

Hi, this video is about identifying entity types based on the system description, based on the description of the data that we are trying to store.  
  
So how do we go about of seeing what the entity types are that we are interested in?  
  
Well, indeed, we start with the description of the data.  
  
The description of the data is usually just some text that describes what should be somehow stored.  
  
And then how do we, based on this description, identify the entities?  
  
Well, we deduce the entities by searching for nouns in the text and for noun phrases.  
  
Be careful, however, because nouns and noun phrases are not only about entity types, but also the attributes describing the entity types will show up as nouns and noun phrases.  
  
So you will have to figure out what's the difference.  
  
And here the difference is that an entity type really is something describing entities, meaning objects that exist and that have properties, whereas attributes are just properties that usually can be just described by just a value or something.  
  
So we should not put attributes as entities.  
  
And the other thing we should avoid is that we shouldn't put what's called global variables.  
  
So we will see in a second an example that describes the data that the bus company wants to store.  
  
Now, the bus company is going to show up as a noun in the description.  
  
The bus company is in fact an object which needs to be described with many different attributes.  
  
But as this is the database for one bus company, there's no point of putting the bus company itself as an entity into our planning because we are not trying to create a database of whatever all the bus companies in the country.  
  
We are trying to describe the database for that bus company.  
  
And so the bus company is the global variable that should not be included in the description of the data.  
  
And then after we identify the identities, we will move on, seeing what the relationships between those entity types are.  
  
And finally we will go back and identify really the attributes.  
  
And then we have all of the information that we can get from the system description.  
  
So now let's move on to the system description.  
  
So here we have the first sentence is a bus company owns a number of buses.  
  
Now, bus company is a noun.  
  
So it's very tempting to put this as an entity.  
  
But just as I said before, bus company is a global variable because this is the data of that bus company of one company.  
  
So we don't include it into our entities for the database.  
  
And it owns a number of buses.  
  
Now, number is a noun, but this is not really the thing that is an entity.  
  
Here.  
  
The entity is probably buses, because we will have, I mean, the bus is an entity here, then the Next sentence.  
  
Each bus is allocated to a particular route, although some routes may have several buses.  
  
So what we see here is we have buses featuring again in a bus company, not too surprising.  
  
And we have now the routes that the buses take.  
  
And I would say that the route is also an entity type because there are several routes and they will have some properties.  
  
Similarly, the buses, there are several buses and they will have different properties.  
  
Each bus, for example, have a registration number.  
  
I guess it comes up in fact in the next sentence.  
  
Each bus has a unique number.  
  
Now we already said bus is an entity type.  
  
Fair enough, but what about unique bus number?  
  
Is it another entity?  
  
A bus number is a noun, could be an entity type.  
  
However, if you think about it, if you want to describe the object bus number, it's just the number itself.  
  
So this is not really an entity type, it's just a property of a bus.  
  
Therefore we will see that later.  
  
It's definitely not an entity type and we will put it later on to the attributes.  
  
It is important to store information about the seating capacity and the make type of all buses.  
  
So we have seating capacity here and we have make type of the buses.  
  
And again, these things, if you think about it, are pieces of information, but they are not entity types.  
  
So you couldn't to describe the seating capacity you need just that number and nothing else.  
  
So it is just a property of a bus and not an entity type by itself.  
  
Similarly, make type, each root is distinguished by a root number.  
  
So root we already said is an entity type.  
  
Root number we will see is not an entity type, will be an attribute passes through a number of towns, number of towns.  
  
Well, number of it just means to several towns and towns here is in fact probably an entity type because to describe a town you need some other pieces of information.  
  
Each town will have some attributes that will allow us to identify what town we are talking about.  
  
Maybe it's also important to know how many people live there.  
  
If you want to organise bus routes and so on.  
  
Information is available on the average number of passengers carried per day for each route.  
  
So it's again, average number of passengers will be not an entity type, it's just a value.  
  
One or more drivers are allocated to each stage of the route.  
  
So now here we enter drivers entered into the picture as noun.  
  
A driver, clearly I would say is an entity type because the driver is something that will exist and we will have to describe it via several attributes.  
  
And drivers are allocated to each stage of a route which corresponds to a journey through some or all of the towns.  
  
Journey, I think, is already kind of described via the notions of stage, of the route and route.  
  
Stage.  
  
However, I would put down as an entity type, because a stage is the thing that the drivers are allocated to and if we just would store information about routes, we would have trouble allocating the drivers to the stages they have to actually service.  
  
So we backtrack, which corresponds to a journey through some or all of the towns on a route.  
  
Drivers have an employee number, a name, address and sometimes a telephone number.  
  
So this last sentence clearly tells you just information about drivers and there are no new entity types.  
  
So to summarise what we now get from this text, here are five entity buses, routes, stages, driver and town.  
  
Sorry, I just noticed that.  
  
Just to point out one mistake that I'm often making but that you should avoid.  
  
The entity type is not buses, any type is bus.  
  
So you always think of it as singular.  
  
The type is bus.  
  
And then there are instances of bus, bus, route, stage, driver, town.  
  
All of them are fairly obvious from the text if you read it again, the one critical bit is the distinction between routes and stages.  
  
But we had to do this because drivers are allocated to stages and we need to store that information somehow what the stage is and from where and where it runs and so on.  
  
So these are the entity types.  
  
Next time we will see what the relationships between those entity types are.  
  
Thanks for watching.  
  
See you next time.