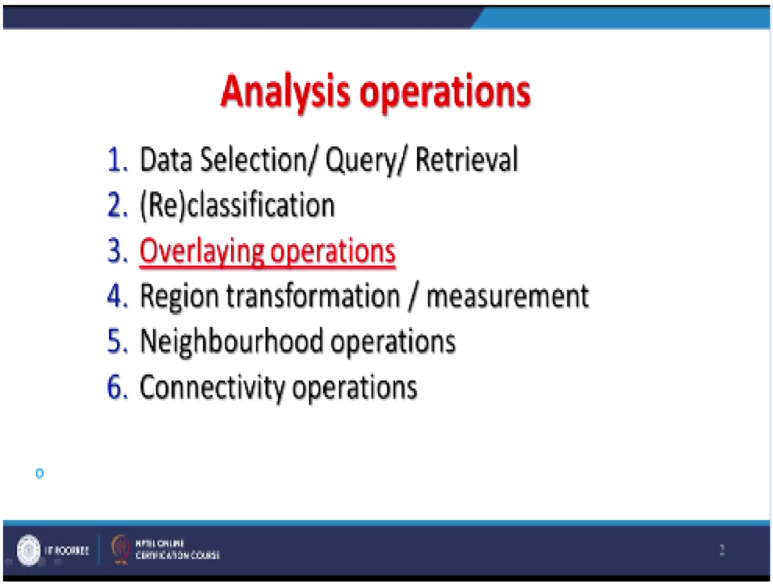
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology - Roorkee

Lecture - 21
GIS Analysis - 02

Hello everyone! and welcome to this discussion on GIS analysis. This is part 2. Few more parts, we will be having for this discussion because as said in the previous lecture that this is basically the heart or key part of GIS operations. So, after the build of a spatial GIS database, we try to do as much as analysis possible whatever the data which is available to us. And of course, as per project requirements and as per the user given specifications, so, this kind of analysis.

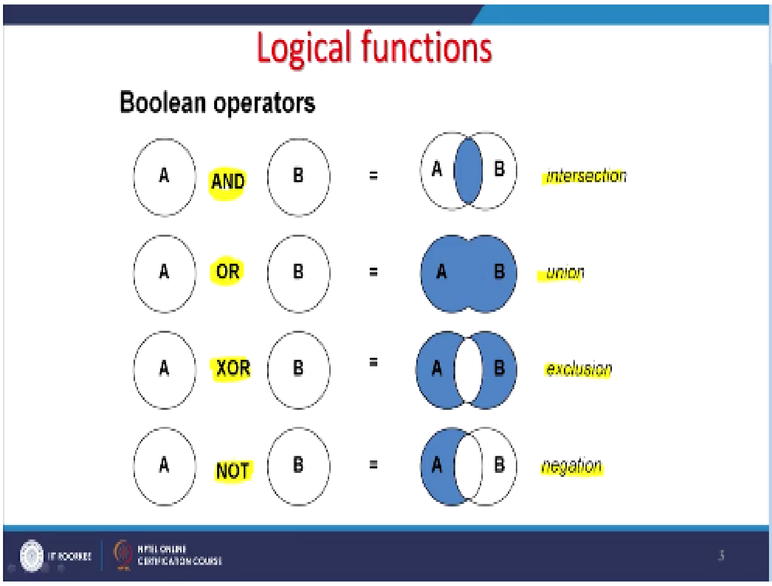
Now in this second part of GIS analysis, we are going to discuss these overlaying operations. So, the first 2 parts; data selection/query/retrieval, we have already discussed and reclassification in the previous discussion.

(Refer Slide Time: 01:25)



Now in this, we will be focusing mainly on the overlaying operations. The basis of overlaying operations is set theory or Boolean logics. If you recall the veins diagram, that is what is the basis but there, things were much simpler like in these Boolean operators or in Boolean logic or these functions, we used to have A intersection B, like A union B, A exclusion B and A negation B.

(Refer Slide Time: 01:49)



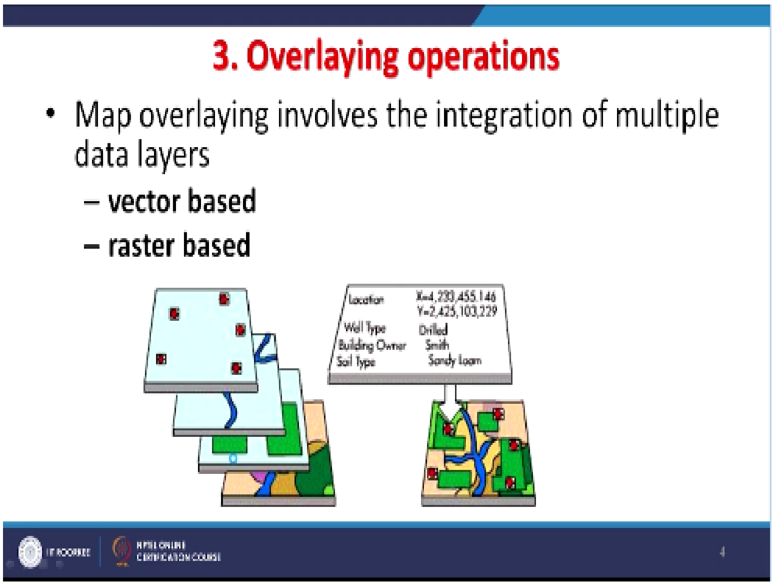
So, when we were introduced the set theory in our 10-12 mathematics then we are introduced to very simple kind of logical functions or this veins diagram. But in GIS, we can perform very complicated one because we are not handling 2 layers having just single polygon each but we might be having multiple layers and each layer might be having larger number of polygons.

But still fundamentally, it would remain same that when we do A intersection B then we are using this operator AND. Or A union B then we are using this OR. XOR that means exclusion and not including or negation we are using. So, these operators and in combination that means single, we can use this operator or multiple operators together, we can perform analysis on our maps.

And these logical functions or these overlaying operations can be performed not only on the vector data but also on the raster data. So, in case of vector data, mainly on polygons but anyway, you can also perform on polyline and points as well. But mainly, it becomes little complicated in case of polygon theme. Whereas in raster; no issue because if all layers, all themes are geo-reference.

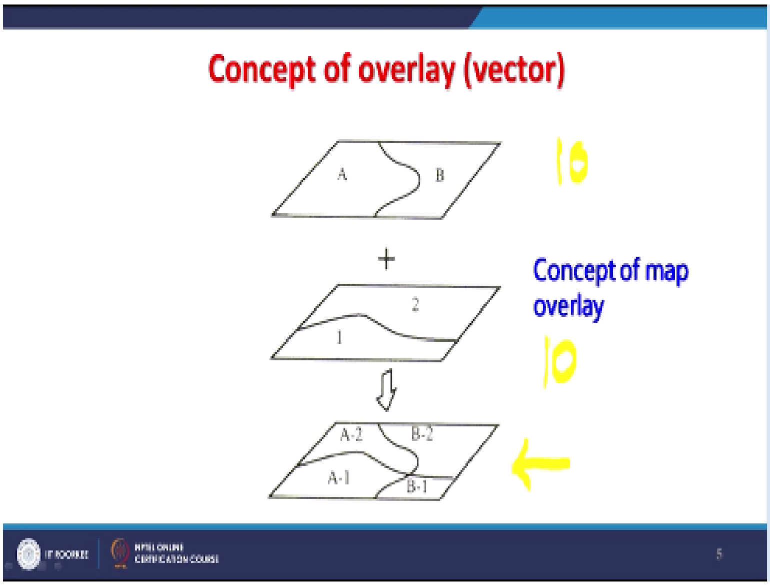
And they may be belonging to a common geographic area then we can perform these Boolean operators.

(Refer Slide Time: 03:46)



So, for using, there are multiple layers we can involve and in combination, we can also use vector and raster layer also together. So, we can perform these overlaying operations on all these.

(Refer Slide Time: 03:59)

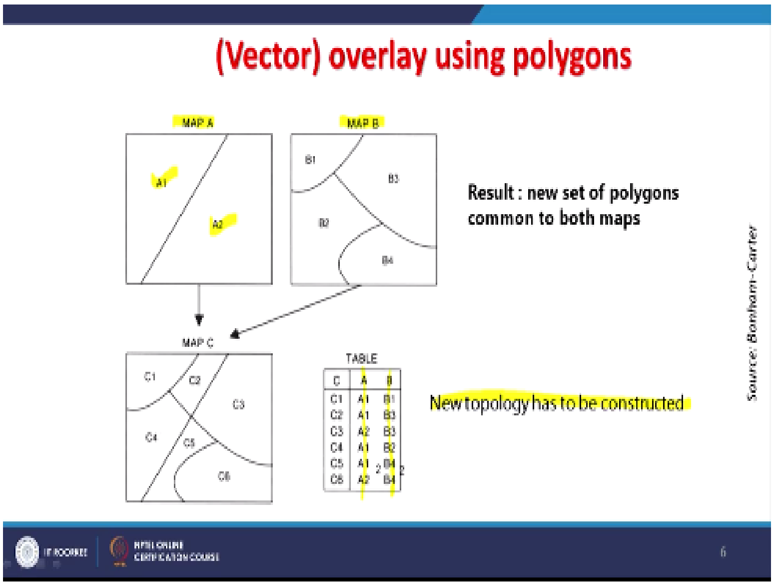


Basically, overlaying operations, here one very simple example is given. On top, we are having one map belonging to the same geographic extent having 2 polygons; A and B. Now for the same area, we are having another polygon map which is having 2 different units 1 and 2. So these polygons 1 and 2 are completely having different shape and size then polygon A and B of the top map.

Now, when we you know intersected these 2; this is what the output is coming. And now in output, we are getting basically 4 polygons. So, why I am discussing this is that when in the real GIS operations, each layer may have 10s of polygons. So, if I am having a 10 polygon here and 10 polygons here, you may end up even with 100 polygons depending on the shape of polygons.

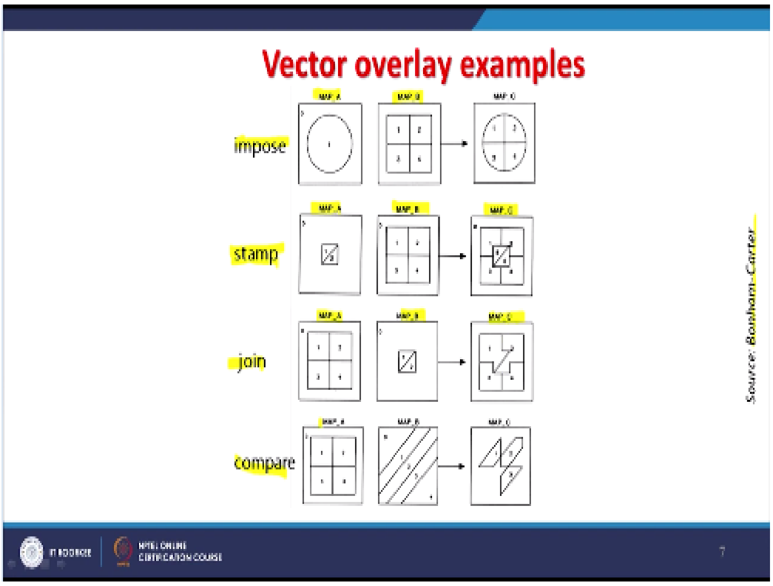
So now, that map may become a little complicated. So, if it becomes complicated and if it is possible to reclassify like we took example in the previous discussion of reclassification then we can reclassify and reduce number of classes and merge some polygons only for our display purpose. So in that way, we can perform such analysis.

(Refer Slide Time: 05:38)



Another example is here that one map say A is having just 2 units and map B is having 4 unit. Now, when we you know intersect with these 2; this is what the result is. So, instead of 2 and 4, now we are having 6 units and new topology would be constructed and that topology which will have extra tabular data or extra field in our attribute like A1, A2 are coming here and here B1, B2, B3 will come as a separate one.

(Refer Slide Time: 06:15)

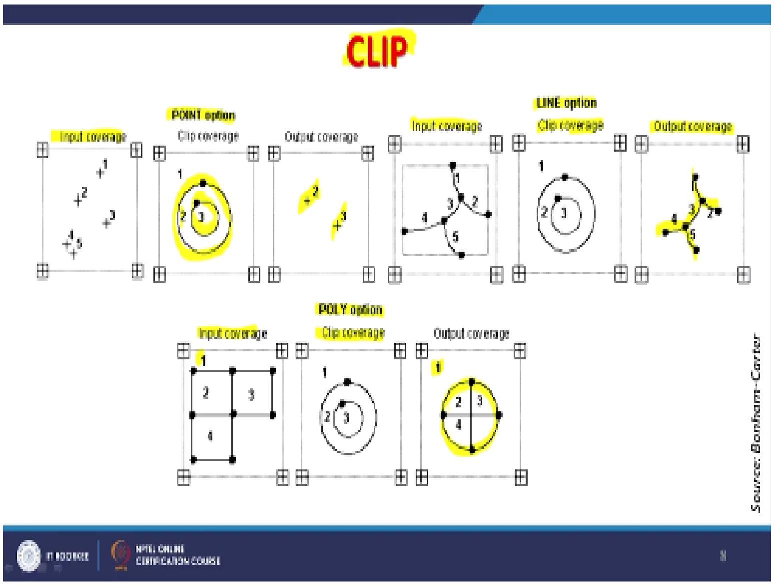


Similarly, in vector overlay operation, lot of possibilities are there like one possibility is about the impose means basically, it is intersecting. So, impose means map A is here, map B is here, output I am getting like this. And why I am showing these terms here because in many softwares or in many literatures like a very famous book on GIS by Bonham and Carter, they have used these terms.

So, many softwares may use these terms but basic operations will remain only those 4; that Intersect, Union, Not and you know Exclusion. So that way, basic operations are same but maybe different terms are used. Now here when we go for a stamping basically, it is combining these 2 maps; A and B and creating a union here.

Similarly, a join, so this is also joining and merging few polygons here like as you can see in this example in map C. Again, compare A and B; when they are compared, the common areas are selected, rest are excluded. So, this is exclusion possibilities there.

(Refer Slide Time: 07:39)

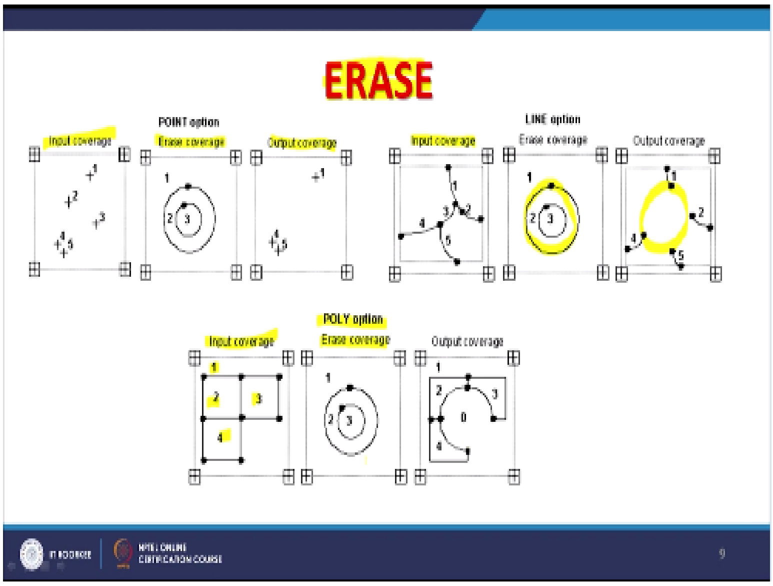


Now, this clip or extract; it is also called extract in many GIS software like in ArcGIS, retrieving data based on basically a polygon boundary. So here, this is what is. If it is in case of point data, this is example that I am having these you know polygons or circles and these are my input coverage or point data so when I overlay these 2, I get only selection of these 2 points.

That means I am basically clipping or extracting only those points which are falling below these 2 circles. Or you know like this point 2 and 3; point 2 and 3 are falling in between these 2 circles. Similarly, I might be having a line or polyline theme. This is my input coverage where I am having many polylines and clip coverage is the same as in case of point theme.

And when I create this intersection, this is how or I get this clip or extraction of polyline only for that circular region. Similarly, if I am having my input as a polygon map where just 3 units are there and one is of course, outside one. Then if I use the same clip coverage to extract then I get something like this as shown here that for this circle, I get this.

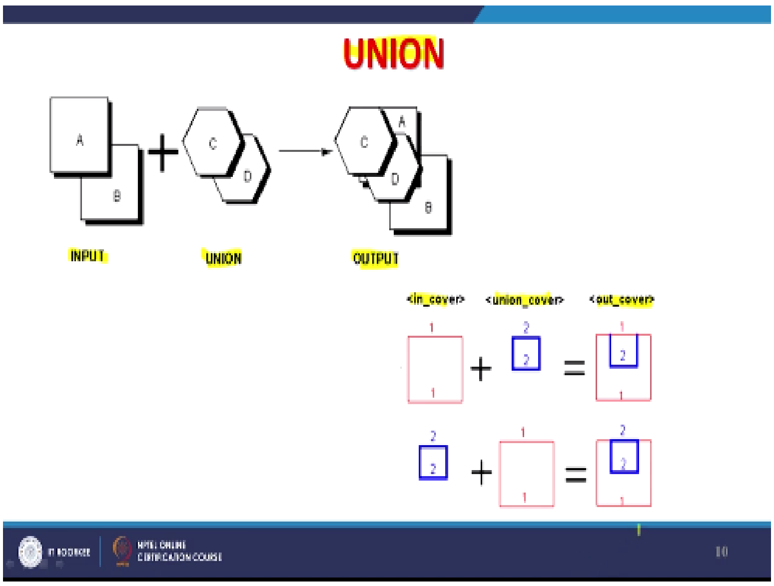
(Refer Slide Time: 09:27)



So, all kinds of possibilities in case of vector data are like this. Similarly, different term is used here is Erase. So, basically it is excluding those areas which are common. So, this is my input coverage, this is erase coverage and this is my output. So, whatever is falling within these circles, is excluded in my output. Similarly for polyline, if I use this erase coverage that now the things which were or polylines which were falling below these circles, have been excluded.

Same with my polygon coverage. My erase coverage is the same in these examples. Input coverage is a polygon having 4 units; 1, 2, 3, 4 and then now, it is excluded. So, the area below circles have been excluded or polygons have been excluded.

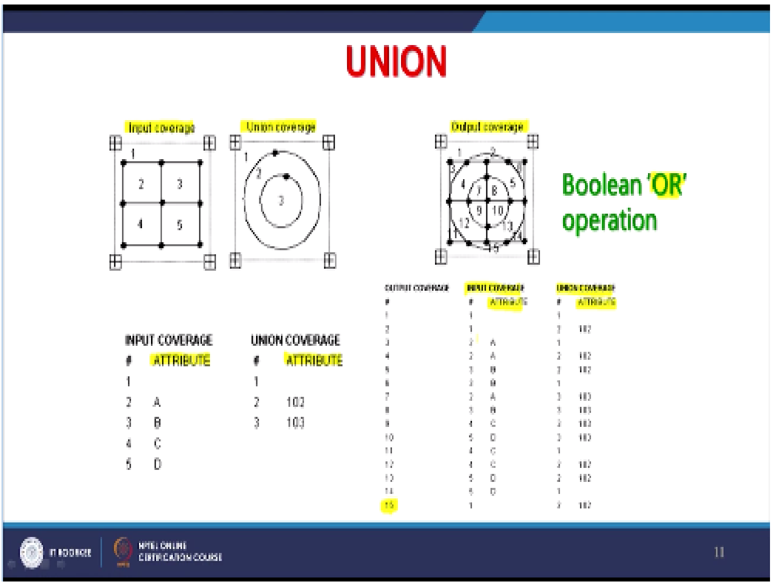
(Refer Slide Time: 10:24)



So, we can have an extraction that is intersection, exclusion, union. Union example is also there. It is combining all. This is my input map having 2 you know, rectangular areas. Another layer which is having hexagonal and when I create a union, I get everything together. In a more realistic one like input cover, union cover and output; something like that I can have.

If I am having you know sequence of themes is different but output is going to be different. Here, the red large box is first input coverage, union coverage is different, though output is like this. Similarly, if input coverage is small blue box or a square, red square is my union coverage but output is going to be the same. So, this was the union example.

(Refer Slide Time: 11:22)



In more, what would happen to this union operator or OR operator. AND for intersection, this OR operator my attribute table will become richer and richer because it is getting information from second layer also. So, this is my union coverage. This is my input coverage where I am having 6 polygons and when I create the union, I am having now large number of polygons in this particular example now. Total number of polygons have become 15.

And of course, input coverage will have an attribute field. Similarly, a union coverage will also carry its attribute field. So originally, these 2 input coverages; input coverage and union coverage have single attributes in this particular example. Now my output coverage, I am having now 2 fields; One attribute is coming from map 1 input coverage, another attribute is coming from my union coverage.

So likewise, I can have not only 2 but I can union many maps having different polygons together and can create one. The only problem is that whatever the output is coming, we should be able to understand and utilise that one. Computer or mathematics point of view or from GIS operations point of view, all complicated these logical operations, nested operations are definitely possible.

(Refer Slide Time: 13:07)

Another example of intersect is that when we use this Boolean or AND operator that when 2 coverages are overlaid, only portion of the input coverage that falls inside the intersect coverage will remain in the output coverage that we have seen these examples. Similarly, while intersect, coverage must be a polygon. This is what the requirement or this clip part that input coverage can be a line, polygon or point coverage.

If you recall the discussion then we have been discussing about the point data. At that time, I mentioned but today again, I will mention that sometimes we get a very big database. And I do not want that my system should be occupied with that database when my work involves only on a small area. For example, I may get all the villages data from 250000 scale toposheets through some portals on the internet.

Now I am focusing only in the villages of say, Uttarakhand. So, in my hard disk, why I should keep the entire database? So, what I would do? I will use a polygon that means for intersect coverage, I will use a polygon which will have the boundary of Uttarakhand and I will do these overlaying operations and operator and extract only villages belongs to the Uttarakhand.

And what would happen that when I use on my point theme; when I use this political boundary of Uttarakhand then output will have only those selected points in my input theme which belongs to the Uttarakhand geographically. Now, what I can do? I can save that selection as a separate file and can remove my entire database from my system.

So, unnecessary I will not keep because for a particular organisation; suppose that organisation is working for Uttarakhand, why to keep the villages data for entire India? You may have if you are having liberty or a lot of spaces available on your hard disk. You can keep as a separate but in the day-to-day operations, when I am focusing only say on

Uttarakhand, why I should all the time display or bring the total file of entire India villages. Because there are 1000s of lakhs of villages. So, why to keep that data?

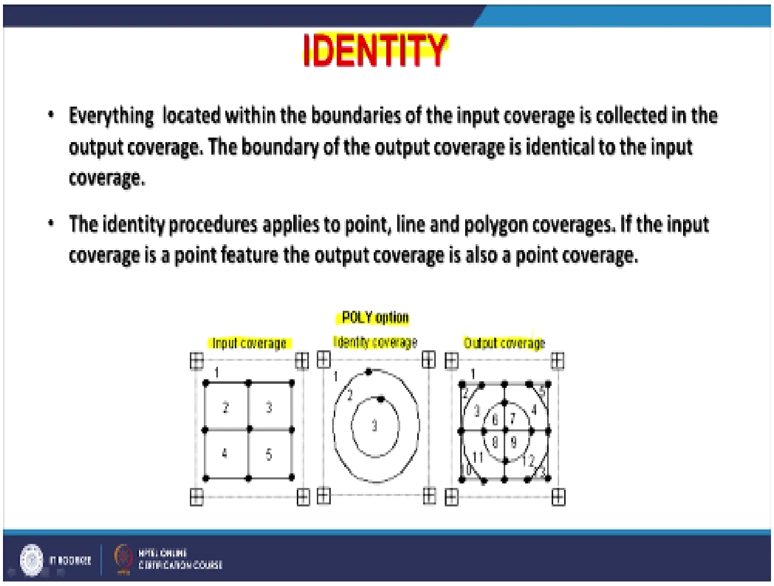
Similarly for a drainage network, I may get some model drainage available on some portals for entire Asia. Now, I am focusing mainly on Ganga basin. So, why I should keep in my hard disk the entire drainage network of entire Asia which has been generated using digital elevation model through surface hydrologic modelling? And again, using 30-meter resolution digital elevation model, that means say it's a huge database and unnecessary every time whenever I will display, it will display for an entire Asia.

But I am focusing mainly on say Ganga basin. So, I want extraction of my data; clip the data or intersect the data only for Ganga basin. So, what do I need? I need a polygon boundary of Ganga basin which I will put over this and intersect, perform this AND operation through this Boolean logic or map calculator and I get only drainage network for Ganga basin. So that way, it helps in basically to reduce the data or reduce space on our hard disk.

Unnecessary data should not reside because unnecessary if I keep displaying that data, it would take a lot of time, one. And in future, there might be the reason of some errors in my GIS analysis which will be based on such themes. So, it is always very much required to extract your data; clip your data as per your requirements from a large database. Like here example is given that input coverage this, same example we have seen.

This is polygon which is you know, intersected coverage. Instead of here circles, I can have a boundary of Uttarakhand. I may have a soil map of entire India. I want to extract only for Uttarakhand like this. So, I will get output of only soil map for Uttarakhand. So likewise, I can create my own sub-database which will be used later and analysis speed will be also much higher rather than performing on an entire database.

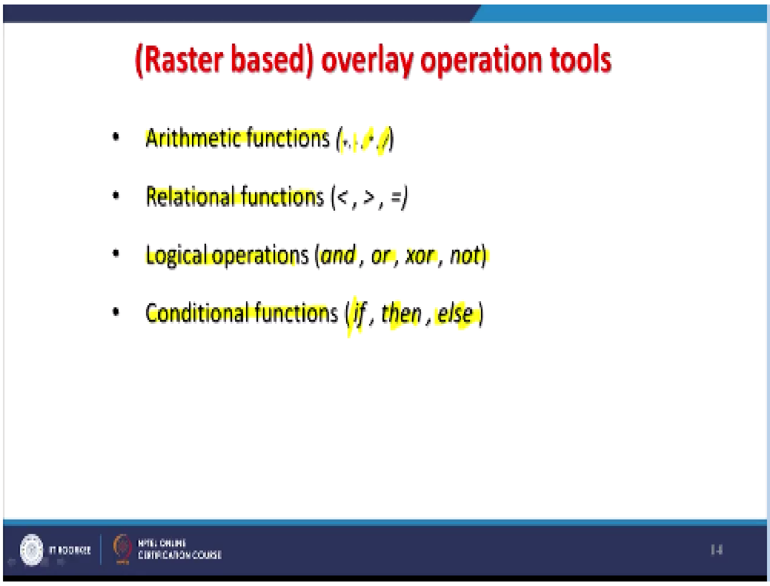
(Refer Slide Time: 18:30)



Similarly, there is an identity. Selection is also there. With that, everything located within the boundaries of input coverage is collected in the output coverage. And the boundary of the output coverage is identical to the input coverage. And like for example, I am having this polygon scenario. Identity coverage again those 2 circles, I am having input coverage like this and when I do, this also creates the output coverage like also.

So, identity procedure applies to the point which you can apply to the point, line, polygon coverage even on the raster, no issue!

(Refer Slide Time: 19:08)



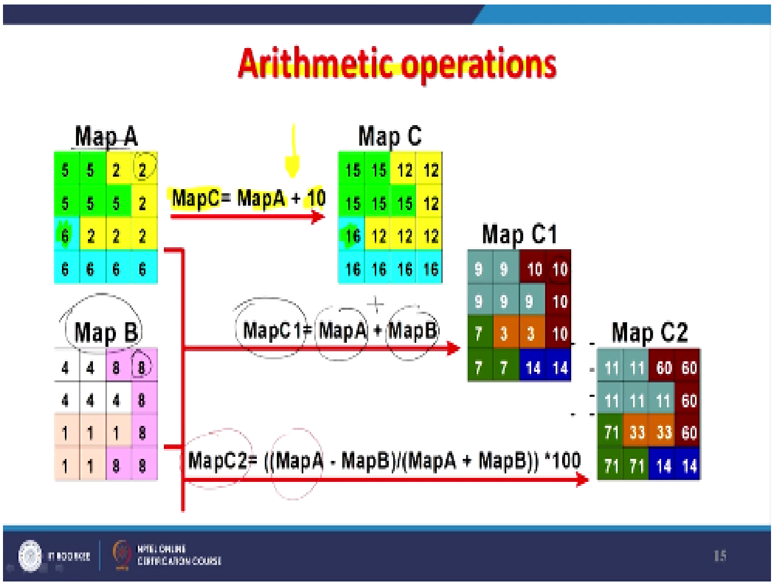
So, raster-based operations; much more operations, you can do on raster also. All arithmetic; this basic arithmetic operations subtraction, addition and your multiplication and division. All

4 basic arithmetic operations, you can perform on raster layers. You can also perform relational functions less than, greater than, equal to or in combination also.

Within the same syntax, within the same query, I can use some arithmetic, some relational or some logical operations together. So, AND, OR, XOR or NOT; all these 4 basic logical operations I can use. And also, conditional functions if, then and else; these I also can use and in combination. So, in one single line query, I can use all operators together or from each category, I can use individual operator.

As I have been repeatedly saying, whatever the output which the system will create, only requirement is that we have to understand. That is the only thing which we require.

(Refer Slide Time: 20:34)



So, like here I am taking the example step by step means from very simple example then little complicated and then we will go for more complications. Using initially only arithmetic operations; that means addition, subtraction, multiplication and division. So here what I am saying to the system that map C is going to be equal to map A + 10. So, as you can see that these different cell values are having different values like 5. When it is added with 10, it becomes 15.

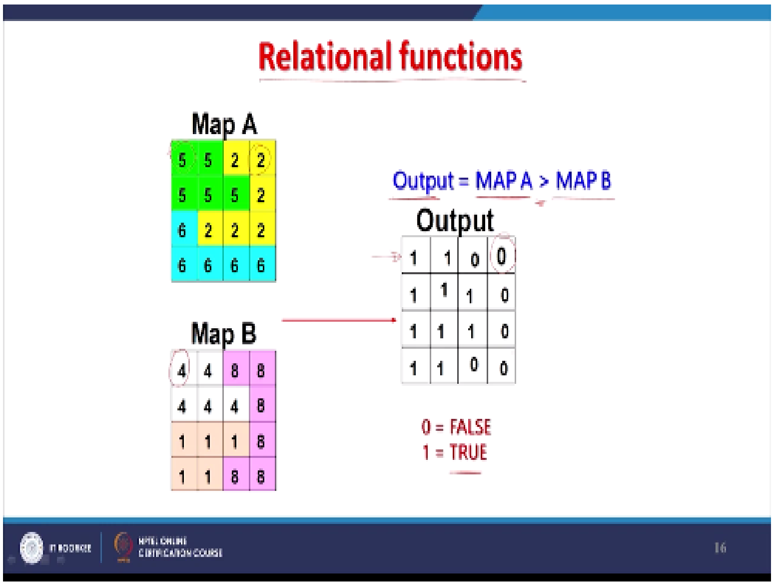
Similarly, 6 added with 10 becomes 16. Very simple arithmetic operation, what I have used only this + one. Similarly, you know I can add these map A and map B like this. So, my map C or map C1 is going to be map A + map B. And as you can see that if I take this corner value for example so 2 + 8 becomes my 10. Very simple arithmetic operations can be done.

Similarly, I can also do other operations like here. This is now map C2 example. So, now here in this one, 3 operators are being used rather all 4 operators are being used.

Then (Map A - map B) / (map A + map B) * 100.

So, all 4 operators have been used and my map C will carry that value. So likewise, you can use all arithmetic operations in one-go or other operations also.

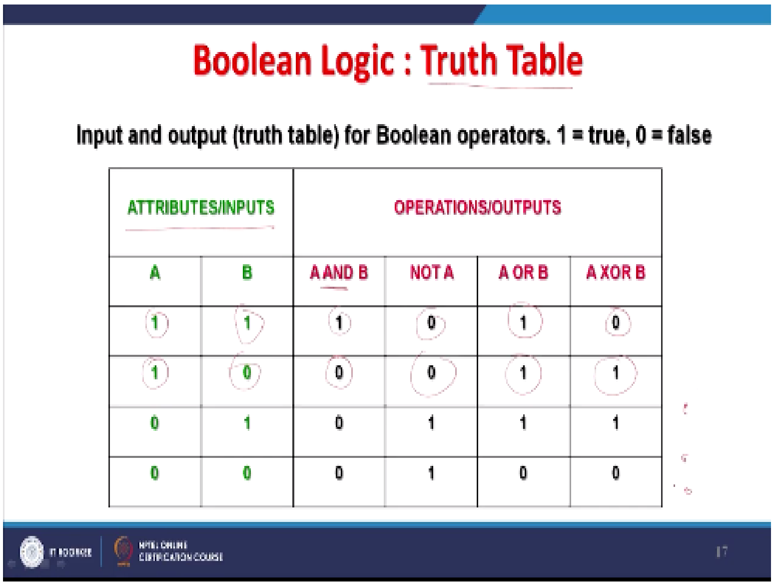
(Refer Slide Time: 22:42)



So, let us see some other examples also. Like relational functions; greater than, less than, equal to or that one. Now here, the condition for the output maps is if map A is greater than the individual cell level, the operation will be done. So, if map A cell > map B then it should fall as true, otherwise it is false. So here, what if I take this value so this value is 5. Map B value is 4 that means map A value is greater than map B and therefore, it is getting value 1.

If I take this example, here the value 2 is less than 8 that means it is false and it is getting 0. Now, this outputs here in this particular example, will become your binary output. So, from 2 coloured maps, you may end up with a binary map but if I require because I want operation something like this which is given in this example, I should be happy with that. That it is creating an output which is just based on the true and false condition. So, this is relational function.

(Refer Slide Time: 24:03)



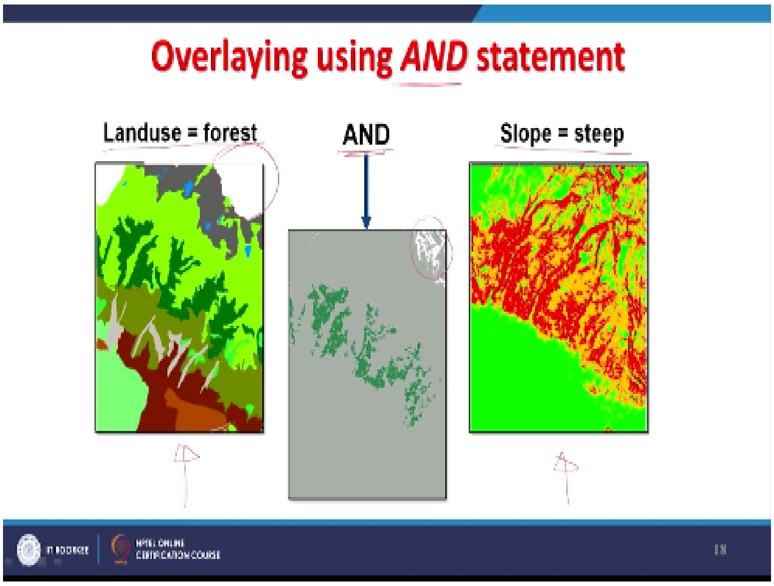
ATTRIBUTES/INPUTS		OPERATIONS/OUTPUTS			
A	B	A AND B	NOT A	A OR B	A XOR B
1	1	1	0	1	0
1	0	0	0	1	1
0	1	0	1	1	1
0	0	0	1	0	0

Now Boolean logic; another thing is we call as a truth table and truth table means basically true and false. So, input and output table for Boolean operators; 1 generally is the true value and 0 is the false value. So, if like here the input attributes are like this for map A and B and likewise. So, if in map A is 1 and in case of B is 1.

So, if I perform this operator that AND or intersection; that A and B then it remains 1. NOT A; then it is false. So, this is in my truth table, it is false. A OR B; anything that means 1 so it is true. A XOR B; it is false, 0. So for all kinds of inputs, all scenarios are given. When A and B are not 1 and 1; that means one is 1, another one is 0 then A AND B of course, it is false. NOT A is false. A OR B; the one is there. So, here is true. And this is true.

So, something like this, we can create a truth table depending on our requirements. But these operations can be perform very easily using either Query Builder or map calculator. The most common one is the map calculator in ArcGIS.

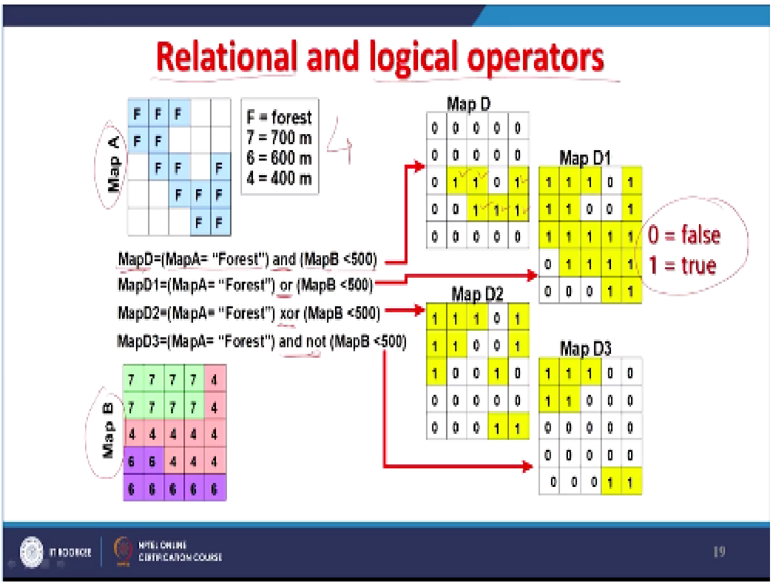
(Refer Slide Time: 25:50)



If we take the real example like using this intersection statement here that when land use = forest and slope = steep. So, 2 input maps are here; one is the land use map, another one is the slope map. An output map will satisfy these 2 conditions that land use = forest and slope = steep. So, when these 2 conditions are there, all those areas will get a true value and rest are false where this is not getting and where you may have the areas where you do not have any value either in 1 map or another.

Like here, you did not have the value. Therefore, the white areas are there or Nodata areas are there. So, in real examples, similar these Boolean logic operations can be done.

(Refer Slide Time: 26:49)

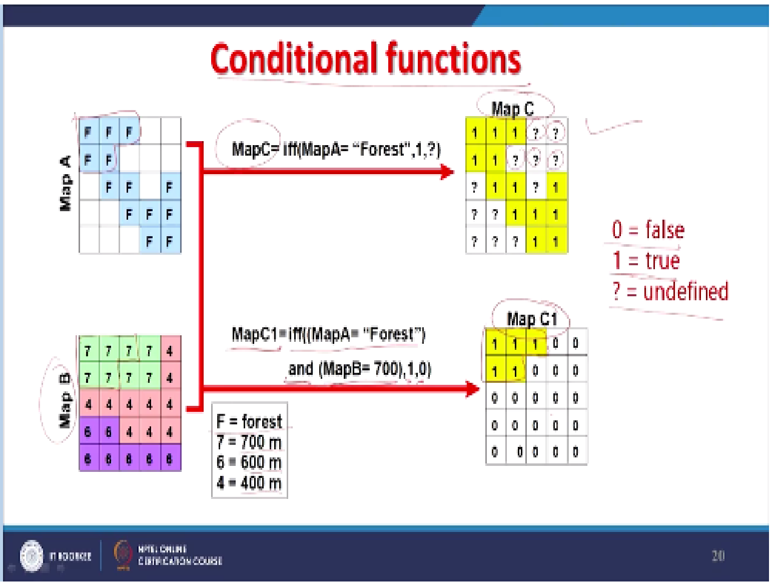


Now if I go for you know in combination of relational or logical operators then similar outputs can be created depending on how many classes in each map are there. Like for example in map A, we are having 4 classes. Similarly in map B, again we are having here only 3 classes. So, the first condition is that the new output map D where map A = forest and the operator is AND; intersection and map B < 500.

So, if these 2 conditions are being satisfied then it would be true otherwise it would be false. So, these 2 conditions so only this wherever you are seeing 1 are getting satisfied otherwise not. Similarly, for OR operator, Union, XOR, NOT or this negation also, it is possible. So, if these conditions are there then they are satisfying and the output we are getting all 4 outputs.

For examples of these Boolean relational or logical operators are in binary form; that is in true or false. There are lot of such applications are there. When we are having very complicated maps and we want to create very simple maps which will tell very simply that what is there and where these conditions are satisfying, where not, you can create such binary output maps.

(Refer Slide Time: 28:37)



Now, conditional functions where if else conditions are coming, so, 3 things are possible. If false then we will get the 0 value. Again, output might be the binary depending on the condition and true; 1 and unidentified will have a question mark or Nodata. So here in map A, we are having 2; 1 is F, another one is having no value or Nodata; unidentified.

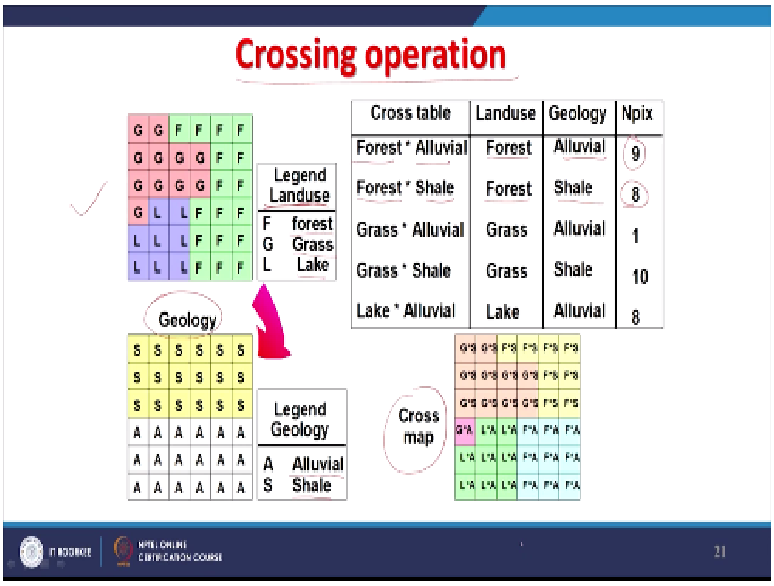
In map B, we are having just 3 classes and all are forest that 700-meter, 600-meter, 400-meter and different values are given. So, if I take this map C example here that if map A = forest then 1, else unidentified because while inputting, I know that there are areas which do not have any other value. Only along with this F, we are having value.

So, my output map will have unidentified areas where things are like this and otherwise, everywhere it will be 1. So, F basically has got converted into 1 or a true value. Now, another example what the condition if that map C1 output; for that output if map A = forest. Same condition as in case of the above example, however one more operator has come that is intersection; AND and map B = 700.

So, if these 2 conditions are getting satisfied then true otherwise false. Any one is not satisfying; it would be false. So, in map C1 output, what we are seeing only this yellow-coloured part is falling as true, rest is going 0. Like for example, map A is forest. So, these areas are having definitely forest. And also map it is 7; 7 stands for 700 so then these are satisfying to their condition. Rest not satisfying and therefore in this binary map, only 1 will show the true, otherwise 0 or false.

So, all kinds of operations in combination. Now here, I am using this intersection. Here, I am using a conditional operation. Here, I am using conditional as well as intersection in the second example. So, all combinations and permutations of these operators are possible through these analytical operations in this overlaying analysis.

(Refer Slide Time: 31:45)



Now here, we also call that cross table or cross operations where we can again create a cross map. Basically, it is through the table and then using the cross-table attributes, we can reclassify our map as done in this case. So again, a map which is showing the land use of 3 types; forest and grassland and lake or water body. And another example is the geology where only 2 types of you know lithology is present.

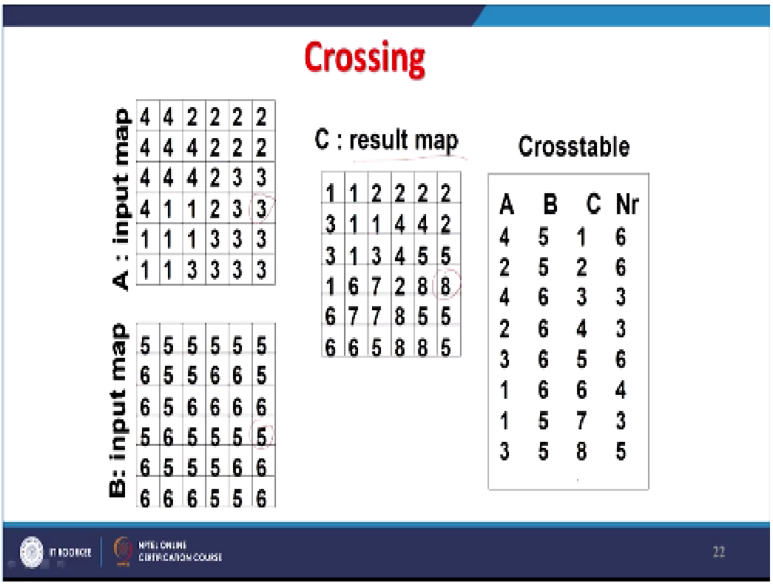
One is the alluvial and other one is the shale. So, the geology one is much simpler than your land use. Does not matter. Now, I can create a cross table. So, where forest and alluvial; 2 different maps, a value can be given 9. When forest and shale are there then forest and shale, then value 8. And likewise, I can assign values. And this cross map is basically having that information.

So, that I can use to keep the information of both maps together in one map through my attribute table while crossing these 2 operators. Such operations we perform when we go for change detection analysis. For example, I am having a land use map of 1990. For the same area, I am having a land use map of say 2000 and another land use map of the same area of 2010 and another 2020 that means I am having 4 maps having 10 years separation.

Now, I can create a cross table that which land use has changed to which one in those 40 years or maybe last 10 years. So, if I take very simple one that how in last 10 years, which land use whether the forest land has become open forest or grassland or wasteland. Similarly, a lake has become you know reclaimed land, sometimes it happens. The water bodies are reclaimed. So, through this cross table, I can very easily understand that where the changes have occurred from which land use to another land use.

So, that can be done through cross table. So, this is very-2 useful operations in real GIS projects where lot of such a change detection are also performed.

(Refer Slide Time: 34:41)



Another example of cross table is here that map A is input map and where values are varying between 1 to 4. Here values are varying only between 5 to 6. Result map is something like this. And for example, if I take this value here 8 so this is 5 + 3, it becomes 8. My cross table will have all information. So likewise, I can perform various kinds of overlaying operations depending of my requirements.

I can use single operator say within arithmetic, I can use only the addition or I can use all operators together and in some very complicated analysis, I can use all different types of operators in a different manner. So, few examples, I discussed with you. So, with this, we come to the end of this discussion about overlaying operations or GIS analysis part 2. Thank you very much.