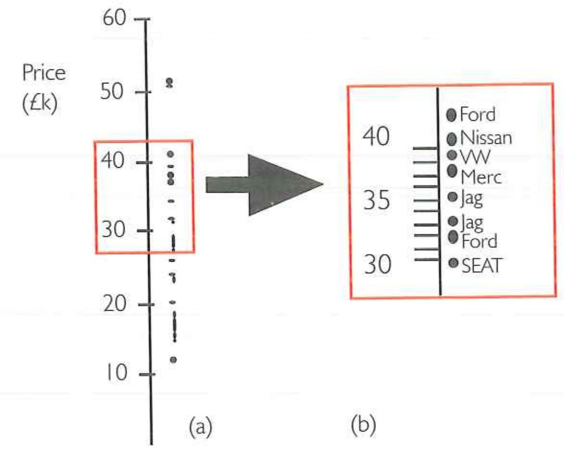


FIGURE 3.21 Semantic zoom reveals data about a second attribute

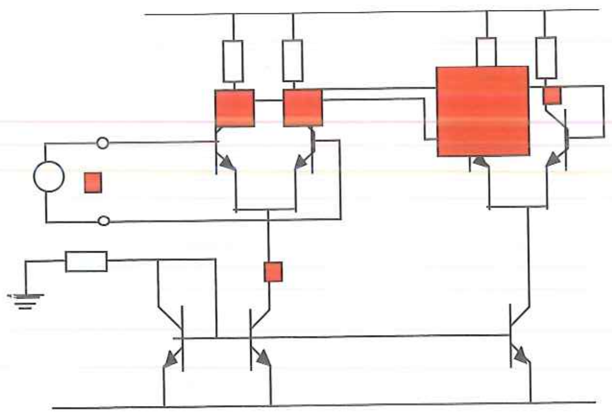


FIGURE 3.22 The area of each red square encodes the value of the voltage occurring at the point in the circuit at which the square is located

been displayed in numerical form. The difference between the two red squares 'pops out'. When numerical values are of interest, a simple mouse click or mouse-over can always replace the red squares with the corresponding numerical values.

A similar technique exists for representing value in such a way as to afford a qualitative understanding. Many people have a mental model of the representation of Australia and New Zealand on a map of the world (Figure 3.23), so that the fact that New Zealanders possess ten times as many bicycles as Australians (if that were true!) could be encoded by correspondingly magnifying the representation of New Zealand on a map, as shown in Figure 3.24. That such magnification encoding can lead to useful insight is demonstrated by *The State of the World Atlas* (Smith, 1999) of which one page, concerned with population density, is shown in Figure 3.25. The familiar large land mass normally located at top left is now replaced by a thin strip representing Canada's low population density, as is the case with Australia. Other pages in this atlas are equally illuminating.

3.1.3 Trivariate data

Since we live in a three-dimensional world there is a temptation to represent data about objects characterized by three attributes by points in three-

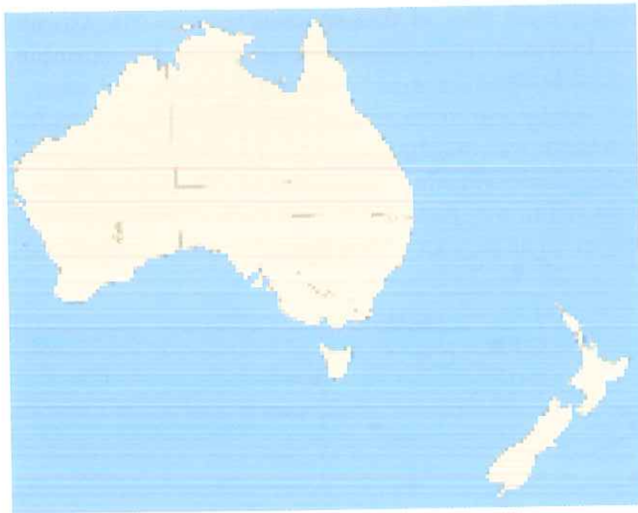


FIGURE 3.23 A representation of Australia and New Zealand on a conventional map

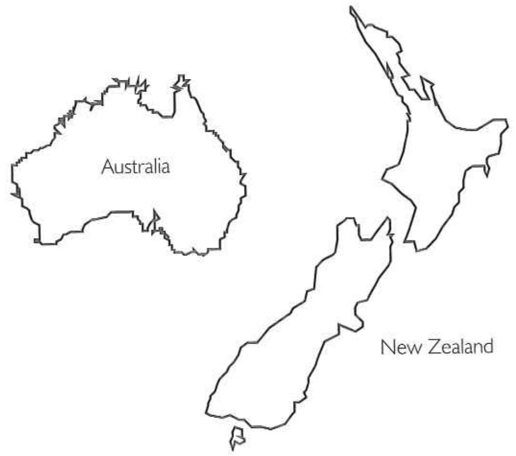


FIGURE 3.24 A representation of Australia and New Zealand indicating that some attribute of New Zealand is ten times its value for Australia

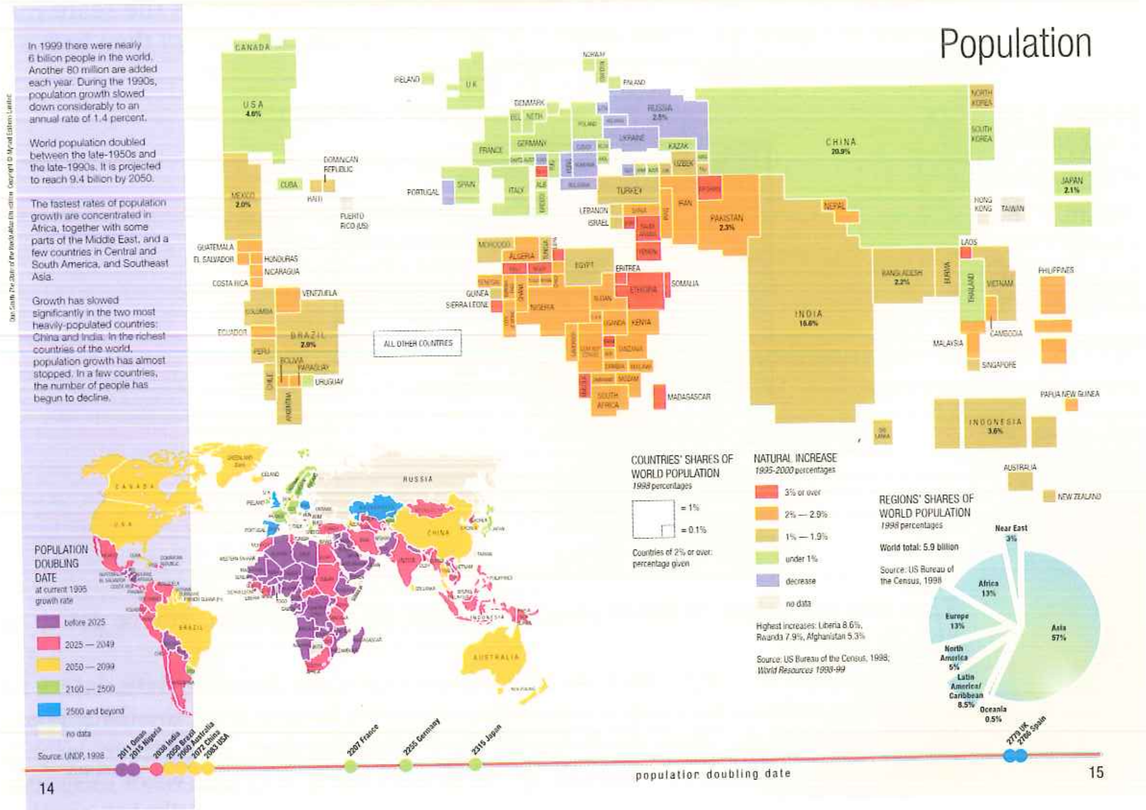


FIGURE 3.25 In *The State of the World Atlas*, magnification encoding is used to give a first impression of population densities. Note the reduced 'size' of Canada and Australia when compared with a conventional map
Source: *Smith (1999)*

dimensional space and to display a 2D view of that space (Figure 3.26). A problem is immediately apparent, however. It is impossible to know, for example, whether house A costs more than house C.

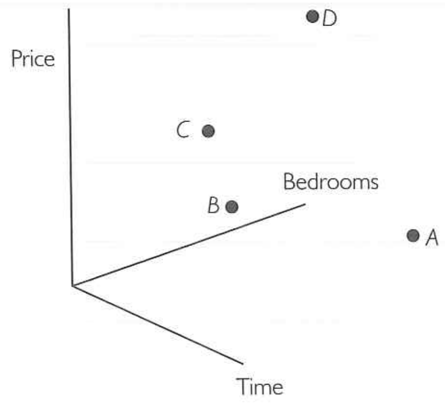


FIGURE 3.26
It is impossible, with a '3D' representation of data, to compare the attributes of Price, Number of bedrooms and Journey time to work values

Scatterplot matrix

One solution is again to bring interaction to the problem and allow a user to reorient the 3D representation so that, for example, the view is the scatterplot shown in Figure 3.27 from which it is clear that house A costs less than house C.

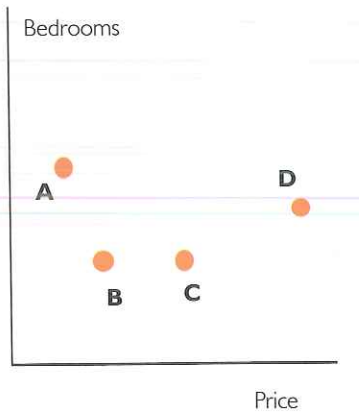


FIGURE 3.27
A projection of the data shown in Figure 3.26, allowing comparison of *price* and *bedrooms* values

This was the origin of the comment that 'for 3D to be useful, you've got to be able to move it'.

For trivariate data – and, as we shall soon see, for hypervariate data – an alternative representation is a systematic structure formed from the three possible 2D views of the data, as shown in the scatterplot matrix in Figure 3.28: in essence, the user has reoriented the 3D plot in three different ways. Although two houses can now be compared easily with respect to three attributes, another problem is present. There are now three times as many points as houses, potentially increasing the cognitive load on the user, and often there will be no room

for labels. Here again interaction can offer a solution by allowing the user to identify, for example, low-cost houses in the Price–Time plane, whereupon the same houses are highlighted in some way in the other two planes (Figure 3.29). Again, we are *brushing* houses from one plane to the other two. But, mindful of the phenomenon of change blindness, we must ensure that the user *notices* the highlighting. Thus, rather than simply changing the colour of a small dot from black to red in all three scatterplots, we might choose to highlight by magnification – even if only temporarily – as well as by colour, as shown in Figure 3.29.

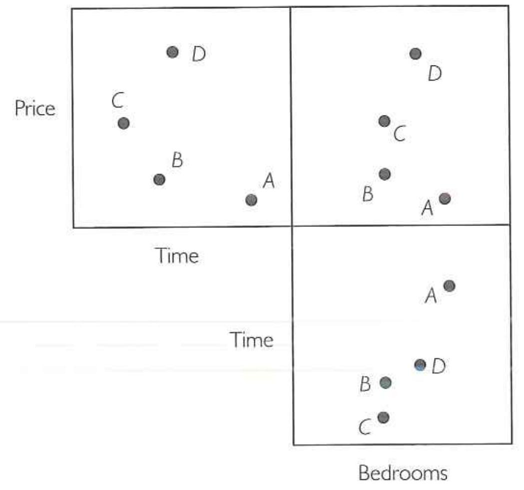


FIGURE 3.28
The scatterplot matrix associated with the data of Figure 3.26

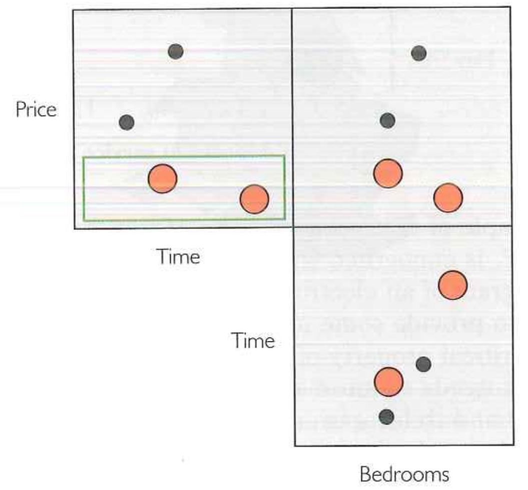


FIGURE 3.29
The highlighting of houses in one plane is brushed into the remaining planes

The above example does not imply that three scatterplots comprising a scatterplot matrix are needed for trivariate data, a fact that can be illustrated by the representation of product fault data. Almost every mass-produced product designed for domestic use will encounter the problem of failure of some kind due to a variety of causes. The fact that a car radio, for example, has failed is easy to record but, and especially in the absence of the faulty product, it is a very expensive matter to find the cause of the fault. Information visualization can help, as shown in the representation of Figure 3.30. Two attributes represent independent variables: one is the month in which the product was manufactured and the other is the number of months it had been in service before the fault occurred. The radius of each circle indicates the number of products manufactured in a particular month and which had been in service for a specific number of months before failing. This representation alone is of considerable help in identifying the types of fault that may be occurring. For example, if many faults occur for one particular month of production irrespective of how long after production the fault occurs (as suggested by Figure 3.30), they are characterized as 'epidemical failure' and can guide the production manager in detecting the underlying cause. If, however, many faults are located on the rising diagonal (again as is the case in Figure 3.30) then a seasonal influence is indicated, since points along such a diagonal refer to products which all fail at roughly the same time.

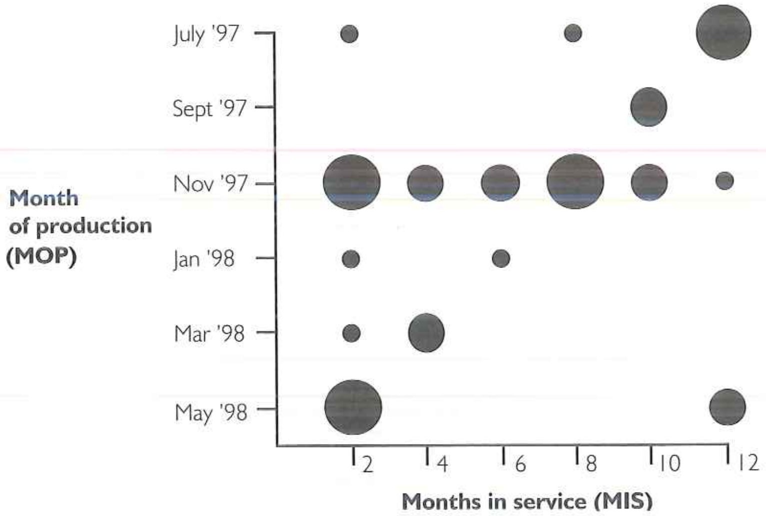


FIGURE 3.30 A representation of reported product failure, based on month of production (MOP) of the failed product and total months in service (MIS) before the fault occurred. The radius of each circle indicates the number of faults reported for a given MOP and MIS

Another example of the representation of trivariate data, and which is similar to Figure 3.22, is concerned with a property of an electronic circuit. Figure 3.31 shows a diagram of an electronic circuit in which circles are superimposed on components to provide some idea as to their importance for – that is, their effect upon – a critical property of the circuit. But if the circuit is a hi-fi amplifier, the designer needs to know this information at a number of frequencies between the bass and treble extremes. It is a simple matter to animate the representation, with circle sizes changing accordingly as a pointer moves continuously up and down the frequency scale.

A special category of trivariate data is to be found in maps showing the location (latitude and longitude) and value of some object or attribute. Figure 3.32 shows the population of major cities in England, Wales and Scotland. By its use of circle size to encode each city's population, we can quickly gain an impression of how population is distributed. We might say that the information in this figure 'pops out'. We do not have to examine numbers to discover where the largest city is and what the relative sizes of the populations are. Thus, if we can arrange that information of interest to a user 'pops out' *without* the need for

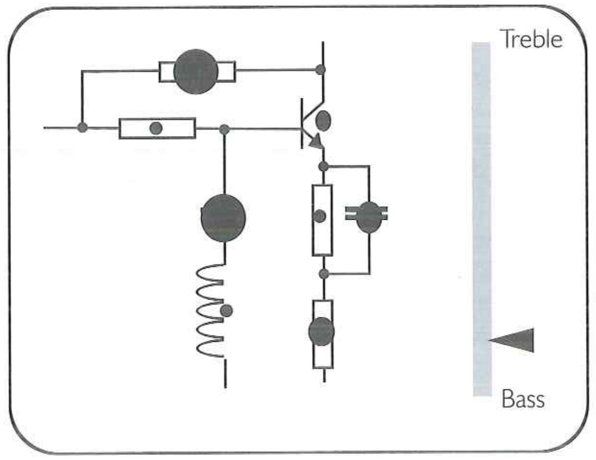


FIGURE 3.31 Circles indicate the extent of the effect of a component on some property of the circuit and change in size as the frequency cycles up and down the range from bass to treble

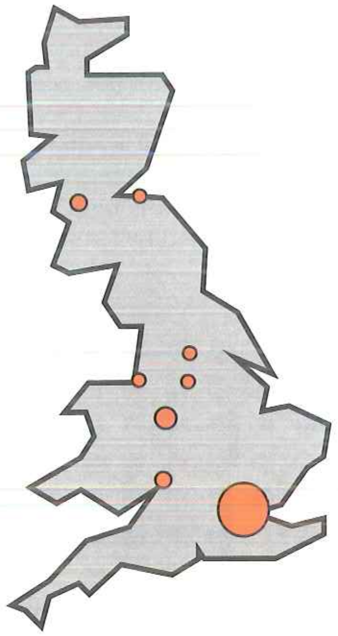


FIGURE 3.32 A representation of the population of major cities in England, Wales and Scotland. Circle area is proportional to population

cognitive effort, then the user is at a huge advantage compared with a situation wherein careful attention has to be given to acquire that insight. How information can be encoded to result in essentially immediate comprehension is closely associated with the human visual system and is discussed immediately below.

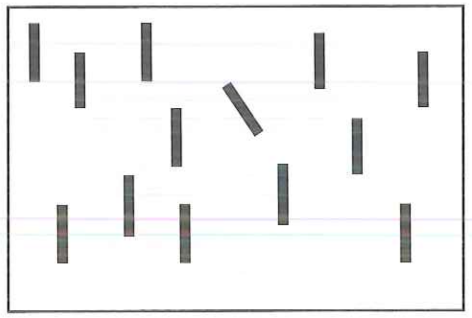
Pre-attentive processing: 'things that pop out'

The desirability of information 'popping out', without the need for careful scrutiny, has been mentioned frequently. As Colin Ware (2004) remarks, 'We can do certain things to symbols to make it much more likely that they will be visually identified even after a very brief exposure. Certain shapes or colours "pop out" from their surroundings. The theoretical mechanism underlying pop-out is called **pre-attentive processing** because, logically, it must occur prior to conscious attention.' The 'very brief exposure' mentioned by Ware is typically in the region from 30 to 300 milliseconds. The importance of pre-attentive processing is reflected in Ware's comment (Ware, 2004, page 149) that:

> *An understanding of what is processed pre-attentively is probably the most important contribution that visual science can make to data visualization.*

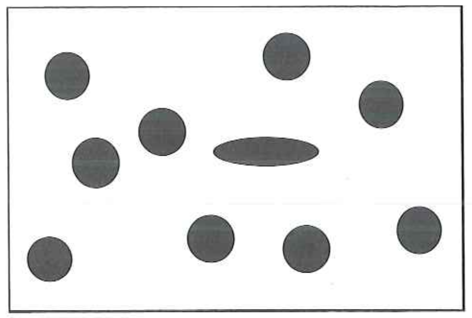
In view of the immense advantage of deriving insight without the need for conscious attention, an obvious question for the interaction designer is, 'How do we *make* information pop out?' An in-depth discussion of pre-attentive processing in the context of the human visual system is available elsewhere (Ware, 2004): here we examine some examples which are sufficient to provide an initial understanding for an interaction designer.

A glance at Figure 3.32 immediately establishes where, in the United Kingdom, the greatest density of city population is to be found (and, for users familiar with a map of the UK, the fact that it *is* a map of the UK). Similarly, in Figure 3.33, the 'odd one out' can quickly be identified, as it can in Figures 3.34 and 3.35. Pre-attentive processing also allows us to quickly identify where, in Figure 3.36, the blue square is located. A cautionary example, however, is provided by the task of locating, in Figure 3.37, the red square in a collection of items which can be either red or blue *and* a square or a circle. Here, pre-attentive processing does not occur; careful scrutiny is required to locate the red square. This is because something which is both square (as opposed to circular) and red (as opposed to blue) is identified by 'conjunction encoding'.



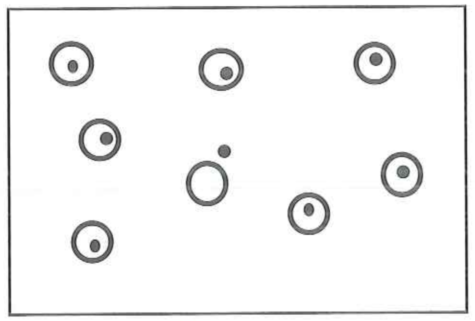
Orientation

FIGURE 3.33 The 'odd one out' can quickly be identified, by pre-attentive processing



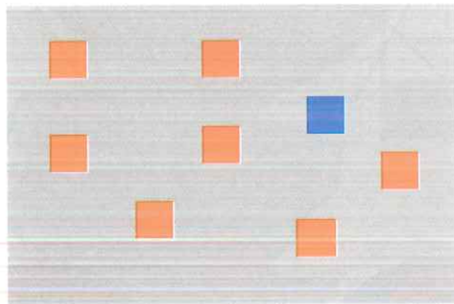
Shape

FIGURE 3.34 Different shapes can often pop out



Enclosure

FIGURE 3.35 A single lack of enclosure can quickly be identified pre-attentively



Colour

FIGURE 3.36 A different colour can be pre-attentively identified

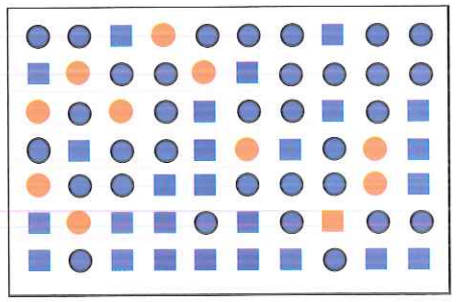


FIGURE 3.37 With conjunction encoding the red square is not pre-attentively identified

The last example does not imply that what might be called 'multiple pop-outs' are not possible. Evidence is provided by a novel representation (Irani and Eskicioglu, 2003) of mobile telephone network cell performance. Each cell is represented as shown in Figure 3.38, with colours and dimensions encoding critical performance measures. The outer boundary of a cell representation is chosen to allow their 'packing', as shown in Figure 3.39. As soon as a specialist user decides upon a specific interest – new call blockage rate, for example – the relevant information pops out without interference from other attribute encodings. An interesting and advantageous aspect of the chosen cell representation is that the potential for and general location of additional cells 'pop out' as a result of the arrows, as seen in the upper region of Figure 3.39.