Leonine
LEARNING

Video Transcript for Module 2: Understanding your Data

Hi guys, so in this movie we'll be looking at how do you understand your data, how can we get a better understanding of the data set we are working it; and looking at that, one of the very useful and proven way of doing this is to leverage the filter button from the table.

So, whenever you create a table, by default it comes with a filter button which is this drop-down arrow showing by the side of each column within the data set so this is called a filter button and you can explore this by going into the table design tab again and you can see that one of the options we have under the style options is the filter button. So, if I deselect it, it goes off, if I need to bring it back, I can simply check on that button.

By using this it helps us understand our data more, so obviously we can have a quick look at the different column headings to understand the kinds of info we have collected. So, we know we have some date information, some unique ID, transaction Id, some store information, store ID information then also you also have like a product ID, product name, product category, you also have the price information, the volume and finally we have the sales value information.

Now we might want to get a better insight on this and a very quick way of doing this before you start analysing is to click on the filter button so by clicking on the filter button very quickly with the date column we can see that we have data from three years so 2016, 2017 and 2018 and if you expand any of these years you will see that we have information from all the months in 2016, all the months in 2017 and then in 2018 we also have for all the months so we have three full years data. So very quickly by using that we can see that; another thing we might want to do is to check the product name and very quickly we can see that there are four unique products

within this data set and if you check the category, we can see that we have the categories for sugar, other income and molasses. So, using your filter button is a very quick way of getting an understanding of your data.

Having said that one other useful thing to do at this stage will be to identify the kinds of fields you have in your data set and broadly speaking there are two main types of fields, we have the facts fields and the dimension field. And a good way to think about your fact or value fields are your numerical fields so for example in this data set we can see that the value field is a monetary field where you have numbers and this is going to be a fact field the volume is also number, it's going to be a fact field and the price is number, a fact or value field.

So, by default generally where you have numbers that would be usually be a fact or value field and where you have texts its usually a dimension fields because that helps you explain some of the numbers you have. So, for example I can explain some of the numbers you have so for example I can explain my revenue by product category like product name so those other fields like the category, product name, even the IDs can serve as a dimension field and then more importantly your data is one of the most important fields and happens to be a dimension field as well.

There are some other considerations in terms of determining whether a field is a fact or value field, but I think for now it suffices to say that your number fields will be your fact or value field whilst your text or date field will serve as your dimension field. And in subsequent movies we will talk about the nuances regarding that. So, see you in the next movie.

