Methods of geographical analysis
N. Spence with A. Owens
GY1 148, 2790 148
2011

Undergraduate study in
Economics, Management,
Finance and the Social Sciences

This is an extract from a subject guide for an undergraduate course offered as part of the University of London International Programmes in Economics, Management, Finance and the Social Sciences. Materials for these programmes are developed by academics at the London School of Economics and Political Science (LSE).

For more information, see: www.londoninternational.ac.uk
This guide was prepared for the University of London International Programmes by:
Professor N. Spence, PhD, LSE, Professor of Geography, Queen Mary, University of London.
Formerly Lecturer, then Reader in Economic Geography at LSE.
Dr A. Owens, PhD, Kings College London; Lecturer in Geography, Queen Mary, University of London.

This is one of a series of subject guides published by the University. We regret that due to pressure of work the authors are unable to enter into any correspondence relating to, or arising from, the guide. If you have any comments on this subject guide, favourable or unfavourable, please use the form at the back of this guide.
## Contents

<table>
<thead>
<tr>
<th>Chapter 1: Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aims and objectives of the course</td>
<td>2</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>3</td>
</tr>
<tr>
<td>Syllabus</td>
<td>3</td>
</tr>
<tr>
<td>The structure of this guide</td>
<td>3</td>
</tr>
<tr>
<td>Reading advice</td>
<td>4</td>
</tr>
<tr>
<td>Essential reading</td>
<td>5</td>
</tr>
<tr>
<td>Further reading</td>
<td>5</td>
</tr>
<tr>
<td>Online study resources</td>
<td>6</td>
</tr>
<tr>
<td>Computer-based work</td>
<td>7</td>
</tr>
<tr>
<td>Examination advice</td>
<td>8</td>
</tr>
<tr>
<td>Chapter 2: The nature, purpose and practice of geography</td>
<td>11</td>
</tr>
<tr>
<td>Essential reading</td>
<td>11</td>
</tr>
<tr>
<td>Further reading</td>
<td>11</td>
</tr>
<tr>
<td>Aims of the chapter</td>
<td>11</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>11</td>
</tr>
<tr>
<td>Why start here?</td>
<td>11</td>
</tr>
<tr>
<td>Geography as science</td>
<td>13</td>
</tr>
<tr>
<td>Induction</td>
<td>15</td>
</tr>
<tr>
<td>Positivism</td>
<td>15</td>
</tr>
<tr>
<td>The cultural turn in human geography</td>
<td>17</td>
</tr>
<tr>
<td>Postmodernism and geography</td>
<td>20</td>
</tr>
<tr>
<td>A reminder of your learning outcomes</td>
<td>21</td>
</tr>
<tr>
<td>Sample examination questions</td>
<td>21</td>
</tr>
<tr>
<td>Chapter 3: Geographical data: human and physical, quantitative and qualitative, primary and secondary</td>
<td>23</td>
</tr>
<tr>
<td>Essential reading</td>
<td>23</td>
</tr>
<tr>
<td>Further reading</td>
<td>23</td>
</tr>
<tr>
<td>Aims of the chapter</td>
<td>23</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>23</td>
</tr>
<tr>
<td>Quantitative and qualitative data</td>
<td>24</td>
</tr>
<tr>
<td>A reminder of your learning outcomes</td>
<td>41</td>
</tr>
<tr>
<td>Sample examination questions</td>
<td>41</td>
</tr>
<tr>
<td>Chapter 4: Presenting data and mapping geographical distributions</td>
<td>43</td>
</tr>
<tr>
<td>Essential reading</td>
<td>43</td>
</tr>
<tr>
<td>Further reading</td>
<td>43</td>
</tr>
<tr>
<td>Aims of the chapter</td>
<td>43</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>44</td>
</tr>
<tr>
<td>Presenting geographical data</td>
<td>44</td>
</tr>
<tr>
<td>Mapping geographical distributions</td>
<td>53</td>
</tr>
<tr>
<td>A reminder of your learning outcomes</td>
<td>57</td>
</tr>
<tr>
<td>Sample examination questions</td>
<td>58</td>
</tr>
<tr>
<td>Chapter 5: Descriptive statistics in geographical applications</td>
<td>59</td>
</tr>
<tr>
<td>Essential reading</td>
<td>59</td>
</tr>
<tr>
<td>Further reading</td>
<td>59</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Methods of geographical analysis is exactly that!

It is a course that is designed to provide an introduction to some of the basic analytical methods employed by geographers when undertaking their research work. These days most degree programmes will include similar courses in their structures. There will be courses (and textbooks) on, say, methods of sociological analysis or methods of biological analysis and these will have their own particular traits depending on whether the subject matter stems from human society or the natural world. Those involving a specific focus on society will be further differentiated according to whether the main emphasis is on the arts and humanities, or more on the social sciences. Those involving the natural world will, too, be differentiated according to whether the physical sciences or the natural sciences provide the emphasis. This is not to say that the methods used in each of these categories cannot be used in another — they can. An elementary text on statistics for biologists will not be that much different, except by way of examples, to the same for statistics for sociologists, for example. In the same way some social sciences, such as economics, might well employ mathematical methods that would not be out of place in the physical sciences. As might be expected, the methods that might be commonplace in the arts and humanities subjects rarely cross over into the armoury of subjects that may be grouped under the collective ‘sciences’, be they social or natural or physical.

Why is it necessary to make these points in an introduction to geographical methods of analysis? It is important that you know what you are going to encounter here. Geography is a subject that straddles the divide between the natural and the social world — this is one of the reasons why it is such an important subject. It is an interface subject between society and the natural world. In order to take up this position it has to have practitioners (geographers) who can engage in serious research in both the natural and the social sciences and, because the social sciences are insufficient to provide every analytical nuance on society, we have also to add research in arts and humanities. Put simply, in geography we have researchers who regard themselves as human geographers and those who are better described as physical geographers. And indeed there have been some — not many — who would like to think that they could wear both hats. Again, put simply, if you have different types of geographer you tend to have different types of methods that they use in their analyses.

So in geography you have the unusual situation of having to become familiar with a whole variety of methods that would not be necessary if you were, say, solely an arts and humanities student or a student from the natural sciences. As a result there is more to learn. This course will provide an overview of the range of methods — some strictly scientific, some much more interpretative even artistic — that contemporary geographers use in their everyday work. You might well use some of them first hand when it comes to, say preparing an independent geographical project or dissertation that you might well opt to do later in this degree.

It is fair to say that physical geographers, in the main, use methods from science and ‘the scientific method’ (and we will see later exactly what this really entails). Some human geographers are also perfectly at home
Aims and objectives of the course

This course aims to offer a basic introduction to some of the methods (and philosophies) that are used in geographical research. The approach here will be wide ranging, attempting to cover the many ways in which research can be undertaken in both physical geography (a natural science) and human geography (some say a social science or, for others, a branch of arts and humanities). The specific aims of the course are to:

• consider the nature, acquisition, presentation and mapping of geographical data
• describe qualitative as well as quantitative approaches
• consider descriptive and inductive statistics, as well as broader modelling themes and the interpretation of qualitative information in an operational geographical context, using real world data and simple computer packages.

04a Statistics 1 must be taken before or at the same time as this course. It provides a thorough grounding in the basics of statistical analysis in a broad social science context. Ideally the two courses should be studied in parallel.

Activity

Although it may seem a little odd to start an activity before we reach the first substantive chapter or have even mentioned the reading for the course, it is good for us to consider some initial common sense aspects of the nature of research, the nature of geography and the nature of geographical research, for this is what the methods that we are going to study in this course are for.

Read pages 1–5 of Chapter 1 in the book by Kitchin and Tate (2000) referenced below. Next write a short note answering the following three (deceptively) simple questions that these authors pose:
1. What is research?

2. Why do research?

3. What's unique about geographical research?

Finally, think about what you have written in the context of what you think research in physical geography is all about (for the above-cited textbook admits to being mainly about research in human geography). (Note Table 1.1 in Kitchen and Tate will be a help because the authors list examples of studies in physical geography as well as human geography and, some would say more interesting, in mixed human and physical geography). These are simple ‘starter for one’ type questions that will put you in the right frame of mind for what is to come. The questions that you might ask of someone (at a party!) who calls themselves ‘a geographer’ will also stand you in good stead when we come to think about interviewing methods later in the course.

**Learning outcomes**

What are the expected learning outcomes of this course? You should have:

- a broad overview of the various philosophies and methodologies that have been used in both human and physical geography and how these have evolved over time
- developed an understanding of the nature and limitations of data used in geographical analysis
- acquired skills in the presentation and mapping of geographical data
- gained experience in applying statistical and some other quantitative methods, as well as qualitative methods, in real geographical contexts and in interpreting the results.

**Syllabus**

The syllabus consists of the following topics:

- Evolution of the principal methodological and philosophical approaches to the study of human and physical geography.
- Quantitative and qualitative data collection in human and physical geography from primary and secondary sources.
- Data presentation and mapping geographical distributions and relationships.
- Descriptive statistics in geographical applications.
- Modelling systems and relationships in real geographical contexts.
- Interpreting qualitative data, case studies and ethical considerations in human geography.

**The structure of this guide**

This **Introduction** comprises Chapter 1 of the guide and sets the scene for the tasks to be covered.

Next is Chapter 2 on the **nature, purpose and practice of geography**. It would also be good to look at these ideas in parallel with the current companion guide **09 Human geography** prepared by Dr Gareth Jones because this, too, contains (in Chapters 2 and 3) a treatment of the nature, scope and methods of geography, with, of course, an emphasis on the human side of the subject.

---

2 Note that this Chapter 2 is a much shortened and edited version of an excellent University of London subject guide by Dr David Demeritt now out of print. Dr Demeritt bears no responsibility for the selection of the passages of his writings chosen here and the heavy editing that has had to be employed. Students who are keen to benefit from his work in greater detail are encouraged to search out his journal publications in this area.
Chapter 3 is about the raw material of the methods to be used, namely geographical data: human and physical, quantitative and qualitative, primary and secondary.

Once collected, geographical data needs to be displayed and Chapter 4 deals with presenting data and mapping geographical distributions.

Chapters 5 and 6 begin our treatment of quantitative analysis first by way of descriptive statistics in geographical applications and then modelling systems and estimating relationships. It is in these two chapters especially that reference will be made to subject guide 04a Statistics 1. The principal idea is to take the statistical theory as given and provide geographical examples to understand its usefulness in interpreting real world data.

Finally, Chapter 7 — Qualitative methods in human geography: depth interviews, observation and ethics — will be of special interest to human geographers. This part of the subject guide has been written by Dr Alastair Owens and usefully brings us back to where we started: contemporary methodological trends in the subject (human aspects).

At the end of the subject guide there is a Sample examination paper and guidance on answering that paper.

Reading advice

The reading specified for this course is unlike that for any other course that you might study. It has been deliberately restricted to a handful of key texts. There will be many fewer than usual references to specific journal articles. However, the emphasis is on methods here and it is anticipated that you will be learning by doing most of the time, although this will inevitably involve reference to and reading the key texts.

Most of you will be studying this course as part of a BSc Geography and Environment degree and we would expect that you are studying course 04a Statistics 1 at the same time. You should have your copy of the subject guide for this course at hand for the tasks to be considered in Chapters 5 and 6 (and it will also be of help in the discussion about data in Chapters 3 and 4).

From time to time you may be sent off to look at a journal and you are reminded that all International Programmes students have free access to the University of London Online Library where you will find a selection of geographical journals providing the full text or an abstract of papers that they contain (see Online resources below).

Finally, internet sources are invaluable in providing, often official, information for geographers to use. Again from time to time you will be referred to some representative sites to access data or other information. If you search out your own sites and sources (which you are encouraged to do) always be aware that there is little quality control on the information publicly available on the internet. It is always worth asking questions about quality of data provided and motives for providing them.

Unless otherwise stated, all websites in this subject guide were accessed in May. We cannot guarantee, however, that they will stay current and you may need to perform an internet search to find the relevant pages.
Essential reading

There is no single textbook for this course. The following books are all very useful and would occupy pride of place in any beginning geographer’s library, furthermore they will be useful for other geographical courses that you might take.

If you purchase one book it should be *The Dictionary of Human Geography* by Johnston et al. (and do not worry if you think that your heart lies in physical geography because this book also has many entries that will be of direct relevance).


It makes no sense to buy both the Rodgerson and the Wheeler et al. texts for they are both essentially statistics books and cover much the same ground. One of these would be useful to buy and, if it helps, more use is made of Wheeler et al. in this subject guide:


or


It also makes no sense at this stage to buy both the Kitchin and Tate, and the Flowerdew and Martin (see Further reading) texts for they are both about doing geographical research and cover broadly the same ground. Both are ideal for those of you thinking about undertaking a project in geographical research, an option for later on in the degree. More use is made of Kitchin and Tate in this subject guide:


Detailed reading references in this subject guide refer to the editions of the set textbooks listed above. New editions of one or more of these textbooks may have been published by the time you study this course. You can use a more recent edition of any of the books; use the detailed chapter and section headings and the index to identify relevant readings. Also check the VLE regularly for updated guidance on readings.

Further reading

Please note that as long as you read the Essential reading you are then free to read around the subject area in any text, paper or online resource. You will need to support your learning by reading as widely as possible and by thinking about how these principles apply in the real world. To help you read extensively, you have free access to the virtual learning environment (VLE) and University of London Online Library (see below).

Other useful texts for this course include:


Online study resources

In addition to the subject guide and the Essential reading, it is crucial that you take advantage of the study resources that are available online for this course, including the virtual learning environment (VLE) and the Online Library.

You can access the VLE, the Online Library and your University of London email account via the Student Portal at:

http://my.londoninternational.ac.uk

You should receive your login details in your study pack. If you have not, or you have forgotten your login details, please email uolia.support@london.ac.uk quoting your student number.

The VLE

The VLE, which complements this subject guide, has been designed to enhance your learning experience, providing additional support and a sense of community. It forms an important part of your study experience with the University of London and you should access it regularly.

The VLE provides a range of resources for EMFSS courses:

• Self-testing activities: Doing these allows you to test your own understanding of subject material.
• Electronic study materials: The printed materials that you receive from the University of London are available to download, including updated reading lists and references.

• Past examination papers and Examiners' commentaries: These provide advice on how each examination question might best be answered.

• A student discussion forum: This is an open space for you to discuss interests and experiences, seek support from your peers, work collaboratively to solve problems and discuss subject material.

• Videos: There are recorded academic introductions to the subject, interviews and debates and, for some courses, audio-visual tutorials and conclusions.

• Recorded lectures: For some courses, where appropriate, the sessions from previous years' Study Weekends have been recorded and made available.

• Study skills: Expert advice on preparing for examinations and developing your digital literacy skills.

• Feedback forms.

Some of these resources are available for certain courses only, but we are expanding our provision all the time and you should check the VLE regularly for updates.

Making use of the Online Library

The Online Library contains a huge array of journal articles and other resources to help you read widely and extensively.

To access the majority of resources via the Online Library you will either need to use your University of London Student Portal login details, or you will be required to register and use an Athens login: http://tinyurl.com/ollathens

The easiest way to locate relevant content and journal articles in the Online Library is to use the Summon search engine.

If you are having trouble finding an article listed in a reading list, try removing any punctuation from the title, such as single quotation marks, question marks and colons.

For further advice, please see the online help pages:
www.external.shl.lon.ac.uk/summon/about.php

Computer-based work

You will be asked to undertake some practical statistical work in this course using a computer. For operational reasons, it has been decided to use the most widely available statistical package, namely that associated with the Microsoft Office spreadsheet product MS Excel. Of course, other statistical packages could have been chosen and you will notice that some of the recommended textbooks have made different choices – such as the Statistical Package for the Social Sciences (SPSS) or MINITAB. It makes no difference, and if you have these packages by all means experiment with them.

Some help will be provided later in using the MS Excel software and reference might need to be made from time to time to the associated HELP facility. It should be noted that it is not the intention of this course to introduce you to understanding your PC or to operationalising computer packages, nor will we spend much time on the workings of spreadsheets,
but there are literally dozens of contemporary texts available on the market that can provide excellent back-up if you feel that you need it. The ‘Dummies’ series of texts helpful in this respect. They are approachable, well-paced, and often highly amusing. For a full list of what is available you can refer to the John Wiley website: www.dummies.com

**Examination advice**

**Resources**

This subject guide provides examples of the types of question that can be set on the various parts of the course and at the end of the guide, you will find a full Sample examination paper.

Furthermore the guide provides an Examiners’ report on this sample examination paper so that you have some idea as to what the Examiners are looking for in your answers.

**What are Examiners looking for?**

We will come in more detail to these important issues later but here are just a few pointers as to what formats will be possible and what examiners are looking for in each part of the syllabus.

For the **nature, purpose and practice of geography** you could be asked an essay-type question to be answered solely on the basis of what you have read, or the task might be to say in a few sentences what the essence is of a series of key methodological and philosophical terms or phrases.

In examining **geographical data: human and physical, quantitative and qualitative, primary and secondary** you could be asked a series of short questions about the nature of and sources of geographical data, perhaps asked to design some questions for a questionnaire or interview schedule, or write a more discursive essay-type piece on your reading about data in general.

In **presenting data and mapping geographical distributions**, again essay-type questions about the nature and techniques of mapping are possible, or you might be given a map or some other representation of geographical data to annotate and comment upon.

When it comes to **descriptive statistics in geographical applications**, the main emphasis will not be on the statistical theory or detailed calculations (as it will be in **04a Statistics 1**) but you are highly likely to be asked to interpret some previously calculated results (often in the form in which they are generated from the computer package). Here the emphasis will be on the real world meaning of what the analytical findings tell us (just how useful are these methods in informing geographical understanding). Of course, to do this well you will need to understand the statistical methods that you have studied and must have benefited from generating some practical results yourself from the course material.

Much the same can be said for the more interesting (or at least less descriptive) **modelling systems and estimating relationships** section of the course.

Finally, when it comes to **qualitative methods in human geography: depth interviews, observation and ethics**, several opportunities for the setting of examination questions arise. In straightforward terms it is always an option to ask you to write an essay
about your reading around topics, such as ethics or the range of qualitative data that can be sourced, or perhaps to discuss the pros and cons of qualitative data versus quantitative data. Or you might be asked to do some practical work analysing and interpreting, say, a depth interview or be asked to think about how to design a participant observation study or identify the ethical implications of a particular research project.

Throughout, the emphasis will be on the geographical. The examination paper will attempt to cover practical interpretative skills (where a knowledge of technique will be critical) as well as more discursive writing skills in the more conventional essay-type questions. Examples from both the human and physical sides of the subject will be covered. Note that you will not be asked to do arithmetic calculations, except perhaps save for those that you might do in your head!

**Format of the examination paper**

**Important:** the information and advice given in the following section are based on the examination structure used at the time this guide was written.

Please note that subject guides may be used for several years. Because of this we strongly advise you to always check both the current Regulations for relevant information about the examination, and the VLE where you should be advised of any forthcoming changes. You should also carefully check the rubric/instructions on the paper you actually sit and follow those instructions.

As this is a half-course, the examination will be an unseen written examination lasting for **two hours**.

There will be six questions, basically one on each component of the course syllabus, although it will be possible for Examiners to ask questions that will require comparison across the syllabus. **Three questions have to be answered, one of which is compulsory.** The compulsory question will be drawn from the first part of the course on the nature, purpose and practice of geography (Chapter 2). This component provides such a broad remit that it is likely that you will be able to draw upon your experience of other components in the course to answer this question.

Therefore you have roughly 40 minutes to answer each question. The Examiners are looking for breadth of work and understanding across the syllabus and this is why you will have to answer questions on at least half of the syllabus.

The emphasis as always should be on the quality and not the quantity of the answer. Several of the questions could require you to write a series of short answers instead of a longer piece of text.

In several senses the preparation for this course may well be different to that for your other courses. There needs to be a greater focus on breadth rather than depth, although of course the really successful examination papers will include some of the latter as well as the former!

You will find a Sample examination paper at the end of this subject guide together with an associated sample Examiners’ report.

Remember, it is important to check the VLE for:

- up-to-date information on examination and assessment arrangements for this course
- where available, past examination papers and Examiners’ commentaries for the course which give advice on how each question might best be answered.
Notes
Chapter 2: The nature, purpose and practice of geography

Essential reading


Further reading


Aims of the chapter

Before tackling the main methods used in geographical analysis it is important to have at least some understanding of the nature, purpose and practice of geography. Geography – encompassing both the human and physical aspects – is a diverse and heterogeneous discipline.

Learning outcomes

By the end of this chapter, and having completed the relevant reading and activities, you should be able to:

• understand that there is no single and unified philosophical approach to the academic subject of geography but attempts have been made to group and devise a chronology for the various paradigms
• realise that much of geography is ‘science’ and that being a science means that specific objectives and method have to be adhered to, although variations in approach are to be seen
• see, for human geography, that in recent years various approaches have been developed that form part of the so-called ‘cultural turn’ and that these too are linked to specific objectives and method
• see, again for human geography, that the most recent popular advances are approaches that can be described as ‘postmodern’ where specific objectives and method seem to be more to do the problem than the solution
• realise that, in practical terms, a language of terminology must be understood before proper understanding of the meaning of these approaches can be acquired.

Why start here?

In this introduction, we will briefly consider a range of theoretical approaches to understanding geographical knowledge, from positivism and empiricism, to hermeneutics, Marxism and postmodernism. These are all terms defined in the essential textbook by Johnston (2000). This is a brilliant work of
Methods of geographical analysis

reference and further reading that not only is valuable here but will prove a ‘goldmine’ for materials relevant to courses in human geography that you will take in future.

It is important for you to understand such philosophical approaches to geography for a number of reasons. First, philosophy provides a foundation for the specific knowledge claims advanced by the various sub-disciplines of geography. As such, this should help you put your reading for the other geography courses into a wider disciplinary and philosophical context. It will help you to appreciate that geography is a complex discipline in which debate and disagreement about the nature of geographical knowledge and the methodology for achieving it are normal.

It is important to understand that what follows in this chapter is only an introduction to what has developed into a large and complicated topic. It could well be that your reading will take you off into areas that seem difficult to follow and are of little relevance for you at this stage of your studies. You will simply have to take it on trust that a little knowledge of these issues will stand you in good stead not only for your future studies in geography but also for what is important here, namely an understanding of the basic methods used in geographical research.

It might be helpful to divide this extensive literature into three basic groups:

- First, there are materials that are concerned with geography as a science.
- Second, there are approaches that have developed as a consequence, the so-called cultural turn in geography.
- Lastly, there are themes that have become popular only in recent times under the umbrella term of postmodernism.

As you might gather, the second and third groups are mainly the preserve of recent human geography. Most physical geographers would describe what they do as ‘science’ and this word can also still be used to describe much work in contemporary human geography. We will limit the detail of the discussion progressively as we move through these materials. Our main emphasis here is to make links to the methods that we employ in our research.

Just one last thing before we begin. There is a deceptively simple question that is often asked of geographers – what is geographical about geography? Concepts of ‘location’ and ‘space’ and ‘region’ and ‘place’ as well as ‘area’, not to mention ‘map’ are all clearly important (and not unproblematic) concepts. It seems that geography incorporates some sort of ‘spatial’ dimension but this can take many forms and certainly is not always mappable. It must be said that it is not always easy to recognise the ‘spatial’ in some research produced by contemporary geographers. The Dictionary of Human Geography has some good material on these themes that will inform not just this but other geography guides. But this is more a guide about the methods that are used in the subject in the light of different philosophical approaches rather than such pure geographical concepts themselves.

We have already briefly looked at the purpose of research in the previous chapter and we have learned that fundamentally it is about discovering things. The purpose of geographical research is about finding things out about geography. We have already said that geography, as a subject is unusual in that it encompasses a whole range of themes ranging from the human to the natural world. In this respect it is different from, say,
Biology or History. So the scope of the natural world is wide and this means that a wide variety of approaches can be adopted. It is also the case that geographical research is undertaken to help solve particular problems facing society. We can argue then that it is a subject that can take on a distinct practical value, or, as some would say, policy relevance. Of course, not all geographical research is policy relevant, but much of it is, at least in comparison to other subjects of study. As a result, geographical research is frequently commissioned by government, business and other societal institutions. This type of research may be said to have an instrumental purpose. It is not just about acquiring knowledge and understanding for its own sake. When this occurs there is a tendency for sponsors to look to a particular mode of discovery or knowledge acquisition. This is the so-called ‘normal silence’ mode of understanding things. This type of research can be replicated by others so that results can be tested. The approach is not dissimilar to conventional medical research where drugs and treatments are tested. It tends to contrast with the types of geographical research produced in other ways involving more cultural or post-modern dimensions that we will discuss in detail shortly. Having said this about the purpose of geographical research, it is the case that some of the ‘non-scientific’ geographical approaches are being used more and more to inform public policy and other decision-making areas of life. It is being realised that ‘science’ does not tell us everything that we want to know that is relevant or of practical value. The implications for more qualitative (as opposed to quantitative) methods will become apparent later.

Geography as science


Activity

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about. As you will see later, this will stand you in good stead for some typical examination questions on the materials contained in this chapter.

Whenever geographers use a research method, be it a soil auger or an in-depth interview, they are making some implicit philosophical assumptions about the nature of knowledge – how it is acquired and validated (or what philosophers call epistemology) – and about the nature of reality (what philosophers call ontology). Many geographers believe that the only knowledge that is legitimate and scientific is knowledge that is based on first-hand observation and measurement. Not all human geographers agree. Many argue that language and discursive categories do not simply represent social reality, they actually construct it. As a consequence different research methods of a more qualitative nature are required. We will look at this in more detail later in the course.

In science there is an important body of literature on what the term actually means: how is knowledge derived from scientific procedures different from understanding derived from common sense? It is usually argued that science employs the scientific method and is undertaken by ‘unbiased’ and ‘sceptical’ scientists. What does this mean for geography
as a science? The answers are not particularly clear-cut, but the following questions are good ones with which to approach your reading and understanding:

- Can and should the sciences be unified by a single methodology? And thus are human and physical geography methodologically similar or based on quite different philosophical principles?
- What are the criteria for scientifically valid knowledge?
- Is science about the search for universal laws (nomothetic)? Or about the accurate description of particular cases (ideographic)?
- Are the social sciences (and human geography) scientific in the same way as natural sciences?
- Is science value-free and objective?
- Does scientific knowledge advance progressively, one discovery building successively on the last?

If you have read some of the references already recommended, you will have seen that both human and physical geography have changed a great deal over the last few decades. How and why does this come about, you might ask. One attempt to answer these questions has involved the notion of **paradigm shifts**, which was advanced by Thomas Kuhn, an eminent philosopher of science. He argued that academic disciplines remain stable for long periods, but then undergo revolutions in thought as new paradigms emerge. A paradigm is best thought of as the combined working assumptions, procedures, practices and findings routinely accepted by a group of researchers that together define a stable scientific community. There have been several attempts to use this idea of paradigm shifts to describe and explain the various changes that have taken place in the discipline.

**Activity**

Here follows a series of propositions that you will be better placed to confront once you have finished studying for this part of the course. It will be instructive to record your own views now to see if your ideas change after undertaking this work. For now, simply note whether you agree or you disagree with these statements:

- Geographers should be careful not to allow political beliefs to influence their interpretation of data.
- Knowledge of something is of little use if you cannot measure it.
- Human geography is a science.
- Physical geography is a science.
- It is possible to conduct competent geographical research without having any philosophical viewpoint influencing it.
- The world exists primarily in relation to an individual’s subjective experience of it.
- The aim of geography should not be to understand the world but to change it.
- Artists convey more insight into people and places in their work than geographers do.
- In order to understand something properly it must be approached without prior assumptions.
- An individual’s position in the world (social class, occupation, gender, age, etc) affects how we see, understand and represent it.
Induction


**Activity**

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about.

The inductive route to scientific explanation is well trodden in geographical research. This is a scientific process whereby explanation can supposedly be induced from observed particulars. Historically, geography has been thought of as an empirical science. In other words knowledge and understanding arise from direct observation and usually some form of measurement. It is interesting to understand the implications of this line of thought with its origins in the enlightenment explorations of the early geographers such as Captain Cook and Charles Darwin. Indeed, most historians insist that the history of geography as a scientific discipline cannot be understood separately from the history of imperialism.

From this it is not difficult to understand the important role that fieldwork has played in the development of geography as an empirical science (and this is a topic that we will return to when we discuss geographical data later in this guide). Recently, however, the foundations of geography in empiricism and in field-based methods have been criticised. Researchers have questioned the political neutrality of scientific observation and exposing the imaginative geographies of far-off places waiting to be explored and represented. In the early times geographers collected facts and described them. This ideographic approach often took the form of descriptive regional geography. However, advocates of positivism (see below) preferred the nomothetic approach of attempting to search for generalised explanatory laws.

The normal route to explanation was inductive logic where, after collecting unordered facts, they are then defined, classified and measured so that ordered facts result. Then a process of inductive generalisation takes place of which scientific laws and theory are the product (and hence explanation). Associated with this process at various stages are quantitative methods, certainly in measurement and description, but also often using inductive statistics that move from the analysis of samples to statements about whole populations. There are a number of problems with this model of scientific progress which you should try to understand (some of which will be dealt with later).

Positivism

**Now read:** The Dictionary of Human Geography has articles on ‘logical positivism’, ‘critical rationalism’, ‘the quantitative revolution’, ‘regional science’, ‘spatial science’, ‘spatial analysis’ and ‘location analysis’.

**Activity**

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about.
Positivism is a complex term, which means different things to different people. The term was first coined by the French sociologist August Comte who saw positivism as a scientific methodology for applying reason and empirical observation to distinguish matters of positive fact from religious belief and superstition. Certainly scientific processes of induction would fall under this heading.

Logical positivism arose out of a concern for the problems of induction and the challenge of going from the particular to the general and thus verifying theory. Through logic and reason, logical positivists hoped to define more closely the scientific method. Starting with a theory, a scientist makes a hypothetical deduction that if this theory is true, then a particular outcome should occur. This hypothetical deduction then provided the basis for testing the theory. In other words, both deductive and inductive methods were used sequentially. But, as ever, there are problems with this line of thought too. So-called critical rationalism was a school of thought that worried about the problems of induction in this logical process. Their solution was a process of falsification though a succession of bold conjectures and refutations. It is fair to say, however, that these approaches had more in common with one another than differences.

This conception of the scientific method is:

• Universalised – it presumed that these rules of method provided a single and universal basis for inference that would apply to both human and physical geography.

• Theory driven – the emphasis here is on the improvement of theory as the driving force of science, as opposed to new empirical findings or practice.

• Normative rather than descriptive – the aim was not to describe how research was conducted but rather how it should be.

Much has been written critiquing positivism, and often the arguments debate the difficulties involved in induction in the absence of theory and logic in the absence of empirical regularities. Other important worries concern value neutrality. Critics complain that in practice science and empirical observation are never value-free and politically neutral. They deny that there is any such thing as pure science that can be divorced from the applications (and implications) of science and technology. By refusing to engage critically with, and condemn as necessary the values science serves, positivist geographers – or so their radical critics claim – simply support the status quo.

Another attack on positivism arises from those humanist geographers who feel that by focusing on the purely material-physical (and therefore universal and predictable) characteristics of things, crucial questions of human meaning are ignored. While such a methodology might be appropriate for studying the world of physical geography, it was entirely inappropriate for a truly human geography that should be sensitive to the particular meanings of places for people and must therefore employ qualitative methods.

But positivism in geography has been about quantitative methods and the revolution that introduced them to both the human and physical parts of the subject. You will be employing some quantitative methods later in this subject guide (as well as some qualitative ones). So you should attempt to develop an overview of the debates about the so-called quantitative revolution in geography, an understanding of the relationship between quantitative methods and positivism, and an appreciation of the different
ways in which quantitative methods were received in human and physical geography. At the time, the paradigm shift that this revolution represents, involved the creation of a sub-field of geography called spatial analysis and some of this type of work is still commonplace in the discipline today, especially in conjunction with recent advances in geographical information systems.

Another offshoot of this revolution on the human side was the discipline of regional science − a field that in conjunction with regional economics is still very strong today. As for physical geography, the analytical modelling (mathematical and statistical) that arose from the quantitative revolution still forms the backbone of the subject.

**Activity**

Have a look at the following website and refer to Michael Pidwirny’s online textbook entitled *The Fundamentals of Physical Geography* (2006) second edition. This is an excellent resource.

www.physicalgeography.net/fundamentals/contents.html

Chapter 3 ‘The Science of Physical Geography’ is especially valuable at this point. Read and make notes on what Professor Pidwirny (University of British Columbia Okanagan) writes, concentrating on the first half of the chapter dealing with scientific methods and changing approaches in physical geography. (Leave the materials on statistics for later. We will come to these topics later in this guide and this reference will again be useful.) Also have a look at Professor Pidwirny’s study guide to each chapter of his book where he summarises what he has been discussing, lists key terms and sends you off to look them up in his valuable glossary. There is also a useful section of further reference web links for each chapter that you might like to follow up.

**Activity**

You might like to read further the chapter by Rob Kitchin ‘Positivistic Geographies and Spatial Science’ in Aitken and Valentine (eds) (2006) pp. 20–29, and write a short note on what it is all about in common sense terms (as far as you can).

**The cultural turn in human geography**

The cultural turn is a broad movement encompassing both researchers’ reactions against positivism and the scientific pretensions of geography. For the most part, it has been human geographers who have been most active in rejecting the idea that the truth of geographical knowledge was guaranteed by its scientific method. Sometimes, as in the case of hermeneutics and interpretative approaches (discussed next), this involves a completely different conception of human geography compared to physical geography. But otherwise, as in the case of Marxist and feminist approaches and critical realism (discussed later), the rejection of positivism does not necessarily mean either a rejection of science or that, methodologically, human and physical geography should be based on the same principles.


**Activity**

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about.
In contrast to empiricists, positivists and critical realists (see below) those advocating a cultural or linguistic turn in human geography – a humanistic human geography – believe that the objective should not be to create an experimental science in search of laws but to focus more on meaning through interpretation. To do this it is necessary to switch to the variety of more qualitative methods on offer, such as discourse analysis, semiotics, in-depth interviewing and ethnography. Hermeneutics is the study of interpretation and meaning. The name comes from the Greek for ‘clarify’ and first rose to prominence as a philosophy for interpretation in the late eighteenth century among biblical scholars struggling to clarify the word of God. This approach recognises that in trying to understand the meaning of something we bring to it a whole set of preconceptions. These provide a context in which we make sense of and interpret the meaning of the text. This practice of reflecting upon and questioning the assumptions and statements of one’s own research is termed reflexivity.

All this makes a hermeneutic approach very different from those approaches that distinguish sharply between fact and fiction. From a hermeneutic perspective these distinctions are much more problematic. As you might guess, this means that the proponents of this approach think that methodologies for human geography and physical geography need to be different. Physical geographers study unconscious physical matter, like climate. By contrast, human geographers study conscious humans and their meaningful actions. Humans studied by human geographers are conscious and capable of changing in response to their own understandings and how we understand them. By engaging hermeneutically with the human subjects they study on such a subject-to-subject basis, human geographers are involved in a very different relationship from the subject-to-object relationship involved in studying the meaningless objects of physical geography and the other natural sciences.

Activity

Just contemplate for a moment what has been said in the above paragraph in respect of a research project on, say, child prostitution in a developing country. Then, add the dimension of, say, wealthy, male, western tourist consumers of such prostitution and you can see how the ‘meaning’ of any potential research finding becomes more complex as you build in your own personal opinions or biases about how you feel about the children and the men who abuse them. Proponents of a hermeneutic approach believe that the only way to understand the subject matter properly is to engage in what has been termed a ‘thick description’ of the subject (meaning a sort of deep dialogue). You should also then begin to see how hermeneutics is based on humanism.

A good question to ask next is what does this mean we can do with geographical knowledge searched out in this way? Some researchers critical of such an approach say that the outputs are descriptive, subjective, non-scientific and non-representative and therefore not useful. The main response is that these criticisms totally miss the point about what it is to be human.

Chapter 2: The nature, purpose and practice of geography

Activity

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about.

Critical or radical geography represents a reaction to apolitical positivism. One of the major intellectual influences on critical geography has been Marxism, and notably the writings of the geographer David Harvey (who interestingly was also in the vanguard advocating the benefits of positivism in his early work). Marxism has often been associated with dialectics, which is a way of understanding things in terms of the processes, flows and relations that create them. Some argue that as a methodology dialectics is perfectly capable of being used to think about the practice of natural sciences as well as the social sciences.

A second major influence on critical geography has been feminism. This began by documenting the experience of women in space, moved onto theorising the relationships between patriarchy and capitalism and more recently has taken a more critical approach to Marxism. Feminists complained that the Marxist emancipatory project was totalising in the sense that it tried to squeeze other axes of social oppression, like sexism and racism, into a single master explanatory frame, namely the Marxist critique of capitalism.

For all their differences, Marxist and feminist geographers shared two common complaints about dominant modes of geographical knowledge:

• First, they complained that, as it was often practised, geography was biased and served the interests of the powerful in society against the poor, vulnerable and oppressed.

• Second, radical critics complained that the pretensions to value-neutrality of both positivistic and humanistic geographers involved a tacit acceptance of the status quo. By contrast, critical geographers thought it should be the role of the geographer to help change society, by critically diagnosing its flaws and prescribing solutions.

Thus, although different critical geographers entertain somewhat different values, they agree that any geographical account, however scientific, must also necessarily be informed by some normative and ethical values. This involves a direct challenge to conventional understandings of science as value-neutral. It has also left critical geographers with something of an ambivalent stance towards science and truth, insisting both that knowledge is always value-laden and partial and yet also wishing to hold onto some notion of context independent truth as the grounds for diagnosing the very real problems of society.

Activities


Finally, in this brief review of reactions to positivism (and interpretivism) we come to critical realism, an approach usually associated with the geographer Andrew Sayer. It is interesting to note that physical geographers also seem to be interested in this methodological approach.
So what is it all about? According to Sayer, this approach lies between ‘pessimistic views’ pointing out that no one interpretation is better than any other and ‘positivistic views’ missing the point at issue.

Critical realists criticise both positivist and interpretivists for their value-neutrality.

They accuse positivists of being naïve about:
- the role of theories and values in observation
- the effects of scientific understandings/knowledge on society

They praise the hermeneutic and interpretative approaches, but complain that:
- hermeneutics’ rejection of the logic of scientific inference is misplaced
- interpretivists’ rejection of science is absurd and relativistic
- interpretivists’ refusal to engage in critique amounts to no change.

**Activity**

You might like to read further the chapter by Andrew Sayer ‘Realism as a Basis for Knowing the World’ in Aitken and Valentine (eds) (2006) pp. 98–106 and write yourself a short note on what it is all about in common sense terms.

---

**Postmodernism and geography**

**Now read:** The Dictionary of Human Geography has articles on ‘modernism’, ‘modernity’, ‘postmodernism’ and ‘postmodernity’.

**Activity**

Read and make your own notes on the topics cited above (some of them have been given quite lengthy treatments). Then read your notes, put them to one side and write in your own words what the terms mean or what the topics are about.

Postmodern, postmoderism and postmodernity are heterogeneous terms, signifying at once a kind of postmodern style of pastiche and juxtaposition in (especially) the arts and architecture; a historical epoch (postmodernity) that comes after and surpasses the modern one; and a variety of social theory (postmodernism).

The key seems to be that postmodernists argue that everything that has gone before (modernism) has been singularly unsuccessful in explaining societal differences. They believe that to search for a single answer or solution is not productive, no method is better than another, no one speaks with more authority about something than another. They believe that there is no absolute truth but that any truth is dependent on interpretation. They believe that it is more productive to read and interpret society than to observe and seek causes. They believe that the best way of working is to deconstruct (meaning intricately tease apart) human society, its practices and cultures. In other words, almost anything goes!

It is complicated, but the human geographers among you will confront these terms more and more as you specialise.

**Activity**

You might like to read further the chapter by David B. Clarke ‘Postmodern Geographies and the Ruins of Modernity’ in Aitken and Valentine (eds) (2006) pp. 107–21 and write yourself a short note on what it is all about in common sense terms.
Lastly, it is worth noting that in Kitchin and Tate (2000) Chapter 1, there is an excellent series of punchy, ‘boxed’, descriptions of what we have been talking about in this chapter. Read the rest of this chapter and especially the boxes and see if they accord with what you have understood from your own reading. You might even want to turn these boxes into ‘flash cards’ for your own revision purposes. And (really lastly this time) perhaps you would like to look up the chapter by A. Stewart Fotheringham ‘Quantification, Evidence and Positivism’ in Aitken and Valentine (eds) (2006) pp.237–50 for this will bring the whole discussion back to methods, which is where we are going next.

A reminder of your learning outcomes

Having completed this chapter, and the relevant reading and activities, you should now:

- understand that there is no single and unified philosophical approach to the academic subject of geography but attempts have been made to group and devise a chronology for the various paradigms
- realise that much of geography is ‘science’ and that being a science means that specific objectives and method have to be adhered to, although variations in approach are to be seen
- see, for human geography, that in recent years various approaches have been developed that form part of the so-called ‘cultural turn’ and that these too are linked to specific objectives and method
- see, again for human geography, that the most recent popular advances are approaches that can be described as ‘postmodern’ where specific objectives and method seem to be more to do the problem than the solution
- realise that, in practical terms, a language of terminology must be understood before proper understanding of the meaning of these approaches can be acquired.

Sample examination questions

1. Explain whether and how geography might be a science.
2. Why has the reaction against positivism been so much more hostile among human geographers than physical geographers?
3. What constitutes an interpretative, hermeneutic approach to human geography? Why do its proponents think human geography must be different from physical geography?
4. What is critical about critical human geography? Can critical human geography ever be scientific?
5. Has postmodernism been liberating for geographers. If not, then why not, and if so, then in what ways?
6. Explain briefly in a paragraph or two the meaning of the following terms:
   - logical positivism
   - hermeneutics.
7. What would you say constitutes the make up of the so-called ‘cultural turn’ in geography?
8. What do qualitative data offer that quantitative data do not, and vice versa?
9. Attempt a chronology of the various important ‘paradigm’ shifts that have occurred in geography since the early days of exploration and discovery.

10. Does the future of geographical analysis lie in data or discourse?
Chapter 3: Geographical data: human and physical, quantitative and qualitative, primary and secondary

Essential reading

Wheeler, D., G. Shaw and S. Barr Statistical Techniques in Geographical Analysis. (London: David Fulton Publishers 2004) third edition [ISBN 9781843121763]. It will be helpful, as background, to undertake a ‘fast read’ of Chapter 2, which discusses statistics and computers in general, although note we will not be using the rather less widely available computer package called MINITAB that these authors use. Chapters 3 and 4 contain very relevant material for issues that we will consider next.

Or if you have purchased Rodgerson P.A. Statistical methods for Geography. A Student's Guide. (London: Sage Publications, 2010) [ISBN 9781848600034 pbk] parts of Chapters 1 and 2 cover the same ground. Note that we will not be using the rather less widely available computer package called SPSS that is used extensively in this book. But it is true that in later chapters Rodgerson also makes use of EXCEL, which is the widely available package that we have chosen for our guide.

Further reading


Aims of the chapter

The aim of this chapter is to introduce you to the basic material upon which we can deploy our methods of geographical analysis, namely geographical data. As we perhaps already know, but certainly as we will see shortly, geographical data can be divided into qualitative and quantitative types. This chapter will focus almost entirely on the latter (with Chapter 7 focusing on various types of qualitative data). We need to know how data are measured and what difficulties are involved. We need to understand what is geographical about geographical data and how we organise our data into a form suitable for analysis. We also need to understand that we will have to collect some data for ourselves and some we will acquire from secondary sources. As regards the former we need to know something of surveys and sampling and as regards the latter we need to have some idea where the potential sources are and what they look like.

Learning outcomes

By the end of this chapter, and having completed the relevant reading and activities, you should be able to:

• understand that data are measured in various ways, some of which are more useful than others, and note that any form of measurement is not without problems associated with error
• put the ‘geography’ into geographical data by considering them in the form(s) of a geographical data matrix (and link this to a spreadsheet); you should be able to put these ideas into practice
• understand that although all data has to be collected, some data has already been collected by others for us to use
• design elementary questions for a questionnaire survey
• understand what sampling means in a geographical context
• search out geographical data (physical and human) collected from secondary sources and experiment with the practical capturing of these data in geographical data matrix and spreadsheet form so that they are ready for analysis.

Quantitative and qualitative data

It must be apparent, after reading Chapter 2, that geographers use both quantitative and qualitative data. By no means all geographers use both, as much depends on the research tasks set. Qualitative data, as might be expected, are mainly (but not exclusively) the preserve of human geographers.

Quantitative data are those types of information that can be measured (as we will see later, in various ways). Qualitative data are those types of information that rely on other forms of portrayal and description (for example, a piece of writing from a newspaper article or a commentary on something or a work of fiction or a historical record or a poem or a narrative about someone’s feelings can be considered as qualitative data, and so too can a photograph or a picture or perhaps even a performance or a piece of music). It all very much depends on the purpose and context in which the piece of information is being used. This is not to say that a piece of writing cannot be analysed quantitatively — it can. There are computer programs that have been designed to do just that by measuring the number of times, say, a particular word or phrase is used, or a situation is developed. The resulting data then become quantitative. It is only when the piece of qualitative data remains what it is — for example, a piece of prose or written comment — but is then interpreted within a research framework without subsequent measurement that it may be considered different from being quantitative. This should all be apparent from the methodological scene that has been set in the previous chapter and it is covered in detail in a later chapter. The rest of this chapter and Chapters 4, 5 and 6 are about methods that use quantitative data of various sorts in both physical and human geography.

Using this course with 04a Statistics 1

In this chapter we are going to draw heavily, but not exclusively, on the other compulsory half course that you are studying if you are studying this course as part of a BSc Degree in Geography: 04a Statistics 1.

Make sure that you have a copy of the subject guide to hand as much of the detailed presentation of the statistical material provided in that subject guide will be relevant to our work in understanding geographical methods. The intention is not to repeat this material but rather to harness some of it for our own geographical purposes. In fact, you will find that the prime intention in this part of the course is to encourage you to use and interpret such methods in a geographical context. Frequent reference will be made to 04a Statistics 1 in what follows and you will be asked to study in detail certain (but not all) of the statistical methods that are
presented there. This will be good for you and, of course, this study will help you to do well on the 04a Statistics 1 course itself. But please note that just because here we do not use all of the material presented in 04a Statistics 1 that you can be selective with your 04a Statistics 1 studies – you can’t! Lastly, if you have flexibility in your study arrangements, it would be better if you were further ahead in your study of 04a Statistics 1 compared to this half course as you work through them both.

Activity

As a start, read or read again Chapters 1 and 2 of the subject guide for 04a Statistics 1. Chapter 1 is a mini revision course in helpful numerical things that you should have done before. Chapter 2 helpfully discusses what the study of Statistics is all about. Both of these chapters provide the context for our geographical work. Read them quickly and let’s move on for we have some real-world geographical analysis to consider.

Let us begin by thinking about measurement.

Measurement

What do we do when we measure something? We ascribe a numerical value to the amount or quantity of something according to clear rules of measurement. Scientists have found that it is helpful to classify efforts to ascribe this value into four different scales of measurement. Each of these scales has its own particular characteristics and, above all, usefulness.

• Nominal scale data
• Ordinal scale data
• Interval scale data
• Ratio scale data

The first of these (nominal) is the lowest measurement scale, which basically involves classifying things into types based on their equality in some respects. For example, we could classify a rock type (say shale) as red or black or grey or green and then assign a number to each of these types, such as 1, 2, 3 or 4. These numbers mean nothing more than a name for the type or category. Nothing would change if the numbers were 2, 1, 4, and 3. Only limited arithmetic operations can be performed on these numbers. We could, for example, count the number of rocks in each class to find the class consuming the largest number. But the number assigned to describe the class itself is immaterial.

One step up (ordinal), the measurement scale is to rank the data. Here the numbers assigned to each class have some order in the sense that subsequent numbers follow a sequence — each being more than in value (or less than) the one previously. For example, we might have collected, say, 10 samples of a type of rock (shale) in a variety of colours. We may classify them in terms of their degree of whiteness or blackness. The most white is equal to 1 and the most black is equal to 10, with the rest somewhere in-between. It is important to recognise that the steps between the numbers in such a scale are not equal, but we can be sure of the ranked nature of the data. Another ordinal measurement type is to classify the 10 rock samples into say, 4, different degrees of whiteness or blackness, where white equals 1, light grey equals 2, dark grey equals 3 and black equals 4. The main distinction in the ordinal measurement type is that questions of a ‘more than’ or ‘less than’ nature can be asked.

Interval data is somewhat more sophisticated. Here you have ranked data but with equality of intervals between the classes. An important
type of interval data in physical geography is temperature. We use a thermometer to measure temperature which involves calibrated equal steps of expansion of an indicator like mercury as temperature rises. The zero point of an interval scale may, however, be arbitrary as in the Fahrenheit and Celsius temperature measurement scales. This arbitrary zero point of interval measurement is problematic for some arithmetic operations and is overcome by the highest measurement scale: ratio data. As its name implies, the critical feature here is the need for equality of ratios. This is achieved by having equality of intervals and a real zero point. If we measure the length or weight of something, say a piece of rock, then we are using a ratio scale from zero to whatever the length or weight is of the particular rock that we are measuring and it does not matter what unit of measure is being employed. This is not the same for measures of temperature using Fahrenheit and Celsius measurement scales. Ratio measurement offers the most opportunities of all the measurement scales to researchers and is by far the most often used. If you measure the length of river A as 1,000 kilometres and that of river B as 250 kilometres (obviously on a ratio scale) then you can be sure that river A is 750 kilometres or four times longer than river B.

Activity

It would be instructive for you to think of some of your own examples of data interesting to geographers – both physical and human – falling into each of the categories described above.

Do not worry too much about the interval scale – exercises like this usually end up talking about temperature! But the other types are more straightforward and well worth thinking about because they occur all the time in geographical analysis.

Measurement error

After you have thought of a variable that would be particularly useful for geographical analysis, your next task is to measure it (and the above scales of measurement are helpful in this respect). But the process of measurement is in itself problematic, basically because of the issue of measurement error. Various sorts of measurement error occur:

Gross errors occur when we, as observers or measurers, make a mistake. We might take our ‘eye off the ball’ and just now and then make a big mistake. These sorts of errors are usually (and hopefully) infrequent, but they can be large. Computers are helpful in spotting these sorts of errors because they tend to stand out. Many automated checking procedures have been devised for computers that help to identify gross errors. They look for ‘outliers’ in the data, which then be subsequently verified or not as the case may be.

Systematic errors are those that tend to produce results that are consistently ‘out’, either too large or too small. They could result from an instrument that has not been properly set up or calibrated.

Errors of method are said to occur when there is a difference between the researcher’s understanding of nature of the variable to be measured and the way in which the variable actually is measured. These sorts of errors occur much more frequently than you would imagine and there is many a debate in geographical research about the exact meaning of variables. For example, to some, Gross Domestic Product (simply the value in monetary terms of the wealth that is generated within country during some period of time) is a good way of measuring levels of development. To others, it is full of errors in the way this variable – development - should
be measured. Those taking a wider view of development than one based on aggregate monetary value of production could well want to look at income distribution, poverty, levels of health and welfare or perhaps education, not to mention some more popular measures of the quality of life of the population. But all these things, too, are difficult to measure.

**Random errors:** there are those errors that occur everywhere and anywhere but are unpredictable and to which it is usually not possible to assign a cause. Furthermore, it is almost impossible to control for such errors by checking systems. Random error is always present and not infrequently is the outcome of the intersection of multiple events only really understood with hindsight.

Finally, it is worth just rehearsing the difference between measurement **accuracy** and measurement **precision**. Here the line is that the former represents a value that is not related the method of measurement. An accurate measure is a true measure of the variable of interest, whereas precision is very much to do with the method of measurement itself. The latter represents just how close to one another repeated measurements of the same variable are. Ideally we are looking for accurate measurements that are both close to the true value and precise.

Some things in geographical analysis are notoriously difficult to measure, and one of them is shape. Shape is a useful variable in both human and physical geography. We might want to measure the shape of, say, a travel-to-work area or a drainage basin area and all we have to go on is an irregular shape on a map. You can imagine them (no need for a map): there are long thin ones, short fat ones, ones with curves of various sorts (concave and convex). We can do some approximations by taking the shape and measuring:

- **Width divided by length:** the width and length can be operationalised as the longest and widest intercepts at right angles to one another. It does not matter if you take the measurement of width or length first as long as you are consistent if measuring many shapes.
- **Diameter of the inscribed circle divided by diameter of the circumscribed circle:** using a set of compasses draw the largest ‘touching’ circle inside the shape and the smallest ‘touching’ circle outside the shape and measure their diameters.
- **Diameter of a circle with shape area divided by length:** estimate the area of the shape and then calculate the diameter of a circle that has this area. You will need to find out the formula for the area of a circle and do some simple sums on a basic calculator. You will also need a measure of length as before.
- **Circumference of a circle with shape area divided by shape perimeter:** estimate the area of the shape and then calculate the circumference of a circle that has this area. You will need to find out the formula for the area of a circle and do some simple sums on a basic calculator. You will also need an estimate of the perimeter of the shape.

All of these measures are going to give different results for what purports to be the same thing: shape.

**Activity**

On a piece of graph paper (or make a sheet of gridded paper yourself), draw a shape, as irregular as you like but keep it fairly centred in the middle of the paper. Then, measure the shape using each of the above four variables. You will need a set of compasses.

**Hint:** approximate the area by counting the squares on the graph paper within the shape.
and devise some rules about squares that fall on the boundary. Perhaps measure the perimeter with a piece of thin string (if you do not have a planimeter to hand)!

This a good exercise as it teaches you much about variables, measurement and error. Remember that the shape that you have just drawn could be a really interesting geographical area that could be crucial to your proposed geographical analysis for some reason – patterns of travel-to-work or stream channel networks in a river basin, etc.

### Geographical data matrix

Now that you know how to measure something that you might be interested in (and some of the difficulties that might be encountered), you should next ask yourselves the question of what makes the data geographical? The answer is simple: location. Geographers are interested in understanding the importance of location or space or place or area or region for a particular topic that they are interested in. A sociologist might be interested in the social relations between a group of people. A geographer's take on such an interest would be the influence of the geographical context on those social relations. Data can be located in various ways such as coordinates, longitude and latitude, regions, post codes, street names and many more. One way of putting the 'geographical' into a set of data is to think in terms of a geographical data matrix. This sounds much more complicated than it actually is: a matrix is simply an ordered array of numbers.

Before we deal with geographical data matrices in more detail it will be helpful first to rehearse what we know about 'spreadsheets' on the computer. As most of us know, spreadsheets are helpful and systematic ways of undertaking sometimes quite complicated, routine calculations using the computer. A spreadsheet package is usually bundled with the software that is supplied with computers these days or at least can be purchased cheaply to work on your computer.

Learning how to use a spreadsheet is straightforward, although it will take some time to become really proficient. If you have not come across spreadsheets before, find one and play with it. (This is not part of the course but will be good for you!) The most widely used spreadsheet is that supplied by Microsoft within its Office product. It is called Excel.

Figures 3.1 – 3.3 are hypothetical spreadsheets from the Excel program showing how geographical data might be organised (x represents individual pieces of data themselves).

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Variable 1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2 2</td>
<td>Region 1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>Region 2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>Region 3</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Region 4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>Region 5</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Figure 3.1: Regions and geographical variables, same date.**

Figure 3.1 shows a conventional data matrix of regions and geographical variables. The axes – vertical and horizontal – of the data matrix in theory are infinite as we can conceive of an infinity of regions (points) across the world and an infinity of variables that we might want to measure at such locations. In practice, of course, things are much more manageable and we know exactly how many regions and how many variables we are dealing with. The Excel spreadsheet has size constraints too! In Figure 3.1 it is not
difficult to see that time is held constant and the temporal snapshot that we are looking at may well be a census date or some such information. In fact, some observers prefer not to think of a geographical data matrix but instead of a geographical data cube where the remaining dimension could be points in time (again, in theory, infinite).

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
</tr>
<tr>
<td>2</td>
<td>Region 1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>Region 2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Region 3</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Region 4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Region 5</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.2: Regions and dates, same geographical variable.**

Figure 3.2 takes a different facet of the data cube. Here we are looking at one variable organised by regions and time.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
</tr>
<tr>
<td>2</td>
<td>Variable 1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>Variable 2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Variable 3</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>Variable 4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>Variable 5</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Figure 3.3: Geographical variables by dates, same region.**

Figure 3.3 similarly looks at a number of variables measured over a variety of time points but for the same region. The key point to note is that you can take any piece of geographical data – from physical or human geography – and locate it within a geographical data matrix of the sorts that we have discussed here.

There is one last thing worth mentioning here and that concerns how we handle flow data. Flow data involve, as the name suggests, flows of some quantity from, in geographical terms, some place to some other place. Migration is a prime and vitally important example of flow data, but so too are international trade and the journey to work. The best way of conceiving of these data is again to use a geographical data matrix. The rows and columns can represent places and the cells of the matrix can represent the values of the flows as, hypothetically, shown in Figure 3.4.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Flow</td>
<td>Region 1</td>
<td>Region 2</td>
<td>Region 3</td>
<td>Region 4</td>
<td>Region 5</td>
</tr>
<tr>
<td>2</td>
<td>Region 1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>Region 2</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Region 3</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>Region 4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>Region 5</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Figure 3.4: Origin: destination flow data for one flow variable.**

This is a simple form of what is called an origin-destination flow matrix where every single flow between a set of places can be recorded. Note the diagonal of this matrix contains the flows (if any) of places to themselves. Another form of flow matrix is the so-called dyadic form where each row represents the values of flows for specific variables (say international trade commodities) for each pair of places (dyads) in the geographical system under analysis. So in Figure 3.5 the flow variables are listed along the horizontal axis and the dyads on the vertical.
Methods of geographical analysis

Figure 3.5: Dyadic flow data for several variables.

The first dyad is the flow from place 1 to itself, the second is from place 1 to place 2, and this is repeated until we reach the last dyad, which is represented by a flow from place N to itself. We will not have chance to revisit flow data in this introductory course but there is no mistaking that flow data are especially interesting and have formed the basis of many a good geographical research project.

Activity

It now worthwhile to try to find some real geographical data and observe them in this particular matrix form. There are lots places where you can find geographical data, which have already been collected, and you will come to some of these shortly, but for our purposes now take a look at a UK website provided (free of charge) by the government statistical service. Here we are going to look at the data contained in an annual publication called Regional Trends (volume 39 refers to 2006) and, as its name suggests, it contains data on a number of themes for the regions of the UK. The address is as follows:

www.statistics.gov.uk/regionaltrends39/

Here you can read all about this publication and even download the various chapters if you like. For this activity you need to download some data. On this web page you are also offered a prompt asking you if you want to ‘Download full report & data’ (which you do) and then another prompt ‘Regional Trends 39 Data (Zip)’ (again which you do). The Zip specification refers to way the file of data is compacted. You are usually asked where you want to store the data that you are going to download – a file on your hard disk perhaps? The download to your computer contains all of the data used in all of the chapters of the report – a lot of data! Let’s have a look at some of it.

First, open up the data in Chapter 2 European Union, table 2.1. To check that you have the right data, the top of the table is shown below as Figure 3.6. In style it looks very much like our Figure 3.1 (regions by variables). And the beauty of it is that it is in Excel spreadsheet form. (It does not look like an Excel spreadsheet because it takes advantage of formatting options to improve the presentation and does not have grid lines – but it is an Excel spreadsheet – try it and see.)
Chapter 3: Geographical data: human and physical, quantitative and qualitative, primary and secondary

2.1 Population and vital statistics, 2002

<table>
<thead>
<tr>
<th>Region</th>
<th>Area (sq km)</th>
<th>Population (thousands)</th>
<th>People sq km</th>
<th>Age under 15</th>
<th>Aged 65</th>
<th>Births (per 1,000 population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR 25</td>
<td>3,164,177</td>
<td>453,757</td>
<td>120</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Austria</td>
<td>83,859</td>
<td>8,053</td>
<td>96</td>
<td>16.5</td>
<td>15.5</td>
<td>9.7</td>
</tr>
<tr>
<td>Österrich</td>
<td>23,790</td>
<td>144</td>
<td>16.6</td>
<td>15.2</td>
<td>9.7</td>
<td>10.6</td>
</tr>
<tr>
<td>Südtirol</td>
<td>25,671</td>
<td>67</td>
<td>16.1</td>
<td>16.5</td>
<td>8.8</td>
<td>9.7</td>
</tr>
<tr>
<td>Österrich</td>
<td>34,938</td>
<td>85</td>
<td>17.9</td>
<td>14.2</td>
<td>10.3</td>
<td>8.1</td>
</tr>
<tr>
<td>Belgium</td>
<td>10,103</td>
<td>339</td>
<td>77.5</td>
<td>16.9</td>
<td>10.8</td>
<td>10.2</td>
</tr>
<tr>
<td>Brussel</td>
<td>161</td>
<td>905</td>
<td>6.0</td>
<td>16.4</td>
<td>14.1</td>
<td>10.4</td>
</tr>
<tr>
<td>Vlaams Geest</td>
<td>13,912</td>
<td>443</td>
<td>16.9</td>
<td>17.1</td>
<td>10.0</td>
<td>9.7</td>
</tr>
<tr>
<td>Région Wallon</td>
<td>16,864</td>
<td>200</td>
<td>18.5</td>
<td>16.8</td>
<td>11.2</td>
<td>11.2</td>
</tr>
<tr>
<td>Cyprus</td>
<td>--</td>
<td>710</td>
<td>--</td>
<td>23.3</td>
<td>11.6</td>
<td>6.9</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>78,800</td>
<td>129</td>
<td>16.1</td>
<td>13.9</td>
<td>9.1</td>
<td>10.4</td>
</tr>
<tr>
<td>Denmark</td>
<td>43,954</td>
<td>125</td>
<td>18.7</td>
<td>14.8</td>
<td>11.9</td>
<td>10.9</td>
</tr>
<tr>
<td>Estonia</td>
<td>45,278</td>
<td>30</td>
<td>17.5</td>
<td>15.2</td>
<td>9.6</td>
<td>13.5</td>
</tr>
<tr>
<td>Finland</td>
<td>30,529</td>
<td>17</td>
<td>18.8</td>
<td>15.2</td>
<td>10.8</td>
<td>5.4</td>
</tr>
<tr>
<td>Manner-Suomi</td>
<td>263,803</td>
<td>17</td>
<td>18.6</td>
<td>15.1</td>
<td>10.8</td>
<td>9.4</td>
</tr>
<tr>
<td>Åland</td>
<td>1,227</td>
<td>26</td>
<td>18.5</td>
<td>16.4</td>
<td>11.6</td>
<td>7.7</td>
</tr>
<tr>
<td>France</td>
<td>543,865</td>
<td>109</td>
<td>17.8</td>
<td>16.2</td>
<td>13.8</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Figure 3.6: Source: National Statistics website: www.statistics.gov.uk.
(Crown copyright material is reproduced with the permission of the Controller of HMSO.)

This is a substantial geographical data matrix providing us with detail on eight variables – population, births, deaths, infant mortality, etc. – for all of the regions (at the so-called NUTS 1 level which is a particular scale of regional focus) of the European Union.

1. Can you name the countries of the European Union?
2. Do you know where the regions within those countries of the European Union are?

Note: do not worry about NUTS – they are just a set of spatial scales used in the EU to produce regions of comparable size, so there are NUTS 1, NUTS 2, NUTS 3 regions, etc.

Next, go back to your downloaded data and have a look at Chapter 3 and specifically table 3.1. (Check against the top of our Figure 3.7. It is not unlike our Figure 3.2 in style, a sort of regions and dates, but a bit more complicated than that.)
### 3.1 Resident population by sex

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>United Kingdom</strong></td>
<td>27,411.6</td>
<td>27,909.0</td>
<td>28,832.4</td>
<td>29,271.0</td>
<td>1.8</td>
</tr>
<tr>
<td>North East</td>
<td>1,283.1</td>
<td>1,253.7</td>
<td>1,232.1</td>
<td>1,237.4</td>
<td>-2.3</td>
</tr>
<tr>
<td>North West</td>
<td>3,367.6</td>
<td>3,311.4</td>
<td>3,288.5</td>
<td>3,326.0</td>
<td>-1.4</td>
</tr>
<tr>
<td>Yorkshire and the Humber</td>
<td>2,386.0</td>
<td>2,396.5</td>
<td>2,421.0</td>
<td>2,459.9</td>
<td>0.1</td>
</tr>
<tr>
<td>East Midlands</td>
<td>1,894.8</td>
<td>1,968.4</td>
<td>2,061.0</td>
<td>2,106.9</td>
<td>3.9</td>
</tr>
<tr>
<td>West Midlands</td>
<td>2,556.6</td>
<td>2,563.9</td>
<td>2,588.8</td>
<td>2,620.2</td>
<td>0.3</td>
</tr>
<tr>
<td>East</td>
<td>2,366.0</td>
<td>2,511.3</td>
<td>2,647.7</td>
<td>2,696.3</td>
<td>6.3</td>
</tr>
<tr>
<td>London</td>
<td>3,277.3</td>
<td>3,296.4</td>
<td>3,597.1</td>
<td>3,673.4</td>
<td>0.6</td>
</tr>
<tr>
<td>South East</td>
<td>3,527.2</td>
<td>3,717.4</td>
<td>3,924.0</td>
<td>3,973.9</td>
<td>5.4</td>
</tr>
<tr>
<td>South West</td>
<td>2,118.5</td>
<td>2,269.6</td>
<td>2,405.5</td>
<td>2,458.9</td>
<td>7.1</td>
</tr>
<tr>
<td>England</td>
<td>22,795.0</td>
<td>23,290.6</td>
<td>24,165.6</td>
<td>24,553.9</td>
<td>2.2</td>
</tr>
<tr>
<td>Wales</td>
<td>1,365.1</td>
<td>1,390.7</td>
<td>1,408.7</td>
<td>1,434.3</td>
<td>1.9</td>
</tr>
<tr>
<td>Scotland</td>
<td>2,494.9</td>
<td>2,444.5</td>
<td>2,433.7</td>
<td>2,445.2</td>
<td>-2.0</td>
</tr>
<tr>
<td>Northern Ireland</td>
<td>756.5</td>
<td>763.2</td>
<td>824.4</td>
<td>836.5</td>
<td>3.5</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>United Kingdom</strong></td>
<td>26,345.9</td>
<td>29,529.7</td>
<td>30,281.1</td>
<td>30,564.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

*Figure 3.7: Source: National Statistics website: www.statistics.gov.uk.*

*(Crown copyright material is reproduced with the permission of the Controller of HMSO.)*

Can you figure it out? It is obviously about population and population change.
Chapter 3: Geographical data: human and physical, quantitative and qualitative, primary and secondary

1.25 Key statistics for London

<table>
<thead>
<tr>
<th></th>
<th>London</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population, 2004² (thousands)</td>
<td>7,429</td>
<td>59,835</td>
</tr>
<tr>
<td>Percentage aged under 18³</td>
<td>19.4</td>
<td>19.5</td>
</tr>
<tr>
<td>Percentage pension age and over²</td>
<td>13.9</td>
<td>18.6</td>
</tr>
<tr>
<td>Standardised Mortality Ratio (UK=100), 2003</td>
<td>97</td>
<td>100</td>
</tr>
<tr>
<td>Infant mortality rate,² 2004</td>
<td>5.2</td>
<td>5.1</td>
</tr>
<tr>
<td>Percentage of pupils achieving 5 or more grades A*-C at GCSE level or equivalent, 2003/04¹</td>
<td>54.4</td>
<td>54.2</td>
</tr>
<tr>
<td>Economic activity rate⁴, spring 2005 (percentages)</td>
<td>74.8</td>
<td>78.6</td>
</tr>
<tr>
<td>Employment rate⁶, spring 2005 (percentages)</td>
<td>69.3</td>
<td>74.4</td>
</tr>
<tr>
<td>Unemployment rate⁴, spring 2005 (percentages)</td>
<td>6.9</td>
<td>4.7</td>
</tr>
<tr>
<td>Median gross weekly earnings: males in full-time employment, April 2005 (£)</td>
<td>574.8</td>
<td>471.5</td>
</tr>
<tr>
<td>Median gross weekly earnings: females in full-time employment, April 2005 (£)</td>
<td>482.9</td>
<td>371.8</td>
</tr>
<tr>
<td>Gross value added, 2004 (£ billion)</td>
<td>165.0</td>
<td>1,005.4</td>
</tr>
<tr>
<td>Gross value added per head index, 2004 (UK=100)</td>
<td>132.2</td>
<td>100</td>
</tr>
<tr>
<td>Total business sites³, March 2004 (thousands)</td>
<td>383.1</td>
<td>2,573.1</td>
</tr>
<tr>
<td>Average dwelling price¹, 2004 (£)</td>
<td>274,860</td>
<td>183,449</td>
</tr>
<tr>
<td>Motor cars currently licensed,¹ 2003 (thousands)</td>
<td>2,450</td>
<td>26,240</td>
</tr>
<tr>
<td>Fatal and serious accidents on roads,¹ 2004 (rates per 100,000 population)</td>
<td>52.1</td>
<td>51.1</td>
</tr>
<tr>
<td>Recorded crime rate¹, 2004/05 (recorded offences per 100,000 population)</td>
<td>13,858</td>
<td>10,508</td>
</tr>
<tr>
<td>Average gross weekly household income, 2001/02 to 2003/04² (£)</td>
<td>740</td>
<td>554</td>
</tr>
<tr>
<td>Average weekly household expenditure, 2001/02 to 2003/04² (£)</td>
<td>485.50</td>
<td>405.20</td>
</tr>
<tr>
<td>Households in receipt of Tax Credits,¹ 2003/04 (percentages)</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

Figure 3.8: Source: National Statistics website: www.statistics.gov.uk. (Crown copyright material is reproduced with the permission of the Controller of HMSO.)

Figure 3.8 above is taken from Chapter 1 Regional Profiles, table 1.25. This is a bit like our Figure 3.3 in style if you only look at the London column (and imagine the United Kingdom column is not there). It is certainly a set of data for one particular region – London – but there is only one column (representing time but really a variety of dates for the range of variables measured). Imagine also this data for London provided in previous Regional Trends for earlier dates. You could add them column by column and then they would look more like our Figure 3.3. Things are becoming complicated, but that is the whole point of looking at data this way.
Finally, Figure 3.9 is taken from the Chapter 11 Environment (table 11.7).

### 11.7 Rivers and canals: by chemical quality

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Fair</th>
<th>Poor E</th>
<th>Bad F</th>
</tr>
</thead>
<tbody>
<tr>
<td>North East</td>
<td>27</td>
<td>31</td>
<td>42</td>
<td>52</td>
</tr>
<tr>
<td>West Midlands</td>
<td>18</td>
<td>26</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>East Midlands</td>
<td>2</td>
<td>15</td>
<td>19</td>
<td>39</td>
</tr>
<tr>
<td>South East</td>
<td>17</td>
<td>25</td>
<td>30</td>
<td>47</td>
</tr>
<tr>
<td>South West</td>
<td>18</td>
<td>26</td>
<td>36</td>
<td>38</td>
</tr>
<tr>
<td>England</td>
<td>14</td>
<td>27</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>Scotland</td>
<td>54</td>
<td>70</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>Northern Ireland</td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>48</td>
</tr>
</tbody>
</table>

**Percentages and Kilometres**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total length surveyed (km)</td>
<td>2,030</td>
<td>2,890</td>
<td>3,530</td>
<td>4,850</td>
</tr>
</tbody>
</table>

**Figure 3.9:** Source: National Statistics website: www.statistics.gov.uk.

(Crown copyright material is reproduced with the permission of the Controller of HMSO.)

Here we are dealing with the chemical quality of UK rivers and canals. The data aggregates rivers to regions and then gives us an idea of whether the chemical quality is good to bad (bad being polluted presumably). It also has some change over time information so we again we have a somewhat complicated matrix here, but it is not difficult to see the structure and organisation.

Now over to you! There are lots of tables of data in the document you have downloaded from this national statistical agency. Browse through them and see if there is anything there to interest you. Pay particular attention to what each section means. (It might also be worth just writing down in a few words what the data in Figures 3.6–3.9 actually represent in the real world, too, as you can be over-focused on structure.)

### Primary and secondary data

Having discussed the nature of geographical data we must next consider where it comes from. Data used in geographical research — human or physical — can be either from primary or secondary sources.

Essentially **primary data** is that which you collect yourself. If you are a social geographer you may well want to design and carry out a survey of respondent’s opinions about something of particular interest. You will have to write and test your questions. You will have to decide whom you are going to ask your questions of. You will have to undertake the fieldwork to carry out the survey and you will have to record and code the results for data analysis. If you are a geomorphologist interested in valley slope profiles you may want to design and carry out a field survey of a particular valley. You will need to decide which slopes you are going to measure and how. You will need to take your instruments into the field and conduct your survey, recording your results in a field notebook. Then, you will have to transfer your results into a form suitable for geographical analysis. These two examples detail the work and effort that is needed to collect
primary data for geographical analysis. The beauty of primary data is that they should be ‘perfect’ for your research purposes in that you have total control over the collection process. The downside is that collecting primary data is incredibly time-consuming.

With **secondary data** somebody else does the collecting. Almost every country in the world has a National Population Census agency and these are the main sources that come to mind but the variety of data sources is staggering. Often large sample sizes are involved in secondary data collection, particularly government surveys. It is advantageous in many ways, especially for statistical analysis. It is also the case that large-scale micro-data about individuals are becoming more and more available. These contrast with the area-based data more usually used in Geography.

Secondary data are immensely useful but are rarely absolutely perfect for your research requirements. This is because they have to serve a myriad purposes of which yours is but one.

This chapter concludes with a brief consideration of primary and secondary data in both human and physical geography.

**Primary data: human geography**

Human geographers can collect primary data in many ways (and Chapter 7 of this subject guide focuses on several of the ways in which qualitative methods can be operationalised). The main quantitative method used is the questionnaire. Questionnaires are not to be taken on lightly for they require careful planning and design, not to mention an awful lot of work. Indeed, they have become so detailed and complicated that a whole subject area has been developed to record the work of researchers in this field. Asking questions is a lot more complicated than it may seem at first sight. Obviously you will need to decide who to ask your questions of and whether the group of people that you choose represents everyone that you want to ask (the whole population in the statistical sense) or just a sample of the people that you could ask. If it is the latter, and it usually is, then you must consider the nature of the sample and how representative it is (including for geographers how geographically representative it is; this will be returned to later). Sometimes questionnaires refer to topics that might be described as ‘sensitive’ and there will be ethical considerations to consider before you actually ask your questions of your respondents. You may be interested to know that in many universities people (even undergraduates) are not allowed to undertake research on individual human subjects without first considering the ethical issues that might be involved. Chapter 7 contains more information on this important topic.

The text by Wheeler et al. (2004), like most others, gives some brief attention to questionnaire design, but if you are going to be serious about implementing a questionnaire survey you will need to consult more specialist texts such as Fowler (2002).

Let us briefly consider the many ways in which questions can be asked in questionnaire surveys with a view to generating quantitative information about a particular topic. Wheeler et al. (2004) outline four main types of questions:

- scaling
- ranking
- contingency
- general information

A **scaling** question asks respondents to scale their response to a question posed. For example, you may ask someone how important they feel...
a particular issue is and they can respond saying that they feel it is important or not important with various opportunities to grade importance in-between. Another scale is the so-called Likert scale where you can ask a respondent whether they agree or disagree to a particular statement about something or another, again with variations in-between. Finally, there is the useful semantic differential scale where respondents are asked to indicate where they stand on a particular issue in-between two extreme positions described by two words.

Let us consider some questions trying to determine attitudes towards the environment, for example.

Please indicate, by ticking one of the following boxes, how important you feel human action is in global warming:

- Not important
- Fairly important
- Important
- Very important
- Do not know

Please indicate, by ticking one of the following boxes, how much you agree with the statement that individuals can do little to influence carbon emissions and thus curtail global warming:

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please indicate, by ticking one of the following boxes, where you stand on the statement that the record of the United States on policies to control greenhouse gases is encouraging or appalling or something in between.

1 = Encouraging
2
3
4
5 = Appalling

As the name suggests, a ranking question asks the respondent to rank their response. For example, a question could ask someone to:

Please rank the following activities in order of importance (from 1 = most important) for their contribution to global warming:

- manufacturing industry
- air transport
- energy production
- household energy consumption
- road transport
- deforestation
- urbanisation
- aerosol use
A contingency question is a closed question, which usually asks for some specific detail about the characteristics of the respondent who is being interviewed. Probably the most usual question concerns income and this could certainly be relevant to public attitudes towards global warming. But other questions could take the form:

If you own motorised transport, please indicate, by ticking one of the following, the size of the engine capacity.

- Less than 100 cc
- 101 cc – 250 cc
- 251 cc – 500 cc
- 501 cc – 1000cc
- 1001 cc – 2000 cc
- Greater than 2001 cc

Please indicate, by ticking one of the following, how many journeys by air do you take each year.

- 0
- 1
- 2 – 5
- 6 – 10
- 11 – 20
- More than 20

And, finally, a general information type question could ask a respondent if, say, he or she would vote for a Green Party candidate at the next election if one was to stand. Or if they would use a recycling facility if one was to be provided.

Once collected, primary data from questionnaire surveys has to be coded into data suitable for analysis. Usually this will involve coding into a spreadsheet, such as Excel, where each row of data might well relate to a particular respondent and each column to their responses to particular questions. If one of the questions that was asked referred to residential location of the respondent it is not difficult to see how aggregate responses for groups of respondents could be drawn up for particular geographical areas covered.

**Activity**

Think of a geographical theme that interests you, or continue with the environmental line of thought if you like, and design some questions (of various sorts) that you could ask of your family or friends or colleagues – don’t forget ethics! Then ask your questions. If people can answer them with ease, you are probably doing a good job. If they answer your question with a question by way of clarification, then there is more work to do. All proper surveys undertake what is called a pilot to trial questions just in this way. Record your results and enter them into a spreadsheet. How will you know what the results mean when you come back to these data in a month’s time?

Hint: you will need a codebook, so make one.
Secondary data: human geography

We have already touched on some secondary data in human geography in the data from Regional Trends above, although the chemical composition of UK rivers will be of interest to physical geographers too. There are literally thousands of sources for secondary data that can be used by human geographers. More and more of these sources are becoming available on the internet. Many of these sources will allow you to access the information, usually – but by no means always – free of charge. And it is especially important to know that the data can often be downloaded straight to a blank Excel spreadsheet on your computer. Wheeler et al. (2004) discuss some of these sources and especially recommend: www.fedstats.gov for spatial information about the USA. It is a brilliant site too! If Europe is more your area, then you should certainly try to find: http://epp.eurostat.ec.europa.eu for information – much of it spatial – on the countries and regions of the European Community including the newly accessed Member States, many from Eastern Europe.

Activity

Locate a computer-based secondary source of some spatial data that is of interest for you as a human geographer (or social survey data that relates to a particular area or region or country). National census bureaus are prime sources but you could also try the United Nations or the OECD or any of the above. Select your data and transfer it to a blank Excel spreadsheet by downloading or copying or cutting and pasting, etc. Find out what your areas mean and define your variables, recording the steps that you have taken and the difficulties that you have encountered. If you are capturing some survey data, then the data might not be coded to areas but relate to individual respondents and that is fine too.

Primary data: physical geography

Most primary data collection for physical geography involves going out into the field and taking measurements about dimensions of the physical environment that are of interest. Time is a particularly important variable in physical geography and much primary data collection involves instruments making repeated measurements automatically. As students undertaking this course will have different, and some very limited, possibilities of fieldwork of this kind detailed tasks like this will not be assigned.

But most students will have access to a topographical map or atlas and so we can envisage some primary data collection using maps as the data source. Here we are interested in collecting information about stream order (for various reasons that can be understood by reference to Strahler, 1954). Look for river systems on the map and see where subsidiary or tributary streams flow into larger channels. Stream order research assigns a code to the various types of stream in a river basin system. Segments of a stream are ordered numerically from the higher reaches of a river to the point where it flows out to sea. At the top of the drainage basin a first order stream has no tributaries (i.e. no other streams flow into a first order streams). Higher order streams begin at river junctions of next lower combinations. A second order stream takes the flow from first order streams. A third order stream takes the flow from second order streams and so on. So, if the stream that flows into the sea is a fourth order stream we know that there will be second and third order streams elsewhere in
the drainage system. But it is possible for streams of a certain order to miss out elements of the hierarchy. A first order stream can for example (unusually) flow directly into a third order stream depending on the network configuration.

The job of primary data collection here is to take a map showing stream and river channels in a landscape and measure the lengths of each of the first order streams. You will need to number each stream and figure out a way of measuring stream length. The result should be spreadsheet of stream length measurements of the first order streams in the territory that you have chosen as your study area. You could have undertaken some fieldwork mapping to do this task of primary data collection but this map-based exercise is just as valuable. This is not a difficult task — finding a good map or atlas might be the most difficult part — but the data generated is certainly primary.

We finish this section with a short discussion about sampling. There are several important points to note. The first is that sampling is just as important to human geography as it is to physical geography. The second is that sampling is fundamentally at the heart of the subject of Statistics.

For an excellent discussion of sampling within the context of (mainly) social science statistics refer to the companion subject guide on 04a Statistics 1 that you have to hand. Chapter 9 is the relevant chapter and the discussion is just as relevant to social survey as it is to proper sampling of physical geography subject matter. Perhaps the only thing that is different for geographers is the spatial dimension. Just how do we ensure representative spatial sampling? The answer, of course, is that we must stratify our samples spatially.

We could, for example, take a map and simply sample locations at random from that map. This could well leave large areas of the map not sampled. We could divide the map into areas of particular interest – say the east or the west of the map and sample more in one part of the map (for some reason) than the other. Better still, if we wanted spatial coverage, we could undertake a systematic sample over a grid space where we could sample, say, at the intersections of grid lines straddling the area of enquiry. Or, further, we could undertake a nested random sample systematically arranged over the grid squares of a systematic sample. It is often helpful, while retaining spatial coverage of an area, to weight the sample according to some underlying distribution or features such as population. Basically we want to be in a position to make statements about the characteristics of the variable that we are measuring over the area that we are looking at. Needless to say great thought has to be placed on sample design.

**Activity**

Revise Chapter 9 of 04a Statistics 1 to ensure that you know the meaning of random sampling, quota sampling stratification and clustering. Then switch into spatial mode and consider the problems associated with design of a spatial sample (e.g. the survey of the opinion of inner city residents about socio-economic wellbeing in the city). Just how would you go about this tricky task and what considerations would you take on board? A good prompt for thinking about this activity is section 3.4.2 of Wheeler et al. (2004) on sampling methods in geographical settings. The question of sample size is important here, but you will need to know more about statistics before we can say more about this important topic. Suffice to say at this stage that the special circumstances of small sampling theory is said to stop at the sample size of 30 plus.
Secondary data: physical geography

Although it is true that there are not quite so many secondary sources that can be accessed for physical geographers, there are still plenty. It is also fair to say that many useful sources for physical geographers do also tend to make charges. Wheeler et al. (2004) discuss some of these sources (free as well as chargeable) and especially recommend the UK Meteorological Office:

www.metoffice.com

for information about weather and climate. They also point us to the UK National River Flow Archive at the Natural Environmental Research Council Centre for Ecology and Hydrology for information about UK rivers; free at:

www.ceh.ac.uk/data/nrfa

If you are more interested in the physical wonders of the USA then you could try the renowned United States Geological Survey at:

www.usgs.gov

which more than likely will see you try the national spatial data infrastructure from the Federal Geographic Data Committee at:

www.fgdc.gov

Activity

This is a superb experience for physical geographers and especially those interested in weather and climate. What follows is an attempt to pilot you through a fantastic but complicated website provided by the UK Meteorological Office. It is not too vital that you find exactly the data that is specified here, for to discover and understand any of the data that are provided will be great. We will focus on UK data, but please note that this site is an archive of some wonderful worldwide climate data that might be much more interesting to you. The main task is to show you the site and encourage you to ‘grab’ some data and put them into a form that can be analysed (an Excel spreadsheet). You may well want to experiment with the capture of other data from this superb source. Anyway hold onto your seats!

Let’s begin by logging on to:

www.metoffice.com

You will receive a nice Welcome to the Met Office.

Search for Education in the menu options at the top of the page called and click to give a page headed Learning centre.

Search the menu options at the top of the page called Higher and click to give a page headed A-level and higher education.

Then under a heading called Observations and charts click a bullet point called weather data for researchers to give a page called Academic Research.

Search for an option bullet point called free data (always good to get something for free!) and click as this will take you further down the page to a series of options, one of which is Climate averages (do not be too concerned that we have not yet discussed averages formally — although most of you will have heard of the term average before).

Under Long-term averages, 1971–2000 click a bullet point called station, district and regional averages to reveal, at last, the data that we are looking for. You are presented with a map that you can click on to obtain the weather information for quite a number of weather stations in the UK. Try clicking on the very most northern station and you will see the results for Lerwick in Scotland. The results consist of monthly averages for Max Temp, Min Temp, Days of Air Frost, Sunshine, Rainfall, Days of Rainfall >= 1mm and Wind at 10m.
The task for this activity is to select, one at a time, 10 weather stations scattered throughout the UK, copy the weather data and paste them into an Excel spreadsheet. Exactly how you do this and what ‘tidying up’ you do is up to you, but the final product should be a workable spreadsheet for 10 places, by 12 months, for 7 variables. You will need to refer back the above discussion on geographical data matrices for some hints and help.

Finally, record the steps that you have taken and the difficulties that you have encountered en route to deriving you secondary source spreadsheet of UK weather station data.

**A reminder of your learning outcomes**

Having completed this chapter, and the relevant reading and activities, you should now:

- understand that data are measured in various ways, some of which are more useful than others, and note that any form of measurement is not without problems associated with error
- put the ‘geography’ into geographical data by considering them in the form(s) of a geographical data matrix (and link this to a spreadsheet); you should be able to put these ideas into practice
- understand that although all data has to be collected, some data has already been collected by others for us to use
- design elementary questions for a questionnaire survey
- understand what sampling means in a geographical context
- search out geographical data (physical and human) collected from secondary sources and experiment with the practical capturing of these data in geographical data matrix and spreadsheet form so that they are ready for analysis.

**Sample examination questions**

1. With respect to geographical data, explain how we measure variables of interest along various scales and according to various degrees of error.
2. In the light of detailed geographical data with which you are familiar, discuss the nature and form that the geographical data matrix can take.
3. Considering any geographical data set that you have looked at – primary or secondary – discuss the detail of the specification of the variables measured and the geographical scale of analysis.
4. Why is it important for geographers to move beyond mere random sampling of geographical information?
5. Discuss the problems that might be encountered in designing questions for inclusion in a questionnaire survey.
Notes
Chapter 4: Presenting data and mapping geographical distributions

Essential reading


Chapter 3 of Wheeler et al. is useful for some of the data acquisition tasks and the drawing of graphs and charts but has little to say about mapping. The dictionary by Johnston et al., however, has useful and substantial entries on ‘map image and map’ and ‘choropleth’ mapping as well as many other useful references to things like map ‘scale’ and map ‘symbolisation’.

Further reading

Clarke, K. Getting Started with Geographic Information Systems. (New Jersey: Prentice Hall, 2003) fourth edition [ISBN 0130460273]. Although this is obviously a text about the new and important area of Geographical Information Systems (GIS), it is useful for cartography and mapping issues.

Aims of the chapter

The purpose of this chapter is to introduce you to the basic ways in which geographical data can be presented. We will also be capturing some data from a remote source as we did in the last chapter.

- We begin by talking about tabular data, which is not often considered, but is probably the main way in which geographical data is presented in research papers and books. We will examine the way in which tables of data can be transformed (not always the case) into usable data in the form of an Excel spreadsheet.

- The second component of the chapter deals with one type of graphical presentation of geographical data: graphs and charts. There are various sorts available depending on whether we want to present data that represents one or several variables.

- Finally, we consider the most important graphical method of data presentation for geographers: mapping. This last theme, especially, is a large topic the full treatment of which is beyond the scope of this half course. We will consider the basic ideas of mapping and point you to some sources where you can find out more.

For each of these three themes we will be using the computer via the Excel spreadsheet to help us with our work.
Learning outcomes

By the end of this chapter, and having completed the relevant reading and activities, you should be able to:

• explain the three main ways in which geographical data can be presented
• show that the design of useful tables of data does not happen accidentally, for considerable thought and effort needs to go into the planning
• consolidate your experience of data capture into machine-readable form from a remote source (skills that you acquired in the last chapter)
• explain the basic types of graphs and charts that can be used to present geographical data in the knowledge that some are more useful for some purposes than others
• review the basic ways in which maps can be drawn, again in the knowledge that the various types of geographical data (points, lines, areas and volumes) are more suited to some mapping formats than others
• present geographical data using the computer by understanding the ways in which the Excel spreadsheet operates.

Presenting geographical data

The results of geographical – or any other type of – analysis need to be presented in a form that is accessible to the reader. In many instances, the written word is the most useful descriptor of research findings and even the various forms of tables, graphs, charts and maps require some form of written annotation.

However, in this chapter we are going to concentrate almost wholly on the various forms of graphic presentation, and we will begin by considering how results can be presented in tabular form. Before we begin, we must stress that although we are considering the notion of ‘presentation’ here, it is important to understand that many graphics used in geographical analysis also form part of the analysis. Commentaries are written on mapped distributions and the maps are not just the final outcome of some analysis (although sometimes they do constitute the final product). Also it should be said that in contemporary human geography research much use is made of photographs and other forms of illustrative material and again, in this case, this usually means that these illustrations are part of the interpretative research process itself.

If you read the geographical research journals you will be struck by the fact that many do not contain graphics at all. It would be fair to say that most contain information presented in tabular form but, depending on the specialism, some do not contain anything other than words!

Tables

We begin our consideration of presenting geographical data with tables. Tables of results should be clear, concise and easily understood; they should be capable of being understood without further commentary.

There are a number of essentials that tables require and in this sense they are not dissimilar to charts and maps:
• A title describing the information presented and its provenance is helpful.
• Footnotes dealing with any technical terms that are used.
• Note that it is usually not necessary to provide full details of any well-known statistics that might have been calculated.

Tables often involve presenting rows and columns of data, and these need to have concise and carefully crafted headings. The numerical content of the table also needs to be carefully considered and when thinking about this, the **appropriateness for purpose** is paramount:

• The precision of the numbers (decimals) needs to be appropriate for purpose.
• The detail of the information also needs to be appropriate for purpose.

The most effective tables for getting a point across are often those that contain a highly selective range of numerical data. But sometimes tables of data are presented in great detail because the prime purpose is to provide the reader with an information source that will be useful for subsequent further analysis.

In Chapter 3 we referred to a very useful, free data source for physical geographers interested in rivers: the UK National River Flow Archive. This source will be accessed here to complete our short discussion about presenting tabular data (and to provide a link to transforming tables to spreadsheets for the next part of our work). The source was:

www.ceh.ac.uk/data/nrfa

When you look at this site you will find a rich source of information about water flowing through a whole range of UK rivers measured at frequent intervals over time by special water recording gauges. Figure 4.1 is a table of data selected from this resource. It does not have a specific title as it is one of many similar tables for many rivers in the UK, measured over a lengthy time series (in fact the title is elsewhere in the wider presentation of River Flow Data). This is obviously a tabular presentation of data for subsequent analysis and consequently it contains a great deal of detail. But notice how clearly presented the information is. We can note:

• that the data refers to the River Tees at a place called Broken Scar and that the data are for 2005
• some locational detail and background information (and there is a lot more available elsewhere on this site concerning this place and this river)
• some fine detail for each day of each month of the daily mean gauged discharges measured in cubic metres per second (we will be considering descriptive statistics in the next chapter)
• some summary statistics about this yearly record
• summary statistics for the previous record (1956–2004) about mean flows runoff and rainfall
• related summary statistics for 2005 and some fine detail on the nature of the site where the data are collected.

The important point to note is that tables of data like this do not happen by accident. A considerable degree of care has gone into their construction to make them the valuable source that they are.
| DAY | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 4   | 33.800 | 9.400 | 20.400 | 7.510 | 15.200 | 8.180 | 4.520 | 3.420 | 3.690 | 4.320 | 70.500 | 35.000 |

**Figure 4.1: Daily mean gauged discharges for the River Tees at Broken Scar 2005.**
(Source: UK National River Flow Archive © Natural Environment Research Council (NERC))

**Tees at Broken Scar**

<table>
<thead>
<tr>
<th>Measuring authority: EA</th>
<th>Grid reference:</th>
<th>Catchment area (sq km):</th>
</tr>
</thead>
<tbody>
<tr>
<td>First year: 1956</td>
<td>(NZ) 259 137</td>
<td>818.4</td>
</tr>
<tr>
<td>Level stn. (m OD):</td>
<td>37.20</td>
<td>Max alt. (m OD):</td>
</tr>
<tr>
<td>Daily mean gauged discharges (cubic metres per second)</td>
<td>4.38</td>
<td>893</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>--------</td>
</tr>
<tr>
<td>30</td>
<td>11.100</td>
<td>12.100</td>
</tr>
</tbody>
</table>

Highest: 340.000  53.600  63.700  62.300  19.700  11.100  17.200  37.100  123.000  82.600  69.900

### Monthly total

<table>
<thead>
<tr>
<th></th>
<th>(million cu m)</th>
<th>116.60</th>
<th>31.91</th>
<th>55.56</th>
<th>45.28</th>
<th>21.38</th>
<th>15.72</th>
<th>12.35</th>
<th>14.23</th>
<th>20.05</th>
<th>43.35</th>
<th>56.34</th>
<th>38.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff (mm)</td>
<td>143</td>
<td>39</td>
<td>68</td>
<td>55</td>
<td>26</td>
<td>19</td>
<td>15</td>
<td>17</td>
<td>24</td>
<td>53</td>
<td>69</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Rainfall (mm)</td>
<td>155</td>
<td>75</td>
<td>65</td>
<td>105</td>
<td>61</td>
<td>47</td>
<td>60</td>
<td>79</td>
<td>94</td>
<td>121</td>
<td>99</td>
<td>71</td>
<td></td>
</tr>
</tbody>
</table>

### Statistics of monthly data for previous record (Oct 1956 to Dec 2004—Incomplete or missing months total 0.1 years)

<table>
<thead>
<tr>
<th>Mean</th>
<th>Date (year)</th>
<th>Low</th>
<th>Date (year)</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg</td>
<td>28.770</td>
<td>2.907</td>
<td>1963</td>
<td>50.090</td>
</tr>
<tr>
<td></td>
<td>25.430</td>
<td>2.803</td>
<td>1963</td>
<td>64.280</td>
</tr>
<tr>
<td></td>
<td>22.030</td>
<td>5.480</td>
<td>1975</td>
<td>68.660</td>
</tr>
<tr>
<td></td>
<td>15.970</td>
<td>2.538</td>
<td>1957</td>
<td>39.040</td>
</tr>
<tr>
<td></td>
<td>9.492</td>
<td>2.099</td>
<td>1959</td>
<td>27.020</td>
</tr>
<tr>
<td></td>
<td>6.969</td>
<td>0.502</td>
<td>1957</td>
<td>21.070</td>
</tr>
<tr>
<td></td>
<td>9.275</td>
<td>0.458</td>
<td>1959</td>
<td>34.100</td>
</tr>
<tr>
<td></td>
<td>10.410</td>
<td>0.637</td>
<td>1959</td>
<td>24.350</td>
</tr>
<tr>
<td></td>
<td>17.640</td>
<td>2.709</td>
<td>1969</td>
<td>53.940</td>
</tr>
<tr>
<td></td>
<td>21.720</td>
<td>4.061</td>
<td>1958</td>
<td>58.950</td>
</tr>
<tr>
<td></td>
<td>27.590</td>
<td>5.780</td>
<td>1971</td>
<td>50.040</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Runoff: Avg</th>
<th>Date (year)</th>
<th>Low</th>
<th>Date (year)</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>1988</td>
<td>10</td>
<td>1998</td>
<td>164</td>
</tr>
<tr>
<td>76</td>
<td>2002</td>
<td>8</td>
<td>1999</td>
<td>96</td>
</tr>
<tr>
<td>72</td>
<td>1997</td>
<td>18</td>
<td>2000</td>
<td>93</td>
</tr>
<tr>
<td>51</td>
<td>1996</td>
<td>8</td>
<td>2001</td>
<td>78</td>
</tr>
<tr>
<td>31</td>
<td>2004</td>
<td>7</td>
<td>2002</td>
<td>73</td>
</tr>
<tr>
<td>22</td>
<td>2004</td>
<td>2</td>
<td>2003</td>
<td>73</td>
</tr>
<tr>
<td>21</td>
<td>2005</td>
<td>6</td>
<td>2004</td>
<td>77</td>
</tr>
<tr>
<td>30</td>
<td>2006</td>
<td>2</td>
<td>2005</td>
<td>95</td>
</tr>
<tr>
<td>33</td>
<td>2007</td>
<td>9</td>
<td>2006</td>
<td>94</td>
</tr>
<tr>
<td>58</td>
<td>2008</td>
<td>13</td>
<td>2007</td>
<td>109</td>
</tr>
<tr>
<td>69</td>
<td>2009</td>
<td>19</td>
<td>2008</td>
<td>113</td>
</tr>
<tr>
<td>90</td>
<td>2010</td>
<td>19</td>
<td>2009</td>
<td>127</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rainfall: Avg</th>
<th>Date (year)</th>
<th>Low</th>
<th>Date (year)</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>1988</td>
<td>16</td>
<td>1989</td>
<td>213</td>
</tr>
<tr>
<td>96</td>
<td>2002</td>
<td>16</td>
<td>2003</td>
<td>246</td>
</tr>
<tr>
<td>93</td>
<td>1997</td>
<td>10</td>
<td>1998</td>
<td>224</td>
</tr>
<tr>
<td>78</td>
<td>1996</td>
<td>18</td>
<td>1997</td>
<td>164</td>
</tr>
<tr>
<td>73</td>
<td>2004</td>
<td>19</td>
<td>2005</td>
<td>167</td>
</tr>
<tr>
<td>73</td>
<td>2005</td>
<td>20</td>
<td>2006</td>
<td>182</td>
</tr>
<tr>
<td>77</td>
<td>2006</td>
<td>12</td>
<td>2007</td>
<td>206</td>
</tr>
<tr>
<td>95</td>
<td>2007</td>
<td>19</td>
<td>2008</td>
<td>256</td>
</tr>
<tr>
<td>94</td>
<td>2008</td>
<td>27</td>
<td>2009</td>
<td>222</td>
</tr>
<tr>
<td>109</td>
<td>2009</td>
<td>25</td>
<td>2010</td>
<td>230</td>
</tr>
<tr>
<td>113</td>
<td>2010</td>
<td>43</td>
<td>2011</td>
<td>239</td>
</tr>
<tr>
<td>127</td>
<td>2011</td>
<td>68</td>
<td>2012</td>
<td>268</td>
</tr>
</tbody>
</table>
### Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>For 2005</th>
<th>For record preceding 2005</th>
<th>As % of $ Reservoir(s) in catchment. pre-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean flow</td>
<td>14.930</td>
<td>16.790</td>
<td>89</td>
</tr>
<tr>
<td>Lowest yearly mean</td>
<td>9.382</td>
<td>1973</td>
<td></td>
</tr>
<tr>
<td>Highest yearly mean</td>
<td>24.000</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Lowest monthly mean</td>
<td>4.611</td>
<td>Jul 0.458 Aug 1959</td>
<td></td>
</tr>
<tr>
<td>Highest monthly mean</td>
<td>43.550</td>
<td>Jan 68.660 Mar 1979</td>
<td></td>
</tr>
<tr>
<td>Lowest daily mean</td>
<td>3.110</td>
<td>17 Aug 0.023 16 Oct 1959</td>
<td></td>
</tr>
<tr>
<td>Highest daily mean</td>
<td>340.000</td>
<td>7 Jan 410.500 31 Jan 1995</td>
<td></td>
</tr>
<tr>
<td>10% exceedance</td>
<td>31.370</td>
<td>40.330</td>
<td>78</td>
</tr>
<tr>
<td>50% exceedance</td>
<td>8.923</td>
<td>8.052</td>
<td>111</td>
</tr>
<tr>
<td>95% exceedance</td>
<td>3.615</td>
<td>1.827</td>
<td>198</td>
</tr>
<tr>
<td>Annual total (million cu m)</td>
<td>470.80</td>
<td>529.90</td>
<td>89</td>
</tr>
<tr>
<td>Annual runoff (mm)</td>
<td>575</td>
<td>647</td>
<td>89</td>
</tr>
<tr>
<td>Annual rainfall (mm)</td>
<td>1032</td>
<td>1151</td>
<td>90</td>
</tr>
</tbody>
</table>

### Factors affecting runoff

- Abstraction for public water supplies
- Augmentation from surface water and/or groundwater.

### Station and catchment description

Compound Crump profile weir with total crest length of 63.9m. Two low-flow crests total 9.1m. Rating review completed by HR Wallingford in 1998. Revised rating serves to reduce high flows and has been used to reprocess flows back to 1982. Significant export of water from direct supply reservoirs and u/s abstraction. Some regulation from Cow Green Res. Transfers of water from Kielder in drought years. A mainly impervious catchment developed on Millstone Grit and Carboniferous Limestone. Headwaters drain the Pennines. Moorland and rough pasture give way to more intensive agriculture in the lower reaches.
It is also possible, with a little effort, to capture much useful numerical information from tables like Figure 4.1. Figure 4.2 is part of an Excel spreadsheet of most of the information that is contained in Figure 4.1. In the activity that follows, we ask you to produce something similar.

Activity

Access the UK National River Archive, find the River Flow Data, select a river of your choice, print a table for that river (similar to Figure 4.1), save that table and convert the information into an Excel spreadsheet suitable for subsequent analysis. The steps are as follows:

1. Find the website (www.ceh.ac.uk/data/nrfa).
3. Click that you accept the conditions of use, then access the excellent location map of the regions and rivers in the UK, or scroll down to find a list of all the regions and rivers in the database.
4. Select a river.
5. Note the background information.
6. Find the Annual Flow Tabulations.
7. Select a year – perhaps 2005 which is the most recent for which data are available.
8. Print out your table of data.
9. Save your data as a Text File
10. Open Excel and open the Text File where the Text Import Wizard will help you to save the basic format of the table using column pointers.

Figure 4.2: Excel spreadsheet of daily mean gauged discharges for the River Tees at Broken Scar 2005.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DAY</td>
<td>JAN</td>
<td>FEB</td>
<td>MAR</td>
<td>APR</td>
<td>MAY</td>
<td>JUN</td>
<td>JUL</td>
<td>AUG</td>
<td>SEP</td>
<td>OCT</td>
<td>NOV</td>
</tr>
<tr>
<td>2</td>
<td>37.6</td>
<td>10.2</td>
<td>14.4</td>
<td>9.3</td>
<td>6.95</td>
<td>5.21</td>
<td>5.06</td>
<td>4.9</td>
<td>5.33</td>
<td>17.8</td>
<td>16.2</td>
<td>14.2</td>
</tr>
<tr>
<td>3</td>
<td>29.3</td>
<td>13.2</td>
<td>7.96</td>
<td>7.06</td>
<td>11.1</td>
<td>4.47</td>
<td>3.95</td>
<td>4.33</td>
<td>9.47</td>
<td>16.3</td>
<td>34.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>39.4</td>
<td>9.62</td>
<td>12.2</td>
<td>7.41</td>
<td>19.7</td>
<td>10.4</td>
<td>4.33</td>
<td>3.93</td>
<td>3.76</td>
<td>6.63</td>
<td>50.6</td>
<td>34.2</td>
</tr>
<tr>
<td>5</td>
<td>33.8</td>
<td>9.4</td>
<td>20.4</td>
<td>7.51</td>
<td>15.2</td>
<td>5.18</td>
<td>4.62</td>
<td>3.42</td>
<td>3.59</td>
<td>4.32</td>
<td>70.5</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>9.82</td>
<td>17.3</td>
<td>16</td>
<td>10.7</td>
<td>10.8</td>
<td>4.91</td>
<td>4.23</td>
<td>3.53</td>
<td>4.6</td>
<td>22.7</td>
<td>18.7</td>
</tr>
</tbody>
</table>
We now have an Excel file of data for one river in the UK (and remember for research purposes you could have many more). Let us now turn to graphical representation of this geographical data. Again we are going to depend on what the Excel spreadsheet offers us.

The first form of graphical representation that we should consider is the simple bar chart. This is especially helpful to record a series of values for a number of observations, for example, if we wanted to construct a hydrograph of stream flow for the River Tees that is recorded in Figure 4.2. The result can be found in Figure 4.3. For those mainly human geographers among you who are concerned about the notion of a hydrograph, refer to the following for clarification:


How was this diagram constructed in Excel? First, call up the spreadsheet of data (Figure 4.2). Then highlight the data that you want to place in the graph – in this case the average monthly values of stream flow. Then pull down the chart wizard menu and click on **bar chart**. (The chart wizard can be found at the top of the page and is a multi-coloured icon looking like a bar chart.) The result is Figure 4.3. Of course, it would be helpful to undertake some proper labelling and assign a tile such as 'Hydrograph of stream flow for the River Tees in 2005'. But this and subsequent graphics have been left in this guide to show simply the raw output. In the activities that follow it would good practice for you to persevere with some proper labelling and titling. Note the bar chart procedure will be extremely useful in the next chapter for the drawing of histograms.

![Figure 4.3](image-url)

*Figure 4.3: Bar chart of average values of daily mean discharges monthly for the River Tees at Broken Scar 2005.*
Activity

Generate a hydrograph for the stream flow of the river that you have selected for the previous activity (and explain briefly the meaning of the graph).

A refinement of the simple bar chart is the divided bar chart. This is useful in the case of two streams of related numbers being compared. In the current example, two such streams of numbers could be the lowest and the highest values of the daily mean gauged discharges in Figure 4.2. The resultant divided bar chart is given in Figure 4.4 and it can be seen how easy it is to compare the relative streams of values month by month.

![Divided bar chart](image)

**Figure 4.4: Divided bar chart of lowest and highest values of daily mean discharges monthly for the River Tees at Broken Scar 2005.**

Activity

Generate a divided bar chart of the lowest and highest values of daily mean discharges for the river that you have selected for the previous activity (and explain briefly the meaning of the graph).

You should by now be able to see the possibilities that the Excel charting options afford you – some charts are possible even in three dimensions! But remember - simplicity is the key.

There is one further useful graphic that geographers and others make much use of and that is the pie chart. This graphic makes most sense in terms of percentages. Take the data in Figure 4.2 regarding monthly totals of discharges (millions of cubic metres) and convert them into the percentages (already undertaken in Figure 4.2). The resultant pie chart is shown in Figure 4.5 which depicts the monthly proportions of total daily gauged discharges where it is clear that two or three months comprise the majority of such values over the year.
Activity

Generate a pie chart of the monthly totals of discharges as a percentage of total discharges for the river that you have selected for the previous activity (and explain briefly the meaning of the chart).

Lastly we should consider the simple line graph as a means of presenting geographical data. Say we wanted to present the two monthly streams of numbers relating to average runoff and rainfall for the River Tees in 2005. The result would be Figure 4.6 where interpretation of the two lines is very much dependent on the data in Figure 4.1 as the two lines cross over.
Activity

Generate a multiple line graph of the monthly averages of runoff and rainfall for the river that you have selected for the previous activity (and explain briefly the meaning of the chart).

Please note again that your equivalents will be much more useful if you take the time to label and title them.

We will return to the graphical presentation of geographical data in each of the next two chapters.

Mapping geographical distributions

When people think of geography or geographers they think of maps. Mapping, or more properly cartography, is by no means a small topic. There are many texts on the subject and in some universities degree courses are offered in this discipline alone. What follows cannot be more than the briefest of introductions to this important topic. We are not going to consider map projections, their history or the reasons why we need them. We are not going to concern ourselves with the fast growing interest in geography and elsewhere in Geographical Information Systems (GIS). Both of these subjects are worthy of extensive study in their own right. In this introductory half course on methods we are just going to cover the basics.

It is important to understand that maps have their own norms and conventions. Maps have a distinctive anatomy consisting of some fundamental basic elements:

- The map needs to be located in its geographical context.
- It needs a north point and a scale to make much sense.
- It needs a title and a legend or key to the various components depicted on the map.

Clarke (2003) discusses all of these things and others in the context of the emerging subject of Geographical Information Systems. These GIS are computer-based mapping systems that allow a range of added value tasks to be undertaken. Maps can be ‘interrogated’ in the computer for a whole range of search tasks that provide many practical solutions in everyday life. The new possibilities for computer mapping have also seen the development of greater skills in map design in terms of the layout of the maps and especially the use of colour.

Clarke is also useful when considering the important matter of choosing a map type. Some types of maps are more suited to a particular purpose than others. Some maps show many things and we have all found such maps useful for everyday use. They may be termed general-purpose maps.

More useful for professional or research purposes, however, are thematic maps which, as their name implies, focus on a particular theme (or themes). Of considerable use to geographers doing research are basic reference maps. These often depict the basic boundaries that geographical data refer to or are collected for. A more general reference map, that we are all familiar with, is the topographic map. There are few countries whose mapping agencies do not produce a version of this. Such maps depict a basic mix of human (settlement and transportation) and physical (rivers and land terrain contours) characteristics. Interestingly, they have become the favoured base maps for Geographical Information Systems.
There are several different types of more specialist maps. **Dot maps** (and their variants in picture symbol maps and proportional circle maps) are useful when specifically point-based data need to be mapped. Typical examples of such mapping are the proportional circle maps representing the population of the various cities or regions in a country. These maps are especially valuable when the quantity to be mapped is an absolute numerical value. **Network maps** depict links and routes and flows. Everyone is familiar with road maps, but other flow networks such as river systems are also mapped. Sometimes in network flow maps, links are drawn proportional to the strength of the flows taking place.

The most important map type in geographical use is the **choropleth map**. This is the familiar shaded map where density of shading relates to some particular feature that the map is intending to portray. Many variables of interest to geographers – say altitude above sea level or population density – can be mapped in this way. Of critical importance are the classes used to categorise the variable selected for mapping (and this links with frequency distributions which are considered in the next chapter), as well as the shading scales chosen to represent these classes. As might be expected there are many conventions – some would say rules – which have been adopted for the drawing of proper choropleth maps.

Whereas choropleth maps may be thought of as two dimensional, even though they may be trying to illustrate three dimensional features, there are a range of maps that have been designed to cover the attributes of volumetric data. The most widely known is, of course, the contour map, which is based on lines of equal value or isolines, but there are other forms. It is interesting that these map types are notoriously difficult to draw by hand, as well as by computer, as much interpolation is required to obtain the correct spacing between the contours or **isolines**.

In summary, it is helpful to realise that different map types most appropriately deal with different map features. Dot maps are good at representing points. Lines are best represented by flow or network maps. Areal data and choropleth mapping goes hand in hand. And finally, volumes are particularly well represented by various forms of isoline maps.

The text by Clarke (2003) has already been mentioned and it helpfully has an associated companion website where several useful things (such as PowerPoint slides of illustrations used in the book) can be found. Particularly useful are Chapters 2 (‘GIS's Roots in Cartography’) and 7 (‘Making Maps with GIS’) and you can enter these titles in the search facility having found the website link below

http://wps.prenhall.com/esm_clarke_gsgis_4

Computer packages for the development of Geographical Information Systems are expensive to buy and complicated to learn. They are far beyond the scope of this half course. But it is important to do some mapping here and so we have selected the somewhat blunt mapping tool in the form of the Excel spreadsheet to provide that introduction. It is important to understand that this is nothing like a GIS, but some progress can be made and some useful products generated using this computer technology.

The key to using Excel for mapping is to understand that it can only map data for which it has a set of areas. The boundary sets that the package contains are limited indeed. We will restrict our activities here to the mapping of data at a national level.

The data that we will use for this small mapping exercise comes from the World Bank and is somewhat time-consuming to acquire. In the meantime, it is an excellent discipline to have to access the source of the data and go
through the process of extracting the data yourself. The following takes you through the steps required to access the data and draw the map. The activity that follows will require a retracing of these steps. First, we need the data, so access the following resource:


This is the site of the World Bank where they provide free of charge to individuals a small subset of the development data that they hold for most of the countries of the world. The data is served up in small amounts where the user can request the countries that are needed, what variables are needed and what dates are needed. It is probably best to keep things simple and go for one variable at a time, one year at a time and lots of countries. This is what has been done here and the steps are as follows:

a. Having accessed the above site you will be faced with a query menu. First you need to select the countries that you require data for. For our purposes it is probably best to select all countries (by holding the shift key and dragging down, then pressing ‘select’).

b. Next you need to select the variables required. In our case we select GNI per capita, Atlas Method (current US $). GNI means Gross National Income. This is a measure of wealth generation (some would link this to levels of development). In this case it is measured per head of population in current US$. It says nothing about the distribution of wealth within the population. The World Bank calculates the measure in a specific way – termed the Atlas Method.

c. Next, a range of years for which data are available is presented. In our case, we will select the year 2000. Any available year could be selected but the choice of 2000 was made in anticipation that most of the data we require will be available by now. Clearly, though, this is an excellent resource for undertaking some change over time analysis.

d. Two further options are then presented. For orientation, select ‘by countries’ (this is just about the way the geographical data matrix is presented). For data export options, select ‘save data as Excel file’. Then view the data that you have captured in Excel form and if it looks as if it’s what you want, save the file and label it.

e. Unfortunately, as with all data sets, invariably there are missing data values. In the mapping procedure that we are going to undertake next we cannot have missing values so these need to be deleted. It is a simple matter in Excel: just delete those rows (countries) with missing values for GNI.

f. Lastly, it will be helpful also to delete the second column, which just repeats the name of the variable selected for each of the countries. The result is a simple spreadsheet containing two columns (country names and values for GNI per capita).

Once we have a cleaned up Excel file of data simply arranged our next task is to draw the map. To do this we use the Excel mapping tool that is a symbol of a globe at the top of the screen adjacent to the Chart symbol that we have used previously in this chapter. The mapping steps are as follows:

a. Highlight the data (the countries column and the variable column)

b. Click on the mapping tool icon that generates a cursor in the form of a cross

c. Drag this cursor into the main body of the spreadsheet and by clicking it with your mouse create an empty square into which the map will be placed
d. Next you will have to resolve unknown geographic data. (This occurs because the list of countries that you have downloaded from the World Bank will not be the same as the list of countries that the blank map in the Excel spreadsheet.) Even the smallest spelling difference in the names of the countries is critical. Most of the problems will be caused by World Bank entries that are not countries at all such as the European Monetary Union. Excel has problems with country aggregates like this because it does not have a boundary for them.

e. In the Resolve Unknown Geographic Data menu you will need to discard anything that Excel finds that it cannot cope with or change a country name that it is obviously confused with to somewhere that it knows. It is not vital that you spend a long time on this but every country that you discard means that a blank space will appear on the map. Once all these queries are resolved the choropleth map is drawn automatically.

f. The mapping tool will then provide you with options to change how the map looks. It is possible to locate bar charts (of various sorts) and pie graphs in the countries for which you have information. Obviously you will need to map more than one variable in the original spreadsheet of data if you are going to opt for a more complicated symbol such as pie chart.

Figures 4.7 and 4.8 are maps drawn in this way for the World Bank variable GNI per capita. The former is a choropleth map and the latter locates (not very clearly) bar charts in the locations of countries for which data are available. You will be interested to learn of how the Excel package decides on the classes that it uses to draw the choropleth map for clearly they do not happen by accident. It is usually the case in automated mapping packages that the well-known Jenks’ method (named after the author) is used. More can be found on this method from the following resource provided by one of the large GIS companies (ESRI) but you will find that you need to have understood the next chapter’s materials at least (and know your 04a Statistics 1 well to fathom the method):

http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=26442

**Activity**

Replicate the choropleth mapping of the World Bank variable GNI per capita as above from start to finish – from data capture right through mapping using Excel. Then repeat the exercise using a different variable of your own choice – there are lots to choose from. You could also experiment with a bar chart map. You will be interested to know that Excel has several blank maps in its locker and there is no need always to produce the map at the world scale. You can draw maps for Europe or Southern Africa, for example, just by selecting them. The maps will be clearer because they will be at a larger scale. Excel also has a number of templates available for mapping at sub-national scales, for example, Australia, Canada, Mexico, the UK and the USA. Of course, you will need data at the appropriate scale, and named in a way that Excel recognises, to use these maps. The Help facility, once you are engaged in attempting some mapping using Excel, is helpful and tells you among many other things that Excel have enlisted the help of another GIS company – MapInfo – to their mapping for them.

After struggling with the automatic mapping of these variables it would be a good idea to attempt the drawing of a choropleth map by hand using just a pencil. A map of any variable would do but it would be easy to look at an Atlas and trace out an outline map of a continent of countries — say Europe or the Americas or Africa — and map, for example, the GNI per capita data that you have captured from the World Bank. Who says computers do not have their uses?
A reminder of your learning outcomes

Having completed this chapter, and the relevant reading and activities, you should now be able to:

- explain the three main ways in which geographical data can be presented
- show that the design of useful tables of data does not happen accidentally, for considerable thought and effort needs to go into the planning
- consolidate your experience of data capture into machine-readable form from a remote source (skills that you acquired in the last chapter)
• explain the basic types of graphs and charts that can be used to present geographical data in the knowledge that some are more useful for some purposes than others
• review the basic ways in which maps can be drawn, again in the knowledge that the various types of geographical data (points, lines, areas and volumes) are more suited to some mapping formats than others
• present geographical data using the computer by understanding the ways in which the Excel spreadsheet operates.

Sample examination questions

1. What sorts of issues need to be considered when selecting appropriate graphs and charts for the presentation of geographical data?

2. Explain why different sorts of geographical data need different sorts of maps for effective presentation.

3. Taking account of issues of map choice and map design, explain what makes for an effective map.

4. Explain how you would go about preparing a choropleth map of GDP per capita by country on a world scale. At the same spatial scale, if you had data on GDP according to source (agriculture, manufacturing and services), explain how you would map them.