Now that we have learned about the importance of data quality and data privacy, in this video we will learn how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data. It's very common when collecting or importing data - whether through manual or automated processes - to get errors and inconsistencies in your data. This can be as simple as spelling mistakes, extra white space, or the wrong case used in text, to empty rows or missing values in your data, to inaccurate or duplicated data. Having these errors and inconsistencies in your data can lead to issues with formulas not working, with unsuccessful sorting and filtering operations and therefore inadequately visualized and presented data findings. These data errors and inconsistencies require you to carry out some form of data-cleaning routine to improve the quality and usability of the data. Let's start off with one of the easier of those tasks, which is spell checking. In Excel, this works in pretty much the same way as you may have already encountered in applications such as Microsoft Word or other common word processing applications. I have some data here relating to the sales of toy vehicles, and the first thing we need to do is select what data we wish to check for spelling; in this case we will try column K which contains the product line data. Then we click Spelling which is on the Review tab. Well that seems to be OK, so let's try the Country information in column T. So, we do have an error here, where a country name has been misspelt, or more likely, mistyped. We just click Change if we are happy with the spelling suggestion, or we could choose another suggestion from the list, or even ignore this error if we know the data is correct, but in this case we will change it. Here's another typo for a country name and here's one more. So, that seems to be all the errors in this column, let's try the final column now which is the deal size in column X. Here is a misspelling of the word small and another for medium. And that seems to be all for this column. The next inconsistency we will look for is empty rows. Empty rows in your data can cause lots of issues relating to moving around your data, working with formulas, and sorting and filtering. Therefore, it's very important to remove them from your data. If you remember from an earlier lesson, when we click CTRL+DOWNARROW, it should take us to the end of that column of data, but notice if we do that in this dataset, the cursor keeps stopping when it gets to an empty row meaning that the dataset if essentially being split into multiple sections, separated by these empty rows. That's not good, so let's resolve that now. We have a couple of options; one option is to just manually scroll down the sheet looking for empty rows and deleting each one, which is easy enough, and fine to do if you only have a small amount of data, but imagine if you were dealing with hundreds, or thousands, or even tens of thousands of rows? That would be a very laborious and time-consuming process. There is a much better way - which involves selecting all our data first, either using the mouse, or the CTRL+SHIFT+END keyboard shortcut. Then we select the Filter icon on the Data tab. We can now see that each column has a filter icon next to the column header. If we then select the Customer Name column-filter in column M then uncheck Select All then scroll down to the bottom of the list, we can check the item called Blanks, and then click OK. This will now show only the empty rows at the top of our sheet; this can be quite hard to see, but if you look in the row numbers, you can see that rows 28,29,65,73,74,75 and 117 are listed at the top and are highlighted in blue text. We can now select these rows, either using the mouse or going to the first cell in the first data row, which is A28, and then using the CTRL+SHIFT+END keyboard shortcut then delete the offending empty rows. We then need to clear the filter and turn it off, so we can view our data again. Now, if we go back to the first row in the top of the datasheet and try the CTRL+DOWN shortcut again, to go to the end of the data column, it will work. The next inconsistency we'll look for is duplicated rows of data; it's quite common for duplicate data rows to exist in your imported data, caused either by human input error, or an error in the import process. There are two ways of doing this in Excel; the first way