**Missing values in SPSS**

In this video I’m going to show you several options of what you can do with missing data in SPSS. So you can see here from my variable called Travel Time, measured in minutes that I have three missing values. Now anything that’s blank, by default is considered missing in SPSS, and technically we don’t have to do anything with it, we can leave it as it is and it would be omitted from any analysis that we did, whether we made a table, a chart graph or did something like regression of t-test- it would be omitted. But if you have a large data set, or a lot of missing values, it might be a good idea to code them so that you know for sure that value is supposed to be missing, because accidents happen, we might accidentally hit the keyboard and delete something we didn’t mean to delete, so it’s a good idea, its good practice to code missing values. Also, there may be more than one reason you have a missing value, especially if you’re collecting data via a questionnaire, so you might have not applicable, maybe you have skip logic, so if someone skips over a question so they couldn’t answer, and then you have people just not answering questions, so you might have multiple reasons.

So let’s look at what we can do. We can code these values and we can do it automatically with SPSS. Now if I’ve got a small data set like this and I just have three missing values. I can enter my code by hand, but if I have a lot of missing values, or a very large data set, we don’t want to do it by hand because it increases the chance of making a mistake. So, always choose a code that is not possible in your data set. I often choose things like 99 or 999, but in this case, because 99 is possible because it’s in minutes, it could take an hour and 38 minutes to commute. I’m going to choose -1 because it’s impossible to have a travel time of -1 minute. So if I go to the transform menu, and recode into same variables, so I don’t want to create a new variable, I just want to replace the missing values with a code. Now the variable I want to recode is travel time, and I’m going to click on old and new values. So the old value frame is where we indicate what the current code or situation of the variable is and the new value where we specify what we want it to be. So the values I want to recode are system missing so I’m going to choose this first option. System missing means its blank. User missing means it’s a missing value code defined by you, the user. So we want the first one here, system missing, and our value is going to be -1. Go ahead and click ‘add’, and then ‘continue’ and ‘okay.’ We can see that my missing values have been replaced with a -1. We’re part way there. If we don’t do anything else, SPSS will now count the -1s as part of our data set which we do not want. We need to define them in the variable view, so click on the variable view tab, and for travel time I want to go to the missing column, and we’re going to define -1 as a discreet missing value. Discreet just means a whole number. So I’m going to put a minus one in and click ‘okay’. Note there are three boxes here which means I can have three distinct missing value codes. If I have more than that which you often see with public surveys and surveys that you download from the internet, they might have a whole range of values, that are defined as missing, and in that case, you would choose the second option here to define a range of values, but I’ve only got one so I’m just going to choose discreet missing values and type in my missing value code and click ‘okay.’ Now if I go back to my data set, if I were to run any analysis, the -1 would tell SPSS this value is missing and it would be omitted from any analysis that I do.

END.