**Enter and define variables in SPSS**

Let’s have a look at how we get started on entering and defining variables in SPSS. For more detail about questions from a questionnaire, please see those specific videos. This is just going to give a general introduction and overview to entering data and defining variables. Now this is a new data file in SPSS. You’ll notice that we have two tabs; one is data view, this is for your data and one is for variable view and this is for your variable names. When you enter variable names in your first row, those names are going to appear across the top as the column headings replacing the var that’s there right now. So let’s have a go with that. My first column if I’ve done a questionnaire is participant ID. So I’m just going to name this variable ‘ID.’ We’ll come to these other columns shortly. Let’s have a look at our data view. You can see that my column heading now says data ID because that’s my variable name. So now in this first column here I would enter data for my ID numbers. So let’s just enter five participants, so one, two, three, four, five. Okay so now I’ve got a bit of data in. The data view is only for data, never put variable names across the first row because you do not analyse variable names. You only analyse data. Also, when you go to enter data it should be numeric data, especially if you’re going to analyse it, it’s just good practice. I know that in SPSS you can kind of get away with not coding nominal variables but its good practice to code them anyway. Because SPSS is quantative software it’s not meant for qualitative data. So let’s go to variable view. We’re going to enter some data, three variables. I’ve got my participant ID, I’m going to enter a categorical variable- gender, and I’m going to enter a scale variable, which is travel time, and that’s how long it takes someone to commute from their home into university. So I’m going to put my variable names in first. My second variable is gender and my third variable is travel time. Now there’s a couple of rules when writing variable names. One, they must begin with a letter. Two, they cannot contain a space, and three; they cannot contain any special characters except for an underscore. So for travel time I could put an underscore in-between and that would be acceptable, but I cannot put a space. The second column is defining the type of data; numeric, string, date, etc. Most of your data will most likely be numeric, occasionally you may have dates. Width- you don’t need to worry about that, that has to do with character width. Number of decimal places, this is how many decimal places you see in the data view, so all of my data doesn’t have any decimals so I’m going to reduce this to zero.

Now you may remember before that my IDs had 1.00, 2.00, had two decimal places. Now those decimal places are gone. It’s much easier to read if you don’t need those decimal places. Now the labels are really important, this is what’s going to appear on all of your outputs for tables, charts and graphs. So be a bit more descriptive. Gender, that doesn’t need to change. Time it takes to travel. It’s a good idea to specify units in your labels as well so my units are minutes. The next column is for value labels, you only need value labels for categorical data so that means you tell SPSS what the codes mean. I’m going to code gender. I’m going to let zero be male and click ‘add.’ And one is going to represent female. These are the codes that I have given. You can use one and two if you want. It doesn’t really matter; I just prefer zeros and ones. So go ahead and click ‘okay.’ My travel time is scale meaning it’s not grouped in anyway so I don’t need to put any value labels and I don’t need anything for participant ID. If you decide to have a code for missing data you can define it here so I’ve coded my missing data as 999. I would put that code in here which would mean that SPSS came across a 999, it would consider it missing. You would need to do that for every variable where you have missing values. The last column we’re going to have a look at is the level of measurement. If you’re not sure about the level of measurement please do have a look at the video I made about the level of measurement in SPSS.

Our ID, you’re never going to use the participant ID in any kind of analysis, so it really doesn’t matter what you put. I’m just going to leave it as scale because it will never be used anyway. Now gender, that is categorical: male, female. The words give it away. Even though we have numerical codes of zero and one, it’s not numeric, it’s not scale, and it’s categorical. Now we have two options for categorical data; ordinal and nominal. SPSS gives you a little bit of help here remembering which one was which. Ordinal has groupings but they have a step-order to them because ordinal data is ordered and nominal is just three bubbles and there’s no particular order because nominal data isn’t ranked or ordered. Now gender: male, female. I can’t rank males and females so it’s going to be nominal. Travel time. Now you can see from the data on the screen that travel time is just the travel time in minutes, how long it takes to commute, so there’s no groupings. That means we can leave it as scale.

Alright, let’s go to our data view and enter some data. We’re going to have a look at the data on your screen, its zero, zero, zero, one, zero, one. Now if yours has zeros and ones, don’t worry. It’s just because my data file has the value labels icon pushed in and clicked. If I untick that I can see my codes. Now this is great because you can see on your data file whether you want to see value labels or numerical codes.

Let’s enter travel time. 35, 15…Okay. Just before we close, we’re going to go back to variable view and make a few pointers. It’s really important to remember to put good value labels in. It’s really important, so the value label, this is what will appear on all of your output. If you take an extra 30 seconds to make a good label, then you don’t have to relabel any of your output. Also remember to be consistent in capitalisation and punctuation and please don’t misspell anything because it looks bad on the graph. The last thing is for missing data. Coding missing data is an option, it’s not compulsory. But if you’ve got a huge data set, or lots of missing values, it’s probably a good idea so you can keep track of your data, so if you do have a missing data code and you enter that into your data set, you can define it here. So we’ve got gender and it has a missing value code of 999.

So let’s say for example this participant, participant two didn’t specify their gender. I can then choose to code that as 999, and since I defined it in my variable view, SPSS will recognise that 999 to be missing. Okay, if you’ve got a questionnaire and you want to see any more specific question types from a questionnaire, please have a look at the relevant videos.

END.