000, 40 to 60,000, 60 to 80,000 and more than 80,000 rupees per month, so these are my 5 categories. You can consider this do not think that always you have to consider only these many categories are you have to take 20,000, 20, 000 to 40,000, you can take as appropriate for a given context, this is only an example.

Similarly based on car ownership, we have created 3 verticals that means 3 classes. Zero car, that means household does not have any car, households have only one car and those households which have 2 or more cars. Now, these 20 household data if we look at this, say for example let us go back, let us take the first household. Income is 15,000 and cars 0, so income 15,000 means income less than 20,000 and car 0.

So, it comes actually the left top most left corner cell, which represent a category with zero household and household income less than 20,000. Similarly if you consider household 2, income is 18,000 and car ownership is 0, so 18,000 again less than 20,000 and car ownership 0,

so 1 and 2 comes under that same category. Similarly household 3 gets a place in a group where the income is greater than 80,000 and 2 or more cars in the household.

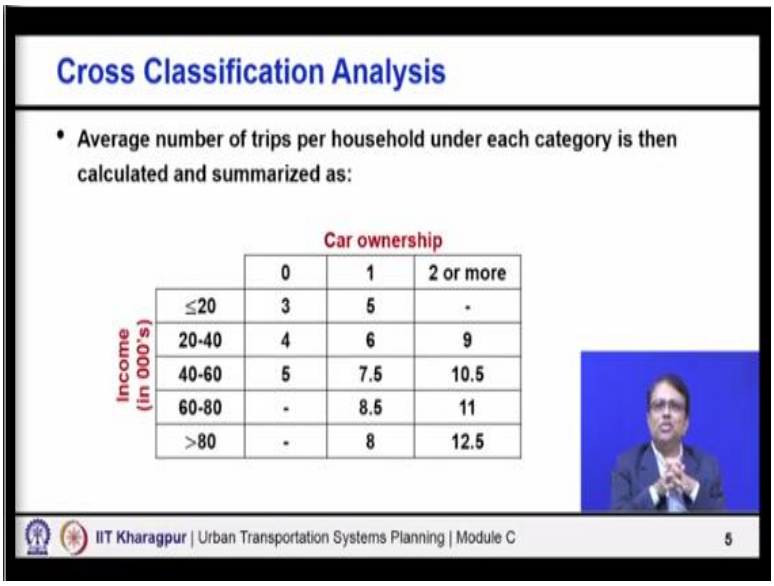
4 gets place in a category where the income is in the range of 40 to 60,000 and zero car ownership. 5 gets a place again in a category with less than 20,000 income and 1 car in the household, like that all 20 households we have distributed. Now because I have considered only 20 households I said you repeatedly house based on 20 household data, you cannot do anything. But this is just a classroom example to demonstrate you and or to explain you the process, so that is why I have taken 20.

So, as usual some cells we got maximum 4 households, one case 4 household then 1, 2, 3, 4, 5 case 2 household 1, 2, 3, 4, 5 case, we got only one household, and 3 cases we did not find any household from this 20 household data does not matter. So, like that in your case, when you do a realistic survey and in real life you try to apply cross classification analysis, you will get so many households.

So, you have to get every sale representative number of data, you cannot go ahead like me in a real world work. But let us say these are ok for the classroom example purpose, and I have identified which household falls under which category. And then all the 20 households have distributed into respective categories, so cell wise I have distributed. So, maybe instead of 20 household, you will probably get data from 5000 households, 6000 households like that.

And all 6000 households same way similar way you will distribute, each household will find a place. There cannot be that one household you are not able to understand where you are fitting, it is impossible. The way the groups have been created or groups are to be created that every household will get a place, so you put it. Next step what we do?

(Refer Slide Time: 09:22)



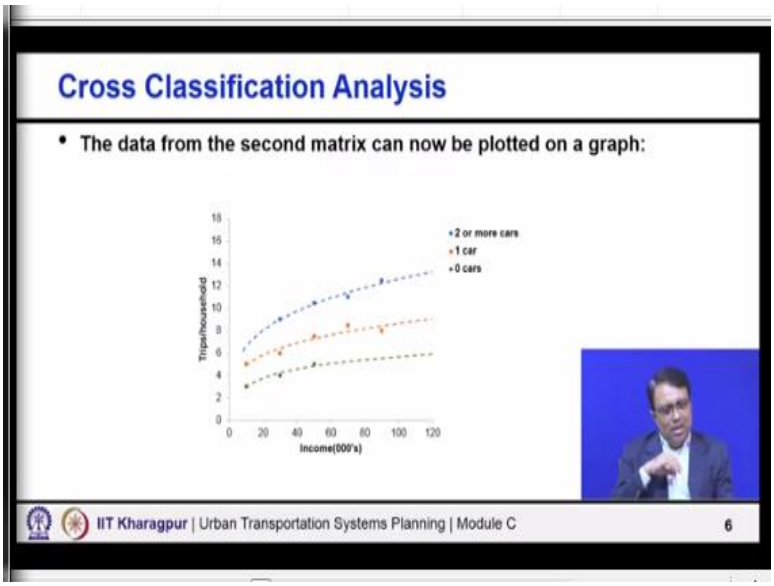
We now calculate the trip rate from those, how? Let us say, the group where the income is less than 20,000 and zero car ownership, I know that household 1 and household 2 are placed there. So, I have 2 household, household 1 makes how many trips? Two trips, household 4 makes how many trips? 4 trips, so total number of trips made by all these households in the category is 6 divided by number of household 2, so 6 by 2, 3.

So, I calculate here you can see, that the trip rate for that category less than 20,000 income and zero car ownership is 3. Like that wherever I had at least one household, I have calculated the trip rate. Again for example purpose with 1 or 2 or 5 household you cannot actually calculate the trip rate that will not be acceptable but here for example purpose it is acceptable. So, let us not bother about the sample size.

So, like this, so I have first distributed the household in suitable appropriate group wherever whichever boxes they are suitable. Actually this household represent whatever characteristics they represent wherever they are supposed to be placed, I have placed them. Like that once you distribute all the households you have surveyed you distribute them in respective categories, then take one by one category and calculate the trip rate. How we calculate the trip rate, totally how many households are there, how many total number of trips they made?

Divide it by the number of households in that category. And in real life it has to be statistically representative sample size. So, let us say now I have calculated the rates, then what we will do the next step.

(Refer Slide Time: 11:50)



Now we go and we plot it, my x axis is income, my y axis is trips per household, x axis I have showed income 20,000, 40,000, 60, 80, 100 like that. And why is the trips per household, so 2, 4, 6, 8, 10 like that. Then whatever cell values I got in the previous table that is there where is trip rate, I have simply first plotted those values. So, I know my one category is less than 20,000, so maybe I have taken it average income is 10,000.

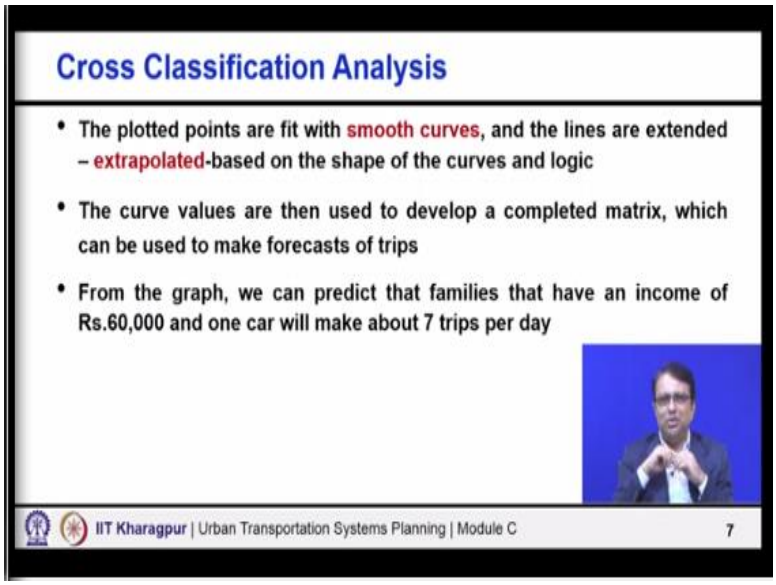
Because 20,000, 0 to 20,000 not everybody will have 20,000. So, it could be any value between 0 to 20,000, I have taken represented that group by an income of 10,000. Similarly 20 to 40,000 I have represented with by a point value of 30,000, 40 to 60,000 I have represented it here using 50,000. And accordingly these 3 lines the bottom line is for what zero car, then the middle line or middle points, the red color therefore one car.

And then the blue colors dots, they represent 2 or more, obviously for any income categories, the zero car households will make less trip, compared to that one car household will make more trips, compared to that 2 car household will make even higher number of trips**,** so that is what is there. So, once you plot these points, then what we do? We simply connect them using smooth

lines, fit the line. And since here in this example, we did not have many values many sales for not having some any household even. So, we have fitted and we assume that this fits or this car represent the data and then from the car we can pick up the value.

So, like that once you established, now if I tell you an income group exactly what is the trip rate, you can calculate it from the graph. Even a different income but within this range within this observed or extrapolated range, we can even get the values. If you wish you can even take that, so it is category wise and also any value of income practically can give you a trip rate following this trend or variation.

(Refer Slide Time: 15:18)



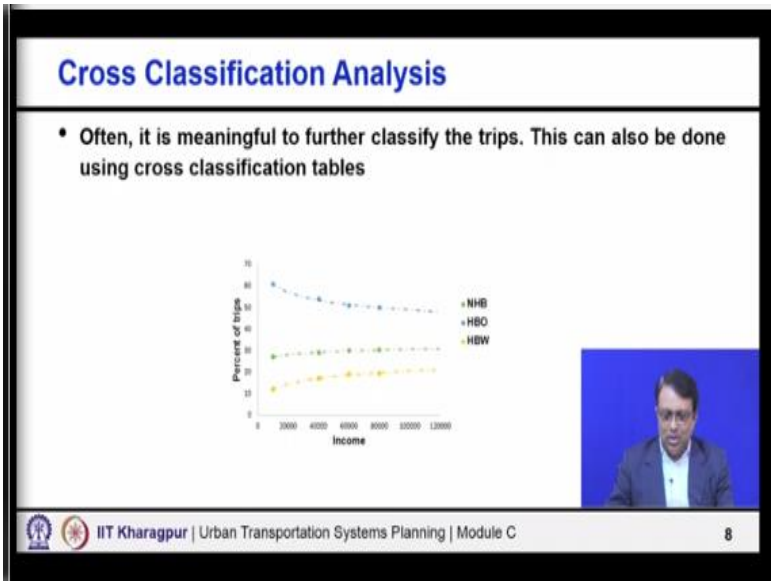
Now as I said the first we plot the points, individual trip rate values what you have calculated in the table and then they are fit or we do a curve fitting, smooth curve fit. And these lines are extended or you can say extrapolated based on the shape of the curves and logic. The curve values are then used to develop a complete matrix which can be used to make forecast of trips. Once we have established these rates either in the form of a table or in the form of logical curves, then the values are there.

Then you want to apply it in the future, you simply say how many households of a particular stratification is likely to be in the future in a zone, take that number assume that the trip rates become stable over time for certain stratifications. So, for that stratification, whatever you have

calibrated in the base here, the trip rate you have got, the strip rate will also be valid for that category in the future.

So, once you know the future number of households in that category, multiply it by the trip rate for that category to get the total number of trips that are likely to be produced for that category of household. Like that submit over now, for all categories which are present in that zone to get the zonal estimates in the future, so that is what is explained here nicely. Now we have now taken a second matrix, that also can be ok.

(Refer Slide Time: 17:25)



So, now often it is meaningful to further classify trips, I have discussed in details during the initial lectures on this trip generation, the trips maybe classified in so many ways. In this example, we have taken one such classification, home based work, home based other, non home based trip. So, if my total trip produced is 100%, then how the shear is likely to vary? Shear means shear for each type of trips, home based work, home based other hand non home based.

Now this is not the only way that you have to always do it in this classification maybe if you have idea, like by purpose, how they are distributed? From the observed data, how we get it? We get it from the same household data, when you have surveyed, you are probably also found out that some household make 3 trips, how many trips are being made for what purpose? or some

household making 5 trips then how many trips are made for what purpose? You must have the data.

In this case, let us assume that similar kind of data we also collected and therefore we developed probably exactly the same way for different income groups. Here again X represents different income categories, and Y represents where in this case what is represented by what percentage of trips. That when every household category based on income, whatever trips will be made, how much percentage of that trip will be home based work shown by red dotted line, and then home based other and non home based.

So the total has to be 100, so you are likely to get a shape like this. For say some categories, if one particular percentage of trip is higher. And for other categories, if that percentage of trip becomes lower, then obviously some other type of trip must have higher share. So, if you are getting a downward trend for a car, you must get upward trend for another car because then the total becomes 100.

So, here suppose, if you have collected the data, then exactly the way we calculated the trip rate, maybe for each category, we also calculated what is going to be the share and we also develop this kind of curve. So, the interesting part here is once you have this kind of curves developed, then using the previous curve once you know this is by stratification based on income and based on car ownership then you can find out.

Then that category of household how many trips are likely to be produced from household that distribute you can get. Then if you also have this representation available, this is also established or calibrated in the base year, then you know for a certain stratification of household certain category of household if total 100 trips are getting produced then how many trips are likely to be home based work, home based other and non home based.

Here also I forgot to mention that the 3 lines, 3 lines represent 3 trip purposes. Now, if you want to find out, no I want to estimate the work trip, I want to estimate the educational trip, I want to estimate the social recreational trip separately, you can always do that procedure is same. So,

what I will do? In that case when I am collecting the data from household I will also take as per that trip purpose classification, how the trips are distributed.

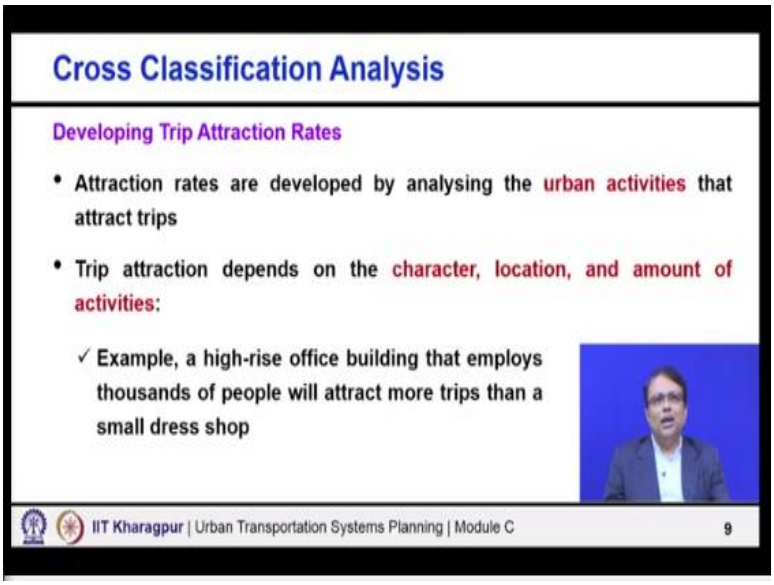
That means if some household has said that I have made 5 trips, then I will also try to get out of those 5 trips, how many are work trips, how many recreational trips, how many educational trips and so on and so forth. So, any kind of trip purpose wise classification you can make, you can make even how many trips are being out of those, how many are being made during the peak hour, morning peak hour between 8 to 10 and how many are made during off peak hours?

So, if it is actually the planners when you were collecting the data you know how and why you are trying to develop these trip rates. So, accordingly you will take the data and whatever data you have collected accordingly you should be able to do some further classification of it. It may be based on like this home based work, home based other, non home base or it may be based on the classification of trips or it may be during when you make the trip.

Also sometimes maybe by mode also, you can take also that, ok so many trips are being made. But then the travel was made by which mode maybe we will find somewhere midway public transport, somewhere made by predominant mode of transport or some or maybe by private vehicle like car and so on so forth, then accordingly I can further classify trips. Any kind of classification you wish you can use it, but what classification will be appropriate for a given context?

You have to decide that what is really meaningful, why you are doing the transportation planning study, what you can really forecast? Because it is to be applied the model is not for just calibration and everything we do in the base year we would not apply in the future. So, future what likely you are able to forecast and what is your objective of doing the fundamental objective of doing this transportation planning study? Say, it is all to be has to make sense end of the day, doable and meaningful.

(Refer Slide Time: 24:46)



So, now similarly we can also develop trip attraction rates, as I have explained. Now attraction rates are developed by analyzing the urban activities that attract trips. And as I said, the character, location and amount of activities all are important, please again remember our discussion what I said to you, when we are talking about the various factors, that amount of urban activity is important, character of urban activity is important.

Because if you say that giving everything means the more employment means more attraction, bigger shopping as compared to smaller shopping area means more attraction. So, the quantity matters, amount of activities better similarly the character of activity matters, not all employments will have equal level of attraction. Now a supermarket and another kind of activities are not going to attract same level of trip.

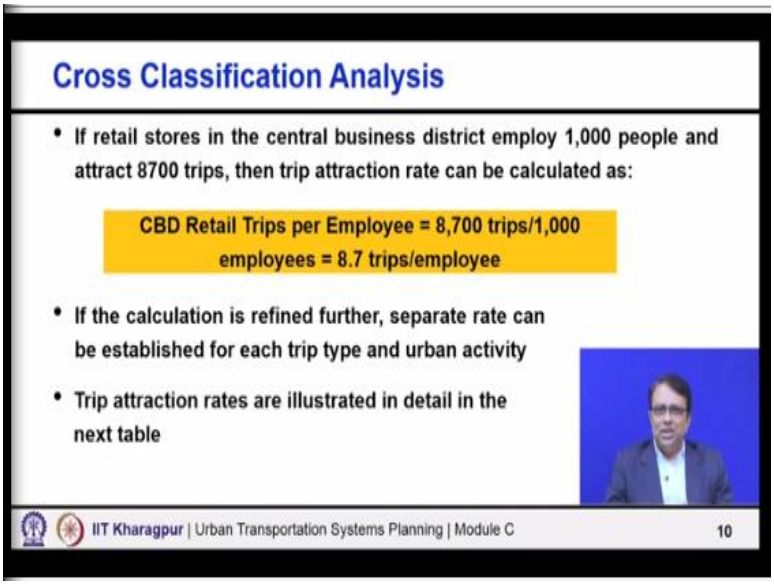
So, the character is also important, so character is important, amount is important and also the location. As I said that a shopping mall located in the heart of the CBD where you have wide roads, enough parking areas well planned and somehow locating in savoir where that the infrastructure is highly deficient, narrow roads and there is you do not know where to park vehicle, there is no off-street facilities.

On street parking is not a feasible option because the roads are so narrow, that if you park your vehicle others cannot go, now the attraction rates can never be same. So, remember that when we

are developing trip attraction rates, we consider activities but we duly consider the character, location and amount. I have given an example here say a high rise office building that employs thousands of people will attract more trip than a small dress shop. Here you can find out right here both the character and amount of activities are different.

Because one case we say high rise building another service is small shop, so the it is mixed up, it is we are selling that the amount wise also they are different. Whatever if you express it in terms of how many square meter area, the square meter area further high rise building for the first rate will be higher. And the square meter area will be much lower in the second case but also the character of activities are different.

(Refer Slide Time: 27:38)



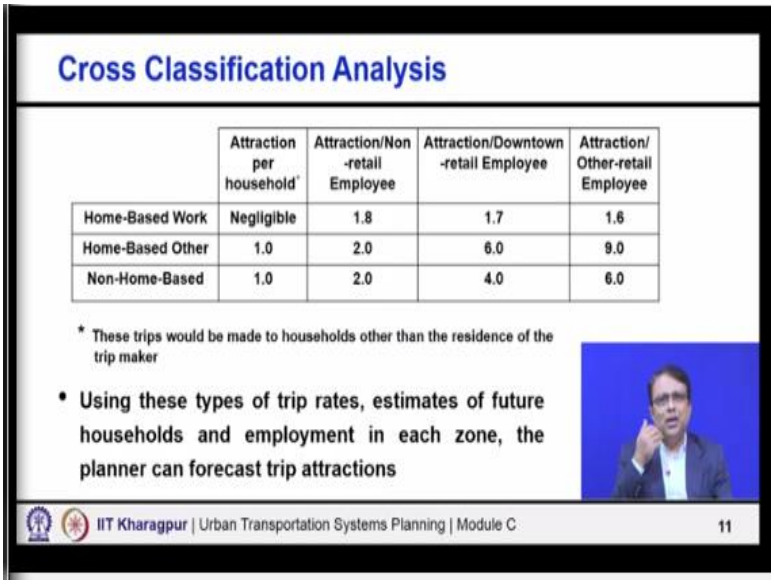
So, like this, then say for example, if retail stores in the central business district employ 1000 people, that is the kind of activities are there. And a survey reveals that 8700 trips are actually attracted to that land use. Because it is after all it is an attraction activity, the employments are an indicator, how many people are working means how big the activities are. And because of that activity 8700 people are getting attracted.

So, if I have to calculate the rate what I will do? I know I can calculate so maybe CBD retail trip rate per employee is 8700 divided by 1000, so 8.7 trips per employee. That means similar kind of land use in comparable areas, if I comparable type of employment, I know that every employee

means that establishment will attract 8.7 trips. So, simply if I can forecast, I will multiply it by this number to get the total.

So, in this case also, if the calculation is refined further, you can also generate separate rates for each trip type and urban activity. As I said earlier we got the total trips produced by household and then we could also get the classification, how much percentage will be for home based work, home based other, non home base. Similarly here also basic trip rates itself we can further classify. You get the total trip and then for different each type of activity, each type of trip we can create a matrix.

(Refer Slide Time: 30:05)



For example, let us say here I am showing you 4 types of attraction, attraction retail employee, other retail employee, retail employee and non retail employee and also the first column if you see attraction per household. Now do not get confused that while talking about attractions, why we are talking about households. Now these trips would be made to households other than the residents of the trip maker, this is not home based trip, these trips are different, these are more like social recreational trips.

So, we are trying to say how many trips are getting attracted, so household also may attract social and recreational trips. So, this is very important, again I am saying these trips would be made in households other than the resident of the trip maker. If it is residents of the trip maker it

is either produced or getting terminated, that becomes the household trip, that is the point of production, not the attraction.

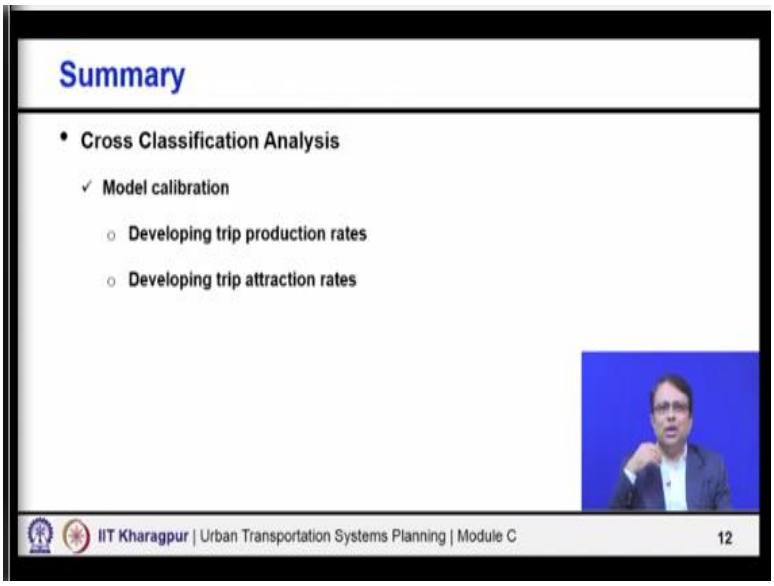
But this is different, that is why we call it as attraction, you may consider. And, so like that 4 categories and type of trip, again I used very similar classification in this example, what I used it for the trip generation, trip rate example. I used home based work, home based other, non home based. So, like that similar way as you have seen I just explained for one case, we establish this trip rate as 8.7 trips per employee.

Just similar following similar procedures, you can actually establish, when you have these 4 categories of land use, you are considering based on attraction and consider 3 types of a trip. Then in each case, you can calculate the trip rate separately, so it is again that way for each category the trip rates are established. Now using these trip rates estimate of future household as well as employment because you have established the trip rates for different categories of household considering income and considering car ownership.

Similarly for attraction, you consider different types of attraction in terms of land use and attractions and also consider different trip type. So, altogether you can actually forecast the trip productions as well as trip attractions in each zone**.** So, that is what the calibration means, calibration means, structure means identifying variables and creation of groups.

So, you know that how many categories, how many sales you have created. Then calibration part what we do? We actually establish the trip rates the way we want as per the created classification, what we do? In that case, we collect data, collect representative sample for each cell, then calculate the trip rate and put it in the table. So, then you can put it in the table, you can also explain or express them graphically and establish the trip rates.

(Refer Slide Time: 33:51)



So, if I have to summarize this lecture, I will say that in the previous lecture, we talk more about the structure of the cross classification techniques of the analysis, structure of the models for trip generation, a trip productions and trip attractions. In this case, we discussed more on the model calibration, that means how to develop to production rates and how to develop the trip attraction rate logically?

By collecting data, analyzing it and then summarizing the finding as per the structure of the model. And with this we are ready, now we have the trip rates, so in future when you forecast the land use. Then accordingly you can use these trip rates to get the zonal trip productions and zonal trip attractions, thank you so much.