Spatial Informatics
Prof. Soumya K. Ghosh
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 16
Spatial Query Demo Tutorial

Hello. So, we are discussing about this SQL on spatial database or querying on spatial data or spatial queries. So, today we will have a say very short demo which we will try to show that how these queries can be written or how these queries can be executed in any standard database supporting spatial extensions right. So, we will be using postgres or postgis which is which supports spatial data and with me is Arindam Das Gupta is there who is also a TA of this particular course and he will show you some small demo how things will be there and those who are interested can basically install in your laptop or desktop and see how things works right.

I just want to mention we have made some synthetic data that is why there may be some issues with projections or what we say SRS that is Spatial Reference System. So, let us not bothered about those things because these are all very handmade data to just show that how query works and we will be looking at it. So, I will request Arindam to start that we. So, what we will do? We will show you a PPT a query and then fall back to this postgres to show that how the query works.

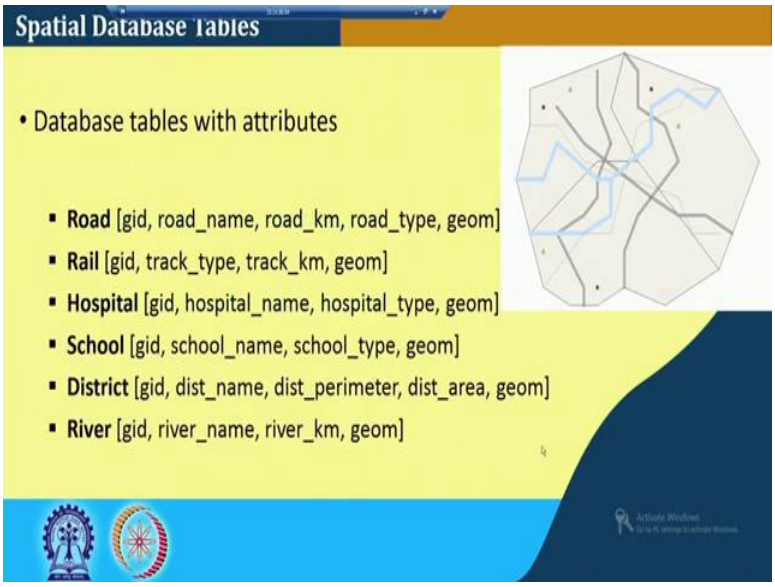
(Refer Slide Time: 01:46)

Hello. Initially we are showing the prerequisite of installing the postgres.

(Refer Slide Time: 01:52)

At first we have to install two softwares; one is postgreSQL 10 and any version you can download. So, this is the URL postgresql.org you can download the software and another is pgAdmin4, it is basically the interface to access the database system, it you can use the database very easily through this interface this is called pgAdmin I will show you in my demo.

(Refer Slide Time: 02:23)



And first of all we are using some synthetic data set, we have developed some synthetic data set. First one is road, rail and hospital, school, district and river and this is the map of these synthetic datasets and you can see the these the blue lines are river and two the black lines showing the road networks and rail networks are showing in very.

Dotted thin line.

Dotted thin line and one two there are two point data sets, one is school and another is this is the school and another is.

Hospital.

Hospital and first of all we are.

(Refer Slide Time: 03:04)

So, what we are having? We are having all 3 type of things; one is polygon like your district and so like that and there are polyline or line type of things like road, rail, river, right.

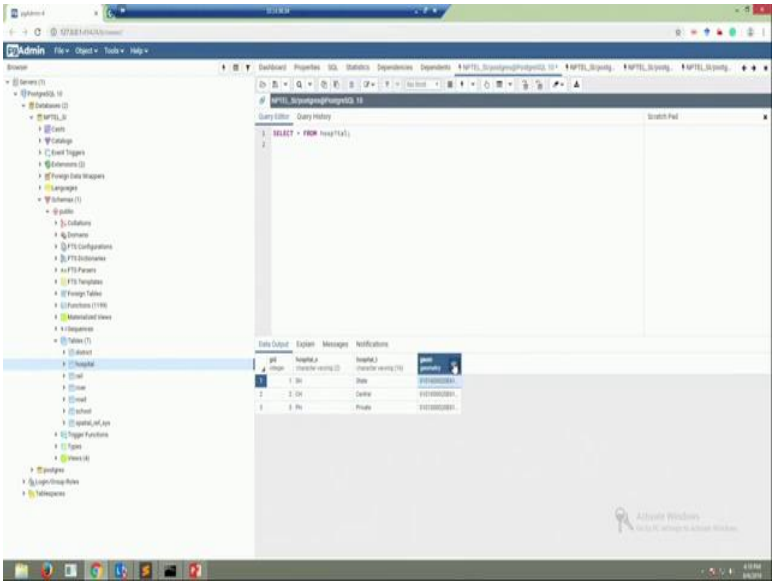
Right.

And we have point feature like hospital, district. So, all 3 type of data things are there and this is again I as mentioned by Arindam I am repeating that this is a synthetic dataset. So, we are not taken care about the projection system. So, if you want to project this thing, it may project anywhere in the world right like we will see that it is projecting

somewhere in Africa etcetera, but we do not bother about those things we let us see the query.

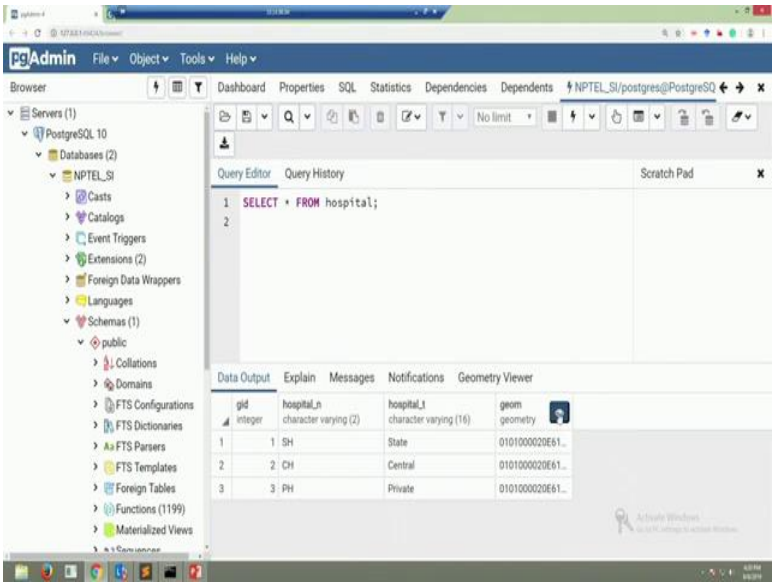
First of all I will show you how to selects a particular table from the database and for this is the formats it is similar to some similar to common database queries; select star from table name and for that we have developed a table, we have already created a table it is called hospital and it contains 4 attributes first one is the gid, the features id and second the name of the hospital third one is the type of hospital and geometry attributes. Hospital is basically the point type of data that is why it contains two types of the information one is x coordinate and another one is y coordinate for demo purpose I will show you the demo.

(Refer Slide Time: 04:25)



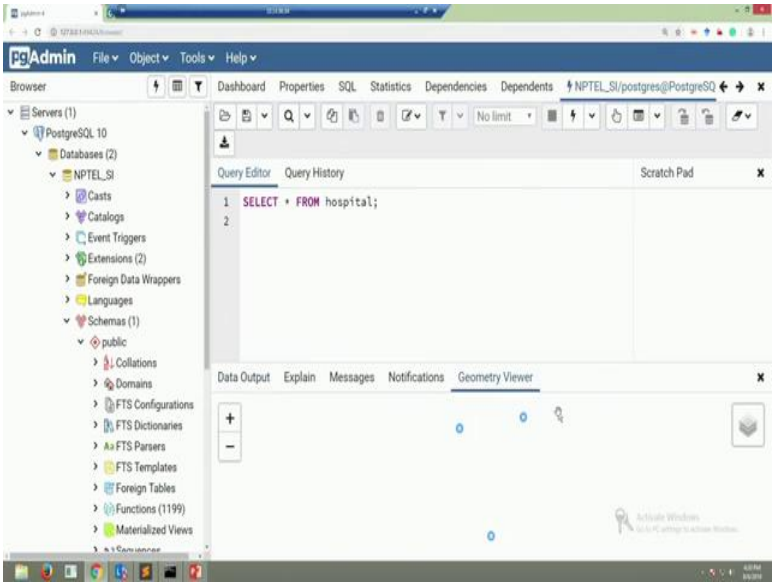
This is the pg admin interface which is connected to the postgres SQL and I will try to fire the query. Select star from hospital, suppose after clicking firing the query this is the output of the output of the query, it contains 3 data sets one is state hospital another one is central hospital one is private hospital. This is the if you want to show the demo.

(Refer Slide Time: 04:56)



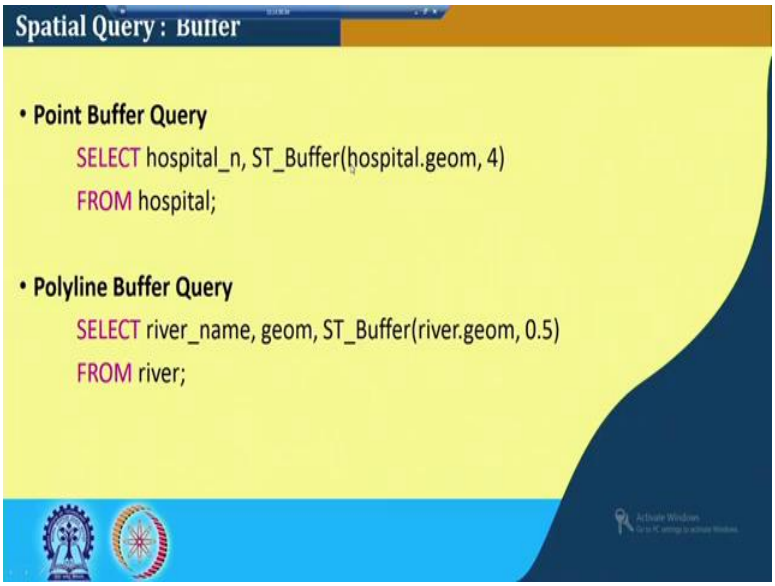
Now, I want to show how to display the contents of hospital table. After now you have first write the query. Select star from hospital then semicolon then click on the execute query it will show a table, it is a table it contains 4 attributes as per our dbt example and I want to show you these points on the map.

(Refer Slide Time: 05:21)



This is the these 3 points on map.

(Refer Slide Time: 05:26)

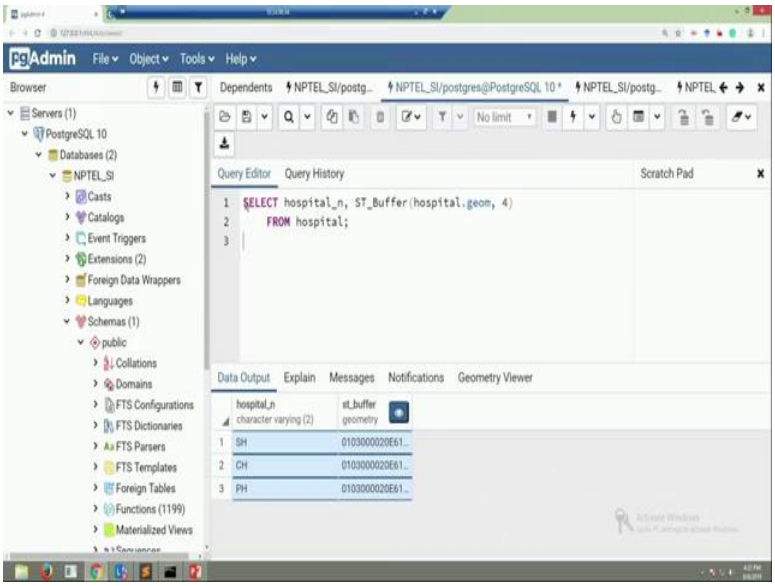


Our next query is buffer query. Suppose I want to know the effective area of a particular point. Suppose there is the point it is hospital I want to know the effective area of means effective area of hospital point within 4 kilometres for that we have to make a query.

This so, this why these are required? Like a suppose you setup a hospital or you set up a school then what we say? What is these overall catchment area type of the things, like a school will serves a population within or the students within 2 kilometre areas, hospital will serve a population within 3 kilometre areas. So, sometimes when you setup the things I need to find out that how much it is thing or in other sense if we look at our SQL sense or spatial query sense, we say that we need to make a buffer of so many unit that exactly what Arindam tying to show select something.

And for that you have to means select the this is the hospital underscore n means hospital name.

(Refer Slide Time: 06:40)



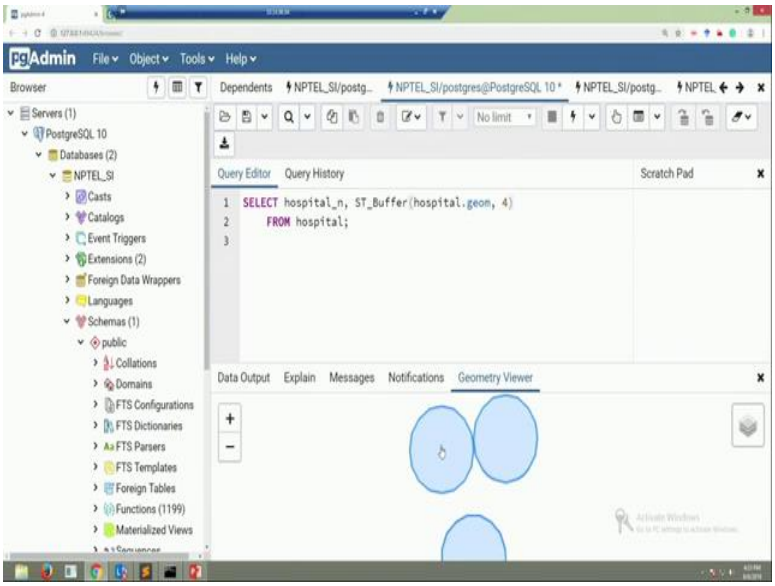
Select the hospital name and select the buffer area of particular the location of that hospital from the hospital table.

Like in this table 4 units 4 kilometres or something.

Yes 4 kilometers some 4 kilometers or.

Yeah, these are the query SELET hospital name ST underscore Buffer is the standard postgis function basically, it contains the geometry attributes and the 4 is the kilometres around the buffer around the hospital and select star from hospital, now I want to execute the query.

(Refer Slide Time: 07:02)



Now, it is showing the buffer area it is showing the buffer of particular hospital.

So, there are 3 hospitals now 3 buffers are there.

3 bufferes are there right and then I want to show the next is the this is the point type buffer I want to show the how to make buffer for line type of geometry. For that suppose I want to know the if river is overflowed how much area will be over flowed if during flood time.

Indentation of the.

Indentation of the.

So, these is the reason that buffering of the.

Buffering of the.

So, this is the line type.

So, as you told it is a point feature now we have a line feature.

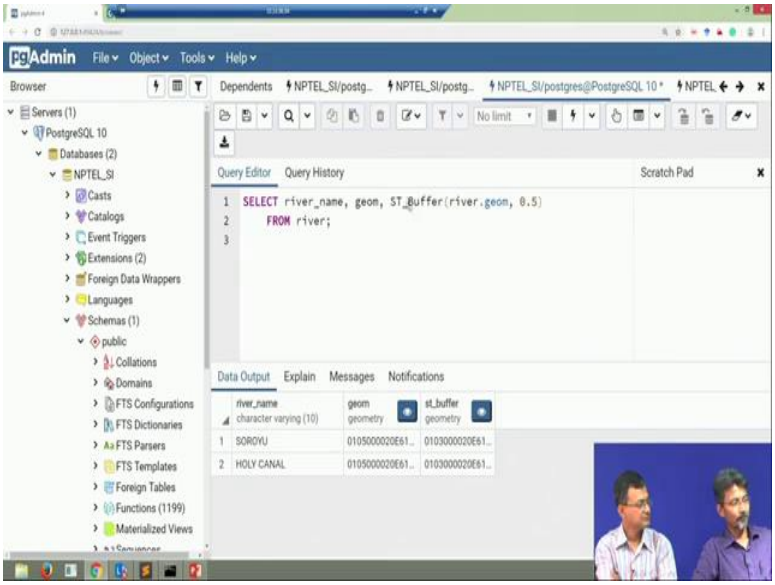
For what is the river name then geometry I want to show which the geometry of the river as well as buffer geometry of that river.

So, if we see that indentation is typically 0.5 kilometer or 500 meter.

500 meter.

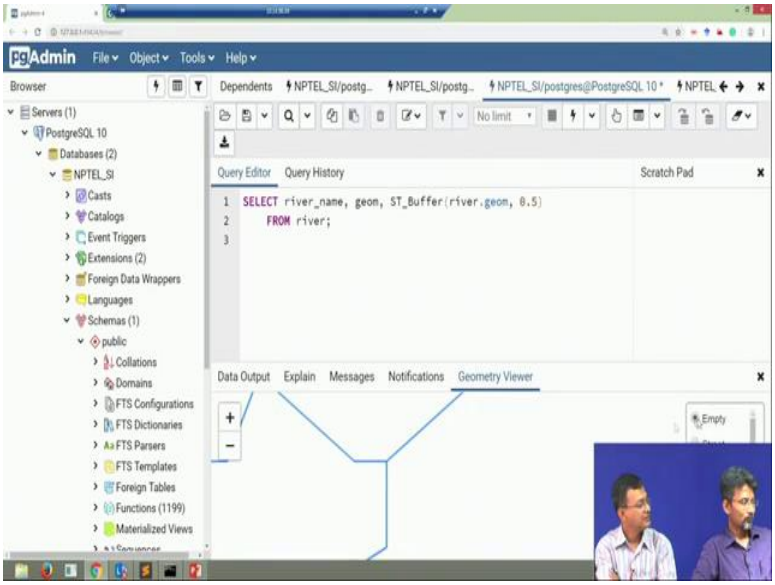
So, we want to find.

(Refer Slide Time: 07:54)



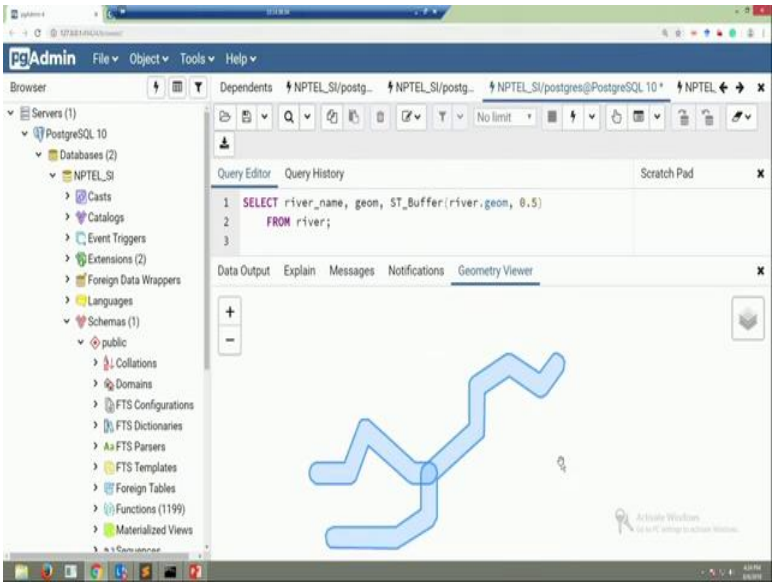
This is the query. SELECT river name. And then geometry of the river and ST Buffer function and through this function we calculate the buffer of the that geometry river geometry. This is the actual geometry of the this is the actual geometry of the river.

(Refer Slide Time: 08:11)



Then I will show you the buffer geometry.

(Refer Slide Time: 08:20)



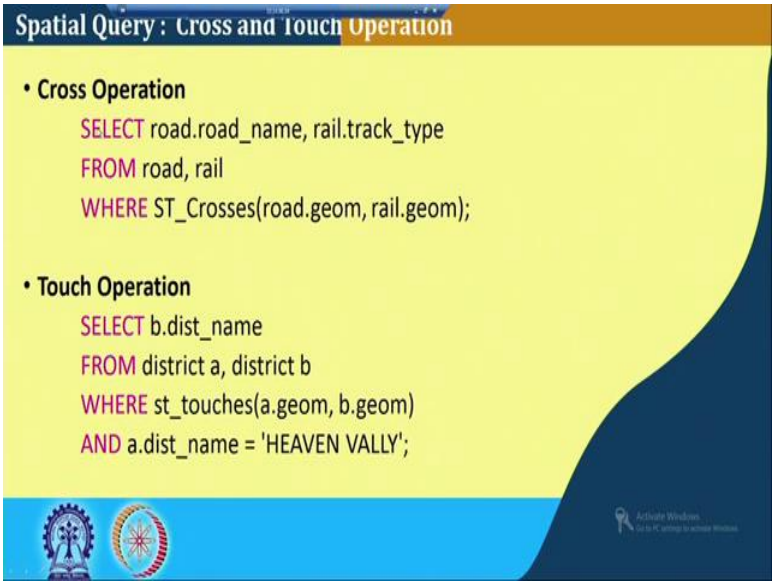
This is the buffer geometry of that same river. Same this is the buffer geometry of the river basically typically polygonal in nature.

So, one point you can see that initially the river was a line object what we have considered, but this buffer thing what we get the output is the polygon object right.

Right.

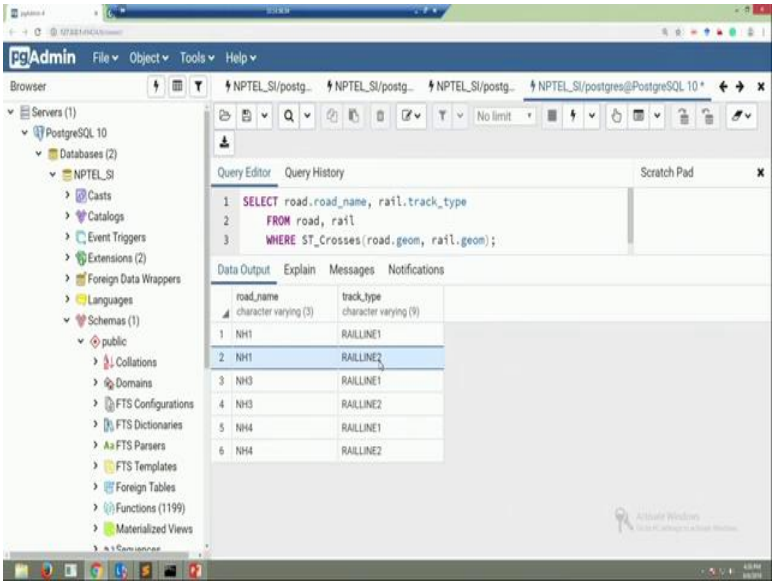
A polygon what we get a polygon thing right as a river ok.

(Refer Slide Time: 08:48)



Next is the next query is the now there is another operation is called cross and touch. Cross means which suppose one which roads cross the another roads or some roads cross the another rail network or roads cross the river network that type of operation can be done through this postgres and for that we have to use the ST Cross functions and here I will show you what roads cross and what type of rail tracks for that the query is like that.

(Refer Slide Time: 09:21)



SELECT road, road name and rail from rail track type and from there are two tables one is road one for road features and another for rail features and I want to check the cross of road geometry and rail geometry. Now, I want to execute the queries after executing queries it will show the actual data sets. Suppose rail line 1 crosses the NH1 both national highway 1 and national highway 1 crossed by rail line1 and 2.

And national highway 3 also crossed by rail line 1 and 2, national highway 4 is rail line 1 and 2, there is no cross on national highway 3 that is why it is not showing here.

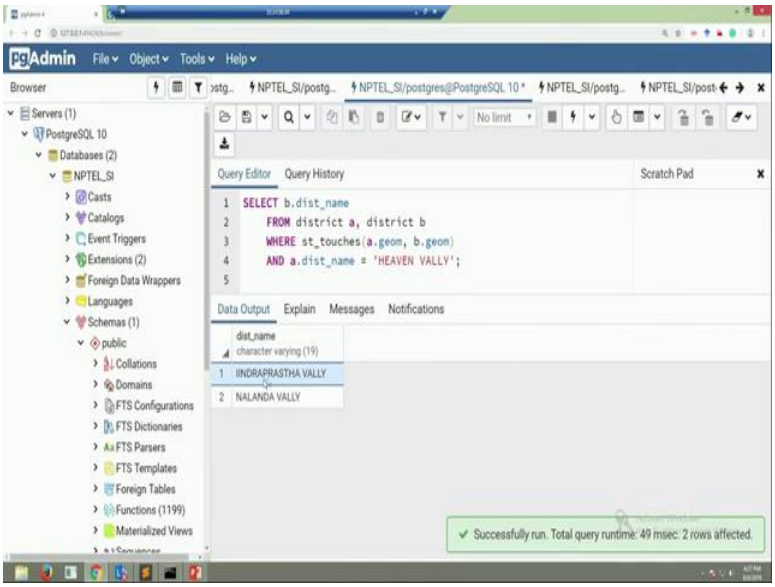
Right.

And then another is this touch operation which district is touched by another district.

Suppose, I want to know the I want to know the there is a hypothetical district called HEAVEN VALLY, I want to know the which district touches the district name HEAVEN VALLY for that you have to execute this queries. First of all you have to SELECT the district name, which district are touched by HEAVEN VALLY district and

from district a and district b same basically both are same district and touches geometry of a geometry means the actual HEAVEN VALLY geometry and b geometry means the other districts geometry.

(Refer Slide Time: 10:43)



After executing this query I will get two there are two districts this is called INDRAPRASTA VALLY and NALANDA VALLY there are two hypothetical district touched by HEAVEN VALLY here, this is the.

INDRAPRASTHA and NALANDA which are.

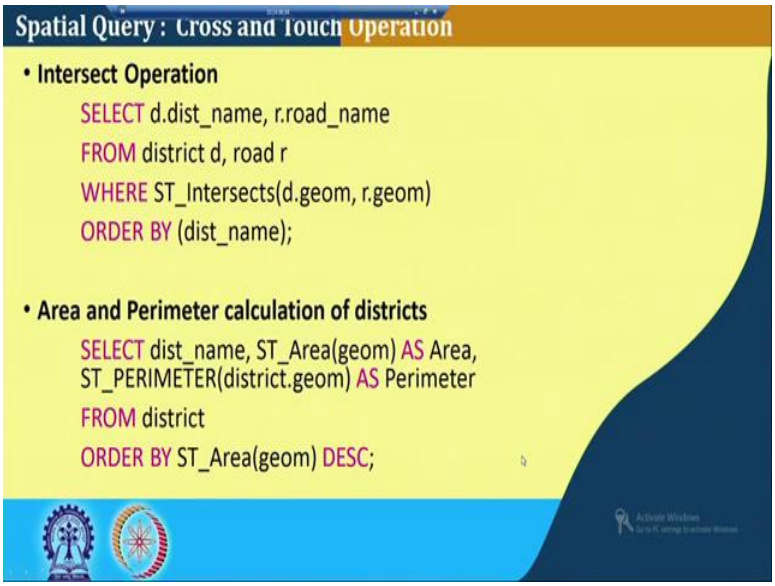
The.

Crossing.

In other sense their geometry are touching each other. Right?

Right.

(Refer Slide Time: 11:02)



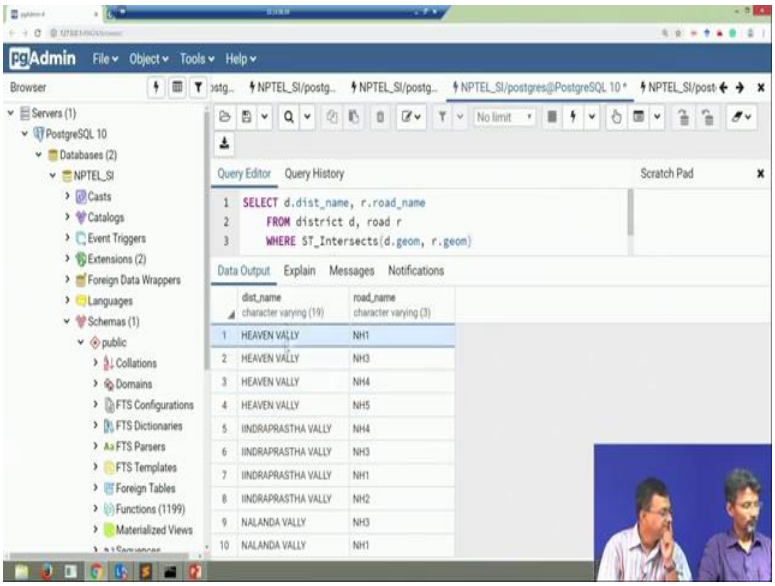
And then there are another operation it is called intersect operation I have to check which roads intersects suppose I want to check which roads intersect another roads or I want to know the which roads intersect particular regions and to check you can check which roads intersect which lake this type of operation can be done through intersect operation.

And for that I have to add my if I my in our first example, I will show the which district. Suppose I want to show the district name which is intersects by a particular road. In that means, we want to intersect a polygonal type of feature with the line type of feature. For that these are query SELECT district name and road name which and from the district d and road r and try to intersect the geometry of district and road network and or I want to show the display the district name according to the name order of name these is the query.

So, that this query you understand that what he is telling that we have now query of a querying between a polygon feature with a line feature. So, how this line feature or the polygon feature is interesting right.

Intersecting.

(Refer Slide Time: 12:10)



So, in this case it is a district versus a road how they are intersecting. So, taken from two tables one is a district table d and road table this. Incidentally all these tables have a geometry property or geom right.

Geom. Property.

Geom. Property.

Geom. Property. So, all are spatial data then this intersect is on the geom property right. Otherwise you see a road and district what if we have not defined anything so, what you have the intersecting points? So, here what is that the road running over the district right. So, we have a geometry class and we want to see that whether the geometry are intersecting right.

Right.

So, that ST Intersects d dot geom that means, district dot geometry and r dot that is the road dot geometry and it is as he mentioned it is ordered by district name.

District name.

So, it is the way it is displayed they ordered by the district name.

Now, you can see the in HEAVEN VALLY there are 3 roads in HEAVEN VALLY road NH1 NH3 NH4 and NH5 crossed by crossed through HEAVEN VALLY Heaven valley district and the INDRAPRASTA VALLY is crossed by only NH4 and NH3 and NH1 and NH2 and NALANDA VALLY is crossed by NH3 NH4 we can get these information from such type of operations.

Right. So, who which road and which.

Which district?

District are this and it is sorted by it is in terms of.

In terms of name of the district.

Name of the district.

Even we can have these queries of finding that overall how many road crosses with maximum number of crossings.

Yeah counts.

And those type of. Then you can use that other SQL construct right counts.

Counts.

Count star and type of things to construct to count the things right.

So, but the interesting part is there, here the intersect is based on the geometry right geometry of the thing right like intersect of the spatial operations in geometry of the thing and this type of spatial operation as we have discussed is are all topological features of spatial data; so, topological operations. So, topological operations we are able to execute and do on a spatial data base right. So, then definitely this spatial data base like as he is showing this postgres or postgis data base those supports things, there are several other data bases which supports spatial data so that it understands that geometry property so on and so forth.

Right.

And to so the how all these are the geometrical operations. Now, I want to calculate something. What? Suppose I want to know the area of particular what is the area of the particular district area or perimeter of the particular district for that postgres offers two functions; one is ST Area here and another is ST PERIMETER; we can calculate the Perimeter of the district.

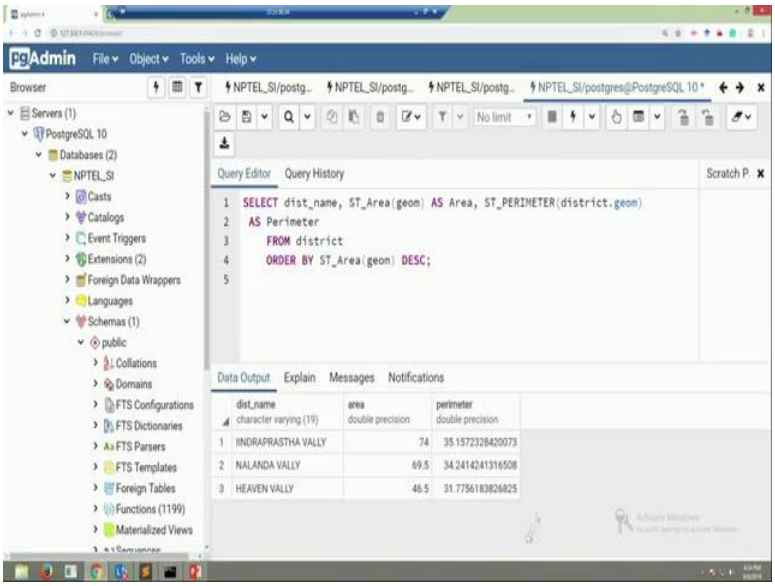
Perimeter of the district means perimeter of any polygonal features we can calculate or area of any polygonal features. For that we have to SELECT the district name what is the area of that the that district and what is the perimeter of that district for that I want to show the.

So, the area and the perimeter is calculating and this.

So, this is if you see these are some what we say it is not a different type of operation. So, one is for intersect other was cross and now we are calculating the area and perimeter type of things right. So, sort of a aggregate operation.

And I want to show the display the output in according to the descending order of the area.

(Refer Slide Time: 15:56)



Now, check the this is my this is my query SELECT district name.

ST underscore Area geom AS Area and ST underscore parameter PERIMETER district dot geom and AS Perimeter.

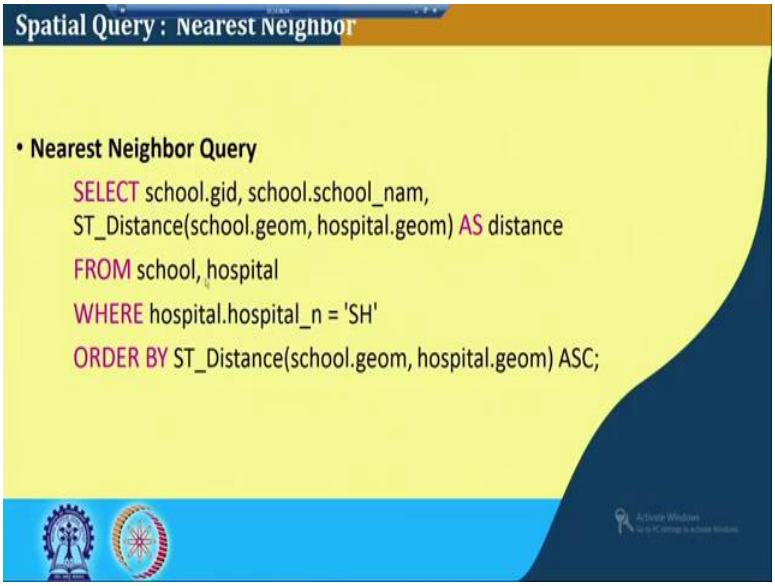
Perimeter.

Perimeter is a renaming operation and then FROM the district table ORDER BY the ST Area according to descending order. After clicking the execution query execution there are 3 district in our data base one is INDRAPRASTA VALLY these its area is 74 square kilometres NANDAN VALLY it is area is 69.5 square kilometers and then HEAVEN VALLY it is area is 46.5 square kilometers and these are the perimeters value all in kilometres 35 34 and 31.

Displayed descending order of the area.

Descending order of the area.

(Refer Slide Time: 16:48)



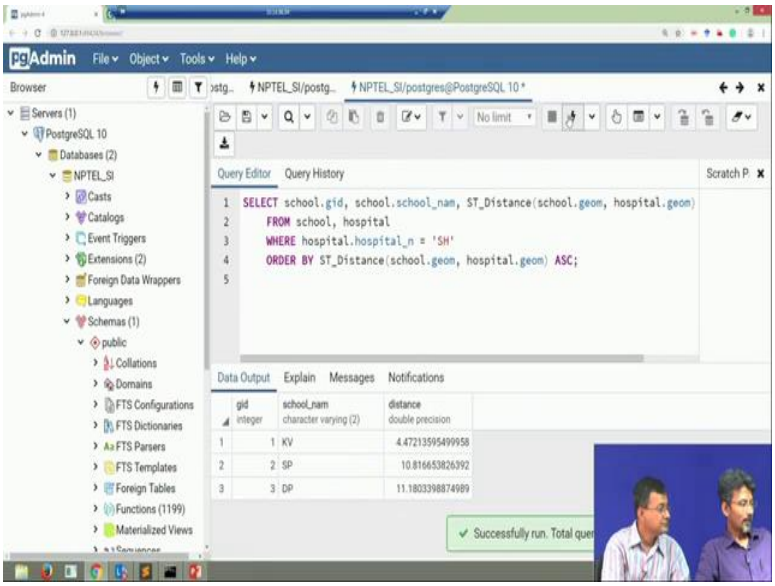
Then there is another query I want to show this is the nearest neighbor query. Nearest neighbor query is a complex query, it is I want to know the which are the schools which are the schools available from a particular hospital.

That means, which the hospitals attached with particular the hospital provides services to which type of schools and this is the query SELECT this is the query SELECT school id and then school name Distance school underscore geom and hospital underscore geom.

That is the calculation writing distance what is the distance between hospital and particular hospital from each school and AS distance then from school and hospital where hospital name is SH.

Then order we will display the order by the distance from the school and hospital.

(Refer Slide Time: 17:45)



This my query.

Ok.

And I want to execute after executing this you can find the there are 3 schools from this hospital SH hospital and one is KV and one is SP and DP 3 schools and these are the distance of the from the ascending order of sh from the sh hospitals these are the output.

Right.

So, here some sort of a what we are taking two types incidentally both are point features.

Point features.

So, even if it is not a point feature if the distance function is defined between a polygon and point it is fine, I may do a centroid of the thing and do the point.

Right.

Nevertheless. So, here what he is doing he has looking at that trying to find out those school name and it is distance from the a nearest hospital of the or a particular hospital called.

SH.

SH right so; that means, in other sense and to find out that given a hospital which are the school it is attached to or in case of any medical crisis, where what is the hospital designated hospital for a particular school right.

Right.

So, these queries.

Find out give display the school name and distance between the two geometry feature or two spatial features and then school name and geometry of the things right

Right.

That means, here what we are doing? We are basically finding a distance operations which you can again it is a distance operations, I can look at it as a something called a distance function in terms of a metric right distance metric operation. And then display the thing in a ascending order.

See now it is the distance is in kilometers here, 4 kilometers, 10 kilometers and 11 kilometers.

So, what we try to do in these particular short demo these are the two open sources right postgre open sources

Postgres postgreSQL 10 and pgAdmin4 are the open source things this you can download and install in your system.

Yes.

And then this supports spatial extension. So, this also supports all non spatial or our attributes type of things along with that it supports a spatial extension; that means, spatial queries like the queries related to topological measure, related to different distance metric and type of things or some of the aggregate functions, these are supported by this

SQL and then next this is the if you look at these are all different type of data types what we; what we did is one is line, one is point and another is hmm.

Polygon.

Polygon type of data type and incidentally all these data types are having a geometry property; that means, we mean to say this has a these are all spatial data types right. So, again I repeat we have taken synthetic data sets that is right.

So, that means, what we say on a paper drawn and type of things data set, but in reality you will be doing you will be getting this data from the different sources. Like if you recollect when we discussed about different types of service oriented architecture in what we do? That say road data bases or road data, is the there is a custodian of the things like in our country it may be the transport department or some of the data is NHA National Highway Authority right or something like that first I rail may be with

Railway department.

Railway department right hospital with some medical or either state or central based on that who is supporting this hospital, that will be there. Schools is with the education department, district is typically what we look at is more of a administrative boundary and maybe with the administration like district boundaries maybe with planning department or district administration and type of things.

And river is maybe with hydrological or water department or in another sense what we are telling is now I require a query say road, if there is a is a river, if it is inundated how much it will be spillover water or given a hospital, which what are the other school city scattering? Now, it requires the data set from different repositories one point and secondly, in need to do some operations either it is something call buffering operation or overlay operation or touch operation and these things right. For what we have discussed in our earlier lectures that these type of operations falls in different category like topological operations Euclidian operation matriculated operation and so, and so, forth those aggregate operations and so, and so, forth. So, those type of things we are able to realize.

For that we require is backend database which understands spatial data that exactly what Arindam showed just now is postgre SQL right.

Right.

Now, this is a typically a hospital table.

Right.

Right now these we have populated based on this synthetic data, but in reality there maybe lot of other things right like the type of hospital whether it is super specialty hospital or not what sort of hospital,l at what level of hospital like is it a primary health center or block health center or district health center or some state level hospitals or something right. It can be hospitals for private hospital with some specialization. So, that all depends on how you design your data base right. So, this is very type things what we wanted to discuss.

And another important thing what he has shown is the geometry right what should be the geometry of that particular thing or in other sense in this case what is the location of that particular hospital and so, and so, forth right.

And so, there are he considered 3 hospital one is state, central and private, but it can the attributes set can be totally dependent on you how you look at that attribute set. Now, we deal with different types of queries one is point query as we discussed that finding a buffer. See this is a straight forward thing what he did, that taking a hospital geometry and creating a buffer around it like 4 units in this case maybe kilometre. So, the requirement as we are told the requirement of such queries are that I want to see that how much I can type of things.

So, there can be polyline buffer query right with a line I want to create a buffer like as he mentioned that river inundated, I want to find out the inundation layer right and we have a even polygon buffer query.

Right. So, I can have a polygon buffer query in this case I could have written a polygon buffer query also right

We can write a small polygon buffer query.

Any polygon we have a polygon? Oh no polygon we have only the district as a polygon right.

District as a polygon.

So, I can have a polygon buffer query.

Same similar thing

Same thing and similar thing.

Just put just put here in st buffer just put the suppose the polygon name is lake, for that just type lake dot geom. Here and I want to the amount of the size of the buffer around that polygon.

If you point if 0.5 then it will create another polygon, in this just 0.5 inches or kilometres large. Then of original polygon

So, the next is. So, other two operation as cross operations like road crossing rail, road crossing maybe road crossing any river any two line by line or it can be some whatever the cross function is allowed for that particular features that we can do, in this case road versus rail crossings are there and we again with that ST crosses. So, ST underscore crosses is a is a which is supported by this postgres.

Postgres.

Postgres. So, that if it if you are using some other things like some SQL extension of some other data bases like oracle or I think mysql is also there.

Mysql is also there, but the function name is different there is.

But never the less you do not have a large number of number of functions said like this and so, you can always check and look for the things right. So, there is another thing that touches the two district ab which they touch each other or not like in our case touching definition or the again the.

Says that it should be

Adjacent distant.

Adjacent distance means they are they are boundaries.

Common boundaries.

Or they will be will be common boundaries at some portion of the things with boundary reduced should be sharing. So, a dot district name equal to heaven valley in these types of things anything could have been there, where we want to find out that where the geom is which are the district which touches that particular things like I want to find out which are the which district which touches west right.

Right.

Where we are now right. So, that can be a thing right. And similarly other operation like intersect to intersecting or area and perimeter calculation of a particular this is for a particular district, I can calculate for any polygon right. A polygon has a area and perimeter. So, you can calculate the area and perimeter and that may be required for different type of purposes right.

So, this is the query and finally, we had a nearest neighbor query there is also very strong and popular query finding out the nearest neighbor based on two feature. In this case school name and school id name and distance from a particular hospital and. So, you want to find out that for a particular hospital called SH which are the schools are there and display them in a ascending order; that means, the nearest school is displayed.

Display at first then.

Right things will be there. So, this query is basically a sort of a nearest neighbor query. So,. So, this is some very few standard or very few query set on postgres. So, I what we believe that this will help you those who are interested in looking at those spatial queries to help you to jumpstart on writing some of the queries, there is a link he has provided for downloading the particular this postgres gis database system. So, that this can work on any database system which supports spatial query sorry spatial extension are there supports definitely spatial query.

You can work on right. So, with this let us conclude our short demo session, thanks to Arindam Jaydeep Shreya other TS who are not in the screen today, but they will be

definitely sometime in the future. So, for setting up this quick, but all in comprehensive demo which will be good for I believe for many of you to work on.

Thank you.