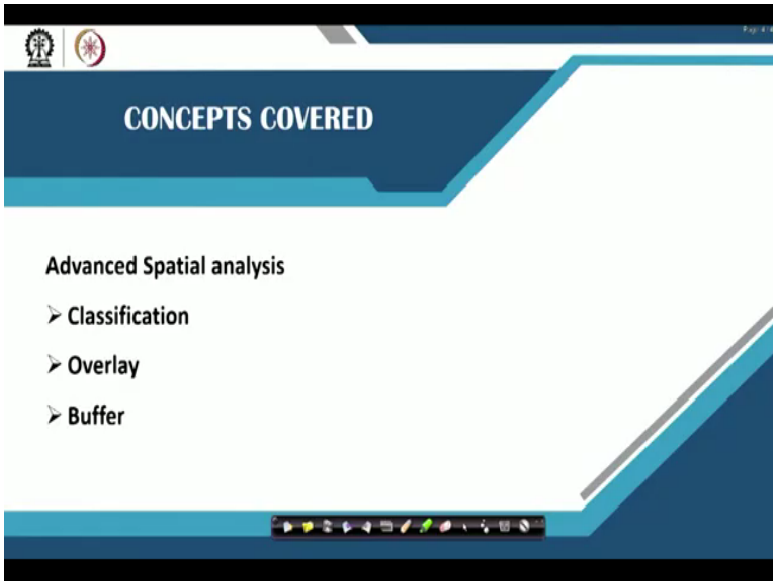
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur

Module-08
Lecture-39
Spatial Analysis (Continued)

Hello, namaste, welcome back to the course on geographic information system. So, in we have discussed about database, we have looked at SQL, in the previous one we looked at very basic spatial analysis and how the database, how the geometry data is stored, how the GIS data is stored and how we can understand different operations using arithmetic logical, statistical operators and Boolean operators.

So, once we have understood all of these things, now, let us look at some spatial analysis based on tools, okay. So, there are certain science behind these tools the way it has been understood. So we will look at some of this spatial analysis and how we use this tools in for various operations.

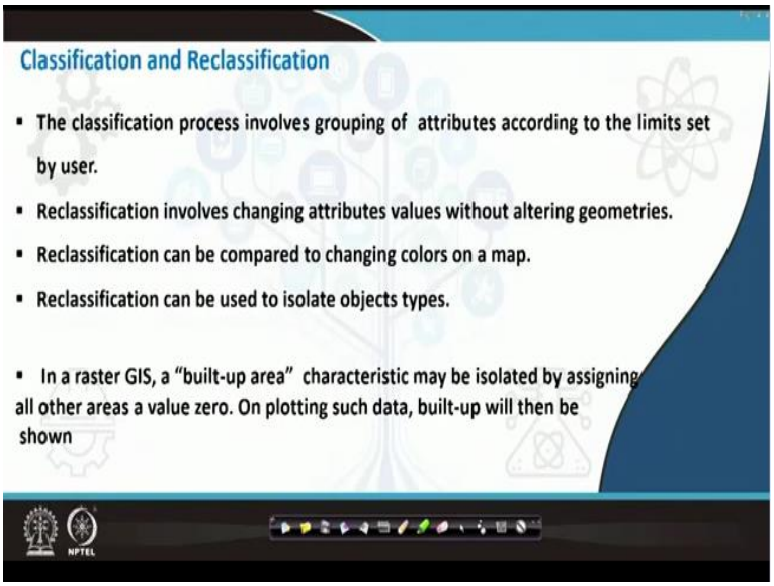
(Refer Slide Time: 01:10)



So, when we are looking at this we look at classification systems, why do we need classification. For example, you will have to you will put out the thematic maps once you have done that analysis, why do we need such analysis. Then we look at overlay, we look at buffer. So, how do

we use buffer, how do we use a function called as overlay. We will look at all of these aspects as we go ahead.

(Refer Slide Time: 01:33)



When we are looking at classification and reclassification, there is a first tool that we will be using, probably in one of my modules may be module 5, module 4, I spoke about the thematic maps. Yes. So when I say thematic maps, these are the ones which have a certain theme okay. If I am representing maybe the hydrological status of various districts in the state of Western Bengal okay.

So I may give some of the districts were probably the water availability rather than hydrological stage let me say it as water availability. Now water in certain districts had asked care in southern district, there is plenty of water. So, if I have to segregate these things, then I would segregate based on theme that is a color. So, if there is severe issues, then probably I would go with representing it with red.

If there are no issues, it is available in plenty, then I would probably represent in the form of green, so RBG format, so blue representing somewhere in between. So, when we are looking at this kind, this is called a thematic representation. So, similarly, you can use any number, any type of semantic representation, but, when we look at this NRC there is NRC theme guide of

how the representation has to be done in form of classification. And there is also the thematic representation style lines by survey of India.

So using those style lines we can basically do this kind of analysis. So, when we are looking at the same thematic maps, there comes your classification reclassification. When you are looking at classification process, it involves attributes according to the limits set by the user, let us say I have certain attributes, which I have to be represented in different forms. Visually it is the color that gives me more information as a user, okay.

Who may not understand what are the different things that are behind a particular value that has been arrived. For example, if you are looking at let us say urban density or the alpha density value that we are looking at, if alpha density has to be reached, there are certain we taken population, then population density, then we look at area. So you are basically user may not need information about the area, the size etc.

But he or she may look at the only the alpha density value. So, higher the alpha density will show in red and lower alpha density will show it in the form of green okay. So now, when you are looking at classification based on alpha density, so we will give a classification based on the values, we will give a range of values which is represented red, the range of values is representing green.

And this all of these scales are prebuilt in any of the GIS software when once you look at it, if you start looking at thematic mapping, it is much easier and the colors are vibrant in terms of display. So, reclassification involves changing attributes values without altering geometries, okay, for example classification involves the normal way of looking at both attributes and geometry which by the user has been set as a limit.

And when you are looking at declassification you are altering the attribute values but not geometry, reclassification can be compared to changing colors on a map okay. So reclassification can be used to isolate object types also okay. In the raster GIS built up area characteristics may

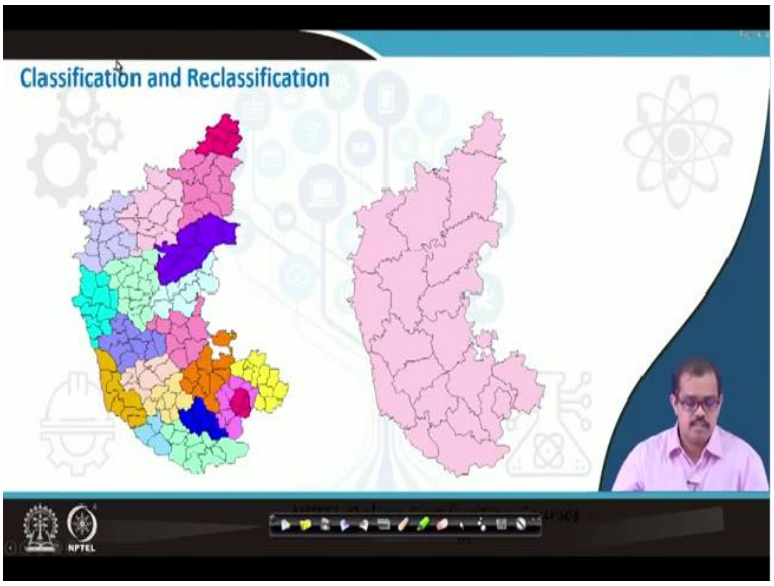
be isolated by assigning zero all our areas of value of zero which is also called as thresholding okay, what do you do is that if you want only the urban area to be given.

Let us say we have 4 land use types, urban, water, vegetation, others, right. So now what I will do is that I will give urban, I will just do something like this okay. Urban that is let us say I have urban as first class, second is vegetation is a second class, third is water and fourth is others, fine. Now what I will do is that I will give 1 = 1 okay, 2 = 0, 3 = 0, 4 = 0 or 1 = 1 star = null is a simple statement, okay.

So, now we just reclassified. Now I have only urban area, other areas are null right. So, this there is a spare we use the same thresholding in terms of classification in terms of raster analysis or another way of classification raster data is using maybe automating in terms of supervised or unsupervised classification. So, using a supervised classification a user gives the input in terms of signatures, which actually trains that classifier in order to in terms of statistical significance,

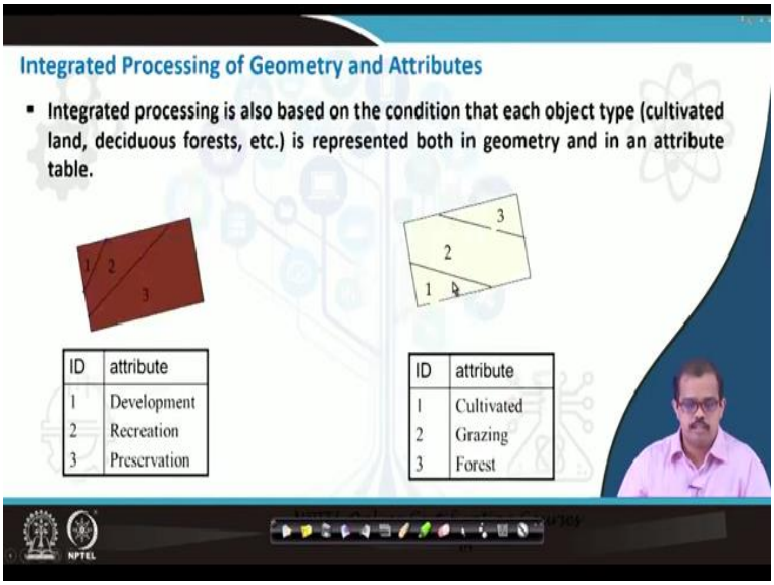
And also statistical mean, standard deviation, variance, covariance etc. Using all these values, there is a classification would happen there are a huge number of classification systems have been developed over a period of time if such classification and similarly, you have other types which is called land cover, land use etc. So, this is about your raster data, I would not get into all those classification techniques okay.

(Refer Slide Time: 07:13)



So just to give you an example this is what it means by reclassification. Now, here I have given classification for each of these, let us say it is each district has been given a different color, this map of Karnataka, so each district I have given a color okay for example Bangalore district is here which is given a particular color and similarly Mysore etc. are given different colors that have been represented here, okay.

(Refer Slide Time: 07:46)



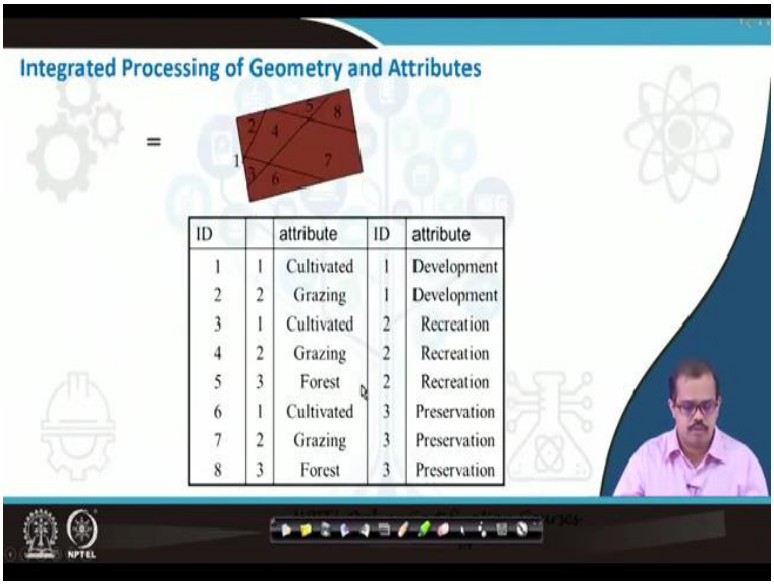
So, this what is mean by classification. Now, once we have understood classification, then how do you actually integrate a geometry and attributes for processing your data. So integrated processing is based on a condition that each object type is represented both in geometry and attributes. So, always when you are looking at integrated processing, you should have both

attribute as geometry also has to have attribute value also, if any one of this is missing, then your integrated processing may not happen in the way it used to happen.

When I say attribute for example, in this example, I would take up in the entire course of time, as for example, the ID is 123 and it is attribute is the as development, recreation and preservation, which means wherever one occurs, it is the land that is allocated development, wherever there is to that is being seen, that is where the recreation can happen, where there is 3 then it is for preservation.

So, similarly, there is an ID for one which is cultivated land. There is 2 for grazing. There is 3 for forest. So, these are 2 classifications that I would take it forward.

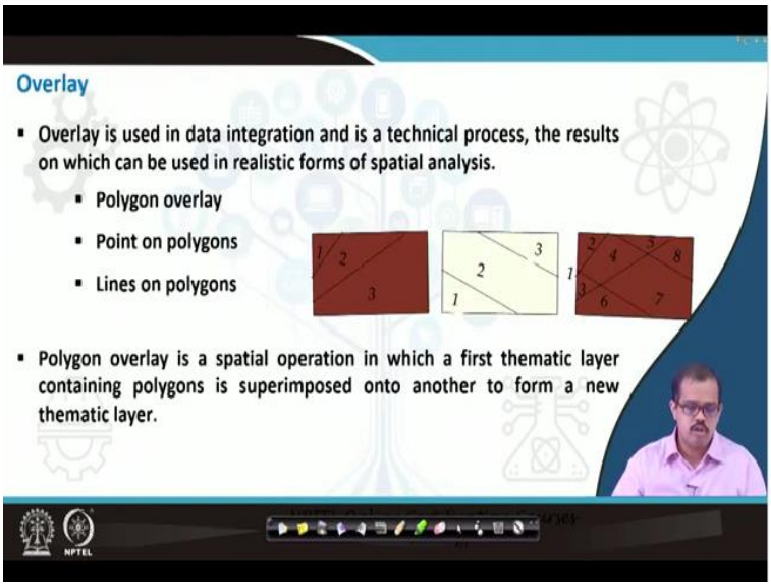
(Refer Slide Time: 09:03)



For example, when you say integrated processing, so, if I use both of these layers okay. So, attribute value is cultivated which is one here and ID here is development, attribute value that is true here as grazing and again ID value is one that is development. So, when it both of this can happen when you are actually using both of the layers, attribute value can be 3. And again ID can be 3 and there can be ID that is 2 which is again recreation.

ID can be 2 which is grazing here again that 2 is recreation. So, these are different combinations of different IDs and different values at in the toe, both the maps with the different geometry and different attribute values okay.

(Refer Slide Time: 09:50)



So that I use. Now the first type of integrated processing is overlay. Now when you are looking at overlay, overlay is a kind of data integration and technical process that results in realistic forms of spatial analysis. That is for example, you have one layer I have shown this example much before also when we are we were discussing about thematic maps etc. So, you have one layer down, one layer on the top another layer on top.

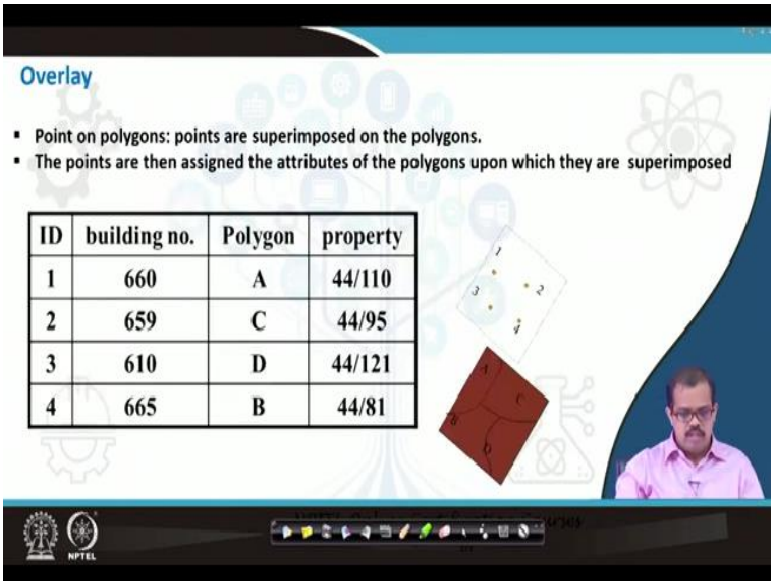
So, now we want to analyze our effects of all of these 3 layers okay. So that is nothing but overlay, you are putting one on the other and trying to understand where for example, here the development can happen okay. So, there are normally 3 kinds of overlay one is the polygon overlay another one is a point on polygons. So when I say point on polygon, it has point that is overlaid on polygon for the analysis.

And I say polygon overlay it is polygon on polygon analysis, there is lines on polygon which means lines is put on polygons for analysis. So all of these can happen when you are looking at overlay function, polygon overlays a special operation in which first thematic layer contains polygons is superimposed into another to form a new thematic layer okay. So now when I say

polygon overlay, so you have all as polygon so you have another thematic polygon, which can be used.

When I am saying polygon here, it is not only one polygon which can be overlaid, any number of polygons until the data that you have, overlay use functions to create some I mean, extremely important outputs, okay.

(Refer Slide Time: 11:33)

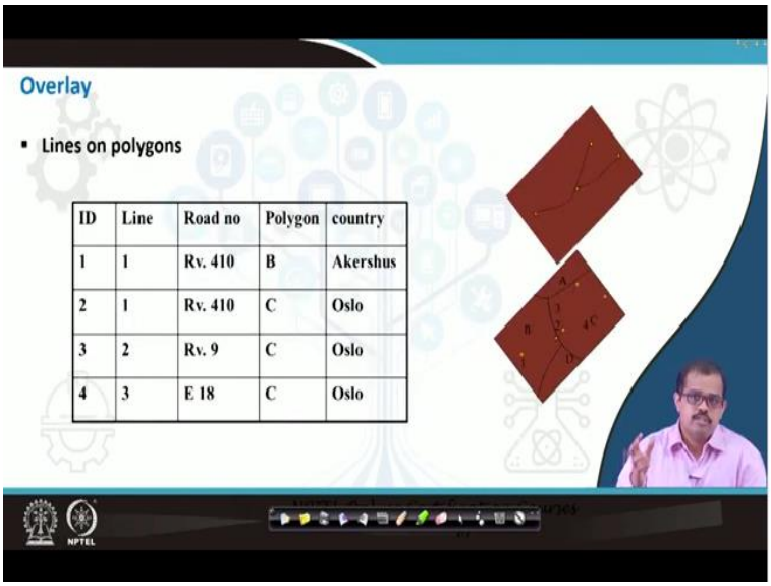


So, when you are looking at point on polygons, here the points are super imposed on polygons. For example, for you to give an example here, there are 4 properties that is therefore piece of property that A B C D and we have 4 IDs which are representing 4 building numbers. So each of this if there is a building number in property A. So, it has certain property number that is 44 by 110 that is represented by 1.

So, now if you want to look at what is if you have allotted to someone and it has to be developed as a property. So, if someone wants to analyze what kind of soil is there, what kind of I mean how the foundation has to be laid etc. So, just using a soil layer on one of and other geographical layers that have been already shown here. So use these 3 layers to extract what is a kind of soil layer in that particular region.

Then once you have extracted it then probably you can find out how the foundation has to be laid that is up to the that civil engineers who are actually looking at that part okay.

(Refer Slide Time: 12:50)

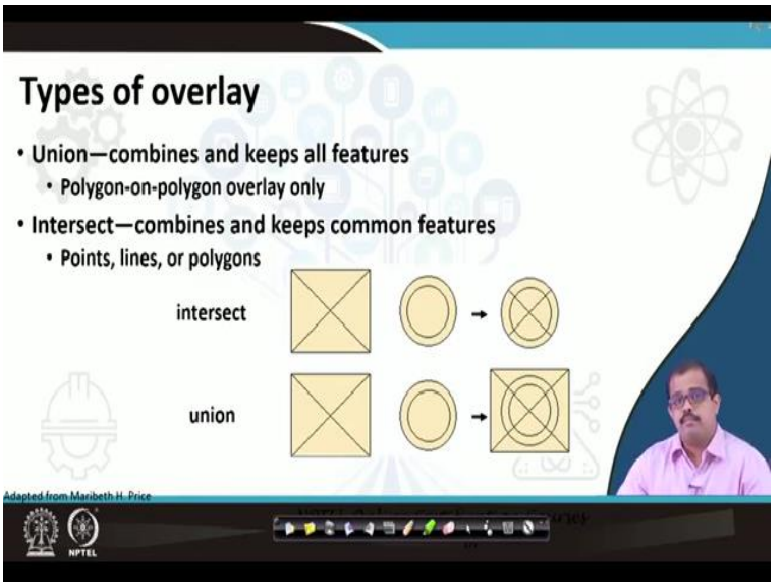


So they can extract information rather than going testing the soil in the ground say if the most of these layers are now readily available online. If you go to Bhuvan you can get soil layers etc. so or and many of this the state remote sensing centers I will be able to give you most of the layers that I am actually explaining. So far if you get those layers so probably you can use it for any kind of analysis.

Then you have lines on polygons, let us say that I have this is a road line, these are the different polygons. Now I have to find out this is the one road line that is now proposed. Now I will have to find out this how where does this road goes which are those areas that are affected. If that area is a particular protected forest or it is an protected area or it is an commercial area or it is an any of the houses etc.

So if I have to find out I will just overlay this lines on those polygons. Once you have put these lines on those polygons, I can find out where this is going, what is the effect that it may have once the road is being built, what is the size of the road and if there is extension what may happen in the future. So, such kind of analysis can be done when your overlay lines on this. So, that is a different way of thinking that usage of overlay analysis.

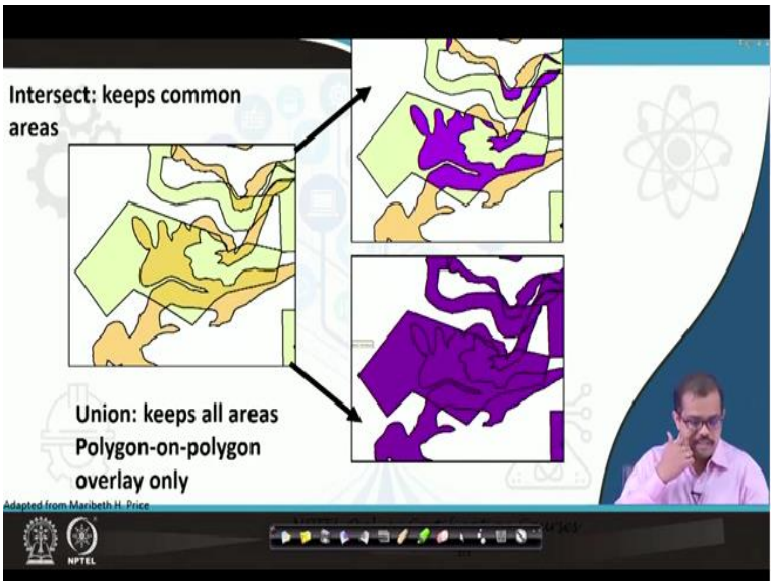
(Refer Slide Time: 14:13)



So, as I said overlay can be many types okay you can use union, you can use intersection. So, union combines and keeps all the features that is polygon on polygon overlay only. So, if this is an polygon and this is a polygon it actually in union has combines both of these normally polygon on polygon overlay only does this, intersection it combines and keeps only the common features. For example, if you look at this and this the common feature is only this.

So, both of this the other areas are cropped and kept, a very good example is that when you want to actually create the your study region based on the boundary, so you crop the other regions and put it out, that is where the overlay helps and it is a very good tool in your GI software to actually look at various aspects.

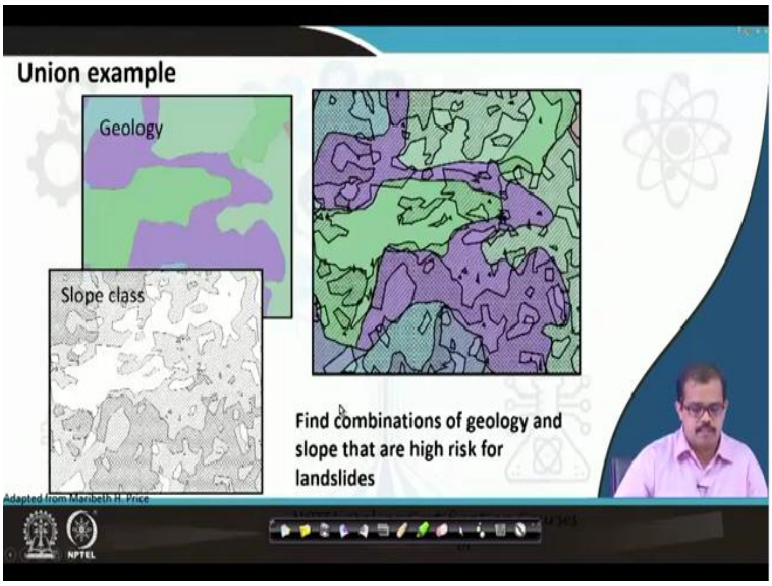
(Refer Slide Time: 15:07)



So, this is one example that from a Maribeth from Maribeth who showed how an intersect tool can help in various analysis. For example, this is one area if you intersect an union tool, so, if this is the region that we are trying to analyze if there was some error or calamity in this particular region. So when you look at intersect what it basically does, it gives you only the region which is affected which is this okay with 2 datasets.

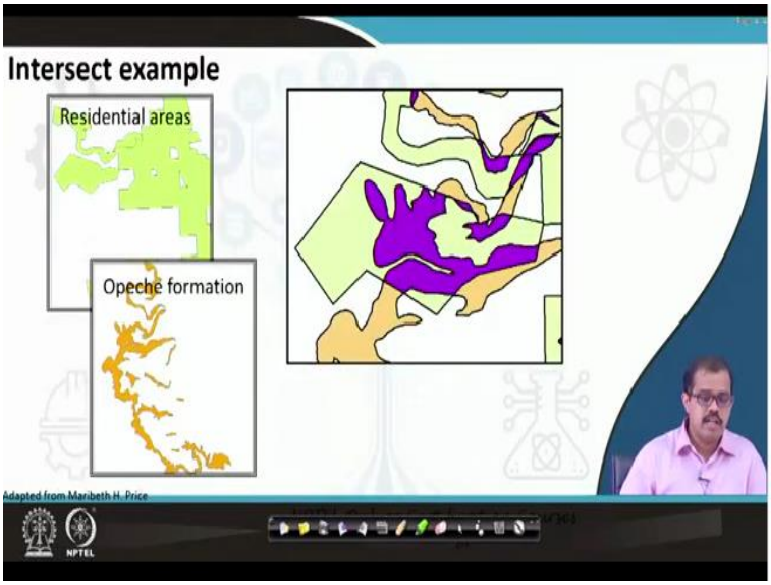
If you look at union both of this will be accommodated here, okay. So, now if you want to look at only the regions that are there here, then you will probably use intersect which is affected by certain phenomena, but if you want to look at the entire region including that region that is affected, then you use union tool that is what it is basically representing.

(Refer Slide Time: 16:08)



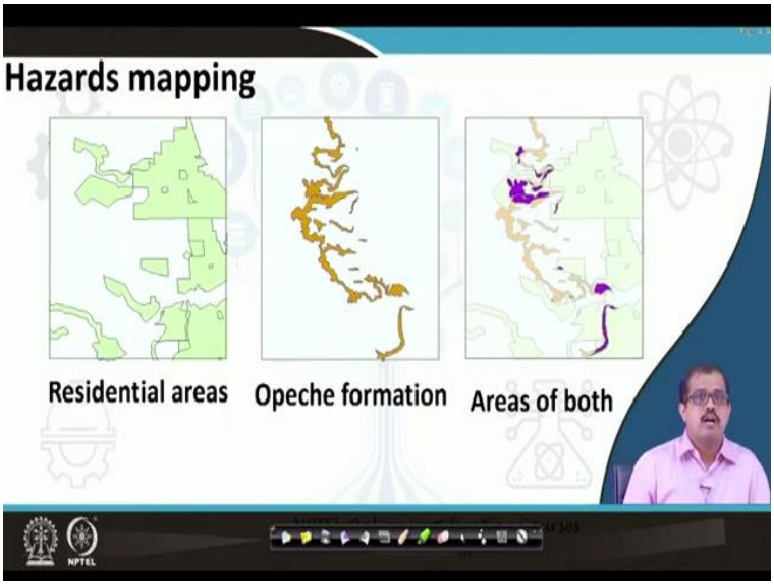
Another example of union is this you have a geology map, you have a slope map. So, if you do union which means to say you get both of these on together. Now, with this you can look at a where there is a high risk of landslides. So, you use both slopes and geology map so, you will say where the landslides are high.

(Refer Slide Time: 16:29)



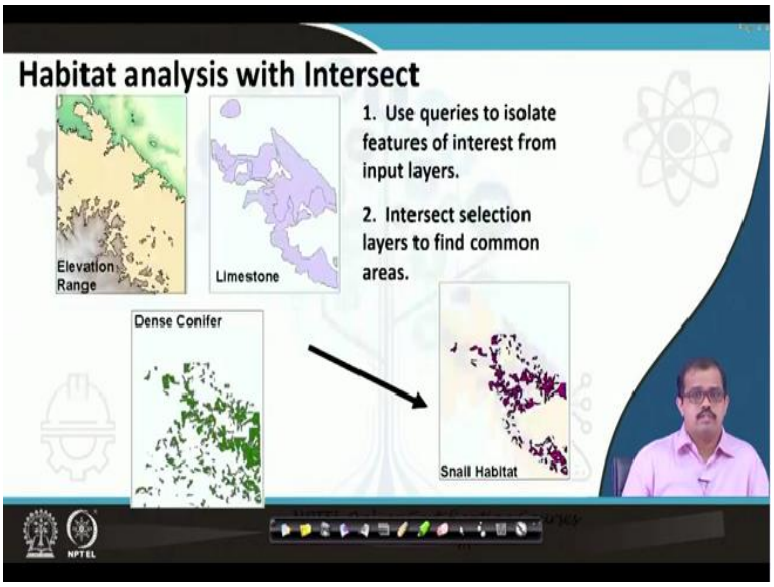
Then intersect for example, you have a residential area here you have an opeche formation that is happening it may affect the residential area. So, where it affects so just intersect you will know where it may affect in a larger scale. So, this is a area that is what the example that I showed previously okay.

(Refer Slide Time: 16:49)



So, hazards mapping is another way of looking at this you can apply this on hazards mapping. These are some examples as adapted from Professor Price. So, if you can see this as some of those, which you can see.

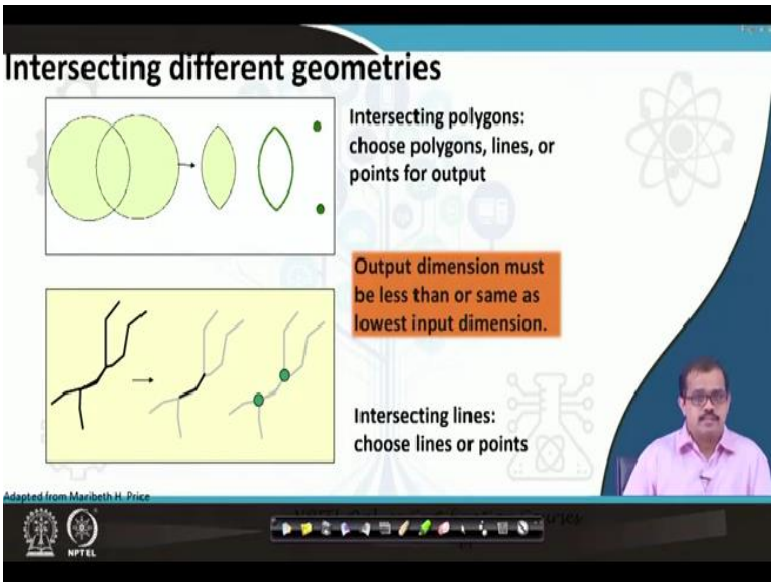
(Refer Slide Time: 17:05)



And habitat analysis this, this is an extensive usage now in terms of how do you look at as habitat. Now here is a snail habitat which we are looking at you have elevations . These are the ones which are normally in a bit of lower elevations and more in terms of dense conifer forest, and where basically there is limestone. So, what we will do is that we will look at elevation, we will look at limestone and we will look at conference forest.

Now, with intersecting all 3 we know that the where the elevation ranges medium and we know that where the limestone is located and where the conifers are all 3 are located. Once there is located that is a probable area where the snails are located. So once you locate those areas, you can find out where the real I mean real locations on the soil. That is why intersect can be used extremely as a useful tool in analysis.

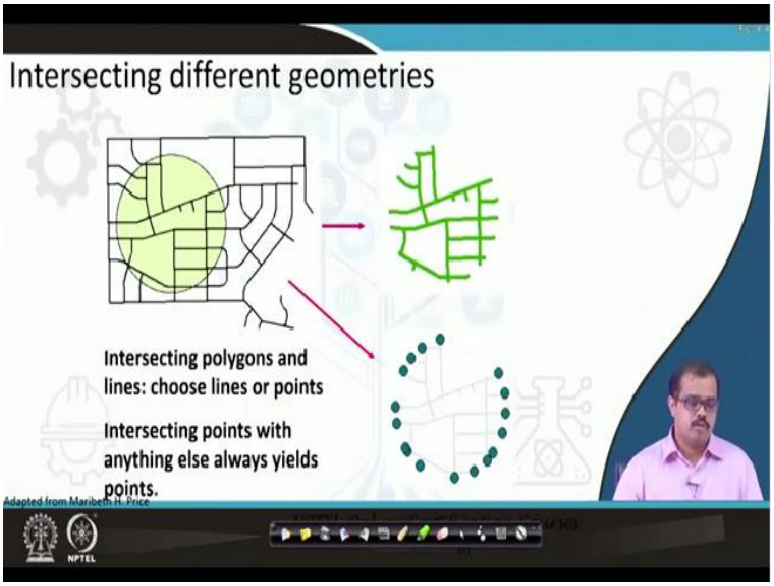
(Refer Slide Time: 18:05)



Intersecting using different geometries. Yes, you can do it. So, there are 2 different geometries that you can see whether it has an intersecting polygons choose polygons lines are points for outputs, how do you want to get your outputs. So, output dimension must be less than our same as the lowest input dimension So, that you should keep in mind So, it will be the same. Then intersecting lines choose lines or points.

If you want output to be in a form of points or you want to have output in the form of lines, how do you want the output. So, look at those ways of we are providing output.

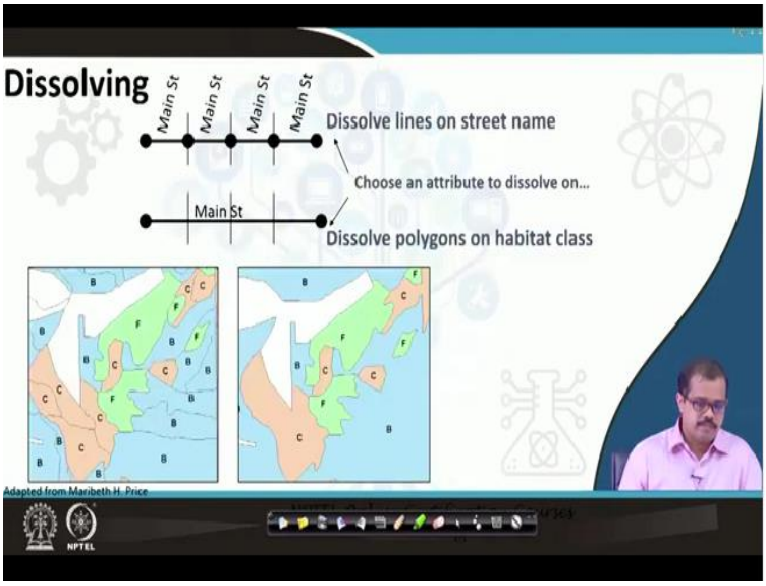
(Refer Slide Time: 18:43)



Then intersecting different geometries. So, this is very important for applications, especially in the urban data. Okay, for example, there is one geometry here and there is an another geometry. So, these intersecting polygons you should choose either it is lines or polygons, why I say this is that you should be extremely careful on what data you are trying to export, what is the report that is being given to the user.

Whether you are giving a report on how denser the street is or you are giving a report on where there are intersections in that particular street network, so intersecting points with anything else always eats points okay, intersecting points with anything else it is only points okay. Intersecting lines can yield both line and points. Intersecting polygons can lead polygon and lines okay. So, having you should be able to understand interpreted as a reverse logic of how the vector model is built.

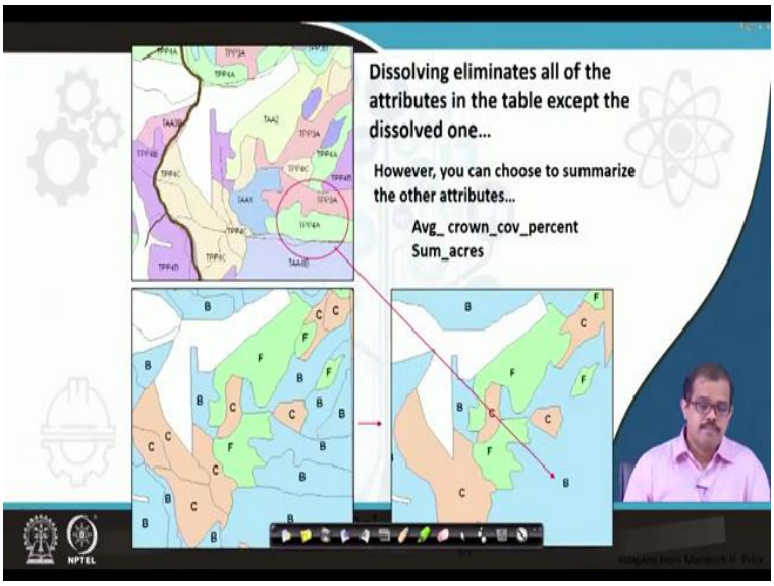
(Refer Slide Time: 19:49)



Then the next tool is dissolving very, very important tool, if you are trying to represent any vector data. For example, let us say that these are different maybe I can say different forms or different fields that are there in this particular vector data okay. So, when this different fields based on this fields each of these geometries are shared. And in case let us say that I want to dissolve all of these into single.

For example, where anything that is B put into single, anything that is A or any other point F is put into single, so it just using dissolve you can put this into single line and if there are many intersections like this, and you want to represent a main street, just using a dissolve can put it into a single line segment with only 2 segments the end and the tail, which is for representing the entire main street. So this is how a dissolve tool can be used and if someone is using RGIS I am sure that he or she would have used this tool extensively okay.

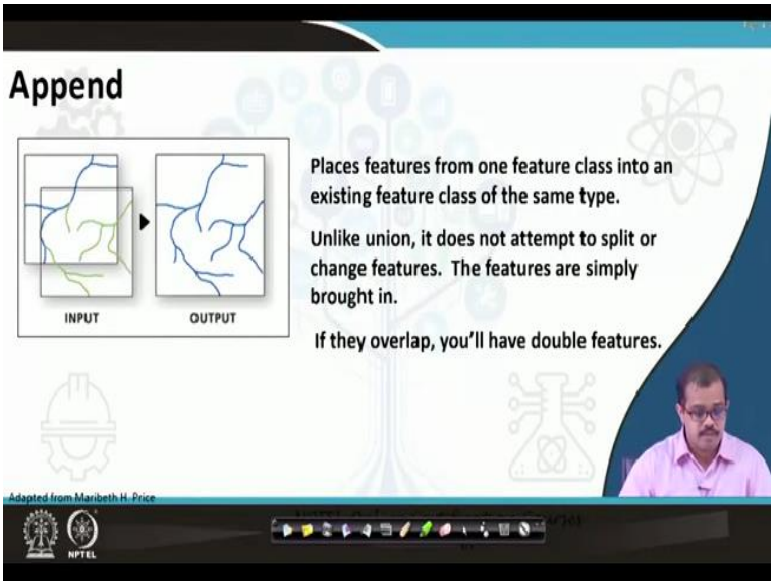
(Refer Slide Time: 21:06)



So, dissolving eliminates all the attributes in the table except the dissolved one. Okay please keep this in mind, this is where you lose the data that is what I said if you give an exclusive access to the user, if they use dissolve and save that data all your data is gone. So look at where he or she saves you she can imported use a different form to do it and save that data instead of saving it the main data. That is how you have to create your entire database okay.

So and you can choose a summarize other attributes which are related as if it has not dissolved. So whatever it has dissolved, you lose those attribute data and only the things that are not dissolved remain as it is.

(Refer Slide Time: 21:55)

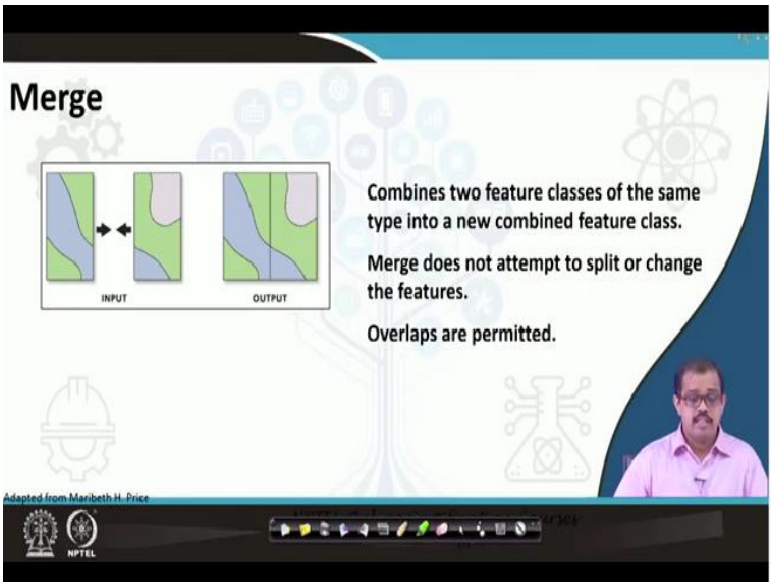


There is something called has append. For example, places with features and a feature class into existing feature class of the same type okay. So for example if there is one vector data that is like this, okay that is represented in blue, there is another input which is in this, if you want to append both of these combine both of the attribute data, combine both of the geometry data and then place it in the output like this.

It does it by default using the append button. Append. unlike union it does not attempt to split or change features. The features have simply bought it, it is like copy paste, instead of union it looks at I mean using both if there is certainly common it is so use that but here it is unlike that it is just a copy paste thing. Okay copies from one layer to the other layer and paste there that is it. Whether it is correct, whether it is not topology is correct or not, it does not look at it.

If they overlap, you will have double features in your database. So look at the status one of the issues in terms of append.

(Refer Slide Time: 23:12)

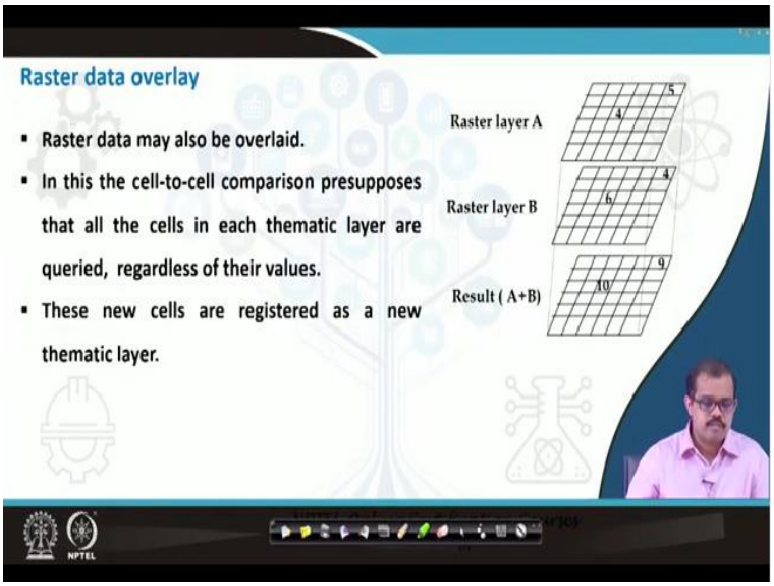


Then there is merge. So for example, you have, let us say you have created 2 layers. And now this happens normally in your satellite data, you have 2 different raster datas. So both representing, if for example, if I am taking Hyderabad data, Hyderabad data maybe in 2 or 3 or 4 images of with respect to a particular spatial resolution. So let us say 4, 2 images that I have, now I have classified both of these images for different issues.

So now I want to combine both to be represented at a single map. So what you do input both of map A map B and then use merge. So what it does, it looks at the common boundaries, if there is an overlap. Okay, for example, this is something this image is covering this area and this image is covering. So the top hand the right hand is actually covering more area than the previous on the left hand.

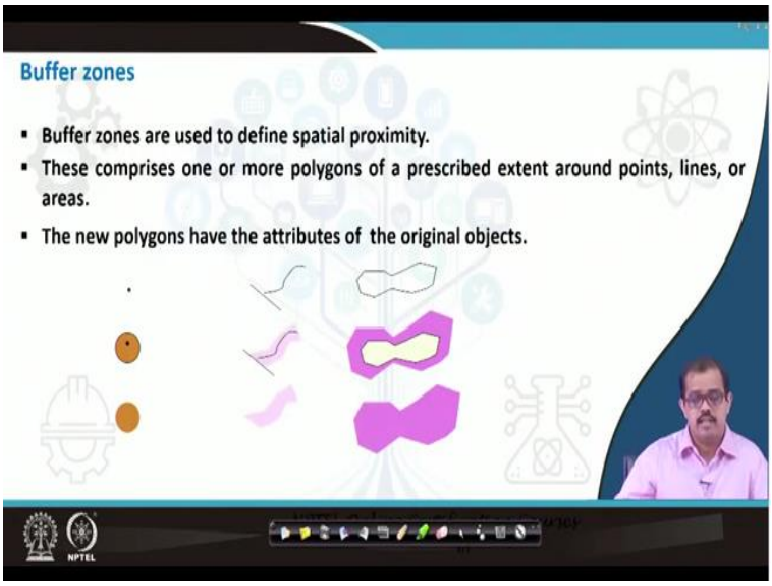
So, what it does it sits on it like this. So, it matches its edges or its latitude and longitude or it i points on these edges. So, that is how the merge happens. So, normally overlaps are always permitted and overlaps are considered to be correct in terms of merging data.

(Refer Slide Time: 24:28)



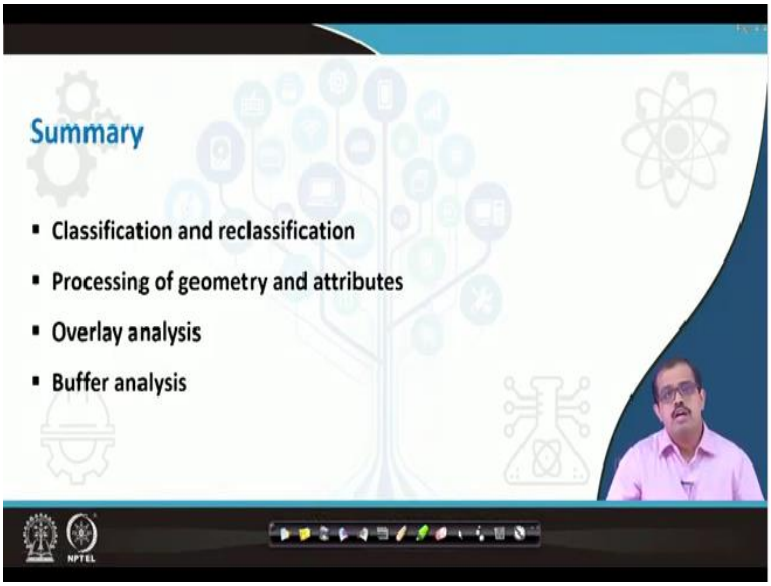
So that the data is not lost. So then you have I spoke about the raster data overlay. So it is cell to cell comparison, which is more of cells based on thematic layer, there are new cells are registered as a new thematic layer whenever you create new cells normally it is a new thematic layer.

(Refer Slide Time: 24:45)



Then the concept of buffer zones, okay. Buffer zones are used to define spatial geometry. So these areas these comprises one or more polygons of a prescribed extent around the points, lines or areas. The new polygons have an attributes of original objects, when you create a buffer, the attribute is of the original objects. Oh, for example, if this is the area that I have taken, so, you have a buffer zone that is linked around it. So, they have the original objects that is linked to it.

(Refer Slide Time: 25:19)



So to summarize this particular class we looked at both classification and reclassification. So, we said the classification is both more is equivalent to thematic representation, we are trying to classify it based on certain attributes that are there in the data set. Then, we looked at how do we process a geometry and attributes. We looked at overlay analysis, overlay analysis is extremely

useful in terms of using it and we said in when you are looking at any of these geometry and attributes, we are looking at union and intersect functions.

Union gives you union of both of the geometries that we take but when you are looking at intersected as the common geometry and we have polygon, polygon intersection, polygon line intersection or polygon point intersection where the basis of polygon part and when you are looking at buffer analysis, we are actually looking at the attribute which was common with the same area where we are looking at the buffer.

Buffer is some as I mean for example, if you are looking at urban analysis tool, if you have already done a land use classification for a certain region, okay. Now, let us say that you should understand what is actually the sprawl area that is out of this urban area, it is simply done as simple as this. So, you can start buffer, so, that region is now you can do the land use classification validated. Now, you know, where the spurt of the urban area has happened, okay.

Extracted look at different networks, different regions of how it is interacting and why the growth has happened. And if the growth has happened logically then you can say these are the reasons or you can use different metrics to quantify that growth and say, how the sprawl has happened over a period of time. So, that is by the buffer analysis may be extremely useful and buffer is normal, I mean for various reasons that buffer analysis queue.

And probably eventually start exploiting will understand why it is being used. Okay, so this is what is for this class. So, I will meet you in the next class. So, we will look at some of the operations in QGIS will go back to QGIS, I would like to take it up in more detail about certain operations that will do it and will perform various tasks in those operations when we come back. Until then, thank you very much.