Making Sense of Quantitative Evidence

Gary J. Kornblith

(from the Making Sense of Evidence series on *History Matters: The U.S. Survey on the Web*, located at http://historymatters.gmu.edu)

Does the very thought of quantitative analysis make you shake in your shoes? Making Sense of Numbers provides a place for students and teachers to begin working with quantitative historical data as a way of understanding the past. Written by Gary J. Kornblith, this guide offers an overview of quantitative methods, how historians use historical data, and step-by-step instructions using actual historical data to determine totals, rates, averages, standard deviations, and coefficients of correlation. Gary J. Kornblith is Professor of History at Oberlin College. He is currently working with Carol Lasser and Patricia Holsworth on a social history of nineteenth-century Oberlin, Ohio, and his previous publications include "Hiram Hill: House Carpenter, Lumber Dealer, Self-Made Man" in Michael A. Morrison, ed., *The Human Tradition in Antebellum America* (2000), *The Industrial Revolution in America* (editor and co-author) (1998); and "Artisan Federalism: New England Mechanics and the Political Economy of the 1790s," in Ronald Hoffman and Peter J. Albert, eds., *Launching the "Extended Republic": The Federalist Era*, (1996).

Reading and Organizing Quantitative Evidence

By trade, historians tend to be skeptics who prefer the specificity of nitty-gritty facts to grand generalizations and fanciful speculations. Such skepticism seems especially appropriate when dealing with claims based on quantitative analysis. Although people often think of numeric data as "hard" evidence, there is also a common perception that experts can make numbers "say" anything they wish. As the aphorism attributed to Mark Twain (among others) declares, "There are three kinds of lies: lies, damned lies, and statistics." One may be tempted to dismiss quantitative analysis because it seems obscure and hence untrustworthy. Yet the information available in numeric form can be too valuable for a good historian to ignore. Quantitative data do not speak for themselves, but with a little coaxing they can sometimes tell us things about the past that we cannot discover in "qualitative" kinds of evidence.

The challenge for beginning historians is twofold: (1) to learn how to pose good questions of available quantitative sources, including both raw and aggregated data; and (2) to learn how to organize and "read" the data yourself to answer the questions you have posed. If you do not like mathematics, you probably will not become a heavy-duty quantitative historian. But you can still use basic quantitative methods in your research, and you can still become a critical reader of complicated quantitative scholarship. There is a range of reasonable positions between that of a true believer, on the one hand, and an anti-numeric nihilist, on the other. The philosophy underlying this guide is that quantitative history is too important to be left exclusively to the mathematically inclined.

What Is Quantitative History?

Put simply, quantitative history is history that involves the use of numeric data—or other evidence that can be counted—as a primary source for analysis and interpretation.

Quantitative history comes in many shapes and sizes. Some quantitative studies focus on small groups of people; others encompass huge populations. Some quantitative studies use data originally collected in numeric form, such as tax assessments or business ledgers; others involve the conversion of non-numeric evidence, such as city directories or church membership lists, into numeric form as a first step of analysis. Some quantitative studies employ rudimentary mathematical techniques (such as addition, subtraction, multiplication, and division) to analyze numeric data; others make use of highly sophisticated statistical procedures and mathematical model-building based on complex theoretical assumptions.

This guide aims to provide a good sense of what quantitative historians do and why they do it, but it does not pretend to offer instruction in upper-level statistics. The hope is that you will gain a basic comfort level with quantitative methods, which will in turn allow you to use numeric sources in your own historical research just as you would use other kinds of sources, such as letters, photographs, and newspaper articles.

Why Historians Started Counting

The emergence of quantitative history as a distinct approach dates to the 1960s, when the convergence of at least three trends prompted historians to turn to numeric data and statistical analysis for help in answering questions and framing interpretations.

First, there was a growing interest in the experiences of ordinary people as distinct from the achievements of "great white men." Political historians, for example, had long focused on the thoughts and actions of presidents, prominent congressmen, and other "movers and shakers" in the national government. But modern polling techniques suggested that voters did not always share the values and views of the public officials they voted for. Leaders could not automatically be considered spokespeople for their followers. So how was a historian to find out what really motivated ordinary people to vote for one person or one party rather than another? In an effort to answer that question, a group of "new political historians" turned to the study of voting behavior, using electoral data and increasingly complicated statistical techniques to determine which factors best explained voting patterns in particular areas during given periods of time. Likewise, "new social historians" set out to study history from "the bottom up." One question they wanted to address was whether the American ideal of equal opportunity was historical fact or fiction. Had it really been possible for poor yet meritorious Americans to rise to positions of wealth and status, or was the American social structure more bounded by hereditary constraints than implied by the "myth of the self-made man" and "the American dream"? Using census records, tax lists, and other quantifiable material, the new social historians sought to determine the extent of social mobility in American history.

A second trend that contributed to the rise of quantitative history was the movement to establish history as a social science dedicated to the rigorous, consistent, and precise application of social theory and social scientific methods in the study of past human behavior. Thus the new political historians of the 1960s borrowed from political science, and the new social historians borrowed from sociology. Yet the most celebrated and most controversial proponents of social-scientific approaches were the "new economic historians," who applied highly mathematical econometric theory and methodology to the study of longstanding historical questions and often came up with unorthodox answers. One famous example (at least within academic circles) will suffice. Conventional wisdom held that the key explanation for the acceleration of American economic growth during the nineteenth century was the advent and expansion of railroads. Robert Fogel decided to test this hypothesis by constructing a "counterfactual" model of what the nineteenth-century American economy would have looked like without railroads. He imagined a network of canals rather than railroads and then, building on limited data and a body of theoretical assumptions, he calculated the probable rate of economic growth under these alternative circumstances. To his avowed surprise, he concluded that canals would have served the economy almost as well as railroads, and hence that railroads were not indispensable to American economic growth in the nineteenth century. Not everybody was convinced, but few could ignore Fogel's audacious approach. He was later awarded the Nobel Prize in Economics in large part because of this pioneering work in the new economic history.

The third factor that encouraged the rise of quantitative history in the United States and elsewhere was the advent of the digital computer. In the early 1960s, huge mainframe computers cost hundreds of thousands of dollars each. The first academic "power users" tended to be natural scientists, but over time social scientists also discovered the advantage of these huge electronic devices for processing large amounts of information and executing elaborate calculations involving many variables and complex manipulation of the data. By comparison to many of their colleagues in related disciplines, historians were rather late in making the transition from note cards to punch cards, the essential input media of the mainframe era. As late as 1965, only a few dozen historians were involved in computerized research projects nationwide. But by the early 1970s, the computer revolution was reaching into history graduate programs, and increasing numbers of young historians learned the basic procedures of data entry and analysis using software such as SPSS. Especially for the study of large populations with hundreds of "data elements," the mainframe computer proved a godsend. Still, most historians continued to shy away from computers until the triumph of the personal computer in the 1980s and the advent of user-friendly software for word processing as well as statistical manipulation. Today the typical personal computer sitting on a faculty desk or in a college computer lab is hundreds of times more powerful and also much easier to use than the enormously expensive mainframes of the founding era of quantitative history.

How Do I Locate Quantitative Data and Assess Reliability?

Americans have long been, in the words of historian Patricia Cline Cohen, "a calculating people." Consequently the sources available for doing quantitative American history are enormously rich and varied. They include census returns, birth and death records, tax lists, membership lists of clubs, churches, and other organizations, business records, social surveys, price lists, city directories, and loads of other quantifiable collections of information.

Yet by comparison to sociologists, psychologists, economists, and political scientists, historians also confront a distinct limitation when they utilize quantitative data. For the most part, historians study dead people and records left behind by dead people. We cannot go back and ask our subjects new questions if we do not like the questions that were asked, say, by a census taker in 1820—or if we cannot read the census taker's handwriting. Likewise, we cannot design our own experiments comparing a test group to a control group to determine if factor "A" really made a historical difference. Instead we often must settle for data that were originally assembled by somebody with a different agenda than our own. But we can still be creative. For example, although we cannot ask voters why they preferred one presidential candidate to another in 1852, by using newspaper reports and government records, we can determine how people voted or at least how a group of people voted in a given electoral district. We can then compare voting patterns to patterns suggested by other available data, such as tax assessments, occupations, and/or the religious and ethnic composition of a particular group of voters. With the help of various statistical techniques, we can then make carefully limited inferences about why people voted the way they did, even though nobody asked them directly to explain their motivations.

So if you are given an assignment to use quantifiable sources, where should you begin? Basically, there are two different (but not mutually exclusive) starting points. You can either (1) begin with an existing data set and think about what question(s) you could answer using that data set; or (2) begin with a question (or set of questions) and look for data that would help you answer it (or them). Which starting point you choose will usually depend on the particular character of the assignment and your relative access to various data sets. You may be provided with a package of pre-assembled documents and data or—at the opposite end of the spectrum—you may be given "free rein" to choose your own topic and locate your own evidence. Between these extremes, you might be directed to use census data available on microfilm, on CD-ROM, or on the Internet. Or you might go to the local courthouse and gather information on the different kinds of criminal cases adjudicated during a given time period.

As you develop your project, keep in mind that not all data are equally reliable. Before you invest a lot of time and energy analyzing a particular set of data, you will want to have a general sense of how the data were collected, by whom, and for what purpose. For example, property assessments in tax lists may or may not represent the market value of taxpayers' estates. Sometimes assessors applied formulas that consistently discounted market prices or left assessments unchanged when market conditions fluctuated. Undoubtedly the meaning and accuracy of tax lists could be affected by whether the assessor simply asked the taxpayer to estimate the value of his or her property or, alternatively, undertook an independent examination of the estate. Likewise, some kinds of property might be exempt from taxation altogether. The more you can find out about the way the data were originally compiled, the better. If you determine that the process was biased, you may still be able to use the data, by taking steps to correct for the distortion in the initial collection procedures.

Can these records be linked to any others?

The special challenges confronting historians extend beyond the impossibility of communicating interactively with dead people. We also run into a problem when we try to trace individuals from one set of records to another. Say, for example, we are interested in comparing the average wealth-holdings of two church congregations, one Lutheran and the other Catholic. We are fortunate to have membership lists for both groups as well as complete tax lists for the larger community. A problem arises, however, when we discover that three Lutherans and two Catholics share the same name: John Williamson. This difficulty is compounded when we find only three John

Williamsons—one with little property, the others more affluent—on the tax list. On the basis of the available documents, we have no way to know which John Williamson belongs to which congregation. To be sure, because there are only two Catholic John Williamsons, we may assume that at least one John Williamson on the tax list is Lutheran—unless there are other congregations in the community or there is a John Williamson who belongs to no church at all. Still, even if we assume that one of the John Williamsons is Lutheran, we cannot determine which one—the relatively poor John Williamson or one of his richer namesakes? Imagine the difficulty when you are trying to make linkages between hundreds of records with dozens of duplicate names. Ironically, the modern solution to this sort of problem is to assign each individual or case a unique number to be used as an identifier across a range of different sets of records. (Notwithstanding federal regulations, your social security number frequently serves this purpose, which is why you may find it on your driver's license, your paycheck, your school transcript, and other important documents.) But the historian cannot assign such numbers retrospectively without having first resolved the confusion they are supposed to avoid.

Sampling: Do I Have to Count Everything?

Another problem historians often encounter has to do with sampling. Statisticians have developed elaborate ways to measure the reliability of samples, and to establish the likelihood that a particular sample appropriately embodies the qualities of the larger population from which it was selected. For a host of reasons, the kind of sample statisticians prefer to work with is a random sample, and this kind of sample can be hard for historians to draw. We often have only the partial remains of a larger body of data to work with—say one month's payroll of a business, two plantations' records of slave births over a given decade, or one neighborhood's building survey before a major fire. Statisticians call this kind of sample a "sample of convenience," but samples of convenience are frequently samples of necessity for historians. Again, we cannot go back and take a new sample of the entire body of data if the surviving records are incomplete.

Yet if the circumstances are favorable, you may want to draw a random sample from a larger population. When dealing with data sets of several hundred records, the advantage of sampling is great: you can save time, energy, and perhaps even your sanity without sacrificing statistical validity. Be sure to note, however, that a random sample is not the same thing as a "systematic sample." Going through a list and selecting every "nth" record is not a random process. To take a random sample, you should use a random number table (commonly found in the back of statistics textbooks) or a random number generator (included in various computer programs) to make sure that the selection process is truly unbiased. Once you have a random sample, you can apply a host of popular statistical tests that are not appropriate for other kinds of samples.

How Do I Locate Patterns? Totals and Rates

To make sense of the past, historians try to identify patterns in the evidence they collect. They look for connections between events or among various circumstances surrounding a single event. They try to determine what changed versus what stayed the same over time. At their most ambitious, they search for an underlying logic that explains historical trends, historical disjunctures (such as political and technological revolutions), and—ultimately—the dynamics of human behavior. Otherwise history would be just "one damn fact after another."

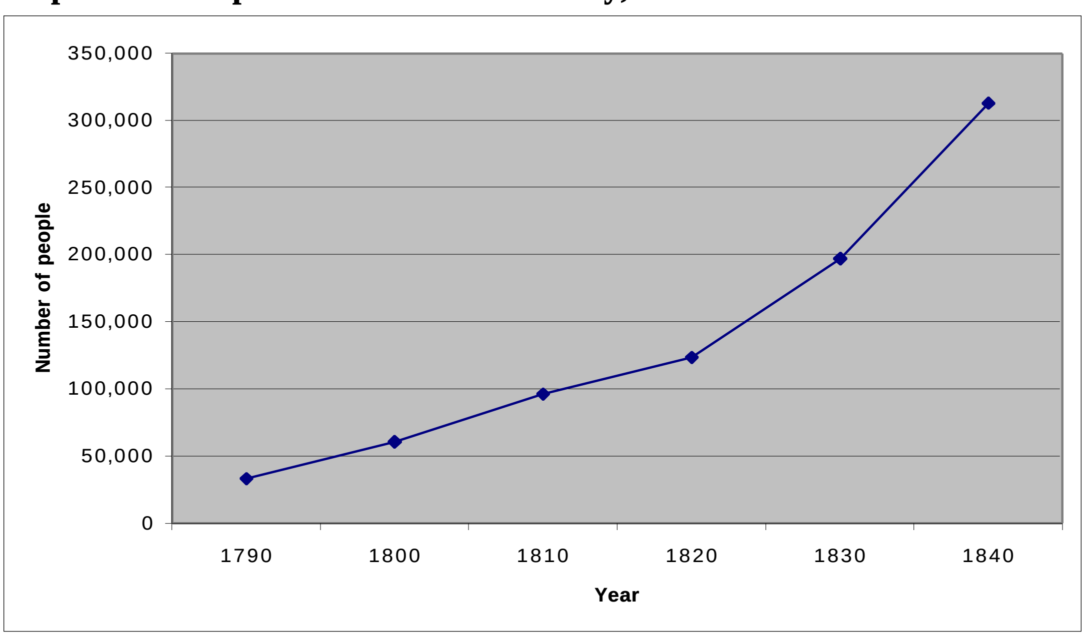
Quantitative historians use a wide variety of tools to help them locate patterns amid a welter of data. Some of these tools you are most likely familiar with from your grade school days, and we will start with them. You may be pleased to discover how much you already know about quantitative methods that will prove useful in your study of American history.

Let us begin with **totals**. Just tracking the total number of people in a particular place over time can be revealing. For example, using federal census records, we can determine how many people resided in New York City at ten-year intervals between 1790 and 1840. Indeed, the census takers themselves did the necessary addition, and the government published the results: in 1790, there were 33,131 residents; in 1800, there were 60,489; in 1810, there were 96,373; in 1820, there were 123,706; in 1830, there were 197,112; and in 1840, 312, 710. Written out in a single sentence, these numbers may not mean much. But if you put them in a table, you should begin to see a pattern. And if you put them in a graph, the pattern may become clearer.

Table A: The Population of New York City, 1790-1840

Year	Population
1790	33,131
1800	60,489
1810	96,373
1820	123,704
1830	197,112
1840	312,710

Graph 1: The Population of New York City, 1790-1840



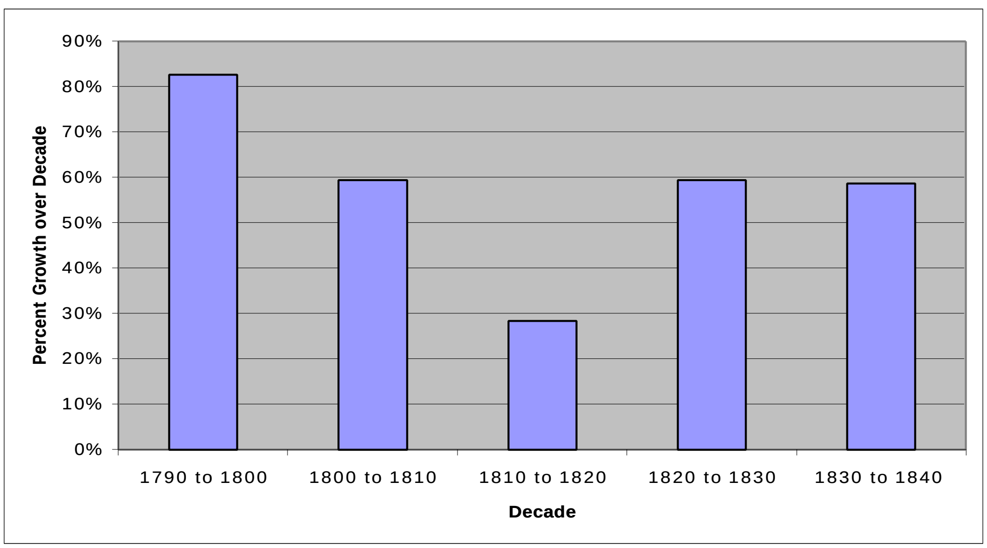
You must still "read" the pattern carefully, however. It is tempting to conclude from the above graph not only that the population of New York City grew substantially between 1790 and 1840, but that it grew at a higher **rate** after 1820 because the line on the graph rises more steeply after 1820 than before. Yet if you calculate population growth per decade as a **percentage** of the city's population at the start of the decade, you may be surprised by the results: by this measure, the decade with the highest rate of growth was 1790-1800. See Table B.

Table B: The Rate of Population Growth by Decade, New York City, 1790-1840

Year	Population	Population growth over decade (absolute difference)	Rate of population growth (increase as percentage of population at decade's start)
1790	33,131		
1800	60,489	27,358	83%
1810	96,373	35,884	59%
1820	123,704	27,331	28%
1830	197,112	73,408	59%
1840	312,710	115,598	59%

Graph 1 illustrates the pattern of population growth by decade in absolute terms. By contrast, Graph 2 below illustrates the change in the rate of population growth by decade.

Graph 2: The Rate of Population Growth by Decade, New York City, 1790-1840



So what can we learn from Graph 2? By itself, it suggests that New York City's population increased continuously between 1790 and 1840, but not at a consistent pace. While the population grew at virtually the same rate (59%) in each of three decades (1800 to 1810, 1820 to 1830, and 1830 to 1840), it grew at a noticeably higher rate in the 1790s and a markedly lower rate in 1810s. Why? Neither Graph 2 nor Table B provides an explanation, but were we to carry our study of New York City forward, we would want to investigate what was distinctive about the two "abnormal" decades.

How Do I Locate Patterns? Averages

Staying within the boundaries of grade-school mathematics, we now turn to **averages**. Recall that you were taught that there are three kinds of averages:

1) the mean or, more technically, the arithmetic mean (the sum of the values divided by the number of cases—what we usually intend when we use the term "average" in everyday conversation);

2) the median (the midpoint in a range of values so that half of the values are higher and half are lower); and

3) the mode (the most often repeated value within a data set).

All three kinds of averages are measures of what statisticians call "central tendency." That is, they represent an effort to identify the center or central number within a range of data, thereby summarizing what the data have in common.

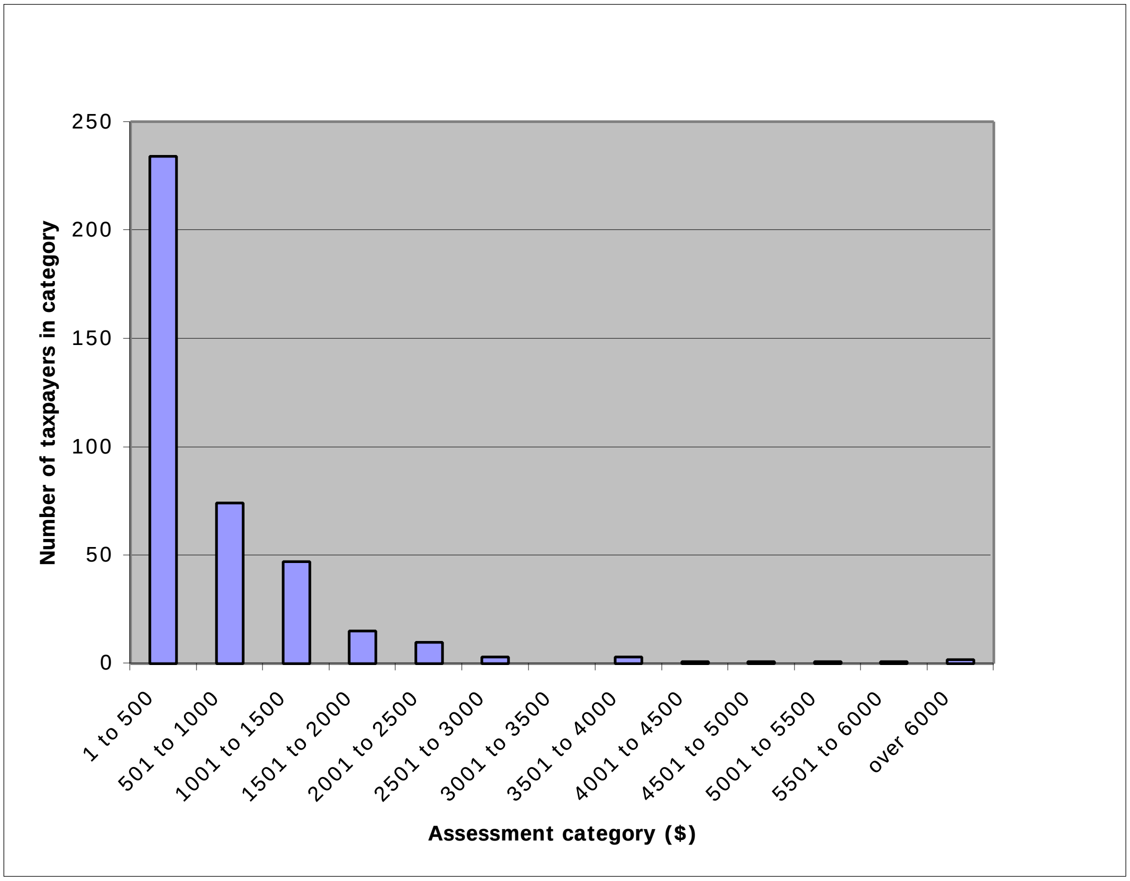
Let's look at an example of historical analysis that uses averages. Our data set is drawn from the tax list of Russia Township, Ohio, in 1850. The tax list provides property assessments for a total of 392 resident taxpayers. With the help of Microsoft Excel, we have calculated the mean, median, and mode:

Mean=$667
Median=$389
Mode=$260

Each of these figures tells us something about the average property holding of a Russia Township taxpayer in 1850. Indeed, depending on what you intend by the term "typical," you could argue that the typical Russia Township taxpayer owned $260, $389, or $667 in assessed property in 1850. It is important to specify which measure you are using when you speak of the "average" or "typical" member of a data set.

Taken together the mean and median tell us something about the larger pattern of property holding in Russia Township that neither reveals on its own. From the fact that the mean is higher than the median, we can infer that the distribution of property assessments is skewed rather than symmetrical. As it happens, the assessments of a small number of quite large property holders raised the mean without affecting the median. This pattern is visually evident in the following graph of the **distribution** of property assessments:

Graph 3: Distribution of Taxpayers by Assessment Category, Russia Township, Ohio, 1850



As Graph 3 makes clear, there was a large range of property assessments in Russia Township in 1850. And just as averages summarize the central tendency of a data set, other measures are useful for summarizing the **dispersion** or **variability** of a data set. The simplest way to specify dispersion is to give the minimum and maximum of the range, $13 and $7358 in the case of Russia Township property assessments. But the minimum and maximum by themselves do not tell us much about how tightly clustered or broadly spread out the bulk of data points are; they just tell us where the extremes lie at either end of the range.

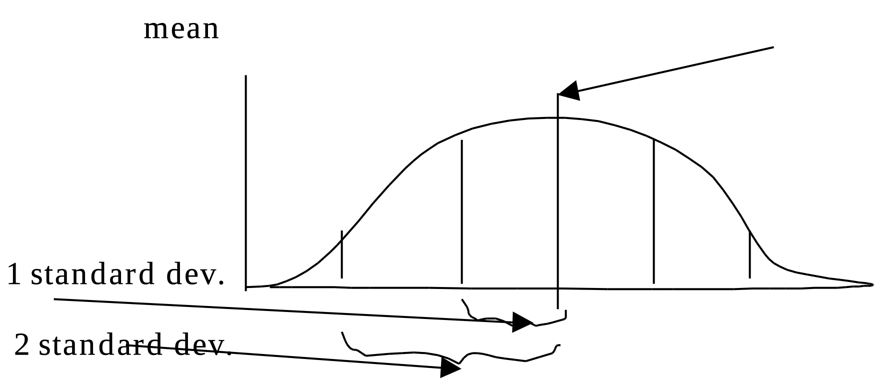
A less intuitive but otherwise more useful summary statistic of dispersion is **standard deviation**. Technically, standard deviation is defined as "the square root of the arithmetic mean of the squared deviations from the mean" [Hubert M. Blalock, Jr., *Social Statistics*, 2d ed. (New York, 1972), 80], and it is calculated according to the following formula:

$$\sigma = \sqrt{\frac{\sum (X-\mu)^2}{N}}$$

where "σ" stands for standard deviation, "X-μ" is the distance between each point in a data set and the mean of the data set, and *N* is the number of points in the data set. (As usual, Σ stands for "sum of" and √ stands for "square root.")

If this seems a bit confusing, don't panic. For common sense purposes, you may wish to conceptualize standard deviation in one of three ways. The first is to think of it as a measure of the average of the distances between each data point and the mean of the data set. The standard deviation is not the *mean* of the distances of the data points from the mean, but it is a kind of average.

A second way to think about standard deviations requires that you imagine a **normal distribution** or the so-called bell curve as pictured below. You are probably familiar with the notion of a normal distribution because aptitude and achievement tests like the SAT are designed so that test scores will be distributed according to such a symmetrical pattern within a large population.



Statisticians have established that in all normal distributions approximately 68 percent of the data will fall within one standard deviation on either side of the mean, and approximately 95 percent of the data will fall within two standard deviations on either side of the mean. That does not mean all normal distributions are identical, however. The bell curve can be flatter or steeper depending on the relative dispersion of the data. If the data are spread out, then the curve will be flatter and the standard deviation larger. If the data are tightly clustered around the mean, then the curve will be sharper and the standard deviation smaller. But the proportion of data within one standard deviation (68 percent) and within two standard deviations (95 percent) remains the same across all normal distributions.

Unfortunately, historical data rarely arrange themselves neatly into a normal distribution. So you may want to think about standard deviation in a third way, by comparing its magnitude to the mean of the data set. As a rule of thumb, when the standard deviation is smaller than the mean, the data are relatively closely clustered and mean is considered a reasonably good representation of the full data set. By contrast, if the standard deviation is greater than the mean, then the data are relatively widely dispersed and the mean is a rather poor representation of the full data set.

If we return to the data set of Russia Township taxpayers in 1850, we can calculate the standard deviation with the help of Microsoft Excel. It is $907, considerably larger than the mean of $667. Here is further evidence that it would be misleading to say, "The typical Russia Township taxpayer was assessed for $667" just because the mean property assessment was $667. In this instance, a better measure of typicality would be the median: $389.

Can these data be organized into categories?

Historians working with evidence that can be counted almost always confront the difficult problem of how to organize the data into categories. But it is only by putting things into categories that we can answer some of the most important and most interesting questions. For example, to say whether "big business" dominated the economy in the nineteenth century, we need to decide, what qualifies as "big." Or to argue that "poverty" decreased in the 1960s, we have to define what incomes qualify as "poor." Or to determine whether Americans in the nineteenth century experienced "upward social mobility," we need to decide what qualifies as "social mobility." Is it more property? How much? Is it a different job? What makes one job "better" than another? Can we put jobs into categories like "blue collar" and "white collar?"

One topic that involves categorizing data is characterizing the experience of American industrialization in the late nineteenth century, especially since historians disagree among themselves about how to characterize this experience. In the following exercise, we will address one aspect of this general issue by examining the profile of manufacturing enterprises in Cleveland, Ohio, in 1880.

We begin with two distinct but related questions: How big were most manufacturing firms? Did most industrial workers work for big or little firms?

Fortunately for us, federal census takers in 1880 collected information on the number of workers employed by each manufacturing firm in Cleveland, then one of the fastest growing cities in the nation. The resulting data set includes information for 1,018 firms.

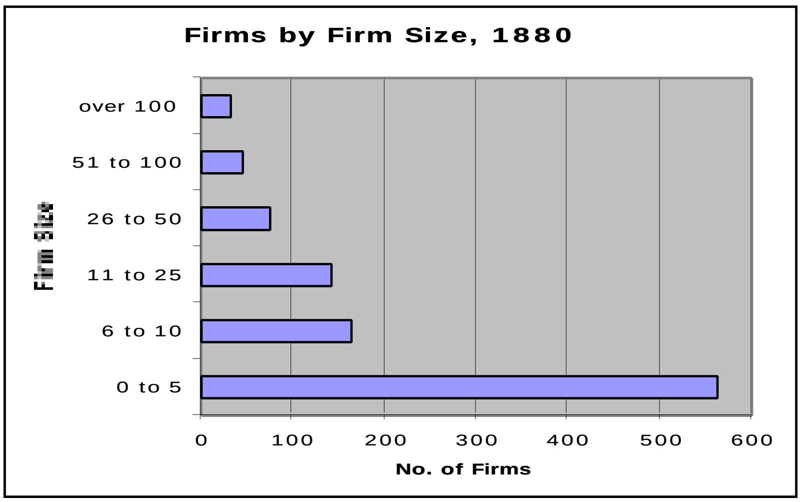
With the help of Microsoft Excel, we can calculate several simple descriptive statistics for our data set. The mean number of employees per firm is 20.0; the median is 4.0; and the mode is 2.0. The range of employees per firm is 0 to 1,935, while the standard deviation is 86.0.

From the median alone, we can arrive at a useful response to the first question posed above: half of the firms employed four or fewer employees. But we can also tell from the difference between the median and the mean and from the magnitude of the standard deviation compared to the mean that there was a wide dispersion of firm sizes. Moreover, we still cannot answer the second question we raised above because we do not yet have a full picture of the distribution of employees across different sizes of firms. To generate this picture, we need first to turn the qualitative terms "big" and "little" into numeric categories and then to graph our data so we can read it properly.

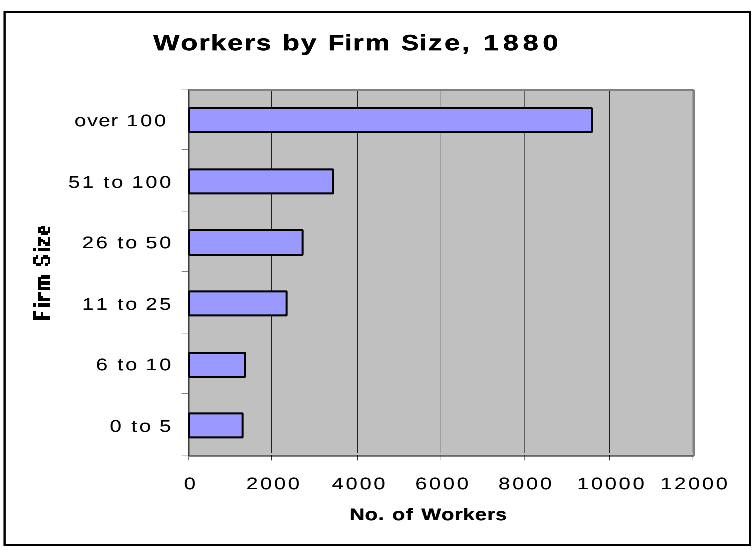
How should we **categorize** (or classify) firms? There are no universally agreed upon definitions of "big" or "little" firms, and to some extent any classification scheme we adopt will be arbitrary. We could just divide our range of firms into two groups, using—for example—either the mean, median, or mode as our "break point" between large and small. But given the difference between these measures of central tendency and given the data set's high standard deviation, we would be better off using a more elaborate classification scheme, one with several categories so as to better represent the distribution of the data. For this reason, we will group our data into six categories of firm size: 0 to 5 employees, 6 to 10 employees, 11 to 25 employees, 26 to 50 employees, 51 to 100 employees, and more than 100 employees. Note that in choosing this scheme we haven't really resolved the issue of what constitutes a "big" or "little" firm. Instead we have kept open the option of using different "break points" when we read the organized data.

Now we are ready to arrange our data and to look for patterns. Examine the two graphs below. Both are based on the data set of manufacturing firms in Cleveland, Ohio, in 1880. Graph 4 displays the distribution of *firms* by the category of firm size (e.g., number of workers employed). Graph 5 displays the distribution of *workers* by the category of firm size.

Graph 4: Number of Manufacturing Firms by Firm Size, Cleveland, Ohio, 1880



Graph 5: Number of Manufacturing Workers by Firm Size, Cleveland, Ohio, 1880



With the help of these graphs, you should be able to answer the questions we raised at the beginning of this exercise:

How big were most manufacturing firms? **[Answer: Most firms employed 5 or fewer workers.]**

Did most manufacturing workers work for big or little firms? **[Answer: Most workers worked for firms with more than 50 employees; almost half worked for firms with more than 100 employees. By the standards of the day, most manufacturing workers worked for big firms.]**

You may also want to consider a third question: Did the typical worker work for the typical firm? **[Answer: No, which may help explain why most employers considered themselves small businessmen while most workers viewed their employers as big businessmen.]**

How Do I Explain Patterns in the Data?

We observed earlier that historians search for patterns in surviving evidence from the past and that descriptive statistics can help in this process. But historians are not always happy just locating a pattern. They frequently want to explain the pattern; they want to know why the pattern emerged and took the shape that it did. In common parlance, they want to know the causes of the historical patterns they identify. And here again quantitative methods can be useful so long as we are careful not to treat statistical measures of association as the equivalent of historical proof. With the help of statistics we can infer the existence of a relationship between two variables or factors, but that does not mean that one factor *caused* the behavior of the other. It remains possible that a third variable was the “prime mover”; by just looking at two variables, we run the risk of mistaking **correlation** for causation. Even complex procedures that statisticians call “multivariate analysis,” which can handle several variables at once, are not powerful enough to prove historical causation to the satisfaction of most scholars. With few exceptions, historians do not believe that historical causation can be reduced to a formula, however complicated and sophisticated the mathematical manipulation of the data may be.

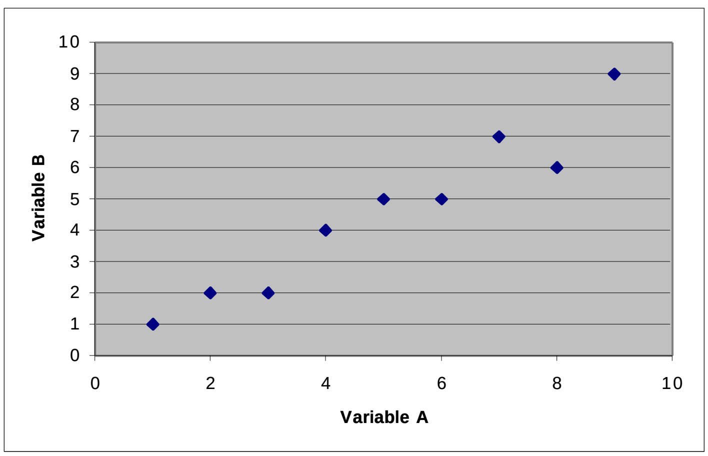
Yet if we cannot prove historical causation by means of statistics alone, we may still be able to use quantitative methods to help substantiate or challenge assertions about historical causation that are commonly expressed in qualitative terms. Although not all qualitative statements about historical causation can be tested quantitatively, some can. The key is to employ statistical tools in a reasonable, restrained, and responsible manner. And do not be surprised if you find that you can undermine a hypothesis more readily than you can come up with an alternative explanation for the data. In history as in many other disciplines, undercutting an argument or an interpretation is often easier than constructing one.

One of the first areas that attracted quantitative historians was the study of voting behavior. The basic story of how Americans had voted in the past was well known, but historians disagreed among themselves about why people voted the way they did. Some historians argued that Americans voted mainly according to their economic interests, pure and simple. Other historians contended that economic interests mattered less than cultural factors, such as ethnic identity, religion, and philosophical outlook. The “new political historians” sought to resolve this debate by using quantitative methods. In particular, they matched various quantifiable variables against the election results and measured the level of **correlation**.

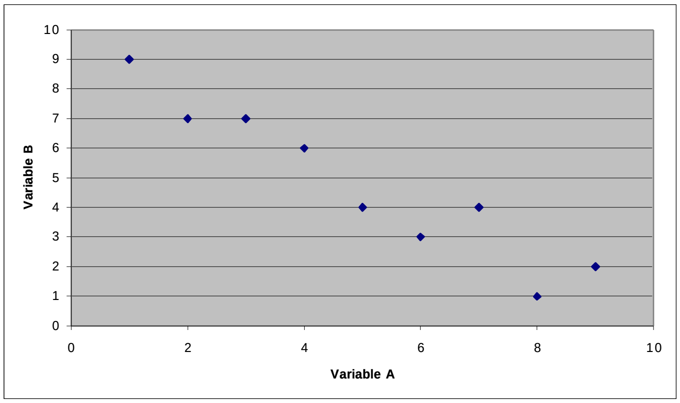
A **coefficient of correlation** indicates the strength of the relationship between two variables. The most commonly used correlation coefficient is the Pearson product-moment coefficient or Pearson *r*. While it is difficult to calculate, it is rather easy to interpret. Essentially, Pearson *r* can fluctuate between 0 and either +1 or – 1. The sign (+ or -) of Pearson *r* indicates the kind of relationship between the two variables. If Pearson *r* is positive, then the two variables behave in tandem and in the same direction: that is, if one goes up, the other goes up and if one goes down, the other goes down. On the other hand, if Pearson *r* is negative, then the two variables behave in tandem but in opposite directions: if one goes up, the other goes down. The closer the coefficient of correlation is to 1 or –1, the stronger the association is between the two variables. If Pearson *r* is 0, then there is no relationship between the two variables.

For those of you who prefer visual thinking, you can “see” different kinds of correlations in the three **scatterplots** below. Graph 6 shows a strong positive correlation between Variable A and Variable B; graph 7 shows a strong negative correlation between Variable A and Variable B; and graph 8 shows no correlation between the two variables.

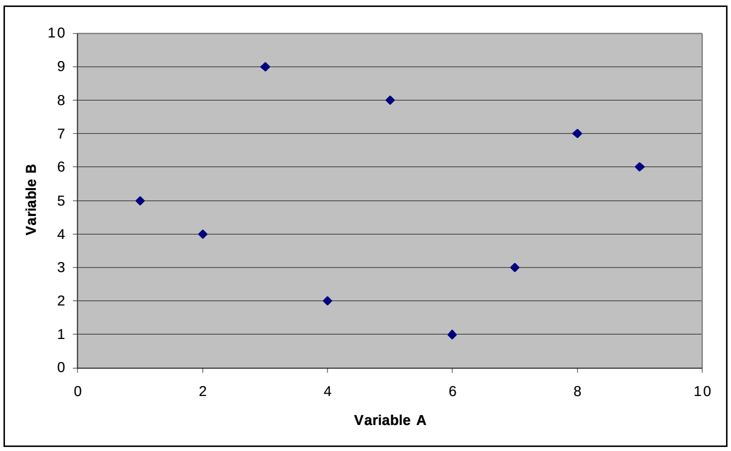
Graph 6: Strong positive correlation (r=+.96)



Graph 7: Strong negative correlation (*r*=-.95)



Graph 8: No correlation (*r*=.00)



Model Interpretation: Voting Patterns in Mississippi, 1860

Following the lead of the new political historians, we will analyze a real historical data set using correlation. Our case study involves the presidential election of 1860. As you may recall, there were four candidates for president in 1860: Abraham Lincoln (Republican), Stephen Douglas (northern Democrat), John C. Breckinridge (southern Democrat), and John Bell (Constitutional Union). We are concerned with the pattern of support for John C. Breckinridge, the most extreme pro-southern candidate, in the state of Mississippi, one of the first states to secede after Lincoln was elected. Statewide Breckinridge received 59.0% of the vote, compared to 4.7% for Douglas, 36.2% for Bell, and 0.0 % for Lincoln. But Breckinridge's support was not distributed evenly across all of Mississippi's counties. What kinds of counties do you think most strongly backed Breckinridge? More specifically, do you think there was a relationship between the density of slaves in a county and its level of support for Breckinridge? If so, what kind of relationship do you think there was?

Below you will find a table that lists the 58 counties in Mississippi for which the following data are available: (1) the percentage of voters who cast their ballots for

Breckinridge in the county, and (2) the percentage of the county's total population that was enslaved.

Table C: Percentage of Votes for Breckinridge and Percentage of Population Enslaved, Mississippi Counties, 1860

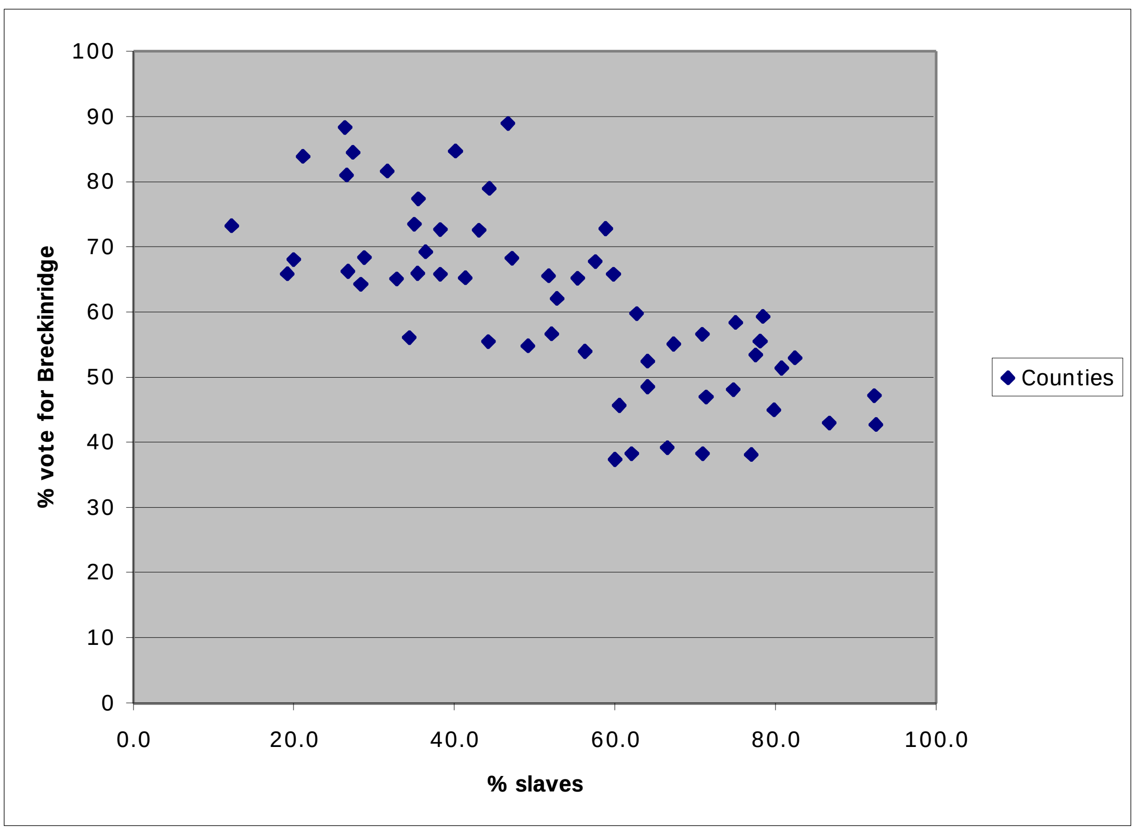
County	% votes for Breckinridge	% slave
Adams	38.3	70.9
Amite	52.5	64.0
Attala	66.0	35.4
Bolivar	43.0	86.7
Calhoun	65.9	19.2
Carroll	59.8	62.7
Chickasaw	65.2	55.3
Choctaw	66.3	26.7
Claiborne	59.3	78.4
Clarke	68.3	47.1
Coahoma	38.2	77.0
Copiah	65.6	51.7
Covington	77.4	35.5
De Soto	37.4	59.9
Franklin	67.8	57.5
Greene	81.6	31.6
Hancock	84.5	27.3
Harrison	83.9	21.1
Hinds	47.0	71.4
Holmes	55.1	67.3
Issaquena	42.8	92.5
Itawamba	68.1	19.9
Jackson	88.3	26.4
Jasper	65.3	41.3
Jefferson	51.4	80.8
Jones	73.3	12.2
Kemper	54.8	49.1
Lafayette	55.5	44.2
Lauderdale	65.8	38.2
Lawrence	84.7	40.1
Leake	65.1	32.8
Lowndes	56.6	70.8
Madison	53.5	77.5
Marion	88.9	46.6
Marshall	45.7	60.5
Monroe	65.8	59.8
Neshoba	81.0	26.5
Newton	73.5	35.0
Noxubee	58.4	75.0
Oktibbeha	72.8	58.8
Panola	38.3	62.0

Perry	64.3	28.3
Pike	79.0	44.3
Pontotoc	56.1	34.4
Rankin	56.7	52.1
Scott	69.3	36.4
Simpson	72.7	38.2
Smith	68.4	28.7
Sunflower	55.6	78.0
Tallahatchie	48.6	64.1
Tunica	45.0	79.8
Warren	39.2	66.5
Washington	47.2	92.3
Wayne	62.1	52.7
Wilkinson	53.0	82.4
Winston	72.6	43.0
Yalobusha	54.0	56.2
Yazoo	48.1	74.7

Source: Great American History Machine

Now let us look at these data graphed as a **scatterplot**.

Graph 9: Relationship Between Percentage of Vote for Breckinridge and Percentage of Population Enslaved in Mississippi Counties, 1860



What does the scatterplot suggest about the relationship between support for Breckinridge and slave density in Mississippi counties in 1860? The scatterplot shows that support for Breckinridge declined as slave density—the proportion of population enslaved—increased.

With the help of Microsoft Excel, we can further specify the relationship by calculating the coefficient of correlation. Pearson *r* is **-.73**. What does this number mean? The fact that the sign of *r* is negative indicates that the variables are related inversely; in other words, as the percentage of the population enslaved rises, the percentage of the vote for Breckinridge declines. Moreover, the magnitude of the correlation is relatively high, especially for historical data. If you square *r,* you get .53. This number, called the coefficient of determination, indicates that 53 percent of "the total variation in one variable is associated with variation in the other variable." [Loren Haskins, Kirk Jeffrey, *Understanding Quantitative History* (Cambridge, 1990), 245.] In other words, more than half of the variation in the vote for Breckrinridge among Mississippi counties in 1860 was associated inversely with the variation in the proportion of slaves in the population of the counties.

Are you surprised? Many students assume that on the eve of the Civil War pro-southern extremism was strongest where slavery was most deeply entrenched. But our analysis of the Mississippi data suggests otherwise. Support for Breckinridge was greatest where slave density was lowest, not highest. By itself the existence of this negative correlation does not explain why Mississippi counties with more slaves were less favorably inclined toward Breckinridge. Nor does it contradict the fact that Mississippi as a state voted overwhelmingly for Breckinridge in 1860. Yet if numbers can speak, this figure cries out for further investigation. As a next step in the research process, we could turn to election statistics for other southern states to see if the same pattern holds for them. [Analysis of data for Alabama and Georgia reveals that in those states, too, the vote for Breckinridge correlated negatively with proportion of slaves in the population but not as strongly as in Mississippi.] Alternatively, we could further explore Mississippi sources, such as newspapers or political speeches, to see what more we can discover there. Either way, we now know better than to assume that the prevalence of slavery alone explained the pattern of southern extremism in 1860.

Conclusion

Even if quantitative methods are better at dispelling myths and challenging simple assumptions than they are at proving arguments about historical causation, they can serve as critical tools in the hands of all kinds of historians. As a beginning historian, try to approach numeric data as you would other types of evidence, with seriousness and skepticism, and devote enough time and energy to mastering the quantitative skills you need to accomplish your research goals. Good luck!

Quantitative Historical Data Online

Quantitative history encompasses many things, from basic statistical skills to analyzing available data to collecting data. As the guide to quantitative history tries to address these various aspects, so too does this guide to online sources. Some of these resources offer individual level data; others offer aggregate data. In addition, some sites provide raw data that must be downloaded and analyzed with statistical software while other sites offer data that are available for online manipulation. This list is intended as a brief overview of the multitude of sources, data, and guides available online, providing links to some of the largest collections and most comprehensive resources on quantitative evidence.

Archives—Records and Quantitative Data

American Family Immigration History Center, Statue of Liberty-Ellis Island Foundation
http://www.ellisisland.org/
Created by a non-profit organization to fund preservation of the Statue of Liberty and Ellis Island, this site provides a searchable database containing records on more than 22 million passengers and ship crewmembers who passed through Ellis Island between 1892 and 1924. In addition to a basic passenger record (name, ethnicity, place of residence, date of arrival, age, marital status, ship of travel, and place of departure), users may view a copy of the original ship manifest (a text version is also available), and a picture of the ship. If no match occurs, the site provides information for names with close or alternate spellings.

Bureau of Economic Analysis, U.S. Department of Commerce
http://www.bea.doc.gov
Comprehensive and summary data estimates concerning national, international, and regional economic activity, and "statistics that influence the decisions made by government officials, business people, households, and individuals." Includes an overview of the economy, providing data on production, purchases by type, price, personal income, government finances, inventories, and balance of payments. This site also offers news releases concerning key economic indicators, descriptions of sources and methodologies used, and articles from the organization's publications. A keyword index to a 1929-2000 set of annual and quarterly national income and product account (NIPA) tables allows users access to data on specific product sales and ways that consumers have spent money.

Online Data Archive, University of Wisconsin, Madison
http://dpls.dacc.wisc.edu/archive_txt.htm
Provides 41 social science statistical data studies on a variety of topics, including 12 studies dealing with Wisconsin-related topics, 14 studies on American subjects, and 15 studies dealing with general or international matters. Subjects pertaining to American history include Slave Movement during the 18th and 19th Centuries; Irish immigrants in Boston in 1847 and 1848; Characteristics of Census Tracts in Nine U.S. Cities, 1940-1960; the growth, consumption habits, and finances of American families in the 1950s and 1960s; financial characteristics of consumers in the early 1960s; premarital sexuality in 1973; Civil Rights volunteers, 1965-1982; urban racial disorders of the 1960s; and the role of the American family in the transmission and maintenance of socioeconomic inequality. This Web site provides data for download; statistical software may be required to analyze and process the data.

FRED, Economic Time-Series Database, Federal Reserve Bank of St. Louis
http://www.stls.frb.org/fred/index.html
Offers national economic and financial data in 12 categories, including: interest rates; consumer price indexes; employment and population; gross domestic product and

components; producer price indexes; trade data; and daily/weekly financial data. Much of the data was compiled monthly. Periods covered vary according to category, and some statistics go back to 1901. Also provides historical and recent statistics for the states of Arkansas, Illinois, Indiana, Kentucky, Mississippi, Missouri, and Tennessee.

Integrated Public Use Microdata Series, University of Minnesota
http://www.ipums.umn.edu/
Currently provides 22 census data samples and 65 million records from 13 federal censuses covering the period 1850-1990. These data "collectively comprise our richest source of quantitative information on long-term changes in the American population." The project has applied uniform codes to previously published and newly created data samples. Rather than offering data in aggregated tabular form, the site offers data on individuals and households, allowing researchers to tailor tabulations to their specific interests. Includes data on fertility, marriage, immigration, internal migration, work, occupational structure, education, ethnicity, and household composition. Offers extensive documentation on procedures used to transform data and includes 13 links to other census-related sites. This site may be somewhat challenging for novices.

Research Data on Voting and Public Opinion, National Election Studies
http://www.umich.edu/~nes/
This site contains a wealth of data from National Election Studies surveys of the American electorate conducted in presidential and congressional election years from 1948 to 1998. Survey information covers public opinion and political participation on topics such as the effectiveness of major political parties, election outcomes, interest in the campaign, and important issues facing voters. The data files and codebooks for each study are available for download, but these large files take considerable time to open and provide complex and highly technical information. More accessible are more than 200 tables and graphs that trace public opinion from 1948 to 1998 on nine topics: Social and Religious characteristics of the Electorate; Partisanship and Evaluation of Political Parties; Ideological Self-Identification; Public Opinion on Public Policy Issues; Support for the Political System; Political Involvement and Partisanship in Politics; Evaluation of Presidential Candidates; Evaluation of Congressional Candidates; and Vote Choice. The site also offers *The NES Guide to Public Opinion and Electoral Behavior,* which offers easily digestible data on the issues drawn from these studies. This site is somewhat difficult to navigate.

US Presidential Election Maps: 1860-1996, University of Virginia Library
http://fisher.lib.virginia.edu/elections/maps/
Maps, color-coded according to presidential candidate, display percent of popular vote the winning candidate in each state received in elections between 1860 and 1996. Currently includes maps showing electoral vote distributions by state for elections between 1900 and 1996. Also contains a chart with the number and percentage of votes each candidate received in each state in the 2000 election. Maps of Virginia show cities and counties won by George W. Bush and Al Gore in the 2000 election, and the percentage of votes that Bush, Gore, Ralph Nader, and all third-party candidates received in each county.

United States Historical Census Data Browser, University of Virginia Library
http://fisher.lib.virginia.edu/census/
Provides data gathered by the Inter-University Consortium for Political and Social Research from census records and other government sources for a study entitled "Historical Demographic, Economic, and Social Data: The United States, 1790-1970." For each decade, users may browse extensive population- and economic-oriented statistical information at state and county levels, arranged according to a variety of categories, including place of birth, age, gender, marital status, race, ethnicity, education, illiteracy, salary levels, housing, and specifics dealing with agriculture, labor, and manufacturing. Allows users to select up to 15 variables when conducting searches and displays both raw data and statistical charts.

Valley of the Shadow: Two Communities in the American Civil War, University of Virginia
http://jefferson.village.virginia.edu/vshadow
A massive, searchable archive relating to two communities, Staunton, Virginia, and Chambersburg, Pennsylvania, before, during, and after the Civil War, including church, agricultural, military, and public records. Public records include population and agricultural censuses, tax digests, and, for Augusta county, records about slave owners and free blacks.

Statistical Guides

HyperStat Online Textbook, David M. Lane
http://davidmlane.com/hyperstat/
An 18-chapter introductory statistics textbook. Each chapter includes links to related articles and books; some include exercises. Provides 16 links to general statistics texts and data sources.

Java Demos for Probability and Statistics
http://www.math.csusb.edu/faculty/stanton/m262/probstat.html
Statisticians, as well as students and instructors, will appreciate this intelligible collection of Java applets. The interactive applets clearly model probability distributions and illustrate other basic statistical concepts. Included are applets that demonstrate hypergeometric distribution, Poisson distribution, normal distribution, bivariate normal distribution, proportions, confidence intervals for means, the central limit theorem, linear regression, and Buffon's Needle. Professor Charles Stanton of California State University, San Bernardino, the Applet developer, provides brief descriptions and instructions for most of the demonstrations.

Selected Annotated Bibliography

Adyelotte, William O., Allan G. Bogue, and Robert William Fogel, eds. *The Dimensions of Quantitative Research in History* (Princeton, NJ: Princeton University Press, 1972).
A collection of essays by leading pioneers of quantitative history that includes case studies in the social, political, and economic development of the United States, France, and Great Britain.

Clubb, Jerome M., Erik W. Austin, and Gordon W. Kirk, Jr. *The Process of Historical Inquiry: Everyday Lives of Working Americans* (New York: Columbia University Press, 1989).
Uses data collected on families of American textile workers in 1888-90 as case study for the application of quantitative historical methods and elementary statistical analysis.

Fogel, Robert William and G. R. Elton, *Which Road to the Past: Two Views of History* (New Haven: Yale University Press, 1983).
Two essays comparing the methods and merits of statistically-oriented "scientific" history and traditional history, one by a champion of quantification and the other by a skeptic.

Gonick, Larry and Woollcott Smith. *The Cartoon Guide to Statistics* (New York: HarperPerennial, 1993).
An entertaining and cleverly illustrated, yet serious, jam-packed introduction to statistics—not a joke book.

Haskins, Loren and Kirk Jeffrey. *Understanding Quantitative History* (Cambridge: M.I.T. Press, 1990).
A "user-friendly" introduction to the application of statistical methods in history that focuses on the basic concepts and skills necessary to read quantitative historical scholarship carefully and critically.

Hudson, Pat. *History by Numbers: An Introduction to Quantitative Approaches* (London: Arnold, 2000).
A comprehensive, sophisticated introduction to statistical methods for historians and the theoretical and empirical issues involved in doing quantitative history; examples drawn mainly from British sources.

Jarausch, Konrad H. and Kenneth A. Hardy, *Quantitative Methods for Historians: A Guide to Research, Data, and Statistics* (Chapel Hill: University of North Carolina Press, 1991).
A solid introduction to doing (not just reading about) quantitative history, especially research involving large databases; the information on computer applications is behind current practice.

Phillips, John L. *How to Think about Statistics*, 6th ed. (New York.: W.H. Freeman, 2000).
A highly accessible primer on fundamental concepts of social statistics, written for undergraduates by a psychologist; examples are drawn from social sciences other than history.

Swierenga, Robert P., ed. *Quantification in American History: Theory and Research* (New York: Atheneum, 1970).
A collection of early essays on quantitative approaches to American history; includes discussions of methodology, an influential critique of quantification, and several examples of political, economic, and social history.

(Click to read a summary of this study HERE). Benefits of the qualitative approachÂ The quantitative approach to gathering information focuses on describing a phenomenon across a larger number of participants thereby providing the possibility of summarizing characteristics across groups or relationships. This approach surveys a large number of individuals and applies statistical techniques to recognize overall patterns in the relations of processes.Â Stronger support for successful training would be evident if using quantitative methods. Integrative Approaches to Qualitative and Quantitative Evidence. Article (PDF Available) Â· January 2004 with 4,601 Reads.