Introduction

Quantification, even of a comparatively simple kind, arouses fear among many students, especially those who come to history with a background in arts subjects. Even those who approach history from a social-science background – students who also study sociology or politics – are easily disheartened when confronted with historical numbers. The same may be true of their teachers, who may not be frightened of introducing some quantitative study into the undergraduate curriculum, but who doubt the extent to which students will be able to use quantitative data with advantage, and who are aware of the resource implications of introducing lab sessions and the like into their courses.

Alternatively, ‘number-crunching’ is seen as the province of economic historians, something to be feared and avoided by ‘mainstream’ history students. Students who encounter regression analyses, correlation coefficients or even simple tabular data in articles they read in conventional history courses are easily alienated, and need to be encouraged to believe in their ability to use the literature to advantage, even where their understanding of the numerical techniques involved is limited. Most numerical data presented by most historians is relatively simple, and should not be difficult for students to use profitably. This ‘Overnight Expert’ guide suggests ways in which students can be helped to appreciate the importance and value – and limitations – of numbers in history, and to overcome their fears of numbers through the integration of small amounts of quantitative work into university curricula.

It also describes some of the resources that are available to historians. The intention is not to set out a curriculum for a course on quantification in history: some universities offer this to undergraduates, and some of the exercises presented in this guide might be of some use in such a course, but the focus here is on less ambitious aims. Some universities also involve students in work with numerical data in courses devoted to history and computing; again, this guide will not offer much advice on how to run a history and computing course. Many history and computing textbooks are available; some are listed in the bibliography below. I will only suggest a few avenues for exploration, and provide a list of further reading.

Numbers are encountered in all areas of history; they are not, and should not, be left to economic historians. In the course of a broad historical curriculum, undergraduate students are likely to encounter demography, prosopography, psephology, and many other areas of study in which quantification is fundamental. Even in histories which are not explicitly quantitative in approach, we find numbers: to take a random example from political history, John Vincent’s classic *Formation of the Liberal Party* (1963) contains tabular data relating to, among other things, the wealth of Liberal peers, the numbers of speeches made by Richard Cobden and John Bright in the House of Commons, the circulations of newspapers, the voting patterns of occupational groups in specific localities, and the appointment...
by different ministries of magistrates in the Duchy of Lancaster. It is a rare or perhaps non-existent professional historian who never counts anything. Moreover, even where numbers are not reported, many historical judgements rest on some kind of underlying, or ‘hidden’, quantification.

The importance of numbers: an exercise

One way of emphasising this to students is by conducting an exercise outlined by Daniel I. Greenstein, in A Historian’s Guide to Computing (1994). Robert W. Fogel, the Nobel Prize-winning economic historian, suggested in 1975 that the sceptic should, in Greenstein’s words, ‘choose a historical text at random, open it to [sic] any page, and count up the quantitative statements found there’. Greenstein uses examples from economic, diplomatic and religious history.

This exercise is a useful one to introduce students to the importance of numbers in history. I have used it as an introductory ‘ice-breaker’ in a training session for postgraduates; colleagues have used it at undergraduate level. My chosen passage is from John Stevenson’s British Society 1914-1945, in the Pelican Social History of Britain series (1984: p. 381). I chose this extract partly because it discusses leisure and recreation, traditionally perceived by students as ‘soft’ subject areas which are less ‘difficult’ than economic history. Of course, historians of these areas are reliant on official and other statistics, as well as contemporaneous impressionistic accounts, when they discern larger trends and patterns in their social histories.

The passage, with the quantitative statements or judgements underlined, would look like this (others might do it differently, or take issue with some of the underlinings):

One of the most important developments in twentieth-century society has been the growth of leisure and recreation. Already by the Edwardian era, many pastimes and pursuits had been fashioned or transformed to meet the needs of a primarily urban and industrial society. In sport, entertainment and private recreations, one of the major driving forces was commercialization, drawing upon the increased spending power of a mass consumer market. Another was the increased leisure time available as a result of shorter working hours, paid holidays, longer life expectancy after retirement, smaller families and, for some, enforced idleness through unemployment. But the growth of leisure illustrates more than commercialism and more free time from work. With the growth of the media, it was part of the development of a more uniform and homogenous society, partaking of an increasingly common culture. Notwithstanding regional and class differences, by 1945 only the remotest parts of Britain were insulated against the pervasive influences of the latest popular tune or major sporting event. In contrast, some aspects of leisure, particularly those centred around hobbies and domestic life, reflected an increasing home-centredness. The two themes of an increasingly common culture, balanced by the cult of domesticity and individual choice, dominated the development of leisure in this period.

Several points might be made about this text. Some of the reliance on quantification is obvious: for example, Stevenson’s comments about longer life expectancy, shorter working hours and increased spending power. Some identifiable trends — for example, the ‘growth’ of the media — reflect changes that are quantifiable, at least to an extent. Others are more subtle. For example, Stevenson refers to the remotest parts of Britain: one defining characteristic of remoteness is low population density, another quantitative measure. The identification of a tune as ‘popular’ or a sporting event as ‘major’ can pose quantitative questions: how many spectators make an event ‘major’? In these cases, the numerical dimension is not the only component — the cultural significance of a sporting event might transcend such considerations — but it is important.

Naturally, this is only an introductory exercise; some teachers of history may consider it rather patronising. It is also rather difficult to integrate into a teaching session other than one devoted to the study of quantitative data. Its main purpose is one of emphasis, and it might be used to overcome suspicion of what a session on quantification might entail. The idea behind the exercise, however, is one that is almost...
universally applicable: specialists in any area of history can easily select a passage, almost at random, and subject it to the same sort of interrogation.

The place of quantification in history

It is recognised that existing courses may not offer any obvious opportunities for the introduction of formal assessment based on numerical analysis or a computing project; however, it is still possible to encourage students, within normal learning and teaching environments, to think about numerical data and its importance. This can be done simply through the selection of appropriate reading material. If students are able to engage with historical arguments that rest on numerical data and its interpretation, they will be encouraged to think beyond the figures towards what they do – and do not – designate and explain. One way of doing this is to conduct a debate on the value and limitations of quantitative methods in history. Such a session could form part of a course on historiography, for example, and has also been used in courses introducing students to history and computing, in which quantification obviously plays an important role. It could also naturally arise from courses on historical topics in which quantification has been important: for example, American slavery, the standard of living during the industrial revolution, British political history in the eighteenth century, and so on.

The emergence of the ‘new economic history’ in the 1960s represented a fundamental challenge to the existing practices of historians. By exploiting new sources – or at least, by identifying patterns revealed in sources which had hitherto remained hidden – the new economic historians used quantification to challenge some established historiographical positions, and to suggest new methods for achieving a better understanding of the past. In many cases this involved bringing new source material to bear on old debates – for example, the use of anthropometric data, which is considered in more detail below – and in others it involved extracting the quantifiable material from other sources with which historians were already familiar.

We should remember, then, that quantitative research can be simply a matter of using established sources in a different way from other historians. Roderick Floud has described the ‘excitement’ of ‘statistical analysis which reveals patterns which no-one had hitherto discerned’. Nevertheless, the excitement generated by detailed empirical work using new technological apparatus did not prevent this work being subjected to serious scrutiny, and it did not prevent widespread scepticism about the value of quantification in history. It is comparatively easy to establish a debate among students about the value of quantification.

It would be possible to select texts from virtually any area of history which rely on quantification. However, Robert W. Fogel and Stanley L. Engerman’s influential and important book *Time on the cross: the economics of American Negro slavery*, first published in 1974, remains a good starting point. The book presented the results of detailed quantitative research into many aspects of the slave economy and slave life, and challenged much of the accepted wisdom on slavery. For example, Fogel and Engerman argued that slave labour was profitable and efficient; that the material deprivations of slaves had been exaggerated; that the antebellum southern economy was growing in the years before the Civil War. Unsurprisingly, the book was heavily criticised, both for its methods and its findings: one historian called it ‘amoral’. Students can be directed to the introduction, which sets out what Fogel and Engerman believed quantitative methods could achieve, and also to their afterword to the 1989 (and 1995) Norton edition, which responds to some of the criticisms of the book.

Fogel, together with the British constitutional historian Geoffrey Elton, contributed to a helpful book which introduces students, in the form of a debate, to the differences between ‘traditional’ historians and the ‘new economic historians’. This book is *Which road to the past?* It should be read with caution, because it is now 20 years old and reflects a situation in which the mutual suspicions of different groups of historians were particularly acute. Students might also note the different subject-matter studied by the two historians: Fogel was an economic and demographic historian, Elton concentrated on high politics and diplomacy,
and perhaps different methodologies were felt appropriate to different subject-matters. There is also little indication of the growing importance of social and cultural history, which complicates the picture much more. An alternative, and much shorter, source is Mark R. Horowitz, ‘Which road to the past?’, History Today 34 (January 1984), pp. 5-10.

It is also useful to counsel students against thoughtless reliance on data that is not comparable over time. The use of numerical data can help students think about the nature of historical source material, and its limitations as well as its strengths. A text which emphasises this, and is critical of the reliance of some historians on long series of data, covering very long periods, and encourages ‘short views’, is D. C. M. Platt, Mickey Mouse numbers in world history: the short view (1989).

General exercises in quantitative history can be assessed in a number of ways. One possibility would be an article report: students can choose a journal article based on quantitative evidence, and write an evaluation, perhaps comparing the findings of the article with alternative sources. Beware, however, of giving students a free choice of articles: you may find yourself with the time-consuming task of reading not only a number of evaluations, but also all the articles themselves! A lecturer can select a few articles, from his/her areas of interest, to which students can be directed.

**AHDS History and other sources: data available to historians**

Some lecturers will want to introduce more ambitious work with numerical data into the curriculum, and this involves finding some data to use. Some may have their own datasets from their research (these may require substantial modification to ensure suitability for a student audience). Alternatively, it might be worth investing the time in inputting data to create a customised dataset appropriate to students’ needs and abilities. A third option is to acquire, and perhaps adapt, existing data that is available from data archives and other sources.

A wide range of quantitative, and other, datasets is available from the Arts and Humanities Data Service for History (AHDS History). (This is the new name for the History Data Service.) Datasets from all periods and all areas are available for download, some to allcomers, some only to those registered. They are normally available free to those working at HE institutions. It is comparatively easy to download the data. The data comes in a variety of formats, but the numerical data is usually in text files, and if teachers wish to convert it into a database or other convenient file format, some knowledge of the relevant software will be necessary. Usually the standard Microsoft products can be used.

At the AHDS website it is possible to browse for datasets by period, place and theme. Each comes with a user’s guide, although the quality of these guides, usually prepared by the depositor, can vary. The staff at AHDS can help with a wide range of inquiries, and can give advice on both technical matters and on the provenance of the datasets they hold.

http://www.ahds.ac.uk/history/

We now look at three particular areas of economic, political and social history where the use of quantitative data and its interpretation might figure in an undergraduate curriculum.

**Economic and social history: the Victorian censuses**

The Victorian censuses offer perhaps the most accessible and easily usable source material for studying British economic and social history in its quantitative dimension. There are census-type sources for earlier periods of history, but the ‘mature’ Victorian censuses of the period 1841–1891 offer the teacher of history unrivalled possibilities for devising ‘hands-on’ exercises. These can be made as easy or as complex as the lecturer wishes. Alternatively, there is a range of resources available on which the teacher can draw. The published reports of the censuses — currently being digitised in a major project which will bring further benefits to academics — can be mined for statistical data on individual places; and the surviving census enumerators’ books, a staple of family history, carry individual-level data. Much of the information from the census enumerators’ books is already available in electronic format, and many historians have used it in quantitative research into the history of Victorian
communities. Some examples are given in the bibliography below.

For a number of years in my department we have used a small computerised section of the Victorian census enumerators’ books to introduce students to the use of databases in history, and to encourage them to undertake a very small amount of basic quantification. This involves no more than simple counts of occupations, ages and so on, and expressing these figures as percentages. They are then asked to consider the implications of their findings for historians of the occupational and social structure of Victorian communities. This exercise also introduces the principles and practice of occupational coding, something which some lecturers might want to explore further with more advanced students.

Historians wishing to use the census in teaching have a range of options. AHDS History has the complete 1881 census enumerators’ books in electronic format, and lecturers can create their own teaching datasets from this. From early 2005 the tabulated data from the 1831 census will be available from AHDS History (see below); some of this is already available via the Victorian Census Project at Staffordshire University (see below). There are also many small datasets available on-line. Examples include the Victorian Census Project, which has free data available from the 1861 census as well as Scottish registration data:

http://www.staffs.ac.uk/schools/humanities_and_soc_sciences/census/vichome.htm

Lecturers will also find useful datasets, as well as on-line teaching material, relating to the Victorian census and other sources, from the Centre for Historical and Contemporary Census Collections (CHCC). There are eight tutorials: seven dealing with themes in Victorian economic and social history, and one with the database and spreadsheet skills required to manipulate and use the census data.

http://chcc.arts.gla.ac.uk

These tutorials give some ideas as to how formal assessments might be constructed using census data. There are also suggestions of how other information — quantifiable and unquantifiable — might modify or enhance the picture we get from the census. Each is broken down into segments, which can be used in teaching sessions or simply put onto reading lists.

Naturally, in devising any element of a course, especially a formal assessment, based on the use of census or any other data in electronic format, consideration must be given to the amount of time required to generate a suitable exercise. A mixture of numerical ‘right/wrong’ questions and more probing questions about the implications of the data should be aimed for. The CHCC tutorials provide examples of the kinds of questions that can be asked, and the themes that might be addressed. The census is one of the easiest sources to use in this way, and possibly the least daunting for students, as well as being of interest in raising issues about the quality of the evidence that is being quantified.

Pollbooks and politics

This is another area in which the use of quantitative data can help students to think about wider historical questions. Prior to the introduction of secret ballots in British parliamentary elections (in 1872), pollbooks recorded who had voted for whom, and enable the historian to investigate electoral patterns and trends. As with many other such sources, the arrival of computer applications from the 1960s made exploitation much easier. Like the census enumerators’ books, pollbooks have been used by family historians: they contain voters’ places of residence, and in some cases an indication of their rank (e.g. ‘Gentleman’).

Eighteenth-century British political historians have made particularly widespread use of pollbooks to examine voting behaviour. A good starting point is W. A. Speck, Tory and Whig: The Struggle for the Constituencies 1701-1715 (1970). Speck, using quantitative analysis of pollbooks, found that party awareness among voters was high, and that voters were inclined to switch frequently between candidates at elections. The existence of a large ‘floating vote’ suggested to Speck that British democracy in this period was ‘participatory’, in other words that voters were aware of issues and made political choices based on their own interest and initiative.
This challenged the dominant view, most famously expressed by Lewis Namier, that voters acted ‘deferentially’, following their social superiors; the evidence marshalled by Namier and others in support of this interpretation was predominantly correspondence which emphasised the provision of transport for voters, the role of bribery and the exertion of influence. Since Speck first worked on the pollbooks, a large literature has grown up around the source, and the availability of data allows lecturers to incorporate some ‘hands-on’ study of pollbooks into the curriculum.

A word of caution should be introduced here. Much of the literature on pollbooks is based on evidence from several elections, which assesses the frequency with which individuals changed their electoral allegiance. This relies on the use of nominal record linkage, and this may not be appropriate at an undergraduate level. Again, however, the data can be presented to undergraduates in a reasonably simplified form, and, in a laboratory setting, simple quantitative questions can be asked. Pollbooks from a number of areas are available from AHDS History (study numbers 3032-3036, 3038, 3165, 3178-3179; see below), including for eighteenth-century Essex, Herefordshire and Northampton. S/N 3908 (Westminster) provides a series of pollbooks and rate books over 70 years and is accompanied by a book: C. Harvey et al., The Westminster Historical Database (Bristol, 1998). Nineteenth-century pollbooks are also available; and there are many studies of nineteenth-century voting behaviour.

**Height, health and history**

In the area of demography, the techniques available to the early practitioners of the ‘new economic history’ and their successors, as well as a range of source material, have enabled historians to use anthropometric data, in other words data on the physique of past populations, especially the height and weight of people in the past. Historians have used this data in a number of productive ways, contributing to the historiographies of France, American slavery, the ‘standard of living debate’ in Britain, and so on. Anthropometric history relates to the history of medicine and health, the history of the human body, as well as to the more traditional concerns of economic and social historians. The methodological issues associated with the use of anthropometric data can also encourage students to think about the nature and imperfections of primary source material.

The study of height data from the past also has the advantage of being something that students can easily understand. As Rodenick Floud and his collaborators pointed out (1990: 5-6), ‘we observe that some of our friends are much taller, some much shorter, than ourselves; in statistical terminology, there is variance in heights within the group’. The meaning and importance of anthropometric data is certainly more apparent to the uninitiated than data on economic growth, wages and earnings, or even literacy and other social/demographic indicators of well-being.

The use of anthropometric data by historians has also provoked some surprisingly fierce exchanges in the pages of academic journals, which in my experience have interested students more than the majority of papers they are assigned to read. A recent lively exchange in the *Economic History Review* between Peter Kirby and Jane Humphries, about the heights of children who worked in early Victorian coal mines, has proved particularly popular in seminars on the standard of living in Britain during the industrial revolution. Students have been surprised to read Humphries accusing Kirby of ‘breathtaking ignorance’! The references to these articles are in the bibliography below; and Peter Kirby’s data on child heights (AHDS History, study number 3108) can be ordered, and is a manageable dataset that could be used for a fairly basic laboratory or seminar teaching session. More complex questions are opened up by recent studies of the impact of smallpox on height: this debate encourages students to consider the nature and implications of the historical evidence.

This feature of records of historical heights, as in the case of the census and other quantitative sources, makes their use in teaching sessions a valuable insight into the processes by which historians use their source materials. Students should be made aware of how anthropometric data has been manipulated. The records of historical heights are quite limited, and we have data only on some groups within the population. We have some data on schoolchildren; and for the nineteenth century,
measurements were regularly collected from military recruits and from convicts. Neither of these were necessarily typical. Convicts were likely to come from more deprived social backgrounds, where we would expect heights to be lower. Military recruiting was often subject to a height limit, below which potential recruits were rejected.

Other problems include the clustering of measurements at convenient ½-inch intervals, the possibility of changing practices over time (e.g. measuring with and without shoes) and the small numbers of observations that can make generalisations questionable. However, all this makes the data more, rather than less, interesting and useful in the teaching of history at undergraduate level. There is a large recent literature on historical heights (some of which has challenged the historical accuracy of the normal height distribution represented above), and a number of datasets are available for downloading from AHDS history.

Some thoughts on assessment

It is difficult to offer more than tentative suggestions about how the study of historical numbers can be integrated into the curriculum in the early years of university history courses. The structure of existing courses, preferred methods of assessment, the availability of resources and the general institutional milieu can all militate against the introduction of any new curricular feature, and perhaps against quantification in particular. However, exercises based on quantitative data can enhance existing courses, and are particularly good at showing students how historians use their source material, and at encouraging them to think about the nature and limitations of that source material. In current practice, anything particularly difficult tends to be left to postgraduate research training courses, and in my experience this is the element of research training which postgraduate students try their hardest to get out of. (It should, however, be added that some of those who approach quantitative research methods courses with the most trepidation actually perform extremely well.)

On the whole, it is probably sensible to keep the more complex statistical techniques out of the mainstream undergraduate history curriculum.

However, as this brief guide has shown, there are many ways in which the study of quantitative data, at a fairly basic level, can be incorporated into mainstream history courses. The start-up costs need not be excessive, provided that computer laboratories are available, and the opportunities in terms of devising innovative assessed exercises are substantial. Assessment has been touched on briefly above, but some final remarks should be made.
Any assessment needs to bear in mind the various intended learning outcomes of the exercise, which are likely to include transferable skills as well as more academic outcomes. Where widely used software is employed, quantification exercises can incorporate assessment of a student’s facility with the relevant package. It is, of course, important to test students’ ability to perform at least the basic software functions and the most basic kinds of calculation, as these are necessary, although insufficient, conditions for presenting and interpreting the findings.

There are four stages in the process of using numbers in history. First, extracting the relevant data or carrying out the necessary calculations. Second, presenting the results in an appropriate form: is a table or a graph the best way to present a particular set of data? How should the data be labelled? What level of detail is required? All these questions need to be asked. Third, the findings must be interpreted: what does the data tell us about the occupational structure of the Victorian population, the results of eighteenth-century elections, the changing height of the population over time, and so on? Fourth, and most important, the interpretations must be fitted into wider historiographical contexts: does the occupational structure of the community reflect what we know about Victorian society more generally? Do these pollbooks suggest a large ‘floating vote’? What does the anthropometric data tell us about the standard of living? The course designer needs to devise exercises which address all four of these stages in the use of evidence.

In my department we have found that producing a worksheet for students to submit, featuring both numerical ‘right/wrong’ answers and more interpretive questions, allows the necessary balance to be struck. The simple questions, requiring the filling in of a table, for example, are followed up with questions requiring one- or two-paragraph answers, for which students are encouraged to read the secondary literature as widely as they would do in the preparation of a traditional history essay. Some examples of how these questions might progress through the four stages outlined above can be found in the CHCC on-line tutorials.

The kind of assessment used, of course, can vary as widely as in any other area of teaching. As suggested above, an article or book review, examining both the methodology and the wider context of a piece of historical writing that uses quantitative data, is one way of helping students to become familiar with historical quantification. Formal or informal skills tests can be used to test the level of acquisition of the key technical and numerical vocabularies. Traditional essays can be set, requiring the use of pre-circulated data. More ambitiously, group exercises, where students interrogate datasets and present their results collaboratively, can be undertaken. These are only a few suggestions; experience will show which are the most appropriate in a given learning environment.

Bibliography

General introductions

The most recent, and best, general textbook on quantitative methods and approaches is Pat Hudson, *History by numbers: an introduction to quantitative approaches* (London: Arnold, 2000).


Because the technology available changes so rapidly, history and computing textbooks date quickly. However, the general principles outlined in these textbooks remain applicable; only the technology available to the teacher and researcher of history has changed.

Quantification and its place in history

The classic text of the ‘new economic history’ was Robert W. Fogel and Stanley L. Engerman, *Time on the cross: the economics of American Negro slavery* (London: Wildwood House, 1974). This controversial book has gone through a number of editions: the Norton


**The Victorian censuses**

The most useful guide for undergraduate students is Edward Higgs, *A Clearer Sense of the Census: Victorian Censuses and Historical Research* (London: HMSO, 1996) - earlier editions were entitled *Making Sense of the Census*.

Other examples of the uses of the census by historians:


There are a great many others!

**Pollbooks and prosopography**


The ‘deference’ argument was most famously made in Lewis Namier, *The structure of politics at the accession of George III* (London: Macmillan, 1929; 2nd ed. 1957, with several reprints).


**Height, health and history**


Bernard Harris, ‘Health, Height and History: An Overview of Recent Developments in Anthropometric History’, *Social History of Medicine*, vol. 7 (1994), pp. 297-320.


For a reply to this, see Jane Humphries, ‘Short Stature among Coal-Mining Children: A Comment’, *Economic History Review*, vol. 50 (1997), pp. 531-7; and for Kirby’s rejoinder, immediately afterwards, see Kirby, ‘Short Stature among Coal-Mining Children: A Rejoinder’, *Economic History Review*, vol. 50 (1997), pp. 538-42.


Several other articles on anthropometric history have appeared in the *Economic History Review* during the past 10-15 years.

I am grateful to Andrew Roach and Sonja Cameron, and also to Donald Spaeth, who has made me aware of much of what is referred to above.