

What will this chapter tell me?

Years at an all-boys' school carefully nurturing a morbid fear of women and a love of heavy metal had made the world of relationships a tricky place for me to inhabit. I'd always dreamt that by my mid-thirties I would have a wife and a cute little child or two to remind me of the important things in life. However, as I took my first furtive and depressing steps into middle age I found myself single, and the closest thing to a child was a ginger cat and this book. However, something remarkable had happened since my teens: rock music had become popular again, and some women liked to talk about Iron Maiden. I needed to capitalize before the ephemeral guillotine of fashion spliced this opportunity from me. I met Zoë, who was not only happy to discuss Iron Maiden, but even owned my favourite of their albums (*Piece of Mind*). She had no aversion to statistics, or soccer, and also happened to be the most beautiful woman ever placed on the face of the earth. Result. 'I'd better marry her before she realizes I'm a balding geek with slight hoarding tendencies', I thought. So that's what we did. A little later than anticipated, my dreams had come true: I started my late thirties with a wife and a cute little ... book about a statistics package called R to which my wife contributed. Mental note to self: next time create a little human, not a big book.

Marriage is a leap of faith into the unknown, a shared adventure full of challenges. A bit like this chapter really, because, upon embarking on it, multilevel linear models were an 'unknown': I knew absolutely nothing about them. If you're reading this section then you probably don't know much about them either. So, we'll learn together – a shared adventure, and, oh boy, will it include some challenges ...

Hierarchical data

In all of the analyses in this book so far we have treated data as though they are organized at a single level. However, in the real world, data are often hierarchical. This means that some variables are clustered or *nested* within other variables. For example, when I'm not writing statistics books I spend most of my time researching how anxiety develops in schoolchildren. When I run research in a school, I test children who have been assigned to different classes, and who are taught by different teachers. The classroom that a child is in could affect my results. Let's imagine I test in two different classrooms. Mr. Nervous, who is very anxious and tells children to be careful, and that things that they do are dangerous, or that they might hurt themselves, teaches the first class. Little Miss Daredevil,¹ who is carefree, tells children not to be scared of things and gives them the freedom to explore new situations, teaches the second class. One day I go into the school with a big animal carrier, which I tell the children contains an animal. I measure whether they will put their hand into the carrier to stroke the animal. Children taught by Mr. Nervous have grown up in an environment that reinforces caution, whereas children taught by Miss Daredevil have been encouraged to embrace new experiences. Therefore, we might expect Mr. Nervous's children to be more reluctant to put their hand in the box because of the classroom experiences that they have had. The classroom is a **contextual variable**.

Also, I might tell some of the children that the animal is a bloodthirsty beast, and tell others that the animal is friendly, expecting that this manipulation will affect the children's enthusiasm for stroking the animal. However, the effect of what I tell the children happens within the context of the classroom to which the child belongs. My threat information ought to have more impact on Mr. Nervous's children than on Miss Daredevil's children. Figure 20.2 illustrates this scenario: children (or cases) is the variable at the bottom of the hierarchy, known as a *level 1* variable. These children are organized by classroom (children are said to be *nested* within classes). The class to which a child belongs is a level up from the participant in the hierarchy and is said to be a *level 2* variable.

A situation with two levels is the simplest hierarchy that you can have. You can have other layers in more complex hierarchies. If we stick with our example, an obvious third level is that classrooms are nested within schools. Therefore, if I ran a study incorporating lots of different schools, as well as different classrooms within those schools, then I would have another level to the hierarchy. We can apply the same logic as before: children in the same school will be more similar to each other than to children in different schools because schools have different teaching environments and also reflect their social demographic (which can differ from school to school). Figure 20.3 shows a three-level hierarchy: the child (level 1), the class to which the child belongs (level 2) and the school within which that class exists (level 3). In this situation we have two contextual variables: school and classroom.

Hierarchical data structures need not apply only to between-participants situations. We can also think of data as being nested within people. In this situation the case, or person, is not at the bottom of the hierarchy (level 1), but is further up. A good example is memory. Imagine that after giving children threat information about my caged animal I asked them a week later to recall everything they could about the animal. Let's say that I originally gave them 15 pieces of information; some children might recall all 15 pieces of information, but others might remember only 2 or 3 bits of information. The bits of information, or memories, are nested within the person and their recall depends on the person. The probability of a given memory being recalled depends on what other memories are available, and the recall of one memory may have knock-on effects for what other memories are recalled. Therefore, memories are not independent units. As such, the person acts as a context within which memories are recalled (Wright, 1998). Figure 20.4 shows this scenario: the child is our level 2 variable, and within each child there are several memories (our level 1 variable). Of course we can also have levels of the hierarchy above the child, for example, the class from which they came could be a level 3 variable. Indeed, we could even include the school again as a level 4 variable. A common situation in which cases are a contextual variable is when we take several measures over time (i.e., a repeated-measures design). In this situation measures at different points in time (level 1) are nested within cases (level 2). We look at this situation in detail in Section 20.7.

¹ Those of you who don't spot the Mr. Men references here, check out <http://www.mrmen.com>. Mr. Nervous used to be called Mr. Jelly and was a pink jelly-shaped blob, which in my opinion was better than his current incarnation.