Chapter Five: Measurements

Dr. Simona Simona

4/12/2020

This chapter will address the following topics:

- The concept of variation

- Measurements

The concept of variation

If somebody is to ask what concept anchors the whole of quantitative methods and statistics, that concept has to be variation for without variation there is no statistics. In fact, it is true to say that without it the whole of social science is null and void. Variations are basically differences in conditions among individuals or objects. The opposite of variation is constant, which means lack of variation or unable to change. Now imagine that the world was full of people of the same age, sex, socioeconomic status, health status, ethnicity, dwelling spaces, consumption, culture, dress code and everything else. What would be there to study? We would know pretty much everything there is to know and there will be no need for any study. Such a situation is of course impractical, but it is importance for us to conceive of it so that we may better appreciate the concept of variation and why it should be the starting point in all research methods courses.

Variation ensures comparability between and among individuals or objects and that is what instigates research especially in the social sciences. All individuals possess attributes that set them apart from others. Developments in science have shown how each one of us has their own biological mekeups called the genome which are specific to us and not present in any other individual to ever exist on earth. We all know how an impression left by the friction ridges of a human finger has become an important piece of criminal investigations because each person has unique fingerprints different from anyone else who has or yet to exist in the world. These are typical examples of the importance of variations. In the social science variations may not be as rigid as the genome or human fingerprints which are unique to individuals. But the differences in social attributes that exist in the world raise very interesting questions that may warrant exploration in terms of social scientific research. Variations in educational status, income levels, labour participation, age, gender, marital status, health status could raise question which make us want to compare one or two attributes. For example, we may want to compare ethinic background and educational outcomes to study whether someone's background could affect how far they go in education. We may may also want to find out whether gender inhibits of bolsters our chances for higher income. These are possible to research on because ethinicity, educational level, gender and income are not constants but vary from one person to the other. These social attributes are called **variables** and I explore them in the next section.

Variables and Variable types

Variables are attributes that can differ from one observation to the other. We have already seen what observations are in Chapter Three. They are basically entities for which we have a series of variables. They are sometimes called **cases**. In the social sciences, we typically use individual persons as observations, who would have several variables such as age, sex, employment status, ethnicity, marital status, educational status, income etc. Each variable will have **values** associated with it. Values are scores that observations will have have on a particular attribute or variable. For instance, sex as a variable may have two values, *male* and *female*. The variable 'age' is likely to have many values depending on the type of study you are conducting. You may have age values like 24, 26, 38, 54 etc. The variable 'educational status' will have values such as *primary*, *secondary* and *tertiary*.

Variables in the social sciences are often distinguished by the role they play in the research question (whether they are dependent, independent, intervening, moderating or control) and the levels at which they are measured (whether it is at nominal, ordinal, interval or ratio). In fact, before we do anything which remotely resemble data management or analysis in any software, we must have a good understanding of the types of variables we will be dealing with and how they are measured. This is because variable types have implications on the nature of analysis we are able to do. That is why we will be spending a lot of time in this chapter to cement that understanding. But before we do that let's first explore how we arrive at the variable that we are able to measure.

Conceptualisation and operationalisation of variables

Where do variables come from and why do we choose to focus on some variables are not others? Well, variables are normally a characteristic of quantitative studies and they come from the questionnaire or any data collection tool that you have used. Before that, variables are simply operational definitions and before they are operationally defined, they are concepts or conceptual definitions that you wish to consider in your study. Before they are conceptually defined, they are your research problem or research topic. This is kind of a rollercoster but I wanted to demonstrate that the linke between your research topic and variables. It is actually uncommon to find your key variables mentioned in your research topic. Some quantitative research instructors will insist that this be the case. I may be interested in studying the effects of education on health outcomes. You see here that my variables of interest (education and health outcome) are in my topic. It becomes very easy for me to remain focussed throughout the research process.

This means you are to think about your variables of interest right at the time you are thinking about your research problem. Therefore, all you have learnt so far about the sources of a research problem applies here. The overarching source of variables in quantitative research are *theories*.

Most of quantitative studies (but not all) begin with theories. Hopefully you understand by now you understand why this is the case from your studies of philosophical foundations of social science research. We use theory to explain, understand and predict social phenomena. The choice of the theory used in our research is often aligned with our world view. So even if theories may appear to pop up innocently in a research paper it is often deep-rooted in the author's own predispositions, beliefs and biases. The important thing though is that the theory that you choose determines the nature of variables you will be researching.

People who study social justice, emancipation or human rights might want to use Pierre Bourdieu's theory of capital and Amartya Sen's capability approach to ground their studues. The theory

of capital considers such variables as income, property ownership, social networks, power and educational status among others, to be important concepts and variables of research. The researcher adopting the theory is inclined to utilise such concepts and variables than one who is studying the causes and consequences of rising crime rates in urban centres.

However, sometimes research is not preconceived. Which means there are no benefits of theories from which to generate variables and drive the research agenda. For instance researchers may be part of a team tasked to execute a multifaceted project commissioned by governmental on nongovernmental agencies which has a research component. In this case, researchers do not determine the nature of variables to be studied and how they are defined. They are only involved in analysing the data with variables that are arising from the main project.

Other scenarios where variables may not be predetermed by researchers is when a study is born out of curiosity and observation of regularities. Patterns of a social phenomena can be established out of curiosity due to consistencies in observed regularities. This can be from the social media, traditional media or even from one's own observations. For example you might notice that there is an abalance of talks about fake posts on social media and you decide to do a study on it. You will not have already made variables but you engage in a study on the basis of hunch believing that you will arrive at useful information in due course

Whatever the sources may be, varibles are often difficult to pinpoint initially because theories and other sources of variables do not bring out refined variables. Theories are notoriously abstract and general. In order for the concepts derived from them to be appropriately measured, they have to be trimmed off of ambiguities that may render them less feasible in scientific studies. In this regard, variables are first given **conceptual definitions**

Conceptual definition

Conceptual definitions tell us literal meanings of concepts and variables that are key to a particular study. Conceptual definitions are important because they provide dictionary definitions of concepts so that we have a good understanding of what the concept means before collecting any data about them. We cannot measure anything without knowing what we are measuring. It is important to note though that sometimes conceptual definitions of our research concepts may depart from their everyday use and even from dictionary definition. This is fine as long as our definition is well understood within our discipline and well described in terms of how it is being applied in our research project. It is not uncommon for specific disciplines to allocate different meanings to every day concepts and even completely manufacture new ones.

In my research, I may want to define the veriable 'income' as an individual's monthly earnings. This gives me a good general picture of the meaning of the variable I am studying. The variable 'income' looks straightfoward but in research terms which requires the highest level of specificity, it can be very difficult to negotiate. If we didn't boil it down by defining it an individual's monthly income, it could have easily meant a variety of things including daily household income, monthly household income, life-time income for the head of the household etc. However, even the currently refined definition is still problematic because it is defined at the conceptual level. If I asked you to go and collect data on people's monthly income, it will not take long before you come back to me and ask for more specifications. Do I mean the net (after tax) or gross (before tax) salary? Is it income that remains after monthly expenditure of before? Is it only income that comes from a salary or can we also income other sources as well? You can see now that conceptual definitions are not adequate when it comes to actually measuring our variables. They are just the begining of definitions that

gives us a general understanding, but we need specific rules and procedures that allow us to measure income consistently from one respondent to the other. For this we need operationalise our variables.

Operational definition

If coneptual definitions tell us the meaning of variables, an operational definition tells us how to measure them. It gives clarity in terms of what we would look for when we go

Dependent vs. independent variables

Other types of variables

Levels of measurements

Measurements of variables are a cricual part of the research process. In fact, if things go wrong here, your study becomes almost irredemable.

There are basically four types of variables and these are nominal, ordinal, interval/ratio and sometimes called scale or continuous variables. Please don't be confused by these terms as that is the nature of quantitative methods. Sometimes nominal and ordinal data are called categorical.In R these refer to different nemes as we have seen in the R-environment chapter. Factors, ordered factors, numeric are the equivalent of nominal, ordinal and scale. The reason we have this confusion is basically because we are now combining two fields names computer science and statistcis.

We mentioned in the previous chapter that data management and manipulation are some of the most important skills to have in the data analyst's tool kit. The data you have is in most cases not originally collected for you but you repurpose it to address the research problem that you have for your study.

Reading in the data

As we can imagine, for us to be analyse any type of data, we first have to bring it into R. When you are a new RuSER, this process may not be that straightfoward. There are different ways of reading in data in the R environment and this dependends on the file format the original data is saved in. Data can be saved in SPSS, Stata, CSV and excel among others, can be