

Hi.  
  
So this is a video about null values.  
  
So what we're doing here is we're continuing the example from last time.  
  
The example is still the employee table, but I changed something in the employee table.  
  
I added a row, which is not so important.  
  
What is important is I added an attribute.  
  
So the attribute is called supervisor.  
  
What this attribute records is who is the supervisor of whom.  
  
And so we have John smith has a supervisor 02 and 02 is the ID of Jim Stewart.  
  
So this means that Jim Stewart is a supervisor of John Smith.  
  
However, Jim Stewart doesn't have a supervisor, it seems.  
  
Now that could have two reasons.  
  
Either we don't know who it is, or we just don't have a supervisor in this case because Jim Stewart is the boss of the whole thing.  
  
So it's either unspecified because there is no supervisor, or it's just unknown.  
  
If whenever something is unspecified or unknown, it gets a value that is called the null value.  
  
So first of all, null values are unavoidable.  
  
We saw in this example there might be the boss of it all who doesn't have a supervisor.  
  
But I also should immediately mention that the the more null values you have in the table, the more likely it is that the table is badly designed.  
  
But let's just accept now that there are null values.  
  
And I want to briefly mention how to query for them.  
  
Because null values follow a special logic.  
  
Whereas we saw conditions before which were boolean true or false, null is really something which is unknown and unknown is not true or not false.  
  
It's just unknown.  
  
And I'm going to give you two examples what that means concretely.  
  
So suppose we want to see the names of all those people who do not have a supervisor.  
  
So what you could be tempted to is writing select last name from the table.  
  
Table is employee.  
  
So we select the last name from the employee table and now we want to say it doesn't have a supervisor.  
  
Now we might be tempted to do the following.  
  
We might be tempted to say where supervisor equals null.  
  
However, this will not give us any results.  
  
Why not?  
  
Because null is not really a value, it just says unknown.  
  
So if I have an equation which says supervisor equals unknown, we just don't know, the supervisor could be equal to the unknown or not equal to the unknown.  
  
That means if I evaluate this expression supervisor equals unknown, then the whole thing becomes the value unknown.  
  
However, the where clause, in order to give a query result, should evaluate to true.  
  
So I only give back the rows where the where clause evaluates to true.  
  
But here, this one will not evaluate to never evaluate to true, always evaluate to unknown.  
  
Therefore, the result for this query is just the empty query.  
  
So in order to do this properly, there's a way, in fact in SQL, to do this where we instead of equals null, we have to write is null.  
  
It's a special language construct just for null values, where we can cheque whether supervisor is null.  
  
And then in fact, you can get all those rows where the supervisor happens to be null, but that's not the same as equals null.  
  
Now, second example of the same kind.  
  
Let's forget about the supervisor, let's just look at the salary.  
  
Always good to forget about the supervisor and look at the salary instead.  
  
So here we have a salary.  
  
Suppose we want to have the name of those people who have a salary which is bigger than 30,000.  
  
Then you get those people where the salary is bigger than 30,000.  
  
And suppose we also want the salaries where the salary is smaller or equal to 37.  
  
Sorry.  
  
In fact, what I wanted to have is not and, but I wanted to have or and would have meant just equals to 30,000, which wouldn't be so interesting.  
  
For my purpose, I need an or here.  
  
So if you look at this, you might be tempted to say salary bigger than 30,000 or smaller equal than 30,000.  
  
Well, this will give you everything, right?  
  
Because clearly the salary will be always either bigger or smaller equal than 30,000.  
  
However, what happens if you run this query is you get all the rows but those where the salary is equal to the null value.  
  
Because again, if the salary, if there was a salary where we don't know what it is, then that expression for salary equals null would evaluate to unknown.  
  
That expression would evaluate to unknown.  
  
So the whole thing would evaluate to unknown.  
  
Therefore we just don't know.  
  
Therefore we don't get this row where the salary is equal to the null.  
  
So this lecture was about null values.  
  
And it's important to take care of the right logic.  
  
With null values, it's neither true nor false, it's just unknown.  
  
Thanks for listening.